MIDPOINTS OF SEGMENTS INDUCED BY A POINT SET

János Pach*
Courant Institute, NYU
and Hungarian Academy of Sciences

Abstract

Applying some well known results in additive number theory, we partially answer two geometric questions due to V. Bálint et al. and F. Hurtado. (1) Let $m(n)$ be the largest integer m with the property that from every set of n points in the plane one can select m elements so that none of them is the midpoint of two others. It is shown that $n^{1-c/\sqrt{\log n}} \leq m(n) \leq n/\log^{c'} n$. (2) Let $\mu(n)$ be the smallest number of distinct midpoints of all segments induced by n points in the plane, no 3 of which are collinear. It is proved that $\lim_{n\to\infty} \mu(n)/n = \infty$ and that $\mu(n) \leq ne^{c''\sqrt{\log n}}$. Here c, c', and c'' denote suitable positive constants.

1 Introduction

Many extremal problems in discrete geometry lead to questions in additive number theory [12]. This is partly due to the fact that their solutions are known or conjectured to be lattice-like, i.e., affinely equivalent to the integer lattice. Here we present two planar examples.

Bálint et al. [1] (see also [10], p. 27.) investigated the following question. A set of points in the plane is said to be *midpoint-free* if it has no pair of elements whose midpoint also belongs to the set. Let $m(n)$

*Supported by NSF grant CCR-00-98246, PSC-CUNY Research Award 64421-0033 and OTKA-T-032452.

denote the largest number m such that every set of n points in the plane has a midpoint-free subset of size m. It was proved in [1] that

$$\lceil \frac{-1+\sqrt{8n+1}}{2} \rceil \le m(n),$$

and it was conjectured that the order of magnitude of this bound cannot be improved, i.e., we have $m(n) = O(\sqrt{n})$. However, it follows from the existence of relatively dense sets of integers containing no 3-term arithmetic progression that this conjecture is wrong.

Theorem 1. *There are positive constants c, c' such that*

$$n^{1-c/\sqrt{\log n}} \le m(n) \le n/\log^{c'} n.$$

F. Hurtado raised the following problem. For any point set P, let $M(P)$ denote the set of midpoints of all the $\binom{n}{2}$ segments spanned by point pairs in P. Determine $\mu(n) = \min_{|P|=n} |M(P)|$, where the minimum is taken over all sets of n points in the plane, no 3 of which are collinear.

Hurtado and Urrutia showed that $\mu(n) = O(n^{\log_2 3}) \approx O(n^{1.585})$, but no superlinear lower bound was known. Using an idea of Behrend and Freiman's theory of set addition, we prove

Theorem 2. *There is a positive constant c such that*

$$\mu(n) \le ne^{c\sqrt{\log n}}.$$

Furthermore, we have $\lim_{n\to\infty} \mu(n)/n = \infty$.

In the next two sections, we establish Theorems 1 and 2, resp., while in the last section some related questions are discussed.

2 Proof of Theorem 1

Consider a set P of n points in the plane with no midpoint-free subset of size larger than $m(n)$. First, choose (e.g., randomly) a straight line ℓ so that the orthogonal projection $\phi : P \to \ell$ takes P into an n-element set P' satisfying the following condition: for any $p_i, p_j, p_k \in P$, the midpoint of the segment p_ip_k is p_j if and only if $\phi(p_i), \phi(p_j)$, and $\phi(p_k)$ (in this order) form an arithmetic progression of length 3. Using simultaneous approximation [8], for any positive integer q, we can replace each point $\phi(p_i)$ by a rational number r_i/q, such that $r_i = r_i(q)$ is an integer and

$$|\phi(p_i) - \frac{r_i}{q}| \leq \frac{1}{q^{1+1/n}}$$

holds for all $1 \leq i \leq n$.

There exists a sufficiently large q satisfying the following condition: each triple $(\phi(p_i), \phi(p_j), \phi(p_k))$ forms an arithmetic progression (in this order) if and only if (r_i, r_j, r_k) does. Indeed, we have

$$|(\phi(p_i) + \phi(p_k) - 2\phi(p_j))q - (r_i + r_k - 2r_j)| \leq$$

$$|q\phi(p_i) - r_i| + |q\phi(p_k) - r_k| + 2|q\phi(p_j) - r_j| \leq \frac{4}{q^{1/n}}.$$

Assuming that $q > 4^n$, if $\phi(p_i)+\phi(p_k)-2\phi(p_j) = 0$ holds for some triple, we obtain that $|r_i+r_k-2r_j| < 1$ so that $r_i+r_k-2r_j = 0$ must also be true. In the reverse direction, assume indirectly that $\phi(p_i) + \phi(p_k) - 2\phi(p_j)$ is not equal to zero, but $r_i(q) + r_k(q) - 2r_j(q) = 0$ holds for infinitely many values of q. For these values, we have

$$|\phi(p_i) + \phi(p_k) - 2\phi(p_j)| \leq \frac{4}{q^{1+1/n}},$$

which leads to a contradiction, as q tends to infinity.

