Incidence Combinatorics

Aaron Anderson

July 5 - 9, 2021

1 Crossing Numbers

Before we get into incidence combinatorics proper, we will develop some tools in graph theory. In particular, we need to understand what it means for a graph to be *planar*, and then we will start to measure how *un*planar a graph is.

Starting from the top, for this class, a *graph* consists of a (finite) vertex set V and an (also finite) edge set E, where each element of E is an unordered pair of distinct elements of V. We typically visualize graphs by drawing the vertices in V as points in the plane, and for each edge (v, w), drawing a curve between v and w. Such a drawing is called *planar* when none of the edges cross each other, and the graph itself is called *planar* when there is some planar way to draw it.

Any planar drawing of a graph breaks the plane into 2d regions known as *faces*. Our main tool for planar graphs is Euler's Formula, which relates the numbers of vertices, edges, and faces. To state it, I need to cover one more definition: A graph is *connected* if you can trace a path from any vertex to any other vertex just by following edges.

Theorem 1 (Euler's Formula)**.** *If X is a finite set, then $|X|$ is the size or cardinality of $|X|$.*

Let G be a connected planar graph with vertex set V and edge set E. Then any planar drawing of G has f faces, where $|V| - |E| + f = 2$.

Problem 1. Prove Euler's formula by induction on the number of faces.

Hint: The connected graphs that can be drawn with $f = 1$ are the *trees*, that is, the connected graphs without cycles. Prove Euler's formula for trees by induction on the number of edges.

In the following, let G be a graph with vertex set V and edge set E.

Problem 2. Show that if G is planar, then $|E| \leq 3|V|$.

Let the *crossing number* $\mathrm{cr}(G)$ of a graph G be the minimum number of pairs of edges that need to cross in order to draw G in the plane.

Problem 3.

1. Show that $\mathrm{cr}(G) = 0$ if and only if G is planar.
2. Show that $\mathrm{cr}(K_5) = \mathrm{cr}(K_{3,3}) = 1$.

Problem 4. Show that $\mathrm{cr}(K_n) \geq \frac{1}{5}\binom{n}{4}$. It is actually known that $\mathrm{cr}(K_n) \leq \frac{3}{8}\binom{n}{4}$, so your lower bound is correct within a factor of 2. (Hint: Look at the contribution from each K_5-shaped subgraph.)

Problem 5. Show that $\mathrm{cr}(G) \geq |E| - 3|V|$.

Now we prove the *Crossing Lemma*:

Lemma 2 (Crossing Lemma)**.** *If* $|E| > 4|V|$*, then* $\operatorname{cr}(G) \geq \frac{|E|^3}{64|V|^2}$.

Problem 6. Let G be a graph where $|E| > 4|V|$. We will use probability to show that $\operatorname{cr}(G) \geq \frac{|E|^3}{64|V|^2}$.

Say you have a loaded coin, which comes up heads with probability p. (We will choose p later, but assume that $0 \leq p \leq 1$.) Now for each vertex $v \in V$, flip the coin, and put v in the set V_H if the coin comes up heads. Now let H be the graph with vertex set V_H, where $v, w \in V_H$ are connected with an edge if and only if they are in G.

1. What is the probability that a given edge e of G is in H?
2. Assume G is drawn in the plane with exactly $\operatorname{cr}(G)$ crossings, and H is drawn the same way, except with some vertices and edges missing. If e_1, e_2 are edges of G that cross, what is the probability that both are in H?
3. Recall that if X is a random variable, $\mathbb{E}[X]$ denotes the *expectation* of X, or the average value it takes. Find the expectation of these three variables:
   - $|V_H|$, the number of vertices of H
   - $|E_H|$, the number of edges of H
   - c_H, The number of crossings in the drawing of H
4. Explain why $\mathbb{E}[c_H] \geq \mathbb{E}[|E_H|] - 3\mathbb{E}[|V_H|]$.
5. Set $p = \frac{4|V|}{|E|}$. Convince yourself that this is a valid probability. Then combine the last two parts of this problem and prove the inequality

$$\operatorname{cr}(G) \geq \frac{|E|^3}{64|V|^2}.$$

Problem 7. Let G be a graph with n vertices each of degree at least 9. Show that the crossing number of G is at least $\frac{4n}{3}$.

