Szemerédi-Trotter theorem and applications

M. Rudnev*

December 6, 2004

Abstract

These notes cover the material of two Applied post-graduate lectures in Bristol, 2004.

The theorem

Szemerédi-Trotter theorem ([1]) is a theorem about the maximum number of incidences between curves and points in the plane. It can be understood in terms of geometric graph theory and is not true on finite fields. The theorem has recently become a powerful tool for dealing with so-called "hard Erdös" problem in combinatorics (some will be discussed in the next section). Many of these problems have analogs in analysis and geometric measure theory. These notes give a brief introduction to the subject. For the state of the art exposition see the book by Matoušek ([2]). In preparation of these notes I have often consulted the expository article of Iosevich ([3]) who once introduced me to the subject.

Consider a set P of cardinality $|P| = n$ of points and a set L of cardinality $|L| = m$ of straight lines in the plane. Call the pair (L, P) an arrangement. Let us characterise the complexity of the arrangement by the number of incidences I:

$$I = \sum_{(p,l)\in P\times L} \delta_{pl} = \sum_{p\in P} m(p) = \sum_{l\in L} n(l). \tag{1}$$

Above $\delta_{pl} = 1$ if the point p lies on the line l and zero otherwise. The notations $m(p)$ and $n(l)$ clearly mean the number of lines passing through a given point and the number of points contained in a given line, respectively. The question is to come up with a non-trivial, i.e. better than the obvious $I \leq mn$ upper bound for the number of incidences I.

The first (easy) bound comes from the fact that there is no more than a single line passing through the same pair of points and the Cauchy-Schwartz inequality. Consider an $m \times n$ matrix A with $a_{lp} = \delta_{lp}$. Then

$$I = \sum_{l\in L} n(l) \leq \sqrt{m}\sqrt{\sum_{l\in L} n^2(l)} = \sqrt{m}\sqrt{\sum_{p,p'\in P}\left(\sum_{l\in L}\delta_{pl}\delta_{p'l}\right)}$$
$$\leq \sqrt{m}\sqrt{I+n^2}.$$

*Contact address: Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK; e-mail: m.rudnev@bris.ac.uk

Indeed, the first term in the estimate corresponds to the case $p = p'$ in the double sum; if $p \neq p'$, then there may be no more than one line of L passing through this given pair of points, the number of such pairs obviously not exceeding n^2. Hence,

$$I \lesssim m + n\sqrt{m} \text{ and } I \lesssim n + m\sqrt{n}. \tag{2}$$

The second estimate will follow from the same argument, only using the fact that there is no more than one point of P where a pair of distinct lines can intersect. Henceforth, the notations $a \lesssim b$ will be used if $a \leq Cb$ for some $C > 0$; also $b \gtrsim a$ if $b > ca$ for some c and $a \approx b$ if $a \lesssim b$ and $a \gtrsim b$.

An interesting case is $m \approx n$, when (2) yields

$$I \lesssim n^{3/2}.$$

Can one do better than that? The answer is the following theorem.

Theorem 1 (Szemerédi-Trotter theorem) *Consider an arrangement (L, P) of m curves and n points in $\mathbb{R}^2$ such that*

(i) each pair of curves intersects at no more than $O(1)$ points;

(ii) there are no more than $O(1)$ curves passing through each pair of points.

Then

$$I(L, P) \lesssim m + n + (mn)^{2/3}. \tag{3}$$

The original proof of Theorem 1 in [1] dealt with the case of straight lines only; it was simplified and generalised in [5]. Then Székely ([4]) gave a new short proof, presented further, which shows that Szemerédi-Trotter theorem is in fact a theorem from geometric graph theory. Clearly, the constant hidden in the $\lesssim$ symbol above depends on the constants hidden in $O(1)$ symbols in the formulation of the theorem. Otherwise it is not difficult to take these quantities into account explicitly, [4].

Proof of Theorem 1. Suppose for simplicity that $O(1) = 1$ in (i) and (ii) in the formulation of the theorem. Build a simple graph $G = (V, E)$, whose set of vertices is $V = P$, $v = n$ being the number of vertices and an edge between a pair of vertices is drawn if they are neighbours on the same line. Then one has $I = e + m$, where e is the number of edges in G. The arrangement (L, P) yields a specific drawing of G in the plane.

