Extremal Results in and out of Additive Combinatorics

Thesis by
Andrei Cosmin Pohoata

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2020
Defended May 21, 2020

Andrei Cosmin Pohoata
ORCID: 0000-0002-3757-2526

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Nets Katz, for showing me the world of additive combinatorics through his lens, for giving the freedom to develop and pursue my personal taste, for his guidance and continued patience.

The Caltech math department has been a warm and welcoming place from the very beginning. I especially want to thank David Conlon, Maksym Radziwiłł, Adam Sheffer and Misha Tyomkyn for making my time at Caltech so enjoyable.

I am also grateful to all my dear friends and collaborators for always being supportive and fun to spend time with over the years. It's been truly a pleasure getting to know you all, my life would be undoubtedly much more unpleasant without some of you in it. You know who you are!

Finally, I would like to thank Karina for her love and support, and my parents for everything they have done for me. This thesis is dedicated to them.

ABSTRACT

In this thesis, we study several related topics in extremal combinatorics, all tied together by various themes from additive combinatorics and combinatorial geometry.

First, we discuss some extremal problems where local properties are used to derive global properties. That is, we consider a given configuration where every small piece of the configuration satisfies some restriction, and use this local property to derive global properties of the entire configuration. We study one such Ramsey problem of Erdős and Shelah, where the configurations are complete graphs with colored edges and every small induced subgraph contains many distinct colors. Our bounds for this Ramsey problem show that the known probabilistic construction is tight in various cases. We study one discrete geometry variant, also by Erdős, where we have a set of points in the plane such that every small subset spans many distinct distances. Finally, we consider an arithmetic variant, suggested by Dvir, where we are given sets of real numbers such that every small subset has a large difference set. In Chapter 2, we derive new bounds for all of the above problems. Along the way, we also essentially answer a question of Erdős and Gyárfás.

Second, we study the behavior of expanding polynomials on sets with additive or multiplicative structure. Given an arbitrary set of real numbers A and a two-variate polynomial f with real coefficients, a remarkable theorem of Elekes and Rónyai from 2000 states that the size $|f(A, A)|$ of the image of f on the cartesian product $A \times A$ grows asymptotically faster than $|A|$, unless f exhibits additive or multiplicative structure. Finding the best quantitative bounds for this intriguing phenomenon (and for variants of it) has generated a lot of interest over the years due to its intimate connection with the *sum-product problem* in additive combinatorics. In Chapter 3, we discuss new bounds for $|f(A, A)|$ when the set A has few sums or few products.

Another central problem in additive combinatorics is the problem of finding good quantitative bounds for maximal progression-free sets in the integers (or various other groups). In 2017, a major breakthrough of Croot, Lev and Pach took the community by surprise with impressive new bounds for the problem in $\mathbb{Z}_4^n$ and in higher order 2-abelian groups. Their new polynomial method was quickly adapted by Ellenberg and Gijswijt to show a similar strong result for the size of the largest three-term progression free subset of $\mathbb{F}_q^n$ where q is an odd prime power, the so-called *cap set problem*. This new set of ideas has subsequently led to very exciting

developments in a vast range of topics. The rest of the thesis will be dedicated to discussing my joint results around these new developents. In Chapter 4, we develop a new multi-layered polynomial method approach to derive improved bounds for the largest three-term progression free set in $\mathbb{Z}_8^n$ (which also improve on the Croot-Lev-Pach bounds for a large family of higher order 2-abelian groups). In Chapter 5, we generalize the Ellenberg-Gijswijt bound for the cap set problem to random progression-free subsets of $\mathbb{F}_q^n$, improving on a theorem of Tao and Vu. A result of this type enables one to find four term progressions-free sets which contain three-term progressions in all of their large subsets (with good quantitative bounds), but which do not contain too many three-term progressions overall. Motivated by this application, in Chapter 6 we continue this investigation and study further the question of determining the maximum total number of 3APs in a given 4AP-free set. We show in general, for all fixed integers $k > s \geq 3$, that if $f_{s,k}(n)$ denotes the maximum possible number of s-term arithmetic progressions in a set of n integers which contains no k-term arithmetic progression, then $f_{s,k}(n) = n^{2-o(1)}$. This answers an old question of Erdős. In Chapter 7, we study some limitations of the Croot-Lev-Pach approach and discuss some problems at the intersection of extremal set theory and combinatorial geometry where one can use additional linear algebraic ideas to go slightly beyond the Croot-Lev-Pach method.

Published Content and Contributions

(Ch2) C. Pohoata, A. Sheffer, *Local properties in colored graphs, distinct distances, and difference sets*, **Combinatorica**, 39 (2019), Issue 3, pp. 705-714.
https://link.springer.com/article/10.1007/s00493-018-3890-2

(Ch2) S. Fish, C. Pohoata, A. Sheffer, *Local properties via color energy graphs and forbidden configurations*, **SIAM Journal on Discrete Mathematics**, 34 (2020), No. 1, pp. 177-187.
https://epubs.siam.org/doi/abs/10.1137/18M1225987

(Ch3) C. Pohoata, *On cartesian products which determine few distinct distances*, **Electronic Journal of Combinatorics**, 26 (2019), Issue 1, P1.7.
https://www.combinatorics.org/ojs/index.php/eljc/article/view/v26i1p7

(Ch3) C. Pohoata, *Expanding polynomials on sets with few products*, **Mathematika**, 66 (2020), pp. 71-78.
https://londmathsoc.onlinelibrary.wiley.com/doi/full/10.1112/mtk.12007

(Ch4) F. Petrov, C. Pohoata, *Improved bounds for progression-free sets in C_8^n*, **Israel Journal of Mathematics**, 236 (2020), Issue 2.
https://link.springer.com/article/10.1007/s11856-020-1977-0

(Ch5) C. Pohoata, O. Roche-Newton, *Four-term progression free sets with three-term progressions in all large subsets*, submitted.
https://arxiv.org/abs/1905.08457.

(Ch6) J. Fox, C. Pohoata, *Sets without k-term progressions can have many shorter progressions*, submitted.
https://arxiv.org/abs/1908.09905.

(Ch7) F. Petrov, C. Pohoata, *On sets of points with prescribed distances*, submitted.
https://arxiv.org/abs/1912.08181.

(Ch7) H. Huang, O. Klurman, C. Pohoata, *On subsets of the hypercube with prescribed Hamming distances*, **Journal of Combinatorial Theory Series A**, 171 (2020), 105156, 21 pp.
https://www.sciencedirect.com/science/article/abs/pii/S0097316519301372

In each of the papers listed above, all authors contributed equally.

TABLE OF CONTENTS

Acknowledgements . . . iii
Abstract . . . iv
Table of Contents . . . vii
Chapter I: Introduction . . . 1
1.1 Local to global properties in colored graphs and difference sets . . . 1
1.2 Sum-product phenomenon and polynomial expansion . . . 5
1.3 Progression-free sets in the integers (and in other groups) . . . 9
1.4 On the Croot-Lev-Pach Lemma and its limitations . . . 17
Chapter II: On the Erdős-Gyárfás problem in graph Ramsey theory . . . 23
2.1 Color energy and forbidden repeats of rainbow stars . . . 25
2.2 Higher order color energy and forbidden cycles . . . 28
Chapter III: Expanding polynomials and sets with additional structure . . . 35
3.1 Algebraic and analytic preliminaries . . . 39
3.2 Expansion when AA is small . . . 40
Chapter IV: Improved Bounds for Progression-Free Sets in $\mathbb{Z}_8^n$ . . . 46
4.1 Regularization and the tensor power trick . . . 49
4.2 Subspaces with zero product in abelian 2-groups . . . 51
4.3 Croot–Lev–Pach bound for C_4^n with group rings . . . 53
4.4 Improved bounds for progression-free sets in C_8^n . . . 54
4.5 A large 3AP-free set in $\mathbb{Z}_8^n$ . . . 60
Chapter V: 4AP-free sets with 3AP's in all large subsets . . . 63
5.1 Hypergraph containers . . . 65
5.2 Supersaturation Results for 3APs . . . 67
5.3 Improved bounds for the cap set problem in random subsets of $\mathbb{F}_q^n$ . . . 69
5.4 An analogous story over integers . . . 74
5.5 Sets with small energy but rich in progressions . . . 78
Chapter VI: How many 3APs can a 4AP-free set actually have? . . . 86
6.1 Proof of Erdos' conjecture . . . 87
Chapter VII: Beyond the Croot-Lev-Pach Lemma . . . 94
7.1 On sets with few distances in $\mathbb{R}^d$ . . . 94
7.2 An algebraic proof of Kleitman's Theorem . . . 96
7.3 Extensions to arbitrary distance sets . . . 103
7.4 On sets in $\mathbb{F}_p^N$ whose difference set avoids the hypercube . . . 109

Chapter 1

INTRODUCTION

1.1 Local to global properties in colored graphs and difference sets

In 1974, Erdős and Shelah [**Erd74**] initiated the study of the following beautiful problem. For positive integers n, k, ℓ, we consider edge colorings of the complete graph K_n such that every induced subgraph over k vertices contains at least ℓ colors. Let $f(n, k, \ell)$ be the minimum possible number of colors in a coloring of K_n with this restriction.

One motivation for studying the asymptotic behavior of this function is that it can be seen as a generalization of the classical Ramsey's theorem, which studies the size of the maximum n with the property that there exists an edge coloring of K_n with c colors where every k of the vertices span at least $\ell = 2$ distinct colors.

In what follows, we allow ℓ to be larger than two, fix the value of k, and look for the minimum c satisfying the above condition as n grows. The Erdős-Shelah problem is attractive in its generality, because there are some ranges where determing $f(n, k, \ell)$ is much easier than Ramsey's theorem. For example, when $k \geq 4$ one can easily prove that

$$f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 2\right) = \Theta\left(n^2\right). \tag{1.1}$$

Indeed, in this case every color can occur at most $\lfloor k/2 \rfloor - 1$ times, since otherwise we will have an induced subgraph with k vertices and at most $\binom{k}{2} - \lfloor k/2 \rfloor + 1$ colors. Since each color repeats a constant number of times, there are $\Theta\left(n^2\right)$ distinct colors.

A few first results for the problem were obtained by Erdős, Elekes, and Füredi [**Erd81**]. Erdős and Gyárfás [**EG97**] started studying the problem more systematically. Using a probabilistic argument based on the Lovász Local Lemma (similar to the best one available for the lower bound construction for diagonal Ramsey numbers), Erdős and Gyárfás [**EG97**] derived the general upper bound

$$f\left(n, k, \ell\right) = O\left(n^{\frac{k-2}{\binom{k}{2}-\ell+1}}\right). \tag{1.2}$$

In the other direction, considering a restriction slightly weaker than in (1.1), they derived the bound $f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 1\right) = \Omega\left(n^{4/3}\right)$, and further asked whether or not the function $f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 1\right)$ is actually quadratic in n when $k \to \infty$.

In joint work with Sara Fish and Adam Sheffer [**FPS20**], we answered the question of Erdős and Gyárfás in the affirmative by showing that

$$f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 1\right) = \Omega(n^{2-8/k}), \tag{1.3}$$

thus confirming that the function grows quadratically in n as $k \to \infty$. Moreover, our bound is almost sharp even for fixed k, since by (1.2) we know that $f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 1\right) = O(n^{2-4/k})$.

The proof of our theorem builds upon ideas from a previous joint work with Adam Sheffer [**PS19**]. In this work, we studied the function $f(n, k, \ell)$ in a different range, since we were motivated by a nice result of Fox, Pach, and Suk [**FPS17**], who showed that for any $\epsilon > 0$

$$\phi\left(n, k, \binom{k}{2} - k + 6\right) = \Omega\left(n^{8/7-\epsilon}\right). \tag{1.4}$$

Here $\phi(n, k, l)$ denotes instead the minimum number of distinct distances that are determined by a planar n point set P with the property that any k points of P determine at least l distinct distances.

It is easy to see that $\phi(n, k, \ell) \geq f(n, k, \ell)$ holds for every positive integers n, k, and ℓ, since in order to lower bound $\phi(n, k, \ell)$ one can simply build a complete graph where every point is a vertex, and every distance corresponds to a distinct edge color. The proof of (1.4), however, is based on the geometry of the problem, so it does not extend to give a lower bound for the corresponding graph theoretic function $f(n, k, \ell)$. Nevertheless, in [**PS19**], we were able to prove the following result.

Theorem 1.1.1 *For any integers $k > m \geq 2$,*

$$f\left(n, k, \binom{k}{2} - m \cdot \left\lfloor \frac{k}{m+1} \right\rfloor + m + 1\right) = \Omega\left(n^{1+\frac{1}{m}}\right).$$

This is a somewhat surprising statement, since not only that it automatically gives an improvement over the Fox-Pach-Suk result for $\phi(n, k, \ell)$ when $\ell \geq \binom{k}{2} - 7 \cdot \lfloor k/8 \rfloor + 8$, but it also gives some unexpectedly strong bounds for $f(n, k, \ell)$, which perhaps didn't feel in reach initially. The bound of Theorem 1.1.1 is asymptotically tight up to sub-polynomial factors for every $m \geq 2$, except possibly for small values of k. Indeed, for every $\varepsilon > 0$ the bound of (1.2) implies

$$f\left(n, k, \binom{k}{2} - m \cdot \left\lfloor \frac{k}{m+1} \right\rfloor + m + 1\right) = O\left(n^{1+\frac{1}{m}+\varepsilon}\right),$$

for every sufficiently large k.

Distinct distances with local properties. The *Erdős distinct distances problem* has been for a long time one of the main problems in discrete geometry. This problem asks for the minimum number of distinct distances spanned by a set of n points in $\mathbb{R}^2$. That is, denoting the distance between two points $p, q \in \mathbb{R}^2$ as $|pq|$, the problem asks for $\min_{|\mathcal{P}|=n} |\{|pq| : p, q \in \mathcal{P}\}|$. Note that n equally spaced points on a line span $n-1$ distinct distances. Erdős [**Erd46**] observed that a $\sqrt{n} \times \sqrt{n}$ section of the integer lattice $\mathbb{Z}^2$ spans $\Theta(n/\sqrt{\log n})$ distinct distances. Proving that every point set determines at least some number of distinct distances turned out to be a deep and challenging problem.

The above problem is just one out of a large family of distinct distances problems, including higher-dimensional variants, structural problems, and many other types of problems (for example, see [**She14**]). The main problems in this family were proposed by Erdős and have been studied for decades. After over 60 years and many works on distinct distances problems, Guth and Katz [**GK15**] almost settled the original question by proving that every set of n points in $\mathbb{R}^2$ spans $\Omega(n/\log n)$ distinct distances. Surprisingly, so far this major discovery was not followed by significant progress in the other main distinct distances problems.

Determining the function $\phi(n, k, l)$ is a very interesting wide open problem for most choices of the parameters k and l. For example, even the value of $\phi(n, 3, 3)$ is somewhat mysterious. First, note that $\phi(n, 3, 3)$ can be also thought of as the minimum number of distinct distances that are determined by a set of n points that do not span any isosceles triangles (including degenerate triangles with three collinear vertices). Since no isosceles triangles are allowed, every point determines $n-1$ distinct distances with the other points of the set, and we thus have $\phi(n, 3, 3) = \Omega(n)$. Erdős [**Erd86**] observed the following upper bound for $\phi(n, 3, 3)$. Behrend [**Beh46**] proved that there exists a set A of positive integers $a_1 < a_2 < \cdots < a_n$ such that no three elements of A determine an arithmetic progression and $a_n < n2^{O(\sqrt{\log n})}$. Therefore, the point set $\mathcal{P}_1 = \{(a_1, 0), (a_2, 0), \ldots, (a_n, 0)\}$ does not span any isosceles triangles. Since $\mathcal{P}_1 \subset \mathcal{P}_2 = \{(1, 0), (2, 0), \ldots, (a_n, 0)\}$ and $D(\mathcal{P}_2) < n2^{O(\sqrt{\log n})}$, we have $\phi(n, 3, 3) < n2^{O(\sqrt{\log n})}$.

The following two result is immediate corollary of Theorem 1.1.1.

Corollary 1.1.2 *For any integers $k > m \geq 2$,*

$$\phi\left(n, k, \binom{k}{2} - m \cdot \left\lfloor \frac{k}{m+1} \right\rfloor + m + 1\right) = \Omega\left(n^{1+\frac{1}{m}}\right).$$

While there are many problems in which the conjectured number of distinct distances is $\Omega(n^{2-\varepsilon})$, proving such strong quantitative bounds has been long considered to be a very challenging endeavor. For example, see [**She14**]. It is therefore perhaps worth emphasizing that the following corollary of the bound from (1.3) represents the first result in this spirit.

Corollary 1.1.3 *For any integers $k > m \geq 2$,*

$$\phi\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 1\right) = \Omega(n^{2-8/k}).$$

Difference sets with local properties. Zeev Dvir recently suggested studying the following additive combinatorics variant of the local properties problem. Given a finite $A \subset \mathbb{R}$, the *difference set* of A is

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

For positive integers n, k, ℓ, we consider sets $A \subset \mathbb{R}$ of size n such that every subset $A' \subset A$ of size k satisfies $|A' - A'| \geq \ell$. Let $g(n, k, \ell)$ be the minimum size of $A - A$ among all sets A that satisfy the above restriction. For simplicity we will ignore non-positive differences. For example, when considering sets with no three-term arithmetic progression we will write $g(n, 3, 3)$ instead of $g(n, 3, 7)$ (we ignore 0 and three negative differences). This notation does not change the problem, and is somewhat more intuitive.

While this seems like a very interesting and natural additive combinatorics problem, we only managed to find one minor and brief mention of it. It is stated in [**EG97**] that Erdős and Sós proved $g(n, 4, 5) \geq \binom{n}{2} - n + 2$, although it seems that this was never published. Like above, the following is a corollary of Theorem 2.0.1.

Corollary 1.1.4 *For any integers $k > m \geq 2$,*

$$g\left(n, k, \binom{k}{2} - m \cdot \left\lfloor \frac{k}{m+1} \right\rfloor + m + 1\right) = \Omega\left(n^{1+\frac{1}{m}}\right).$$

To prove this corollary one can build a complete graph where every element of A is a vertex, and every difference corresponds to a distinct edge color.

We now present an example illustrating that $g(n,k,\ell)$ and $\phi(n,k,\ell)$ may be very different problems. Currently, nothing is known about $\phi(n,4,5)$ beyond the trivial bounds $\phi(n,4,5)=\Omega(n)$ and $\phi(n,4,5)=O(n^2)$ –this is considered to be a difficult open problem. On the other hand, it can be easily shown that $g(n,4,5)=\Theta(n^2)$. Indeed, consider a set A of n real numbers and with every $A'\subset A$ of size four satisfying $|A'-A'|\geq 5$. Assume that there exist four distinct reals $a_1,a_2,a_3,a_4\in A$ such that $a_1-a_2=a_3-a_4$. This implies that $a_1-a_3=a_2-a_4$, and thus these four points span at most four differences. The above contradiction implies that no difference repeats more than twice (it is still possible that $a_1-a_2=a_3-a_1$). We conclude that $|A-A|=\Theta(n^2)$.

In [**FPS20**], we also derive some significantly stronger lower bounds for $g(n,k,\ell)$ than what Theorem 1.1.3 gives.

Theorem 1.1.5 *For all $k>r\geq 2$,*

$$g\left(n,2rk,\binom{2rk}{2}-\binom{2k}{2}\cdot\left[\binom{r}{2}+(r-1)\right]+1\right)=\Omega\left(n^{\frac{r}{r-1}\cdot\frac{k-1}{k}}\right).$$

For example, by setting $r=2$ in Theorem 1.1.5, we get that for every even $k\geq 4$,

$$g\left(n,k,\binom{k}{2}-2\cdot\binom{k/2}{2}+1\right)=\Omega\left(n^{2-\frac{8}{k}}\right).$$

For large k, the expression $2\cdot\binom{k/2}{2}$ is almost half of $\binom{k}{2}$. That is, the number of allowed difference repetitions is about half of the total number of pairs. This behavior is very different than the behavior of $f(n,k,\ell)$, where the linear threshold occurs already when there are about k repetitions. As we increase r in Theorem 1.1.5, the number of allowed repetitions increases while the lower bound for the number of differences decreases. We will discuss the proofs of the results stated in this section in Chapter 2 of this thesis.

1.2 Sum-product phenomenon and polynomial expansion

Given a finite subset A in a field $\mathbb{F}$, in the previous section we defined the difference set of A as

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

Similarly, one can define the *sum set* and *product set* of A by

$$A - A = \{a + a' \ : \ a, a' \in A\} \ \text{ and } \ AA = \{aa' \ : \ a, a' \in A\}.$$

As before, let us work over the reals. To get some intuition, first note a trivial upper bound for the size of the sum set, $|A + A| \leq (|A|^2 + |A|)/2$. Indeed, the right-hand side is simply $|A|$, the number of sums of the form $a + a$ with $a \in A$, plus the total number of subsets of A of size two, and equality is obtained when every pair of elements of A gives a distinct sum. If we build A by taking real numbers at random, then we expect $|A + A|$ to be very close to this upper bound, since the probability that $a + b = c + d$ is very small. On the other hand, if $A = \{1, \ldots, n\}$ then

$$|A + A| = |\{2, 3, \ldots, 2n\}| = 2|A| - 1.$$

The same bound holds a little bit more generally whenever A is an arithmetic progression, and it is not too difficult to check that this is indeed the only class of extremal configurations. In other words, for every $A \subset \mathbb{R}$ we have that $|A + A| \geq 2|A| - 1$, with equality if and only if A is an arithmetic progression. This can be argued as follows (over $\mathbb{R}$). Denote the elements of A as $a_1 < a_2 < \ldots < a_{|A|}$. Then $A + A$ contains the following $2|A| - 1$ distinct elements:

$$a_1 + a_1 < a_1 + a_2 < a_1 + a_3 < a_1 + a_4 < \ldots < a_1 + a_{|A|} < a_2 + a_{|A|} < \ldots < a_{|A|+|A|}.$$

Hence $|A + A| \geq 2|A| - 1$. If $|A + A| = 2|A| - 1$, these must be the only elements inside $A + A$. However, note that

$$a_1 + a_1 < a_1 + a_2 < a_2 + a_2 < a_2 + a_3 < \ldots < a_{|A|-1} + a_{|A|-1} < a_{|A|-1} + a_{|A|} < a_{|A|} + a_{|A|}$$

is also a list of $2|A| - 1$ distinct elements in $A + A$. It follows that these two lists must coincidence term by term, which easily implies that A is an arithmetic progression. A similar dychotomy holds for the difference set $A - A$. One of the main problems of additive combinatorics is characterizing the finite sets A (in either $\mathbb{R}$ or other additive groups) for which $|A + A|$ (or $|A - A|$) is small with respect to $|A|$. When A is a set which satisfies $|A + A| = o(|A|^2)$ or $|A - A| = o(|A|^2)$ we will losely say that A has *additive structure*.

Just as with sum sets, a trivial bound for the size of A is $|AA| \leq (|A|^2 + |A|)/2$. And once again, if A is obtained by choosing elements of $\mathbb{R}$ at random, then we expect $|AA|$ to be very close to this upper bound. On the other hand, if $A = \left\{2^1, 2^2, \ldots, 2^n\right\}$, then

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

The same bound applies when A is a geometric progression with $0 \notin A$. For every finite set A with $0 \notin A$, one can similarly argue like above that $|AA| \geq 2|A| - 1$. When $0 \in A$, it is possible to obtain $|AA| = 2|A| - 2$.

One of the central questions in additive combinatorics is the so-called *sum-product conjecture* of Erdős and Szemerédi [**ES83**] which states that no set has both a small sum set and a small product set.

Conjecture 1.2.1 *For $\epsilon > 0$, there exists n_0 such that every finite set $A \subset \mathbb{R}$ with $|A| \geq n_0$ satisfies the inequality*

$$\max\{|A + A|, |AA|\} \geq c|A|^{2-\epsilon}$$

for some absolute constant $c > 0$.

Erdős and Szemerédi [**ES83**] also showed how to construct arbitrarily large sets A of size n with $\max\{|A + A|, |AA|\} = O\left(|A|^{2-\frac{c}{\log\log n}}\right)$ for some constant $c > 0$. This example shows that Conjecture 1.2.1 is false without the extra ϵ in the exponent. For the record, it is also perhaps important to mention that in [**ES83**], Erdős and Szemerédi originally asked to prove Conjecture 1.2.1 for subsets of the integers rather than reals, but the question stands for all subsets of the reals and all the recent works study it in this generality. The problem has a long and beautiful history, with a lot of exciting progress and many equally difficult variants. See [**KS15**], [**KS16**], [**Sha19**], and [**SS19**].

Expanding polynomials and sets with additional structure. My work so far on this topic has mostly focused on polynomial expansion over the reals. Given polynomials $f \in \mathbb{R}[x]$ and $g \in \mathbb{R}[x, y]$, and sets $A, B \subset \mathbb{R}$, we write

$$f(A) = \{f(a) : a \in A\} \text{ and } g(A, B) = \{g(a, b) : a \in A,\ b \in B\}.$$

That is, $g(A, B)$ is the set of distinct values that can be obtained by applying g on the cartesian product $A \times B$. When $g(x, y) = x + y$ or $g(x, y) = xy$, to be consistent with the sum-product world, the more convenient notation $g(A, B) = A+B$ or $g(A, B) = AB$ is generally preferred. Since the story is over the reals, my contributions naturally revolve around the result of Elekes and Rónyai, who in [**ER00**] uncovered the remarkable fact that $|f(A, B)|$ must be asymptotically larger than $|A|$ or $|B|$, if the polynomial $f \in \mathbb{R}[x, y]$ does not have one of the special

forms $f = h(g_1(x) + g_2(y))$ or $f = h(g_1(x) \cdot g_2(y))$, for some $h, g_1, g_2 \in \mathbb{R}[x]$. The current best bound for this problem is the following one by Raz, Sharir, and Solymosi [**RSS16**].

Theorem 1.2.2 *Let d be a positive integer, let $A, B \subset \mathbb{R}$ be finite sets, and let $f \in \mathbb{R}[x, y]$ be of degree d. Then, unless $f = h(g_1(x)+g_2(y))$ or $f = h(g_1(x)\cdot g_2(y))$ for some $h, g_1, g_2 \in \mathbb{R}[x]$, we have*

$$f(A, B) = \Omega_d \left(\min \left\{ |A|^{2/3}|B|^{2/3}, |A|^2, |B|^2 \right\} \right).$$

Theorem 1.2.2 generalizes many problems from discrete geometry and additive combinatorics –in particular, the bipartite distinct distances problem for sets of points lying on two lines, which, despite being a particular case of the Elekes-Ronyai problem, is also very likely to be the best model problem to look at if one would like to improve on the bound from Theorem 1.2.2.

This general problem has also been studied successfully over $\mathbb{F}_p$ as well, where many strong results are available. See for instance [**BT12**] or [**Tao15**]. It is perhaps important to point out that the first result with good quantitative bounds for polynomial expansion was in fact over finite fields: in [**Vu08**], Vu classified the two variable polynomials $f(x, y) \in \mathbb{F}_p[x, y]$ such that $|f(A, A)|$ is large whenever $|A + A|$ is small. Motivated by [**Vu08**], Shen then considered in [**She12**] the analogous question over the reals and used tools from incidence geometry to prove the following beautiful result, which preceeded and inspired Theorem 1.2.2.

Theorem 1.2.3 *Let d be a positive integer and let $f \in \mathbb{R}[x, y]$ be a polynomial of degree d that is not of the form $g(L(x, y))$ for some linear polynomial L and some univariate polynomial g. If A is a finite set of real numbers, then*

$$|A + A||f(A, A)| = \Omega_d \left(|A|^{5/2} \right).$$

In particular, Theorem 3 implies that when $A \subset \mathbb{R}$ satisfies $|A + A| = K|A|$, we have that $|f(A, A)| = \Omega_{d,K}(|A|^{3/2})$ for every polynomial that is not of the form $g(L(x, y))$ for some linear polynomial L and some univariate polynomial g. Like Theorem 1.2.2, however, this is not optimal and it is widely believed that the exponent $3/2$ could probably be replaced with $2 - \epsilon$ for every $\epsilon > 0$ in general (as it is the case for $f(x, y) = xy$; see for example [**Sol09**] for more details).

In joint work with Adam Sheffer [**PS17**], by using more advanced incidence theorems and a more careful analysis, we managed to improve on Theorem 3 when $|A + A|$ is small by showing the following result.

Theorem 1.2.4 *Let $A \subset \mathbb{R}$ be a finite set with $|A + A| = K|A|$ and let $f \in \mathbb{R}[x, y]$ be a polynomial of degree at most d that is not of form $g(L(x, y))$ for some linear polynomial L and some univariate polynomial g. Then, for any $\epsilon > 0$, we have*

$$|f(A, A)| = \Omega_{d,K}\left(|A|^{14/9-\epsilon}\right).$$

We will discuss this a bit more in Chapter 3. In a more recent follow up work [**Poh20**], we also addressed the natural "dual" problem of classifying the two variable polynomials $f(x, y) \in \mathbb{R}[x, y]$ such that $|f(A, A)|$ is large whenever $|AA|$ is small. As in the additive case, it is easy to check that there are some polynomials for which this implication does not hold; however, perhaps a bit surprisingly, it turns out we can prove a much stronger expansion theorem than in the additive case.

Theorem 1.2.5 *Let d be a positive integer and let $f \in \mathbb{R}[x, y]$ be a polynomial of degree d that is not of the form $g(M(x, y))$ for some single monomial $M(x, y)$ and some univariate polynomial g. If A is a finite set of real numbers such that $|AA| = K|A|$, then*

$$|f(A, A)| = \Omega_{d,K}\left(|A|^2\right).$$

This is optimal up to the dependence on d and K, and when $f(x, y) = x+y$ it recovers a result of Chang [**Cha06**]. The proof of Theorem 3.0.5 is in some sense in the spirit of the proofs of Theorem 1.2.2, Theorem 1.2.3, and Theorem 1.2.4, but the reason it does so well quantitatively is the fact that it does not rely on any incidence geometry. The main input is instead a (quantitative) version of the celebrated Schmidt subspace theorem [**Sch72**] due to Amoroso and Viada [**AV09**], which we use to control the number of "nondegenerate" solutions to certain polynomial equations. We will discuss this proof in detail in Chapter 3.

1.3 Progression-free sets in the integers (and in other groups)

For a positive integer $k \geq 3$ and a set A inside some additive group, we define $f_k(A)$ to be size of the largest k-AP free subset of A, i.e. the size of the largest subset of A

without any sequence of the form $\{x, x+d, \ldots, x+(k-1)d\}$, with $d \neq 0$. Following the standard notation, we will write $r_k(N) := f_k(\{1, \ldots, N\})$. Estimating $r_k(N)$ has been a central topic in combinatorics for almost a century. The question was originally inspired by a theorem of van der Waerden [**Wae27**], one of the earliest results in Ramsey theory: if the integers are partitioned into finitely many parts then one of the parts must contain arbitrarily long arithmetic progressions. [1]

In the 1930's, Erdős and Turán [**ET30**] conjectured that this was true simply because one of the sets in such a partition must contain a positive proportion of the integers, and that any set of similar density must also contain arbitrarily long arithmetic progressions. In other words, $r_k(N) = o(N)$, as N tends to infinity. In 1953, Roth [**Rot53**] resolved this conjecture for $k = 3$, and in 1975 Szemerédi [**Sze75**] settled the full question, showing that indeed $r_k(N) = o(N)$ holds for every $k \geq 3$, as N tends to infinity. These theorems (and their proofs) have had a profound impact on modern combinatorics and modern mathematics in general. Several different proofs of Szemerédi's theorem have since been discovered, and some of them have blossomed into rich areas of mathematical research. Here are some the most influential modern proofs of Szemerédi's theorem (in historical order):

- The ergodic theoretic approach (Furstenberg [**Fur77**]).
- Higher-order Fourier analysis (Gowers [**Gow01**]).
- Hypergraph regularity lemma (Rődl et al. [**RNSSK05**], Gowers [**Gow07**]).

Another modern proof of Szemerédi's theorem results from the *density Hales–Jewett theorem*, which was originally proved by Furstenberg and Katznelson [**FK91**] using ergodic theory, and subsequently a new combinatorial proof was found in the first successful Polymath Project, an online collaborative project initiated by Gowers [**Pol12**][2].

The relationships between these disparate approaches are not yet completely understood, and there are many open problems, especially regarding quantitative bounds. A unifying theme underlying all known approaches to Szemerédi's theorem is the *dichotomy between structure and pseudorandomness*.

[1] Having arbitrarily long arithmetic progressions is very different from having infinitely long arithmetic progressions. For instance, it is not too difficult to show that one can color the integers using just two colors so that there are no infinitely long monochromatic arithmetic progressions.

[2] All subsequent Polymath Project papers are written under the pseudonym D. H. J. Polymath, whose initials stand for "density Hales–Jewett."

One of the most intensely studied problems since the work of Roth and Szemerédi has been the question of determining the exact growth of $r_3(N)$ as a function of N. Roth's beautiful Fourier analytic method was powerful enough to give the explicit quantitative bound $r_3(N) \ll N/\log\log N$. Szemerédi [**Sze90**], Heath-Brown [**HB87**], Bourgain [**Bou99**], [**Bou08**], Sanders [**San11**], and Bloom [**Blo16**] subsequently improved on this bound by pushing to new heights the analytic framework provided by Roth, now commonly referred to as the density increment method. The current state-of-the-art is that

$$\frac{\log^{1/4} N}{2^{2\sqrt{2}\sqrt{\log N}}} \cdot N \ll r_3(N) \ll \frac{(\log\log N)^4}{\log N} \cdot N. \tag{1.5}$$

The upper bound in (1.5) is from [**Blo16**], while the lower bound comes from Elkin [**Elk11**], who slightly improved upon the celebrated construction of Behrend [**Beh46**]. It is a common belief that the lower bound is closer to the true magnitude of $r_3(N)$, but this is currently out of reach with the current methods. A target within reach is however the following more refined conjecture of Erdős and Turan, commonly stated as follows (for progressions of length three).

Conjecture 1.3.1 *(Erdős-Turan) Suppose $A \subset \mathbb{N}$ is a sequence such that $\sum_{a\in A} \frac{1}{a} = \infty$. Then, A contains non-trivial three-term arithmetic progressions.*

By partial summation, it is not too hard to see that this would be implied by the fact that $r_3(N) \ll N/(\log N)^{1+\epsilon}$ for some $\epsilon > 0$, which is not too far from the bound in (1.5).

In the meantime, there has also been some spectacular progress on the cap set problem from an unexpected different direction. In a recent breakthrough paper, Croot, Lev and Pach [**CLP17**] used a novel polynomial method approach to show that

$$f_3(\mathbb{Z}_4^n) \leqslant 4^{\gamma n} \approx (3.611)^n, \tag{1.6}$$

for some explicit constant γ which involves the binary entropy function $\mathcal{H}_2(p)$. This was a remarkable improvement on the previous known bounds for $f_3(\mathbb{Z}_4^n)$, the prior record due to Sanders [**San09**] being of the form $f_3(\mathbb{Z}_4^n) \ll \frac{4^n}{n(\log n)^\epsilon}$, for an absolute constant $\epsilon > 0$, and, next to the paper of Bateman and Katz, representing another culmination of the density increment method. Soon after Croot-Lev-Pach, Ellenberg and Gijswijt [**EG17**] used the method from [**CLP17**] to also prove a new

bound for the size of the largest three-term progression free subset of $\mathbb{F}_q^n$, where q is an odd prime power.

$$f_3(\mathbb{F}_q^n) \ll q^{n(1-c_q)}, \tag{1.7}$$

where $c_q > 0$ can be calculated explicitly (and satisfies $c_q = \Theta((\log q)^{-1})$ as q grows large). This new set of ideas has subsequently led to very exciting developments in a vast range of topics. See for instance [**Gro19**] for a somewhat chronological account of some of the earlier applications.

One of the big outstanding questions has naturally become the following:

Question. *Is there any hope to use insights from the recent polynomial method developments to say something new about $r_3(N)$ or to refine further any of the ingredients involved in the Fourier analytic approach?*

This line one inquiry motivated, in some sense, most of the results we shall discuss in the second part of this thesis. I believe that one of the best places to start looking for a connection between the two worlds is once again, perhaps a bit ironically, the problem of determining $f_3(\mathbb{Z}_4^n)$, which was the source of the new revolution to begin with. This is a rather special model problem for the one of determing $r_3(N)$, since, unlike the finite field models, the best lower bound construction does not come from a product construction in the style of Edel [**Ede04**], but rather a Behrend-type example, in the same spirit with the lower bound construction for $r_3(N)$. See [**San09**] or [**PP20**]. Since it is short and sweet, we quickly revisit the construction below (attributed in [**San09**] to an unpublished manuscript of Elsholtz).

Consider a large subset of the grid $\{0,1,2\}^n$ which does not contain three points which lie on the same line, for instance

$$S_n := \{x \in \{0,1,2\}^n : x_i = 1 \text{ for } \lfloor (n+1)/3 \rfloor \text{ values of } i \in [n]\}.$$

It is easy to see that this is indeed a subset of the hypersphere with center at $(1,\dots,1)$ and radius $(n - \lfloor (n+1)/3 \rfloor)^{1/2}$. Furthermore,

$$|S_n| = \binom{n}{\lfloor (n+1)/3 \rfloor} 2^{n-\lfloor (n+1)/3 \rfloor} \approx \left(\frac{9}{4\pi}\right)^{1/2} \cdot \frac{3^n}{\sqrt{n}}$$

as $n \to \infty$, by Stirling's formula. Consider the trivial embedding of $\{0,1,2\}^n$ into $\mathbb{Z}_4^n$ and let the image of S_n be S_n'. We claim that S_n' serves as an example of a subset of $\mathbb{Z}_4^n$ which is free of nontrivial three-term arithmetic progressions. Suppose $x + y = 2z$ occurs for distinct $x, y, z \in S_n' \subset \mathbb{Z}_4^n$. Then, note that there must be some minimal

index $i \in \{1, \ldots, n\}$ where $x_i + y_i = 2z_i$ holds in $\mathbb{Z}_4$ but $x_i + y_i \neq 2z_i$ in $\mathbb{Z}$. Indeed, the preimages of x, y, z would otherwise represent a three-term arithmetic progression of points in S_n, which is impossible by design (since S_n does not contain collinear triples). On the other hand, if $x_i + y_i = 2z_i$ holds in $\mathbb{Z}_4$ but $x_i + y_i \neq 2z_i$ in $\mathbb{Z}$, it must be that

$$(x_i, y_i, z_i) \in \{(2, 2, 0), (0, 0, 2)\}.$$

This is definitely possible; however, note that in this case the new triple $(x, y, z^{(i)})$ in $S'_n \times S'_n \times S'_n$, where $z^{(i)} \in S'_n$ is the vector obtained from z by switching the i-th coordinate from 0 to 2 (or from 2 to 0) projects to a true three-term arithmetic progression in $\mathbb{Z}$ onto the i-th coordinate. It is perhaps worth highlighling the fact that by switching the i-th coordinate of z from 0 to 2 or vice versa, one indeed successfully obtains a point which is still on the sphere S_n, because S_n was chosen to centered at $(1, \ldots, 1)$. At this point, either the preimages of x, y, $z^{(i)}$ represent a three-term arithmetic progression of points in S_n, which is impossible by design (since S_n does not contain collinear triples), or we have some $j \in \{1, \ldots, n\}$ with $j > i$, where $x_j + y_j = 2z_j$ holds in $\mathbb{Z}_4$ but $x_j + y_j \neq 2z_j$ in $\mathbb{Z}$. By repeating the operation, we eventually reach a triple $(x, y, f(z)) \in S'_n \times S'_n \times S'_n$ which comes from a non-trivial three-term arithmetic progression in $\mathbb{Z}^n$, a contradiction.