2 Asymptotic (Big O) Notation

Before we move on to the "Incidences" in the title of this class, we need to develop some notation. Let's say we have a function, $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$, and we'd like to know *roughly* how fast it grows. (In this class, our function f will almost always be a function $f : \mathbb{N} \to \mathbb{N}$ that counts something.) Our function might be hard to calculate and understand exactly, and might jump around erratically, but we can get the big-picture idea by comparing it to better-behaved functions with *Big O notation*.

Definition 1. Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ and $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be functions. We say that $f(n) = O(g(n))$ (pronounced "f is big O of g") when there exists a positive constant $C > 0$ and a constant $N \in \mathbb{N}$ such that for all $n \geq N$, $f(n) \leq C \cdot g(n)$.

Big O notation allows us to provide a loose *upper bound* for $f(n)$, with $f(n) = O(g(n))$ saying that for some choice of C, the function $C \cdot g(n)$ is an upper bound for $f(n)$ eventually. This notation can save us a lot of work in calculating the exact constants C and N. It's basically a generalization of $f(n) \leq g(n)$, so the use of an $=$ sign in the notation is a bit strange, but it's standard, so here we are.

Problem 8 (Big-O arithmetic, borrowed from Viv)**.** Let f_1, f_2, g_1, and g_2 be functions from $\mathbb{N} \to \mathbb{R}_{\geq 0}$.

1. Assume that $f_1(n) = O(g_1(n))$ and $f_2(n) = O(g_2(n))$. Show that $f_1 \cdot f_2(n) = O(g_1 \cdot g_2(n))$, where $f_1 \cdot f_2(n) = f_1(n) \cdot f_2(n)$. In particular, show that $f_1 \cdot f_2(n) = O(f_1 \cdot g_2(n))$.
2. Show that $(f_1 + f_2)(n) = O(\max\{f_1, f_2\}(n))$ and that $\max\{f_1, f_2\}(n) = O((f_1 + f_2)(n))$.
3. Assume that $f_1(n) = O(g_1(n))$ and $f_2(n) = O(g_2(n))$. Show that $(f_1+f_2)(n) = O(\max\{g_1, g_2\}(n))$.
4. Assume that $f_1(n) = O(g_1(n))$ and $C \in \mathbb{R}$ is any constant. Show that $C * f_1(n) = O(g_1(n))$.
5. Assume that $\lim_{n\to\infty} f_1(n) = \infty$. Show that $f_1(n) + 1 = O(f_1(n))$.

Problem 9. Let $r \in \mathbb{R}$. Show that for all $\varepsilon > 0$, $n^r \log n = O(n^{r+\varepsilon})$.

Definition 2. Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ and $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be functions. We say that $f(n) = \Omega(g(n))$ when there exists a positive constant $C > 0$ and a constant $N \in \mathbb{N}$ such that for all $n \geq N$, $f(n) \geq C \cdot g(n)$.

Problem 10. Show that $f(n) = O(g(n))$ if and only if $g(n) = \Omega(f(n))$.

Definition 3. Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ and $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be functions. We say that $f(n) = \Omega(g(n))$ when there exists a positive constant $C > 0$ and a constant $N \in \mathbb{N}$ such that for all $n \geq N$, $f(n) \geq C \cdot g(n)$.

The idea here is $f(n) = \Omega(g(n))$ when g is an approximate *lower* bound for f.

Problem 11. Show that $f(n) = O(g(n))$ if and only if $g(n) = \Omega(f(n))$.

Definition 4. Let $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ and $g : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be functions. We say that $f(n) = \Theta(g(n))$ when both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

Problem 12. Show that for all k, $\binom{n}{k} = \Theta(n^k)$. (Here, we're thinking of $\binom{n}{k}$ as a function $f(n)$ for a fixed k.)

