Almost Empty Polygons

HELENA NYKLOVÁ*

Department of Applied Mathematics, Charles University
Malostranské nám. 25, 118 00 Praha 1
Czech Republic
e-mail: nyklova@kam.ms.mff.cuni.cz

June 14, 2001

Abstract

In this paper we study a problem related to the classical Erdős–Szekeres Theorem on finding points in convex position in planar point sets. We study for which n and k there exists a number $h(n,k)$ such that in every planar point set X of size $h(n,k)$ or larger, no three points on a line, we can find n points forming a vertex set of a convex n-gon with at most k points of X in its interior. Recall that $h(n,0)$ does not exist for $n \geq 7$ by a result of Horton.

In this paper we prove the following results. First, using Horton's construction with no empty 7-gon we obtain that $h(n,k)$ does not exist for $k \leq c \cdot 2^{\frac{n}{4}}$. Then we give some exact results for convex hexagons: every point set containing a convex hexagon contains a convex hexagon with at most seven points inside it and any such set of at least 19 points contains a convex hexagon with at most five points inside it. Thus, $h(6,5) = 19$.

1 Introduction

In 1935 Erdős and Szekeres [ES35] proved that for any given integer n there exists a number $f(n)$ such that for any set containing at least $f(n)$ points in the plane, no three on a line, it is possible to select n points forming a convex n-gon. Later Erdős conjectured that for a given n there exists an integer $h(n)$ such that every point set X of size at least $h(n)$ in the plane contains a vertex set of a convex n-gon with no other point of X in its interior.

Partial answer to the last mentioned question was given in 1978 by Harborth [Har78]. He proved that an empty convex pentagon can be found in every set of at least ten points. Five years later, Horton constructed an arbitrarily large set without an empty convex heptagon (see [Hor83]). Therefore $h(n)$ does not exist for any $n \geq 7$. The problem of empty hexagons remains open.

We generalize Erdős's question and study whether there exists a number $h(n,k)$ such that in every planar point set X of size $h(n,k)$ or larger, no three points on a line, we can find n points forming a convex polygon with at most k points of X in its interior.

We talk about point sets in the plane and we always assume the set to be in *general position*; i.e. no three points collinear. We say that a point set P is in *convex position* if $x \notin \mathrm{conv}(P \setminus \{x\})$ for every $x \in P$.

The main and basic theorem we use is the following Erdős-Szekeres Theorem [ES35].

*Supported by Charles University grants No. 99/158 and 99/159 and by project LN00A056 of the Ministry of Education of the Czech Republic.

Theorem 1 (Erdős–Szekeres Theorem [ES35]) *For every $n \in \mathbb{N}$, there exists an integer $f(n)$ such that any set X of at least $f(n)$ points in general position in the plane contains n points in convex position.*

From now on, $f(n)$ will denote the smallest number satisfying the Erdős–Szekeres Theorem. The best known bounds for $f(n)$ are due to Erdős and Szekeres [ES35, ES60] (the lower bound) and Tóth and Valtr [TV97] (the upper bound):

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

The lower bound is sharp for $n \leq 5$ and is conjectured to be sharp for all n.

Definition 2 *We define $h(n,k)$ as the smallest number such that for each set X of at least $h(n,k)$ points in general position in the plane there exists a set $P \subseteq X$ of n points in convex position such that $|Int(\operatorname{conv} P) \cap X| \leq k$.*

We use the Horton's construction [Hor83] with no empty heptagon to prove the non-existence of $h(n,k)$ for relatively small k.

Theorem 3 *For every $n \in \mathbb{N}$, $h\left(n, c \cdot 2^{\frac{n}{4}} + O(n)\right)$ does not exist, where $c \in \mathbb{R}$ is a constant, $c \doteq 9$.*

We prove three results about points in the interior of convex hexagons.

Theorem 4 *In every set X of points in general position in the plane, such that there is a subset $P \subseteq X$ of 6 points in convex position, there is a subset $P' \subseteq X$ of 6 points with at most 7 points in its interior. As a consequence $h(6,7) = f(6)$.*

Theorem 5

$$h(6,6) = f(6).$$

Theorem 6

$$h(6,5) = max\{f(6), 19\}.$$

We prove Theorem 3 in Section 2. In Section 3 we prove Theorems 4, and 5 and sketch the proof of Theorem 6.

2 Almost empty n-gons

We construct an arbitrarily large set, which does not contain a vertex set of a convex n-gon with fewer than k points inside, to prove that $h(n,k)$ does not exist for $k \leq c_i \cdot 2^{\frac{n}{4}} - n - 3$. $c_i \doteq 3$ is a constant dependent on $i \equiv n \pmod 4$. The construction we use is due to Horton [Hor83] and Valtr [Val92a].

Let X and Y be finite sets of points in general position in the plane. We say that X lies *high above* Y (and Y lies *deep below* X) if the following three conditions are satisfied:

(i) no two points in $X \cup Y$ have the same x-coordinate,

(ii) Y lies entirely below every line determined by two points of X,

(iii) X lies entirely above every line determined by two points of Y.

Consider a planar point set $H = \{(x_1, y_1),\ (x_2, y_2), \ldots, (x_N, y_N)\}$, with the points ordered by the increasing x-coordinate. $H_1 = \{(x_1, y_1), (x_3, y_3) \ldots\}$ denotes the points with odd indices, and $H_2 = \{(x_2, y_2), (x_4, y_4), \ldots\}$ the points with even indices. We say that H is a *Horton set* if H_2 lies high above H_1 (or H_2 lies deep below H_1) and both H_1 and H_2 are Horton sets. Every set H of size $|H| \leq 2$ is a Horton set.

Theorem 7 ([Val92a]) *For every $M \in \mathbb{N}$ there exists a Horton set of size M.*

We say that points $c_1, c_2, \ldots, c_\ell$ (sorted by the increasing x-coordinate) form a *convex ℓ-chain* if they are in convex position and all the points $c_2, \ldots, c_{\ell-1}$ lie below the line $c_1 c_\ell$. They form a *concave ℓ-chain* if they are in convex position and the points $c_2, \ldots, c_{\ell-1}$ all lie above the line $c_1 c_\ell$.

We say that a point $a = (x, y)$ *lies above* a convex ℓ-chain (or *lies below* a concave ℓ-chain, respectively) $c_1 = (x_1, y_1)$, $c_2 = (x_2, y_2), \ldots$, $c_\ell = (x_\ell, y_\ell)$, if $x_1 < x < x_\ell$ and a lies above (or a lies below) all the lines $c_i c_{i+1}$ for $i = 1, \ldots l - 1$.

The question is how many points lie above any convex ℓ-chain and below any concave ℓ-chain in an arbitrary Horton set H. Consider any Horton set H containing a convex ℓ-chain. $c(\ell)$ denotes the smallest number of points of H lying above a convex ℓ-chain $c_1, \ldots, c_\ell \subset H$. Due to symmetry, $c(\ell)$ also denotes the smallest number of points of H lying below a concave ℓ-chain in H.

In any Horton set, the indices of points alternate between the 'upper' and the 'lower' Horton subsets. This property is given by the definition of the Horton set. Therefore the value of $c(\ell)$ is independent of the choice of H.

