More distinct distances under local conditions

Jacob Fox* János Pach† Andrew Suk‡

Abstract

We establish the following result related to Erdős's problem on distinct distances. Let V be an n-element planar point set such that any p members of V determine at least $\binom{p}{2} - p + 6$ distinct distances. Then V determines at least $n^{\frac{8}{7}-o(1)}$ distinct distances, as n tends to infinity.

1 Introduction

In his classic 1946 paper [4], Erdős asked to determine or estimate the minimum number of distinct distances determined by an n-element planar point set V. He showed that a $\sqrt{n} \times \sqrt{n}$ integer lattice determines $\Theta(n/\sqrt{\log n})$ distinct distances, and conjectured that any n-element point set determines at least $n^{1-o(1)}$ distinct distances. Several authors established lower bounds for this problem, and Guth and Katz [10] answered Erdős's question by proving that any n-element planar point set determines at least $\Omega(n/\log n)$ distinct distances.

In [8], Erdős and Gyárfás studied the following generalization. For integers p and q with $q \leq \binom{p}{2}$, let $D(n,p,q)$ denote the minimum number of distinct distances determined by a planar n-element point set V with the property that any p points from V determine at least q distinct distances. Trivially, we have $D\left(n,p,\binom{p}{2}\right) = \Theta(n^2)$, and it follows from the Guth-Katz result that $D(n,p,q) = \Omega(n/\log n)$ for every p and q.

By considering the $\sqrt{n}\times\sqrt{n}$ integer lattice, we get $D(n,3,2) = O(n/\sqrt{\log n})$, and the Guth-Katz result gives $D(n,3,2) = \Omega(n/\log n)$.

For the value $D(n,3,3)$, it is easy to see that $D(n,3,3) \geq n-1$. In this setting, no three points form an isosceles triangle. Thus, all distances between an arbitrarily fixed point and the remaining $n-1$ points are distinct. It is not known whether $D(n,3,3) = O(n)$. This problem is closely related to another classical question: What is the largest number of elements one can select from $\{1,2,\ldots,n\}$ without choosing 3 numbers that form an *arithmetic progression*? Suppose we can select δn numbers satisfying this condition, for some $\delta > 0$. Regarding them as points in the plane, they induce no isosceles triangle, and altogether, the number of distinct distance determined by them is at most $n-1$. Thus, we would obtain that $D(\delta n,3,3) < n$, that is, $D(n,3,3) \leq (1/\delta)n = O(n)$. However, Roth [12] and, more generally, Szemerédi [16] showed that no such δ exists. The best known upper bound, $D(n,3,3) = ne^{O(\sqrt{\log n})}$, follows from a 1-dimensional construction of Behrend [1] and a proper 2-dimensional one of Erdős, Füredi, Pach,

*Stanford University, Stanford, CA. Supported by a Packard Fellowship, by NSF CAREER award DMS-1352121, and by an Alfred P. Sloan Fellowship. Email: jacobfox@stanford.edu.

†EPFL, Lausanne and Rényi Institute, Budapest. Research partially supported by Swiss National Science Foundation grants 200020-165977 and 200021-162884. Email: pach@cims.nyu.edu.

‡University of Illinois at Chicago, Chicago, IL. Supported by NSF grant DMS-1500153. Email: suk@uic.edu.

and Ruzsa [7]. Erdős conjectured that

$$\lim_{n\to\infty} \frac{D(n,3,3)}{n} = \infty,$$

and this is still open.

For larger values of p, the problem becomes increasingly complicated. Clearly, $D(n,4,3) = O(n/\sqrt{\log n})$, see, e.g., Sheffer [14]. Dumitrescu [3] observed that $D(n,4,4) = ne^{O(\sqrt{\log n})}$. Erdős [5] also conjectured that $D(n,4,5)$ grows quadratically in n, but the best known lower and upper bounds are only $\Omega(n)$ and $O(n^2)$.

