Long Alternating Paths in Bicolored Point Sets *

Jan Kynčl† János Pach‡ Géza Tóth§

Abstract

Given $n$ red and $n$ blue points in convex position in the plane, we show that there exists a noncrossing alternating path of length $n + c\sqrt{\frac{n}{\log n}}$. We disprove a conjecture of Erdős by constructing an example without any such path of length greater than $4n/3 + c'\sqrt{n}$.

1 Introduction

It is a basic problem in geometric graph theory to decide which graphs can be drawn on a given point set with noncrossing straight-line edges. For instance, it is known that every outerplanar graph (i.e., triangulated cycle) $G$ of $n$ vertices can be drawn on any set of $n$ points in general position in the plane [GMPP91]. Moreover, if $G$ is a rooted tree, one can find such an embedding even if the image of the root is specified [IPTT94], [T96]. An unsolved problem of this kind is to find the size of the smallest “universal” set in the plane, on which one can draw every planar graph of $n$ vertices with noncrossing straight-line edges [dFPP90], [CK89].

*János Pach has been supported by NSF Grant CCR-00-98246, by PSC-CUNY Research Award 65392-0034, and by OTKA T-032452. Géza Tóth has been supported by OTKA-T-038397.
†Department of Applied Mathematics and Institute for Theoretical Computer Science, Charles University, Prague, Czech Republic; jankync1@centrum.cz
‡City College, CUNY and Courant Institute of Mathematical Sciences, New York University, New York, NY, USA; pach@cims.nyu.edu
§Rényi Institute, Hungarian Academy of Sciences, Budapest, Hungary; geza@renyi.hu
We obtain many interesting new questions by considering colored point sets. It is a well known mathematics contest problem to prove that between any set $R$ of $n$ red and any set $B$ of $n$ blue points in general position in the plane there is a noncrossing matching, i.e., a one-to-one correspondence between their elements so that the segments connecting the corresponding point pairs are pairwise disjoint. Moreover, if $R$ and $B$ are separated by a line, one can also find an alternating Hamilton path, i.e., a noncrossing polygonal path passing through every element of $R \cup B$ such that any two consecutive vertices have opposite colors [AGH97]. For $n \geq 8$, the last statement is known to be false even if we assume that $R \cup B$ is in convex position, i.e., its elements form the vertex set of a convex $2n$-gon. The second named author heard the following problem from Erdös in the late nineteen-eighties.

**Problem.** Determine or estimate the largest number $\ell = \ell(n)$ such that, for every set of $n$ red and $n$ blue points on a circle, there exists a noncrossing alternating path consisting of $\ell$ vertices.

Of course, the condition that the points are on a circle is equivalent to the assumption that they are in convex position.

The conjectured asymptotically extremal configuration was the following. Suppose $n$ is divisible by four. Cut the circle into four intervals, and place in them $\frac{n}{2}$ red, $\frac{n}{4}$ blue, $\frac{n}{2}$ red, and $\frac{3}{4}n$ blue points, respectively. It is easy to see that in this construction the number of vertices in the longest noncrossing alternating path is $\frac{3}{2}n + 2$. That is, we have $\ell(n) \leq \frac{3}{2}n + 2$. The main aim of this note is to disprove Erdös’s conjecture by exhibiting a better construction in Section 2.

From the other direction, it is easy to argue that $\ell(n) \geq n$. Indeed, divide the circle into two arcs, each containing $n$ points. At least half of the points belonging to the first arc are of the same color, say, red. Then the second arc must contain the same number of blue points. Enumerate the red (resp. blue) points of the first (resp. second) arc in clockwise (resp. counterclockwise) order. Starting with the first red point on the first arc, and connecting each point with the next available element of opposite color on the other arc, we obtain a noncrossing alternating path of length $2\lceil \frac{n}{2} \rceil \geq n$. In Section 3, we improve this bound by a term that tends to infinity. Our results can be summarized as follows.
Theorem 1. There exist suitable constants $c, c' > 0$ such that
\[ n + c\sqrt{n \log n} < \ell(n) < \frac{4}{3}n + c'\sqrt{n}. \]

It is an annoying feature of this problem that it is not clear whether the assumption that the points are in convex position plays any significant role. In particular, the above argument for finding an alternating path of length $n$ easily generalizes to arbitrary 2-colored sets. We conjecture that the upper bound in Theorem 1 is asymptotically tight.

