Point-sets with few k-sets

Helmut Alt* Stefan Felsner* Ferran Hurtado† Marc Noy†
Emo Welzl‡

22nd May 1998

Abstract

A k-set of a finite set S of points in the plane is a subset of cardinality k that can be separated from the rest by a straight line. The question of how many k-sets a set of n points can contain is a long-standing open problem where a lower bound of $\Omega(n \log k)$ and an upper bound of $O(nk^{1/3})$ are known today.

Under certain restrictions on the set S, for example, if all points lie on a convex curve, the number of k-sets is *linear*. We generalize this observation by showing that if the points of S lie on a constant number of convex curves, the number of k-sets remains linear in n.

Keywords: convex curve, k-set, Lovász' procedure.
AMS Subject Classification: 52C10, 52A10, 68R99, 05C99.

1 Introduction and definitions

Let S be a set of n points in the plane. A *k-set* $(1 \leq k \leq n-1)$ is a subset of S of cardinality k that can be separated from the rest by a straight line. The simple and natural question of how many k-sets a set of n points can contain has been considered for more than 25 years and has inspired considerable research. Nevertheless, it is still not solved completely and remains one of the most prominent open problems in combinatorial geometry.

The first upper bound is due to Lovász [6] who showed that the number of k-sets can be at most $O(n\sqrt{k})$ which is $O(n^{3/2})$. Erdős et al [5] constructed a family of sets having $\Omega(n \log k)$ k-sets which is $\Omega(n \log n)$ for suitable values of k. For a long time the gap between lower and upper bound could not be

*Institut für Informatik, Freie Universität Berlin, Takustr. 9, D-14195 Berlin, E-mail: *name*@inf.fu-berlin.de
†Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, Pau Gargallo 5, 08028-Barcelona, España. Partially supported by Proyecto MEC-DGES-SEUID PB96-0005-C02-02.
‡Theoretische Informatik, ETH Zentrum, CH-8092 Zürich, Switzerland

narrowed until Pach, Steiger and Szemerédi [7] showed an upper bound of $O(n\sqrt{k}/\log^* k)$. More significant progress was made only recently by Dey [3] who proved an upper bound of $O(n\sqrt[3]{k})$ which is $O(n^{4/3})$. An alternative proof of this bound is given in [2].

Various generalizations of the k-set problem or its dual formulation, namely determining the number of cells at the *k-level* of the arrangement of a set of n straight lines, have been considered. Among them are k-sets in higher dimensions or k-levels of curves in two dimensions and surfaces in three dimensions [1, 8].

Under certain restrictions on the set S it is possible to prove better upper bounds on the number of k-sets. For example, if all points of S lie on a straight line, there are only 2 k-sets. If all points of S lie on a convex curve, S contains at most n k-sets. Here, we will generalize the latter observation by showing that if the points of S lie on a constant number of convex curves, the number of k-sets remains *linear* in n.

In the following let S be a fixed set of n points in the plane and let $k \leq n-2$. An oriented straight line l is called a *k-line* of S exactly if the open halfplane right of l contains k points of S. Of special interest are k-lines determined by two points $p, q \in S$. The segment $\overline{pq}$ oriented from p to q of such a line is called a *k-segment* of S.

Let K be k-set of S and let l be a line separating K from $S \setminus K$. Allow l to move and rotate subject to the restriction that no point of S changes from one side to the other and note that l can reach exactly one position where it contains one point from $p \in K$ and one point $q \in S \setminus K$ such that the corresponding $(k-1)$-segment is oriented from p to q. This is a bijection between k-sets and $(k-1)$-segments, hence, the number of k-sets equals the number of $(k-1)$-segments.

To obtain bounds on the number of k-segments we make use of two powerful results from the classical article [5].

Alternation Lemma: *If we order all lines supporting k-segments containing a point $p \in S$ in counterclockwise order, then between any two lines supporting outgoing segments there is a line supporting an incoming segment and between any two lines supporting incoming segments there is a line supporting an outgoing segment.*

For the proof consider a line rotating around p. Just before an outgoing segment the line has $(k+1)$ points on its positive side and just after the segment it has k points on its positive side. So, between two outgoing segments the number of points on the positive side must change from k to $k+1$. This is only possible if the line scans some point q on the opposite side. So $\overline{qp}$ forms an ingoing k-segment between the two outgoing ones. An analogous argument holds for the symmetric case.

