A Note on the Number of Bichromatic Lines

Daniel J. Kleitman and Rom Pinchasi

Massachusetts Institute of Technology

August 30, 2014

Abstract

Let G be a set of points in the plane which are colored red or green. If the number of red points differ from the number of green points by at most 1 and no color class is collinear, then there are at least $|G| - 3$ bichromatic lines which are determined by G. We conjecture that the tight lower bound is in fact $|G| - 1$.

1 Introduction

We say that a set of points G determines a line L, if L passes through at least two points of G.

The celebrated theorem of Gallai and Sylvester ([S93], [G44]) asserts that any finite set of points in the real affine plane which is not collinear determines a line which passes through precisely two points of the set. An immediate corollary of this is an old conjecture of Erdös which is that any non-collinear set of n points in the plane determines at least n lines. There is more than one proof for this well known theorem, and some of the most striking proofs are algebraic ([1]).

Bichromatic set of points have been studied by many authors. A beautiful result of Motzkin ([2]) asserts that any set G of red and green points in the plane which is not collinear, determines a monochromatic line, that is, a line which passes through points from one color class only.

Analogies to Gallai-Sylvester theorem for bichromatic sets have also been studied. It is not difficult to come up with an example for a bichromatic set which is not collinear in which no line passes through precisely one red point and one green point, even when the number of red points equals the number of green points. In [PP00] it is shown that if G is a set of n red points and n green points that is not collinear, then one can always find a bichromatic line determined by G which contains at most two red points and at most two green points. From this one can easily show by induction that a set of n red points and n green points in the plane which is not collinear

determines at least $n + 1$ bichromatic lines. This is in fact tight as shown by taking $n - 1$ red points and n green points on one line and another red point outside of this line.

Can we get a stronger result under some natural conditions? In this paper we show that if one assumes that G is a set of n red points and n or $n - 1$ green points so that no color class is contained in a line, then G determines at least $|G| - 3$ bichromatic lines. The example of $n - 1$ red points and $n - 1$ green points on one line and another two points red and green which are collinear with one of the other points, shows that the best lower bound cannot be better than $|G| - 1$. Our proof will, in fact, imply this tight bound assuming that this bound is true for every G such that $|G| \leq 21$. We now state our main theorem:

Theorem 1.1. *Let G be a set of n red points and n or $n - 1$ green points. If no color class is contained in a line, then G determines at least $|G| - 3$ bichromatic lines.*

Conjecture 1.2. *Let G be a set of n red points and n or $n - 1$ green points. If no color class is contained in a line, then G determines at least $|G| - 1$ bichromatic lines.*

A proof of Conjecture 1.2 will lead to a colorful version of the above mentioned conjecture of Erdös. Indeed, take any set G of points in the plane, which is not contained in a line. One can always (except in the case where G has $|G| - 1$ collinear points) color it by two colors such the each color class is not collinear and the difference between the cardinalities of the two color classes is at most 1. Then, Conjecture 1.2, will guarantee that G determines at least $|G| - 1$ bichromatic lines and, by Motzkin theorem which we mentioned earlier, there is at least one monochromatic line, determined by G. We thus get a total of at least $|G|$ lines determined by G.

It is interesting to note that Theorem 1.1 is not valid if we omit the condition that the difference between the number of red and green points is small. This is shown by the following construction. Fix $d \geq 3$ and $k \geq d$. Consider the set $\{(x_1, \ldots x_d) \in R^d | 0 \leq x_i < k, x_i \in N\}$ in R^d, and color these points red. Then think of R^d as being embedded in the projective space of dimension d. Take the d points at infinity which correspond to the directions of the axis. Color these points green. Now project this configuration to the two dimensional affine plane by a projection which is in general position with respect to the configuration. The resulting planar set consists of k^d red points and d green points, and no color class is collinear. The number of bichromatic lines is precisely dk^{d-1}, since every green point lies on just k^{d-1} bichromatic lines. This number is much smaller than $k^d + d - 3$, which is the cardinality of the set of points minus 3.

