An Erdős-Szekeres type problem in the plane

Gyula Károlyi[*]
Eötvös University, Budapest
e-mail: karolyi@cs.elte.hu

Géza Tóth[†]
Massachusetts Institute of Technology and Hungarian Academy of Sciences
e-mail: toth@math.mit.edu

Abstract

Let $f(k,n)$, $n \geq k \geq 3$, denote the smallest positive integer such that any set of $f(k,n)$ points, in general position in the plane, contains n points whose convex hull has at least k vertices. We give lower and upper estimates on $f(k,n)$, both in the form $\Theta(kn) + 2^{\Theta(k)}$.

1 Introduction

A classical result of Erdős and Szekeres [ES1] states that, for every integer $n \geq 3$ there is a smallest positive integer $g(n)$ such that among any $g(n)$ points, in general position in the plane, there exist n points in convex position. The best known bounds for $g(n)$ are the following.

Theorem 1.1. [ES2, TV]

$$2^{n-2} + 1 \leq g(n) \leq \binom{2n-5}{n-2} + 2 \ .$$

The following generalization was motivated in [K]. For integers $n \geq k \geq 3$, let $f(k,n)$ be the smallest number with the property that among any $f(k,n)$ points in general position

[*]Supported by Hungarian research grants OTKA F030822 and FKFP 0151/1999.
[†]Supported by NSF grant DMS-99-70071, OTKA T020914 and OTKA F22234.

in the plane, there exist n points whose convex hull has at least k vertices. Clearly $f(k,n)$ exists and satisfies $g(k) \le f(k,n) \le g(n)$.

It follows from a canonical version of the Erdős-Szekeres theorem (see [BV, PS]) that, for any fixed k, $f(k,n)$ is a linear function of n. The coefficient of n however is of order $2^{\Omega(k^2)}$.

In this note we obtain the following improvements.

Theorem 1.2. *For arbitrary integers $n \ge k \ge 3$,*

$$\frac{(k-1)(n-1)}{2} + 2^{k/2-4} \le f(k,n) \le 2kn + 2^{8k} \ .$$

Better results are available for small values of k.

Theorem 1.3. $f(4,n) = \lceil 3n/2 \rceil - 1$.

Theorem 1.4. $2n - 1 \le f(5,n) \le 7n - 23$.

We prove these results in Sections 2 and 3, respectively. Section 4 contains the proof of the upper bound in Theorem 1.2, while the lower bound is proved in Section 5.

2 The case $k = 4$

Proof of Theorem 1.3. The lower bound follows from Theorem 1.2. To prove the upper bound, let P denote any set of at least $\lceil 3n/2 \rceil - 1$ points, in general position in the plane. If $\mathrm{conv}(P)$ has at least 4 vertices, then we are done. Therefore we may assume that $\mathrm{conv}(P)$ has only 3 vertices which we denote, in counter-clockwise order, by p_1, p_2, p_3. Let $q_1 = p_1$. We define the points $q_2, \dots, q_{\lceil n/2 \rceil}$ recursively as follows. Suppose that, for some $i \le \lceil n/2 \rceil - 1$ the convex hull of $P \setminus \{q_1, \dots, q_i\}$ has at least 4 vertices. In this case we have found at least $\lceil 3n/2 \rceil - 1 - i \ge n$ points whose convex hull has at least 4 vertices, and we are done. Thus, we may assume that the convex hull of $P \setminus \{q_1, \dots, q_i\}$ is a triangle $p_2 p_3 q_{i+1}$.

This way we have obtained points $q_1 = p_1, q_2, \dots, q_{\lceil n/2 \rceil}$ such that

$$P' = P \setminus \{q_1, \dots, q_{\lceil n/2 \rceil}, p_2, p_3\} \subset \triangle p_2 p_3 q_{\lceil n/2 \rceil} \subset \triangle p_2 p_3 q_{\lceil n/2 \rceil - 1} \subset \dots \subset \triangle p_2 p_3 q_1 \ .$$

