

Improved Bounds for Covering Paths and Trees in the Plane

Ahmad Biniáz ✉

School of Computer Science, University of Windsor, Canada

Abstract

A *covering path* for a planar point set is a path drawn in the plane with straight-line edges such that every point lies at a vertex or on an edge of the path. A *covering tree* is defined analogously. Let $\pi(n)$ be the minimum number such that every set of n points in the plane can be covered by a noncrossing path with at most $\pi(n)$ edges. Let $\tau(n)$ be the analogous number for noncrossing covering trees. Dumitrescu, Gerbner, Keszegh, and Tóth (Discrete & Computational Geometry, 2014) established the following inequalities:

$$\frac{5n}{9} - O(1) < \pi(n) < \left(1 - \frac{1}{601080391}\right)n, \quad \text{and} \quad \frac{9n}{17} - O(1) < \tau(n) \leq \left\lfloor \frac{5n}{6} \right\rfloor.$$

We report the following improved upper bounds:

$$\pi(n) \leq \left(1 - \frac{1}{22}\right)n, \quad \text{and} \quad \tau(n) \leq \left\lfloor \frac{4n}{5} \right\rfloor.$$

In the same context we study rainbow polygons. For a set of colored points in the plane, a *perfect rainbow polygon* is a simple polygon that contains exactly one point of each color in its interior or on its boundary. Let $\rho(k)$ be the minimum number such that every k -colored point set in the plane admits a perfect rainbow polygon of size $\rho(k)$. Flores-Peñaloza, Kano, Martínez-Sandoval, Orden, Tejel, Tóth, Urrutia, and Vogtenhuber (Discrete Mathematics, 2021) proved that $20k/19 - O(1) < \rho(k) < 10k/7 + O(1)$. We report the improved upper bound of $\rho(k) < 7k/5 + O(1)$.

To obtain the improved bounds we present simple $O(n \log n)$ -time algorithms that achieve paths, trees, and polygons with our desired number of edges.

2012 ACM Subject Classification Theory of computation → Computational geometry; Mathematics of computing → Discrete mathematics

Keywords and phrases planar point sets, covering paths, covering trees, rainbow polygons

Digital Object Identifier 10.4230/LIPIcs.SoCG.2023.19

Related Version *Full Version*: <https://arxiv.org/abs/2303.04350>

Funding Research supported by NSERC.

1 Introduction

Traversing a set of points in the plane by a polygonal path possessing some desired properties has a rich background. For example the famous traveling salesperson path problem asks for a polygonal path with minimum total edge length [6, 29]. In recent years there has been an increased interest in paths with properties such as being noncrossing [2, 9], minimizing the longest edge length [8], maximizing the shortest edge length [4], minimizing the total or the largest turning angle [1, 18], and minimizing the number of turns (which is the same as minimizing the number of edges) [15, 30] to name a few.

The main focus of this paper is polygonal paths with a small number of edges. It is related to the classical nine dots puzzle which asks for covering the vertices of a 3×3 grid by a polygonal path with no more than 4 segments (Figure 1). It appears in Sam Loyd’s Cyclopedia of Puzzles from 1914 [27].



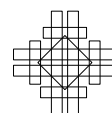
© Ahmad Biniáz;
licensed under Creative Commons License CC-BY 4.0
39th International Symposium on Computational Geometry (SoCG 2023).

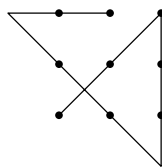
Editors: Erin W. Chambers and Joachim Gudmundsson; Article No. 19; pp. 19:1–19:15



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany





■ **Figure 1** The nine dots puzzle.

Let P be a set of n points in the plane. A *spanning path* for P is a path drawn in the plane with straight-line edges such that every point of P lies at a vertex of the path and every vertex of the path lies at a point of P . In other words, it is a Hamiltonian path which has exactly $n - 1$ edges. The path in the figure above is not a spanning path because two of its vertices do not lie on given points. A *covering path* for P is a path drawn in the plane with straight-line edges such that every point of P lies at a vertex or on an edge of the path. A vertex of a covering path can be any point in the plane (not necessarily in P). The path in the figure above is a covering path with 4 edges. With these definitions, any spanning path is also a covering path, but a covering path may not be a spanning path. A *covering tree* for P is defined analogously as a tree drawn in the plane with straight-line edges such that every point of P lies at a vertex or on an edge of the tree. A covering path or a tree is called *noncrossing* if its edges do not cross each other. The edges of covering paths and trees are also referred to as *links* in the literature [5].

Covering paths and trees have received considerable attention in recent years, see e.g. [5, 15, 25]. In particular covering paths with a small number of edges find applications in robotics and heavy machinery for which turning is an expensive operation [30]. Covering trees with a small number of edges are useful in red-blue separation [20] and in constructing rainbow polygons [19]. In 2010 F. Morić [14] and later Dumitrescu, Gerbner, Keszegh, and Tóth [15] raised many challenging questions about covering paths and trees. Specifically they asked the following two questions which are the main topics of this paper. As noted in [14], analogous questions were asked by E. Welzl in Gremo's Workshop on Open Problems 2011.

1. What is the minimum number $\pi(n)$ such that every set of n points in the plane can be covered by a noncrossing path with at most $\pi(n)$ edges?
2. What is the minimum number $\tau(n)$ such that every set of n points in the plane can be covered by a noncrossing tree with at most $\tau(n)$ edges?

For both $\pi(n)$ and $\tau(n)$, a trivial upper bound is $n - 1$ (which comes from the existence of a noncrossing spanning path) and a trivial lower bound is $\lceil \frac{n}{2} \rceil$ (because if no three points are collinear then each edge covers at most two points). In 2014, Dumitrescu et al. [15] established, among other interesting results, the following nontrivial bounds:

$$\frac{5n}{9} - O(1) < \pi(n) < \left(1 - \frac{1}{601080391}\right)n, \quad \text{and} \quad \frac{9n}{17} - O(1) < \tau(n) \leq \left\lfloor \frac{5n}{6} \right\rfloor.$$

The following is a related question that has recently been raised by Flores-Peñaloza, Kano, Martínez-Sandoval, Orden, Tejel, Tóth, Urrutia, and Vogtenhuber [19] in the context of rainbow polygons. For a set of colored points in the plane, a *rainbow polygon* is a simple polygon that contains at most one point of each color in its interior or on its boundary. A rainbow polygon is called *perfect* if it contains exactly one point of each color. The *size* of a polygon is the number of its edges (which is the same as the number of its vertices).

3. What is the minimum number $\rho(k)$ (known as *the rainbow index*) such that every k -colored point set in the plane, with no three collinear points, admits a perfect rainbow polygon of size $\rho(k)$?

