The Maximum Size of a Convex Polygon in a Restricted Set of Points in the Plane*

N. Alon,[1] M. Katchalski,[2] and W. R. Pulleyblank[3]

[1] Department of Mathematics, Tel Aviv University, Tel Aviv, Israel

[2] Department of Mathematics, Technion-Israel Institute of Technology, Haifa, Israel

[3] Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Canada

Abstract. Assume we have k points in general position in the plane such that the ratio between the maximum distance of any pair of points to the minimum distance of any pair of points is at most $\alpha\sqrt{k}$, for some positive constant α. We show that there exist at least $\beta k^{1/4}$ of these points which are the vertices of a convex polygon, for some positive constant $\beta = \beta(\alpha)$. On the other hand, we show that for every fixed $\varepsilon > 0$, if $k > k(\varepsilon)$, then there is a set of k points in the plane for which the above ratio is at most $4\sqrt{k}$, which does not contain a convex polygon of more than $k^{1/3+\varepsilon}$ vertices.

1. Introduction

For any positive integer n, let $f(n)$ be the smallest integer such that from every set of $f(n)$ points in general position in the plane (i.e., no three are on a line), it is always possible to select n points which are the vertices of a convex n-gon. Erdös and Szekeres [3], [4] proved that $f(n)$ exists and showed that

$$2^{n-2} + 1 \leq f(n) \leq \binom{2n-4}{n-2} + 1.$$

In this paper we study the effect of restricting the geometry of the configurations. The configurations which establish the lower bound on $f(n)$ have some pairs of points very close together and others very far apart. We investigate the effect of placing an upper bound on the ratio of the maximum and minimum distances and show that, in this case, substantially stronger results are true.

Let $A = \{a_1, \ldots, a_k\}$ be a set of k points in general position in the plane, i.e., no three lie on a line. Let $d_{ij} = d(a_i, a_j)$ denote the Euclidean distance between a_i and a_j, We define

$$q(A) = \frac{\max\{d_{ij}: 1 \le i < j \le k\}}{\min\{d_{ij}: 1 \le i < j \le k\}}.$$

That is, $q(A)$ is the ratio of the diameter to the minimum distance between any pair of points. We can assume without loss of generality that

$$\min\{d_{ij}: 1 \le i < j \le k\} = 1.$$

In this case, $q(A)$ is just the diameter of A, and if we construct a disk D_i of radius $\frac{1}{2}$ centered at a_i for $1, 2, \ldots, k$, then the set $\{D_1, D_2, \ldots, D_k\}$ is a set of nonoverlapping unit disks. The total area of the disks is $k\pi/4$ so they cannot be contained in a circle of radius less than $\sqrt{k}/2$. Therefore $q(A) \ge c\sqrt{k}$ for some positive constant c. We consider the effect of requiring $q(A) \le \alpha\sqrt{k}$ for some constant α. That is, the diameter of A is bounded by a constant multiplied by the square root of the number of points. In this case we show that there exists a constant $\beta > 0$ such that A must contain a convex polygon containing $\beta k^{1/4}$ points. On the other hand, for every fixed $\varepsilon > 0$, if $k > k(\varepsilon)$, there exists a set A of k points which satisfies $q(A) \le 4\sqrt{k}$ but does not contain a convex polygon of more than $2k^{1/3+\varepsilon}$ vertices.

Let $c(A)$ denote the size of the largest subset of A which forms the vertex set of a convex polygon. We illustrate the effect on $c(A)$ of bounding $q(A)$ by considering the case $|A| = 4$. If $c(A) = 3$ then one point p_1 is contained in the convex hull of the other three, p_2, p_3, p_4. So one angle formed at p_1, say $\angle p_2 p_1 p_3$ is at least $2\pi/3$. Therefore, if $d_{ij} \ge 1$ for $1 \le i < j \le 4$, we must have $d_{23} \ge \sqrt{3}$. Therefore $q(A) \ge \sqrt{3}$. Consequently, if we require $q(A) < \sqrt{3}$ we can ensure that the points form a convex 4-gon. (Note that $\sqrt{3}$ is best possible, for if p_1 is the center of an equilateral triangle with vertices p_2, p_3, p_4, then $q(A) = \sqrt{3}$.)

