SZEMEREDI-TROTTER INCIDENCE THEOREM, RELATED RESULTS AND SOME AMUSING CONSEQUENCES

Alex Iosevich

June 24, 2002

Geometric combinatorics is an old and beautiful subject. We will not even attempt to cover anything resembling a significant slice of this broad and influential discipline. See, for example, [AgPa95] for a thorough description of this subject. The purpose of this article is to give an elementary proof of Szemeredi-Trotter incidence theorem ([ST83]), a result that found a tremendous number of applications in combinatorics, analysis, and analytic number theory. We shall describe some of the consequences of this seminal result.

Definition. An incidence of a point and a line is a pair (p, l), where p is a point, l is a line, and p lies on l.

Theorem 0.1. *(Szemeredi-Trotter) Let I denote the number of incidences of a set of n points and m lines (or m strictly convex closed curves). Then*

$$I \lesssim n + m + (nm)^{\frac{2}{3}}, \tag{0.1}$$

where here and throughout the paper, $A \lesssim B$ means that there exists a positive constant C such that $A \leq CB$.

Corollary 0.2. *Let S be a subset of $\mathbb{R}^2$ of cardinality n. Let $\Delta(S) = \{|x-y| : x, y \in S\}$, where $|\cdot|$ denotes the Euclidean norm. Then*

$$\#\Delta(S) \gtrsim n^{\frac{2}{3}}. \tag{0.2}$$

This estimate is not sharp. It is conjectured to hold with the exponent 1 in place of $\frac{2}{3}$. The best known exponent to date, $\frac{6}{7}$ is due to Solomosi and Toth ([SolToth2001]). However, Corollary 0.2 is still quite useful as we shall see in the final section of this paper.

Research supported in part by an NSF grants

Corollary 0.3. *Let A be a subset of $\mathbb{R}$ of cardinality n. Then either $A+A=\{a+a' : a,a' \in A\}$ or $A \cdot A = \{aa' : a,a' \in A\}$ has cardinality $\gtrsim n^{\frac{5}{4}}$.*

This paper is organized as follows. In the first section we prove Theorem 0.1. In the second section we deduce the corollaries. In the last section we give some applications of Theorem 0.1 to problems in analysis and convex geometry, including the non-existence of orthogonal Fourier bases of the disc. In the final section of the paper, we describe a combinatorial principle due to Moser which allows us to prove non-existence of orthogonal Fourier bases for a ball in any dimension greater than 1.

Section I: Proof of Theorem 0.1

We shall deduce Theorem 0.1 from the following graph theoretic result due to Ajtai et al ([Ajtai86]), and, independently, to Leighton.

Definition. The crossing number of a graph, $cr(G)$, is the minimal number of crossings over all the planar drawings of this graph. A crossing is an intersection of two edges not at a vertex.

Theorem 1.1. *Let G be a graph with n vertices and e edges. Suppose that $e \geq 4n$. Then*

$$cr(G) \gtrsim \frac{e^3}{n^2}. \tag{1.1}$$

Before proving Theorem 1.1, we show how it implies Theorem 0.1. Take the points in the statement of the Theorem as vertices of a graph. Connect two vertices with an edge if the two corresponding points are consecutive on some line. It follows that

$$e = I - m. \tag{1.2}$$

If $e > 4n$ we get $I < 4n+m$, which is fine with us. If $e \geq 4n$, we invoke Theorem 1.1 to see that

$$cr(G) \gtrsim \frac{e^3}{n^2} = \frac{(I-m)^3}{n^2}. \tag{1.3}$$

Combining (1.3) with the obvious estimate $cr(G) \leq m^2$, we complete the proof of Theorem 0.1.

