On the Number of Directions Determined by a Three-Dimensional Points Set*

János Pach† Rom Pinchasi‡ Micha Sharir§

March 9, 2004

Abstract

Let P be a set of n points in $\mathbb{R}^3$, not all of which are in a plane and no three on a line. We partially answer a question of Scott (1970) by showing that the connecting lines of P assume at least $2n-3$ different directions if n is even and at least $2n-2$ if n is odd. These bounds are sharp. The proof is based on a far-reaching generalization of Ungar's theorem concerning the analogous problem in the plane.

1 Introduction

Erdős [7] pointed out the following immediate consequence of the celebrated Gallai-Sylvester theorem on ordinary lines (see Borwein and Moser [4] for a survey): n non-collinear points in the plane determine at least n different connecting lines. Equality is attained if and only if all but one of the points are collinear.

In the same spirit, Scott [16] posed two similar questions in 1970:

1. Is it true that the minimum number of different directions assumed by the connecting lines of $n \geq 4$ non-collinear points in the plane is $2\lfloor n/2 \rfloor$?

*Work on this paper by János Pach and Micha Sharir has been supported by NSF Grants CCR-97-32101 and CCR-00-98246, and by a joint grant from the U.S.–Israel Binational Science Foundation. Work by János Pach has also been supported by PSC-CUNY Research Award 63382-0032 and by Hungarian Science Foundation grant OTKA T-043520. Work by Micha Sharir has also been supported by a grant from the Israeli Academy of Sciences for a Center of Excellence in Geometric Computing at Tel Aviv University, and by the Hermann Minkowski–MINERVA Center for Geometry at Tel Aviv University.

†City College, CUNY and Courant Institute of Mathematical Sciences, NYU, 251 Mercer Street, New York, NY 10012, USA; `pach@cims.nyu.edu`

‡Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; `room@math.mit.edu`

§School of Computer Science, Tel Aviv University, Tel Aviv 69 978, Israel, and Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA; `michas@tau.ac.il`

2. Is it true that the minimum number of different directions assumed by the connecting lines of $n \geq 6$ non-coplanar points in 3-space is $2n - 3$ if n is even and $2n - 2$ if n is odd?

Twelve years later, the first question was answered in the affirmative by Ungar [18]. His proof is a real gem, a brilliant application of the method of *allowable sequences* invented by Goodman and Pollack [9], [10]. Moreover, it solves the problem in an elegant combinatorial setting, for "pseudolines", as was suggested independently by Goodman and Pollack and by Cordovil [6]. For even n, Ungar's theorem generalizes Erdős's above mentioned result. However, in contrast to Erdős's result, here there is an overwhelming diversity of extremal configurations, for which equality is attained. Four infinite families and more than one hundred sporadic configurations were catalogued by Jamison and Hill [14] (see also [13] for an excellent survey).

Progress on the second question of Scott has been much slower. As Jamison [13] noticed, unless we impose some further restriction on the point set, for odd n, the number of directions determined by n points in 3-space can be as small as $2n - 5$. Indeed, equality is attained, e.g., for the n-element set obtained from the vertex set of a regular $(n - 3)$-gon P_{n-3} (or from any other centrally symmetric extremal configuration for the planar problem) by adding its center c and two other points whose midpoint is c and whose connecting line is orthogonal to the plane of P_{n-3}.

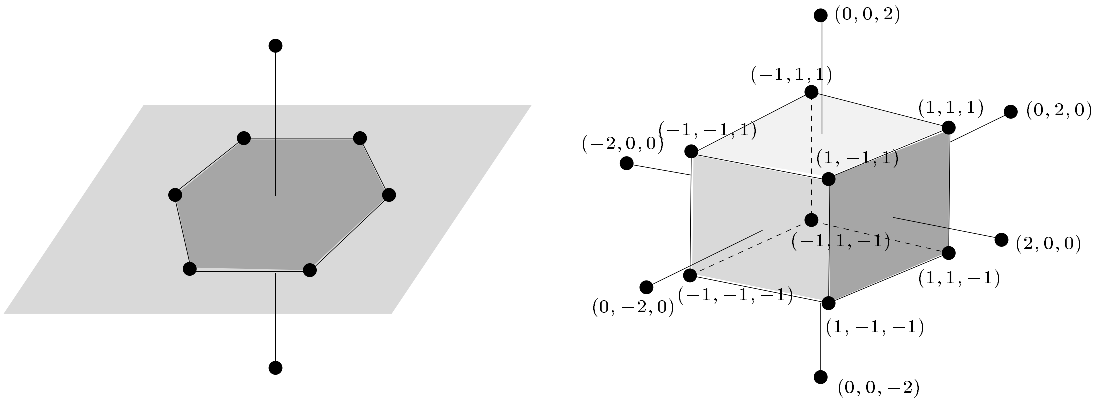
Blokhuis and Seress [3] introduced a natural condition excluding the above configurations: they assumed that no three points are collinear. Under this assumption, they proved that every non-coplanar set of n points in 3-space determines at least $1.75n - 2$ different directions.

The aim of the present paper is to answer Scott's second question in the affirmative, using the same assumption as Blokhuis and Seress.

Theorem 1.1. *Every set of $n \geq 6$ points in $\mathbb{R}^3$, not all of which are on a plane and no three are on a line, determine at least $n + 2\lceil n/2 \rceil - 3$ different directions. This bound is sharp.*

Removing the center c from the configuration described above that determines $2n - 5$ directions, we obtain a set of even size $n' = n - 1$ with $2n' - 3$ directions and no three collinear points (see Figure 1(a)). If the number of points is even, then this construction provides the only known infinite family for which Theorem 1.1 is sharp. In addition, there are four known sporadic extremal configurations, each of which is a subset of the 14-element set depicted in Figure 1(b).