For any graph $\mathcal{G}$ let $Cr\,\mathcal{G}$ be the number of edge crossings on this drawing. A crossing is counted whenever two edges intersect on the drawing, and the point of their intersection is not a vertex. Also define $Cr_*\mathcal{G}$ as the minimum number of crossings over all possible drawings of $\mathcal{G}$.

If $\mathcal{G}$ is planar graph for instance, then $Cr_*\mathcal{G} = 0$.

Lemma 2 *For any $\mathcal{G}$, one has*

$$Cr_*\mathcal{G} > e - 3v.$$

Proof: If $\mathcal{G}$ were planar, its Euler characteristic

$$f - e + v = 2,$$

where f is the number of faces of $\mathcal{G}$, counting the exterior. Without loss of generality, one can assume that each edge separates two faces. Then $f \leq \frac{2}{3}e$ (the extreme case being $\mathcal{G}$ – a triangle) as each face has at least three edges. This implies $e < 3v$ for a planar graph. This yields the lemma, for if $Cr_*\mathcal{G} \leq e - 3v$ were true, one could remove $Cr_*\mathcal{G}$ edges from $\mathcal{G}$, whereupon $\mathcal{G}$ would have become planar. $\Box$.

Lemma 3 *Either* $e \lesssim v$ *or* $Cr_*\mathcal{G} \gtrsim \frac{e^3}{v^2}$.

Proof. Let $0 < p < 1$, take a biased coin with the probability p for heads. For each vertex of $\mathcal{G}$ toss a coin, and if the coin falls tails, delete this vertex and all the edges incident to it. Let $\mathcal{G}_p$ be a subgraph that remains. Given a drawing of $\mathcal{G}$, the quantities $v_p, e_p, Cr\,\mathcal{G}_p$ are random variables on the probability space $(\mathbb{Z}_2^v, p)$. For each realisation $\mathcal{G}_p$ one has Lemma 2. Taking expectations, one has

$$\mathbf{E}[Cr\,\mathcal{G}_p] > \mathbf{E}[e_p] - 3\mathbf{E}[v_p] \quad \Rightarrow \quad p^4 Cr\,\mathcal{G} > p^2 e - 3pv.$$

Choosing $p \sim \frac{v}{e}$ (e.g. $p = 4\frac{v}{e}$, which makes sense only if $e > 4v$) such that the right hand side is still positive, completes the proof of the lemma. □

To prove Theorem 1 now return to the incidence graph G (where $e = I - m$) and note that by construction $Cr\,G < m^2$. Now apply Lemma 3. That is either $m^2 < Cr\,G \lesssim \dfrac{(I-m)^3}{n^2}$ or $I - m \lesssim n$. The theorem follows. □

Exercise: Work out all the constants in the proof of Theorem 1. In the original proof of Szemerédi-Trotter theorem ([1]) these constants were dozens of orders of magnitude greater.

Proposition 4 *Bound (3) in Szemerédi-Trotter theorem is tight.*

Proof. Let us prove that $I(n,n) \gtrsim n^{4/3}$ for some arrangement (L, P) of points and straight lines to be constructed. Let $n = 4k^3$, for some integer k. For P take the grid $\{0, 1, \ldots, k-1\} \times \{0, 1, \ldots, 4k^2 - 1\}$. For L take all lines $y = ax + b$, with $(a, b) \in \{0, 1, \ldots, 2k-1\} \times \{0, 1, \ldots, 2k^2 - 1\}$. Then for $x \in [0, k)$ one has $ax + b < ak + b < 2k^2 + 2k^2 = 4k^2$, so for each $i = 0, \ldots, k-1$ each line contains a point of P with $x = i \in \{0, 1, \ldots k\}$. Thus $I \approx k^4 \approx n^{4/3}$. □

Exercise: Work out the same example for parabolas rather than straight lines.