The lower bound for $c(A)$ is proved by geometric arguments (in Section 2) and the upper bound is established by probabilistic arguments (in Section 3).

2. The Lower Bound

In this section we prove:

Theorem 2.1.   *For any positive constant α, there exists a positive constant $\beta = \beta(\alpha)$ such that if A is a set of k points in general position in the plane such that $q(A) \le \alpha\sqrt{k}$, then $c(A) \ge \beta k^{1/4}$.*

The proof relies on two lemmas.

Lemma 2.1. *Let C be a circle of radius $r > 4$ with centre O. Let Δ satisfy $2/\sqrt{r} < \Delta < 1$. Let $x_1, x_2, \ldots, x_m$ be points arranged in a cyclic order on C such that $\angle x_i O x_{i+1} \geq \Delta$ and $\angle x_m O x_1 \geq \Delta$. For $i = 1, 2, \ldots, m$, let D_i be a disk of radius $\frac{1}{2}$ centered at x_i and let $a_i \in D_i$. Then $a_1, a_2, \ldots, a_m$ are the vertices of a convex m-gon.*

Proof. The points $x_1, x_2, \ldots, x_m$ are the vertices of a convex m-gon. Since the distance between a_i and x_i is at most $\frac{1}{2}$, it is sufficient to prove that the distance between x_j and the line joining x_{j-1} and x_{j+1} is more than 1. Let h denote this distance. We have

$$h \geq r(1 - \cos \Delta) = 2r \sin^2 \frac{\Delta}{2} \geq 2r \left(\frac{\Delta}{2}\right)^2 \times \frac{\sin^2(\pi/3)}{(\pi/3)^2} \geq 2 \times 1 \times \frac{27}{4\pi^2} > 1$$

as required.   □

Lemma 2.2. *Let D_1 and D_2 be two nonoverlapping disks of radius $\frac{1}{2}$ with centers a_1, a_2 respectively, each of which intersects a fixed circle C of radius $r > 7$ centered at O. Suppose either both a_1 and a_2 lie outside or on C or else both lie inside or on C. Then $\theta = \angle a_1 O a_2 > 0.8/r$.*

Proof. Assume $r\theta \leq 0.8$ and that, say, a_1 and a_2 are both on or outside C. Consider an appropriate arc on C, of length 0.8, and the corresponding one on the circle concentric to C, of radius $r + \frac{1}{2}$. The annulus sector defined by these arcs contains a_1 and a_2, and thus has diameter ≥ 1, but this is a contradiction for large r, as the sector is then nearly a 0.8×0.5 rectangle. (Consult [1] for similar arguments.)   □

Proof of Theorem 2.1. Assume that the minimum distance between any pair of points of A is 1. Then $q(A)$ is the diameter of A. By Jung's theorem (see Chapter 16 of [7] or Problem 31 of [5]) the points of A are contained in a disk Q of radius at most $(\sqrt{3}/3)q(A) \leq (\sqrt{3}\alpha/3)\sqrt{k}$ centered at a point O. For each $a_i \in A$, let D_i be a disk with radius $\frac{1}{2}$ centered at a_i. Consider the concentric circles centered at O with radii $1, 2, \ldots, \lfloor(3\alpha/3)\sqrt{k}\rfloor$. At least one, say C of radius r, will intersect at least $(\sqrt{3}/\alpha)\sqrt{k}$ of the D_i. All of these disks lie completely within the annulus determined by the circles of radii $r - 1$ and $r + 1$, centered at O. The area of this annulus is $4\pi r$ and the area of the D_i in this annulus is at least $(\sqrt{3}\pi/(4\alpha))\sqrt{k}$. Therefore $r \geq (\sqrt{3}/(16\alpha))\sqrt{k}$.

