On sets with small doubling *

Shkredov I.D.

Annotation.

Let \( G \) be an arbitrary Abelian group and let \( A \) be a finite subset of \( G \). \( A \) has small additive doubling if \( |A + A| \leq K|A| \) for some \( K > 0 \). These sets were studied in papers of G.A. Freiman, Y. Bilu, I. Ruzsa, M.C.–Chang, B. Green and T. Tao. In the article we prove that if we have some minor restrictions on \( K \), then for any set with small doubling there exists a set \( \Lambda \), \( \Lambda \ll \varepsilon K \log |A| \) such that \( |A \cap \Lambda| \gg |A|/K^{1/2+\varepsilon} \), where \( \varepsilon > 0 \). In contrast to the previous results our theorem is nontrivial for large \( K \). For example one can take \( K = |A| \eta \), where \( \eta > 0 \). We use an elementary method in our proof.

1. Introduction.

Let \( G \) be an arbitrary Abelian group with additive group operation \(+\). Suppose that \( A, B \) are two finite subsets of \( G \) and define their sumset \( A + B \) to be the set of all pairwise sums \( a + b \) with \( a \in A, b \in B \). Let \( \log \) stand for the logarithm to base 2.

Suppose that \( A \) is a set such that \( |A + A| \leq K|A| \), where \( K \geq 1 \) is small (for example \( K = \log \log |A| \) or \( K = 2 \)). These sets are called sets with small doubling. The properties of such sets were studied in papers [4, 5, 6, 7, 9, 10, 13, 14, 15, 16]. G.A. Freiman (see [4]) proved the following wonderful result on the structure of these sets.

Recall that a set \( Q \subseteq G \) is called a \( d \)-dimensional arithmetic progression if
\[
Q = \{n_0 + n_1 \lambda_1 + \cdots + n_d \lambda_d : 0 \leq \lambda_i < m_i\},
\]
where \( m_i, n_i \in \mathbb{Z} \) and \( m_i \geq 0 \).

Let \( G = \mathbb{Z} \).

**Theorem 1.1 (Freiman)** Let \( K \geq 1 \) be a real number, and \( A \subseteq \mathbb{Z} \) be a finite set. Let also \( |A + A| \leq K|A| \). Then there exist numbers \( d = d(K) \) and \( C = C(K) \) depend on \( K \) only and \( d \)-dimensional arithmetic progression \( Q \) such that \( |Q| \leq C|A| \) and \( A \subseteq Q \).

The functions \( d = d(K) \) and \( C = C(K) \) were studied in [6, 7]. In paper [7] M.–C. Chang proved that \( d = O(K^2 \log^2 K) \) and \( C = \exp(O(K^2 \log^2 K)) \) (as usual we use \( X = O(Y) \) or \( X \ll Y \) to denote an estimate of the form \( X \leq MY \) for some absolute constant \( M \)).

Let \( n \) be a positive integer. Sets with small doubling in groups \( G = (\mathbb{Z}/2\mathbb{Z})^n \) were considered in [9, 17, 13, 14, 16]. For example we formulate a theorem from [9]. Note that \((\mathbb{Z}/2\mathbb{Z})^n\) is a vector space.

---

*This work was supported by RFFI grant no. 06-01-00383, President's of Russian Federation grant N 1726.2006.1 and INTAS (grant no. 03–51–5–70).
Theorem 1.2 Let $K \geq 1$ be a real number. Let $A \subseteq (\mathbb{Z}/2\mathbb{Z})^n$ be a set such that $|A + A| \leq K|A|$. Then $A$ is contained in a subspace $H$ with $|H| \leq K^22^{K^2}|A|$.

There are another structural results on sets with small doubling. Let $A$ be a set with small doubling and $d$ is a small positive integer. Is it true that $A$ has large intersection with some $d$-dimensional arithmetic progression? It is known that there is a positive answer at the question and we give two examples of such results. In [14] the following theorem was proved.

Theorem 1.3 Let $K \geq 1$ be a real number. Suppose that $A \subseteq (\mathbb{Z}/2\mathbb{Z})^n$ is a set such that $|A + A| \leq K|A|$. Then there exists a subspace $H$ such that $|H| \ll K^{O(1)}|A|$ and $|A \cap H| \gg \exp(-K^{O(1)})|A|$.

Finally, in recent paper [16] B. Green and T. Tao proved the following theorem.

Theorem 1.4 Let $K \geq 1$ be a real number. Let $A \subseteq (\mathbb{Z}/2\mathbb{Z})^n$ be a set such that $|A + A| \leq K|A|$. Then there exists a subspace $H$ and $x \in (\mathbb{Z}/2\mathbb{Z})^n$ such that $|H| \gg K^{-O(\sqrt{K})}|A|$ and $|A \cap (x + H)| \geq \frac{1}{32}|H|$.

Let us formulate our main result. Let $E = \{e_1, \ldots, e_{|E|}\} \subseteq G$ be a finite set. By $\text{Span} E$ denote the set $\text{Span} E = \{ \sum_{i=1}^{\infty} e_i : \varepsilon_i \in \{-1, 0, 1\}\}.$

Theorem 1.5 Let $G$ be an Abelian group. Let $K, \varepsilon$ be real numbers, $\varepsilon \in (0, 1/2]$, $A \subseteq G$ be a finite set, $|A| \geq 2^{32/\varepsilon}$, $1 \leq K \leq \min\{ (2^{-58}e^{-4}|A|^{3/2+\varepsilon})^{-1}, |A|^\varepsilon \}$. Let also $A$ contains at least $|A|^3/K$ quadruples with $a_1 + a_2 = a_3 + a_4$. Then there exists a set $\Lambda$ such that $|\text{Span} \Lambda \cap A| \geq \frac{1}{2} \cdot \frac{|A|}{|\Lambda|}$ and $|\Lambda| \leq 2^{30}e^{-2}K\log |A|$.

It is easy to see that the number $K$ in Theorems 1.1,1.2,1.3,1.4 cannot be too large. For example Theorem 1.4 is trivial if $K \gg \left( \frac{\log |A|}{\log \log |A|} \right)^2$. In contrast to these results our Theorem 1.5 is nontrivial for large $K$ (for example one can take $K = |A|^\eta$, where $\eta > 0$ is a sufficiently small number). On the other hand the cardinality of the set $\Lambda$ depends on $|A|$. This fact differnces our main result from Theorems 1.1—1.4.

This paper is organized as follows. In §2 we study so-called ”connected sets” in Abelian groups. We prove that any such set has large intersection with $\text{Span} \Lambda$ for some small set $\Lambda$ (in more detail see Propositions 2.7, 2.9). Besides in the section we show that any set contains large connected subset. These two facts imply Theorem 1.5. We give its proof in §3. In the last section we discuss some relations between our definition of connectedness and a graph–theoretical definition of connectedness from paper [8].

The author is grateful to Professor N.G. Moshchevitin for constant attention to this work.

2. On connected sets in Abelian groups.

Let $G$ be an arbitrary Abelian group with additive group operation $+$. Let $A \subseteq G$ be a finite set, and $k \geq 2$ be a positive integer. By $T_k(A)$ denote the following number

$$T_k(A) := \left| \{ a_1 + \cdots + a_k = a'_1 + \cdots + a'_k : a_1, \ldots, a_k, a'_1, \ldots, a'_k \in A \} \right|.$$ 

Denote by the same letter $A$ the characteristic function of $A$. Thus $A(x) = 1$ if $x \in A$ and $A(x) = 0$ otherwise. We shall write $\sum_A$ instead of $\sum_{x \in G}$.

Definition 2.1 Let $k \geq 2$ be a positive integer, and $\beta \in [0, 1]$ be a real number. Suppose that $A$ is a finite nonempty set $A \subseteq G$. $A$ is called $\beta$–connected of degree $k$ if there is an absolute constant $C \in (0, 1]$ such that for any $B \subseteq A$, $|B| \geq \beta|A|$ we have

$$T_k(B) \geq C^{2k} \left( \frac{|B|}{|A|} \right)^{2k} T_k(A).$$  \hspace{1cm} (1)
If $\beta = 0$ then $A$ is connected of degree $k$.

