Crossing Families

Boris Aronov[1], Paul Erdős[2], Wayne Goddard[3], Daniel J. Kleitman[3],

Michael Klugerman[3], János Pach[2,4], Leonard J. Schulman[3]

Abstract

Given n points in the plane, a crossing family is a collection of line segments, each joining two of the points, such that any two line segments intersect internally. We show that any n points in general position possess a crossing family of size at least $\sqrt{n/12}$, *and describe an* $O(n \log n)$*-time algorithm for finding one.*

Introduction

Consider the following problem. You are given n points in the plane in general position (no three points are collinear). The task is to find a large collection of line segments, each joining two of the given points, such that any two line segments intersect internally. We call such a collection of segments a *crossing family*. A natural variation on this is the *two-colored* version: the n points are divided into two equally-sized color classes, and each segment must join points of distinct colors.

A natural approach to this problem is to look for a large convex set. As a well-known result of the second author and Szekeres [4] states that in any set of n points there is a convex subset of size $\log_2 n$, we are thus guaranteed a crossing family of size $\lfloor(\log_2 n)/2\rfloor$ in the uncolored case. It is also possible, with some ingenuity, to produce a family of size $\Omega(\log n)$ when the red and blue points each form a convex set of size $\log n$.

In this paper we show that $\Omega(\sqrt{n})$-size crossing families exist in both the colored and uncolored versions of the problem. Our proof is constructive, and can be implemented to find such a family in time $O(n \log n)$.

We will use the following terminology. We say that two disjoint sets A and B can be *crossed* if there exists a crossing family in which each line segment connects a point in A and one in B, and where all the points in the two sets are used. Also, we say that A *avoids* B if no line (not line segment) subtended by a pair of points in A intersects the convex hull of B. A and B are *mutually avoiding* if A avoids B and B avoids A.

We will show (cf. Corollary 1) that if equal-sized A and B are mutually avoiding then they can be crossed. Indeed, our strategy for constructing crossing families is to find a pair of large mutually avoiding sets. Though we have not determined what the true size of a maximum crossing family is, the results of Section 1 suggest that this condition of mutual avoidance is too strong.

We draw the reader's attention to two results on collections of m line segments among n (uncolored) points in the plane. Alon and the second author [1] showed that if $m \geq 6n-5$ then there are always three mutually disjoint line segments. Capoyleas and the sixth author [2] showed that for $k \leq n/2$ if the points are in convex position and $m > (k-1)(2n+1-2k)$, then there is a crossing family of size k, and that this is best possible.

We will proceed as follows. In the first section we discuss some characterizations of crossing families and related objects. In Section 2 we show how to construct a pair of mutually avoiding sets of size $\Omega(\sqrt{n})$. Upper bounds on the size of a crossing family are considered in Section 3. In Section 4 we examine "parallel" families of line segments: two line segments are parallel if the lines they subtend intersect outside both line segments.

[1] Department of Computer Science, Polytechnic University

[2] Mathematical Institute of the Hungarian Academy of Sciences

[3] Department of Mathematics, Massachusetts Institute of Technology

[4] Courant Institute, New York University

[5] Research supported in part by DARPA grant N00014-89-J-1988, Air Force grant AFOSR-89-0271, NSF grant DMS-8606225, and an ONR graduate fellowship. Further, part of this work was conducted at and supported by DIMACS (Center for Discrete Mathematics and Theoretical Computer Science), a National Science Foundation Science and Technology Center – NSF-STC88-09648.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

© 1991 ACM 0-89791-426-0/91/0006/0351 $1.50

Finally in Section 5 we describe another approach to constructing mutually avoiding sets, which generalizes to higher dimensions.

1 A Characterization

In this section we examine conditions which characterize when two sets can be crossed.