At least one half, i.e., $(\sqrt{3}/(2\alpha))\sqrt{k}$, of these D_i must lie either within or on C or else outside or on C. Assume we have the latter situation; the former is handled identically. Order these disks in accordance with the order in which they meet C. By Lemma 2.2, if a_j and a_l are the centers of consecutive disks, then $\angle a_j O a_l > 0.8/r$. Arbitrarily choose one of the centers of these disks, say b_1, as a starting point and then, traveling clockwise, let b_2 be the center of the $\lceil 2.5\sqrt{r}\rceil$th disk, let b_3 be the center of the $2\lceil 2.5\sqrt{r}\rceil$th disk and, in general, let b_i be the

center of the $(i-1)\lceil 2.5\sqrt{r}\rceil$th disk. Then $\measuredangle b_iOb_{i+1} \geq (0.8/r)2.5\sqrt{r} = 2/\sqrt{r}$. By Lemma 2.1, these points are the vertices of a convex polygon. The number of these points is at least $(\sqrt{3}k/(2\alpha))/(\lceil 2.5\sqrt{r}\rceil)$. For $r \geq 10$, $\lceil 2.5\sqrt{r}\rceil \leq 3\sqrt{r}$. We also have $r \leq \alpha\sqrt{k}/\sqrt{3}$. Therefore the number of these points is at least $k^{1/4}/(3^{1/4}2\alpha^{1.5})$. Hence setting $\beta = 1/(3^{1/4}2\alpha^{1.5})$ gives the result, for k sufficiently large. $\qquad\square$

Arguments similar to those used to prove Theorem 2.1 can be used to prove the following:

Theorem 2.2. *For any δ satisfying $\frac{1}{2} \leq \delta < \frac{2}{3}$ for any positive constant α, there exists a positive constant $\beta = \beta(\alpha, \delta)$ such that if A is a set of k points in general position in the plane where $q(A) \leq \alpha k^\delta$, then $c(A) \geq \beta k^{1-1.5\delta}$.*

3. The Upper Bound

In this section we prove the following result:

Theorem 3.1. *For every $\varepsilon > 0$ there is an integer $k(\varepsilon)$ such that for every $k > k(\varepsilon)$ there exists a set A of k points in the plane satisfying $q(A) < 4\sqrt{k}$ and $c(A) < 2k^{1/3+\varepsilon}$.*

The proof is probabilistic. We start with a geometric lemma.

Lemma 3.1. *Let P be a convex polygon on m vertices and let q denote the diameter of P, i.e., the maximum distance between a pair of vertices of P. Then for every t in the range $0 < t < m/15$ there are t consecutive vertices $p_1, p_2, \ldots, p_t$ of P such that:*

(a) *The distance between p_i and p_{i+1} is less than $10tq/m$ for all $1 \leq i \leq t-1$.*
(b) *$\pi \geq \measuredangle p_i p_{i+1} p_{i+2} \geq \pi - 15t/m$ for all $1 \leq i \leq t-2$.*

Proof. Let $q_0, q_1, \ldots, q_m = q_0$ be the vertices of P in accordance with their cyclic order. By Jung's theorem there exists a disk of radius $\sqrt{3}/3 \cdot q$ which contains P. The perimeter of this disk is $2\pi\sqrt{3}/3 \cdot q < 5q$ and hence the perimeter of P is smaller than $5q$. Thus the sum of the lengths of the edges q_iq_{i+1} is smaller than $5q$. Put

$$B = \{i: q_iq_{i+1} > 10tq/m\}.$$

Since $|B|10tq/m \leq \sum_{i \in B} q_iq_{i+1} < 5q$, we conclude that

$$|B| < m/2t. \tag{3.1}$$

For every index i, let ε_i denote the quantity $\pi - \measuredangle q_iq_{i+1}q_{i+2}$. Clearly, $\sum_{i=1}^{n} \varepsilon_i = 2\pi$ and each ε_i is nonnegative. Put

$$C = \{i: \varepsilon_i > 15t/m\}.$$

Since $|C|15t/m < \sum_{i \in C} \varepsilon_i \leq 2\pi$, we conclude that

$$|C| < 2\pi m/15t < m/2t. \tag{3.2}$$

By (3.1) and (3.2) there are less than m/t indices i with $i \in B \cup C$. Hence there are t consecutive i's such that $i \notin B \cup C$, i.e., such that $q_i q_{i+1} < 10tq/m$ and $\sphericalangle q_i q_{i+1} q_{i+2} \geq \pi - 15t/m$ for each of them. Let $p_1, p_2, \ldots, p_t$ be the corresponding points of P. Obviously these satisfy (a) and (b). This completes the proof.   □

