Connected Matchings*

Oswin Aichholzer[1], Sergio Cabello[2], Viola Mészáros[3], and Jan Soukup[4]

1 **Institute of Software Technology, Graz University of Technology, Graz, Austria**
`oaich@ist.tugraz.at`
2 **Faculty of Mathematics and Physics, University of Ljubljana, Slovenia**
Institute of Mathematics, Physics and Mechanics, Slovenia
`sergio.cabello@fmf.uni-lj.si`
3 **Bolyai Institute, University of Szeged, Szeged, Hungary**
`meszaros.viola@gmail.com`
4 **Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic**
`soukup@kam.mff.cuni.cz`

Abstract

We show that each set of $n \geqslant 2$ points in the plane in general position has a straight-line matching with at least $(5n+1)/27$ edges whose segments form a connected set, while for some point sets the best one can achieve is $\lceil \frac{n-1}{3} \rceil$.

1 Introduction

Consider a set P of n points in the plane in general position, meaning that no three points of P are collinear. A (straight line) *matching* M for P is a set of segments with endpoints in P such that no two segments share an endpoint. A matching M for P is *connected* (via their crossings) if the union of the segments of M forms a connected set. Equivalently, a matching is connected when the intersection graph of its segments is connected. The *size* of the matching M is the number of edges (or segments) in M. In this paper, we study the following problem.

▶ **Question 1.1** (Connected Matching). *Find for each n the largest value $f(n)$ with the following property: each set of n points in general position in the plane has a connected matching with $f(n)$ edges.*

We provide upper and lower bounds for the function $f(n)$. Our upper bounds are constructive and lead to effective algorithms to compute the connected matching. In this short version we focus on our constructions. Missing details and proofs of the algorithmic claims will appear in the full paper, where we also consider a colored version of the problem.

The problem can be seen as a variant of the problem on crossing families of Aronov et al. [3], where one wants to find as many segments as possible with endpoints in P such

* Research on this work has been initiated at the 17th European Geometric Graph Week which was held from August 15th to 19th in Leipzig. We thank all participants for the good atmosphere as well as for inspiring discussions on the topic. O. A. supported by FWF grant W1230. Research funded in part by the Slovenian Research and Innovation Agency (P1-0297, J1-2452, N1-0218, N1-0285). Research funded in part by the European Union (ERC, KARST, project number 101071836). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. J. S. supported by the grant SVV–2023–260699.

40th European Workshop on Computational Geometry, Ioannina, Greece, March 13–15, 2024.
This is an extended abstract of a presentation given at EuroCG'24. It has been made public for the benefit of the community and should be considered a preprint rather than a formally reviewed paper. Thus, this work is expected to appear eventually in more final form at a conference with formal proceedings and/or in a journal.

that any pair of segments crosses in their interior. While in our setting we are asking for a connected subgraph in the intersection graph of the segments, the crossing families problem asks that the intersection graph is a complete graph. The best lower bound, showing an almost linear lower bound for crossing families, has been a recent breakthrough by Pach, Rubin and Tardos [5]. Aichholzer et al. [2] have the currently best upper bound.

2 Balanced separation with a short path

In this section we provide a structural result about splitting the convex hull of a point set with a single edge or with a 2-edge path in such a way that both sides contain a large fraction of the point set. A very similar result can be found in Ábrego and Fernández-Merchant [1, Lemma 2]. We include a proof because their bound has a small error[1], our approach is different in the treatment of the triangular case (Theorem 2.1), and we discuss the algorithmic counterpart, a part that is not considered in [1] and that forces us to rework a proof.

We first consider the case when the convex hull is a triangle and the partition can be with different number of points. This will be a tool for the general case. See Figure 1.

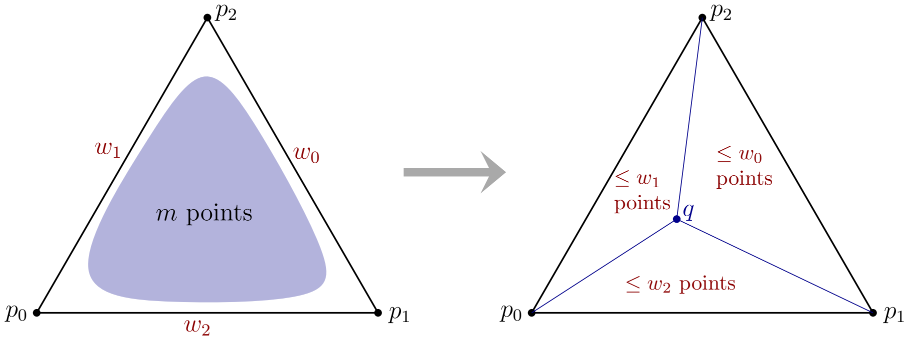

Figure 1 Statement in Theorem 2.1.