The class of connected sets is wide enough. On the one hand very structured sets like arithmetic progressions, multidimensional arithmetic progressions, subspaces are connected sets (see Corollary 2.4 below). On the other hand any so-called dissociated set (see Definition 2.5) belongs to this class. Other examples of connected of degree $k$ sets will be considered in section 4.

**Definition 2.2** Let $f, g : G \to \mathbb{R}$ be arbitrary functions. Denote by $(f * g)(x)$ the function

$$
(f * g)(x) = \sum_s f(s)g(x - s).
$$

Clearly, $(f * g)(x) = (g * f)(x)$, $x \in G$. By $(f \circ g)(x)$ denote the function

$$
(f \circ g)(x) = \sum_s f(s)g(s - x).
$$

Obviously, $(f \circ g)(x) = (g \circ f)(-x)$, $x \in G$.

Suppose that $A, B \subseteq G$ are any sets. Then $(A * B)(x) \neq 0$ iff $x \in A + B$ and $(A \circ B)(x) \neq 0$ iff $x \in A - B$. Hence $T_2(A) = \sum_x (A * A)^2(x)$. Further denote by $*_k$ the composition of $k$ operations $\ast$, $k \geq 1$. Then $T_k(A) = \sum_x (A *_{k-1} A)^2(x)$, $k \geq 2$. Since $T_2(A) := |\{a_1 + a_2 = a_1' + a_2' : a_1, a_2, a_1', a_2' \in A\}| = |\{a_1 - a_1' = a_2 - a_2' : a_1, a_2, a_1', a_2' \in A\}|$ it follows that $T_2(A) = \sum_x (A \circ A)^2(x)$.

Let $f : G \to \mathbb{R}$ be a function. By $T_k(f)$ denote the quantity $T_k(f) = \sum_x (f *_{k-1} f)^2(x)$. Let us prove the following simple lemma.

**Lemma 2.3** Let $p_1, p_2$ be positive integers, and $k_1 = 2^{p_1}$, $k_2 = 2^{p_2}$. Let also $f_1, \ldots, f_{k_1}, g_1, \ldots, g_{k_2} : G \to \mathbb{R}$ be functions. Then

$$
\left| \sum_x (f_1 \ast \cdots \ast f_{k_1})(x) \cdot (g_1 * \cdots * g_{k_2})(x) \right| \leq (T_{k_1}(f_1))^{1/2k_1} \ldots (T_{k_1}(f_{k_1}))^{1/2k_1} (T_{k_2}(g_1))^{1/2k_2} \ldots (T_{k_2}(g_{k_2}))^{1/2k_2}. \tag{4}
$$

**Proof.** First of all let us suppose that $k_1 = k_2 = k = 2^p$, where $p$ is a positive integer. We prove the lemma by induction. Put $\sigma = \sum_x (f_1 \ast \cdots \ast f_k)(x) \cdot (g_1 * \cdots * g_k)(x)$.

Using the Cauchy–Schwartz inequality, we get

$$
\sigma^2 \leq \sum_x (f_1 \ast \cdots \ast f_k)^2(x) \cdot \sum_x (g_1 * \cdots * g_k)^2(x) = \sigma_1 \sigma_2. \tag{5}
$$

Consider the sum $\sigma_1$. By definitions of $\ast, \circ$, we obtain

$$
\sigma_1 = \sum_x ((f_1 \circ f_1) \ast \cdots \ast (f_{2^p-1} \circ f_{2^p-1}))(x) \cdot ((f_{2^p-1+1} \circ f_{2^p-1+1}) \ast \cdots \ast (f_k \circ f_k))(x).
$$

By the induction hypothesis, we have

$$
\sigma_1 \leq (T_{2^p-1}(f_1 \circ f_1))^{1/k} \ldots (T_{2^p-1}(f_k \circ f_k))^{1/k}. \tag{6}
$$

Besides, $T_{2^p-1}(f_1 \circ f_1) = T_k(f_1)$. Hence

$$
\sigma_1 \leq (T_k(f_1))^{1/k} \ldots (T_k(f_k))^{1/k}. \tag{7}
$$
In the same way
\[ \sigma_2 \leq (T_k(g_1))^{1/k} \cdots (T_k(g_k))^{1/k}. \] (8)
Combining (7), (8) and (5), we obtain that (4) holds.

Let now \( k_1 = 2p_1, \ k_2 = 2p_2, \) and \( p_1 \neq p_2. \) Put \( \sigma' = \sum_x (f_1 \cdots f_{k_1})(x) \cdot (g_1 \cdots g_{k_2})(x). \) Using the Cauchy–Schwartz inequality, we get
\[ \sigma'^2 \leq \sum_x (f_1 \cdots f_{k_1})^2(x) \cdot \sum_x (g_1 \cdots g_{k_2})^2(x) = \sigma_1' \sigma_2'. \] (9)
Using (4) for \( \sigma'_1, \ \sigma'_2, \) we have
\[ |\sigma'| \leq (T_k(f_1))^{1/2k_1} \cdots (T_k(f_{k_1}))^{1/2k_1}(T_k(g_1))^{1/2k_2} \cdots (T_k(g_{k_2}))^{1/2k_2}. \] (10)
This completes the proof.

Let us derive a corollary from Lemma 2.3. Let \( n \) be positive integer, \( q \) be a prime and let \( G = (\mathbb{Z}/q\mathbb{Z})^n. \) As was noted above \( G \) is a vector space.

**Corollary 2.4** Let \( n, \ p \) be positive integers, \( k = 2^p, \ q \) be a prime, and \( G = (\mathbb{Z}/q\mathbb{Z})^n. \) Let also \( P \) be a subspace of \( G. \) Then \( P \) is a connected of degree \( k \) set and (1) is true for \( C = 1. \)

**Proof.** Let \( B \subseteq P \) be a set and let \( \sigma(B) := \sum_x (B * P *_{k-2} P)(x) \cdot (P *_{k-1} P)(x). \) The sum \( \sigma(B) \) equals the number of solutions of the equation \( b + p_2 + \cdots + p_k = p'_1 + \cdots + p'_k, \) where \( b \in B \) and \( p_2, \ldots, p_k, p'_1, \ldots, p'_k \in P. \) Since \( P \) is a subspace of \( (\mathbb{Z}/q\mathbb{Z})^n \) it follows that \( b + p_2 + \cdots + p_k - p'_1 - \cdots - p'_k \in P. \) Hence \( \sigma(B) \geq |B||P|^{k-2}. \) In particular \( \sigma(P) = T_k(B) \geq |P|^{2k-1}. \) Since \( T_k(P) \leq |P|^{2k-1} \) it follows that \( T_k(P) = |P|^{2k-1}. \) Using Lemma 2.3 with \( f_1 = B, f_2 = \cdots = f_k = g_1 = \cdots = g_k = A, \) we get
\[ \sigma^{2k}(B) \leq T_k(B) \cdot T_k^{2k-1}(A). \] (11)
Combining the last inequality and the lower bound for \( \sigma(B) \), we obtain
\[ T_k(B) \geq \frac{\sigma^{2k}(B)}{T_k^{2k-1}(A)} \geq \frac{|B|^{2k} |P|^{(2k-2)2k}}{|P|^{2k-1}^2} = \left( \frac{|B|}{|P|} \right)^{2k} \frac{|P|^{2k-1}}{|P|^{2k-1}} = \left( \frac{|B|}{|P|} \right)^{2k} T_k(P). \]
This completes the proof.

Thus very structured sets like subspaces are connected sets. Consider another examples of connected sets.

We need in the following definition (see [18] or [7]).

**Definition 2.5** We say that \( \Lambda = \{\lambda_1, \ldots, \lambda_{|\Lambda|}\} \subseteq G \) is **dissociated** if the equality
\[ \sum_{i=1}^{|\Lambda|} \varepsilon_i \lambda_i = 0, \] (12)
where \( \varepsilon_i \in \{-1, 0, 1\} \) implies that all \( \varepsilon_i \) are equal to zero.

If \( \Lambda \) is a dissociated set then there exists a good upper bound for \( T_k(\Lambda) \) (see [18] and also [11, 19]).

**Statement 2.6** There is an absolute constant \( M > 0 \) such that for any dissociated set \( \Lambda \subseteq G \) and any positive integer \( k \geq 2, \) we have
\[ T_k(\Lambda) \leq M^k k^k |\Lambda|^k, \] (13)
where $M \leq 288$.

