Point-Line Incidences.

F. Nilov, A. Polyanskii, N. Polyanskii

Figure 1: Paul Erdos

Once upon a time there was a mathematician named Erdos, who invented a lot of striking problems. One day he gave us the following riddle.

Given a finite set $\mathcal{L}$ of lines and a finite set P of points, let $I(\mathcal{L}, P)$ denote the number of incidences between $\mathcal{L}$ and P, i.e., the number of pairs (l, p), $l \in \mathcal{L}$, $p \in P$, such that $p \in l$.

By $I(n,m)$ denote the largest $I(\mathcal{L}, P)$ among all pairs $(\mathcal{L}, P)$ provided that $|\mathcal{L}| = n$, $|P| = m$. (Here and below the symbol $|A|$ denotes the cardinality of the set A.)

Main question: How to estimate $I(n, m)$?

P. Erdos conjectured: there exists a constant C such that

$$I(n, n) \leqslant C\left(n^{4/3}\right).$$

Notice that this bound is evidently better than the trivial one $I(n, n) \leqslant n^2$.

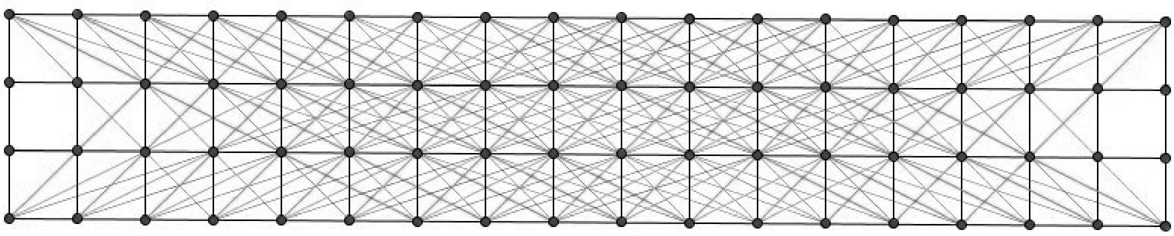

Figure 2: Many point-line incidences

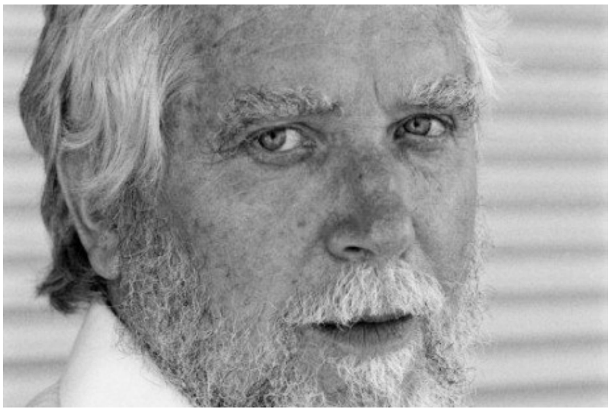

Figure 3: Endre Szemeredi

Figure 4: William Trotter

In 1983, Szemeredi and Trotter managed to prove the Erdos conjecture. This result is called the **Szemeredi–Trotter theorem**.

Our main aim is to prove it using two different approaches. Along the way, we will discover some useful tricks and handle a lot of geometric and combinatorial problems.

1 Combinatorics

1.1 Introduction. Some Combinatorial Problems in Geometry.

1.1. Consider a finite set A of points in the plane. Prove that there exists a line, which divides the plane into two half-planes such that each open half-plane contains at most $\lfloor |A|/2 \rfloor$ points of A.

1.2. Consider $2n$ points in the plane. Prove that the points can be divided into n pairs in such a way that there is no intersection between line segments joining points in pairs.

1.3. Consider n points in the plane so that no three points lie on a line. Construct a non self-intersecting polygonal line joining the given points.

1.4. Consider n points and n pairwise non-parallel straight lines in the plane. Prove that points and lines can be enumerated from 1 to n in such a way that any two segments of perpendiculars from the corresponding points onto the lines don't intersect each other.

