DISCRETE ANALYSIS, 2018:12, 10 pp.
www.discreteanalysisjournal.com

The growth rate of tri-colored sum-free sets

Robert Kleinberg David E Speyer Will Sawin

Received 17 May 2017; Revised 6 June 2018; Published 6 July 2018

arXiv:1607.00047v3 [math.CO] 6 Jul 2018

Abstract: Let G be an abelian group. A tri-colored sum-free set in G is a collection of triples $(\boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{c}_i)$ in G such that $\boldsymbol{a}_i + \boldsymbol{b}_j + \boldsymbol{c}_k = 0$ if and only if $i = j = k$. Fix a prime q and let C_q be the cyclic group of order q. Let $\theta = \min_{\rho>0}(1+\rho+\cdots+\rho^{q-1})\rho^{-(q-1)/3}$. Blasiak, Church, Cohn, Grochow, Naslund, Sawin, and Umans (building on previous work of Croot, Lev and Pach, and of Ellenberg and Gijswijt) showed that a tri-colored sum-free set in C_q^n has size at most $3\theta^n$. Between this paper and a paper of Pebody, we will show that, for any $\delta > 0$, and n sufficiently large, there are tri-colored sum-free sets in C_q^n of size $(\theta-\delta)^n$. Our construction also works when q is not prime.

1 Introduction

Let G be an abelian group. Let $\boldsymbol{t} \in G^n$. We make the following slightly nonstandard definition: a ***sum-free set in G^n with target $\boldsymbol{t}$*** is a collection of triples $(\boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{c}_i)$ in $G^n \times G^n \times G^n$ such that $\boldsymbol{a}_i + \boldsymbol{b}_j + \boldsymbol{c}_k = \boldsymbol{t}$ if and only if $i = j = k$. We may always replace $(\boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{c}_i)$ by $(\boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{c}_i - \boldsymbol{t})$ to make the target $\boldsymbol{0}$ (as we did in the abstract, and as is more standard), but allowing an arbitrary target will simplify our notation. The usual terminology is "tri-colored sum-free set", but we omit the reference to the coloring as we never consider any other kind.

If $X \subset G^n$ is a set with no three-term arithmetic progressions, then $\{(\boldsymbol{x}, \boldsymbol{x}, -2\boldsymbol{x}) : \boldsymbol{x} \in X\}$ is sum-free with target $\boldsymbol{0}$, so lower bounds on sets without three-term arithmetic progressions are also bounds on sum-free sets. The reverse does not hold: the largest known three-term arithmetic progression free subsets of C_3^n (where C_q is the cyclic group of order q) are of size 2.217^n [10]. Before this paper, the largest known sum-free sets in C_3^n were of size 2.519^n [1]; this paper will raise the bound to 2.755^n and show that this bound is tight.

Letting $r_3(G^n)$ denote the largest subset of G^n with no three-term arithmetic progressions, the question of whether $\limsup_{n\to\infty} r_3(G^n)^{1/n} < |G|$ was open, until recently, for every abelian G containing elements of order greater than two. The breakthrough work of Croot, Lev, and Pach [9] introduced a polynomial method to prove that strict inequality holds when G is cyclic of order 4, and Ellenberg and Gijswijt [12] built upon their ideas to prove it for cyclic groups of odd prime order. Blasiak et al. [5] applied the same method to prove upper bounds for sum-free sets in G^n for any fixed finite abelian group G.

We recall here one case of their bound. Let C_q be the cyclic group of order q. Let $\theta = \min_{\beta>0}(1+\beta+\cdots+\beta^{q-1})\beta^{-(q-1)/3}$ and let ρ be the value of β at which the minimum is attained. We note that the minimum is attained at a unique point which belongs to $(0,1)$ because $(1+\beta+\cdots+\beta^{q-1})\beta^{-(q-1)/3}$ approaches ∞ as

© 2018 Robert Kleinberg, Will Sawin, and David E Speyer
Licensed under a Creative Commons Attribution License (CC-BY) DOI: 10.19086/da.3734

β goes to 0 from above, is increasing on the interval $[1,\infty)$, and has increasing first derivative on the interval $(0,1)$.

The following result of [5] is closely related to the results of [12] (for primes) and [21, Theorem 4] (for prime powers). (What we denote θ is called $qJ(q)$ in [5].)

Theorem 1 ([5, Theorem 4.14])**.** *If q is a prime power, then sum-free sets in C_q^n have size at most $3\theta^n$.*

Prior to this paper, it was not clear whether any of these applications of the polynomial method yielded tight bounds. In fact, Theorem 1 is tight to within a subexponential factor.

Theorem 2. *Fix an integer $q \geq 2$. Define θ as above. For n sufficiently large, there are sum-free sets in C_q^n with size $\geq \theta^n e^{-2\sqrt{(2\log 2\log\theta)n}-O_q(\log n)}$.*

In this paper, we show Theorem 2 except for a hypothesis on the existence of a probability distribution satisfying certain properties (Theorem 4). In [20], Pebody will verify Theorem 4, completing the proof of Theorem 2.[1]

The question of whether Theorem 1 also yields a tight bound for $\limsup r_3(G^n)^{1/n}$ remains open.