Problem 13. Show that $f(n) = \Theta(g(n))$ is an *equivalence relation*, that is, the following three properties hold:

- **Reflexivity:** For all $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$, $f(n) = \Theta(f(n))$.
- **Symmetry:** For all $f, g : \mathbb{N} \to \mathbb{R}_{\geq 0}$, $f(n) = \Theta(g(n))$ if and only if $g(n) = \Theta(g(n))$.
- **Transitivity:** For all $f, g, h : \mathbb{N} \to \mathbb{R}_{\geq 0}$, $f(n) = \Theta(g(n))$ and $g(n) = \Theta(h(n))$ implies $f(n) = \Theta(h(n))$.

Definition 5. Let $f, g : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ be functions on two variables. Then we can still define O, Ω, and Θ pretty much as before:

- Let $f(m, n) = O(g(m, n))$ when there exist $C > 0$ and $N \in \mathbb{N}$ such that for all $m, n \geq N$, $f(m, n) \leq C \cdot g(m, n)$.
- Let $f(m, n) = \Omega(g(m, n))$ when there exist $C > 0$ and $N \in \mathbb{N}$ such that for all $m, n \geq N$, $f(m, n) \geq C \cdot g(m, n)$.
- Let $f(m, n) = \Theta(g(m, n))$ when $f(m, n) = O(g(m, n))$ and $f(m, n) = \Omega(g(m, n))$.

Problem 14. Pick your two favorite parts of Problem 8 and check that they still work when f_1, f_2, g_1, g_2 are functions on two variables.

3 Review

It will be useful to remember this lemma, and this definition, from yesterday:

Lemma 3 (Crossing Lemma). *If $|E| > 4|V|$, then* $\mathrm{cr}(G) \geq \frac{|E|^3}{64|V|^2}$.

Definition 6. Let $f, g : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ be functions on two variables.

- Let $f(m,n) = O(g(m,n))$ when there exist $C > 0$ and $N \in \mathbb{N}$ such that for all $m, n \geq N$, $f(m,n) \leq C \cdot g(m,n)$.
- Let $f(m,n) = \Omega(g(m,n))$ when there exist $C > 0$ and $N \in \mathbb{N}$ such that for all $m, n \geq N$, $f(m,n) \geq C \cdot g(m,n)$.
- Let $f(m,n) = \Theta(g(m,n))$ when $f(m,n) = O(g(m,n))$ and $f(m,n) = \Omega(g(m,n))$.

4 Counting Incidences with the Crossing Lemma

Let P be a set of n points in $\mathbb{R}^2$, and let L be a set of m lines. An *incidence* of P on L is defined to be an ordered pair (p, ℓ) where $p \in P, \ell \in L$, and the point p lies on the line ℓ. We will care about the number of incidences of P on L, which we denote by $I(P, L)$. The critical theorem that we will use, our hammer that makes every problem in discrete geometry look like a nail, is the *Szemerédi-Trotter* Theorem:

Theorem 4. *If P is a set of n points and L a set of m lines in $\mathbb{R}^2$, then* $I(P, L) = O\left(m^{2/3}n^{2/3} + m + n\right)$.

We will prove this in Problem 16, but first let's find some examples to explain the terms in this bound.

Problem 15. This problem will justify the m and n terms in $O\left(m^{2/3}n^{2/3} + m + n\right)$:

- For every $m, n > 0$, find a set L of m lines and a set P of n points in $\mathbb{R}^2$ such that $I(P, L) \geq m$.
- Similarly, for every $m, n > 0$, find a set L of m lines and a set P of n points in $\mathbb{R}^2$ such that $I(P, L) \geq n$.

We now prove Theorem 4:

Problem 16. Let P be a finite set of points and L a finite set of lines in the plane.