We have therefore shown that

$$f_3(\mathbb{Z}_4^n) = \Omega\left(\frac{3^n}{\sqrt{n}}\right).$$

It is quite likely that, up to constants, the size of the *every* 3AP-free subset of $\mathbb{Z}_4^n$ is at most $3^n/\sqrt{n}$ for n sufficiently large, i.e. $f_3(\mathbb{Z}_4^n) = \Theta\left(3^n/\sqrt{n}\right)$. Proving a bound of this quality would constitute a major breakthrough.

Improved Bounds for Progression-Free Sets in $\mathbb{Z}_8^n$. In joint work with Fedor Petrov [**PP20**], we studied the question of improving the best bound for the largest three-term progression free set in $\mathbb{Z}_8^n$. Since $\mathbb{Z}_8^n$ is a union of 2^n cosets of a subgroup isomorphic to $\mathbb{Z}_4^n$, a corollary of the Croot-Lev-Pach theorem (1.6) is that $f_3(\mathbb{Z}_8^n) \leq 2^n \cdot (3.611)^n = (7.222)^n$. Using a rather delicate "two layered" polynomial method approach, we managed to improve this bound by an exponential factor, showing that

$$f_3(\mathbb{Z}_8^n) \leq (7.0899)^n.$$

The precise description of the explicit constant is even more complicated than the one for the constant from (1.6), since not only that it involves the binary entropy

function but also a higher order version of it. We include below a formal statement for the record.

Theorem 1.3.2 *If $A \subset \mathbb{Z}_8^n$ is a set without non-trivial three-term progressions, then*

$$|A| \leqslant \left(2 \cdot 2^{\mathcal{H}_4(\rho_0)}\right)^n \approx (7.0899)^n,$$

where

- $\mathcal{H}_4(\rho) = \log_2 \min_{0<x\leq 1} \left\{x^{-3\rho}(1 + x + x^2 + x^3)\right\}$, *and*
- $\rho_0 \approx 0.32$ *solves the system*

$$\mathcal{H}_4(\rho) = \mathcal{H}_2(\theta_1) + \mathcal{H}_2(1 - 2\theta_1),\ \mathcal{H}_4(1 - 2\rho) = 1 + \mathcal{H}_2(1 - 2\theta_1).$$

In our work, we also gave a new proof of (1.6) by further refining the group ring approach started off by Petrov (which previously was only known to work to reprove (1.7) and couldn't handle groups of even order). We believe that improving further on the bound for $f_3(\mathbb{Z}_4^n)$ could be the key to (perhaps more substantial) progress on the quantitative bounds for $r_3(N)$. We will discuss these results in Chapter 4.

Four-term progression free sets rich in 3APs. Getting close to the $N/(\log N)$ threshold with $r_3(N)$ is a delicate problem with a long and beautiful history. There are a few remarkable instances where this barrier was approached and even bypassed, in the presence of additional structure (or randomness). For instance, in 1996, Kohayakawa, Luczak and Rődl [**KLR96**] proved the following random version of Roth's theorem.

Theorem 1.3.3 *There exist a $C > 0$ such that the following holds: for all $p > CN^{-1/2}$, if A is random subset of $[N]$ with the events $x \in A$ being independent with probability $\mathbb{P}(x \in A) = p$, then with probability $1 - o_{N\to\infty}(1)$, we have that*

$$f_3(A) = o(|A|).$$

The proof of Theorem 1.3.3 made use of the so-called Szemerédi regularity lemma, so it provides poor quantitative bound for $f_3(A)$, but it allows one to detect three-term progressions in subsets of $\{1, \ldots, N\}$ with extremely low densities, provided that those subsets have large relative density compared to a random set. Moreover,

the result extends to arbitrary finite abelian groups of odd order. In 2003, Green [**Gre05**] also famously proved that positive density subsets of the set of primes in $\{1, \dots, N\}$ already must always contain three-term progressions, if N is sufficiently large. The Fourier analytic transference principle behind this result then led to a vast amount of exciting developments, such as the Green-Tao theorem [**GT08**]. Among (many) other things, this method was also used by Tao and Vu in [**TV06**] to improve on the variant of Theorem 1.3.3 for $\mathbb{F}_q^n$, which allowed them to provide Meshulam/Roth type quantitative bounds for $f_3(A)$ (but with the price of getting the result only for a very suboptimal range for the probability p).

In joint work with Oliver Roche-Newton [**PRN20**], we pushed this line of work further by putting together the new polynomial method developments with another recent breakthrough in combinatorics, the powerful hypergraph container method (applied in the iterative style of Balogh and Solymosi [**BS19**]). Combining the two, we were able to prove a strong generalization of the Ellenberg-Gijswijt theorem in the random setting.

Theorem 1.3.4 *Let $\beta > 0$, $t < c_q(1 - 2\beta)$ and let p be a positive real number satisfying*

$$q^{n\left(-\frac{1}{2}+\frac{t(C_q-1)}{2}\right)} \leq p \leq 1.$$

Let B be a random subset of $\mathbb{F}_q^n$ with the events $x \in B$ being independent with probability $\mathbb{P}(x \in B) = p$. Then, with probability $1 - o_{n\to\infty}(1)$ we have that

$$f_3(B) \ll pq^{n(1-t+2\beta)}.$$

In particular, for all $\epsilon > 0$, there exists $\delta(\epsilon, q) := \delta > 0$ such that if B is defined as above with $p = q^{n(-\frac{1}{2}+\epsilon)}$, then with probability $1 - o_{n\to\infty}(1)$,

$$f_3(B) \ll |B|^{1-\delta}.$$

This generalization specifically recovers quite precisely the bound (1.7) when $p = 1$, and our result improves the quantitative bound of Tao and Vu for $f_3(A)$ for the finite field variant of Theorem 1.3.3, without losses of any sort for the range of the allowed probability. Also, this result has a couple of interesting consequences that are perhaps worth mentioning. For instance, it is not hard to see that Theorem 1.3.4 implies the following result.

Theorem 1.3.5 *For all $\beta > 0$, there exists $n_0 = n_0(\beta)$ such that the following statement holds for all $n \geq n_0$ and for any prime power q: there exists a 4-AP free set $A \subset \mathbb{F}_q^n$ with the property that $f_3(A) \leq |A|^{1-\frac{1}{2(Cq-2)}+\beta}$.*

That is, there is a set $A \subset \mathbb{F}_q^n$ which does not contain a non-trivial 4-AP but for which every large subset $A' \subset A$ contains a non-trivial 3-AP. The positive constant C_q depends on the aforementioned constant c_q via $C_q = 1 + \frac{1}{c_q}$. For a concrete example, one can calculate that $C_5 \approx 15.12589$, meaning that every set $A' \subset A$ larger than $|A|^{0.962}$ contains a 3-AP. An analogue phenomenon holds in the integers. For instance, for all $\alpha > 0$ and for all $N \in \mathbb{N}$ sufficiently large (depending on α), there exists a set of integers A with $|A| = N$ which does not contain any nontrivial 4-APs, and for which $f_3(A) \ll N/(\log N)^{1-\alpha}$. The upper bound on $f_3(A)$ is of course not of the same quality as the one from 5.0.1, since the latter incorporated (1.7), but it is still maybe interesting since it shows that there exist sets of N integers without non-trivial 4-APs for which the size of the largest 3-AP free subset is smaller than roughly the best upper bound known for $r_3(N)$. It is a bit surprising that such a set A doesn't actually contain that many three-term progressions as one would be inclined to guess, so we were also able to modify the argument of Green and Wolf from [**GW10**] to find asymptotically larger three-term progression-free sets in this special set A than we know exist in $\{1, \ldots, N\}$ (from [**Elk11**]). We will discuss all these results in Chapter 5.

In Chapter 6, we will also discuss a more recent joint work with Jacob Fox [**FP20**], where we continued this investigation and further studied the quantity $f_{s,k}(n)$, defined to be the maximum possible number of s-term arithmetic progressions in a set of n integers which contains no k-term arithmetic progression. For all fixed integers $k > s \geq 3$, we were able to prove that $f_{s,k}(n) = n^{2-o(1)}$, which answered an old question of Erdős from [**Erd73**]. More precisely, our methods established the following upper and lower bounds for $f_{s,k}(n)$, which show that its growth is closely related to the bounds in Szemerédi's theorem.

Theorem 1.3.6 *There exist absolute positive constants c and C such that, for integers $k > s \geq 3$ and every sufficiently large integer n, we have*

$$\left(\frac{c \cdot r_k(n)}{n}\right)^{2(s-2)} \cdot n^2 \leq f_{s,k}(n) \leq \left(\frac{r_k(n)}{n}\right)^{C} \cdot n^2.$$

1.4 On the Croot-Lev-Pach Lemma and its limitations

Besides its broad applicability, the most appealing aspect of the work of Croot, Lev and Pach from [**CLP17**] is that relies on a very simple observation.

Lemma 1.4.1 *Let $P \in \mathbb{F}[x_1, \cdots, x_n]$ be a multilinear polynomial of degree at most d over a field $\mathbb{F}$, and let M denote the $|\mathbb{F}|^n \times |\mathbb{F}|^n$ matrix with entries $M_{\vec{x},\vec{y}} = P(\vec{x}+\vec{y})$ for $\vec{x}, \vec{y} \in \mathbb{F}^n$. Then*

$$\text{rank}_{\mathbb{F}}(M) \leq 2 \sum_{i=0}^{\lfloor d/2 \rfloor} \binom{n}{i}.$$

During the last few years, this has become known as the *Croot-Lev-Pach Lemma*. The claim can be checked as follows: let $m := \sum_{i=0}^{\lfloor d/2 \rfloor} \binom{n}{i}$, and let $\mathcal{K} = \{K_1, \ldots, K_m\}$ be the collection of all sets $K \subset \{1, \ldots, n\}$ with $|K| \leq d/2$. Using the notation,

$$x^I := \prod_{i=1}^{I} x_i, \quad \vec{x} = (x_1, \ldots, x_n) \in \mathbb{F}^n, \ I \subset [n],$$

there exist coefficients $C_{I,J} \in \mathbb{F}$ $(I, J \subset [n])$ depending only on the polynomial P such that for all $x, y \in \mathbb{F}^n$, we have

$$\begin{aligned} P(\vec{x}+\vec{y}) &= \sum_{\substack{I,J \subset [n] \\ I \cap J = \emptyset \\ |I|+|J| \leq d}} C_{I,J} x^I y^J \\ &= \sum_{I \in \mathcal{K}} x^I \sum_{\substack{J \subset [n] \setminus I \\ |J| \leq d - |I|}} C_{I,J} y^I + \sum_{J \in \mathcal{K}} \left(\sum_{\substack{I \subset [n] \setminus J \\ d/2 < |I| \leq d - |J|}} C_{I,J} x^I \right) y^J. \end{aligned}$$

Note that the expression from the right-hand side can be interpreted as the scalar product of the vectors $u(x), v(y) \in \mathbb{F}^{2m}$ defined by

$$u_i(x) = x^{K_i}, \ \ u_{m+i}(x) = \sum_{\substack{I \subset [n] \setminus K_i \\ d/2 < |I| \leq d - |K_i|}} C_{I,K_i} x^I$$

and

$$v_i(y) = \sum_{\substack{J \subset [n] \setminus K_i \\ |J| \leq d - |K_i|}} C_{K_i,J} y^J, \ \ v_{m+i}(y) = y^{K_i}$$

for all $1 \leq i \leq m$. This gives a representation of M as a sum of $2m$ matrices of rank 1, so $\text{rank}_{\mathbb{F}}(M) \leq 2m$, as claimed.

One can then extend the above argument to show that if the field $\mathbb{F}$ is finite and $P \in \mathbb{F}[x_1, \cdots, x_n]$ is an $\mathbb{F}$-linear combination of monomials from the set

$$\mathcal{M}_d(\mathbb{F}, d) = \left\{ x_1^{a_1} \dots x_n^{a_n} \ : 0 \le a_i \le |\mathbb{F}| - 1, \ \sum_{i=1}^{n} a_i = d \right\},$$

then the rank above satisfies $\text{rank}_{\mathbb{F}}(M) \le 2|\mathcal{M}_{\lfloor d/2 \rfloor}(\mathbb{F}, d)$. It is precisely this slight generalization (together with a novel dimension argument) that enabled Ellenberg and Gijswijt in [**EG17**] to prove (1.7), and later others to prove other interesting (and quite unexpected) things.

Besides finding more applications, it is however also important to understand the limitations of the Croot-Lev-Pach method. One of the most tantalizing problems in additive combinatorics has become to extend the existing set of ideas to perhaps say something new about k-term progression free sets in $\mathbb{F}_q^n$, where $q \ge k$. This is a problem where the Croot-Lev-Pach method can be implemented, yet it only produces trivial bounds. See for instance the discussions from from [**Tao**] and [**TS**]. There are quite a few other problems which exhibit similar behavior. In Chapter 7, we will aim to make some progress in this direction by studying a few such problematic applications, where it turns out one can go slightly beyond the Croot-Lev-Pach Lemma by using some additional ideas.

Sets of points $\mathbb{R}^d$ with few distinct distances. Given a positive integer s, a finite subset A in a metric space M is called an s-distance set in M if there are s positive real numbers $d_1, \dots, d_s$ such that all the pairwise distances determined by the points in M are among these numbers, and each d_i is realized. Upper bounding the size of such sets is a famous problem in combinatorial geometry, with a lot of activity around the various possible variants. See for instance [**GY18**] and the references therein. When M is $\mathbb{R}^d$, with the usual Euclidean distance, the classical result in the area is the following beautiful result due to Bannai, Bannai and Stanton [**BBS83**] from 1983.

Theorem 1.4.2 *If A is an s-distance subset in $\mathbb{R}^d$, then*

$$|A| \le \binom{d+s}{s}.$$

The proof of Theorem 1 from [**BBS83**] builds upon the linear independence argument introduced for this problem by Larman, Rogers and Seidel in [**LRS77**]. In

[**LRS77**], the authors proved that when $s = 2$, the inequality $|A| \leq (d+1)(d+4)/2$ follows from the fact that to each a point in A one can associate a polynomial $f_a \in \mathbb{R}[x_1, \ldots, x_d]$ such that $\{f_a, a \in A\}$ is a set of linearly independent polynomials over the reals, which also happens to lie in a subspace of $\mathbb{R}[x_1, \ldots, x_d]$ of dimension $(d+1)(d+4)/2$. This argument was later amplified by Blokhuis [**Blo81**] who showed that one can further add a list of $d+1$ other polynomials to $\{f_a : a \in A\}$ and get an even larger list of linearly independent polynomials that lie in the same vector space of dimension $(d+1)(d+4)/2$. This led to $|A| \leq (d+1)(d+4)/2 - (d+1) = \binom{d+2}{2}$, which established the important first case $s = 2$ of Theorem 7.1.1. This story was successfully generalized by Bannai-Bannai-Stanton in [**BBS83**], but for larger s the argument to show that one can add a new list of (higher degree) polynomials to the old list and still get a set of linearly independent elements in the same vector space is significantly more technical.

We give a new simple proof of Theorem 7.1.1 via an improved version of the so-called Croot-Lev-Pach Lemma [**CLP17**] over the reals, which may be of independent interest and could possibly have other applications. We state this in a general form, which captures the original version of the Croot-Lev-Pach Lemma as well.

Theorem 1.4.3 *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and $A \subset V$ be a finite set. Let s be a nonnegative integer and let $p(\overrightarrow{x}, \overrightarrow{y})$ be a $2 \cdot \dim V$-variate polynomial wih coefficients in $\mathbb{F}$ and of degree at most $2s+1$. Consider the matrix $M_{p,A}$ with rows and columns indexed by A and entries $p(\cdot,\cdot)$. It corresponds to a (not necessary symmetric) bilinear form on $\mathbb{F}^A$ by a formula*

$$\Phi_p(f,g) = \sum_{a,b \in A} p(\overrightarrow{a}, \overrightarrow{b}) f(a) g(b), \text{ for } f, g : A \to \mathbb{F},$$

which in turn defines a quadratic form $\Phi_p(f,f)$. Denote by $\operatorname{rank}(p, A)$ the rank of matrix $M_{p,A}$; if $\mathbb{F} = \mathbb{R}$ denote also by $r_+(p), r_-(p)$ the inertia indices of the quadratic form $\Phi_p(f,f)$. Finally, denote by $\dim_s(A)$ the dimension of the space of polynomials of degree at most s considered as functions on A. Then:

1) $\operatorname{rank}(p, A) \leqslant 2 \dim_s(A)$.

2) if $\mathbb{F} = \mathbb{R}$, *then* $\max\{r_+(p, A), r_-(p, A)\} \leqslant \dim_s(A)$.

We prove this Theorem in Chapter 7. This is based on a recent joint work with Fedor Petrov [**PP20**].

Subsets of the hypercube with prescribed Hamming distances. To put the story into some context, we begin with a rough version of the isodiametric inequality which states that in $\mathbb{R}^n$, among all bodies of a given diameter, the n-dimensional ball has the largest volume. Solving a conjecture of Erdős, in [**Kle86**] Kleitman proved the following important theorem in extremal set theory, which can be thought of as an analogue of the isodiametric inequality in the discrete setting.

Theorem 1.4.4 *Suppose $\mathcal{F}$ is a collection of binary vectors in $\{0,1\}^n$, such that the Hamming distance between any two vectors is at most $d < n$. Then*

$$|\mathcal{F}| \leq \begin{cases} \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t}, & \text{for } d = 2t; \\ 2\left(\binom{n-1}{0} + \cdots + \binom{n-1}{t}\right), & \text{for } d = 2t+1. \end{cases}.$$

In addition to being an interesting result in its own right, it is perhaps important to highlight that Theorem 1.4.4 also represents the main input in several surprising constructions which show all sorts of different results across combinatorics are optimal in a certain way. For example, in [**KS03**] Kostochka and Sudakov use Theorem 1.4.4 to show that for infinitely many n there is a graph G on n vertices with at least $n(n-2)/4$ edges such that any subset of G of linear size contains a pair of vertices with at most $o(n)$ common neighbors (thus providing evidence that the main dependent random choice lemma from probabilistic combinatorics cannot be improved in some sense; see [**FS11**] for more details); in [**MS96**], Matousek and Spencer use Kleitman's theorem to provide an example of a two coloring of the first n positive integers such that all the subsets which are arithmetic progressions have discrepancy $O(n^{1/4})$, thus proving that a classical result of Roth is best possible, up to constants.

In joint work with Hao Huang and Oleksiy Klurman [**HKP20**], we gave a new proof of Kleitman's theorem, based on the following bound by Cvetković: the independence number $\alpha(G)$ is bounded from above by the number of non-negative (resp. non-positive) eigenvalues of the adjacency matrix of G. This result, when extended to pseudo-adjacency matrices, is still correct. A careful choice of a proper pseudo-adjacency matrix (intimiately connected with tensors appearing in Croot-Lev-Pach based approaches for certain controlled variants of the problem) then led to an algebraic proof of Kleitman's isodiametric theorem. This method also allowed us to prove a slightly more general version of Kleitman's result, when the allowed

Hamming distances between two vectors lie in an arbitrary interval. In particular, we showed the following new estimate.

Theorem 1.4.5 *For given integers $t > s \geq 0$, suppose $\mathcal{F}$ is a collection of binary vectors in $\{0,1\}^n$, such that for every $x, y \in \mathcal{F}$, $d(x,y) \in \mathcal{L}$, with $\mathcal{L} = \{2s+1, \cdots, 2t\}$, then for n sufficiently large,*

$$|\mathcal{F}| \leq \binom{n}{t-s} + \binom{n}{t-s+1} + \cdots + \binom{n}{0}.$$

Similarly, if $\mathcal{L} = \{2s+1, \cdots, 2t+1\}$, then $|\mathcal{F}| \leq (2+o(1))\binom{n}{t-s}$.

Subsets of the hypercube with Hamming distances in a prescribed set of consecutive integers appear in the coding theory literature in the regime when s and t are linear in n, for instance in the context of ϵ-balanced codes (of length n). These are subsets $\mathcal{F}$ of $\{0,1\}^n$ with pairwise Hamming distances between $\frac{1-\epsilon}{2}n$ and $\frac{1+\epsilon}{2}n$. By mapping them to vectors on the unit sphere in $\mathbb{R}^n$ via

$$(v_1, \ldots, v_n) \mapsto \frac{1}{\sqrt{n}} \cdot ((-1)^{v_1}, (-1)^{v_2}, \ldots, (-1)^{v_n}) \in \left\{-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right\}^n,$$

one can easily note that in this case estimating $|\mathcal{F}|$ amounts to estimating the length of a certain spherical code, for which other methods are useful. We refer to [**Alo09**] for more details. For our general range (in particular when s and t are small compared to n), the problem of upper bounding $|\mathcal{F}|$ is of a different nature, and results about spherical codes do not apply.

Subsets of $\mathbb{F}_p^N$ without differences in $\{0,1\}^N$. Using these techniques, in our paper [**HKP20**] we also considered a somewhat different extremal set theory problem, which in fact falls much closer to the realm of additive combinatorics. Before I describe our result, we introduce some terminology (and a little context). A set $H \subset \mathbb{Z}^+$ is *i*ntersective if whenever A is a subset of positive upper density of $\mathbb{Z}$, we have $(A-A) \cap H \neq \emptyset$. In the late 1970s, Sárközy [**S\'ar78**], and independently Furstenberg [**Fur77**, **Furr81**], proved that the set of nonzero perfect squares is intersective. A quantitative version has also been considered. Denote by $D(H,N)$ the maximum size of a subset $A \subset \{1, \cdots, N\}$ such that $(A-A) \cap H = \emptyset$. It is not hard to see that a set is intersective if and only if $D(H,N) = o(N)$. The order of growth of the function $D(H,N)$ has been intensely studied for many intersective sets H; we refer the readers to a survey of Lê [**Le14**].

One particular development which motivated our line of inquiry was a recent result of Green for a variant of Sarkőzy's theorem in function fields [**Gre16**]. Using the recent Croot-Lev-Pach Lemma, Green gave a short proof that if $k \geq 2$ and $A \subset \mathbb{F}_q[x]_{\deg<n}$ contains no distinct $p, q \in A$ such that $p - q = b^k$, then $|A| \ll q^{(1-c_{k,q})n}$, for some explicit constant $c_{k,q}$. In [**HKP20**], we wanted to gain over Croot-Lev-Pach bounds for a different intersective set. Let $J = \{0,1\}^N$ and define $D_{\mathbb{F}_p}(J,N)$ to be the maximum cardinality of $A \subset \mathbb{F}_p^N$ such that $(A-A) \cap J = \emptyset$. It is not too difficult to use Sperner's Theorem to prove that $D_{\mathbb{F}_p}(J,N) = o(p^N)$. In [**Le14**], it is mentioned that Alon improved on this bound by showing the following result:

$$\frac{(p-1)^N}{p\sqrt{N}} \ll D_{\mathbb{F}_p}(J,N) \leq (p-1)^N.$$

The second inequality used the polynomial method and in fact can be reproved by using the Croot-Lev-Pach Lemma from above. In our work, we used our spectral method to slightly improve on Alon's upper bound.

Theorem 1.4.6

$$D_{\mathbb{F}_p}(J,N) \leq \left(1 - \frac{1}{2}\left(1 - \frac{1}{p-1}\right)^p\right)(p-1)^N.$$

It is perhaps important to mention that, beyond our applications, our main contribution from [**HKP20**] consists in a rather fresh perspective on how to construct pseudoadjacency matrices for which the so-called Cauchy interlacing theorem gives new information (in our case to specifically beat Croot-Lev-Pach based bounds). This recently helped inspire Huang's recent breakthrough paper on the sensitivity conjecture (see for instance [**Bis**]), and might have more potential for applications.

Chapter 2

ON THE ERDŐS-GYÁRFÁS PROBLEM IN GRAPH RAMSEY THEORY

In this chapter, we will prove the theorems mentioned in Section 1.1. We begin by recalling the definition of the main function that will appear throughout this chapter: given positive integers n, k, and ℓ, $f(n, k, \ell)$ denotes the minimum possible number of colors in a coloring of K_n with the property that every induced subgraph over k vertices contains at least ℓ colors.

To set things up, we also recall a few quick things from the introduction. For example, $f(n, 3, 3)$ denotes the minimum consider edge colorings of K_n where every triangle contains three distinct colors. In particular, no vertex can be adjacent to two edges with the same color, so we immediately have that $f(n, 3, 3) \geq n - 1$.

As another important example, recall that for any $k \geq 4$ we had that

$$f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 2\right) = \Theta\left(n^2\right). \tag{2.1}$$

Indeed, in this case every color can occur at most $\lfloor k/2 \rfloor - 1$ times, since otherwise we will have an induced subgraph with k vertices and at most $\binom{k}{2} - \lfloor k/2 \rfloor + 1$ colors. Since each color repeats a constant number of times, there are $\Theta\left(n^2\right)$ distinct colors. Erdős and Gyárfás [**EG97**] asked what happens when we move one away from the quadratic threshold. That is, they studied the case of $\ell = \binom{k}{2} - \lfloor \frac{k}{2} \rfloor + 1$, and derived the bound $f(n, k, \ell) = \Omega(n^{4/3})$. This chapter is in some sense a collection of stories around this question.

In Section 2.1, we will first start with Theorem 1.1.1, which we recall for the reader's convenience.

Theorem 2.0.1 *For any integers $k > m \geq 2$,*

$$f\left(n, k, \binom{k}{2} - m \cdot \left\lfloor \frac{k}{m+1} \right\rfloor + m + 1\right) = \Omega\left(n^{1+\frac{1}{m}}\right).$$

To get a sense of the statement, note that by applying Theorem 2.0.1 with $m = 2$ we get

$$f\left(n, k, \binom{k}{2} - 2 \cdot \left\lfloor \frac{k}{3} \right\rfloor + 3\right) = \Omega\left(n^{3/2}\right).$$

In [**EG97**], Erdős and Gyárfás derived a bound of $\Omega(n^{4/3})$ colors when ℓ is one unit away from the trivial case. Theorem 2.0.1 implies this bound already when $\ell \geq \binom{k}{2} - 3 \cdot \lfloor k/4 \rfloor + 4$.

As mentioned in Section 1.1, the bound of Theorem 2.0.1 is asymptotically tight up to sub-polynomial factors for every $m \geq 2$, except possibly for small values of k. Indeed, for every $\varepsilon > 0$ the Lovász Local Lemma upper bound of (1.2) implies

$$f\left(n, k, \binom{k}{2} - m \cdot \left\lfloor \frac{k}{m+1} \right\rfloor + m + 1\right) = O\left(n^{1+\frac{1}{m}+\varepsilon}\right),$$

for every sufficiently large k.

Our proof technique. To prove Theorem 2.0.1, we define a new abstract variant of the concept of additive energy, which is a main tool in additive combinatorics.

Given a finite $A \subset \mathbb{R}$, the *sum set* of A is

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

A uniformly chosen set A of a fixed finite size is expected to satisfy $|A+A| = \Theta(|A|^2)$. On the other hand, there are sets that satisfy $|A + A| = \Theta(|A|)$, such as arithmetic progressions. We say that such sets have an "additive structure" that leads to a small sum set. The *polynomial Freiman–Ruzsa* conjecture is a main open problem in Additive Combinatorics, asking to characterize the sets that have a small sum set.

One main tool for studying the additive structure of a finite $A \subset \mathbb{R}$ is the *additive energy* of A:

$$E(A) = \left|\left\{(a, b, c, d) \in A^4 : \ a + b = c + d\right\}\right|.$$

While the size of the sum set of A is at least linear in $|A|$ and at most quadratic in $|A|$, the additive energy of A is at least quadratic in $|A|$ and at most cubic in $|A|$.

A small sum set implies a large additive energy. In particular, a simple use of the Cauchy-Schwarz inequality implies $E(A) \geq \frac{|A|^4}{|A+A|}$. In the other direction, the *Balog–Szemerédi–Gowers theorem* implies that if $E(A)$ is large then there exists a large subset $A' \subset A$ such that $A' + A'$ is small. For more information, see for example the book "Additive Combinatorics" by Tao and Vu [**TV06**].

For Theorem 2.0.1, we define the following abstract graph variant of additive energy. Given a graph $G = (V, E)$ with colored edges, we denote by $c(u, v)$ the color of the

edge $(u, v) \in E$. We define the *color energy* of G as[1]

$$\mathbb{E}(G) = \left|\left\{(v_1, u_1, v_2, u_2) \in V^4 : \ c(v_1, u_1) = c(v_2, u_2)\right\}\right|.$$

That is, instead of the number of quadruples that satisfy an additive relation, we ask for the number of quadruples that satisfy a color relation.

There exist energy variants for Cayley graphs that are based on the corresponding group action (for example, see Gowers [**Gow08**]). However, as far as we know this is the first use of such a non-algebraic energy variant. Using this technique, we will also show in Section 2.2 that $f(n, k, \ell)$ becomes arbitrarily close to n^2 as k grows, when $\ell = \binom{k}{2} - \lfloor \frac{k}{2} \rfloor + 2$, namely

Theorem 2.0.2 *For every $k \geq 8$, if $r = \lfloor k/4 \rfloor$ then*

$$f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 1\right) = \Omega(n^{2-2/r}).$$

We will prove Theorem 2.0.2 by extending the color energy technique from the proof of Theorem 2.0.2. In Section 2.3, we push the methods further in the arithmetic setup and use so-called "higher order color energies" to prove Theorem 1.1.5.

2.1 Color energy and forbidden repeats of rainbow stars

We begin with the proof of Theorem 2.0.1, as promised. To prove the result, we will rely on the following simple counting lemma (see [**Erd64**] and [**Juk11**]).

Lemma 2.1.1 *Let A be a set of n elements and let $d \geq 2$ be an integer. Let $A_1, \ldots, A_k$ be subsets of A, each of size at least m. If $k \geq 2dn^d/m^d$ then there exist $1 \leq j_1 < \ldots < j_d \leq k$ such that $|A_{j_1} \cap \ldots \cap A_{j_d}| \geq \frac{m^d}{2n^{d-1}}$.*

Proof of Theorem 2.0.1. To prove the theorem we will prove that for any integers $a, b \geq 2$

$$f\left(n, a(b+1), \binom{a(b+1)}{2} - ba + b + 1\right) = \Omega\left(n^{1+\frac{1}{b}}\right). \tag{2.2}$$

When k is divisible by $m+1$, we obtain the statement of the theorem by setting $b = m$ and $a = k/(m+1)$. When k is not divisible by $m+1$, we write $a = \lfloor k/(m+1) \rfloor$,

[1] We use $\mathbb{E}(G)$ rather than $E(G)$ since the latter is a standard notation for the set of edges of G.

and rewrite (2.2) as $f\left(n, k, \binom{k}{2} - ba + b + 1\right) = \Omega\left(n^{1+\frac{1}{b}}\right)$. The rest of the proof is the same in both cases. However, when reading the proof for the first time we recommend assuming that k is divisible by $m + 1$.

Let $G = (V, E)$ be a copy of K_n with colored edges, such that every induced $K_{a(b+1)}$ contains at least $\binom{a(b+1)}{2} - ba + b + 1$ distinct colors. We denote the set of colors as $C = \{c_1, c_2, \ldots\}$, and the color of an edge $(v, u) \in E$ as $c(v, u)$. Our goal is to prove that $|C| = \Omega\left(n^{1+1/b}\right)$, and we begin by studying some configurations that cannot occur in G.

Forbidden configurations. We first show that no vertex can be adjacent to many edges of the same color. Assume for contradiction that there exists a vertex $v \in V$ and color $c \in C$, such that at least $ba - b + 1$ vertices $u \in V$ satisfy $c(v, u) = c$. Let $V' \subset V$ consist of v, of $ba - b + 1$ vertices satisfying $c(v, u) = c$, and of $b + a - 2$ additional vertices. Then V' is a set of $a(b + 1)$ vertices, and the induced subgraph on V' contains at most $\binom{a(b+1)}{2} - ba + b$ distinct colors, contradicting the assumption on G. This contradiction implies that for every $v \in V$ and $c \in C$, at most $ba - b$ of the edges incident to v have color c. This in turn implies that every color appears at most $b(a - 1)n/2$ times.

We next show that there cannot be a vertices that are adjacent to the same b "popular" colors. Let $V_c \subset V$ be the set of endpoints of edges of color c. For an integer j, let C_j be the set of colors that appear at least 2^j times. For j with $2^j \geq a$, assume for contradiction that there exist $c_1, \ldots, c_b \in C_j$ that satisfy $|V_{c_1} \cap \cdots \cap V_{c_b}| \geq a$. Let $V' = \{v_1, v_2, \ldots, v_a\}$ be a set of a vertices from $V_{c_1} \cap \cdots \cap V_{c_b}$. That is, for every vertex $v \in V'$ and every color $c \in \{c, \ldots, c_b\}$ there exists $u \in V$ satisfying $c(v, u) = c$. An example is depicted in Figure 2.1.

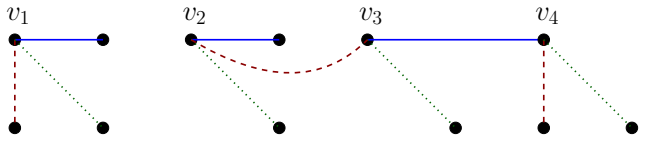

Figure 2.1: The three colors are represented as solid, dashed, and dotted edges. Every vertex of $V' = \{v_1, v_2, v_3, v_4\}$ is incident to an edge of every color.

We construct a subset $V'' \subset V$ as follows: For every $v \in V'$ and color $c \in C_j$, if there exist vertices $u \in V \setminus V'$ such that $c(v, u) = c$ then we add one such vertex

to V''. If a vertex $u \in V \setminus V'$ was added more than once to V'', we consider V'' as containing a single copy of u (that is, V'' is not a multiset). Note that if for $v \in V'$ and $c \in C_j$ there is no $u \in V \setminus V'$ satisfying $c(v,u) = c$ then there exists $u \in V'$ satisfying $c(v,u) = c$. For $1 \leq i \leq b$, let r_i denote the number of vertices of V' that are not connected to any vertex of $V \setminus V'$ with an edge of color c_i. For every $1 \leq i \leq b$, at least $r_i/2$ edges in the induced subgraph of V' have color c_i. Thus, $V^* = V' \cup V''$ is a set of at most $(b+1)a - \sum_{i=1}^{b} r_i$ vertices of V and at least $a - r_i/2$ edges with color c_i. For $1 \leq i \leq b$, we add $\lfloor r_i/2 \rfloor$ additional edges of color c_i by adding at most r_i vertices of V to V^*. This is always possible since c_i is in C_j and $2^j \geq a$. Since the resulting set V^* contradicts the assumption about G, no distinct $c_1, \ldots, c_b \in C_j$ satisfy $|V_{c_1} \cap \cdots \cap V_{c_b}| \geq a$.

Popular colors. Let $k_j = |C_j|$. We now derive two upper bounds for k_j.

Fix an integer j satisfying $2^j \geq b(2a^{b+1}n^{b-1})^{1/b}$ and let $c \in C_j$. Since a single vertex is incident to at most $ba - b$ edges with color c, we get that $|V_c| \geq \frac{2^{j+1}}{ba-b} \geq \frac{2^j}{ba}$. Since no b sets V_c have a large intersection, we use the contrapositive of Lemma 2.1.1 to obtain that the number of sets is not large. By the lower bound for 2^j, every b sets V_c intersect in less than $a \leq \frac{2^{jb}}{2a^b b^b n^{b-1}}$ vertices. Since the size of each such set is at least $m = \frac{2^j}{ba}$, the contrapositive of the lemma implies that the number of sets is

$$k_j < \frac{2bn^b}{m^b} = \frac{2bn^b}{(2^j/ba)^b} = \frac{2n^b b^{b+1} a^b}{2^{jb}}. \tag{2.3}$$

When $2^j < b(2a^{b+1}n^{b-1})^{1/b}$ we will rely on the straightforward bound

$$k_j < n^2/2^j. \tag{2.4}$$

An energy argument. Let m_c be the number of edges with color c. In the beginning of the forbidden configurations part above, we proved that $m_c \leq b(a-1)n$ for every $c \in C$. Since every edge contributes to exactly one m_c, we have $\sum_{c \in C} m_c = \binom{n}{2}$.

Recall that the color energy of G is defined as

$$\mathbb{E}(G) = \left| \left\{ (v_1, u_1, v_2, u_2) \in V^4 : \; c(v_1, u_1) = c(v_2, u_2) \right\} \right|.$$

In other words, $\mathbb{E}(G)$ is the number of pairs of edges with the same color. Since the graph is undirected, when computing $\mathbb{E}(G)$ we consider (v_1, u_1, v_2, u_2) and (u_1, v_1, v_2, u_2) as the same quadruple. On the other hand, we count (v_1, u_1, v_2, u_2)

and (v_2, u_2, v_1, u_1) as two separate quadruples. We can also think of $\mathbb{E}(G)$ as the square of the ℓ_2-norm of the color frequencies, since

$$\mathbb{E}(G) = \sum_{c \in C} m_c^2.$$

By the Cauchy-Schwarz inequality, we have

$$\mathbb{E}(G) = \sum_{c \in C} m_c^2 \geq \frac{\left(\sum_{c \in C} m_c\right)^2}{|C|} = \frac{\binom{n}{2}^2}{|C|} = \Omega\left(\frac{n^4}{|C|}\right).$$

Let $t = \lfloor \log b(2a^{b+1}n^{b-1})^{1/b} \rfloor$. By dyadic pigeonholing together with (2.3) and (2.4), we obtain

$$\begin{aligned}
\mathbb{E}(G) = \sum_{c \in C} m_c^2 &= \sum_{j=0}^{\log(ban)} \sum_{\substack{c \in C \\ 2^j \leq m_c < 2^{j+1}}} m_c^2 < \sum_{j=0}^{\log(ban)} k_j \left(2^{j+1}\right)^2 \\
&= \sum_{j=0}^{t} k_j 2^{2j+2} + \sum_{j=t+1}^{\log(ban)} k_j 2^{2j+2} \\
&= \sum_{j=0}^{t} \frac{n^2}{2^j} \cdot 2^{2j+2} + \sum_{j=t+1}^{\log(ban)} \frac{2n^b b^{b+1} a^b}{2^{jb}} \cdot 2^{2j+2} \\
&= O\left(n^{3-1/b}\right) + O\left(n^{3-2/b} \log n\right) = O\left(n^{3-1/b}\right).
\end{aligned}$$

The $\log n$ in the last line exists only when $b = 2$, but it does not affect the final bound in any case. Combining the two above bounds on $\mathbb{E}(G)$ yields $|C| = \Omega(n^{1+1/b})$, as required. This completes the proof.