Lemma 8 *For every odd natural ℓ, $c(\ell) = 2^{\frac{\ell+1}{2}} - \ell - 1$. For every even natural ℓ, $c(\ell) = \frac{3}{2} 2^{\frac{\ell}{2}} - \ell - 1$.*

Proof. We proceed by induction on ℓ. We have a Horton set $H = \{h_1, \ldots, h_N\} = H_1 \cup H_2$ containing a convex ℓ-chain $C = \{c_1, c_2, \ldots, c_\ell\}$. Without loss of generality we can suppose H_2 lies high above H_1.

For $\ell = 2$ we take $C = \{h_i, h_{i+1}\}$, any two consecutive points. There is no point above nor below the segment $h_i h_{i+1}$, thus $c(2) = 0$. For $\ell = 3$ we take three consecutive points $\{h_{2i}, h_{2i+1},\ h_{2i+2}\}$ (beginning in the upper set). We have a convex 3-chain with no point above it and $c(3) = 0$.

Now we suppose that the identity holds for any $\ell' < \ell$. Choose a convex ℓ-chain $C = \{c_1, c_2, \ldots, c_\ell\} \subset H$ with the smallest number of points above it. Without loss of generality suppose $C \cap H_1 \neq \emptyset$ and $C \cap H_2 \neq \emptyset$. If $C \subseteq H_i$, take the smaller Horton set H_i with possibly fewer points above C. As the set H_2 is far above H_1, there can be at most two points of C, c_1 and c_ℓ, in $C \cap H_2$. $c_2, \ldots, c_{\ell-1}$ form a convex $(\ell-2)$-chain C' in a Horton set H_1. We know by the induction hypothesis that there are $c(\ell-2)$ points above any $(\ell-2)$-chain in H_1. H_1 contains the points with odd indices and in between any two consecutive points with odd indices, there is one point with even index. So we have at least $\ell - 2$ points of C' and $c(\ell-2)$ points above C' in H_1, that implies at least $c(\ell-2) + \ell - 2 - 1$ points above them in H_2. This is altogether $c(\ell) \geq 2c(\ell-2) + \ell - 3$.

Now we distinguish the two cases of ℓ being odd or even.

- The induction hypothesis for an odd ℓ is $c(\ell-2) = 2^{\frac{\ell-1}{2}} - \ell + 1$.

$$c(\ell) \geq 2c(\ell-2) + \ell - 3 = 2\left(2^{\frac{\ell-1}{2}} - \ell + 1\right) + \ell - 3 = 2^{\frac{\ell+1}{2}} - \ell - 1$$

- The induction hypothesis for an even ℓ is $c(\ell-2) = \frac{3}{2} 2^{\frac{\ell-2}{2}} - \ell + 1$.

$$c(\ell) \geq 2c(\ell-2) + \ell - 3 = 2\left(\frac{3}{2} \cdot 2^{\frac{\ell-2}{2}} - \ell + 1\right) + \ell - 3 = \frac{3}{2} \cdot 2^{\frac{\ell}{2}} - \ell - 1$$

Now we have proven the lower bound for $c(\ell)$. It is possible to construct (inductively) a Horton set consisting of a convex ℓ-chain and exactly the desired number of points above it. Therefore the equality holds. $\square$

Proof of Theorem 3. For every n we can — according to the Erdős–Szekeres Theorem — take a Horton set H, which contains a convex n-gon. We count how many points lie in each convex n-gon in H. We are able to construct an arbitrarily large Horton set and if we knew that in each convex n-gon there are at least N points inside, then $h(n, N-1)$ would not exist.

We choose a convex n-gon P with the vertices in H and with the smallest number of points inside. The Horton set H can be split into the two sets H_1 deep below H_2. We can suppose that P has a nonempty intersection with both of them. $P \cap H_2$ is a concave m-chain and $P \cap H_1$ a convex k-chain, where $k+m=n$. By Lemma 8, there are $c(m)$ points below any concave m-chain of the Horton set H_2 and $c(k)$ points of H_1 above any convex k-chain in H_1. As H_2 lies high above H_1, there are $c(k)+c(m)$ points of H in the convex hull of P. As c grows exponentially, the smallest number of points is contained in the polygon P with $k = \left\lceil \frac{n}{2} \right\rceil$ and $m = \left\lfloor \frac{n}{2} \right\rfloor$.

We solve the equation $N = c\left(\left\lceil \frac{n}{2} \right\rceil\right) + c\left(\left\lfloor \frac{n}{2} \right\rfloor\right)$ for the 4 cases of n, according to its remainder modulo 4. For example

- if $n \equiv 0 \pmod 4$, then $h(n, N-1)$ does not exist for

$$N - 1 \le 2 \cdot c\left(\frac{n}{2}\right) - 1 = 2 \cdot \frac{3}{2} \cdot 2^{\frac{n}{4}} - 2 \cdot \frac{n}{2} - 2 - 1 = 3 \cdot 2^{\frac{n}{4}} - n - 3,$$

- if $n \equiv 1 \pmod 4$, then $h(n, N-1)$ does not exist for

$$N - 1 \le c\left(\frac{n-1}{2}\right) + c\left(\frac{n+1}{2}\right) - 1 =$$
$$2 \cdot \frac{3}{2} \cdot 2^{\frac{n-1}{4}} - \frac{n-1}{2} - 1 + 2^{\frac{n+1}{4}+\frac{1}{2}} - \frac{n+1}{2} - 1 - 1 = 7 \cdot 2^{\frac{n-5}{4}} - n - 3,$$

and similarly the remaining two cases

- if $n \equiv 2 \pmod 4$, then $h(n, N-1)$ does not exist for $N - 1 \le 2^{\frac{n+6}{4}} - n - 3$,
- if $n \equiv 3 \pmod 4$, then $h(n, N-1)$ does not exist for $N - 1 \le 5 \cdot 2^{\frac{n-3}{4}} - n - 3$.

In general we have proved $h\left(n, c_i \cdot 2^{\frac{n}{4}} - n - 3\right)$ does not exist. The constant $c_i \doteq 3$ depends only on $i \equiv n \pmod 4$, the remainder of n when divided by 4.

A slightly better constant $c'_i \doteq 3c_i \doteq 9$ can be obtained when we replace each point in the original Horton set by a copy of another small Horton set.

Suppose there is no horizontal line spanned by a pair points of a sufficiently large Horton set H. Construct the maximum Horton set H_0 containing no convex nor concave $(m+1)$-chain. Its size is denoted by $H(m)$. Transform H_0 such that its diameter is much smaller than any distance of any two points of H, and any line spanned by the points of H_0 is almost horizontal. Replace each point of H by a copy H_0 and construct $H' = H + H_0$.

It can be proven inductively that $H(m) = 2^{\frac{m}{2}+1} - 2$ for any even m and $H(m) = 2^{\frac{m+1}{2}} + 2^{\frac{m-1}{2}} - 2$ for m odd. Then we have $c_m(\ell) = H(m) \cdot c(\ell - m + 1) + c(m)$ points above any convex ℓ-chain in H'. (There is a convex m-chain in each copy of H_0, but as all the lines are almost horizontal, we can use the bottom one only.) For large n and $m = \varepsilon n$ this gives us asymptotically 3 times more points in any convex polygon in H' (compared to the number of points in a polygon of the same size in H). It is not possible to improve the exponent of 2 this way. $\square$

The following table shows what is the maximal k such that $h(n,k)$ does not exist by Theorem 3, for some small n.

n	6	7	8	9	10	11	12	13	14	15	16	17	18	20	25
k	-1	0	1	2	3	6	9	13	19	27	39	51	63	119	373

3 Points in convex hexagons

3.1 More than six points inside

Every set K of at least $f(6)$ points in general position in the plane contains a convex hexagon. First we show that every such set contains a hexagon with at most seven points inside.

