ISSN 2220-5438

Reprint from

Moscow Journal of Combinatorics and Number Theory

Moscow Journal of Combinatorics and Number Theory Volume 7 • Issue 2 2017

URSS

МФТИ

URSS

Volume 7 • Issue 2 2017

Moscow Journal
of Combinatorics
and Number Theory
2017, vol. 7, iss. 2,
pp. 74–93, [pp. 162–181]

The degree of regularity of the equation $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i + b$

Tomasz Schoen (Poznań), Katarzyna Taczała (Poznań)

Abstract: We confirm a conjecture of Fox and Kleitman from [3] on the maximal degree of regularity of the equation $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i + b$, for $b \in \mathbb{N}$. To establish this result we prove a generalization of a theorem of Eberhard, Green and Manners [2] on the sets with doubling less than 4.

Keywords: Fox-Kleitman conjecture, inverse problems, arithmetic regularity lemma

AMS Subject Classification: 05D10, 11P70

Received: 09.01.2017

1. Introduction

A linear equation is called r-regular if for every r-coloring of $\mathbb{N}$ there exists a monochromatic solution to this equation. The equation is called regular if it is r-regular for every positive integer r. Otherwise, the largest integer such that the equation is r-regular is called the degree of regularity of this equation. Kleitman and Fox studied regularity of various linear equation in [3], in particular

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i + b, \tag{1}$$

where b is a positive integer. It is easy to observe that for every positive integer b one can define a $2n$-coloring $c : \mathbb{N} \to \{0, \dots, 2n-1\}$ by $c(z) = j$ if $\frac{j}{2n} \le \frac{z}{2b} - \lfloor \frac{z}{2b} \rfloor < \frac{j+1}{2n}$ without any monochromatic solution to (1), thus the degree of regularity of this equation is at most $2n-1$. On the other hand, Straus proved in [6] that for every n one can always find an integer $b = b_n$ such that the equation (1) is $\Omega(\log n)$-regular. Kleitman and Fox [3] conjectured that their upper bound is tight.

CONJECTURE 1. For $n \in \mathbb{N}$, there is a positive integer b_n such that the equation

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i + b_n \tag{2}$$

is $(2n-1)$-regular.

The aim of this paper is to confirm Conjecture 1. Our approach will rely on a certain structural theorem.

For a set $A \subseteq \{1, \dots, N\}$ we denote by $2A$ the sumset $A + A = \{a_1 + a_2 : a_1, a_2 \in A\}$ and, more generally, by kA we denote the k-fold sumset $A + \dots + A = \{a_1 + \dots + a_k : a_i \in A \text{ for } 1 \le i \le k\}$. Clearly, we have that $2|A| - 1 \le |2A| \le \frac{|A|(|A|+1)}{2}$, where the lower bound is only achieved for the arithmetic progressions. The structure of the sets with small doubling was intensively studied. The well known $3k-4$ theorem of Freiman from [4] states that if $|2A| \le 3|A| - 4$, then A is contained in an arithmetic progression of length at most $|A + A| - |A| + 1$. In this paper we focus on another structural theorem proved by Eberhard, Green and Manners in [2]. Denote by $D_\delta(A) = \{x : 1_A * 1_{-A}(x) \ge \delta\}$ the set of δ-popular differences, where for any two functions $f, g : \{1, \dots, N\} \to \mathbb{C}$ their convolution is defined by $f * g(n) = \frac{1}{N} \sum_m f(m)g(n-m)$.

THEOREM 1.1. *For every $\varepsilon > 0$ there is some $\delta \gg_\varepsilon 1$ such that the following holds. If $A \subseteq \{1, \dots, N\}$ is a set with $|D_\delta(A)| \le 4|A| - \varepsilon N$, then there is an arithmetic progression $P \subseteq \{1, \dots, N\}$ of length $|P| \gg_\varepsilon N$ such that $|A \cap P| \ge (\frac{1}{2} + \frac{1}{5}\varepsilon)|P|$.*

Following the method of Eberhard, Green and Manners, we will generalize this theorem for sets A which have small k-fold sumset for some $k \ge 2$.

THEOREM 1.2. *For every $\varepsilon > 0$, $k \ge 2$ and every set $A \subseteq \{1, \dots, N\}$ if $|kA| \le k^2|A| - \varepsilon N$, then there is an arithmetic progression $P \subseteq \{1, \dots, N\}$ of length $|P| \gg_{\varepsilon,k} N$ such that $|A \cap P| \ge \frac{1}{k}(1 + \frac{4}{5k}\varepsilon)|P|$. In particular, there exist*

$1 \le j \le k-1$ *and an arithmetic progression* $P \subseteq \{1,\dots,jN\}$ *of length* $|P| \gg_{\varepsilon,k} N$ *such that* $|jA \cap P| \ge (\frac{1}{2} + \frac{1}{j^2+1}\varepsilon)|P|$.

The key ingredient in proof of this theorem is the arithmetic regularity lemma. We give an exact statement of this lemma in the next sections, along with some other auxiliary lemmas from [2].

2. Auxiliary lemmas

In this section we will quote various auxiliary lemmas and definitions mainly from [2]. The crucial result we use is the arithmetic regularity lemma proven by Green and Tao in [5]. The proof of the version we need was also given by Eberhard in [1]. This lemma allows us to decompose a function $f : \{1,\dots,N\} \to [0,1]$ into a structured part and parts small with respect to $\ell_2(N)$ and the Gowers $U^2(N)$ norms. We recall first some definitions.

Definition 1. *Let* $f : \{1,\dots,N\} \to \mathbb{C}$ *be a function. Then we define the Gowers* $U^2(N)$ *norm*

$$||f||_{U^2(N)} = ||f||_{U^2(G)}/||1_{\{1,\dots,N\}}||_{U^2(G)},$$

where $G = \mathbb{Z}/N'\mathbb{Z}$ *for some arbitrary* $N' > 4N$ *and*

$$||f||^4_{U^2(G)} = \mathbb{E}_{x,h_1,h_2\in G} f(x)\overline{f(x+h_1)f(x+h_2)}f(x+h_1+h_2).$$

Above and throughout the paper for any function $f : \{1,\dots,N\} \to \mathbb{C}$ and any subset $K \subseteq \{1,\dots,N\}$ we use the notation $\mathbb{E}_{n\in K} f(x) = \frac{1}{K}\sum_{n\in K} f(x)$.

Definition 2. *Let* $f : \{1,\dots,N\} \to \mathbb{C}$ *be a function. Then we define*

$$||f||_{\ell_2(N)} = \left(\frac{1}{N}\sum_{n\le N}|f(n)|^2\right)^{\frac{1}{2}}.$$

Definition 3. *Suppose that* $\theta \in \mathbb{R}^m$. *Let* $N \ge 1$ *be an integer and let* $A > 0$ *be some real parameter. We say that* θ *is* (A,N)*-irrational if whenever* $q_1,\dots,q_m$ *are integers, not all zero, with* $\sum_i |q_i| \le A$ *then* $||q_1\theta_1 + \dots + q_m\theta_m||_{\mathbb{R}/\mathbb{Z}} \ge A/N$.

Definition 4. *Suppose that* X *is a compact metric abelian group endowed with translation invariant metric* d *and* $f : X \to \mathbb{C}$ *is a function. If there exists a constant* L

such that $|f(x) - f(x')| \le Ld(x, x')$ for all $x, x' \in X$, then the function f is called a Lipschitz function. The Lipschitz constant $||f||_{Lip}$ of the function f is defined by

$$||f||_{Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}.$$

Throughout the paper let X be the product of spaces $\mathbb{Z}/q\mathbb{Z}$, $[0, 1]$, $(\mathbb{R}/\mathbb{Z})^m$ with the suitable product metric of the discrete metric on $\mathbb{Z}/q\mathbb{Z}$ and the Euclidean metrics on $[0, 1]$ and $(\mathbb{R}/\mathbb{Z})^m$. We will also denote $(\mathbb{R}/\mathbb{Z})^m$ by $\mathbb{T}$. We can now formulate the arithmetic regularity lemma.