According to a beautiful result of Motzkin [15], Rabin, and Chakerian [5] (see also [1]), any set of n non-collinear points in the plane, colored with two colors *red* and *green*, determines a monochromatic line. Motzkin and Grünbaum [11] initiated the investigation of *biased* colorings, i.e., colorings without monochromatic red lines. Their motivation was to justify the intuitive feeling that if there are many red points in such a coloring and not all of them are collinear, then the number of green points must also be rather large. Denoting the sets of red and green points by R and G,



(a) (b)

Figure 1: Two examples of sets with an even number n of points, not all on a plane and no three coplanar, that determine $2n-3$ different directions.

respectively, it is a challenging unsolved question to decide whether the "surplus" $|R|-|G|$ of the coloring can be arbitrarily large. We do not know any example where this quantity exceeds 6 [12].

The problem of biased colorings was rediscovered by Erdős and Purdy [8], who formulated it as follows: What is the smallest number $m(n)$ of points necessary to represent (i.e., stab) all lines spanned by n non-collinear points in the plane, if the generating points cannot be used. An $\Omega(n)$ lower bound follows from the "weak Dirac conjecture" proved by Szemerédi and Trotter [17] and Beck [2], according to which there is a point that lies on $\Omega(n)$ different connecting lines. Each of these connecting lines has to be represented by a different point.

In Section 2, we reduce Theorem 1.1 to a statement (Theorem 2.2) showing that under some further restrictions the surplus is indeed bounded. More precisely, if there is no connecting line whose leftmost and rightmost points are both red, then we have $|G| \geq 2\lfloor |R|/2 \rfloor$, so in particular $|R|-|G| \leq 1$.

Another way of rephrasing Ungar's theorem is that from all closed segments whose endpoints belong to a non-collinear set of n points in the plane, one can always select at least $2\lfloor n/2 \rfloor$ such that no two of them are parallel. Unless we explicitly state it otherwise, every *segment* used in this paper is assumed to be *closed*. Our proof of Theorem 2.2 is based on a far-reaching generalization of Ungar's result. To formulate this statement, we need to relax the condition of two segments being *parallel*.

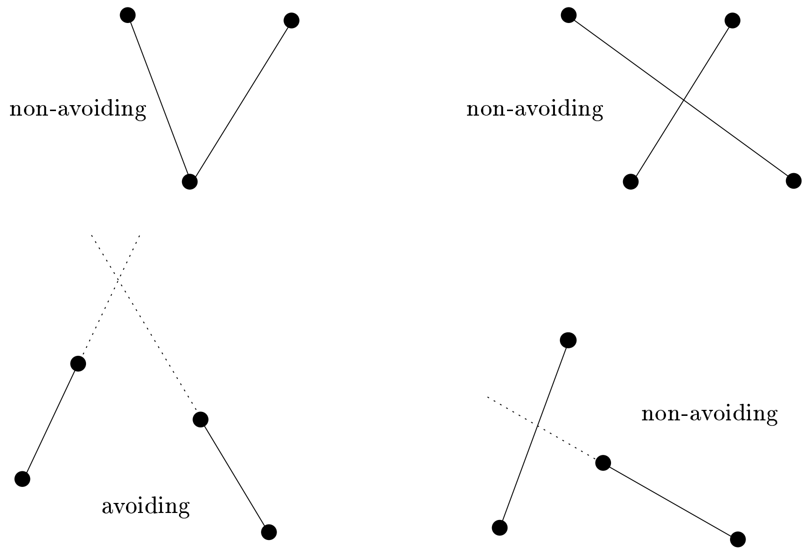


Figure 2: Avoiding and non-avoiding segments.

Definition 1.2. *Two segments belonging to distinct lines are called* avoiding *if one of the following two conditions is satisfied (see Figure 2):*

(i) they are parallel, or

(ii) the intersection of their supporting lines does not belong to any of the segments.

An alternative definition is that two segments are avoiding if and only if they are disjoint and their convex hull is a quadrilateral.

The main result of this paper, which implies Theorem 1.1 and Theorem 2.2 (stated in the next section), is the following strengthening of Ungar's theorem, which is of independent interest.

Theorem 1.3. *From all closed segments determined by a set of n non-collinear points in the plane, one can always select at least $2\lfloor n/2 \rfloor$ pairwise non-avoiding ones, lying on distinct lines.*

Theorem 1.3 is established in Sections 3 and 4.

This paper leaves open the problem of extending Theorem 1.1 to the general case, where the given point set may contain triples of collinear points.

2 Reduction of Theorem 1.1 to a Planar Problem

Let P be a set of n points in $\mathbb{R}^3$ such that not all of them lie in a common plane and no three of them are collinear. Let p_0 be an *extreme* point of P, i.e., a vertex of the convex hull of P. Consider a supporting plane to P at p_0, and translate it into a new

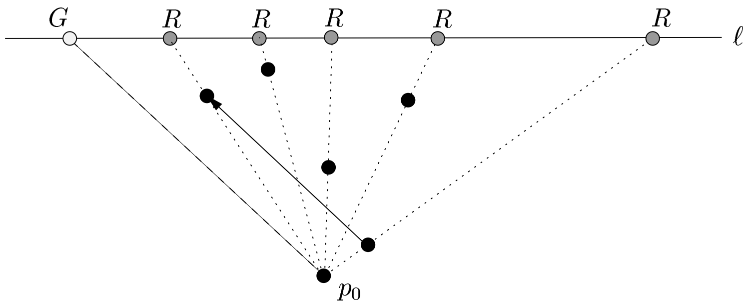


Figure 3: Proving the existence of a green point on ℓ, in extreme position.

position π so that P lies in the slab bounded by these two planes. Project from p_0 all points of $P \setminus \{p_0\}$ onto π. We obtain a set R of $n-1$ distinct points in π, not all on a line, and we will refer to the elements of R as *red* points. Each red point corresponds to a direction determined by p_0 and some other point of P.

For each pair of elements $p, p' \in P \setminus \{p_0\}$, take a line parallel to pp' that passes through p_0. Color with *green* the intersection point of this line with π, unless it has already been colored red. The set of all green points is denoted by G. By definition, we have $R \cap G = \emptyset$.