▶ **Theorem 2.1.** *Assume that we have a triangle with vertices p_0, p_1 and p_2 and in its interior there is a set P of $m \geqslant 1$ points such that $P \cup \{p_0, p_1, p_2\}$ is in general position. For any integer weights w_0, w_1, w_2 such $0 \leqslant w_0, w_1, w_2 < m$ and $\ell := w_0 + w_1 + w_2 > 2m - 3$, there exist at least $\ell - 2m + 3 > 0$ points $q \in P$ such that, for each $i \in \{0, 1, 2\}$, the triangle $\triangle(p_i q p_{i+1})$ contains at most w_{i+2} points of P in its interior, where all indices are modulo 3.*

We can find $\ell - 2m + 3$ points with this property in linear time.

Proof. In this proof, all indices are modulo 3. For $i \in \{0, 1, 2\}$, consider a ray r_i that starts at p_{i-1} and goes through p_i. We rotate r_i around p_{i-1} in the direction towards p_{i+1} until we pass r_i over $m - w_i - 1$ points of P. See Figure 2, left, to visualize the case $i = 1$. For any of the points $q \in P$ we did not scan over, the triangle $\triangle(p_{i-1} q p_{i+1})$ contains at most w_i points of P in its interior; note that q is not in the interior of $\triangle(p_{i-1} q p_{i+1})$.

Some points of P may be scanned more than once, but in total we scan at most $3m - w_1 - w_2 - w_3 - 3 = 3m - \ell - 3$ points. So there are at least $m - (3m - \ell - 3) = \ell - 2m + 3 > 0$ points remaining, and each of them satisfies the desired property. ◀

As a special case we state the following corollary, which might be of its own interest.

[1] Lemma 2 in [1] is not correct for $n = 4$.

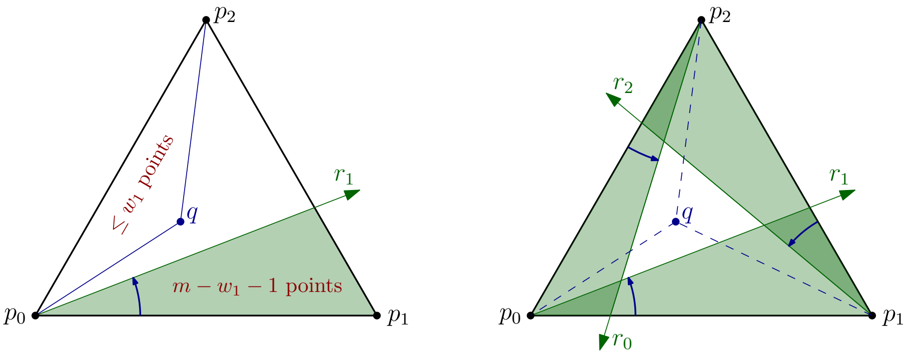

Figure 2 Left: rotating r_1 until we pass over $m - w_1 - 1$ points. Right: the part of the triangle that is not shadowed contains at least $\ell - 2m + 3$ points.

▶ **Corollary 2.2.** *Let Δ be a triangle with a set P of $m \geqslant 1$ points in its interior. Then there is a point of P that splits Δ into three triangles, such that none of these parts contains more than $\lceil (2m-2)/3 \rceil$ points of P in its interior.*

This result resembles the classical Centerpoint theorem [4, Section 1.4], which tells that for each set P of n points in the plane there exists a so-called centerpoint q with the property that each open halfplane that does not contain q has at most $2n/3$ of the points of P inside. However, the centerpoint does not need to be a point of P, and for some point sets it cannot be an element of P.

Denote by $CH(P)$ the convex hull of P. A point $p \in P$ is *extremal* for P if it lies on the boundary of $CH(P)$. A *k-separating path* for P is a plane path π spanned by vertices of P and connecting two different extremal points of P such that $CH(P) \setminus \pi$ has two parts, each containing at least k points; the points on the path are counted in no part. See Figure 3. The *length* of such a path is its number of edges.