The problem of covering a set of $n$ red and $n$ blue points with several noncrossing alternating paths was discussed by Kaneko, Kano, and Suzuki [KKS04]. Alternating Hamiltonian cycles with at most $n - 1$ crossings were found by Kaneko, Kano, and Yoshimoto [KKY00]. Their result cannot be improved. Many other interesting questions about partitioning the plane into a given number of convex pieces, each containing roughly or exactly the same number of red and blue points, were studied in [BKS00], [BM01], [S02]. Analogous questions can be asked when we color by red and blue all the $\binom{n}{2}$ segments between $n$ points in general position in the plane. Furthermore, instead of long alternating paths, we may be interested in finding long monochromatic ones [KPTV98]. For a recent survey, see [KK04].

2 Upper Bound

Consider an ‘2-equicolored’ set $C$ of $2n$ points in convex position in the plane. That is, let half of the elements of $C$ be colored red and half of them blue. An uninterrupted run (or, in short, run) is a maximal set of consecutive points of $C$ that have the same color. The length of a run is the number of its elements. We say that $C$ is a $k$-configuration if it consists of $k$ red and $k$ blue runs.

A set of pairwise disjoint segments, each of which connects two points of different colors, is called a matching. The size of a matching is defined as the total number of points participating in it, that is, twice the number of segments. A matching is said to be separated if there is a straight-line that intersects the interior of each of its segments.

Lemma 2.1. Let $C$ be a $k$-configuration for some $k > 0$, which has a noncrossing alternating path of length $l$. Then $C$ has a separated matching
whose size is at least $l - 4k - 1$.

**Proof.** Suppose without loss of generality that all elements of $C$ lie on a circle. Consider a noncrossing alternating path $p$ of length $l$. Fix a chord $c$ of the circle that crosses the first and last segments along $p$, but does not pass through any point of $C$. Let $M_1$ denote the matching consisting of all odd-numbered segments of $p$. Clearly, the size of $M_1$ is at least $l - 1$. Let $M_2 \subseteq M_1$ be the set of all segments in $M_1$ that cross $c$. By definition, $M_2$ is a separated matching.

To establish the lemma, it is enough to show that the number of elements of $M_1$ that do not cross $c$ is at most $2k$. Let us call these segments *outer* segments. For each pair of consecutive points of $C$ that have different colors, pick a point between them on the circle. Any two consecutive runs of $C$ are separated by at least one such point, so the number of points we selected is precisely $2k$. Every outer segment $s$ divides the circle into two (closed) arcs. One of them, $I(s)$, contains both endpoints of $c$; let the other one be denoted by $J(s)$. Since $s$ connects two points of different colors, $J(s)$ must contain at least one of the selected points. On the other hand, both endpoints of the alternating path $p$ belong to $I(s)$, so $J(s)$ cannot contain the endpoints of any outer segment other than the endpoints of $s$. Thus, for any two outer segments, $s$ and $s'$, $J(s)$ and $J(s')$ are disjoint, so the selected points lying in the corresponding arcs $J(s)$ and $J(s')$, resp., are different. Hence, the number of outer segments cannot exceed the total number of selected points, which is $2k$. See Fig. 1.

\[ J(s) \]

Figure 1: $J(s)$ and $J(s')$ are disjoint.
Any $k$-configuration with runs $S_1, S_2, \ldots, S_{2k}$ can be represented by a sequence $(s_1, s_2, \ldots, s_{2k})$, where $s_i = |S_i| \frac{3k-2}{n}$. We assume that the odd-numbered runs are red and the even-numbered are blue.

**Theorem 2.2.** Let $k \geq 2$ and assume that $n$ is divisible by $3k-2$. Then the size of any separated matching in the $k$-configuration $(k, 2k - 1, k, 1, \ldots, 1)$ is at most $2n \frac{2k-1}{3k-2}$.

![Diagram of the k-configuration](image)

**Figure 2:** The $k$-configuration $(k, 2k - 1, k, 1, \ldots, 1)$.

**Proof.** Let $S_1, S_2, \ldots, S_{2k}$ denote the consecutive runs of the $k$-configuration $(k, 2k - 1, k, 1, \ldots, 1)$ (see Fig. 2). Let $M$ be a separated matching. We distinguish five cases, according to the set of runs that are connected to $S_2$ by at least one edge in $M$.

