Discrete & Computational Geometry
https://doi.org/10.1007/s00454-022-00438-0

An Almost Optimal Bound on the Number of Intersections of Two Simple Polygons

Eyal Ackerman[1] · **Balázs Keszegh**[2,3] · **Günter Rote**[4]

Received: 16 August 2020 / Revised: 31 July 2022 / Accepted: 31 July 2022

Abstract
What is the maximum number of intersections of the boundaries of a simple m-gon and a simple n-gon? This is a basic question in combinatorial geometry, and the answer is easy if at least one of m and n is even: If both m and n are even, then every pair of sides may cross and so the answer is mn. If exactly one polygon, say the n-gon, has an odd number of sides, it can intersect each side of the m-gon polygon at most $n-1$ times; hence there are at most $mn-m$ intersections. It is not hard to construct examples that meet these bounds. If both m and n are odd, the best known construction has $mn-(m+n)+3$ intersections, and it is conjectured that this is the maximum. However, the best known upper bound is only $mn-(m+\lceil n/6\rceil)$, for $m \geq n$. We prove a new upper bound of $mn-(m+n)+C$ for some constant C, which is optimal apart from the value of C.

Keywords Simple polygon · Ramsey theory · Combinatorial geometry

Mathematics Subject Classification 52C45

Editor in Charge: Kenneth Clarkson

Eyal Ackerman
ackerman@sci.haifa.ac.il

Balázs Keszegh
keszegh@renyi.hu

Günter Rote
rote@inf.fu-berlin.de

1 Department of Mathematics, Physics, and Computer Science, University of Haifa at Oranim, 36006 Tivon, Israel

2 Alfréd Rényi Institute of Mathematics, 1053 Budapest, Hungary

3 MTA-ELTE Lendület Combinatorial Geometry Research Group, ELTE Eötvös Loránd University, Budapest, Hungary

4 Department of Computer Science, Freie Universität Berlin, Takustr. 9, 14195 Berlin, Germany

Published online: 12 October 2022

1 Introduction

To determine the union of two or more geometric objects in the plane is one of the basic computational geometric problems. In strong relation to that, determining the maximum complexity of the union of two or more geometric objects is a basic extremal geometric problem. We study this problem when the two objects are simple polygons.

Let P and Q be two simple polygons with m and n sides, respectively, where $m, n \geq 3$. We are interested in the maximum number of intersection points of the boundaries of P and Q, hence we assume that this number is finite, that is, there are no sides of P and Q that overlap.

This problem was first studied in 1993 by Dillencourt et al. [3]. The cases when m or n is even are solved there. If m and n are both even, then every pair of sides may cross and so the answer is mn. Figure 1a shows one of many ways to achieve this number. If one polygon, say Q, has an odd number n of sides, no line segment s can be intersected n times by Q, because otherwise each side of Q would have to flip from one side of s to the other side. Thus, each side of the m-gon P is intersected at most $n-1$ times, for a total of at most $mn-m$ intersections. It is easy to see that this bound is tight when P has an even number of sides, see Fig. 1b.

When both m and n are odd, the situation is more difficult; the bound that is obtained by the above argument remains at $mn - \max\{m, n\}$, because the set of m intersections that are necessarily "missing" due to the odd parity of n might conceivably overlap with the n intersections that are "missing" due to the odd parity of m. However, the best known family of examples gives only $mn-(m+n)+3 = (m-1)(n-1)+2$ intersection points, see Fig. 1c.

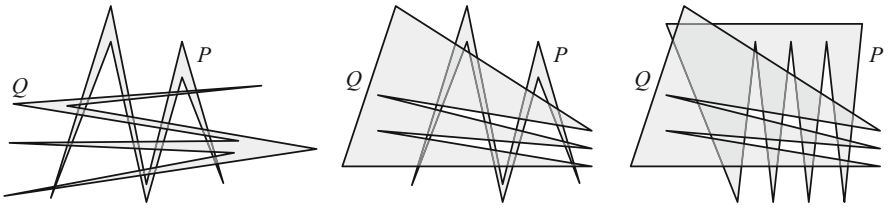
Conjecture 1 *Let P and Q be simple polygons with m and n sides, respectively, such that $m, n \geq 3$ are odd numbers. If they intersect in a finite number of points, then there are at most $mn-(m+n)+3$ intersection points between sides of P and sides of Q.*

In [3] an unrecoverable error appears in a claimed proof of Conjecture 1. Another attempted proof [8] also turned out to have a fault. The only correct improvement over the trivial upper bound is an upper bound of $mn-(m+\lceil n/6\rceil)$ for $m \geq n$, due to Černý et al. [2]. We will briefly discuss their proof in Sect. 2. We improve the upper bound to $mn-(m+n)+O(1)$, which is optimal apart from an additional constant:

Theorem 1 *There is an absolute constant C such that the following holds. Suppose that P and Q are simple polygons with m and n sides, respectively, such that m and n are odd numbers. If they intersect in a finite number of points, then there are at least $m+n-C$ pairs of a side of P and a side of Q that do not intersect. Hence, there are at most $mn-(m+n)+C$ intersections.*

The value of the constant C that we obtain in our proof is around $2^{2^{67}}$. We did not make a large effort to optimize this value, and obviously, there is ample space for improvement.

A preliminary version of this paper was presented at the *36th International Symposium on Computational Geometry* (SoCG 2020) in June 2020 [1].



(a) (b) (c)

Fig. 1 **a** Optimal construction for $m = n = 8$, with $8 \times 8 = 64$ intersections. **b** Optimal construction for $m = 8$, $n = 7$, with $8 \times 6 = 48$ intersections. **c** Lower-bound construction for $m = 9$, $n = 7$. There are $8 \times 6 + 2 = 50$ intersections

2 Overview of the Proof

First we establish the crucial statement that the odd parity of m and n allows us to *associate* to any two consecutive sides of one polygon a pair of consecutive sides of the other polygon with a restricted intersection pattern among the four involved sides (Lemma 5 and Fig. 5). This is the only place where we use the fact that P and Q are polygons, rather than merely two sets of pairwise disjoint segments, and that each of them has an odd number of sides.

A simple observation (Observation 2) relates the bound on C in Theorem 1 to the number of connected components of the bipartite "disjointness graph" between the polygon sides of P and Q. Our goal is therefore to show that there are few connected components.

We proceed to consider *two* pairs of associated pairs of sides (four consecutive pairs with eight sides in total). Unless they form a special structure, they cannot belong to four different connected components (Lemma 7). (Four is the maximum number of components that they could conceivably have.) The proof involves a case analysis with a moderate number of cases. This structural statement allows us to reduce the bound on the number of components by a constant factor, and thereby, we can already improve the best previous result on the number of intersections (Proposition 9 in Sect. 6).

Finally, to get a constant bound on the number of components, our strategy is to use Ramsey-theoretic arguments like the Erdős–Szekeres Theorem on caps and cups or the pigeonhole principle (see Sect. 7) in order to impose additional structure on the configurations that we have to analyze. This is the place in the argument where we give up control over the constant C in exchange for useful properties that allow us to derive a contradiction. This eventually boils down again to a moderate number of cases (Sect. 8.2).

By contrast, the proof of the bound $mn - (m + \lceil n/6 \rceil)$ for $m \geq n$ by Černý et al. proceeds in a more local manner. The core of their argument [2, Lemma 3], which is proved by case analysis, is that it is impossible to have six consecutive sides of one polygon together with six distinct sides of the other polygon forming a perfect matching in the disjointness graph. This statement is used to bound the number of components of the disjointness graph. (Lemma 8 below uses a similar argument.)

3 General Assumptions and Notations

Let P and Q be two simple polygons with sides $p_0, p_1, \ldots, p_{m-1}$ and $q_0, q_1, \ldots, q_{n-1}$. We assume that $m \geq 3$ and $n \geq 3$ are odd numbers. Indices are taken modulo m or n, respectively. We consider the sides p_i and q_j as closed line segments. The condition that the polygon P is simple means that its edges are pairwise disjoint except for the unavoidable common endpoints between *consecutive* sides p_i and p_{i+1}. Throughout this paper, unless stated otherwise, we regard a polygon as a piecewise linear closed curve, and we disregard the region that it encloses. Thus, by intersections between P and Q, we mean intersection points between the polygon *boundaries*.

General Position. The basic assumption that is made throughout the paper is that P and Q have only finitely many intersection points, i.e., there are no overlapping edges. In addition, we will assume from now on that no vertex of one polygon lies on the other polygon. This assumption will be justified in Sect. 9 by a non-trivial perturbation. Thus, every intersection point between P and Q is an interior point of two polygon sides.

Furthermore, we assume for convenience that no three vertices of P and Q combined lie on a line, and no two sides of P and Q combined are parallel. Indeed, given that no vertex of one polygon lies on the other polygon, this can be achieved easily by another small perturbation of the sides without changing the intersection pattern.

The Disjointness Graph. As in [2], our basic tool of analysis is the *disjointness graph* of P and Q, which we denote by $G^{\mathrm{D}} = (V^{\mathrm{D}}, E^{\mathrm{D}})$. (Its original name in [2] is *non-intersection graph*.) It is a bipartite graph with node set $V^{\mathrm{D}} = \{p_0, p_1, \ldots, p_{m-1}\} \cup \{q_0, q_1, \ldots, q_{n-1}\}$ and edge set $E^{\mathrm{D}} = \{(p_i, q_j) \mid p_i \cap q_j = \emptyset\}$. (Since we are interested in the situation where almost all pairs of edges intersect, the disjointness graph is more useful than its more commonly used complement, the intersection graph.) Our goal is to bound from above the number of connected components of G^{D}:

Observation 2 *If G^{D} has at most C connected components, then G^{D} has at least $m+n-C$ edges. Thus, there are at least $m+n-C$ pairs of a side of P and a side of Q that do not intersect, and there are at most $mn-(m+n)+C$ crossings between P and Q.*

For a polygon side s of P or Q, $\mathrm{CC}(s)$ denotes the connected component of the disjointness graph G^{D} to which s belongs.

Geometric Notions. Let s and s' be two line segments. We denote by $\ell(s)$ the line through s and by $I(s, s')$ the intersection of $\ell(s)$ and $\ell(s')$, see Fig. 4. We say that s and s' are *avoiding* if neither of them contains $I(s, s')$. (This requirement is stronger than just disjointness.) If s and s' are avoiding or share an endpoint, we denote by $\mathbf{r}_{s'}(s)$ the ray from $I(s, s')$ to infinity that contains s, and by $\mathbf{r}_s(s')$ the ray from $I(s, s')$ to infinity that contains s'. Moreover, we denote by $\mathrm{Cone}(s, s')$ the convex cone with apex $I(s, s')$ between these two rays.