Finally, let us consider Szemerédi-Trotter theorem in the context of vector spaces over finite fields. More precisely, let $\mathbb{Z}_q = \{0, 1, \ldots, q-1\}$ (q is a prime) be the finite field of q elements, by addition and multiplication modulo q. Let $\mathbb{Z}_q^2$ be the two dimensional vector space over $\mathbb{Z}_q$. Given $a, b \in \mathbb{Z}_q^2$, with $a \neq (0,0)$, a line in $\mathbb{Z}_q^2$ is the set of points $\{a + tb,\, t \in \mathbb{Z}_q\}$. So each line contains exactly q points. It is easy to show that Szemerédi-Trotter theorem is not true in $\mathbb{Z}_q^2$.

Proposition 5 *Upper bound (2) is sharp in* $\mathbb{Z}_q^2$.

Proof. Note that the basic assumptions (i), (ii) of Theorem 1, which were also used to get (2) are satisfied. So is (2) which was their combinatorial (rather than geometrical, as was the proof of Theorem 1) consequence. Now take $P = \mathbb{Z}_q^2$ and L as the set of all distinct lines in $\mathbb{Z}_2^q$, so $|L|, |P| \approx q^2 \equiv n$. Consider the number of incidences. Each line contains exactly q points, so $I(n,n) \approx n^{3/2}$, which meets the upper bound (2). □

Let us note finally that higher-dimensional generalisations of Szemerédi-Trotter theorem (e.g for incidences between points and surfaces) are problematic. The proof presented above is essentially two-dimensional. There is a technique though of Clarkson et al. ([5]) which does extend to higher dimension, at least if one deals with hyperplanes or spheres. However, a serious obstacle which arises is the necessity of some nondegeneracy conditions, imposed by the fact that even in $\mathbb{R}^3$ one can have rich families of surfaces, which would all intersect along the same curve. E.g. it's easy to come up with a family of translates of a paraboloid $z = x^2 + y^2$ in $\mathbb{R}^3$, which would all intersect along the parabola, say $z = x^2$. Then if the points P all lie on this parabola, each paraboloid contains each point. Or, consider n points on the unit circle in the x_1x_2 plane in $\mathbb{R}^4$ along with m three-spheres of radius $\sqrt{2}$, whose centers all lie on the unit circle in the x_3x_4 plane. Then the number of incidences is trivially mn, as each sphere contains each point.

Some applications

Take a strictly convex curve γ, i.e. every point of γ is an extreme point of its convex interior Ω_γ. Let the origin $O \in \Omega_\gamma$, consider now the domain $t\Omega$, the image of Ω through the homothety centered at O. How many integer lattice points are there on the boundary $t\gamma$?

Proposition 6 *There is a tight upper bound* $|\mathbb{Z}^2 \cap t\gamma| \lesssim t^{2/3}$.

Proof. As L, take the curve $t\gamma$ and all its translates from the origin, to each lattice point inside $t\Omega$. For P take the union of all the lattice points inside all the translates above. Then $|P| \approx |L| \approx t^2$, so the number of incidences $I \lesssim t^{8/3}$, which means at most $t^{2/3}$ incidences per curve, as everything is translation-invariant.

Tightness of the estimate comes from the well established fact that a convex hull of all the integer points inside a large circle of radius t has $\approx t^{2/3}$ vertices. For the proof of this fact (not presented here) and its generalisations, in particular for the case of d dimensions, see [6]. □

Here is another simple application.

Proposition 7 *For any finite point set $A \subset \mathbb{R}$, either the sum set $A + A$ or the product set $A \cdot A$ has $\gtrsim |A|^{5/4}$ elements.*

Proof. For P take the set $(A+A) \times (A \cdot A) \subset \mathbb{R}^2$. For L, take all the lines in the form $x = a+t, y = a't$, $a, a' \in A$. There are $|A|^2$ lines, and an incidence with P occurs if and only if $t \in A$. Hence,

$$I = |A|^3 \lesssim (|A|^2|P|)^{2/3}, \quad \text{so} \quad |P| \gtrsim |A|^{5/2}.$$