LEMMA 2.1 (Arithmetic regularity lemma). *Suppose we are given a parameter $\delta > 0$ and a growth function $\mathcal{F} : \mathbb{N} \to \mathbb{R}_+$. Then for any function $f : \{1, \dots, N\} \to [0, 1]$ there is an $M \ll_{\delta,\mathcal{F}} 1$ and a decomposition $f = f_{tor} + f_{unf} + f_{sml}$ into functions taking values in $[-1, 1]$, where $||f_{sml}||_{\ell_2(N)} \le \delta$, $||f_{unf}||_{U^2(N)} \le 1/\mathcal{F}(M)$ and $f_{tor}(n) = F(n(\text{mod } q), n/N, \theta n)$ for some $q, m \le M$ and some function $F : \mathbb{Z}/q\mathbb{Z} \times [0, 1] \times \mathbb{T} \to [0, 1]$ with Lipschitz constant at most M. Furthermore, the element $\theta \in \mathbb{T}$ may be taken to be $(\mathcal{F}(M), N)$-irrational.*

We will also make use of various properties of the functions obtained through such decomposition.

LEMMA 2.2. *Suppose that $\theta \in \mathbb{T}$ is (A, N)-irrational and let $F : \mathbb{T} \to \mathbb{C}$ be a function with Lipschitz constant at most M. Suppose that $P \subseteq \{1, \dots, N\}$ is an arithmetic progression of length at least ηN. Let μ be a normalized Lebesgue measure on $\mathbb{T}$. Then*

$$\left| \mathbb{E}_{n \in P} F(\theta n) - \int F d\mu \right| \le \delta,$$

provided that $A > A_0(M, d, \eta, \delta)$ is large enough.

LEMMA 2.3. *Suppose that $\theta \in \mathbb{T}$ is (A, N)-irrational. Let $q \in \mathbb{N}$ and let $F : \mathbb{Z}/q\mathbb{Z} \times \times [0, 1] \times \mathbb{T} \to [0, 1]$ be a function with Lipschitz constant at most M. Let μ be the product of the uniform measure on $\mathbb{Z}/q\mathbb{Z}$ and the normalized Lebesgue measure on $[0, 1]$ and $\mathbb{T}$. If $\delta > 0$ is arbitrary, then*

$$\left| \mathbb{E}_{n \le N} F(n(\text{mod } q), n/N, \theta n) - \int F d\mu \right| \le \delta,$$

provided that $A > A_0(M, q, d, \delta)$ and $N > N_0(M, q, d, \delta)$ are large enough.

LEMMA 2.4. *Let X be a compact metric abelian group endowed with translation invariant metric d and a translation-invariant probability measure μ. Suppose that $f : X \to \mathbb{C}$ is a function with Lipschitz constant at most K. Let $g : X \to \mathbb{C}$ be any continuous function with $||g||_\infty \leq 1$. Then the convolution $f * g(x) = \int f(y)g(x-y)d\mu(y)$ also has Lipschitz constant at most K.*

Next, we quote some useful lemmas for dealing with functions with small Gowers norm.

LEMMA 2.5. *Suppose that $f : \{1, \dots, N\} \to \mathbb{C}$ is a function. Then $|\mathbb{E}_{n\leq N} f(n)| \ll \ll ||f||_{U^2(N)}$. More generally suppose that $P \subseteq \{1, \dots, N\}$ is an arithmetic progression of length at least ηN. Then $|\mathbb{E}_{n\in P} f(n)|| \ll \eta^{-1}||f||_{U^2(N)}$.*

LEMMA 2.6. *Let $f, \widetilde{f}, g : \{1, \dots, N\} \to [-1, 1]$ be functions such that $||\widetilde{f} - f||_{U^2(N)} \leq \leq \delta$. Then $|\widetilde{f} * g(d) - f * g(d)| \leq 4\delta^{1/2}$ for all except at most $40\delta N$ values of d.*

Finally, we also need a lemma to bound the values of convolution of functions.

LEMMA 2.7. *Let $P, P' \subseteq \{1, \dots, N\}$ be arithmetic progressions with the same length. Let $f : P \to \mathbb{C}$ and $g : P' \to \mathbb{C}$ be two functions. Suppose that both are bounded pointwise by 1 and that either $\mathbb{E}_{n\in P}|f(n)| \leq \eta$ or $\mathbb{E}_{n\in P'}|g(n)| \leq \eta$. Then $||f * g||_\infty \leq \eta|P|/N$.*

3. Proof of theorem 1.2

Our argument closely follows the proof of theorem 1.1 from [2]. In the course of the proof we will always assume that N is large enough depending on ε, k. For small N the theorem is true trivially by taking any two elements of A to be an arithmetic progression P (or one if $|A| = 1$). We will use the induction on k. For $k = 2$ the lemma follows from Theorem 1.1. Let $k \geq 3$. We assume that the theorem is true for $k - 1$ and that $A \subseteq \{1, \dots, N\}$ is a set such that $|kA| \leq k^2|A| - \varepsilon N$. Let $\widetilde{\mathcal{F}} : \mathbb{N} \to \mathbb{R}_+$ be a growth function depending on ε and k to be chosen later. Let $\widetilde{\varepsilon} = \min(\varepsilon, \frac{1}{100k^2})$. Then we can apply the arithmetic regularity lemma to the function 1_A to obtain a parameter $\widetilde{M} \ll_{\varepsilon,k,\widetilde{\mathcal{F}}} 1$ and a decomposition

$$1_A = f_{tor} + \widetilde{f}_{sml} + f_{unf},$$

where $||\widetilde{f}_{sml}||_{\ell_2(N)} \leq \widetilde{\varepsilon}^{4k(k+2)}$, $||f_{unf}||_{U^2(N)} \leq \frac{1}{\widetilde{\mathcal{F}}(\widetilde{M})}$ and

$$f_{tor} = F(n \text{ (mod } q), \frac{n}{N}, \theta n)$$

for some $F : \mathbb{Z}/q\mathbb{Z} \times [0, 1] \times (\mathbb{R}/\mathbb{Z})^m \to [0, 1]$ such that $q, m, ||F||_{Lip} \leq \widetilde{M}$ and for some $(\widetilde{\mathcal{F}}(\widetilde{M}), N)$-irrational $\theta \in (\mathbb{R}/\mathbb{Z})^m$.

We will study the density of A on some arithmetic progressions in $\{1, \ldots, N\}$. To do this we define a parameter $M = \lceil \widetilde{\varepsilon}^{-4k(k+2)} \widetilde{M} \rceil$ and progressions:

$$I_{a,i} = \left\{ n \in \left(\frac{(i-1)N}{M}, \frac{iN}{M} \right] : n \equiv a \text{ (mod } q) \right\}$$

for $a \in \mathbb{Z}/q\mathbb{Z}$ and $i \in \{1, \ldots, M\}$. We proceed as in [2] by defining functions $F_{a,i} : \mathbb{T} \to [0, 1]$ in the following way: $F_{a,i}(x) = F(a, i/M, x)$ for $a \in \mathbb{Z}/q\mathbb{Z}$ and $i \in \{1, \ldots, M\}$. Now we can replace function f_{tor} by a function $f_{str} : \{1, \ldots, N\} \to$ $\to [0, 1]$ such that

$$f_{str}(x) = \sum_{a(\text{mod } q)} \sum_{i=1}^{M} 1_{I_{a,i}}(x) F_{a,i}(\theta n).$$