Question 3 is related to covering trees in the sense that (as we will see later in Section 4) particular covering trees could lead to better upper bounds for $\rho(k)$. Flores-Peñaloza et al. [19] established the following inequalities:

$$\frac{20k}{19} - O(1) < \rho(k) < \frac{10k}{7} + O(1).$$

The upper bounds on $\pi(n)$, $\tau(n)$, and $\rho(n)$ are universal (i.e., any point set admits these bounds) and they are obtained by algorithms that achieve paths, trees, and polygons of certain size [15, 19]. The lower bounds, however, are existential (i.e., there exist point sets that achieve these bounds) and they are obtained by the same point set that is exhibited in [15]. Perhaps there should be configurations of points that achieve better lower bounds for each specific number.

1.1 Our contributions

Narrowing the gaps between the lower and upper bounds for $\pi(n)$, $\tau(n)$, and $\rho(n)$ are open problems which are explicitly mentioned in [15, 19]. In this paper we report the following improved upper bounds for the three numbers:

$$\pi(n) \leq \left(1 - \frac{1}{22}\right)n, \quad \tau(n) \leq \left\lceil \frac{4n}{5} \right\rceil, \quad \text{and} \quad \rho(k) < \frac{7k}{5} + O(1).$$

The new bounds for $\pi(n)$ and $\tau(n)$ are the first improvements in 8 years. To obtain these bounds we present algorithms that achieve noncrossing covering paths, noncrossing covering trees, and rainbow polygons with our desired number of edges. The algorithms are simple and run in $O(n \log n)$ time where n is the number of input points. The running time is optimal for paths because computing a noncrossing covering path has an $\Omega(n \log n)$ lower bound [15]. A noncrossing covering tree, however, can be computed in $O(n)$ time by taking a spanning star. We extend our path algorithm and achieve an upper bound of $(1 - \frac{1}{22})n + 2$ for noncrossing *covering cycles*. This is a natural variant of the traveling salesperson tour problem with the objective of minimizing the number of links, which is NP-hard [5].

Our algorithms share some similarities with previous algorithms in the sense that both are iterative and use the standard plane sweep technique which scans the points from left to right. However, to achieve the new bounds we employ new geometric insights and make use of convex layers and the Erdős-Szekeres theorem [16].

Regardless of algorithmic implications, our results are important because they provide new information on universal numbers $\pi(n)$, $\tau(n)$, and $\rho(n)$ similar to the crossing numbers [3, 13, 22], the size of crossing families (pairwise crossing edges) [28], the Steiner ratio [6, 23], and other numbers and constants studied in discrete geometry (such as [8, 10, 17]).

Remark. Collinear points are beneficial for covering paths and trees as they usually lead to paths and trees with fewer edges. To avoid the interruption of our arguments we first describe our algorithms for point sets with no three collinear points. In the end we show how to handle collinearities.

1.2 Related problems and results

If we drop the noncrossing property, Dumitrescu et al. [15] showed that every set of n points in the plane admits a (possibly self-crossing) covering path with $n/2 + O(n/\log n)$ edges. Covering paths have also been studied from the optimization point of view. The problem of

computing a covering path with minimum number of edges for a set of points in the plane (also known as the *minimum-link* covering path problem and the *minimum-bend* covering path problem) is shown to be NP-hard by Arkin et al. [5]. Stein and Wagner [30] presented an $O(\log z)$ -approximation algorithm where z is the maximum number of collinear points.

Keszegh [25] determined exact values of $\pi(n)$ and $\tau(n)$ for vertices of square grids. The axis-aligned version of covering paths is also well-studied and various lower bounds, upper bounds, and approximation algorithms are presented to minimize the number of edges of such paths; see e.g. [7, 12, 24]. Covering trees are studied also in the context of separating red and blue points in the plane [20]. The problem of covering points in the plane with minimum number of lines is another related problem which is also well-studied, see e.g. [11, 21, 26].

For problems and results related to rainbow polygons we refer the reader to the paper of Flores-Peñaloza et al. [19]. In particular, they determine the exact rainbow indices for small values of k by showing that $\rho(k) = k$ for $k \in \{3, 4, 5, 6\}$ and $\rho(7) = 8$.

1.3 Preliminaries

For two points p and q in the plane we denote by $\ell(p, q)$ the line through p and q , and by pq the line segment with endpoints p and q . For two paths δ_1 and δ_2 , where δ_1 ends at the same vertex at which δ_2 starts, we denote their concatenation by $\delta_1 \oplus \delta_2$.

A point set P is said to be in *general position* if no three points of P are collinear. We denote the convex hull of P by $\text{CH}(P)$. A set K of k points in the plane in convex position, with no two points on a vertical line, is a *k-cap* (resp. a *k-cup*) if all points of K lie on or above (resp. below) the line through the leftmost and rightmost points of K . A classical result of Erdős and Szekeres [16] implies that every set of at least $\binom{2k-4}{k-2} + 1$ points in the plane in general position, with no two points on a vertical line, contains a *k-cap* or a *k-cup*. This bound is tight in the sense that there are point sets of size $\binom{2k-4}{k-2}$ that do not contain any *k-cap* or *k-cup* [16].

2 Noncrossing covering paths

In this section we prove that $\pi(n) \leq (1 - 1/22)n$. We start by the following folklore result on the existence of noncrossing polygonal paths among points in the plane; see e.g. [15, 20].

► **Lemma 1.** *Let P be a set of points in the plane in the interior of a convex region C , and let p and q be two points on the boundary of C . Then $P \cup \{p, q\}$ admits a noncrossing spanning path with $|P| + 1$ edges such that its endpoints are p and q , and its relative interior lies in the interior of C .*

In fact the spanning path that is obtained by Lemma 1 is a noncrossing covering path for $P \cup \{p, q\}$ and it lies in the convex hull of $P \cup \{p, q\}$. The following lemma shows that any set of 23 points can be covered by a noncrossing path with 21 edges.

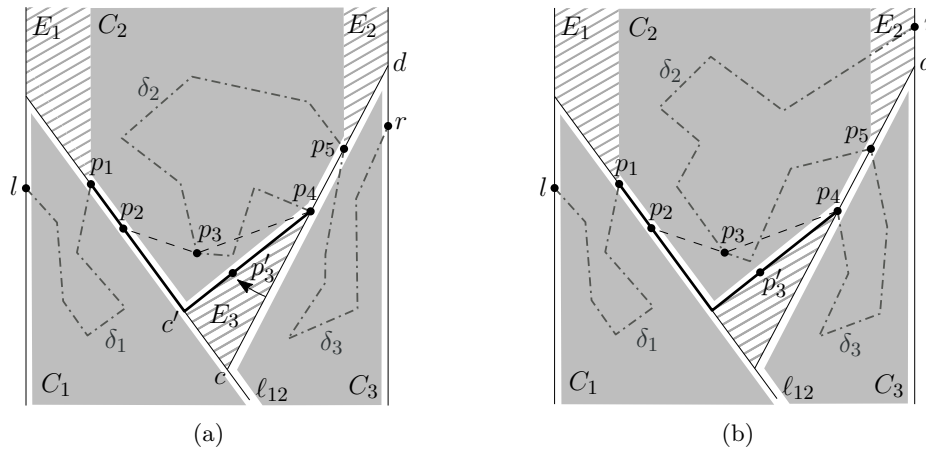
► **Lemma 2.** *Let P be a set of at least 23 points in the plane such that no two points have the same x -coordinate. Let H be the vertical strip bounded by the vertical lines through the leftmost and rightmost points of P . Then there exists a noncrossing covering path for P with $|P| - 2$ edges that is contained in H and its endpoints are the leftmost and rightmost points of P .*

