On the light side of geometric graphs

Eyal Ackerman* Rom Pinchasi†

August 14, 2011

Abstract

Let G be a geometric graph on n vertices in general position in the plane. Suppose that for every line ℓ in the plane the subgraph of G induced by the set of vertices in one of the two half-planes bounded by ℓ has at most k edges. Then G has at most $O(n\sqrt{k})$ edges. This bound is best possible.

1 Introduction

Let G be an n-vertex *geometric graph*. That is, a graph drawn in the plane such that its vertices are distinct points and its edges are straight-line segments connecting corresponding vertices. It is usually assumed, as we will assume in this paper, that the set of vertices of G is in general position in the sense that no three of them lie on a line.

Let ℓ be a line that does not contain any vertex of G (unless stated otherwise, we consider only such lines). Every edge of G either crosses ℓ, or is contained in one of the two half-planes bounded by ℓ. We say that G has a *k-light side* with respect to ℓ, if one of these half-planes contains at most k edges of G. If G has a *k-light side* with respect to *every* line ℓ, then G is *k-near bipartite*. We consider the following problem: What is the maximum number of edges of an n-vertex k-near bipartite geometric graph?

We will think of k as a function of n, that is $k = k(n)$, so obviously this question is interesting only when $k(n) = o(n^2)$. The following simple construction shows an $n\sqrt{k}$ lower bound. Let G be the geometric graph whose vertices is the set of vertices of a regular n-gon P. We denote the vertices of P (and of G) by $v_0, \ldots, v_{n-1}$, indexed in a clockwise order. The *cyclic distance* between two vertices, v_i, v_j, $i < j$, is defined as $\min\{j-i, i+n-j\}$. The edge set of G consists of all edges (v_i, v_j) such that the cyclic distance between v_i and v_j is at least $\lfloor n/2 - \sqrt{k} \rfloor$. The number of edges in G is at least $n\sqrt{k}$ as each vertex has degree at least $2\sqrt{k}$. One can easily verify that each half-plane bounded by a line that pass through the center of P contains at most k edges of G. It follows that if ℓ is a line not passing through the center of P, then the half-plane that is bounded by ℓ and does not contain the center of P also contains at most k edges. Therefore G is k-near bipartite.

Our main result shows that this construction is essentially best possible.

Theorem 1. *Every n-vertex k-near bipartite geometric graph has at most $O(n\sqrt{k})$ edges.*

*Department of Mathematics, Physics, and Computer Science, University of Haifa at Oranim, Tivon 36006, Israel. ackerman@sci.haifa.ac.il.

†Mathematics Department, Technion—Israel Institute of Technology, Haifa 32000, Israel. room@math.technion.ac.il

Related work. It is a common technique when studying Turán-type problems in geometric graphs to split the edge set into ones that are crossed by a certain line and to ones that are not and then claim (usually by induction) that the number of edges not crossed by the line is small (see, e.g., [1, 2, 3, 4, 5]). Fulek and Suk [1] studied geometric graphs that do not contain two disjoint copies of a certain geometric pattern. If there is a constant c such that an n-vertex geometric graph with at least cn edges must contain one copy of a certain geometric pattern, then a graph avoiding two disjoint copies of this pattern is cn-near bipartite and hence by Theorem 1 has $O(n^{3/2})$ edges. However, this bound is inferior to the $O(n \log n)$ bound found for this case in [1].

2 Proof of Theorem 1

Our proof requires that every k-near bipartite graph has a subgraph in which the degree of every vertex is $O(k)$ and the number of edges is high. When $k(n) = \Omega(n)$ the graph itself satisfies this property. For $k(n) = o(n)$ we use the following lemma, whose proof we postpone.

Lemma 2.1. *There are constants $c, d > 0$ such that the following holds. Let $G = (V, E)$ be a geometric graph on n vertices that is k-near bipartite. Then there exists a subgraph of G that has at least $c|E| - O(n)$ edges and the degree of each of its vertices is at most dk.*

Theorem 1 follows immediately from Lemmas 2.1 and 2.2.

