IMPROVED BOUNDS
FOR PROGRESSION-FREE SETS IN $C^n_8$

BY

Fedor Petrov*

Department of Mathematics and Computer Science
St. Petersburg State University
St. Petersburg 199178, Russia
e-mail: f.v.petrov@spbu.ru

AND

Cosmin Pohoata

The Division of Physics, Mathematics and Astronomy
California Institute of Technology
1200 E California Blvd, Pasadena CA 91125, USA
e-mail: apohoata@caltech.edu

ABSTRACT

Let $G$ be a finite group, and let $r_3(G)$ represent the size of the largest subset of $G$ without non-trivial three-term progressions. In a recent breakthrough, Croot, Lev and Pach proved that $r_3(C^n_4) \leq (3.611)^n$, where $C_m$ denotes the cyclic group of order $m$. For finite abelian groups $G \cong \prod_{i=1}^{n} C_{m_i}$, where $m_1, \ldots, m_n$ denote positive integers such that $m_1 | \cdots | m_n$, this also yields a bound of the form $r_3(G) \leq (0.903)^{\text{rk}_4(G)} |G|$, with $\text{rk}_4(G)$ representing the number of indices $i \in \{1, \ldots, n\}$ with $4 | m_i$. In particular, $r_3(C^n_8) \leq (7.222)^n$. In this paper, we provide an exponential improvement for this bound, namely $r_3(C^n_8) \leq (7.0899)^n$.

* Research supported by Russian Science Foundation grant 17-71-20153.
Received January 14, 2019 and in revised form April 16, 2019
1. Introduction

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 $r_3(G)$ 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, the first important estimate being established by Roth in [14]. Currently the best known upper bound is due to Bloom [5], who proved that

$$r_3(C_n) \ll \frac{(\log \log n)^4}{\log n} n.$$ 

The best known lower bound is of the form

$$r_3(C_n) \gg n \exp(-c \sqrt{\log n})$$

for some absolute constant $c > 0$ and is due to Behrend [3]. In particular, $r_3(C_n)$ grows faster than $n^{1-\epsilon}$ for any fixed $\epsilon > 0$.

For other groups $G$, $r_3(G)$ 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 [6], where they showed that

$$r_3(C_4^n) \leq 4^{\gamma n} \approx (3.611)^n.$$ 

The constant $\gamma$ in their paper is given by

$$\gamma := \max \left\{ \frac{1}{2} (H_2(0.5 - \epsilon) + H_2(2\epsilon)) : 0 < \epsilon < 0.25 \right\} \approx 0.926,$$

where $H_2(\theta)$ denotes the binary entropy function

$$H_2(\theta) = -\theta \log_2 \theta - (1 - \theta) \log_2 (1 - \theta), \quad \theta \in (0, 1).$$

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 [15] being of the form

$$r_3(C_4^n) \ll \frac{4^n}{n(\log n)^\epsilon}$$

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 [8] proved that $r_3(C_p^n) \leq \kappa_p^n$ for all odd primes $p$, where $\kappa_n$ generally stands for

$$\kappa_n := \min \{ x^{(1-n)/3} (1 + x + \cdots + x^{n-1}) : x > 0 \}.$$
This was another major result, as it improved dramatically the celebrated estimate
\[ r_3(C_3^n) \ll \frac{3^n}{n^{1+\epsilon}} \]
of Bateman and Katz [2]. This was further adapted by three different teams to prove that for all odd prime powers \( q \), \( r_3(C_q^n) \leq \kappa_q n \) ([4], [17], and [13]), and also later on by various other authors to prove several other different results in extremal combinatorics.

The group algebra approach from [13] allows one to estimate \( r_3(G) \) 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 \( r_3(G) \). The first aim of this paper is to give a group algebra proof of the fact that \( r_3(C_4^n) \leq (3.611)^n \), 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 (1) with the one from (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 is the main goal of our paper.

