Discrepancy of Sums of two Arithmetic Progressions

Nils Hebbinghaus

Max-Planck-Institut für Informatik,
Saarbrücken, Germany.

Abstract

Estimating the discrepancy of the hypergraph of all arithmetic progressions in the set $[N] = \{1, 2, \ldots, N\}$ was one of the famous open problems in combinatorial discrepancy theory for a long time. An extension of this classical hypergraph is the hypergraph of sums of k ($k \geq 1$ fixed) arithmetic progressions. The hyperedges of this hypergraph are of the form $A_1 + A_2 + \ldots + A_k$ in $[N]$, where the A_i are arithmetic progressions. For this hypergraph Hebbinghaus (2004) proved a lower bound of $\Omega(N^{k/(2k+2)})$. Note that the probabilistic method gives an upper bound of order $O((N \log N)^{1/2})$ for all fixed k. Přívětivý improved the lower bound for all $k \geq 3$ to $\Omega(N^{1/2})$ in 2005. Thus, the case $k = 2$ (hypergraph of sums of two arithmetic progressions) remained the only case with a large gap between the known upper and lower bound. We bridge this gap (up to a logarithmic factor) by proving a lower bound of order $\Omega(N^{1/2})$ for the discrepancy of the hypergraph of sums of two arithmetic progressions.

1   Introduction

A finite hypergraph $\mathcal{H} = (V, \mathcal{E})$ consists of a finite set V and a set $\mathcal{E}$ of subsets of V. The elements of V are called vertices and those of $\mathcal{E}$ hyperedges of the hypergraph $\mathcal{H}$. If we 2–partition the set of vertices V, this 2–partition

1

clearly induces a 2–partition in every hyperedge $E \in \mathcal{E}$. The discrepancy of $\mathcal{H}$ is a non-negative integer that indicates how ballanced $\mathcal{H}$ can be 2–partitioned with respect to all its hyperedges $E \in \mathcal{E}$. We make this more precise and express the 2–partition through a coloring $\chi \colon V \to \{-1, 1\}$ of the vertices of $\mathcal{H}$ with the two "colors" -1 and 1. Now for every hyperedge $E \in \mathcal{E}$ the imbalance due to the coloring χ can be calculated as follows. Let $\chi(E) := \sum_{x \in E} \chi(x)$. Then $|\chi(E)|$ is the absolute difference between the number of vertices in E colored with "color" -1 and the number of vertices in E colored with "color" 1. The discrepancy of $\mathcal{H}$ with respect to the (specific) coloring χ is defined as

$$\mathrm{disc}(\mathcal{H}, \chi) := \max_{E \in \mathcal{E}} |\chi(E)|.$$

In other words, $\mathrm{disc}(\mathcal{H}, \chi)$ is the maximal imbalance of any hyperedge $E \in \mathcal{E}$ under the coloring χ. Now the discrepancy of $\mathcal{H}$ is defined as

$$\mathrm{disc}(\mathcal{H}) := \min_{\chi \colon V \to \{-1, 1\}} \mathrm{disc}(\mathcal{H}, \chi),$$

where the minimum is taken over all $2^{|V|}$ possible colorings $\chi \colon V \to \{-1, 1\}$ of the set of vertices V. Thus, $\mathrm{disc}(\mathcal{H})$ is the least possible imbalance of any hyperedge $E \in \mathcal{E}$ that can not be avoided under any coloring $\chi \colon V \to \{-1, 1\}$.

One of the famoust long-standing open problems in (combinatorial) discrepancy theory was to determine the right order for the discrepancy of the hypergraph of arithmetic progressions in the first N natural numbers ($N \in \mathbb{N}$). Before we give a brief overview over the history of this problem, we introduce the hypergraph $\mathcal{H}_{AP}$ of arithmetic progressions. For convenience, let us define for every interval $I \subseteq \mathbb{R}$ the set

$$I_{\mathbb{Z}} := \{z \in \mathbb{Z} \mid z \in I\}$$

of all integers in the intervall I. In particular, we introduce the abreviation

$$[x] := [1, x]_{\mathbb{Z}}$$

for the set of all natural numbers n with $1 \le n \le x$ ($x \in \mathbb{R}$). Let $N \in \mathbb{N}$. An arithmetic progression in $[N]$ is a subset of $[N]$ of the form

$$A_{a, \delta, L} := \{a + j\delta \mid j \in [0, L-1]_{\mathbb{Z}}\}.$$

2

Now we can define the hypergraph $\mathcal{H}_{AP} = ([N], \mathcal{E}_{AP})$. The set of vertices of $\mathcal{H}_{AP}$ is the set $[N]$ and the set of hyperedges is

$$\mathcal{E}_{AP} := \{A_{a,\delta,L} \mid a, \delta \in [N], L \in [\frac{N-a}{\delta} + 1]\},$$

where $L \in [\frac{N-a}{\delta} + 1]$ just ensures that $A_{a,\delta,L} \subseteq [N]$.