1.5. Consider a non-convex polygon $A_1A_2\dots A_n$. Suppose two non-adjacent vertices A_i and A_j satisfy the following property: the polygon lies completely on one closed half-plane bounded by the line A_iA_j; then we reflect either the polyline $(A_i\dots A_j)$ or $(A_j\dots A_i)$ through the center of the segment A_iA_j. Prove that applying a finite number of reflections we must come to a convex polygon.

1.2 Incidence Problems

1.6. Let $\mathcal{M} = \{M_1, \ldots, M_s\}, \quad M_i \subset \{1, 2, \ldots, n\}, |M_i| = 3$, be an arbitrary collection of subsets of n-set. Suppose $|M_i \cap M_j| \neq 1$; then find the largest possible s.

1.7. Let $\mathcal{M} = \{M_1, \ldots, M_s\}, \quad M_i \subset \{1, 2, \ldots, n\}, |M_i| = 4$, be an arbitrary collection of subsets of n-set such that $|M_i \cap M_j| \neq 2$. Denote by $S(n)$ the largest possible s. Prove that $\lfloor n/4 \rfloor^2 \leqslant S(n) \leqslant n(n-1)/4$.

Definition. Let $\mathcal{L} = \{l_1, \ldots, l_n\}, \quad l_i \in P$, be an arbitrary collection of subsets of m-set $P = \{p_1, \ldots p_m\}$. A pair (l_i, p_k) is said to be an *incidence* if $p_k \in l_i$. Denote by $I(\mathcal{L}, P)$ the number of incidences formed by the elements of $\mathcal{L}$ and P. Define $I(n, m) = \max\limits_{|\mathcal{L}|=n,\, |P|=m} I(\mathcal{L}, P)$.

Our goal is "how to estimate $I(n, m)$ when P and $\mathcal{L}$ satisfy certain properties". Otherwise, we have only a trivial bound $I(\mathcal{L}, P) \leqslant mn$.

A typical interpretation of incidences is as follows:

Consider a blank matrix X with n rows and m columns. Rows correspond to elements of the set $\{l_1, \ldots, l_n\}$, and columns correspond to elements of the set $\{p_1, \ldots p_m\}$. We will put a star in the entry $x_{i,j}$ iff $p_j \in l_i$. Now we can reformulate the problem: "how many stars are there in the matrix?"

Use this interpretation if it is convenient. (see problem **1.11.**)

Further, it is supposed that $n, m, r \in \mathbb{N}$, and $\mathcal{L} = \{l_1, \ldots, l_n\}$ is an arbitrary collection of subsets of m-set $P = \{p_1, \ldots p_m\}$ such that $|l_i \cap l_j| \leqslant r$ for any $i \neq j$.

1.8. Let $r = 1$. Prove that

(a) $I(\mathcal{L}, P) \leqslant n^2 + m$, $I(\mathcal{L}, P) \leqslant m^2 + n$.

(b) $I(\mathcal{L}, P) \leqslant \sqrt{m(n^2 - n)} + m$, $I(n, m) \leqslant \sqrt{n(m^2 - m)} + n$.

1.9. Prove that

$$I(\mathcal{L}, P) \leqslant \sqrt{mr(n^2 - n)} + m,\ I(\mathcal{L}, P) \leqslant \sqrt{nr(m^2 - m)} + n.$$

1.10. Let $r = 1$.

(a) Find $\max I(\mathcal{L}, P)$ if $n \leqslant 3$.

(b) Find $\max I(\mathcal{L}, P)$ if $m \geqslant C_n^2$. Find a configuration such for which the maximal bound is obtained.

1.11. Let 13×13 square be divided into the unit squares. Some centers of unit squares are marked in such a way that there is no rectangle with vertices being in marked points, sides of which are parallel to the sides of the square. Find the largest possible number of marked points.

1.12. 100 (a) little mouses are nibbling 1000 (b) pieces of cheese. Each mouse eat some pieces making holes. But any two mouses leave holes in at most 10 (c) common pieces of cheese.