Consider red points X and blue points Y separated by a line $\mathcal{L}$. We say a red point x sees a blue point y at rank i if y is the i^{th} blue point counterclockwise as seen from x. And vice versa. Then we say X and Y obey the *rank condition* if there exist labelings $x_1, \ldots, x_s$ and $y_1, \ldots, y_s$ of X and Y such that for all i, x_i sees y_i at rank i and vice versa. For the *strong rank condition*, the labelings must be such that x_i sees y_j at rank j for all i and j. For the *weak rank condition*, we require only that x_i sees y_i at the same rank that y_i sees x_i.

Theorem 1 *Let X and Y be s red and s blue points separated by a line. Then:*
(1) The weak rank condition is equivalent to the rank condition.
(2) X and Y can be crossed if and only if they obey the rank condition.
(3) X and Y are mutually avoiding if and only if they obey the strong rank condition.

Proof: Assume for ease of description that the line $\mathcal{L}$ is vertical, with reds X on the left and blues Y on the right.

(1) Let X and Y obey the weak rank condition. We claim that the labeling must be such as to satisfy the rank condition. To establish this, it is sufficient to show that, if a_i denotes the rank at which x_i and y_i see each other, then all the a_i are distinct. Consider any i and j, and let ℓ_i and ℓ_j be the lines induced by $x_i y_i$ and $x_j y_j$. Say these lines intersect to the left of $\mathcal{L}$, and that ℓ_i has greater slope than ℓ_j. Then there are exactly $a_i - 1$ blue points below ℓ_i. But this includes y_j, so there are at most $a_i - 2$ blue points below ℓ_j and thus $a_j < a_i$.

(2) Assume first that X and Y can be crossed. Let $l_1, \ldots, l_s$ denote the line segments of a complete crossing family in order of increasing slope. Label the red endpoint of l_i, x_i, and the blue endpoint y_i. Since $l_1, \ldots, l_{i-1}$ are of lesser slope than l_i, and intersect it, x_i sees $y_1, \ldots, y_{i-1}$ before it sees y_i. Similarly, x_i sees $y_{i+1}, \ldots, y_s$ after y_i and thus it sees y_i at rank i. For the same reason, y_i sees x_i at rank i.

Assume now that there exist such labelings. We prove by induction on s that the family $\{x_i y_i\}_i$ of line segments is a crossing family. The case $s = 1$ is trivial.

Consider the line segment $x_s y_s$ subtending line ℓ_s. By the rank conditions, $X - x_s$ lies above this line, while $Y - y_s$ lies below it. This implies that:

> **Observation:** *$x_s y_s$ crosses the line segment $x_i y_i$ iff it crosses the line subtended by $x_i y_i$.*

It also implies that the slope of $x_i y_i$ is less than that of $x_s y_s$ for all $i < s$.

Let A be the set of all line segments that the line segment $x_s y_s$ does not cross, and order them with respect to their $\mathcal{L}$-intercepts. *Suppose* that A contains a line segment $x_a y_a$ whose $\mathcal{L}$-intercept is above $x_s y_s$; then choose such a such that $x_a y_a$ has the highest $\mathcal{L}$-intercept. See Figure 1. Note that the line ℓ_a does not intersect the line segment $x_s y_s$.

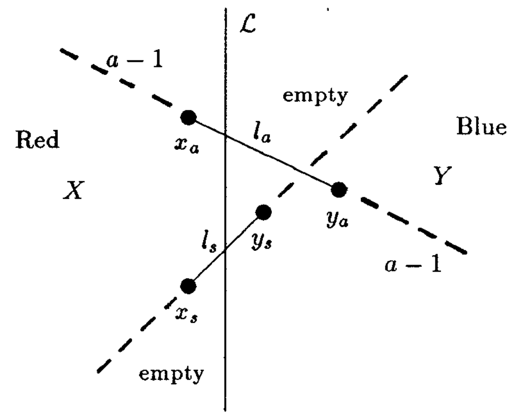

Figure 1: The picture when $x_a y_a$ is above $x_s y_s$

So there exist $a - 1$ red points above ℓ_a and $a - 1$ blue points below it. But y_s counts while x_s doesn't. Therefore there exists b such that x_b and y_b are *both* above ℓ_a. Then $x_b y_b$ does not cross $x_s y_s$ so, by the observation, neither does ℓ_b, the line through these two points. In fact, $x_s y_s$ is also below ℓ_b while the intercept of ℓ_b with $\mathcal{L}$ is above that of ℓ_a, a contradiction.