Remark. One way to improve the proof of Theorem 2.0.1 might be to derive an upper bound on k_j stronger than the straightforward bound of (2.4) when $2^j \approx n^{(b-1)/b}$.

2.2 Higher order color energy and forbidden cycles

In this section, we will prove Theorem 2.0.2 and Theorem 1.1.5, after [**FPS20**]. The arguments demonstrate how color energy can be combined with upper bounds on $\mathrm{ex}(n, C_{2r})$, the extremal number of the $2r$-cycle. We will first require a simple lemma, obtained by using straightforward probabilistic arguments (see for example [**AS04**]). We thus only provide a brief proof sketch.

Lemma 2.2.1 *Consider a graph $G = (V, E)$ with $|V| = n$, for a sufficiently large n.*

(a) The set V can be partitioned into disjoint sets V_1, V_2 such that $|V_1| = \lceil n/2 \rceil, |V_2| = \lfloor n/2 \rfloor$, and at least $|E|/3$ of the edges of E do not have both of their endpoints in the same V_j.

(b) For an integer $r \geq 2$, let $T \subset E^r$. Then V can be partitioned into disjoint sets $V_1, \ldots, V_r$, each of size $\lceil n/r \rceil$ or $\lfloor n/r \rfloor$, such that $\Omega_r(|T|)$ of the tuples of T contains only edges having both of their endpoints in the same V_j.

Proof Sketch of Lemma 2.2.1. (a) Pick a uniform random partition of V among the set of partitions satisfying $|V_1| = \lceil n/2 \rceil$ and $|V_2| = \lfloor n/2 \rfloor$. Let X denote the number of edges of E that do not have both of their endpoints in the same V_j. Since the expected size of X is larger than $|E|/3$, there exists at least one partition for which $X > |E|/3$.

(b) Pick a uniform random partition $V_1, \ldots, V_r$ of V among the set of partitions into r parts of size $\lceil n/r \rceil$ or $\lfloor n/r \rfloor$. Let X denote the number of r-tuples of T that contain only edges with both their endpoints in the same V_j. It is not difficult to show that the probability of $2r$ specific vertices being in a specific V_j is at least $(4r)^{-2r}$. This can in turn be used to show that the expected size of X is larger than $|T|(4r)^{-2r}$. This proves Lemma 2.2.1.

Higher energy graphs. Like before, consider a copy of K_n denoted as $G = (V, E)$ and a coloring $\chi : E \to C$. In general, for an integer $r \geq 2$, we define the *r-th color energy* of the coloring as

$$\mathbb{E}_r(\chi) = \left|\left\{(a_1, a_2, \cdots, a_{2r}) \in V^{2r},\ \chi(a_1, a_2) = \chi(a_3, a_4) = \cdots = \chi(a_{2r-1}, a_{2r})\right\}\right|.$$

For a color $c \in C$, we set

$$m_c = \left|\left\{(v_1, v_2) \in V^2 :\ v_1 \neq v_2 \text{ and } \chi(v_1, v_2) = c\right\}\right|.$$

Since every ordered pair of distinct vertices in V^2 contributes to m_c for exactly one c, we get that $\sum_{c \in C} m_c = n(n-1)$. The number of $2r$-tuples that contribute to $\mathbb{E}_r(\chi)$ and correspond to the color c is exactly m_c^r. This implies that $\mathbb{E}_r(\chi) = \sum_{c \in C} m_c^r$. By Hölder's inequality,

$$\mathbb{E}_r(\chi) = \sum_{c \in C} m_c^r \geq \frac{\left(\sum_{c \in C} m_c\right)^r}{\left(\sum_{c \in C} 1\right)^{r-1}} = \frac{n^r (n-1)^r}{|C|^{r-1}}. \tag{2.5}$$

Note that the "standard" color energy $\mathbb{E}(\chi)$ is the second color energy $\mathbb{E}_2(\chi)$. By (2.5), to obtain a lower bound for the number of colors, it suffices to derive an upper bound for $\mathbb{E}_r(\chi)$.

By Lemma 2.2.1(b), there exists a partition of V into r disjoint subsets $V_1, \ldots, V_r$, each of size $\Theta(n)$, with the following property. When removing from E every edge that does not have both of its endpoints in the same V_j, the energy $\mathbb{E}_r(\chi)$ does not change asymptotically. (That is, after removing an edge $(u, v) \in E$, we remove from $\mathbb{E}_r(\chi)$ the contribution coming from $2r$-tuples that involve $\chi(u, v)$.) We indeed remove from G every such edge.

An *r-th energy graph* of G and χ, denoted $\hat{G}^r(\chi)$, is defined as follows. Each such graph corresponds to a different partition of V into $V_1, \ldots, V_r$ of the form described in Lemma 2.2.1(b). The set of vertices is $V(\hat{G}^r) = V_1 \times V_2 \times \cdots \times V_r$. An edge between $(v_1, \ldots, v_r), (v'_1, \ldots, v'_r) \in V_1 \times \cdots \times V_r$ is in $E(\hat{G}^r)$ if and only if $\chi(v_1, v'_1) = \chi(v_2, v'_2) = \cdots = \chi(v_r, v'_r)$. Note that $\mathbb{E}_r(\chi) = \Theta(|E(\hat{G}^r)|)$. Thus, to obtain a lower bound for the number of colors, we can derive an upper bound on the number of edges in an r-th energy graph of K_n.

We remove "unpopular" colors from $\hat{G}^r$, as follows. Every edge $e \in E(\hat{G}^r)$ corresponds to several edges of E that have the same color. We say that e also has this color. For every color $c \in C$ that appears in E at most $\log n$ times, we remove from $E(\hat{G}^r)$ every edge that is associated with c. Note that every such color is associated with $O(\log^r n)$ edges of $E(\hat{G}^r)$. Since $|C| = O(n^2)$, this step removes $O(n^2 \log^r n)$ edges from the r-th energy graph. This number is too small to have an effect on any of our proofs. We refer to the resulting graph as a *pruned r-th energy graph* of G, and denote it as G^r.

We are now ready to move on to the proof of Theorem 2.0.2, which we restate below for the reader's convenience.

Theorem 2.0.2. *For every $k \geq 8$, if $r = \lfloor k/4 \rfloor$ then*

$$f\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 1\right) = \Omega(n^{2-2/r}).$$

Proof of Theorem 2.0.2. Consider a complete graph K_n denoted as $G = (V, E)$ and a coloring $\chi : E \to C$, such that every induced K_k contains at least $\binom{k}{2} - k/2 + 1$

colors. Consider also the color energy $\mathbb{E}_2(\chi)$, as defined and pruned on the previous page. By (2.5), to obtain a lower bound for $|C|$ it suffices to derive an upper bound for $\mathbb{E}_2(\chi)$. Let G^2 be a pruned second energy graph of χ. By the above discussion, it suffices to derive an upper bound on $|E(G^2)|$. The definition of G^2 relies on a partition of V into parts V_1, V_2, but we will not rely on this property in the current proof.

To bound $|E(G^2)|$, we wish to apply Theorem 2.2.2 on G^2. For this purpose, we assume for contradiction that G^2 contains a simple cycle γ of length $2r$. We write $\gamma = (a_1, b_1), \cdots, (a_{2r}, b_{2r})$, where the vertices $a_1, \ldots, a_{2r}, b_1, \ldots, b_{2r} \in V$ may not be distinct. Let S be the set of these vertices, so $|S| \leq 4r \leq k$.

For some intuition, we first consider the case where S consists of $2r$ distinct vertices of V. If necessary, we arbitrarily add vertices to S until $|S| = k$. For every $1 \leq j \leq 2r$, the edge between (a_j, b_j) and (a_{j+1}, b_{j+1}) implies that $\chi(a_j, a_{j+1}) = \chi(b_j, b_{j+1})$ (where $a_{2r+1} = a_1$ and $b_{2r+1} = b_1$). This in turn implies that the number of distinct colors spanned by the vertices of S is at most $\binom{k}{2} - 2r \leq \binom{k}{2} - \lfloor k/2 \rfloor$. We obtained a contradiction to the assumption that every k vertices of V span at least $\binom{k}{2} - \lfloor k/2 \rfloor + 1$ colors.

We next consider the general case, where S might contain fewer than $2r$ vertices of V. We go one-by-one over the edges of γ. In particular, in the j-th step we consider the edge between (a_j, b_j) and (a_{j+1}, b_{j+1}) (as before, $a_{2r+1} = a_1$ and $b_{2r+1} = b_1$). At each step, we have $\chi(a_j, a_{j+1}) = \chi(b_j, b_{j+1})$ and this is either a new color repetition or a repetition we already counted in one of the previous steps. If we are in the latter case, that means that both a_j and b_j already appeared in previous edges of the cycle.

Let m mark the number of steps in which we did not find a new color repetition. In other words, there are at least $2r - m$ distinct color repetitions. In each of the m steps without a new repetition we also had two repeating vertices, so $|S| \leq 2r - 2m$. Let $c = \chi(a_1, a_2)$. We add to S the endpoints of m more edges with color c, and note that $|S| \leq 2r$. By the definition of G^2, the color c has at least $\log n$ edges associated with it, so this is always possible. If necessary, we add to S additional arbitrary vertices until it is of size k. Since the vertices of S span at most $\binom{k}{2} - 2r \leq \binom{k}{2} - \lfloor k/2 \rfloor$ colors, we again obtain a contradiction.

The above contradiction implies that G^2 does not contain a cycle of length $2r$. At this point, it is important to recall a classical result about graphs with no cycles of a given length.

Theorem 2.2.2 *For every integer $k \geq 2$, we have that* $\mathrm{ex}(n, C_{2k}) = O(n^{1+1/k})$.

This was originally stated by Erdős [**Erd63**] without proof. For an elegant proof, see the following nice note of Naor and Verstraete [**NV05**]. By Theorem 2.2.2, we obtain

$$\mathbb{E}_2(\chi) = O(|E(G^2)|) = O\left(\mathrm{ex}(V(G^2), C_{2r})\right) = O\left(\left(n^2\right)^{1+1/r}\right) = O(n^{2+2/r}).$$

Combining this with (2.5) implies the asserted bound $|C| = \Omega(n^{2-2/r})$. This completes the proof of Theorem.

Remark. One may wonder what happens when, in the proof of Theorem 2.0.2, we replace Theorem 2.2.2 with a bound on the extremal number of some other subgraph H. With a slightly modified analysis, one tends to get result for various other parameters. For instance, if one forbids a subdivision H_t in the energy graph instead of a cycle, this change leads to the inequality $f\left(n, k, \binom{k}{2} - k + 2\sqrt{k} + 1\right) = \Omega\left(n^{1+1/(2\sqrt{k}-2)}\right)$. This bound is subsumed by a result of Sárközy and Selkow [**SS01**].

Improved bounds for difference sets with local properties. The goal of this subsection is to prove Theorem 1.1.5, which we restate below for the reader's convenience.

Theorem 1.1.5. *For $k > r \geq 2$, let $\ell = \binom{2rk}{2} - \binom{2k}{2} \cdot \left[\binom{r}{2} + (r-1)\right] + 1$. Then*

$$g(n, 2rk, \ell) = \Omega\left(n^{\frac{r}{r-1} \cdot \frac{k-1}{k}}\right).$$

Proof of Theorem 1.1.5. Let A be a set of n real numbers such that every subset $A' \subset A$ of size $2rk$ satisfies $|A' - A'| \geq \ell$. Let $G = (V, E)$ be a copy of K_n with a vertex corresponding to each element of A. We associate a color with each element of $A - A$ and color an edge (a, a') with the color associated with $|a - a'|$ (recall that we define $A - A$ as containing only positive differences). Let C be the set of colors and note that $|C| = |A - A|$. By (2.5), to obtain a lower bound for $|C|$ it suffices to derive an upper bound for $\mathbb{E}_r(\chi)$. Let G^r be a pruned r-th energy graph of χ. By the discussion from the subsection on higher energy graphs, it suffices to derive an upper bound on $|E(G^r)|$.

Consider an edge $((v_1, \ldots, v_r), (v'_1, \ldots, v'_r)) \in E(G^r)$. Thinking of the vertices $v_j \in V$ as their corresponding elements in A, we have $|v_1 - v'_1| = |v_2 - v'_2| = \cdots = |v_r - v'_r|$. We associate this edge with a sequence of $r-1$ symbols from $\{+,-\}$, as follows (that is, we associate the edge with an element of $\{+,-\}^{r-1}$). For every $2 \leq j \leq r$, the $(j-1)$-th element of the sequence is '+' if $v_1 - v'_1 = v_j - v'_j$, and '-' if $v_1 - v'_1 = v'_j - v_j$. That is, the associated symbol encodes how to remove the absolute values from $|v_1 - v'_1| = |v_2 - v'_2| = \cdots = |v_r - v'_r|$.

We partition G^r into 2^{r-1} graphs, as follows. Each graph contains the same set of vertices $V(G^r)$, and each graph corresponds to one of the 2^{r-1} sequences of $\{+,-\}^{r-1}$. A graph that corresponds to a specific sequence $s \in \{+,-\}^{r-1}$ contains the edges of $E(G^r)$ that are associated with s. Note that every edge of G^r corresponds to exactly one of the 2^{r-1} graphs. Thus, to obtain an upper bound for $|E(G^r)|$ it suffices to bound the number of edges in each of these graphs.

Let $H = (V(G^r), E_H)$ be one of the 2^{r-1} graphs constructed in the preceding paragraph. Assume for contradiction that H contains a cycle γ of length $2k$, and denote the vertices of γ as

$$(v_{1,1}, \ldots, v_{1,r}), (v_{2,1}, \ldots, v_{2,r}), \ldots, (v_{2k,1}, \ldots, v_{2k,r}).$$

Recall that every edge of E_H corresponds to the same symbol $s \in \{+,-\}^{r-1}$. By the definition of an edge in an energy graph, for every $1 \leq j \leq 2k$ we have

$$v_{j,1} - v_{j+1,1} = s_2(v_{j,2} - v_{j+1,2}) = \cdots = s_r(v_{j,r} - v_{j+1,r}). \tag{2.6}$$

For any $1 \leq j_1 < j_2 \leq 2k$, summing (2.6) for $j_1 \leq j < j_2$ yields

$$v_{j_1,1} - v_{j_2,1} = s_2(v_{j_1,2} - v_{j_2,2}) = \cdots = s_r(v_{j_1,r} - v_{j_2,r}). \tag{2.7}$$

By the above, the vertices of the cycle γ form a clique K_{2k} in G^r. We next claim that the $2kr$ vertices $v_{j,\ell} \in V$ used to define the vertices of γ are all distinct. By the definition of $V_1, \ldots, V_r$, if $j \neq j'$ then $v_{j,\ell}$ and $v_{j',\ell'}$ must correspond to different elements of V. Assume for contradiction that $v_{j,\ell} = v_{j,\ell'}$ for some $\ell \neq \ell'$. By (2.7) with $j_1 = \ell$ and $j_2 = \ell'$, we obtain that $(v_{\ell,1}, \ldots, v_{\ell,r}) = (v_{\ell',1}, \ldots, v_{\ell',r})$. This contradicts γ being a simple cycle, which implies that the $2kr$ vertices $v_{j,\ell} \in V$ are indeed distinct.

Consider the set S consisting of the $2kr$ vertices $v_{j,\ell} \in V$ used to define the vertices of γ. By the preceding paragraph $|S| = 2kr$. By (2.6), for each of the $\binom{2k}{2}$ choices

for j_1 and j_2 we have $r-1$ distinct color repetitions. Consider one such repetition $v_{j_1,\ell} - v_{j_2,\ell} = v_{j_1,\ell'} - v_{j_2,\ell'}$ with $\ell \neq \ell'$ and note that it leads to a second repetition $v_{j_1,\ell} - v_{j_1,\ell'} = v_{j_2,\ell} - v_{j_2,\ell'}$ (if instead we start with $v_{j_1,\ell} - v_{j_2,\ell} = v_{j_2,\ell'} - v_{j_1,\ell'}$ then we have the second repetition $v_{j_1,\ell} - v_{j_2,\ell'} = v_{j_2,\ell} - v_{j_1,\ell'}$). In addition to the $r-1$ repetitions from the edge definition, we obtain $\binom{r}{2}$ repetitions this way. Thus, for each of the $\binom{2k}{2}$ choices for j_1 and j_2 we actually have $r - 1 + \binom{r}{2}$ distinct color repetitions. This contradicts the local property assumption, so H does not contain a cycle of length $2k$.

Since H does not contain a cycle of length $2k$, Theorem 2.2.2 implies

$$|E_H| \leq \mathrm{ex}(|V(G^r)|, C_{2k}) = O\left(|V(G^r)|^{1+1/k}\right) = O(n^{r+r/k}).$$

Recall that $E(G^r)$ is partitioned into 2^{r-1} subsets, each satisfying the above upper bound for $|E_H|$. We thus have

$$\mathbb{E}_r(\chi) = \Theta\left(|E(G^r)|\right) = O(n^{r+r/k}).$$

Combining this upper bound for $\mathbb{E}_r(\chi)$ with (2.5) gives

$$|A - A| = |C| = \Omega\left(n^{\frac{r}{r-1}\cdot\frac{k-1}{k}}\right).$$

This completes the proof of Theorem 1.1.5.

Chapter 3

EXPANDING POLYNOMIALS AND SETS WITH ADDITIONAL STRUCTURE

In this chapter, we discuss results about strong expansion under the presence of additional additive or multiplicative structure, as discussed in Section 1.2. This is a more technical chapter, so as a warm-up we begin with a simple inverse result, which also happens to be subject of the first paper I wrote as PhD student, [**Poh19**].

Theorem 3.0.1 *Suppose A is a finite set of real numbers and let $\Delta(A \times A)$ be the set of distances spanned by $A \times A$. Then,*

$$|A - A| \ll |\Delta(A \times A)|^{\frac{6}{7}} \log^{\frac{1}{7}} |A|,$$

or equivalently $|D| \ll |D^2 + D^2|^{\frac{6}{7}} \log^{\frac{1}{7}} |D|$, where D denotes the difference set $A - A$.

In other words, this states that cartesian products $A \times A \subset \mathbb{R}^2$ which determine few distinct distances must have additive structure. This improved on a recent theorem by Hanson [**Han16**], who showed that if $A \subset \mathbb{R}$ is so that $A \times A$ determines $O(|A|^2)$ distinct distances, then $|A - A| \ll |A|^{2-\frac{1}{8}}$. In the meantime this was also sharpened by Roche-Newton in [**RN**], who showed under the same hypothesis that $|A - A| \ll |A|^{2-\frac{2}{11}}$. Theorem 3.0.1 above implies the stronger result that $|A - A| \ll |A|^{2-\frac{2}{7}} \log^{\frac{1}{7}} |A|$. The proof relies on the following beautiful sum-product estimate of Solymosi from [**Sol09**], which is arguably the most important result for the sum-product problem over the reals.

Theorem 3.0.2 *Let $S \subset \mathbb{R}$ be a set. Then,*

$$|S + S|^2 |SS| \geq \frac{|S|^4}{4\lceil \log |S| \rceil}.$$

We will also need the classical Plűnnecke-Ruzsa inequality, for which a simple proof can be found in [**Pet12**]

Lemma 3.0.3 *Let $A \subset \mathbb{R}$ be a finite set. Then*

$$|kA - \ell A| \leq \frac{|A + A|^{k+\ell}}{|A|^{k+\ell-1}}.$$

We are now ready for the proof of Theorem 3.0.1.

Proof of Theorem 3.0.1. If we let $D := A - A$, then the number of distinct distances determined by $A \times A$ as a point set in $\mathbb{R}^2$ is given by $|\Delta(A \times A)| = |D^2 + D^2|$, where $D^2 = \left\{(x-y)^2 : x, y \in A\right\}$. We claim that

$$|D| \ll |D^2 + D^2|^{\frac{6}{7}} \log^{\frac{1}{7}} |D|.$$

We apply Theorem 3.0.2 for the set $S := D^2$. Using the observation that $|D^2D^2|$ is equal to $|DD|$ (up to a small constant), this yields

$$|D^2 + D^2|^2|DD| \geq |D^2 + D^2|^2|D^2D^2| \geq \frac{|D^2|^4}{4\lceil \log |D^2| \rceil} \gg \frac{|D|^4}{\log |D|}.$$

On the other hand, note that for every four real numbers a_1, a_2, b_1, b_2, the following identity holds

$$(b_1 - a_1)^2 + (b_2 - a_2)^2 - (b_1 - a_2)^2 - (b_2 - a_1)^2 = 2(a_2 - a_1)(b_1 - b_2).$$

This yields the inclusion

$$2 \cdot DD \subset 2D^2 - 2D^2.$$

We emphasize here that for $X \subset \mathbb{R}$ and $c \in \mathbb{Z}_{>0}$, the set $c \cdot X$ denotes the set of scalar multiples $\{cx : x \in X\}$, whereas cX denotes the sumset $\sum_{i=1}^{c} X$. The inclusion together with Lemma 3.0.3 then yields

$$\begin{aligned} |D^2 + D^2|^2|DD| &= |D^2 + D^2|^2|2 \cdot DD| \\ &\leq |D^2 + D^2|^2|2D^2 - 2D^2| \\ &\ll |D^2 + D^2|^2 \left(\frac{|D^2 + D^2|^4}{|D|^3}\right). \end{aligned}$$

Putting the two bounds together, we conclude that

$$\frac{|D^2 + D^2|^6}{|D|^3} \gg \frac{|D|^4}{\log |D|},$$

which yields

$$|D| \ll |D^2 + D^2|^{\frac{6}{7}} \log^{\frac{1}{7}} |D|.$$

This completes the proof of Theorem 3.0.1.

There seems to be a mysterious connection between the sum-product problem and the problem of establishing structure in planar points sets which determine few

distinct distances (i.e. sets of points which achieve equality up to constants in the Guth-Katz bound [**GK15**]). In addition to Theorem 3.0.1, there's also an interesting result of Sheffer, Zahl, and de Zeeuw [**SZZ16**], which shows that in such extremal configurations P, one cannot have lines (or circles) containing more than $|P|^{7/8}$ (or $|P|^{4/5}$) points from P. Their method makes use of incidence results between some suitably defined points and curves, where the curves do not have large multiplicity thanks to a sum-product result. Understanding the intriguing connection between the two worlds beyond these two rather isolated data points would be very interesting.

We now move on to the expansion phenomenon discussed in Section 1.2, which lies in some sense at the opposite end of the spectrum. Starting with a finite set of reals A with strong additive or multiplicative structure, we will now see that $f(A, A)$ expands rather dramatically (i.e. beyond linearly) for all polynomials $f \in \mathbb{R}[x, y]$ which do not "encourage" this additive or multiplicative structure of the set A, respectively. The first such result we proved in [**PS17**] was Theorem 3.0.4, which we restate below.

Theorem 3.0.4 *Let $A \subset \mathbb{R}$ be a finite set with $|A + A| = K|A|$ and let $f \in \mathbb{R}[x, y]$ be a polynomial of degree at most d that is not of form $g(L(x, y))$ for some linear polynomial $L \in \mathbb{R}[x, y]$ and some univariate polynomial g. Then, for any $\epsilon > 0$, we have*

$$|f(A, A)| = \Omega_{d,K}\left(|A|^{14/9-\epsilon}\right).$$

Recall that this improved on Theorem of Shen [**She12**] which said that if $f \in \mathbb{R}[x, y]$ is not of the form $g(L(x, y))$ for some linear polynomial L and some univariate polynomial g, then $|A + A| = K|A|$ implies that $|f(A, A)| = \Omega_{d,K}(|A|^{3/2})$ for every polynomial that is not of the form $g(L(x, y))$ for some linear polynomial L and some univariate polynomial g. The hypothesis on f is necessary: if $A = \{1, \ldots, N\}$, then it is not too difficult to see that a polynomial of the form $f(x, y) = g(L(x, y))$ satisfies $|f(A, A)| = \Theta_{d,K}(|A|)$. Take for instance $f(x, y) = x + y$, for which $|f(A, A)| = |A + A| = K|A|$.

The proof of Theorem 3.0.4 builds upon the proof of Shen's theorem from [**She12**], which is an elegant combination of Vu's argument from [**Vu08**] (for the corresponding expansion result over finite fields) and Elekes' argument from [**Ele97**] (for the sum-product problem over the reals). For the record, it is perhaps also important to emphasize that Elekes' proof relied on a surprisingly simple application of the

Szemerédi-Trotter theorem about incidences between points and lines in the plane, whose success lead to an entire revolution in the field and inspired multiple impressive results over the reals (such as Solymosi's Theorem 3.0.2 above). In order to prove Theorem 3, Shen was able to replace the spectral graph theory from the finite field case with more general tools from incidence geometry over the reals. In joint work with Adam Sheffer [**PS17**], we proved Theorem 3.0.4 by using more advanced incidence theorems and a more careful analysis involving some of the high energy ideas from Chapter 2. See [**PS17**] for more details.

For the reader's convenience, we will now take a moment to recall the main result from [**Poh20**] whose proof we will discuss in detail in this chapter. In [**Poh20**], I addressed the dual problem of classying the two variable polynomials $f(x, y) \in \mathbb{R}[x, y]$ such that $|f(A, A)|$ is large whenever $|AA|$ is small. As in the additive case, it is easy to check that there are some polynomials for which this implication does not hold. For instance, consider $A = \left\{2, 2^2, \ldots, 2^N\right\}$ for which $|AA| = 2N - 1$. If we let f be a single monomial such as $f(x, y) = x^2y^3$, then it is easy to check that $f(A, A) = 5N - 4$. More generally, if we choose g to be a polynomial in one variable and $M(x, y)$ to be a single monomial, the $f(A, A)$ will also be small for $f(x, y) = g(M(x, y))$ in general. Indeed, consider say $f(x, y) = xy + x^2y^2$; then we also have that $f(A, A) = 2N - 1$. Our main result shows that $g(M(x, y))$ is the *only* real enemy, in the following strong sense.

Theorem 3.0.5 *Let d be a positive integer and let $f \in \mathbb{R}[x, y]$ be a polynomial of degree d that is not of the form $g(M(x, y))$ for some single monomial $M(x, y)$ and some univariate polynomial g. If A is a finite set of real numbers such that $|AA| = K|A|$, then*

$$|f(A, A)| = \Omega_{d,K}\left(|A|^2\right).$$

It is a bit surprising that one can prove a much stronger expansion theorem in the multiplicative case than in the additive case. Theorem 3.0.5 is optimal up to the dependence on d and K. When $f(x, y) = x + y$ it recovers a result of Chang [**Cha06**]. We prove Theorem 3.0.5 in Section 3.3, after introducing the required ingredients in the upcoming Section 3.2.

3.1 Algebraic and analytic preliminaries

The proof of Theorem 3.0.5 is in some sense in the spirit of the proofs of Theorem 1.2.2, Theorem 3, and Theorem 3.0.4, but the reason it does so well quantitatively is the fact that it does not rely on any incidence geometry. The main input is instead a (quantitative) version of the celebrated Schmidt subspace theorem [**Sch72**] due to Amoroso and Viada [**AV09**], which we use to control the number of "nondegenerate" solutions to certain polynomial equations.

Theorem 3.1.1 *Let $a_1, \ldots, a_n \in K$ be nonzero elements of an algebraically closed field K, and let Γ be a subgroup of K of finite rank r. Then, the number of solutions of the equation*

$$a_1 z_1 + \ldots + a_n z_n = 1$$

with $z_i \in \Gamma$ and no subsum on the left hand side vanishing is at most

$$C(n, r) := (8n)^{4n^4(n+nr+1)}.$$

Schmidt's subspace theorem (together with its different variants) represents a powerful result in number theory, particularly famous for its many applications in diophantine approximation and complexity of algebraic numbers. Many excellent surveys have been written about it, so for instance [**Bil07**] and [**SS14**]. In fact, Theorem 3.1.1 has already manifested itself in additive combinatorics as well in [**Cha06**], where Chang noticed that one can use it to prove that $|AA| = O(|A|)$ implies $|A + A| = \Omega(|A|^2)$. Theorem 3.0.5 can therefore also be seen as a generalization of this phenomenon.

The next ingredient is a multiplicative version of a somewhat more unusual version of Freiman's theorem from additive combinatorics, which is a combination of [**Cha94**] and Freiman's Lemma [**Fre73**]. See [**GT06**] for more details.

Theorem 3.1.2 *Let $t \geq 1$ be an integer, let $\epsilon > 0$, and let A be a finite set of real numbers with $|AA| = K|A|$ and $|A| \geq CK^2/\epsilon$ for some absolute constant $C > 0$. Then, A is a subset of a set G, which is of the form*

$$G := g_1^{[H_1]} \cdot \ldots \cdot g_r^{[H_r]} = \left\{ \prod_{i=1}^{r} g_i^{\mu_i} : \ \mu_i \in \mathbb{Z}, \mu_i \in [H_i] \right\}^{1},$$

[1] If n is a positive integer, $[n]$ denotes the set $\{0, 1, \ldots, n-1\}$.

where $r \leq \lfloor K - 1 + \epsilon \rfloor$, all the products in

$$G^{(t)} := \left\{ \prod_{i=1}^{r} g_i^{\mu_i} : \ \mu_i \in \mathbb{Z}, \mu_i \in [tH_i] \right\}$$

are pairwise distinct, and

$$|G| = H_1 \cdot \ldots \cdot H_r \leq t^K \exp(CK^2 \log^3 K)|B|.$$

The last two ingredients are more algebraic in nature. First, recall that a polynomial $f \in \mathbb{R}[x, y]$ is said to be *reducible* if there exist polynomials $f_1, f_2 \in \mathbb{R}[x, y]$ of positive degrees such that $f(x, y) = f_1(x, y) \cdot f_2(x, y)$. A polynomial that is not reducible is said to be *irreducible*. Furthermore, we say that a polynomial $p \in \mathbb{R}[x, y]$ is *decomposable* if there exists a univariate polynomial p_1 of degree at least two and $p_2 \in \mathbb{R}[x, y]$ such that $p(x, y) = p_1(p_2(x, y))$. Similarly, a polynomial that is not decomposable is said to be *indecomposable*.

We will need a consequence of a theorem of Stein [**Ste89**], which follows from the main result of [**Aya02**]. See [**RSS16**] for more details.

Theorem 3.1.3 *If $f \in \mathbb{R}[x, y]$ is indecomposable, then the polynomial $f(x, y) - \lambda$ is reducible for at most* $\deg f$ *values of $\lambda \in \mathbb{R}$.*

Last but not least, we will also need the classical Bézout theorem [**Cox05**], which again we only state for real polynomials.

Theorem 3.1.4 *Let f and g be two polynomials in $\mathbb{R}[x, y]$. If f and g vanish simultaneously on more than* $(\deg f)(\deg g)$ *points of $\mathbb{R}^2$, then f and g have a common non-trivial factor.*

3.2 Expansion when AA is small

We prove Theorem 3.0.5 by following the presentation from [**Poh20**]. Let $f \in \mathbb{R}[x, y]$ be a polynomial that is not of the form $g(M(x, y))$ for some single monomial M and some univariate polynomial g, and let d be the degree of f. We will prove that

$$|f(A, A)| = \Omega_{d,K}(|A|^2)$$

whenever $A \subset \mathbb{R}$ satisfies $|AA| = K|A|$. The dependence on d and K is going to be explicit, but since it is not a priority from time to time we will reserve the right to hide certain expressions under the asymptotic notation whenever it is more convenient.

First, recall that if f is decomposable, then there exist a univariate f_1 of degree at least two and $f_2 \in \mathbb{R}[x, y]$ such that $f(x, y) = f_1(f_2(x, y))$. Let (f_1, f_2) be a pair of such polynomials that minimizes the degree of f_2. In particular, this implies that f_2 is indecomposable. Since f is of degree at most d, so are f_1 and f_2. Since f_1 is univariate, for every $a \in \mathbb{R}$ there exist at most d numbers $b \in \mathbb{R}$ such that $f_1(b) = a$. Thus, if $|f_2(A, A)| \geq T$ holds for some positive quantity T, then $|f(A, A)| \geq T/d$. It then remains to derive the lower bound for the indecomposable f_2, which we also know it is not a single monomial $M(x, y)$ from the hypothesis. Abusing of notation, we will refer to f_2 as f from now on, and therefore assume without loss of generality that f is indecomposable and not a single monomial as well.

Next, we naturally define the following polynomial energy of A by

$$E_f(A) := \left|\left\{(x, y, x', y') \in A^4 : \ f(x, y) = f(x', y')\right\}\right|.$$

For each $\alpha \in f(A, A)$, we also let $m_A(\alpha)$ denote the number of pairs $(x, y) \in A \times A$ such that $f(x, y) = \alpha$. In particular,

$$m_A(\alpha)^2 = \left|\left\{(x, y, x', y') \in A^4 : \ f(x, y) = f(x', y') = \alpha\right\}\right|,$$

so by Cauchy-Schwarz,

$$E_f(A) = \sum_{\alpha \in f(A,A)} m_A(\alpha)^2 \geq \frac{1}{|f(A, A)|} \cdot \left(\sum_{\alpha \in f(A,A)} m_A(\alpha)\right)^2 = \frac{|A|^4}{|f(A, A)|}.$$

In order to prove that $|f(A, A)| = \Omega_{d,K}(|A|^2)$, it thus suffices to show that $E_f(A) = O_{d,K}(|A|^2)$ instead. To achieve this, we will show that for most values of $\alpha \in f(A, A)$, the number of solutions in $A \times A$ to the equation $f(x, y) = \alpha$ is at most a constant which depends solely on d and K. More precisely, we claim that

$$\left|\left\{\alpha \in f(A, A) : \ m_A(\alpha) > C\left(\binom{d+2}{2}, K\right) + d^2 2^{\binom{d+2}{2}}\right\}\right| = O_d(1), \tag{3.1}$$

where $C\left(\binom{d+2}{2}, K\right)$ is the explicit constant from Theorem 3.1.1.

Let us first check that this claim implies that $E_f(A) = O_{d,K}(|A|^2)$. For convenience, let

$$\Upsilon(A) := \left\{\alpha \in f(A, A) : \ m_A(\alpha) > C\left(\binom{d+2}{2}, K\right) + d^2 2^{\binom{d+2}{2}}\right\},$$

and write

$$E_f(A) = \sum_{\alpha \in \Upsilon(A)} m_A(\alpha)^2 + \sum_{\alpha \in f(A,A)\setminus\Upsilon(A)} m_A(\alpha)^2. \tag{3.2}$$

For every $\alpha \in f(A, A)$, it is easy to see that $m_A(\alpha)^2 = O_d(|A|^2)$. Indeed, recall that this quantity is the number of solutions in A^4 to $f(x, y) = f(x', y') = \alpha$, so once x and x' are chosen in A, there are at most d value for each y and y' that can satisfy the equality. In particular, if $|\Upsilon(A)| = O_d(1)$, this implies that the first term in (3.2) satisfies

$$\sum_{\alpha \in \Upsilon(A)} m_A(\alpha)^2 = O_d(|A|^2).$$

For the second term, note that if $\alpha \notin \Upsilon(A)$ then

$$M := \max_{\alpha \in f(A,A)\setminus\Upsilon(A)} m_A(\alpha) \leq C\left(\binom{d+2}{2}, K\right) + d^2 2^{\binom{d+2}{2}} = O_{d,K}(1),$$

therefore

$$\sum_{\alpha \in f(A,A)\setminus\Upsilon(A)} m_A(\alpha)^2 \leq M \sum_{\alpha \in f(A,A)\setminus\Upsilon(A)} m_A(\alpha) \leq M|A|^2 = O_{d,K}(|A|^2).$$

Putting these two estimates together, we indeed get that $E_f(A) = O_{d,K}(|A|^2)$. We are now left to prove (3.1), which will require the tools from Section 2.

Recall that A satisfies $|AA| = K|A|$. If the size of A is upper bounded by a constant in terms of K, then there is nothing to prove since $|f(A, A)| = \Omega_{d,K}(|A|^2)$ is trivially true, so we can safely apply Theorem 3.1.2 with $\epsilon = 1$ and $t = d$. This implies that A is a subset of a set G, which is of the form

$$G := g_1^{[H_1]} \cdot \ldots \cdot g_r^{[H_r]} = \left\{ \prod_{i=1}^{r} g_i^{\mu_i} : \ \mu_i \in \mathbb{Z}, \mu_i \in [H_i] \right\},$$

where $r \leq \lfloor K \rfloor \leq K$ and all the products all the products in

$$G^{(d)} := \left\{ \prod_{i=1}^{r} g_i^{\mu_i} : \ \mu_i \in \mathbb{Z}, \mu_i \in [dH_i] \right\}$$

are pairwise distinct. We also have a quantitative estimate for $|G|$, but it is not required.

For each $\alpha \in f(A, A)$, we now analyze the number of solutions in $A \times A$ to $f(x, y) = \alpha$. Write f explicitly as

$$f(x, y) := \sum_{(i,j)\in S} a_{i,j} x^i y^j,$$

where S is some subset of the set of pairs $\{(i,j) : i,j \geq 0, i+j \leq d\}$ and $a_{i,j}$ is a real coefficient for each $(i,j) \in S$.

We begin with a first key lemma.

Lemma 3.2.1 *For every $\alpha \in f(A,A)$, the number of solutions in $A \times A$ to*

$$\sum_{(i,j)\in S} a_{i,j} x^i y^j = \alpha$$

with no subsum on the left hand side vanishing is at most $C\left(\binom{d+2}{2}, K\right)$.

