Discrete Comput Geom 1:95–100 (1986)

Discrete & Computational Geometry

More on k-Sets of Finite Sets in the Plane*

Emo Welzl

Institutes for Information Processing, IIG, Technical University of Graz and Austrian Computer Society, Schießstattgasse 4a, A-8010 Graz, Austria

Abstract. For a set S of n points in the plane and for $K \subseteq \{1,2,\ldots,\lfloor \frac{1}{2}n \rfloor\}$, let $f_K(S)$ denote the number of subsets of S with cardinality $k \in K$ which can be cut off S by a straight line. We show that there is a positive constant c such that $f_K(S) < cn(\sum_{k \in K} k)^{1/2}$.

Introduction

For a set S of n points in the Euclidean plane E^2, we call a subset S' of S a *k-set of S*, $1 \le k \le n-1$, if S' contains exactly k points and it can be cut off S by a straight line (disjoint from S). Let $f_k(S)$ be the number of k-sets realized by S, and let $f_k(n) = \max\{f_k(T) \mid T$ a set of n points in $E^2\}$ for k, $1 \le k \le n$. It is trivial to observe that $f_k(n) = f_{n-k}(n)$; hence we restrict our attention to values k in the range $1 \le k \le \lfloor \frac{1}{2}n \rfloor$.

The only values of k for which $f_k(n)$ is known exactly are $f_1(n) = n$ which is easy to determine, and $f_2(n) = \lfloor 3n/2 \rfloor$ (see [3,4]; for more about $f_k(n)$ for small k see [9]). For arbitrary k, [6] proves the existence of positive constants c_1, c_2, and n_0 such that

$$c_1 n \log_2(k+1) \le f_k(n) \le c_2 n k^{1/2} \qquad \forall n \ge n_0 \tag{1}$$

(see [4] for an independent development of these bounds and also for some applications of the considered functions to problems in computational geometry).

In this article we are interested in the sum over $k \in K$ of the number of k-sets which can be realized by a set of n points in the plane, where K is a subset of

*This research was carried out during the author's stay at the University of Leiden, The Netherlands.

$\{1,2,\ldots,\lfloor\frac{1}{2}n\rfloor\}$. More specifically, for a set S of n points in the Euclidean plane and for $K\subseteq\{1,2,\ldots,\lfloor\frac{1}{2}n\rfloor\}$, we define $f_K(S)=\sum_{k\in K}f_k(S)$ and $f_K(n)=\max\{f_K(T)|T$ a set of n points in $E^2\}$.

Goodman and Pollack [7] showed that if $k<\frac{1}{2}n$, then $\sum_{i=1}^{k}f_i(S)\le 2nk-2k^2-k$ for a point set S of cardinality n. This bound was then improved to

$$\sum_{i=1}^{k} f_i(S) \le nk \quad \text{for } k<\tfrac{1}{2}n \tag{2}$$

in [1]. Moreover, the bound in (2) is tight so that

$$f_K(n) = nk \quad \text{for } K=\{1,2,\ldots,k\},\ k<\tfrac{1}{2}n. \tag{3}$$

Note that (3) does not hold for $k=\frac{1}{2}n$ if n is even. Otherwise we could calculate for even n:

$$f_{n/2}(n) = 2f_K(n)-n(n-1) \quad \text{for } K=\{1,2,\ldots,\tfrac{1}{2}n\}$$

which *would* give $f_{n/2}(n)=n$ if we apply (3); a contradiction to (1).

For the so-called half-planar range estimation problem we are interested in values $f_K(n)$, where $K=\{\alpha,2\alpha,\ldots,\lfloor n/(2\alpha)\rfloor\alpha\}$ for a positive integer α, $\alpha\le\frac{1}{2}n$ (see [5]). So far the only bound known was the one obtained using (1) directly. However, e.g., for $\alpha=n^{1/3}$, this gives already an upper bound which is (for large enough n) worse than the trivial one of n^2-n ($\ge\sum_{i=1}^{n-1}f_i(S)$ for a set S of n points). We show the following:

Theorem 1. *Let n be a positive integer and let $\phi\ne K\subseteq\{1,2,\ldots,\lfloor\frac{1}{2}n\rfloor\}$. Then*

$$f_K(n) < 2^{3/2}n\left(\sum_{k\in K}k\right)^{1/2}.$$

This includes the bound in (1) for $K=\{k\}$, and asymptotically also the bound in (3) for $K=\{1,2,\ldots,k\}$. For the above-mentioned case we get

Corollary 2. *There is a constant c_3, such that for $n\ge1$, for $1\le\alpha\le\lfloor\frac{1}{2}n\rfloor$, and for $K=\{\alpha,2\alpha,\ldots,\lfloor n/(2\alpha)\rfloor\alpha\}$,*

$$f_K(n) < c_3n^2\alpha^{-1/2}.$$

Thus the bound is better than the trivial quadratic one, unless α is a "constant." (We mean here that if α is a function in n which is monotonically increasing and unbounded, then Corollary 2 gives a nontrivial bound for large enough n.)

More general than Corollary 2 we get the following bound for $f_K(n)$ expressed in the cardinality of K (and in n).

Corollary 3. *There is a constant c_4 such that*

$$f_K(n) \leq c_4 n^{3/2}(\#K)^{1/2}.$$

($\#K$ denotes the cardinality of a finite set K.)

Proof of the Result

First we briefly describe the concept of (circular) n-sequences which is due to [7]. For a discussion more careful than here we refer to this original source or to [2].

Let S be a set of n points in general position, i.e., no three points are collinear and no two connecting lines are parallel. Consider now a directed line L which is not orthogonal to any line determined by two points in S. We label the points in S by $1,2,\ldots,n$ in such a way that their orthogonal projections on L form an increasing sequence. Let us observe these projections, while L rotates (counterclockwise, say). Then, whenever L passes through a direction orthogonal to a line connecting points i and j, the order of the projections on L changes by having indices i and j interchanged. After a rotation of 180°, the indices appear in reverse order, i.e., they form a decreasing sequence.

In this way we obtain a sequence of permutations

$$P(S) = P_0, P_1, \ldots, P_N$$

which we call the *n-sequence induced by* S and which has the following obvious properties

(i) $P_0 = 12\cdots n, \qquad P_N = n(n-1)\cdots 1, \quad \text{and} \quad N = \binom{n}{2}.$

(ii) For q, $1 \leq q \leq N$, P_q differs from P_{q-1} only by the interchanges of two *adjacent* indices i and j, where, for $i < j$, ij appears in P_{q-1} and ji appears in P_q. (That is, i may jump to the right only over a j with $j > i$.) We call ij the *switch (from P_{q-1} to P_q)* and if i is in the kth position in P_{q-1}, then we say that switch ij *occurs in position* k and we write $\mathrm{oc}(ij) = k$.

Every sequence of permutations which satisfies (i) and (ii) is called an *n-sequence*. (As we have seen, every point set induces such an n-sequence; however, there are n-sequences which are not induced by any point set, see [7].)

It is now easily seen that every k-set of S forms either a prefix or a suffix of length k in some permutation P_q, $1 \leq q \leq N$. Every switch which occurs in position k produces a new prefix of length k and a new suffix of length $n-k$ ("new" with respect to the occurring numbers). In an n-sequence only prefixes (suffixes) of P_0 of length k involve the same numbers as any suffix (prefix, respectively) of length k in a permutation. Hence, if we denote by $g_k(P)$ the number of switches which occur in position k in an n-sequence P, then we get the following:

Observation 4. Let S be a set of n points in general position in the plane and let $P(S)$ be an n-sequence induced by S. Then, for k, $1 \leq k \leq n-1$,

$$f_k(S) = g_k(P(S)) + g_{n-k}(P(S)).$$

Note that for even n and $k = \frac{1}{2}n$ the observation is also true, since then every switch in position $\frac{1}{2}n$ determines *two* new $\frac{1}{2}n$-sets, namely, the prefix *and* the suffix of length $\frac{1}{2}n$ in the following permutation.