- Assume for the moment that there are at least 2 points from P on every line in L. Now use P and L to construct a graph G. Find upper and lower bounds for the crossing number of G, and use this to show that $I(P, L) = O(|P|^{2/3}|L|^{2/3} + |P|)$. (Hint: The $|E| < 4|V|$ assumption of the crossing lemma will require you to do a bit of casework.)
- Prove that $I(P, L) = O(|P|^{2/3}|L|^{2/3} + |P| + |L|)$.

Problem 17. This problem will justify the $m^{2/3}n^{2/3}$ term in Theorem 4:

Let $h, w \in \mathbb{N}$, and let P be the $h \times w$-grid $\{1, 2, \ldots, w\} \times \{1, 2, \ldots, h\}$, consisting of h rows of w points in the plane. Let L be the set of all lines of negative slope that pass through exactly one point of each row, as depicted in the following figure.

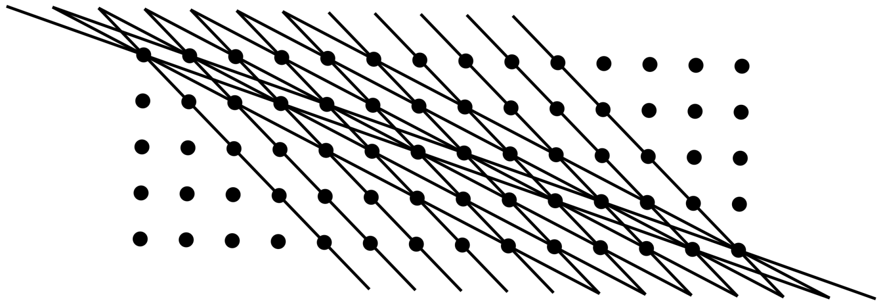

Figure 1: Picture by Adam Sheffer

- Show that there are approximately $\frac{w}{h}$ slopes of lines in L.
- Given a slope s, (approximately) how many lines of slope s are there in L?
- Show that when h, w are both large, $|L| = \Theta(\frac{h^2}{w})$.
- Show that $I(P, L) = \Theta(|P|^{2/3}|L|^{2/3})$

Hint: This problem, if done precisely, has several off-by-one issues, that is, lots of terms that look like $h-1$ or $w-1$, where you'd really rather it was just h or w. As we're only going for approximate and asymptotic answers here, that difference won't really matter, as we care about the case when h and w are really large, and the -1 is tiny in comparison. Thus I'd just ignore those -1s at first, and then afterwards, if you have time, go back and see how to rigorously take care of them by changing the constants of your Θs.

Problem 18. If P is a finite set of points in the plane, let $T(P)$ be the number of right triangles whose vertices are all in P, and let $U(P)$ be the number of triangles whose vertices are all in P that have area 1 (the U stands for *unit* area). Use Theorem 4 to prove that $U(P) = O(|P|^{7/3})$.

4.1 Rich Lines

Let P be a set of n points in the plane, and for $r \geq 2$, let m_r be the number of lines through at least r points in P. These are called the r-rich lines.

Problem 19. Prove that there is a positive constant C such that for all $r \geq 2$, $m_r \leq C\left(\frac{n^2}{r^3} + \frac{n}{r}\right)$.

Note: This is a bit stronger than saying that $m_r = O\left(\frac{n^2}{r^3} + \frac{n}{r}\right)$, but if you can prove the Big-O version, you're well on your way.

Problem 20. Show that if $A \subset \mathbb{R}$ has $|A| = n$, then there are $O(n^4 \log(n))$ collinear triples in $A \times A$.

4.2 Higher-Degree Curves

Problem 21. Let P be a set of n points in the plane, and Γ a set of n hyperbolas, each defined by an equation of the form $(x-a)^2 - (y-b)^2 = 1$. Show that $I(P, \Gamma) = O(n^{4/3})$.

Problem 22. Construct a set P of m points and a set Γ of n parabolas, each defined by an equation of the form $y = ax^2 + bx + c$, such that $I(P, \Gamma) = \Omega(m^{1/2}n^{5/6})$. This shows that Szemerédi-Trotter doesn't apply to all conic sections.

