$|E| \geq 5m|V|$, then $\mathrm{cross}(G) = \Omega(\frac{|E|^3}{|V|^2 m})$. In particular we have $|E| = O(m|V| + m^{1/3}|V|^{2/3}\mathrm{cross}(G)^{1/3})$.

8.2 The Szemerédi–Trotter theorem

Given a finite collection of points P and lines L, a basic question is to bound the number

$$I(P, L) := |\{(p, l) \in P \times L : p \in l\}|$$

of incidences between P and L. Clearly we can make $I(P, L)$ as small as zero without any difficulty, so the interesting question is to maximize $I(P, L)$ for fixed cardinalities $|P|$ and $|L|$. One of course has the trivial bound $I(P, L) \leq |P||L|$, and one can improve this further without difficulty to

$$I(P, L) \leq \min(|P|^{1/2}|L| + |P|, |L|^{1/2}|P| + |L|); \tag{8.2}$$

see exercises. In [348], Szemerédi and Trotter proved the following stronger estimate, which is sharp up to constants.

Theorem 8.3 (Szemerédi–Trotter theorem) *Let P be a finite set of points and let L be a finite set of lines. Then we have*

$$I(P, L) \leq 4|P|^{2/3}|L|^{2/3} + 4|P| + |L|.$$

Proof We may remove those lines $l \in L$ which do not contain any points in P, as they contribute nothing to the left-hand side. Thus we may assume that every line in L contains at least one point in P. Now let $G = G(P, E)$ be the graph whose vertices are the points in P, and two points a and b are connected if and only if the open line segment from a to b lies in a line in L and contains no points in P.

We now apply the double counting method to $|E|$, the number of edges. Observe that if a line l in L contains $k \geq 1$ points in P, then l contributes $k - 1$ edges to E. Summing over $l \in L$, we conclude

$$|E| = I(P, L) - |L|.$$

On the other hand, observe that G has a tautological drawing, with the vertices in P mapping to themselves, and the edge $[a, b]$ mapping to the line segment from a to b. Since any two lines in L can intersect in at most one point, we conclude that $\mathrm{cross}(G) \leq |L|^2$. Applying the crossing number inequality, we conclude that either $|E| \leq 4|P|$ or $\mathrm{cross}(G) \geq |E|^3/64|P|^2$. Thus $|E| \leq \max(4|P|, 4|P|^{2/3}|L|^{2/3})$, and the claim follows. $\square$

Remark 8.4 The above proof is due to Székely [342]; the original proof of Szemerédi and Trotter is quite different (see Exercise 8.4.7 for a proof closer in spirit to that). The symmetry between P and L can be explained by projective duality; if we embed the plane $\mathbf{R}^2$ into the projective space of $\mathbf{R}^3$, then points become associated to subspaces of $\mathbf{R}^3$ of dimension 1, while lines are associated to subspaces of codimension 1.

Let us now derive a few corollaries from the theorem. An immediate consequence, which we leave as a exercise, allows us to bound the number of lines which are "rich" in the sense that they contain many elements of a given set P of points.

Corollary 8.5 (Rich lines) *If P is any finite set of points and $k \geq 2$, then*

$$|\{l \text{ a line} : |l \cap P| \geq k\}| = O\left(\max\left(\frac{|P|^2}{k^3}, \frac{|P|}{k}\right)\right).$$

Dually, for any finite set L of lines, we have

$$|\{p \in \mathbf{R}^2 | \{l \in L : p \in l\}| \geq k\}| = O\left(\max\left(\frac{|L|^2}{k^3}, \frac{|L|}{k}\right)\right).$$

Remark 8.6 In typical applications, such as those below, $k \leq |P|^{1/2}$ so the term $\frac{|P|^2}{k^3}$ is dominating. The case $k > |P|^{1/2}$ can be treated by the cruder estimate (8.2). Similarly for the second half of the corollary.

Next, we bound the number of pairs of points which are connected by a rich line.

Corollary 8.7 (Rich pairs) *If P is any finite set of points and $k \geq 1$, then*

$$|\{(p, q) \in P \times P : p \neq q; k \leq |l_{p,q} \cap P| \leq 2k\}| = O\left(\max\left(\frac{|P|^2}{k}, |P|k\right)\right)$$

where $l_{p,q}$ is the unique line connecting p and q. In particular, if $1 \leq k \leq |P|^{1/2}$, then

$$\left|\{(p, q) \in P \times P : p \neq q; k \leq |l_{p,q} \cap P| \leq |P|^{1/2}\}\right| = O\left(\frac{|P|^2}{k}\right).$$

Proof For the first bound, we observe that each line l with $k \leq |l \cap P| \leq 2k$ contributes at most $O(k^2)$ pairs to the left-hand side, so the claim follows from Corollary 8.5. The second bound follows from the first by a standard dyadic decomposition argument. □

An easy modification of this argument, which we leave as an exercise, allows us to also control collinear triples that are not on too rich of a line:

Corollary 8.8 (Collinear triples) *Let P be a finite set of points. Then the number of triples (u, v, w) where u, v, w are three collinear distinct points in P, whose line contains at most $|P|^{1/2}$ points in P, is at most $O(|P|^2 \log |P|)$.*

Applying this in particular to Cartesian products $P = A \times B$, where A, B are sets of real numbers with $|A| = |B| = m$, we observe that $|P| = m^2$ and no line intersects P in more than $|P|^{1/2} = m$ points. We conclude

Corollary 8.9 *Let A and B be sets of real numbers of cardinality m. Then $A \times B$ contains at most $O(m^4 \log m)$ collinear triples.*

It is an easy matter to extend the Szemerédi–Trotter theorem to more general curves than lines.

