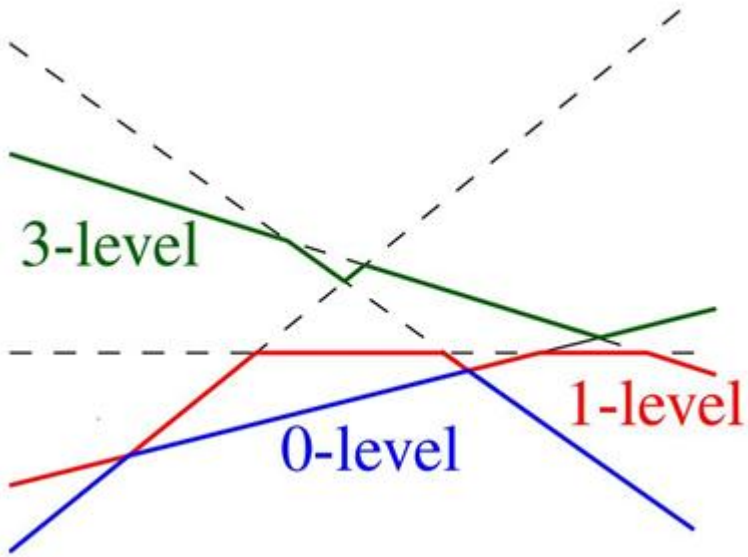
Depth Estimation via Sampling

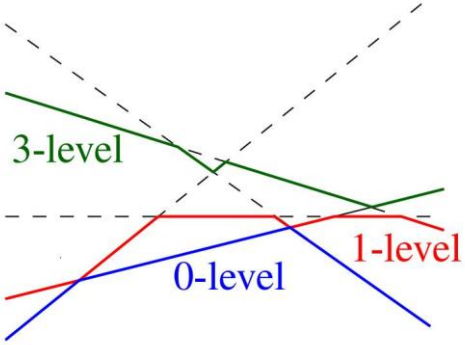
Eliad Tsfadia

The at most k-levels

- Let L be n lines in the plane.
- Every intersection point between two line is called a vertex.
- A point $p \in \bigcup_{l \in L} l$ is of **level $\boldsymbol{k}$** if there are exactly k lines of L strictly below it.
- The **$\boldsymbol{k}$-level** is a curve along the lines of L that passes through points in level k.



Questions (Upper Bounds)



- What is the number of vertices in the 0-level ?

The 0-level as at most $n-1$ vertices (each line might contribute at most one segment to the 0-level)

- What is the number of vertices in the k-level ?

This is a surprisingly hard question.

- What is the number of vertices of level $\leq k$?

<u>Theorem 1</u>: For $k > 1$, The number of vertices in $A(L)$ of level $\leq k$ is $O(nk)$.
<u>Proof</u>:

- Let $V_{\leq k} = V_{\leq k}(L)$ be the set of all vertices of level $\leq k$ in $A(L)$
- Pick a random sample R of L, by picking each line to be in the sample with probability $\frac{1}{k}$. Observe that $E[|R|] = \frac{n}{k}$
- For any vertex $v \in A(L)$, Let X_v be an indicator variable which is 1 if v is a vertex of 0-level in $A(R)$, and 0 otherwise.
- For $v \in V_{\leq k}$ $(assume\ v$ is a vertex in level j, for $0 \leq j \leq k)$ what is the probability that X_v = 1 ?
- $\Pr[X_v = 1] = (1 - \frac{1}{k})^j (\frac{1}{k})^2 \geq (1 - \frac{1}{k})^k \frac{1}{k^2} \geq e^{\left(-2\frac{k}{k}\right)} \frac{1}{k^2} = \frac{1}{e^2 k^2}$

(note that $1 - x \geq e^{-2x}$ for $0 < x \leq \frac{1}{2}$)

