SZEMERÉDI-TROTTER INCIDENCE THEOREM AND APPLICATIONS

PAUL H. KOESTER

Given a finite set of points $P \subset \mathbb{R}^2$ and a finite family of lines L, we define the set of incidences of P and L by

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

In words, $I(P, L)$ consists of all occurrences of a point $p \in P$ lying on a line $l \in L$. A basic problem is to determine how many such occurrences there may be. Unless P and L are given explicitly this task is impossible, but we can try to find upper bounds on how big $I(P, L)$ may be, in terms of the sizes of P and L. (Finding lower bounds is trivial; we may choose sets P and L with any size with $I(P, L) = \emptyset$.)

The first upper we find is

$$|I(P, L)| \leq |P||L|. \tag{1}$$

To prove this, we observe that each point p can belong to at most $|L|$ pairs (p, l) since there are only $|L|$ possible lines. As there are $|P|$ points, we therefore obtain the advertised bound.

The argument we just gave for (1) leaves room for quite a bit of improvement. Before reading on you should try playing around with some simple cases, like when $|P|$ and $|L|$ are 3 or 4 and try to convince yourself that $|I(P, L)|$ can never be as big as $|P||L|$, unless one of P or L has size one.

The problem is that the above argument did not utilize any specific information about the behavior of points and lines. For example, suppose P and L are sets of points and sets of lines, and suppose we know that every point of P lies on the line l. Then we know that at most one point of P may lie on any of the other lines in L. This is because a line is determined by two points, and therefore two different lines may intersect in at most one point.

The previous paragraph suggests that we should be able to do better than

$$|I(P, L)| \leq |P||L|,$$

Date: October 16, 2006.

These are notes to accompany the author's talk at the WUSTL Math Club.

so the question now is how much better? The correct inequality was first proved by Szemerédi and Trotter [ST83]in the early 1980's. Their proof is long and complicated but a much shorter proof was found by Székely in [Szé97].

Szemerédi-Trotter Incidence Theorem. *There exists a constant $C > 0$ so that*

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

for any choice of P and L.

In most applications $|P|$ and $|L|$ will be very large and roughly the same size, in which case we may use the estimate

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

for a somewhat larger constant.

That this is the correct estimate was conjectured by Erdős, and is based on the following example: Let N be any positive integer and

$$P = \{(n, m) \in \{1, 2, \dots, N\} \times \{1, 2, \dots, N^2\}.$$

For each point $(n, 1)$ with $n \in \{1, 2, \dots, N^2\}$ we make a judicious choice of N lines passing through $(n, 1)$. Regardless of the choice of lines, we have $|P| = N^3$ and $|L| = N^3$. Choosing L correctly, we will have that each line in L has exactly N points of P from which we deduce

$$I(P, L) = N^4 = |P|^{2/3}|L|^{2/3}.$$

This example shows that, apart from the specific value of C, the Szemerédi-Trotter Theorem is sharp.

This theorem has found applications in many problems in discrete geometry, harmonic analysis, and number theory. We mention two such applications here.

1. Sum-Product Theorem

If A is a finite subset of integers then we define the sumset of A as

$$A + A = \{a + b : a, b \in A\} \tag{2}$$

and the product set

$$A \cdot A = \{ab : a, b \in A\}. \tag{3}$$

It is fairly easy to check

$$2|A| - 1 \leq |A + A|, |A \cdot A| \leq |A|^2.$$

If A is an arithmetic progression, say

$$A = \{0, 1, 2, \ldots, M\}$$

then $|A + A| = 2M + 1$. That is, A has a small sumset, but the product set appears to be large. On the other hand, if A is a geometric progression, say

$$A = \{1, 2, 4, \ldots, 2^M\}$$

then $|A \cdot A| = 2M + 1$. That is, A has a small product set, but the sumset appears to be quite large. Based on these observations, Paul Erdős and Endré Szemerédi conjectured in [ES83] that this is always true:

Erdős-Szemerédi Conjecture. *For every $\epsilon > 0$, there exists a constant $C_\epsilon > 0$ so that we have*

$$\max\{|A + A|, |A \cdot A|\} \geq C_\epsilon |A|^{2-\epsilon}$$

provided $|A|$ is sufficiently large.

That is, for any set finite set A, at least one of $|A + A|$ or $|A \cdot A|$ is almost as large as possible.

We are a long way away from solving this conjecture completely. We do have a partial result, due to Elekes [Ele97], which uses the Szemerédi-Trotter theorem in a very nice way.

Elekes' Theorem. *There exists a positive constant C so that*

$$\max\{|A + A|, |A \cdot A|\} \geq |A|^{5/4}.$$

Joszef Solymosi [Sol05] has improved the 5/4 to 14/11 but the argument is much more intricate (but still uses incidence geometry).