The second result is due to Lovász [6, 5].

Lovász' Lemma: *Let S be a set of n points and l a straight line containing no points of S and dividing S into two subsets of m and $n-m$ points, respectively. Then for any $k \leq n-2$ the number of k-segments intersecting l is at most $2\min\{m, n-m, k+1\}$.*

A simple consequence of Lovász Lemma is that the number of k-segments intersecting a straight line l is $O(n)$ (in fact, it is at most n). The proof of the lemma is based on a procedure that enumerates all k- and $(k-1)$-segments.

2 The main result

In the following we will consider point sets lying on a fixed collection of curves in the plane. Each curve can be closed or not, bounded or unbounded. We call a curve *convex* if it is a piece of the boundary of its convex hull Γ. Let γ be a closed convex curve, slightly stretching notation we will say a point is inside γ when the point is an element of Γ.

Our main result is stated in the following theorem.

Theorem 1 *Consider a fixed, finite collection of v pairwise disjoint convex curves in the plane. Then there is a constant $c_v > 0$ such that the number of k-sets of any set S of n points, each lying on one of the curves, is at most $4vn + c_v$.*

We first observe that without loss of generality we may assume that the points in S are in *general position* in the sense that no three of them lie on a straight line. In fact, let $K \subset S$ be a k-set. Then there exists a line l not containing any point of S which separates K and $S \setminus K$. Consider the arrangement of such lines for all $k-$sets. Then each point lies in the interior of a cell of this arrangement so that under slight perturbations of the points in S any k-set K still remains a k-set. So the points of S can be brought into general position without decreasing the number of k-sets.

The main step towards the proof of the theorem is the following lemma.

Lemma 1 *Let $\gamma_1 \dots, \gamma_t$ be closed convex curves where γ_i is contained inside γ_{i-1}, $i = 2, \dots, t$. The points of S may lie on the curves (subset C) or outside of γ_1 (subset A) see Figure 1. Then*

$$r_C \leq (2t-1)|C| + (t+1/2)r_A$$

where r_C denotes the number of k-segments incident to at least one point in C and r_A the number of k-segments with one endpoint in A and one in C.

Proof. Let I_k be the set of all k-segments $\overline{pq}$ such that q is on some curve γ_i and p is outside or on γ_i. Dually O_k is the set of k-segments $\overline{pq}$ such that q is on some curve γ_i and p is inside or on γ_i. The segments counted by r_A

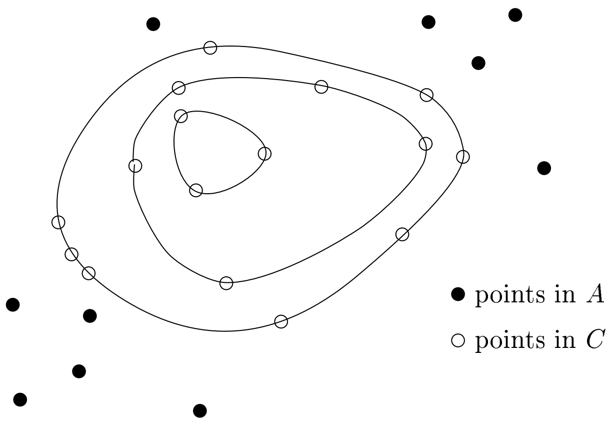

Figure 1: Nested convex curves.

are partitioned into those which are in I_k, let their number be r_A^I, and those in O_k, let their number be r_A^O. We next produce a bound on $|I_k|$.

For a point $p \in S$ on γ_i we partition the k-segments in I_k containing p into incoming and outgoing segments. The sizes of these classes are denoted as follows.

$\epsilon(p) := \#$ k-segments in I_k ending in p.
$\sigma(p) := \#$ k-segments in I_k starting in p.

Claim 1. $\sigma(p) - 1 \leq \epsilon(p)$ for all $p \in C$.

Proof. Consider a line l tangent to γ_i at p. W.l.o.g. we assume that l is vertical and γ_i is left of l. All k-segments counted by $\sigma(p)$ lead from p to some point left of l. As a consequence of the alternation lemma there are at least $\sigma(p) - 1$ k-segments starting in the right halfplane determined by l and ending in p. All these k-segments are counted by $\epsilon(p)$. △

The claim shows that we can match all but at most one of the k-segments leaving p towards the interior with a k-segment directed from outside towards p. Therefore, the k-segments in I_k can be partitioned into directed chains such that every chain either starts with a segment incident to some point of A or it is the unique chain spawn at p. Since a chain starting in A contains at most $t+1$ segments and a chain starting at a point in C contains at most t segments we obtain $|I_k| \leq t|C| + (t+1)r_A^I$.