For any $p \geq 4$, we have

$$D\left(n, p, \binom{p}{2} - \lfloor p/2 \rfloor + 2\right) \geq \Omega(n^2).$$

To see this, it is enough to notice that in this setting no distance can occur $\lfloor \frac{p}{2} \rfloor$ times, because otherwise any p-element set of points containing all endpoints of the corresponding segments would determine only at most $\binom{p}{2} - (\lceil \frac{p}{2} \rceil - 1)$ distinct distances. Erdős and Gyárfás [8] proved that even if we reduce by one the required number of distinct distances among any p points to $q = \binom{p}{2} - \lfloor \frac{p}{2} \rfloor + 1$, the number of distinct distances in the whole n-element set must be superlinear in n. Specifically, we have

$$D\left(n, p, \binom{p}{2} - \lfloor p/2 \rfloor + 1\right) = \Omega\left(n^{4/3}\right).$$

Furthermore, a result of Sárkőzy and Selkow [13] implies that for every $p \geq 6$ there exists $\epsilon = \epsilon(p) > 0$ with

$$D\left(n, p, \binom{p}{2} - p + \lceil \log p \rceil + 4\right) = \Omega\left(n^{1+\epsilon}\right).$$

The last two results were established in the following Ramsey-theoretic framework. We color all point pairs that determine the same distance with the same color. Then any p-element set contains pairs of at least q distinct colors. Using the last assumption alone, one can prove that the total number of colors cannot be too small.

The aim of the present note is to improve the Sárközy-Selkow bound by exploring the special properties of the above coloring that can be deduced from the geometric constraints. Our main result is the following.

Theorem 1. *Let $p \geq 6$ be an integer. Then the minimum number of distinct distances determined by n points in the plane with the property that any p of them induce at least $\binom{p}{2} - p + 6$ distinct distances, satisfies*

$$D\left(n, p, \binom{p}{2} - p + 6\right) \geq n^{\frac{8}{7} - o(1)},$$

as n tends to infinity.

Let us remark that for $p < 9$, one can obtain a better bound of $\Omega(n^2)$ by the simple argument stated above.

For a fixed p, define $q_1(p)$ to be the *largest* integer q for which $D(n,p,q) = O(n)$. Likewise, let $q_2(p)$ denote the *smallest* integer q for which $D(n,p,q) = \Omega(n^2)$. By the Guth-Katz result, we have $q_1(p) \geq \Omega(p/\log p)$. As we have seen above, $q_2(p) \leq \binom{p}{2} - \lfloor \frac{p}{2} \rfloor + 2$, and it was observed by Sheffer [14] that $q_2(p) \geq 2\lfloor \frac{p}{2} \rfloor$.

2 Graph-theoretic tools

Before we prove Theorem 1, we list several results that we will use. Let V be an ordered point set in $\mathbb{R}^d$, and let $E \subset \binom{V}{2}$. We say that E is a *semi-algebraic* relation on V with *complexity* at most t if there are at most t polynomials $g_1, \ldots, g_s \in \mathbb{R}[x_1, \ldots, x_{2d}]$, $s \leq t$, of degree at most t and a Boolean formula Φ such that for vertices $u, v \in V$ such that u comes before v in the ordering,

$$(u, v) \in E \quad \Leftrightarrow \quad \Phi(g_1(u, v) \geq 0; \ldots; g_s(u, v) \geq 0) = 1.$$

At the evaluation of $g_\ell(u, v)$, we substitute the variables $x_1, \ldots, x_d$ with the coordinates of u, and the variables $x_{d+1}, \ldots, x_{2d}$ with the coordinates of v. Here we only consider *symmetric* relations E, that is, $(u, v) \in E$ if and only if $(v, u) \in E$.

A classical result due to Kővári, Sós, and Turán, and independently Erdős, in extremal graph theory states the following.

Theorem 2.1 (Kővári-Sós-Turán [11], Erdős)**.** *Let $G = (U, V, E)$ be a bipartite graph. If G does not contain the subgraph $K_{2,r}$ with 2 vertices in U and r vertices in V, then*

$$|E(G)| = O(|U|\sqrt{|V|} + |V|),$$

where the hidden constant depends on r.

In particular, noting that every graph has a bipartite subgraph with at least half of its edges, we have that for any fixed r, all $K_{2,r}$-free graphs on $|V|$ vertices have $O(|V|^{3/2})$ edges. The next result improves this upper bound under the additional condition that the edge set E of the graph is a semi-algebraic relation with bounded description complexity.