Because F is $\widetilde{M}$-Lipschitz, so is $F_{a,i}$ and therefore f_{str} differs from f_{tor} by at most $\widetilde{\varepsilon}^{4k(k+2)}$ as

$$|f_{str}(n) - f_{tor}(n)| = \left| F(a, \frac{i}{M}, \theta n) - F(a, \frac{n}{N}, \theta n) \right| \leq \widetilde{M} \left| \frac{i}{M} - \frac{n}{N} \right| \leq \widetilde{M} \frac{1}{M} \leq \widetilde{\varepsilon}^{4k(k+2)}$$

for $n \in I_{a,i}$. We can deal with this difference by defining a function $f_{sml} = \widetilde{f}_{sml} +$ $+ f_{tor} - f_{str}$, which still has small $\ell_2(N)$ norm $||f_{sml}||_{\ell_2(N)} \leq 2\widetilde{\varepsilon}^{4k(k+2)}$. Observe that we can always choose the function $\widetilde{\mathcal{F}}$ in such a way that $\widetilde{\mathcal{F}}(\widetilde{M}) \geq \mathcal{F}(M)$ for some arbitrary growth function $\mathcal{F}$ depending on ε, k. Therefore, we obtain a decomposition

$$1_A = f_{str} + f_{unf} + f_{sml},$$

where $||f_{sml}||_{\ell_2(N)} \leq 2\widetilde{\varepsilon}^{4k(k+2)}$, $||f_{unf}||_{U^2(N)} \leq \frac{1}{\mathcal{F}(M)}$ and

$$f_{str}(n) = \sum_{a(\text{mod} q)} \sum_{i=1}^{\widetilde{M}} 1_{I_{a,i}}(n) F_{a,i}(\theta n)$$

with θ being $(\mathcal{F}(M), N)$-irrational. We denote by $\alpha(a, i)$ the density of A on $I_{a,i}$ and put

$$E = \{(a, i) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \dots, M\} : \mathbb{E}_{n \in I_{a,i}} |f_{sml}(n)| > \widetilde{\varepsilon}^{2k(k+2)}\}.$$

We quote here some lemmas from [2].

LEMMA 3.1. $|E| \leq \widetilde{\varepsilon}^{2k(k+2)-1} qM$.

LEMMA 3.2. *For all $(a, i) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \dots, M\}$ outside E we have $\int_{\mathbb{T}} F_{a,i} \geq \alpha(a, i) - \widetilde{\varepsilon}^4$ provided that $\mathcal{F}(M)$ is large enough depending on ε, k, M and N is large enough depending on q, M.*

LEMMA 3.3. *Let $\eta > 0$. If $\mathcal{F}$ grows sufficiently rapidly depending on η, then the following is true. If $F : \mathbb{T} \to [0, 1]$ is M-Lipschitz, θ is $(\mathcal{F}(M), N)$-irrational and $I \subseteq \{1, \dots, N\}$ is any progression of length at least $\frac{N}{M^2}$, then the proportion of $n \in I$ such that $F(\theta n) > \eta$ is at least $\mu(\{x \in \mathbb{T} : F(x) > 2\eta\}) - \eta$.*

We will also need the following result of Tao [7].

LEMMA 3.4. *If $S_1, S_2 \subseteq \mathbb{T}$ are open and $0 \leq t \leq \min(\mu(S_1), \mu(S_2))$, then*

$$\int_{\mathbb{T}} \min(1_{S_1} * 1_{S_2}, t) d\mu \geq t \min(\mu(S_1) + \mu(S_2) - t, 1).$$

Next lemma is a generalization of Lemma 4.6 from [2].

LEMMA 3.5. *Let $k \geq 2$ and $0 < \eta < (3k-4)^{-k(k+1)}$. Suppose that $F_1, F_2, \dots, F_k : \mathbb{T} \to [0, 1]$ are M-Lipschitz functions such that $\int F_1, \dots, \int F_k \geq 3\eta^{1/(k(k+1))}$. Then the measure of the set of x for which $F_1 * \dots * F_k(x) > \eta$ is at least*

$$\min\left(\int F_1 + \dots + \int F_k, 1\right) - (3k-2)\eta^{1/(k(k+1))}.$$

PROOF. We will use the induction on k. For $k = 2$ our lemma is just Lemma 4.6 from [2]. We assume that the lemma is true for k and that the functions $F_1, \dots, F_{k+1}$ satisfy the conditions of the lemma for $k+1$. Let $\widetilde{\eta} = \eta^{k/(k+2)}$ and denote the convolution $F_1 * \dots * F_k$ by F and $\frac{1}{(k+1)(k+2)}$ by β. We define an open set $S = \{x \in \mathbb{T} : F(x) > \eta^{\beta + k/(k+2)}\}$. Since $\eta^{\beta + k/(k+2)} < \widetilde{\eta}$ we can estimate the measure of the set S from below by the measure of the set

$\{x \in \mathbb{T} : F(x) \geq \widetilde{\eta}\}$. By assumption $\int F_1, \ldots, \int F_k \geq 3\eta^\beta = 3\widetilde{\eta}^{1/k(k+1)}$ and $\widetilde{\eta} = \eta^{k/(k+2)} < (3k-1)^{-(k+1)(k+2)k/(k+2)} < (3k-1)^{-k(k+1)} < (3k-4)^{-k(k+1)}$. By the inductive hypothesis

$$\mu(S) \geq \mu(\{x \in \mathbb{T} : F(x) \geq \widetilde{\eta}\}) \geq \min\left(\int F_1 + \ldots + \int F_k, 1\right) - (3k-2)\widetilde{\eta}^{1/k(k+1)}$$
$$= \min\left(\int F_1 + \ldots + \int F_k, 1\right) - (3k-2)\eta^\beta.$$

In fact we can show that $\mu(S) \geq \eta^\beta$. Indeed, in view of $\eta < (3k-1)^{-(k+1)(k+2)}$ and $\int F_i \geq 3\eta^\beta$ we have

$$\min\left(\int F_1 + \ldots + \int F_k, 1\right) \geq \min\left(3k\eta^\beta, (3k-1)\eta^\beta\right) \geq (3k-1)\eta^\beta.$$

Put $T = \{x \in \mathbb{T} : F_{k+1}(x) > \eta^{1/(k+2)}\}$ and denote by T^c the complement of T in $\mathbb{T}$. Since

$$\int_{\mathbb{T}} F_{k+1} = \int_T F_{k+1} + \int_{T^c} F_{k+1} \leq \int_T 1 + \int_{T^c} \eta^{1/(k+2)}$$
$$\leq \mu(T) + \eta^{1/(k+2)}\mu(T^c) \leq \mu(T) + \eta^{1/(k+2)}$$

we can estimate the measure of T from below by

$$\mu(T) \geq \int F_{k+1} - \eta^{1/(k+2)} \geq 3\eta^\beta - \eta^\beta \geq \eta^\beta.$$

Further, by assumption, the functions $F_1, \ldots, F_{k+1}$ are M-Lipschitz, hence by Lemma 2.4 so is the convolution F. Consequently, F and F_{k+1} are continuous functions, so that the sets S and T are open. Therefore we can use the Lemma 3.4 with $t = \eta^\beta$. We have

$$\int_{\mathbb{T}} \min(1_S * 1_T, \eta^\beta)d\mu \geq \eta^\beta \min(\mu(S) + \mu(T) - \eta^\beta, 1),$$

so

$$\int_{\mathbb{T}} \frac{\min(1_S * 1_T, \eta^\beta)}{\eta^\beta} d\mu \geq \min(\mu(S) + \mu(T) - \eta^\beta, 1).$$

We estimate the right-hand side using the integrals of functions $F_1, \dots, F_{k+1}$

$$\min\left(\min\left(\int F_1 + \ldots + \int F_k, 1\right) - (3k-2)\eta^\beta + \int F_{k+1} - \eta^\beta - \eta^\beta, 1\right)$$
$$\geq \min\left(\min\left(\int F_1 + \ldots + \int F_k, 1\right) + \int F_{k+1}, 1\right) - 3k\eta^\beta$$
$$\geq \min\left(\int F_1 + \ldots + \int F_k + \int F_{k+1}, 1\right) - 3k\eta^\beta.$$