- We saw that for $v \in V_{\leq k}$, $E[X_v] = \Pr[X_v = 1] \geq \frac{1}{e^2 k^2}$
- On the other hand, the number of vertices on the 0-level of $\mathrm{A}(R)$ is at most $|R| - 1$.
- As such: $\sum_{v \in V_{\leq k}} X_v \leq \sum_{v \in A(L)} X_v \leq |R| - 1$
- In expectation: $\mathrm{E}[\sum_{v \in V_{\leq k}} X_v] \leq E[|R| - 1] = \frac{n}{k} - 1 \leq \frac{n}{k}$
- On the other hand: $\mathrm{E}[\sum_{v \in V_{\leq k}} X_v] = \sum_{v \in V_{\leq k}} E[X_v] \geq \frac{|V_{\leq k}|}{e^2 k^2}$
- Finally, we get: $\frac{|V_{\leq k}|}{e^2 k^2} \leq \frac{n}{k}$. Namely, $|V_{\leq k}| \leq e^2 nk = O(nk)$ ■

Extention of Theorem 1 - Clarkson Shor (89):

- Suppose we have a set S of n objects (lines).
- The objects defines configurations (vertices).
- Each configuration is defined by at most d objects ($d = 2$).
- Each configuration is in conflict with some of the objects (a vertex is in conflict with all the lines that below it)
- The weight/depth of a configuration is the number of objects that are in conflict with it (The level of a vertex).
- $N_{\leq k}(S) =$ The number of configurations in S with weight $\leq k$ ($|V_{\leq k}|$).
- $N_{\leq k}(n) = max_{|S|=n} N_{\leq k}(S)$ (the maximum of $|V_{\leq k}(L)|$ for every $|L| = n$).
- Then: $N_{\leq k}(n) = O\left(k^d N_0\left(\frac{n}{k}\right)\right)$ ($N_{\leq k} = O\left(k^2 N_0\left(\frac{n}{k}\right)\right) = O\left(k^2 \cdot \frac{n}{k}\right) = O(nk)$).

<u>Proof (Reproduce of the proof of Theorem 1)</u>:

- Let S be a set of n objects.
- Let $C_{\leq k} = C_{\leq k}(S)$ be the set of all configurations of weight $\leq$ k .
- Pick a random sample R of S, by picking each line to be in the sample with probability $\frac{1}{k}$. Observe that $E[|R|] = \frac{n}{k}$.
- For any configuration c, Let X_c be an indicator variable which is 1 if c is a configuration of weight 0 in R, and 0 otherwise.
- For $c \in C_{\leq k}$ (assume c is a configuration with weight j in S, for $0 \leq j \leq k$) what is the probability that X_c = 1 ?
- $\Pr[X_c = 1] = (1 - \frac{1}{k})^j (\frac{1}{k})^d \geq (1 - \frac{1}{k})^k \frac{1}{k^d} \geq e^{\left(-2\frac{k}{k}\right)} \frac{1}{k^d} = \frac{1}{e^2 k^d}$

(note that $1 - x \geq e^{-2x}$ for $0 < x \leq \frac{1}{2}$)

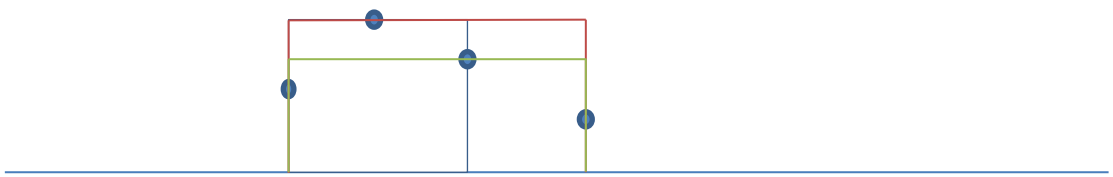
- We saw that $E[X_c] = \Pr[X_c = 1] \geq \frac{1}{e^2 k^d}$
- As such: $\sum_{c \in C_{\leq k}} X_c \leq \sum_{c \in C_{\leq n}} X_c = N_0(R) \leq N_0(|R|)$
- In expectation: $\mathrm{E}[\sum_{c \in C_{\leq k}} X_c] \leq E[N_0(|R|)]$
- On the other hand: $\mathrm{E}[\sum_{c \in C_{\leq k}} X_c] = \sum_{c \in C_{\leq k}} E[X_c] \geq \frac{|C_{\leq k}|}{e^2 k^d}$
- We get: $\frac{|C_{\leq k}|}{e^2 k^d} \leq E[N_0(|R|)]$. Namely, $N_{\leq k}(S) = |C_{\leq k}| = O(k^d E[N_0(|R|)])$
- Does $E[N_0(|R|)] = O\left(N_0\left(\frac{n}{k}\right)\right)$?