Theorem 2.2 (Theorem 1.2 in [9])**.** *For fixed $d \geq 4$, $r \geq 2$ and $t \geq 1$, let $U, V \subset \mathbb{R}^d$ be finite point sets such that $|U| \leq |V|$, and let $E \subset U \times V$ be a semi-algebraic relation with complexity at most t. If the bipartite graph $G = (U \cup V, E)$ is $K_{2,r}$-free, then $|E| \leq |V|^{\frac{3}{2} - \frac{1}{4d-2} + o(1)}$.*

Let us remark that Theorem 1.2 in [9] is stated for incidences between points and varieties, but the proof remains valid for semi-algebraic relations up to a constant factor depending on r. We also note that a result of Sheffer [15] shows that Theorem 2.2 is tight up to the $o(1)$ factor in the exponent. We will use the $d = 4$ special case of Theorem 2.2. We also need Vizing's theorem.

Lemma 2.3 ([17])**.** *Let $G = (V, E)$ be a graph with maximum degree p. Then the edges of G can be partitioned into $p + 1$ matchings.*

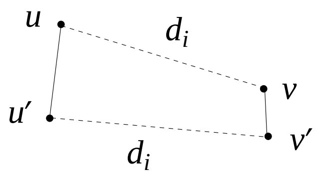
We are now ready to prove Theorem 1.

3 Proof of Theorem 1

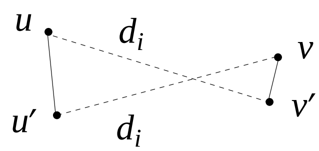
Let $p \geq 6$ be a fixed integer. We want to show that

$$D\left(n, p, \binom{p}{2} - p + 6\right) = \Omega\left(n^{1 + \frac{1}{7+\delta}}\right),$$

where δ is an arbitrarily small constant. Let V be an n-element planar point set such that any p points from V determine at least $q = \binom{p}{2} - p + 6$ distinct distances. Suppose V determines x distinct distances $d_1, \ldots, d_x$ with multiplicity $m_1, \ldots, m_x$, respectively, where $m_1 + \cdots + m_x = \binom{n}{2}$. Notice we have the following simple claim.



(a) $|u-v| = |u'-v'|$.



(b) $|u-v'| = |u'-v|$.

Figure 1: Edge (uu', vv') in G.

Claim 3.1. For any i and for any $u \in V$, there are at most $p-5$ points of V at distance d_i from u.

Together with Vizing's theorem, we have the following.

Corollary 3.2. *For any i, the pairs of points in V at distance d_i can be partitioned into at most $p-5+1 < p$ matchings.*

By partitioning the pairs of points in V at distance d_i into p matchings, let $m_{i,j}$ denote the size of the jth matching ($0 \leq m_{i,j} \leq n/2$).

Let us fix the lexicographic ordering of the points in V. We define a new point set $W = \binom{V}{2}$ in $\mathbb{R}^4$, where $uv \in W$ if and only if $u, v \in V$ and u comes before v in the ordering. Let G be the graph with vertex set $W = \binom{V}{2}$, and $(uu', vv') \in E(G)$ if and only if u, u', v, v' are distinct elements of V, and $|u-v| = |u'-v'|$ or $|u-v'| = |u'-v|$. See Figure 1. Clearly, $E(G)$ is a semi-algebraic relation with description complexity at most four. We can assume that $x < \binom{n}{2}/(10p)$, since otherwise we are done. Therefore, by the Jensen's inequality, we have

$$|E(G)| \geq \sum_{i=1}^{x}\sum_{j=1}^{p}\binom{m_{i,j}}{2} \geq xp\binom{\sum_i\sum_j \frac{m_{i,j}}{xp}}{2} \geq xp\binom{\binom{n}{2}/(xp)}{2} \geq \frac{n^4}{9xp},$$

provided that n is sufficiently large. Hence, we have

$$x \geq \frac{n^4}{9p} \cdot \frac{1}{|E(G)|}, \tag{1}$$

and it is sufficient to bound $|E(G)|$ from above. By a standard probabilistic argument, we can partition $W = W_1 \cup W_2$ such that at least half of the edges in G are between W_1 and W_2 and $|W_1|, |W_2| \geq \lfloor |W|/2 \rfloor$. Let G' be the bipartite graph with parts W_1, W_2, such that $E(G') = \{(uu', vv') \in E(G) : uu' \in W_1, vv' \in W_2\}$. Therefore, it is enough to bound the number of edges in G'.