We now turn our attention to the proof of Theorem 1.1. Let G be a planar graph with n vertices, e edges and f faces. Euler's formula (easy to prove...) says that

$$n - e + f = 2. \tag{1.4}$$

Combined with the observation that $3f \leq 2e$, we see that in such a planar graph

$$e \geq 3n - 6. \tag{1.5}$$

It follows that if G is any graph, then

$$cr(G) \geq e - 3n. \tag{1.6}$$

We now convert this linear estimate into the estimate we want by randomization. More precisely, let G be as in the statement of Theorem 1.1 and let H be a random subgraph of G formed by chosing each vertex with probability p to be chosen later. Naturally, we keep an edge if and only if both vertices survive the random selection. Let $\mathbb{E}()$ denote the usual expected value. An easy computation yields

$$\mathbb{E}(vertices) = np, \tag{1.7}$$

$$\mathbb{E}(edges) = ep^2, \tag{1.8}$$

$$\mathbb{E}(crossing\ number\ of\ H) \leq p^4 cr(G). \tag{1.9}$$

It follows that

$$cr(G) \geq \frac{e}{p^2} - \frac{3n}{p^3}. \tag{1.10}$$

Chosing $p = \frac{4n}{e}$ we complete the proof of Theorem 1.1.

Proof of Corollary 0.2 and Corollary 0.3

To prove Corollary 0.3, draw a circle of fixed radius around each point in S. By Theorem 0.1, the number of incidences is $\lesssim n^{\frac{4}{3}}$. This means that a single distance cannot repeat more than $\approx n^{\frac{4}{3}}$ times. It follows that there must be at least $\approx n^{\frac{2}{3}}$ distinct distances since the total number of distances is $\approx n^2$. In other words, we just proved that $\Delta(S) \gtrsim n^{\frac{2}{3}}$ as promised.

The choice of lines and points is less obvious in the proof of Corollary 0.3. Let $P = (A+A) \times (A \cdot A)$. Let L be the set of lines of the form $\{(ax, a' + x) : a, a' \in A\}$. We have

$$\#P = \#(A+A) \times \#(A \cdot A), \tag{2.1}$$

$$\#L = n^2, \tag{2.2}$$

while the number of incidences is clearly $n \times n^2 = n^3$. It follows that

$$n^3 \lesssim (\#P)^{\frac{2}{3}} n^{\frac{4}{3}}, \tag{2.3}$$

which means that

$$\#P \gtrsim n^{\frac{5}{3}}. \tag{2.4}$$

It follows that either $\#(A+A)$ or $\#(A \cdot A)$ exceeds a constant multiple of $n^{\frac{5}{4}}$. This completes the proof of Corollary 0.3.

Definition. We say that a domain $\Omega \subset \mathbb{R}^d$ is spectral if $L^2(\Omega)$ has an orthogonal basis of the form $\{e^{2\pi i x\cdot a}\}_{a\in A}$.

The following result is due to Fuglede ([Fuglede74]). It was also proved in higher dimensions by Iosevich, Katz and Pedersen ([IKT99]).

Theorem 2.1. *A disc, $D = \{x \in \mathbb{R}^2 : |x| \leq r\}$, is not spectral.*

Proof of Theorem 2.1. Let A denote a putative spectrum. We need the following basic lemmas:

Lemma 2.2. *A is separatee in the sense that there exists $c > 0$ such that $|a - a'| \geq c$ for all $a, a' \in A$.*

Lemma 2.3. *There exists $s > 0$ such that any square of side-length s contains at least one element of A.*

For a sharper version of Lemma 2.3 see [IP2000].

The proof of Lemma 2.2 is straight-forward. Orthogonality implies that

$$\int_D e^{2\pi i x\cdot(a-a')}dx = 0, \tag{2.1}$$

whenever $a \neq a' \in A$. Since $\int_D dx = 2\pi r$ and the function $\int_D e^{2\pi i x\cdot\xi}dx$ is continuous, the left hand side of (2.1) would have to be strictly positive if $|a - a'|$ were small enough. This implies that $|a - a'|$ can never be smaller than a positive constant depending on r.

The proof of Lemma 2.3 is a bit more interesting. By Bessel's inequality we have

$$\sum_A |\widehat{\chi}_D(\xi + a)|^2 \equiv |D|^2, \tag{2.2}$$

for almost every $\xi \in \mathbb{R}^d$, since the left hand side is a sum of squares of Fourier coefficients of the exponential with the frequency ξ with respect to the putative orthogonal basis $\{e^{2\pi i x\cdot a}\}_{a\in A}$. We have

$$\sum_{A_\xi} |\widehat{\chi}_D(a)|^2 = \sum_{A_\xi\cap Q_s} + \sum_{A_\xi\cap Q_s^c} = I + II, \tag{2.3}$$

where $A_\xi = A - \xi$ and Q_s is a square of side-length s centered at the origin.