We need the following simple property of the sets R and G, which implies that along every line passing through at least two red points either the leftmost or the rightmost point belonging to $R \cup G$ is green.

Lemma 2.1. *Every line connecting two red points $r, r' \in R$ passes through at least one green point $g \in G$ that does not belong to the (closed) segment rr'.*

Proof: Let ℓ be a line in π passing through at least two red points $r, r' \in R$. Assume without loss of generality that r and r' are the leftmost and rightmost red points along ℓ. Let p and p' denote those elements of P whose projections to π are r and r', respectively. Observe that in the plane induced by p_0 and ℓ, the direction of pp' does not belong to the convex cone enclosed by the rays p_0p and p_0p', so the line through p_0 parallel to pp' will cross ℓ in a green point g meeting the requirements. See Figure 3. $\square$

To establish Theorem 1.1, it is sufficient to verify the following result.

Theorem 2.2. *Let R be a set of n red points in the plane, not all collinear, and let G be a set of m green points such that $R \cap G = \emptyset$ and every line ℓ connecting at least two red points in R passes through a green point $g \in G$ that does not belong to any segment rr', for $r, r' \in R \cap \ell$.*

Then we have $m \geq 2\lfloor n/2 \rfloor$.

Indeed, to prove Theorem 1.1 it is enough to notice that in our setting we have

$|R| = n - 1$ and that the number of different directions determined by P is equal to

$$|R| + |G| \geq n - 1 + 2\left\lfloor \frac{n-1}{2} \right\rfloor = n + 2\left\lceil \frac{n}{2} \right\rceil - 3.$$

Thus, applying Theorem 2.2, Theorem 1.1 immediately follows.

It is interesting to note that Theorem 2.2 also implies Ungar's above-mentioned theorem. To see this, regard the elements of our given planar point set as *red*, and the directions determined by them as green points on the line at infinity, and apply Theorem 2.2. (If we wish, we can perform a projective transformation and avoid the use of points at infinity.)

It remains to prove Theorem 2.2. However, as mentioned in the introduction, this result can be easily deduced from Theorem 1.3, which is a further extension of Ungar's theorem:

Proof of Theorem 2.2 (using Theorem 1.3): Applying Theorem 1.3 to the set R, we obtain $2\lfloor n/2 \rfloor$ segments with red endpoints that lie in distinct lines and no pair of them are avoiding. By the condition in Theorem 2.2, the continuation of each of these segments passes through a green point. Assign such a green point to each segment. Observe that these points are all distinct. Indeed, if we can assign the same green point to two different segments, then they must be avoiding, by definition. This completes the proof of Theorem 2.2 and hence of Theorem 1.1. □

3 Junctions and Stations – Proof of Theorem 1.3

The aim of this and the next section is to establish an equivalent *dual* version of Theorem 1.3. Fix an (x, y)-coordinate system in the plane. We apply a standard duality transform that maps a point $p = (p_1, p_2)$ to the line p^* with equation $y + p_1 x + p_2 = 0$. Vice versa, a non-vertical line l with equation $y + l_1 x + l_2 = 0$ is mapped to the point $l^* = (l_1, l_2)$. Consequently, any two parallel lines are mapped into points having the same x-coordinate. It is often convenient to imagine that the dual picture lies in another, so-called *dual*, plane, different from the original one, which is referred to as the *primal* plane.

The above mapping is incidence and order preserving, in the sense that p lies above, on, or below ℓ if and only ℓ^* lies above, on, or below p^*, respectively. The points of a segment $e = ab$ in the primal plane are mapped to the set of all lines in the closed *double wedge* e^*, which is bounded by a^* and b^* and does not contain the vertical direction. All of these lines pass through the point $q = a^* \cap b^*$, which is called the *apex* of the double wedge e^*. All double wedges used in this paper are assumed to be closed, and they never contain the vertical direction.

Definition 3.1. *We call two double wedges* avoiding *if their apices are distinct and the apex of neither of them is contained in the other (see Figure 4).*

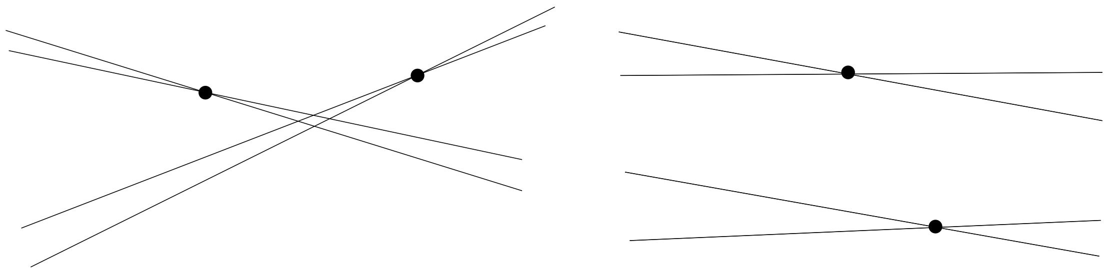


Figure 4: Two possible kinds of avoiding double wedges.

It is easy to see that, according to this definition, two non-collinear segments in the primal plane are avoiding if and only if they are mapped to avoiding double wedges.

Switching to the dual plane, Theorem 1.3 can now be reformulated as follows.

Theorem 3.2. *Let L be a set of n pairwise non-parallel lines in the plane, not all of which pass through the same point. Then the set of all double wedges bounded by pairs of lines in L has at least $2\lfloor n/2 \rfloor$ pairwise non-avoiding elements with different apices.*

Note that the definition of double wedges depends on the choice of the coordinate system, so *a priori* Theorem 3.2 gives a separate statement in each coordinate frame. However, each of these statements is equivalent to Theorem 1.3, and that result does not depend on coordinates. Therefore, we are free to use whatever coordinate system we like. In the final part of the analysis (given in Section 4), we will exploit this property. But until then, no restriction on the coordinate system is imposed.