2   Proof of the Main Theorem

We prove Theorem 1.1 by induction on $|G|$. Hence we may assume that there is no bichromatic line l which passes through exactly one red point and exactly one green

point. Indeed, otherwise if there exists such a line l, we can remove the red point on l thus killing at least one bichromatic line and conclude the theorem by induction (possibly by switching the roles of red and green points). Here one should be a little careful though, since it may happen that after removing one red point the set of red points become collinear. Nevertheless, this case can be easily resolved by inspection.

We will consider the case where there are n green points and the case where there are $n-1$ green points simultaneously. Let D denote the number of unordered pairs of points of G having different color, and Let S denote the number of unordered pairs of points of G having the same color.

If there are n green points, then $S - D = 2\binom{n}{2} - n^2 = -n$. If there are $n-1$ green points, then $S - D = \binom{n}{2} + \binom{n-1}{2} - n(n-1) = -n+1$. Therefore, we have $S - D \le -n+1$. We now estimate $S - D$ in a different way. We will count the contribution of each bichromatic line to the difference $S - D$. To be more precise, if L is a bichromatic line with r red points and g green points, then it contains rg pairs of points with different colors and $\binom{r}{2} + \binom{g}{2}$ pairs of points with the same color. Any such pair of points will not be counted by any other bichromatic line. Let $L_{r,g}$ $(r, g \ge 1)$ denote the number of bichromatic lines which contain exactly r red points and exactly g green points. By considering the contribution of every bichromatic line to the difference $S - D$, we will obtain a lower bound for $S - D$ since it may be that there are monochromatic lines which then contribute just to S.

We have:

$$S \ge \sum_{r,g \ge 1} \left(\binom{r}{2} + \binom{g}{2} \right) L_{r,g} \tag{1}$$

and

$$D = \sum_{r,g \ge 1} rg L_{r,g}. \tag{2}$$

The weak inequality in (1) is in fact a strict inequality. Indeed, there must be monochromatic lines. We know already that there are lines determined by G which contain exactly two points. Since by our assumption there is no such line which is bichromatic, all these lines must be monochromatic and thus contribute to S.

Let t_k denote the number of lines determined by G which pass through precisely k points. A classical result known as Cauchy's formula, which in fact follows very easily from Euler polyhedral formula, asserts that $t_2 \ge 3 + \sum_{k \ge 3} (k-3) t_k$.

Therefore, we may conclude:

$$t_2 \ge 3 + \sum_{r+g \ge 3} (r+g-3) L_{r,g}. \tag{3}$$

Since t_2 counts only monochromatic lines we get

$$S - D \geq 3 + \sum_{r+g \geq 3} (r + g - 3)L_{r,g} + \sum_{r,g \geq 1}(\binom{r}{2} + \binom{g}{2} - rg)L_{r,g} =$$

$$= 3 + \sum_{r+g \geq 3} (\frac{(r-g)^2 + r + g}{2} - 3)L_{r,g} \tag{4}$$

Combining this with $D - S \leq -n + 1$ we obtain

$$\sum_{r+g \geq 3} L_{r,g} \geq n + 2 + \sum_{r+g \geq 3} (((r-g)^2 + r + g)/2 - 2)L_{r,g} \tag{5}$$

Denote by X the left hand side of (5). X is in fact the number of bichromatic lines determined by G. We wish to show that $X \geq 2n - 3$. Observe that if $3 \geq r + g \leq 4$, then $((r-g)^2 + r + g)/2 - 2 \geq 0$. Hence, we have

$$X \geq n + 2 + \sum_{r+g \geq 5} (((r-g)^2 + r + g)/2 - 2)L_{r,g} \tag{6}$$

Let us consider once more the quantity D. We know that $D \geq n(n-1)$ (this is the worst case when there are $n-1$ green points and n red points). Therefore, from (2) we get:

$$\sum_{r+g \geq 3} rgL_{r,g} \geq n(n-1).$$

(The fact that the sum goes over all r, g such that $r + g \geq 3$ rather than $r, g \geq 1$ does not make any difference since we assume that $L_{1,1} = 0$.)