- **<u>Lemma</u>**: If $N_0(n) = O(poly(n))$, Then $E[N_0(|\mathrm{R}|)] = O\left(\mathrm{N}_0\left(\frac{\mathrm{n}}{\mathrm{k}}\right)\right)$.
- If we assume the condition in this lemma, we get the final result:

$$N_{\leq k}(S) = O\left(k^d N_0\left(\frac{n}{k}\right)\right)$$

- This is a bound for every S. Thus, $N_{\leq k}(n) = O\left(k^d \mathrm{N}_0\left(\frac{\mathrm{n}}{\mathrm{k}}\right)\right)$ ■

- Clarkson-Shor: $N_{\leq k}(n) = O\left(k^d N_0\left(\frac{n}{k}\right)\right)$
- **<u>Example</u>**: Let $P \subseteq \mathbb{R}^2$ be a set of n points above the x-axis. Let $\mathcal{R}$ be the set of rectangles that their bottom edge is lying on the x-axis and that contain a point of P on each of their other three edges.



- What is the number of rectangle in $\mathcal{R}$ that contain at most k points of P?
- **<u>Solution</u>**: Can we bound $N_0(n)$? (the number of empty rectangles)
- $N_0(n) \leq n = O(n)$ (each point $p \in P$ can have at most one such empty rectangle containing p on its upper edge)
- By Clarkson-Shor we get: $N_{\leq k}(n) = O\left(k^3 \cdot \frac{n}{k}\right) = O(k^2 n)$.

- **<u>Another Example</u>**: Let P be a set of n points in $\mathbb{R}^d$. A region is a halfspace with d points on its boundary (the halfspace above the points). What is the number of different halfspaces containing at most k points of P ?
- **<u>Solution</u>**:
- <u>Theorem without proof</u>: the complexity of the convex hull of n points in d dimensions is bounded by $O(n^{\lfloor \frac{d}{2} \rfloor})$.
- By the theorem we get: $N_0(n) = O(n^{\lfloor \frac{d}{2} \rfloor})$.
- By Clarkson-Shor we get: $N_{\leq k}(n) = O\left(k^d \left(\frac{n}{k}\right)^{\lfloor \frac{d}{2} \rfloor}\right) = O\left(k^{\lceil \frac{d}{2} \rceil} n^{\lfloor \frac{d}{2} \rfloor}\right)$.

- We saw that by Clarkson-Shor we can give a good upper bound for $N_{\leq k}(n)$ in many different problems, If we have a good bound for $N_0(n)$.
- What about $N_k(n)$?
- If we stay at the first example (the k-level problem), can we give a good upper bound for the number of vertices in the k-level which is smaller than $O(nk)$?
- We will see later a bound of $O(nk^{\frac{1}{3}})$ to the number of vertices in the k-level.

The Crossing Lemma

- **<u>Theorem 2 (Euler's formula)</u>**: For a connected planar graph G with n vertices, m edges and f faces we have: $n + f = m + 2$.

  If G is not necessarily connected, we have: $n + f \geq m + 2$.

- **<u>Lemma 3</u>**: If G is a simple planar graph, then $m \leq 3n - 6$ and $f \leq 2n - 4$.

- **<u>Definition</u>**: For any simple graph G, denote as $c(G)$ as the minimal number of edge crossing in any drawing of G.

- **<u>Claim 4</u>**: For any simple graph G, $c(G) \geq m - 3n + 6$

<u>Claim 4</u>: For any simple graph G, $c(G) \geq m - 3n + 6$

<u>Proof</u>:

- Draw G with $c(G)$ crossings.
- Let H be the graph resulting from G by removing one of the edges from each pair of crossing edges.
- We have $m(H) \geq m - c(G)$.
- Note that H is Planar. By Lemma 3 we get $m(H) \leq 3n(H) - 6 = 3n - 6$
- We get: $m - c(G) \leq 3n - 6$, Which is equivalent to

  $c(G) \geq m - 3n + 6$. ■

<u>Lemma 5 (The Crossing Lemma)</u>: For a simple graph G with $m \geq 6n$ we have $c(G) = \Omega(\frac{m^3}{n^2})$.