For a finite abelian group
\[ G \cong \prod_{i=1}^{n} C_{m_i} \]
with positive integers \( m_1 | \cdots | m_n \), denote by \( \text{rk}_4(G) \) the number of indices \( i \in \{1, \ldots, n\} \) with \( 4|m_i \). Since \( G \) is a union of \( 4^{-\text{rk}_4(G)}|G| \) cosets of a subgroup isomorphic to \( C_{4^{\text{rk}_4(G)}} \), this yields a bound of the form
\[ r_3(G) \leq 4^{-(1-\gamma)\text{rk}_4(G)}|G| \approx (0.903)^{\text{rk}_4(G)}|G|. \]
This is the content of Corollary 1 in [6]. For instance, if \( G = C_8^n \), the above gives
\[ r_3(C_8^n) \leq 2^n \cdot r_3(C_4^n) \leq (7.222)^n. \]
In Section 5, we improve on this estimate and show the following:

\(^1\) Multiplicative matchings coincide with what initially were called tricolored sum-free sets in [12]; the updated term is adopted from Aaronson [1] and Sawin [16].
Theorem 1.1: If $A \subset C^n_8$ is a set without non-trivial three-term progressions, then

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

where $2^{\mathcal{H}_4(\rho)}$ represents a weighted version of $\kappa_4$ given by

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

and $\rho_0 \approx 0.32$ solves the system

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

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,

$$r_3(C^n_8) \leq (7.0899)^n.$$

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

Corollary 1.2: If a finite abelian group $G$ is written as

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

where $m_1 | \cdots | m_n$, then

$$r_3(G) \leq (0.8863)^{\text{rank}(G)}|G|,$$

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

This is of course similar in spirit to Corollary 1 from [6] and constitutes an improvement in various other cases beyond Theorem 1.1. Like before, it follows immediately from Theorem 1.1 due to the simple fact that if $n := \text{rank}(G)$, then the group $G$ is a union of $8^{-n}|G|$ cosets of a subgroup isomorphic to $C^n_8$.

2. 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 $C^n_8$ without non-trivial three-term progressions to upper bounding the size of the largest three-term progression-free subset of $C^n_4$ which has the further property that it roughly intersects each of the $2^n$ cosets of $C^n_2$ in the same number of elements.
Let $\Omega$ 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 **weight** of a subset $B \subset A$ by

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

**Lemma 2.1:** 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

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

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

Now we assume that the universe $\Omega$ 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 $G$ contains either 0 or $k'$ species presented in $A$.

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

**Lemma 2.2:** In the above setting, any set $A \subset \Omega$ contains a super-regular subset of weight at least $w(A)/(H_m H_{m'})$.

**Proof.** By Lemma 2.1, 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 2.1 to the partition of these species into genera. \[\blacksquare\]

A similar statement holds for the higher taxonomic hierarchy, and is proved in the same way.
Throughout the paper we apply both Lemmas 2.1 2.2 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 2.1 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 $g \sim h$ if $g^2 = h^2$ and $g \sim h$ if $g^4 = h^4$, respectively. Then by Lemma 2.2 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

$$r_3(G_1 \times G_2) \geq r_3(G_1)r_3(G_2).$$

In particular, by Fekete’s Lemma [11],

$$\lim_{n \to \infty} (r_3(G^n))^{1/n} = \sup_{n > 0} (r_3(G^n))^{1/n}. \tag{4}$$

This implies that any estimate of the form $r_3(G^n) \leq c^{n+o(n)}$ automatically yields $r_3(G^n) \leq c^n$. In particular, Lemmas 2.1 and 2.2 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 $\leq c^{n+o(n)}$.