Observation 3 *If a segment s'' that does not go through $I(s, s')$ has one of its endpoints in the interior of $\mathrm{Cone}(s, s')$, then s'' cannot intersect both $\mathbf{r}_{s'}(s)$ and $\mathbf{r}_s(s')$. In particular, it cannot intersect both s and s'.*

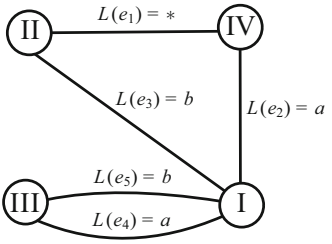


Fig. 2 The edge-labeled multigraph G_0 in Proposition 4

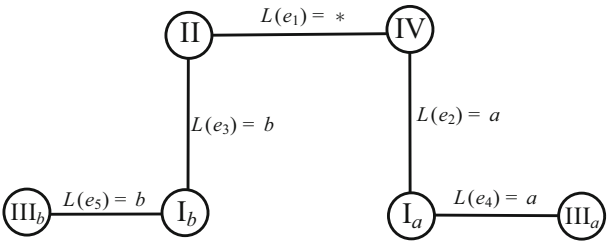


Fig. 3 The unfolded graph G_0'

4 Associated Pairs of Consecutive Sides

We begin with the following seemingly unrelated claim concerning a specific small edge-labeled multigraph. Let $G_0 = (V_0, E_0)$ be the undirected multigraph shown in Fig. 2. It has four nodes $V_0 = \{\text{I}, \text{II}, \text{III}, \text{IV}\}$ and five edges $E_0 = \{e_1 = \{\text{II}, \text{IV}\}, e_2 = \{\text{I}, \text{IV}\}, e_3 = \{\text{I}, \text{II}\}, e_4 = \{\text{I}, \text{III}\}, e_5 = \{\text{I}, \text{III}\}\}$. Every edge $e_i \in E_0$ has a label $L(e_i) \in \{a, b, *\}$ as follows: $L(e_1) = *$, $L(e_2) = L(e_4) = a$, $L(e_3) = L(e_5) = b$.

Proposition 4 *If W is a closed walk in G_0 of odd length, then W contains two cyclically consecutive edges of labels a and b.*

Proof Suppose for contradiction that W does not contain two consecutive edges of labels a and b. Since W cannot switch between the a-edges and the b-edges in I or III, we can split I (resp., III) into two nodes I_a and I_b (resp., III_a and III_b) such that every a-labeled edge that is incident to I (resp., III) in G_0 becomes incident to I_a (resp., III_a) and every b-labeled edge that is incident to I (resp., III) in G_0 becomes incident to I_b (resp., III_b). In the resulting graph G_0', which is shown in Fig. 3, we can find a closed walk W' that corresponds to W and that uses the edges with the same name as W. Since G_0' is a path, every closed walk has even length. Thus, W cannot have odd length. □

Lemma 5 *Let p_a and p_b be two sides of P that are either consecutive or avoiding such that $\text{CC}(p_a) \neq \text{CC}(p_b)$. Then there are two consecutive sides $q_i, q_{i\pm1}$ of Q such that $(p_a, q_i), (p_b, q_{i\pm1}) \in E^{\text{D}}$ and $(p_a, q_{i\pm1}), (p_b, q_i) \notin E^{\text{D}}$. Furthermore, $I(p_a, p_b) \in \text{Cone}(q_i, q_{i\pm1})$ or $I(q_i, q_{i\pm1}) \in \text{Cone}(p_a, p_b)$.*

The sign '$\pm$' is needed since we do not know which of the consecutive sides intersects p_i and is disjoint from p_{i+1}.

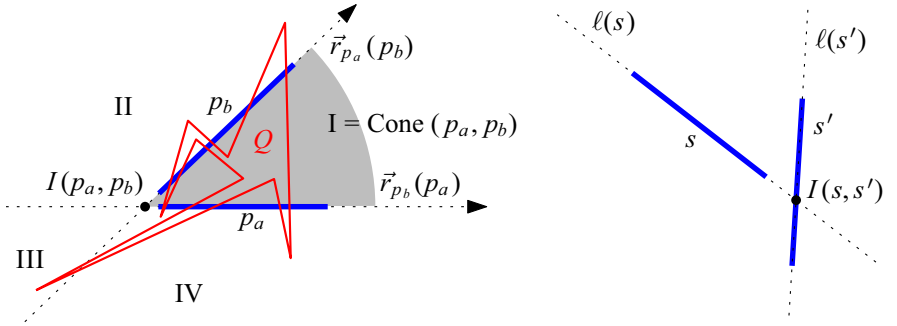


Fig. 4 How an odd polygon Q can intersect two segments. The segments p_a and p_b are avoiding, whereas s and s' are disjoint but non-avoiding. In this situation, we say that s *stabs* s'

Proof We may assume without loss of generality that $I(p_a, p_b)$ is the origin, p_a lies on the positive x-axis and the interior of p_b is above the x-axis. The lines $\ell(p_a)$ and $\ell(p_b)$ partition the plane into four convex cones ("quadrants"). Denote them in counterclockwise order by I, II, III, IV, starting with $\mathrm{I} = \mathrm{Cone}(p_a, p_b)$, see Fig. 4.

Every side of Q must intersect p_a or p_b (maybe both), since $\mathrm{CC}(p_a) \neq \mathrm{CC}(p_b)$. One can now check that traversing the sides of Q in order generates a closed walk W in the graph G_0 of Fig. 2. For example, a side of Q that we traverse from its endpoint in I to its endpoint in III and that intersects p_a corresponds to traversing the edge $e_4 = \{\mathrm{I}, \mathrm{III}\}$ from I to III, whose label is $L(e_4) = a$. We do not care which of p_a and p_b are crossed when we move between II and IV.

It follows from Proposition 4 that Q has two consecutive sides $q_i, q_{i\pm1}$ such that q_i intersects p_b and does not intersect p_a, while $q_{i\pm1}$ intersects p_a and does not intersect p_b. Hence, $(p_a, q_i), (p_b, q_{i\pm1}) \in E^{\mathrm{D}}$ and $(p_a, q_{i\pm1}), (p_b, q_i) \notin E^{\mathrm{D}}$. Furthermore, $I(q_i, q_{i\pm1})$ must be either in I or III as these are the only nodes in G_0 that are incident both to an edge labeled a and an edge labeled b. In the latter case $I(p_a, p_b) \in \mathrm{Cone}(q_i, q_{i\pm1})$, and in the former case $I(q_i, q_{i\pm1}) \in \mathrm{Cone}(p_a, p_b)$. □

Let p_i, p_{i+1} be two sides of P such that $\mathrm{CC}(p_i) \neq \mathrm{CC}(p_{i+1})$. Then by Lemma 5 there are sides $q_j, q_{j\pm1}$ of Q such that $(p_i, q_j), (p_{i+1}, q_{j\pm1}) \in E^{\mathrm{D}}$. We say that such a pair $q_j, q_{j\pm1}$ is *associated* to p_i, p_{i+1}. By Lemma 5 we have $I(q_j, q_{j\pm1}) \in \mathrm{Cone}(p_i, p_{i+1})$ or $I(p_i, p_{i+1}) \in \mathrm{Cone}(q_j, q_{j\pm1})$. If the first condition holds we say that p_i, p_{i+1} is *hooking* and $q_j, q_{j\pm1}$ is *hooked*, see Fig. 5. In the second case we say that p_i, p_{i+1} is *hooked* and $q_j, q_{j\pm1}$ is *hooking*. Note that it is possible that a pair of consecutive sides is associated with several pairs and that it is both hooking and hooked (with respect to two different pairs from the other polygon or even with respect to a single pair, as in Fig. 5c).

Observation 6 (Axis Property) *If the pair p_i, p_{i+1} and the pair $q_j, q_{j\pm1}$ are associated such that $(p_i, q_j), (p_{i+1}, q_{j\pm1}) \in E^{\mathrm{D}}$, then the line through $I(p_i, p_{i+1})$ and $I(q_j, q_{j\pm1})$ separates p_i and $q_{j\pm1}$ on the one side from p_{i+1} and q_j on the other side.*

We call this line the *axis* of the associated pairs. In our figures it appears as a dotted line when it is shown.

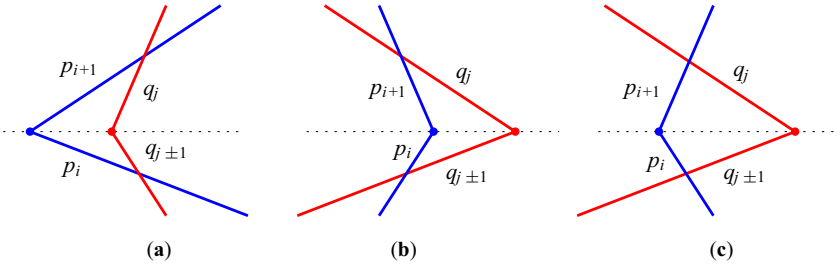


Fig. 5 Hooking and hooked pairs of consecutive sides. **a** The pair p_i, p_{i+1} is hooking and the associated pair q_j, $q_{j\pm1}$ is hooked. **b** Vice versa. **c** Both pairs are both hooking and hooked

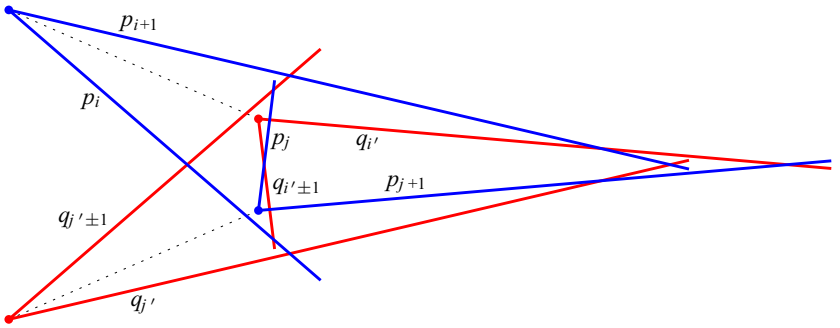


Fig. 6 The pair p_i, p_{i+1} is hooking with respect to the pair $q_{i'}$, $q_{i'\pm1}$, and p_j, p_{j+1} is hooked with respect to $q_{j'}$, $q_{j'\pm1}$

5 The Principal Structure Lemma About Pairs of Associated Pairs

Lemma 7 *Let p_i, p_{i+1}, p_j, p_{j+1} be two pairs of consecutive sides of P that belong to four different connected components of G^{D}. Then it is impossible that both p_i, p_{i+1} and p_j, p_{j+1} are hooked or that both pairs are hooking.*

Figure 6 shows a scenario with four different components, together with the associated pairs of Q. The combinatorial structure of such a configuration is unique up to relabeling.

Proof Suppose by contradiction that both pairs p_i, p_{i+1} and p_j, p_{j+1} are hooking or both of these pairs are hooked. Let $q_{i'}$, $q_{i'\pm1}$ and $q_{j'}$, $q_{j'\pm1}$ be associated (hooked or hooking) pairs of p_i, p_{i+1} and p_j, p_{j+1}, respectively, such that $(p_i, q_{i'})$, $(p_{i+1}, q_{i'\pm1}) \in E^{\mathrm{D}}$ and $(p_j, q_{j'})$, $(p_{j+1}, q_{j'\pm1}) \in E^{\mathrm{D}}$. Recall that the existence of such pairs follows from Lemma 5.