Proof of Lemma 3.2.1. Let Γ be multiplicative subgroup of $\mathbb{C}^*$ generated by $g_1, \ldots, g_r$, which has rank $r \leq K$ and contains G (and thus also A). The number of solutions to

$$\sum_{(i,j)\in S} a_{i,j} z_{i,j} = \alpha \tag{3.3}$$

with $z_{i,j} \in G$ for each $(i,j) \in S$ and no subsum on the left hand side vanishing is at most the number of solutions to (3.3) with the $z_{i,j}$ in Γ and no subsum on the left hand side vanishing, so by Theorem 3.1.1 it is at most $C\left(\binom{d+2}{2}, K\right)$. If we also can argue that for each such solution $(z_{i,j})_{(i,j)\in S}$ to (3.3), there is at most one solution in $\Gamma \times \Gamma$ (and thus in $A \times A$) to the system of equations

$$x^i y^j = z_{i,j} \quad \text{for each } (i,j) \in S, \tag{3.4}$$

then the claim follows.

For each $z \in G$ write $z = g_1^{z_1} \cdot \ldots \cdot g_r^{z_r}$ for some $z_i \in [H_i]$. Plugging this into (3.4), we get

$$g_1^{ix_1+jy_1} \cdot \ldots \cdot g_r^{ix_r+jy_r} = g_1^{z_{i,j,1}} \cdot \ldots \cdot g_r^{z_{i,j,r}} \quad \text{for each } (i,j) \in S.$$

Here all $x_k, y_k, z_{i,j,k} \in [H_k]$ for all $k \in \{1, \ldots, r\}$. Furthermore, since $i+j \leq d$, we also have that $ix_k + jy_k \in [dH_k]$ for each k, so by the fact that $G^{(d)}$ has all its products pairwise distinct, it follows that (3.4) translates into the following system of equalities, call it $\mathcal{S}_{i,j}$, satisfied by the exponents above for each $(i,j) \in S$:

$$ix_k + jy_k = z_{i,j,k} \quad \text{for all } k \in \{1, \ldots, r\}.$$

At this point, recall that f is indecomposable by our assumption and is also not a single monomial, so it must contain at least two monomials, say $x^i y^j$ and $x^{i'} y^{j'}$, for

which the two-dimensional vectors (i, j) and (i', j') are not a scalar multiple of each other. In particular, if a pair $(x, y) \in A \times A \subset G \times G$ satisfies both $\mathcal{S}_{i,j}$ and $\mathcal{S}_{i',j'}$, then each pair (x_k, y_k) is uniquely determined in terms of i, j, i', j' and $z_{i,j,k}$, $z_{i',j',k}$ for each $k \in \{1, \ldots, r\}$, which implies that (x, y) is then uniquely determined. This proves the claim.

We now analyze what happens if there are vanishing subsums on the left hand side of $f(x, y) = \alpha$. In this sense, we prove the following second key lemma.

Lemma 3.2.2 *For all but possibly at most $d + 1$ values of $\alpha \in f(A, A)$, the number of pairs $(x, y) \in A \times A$ satisfying $f(x, y) = \alpha$ with some vanishing subsum on the left hand side is at most $d^2 2^{\binom{d+2}{2}}$.*

Proof of Lemma 3.2.2. Recall $f(x, y) := \sum_{(i,j)\in S} a_{i,j} x^i y^j$ with $|S| \geq 2$, and now suppose that

$$\sum_{(i',j')\in S'} a_{i',j'} x^{i'} y^{j'} = 0$$

for some nontrivial subset $S' \subset S$. Let $N_{S'}(\alpha)$ be number of common solutions in $A \times A$ to

$$f(x, y) - \alpha = 0 \text{ and } g_{S'}(x, y) = 0, \tag{3.5}$$

where $g_{S'} \in \mathbb{R}[x, y]$ is the polynomial defined by

$$g_{S'}(x, y) := \sum_{(i',j')\in S'} a_{i',j'} x^{i'} y^{j'}$$

for a nontrivial subset S' of S. It suffices to prove that

$$\sum_{S'\subset S} N_{S'}(\alpha) \leq d^2 2^{\binom{d+2}{2}}$$

holds for all but possibly at most $d + 1$ values of $\alpha \in f(A, A)$.

For each $S' \subset S$, note that $g_{S'}$ has degree at most d, since $\deg f = d$. By Theorem 3.1.3 there are at most d values of α for which $f(x, y) - \alpha$ is reducible, and at most one value for which $f(x, y) - \alpha$ may be identical to $g_{S'}(x, y)$, for some $S' \subset S$ (α may be equal to the free term in f). For each of the other $\alpha \in f(A, A)$, we have that $N_{S'}(\alpha) \leq d^2$ for every proper $S' \subset S$. Indeed, if α is such that the polynomial $f(x, y) - \alpha$ is irreducible in $\mathbb{R}[x, y]$ and does not coincide with $g_{S'}(x, y)$, then this

simply follows from Theorem 3.1.4, since (3.5) must have at most d^2 solutions if there is no common factor. Therefore,

$$\sum_{S' \subset S} N_{S'}(\alpha) \leq d^2 2^{|S|} \leq d^2 2^{\binom{d+2}{2}}.$$

This completes the proof of Lemma 3.2.2.

Claim (3.1) now follows by combining Lemma 3.2.1 and Lemma 3.2.2. Indeed, together these two imply that for all but possibly at most $d+1$ values of $\alpha \in f(A, A)$, the number of pairs $(x, y) \in A \times A$ with $f(x, y) = \alpha$ is at most $C\left(\binom{d+2}{2}, K\right) + d^2 2^{\binom{d+2}{2}}$. In other words, $|\Upsilon(A)| \leq d + 1$, where

$$\Upsilon(A) := \left\{ \alpha \in f(A, A) : \ m_A(\alpha) > C\left(\binom{d+2}{2}, K\right) + d^2 2^{\binom{d+2}{2}} \right\}.$$

This completes the proof of Theorem 3.0.5.

Chapter 4

IMPROVED BOUNDS FOR PROGRESSION-FREE SETS IN $\mathbb{Z}_8^N$

In this chapter, we will discuss the results from my joint work with Fedor Petrov [**PP20**]. Let G be a finite group. A non-trivial three-term progression in G is an ordered triple $(a, b, c) \in G^3$ of *mutually distinct* elements such that $ac = b^2$. Let $f_3(G)$ be the size of the largest $A \subset G$ without non-trivial three-term progressions.

The problem of upper bounding $r_3(C_n)$ has a long history, which we briefly have already reviewed in Section 1.3. Nonetheless, let us recall some things in order to also get used to the multiplicative notation. The first important estimate is the one of Roth from [**Rot53**]. Currently, the best known upper bound is due to Bloom [**Blo16**], who proved that

$$f_3(C_n) \ll \frac{(\log\log n)^4}{\log n} n.$$

The best known lower bound is of the form

$$f_3(C_n) \gg n \exp(-c\sqrt{\log n})$$

for some absolute constant $c > 0$ and is due to Behrend [**Beh46**]. In particular, $f_3(C_n)$ grows faster than $n^{1-\epsilon}$ for any fixed $\epsilon > 0$.

For other groups G, $f_3(G)$ turns out to be much smaller than $|G|$. The first result of this kind was obtained by Croot, Lev and Pach in their recent breakthrough paper [**CLP17**], where they showed that

$$f_3(C_4^n) \leqslant 4^{\gamma n} \approx (3.611)^n.$$

The constant γ in their paper is given by

$$\gamma := \max\left\{\frac{1}{2}\left(\mathcal{H}_2(0.5-\epsilon) + \mathcal{H}_2(2\epsilon)\right) : \ 0 < \epsilon < 0.25\right\} \approx 0.926, \tag{4.1}$$

where $\mathcal{H}_2(\theta)$ denotes the binary entropy function

$$\mathcal{H}_2(\theta) = -\theta\log_2\theta - (1-\theta)\log_2(1-\theta), \ \ \theta \in (0,1).$$

This constant arises naturally in their polynomial method proof, which makes clever use of the group structure of C_4^n. This was a remarkable improvement on the previous

known bounds for $G = C_4^n$, the prior record due to Sanders [**San09**] being of the form

$$f_3(C_4^n) \ll \frac{4^n}{n(\log n)^\epsilon}$$

with an absolute constant $\epsilon > 0$. Soon after, their method was adapted and simplified in setups with more pleasant group structure. First, Ellenberg and Gijswijt in [**EG17**] proved that $f_3(C_p^n) \leqslant \kappa_p^n$ for all odd primes p, where κ_n generally stands for

$$\kappa_n := \min\left\{x^{(1-n)/3}(1 + x + \cdots + x^{n-1}) : x > 0)\right\}. \tag{4.2}$$

This was another major result, as it improved dramatically the celebrated estimate

$$f_3(C_3^n) \ll \frac{3^n}{n^{1+\epsilon}}$$

of Bateman and Katz [**BK12**]. This was further adapted by three different teams to prove that for all odd prime powers q, $f_3(C_q^n) \leqslant \kappa_q^n$ ([**BCCGNSU17**], [**Spe16**], and [**Pet16**]), and also later on by various other authors to prove several other different results in extremal combinatorics.

The group algebra approach introduced in [**Pet16**] allows one to estimate $f_3(G)$ for groups which are not necessarily abelian. Nonetheless, all such extensions have been so far about groups of odd order. One of the difficulties about groups of even order consists of the fact that they may contain "semi-trivial" progressions (a, b, a) with $a^2 = b^2$ and $a \neq b$. In particular, an estimate for the number of so called *multiplicative matchings*[1] is no longer an estimate for $f_3(G)$. In the upcoming sections, the first aim is to give a group algebra proof of the fact that $f_3(C_4^n) \leqslant (3.611)^n$, with a more motivated account for the constant $\kappa_4 \approx 3.61$. The purpose of this is two-fold. First, it will reconcile the expression from (4.1) with the one from (4.2), thus showing a clear analogy between C_4^n and the odd prime power regime. Second, it will provide a framework that will allow us to give improved bounds for progression-free sets in other (abelian) 2-groups, which will constitute the main result in this chapter.

For a finite abelian group $G \cong \prod_{i=1}^n C_{m_i}$ with positive integers $m_1 | \ldots | m_n$, denote by $\mathrm{rk}_4(G)$ the number of indices $i \in \{1, \ldots, n\}$ with $4|m_i$. Since G is a union of $4^{-\mathrm{rk}_4(G)}|G|$ cosets of a subgroup isomorphic to $C_4^{\mathrm{rk}_4(G)}$, this yields a bound of the form

$$f_3(G) \leqslant 4^{-(1-\gamma)\,\mathrm{rk}_4(G)}|G| \approx (0.903)^{\mathrm{rk}_4(G)}|G|. \tag{4.3}$$

[1] Multiplicative matchings coincide with what initially were called *tricolored sum-free sets* in [**KSS18**]; the updated term is adopted from Sawin [**Saw18**].

This is the content of Corollary 1 in [**CLP17**]. For instance, if $G = C_8^n$, the above gives

$$f_3(C_8^n) \leqslant 2^n \cdot r_3(C_4^n) \leq (7.222)^n.$$

In Section 4.4, we improve on this estimate and show the following:

Theorem 4.0.1 *If $A \subset C_8^n$ is a set without non-trivial three-term progressions, then*

$$|A| \leqslant \left(2 \cdot 2^{\mathcal{H}_4(\rho_0)}\right)^n \approx (7.0899)^n,$$

where $2^{\mathcal{H}_4(\rho)}$ represents a weighted version of κ_4 given by

$$2^{\mathcal{H}_4(\rho)} = \min_{x>0} \left\{x^{-3\rho}(1 + x + x^2 + x^3)\right\},$$

and $\rho_0 \approx 0.32$ solves the system

$$\mathcal{H}_4(\rho) = \mathcal{H}_2(\theta_1) + \mathcal{H}_2(1 - 2\theta_1),\ \mathcal{H}_4(1 - 2\rho) = 1 + \mathcal{H}_2(1 - 2\theta_1)$$

for $\theta_1 \in [x_0, 1]$ and $\rho \in [1/4, 1/2]$. Here, the constant x_0 stands for the unique maximum point of the function $\mathcal{H}_2(1 - 2x) + \mathcal{H}_2(x)$ in $[1/4, 1/2]$. In particular,

$$f_3(C_8^n) \leqslant (7.0899)^n.$$

For finite abelian groups, it is also worth mentioning the following consequence.

Corollary 4.0.2 *If a finite abelian group G is written as*

$$G \cong \prod_{i=1}^{n} C_{m_i},$$

where $m_1 | \ldots | m_n$, then

$$f_3(G) \leqslant (0.8863)^{\mathrm{rk}_8(G)}|G|,$$

where $\mathrm{rk}_8(G)$ denotes the number of indices $i \in \{1, \ldots, n\}$ with $8|m_i$.

This is similar in spirit with Corollary 1 from [**CLP17**] and constitutes an improvement in many other cases beyond Theorem 4.0.1. Like before, it follows immediately from Theorem 4.0.1 due to the simple fact that if $n := \mathrm{rk}_8(G)$, then the group G is a union of $8^{-n}|G|$ cosets of a subgroup isomorphic to C_8^n.

4.1 Regularization and the tensor power trick

Before we begin, we will first prove a couple of lemmas which will allow us to reduce the problem of upper bounding the size of the largest subset of C_4^n without non-trivial three-term progressions to upper bounding the size of the largest three-term progression-free subset of C_4^n which has the further property that it roughly intersects each of the 2^n cosets of C_2^n in the same number of elements.

Let Ω be a finite set which is partitioned into classes of size at most m. A subset $A \subset \Omega$ is called *regular* if there exists an integer k such that $|A \cap C| \in \{0, k\}$ for every class C. Suppose further that each element $x \in A$ has a non-negative *weight* $w(x)$, and define the the weight of a subset $B \subset A$ by

$$w(B) := \sum_{x \in B} w(x).$$

Lemma 4.1.1 *Every set $A \subset \Omega$ contains a regular subset of weight at least $w(A)/H_m$, where $H_m = 1 + 1/2 + \cdots + 1/m$.*

Proof. Assume the contrary: A does not contain such a regular subset. For each $i \in \{1, 2, \ldots, m\}$, each class C with at least i elements of A contains a subset of A of size i and weight at least $\frac{i}{|C \cap A|} w(C \cap A)$. Thus by our assumption

$$i \cdot \sum_{C : |C \cap A| \geqslant i} \frac{w(C \cap A)}{|C \cap A|} < \frac{w(A)}{H_m}.$$

Divide this inequality by i and sum up over all $i = 1, 2, \ldots, m$. We get $w(A) < w(A)$, a contradiction. This proves Lemma 4.1.1.

Now we assume that the universe Ω consists of animals in a certain forest. They may be of different species, and there are at most m animals of each species. Next, the species belong to different genera, each genus consists of at most m' different species. A subset $A \subset \Omega$ is called *super-regular*, if there exist integers k, k' such that for every species C, the set A contains 0 or k animals of species C (in the latter case we say that the species C is presented in A), and each genus $\mathcal{G}$ contains either 0 or k' species presented in A.

Also, each animal $a \in \Omega$ has its weight $w(a) \geqslant 0$.

Lemma 4.1.2 *In the above setting, any set $A \subset \Omega$ contains a super-regular subset of weight at least $w(A)/(H_m H_{m'})$.*

Indeed, by Lemma 4.1.1, we find a regular subset $B \subset A$ (with respect to the partition into species) of weight at least $w(A)/H_m$. Consider the species presented in B; their weights are well-defined, so the conclusion follows by applying again Lemma 4.1.1 to the partition of these species into genera.

A similar statement holds for the higher taxonomic hierarchy and is proved in the same way.

In what follows, we will apply both Lemmas 4.1.1, 4.1.2 for the weight function equal to 1 everywhere. For the group $G = C_4^n$, we consider the subgroup generated by its involutions, i.e. the image and the kernel of the endomorphism of C_4^n defined by $g \mapsto g^2$; this is a copy of C_2^n, so we can partition C_4^n into 2^n cosets modulo the subgroup $C_2^n = \{g^2 : g \in G\}$. Thus by Lemma 4.1.1 every subset $A \subset C_4^n$ contains a regular subset B of size at least $|A|/H_{2^n}$. For the group C_8^n, define the species and genera as equivalence classes of the relations

$$g \sim h \text{ if } g^2 = h^2 \text{ and } g \sim h \text{ if } g^4 = h^4,$$

respectively. Then by Lemma 4.1.2 every subset $A \subset C_8^n$ contains a super-regular subset B of size at least $|A| \cdot (H_{2^n})^{-2}$.

Returning to sets without three-term progressions, note that for arbitrary groups G_1, G_2, the product of two such sets $A_1 \subset G_1$, $A_2 \subset G_2$ is itself a subset in $G_1 \times G_2$ without three-term progressions. Hence

$$f_3(G_1 \times G_2) \geqslant f_3(G_1) f_3(G_2).$$

In particular, by Fekete's Lemma [**MF23**],

$$\lim_{n\to\infty} (f_3(G^n))^{1/n} = \sup_{n>0} (f_3(G^n))^{1/n}. \tag{4.4}$$

This implies that any estimate of the form $f_3(G^n) \leqslant c^{n+o(n)}$ automatically yields $f_3(G^n) \leqslant c^n$. In particular, Lemma 4.1.1 and Lemma 4.1.2 above reduce the problem of proving subexponential upper bounds c^n for the size of the largest subset of C_4^n or C_8^n without three-term progressions to proving that regular (respectively, super-regular) three-term progression-free subsets of the group C_4^n (respectively, C_8^n) have size $\leqslant c^{n+o(n)}$.

4.2 Subspaces with zero product in abelian 2-groups

In this section, we build the general framework that we will use for the proof of Theorem 4.0.1. Along the way, we explain the natural relationship between

$$\frac{1}{2} \cdot \max_{0<\epsilon<0.25} \{\mathcal{H}(0.5-\epsilon) + \mathcal{H}(2\epsilon)\} \approx 0.926$$

and

$$\kappa_4 := \min_{x>0} x^{-1}(1 + x + x^2 + x^3) \approx 3.611.$$

Let $G = \prod_{i=1}^{n} C_{2^{m_i}}$ be an abelian 2-group. If $X_1, \ldots, X_k$ are subspaces of $\mathbb{F}_2[G]$, we will denote by $X_1 \cdot \ldots \cdot X_k$ the product set $\{x_1 \cdot \ldots \cdot x_k : x_i \in X_i \text{ for every } i \in \{1, \ldots, k\}\}$. In this Section, we will be interested in subspaces whose product set equals zero. Here $\mathbb{F}_2[G]$ represents the group ring of G over $\mathbb{F}_2$, namely

$$\mathbb{F}_2[G] := \mathbb{F}_2[\tau_1, \ldots, \tau_n]/\langle \tau_i^{2^{m_i}} = 0, i = 1, \ldots, n\rangle.$$

The nilpotent elements τ_i have form $1 + g_i$, where g_i are generators of the cyclic groups $C_{2^{m_i}}$. Therefore $\mathbb{F}_2[G]$ is linearly generated by the monomials $\prod_{i=1}^{n} \tau_i^{\lambda_i}, 0 \leqslant \lambda_i < 2^{m_i}$. Introducing the positive weights $w_i, i = 1, \ldots, n$, we define the power of a monomial $\prod_{i=1}^{n} \tau_i^{\lambda_i}$ as $\sum_{i=1}^{n} w_i \lambda_i$. Then if the sum of degrees of several monomials exceeds

$$\deg_{\max} := \sum_{i=1}^{n} w_i (2^{m_i} - 1),$$

their product equals to zero. This allows to get quite large subspaces in $\mathbb{F}_2[G]$ with zero product. Namely, denote by $X(\theta)$ the span of all monomials of degree strictly greater than $\theta \deg_{\max}$. Then $X(\theta_1)X(\theta_2) \cdot \ldots X(\theta_k) = 0$ provided that $\sum \theta_i \geqslant 1$. Note that $\operatorname{codim} X(\theta)$ equals to the number of monomials of degree at most $\theta \deg_{\max}$. To estimate the number of such monomials, we may use a Chernoff type argument, as follows. If $0 < x \leqslant 1$, we get

$$\operatorname{codim} X(\theta) \leqslant x^{-\theta \deg_{\max}} \prod_{i=1}^{n} (1 + x^{w_i} + x^{2w_i} + \cdots + x^{(2^{m_i}-1)w_i}) =: \Phi_\theta(x).$$

This may be seen from opening the brackets on the right hand side: each monomial $\prod_{i=1}^{n} \tau_i^{\lambda_i}$ of degree at most $\theta \deg_{\max}$ corresponds to a contribution $x^{\sum w_i \lambda_i - \theta \deg_{\max}} \geqslant 1$. Note that if $\theta \leqslant 1/2$, we have $\Phi_\theta(x) \leqslant \Phi_\theta(1/x)$ for $x \geqslant 1$. Thus the minimum of $\Phi_\theta(x)$ over all positive x is attained on $(0, 1]$. Therefore in this case we may write $\operatorname{codim} X(\theta) \leqslant \min_{x \in (0,1]} \Phi_\theta(x)$.

When $G = C_k^n$ for k equal to some power of 2, we may choose the weights $w_1 = w_2 = \cdots = w_n = 1$, so this gives

$$\operatorname{codim} X(\theta) \leqslant \left(\min_{x>0} x^{-\theta(k-1)}(1 + x + x^2 + \cdots + x^{k-1})\right)^n.$$

Using the notation $\mathcal{H}_k(\theta) := \log_2 \min_{x>0} x^{-\theta(k-1)}(1 + x + x^2 + \cdots + x^{k-1})$, this rewrites as

$$\operatorname{codim} X(\theta) \leqslant 2^{n\mathcal{H}_k(\theta)}. \tag{4.5}$$

We will use this quite a few times in this chapter. We note that for $k = 2$ this is the usual binary entropy function

$$\mathcal{H}_2(\theta) = \min_{x>0} -\theta \log_2 x + \log_2(1 + x) = -\theta \log_2 \theta - (1 - \theta) \log_2(1 - \theta).$$

We also note that with all of these notations we may rewrite (4.2) as $\log_2 \kappa_p = \mathcal{H}_p(1/3)$.

Now, consider $\kappa_4 = \min_{x>0} x^{-1}(1 + x + x^2 + x^3) = \min_{x>0} x^{-1}(1 + x)(1 + x^2)$, and let $x_0 > 0$ be a minimizer; that is,

$$\kappa_4 = x_0^{-1}(1 + x_0)(1 + x_0^2) = \left(x_0^{-2\theta}(1 + x_0^2)\right) \cdot \left(x_0^{2\theta-1}(1 + x_0)\right).$$

Here $\theta \in [1/4, 1/2]$ is arbitrary. It follows that $\log_2 \kappa_4 \geqslant \mathcal{H}_2(\theta) + \mathcal{H}_2(1 - 2\theta)$. Taking the maximum over θ we get

$$\log_2 \kappa_4 \geqslant \max_{\theta\in[1/4,1/2]} \mathcal{H}_2(\theta) + \mathcal{H}_2(1 - 2\theta). \tag{4.6}$$

Actually we have an equality in (4.6). This may be explained as follows: choose θ such that the minimum of $x^{-2\theta}(1 + x^2)$ is attained at x_0, this gives $\mathcal{H}_2(\theta) = \log_2 x_0^{-2\theta}(1 + x_0^2)$. Then both the product

$$\left(x^{-2\theta}(1 + x^2)\right) \cdot \left(x^{2\theta-1}(1 + x)\right) = x^{-1}(1 + x + x^2 + x^3)$$

and the first multiple have a critical point at x_0. Thus so does the second multiple, and it is easy to see that it actually attains its minimum at x_0. Therefore $\mathcal{H}_2(1-2\theta) = \log_2 x_0^{2\theta-1}(1 + x_0)$. Hence for this specific value of θ we get $\mathcal{H}_2(\theta) + \mathcal{H}_2(1 - 2\theta) = \log_2 \kappa_4$, and the maximum over all possible values of θ is not less than $\log_2 \kappa_4$, or in other words, (4.6) is an equality. In particular,

$$\log_2 \kappa_4 = \max_{\theta\in[1/4,1/2]} \{\mathcal{H}_2(\theta) + \mathcal{H}_2(1 - 2\theta)\} = \max_{\epsilon\in(0,1/4)} \{\mathcal{H}_2(0.5 - \epsilon) + \mathcal{H}_2(2\epsilon)\},$$

i.e. $\kappa_4 = 4^\gamma$, where

$$\gamma = \frac{1}{2} \cdot \max_{0<\epsilon<0.25} \{\mathcal{H}(0.5 - \epsilon) + \mathcal{H}(2\epsilon)\}.$$

4.3 Croot–Lev–Pach bound for C_4^n with group rings

In this section, we use the subspaces with vanishing product from the previous section to give the promised alternate proof of

$$f_3(C_4^n) \leq (3.611)^n.$$

For reference purposes, we state this formally one more time.

Theorem 4.3.1 *If $A \subset C_4^n$ is a set without non-trivial three-term progressions, then*

$$|A| \leqslant \kappa_4^n,$$

where $\kappa_4 = \min_{x>0} x^{-1}(1 + x + x^2 + x^3) \approx 3.611$.

Proof. From the regularization argument (Lemma 4.1.1) and tensor power trick from Section 2, it is enough to prove that $|A| \leqslant 2\kappa_4^n$ holds whenever A is a regular subset of C_4^n. To do this, we will proceed by contradiction. Assume that $|A| > 2\kappa_4^n$, and let $\alpha > 0$ and $\beta \in [1, 2]$ be such that $\kappa_4 = \alpha \cdot \beta$ and suppose that there are more than β^n classes modulo C_2^n present in A with the property that each class contains more than $2\alpha^n$ elements of A.

We would like to emphasize at this early point that if $a \in A$ belongs to such a class $g_0 \cdot C_2^n$, then $g_0^2 = a^2$, so $\beta^n < |A^2|$, where A^2 denotes the set $\left\{x^2 : x \in A\right\}$. Next, choose $\theta \in [1/4, 1/2]$ such that $\log_2 \beta = \mathcal{H}_2(1 - 2\theta)$. Consider the subspaces $X(1 - 2\theta), X(\theta)$ in $\mathbb{F}_2[C_2^n]$ defined in Section 3. By (4.5),

$$\operatorname{codim} X(1 - 2\theta) \leqslant 2^{n\mathcal{H}_2(1-2\theta)} = \beta^n < |A^2|,$$

$X(1 - 2\theta)$ must have a common non-zero element with the subspace of $\mathbb{F}_2$-valued functions supported on $A^{-2} = \{h^{-1} \mid h \in A^2\}$. In other words, there exists a non-zero element of the form

$$\sum_{h \in A^2} \eta(h^{-1})h^{-1} \in X(1 - 2\theta). \tag{4.7}$$

Fix $g_0 \in A$ such that $\eta(g_0^{-2}) \neq 0$, and let $C = g_0 \cdot C_2^n \cap A$; by our assumption on A, we know that $|C| > 2\alpha^n$.

Consider the product

$$\left(\sum_{g \in C} \varphi(g)g_0^{-1}g\right)\left(\sum_{g \in C} \psi(g)g_0^{-1}g\right)\left(\sum_{h \in A^2} \eta(h^{-1})h^{-1}\right) \tag{4.8}$$

inside the group algebra $\mathbb{F}_2[C_4^n]$, where the functions $\varphi, \psi : C \to \mathbb{F}_2$ are chosen so that

$$\sum_{g \in C} \varphi(g) g_0^{-1} g, \ \sum_{g \in C} \psi(g) g_0^{-1} g \in X(\theta). \tag{4.9}$$

This product equals to 0, since $X(\theta)X(\theta)X(1-2\theta) = 0$. On the other hand, A does not contain non-trivial three-term progressions, so the coefficient of g_0^{-2} in this product also equals $\sum_{g \in C} \varphi(g)\psi(g)\eta(g_0^{-2})$. Together with $\eta(g_0^{-2}) \neq 0$, this yields

$$\sum_{g \in C} \varphi(g)\psi(g) = 0$$

for every φ, ψ satisfying (4.9). However, the vector subspace of $\mathbb{F}_2^C$ spanned by the functions φ with $\sum_{g \in C} \varphi(g) g_0^{-1} g \in X(\theta)$ has codimension at most $\operatorname{codim} X(\theta)$, and so does the subspace spanned by the functions ψ such that $\sum_{g \in C} \psi(g) g_0^{-1} g \in X(\theta)$. By (4.5), the sum of their codimensions is at most $2 \cdot \operatorname{codim} X(\theta) \leqslant 2 \cdot 2^{n\mathcal{H}_2(\theta)}$, while $\log_2 \beta = \mathcal{H}_2(1-2\theta)$, which by (4.6) yields $\log_2 \alpha \geqslant \mathcal{H}_2(\theta)$. Putting these together, we conclude that the sum of the codimensions of these spaces is at most

$$2 \cdot \operatorname{codim} X(\theta) \leqslant 2 \cdot 2^{n\mathcal{H}_2(\theta)} \leqslant 2\alpha^n < |C|,$$

which is a contradiction, since this means the subspaces can't be orthogonal with respect to the bilinear form $\sum_{g \in C} \varphi(g)\psi(g)$. This completes the proof of Theorem 4.3.1.

For the proof of Theorem 4.0.1, we will require a few additional tools.

4.4 Improved bounds for progression-free sets in C_8^n

In this section, we present the proof of Theorem 4.0.1. We begin with some further linear algebraic preliminaries.

Lemma 4.4.1 *If X_1, X_2 are the subspaces of a linear space X over certain field, the codimension of the subspace $X_1 \cap X_2$ in X_2 does not exceed* $\operatorname{codim} X_1$.

Proof of Lemma 4.4.1. The space X_1 is a set of vectors in X satisfying certain $m := \operatorname{codim} X_1$ linear equations. The vectors in X_2 satisfying these m equations form a subspace of X_2 of codimension at most m. This proves Lemma 4.4.1.

Let Ω be a finite set, K be a fixed field and K^Ω a space of K-valued functions on Ω. For a function $f \in K^\Omega$ denote by $\operatorname{supp}(f) = \{x \in \Omega : f(x) \neq 0\}$ the support of f.

Lemma 4.4.2 *Suppose that $X \subset K^\Omega$ is a space of dimension d. Then X contains a function f with $|\operatorname{supp}(f)| \geqslant d$.*

While simple, this observation was an important step in the Ellenberg-Gijswijt argument from [**EG17**]. We record the short proof here for the reader's convenience.

Proof of Lemma 4.4.2. Consider $f \in X$ with maximal value of $|\operatorname{supp}(f)|$. If $|\operatorname{supp}(f)| < d$, the number of equations $g(x) = 0$ for $x \in \operatorname{supp}(f)$ is less than the dimension of X; in particular, there exists a non-zero function $g \in X$ which vanishes on $\operatorname{supp}(f)$. But then

$$|\operatorname{supp}(f+g)| > |\operatorname{supp}(f)|,$$

which contradicts the choice of f. This proves Lemma 4.4.2.

Last but not least, we will also need a generalization of a fact which we used at the end of the proof of Theorem 4.3.1.

Lemma 4.4.3 *Suppose that $a \in K^\Omega$ is a function for which the subspaces $X, Y \subset K^\Omega$ satisfy the condition $\sum_{x\in\Omega} a(x)f(x)g(x) = 0$ for all $f \in X, g \in Y$. Then,*

$$\operatorname{codim} X + \operatorname{codim} Y \geqslant |\operatorname{supp}(a)|.$$

Proof of Lemma 4.4.3. Denote $\Omega_0 = \operatorname{supp}(a)$. There is a natural embedding of K^{Ω_0} into K^Ω. By Lemma 4.4.1, the subspaces $X_0 = X \cap K^{\Omega_0}$, $Y_0 = Y \cap K^{\Omega_0}$ have codimensions in K^{Ω_0} at most $\operatorname{codim} X, \operatorname{codim} Y$ respectively. But they are orthogonal subspaces with respect to the full rank bilinear form

$$\langle f, g\rangle := \sum_{x\in\Omega_0} a(x)f(x)g(x).$$

Thus the sum of their codimensions is at least $|\Omega_0| = |\operatorname{supp}(a)|$, and the statement of Lemma 4.4.3 is proved.

Using the subgroup generated by squares. We move on to showing a general lemma about progression-free sets in finite groups, which is the key to our arguments and which may be of independent interest.

Let G be a finite group, and let $H = \{g^2 : g \in G\}$. We assume that H is a subgroup of G (in particular, this is so in the abelian case, or for the groups of odd order,

when simply $H = G$). In this case, H is a normal subgroup due to the identity $hg^2h^{-1} = (hgh^{-1})^2$. Furthermore, fix an arbitrary field K. For a subset $A \subset G$, we identify A^K with the span of A as a subset of the group algebra $K[G]$. In particular, we have that $H^K = K[H]$.

Lemma 4.4.4 *Let X, Y, Z be subspaces of $K[H]$ which satisfy $XYZ = 0$. Suppose $A \subset G$ satisfies the following conditions:*

(i) $|A^2| \geq 5 \cdot \operatorname{codim} Y$;

(ii) all elements of A^2 have the same number of square roots in A;

(iii) each H-coset contains either no elements of A or more than $\frac{5}{4}(\operatorname{codim} X + \operatorname{codim} Z)$ elements of A.

Then, A contains a three-term progression.

Proof of Lemma 4.4.4. Suppose that A does not contain three-term progressions. First, note that $A^2 \subset H$. If A^{-2} once again denotes the set $\left\{h^{-1} \mid h \in A^2\right\}$, fix a function $\eta : A^{-2} \to K$ such that $\sum_{c \in A^{-2}} \eta(c)c$ belongs to Y and with the property that

$$|\operatorname{supp}(\eta)| \geqslant |A^{-2}| - \operatorname{codim} Y \geqslant \frac{4}{5}|A^{-2}|.$$

Such a map η exists by Lemma 4.4.2. In the second inequality, we made use of condition (i). For convenience, let $y_0 := \sum_{c \in A^{-2}} \eta(c)c$. Furthermore, consider an arbitrary coset $g_0H = Hg_0$ and choose two arbitrary functions $\varphi : A \cap g_0H \to K$, $\psi : A \cap g_0H \to K$ such that

$$x := \sum_{a \in g_0H} \varphi(a) g_0^{-1} a \in X \quad \text{and} \quad z := \sum_{b \in Hg_0} \psi(b) b g_0^{-1} \in Z.$$

Since $XYZ = 0$, we have that $xy_0z = 0$, so the coefficient of g_0^{-2} in this product equals 0. On the other hand, it equals

$$\sum_{a \in g_0H \cap A, b \in g_0H \cap A, c \in A^{-2} : acb = 1} \varphi(a)\eta(c)\psi(b).$$

However, A does not contain three-term progressions, so $acb = 1$ implies that $a = b, c = a^{-2}$. In particular, we get that

$$\sum_{a \in g_0H \cap A} \varphi(a)\psi(a)\eta(a^{-2}) = 0. \tag{4.10}$$

We claim that for a certain $g_0 \in G$, this is a contradiction with Lemma 4.4.3. To see this, recall first that the choice of η and (ii) assured us that at least $\frac{4}{5}|A|$ elements $a \in A$ are such that $a^{-2} \in \mathrm{supp}(\eta)$. By the pigeonhole principle, this means that there exists a coset g_0H such that at least $\frac{4}{5}|A \cap g_0H|$ elements of $A \cap g_0H$ satisfy this condition. Since the vector space spanned by the functions $\Psi \in K^{A\cap g_0H}$ such that $\sum_{a\in g_0H} \Psi(a)g_0^{-1}a \in X$ (respectively, such that $\sum_{b\in Hg_0} \Psi(b)bg_0^{-1} \in Z$) has codimension at most $\mathrm{codim}\, X$ (respectively, $\mathrm{codim}\, Z$), Lemma 5.1 and (4.10) imply that

$$\mathrm{codim}\, X + \mathrm{codim}\, Z \geqslant \left|\left\{a \in A \cap g_0H \ : \ a^{-2} \in \mathrm{supp}(\eta)\right\}\right| \geqslant \frac{4}{5}|A \cap g_0H|,$$

a contradiction with (iii). This proves Lemma 4.4.4.

We will use this lemma to first complete the proof of Theorem 4.0.1.

Proof of Theorem 4.0.1. From the regularization argument (Lemma 4.1.2) and the tensor power trick from Section 2, it is enough to prove $|A| \leqslant 10.05 \cdot (7.09)^n$ in the case when A is a super-regular subset of C_8^n without three-term progressions. Accordingly, suppose that A is covered by $4.02 \cdot \gamma^n$ classes modulo $C_4^n = G^2$, where each such class contains itself $\frac{5}{4.02} \cdot \beta^n$ classes modulo $C_2^n = G^4$, with the property that each subclass modulo C_2^n intersects A in precisely $2.01 \cdot \alpha^n$ elements. In particular, $|A| = 10.05 \cdot (\alpha\beta\gamma)^n, |A^2| = 5 \cdot (\gamma\beta)^n, |A^4| = 4.02 \cdot \gamma^n$. In this setup, note that we may also assume that α, β, γ are all in the interval $(1, 2]$. Indeed, the fact that $\alpha, \beta, \gamma \leqslant 2$ is clear since there are at most 2^n cosets of C_4^n inside C_8^n, and at most 2^n cosets of C_2^n inside C_4^n (and the C_2^n-cosets meet A in at most 2^n elements). Also, if $\min\{\alpha, \beta, \gamma\} \leqslant 1$, we get that $|A| = O(4^n)$, so we can assume from now on that $\alpha, \beta, \gamma \in (1, 2]$. Furthermore, note that for each class C modulo C_4^n which intersects A, we already have an upper bound for $|A \cap C|$. By shifting $A \cap C$ by a suitable element of C_8^n, we can send $A \cap C$ inside the trivial coset of C_4^n inside C_8^n. This operation preserves the property of not containing three-term progressions, so we can apply Theorem 4.3.1 to write $|A \cap C| \leqslant (\kappa_4)^n$. The same bound also follows trivially from the super-regularity of A, since $A \cap C$ already has the same size as the intersection of A with any coset of C_4^n inside C_8^n, however we can use the super-regular structure of A more efficiently.

In light of the above, suppose without loss of generality that $A \cap C \subset C_4^n$, and let $\theta \in [1/4, 1/2]$ be such that $\log_2 \beta = \mathcal{H}_2(1-2\theta)$. We first claim that $\log_2 \alpha < \mathcal{H}_2(\theta)$.

This follows in fact by applying the argument from Section 4 to $A \cap C \subset C_4^n$. Indeed, if $\log_2 \alpha \geq \mathcal{H}_2(\theta)$ we have that

$$\operatorname{codim} X(1-2\theta) \leq 2^{n\mathcal{H}_2(1-2\theta)} = \beta^n \quad \text{and} \quad \operatorname{codim} X(\theta) \leq 2^{n\mathcal{H}_2(\theta)} \leq \alpha^n,$$

so we can consider once again the (zero) product from (4.8) for a suitable intersection A_{g_0} of $A \cap C$ with a coset of C_2^n. Similarly, the fact that $A \cap C$ has no three-term progressions then produces two spaces of functions, Φ and Ψ, each with codimension at most $\operatorname{codim} X(\theta)$ in $\mathbb{F}_2^{A_{g_0}}$, which must also be orthogonal with respect to the bilinear form $\sum_{g \in A_{g_0}} \varphi(g)\Psi(g)$. However, the super-regularity of A and Lemma 5.1 then imply

$$2.01 \cdot \alpha^n = |A_{g_0}| \leq \operatorname{codim} \Phi + \operatorname{codim} \Psi \leq 2 \cdot \operatorname{codim} X(\theta) \leq 2\alpha^n,$$

which is a contradiction. Consequently, $\log_2 \alpha < \mathcal{H}_2(\theta)$, as claimed.