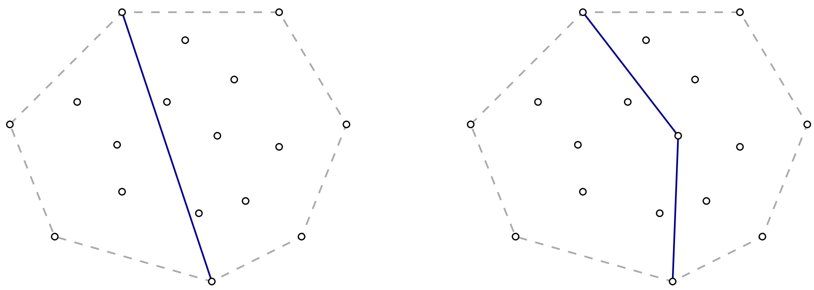

Figure 3 Left: 5-separating path of length 1. Right: 7-separating path of length 2.

▶ **Theorem 2.3.** *Let P be a set of $n \geqslant 2$ points in general position in the plane. There exists a $\lceil \frac{n-4}{3} \rceil$-separating path for P of length 1 or 2 and it can be found in time linear in n.*

Proof sketch. For $n \leqslant 4$ the statement is trivially true. So for the reminder of the proof assume that $n \geqslant 5$. Let us set $r = \lceil (2(n-3)-2)/3 \rceil = \lceil (2n-8)/3 \rceil$. The intuition is that r is the bound of Corollary 2.2 for $n - 3 \geqslant 1$ points; in our current setting, n is also counting the vertices of the triangle. We also set $k = \lceil (n-4)/3 \rceil \geqslant 1$ as $n \geqslant 5$.

Choose an extremal point $q_0 \in P$ with the smallest y-coordinate. Let $q_1, \ldots, q_{n-1}$ be the points $P \setminus \{q_0\}$ sorted increasingly by the angle $\overline{q_0 q_i}$ makes with the horizontal rightward ray from q_0. See Figure 4, left.

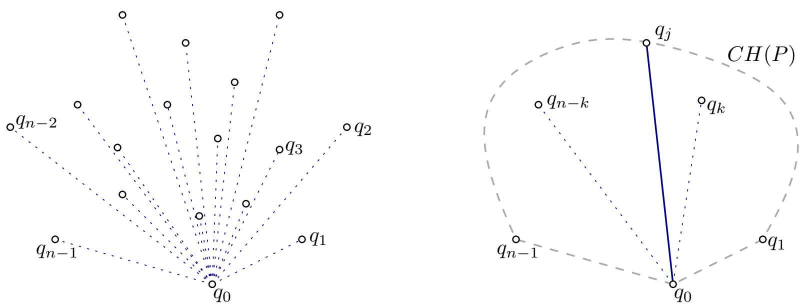

Figure 4 Proof of Theorem 2.3.

If between q_k and q_{n-k} there is some extremal point q_j for P, which implies that $k < j < n-k$, then the segment q_0q_j is a k-separating path of length 1 and we are done. See Figure 4, right. Otherwise, the rays q_0q_k and q_0q_{n-k} intersect the same edge e of $CH(P)$. Let q_aq_b be the edge e, with $a < b$. This means $a \leqslant k < n-k \leqslant b$ and the triangle $\triangle(q_0q_aq_b)$ has exactly $b-a-1$ points in its interior. See Figure 5, left.

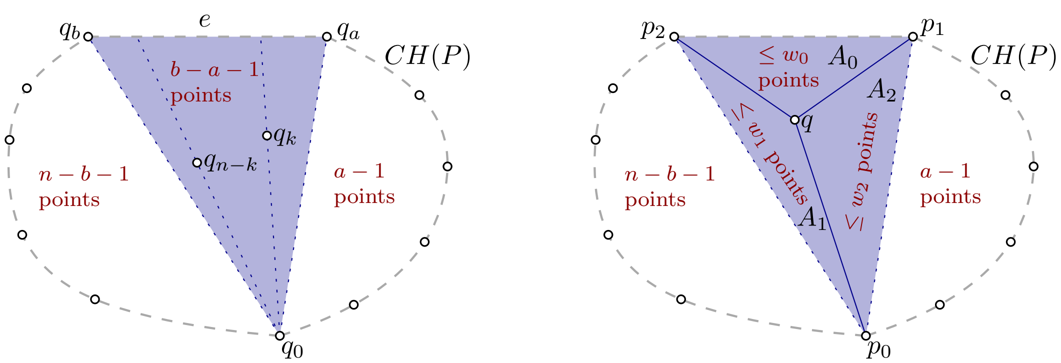

Figure 5 Continuation of the proof of Theorem 2.3.