For better readability, we rename p_i, p_{i+1} and $q_{i'}$, $q_{i'\pm1}$ as a, b and A, B, and we rename p_j, p_{j+1} and $q_{j'}$, $q_{j'\pm1}$ as a', b' and A', B'. The small letters denote sides of P and the capital letters denote sides of Q. In the new notation, a, b are consecutive sides of P with an associated pair A, B of consecutive sides of Q, and a', b' are two other consecutive sides of P with an associated pair A', B' of consecutive sides of Q. The disjointness graph G^{D} contains the edges (a, A), (b, B), (a', A'), (b', B'). Since

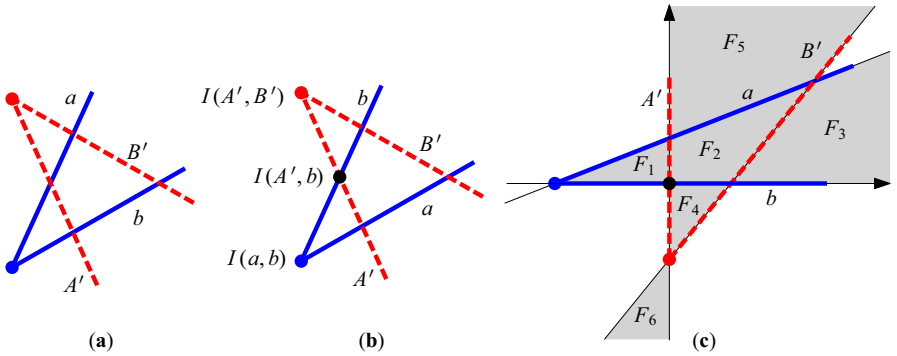


Fig. 7 Normalizing the position of a, b, A', B'

a, b, a', b' belong to different connected components of G^{D}, it follows that the nodes A, B, A', B', to which they are connected, belong to the same four different connected components. There can be no more edges among these eight nodes, and they induce a matching in G^{D}. One can remember as a rule that every side of P intersects every side of Q among the eight involved sides, except when their names differ only in their capitalization.

Suppose first that both a, b and a', b' are hooking. Hence, $I(A, B) \in \mathrm{Cone}(a, b)$ and $I(A', B') \in \mathrm{Cone}(a', b')$. Since each of A' and B' intersects each of a and b they must lie as in Fig. 7a. To facilitate the future discussion, we will now normalize the positions of these four sides.

We first ensure that the intersection $I(A', b)$ is directly adjacent to the two polygon vertices $I(a, b)$ and $I(A', B')$ in the arrangement of the four sides, as shown in Fig. 7b. This can be achieved by swapping the labels a, A with the labels b, B if necessary, and by independently swapping the labels a', A' with b', B' if necessary. Our assumptions are invariant under these swaps.

By an affine transformation we may finally assume that $I(A', b)$ is the origin; b lies on the x-axis and is directed to the right; and A' lies on the y-axis and is directed upwards. Then a has a positive slope and its interior is in the upper half-plane, and B' has a positive slope and its interior is to the right of the y-axis, see Fig. 7c.

The arrangement of the lines through a, b, A', B' has 11 faces, some of which are marked as $F_1, \ldots, F_6$ in Fig. 7. Our current assumption is that both a, b and a', b' are hooking: The hooking of a, b means that $I(A, B) \in \mathrm{Cone}(a, b) = F_1 \cup F_2 \cup F_3$. By the Axis Property (Observation 6), the line through $I(A', B')$ and $I(a', b')$ must separate A' from B'. Therefore, the vertex $I(a', b')$ can lie only in $F_2 \cup F_4 \cup F_5 \cup F_6$. Thus, based on the faces that contain $I(A, B)$ and $I(a', b')$, there are 12 cases to consider. Some of these cases are symmetric, and all can be easily dismissed, as follows.

In the figures, the four sides a', b', A', B', which are associated to the second associated pair are dashed. All dashed sides of one polygon must intersect all solid sides of the other polygon.

1. $I(A, B) \in F_1$ and $I(a', b') \in F_2$, see Fig. 8 (symmetric to $I(A, B) \in F_2$ and $I(a', b') \in F_4$). Let r_a (resp., r_b) be the ray on $\ell(a)$ (resp., $\ell(b)$) that goes from the right endpoint of a (resp., b) to the right. Since a' is not allowed to cross b, the

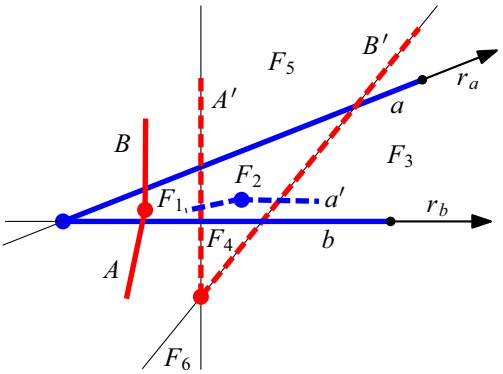


Fig. 8 Case 1: $I(A, B) \in F_1$, $I(a', b') \in F_2$

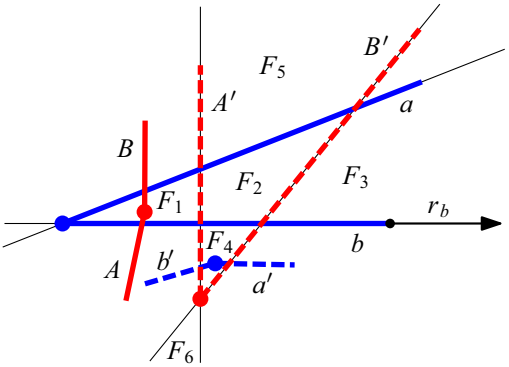


Fig. 9 Case 2: $I(A, B) \in F_1$, $I(a', b') \in F_4$

only way for a' to intersect A is by crossing r_b. Similarly, in order to intersect B, a' has to cross r_a. However, it cannot intersect both r_a and r_b, by Observation 3. Since we did not use the assumption that A, B are hooked, the analysis holds for the symmetric case 6, $I(A, B) \in F_2$ and $I(a', b') \in F_4$, as well.

2. $I(A, B) \in F_1$ and $I(a', b') \in F_4$, see Fig. 9. Since a' is not allowed to cross b, the only way for a' to intersect B is by crossing r_b. However, in this case a' cannot intersect A.
3. $I(A, B) \in F_1$ and $I(a', b') \in F_5$, see Fig. 10 (symmetric to $I(A, B) \in F_3$ and $I(a', b') \in F_4$). Both a' and b' must intersect A, and they have to go below the line $\ell(b)$ to do so. However, a' can only cross $\ell(b)$ to the right of b, and b' can only cross $\ell(b)$ to the left of b, and therefore they cross A from different sides. This is impossible, because a' and b' start from the same point.
4. $I(A, B) \in F_1$ and $I(a', b') \in F_6$. If one of the polygon sides a' and b' has an endpoint in F_4 (see Fig. 11a), then this side cannot intersect B. So assume otherwise, see Fig. 11b. The side a' intersects B' and is disjoint from A', while b' is disjoint from B' and intersects A'. (Due to space limitation some line segments are drawn schematically as curves.) Thus, each of a' and b' has an endpoint in

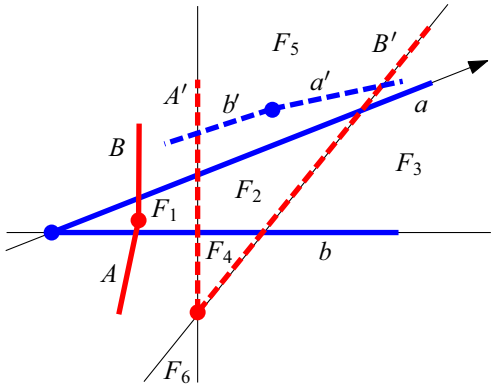
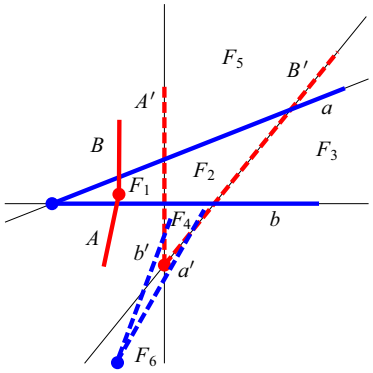
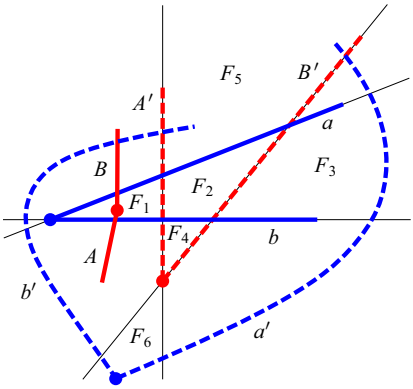


Fig. 10 Case 3: $I(A, B) \in F_1$ and $I(a', b') \in F_5$



(**a**) At least one of the sides a' and b' has an endpoint in F_4.



(**b**) None of the sides a' and b' has an endpoint in F_4.

Fig. 11 Case 4: $I(A, B) \in F_1$ (or $I(A, B) \in F_2$, which is similar) and $I(a', b') \in F_6$

$F_2 \cup F_5$. But then $I(A, B) \in \text{Cone}(a', b')$ and it follows from Observation 3 that neither A nor B can intersect both a' and b'.

5. $I(A, B) \in F_2$ and $I(a', b') \in F_2$, see Fig. 12. Since a', b' is hooking, $I(A', B') \in \text{Cone}(a', b')$, and the line segments a', b', A', b, B' enclose a convex pentagon. The polygon side A must intersect b, a', and b', but it is restricted to $F_2 \cup F_4$. It follows that A must intersect three sides of the pentagon, which is impossible. (This is in fact the only place where we need the assumption that a', b' is hooking.)
6. $I(A, B) \in F_2$ and $I(a', b') \in F_4$. This is symmetric to case 1.
7. $I(A, B) \in F_2$ and $I(a', b') \in F_5$, see Fig. 13 (symmetric to $I(A, B) \in F_3$ and $I(a', b') \in F_2$). Then A is restricted to $F_2 \cup F_4$, while a' and b' do not intersect F_2 and F_4. Therefore A can intersect neither a' nor b'.
8. $I(A, B) \in F_2$ and $I(a', b') \in F_6$. This case is very similar to case 4, where $I(A, B) \in F_1$ and $I(a', b') \in F_6$, see Fig. 11. If one of the polygon sides a'

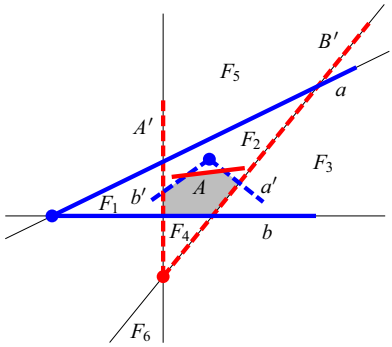


Fig. 12 Case 5: $I(A, B) \in F_2$, $I(a', b') \in F_2$

and b' has an endpoint in F_4, then it cannot intersect B. Otherwise, $I(A, B) \in \text{Cone}(a', b')$ and therefore, neither A nor B can intersect both a' and b'.