From now on K denotes a set of points in general position in the plane, which contains a convex hexagon. In K, we fix a convex hexagon $P = p_1p_2p_3p_4p_5p_6$ with the smallest number m of points inside. Its edges will be denoted by $e_1 = p_1p_2, \ldots, e_6 = p_6p_1$ and the points inside P by $Q = \{q_1, q_2, \ldots q_m\}$.

The points in P must satisfy the following two properties:

Property I: there is no point in the convex hull of any three consecutive vertices of P. If for example $p'_6 \in \mathrm{conv}\{p_5, p_6, p_1\}$ then the convex hexagon $p_1p_2p_3p_4p_5p'_6$ has fewer points inside.

Property II: any line determined by two distinct points inside P partitions the vertex set of P into two triples, otherwise there is a convex hexagon with fewer points inside (at least by two). For example if the line $p'_3p'_4$ intersects edges e_2 and e_4 there is a convex hexagon $p_1p_2p'_3p'_4p_5p_6$, which has fewer points inside. See Figure 1a.

We define the following description of the point configuration K. We define a table $T = \{t_{ij}\}_{i,j=1}^{m}$; $t_{ij} \in \{1, \ldots, 6\}$, $t_{ij} = k$ if the half-line $\overrightarrow{q_iq_j}$ intersects the edge e_k. The diagonal entries t_{ii} are not defined. T_i denotes the set of values of t_{ij}, where i is fixed and $i \neq j \in \{1, \ldots, m\}$.

Remark: when using (mod 6)*, we use 6 instead of 0; beginning with 1 seems to be more natural for numbering.*

Lemma 9 *Let Q be a set of points inside the convex hull of the convex hexagon P. If there is no other convex hexagon in $P \cup Q$, then the table T (defined in the previous paragraph) has the following properties:*

1. $t_{ij} \equiv t_{ji} + 3 \pmod 6$, *for each pair $i \neq j$,*
2. *for every i, there is no pair j, j' such that $t_{ij} \equiv t_{ij'} + 3 \pmod 6$,*
3. $|T_i| \leq 3$,
4. *if $i \neq j$ then $|T_i \triangle T_j| \geq 2$, where $T_i \triangle T_j = (T_i \setminus T_j) \cup (T_j \setminus T_i)$ is the symmetric difference,*
5. *if $T_i \neq \{1,3,5\}$ and $T_i \neq \{2,4,6\}$, then q_i lies on the convex hull of $\{q_1, \ldots, q_m\}$.*

Proof of 1. Easily follows from Property II: every line induced by two points inside a hexagon must intersect the opposite edges of P. See Figure 1a.

Proof of 2 and 3. Suppose to a contrary that there is a triple i, j, j' violating **2**, and suppose that the half-line $\overrightarrow{q_iq_j}$ intersects e_1 and $\overrightarrow{q_iq_{j'}}$ intersects e_4. Then either $\{p_2, p_3, p_4, q_{j'}, q_i, q_j\}$ or $\{p_1, q_j, q_i, q_{j'}, p_5, p_6\}$ is in convex position. Again we have a convex hexagon with fewer points inside. See Figure 1b.

Property **3** is an immediate consequence of **2**. T_i can contain at most one element of each pair : $\{1,4\}$, $\{2,5\}$ and $\{3,6\}$; that is at most three elements.

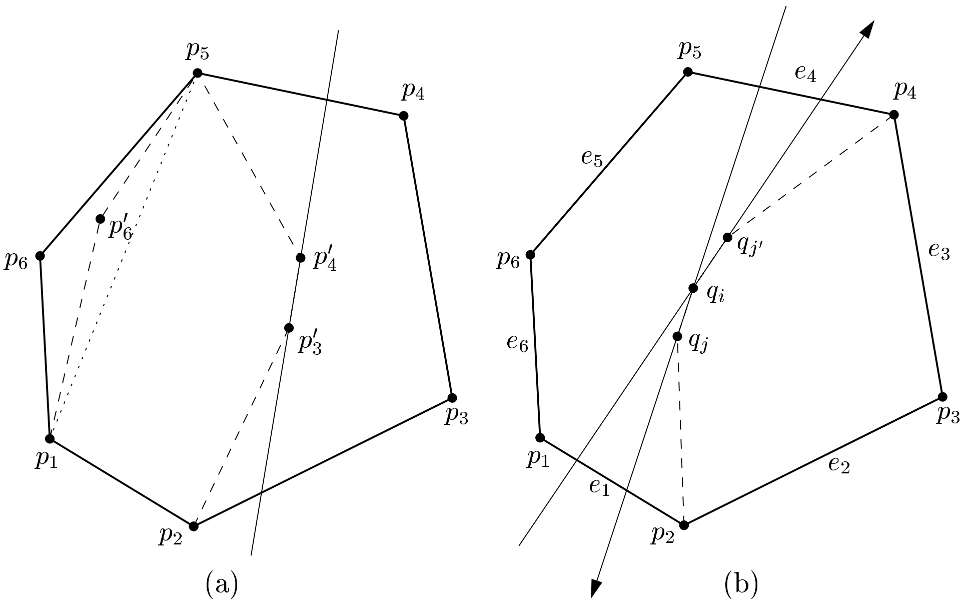

Figure 1: Illustration to the proof of Lemma 9.

Proof of 4. By **1**, $t_{ji} \equiv t_{ij} + 3 \pmod 6$. We know by **2** that $t_{ij} + 3 \pmod 6$ cannot be in T_i, but t_{ji} certainly is in T_j. Symmetrically $t_{ji} \neq t_{ij} \notin T_j$, but $t_{ij} \in T_i$. Thus $\{t_{ij}, t_{ji}\} \subseteq T_i \triangle T_j$.

Proof of 5. T_i is a subset of three consecutive edges. We can suppose without loss of generality that $T_i \subseteq \{1, 2, 3\}$. Therefore all other points q_j lie inside the angle $\alpha = \sphericalangle\, p_1 q_i p_4$.

If α is convex, q_i lies on the convex hull of $\{q_1, ..., q_m\}$. Suppose that α is concave and for a contradiction, there are points q_j and q_k on the convex hull of Q such that the segment $q_j q_k$ intersects the segments $p_1 q_i$ and $q_i p_4$. But then the line $q_j q_k$ does not satisfy **1** and that is a contradiction. □

Proof of Theorem 4. To prove Theorem 4 we use the previous Lemma 9.