Sum-free sets have applications in theoretical computer science, especially the circle of ideas surrounding fast matrix multiplication algorithms. The $O(n^{2.41})$ algorithm of Coppersmith and Winograd [8] rests on a combinatorial construction that can, in hindsight, be interpreted[2] as a large sum-free set in $\mathbb{F}_2^n$. In the same paper they presented a conjecture in additive combinatorics that, if true, would imply that the exponent of matrix multiplication is 2, i.e., that there exist matrix multiplication algorithms with running time $O(n^{2+\varepsilon})$ for any $\varepsilon > 0$. This conjecture, along with another conjecture by Cohn et al. [7] that also implies the exponent of matrix multiplication is 2, was shown by Alon, Shpilka, and Umans [1] to necessitate the existence of sum-free sets of size $3^{n-o(n)}$ in $\mathbb{F}_3^n$. The upper bound on sum-free sets by Blasiak et al. [5] thus refutes both of these conjectures. Furthermore, Blasiak et al. [5] show that a more general family of proposed fast matrix multiplication algorithms based on the "simultaneous triple product property" (STPP) [7] in an abelian group H necessitates the existence of sum-free sets of size $|H|^{1-o(1)}$. Their upper bound on sum-free sets in abelian groups of bounded exponent thus precludes achieving matrix multiplication exponent 2 using STPP constructions in such groups.

A second application of sum-free sets in theoretical computer science concerns property testing, the study of randomized algorithms for distinguishing functions f having a specified property from those which have large Hamming distance from every function that satisfies the property. A famous example is the Blum-Luby-Rubinfeld (BLR) linearity tester [6], which queries the function value at only $O(1/\delta)$ points and succeeds, with error probability less than $1/3$, in distinguishing linear functions on $\mathbb{F}_2^n$ from those that have distance $\delta \cdot 2^n$ from any linear function. Testers which can distinguish low-degree polynomials on $\mathbb{F}_2^n$ from those that are far from any low-degree polynomial constitute an important ingredient in the celebrated PCP Theorem [2]. Bhattacharya and Xie [4] demonstrated that constructions of large sum-free sets in $\mathbb{F}_2^n$ could be used to derive lower bounds on the complexity of testing certain linear-invariant properties of Boolean functions.

Finally, sum-free sets have applications to removal lemmas in additive combinatorics, a topic that is heavily intertwined with property testing. In particular, Green [16] proved an "arithmetic removal lemma" for abelian groups which implies that for every $\varepsilon > 0$, there is a $\delta > 0$ such that for any abelian group G and three subsets A,B,C, either there are at least $\delta|G|^2$ distinct triples $(a,b,c) \in A\times B\times C$ satisfying $a+b+c=0$, or one can eliminate all such triples by deleting at most $\varepsilon|G|$ elements from each of A,B, and C. Green's argument yields an upper bound for δ^{-1} which is a tower of twos of height polynomial in ε^{-1}. This bound

[1] [20] was written in response to a preprint version of the paper which stated Theorem 4 as a conjecture, and we have chosen to preserve the chronology here, though otherwise updating the paper to reflect his result.

[2] This interpretation was made explicit by Fu and Kleinberg [15].

can be improved using combinatorial[3] or Fourier analytic[4] techniques, but for general abelian groups G the value of δ is not bounded below by any polynomial function of ε. However, when G is the group $\mathbb{F}_q^n$, Fox and Lovasz [14] have applied our nearly-tight construction of sum-free sets in G to obtain bounds of the form

$$\varepsilon^{-C_q+o(1)} < \delta^{-1} < (\varepsilon/3)^{-C_q},$$

where C_q is a constant depending on q but not n, and where $o(1)$ goes to 0 as ε goes to 0 for any fixed q.

2 Notation

Throughout this paper, we will use the following conventions: Lower case Roman letters denote integers, elements of cyclic groups (denoted C_q), of finite fields (denoted $\mathbb{F}_q$), or general finite sets. Lower case Roman letters in boldface denote elements of $\mathbb{Z}_{\geq 0}^m$ (for any m), C_q^m or $\mathbb{F}_q^m$. Capital Roman letters denote subsets of $\mathbb{Z}_{\geq 0}^m$, C_q^m or $\mathbb{F}_q^m$. Lower case Greek letters denote real numbers; lower case Greek letters in boldface denote elements of $\mathbb{R}^m$. A notation such as $\alpha(x)$ or $\boldsymbol{\alpha}(x)$ refers to a function of x valued in real numbers, or real vectors. For any sets U and V, we write U^V for the set of U-valued functions on V. All logarithms are to base e.

We fix a positive integer q. In section 4, we will fix n to be a positive integer divisible by 3. The notation $O_q(\ )$ will always refer to bounds as $n \to \infty$ through integers divisible by 3, with q fixed. Let $\boldsymbol{t} = (q-1, q-1, \ldots, q-1) \in \mathbb{Z}_{\geq 0}^n$.

We define the following sets of lattice points:

$$\begin{aligned} I &= \{0, 1, \ldots, q-1\} \subset \mathbb{Z}_{\geq 0} \\ T &= \{(a,b,c) \in I^3 : a+b+c = q-1\} \end{aligned}$$

3 Entropy

Let A be a finite set and let $\boldsymbol{e} = (e_1, e_2, \ldots, e_n) \in A^n$. We define the probability distribution $\boldsymbol{\sigma}(\boldsymbol{e})$ on A by $\boldsymbol{\sigma}_a(\boldsymbol{e}) = \#\{r : e_r = a\}/n$. In other words, $\boldsymbol{\sigma}(\boldsymbol{e})$ is the probability distribution of uniformly randomly selecting an element of $\boldsymbol{e}$.