3. 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 1.1. 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 $X_1 \cdots X_k$ the product set
\[
\{x_1 \cdot \cdots \cdot x_k : x_i \in X_i \text{ for every } i \in \{1, \ldots, k\}\}.
\]
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 $\tau_i$ have the 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 $\prod_{i=1}^{n} \tau_i^{\lambda_i}$, $0 \leq \lambda_i < 2^{m_i}$. 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
\[
\deg_{\text{max}} := \sum_{i=1}^{n} w_i (2^{m_i} - 1),
\]
their product equals 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 $\theta \deg_{\text{max}}$. Then $X(\theta_1)X(\theta_2) \cdots X(\theta_k) = 0$ provided that $\sum \theta_i \geq 1$. Note that codim $X(\theta)$ equals the number of monomials of degree at most $\theta \deg_{\text{max}}$. To estimate the number of such monomials, we may use a Chernoff type argument, as follows. If $0 < x \leq 1$, we get
\[
\text{codim } X(\theta) \leq x^{-\theta \deg_{\text{max}}} \prod_{i=1}^{n} (1 + x^{w_i} + x^{2w_i} + \cdots + x^{(2^{m_i} - 1)w_i}) =: \Phi(\theta)(x).
\]
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 $\theta \deg_{\text{max}}$ corresponds to a contribution $x^{\sum w_i \lambda_i - \theta \deg_{\text{max}}} \geq 1$. Note that if $\theta \leq 1/2$, we have $\Phi(\theta)(x) \leq \Phi(1/x)$ for $x \geq 1$. Thus the minimum of $\Phi(\theta)(x)$ over all positive $x$ is attained on $(0, 1]$. Therefore in this case we may write codim $X(\theta) \leq \min_{x \in (0, 1]} \Phi(\theta)(x)$.

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
\[
\text{codim } X(\theta) \leq (\min_{x>0} x^{-\theta(k-1)} (1 + x + x^2 + \cdots + x^{k-1}))^n.
\]
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 is rewritten as
\[
\text{(5)} \quad \text{codim } X(\theta) \leq 2^{n \mathcal{H}_k(\theta)},
\]
which we will use repeatedly throughout the paper. We note that for $k = 2$ this is the usual binary entropy function

$$
H_2(\theta) = \min_{x > 0} -\theta \log_2 x + \log_2 (1 + x) = -\theta \log_2 \theta - (1 - \theta) \log_2 (1 - \theta).
$$

We also note that with all of these notations we may rewrite (2) as

$$
\log_2 \kappa_p = H_p(1/3).
$$

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,

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

Here $\theta \in [1/4, 1/2]$ is arbitrary. It follows that $\log_2 \kappa_4 \geq H_2(\theta) + H_2(1 - 2\theta)$. Taking the maximum over $\theta$ we get

$$
(6) \quad \log_2 \kappa_4 \geq \max_{\theta \in [1/4, 1/2]} \{H_2(\theta) + H_2(1 - 2\theta)\}.
$$

Actually we have an equality in (6). This may be explained as follows: choose $\theta$ such that the minimum of $x^{-2\theta}(1 + x^2)$ is attained at $x_0$; this gives

$$
H_2(\theta) = \log_2 x_0^{-2\theta}(1 + x_0^2).
$$

Then both the product

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

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 $H_2(1 - 2\theta) = \log_2 x_0^{2\theta - 1}(1 + x_0)$. Hence for this specific value of $\theta$ we get $H_2(\theta) + H_2(1 - 2\theta) = \log_2 \kappa_4$, and the maximum over all possible values of $\theta$ is not less than $\log_2 \kappa_4$, or in other words, (6) is an equality. In particular,

$$
\log_2 \kappa_4 = \max_{\theta \in [1/4, 1/2]} \{H_2(\theta) + H_2(1 - 2\theta)\} = \max_{\epsilon \in (0, 1/4)} \{H_2(0.5 - \epsilon) + H_2(2\epsilon)\},
$$

i.e., $\kappa_4 = 4^7$, where

$$
\gamma = \frac{1}{2} \cdot \max_{0 < \epsilon < 0.25} \{H(0.5 - \epsilon) + H(2\epsilon)\}.
$$
4. The Croot–Lev–Pach bound for $C_4^n$ with group rings

In this section, we use the subspaces with vanishing product from Section 3 to give the promised alternate proof of

$$r_3(C_4^n) \leq (3.611)^n.$$ 

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

**Theorem 4.1:** If $A \subset C_4^n$ is a set without non-trivial three-term progressions, then

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

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

**Proof.** From the regularization argument (Lemma 2.1) and tensor power trick from Section 2, it is enough to prove that $|A| \leq 2\kappa_4^n$ 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 $\beta^n$ 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 \( \{x^2 : x \in A\} \). Next, choose $\theta \in [1/4, 1/2]$ such that $\log_2 \beta = H_2(1 - 2\theta)$. Consider the subspaces $X(1 - 2\theta), X(\theta)$ in $F_2[C_2^n]$ defined in Section 3. By (5),

$$\text{codim } X(1 - 2\theta) \leq 2^n H_2(1 - 2\theta) = \beta^n < |A^2|,$$

$X(1 - 2\theta)$ must have a common non-zero element with the subspace of $F_2$-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