Given a (large) positive perfect square $k = r^2$, we construct a random set $A = A_k$ in the plane as follows. For $0 \leq i < 2r$ and $0 \leq j < 2r$, let $S_{i,j}$ be the unit square given by $S_{i,j} = \{(i+x, j+y): 0 \leq x < 1, 0 \leq y < 1\}$. For each even i and even j, $0 \leq 1 < 2r$, $0 \leq j < 2r$, independently, let us choose one random point $p_{i,j}$ in $S_{i,j}$, according to a uniform distribution. $A = A_k$ is the set consisting of these $k = r^2$ points. Clearly, the distance between any two points of A is at least 1, and the diameter of A is at most $2\sqrt{2}\sqrt{k} < 3\sqrt{k}$. We next show that, for every fixed $\varepsilon > 0$, if $k = r^2$ is sufficiently large, then with positive probability (and, in fact, with probability that tends to 1 as k tends to infinity), A contains no convex polygon of size greater than $k^{1/3+\varepsilon}$. Indeed, by Lemma 3.1, with $m = \lfloor k^{1/3+\varepsilon} \rfloor$, $t = \lceil 2/\varepsilon \rceil$, and $q = 3\sqrt{k}$ we conclude that every such polygon must contain t consecutive vertices $p_1, p_2, \ldots, p_t$ such that

$$p_i p_{i+1} \leq 10tq/m = O(k^{1/6-\varepsilon}) \qquad \text{for all} \quad 1 \leq i < t, \tag{3.3}$$

$$\pi \geq \sphericalangle p_i p_{i+1} p_{i+2} \geq \pi - 15t/m = \pi - O(1/k^{1/3+\varepsilon}). \tag{3.4}$$

Put $x = 10tq/m$, $\sigma = \pi - 15t/m$. Let us estimate the probability that A contains such a sequence of t points.

We claim that the number of choices for the squares $S_{i,j}$, from which such a sequence can be formed, is bounded by $O(kx^t)$. (In all the O notation here we mean that ε (and hence t) are fixed, and k (and hence m) tend to infinity.) Indeed, the first $S_{i,j}$, denoted by S_1, can be any of the k possible squares, and the last $S_{i,j}$, denoted by S_t, must be within a distance of tx from the first, and thus the number of ways to choose these two is $O(kx^2)$. Given S_1 and S_t, each other candidate $S_{i,j}$ must be a square which is very close to the strip which is the convex hull of S_1 and S_t. Indeed, if $p_1, p_2, \ldots, p_t$ satisfy (3.4), then, for all $1 < i < t$, $\sphericalangle p_1 p_i p_t \geq \pi - 15t(t-2)/m = \pi - O(1/k^{1/3+\varepsilon})$, and hence the distance between p_i and the line $p_1 p_t$ is at most $O(txt^2/m) = o(1)$. There are $O(tx) = O(x)$ squares $S_{i,j}$ satisfying these requirements and the number of ways of choosing $(t-2)$ of them is $O(x^{t-2})$. This establishes the claim.

Given a fixed set of t distinct squares $S_1, S_2, \ldots, S_t$, in which it is possible to choose $P_i \in S_i$ which satisfy (3.3) and (3.4), we next claim that the probability that indeed the chosen p_i's will satisfy (3.3) and (3.4) is bounded by

$$O((\sin \sigma \cdot x)^{t-2}) = O(1/k^{(1/6+2\varepsilon)(t-2)}).$$

Indeed, given p_i and p_{i+1}, and given S_{i+2}, in order to satisfy $\sphericalangle p_i p_{i+1} p_{i+2} \geq \sigma$, p_{i+2} must lie in a region of area $O(\sin \sigma \cdot x) = O(1/k^{1/6+2\varepsilon})$ inside S_{i+2}. By the independence of our choices the claim follows.