Lemma 2.2. *Let $d > 0$ be a constant and let G be an n-vertex k-near bipartite graph such that the degree of every vertex in G is at most dk. Then there is another constant $a = a(d)$ such that G has at most $an\sqrt{k}$ edges.*

Proof. Denote by $A(\ell)$ and $B(\ell)$ the sets of vertices of G that lie above and below a line ℓ, respectively. If an edge e crosses two (non-vertical) lines ℓ and ℓ', then either one endpoint of e is in $A(\ell) \cap B(\ell')$ and the other is in $A(\ell') \cap B(\ell)$, or one endpoint of e is in $A(\ell) \cap A(\ell')$ and the other is in $B(\ell') \cap B(\ell)$. We say that in the first case the crossing is of type $ABAB$, while in the second case it is of type $AABB$.

Call a line ℓ *almost balanced* if each of the two half-planes bounded by ℓ contains at most $(d + 1)k$ edges of G. Notice that if ℓ is almost balanced, then there are at most $k + (d + 1)k = (d + 2)k$ edges of G not crossing ℓ. We first show that there is an almost balanced line with any given slope, and that for every almost balanced line there is another almost balanced line separating almost the same sets of vertices.

Proposition 2.3. *For every line ℓ there is a line ℓ' parallel to ℓ such that ℓ' is almost balanced.*

Proof. The proof is in fact just a continuity argument. Without loss of generality assume that ℓ is horizontal and that it contains at most k edges of G in the half-plane below it. Start translating ℓ upwards keeping track of the number of edges of G below it. Clearly, this number only increases and changes only when ℓ goes past a vertex of G. There is a first time where this number must be greater than k or else the number of edges of G is at most k and the lemma follows trivially (recall that $k = o(n^2)$). Assume therefore that as ℓ goes above a vertex x the number of edges below ℓ becomes greater than k. Since the degree of x is at most dk, the number of edges of G below ℓ is at most $(d + 1)k$. On the other hand because G is k-near bipartite and the number of edges of G below ℓ is greater than k it must be that the number of edges of G above ℓ is at most k. Hence we can take ℓ' to be this translation of the line ℓ. □

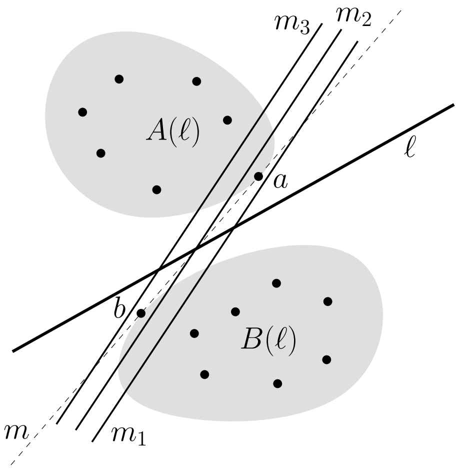


Figure 1: An illustration for the proof of Proposition 2.4.

Proposition 2.4. *Let ℓ be an almost balanced line. Then there exists an almost balanced line ℓ' such that $1 \leq |A(\ell) \cap B(\ell')| + |A(\ell') \cap B(\ell)| \leq 2$.*

Proof. Let m be the common tangent to the convex hulls of $A(\ell)$ and $B(\ell)$ that separates them such that the slope of m is greater than the slope of ℓ. Let $a \in A(\ell)$ and $b \in B(\ell)$ be the points that determine the line m. Rotate m counterclockwise just a little bit to obtain a line m'. Let m_1 be a line very close to m that is parallel to m' and has both a and b above it. Let m_2 be a line very close to m that is parallel to m' and separates a and b. Finally let m_3 be a line very close to m that is parallel to m' and has both a and b below it (see Figure 1). Notice that $1 \leq |A(\ell) \cap B(m_i)| + |A(m_i) \cap B(\ell)| \leq 2$, for $i = 1, 2, 3$. We will conclude the proof of the lemma by showing that one of m_1, m_2, and m_3 is almost balanced.