Fix a vertex $u_1u_2 \in W_1$, and let $N(u_1u_2) \subset W_2$ such that

$$N(u_1u_2) = \{v_1v_2 \in W_2 : (u_1u_2, v_1v_2) \in E(G')\}.$$

Consider the graph G_0 with $V(G_0) = V$ and $E(G_0) = N(u_1u_2)$. Then, applying the following lemma with $r = p$, we obtain that the maximum degree of the vertices of G_0 is less than $p-3$.

Lemma 3.3. *For $r \geq 4$, suppose there is a vertex v with degree $r-3$ in G_0 with neighbors $w_1, \ldots, w_{r-3}$. Then the points $u_1, u_2, v, w_1, \ldots w_{r-3}$ determine at most $\binom{r}{2} - r + 3$ distinct distances.*

Proof. We proceed by induction on r. The base case $r = 4$ follows since $(u_1u_2, vw_1) \in E(G')$. Now assume that the statement holds up to $r-1$. By the induction hypothesis, $u_1, u_2, v, w_1, \ldots, w_{r-4}$ determine at most

$$\binom{r-1}{2} - (r-1) + 3 = \binom{r}{2} - 2(r-1) + 3$$

distinct distances. Since $(u_1u_2, vw_{r-3}) \in E(G')$, we have either $|u_1 - v| = |u_2 - w_{r-3}|$ or $|u_1 - w_{r-3}| = |u_2 - v|$. Thus, adding w_{r-3} introduces at most $r-2$ new distances. Therefore, $u_1, u_2, v, w_1, \ldots w_{r-3}$ determine at most

$$\binom{r}{2} - 2(r-1) + 3 + (r-2) = \binom{r}{2} - r + 3$$

distinct distances, as required. ☐

Lemma 3.4. *Suppose that the vertices $u_1u_2, u_3u_4 \in W_1$ and $v_1v_2, v_3v_4, \ldots, v_{2r-1}v_{2r} \in W_2$ induce a $K_{2,r}$ in G', such that $v_1, v_2, \ldots, v_{2r-1}, v_{2r}$ are distinct points of V. Then there are $2r+4$ points in V that determine at most $\binom{2r+4}{2} - 2r$ distinct distances.*

Proof. The proof falls into two cases: either u_1, u_2, u_3, u_4 are distinct, or we can assume without loss of generality that $u_2 = u_3$, say.

Case 1. Suppose u_1, u_2, u_3, u_4 are all distinct. Since $(u_1u_2, v_iv_{i+1}), (u_3u_4, v_iv_{i+1}) \in E(G')$, for every odd integer $i \in \{1, 3, 5, \ldots, 2r-1\}$, we get $2r$ elements of $E(G')$, and each such element gives a repeated distance. Hence, the number of repetitions is at least $2r$, so $u_1, u_2, u_3, u_4, v_1, \ldots v_{2r}$ determine at most $\binom{2r+4}{2} - 2r$ distinct distances.

Case 2. Suppose $u_2 = u_3$. In view of Lemma 3.3 (with $r = 5$), the five points u_1, u_2, u_4, and v_i, v_{i+1}, for i odd, determine at most $\binom{5}{2} - 2$ distinct distances. Hence, $u_1, u_2, u_4, v_1, \ldots v_{2r}$ determine at most

$$\binom{2r+3}{2} - 2r$$

distinct distances. Now adding any point w to $u_1, u_2, u_4, v_1, \ldots v_{2r}$ gives us $2r+4$ points that determine at most

$$\binom{2r+3}{2} - 2r + (2r+3) = \binom{2r+4}{2} - 2r$$

distinct distances. ☐

Suppose that p is even and recall that $p \geq 6$. Then the bipartite graph G' is $K_{2,\frac{(p-3)(p-4)}{2}}$-free. Indeed, otherwise by Lemmas 3.3 and 2.3, we would have vertices $u_1u_2, u_3u_4 \in W_1$ and $v_1v_2, \ldots, v_{p-3}v_{p-4} \in W_2$ that induce a $K_{2,\frac{p-4}{2}}$ in G', such that the points $v_1, v_2, \ldots, v_{p-3}, v_{p-4}$ are distinct elements of V. Then by Lemma 3.4, we would have p points in V that determine at most $\binom{p}{2} - p + 4$ distinct distances, which is a contradiction. Therefore, applying Theorem 2.2 with $d = 4$ we obtain

$$|E(G)| \leq 2 \cdot |E(G')| \leq O\left(|W|^{\frac{3}{2} - \frac{1}{14+\varepsilon}}\right) \leq O\left(n^{3 - \frac{2}{14+\varepsilon}}\right),$$

where $\varepsilon = 2\delta$. Together with (1), we get

$$x \geq \Omega\left(\frac{n^4}{|E(G)|}\right) \geq \Omega\left(n^{1+\frac{2}{14+\varepsilon}}\right).$$