<u>Proof</u>:

- Draw G with $c(G)$ crossings, and let U be a random subset of $V(G)$ selected by choosing each vertex with probability p (We will choose p later).
- Let $H = G_U = (U, E')$ be the induced subgraph of G over U.

  $E' = \{uv | uv \in E(G) \text{ and } u, v \in U\}$.
- What is the Probability that a vertex v survived?

Exactly p

- What is the probability that an edge of G survived?

Exactly p^2

- What is the probability of a crossing to survived?

Exactly p^4

- Let N_v , N_e and N_c denote the number of vertices, edges and crossing that survived, respectively. Then, $E[N_v] = np$, $E[N_e] = mp^2$, $E[N_c] = c(G)p^4$
- By Claim 4 we have: $N_c \geq c(H) \geq N_e - 3N_v + 6$
- In particular: $E[N_c] \geq E[N_e] - 3E[N_v] + 6$
- This implies that: $c(G)p^4 \geq mp^2 - 3np + 6$
- We get: $c(G) \geq \frac{m}{p^2} - \frac{3n}{p^3} + \frac{6}{p^4} \geq \frac{m}{p^2} - \frac{3n}{p^3}$
- In particular, by setting $p = \frac{6n}{m} \leq 1$ we have that:

$$c(G) \geq \frac{m}{\left(\frac{6n}{m}\right)^2} - \frac{3n}{\left(\frac{6n}{m}\right)^3} = \frac{m^3}{36n^2} - \frac{m^3}{72n^2} = \frac{m^3}{72n^2} = \Omega\left(\frac{m^3}{n^2}\right) \blacksquare$$

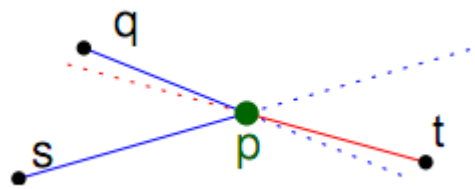
Duality

- A point $p = (a, b)$ is transferred to the line $p^* = \{y = ax + b\}$.
- A line $l = \{y = cx + d\}$ is transferred to the point $l^* = (-c, d)$.
- We can extent this duality to dill with lines in the form $\{x = a\}$ in the projective plane with Homogeneous coordinates. For the sake of simplicity, assume there aren't such lines.
- $p = (a, b)$ is on $l = \{y = cx + d\} \Leftrightarrow$
  $ac + d = b \Leftrightarrow -ac + b = d \Leftrightarrow$
  $\boldsymbol{l}^* = (-c, d)$ is on $p^* = \{y = ax + b\}$
- $p = (a, b)$ is above $l = \{y = cx + d\} \Leftrightarrow$
  $ac + d < b \Leftrightarrow -ac + b > d \Leftrightarrow$
  $p^* = \{y = ax + b\}$ is above $\boldsymbol{l}^* = (-c, d)$

On the number of k-sets
(Application of the crossing lemma)

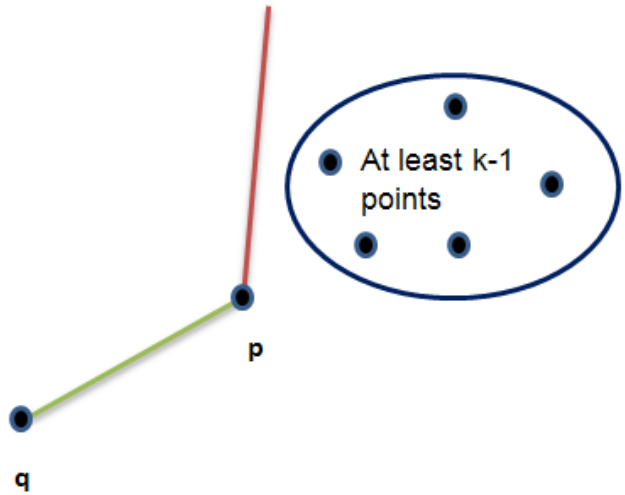
- Let P be a set of n points in general position.
- A pair of points $p, q \in P$ form a k-set if there are exactly k points in the (closed) halfplane below the line passing through p and q.
- Consider the graph $G = (P, E)$ that has an edge for every k-set.
- We want to bound the size of E as a function of n and k.
- Observe that via duality we have that the number of k-sets is exactly the number of vertices in the $(k-2)-$level in the dual arrangement $A(P^*)$.
- Thus, if we can bound the size of E, we also can bound the number of vertices in the k-level.