Thus, we have reduced the problem to the following: determine the largest positive integer $m_3'(n)$ such that every set of n integers has a subset of size $m_3'(n)$ which contains no arithmetic progression of length 3.

Let $m_3(n)$ denote the largest number of elements that can be chosen from the first n positive integers without containing a 3-term arithmetic progression. Clearly, we have $m_3'(n) \leq m_3(n)$ for every n. It was proved by Komlós, Sulyok, and Szemerédi [11] in a more general setting that there exists a constant $c > 0$ such that $m_3'(n) \geq cm_3(n)$. Thus, Theorem 2 immediately follows from well known estimates on $m_3(n)$, due to Behrend [2], Heath-Brown [9], and Szemerédi [14].

Note that the same argument can be applied in higher dimensions.

3 Proof of Theorem 2

First we establish the upper bound, by adapting the arguments in [5]. Assume, for the sake of simplicity, that $n = \lfloor \frac{2^{d(d-2)}}{d} \rfloor$ for some natural number $d \geq 4$. Consider the set L of all lattice points $(x_1, \ldots, x_d) \in \mathbf{R^d}$ with integer coordinates $0 \leq x_i < 2^d$. The number of distinct distances determined by L is at most $d(2^d)^2$, because there are at most that many numbers of the form $(\sum_{i=1}^d (x_i - x_i')^2)^{1/2}$, where $0 \leq x_i, x_i' < 2^d$. In particular, there is a sphere around the origin which contains at least

$$\frac{|L|}{d(2^d)^2} = \frac{(2^d)^d}{d(2^d)^2} \geq \lfloor \frac{2^{d(d-2)}}{d} \rfloor = n$$

elements of L. Let P denote the set of these points.

Let $M(P)$ denote the set of midpoints of all segments determined by P. Clearly, we have $|M(P)| = |P+P|$, where $P+P = \{\, p_1+p_2 \mid p_1, p_2 \in P \,\}$. Observe that every element of $P+P$ is a vector $(x_1, \ldots, x_d) \in \mathbf{R^d}$ with integer coordinates $0 \leq x_i < 2^{d+1}$, hence

$$|M(P)| = |P+P| \leq (2^{d+1})^d < n2^{8\sqrt{\log n}}\,.$$

Fix a 2-dimensional plane Π in $\mathbf{R^d}$, and for any $p \in P$ let p' denote the orthogonal projection of p into Π. Evidently, we can choose Π so as to meet the following two conditions: (i) the projections of no two elements of P coincide, (ii) no 3 elements of P' are collinear. In view of the fact that $p_1 + p_2 = p_3 + p_4$ implies $|p_1' + p_2'| = |p_3' + p_4'|$, we have that the number of distinct midpoints of all segments determined by P' satisfies

$$|M(P')| = |P' + P'| \leq |P+P| < n2^{8\sqrt{\log n}}\,,$$

as required. This argument easily extends to the general case when n can take any positive integer value.

We prove the second part of Theorem 2 by contradiction. Assume that for infinitely many values of n there are n-element point sets P_n with no 3 collinear points in the plane such that the the number of midpoints of all segments spanned by P_n satisfies $|M(P_n)| = |P_n + P_n| < Cn$, for an absolute constant C.

We need the following well known result of Freiman [6]: For any integer C, there exists C' with the property that any n-element set P_n

in the plane with $|P_n + P_n| < Cn$ can be covered by the projection of a lattice of dimension C and size $C'n$. That is,

$$P_n \subseteq \{v_0 + m_1 v_1 + \cdots + m_C v_C \mid 1 \le m_i \le n_i\},$$

for suitable vectors $v_i \in \mathbf{R}^2$ and natural numbers n_i satisfying $\prod_{i=1}^{C} n_i \le C'n$. (See Ruzsa [13] for a simple proof.)

