New results on the distribution of distances determined by separated point sets *

E. Makai, Jr.[1], J. Pach[2], and J. Spencer[3]

Abstract. In this mini-survey, we summarize some joint results found by Paul Erdős and the authors during the past decade on the set of distances between n points in Euclidean space. We concentrate on two types of questions: (1) How uniformly can the $\binom{n}{2}$ distances determined by such a point set be distributed? (2) What is the maximum number of distances that can lie in the union of k intervals of length 1, provided that the minimal distance is at least 1?

1 Introduction

Ever since Paul Erdős rediscovered Sylvester's conjecture [35], when he was a student, he was fascinated by geometric problems with a combinatorial flavor. According to this celebrated result, today known as the Gallai-Sylvester theorem [28], any finite set of non-collinear points in the plane has two elements such that the line determined by them does not pass through any other element. In the dual setting: any finite set of non-concurrent straight lines (no two of which are parallel) determines an intersection point incident to precisely two elements of the set.

This result has inspired a lot of research. Recently, Rom Pinchasi [32] verified an analogous conjecture of András Bezdek [4, 5] for circles. He proved that any set of at least 5 pairwise intersecting unit circles in the plane determines an intersection point incident to precisely *two* elements of the set. It is conjectured that if we only assume that at least two unit circles have a point in common, then there exists an intersection point incident to at most *three* elements of the set. For some other results in this direction, see [1].

*Research supported by Hungarian National Foundation for Scientific Research, grants no. T-030012, T-020914, and by NSF grants.

We mention two other well-known conjectures related to the Sylvester-Gallai theorem.

Dirac's Conjecture [15]: *Any sufficiently large set of n non-collinear points determines at least $n/2$ lines, each of which passes through precisely two elements of the set.*

Fukuda's Conjecture: [14]: *Any non-collinear set of n red and n blue points in the plane, separated by a line, contains a red point and a blue point such that their connecting line does not pass through any other element of the set.*

The best known results concerning these questions are due to Csima and Sawyer [12, 13] and Pach and Pinchasi [31], respectively.

Erdős often said that even Euclid would appreciate these problems, because they involved only the most elementary concepts of geometry. Incidence relations between points and lines in the Euclidean and projective planes came to new life as objects of combinatorial study. The Gallai-Sylvester theorem has also contributed a great deal to the development of the theory of combinatorial designs [2].

In 1946, in a paper published in the *American Mathematical Monthly* [17], Erdős raised a new type of question: At most how many times can the unit distance occur among n points in the plane (more generally, in Euclidean or other metric spaces)?

It is conjectured that in the plane the maximum is $n^{1+\text{const}/\log\log n}$, which is attained by a $\sqrt{n} \times \sqrt{n}$ piece of a square lattice, but the best currently known upper estimate, due to Spencer, Szemerédi, and Trotter [34], is only $O(n^{4/3})$ (see [36] for a simple proof). In 3-space, the best known upper bound is $n^{3/2}\beta(n)$, where $\beta(n)$ is an extremely slowly increasing function, closely related to the *inverse Ackermann function* [11], but the truth is probably closer to $n^{4/3}$. In higher dimensions, we know the asymptotically tight answer [20], [26], [30]. The maximum number of times that the same distance can occur among n distinct points in k-space is

$$\begin{array}{ll} n^2\left(\frac{1}{2}-\frac{1}{k}\right)+O(n) & \text{if } k \geq 4 \text{ is even;} \\ n^2\left(\frac{1}{2}-\frac{1}{k-1}\right)+O(n^{4/3}) & \text{if } k \geq 5 \text{ is odd.} \end{array}$$

he unit distance problem is essentially equivalent to the following question about incidences: What is the maximum number of incidences between n points and n unit circles in the plane (spheres in space)?

