SIAM J. DISC. MATH.
Vol. 4, No. 3, pp. 329–341, August 1991

MIDPOINTS OF DIAGONALS OF CONVEX n-GONS*

PAUL ERDÖS†, PETER FISHBURN‡, AND ZOLTAN FÜREDI†

Abstract. Let $f(n)$ be the minimum over all convex planar n-gons of the number of different midpoints of the $\binom{n}{2}$ line segments, or diagonals, between distinct vertices. It is proved that $f(n)$ is between approximately $0.8\binom{n}{2}$ and $0.9\binom{n}{2}$. The upper bound uses the fact that the number of multiple midpoints, shared by two or more diagonals, can be as great as about $\binom{n}{2}/10$. Cases for which the number of midpoints is at least $\lceil n(n-2)/2 \rceil + 1$, the number for a regular n-gon when n is even, are noted.

Key words. convex n-gons, diagonal midpoints, multiple midpoints

AMS(MOS) subject classifications. 52A10, 52A25, 51M20

1. Introduction. Let M denote the set of midpoints of the $\binom{n}{2}$ line segments between distinct vertices of a convex n-gon in the plane. Let $f(n) = \min |M|$, taken over all convex n-gons. We prove that $f(n)$ is between about $0.40n^2$ and $0.45n^2$.

THEOREM 1. *For all* $n \geqq 3$,

$$\binom{n}{2} - \left\lfloor \frac{n(n+1)(1-e^{-1/2})}{4} \right\rfloor \leqq f(n) \leqq \binom{n}{2} - \left\lfloor \frac{n^2-2n+12}{20} \right\rfloor.$$

The lower bound proof, in § 2, is based in part on the following lemma.

PARALLELOGRAM LEMMA (Euclid). *Two finite crossing line segments in the plane have the same midpoint if and only if the ends of the segments are the vertices of a parallelogram.*

Section 2 also uses the notion of a multiple midpoint. Call a point in M *multiple* if it is the midpoint of two or more of the $\binom{n}{2}$ line segments between vertices of the convex n-gon. We let $\mathbf{M}$ denote the set of multiple midpoints.

Let $g(n) = \max |\mathbf{M}|$, taken over all convex n-gons. Clearly $f(n) + g(n) \leqq \binom{n}{2}$. The upper bound on $f(n)$ in Theorem 1 is a corollary of the following theorem.

THEOREM 2. *For all* $n \geqq 3$,

$$g(n) \geqq \left\lfloor \frac{n^2-2n+12}{20} \right\rfloor.$$

This quadratic lower bound on $g(n)$ is the largest lower bound presently known for $n \geqq 18$, but for most $n \leqq 17$ it is exceeded as follows:

n	5	6	7	8	9	10	11	12	13	14	15	16	17
$\left\lfloor \frac{n^2-2n+12}{20} \right\rfloor$	1	1	2	3	3	4	5	6	7	9	10	11	13
$g(n) \geqq$	1	2	3	3	4	5	6	8	9	10	11	13	14.

The construction for the improved lower bound on $g(n)$ is described in § 4.

* Received by the editors April 16, 1990; accepted for publication August 10, 1990. This research was supported by DIMACS (Center for Discrete Mathematics and Theoretical Computer Science), a National Science Foundation Science and Technology Center grant NSF-STC88-09648, and by Hungarian National Science Foundation grant 1817.

† Mathematical Institute, Hungarian Academy of Sciences, Budapest, Hungary.

‡ AT&T Bell Laboratories, Murray Hill, New Jersey 07974.

Section 5 concludes our study of M with remarks on $|M|$ when the number of multiple midpoints is small. Its main result, which includes all regular n-gons for even n, is the following theorem.

THEOREM 3. *If the number of multiple midpoints is less than* 3, *or if one vertex of the convex n-gon is an endpoint of diagonals whose midpoints include all multiple midpoints, then*

$$|M| \geqq \left\lceil \frac{n(n-2)}{2} \right\rceil + 1.$$

This inequality can fail when $|\mathbf{M}| = 3$.

We are not aware of previous contributions to the problems investigated here. Some time ago, Behrend [1] looked at sets of integers that contain no element midway between two others. More recently, Freiman [4], [5] obtained many results involving midpoints in additive number theory. One of these says that if $2 \leqq \lambda < 2^m$, $m \geqq 2$, then there is a constant $c_\lambda > 0$ such that every sufficiently large finite $X \subseteq \mathbb{R}^m$ whose points determine no more than $\lambda |X|$ midpoints has at least $c_\lambda |X|$ of its points in some hyperplane in $\mathbb{R}^m$. Fishburn [2], [3] gives an elementary proof of the planar version of Freiman's result and finds nearly best values of c_λ for $2 \leqq \lambda < 4$. The latter work uses results in the present paper.