We will prove that for every n-sequence P and every $\phi \neq K \subseteq \{1,2,\ldots,\lfloor \frac{1}{2}n \rfloor\}$:

$$\sum_{k \in K} g_k(P) < 2^{1/2} n \Big(\sum_{k \in K} k \Big)^{1/2}. \tag{4}$$

Since an upper bound for $\Sigma_{k \in K} g_k(P)$ is also an upper bound for $\Sigma_{k \in K} g_{n-k}(P)$, this will certainly yield our result for point sets in general position (by Observation 4). It is actually a more general result since there are n-sequences which are not induced by point sets. Moreover, since a sufficiently small perturbation of the points in a set S does not decrease $f_k(S)$ for any k, Theorem 1 will follow for all point sets.

Let P be a fixed n-sequence. For k, $1 \le k \le \lfloor \frac{1}{2}n \rfloor$, and i, $1 \le i \le n$, we define $L_k(i) = \{ji \mid j < i,\ \mathrm{oc}(ji) = k\}$ and $R_k(i) = \{ij \mid i < j,\ \mathrm{oc}(ij) = k\}$. Moreover, we set $l_k(i) = \#L_k(i)$ and $r_k(i) = \#R_k(i)$. Intuitively speaking, $L_k(i)$ is the set of switches in position k, where i "jumps" to the left, i.e., $l_k(i)$ is the number of "jumps" of i from position $k+1$ to position k.

Clearly, $g_k(P) = \Sigma_{i=1}^{n} l_k(i)$, and so

$$\sum_{k \in K} g_k(P) = \sum_{i=1}^{n} \sum_{k \in K} l_k(i) \tag{5}$$

for $K \subseteq \{1,2,\ldots,\lfloor \frac{1}{2}n \rfloor\}$. Our goal is to prove that $\Sigma_{i=1}^{n}(\Sigma_{k \in K} l_k(i))^2 < 2n \Sigma_{k \in K} k$, from which (4) will follow via (5).

Let $1 \le i \le k \le \lfloor \frac{1}{2}n \rfloor$. Then i starts in a position $\le k$ (in P_0) and i ends up in a position $> k$ (in P_N). Hence, i has to make one more jump from k to $k+1$ than it makes jumps from $k+1$ to k. This shows that $r_k(i) - l_k(i) = 1$. Analogously, the rest of the following observation can be seen: for i, $1 \le i \le n$, and for k, $1 \le k \le \lfloor \frac{1}{2}n \rfloor$

$$r_k(i) - l_k(i) = \begin{cases} +1 & \text{if } 1 \le i \le k \\ 0 & \text{if } k < i < n-k+1 \\ -1 & \text{if } n-k+1 \le i \le n. \end{cases} \tag{6}$$

In a next step, for m, $1 \le m \le n-1$, and for k, $1 \le k \le \lfloor \frac{1}{2}n \rfloor$, let $C_k(m)$ be the set of switches ij with $\mathrm{oc}(ij) = k$ and $i \le m < j$, and let $c_k(m) = \#C_k(m)$. Observe that $C_k(m)$ contains the switches in $C_k(m-1)$ *plus* the switches mj in position k *minus* the switches im in position k. Therefore, $c_k(m) = c_k(m-1) + r_k(m) - l_k(m)$. Thus, since $c_k(1) = 1$, we can infer from (6) that $c_k(2) = 2$, $c_k(3) = 3, \ldots, c_k(m) = m$ for $m \le k$, that $c_k(m) = k$ for $k \le m \le n-k$, and, finally, that $c_k(m) = n-m$ for $n-k \le m \le n-1$. Hence we get for all k

$$\sum_{m=1}^{n-1} c_k(m) = \sum_{m=1}^{k-1} m + \sum_{m=k}^{n-k} k + \sum_{m=n-k+1}^{n-1} (n-m) = nk - k^2. \tag{7}$$

For a switch $s = ij$, $i < j$, let $j - i$ be its *width* which we denote by $|s|$. Obviously, a switch s of width w and with $\mathrm{oc}(s) = k$ contributes exactly w to the sum in (7). Consequently,

$$\sum_{\mathrm{oc}(s)=k} |s| = nk - k^2 < nk. \tag{8}$$

For $K \subseteq \{1,2,\ldots,\lfloor \frac{1}{2}n \rfloor\}$, we extend the definition of $L_k(i)$ and $l_k(i)$ to $L_K(i) = \bigcup_{k \in K} L_k(i)$ and $l_k(i) = \sum_{k \in K} l_k(i)$. We get

$$\begin{aligned} \sum_{\mathrm{oc}(s)\in K} |s| &= \sum_{i=1}^{n} \sum_{s \in L_K(i)} |s| \\ &\geq \sum_{i=1}^{n} \sum_{j=1}^{l_K(i)} j \geq \sum_{i=1}^{n} \tfrac{1}{2}(l_K(i))^2. \end{aligned} \tag{9}$$