9. $I(A, B) \in F_3$ and $I(a', b') \in F_2$. This is symmetric to case 7.
10. $I(A, B) \in F_3$ and $I(a', b') \in F_4$. This is symmetric to case 3.
11. $I(A, B) \in F_3$ and $I(a', b') \in F_5$, see Fig. 14. Then the intersection of b' and A can lie only in the lower left quadrant. It follows that the triangle whose vertices are $I(a', b')$, $I(a', A)$, and $I(A, b')$ contains a and does not contain $I(A, B)$. This in turn implies that B cannot intersect both b' and a, without intersecting B'.
12. $I(A, B) \in F_3$ and $I(a', b') \in F_6$, see Fig. 15. As in case 4, we may assume that neither a' nor b' has an endpoint in F_4, since then this side could not intersect B. We may also assume that $I(A, B) \notin \text{Cone}(a', b')$ for otherwise neither A nor B intersects both of a' and b', according to Observation 3. If a' has an endpoint in F_2, then it cannot intersect B (see Fig. 15a). Otherwise, if a' has an endpoint in F_5, then B cannot intersect b' (Fig. 15b).

We have finished the case that a, b and a', b' are hooking. Suppose now that a, b and a', b' are hooked with respect to A, B and A', B', respectively. Then A, B is hooking with respect to a, b and A', B' is hooking with respect to a', b'. Recall that A, B, A', and B' belong to four different connected components. Hence, this case can be handled as above, after exchanging the capital letters with the small letters (i.e., exchanging P and Q). □

6 A Weaker Bound

The principal structure lemma is already powerful enough to get an improvement over the previous best bound:

Lemma 8 *G^D has at most $(n + 5)/2$ connected components.*

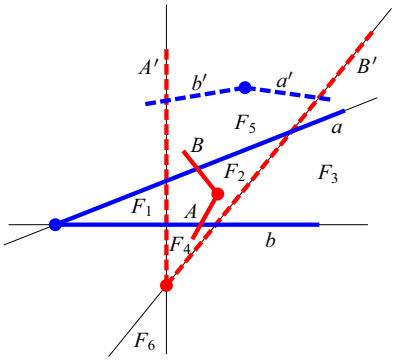


Fig. 13 Case 7: $I(A, B) \in F_2$, $I(a', b') \in F_5$

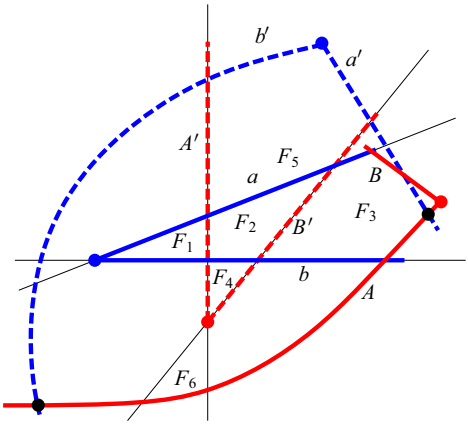
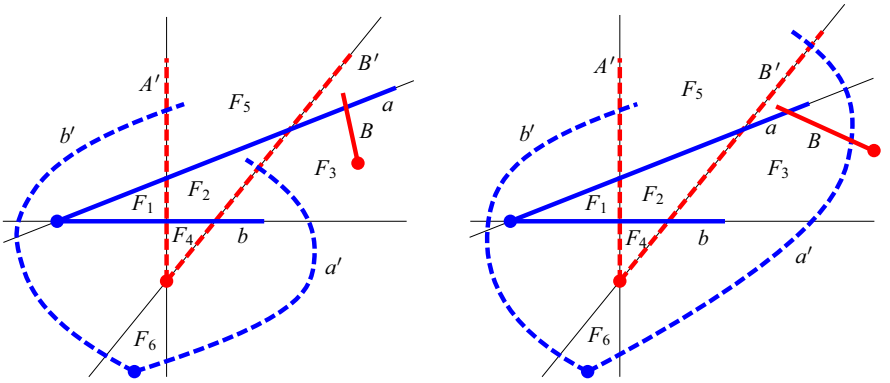


Fig. 14 Case 11: $I(A, B) \in F_3$ and $I(a', b') \in F_5$

Proof Partition the sides $q_0, q_1, \ldots, q_{n-1}$ of Q into $(n-1)/2$ disjoint pairs q_{2i}, q_{2i+1}, discarding the last side q_{n-1}. Let H_+ denote the subset of these pairs that are hooked. Suppose first that this set contains some pair q_{2i_0}, q_{2i_0+1} of sides that are in two different connected components. Combining q_{2i_0}, q_{2i_0+1} with any of the remaining pairs q_{2i}, q_{2i+1} of H_+, Lemma 7 tells us that the sides q_{2i} and q_{2i+1} must either belong to the same connected component, or one of them must belong to $\mathrm{CC}(q_{2i_0})$ or $\mathrm{CC}(q_{2i_0+1})$. In other words, each remaining pair contributes at most one "new" connected component, and it follows that the sides in H_+ belong to at most $|H_+| + 1$ connected components. This conclusion holds also in the case that H_+ contains no pair q_{2i_0}, q_{2i_0+1} of sides that are in different connected components.

The same argument works for the complementary subset H_- of pairs that are not hooked, but hooking. Along with $\mathrm{CC}(q_{n-1})$ there are at most $(|H_+|+1)+(|H_-|+1)+1 = (n-1)/2+3 = (n+5)/2$ components. □

Together with Observation 2, this already improves the previous bound of $mn - (m + \lceil n/6 \rceil)$ for a large range of parameters, namely when $m \geq n \geq 11$:



(a) If a' has an endpoint in F_2, then it cannot intersect B.

(b) If a' has an endpoint in F_5, then B cannot intersect b'.

Fig. 15 Case 12: $I(A, B) \in F_3$ and $I(a', b') \in F_6$

Proposition 9 *Let P and Q be simple polygons with m and n sides, respectively, such that m and n are odd and $m \geq n \geq 3$. Then there are at most $mn - (m + (n-5)/2)$ intersection points between P and Q.*

7 Ramsey-Theoretic Tools

We recall some classic results. A tournament is a directed graph that contains between every pair of nodes x, y either the arc (x, y) or the arc (y, x) but not both. A tournament is *transitive* if for every three nodes x, y, z the existence of the arcs (x, y) and (y, z) implies the existence of the arc (x, z). Equivalently, the nodes can be ordered on a line such that all arcs are in the same direction. The following is easy to prove by induction.

Lemma 10 (Erdős and Moser [5]) *Every tournament on a node set V contains a transitive sub-tournament on $1 + \lfloor \log_2 |V| \rfloor$ nodes.*

Proof Choose $v \in V$ arbitrarily, and let $N \subseteq V - \{v\}$ with $|N| \geq (|V| - 1)/2$ be the set of in-neighbors of v or the set of out-neighbors of v, whichever is larger. Then v together with a transitive sub-tournament of N gives a transitive sub-tournament of size one larger. □

A set of points $p_1, p_2, \ldots, p_r$ in the plane sorted by x-coordinates (and with distinct x-coordinates) forms an *r-cup* (resp., *r-cap*) if p_i is below (resp., above) the line through p_{i-1} and p_{i+1} for every i with $1 < i < r$.

Theorem 11 (Erdős–Szekeres Theorem for caps and cups in point sets [6]) *For any two integers $r \geq 2$ and $s \geq 2$, the value $\mathrm{ES}(r, s) := \binom{r+s-4}{r-2}$ fulfills the following statement:*

Suppose that P is a set of $\mathrm{ES}(r,s)+1$ points in the plane with distinct x-coordinates such that no three points of P lie on a line. Then P contains an r-cup or an s-cap.

Moreover, $\mathrm{ES}(r,s)$ is the smallest value that fulfills the statement.

A similar statement holds for lines by the standard point-line duality. A set of lines $\ell_1, \ell_2, \ldots, \ell_r$ sorted by slope forms an *r-cup* (resp., *r-cap*) if ℓ_{i-1} and ℓ_{i+1} intersect below (resp., above) ℓ_i for every $1 < i < r$.

Theorem 12 (Erdős–Szekeres Theorem for lines) *For the numbers $\mathrm{ES}(r,s)$ from Theorem 11, the following statement holds for any two integers $r \geq 2$ and $s \geq 2$:*

If L is a set of $\mathrm{ES}(r,s)+1$ non-vertical lines in the plane no two of which are parallel and no three of which intersect at a common point, then L contains an r-cup or an s-cap.

Theorem 13 (Erdős–Szekeres Theorem for monotone subsequences [6]) *For any integer $r \geq 0$, a sequence of r^2+1 distinct numbers contains either an increasing subsequence of length $r+1$ or a decreasing subsequence of length $r+1$.*

8 Proof of Theorem 1

8.1 Imposing More Structure on the Examples

Going back to the proof of Theorem 1, recall that in light of Observation 2 it is enough to prove that G^{D}, the disjointness graph of P and Q, has at most constantly many connected components.

We will use the following constants: $C_6 := 6$; $C_5 := (C_6)^2 + 1 = 37$; $C_4 := \mathrm{ES}(C_5, C_5) + 1 = \binom{70}{35} + 1 = 112{,}186{,}277{,}816{,}662{,}845{,}433 < 2^{67}$; $C_3 := 2^{C_4-1}$; $C_2 := C_3 + 5$; $C_1 := 8C_2$; $C := C_1 - 1 < 2^{2^{67}}$.

We claim that G^{D} has at most C connected components. Suppose that G^{D} has at least $C_1 = C + 1$ connected components, numbered as $1, 2, \ldots, C_1$. For each connected component j, we find two consecutive sides q_{i_j}, q_{i_j+1} of Q such that $\mathrm{CC}(q_{i_j}) = j$ and $\mathrm{CC}(q_{i_j+1}) \neq j$. We call q_{i_j} the *primary* side and q_{i_j+1} the *companion* side of the pair. We take these C_1 consecutive pairs in their cyclic order along Q and remove every second pair. This ensures that the remaining $C_1/2$ pairs are disjoint in the sense that no side of Q belongs to two different pairs.

We apply Lemma 5 to each of the remaining $C_1/2$ pairs q_{i_j}, q_{i_j+1} and find an associated pair $p_{k_j}, p_{k_j \pm 1}$ such that $(q_{i_j}, p_{k_j}), (q_{i_j+1}, p_{k_j \pm 1}) \in E^{\mathrm{D}}$. Therefore, $\mathrm{CC}(q_{i_j}) = \mathrm{CC}(p_{k_j})$ and $\mathrm{CC}(q_{i_j+1}) = \mathrm{CC}(p_{k_j \pm 1}) \neq \mathrm{CC}(q_{i_j})$. Again, we call p_{k_j} the primary side and $p_{k_j \pm 1}$ the companion side. As before, we delete half of the pairs $p_{k_j}, p_{k_j \pm 1}$ in cyclic order along P, along with their associated pairs from Q, and thus we ensure that the remaining $C_1/4$ pairs are disjoint also on P.

At least $C_1/8$ of the remaining pairs q_{i_j}, q_{i_j+1} are hooking or at least $C_1/8$ of them are hooked. We may assume that at least $C_2 = C_1/8$ of the pairs q_{i_j}, q_{i_j+1} are hooking with respect to their associated pair, $p_{k_j}, p_{k_j \pm 1}$, for otherwise, $p_{k_j}, p_{k_j \pm 1}$ is

hooking with respect to q_{i_j}, q_{i_j+1} and we may switch the roles of P and Q. Let us denote by Q_2 the set of C_2 hooking consecutive pairs $(q_{i_j}, q_{i_j\pm1})$ at which we have arrived. (Because of the potential switch, we have to denote the companion side by $q_{i_j\pm1}$ instead of q_{i_j+1} from now on.)