We invoke the following basic fact. See, for example, [Stein93]. We have

$$|\widehat{\chi}_D(\xi)| \lesssim |\xi|^{-\frac{3}{2}}. \tag{2.4}$$

It follows that

$$II \lesssim \sum_{A_\xi \cap Q_s^c} |a|^{-3} \lesssim s^{-1}. \tag{2.5}$$

Chosing s big enough so that $s^{-1} << |D|^2$, we see that $I \neq 0$, and, consequently, that $A_\xi \cap Q_s$ is not empty. This completes the proof of Lemma 2.3.

We are now ready to complete the proof of Theorem 2.1. Intersect A with a large disc of radius R. By Lemma 2.2 and Lemma 2.3, this disc contains $\approx R^2$ points of A. We need another basic fact about $\widehat{\chi}_D(\xi)$, that it is radial, and in fact equals, up to a constant, to $|\xi|^{-1} J_1(2\pi|\xi|)$, where J_1 is the Bessel function of order 1. We also need to know that zeros of Bessel functions are separated in the sense of Lemma 2.2. This fact is contained in any text on special functions. See also [SteinWeiss71].

With this information in tow, recall that orthogonality implies that $|a-a'|$ is a zero of J_1. Since the largest distance in the disc of radius R is $2R$ and zeros of J_1 are separated, we see that the total number of distinct distances between the elements of A in the disc or radius R is at most $\approx R$. This is a contradiction since Corollary 0.2 says that R^2 points determine at least $R^{\frac{4}{3}}$ distinct distances. This completes the proof of Theorem 2.1.

Application to convex geometry

The following result is due to Andrews ([Andrews61]).

Theorem 3.1. *Let Q be a convex polygon with n integer vertices. Then $n \leq |Q|^{\frac{1}{3}}$.*

Proof of Theorem 2.1. Let $\mathcal{C}$ denote a strictly convex curve running through the vertices of Q. Let Ω denote the convex domain bounded by $\mathcal{C}$. Let L denote the set of strictly convex curves obtained by translating $\mathcal{C}$ by every lattice point inside Ω. Let P denote the set of lattice points contained in the union of all those translates. By Theorem 0.1 the number incidences between the elements of P and elements of L is $\lesssim |\Omega|^{\frac{4}{3}}$ since $\#L \approx \#P \approx |\Omega|$. Since each translate of $\mathcal{C}$ contains exactly the same number of lattice points,

$$\#\mathcal{C} \cap \mathbb{Z}^2 \lesssim \frac{|\Omega|^{\frac{4}{3}}}{|\Omega|} = |\Omega|^{\frac{1}{3}}. \tag{3.1}$$

This completes the proof of Theorem 3.1. See [Iosevich2002] for a more general version of this result.

What sort of an incidence theorem would be required to prove a more general version of this result? Well, suppose we had a theorem which said that the number of incidences between n points and n strictly convex hypersurfaces in general position in $\mathbb{R}^d$ is $\lesssim n^\alpha$. By general position we simply mean that intersection of any d of these hypersurfaces contains at most 2 points. Repeating the argument above, we would arrive at the conclusion that if P is a convex polyhedron with N lattice vertices, then

$$|P| \gtrsim N^{\frac{1}{\alpha-1}}. \tag{3.2}$$

However, a higher dimensional version of the aforementioned theorem of Andrews says that $|P| \gtrsim N^{\frac{d+1}{d-1}}$. This leads us to conjecture the following.

Conjecture 3.2. *The number of incidences between n points and n strictly convex hypersurfaces in general position is $\lesssim n^{2-\frac{2}{d+1}}$.*

This result would be sharp in view of (3.2) and the following result due to Barany and Larman ([BL98]).

Theorem 3.3. *The number of vertices of P_R, the convex hull of the lattice points contained in the ball of radius $R >> 1$ centered at the origin is $\approx R^{d\frac{d-1}{d+1}}$.*

Higher dimensions

Definition. We say that $A \subset \mathbb{R}^d$ is well-distributed if the conclusions of Lemma 2.2 and Lemma 2.3 hold for A.