(a) Prove that the number of holes is at most 11000 ($b + a\sqrt{bc}$).

(b) Prove that the number of holes is at most 10500 $\left(\frac{b+\sqrt{b^2+4bca(a-1)}}{2}\right)$.

1.3 Ham Sandwich Theorem.

1.13. Prove that for a finite set of points in general position in the plane each colored "red" or "blue", there is a line that simultaneously bisects the red points and bisects the blue points. ("Bisecting" means that each side contains less than half of the total number of points.)

1.14. Prove that for a finite set of points in general position in space each colored "red", "blue" or "green", there is a plane that simultaneously bisects each coloured group, that is, the number of points of any color on either side of the plane is less than half of the total number of points.

1.15. Consider $2m$ points, in general position in the plane, such that m points are "red" and the others are "blue". Prove that the points can be divided into n pairs in such a way that each pair consists of a red point and a blue point, and no two segments joining points in pairs intersect each other.

By $\mathbb{R} = \mathbb{R}^1$, $\mathbb{R}^2$ and $\mathbb{R}^3$ denote, respectively, a line, a plane and an 3-dimensional space. A point of $\mathbb{R} = \mathbb{R}^1$ is a real number, a point of $\mathbb{R}^2$ is a pair $\mathbf{x} = (x_1, x_2)$, where $x_1, x_2 \in \mathbb{R}$, and a point of $\mathbb{R}^3$ is a vector $\mathbf{x} = (x_1, x_2, x_3)$, where $x_1, x_2, x_3 \in \mathbb{R}$. An element of $\mathbb{R}^n$ is a point (a vector) $\mathbf{x} = (x_1, \dots, x_n)$, where $x_i \in \mathbb{R}$, $i = 1,\dots,n$ (x_i is a coordinate).

A hyperplane h of $\mathbb{R}^d$ is a set of points $\mathbf{x} = (x_1, \dots, x_d)$ that can be described with a single linear equation of the following form (where at least one of the a_i's is non-zero):

$$\sum_{i=1}^{d} a_i x_i = a_0.$$

Checking understanding. What does mean a hyperplane in $\mathbb{R}^2$ and $\mathbb{R}^3$?

We will say that a hyperplane h *bisects* a finite set A if neither of the two open half-spaces bounded by h contains more than $\lfloor |A|/2 \rfloor$ points of A.

Ham Sandwich Theorem. Every d finite sets $A_1, \dots, A_d \subset \mathbb{R}^d$ can be simultaneously bisected by a hyperplane.

The ham sandwich theorem takes its name from the case when $d = 3$ and the three objects of any shape are a chunk of ham, a piece of cheese and a chunk of bread — notionally, a sandwich — which can then all be simultaneously bisected with a single cut (i.e., a plane).

1.16. Consider $3n$ points in general position in the space such that n points are "red", other n points are "blue" and the others are "green". Prove that

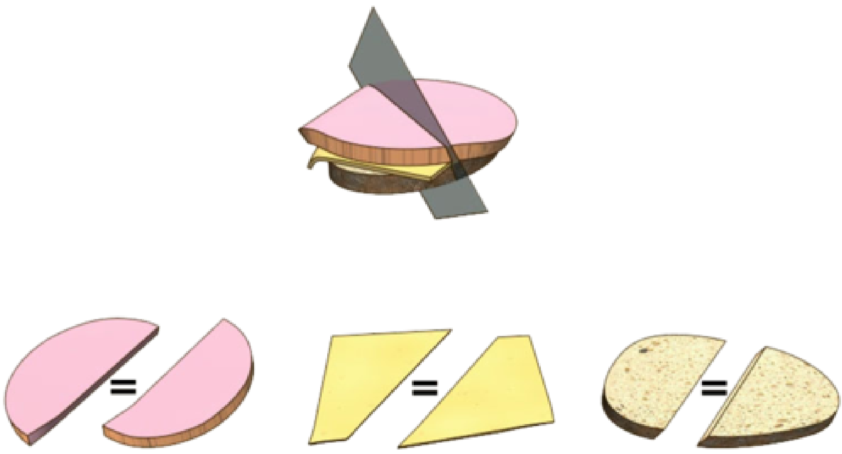

Figure 5: Sandwich cutting

the points can be portioned into “rainbow” 3-tuples in such a way that no two triangles with vertices on points of triples intersect each other.