Next, consider $\rho \in [1/4, 1/2]$ such that $\log_2 \gamma\beta = \mathcal{H}_4(1-2\rho)$. Note that

$$|A^2| = 5 \cdot (\gamma\beta)^n = 5 \cdot 2^{n\mathcal{H}_4(1-2\rho)} \geq 5 \cdot \operatorname{codim} Y.$$

Applied for $\mathbb{F}_2[C_4^n]$ and the subspaces $Y = X(1-2\rho)$, $X = Z = X(\rho)$, Lemma 4.4.4 thus yields

$$\frac{5}{2} \cdot \alpha^n \beta^n = (2.01 \cdot \alpha^n) \left(\frac{5}{4.02} \cdot \beta^n \right) \leq \frac{5}{2} \cdot \operatorname{codim} X(\rho) \leq \frac{5}{2} \cdot 2^{n\mathcal{H}_4(\rho)},$$

which implies

$$\log_2 \alpha\beta \leqslant \mathcal{H}_4(\rho).$$

This condition imposes a special further constraint on α, β, γ, and maximizing the product $\alpha\beta\gamma$ requires a delicate analysis which will be covered in the next subsection. For now, let us just argue

$$\log_2 \alpha\beta\gamma < c < 1 + \log_2 \kappa_4,$$

for certain c, i.e. $r_3(C_8^n) = O(c^n)$, where $c < 2\kappa_4 \approx 7.222$. The analysis below will show roughly that if $\log_2 \alpha\beta\gamma$ is close to $1 + \log_2 \kappa_4$, then γ must be close to 2, while $\alpha\beta$ must be close to $2^{\kappa_4} = \mathcal{H}_4(1/3)$. Therefore $\log_2 \alpha\beta < \mathcal{H}_4(\rho)$ implies that $\rho \geqslant 1/3 + o(1)$, and

$$\log_2 2\beta = \log_2 \gamma\beta + o(1) = \mathcal{H}_4(1-2\rho) + o(1) \leqslant \mathcal{H}_4(1/3) + o(1) = \log_2 \alpha\beta + o(1),$$

which represents a contradiction.

Maximizing $\alpha\beta\gamma$. For the reader's convenience, let us first recall the restrictions we have on α, β and γ; in the previous subsection, we showed that there exist positive reals $\rho, \theta \in [1/4, 1/2]$ such that

$$\begin{cases} \log_2 \beta & = \mathcal{H}_2(1-2\theta) \\ \log_2 \alpha & \leqslant \mathcal{H}_2(\theta) \\ \log_2 \gamma\beta & = \mathcal{H}_4(1-2\rho) \\ \log_2 \alpha\beta & \leqslant \mathcal{H}_4(\rho). \end{cases} \tag{4.11}$$

The maximal value of $\alpha\beta\gamma$ for $\rho, \theta \in [1/4, 1/2]$, $\alpha, \beta, \gamma \in [1, 2]$ and (4.11) is achieved. Denote the corresponding point $(\rho_0, \theta_0, \alpha_0, \beta_0, \gamma_0)$. Assume that $\gamma_0 < 2$. If $\beta_0 = 1$, then we have $\alpha_0\beta_0\gamma_0 \leqslant 4$, which is definitely not a maximum, thus $\beta_0 > 1$. Choose γ slightly greater than γ_0 and $\beta < \beta_0$ so that $\beta_0\gamma_0 = \beta\gamma$. Since the binary entropy function $\mathcal{H}_2$ is increasing on $[0, 1/2]$, the new θ such that $\log_2 \beta = \mathcal{H}_2(1-2\theta)$ satisfies $\theta > \theta_0$. In particular, this means that there exists $\alpha > \alpha_0$ such that $\log_2 \alpha \leqslant \mathcal{H}_2(\theta)$ and $\alpha\beta \leqslant \alpha_0\beta_0$. We have $\alpha\beta\gamma = \alpha\beta_0\gamma_0 > \alpha_0\beta_0\gamma_0$, a contradiction with maximality. Therefore the maximum is achieved for $\gamma_0 = 2$. If $\alpha_0 = 2$, we get $\theta_0 = 1/2$, $\beta_0 = 1$ and $\alpha_0\beta_0\gamma_0 = 4$, too small for a maximum.

Next, we claim that for the point $(\rho_0, \theta_0, \alpha_0, \beta_0, 2)$ which maximizes the value $\alpha\beta\gamma$ the second inequality from (4.11) must be an equality. We argue this again by contradiction; suppose that $\log_2 \alpha_0 < \mathcal{H}_2(\theta_0)$. Then we may choose β slightly less than β_0, define ρ by $\log_2 2\beta = \mathcal{H}_4(1-2\rho)$ and θ by $\log_2 \beta = \mathcal{H}_2(1-2\theta)$. After that we may choose $\alpha \in (\alpha_0\beta_0/\beta, 2)$ so that (4.11) still holds for α, β (and $\gamma = \gamma_0 = 2$), which yields a contradiction. This is indeed clear when $\log_2 \alpha_0\beta_0 < \mathcal{H}_4(\rho_0)$, but even if we had equality in the last line from (4.11), namely $\log_2 \alpha_0\beta_0 = \mathcal{H}_4(\rho_0)$, then we can choose α so that

$$\log_2 \alpha\beta = \mathcal{H}_4(\rho) > \mathcal{H}_4(\rho_0) = \log_2 \alpha_0\beta_0.$$

Therefore, $\log_2 \alpha_0 = \mathcal{H}_2(\theta_0)$. We also claim that equality must hold in the last inequality from (4.11). Suppose that $\log_2 \alpha_0\beta_0 < \mathcal{H}_4(\rho_0)$. The function $\mathcal{H}_2(1-2x) + \mathcal{H}_2(x)$ is concave on $[1/4, 1/2]$, so it has a unique point of maximum, which we call x_0 just like in Section 3. If $\theta_0 \neq x_0$, we may perturb the pair (α, β) slightly so that the product $\alpha\beta$ increases and the conditions from (4.11) still hold (with $\log_2 \alpha = \mathcal{H}_2(\theta)$). If $\theta_0 = x_0$, we have

$$\log_2 \alpha_0\beta_0 = \max_{x \in [1/4, 1/2]} \{\mathcal{H}_2(1-2x) + \mathcal{H}_2(x)\} = \mathcal{H}_4(1/3),$$

so $\rho_0 \geqslant 1/3$, but then by the analysis from Section 3

$$\mathcal{H}_4(1/3) \geqslant \mathcal{H}_4(1-2\rho) \geqslant \log_2 \gamma_0\beta_0 = \log_2 2\beta_0 > \log_2 \alpha_0\beta_0 = \mathcal{H}_4(1/3),$$

which is once again a contradiction.

We have thus proved that $\gamma_0 = 2$, $\log_2 \alpha_0 = \mathcal{H}_2(\theta_0)$, $\log_2 \alpha_0\beta_0 = \mathcal{H}_4(\rho_0)$. Finally, let us assume that we found certain $\theta_1 \in [x_0, 1/2]$ and $\rho_1 \in [1/4, 1/2]$ satisfying

$$\mathcal{H}_4(\rho_1) = \mathcal{H}_2(\theta_1) + \mathcal{H}_2(1-2\theta_1),\ \mathcal{H}_4(1-2\rho_1) = 1 + \mathcal{H}_2(1-2\theta_1). \qquad (4.12)$$

We claim that $\theta_1 = \theta_0, \rho_1 = \rho_0$. We argue this one last time by contradiction. If $\theta_0 < x_0 \leqslant \theta_1$, note that we get

$$\mathcal{H}_4(1-2\rho_1) = 1 + \mathcal{H}_2(1-2\theta_1) < 1 + \mathcal{H}_2(1-2\theta_0) = \mathcal{H}_4(1-2\rho_0),$$

therefore $\rho_1 > \rho_0$, and we may replace α_0 and β_0 by α and β defined by $\log_2 \alpha = \mathcal{H}_2(\theta_1)$, $\log_2 \beta = \mathcal{H}_2(1-2\theta_1)$, with $\alpha\beta > \alpha_0\beta_0$, contradicting the maximality of $\alpha_0\beta_0$. If $\theta_0 \geqslant x_0$, both functions $\mathcal{H}_2(x) + \mathcal{H}_2(1-2x)$ and $1 + \mathcal{H}_2(1-2x)$ decrease on the segment $[x_0, 1/2]$ containing both θ_0 and θ_1. This implies that if, say, $\theta_0 < \theta_1$, we get $\rho_0 > \rho_1$ and $1 - 2\rho_0 > 1 - 2\rho_1$, which is also impossible.

To pinpoint our optimizer $(\rho_0, \theta_0, \alpha_0, \beta_0, \gamma_0)$, we therefore look for $\theta_1 \in [x_0, 1/2]$ and $\rho_1 \in [1/4, 1/2]$ satisfying (4.12). The first equation defines ρ_1 as a (strictly) decreasing function of θ_1, whereas the second equations represents it as an increasing one. Thus such θ_1 is (a priori at most) unique and the approximate estimates may be specified by Intermediate Value Theorem. Numerically, the values of θ_1, ρ_1 and $2^{\mathcal{H}_4(\rho_1)+1} = 2\alpha_0\beta_0$ are about $\theta_1 \approx 0.343$, $\rho_1 \approx 0.32$, $2\alpha_0\beta_0 \approx 7.0899$. Putting everything together, we can finally conclude that

$$|A| = 10.05 \cdot (\alpha\beta\gamma)^n \leq 10.05 \cdot (7.0899)^n,$$

which completes the proof of Theorem 4.0.1.

4.5 A large 3AP-free set in $\mathbb{Z}_8^n$

Finding examples of large sets inside C_8^n without non-trivial three-term progressions is also quite an interesting problem. As with C_3^n, where the best lower bound is due to Edel [**Ede04**], one would be tempted to find the largest possible three-term progression free set in C_8^k for a few small values of k, and then output the best

cartesian product construction. We believe all such attempts lead to lower bounds of the form

$$r_3(C_8^n) = \Omega\left(c^n\right),$$

where $c < 5$. We can do better by using a Behrend-type construction. We switch to additive notation for convenience.

Theorem 4.5.1 *Suppose that $G = \mathbb{Z}_8^n$. Then there is a set $A \subset G$ with no three-term progression and*

$$|A| = \Omega\left(5^n/\sqrt{\log n}\right).$$

In other words, $f_3(\mathbb{Z}_8^n) = \Omega\left(5^n/\sqrt{\log n}\right)$.

The construction is similar to the lower bound construction for $f_3(\mathbb{Z}_4^n)$ described in Section 1.3.

Proof of Theorem 4.5.1. Consider the set $S \subset \mathbb{Z}^n$ consisting of the points $(x_1, \ldots, x_n) \in \{0, 1, 2, 3, 4\}^n$ with the property that

$$\sum_{i=1}^{n} (x_i - 2)^2 = 2n.$$

In other words, S is the intersection of $\{0, 1, 2, 3, 4\}^n$ with the n-dimensional hypersphere centered at $(2, \ldots, 2)$ and radius $n\sqrt{2}$. In particular, no three points in S are collinear. Moreover, the size of $|S|$ is at $\Omega(5^n/\sqrt{n})$, as one can easily see from the Central Limit Theorem. Indeed, let X be the random variable which takes values $0, 1, 2, 3, 4$ with probability $1/5$ each; let $X_1, \ldots, X_n$ be n independent copies of X and let $Y_i = (X_i - 2)^2$ for each $i = 1, \ldots n$. It is easy to see that $\mathbb{E}[Y_i] = 2$, so $|S|/5^n$ is the probability that that $Y_1 + \ldots + Y_n = 2n$.

Consider the identity map $\Psi : \{0, 1, 2, 3, 4\}^n \to (\mathbb{Z}/8\mathbb{Z})^n$ and let A denote the image of S. We claim that A does not contain non-trivial three-term $(\mathbb{Z}/8\mathbb{Z})^n$ arithmetic progressions. To see this, note that if $a + c = 2b$, with $a \neq c$, then either $\Psi^{-1}(a)$, $\Psi^{-1}(b)$, $\Psi^{-1}(c)$ is a three-term progression in $\mathbb{Z}^n$ or there must be a nonempty subset $I \subset \{1, \ldots, n\}$ such that

$$(\Psi^{-1}(a)_i, \Psi^{-1}(b)_i, \Psi^{-1}(c)_i) \in \{(4, 0, 4), (0, 4, 0)\}$$

for every $i \in I$. The former scenario is impossible, since S does not contain three points in arithmetic progression. If the latter happens, we let $a', b', c' \in \{0, 1, 2, 3, 4\}^n$

be the points obtained from a, b, c by swapping $(4, 0, 4)$ with $(4, 4, 4)$ and/or by swapping $(0, 4, 0)$ with $(0, 0, 0)$ for each coordinate i where $\Psi^{-1}(a)_i, \Psi^{-1}(b)_i, \Psi^{-1}(c)_i$ is a three-term progression in $\mathbb{Z}/8\mathbb{Z}$ but not in $\mathbb{Z}$. Note that if a, b, c lie on a hypersphere centered at $(2, \ldots, 2)$, the points a', b', c' must also lie on the same hypersphere. However, if $a + c = 2b$ holds in $(\mathbb{Z}/8\mathbb{Z})^n$ then $a' + c' = 2b'$ must also hold in $\mathbb{Z}^n$, and this is again impossible. This proves Theorem 4.5.1.

We end this chapter with an intriguing open problem.

Question. *Are there examples of subsets A of $\mathbb{Z}_8^n$ which do not contain nontrivial three-term progressions and for which $|A| = \Omega\left(6^n/\sqrt{\log n}\right)$?*

Chapter 5

4AP-FREE SETS WITH 3AP'S IN ALL LARGE SUBSETS

In this chapter, we consider a problem in this direction but with a slightly different flavour. Let $k \geq 3$ be an integer and suppose that we have a set A in a group G which does not contain any $(k+1)$-APs. Is it always possible to find a large subset of A which does not contain any k-APs? Or using the notation we have established, is it always the case that $f_k(A)$ is large when A is $(k+1)$-AP free?

Perhaps a first intuitive guess is that the answer should be "yes", and that all k-APs can be destroyed by deleting a relatively small number of elements of A. Focusing on the situation when $k = 3$ and G is $\mathbb{Z}$ or $\mathbb{F}_q^n$, the results of this chapter, based on my joint work with Oliver Roche-Newton [**PRN20**], give quantitative answers to this question in the negative direction. Our main result is perhaps the following.

Theorem 5.0.1 *For all $\beta > 0$, there exists $n_0 = n_0(\beta)$ such that the following statement holds for all $n \geq n_0$ and for any prime power q. There exists a four-term progression free set $A \subset \mathbb{F}_q^n$ such that*

$$f_3(A) \leq |A|^{1-\frac{1}{2(C_q-2)}+\beta}.$$

That is, we show the existence of a set $A \subset \mathbb{F}_q^n$ which does not contain a non-trivial 4-AP but for which every large subset $A' \subset A$ contains a non-trivial 3-AP. The positive constant C_q depends on the aforementioned constant c_q via

$$C_q = 1 + \frac{1}{c_q}.$$

For a concrete example, one can calculate that $C_5 \approx 15.12589$, meaning that every set $A' \subset A$ larger than $|A|^{0.962}$ contains a 3-AP.

Our proof relies on an iterated application of the so-called hypergraph container theorem, which we will describe in the next section, and which takes as input a supersaturated version of the subexponential Ellenberg-Gijswijt upper bound for 3-AP free subsets of $\mathbb{F}_q^n$ from (1.7). In fact, we will derive Theorem 5.0.1 from a more general result about random subsets of $\mathbb{F}_q^n$, in the spirit of Kohayakawa-Luczak-Rődl [**KLR96**] and Conlon-Gowers [**CG16**].

Theorem 5.0.2 *Let $\beta > 0$, $t < c_q(1-2\beta)$ and let p be a positive real number satisfying*

$$q^{n\left(-\frac{1}{2}+\frac{t(C_q-1)}{2}\right)} \leq p \leq 1.$$

Let B be a random subset of $\mathbb{F}_q^n$ with the events $x \in B$ being independent with probability $\mathbb{P}(x \in B) = p$. Then, with probability $1 - o_{n\to\infty}(1)$ we have that

$$f_3(B) \ll pq^{n(1-t+2\beta)}.$$

In particular, for all $\epsilon > 0$, there exists $\delta(\epsilon, q) := \delta > 0$ such that if B is defined as above with $p = q^{n(-\frac{1}{2}+\epsilon)}$, then with probability $1 - o_{n\to\infty}(1)$,

$$f_3(B) \ll |B|^{1-\delta}.$$

This allows us to detect three-term arithmetic progressions in subsets of $\mathbb{F}_q^n$ of size as small as $q^{n\left(\frac{1}{2}+\epsilon\right)}$, which is beyond the reach of the Ellenberg-Gijswijt bound (1.7), provided that those subsets have large relative density compared to a random set. This improves a result of Tao and Vu from [**TV06**]. It is also worth pointing out that the range for p in Theorem 5.0.2 is optimal. Indeed, if $p = q^{-n/2}/2$ then the expected number of three-term progressions in a random subset B of $\mathbb{F}_q^n$ (where each element in B chosen independently with probability p) is less than $q^{n/2}/8$, while the expected number of elements in B is $q^{n/2}/2$. Therefore, one can almost always remove an element from each progression and still be left with at least half the elements of B.

We also consider the analogue of Theorem 5.0.1 in the integer setting, where we obtain the following result.

Theorem 5.0.3 *For all $\alpha > 0$ and for all $N \in \mathbb{N}$ sufficiently large (depending on α), there exists a set of integers A with $|A| = N$ which does not contain any nontrivial four-term arithmetic progression, and for which*

$$f_3(A) \ll \frac{1}{(\log N)^{1-\alpha}} \cdot N. \tag{5.1}$$

It is important to mention that in the integer setting, if merely a sublinear upper bound on $f_3(A)$ would be the goal, one could could pretty easily explicitly describe a set of integers A with no four-term progressions for which the powerful density Hales-Jewett theorem ensures that all of its relatively dense subsets share the same property; consider, for instance, the subset of the first N integers with only digits

0, 1 or 2 in base 6. This is a 4-AP-free sets A for which indeed $f_3(A) = o(|A|)$ but the asymptotic notation doesn't hide good bounds. In the (non-quantitative) direction, a much more general statement was also recently established by Balogh, Liu and Sharifzadeh in [**BLS17**], who show that for all $k \geq 3$, there exists a set S of primes such that S is $(k+1)$-AP free, and $f_k(S) = o(|S|)$. Theorem 5.0.3 should perhaps be thought of as follows: there exist sets of N integers without non-trivial four-term progressions for which the size of the largest 3-AP free subset is smaller than roughly the best upper bound known for $r_3(N)$.

After discussing the required ingredients in Sections 5.2 and 5.3, we prove Theorems 5.0.1 and 5.0.3 in Sections 5.4 and 5.5, respectively. In Section 5.6, we will discuss another application of our methods, showing that for sets (in $\mathbb{F}_q^n$ or $\mathbb{Z}$) the property of "having nontrivial three-term progressions in all large subsets" is almost entirely uncorrelated with the property of "having large additive energy". In particular, we prove the existence of sets A with minimal additive energy and small $f_3(A)$.

5.1 Hypergraph containers

A critical tool in this chapter comes from the theory of hypergraph containers. The statement that we use is rather technical, but it can be roughly summarised as follows: if a hypergraph $H = (V, E)$ has a good edge distribution (in the sense that no vertices have unusually large degree, and more generally the elements of any set of vertices do not share too many common edges) then we obtain strong information about the independent sets of the hypergraph. This strong information is that there is a family C of subsets of V such that

- For every independent set $X \subset V$, there is some $A \in C$ such that $X \subset A$,
- C is not too large,
- Each $A \in C$ does not have too many edges.

The wonderful theory of hypergraph containers was developed independently by Balogh, Morris and Samotij [**BMS15**] and Saxton and Thomasson [**ST15**]. For a recent survey on this topic, see [**BMS18**]. This method has led to several significant breakthroughs in combinatorics in recent years, most notably in the field of extremal graph theory. However, this purely combinatorial tool has also led to new results in additive combinatorics. For example, it was proven by Balogh, Liu and Sharifzadeh

[**BLS17**] that, for infinitely many $N \in \mathbb{N}$ there are at most $2^{O(r_k(N))}$ subsets of $[N]$ which do not contain a k-AP. Note that this is almost best possible, since any subset of a k-AP free set is k-AP free, and so the subsets of the largest k-AP free set give at least $2^{r_k(N)}$ sets which are k-AP free. Another application of containers closely related to (and which inspired) this project can be found in Balogh and Solymosi [**BS19**], where it was proven that there exists a set P of N points in the plane such that P does not contain any collinear quadruples, but any subset of P of size larger than $N^{5/6+o(1)}$ contains a collinear triple.

In order to state the required hypergraph container result formally, we need to introduce some more notation. Let $H = (V, E)$ be an r-uniform hypergraph. Write $e(H) = |E|$. For any $S \subset V$, the subhypergraph induced by S is denoted $H[S]$. The co-degree of S is the quantity

$$d(S) := |\{e \in E : S \subseteq e\}|.$$

In the case when $S = \{v\}$ is a singleton, we simply write $d(v)$. The average degree of a vertex in H is denoted by d, that is,

$$d = \frac{1}{|V|}\sum_{v\in V} d(v) = \frac{r|E|}{|V|}.$$

For each $2 \leq j \leq r$, denote

$$\Delta_j(H) := \max_{S\subset V:|S|=j} d(S).$$

For $0 < \tau < 1$, define the function

$$\Delta(H,\tau) = 2^{\binom{r}{2}-1}\sum_{j=2}^{r}\frac{\Delta_j(H)}{2^{\binom{j-1}{2}}d\tau^{j-1}}.$$

This function gives a measure of how well-distributed the edges of H are. In this thesis, we will only consider 3-uniform hypergraphs, in which case the function can be expressed more straightforwardly:

$$\Delta(H,\tau) = \frac{4\Delta_2(H)}{d\tau} + \frac{2\Delta_3(H)}{d\tau^2}.$$

The exact result that we will use is Corollary 3.6 in [**ST15**].

Theorem 5.1.1 *Let $H = (V, E)$ be and r-uniform hypergraph with $|V| = N$. Let $0 < \epsilon, \tau < 1/2$ satisfy the conditions that*

- $\tau < 1/(200 \cdot r \cdot r!^2)$,
- $\Delta(H, \tau) \leq \frac{\epsilon}{12r!}$.

Then there exists $c = c(r) \leq 1000 \cdot r \cdot r!^3$ *and a collection* $\mathcal{C}$ *of subsets of* $V(H)$ *such that*

- *If* $X \subseteq V$ *is an independent set then there is some* $A \in \mathcal{C}$ *such that* $X \subseteq A$,
- *for every* $A \in \mathcal{C}$, $e(H[A]) \leq \epsilon e(H)$,
- $\log |\mathcal{C}| \leq cN\tau \cdot \log(1/\epsilon) \cdot \log(1/\tau)$.

5.2 Supersaturation Results for 3APs

In most applications of the container method, a crucial ingredient is a so-called Supersaturation Lemma. Extremal results in combinatorics often state that sufficiently large subsets of a given set contain at least one copy of some special structure. A supersaturation result goes further, and says that sufficiently dense subsets of a given set contain *many* copies of certain structures.

In our particular setting we can be more concrete. We need to prove that sufficiently large subsets of $\mathbb{F}_q^n$ and $[N]$ contain many 3-APs. The results and techniques in these two different settings differ significantly, particularly in light of recent developments concerning the size of the largest 3AP-free set in $\mathbb{F}_q^n$ in [**CLP17**] and [**EG17**].

Supersaturation in $\mathbb{F}_q^n$. We begin by finally defining the previously mentioned constant c_q by

$$q^{1-c_q} = \inf_{0<y<1} \frac{1 + y + \cdots + y^{q-1}}{y^{(q-1)/3}}.$$

Also, recall that $C_q := 1 + \frac{1}{c_q}$. For a fixed q, these constants c_q and C_q can be calculated explicitly.

Define a *triangle* in $\mathbb{F}_q^n$ to be a triple $(x, y, z) \in \mathbb{F}_q^n \times \mathbb{F}_q^n \times \mathbb{F}_q^n$ such that $x + y + z = 0$. To obtain a supersaturation result for arithmetic progressions in $\mathbb{F}_q^n$, we will make use of the following result of Fox and Lovász [**FL17**].

Theorem 5.2.1 *Let* $0 < \epsilon < 1$ *and* $\delta = (\epsilon/3)^{C_q}$. *If* $X, Y, Z \subset \mathbb{F}_q^n$ *with less than* δq^{2n} *triangles in* $X \times Y \times Z$, *then we can remove* ϵq^n *elements from* $X \cup Y \cup Z$ *so that no triangle remains.*

This implies the following corollary.

Corollary 5.2.2 *Let $A \subset \mathbb{F}_q^n$ with $|A| = q^{n(1-s)}$, $0 \leq s < c_q$ and suppose that n is sufficiently large. Then A contains $\Omega_q(q^{n(2-sC_q)})$ non-trivial three term arithmetic progressions.*

Proof of Corollary 5.2.2. Applying the bound (1.7), we know that for some constant k, every subset of A with size greater than $kq^{n(1-c_q)}$ contains a three term arithmetic progression. Let $\epsilon = \frac{1}{2q^{ns}}$. It therefore follows that, for n sufficiently large,

$$|A| - \epsilon q^n = \frac{q^{n(1-s)}}{2} \geq kq^{n(1-c_q)}.$$

In particular, any subset of A of size $|A| - \epsilon q^n$ contains a non-trivial 3-AP. To put it another way, if we remove ϵq^n elements from A, the resulting set still contains a 3-AP.

Now we can apply Theorem 5.2.1 in its contrapositive form with $X = Y = A$ and $Z = -2A$, so that the property of being triangle free is the same as that of being 3-AP free. It follows that $A \times A \times (-2A)$ contains at least

$$\delta q^{2n} = \left(\frac{1}{6q^{ns}}\right)^{C_q} q^{2n} = k'(q)q^{n(2-sC_q)}$$

triangles. Some of these triangles may correspond to trivial arithmetic progression, but the number of such progressions in negligible and the proof is complete.

Supersaturation in the integers. A supersaturation lemma for three term arithmetic progressions in $[N]$ is already standard, in the form of Varnavides' Theorem. We will use the following formulation, which can be derived from Lemma 3.1 in [**CS09**].

Theorem 5.2.3 *Suppose that for all $N \in \mathbb{N}$ we have $r_3(N) \leq \frac{N}{h(N)}$ for some invertible function $h : \mathbb{R}^+ \to \mathbb{R}^+$. Then every subset $A \subset [N]$ with cardinality $|A| = \eta N$, such that*

$$1 \leq \left\lfloor h^{-1}\left(\frac{4}{\eta}\right)\right\rfloor \leq N,$$

must contain at least

$$\left(\frac{\eta}{2(h^{-1}(\frac{4}{\eta}))^4}\right)N^2,$$

non-trivial three term arithmetic progressions.

5.3 Improved bounds for the cap set problem in random subsets of $\mathbb{F}_q^n$

The proof of Theorem 5.0.2 begins by iteratively applying the container theorem to subsets of $\mathbb{F}_q^n$ in order to establish the existence of a convenient family of sets C which contain all 3-AP free subsets of $\mathbb{F}_q^n$. This results in the following container lemma.

Lemma 5.3.1 *For all $\beta > 0$ and for all $0 \leq t \leq c_q(1-3\beta)$ there exists a constant $c = c(q,\beta)$ such that there exists a family C of subsets of $\mathbb{F}_q^n$ with the following properties:*

- $|C| \leq 2^{n^2 c(q,\beta) q^{n\left(\frac{1}{2}+\beta+\frac{t(Cq-3)}{2}\right)}}$,
- *for all $A \in C$, $|A| \leq q^{n(1-t)}$,*
- *If $X \subset \mathbb{F}_q^n$ is 3-AP free then there exists $A \in C$ such that $X \subseteq A$.*

Proof of Lemma 5.3.1. At the outset, this problem is converted into a graph theoretic situation in order to setup an application of Theorem 5.1.1. Given $A \subset \mathbb{F}_q^n$, define a 3-uniform $H(A) = (V, E)$ hypergraph with vertex set $V = A$. Three distinct vertices form an edge in H if and only if they form a three term arithmetic progression.

The aim is to find a good set of containers for the hypergraph $H(\mathbb{F}_q^n)$. We will eventually obtain a family C of subsets of $\mathbb{F}_q^n$ such that

- $|C| \leq 2^{n^2 c(q,\beta) q^{n\left(\frac{1}{2}+\beta+\frac{t(Cq-3)}{2}\right)}}$,
- for all $A \in C$, $|A| \leq q^{n(1-t)}$,
- if X is an independent set in the hypergraph $H(\mathbb{F}_q^n)$, then there is some $A \in C$ such that $X \subseteq A$.

Once the existence of such a family C has been established, the proof of Lemma 5.3.1 will be complete.

We will iteratively apply the container theorem to subsets of $\mathbb{F}_q^n$. We begin by applying Theorem 5.1.1 to the graph $H(\mathbb{F}_q^n)$. As a result, we obtain a set C_1 of containers. We iterate by considering each $A \in C_1$. If A is not small enough, then we apply Theorem 5.1.1 to the graph $H(A)$ to get a family of containers C_A. If A

is sufficiently small then we put this A into a final set C of containers (or to put it another way, we write $C_A = A$).

Repeating this for all $A \in C_1$ we obtain a new set of containers

$$C_2 = \bigcup_{A \in C_1} C_A.$$

Note that C_2 is a container set for $H(\mathbb{F}_q^n)$. Indeed, suppose that X is an independent set in $H(\mathbb{F}_q^n)$. Then there is some $A \in C_1$ such that $X \subset A$. Also, X is an independent set in the hypergraph $H(A)$, which implies that $X \subset A'$ for some $A' \in C_A \subset C_2$.

We then repeat this process, defining

$$C_i = \bigcup_{A \in C_{i-1}} C_A.$$

By choosing the values of τ and ϵ appropriately, we can ensure that after relatively few steps we have all of the elements of C_k sufficiently small. We then declare $C = C_k$. It turns out that, because of k being reasonably small, $|C|$ is also fairly small.

Now we give more precise details of how to run this argument. Let $A \in C_j$, with $j \leq k$, and write $|A| = q^{n(1-s)}$. If $s \leq t$, then apply the container theorem to $H(A)$ with

$$\epsilon = q^{-\beta n}, \qquad \tau = q^{\frac{n}{2}(2\beta - 1 + s(C_q - 1))}.$$

In order to apply the container theorem, we need to check that the conditions $\tau < 1/(200 \cdot 3 \cdot 3!^2) = 1/21600$, and $\Delta(H, \tau) \leq \frac{\epsilon}{72}$ hold. The first of these conditions will hold if we take n sufficiently large. This follows from the condition that $s \leq t \leq c_q(1 - 3\beta)$.

For the second condition, we need to verify that

$$\frac{4\Delta_2}{d\tau} + \frac{2\Delta_3}{d\tau^2} \leq \frac{\epsilon}{72}. \tag{5.2}$$

Observe that, for any subset $A \subset \mathbb{F}_q^n$, $\Delta_2(H(A)) \leq 3$, since for any two distinct elements $a_1, a_2 \in A$, there are at most three possible choices of a third element $a_3 \in A$ such that $\{a_1, a_2, a_3\}$ forms an arithmetic progression. We also have $\Delta_3(H(A)) \leq 1$.

To bound the average vertex degree d, we use Theorem 5.2.2. The set A has cardinality $q^{n(1-s)}$, implying that it contains $\Omega_q(q^{n(2-sC_q)})$ non-trivial three-term arithmetic progressions. Therefore,

$$d = \frac{3|E(H(A))|}{|A|} \gg_q \frac{q^{n(2-sC_q)}}{q^{n(1-s)}} = q^{n(1-s(C_q-1))}.$$

Therefore, it follows that, for some constant c_0 depending on q,

$$\frac{4\Delta_2}{d\tau} + \frac{2\Delta_3}{d\tau^2} \leq \frac{12}{d\tau} + \frac{2}{d\tau^2} \leq \frac{14}{d\tau^2} < \frac{c_0}{q^{2\beta n}} \leq \frac{\epsilon}{72},$$

where the last inequality holds for all n sufficiently large. This verifies the condition (5.2), and so we can apply Theorem 5.1.1 and obtain a set of containers C_A with

$$|C_A| \leq 2^{cq^{n(1-s)}\tau \cdot \log(1/\epsilon) \cdot \log(1/\tau)} \leq 2^{c(n\log q)^2 q^{\frac{n}{2}(1+s(C_q-3)+2\beta)}}.$$

Since $s \leq t$, it follows that we have the bound

$$|C_A| \leq 2^{c(n\log q)^2 q^{\frac{n}{2}(1+t(C_q-3)+2\beta)}}.$$

We also know that, for each $B \in C_A$,

$$e(H(B)) \leq \epsilon e(H(A)) = q^{-\beta n} e(H(A)).$$

Therefore, at the ith level of this iterative procedure, a container $B \in C_i$ satisfies

$$e(H(B)) \leq q^{n(2-i\beta)}.$$

This is good, because after $c(\beta)$ steps we can ensure that $e(H(B))$ is sufficiently small so that we can apply Theorem 5.2.2 and deduce that $|B| \leq q^{n(1-t)}$. In particular, if we take

$$k := \left\lceil \frac{tC_q}{\beta} + 1 \right\rceil$$

then Theorem 5.1.1 tells us that for each $B \in C_k$, $|B| \leq q^{n(1-t)}$

So, the process terminates after at most k steps. This implies that the final set of containers $C = C_k$ has cardinality

$$|C| \leq 2^{c(n\log q)^2 k q^{\frac{n}{2}(1+t(C_q-3)+2\beta)}} = 2^{n^2 c(q,\beta) q^{\frac{n}{2}(1+t(C_q-3)+2\beta)}},$$

as claimed. This completes the proof.

The set of containers established in Lemma 5.3.1 can now be used to deduce Theorem 5.0.2, which we recall for the reader's convenience.

Theorem 5.0.2. Let $\beta > 0$, $t \leq c_q(1-3\beta)$ and let p be a positive real number satisfying

$$q^{n\left(-\frac{1}{2}+\frac{t(C_q-1)}{2}-\frac{\beta}{2}\right)} \leq p \leq 1.$$

Let B be a random subset of $\mathbb{F}_q^n$ with the events $x \in B$ being independent with probability $\mathbb{P}(x \in B) = p$. Then, with probability $1 - o_{n\to\infty}(1)$ we have that

$$f_3(B) \ll pq^{n(1-t+2\beta)}.$$

In particular, for all $\epsilon > 0$, there exists $\delta(\epsilon, q) := \delta > 0$ such that if B is defined as above with $p = q^{n(-\frac{1}{2}+\epsilon)}$, then with probability $1 - o_{n\to\infty}(1)$,

$$f_3(B) \ll |B|^{1-\delta}.$$

Proof of Theorem 5.0.2. For convenience, define $m = pq^{n(1-t+2\beta)}$, and let C be the container set guaranteed by Lemma 5.3.1. We first note that the probability that B contains a three-term progression-free subset of size at least m is upper bounded by

$$|C|\binom{q^{n(1-t)}}{m}p^m. \tag{5.3}$$

This is because a 3-AP free set of size m must be contained in some $A \in C$, and each subset of size m belongs to the random subset B with probability p^m. Every $A \in C$ has size

$$|A| \le q^{n(1-t)},$$

and so the number of possible candidates for a 3-AP free set of size m is at most

$$|C|\binom{q^{n(1-t)}}{m}.$$

An application of the union bound then gives (5.3). Using the bound

$$|C| \le 2^{n^2 c(q,\beta) q^{n\left(\frac{1}{2}+\frac{t(Cq-3)}{2}+\beta\right)}},$$

and the standard binomial coefficient estimate $\binom{s}{t} \le \left(\frac{es}{t}\right)^t$ gives

$$\begin{aligned} |C|\binom{q^{n(1-t)}}{m}p^m &\le 2^{n^2 c(q,\beta) q^{n\left(\frac{1}{2}+\frac{t(Cq-3)}{2}+\beta\right)}} \left(\frac{epq^{n(1-t)}}{m}\right)^m \\ &= 2^{n^2 c(q,\beta) q^{n\left(\frac{1}{2}+\frac{t(Cq-3)}{2}+\beta\right)}} \left(\frac{e}{q^{2\beta n}}\right)^m \\ &\le \left(\frac{2e}{q^{2\beta n}}\right)^m. \end{aligned} \tag{5.4}$$

In the last inequality above, we have used the fact that for n sufficiently large,

$$m = pq^{n(1-t+2\beta)} \ge q^{n\left(\frac{1}{2}+\frac{t(Cq-3)}{2}+\frac{3}{2}\beta\right)} \ge n^2 c(q,\beta) q^{n\left(\frac{1}{2}+\frac{t(Cq-3)}{2}+\beta\right)}.$$

The lower bound on p in the statement of the theorem was used here. The quantity in (5.4) tends to zero as n goes to infinity, which completes the proof of the first part of the statement.

The second statement follows from the first by taking

$$t = \frac{2\epsilon}{C_q - 1}, \quad \beta = t/4.$$

Indeed, for suitably chosen constants $c, C > 0$, the statement

$$cpq^n \leq |B| \leq Cpq^n = Cq^{n(\frac{1}{2}+\epsilon)}$$

is true with probability $1 - o_{n\to\infty}(1)$. Therefore, with probability $1 - o_{n\to\infty}(1)$, we have

$$f_3(B) \ll pq^{n(1-\frac{t}{2})} \ll |B|q^{n(-\frac{t}{2})} \ll_\epsilon |B|^{1-\delta(\epsilon)}.$$

We finally use Theorem 5.0.2 to deduce Theorem 5.0.1.