Theorem 4.1. *If $R > 0$ is sufficiently large, then*

$$\#(\Delta(A \cap [-R, R]^d) \gtrsim R^{2-\frac{1}{d}}. \tag{4.1}$$

Corollary 4.2. *The ball $B_d = \{x : |x| \leq 1\}$ is not spectral in any dimenion greater than* 1.

Corollary 4.2 follows from Theorem 4.1 in the same way as Theorem 2.1 follows from Corollary 0.2. Lemma 2.2 and Lemma 2.3 go through without change except that in $\mathbb{R}^d$,

$$|\widehat{\chi}_{B_d}(\xi)| \lesssim |\xi|^{-\frac{d+1}{2}}, \tag{4.2}$$

$\widehat{\chi}_{B_d}(\xi)$ is a constant multiple of

$$|\xi|^{-\frac{d}{2}} J_{\frac{d}{2}}(2\pi|\xi|), \tag{4.3}$$

and the zeroes of $J_{\frac{d}{2}}$ are still separated.

See [Stein93], [SteinWeiss71] and/or any text on special functions for the details. Better yet, prove it yourself- it is not very difficult.

We are left to prove Theorem 4.1. Since A is well-distributed, there is $s > 0$ such that every cube of side-length s contains at least one point of A. Fix a reference cube of side-length s and consider a row of consecutive cubes in each of the coordinate directions with respect to the reference cube. Chose a point of A in the 10th cube in each coordinate direction. Name those points $P_1, P_2, \dots, P_d$. Let O denote the center of the reference cube. Construct a system of annuli centered at O of width Md, with the first annulus of radius $\approx R$. Construct $\approx R$ such annuli.

It follows from the assumption that A is well distributed that each constructed annulus $\mathcal{A}$ has $\approx R^{d-1}$ points of A. Let

$$\cup_{i=1}^{d}\{|x-P_i| : x \in \mathcal{A}\} = \{d_1, \dots, d_k\}. \tag{4.4}$$

Let

$$A_j^l = \{x \in \mathcal{A} \cap A : |x-P_l| = d_j\}. \tag{4.5}$$

It is not hard to see that

$$A_j^l = \cup_{1 \le j_m \le k} \cup_{m=1}^{d-1} A_j^l \cap_{l' \neq l} A_{j_m}^{l'}. \tag{4.6}$$

Taking unions of both sides in j and counting, we see that

$$R^{d-1} \lesssim k^d, \tag{4.7}$$

where we have used the fact that the intersection of d spheres in question consists of at most two points. Taking d'th roots and using the fact that we have $\approx R$ annuli with $\approx R^{d-1}$ point of A, we conclude that

$$\#\Delta(A \cap [-R,R]^d) \gtrsim R^{1+\frac{d-1}{d}} = R^{2-\frac{1}{d}}, \tag{4.8}$$

as desrired.

Some comments on finite fields

In this section we consider incidence theorems in the context of finite fields. More precisely, let F_q denote the finite field of q elements. Let F_q^d denote the d-dimensional vector space over F_q. A line in F_q^d is a set of points $\{x+tv : t \in F_q\}$ where $x \in F_q^d$ and $v \in F_q^d \setminus (0, \dots, 0)$. A hyperplane in F_q^d is a set of points $(x_1, \dots, x_d)$ satisfying the equation $A_1x_1 + \cdots + A_dx_d = D$, where $A_1, \dots, A_d, D \in F_q$ and not all A_j's are 0.

It is clear that without further assumptions, the number of incidences between n hyperplanes and n points is $\approx n^2$ and no better, since we can take all n plnaes to be rotates of the same plane about a line where all the points are located. We shall remove this "difficulty" by operating under the following non-degeneracy assumption.

Definition. We say that a family of hyperplanes in F_q^d is non-degenerate if the intersection of any d (or fewer) of the hyperplanes in the family contains at most one point.