- **<u>Lemma 6</u>**: Let qp and sp be two k-set edges of G, with q and s to the left of p. Then there exists a point $t \in P$ to the right of p such that pt is a k-set, and $line(p,t)$ lies between $line(q,p)$ and $line(s,p)$



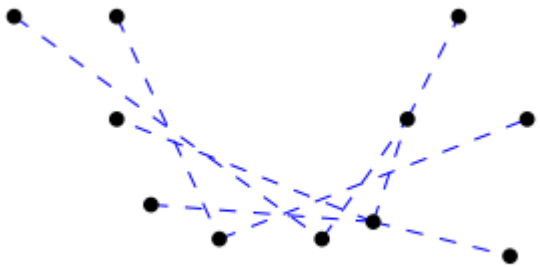
- **<u>Proof sketch:</u>**
- Below (and include) each of $line(q,p)$ and $line(s,p)$ there are exactly k points of P.
- If we increase by small ε the slope of $line(q,p)$ (around p) we get that there are exactly $k-1$ points below this line (because it won't include q)
- If we decrease by small ε the slope of $line(s,p)$ (around p) we get that there are exactly k points below this line (because it will still include s).
- If we continue increase the slope of $line(q,p)$, we have to get from value $k-1$ to value k somewhere before we get to s.
- These means that we got to a new point t that has to be on the right of p (why?) and below $line(p,t)$ there are exactly k points.

- **<u>Lemma 7</u>**: Let $p \in P$ and Let $q \in P$ be a point to the left of p so that $qp \in E$. Furthermore, assume that there are at least $k-1$ points of P to the right of p. Then, there exists a point $t \in P$ such that $pt \in E$ and pt has larger slope than qp.
- **<u>Proof sketch</u>**:

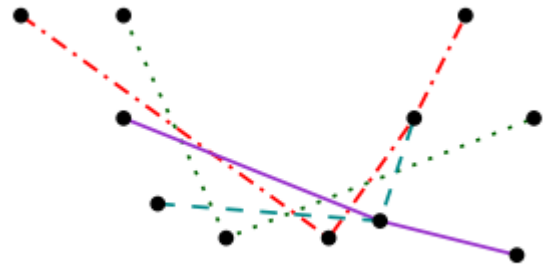


- Below (and include) $line(q,p)$ there are exactly k points of P.
- If we increase by small ε the slope of $line(q,p)$ (through p) we get value $k-1$
- If we increase the slope to infinity, we get value $\geq$ k.
- If we continue increasing the slope from qp we have to get from value $k-1$ to k.
- These means that we got to a new point t in the right of p and below $line(p,t)$ there are exactly k points, and slope(pt) > slope(qp)

- **<u>Lemma 8</u>**: The edges of G can be decomposed into $k-1$ convex chains $C_1, \dots, C_{k-1}$.