5 Distinct and Unit Distances Problems

In this section, we will investigate two questions posed by Erdős in 1946. They have stumped mathematicians for decades, but we can make some progress today!

Let P be a finite set of points in the plane.

5.1 Unit Distances

We define the *unit distance pairs* in P to be pairs $\{p, q\} \subseteq P$ such that $|pq| = 1$. We define $U(P)$ to be the number of unit pairs in P, and for a natural number n, define $U(n)$ to be the largest value of $U(P)$ for some set P with $|P| = n$. In the same 1946 paper, Erdős asked for upper and lower bounds on $U(n)$. Erdős's original upper bound has only been improved once, in 1984, and we will prove that best-known result today!

Problem 23. Find a sequence of sets $P_m \subset \mathbb{R}^2$ for all natural numbers m such that $|P_n| = 2^m$, and the number of unit distances in P_m is $m2^m$. Conclude that $U(n) = \Omega(n \log n)$. (This is close to the best-known lower bound.)

Problem 24. Let C be a finite set of unit-radius circles. Let $I(P, C)$ be the number of incidences of points in P on circles in C. Taking inspiration from our proof of Szemerédi-Trotter, construct a graph with vertex set P, and use it to prove that $I(P, C) = O\left(|P|^{2/3}|C|^{2/3} + |P| + |C|\right)$.

Problem 25. Construct a set C of unit-radius circles such that $I(P, C)/2$ is the number of unit distance pairs in P, and use it to put an upper bound on the number of unit distance pairs.

Problem 26. The Unit Distances Problem can be posed in more dimensions also. For any $k \in \mathbb{N}$ and any finite set P in $\mathbb{R}^k$, we can define $U_k(P)$ to be the number of unit pairs in P, and define $U_k(n)$ to be the largest value of $U(P)$ over all sets $P \subseteq \mathbb{R}^k$ with $|P| = n$, so that $U_2(n) = U(n)$. The 3-dimensional version is also wide open, but we can actually solve it for $k \geq 4$ (up to a constant factor, as usual). Prove that for $k \geq 4$, $U_k(n) = \Theta(n^2)$.

5.2 Distinct Distances

We define the number of *distinct distances* in P to be the number of real numbers r such that there are points $p, q \in P$ such that $|pq| = r$. Erdős's *Distinct Distances Problem* asks to find upper and lower bounds on the function $d : \mathbb{N} \to \mathbb{N}$, where $d(n)$ is defined to be the minimum number of distinct distances in a set P with $|P| = n$. This problem was a famous open problem for decades, and was not fully solved until 2015 (well, actually, it's only been aaaaalmost solved, but people are pretty happy with the 2015 paper).

Problem 27. For each point $p \in P$, let C_p be the set of circles centered at P that pass through other points in P. Let $p, q \in P$ be distinct points. Show that $2|C_p||C_q| \geq |P| - 2$. Conclude that $d(n) = \Omega(n^{1/2})$.

Problem 28. The Distinct Distances Problem has been posed in more dimensions also. If k is a positive integer, let $d_k(n)$ be the minimum number of distinct distances in a set of n points in $\mathbb{R}^k$. Show that for each k, $d_k(n) = \Omega(n^{1/k})$ and $d_k(n) = O(n^{2/k})$.

Hint: To show that $d_k(n) = O(n^{2/k})$, you want to actually provide concrete examples.

Problem 29. We now have an upper bound on the number of distinct distances, but we can improve that upper bound in the two-dimensional case. Use the following Theorem to improve your upper bound, by measuring the number of distinct distances in an $n \times n$ grid of points in the plane.

Theorem 5 (Landau-Ramanujan). *The number of integers in* $[1, m]$ *that are the sum of two perfect squares is* $\theta(m/\sqrt{\log m})$.

We can also use incidences to improve our lower bound:

Problem 30. Using our version of Szemerédi-Trotter for unit circles, show that $d(n) = \Omega(n^{2/3})$.