In a more general setting, Aronov and Sharir [3] have recently proved that the maximum number of incidences between m points and n not necessarily congruent circles in the plane is

$$\begin{array}{ll} O(m^{2/3}n^{2/3}+m) & \text{if } m \geq n^{3/2}; \\ O(m^{4/7}n^{17/21}+n) & \text{if } m \leq n^{3/2}, \end{array}$$

and the first estimate cannot be improved. Incidence estimates of this kind may lead to an improvement of Székely's result [36] that every set of n points in the plane determines at least constant times $n^{4/5}$ distinct distances (which is only slightly better than the best previously known bound of Chung, Szemerédi, and Trotter [10]). We note that most recently const $\cdot n^{4/5}$ was improved to const $\cdot n^{6/7}$ by J. Solymosi, Cs. Tóth [39], and subsequently this was improved to $\text{const}_\varepsilon \cdot n^{4/(5-e^{-1})-\varepsilon}$ ($\varepsilon > 0$ arbitrary), by G. Tardos [40]. It is conjectured that the optimal configuration is lattice-like, and it gives roughly constant times $n/\sqrt{\log n}$ (more exactly $\Theta(n/\sqrt{\log n})$) distinct distances.

In the last decade Erdős's attention turned to some interesting new variants of the above questions, which *cannot* be reformulated in terms of incidences between points and curves (surfaces). For instance, what is the minimum diameter of a set of n points in the plane with the property that any two distances determined by them are either equal or differ by at least *one*? (It follows from the above mentioned result of Tardos [40] that this number, the largest distance determined by such a point set, is at least constant times $n^{4/(5-e^{-1})-\varepsilon}$, but it is possible that the truth is $n-1$ or at least linear in n.)

In the present mini-survey, we discuss two questions posed by Erdős that belong to this category. It seems that they are related to problems in integral geometry and in the theory of packing and covering rather than to incidences. Most of the results were obtained in collaboration with Erdős. The proofs will be published elsewhere.

2 Uniformity of distance distributions

Given n points in $\mathbf{R^k}$, the k-dimensional Euclidean space, let $d_1 \leq ... \leq d_{\binom{n}{2}}$ denote all the distances determined by them, with multiplicity. Paul Erdős asked the following question. Can it occur that these distances are, in some sense, "uniformly distributed" in the interval $[d_1, d_{\binom{n}{2}}]$? To make his question more concrete, Erdős asked for the minimum of

$$S := \sum_{j=1}^{\binom{n}{2}-1} (d_{j+1}-d_j)^2,$$

provided that $d_1 \geq 1$, i.e., the point set is *separated.* Obviously, here we can and will assume without loss of generality that $d_1 = 1$. Observe, that if the distances could be "uniformly distributed" in $[d_1, d_{\binom{n}{2}}]$, i.e., they could form an arithmetic progression, then it would follow from the arithmetic–quadratic mean inequality that the minimum is

$$\frac{\left(d_{\binom{n}{2}} - 1\right)^2}{\binom{n}{2} - 1}.$$

Since our point set is separated, we have that $d_{\binom{n}{2}}$ is at least constant times $n^{1/k}$. Therefore, $\min S \geq \text{const} \cdot n^{-2(1-1/k)}$.

However, this estimate is far from being sharp. Instead, we have

Theorem 1.1 [25] *For every separated set of n points on the line ($k = 1$), we have*

$$S \geq \text{const} \cdot \log n.$$

This inequality is tight up to the value of the constant.

That is, in the 1-dimensional case, the sum of the squares of the gaps between the consecutive distances, somewhat surprisingly, tends to infinity as $n \to \infty$. This is not the case in higher dimensions.

Theorem 1.2 [33] *For every separated set of n points in the plane ($k = 2$), we have*

$$S \geq \text{const} \cdot n^{-6/7}.$$

This result is tight up to the value of the constant.

In place of the *squares* of the gaps between the distances, one could use any other *increasing convex function* of the gaps to define a competing measure of uniformity. In this case, the arithmetic–quadratic mean inequality has to be replaced by Jensen's inequality. In particular, for any $c > 1$, we may wish to minimize the quantity

$$S_c^k := \sum_{j=1}^{\binom{n}{2}-1} (d_{j+1} - d_j)^c.$$

The linear case is essentially settled by

Theorem 1.3 [25] *For every separated set of n points on the line ($k = 1$), we have*

$$S_c^1 \geq \text{const}_c \cdot \begin{cases} n^{2-c} & \textit{if} \quad 1 < c < 2; \\ \log n & \textit{if} \quad c = 2; \\ 1 & \textit{if} \quad c > 2. \end{cases}$$

For every fixed c, this bound is tight up to the value of the constant.

This result can be generalized to higher dimensions, as follows.