Proof. Our proof is constructive. Let l and r be the leftmost and rightmost points of P , respectively. Let $P' = P \setminus \{l, r\}$, and notice that $|P'| \geq 21$. We assume that P' is in general position. In the end of the proof we briefly describe how to handle collinearities. By the

result of [16] the set P' has a 5-cap or a 5-cup. After a suitable reflection we may assume that it has a 5-cup K with points p_1, p_2, p_3, p_4, p_5 from left to right, as in Figure 2(a). Among all 5-cups in P' we may assume that K is one for which p_1 is the leftmost possible point. Also among all such 5-cups (with leftmost point p_1) we may assume that K is the one for which p_5 is the rightmost possible point. This choice of K implies that the region that is the intersection of H with the halfplane above $\ell(p_1, p_2)$ and the halfplane to the left of the vertical line through p_1 is empty of points of P' ; this region is denoted by E_1 in Figure 2(a). Similarly the region that is the intersection of H with the halfplane above $\ell(p_4, p_5)$ and the halfplane to the right of the vertical line through p_5 is empty of points of P' ; this region is denoted by E_2 in Figure 2(a).

For brevity let $\ell_{12} = \ell(p_1, p_2)$ and $\ell_{45} = \ell(p_4, p_5)$. We distinguish two cases: (i) l lies below ℓ_{12} or r lies below ℓ_{45} , and (ii) l lies above ℓ_{12} and r lies above ℓ_{45} .

(i) In this case we may assume, up to symmetry, that l lies below ℓ_{12} as in Figure 2. Let c be the intersection point of ℓ_{12} with ℓ_{45} , and d be the intersection point of ℓ_{45} with the right boundary of H . Since K is a cup, c lies below K and hence in H . Consider the ray emanating from p_4 and passing through c . Rotate this ray clockwise around p_4 and stop as soon as hitting a point in the triangle $\Delta p_2 p_4$; see Figure 2(a). Notice that such a point exists because p_3 is in $\Delta p_2 p_4$. Denote this first hit by p'_3 (it might be the case that $p'_3 = p_3$). Then p_1, p_2, p'_3, p_4, p_5 is a 5-cup which we denote by K' (again, it might be the case that $K' = K$). Let c' be the intersection point of the rotated ray with ℓ_{12} . Our choice of p'_3 implies that the triangle $\Delta c p_4 c'$ is empty, i.e. its interior has no points of P ; this triangle is denoted by E_3 in Figure 2(a).

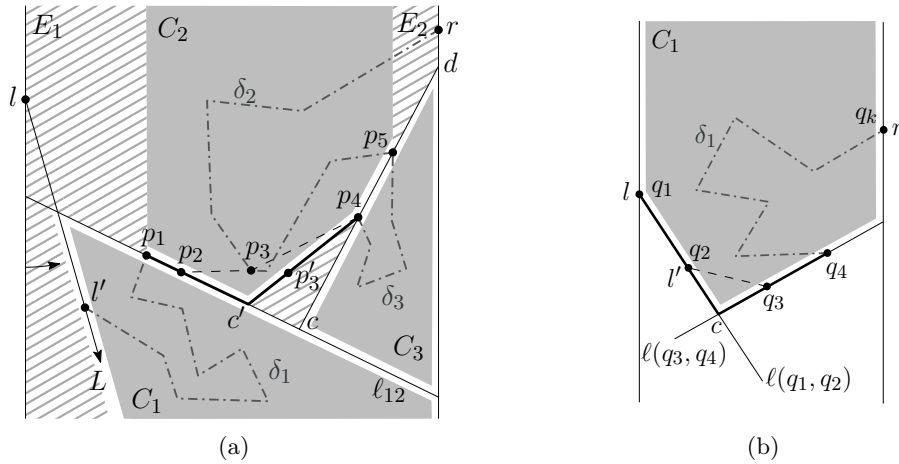


■ **Figure 2** Illustration of the proof of Lemma 2. (a) l lies below ℓ_{12} and r lies below ℓ_{45} . (b) l lies below ℓ_{12} and r lies above ℓ_{45} .

The points of P' lie in the interior or on the boundary of three convex regions C_1, C_2, C_3 as depicted in Figures 2(a) and 2(b). The region C_1 is the intersection of H and the halfplane below ℓ_{12} . The region C_3 is the intersection of H and the halfplane above ℓ_{12} and the halfplane below ℓ_{45} . The region C_2 is the intersection of H and five halfplanes (the halfplanes above the lines $\ell_{12}, \ell_{45}, \ell(c', p_4)$, the halfplane to the right of the vertical line through p_1 , and the halfplane to the left of the vertical line through p_5). Let P_i be the set of points of P in the interior (but not on the boundary) of each C_i . Then $P_1 \cup P_2 \cup P_3 = P \setminus \{l, p_1, p_2, p'_3, p_4, p_5, r\}$, and thus $|P_1| + |P_2| + |P_3| = |P| - 7$.

We construct a covering path for P as follows. The four points p_1, p_2, p'_3, p_4 can be covered by the path (p_1, c', p_4) which has two edges p_1c' and $c'p_4$. Let δ_1 be the noncrossing path with $|P_1| + 1$ edges that is obtained by applying Lemma 1 on (P_1, C_1, l, p_1) where l and p_1 play the roles of p and q in the lemma. We now consider two subcases.

- r lies below ℓ_{45} . In this case r is on the boundary of C_3 , as in Figure 2(a). Let δ_2 and δ_3 be the noncrossing paths with $|P_2| + 1$ and $|P_3| + 1$ edges that are obtained by applying Lemma 1 on (P_2, C_2, p_4, p_5) and (P_3, C_3, p_5, r) , respectively. By interconnecting these paths we obtain a noncrossing covering path $\delta_1 \oplus (p_1, c', p_4) \oplus \delta_2 \oplus \delta_3$ for P . This path has $(|P_1| + 1) + 2 + (|P_2| + 1) + (|P_3| + 1) = |P| - 2$ edges, and it lies in H .
- r lies above ℓ_{45} . In this case r is on the boundary of the convex region $C_2 \cup E_2$, as in Figure 2(b). Let δ_2 and δ_3 be the noncrossing paths obtained by applying Lemma 1 on $(P_2, C_2 \cup E_2, p_5, r)$ and (P_3, C_3, p_4, p_5) , respectively. Then $\delta_1 \oplus (p_1, c', p_4) \oplus \delta_3 \oplus \delta_2$ is a noncrossing covering path for P . This path has $|P| - 2$ edges, and it lies in H .