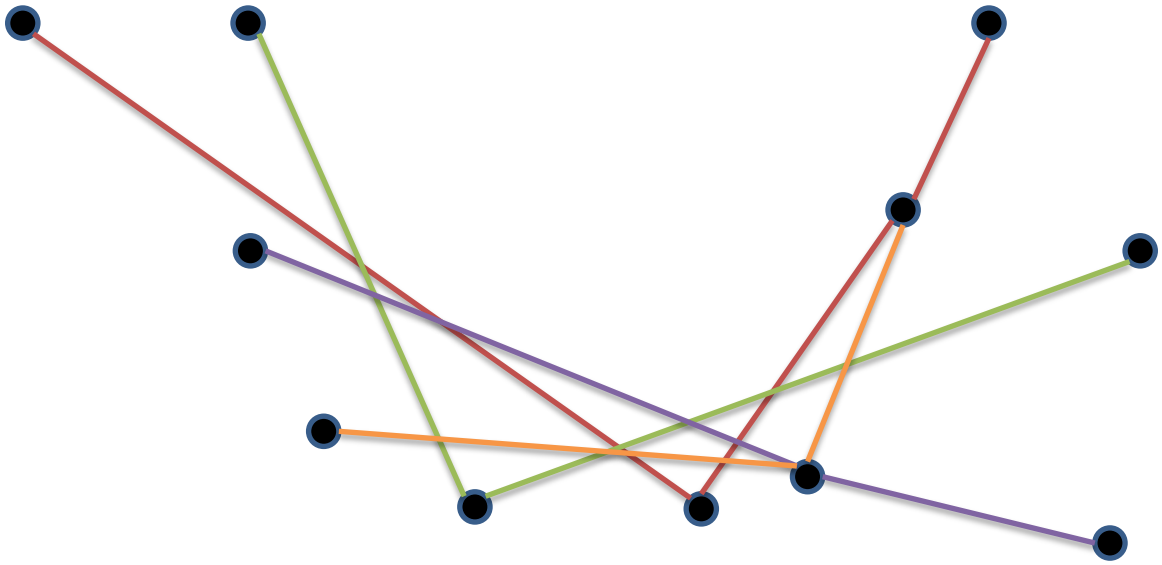


all 5-sets

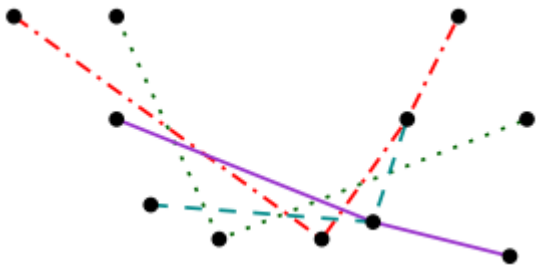


convex chains

- **<u>Proof (by iterative algorithm):</u>**
- The algorithm will have $k-1$ iterations.
- In each iteration, choose the left most point that still has an edge in E that hasn't been chosen yet. Suppose the point is q and the edge is $e = qp \in E$.
- Rotate a line around p (counterclockwise) till we encounter an edge $e' = ps \in E$ where s is a point to the right of p.
- We can now walk from e to e' and continue walking in this way, forming a chain of edges in G.
- **<u>Claim:</u>** in the end of the algorithm, we get $k-1$ edge disjoint convex chains that contain all the edges in G.



- **<u>Claim:</u>** In the end of the algorithm, we get $k-1$ edge disjoint convex chains that contain all the edges in G.



- **<u>Proof</u>**:

(1) The chains are convex since we rotate counterclockwise as we walk along a chain.

(2) By Lemma 6, no two chains can be "merged" into using the same edge.

(3) By Lemma 6, no two chains can end in the same point.

(4) By Lemma 7, such a chain can end only in the last $k-1$ points of P.

(5) If there was an edge $e = pq \in E$ that hasn't been chosen, we could create a convex chain from this edge, in contradiction to (2) and (3).

- **<u>Lemma 8</u>**: The edges of G can be decomposed into $k-1$ convex chains $C_1, \dots, C_{k-1}$.
- Similary, the edges of G can be decomposed into $n-k+1$ concave chains $D_1, \dots, D_{n-k+1}$.

- **<u>Proof of second part</u>:**
- Rotate the plane by $180°$. Every k-set is now $(n-k+2)$-set.
- By the first argumentation, the edges of G can be decomposed into $n-k+1$ convex chains, which are concave in the original orientation. ■

- **<u>Theorem 9</u>**: $|E| = O(nk^{\frac{1}{3}})$, which means that the number of k-sets defined by a set of n points in the plane is $O(nk^{\frac{1}{3}})$.
- **<u>Proof</u> :**
- The graph G has $|P| = n$ vertices. Let $m = |E(G)|$ be the number of k-sets.
- By Lemma 8, the edges can be decomposed into convex chains $C_1, \dots, C_{k-1}$ and can be decompose also to concave chains $D_1, \dots, D_{n-k+1}$.
- Any crossing of two edges of G is an intersection point of one convex chain of $C_1, \dots, C_{k-1}$ to one concave chain of $D_1, \dots, D_{n-k+1}$.
- Since a convex chain and a concave chain can have at most two intersections, we conclude that there are at most $2(k-1)(n-k+1) = O(nk)$ crossing in G.
- By the crossing lemma (Lemma 5) there are at least $\Omega\left(\frac{m^3}{n^2}\right)$ crossings in G.
- Putting this two together, we conclude $\frac{m^3}{n^2} = O(nk) \Rightarrow m = O(nk^{\frac{1}{3}})$ ■