Theorem 1.4 [23] *For any positive integer k, for every separated set of n points in* $\mathbf{R^k}$*, we have*

$$S_c^k \geq \text{const}_{c,k} \cdot \begin{cases} n^{\frac{c}{k}-2(c-1)} & \text{if} \quad 1 < c < \frac{k+3}{k+1}; \\ n^{-\frac{3(k-1)}{k(k+1)}} \cdot \log^{\frac{k+3}{4k}} n & \text{if} \quad c = \frac{k+3}{k+1}; \\ n^{-\frac{6(k-1)(c-1)^2}{3(k-1)(c-1)+2c}} & \text{if} \quad c > \frac{k+3}{k+1}. \end{cases}$$

According to Theorem 1.3, the last result is asymptotically tight for $k = 1$. By Theorem 1.2, it is also tight in the case $k = 2$, $c = 2$. We conjecture that, at least for small values of c, Theorem 1.4 is not far from being optimal.

To prove our lower bounds, we had to strengthen the following classical result of Bieberbach [9] (Ch. 44, 54): the volume of any k-dimensional convex body of diameter d is at most as large as the volume of a k-dimensional ball with the same diameter. Our volume estimate depends on one further parameter e, and Bieberbach's inequality can be obtained from it by integration with respect to e. More precisely, we have

Theorem 1.5 [23] *Let* $D \subset \mathbf{R}^k$ *be a convex body of diameter d, and let* $0 < e < d$*. Then*

$$\underset{|p-q|=e}{\int_D \int_D} dp\, dq \leq e^{k-1} \cdot k\kappa_k \cdot \text{Vol}(B(p_0, d/2) \cap B(q_0, d/2)),$$

where p_0 *and* q_0 *are any two points at distance e from each other,* $B(x,r)$ *denotes the ball of radius r centered at x, and* κ_k *is the volume of the unit ball* $B(0,1) \subset \mathbf{R}^k$*.*

Equality holds here if and only if D is a ball of diameter d.

The tightness of Theorems 1.1 and 1.3 can be established by explicit constructions. In the case of Theorem 1.2, we had to start with a random construction given by a Poisson process and modify it "by hand."

3 Nearly equal distances

During the past fifty years, Erdős almost "systematically" raised and partially answered a wide range of questions about the distribution of all *sums,* *differences,* or *distances* determined by a set of n *numbers* or *points* in various metric spaces [21]. A famous example is the following problem due to Littlewood and Offord [16]. Given a set of n non-zero numbers, at most how many of the 2^n partial sums formed by them can coincide? The maximum, determined in [16], is $\binom{n}{\lfloor n/2 \rfloor}$. Interestingly, it follows from the proof that if every number has absolute value at least 1, and we wish to maximize the number of partial sums that are *nearly equal*, i.e., which fall into an open interval of length 1, the answer remains unchanged.

Another famous question of this type is the one mentioned in the Introduction: Given n points in the plane (or, more generally, in k-space), at most how many of the $\binom{n}{2}$ interpoint distances can coincide? Erdős observed that, unlike in the case of the Littlewood-Offord problem, the answer to the above question does not remain the same if we count the number of distances that are *nearly equal.* To exclude the trivial examples when all distances are nearly equal to 0, we consider only *separated* point sets, i.e., we assume that the minimum distance between two points is at least 1 (which can be attained by proper scaling). What is the maximum number of pairs of points in an n-element set with distance in a closed unit interval $[t, t+1]$?

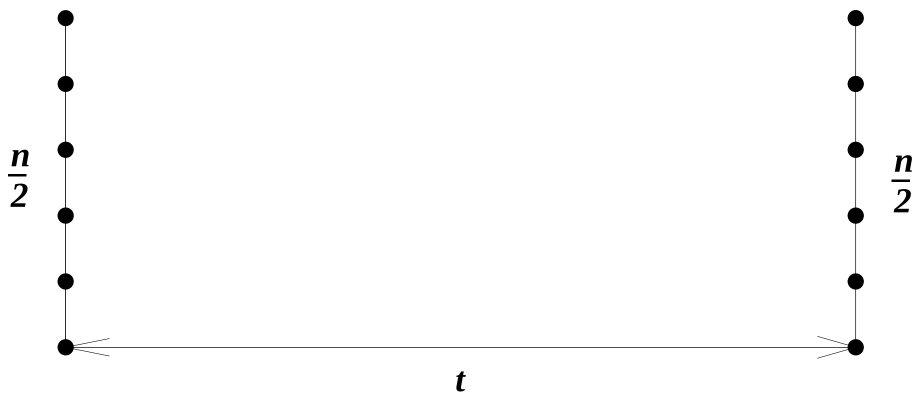

Figure 1.