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

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

$$\left(\sum_{g \in C} \varphi(g)g_0^{-1}g\right)\left(\sum_{g \in C} \psi(g)g_0^{-1}g\right)\left(\sum_{h \in A^2} \eta(h^{-1})h^{-1}\right)$$
inside the group algebra $\mathbb{F}_2[C_n^4]$, where the functions $\varphi, \psi : C \rightarrow \mathbb{F}_2$ are chosen so that

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

This product equals 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 $\varphi, \psi$ satisfying (9). However, the vector subspace of $\mathbb{F}_2^C$ spanned by the functions $\varphi$ with $\sum_{g \in C} \varphi(g) g_0^{-1} g \in X(\theta)$ has codimension at most $\text{codim} X(\theta)$, and so does the subspace spanned by the functions $\psi$ such that $\sum_{g \in C} \psi(g) g_0^{-1} g \in X(\theta)$. By (5), the sum of their codimensions is at most $2 \cdot \text{codim} X(\theta) \leq 2 \cdot 2^n H_2(\theta)$, while $\log_2 \beta = H_2(1-2\theta)$, which by (6) yields $\log_2 \alpha \geq H_2(\theta)$. Putting these together, we conclude that the sum of the codimensions of these spaces is at most

$$
2 \cdot \text{codim} X(\theta) \leq 2 \cdot 2^n H_2(\theta) \leq 2\alpha^n < |C|,
$$

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

For the proof of Theorem 1.1, we will require a few additional tools.

5. Improved bounds for progression-free sets in $C_n^8$

In this section, we present the proof of Theorem 1.1. We begin with some further linear algebraic preliminaries.

**Lemma 5.1:** If $X_1, X_2$ are the subspaces of a linear space $X$ over a certain field, the codimension of the subspace $X_1 \cap X_2$ in $X_2$ does not exceed $\text{codim} X_1$.

**Proof.** The space $X_1$ is a set of vectors in $X$ satisfying certain $m := \text{codim} X_1$ linear equations. The vectors in $X_2$ satisfying these $m$ equations form a subspace of $X_2$ of codimension at most $m$. ■
Let $\Omega$ be a finite set, $K$ be a fixed field and $K^\Omega$ a space of $K$-valued functions on $\Omega$. For a function $f \in K^\Omega$ denote by $\text{supp}(f) = \{x \in \Omega : f(x) \neq 0\}$ the support of $f$.

**Lemma 5.2:** Suppose that $X \subset K^\Omega$ is a space of dimension $d$. Then $X$ contains a function $f$ with $|\text{supp}(f)| \geq d$.

While simple, this observation was an important step in the Ellenberg–Gijswijt argument from [8]. We record the short proof here for the reader’s convenience.

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

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

which contradicts the choice of $f$. 

Last but not least, we will also need a generalization of a fact which we used at the end of the proof of Theorem 4.1.

**Lemma 5.3:** 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,

$$\text{codim} X + \text{codim} Y \geq |\text{supp}(a)|.$$

**Proof.** Denote $\Omega_0 = \text{supp}(a)$. There is a natural embedding of $K^{\Omega_0}$ into $K^\Omega$. By Lemma 5.1, the subspaces $X_0 = X \cap K^{\Omega_0}, Y_0 = Y \cap K^{\Omega_0}$ have codimensions in $K^{\Omega_0}$ at most $\text{codim} X, \text{codim} Y$ 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 $|\Omega_0| = |\text{supp}(a)|$, and the statement of Lemma 5.3 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 $h g^2 h^{-1} = (h g h^{-1})^2$. Furthermore, fix an arbitrary field $K$. For a subset $A \subset G$, we identify $A^K$ with a span of $A$ as a subset of the group algebra $K[G]$. In particular, we have that $H^K = K[H]$.