1.17.* Two thieves have stolen a precious necklace of nearly immeasurable value, not only because of the precious stones (diamonds, sapphires, rubies, etc.), but also because these are set in pure platinum. The thieves do not know the values of the stones of various kinds, and so they want to divide the stones of each kind evenly. In order to waste as little platinum as possible, they want to achieve this by as few cuts as possible. We assume that the necklace is open (with two ends) and that there are d different kinds of stones, an even number of each kind. How many cuts are necessary to do so?

(a) $(d-1)$ cuts?

(b) d cuts?

Comment. Solve exercises on the Veronese map for better understanding of this problem.

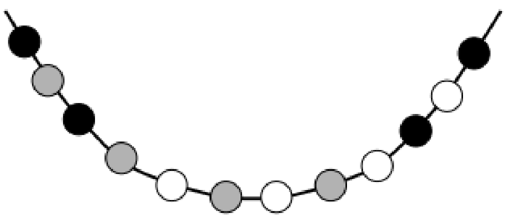

Figure 6: For the necklace here, 3 cuts should suffice

2 Geometric Constructions

We will say that a set P of n points and a set $\mathcal{L}$ of n lines in the plane form a *configuration* n_d if, for any line of $\mathcal{L}$, there exist exactly d points of P lying on this line, and, for any point of P, there exist exactly d lines of $\mathcal{L}$ passing through this point. This implies that $I(\mathcal{L},P) = nd$

2.1. Construct an example of configuration 9_3.

Hint: This is a well-known geometric theorem.

2.2. Construct an example of configuration 9_3 that is different from the previous one.

2.3. Construct an example of configuration 10_3.

Hint: This is a well-known geometric theorem.

2.4. Construct an example of configuration 10_3 that is different from the previous one.

2.5.* Prove that there exist a constant factor C such that for any N we can find a set $\mathcal{L}$ of $n, n > N$, lines and a set P of $n, n > N$, points with $I(\mathcal{L},P) > Cn^{4/3}$.

Hint: Look at the figure 2 on the page 1.

3 Algebraic Geometry & the polynomial method

Let f and g be two functions defined on some subset of the real numbers. One writes $g(x) = O(f(x))$ if and only if there exists a positive constant $C > 0$ such that $|g(x)| \leqslant Cf(x)$ for all x.

Given a polynomial $f(x, y) = \sum_{i,j\leqslant 0} a_{i,j}x^iy^j$. The *z*ero set Z_f of $f(x, y)$ is the subset of points (x, y) in $\mathbb{R}^2$ on which $f(x, y) = 0$. The polynomial $f(x,y)$ is of degree p ($\deg f = p$) if $p = \max\{i + j \mid a_{i,j} \neq 0\}$.

A polynomial $f(x,y)$ is called a *monomial* if $a_{i,j} = 1$ for exactly one pair (i,j) and $a_{i,j} = 0$ otherwise.

3.1. Consider $f(x, y)$ of degree d and an arbitrary line l. Prove that either $l \in Z_f$, or $|l \cap Z_f| \leqslant D$.

3.2. Consider a polynomial $f(x, y)$ of degree d. Prove that the number of distinct lines contained in the set Z_f does not exceed d.

3.3. Show that the number of bivariate monomials of degree at most d equals $\binom{d+2}{2}$.

Hint. In other words, we must find the number of pairs (i,j) of nonnegative integers with $i + j \leqslant D$.

Let P be a set of n points in the plane, and let r be a parameter, $1 < r \leqslant n$. We say that $f(x, y)$ is an *r-partitioning polynomial* for P if no connected component of $\mathbb{R}^2 \setminus Z_f$ contains more than n/r points of A.