The following construction shows that in the plane, for a suitable value of t, this number is at least $\lfloor n^2/4 \rfloor$ (see Fig. 1). Take a set V of n points with coordinates

$$(0,1),(0,2),\ldots,(0,\lfloor n/2 \rfloor);(t,1),(t,2),\ldots,(t,\lceil n/2 \rceil).$$

The distance between any element of V whose x-coordinate is 0 and any other element whose x-coordinate is t belongs to an interval $[t, t+O(1/t)]$.

Thus, if t is large enough, $\lfloor n^2/4 \rfloor$ pairs will have a distance lying in $[t, t+1]$. Erdős conjectured that this configuration is optimal, provided that n is sufficiently large.

Theorem 2.1 [25] *Let P be a separated set of n points in the plane. If n is sufficiently large, then for every $t > 0$, the number of pairs of points in P whose distance lies in the interval $[t, t+1]$ is at most $\lfloor n^2/4 \rfloor$. This bound can be attained for every $t \geq t(n)$, where $t(n)$ is a suitable function of n.*

Let $T(k, n)$ denote the number of edges of a *balanced complete k-partite graph* on n vertices, i.e., a graph whose vertices are partitioned into k groups of size $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$, and two vertices are connected by an edge if and only if they belong to different groups. Clearly, we have $T(k, n) = \frac{n^2}{2}\left(1 - \frac{1}{k} + o(1)\right)$.

To generalize Theorem 2.1 to higher dimensions, we need a sharper version of the following result of Turán [38] (see also [30], [8]). Any graph of n vertices, which does not contain a complete subgraph with $k+1$ vertices, has at most $T(k, n)$ edges.

Theorem 2.2 [25] *Let $k \geq 2$ be an integer, and let P be a separated set of n points in $\mathbf{R^k}$. If n is sufficiently large, then for every $t > 0$, the number of pairs of points in P whose distance lies in the interval $[t, t+1]$ is at most $T(k, n)$. This bound can be attained for every $t \geq t(k, n)$, where $t(k, n)$ is a suitable function of k and n.*

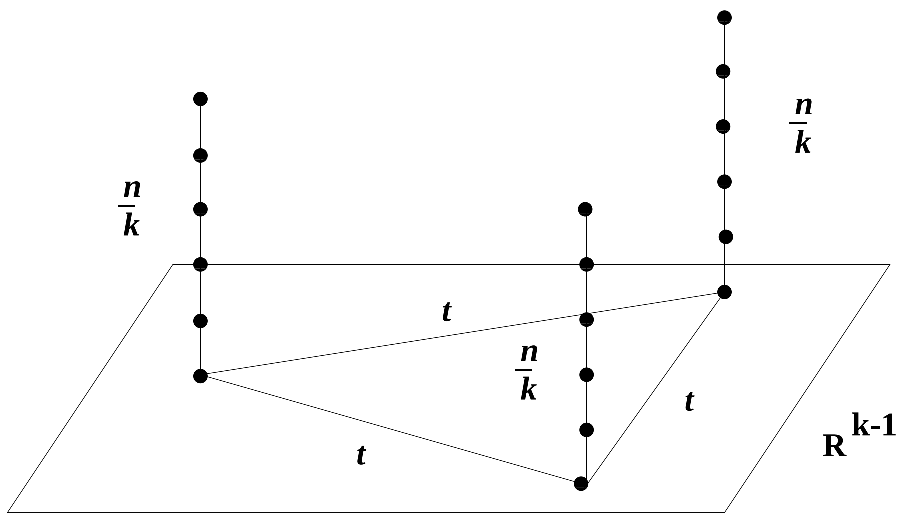

Figure 2.

The tightness of the bound in Theorem 2.2 can be shown by straightforward generalization of the construction preceding Theorem 2.1. Let t be a sufficiently large number, and let $v_1, v_2, \ldots, v_k$ be the vertices of a regular $(k-1)$-dimensional simplex in the hyperplane $x_k = 0$, with edge-length t. At each v_i draw a perpendicular to the hyperplane $x_k = 0$, and on each of these lines pick $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$ distinct points whose x_k-coordinates are integers between 0 and n/k, so that the total number of points is n (see Fig. 2). If t is sufficiently large, depending on k and n, the distance between any pair of points selected on different perpendiculars belongs to the interval $[t, t+1]$.