Suppose that a set of $2\lfloor n/2 \rfloor$ double wedges meets the conditions in Theorem 3.2. Clearly, we can replace each element of this set, bounded by a pair of lines $\ell_1, \ell_2 \in L$, by the *maximal* double wedge with the same apex, i.e., the double wedge bounded by those lines through $\ell_1 \cap \ell_2$ which have the *smallest* and *largest* slopes. If every pair of double wedges in the original set was non-avoiding, then this property remains valid after the replacement.

It is sufficient to prove Theorem 3.2 for the case when n is even, because for odd n the statement trivially follows.

The proof is constructive. Let $\mathcal{A}(L)$ denote the *arrangement* of L, consisting of all vertices, edges, and faces of the planar map induced by L. We will construct a set of n vertices of $\mathcal{A}(L)$ with distinct x-coordinates, and show that the maximal double wedges whose apices belong to this set are pairwise non-avoiding.

We start by defining a sequence J of vertices $v_1, v_2, \ldots$, which will be referred to as *junctions.* Let L^- (resp., L^+) denote the subset of L consisting of the $n/2$ lines with the smallest (resp., largest) slopes. If we wish to simplify the picture, we can apply an affine transformation that keeps the vertical direction fixed and carries the elements of L^- and L^+ to lines of negative and positive slopes, respectively (whence

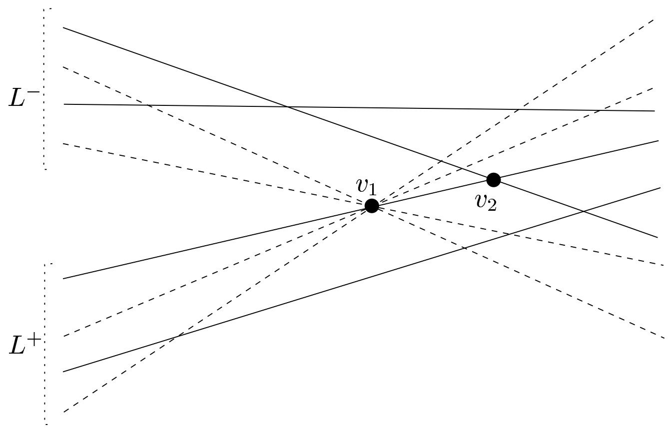


Figure 5: Choosing the first junction v_1 in J. The dashed lines, two from L^- and two from L^+, are removed. The next junction v_2 is also shown.

the choice of notation). However, we will never use this property explicitly (although the figures will reflect this convention).

The construction proceeds as follows.

STEP 1: Set $i := 1$ and $L_1^- := L^-$, $L_1^+ := L^+$.

STEP 2: If $L_i^- = L_i^+ = \emptyset$, the construction of J terminates. Otherwise, as we will see, neither set is empty. Let v_i be the *leftmost* intersection point between a line in L_i^- and a line in L_i^+. Let d_i^- (and d_i^+) denote the number of elements of L_i^- (and L_i^+, respectively) incident to v_i, and put $d_i = \min\{d_i^-, d_i^+\}$. Define L_{i+1}^- (and L_{i+1}^+) as the set of lines obtained from L_i^- (resp., L_i^+) by deleting from it the d_i elements that are incident to v_i and have the smallest (resp., largest) slopes among those incident lines. (That is, if $d_i^- = d_i^+$, then all lines incident to v_i are deleted; otherwise, if, say, $d_i^- > d_i^+$, we are left with $d_i^- - d_i^+$ lines through v_i that belong to L_i^- and separate the deleted elements of L_i^- from the deleted elements of L_i^+. See Figure 5.) Set $i := i+1$, and repeat Step 2.

Let $J = \langle v_1, v_2, \ldots, v_k\rangle$ denote the resulting sequence.

It is easy to verify the following properties of this construction.

Claim 3.3. *(i)* $|L_i^-| = |L_i^+|$, *for each* $i = 1, \ldots, k$.

(ii) For every $1 \le i < j \le k$, *the junction* v_i *lies in the left unbounded face* f_j *of* $\mathcal{A}(L_j^- \cup L_j^+)$ *which separates* L_j^- *and* L_j^+ *at* $x = -\infty$ *(whose rightmost vertex is* v_j*).* v_i *lies in the interior of* f_j *if* $d_i^- = d_i^+$*; otherwise it may lie on the boundary of* f_j*.*

(iii) $\sum_{i=1}^k d_i = n/2$. $\square$

Next, between any two consecutive junctions v_i and v_{i+1}, for $1 \le i < k$, we specify $d_i + d_{i+1} - 1$ further vertices of $\mathcal{A}(L)$, called *stations*.

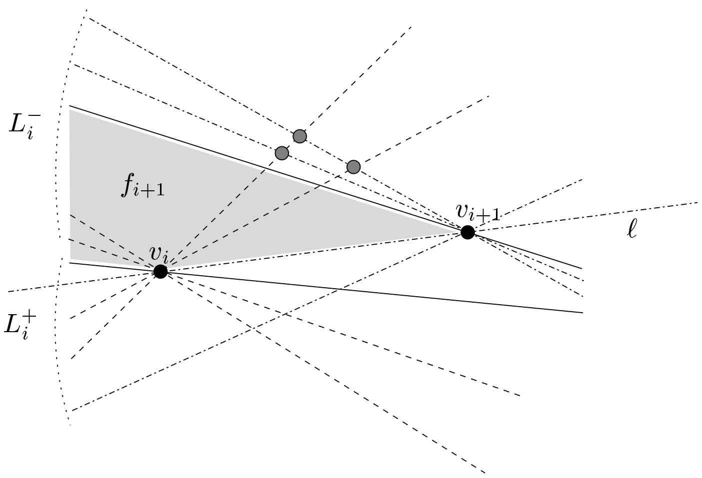


Figure 6: Collecting stations (shown highlighted) between v_i and v_{i+1}. The dashed lines are those removed at v_i, and the dashed-dotted ones are those removed at v_{i+1}. The figure depicts the subcase where a line ℓ that has been removed at v_{i+1} also passes through v_i. In this case the lines of L_i^- deleted at v_i and the lines of L_{i+1}^+ deleted at v_{i+1} *do not* generate enough stations.