Proof of Theorem 5.0.1. Construct a subset $P \subset \mathbb{F}_q^n$ by choosing elements independently at random with probability $p = \frac{1}{100}q^{-n/3}$. The expected number of elements in P is $pq^n = \frac{1}{100}q^{2n/3}$, while the expected number of nontrivial four-term progressions is at most $p^4q^{2n} = 10^{-8}q^{2n/3}$. Indeed, the latter follows from the fact that $\mathbb{F}_q^n$ contains less than q^{2n} non-trivial 4-APs and each one survives the random process with probability p^4. In particular, the expected number of elements of P is considerably larger than the expected number of 4-APs. Therefore, with high probability both

$$|P| \geq \frac{1}{1000}q^{2n/3}$$

and

$$|\{\text{all non-trivial 4-APs in } P\}| \leq \frac{1}{2000}q^{2n/3}$$

hold. We can then delete one element from each 4-AP and obtain a set P' with size $\Omega(q^{2n/3})$ which has no nontrivial four-term progressions.

On the other hand, we can apply Theorem 5.0.2 with $t = \frac{1}{3(C_q-1)}$ and the above choice of p, as these values satisfy the required conditions provided that n is sufficiently large. Therefore, with probability tending to 1 as n goes to infinity, the randomly constructed set P satisfies

$$f_3(P) \leq pq^{n\left(1-\frac{1}{3(C_q-1)}+2\beta\right)} \ll q^{n\left(\frac{2}{3}-\frac{1}{3(C_q-2)}+2\beta\right)}.$$

Now, for every positive integer m, P' contains a three-term progression-free set of size m only if P also does. That is, $f_3(P') \leq f_3(P)$. Therefore,

$$f_3(P') \leq f_3(P) \ll q^{n\left(\frac{2}{3}-\frac{1}{3(Cq-2)}+2\beta\right)} \ll |P'|^{1-\frac{1}{2(Cq-2)}+3\beta}.$$

This completes the proof.

5.4 An analogous story over integers

We will prove the following more general result which involves the parameter $r_3(N)$.

Proposition 5.4.1 *Suppose that for all sufficiently large $N \in \mathbb{N}$ we have $r_3(N) \leq \frac{N}{h(N)}$ for some monotone increasing and invertible function $h : [1,\infty) \to [1,\infty)$. Suppose also that h satisfies the following technical conditions:*

- *For all $x \in [1,\infty)$, $h(x) \leq x$.*
- *There exists an absolute constant γ such that for all N sufficiently large*

$$h\left(\frac{N^{1/5}}{1000}\right) \geq 4h(N^{\gamma}) \tag{5.5}$$

$$N^{1/10} \geq [h(N^{\gamma})]^{3/2}[h^{-1}(4h(N^{\gamma}))]^2. \tag{5.6}$$

Then for all $\alpha > 0$ and for all n sufficiently large (depending on α), there exists a four-term progression-free set $A \subset \mathbb{N}$ with cardinality n such that

$$f_3(A) \ll \frac{n}{[h(n^{\frac{3}{2}\gamma})]^{1-\alpha}}.$$

Note that the rather complicated looking statement of Proposition 5.4.1 does imply the upper bound from Theorem 5.0.3. Indeed, because of Bloom's upper bound on $r_3(N)$ in (1.5), we can carelessly bound $r_3(N)$ by

$$r_3(N) \leq C\frac{N}{(\log N)^{1-\alpha}}$$

for any $\alpha > 0$ and for some positive absolute constant C. We can then apply Proposition 5.4.1 with $h(x) = \frac{1}{C}(\log x)^{1-\alpha}$. It is a tedious calculation to check that h does indeed satisfy the conditions of Theorem 5.4.1, with room to spare, if we take $\gamma = \frac{1}{24\cdot 4^{\frac{1}{1-\alpha}}}$, which gives the required bound.

Proof of Proposition 5.4.1. The proof is similar to that of Theorem 5.0.1, although the calculations are more taxing. On the other hand, this proof is a little more straightforward, since we make just a single application of the container theorem. We remark that this approach with a single application was also possible in the proof of Theorem 5.0.1, but the iterative approach gave a better quantitative result. However, the quantitative gains of the iterative approach seem to be negligible in the integer case.

Once again, we define a 3-uniform hypergraph which encodes three term arithmetic progressions. This hypergraph H has vertex set $[N]$, and three distinct elements of $[N]$ form an edge if they form an arithmetic progression.

Note that the average degree d of this hypergraph is at least $N/9$, since there are at least $N^2/9$ edges. Indeed, if we take any two distinct integers $a, b \in [1, N/2]$ with $a < b$, there exists a third integer $c = 2b - a \in [1, N]$ such that $\{a, b, c\}$ forms an arithmetic progression. This shows the existence of at least

$$\binom{\lfloor \frac{N}{2} \rfloor}{2} > \frac{N^2}{9},$$

non-trivial 3-APs, where the latter inequality holds provided that N is sufficiently large. Also, as in the proof of Theorem 5.0.1, we have $\Delta_2 \leq 3$ and $\Delta_3 = 1$.

Fix

$$\eta := \frac{1}{h(N^\gamma)},$$

where γ is the constant in the statement of Proposition 5.4.1. Define

$$\epsilon := \frac{\eta}{\left(h^{-1}(\frac{4}{\eta})\right)^4}, \qquad \tau := \frac{100}{(N\epsilon)^{1/2}}.$$

We would like to apply the Theorem 5.1.1 with these parameters. In order to do this, we need to check that the conditions

$$\tau < 1/(200 \cdot 3 \cdot 3!^2) = 1/21600 \tag{5.7}$$

and

$$\Delta(H, \tau) \leq \frac{\epsilon}{72} \tag{5.8}$$

hold.

For (5.7) to hold, it would be enough to verify that

$$N\epsilon \geq 10^{12}. \tag{5.9}$$

That is,

$$\frac{\eta}{\left(h^{-1}(\frac{4}{\eta})\right)^4} \geq \frac{10^{12}}{N}. \tag{5.10}$$

Because of the assumption that $h(x) \leq x$ for all $x \in \mathbb{R}^+$, it follows that that $h^{-1}(x) \geq x$ and in particular

$$\frac{1}{x} \geq \frac{1}{h^{-1}(x)}. \tag{5.11}$$

Applying (5.11) with $x = \frac{4}{\eta}$, it follows that

$$\frac{\eta}{\left(h^{-1}(\frac{4}{\eta})\right)^4} = 4\frac{\frac{\eta}{4}}{\left(h^{-1}(\frac{4}{\eta})\right)^4} \geq 4\frac{1}{\left(h^{-1}(\frac{4}{\eta})\right)^5},$$

so that (5.10) would hold as long as

$$\frac{1}{\left(h^{-1}(\frac{4}{\eta})\right)^5} \geq \frac{10^{12}}{N}.$$

Since h is monotone increasing, this can be rearranged to give

$$\eta \geq \frac{4}{h\left(\frac{N^{1/5}}{10^{12/5}}\right)}.$$

The latter inequality holds for our choice of η. Here we have used the condition (5.5) in the statement of the theorem. This implies that (5.10) holds, and therefore so does (5.7).

For (5.8) to hold, we need to verify that

$$\frac{4 \cdot 9 \cdot 3}{N\tau} + \frac{2 \cdot 9}{N\tau^2} \leq \frac{\epsilon}{72}. \tag{5.12}$$

Since $\tau < 1$, it will be sufficient to check that $\frac{126}{N\tau^2} \leq \frac{\epsilon}{72}$. By the earlier choice of τ, this is equivalent to $(100)^2 \geq 72 \cdot 126$, which is indeed true.

Theorem 5.1.1 then gives a collection C of subsets of $[N]$ such that

- $|C| \leq 2^{c\tau N \log(\frac{1}{\tau}) \log(\frac{1}{\epsilon})}$,
- for all $A \in C$, $e(H[A]) \leq \epsilon e(H)$,
- if $X \subseteq [N]$ is an independent set in H, then there is some $A \in C$ such that $X \subseteq A$.

It follows from the second fact above and Theorem 5.2.3 that $|A| \leq \eta N$ for all $A \in C$. Note here that the condition of Theorem 5.2.3 follows from condition (5.6).

Observe that, for N sufficiently large

$$\frac{1}{\tau}, \frac{1}{\epsilon} \leq N.$$

The first of these inequalities follows from the fact that $\epsilon < \frac{1}{2}$, while the second is a consequence of (5.9). Using these two inequalities and the definition of τ, gives the bound

$$|C| \leq 2^{c'(\log N)^2 (\frac{N}{\epsilon})^{1/2}}. \tag{5.13}$$

Construct a subset $P \subset [N]$ by choosing elements independently at random with probability p. The expected number of elements in P is pN. The expected number of four-term arithmetic progressions is at most $p^4 N^2$. Therefore, if we choose $p = \frac{1}{100} N^{-1/3}$ then with high probability the number of elements will be much larger than the number of four-term arithmetic progressions. We can then delete one element from each 4-AP and obtain a set P' with size $\Theta(N^{2/3})$ which has no 4-APs. Just as was the case in the proof of Theorem 5.0.1, note here that $f_3(P') \leq f_3(P)$.

Now, we claim that it is unlikely that $H(P)$ contains an independent set of cardinality $m = N^{2/3} \eta^{1-\alpha}$. Indeed, note that

$$\mathbb{P}[H(P') \text{ contains an independent set of size } m] \leq |C| \binom{N\eta}{m} p^m,$$

whereas, by using the bound on $|C|$ from (5.13) together with standard binomial coefficient estimates, we also have that

$$|C| \binom{N\eta}{m} p^m \leq 2^{c'(\log N)^2 (\frac{N}{\epsilon})^{1/2}} \left(\frac{eN\eta}{mN^{1/3}} \right)^m = 2^{c'(\log N)^2 (\frac{N}{\epsilon})^{1/2}} (e\eta^{\alpha})^m.$$

With the choices we have made for η and m, it follows that the bound

$$c'(\log N)^2 \left(\frac{N}{\epsilon} \right)^{1/2} \leq m$$

holds for N sufficiently large. At this is the point, we have used the technical condition (5.6) in the statement of Proposition 5.4.1. Therefore, the probability that $H(P')$ contains an independent set of size m is less than $(2e\eta^{\alpha})^m$, which becomes arbitrarily small as N gets arbitrarily large.

It follows that there exists a 4-AP free set P' of size $\Theta(N^{2/3})$ with the property that all of its subsets of size at least $N^{2/3}\eta^{1-\alpha}$ contain a 3-AP. That is,

$$f(|P'|) \ll |P'|\eta^{1-\alpha} \approx \frac{|P'|}{\left(h(|P'|^{\frac{3}{2}\gamma})\right)^{1-\alpha}}.$$

This completes the proof of Proposition 5.4.1.

5.5 Sets with small energy but rich in progressions

In this section, we discuss another application of Theorem 5.0.2, in connection with a different type of generalization of Roth's theorem, first observed by Sanders [**San09b**].

Theorem 5.5.1 *Let $\delta > 0$ and suppose that $A \subset \mathbb{Z}$ has at least $\delta|A|^3$ additive quadruples. Then, there exist absolute constants $c, C > 0$ such that A contains at least $\exp(-C\delta^{-c}) \cdot |A|^2$ three-term arithmetic progressions.*

Here an additive quadruple means a solution to $a + b = c + d$ with all a, b, c, d in A. The number of such quadruples is usually denoted by $E(A)$ and called the *additive energy* of A. Theorem 5.5.1 says that sets with large energy have many three-term arithmetic progressions. This follows from the Balog-Szemerédi-Gowers theorem (see [**Gow98**] or [**TV06**]) and the fact that sets with small sumsets have many three-term arithmetic progressions, a consequence of Roth's theorem. Results like the latter hold in general abelian groups G and quantitative versions were also studied by Henriot in [**Hen16**]. For our purposes, the groups of interest are $G = \mathbb{Z}$ and $G = \mathbb{F}_q^n$, so we begin by recording an improvement (and generalisation) of a theorem of Henriot [**Hen16**], which may be of independent interest, and which is meant to illustrate a phenomenon similar to the one described by Theorem 5.5.1 (with better quantitative bounds).

Theorem 5.5.2 *Let $A \subset \mathbb{F}_q^n$ be such that $|A + A| \leq K|A|$ for some $K > 0$. Then, A contains at least $(qK^4)^{2-C_q} \cdot |A|^2$ three-term arithmetic progressions.*

Proof of Theorem 5.5.1. For the reader's convenience, we recall that for any two commutative groups G_1, G_2 two sets $S \subset G_1$ and $T \subset G_2$ are said to be Freiman

s-isomorphic if there exists a one to one map $\phi : S \to T$ such that for every $x_1, \ldots, x_s, y_1, \ldots, y_s$ in S (not necessarily distinct) the equation

$$x_1 + \ldots + x_s = y_1 + \ldots + y_s$$

holds if and only if

$$\phi(x_1) + \ldots + \phi(x_s) = \phi(y_1) + \ldots + \phi(y_s).$$

Let $K = |A + A|/|A|$. By a finite field version of the so-called Freiman-Ruzsa modelling lemma (see for instance [**Sis18**] for more details), A is Freiman 2-isomorphic to a subset of $G = \mathbb{F}_q^m$, where $|G| \leq q \cdot K^4|A|$. We identify this subset with A since the Freiman 2-isomorphisms preserves three-term progressions. By Corollary 5.2.2 applied inside G, it follows that A contains at least $|A|^2(qK^4)^{2-C_q}$ three-term arithmetic progressions, as claimed.

Theorem 5.5.2, combined with the Balog-Szemerédi-Gowers theorem, shows that subsets $A \subset \mathbb{F}_q^n$ must have many three-term progressions even if $E(A) \gg |A|^{3-\epsilon}$ for some $\epsilon > 0$ (which depends on q). A natural question now seems to be: if A has large additive energy, does it also mean that A must have nontrivial three-term progressions in all large subsets? A naive view is that Theorem 5.5.1 and Theorem 5.5.2 suggest that the answer could be yes. However, a simple counterexample already points towards the contrary: consider a set of A where half of the elements form an additively structured set (like an arithmetic progression), while the other half consists of random elements. It is easy to check that $E(A) \gg |A|^3$ because the additively structured part has large energy, while there is no reason why the random part should contain any non-trivial three-term progressions.

We will push this observation one step further and show next that for sets in $\mathbb{F}_q^n$ or $\mathbb{Z}$ the property of "having nontrivial three-term progressions in all large subsets" is in fact *entirely uncorrelated* with the property of "having large additive energy".

Theorem 5.5.3 *For all $\epsilon > 0$ and any prime power q there exists $\delta(\epsilon, q) := \delta > 0$ and $n_0 = n_0(\epsilon, q)$ such the following statement holds. For all $n \geq n_0$ there exists a set $A \subset \mathbb{F}_q^n$ with*

$$E(A) \leq |A|^{2+\epsilon}$$

and

$$f_3(A) \ll |A|^{1-\delta}.$$

In other words, not only that sets with large additive energy may have large subsets with no proper three-term progressions, but there also exist sets with low energy with the property that all their large subsets contain nontrivial three-term progressions. The proof uses again Theorem 5.0.2 and is similar to the proof of Theorem 5.0.1.

Proof of Theorem 5.5.3. Construct a subset $P \subset \mathbb{F}_q^n$ by choosing elements independently at random with probability $p = q^{n\left(-\frac{1}{2}+\frac{\epsilon}{4-2\epsilon}\right)}$. The expected number of elements in P is $pq^n = q^{n\left(\frac{1}{2}+\frac{\epsilon}{4-2\epsilon}\right)}$.

The expected size of $E(P)$ is $p^4q^{3n} = q^{n\left(1+\frac{4\epsilon}{4-2\epsilon}\right)}$. Indeed, this follows from the fact that there are q^{3n} solutions to the equation

$$a + b = c + d, \qquad a, b, c, d \in \mathbb{F}_q^n$$

and each solution survives the random process with probability p^4. Therefore, with high probability both

$$|P| \geq \frac{1}{100} q^{n\left(\frac{1}{2}+\frac{\epsilon}{4-2\epsilon}\right)}$$

and

$$E(P) \leq 100 q^{n\left(1+\frac{4\epsilon}{4-2\epsilon}\right)}$$

hold. In particular, with high probability,

$$E(P) \ll |P|^{2+\epsilon}.$$

On the other hand, we can apply Theorem 5.0.2 with

$$t = \frac{2\epsilon}{(4-2\epsilon)(C_q-1)}, \qquad \beta = \frac{t}{4}$$

The above choice of p is admissible for these choices of t and β. Therefore, with probability tending to 1 as n goes to infinity, the randomly constructed set P satisfies

$$f_3(P) \ll pq^{n(1-t+2\beta)} = pq^{n(1-\frac{t}{2})} = q^{n(\frac{1}{2}-\frac{\epsilon}{(4-2\epsilon)(C_q-1)})} \ll |P|^{1-\delta},$$

where

$$\delta = \frac{\epsilon C_q}{2(C_q-1)}.$$

This completes the proof.

A similar statement can be established in the integer case, which we state without proof as follows.

Theorem 5.5.4 *For all $\alpha, \epsilon > 0$ there exists a set $A \subset \mathbb{N}$ such that*

$$E(A) \ll |A|^{2+\epsilon}$$

and

$$f_3(A) \ll_\epsilon \frac{|A|}{(\log |A|)^{1-\alpha}}.$$

We end this section with a discussion regarding the optimality of Theorem 5.5.4. For this purpose, we recall a theorem of Komlós, Sulyok and Szemerédi [**KSS75**].

Theorem 5.5.5 *There is an absolute constant $c > 0$ such that for any sufficiently large set $A \subset \mathbb{Z}$,*

$$f_3(A) \geq c \cdot f_3(\{1, \ldots, |A|\}) = c \cdot r_3(|A|).$$

Essentially, Theorem 5.5.5 tells us that $f_3(A)$ is minimal as a function of $|A|$ when A is an interval.[1] Combining this with Elkin's theorem

$$r_3(N) \gg \frac{\log^{1/4} N}{2^{2\sqrt{2}\sqrt{\log N}}} \cdot N,$$

it follows that *every* sufficiently large set $A \subset \mathbb{Z}$ contains a three-term progression free subset of cardinality at least

$$\Omega\left(\frac{\log^{1/4} |A|}{2^{2\sqrt{2}\sqrt{\log |A|}}} \cdot |A|\right). \tag{5.14}$$

So Theorem 5.5.4 is as close to optimal as the upper bound for $r_3(N)$ in (1.5) is close to optimal. Note however that in this observation we have not used the additional hypothesis that A has low additive energy. The next natural question therefore seems to be: is it possible to get a significantly better bound than

$$f_3(A) \gg \frac{\log^{1/4} |A|}{2^{2\sqrt{2}\sqrt{\log |A|}}} \cdot |A| \tag{5.15}$$

for *all* sets $A \subset \mathbb{Z}$ satisfying $E(A) \ll |A|^{2+\epsilon}$ for some (or even all) $0 < \epsilon < 1$? This time, the answer turns out to be (a modest) *yes*.

[1] In fact, [**KSS75**] gives much more general information about systems of linear equations, but the version stated as Theorem 5.5.5 corresponds to the case we are interested in.

Theorem 5.5.6 *Let $0 < \epsilon < 1$ and let $A \subset \mathbb{Z}$ be such that $E(A) \ll |A|^{2+\epsilon}$. Then,*

$$f_3(A) \gg \frac{\log^{1/4} |A|}{2^{2\sqrt{(1+\epsilon)\log N}}} \cdot N.$$

In particular, *all* sets with $E(A) \ll |A|^{2+\epsilon}$ for all $\epsilon > 0$ have slightly larger 3AP-free sets than we know $\{1, \ldots, N\}$ must have. Our argument follows closely the alternative proof of Elkin's bound due to Green and Wolf from [**GW10**], which can be easily modified to start with a general set of N integers instead of the interval $\{1, \ldots, N\}$. The main observation is that for a set A with $E(A) \ll |A|^{2+\epsilon}$ for some $0 < \epsilon < 1$, we have a power saving on the total number $T(A)$ of three-term progressions with elements in A. Indeed, for each element $s \in A + A$, let $r_{A+A}(s)$ denote the number of pairs $(x, y) \in A \times A$ such that $x + y = s$. For each $b \in A$, note that $r_{A+A}(2b)$ represents the number of three-term progressions centered at b. By Cauchy-Schwarz,

$$T(A)^2 = \left(\sum_{b \in A} r_{A+A}(2b) \right)^2 \leq |A| \left(\sum_{b \in A} r_{A+A}^2(2b) \right).$$

Since

$$\sum_{b \in A} r_{A+A}^2(2b) \leq \sum_{s \in A+A} r_{A+A}^2(s) = E(A),$$

it follows that $T(A)^2 \leq |A| \cdot E(A) \ll |A|^{3+\epsilon}$, i.e. $T(A) \ll |A|^{(3+\epsilon)/2}$. Theorem 5.5.6 will then follow from the following more general result.

Proposition 5.5.7 *Let $A \subset \mathbb{Z}$ be a set of size N such that the number of three-term progressions satisfies $T(A) = N^2/t(A)$. Then A contains a three-term progression free subset A' such that*

$$|A'| \gg N \cdot \frac{\left[\log \left(\frac{N}{t(A)} \right) \right]^{1/4}}{2^{2\sqrt{2 \log_2 \left(\frac{N}{t(A)} \right)}}}.$$

Proof. Let N be a sufficiently large positive integer and let A be some four-term progression free set of size N. Let d be a positive integer to be precisely determined later (but which we shall think of as sufficiently large for the time being), and let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ denote the d-dimensional torus. For each $\theta, \alpha \in \mathbb{T}^d$, let $\Psi_{\theta,\alpha} : A \to \mathbb{T}^d$ be the map defined by

$$n \mapsto \theta n + \alpha \mod 1. \tag{5.16}$$

For a fixed n integer, as we let θ, α vary uniformly and independently over $\mathbb{T}^d$, the image $\Psi_{\theta,\alpha}$ is uniformly distributed on the d-dimensional torus. Moreover, it is also true that the pair of points

$$(\Psi_{\theta,\alpha}(n), \Psi_{\theta,\alpha}(n')) \text{ is uniformly distributed on } \mathbb{T}^d \times \mathbb{T}^d \tag{5.17}$$

as θ, α vary uniformly and independently over $\mathbb{T}^d$, provided that integers n and n' are distinct. Indeed,

$$\int e^{2\pi i(k\cdot(\theta n+\alpha)+k'\cdot(\theta n'+\alpha))} d\theta d\alpha = 0$$

unless $k + k' = kn + k'n' = 0$, which is however impossible if n and n' are distinct. Since the exponentials $e^{2\pi i(kx+k'x')}$ are dense in $L^2(\mathbb{T}^d \times \mathbb{T}^d)$, the claim checks out.

Fix δ to be a positive constant which we will declare later. We identify the d-dimensional torus $\mathbb{T}^d$ with $[0, 1)^d$, and for each $r \leq \frac{1}{2}\sqrt{d}$, we define the annulus

$$S(r) := \left\{ x \in [0, 1/2]^d : r - \delta \leq \|x\|_2 \leq r \right\}.$$

Like in Lemma 2.2 from [**GW**], out of all of the possible values of r, we choose the one for which $S := S(r)$ satisfies

$$\text{vol}(S(r)) \geq c\delta 2^{-d}, \tag{5.18}$$

for some absolute constant c.

Finally, for each θ, α chosen uniformly and independently at random on $\mathbb{T}^d$, we let $A_{\theta,\alpha}$ be the subset of A defined by

$$A_{\theta,\alpha} := \{n \in A : \Psi_{\theta,\alpha}(n) \in S\},$$

where $\Psi_{\theta,\alpha}$ is the map from (5.16). By (5.17), the expected size of $A_{\theta,\alpha}$ satisfies

$$\mathbb{E}_{\theta,\alpha}|A_{\theta,\alpha}| = N \cdot \text{vol}(S), \tag{5.19}$$

while the expected number $T(A_{\theta,\alpha})$ of three term progressions in $A_{\theta,\alpha}$ is

$$\mathbb{E}_{\theta,\alpha}T(A_{\theta,\alpha}) = T(A) \cdot \text{vol}(\Upsilon). \tag{5.20}$$

Here Υ represents the set points $(x, y) \in \mathbb{T}^d \times \mathbb{T}^d$ so that $x - y$, x and $x + y$ all lie in S.

We can upper bound the volume of Υ as follows. By the parallelogram law

$$2\|x\|^2 + 2\|y\|_2^2 = \|x + y\|_2^2 + \|x - y\|_2^2,$$

so

$$\|y\|_2 \leq \sqrt{r^2 - (r-\delta)^2} \leq \sqrt{2\delta r}.$$

If V_d denotes the volume of the unit ball in $\mathbb{R}^d$, then this implies

$$\mathrm{vol}(\Upsilon) \leq \mathrm{vol}(S) \cdot (\sqrt{2\delta r})^d V_d.$$

On the other hand, we have the estimate

$$V_d \ll 10^d d^{-d/2};$$

therefore

$$\mathrm{vol}(\Upsilon) \leq \mathrm{vol}(S) \cdot (\sqrt{2\delta r})^d 10^d d^{-d/2} \leq \mathrm{vol}(S) \cdot 10^d \left(\frac{\delta}{\sqrt{d}}\right)^{d/2}.$$

By (5.20), this estimate implies

$$\mathbb{E}_{\theta,\alpha} T(A_{\theta,\alpha}) = T(A) \cdot \mathrm{vol}(\Upsilon) \leq C \frac{N^2}{t(A)} \cdot \mathrm{vol}(S) \cdot 10^d \left(\frac{\delta}{\sqrt{d}}\right)^{d/2},$$

for an absolute constant $C > 0$. Now, if we choose δ and d so that

$$10^d \left(\frac{\delta}{\sqrt{d}}\right)^{d/2} \leq \frac{1}{3C} \cdot \frac{t(A)}{N} \tag{5.21}$$

then by (5.19)

$$\mathbb{E}_{\theta,\alpha} T(A_{\theta,\alpha}) = T(A) \cdot \mathrm{vol}(\Upsilon) \leq \frac{1}{3} \cdot N \cdot \mathrm{vol}(S) = \frac{1}{3} \cdot \mathbb{E}_{\theta,\alpha} |A_{\theta,\alpha}|.$$

Consequently, by deleting one element from each progression appearing in $A_{\theta,\alpha}$, the remaining subset $A'_{\theta,\alpha} \subset A_{\theta,\alpha} \subset A$ is three-term progression-free. Moreover, $A'_{\theta,\alpha}$ has expected size

$$\mathbb{E}_{\theta,\alpha} |A'_{\theta,\alpha}| \geq \frac{2}{3} \cdot \mathbb{E}_{\theta,\alpha} |A_{\theta,\alpha}| \geq \frac{2}{3} \cdot N \cdot \mathrm{vol}(S) \gg N\delta 2^{-d},$$

where the last inequality follows from (5.18). In particular, there exists a specific choice of $\theta, \alpha \in \mathbb{T}^d$ so that $A' := A'_{\theta,\alpha}$ is a three-term progression free subset of A for which

$$|A'| \gg N\delta 2^{-d}.$$

Finally, take

$$\delta := C' \sqrt{d} \cdot \left(\frac{t(A)}{N}\right)^{2/d}$$

for some absolute constant $C' > 0$ so that (5.21) is achieved. For this choice, we have

$$|A'| \gg N\delta 2^{-d} \gg \sqrt{d} \cdot t(A)^{2/d} N^{1-2/d} \cdot 2^{-d}.$$

Set

$$d = \left\lceil \sqrt{2 \log_2 \left(\frac{N}{t(A)} \right)} \right\rceil.$$

It then follows that

$$|A'| \gg N \cdot \frac{\left[\log_2 \left(\frac{N}{t(A)} \right) \right]^{1/4}}{2^{2\sqrt{2 \log_2 \left(\frac{N}{t(A)} \right)}}}.$$

This concludes the proof of Proposition 5.5.7 and thus that of Theorem 5.5.6 (one can check that taking $t(A) = \Theta(N^{(1-\epsilon)/2})$ in Proposition 5.5.7 yields the bound from Theorem 5.5.6, as claimed).

Chapter 6

HOW MANY 3APS CAN A 4AP-FREE SET ACTUALLY HAVE?

Let $k \geq 3$ be an integer. In this chapter, a k-term arithmetic progression of integers will denote as usual a set of the form $\{x, x+d, \ldots, x+(k-1)d\}$. If $d \neq 0$, then we say that the progression is *non-trivial*. If a set A does not contain any non-trivial k-term arithmetic progressions, we say that A is *k-AP free*. The study of k-AP free sets in the integers and other groups has been a central topic in additive combinatorics. Following the standard notation, we will denote by $r_k(n)$ the size of the largest k-AP free subset of $\{1, \ldots, n\}$. The seminal result on this topic is Szemerédi's Theorem [**Sze75**], which states that sets of integers with positive density contain arbitrarily long arithmetic progressions, or using the notation above $r_k(n) = o(n)$.

Since Szemerédi, the problem of finding better quantitative bounds for $r_k(n)$ has received a lot of attention, with impressive progress that led to many important tools, which in the meantime have become standard. For our application, we won't need the best bounds for each k, so we will limit ourselves to only mentioning Gowers' theorem [**Gow98**, **Gow01**] that for each $k \geq 3$ there exists an absolute constant $c_k > 0$ such that

$$r_k(n) \ll \frac{n}{(\log\log n)^{c_k}}. \tag{6.1}$$

Regarding lower bounds, Rankin [**Ran60**] showed that there exists a constant $c'_k > 0$ such that

$$r_k(n) \gg \frac{n}{2^{c'_k(\log n)^{1/\lceil \log k \rceil}}}. \tag{6.2}$$

Throughout this chapter, all logarithms are base 2 and the signs $\ll$ and $\gg$ are the usual Vinogradov symbols.

Let $\mathcal{A}_k(n)$ be the set of n-term nonnegative integer sequences which contain no k-term arithmetic progression as a subsequence. Furthermore, let $f_s(A)$ denote the number of s-term arithmetic progressions in A, and finally let $f_{s,k}(n) = \max_{A \in \mathcal{A}_k(n)} f_s(A)$. In [**E73**, **page 119**], Erdős observed that

$$\frac{\log f_{3,4}(n)}{\log n} > 1.4649$$

holds for infinitely many n by constructing examples of sequences $A \in \mathcal{A}_4(3^s)$ for which $f(A) = 3^{s-1}$. Furthermore, he noticed that for each $k > 3$ the limit

$\lim_{n\to\infty} \log f_{3,k}(n)/\log n := f_{3,k}$ exists, and asked whether or not $f_{3,k}$ is always less than 2. In [**SA77**], Simmons and Abbott improved on Erdős' observation by showing that $f_{3,4}(n) \geq n^{1.623}$ holds infinitely often, and also proved that $s_{3,k} \to 2$ as k goes to infinity. Nonetheless, in the regime when k is fixed, there has been no further progress on understanding the limit $f_{3,k}$ as far as we are aware of. In this note, we settle Erdős' question in the negative by proving the following more general result.

Theorem 6.0.1 *For all integers $k > s \geq 3$, we have*

$$\lim_{n\to\infty} \frac{\log f_{s,k}(n)}{\log n} = 2.$$

In fact, we prove upper and lower bounds for $f_{s,k}(n)$ which show that its growth is closely related to the bounds in Szemerédi's theorem.

Theorem 6.0.2 *There exist absolute positive constants c and C such that, for integers $k > s \geq 3$ and every sufficiently large integer n, we have*

$$\left(\frac{c \cdot r_k(n)}{n}\right)^{2(s-2)} \cdot n^2 \leq f_{s,k}(n) \leq \left(\frac{r_k(n)}{n}\right)^{C} \cdot n^2.$$

In light of the bounds on $r_k(n)/n$ provided by (6.1) and (6.2), it is easy to check that Theorem 6.0.1 follows from Theorem 6.0.2; therefore, it suffices to prove the latter. We will do this already in Section 2. The proof of Theorem 6.0.2 will require a few ingredients from additive combinatorics, but we will state them in full as we will get to apply them, as they do not require much preparation.

6.1 Proof of Erdos' conjecture

We first prove the desired upper bound on $f_{s,k}(n)$. For $s \geq 3$, we have $f_{s,k}(n) \leq f_{3,k}(n)$, so in order to prove the upper bound it suffices to show that

$$f_{3,k}(n) \leq \left(\frac{r_k(n)}{n}\right)^{C} n^2$$

holds for some absolute constant $C > 0$ and sufficiently large n. We will in fact show this claim for $C = 1/25$. Let $A \in \mathcal{A}_k(n)$ and let pn^2 denote the number

of three-term arithmetic progressions in A, where p is some positive real number (which is strictly less than 1); i.e. $f_3(A) = pn^2$.

To upper bound p, we will require the following variant of the Balog-Szemerédi-Gowers theorem (see [**Go98**] or [**FS11**]).

Theorem 6.1.1 *If A and B are sets of n integers and G is a bipartite graph between A and B with pn^2 edges such that partial sumset $A +_G B$ has size at most $K|A|$, then there is a subset A' of A with $|A'| \geq pn/4$ and*

$$|A' - A'| \ll K^4 p^{-5} n.$$

Here $A +_G B$ denotes as usual the sumset restricted to the edges coming from G, namely

$$A +_G B = \{a + b : \ a \in A, b \in B, (a, b) \in E(G)\}.$$

It is perhaps important to mention that Theorem 6.1.1 is a somewhat nonstandard version of the Balog-Szemerédi-Gowers theorem, which outputs directly a large set $A' \subset A$ with small difference set, without applying any Ruzsa-type inequality. We refer the reader to the proof of [**FS11**, **Lemma 5.2**, **page 9**], from which the following statement can also be extracted.

Lemma 6.1.2 *If a bipartite graph $G = (A, B, E)$ with $|A| = |B| = n$ has pn^2 edges, then there is a subset A' of A of size at least $pn/4$ such that every pair of vertices in A have at least $\Omega(p^5 n^3)$ paths of length four connecting them.*

Using Lemma 6.1.2, one can then deduce Theorem 6.1.1 in the usual way. Applied to the graph from the setup of Theorem 6.1.1, Lemma 6.1.2 produces $A' \subset A$ of size at least $pn/4$ such that every pair of vertices in A have at least $\Omega(p^5 n^3)$ paths of length four connecting them. This set happens to also satisfy $|A' - A'| \ll K^4 p^{-5} n$. Indeed, for each $a, a' \in A'$, consider a path of length four in G between them, say (a, b, a'', b', a'). For $y := a - a' \in A' - A'$, we can then write

$$a - a' = (a + b) - (a'' + b) + (a'' + b') - (a' + b') = x_1 - x_2 + x_3 - x_4,$$

where $x_1 = a + b$, $x_2 = a'' + b$, $x_3 = a'' + b'$, and $x_4 = a' + b'$ are all elements of $A +_G B$. Since for every $a, a' \in A'$ there are at least $\Omega(p^5 n^3)$ paths of length four between a and a', this means every $y \in A' - A'$ can be written as $x_1 - x_2 + x_3 - x_4$ for

at least $\Omega(p^5n^3)$ quadruples $(x_1, x_2, x_3, x_4) \in (A +_G B)^4$. However, $|A +_G B| \leq Kn$ holds by assumption, so there are at most K^4n^4 such quadruples. By the pigeonhole principle, it then follows that the number of distinct elements $y \in A' - A'$ is at most $O(K^4p^{-5}n)$, as claimed.

Returning to the task of deriving the upper bound from Theorem 6.0.2, we apply Theorem 6.1.1 to the graph G where A and B are chosen to be two copies of our k-AP free A and with an edge between $(a, b) \in A \times A$ if $a + b = 2c$ for some $c \in A$. This graph has precisely pn^2 edges and we can apply Theorem 6.1.1 to it with $K = 1$ since

$$|A +_G A| = |\{2a : \ a \in A\}| = |A|.$$

This yields a subset $A' \subset A$ with $|A'| \geq p|A|/4$ and $|A' - A'| \ll p^{-5}n \ll p^{-6}|A'|$. At this point, we recall a version of the so-called Freiman-Ruzsa modelling lemma (see for instance [**Ru09**, **Theorem 2.3.5**, **page 127**]).

Lemma 6.1.3 *Let S be a finite set of integers and let $r \geq 2$ be an arbitrary integer. Then, there is a set $S^* \subset S$ with $|S^*| \geq |S|/r^2$ which is Freiman r-isomorphic to a set of integers T such that*

$$T \subset \left\{1, 2 \ldots, \left\lceil \frac{1}{r} \cdot |rS - rS| \right\rceil\right\}.$$

Here $rS - rS$ denotes the sumset $S + \ldots + S - S - \ldots - S$, where S appears $2r$ times. For the reader's convenience, we also recall that for any two commutative groups G_1, G_2 two sets $S \subset G_1$ and $T \subset G_2$ are said to be Freiman r-isomorphic if there exists a one to one map $\phi : S \to T$ such that for every $x_1, \ldots, x_r, y_1, \ldots, y_r$ in S (not necessarily distinct) the equation

$$x_1 + \ldots + x_r = y_1 + \ldots + y_r$$

holds if and only if

$$\phi(x_1) + \ldots + \phi(x_r) = \phi(y_1) + \ldots + \phi(y_r).$$

We combine Lemma 6.1.3 with the Plünnecke-Ruzsa inequality, which we saw already in Chapter 3. We recall the statement (with a convenient slight change in notation).

Lemma 6.1.4 *Let S and T be finite sets of reals such that $|S + T| \leq \alpha|S|$, and let r, r' be positive integers. Then*

$$|rT - r'T| \leq \alpha^{r+r'}|S|.$$

Indeed, if we apply this with $S = A'$, $T = -A'$, $r = r' = 2$, and $\alpha = p^{-6}$, we have

$$|2A' - 2A'| \leq p^{-24}|A'| \leq p^{-24}n.$$

Therefore, by Lemma 6.1.3, there is a subset $A^* \subset A'$ with $|A^*| \gg pn$ which is Freiman 2-isomorphic to a set of integers $\phi(A^*)$ contained in the interval

$$\left\{1, \ldots, \left\lceil p^{-24}n \right\rceil\right\}.$$

In particular, since ϕ preserves k-term arithmetic progressions,

$$pn \ll |A^*| = |\phi(A^*)| \leq r_k(\left\lceil p^{-24}n \right\rceil).$$

Lastly, recall that $r_k(n)$ is subadditive as a function of n, namely the inequality

$$r_k(n + n') \leq r_k(n) + r_k(n')$$

holds for all positive integers n, n'. In particular, $r_k(\left\lceil p^{-24}n \right\rceil) \ll p^{-24}r_k(n)$, hence $pn \ll p^{-24}r_k(n)$, or equivalently $p^{25} \ll r_k(n)/n$. This means that A contains at most $(r_k(n)/n)^{1/25}\, n^2$ three-term arithmetic progressions. This completes the proof of the upper bound.