To estimate the left-hand side let us define $X = \{x \in \mathbb{T} : 1_S * 1_T(x) \geq \eta^{1/(k+2)-\beta}\}$. We have

$$\int_{\mathbb{T}} \frac{\min(1_S * 1_T, \eta^b)}{\eta^b} d\mu \leq \int_X \frac{\min(1_S * 1_T, \eta^b)}{\eta^b} d\mu + \int_{X^c} \frac{\min(1_S * 1_T, \eta^b)}{\eta^b} d\mu$$
$$\leq \int_X \frac{\eta^b}{\eta^b} d\mu + \int_{X^c} \frac{\eta^{1/(k+2)-b}}{\eta^b} d\mu$$
$$\leq \mu(X) + \int_{X^c} \eta^{1/(k+2)-2b} d\mu \leq \mu(X) + \eta^b$$

where the last inequality follows from the fact that $\frac{1}{k+2} - 2\beta \geq \beta$ for $k \geq 2$. Observe that for $x \in X$ we have

$$F * F_{k+1}(x) \geq \eta^{\beta + k/(k+2)} 1_S * \eta^{1/(k+2)} 1_T(x) \geq \eta^{\beta+k/(k+2)} \cdot \eta^{1/(k+2)} \cdot \eta^{1/(k+2)-\beta} = \eta$$

where the first inequality follows from the definitions of the sets S and T and the second inequality follows from the definition of the set X. Comparing these estimates we get

$$\mu(\{x \in \mathbb{T} : F * F_{k+1}(x) \geq \eta\}) \geq \min\left(\int F_1 + \ldots + \int F_k + \int F_{k+1}, 1\right) - (3k+1)\eta^\beta,$$

which completes the proof. □

With this lemma we will be able to link the size of kA with the densities $\alpha(a, i)$.

LEMMA 3.6. *Let* $(a_1, i_1), \dots, (a_k, i_k) \notin E$ *and* $\alpha(a_1, i_1), \dots, \alpha(a_k, i_k) \geq 4\widetilde{\varepsilon}^2$. *Then*

$$|kA \cap I_{a_1+\ldots+a_k, i_1+\ldots+i_k-j}| \geq \frac{N}{qM} \min(\alpha(a_1, i_1) + \ldots + \alpha(a_k, i_k), 1) - 7k\widetilde{\varepsilon}^2 \frac{N}{qM}$$

for any $0 \leq j \leq k-1$.

PROOF. Recall that

$$I_{a_1+\ldots+a_k, i_1+\ldots+i_k-j} = \left\{ n \in \left(\frac{(\sum_{h=1}^{k} i_h - j - 1)N}{M}, \frac{(\sum_{h=1}^{k} i_h - j)N}{M} \right] : n \equiv \sum_{h=1}^{k} a_h \pmod{q} \right\}.$$

Let us notice first that in view of Lemma 3.2 it is sufficient to prove that

$$|kA \cap I_{a_1+\ldots+a_k, i_1+\ldots+i_k-j}| \geq \frac{N}{qM} \min \left(\int F_{a_1,i_1} + \ldots + \int F_{a_k,i_k}, 1 \right) - 6k\widetilde{\varepsilon}^2 \frac{N}{qM}$$

for $(a_1, i_1), \ldots, (a_k, i_k) \notin E$ and $\int F_{a_1,i_1}, \ldots, \int F_{a_k,i_k} \geq 3\widetilde{\varepsilon}^2$. To show that $d \in kA \cap I_{a_1+\ldots+a_k, i_1+\ldots+i_k-j}$ we have to verify that $1_{A|I_{a_1,i_1}} * \ldots * 1_{A|I_{a_k,i_k}}(d) > 0$ for $d \in I_{a_1+\ldots+a_k, i_1+\ldots+i_k-j}$. We will write F_h for F_{a_h,i_h} for $1 \leq h \leq k$. We will focus on the case $j = 0$. Now, let $d_h \in I_{a_1+\ldots+a_h, i_1+\ldots+i_h}$ satisfy $d_h \leq \frac{i_1+\ldots+i_h-2h\widetilde{\varepsilon}^2}{M} N$ for some $h \leq k$. We will use the induction on h. As in [2] one can prove that

$$f_{str|I_{a_1,i_1}} * f_{str|I_{a_2,i_2}}(d_2) \geq \frac{\widetilde{\varepsilon}^2}{qM}(F_1 * F_2(\theta d_2) - \frac{6}{4 \cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}).$$

Let us denote $F_1 * \ldots * F_{h-1}$ by F and assume that

$$f_{str|I_{a_1,i_1}} * \ldots * f_{str|I_{a_{h-1},i_{h-1}}}(d_{h-1}) \geq \left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{h-2} (F(\theta d_{h-1}) - \frac{6^{h-1}}{4 \cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}).$$

Furthermore, we define

$$X = \left\{ n \in I_{a_1+\ldots+a_{h-1}, i_1+\ldots+i_{h-1}} : \frac{\sum_{1 \leq j \leq h-1} i_j - 2h\widetilde{\varepsilon}^2}{M} N \leq n \leq \frac{\sum_{1 \leq j \leq h-1} i_j - 2(h-1)\widetilde{\varepsilon}^2}{M} N \right\}.$$

Note that $d_h - n \in I_{a_h,i_h}$ for every $n \in X$ and $\frac{\widetilde{\varepsilon}^2 N}{qM} \leq |X| \leq 3\frac{\widetilde{\varepsilon}^2 N}{qM}$ provided that $N(q, M)$ is large enough. Therefore

$$f_{str|I_{a_1,i_1}} * \ldots * f_{str|I_{a_h,i_h}}(d_h) \geq \frac{1}{N} \sum_{n \in X} f_{str|I_{a_1,i_1}} * \ldots * f_{str|I_{a_{h-1},i_{h-1}}}(n) f_{str|I_{a_h,i_h}}(d_h - n)$$

$$\geq \quad \frac{1}{N} \sum_{n \in X} \left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{h-2} \left(F(\theta n) - \frac{6^{h-1}}{4 \cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}\right) F_h(\theta(d_h - n))$$

$$
\begin{aligned}
&\geq \frac{1}{N}\left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{h-2}\left[\frac{|X|}{|X|}\sum_{n\in X}F(\theta n)F_h(\theta(d_h-n))-\sum_{n\in X}\frac{6^{h-1}}{4\cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}F_h(\theta(d_h-n))\right]\\
&\geq \frac{1}{N}\left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{h-2}\left[|X|\left(\int F(\theta n)F_h(\theta(d_h-n))d\mu-\frac{6^{h-1}}{4\cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}-\frac{6^{h-1}}{4\cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}\right)\right]\\
&\geq \frac{1}{N}\left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{h-2}\left[\frac{\widetilde{\varepsilon}^2 N}{qM}F*F_h(\theta d_h)-2\cdot 3\cdot\frac{6^{h-1}}{4\cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}\frac{\widetilde{\varepsilon}^2 N}{qM}\right]\\
&\geq \left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{h-1}\left[F*F_h(\theta d_h)-\frac{6^h}{4\cdot 6^k}\widetilde{\varepsilon}^{2k(k+1)}\right].
\end{aligned}
$$