In 1964, Roth [R64] proved a lower bound for the discrepancy of the hypergraph $\mathcal{H}_{AP}$ of order $\Omega(N^{\frac{1}{4}})$. Using a random coloring of the vertices of $\mathcal{H}_{AP}$, one can easily show an upper bound of order $O((N \log N)^{\frac{1}{2}})$ for the discrepancy of the hypergraph $\mathcal{H}_{AP}$. The first non-trivial upper bound is due to Sárközy. 1973 he proved $\mathrm{disc}(\mathcal{H}_{AP}) = O(N^{1/3} \log^{1/3} N)$ A sketch of his beautiful proof can be found in the book *Probabilistic Methods in Combinatorics* by Erdős and Spencer [ES74]. Beck [B81] showed in 1981 (inventing the famous partial coloring method) that Roth's lower bound is almost sharp. His upper bound of order $O(N^{1/4} \log^{5/4} N)$ was finally improved by Matoušek and Spencer [MS96] in 1996. They showed by a refinement of the partial coloring method — the entropy method — that the discrepancy of the hypergraph $\mathcal{H}_{AP}$ is exactly of order $\Theta(N^{1/4/})$.

Therefore, after 32 years, this open problem was solved. In the next years several extensions of this discrepancy problem were studied. Doerr, Srivastav and Wehr [DSW] determined the discrepancy of d–dimensional arithmetic progressions. For the hypergraph $\mathcal{H}_{AP,d} = ([N]^d, \mathcal{E}_{AP,d})$, where $\mathcal{E}_{AP,d} := \{\prod_{i=1}^d E_i \mid E_i \in \mathcal{E}_{AP}\}$, they proved $\mathrm{disc}(\mathcal{H}_{AP,d}) = \Theta(N^{d/4})$. Another related hypergraph — the hypergraph of all 1–dimensional arithmetic progressions in the d–dimensional grid $[N]^d$ was studied by Valko [V2002]. He proved for the discrepancy of this hypergraph a lower bound of order $\Omega(N^{d/(2d+2)})$ and an upper bound of order $O(N^{d/(2d+2)} \log^{5/2} N)$.

The hypergraph that we consider in this paper was introduced by Hebbinghaus [H2004] in a generalized version. Let $k \in \mathbb{N}$ and $N \in \mathbb{N}$. The hypergraph $\mathcal{H}_{kAP} = ([N], \mathcal{E}_{kAP})$ of sums of k arithmetic progressions is defined as follows. The vertices of $\mathcal{H}_{kAP}$ are the first N natural numbers. And the set of hyperedges $\mathcal{E}_{kAP}$ is defined as

$$\mathcal{E}_{kAP} := \left\{ \left(\sum_{i=1}^k A_{a_i, \delta_i, L_i} \right) \cup [N] \,\middle|\, a_i \in \mathbb{Z}, \delta_i, L_i \in [N] (i \in [k]) \right\},$$

where the sum of k sets M_i $(i \in [k])$ is

$$\sum_{i=1}^{k} M_i = \left\{ \sum_{i=1}^{k} \middle| m_i \in M_i (i \in [k]) \right\}.$$

For the hypergraph $\mathcal{H}_{kAP}$ of sums of k arithmetic progressions in $[N]$ Hebbinghaus [H2004] proved a lower bound of order $\omega(N^{k/(2k+2)})$ in 2004. But there remained a large gap between this bound and the upper bound of order $O(N^{1/2} \log^{1/2} N)$ from the random coloring method. In 2006 Přívětivý [P2006] nearly closed this gap for $k \geq 3$ by proving a lower bound of order $\Omega(N^{1/2})$ for the discrepancy of the hypergraph $\mathcal{H}_{3AP}$ of sums of three arithmetic progressions. This lower bound clearly extends to all hypergraphs $\mathcal{H}_{kAP}$ for all $k \geq 3$. Thus, the case $k = 2$ was the last with a large gap between the lower and the upper bound for the discrepancy. In this paper we improve the lower bound for the discrepancy of the hypergraph $\mathcal{H}_{2AP}$ of sums of two arithmetic progressions from the order $\Omega(N^{1/3})$ to the order $\Omega(N^{1/2})$. This result shows that the upper bound of order $O(N^{1/2} \log^{1/2} N)$ for the discrepancy of $\mathcal{H}_{2AP}$ determined by the random coloring method is almost sharp. We will prove the following theorem.

Theorem 1. *Let $N \in \mathbb{N}$. For the hypergraph $\mathcal{H}_{2AP}$ of sums of two arithmetic progressions we obtain the following bounds.*

(i) $\mathrm{disc}(\mathcal{H}_{2AP}) = \Omega(N^{1/2})$.

(ii) $\mathrm{disc}(\mathcal{H}_{2AP}) = O(N^{1/2} \log^{1/2} N)$.

Since $|\mathcal{E}_{2AP}| = O(N^6)$ the second assertion is a direct consequence of the general upper bound for a hypergraph $\mathcal{H}$ with n vertices and m hyperedges $\mathrm{disc}(\mathcal{H}) = O(\sqrt{n \log m})$ derived by the random coloring method.

2   A Special Set of Hyperedges

In this section we define a special subset $\mathcal{E}_0$ of the set $\mathcal{E}_{2AP}$ of all sums of two arithmetic progressions in $[N]$. The elements of this set $\mathcal{E}_0$ and all their

translates build the set of hyperedges in which we will find for every coloring $\chi\colon V \to \{-1, 1\}$ a hyperedge with discrepancy of order $\Omega(N^{\frac{1}{2}})$.