■ **Figure 3** Illustration of the proof of Lemma 2 where l lies above ℓ_{12} and r lies above ℓ_{45} . (a) $l' \neq p_1$, and (b) $l' = p_1$ and $r' = p_5$ (here $p_1 = q_2$ and $p_5 = q_{k-1}$).

(ii) In this case l lies above ℓ_{12} and r lies above ℓ_{45} . Let L and R be the downward rays emanating from l and r , respectively. Rotate L counterclockwise around l and stop as soon as hitting a point l' of P . Since E_1 is empty, l' is either p_1 or a point below ℓ_{12} ; see Figure 3(a). Rotate R clockwise around r and stop as soon as hitting a point r' of P . Since E_2 is empty, r' is either p_5 or a point below ℓ_{45} . We distinguish two subcases.

- $l' \neq p_1$ or $r' \neq p_5$. Up to symmetry we assume that $l' \neq p_1$ as depicted in Figure 3(a). Define c, c', d, p'_3 and the 5-cup K' as in case (i), and recall that the triangle $\Delta cp_4c'$ is empty. The points of P' lie in the interior or on the boundary of three convex regions C_1, C_2, C_3 as depicted in Figures 3(a). The region C_1 is the intersection of H and the halfplane below ℓ_{12} and the halfplane above $\ell(l, l')$. The regions C_2 and C_3 are defined as in case (i). Let P_i be the set of points of P in the interior (but not on the boundary) of each C_i . Then $P_1 \cup P_2 \cup P_3 = P \setminus \{l, l', p_1, p_2, p'_3, p_4, p_5, r\}$, and thus $|P_1| + |P_2| + |P_3| = |P| - 8$. We cover l and l' by the edge (l, l') and cover the four points p_1, p_2, p'_3, p_4 by the path (p_1, c', p_4) which has two edges. Let δ_1, δ_2 , and δ_3 be the noncrossing paths with $|P_1| + 1, |P_2| + 1$, and $|P_3| + 1$ edges obtained by applying Lemma 1 on (P_1, C_1, l', p_1) , $(P_2, C_2 \cup E_2, p_5, r)$, and (P_3, C_3, p_4, p_5) , respectively; see Figures 3(a). By interconnecting these paths we obtain a noncrossing covering path $(l, l') \oplus \delta_1 \oplus (p_1, c', p_4) \oplus \delta_3 \oplus \delta_2$ for P . This path has $1 + (|P_1| + 1) + 2 + (|P_3| + 1) + (|P_2| + 1) = |P| - 2$ edges, and it lies in H .

- $l' = p_1$ and $r' = p_5$. In this case the lower chain on the boundary of $\text{CH}(P)$ has at least 5 vertices, including l, l', r, r' , and a point in the triangle formed by L, R , and $\ell(p_1, p_5)$. Let $k \geq 5$ be the number of vertices of this chain. Let q_1, q_2, \dots, q_k denote the vertices of this chain that appear in this order from left to right, as in Figure 3(b). Then $q_1 = l, q_2 = l' = p_1, q_k = r$, and $q_{k-1} = r' = p_5$.

Let c be the intersection point of $\ell(q_1, q_2)$ and $\ell(q_3, q_4)$, which lies in H . Then, the four points q_1, q_2, q_3, q_4 can be covered by the path (q_1, c, q_4) . All points of P lie in the interior or on the boundary of a convex region C_1 that is the intersection of H with the halfplanes above $\ell(q_1, q_2)$ and $\ell(q_3, q_4)$; this region is shaded in Figure 3(b). Let P_1 be the points of P that lie in the interior (but not on the boundary) of C_1 . Then $P_1 = P \setminus \{q_1, q_2, q_3, q_4, q_k\}$ and $|P_1| = |P| - 5$. Let δ_1 be the covering path with $|P_1| + 1$ edges that is obtained by applying Lemma 1 on (P_1, C_1, q_4, q_k) where q_4 and q_k play the roles of p and q in the lemma. Then $(q_1, c, q_4) \oplus \delta_1$ is a noncrossing covering path for P . This path has $2 + (|P_1| + 1) = |P| - 2$ edges, and it lies in H .

This is the end of our proof (for P' being in general position).

One can simply adjust the above construction to work even if P' is not in general position. For the sake of completeness here we give a brief description of an alternative (and perhaps simpler) construction when P' has three or more collinear points. Let p_1, p_2, p_3 be three collinear points in P' from left to right and let ℓ_{13} be the line through these points. We choose p_1, p_2, p_3 in such a way that there is no point of P' on ℓ_{13} to the left of p_1 or to the right of p_3 . Up to symmetry we have two cases: (i) l lies on or above ℓ_{13} and r lies on or below ℓ_{13} , and (ii) both l and r lie below ℓ_{13} .

In case (i) we first obtain a path by applying Lemma 1 on l, p_1 and all points above ℓ_{13} . Then we connect p_1 and p_3 by one edge which also covers p_2 . Then we obtain another path by applying Lemma 1 on r, p_3 and all points below ℓ_{13} . This gives a covering path with $|P| - 2$ edges.

In case (ii) we start from l and walk on the vertices of $\text{CH}(P)$ in clockwise direction (and at the same time cover the visited vertices) and stop at the first vertex, say p_0 , for which the next vertex, say x , is on or above ℓ_{13} (it could be the case that $p_0 = l$). Denote the traversed path between l and p_0 by δ_l . First assume that x is above ℓ_{13} . We connect p_0 to x . Then we obtain a path by applying Lemma 1 on x, p_1 and all points above ℓ_{13} . Then we connect p_1 to p_3 by one edge which also covers p_2 . Then we extend the current path to a covering path for P by applying Lemma 1 on p_3, r and the remaining points below ℓ_{13} . Now assume that x is on ℓ_{13} , in which case $x = p_1$. If there is no point of P' above ℓ_{13} then we connect p_1 to p_3 by one edge and then extend it to a covering path for P by applying Lemma 1 on p_3, r and the remaining points below ℓ_{13} . Assume that there are points above ℓ_{13} . We repeat the above process from r by a counterclockwise walk on the vertices of $\text{CH}(P)$, and due to symmetry, assume that p_3 is the first visited vertex that lies on or above ℓ_{13} . Let p_4 denote the vertex of $\text{CH}(P)$ after p_3 . Notice that p_4 lies above ℓ_{13} . Let c be the intersection point of the lines $\ell(p_0, p_1)$ and $\ell(p_3, p_4)$. To obtain a covering path, we start with δ_l , connect its endpoint p_0 to c , and connect c to p_3 ; these two edges cover p_1 and p_4 . Then we continue by a path obtained from Lemma 1 applied on p_3, r and the remaining points. ◀

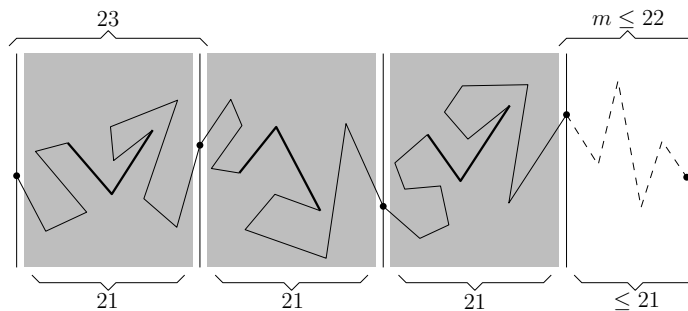
The following corollary, although very simple, will be helpful in the analysis of our algorithm.