Clearly then one of the projections of P exceeds $|A|^{5/4}$ in size. □

Last but not least, for a point set $A \subset \mathbb{R}^2$ define its distance set

$$\Delta(A) = \{\|a - a'\|,\, a, a' \in A\},$$

with respect to the Euclidean distance $\|\cdot\|$. For $t \in \Delta(A)$ define its multiplicity $\nu(t)$ as a number of non-ordered pairs $a, a' \in A$ such that $\|a - a'\| = t$. It was conjectured by Erdös that for any A, t and any $\varepsilon > 0$, $\exists C_\varepsilon$ such that $\nu(t) \leq C_\varepsilon |A|^{1+\varepsilon}$. Note that the conjecture must take advantage of the fact that the Euclidean distance is determined in terms of circles, for otherwise it would contradict Proposition 6. Curiously, there has been no improvement to the following simple corollary of Theorem 1.

Proposition 8 *For any finite point set $A \subset \mathbb{R}^2$, $\sup_{t\in\Delta(A)} \nu(t) \lesssim |A|^{4/3}$.*

Proof. For any t consider $P = A$ and L the set of circles of radius t, centered at points of A. Then the number of incidences I is the number of occurrences $\nu(t)$ of the distance t and $I \lesssim n^{4/3}$, the latter bound being independent of t. □

Proposition 8 implies that $\Delta(A)$ has $\gtrsim |A|^{2/3}$ (distinct) elements, which is the first step towards a weaker, yet a more famous conjecture, also by Erdös, that $|\Delta(A)| \gtrsim C_\varepsilon |A|^{1-\varepsilon}$. Here the best exponent known is slightly larger than 6/7, after a technique developed by Solymosi and Toth ([7]). The technique in essence consists in several divide-and-concur consecutive applications of Szemerédi-Trotter theorem. The Erdös distance conjecture has the following continuous analog due to Falconer. Take a compact Borel set $A \subset \mathbb{R}^d$ of Hausdorff dimension greater than $d/2$. Then the distance set $\Delta(A) \subset \mathbb{R}$ has positive Lebesgue measure.

Convexity and sumsets

This last section is dedicated to sumsets generated by convex functions. Namely, let $f(t)$ be strictly convex (i.e Jensen's inequality is always strict for f), let $B = \{1, 2, \ldots, N\}$, $S = \{f(i), i \in B\}$ and

$$dS = \{x = f(i_1) + \ldots + f(i_d), (i_1, \ldots, i_d) \in B^d\}$$

be the dth sumset of f, with the multiplicity $\nu_d(x)$ be the number of realisations of x, i.e. non-ordered d-tuples $s_1, \ldots, s_d \in S$, such that $x = s_1 + \ldots + s_d$. Clearly, the L_1 norm $\|\nu_d\|_1 = \sum_{x \in dS} \nu_d(x) = N^d$, and the issue is to estimate higher moments $\|\nu_d\|_p^p = \sum_{x \in dS} \nu_d^p(x)$, for $p > 1$ in terms of N, independent of f as long as it is strictly convex. It seems likely that for $p = \infty$, by Proposition 6 in $d = 2$ or the results of [6] for $d > 2$ (which do not apply directly however) one cannot do better than $\|\nu_d\|_\infty \lesssim N^{d\frac{d-1}{d+1}}$, which would yield the first non-trivial bound on all the pth moments.

For the future consider $p = 2$. Knowing the pth moment enables one to render judgement on the support of $\nu(t)$. E.g. by Cauchy-Schwartz:

$$N^{2d} = \|\nu_d\|_1^2 \leq |dS| \cdot \|\nu_q\|_2^2, \quad \text{so} \quad |dS| \geq \frac{N^{2d}}{\|\nu_q\|_2^2}. \tag{4}$$

Observe that the second moment $\|\nu_d\|_2^2$ gives the number of non-ordered d-tuples, solutions of the diophantine equation

$$s_1 + \ldots + s_d = s_{d+1} + \ldots + s_{2d}.$$

Another useful interpretation of the problem is that $\nu_d(x)$ is the number of integer lattice points on the hypersurface $f(x_1) + \ldots + f(x_d) = x$ in $\mathbb{R}^d$.

