9. Reflections on the Szemeredi-Trotter theorem

Thur March 6

There is an important analogy between the Szemeredi-Trotter theorem and the exceptional set problem in projection theory. The Szemeredi-Trotter theorem can be viewed as the sharp projection theorem for finite sets of points in $\mathbb{R}^2$. The exceptional set problem concerns the projection theory of a finite set of balls in $\mathbb{R}^2$ subject to a natural spacing condition. The sharp answers to both problems are essentially the same – based on integer grids. This analogy was noticed by Tom Wolff in the late 1990s. He adapted proof methods from combinatorial geometry to problems in geometric measure theory and harmonic analysis, with striking results. He tried hard to adapt the proof of Szemeredi-Trotter to the exceptional set problem and the Furstenberg set conjecture, but he was not able to prove sharp results.

The proof of Szemeredi-Trotter using topological methods is elegant and important, but there are several important questions that it does not address. In this class we will discuss them.

First let's recall the statement of Szemerédi-Trotter theorem. Let X be a set of points, L be a set of lines (both in $\mathbb{R}^2$), we use $I(X, L)$ to denote the set of incidences between them:

$$I(X, L) = \{(p, l) \in X \times L : p \in L\}.$$

Szemerédi-Trotter claims that

$$|I(X, L)| \lesssim |X| + |L| + |X|^{2/3}|L|^{2/3}.$$

All the current proofs of this theorem, like the cell decomposition method we discussed in the previous lectures, used the topology of Euclidean plane. This is not surprising, as the conclusion of this theorem is indeed related to the structure of the base field. If we replace $\mathbb{R}^2$ by $\mathbb{F}_p^2$, Szemerédi-Trotter bound will fail as one can see by taking L to be all the lines in $\mathbb{F}_p^2$.

On the other hand, the current methods provide little information on some closely related problems, such as:

1. Projection theory over finite fields.
2. Structure of sharp examples for Szemerédi-Trotter.
3. Projection theory of unit balls, instead of points, in $\mathbb{R}^2$ with spacing conditions (lots of attempts by Wolff).

Structure of Sharp Examples. Let's stare at the Szemerédi-Trotter bound:

$$|I(X, L)| \lesssim |X| + |L| + |X|^{2/3}|L|^{2/3}.$$

There are three terms on the RHS. The first two terms are given by double counting which generalizes to other fields. They dominate when there are too many points or lines, in which case the structure of the sharp examples example is not very rigid, giving us many degrees of freedom. To be more specific, when the first dominates we have $|X| \gg |L|^2$, which means that the number of points has already exceeded the total number of intersections among the lines. In this case, the upper bound is tight if each point has a line passing through it. The typical sharp example looks like some chains of beads. Similarly, when the second term dominates the sharp examples look like a bunch of stars, where each line doesn't have much chance to pass through too many points.

The case where the third term dominates is the most interesting one. The known sharp examples are integer grids and their variation R-grids, where R is the integer rings of number fields. We expect that the sharp examples in this case are highly structured. To see what information about the sharp examples the proof of Szemerédi-Trotter theorem tells us, let's briefly review the cell decomposition proof: Divide $\mathbb{R}^2$ into s^2 cells. In each cell there are $|X|/s^2$ points and (in average) $|L|/s$ lines. By choosing s to be large enough we will have $|L|/s \gtrsim (|X|/s^2)^2$ and then apply the double counting bound. The proof doesn't tell us much information on the structure of the sharp example unless we can figure out the way our cells interact. Unfortunately the proof of cell decomposition is not very constructive and based on existence theorems from topology.

Remark 9.1. *In the projective plane $\mathbb{PR}^2$ there is something called point-line duality. It preserves the incidence relationship between points and lines. In fact, the statement that a point with coordinates $[a_0, a_1, a_2]$ lies on a line with coefficients $[b_0, b_1, b_2]$ simply means $a_0b_0 + a_1b_1 + a_2b_2 = 0$, where the roles of a_i and b_i are interchangeable. The chain example and the star examples are mapped to each other via the point-line duality, while the grid example will be mapped to something different.*

There are also some interesting variations of this problem. For example, one may ask about the structure of X which maximizes the projections for some particular D. Define $S_D(N) = \min_{|X|=N} S(X, D)$. For an arbitrary D, what can we tell about the structure of X achieving this maximum? All the known examples are for direction sets with special structures.