By **5** if T_i is neither $\{1, 3, 5\}$ nor $\{2, 4, 6\}$ then q_i lies on the convex hull of $\{q_1, \ldots, q_m\}$. Thus we can choose at most five sets different from $\{1, 3, 5\}$ and $\{2, 4, 6\}$, otherwise there are six or more vertices of the convex hull of Q and a convex hexagon and with fewer points inside. All the rows of T must differ at least by two elements, therefore the sets $\{1, 3, 5\}$ and $\{2, 4, 6\}$ can be used at most once each. And that is the conclusion — in every set containing a convex hexagon we can find a convex hexagon with at most seven points inside. An example of a configuration of a hexagon with seven points inside and no other convex hexagon present is given in Figure 2. □

Here is the corresponding table T:

	1	**2**	**3**	**4**	**5**	**6**	**7**	T_i
1	—	1	1	1	2	2	1	{1,2}
2	4	—	2	2	4	4	3	{2,3,4}
3	4	5	—	3	4	4	5	{3,4,5}
4	4	5	6	—	4	4	5	{4,5,6}
5	5	1	1	1	—	6	1	{5,6,1}
6	5	1	1	1	3	—	1	{1,3,5}
7	4	6	2	2	4	4	—	{2,4,6}

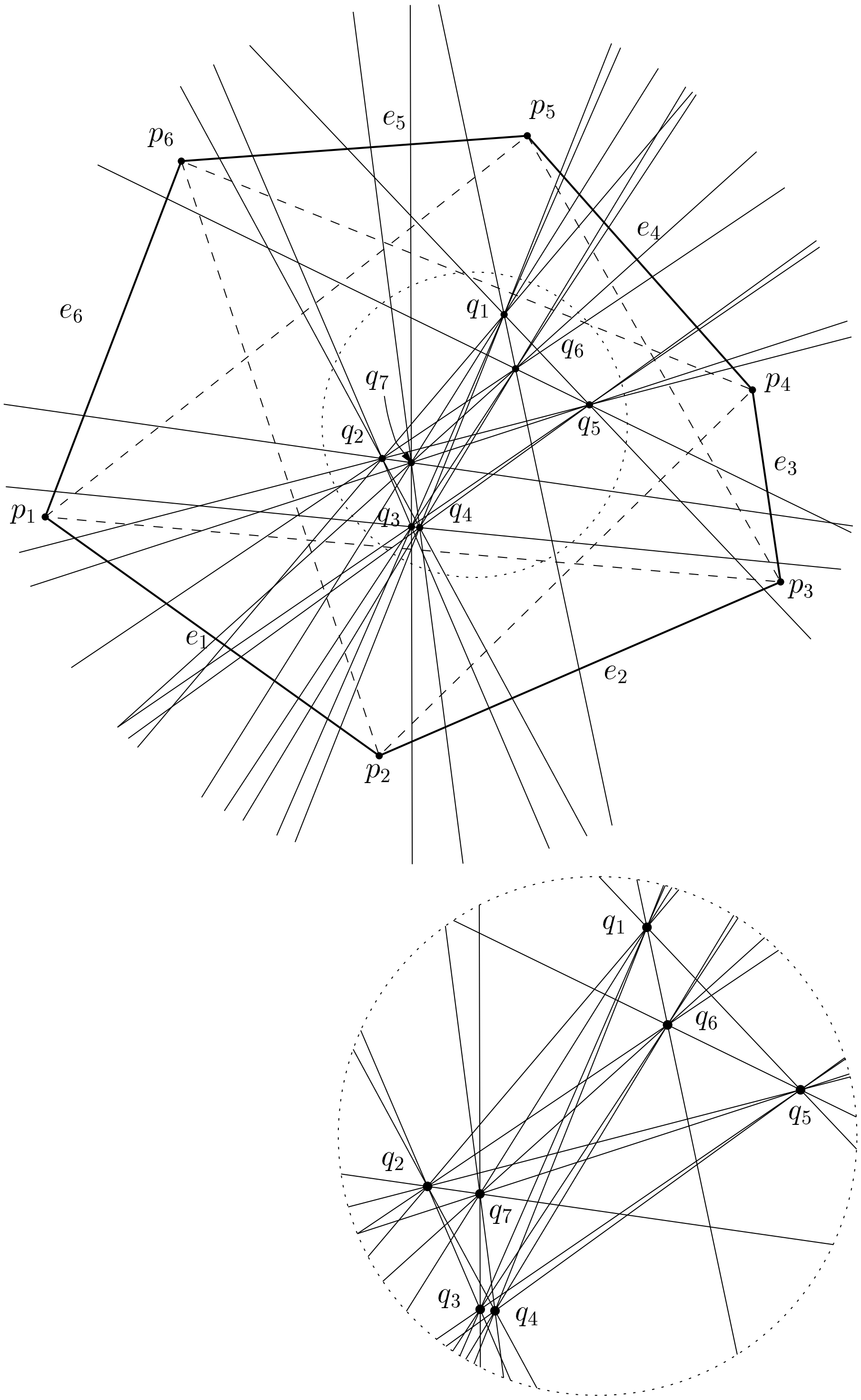

Figure 2: Example of a configuration of seven points inside a convex hexagon

3.2 Six points inside

Proof of Theorem 5. In the following we use the two general expressions with a special meaning. When we talk about a *configuration* we always mean a convex hexagon P with 7 (in the next section with 6) points inside its convex hull and possibly some other points outside, all in general position in the plane, such that there is no convex hexagon with fewer points inside its convex hull. When talking about *a hexagon* we mean the vertex set of a convex hexagon. We also use this expression in the meaning of the convex hull of such point set (e.g. a point p lies inside a hexagon), but this should not be confusing.

We know from the previous section that in every planar point set of size at least $f(6)$, a hexagon with at most seven points in the interior of its convex hull can be found. Now we examine what is the maximal size of a configuration with no hexagon with fewer than seven points inside. As a consequence, we prove Theorem 5.

Let $P = \{p_1, \ldots, p_6\}$ denote the vertices of the convex hexagon and $Q = \{q_1, \ldots, q_7\}$ denote the points inside. The remaining points form a set X.

In the following we often talk about half-planes; $(ab)^R$ denotes the half-plane bounded by the line ab, which lies on the right side when going from a to b. Symmetrically $(ab)^L$ denotes the half-plane on the left side of $\overrightarrow{ab}$. For five points of K, which are in convex position, *region* $a_1a_2a_3a_4a_5$ denotes the intersection of half-planes $(a_1a_2)^R \cap (a_2a_3)^R \cap (a_3a_4)^R \cap (a_4a_5)^R \setminus \operatorname{conv} a_1a_2a_3a_4a_5$, where $a_1, \ldots, a_5$ are indexed in the clockwise order.

We know by the previous Theorem 4 that the points of Q form the following structure. Five points lie on the convex hull of Q; let these five points be denoted by $q_1, \ldots, q_5$ and number them in the counter-clockwise order. The allowed directions for the remaining two points q_6 and q_7 are w.l.o.g. $T_6 = \{1, 3, 5\}$ and $T_7 = \{2, 4, 6\}$. No point of Q can lie in a triangle induced by three consecutive vertices of P. Altogether we have only four areas (marked gray in Figure 3), where $q_1, \ldots, q_5$ can lie.

Each of these four areas contains at least one point, otherwise q_6 or q_7 lies on the convex hull of Q and we have a hexagon with only one point inside. Therefore one of these four areas contains two of the points Q. Let these two points be q_1 and q_5. See Figure 3.

We now distinguish the two cases:
Case 1: The half-lines $\overrightarrow{q_1q_4}$, $\overrightarrow{q_2q_3}$ intersect the same edge of P,
Case 2: the half-lines $\overrightarrow{q_1q_4}$, $\overrightarrow{q_2q_3}$ intersect distinct edges of P.

Case 1: The half-lines $\overrightarrow{q_1q_4}$, $\overrightarrow{q_2q_3}$ intersect the same edge of P.