Lemma 5.4: Let $X, Y, Z$ be subspaces of $K[H]$ which satisfy $XYZ = 0$. Suppose $A \subset G$ satisfies the following conditions:

(i) $|A^2| \geq 5 \cdot \text{codim} Y$;
(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 $\frac{5}{4}(\text{codim } X + \text{codim } Z)$ elements of $A$.

Then, $A$ contains a three-term progression.

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

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

Such a map $\eta$ exists by Lemma 5.2. 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 $g_0 H = H g_0$ and choose two arbitrary functions $\varphi : A \cap g_0 H \rightarrow K, \psi : A \cap g_0 H \rightarrow K$ such that

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

Since $XYZ = 0$, we have that $xy_0 z = 0$, so the coefficient of $g_0^{-2}$ in this product equals 0. On the other hand, it equals

$$\sum_{a \in g_0 H \cap A, b \in g_0 H \cap A, c \in A^{-2}, abc = 1} \varphi(a) \eta(c) \psi(b).$$

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

$$\sum_{a \in g_0 H \cap A} \varphi(a) \psi(a) \eta(a^{-2}) = 0.\quad (10)$$
We claim that for a certain $g_0 \in G$, this is a contradiction to Lemma 5.3. To see this, recall first that the choice of $\eta$ and (ii) assured us that at least $\frac{4}{5}|A|$ elements $a \in A$ are such that $a^{-2} \in \text{supp}(\eta)$. By the pigeonhole principle, this means that there exists a coset $g_0 H$ such that at least $\frac{4}{5}|A \cap g_0 H|$ elements of $A \cap g_0 H$ satisfy this condition. Since the vector space spanned by the functions $\Psi \in K^{A \cap g_0 H}$ such that $\sum_{a \in g_0 H} \Psi(a)g_0^{-1}a \in X$ (respectively, such that $\sum_{b \in Hg_0} \Psi(b)bg_0^{-1} \in Z$) has codimension at most $\text{codim} X$ (respectively, $\text{codim} Z$), Lemma 5.1 and (10) imply that
\[
\text{codim} X + \text{codim} Z \geq |\{a \in A \cap g_0 H : a^{-2} \in \text{supp}(\eta)\}| \geq \frac{4}{5}|A \cap g_0 H|,
\]
a contradiction to (iii).

We will use this lemma to first complete the proof of Theorem 1.1.

**Proof of Theorem 1.1.** From the regularization argument (Lemma 2.2) and the tensor power trick from Section 2, it is enough to prove $|A| \leq 10.05 \cdot (7.09)^n$ 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,
\[|A| = 10.05 \cdot (\alpha \beta \gamma)^n, \quad |A^2| = 5 \cdot (\gamma \beta)^n, \quad |A^4| = 4.02 \cdot \gamma^n.
\]
In this setup, note that we may also assume that $\alpha, \beta, \gamma$ are all in the interval $(1, 2]$. Indeed, the fact that $\alpha, \beta, \gamma \leq 2$ 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 $\min\{\alpha, \beta, \gamma\} \leq 1$, 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 4.1 to write $|A \cap C| \leq (\kappa_4)^n$. 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 \( A \cap C \subset C^n_4 \), 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 \( A \cap C \subset C^n_4 \). Indeed, if \( \log_2 \alpha \geq \mathcal{H}_2(\theta) \) we have that

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

so we can consider once again the (zero) product from (8) for a suitable intersection \( A_{g_0} \) of \( A \cap C \) with a coset of \( C^n_2 \). Similarly, the fact that \( A \cap C \) has no three-term progressions then produces two spaces of functions, \( \Phi \) and \( \Psi \), each with codimension at most \( \text{codim } X(\theta) \) 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

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

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

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

Applied for \( \mathbb{F}_2[C^n_4] \) and the subspaces \( Y = X(1 - 2\rho), X = Z = X(\rho) \), Lemma 5.4 thus yields

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

which implies

\[
\log_2 \alpha \beta \leq \mathcal{H}_4(\rho).
\]

This condition imposes a special further constraint on \( \alpha, \beta, \gamma \), 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., \( r_3(C^n_8) = O(c^n) \), 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 \( \gamma \) 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 \( \rho \geq 1/3 + o(1) \), and