2. Lower bounds on *f*(*n*). Let $f(V) = |M|$ for a convex n-gon with vertex set V and nonempty multiple midpoint set $\mathbf{M}$. For each $\mu \in \mathbf{M}$ let

$$V(\mu) = \{x \in V : \mu = (x+y)/2 \text{ for some } y \in V\},$$

$$E(\mu) = \{\text{all diagonals with midpoint } \mu\},$$

$$D(\mu) = \{\text{all diagonals for } V(\mu) \text{ except those in } E(\mu)\}.$$

Thus $V(\mu)$ is the vertex set of $E(\mu)$, $|V(\mu)| = 2|E(\mu)|$, and $|D(\mu)| + |E(\mu)| = \binom{|V(\mu)|}{2}$. Let $\mu^* = |E(\mu)|$. Then

$$|D(\mu)| = 2\mu^*(\mu^* - 1).$$

Clearly, $E(\mu) \cap E(\lambda) = \varnothing$ when $\mu \neq \lambda$, $\mu, \lambda \in \mathbf{M}$, and the same hypotheses and the Parallelogram Lemma are easily seen to imply $D(\mu) \cap D(\lambda) = \varnothing$. Obviously,

$$f(V) = \binom{n}{2} - \sum_{\mu \in \mathbf{M}} (\mu^* - 1).$$

We observe in passing that for $n \geqq 3$

$$f(V) \geqq \binom{n}{2} - \left\lfloor \frac{n^2 - 2n}{8} \right\rfloor$$

so that $f(n)$ is at least as great as about $(\frac{3}{8})n^2$. Observe that $\sum |D(\mu)| \leqq \binom{n}{2} - \lceil n/2 \rceil$ since for every $x \in V$ there is a $y \in V \setminus \{x\}$ across the n-gon from x such that $[x, y]$ is not a side of a parallelogram on four vertices of V. Therefore

$$\sum_{\mu \in \mathbf{M}} (\mu^* - 1) = \sum \frac{|D(\mu)|}{2\mu^*} \leqq \frac{1}{4} \sum |D(\mu)| \leqq \frac{1}{4} \left\{ \binom{n}{2} - \lceil n/2 \rceil \right\},$$

and the given inequality for $f(V)$ follows from this and the concluding equality of the preceding paragraph.

The rest of this section is devoted to the better lower bound specified in the following lemma.

LEMMA 1.

$$f(V) \geqq \binom{n}{2} - \left\lfloor \frac{n(n+1)(1-e^{-1/2})}{4} \right\rfloor.$$

This shows that $f(n) > 0.4016n^2$ for all large n.

A new definition is needed. Let

$$C(x) = \{[y,z] \in E(\mu) : x \in V(\mu) \quad \text{and} \quad x \notin \{y,z\}\}$$

for each $x \in V$: see Fig. 1(a). The following result is central.

LEMMA 2. *Every two diagonals in $C(x)$ intersect in the interior of the n-gon.*

Proof. Suppose otherwise for $[y, z]$, $[a, b] \in C(x)$. Let μ be the midpoint of $[y, z]$ and of $[x, w]$, let α be the midpoint of $[a, b]$ and $[x, c]$, and suppose with no loss of generality that α lies in the x direction from $[y, z]$. Then a and b must lie in the three-sided dashed regions shown in Fig. 1(b), one in each region, or else convexity will be violated.

Assume that a is in the upper dashed region and b is in the lower dashed region. Suppose $a = y$: see Fig. 1(c). Then $b \neq z$ by our initial supposition, and since $[x, y]$, $[b, c]$, and $[z, w]$ are mutually parallel by the Parallelogram Lemma, we violate convexity. Therefore $a \neq y$. Similarly, $b \neq z$.

It follows that a and b are interior to their regions. Position a accordingly, anywhere in its region: see Fig. 1(d). Then convexity forces b to be interior to the shaded triangular

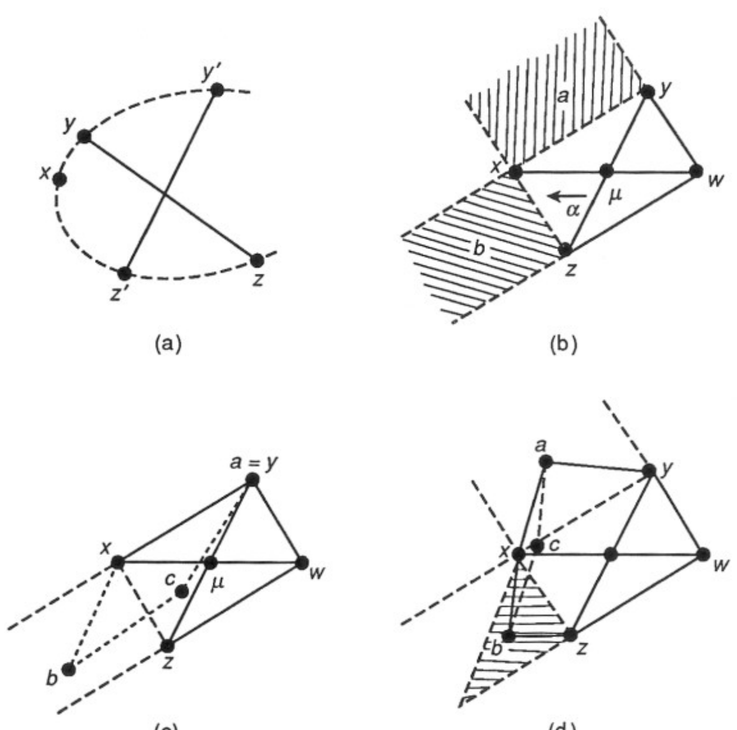


FIG. 1

region. But c, the fourth vertex of the parallelogram for α, will then lie in the interior of the hexagon with vertices $axbzwy$, which gives another violation of convexity. □

For each $v \in V$ let

$$c_v = |C(v)| = \sum_{\{\mu : v \in V(\mu)\}} (\mu^* - 1),$$

and for each diagonal $[x, y]$ of the n-gon define its "length" by

$$l(x,y) = 1 + \min\{\text{number of } V \text{ points on one side of } xy \text{ line},$$
$$\text{number of } V \text{ points on the other side of } xy \text{ line}\},$$

so that $1 \leqq l(x, y) \leqq \lfloor n/2 \rfloor$. If n is odd, there are n diagonals for each $l \in \{1, \cdots, (n-1)/2\}$; if n is even, there are n diagonals for each $l \in \{1, \cdots, (n-2)/2\}$ and $n/2$ diagonals with $l = n/2$. The following connection between c and l is immediate from Lemma 2.