3.4. For any $r > 1$, every finite point set admits an r-partitioning polynomial of degree at most r.

Definition. Given an arbitrary integer d, and let $D = \binom{d+2}{2} - 1$. A map $\varphi : \mathbb{R}^2 \to \mathbb{R}^D$ is said to be *the Veronese map of degree* d if this map is given by the following formula:

$$\varphi(x,y) := (x^iy^j)_{(i,j)|1\leqslant i+j\leqslant d} \in \mathbb{R}^D.$$

(We think of the coordinates in $\mathbb{R}^l$ as indexed by pairs (i,j) with $1 \leqslant i+j \leqslant D$.)

Note that we can apply the Veronese map of degree 2 to prove that any 5 points in the plane such that no 4 points are collinear uniquely determines a conic passing through the given points (A *c*onic is a zero set of a bivariate polynomial of degree 2). In order to check it, we should consider an appropriate hyperplane containing the images of the given points under the Veronese map (we have D=5 in the considered case, and one can show that there exists a unique hyperplane on which all five images of the points lye). It is easy to check that the conic in the plane which corresponds the hyperplane is required.

3.5. Given finite sets $A_1, \dots, A_l$. Let k be an integer such that $\binom{k+2}{2} - 1 \geqslant l$. Prove that there exists a polynomial of degree at most k that is an 2-partitioning polynomial for every one of these sets.

3.6. For every $r > 1$, every finite point set $P \subset \mathbb{R}^2$ admits an r-partitioning polynomial f of degree at most $O(\sqrt{r})$.

4 The First Proof of the Szemeredi-Trotter Theorem

4.1. Prove that $I(m,n) = I(n,m)$.

In what follows, we assume $|L| = |P| = n$. First of all, we construct an r-partitioning polynomial $f(x,y)$ for the given set P.

By $L_0 \subset L$ denote a subset of lines $l \in L$, $Z_l \subset Z_f$, by $P_0 \subset P$ denote a set of points $p \in P \cap Z_f$. Suppose Z_f bisects the plane into s parts. By P_i denote points of P lying in the i-th part of the plane, and by L_i denote lines of $\mathcal{L}$ passing through the i-th part of the plane.

4.2. Prove that there exist constants C_1, C_2, C_3 such that

a) $I(L_0, P_0) \leqslant C_1 n\sqrt{r}$;

b) $I(L \setminus L_0, P_0) \leqslant C_2 n\sqrt{r}$;

c) $\sum\limits_{i=1}^{s} I(L_i, P_i) \leqslant C_3 n^2/r$.

4.3. Having chosen a certain r, prove the Szemeredi-Trotter theorem.

4.4. Prove the the Szemeredi-Trotter theorem in the general case:

Szemeredi-Trotter theorem. $I(n,m) = O((nm)^{2/3} + n + m)$.

In order to check this, you have to obtain bounds, similar to those in **4.2.**

5 Applications of the Szemeredi-Trotter Theorem

5.1. (a) Given a finite set P, $|P| = m$, in the plane. Prove that the number of lines passing through k distinct points of P is at most $C(m^2/k^3 + m/k)$, where C is some constant, which does not depend on m, k.

(b) Prove that such lines define at most $D(m^2/k^2 + m)$ incidences with P, where D is some constant, which does not depend on m, k.

5.2. Let P be a set of points in the plane and let L be the set of lines containing at least 2 points in P. Then there exist such $c_1, c_2 > 0$ that one of these two cases must hold:

1. There exists a line in L that contains at least $c_1|P|$ points.
2. $|L| \geqslant c_2|P|^2$.

5.3. Let $A \subset \mathbb{R}$ be a finite set. Then there exists $c > 0$ such that

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

Here

$$A + A = \{a_1 + a_2 : a_1 \in A, a_2 \in A\}, A \cdot A = \{a_1 \dot{a}_2 : a_1 \in A, a_2 \in A\}.$$