A NEW UPPER BOUND FOR THE SIZE OF 
s-DISTANCE SETS IN BOXES

G. HEGEDŰS
Óbuda University, Bécsi út 96, H-1037 Budapest, Hungary
e-mail: hegedus.gabor@nik.uni-obuda.hu

(Received February 22, 2019; revised July 25, 2019; accepted July 26, 2019)

Abstract. Let \( G \subseteq \mathbb{R}^n \) be an arbitrary subset. We denote by \( d(G) \) the set of (non-zero) Euclidean distances among points of \( G \):

\[ d(G) := \{ d(p_1, p_2) : p_1, p_2 \in G, \ p_1 \neq p_2 \}. \]

An \( s \)-distance set is any subset \( H \subseteq \mathbb{R}^n \) such that \( |d(H)| \leq s \).

Our main result is a new upper bound for the size of \( s \)-distance sets in the direct product of finite sets of points in the Euclidean space. We use Tao’s slice rank method in our proof.

1. Introduction

In this manuscript we give upper bounds for \( s \)-distance sets in the direct product of finite sets of points in the Euclidean space. Here an \( s \)-distance set is by definition any subset \( G \) of the Euclidean space \( \mathbb{R}^n \) such that the size of the set of distances among points of \( G \) is at most \( s \). Two independent research directions motivated our results.

Our first motivation comes from the following question of Erdős: Given \( n \) points in the plane, what is the smallest number of distinct distances they can determine?

Erdős proved the following result.

**Theorem 1.1** (Erdős [7]). The minimum number of \( f(n) \) of distances determined by \( n \) points of the plane satisfies the inequalities

\[ \left( n - \frac{3}{4} \right)^{1/2} - \frac{1}{2} \leq f(n) \leq \frac{cn}{(\log n)^{1/2}}. \]

Erdős conjectured that the square grid is essentially the extremal example, consequently the upper bound for the function \( f(n) \) is sharp.

Key words and phrases: \( s \)-distance set, distance problem, slice rank method.

Mathematics Subject Classification: 52C45, 12Y99, 05D99.
Our second motivation comes from the investigation of different types of $s$-distance sets in the $d$-dimensional Euclidean space.

First we introduce some notation. Let $\mathbb{R}^d$ denote the $d$-dimensional Euclidean space. Let $\mathcal{G} \subseteq \mathbb{R}^n$ be an arbitrary set. Here we can consider naturally the points of $\mathcal{G}$ as vectors. We denote by $d(\mathcal{G})$ the set of (non-zero) distances among points of $\mathcal{G}$:

$$d(\mathcal{G}) := \{ d(p_1, p_2) : p_1, p_2 \in \mathcal{G}, \ p_1 \neq p_2 \}.$$ 

Let $(x, y)$ stand for the standard scalar product. Let $\mathcal{G} \subseteq \mathbb{R}^n$ be an arbitrary set. We denote by $s(\mathcal{G})$ the set of scalar products between distinct points of $\mathcal{G}$:

$$s(\mathcal{G}) := \{ (p_1, p_2) : p_1, p_2 \in \mathcal{G}, \ p_1 \neq p_2 \}.$$ 

Bannai, Bannai and Stanton proved the following result.

**Theorem 1.2** (Bannai, Bannai and Stanton [1, Theorem 1]). Suppose that $\mathcal{F} \subseteq \mathbb{R}^n$ is a set satisfying $|d(\mathcal{F})| \leq s$. Then

$$|\mathcal{F}| \leq \binom{n + s}{s}.$$ 

Delsarte, Goethals and Seidel investigated the spherical $s$-distance sets. They proved the following theorem.

**Theorem 1.3** (Delsarte, Goethals and Seidel [6]). Suppose that $\mathcal{F} \subseteq \mathbb{S}^{n-1}$ is a set satisfying $|d(\mathcal{F})| \leq s$. Then

$$|\mathcal{F}| \leq \binom{n + s - 1}{s} + \binom{n + s - 2}{s - 1}.$$ 

We state here our main results.