It is not hard to extend Theorem 2.1 in another direction. Suppose that, instead of maximizing the number of "nearly equal" distances, we want to estimate the largest number of distances "nearly equal" to (at least) one of p preselected values.

Theorem 2.3 [22] *Let P be a separated set of n points in the plane, and let p be a positive integer.*

For any $t_1, t_2, \ldots, t_p > 0$, the number of pairs of points in P whose distance lies in $[t_1, t_1+1] \cup [t_2, t_2+1] \cup \ldots \cup [t_p, t_p+1]$ is at most

$$\frac{n^2}{2}\left(1 - \frac{1}{p+1} + o(1)\right).$$

This estimate is tight for every fixed p and for some $t_1 = t_1(p,n), t_2 = t_2(p,n), \ldots, t_p = t_p(p,n)$.

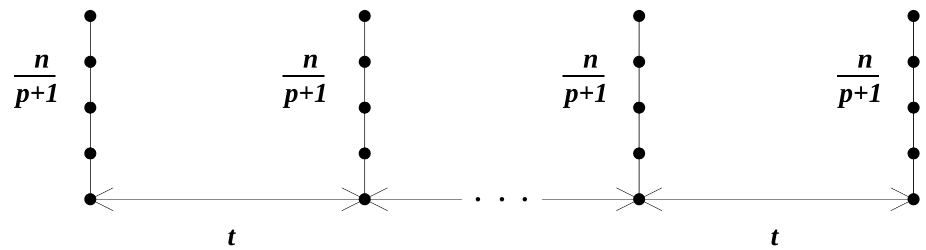

Figure 3.

To show the tightness, now consider the points

$$(0,1), (0,2), \ldots, (0,n_0); (t,1), (t,2), \ldots, (t,n_1); \ldots; (pt,1), (pt,2), \ldots, (pt,n_p),$$

where $n_0, ..., n_p$ are equal either to $\lfloor n/(p+1) \rfloor$ or $\lceil n/(p+1) \rceil$, and $n_0 + ... + n_p = n$ (see Fig. 3). If t is large enough, all distances between points in different groups lie in $[t, t+1] \cup [2t, 2t+1] \cup \ldots \cup [pt, pt+1]$. It is conjectured that this construction is not only *asymptotically* optimal, but it also gives the *exact* maximum, provided that n is sufficiently large.

It appears to be technically difficult to generalize Theorem 2.3 to higher dimensions. The only case we can settle is $p = 2$.

For any $k \geq 1$, let m_k denote the largest size of a point set in $\mathbf{R}^\mathbf{k}$ determining *two* distinct distances. Clearly, $m_1 = 3$ and $m_2 = 5$. We know the exact values of m_k for $k \leq 8$ (see [29]), and it is also known [6], [7] that

$$\binom{k+1}{2} \leq m_k \leq \binom{k+2}{2}$$

holds for every k.

Theorem 2.4 [24] *Let $k \geq 2$ be an integer, and let P be a separated set of n points in $\mathbf{R}^\mathbf{k}$.*

For any $t_1, t_2 > 0$, the number of pairs of points in P whose distance lies in $[t_1, t_1 + 1] \cup [t_2, t_2 + 1]$ is at most

$$\frac{n^2}{2}\left(1 - \frac{1}{m_{k-1}} + o(1)\right).$$

For every fixed k and for some $t_1 = t_1(k, n)$, $t_2 = t_2(k, n)$, this estimate is tight.