► **Corollary 3.** *Let Q be a set of at least 22 points in the plane and let l be its leftmost point. Then there exists a noncrossing covering path for Q with $|Q| - 2$ edges that lies to the right of the vertical line through l and has l as an endpoint.*

Proof. We add a dummy point r to the right of all points in Q . Let $P = Q \cup \{r\}$. We obtain a noncrossing covering path δ for P with $|P| - 2$ edges by Lemma 2. Recall that r is an endpoint of δ . Also recall from the proof of Lemma 2 that in all cases r gets connected to δ by a path that is obtained from Lemma 1. No edge of this path has a point of P in its interior (even if two consecutive edges happen to be collinear, we treat them as two different edges). Thus, the edge of δ that covers r has no point in its interior. Therefore, by removing r and its incident edge from δ we obtain a covering path with $|Q| - 2$ edges for Q that satisfies the conditions of the corollary. ◀

► **Theorem 4.** *Every set of n points in the plane admits a noncrossing covering path with at most $\lceil 21n/22 \rceil - 1$ edges. Thus, $\pi(n) \leq (1 - 1/22)n$. Such a path can be computed in $O(n \log n)$ time.*

Proof. Let P be a set of n points in the plane. After a suitable rotation we may assume that no two points of P have the same x -coordinate. Draw vertical lines in the plane such that each line goes through a point of P , there are exactly 21 points of P between any pair of consecutive lines, no point of P lies to the left of the leftmost line, and at most 21 points of P lie to the right of the rightmost line; see Figure 4. Each pair of consecutive lines defines a vertical strip containing 23 points; 21 points in its interior and 2 points on its boundary (the point on the boundary of two consecutive strips is counted for both strips). For the 23 points in each strip we obtain a noncrossing covering path with 21 edges using Lemma 2. Each path lies in its corresponding strip and its endpoints are the two points on the boundary of the strip. By assigning to each strip the point on its left boundary, it turns out that for every 22 points we get a path with 21 edges.



■ **Figure 4** Illustration of the proof of Theorem 4.

Let m be the number of points on or to the right of the rightmost line, and notice that $m \leq 22$. We distinguish between two cases $m = 22$ and $m < 22$.

If $m = 22$ (in this case n is divisible by 22) then we cover these 22 points by a noncrossing path with 20 edges using Corollary 3. The union of this path and the paths constructed within the strips is a noncrossing covering path for P . The total number of edges in this path is $21n/22 - 1$.

If $m < 22$ then $m = n - 22\lfloor n/22 \rfloor$. In this case we cover the m points by an x -monotone path with $m - 1$ edges (dashed segments in Figure 4). Again, the union of this path and the paths constructed within the strips is a noncrossing covering path for P . The total number of edges in this path is $21\lfloor n/22 \rfloor + m - 1 = n - \lfloor n/22 \rfloor - 1 = \lceil 21n/22 \rceil - 1$.

Each call to Lemma 2 and Corollary 3 takes constant time. Therefore, after rotating and sorting the points in $O(n \log n)$ time, the rest of the algorithm takes linear time. ◀

Our path construction in Theorem 4 achieves a similar bound for covering cycles.

► **Corollary 5.** *Every set of n points in the plane admits a noncrossing covering cycle with at most $\lceil 21n/22 \rceil + 1$ edges. Such a cycle can be computed in $O(n \log n)$ time.*

Proof. Let δ be the path constructed by Theorem 4 on a point set P of size n . Recall m from the proof of this theorem. If $m < 22$ then the two endpoints of δ are the leftmost and rightmost points of P . Thus, by introducing a new point p with a sufficiently large y -coordinate and connecting it to the two endpoints of δ , we obtain a noncrossing covering cycle for P . If $m = 22$ then the dummy point that was introduced in Corollary 3 could be chosen suitably to play the role of p . ◀

3 Noncrossing covering trees

In this section we prove the following theorem which gives an algorithm for computing a noncrossing covering tree with roughly $4n/5$ edges. We should clarify that the number of *edges* of a tree is different from the number of its *segments* (where each segment is either a single edge or a chain of several collinear edges of the tree). For example the tree in Figure 6(b) has 10 edges and 7 segments, where the segments p_1p_7 and p_5p_8 consist of 3 and 2 collinear edges, respectively.

► **Theorem 6.** *Every set of n points in the plane admits a noncrossing covering tree with at most $\lceil 4n/5 \rceil$ edges. Thus, $\tau(n) \leq \lceil 4n/5 \rceil$. Such a tree can be computed in $O(n \log n)$ time.*

Proof. Let P be a set of n points in the plane. After a suitable rotation we may assume that no two points of P have the same x -coordinate. We present an iterative algorithm to compute a noncrossing covering tree for P that consists of at most $\lceil 4n/5 \rceil$ edges. In a nutshell, the algorithm scans the points from left to right and in every iteration (except possibly the last iteration) it considers 4 or 5 new points and covers them with 3 or 4 new edges, respectively. Thus the ratio of the number of new edges to the number of covered points would be at most $4/5$. We begin by describing an intermediate iteration of the algorithm; the first and last iterations will be described later. We assume that the scanned points in each iteration are in general position. In the end of the proof we describe how to handle collinearities. Let m be the number of points that have been scanned so far and let l be the rightmost scanned point (our choice of the letter l will become clear shortly). We maintain the following invariant at the beginning of every intermediate iteration.