**Theorem 1.4.** Let $A_i \subseteq \mathbb{R}, \ |A_i| = q \geq 2$ for each $1 \leq i \leq n$. Consider the box $B := \prod_{i=1}^n A_i \subseteq \mathbb{R}^n$. Suppose that $\mathcal{F} \subseteq B$ is a set such that $|d(\mathcal{F})| \leq s$. Then

$$|\mathcal{F}| \leq 2 \cdot \left| \left\{ x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n} : 0 \leq \alpha_i \leq q - 1 \text{ for each } i, \ \sum_i \alpha_i \leq s \right\} \right|.$$ 

**Remark.** If $q = 2$, then

$$\left| \left\{ x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n} : 0 \leq \alpha_i \leq q - 1 \text{ for each } i, \ \sum_i \alpha_i \leq s \right\} \right| = \sum_{j=0}^s \binom{n}{j},$$
hence we get the upper bound

\[ |\mathcal{F}| \leq 2 \sum_{j=0}^{s} \binom{n}{j}, \]

which is clearly a better bound than the Bannai, Bannai and Stanton’s bound \( \binom{n+s}{s} \).

We use Tao’s slice rank method in our proof (see the blog post [10]). This method is a symmetric reformulation of the original Croot–Lev–Pach polynomial method (see [4]). This proof technique simplified the previous proof of Ellenberg and Gijswijt’s breakthrough about the upper bounds for the size of subsets \( A \) in \( (\mathbb{Z}_p)^n \) without three-term arithmetic progressions (see [8]).

Let \( t, d \geq 2 \) be integers. Define

\[ J(t, d) := \frac{1}{t} \left( \min_{0 \leq x < 1} \frac{1 - x^t}{1 - x} x^{-\frac{t-1}{d}} \right), \]

and for brevity let \( J(q) := J(q,3) \) for each \( q > 1 \).

The constant \( J(q) \) appeared in the proof of Ellenberg and Gijswijt’s bound for the size of three-term progression-free sets (see [8]). It was proved in [2, Proposition 4.12], that \( J(q) \) is a decreasing function of \( q \) and

\[ \lim_{q \to \infty} J(q) = \inf_{z > 3} \frac{z - z^{-2}}{3 \log(z)} = 0.8414 \ldots. \]

We can verify easily that \( J(3) \approx 0.9184 \), consequently \( J(q) \) lies in the range

\[ 0.8414 \leq J(q) \leq 0.9184 \]

for each \( q \geq 3 \). The following corollary gives a more appropriate upper bound for the size of \( s \)-distance sets in boxes.

**Corollary 1.5.** Let \( A_i \subseteq \mathbb{R}, |A_i| = q \geq 2 \) for each \( 1 \leq i \leq n \). Consider the box \( \mathcal{B} := \prod_{i=1}^{n} A_i \subseteq \mathbb{R}^n \). Suppose that \( \mathcal{F} \subseteq \mathcal{B} \) is a set such that \( |d(\mathcal{F})| \leq s \). Let \( d := \frac{n(q-1)}{s} \). Then

\[ |\mathcal{F}| \leq 2(q J(q,d))^n. \]

Deza and Frankl proved the following statement.

**Theorem 1.6** (Deza and Frankl [5, Theorem 1.4]). Suppose that \( \mathcal{F} \subseteq \mathbb{R}^n \) is an arbitrary set such that \( |s(\mathcal{F})| \leq s \). Then

\[ |\mathcal{F}| \leq \binom{n+s}{s}. \]
It is easy to prove the following result using a slight modification of Tao’s slice rank method. We give here only a sketch of the proof.

**Theorem 1.7.** Let $A_i \subseteq \mathbb{R}$, $|A_i| = q \geq 2$ for each $1 \leq i \leq n$. Consider the box $B := \prod_{i=1}^{n} A_i \subseteq \mathbb{R}^n$. Suppose that $F \subseteq B$ is an arbitrary set which satisfies the following properties:

(i) $(f, f) \notin s(F)$ for each $f \in F$;
(ii) $|s(F)| \leq s$.

Then

$$|F| \leq \left| \left\{ x_1^{\alpha_1} \cdots x_n^{\alpha_n} : 0 \leq \alpha_i \leq q - 1 \text{ for each } i, \sum_i \alpha_i \leq s \right\} \right|.$$  

**2. Proofs**

The proofs of our main results are built on the idea of the proof of the following proposition.

**Proposition 2.1** [4, Lemma 1]. Suppose that $n \geq 1$ and $d \geq 0$ are integers, $P$ is a multilinear polynomial in $n$ variables of total degree at most $d$ over a field $\mathbb{F}$, and $F \subseteq \mathbb{F}^n$ is a subset with

$$|F| > 2 \sum_{i=0}^{[d/2]} \binom{n}{i}.$$  

If $P(a - b) = 0$ for all $a, b \in F$, $a \neq b$, then $P(0) = 0$.