Denote the directions in $\mathbb{R}^2$ by elements of $\mathbb{R}\cup\{\infty\}$ with corresponding projections $\pi_t(x) = x_1 + tx_2$ for $t \in \mathbb{R}$, $\pi_\infty(x) = x_2$. Since we can use a projective transformation to map any three directions to any three specified directions without changing the incidence structure, let's begin with $|D| = 4$. Without of loss of generality we may assume $D = \{0, 1, t, \infty\}$. When t is rational with small denominator the grid example still works. Things become more interesting when t is transcendental. For example, we may take $P_{k,s} = \{a_0 + a_1t + \cdots + a_{k-1}t^{k-1} : a_i \in \mathbb{Z}, 0 \le a_i \le s-1\}$ be a set of polynomials in t and let $X = P_{k,s} \times P_{k,s}$. Then $|X| = s^{2k}$. We have $\pi_0(X) = \pi_\infty(X) = P_{k,s}$, $\pi_1(X) \subset P_{k,2s}$, $\pi_t(X) \subset P_{k+1,2s}$. Since t is transcendental, $S(X, D) \sim |\pi_t(X)| \sim 2^k|X|^{1/2k}|X|^{1/2}$. Choose $k = (\log_2(|X|)/2^{1/2}$ to maximize the RHS, we obtain that is this case $S_D(X) \sim e^{c(\log|X|)^{1/2}}|X|^{1/2}$. It would also be interesting to analyze $S(X, D)$ for other D's, like $D = \{0, 1, \infty, t_1, \ldots, t_k\}$ where t_j's are algebraic independent over $\mathbb{Q}$.

Projection Theory over Finite Fields. We have seen that projection theory over finite fields may be different from that over the reals. Let X be a set of points, D be a set of directions. Let $S = \max_{\theta \in D} \pi_\theta(X)$. Our conjecture is that for $|S| \leq p/2$,

$$|S| \gtrsim |D|^{1/2}|X|^{1/2}.$$

It would be attempting to investigate the structure of sharp examples for this bound, and one may conjecture they are essentially grids.

Remark 9.2. *Let's give an example showing that the original version of Szemerédi-Trotter bound fails over complex field. Again, let X be a set of unit balls in $B_R^{\mathbb{C}^2} \subset \mathbb{C}^2$, $D \subset B_1^{\mathbb{C}} \subset \mathbb{C}$ be an R^{-1}-separated set of directions. For $t \in D$, let $\pi_t : \mathbb{C}^2 \to \mathbb{C}$ be the map $(z_1, z_2) \mapsto z_1 + tz_2$. For our example, choose X to be a maximal set of R^{-1}-separated unit balls with centers in $\mathbb{R}^2$, and D to be a maximal R^{-1}-separated subset of $\mathbb{R} \cap B_R^{\mathbb{C}}$. Then $|X| \sim R^2$, $|D| \sim R$ satisfy the Hausdorff spacing condition, while $S(X, D) \sim R \ll |X|^{1/2}|D|^{1/2}$.*

Projection Theory of Unit Balls. Let X be a set of unit balls in B_R, D be a set of $1/R$-separated directions. Define $N_X(r) = \max_{c \in B_R} |X \cap B(c, r)|$, $N_D(\rho) = \max_{\gamma \subset \mathbb{S}^1, |\gamma| = \rho} |D \cap \gamma|$. We will assume that X has Hausdorff spacing, which means there exists $0 \leq \alpha \leq 1$ such that $|X| \sim R^\alpha$, $N_X(r) \lesssim r^\alpha$. Similarly we will also assume that $|D| \sim R^\beta$, $N_D(r/R) \lesssim r^\beta$. The following conjecture by Furstenberg was recently proved by Orponen, Shmerkin, Ren and Wang:

Theorem 9.3. *Under the above assumption, we have*

$$|D| \lessapprox |S|^2/|X|$$

if $|D| \lesssim R^{-\epsilon} \min(R, |X|)$.

We will discuss briefly why cell decomposition doesn't work in this case. Suppose that we have divided B_R into s^2 cells. There is no guarantee on the shape of each cell O_i, but in one important scenario, most cells are roughly balls of some radius r so that it is possible for us to apply induction hypothesis. (At first one might think $r = R/s$, but this may not be the case. It may be that most of the balls of radius r cover only a fractal subset of B_R which contains our set X.) The problem is, the R^{-1}-separated directions may look indistinguishable at smaller scales. In each cell we have to choose an r-separated subset $D_i \subset D$. By the Hausdorff assumption, in a typical cell we will have $|X_i| \sim r^\alpha$, $|D_i| \sim r^\beta$. So we can not force $|X_i|^2$ to be smaller than $|L_i|$ by simply passing to smaller balls.

18.156 Projection Theory
Spring 2025