Consider the points $q_2, \dots, q_{\lceil n/2 \rceil}$, in counter-clockwise order of visibility from p_1, and denote by r_1 and r_2 the first and the last points, respectively. Let s_i $(i = 1, 2)$ denote the intersection point of line $p_1 r_i$ with segment $p_2 p_3$. Note that $|P'| \ge n - 3$. Thus, we may assume, without any loss of generality, that the convex quadrilateral $p_2 r_1 r_2 s_2$ contains at least $\lceil (n-3)/2 \rceil$ points of P'. Denote the set of these points by P''. In this case p_1, p_2 and

r_2 are extremal points of the set $P^* = P'' \cup \{q_1, q_2, \dots, q_{\lceil n/2 \rceil}, p_2\}$, which has at least $\lceil (n-3)/2 \rceil + \lceil n/2 \rceil + 1 = n$ elements. Moreover, every point of P' lies inside triangle $p_2 r_2 p_3$, consequently, every point of P'' lies inside triangle $p_2 r_2 s_2$. Thus, P^* has at least one more extremal point. This completes the proof of the theorem.

3 The method of convex and concave chains

Theorem 3.1. *For arbitrary integers $n \geq k \geq 3$,*

$$f(k,n) \leq \binom{2k-5}{k-2} n \,.$$

Proof. Fix k an n. Let P denote a set of points, in general position in the plane, whose cardinality N is large enough. Let p denote one of its extremal points, and number the other points of P as $p_1, p_2, \dots, p_{N-1}$, in clockwise order of visibility from p. A *convex chain of length* ℓ with left (resp. right) endpoint p_{i_1} (resp. $p_{i_{\ell'}}$) is any sequence of $\ell' \geq \ell$ points $p_{i_1}, p_{i_2}, \dots, p_{i_{\ell'}}$ $(i_1 < i_2 < \dots < i_{\ell'})$, such that $pp_{i_1}p_{i_2}\dots p_{i_{\ell'}}$ is a convex $(\ell'+1)$-gon which contains at least $n-k-\ell'+\ell$ points of P in its interior. Similarly, a *concave chain of length* ℓ with left (resp. right) endpoint p_{i_1} (resp. $p_{i_{\ell'}}$) is any sequence of $\ell' \geq \ell$ points $p_{i_1}, p_{i_2}, \dots, p_{i_{\ell'}}$ $(i_1 < i_2 < \dots < i_{\ell'})$, such that the region bounded by the segments $p_{i_j}p_{i_{j+1}}$ $(1 \leq j \leq \ell-1)$ and the rays starting at point p and incident to points p_1 and p_ℓ, respectively, is an unbounded convex region which contains at least $n-k-\ell'+\ell$ points of P in its interior.

For $i,j \geq 2$, let $g_{k,n}(i,j)$ denote the smallest integer such that, for an arbitrary set P with N large enough, and for an arbitrary choice of its extremal point p, any $g_{k,n}(i,j)$-element subset of $\{p_1, p_2, \dots, p_{N-1}\}$ contains either a concave chain of length i or a convex chain of length j. When it does not cause any ambiguity, we simply write $g(i,j)$ for $g_{k,n}(i,j)$. It is immediate, that $g_{k,n}(2,j) = g_{k,n}(i,2) = n-k+2$ for any $i,j \geq 2$.

Lemma 3.2. *For $i,j \geq 3$, we have $g_{k,n}(i,j) \leq g_{k,n}(i-1,j) + g_{k,n}(i,j-1) - 1$.*

Proof. The proof is analogous to one of the original proofs of the Erdős-Szekeres theorem [ES1]. Suppose that N is large enough, and let $S \subset \{p_1, p_2, \dots, p_{N-1}\}$, $|S| = g(i-1,j) + g(i,j-1) - 1$. If S contains a concave chain of length i, we are done. Otherwise, since $|S| \geq g(i,j-1)$, it contains a convex chain of length $j-1$. Delete its left endpoint from S. Since we still have at least $g(i,j-1)$ points, there is another convex chain of length $j-1$. Delete its left endpoint from S again and continue as long as the remaining set has at least $g(i,j-1)$ points. We deleted $g(i-1,j)$ points of S, all of them are left endpoints of a convex chain of length $j-1$. By definition of $g(i-1,j)$, the set of deleted points contains either a convex chain of size j or a concave chain of size $i-1$. In the first case we are done. In the second case, let q be the right endpoint of that concave chain and let r be its second point from the right. q is also the left endpoint of some convex chain of length $j-1$, let s be its second point from the left. Now it is easy to see that depending on the angle $\angle rqs$, either the concave chain can be extended by s or the convex chain can be extended by r, concluding the proof of the lemma.