In the estimate above the second inequality follows from the inductive hypothesis. In the fourth inequality we used Lemma 2.4 and the fact that the product of two M-Lipschitz functions bounded pointwise by 1 is $2M$-Lipschitz, hence we can use Lemma 2.2 provided $\mathcal{F}(M)$ is large enough depending on $M, m, \widetilde{\varepsilon}, h, k$. Repeating this argument until $h = k$, we get

$$
f_{str|I_{a_1,i_1}} * \ldots * f_{str|I_{a_k,i_k}}(d_k) \geq \left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{k-1}\left[F_1 * \ldots * F_k(\theta d_k) - \frac{1}{4}\widetilde{\varepsilon}^{2k(k+1)}\right]
$$

for at least $\frac{N}{qM}(1-2(k+1)\widetilde{\varepsilon}^2)$ values of d_k. But if $\mathcal{F}$ grows sufficiently rapidly depending on ε, M, then by Lemma 3.3 the proportion of $d \in I_{a_1+\ldots+a_k, i_1+\ldots+i_k}$ such that $F_1 * \ldots * F_k(\theta d) > \frac{1}{2}\widetilde{\varepsilon}^{2k(k+1)}$ is at least $\mu(Y) - \frac{1}{2}\widetilde{\varepsilon}^{2k(k+1)}$, where $Y = \{x \in \mathbb{T} : F_1 * \ldots * F_k(x) > \widetilde{\varepsilon}^{2k(k+1)}\}$. We can estimate the measure of Y using Lemma 3.5 with $\eta = \widetilde{\varepsilon}^{2k(k+1)}$. Recall that $\widetilde{\varepsilon} = \min(\varepsilon, \frac{1}{100k^2})$. Then

$$
0 < \eta \leq \left(\frac{1}{100k^2}\right)^{2k(k+1)} = \left(\frac{1}{10^4 k^4}\right)^{k(k+1)} < (3k-4)^{-k(k+1)}
$$

and $\int F_1, \ldots, \int F_k \geq 3\widetilde{\varepsilon}^2 = 3\eta^{1/(k(k+1))}$. Therefore by Lemma 4.6

$$
\mu(Y) \geq \min\left(\int F_1 + \ldots + \int F_k, 1\right) - (3k-2)\widetilde{\varepsilon}^2.
$$

This proves that

$$
f_{str|I_{a_1,i_1}} * \ldots * f_{str|I_{a_k,i_k}}(d_k) \geq \frac{1}{4}\left(\frac{\widetilde{\varepsilon}^2}{qM}\right)^{k-1}\widetilde{\varepsilon}^{2k(k+1)} = \frac{1}{4}\left(\frac{1}{qM}\right)^{k-1}\widetilde{\varepsilon}^{2k^2+4k-2}
$$

for at least

$$\frac{N}{qM}\min\Big(\int F_1 + \ldots + \int F_k, 1\Big) - 5k\frac{N}{qM}\tilde{\varepsilon}^2$$

values of $d \in I_{a_1+\ldots+a_k, i_1+\ldots+i_k}$.

Next we will bound the contribution of the functions f_{sml} and f_{unf}. Let us first recall that by Lemma 3.1 for $1 \leq h \leq k$ $\mathbb{E}_{n\in I_{a_h,i_h}}|f_{sml|I_{a_h,i_h}}(n)| \leq \tilde{\varepsilon}^{2k(k+2)}$ since $(a_h, i_h) \notin E$. For $2 \leq j \leq k$ let $\mathcal{B}_j = \{b_1, b_2, \ldots, b_j\}$ be any subset of size j of the pairs $\mathcal{A} = \{(a_1, i_1), \ldots, (a_k, i_k)\}$. We will show, by induction, the analogue of Lemma 2.7

$$|f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_j}}(n)| \leq \frac{1}{100k^{2j}}\frac{1}{(qM)^{j-1}}\tilde{\varepsilon}^{2k^2+4k-2}.$$

For $j = 2$, $\mathcal{B}_2 = \{b_1, b_2\}$ and any n we have

$$\begin{aligned}|f_{sml|I_{b_1}} * f_{sml|I_{b_2}}(n)| &\leq \frac{1}{N}\sum_d |f_{sml|I_{b_1}}(n-d)||f_{sml|I_{b_2}}(d)| \leq \frac{|I_{b_2}|}{N}\mathbb{E}_{d\in I_{b_2}}|f_{sml|I_{b_2}}(d)| \\ &\leq \frac{1}{N}\frac{2N}{qM}\tilde{\varepsilon}^{2k^2+4k} \leq \frac{2}{qM}\frac{1}{10^4k^4}\tilde{\varepsilon}^{2k^2+4k-2} \leq \frac{1}{100k^4}\frac{1}{qM}\tilde{\varepsilon}^{2k^2+4k-2}.\end{aligned}$$

Now, we assume that $|f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_{j-1}}}(n)| \leq \frac{1}{100k^{2j-2}}\frac{1}{(qM)^{j-2}}\tilde{\varepsilon}^{2k^2+4k-2}$ for any $\mathcal{B}'_j = \{b'_1, \ldots, b'_{j-1}\} \subseteq \mathcal{A}$ and any n. Then for $\mathcal{B}_j = \{b_1, \ldots, b_j\} \subseteq \mathcal{A}$ and any n we have

$$\begin{aligned}|f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_j}}(n)| &\leq \frac{1}{N}\sum_d |f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_{j-1}}}(n-d)||f_{sml|I_{b_j}}(d)| \\ &\leq \frac{1}{100k^{2j-2}}\frac{1}{(qM)^{j-2}}\tilde{\varepsilon}^{2k^2+4k-2}\frac{|I_{b_j}|}{N}\mathbb{E}_{d\in I_{b_j}}|f_{sml|I_{b_j}}(d)| \\ &\leq \frac{1}{100k^{2j}}\frac{1}{(qM)^{j-1}}\tilde{\varepsilon}^{2k^2+4k-2}.\end{aligned}$$

Analogously, we can show that for $c := k^{1/(k-1)}$, $1 \leq j \leq k-1$, $\mathcal{C}_j = \{c_1, \ldots, c_j\} \subseteq \mathcal{A}$ and any n

$$|f_{str|I_{c_1}} * \ldots * |f_{str|I_{c_j}}(n)| \leq \Big(\frac{c}{qM}\Big)^{h-1}$$

as $\mathbb{E}_{n\in I_{c_j}}|f_{str|I_{c_j}}(n)| \leq 1$ and $|I_{c_j}| \leq c\frac{N}{qM}$ for N sufficiently large. Combining these estminates for $\mathcal{B}_j$ and $\mathcal{C}_h$ such that $2 \leq j \leq k$, $h = k-j$, $\mathcal{B}_j \cup \mathcal{C}_h = \mathcal{A}$ and

$\mathcal{B}_j \cap \mathcal{C}_h = \varnothing$ we get

$$|f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_j}} \; * \; f_{str|I_{c_1}} * \ldots * f_{str|I_{c_h}}(n)|$$
$$\leq \frac{1}{N}\sum_d |f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_j}}(d)||f_{str|I_{c_1}} * \ldots * f_{str|I_{c_h}}(n-d)|$$
$$\leq \frac{1}{N}\left(\frac{c}{qM}\right)^{h-1} \sum_{d \in I_{b_1}+\ldots+I_{b_j}} |f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_j}}(d)|$$
$$\leq \frac{1}{N}\left(\frac{c}{qM}\right)^{h-1} \frac{1}{100k^{2j}} \frac{1}{(qM)^{j-1}} \widetilde{\varepsilon}^{2k^2+4k-2} |I_{b_1}+\ldots+I_{b_j}|$$
$$\leq \frac{1}{N}\left(\frac{c}{qM}\right)^{h-1} \frac{1}{100k^{2j}} \frac{1}{(qM)^{j-1}} \widetilde{\varepsilon}^{2k^2+4k-2} jc\frac{N}{qM}$$
$$\leq c^h \left(\frac{1}{qM}\right)^{k-1} \frac{j}{100k^{2j}} \widetilde{\varepsilon}^{2k^2+4k-2}.$$