Similarly A contains no line segment whose $\mathcal{L}$-intercept is below $x_s y_s$ and thus we have established that $x_s y_s$ crosses all the other line segments. Since the slope of $x_i y_i$ $(i < s)$ is less than that of $x_s y_s$, this implies that every red point x_i sees y_s after y_i, and vice versa. Hence the deletion of x_s and y_s preserves the rank condition for $x_1, \ldots, x_{s-1}$ and $y_1, \ldots, y_{s-1}$. Therefore, by the induction hypothesis, $x_i y_i$ intersects $x_j y_j$ for $1 \leq i < j \leq s-1$.

(3) Assume first that X and Y are mutually avoiding. Let $x \in X$ be arbitrary, and let $y_1, \ldots, y_s$ be the ranking of points as seen from x. We claim that every point in X sees Y the same way. For, move x' smoothly along a straight line from x to x_i; if at some stage the picture of Y were to change, then x' would have to cross a line subtended by a pair of points in Y. Similarly, all of Y sees X the same way.

Assume second that X and Y are not mutually avoiding. Say X does not avoid Y. Then there is a line ℓ, subtended by two points x_i and x_j in X, such that two points y_k and y_l in Y are on either side of ℓ. But then y_k and y_l see x_i and x_j in a different order. □

Despite the above characterizations, we are only able to exploit:

Corollary 1 *If X and Y of equal cardinality are mutually avoiding then they can be crossed.*

2 Construction of an $\Omega(\sqrt{n})$ Crossing Family

The game-plan here is to find a pair of mutually avoiding sets X' and Y' of size $\Omega(\sqrt{n})$. This is achieved by finding red and blue subsets X and Y such that X avoids Y, and then finding subsets $X' \subseteq X$ and $Y' \subseteq Y$ such that Y' avoids X'. By Corollary 1 we thus obtain a crossing family of cardinality $\min\{|X'|, |Y'|\}$.

We will use the following well-known results:

Lemma 1 *For any line $\mathcal{L}$ in the plane and finite set of points, it is possible to find another line $\mathcal{M}$ which simultaneously splits the points in both halfplanes in any desired proportions.*

Lemma 2 [4] *For any sequence of real numbers of length n, there is either an ascending or a descending subsequence of length $\sqrt{n}$.*

Our strategy has three steps. We assume that there are $n/2$ red and $n/2$ blue points.

Step 1

Step 1 is a preliminary one where the plane is partitioned by three lines (as depicted in Figure 2) so that particular regions have linearly many points of a particular color.

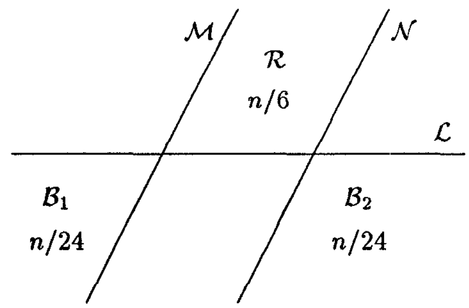

Figure 2: The H-picture

Specifically, first find a line $\mathcal{L}$ such that at least $n/4$ of the reds are on one side and at least $n/4$ blues on the other by moving a horizontal $\mathcal{L}$ down from $y = +\infty$ until $n/4$ of the first color, say red, are above it. Discard the blue points above $\mathcal{L}$ and the red points below it. Second, use Lemma 1 to find a line $\mathcal{M}$ such that exactly $n/24$ of the red and $n/24$ of the blue points are to the left of $\mathcal{M}$. Finally, take a line $\mathcal{N}$ parallel to $\mathcal{M}$ at $x = +\infty$ and move it to the left until $n/24$ of the first color, say blue, are on its right. See Figure 2. Note that the region $\mathcal{R}$ contains at least $n/6$ red points, and the regions $\mathcal{B}_1$ and $\mathcal{B}_2$ both at least $n/24$ blue points.