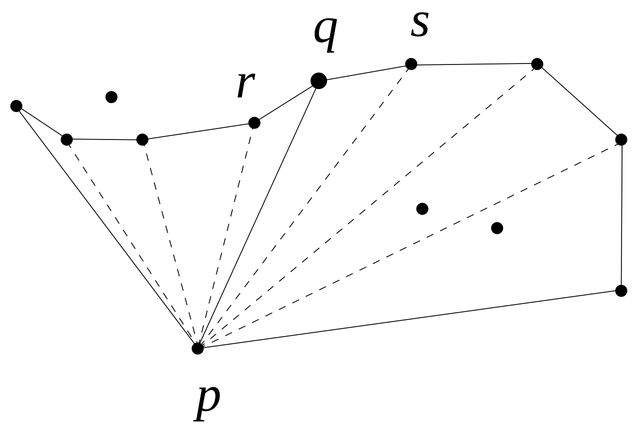


Figure 1.

Since $g_{k,n}(i,2) = g_{k,n}(2,j) < n$, it follows by induction that $g_{k,n}(i,j) < \binom{i+j-4}{i-2}n$, in particular, $g_{k,n}(k,k-1) \leq \binom{2k-5}{k-2}n - 1$. Consequently, if $N \geq \binom{2k-5}{k-2}n$, then either P contains a concave chain of length k, or it contains a convex chain of length $k-1$, and the result follows.

Proof of Theorem 1.4. The lower bound follows from Theorem 1.2. To prove the upper bound, notice first that $g_{k,n}(3,3) = n-k+3$. By repeated application of Lemma 3.2 we obtain

$$\begin{aligned} g_{k,n}(5,4) = g(5,4) &\leq g(4,4) + g(5,3) - 1 \\ &\leq g(3,4) + 2g(4,3) + g(5,2) - 3 \\ &\leq g(2,4) + 3g(3,3) + 2g(4,2) + g(5,2) - 6 \\ &= 3(n-k+3) + 4(n-k+2) - 6 \\ &= 7n - 7k + 11\ . \end{aligned}$$

Consequently, $f(5,n) \leq g_{5,n}(5,4) + 1 \leq 7n - 23$.

4 The upper bound

Proof of Theorem 1.2 (upper bound). Obviously, $f(3,n) = n$. Thus, in the sequel we assume $k \geq 4$. We prove the following estimate:

$$f(k,n) \leq \max\{(k-1)(2n-8k+19), 0\} + \max\{n-k+1, g(4k-10)\}\ .$$

Combining this with Theorem 1.1 the upper bound in Theorem 1.2 follows.

Let P be any set of $N \geq \max\{(k-1)(2n-8k+19), 0\} + \max\{n-k+1, g(4k-10)\}$ points, in general position in the plane. Peel off convex layers from P as follows. Let $P_0 = P$ and Q_0 be the vertices of the convex hull of P. If we already have P_i and Q_i, let $P_{i+1} = P_i \setminus Q_i$ and let Q_{i+1} be the set of vertices of the convex hull of P_{i+1}. If there is a smallest integer $i \leq 2n - 8k + 19$ such that $|Q_i| \geq k$, then it is easy to check that $|P_i| \geq n$. That is, we found at least n points whose convex hull has at least k vertices, and we are done.

We can therefore assume that $|Q_i| \leq k-1$ for $1 \leq i \leq t = \max\{2n-8k+19, 0\}$, implying that $P' = P_{t+1}$ has at least $g(4k-10)$ points. Consequently, P' contains the vertex set of a convex polygon $K = p_1, p_2, \ldots, p_{4k-10}$, in counter-clockwise order. The segments $p_{k-2}p_{k-1}$ and $p_{3k-7}p_{3k-6}$ are opposite sides of the polygon K, and we may assume, without any loss of generality, that rays r_1, starting at p_{k-2} and passing through p_{k-1}, and r_2, starting at p_{3k-6} and passing through p_{3k-7}, do not intersect each other.

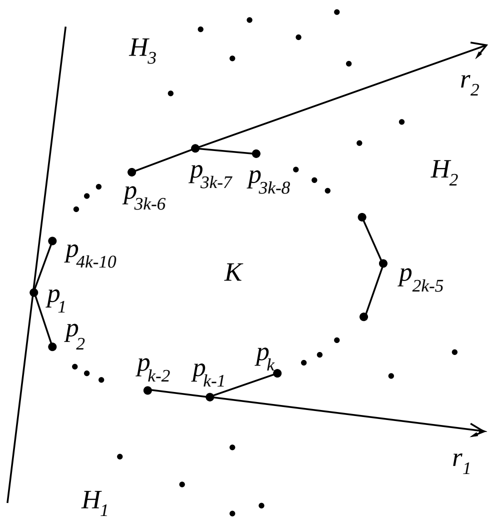


Figure 2.