By a completely symmetric argument we obtain $|O_k| \leq t|C| + (t+1)r_A^O$. The bound obtained for r_C by adding the bounds on I_k and O_k together can be slightly improved. A chain of segments of the maximal length t or $t+1$ respectively ends with a segment leading from γ_t to γ_t, i.e., with a segment in $I_k \cap O_k$. Suppose $r_A^O \geq r_A^I$ and let O'_k be O_k without segments leading from γ_t to γ_t it follows $|O'_k| \leq (t-1)|C| + tr_A^O$. Since $r_C \leq |I_k| + |O'_k|$ and $r_A = r_A^I + r_A^O$ we obtain the bound stated in the lemma. □

Next assume that the points of the set S lie on a finite number of closed convex curves which form u nested "clusters" each of the form described in Lemma 1 (see Figure 2).

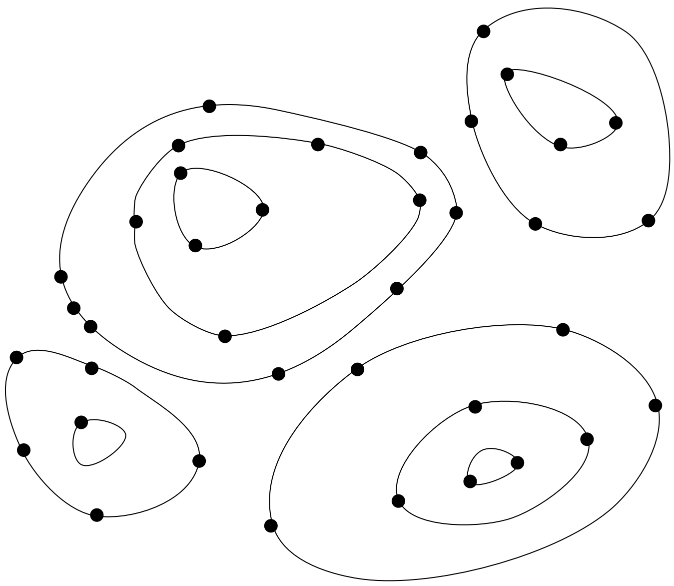

Figure 2: Four clusters of nested convex curves.

Proposition 1 *If the points of S lie on u disjoint nested clusters of closed convex curves such that each cluster consists of at most t curves, then the number of k-segments of S is at most $(2t-1)n + 12t(k+1)(u-2)$.*

We believe that this proposition is of independent interest. Especially in the case that the points lie on a constant number of unnested convex curves ($t = 1$) it shows that there are $n + O(1)$ k-sets for any constant k.

Proof. Consider one cluster C of the u clusters and let r_C be the number of segments with at least one endpoint on C. If C can be separated from all the other clusters by h_C straight lines, then by Lovász' Lemma there can be only $2(k+1)h_C$ segments with one endpoint in C and one endpoint outside. This is a bound for the number r_A of Lemma 1 of segments with one endpoint in C and one outside of C. From Lemma 1 we obtain $r_C \leq (2t-1)|C| + (t+\frac{1}{2})2(k+1)h_C$.

For a bound on the total number of k-segments of S we, therefore, need a bound on $\sum_C h_C$. Such a bound can be obtained from a result of Edelsbrunner et. al. [4], they show that $u \geq 3$ convex sets can be covered by a set of u disjoint g_i-gons with $\sum_i g_i \leq 6u-9$. A more careful glance at their construction shows that we can spare 3 more edges if we only require that any two sets are separated by polygonal chains. Thus, we obtain a set of straight lines with $\sum_C h_C \leq 6(u-2)$. Hence, $\sum r_C \leq (2t-1)n + (t+1/2)2(k+1)6(u-2)$.