COROLLARY 1. *$c_v \leqq l(x, y)$ for all $v \in V$ and all $[x, y] \in C(v)$.*

We now construct an $\binom{n}{2} \times n$ 0-1 matrix $A(V)$ that will be manipulated to yield the conclusion of Lemma 1. The $\binom{n}{2}$ rows of $A(V)$ are labeled by the diagonals in nonincreasing order of their l values: the final n rows have $l = 1$. The n columns of $A = A(V)$ are labeled by the vertices in nonincreasing order of their c_v. Write $i \to [x, y]$ when row i has label $[x, y]$, and $j \to v$ when column j has label v. We define A's entries by the following: when $i \to [x, y]$ and $j \to v$,

$$A_{ij} = \begin{cases} 1 & \text{if } [x,y] \in C(v), \\ 0 & \text{otherwise.} \end{cases}$$

When $j \to v$, $c_v = \sum_i A_{ij}$ and $\sum c_v = \sum_{\mathbf{M}} 2\mu^*(\mu^* - 1)$. Let $r_i = \sum_j A_{ij}$ for row i. When $i \to [x, y]$, $r_i = 0$ if $[x, y] \notin \cup_{\mathbf{M}} E(\mu)$, but if $[x, y] \in E(\mu)$ then

$$r_i = |\{v : [x,y] \in C(v)\}| = 2(\mu^* - 1).$$

Since μ^* rows have labels in $E(\mu)$,

$$\sum_{i=1}^{\binom{n}{2}} \frac{r_i}{r_i + 2} = \sum_{\mathbf{M}} \left[\frac{2(\mu^* - 1)}{2\mu^*}\right]\mu^* = \sum_{\mathbf{M}} (\mu^* - 1).$$

Therefore

$$f(V) = \binom{n}{2} - \sum_i \frac{r_i}{r_i + 2}.$$

Our lower bound on $f(V)$ is obtained from an upper bound on $\sum r_i/(r_i + 2)$.

Assume until later that n is odd. Then, by Corollary 1 and the nonincreasing order of rows by l,

$$\text{if } A_{ij} = 1 \quad \text{then } c_j \leqq \frac{n-1}{2} - \left\lfloor \frac{i-1}{n} \right\rfloor,$$

where c_j is c_v when $j \to v$.

Let $\mathscr{A}$ be the set of all $\binom{n}{2} \times n$ nonnegative integer matrices with column sums $c_1 \geqq c_2 \geqq \cdots \geqq c_n$, row sums $r_1, r_2, \cdots$, and

$$c_j \leqq \frac{n-1}{2} - \left\lfloor \frac{i-1}{n} \right\rfloor \quad \text{whenever entry } (i, j) \neq 0.$$

Clearly $A(V) \in \mathscr{A}$. Suppose $B \in \mathscr{A}$, $i < a$, $j < b$, and $B_{ib}B_{aj} > 0$. Let B' equal B except on $\{i, a\} \times \{j, b\}$, where

$$B'_{ij} = B_{ij} + 1,$$

$$B'_{ib} = B_{ib} - 1,$$

$$B'_{aj} = B_{aj} - 1,$$

$$B'_{ab} = B_{ab} + 1.$$

Then $B' \in \mathscr{A}$ since we have changed neither the column nor row sums, and in going to B' we need

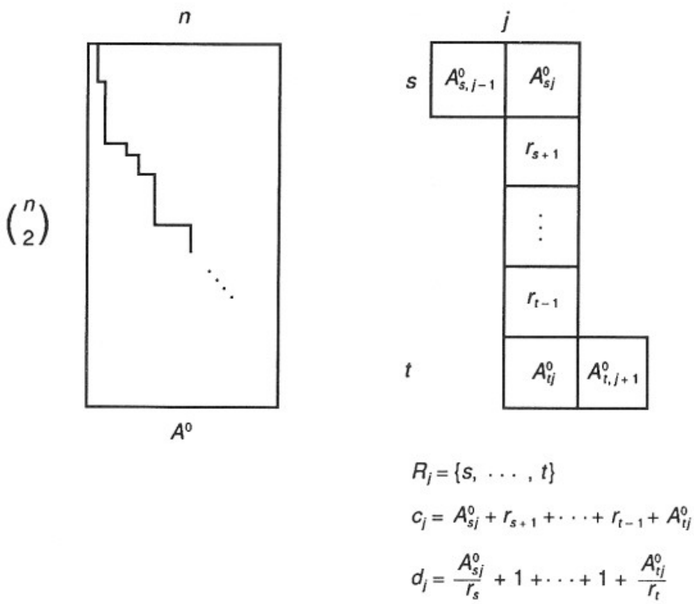
$$c_j \leqq \frac{n-1}{2} - \left\lfloor \frac{i-1}{n} \right\rfloor \qquad \text{and} \qquad c_b \leqq \frac{n-1}{2} - \left\lfloor \frac{a-1}{n} \right\rfloor.$$

The first of these is true since $i < a$ and for B we had $c_j \leqq (n-1)/2 - \lfloor (a-1)/n \rfloor$. The second is true since $j < b \Rightarrow c_b \leqq c_j$.