Fix an index $1 \leq i < k$, and consider the vertical slab between v_i and v_{i+1}. By Claim 3.3 (ii), v_i lies inside or on the boundary of the face f_{i+1} of $\mathcal{A}(L_{i+1}^- \cup L_{i+1}^+)$, whose rightmost vertex is v_{i+1}. See Figure 6. Hence, the segment $e = v_i v_{i+1}$ is contained in the closure of f_{i+1}. Now at least one of the following two conditions is satisfied: (a) all the d_i lines removed from L_i^+ and all the d_{i+1} lines removed from L_{i+1}^- pass above e, or (b) all the d_i lines removed from L_i^- and all the d_{i+1} lines removed from L_{i+1}^+ pass below e. (We caution the reader that this statement is not totally obvious when e belongs to the boundary of f_{i+1}.)

Assume, by symmetry, that (a) holds. Denote the lines removed from L_i^+ by $\ell_1^+, \ldots, \ell_{d_i}^+$, listed according to increasing slopes, and those removed from L_{i+1}^- by $\ell_1^-, \ldots, \ell_{d_{i+1}}^-$, listed according to decreasing slopes. Define the set of *stations* S_i in the vertical slab between v_i and v_{i+1} as the collection of all intersection points of $\ell_{d_i}^+$ with the lines $\ell_1^-, \ldots, \ell_{d_{i+1}}^-$, and all intersection points of $\ell_{d_{i+1}}^-$ with the lines $\ell_1^+, \ldots, \ell_{d_i}^+$. Clearly, we have $|S_i| = d_i + d_{i+1} - 1$ such points; see Figure 6.

Finally, we consider the portions of the plane to the left of v_1 and to the right of v_k and collect there a set S_k of $d_k + d_1 - 1$ additional *stations*. Actually, exploiting the fact that we can (almost) freely select the coordinate system used for the duality transform, we will be able to select $d_k + d_1 - 1$ suitable stations, so that all of them, or all but one, lie to the left of v_1. The proper choice of the coordinate system as well as the details of the construction of S_k are described in the next section.

Let $Q = J \cup \left(\cup_{i=1}^k S_i\right)$. In view of Claim 3.3 (iii), the total number $|Q|$ of junctions and stations equals

$$\begin{aligned}
|Q| &= |J| + \sum_{i=1}^{k} |S_i| \\
&= k + \sum_{i=1}^{k-1} (d_i + d_{i+1} - 1) + (d_k + d_1 - 1) \\
&= 2\sum_{i=1}^{k} d_i = n.
\end{aligned}$$

To complete the proof of Theorem 3.2 (and hence of Theorem 1.3), we need to verify

Claim 3.4. *Associate with each element $q \in Q$ the* maximal *double wedge $W(q)$ (not containing the vertical line through q), which is bounded by a pair of lines passing through q. Then the resulting set of n double wedges has no two avoiding elements.*

We close this section by verifying the last claim for the set of wedges $\{W(q) | q \in Q \setminus S_k\}$. The extension to the general case is postponed to the last section, where S_k is defined.

Let $u, v \in Q \setminus S_k$ with u lying to the left of v. We distinguish three cases:

Case A: Both u and v are junctions.

Put $u = v_i$ and $v = v_j$, with $i < j$. Then $W(v)$ is bounded by a line $\ell \in L_j^-$ and by a line $\ell' \in L_j^+$. By Claim 3.3(ii), v_i lies between these two lines, and thus belongs to $W(v)$.

Case B: u is a junction and v is a station not in S_k.

Put $u = v_i$ and let S_j be the set of stations that contains v, where $i \le j$. Then $W(v)$ is bounded by two lines ℓ, ℓ', where either $\ell \in L_j^-$ and $\ell' \in L_{j+1}^+$, or $\ell \in L_{j+1}^-$ and $\ell' \in L_j^+$. By construction, we have in both cases $\ell \in L_j^-$ and $\ell' \in L_j^+$, and the analysis is completed as in Case A.

Case C: u is a station not in S_k and v is a junction or a station not in S_k.

Let S_i be the set of stations containing u. The arguments in Case A and Case B imply that $v_i \in W(v)$. If v is also a station in S_i or $v = v_{i+1}$ then it is easy to verify, by construction, that $W(u)$ and $W(v)$ are non-avoiding (see Figure 6). Suppose then that v lies to the right of v_{i+1}. Then both v_i and v_{i+1} lie in the left wedge of $W(v)$, and u is incident to a line λ of positive slope that passes through v_i and to a line λ' of negative slope that passes through v_{i+1}. If $u \notin W(v)$ then a boundary line of $W(v)$ must separate u from v_i and v_{i+1}, in which case $v \in W(u)$; see Figure 7.

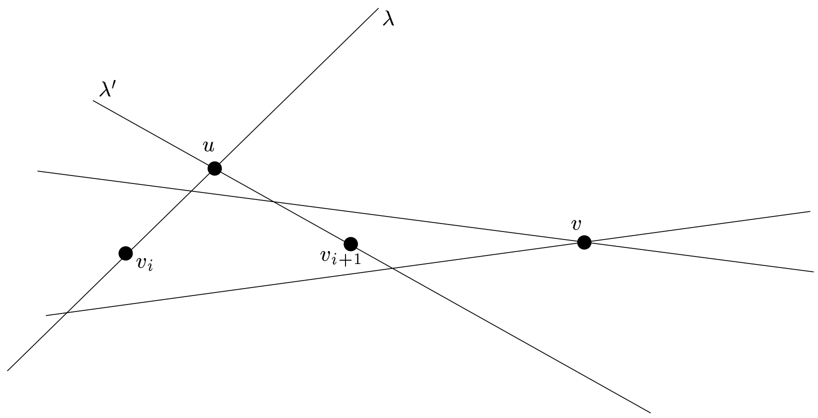


Figure 7: Illustrating Case C of the proof that $W(u)$ and $W(v)$ cannot be avoiding.