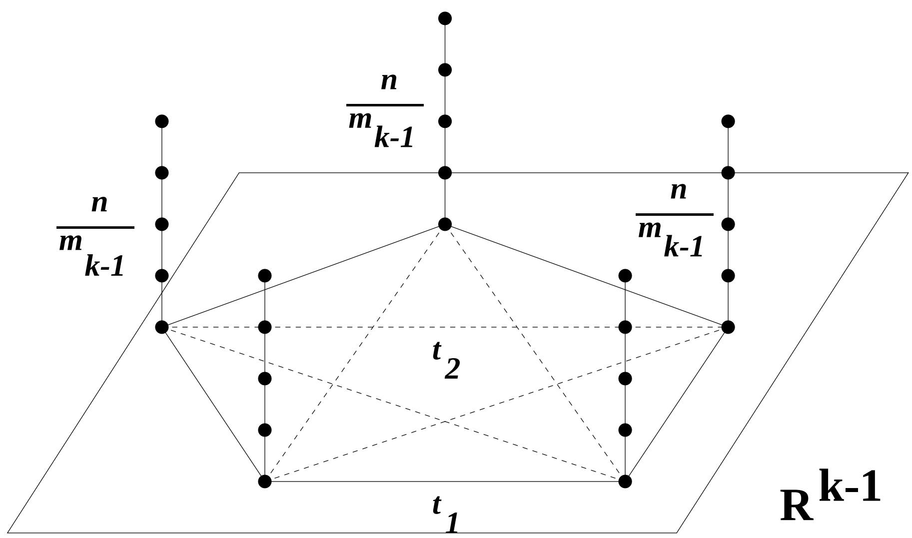

Figure 4.

It is not hard to modify the previous constructions to obtain an example proving that the last result is also tight. Choose an m_{k-1}-element set X in the hyperplane $x_k = 0$, which determines *two* (non-zero) distances. By proper scaling, we can achieve that even the smaller distance is arbitrarily large. At each point of X draw a perpendicular to the hyperplane $x_k = 0$, and on each of these lines pick $\lfloor n/m_{k-1} \rfloor$ or $\lceil n/m_{k-1} \rceil$ distinct points whose x_k-coordinates are integers between 0 and n/m_{k-1}, so that the total number of points is n (see Fig. 4).

Note that if $k \neq 4, 5$ and n is sufficiently large, we can prove the stronger statement that the number of point pairs whose distance lies in $[t_1, t_1 + 1] \cup [t_2, t_2 + 1]$ is at most $T(m_{k-1}, n)$. It is quite possible that the same is true for $k = 4$ and 5.

The proofs of all statements in this section are based on some standard results in extremal graph [8] and hypergraph theory [19], such as relatives of the Erdős-Stone theorem [27] and Szemerédi's Regularity Lemma [37].

References

[1] N. Alon, H. Last, R. Pinchasi, and M. Sharir, On the complexity of arrangements of circles in the plane, *Discrete Comput. Geom.*, to appear.

[2] I. Anderson, *Combinatorial Designs and Tournaments,* Oxford Lecture Series in Mathematics and its Applications **6**, Oxford University Press, New York, 1997.

[3] B. Aronov and M. Sharir, Improved bounds for incidences and complexity of many faces in arrangements of circles (manuscript), Tel Aviv University, 2000.

[4] A. Bezdek, On the density of dual circle coverings, *Geometriae Dedicata* **33** (1990), 227–238.

[5] A. Bezdek, On the intersection points of unit circles (Open problem), *Amer. Math. Monthly* **99** (1992), 779–780.

[6] A. Blokhuis, A new upper bound for the cardinality of 2-distance sets in Euclidean space, in: *Convexity and Graph Theory (Jerusalem, 1981), North-Holland Math. Stud.* **87**, North-Holland, Amsterdam-New York, 1984, 65–66.

[7] A. Blokhuis, J. J. Seidel, Few-distance sets in $R^{p,q}$, in: *Symposia Mathematica* Vol. **XXVIII** *(Rome, 1983)*, Academic Press, London-New York, 1986, 145–158.

[8] B. Bollobás, *Extremal Graph Theory,* London Mathematical Society Monographs **11**, Academic Press, London, 1978.

[9] T. Bonnesen, W. Fenchel, *Theorie der konvexen Körper,* Berichtigter Reprint, Springer-Verlag, Berlin-New York, 1974

[10] F. Chung, E. Szemerédi, and W. T. Trotter, The number of different distances determined by a set of points in the Euclidean plane, *Discrete Comput. Geom.* **7** (1992), 1–11.

[11] K. L. Clarkson, H. Edelsbrunner, L. J. Guibas, M. Sharir, E. Welzl, Combinatorial complexity bounds for arrangements of curves and surfaces, *Discr. Comput. Geom.* **5** (1990), 99–160.

[12] J. Csima and E. T. Sawyer, There exist $6n/13$ ordinary points, *Discrete Comput. Geom.* **9** (1993), 187–202.