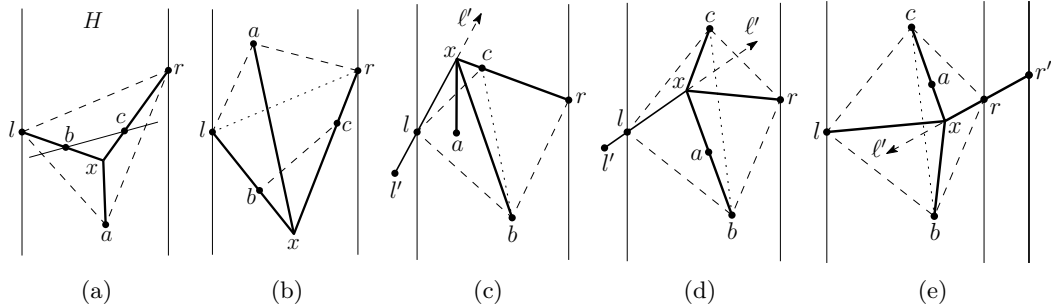
Invariant. All the m points that have been scanned so far, are covered by a noncrossing tree T with at most $4\lceil m/5 \rceil$ edges. The tree T lies to the left of the vertical line through l and the degree of l in T is one.

In the current (intermediate) iteration we scan four new points, namely a, b, c , and r where r is the rightmost point. Let H be the vertical strip bounded by the vertical lines through l and r (l is the leftmost point and r is the rightmost point in H); see Figure 5(a). Let $Q = \{l, a, b, c, r\}$. We consider three cases depending on the number of vertices of $\text{CH}(Q)$. Notice that r and l are two vertices of $\text{CH}(Q)$.

- $\text{CH}(Q)$ has three vertices. Let a be the third vertex of $\text{CH}(Q)$. Then b and c lie in the interior of $\text{CH}(Q)$, as in Figure 5(a). In this case two vertices of $\text{CH}(Q)$, say l and r , lie on the same side of $\ell(b, c)$. Thus l, b, c , and r form a convex quadrilateral. After a suitable relabeling assume that l, b, c, r appear in this order along the boundary of the quadrilateral. Let x be the intersection point of $\ell(l, b)$ and $\ell(r, c)$, which lies in the

19:10 Improved Bounds for Covering Paths and Trees in the Plane

triangle $\triangle lra$. We cover the four scanned points a , b , c , and r by three edges xl , xr , and xa which lie in H . We add these edges to T . The degree of r is one in the new tree (no matter which two vertices of $\text{CH}(Q)$ lay on the same side of $\ell(b, c)$). The invariant holds and we proceed to the next iteration.



■ **Figure 5** Illustration of the proof of Theorem 6. (a) $\text{CH}(Q)$ has three vertices. (b) $\text{CH}(Q)$ has five vertices. (c)-(e) $\text{CH}(Q)$ has four vertices.

- $\text{CH}(Q)$ has five vertices. We explain this case first as our argument is shorter, and also it will be used for the next case. In this case Q contains a 4-cap or a 4-cup with endpoints l and r . After a suitable reflection and relabeling assume it has the 4-cup l, b, c, r as in Figure 5(b). Let x be the intersection point of $\ell(l, b)$ and $\ell(r, c)$, and observe that it lies in H . We cover a, b, c , and r by three edges xl , xr , and xa which lie in H . We add these edges to T . The degree of r is one in the new tree. The invariant holds for the next iteration.
- $\text{CH}(Q)$ has four vertices. After a suitable relabeling assume that b and c are two vertices of $\text{CH}(Q)$ (other than l and r). Thus a lies in the interior of $\text{CH}(Q)$. If both b and c lie above or below $\ell(l, r)$ then l, b, c, r form a 4-cap or a 4-cup, in which case we cover the points as in the previous case. Therefore we may assume that one point, say b , lies below $\ell(l, r)$ and c lies above $\ell(l, r)$ as in Figures 5(c)-(e). We consider two subcases.
 - a lies in the triangle $\triangle lbc$. By the invariant, l has degree one in T . Let l' be the neighboring vertex of l in T . Let ℓ' be the ray emanating from l' and passing through l . We consider two subcases: (i) the segment bc does not intersect ℓ' and (ii) the segment bc intersects ℓ' .

In case (i) the segment bc lies below or above ℓ' . By symmetry assume that it lies below ℓ' . Then a and r also lie below ℓ' , as in Figure 5(c). In this case $\ell(r, c)$ intersects ℓ' . Let x be their intersection point, and observe that it lies in H . We replace the edge $l'l$ of T by $l'x$ (this does not increase the number of edges because l has degree one). Notice that $l'x$ contains l . Then we cover a, b, c , and r by adding three edges xr , xb , and xa to T . Therefore the number of edges of T is increased by 3. Moreover, r has degree one in the new tree, and all the newly introduced edges lie to the left of the vertical line through r . Thus the invariant holds for the next iteration.

In case (ii) the ray ℓ' goes through $\triangle lbc$. The point a lies below or above ℓ' . By symmetry assume that it lies below ℓ' , as in Figure 5(d). Let x be the intersection point of $\ell(a, b)$ and ℓ' , which lies in $\triangle lbc$. We replace the edge $l'l$ of T by $l'x$. Then we cover a, b, c , and r by adding three edges xr , xb , and xc to T . Thus, the number of edges of T is increased by 3, the vertex r has degree one in the new tree, and all new edges lie to the left of the vertical line through r . The invariant holds for the next iteration.

- *a* lies in the triangle $\triangle rbc$. Here is the place where we use four new edges to cover five vertices. In fact the ratio $4/5$ comes from this case (In previous cases we were able to cover four points by three new edges). In this case we scan the next point after r which we denote by r' , as in Figure 5(e). Now let ℓ' be the ray emanating from r' and passing through r . The current setting is essentially the vertical reflection of the previous case where r and r' play the roles of l and l' , respectively. We handle this case analogous to the previous case. Our analysis is also analogous except that now we consider the edge $r'x$ as a new edge. Thus we use four new edges to cover five points a, b, c, r , and r' . All new edges lie to the left of the vertical line through r' , and the degree of r' is one in the new tree. Thus the invariant holds for the next iteration.

This is the end of an intermediate iteration. The noncrossing property of the resulting tree follows from our construction. This iteration suggests a covering tree with roughly $4n/5$ edges. To get the exact claimed bound we need to have a closer look at the first and last iterations of the algorithm.

For the first iteration of the algorithm we scan only the leftmost input point. This point will play the role of l for the second iteration (which is the first intermediate iteration). The invariant holds for the second iteration because the tree has no edges at this point. If we happen to use the edge $l'l$ in the second iteration, then we take $l' = l$ and give the ray ℓ' an arbitrary direction to the right. Based on the above construction this could happen only when we scan four points (a, b, c, r) in the second iteration. In this case the first five points (l, a, b, c, r) are covered by four edges, and thus the invariant holds for the following iteration. In the last iteration of the algorithm we are left with $w \leq 4$ points that are not being scanned. We connect these w points by w edges to the rightmost scanned point. Therefore, the algorithm covers all points by a noncrossing tree with at most $\lceil 4n/5 \rceil$ edges.