Let A be a finite set and $\boldsymbol{\lambda} \in \mathbb{R}_{\geq 0}^A$ a probability distribution on A. The ***entropy***, $\eta(\boldsymbol{\lambda})$, is defined by

$$\eta(\boldsymbol{\lambda}) = -\sum_{a \in A} \boldsymbol{\lambda}_a \log(\boldsymbol{\lambda}_a)$$

where $0 \log 0$ is considered to be 0. The importance of the entropy function in our situation is the following:

Lemma 3. *Let A be a finite set, and let $\boldsymbol{e}_0 \in A^n$. Then*

$$n\eta(\boldsymbol{\sigma}(\boldsymbol{e}_0)) - O_{|A|}(\log n) \leq \log\left(\#\{\boldsymbol{e} \in A^n : \boldsymbol{\sigma}(\boldsymbol{e}) = \boldsymbol{\sigma}(\boldsymbol{e}_0)\}\right) \leq n\eta(\boldsymbol{\sigma}(\boldsymbol{e}_0)).$$

The implied constant in O depends only on $|A|$ and not on n or $\boldsymbol{e}_0$.

Proof. For $a \in A$, let $n_a = n\boldsymbol{\sigma}_a(\boldsymbol{e}_0)$ be the number of times a appears in $\boldsymbol{e}_0$.

The number of $\boldsymbol{e} \in A^n$ such that $\boldsymbol{\sigma}(\boldsymbol{e}) = \boldsymbol{\sigma}(\boldsymbol{e}_0)$ is equal to the multinomial coefficient

$$\binom{n}{(n_a)_{a \in A}} := \frac{n!}{\prod_{a \in A} n_a!}.$$

[3] See [13], building upon the combinatorial proof of Green's result in [18].

[4] See [17], which pertains to the case $G = \mathbb{F}_2^n$ and adapts the proof idea of [13] to the analytic setting.

For the upper bound, we take one term from the multinomial formula

$$n^n = \left(\sum_{a\in A} n_a\right)^n \geq \binom{n}{(n_a)_{a\in A}} \prod_{a\in A} n_a^{n_a},$$

so

$$\binom{n}{(n_a)_{a\in A}} \leq \prod_{a\in A}\left(\frac{n}{n_a}\right)^{n_a} = \exp(n\eta(\boldsymbol{\sigma}(e_0))).$$

For the lower bound, we use the following version of Stirling's formula. (See, e.g., [22].)

$$(n+\tfrac{1}{2})\log(n) - n + \tfrac{1}{2}\log(2\pi) \;<\; \log(n!) \;<\; (n+\tfrac{1}{2})\log(n) - n + \tfrac{1}{2}\log(2\pi) + \tfrac{1}{12}$$

Applying this estimate to each of the factorial terms, and using $\sum_{a\in A} n_a = n$ we find that

$$\left|\log\binom{n}{(n_a)_{a\in A}} - \sum_{a\in A} n_a \log\left(\frac{n}{n_a}\right)\right| \leq |A|\left[\log(n) + \log(2\pi) + \frac{1}{6}\right].$$

Note that $\eta(\boldsymbol{\sigma}(\boldsymbol{e}_0)) = \sum_{a\in A} \frac{n_a}{n}\log\left(\frac{n}{n_a}\right)$, so this gives

$$\left|\log\binom{n}{(n_a)_{a\in A}} - n\eta(\boldsymbol{\sigma}(\boldsymbol{e}_0))\right| \leq |A|\left[\log(n) + \log(2\pi) + \frac{1}{6}\right]. \qquad \square$$

If A and B are finite sets, $f : A \to B$ is a map and $\boldsymbol{\lambda}$ is a probability distribution on A, then we define the probability distribution $f_*\boldsymbol{\lambda}$ on B by

$$(f_*\boldsymbol{\lambda})_b = \sum_{a\in f^{-1}(b)} \boldsymbol{\lambda}_a.$$

It is well known that $\eta(f_*\boldsymbol{\lambda}) \leq \eta(\boldsymbol{\lambda})$, with strict inequality if there are distinct elements a_1 and $a_2 \in A$ with $f(a_1) = f(a_2)$ and $\boldsymbol{\lambda}_{a_1}, \boldsymbol{\lambda}_{a_2} > 0$.

With ρ and θ as defined before, define a probability distribution $\boldsymbol{\psi}$ on I by

$$\boldsymbol{\psi}_k = \frac{\rho^k}{1+\rho+\cdots+\rho^{q-1}}.$$

Let $f : T \to I$ be the map $f((i,j,k)) = k$. The following is proved in [20].[5]

Theorem 4 ([20, Theorem 4])**.** *There is an S_3-symmetric probability distribution $\boldsymbol{\pi}$ on T with $f_*(\boldsymbol{\pi}) = \boldsymbol{\psi}$.*

More precisely, [20] proves that $\boldsymbol{\psi}, \boldsymbol{\psi}, \boldsymbol{\psi}$ are compatible in the sense that there are random variables X_1, X_2, X_3 whose distributions are each $\boldsymbol{\psi}$ and such that $X_1 + X_2 + X_3$ is constant. As each variable has expectation $(p-1)/3$, that constant is certainly $p-1$, so (X_1, X_2, X_3) is a random T-valued variable. Its probability distribution is a probability distribution on T whose three projections are each $\boldsymbol{\psi}$. Symmetrizing it, we obtain an S_3-symmetric probability distribution on T whose projection under f is $\boldsymbol{\psi}$, as stated in Theorem 4.