Any connected of degree $k$ set has the following property.

**Proposition 2.7** Let $k \geq 2$ be a positive integer. Suppose that $A \subseteq G$ is a connected of degree $k$ set and for $C > 0$ inequality (1) holds. Then there exists a set $\Lambda \subseteq A$, $|\Lambda| \leq 288C^{-2k} \frac{|A|^2}{T_k^{1/k}(A)}$ such that any $a \in A$ can be expressed in the form

$$a = \sum_{i=1}^{|\Lambda|} \varepsilon_i \lambda_i,$$

where $\varepsilon_i \in \{-1, 0, 1\}$.

**Proof.** Let $\Lambda$ be a maximal dissociated subset of $A$. Prove that any $a \in A$ can be expressed in the form

$$a = \sum_{i=1}^{|\Lambda|} \varepsilon_i \lambda_i,$$

where $\varepsilon_i \in \{-1, 0, 1\}$. If $a = 0$ then (15) holds. Let $a$ be an arbitrary element of $A \setminus \Lambda$, $a \neq 0$. Consider all equations $\sum_{i=1}^{|\Lambda|+1} \varepsilon_i \tilde{\lambda}_i = 0$, where $\tilde{\lambda}_i \in A \setminus \{a\}$ and $\varepsilon_i \in \{-1, 0, 1\}$, $i \in \{1, 2, \ldots, |\Lambda|+1\}$. If all these equations are trivial, i.e. we have $\varepsilon_i = 0$, $i \in \{1, 2, \ldots, |\Lambda|+1\}$ then we obtain a contradiction with the maximality of $\Lambda$. It follows that there exists non-trivial equation $\varepsilon a + \sum_{i=1}^{|\Lambda|} \varepsilon_i \lambda_i = 0$, $\varepsilon, \varepsilon_i \in \{-1, 0, 1\}$ such that not all $\varepsilon, \varepsilon_i$ are equal to zero. Note that $\varepsilon \neq 0$. Whence any $a \in A$ is involved in some equation (14).

Let us prove that $|\Lambda| \leq 288C^{-2k} \frac{|A|^2}{T_k^{1/k}(A)}$. Using Statement 2.6, we have $T_k(\Lambda) \leq (288)^k k^k |\Lambda|^k$. On the other hand, the set $A$ is connected of degree $k$. Hence $T_k(\Lambda) \geq C^{2k}(|\Lambda|/|A|)^2k \cdot T_k(A)$. It follows that $|\Lambda| \leq 288C^{-2k} \frac{|A|^2}{T_k^{1/k}(A)}$. This completes the proof.

We need in a more delicate definition of connectedness.

**Definition 2.8** Let $k \geq 2$ be a positive integer, and $\beta_1, \beta_2 \in [0, 1]$ be real numbers, $\beta_1 \leq \beta_2$. Nonempty set $A \subseteq G$ is called $(\beta_1, \beta_2)$-connected of degree $k$ if there exists an absolute constant $C \in (0, 1]$ such that for any $B \subseteq A$, $\beta_1 |A| \leq |B| \leq \beta_2 |A|$ we have

$$T_k(B) \geq C^{2k} \left( \frac{|B|}{|A|} \right)^{2k} T_k(A).$$

Clearly, any $\beta$-connected of degree $k$ set is a $(\beta, \beta_2)$-connected of degree $k$, where $\beta_2 \in [\beta, 1]$ is an arbitrary number. Nevertheless we have the following weak analog of Proposition 2.7 for $(\beta_1, \beta_2)$-connected of degree $k$ sets.

**Proposition 2.9** Let $k \geq 2$ be a positive integer, $0 < \beta_1 \leq \beta_2$ be real numbers. Let also $A \subseteq G$ be a $(\beta_1, \beta_2)$-connected of degree $k$ set and for $C > 0$ inequality (1) holds. Suppose that $\beta_2 \geq \beta_1 + 1/|A|$, $T_k(A) \geq 2^{4k} C^{-2k} k^k |A|^k$ and $|A| \geq 1/\beta_1$. Then there exists a set $\Lambda \subseteq A$ such that

$$|\Lambda| \leq 2^{13} C^{-2k} \frac{|A|^2}{T_k^{1/k}(A)},$$

and $|\text{Span} \Lambda \cap A| \geq (1 - \beta_1)|A|$.

**Proof.** The proof of the proposition is a sort of inductive process. Let $\Lambda_1$ be a dissociated subset of $A$ such that $|\text{Span} \Lambda_1 \cap A| \geq (1 - \beta_1)|A|$. Clearly, there exists such set $\Lambda_1$, for example one put $\Lambda_1$ to be a maximal dissociated subset of $A$. Let $l = 2^{13} C^{-2k} \frac{|A|^2}{T_k^{1/k}(A)}$. If
|Λ1| \leq l \text{ then the proposition is proved. Suppose that } |Λ_1| > l. \text{ Let } Λ'_1 \subseteq Λ_1 \text{ be an arbitrary set of the cardinality } l. \text{ Obviously, that } Λ'_1 \text{ is a dissociated set. Consider the set } A_1 = A \setminus Λ'_1. \text{ If } |A_1| < (1 - β_1)|A| \text{ then we stop our algorithm. If } |A_1| \geq (1 - β_1)|A| \text{ then let } Λ_2 \text{ be a dissociated subset of } A_1 \text{ such that } |\text{Span } Λ_2 \cap A_1| \geq (1 - β_1)|A|. \text{ Suppose that } |Λ_2| \leq l. \text{ Then } |\text{Span } Λ_2 \cap A| \geq |\text{Span } Λ_2 \cap A_1| \geq (1 - β_1)|A| \text{ and we are done. It follows that } |Λ_2| > l. \text{ Let } Λ'_2 \subseteq Λ_2 \text{ be an arbitrary set of the cardinality } l \text{ and consider the set } A_2 = A_1 \setminus Λ'_2. \text{ An so on. We get the sets } A_0 = A, A_1, A_2, \ldots, A_s \text{ and disjoint dissociated sets } Λ'_1, \ldots, Λ'_s \text{ from } A. \text{ We have } |A_s| < (1 - β_1)|A|. \text{ Since for all } l = 1, 2, \ldots, s \text{ the following holds } A_l = A \setminus \bigcup_{i=1}^{l} Λ'_i \text{ it follows that } \sum_{i=1}^{s} |Λ'_i| = |A| - |A_s| > β_1|A|. \text{ Let } B = \bigcup_{i=1}^{s} Λ'_i. \text{ Then } |B| > β_1|A|. \text{ We can remove some elements from } Λ'_s \text{ and assume that the cardinality of } \bigcup_{i=1}^{s} Λ'_i \text{ equals } [β_1|A]| + 1. \text{ Denote by the same letter } B \text{ our modified set. We have } B \subseteq A \text{ and } |B| \geq β_1|A|. \text{ Since } β_2 \geq β_1 + 1/|A| \text{ it follows that } |B| \leq β_2|A|. \text{ By assumption the set } A \text{ is } (β_1, β_2) \text{ connected of degree } k. \text{ Hence}

\begin{equation}
T_k(B) \geq C^{2k} β_1^{2k} T_k(A).
\end{equation}

On the other hand

\begin{equation}
T_k(B) \leq T_k \left( \bigcup_{i=1}^{s} Λ'_i \right) = \sum_{i_1, \ldots, i_k=1}^{s} \sum_{j_1, \ldots, j_k=1}^{s} \sum_{x} (Λ'_{i_1} \ast \cdots \ast Λ'_{i_k})(x) \cdot (Λ'_{j_1} \ast \cdots \ast Λ'_{j_k})(x).
\end{equation}

Using Lemma 2.3, Statement 2.6 and (19), we get

\begin{equation}
T_k(B) \leq s^{2k} \max_{i=1, \ldots, s} T_k(Λ'_i) \leq s^{2k}(288)^k k^k k^l.
\end{equation}

By assumption \( T_k(A) \geq 2^{14k} C^{-2k} k|A|^k \). \text{ Whence } |A| \geq 2^{14k} C^{-2k} k T_k(A)^{1/k} = 2l \text{ and } s \geq 2. \text{ Since } \bigcup_{i=1}^{s-1} Λ'_i \subseteq B \text{ and } |A| \geq 1/β_1 \text{ it follows that } sl/2 \leq (s - 1)l \leq |B| \leq 2β_1|A|. \text{ Hence } s \leq 4β_1|A|/l. \text{ Combining the last inequality and (20), we have}

\begin{equation}
T_k(B) \leq 2^{4k} β_1^{2k}(288)^k k^k l|A|^{2k} k^k l.
\end{equation}

This contradicts with (18) and we obtain the required result.