*Case 1:* No edge of $M$ has an endpoint in $S_2$.

Then $M$ uses at most $n - n \frac{2k-1}{3k-2}$ blue points, so its size is at most $2n - 2n \frac{2k-1}{3k-2} = 2n \frac{k-1}{3k-2} < 2n \frac{2k-1}{3k-2}$.

*Case 2:* No edge of $M$ runs between $S_2$ and $S_1 \cup S_3$.

The points in $S_2$ can be connected only to the elements of $S_5, S_7, \ldots, S_{2k-1}$. Hence, at most $n \frac{k-1}{3k-2}$ points of $S_2$ are matched and at least $n \frac{k}{3k-2}$ are missed by $M$. The size of $M$ is at most $2n - 2n \frac{k}{3k-2} = 2n \frac{2k-2}{3k-2}$.

*Case 3:* $S_2$ is connected by an edge of $M$ to both $S_1$ and $S_3$. 


Since $M$ is a separated matching, the blue points in $S_4, S_6, \ldots, S_{2k}$ are not matched, so $M$ uses at most $n \frac{2k-1}{3k-2}$ blue points. Thus, the size of $M$ is at most $2n \frac{2k-1}{3k-2}$.

**Case 4:** $S_2$ is connected by an edge of $M$ to $S_1$, but not to $S_3$.
Suppose that the size of $M$ exceeds $2n \frac{2k-1}{3k-2}$. Then $M$ matches more than $n \frac{k}{2k-3}$ points of $S_2$, so at least one edge of $M$ must connect $S_2$ to a red run different from $S_1$. Let $i > 1$ denote the smallest integer such that there is an edge of $M$ between $S_2$ and $S_{2i+1}$. Then $M$ matches at most $n \frac{k+(k-1)}{3k-2} = n \frac{2k-i}{3k-2}$ blue points from $S_2$ and misses the $n \frac{k-i}{3k-2}$ blue points in $S_{2i+2}, S_{2i+4}, \ldots, S_{2k}$. Therefore, it does not match at least $n \left(\frac{2k-1}{3k-2} - \frac{k+(k-1)}{3k-2} \right) = n \frac{k-1}{3k-2}$ blue points, and its size is at most $2n \frac{2k-1}{3k-2}$.

**Case 5:** $S_2$ is connected by an edge of $M$ to $S_3$, but not to $S_1$.
By symmetry, the same argument applies as in the previous case.

**Lemma 2.3.** For any positive integers $k$ and $n \geq k$, there exists a $k$-configuration of $2n$ points with no alternating path longer than $2n \frac{2k-1}{3k-2} + 16k$.

**Proof.** The statement is trivial for $k = 1$, and also for $n \leq 8k$. Suppose that $k \geq 2$, and $n > 8k$. Let $n_0 \leq n$ be the largest integer divisible by $3k-2$. Let $C_0$ denote a $k$-configuration consisting of $n_0$ red and $n_0$ blue points, considered in Theorem 2.2. Add $n - n_0$ red points to $S_1$ and $n - n_0$ blue points to $S_2$, and denote the resulting $k$-configuration by $C$.

We claim that $C$ satisfies the requirement of the lemma. Let $p$ be an alternating path of length $l(p)$ in $C$. By Lemma 2.1, there is a separated matching $M_1$ in $C$, whose size is $l(M_1) \geq l(p) - 4k - 1$. Remove from $M_1$ the $2n - 2n_0$ points that were added later and all edges in $M_1$ incident to them. We obtain a separated matching $M_0$ of $C_0$ of size $l(M_0) \geq l(M_1) - 4(n - n_0) \geq l(M_1) - 4(3k-2) = l(M_1) - 12k + 8 \geq l(p) - 16k$. By Theorem 2.2, we have $l(M_0) \leq 2n_0 \frac{2k-1}{3k-2}$, so that $l(p) \leq 2n \frac{2k-1}{3k-2} + 16k$.