Tao’s slice rank method (see the blog [10] and an other proof in [2, Section 4]) gives easily the following slight generalization of Proposition 2.1 to boxes, which is a special case of [9, Corollary 1.3].

**Theorem 2.2.** Let $\mathbb{F}$ be an arbitrary field. Let $A_i \subseteq \mathbb{F}$ be fixed subsets such that $|A_1| = \ldots = |A_n| = t > 0$. Let $F \subseteq \prod_{i=1}^{n} A_i$ be an arbitrary subset. Suppose that there exists a polynomial

$$P(x_1, \ldots, x_n, y_1, \ldots, y_n) \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots y_n]$$

satisfying the following conditions:

(i) $P(a, a) \neq 0$ for each $a \in F$;
(ii) if $a, b \in F$, $a \neq b$ are arbitrary vectors, then $P(a, b) = 0$.

Then

$$|F| \leq 2 \left| \left\{ x_1^{\alpha_1} \cdots x_n^{\alpha_n} : 0 \leq \alpha_i \leq t - 1 \text{ for each } 1 \leq i \leq n \right. \right.$$  

$$\left. \text{ and } \sum_{j=1}^{n} \alpha_j \leq \frac{\deg(P)}{2} \right|.$$
As an easy consequence, we proved the following result.

**Corollary 2.3** [9, Corollary 1.5]. Let $\mathbb{F}$ be an arbitrary field. Let $A_i \subseteq \mathbb{F}$ be fixed subsets such that $|A_1| = \ldots = |A_n| = t > 0$. Let $\mathcal{F} \subseteq \prod_{i=1}^n A_i$ be an arbitrary subset. Suppose that there exists a polynomial $P(x_1, \ldots, x_n, y_1, \ldots, y_n) \in \mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ satisfying the following conditions:

(i) $P(a, a) \neq 0$ for each $a \in \mathcal{F}$;

(ii) if $a, b \in \mathcal{F}$, $a \neq b$ are arbitrary vectors, then $P(a, b) = 0$.

Let $d := \frac{2n(t-1)}{\deg(P)}$. Then

$$|\mathcal{F}| \leq 2(tJ(t, d))^n.$$ 

**Proof.** We use the following inequality.

**Theorem 2.4.** Let $n \geq 1$, $t, d \geq 2$ be integers and $B := \{0, 1, \ldots, t - 1\}$. Consider the set

$$B(n, d, t) := \{v = (v_1, \ldots, v_n) \in B^n : \sum_i v_i \leq \frac{n(t-1)}{d}\}.$$ 

Then $|B(n, d, t)| \leq (tJ(t, d))^n$.

The proof of Theorem 2.4 uses Markov’s inequality. We omit here this proof.

Corollary 2.3 follows easily from Theorem 2.2 and Theorem 2.4. Namely

\begin{equation}
|\mathcal{F}| \leq 2 \left| \left\{x_1^{\alpha_1} \ldots x_n^{\alpha_n} : 0 \leq \alpha_i \leq t - 1 \text{ for each } 1 \leq i \leq n \text{ and } \sum_{j=1}^n \alpha_j \leq \frac{\deg(P)}{2} \right\} \right|
\end{equation}

by Corollary 2.2. But

$$\frac{\deg(P)}{2} = \frac{n(t-1)}{d},$$

hence

\begin{equation}
|\mathcal{F}| \leq 2 \left| \left\{x_1^{\alpha_1} \ldots x_n^{\alpha_n} : 0 \leq \alpha_i \leq t - 1 \text{ for each } 1 \leq i \leq n \text{ and } \sum_{j=1}^n \alpha_j \leq \frac{n(t-1)}{d} \right\} \right| = 2|B(n, d, t)|.
\end{equation}
Consequently 
\[ |\mathcal{F}| \leq 2|B(n, d, t)| \leq 2(tJ(t, d))^n \]
by Theorem 2.4. □

**Proof of Theorem 1.4.** Consider the set \( d(\mathcal{F}) = \{d_1, \ldots, d_s\} \). Clearly \( d_i \neq 0 \) for each \( i \).