Let us prove now that any \( A \subseteq G \text{ contains some large } (β_1, β_2) \text{–connected of degree } k \text{ set. We begin with some notation.}

\text{Definition 2.10 Let } A \subseteq G \text{ be an arbitrary finite set, } |A| \geq 2, \text{ and } k \geq 2 \text{ be a positive integer. By } ζ_k(A) \text{ denote the quantity}

\begin{equation}
ζ_k = ζ_k(A) := \frac{\log T_k(A)}{\log |A|}.
\end{equation}

In other words } T_k(A) = |A|^{ζ_k}. \text{ Clearly, for any set } A, \text{ we have } k \leq ζ_k(A) \leq 2k - 1.

\text{Let } A \subseteq G \text{ be a finite set, } |A| = m \geq 2, \text{ } p \text{ be a positive integer, and } k = 2^p. \text{ Write } ζ \text{ for } ζ_k(A).

\text{Theorem 2.11 Let } β_1, β_2 \in (0, 1) \text{ be real numbers, } β_1 \leq β_2. \text{ Then there exists a set } A' \subseteq A \text{ such that}
\begin{enumerate}
\item \( A' \text{ is } (β_1, β_2) \text{–connected of degree } k \text{ set such that (16) holds for any } C \leq 1/32. \)
\item \( |A'| \geq m \cdot 2^{\frac{\log(2^k - 1)/ζ \log(1 - β_1)}{\log m}} k \log(1 - β_2), \text{ where } κ = \frac{\log((1 - β_1)^{-1})}{\log m} (1 - 16C). \)
\item \( ζ_k(A') \geq ζ_k(A). \)
\end{enumerate}
Proof. Let $C \leq 1/32$ be a real number. The proof of Theorem 2.11 is a sort algorithm. If $A$ is $(\beta_1, \beta_2)$–connected of degree $k$ and (16) is true with the constant $C$ then there is nothing to prove. Suppose that $A$ is not $(\beta_1, \beta_2)$–connected of degree $k$ set (with the constant $C$). Then there exists a set $B \subseteq A$, $\beta_1|A| \leq |B| \leq \beta_2|A|$ such that (16) does not hold. Note that $|A| > 2$. Let $\overline{B} = A \setminus B$. We have

$$T_k(A) = \sum_x (A *_{k-1} A)^2(x) =$$

$$= \sum_x (B * A *_{k-2} A)(x)(A *_{k-1} A)(x) + \sum_x (B * A *_{k-2} A)(x)(A *_{k-1} A)(x) = \sigma_1 + \sigma_2. \quad (21)$$

Using Lemma 2.3 with $f_1 = B$, $f_2 = \cdots = f_k = g_1 = \cdots = g_k = A$, we obtain

$$\sigma_1^{2k} \leq T_k(B) \cdot T_k^{2k-1}(A). \quad (22)$$

In the same way

$$\sigma_2^{2k} \leq T_k(\overline{B}) \cdot T_k^{2k-1}(A). \quad (23)$$

Let $c_B = |B|/|A|$. Combining $T_k(B) < C^{2k}c_B^kT_k(A)$ and (22), we have $\sigma_1 < Cc_BT_k(A)$. Using the last inequality, (21) and (23), we get

$$T_k(\overline{B}) > T_k(A)(1 - Cc_B)^{2k}. \quad (24)$$

Let $\overline{\zeta} = \zeta_k(\overline{B})$, $b = |B|$ and $\overline{b} = |\overline{B}| = m - b$. Using (24), we obtain

$$\overline{\zeta} \log \overline{b} > \zeta \log m + 2k \log(1 - Cc_B).$$

Hence

$$\overline{\zeta} > \frac{\zeta \log m + 2k \log(1 - Cc_B)}{\log \overline{b}} = \frac{\zeta \log m + 2k \log(1 - Cc_B)}{\log m + \log(1 - b/m)} = \frac{\zeta + 2k \log((1 - Cc_B)/m)}{1 + \log((1 - c_B)/m)} \geq$$

$$\geq \left(\zeta + 2k \frac{\log((1 - Cc_B)/m)}{\log m}\right) \left(1 - \frac{\log(1 - c_B)}{\log m}\right) \geq \zeta + \frac{\log((1 - Cc_B)^{-1})}{\log m} (1 - 16C) \geq$$

$$\geq \zeta (1 + \frac{\log((1 - \beta_1)^{-1})}{\log m} (1 - 16C)) = \zeta (1 + \kappa), \quad (25)$$

where $\kappa = \frac{\log((1 - \beta_1)^{-1})}{\log m} (1 - 16C) > 0$. Besides, by the definition of $(\beta_1, \beta_2)$–connectedness of degree $k$, we have

$$|\overline{B}| \geq (1 - \beta_2)m = (1 - \beta_2)|A|. \quad (26)$$

Thus if the set $A$ is not $(\beta_1, \beta_2)$–connected of degree $k$ then there is a set $\overline{B} \subseteq A$ such that (25), (26) hold. Put $A_1 = \overline{B}$ and apply the arguments above to $A_1$. And so on. We get the sets $A_0 = A, A_1, A_2, \ldots, A_s$. Clearly, for any $A_i$, we have $\zeta(A_i) \leq 2k - 1$. Using this and (25), we obtain that the total number of steps of our algorithm does not exceed $\frac{\log((2k-1)/\zeta)}{\log(1 + \kappa)}$. At the last step of the algorithm, we find the set $A' = A_s \subseteq A$ such that $A'$ is $(\beta_1, \beta_2)$–connected of degree $k$ and $\zeta_k(A') \geq \zeta = \zeta_k(A)$. Thus $A'$ has the properties 1) and 3) of the theorem. Let us prove 2). Using (26), we obtain

$$|A'| \geq (1 - \beta_2)^s m \geq m \cdot 2^{\frac{\log((2k-1)/\zeta)}{\log(1 + \kappa)} \log(1 - \beta_2)}. \quad \text{p. 7}$$
This completes the proof.

**Corollary 2.12** Let $G$ be an Abelian group, $\varepsilon$, $\delta$ be real numbers, $\varepsilon \in (0, 1/8]$, $\delta \in (0, 1]$, $\delta \geq |G|^{-\varepsilon}$ and let $A \subseteq G$ be a set, $|A| \geq \delta|G| \geq 2$. Let also $p$ be a positive integer, $k = 2^p$, and $\beta_1, \beta_2 \in (0, 1)$ be real numbers, $\beta_1 \leq \beta_2$, $\beta_1 \leq 1 - |A|^{-2\varepsilon}$. Then there exists a set $A' \subseteq A$ such that
1) $A'$ is $(\beta_1, \beta_2)$-connected of degree $k$ and (16) is true for any $C \leq 1/32$.
2) $|A'| \geq |G| \cdot 2^{(\frac{3}{2} - 1) + 32\varepsilon} \frac{\log(1 - \beta_2)}{\log(1 - \beta_1)} + 1$.
3) $\zeta_k(A') \geq \zeta_k(A)$.

In particular, if $\beta_2 = \beta_1$, $k = 2$ and $\varepsilon = 1/8$ then the cardinality of $|A'|$ is at least $\delta^6|G|$.