We will need to compute:

Lemma 5. *With notation as above, $\eta(\boldsymbol{\psi}) = \log\theta$.*

[5] A proof was also claimed in a preprint [19], but we are unable to confirm all the steps in the argument.

Proof. Note that
$$\boldsymbol{\psi}_k = \frac{\rho^{k-(q-1)/3}}{\theta}.$$

We have
$$\eta(\boldsymbol{\psi}) = -\sum_{k\in I} \boldsymbol{\psi}_k \log \frac{\rho^{k-(q-1)/3}}{\theta} = \left(\sum_{k\in I} \boldsymbol{\psi}_k\right) \log\theta - \left(\sum_{k\in I} (k-(q-1)/3)\boldsymbol{\psi}_k\right) \log\rho. \tag{1}$$

The result follows by substituting
$$\sum_{k\in I} \boldsymbol{\psi}_k = 1$$
$$\sum_{k\in I} (k-(q-1)/3)\boldsymbol{\psi}_k = \frac{\rho}{\theta} \cdot \frac{d}{d\beta} \left[(1+\beta+\cdots+\beta^{q-1})\beta^{-(q-1)/3}\right]_{\beta=\rho} = 0,$$

into (1). □

Remark 6. If $\boldsymbol{\pi}$ is any S_3-symmetric probability distribution on T then $f_*(\boldsymbol{\pi})$ has expected value $\frac{q-1}{3}$. Of all probability distributions on I with expected value $\frac{q-1}{3}$, the distribution $\boldsymbol{\psi}$ has the greatest entropy.

4 The construction

Let $\boldsymbol{\pi}$ be the probability distribution on T guaranteed by Theorem 4. Fix n divisible by 3, so that when S_3 acts on the lattice $\mathbb{Z}^T$ by permuting the coordinates according to the S_3 action on T, the fixed point set of the action includes lattice vectors whose coordinates sum up to n. We can approximate $\boldsymbol{\pi}$ to within $O_q(1/n)$ by an S_3-symmetric distribution $\boldsymbol{\pi}'$ where the probability of each element is an integer multiple of $1/n$; such a $\boldsymbol{\pi}'$ can be found by scaling down $\mathbb{Z}^T$ by $1/n$, taking the set of S_3-fixed points that belong to the probability simplex, and selecting the closest such point to $\boldsymbol{\pi}$. Then the marginal distribution $\boldsymbol{\psi}'$ will be within $O_q(1/n)$ of $\boldsymbol{\psi}$. The entropy function of a probability distribution, viewed as function of the vector of the probabilities of the elements, is a differentiable function on the open set of probability distributions assigning positive probability to every element. Thus, because $\boldsymbol{\psi}$ assigns positive probability to each element, the entropy is Lipschitz in a neighborhood of $\boldsymbol{\psi}$. For large enough n, $\boldsymbol{\psi}'$ is in that neighborhood, so
$$\eta(\boldsymbol{\psi}') = \eta(\boldsymbol{\psi}) - O_q(1/n) = \log\theta - O_q(1/n). \tag{2}$$

(The second equality is Lemma 5.)

Define the following sets:
$$W = \{\boldsymbol{a} \in I^n : \boldsymbol{\sigma}(\boldsymbol{a}) = \boldsymbol{\psi}'\}$$
$$V = \{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in W^3 : \boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c} = \boldsymbol{t}\}.$$

We will show in Lemma 10 that $|V|$ and $|W|$ grow exponentially in n, with $|V|$ having the faster growth rate. Our sum-free set in C_q^n will be a subset of V.

Let p be a prime number between $4|V|/|W|$ and $8|V|/|W|$ (such a prime exists by Bertrand's postulate). Since $|V|$ grows faster than $|W|$, the prime p goes to ∞ as n does. Let S be a subset of $\mathbb{F}_p$ having no three distinct elements in arithmetic progression. Behrend's construction [3], with Elkin's improvement [11], implies that, for p sufficiently large one can choose such a set whose cardinality is at least $p \cdot e^{-2\sqrt{2\log 2\log p}}$.

Let $h : \mathbb{Z}^{n+2} \to \mathbb{F}_p$ be a linear map, chosen uniformly at random from all such linear maps. For any $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V$, the sequence
$$h(0,1,\boldsymbol{a}),\ \tfrac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}),\ h(1,0,\boldsymbol{c})$$

constitutes a (possibly degenerate) arithmetic progression in $\mathbb{F}_p$. Thus, this arithmetic progression is contained in S if and only if its three terms are all equal to one another and lie in S. Define V' to be the subset of V given by

$$V' = \left\{ (\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in W^3 : \begin{array}{l} \boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=\boldsymbol{t} \\ h(0,1,\boldsymbol{a}) = \frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}) = h(1,0,\boldsymbol{c}) \in S \end{array} \right\}.$$

Define V'' to be the set of all $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V'$ such that every other $(\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c}') \in V'$ obeys $\boldsymbol{a}' \neq \boldsymbol{a}, \boldsymbol{b}' \neq \boldsymbol{b}, \boldsymbol{c}' \neq \boldsymbol{c}$.