The main result of this section is the following:

Theorem 5.1. *Suppose that a family $\mathcal{F}$ of n hyperplanes in F_q^d is non-degerate. Let $\mathcal{P}$ denote a family of n points in F_q^d. The the number of incidences between the elements of $\mathcal{F}$ and $\mathcal{P}$ is $\lesssim n^{2-\frac{1}{d}}$. Moreover, this estimate is sharp.*

We prove sharpness first. Let $\mathcal{F}$ denote the set of all the hyperplanes in F_q^d and $\mathcal{P}$ denote the set of all the points in F_q^d. It is clear that $\#\mathcal{F} \approx \mathcal{P} \approx q^d$. On the other hand, the number of incidences is simply the number of hyperplanes times the number of points on each hyperplanes, which is $\approx q^{2d-1}$. Since $q^{2d-1} = (q^d)^{2-\frac{1}{d}}$, the sharpness of the Theorem 5.1 is proved.

We now prove the positive result. Consider an n by n matrix whose (i,j) entry if 1 if i'th point lies on j's line, and 0 otherwise. The non-degeneracy condition implies that this matrix does not contain a d by 2 submatrix consisting of 1's. Using Holder's inequality we see that the number of incidences,

$$I = \sum_{i,j} I_{ij} \leq \left(\sum_i \left(\sum_j I_{ij} \right)^d \right)^{\frac{1}{d}} \times n^{\frac{d-1}{d}} \tag{5.1}$$

$$= \left(\sum_i \sum_{j_1,\dots,j_d} I_{ij_1} \dots I_{ij_d} \right)^{\frac{1}{d}} \times n^{\frac{d-1}{d}} \lesssim n \times n^{\frac{d-1}{d}} = n^{2-\frac{1}{d}}, \tag{5.2}$$

because when j_k's are distinct, $I_{ij_1} \dots I_{ij_d}$ can be non-zero for at most one value of i due to the non-degeneracy assumption. If j_k's are not distinct, we win for the same reason. This completes the proof of Theorem 5.1.

Why should the finite field case be different from the Euclidean case? The proof of Szemeredi-Trotter theorem given above suggests that main result may be the notion of order. In the proof of Szemeredi-Trotter we used the fact that points on a line may be ordered. However, no such notion exists in a finite field. Nevertheless, Tom Wolff conjectured that if q is a prime, then there exists $\epsilon > 0$ such that the number of incidences between n points and n lines in F_q^2 should not exceed $n^{\frac{3}{2}-\epsilon}$ for $n \approx q$. There has been no progress on this conjecture as far as we know. The notion of order in F_q^2 when q is a prime is partially addressed in [IosevichII2002].

References

[AgPa95] *Combinatorial geometry*, A Wiley-Interscience publication, 1995.

[Ajtai86] M. Ajtai, V. Chvatal, M. Newborn, and E. Szemeredi, *Crossing-free subgraphs*, North-Holland Math Studies **60** (1986).

[Andrews63] G. Andrews, *A lower bound for the volume of strictly convex bodies with many boundary lattice points*, Trans. Amer. Math. Soc. **106** (1963), 270-279.

[BL98] I. Barany and D. Larman, *The convex hull of the integer points in a large ball*, Math. Ann. **312** (1999), 167-181.

[Iosevich2002] A. Iosevich, *A generalization of Andrew's theorem*, (submitted for publication) (2002).

[IosevichII2002] A. Iosevich, *The notion of order in a two-dimensional vector space over a finite field with prime number of elements.*, (in preparation). (2002).

[IKT99] A. Iosevich, N. Katz, and S. Pedersen, *Fourier bases and a distance problem of Erdos*, Math. Res. Lett. **6** (1999), 251-255.

[IP2000] A. Iosevich and S. Pedersen, *How large are the spectral gaps?*, Pacific J. Math. **192** (2000).

[Fuglede74] B. Fuglede, *Commuting self-adjoint partial differential operators and a group theoretic problem*, J. Func. Anal. **16** (1974), 101-121.

[SolToth2001] J. Solymosi and Cs. D. Toth, *Distinct distances in the plane*, Discrete Comp. Geom. (Misha Sharir birthday issue) **25** (2001), 629-634.

[Stein93] E. M. Stein, *Harmonic Analysis*, Princeton University Press, 1993.

[ST83] E. Szemeredi and W. Trotter, *Extremal problems in discrete geometry*, Combinatorica **3** (1983), 381-392.

[SteinWeiss71] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, 1971.