Step 2

For convenience, apply an affine transformation such that $\mathcal{M}$ and $\mathcal{N}$ are vertical. Order the reds in $\mathcal{R}$ from left to right. By Lemma 2 there exists either an ascending or a descending subsequence R of length $\sqrt{n/6}$. Without loss of generality assume that R is *descending*. Then observe that R avoids $\mathcal{B}_1$.

Step 3

Consider the middle point x of R breaking it into two parts R_1 and R_2, each a descending sequence of length $\sqrt{n/24}$. See Figure 3.

Consider the positions of the blue points in $\mathcal{B}_1$ expressed in polar coordinates (r, θ) with x as the origin (and θ measured counterclockwise), and order them as $\{b_i\}$ (for $i = 1, \ldots, n/24$) in decreasing distance r_i from

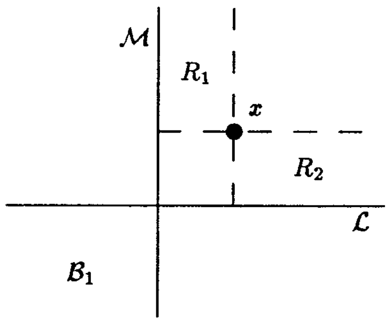

Figure 3: x splits R into two parts

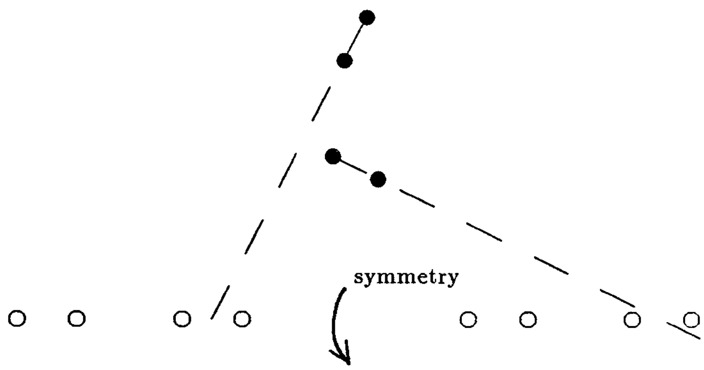

Figure 4: Arrangement for upper bound for two-colored case: 4 more black points are located symmetrically below the whites

x. By Lemma 2 there exists a subsequence $B = \{b_{k_i}\}$ (for $i = 1, \ldots, \sqrt{n/24}$) whose angles θ_{k_i} are either decreasing or increasing. Say they are *increasing*. We claim that B avoids R_1. For consider two points b_{k_i} and b_{k_j} of B with $i < j$. From the conditions on B it follows that b_{k_j} is to the right of b_{k_i} and below the line subtended by x and b_{k_i}. Thus the line spanned by b_{k_i} and b_{k_j} avoids the region containing R_1.

Applying Corollary 1 we obtain:

Theorem 2
(i) Given $n/2$ red and $n/2$ blue points, there exists a crossing family of size at least $\sqrt{n/24}$.
(ii) Given n uncolored points, there exists a crossing family of size at least $\sqrt{n/12}$.

Proof: We proved the two-colored case above. The only change for the uncolored case is that $\mathcal{L}$ may be found without discarding half the points. □

Note that the above procedure provides a polynomial-time algorithm. In fact, it can be performed in $O(n \log n)$ time. One need only verify that one can apply Lemmas 1 and 2 in this time, but this is well-known. (For Lemma 1 cf. [3].)

3 Upper Bounds

The best upper bounds we have found are linear. For the uncolored case we get a 1/2 crossing ratio (points used/points total) with four non-convex points. For the two-colored case, we can get a 3/8 crossing ratio. Consider, for example, the arrangement X of sixteen points described in Figure 4.