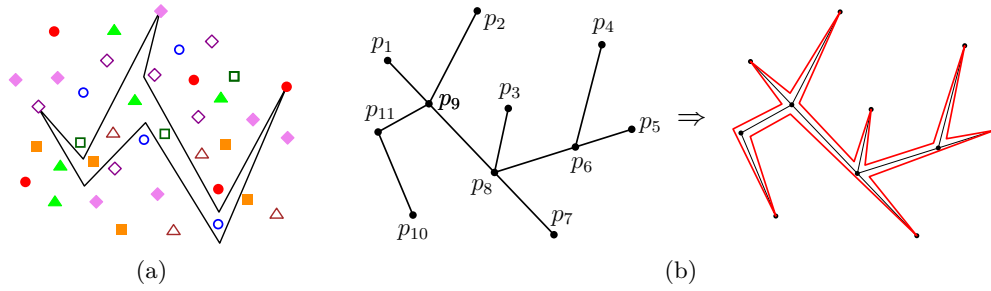
If three or more of the scanned points are collinear then cover all collinear points by one edge and connect the left endpoint of this edge to l . Then we connect every remaining scanned point to l . The number of new edges is at most 3 (for 4 scanned points) and 4 (for 5 scanned points).

Each iteration takes constant time. Therefore, after rotating and sorting the points in $O(n \log n)$ time, the rest of the algorithm takes linear time. ◀

4 Perfect rainbow polygons

Recall that a perfect rainbow polygon for a set of colored points, is a simple polygon that contains exactly one point of each color in its interior or on its boundary. Figure 6(a) shows a perfect rainbow polygon of size 9 (nine edges) for an 8-colored point set (i.e. colored by 8 different colors). There is a relation (as described below) between rainbow polygons and noncrossing covering trees. We employ this relation (similar to [19]) and present an algorithm that achieves a perfect rainbow polygon of size at most $7k/5 + O(1)$ for any k -colored point set.

We adopt the following notation and definitions from [19]. Let T be a noncrossing geometric tree. Recall that a segment of T is a chain of collinear edges in T . Let M be a partition of the edges of T into a minimal number of pairwise noncrossing segments. Let s denote the number of segments in M . A *fork* of T (with respect to M) is a vertex f that lies in the interior of a segment $ab \in M$ and it is an endpoint of another segment of M . The *multiplicity* of f is a number in $\{1, 2\}$ that is determined as follows. If the segments that have f as an endpoint lie on both sides of $\ell(a, b)$ then f has multiplicity 2, otherwise (the



■ **Figure 6** (a) A perfect rainbow polygon of size 9 for an 8-colored point set. (b) Left: A noncrossing tree with ten edges that can be partitioned into seven segments $\{p_1p_7, p_2p_9, p_3p_8, p_4p_6, p_5p_8, p_{10}p_{11}, p_{11}p_9\}$; p_6 and p_8 are two forks with multiplicity 1 and p_9 is a fork with multiplicity 2. Right: Obtaining a simple enclosing polygon from the tree.

segments lie on one side of the line) f has multiplicity 1. See the tree in Figure 6(b) for an example. Let t denote the sum of multiplicities of all forks in T . The following lemma expresses the size of a polygon enclosing T in terms of s and t .

► **Lemma 7** (Flores-Peñaloza et al. [19]). *Let T be a noncrossing geometric tree and M be a partition of its edges into a minimal number of pairwise noncrossing segments. Let s be the number of segments in M and t be the total multiplicity of forks in T . If $s \geq 2$ and $t \geq 0$, then for every $\varepsilon > 0$ there exists a simple polygon of size $2s + t$ and of area at most ε that encloses T .*

There are simple intuitions behind Lemma 7. For example if we cut out the tree T from the plane, then the resulting hole could be expressed as a desired polygon. Alternatively, if we start from a vertex of T and walk around T (arbitrary close to its edges) until we come back to the starting vertex, then the traversed tour could be represented as a desired polygon. See Figure 6(b).

In view of Lemma 7, a better covering tree (i.e. for which $2s + t$ is smaller) leads to a better polygon (i.e. with fewer edges). The following theorem (proven in the full version of the paper at <https://arxiv.org/abs/2303.04350>) gives a covering tree for which $2s + t$ is smaller (compared to that of [19]). Our construction in the proof of this theorem shares some similarities with the construction in our proof of Theorem 6. However, the details of the two constructions are different because they have different objectives.

► **Theorem 8.** *Let K be a set of k points in the plane in general position. Then, in $O(k \log k)$ time, one can construct a noncrossing covering tree for K consisting of at most $\lceil \frac{3k}{5} \rceil + 2$ pairwise noncrossing segments with at most $\lceil \frac{k}{5} \rceil$ forks of multiplicity 1.*

With this lemma and theorem in hand, we present our algorithm for computing a perfect rainbow polygon.

Algorithm. (in a nutshell). Let P be a set of n points in the plane in general position that are colored by k distinct colors. The algorithm picks one point from each color (arbitrarily), covers the chosen points by a noncrossing tree (using Theorem 8), and then obtains a perfect rainbow polygon from the tree (using Lemma 7).

Analysis. Let K be the set of k chosen points, and let T be the covering tree for K obtained by Theorem 8. Then $s \leq \lceil \frac{3k}{5} \rceil + 2$ and $t \leq \lceil \frac{k}{5} \rceil$. Thus, the perfect rainbow polygon obtained by Lemma 7 has size

$$2s + t \leq 2 \left(\left\lceil \frac{3k}{5} \right\rceil + 2 \right) + \left\lceil \frac{k}{5} \right\rceil \leq \left\lceil \frac{7k}{5} \right\rceil + 6.$$

The tree T can be obtained in $O(k \log k)$ time, by Theorem 8. To obtain a polygon (avoiding points of $P \setminus K$) from T we need to choose a suitable ε in Lemma 7. As noted in [19], half of the minimum distance between the edges of T and the points of $P \setminus K$ is a suitable ε , which can be found in $O(n \log n)$ time by computing the Voronoi diagram of the edges of T together with the points of $P \setminus K$. Thus the total running time of the algorithm is $O(n \log n)$. The following theorem summarizes our result in this section.

► **Theorem 9.** *Every k -colored point set of size n in the plane in general position admits a perfect rainbow polygon of size at most $\lceil 7k/5 \rceil + 6$. Thus, $\rho(k) \leq \lceil 7k/5 \rceil + 6$. Such a polygon can be computed in $O(n \log n)$ time.*

Remark. The general position assumption is necessary for our algorithm because if a non-selected point (i.e. a point of $P \setminus K$) lies on a segment of T then the resulting polygon is not a valid rainbow polygon as it contains two or more points of the same color.