It follows from a finite sequence of switches as just described that $A(V)$ can be transformed into $A^0 \in \mathscr{A}$ so that no positive entry of A^0 is northeast or southwest of another positive entry. This implies that all positive entries of A^0 lie on a rectilinear staircase path as shown in Fig. 2(a). We suppose for convenience that all entries of A^0 on the path are positive: this is not needed for the desired conclusion, but it simplifies calculations by avoiding special notation that would continually refer to the set of all rows for which $r_i > 0$.

Let $W = \sum r_i/(r_i + 2)$ and let R_j be the set of rows i for which $A^0_{ij} > 0$. The staircase pattern gives $R_1 \leqq R_2 \leqq \cdots \leqq R_n$. Also let

$$W_j = \sum_{i \in R_j} \frac{A^0_{ij}}{r_i + 2} \qquad j = 1, \cdots, n.$$



(a) (b)

FIG. 2

Then

$$W=\sum_{i}\frac{r_i}{r_i+2}=\sum_{i}\left(\sum_{j=1}^{n}\frac{A_{ij}^{0}}{r_i}\right)\frac{r_i}{r_i+2}$$

$$=\sum_{j=1}^{n}\sum_{i\in R_j}\frac{A_{ij}^{0}}{r_i+2}=\sum_{j}W_j.$$

We assign a fractional number of rows, d_j, to column j in which $A_{ij}^{0}>0$, as follows:

$$d_j=\sum_{i\in R_j}A_{ij}^{0}/r_i.$$

If i is the first or last member of R_j, $0<A_{ij}^{0}/r_i\leqq 1$, and if i is between the first and last members of R_j then $A_{ij}^{0}/r_i=1$ since the r_i total for row i is all in column j: see Fig. 2(b).

Suppose $t=\max R_j$. Then

$$d_1+\cdots+d_j=t-1+\sum_{k\leqq j}A_{tk}^{0}/r_i\leqq t$$

and, by our earlier bound on c_j,

$$c_j\leqq\frac{n-1}{2}-\left\lfloor\frac{d_1+\cdots+d_j-1}{n}\right\rfloor.$$

LEMMA 3. *For all* j, $c_j/d_j\geqq 2$ *and*

$$W_j\leqq\frac{c_jd_j}{c_j+2d_j}.$$

Proof. For each paired piece of c_j and d_j as shown in Fig. 2(b), $A_{sj}^{0}\geqq(A_{sj}^{0}/r_s)2$, $r_{s+1}\geqq 1(2)$, $\cdots$, so summation gives $c_j\geqq 2d_j$. When $\|R_j\|=m$, the inequality

$$W_j\leqq d_j\frac{c_j/d_j}{c_j/d_j+2}$$

can be put in the form

$$\sum_{k=1}^{m}p_k(\bar{r}-r_k)/(r_k+2)\geqq 0,$$

where $p_k>0$, $\sum p_k=1$ and $\bar{r}=\sum p_kr_k$. When multiplied by the product of the (r_k+2), this inequality becomes

$$\sum_{i<j}p_ip_j\left[\prod_{k\notin\{i,j\}}(r_k+2)\right](r_i-r_j)^2\geqq 0,$$

which is true. □

By Lemma 3, $\sum_{\mathbf{M}}(\mu^*-1)=W\leqq\sum_j c_jd_j/(c_j+2d_j)$. Denote the latter sum by $F(c,d)$, $c=(c_1,\cdots,c_n)$ and $d=(d_1,\cdots,d_n)$, and consider the following problem:

$$\text{maximize } F(c,d)=\sum_{j=1}^{n}\frac{c_jd_j}{c_j+2d_j}$$

subject to $c_1 \geqq c_2 \geqq \cdots \geqq c_n$ and, for $j = 1, \cdots, n$,

$$d_j > 0, \quad c_j/d_j \geqq 2, \quad c_j \leqq \frac{n-1}{2} - \left\lfloor \frac{d_1 + \cdots + d_j - 1}{n} \right\rfloor.$$

We replace the final constraint by the weaker but smooth $c_j \leqq (n+1)/2 - (d_1 + \cdots + d_j)/n$, observe that F increases in each c_j, and therefore take c_j as large as possible:

$$c_j^* = \frac{n+1}{2} - \frac{d_1 + \cdots + d_j}{n}.$$

Thus max $F(c, d) \leqq$ max $F(c^*, d)$ subject to $d_j > 0$ and $c_j^* \geqq 2d_j$.