Define the polynomial

\[ P(x_1, \ldots, x_n, y_1, \ldots, y_n) := \prod_{i=1}^{s} \left( \sum_{j=1}^{n} (x_j - y_j)^2 - d_i^2 \right) \]
\[ \in \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_n]. \]

Clearly \( \deg(P) = 2s \). Then

\[ P(a, a) = \prod_{i=1}^{s} (-d_i^2) \neq 0 \]
for each \( a \in \mathcal{F} \). On the other hand if \( a, b \in \mathcal{F}, a \neq b \), then

\[ P(a, b) = \prod_{i=1}^{s} \left( \sum_{j=1}^{n} (a_j - b_j)^2 - d_i^2 \right) = \prod_{i=1}^{s} (d(a, b)^2 - d_i^2). \]

But then there exists an \( i \) such that \( d(a, b) = d_i \), because \( d(\mathcal{F}) = \{d_1, \ldots, d_s\} \). Hence \( d(a, b)^2 = d_i^2 \), so \( P(a, b) = 0 \).

Finally we can apply Theorem 2.2 with the choices \( F = \mathbb{R} \) and \( t = q \). □

**Proof of Corollary 1.5.** If we use the same polynomial (4), then we get an analogous proof of Corollary 1.5 from Corollary 2.3 as the previous proof of Theorem 1.4. □

**Proof of Theorem 1.7.** Let \( s(\mathcal{F}) := \{w_1, \ldots, w_s\} \). Here the same proof works as in Theorem 1.4, but we need to use the polynomial

\[ P(x_1, \ldots, x_n, y_1, \ldots, y_n) := \prod_{i=1}^{s} \left( \sum_{j=1}^{n} x_jy_j - w_i \right) \]
\[ \in \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_n]. \] □

3. Concluding remarks

We conjecture that the following bound is a sharp upper bound.
Conjecture. Let $A_i \subseteq \mathbb{R}$, $|A_i| = q \geq 2$ for each $1 \leq i \leq n$. Consider the box $B := \prod_{i=1}^{n} A_i \subseteq \mathbb{R}^n$. Suppose that $\mathcal{F} \subseteq B$ is a set with $|d(\mathcal{F})| \leq s$. Then

$$|\mathcal{F}| \leq \left| \left\{ x_1^{\alpha_1} \cdots x_n^{\alpha_n} : 0 \leq \alpha_i \leq q - 1 \text{ for each } i, \sum_i \alpha_i \leq s \right\} \right|. $$

As a lower bound, we give the following construction in the case $q = 2$. If $A \subseteq [n]$, then denote by $\mathbf{v}_A$ the characteristic vector of $A$. Consider the set

$$\mathcal{F} := \{ \mathbf{v}_A : A \subseteq [n], |A| = s \} \subseteq \{0, 1\}^n \subseteq \mathbb{R}^n. $$

Then $\mathcal{F}$ is an $s$-distance set with $|\mathcal{F}| = \binom{n}{s}$. 

References

[1] E. Bannai, E. Bannai and D. Stanton, An upper bound for the cardinality of an $s$-distance subset in real Euclidean space. II, Combinatorica, 3 (1983), 147–152.

[2] J. Blasiak, T. Church, H. Cohn, J. A. Grochow and C. Umans, On cap sets and the group-theoretic approach to matrix multiplication, Discrete Anal., 2017 (2017), 1–27.

[3] A. Blokhuis, A new upper bound for the cardinality of 2-distance sets in Euclidean space, Ann. Discrete Math., 20 (1984), 65–66.

[4] E. Croot, V. Lev and P. Pach, Progression-free sets in $\mathbb{Z}_4^n$. Ann. of Math., 185 (2017), 331–337.

[5] M. Deza and P. Frankl, Bounds on the maximum number of vectors with given scalar products, Proc. Amer. Math. Soc., 95 (1985), 323-329.

[6] P. Delsarte, J. M. Goethals and J. J. Seidel, Spherical codes and designs, Geom. Dedicata, 6 (1977), 363–388.

[7] P. Erdős, On sets of distances of $n$ points, Amer. Math. Monthly, 53 (1946), 248–250.

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

[9] G. Hegedüs, A generalization of Croot–Lev–Pach’s Lemma and a new upper bound for the size of difference sets in polynomial rings, preprint, arXiv:1803.05308 (2018).

[10] T. Tao, A symmetric formulation of the Croot–Lev–Pach–Ellenberg–Gijswijt capset bound (2016), https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/