We want to apply Theorem 2.1 to $\triangle(q_0q_aq_b)$ and the $m = b-a-1 \geqslant (n+1)/3 \geqslant 1$ points of P in its interior. To this end, set $p_0 = q_0$, $p_1 = q_a$, $p_2 = q_b$, $w_0 = r$, $w_1 = r-(n-b-1)$, and $w_2 = r-(a-1)$. See Figure 5, right. After checking that indeed $w_0+w_1+w_2 > 2m-3$, which we skip, Theorem 2.1 implies the existence of a point $q \in P$ in the interior of $\triangle(p_0p_1p_2) = \triangle(q_0q_aq_b)$ that splits it into three triangular pieces such that the interior of the triangle $\triangle(p_{i-1}qp_{i+1})$ has at most w_i points of P (for $i = 0,1,2$ and indices modulo 3).

We split $CH(P)$ into three parts A_0, A_1, A_2 by removing the segments $qq_0 = qp_0$, $qq_a = qp_1$ and $qq_b = qp_2$. See Figure 5, right. The points q, q_0, q_a, q_b belong to no part, while all the other points of P belong to exactly one part. From the choices of weights w_i, each part contains at most r points of P. Any part among A_0, A_1, A_2 with most points has at least $\lceil (n-4)/3 \rceil = k$ points and its boundary defines a k-separating path of length 2. ◀

3 Upper bound

Consider n points split into three sets A_0, A_1, A_2 of size $\sim \frac{n}{3}$, where each A_i lies on its own slightly curved blade of a three-bladed windmill; see Figure 6. We use indices modulo three in the discussion. We can form such a configuration so that each line determined by two points of A_i separates A_{i+1} from A_{i+2}, and no segment connecting one point of A_i with one point of A_{i+1} crosses any segment connecting two points of A_{i+1}. Hence, the set of all segments is separated into three parts where each part consists of segments connecting two points of A_i or one point of A_i and one point of A_{i+1}, and segments from different parts do not cross. Clearly, the size of the largest matching spanning $A_i \cup A_{i+1}$, if their sizes differ by at most one, is $\min\{|A_i|, |A_{i+1}|\}$, and the largest of those values over $i \in \{0, 1, 2\}$ gives the largest connected matching. Treating carefully the modulus of n, we get for each $n \geqslant 1$ a point set where the maximum connected matching has size $\lceil \frac{n-1}{3} \rceil$.

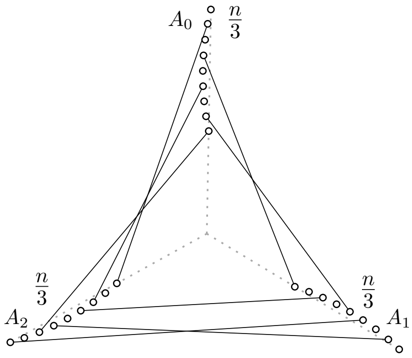

Figure 6 Upper bound for connected matchings.

4 Lower bound

We first consider the following special setting, depicted in Figure 7, left.

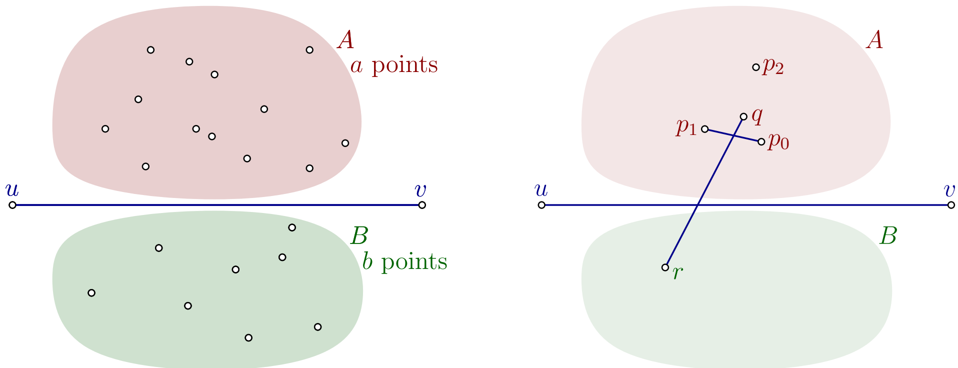

Figure 7 Left: Situation in Lemma 4.1. Right: edges added to the matching when A has four points not in convex position.

