CROSSING NUMBERS AND DISTINCT DISTANCES

The Szemerédi-Trotter theorem plays a fundamental role in incidence geometry in the plane. It can be rephrased in several equivalent ways, and it helps to know the different ways. We recall three standard phrasings here.

If $\mathfrak{S}$ is a set of points and $\mathfrak{L}$ is a set of lines (or curves), recall that the set of incidences $I(\mathfrak{S}, \mathfrak{L}) = \{(p,l) \in \mathfrak{S} \times \mathfrak{L} | p \in l\}$. We now give three versions of Szemerédi-Trotter.

Version 1. *If $\mathfrak{S}$ is a set of S points in the plane, and $\mathfrak{L}$ is a set of L lines in the plane, then the number of incidences is bounded as follows:*

$$|I(\mathfrak{S}, \mathfrak{L})| \le C(S^{2/3}L^{2/3} + S + L).$$

Version 2. *Suppose that $\mathfrak{L}$ is a set of L lines in the plane, and let P_k be the set of points that lie on $\ge k$ lines of $\mathfrak{L}$.*

Then $|P_k| \le C(L^2 k^{-3} + L k^{-1})$.

Version 3. *Suppose that $\mathfrak{S}$ is a set of S points in the plane, and let $\mathfrak{L}_r$ be the set of lines that contain $\ge r$ points of $\mathfrak{S}$.*

Then $|\mathfrak{L}_r| \le C(S^2 r^{-3} + S r^{-1})$.

In an earlier lecture, we proved Version 2 using the crossing number theorem. With tiny modifications, the same argument proves any version of the theorem. Also, any version above implies any other version by a short counting argument.

1. Distinct distances

Theorem 1.1. *(Székely) If we have N distinct points in the plane, then they determine $\ge cN^{4/5}$ distinct distances. In fact, there is one point p in the set so that the distance from p takes $\ge cN^{4/5}$ distinct values.*

Proof. Suppose that for each point p in our set $\mathfrak{S}$, the set of distances $\{dist(p,q)\}_{q \in \mathfrak{S}}$ takes on $\le t$ different values. We assume $t \le cN^{4/5}$ and we will get a contradiction. We can choose Nt circles so that each point of the set lies in $N-1$ circles. We draw all these circles, leaving out circles with ≤ 4 points. We make a multigraph G whose vertices are the points of $\mathfrak{S}$ and whose edges are arcs between consecutive points on one of the circles.

This multigraph has $V = N$ vertices. It has $E \ge (1/2)N^2$ edges. (Before removing unpopular circles, it would have $N^2 - N$ edges. We removed edges that were on the unpopular circles, but these circles contribute a total of $\leq 4Nt \leq (1/200)N^2$ edges.) It has crossing number $\leq 2(Nt)^2$, because a pair circles intersects in ≤ 2 points.

The multigraph G may have very high multiplicity. Our strategy will be to estimate how many high-multiplicity edges G can have, and trim edges from G to reduce the multiplicity.

Lemma 1.2. *The number of edges of G with multiplicity $\geq M$ is at most $C[N^2 M^{-2} t + N \log Nt]$.*

Proof. Consider edges from a vertex p_1 to a vertex p_2. Each edge is the arc of a circle, and the center of the circle must lie on the perpendicular bisector of p_1 and p_2. If there are many edges from p_1 to p_2, then there must be many points of our set along the perpendicular bisector.

We define a map from edges of our multigraph to lines, sending an edge to the corresponding perpendicular bisector. A line containing A points of $\mathfrak{G}$ contributes $\leq 2At$ edges of the multigraph, each with multiplicity $\leq A$.

Let $\mathfrak{L}_j$ denote the set of lines in the plane which contain $\sim 2^j$ points of $\mathfrak{G}$. (More precisely, the number of points is greater than 2^{j-1} and at least 2^j.) The number of edges with multiplicity at least M is bounded by

$$\sum_{2^j \geq M} |\mathfrak{L}_j| 2 \cdot 2^j t.$$

The size of $\mathfrak{L}_j$ is bounded by the Szemerédi-Trotter theorem (see Version 3 above). Plugging in, we get:

$$\leq \sum_{2^j \geq M} C(N^2 2^{-3j} + N 2^{-j}) 2^j t.$$

The $N^2 2^{-3j}$ term decays exponentially in j, and the total is $\leq CN^2 M^{-2} t$. The second term is independent of j, and we need to sum over $\sim \log N$ values of j, so the total is $\leq CN \log Nt$. □

We choose $M = \alpha t^{1/2}$ for a large constant α. By choosing α large enough, we can arrange that the number of edges of multiplicity $\geq M$ is at most $(1/10)N^2$. We let $G' \subset G$ be the multigraph given by deleting all edges of G with multiplicity $\geq M$. The graph G' still has $\geq (1/3)N^2$ edges, and it now has multiplicity at most $M \lesssim t^{1/2}$.