Consider any open half plane H whose supporting line is incident to p_1 such that H contains points $p_2, p_3, \ldots p_{4k-10}$. The polygonal chains $(\cup_{i=1}^{k-3} p_i p_{i+1}) \cup r_1$ and $p_1 p_{4k-10} \cup (\cup_{i=3k-6}^{4k-9} p_i p_{i+1}) \cup r_2$ divide H into 3 open regions H_1, H_2, H_3, of which the middle one, H_2, contains vertices $p_k, p_{k+1}, \ldots, p_{3k-8}$ of K (see the Figure). Thus, $|P \cap H_2| \geq 2k - 7$.

It follows from the construction of the convex layers Q_i that $H \cap Q_i \neq \emptyset$ for $i = 1, 2, \ldots, t$. Consequently, $|P \cap H| \geq t + 4k - 11$. Define

$$R_1 = (P \cap (H_1 \cup H_2)) \cup \{p_2, p_3, \ldots, p_{k-1}\}$$

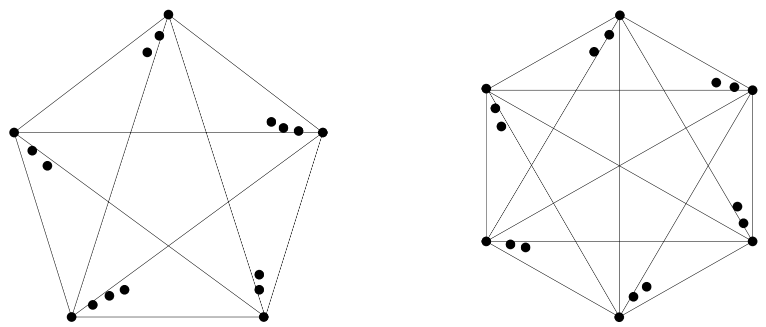
and

$$R_2 = (P \cap (H_3 \cup H_2)) \cup \{p_{3k-7}, p_{3k-6}, \dots, p_{4k-10}\} ,$$

then we have $|R_1|+|R_2| = |P \cap H|+|P \cap H_2| \geq t+6k-18 \geq 2n-2k+1$. If $|R_1| \geq n-k+1$, then $R_1 \cup \{p_1, p_2, \dots, p_{k-1}\}$ contains at least n points, and has at least k extremal points, including $p_1, p_2, \dots, p_{k-1}$. We argue similarly if $|R_2| \geq n-k+1$.

5 The construction

Proof of Theorem 1.2 (lower bound). In fact, we prove that $\left\lfloor \frac{(k-1)(n-1)}{2} \right\rfloor + a_k \leq f(k,n)$, where $a_k = 2^{\lfloor k/2 \rfloor - 3} + 1$ if $k \geq 6$ and $a_k = 1$ otherwise. First, for any $n \geq k \geq 4$ we obtain a set $P_{k,n}$ of $\left\lfloor \frac{(k-1)(n-1)}{2} \right\rfloor$ points, in general position in the plane, which does not contain n points whose convex hull has at least k vertices. Let $v_1, v_2, \dots, v_{k-1}$ denote, in this order, the vertices of a regular $(k-1)$-gon. Write $v_0 = v_{k-1}$, $v_k = v_1$ and $v_{k+1} = v_2$. For every $1 \leq i \leq k-1$, construct points $v_{i1} = v_i, v_{i2}, \dots, v_{it_i}$, where $t_i = \lfloor (n-1)/2 \rfloor$ if i is odd, and $t_i = \lceil (n-1)/2 \rceil$ if i is even, such that $v_i v_{i2} v_{i3} \dots v_{it_i} v_{i+1}$ is a convex polygon lying in the intersection of triangles $v_{i-1} v_i v_{i+1}$ and $v_i v_{i+1} v_{i+2}$ and, with the notation $K_i = \{v_i, v_{i2}, \dots, v_{it_i}\}$, every line $v_{ij} v_{ik}$ separates K_{i+1} from v_{i-1}.