▶ **Lemma 4.1.** *Assume that we have a horizontal segment uv, a set A of a points above the line supporting uv, and a set B of $b \leqslant a$ points below the line supporting uv such that, for all $(a, b) \in A \times B$, the segment ab intersects uv, and $A \cup B \cup \{u, v\}$ consists of $a + b + 2$ points in general position. Then, $A \cup B \cup \{u, v\}$ has a connected matching of size at least*

$$m(a,b) := \begin{cases} 1+b & \text{if } b \leqslant a \leqslant 2b+3, \\ (a+3b+2)/5, & \text{if } 2b+3 \leqslant a \leqslant 7b+3, \\ 1+2b, & \text{if } a \geqslant 7b+3. \end{cases}$$

Such a connected matching can be computed in $O(1 + a \log a)$ *time.*

Proof sketch. We first make two easy observations that will come in handy:

(a) A matching of B onto A with b edges together with the edge uv to “connect” them is a connected matching of size $b + 1$. We want to improve upon this when the sides are unbalanced, in particular when a is larger than $2b \pm O(1)$.
(b) If A has a large subset A' in convex position, then we can get a connected matching of size $\lfloor \frac{|A'|}{2} \rfloor$, for example by connecting “antipodal” points along the boundary of $CH(A')$.

We construct a connected matching M iteratively as follows. At the start we add uv to M. While $|A| > |B| > 0$ and A has four points p_0, p_1, p_2, q such that q is in the interior of $\triangle(p_0p_1p_2)$, we take an arbitrary point $r \in B$, add the edge qr to M, and to M the edge pp' of $\triangle(p_0p_1p_2)$ crossed by qr. See Figure 7, right. Note that $\{pp', uv, qr\}$ is a connected matching. Then we remove p, p', q from A, and r from B. With each repetition of this operation, we increase the size of the matching by two, remove three points from A, and remove a point from B. We repeat this operation until B is empty, $|A| \leqslant |B|$, or A is in convex position, whatever happens first. Let k be the number of repetitions of this operation, let A' and B' be the subsets of A and B, respectively, that remain at the end. Therefore, M is a connected matching with $1 + 2k$ edges, A' has $a - 3k$ points, and B' has $b - k$ points.

We now consider the different conditions that hold at the end:

- If we finish because B' is empty, then $k = b$ and the matching M has $1 + 2b$ edges.
- If we finish because $|A'| \leqslant |B'|$, we match the remaining points of A' to B' arbitrarily and add those edges $|A'|$ to M; since they cross uv, M keeps being a connected matching. Using that the cardinality of A decreases at steps of size 3 and the cardinality of B decreases at steps of size 1, it is possible to show that the size of the connected matching M is in this case $1 + \lfloor (a + b)/2 \rfloor$.
- If we finish because A' does not have any 4 points with the desired condition, the key observation is to note that A' is in convex position. (This is also true if $|A'| \leqslant 3$.) We consider two connected matchings and take the best of both.
  The first matching is obtained by adding to M a matching between all the vertices of B' and any subset of A' with $|B'|$ points. The second matching, which we denote by M', is obtained by taking a connected matching of the points A'; they are in convex position.
  A comparison between M and M' shows that the larger one has size at least

$$\begin{cases} 1 + b & \text{if } b \leqslant a \leqslant 2b + 3, \\ (a + 3b + 2)/5, & \text{if } 2b + 3 \leqslant a \leqslant 7b + 3, \\ (a - 3b - 1)/2, & \text{if } a \geqslant 7b + 3. \end{cases}$$

Since we have given a construction that can finish with 3 different conditions, one has to show that in each scenario $m(a, b)$ is a lower bound on the size of the connected matching. We skip this computation. ◀

Note that the bound $m(a, b)$ of Lemma 4.1 is monotone increasing in a and in b, also when we take a and b as real values (with $b \leqslant a$ always.) Moreover, when $a + b$ is kept constant, $m(a, b)$ is larger for larger b. This means that $m(a, b) \leqslant m(a-1, b+1)$, if $b \leqslant a-2$.