All elements of $\mathcal{E}_0$ are sums of two arithmetic progressions with starting point 0. Thus, we can characterize them by the difference and length of the two arithmetic progressions. We define for all $\delta_1, \delta_2, L_1, L_2 \in \mathbb{N}$:

$$E_{\delta_1, L_1, \delta_2, L_2} := \{j_1\delta_1 + j_2\delta_2 \mid j_1 \in [0, L_1 - 1], j_2 \in [0, L_2 - 1]\}.$$

Before specifying the set $\mathcal{E}_0$ we should mention that due to a case distinction in the proof of the Main Lemma the set $\mathcal{E}_0$ is the union of three subsets $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_3$, each of them corresponding to one of the cases. The first two sets $\mathcal{E}_1$ and $\mathcal{E}_2$ are easy to define. We set

$$\mathcal{E}_1 := \{E_{\delta_1, L_1, \delta_2, L_2} \mid \delta_1 \in [24], L_1 = \lceil \tfrac{N}{6\delta_1} \rceil, \delta_2 = 1, L_2 = 1\},$$

and

$$\mathcal{E}_2 := \{E_{\delta_1, L_1, \delta_2, L_2} \mid \delta_1 \in [25, N^{\frac{1}{2}}]_{\mathbb{Z}}, L_1 = \lceil \tfrac{N}{12\delta_1} \rceil, \delta_2 \in [\delta_1 - 1], L_2 = \lceil \tfrac{\delta_1 - 1}{12} \rceil \}.$$

The definition of the last set $\mathcal{E}_3$ is not straightforward. For every difference δ_1 of the first arithmetic progression we have to determine a set of differences δ_2 for the second arithmetic progression. Let $\delta_1 \in [N^{\frac{1}{2}}]$ and let

$$B(\delta_1) := \{b \in [\delta_1] \mid (b, \delta_1) = 1\}$$

be the set of all elements of $[\delta_1]$ that are relatively prime to δ_1. Here (b, δ_1) denotes the greatest common divisor of b and δ_1. Let $b \in B(\delta_1)$. Set $\bar{k} := \lfloor \log(N^{\frac{1}{2}}\delta_1^{-1}) \rfloor$. We define for all $0 \leq k \leq \bar{k}$ sets $M(b, k)$ of distances for the second arithmetic progression. The set $M(b, k)$ should cover the range of possible differences for the second arithmetic progression for the interval $(2^k N^{\frac{1}{2}}, 2^{k+1} N^{\frac{1}{2}}]$. We define

$$M(b, k) := (b + 2^{2k}\delta_1 \mathbb{Z}) \cap (2^k N^{\frac{1}{2}}, 2^{k+1} N^{\frac{1}{2}} + 2^{2k}\delta_1).$$

For all $0 \leq k \leq \bar{k}$, we set $M_{\delta_1}(k) := \bigcup_{b \in B(\delta_1)} M(b, k)$. Now we are able to define the third set $\mathcal{E}_3$. Let

$$\mathcal{E}_3 := \bigcup_{\delta_1 \in [N^{\frac{1}{2}}]} \bigcup_{k=0}^{\bar{k}} \left\{ E_{\delta_1, L_1, \delta_2, L_2} \mid L_1 = \left\lceil \tfrac{2^k N^{\frac{1}{2}}}{12} \right\rceil, \delta_2 \in M_{\delta_1}(k), L_2 = \left\lceil \tfrac{2^{-k} N^{\frac{1}{2}}}{12} \right\rceil \right\}.$$

5

In the next lemma we prove that the cardinality of the set $\mathcal{E}_0$ is of order $O(N)$. This is an essential property of the set $\mathcal{E}_0$ for the proof of the lower bound of the discrepancy of the hypergraph of sums of two arithmetic progressions.

Lemma 2. *We have $|\mathcal{E}_3| \leq 6N$ and thus $|\mathcal{E}_0| \leq 7N$.*

Proof. We have to estimate $|\mathcal{E}_3| = \sum_{\delta_1 \in \left[N^{\frac{1}{2}}\right]} \sum_{k=0}^{\bar{k}} |M_{\delta_1}(k)|$. For this pupose we look for $|M(b,k)|$ for all $b \in B(\delta_1)$ and all $0 \leq k \leq \bar{k}$. We first show that the difference $2^{2k}\delta_1$ of two consecutive elements of $M(b,k)$ is at most $2^k N^{\frac{1}{2}}$.

$$2^{2k}\delta_1 \leq 2^k 2^{\log\left(N^{\frac{1}{2}}\delta_1^{-1}\right)}\delta_1 = 2^k N^{\frac{1}{2}}.$$

Hence,

$$|M(b,k)| \leq \frac{3 \cdot 2^k N^{\frac{1}{2}}}{2^{2k}\delta_1} = 3 \cdot 2^{-k} N^{\frac{1}{2}} \delta_1^{-1}.$$

Since $M_{\delta_1}(k) = \bigcup_{b \in B(\delta_1)} M(b,k)$, this yields $|M_{\delta_1}(k)| \leq \delta_1 |M(b,k)| \leq 3 \cdot 2^{-k} N^{\frac{1}{2}}$.