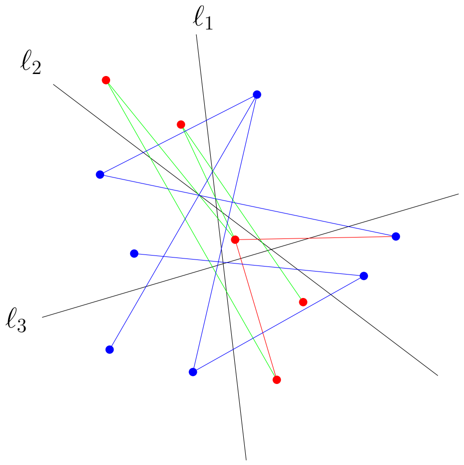
Because ℓ is almost balanced, then the number of edges of G below m_1 is at most $(d+1)k$. If it is more than k, then, as G is k-near bipartite, the number of edges of G above m_1 is at most k and therefore m_1 is almost balanced.

Hence assume that the number of edges of G below m_1 is at most k. It follows that the number of edges of G below m_2 is at most $(k+1)d$. If it is more than k, then, as G is k-near bipartite, the number of edges of G above m_2 is at most k and therefore m_2 is almost balanced.

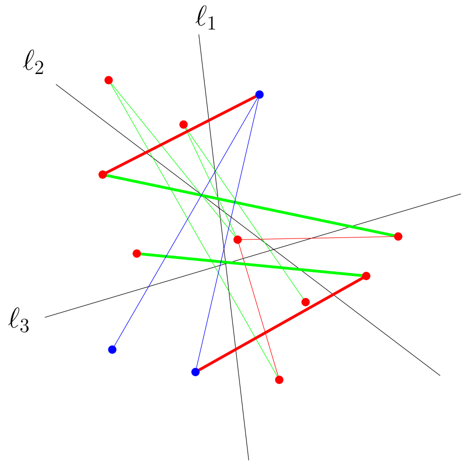
Hence assume that the number of edges of G below m_2 is at most k. It follows that the number of edges of G below m_3 is at most $(d+1)k$. If it is more than k, then, as G is k-near bipartite, the number of edges of G above m_3 is at most k and therefore m_3 is almost balanced. Therefore the number of edges of G below m_3 is at most k. On the other hand the number of edges of G above m_3 is less than the number of edges of G above ℓ which is at most $(d+1)k$. Therefore, m_3 is almost balanced. □

The strategy in the rest of the proof is to find roughly $n/\sqrt{k}$ almost balanced lines $\ell_1, \ldots, \ell_t$ with distinct directions. The number of those edges that are not crossed by at least one of these lines is at most $(d+2)kt$. To estimate from above those edges that cross all lines $\ell_1, \ldots, \ell_t$ observe that such edges have both of their vertices in two "opposite" unbounded faces of the arrangement of lines $\ell_1, \ldots, \ell_t$. We will choose the lines $\ell_1, \ldots, \ell_t$ so that the number of such edges will be small.

We choose the lines ℓ_i one by one, and with increasing slopes. Initially, all the edges and vertices of G are colored blue. Recall that by Proposition 2.3 there is an almost balanced line with any given slope and let ℓ_1 be an almost balanced line with a very small slope, such



(a) before ℓ_3 is added



(b) after ℓ_3 is added

Figure 2: An example for adding a new line

that there are no two vertices of G the line determined by which has a slope smaller than the slope of ℓ_1. We recolor all the blue edges of G not crossing ℓ_1 with red.

Suppose that we have already chosen the lines $\ell_1, \ldots, \ell_i$. If the number of remaining blue edges is at most $(3d+2)k$, we stop. Otherwise we choose a new line ℓ_{i+1} with a slope that is greater than the slope of ℓ_i (in a way that is specified later). Once ℓ_{i+1} is chosen, all the blue edges that ABAB-cross ℓ_i and ℓ_{i+1} are colored green. The blue vertices in $(A(\ell_i) \cap B(\ell_{i+1})) \cup (A(\ell_{i+1}) \cap B(\ell_i))$ are colored red, as well as any blue edge that is adjacent to one of them. See Figure 2 for an example. It is not hard to see that the following invariants are maintained after ℓ_{i+1} is added:

(1) The endpoints of any remaining blue edge are blue.