LEMMA 4.

$$\max F(c^*, d) \leqq \frac{n(n+1)}{4}\left(1 - \frac{1}{\sqrt{e}}\right).$$

Proof. Let $c_0 = (n+1)/2$ and omit * on c_j. Also let $x_j = c_j/d_j \geqq 2$. By the definition of c_j^* we have $c_j = [nx_j/(1 + nx_j)]c_{j-1}$. Therefore

$$c_j = c_0 \prod_{i=1}^{j} \frac{nx_i}{1+nx_i},$$

$$d_j = c_0 \left(\prod_{i=1}^{j-1} \frac{nx_i}{1+nx_i} \right) \frac{n}{1+nx_j},$$

and

$$F(c,d) = c_0 \sum_{j=1}^{n} \left(\prod_{i=1}^{j} \frac{nx_i}{1+nx_i} \right) \frac{1}{2+x_j}$$

with each term in the sum $\leqq \frac{1}{4}$ since $x_j \geqq 2$. For $a, b > 0$

$$\left(\frac{na}{1+na}\right)\frac{1}{2+a} + \left(\frac{na}{1+na} \cdot \frac{nb}{1+nb}\right)\frac{1}{2+b} \geqq \left(\frac{nb}{1+nb}\right)\frac{1}{2+b} + \left(\frac{nb}{1+nb} \cdot \frac{na}{1+na}\right)\frac{1}{2+a}$$

if and only if $a \geqq b$. It follows that F is maximized when $x_1 \geqq x_2 \geqq \cdots \geqq x_n \geqq 2$, so assume the following.

Fix x_2 through x_n. Let $x = x_1$. Then

$$\frac{1}{c_0} F(c,d) = \left(\frac{nx}{1+nx}\right)\frac{1}{2+x} + \left(\frac{nx}{1+nx}\right) S,$$

where $S \leqq (n-1)/4$. Differentiation shows that the right-hand side decreases when

$$2 - nx^2 + S(2+x)^2 < 0,$$

which is true when $x \geqq 2 + 1/n$. We may therefore suppose that $x < 2 + 1/n$. But then S is much smaller than $n/4$, and the preceding inequality holds for all $x \geqq 2$. This implies that F is maximized at $x_1 = 2$, hence at $x_j = 2$ for all j, where

$$F = \frac{c_0}{4}\left[\frac{2n}{2n+1} + \left(\frac{2n}{2n+1}\right)^2 + \cdots + \left(\frac{2n}{2n+1}\right)^n\right]$$

$$= \frac{n(n+1)}{4}\left[1 - \left(\frac{2n}{2n+1}\right)^n\right] < \frac{n(n+1)}{4}\left(1 - \frac{1}{\sqrt{e}}\right). \qquad \square$$

Lemma 4 completes our proof of Lemma 1 when n is odd. When n is even, the preceding analysis is modified by replacing the bound on c_j obtained from Corollary 1 by

$$c_j \leqq \frac{n}{2} - \left\lfloor \frac{i+n/2-1}{n} \right\rfloor,$$

which corresponds to the remark on l for n even that precedes Corollary 1. Then c_j^* preceding Lemma 4 can be replaced by

$$c_j^* = \frac{n+1}{2} + \frac{1}{n} - \frac{d_1 + \cdots + d_j}{n}.$$

The only effect this has on the proof of Lemma 4 is to change c_0 there to $c_0 = (n+1)/2 + 1/n$. This changes the final equation in that proof to

$$F = \left[\frac{n(n+1)}{4} + \frac{1}{2}\right]\left[1 - \left(\frac{2n}{2n+1}\right)^n\right].$$

It is easily checked that this is less than $n(n+1)(1-e^{-1/2})/4$ when $n \geqq 10$. Therefore Lemma 4 holds for all $n \geqq 3$ except for $n \in \{4, 6, 8\}$. Lemma 1 claims for these three that $f(n=4) \geqq 5$, $f(n=6) \geqq 11$, and $f(n=8) \geqq 21$. Since $f(4) = 5$, $f(6) = 13$, and $f(8) \in \{24, 25\}$, Lemma 1 holds for all $n \geqq 3$.

3. Lower bound on $g(n)$.

THEOREM 2.

$$g(n) \geqq \left\lfloor \frac{(n^2 - 2n + 12)}{20} \right\rfloor \quad \textit{for } n \geqq 3.$$

Proof. Let $a_k = k^2$ for $k = 4, 5, \cdots, 3m-1$ and $b_k = 3k^2$ for $k = 1, 2, \cdots, m$ with $m \geqq 2$. Take $N > 12m^2$ and construct the convex $(4m-4)$-gon that has m lower-left vertices $(-k, b_k)$ for $k = 1, \cdots, m$, and $3m-4$ upper-right vertices $(k, N - a_k)$ for $k = 4, \cdots, 3m-1$. For every $1 \leqq i < j \leqq m$ it is easily checked that

$$(*) \qquad b_j - b_i = a_{i+2j} - a_{2i+j}.$$

Since $j - i = (i+2j) - (2i+j)$, it follows that $[(-j, b_j), (i+2j, N - a_{i+2j})]$ and $[(-i, b_i), (2i+j, N - a_{2i+j})]$ have the common midpoint

$$c_{ij} = ((i+j)/2, (N - i^2 - j^2 - 4ij)/2).$$

Moreover, if $i \neq k$, $i < j$, $k < l$, and $i + j = k + l$, then the vertical components of c_{ij} and c_{kl} are distinct. Therefore every multiple midpoint c_{ij} is distinct, so

$$g(4m-4) \geqq \binom{m}{2}.$$

The lower bound ratio of $g(n)/n^2$ in this case is approximately $(m^2/2)/(4m)^2 = 1/32$.