$P_{6,8}$ $\qquad$ $P_{7,7}$

Figure 3.

Claim 5.1. *Suppose $u_1, u_2, \dots, u_r \in P_{k,n}$ are vertices of a convex polygon K.*

- *(i) If three of the u_α are vertices of some K_i, then K lies in the triangle $v_i v_{i+1} v_{i+2}$, with vertex v_{i+2} omitted.*
- *(ii) If two of the u_α are vertices of some K_i and K does not lie in triangle $v_i v_{i+1} v_{i+2}$, with vertex v_{i+2} omitted, then none of the u_α is of the form $v_{(i+1)j}$.*

It follows that if a subset of $P_{k,n}$ does not lie in triangle $v_iv_{i+1}v_{i+2}$, with vertex v_{i+2} omitted, then its convex hull may have at most $k-1$ vertices. On the other hand if a subset of $P_{k,n}$ does lie in triangle $v_iv_{i+1}v_{i+2}$, with vertex v_{i+2} omitted, then it has at most $n-1$ points. Thus, $P_{k,n}$ does not contain n points whose convex hull has at least k vertices, as we claimed above. This proves the lower bound in the case $k < 6$.

If $k \geq 6$ we can extend $P_{k,n}$ with $2^{\lfloor k/2 \rfloor - 3}$ points as follows. The segments v_iv_j divide the convex polygon $v_1v_2 \ldots v_{k-1}$ into finitely many regions. Denote by S the region which contains the centre of the polygon if $k-1$ is odd. If $k-1$ is even, then there are several regions which have the centre of the polygon on their boundary, let in this case S be one of these regions.

Claim 5.2. *Any line through an inner point of S which is not incident to any v_i separates $k/2$ of the v_i from the others if k is even; and separates either $\lfloor k/2 \rfloor$ or $\lceil k/2 \rceil$ of the v_i from the others if k is odd.*

In view of Theorem 1.1, there is a set E_k of $2^{\lfloor k/2 \rfloor - 3}$ points, in general position in the plane, which does not contain $\lfloor k/2 \rfloor - 1$ points in convex position. Let S_k denote the image of E_k under a suitable similarity such that S_k lies inside the region S and $P'_{k,n} = P_{k,n} \cup S_k$ is in general position. We claim that $P'_{k,n}$ does not contain n points whose convex hull has at least k vertices.

Notice first that S, and so S_k, too, is disjoint from any triangle $v_iv_{i+1}v_{i+2}$. Thus, if the vertex set of the convex hull of some subset of $P'_{k,n}$ is disjoint from S_k, then either the convex hull has at most $k-1$ vertices or the subset itself has less than n points. On the other hand, if the vertex set of the convex hull of some subset of $P'_{k,n}$ is not disjoint from S_k, then the convex hull has at most $\lfloor k/2 \rfloor - 2$ vertices in S_k. Moreover, it follows from Claims 5.2 and 5.1(ii) that the convex hull cannot have more than $\lceil k/2 \rceil + 1$ vertices in $P_{k,n}$. Altogether, it cannot have more than $k-1$ vertices in $P'_{k,n}$.

This completes the proof of the theorem.

Remark. Most likely S_k can be replaced by a larger set, maybe even of the size $c2^{2k}$, but any essential improvement would certainly require a lot of technical details.

Acknowledgments. Part of this research has been done while the first author visited the Institute for Theoretical Computer Science at the ETH Zurich. He is very grateful for the stimulating atmosphere and the hospitality of the Institute.

References

[BV] I. Bárány and P. Valtr, A positive fraction Erdős-Szekeres theorem, *Discrete and Computational Geometry* **19** (1998), 335–342.

[ES1] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Mathematica* **2** (1935), 463–470.

[ES2] P. Erdős and G. Szekeres, On some extremum problems in elementary geometry, *Ann. Universitatis Scientiarum Budapestinensis, Eötvös, Sectio Mathematica* **3/4** (1960–61), 53–62.

[K] Gy. Károlyi, Ramsey-remainder for convex sets and the Erdős-Szekeres theorem, *Discrete Applied Mathematics*, to appear.

[PS] J. Pach and J. Solymosi, Canonical theorems for convex sets, *Discrete and Computational Geometry* **19** (1998), 427–435.

[TV] G. Tóth and P. Valtr, Note on the Erdős-Szekeres theorem, *Discrete and Computational Geometry* **19** (1998), 457–459.