Thus, we get

$$
\begin{aligned}
|\mathcal{E}_3| &= \sum_{\delta_1 \in \left[N^{\frac{1}{2}}\right]} \sum_{k=0}^{\bar{k}} |M_{\delta_1}(k)| \\
&\leq \sum_{\delta_1 \in \left[N^{\frac{1}{2}}\right]} \sum_{k=0}^{\bar{k}} 3 \cdot 2^{-k} N^{\frac{1}{2}} \\
&< 3N \sum_{k=0}^{\infty} 2^{-k} \\
&\leq 6N
\end{aligned}
$$

It is easy to see that $|\mathcal{E}_1 \cup \mathcal{E}_2| < N$. This proves the lemma. $\square$

3   Discrete Fourier Analysis

The purpose of this section is discrete Fourier analysis on the additive group $(\mathbb{Z}, +)$ and its connection to the discrepancy of the hypergraph $\mathcal{H}_{2AP}$. First

of all, let us extend the coloring χ to the set of all integers as follows. We keep the "old" color values for the set $[N]$ and set $\chi(z) := 0$ for all $z \in \mathbb{Z} \setminus [N]$. Thus, the (extended) coloring $\chi \colon \mathbb{Z} \to \{-1, 0, 1\}$ satisfies the condition: $\chi(z) = 0$, if and only if $z \in \mathbb{Z} \setminus [N]$. For every set $E \subseteq \mathbb{Z}$ we define its color value $\chi(E) := \sum_{x \in E} \chi(x)$. One can easily verify that we can express the coloring value of the set $E_a := a + E = \{a + x \mid x \in E\}$ as convolution of χ and the indicator function $\mathbb{1}_{-E}$ of the set $-E = \{-x \mid x \in E\}$ evaluated at a. For all $a \in \mathbb{Z}$, we have

$$\chi(E_a) = (\chi * \mathbb{1}_{-E})(a). \tag{1}$$

Thus, for all $E \subseteq \mathbb{Z}$

$$\sum_{a \in \mathbb{Z}} |\chi(E_a)|^2 = \|\chi * \mathbb{1}_{-E}\|_2^2.$$

For the proof of the lower bound in Theorem 1 we use a 2–norm approach. More precisely, we will estimate the sum of squared discrepancies

$$\sum_{E \in \mathcal{E}_0} \sum_{a \in \mathbb{Z}} |\chi(E_a)|^2 = \sum_{E \in \mathcal{E}_0} \|\chi * \mathbb{1}_{-E}\|_2^2. \tag{2}$$

Using two well-known facts from Fourier analysis, the Plancherel Theorem and the multiplicity of the Fourier transform, we will lower bound this sum of squared discrepancies. Afterwards an averaging argument will yield the existence of a hyperedge E with a discrepancy of order $\Omega(N^{\frac{1}{2}})$. But first of all we introduce the Fourier transform of a function $f \colon \mathbb{Z} \to \mathbb{C}$. The Fourier transform of f is defined as

$$\widehat{f} \colon [0, 1) \to \mathbb{C}, \quad \alpha \mapsto \sum_{z \in \mathbb{Z}} f(z) e^{2\pi i z \alpha}.$$

In the following lemma we list the two facts from Fourier analysis on the additive group $(\mathbb{Z}, +)$ that we will need for our calculations.

Lemma 3. *Let $f, g \colon \mathbb{Z} \to \mathbb{C}$ two square integrable functions. For the Fourier transform of f and g we get*

(i) $\|\widehat{f}\|_2^2 = \|f\|_2^2$ *(Plancherel Theorem)*,

(ii) $\widehat{f * g} = \widehat{f}\widehat{g}$.

4 Proof of the Lower Bound

Before we prove the lower bound for the discrepancy of the hypergraph of sums of two arithmetic progressions, we state the following lemma.

Lemma 4 (Main Lemma). *For every $\alpha \in [0,1)$, there exists an $E \in \mathcal{E}_0$ such that*

$$|\widehat{\mathbb{1}}_{-E}(\alpha)| \geq \frac{1}{300}N.$$

Applying this lemma, we are able to give the lower bound proof.

Proof of Theorem 1. Using the equation (2) and Lemma 3, we get

$$
\begin{aligned}
\sum_{E \in \mathcal{E}_0} \sum_{a \in \mathbb{Z}} |\chi(E_a)|^2
&= \sum_{E \in \mathcal{E}_0} \|\chi * \mathbb{1}_{-E}\|_2^2 \\
&= \sum_{E \in \mathcal{E}_0} \|\widehat{\chi * \mathbb{1}_{-E}}\|_2^2 \\
&= \sum_{E \in \mathcal{E}_0} \|\widehat{\chi}\widehat{\mathbb{1}}_{-E}\|_2^2 \\
&= \sum_{E \in \mathcal{E}_0} \int_0^1 |\widehat{\chi}(\alpha)|^2 |\widehat{\mathbb{1}}_{-E}(\alpha)|^2 d\alpha \\
&= \int_0^1 |\widehat{\chi}(\alpha)|^2 \left(\sum_{E \in \mathcal{E}_0} |\widehat{\mathbb{1}}_{-E}(\alpha)|^2 \right) d\alpha.
\end{aligned}
$$