We get a larger ratio by deleting vertices at both ends of the upper-right part of the construction since the a_k pairs that match the b_k pairs as in $(*)$ are denser in the middle of the a_k sequence. A crude calculation for the quadratic terms shows that if we delete K vertices at each end of the a_k sequence then we lose about $K^2/3$ of the c_{ij}. This also removes $2K$ points from n, so the new ratio for $g(n)/n^2$ is approximately

$$\frac{m^2/2 - K^2/3}{(4m-2K)^2},$$

which is maximized through differentiation with respect to K at $K = 3m/4$, where the ratio is $1/20$.

To be more precise, suppose T vertices are removed from the a_k sequence, $\lfloor T/2 \rfloor$ at one end and $\lceil T/2 \rceil$ at the other end. Then, with details omitted, we get

$$g(4m-4-T) \geqq \binom{m}{2} - \left\lceil \frac{T}{6} \right\rceil \left(T+3-3\left\lceil \frac{T}{6} \right\rceil \right).$$

Given n, we then consider the (m, T) pairs that satisfy $n = 4m - 4 - T$ to determine the pair that maximizes the right-hand side of the preceding inequality. Further calculations show that the maximum is $\lfloor (n^2 - 2n + 12)/20 \rfloor$, as claimed in Theorem 2. □

4. Another construction for $g(n)$. The a_k and b_k of the preceding proof were chosen in an attempt to minimize n, given that each of the $\binom{m}{2}$ pairs from the lower left is matched by a pair from the upper right to yield a different multiple midpoint. We examined variations to this construction, but their lower bounds on $g(n)/n^2$ were smaller than $1/20$.

However, as mentioned earlier, a different construction gives larger lower bounds on $g(n)$ for most $n \leqq 17$. This other construction yields a lower bound on $g(n)/n$ of approximately $7/6$ for large n, as compared to $n/20$ for the quadratic construction used to prove Theorem 2, and is therefore much less powerful than the quadratic bound for large n.

Figure 3 illustrates the other construction that yields the largest lower bounds on $g(n)$ for small n that are presently known. Its 18 vertices are numbered in the order in which they enter the construction. We begin with the tall narrow rectangle for vertices

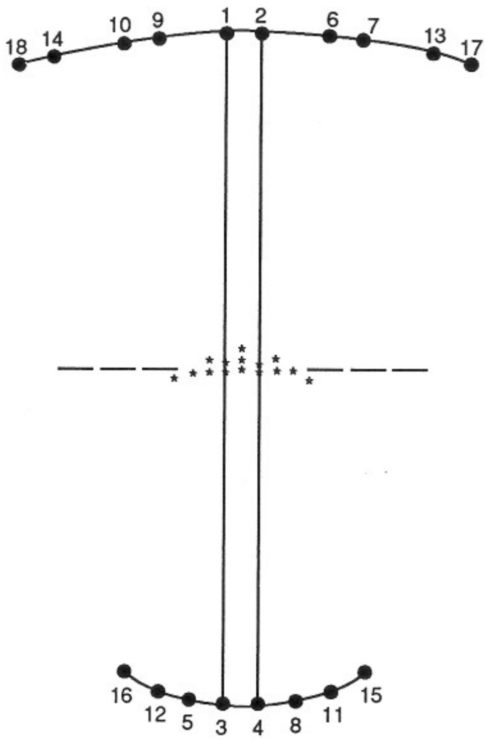


FIG. 3. *'s *denote multiple midpoints*.

1 through 4, then position 5 to the left of 3 so that the horizontal distances from 5 to 3 and from 3 to 4 are equal, with 5 slightly above the line through 3 and 4.

The other points are then positioned by midpoint restrictions and symmetry. Let $\mu(i,j)$ denote the midpoint between i and j. A complete account of multiple midpoints is shown in the following construction routine:

$\mu(1,4)=\mu(2,3)$
5: position as described above
6: $\mu(5,6)=\mu(2,4)$
7: $\mu(5,7)=\mu(3,6)$
8: position horizontally symmetric to 5
9: $\mu(8,9)=\mu(1,3)$
10: $\mu(8,10)=\mu(4,9)$
11: $\mu(10,11)=\mu(2,5)$
12: $\mu(7,12)=\mu(1,8)$ and $\mu(6,12)=\mu(9,11)$
13: $\mu(12,13)=\mu(3,7)$
14: $\mu(11,14)=\mu(4,10)$
15: $\mu(14,15)=\mu(2,12)$
16: $\mu(13,16)=\mu(1,11)$ and $\mu(7,16)=\mu(10,15)$
17: $\mu(12,17)=\mu(5,13)$
18: $\mu(11,18)=\mu(8,14)$.

Each μ equation here identifies a different multiple midpoint. The lower bound on $g(n)$ is the number of μ equations in place after point n is added.

The preceding construction shows that every vertex can be at the end of two or more diagonals whose midpoints are in $\mathbf{M}$. This occurs for the first time in the construction at $n=12$. Moreover, by Theorem 2, the smallest n presently known for which $g(n)>n$ is $n=23$. We do not know whether a smaller n suffices in either case.

5. Small numbers of multiple midpoints. This section focuses on situations that force $|M|$ to be much larger than the upper bound on $f(n)$ in Theorem 1. For each $\mu\in M$ let $V(\mu)$ be the set of all vertices at ends of diagonals that have midpoint μ. If $\mu\in\mathbf{M}$ then $|V(\mu)|\in\{4,6,8,\cdots\}$.