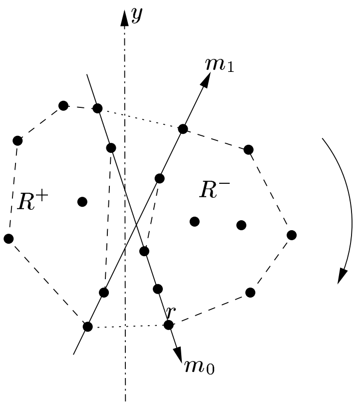


Figure 8: The primal construction of R^- and R^+.

4 Wrapping Up – The End of the Proof

In this section, we define the missing set of stations S_k, and extend the proof of Claim 3.4 to handle also elements of S_k. We need an elementary geometric fact that is easier to formulate in the primal setting.

Lemma 4.1. *Let R be a set of n non-collinear points in the plane, let n be even, and let r be any vertex of the convex hull of R. Then there exists a partition of R into two $n/2$-element subsets, R^- and R^+, whose convex hulls are disjoint and which have a common inner tangent m_0 passing through r.*

Proof: Rotate a directed line ℓ counterclockwise about r, starting with all the points of $R \setminus \{r\}$ lying to the left of ℓ, until the closed halfplane to the right of ℓ contains for the first time more than $n/2$ points. Define R^- to be the set R_0 of points in the open halfplane to the right of ℓ, plus the first $n/2 - |R_0|$ points of $\ell \cap R$ along ℓ. m_0 coincides with the final position of ℓ. See Figure 8. □

Let m_1 denote the other inner tangent of the convex hulls of R^- and R^+. Now choose an orthogonal (x, y)-coordinate system whose y-axis is a line strictly separating R^- and R^+. Suppose without loss of generality that

(a) R^+ and R^- are to the left and to the right of the y-axis, respectively,

(b) $r \in R^-$, and

(c) m_0 is oriented from r away from the other contact point(s), and the positive y-direction lies counterclockwise to it. See Figure 8.

In the dual picture, R^- and R^+ become $n/2$-element sets of lines, L^- and L^+, having negative and positive slopes, respectively. Applying the construction described in the previous section to $L := L^- \cup L^+$, we obtain a sequence of junctions $J = \langle v_1, v_2, \ldots, v_k \rangle$ and sets of stations $S_1, \ldots, S_{k-1}$.

Since m_1 is the line with the *largest* slope connecting a point of R^+ and a point of R^-, our duality implies that m_1^*, the dual of m_1, is the *leftmost* intersection point between a line of L^+ and a line of L^-. Hence, we have $v_1 = m_1^*$. As our construction sweeps the dual plane from left to right, we collect junctions and stations whose dual lines rotate clockwise from m_1 onwards.

Claim 4.2. *At least one of the following two conditions will be satisfied:*

(i) The last junction, v_k, is identical to m_0^*, the dual of m_0.

(ii) r^*, the dual of $r \in R^-$, passes through v_k and is the unique element of L^- deleted during the procedure at v_k (so that $d_k = 1$).

Proof: Suppose that during the procedure r^* is deleted at a junction v_j, for some $j \leq k$. Clearly, v_j^* passes through r and at least one point $t \in R^+$.

If in the primal plane v_j^* passes through another point $r' \neq r$ of R^-, then $v_j^* = m_0$ (otherwise it has to lie clockwise to m_0 and then it cannot meet any point of R^+). In this case, in the dual plane there cannot be any intersection point between a line of L^- and a line of L^+ to the right of v_j, so that $j = k$. That is, we have $v_k^* = m_0$, and (i) holds.

If in the primal plane v_j^* does not pass through any element $r' \in R^-$ other than r, then we have $d_j = 1$. If $j = k$, then condition (ii) is satisfied. Let us assume, by contradiction, that $j < k$ and $v_k^* \neq m_0$. Take any two lines $\ell_- \in L^-$ and $\ell_+ \in L^+$ in the dual plane that are deleted during the procedure at the last junction v_k. By assumption and construction, we have $\ell_-^* \neq r$, and the slope of the segment $\ell_+^* \ell_-^* \subset v_k^*$ connecting their duals in the primal plane (i.e., the slope of v_k^*) is smaller than that of the segment tr. We claim that the two segments $\ell_+^* \ell_-^* \subset v_k^*$ and $tr \subset v_j^*$ are avoiding. Indeed, $\ell_+^* \ell_-^*$ must meet m_0 to the left of r, or else r would not be an extreme point of R (see Figure 9). For a similar reason, ℓ_-^* must lie above v_j^*. These facts, together with the slope relationship between v_j^* and v_k^*, imply that the two segments are avoiding. This, in turn, implies that the wedges $W(v_k)$ and $W(v_j)$ are avoiding, contradicting Claim 3.4 (Case A). □

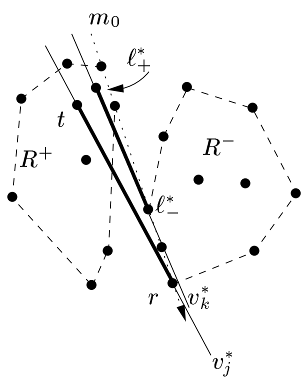


Figure 9: The segments tr and $\ell_+^*\ell_-^*$ must be avoiding.

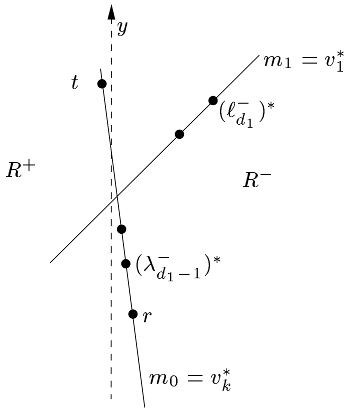
The above argument is valid for any coordinate system whose y-axis strictly separates the sets R^- and R^+. We specify a coordinate system with this property as follows.