**Proof.** Using Theorem 2.11 with $C = 1/32$, we find a set $A' \subseteq A$ with properties 1)–3) guaranteed by the theorem. Let us prove that $|A'| \geq |G| \cdot 2^{(\frac{3}{2} - 1) + 32\varepsilon} \frac{\log(1 - \beta_2)}{\log(1 - \beta_1)} + 1$. Let $N = |G|$, $m = |A|$, and $\zeta = \zeta_k(A)$. Clearly, $T_k(A) \geq \delta^2 N^{2k-1}$. Hence

$$\zeta \geq 2k - 1 + (2k - 1) \frac{\log(1/\delta)}{\log N} - \frac{2k}{1 - \varepsilon} \frac{\log(1/\delta)}{\log N}. \quad (27)$$

Since $\delta \geq N^{-\varepsilon}$ it follows that

$$2k - 1 - \zeta \leq \frac{\log(1/\delta)}{\log N} (1 + 4k\varepsilon) \quad \text{and} \quad \zeta \geq (2k - 1)(1 - 5\varepsilon). \quad (28)$$

By Theorem 2.11, we have $|A'| \geq m \cdot 2^k \frac{\log(1 - \beta_2)}{\log(1 - \beta_1)} \log(1 - \beta_2) + 1$, where $\kappa = \frac{\log(1 - \beta_1)^{-1}}{2 \log m}$. Using the last inequality, $\beta_1 \leq 1 - |A|^{-2\varepsilon}$ and (27), (28), we get

$$|A'| \geq m \cdot 2^k \frac{\log(1 - \beta_2)}{\log(1 - \beta_1)} \log(1 - \beta_2) \log m \geq N \cdot 2^k \frac{\log(1 - \beta_2)}{\log(1 - \beta_1)} + 1.$$ 

This completes the proof.

**Note 2.13** Certainly, the constant 32 at the second point of Corollary 2.12 can be decreased. The constant 2 in the numerator of $\frac{2}{2k-1}$ depends on an upper bound for $C$. If $C$ is less than 1/32 then the number 2 is also decreases.

**3. The proof of main result.**

**Lemma 3.1** Let $A$ be a finite nonempty set, and $k$ be a positive integer, $k \geq 2$. Then $T_k(A) \geq T_k(A)/|A|^{k-2}$.

**Proof.** The proof is trivial.

**The proof of Theorem 1.5** Let $m = |A|$, $\beta_1 = 1/2$, $\beta_2 = \beta_1 + 1/\log m$, $C = \varepsilon 2^{-7}$, $k = 2^p$, $p = [\log \ln m] + 1$. Clearly, $C \leq 1/32$. Using Theorem 2.11, we find $A' \subseteq A$ such that 1)–3) hold. By assumption $T_2(A) \geq |A|^3/K$. Using Lemma 3.1, we get $T_k(A) \geq T_k(A)/|A|^{k-2} \geq |A|^{2k-1}/K^{k-1}$. Whence

$$\zeta = \zeta_k(A) \geq 2k - 1 - (k - 1) \frac{\log K}{\log m}. \quad (29)$$

Using $K \leq m^\varepsilon$ and (29), we obtain

$$\zeta \geq (2k - 1) \left(1 - \frac{k - 1}{2k - 1} \frac{\log K}{\log m}\right) \geq (2k - 1) \left(1 - \frac{\varepsilon}{2}\right). \quad (30)$$

By 2) of Theorem 2.11, we have

$$|A'| \geq m \cdot 2^k \frac{\log(2k-1)/\zeta_k(A)}{\log(1 - \beta_2)^{-1}} = m2^{-\sigma}, \quad (31)$$
where \( \kappa = \frac{\log(1-\beta_1^{-1})}{\log m} (1-16C) \). Let us obtain an upper bound on \( \sigma \). Using simple inequalities 
\[
\log(1+x) \leq \frac{x}{\ln 2}, \quad \log(1+x) \geq \frac{1}{\ln 2} (x-x^2/2), \quad x \geq 0, \quad \text{and inequalities } m \geq 2^{32/\varepsilon}, \ (29), \ (30), \text{we get }
\]
\[
\sigma \leq \log \left(1 + \frac{2k-1-\zeta}{\zeta} \right) \frac{\ln 2}{\kappa} (1+\kappa) \log((1-\beta_2)^{-1}) \leq \frac{k-1}{2k-1} \frac{\log K}{\log m} (1+\varepsilon) \frac{\log m}{1-16C} \left(1 + \frac{1}{\log m} \right) \log((1-\beta_2)^{-1}) \leq \frac{k-1}{2k-1} \frac{\log K}{\log m} (1+\varepsilon) (1+32C) \left(1 + \frac{8}{\log m} \right) \leq \log K^{\frac{1}{2}} + \frac{\varepsilon}{16}.
\]
Hence \( |A'| \geq \frac{m}{K^{1/2+\varepsilon}} \). Since \( \zeta_k(A') \geq \zeta_k(A) \) it follows that
\[
T_k(A') \geq \frac{|A'|^{2k-1}}{K^{1/k}} \geq \frac{|A'|^{2k-1}}{K^k}.
\] (32)

Using \( |A'| \geq \frac{m}{K^{1/2+\varepsilon}}, \ (32) \) and \( K \leq (2^{-58} \varepsilon^{-4} \frac{|A|}{\log |A|})^{(3/2+\varepsilon)^{-1}} \) it is easy to see that \( T_k(A') \geq 2^{4kC-2k}k^k|A|^k \). Using Proposition 2.9, we find a set \( \Lambda \) such that \( |\text{Span } \Lambda \cap A'| \geq |A'|/2 \) and
\[
|A| \leq 2^{27} \varepsilon^{-2k} \left| \frac{|A|^2}{T_k^{1/k}(A')} \right|.
\] (33)

We have
\[
|\text{Span } \Lambda \cap A| \geq |\text{Span } \Lambda \cap A'| \geq \frac{|A'|}{2} \geq \frac{1}{2} \cdot \frac{m}{K^{1/2+\varepsilon}}.
\] (34)

Let us prove that \( |A| \leq 2^{30} \varepsilon^{-2} K \log m \). Combining (33) and (32), we get
\[
|A| \leq 2^{27} \varepsilon^{-2} K k^k |A|^k \leq 2^{27} \varepsilon^{-2} K k^k |A|^k \leq 2^{27} \varepsilon^{-2} K k^k |A|^k.
\]

Recall that \( k = 2^p, \ p = \lfloor \log \ln m \rfloor + 1 \), we finally obtain \( |A| \leq 2^{30} \varepsilon^{-2} K \log m \). This completes the proof.

**Corollary 3.2** Let \( G \) be an Abelian group. Let \( K, \varepsilon \) be real numbers, \( \varepsilon \in (0,1/2], \ A \subseteq G \) be an arbitrary set, \( |A| \geq 2^{32/\varepsilon}, \ 1 \leq K \leq \min\left\{ (2^{-58} \varepsilon^{-4} \frac{|A|}{\log |A|})^{(3/2+\varepsilon)^{-1}}, |A|^{\varepsilon} \right\} \). Let also \( |A+A| \leq K|A| \). Then there exists a set \( \Lambda \) such that \( |\text{Span } \Lambda \cap A| \geq \frac{1}{2} \cdot \frac{|A|}{K^{1/2+\varepsilon}} \) and \( |A| \leq 2^{30} \varepsilon^{-2} K \log |A| \).

**Proof.** We have \( |A+A| \leq K|A| \). By the Cauchy–Schwarz inequality
\[
|A|^4 = \left( \sum_x (A \ast A)(x) \right)^2 \leq \sum_x (A \ast A)(x)^2 \cdot |A+A| \leq T_2(A) \cdot K|A|.
\]
Hence \( T_2(A) \geq |A|^3/K \). Using Theorem 1.5, we obtain the required result.

**4. Another definitions of connectedness.**

In the section we discuss some relations between our definition of connectedness and a graph–theoretical definition of connectedness.