Combining the two claims, we conclude that the probability that A contains t points $p_1, p_2, \ldots, p_t$ which satisfy (3.3) and (3.4), is bounded by

$$O((kx^t) \cdot 1/k^{(1/6+2\varepsilon)(t-2)}) = O(k^{1-3\varepsilon t+2(1/6+2\varepsilon)}).$$

Substituting $t = \lceil 2/\varepsilon \rceil$ and assuming, say, that $\varepsilon < \frac{1}{6}$, gives $O(1/k^4)$ and hence if k is sufficiently large, this probability is smaller than 1.

We thus proved that for every $\varepsilon > 0$ there is a $k(\varepsilon)$ such that for every perfect square $k > k(\varepsilon)$ there is a set $A = A_k$ of k points in the plane with $q(A) < 3\sqrt{k}$ and $c(A) < k^{1/3+\varepsilon}$. If $k > k(\varepsilon)$ is not a perfect square, put $K = \lceil \sqrt{k} \rceil^2$ and take any subset of cardinality k of the set A_k. This set has k points and satisfies $q(A) < 3\sqrt{K} < 4\sqrt{k}$ and $c(A) < K^{1/3+\varepsilon} < 2k^{1/3+\varepsilon}$. This completes the proof of Theorem 3.1.   □

It is interesting to note that if $q(A)$ is bounded by $\alpha\sqrt{k}$ then $c(A)$ is also bounded from above by $\gamma\sqrt{k}$:

Theorem 3.2.   *For any positive constant α, there exists a positive constant $\gamma = \gamma(\alpha)$ such that if A is a set of k points in a general position in the plane such that $q(A) \leq \alpha\sqrt{k}$, then $c(A) \leq \gamma\sqrt{k}$.*

Proof.   Again, assume that each pair of points of A is at least one unit apart. By Jung's theorem there exists a disk Q of radius at most $(\sqrt{3}/3)q(A) \leq (\sqrt{3}\alpha/3)\sqrt{k}$ which contains all points of A. The perimeter of this disk is at most $(2\pi\sqrt{3}\alpha/3)\sqrt{k}$. Let P be a convex polygon whose vertices belong to A which has $c(A)$ vertices. The length of P is therefore at most $(2\pi\sqrt{3}\alpha/3)\sqrt{k}$. Since each pair of vertices is at least distance 1 apart, P contains at most $(2\pi\sqrt{3}\alpha/3)\sqrt{k}$ vertices which gives the result.   □

Finally, a related problem: suppose that $q(a) \leq \alpha\sqrt{k}$ for some $\alpha > 0$ and k is large. Does it follow that A contains the vertices of an empty convex 7-gon, that is, a convex 7-gon which does not contain any members of A apart from its seven vertices? If the restriction on $q(A)$ is removed, then Horton [6] has constructed such sets A having no empty 7-gon.

References

1. P. Bateman and P. Erdös, Geometrical extrema suggested by a lemma of Besicovitch, *Amer. Math. Monthly* **58** (1951), 306–314; reprinted in [2], 667–675.
2. P. Erdös, *The Art of Counting*, MIT Press, Cambridge, MA, 1973.
3. P. Erdös and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463–470; reprinted in [2], 5–12.

4. P. Erdös and G. Szekeres, On some extremum problems in elementary geometry, *Ann. Univ. Sci. Budapest* **3/4** (1960/61), 53-62; reprinted in [2], 680-689.
5. H. Hadwiger, H. Debrunner, and V. Klee, *Combinatorial Geometry in the Plane*, Holt, Rinehart, Winston, New York, 1964 (English translation).
6. J. D. Horton, Sets with no empty convex 7-gons, *Canad. Math. Bull.* **26** (1983), 482-484.
7. H. Rademacher and O. Toeplitz, *The Enjoyment of Mathematics*, translated by H. Zuckerman, Princeton University Press, Princeton, NJ, 1957.