The upper bound in Theorem 1 immediately follows from Lemma 2.3. For any $n$, set $k = \lfloor \sqrt{n} \rfloor$. Applying Lemma 2.3, we obtain a configuration of $n$ red and $n$ blue points in which the length of any noncrossing alternating path is at most $\frac{2k-1}{3k-2} \cdot 2n + 16k < \frac{4n}{3} + 20\sqrt{n}$, as required.
3 Lower Bound

As before, let $C$ be the vertex set of a convex $2n$-gon, with $n$ red and $n$ blue elements. Suppose without loss of generality that the elements of $C$ lie on a circle. A set of consecutive vertices of $C$ (of not necessarily the same color) is said to be an interval. The length of an interval is its cardinality.

Assume that we can find a separated matching $M$ of size $2l$, all of whose segments are crossed by a chord $c$. Then we can easily construct a non-crossing alternating path of length $2l$. To see this, enumerate the segments $s_1, s_2, \ldots, s_l$ of $M$ according to the order of their intersection points with $c$. Let $r_i$ and $b_i$ denote the red and blue endpoints of $s_i$, respectively. Then $p = (r_1 b_1, b_1 r_2, r_2 b_2, \ldots, b_{l-1} r_l, r_l b_l)$ is a noncrossing alternating path of length $2l$.

Therefore, it is sufficient to establish a lower bound on the size of a separated matching in a $k$-configuration of $2n$ points. We divide the proof into two steps. Lemmas 3.1 and 3.2 provide reasonably good bounds when $k$ is small and when $k$ is large, respectively. Their combination implies the general lower bound in Theorem 1.

Lemma 3.1. Let $k, m, n$ be positive integers such that $k = 2^m$ divides $n$. Let $C$ be a $k$-configuration of $2n$ points. If $C$ contains a run of length at least $\frac{n}{k}$, then it determines a separated matching whose size is at least $n \left( 1 + \frac{1}{k(m+1)} \right)$.

Proof. Let $S$ be the run of length at least $\frac{n}{k}$. Let $I_0$ denote a monochromatic interval in $S$, whose length is precisely $\frac{n}{k}$. For $1 \leq i \leq m + 1$, let $I_i$ be an interval of length $2^{i-1} \frac{n}{k}$ such that $I_0, I_1, \ldots, I_{m+1}$ are consecutive in the clockwise direction (see Fig. 3). These intervals form a partition of the underlying set $C$ consisting of all $2n$ vertices. Assume without loss of generality that all elements of $I_0$ are blue.

![Figure 3: $I_0, I_1, \ldots, I_{m+1}$ form a partition of the vertices.](image)

Suppose for contradiction that there is no separated matching whose size
is at least $n \left(1 + \frac{1}{k(m+1)} \right)$.

**Claim.** For every $0 \leq i \leq m+1$, the interval $J_i = I_0 \cup I_1 \cup \ldots \cup I_i$ has at least $\left(2^{i-1} \frac{1}{k} + \frac{1}{2k} - \frac{i}{2k(m+1)} \right) n$ blue points. Moreover, strict inequality holds if $i > 0$.

Note that Lemma 3.1 immediately follows from the Claim. Indeed, for $i = m + 1$, we obtain that there are more than $\left(2^{m+1} \frac{1}{k} + \frac{1}{2k} - \frac{m+1}{2k(m+1)} \right) n = n$ blue points on the circle, which is a contradiction.

**Proof of Claim.** We proceed by induction on $i$. For $i = 0$, the statement obviously holds. Assume that for some $i \in \{0, 1, 2, \ldots, m\}$, there are at least $\left(2^{i-1} \frac{1}{k} + \frac{1}{2k} - \frac{i}{2k(m+1)} \right) n$ blue points in the interval $J_i$. We show that there are more than $\left(2^{i-1} \frac{1}{k} - \frac{i}{2k(m+1)} \right) n$ blue points in $I_{i+1}$.

Suppose this is not the case. Then there are at least $l = \left(2^{i-1} \frac{1}{k} + \frac{1}{2k} - \frac{i}{2k(m+1)} \right) n$ red points in $I_{i+1}$. Since we have $2^{i-1} \frac{1}{k} + \frac{1}{2k} - \frac{i}{2k(m+1)} \geq 2^{i-1} \frac{1}{k} + \frac{1}{2k(m+1)}$, the number of blue points in $J_i$ is at least $l$. Thus, there is a separated matching of size $2l$ between the blue points of $J_i$ and the red points of $I_{i+1}$. Let $M$ be the most ‘economical’ such matching of size $2l$. That is, if $b_1, b_2, \ldots, b_s$ ($s \geq l$) denote the blue points of $J_i$ listed in counterclockwise order and $r_1, r_2, \ldots, r_t$ ($t \geq l$) denote the red points of $I_{i+1}$ listed in clockwise order, then let $M$ consist of the segments $b_1r_1, b_2r_2, \ldots, b_tr_t$.