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

which represents a contradiction.
MAXIMIZING $\alpha \beta \gamma$. For the reader’s convenience, let us first recall the restrictions we have on $\alpha$, $\beta$ and $\gamma$; in the previous subsection, we showed that there exist positive reals $\rho, \theta \in [1/4, 1/2]$ such that

$$
\begin{align*}
\log_2 \beta &= H_2(1 - 2\theta), \\
\log_2 \alpha &\leq H_2(\theta), \\
\log_2 \gamma \beta &= H_4(1 - 2\rho), \\
\log_2 \alpha \beta &\leq H_4(\rho).
\end{align*}
$$

The maximal value of $\alpha \beta \gamma$ for $\rho, \theta \in [1/4, 1/2]$, $\alpha, \beta, \gamma \in [1, 2]$ and (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 $\alpha_0 \beta_0 \gamma_0 \leq 4$, which is definitely not a maximum, thus $\beta_0 > 1$. Choose $\gamma$ slightly greater than $\gamma_0$ and $\beta < \beta_0$ so that $\beta_0 \gamma_0 = \beta \gamma$. Since the binary entropy function $H_2$ is increasing on $[0, 1/2]$, the new $\theta$ such that $\log_2 \beta = H_2(1 - 2\theta)$ satisfies $\theta > \theta_0$. In particular, this means that there exists $\alpha > \alpha_0$ such that $\log_2 \alpha \leq H_2(\theta)$ and $\alpha \beta \leq \alpha_0 \beta_0$. We have

$$
\alpha \beta \gamma = \alpha \beta_0 \gamma_0 > \alpha_0 \beta_0 \gamma_0.
$$

a contradiction to 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 (11) must be an equality. We argue this again by contradiction; suppose that $\log_2 \alpha_0 < H_2(\theta_0)$. Then we may choose $\beta$ slightly less than $\beta_0$, define $\rho$ by $\log_2 2\beta = H_4(1 - 2\rho)$ and $\theta$ by $\log_2 \beta = H_2(1 - 2\theta)$. After that we may choose $\alpha \in (\alpha_0 \beta_0 / \beta, 2)$ so that (11) still holds for $\alpha, \beta$ (and $\gamma = \gamma_0 = 2$), which yields a contradiction. This is indeed clear when $\log_2 \alpha_0 \beta_0 < H_4(\rho_0)$, but even if we had equality in the last line from (11), namely $\log_2 \alpha_0 \beta_0 = H_4(\rho_0)$, then we can choose $\alpha$ so that

$$
\log_2 \alpha \beta = H_4(\rho) > H_4(\rho_0) = \log_2 \alpha_0 \beta_0.
$$

Therefore, $\log_2 \alpha_0 = H_2(\theta_0)$. We also claim that equality must hold in the last inequality from (11). Suppose that $\log_2 \alpha_0 \beta_0 < H_4(\rho_0)$. The function $H_2(1 - 2x) + H_2(x)$ 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 $(\alpha, \beta)$ slightly so that the product $\alpha \beta$ increases and the conditions from (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 \(\rho_0 \geq 1/3\), but then by the analysis from Section 3
\[
\mathcal{H}_4(1/3) \geq \mathcal{H}_4(1 - 2\rho) \geq \log_2 \gamma_0 \beta_0 = \log_2 2\beta_0 > \log_2 \alpha_0 \beta_0 = \mathcal{H}_4(1/3),
\]
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
\[
(12) \quad \mathcal{H}_4(\rho_1) = \mathcal{H}_2(\theta_1) + \mathcal{H}_2(1 - 2\theta_1), \quad \mathcal{H}_4(1 - 2\rho_1) = 1 + \mathcal{H}_2(1 - 2\theta_1).
\]
We claim that \(\theta_1 = \theta_0, \rho_1 = \rho_0\). We argue this one last time by contradiction. If \(\theta_0 < x_0 \leq \theta_1\), 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 \(\alpha_0\) and \(\beta_0\) by \(\alpha\) and \(\beta\) defined by
\[
\log_2 \alpha = \mathcal{H}_2(\theta_1), \quad \log_2 \beta = \mathcal{H}_2(1 - 2\theta_1),
\]
with \(\alpha \beta > \alpha_0 \beta_0\), contradicting the maximality of \(\alpha_0 \beta_0\). If \(\theta_0 \geq x_0\), 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 \(\theta_0\) and \(\theta_1\). 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 (12). The first equation defines \(\rho_1\) as a (strictly) decreasing function of \(\theta_1\), whereas the second equation represents it as an increasing one. Thus such \(\theta_1\) is (a priori at most) unique and the approximate estimates may be specified by the Intermediate Value Theorem. Numerically, the values of \(\theta_1, \rho_1\) 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 1.1. ■
6. Concluding remarks