The Main Lemma yields for every $\alpha \in [0,1)$ the existence of an $E \in \mathcal{E}_0$ such that $|\widehat{\mathbb{1}}_{-E}(\alpha)| \geq \frac{1}{300}N$. Thus, we get for every $\alpha \in [0,1)$

$$\sum_{E \in \mathcal{E}_0} |\widehat{\mathbb{1}}_{-E}(\alpha)|^2 \geq \tfrac{1}{90000}N^2.$$

Hence, we can continue the estimation of the sum of squarred discrepancies as follows.

$$
\begin{aligned}
\sum_{E \in \mathcal{E}_0} \sum_{a \in \mathbb{Z}} |\chi(E_a)|^2
&= \int_0^1 |\widehat{\chi}(\alpha)|^2 \left(\sum_{E \in \mathcal{E}_0} |\widehat{\mathbb{1}}_{-E}(\alpha)|^2 \right) d\alpha \\
&\geq \tfrac{1}{90000}N^2\|\widehat{\chi}\|_2^2 \\
&= \tfrac{1}{90000}N^2\|\chi\|_2^2 \\
&= \tfrac{1}{90000}N^3.
\end{aligned}
$$

Since every $E \in \mathcal{E}_0$ satisfies $E \subseteq [0, N-1]_\mathbb{Z}$, we get for every $a \in \mathbb{Z} \setminus [-N + 1, N]_\mathbb{Z}$ that $E \cap [N] = \emptyset$ and thus $\chi(E_a) = 0$. Therefore, $\sum_{E \in \mathcal{E}_0} \sum_{a \in \mathbb{Z}} |\chi(E_a)|^2$ is the sum of at most $2N|\mathcal{E}_0| \leq 14N^2$ non-trivial elements (Lemma 2). Hence, there exists an $E \in \mathcal{E}_0$ and an $a \in [-N + 1, N]_\mathbb{Z}$ such that

$$|\chi(E_a)|^2 \geq \tfrac{1}{1260000} N.$$

Thus, we have proven

$$\mathrm{disc}(\mathcal{H}_{2AP}) \geq |\chi(E_a)| > \tfrac{1}{1200} N^{\frac{1}{2}}.$$

$\square$

Before we can prove the Main Lemma, we have to state and prove the following four lemmas.

Lemma 5. *For every $\alpha \in [0, 1)$ and every $k \in \mathbb{N}$, there exists a $\delta \in [k]$ and an $a \in \mathbb{Z}$ such that*
$$|\delta\alpha - a| < \frac{1}{k}.$$

Proof. For all $j \in [k]$, we define

$$M_j := \left\{ \delta \in [k] \colon \delta\alpha - \lfloor \delta\alpha \rfloor \in \left[\frac{j-1}{k}, \frac{j}{k} \right) \right\}.$$

For every $\delta \in M_1$, holds $|\delta\alpha - \lfloor \delta\alpha \rfloor| < \frac{1}{k}$. Thus, we can assume $M_1 = \emptyset$. By the pigeon hole principle, there exists a $j \in [k] \setminus \{1\}$ with $|M_j| \geq 2$. Let $\delta_1, \delta_2 \in M_j$ with $\delta_1 < \delta_2$. Set $\delta := \delta_2 - \delta_1$. Using $\delta_1, \delta_2 \in M_j$, we get

$$|\delta - (\lfloor \delta_2\alpha \rfloor - \lfloor \delta_1\alpha \rfloor)| = |(\delta_2 - \lfloor \delta_2\alpha \rfloor) - (\delta_1 - \lfloor \delta_1\alpha \rfloor)| < \frac{1}{k}.$$

$\square$

Lemma 6. *Let $a, \delta \in \mathbb{N}$ with $(a, \delta) = 1$. There exists a $k \in [\delta - 1]$ such that*

$$ka \equiv 1 \ (mod \ \delta).$$

Moreover, $(\delta - k)a \equiv -1 \ (mod \ \delta)$. It holds $(k, \delta) = (\delta - k, \delta) = 1$.

Proof. Since $(a, \delta) = 1$, there exist $k, \ell \in \mathbb{Z}$ with

$$ka + \ell\delta = 1.$$

Thus, $ka \equiv 1 \pmod{\delta}$. Obviously, k can be choosen from the set $[\delta - 1]$. The second assertion follows from

$$ka + (\delta - k)a = \delta a \equiv 0 \pmod{\delta}.$$

Finally, the equation $ka + \ell\delta = 1$ proves also $(k, \delta) = 1$. But this implies $(\delta - k, \delta) = 1$. $\qquad\square$

Lemma 7. *Let $\alpha \in [0, 1)$, $\delta_1, \delta_2, L_1, L_2 \in \mathbb{N}$ with $L_1 \neq 1 \neq L_2$ be chosen such that for suitable $a_1, a_2 \in \mathbb{Z}$ we have*

$$|\delta_j\alpha - a_j| \leq \frac{1}{12(L_j - 1)}, \quad (j = 1, 2).$$

Set $E := \{j_1\delta_1 + j_2\delta_2 \colon j_1 \in [0, L_1 - 1], j_2 \in [0, L_2 - 1]\}$. For the Fourier transform of the indicator function $\mathbb{1}_{-E}$ of the set $-E$ we get

$$|\widehat{\mathbb{1}}_{-E}(\alpha)| \geq \frac{|E|}{2}.$$

Proof. The Fourier transform of a function $f \colon \mathbb{Z} \to \mathbb{C}$ is given as $\widehat{f} \colon [0, 1) \to \mathbb{C}$, $\alpha \mapsto \sum_{z \in \mathbb{Z}} f(z)e^{-2\pi i z\alpha}$. Thus,