Now we apply the crossing number theorem for multigraphs. We recall the statement from last lecture.

Theorem 1.3. *(Crossing number estimate for multigraphs) If G is a multigraph with V vertices and E edges and with multiplicity $\leq M$, and if $E \geq 100MV$, then the crossing number of G is at least $cE^3 V^{-2} M^{-1}$.*

Our graph G' has crossing number at most $\sim N^2 t^2$. But by the theorem above, we have

$$N^2 t^2 \gtrsim k(G') \gtrsim E^3 V^{-2} M^{-1} \sim N^6 N^{-2} t^{-1/2}.$$

Rearranging gives $t^{5/2} \gtrsim N^2$ and so $t \gtrsim N^{4/5}$ as desired.

□

Building on the crossing number approach introduced by Székely, Solymosi-Toth and then Katz-Tardos improved the estimates in the distinct distance problem. Katz-Tardos proved that for any N points in the plane, one of the points determines $\geq cN^{.864}$ distances with the other points. This approach gave the best estimate in the distinct distance problem before the polynomial method approach.

Using the polynomial method we will prove that the number of distinct distances given by N points is $\geq cN(\log N)^{-1}$. However, this approach does not bound the number of distances from a single point. It looks completely plausible that for any N points in the plane, one of the points determines $\geq cN(\log N)^{-1}$ (or even $\geq cN(\log N)^{-1/2}$)) distances with the other points. This would be a better theorem if it's true.

2. What about three dimensions?

In the last few sections, we have had a brief but substantial introduction to incidence geometry in two dimensions. What happens in three dimensions? We will brainstorm some questions below. Three dimensions are more complicated, there are more questions, and it's less clear which are the fundamental questions.

Question 1. *Given S points and L lines in $\mathbb{R}^3$, what is the maximum possible number of incidences?*

It turns out that this question is equivalent to the corresponding question in $\mathbb{R}^2$. Since $\mathbb{R}^2 \subset \mathbb{R}^3$, the maximum must be at least as big in two dimensions. But given an arrangement of points and lines in $\mathbb{R}^3$, we can project to a generic plane. The projection gives S distinct points and L distinct lines in the plane, and it has at least as many incidences. To try to make the question more interesting, we may bound the number of points or lines in a plane. For example, we may consider the following question.

Question 2. *Given S points and L lines in $\mathbb{R}^3$, with $\leq B$ lines in any plane, what is the maximum possible number of incidences?*

Besides points and lines, $\mathbb{R}^3$ also contains planes. We can try to make similar incidence questions also using planes.

Question 3. *Given S points and P planes in $\mathbb{R}^3$, what is the maximum possible number of incidences?*

This question has a simple answer. Take all P planes containing a line l, and all S points in l. Then the number of incidences is SP, which is the maximum possible. To try to make the question more interesting, we may rule out this example by bounding the number of planes containing any line.

Question 4. *Given S points and P planes in $\mathbb{R}^3$, with the restriction that any line lies in $\leq B$ of the planes, what is the maximum possible number of incidences?*

(I don't know the answer to this question... it may be open.)

We can combine lines and planes.

Question 5. *Given L lines and P planes in $\mathbb{R}^3$, what is the maximum possible number of pairs (l, π) where the line l is contained in the plane π?*

By duality, this question is equivalent to question 1, and so it is answered by the Szemerédi-Trotter theorem (up to a constant factor).

We may then combine points, lines, and planes:

Question 6. *Given S points, L lines, and P planes in $\mathbb{R}^3$, what is the maximum possible number of triples (p, l, π) with the point p in the line l in the plane π?*

(I don't know the answer to this question... it may be open.)

We will come back to question 2 later on, using the polynomial method. The Szemerédi-Trotter theorem plays a central and fundamental role in two dimensions. There may not be any one result in three dimensions which is so central. And there are definitely many more questions besides the ones listed here.