Without loss of generality we can suppose that $\overrightarrow{q_1q_4}$ and $\overrightarrow{q_2q_3}$ intersect e_5 and $q_1, q_5 \in \triangle p_1p_2q_6 \cap \triangle p_2p_3q_7$. Consequently $\overrightarrow{q_1q_5}$ must intersect e_6: If $\overrightarrow{q_1q_5}$ intersected e_5, we would have a hexagon $p_3p_4p_5q_4q_5q_1$ with only four points inside. If the intersected edge was e_1, we would have $q_5q_1q_6p_5p_6p_1$ with three points inside.

We now look for the 'prohibited areas' — parts of the plane, where no point can lie. For a better orientation in the letters, see Figure 4.

(1) No point $x \in X$ can lie in the region $p_1q_4q_7q_1p_2$, otherwise we have a convex hexagon $xp_1q_4q_7q_1p_2$ with one point (q_5) inside.

No point can lie in the following regions, otherwise we have an empty hexagon:

(2) region $p_4q_2q_6q_3p_5$,

(3) region $p_2q_1q_6q_2p_3$,

(4) region $p_6q_3q_6q_4p_1$,

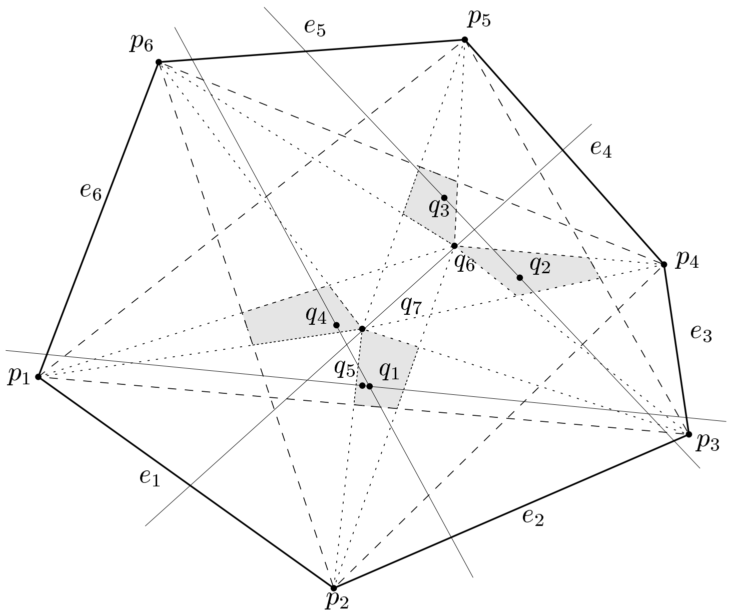

Figure 3: Notation of K.

(5) region $p_5q_3q_7q_4p_6$,

(6) region $p_3q_1q_7q_2p_4$.

Now only two areas remain, where a point of X could lie: intersection of half-planes $(q_2p_3)^L \cap (q_1p_3)^R$ and intersection of half-planes $(q_3p_6)^R \cap (q_4p_6)^L$. There are three more restrictions.

(7) No point can lie in the region $p_6q_4q_1p_2p_1$,

(8) region $p_3q_2q_3p_5p_4$,

(9) region $p_2p_1q_5q_1p_3$, otherwise we have a hexagon with at most one point inside.

In some cases of position of P and Q we still have two areas, where the points of X can lie, area $A = (p_1p_2)^R \cap (q_2p_3)^L \cap (p_5p_4)^L$ and area $B = (q_3p_6)^R \cap (p_2p_1)^L \cap (q_4p_6)^L$. The points of $X \cap A$ will be denoted by $x_1, x_2, \dots$ and the points $X \cap B$ by $y_1, y_2, \dots$, and number them according to their distance to the hexagon P.

The line x_1x_2 has to intersect the convex hull of P, otherwise there would be a convex hexagon $\{x_1x_2\}p_4q_6q_7p_2$ with at most three points inside. The half-line $\overrightarrow{x_2x_1}$ cannot intersect P on the right of p_3, i.e. $p_3 \notin (x_1x_2)^L$, otherwise we have a hexagon $p_1q_5q_1p_3x_1x_2$ with only one point inside, and $p_3 \notin (x_1x_2)^R$ neither: the hexagon $p_5q_3q_2p_3x_1x_2$ would contain only one point inside. Therefore there is no point x_2 and $|X \cap A| \leq 1$.

The situation is similar in the area B. Every line $\overrightarrow{y_jy_i}$, where $1 \leq i < j \leq 5$, must intersect $\operatorname{conv} P$, otherwise we have convex hexagon $\{y_iy_j\}p_5q_6q_7p_1$ with at most five points inside. $\overrightarrow{y_2y_1}$ must intersect $\operatorname{conv} P$ on the right of p_6, otherwise we have hexagon $p_2q_1q_4p_6y_1y_2$ with at most two points inside. There can be no point in the half-plane $(y_2y_1)^L$, because we would have hexagon $p_4q_3p_6y_1y_2y_3$ with only one point inside. Thus y_3 lies in $(y_2y_1)^R$ and no other point can lie in B: no other point lies in the half-plane $(y_3y_2)^L$ for the same reason as above and no other point can lie in the opposite half-plane, otherwise we have empty hexagon $p_2p_1y_1y_2y_3y_4$.

Altogether at most four points of X can lie in $A \cup B$, but if we allow all of them, we have empty hexagon $x_1p_2p_1y_1y_2y_3$. Therefore, we can place at most three points outside P, without obtaining a convex hexagon with fewer than seven points inside.

Case 2: The half-lines $\overrightarrow{q_1q_4}$ and $\overrightarrow{q_2q_3}$ intersect distinct edges of P.

The second case which we have to discuss is that $\overrightarrow{q_2q_3}$ and $\overrightarrow{q_1q_4}$ intersect distinct edges of P; that is without loss of generality $\overrightarrow{q_2q_3}$ intersects e_6, $\overrightarrow{q_1q_4}$ intersects e_5 and $\overrightarrow{q_1q_5}$ intersects e_6.

The prohibited areas (1),...,(6) remain the same as in the previous case and there are also three more restrictions for this case:

(7) there is no point in the region $p_6q_4q_1p_2p_1$,

(8) region $p_5p_4q_2q_3p_6$,

(9) region $p_2p_1q_5q_1p_3$.

As in the first case we have only two areas left, where any points of X can lie. First let us discuss the area $A = (q_3p_6)^R \cap (q_4p_6)^L \cap (p_2p_1)^L \cap (p_4p_5)^R$.

There can be at most one point in $X \cap A$: suppose there are two points $x_1, x_2 \in X \cap A$ and the line $\overrightarrow{x_2x_1}$ intersects $\operatorname{conv} P$ on the left of p_6, then the hexagon $p_2q_1q_4p_6x_1x_2$ has at most two points inside. If it intersected $\operatorname{conv} P$ on the right of p_6, there would be only one point in $p_4q_2q_3p_6x_1x_2$ — a contradiction. Thus $|X \cap A| \leq 1$.

Now we examine the area $B = (q_1p_3)^R \cap (q_2p_3)^L \cap (p_1p_2)^R$. The line y_2y_1 intersects $\operatorname{conv} P$ on the left of p_3 otherwise there is $p_1q_5q_1p_3y_1y_2$ with only one point inside. No point can lie in the half-plane $(y_2y_1)^R$, because of $p_5q_2p_3y_1y_2y_3$. Therefore y_3 lies on the left of $\overrightarrow{y_2y_1}$. No point can lie on the right of y_3y_2, because there would be only one point in $p_5q_2p_3y_2y_3y_4$, and no point can lie on the left of this line, because there would be empty hexagon $p_1p_2y_1y_2y_3y_4$. That is $|X \cap C| \leq 3$.