Choose the y-axis to be very close to m_0, so that, in the dual plane the slope of every line of L passing through m_0^* has smaller absolute value than the slope of any other line of L; that is, the x-coordinates of the points of $m_0 \cap R$ have smaller absolute values than those of any other point of R. See Figures 10(a) and 11(a).

Now we are in a position to define the set of stations S_k. Pass to the dual plane. The first junction, v_1, lies inside or on the boundary of the face f_k of $\mathcal{A}(L_k^- \cup L_k^+)$, whose rightmost vertex is v_k, so that the segment $e = v_1 v_k$ is contained in the closure of f_k.

Suppose first that $v_k = m_0^*$. We can assume by symmetry that in the dual plane all the d_1 lines removed from $L_1^- = L^-$ during the procedure pass below e, and all the d_k lines of L_k^- pass above e (as in the preceding section, this statement is not totally obvious when e lies on the boundary of f). Let $\ell_1^-, \ldots, \ell_{d_1}^-$ and $\lambda_1^-, \ldots, \lambda_{d_k}^-$ denote the removed lines of L_1^- and of L_k^-, respectively, listed in the decreasing order of their slopes. By the special choice of our coordinate system, each line ℓ_i^- intersects every line λ_j^- to the left of v_1. Indeed, the slope of the primal segment $(\lambda_j^-)^*(\ell_i^-)^*$ is larger than that of m_1, because $(\lambda_j^-)^* \in m_0$ lies below m_1 and to the left of $(\ell_i^-)^* \in m_1$; see Figure 10(a). (We note that the assumption that all lines in L_1^- pass strictly below v_k implies that $c := m_0 \cap m_1$ is not dual to any line in L_1^-, implying that each $(\lambda_j^-)^*$ does indeed lie to the left of every $(\ell_i^-)^*$.) Define the last set of *stations*, S_k, as the collection of all intersection points of $\ell_{d_1}^-$ with the lines $\lambda_1^-, \ldots, \lambda_{d_k}^-$, and all intersection points of $\lambda_{d_k}^-$ with the lines $\ell_1^-, \ldots, \ell_{d_1}^-$. See Figure 10(b). Clearly, we have $|S_k| = d_k + d_1 - 1$ such points, all lying to the left of v_1.

Suppose next that $v_k \neq m_0^*$. In this case, according to Claim 4.2, v_k lies on r^* and $d_k = 1$. Refer to Figure 11. Again, let $\ell_1^-, \ldots, \ell_{d_1}^-$ denote the lines removed from $L_1^- = L^-$ at v_1, listed in the decreasing order of their slopes. In the dual plane, the line r^* passes above v_1 and, by the choice of the coordinate system, it intersects every



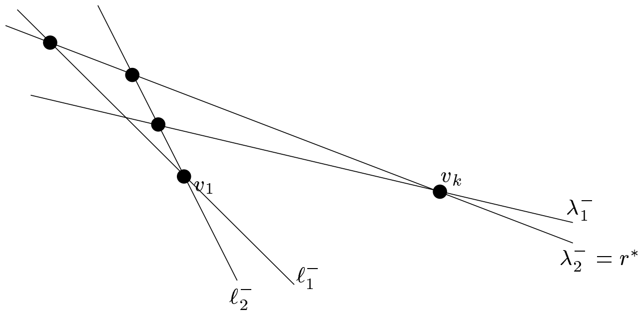


Figure 10: The case $v_k = m_0^*$ of the construction of S_k. (a) The primal structure. (b) The stations in S_k (highlighted to the left of v_1).

ℓ_i^- to the left of v_1, with the possible exception of ℓ_1^-. The intersection $r^* \cap \ell_1^-$ can lie to the right of v_1 (and of v_k) only if the point $c := m_0 \cap m_1$ belongs to R^- and is dual to a line removed at v_1, in which case that line must be $\ell_1^- = c^*$. Note that in this case $r^* \cap \ell_1^- = r^* \cap c^*$ is identical to the point m_0^* dual to m_0, and the choice of the coordinate system implies that this is the rightmost vertex of $\mathcal{A}(L)$ on r^*. We define S_k to be the set of intersection points between the lines $\ell_1^-, \ldots, \ell_{d_1}^-$ and r^*.[1] Thus, either all points of S_k, or all but one (namely, m_0^*) lie to the left of v_1. Clearly, we have $|S_k| = d_1 = d_k + d_1 - 1$, as required.

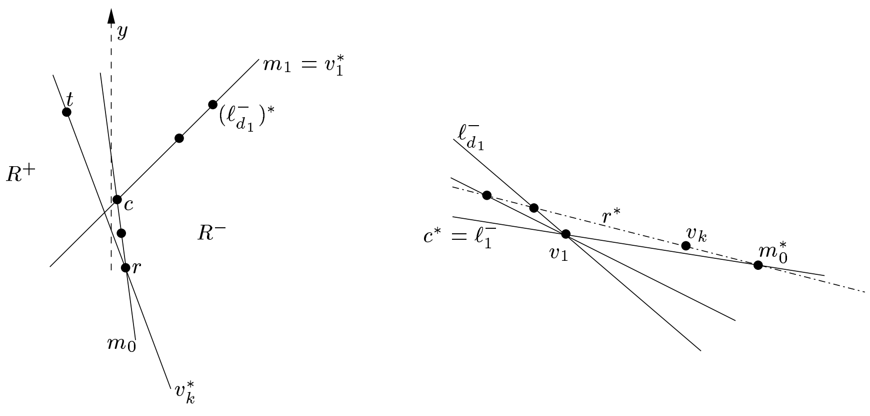
We have to complete the proof of Claim 3.4. It remains to show the following:

Claim 4.3. *For any $u \in Q$ and any $v \in S_k$, the maximal wedges $W(u)$ and $W(v)$ associated with them are non-avoiding.*

Proof: If both u and v belong to S_k, then the claim is obviously true. From now on suppose that $u \notin S_k$. Then we have $u \in \{v_i\} \cup S_i \cup \{v_{i+1}\}$, for some $1 \le i < k$.