The conjecture for the upper bound on $\|\nu_d\|_2^2$ comes from an obvious test choice $f(t) = t^2$. It is well known that if $d \geq 5$, a sphere of radius r containing one integer lattice point, actually contains some r^{d-2} points. If $d = 2$ this estimate gets modified by a power of $\log r$. So it is reasonable to conjecture that

$$\|\nu_d\|_2^2 \leq C_\varepsilon N^{2d-2+\varepsilon}, \tag{5}$$

where $C_\varepsilon(d)$ is the same for all strictly convex functions. Note that in such a set-up it is unlikely that one can benefit by a number-theoretical approach. Recent results of Iosevich et al. ([8]) show that one can take ε as small as 2^{-d+1}.

Conjecture (5) gets partially motivated by estimating L_{2d} norms of trigonometric polynomials. Without loss of generality ([8]) one can think that f is integer-valued. If so, let $\theta \in \mathbb{R}/\mathbb{Z}$ and

$$P(\theta) = \sum_{j=1}^{N} \gamma_j e^{2\pi \iota s_j \theta}, \quad \text{with} \quad |\gamma_j| = 1, \forall j.$$

Then it is easy to see that

$$\|\nu_d\|_2^2 = \int_0^1 |P(\theta)|^{2d} d\theta,$$

so (5) is the conjecture about exponential sums.

The following theorem was proved in [8].

Theorem 2 *For $d \geq 2$, let $\alpha_d = 2(1 - 2^{-d})$. Then*

$$\|\nu_d\|_2^2 \lesssim N^{2d-\alpha}. \tag{6}$$

Note that by (4) it follows that

$$|dS| \gtrsim N^{\alpha}. \tag{7}$$

The latter bound is due to Elekes et al. ([9]).

Both bounds (6) and (7) required Szemerédi-Trotter theorem to prove them. However, the former estimate is harder and needs the following weighted version of Theorem 1, [8].

Theorem 3 *Suppose each point p in the arrangement (L, P) has weight $\mu(p) \geq 0$ and each line l has weight $\mu(l) \geq 0$, with the maximum possible weights being $\hat{\mu}_P$ and $\hat{\mu}_L$ respectively. Let (m, n) be net weights of all lines and points, respectively. Count weighted incidences as*

$$I = \sum_{p \in P, l \in L} \mu(p)\mu(l)\delta_{pl}.$$

Then

$$I \lesssim \hat{\mu}_P \hat{\mu}_L \left[\left(\frac{m}{\mu_L} \frac{n}{\mu_P} \right)^{\frac{2}{3}} + \frac{m}{\mu_L} + \frac{n}{\mu_P} \right]. \tag{8}$$

In other words, the maximum number of weighted incidences is achieved when weights are distributed uniformly, the number of distinct lines (points) in the arrangement - note that it does not enter estimate (8) - being the smallest possible.

Let us prove Theorem 2 in the case $d = 2$ and indicate how to extend it to $d > 2$. Let $2B = \{1, \ldots, 2N\}$ and γ be a curve $\{(t, f(t)),\, t \in [1, N]\} \subset \mathbb{R}^2$. Let $(i, u) \in B \times S$, consider the set L of translates γ_{iu} of γ by all vectors (i, u). Now, for $(j, x) \in 2B \times 2S$ consider the system of equations

$$\left\{ \begin{array}{rcl} t + i & = & j, \\ f(t) + u & = & x. \end{array} \right. \tag{9}$$