Here the first equality and the last inequality are immediate, while the first inequality follows from the fact that, for fixed i, $L_K(i)$ contains at most one switch of width 1, at most one switch of width 2, and so on. Hence, the sorted sequence of widths occurring in $L_K(i)$ dominates the sequence $1,2,\ldots,l_K(i)$.

We infer from (8) and (9) that

$$\sum_{i=1}^{n} (l_K(i))^2 < 2n \sum_{k \in K} k$$

which allows the maximum in (5) if the values $l_K(i)$ are the same for all i. Thus

$$\sum_{i=1}^{n} l_K(i) < 2^{1/2} n \Big(\sum_{k \in K} k \Big)^{1/2}$$

(The reader might prefer to make this conclusion by applying Schwarz's inequality.) By (5) we have now

Theorem 5. *Let P be an n-sequence and let $\phi \neq K \subseteq \{1,2,\ldots,\lfloor \frac{1}{2}n \rfloor\}$. Then*

$$\sum_{k \in K} g_k(P) < 2^{1/2} n \Big(\sum_{k \in K} k \Big)^{1/2}$$

and

$$\sum_{k \in K} (g_k(P) + g_{n-k}(P)) < 2^{3/2} n \Big(\sum_{k \in K} k \Big)^{1/2}$$

This completes also the proof of Theorem 1 (see Observation 4).

Remark. If $\frac{1}{2}n \notin K$ (which is always true for odd n), the bound can be improved to

$$\sum_{k \in K} (g_k(P) + g_{n-k}(P)) < 2n\left(\sum_{k \in K} k\right)^{1/2}.$$

This can be obtained if one goes through the proof directly for $\Sigma_{k \in K}(g_k(P) + g_{n-k}(P))$ instead of taking twice a bound for $\Sigma_{k \in K} g_k(P)$.

If n is even and if we use $\Sigma_{\mathrm{oc}(s)=n/2}|s| = \frac{1}{4}n^2$ (see (8)), then we can easily (along the lines above) obtain

$$g_{n/2}(P) + g_{n-n/2}(P) = 2g_{n/2}(P) < 2^{1/2}n^{3/2}.$$

Since we do not expect the bound to be asymptotically tight we decided not to elaborate on the constant.

References

1. N. Alon and E. Győri, The number of small semispaces of a finite set of points in the plane, J. Combin Theory Ser. A, to appear.
2. H. Edelsbrunner, Arrangements and Geometric Computation, forthcoming book.
3. H. Edelsbrunner and G. Stőckl, The number of extreme pairs of finite point-sets in Euclidean spaces, Institutes for Information Processing, Technical University of Graz, Austria, Report F142 (1984).
4. H. Edelsbrunner and E. Welzl, On the number of line separations of a finite set in the plane, J. Combin. Theory Ser. A 38 (1985) 15–29.
5. H. Edelsbrunner and E. Welzl, Constructing belts in two-dimensional arrangements with applications, SIAM J. Comput., to appear.
6. P. Erdős, L. Lovász, A. Simmons, and E. G. Straus, Dissection graphs of planar point sets, in A Survey of Combinatorical Theory (J. N. Strivastava et al., eds.), North Holland, Amsterdam, 1973.
7. J. E. Goodman and R. Pollack, On the combinatorial classification of nondegenerate configurations in the plane, J. Combin. Theory Ser. A 29 (1980) 220–235.
8. J. E. Goodman and R. Pollack, On the number of k-subsets of a set of n points in the plane, J. Combin. Theory Ser. A 36 (1984) 101–104.
9. G. Stőckl, Gesammelte und neue Ergebnisse über extreme k-Mengen für ebene Punktmengen, Diplomarbeit, Institutes for Information Processing, Technical University of Graz (1984).

Received April 10, 1985.