Observing that all k-segments with endpoints in different clusters have been counted twice we obtain the bound stated in the lemma. □

Proof [**Theorem**]. Consider v disjoint convex curves containing all points of S. Since S is finite we can assume that all curves are bounded. To the curves we add all vertical lines through their endpoints and all vertical tangents (see Figure 3) This produces at most $4v$ vertical lines decomposing the plane into

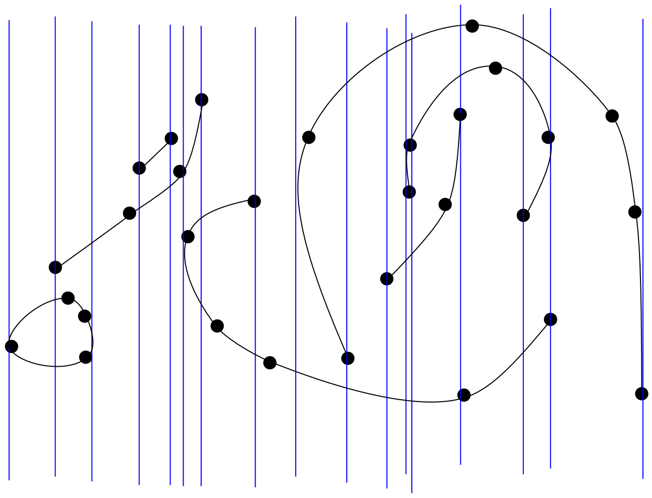

Figure 3: Vertical lines added to a set of convex curves.

at most $4v - 1$ slabs containing all points of S. Each slab contains at most $2v$ upward or downward convex pieces of curves (see Figure 4). Extending the pieces of curves in a slab we transform each slab into a into a set of disjoint clusters of convex curves (see Figure 4). The number of clusters in a slab and the number t of curves within each cluster are both bounded by $2v$. Overall we obtain $u \leq 8v^2$ clusters which -by construction- can be covered by the same number of disjoint 4-gons. Very similar to the argument proving Proposition 1 we obtain the bound $(2t-1)n + 8t(k+1)u$ for the number of k-segments. Replacing t and u by the above estimates we obtain the bound $4vn + 128(k+1)v^3$. □

References

[1] P. K. Agarwal, B. Aronov, T. Chan, and M. Sharir, *On levels in arrangements of lines, segments, planes, and triangles*, Discrete Comput. Geom., (1997), p. to appear.

[2] A. Andrzejak, B. aronov, S. Har-Peled, R. Seidel, and E. Welzl, *Results on k-sets and j-facets via continuous motions*, in Proc. 14th Annu. ACM Sympos. Comput. Geom., 1998.

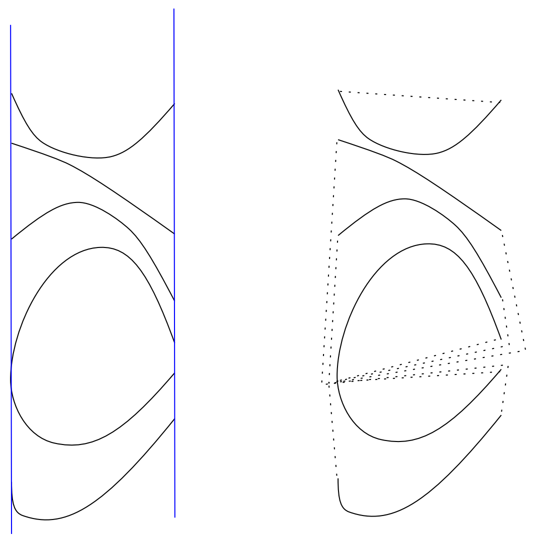

Figure 4: Curves restricted to a slab and their extension.

[3] T. Dey, *Improved bounds on planar k-sets and k-levels*, in Proc. 38rd Annu. IEEE Sympos. Found. Comput. Sci., 1997.

[4] H. Edelsbrunner, A. D. Robison, and X.-J. Shen, *Covering convex sets with non-overlapping polygons*, Discr. Math., 81 (1990), pp. 153–164.

[5] P. Erdös, L. Lovász, A. Simmons, and E. G. Straus, *Dissection graphs of planar point sets*, in A Survey of Combinatorial Theory, North-Holland, 1973, pp. 139–149.

[6] L. Lovász, *On the number of halving lines*, Ann. Univ. Sci. Budapest, Eötvös, Sect. Math., 14 (1971), pp. 107–108.

[7] J. Pach, W. Steiger, and E. Szemerédi, *An upper bound on the number of planar k-sets*, Discrete Comput. Geom., 7 (1992), pp. 109–123.

[8] M. Sharir, *On k-sets in arrangements of curves and surfaces*, Discrete Comput. Geom., 6 (1991), pp. 593–613.