The relevance of point-line incidence geometry to this theorem is suggested by the equation of a line:

$$y = a \cdot (x - b).$$

The right hand side involves both addition and multiplication. In order to apply Szemerédi-Trotter to the sum product theorem we need to find the correct choice of P and L. We let

$$P = (A + A) \times (A \cdot A)$$

and

$$L = \{l_{a,b} : a, b \in A\}$$

where

$$l_{a,b} : y = a \cdot (x - b).$$

Note that $|L| = |A|^2$.

For each line $l_{a,b}$ with $a, b \in A$ we have the point $(c + b, a \cdot c) \in l_{a,b}$ for each $c \in A$. This follows from the calculation

$$a \cdot ((c + b) - b) = a \cdot c.$$

Thus, every line in L has at least $|A|$ points from P and so

$$|I(P, L)| \geq |L| = |A|^3.$$

On the other hand, Szemerédi-Trotter implies [1]

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

Combining these inequalities we find

$$|A|^{5/2} \leq C|A + A||A \cdot A|$$

and so

$$\max\{|A + A|, |A \cdot A|\} \geq C|A|^{5/4}.$$

2. The Number of Lattice Points on a Convex Curve

Let Γ denote a fixed strictly convex curve in the plane, and suppose this curve lies in the box $[1, n]^2$ for some integer n. Is it possible to estimate the number of integer points which lie on Γ? Since $\Gamma \subset [1, n]^2$ and $[1, n]^2$ has only n^2 integer points we certainly can't have more than n^2 integer points on Γ. A better argument, which invokes the hypothesis that Γ is strictly convex, yields the estimate that there are no more than $2n$ integer points on Γ. Using a variant of Szemerédi-Trotter we can push this down to $Cn^{2/3}$. This is the 2 dimensional version of a theorem of George Andrews [And63]. The proof using Szemerédi-Trotter was communicated to us by Alex Iosevich.

First, we remark that the Szemerédi-Trotter Theorem applies to incidences between points and object referred to as pseudo-lines. The point being that the proof of Szemerédi-Trotter only uses the fact that a line is determined by a finite number of points. A family of r-pseudo-lines is a collection of subsets of the plane so that any such subset is determined by r points. Thus, if L is a family of r-psuedo-lines and P is a set of points, then

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

[1] We use a standard convention in this subject, that C always denotes a positive constant which may change from line to line.

where the value of C may depend on r.

A family of algebraic curves of degree not exceeding $r-1$ is a family of r-pseudo-lines.

The relevant example of a family of 2-pseudo-lines for us consists of a collection of translates of a single strictly convex curve;

$$L = \{\Gamma + (x, y) : (x, y) \in \{1, 2 \ldots, n\} \times \{1, 2 \ldots, n\}\}.$$

We have $|L| = n^2$. We take our point set P to be

$$P = \{1, 2, \ldots, 2n\} \times \{1, 2, \ldots, 2n\},$$

and so $|P| = 4n^2$. We therefore have

$$|I(P, L)| \leq C|P|^{2/3}|L|^{2/3} = 16^{1/3}Cn^{8/3}.$$

On the other hand, for each $(x, y) \in \{1, 2 \ldots, n\} \times \{1, 2 \ldots, n\}$ the curve $\Gamma + (x, y)$ hits exactly $|\Gamma \cap \mathbb{Z}^2|$ points of P. This yields

$$|I(P, L)| \geq |\Gamma \cap \mathbb{Z}^2| n^2.$$

Combining the inequalities, we obtain

$$|\Gamma \cap \mathbb{Z}^2| \leq 16^{1/3}Cn^{2/3}.$$

References

[And63] George E. Andrews. A lower bound for the volume of strictly convex bodies with many boundary lattice points. *Trans. Amer. Math. Soc.*, 106:270–279, 1963.

[Ele97] György Elekes. On the number of sums and products. *Acta Arith.*, 81(4):365–367, 1997.

[ES83] P. Erdős and E. Szemerédi. On sums and products of integers. In *Studies in pure mathematics*, pages 213–218. Birkhäuser, Basel, 1983.

[Sol05] József Solymosi. On the number of sums and products. *Bull. London Math. Soc.*, 37(4):491–494, 2005.

[ST83] Endre Szemerédi and William T. Trotter, Jr. Extremal problems in discrete geometry. *Combinatorica*, 3(3-4):381–392, 1983.

[Szé97] László A. Székely. Crossing numbers and hard Erdős problems in discrete geometry. *Combin. Probab. Comput.*, 6(3):353–358, 1997.

Department of Mathematics, Washington University, St. Louis, MO 63130, U. S. A.

E-mail address: `paul@math.wustl.edu`