[13] J. Csima and E. T. Sawyer, The $6n/13$ theorem revisited, in: *Graph Theory, Combinatorics, and Algorithms, Vol. 1, 2 (Kalamazoo, MI, 1992)*, Wiley-Intersci. Publ., Wiley, New York, 1995, 235–249.

[14] I. P. F. Da Silva and K. Fukuda, Isolating points by lines in the plane, *J. Geom.* **62** (1998), 48–65.

[15] G. Dirac, Collinearity properties of sets of points, *Quarterly J. Math. (2)* **2** (1951), 221–227.

[16] P. Erdős, On a lemma of Littlewood and Offord, *Bull. Amer. Math. Soc.* **51** (1945), 898–902.

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

[18] P. Erdős, On sets of distances of n points in Euclidean space, *Magyar Tud. Akad. Mat. Kutató Int. Közl. (Publ. Math. Inst. Hungar. Acad. Sci.)*, **5** (1960), 165–169.

[19] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190.

[20] P. Erdős, On some applications of graph theory to geometry, *Canad. J. Math.* **19** (1967), 968–971.

[21] P. Erdős, *The Art of Counting: Selected Writings (J. Spencer, ed.)*, M. I. T. Press, Cambridge, MA, 1973.

[22] P. Erdős, E. Makai, J. Pach, Nearly equal distances in the plane, *Combinatorics, Probability and Computing* **2** (1993), 401–408.

[23] P. Erdős, E. Makai, Jr., J. Pach, unpublished manuscript.

[24] P. Erdős, E. Makai, Jr., J. Pach, unpublished manuscript.

[25] P. Erdős, E. Makai, J. Pach, J. Spencer, Gaps in difference sets, and the graph of nearly equal distances, *Appl. Geom. and Discr. Math., The V. Klee Festschrift,* 1991, 265–273.

[26] P. Erdős, J. Pach, Variations on the theme of repeated distances, *Combinatorica* **10** (1990), 261–269.

[27] P. Erdős, A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091.

[28] T. Gallai, Solution of problem 4065, *Amer. Math. Monthly* **51** (1944), 169–171.

[29] P. Lisoněk, New maximal two-distance sets, *J. Combin. Theory, Ser. A* **77** (1997), 318–338.

[30] J. Pach, P. K. Agarwal, *Combinatorial Geometry,* J. Wiley and Sons, New York, 1995.

[31] J. Pach and R. Pinchasi, Bichromatic lines with few points, *J. Combin. Theory Ser. A* **90** (2000), 326–335.

[32] R. Pinchasi, Gallai-Sylvester theorem for pairwise intersecting unit circles in the plane (manuscript), Hebrew University, 1999.

[33] J. Pach, J. Spencer, Uniformly distributed distances — a geometric application of Janson's inequality, *Combinatorica* **19** (1999), 111–124.

[34] J. Spencer, E. Szemerédi, W. T. Trotter, Jr., Unit distances in the Euclidean plane, *Graph Theory and Combinatorics,* Academic Press, London–New York, 1984, 293–303.

[35] J. Sylvester, Mathematical question 11851, *Educational Times* **59** (1893), 98–99.

[36] L. Székely, Crossing numbers and hard Erdős problems in discrete geometry, *Combin. Probab. Comput.* **6** (1997), 353–358.

[37] E. Szemerédi, Regular partitions of graphs, in: *Problèmes Combinatoires et Théorie des Graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, Colloques Internat. CNRS Vol. **260**, CNRS, Paris, 1978, 399–401.

[38] P. Turán, On the theory of graphs, *Colloquium Math.* **3** (1954), 19–30.

[39] J. Solymosi, Cs. Tóth, Distinct distances in the plane, *Discr. Comput. Geom.* **25** (2001), 629-634.

[40] G. Tardos, in preparation.

[1] Alfréd Rényi Mathematical Institute, Hungarian Academy of Sciences, H-1364 Budapest, Pf. 127, HUNGARY

[2] Alfréd Rényi Mathematical Institute, Hungarian Academy of Sciences, H-1364 Budapest, Pf. 127, HUNGARY, and Courant Institute of Mathematical Sciences, New York University

[3] Courant Institute of Mathematical Sciences, New York University

E-mail: `makai@math-inst.hu, pach@cims.nyu.edu, spencer@cs.nyu.edu`