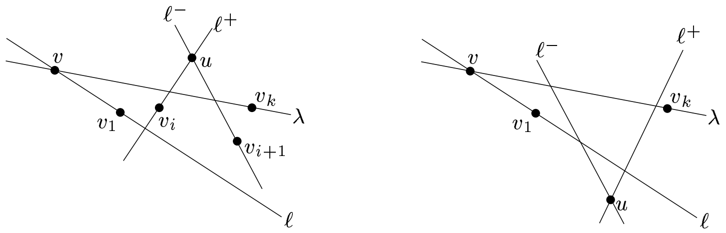
We start with the case $v_k = m_0^*$. Let $v \in S_k$ be the intersection point of two lines, ℓ and λ, passing through v_1 and v_k, respectively, which, without loss of generality, we assume to belong to L^-, If u is contained in the double wedge bounded by ℓ and λ, then $u \in W(v)$, so that $W(u)$ and $W(v)$ are non-avoiding. Otherwise, since v lies to the left of v_1, u lies either above λ or below ℓ. If u is above λ, then it is not a junction, so it must be the crossing point of a line $\ell^+ \in L^+$ and a line $\ell^- \in L^-$ which are removed during the procedure at junction v_i and at junction v_{i+1}, respectively. See Figure 12(a). Both v_i and v_{i+1} lie on or below λ, so that the left portion of the double wedge bounded by ℓ^- and ℓ^+ contains v. Thus, we have $v \in W(u)$. If, on the

[1]Note the assymmetry between this case, where the stations are constructed using lines in L^- only, and the previous case, where the stations can be constructed using either lines of L^- or lines of L^+, depending on the relative position of the lines incident to v_1 and v_k.



(a) (b)

Figure 11: The case $v_k \neq m_0^*$ of the construction of S_k. (a) The choice of the coordinate frame. (b) The dual picture.



(a) (b)

Figure 12: The proof that $W(u)$ and $W(v)$ are non-avoiding when v is a station to the left of v_1.

other hand, u is below ℓ, as in Figure 12(b), then it is either a junction or a station, and it is the crossing point of a line $\ell^- \in L^-$ and a line $\ell^+ \in L^+$, each of which is removed at junction v_i or at junction v_{i+1}. Now ℓ^- must pass above (or through) v_1 and hence above v, while ℓ^+ must pass below v. Again we can conclude that the left portion of the double wedge bounded by ℓ^- and ℓ^+, and thus $W(u)$, contains v.

If $v_k \neq m_0^*$, the above argument can be repeated *verbatim*, unless $m_0^* \in S_k$ and $v = m_0^*$; so assume this to be the case. Now it is simplest to establish the claim in the primal plane, by noting that the segment dual to $W(v)$ lies on the line m_0, and that, by construction (since $u \notin S_k$), the segment dual to $W(u)$ must connect a point of R^- to a point of R^+, and thus must intersect m_0, showing that these two segments are non-avoiding. $\Box$

By verifying the last claim, we have completed the proof of Claim 3.4 and hence of Theorem 3.2. This was our last debt.

References

[1] M. Aigner and G. Ziegler, *Proofs from The Book (including illustrations by Karl H. Hofmann), 2nd ed.* Springer-Verlag, Berlin, 2001.

[2] J. Beck, On the lattice property of the plane and some problems of Dirac, Motzkin and Erdős in combinatorial geometry, *Combinatorica* **3** (1983), 281–297.

[3] A. Blokhuis and Á. Seress, The number of directions determined by points in the three-dimensional Euclidean space, *Discrete Comput. Geom.* **28** (2002), 491–494.

[4] P. Borwein and W. O. J. Moser, A survey of Sylvester's problem and its generalizations, *Aequationes Math.* **40** (1990), 111–135.

[5] G. D. Chakerian, Sylvester's problem on collinear points and a relative, *Amer. Math. Monthly* **77** (1970), 164–167.

[6] R. Cordovil, The directions determined by n points in the plane, a matroidal generalization, *Discrete Math.* **43** (1983), 131–137.

[7] P. Erdős, Solution to Problem Nr. 4065, *Amer. Math. Monthly* **51** (1944), 169.

[8] P. Erdős and G. Purdy, Some combinatorial problems in the plane, *J. Combinatorial Theory, Ser. A* **25** (1978), 205–210.

[9] J. E. Goodman and R. Pollack, A combinatorial perspective on some problems in geometry, *Congr. Numer.* **32** (1981), 383–394.

[10] J. E. Goodman and R. Pollack, Allowable sequences and order types in discrete and computational geometry, in: *New Trends in Discrete and Computational Geometry (J. Pach, ed.), Algorithms Combin.* **10**, Springer, Berlin, 1993, 103–134.

[11] B. Grünbaum, Arrangements of colored lines, Abstract 720-50-5, *Notices Amer. Math. Soc.* **22** (1975), A-200.

[12] B. Grünbaum, Monochromatic intersection points in families of colored lines, *Geombinatorics* **IX** (1999), 3–9.

[13] R. E. Jamison, Survey of the slope problem, in: *Discrete Geometry and Convexity, Ann. New York Acad. Sci.* **440**, New York Acad. Sci., New York, 1985, 34–51.

[14] R. E. Jamison and D. Hill, A catalogue of sporadic slope-critical configurations. in: *Proceedings of the Fourteenth Southeastern Conference on Combinatorics, Graph Theory and Computing (Boca Raton, Fla., 1983), Congr. Numer.* **40** (1983), 101–125.

[15] T. S. Motzkin, Nonmixed connecting lines, Abstract 67T 605, *Notices Amer. Math. Soc.* **14** (1967), 837.

[16] P. R. Scott, On the sets of directions determined by n points, *Amer. Math. Monthly* **77** (1970), 502–505.

[17] E. Szemerédi and W.T. Trotter, Jr., Extremal problems in discrete geometry, *Combinatorica* **3** (1983), 381–392.

[18] P. Ungar, $2N$ noncollinear points determine at least $2N$ directions, *J. Combin. Theory, Ser. A* **33** (1982), 343–347.