5.3 Alternate Metrics

Both of these problems revolve around a notion of "distance" of points in the plane. What happens if we change that too?

One alternate notion of distance is called the L_1 or *Manhattan* metric. In this context, we redefine the distance between two points of the plane, (x_1, y_1) and (x_2, y_2), to be $|x_1 - y_1| + |x_2 - y_2|$. To get an intuition for this, imagine you're in a city with a perfectly North-South/East-West grid (more or less like Manhattan). If you want to get from one point to another, the fastest way to get there is by walking due North/South until you're on the same street as your destination, and then walking East/West until you get there. This means that the total distance you cover is the sum of the horizontal distance between the two locations and the vertical distance - $|x_1 - y_1| + |x_2 - y_2|$.

Problem 31. Draw the set of all points in the plane that have L_1-distance 1 from the origin. This is the L_1 version of the unit circle.

Problem 32. Given a finite set $P \subset \mathbb{R}^2$, define $U_{L_1}(P)$ to be the number of unit-distance pairs of points in P, where by "unit-distance", I now mean that the L_1-distance between the two points is 1. For $n \in \mathbb{N}$, define $U_{L_1}(n)$ as before to be the maximum of $U_{L_1}(P)$ over all P with $|P| = n$. Up to a constant factor, solve for $U_{L_1}(n)$.

Problem 33. Given a finite set $P \subset \mathbb{R}^2$, define $d_{L_1}(P)$ to be the number of distinct L_1-distances between pairs of points in P. For $n \in \mathbb{N}$, define $d_{L_1}(n)$ as before to be the minimum of $d_{L_1}(P)$ over all P with $|P| = n$. Up to a constant factor, solve for $d_{L_1}(n)$.

6 Alternate Metrics

Both of these problems revolve around a notion of "distance" of points in the plane. What happens if we change that too?

One alternate notion of distance is called the L_1 or *Manhattan* metric. In this context, we redefine the distance between two points of the plane, (x_1, y_1) and (x_2, y_2), to be $|x_1 - y_1| + |x_2 - y_2|$. To get an intuition for this, imagine you're in a city with a perfectly North-South/East-West grid (more or less like Manhattan). If you want to get from one point to another, the fastest way to get there is by walking due North/South until you're on the same street as your destination, and then walking East/West until you get there. This means that the total distance you cover is the sum of the horizontal distance between the two locations and the vertical distance - $|x_1 - y_1| + |x_2 - y_2|$.

Problem 34. Draw the set of all points in the plane that have L_1-distance 1 from the origin. This is the L_1 version of the unit circle.

Problem 35. Given a finite set $P \subset \mathbb{R}^2$, define $U_{L_1}(P)$ to be the number of unit-distance pairs of points in P, where by "unit-distance", I now mean that the L_1-distance between the two points is 1. For $n \in \mathbb{N}$, define $U_{L_1}(n)$ as before to be the maximum of $U_{L_1}(P)$ over all P with $|P| = n$. Up to a constant factor, solve for $U_{L_1}(n)$.

Problem 36. Given a finite set $P \subset \mathbb{R}^2$, define $d_{L_1}(P)$ to be the number of distinct L_1-distances between pairs of points in P. For $n \in \mathbb{N}$, define $d_{L_1}(n)$ as before to be the minimum of $d_{L_1}(P)$ over all P with $|P| = n$. Up to a constant factor, solve for $d_{L_1}(n)$.

7 Additive Combinatorics

Let $A, B \subset \mathbb{R}$ be finite sets. Then we define $A+B$ to be $\{a+b : a \in A, b \in B\}$ and $AB = \{ab : a \in A, b \in B\}$.

Problem 37. If $A \subset \mathbb{R}$ has size n, what are the minimum and maximum possible values of $|A+A|$ and $|AA|$? Given some n, can you find a set A of size n where $|A+A|$ and $|AA|$ are both maximized?

Problem 38. If $A \subset \mathbb{R}$ is finite, show that $\max(|A+A|, |AA|) = \Omega(|A|^{5/4})$.