Without loss of generality assume that $n_1 \ge n^{1/C}$. Obviously, we can fix some values $\bar{m}_2, \ldots, \bar{m}_C$ so that

$$v_0 + m_1 v_1 + \bar{m}_2 v_2 + \cdots + \bar{m}_C v_C \in P_n$$

for at least

$$\frac{n}{n_2 n_3 \cdots n_C} \ge \frac{n_1}{C'} \ge \frac{n^{1/C}}{C'}$$

different integers m_1. However, the corresponding points of P_n are all on a line, contradicting our assumption.

4 Related problems

4.1. It was noticed by Cockayne and Hedetniemi [3] that the problem of placing queens on the diagonal of an $n \times n$ chessboard so as to cover all squares is equivalent to the problem of finding a midpoint-free set of integers up to $n/2$, i.e., one containing no 3-term arithmetic progression.

4.2. Erdős raised the following problem related to Theorem 1. Determine the largest integer $\alpha(n)$ such that every set of n points in the plane, no four on a line, has an $\alpha(n)$-element subset with no collinear triples. The best known bounds, due to Füredi [7], leave plenty of room for improvement:

$$\Omega(\sqrt{n \log n}) \le \alpha(n) \le o(n).$$

4.3. Erdős, Fishburn, and Füredi [4] studied the following question, strongly related to Theorem 2. Given a set P of n points in *convex position* in the plane, let $M(P)$ denote the set of midpoints of its $\binom{n}{2}$ sides and diagonals. How small can the cardinality $\mu_c(n)$ of M be for fixed n? One might guess that the answer is $(0.5 - o(1))n^2$. However, it was shown in [4] that this minimum is somewhere between $0.40\, n^2$ and $0.45\, n^2$. In fact, we have

$$\binom{n}{2} - \lfloor \frac{n(n+1)(1-e^{-1/2})}{4} \rfloor \leq \mu_c(n) \leq \binom{n}{2} - \lfloor \frac{n^2-2n+12}{20} \rfloor,$$

for all $n \geq 3$. The upper bound follows from the fact that the number of multiple midpoints can be as large as $\lfloor (n^2-2n+12)/20 \rfloor$.

Woodall [15] solved a similar problem of R. Hall, concerning the minimum number of midpoints induced by an n-element subset of the vertex set of a d-dimensional cube ($n \leq 2^d$).

References

[1] V. Bálint, M. Branická, P. Grešák, I. Hrinko, P. Novotný, and M. Stacho, Several remarks about midpoint-free subsets, *Studies of University Transport and Communication in Žilina, Math.-Phys. Series* **10**, 3-10.

[2] F. A. Behrend: On sets of integers which contain no three terms in arithmetical progression, *Proc. National Academy of Sciences, USA* **28** (1942), 561–563.

[3] E. J. Cockayne and S. T. Hedetniemi, On the diagonal queens domination problem. *J. Combin. Theory, Ser. A* **42** (1986), 137–139.

[4] P. Erdős, P. Fishburn, and Z. Füredi, Midpoints of diagonals of convex n-gons, *SIAM J. Discrete Math.* **4** (1991), 329–341.

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

[6] G. Freiman, *Foundations of a Structural Theory of Set Addition*, *Translations of Mathematical Monographs* **37**, American Mathematical Society, Providence, 1973.

[7] Z. Füredi, Maximal independent subsets in Steiner systems and in planar sets, *SIAM J. Discrete Mathematics* **4** (1991), 196–199.

[8] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers, 4th ed.*, Clarendon Press, Oxford, 1960.

[9] D. R. Heath-Brown, Integer sets containing no arithmetic progressions, *J. London Math. Soc.* (2) **35** (1987), 385–394.

[10] V. Klee and S. Wagon, *Old and New Unsolved Problems in Plane Geometry and Number Theory, Dolciani Mathematical Expositions* **11**, Mathematical Association of America, 1991.

[11] J. Komlós, M. Sulyok, and E. Szemerédi, Linear problems in combinatorial number theory, *Acta Mathematica Academiae Scientiarum Hungaricae* **26** (1975), 113–121.

[12] J. Pach and P. K. Agarwal, *Combinatorial Geometry,* Wiley-Interscience, New York, 1995.

[13] I. Z. Ruzsa, Arithmetical progressions and the number of sums, *Periodica Mathematica Hungarica* **25** (1992), 105–111.

[14] E. Szemerédi, Integer sets containing no arithmetic progression, *Acta Math. Hungar.* **56** (1990), 155–158.

[15] D. R. Woodall, A theorem on cubes, *Mathematika* **24** (1977), 60–62.