Remark 7. For this remark, assume q is odd. Define a tri-colored 3-AP-free set in C_q^n to be a set of triples $(\boldsymbol{a}_i,\boldsymbol{b}_i',\boldsymbol{c}_i)$ in $(C_q^n)^3$ such that $\boldsymbol{a}_i+\boldsymbol{c}_k = 2\boldsymbol{b}_j'$ if and only if $i=j=k$. Replacing $(\boldsymbol{a}_i,\boldsymbol{b}_i,\boldsymbol{c}_i)$ with $(\boldsymbol{a}_i, \frac{1}{2}(\boldsymbol{t}-\boldsymbol{b}) \bmod q, \boldsymbol{c}_j)$ turns any tri-colored sum-free set into a tri-colored 3-AP-free set. In our set V'', each of $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ has entries distributed over I with probability distribution $\boldsymbol{\psi}$. Therefore in the tri-colored 3-AP free set, the entries of $\boldsymbol{a}$ and $\boldsymbol{c}$ will be distributed with probability $\boldsymbol{\psi}$, but the entries of $\boldsymbol{b}$ will be distributed with the different distribution $g_*\boldsymbol{\psi}$ where $g: I \to I$ is the map $g(b) = \frac{1}{2}(q-1-b) \bmod q$. By contrast, if $X \subset C_q^n$ is a 3-AP-free set in the standard sense, then $\{(\boldsymbol{x},\boldsymbol{x},\boldsymbol{x}) : \boldsymbol{x} \in X\}$ is a tri-colored 3-AP-free set but, for this tri-colored 3-AP-free set, each of the three components has the same distribution. This discrepancy suggests that it may be hard to lift our constructions out of the colored setting.

The set V'' will be our sum-free set. We verify that it is sum-free in Lemma 9.

Lemma 8. *For any $\boldsymbol{a} = (a_1,a_2,\ldots,a_n) \in W$, we have $\sum a_i = n(q-1)/3$.*

Proof. By definition, $\boldsymbol{\sigma}(\boldsymbol{a}) = \boldsymbol{\psi}'$, so we want to show the expected value of the distribution $\boldsymbol{\psi}'$ is $(q-1)/3$. But $\boldsymbol{\psi}'$ is the marginal of the S_3 symmetric distribution $\boldsymbol{\pi}'$ on T. As $\boldsymbol{\pi}'$ is a symmetric distribution for a triple of random variables summing to $q-1$, the expectation of each variable must be $(q-1)/3$. □

Lemma 9. *For any choice of the map h, the set V'' is a sum-free set with target $\boldsymbol{t}$ in C_q^n.*

Proof. Suppose that we have three (not necessarily distinct) triples $(\boldsymbol{a}_i,\boldsymbol{b}_i,\boldsymbol{c}_i)\,(i=0,1,2)$ in V'' such that $\boldsymbol{a}_0+\boldsymbol{b}_1+\boldsymbol{c}_2=\boldsymbol{t}$ in C_q^n.

We claim that we also have $\boldsymbol{a}_0+\boldsymbol{b}_1+\boldsymbol{c}_2=\boldsymbol{t}$ in $\mathbb{Z}^n$. By Lemma 8, the entries of $\boldsymbol{a}_0$, $\boldsymbol{b}_1$ and $\boldsymbol{c}_2$ each sum to $n(q-1)/3$ (in $\mathbb{Z}$) so the sum of all the entries of $\boldsymbol{a}_0+\boldsymbol{b}_1+\boldsymbol{c}_2$ (with the sum taken in $\mathbb{Z}$) must be $n(q-1)$. Now the sum $\boldsymbol{a}_0+\boldsymbol{b}_1+\boldsymbol{c}_2$ in $\mathbb{Z}^n$ has each entry congruent to $q-1 \bmod q$, by the assumption $\boldsymbol{a}_0+\boldsymbol{b}_1+\boldsymbol{c}_2=\boldsymbol{t}$ in C_q^n, and each entry is nonnegative, because the entries of $\boldsymbol{a}_0,\boldsymbol{b}_1$, and $\boldsymbol{c}_2$ are nonnegative. So each entry is at least $q-1$. We just saw that the sum of all the entries is $n(q-1)$, so each entry is exactly $q-1$, as claimed.

Now that we know $\boldsymbol{a}_0+\boldsymbol{b}_1+\boldsymbol{c}_2=\boldsymbol{t}$ in $\mathbb{Z}^n$, we deduce that $\left(h(0,1,\boldsymbol{a}_0), \frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}_1), h(1,0,\boldsymbol{c}_2)\right)$ is an arithmetic progression in $\mathbb{F}_p$. Since $(\boldsymbol{a}_0,\boldsymbol{b}_0,\boldsymbol{c}_0) \in V'$, we have $\boldsymbol{a}_0 \in W$ and $h(0,1,\boldsymbol{a}_0) \in S$. Similarly, $\boldsymbol{b}_1, \boldsymbol{c}_2 \in W$ and $\frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}_1), h(1,0,\boldsymbol{c}_2) \in S$. So $\left(h(0,1,\boldsymbol{a}_0), \frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}_1), h(1,0,\boldsymbol{c}_2)\right)$ is a (possibly degenerate) arithmetic progression in S. As S is arithmetic-progression-free, we must have $h(0,1,\boldsymbol{a}_0) = \frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}_1) = h(1,0,\boldsymbol{c}_2) \in S$. We have now checked that $(\boldsymbol{a}_0,\boldsymbol{b}_1,\boldsymbol{c}_2)$ obeys all the conditions to be an element of V'.