$$\widehat{\mathbb{1}}_{-E}(\alpha) = \sum_{z \in E} e^{2\pi i z\alpha}.$$

Let $z \in E$. There exists a $j_1 \in [0, L_1 - 1]$ and a $j_2 \in [0, L_2 - 1]$ with $z = j_1\delta_1 + j_2\delta_2$. Hence,

$$
\begin{aligned}
e^{2\pi i z\alpha} &= e^{2\pi i(j_1\delta_1 + j_2\delta_2)\alpha} \\
&= e^{2\pi i[j_1(\delta_1\alpha - a_1) + j_2(\delta_2\alpha - a_2)]}e^{2\pi i(j_1 a_1 + j_2 a_2)} \\
&= e^{2\pi i[j_1(\delta_1\alpha - a_1) + j_2(\delta_2\alpha - a_2)]}.
\end{aligned}
$$

Using $|j_1(\delta_1\alpha - a_1) + j_2(\delta_2\alpha - a_2)| \leq \frac{L_1 - 1}{12(L_1 - 1)} + \frac{L_2 - 1}{12(L_2 - 1)} = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$, we get $\Re(e^{2\pi i z\alpha}) \geq \frac{1}{2}$. This proves

$$|\widehat{\mathbb{1}}_{-E}(\alpha)| = \sum_{z \in E} e^{2\pi i z\alpha} \geq \Re(\sum_{z \in E} e^{2\pi i z\alpha}) \geq \frac{|E|}{2}.$$

$\qquad\square$

Lemma 8. *Let $\delta_1, \delta_2, L_1, L_2 \in \mathbb{N}$. If $L_1 \leq \frac{\delta_2}{(\delta_1, \delta_2)}$ then*

$$|\{j_1\delta_1 + j_2\delta_2 \colon j_1 \in [0, L_1 - 1], j_2 \in [0, L_2 - 1]\}| = L_1L_2.$$

Proof. Assume there are $(j_1, j_2), (j_1', j_2') \in [0, L_1 - 1] \times [0, L_2 - 1]$ such that $(j_1, j_2) \neq (j_1', j_2')$ and

$$j_1\delta_1 + j_2\delta_2 = j_1'\delta_1 + j_2'\delta_2.$$

Clearly, $j_1 \neq j_1'$ and $j_2 \neq j_2'$. Since $(j_1 - j_1')\delta_1 = (j_2' - j_2)\delta_2$ is divisible by δ_1 and δ_2 and thus also by their least common multiple $\mathrm{lcm}(\delta_1, \delta_2) = \frac{\delta_1\delta_2}{(\delta_1, \delta_2)}$, we get

$$L_1 > |j_1 - j_1'| \geq \frac{\delta_2}{(\delta_1, \delta_2)}.$$

This contradiction shows that the function

$$f \colon [0, L_1 - 1] \times [0, L_1 - 1] \to \mathbb{Z}, \quad (j_1, j_2) \mapsto j_1\delta_1 + j_2\delta_2$$

is injective which proves the assumption. $\qquad\square$

By combining Lemma 5, Lemma 6, Lemma 7, and Lemma 8, we are able to prove the Main Lemma. Recall that we proved the lower bound for the discrepancy of the hypergraph of all sums of two arithmetic progressions just by applying the Main Lemma.

Proof of the Main Lemma. Using Lemma 5, we can find a $\delta_1 \in [N^{\frac{1}{2}}]$ such that for an appropriate $a_1 \in \mathbb{Z}$ it holds $|\delta_1\alpha - a_1| < N^{-\frac{1}{2}}$. Dividing by δ_1, we get

$$|\alpha - \frac{a_1}{\delta_1}| < N^{-\frac{1}{2}}\delta_1^{-1}. \tag{3}$$

We can choose δ_1 and a_1 in such a way that $\frac{a_1}{\delta_1}$ is an irreducible fraction. We distinguish three cases.

Case 1: $|\alpha - \frac{a_1}{\delta_1}| < N^{-1}$ and $\delta_1 \leq 24$.

Set $L_1 := \lceil \frac{N}{6\delta_1} \rceil$, $\delta_2 := 1$, and $L_2 := 1$. The set $E := E_{\delta_1, L_1, \delta_2, L_2}$ is an element of the special set of hyperedges $\mathcal{E}_0$. More precisely, $E \in \mathcal{E}_1$. Arguments

11

similar to those used in the proof of Lemma 7 show

$$
\begin{aligned}
|\widehat{\mathbb{1}}_{-E}(\alpha)| \;\geq\;& \Re\left(\sum_{z\in E} e^{2\pi i z\alpha}\right) \\
=\;& \sum_{j_1=0}^{L_1-1} \Re\left(e^{2\pi i j_1 \delta_1 \alpha}\right) \\
\geq\;& L_1 \Re\left(e^{\frac{2\pi i}{6}}\right) \\
\geq\;& \frac{N}{288}.
\end{aligned}
$$

Case 2: $|\alpha - \frac{a_1}{\delta_1}| < N^{-1}$ and $\delta_1 > 24$.