(2) Every remaining blue edge crosses all the lines $\ell_1, \ldots, \ell_{i+1}$.

(3) Every red edge does not cross at least one of the lines $\ell_1, \ldots, \ell_{i+1}$.

(4) Each vertex in $\bigcap_{j=1}^{i+1} A(\ell_j)$ and $\bigcap_{j=1}^{i+1} B(\ell_j)$ is blue. The rest of the vertices are red.

Invariants (1)–(3) are straightforward. The last invariant can be proved by induction on the number of lines: It clearly holds after ℓ_1 is chosen. If a vertex v is colored red, then there was $1 < j \leq i+1$ such that $v \in (A(\ell_{j-1}) \cap B(\ell_j)) \cup (A(\ell_j) \cap B(\ell_{j-1}))$, therefore $v \notin \bigcap_{j=1}^{i+1} A(\ell_j) \cup \bigcap_{j=1}^{i+1} B(\ell_j)$. Conversely, if $v \notin \bigcap_{j=1}^{i+1} A(\ell_j) \cup \bigcap_{j=1}^{i+1} B(\ell_j)$, then there are $1 \leq x, y \leq i+1$ such that $v \in A(\ell_x) \cap B(\ell_y)$. Assume without loss of generality that $x < y$. If $y < i+1$ then v is red by the induction hypothesis. Otherwise if $y = i+1$ and v is blue, then by the induction hypothesis $v \in A(\ell_i)$ (because $v \notin B(\ell_x)$ it is impossible that $v \in \bigcap_{j=1}^{i} B(\ell_j)$) and therefore it will be colored red after ℓ_{i+1} is added.

The line ℓ_{i+1} is chosen using the following proposition.

Proposition 2.5. *Suppose that the number of blue edges is greater than $(3d+2)k$. Then there exists an almost balanced line ℓ_{i+1} with the following properties.*

1. *The slope of ℓ_{i+1} is greater than the slope of ℓ_i.*

2. *The number of blue edges that ABAB-cross ℓ_i and ℓ_{i+1} is at most $4dk$.*

3. *The number of blue vertices in $(A(\ell_{i+1}) \cap B(\ell_i)) \cup (A(\ell_i) \cap B(\ell_{i+1}))$ is at least $\sqrt{2dk}$.*

Proof. We repeatedly apply (the proof of) Proposition 2.4 and find lines m_j, $j = 1, 2, \ldots$, with increasing slopes that are almost balanced and at each step the number of vertices of G in $(A(m_j) \cap B(\ell_i)) \cup (A(\ell_i) \cap B(m_j))$ changes by at most 2. Hence the number of blue edges that ABAB-cross ℓ_i and m_j changes at each step by at most $2dk$. Once the number of these blue edges is greater than $2dk$ (and is therefore at most $4dk$), we stop and set $\ell_{i+1} = m_j$. Notice that upon stopping the number of blue vertices in $(A(\ell_i) \cap B(\ell_{i+1})) \cup (A(\ell_{i+1}) \cap B(\ell_i))$ is at least $\sqrt{2dk}$. This is because the vertices of every blue edge are both blue.

It remains to show that we indeed stop at some point. Suppose we do not and let j be the smallest index such that m_j has a positive slope while m_{j+1} has a negative slope. There are at most $(d+2)k$ blue edges that do not cross m_{j+1} since m_{j+1} is almost balanced. Let e be a blue edge that ABAB-crosses ℓ_1 and m_j. It follows from Invariants (1),(2), and (4) above, that e also ABAB-crosses ℓ_i and m_j and therefore there are at most $2dk$ such edges. Since there are more than $(3d+2)k$ remaining blue edges, there must be a blue edge (u, v) that AABB-crosses ℓ_1 and m_j, that is $u \in A(\ell_1) \cap A(m_j)$ and $v \in B(\ell_1) \cap B(m_j)$. The slope of the line determined by u and v cannot be negative since then it will be smaller than the slope of ℓ_1. This implies that when applying (the proof of) Proposition 2.4 the next line m_{j+1} would have a positive slope, which is a contradiction. □