If $P = \{(j, x)\}$, (9) is the (non-weighted) incidence problem for the arrangement (L, P), which satisfies the assumptions of Theorem 1, by convexity of the curves involved. Let us order $2S = \{x_1, x_2, \ldots\}$ by non-increasing multiplicity $\nu(x)$. Suppose, P_τ is the subset of P of points p such that the number of lines incident to it $m(p) \geq \tau$. Then one has

$$\tau |P_\tau| \leq I(L, P_\tau) \lesssim (N^2 |P_\tau|)^{2/3},$$

where N^2 is the total number of lines. (The first inequality is the definition of P_τ, the second one is (3) where non-interesting linear terms have been omitted.) As $|P_\tau| \approx N|2S|_\tau$, where $2S_\tau$ is the subset $\{x \in 2S,\, \nu_2(x) \geq \tau\}$, of $2S$, one gets

$$|2S_\tau| \lesssim \frac{N^3}{\tau^3}.$$

By ordering in $2S$ this implies, taking the inverse function, that

$$\nu_2(x_t) \lesssim N t^{-1/3} = M_2(t). \tag{10}$$

Expression (10) is a majorant for the multiplicity distribution $\nu_2(x_t)$ in $2S$ and is to be used as follows. As naturally the L_1 norm $\|\nu_2\|_1 = N^2$, define a quantity C_2 implicitly via

$$\int_1^{C_2} M(t)dt = N^2, \tag{11}$$

so $C_2 \approx N^{3/2}$. By (10) the quantity C_2 gives the lower bound for cardinality of $2S$. Now define the upper bound for the mean multiplicity $\bar{\nu}_2$ as

$$\bar{\nu}_2 = \frac{\|\nu_2\|_1}{C_2} = \sqrt{N} \tag{12}$$

and partition $2S$ to $2S_{\bar{\nu}_2}$, where $\nu_2(x) \geq \bar{\nu}_2$, and its complement $2S^c_{\bar{\nu}_2}$. Now

$$\begin{array}{rcccl} \sum_{x\in 2S_{\bar{\nu}_2}} \nu_2^2(x) & \lesssim & \int_1^{C_2} M^2(t)dt & \approx & N^{5/2}, \\ \sum_{x\in 2S^c_{\bar{\nu}_2}} \nu_2^2(x) & \lesssim & \bar{\nu}_2 \sum_{x\in 2S} \nu(x) & = & N^{5/2}. \end{array} \tag{13}$$

This proves Theorem 2 for $d = 2$.

In order to continue by induction, for say $d = 3$, one returns to (9), where u however will now live in $2S$ and is endowed with a weight $\nu_2(u)$. The weight distribution for the (ordered) set of u's is majorated by (10) and one can prove that Theorem 3 applies to estimate the number of weighted incidences, with $\hat{\mu}_P = 1$ and $\hat{\mu}_L = \bar{\nu}_2$, cf. (12). I.e. instead of the maximum L_∞ bound for the weights - which would be trivially N for $d = 2$ (or less trivially $N^{2/3}$ by Proposition 6) - one can use a smaller L_1 bound $\sqrt{N}$. To prove it, one has to partition the set $2S$ onto some $\log\log N$ pieces and apply Theorem 3 for each piece. For more detail about this divide-and-concur approach see [8].

References

[1] E. Szemerédi, W. Trotter. Extremal problems in discrete geometry. *Combinatorica* **3** (1983) 381–392.

[2] J. Matoušek. Lectures on Discrete Geometry. *Springer-Verlag, New York* 2002.

[3] A. Iosevich. Fourier analysis and geometric combinatorics. *To appear, available at* http://www.math.missouri.edu/∼iosevich.

[4] L. Székely. Crossing numbers and hard Erdös problems in discrete geometry. *Combin. Prob. Comput.* **6** (1997) 353–358.

[5] K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, E. Welzl. Combinatorial complexity bounds for arrangements of curves and surfaces. *Disc. Comp. Geom.* **5** (1990) 99–160.

[6] I. Bàràny, D. G. Larman. The convex hull of the integer points in a large ball. *Math. Ann.* **312** (1998) 167–181.

[7] J. Solymosi and C. Toth. Distinct distances in the plane. *Disc. Comp. Geom.* **25** (2001), 629-634.

[8] A. Iosevich, S. Konyagin, M. Rudnev, V. Ten. On combinatorial complexity of convex sequences. *Preprint, available at* http://www.maths.bris.ac.uk/∼maxmr.

[9] G. Elekes, M. Nathanson, I. Ruzsa. Convexity and sumsets. *J. Number Theory* **83** (1999) 194–201.