Set $L_1 := \lceil \frac{N}{12\delta_1} \rceil$. Using again Lemma 5, there is a $\delta_2 \in [\delta_1 - 1]$ such that for a suitable $a_2 \in \mathbb{Z}$ it holds $|\delta_2 \alpha - a_2| \leq \frac{1}{\delta_1-1}$. Hence,

$$
|\alpha - \frac{a_2}{\delta_2}| \leq \frac{1}{(\delta_1 - 1)\delta_2}. \tag{4}
$$

Set $L_2 := \lceil \frac{\delta_1-1}{12} \rceil$. Since $\frac{a_1}{\delta_1}$ is an irreducible fraction and $\delta_2 < \delta_1$ we get $\frac{a_1}{\delta_1} \neq \frac{a_2}{\delta_2}$. Thus,

$$
|\frac{a_1}{\delta_1} - \frac{a_2}{\delta_2}| \geq \frac{1}{\mathrm{lcm}(\delta_1, \delta_2)}. \tag{5}
$$

On the other hand, using (3) and (4) we get

$$
\begin{aligned}
|\frac{a_1}{\delta_1} - \frac{a_2}{\delta_2}| \leq |\frac{a_1}{\delta_1} - \alpha| + |\alpha - \frac{a_2}{\delta_2}| \;\leq\;& \frac{1}{N\delta_1} + \frac{1}{(\delta_1 - 1)\delta_2} \\
\leq\;& \left(\frac{1}{\delta_1} + \frac{\delta_1}{\delta_1 - 1}\right)\frac{1}{\delta_1 \delta_2} \\
<\;& \frac{13}{12}\frac{1}{\delta_1 \delta_2}
\end{aligned} \tag{6}
$$

Combining (5) and (6) gives

$$
\frac{1}{\mathrm{lcm}(\delta_1, \delta_2)} < \frac{13}{12}\frac{1}{\delta_1 \delta_2}.
$$

12

But this implies $(\delta_1, \delta_2) = \frac{\delta_1 \delta_2}{\text{lcm}(\delta_1, \delta_2)} < \frac{13}{12}$ and thus $(\delta_1, \delta_2) = 1$. Define

$$E := E_{\delta_1, L_1, \delta_2, L_2} = \{j_1 \delta_1 + j_2 \delta_2 : j_1 \in [0, L_1 - 1], j_2 \in [0, L_2 - 1]\}.$$

We have $E \in \mathcal{E}_2 \subseteq \mathcal{E}_0$. Since $L_2 = \lceil \frac{\delta_1 - 1}{12} \rceil < \frac{\delta_1}{6} < \frac{\delta_1}{(\delta_1, \delta_2)}$, we can apply Lemma 8 and get $|E| = L_1 L_2 \geq \frac{N}{12\delta_1} \frac{\delta_1 - 1}{12} \geq \frac{1}{150} N$. Furthermore, $\delta_1, \delta_2, L_1, L_2$ satisfy the conditions of Lemma 7. Thus,

$$|\widehat{\mathbb{1}}_{-E}(\alpha)| \geq \frac{|E|}{2} \geq \frac{1}{300} N.$$

Case 3: $|\alpha - \frac{a_1}{\delta_1}| \geq N^{-1}$.

Choose k such that

$$\left| \alpha - \frac{a_1}{\delta_1} \right| \in [2^{-k-1} N^{-\frac{1}{2}} \delta_1^{-1}, 2^{-k} N^{-\frac{1}{2}} \delta_1^{-1}). \tag{7}$$

Since $|\alpha - \frac{a_1}{\delta_1}|$ is lower bounded by N^{-1} and from above by $N^{-\frac{1}{2}} \delta_1^{-1}$ (by Inequality (3)), it holds $0 \leq k \leq \log(N^{\frac{1}{2}} \delta_1^{-1})$. Set $L_1 := \lceil \frac{2^k N^{\frac{1}{2}}}{12} \rceil$. Using $(a_1, \delta_1) = 1$, we can apply Lemma 6, which yields the existence of a $\gamma \in [\delta_1 - 1]$ such that

(i) $\gamma a_1 \equiv 1 \pmod{\delta_1}$,

(ii) $(\delta_1 - \gamma) a_1 \equiv -1 \pmod{\delta_1}$.

Let $s := (\alpha - \frac{a_1}{\delta_1})|\alpha - \frac{a_1}{\delta_1}|^{-1}$, i.e., s is the algebraic sign of $(\alpha - \frac{a_1}{\delta_1})$. If $s = 1$ we set $b := \delta_1 - \gamma$, otherwise we set $b := \gamma$. In both cases there exists a $\mu \in \mathbb{Z}$ such that

$$b \frac{a_1}{\delta_1} = \mu - \frac{s}{\delta_1}.$$

Define $d := |\alpha - \frac{a_1}{\delta_1}|^{-1} 2^{-2k} \delta_1^{-2} - b 2^{-2k} \delta_1^{-1}$ and $\delta_2 := b + \lceil d \rceil 2^{2k} \delta_1$. Then

$$
\begin{aligned}
\delta_2 \alpha &= (b + \lceil d \rceil 2^{2k} \delta_1) \frac{a_1}{\delta_1} + (b + \lceil d \rceil 2^{2k} \delta_1) \left(\alpha - \frac{a_1}{\delta_1} \right) \\
&= (b + \lceil d \rceil 2^{2k} \delta_1) \frac{a_1}{\delta_1} + (b + d 2^{2k} \delta_1) \left(\alpha - \frac{a_1}{\delta_1} \right) + (\lceil d \rceil - d)(\delta_1 \alpha - a_1) \\
&= \mu - \frac{s}{\delta_1} + \lceil d \rceil 2^{2k} a_1 + \frac{s}{\delta_1} + (\lceil d \rceil - d) 2^{2k} s |\delta_1 \alpha - a_1| \\
&= \mu + \lceil d \rceil 2^{2k} a_1 + (\lceil d \rceil - d) 2^{2k} s |\delta_1 \alpha - a_1|
\end{aligned}
$$