We are now ready to complete the proof of Lemma 2.2. Suppose that the process described above stops after t lines have been chosen. Every edge of G is either blue, red, or green. The number of blue edges is at most $3d + 2$. The number of green edges is at most $4dkt$. The number of red edges is at most $(d+2)kt$, since each of the lines we choose is almost balanced and therefore there at most $(d+2)k$ edges that do not cross it. Because we color red at least $\sqrt{2dk}$ vertices of G when adding a new line, it follows that $t \leq \frac{n}{\sqrt{2dk}}$. Therefore the number of edges of G is at most $(3d+2) + (5d+2)\sqrt{\frac{k}{2d}} \cdot n$. This concludes the proof of Lemma 2.2. □

It remains to prove Lemma 2.1.

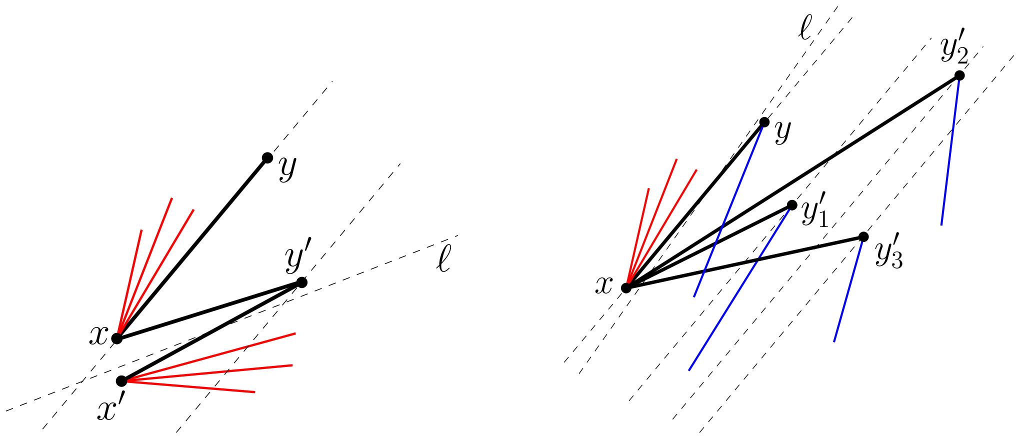
Proof of Lemma 2.1: Let $G = (V, E)$ be a geometric graph on n vertices that is k-near bipartite. We will show that there exists a subgraph of G that has at least $|E|/20 - 4n$ edges and the degree of each of its vertices is at most $12k$ (in order to simplify the presentation we do not attepmt to optimize these constants).

For every vertex x of G denote by $d(x)$ the degree of x in G. Divide the edges adjacent to x into two sets, those that go to the left and those that go to the right. Color red the $\lceil \frac{1}{10} d(x) \rceil$ edges going to the right from x that have the largest slopes, as well as the $\lceil \frac{1}{10} d(x) \rceil$ edges going to the right from x that have the smallest slopes. Do the same for the edges going to the left from x. The number of edges colored red is at most $\sum_x 4\lceil \frac{1}{10} d(x) \rceil \leq \sum_x 4(\frac{1}{10} d(x) + 1) = \frac{4}{5}|E| + 4n$. Remove all the red edges from G to obtain a subgraph $G_1 = (V, E_1)$ with at least $|E|/5 - 4n$ edges.

Let P denote the set of vertices whose degree in G_1 is at least $12k$ and let $Q = V \setminus P$. Of course, if P is empty, then we are done and G_1 is the desired subgraph.

Proposition 2.6. *There is no edge $(x, y) \in E_1$ such that $x, y \in P$.*

Proof. Assume to the contrary that $x, y \in P$ are connected by an edge e in G_1. Since $d_1(x), d_1(y) \geq 12k$ it follows that $d(x), d(y) \geq 20k$. Without loss of generality assume that x is to the left of y and the slope of e is positive. Because e was not colored red as an edge adjacent to x nor as an edge adjacent to y, we can conclude that in G there are at least $20k/10 = 2k$ edges adjacent to x going to the right with a greater slope than the slope of e and there are at least $2k$ edges adjacent to y going to the left with a greater slope than the slope of e. Consider the line ℓ containing e and slightly rotate it counterclockwise around the



(a) If the slope of (x, y) is greater or equal to the slope of (x', y') then G does not have a k-light side w.r.t. ℓ.