We next prove the desired lower bound on $f_{s,k}(n)$ in Theorem 6.0.2. We begin by revisiting some further simple properties of $r_k(n)$ as a function of n. In addition to being subadditive, we also recall that $r_k(n)$ is an increasing function, so $r_k(m) \leq r_k(n)$ if $m \leq n$. Together these imply that if $n \geq m$, we have $r_k(n) \leq \lceil \frac{n}{m} \rceil r_k(m) \leq \frac{2n}{m} r_k(m)$, so

$$\frac{r_k(n)}{2n} \leq \frac{r_k(m)}{m}. \tag{6.3}$$

For all positive integers m and n, we have

$$r_k(2mn) \geq r_k(m)r_k(n). \tag{6.4}$$

Indeed, if U is a subset of $\{1, \ldots, m\}$ without a k-term arithmetic progression and V is a subset of $\{1, \ldots, n\}$ without a k-term arithmetic progression, then the set

$$W = \{2u(n-1) + v : \ u \in U, v \in V\}$$

is a k-AP free subset of $\{1, \ldots, 2mn\}$ of size $|U||V|$, so (6.4) follows.

In particular, if $n \geq N^{1/2}$, letting $m = \lfloor \frac{N}{2n} \rfloor$, we have

$$r_k(N) \geq r_k(2mn) \geq r_k(n)r_k(m) \geq r_k(n)\frac{m}{2n}r_k(n) \geq \frac{N}{8}\left(\frac{r_k(n)}{n}\right)^2,$$

where the first inequality follows from $r_k(n)$ being an increasing function, the second inequality is by (6.4), the third inequality is by (6.3) using $n \geq m$, and finally the fourth inequality is by substituting in $n \leq 4mN$. It thus follows that

$$\frac{r_k(N)}{N} \geq \frac{1}{8}\left(\frac{r_k(n)}{n}\right)^2. \tag{6.5}$$

Let $N = N_{n,k,s}$ be the least positive integer such that $r_k(N) = \lfloor n/s \rfloor$. Such an N exists since, for every m, $r_k(m+1) = r_k(m)$ or $r_k(m)+1$ and $\lim_{m\to\infty} r_k(m) = \infty$. We will show that for $k > s \geq 3$ and n sufficiently large in terms of k, we have

$$f_{s,k}(n) \geq \left(\frac{n}{300sN}\right)^{s-2} n^2. \tag{6.6}$$

For n sufficiently large in terms of k, we have $n \geq N^{1/2}$ holds (for instance by (6.2)), so (6.5) implies that $n/N \geq s \cdot r_k(N)/N \geq s \cdot (1/8) \cdot (r_k(n)/n)^2$, and hence the lower bound from Theorem 6.0.2 follows from (6.6). We next prove (6.6) using a probabilistic construction of a k-AP free set A of n integers with many s-term arithmetic progressions.

For each $1 \leq i \leq s$, let d_i be an integer chosen uniformly and independently at random from the set $\{1, \ldots, 2N\}$. Let $S \subset \{1, \ldots, N\}$ be a k-AP free set of cardinality $r_k(N) = \lfloor n/s \rfloor$, and S_i denote the translate $\{x + 6(i-1)N - 1 + d_i : x \in S\}$, i.e. $S_i := S + \{6(i-1)N - 1 + d_i\}$.

Finally, let us consider the set $A \subset \{1, \ldots, 6sN\}$ defined by

$$A := \bigcup_{i=1}^{s} S_i.$$

We first check that such a (random) set must be k-AP free. Indeed, the sets $S_1, \ldots, S_s$ are pairwise disjoint since, for each $1 \leq i \leq s$, we have

$$S_i \subset \{6(i-1)N + 1, \ldots, 6(i-1)N + 3N - 1\}.$$

Furthermore, these sets are spaced out so that if an arithmetic progression contains an element from S_i and an element of S_j with $i \neq j$, then its common difference is at least $3N + 2$, in which case the arithmetic progression cannot contain two elements in the same S_i. In particular, every arithmetic progressions in A of length longer than s must be a subset one of the S_i, and hence A is k-AP free. Finally, $|A| = s|S| = s\lfloor \frac{n}{s} \rfloor \leq n$, so A is indeed in $\mathcal{A}_k(n)$, or it can be artificially augmented

to a set in $\mathcal{A}_k(n)$ by adding some elements that do not create k-term arithmetic progressions.

We next lower bound the expected number of s-term arithmetic progressions in A. The number of s-term arithmetic progressions $a, a+D, \ldots, a+(s-1)D$ with $a+(i-1)D \in \{6(i-1)N+N+1, \ldots, 6(i-1)N+2N\}$ for $1 \le i \le s$ is the same as the number of s-term arithmetic progressions in $\{1, \ldots, N\}$ with any integer common difference, which is

$$N + 2\sum_{a=1}^{N-1} \left\lfloor \frac{N-a}{s} \right\rfloor \ge \frac{1}{s}\binom{N}{2}.$$

For each such s-term arithmetic progression $a, a+D, \ldots, a+(s-1)D$ and for each sequence $(a_1, \ldots, a_s)$ of s elements from S, there is a choice of $d_1, \ldots, d_s \in \{1, \ldots, 2N\}$ such that $a_i + 6(i-1)N - 1 + d_i = a + (i-1)D$ for $1 \le i \le s$. Hence, the expected number of s-term arithmetic progressions in A is at least

$$\frac{1}{s}\binom{N}{2}|S|^s(2N)^{-s} \ge \frac{1}{4s}N^2\left(\frac{\lfloor n/s \rfloor}{2N}\right)^s \ge \left(\frac{n}{300sN}\right)^{s-2} n^2.$$

Thus, there must exist a choice of such an A for which the number of s-term arithmetic progressions is at least this lower bound on the expected number, which completes the proof of (6.6) and hence Theorem 6.0.2 is proved.

An intriguing open problem. We would like to end this chapter with a few more words on the upper bound from Theorem 5.0.3 from Chapter 5. In light of Theorem 5.5.5, this is as good in some sense as the upper bound for $r_3(N)$ from (1.5) but, like in the second part of Section 6, one can then similarly ask whether it is possible to always improve on the bound

$$f_3(A) \gg \frac{\log^{1/4}|A|}{2^{2\sqrt{2}\sqrt{\log|A|}}} \cdot |A| \tag{6.7}$$

for *all* sets A without nontrivial four-term progressions. In Theorem 5.0.3, the special 4-AP-free set A we constructed with

$$f_3(A) \ll \frac{1}{(\log N)^{1-\epsilon}} \cdot N$$

also happened to satisfy the property that $T(A) = \Theta(|A|^{3/2})$, so by Proposition 5.5.7 it also has larger three-term progression free sets than we know $\{1, \ldots, N\}$ must have, namely

$$f_3(A) \gg \frac{\log^{1/4} N}{2^{2\sqrt{\log N}}} \cdot N.$$

In [**GR12**], Gyarmati and Ruzsa also improved on (6.7) when $A = \left\{1, 2^2, \ldots, N^2\right\}$ by more number theoretic means that are quite specific to perfect squares. However, is it possible to get a bound better bound than (6.7) for *all* 4-AP free sets A? Theorem 6.0.2 shows that four-term progression free sets of size N may sometimes contain $\gg N^2/2^{3(\log N)^{1/3}}$ three-term progressions, so Proposition 5.5.7 doesn't yield any asymptotic gain over the Elkin lower bound in general. It would be interesting if other methods would be able to provide such a result.

Chapter 7

BEYOND THE CROOT-LEV-PACH LEMMA

In this chapter, we will prove the results mentioned in Section 1.4. We recall the statement of the original Croot-Lev-Pach Lemma from [**CLP17**].

Lemma 7.0.1 *Let $P \in \mathbb{F}[x_1, \cdots, x_n]$ be a multilinear polynomial of degree at most d over a field $\mathbb{F}$, and let M denote the $|\mathbb{F}|^n \times |\mathbb{F}|^n$ matrix with entries $M_{\vec{x},\vec{y}} = P(\vec{x}+\vec{y})$ for $\vec{x}, \vec{y} \in \mathbb{F}^n$. Then*

$$\mathrm{rank}_{\mathbb{F}}(M) \leq 2 \sum_{i=0}^{\lfloor d/2 \rfloor} \binom{n}{i}.$$

7.1 On sets with few distances in $\mathbb{R}^d$

Given a positive integer s, recall that a finite subset A in a metric space M is called an s-distance set in M if there are s positive real numbers $d_1, \ldots, d_s$ such that all the pairwise distances determined by the points in M are among these numbers, and each d_i is realized. In this section, we will give an alternate proof of the following celebrated result for the case $M = R^d$ (with the usual Euclidean distance), which is due Bannai, Bannai and Stanton [**BBS**] from 1983.

Theorem 7.1.1 *If A is an s-distance subset in $\mathbb{R}^d$, then*

$$|A| \leq \binom{d+s}{s}.$$

The proof goes via a slightly improved version of the so-called Croot-Lev-Pach Lemma [**CLP17**] over the reals, which may be of independent interest. We state this in a general form, which captures the original version of the Croot-Lev-Pach Lemma as well.

Theorem 7.1.2 *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and $A \subset V$ be a finite set. Let s be a nonnegative integer and let $p(\overrightarrow{x}, \overrightarrow{y})$ be a $2 \cdot \dim V$-variate*

polynomial wih coefficients in $\mathbb{F}$ *and of degree at most* $2s+1$. *Consider the matrix* $M_{p,A}$ *with rows and columns indexed by* A *and entries* $p(\cdot,\cdot)$. *It corresponds to a (not necessary symmetric) bilinear form on* $\mathbb{F}^A$ *by a formula*

$$\Phi_p(f,g) = \sum_{a,b\in A} p(\overrightarrow{a},\overrightarrow{b})f(a)g(b), \text{ for } f,g : A \to \mathbb{F},$$

which in turn defines a quadratic form $\Phi_p(f,f)$. *Denote by* $\operatorname{rank}(p,A)$ *the rank of matrix* $M_{p,A}$; *if* $\mathbb{F} = \mathbb{R}$ *denote also by* $r_+(p), r_-(p)$ *the inertia indices of the quadratic form* $\Phi_p(f,f)$. *Finally, denote by* $\dim_s(A)$ *the dimension of the space of polynomials of degree at most* s *considered as functions on* A. *Then:*

1) $\operatorname{rank}(p,A) \leqslant 2\dim_s(A)$.

2) if $\mathbb{F} = \mathbb{R}$, *then* $\max\{r_+(p,A), r_-(p,A)\} \leqslant \dim_s(A)$.

We will first prove Theorem 7.1.2, and then we will use it to deduce Theorem 7.1.1. We will need only part 2) of the Lemma above, since part 1) is more or less the original Croot-Lev-Pach lemma in disguise (which doesn't help directly), but we will include nonetheeless a quick new proof of part 1) as well since it motivated part 2).

Proof of Theorem 7.1.2. Endow the space $\mathbb{F}^A$ with a natural inner product $\langle f,g\rangle = \sum_{a\in A} f(a)g(a)$.

Consider the space $\Omega \subset \mathbb{F}^A$ of functions f on A satisfying $\langle f,\phi\rangle = 0$ for all polynomials ϕ of degree at most s. It is easy to see that the dimension of Ω as a vector space over $\mathbb{F}$ is at least $|A| - \dim_s(A)$.

The key observation is that $\Phi_p(f,g) = 0$ whenever $f,g \in \Omega$. Indeed, for any monomial $x^\alpha y^\beta$ in the polynomial $p(\overrightarrow{x},\overrightarrow{y})$ (here α,β are multi-indices with sum of degrees at most $2s+1$) we have

$$\sum_{a,b\in A} a^\alpha b^\beta f(a)g(b) = \left(\sum_{a\in A} a^\alpha f(a)\right)\cdot\left(\sum_{b\in B} b^\beta g(b)\right) = 0,$$

since either α or β have degree at most s and f,g are choosing from Ω.

We will now prove both claims of Theorem 7.1.2 by using dimension arguments.

Indeed, the bilinear form $\Phi_p[\cdot,\cdot]$ on $\mathbb{F}^A$ takes zero values on $\Omega\times\Omega$, thus all non-zero entries of its matrix in appropriate basis (which includes the basis of Ω and

any other $|A| - \dim\ \Omega$ basis vectors) may be covered by $|A| - \dim \Omega$ rows and $|A| - \dim \Omega$ columns. This implies that every minor of $M_{p,A}$ of dimension at least $2(|A| - \dim \Omega) + 1$ must vanish. Therefore,

$$\operatorname{rank}(p, A) \leqslant 2(|A| - \dim \Omega) \leqslant 2 \dim_s(A).$$

This proves the first claim of Theorem 7.1.2.

If $\mathbb{F} = \mathbb{R}$, by Sylvester's Law of Inertia, we may choose a subspace $Y \subset \mathbb{F}^A$ of dimension $r_+(p, A)$ such that the quadratic form $\Phi_p(f, f)$ restricted to Y is positive definite. If $f \in Y \cap \Omega$ and $f \neq 0$, we have $0 = \Phi_p(f, f) > 0$, which is impossible. Therefore, $Y \cap \Omega = \{0\}$ and $\dim Y + \dim \Omega \leqslant |A|$, which yields that $r_+(p, A) = \dim Y \leqslant |A| - \dim \Omega \leqslant \dim_s(A)$. Analogously, we also have that $r_-(p, A) \leqslant \dim_s(A)$. This completes the proof of Theorem 7.1.2.

We now deduce Theorem 7.1.1 from Theorem 7.1.2.

If A is an s-distance subset in $\mathbb{R}^d$ and S is the set of distinct distances it determines, consider the $2d$-variate polynomial p with real coefficients defined by

$$p(\vec{x}, \vec{y}) = \prod_{d \in S} (d^2 - \|\vec{x} - \vec{y}\|^2).$$

The matrix $M_{p,A}$ from Theorem 7.1.2 is then a positive scalar matrix for this polynomial; therefore, $r_+(p, A) = |A|$, and so part 2) of Theorem 7.1.2 implies that

$$|A| = r_+(p, A) \leqslant \dim_s(A) \leqslant \dim_s(\mathbb{R}^d) = \binom{s+d}{d}.$$

This completes the proof of Theorem 7.1.1.

7.2 An algebraic proof of Kleitman's Theorem

To put things into the appropriate context, we start this section with an algebraic argument that comes very close to proving Theorem 1.4.4, but only ends up giving a weaker bound.

For this upcoming discussion, we only address the case $d = 2t$. We enumerate the elements of $\mathbb{F}_2^n$ and consider the $2^n \times 2^n$ matrix M defined by

$$M_{\vec{x},\vec{y}} = \binom{d(\vec{x}, \vec{y}) - 1}{2t} := \frac{(d(\vec{x}, \vec{y}) - 1) \cdots (d(\vec{x}, \vec{y}) - 2t)}{(2t)!}$$

for every $\vec{x}, \vec{y} \in \mathbb{F}_2^n$. Let M' denote the $2^n \times 2^n$ binary matrix obtained from M by reducing each element modulo 2. Note that every two distinct vectors $\vec{x}, \vec{y} \in \mathcal{F}$ has Hamming distance in $\{1, \cdots, 2t\}$, and $\binom{z-1}{2t}$ equals 0 for $z \in \{1, \cdots, 2t\}$, and non-zero for $z = 0$. Therefore the matrix M' restricts on $\mathcal{F} \times \mathcal{F}$ to a full-rank submatrix, and thus $\operatorname{rank} M \geq \operatorname{rank} M' \geq |\mathcal{F}|$. On the other hand, there's a polynomial $p \in \mathbb{F}_2[t_1, .., t_n]$ with $\deg p \leq 2t$ so that

$$p(\vec{x} - \vec{y}) = \binom{d(\vec{x}, \vec{y}) - 1}{2t} \mod 2, \quad \text{for every } \vec{x}, \vec{y} \in \mathbb{F}_2^n$$

This polynomial is given explicitly by

$$p(t_1, \ldots, t_n) = \sum_{S \subset \{1,..,n\}, |S| \leq 2t} \prod_{i \in S} t_i.$$

Indeed, note that for every $x, y \in \mathbb{F}_2^n$,

$$\binom{d(\vec{x}, \vec{y}) - 1}{2t} = \sum_{\ell=0}^{2t} (-1)^\ell \binom{d(\vec{x}, \vec{y})}{\ell}.$$

Furthermore, in $\mathbb{F}_2$ we also have that

$$\sum_{|S|=\ell} \prod_{i \in S} (x_i - y_i) = \binom{d(\vec{x}, \vec{y})}{\ell},$$

so

$$\binom{d(\vec{x}, \vec{y}) - 1}{2t} = \sum_{S \subset \{1,..,n\}, |S| \leq 2t} (-1)^{|S|} \prod_{i \in S} (x_i - y_i) = \sum_{S \subset \{1,..,n\}, |S| \leq 2t} \prod_{i \in S} (x_i - y_i),$$

as claimed. Lemma 7.0.1 then immediately implies

$$|\mathcal{F}| \leq 2 \sum_{i=0}^{t} \binom{n}{i}.$$

When $d = 2t + 1$, it is not to hard to adapt the above argument to show that

$$|\mathcal{F}| \leq 4 \left(\binom{n-1}{0} + \cdots + \binom{n-1}{t} \right).$$

One can however improve on this rank argument and establish the precise version of Theorem 1.4.4. In fact, we will prove Theorem 7.3.1, but to keep things simple for the rest of this section we will stick to the case when $s = 0$ which recovers Theorem 1.4.4. We start with a few lemmas involving simple linear algebra.

Let $M_{n,k}$ be a $2^n \times 2^n$ matrix, whose rows and columns are indexed by vectors in $\{0,1\}^n$. The $(\vec{x}, \vec{y})$-th entry of $M_{n,k}$ is equal to 1 if and only if $\vec{x}$ and $\vec{y}$ differ in exactly k coordinates, and 0 otherwise. For example, $M_{n,1}$ is the adjacency matrix of the n-dimensional hypercube, and $M_{n,k}$ is the adjacency matrix of a Hamming-type graph in which two vertices are adjacent if they are at distance k. The following lemma determines the spectrum of all $M_{n,k}$ for all $1 \le k \le n$.

Lemma 7.2.1 *The spectrum of $M_{n,k}$ consists of $K_k(i;n)$ with multiplicity $\binom{n}{i}$, for $i = 0, \cdots, n$. Here $K_k(i;n)$ is the Krawtchouk polynomial with parameter 2:*

$$K_k(i;n) = \sum_{j=0}^{k} (-1)^j \binom{i}{j}\binom{n-i}{k-j}.$$

For example, when $k = 1$, it is easy to check that the eigenvalues of the n-dimensional hypercube are $K_1(i;n) = n - 2i$ with multiplicity $\binom{n}{i}$. The lemma can be found in [**LW01**]. For completeness, we include its proof below using the Fourier transform on hypercubes as eigenvectors. Throughout the proof we use the notation $d(U,V)$ for the Hamming distance between the indicator vectors of U and V, for two subsets $U, V \subset [n]$.

Proof of Lemma 7.2.1. Let $\vec{v}_S$ be a vector in $\mathbb{R}^{2^n}$ defined as (with its 2^n coordinates viewed as subsets of $[n]$):

$$(\vec{v}_S)_T = (-1)^{|S \cap T|}.$$

It is not hard to show that $\{\vec{v}_S\}_{S \subset [n]}$ form an orthogonal basis. On the other hand,

$$(M_{n,k}\vec{v}_S)_T = \sum_{U \subset [n]} (M_{n,k})_{T,U}(\vec{v}_S)_U = \sum_{U : d(U,T)=k} (-1)^{|S \cap U|}.$$

Note that the number of sets U with the property that U and T differ in j coordinates in S is equal to $\binom{|S|}{j}\binom{n-|S|}{k-j}$. For each of such U,

$$(-1)^{|S \cap U|} = (-1)^{|S \cap T|} \cdot (-1)^j,$$

since $|S \cap U| = |S \cap T| + |(T \Delta U) \cap S| - 2|S \cap T \cap \overline{U}|$. Therefore

$$(M_{n,k}\vec{v}_S)_T = (-1)^{|S \cap T|} \cdot \sum_{j=0}^{|S|} (-1)^j \binom{|S|}{j}\binom{n-|S|}{k-j} = K_k(|S|;n)(\vec{v}_S)_T.$$

This immediately shows that $K_k(i;n)$ are eigenvalues of $M_{n,k}$ with multiplicity $\binom{n}{i}$. This proves Lemma 7.2.1.

From the proof, observe that for fixed n, the eigenspace decomposition of $M_{n,k}$ is the same for every k. Hence it is straightforward to establish the following result.

Lemma 7.2.2 *If $f(1), \cdots, f(n)$ is a sequence of real numbers and let*

$$M = \sum_{k=1}^{n} f(k) M_{n,k}.$$

Then the spectrum of M consists of

$$\lambda_i = \sum_{k=1}^{n} f(k) K_k(i; n)$$

with multiplicity $\binom{n}{i}$, for $i = 0, \cdots, n$.

The following well-known theorem studies the relation between the spectrum of a symmetric matrix and that of its principal minor.

Lemma 7.2.3 *(Cauchy's Interlacing Theorem) Let A be a symmetric matrix of size n, and B is a principal minor of A of size $m \leq n$. Suppose the eigenvalues of A are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and the eigenvalues of B are $\mu_1 \geq \cdots \geq \mu_m$. Then for $1 \leq i \leq m$, we have*

$$\lambda_{i+n-m} \leq \mu_i \leq \lambda_i.$$

The following corollary of the Cauchy's Interlacing Theorem was discovered earlier by Cvetković [**cvetkovic**]. It provides a useful technique to bound the independence number of a graph.

Corollary 7.2.4 *Let G be a n-vertex graph, and M be a symmetric $n \times n$ matrix such that $M_{ij} = 0$ whenever $ij \notin E(G)$ (such M is often called a pseudo-adjacency matrix of G). Let $n_{\leq 0}(M)$ (resp. $n_{\geq 0}(M)$) be the number of non-positive (resp. non-negative) eigenvalues of M. Then the independence number of G satisfies*

$$\alpha(G) \leq \min\{n_{\leq 0}(M), n_{\geq 0}(M)\}$$

Proof of Corollary 7.2.4. Suppose I is a maximum independent set of G with $|I| = \alpha(G)$. Then I naturally corresponds to an all-zero principal minor B of M.

And the eigenvalues of B are $\mu_1 = \cdots = \mu_{\alpha(G)} = 0$. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of M. By Cauchy's Interlacing Theorem,

$$0 = \mu_{\alpha(G)} \leq \lambda_{\alpha(G)}.$$

So M has at least $\alpha(G)$ non-negative eigenvalues. Similarly,

$$\lambda_{1+n-\alpha(G)} \leq \mu_1 = 0,$$

which implies that M has at least $\alpha(G)$ non-positive eigenvalues. This proves Corollary 7.2.4.
Now we are ready to prove Theorem 1.4.4.

Proof of Theorem 1.4.4. For given n, d, we define a graph G whose vertex set $V(G) = \{0,1\}^n$, and two vertices are adjacent if their Hamming distance is at least $d+1$. Kleitman's problem is now equivalent to determining the independence number $\alpha(G)$.

We start with the even case $d = 2t$. By Corollary 6.5 applied to G, it suffices to find real numbers $f(k)$ for $k = 2t+1, \cdots, n$ and define $M = \sum_{k=2t+1}^{n} f(k) M_{n,k}$, such that either the number of non-positive or non-negative eigenvalues of M is at most $\sum_{i=0}^{t} \binom{n}{i}$.

At this point, it is perhaps important to mention that choosing $f(k) = \binom{k-1}{t}$ recovers the $2^n \times 2^n$ symmetric matrix M defined by $M_{x,y} = \binom{d(x,y)-1}{2t}$ from the Croot-Lev-Pach approach, however this is *not* going to be the choice we are going to make for the sequence $f(1), \ldots, f(n)$. We choose $f(k) = \binom{\ell}{t}$ if $k = 2\ell+1$ or $k = 2\ell+2$. Equivalently $f(k) = \binom{\lfloor (k-1)/2 \rfloor}{t}$. By Lemma 7.2.2, the eigenvalue of M with multiplicity $\binom{n}{i}$ is equal to

$$\lambda_i = \sum_{k=2t+1}^{n} f(k) \sum_{j=0}^{k} (-1)^j \binom{i}{j}\binom{n-i}{k-j},$$

for $i = 0, \cdots, n$. Although computing the exact value of λ_i's might not be easy, it turns out that we can determine their signs in a rather straightforward way. We claim that for every t, we have

- $(-1)^i \lambda_i > 0$ for $i = 0, \cdots, t$.
- $\lambda_{n-i} = \lambda_{i+1}$ for $i = 0, \cdots, t-1$.
- $\lambda_{t+1} = \lambda_{t+2} = \cdots = \lambda_{n-t} = (-1)^{t+1}$.

To show the above claims, we use generating functions and observe that λ_i is equal to the constant term of the following formal power series:

$$\left(\sum_{k=2t+1}^{n} f(k)x^{-k}\right)\left(\sum_{j=0}^{i}(-1)^j\binom{i}{j}x^j\right)\left(\sum_{\ell=0}^{n-i}\binom{n-i}{\ell}x^\ell\right).$$

Here in the generating function, we may extend the sum and the domain of f to all the integers greater or equal to $2t+1$, with $f(k) = \binom{\lfloor (k-1)/2 \rfloor}{t}$ as before. This would not affect the constant term since for $\binom{i}{j}$ and $\binom{n-i}{\ell}$ to be non-zero, one must have $j \le i$ and $\ell \le n-i$. So $f(k)$ for only those k up to n may contribute to the constant term.

A quick calculation shows that

$$\begin{aligned}
\sum_{k=2t+1}^{\infty} f(k)x^{-k} &= \binom{t}{t}(x^{-(2t+1)} + x^{-(2t+2)}) + \binom{t+1}{t}(x^{-(2t+3)} + x^{-(2t+4)}) + \cdots \\
&= x^{-(2t+1)}(1+x^{-1})\left(\binom{t}{t} + \binom{t+1}{t}x^{-2} + \cdots\right) \\
&= x^{-(2t+1)}(1+x^{-1})(1-x^{-2})^{-(t+1)} \\
&= \frac{x+1}{(x^2-1)^{t+1}}.
\end{aligned}$$

Note that

$$\sum_{j=0}^{n}(-1)^j\binom{i}{j}x^j = (1-x)^i \quad \text{and} \quad \sum_{\ell=0}^{n-i}\binom{n-i}{\ell}x^\ell = (1+x)^{n-i}.$$

Therefore, λ_i is equal to the constant term of the following power series:

$$\frac{x+1}{(x^2-1)^{t+1}} \cdot (1-x)^i(1+x)^{n-i} = (-1)^{t+1}(1+x)^{n-i-t}(1-x)^{i-t-1}.$$

For $t+1 \le i \le n-t$, both $n-i-t$ and $i-t-1$ are nonnegative, so the constant term is equal to $(-1)^{t+1}$. For $0 \le i \le t$, one needs to consider the constant term of

$$(-1)^{t+1}\frac{(1+x)^{n-i-t}}{(1-x)^{t+1-i}} = (-1)^i\frac{(1+x)^{n-i-t}}{(x-1)^{t+1-i}}.$$

The latter expression rewrites as $(-1)^i x^{-(t+1-i)}(1+x)^{n-i-t}(1-\frac{1}{x})^{-(t+1-i)}$. Obviously in the expansion of $(1+x)^{n-i-t}(1-\frac{1}{x})^{-(t+1-i)}$ for $x > 1$, all the coefficients are positive. So $(-1)^i\lambda_i$ is positive since $t+1-i \le n-i-t$. For $i > n-t$, note that in a power series, substituting x by $-x$ does not change the constant term. Therefore letting $i = n+1-j$, λ_i is equal to the constant term of

$$(-1)^{t+1}\frac{(1+x)^{i-(t+1)}}{(1-x)^{i-(n-t)}} = (-1)^{t+1}\frac{(1+x)^{(n-t)-j}}{(1-x)^{(t+1)-j}},$$

which is exactly $\lambda_j = \lambda_{n+1-i}$.

Therefore for even $t = 2m$, the only non-negative eigenvalues are $\lambda_0, \lambda_2, \cdots, \lambda_{2m}$, $\lambda_{n-1}, \lambda_{n-3}$,
$\cdots, \lambda_{n-(2m-1)}$, and their multiplicities add up to $\sum_{i=0}^{t} \binom{n}{i}$. Similarly when $t = 2m+1$, the only non-positive eigenvalues are $\lambda_1, \cdots, \lambda_{2m+1}, \lambda_n, \cdots, \lambda_{n-2m}$, and their total multiplicity equals $\sum_{i=0}^{t} \binom{n}{i}$ as well. This finishes the proof for the even case.

The proof for the odd case $d = 2t + 1$ works in a similar fashion, except that we have to choose $f(k)$ for $k = 2t+2, \cdots, n$ in a slightly different way. Here we define $f(k) = 0$ for odd k, and $f(k) = \binom{k/2-1}{t}$ for even k. By a similar argument, the eigenvalue λ_i with multiplicity $\binom{n}{i}$ is equal to the constant term of the following formal sum:

$$\left(\sum_{k=2t+2}^{\infty} f(k)x^{-k}\right)\left(\sum_{j=0}^{i}(-1)^j\binom{i}{j}x^j\right)\left(\sum_{l=0}^{n-i}\binom{n-i}{l}x^l\right).$$

It is equal to

$$\begin{aligned}
&x^{-(2t+2)}\left(\binom{t}{t} + \binom{t+1}{t}x^{-2} + \binom{t+2}{t}x^{-4} + \cdots\right)(1-x)^i(1+x)^{n-i} \\
&= x^{-(2t+2)}(1-x^{-2})^{-(t+1)}(1-x)^i(1+x)^{n-i} \\
&= (-1)^{t+1}(1-x)^{i-t-1}(1+x)^{n-i-t-1}.
\end{aligned}$$

Once again, for $t+1 \le i \le n-t-1$, the constant term equals $(-1)^{t+1}$. For $0 \le i \le t$, it is equal to

$$(-1)^{t+1}(1-x)^{-(t+1-i)}(1+x)^{n-i-t-1} = (-1)^i x^{-(t+1-i)}(1+x)^{n-i-t-1}(1-\frac{1}{x})^{-(t+1-i)}.$$

Again note that the expansions of both $(1+x)^{n-i-t-1}$ and $(1-\frac{1}{x})^{-(t+1-i)}$ only consist of positive coefficients. Therefore $(-1)^i\lambda_i > 0$. Similar as before, one can show that for $n - t \le i \le n$, $\lambda_i = \lambda_{n-i}$. Now we apply Corollary 6.5 once again. Note that none of λ_i's is zero. We only need to show that either the number of positive or negative eigenvalues is small. For even $t = 2m$, the only positive λ_i are $\lambda_0, \lambda_2, \cdots, \lambda_{2m}$ and $\lambda_n, \lambda_{n-2}, \cdots, \lambda_{n-2m}$, whose total multiplicity is equal to $2\sum_{i=0}^{m}\binom{n}{2i}$. For odd $t = 2m+1$, the only negative eigenvalues are $\lambda_1, \lambda_3, \cdots, \lambda_{2m+1}$ and $\lambda_{n-1}, \lambda_{n-3}, \cdots, \lambda_{n-(2m+1)}$, whose multiplicity is $2\sum_{i=0}^{m}\binom{n}{2i+1}$. Finally, it is easy to check both sum equal the sum in Kleitman's Theorem for the case $d = 2t + 1$, noting that for even $t = 2m$,

$$2\sum_{i=0}^{m}\binom{n}{2i} = 2\left(\sum_{i=0}^{m}\left(\binom{n-1}{2i-1} + \binom{n-1}{2i}\right)\right) = 2\sum_{i=0}^{t}\binom{n-1}{i},$$

and for odd $t = 2m + 1$,

$$2\sum_{i=0}^{m}\binom{n}{2i+1} = 2\left(\sum_{i=0}^{m}\left(\binom{n-1}{2i} + \binom{n-1}{2i+1}\right)\right) = 2\sum_{i=0}^{t}\binom{n-1}{i}.$$

This proves Theorem 1.4.4.

7.3 Extensions to arbitrary distance sets

In this section, we discuss a few generalizations of Kleitman's theorem to other sets of allowed distances. The next theorem shows that a bound similar to Kleitman's holds for all n when the set of allowed distances is an interval.

Theorem 7.3.1 *For given integers $t > s \geq 0$, suppose $\mathcal{F}$ is a collection of binary vectors in $\{0,1\}^n$, such that for every $\vec{x}, \vec{y} \in \mathcal{F}$, $d(\vec{x}, \vec{y}) \in \mathcal{L}$, with $\mathcal{L} = \{2s + 1, \cdots, 2t\}$, then for all n,*

$$|\mathcal{F}| \leq \binom{n}{t-s} + 2\binom{n}{t-s+1} + \cdots + 2\binom{n}{0}.$$

Proof of Theorem 7.3.1. We follow the proof of Theorem 1.4.4 for a different pseudo-adjacency matrix $M = \sum_{k=1}^{n} f(k)M_{n,k}$. Here for every integer $\ell \geq 0$, we take

$$f(2\ell+1) = f(2\ell+2) = \binom{\ell - s}{t-s}.$$

By extending the definition of binomial coefficients to the whole set of integers, we have that $f(k) \neq 0$ if $k \geq 2t+1$, or $1 \leq k \leq 2s$. Therefore M is a pseudo-adjacency matrix for our purpose of bounding independence number.

The remaining task is to calculate the eigenvalues of M. Using similar arguments as in Theorem 1.4.4, we have that M has an eigenvalue λ_i of multiplicity $\binom{n}{i}$, and λ_i is equal to the constant term in the following formal power series, for $|x| > 1$,

$$\left(\sum_{k=1}^{\infty} f(k)x^{-k}\right)(1-x)^i(1+x)^{n-i}.$$

Let $g_s(x) = \sum_{k=0}^{\infty}\binom{k-s}{t-s}x^k$, we will first show by induction (on s) that for $|x| < 1$, it converges to

$$h_s(x) = \frac{\sum_{j=s}^{t}\binom{t}{j}(x-1)^{t-j}}{(1-x)^{t-s+1}}.$$

For $s = 0$,

$$g_s(x) = \binom{t}{t}x^t + \binom{t+1}{t}x^{t+1} + \cdots = x^t/(1-x)^{t+1},$$

which equals $h_s(x)$. Assume for $s \geq 0$, $g_s(x) = h_s(x)$, then

$$\begin{aligned} h_{s+1}(x) &= h_s(x)(1-x) - (-1)^{t-s}\binom{t}{s} = g_s(x)(1-x) - (-1)^{t-s}\binom{t}{s} \\ &= \sum_{k=0}^{\infty}\binom{k-s}{t-s}x^k - \sum_{k=0}^{\infty}\binom{k-s}{t-s}x^{k+1} - (-1)^{t-s}\binom{t}{s} \\ &= \sum_{k=0}^{\infty}\binom{k-s}{t-s}x^k - \sum_{k=1}^{\infty}\binom{k-s-1}{t-s}x^k - (-1)^{t-s}\binom{t}{s} \\ &= \binom{-s-1}{t-s} + \sum_{k=0}^{\infty}\binom{k-s-1}{t-s-1}x^k - (-1)^{t-s}\binom{t}{s} = g_{s+1}(x). \end{aligned}$$

This completes the proof that $g_s(x) = h_s(x)$. Now we have

$$\begin{aligned} \left(\sum_{k=1}^{\infty} f(k)x^{-k}\right)(1-x)^i(1+x)^{n-i} &= (1-x)^i(1+x)^{n-i}(x^{-1}+x^{-2})g_s(x^{-2}) \\ &= (1-x)^i(1+x)^{n-i+1}x^{-2}\frac{\sum_{j=s}^{t}\binom{t}{j}(x^{-2}-1)^{t-j}}{(1-x^{-2})^{t-s+1}}. \end{aligned}$$

After rearranging, this equals $(-1)^{t-s+1}\sum_{j=s}^{t}\binom{t}{j}(1+x)^{n-i-j+s}(1-x)^{i-j+s-1}x^{2(j-s)}$. Recall that the eigenvalue λ_i of multiplicity $\binom{n}{i}$ is equal to the constant term in the power series. Note that the sum is over all integers j between s and t. In this range, $n-i-j+s \geq n-(t-s)-i$, $i-j+s-1 \geq i-1-(t-s)$, and $2(j-s)$ is strictly greater than 0 except for $j = s$. So for $(t-s)+1 \leq i \leq n-(t-s)$, the product in the sum is a polynomial divisible by x and thus its constant term is 0. Only $j = s$ would contribute to the constant term. Therefore for i in this range,

$$\lambda_i = (-1)^{t-s+1}\binom{t}{s}.$$

In other words, for $(t-s)+1 \leq i \leq n-(t-s)$, λ_i have the same sign. This would immediately imply

$$\begin{aligned} \alpha(G) \leq \min\{n_{\leq 0}(M), n_{\geq 0}(M)\} &\leq \sum_{i=0}^{t-s}\binom{n}{i} + \sum_{i=n-(t-s)+1}^{n}\binom{n}{i} \\ &= \binom{n}{t-s} + 2\binom{n}{t-s+1} + \cdots + 2\binom{n}{0}. \end{aligned}$$

This completes the proof of Theorem 7.3.1.

Remark. For sufficiently large n, by a slightly more careful analysis, one can actually remove the factors of 2 in the statement of Theorem 7.3.1, and show that

$$|\mathcal{F}| \leq \binom{n}{t-s} + \binom{n}{t-s+1} + \cdots + \binom{n}{0},$$

which gives the exact same bound as in Theorem 7.3.1. To achieve this goal, one can show that when $t-s$ is odd, $\lambda_{2i} > 0$ whenever $0 \leq 2i \leq t-s$ and $\lambda_{n-2i-1} > 0$ whenever $n-2i-1 > n-(t-s)$; and when $t-s$ is even, $\lambda_{2i+1} < 0$ whenever $0 \leq 2i+1 \leq t-s$, and $\lambda_{n-2i} < 0$ whenever $n-2i > n-(t-s)$. Then applying Corollary 6.5 gives the desire upper bound. The calculations are a bit tedious, so we decided to omit the details, since it still gives a upper bound that is asymptotically the same, $(1+o(1))\binom{n}{t-s}$. Moreover, using similar techniques, one can show that if the set of allowed distances are $\{2s+1, \cdots, 2t+1\}$, then $|\mathcal{F}| \leq (2+o(1))\binom{n}{t-s}$, generalizing Kleitman's Theorem for odd diameters.

When $\mathcal{L} = \{2s+1, \cdots, 2t\}$, Theorem 7.3.1 gives an upper bound which is $O(n^{t-s})$ for fixed s, t and large n. The following theorem shows that this upper bound is tight up to a constant factor.