Now, recalling the definition of V'' and the fact that $(\boldsymbol{a}_i,\boldsymbol{b}_i,\boldsymbol{c}_i) \in V'$ for $i=0$, 1, 2, we may conclude that $(\boldsymbol{a}_i,\boldsymbol{b}_i,\boldsymbol{c}_i) = (\boldsymbol{a}_0,\boldsymbol{b}_1,\boldsymbol{c}_2)$ for $i=0$, 1, 2. In other words, the three triples $(\boldsymbol{a}_0,\boldsymbol{b}_0,\boldsymbol{c}_0)$, $(\boldsymbol{a}_1,\boldsymbol{b}_1,\boldsymbol{c}_1)$ and $(\boldsymbol{a}_2,\boldsymbol{b}_2,\boldsymbol{c}_2)$ are all equal to one another. □

We will now begin to estimate the expected value of $|V''|$.

Lemma 10. *We have*

$$|V| \;\geq\; \exp(\eta(\boldsymbol{\pi}')n - O_q(\log n))$$

and

$$\exp(\eta(\boldsymbol{\psi}')n) \;\geq\; |W| \;\geq\; \exp(\eta(\boldsymbol{\psi}')n - O_q(\log n)).$$

Since $\boldsymbol{\psi'} = f_*\boldsymbol{\pi'}$, we have $\eta(\boldsymbol{\pi'}) \geq \eta(\boldsymbol{\psi'})$. Moreover, if n is large enough that the distribution $\boldsymbol{\pi'}$ is not a point-mass on $(\frac{q-1}{3}, \frac{q-1}{3}, \frac{q-1}{3})$, then we have strict inequality since $\boldsymbol{\pi'}$ is S_3-symmetric, so $\boldsymbol{\pi}'_{ijk} > 0$ implies $\boldsymbol{\pi}'_{jik} > 0$. This establishes the previous claim that $|V|$ and $|W|$ grow exponentially, with $|V|$ having the faster rate.

Proof. Since $W = \{\boldsymbol{e} \in I^n : \boldsymbol{\sigma}(\boldsymbol{e}) = \boldsymbol{\psi'}\}$, the lower and upper bounds for $|W|$ follow from Lemma 3. We now need to establish the lower bound for V.

Let $V_0 = \{\boldsymbol{f} \in T^n : \boldsymbol{\sigma}(\boldsymbol{f}) = \boldsymbol{\pi'}\}$. An element of T^n is an n-tuple of triples of integers $((a_1,b_1,c_1),(a_2,b_2,c_2),\ldots,(a_n,b_n,c_n))$ with $a_i + b_i + c_i = q-1$. Reorganizing these integers as $((a_1,a_2,\ldots,a_n),(b_1,b_2,\ldots,b_n),(c_1,c_2,\ldots,c_n))$, we obtain a triple of length n vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ with $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c} = \boldsymbol{t}$. Let us apply this construction to some $\boldsymbol{f}$ in V_0 to get some $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$. Since $\boldsymbol{\pi'}$ is S_3 symmetric, we have $\boldsymbol{\sigma}(\boldsymbol{a}) = \boldsymbol{\sigma}(\boldsymbol{b}) = \boldsymbol{\sigma}(\boldsymbol{c}) = \boldsymbol{\psi'}$ so $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ lie in W and $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V$. This construction gives an injection from V_0 into V, so $|V| \geq |V_0|$.

By Lemma 3, $|V_0| = \exp(\eta(\boldsymbol{\pi'})n - O_q(\log n))$, so $|V| \geq \exp(\eta(\boldsymbol{\pi'})n - O_q(\log n))$ as desired. □

Lemma 11. *Suppose $p > q$. For any two distinct elements $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$, $(\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c}') \in V$, the $(n+2)\times 6$-matrix over $\mathbb{F}_p$ given by*

$$M = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 & 1 & 1 \\ 1 & 1 & 1/2 & 1/2 & 0 & 0 \\ \boldsymbol{a} & \boldsymbol{a}' & (\boldsymbol{t}-\boldsymbol{b})/2 & (\boldsymbol{t}-\boldsymbol{b}')/2 & \boldsymbol{c} & \boldsymbol{c}' \end{pmatrix}$$

has rank at least 3.

Proof. The first two rows already have rank 2, so we simply must show that the bottom n rows are not all in the span of the first two. If the bottom n rows were in the span of the first two, then modulo p the first column would equal the second, the third column equal the fourth, and the fifth column equal the sixth. Since the entries of the matrix are between 0 and $q-1$, and $p > q$, equality of columns modulo p implies outright equality. This gives $\boldsymbol{a} = \boldsymbol{a}'$, $\boldsymbol{b} = \boldsymbol{b}'$ and $\boldsymbol{c} = \boldsymbol{c}'$, contrary to our assumption that $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$ and $(\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c}')$ are distinct. □

Lemma 12. *When $p > q$ and h is a uniformly random homomorphism of $\mathbb{Z}^{n+2}$ to $\mathbb{F}_p$, the expected cardinality of V'' is at least $\frac{1}{32}e^{-2\sqrt{2\log 2 \log p}} \cdot |W|$.*