For $r + g \leq 4$, we have $rg \leq 4$. Moreover, since $\sum_{r+g \leq 4} L_{r,g} \leq X$, we get $4X + \sum_{r+g \geq 5} rgL_{r,g} \geq n(n-1)$, or more conveniently:

$$\sum_{r+g \geq 5} rgL_{r,g} \geq n(n-1) - 4X. \tag{7}$$

We wish to minimize the right hand side of (6) under the constraint in (7). This is a linear programming problem in the nonnegative variables $L_{r,g}$ for any two integer numbers r and g such that $r + g \geq 5$ and $r, g \geq 1$. We assume by contradiction that $X \leq 2n - 2$ (so that this inductive argument will in fact be valid also when we wish to prove a lower bound of $|G| - 1$ for the number of bichromatic lines determined by G, as long as the basis of induction is verified).

4

Claim 2.1. *There is no (bichromatic) line, determined by G, which contains n points or more.*

Proof. Clearly, if a line determines by G contains n points, then it must be bichromatic since otherwise there is a color class which is collinear. Suppose there is such a bichromatic line M, containing A red points and B green points, where $A + B \geq n$. If there are at least 2 red points and at least 2 green points not on M, then we obtain at least $A + B + (A - 1) + (B - 1)$ bichromatic lines connecting the points on M to one of the 4 points (2 red and 2 green) outside M. Together with the line M itself we obtain at least $2n - 1$ bichromatic lines.

Suppose there is only one red point outside M and at least 2 green points g_1, g_2 outside M. Then there are $n-1$ red points on M. There are at least $2n-3$ bichromatic lines connecting g_1 and g_2 to the $n-1$ red points on M. There is also the line M itself and a line which connects the red point outside M to a green point on M. Together we get a total of at least $2n - 1$ bichromatic lines.

Suppose there is only one green point outside M and at least 2 red points ,r_1 and r_2, outside M. If G has n green points, then this case is symmetric to the previous one. If G has only $n - 1$ green points, then there are $n - 2$ green points on M and at least 2 red points on M. There are at least $2n - 5$ bichromatic lines connecting r_1 and r_2 to the $n - 2$ red points on M. There is also the line M itself and two lines which connect the green point outside M to two red point on M. Together we get a total of at least $2n - 2 = |G| - 1$ bichromatic lines.

If there is only one red point outside M and only one green point outside M, we obtain $A + B$ bichromatic lines connecting the points on M to one of the two points not on M. Together with the line M itself, we have at least $A + B + 1 = |G| - 1$ bichromatic lines ∎

We introduce another variable (if it does not exist already) which is $L_{k,k}$ where $k = n/2$, and we will consider the inequality

$$X \geq n + 2 + L_{k,k}(k - 2) + \sum_{r+g \geq 5} L_{r,g}(((r - g)^2 + r + g)/2 - 2) \tag{8}$$

under the constrain:

$$L_{k,k}k^2 + \sum_{r+g \geq 5} L_{r,g}rg \geq n(n - 1) - 4X. \tag{9}$$

And assuming that $L_{k,k} \geq 0$ and real. Clearly, if we can show that X is always greater than $2n - 2$, then we are done. Indeed, by taking $L_{k,k} = 0$ we just get the original expression of the right hand side of (2).

By Claim 2.1, we may (and will) assume $L_{r,g} = 0$ for $r + g \geq n$. We do allow, however, to $L_{k,k}$ to be nonzero. It is not hard to see that the right hand side of (8) is

minimized when the only $L_{a,b}$ which is not zero is the one for which $\frac{((a-b)^2+a+b)/2-2}{ab}$ is minimum. This $L_{a,b}$ is usually $L_{k,k}$ unless $k < 6$ and then $L_{a,b}$ will be $L_{3,3}$.