(b) If $d_3(x) \geq 12k$ then G does not have a k-light side w.r.t. ℓ.

Figure 3: Illustrations for the proof of Lemma 2.1.

midpoint of e. Then there are at least $2k$ edges of G in each of the two half-planes bounded by ℓ. This is a contradiction to the assumption that G is k-near bipartite. $\square$

We may assume, without loss of generality, that at least $1/4$ of the edges in G_1 that connect a vertex $p \in P$ and a vertex $q \in Q$ are such that p is to the left of q and the edge (p, q) has a positive slope. Thus, by removing all the other edges connecting a vertex in P and a vertex in Q, we obtain a subgraph $G_2 = (V, E_2)$ such that $|E_2| \geq |E_1|/4 \geq (|E|/5 - 4n)/4 = |E|/20 - n$.

We will show that by removing at most n edges from G_2, we obtain the desired graph. To this end, the following observation will be useful.

Proposition 2.7. *Let $(x, y), (x, y'), (x', y')$ be edges in G_2 such that $x, x' \in P$, $y, y' \in Q$, and the slopes of both (x, y) and (x', y') are greater than the slope of (x, y'). Then (x', y') has a greater slope than (x, y).*

Proof. Suppose that the slope of (x, y) is greater or equal to the slope of (x', y') (see Figure 3(a)). Note that if we slightly rotate clockwise the line containing (x', y') around the midpoint of (x', y') then the resulting line ℓ separates (red) edges in G going right from x with a slope that is greater than the slope of (x, y) and edges going right from x' with a slope is smaller than the slope of (x', y'). However, each of these two sets of edges contains at least $2k$ red edges since $x, x' \in P$, and therefore G does not have a k-light side w.r.t. ℓ, which is a contradiction. $\square$

Next, for every vertex x in G_2 we color blue the edge with the greatest slope that is adjacent to x. Let $G_3 = (V, E_3)$ be the subgraph we obtain by removing all the blue edges from G_2. Then $|E_3| \geq |E_2| - n \geq |E|/20 - 2n$. Denote by $d_3(x)$ the degree in G_3 of a vertex x. We claim that $d_3(x) \leq 12k$ for every vertex x and therefore G_3 is the desired graph.

Suppose that G_3 contains a vertex x such that $d_3(x) \geq 12k$. Therefore, $x \in P$. Let (x, y) be the edge with the greatest slope that is adjacent to x in G_3, and let (x, y') be a different edge (with a smaller slope). It follows from Proposition 2.7 that every edge (x', y') in G_2 has a greater slope than the slope of (x, y). Therefore, for every neighbor y' of x there is at least one blue edge (x', y') whose slope is greater than the slope of (x, y). If we slightly rotate counterclockwise the line containing (x, y) around the midpoint of (x, y), then the resulting line ℓ separates these edges and red edges in G that are going right from x and whose slope is greater than the slope of (x, y) (see Figure 3(b)). However, each of these two sets of edges contains at least $2k$ edges, therefore G does not have a k-light side w.r.t. ℓ, which is a contradiction. This concludes the proof of Lemma 2.1. □

References

[1] R. Fulek and A. Suk, On disjoint crossing-families in geometric graphs, *EuroComb* 2011, to appear.

[2] J. Pach, R. Pinchasi, G. Tardos, G. Tóth, Geometric graphs with no self-intersecting path of length three, *European J. Combinatorics* **25** (2004), no. 6, 793–811.

[3] G. Tóth and P. Valtr, Geometric graphs with few disjoint edges, *Discrete Comput. Geom.* **22** (1999), 633–642.

[4] P. Valtr, Graph drawings with no k pairwise crossing edges, In G. D. Battista, editor, *Graph Drawing*, volume 1353 of *Lecture Notes in Computer Science*, 205–218, Springer, 1997.

[5] P. Valtr, On geometric graphs with no k pairwise parallel edges, *Discrete Comput. Geom.* **19** (1998), no. 3, 461–469.