By construction, all C_2 primary sides q_{i_j} of these pairs belong to distinct components. We now argue that all C_2 adjacent companion sides $q_{i_j\pm1}$ with at most one exception lie in the same connected component, provided that $C_2 \geq 4$.

We model the problem by a graph whose nodes are the connected components of G^{D}. For each of the C_2 pairs $q_{i_j}, q_{i_j\pm1}$, we insert an edge between $\mathrm{CC}(q_{i_j})$ and $\mathrm{CC}(q_{i_j\pm1})$. The result is a multigraph with C_2 edges and without loops. Two disjoint edges would represent two consecutive pairs of the form $(q_{i_j}, q_{i_j\pm1})$ whose four sides are in four distinct connected components, but this is a contradiction to Lemma 7. Thus, the graph has no two disjoint edges, and such graphs are easily classified: they are the triangle (cycle on three vertices) and the star graphs K_{1t}, possibly with multiple edges. Overall, the graph involves at least $C_2 \geq 4$ distinct connected components $\mathrm{CC}(q_{i_j})$, and therefore the triangle graph is excluded. Let v be the central vertex of the star. There can be at most one j with $\mathrm{CC}(q_{i_j}) = v$, and we discard it. All other sides q_{i_j} have $\mathrm{CC}(q_{i_j}) \neq v$, and therefore $\mathrm{CC}(q_{i_j\pm1})$ must be the other endpoint of the edge, that is, v.

In summary, we have found $C_2 - 1$ adjacent pairs $q_{i_j}, q_{i_j\pm1}$ with the following properties.

- The primary sides q_{i_j} belong to $C_2 - 1$ distinct components.
- All companion sides $q_{i_j\pm1}$ belong to the same component, distinct from the other $C_2 - 1$ components.
- All $2C_2 - 2$ sides of the pairs $q_{i_j}, q_{i_j\pm1}$ are distinct.
- Each $q_{i_j}, q_{i_j\pm1}$ is hooking with respect to an associated pair $p_{k_j}, p_{k_j\pm1}$.
- All $2C_2 - 2$ sides of the pairs $p_{k_j}, p_{k_j\pm1}$ are distinct.

Let us denote by Q_2' the set of $C_2 - 1$ sides q_{i_j}.

Proposition 14 *There are no six distinct sides $q_a, q_b, q_c, q_d, q_e, q_f$ among the $C_2 - 1$ sides $q_{i_j} \in Q_2'$ such that q_a, q_b are avoiding or consecutive, q_c, q_d are avoiding or consecutive, and q_e, q_f are avoiding or consecutive.*

Proof Suppose for contradiction that six such sides exist. It follows from Lemma 5 that there are two consecutive sides $p_{a'}$ and $p_{b'}$ of P such that $\mathrm{CC}(p_{a'}) = \mathrm{CC}(q_a)$ and $\mathrm{CC}(p_{b'}) = \mathrm{CC}(q_b)$.

Similarly, we find a pair of consecutive sides $p_{c'}$ and $p_{d'}$ of P such that $\mathrm{CC}(p_{c'}) = \mathrm{CC}(q_c)$ and $\mathrm{CC}(p_{d'}) = \mathrm{CC}(q_d)$, and the same story for e and f. By the pigeonhole principle, two of the three consecutive pairs $(p_{a'}, p_{b'})$, $(p_{c'}, p_{d'})$, $(p_{e'}, p_{f'})$ are hooking or two of them are hooked. This contradicts Lemma 7. □

Define a complete graph whose nodes are the $C_2 - 1$ sides $q_{i_j} \in Q_2'$, and color an edge (q_{i_j}, q_{i_k}) red if q_{i_j} and q_{i_k} are avoiding or consecutive and blue otherwise. Proposition 14 says that this graph contains no red matching of size three. This means that we can get rid of all red edges by removing at most four nodes. To see this, pick any red edge and remove its two nodes from the graph. If any red edge remains, remove its

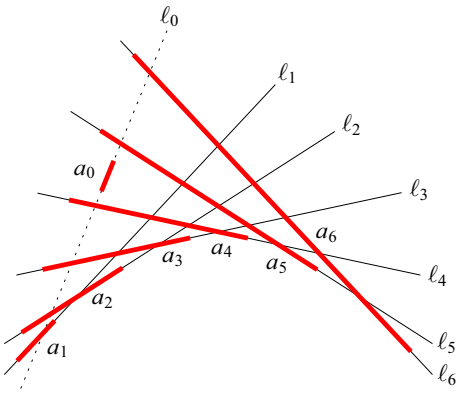


Fig. 16 The seven sides $a_0, a_1, \ldots, a_6$. The lines $\ell_0, \ldots, \ell_6$ form a 7-cap. The side a_0 could be between any two consecutive intersection points of l_0 with the rest of the lines

two nodes. Then all red edges are gone, because otherwise we would find a matching with three red edges.

We conclude that there is a blue clique of size $C_3 = C_2 - 5$, i.e., there is a set $Q_3 \subset Q_2'$ of C_3 polygon sides among the $C_2 - 1$ sides $q_{i_j} \in Q_2'$ that are pairwise non-avoiding and disjoint, i.e., they do not share a common endpoint.

Our next goal is to find a subset of seven segments in Q_3 that are arranged as in Fig. 16. To define this precisely, we say for two segments s and s' that s *stabs* s' if $I(s, s') \in s'$, see Fig. 4. Among any two non-avoiding and non-consecutive sides s and s', either s stabs s' or s' stabs s, but not both. Define a tournament T whose nodes are the C_3 sides $q_{i_j} \in Q_3$, and the arc between each pair of nodes is oriented towards the stabbed side. It follows from Lemma 10 that T has a transitive sub-tournament of size $1 + \lfloor \log_2 C_3 \rfloor = C_4$.

Furthermore, since $C_4 = \mathrm{ES}(C_5, C_5) + 1$, it follows from Theorem 12 that there is a subset of C_5 sides such that the lines through them form a C_5-cup or a C_5-cap. By a vertical reflection if needed, we may assume that they form a C_5-cap.

We now reorder these C_5 sides q_{i_j} of Q in stabbing order, according to the transitive sub-tournament mentioned above. By the Erdős–Szekeres Theorem on monotone subsequences (Theorem 13), there is a subsequence of size $C_6 + 1 = \sqrt{C_5 - 1} + 1 = 7$ such that their slopes form a monotone sequence. By a horizontal reflection if needed, we may assume that they have decreasing slopes.

We rename these seven segments to $a_0, a_1, \ldots, a_6$, and we denote the line $\ell(a_i)$ by ℓ_i, see Fig. 16. We have achieved the following properties:

- The lines $\ell_0, \ldots, \ell_6$ form a 7-cap, with decreasing slopes in this order.
- The segments a_i are pairwise disjoint and non-avoiding.
- a_i stabs a_j for every $i < j$.

These properties allow a_0 to lie between any two consecutive intersections on ℓ_0. There is no such flexibility for the other sides: Every side a_j is stabbed by every preceding side a_i. For $1 \le i < j$, a_i cannot stab a_j from the right, because then a_0 would not be able to stab a_i. Hence, the arrangement of the sides $a_1, \ldots, a_6$ must be exactly as

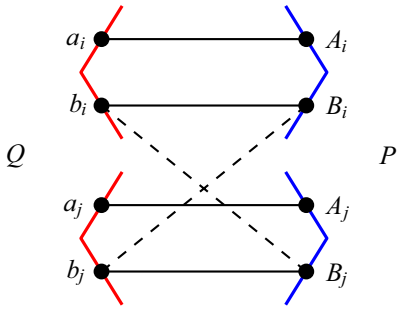


Fig. 17 The subgraph of G^{D} induced on two pairs of consecutive sides a_i, b_i and a_j, b_j of P and their associated partner pairs A_i, B_i and A_j, B_j of Q. Parts of P and Q are shown to indicate consecutive sides. The dashed edges may or may not be present

shown in Fig. 16, in the sense that the order of endpoints and intersection points along each line ℓ_i is fixed. We will ignore a_0 from now on.

8.2 Finalizing the Analysis

Recall that every a_i is the primary side of two consecutive sides a_i, b_i of Q that are hooking with respect to an associated pair A_i, B_i of consecutive sides of P. The sides a_i and A_i are the primary sides and b_i and B_i are the companion sides. All these 4×6 sides are distinct, and they intersect as follows: a_i intersects B_i and is disjoint from A_i; b_i intersects A_i and is disjoint from B_i; and $I(A_i, B_i) \in \mathrm{Cone}(a_i, b_i)$.

Figure 17 summarizes the intersection pattern among these sides. A side A_i must intersect every side a_j with $j \neq i$ and every side b_j since $\mathrm{CC}(A_i) = \mathrm{CC}(a_i) \neq \mathrm{CC}(a_j)$ and $\mathrm{CC}(A_i) = \mathrm{CC}(a_i) \neq \mathrm{CC}(b_i) = \mathrm{CC}(b_j)$. (Recall that all companion sides b_i belong to the same component.) Similarly, every side B_i must intersect every side a_j. We have no information about the intersection between B_i and b_j, as these sides belong to the same connected component. We will now derive a contradiction through a series of case distinctions.

Case 1: There are three segments A_i with the property that A_i crosses ℓ_i to the left of a_i. Without loss of generality, assume that these segments are A_1, A_2, A_3, see Fig. 18. The segments A_1, A_2, A_3 must not cross because P is a simple polygon. Therefore A_1 intersects a_2 to the right of $I(a_1, a_2)$ because otherwise A_1 would cross A_2 on the way between its intersections with ℓ_2 and with a_1. A_3 must cross ℓ_3, a_2, a_1 in this order, as shown. But then A_1 and A_3 (and a_2) block A_2 from intersecting a_3.

Case 2: There at most two segments A_i with the property that A_i crosses ℓ_i to the left of a_i. In this case, we simply discard these segments. We select four of the remaining segments and renumber them from 1 to 4. From now on, we can make the following assumption:

General Assumption: For every $1 \leq i \leq 4$, the segment A_i does not cross ℓ_i at all, or it crosses ℓ_i to the right of a_i.