If p is odd, then G' is $K_{2,\frac{(p-3)(p-5)}{2}}$-free. Indeed, otherwise by Lemmas 3.3 and 2.3, we would have vertices $u_1u_2, u_3u_4 \in W_1$ and $v_1v_2, \ldots, v_{p-4}v_{p-5} \in W_2$ that induce a $K_{2,\frac{p-5}{2}}$ in G', such that the points $v_1, v_2, \ldots, v_{p-4}, v_{p-5}$ are all distinct. Then by Lemma 3.4, we would have $p-1$ points that determine at most

$$\binom{p-1}{2} - p + 5 = \binom{p}{2} - 2p + 6$$

distinct distances. Adding any point to our collection would give us p points that determine at most $\binom{p}{2} - p + 5$ distinct distances, a contradiction. Just as above, we have

$$|E(G)| \leq O\left(|W|^{\frac{3}{2}-\frac{1}{14+\varepsilon}}\right) \leq O\left(n^{3-\frac{2}{14+\varepsilon}}\right).$$

Combining this with (1), we obtain

$$x \geq \Omega\left(\frac{n^4}{|E(G)|}\right) \geq \Omega\left(n^{1+\frac{2}{14+\varepsilon}}\right) = \Omega\left(n^{1+\frac{1}{7+\delta}}\right).$$

This completes the proof of Theorem 1. □

References

[1] F. A. Behrend, On sets of integers which contain no three terms in arithmetical progression, *Proc. Nat. Acad. Sci. U. S. A.* **32** (1946), 331–332.

[2] P. Brass, W. Moser, and J. Pach, *Research Problems in Discrete Geometry*, Berlin, Germany, Springer-Verlag, 2005.

[3] A. Dumitrescu, On distinct distances among points in general position and other related problems, *Period. Math. Hungar.* **57** (2008), 165–176.

[4] P. Erdős, On sets of distances of n points, *Amer. Math. Monthly* **53** (1946), 248–150.

[5] P. Erdős, On some metric and combinatorial geometric problems, *Discrete Math.* **60** (1986), 147–153.

[6] P. Erdős, Some of my recent problems in combinatorial number theory, geometry and combinatorics, in: *Graph Theory, Combinatorics, Algorithms and Applications, vol. 1* (Y. Alavi et al., eds.), Wiley 1995, 335–349.

[7] P. Erdős, Z. Füredi, J. Pach, and I.Z. Ruzsa, The grid revisited, *Discrete Math.* **111** (1993), 189–196.

[8] P. Erdős and A. Gyárfás, A variant of the classical Ramsey problem, *Combinatorica* **17** (1997), 459–467.

[9] J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl, A semi-algebraic version of Zarankiewicz's problem, *J. Eur. Math. Soc.*, to appear. Preprint arXiv:1407.5705, 2015.

[10] L. Guth, and N. Katz, On the Erdős distinct distances problem in the plane, *Ann. Math.* **181** (2015), 155–190.

[11] P. Kővári, V. Sós, and P. Turán, On a problem of Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.

[12] K. F. Roth, On certain sets of integers, *J. London Math. Soc.* **28** (1953), 104–109.

[13] G. Sárkőzy and S. Selkow, On edge colorings with at least q colors in every subset of p vertices, *Electron. J. Combin.* **8** (2001), no. 1.

[14] A. Sheffer, Distinct distances: open problems and current bounds, preprint arXiv:1406.1949, 2015.

[15] A. Sheffer, Lower bounds for incidences with hypersurfaces, preprint arXiv:1511:03298, 2015.

[16] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.

[17] V. Vizing, On an estimate of the chromatic class of a p-graph, *Diskret. Analiz.* **3** (1964), 25–30.