Conclusion

As the conclusion of the analysis of the two cases we obtain the following claim. In every set K, containing a vertex set of a convex hexagon, and of size at least $7+6+3+1 = 17$ points, we can find a convex hexagon with at most 6 points inside. But we know from Section 1 that $f(6) \geq 2^{6-2} + 1 = 17$ and $h(6,k) \geq f(6)$, thus $h(6,6) = f(6)$. □

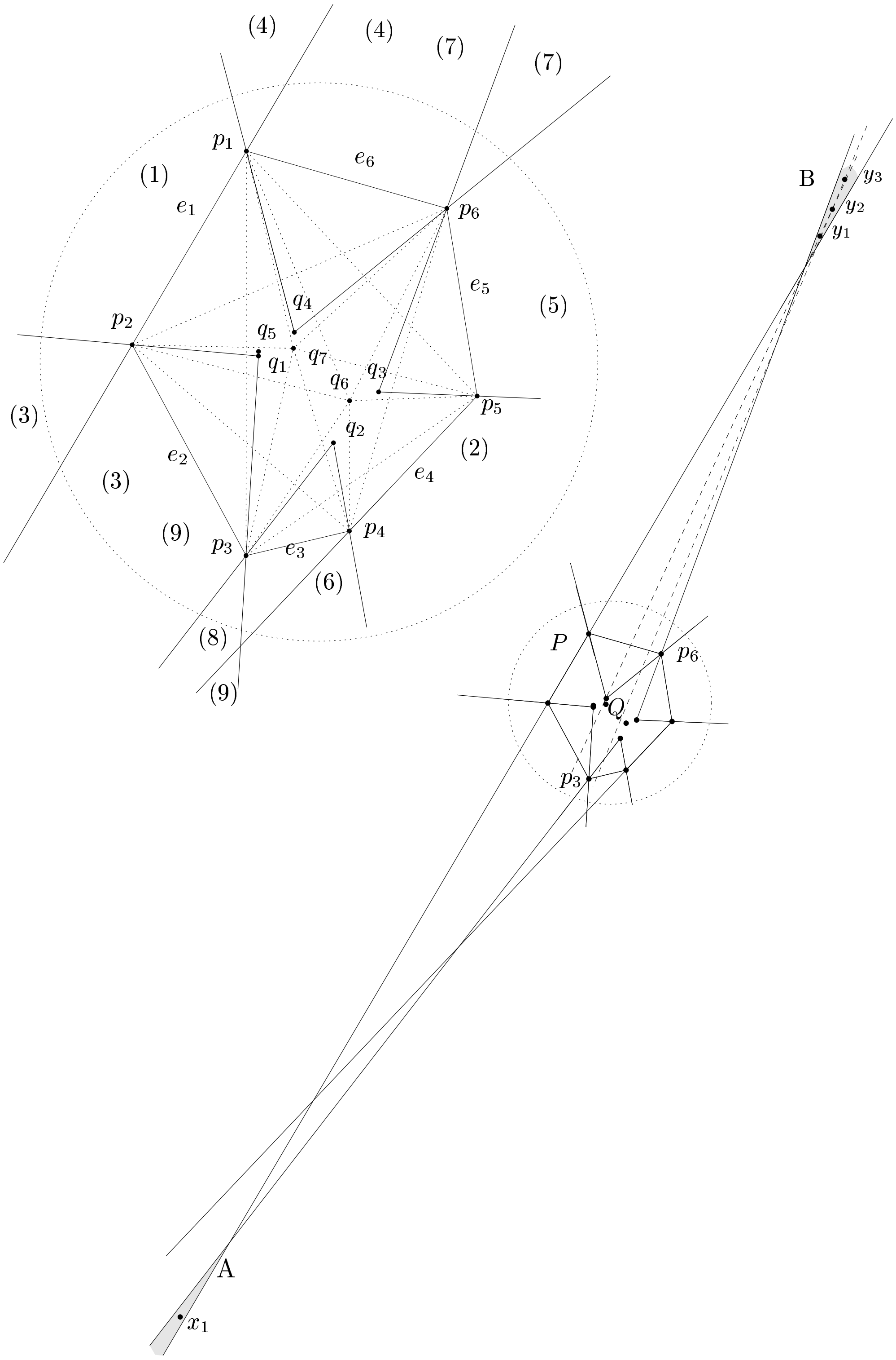

Figure 4: Construction for $\overrightarrow{q_1 q_4}$ and $\overrightarrow{q_2 q_3}$ intersecting the same edge

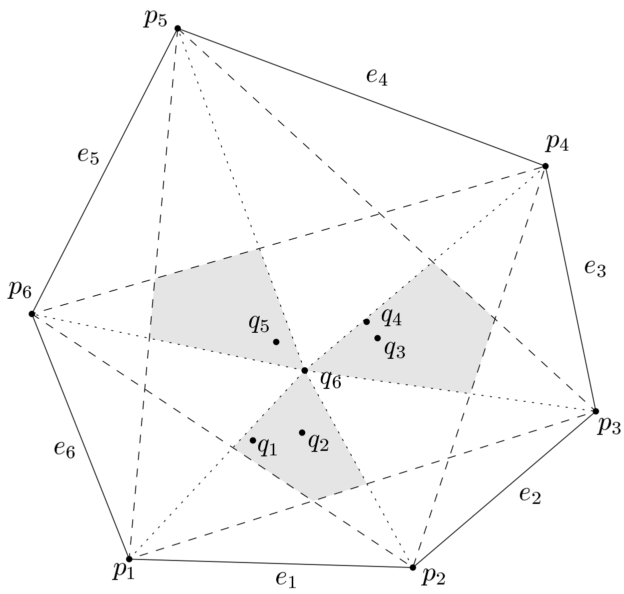

Figure 5: Notation of K in Case 2.1.

3.3 Five points inside

Proof of Theorem 6. We use exactly the same approach as in the proof of Theorem 5, only the number of cases to analyze is much larger. We will not discuss all of them in this paper.

Let us take a configuration K. If we have 7 points inside the hexagon P, we can use exactly the same arguments as in the proof of the previous Theorem 5, therefore the maximum size of the configuration K would be 16. Therefore we suppose we have a configuration containing a hexagon with six points inside.

It follows from Lemma 9 that we have two possibilities — 4 or 5 points lying on the convex hull of the inner points Q.

The first case of $4 = |\text{conv } Q|$ is very similar to the proof of the previous Theorem 5. The only difference is that we lose the point q_5 and the restrictions induced by this point. The first sub-case gives us at most 18 points in K. The resulting configuration is given in Figure 7. In the second sub-case the largest K consists of 16 points only.

Now we discuss the case of conv Q being a pentagon. We will call the one interior point q_6. According to Lemma 9, the remaining points of Q are allowed to lie in three non-consecutive directions from q_6; i.e. without loss of generality in the three quadrilaterals marked gray in Figure 5. Each of these three areas has to contain at least one point, otherwise q_6 lies on the convex hull of Q and we have an empty hexagon Q. There are two possibilities, how the points $q_1, \ldots, q_5$ can be distributed to the allowed areas:
Case 2.1: Exactly two points lie in two of the allowed areas and a single point in the remaining one.
Case 2.2: Three points lie in one area and one point in each of the remaining two.