In the first case we get, $L_{k,k} \geq (n(n-1) - 4X)/(\frac{n}{2})^2$ and from (8)

$$X \geq n + 2 + \frac{n(n-1) - 4X}{(\frac{n}{2})^2}(\frac{n}{2} - 2)$$

This implies $X > 2n - 2$, assuming that $n \geq 20$.

In the second case $L_{3,3} = (n^2 - 4X)/9$ and from (8) we obtain

$$X \geq n + 2 + \frac{n(n-1) - 4X}{9}.$$

This again implies $X > 2n - 2$, if we assume $n \geq 20$.

It is left to to show that Theorem 1.1 is true for $n < 20$. It is a bit difficult to check all these cases by hand. However, we are just confronted with a finite linear programming problem.

Consider first that case where there are n red points and n green points. Denote by $L_{r,g}$ the number of lines with exactly r red points and exactly g green points (this time we allow $r = 0$ or $g = 0$). By Claim 2.1 we may assume that $L_{r,g} = 0$ for $r + g \geq n$. Then we have the following constraints:

$$
\begin{aligned}
\sum_{r+g<n} L_{r,g}\binom{r}{2} &= \binom{n}{2} \\
\sum_{r+g<n} L_{r,g}\binom{g}{2} &= \binom{n}{2} \\
\sum_{r+g<n} L_{r,g}rg &= n^2 \\
\sum_{r+g<n} L_{r,g}(r + g - 3) &\leq -3
\end{aligned}
$$

(10)

The last equation is just a reformulation of Cauchy's formula. In addition we know that each $L_{r,g}$ is a nonnegative integer. Under these constrains we wish to minimize the number of bichromatic lines. Our linear objective function is $\sum_{r\geq1,g\geq1,r+g<n} L_{r,g}$. We can impose the additional restriction that $L_{1,1} = 0$ since we assume that there are no bichromatic lines through just two points.

Using a computer, this linear programming can be solved easily for every $n < 20$, to verify the lower bound of $2n - 3$ for the objective function.

We can do the same thing in the case where there are n red points and $n-1$ green points. We get the following constrains:

6

$$\sum_{r+g<n} L_{r,g}\binom{r}{2} = \binom{n}{2}$$

$$\sum_{r+g<n} L_{r,g}\binom{g}{2} = \binom{n-1}{2}$$

$$\sum_{r+g<n} L_{r,g}rg = n(n-1)$$

$$\sum_{r+g<n} L_{r,g}(r+g-3) \le -3$$

$$(11)$$

with the same linear objective function. Here again, it is straight forward to show, by using a computer, that for $n < 20$ the linear objective function is at least $2n - 4$ as required. This completes the proof of Theorem 1.1. ■

3   Concluding Remarks

It is readily seen from the proof of the inductive step in Theorem 1.1, that even a better bound could be proved as long as the small configurations of up to 40 points satisfy it. Unfortunately, it seems that the linear constraints presented in the proof are not enough to show the validity of the statement in Conjecture 1.2 for these small point sets. The largest n for which this obstacle occurs is $n = 11$ (namely, for a set with 22 or 21 points). Thus if one could verify the validity of the statement in Conjecture 1.2 for a set of points of size 22 and 21, then the conjecture would be true for every larger set of points.

On the other hand for those who do not trust computer for proving mathematical results, we just note that a weaker bound of $|G| - 5$ in Theorem 1.1 can be shown by hand using the same inductive proof given here.

Acknowledgments

References

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

[G44] T. Gallai, Solution of problem 4065, *American Mathematical Monthly* **51** (1944), 169–171.

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

[PP00] J. Pach and R. Pinchasi, Bichromatic lines with few points, *Journal of Combinatorial Theory* A **90** (2000), 326–335.

[S93] J. Sylvester, Mathematical question 11851, *Educational Times* **59** (1893), 98–99.