In the similar way, we see that for any $b \in \mathcal{A}$ and $\mathcal{C}_{k-1} = \mathcal{A} \setminus \{b\}$ and any n

$$|f_{sml|I_b} * f_{str|I_{c_1}} * \ldots * f_{str|I_{c_{k-1}}}(n)| \leq \frac{1}{N}\sum_d |f_{sml|I_b}(d)||f_{str|I_{c_1}} * \ldots * f_{str|I_{c_{k-1}}}(n-d)|$$
$$\leq \frac{1}{N}\left(\frac{c}{qM}\right)^{k-2} |I_b| \mathbb{E}_{d \in I_b} |f_{sml|I_b}(d)| \leq \left(\frac{c}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2} \frac{1}{100k^4}.$$

Finally, we have

$$(f_{str|I_{a_1,i_1}} + f_{sml|I_{a_1,i_1}}) * \ldots * (f_{str|I_{a_k,i_k}} + f_{sml|I_{a_k,i_k}})(d)$$
$$= f_{str|I_{a_1,i_1}} * \ldots * f_{str|I_{a_k,i_k}}(d) + \sum_{j=1}^{k} \sum_{(\mathcal{B}_j, \mathcal{C}_{k-j})} f_{sml|I_{b_1}} * \ldots * f_{sml|I_{b_j}} * f_{str|I_{c_1}} * \ldots * f_{str|I_{c_{k-j}}}(d)$$
$$\geq \frac{1}{4}\left(\frac{1}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2} - \left(k\left(\frac{c}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2} \frac{1}{100k^4} + \sum_{j=2}^{k} \binom{k}{j} c^{k-j} \left(\frac{1}{qM}\right)^{k-1} \frac{j}{100k^{2j}}\right)$$
$$\geq \frac{1}{4}\left(\frac{1}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2} - \left[\left(\frac{c}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2} \left(\frac{1}{100k^3} + \frac{1}{100k}\right)\right]$$
$$\geq \frac{1}{4}\left(\frac{1}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2} - \frac{1}{50}\left(\frac{1}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2} \geq \frac{1}{5}\left(\frac{1}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2}$$

for at least

$$\frac{N}{qM} \min(\int F_1 + \ldots + \int F_k, 1) - 5k\frac{N}{qM}\widetilde{\varepsilon}^2$$

values of $d \in I_{a_1+\ldots+a_k, i_1+\ldots+i_k}$.
Since $\|1_{A|I_{a_j,i_j}} - (f_{str|I_{a_j,i_j}} + f_{sml|I_{a_j,i_j}})\|_{U^2(N)} = \|f_{unf|I_{a_j,i_j}}\|_{U^2(N)} \leq \frac{1}{\mathcal{F}(M)}$ we can use the Lemma 2.6 to include the contribution of f_{unf} provided that $\mathcal{F}$ grows sufficiently rapidly depending on $k, \widetilde{\varepsilon}, q, M$. Then

$$1_{A|I_{a_1,i_1}} * \ldots * 1_{A|I_{a_k,i_k}}(d) \geq \frac{1}{8}\left(\frac{1}{qM}\right)^{k-1} \widetilde{\varepsilon}^{2k^2+4k-2}$$

for at least

$$\frac{N}{qM}\min\left(\int F_1 + \ldots + \int F_k, 1\right) - 6k\frac{N}{qM}\widetilde{\varepsilon}^2$$

values of $d \in I_{a_1+\ldots+a_k, i_1+\ldots+i_k}$. □

To prove the next lemma we will need the Brunn-Minkowski theorem.

THEOREM 3.1. *Let X, Y be open subsets of $\mathbb{R}^m$. Then we have $\lambda(X+Y)^{1/m} \geq$ $\geq \lambda(X)^{1/m} + \lambda(Y)^{1/m}$, where λ is the Lebesgue measure.*

LEMMA 3.7. *For a given function $\alpha : \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, M\} \to [0, 1]$ and $(x, y) \in$ $\in \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, kM\}$ we define $\widetilde{\alpha}(x, y) = \max(\alpha(a_1, i_1) + \ldots + \alpha(a_k, i_k))$, where the maximum is taken over all pairs $(a_1, i_1), \ldots, (a_k, i_k) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, M\}$ such that $a_1 + \ldots + a_k = x$ and $i_1 + \ldots + i_k - j = y$ for some $j \in \{0, \ldots, k-1\}$. Then*

$$\sum_{x,y} \widetilde{\alpha}(x, y) \geq k^2 \sum_{a,i} \alpha(a, i)$$

PROOF. We define the following open set $X \subset \mathbb{Z}/q\mathbb{Z} \times \mathbb{R}^2$

$$X = \bigcup_{(a,i)\in\mathbb{Z}/q\mathbb{Z}\times\{1,\ldots,M\}} \{a\} \times (i-1, i) \times (0, \alpha(a, i)).$$

Let us observe that

$$X + X = \bigcup_{(a,i),(a',i')} \{a + a'\} \times (i + i' - 2, i + i') \times (0, \alpha(a, i) + \alpha(a', i'))$$

hence

$$kX = \bigcup_{(a_1,i_1),\ldots,(a_k,i_k)} \left\{\sum_{j=1}^{k} a_j\right\} \times \left(\sum_{j=1}^{k} i_j - k, \sum_{j=1}^{k} i_j\right) \times \left(0, \sum_{j=1}^{k} \alpha(a_j, i_j)\right)$$

$$= \bigcup_{(a_1,i_1),\dots,(a_k,i_k)} \left(\left\{ \sum_{j=1}^{k} a_j \right\} \times \left(\sum_{j=1}^{k} i_j - k, \sum_{j=1}^{k} i_j - (k-1) \right) \times \left(0, \sum_{j=1}^{k} \alpha(a_j, i_j) \right) \right) \cup \dots$$

$$\cup \left(\left\{ \sum_{j=1}^{k} a_j \right\} \times \left(\sum_{j=1}^{k} i_j - 1, \sum_{j=1}^{k} i_j \right) \times \left(0, \sum_{j=1}^{k} \alpha(a_j, i_j) \right) \right)$$

$$= \bigcup_{(x,y)\in \mathbb{Z}/q\mathbb{Z}\times\{1,\dots,kM\}} \{x\} \times (y-1, y) \times (0, \widetilde{\alpha}(x,y)),$$

if we ignore the sets of measure zero in the second equality.

Now, let ν be a a product of counting measure on $\mathbb{Z}/q\mathbb{Z}$ and Lebesgue measure λ on $\mathbb{R}^2$. Then $\nu(X) = \sum_{a,i} \alpha(a, i)$ and $\nu(kX) = \sum_{x,y} \widetilde{\alpha}(x, y)$. We want to show that $\nu(kX) \geq k^2\nu(X)$. If $q = 1$, then we simply apply the Brunn-Minkowski inequality for $m = 2$ and the induction on k. Otherwise, we use fibres X_a of X over $a \in \mathbb{Z}/q\mathbb{Z}$. These fibres are open subsets of $\mathbb{R}^2$. Let a_* be an element of $\mathbb{Z}/q\mathbb{Z}$ such that $\lambda(X_{a_*}) = \sum_i \alpha(a_*, i) = \max_a \lambda(X_a)$. Then for any a such that $X_a \neq \varnothing$ Brun-Minkowski inequality implies that

$$\lambda(X_a + X_{a_*}) \geq (\sqrt{\lambda(X_a)} + \sqrt{\lambda(X_{a_*})})^2 \geq 4\lambda(X_a).$$

We then use induction. Let us assume that $\lambda(X_a + (j-1)X_{a_*}) \geq j^2\lambda(X_a)$. Then

$$\begin{aligned}\lambda(X_a + jX_{a_*}) &\geq (\sqrt{\lambda(X_a + (j-1)X_{a_*})} + \sqrt{\lambda(X_{a_*})})^2 \\ &= \lambda(X_a + (j-1)X_{a_*}) + 2\sqrt{\lambda(X_a + (j-1)X_{a_*})\lambda(X_{a_*})} + \lambda(X_{a_*}) \\ &\geq j^2\lambda(X_a) + 2j\lambda(X_a) + \lambda(X_a) = (j+1)^2\lambda(X_a).\end{aligned}$$