5 Concluding remarks

A natural open problem is to improve the presented upper bounds or the known lower bounds for $\pi(n)$, $\tau(n)$, and $\rho(k)$. Here are some directions for further improvements:

- For the proof of Lemma 2 we used a 5-cap or a 5-cup which forced us to scan 21 points in each iteration (due to the result of Erdős and Szekeres). If one could manage to use a 4-cap or a 4-cup instead, then it could improve the upper bound for $\pi(n)$ further.
- Our iterative algorithm in the proof of Theorem 6, covers 4 points by 3 edges in all cases except in the last case (where $\text{CH}(Q)$ has four vertices and a lies in $\triangle rbc$) for which it covers 5 points by 4 edges. The upper bound $4n/5$ for $\tau(n)$ comes from this case. If one could argue that this case won't happen often (for example by showing that it won't happen in three consecutive iterations or by choosing a different ordering for points), then it would lead to a slightly improved upper bound for $\tau(n)$.

References

- 1 Alok Aggarwal, Don Coppersmith, Sanjeev Khanna, Rajeev Motwani, and Baruch Schieber. The angular-metric traveling salesman problem. *SIAM Journal on Computing*, 29(3):697–711, 1999. Also in *SODA '97*.
- 2 Oswin Aichholzer, Sergio Cabello, Ruy Fabila Monroy, David Flores-Peñaloza, Thomas Hackl, Clemens Huemer, Ferran Hurtado, and David R. Wood. Edge-removal and non-crossing configurations in geometric graphs. *Discrete Mathematics & Theoretical Computer Science*, 12(1):75–86, 2010.
- 3 Oswin Aichholzer, Frank Duque, Ruy Fabila Monroy, Oscar E. García-Quintero, and Carlos Hidalgo-Toscano. An ongoing project to improve the rectilinear and the pseudolinear crossing constants. *Journal of Graph Algorithms and Applications*, 24(3):421–432, 2020.
- 4 Esther M. Arkin, Yi-Jen Chiang, Joseph S. B. Mitchell, Steven Skiena, and Tae-Cheon Yang. On the maximum scatter traveling salesperson problem. *SIAM Journal on Computing*, 29(2):515–544, 1999. Also in *SODA '97*.

- 5 Esther M. Arkin, Joseph S. B. Mitchell, and Christine D. Piatko. Minimum-link watchman tours. *Information Processing Letters*, 86(4):203–207, 2003.
- 6 Sanjeev Arora. Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. *Journal of the ACM*, 45(5):753–782, 1998.
- 7 Sergey Bereg, Prosenjit Bose, Adrian Dumitrescu, Ferran Hurtado, and Pavel Valtr. Traversing a set of points with a minimum number of turns. *Discrete & Computational Geometry*, 41(4):513–532, 2009. Also in *SoCG'07*.
- 8 Ahmad Biniiaz. Euclidean bottleneck bounded-degree spanning tree ratios. *Discrete & Computational Geometry*, 67(1):311–327, 2022. Also in *SODA'20*.
- 9 Jakub Cerný, Zdenek Dvorák, Vít Jelínek, and Jan Kára. Noncrossing Hamiltonian paths in geometric graphs. *Discrete Applied Mathematics*, 155(9):1096–1105, 2007.
- 10 Timothy M. Chan. Euclidean bounded-degree spanning tree ratios. *Discrete & Computational Geometry*, 32(2):177–194, 2004. Also in *SoCG 2003*.
- 11 Jianer Chen, Qin Huang, Iyad Kanj, and Ge Xia. Near-optimal algorithms for point-line covering problems. *CoRR*, abs/2012.02363, 2020.
- 12 Michael J. Collins. Covering a set of points with a minimum number of turns. *International Journal of Computational Geometry & Applications*, 14(1-2):105–114, 2004.
- 13 Éva Czabarka, Ondrej Sýkora, László A. Székely, and Imrich Vrto. Biplanar crossing numbers. II. Comparing crossing numbers and biplanar crossing numbers using the probabilistic method. *Random Structures & Algorithms*, 33(4):480–496, 2008.
- 14 Erik D. Demaine and Joseph O'Rourke. Open problems from CCCG 2010. In *Proceedings of the 22nd Canadian Conference on Computational Geometry*, 2011.
- 15 Adrian Dumitrescu, Dániel Gerbner, Balázs Keszegh, and Csaba D. Tóth. Covering paths for planar point sets. *Discrete & Computational Geometry*, 51(2):462–484, 2014.
- 16 P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.
- 17 Sándor P. Fekete and Henk Meijer. On minimum stars and maximum matchings. *Discrete & Computational Geometry*, 23(3):389–407, 2000. Also in *SoCG 1999*.
- 18 Sándor P. Fekete and Gerhard J. Woeginger. Angle-restricted tours in the plane. *Computational Geometry: Theory and Applications*, 8:195–218, 1997.
- 19 David Flores-Peñalosa, Mikio Kano, Leonardo Martínez-Sandoval, David Orden, Javier Tejel, Csaba D. Tóth, Jorge Urrutia, and Birgit Vogtenhuber. Rainbow polygons for colored point sets in the plane. *Discrete Mathematics*, 344(7):112406, 2021.
- 20 Radoslav Fulek, Balázs Keszegh, Filip Morić, and Igor Uljarević. On polygons excluding point sets. *Graphs and Combinatorics*, 29(6):1741–1753, 2013.
- 21 Magdalene Grantson and Christos Levcopoulos. Covering a set of points with a minimum number of lines. In *Proceedings of the 6th International Conference on Algorithms and Complexity (CIAC)*, pages 6–17, 2006.
- 22 Frank Harary and Anthony Hill. On the number of crossings in a complete graph. *Proceedings of the Edinburgh Mathematical Society*, 13:333–338, 1963.
- 23 Alexandr O. Ivanov and Alexey A. Tuzhilin. The Steiner ratio Gilbert–Pollak conjecture is still open: Clarification statement. *Algorithmica*, 62(1-2):630–632, 2012.
- 24 Minghui Jiang. On covering points with minimum turns. *International Journal of Computational Geometry & Applications*, 25(1):1–10, 2015.
- 25 Balázs Keszegh. Covering paths and trees for planar grids. *Geombinatorics Quarterly*, 24, 2014.
- 26 Stefan Langerman and Pat Morin. Covering things with things. *Discrete & Computational Geometry*, 33(4):717–729, 2005. Also in *ESA'02*.
- 27 Samuel Loyd. *Cyclopedia of 5000 Puzzles, Tricks & Conundrums*. The Lamb Publishing Company, 1914.
- 28 János Pach, Natan Rubin, and Gábor Tardos. Planar point sets determine many pairwise crossing segments. *Advances in Mathematics*, 386:107779, 2021. Also in *STOC'19*.

- 29 Christos H. Papadimitriou. The Euclidean traveling salesman problem is NP-complete. *Theoretical Computer Science*, 4(3):237–244, 1977.
- 30 Clifford Stein and David P. Wagner. Approximation algorithms for the minimum bends traveling salesman problem. In *Proceedings of the 8th International Conference on Integer Programming and Combinatorial Optimization (IPCO)*, pages 406–422, 2001.