Problem 39. Prove that $|A + AA| = \Omega(|A|^{3/2})$.

Problem 40. Prove that $|A+A+A||AAA| = \Omega(|A|^{3/2}|A+A|^{1/2}|AA|^{1/2})$.

8 Cuttings

To move further in incidence combinatorics, we need to develop new tools. Probably the most successful technique in the area is *partitioning*. If you want to understand the incidences of a set of points on a set of curves, it is often helpful to carve up the plane (or a higher-dimensional space) into carefully-chosen pieces, bound the number of incidences in each piece separately, and then add it all back up. In particular, one of the key ideas behind the proof of Distinct Distances is a partitioning theorem, that leads to a whole "Polynomial Method". As for this class, we're going to look at how to prove Szemerédi-Trotter using this partitioning philosophy.

First we'll look at general graph theory result which gives us some quick incidence bounds. For the statement of this theorem, we'll need a bit of nomenclature about bipartite graphs. A *bipartite graph* is a graph whose vertex set can be split into two disjoint subsets V_1 and V_2 such that each edge connects a vertex in V_1 to a vertex in V_2. For $s, t \in \mathbb{N}$, we define the *complete bipartite graph* $K_{s,t}$ to be a bipartite graph where $|V_1| = s$ and $|V_2| = t$, where every vertex in V_1 is connected to every vertex in V_2.

Problem 41. How many edges does $K_{s,t}$ have?

If G is a bipartite graph with its vertex set partitioned into V_1 and V_2, then a *copy of* $K_{s,t}$ consists of a set $V_1' \subseteq V_1$ of s vertices in V_1 and a set $V_2' \subseteq V_2$ of t vertices in V_2, such that every vertex of V_1' is connected to every vertex of V_2'.

Theorem 8.1 (Kövári-Sós-Turán)**.** *Let G be a bipartite graph with vertex set partitions V_1 and V_2, with $|V_1| = m$ and $|V_2| = n$, and edge set E. Assume G contains no copy of $K_{s,t}$ with the s vertices in V_1 and the t vertices in V_2. Then $|E| = O(mn^{1-1/s} + n)$.*

Problem 42. Let P be a finite set of points in the plane, and L a finite set of lines. Using Theorem 8.1, show that $I(P, L) = O(|P||L|^{1/2} + |L|)$ and $I(P, L) = O(|L||P|^{1/2} + |P|)$.

Problem 43. Let P be a finite set of points in the plane, and C a finite set of not-necessarily-unit circles. What bounds can you find on $I(P, C)$? What if C is instead a finite set of curves, each of which is the graph of a polynomial f of degree at most k?

Now let's introduce a partitioning tool that will let us divide up the plane and then conquer with Theorem 8.1. Later on, we will prove a slightly weaker version of it.

Theorem 8.2 (Cutting Lemma for Lines)**.** *Let L be a set of n lines in the plane, and let r be a parameter, $1 < r < n$. Then the plane can be subdivided into t generalized triangles (this means intersections of three half-planes) $\Delta_1, \Delta_2, \ldots, \Delta_t$ in such a way that the interior of each Δ_i is intersected by at most n/r lines of L, and we have $t = O(r^2)$.*

This is a cutting lemma designed specifically for a set of n lines in the plane, so we will use it for another proof of Szemerédi-Trotter.

Problem 44. Let P be a set of n points and L a set of n lines in the plane. Let's prove that $I(P, L) = O(n^{2/3})$ using Theorem 8.2 and Kövári-Sós-Turán.

Specifically, for some yet-to-be-determined value of r, apply Theorem 8.2 to L, and use Kövári-Sós-Turán to bound the incidences in each triangle Δ_i separately (you'll also want to consider the vertices of these triangles separately, because such a vertex could belong to arbitrarily many triangles). Now add up what you get, and find a value of r that will give you the desired bound.

(If you want to get the same $I(P, L) = O(m^{2/3}n^{2/3} + m + n)$, then you will have to vary r depending on the exact ratio of m to n, but I think you get the idea.)