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 [7], 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

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

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

**Theorem 6.1:** Suppose that $G = (\mathbb{Z}/8\mathbb{Z})^n$. Then there is a set $A \subset G$ with no three-term progression and

$$|A| = \Omega(|G|^{\log 5/\log 8}/\sqrt{\log |G|}).$$

**Proof.** 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 $Y_1 + \cdots + Y_n = 2n$.

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 non-empty 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.

A similar story holds for three-term progression-free sets inside $(\mathbb{Z}/4\mathbb{Z})^n$, where the product constructions seemingly lead only to lower bounds of the form

$$r_3(G) = \Omega(C^n),$$

where $C < 3$. One can easily adapt the above construction to get the following lower bound.

**Theorem 6.2:** Suppose that $G = (\mathbb{Z}/4\mathbb{Z})^n$. Then there is a set $A \subset G$ with no three-term progression and

$$|A| = \Omega(|G|^3/\log 4/\sqrt{\log |G|}).$$

A similar construction for $(\mathbb{Z}/4\mathbb{Z})^n$, due to Elsholtz, is described by Sanders in [15], who cites an unpublished manuscript which we couldn’t access. In the meantime, Elsholtz and Pach independently generalized further this lower bound story and produced similar examples in $(\mathbb{Z}/m\mathbb{Z})^n$. See [10].

**References**

[1] J. Aaronson, *A connection between matchings and removal in abelian groups*, https://arxiv.org/abs/1612.04172.

[2] M. Bateman and N. H. Katz, *New bounds on cap sets*, Journal of the American Mathematical Society 25 (2012), 585–613.

[3] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proceedings of the National Academy of Sciences of the United States of America 32 (1946), 331–332.

[4] 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*, Discrete Analysis (2017), Article no. 3.
[5] T. F. Bloom, A quantitative improvement for Roth’s theorem on arithmetic progressions, Journal of the London Mathematical Society 93 (2016), 643–663.

[6] E. Croot, V. Lev and P. Pach, Progression-free sets in $\mathbb{Z}_3^n$ are exponentially small, Annals of Mathematics 185 (2017), 331–337.

[7] Y. Edel, Extensions of generalized product caps, Designs, Codes and Cryptography 31 (2004), 5–14.

[8] J. Ellenberg and D. Gijswijt, On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, Annals of Mathematics 185 (2017), 339–343.

[9] Y. Edel, Extensions of generalized product caps, Designs, Codes and Cryptography 31 (2004), 5–14.

[10] C. Elsholtz and P. Pach, Caps and progression-free sets in $\mathbb{Z}_m^n$, preprint.

[11] M. Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, Mathematische Zeitschrift 17 (1923), 228–249.

[12] R. Kleinberg, W.F. Sawin, D.E. Speyer, The growth rate of tri-colored sum-free sets, preprint at arXiv:1607.00047, 2016.

[13] F. Petrov, Combinatorial results implied by many zero divisors in a group ring, https://arxiv.org/abs/1606.03256,

[14] K. F. Roth, On certain sets of integers, Journal of the London Mathematical Society 28 (1953), 104–109.

[15] T. Sanders, Roth’s theorem in $\mathbb{Z}_4^n$, Analysis & PDE 2 (2009), 211–234.

[16] W. Sawin, Bounds for matchings in nonabelian groups, Electronic Journal of Combinatorics 25 (2018), Article no. 4.23.

[17] D. Speyer, https://sbseminar.wordpress.com/2016/07/08/bounds-for-sum-free-sets-in-prime-power-cyclic-groups-three-ways/.