That these ratios can be obtained for arbitrarily large n is shown by the following. Consider splitting each point $x \in X$ almost vertically into two nearby twins x^1 and x^2. This is performed so that, for any i, j, k, l, xy crosses zt iff $x^i y^j$ crosses $z^k t^l$. This process yields a new arrangement X'.

Lemma 3 *The crossing ratio for X' is the same as that for X.*

Proof: Consider first a crossing family $\mathcal{F}$ on X. Then every line segment xy in $\mathcal{F}$ may be replaced with two intersecting line segments matching x's and y's replacements. Thus the crossing ratio for X' is at least that for X.

Consider now a crossing family $\mathcal{F}$ on X'. Let G be the graph on vertex set X such that x and y are adjacent iff $x^i y^j \in \mathcal{F}$ for some i and j. Then G has maximum degree 2.

G is obviously bipartite when X is two-colored, but it also bipartite even if X is not colored: as each pair of twins x^1 and x^2 cannot have line segments going both to the right and to the left, the vertices of G are two-colorable via "left" and "right". Now since G has maximum degree 2, it has a matching $\mathcal{M}$ of cardinality at least $|\mathcal{F}|/2$. But $\mathcal{M}$ represents a crossing family for X. □

4 Parallel Families

Recall that two line segments are *parallel* if the lines they subtend intersect outside both line segments. Parallel families have several similarities with crossing families. For example, the following is established using the methods of Section 1.

Theorem 3 *Let X and Y be two sets, of s points each, separated by a line. Then X and Y can be paired up to form a parallel family iff there exist labelings $x_1, \ldots, x_s$ and $y_1, \ldots, y_s$ of X and Y such that for all i, x_i sees y_i at rank i and y_i sees x_i at rank $s+1-i$. In particular, if X and Y are mutually avoiding then X and Y can be so paired.*

However the most interesting fact is that the parallel and crossing family problems are equivalent.

Theorem 4 *Let $c(n)$ (resp. $p(n)$) denote the minimum number of segments in a maximum crossing (parallel) family, among all configurations of n blue and n red points separated by a line, with all $2n$ points in general position. Then $c(n) = p(n)$.*

Proof: We use polarization arguments [5]. Recall that polarization in a circle of radius r about the origin $\bar{o}$ takes a point $p \neq \bar{o}$ to the line $\ell = \{ z \in \mathbf{R}^2 : z \cdot p = r^2 \}$; while taking ℓ to that point p. (Note that ℓ intersects, and is perpendicular to, the ray $\vec{op}$). Further recall that any segment of a line ℓ is carried to a double wedge not containing the origin, namely an interval of lines all passing though the polar point of ℓ. Further, two segments cross iff each polar double wedge contains the center of the other, and they are parallel iff neither polar double wedge contains the center of the other.

Let B be a set of n blue points, R a set of n red points, and Y a vertical line separating them with B on the right. We show how to construct a pair of sets B' and R' of n blue and n red points in general position separated by a line X, and a map s carrying B to B' and R to R' bijectively, such that if segments b_1r_1 and b_2r_2 ($b_1, b_2 \in B, r_1, r_2 \in R$) intersect then segments $s(b_1)s(r_1)$ and $s(b_2)s(r_2)$ are parallel, while if b_1r_1 and b_2r_2 are parallel then $s(b_1)s(r_1)$ and $s(b_2)s(r_2)$ intersect. Our theorem will thus follow.

Let o_1 be a point on Y far below $B \cup R$ and let X be the line perpendicular to Y through o_1. Now let s_1 be the polarization map in a circle centered at o_1. Then s_1 carries points of B to lines with negative slopes and positive Y-intercepts, and points of R to lines with positive slopes and positive Y-intercepts. Let o_2 be a point on X which is to the right of the X-intercepts of all lines $s_1(b)$ for $b \in B$. Let s_2 be the polarization map in a circle centered at o_2. Then s_2 carries lines of $s_1(B)$ to points below X and lines of $s_1(R)$ to points above X.