Proof. For any $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V$, we want to compute the probability that there exists $s \in S$ such that

$$h(0,1,\boldsymbol{a}) = \tfrac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}) = h(1,0,\boldsymbol{c}) = s. \tag{3}$$

Furthermore, since $h(0,1,\boldsymbol{a})$, $\frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b})$, $h(1,0,\boldsymbol{c})$ always form a (possibly degenerate) arithmetic progression, if any two of these values are equal to s then the third one equals s as well. The vectors $(0,1,\boldsymbol{a})$ and $(1,0,\boldsymbol{c})$ are linearly independent modulo p, so the pair $(h(0,1,\boldsymbol{a}),h(1,0,\boldsymbol{c}))$ is uniformly distributed in $\mathbb{F}_p^2$ and the probability that (3) is satisfied for a fixed $s \in S$ is p^{-2}. Summing over all $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V$ and $s \in S$ we obtain

$$\mathbb{E}(|V'|) = \frac{|V||S|}{p^2}. \tag{4}$$

An element $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V'$ belongs to V'' unless there exists some other $(\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c}') \in V'$ such that one of the equations $\boldsymbol{a} = \boldsymbol{a}'$, $\boldsymbol{b} = \boldsymbol{b}'$, or $\boldsymbol{c} = \boldsymbol{c}'$ holds. In order for any such equation to hold, it must be the case that there is a single element $s \in S$ such that

$$s = h(0,1,\boldsymbol{a}) = h(0,1,\boldsymbol{a}') = \tfrac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}) = \tfrac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}') = h(1,0,\boldsymbol{c}) = h(1,0,\boldsymbol{c}'). \tag{5}$$

By Lemma 11, the six-tuple $(h(0,1,\boldsymbol{a}), h(0,1,\boldsymbol{a}'), \frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}), \frac{1}{2}h(1,1,\boldsymbol{t}-\boldsymbol{b}'), h(1,0,\boldsymbol{c}), h(1,0,\boldsymbol{c}'))$ is uniformly distributed on a subspace of $\mathbb{F}_p^6$ of dimension at least 3. Hence, for any $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}),(\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c}') \in V$

and for a fixed s, the probability that (5) holds is at most p^{-3}. The probability that there exists some s for which (5) holds is thus bounded by $|S|p^{-3}$.

For any $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V$, the number of elements $(\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c}') \in V$ such that $\boldsymbol{a}' = \boldsymbol{a}$ is equal to $|V|/|W|$. (To see this, note that the group S_n acts on V and W by permuting the coordinates of vectors. These actions are compatible with the projection map $V \to W$ defined by $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \mapsto \boldsymbol{a}$. The fibers of this projection map must be equinumerous because the action of S_n on W is transitive.) Thus, for any $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in V$ the probability that $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$ belongs to V' but not V'' because it "collides" with another ordered triple of the form $(\boldsymbol{a},\boldsymbol{b}',\boldsymbol{c}')$ in V' is bounded above by $\frac{|V|}{|W|}|S|p^{-3}$. The analogous counting argument applies to collisions with triples of the form $(\boldsymbol{a}',\boldsymbol{b},\boldsymbol{c}')$ and $(\boldsymbol{a}',\boldsymbol{b}',\boldsymbol{c})$. Summing over $|V|$ choices of $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$, we find that the expected cardinality of $V' \setminus V''$ is bounded above by

$$3|V|\frac{|V|}{|W|}|S|p^{-3} = \frac{3|V|}{p|W|} \cdot \frac{|V||S|}{p^2} < \frac{3}{4} \cdot \mathbb{E}(|V'|).$$

Thus,

$$\mathbb{E}(|V''|) \geq \frac{1}{4}\mathbb{E}(|V'|) = \frac{|V||S|}{4p^2} = \frac{1}{4} \cdot \frac{|V|}{p} \cdot \frac{|S|}{p} > \frac{e^{-2\sqrt{2\log 2\log p}}}{32} \cdot |W|.$$

$\square$

We now prove our main theorem.

Theorem 13. *If n is sufficiently large then there exists a sum-free set in C_q^n with target $\boldsymbol{t}$ whose size is greater than $\theta^n e^{-2\sqrt{2\log 2\log\theta\ n}-O_q(\log n)}$.*

Proof. The random set V'' constructed above is a sum-free set in C_q^n with target $\boldsymbol{t}$ (Lemma 9) and its expected size is greater than $\frac{1}{32}e^{-2\sqrt{2\log 2\log p}} \cdot |W|$ (Lemma 12), because we may take n large enough that $p > q$. Using Lemma 10 we have

$$|W| \geq \exp(\eta(\boldsymbol{\psi}')n - O_q(\log n)) \geq \exp((\log\theta - O_q(1/n))n - O_q(\log n)) \geq \theta^n \exp(-O_q(\log n))$$

for all sufficiently large n. The inequality $|V| \leq |W|^2$ holds because the projection map $V \to W^2$ defined by $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \mapsto (\boldsymbol{a},\boldsymbol{b})$ is one-to-one. This justifies the second inequality in

$$p < 8\frac{|V|}{|W|} \leq 8|W| < 8\exp(\eta(\boldsymbol{\psi}')n) \leq 8\exp((\log\theta + O_q(1/n))n),$$

while the third inequality follows from Lemma 10. Taking logarithms of both sides, we deduce that $\log p < n\log\theta + O_q(1)$, and hence

$$e^{-2\sqrt{2\log 2\log p}} > e^{-2\sqrt{2\log 2(n\log\theta + O_q(1))}} > e^{-2\sqrt{2\log 2\log\theta\ n} - O_q(1/\sqrt{n})}.$$