▶ **Theorem 4.2.** *Let* P *be a set of* $n \geqslant 2$ *points in general position in the plane. Then* P *has a connected matching set of size at least* $(5n + 1)/27$ *and it can be computed in* $O(n \log n)$ *time.*

Proof sketch. By Theorem 2.3 we know that there is a $\lceil \frac{n-4}{3} \rceil$-separating path P of length 1 or 2 for P. Let A and B be the sets of points of P on each side of π, such that $|A| \geqslant |B|$. Note that the vertices of π do not go to any of the sides, which means that $n-3 \leqslant |A|+|B| \leqslant n-2$ and $\lceil \frac{n-4}{3} \rceil \leqslant |B| \leqslant |A|$. Each edge connecting a point of A to a point of B crosses π.

If π consists of a single edge e, then we match all points of B to points of A arbitrarily, and include e also in the matching. Since all these edges intersect e, they form a connected matching of size $1 + |B| \geqslant \lceil \frac{n-1}{3} \rceil \geqslant \frac{5n+1}{27}$. (This last inequality holds for $n \geqslant 2$.)

For the remainder of this proof we assume that π has length two, and denote its edges by e_1 and e_2. We build a *maximal* matching M_1 from $B_1 \subseteq B$ to $A_1 \subseteq A$ with edges that cross e_1. This means that $|A_1| = |B_1|$ and there is no point in $A \setminus A_1$ that can be connected to a point in $B \setminus B_1$ by crossing e_1. Set $A_2 = A \setminus A_1$ and $B_2 = B \setminus B_1$. Each segment connecting a point in A_2 to a point of B_2 must cross e_2 because it does not cross e_1. We make an arbitrary matching M_2 connecting each point of B_2 to points of A_2; this can be done because $|B_2| = |B| - |M_1| \leqslant |A| - |M_1| = |A_2|$. We add e_1 to M_1 and e_2 to M_2 so that M_1 and M_2 become connected matchings with $|M_1| + |M_2| = 2 + |B|$.

If M_1 or M_2 has size at least $\frac{5n+1}{27}$, then we are done. Therefore, we restrict our attention to the case when $|M_1|, |M_2| \leqslant \frac{5n+1}{27}$. Since $|A_1| = |B_1| = |M_1| - 1 \leqslant \frac{5n-26}{27}$, we have

$$|B_2| \;=\; |B| - |B_1| \;\geqslant\; \left\lceil \frac{n-4}{3} \right\rceil - \frac{5n-26}{27} \;\geqslant\; \frac{4n-10}{27}.$$

We apply Lemma 4.1 to the segment e_2 with A_2 and B_2 to get a connected matching, where $a = |A_2|$ and $b = |B_2|$. Using properties of the lower bound $m(a,b)$ of Lemma 4.1, we can see that a worst-case lower bound is obtained evaluating $m(a,b)$ at the values $a = \frac{13n-19}{27}$ and $b = \frac{4n-10}{27}$. With these values one obtains the lower bound $(5n+1)/27$. ◀

References

1 Bernardo M. Ábrego and Silvia Fernández-Merchant. The rectilinear local crossing number of K_n. *J. Comb. Theory, Ser. A*, 151:131–145, 2017. doi:10.1016/j.jcta.2017.04.003.

2 Oswin Aichholzer, Jan Kyncl, Manfred Scheucher, Birgit Vogtenhuber, and Pavel Valtr. On crossing-families in planar point sets. *Comput. Geom.*, 107:101899, 2022. doi:10.1016/j.comgeo.2022.101899.

3 Boris Aronov, Paul Erdös, Wayne Goddard, Daniel J. Kleitman, Michael Klugerman, János Pach, and Leonard J. Schulman. Crossing families. *Comb.*, 14(2):127–134, 1994. doi:10.1007/BF01215345.

4 Jiří Matoušek. *Lectures on discrete geometry*, volume 212 of *Graduate texts in mathematics*. Springer, 2002. doi:10.1007/978-1-4613-0039-7.

5 János Pach, Natan Rubin, and Gábor Tardos. Planar point sets determine many pairwise crossing segments. *Adv. Math.*, 386:Paper No. 107779, 21, 2021. doi:10.1016/j.aim.2021.107779.