There's also a cutting lemma for circles:

Theorem 8.3 (Cutting Lemma for Circles). *Let C be a set of n circles in the plane, and let r be a parameter, $1 < r < n$. Then the plane can be subdivided into t sets $\Delta_1, \Delta_2, \ldots, \Delta_t$ in such a way that the interior of each Δ_i is intersected by at most n/r circles of C, and we have $t = O(r^2 \log^2(n))$. Each Δ_i can be selected to be a "circular trapezoid" - that is, its boundary consists of at most two vertical line segments and at most two circular arcs - and the interiors of the Δ_is are disjoint.*

You can get rid of the pesky $\log^2(n)$ factor, but it makes this easier to prove, and then it is no longer guaranteed that the Δ_is are nonoverlapping. It's not actually horrible if they're overlapping, but it makes the problem that follows a little too hard for this class.

Problem 45. Using the cutting lemma for circles together with your previous incidence bounds to show that if P is a set of n points in the plane and C is a set of n circles, then $I(P, C) = O(n^{1.4} \log^c(n))$ for some constant c.

Actually, more is true. We showed that if P is a set of points and C a set of circles with $|P| = |C| = n$, we have $I(P, C) = O(n^{1.4} \log^c(n))$. If we had more time, we'd be able to show that in general, even if $|P|$ and $|C|$ have different sizes, $I(P, C) = O(|P|^{3/5}|C|^{4/5} + |P| + |C|)$. When $|P| = |C| = n$, this reduces to the bound we proved, just without that pesky log factor.

Problem 46. Improve the lower bound on distinct distances: $d(n) = \Omega(n^{3/4})$.

Problem 47. Now let's try to prove a slightly weaker version of Theorem 8.2. The only difference is the very last line, where we have $t = O(r^2 \log^2(n))$ instead of $t = O(r^2)$.

Let L be a set of n lines in the plane, and let r be a parameter, $1 < r < n$. Then we will show that the plane can be subdivided into t generalized triangles (this means intersections of three half-planes) $\Delta_1, \Delta_2, \ldots, \Delta_t$ in such a way that the interior of each Δ_i is intersected by at most n/r lines of L, and we have $t = O(r^2 \log^2(n))$.

To show this, we carefully choose a number s, and then randomly select (with replacement) s lines from L, which we collect in a set $S \subseteq L$. Then we use the at most s lines in S, and use them to cut the plane up into polygonal cells. If any of these cells has too many sides, we cut it along diagonals until we are left with only triangles. Now find a value of s such that the number of triangles in this decomposition is is $O(r^2 \log^2(n))$ while the probability that the interior of each triangle is intersected by at most n/r lines of L is positive.

Problem 48. Now let's prove Theorem 8.1. To do this, I'd recommend bounding above and below the number of copies of $K_{s,1}$ in the graph G. Then you'll also want to know *Jensen's Inequality*, which says that if $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then for any $x_1, \ldots, x_k$, the average value of f at those points is at least the value of f at the average, that is: $\frac{1}{k}\sum_{i=1}^k f(x_i) \geq f\left(\frac{1}{k}\sum_{i=1}^k x_i\right)$.

Problem 49. Prove the cutting lemma for circles.

Hint: As with the cutting lemma for lines, choose $S \subseteq C$ with s independent random draws. Then partition the plane with these circles, and from every intersection point, and from the leftmost and rightmost points on each circle, draw a vertical line segment up to the next circle (or all the way to infinity) and down to the next circle (or infinity).

Show that for the correct choice of s, these circles and line segments partition the plane into the $\Delta_1, \ldots, \Delta_t$ that we want with positive probability.

9 References

Most of this material comes from either Adam Sheffer's website or the notes I took in his classes. Check out `https://adamsheffer.wordpress.com/pdf-files/` if you want to learn more about this area!

The material in the last section, however, draws mostly from Andrew Suk's notes at `www.math.ucsd.edu/~asuk/Lecture4.pdf`