Hence,

$$\mathbb{E}(|V''|) > \frac{1}{32}e^{-2\sqrt{2\log 2\log\theta\ n} - O_q(1/\sqrt{n})} \cdot |W| \geq \theta^n e^{-2\sqrt{2\log 2\log\theta\ n} - O_q(\log n)}$$

for sufficiently large n. The theorem follows because there must exist at least one choice of h for which the cardinality of the random set V'' is at least as large as its expected value. $\square$

It follows from Roth's theorem that our construction produces sum-free sets $V'' \subseteq V$ of size $\mathbb{E}(|V''|) \leq \mathbb{E}(V') = \frac{V|S|}{p^2} = o(|W|)$ regardless of how we choose S. We do not know if an arbitrary sum-free set contained in V must have size $o(|W|)$, only the trivial bound $|W|$. It would be interesting to improve this situation.

Acknowledgements

We would like to thank Henry Cohn, Jacob Fox, László Miklós Lovász, and Terence Tao for helpful conversations. We would particularly like to thank Jordan Ellenberg for a series of blog posts which drew our attention to this problem and sparked our collaboration. The first author was employed at Microsoft Research New England at the time these results were discovered, and he is grateful to Microsoft Corporation for their support of this research. The second author was supported by Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zurich Foundation. The third author was partially supported by NSF grant DMS-1600223.

References

[1] Noga Alon, Amir Shpilka, and Christopher Umans. On sunflowers and matrix multiplication. *Comput. Complexity*, 22(2):219–243, 2013. 1, 2

[2] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998. 2

[3] Felix A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci.*, 32:331–332, 1946. 5

[4] Arnab Bhattacharya and Ning Xie. Lower bounds for testing triangle freeness in Boolean functions. *Computational Complexity*, 24(1):65–101, 2015. 2

[5] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans. On cap sets and the group-theoretic approach to matrix multiplication. *Discrete Analysis*:3, 2017. 1, 2

[6] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences*, 47:549–595, 1993. 2

[7] Henry Cohn, Robert Kleinberg, Balazs Szegedy, and Christopher Umans. Group-theoretic algorithms for matrix multiplication. In *Proceeding of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 379–388, 2005. 2

[8] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *J. Symbolic Computation*, 9(3):250–280, 1990. 2

[9] Ernie Croot, Vsevolod Lev, and Peter Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Ann. of Math.*, 185(1):331–337, 2017. 1

[10] Yves Edel. Extensions of generalized product caps. *Des. Codes Cryptogr.*, 31(1):5–14, 2004. 1

[11] Michael Elkin. An improved construction of progression-free sets. *Israel J. Math*, 184(1):93–128, 2011. 5

[12] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math.*, 185(1):339–343, 2017. 1, 2

[13] Jacob Fox. A new proof of the graph removal lemma. *Ann. of Math.*, 174(1):561–579, 2011. 3

[14] Jacob Fox and László Miklós Lovász. A tight bound for Green's arithmetic triangle removal lemma in vector spaces. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '17, pages 1612–1617, Philadelphia, PA, USA, 2017. Society for Industrial and Applied Mathematics. 3

[15] Hu Fu and Robert Kleinberg. Improved lower bounds for testing triangle-freeness in boolean functions via fast matrix multiplication. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 17th International Workshop, APPROX 2014, and 18th International Workshop, RANDOM 2014*. 2014. arXiv:1308.1643 [cs.CC]. 2

[16] Ben Green. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geometric & Functional Analysis (GAFA)*, 15(2):340–376, 2005. 2

[17] Pooya Hatami, Sushant Sachdeva, and Madhur Tulsiani. An arithmetic analogue of Fox'ĂŹs triangle removal argument. *Online J. Anal. Comb.*, 11, 2016. 3

[18] Daniel Král', Oriol Serra, and Lluís Vena. A removal lemma for systems of linear equations over finite fields. *Israel Journal of Mathematics*, 187(1):193–207, 2012. 3

[19] Sergey Norin. A distribution on triples with maximum entropy marginal. 2016. arXiv:1608.00243 [math.CO]. 4

[20] Luke Pebody. Proof of a conjecture of Kleinberg-Sawin-Speyer. *Discrete Analysis*:13, 2018. 2, 4

[21] Fedor Petrov. Combinatorial results implied by many zero divisors in a group ring. 2016. arXiv:1606.03256 [math.CO]. 2

[22] Herbert Robbins. A remark on Stirling's formula. *Amer. Math. Monthly*, 62(1):26–29, 1955. 4

AUTHORS

Robert Kleinberg
Department of Computer Science
Cornell University
Ithaca, NY 14853, USA
robert [dot] kleinberg [at] cornell [dot] edu

Will Sawin
ETH Institute for Theoretical Studies
ETH Zurich
8092 Zürich, Switzerland
william [dot] sawin [at] math [dot] ethz [dot] ch

David E Speyer
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109, USA
speyer [at] umich [dot] edu