By induction we conclude that $\lambda(X_a + (k-1)X_{a_*}) \geq k^2\lambda(X_a)$ And since $X_a + + (k-1)X_{a_*}$ are disjoint as a ranges over $\mathbb{Z}/q\mathbb{Z}$, we have

$$\nu(kX) \geq \sum_{a:X_a\neq\varnothing} \lambda(X_a + (k-1)X_{a_*}) \geq \sum_{a:X_a\neq\varnothing} k^2\lambda(X_a) = k^2\nu(X)$$

and finally

$$\sum_{x,y} \widetilde{\alpha}(x, y) = \nu(kX) \geq k^2\nu(X) = k^2 \sum_{a,i} \alpha(a, i). \qquad \square$$

LEMMA 3.8. *For a given function* $\alpha : \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, M\} \to [0, 1]$, $(x, y) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, kM\}$ *and* $\eta > 0$ *we define* $\widetilde{\alpha}(x, y) = \max(\alpha(a_1, i_1) + \ldots + \alpha(a_k, i_k))$, *where the maximum is taken over all pairs* $(a_1, i_1), \ldots, (a_k, i_k) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, M\}$ *such that* $\alpha(a_1, i_1), \ldots, \alpha(a_k, i_k) > \eta$, $a_1 + \ldots + a_k = x$ *and* $i_1 + \ldots + i_k - j = y$ *for some* $j \in \{0, \ldots, k-1\}$. *Then*

$$\sum_{x,y} \widetilde{\alpha}(x, y) \geq k^2 \sum_{a,i} \alpha(a, i) - k^2 \eta q M.$$

PROOF. Let us define a set $S = \{(a, i) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, M\} : \alpha(a, i) \leq \eta\}$ and a function

$$\alpha^{\dagger}(a, i) = \begin{cases} \alpha(a, i), \text{ if } (a, i) \notin X, \\ 0 \text{ if } (a, i) \in X. \end{cases}$$

We see that $\widetilde{\alpha}(x, y) = \widetilde{\alpha}^{\dagger}(x, y)$, where $\widetilde{\alpha}^{\dagger}(x, y)$ can be interpreted as a maximum in the sense of both this and the previous lemma. Then using the previous lemma we get

$$\sum_{x,y} \widetilde{\alpha}(x,y) = \sum_{x,y} \widetilde{\alpha}^{\dagger}(x,y) \geq k^2 \sum_{a,i} \alpha^{\dagger}(a,i) = k^2 \left(\sum_{(a,i) \notin S} \alpha(a,i) - \sum_{(a,i) \in S} \alpha(a,i) \right)$$
$$\geq k^2 \sum_{a,i} \alpha(a,i) - k^2 \eta q M. \qquad \square$$

We will continue the proof of Theorem 1.2. Put

$$\alpha'(a, i) = \begin{cases} \alpha(a, i) \text{ if } (a, i) \notin E, \\ 0 \text{ if } (a, i) \in E. \end{cases}$$

Then by Lemma 3.6 applied with $\alpha'(a, i)$

$$|kA \cap I_{a_1 + \ldots + a_k, i_1 + \ldots + i_k - j}| \geq \frac{N}{qM} \min(\alpha'(a_1, i_1) + \ldots + \alpha'(a_k, i_k), 1) - 7k\widetilde{\varepsilon}^2 \frac{N}{qM}$$

for any $0 \le j \le k-1$ provided $\alpha'(a_h, i_h) \ge 4\widetilde{\varepsilon}^2$ for all $1 \le h \le k$. Summing up for different values of $x = a_1 + \ldots + a_k$ and $y = i_1 + \ldots + i_k - j$ we obtain

$$|kA| \ge \frac{N}{qM} \sum_{x,y} \min(\widetilde{\alpha}'(x,y), 1) - 7k\widetilde{\varepsilon}^2 N$$

where $\widetilde{\alpha}'$ is the maximum defined in Lemma 3.8 with $\eta = 4\widetilde{\varepsilon}^2$. But then

$$|kA| \ge \frac{N}{qM} \sum_{x,y} \min(\widetilde{\alpha}'(x,y), 1 + \frac{4}{5k}\varepsilon) - \frac{4}{5}\varepsilon N - 7k\widetilde{\varepsilon}^2 N.$$

Suppose that $\widetilde{\alpha}'(x,y) < 1 + \frac{4}{5k}\varepsilon$ for all (x,y). Then by Lemma 3.8

$$\begin{aligned} |kA| &\ge \frac{N}{qM}\sum_{x,y}\widetilde{\alpha}'(x,y) - \frac{4}{5}\varepsilon N - 7k\widetilde{\varepsilon}^2 N \ge \frac{N}{qM}k^2\sum_{a,i}\alpha'(a,i) - 4\widetilde{\varepsilon}^2 k^2 N - \frac{4}{5}\varepsilon N - 7k\widetilde{\varepsilon}^2 N \\ &\ge \frac{N}{qM}k^2\sum_{a,i}\alpha(a,i) - 5\widetilde{\varepsilon}^2 k^2 N - \frac{4}{5}\varepsilon N - 7k\widetilde{\varepsilon}^2 N \ge k^2|A| - 6\widetilde{\varepsilon}^2 k^2 N - \frac{4}{5}\varepsilon N - 7k\widetilde{\varepsilon}^2 N \\ &\ge k^2|A| - (6k^2 + 7k)\frac{\varepsilon}{100k^2}N - \frac{4}{5}\varepsilon > k^2|A| - \varepsilon N \end{aligned}$$

which gives us a contradiction with the assumptions of the theorem.
Therefore there exists $(x,y) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, kM\}$ such that $\widetilde{\alpha}'(x,y) \ge 1 + \frac{4}{5k}\varepsilon$ and consequently there are $(a_1, i_1), \ldots, (a_k, i_k) \in \mathbb{Z}/q\mathbb{Z} \times \{1, \ldots, M\}$ such that $\alpha(a_1, i_1) + \ldots + \alpha(a_k, i_k) \ge 1 + \frac{4}{5k}\varepsilon$. But then there exists an arithmetic progression $P \subseteq \{1, \ldots, N\}$ of length at least cN (where $c = c(q, M)$) such that $|A \cap P| \ge \frac{1}{k}(1 + \frac{4}{5k}\varepsilon)|P|$. Let $P = \{x + d, x + 2d, \ldots, x + |P|d\}$ be this arithmetic progression and let $A' = \frac{(A \cap P) - x}{d} = \{\frac{a-x}{d} : a \in A \cap P\}$. Then, clearly $A' \subseteq \{1, \ldots, |P|\}$ and $|A'| \ge \frac{1}{k}(1 + \frac{4}{5k}\varepsilon)|P|$. We will consider two cases:

- $|(k-1)A'| > (\frac{1}{2} + \frac{1}{k^2+1}\varepsilon)((k-1)|P| - k + 2)$.
  Thus, $(k-1)$-fold sumset of A' has density $(\frac{1}{2} + \frac{1}{k^2+1}\varepsilon)$ on the arithmetic progression $Q = \{k-1, \ldots, (k-1)|P|\}$. Observe that $(k-1)A' \cdot d + (k-1)x \subseteq (k-1)A$, $Q \cdot d + (k-1)x \subseteq (k-1)P$ and that $(k-1)P$ is an arithmetic progression contained in $\{1, \ldots, (k-1)N\}$ of length

$$|(k-1)P| \ge (k-1)|P| - k + 2 \ge (k-1)cN - k + 2 \ge c(\varepsilon, k)N.$$

  Therefore the statement of the theorem is satisfied with $j = k-1$ and $P' = (k-1)P$.