The following lemma is an easy consequence of the Parallelogram Lemma.

LEMMA 5. *If $\mu\in\mathbf{M}$ then all midpoints of line segments between points in $V(\mu)$ that differ from μ are different from each other. If $\lambda,\mu\in\mathbf{M}$, $\lambda\neq\mu$, and L is the line through λ and μ, then $L\cap V(\lambda)\cap V(\mu)=\varnothing$, $|V(\lambda)\cap V(\mu)|\leqq 2$ and, if $|V(\lambda)\cap V(\mu)|=2$, then λ or μ is the midpoint between the points in $V(\lambda)\cap V(\mu)$.*

Let R_n be a regular n-gon for even n. It follows immediately from the first part of Lemma 5 that

$$|M(R_n)|=\binom{n}{2}-\left(\frac{n}{2}-1\right)=\frac{n(n-2)}{2}+1.$$

An easy proof also shows that the maximum number of parallelograms that can be formed from the vertices of a convex n-gon for even n occurs at R_n and equals $n(n-2)/4$. The two diagonals of each parallelogram cross at the one multiple midpoint of R_n.

The initial observation in the preceding paragraph generalizes to the following theorem.

THEOREM 3. *If either $\|\mathbf{M}\| \leqq 2$ or $\bigcap_{\mathbf{M}} V(\mu) \neq \varnothing$ then*

$$\|M\| \geqq \left\lceil \frac{n(n-2)}{2} \right\rceil + 1.$$

The conclusion can fail if $\|\mathbf{M}\| = 3$.

Figure 4 verifies the final statement of the theorem. With $n = 11$ and $\mathbf{M} = \{\mu_1, \mu_2, \mu_3\}$, we have

$$V(\mu_1) = \{1, 2, 3, 6, 8, 9\},$$

$$V(\mu_2) = \{2, 4, 6, 11\},$$

$$V(\mu_3) = \{3, 4, 5, 7, 10, 11\}$$

so that $V(\mu_1) \cap V(\mu_2) = \{2, 6\}$, $V(\mu_2) \cap V(\mu_3) = \{4, 11\}$, and $V(\mu_1) \cap V(\mu_3) = \{3\}$. Starting with all diagonals in place, we must remove five (two for μ_1, one for μ_2, two for μ_3) to have no multiple midpoint, so $\|M\| = \binom{11}{2} - 5 = 50$. On the other hand, $\lceil (11)(9)/2 \rceil + 1 = 51$.

The conclusion of Theorem 3 for $\|\mathbf{M}\| \leqq 2$ follows from Lemma 5. For example, if $\mathbf{M} = \{\lambda, \mu\}$ and $\|V(\lambda) \cap V(\mu)\| = 2$, then with $2c_\lambda = \|V(\lambda)\|$ and $2c_\mu = \|V(\mu)\|$ we have $2(c_\lambda + c_\mu) \leqq n + 2$. The removal of $(c_\lambda - 1) + (c_\mu - 1)$ diagonals gives a configuration in which no two remaining diagonals have the same midpoint, so

$$\|M\| \geqq \binom{n}{2} - (c_\lambda + c_\mu - 2) \geqq \binom{n}{2} + 2 - \left\lfloor \frac{n+2}{2} \right\rfloor$$

$$= \binom{n}{2} + 1 - \left\lfloor \frac{n}{2} \right\rfloor = \left\lceil \frac{n(n-2)}{2} \right\rceil + 1.$$

The following lemma is needed for the $\cap V(\mu) \neq \varnothing$ part of Theorem 3.

LEMMA 6. *Suppose $\|\mathbf{M}\| = t \geqq 3$ and $\bigcap_{\mathbf{M}} V(\mu) \neq \varnothing$. Let α_k be the number of vertices in exactly k of the $V(\mu)$ for $\mu \in \mathbf{M}$. Then $\alpha_t = 1$, $\alpha_k = 0$ for $2 < k < t$, and $\alpha_2 \leqq t - 1$.*

Proof. Take $x \in \bigcap_{\mathbf{M}} V(\mu)$ and $\mathbf{M} = \{\mu_1, \mu_2, \cdots, \mu_t\}$ as shown in Fig. 5(a). If vertex $y \neq x$ is in at least two $V(\mu)$, say $V(\mu_i)$ and $V(\mu_j)$, then Lemma 5 requires $y \in \{v_i, v_j\}$. It follows that $\alpha_t = 1$ and $\alpha_k = 0$ for all $2 < k < t$.