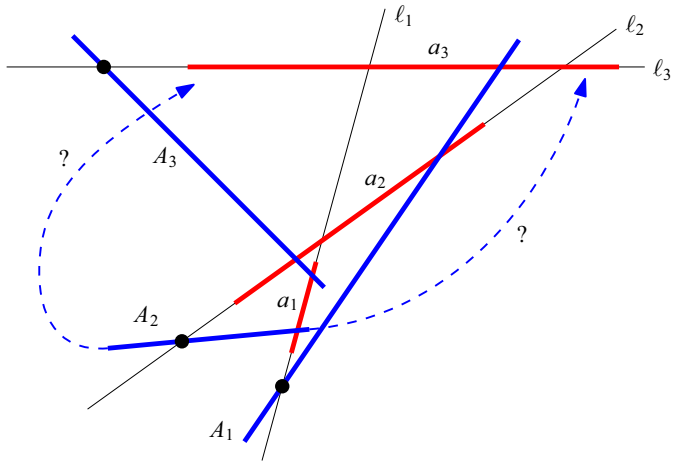


Fig. 18 The assumed intersection points between A_i and ℓ_i are marked

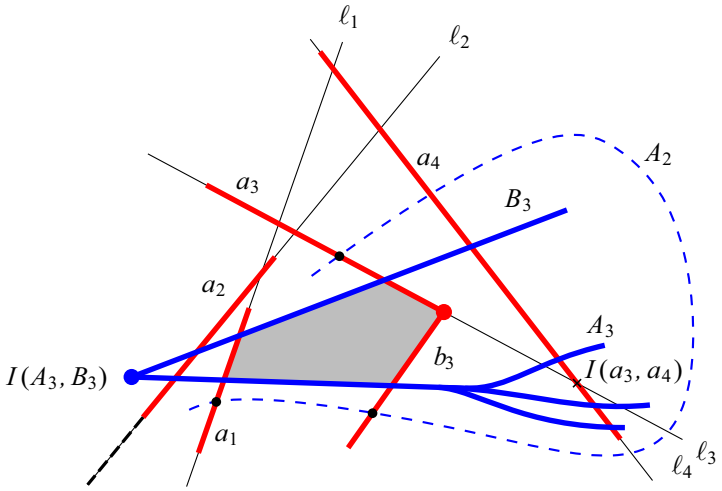


Fig. 19 Case 2.1, $I(A_3, B_3) = \text{left}(A_3)$ and $I(a_3, b_3) = \text{right}(a_3)$. A hypothetical segment A_2 is shown as a dashed curve. The side a_2 and the part of ℓ_2 to the left of a_2 is blocked for A_2

This implies that A_3 must intersect the sides a_2, a_1, a_4 in this order, and it is determined in which cell of the arrangement of the lines $\ell_1, \ell_2, \ell_3, \ell_4$ the left endpoint of A_3 lies (see Figs. 16 and 19). For the right endpoint, we have a choice of two cells, depending on whether A_3 intersects ℓ_3 or not.

We denote by $\text{left}(s)$ and $\text{right}(s)$ the left and right endpoints of a segment s. We distinguish four cases, based on whether the common endpoint of A_3 and B_3 lies at $\text{left}(A_3)$ or $\text{right}(A_3)$, and whether the common endpoint of a_3 and b_3 lies at $\text{left}(a_3)$ or $\text{right}(a_3)$.

Case 2.1: $I(A_3, B_3) = \text{left}(A_3)$ and $I(a_3, b_3) = \text{right}(a_3)$, see Fig. 19.

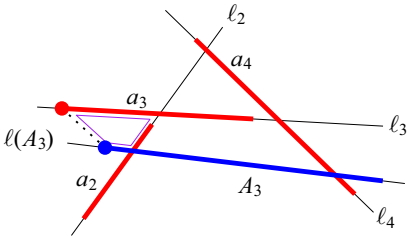


Fig. 20 Case 2.2. $I(A_3, B_3) = \text{left}(A_3)$, $I(a_3, b_3) = \text{left}(a_3)$, $\ell(A_3)$ does not intersect a_3

As indicated in the figure, we leave it open whether and where A_3 intersects ℓ_3. We know that b_3 must lie below ℓ_3 because $I(A_3, B_3) \in \text{Cone}(a_3, b_3)$. We claim that A_2 cannot have the required intersections with a_1, a_3, and b_3. Let us first consider a_1: It is cut into three pieces by A_3 and B_3.

If A_2 intersects the middle piece of a_1 in the wedge between A_3 and B_3, then A_2 intersects exactly one of a_3 and b_3 inside the wedge, as these parts together with a_1 are three sides of a convex pentagon. If A_2 intersects a_3, then it has crossed ℓ_3 and it cannot cross b_3 thereafter. If A_2 intersects b_3, it must cross ℓ_4 before leaving the wedge, and then it cannot cross a_3 thereafter.

Suppose now that A_2 crosses the bottom piece of a_1. Then it cannot go around A_3 and B_3 to the right in order to reach a_3 because it would have to intersect ℓ_4 twice. A_2 also cannot pass to the left of A_3 and B_3 because it cannot cross ℓ_2 through a_2 or, by the general assumption, to the Suppose finally that A_2 crosses the top piece of a_1. Then it would have to cross ℓ_3 twice before reaching b_3.

Case 2.2: $I(A_3, B_3) = \text{left}(A_3)$ and $I(a_3, b_3) = \text{left}(a_3)$.

If $\ell(A_3)$ does not intersect a_3, we derive a contradiction as follows, see Fig. 20. We know that the sides a_2, a_3, a_4 must be arranged as shown. The segment A_3 crosses a_2 but not a_3. Now, the parts of a_3 and A_3 to the left of ℓ_2 form two opposite sides of a quadrilateral, as shown in the figure. If this quadrilateral were not convex, then either $\ell(A_3)$ would intersect a_3, which we have excluded by assumption, or ℓ_3 would intersect A_3 left of a_3, contradicting the General Assumption. Thus, the sides a_3 and A_3 violate the Axis Property (Observation 6), which requires a_3 and A_3 to lie on different sides of the line through $I(A_3, B_3)$ and $I(a_3, b_3)$.

Looking back at the proofs so far, we have seen that the configuration of the segments a_1, a_2, a_3, a_4 according to Fig. 16 in connection with the particular case assumptions make the situation sufficiently constrained that the case can be dismissed by looking at the drawing. The treatment of the other cases will be proofs by picture in a similar way, but we will not always spell out the arguments in such detail.

If $\ell(A_3)$ intersects a_3, the situation must be as shown in Fig. 22: the pair A_3, B_3 is hooked by a_3 and b_3. The analysis of Case 2.1 (Fig. 19) applies verbatim, except that the word "pentagon" must be replaced by "hexagon".

Case 2.3: $I(A_3, B_3) = \text{right}(A_3)$ and $I(a_3, b_3) = \text{right}(a_3)$.

If A_3 lies entirely below ℓ_3, then A_3 together with a_3 violates the Axis Property (Observation 6), see Fig. 21.

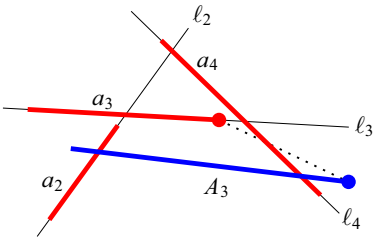


Fig. 21 Case 2.3. $I(A_3, B_3) = \text{right}(A_3)$, and $I(a_3, b_3) = \text{right}(a_3)$, A_3 lies below ℓ_3

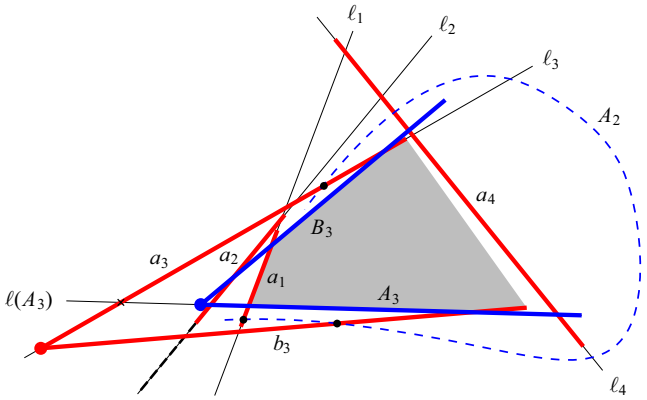


Fig. 22 Case 2.2, $I(A_3, B_3) = \text{left}(A_3)$, $I(a_3, b_3) = \text{left}(a_3)$, and $\ell(A_3)$ intersects A_3. A hypothetical segment A_2 is shown as a dashed curve

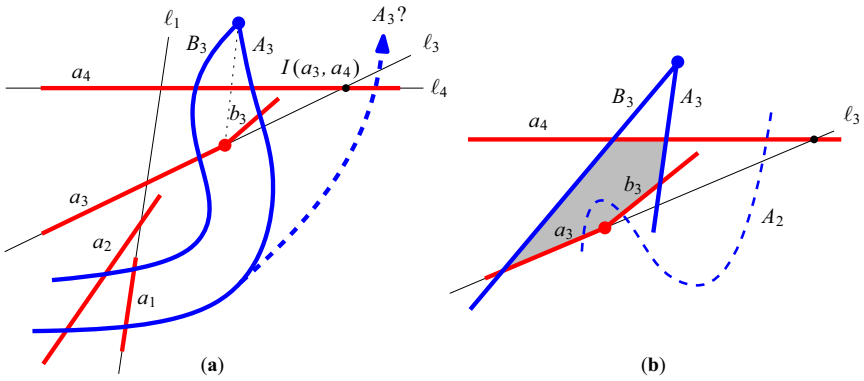


Fig. 23 Case 2.3. A_3 intersects ℓ_3

Let us therefore assume that A_3 intersects ℓ_3 (to the right of a_3), and thus $\text{right}(A_3) = I(A_3, B_3)$ lies above ℓ_3, see Fig. 23a. Then b_3 must also lie above ℓ_3, because a_3, b_3 is supposed to be hooking, that is, $I(A_3, B_3) \in \text{Cone}(a_3, b_3)$.

It follows that A_3 cannot intersect ℓ_3 to the right of $I(a_3, a_4)$ (the option shown as a dashed curve), because otherwise it would miss b_3: b_3 is blocked by a_4.

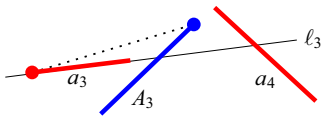


Fig. 24 Case 2.4. A_3 intersects ℓ_3

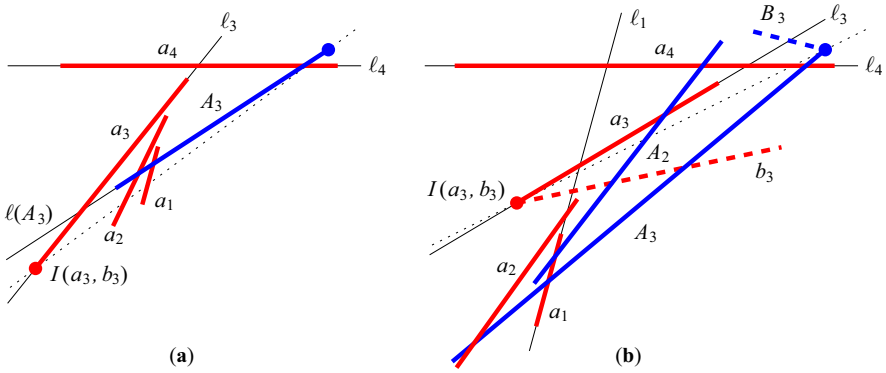


Fig. 25 Case 2.4. A_3 lies below ℓ_3

Therefore, the situation looks as shown in Fig. 23a. Figure 23b shows the position of the relevant pieces. The segments a_4, B_3, a_3, b_3, A_3 enclose a convex pentagon. Now, the segment A_2 should intersect a_3, b_3, and a_4 without crossing A_3 and B_3, like the dashed curve in the figure. This is impossible.

Case 2.4: $I(A_3, B_3) = \text{right}(A_3)$ and $I(a_3, b_3) = \text{left}(a_3)$.

If A_3 intersects ℓ_3 (to the right of a_3), then A_3 together with a_3 violates the Axis Property (Observation 6), see Fig. 24. We thus assume that A_3 lies entirely below ℓ_3.