- $|(k-1)A'| \leq (\frac{1}{2} + \frac{1}{k^2+1}\varepsilon)((k-1)|P| - k + 2)$.
  Then we have

$$\begin{aligned}|(k-1)A'| &\leq (\frac{1}{2} + \frac{1}{k^2+1}\varepsilon)((k-1)|P| - k + 2) \\ &\leq \frac{1}{2}(1 + \frac{4}{5k}\varepsilon)(k-1)|P| - \frac{2k^2+2-5k}{5k(k^2+1)}\varepsilon(k-1)|P| \\ &\leq \frac{1}{2}(1 + \frac{4}{5k}\varepsilon)(k-1)\frac{k}{1+\frac{4}{5k}\varepsilon}|A'| - \frac{1}{15}\varepsilon|P| \leq (k-1)^2|A'| - \frac{1}{15}\varepsilon|P|.\end{aligned}$$

So we can use the induction hypothesis for the main theorem for the set A', $N = |P|$, $k-1$ and $\frac{1}{15}\varepsilon$. It implies that there exists an arithmetic progression $Q \subseteq \{1, \ldots, j|P|\}$ of length $|Q| \geq c(\frac{1}{15}\varepsilon, k-1)|P| \geq c(\varepsilon, k)N$ such that $|jA' \cap Q| \geq (\frac{1}{2} + \frac{1}{j^2+1}\varepsilon)|Q|$ for some $1 \leq j \leq k-2$. Therefore, $|jA \cap (Q \cdot d + jx)| \geq (\frac{1}{2} + \frac{1}{j^2+1}\varepsilon)|Q|$ and the assertion follows.

4. The conjecture of Fox and Kleitman

In this section we will study the degree of regularity of the equation (2). We begin with a simple lemma, which guarantees the existence of the long arithmetic progression contained in the difference set $A - A$ for sets A with density bigger than $1/2$.

LEMMA 4.1. *Let P be an arithmetic progression of difference d and A be a subset of integers. For every $\varepsilon > 0$ if $|A \cap P| > (\frac{1}{2} + \varepsilon)|P|$, then $\pm kd \in A - A$ for $0 \leq k \leq \lfloor 2\varepsilon|P| \rfloor$.*

PROOF. Let $P = \{x + d, x + 2d \ldots, x + ld\}$, $A' = A \cap P$ and let $0 \leq k \leq \lfloor 2\varepsilon|P| \rfloor$. To show that $kd \in A - A$ it is sufficient to show that $|(A' + kd) \cap A'| > 0$. We have

$$|(A' + kd) \cap A'| = |A' + kd| + |A'| - |(A' + kd) \cup A'| > 2(\frac{1}{2} + \varepsilon)|P| - (|P| + k) \geq 0,$$

because $(A' + kd) \cup A' \subseteq \{x + d, x + 2d \ldots, x + (l + k)d\}$. □

Now we are ready to estimate from below the degree of regularity of the equation

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i + b_n.$$

THEOREM 4.1. *Let $n \geq 2$. Then there exists an integer b_n such that for every N big enough and every $(2n-1)$-colouring of $\{1,\ldots,N\}$ there exists a monochromatic solution of the equation $x_1+\ldots+x_n = y_1+\ldots+y_n+b_n$.*

PROOF. Let us denote $r = 2n-1$ and $\varepsilon = \frac{1}{2r}$. Let $\delta(\varepsilon)$ be the minimum of constants estimating the length of arithmetic progressions in Theorem 1.2 for $k \in \{2,\ldots,n\}$ and ε. Put $b_n = \left(\left\lfloor \frac{2n}{\delta(\varepsilon)} \right\rfloor\right)!$ and suppose N is large enough. Then for any colouring of $\{1,\ldots,N\}$ with r colours we can find a monochromatic subset $A \subseteq \{1,\ldots,N\}$ of size at least $\frac{N}{r}$.

We will now show that there exist an integer $1 \leq j \leq n$ and an arithmetic progression $P \subseteq \{1,\ldots,jN\}$ of length $|P| \geq \delta(\varepsilon)N$ such that $|jA \cap P| \geq \left(\frac{1}{2}+\frac{1}{j^2+1}\varepsilon\right)|P|$. Suppose that this is not true. Then by Theorem 1.2 applied twice

$$|nA-nA| > 4|nA| - \varepsilon nN > 4(n^2|A| - \varepsilon N) - \varepsilon nN \geq 4n^2\frac{N}{2n-1} - \frac{4+n}{4n-2}N$$
$$\geq \frac{4n^2-2n+2n}{2n-1}N - \frac{4+n}{4n-2}N \geq 2nN + \frac{4n-4-n}{4n-2}N > 2nN.$$

On the other hand, $|nA-nA| \leq 2nN$, which gives us a contradiction. Hence there exists an arithmetic progression P of length $|P| \geq \delta(\varepsilon)N$ such that $|jA \cap P| \geq$ $\geq (\frac{1}{2}+\frac{1}{j^2+1}\varepsilon)|P|$ for some $1 \leq j \leq n$. Let d be the common difference of P. Lemma 4.1 implies that $\pm id \in jA - jA$ for $0 \leq i \leq \lfloor 2\frac{1}{j^2+1}\varepsilon|P| \rfloor$. Since $P \subseteq \{1,\ldots,nN\}$ we have

$$(|P|-1)d \leq nN \Rightarrow d \leq \frac{nN}{|P|-1} \leq \frac{nN}{\delta(\varepsilon)N-1} \leq \frac{nN}{\frac{1}{2}\delta(\varepsilon)N} = \frac{2n}{\delta(\varepsilon)}.$$

But d is an integer, so $d \leq \lfloor \frac{2n}{\delta(\varepsilon)} \rfloor$. This means that for N big enough there exists $1 \leq i \leq \lfloor 2\frac{1}{j^2+1}\varepsilon|P| \rfloor$ such that $id = \left(\left\lfloor \frac{2n}{\delta(\varepsilon)} \right\rfloor\right)!$. So we can find elements $x_1,\ldots x_j, y_1,\ldots,y_j \in A$ such that

$$x_1+\ldots+x_j+(n-j)x_j - (y_1+\ldots+y_j+(n-j)x_j) = b_n,$$

which gives us a monochromatic solution to the equation $x_1+\ldots+x_n = y_1+\ldots+y_n+b_n$. □

Acknowledgements

The authors are partially supported by National Science Centre, Poland grant 2016/21/B/ST1/00307.

Bibliography

1. ***S. Eberhard,*** *The abelian arithmetic regularity lemma*, https://www.dpmms.cam.ac.uk/∼se288/abelianregularity.pdf
2. ***S. Eberhard, B. Green, F. Manners,*** *Sets of integers with no large sum-free subset*, Annals of Math. **180** (2014), 621–652.
3. ***J. Fox, D. Kleitman,*** *On Rado's Boundedness Conjecture*, Journal of Combinatorial Theory Series A **113** (2006), 84–100.
4. ***G. Freiman,*** *On the addition of finite sets. I*, Izv. Vyssh. Uchebn. Zaved. Mat. **13** (6) (1959), 202–213.
5. ***B. Green, T. Tao,*** *An arithmetic regularity lemma*, *an associated counting lemma*, *and applications*, An irregular mind, Bolyai Soc. Math. Stud., 21, Janos Bolyai Math. Soc., Budapest, (2010), 261–334.
6. ***E. G. Straus,*** *A Combinatorial Theorem in Group Theory*, Mathematics of Computation **29** (1975), 303–309.
7. ***T. Tao,*** *Spending Symmetry*, American Mathematical Society, Providence, R. I., available at https://terrytao.files.wordpress.com/2012/11/blog-book.pdf

Tomasz Schoen

Faculty of Mathematics
and Computer Science,
Adam Mickiewicz University,
Umultowska 87,
61–614 Poznań, Poland
schoen@amu.edu.pl

Katarzyna Taczała

Faculty of Mathematics
and Computer Science,
Adam Mickiewicz University,
Umultowska 87,
61–614 Poznań, Poland
katarzyna.taczala@amu.edu.pl