Let $K$ denote the interval between $b_l$ and $r_t$, oriented clockwise. All blue points from $K$ not matched by $M$ lie in $I_{i+1}$, so their number is at most $u_b = \left(2^{i-1} \frac{1}{k} - \frac{1}{2k(m+1)} \right) n$. All red points of $K$ not matched by $M$ lie in $J_i$, so their number is at most

$$u_r = \left(2^i \frac{1}{k} - \left(2^{i-1} \frac{1}{k} + \frac{1}{2k} - \frac{i}{2k(m+1)} \right) \right) n = \left(2^{i-1} \frac{1}{k} - \frac{1}{2k} + \frac{i}{2k(m+1)} \right) n.$$

Using the fact that $u_b - u_r = \left(\frac{1}{2k} - \frac{i+1}{2k(m+1)} \right) n \geq 0$, we obtain that $u_b \geq u_r$.

Let $L$ denote the complement of $K$ in the set $C$ of all points. Clearly, $L$ has at least $n_0 = n - l - u_b = \left(1 - 2^{i} \frac{1}{k} \right) n$ points of each color. Divide $L$ into two intervals $L_1, L_2$, each of length at least $n_0$. Obviously, at least one of the following two conditions is satisfied:

(a) There are at least $\frac{n_0}{2}$ blue points in $L_1$ and at least $\frac{n_0}{2}$ red points in $L_2$. 

8
(b) There are at least \( \frac{n_0}{2} \) red points in \( L_1 \) and at least \( \frac{n_0}{2} \) blue points in \( L_2 \).

We can assume without loss of generality that (a) is true. Then there exists a separated matching \( M' \) of size at least \( n_0 \) between intervals \( L_1 \) and \( L_2 \). The union of \( M \) and \( M' \) is also a separated matching. (One endpoint of the corresponding chord lies between the intervals \( J_i \) and \( I_{i+1} \), and the other between \( L_1 \) and \( L_2 \).) The size of \( M \cup M' \) is at least

\[
2l + n_0 = \left( 2^{i-1} \frac{1}{k} + \frac{1}{2k(m+1)} \right) 2n + \left( 1 - 2^i \frac{1}{k} \right) n = \left( 1 + \frac{1}{k(m+1)} \right) n,
\]

which is a contradiction. Hence, our assumption was wrong: there are more than \( \left( 2^{i-1} \frac{1}{k} - \frac{1}{2k(m+1)} \right) n \) blue points in \( I_{i+1} \).

Consequently, the number of blue points in the interval \( J_{i+1} = I_0 \cup I_1 \cup \ldots \cup I_{i+1} \) is larger than

\[
\left( 2^{i-1} \frac{1}{k} + \frac{1}{2k} - \frac{i}{2k(m+1)} \right) n + \left( 2^{i-1} \frac{1}{k} - \frac{1}{2k(m+1)} \right) n =
\]

\[
\left( 2^i \frac{1}{k} + \frac{1}{2k} - \frac{i + 1}{2k(m+1)} \right) n,
\]

completing the induction step, the proof of the Claim, and hence the proof of Lemma 3.1.

**Lemma 3.2.** For \( n \geq k \geq 1 \), every \( k \)-configuration of \( 2n \) points admits an alternating path whose length is at least \( n + k - 1 \).

**Proof.** Let \( v_1, v_2, \ldots, v_{2n} \) be the vertices of a \( k \)-configuration, in clockwise direction. Assume that \( v_1 \) is red. For any \( 1 \leq i < 2n \), if \( v_i \) is red and \( v_{i+1} \) is blue, then \( v_i \) and \( v_{i+1} \) are called special vertices. There are \( 2k \) special and \( 2n - 2k \) non-special vertices. Moreover, exactly half of the special and exactly half of the non-special vertices are red, and blue. Let \( m \) be the smallest number such that there are \( n - k \) non-special vertices in the set \( \{ v_i \mid 1 \leq i \leq m \} \). Assume that \( t \geq \lceil \frac{n-k}{2} \rceil \) of them are red. (The other case can be settled analogously.) Denote those red points by \( u_1, u_2, \ldots, u_t \), in clockwise direction. Then the set \( \{ v_i \mid m + 1 \leq i \leq 2n \} \) also contains \( n - k \) non-special vertices, and \( t \) of them are blue. Denote those blue points by \( w_1, w_2, \ldots, w_t \), in counterclockwise direction.