Theorem 7.3.2 *For sufficiently large n, there exists a family $\mathcal{F}$ of $(1/\binom{t}{s} - o(1))\binom{n}{t-s}$ binary vectors in $\{0,1\}^n$, such that for every two vectors $\vec{x}, \vec{y} \in \mathcal{F}$,*

$$d(\vec{x}, \vec{y}) \in \mathcal{L} = \{2s+1, \cdots, 2t\}.$$

Proof of Theorem 7.3.2. We will define a family $\mathcal{F}$ consisting of some vectors with $2t$ 1-coordinates. For two such vectors $\vec{x}$ and $\vec{y}$, denote by X and Y the t-sets they naturally correspond to. Then $d(\vec{x}, \vec{y}) \in \{2s+1, \cdots, 2t\}$ is equivalent to $4t - 2|X \cap Y| \in \{2s+1, \cdots, 2t\}$, i.e. $|X \cap Y| \in \{t, \cdots, 2t-s-1\}$. By the famous result of Rödl [**Rod85**] on the Erdős-Hanani Conjecture [**EH63**], for sufficiently large n, there exists a packing of $m = (1-o(1))\binom{n-t}{t-s}/\binom{t}{t-s}$ copies of complete $(t-s)$-uniform hypergraphs K_t^{t-s} in K_{n-t}^{t-s}. Suppose the vertex set of these hypercliques are $V_1, \cdots, V_m$. Then $|V_i| = t$ and $|V_i \cap V_j| \in \{0, \cdots, t-s-1\}$. Take $F_i = V_i \cup \{n-t+1, \cdots, n\}$ and $\mathcal{F} = \{F_1, \cdots, F_m\}$. It is easy to check that $|F_i| = 2t$ and $|F_i \cap F_j| \in \{t, \cdots, 2t-s-1\}$. This proves Theorem 7.3.2.

For a set $\mathcal{L}$ of integers, let $f_{\mathcal{L}}(n)$ be the maximum number of binary vectors in $\{0,1\}^n$ with pairwise Hamming distance in $\mathcal{L}$. The theorems above show that

$$(1-o(1))\binom{n}{t-s}/\binom{t}{s} \le f_{\{2s+1,\cdots,2t\}}(n) \le (1+o(1))\binom{n}{t-s}.$$

For $s = 0$ the upper and lower bounds agree, as shown by Theorem 1.4.4. For general s and t, it is plausible that the lower bound is asymptotically tight. We are able to verify this conjecture for the special case $\mathcal{L} = \{2s+1, 2s+2\}$. We start with the following lemma on subsets of restricted intersection sizes.

Lemma 7.3.3 *Given integers $i > j \ge 1$, if $\mathcal{F}$ is a collection of i-subsets of $[n]$ whose pairwise intersections have size exactly j, then $|\mathcal{F}| \le (1+o(1))n/(i-j)$.*

Proof of Lemma 7.3.3. Suppose $\mathcal{F} = \{F_1, \cdots, F_m\}$, and without loss of generality assume $F_1 = \{1, \cdots, i\}$. For every j-subset S of $[i]$, let $\mathcal{F}_S = \{F : F \cap [i] = S\}$, then by the assumption $\mathcal{F} = \{F_1\} \cup (\bigcup_S \mathcal{F}_S)$. Let $\mathcal{F}'_S = \{F \setminus S : F \in \mathcal{F}_S\}$. Then each non-empty $\mathcal{F}'_S$ consists of pairwise disjoint $(i-j)$-subsets of $\{i+1, \cdots, n\}$. This immediately gives $|\mathcal{F}_S| \le (n-i)/(i-j)$.

We claim that for two distinct j-sets S and T, if both $\mathcal{F}_S$ and $\mathcal{F}_T$ are non-empty, then they both contain at most $i-j$ sets. This is because $|S \cap T| < j$ and thus $\mathcal{F}'_S$ and $\mathcal{F}'_T$ are cross-intersecting, and that a set $U \in \mathcal{F}'_S$ of size $i-j$ can only intersect at most $i-j$ pairwise disjoint subsets. Therefore for all but at most one set S, $|\mathcal{F}_S| \le i-j$. Hence

$$|\mathcal{F}| \le 1 + \sum_{S: S\subset[i], |S|=j} |\mathcal{F}_S| \le 1 + \left(\binom{i}{j} - 1\right)(i-j) + \frac{n-i}{i-j} = (1+o(1))\frac{n}{i-j}.$$

This completes the proof of Lemma 7.3.3.

Theorem 7.3.4 *For integers $s \ge 0$,*

$$f_{\{2s+1,2s+2\}}(n) = (1+o(1))\frac{n}{s+1}.$$

Proof of Theorem 7.3.4. Let $\mathcal{F}$ be a family of m vectors in $\{0,1\}^n$ with pairwise Hamming distance either $2s+1$ or $2s+2$. Without loss of generality assume one of these vectors is the all-zero vector, then the remaining $m-1$ vectors are the indicator vectors of subsets of $[n]$ of size $2s+1$ or $2s+2$. Denote by $\mathcal{A}$ the family of these $(2s+1)$-sets, and $\mathcal{B}$ the family of $(2s+2)$-sets. For two sets $A_1, A_2 \in \mathcal{A}$, we have

$$|A_1| + |A_2| - 2|A_1 \cap A_2| = |A_1 \Delta A_2| \in \{2s+1, 2s+2\}.$$

By considering the parity, this gives $|A_1 \cap A_2| = s$. Similar arguments show that for two sets $B_1, B_2 \in \mathcal{B}$, $|B_1 \cap B_2| = s+1$. And for $A \in \mathcal{A}$, $B \in \mathcal{B}$, $|A \cap B| = s+1$. Now we construct a new family $\mathcal{A}'$ of subsets of $[n+1]$, by adding the element $n+1$ to each set in $\mathcal{A}$. It is straightforward to check that $C = \mathcal{A}' \cup \mathcal{B}$ satisfies the property that every set contains $2s+2$ elements, while every two subsets intersect in exactly $s+1$ elements. Now applying Lemma 7.3.3 for C, we have $|C| \leq (1+o(1))n/(s+1)$, and the same upper bound on $|\mathcal{F}|$ and $f_{\{2s+1,2s+2\}}(n)$ follows. On the other hand, Theorem 7.3.2 with $t = s+1$ gives

$$f_{\{2s+1,2s+2\}}(n) \geq (1-o(1))n/(s+1),$$

so this completes the proof of Theorem 7.3.4.

Note that $\{2s+1, \cdots, 2t\}$ is a set consisting of $2(t-s)$ integers. It is tempting to speculate that the order of magnitude of $f_{\mathcal{L}}(n)$ solely depends on the size of the set $\mathcal{L}$ of allowed distances. However this is false. For example, suppose $\mathcal{L}$ only consists of odd distances. Then $f_{\mathcal{L}}(n) \leq 2$ since if the family contains three vectors, their corresponding subsets A, B, C satisfy

$$2(|A \cup B \cup C| - |A \cap B \cap C|) = |A \Delta B| + |A \Delta C| + |B \Delta C| \equiv 1 \pmod 2,$$

resulting in a contradiction. This observation immediately leads to the following simple upper bound for general $\mathcal{L}$.

Theorem 7.3.5 *If $\mathcal{L}$ is a set of positive integers, c of which are even, then when n tends to infinity, $f_{\mathcal{L}}(n) = O(n^c)$.*

Proof of Theorem 7.3.5. Suppose $\mathcal{L} = \{\ell_1, \cdots, \ell_s\}$, and $\mathcal{F}$ is a family of vectors in $\{0,1\}^n$ with pairwise Hamming distances in $\mathcal{L}$. Without loss of generality assume $\vec{0} \in \mathcal{F}$, and the rest of the vectors correspond to subsets in a family $\mathcal{A}$. Then every subset in $\mathcal{A}$ has size in $\mathcal{L} = \{\ell_1, \cdots, \ell_s\}$. Define $\mathcal{A}_i = \{A : |A| = \ell_i, A \in \mathcal{A}\}$, for $i = 1, \cdots, s$. Then for two distinct subsets X, Y in $\mathcal{A}_i$, their corresponding vectors have Hamming distance equal to

$$2\ell_i - 2|X \cap Y| = |X| + |Y| - 2|X \cap Y| = |X \Delta Y| \in \mathcal{L}.$$

Since there are c even numbers in $\mathcal{L}$, $|X \cap Y|$ belongs to a set of at most c possible intersection sizes. By the Frankl-Wilson Theorem [**fw**], $|\mathcal{A}_i| \leq \binom{n}{c} + \cdots + \binom{n}{0}$, and therefore

$$|\mathcal{F}| = 1 + |\mathcal{A}| = 1 + \sum_{i=1}^{s} |\mathcal{A}_i| = O(n^c),$$

as claimed.

Although the problem of determining the order of magnitude for every fixed distance set $\mathcal{L}$ and sufficiently large n seems beyond our reach, we can still establish asymptotically sharp bounds for some other special distance sets. In fact, we have already established one at the beginning of Section 2. Recall that we started by using the Croot-Lev-Pach Lemma to show a weaker version of Theorem 1.4.4. Similarly, we can also prove the following asymptotically sharp estimate for a different type of arithmetic constraint on $\mathcal{L}$.

Theorem 7.3.6 *For given integers n, k such that $n \geq 2^k$, let $\mathcal{L}$ consist of all the integers between 1 and n that are not divisible by 2^k. Then, for n sufficiently large, we have*

$$f_{\mathcal{L}}(n) = (2 + o(1))\binom{n}{2^{k-1} - 1}.$$

The reader should compare this to Theorem 7.3.5. This was also recorded independently in a blog post by Ellenberg [**Ell**].

Proof. We start by proving the upper bound. Take a $2^n \times 2^n$ matrix M, whose rows and columns correspond to n-dimensional binary vectors, and $M_{\vec{x},\vec{y}} = g(\vec{x} - \vec{y})$. Here $g : \mathbb{F}_2^n \to \mathbb{F}_2$ is the following polynomial:

$$g(\vec{z}) = \prod_{j=0}^{k-1} \left(1 - \binom{\|z\|}{2^j}\right).$$

Here $\|\cdot\|$ is the Hamming norm, so $d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$. Suppose $\mathcal{F}$ is a family of vectors such that their pairwise Hamming distance is not divisible by 2^k. Therefore for distinct $\vec{x}, \vec{y} \in \mathcal{F}$, in the binary representation of $\|\vec{x} - \vec{y}\|$, the last k digits are not all 0.

At this point, we recall the classical Lucas' theorem.

Lemma 7.3.7 *Given a prime number p and two positive integers $A \geq B$ with p-ary representations $A = \sum_{i=0}^{s} a_i p^i$ and $B = \sum_{i=0}^{s} b_i p^i$, then*

$$\binom{A}{B} \equiv \prod_{i=0}^{s} \binom{a_i}{b_i} \pmod{p}.$$

By Lemma 7.3.7, for some $j \in \{0, \cdots, k-1\}$, $\binom{\|z\|}{2^j} \equiv 1 \pmod 2$. Therefore $M_{\vec{x},\vec{y}} = g(\vec{x}-\vec{y}) \equiv 0 \pmod 2$. On the other hand, obviously $M_{\vec{x},\vec{x}} \equiv 1 \pmod 2$. Therefore the family $\mathcal{F}$ naturally induces a submatrix of M, which is a unit matrix in $\mathbb{F}_2$ and has full rank. As a consequence, $|\mathcal{F}|$ is upper-bounded by the $\mathbb{F}_2$-rank of M. Note that $\deg(g) = \sum_{j=0}^{k-1} 2^j = 2^k - 1$. Lemma 7.0.1 immediately implies

$$|\mathcal{F}| \leq 2 \sum_{i=0}^{2^{k-1}-1} \binom{n}{i}.$$

The lower bound can be obtained again by the same extremal construction for Kleitman's theorem, when the allowed distance set is $\{1, \cdots, 2^k - 1\}$. Theorem 1.4.4 gives

$$f_{\mathcal{L}}(n) \geq 2 \sum_{j=0}^{2^{k-1}-1} \binom{n-1}{j}.$$

7.4 On sets in $\mathbb{F}_p^N$ whose difference set avoids the hypercube

In this section, we prove Theorem 1.4.6. Here we briefly sketch the idea. We construct a graph G with vertex set $\mathbb{F}_p^N$, two vertices $\vec{x}$ and $\vec{y}$ are adjacent if $\vec{x}-\vec{y}$ or $\vec{y}-\vec{x}$ is in $J = \{0,1\}^N$. Then $\alpha(G) = D_{\mathbb{F}_p}(J, N)$. We will choose a pseudo-adjacency matrix for G and apply Corollary 6.5.

The following lemma computes the spectrum of a family of matrices, that are natural candidates for the pseudo-adjacency matrix of G.

Lemma 7.4.1 *Let $\omega = e^{2i\pi/p}$ and M be a $p^N \times p^N$ matrix whose rows and columns are indexed by vectors in $\mathbb{F}_p^N$, and $M_{\vec{u},\vec{v}} = f(\vec{u}-\vec{v})$, where f is a function mapping $\mathbb{F}_p^N$ to $\mathbb{R}$. Then the function $\chi_{\vec{v}} : \mathbb{F}_p^N \to \mathbb{C}$ with $\chi_{\vec{v}}(\vec{u}) = \omega^{\langle \vec{u},\vec{v}\rangle}$, when viewed as a vector, is an eigenvector of M, corresponding to the eigenvalue*

$$\sum_{\vec{x}} f(\vec{x}) \omega^{-\langle \vec{v},\vec{x}\rangle}.$$

Moreover all of them form a basis of $\mathbb{C}^{(p^N)}$.

Proof of Lemma 7.4.1. We first verify $\chi_{\vec{v}}$ is an eigenvector of M. We have

$$\begin{aligned}(M\chi_{\vec{v}})_{\vec{z}} &= \sum_{\vec{y}} M_{\vec{z},\vec{y}} \cdot \chi_{\vec{v}}(\vec{y}) = \sum_{\vec{y}} f(\vec{z}-\vec{y}) \cdot \omega^{\langle \vec{v},\vec{y}\rangle} \\ &= \sum_{\vec{x}} f(\vec{x})\omega^{\langle \vec{v},\vec{z}-\vec{x}\rangle} = \omega^{\langle \vec{v},\vec{z}\rangle} \cdot \sum_{\vec{x}} f(\vec{x})\omega^{-\langle \vec{v},\vec{x}\rangle} \\ &= \chi_{\vec{v}}(\vec{z}) \cdot \left(\sum_{\vec{x}} f(\vec{x})\omega^{-\langle \vec{v},\vec{x}\rangle} \right)\end{aligned}$$

It is straightforward to show that $\chi_{\vec{v}}$ are linearly independent, which completes the proof of Lemma 7.4.1.

Now we are ready to prove Theorem 1.4.6.

Proof of Theorem 1.4.6. From the discussions at the beginning of this section, we only need to upper bound the independence number of G.

We define M to be a $p^N \times p^N$ matrix with rows and columns indexed by vectors in $\mathbb{F}_p^N$. We let $M_{\vec{u},\vec{v}} = (-1)^{c(\vec{u}-\vec{v})}$, for vectors $\vec{u} \neq \vec{v}$ with either $\vec{u}-\vec{v}$ or $\vec{v}-\vec{u}$ in $\{0,1\}^N$; and 0 otherwise. Here the function c maps a vector in $\mathbb{F}_p^N$ to its number of non-zero coordinates. Clearly M is a pseudo-adjacency matrix of G. By Lemma 7.4.1, for every $\vec{v} = (v_1, \cdots, v_N) \in \mathbb{F}_p^N$, $\chi_{\vec{v}}$ is an eigenvector of M with eigenvalue equal to

$$\sum_{\vec{x}\in\{0,1\}^N\setminus\vec{0}} (-1)^{c(\vec{x})}\omega^{-\langle \vec{v},\vec{x}\rangle} + \sum_{\vec{x}\in\{0,1\}^N\setminus\vec{0}} (-1)^{-c(\vec{x})}\omega^{\langle \vec{v},\vec{x}\rangle}$$

Note that

$$\begin{aligned}\sum_{\vec{x}\in\{0,1\}^N} (-1)^{c(\vec{x})}\omega^{-\langle \vec{v},\vec{x}\rangle} &= \sum_{\vec{x}\in\{0,1\}^N} (-1)^{\sum_{i=1}^N x_i}\omega^{-\langle \vec{v},\vec{x}\rangle} = \sum_{\vec{x}\in\{0,1\}^N} \prod_{i=1}^N (-1)^{x_i}\omega^{-v_i x_i} \\ &= \prod_{i=1}^N (1-\omega^{-v_i}).\end{aligned}$$

Similarly one can show that

$$\sum_{\vec{x}\in\{0,1\}^N} (-1)^{-c(\vec{x})}\omega^{\langle \vec{v},\vec{x}\rangle} = \prod_{i=1}^N (1-\omega^{v_i}).$$

Therefore $\chi_{\vec{v}}$ corresponds to the eigenvalue

$$\prod_{i=1}^N (1-\omega^{-v_i}) + \prod_{i=1}^N (1-\omega^{v_i}) - 2$$

When $v_j = 0$ for some index j, $\omega^{v_j} = \omega^{-v_j} = 1$, so this gives eigenvalue -2. Otherwise all the $v_j \in \{1, \cdots, p-1\}$. This already shows that the number of non-negative eigenvalues is at most $(p-1)^N$, and Corollary 6.5 gives an upper bound matching Alon's bound. But in fact we can estimate the number of non-negative eigenvalues more carefully. Note that

$$\begin{aligned}\prod_{j=1}^{N}(1-\omega^{-v_j}) &= \prod_{j=1}^{N}(1-\cos(2\pi v_j/p)+i\sin(2\pi v_j/p)) \\ &= \prod_{j=1}^{N} 2\sin(\pi v_j/p)\cdot e^{i(\pi/2-\pi v_j/p)} \\ &= \left(\prod_{j=1}^{N} 2\sin(\pi v_j/p)\right)\cdot e^{i(\pi N/2-\pi\sum_{j=1}^{N} v_j/p)}\end{aligned}$$

Similarly,

$$\prod_{j=1}^{N}(1-\omega^{v_j}) = \left(\prod_{j=1}^{N} 2\sin(\pi v_j/p)\right)\cdot e^{-i(\pi N/2-\pi\sum_{j=1}^{N} v_j/p)}.$$

Therefore

$$\prod_{i=1}^{N}(1-\omega^{-v_i}) + \prod_{i=1}^{N}(1-\omega^{v_i}) - 2 = 2\left(\prod_{j=1}^{N} 2\sin(\pi v_j/p)\right)\cos\left(\pi N/2 - \pi\sum_{j=1}^{N} v_j/p\right) - 2.$$

If it is non-negative, then since $\sin(\pi v_j/p) \geq 0$ for $v_j \in \{1, \cdots, p-1\}$, it must hold that

$$\cos\left(\pi N/2 - \pi\sum_{j=1}^{N} v_j/p\right) > 0.$$

Note that this inequality cannot hold for both $(v_1, \cdots, v_N)$ and $(u_1, \cdots, u_N) = (p-1-v_1, \cdots, p-1-v_p, p-v_{p+1}, \cdots, p-v_N)$. Since

$$\cos\left(\pi N/2 - \pi\sum_{j=1}^{N} u_j/p\right) = -\cos\left(\pi N/2 - \pi\sum_{j=1}^{N} v_j/p\right).$$

Therefore there are at least half of those $(v_1, \cdots, v_N) \in [p-2]^p \times [p-1]^{N-p}$ correspond to negative eigenvalues. Therefore

$$\begin{aligned}\alpha(G) \leq n_{\geq 0}(M) &\leq (p-1)^N - \frac{1}{2}(p-2)^p(p-1)^{N-p} \\ &= \left(1 - \frac{1}{2}\left(1-\frac{1}{p-1}\right)^p\right)(p-1)^N.\end{aligned}$$

When $p \to \infty$ the constant factor tends to $1 - 1/(2e)$.

Remark. We believe that for general p, a more careful analysis of the signs of these eigenvalues should show that for at most half of $(v_1, \cdots, v_N) \in [p-1]^N$, $\cos\left(\pi N/2 - \pi \sum_{j=1}^{N} v_j/p\right) > 0$. This would improve the constant to $1/2$. For some small values of p, we can actually obtain better constants. For example, when $p = 3$, the same method gives $\alpha(G) \leq (1/3 + o(1))2^N$.

BIBLIOGRAPHY

[Alo09] N. Alon, Perturbed identity matrices have high rank: proof and applications, *Combin. Probab. Comput.* **1**8 (2009), 3–15.

[Bis] A. Bishnoi, Spectral proofs of theorems on the boolean hypercube, blog post at https://anuragbishnoi.wordpress.com/2019/07/27/spectral-proofs-of-theorems-on-the-boolean-hypercube.

[AV09] F. Amoroso and E. Viada, Small points on subvarieties of a torus, *Duke Math. Journal*, **150**(3): 407–442, 2009.

[BBS83] E. Bannai, Et. Bannai, D. Stanton, An upper bound for the cardinality of an s-distance subset in real Euclidean space, II, *Combinatorica* **3** (1983) 147–152.

[Beh46] F. A. Behrend, On sets of integers which contain no three terms in arithmetical progression, *P*roc. Nat. Acad. Sci. U.S.A, 32 (1946): 331-332.

[BCCGNSU17] J. Blasiak, T. Church, H. Cohn, J. A. Grochow, E. Naslund, W. F. Sawin, and C. Umans, On cap sets and the group-theoretic approach to matrix multiplication, *D*iscrete Analysis 2017:3.

[Blo81] A. Blokhuis, A new upper bound for the cardinality of 2-distance sets in Euclidean space, *Combinatorica*, **1** (1981), 99-102.

[BLS17] J. Balogh, H. Liu and M. Sharifzadeh, The Number of Subsets of Integers with No k-Term Arithmetic Progression, *Int. Math. Res. Not. IMRN* (2017), no. 20, 6168-6186.

[BMS15] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.* 28 (2015), no. 3, 669-709.

[BMS18] J. Balogh, R. Morris and W. Samotij, The method of hypergraph containers, *arXiv:1801.04584* (2018).

[BS19] J. Balogh, J. Solymosi, On the number of points in general position in the plane, *Discrete Anal.*, Paper No. 16, 20 pp.

[BK11] M. Bateman, N. H. Katz, Structure in additively nonsmoothing sets, preprint at https://arxiv.org/abs/1104.286.

[BK12] M. Bateman, N. H. Katz, New bounds on cap sets, *J.* Amer. Math. Soc. **25** (2012), no. 2, 585–613.

[Blo16] T. F. Bloom, A quantitative improvement for Roth's theorem on arithmetic progressions, *J. Lond. Math. Soc.* (2) **93** (2016), no. 3, 643-663.

[Bou99] J. Bourgain, On triples in arithmetic progression, *Geom. Funct. Anal.* **9** (1999): 968–984.

[Bou08] J. Bourgain, Roth's theorem on progressions revisited, *J. Anal. Math.* **104** (2008): 155–192.

[Bra85] L. Branges, A proof of the Bieberbach conjecture, *Acta Mathematica*, **154**, Number 1-2 (1985), 137-152.

[BT12] B. Bukh, J. Tsimerman, Sum-product estimates for rational functions, *Proc. Lond. Math. Soc.* (3) 104 (2012), no. 1, 1-26

[CR14] D. Carmon, Z. Rudnick, The autocorrelation of the Mőbius function and Chowla's conjecture for the rational function field, *Quaterly J. Math.*, **65** (2014), 53-61.

[Cha06] M.-C. Chang, Sum and product of different sets, *Contributions to Discrete Mathematics*, **1** (1), 2006.

[CG16] D. Conlon and W. T. Gowers, Combinatorial theorems in sparse random sets, *Ann. of Math.* (2) 184 (2016), no. 2, 367-454.

[CLP17] E. Croot, V. Lev and P. P. Pach, Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small, *Ann. of Math.* (2) 185 (2017), no. 1, 331-337.

[CRS18] P. Candela, J. Rué, O. Serra, Memorial to Javier Cilleruelo: a problem list, *Integers*, **18**: Paper No. A28, 9, 2018.

[CS09] E. Croot, O. Sisask, A new proof of Roth's theorem on arithmetic progressions, *Proc. Amer. Math. Soc.* 137 (2009), no. 3, 805-809.

[Ede04] Y. Edel. Extensions of generalized product caps, *Des. Codes Cryptogr.*, **31** (1): 5–14, 2004.

[Ele97] G. Elekes, On the number of sums and products, *Acta Arith.*, **81** (1997) 365–367.

[Ell] J. S. Ellenberg, Difference sets missing a hamming sphere, blog post at https://quomodocumque.wordpress.com/2017/02/11/difference-sets-missing-a-hamming-sphere/.

[Elk11] M. Elkin, An improved construction of progression-free sets, *Israel Journal of Math.* 184 (2011), 93-128.

[EG17] J. Ellenberg and D. Gijswijt, On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, *Ann. of Math.* (2) 185 (2017), no. 1, 339-343.

[EH63] P. Erdős, H. Hanani, On a limit theorem in combinatorial analysis, *Publ. Math. Debrecen*, **1**0 (1963), 10-13.

[ER00] G. Elekes, L. Rónyai, A combinatorial problem on polynomials and rational functions, *J. Combin. Theory Ser. A* **8**9 (2000), 1-20.

[Erd73] P. Erdős, 'Problems and results on combinatorial number theory', Chapter 12 in *A survey of Combinatorial Theory* edited by J. Srivastava, North Holland, 1973.

[Erd74] P. Erdős, Problems and results on finite and infinite graphs, *Recent advances in graph theory* (1974), 183–192.

[Erd81] P. Erdős, Solved and unsolved problems in combinatorics and combinatorial number theory, *Proc. Twelfth Southeastern Conference on Combinatorics, Graph Theory and Compu Congr. Numer.* **3**2 (1981), 49–62.

[EG97] P. Erdős and A. Gyárfás, A variant of the classical Ramsey problem, *Combinatorica* **1**7 (1997), 459–467.

[ES83] P. Erdős, E. Szemerédi, On sums and products of integers, *Studies in pure mathematics*, 213-218, Birkhäuser, Basel, 1983.

[FL17] J. Fox and L. M. Lovász, A tight bound for Green's arithmetic triangle removal lemma in vector spaces, *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*.

[FP20] J. Fox, C. Pohoata, Sets without k-term progressions can have many shorter progressions, preprint at https://arxiv.org/abs/1908.09905.

[FPS17] J. Fox, J. Pach, and A. Suk, More distinct distances under local conditions, *Combinatorica*, **38** (2018), 501-509.

[FPS20] S. Fish, C. Pohoata, A. Sheffer, Local properties via color energy graphs and forbidden configurations, *SIAM Journal on Discrete Mathematics*, 34 (2020), No. 1, pp. 177-187.

[FS11] J. Fox, B. Sudakov, 'Dependent random choice', *Random Structures Algorithms*, 38:68–99, 2011.

[Fur77] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. D Analyse Math*, **7**1 (1977),204–256.

[Fur81] H. Furstenberg, Recurrence in Ergodic Theory and Combinatorial Number Theory, Princeton Univ. Press, 1981.

[GR12] K. Gyarmati, I. Z. Ruzsa, A set of squares without arithmetic progressions, *Acta Arithmetica*, **155** (2012), no. 1, 109–115.

[Gre05] B. Green, Roth's theorem in the primes, *Ann. of Math.* (2) 161 (2005), 1609–1636.

[Gre16] B. Green, Sarkozy's theorem in function fields, *Quarterly J. of Math.*, 68, 2016.

[Gre05] B. Green, Finite field models in additive combinatorics. In Bridget S Webb, editor, *Surveys in combinatorics 2005*, pages 1–27. Cambridge Univ. Press, Cambridge, Cambridge, 2005.

[Gre] B. Green, Some open problems, unpublished manuscript.

[Gro19] J. A. Grochow, New applications of the polynomial method: The cap set conjecture and beyond, *Bull. Amer. Math. Soc.* **56** (2019), 29-64.

[GT08] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. of Math.* (2) **167** (2008), no. 2, 481-547.

[GT17] B. Green, T. Tao, New bounds for Szemerédi's theorem, III: A polylogarithmic bound for $r_4(N)$, *Mathematika*, (63), Issue 3, 2017, 944-1040.

[GY18] A. Glazyrin, W.-H. Yu, Upper bounds for s-distance sets and equiangular lines, *Advances in Mathematics*, Volume 330, 2018, 810-833.

[GW10] B. Green and J. Wolf, A note on Elkin's improvement of Behrend's construction, in *Additive number theory: Festschrift in honor of the sixtieth birthday of Melvyn B. Nathanson*, pages 141–144. Springer-Verlag, 1st edition, 2010.

[Gow98] W. T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, *Geom. Funct. Analysis* **8** (1998), 529-551.

[Gow07] W. T. Gowers, Hypergraph Regularity and the multidimensional Szemerédi Theorem, *Ann. Math.* 166 (2007), 897–946.

[GJ20] W. T. Gowers, O. Janzer, Improved bounds for the Erdős-Rogers function, *Advances in Combinatorics*, 2020:3, 27pp.

[GK15] L. Guth and N. H. Katz, On the Erdős distinct distances problem in the plane, *Ann. of Math.* (2) 181 (2015), no. 1, 155-190.

[Hal71] G. Halász, On the distribution of additive and the mean values of multiplicative arithmetic functions, *Studia Sci. Math. Hungar.*, **6** (1971), 211-233.

[Hal96] R. Hall, Proof of a conjecture of Heath-Brown concerning quadratic residues, *Proc. Edinburgh Math. Soc. (2)*, **39** (1996), 581-588.

[Han18] B. Hanson, *The Additive Structure of Cartesian Products*, Combinatorica **38** (2018), 1095-1100.

[HB87] D. R. Heath-Brown, Integer sets containing no arithmetic progressions, *J. London Math. Soc.* **35** (1987): 385–394.

[Hen16] K Henriot, Arithmetic progressions in sets of small doubling, *Mathematika* **62** (2016), no. 2, 587-613.

[HKP20] H. Huang, O. Klurman, C. Pohoata, *On subsets of the hypercube with prescribed Hamming distances*, to appear in *J. Combin. Theory Ser. A*.

[Kat64] G. O. H. Katona, Intersection theorems for systems of finite sets, *Acta Math. Hungar.* **1**5 (1964), 329–337.

[Kle86] D. Kleitman, *On a combinatorial conjecture of Erdős*, *J. Combin. Theory Ser. A* **4**3 (1986), 85–90.

[KLR96] Y. Kohayakawa, T. Luczak and V. Rődl, Arithmetic progressions of length three in subsets of a random set, *Acta Arith.*, **75** (1996) (2), 133-163.

[KS15] S.V. Konyagin and I.D. Shkredov, On sum sets of sets, having small product set, *Proc. Steklov Inst. Math.* **290** (2015), 288–299.

[KS16] S.V. Konyagin and I.D. Shkredov, New results on sum–products in $\mathbb{R}$, *Proc. Steklov Inst. Math.* **294** (2016), 87–98.

[Kle86] D. Kleitman, On a combinatorial conjecture of Erdős, *J. Combin. Theory Ser. A* **4**3 (1986), 85–90.

[KM18] O. Klurman, A. P. Mangerel, Rigidity theorems for multiplicative functions, *Mathematische Annalen*, **372** (1-2): 651–697, 2018.

[KS03] A. Kostochka and B. Sudakov, On Ramsey numbers of sparse graphs, *Combin. Probab. Comput.* **12** (2003), 627–641.

[KSS75] J. Komlós, M. Sulyok and E. Szemerédi, Linear problems in combinatorial number theory, *Acta Math. Acad. Sci. Hungar.* 26 (1975), 113-121.

[KSS18] R. Kleinberg, W.F. Sawin, D.E. Speyer, The growth rate of tri-colored sum-free sets, *Discrete Analysis*, 2018:12, 10 pp.

[Lê14] T. H. Lê, Problems and results on intersective sets, Combinatorial and additive number theory, *Springer Proceedings in Mathematics & Statistics* **1**01 (2014), 115–128.

[LRS77] D. G. Larman, C. A. Rogers, J. J. Seidel, On two-distance sets in Euclidean space, *Bull. London Math. Soc.* **9** (1977), 261-267.

[LW01] J. H. van Lint, R. M. Wilson, A Course in Combinatorics, *Cambridge University Press*, 2001.

[May19] J. Maynard, Primes with restricted digits, *Inventiones mathematicae*, (2019) 217: 127.

[MF23] M. Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, *M*ath. Z., 17, (1) 228 (1923).

[MS96] J. Matousek, J. Spencer, Discrepancy in Arithmetic Progressions, *J. Amer. Math. Soc.* **9** (1996), 195-204.

[Pet16] F. Petrov, Combinatorial results implied by many zero divisors in a group ring, preprint at https://arxiv.org/abs/1606.03256.

[PT06] J. Pach and G. Tardos, Forbidden paths and cycles in ordered graphs and matrices, *Israel J. Math.* **155** (2006), 359–380.

[PP20] F. Petrov, C. Pohoata, *Improved bounds for progression-free sets in* C_8^n, *Israel J. Math.*, 236 (2020), Issue 2.

[Poh19] C. Pohoata, On cartesian products which determine few distinct distances, *Electron. J. Combin.*, **26** (2019), Issue 1, P1.7.

[Poh20] C. Pohoata, Expanding polynomials on sets with few products, *Mathematika*, 66 (2020), pp. 71-78.

[PRN20] C. Pohoata, O. Roche-Newton, Four-term progression free sets with three-term progressions in all large subsets, preprint at https://arxiv.org/abs/1905.08457.

[PS17] C. Pohoata, A. Sheffer, Higher distances energies and expanders with structure, preprint at https://arxiv.org/abs/1709.06696.

[PS19] C. Pohoata, A. Sheffer, Local properties in colored graphs, distinct distances, and difference sets, *Combinatorica*, **39** (2019), Issue 3, pp 705-714.

[PSS88] J. Pintz, W. L. Steiger, E. Szemerédi, On sets of natural numbers whose difference set contains no squares, *J. London Math. Soc.* **37** (1988), pp. 219-231.

[Ran60] R. A. Rankin, Sets of integers containing not more than a given number of terms in arithmetical progression, *Proc. Roy. Soc. Edinburgh*, Sect. A, (65) 1960/1961, 332-344.

[RN17] O. Roche-Newton, On sets with few distinct distances, preprint at https://arxiv.org/abs/1608.02775.

[RSS16] O. E. Raz, M. Sharir, and J. Solymosi, Polynomials vanishing on grids: The Elekes-Rónyai problem revisited, *Amer. J. Math.* **1**38 (2016), 1029-1065.

[RSZ16] O. E. Raz, M. Sharir, and F. de Zeeuw, Polynomials vanishing on Cartesian products: The Elekes-Szabó Theorem revisited, *Duke Math. J.*, **165** (2016), 3517–3566.

[Rod85] V. Rödl, On a packing and covering problem, *European J. Combin.* **6** (1985), 69–78.

[Rot53] K. F. Roth, On certain sets of integers, *J. London Math. Soc.* **28** (1953), 104-109.

[Rud18] M. Rudnev, On the number of incidences between planes and points in three dimensions, *Combinatorica*, 38 (2018), pp. 219–254.

[Ruz94] I. Z. Ruzsa, Generalized arithmetical progressions and sumsets, *Acta Math. Hungar.*, **65** (1994), no. 4, 379–388.

[Ruz09] I. Z. Ruzsa, 'Sumsets and structure', pp. 87-210 in *Combinatorial number theory and additive group theory*, edited by A. Geroldinger and I. Z. Ruzsa, Birkhäuser, Basel, 2009.

[San09] T. Sanders, Roth's theorem in $\mathbb{Z}_4^n$, *Anal. PDE*, **2** (2009), no. 2, 211-234.

[San09b] T. Sanders, Three-term arithmetic progressions and sumsets, *Proc. Edinb. Math. Soc.* (2) **52** (2009), no. 1, 211-233.

[San11] T. Sanders, On Roth's theorem on progressions, *Ann. of Math.* **174** (2011): 619–636.

[SA77] G. J. Simmons, H. L. Abbott, How many 3-term arithmetic progressions can there be if there are no long ones?, *The Amer. Math. Monthly*, Vol. 84, No. 8 (Oct., 1977), pp. 633-635.

[Sár78] A. Sárközy, On difference sets of sequences of integers, I., *Acta Math. Acad. Sci. Hungar.* **3**1 (1978), 125–149.

[Saw18] W. Sawin, Bounds for Matchings in Nonabelian Groups, *Electron. J. Combin.* **4** (2018), P4.23, 21pp.

[Sch72] W. M. Schmidt, Norm form equations, *Ann. of Math.*, **96**(3): 526–551, 1972.

[Sha19] G. Shakan, On higher energy decomposition and the sum–product phenomenon, *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 167, Issue 3, pp. 599-617.

[SS19] G. Shakan, I. D. Shkredov, Breaking the 6/5 threshold for sums and products modulo a prime, preprint at https://arxiv.org/abs/1806.07091.

[SZZ16] A. Sheffer, J. Zahl, and F. de Zeeuw, Few distinct distances implies no heavy lines or circles, *Combinatorica* **36** (2016), 349–364.

[She12] C. Shen, Algebraic methods in sum-product phenomena, *Israel J. Math.* **1**88 (2012), 123-130.

[1] O. Sisask, Convolutions of sets with bounded VC-dimension are uniformly continuous, preprint at https://arxiv.org/abs/1802.02836.

[Sol09] J. Solymosi, Bounding multiplicative energy by the sumset, *Advances in Mathematics*, Volume 222, **2** (2009), 402-408.

[Spe16] D. Speyer, blog post at https://sbseminar.wordpress.com/2016/07/08/bounds-for-sum-free-sets-in-prime-power-cyclic-groups-three-ways/.

[ST15] D. Saxton and A. Thomasson, Hypergraph containers, *Invent. Math.* 201 (2015), no. 3, 925-992.

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

[Sze90] E. Szemerédi, Integer sets containing no arithmetic progressions, *Acta Math. Hungar.* **56** (1990): 155–158.

[Tao] T. Tao, A symmetric formulation of the croot– lev–pach–ellenberg-gijswijt capset bound, blog post at https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capsetbound/.

[Tao15] T. Tao, Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets, *Contributions to Discrete Mathematics* **10** (2015), 22–98.

[Tao16] T. Tao. The Erdos discrepancy problem, *Discrete Analysis*, 1:29 pp, 2016.

[TV06] T. Tao, V. Vu. Additive combinatorics, *Cambridge University Press* (2006).

[TS] T. Tao, W. Sawin, Notes on the slice rank of tensors, blog post at https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/.

[Vu08] V. Vu, Sum-product estimates via directed expanders, *Math. Res. Lett.*, **15** (2008), 375–388.

[Wir67] E. Wirsing, Das asymptotische Verhalten von Summen űber multiplikative Funktionen. II, *Acta Math. Acad. Sci. Hungar.*, **18** (1967), 411-467.

[Wol13] G. Wolfovitz, K_4-free graphs without large induced triangle-free subgraphs, *Combinatorica* **33** (2013), no. 5, 623-631.