Suppose that \( \Gamma = (V, f) \) is a graph, where \( V \) is the set of vertices of \( \Gamma \) and \( f \) is the characteristic function of a symmetric subset of \( V \times V \). Let \( X, Y \subseteq V \) be arbitrary sets. By \( e(X,Y) \) denote the number of vertices between \( X \) and \( Y \). In other words \( e(X,Y) = \sum_{x \in X} \sum_{y \in Y} f(x,y) \). Recall that \( \Gamma \) is connected if for any vertex \( x \), we have \( e(x, V \setminus \{x\}) > 0 \). In [8] I. Ruzsa and G. Elekes gave the following definition.
Definition 4.1 Let $\alpha \in (0,1]$ be a real number. A graph $\Gamma = (V,f)$ is called $\alpha$–dense–connected if for any partition of the set of vertices into two disjoint parts, say $E \sqcup F = V$, we have
\[ e(E, F) \geq \alpha |E||F|. \]

We give an analog of the definition above for subsets of Abelian groups. Let $G$ be an Abelian group, and $A \subseteq G$ be a finite set. In papers [2, 3, 7] the graph of "popular differences" of $A$ was considered. This graph $\Gamma_A = (V_A, f_A)$ played a significant role in various problems of combinatorial number theory (see articles [2, 3, 7] and book [21]). The vertex set $V_A$ of the graph $\Gamma_A$ is $A$, and the function $f_A$ is the characteristic function of the symmetric set of "popular differences"
\[ f(x, y) = \begin{cases} 1, & \text{if } |\{x - y = a_1 - a_2 : a_1, a_2 \in A\}| \geq h, \\ 0, & \text{otherwise.} \end{cases} \]

Here $h$ is a number, $0 \leq h \leq |A|$. In many problems of combinatorial number theory $h$ was taken approximately $T_2(A)/|A|^2$. Thus the function $f(x, y)$ equals 1 if $(A \circ A)(x - y) \geq h$ and equals 0 otherwise. Ruzsa and Elekes applied $\alpha$–dense–connected subgraphs of $\Gamma_A$ to prove some results on sumsets (see details in [8]).

In the article we define a new (generalized) graph $\Gamma'_A = (V'_A, f'_A)$, where $f'_A$ is a symmetric function but not the characteristic function of some subset of $V'_A \times V'_A$. Put $V'_A := A$ and $f'_A(x, y) := (A \circ A)(x - y)$. The constructed graph $\Gamma'_A$ is an "approximation" of the graph $\Gamma_A$ in the sense that the function $f_A$ is a normalized and truncated version of the function $f'_A : f_A(x,y) = \theta(f'_A(x,y)/h)$, where $\theta$ is a shifted Heaviside's function: $\theta(x) = 1$ if $x \geq 1$ and $\theta(x) = 0$ if $x < 1$. Then the graph $\Gamma'_A$ is (generalized) $\alpha$–dense–connected if for any partition of the set of vertices into two disjoint parts $E$ and $V$, $E \sqcup F = A$, we have
\[ e(E, F) = \sum_{x \in E} \sum_{y \in F} (A \circ A)(x - y) = \sum_z (E \circ F)(z) \cdot (A \circ A)(z) \geq \alpha |E||F|. \] (35)

We shall call a set $A$ is strongly connected if inequality (35) holds. As was noted above in many problems of combinatorial number theory the order of the number $h$ was $T_2(A)/|A|^2$. We also put $\alpha = C \cdot T_2(A)/|A|^2$, where $C > 0$ is a constant.

Definition 4.2 Let $k \geq 2$ be a positive integer. An arbitrary nonempty finite set $A \subseteq G$ is called strongly connected of degree $k$ if there is an absolute constant $C \in (0,1]$ such that for any disjoint sets $E, F \subseteq A$, $E \sqcup F = A$, we have
\[ \sum_x (E \circ F)(x) \cdot ((A *_{k-2} A) \circ (A *_{k-2} A))(x) \geq C e_{E} c_{F} T_{k}(A), \] (36)
where $e_{E} = |E|/|A|$, $c_{F} = |F|/|A|$.

First of all let us show that any strongly connected set is a connected set.

Statement 4.3 Let $p$ be a positive integer, and $k = 2^p$. Suppose that $A$ is a strongly connected of degree $k$ set such that (36) holds with some constant $C$. Then $A$ is connected of degree $k$ and inequality (1) holds with $C/8$.

Proof. If the cardinality of $A$ is less than two then there is nothing to prove. Let $|A| \geq 3$, $B$ be an arbitrary subset of $A$, and $\overline{B} = A \setminus B$. Let also
\[ \sigma = \sum_{x} (B \circ \overline{B})(x) \cdot ((A *_{k-2} A) \circ (A *_{k-2} A))(x). \] (37)
Using Lemma 2.3, we get \( \sigma(38) \), we obtain,

\[ i < j \]

Then there is an element \( a \) generating by \( H \). If there are two nonempty intersections of cosets of \( H \) with \( A \), say, \( A_i \) and \( A_j \), \( i < j, i, j \in \{1, \ldots, r\} \) then put \( E = \bigcup_{i=1}^{n} A_i \) and \( F = A \setminus E \). Clearly, \( E \) and \( F \) are nonempty sets. Since for any \( e \in E \) and \( f \in F \), we have \( e - f \notin H \), and, consequently, \( e - f \notin S \) it follows that

\[
\sum_{x}(E \circ F)(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) \leq \sum_{x \notin S}(E \circ F)(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) < C|E||F| \cdot \frac{T_k(A)}{|A|^2}.
\]

Since \( A \) is a strongly connected of degree \( k \) it follows that

\[
\sigma \geq C \frac{|B| \sqrt{|B|}}{|A|} \cdot T_k(A).
\]

We have

\[
\sigma = \sum_{x} (B \ast A \ast_{k-2} A)(x) \cdot (\overline{B} \ast A \ast_{k-2} A)(x).
\]

Using Lemma 2.3, we get \( \sigma^{2k} \leq T_k(B)T_k(\overline{B})T_k^{2k-2}(A) \). Combining the last inequality and (38), we obtain

\[
T_k(B)T_k(\overline{B}) \geq C^{2k} \frac{|B|^{2k}}{|A|^{2k}} \cdot \frac{|\overline{B}|^{2k}}{|A|^{2k}} T_k^2(A).
\]

If \( |B| \leq |A|/2 \) then \( |\overline{B}| \geq |A|/2 \). Using this lower bound for \( |\overline{B}| \), we get

\[
T_k(B) \geq \left( \frac{C}{2} \right)^{2k} \left( \frac{|B|}{|A|} \right)^{2k} T_k(A)
\]

and the statement is proved. Suppose that \( |B| > |A|/2 \). Then let \( B_1 \) be an arbitrary subset of \( B \) of the cardinality \( |A|/2 \). Clearly, \( |B| \leq 4|B_1| \). By (41), we have

\[
T_k(B) \geq T_k(B_1) \geq \left( \frac{C}{2} \right)^{2k} \left( \frac{|B_1|}{|A|} \right)^{2k} T_k(A) \geq \left( \frac{C}{8} \right)^{2k} \left( \frac{|B|}{|A|} \right)^{2k} T_k(A).
\]

This completes the proof.

Thus any strongly connected set is connected. In particular, Proposition 2.7 is true for an arbitrary strongly connected set and therefore any strongly connected set is contained in \( \text{Span} \ A \) for some \( A \) with small cardinality. Apparently, it was S.V. Konyagin (see [20]) who first proved that an arbitrary strongly connected set is economically contained in some special subgroup (see also another variant of his statement in book [21] p. 114, ex. 2.6.10). We formulate his result in our terms and give the proof for the sake of completeness.

**Statement 4.4** Let \( k \geq 2 \) be a positive integer. Let also \( A \subseteq G \) be a strongly connected of degree \( k \) set such that (36) holds with some constant \( C \). Let

\[
S = \left\{ h \in G : ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) \geq C \frac{T_k(A)}{|A|^2} \right\}.
\]

Then there is an element \( a \in G \) such that \( A \subseteq \langle S \rangle + a \), where \( \langle S \rangle \) is the subgroup of \( G \) generating by \( S \).

**Proof.** Assume the converse. Let \( H = \langle S \rangle \) and let \( A_1, \ldots, A_r \subseteq A \) be intersections of \( A \) with cosets of \( H \). If there are two nonempty intersections of cosets of \( H \) with \( A \), say, \( A_i \) and \( A_j \), \( i < j \), \( i, j \in \{1, \ldots, r\} \) then put \( E = \bigcup_{i=1}^{n} A_i \) and \( F = A \setminus E \). Clearly, \( E \) and \( F \) are nonempty sets. Since for any \( e \in E \) and \( f \in F \), we have \( e - f \notin H \), and, consequently, \( e - f \notin S \) it follows that

\[
\sum_{x}(E \circ F)(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) \leq \sum_{x \notin S}(E \circ F)(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) < C|E||F| \cdot \frac{T_k(A)}{|A|^2}.
\]
with contradiction. This completes the proof.