If $\ell(A_3)$ passes above $I(a_3, b_3) = \text{left}(a_3)$, the sides a_3 and A_3 violate the Axis Property, see Fig. 25a. On the other hand, if $\ell(A_3)$ passes below $I(a_3, b_3) = \text{left}(a_3)$, as shown in Fig. 25b, then b_3 must cross ℓ_1 to the right of a_1 in order to reach A_2. Again by the Axis Property, B_3 must remain above the dotted axis line through $I(A_3, B_3) = \text{right}(A_3)$ and $I(a_3, b_3) = \text{left}(a_3)$. On ℓ_1, b_3 separates a_1 from the axis line, and hence a_1 lies below the axis line. Therefore B_3 and a_1 cannot intersect. This concludes the proof of Theorem 1.

9 Degenerate Cases

We will now justify our general-position assumption that we introduced in Sect. 3 stating that no vertex of one polygon lies on the boundary of the other polygon. That is, we will show that one can get rid of such degenerate cases without decreasing the number of intersection points. There are two types of degenerate cases:

vertex-vertex: Two vertices from different polygons coincide.

vertex-side: A vertex of one polygon lies on a side of the other polygon.

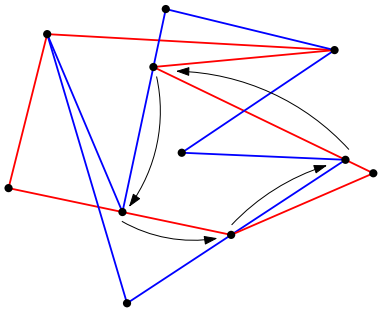


Fig. 26 The dependence graph may contain cycles

We will first use a global perturbation governed by a parametric system of linear equations to get rid of vertex-vertex degeneracies. Vertex-side degeneracies can then be treated locally by shifting sides one by one.

9.1 Coincident Vertices

Finding a good perturbation is delicate because a side may be involved in multiple degeneracies. We will simultaneously perturb all vertex-vertex degeneracies while leaving vertex-side degeneracies intact.

For each vertex-side degeneracy, some point A of one polygon lies on a side BC of the other polygon: There is a ratio $\lambda = \lambda_{ABC} \in (0, 1)$ such that

$$A = \lambda B + (1 - \lambda)C. \tag{1}$$

Our perturbation will maintain this relation, with the same value λ. We say that A *depends* on B and C. We can set up a *dependence graph* that has arcs (A, B) and (A, C), for each such vertex-side degeneracy. This graph may contain cycles. The polygon vertices that don't depend on other vertices are called *free*.

We specify the movement of the free vertices as follows: Free vertices that are not involved in a degeneracy stay where they are. For each pair (P_i, Q_j) of coincident vertices, we leave Q_j stationary but we move $P_i = P_i(\alpha) := P_i^0 + \alpha E_i$ linearly, with constant speed in an appropriate direction E_i. We choose each direction E_i independently so that the two polygons have at least one nondegenerate intersection in the neighborhood of the original configuration, assuming that all other points stay where they are.

Theorem 17 below shows that we can perform such a motion in a continuous way while simultaneously maintaining all vertex-side degeneracies (1).

In this theorem, we consider a general system of finitely many points $A_1, A_2, \ldots, A_N$, which are partitioned into *dependent* points and *free* points. We denote the indices of the dependent points by $D \subseteq [N]$ and those of the free points by $F = [N] \setminus D$. Each dependent point A_i is a convex combination of other points:

$$A_i = \sum_{j \in [N] \setminus \{i\}} \lambda_{ij} A_j \tag{2}$$

with

$$\sum_{j \in [N] \setminus \{i\}} \lambda_{ij} = 1,$$

whereas each free point A_i is given "externally", depending linearly on some parameter α:

$$A_i = A_i^0 + \alpha E_i, \qquad i \in F. \tag{3}$$

We will use the following simple observation.

Observation 15 *If* (2) *holds for a dependent point* A_i*, then either*

(i) *all points* A_j *on which* A_i *depends coincide with* A_i*, or*
(ii) *some point* A_j *on which* A_i *depends is strictly larger than* A_i*, in the lexicographic order according to the coordinates* (x, y).

The *dependence graph* contains a node for every point A_i, and contains the arc (i, j) whenever $\lambda_{ij} > 0$. The next lemma and theorem hold for a general system of the form (2)–(3), under the following assumption about the dependence graph: From every dependent point, one must be able to reach a free point. In other words, there must be no component of dependent points that depend only on each other. It is easy to show that this condition is fulfilled in our setting, see Lemma 18.

Lemma 16 *Suppose that from every dependent point, one can reach a free point in the dependence graph. Then, in every solution of the system of equations* (2)*, all points lie in the convex hull of the free points.*

Proof We start with some dependent point A_i. Our first goal is to reach a free point that is lexicographically at least as large as A_i. We apply Observation 15 to A_i. In case (i), we repeatedly apply Observation 15 to all points A_j on which A_i depends. Eventually, we conclude that all points that are reachable from A_i coincide with A_i, or we get a point A_k that is lexicographically strictly larger. In the first case, there is, among the points that coincide with A_i, a free point, by assumption. In the second case, we repeat the process with A_k until we eventually arrive at a free point.

Thus, we conclude that the x-coordinate of A_i is not larger than the largest x-coordinate of the free points. Since the proof is not affected by rotations, (or in fact, arbitrary affine transformations) we obtain the claimed result. □

Theorem 17 *Suppose that from every dependent point, one can reach a free point in the dependence graph. If the system of equations* (2)–(3) *has a solution for* $\alpha = 0$*, then it has a unique solution for all* α*, and the solution depends linearly on* α.

Proof The system of equations (2)–(3) actually decomposes into two independent systems of linear equations: one for the x-coordinates, and one for the y-coordinates. (In d dimensions, there would be d such systems.)

If we substitute the explicit values from (3) into (2), we get a $|D| \times |D|$ system of linear equations with parametric right-hand sides where the unknowns are the x-coordinates (resp., y-coordinates) of the $|D|$ dependent points. We are done if we show

that the coefficient matrix of this system, whose entries are the coefficients λ_{ij}, has non-zero determinant. A system of linear equations with zero determinant has either no solution or infinitely many solutions that form an affine subspace. Let us look at the set of solutions for $\alpha = 0$. We can exclude the first possibility, because we have assumed a solution for $\alpha = 0$. To exclude the second possibility, assume that there are infinitely many solutions. Since they form an affine subspace, there must be solutions where some coordinate of some point takes arbitrarily high values. This contradicts Lemma 16. □

In order to apply Theorem 17, we have to show that the assumption is fulfilled in our system:

Lemma 18 *In the dependence graph derived from the system of dependencies* (1) *describing the vertex-side degeneracies, one can reach a free point from every dependent point.*

Proof We look at our given (initial) configuration. We take some dependent point A_i and apply Observation 15 to it. We know that the points on which A_i depends don't coincide with A_i, and thus we find a lexicographically strictly larger point. Repeating the argument as long as necessary, we must eventually arrive at a free point that is reachable from A_i. □

One final point: We have chosen each movement $P_i = P_i^0 + \alpha E_i$ of a free point that is involved in a vertex-vertex degeneracy in such a way that there are no degeneracies in the vicinity of P_i, *assuming that the other points do not move*. In reality, the other endpoints of the four involved sides may move on linear trajectories. However, as we start from coinciding points, the movement of the other endpoints has a lower-order effect. To see this, consider the marked distance from Q_j to the intersection between $Q_{j-1}Q_j$ and P_iP_{i-1} in Fig. 27. P_i and Q_j are free, and Q_j remains stationary. If P_i moves linearly as αE_1, the marked distance changes like $k\alpha + O(\alpha^2)$ for some constant $k > 0$ if all other points remain fixed, see Fig. 27b: The distance would have a linear growth of the form $k\alpha$ if the sides would undergo a parallel shift, but in fact the sides are rotated. If one or both endpoints of an edge move by a distance $O(\alpha)$, the rotation angle is bounded by $O(\alpha)$. Since the leading side αE_i of the shaded triangle is of length $O(\alpha)$, the effect of the rotation on the side lengths of the triangle is of order $O(\alpha^2)$. (This can be checked by using the sine law.) An additional linear movement of P_{i-1} (see Fig. 27c) or Q_{j-1} by a distance $O(\alpha)$ will cause an additional change of direction by $O(\alpha)$. Still, the side lengths of the triangle change only by $O(\alpha^2)$. Thus, for small enough α, the linear growth $k\alpha$ dominates, and there will be no degenerate situations.

The morphing procedure goes back to Floater and Gotsman [7]. The idea for proving unique solvability in Theorem 17 was used, in a different context, in [4].

9.2 Vertex-Side Degeneracies

Once there are no more coincident vertices, it is easy to get rid of the remaining degeneracies by looking at the sides one by one, see Fig. 28.

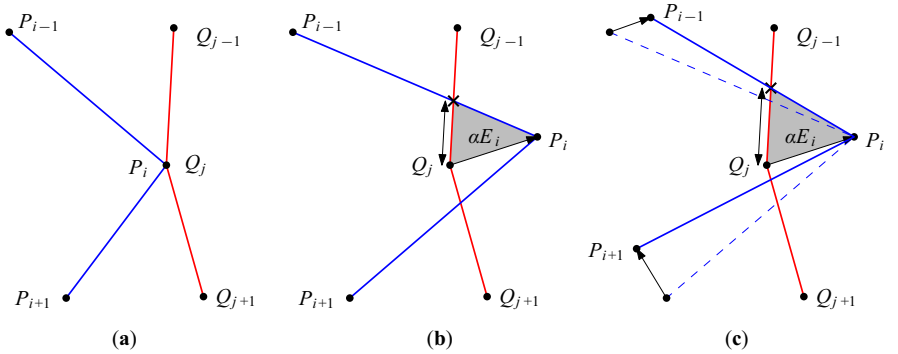


Fig. 27 Lower-order effects due to movement of endpoints

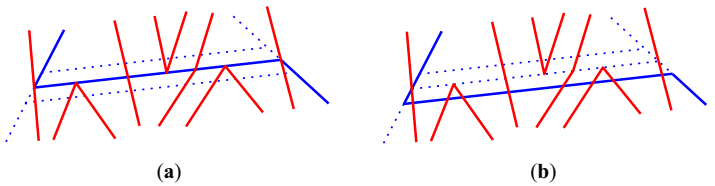


Fig. 28 Perturbing a blue polygon side to get rid of collinear vertices without decreasing the number of intersection points: **a** before, **b** after

If we perform a small parallel shift of a side, an intersection point may disappear or multiply into two intersection points, depending on the direction of the shift. However, the average number of intersections over the two possible directions is always the same as before. Thus, there exists a small parallel shift that does not decrease the number of intersection points.

9.3 The Combinatorial Counting Model

Instead of counting intersection points, one could count pairs of sides that intersect. A vertex-side degeneracy is then counted twice, and a vertex-vertex degeneracy is counted four times. We leave this variation for future research.

In [3], a more refined counting model is used: a further distinction is made for vertex-side degeneracies, in accordance with the number of intersection points that can be achieved by perturbation: if one polygon *touches* the other (i.e., two adjacent sides lie on the same side of the side of the other polygon), the intersection is counted twice. If one polygon *crosses* the other, it is counted as one intersection. In vertex-vertex degeneracies, depending on the angles in which the sides meet, the number of intersections that one can get after perturbation can be any number between 1 and 4. With this type of accounting, our results carry over.