13

Using (7), we get

$$|\delta_2 \alpha - (\mu + \lceil d \rceil 2^{2k} a_1)| \in [0, 2^k N^{-\frac{1}{2}}).$$

Since $\lceil d \rceil < d+1 = |\alpha - \frac{a_1}{\delta_1}|^{-1} 2^{-2k} \delta_1^{-2} - b 2^{-2k} \delta_1^{-1} + 1$, (7) yields the estimation

$$\delta_2 = b + \lceil d \rceil 2^{2k} \delta_1 < |\alpha - \frac{a_1}{\delta_1}|^{-1} \delta_1^{-1} + 2^{2k} \delta_1 \le 2^{k+1} N^{\frac{1}{2}} + 2^{2k} \delta_1.$$

On the other hand $\delta_2 \ge b + d 2^{2k} \delta_1 = |\alpha - \frac{a_1}{\delta_1}|^{-1} \delta_1^{-1} > 2^k N^{\frac{1}{2}}$. Thus, $\delta_2 \in M(b,k) \subseteq M_{\delta_1}(k)$. Set $L_2 := \lceil 2^{-k} \frac{N^{\frac{1}{2}}}{12} \rceil$. Then the set $E := E_{\delta_1, L_1, \delta_2, L_2}$ is an element of $\mathcal{E}_3$ and thus $E \in \mathcal{E}_0$.

Before we can apply Lemma 8, we have to verify its conditions for the quadruple $(\delta_1, L_1, \delta_2, L_2)$. Since $(b, \delta_1) = 1$, also $(\delta_1, \delta_2) = 1$. Moreover,

$$
\begin{aligned}
\delta_2 \ge b + d \delta_1 2^{2k} &= |\alpha - \frac{a_1}{\delta_1}|^{-1} \delta_1^{-1} \\
&> (2^{-k} N^{-\frac{1}{2}} \delta_1^{-1})^{-1} \delta_1^{-1} \\
&= 2^k N^{\frac{1}{2}} \\
&> \lceil \frac{2^k N^{\frac{1}{2}}}{12} \rceil = L_1.
\end{aligned}
$$

Thus, the conditions of Lemma 8 are satisfied and the cardinality of the set $E := \{ j_1 \delta_1 + j_2 \delta_2 \mid j_1 \in [0, L_1 - 1], j_2 \in [0, L_2 - 1] \}$ can be estimated as follows: $|E| = L_1 L_2 \ge \frac{2^k N^{\frac{1}{2}}}{12} \frac{2^{-k} N^{\frac{1}{2}}}{12} = \frac{N}{144}$. Therefore, Lemma 7 proves

$$|\widehat{\mathbb{1}}_{-E}(\alpha)| \ge \frac{|E|}{2} \ge \frac{N}{288}.$$

$\square$

References

[AS00] N. Alon, J. Spencer, and P. Erdős. *The Probabilistic Method* (Second Edition). John Wiley & Sons, Inc., New York, 2000.

[B81] J. Beck. Roth's Estimate of the Discrepancy of Integer Sequences is Nearly Sharp. *Combinatorica* **1**(4) (1981), 319-325.

[BS95] J. Beck and V. T. Sós. *Discrepancy theory*. In R. Graham, M. Gröschel and L. Lovász, editors, Handbook of Combinatorics, 1405-1446. 1995.

[DSW] B. Doerr, A. Srivastav, P. Wehr. Discrepancies of Cartesian Products of Arithmetic Progressions. *Electronic Journal of Combinatorics*, to appear.

[ES74] P. Erdős, J. Spencer. *Probabilistic Methods in Combinatorics*. Akademiai Kiado, Budapest, 1974.

[H2004] N. Hebbinghaus. Discrepancy of Sums of Arithmetic Progressions. *Electronic Notes in Discrete Mathematics* **17C** (2004), pages 185-189.

[Mat99] J. Matoušek. *Geometric Discrepancy*. Springer, Berlin, 1999.

[MS96] J. Matoušek, J. Spencer. Discrepancy in Arithmetic Progressions. *Journal of the American Mathematical Society* **9**(1) (1996), 195-204.

[P2006] A. Přívětivý. Discrepancy of Sums of Three Arithmetic Progressions. *The Electronic Journal of Combinatorics* **13**(1) (2006), R5.

[R64] K. F. Roth. Remark Concerning Integer Sequences. *Acta Arithmetica* **9** (1964), 257-260.

[V2002] B. Valkó. Discrepancy of Arithmetic Progressions in Higher Dimensions. *Journal of Number Theory* **92** (2002), 117-130.