We consider the case of points distributed 2+2+1 around q_6. Without loss of generality we can assume $\overrightarrow{q_6q_1}$ and $\overrightarrow{q_6q_2}$ intersect the same edge e_1, $\overrightarrow{q_6q_3}$ and $\overrightarrow{q_6q_4}$ intersect e_3 and $\overrightarrow{q_6q_5}$ intersects e_5. Only one of the half-lines $\overrightarrow{q_5q_4}$ and $\overrightarrow{q_5q_1}$ can intersect e_2, otherwise there is a hexagon with fewer points inside. Thus without loss of generality $\overrightarrow{q_5q_4}$ intersects e_3.

Prohibited areas for Case 2.1 are:

(1) region $p_2q_2q_6q_3p_3$,

(2) region $p_4q_4q_6q_5p_5$ and

(3) region $p_6q_5q_6q_1p_1$.

The intersection of half-planes $(q_1p_1)^L \cap (q_2p_2)^R$ will be denoted A, $(q_3p_3)^L \cap (q_4p_4)^R$ by B and $C = (q_5p_5)^L \cap (q_5p_6)^R$. The points of $X \cap A$ will be denoted by $x_1, x_2, \dots$, points of $X \cap B$ by $y_1, y_2, \dots$ and $X \cap C$ by $z_1, z_2, \dots$. The points are numbered increasingly with their distance to P. We have to explore these three areas with respect to whether they intersect or not. But not to bore the reader we analyze only $A \setminus (B \cup C)$ as an example. The remaining cases are very similar.

A is restricted by three more prohibited areas:

(4) region $p_1q_6q_3p_3p_2$,

(5) region $p_1p_6q_5q_6p_2$

and according to the direction of the line $\overrightarrow{q_1q_2}$

(6) region $p_1q_1q_2p_3p_2$ or

(6') region $p_1p_6q_1q_2p_2$.

Without loss of generality we can assume $\overrightarrow{q_1q_2}$ intersects e_3 and (6)= region $p_1q_1q_2p_3p_2$ is the prohibited area.

Lines q_6q_2 and q_6q_1 partition A into three parts. $A_1 = A \cap (q_6q_2)^R \cap (q_6q_1)^L$, $A_2 = A \cap (q_6q_2)^L$, $A_3 = A \cap (q_6q_1)^R$. When we consider the restrictions (1) to (6) we have:

$$A_1 = (q_6q_2)^R \cap (q_6q_1)^L \cap (q_3p_3)^R \cap (q_5p_6)^L \cap (p_6p_1)^R \cap (p_3p_2)^L,$$

$$A_2 = \left((q_2p_2)^R \cap (q_6q_2)^L \cap (q_3p_3)^R \cap (q_5p_6)^L \cap (p_3p_2)^L\right) \setminus \left((q_6p_2)^R \cap (p_6p_1)^L\right),$$

$$A_3 = (q_6q_1)^R \cap (q_1p_1)^L \cap (q_3p_3)^R \cap (q_5p_6)^L \cap (p_6p_1)^R \cap (p_3p_2)^L.$$

Every line $\overrightarrow{x_jx_i}$, where $1 \le i < j \le 7$ has to intersect p_1p_2, otherwise there is hexagon $p_1q_1q_2p_2x_ix_j$ with at most $j-2 < 6$ points inside. If x_i lies in A_1 then every line $\overrightarrow{x_jx_i}$ must intersect the segment q_1q_2, otherwise one of hexagons $p_3q_3q_6q_1x_ix_j$ and $p_6q_5q_6q_2x_ix_j$ contains at most 2 points only. If $x_i \in A_2$ then $\overrightarrow{x_jx_i}$ has to intersect P on the right of q_2, otherwise there is hexagon $p_3q_3q_6q_2x_ix_j$ with at most $j-1 < 6$ points inside, and if $x_i \in A_3$ then $\overrightarrow{x_jx_i}$ goes to the left of q_1: hexagon $p_6q_5q_6q_1x_ix_j$ would contain at most $j-1 < 6$ points only.

We discuss the cases of x_1 lying in the distinct areas A_1, A_2, A_3.

$x_1 \in A_1$ The line $\overrightarrow{x_2x_1}$ has to intersect q_1q_2. Suppose q_6 lies on the left of $\overrightarrow{x_2x_1}$. Then no other point x_i of X lies on the left of this line: only two points would lie in the hexagon $p_6q_5q_6x_1x_2x_i$. Therefore x_3 lies on the right of $\overrightarrow{x_2x_1}$ and no other point lies on the right of $\overrightarrow{x_3x_2}$, otherwise there is empty hexagon $p_3p_2x_1x_2x_3x_i$. If x_4 lies on the right of $\overrightarrow{x_3x_1}$, only four points can lie in A, otherwise one of the hexagons $p_3p_2x_1x_3x_4x_i$ and $p_6p_1x_2x_3x_4x_i$ is empty. If x_4 lies on the left of $\overrightarrow{x_3x_1}$, we can have one more point x_5 in between the lines $\overrightarrow{x_4x_2}$ and $\overrightarrow{x_4x_3}$. No other point can lie in A, otherwise $p_3p_2x_3x_4x_5x_6$ or $p_6p_1x_2x_4x_5x_6$ is empty. See Figure 6.

If q_6 lies on the right of $\overrightarrow{x_2x_1}$, no other point of X can lie on the right of this line and the situation is symmetric. We can place at most five points to A.

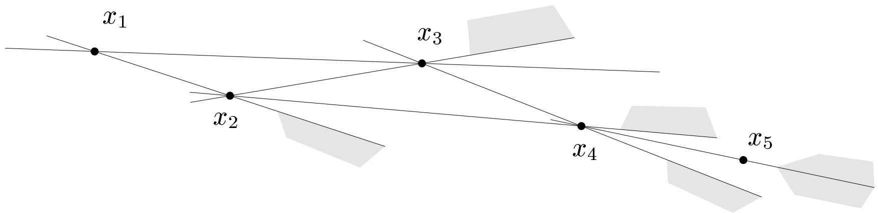

Figure 6

$x_1 \in A_2$ $\overrightarrow{x_2x_1}$ intersects q_2p_2 and no other point x_i lies on the left of $\overrightarrow{x_2x_1}$, otherwise there is empty hexagon $p_1q_1q_2x_1x_2x_i$. Therefore $x_3 \in (x_2x_1)^R$. No other point can lie on the left of $\overrightarrow{x_3x_1}$ as well and no other point can lie on the right of $\overrightarrow{x_3x_2}$, otherwise we have empty hexagon $p_3p_2x_1x_2x_3x_i$. If $x_2 \in A_2$, no point can lie on the left of $\overrightarrow{x_3x_2}$ either, therefore suppose x_3 lies in $A_3 \cup A_1$ and x_4 lies in between the lines $\overrightarrow{x_3x_2}$ and $\overrightarrow{x_3x_1}$. No point can lie on the right of $\overrightarrow{x_4x_3}$, because there would be hexagon $p_3p_2x_1x_3x_4x_i$ with only 1 point inside, and no other point can lie on the left of this line: there would be an empty hexagon $p_6p_1x_2x_3x_4x_i$.