Let $s = s_2 \circ s_1$. Note that o_1 lies outside every double wedge $s_1(b)s_1(r)$ (for $b \in B, r \in R$) while o_2 lies inside every such double wedge. Therefore s maps every line segment br to the complement of the line segment $s(b)s(r)$. Thus, s takes intersecting segments to intersecting complements of segments (i.e. the corresponding segments are parallel), and conversely. □

5 Another Approach to Mutually Avoiding Sets

In this section we describe another approach to producing a pair of polynomial-size mutually avoiding sets. Unfortunately, this approach yields only sets of size $\Omega(n^{1/3})$.

We say that a line *stabs* a set if it intersects the convex hull of that set. The *stabbing number* of a collection of sets is the maximum number of sets that any line stabs.

We will need the following special case of a result of Matoušek:

Lemma 4 [5] *Let P be a set of n points in the plane and let $r \leq n$. Then there exists a subset $P' \subseteq P$ of at least $n/2$ points and a partition $\{P_1, \ldots, P_m\}$ of P' with $|P_i| = \lfloor n/r \rfloor$ for all i and with stabbing number $O(r^{1/2})$.*

Theorem 5 *Any set P of n points in the plane contains a pair of mutually avoiding subsets each of size $\Omega(n^{1/3})$.*

Proof: Say we apply the above result with parameter r yielding a partition of P' into subsets $P_1, \ldots, P_m$. Note that m is $\Theta(r)$. Each P_i generates $O(n^2/r^2)$ lines, and each of these lines stabs $O(r^{1/2})$ other subsets. Thus there are at most $O(n^2/r^{1/2})$ stabbings in all. As there are $\Theta(r^2)$ pairs of subsets, some pair has $O(n^2/r^{5/2})$ mutual stabbings. If r is such that this quantity is slightly less than 1, then there exists a pair of mutually avoiding sets, and these are of size $n/r = \Omega(n^{1/5})$.

We can, however, use the deletion method to achieve a better result. Let $r \approx (2n)^{2/3}$ so that some two subsets P_i, P_j have at most $n/2r$ mutual stabbings. Each stabbing is created by a line which can be eliminated by removing one of the points from P_i or P_j (as the case may be). The depleted versions of P_i and P_j are each of size at least $n/2r = \Omega(n^{1/3})$ and are mutually avoiding. □

Minor changes suffice to handle the two-colored case.

It may be that further progress can be made on the problem using this approach (if $\Theta(\sqrt{n})$ is not the correct answer). In this regard note that Matoušek's result itself is essentially optimal.

An advantage of this approach, where one partitions the points and counts the total stabbings, is that it can be extended to higher dimensions. In dimension d a crossing family is a collection of d-simplices such that every two simplices intersect and have disjoint vertex sets. An inductive argument shows the existence of n^{ϵ}-sized crossing families in any fixed dimension d, but the values of ε are unspectacular.

6 Comments

The most obvious open question is whether a linear-sized crossing family always exists. For some special cases this is true; for example, if the input is $n/2$ reds and $n/2$ blues, each a convex set. One can also show that, given n (colored or not) random points in the unit circle say, almost surely a linear-sized crossing family exists. We omit the details in both cases.

References

[1] N. Alon and P. Erdős, Disjoint edges in geometric graphs, *Discrete Comput. Geom.* 4 (1989), 287–290.

[2] V. Capoyleas and J. Pach, A Turán-type theorem on chords of a convex polygon, DIMACS Technical Report 90-30, 1990, to appear in *J. Combin. Theory Ser. A.*

[3] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, Springer-Verlag, Berlin, 1987.

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

[5] J. Matoušek, Efficient Partition Trees, *These Proceedings.*

[6] E.G. Rees, *Notes on Geometry*, Springer-Verlag, Berlin, 1983.