On the number of incidences
(Another application of the crossing lemma)

- Let P be a set of n distinct points in the plane.
- Let L be a set of m distinct lines in the plane (we do not assume general position here).
- Let $I(P, L)$ denote the number of point/line pairs (p, l) where $p \in P$, $l \in L$ and $p \in l$.
- Let $I(n, m) = max_{|P|=n\,,|L|=m} I(P, L)$
- Clearly, $I(n, m) \leq nm$.
- **<u>Lemma 6</u>**: $I(n, m) = O(n^{\frac{2}{3}} m^{\frac{2}{3}} + n + m)$

- **<u>Lemma 10</u>**: $I(n,m) = O(n^{\frac{2}{3}}m^{\frac{2}{3}} + n + m)$

- **<u>Proof</u>**: Let P and L be the set of n points and m lines realizing $I(n,m)$.

- Note that each line contain at least one point, and $I(n,m) \geq \max(n,m)$.

- Let G be a graph over the points of P. We connect two points if they lie consecutively on a common line of L. What is $n(G)$ and $m(G)$?

- Clearly, $n(G) = n$ and $m(G) = I - m$, where $I = I(n,m)$.

- Note that $c(G) \leq m^2$ (trivial bound)

- On the other hand, by lemma 5, if $m(G) \geq 6n(G)$ (i.e. $I - m \geq 6n$)

  We have $c(G) \geq c_1 \frac{m(G)^3}{n(G)^2}$.

- Thus, $c_1 \frac{(I-m)^3}{n^2} = c_1 \frac{m(G)^3}{n(G)^2} \leq c(G) \leq m^2$

- Thus, $I = c_1^{-\frac{1}{3}} n^{\frac{2}{3}} m^{\frac{2}{3}} + m = O(n^{\frac{2}{3}}m^{\frac{2}{3}} + m)$

- We proved: $m(G) \geq 6n(G) \Rightarrow I(n,m) = O(n^{\frac{2}{3}}m^{\frac{2}{3}} + m)$.
- If $m(G) < 6n(G)$ we get $I - m \leq 6\mathrm{n} \Rightarrow I = O(n + m)$
- Putting the two together, we have that $I(n,m) = O(n^{\frac{2}{3}}m^{\frac{2}{3}} + n + m)$ ■

- **<u>Lemma 11</u>**: $I(n,m) = \Omega(n^{\frac{2}{3}}m^{\frac{2}{3}} + n + m)$
- **<u>Proof</u>:** It's clear that $I(n,m) = \Omega(n+m)$.
- For a positive integer k, let $[\![k]\!] = \{1, \dots, k\}$.
- Denote $N_1 = \frac{1}{2}n^{\frac{2}{3}}m^{-\frac{1}{3}},\ N_2 = n^{-\frac{1}{3}}m^{\frac{2}{3}},\ M = n^{\frac{1}{3}}m^{\frac{1}{3}}$.
- For the sake of simplicity of exposition, assume $N_1,\ N_2,\ N_3$ are integers.
- Let $P = [\![N_1]\!]\text{X}[\![2M]\!]$ and $\mathrm{L} = \{y = ax + b | a \in [\![N_2]\!] \text{ and } b \in [\![M]\!]\}$.
- Note that $|P| = N_1 \cdot 2M = n$ and $|L| = N_2 \cdot M = m$.
- For any $x \in [\![N_1]\!]$, and for any $a \in [\![N_2]\!]$ and $b \in [\![M]\!]$ we have that:

  $y = ax + b \leq N_1N_2 + M = \frac{1}{2}M + M \leq 2M$.
- Namely, any line of L is incident to N_1 points of P.
- Namely, $I(P,L) = N_1|L| = \frac{1}{2}n^{\frac{2}{3}}m^{-\frac{1}{3}} \cdot m = \frac{1}{2}n^{\frac{2}{3}}m^{\frac{2}{3}} \Rightarrow I(n,m) = \Omega(n^{\frac{2}{3}}m^{\frac{2}{3}})$ ■