Theorem 8.10 (Generalized Szemerédi–Trotter theorem) *[342] Let P be a finite collection of points in $\mathbf{R}^2$, and let L be a finite collection of curves in $\mathbf{R}^2$. Suppose that any two curves in L intersect in at most α points, and any two points in P are simultaneously incident to at most β lines; then*

$$|\{(p, l) \in P \times L : p \in l\}| = O\big(\alpha^{1/3}\beta^{1/3}|P|^{2/3}|L|^{2/3} + |L| + \beta|P|\big).$$

As an application of this theorem we prove the following remarkable result of Andrews [13].

Theorem 8.11 *Let $\Gamma \subset \mathbf{R}^2$ be a lattice (e.g. $\Gamma = \mathbf{Z}^2$). If C is a convex n-gon with vertices in Γ, then the interior of C contains $\Omega(n^3)$ lattice points.*

Proof Let $\mathcal{C}$ be the boundary of C and F be collection of (piecewise linear) curves obtained by translating $\mathcal{C}$ by the lattice points inside C. Let P be the set of lattice points covered by the union of the curves in F and m be the number of lattice points inside C. We have $|F| = m$ and $|P| = \Theta(m)$ (cf. (3.10)).

We apply the double counting method to the number of incidences between P and F. On the one hand, the generalized Szemerédi–Trotter theorem gives an upper bound of $O(m^{4/3})$ for these incidences. On the other hand, each translate of $\mathcal{C}$ contains exactly n points, so the number of incidences is at least nm. Comparing these bounds we obtain $m = \Omega(n^3)$ as desired. □

Remark 8.12 The above theorem generalizes for $\mathbf{R}^d$. For any fixed d, Andrews proved that a convex polytope in $\mathbf{R}^d$ with n non-coplanar integral points on its boundary has volume $\Omega(n^{(d+1)/(d-1)})$. The above proof, however, does not generalize for higher dimensions.

An important open problem is to extend the Szemerédi–Trotter theorem to planes over other fields, for instance the complex plane $\mathbf{C}^2$ or the finite field planes F_p^2. The crude estimate (8.2) applies in all of these situations, but one would like to improve this bound. In the case of F_p^2 it was shown that $I(P, L) = O_\delta(\max(|P|, |L|)^{3/2-\varepsilon(\delta)})$ whenever $|P|, |L| \leq p^{2-\delta}$ for all $\delta > 0$ and some $\varepsilon(\delta) > 0$ depending on δ; see [43], [44]. The main ingredients in this argument was the sum-product estimate in Corollary 2.58 and the Balog–Szemerédi–Gowers theorem (Theorem 2.29).

Exercises

8.2.1 Using only the basic facts that two distinct points determine at most one line, and two distinct lines intersect in at most one point, together with the Cauchy–Schwarz inequality, prove (8.2). Observe that this argument works over any field, not just $\mathbf{R}$. In the case where the field is F_{p^2}, show that the bound can be sharp when $|P| = |L| = p^2$, or when $|P| = |L| = p^4$.

8.2.2 Let $n, m \geq 1$ be given. Find an example of a set of points P and a set of lines L such that $|P| = n$, $|L| = m$, and the number of incidences between P and L is $\Theta(n^{2/3}m^{2/3} + n + m)$, thus demonstrating that the Szemerédi–Trotter theorem is sharp up to constants. (Hint: consider sets P of the form $P = [1, a] \times [1, ab]$ for various parameters a, b.)

8.2.3 Prove Corollary 8.5.

8.2.4 Prove Corollary 8.8.

8.2.5 Let P be a finite set of points, and let $k \geq 2$. Show that

$$|\{(p, l) : p \in P; l \text{ a line}; p \in l; |l \cap P| \geq k\}| = O\left(\frac{|P|^2}{k^2} + |P| \log |P|\right).$$

8.2.6 (Beck's theorem) [19] Let P be a finite set of points. Show that either there exists a line that is incident to $\Theta(|P|)$ points in P, or there exist $\Theta(|P|^2)$ lines that are each incident to exactly two points in P.

8.2.7 (Sylvester–Gallai theorem) Let P be a finite set of points, not all of which are collinear. Show that there exists a line that contains exactly two points in P. (Hint: minimize the quantity $\mathrm{dist}(p, l)$, where l is a line containing two or more points in P and $p \in P \backslash l$. Using elementary geometry, show that this quantity is minimized only when l contains exactly two points from P.)

8.2.8 Prove Theorem 8.10. (Hint: use Exercise 8.1.5.)

8.2.9 Let γ be a strictly convex curve in $\mathbf{R}^2$. Show that $|(R \cdot \gamma) \cap \Gamma| = O_\gamma(R^{2/3})$ for all $R \geq 1$ and all lattices Γ.

8.2.10 Let γ be a strictly convex curve in $\mathbf{R}^2$, and let A be a finite set in $\mathbf{R}^2$. Show that $|\{(a, a') \in A \times A : a - a \in \gamma\}| = O(|A|^{4/3})$. Deduce from this that $|\{|x - y| : x, y \in A\}| = \Omega(|A|^{2/3})$.