10 Discussion

We have shown that the number of intersections of a simple n-gon and a simple m-gon, for odd n and m, is at most $mn - (m + n) + C$, for some constant C. This bound matches the best lower bound $mn - (m + n) + 3$ up to a constant. However, our constant C is huge, and we have not made much effort to improve it. For example, it can be improved by replacing Lemma 10 with a less general result which suffices for our purposes. Recall that we use Lemma 10 to obtain a transitive sub-tournament in a tournament whose vertices represent pairwise disjoint segments and whose edges represent stabbing relations. Instead of the logarithmic-size transitive sub-tournament obtained using Lemma 10, one can get a polynomial-size transitive sub-tournament using Dilworth's Theorem as follows (see [9] for a similar argument): Let s and s' be two disjoint segments such that s stabs s' or vice versa. Let $s \prec_1 s'$ (resp., $s \prec_2 s'$) if s stabs s' such that the right extension of s intersects s' and the slope of s is smaller (resp., larger) than the slope of s'. Similarly, let $s \prec_3 s'$ (resp., $s \prec_4 s'$) if s stabs s' such that the left extension of s intersects s' and the slope of s is smaller (resp., larger) than the slope of s'. Every pair of segments is comparable by exactly one of these four relations and this implies that each relation $\prec_i$ is transitive. From Dilworth's Theorem it follows that the above-mentioned tournament has $N^{1/4}$ vertices such that the relation between each pair of their corresponding segments is $\prec_1$, or $N^{3/4}$ vertices such that the relation between each pair of their corresponding segments is not $\prec_1$ (where N is the size of the tournament). In the second case we can apply Dilworth's Theorem once or twice more in a similar way, and conclude that there are always $i \in \{1, 2, 3, 4\}$ and $N^{1/4}$ vertices such that the relation between each pair of their corresponding segments is $\prec_i$. Hence, the tournament has a transitive sub-tournament of size $N^{1/4}$.[1]

This gives $C < (2^{67})^4 = 2^{268}$, which is obviously still very far from the truth, so determining the exact answer for this very basic question remains an interesting open problem. We believe that our approach to bound the number of connected components of the disjointness graph is a right approach, that is, we conjecture:

Conjecture 2 *Suppose that P and Q are simple polygons with m and n sides, respectively, such that m and n are odd numbers. Then there are at most three connected components in the disjointness graph of the sides of P and Q.*

If both polygons are allowed to self-intersect, then a simple construction shows that the upper bound $mn - \max\{m, n\}$ is the correct one [3, 8]. However, if only one of the polygons is allowed to self-intersect then we do not know of such a construction.

Our results should carry over to *pseudolinear* polygons, that is, polygons whose sides can be extended to an arrangement of pseudolines, but we did not check the details of this extension. In fact, we already used the freedom to draw curved edges in our schematic drawings, like in Figs. 11b, 14, and 15. Moreover, the computer verifications mentioned below in Sect. 10.2 use none of the properties of straight

[1] It may be interesting to determine the maximum size of such a sub-tournament that one can guarantee for every set of N disjoint segments where for each pair of them one segment stabs the other. A recursive construction starting with three segments whose corresponding tournament is a directed cycle, then replacing each segment with the same "flattened" construction shows that this number is at most $O(N^{\log_3 2}) \approx O(N^{0.631})$.

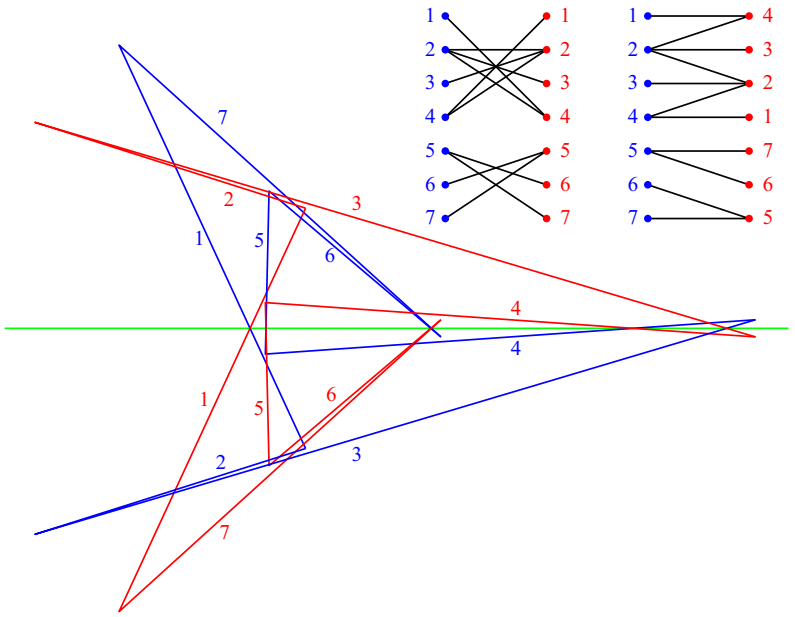


Fig. 29 A best lower bound construction for $m = n = 7$. It is symmetric with respect to the horizontal green axis. In the left drawing of the disjointness graph, the nodes are arranged in the natural order as the edges appear along the polygon boundaries. The right drawing makes the structure as a forest clearer

lines, as opposed to pseudoline arrangements. Obviously, Conjectures 1 and 2 extend to pseudolinear polygons as well.

10.1 Alternative Lower-Bound Constructions

Studying the best lower bound constructions (conjectured to be optimal) may provide insights that may help to improve the upper bound. For example, the disjointness graph of the construction on Fig. 1c has three components: one edge and two stars. If we replace a side that corresponds to a leaf in the disjointness graph by a narrow zigzag of three sides, the number of intersections increases by $2(m-1)$ or $2(n-1)$, depending of which polygon side is replaced. This way we can get a best lower bound construction for $m, n+2$ or $m+2, n$ from a best lower bound construction for odd m and n. Therefore, there are many ways to create best lower bound constructions. Note that each step as above replaces one leaf by three leaves in the disjointness graph, and thus repeating it would yield the existence of large-degree vertices and constantly many vertices that dominate the disjointness graph. Thus, it is tempting to conjecture that an optimal construction must posses these properties apart from having three connected components (which are necessarily trees). However, this does not seem to be true. We have a construction for $m = n = 7$ which is significantly different from the one on Fig. 1c and whose disjointness graph has maximum degree 3, see Fig. 29. We believe that this construction can be extended for greater values of $m = n$ while maintaining maximum degree 3, but the figure becomes too complicated to draw, so we do not go into further details.

10.2 Computer-Supported Proofs

We have independently confirmed Lemma 7 concerning the intersection structure of pairs of associated pairs of sides, which is our principal workhorse, by a computer enumeration.

Lemma 7 involves eight polygon sides. Thus, we enumerated all 158830 arrangements of eight lines. Actually, we enumerated pseudoline arrangements, but it is known that all pseudoline arrangements with less than nine lines are stretchable to line arrangements. For each arrangement, we tried all possible ways of assigning the sides p_i, p_{i+1}, p_j, p_{j+1} and $q_{i'}$, $q_{i'\pm1}$, $q_{j'}$, $q_{j'\pm1}$ (in the notation of Lemma 7) to these lines. Since there are symmetries in the configuration of the eight sides, we did not have to try all $8! = 40320$ assignments, but only 840. Once this assignment is given, we know where the vertices between consecutive sides, like p_i, p_{i+1}, are in the arrangement. This specifies one endpoint of each side. It is then easy to check whether it is possible to realize the required intersection pattern of a perfect matching among the eight involved sides. We first took into account the *required* intersections among the sides, for example, between p_i and $q_{j'}$. On each line, these required intersection points must lie on one side of the vertex that we already know. If this is the case, we get a minimum range on the line that the side must cover. We then check if all *forbidden* intersections (for example, between p_i and $q_{i'}$, or between p_i and p_j) are avoided.

After a few days of computing (and several weeks of programming), the result was that there were, up to isomorphism, 14 configurations with the required intersection pattern. In 13 of them, some associated pair did not conform to the hooking/hooked pattern of Lemma 5. The remaining configuration is the one in Fig. 6.

We had some hope to attack the case of *three* pairs of associated pairs of sides and obtain a more powerful statement than Proposition 14. However, since there are too many (pseudo-)line arrangements with 12 lines, this is out of reach for the simple approach outlined above.

10.3 The Union of the Polygon Areas

Instead of polygon boundaries, we can consider polygons as regions and ask about the complexity of the union of the polygon areas. The bound on boundary intersections gives an upper bound on the boundary of this union: Every boundary intersection contributes to the complexity of the union, and in the worst case all the $m+n$ vertices of the two polygons contribute to the complexity of the boundary in addition.

In the examples of Fig. 1, this worst case arises: all polygon vertices contribute to the boundary of the union of the polygon areas. This gives a lower bound of $mn-3$ for the maximum complexity of the union of an m-gon and an n-gon and it is conjectured that $mn-3$ is in fact the exact value. Theorem 1 implies the upper bound $mn+C$ for some absolute constant C.

One can also ask about the *intersection* of the polygon areas. Here, the relation to our problem is not clear.

Acknowledgements We thank the reviewers for helpful suggestions.

Funding Open access funding provided by ELKH Alfréd Rényi Institute of Mathematics. Eyal Ackerman: The main part of this work was performed during a visit to Freie Universität Berlin which was supported by the Freie Universität Alumni Program. Balázs Keszegh: Research supported by the Lendület program of the Hungarian Academy of Sciences (MTA), under the grant LP2017-19/2017, by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, by the National Research, Development and Innovation Office—NKFIH under the grant K 116769, K 132696, and FK 132060, and by the ÚNKP-21-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

References

1. Ackerman, E., Keszegh, B., Rote, G.: An almost optimal bound on the number of intersections of two simple polygons. In: 36th International Symposium on Computational Geometry. Leibniz Int. Proc. Inform., vol. 164, # 1. Leibniz-Zent. Inform., Wadern (2020)
2. Černý, J., Kára, J., Král', D., Podbrdský, P., Sotáková, M., Šámal, R.: On the number of intersections of two polygons. Comment. Math. Univ. Carolin. **44**(2), 217–228 (2003)
3. Dillencourt, M.B., Mount, D.M., Saalfeld, A.: On the maximum number of intersections of two polyhedra in 2 and 3 dimensions. In: 5th Canadian Conference on Computational Geometry (Waterloo 1993), pp. 49–54. University of Waterloo (1993)
4. Dujmović, V., Frati, F., Gonçalves, D., Morin, P., Rote, G.: Every collinear set in a planar graph is free. Discrete Comput. Geom. **65**(4), 999–1027 (2021)
5. Erdős, P., Moser, L.: On the representation of directed graphs as unions of orderings. Magyar Tud. Akad. Mat. Kutató Int. Közl. **9**, 125–132 (1964)
6. Erdös, P., Szekeres, G.: A combinatorial problem in geometry. Compos. Math. **2**, 463–470 (1935)
7. Floater, M.S., Gotsman, C.: How to morph tilings injectively. J. Comput. Appl. Math. **101**(1–2), 117–129 (1999)
8. Günther, F.: The maximum number of intersections of two polygons (2012). arXiv:1207.0996
9. Pach, J., Törőcsik, J.: Some geometric applications of Dilworth's theorem. Discrete Comput. Geom. **12**(1), 1–7 (1994)