We prove an analog of Theorem 2.11 for strongly connected sets.

Let \( E, F \subseteq A \) be sets. Denote by \( e(E, F) \) the quantity \( \sum_x (E \circ F)(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) \). Clearly, \( e(E \cup E, F) = e(E, F) + e(E, F) \). Suppose that \( E \subseteq A \) is an arbitrary set. By \( c_E \) denote the ratio \( |E|/|A| \).

We need in the following technical definition of strongly connected of degree \( k \) sets.

**Definition 4.5** Let \( k \geq 2 \) be a positive integer. An arbitrary nonempty finite set \( A \subseteq G \) is called \( \beta - \text{strongly connected of degree } k \) if there is an absolute constant \( C \in (0, 1] \) and a set \( B \subseteq A, |B| \geq \beta |A| \) such that for any disjoint sets \( E, F \subseteq B, E \cup F = B \), we have

\[
\sum_{x} (E \circ F)(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) \geq Cc_Ec_FT_k(A).
\]

(42)

Our next statement can be proved likewise Statement 4.3.

**Statement 4.6** Let \( p \) be a positive integer, \( k = 2^p \), and \( \beta \in [0, 1] \) be a real number. Let \( A \) be a strongly \( \beta - \text{connected of degree } k \) set and (42) is true with some \( C \) and some \( B \subseteq A, |B| \geq \beta |A| \). Then \( B \) is connected of degree \( k \) set and inequality (1) holds with \( C\beta^2/8 \).

A graph–theoretical variant of Lemma 4.7 below was proved in [8].

**Lemma 4.7** Let \( k \geq 2 \) be a positive integer, \( \varepsilon_1 \in [0, 1] \) be a real number, and let \( A \subseteq G \) be a finite set. Then there exists a partition of \( A \) into disjoint sets \( A_1, \ldots, A_l \) such that

1) For all \( i, j \in \{1, \ldots, l\}, i \neq j \), we have \( e(A_i, A_j) \leq \varepsilon_1 c_{A_i} c_{A_j} T_k(A) \).

2) For any \( i \in \{1, \ldots, l\} \) the set \( A_i \) has the following property: for any disjoint sets \( E, F \subseteq A_i, E \cup F = A_i \), we have \( e(E, F) \geq \varepsilon_1 c_E c_FT_k(A) \).

Besides, the following inequality holds

3) \( \sum_{i=1}^l T_k(A_i) \geq T_k(A) \cdot (1 - (2k - 1)\varepsilon_1) \).

**Proof.** Consider all partitions of \( A \) into disjoint subsets \( A_1, \ldots, A_s \), where \( s \) is an arbitrary positive integer. Select one for which the sum

\[
\sigma(A_1, \ldots, A_s) = \sum_{1 \leq i < j \leq s} (e(A_i, A_j) - \varepsilon_1 c_{A_i} c_{A_j} T_k(A))
\]

(43)

is minimal. By minimality of this partition, say \( \{A_1, \ldots, A_l\} \), we have 2).

Let us prove 1). Suppose that for some \( i, j \in \{1, \ldots, l\}, i \neq j \) the following holds \( e(A_i, A_j) > \varepsilon_1 c_{A_i} c_{A_j} T_k(A) \). Constructing the new partition \( \mathcal{P} \) of the set \( A \), \( \mathcal{P} = \{A_r\}_{r \neq i,j} \cup (A_i \cup A_j) \) and using the last inequality, we get

\[
\sigma(\mathcal{P}) = \sigma(A_1, \ldots, A_s) - (e(A_i, A_j) - \varepsilon_1 c_{A_i} c_{A_j} T_k(A)) < \sigma(A_1, \ldots, A_s).
\]

with contradiction.

Prove that 1) implies 3). Indeed,

\[
T_k(A) = \sum_x (A \ast_{k-1} A)^2(x) = \sum_{i,j=1}^l \sum_x (A_i \ast A \ast_{k-2} A)(x) \cdot (A_j \ast A \ast_{k-2} A)(x) =
\]

\[
= \sum_{i=1}^l \sum_x (A_i \ast A \ast_{k-2} A)(x) \cdot (A_i \ast A \ast_{k-2} A)(x) + \sum_{i,j=1}^l \sum_{i \neq j} x (A_i \ast A \ast_{k-2} A)(x) \cdot (A_j \ast A \ast_{k-2} A)(x)
\]

\[
= \sum_{i=1}^l \sum_x (A_i \ast A \ast_{k-2} A)(x) \cdot (A_i \ast A \ast_{k-2} A)(x) + \sum_{i,j=1}^l \sum_{i \neq j} x (A_i \circ A_j)(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x)
\]

12
And so on. Finally, we obtain

\[ \sum_{i=1}^{l} \sum_{x} (A_{i} \ast A \ast_{k-2} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) + \varepsilon_{1} \sum_{i,j=1}^{l} c_{A_{i}} c_{A_{j}} T_{k}(A) \leq \]

\[ \leq \sum_{i=1}^{l} \sum_{x} (A_{i} \ast A \ast_{k-2} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) + \varepsilon_{1} T_{k}(A) . \]

Hence \( \sum_{i=1}^{l} \sum_{x} (A_{i} \ast A \ast_{k-2} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) \geq (1 - \varepsilon_{1}) T_{k}(A) \). Similarly,

\[ \sum_{i=1}^{l} \sum_{j=1}^{l} \sum_{x} (A_{i} \ast A_{j} \ast A \ast_{k-3} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) \leq \]

\[ \leq \sum_{i=1}^{l} \sum_{x} (A_{i} \ast A \ast_{k-3} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) + \]

\[ + \sum_{i=1}^{l} \sum_{j=1}^{l} \sum_{x} (A \ast A_{j} \ast A \ast_{k-3} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) \leq \]

\[ \leq \sum_{i=1}^{l} \sum_{x} (A_{i} \ast A \ast_{k-3} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) + \]

\[ + \sum_{i=1}^{l} \sum_{j=1, j \neq i}^{l} \sum_{x} (A_{i} \circ A_{j})(x) \cdot ((A \ast_{k-2} A) \circ (A \ast_{k-2} A))(x) \leq \]

\[ \leq \sum_{i=1}^{l} \sum_{x} (A_{i} \ast A \ast_{k-3} A)(x) \cdot (A_{i} \ast A \ast_{k-2} A)(x) + \varepsilon_{1} T_{k}(A) . \]

And so on. Finally, we obtain

\[ \sum_{i=1}^{l} T_{k}(A_{i}) = \sum_{i=1}^{l} \sum_{x} (A_{i} \ast_{k-1} A_{i})(x) \cdot (A_{i} \ast_{k-1} A_{i})(x) \geq (1 - (2k - 1)\varepsilon_{1}) \cdot T_{k}(A) . \]

This concludes the proof.

**Note 4.8** It is easy to see that the third property of the constructed partition implies that there is \( i_{0} \in \{1, \ldots, l\} \) such that \( |A_{i_{0}}| \geq (1 - (2k - 1)\varepsilon_{1}) \cdot m^{\frac{\zeta(A)-1}{2k-2}} \geq (1 - (2k - 1)\varepsilon_{1}) \cdot m^{1/2} \).

In fact, we have

\[ m^{\zeta(A)}(1 - (2k - 1)\varepsilon_{1}) \leq \sum_{i=1}^{l} T_{k}(A_{i}) \leq (\max_{i=1, \ldots, l} |A_{i}|)^{2k-2} \sum_{i=1}^{l} |A_{i}| \leq (\max_{i=1, \ldots, l} |A_{i}|)^{2k-2} m . \]

This yields that if we put \( \beta = (1 - (2k - 1)\varepsilon_{1}) m^{-1/2} \) then any set \( A \subseteq G, |A| = m \) is strongly \( \beta \)--connected of degree \( k \) and inequality (42) holds with any \( C \leq 1/(2k - 1) \). Thus to obtain nontrivial results on the structure of \( A \) one should prove that \( A \) is strongly \( \beta \)--connected for large \( \beta \).