The vertices \( u_1, u_2, \ldots, u_t \) divide the set \( \{ v_i \mid 1 \leq i \leq m \} \) into \( t + 1 \) intervals of consecutive vertices, denote them by \( I_0, I_1, \ldots, I_t \), in clockwise
direction. For $0 \leq 1 \leq t$, if $I_i$ contains some special vertices, denote them by $u_{i,1}, u_{i,2}, \ldots u_{i,\alpha_i}$, in clockwise direction. Since $u_1, u_2, \ldots, u_t$ are non-special, $u_{i,j}$ is red if $j$ is odd, and blue if $j$ is even.

Similarly, the vertices $w_1, w_2, \ldots, w_t$ divide the set $\{v_i \mid m + 1 \leq i \leq 2n\}$ into $t + 1$ intervals of consecutive vertices, denote them by $J_0, J_1, \ldots, J_t$, in counterclockwise direction. For $0 \leq 1 \leq t$, if $J_i$ contains some special vertices, denote them by $w_{i,1}, w_{i,2}, \ldots w_{i,\beta_i}$, also in counterclockwise direction. Now $w_{i,j}$ is blue if $j$ is odd, and red if $j$ is even.

Finally, consider the following path:

$$u_{0,1}, u_{0,2}, \ldots u_{0,\alpha_0}, u_1, w_{0,1}, w_{0,2}, \ldots w_{0,\delta_0}, w_1, u_{1,1}, u_{1,2}, \ldots u_{1,\alpha_1}, u_2, \ w_{1,1}, w_{1,2}, \ldots w_{1,\delta_1}, w_2, \ldots, u_{t,1}, u_{t,2}, \ldots u_{t,(\alpha_t-1)}, w_{t,1}, w_{t,2}, \ldots w_{t,\beta_t}.$$ 

It is a noncrossing, alternating path of length $2t + 2k - 1 \geq n - k + 2k - 1 = n + k - 1$. This concludes the proof of the Lemma.

Now we are ready to prove the lower bound in Theorem 1. Suppose that we have a $k$-configuration of $2n$ points. We distinguish two cases. If $k \geq \frac{1}{10} \sqrt{\frac{n}{\log n}}$, then, by Lemma 3.2, there exists an alternating path of length at least $n + \frac{1}{10} \sqrt{n} - 1$.

We are left with the case $k < \frac{1}{10} \sqrt{\frac{n}{\log n}}$. Let $m$ be the least positive integer such that $k \leq 2^m$. Then $m < 1 + \log k$. Let $n' = 2m \left\lfloor \frac{n}{2^m} \right\rfloor \geq n - 2^m$ and choose any subconfiguration $C'$ of $n'$ red and $n'$ blue points from $C$. $C'$ is a $k_0$-configuration for some $k_0 \leq k$. So it has a run of length at least $\frac{n'}{k_0} \geq \frac{n'}{k} \geq \frac{n'}{2^m}$. Now, according to Lemma 3.1, $C$ has a separated matching (and also an alternating path) whose size is at least

$$n' + \frac{n'}{2^m(m + 1)} \geq n - 2k + \frac{n}{2^{m+1}(m + 1)} \geq$$

$$n - 2k + \frac{n}{4k(\log k + 2)} \geq n - 2k + \frac{n}{2k \log n} \geq$$

$$n - \frac{1}{5} \sqrt{n} + \frac{5n}{\log n} \sqrt{\frac{\log n}{n}} \geq n + 4 \sqrt{\frac{n}{\log n}}.$$ 

This completes the proof.

**Conjecture.** For any fixed $k$ and large $n$, every $k$-configuration of $2n$ points admits a separated matching of size at least $2n^{\frac{2k-1}{3k-2}} + o(n)$.
References