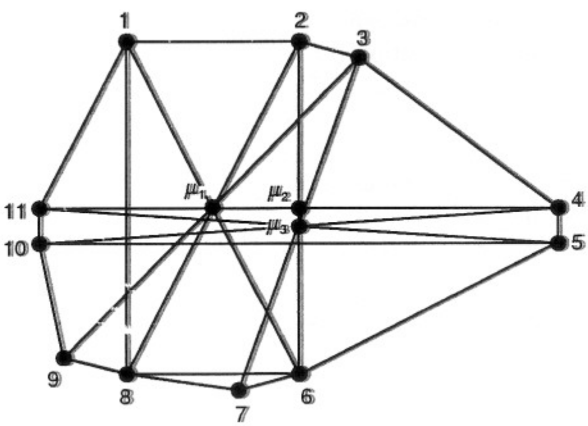


FIG. 4

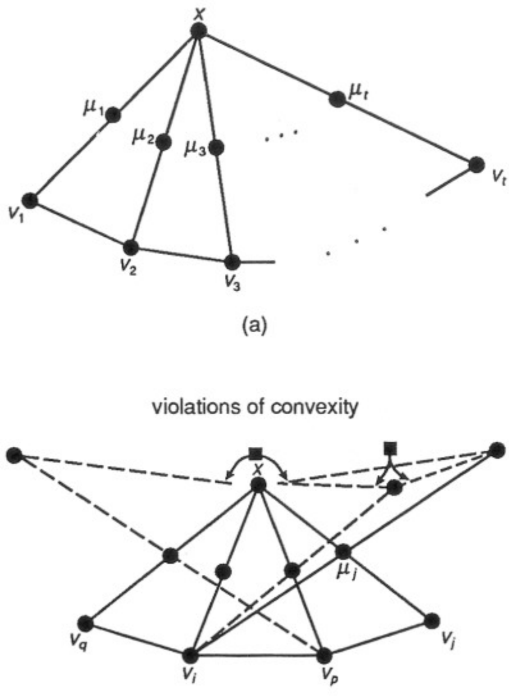


FIG. 5

Lemma 5 and convexity imply the following: see Fig. 5(b).

Rule 1. $v_i \in V(\mu_j)$ for at most one $j \neq i$.

Rule 2. $[v_i \in V(\mu_j),\, i < j] \Rightarrow v_p \notin V(\mu_q)$ if $q \leqq i < p \leqq j$.

Rule 1 implies $\alpha_2 \leqq t$, and it follows easily from Rule 2 and its dual for $j < i$ that $\alpha_2 \leqq 2$ when $t = 3$. We use induction on t in what follows.

Suppose $\alpha_2 \leqq t - 1$ for $t = 3, \cdots, r-1$ with $r \geqq 4$. Contrary to the lemma, suppose $\alpha_2 = r$ when $|\mathbf{M}| = r$. Suppose then that $v_i \in V(\mu_r)$ for some $i < r$. By Rule 2, $v_p \notin V(\mu_q)$ when $q \leqq i < p \leqq r$. Since $\alpha_2 = r$ and Rule 1 require every v_p to be in a $V(\mu_q)$, $q \neq p$, it follows that each of v_{i+1} through v_r is in one of $V(\mu_{i+1})$ through $V(\mu_r)$ that has a different index. But this is impossible by the induction hypothesis if $i < r - 2$, by Lemma 2 if $i = r - 2$, and by definition if $i = r - 1$. Hence no v_i for $i < r$ is in $V(\mu_r)$, a contradiction. Therefore $\alpha_2 \leqq r - 1$. $\square$

Proof Completion (Theorem 3). Let $|\mathbf{M}| = t \geqq 3$ with $\mathbf{M} = \{\mu_1, \cdots, \mu_t\}$ and $|V(\mu_k)| = 2c_k$. Suppose $\cap_{\mathbf{M}} V(\mu) \neq \varnothing$. Since $(\alpha_2, \cdots, \alpha_t) \leqq (t-1, 0, \cdots, 0, 1)$ by Lemma 6, $\sum (2c_k) \leqq n + 2(t-1)$. Excision of $\sum (c_k - 1)$ diagonals gives a configuration in which no remaining diagonals have the same midpoint. A calculation similar to that preceding Lemma 6 yields $|M| \geqq \lceil n(n-2)/2 \rceil + 1$. $\square$

6. Discussion. We have shown that $f(n)$ lies between about $0.8\binom{n}{2}$ and $0.9\binom{n}{2}$, that $|\mathbf{M}|$ can be as large as about $0.1\binom{n}{2}$, and that if either $|\mathbf{M}| \leqq 2$ or if some vertex lies at the ends of diagonals whose midpoints cover $\mathbf{M}$, then the corresponding n-gon has $|M| \geqq \lceil n(n-2)/2 \rceil + 1$.

Several open problems, in addition to exact values of $f(n)$ and $g(n)$, are suggested by our study. Does $\lim f(n)/n^2$ exist and, if so, what is its value? We ask a similar question for $g(n)/n^2$. Let $\mathbf{M}_3$ denote the set of all midpoints shared in common by at least three diagonals, and let $h(n) = \max |\mathbf{M}_3|$ over all convex n-gons. Does there exist $c > 0$ such that $h(n) > cn^2$ for all large n? If so, does a similar conclusion hold for midpoints with multiplicities that exceed 3?

REFERENCES

[1] F. A. BEHREND, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. USA, 32 (1946), pp. 331–333.

[2] P. C. FISHBURN, *On a contribution of Freiman to additive number theory*, J. Number Theory, 35 (1990), pp. 325–334.

[3] ———, *Sum set cardinalities of line restricted planar sets*, AT&T Bell Laboratories, Murray Hill, NJ, 1989.

[4] G. A. FREIMAN, *Foundations of a Structural Theory of Set Addition*, Vol. 37, Transl. Math. Monographs, American Mathematical Society, Providence, RI, 1973.

[5] ———, *What is the structure of K if K + K is small*? in Number Theory, New York, 1984–1985, Lecture Notes in Math., 1240, Springer-Verlag, New York, 1987, pp. 109–134.