**Theorem 4.9** Let \( A \subseteq G \) be a set. Let also \( \varepsilon, \beta \in (0, 1) \) be real numbers, and \( |A| \geq \varepsilon/(2\beta^{2}) \). Then there exists a partition of \( A \) into disjoint sets \( A_{1}, \ldots, A_{t}, \Omega \) such that

1) Any set \( A_{i}, i = 1, \ldots, t \) is strongly \( \beta \)--connected of degree 2 and inequality (42) holds with
any $C \leq \varepsilon \log(1/\beta)/(3 \log(2|A|/\varepsilon))$.

2) $\sum_{i=1}^{t} T_2(A_i) \geq (1 - \varepsilon) \cdot T_2(A)$.

**Proof.** Let $m = |A|$, $s_0 = \log(2m/\varepsilon)/(2 \log(1/\beta)) \geq 1$, and $\varepsilon' = \varepsilon/(6s_0)$. Let $C \leq \varepsilon'$ be a real number. The proof of Theorem 4.9 is a sort algorithm. If $A$ is strongly $\beta$—connected of degree 2 and (42) is true with the constant $C$ then there is nothing to prove. Suppose that $A$ is not strongly $\beta$—connected of degree 2 set (with the constant $C$). Using Lemma 4.7 with $\varepsilon_1 = \varepsilon'$, we get the partition $\mathcal{P}^{(1)}$ of $A$ into $A_1, \ldots, A_t$ satisfy properties 1) — 3) of the lemma. Since $A$ is not strongly $\beta$—connected of degree 2 it follows that for any $l \in \{1, \ldots, l\}$, we have $|A_i| < \beta|A|$. Using the third property of the partition $\mathcal{P}^{(1)}$, we obtain

$$\sum_{A \in \mathcal{P}^{(1)}} T_2(A) = \sum_{i=1}^{l} T_2(A_i) \geq (1 - 3\varepsilon')T_2(A).$$

Let $B^{(1)} = \{A_i \mid \text{is not strongly } \beta \text{—connected of degree 2}\}$, and $G^{(1)}$ be the collection of all other sets of the partition $\mathcal{P}^{(1)}$. Let us construct a new partition of $A$. We do not change the sets $A_i$ from $G^{(1)}$. Further, for any $A_i$ belongs to $B^{(1)}$, we use Lemma 4.7 with $\varepsilon_1 = \varepsilon'$. We get a new partition of $A_i$ into subsets $A_{ij}$, $j \in \{1, \ldots, l(i)\}$. So we construct a new partition $\mathcal{P}^{(2)}$ of the set $A$. For any $A_i \in B^{(1)}$ the following holds $\sum_{j=1}^{l(i)} T_2(A_{ij}) \geq (1 - 3\varepsilon')T_2(A_i)$. Hence

$$\sum_{A \in \mathcal{P}^{(2)}} T_2(A) \geq (1 - 3\varepsilon')T_2(A).$$

Let $B^{(2)} = \{A_{ij} \mid \text{is not strongly } \beta \text{—connected of degree 2}\}$. For an arbitrary $A_{ij} \in B^{(2)}$, we use Lemma 4.7 with $\varepsilon_1 = \varepsilon'$. We get a new partitions of the sets $A_{ij}$ into disjoint subsets $A_{ijr}$. And so on. At $s$—th step of the algorithm, we construct the partition $\mathcal{P}^{(s)}$ such that

$$\sum_{A \in \mathcal{P}^{(s)}} T_2(A) \geq (1 - 3\varepsilon'^s)T_2(A) \geq (1 - 3\varepsilon's) \cdot T_2(A).$$

It is easy to see that if for some $s \leq s_0$ the following holds

$$\sum_{A \in \mathcal{P}^{(s)} \setminus B^{(s)}} T_2(A) \geq (1 - \varepsilon) \cdot T_2(A),$$

then we are done. Indeed, just put $\Omega = \bigcup_{A \in B^{(s)}} A$. Suppose that for all $s \leq s_0$ inequality (46) does not hold. Using inequality $s \leq s_0$ and (45), we get $\sum_{A \in \mathcal{P}^{(s)}} T_2(A) \geq (1 - \varepsilon/2) \cdot T_2(A).$ Hence

$$\sum_{A \in B^{(s)}} T_2(A) \geq \frac{\varepsilon}{2} \cdot T_2(A) \geq \frac{\varepsilon m^2}{2}.$$ 

For any $A \in B^{(s)}$, we have $|A| < \beta^s m$. Whence

$$\sum_{A \in B^{(s)}} T_2(A) < (\beta^s m)^2 \sum_{A \in \mathcal{P}^{(s)}} |A| = \beta^{2s} m^3.$$ 

If $s = s_0$ then the last inequality contradicts (47). So for some $s < s_0$ inequality (46) holds. This completes the proof.

There is a difference between Theorem 2.11 and Theorem 4.9. In Theorem 4.9 we prove that there exists a partition of $A$ into strongly $\beta$—connected components and some exceptional set $\Omega$ while Theorem 2.11 states that there is one connected subset of $A$. Besides, Theorem 4.9 implies that the remaining set $\Omega$ has small $T_2(\Omega)$. Indeed, by the property 2), we have $\sum_{i=1}^{t} T_2(A_i) \geq (1 - \varepsilon) \cdot T_2(A)$, whence $T_2(\Omega) \leq \varepsilon T_2(A)$. 

14
References

[1] Balog A., Szemerédi E. A statistical theorem of set addition // Combinatorica 14 (1994), 263–268.

[2] Gowers W. T. A new proof of Szemerédi’s theorem for arithmetic progressions of length four // Geom. Funct. Anal. 8 (1998), 529–551.

[3] Gowers W. T. A new proof of Szemerédi’s theorem // Geom. Funct. Anal. 11 (2001), 465–588.

[4] Freiman G. A. Foundations of a Structural Theory of Set Addition / Kazan Gos. Ped. Inst., Kazan, 1966, in Russian.

[5] Bilu Y. Structure of sets with small sumset // Structure Theory of Sets Addition, Astérisque, Soc. Math. France, Montrouge 258 (1999), 77–108.

[6] Ruzsa I. Generalized arithmetic progressions and sumsets // Acta Math. Hungar. 65 (1994), 379–388.

[7] Chang M.–C., A polynomial bound in Freiman’s theorem // Duke Math. J. 113 (2002) no. 3, 399–419.

[8] Elekes G., Ruzsa I. The structure of sets with few sums along a graph // http://www.cs.elte.hu/~elekes/Abstracts/alog.ps, submitted for publication.

[9] Ruzsa I. An analog of Freiman’s theorem in groups // Structure theory of set addition // Astérisque No. 258 (1999), 323–326.

[10] Green B., Ruzsa I. An analogue of Freiman’s theorem in an arbitrary abelian group // J. London Math. Soc., submitted for publication.

[11] Green B. Spectral structure of sets of integers // Fourier analysis and convexity (survey article, Milan 2001), Appl. Numer. Harmon. Anal., Birkhauser Boston, Boston, MA (2004), 83–96.

[12] Green B. Finite field model in additive combinatorics // Surveys in Combinatorics 2005, LMS Lecture Notes 329, 1–29.

[13] Green B. The polynomial Freiman–Ruzsa conjecture // http://www.dpmms.cam.ac.uk/~bjg23.

[14] Green B. Boolean functions with small spectral norm // Geom. Funct. Anal., submitted for publication.

[15] Green B. An inverse theorem for the Gowers $U^3$–norm, with applications // Proc. Edin. Math. Soc., submitted for publication.

[16] Green B. A note on the Freiman and Balog–Szemerédi–Gowers theorems in finite fields // http://www.arXiv:math.CO/0701585 v1, submitted for publication.

[17] Sanders T. A note on Freiman’s theorem in vector spaces // http://www.arXiv:math.NT/0605523.
[18] Rudin W. Fourier analysis on groups / Wiley 1990 (reprint of the 1962 original).

[19] Rudin W. Trigonometric series with gaps // J. Math. Mech. 9 (1960), 203–227.

[20] Bourgain J., Konygin S. Estimates for the number of sums and products and for exponential sums over subgroups in fields of prime order // C. R. Acad. Sci. Paris, Ser. I 337 (2003), 75–80.

[21] Tao T., Vu V. Additive combinatorics / Cambridge University Press 2006.