$x_1 \in A_3$ $\overrightarrow{x_2x_1}$ intersects P on the left of q_1. No point can lie on the right of $\overrightarrow{x_2x_1}$, otherwise $x_3x_2x_1q_1q_2p_2$ is empty. x_3 lies on the left of $\overrightarrow{x_2x_1}$ and no point can lie on the left of $\overrightarrow{x_3x_2}$ (hexagon $p_6p_1x_1x_2x_3x_i$ would be empty). If x_2 lies in A_3 then no point can lie on the right of $\overrightarrow{x_3x_2}$ either. Therefore suppose $x_2 \in A_1 \cup A_2$ and x_4 lies in $(x_3x_2)^R \cap (x_3x_1)^L$. No other point can lie in A, otherwise there is empty hexagon $p_3p_2x_2x_3x_4x_i$ or $p_6p_1x_1x_3x_4x_i$.

We have the maximal number of points lying in A: $|X \cap A| \leq 5$. The similar analysis needs to be done for B and C, with respect to whether the areas intersect or not. The maximum size of K is 17.

The analysis of Case 2.2 is similar and the result is the same: the maximum size of K is 17.

Now we have discussed the four sub-cases with the result that the largest number of points is obtained in Case 1.1 (that is four points on the convex hull of Q and the half-lines $\overrightarrow{q_1q_4}$, $\overrightarrow{q_2q_3}$ intersecting the same edge of P), where we can place 6 additional points without obtaining a convex hexagon with less than six points inside. That is $h(6,6) \leq 19$ with respect to the fact that $h(6,6) \geq f(6)$, and we conclude $h(6,5) = max\{19, f(6)\}$.

□

Acknowledgement

I would like to thank my PhD advisor Pavel Valtr for spending his precious time in consultations with me and also for correcting the many mistakes I made while writing this text.

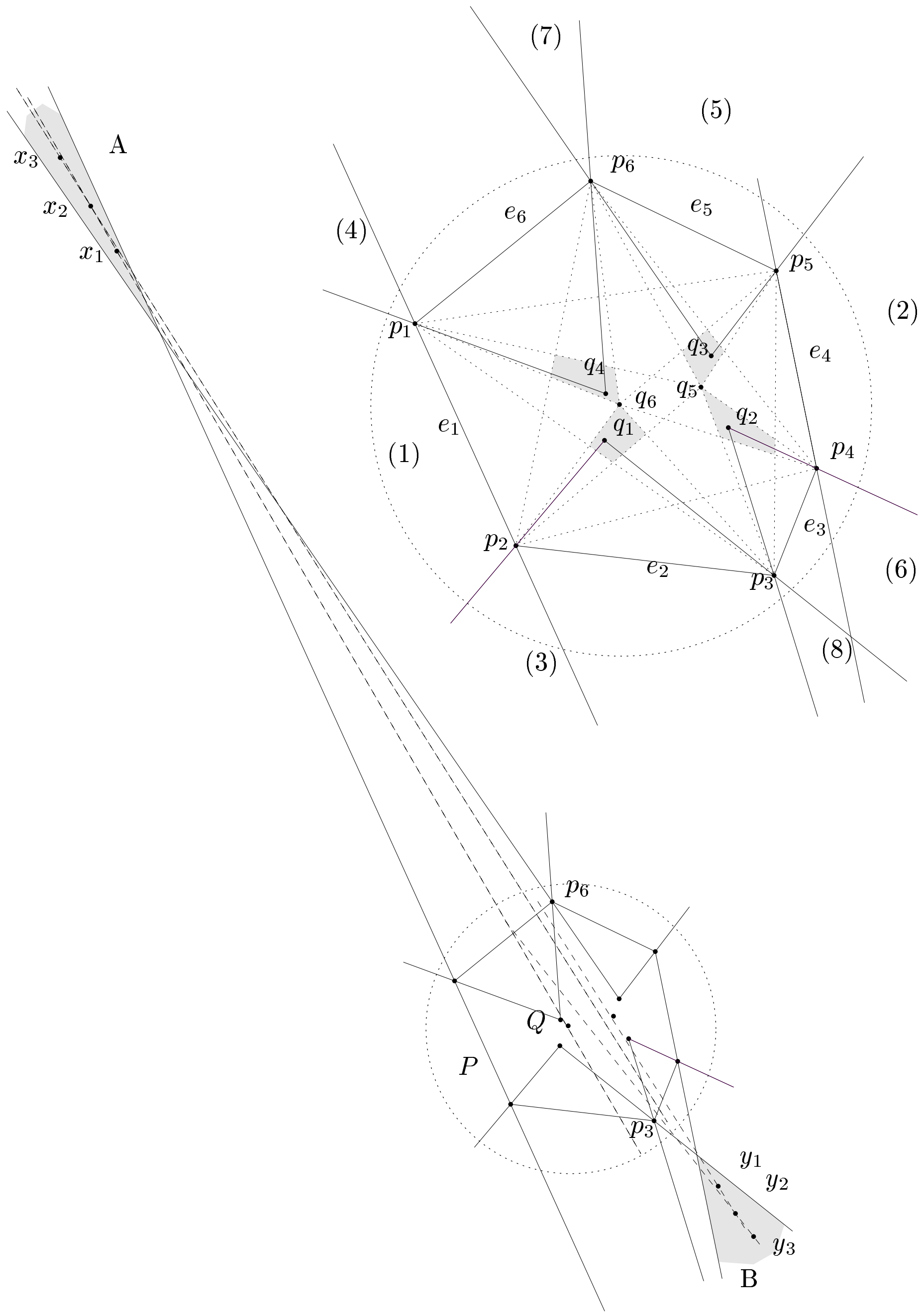

Figure 7: An example of a configuration of 18 points.

References

[ES35] P. Erdős, G. Szekeres: A combinatorial problem in geometry *Compositio Mathematica* **2** (1935), 463-470

[ES60] P. Erdős, G. Szekeres: On some extremum problems in elementary geometry *Ann. Univ. Sci. Budapest* **3/4** (1960/61), 53-62

[Har78] H. Harborth: Konvexe Fünfecke in ebenen Puktmengen, *Elem. Math.* **33** (1978), 116-118

[Hor83] J. D. Horton: Sets with no empty convex 7-gons *Canad. Math. Bull.* **26** (1983), 482-484

[HKU01] K. Hosono, Gy. Károlyi, M. Urabe: On the existence of a convex polygon with a specified number of interior points *Proceedings of the* 2^{nd} *Japanese–Hungarian Symposium on Discrete Mathematics and its Applications* (2001),

[KPT01] Gy. Károlyi, J. Pach, G. Tóth: A modular version of the Erdős–Szekeres Theorem, to appear in *Studia Sci. Math. Hungar.* (2001)

[KT99] Gy. Károlyi, G. Tóth: An Erdős–Szekeres type problem in the plane *Periodica Math. Hungar.* **39** (1999), 153-159

[Nyk99] H. Nyklová: Empty and almost empty polygons (Master Thesis) *Charles University, Prague* (1999)

[TV97] G. Tóth, P. Valtr: Note on the Erdős–Szekeres theorem *Discrete Comput. Geom.* **19** (1998), 457-459

[Val92a] P. Valtr: Convex independent sets and 7-holes in restricted planar point sets *Discrete Comput. Geom.* **7** (1992), 135-152

[Val92b] P. Valtr: Sets in $\mathbb{R}^d$ with no large empty convex subsets *Discrete Mathematics* **108** (1992), 115-124

[Val01] P. Valtr: A sufficient condition for the existence of large empty convex polygons, submitted (2001)