Study texts for

DIMATIA - DIMACS
Research Experience for Undergraduates

July 15. – August 5. 1999

Prague

Supported by NSF - AV ČR grant
KONTAKT ME 337

and

GAČR 201/99/0242

Piercing and selection theorems in convexity (lecture notes)

Jiří Matoušek

Department of Applied Mathematics
Charles University
Malostranské nám. 25, 118 00 Praha 1
Czech Republic

Rev. Jan/8/1999

We will discuss several more advanced theorems concerning families of simplices or convex sets in $\mathbf{R}^d$. Some of them can be seen as generalizations of Helly's theorem, asserting that under suitable conditions, a "large" subfamily of a given family has a nonempty intersection. Others claim that a given family can be pierced by few points, where "piercing" means that each set of the family contains at least one of the points.

Readers who don't like higher dimensions may want to consider the dimensions 2 and 3 only. The results are usually no less interesting there.

Glossary of some theorems needed in the sequel.

Separation Theorem: Any two disjoint convex sets $C, D \subset \mathbf{R}^d$ can be (non-strictly) separated by a hyperplane; that is, there is $a \in \mathbf{R}^d \setminus \{0\}$ and $b \in \mathbf{R}$ such that $\langle a, x\rangle \leq b$ for all $x \in C$ and $\langle a, x\rangle \geq b$ for all $x \in D$. If C is compact and D is closed then the separation can be required to be strict (strict inequalities).

Helly's Theorem: If $C_1, C_2, \ldots, C_n$ is a collection of convex sets in $\mathbf{R}^d$ such that any at most $d+1$ of them intersect, then all of them intersect.

Radon's Theorem: If A is a set of $d+2$ points in $\mathbf{R}^d$, then there are disjoint subsets $A_1, A_2 \subset A$ such that $\mathrm{conv}(A_1) \cap \mathrm{conv}(A_2) \neq \emptyset$.

Carathéodory's Theorem: If $X \subseteq \mathbf{R}^d$ and $x \in \mathrm{conv}(X)$ then $x \in \mathrm{conv}(Y)$ for some at most $(d+1)$-point $Y \subseteq X$.

Ham-Sandwich Cut Theorem: Any d finite sets in $\mathbf{R}^d$ can be simultaneously bisected by a hyperplane. A hyperplane h *bisects* a finite set A if each of the open halfspaces defined by h contains at most $\lfloor |A|/2 \rfloor$ points of A.

1 Fractional Helly Theorem

Helly's Theorem says that if any at most $d+1$ sets of a finite family of convex sets in $\mathbf{R}^d$ intersect then all the sets of the family intersect. What if not necessarily all, but a large

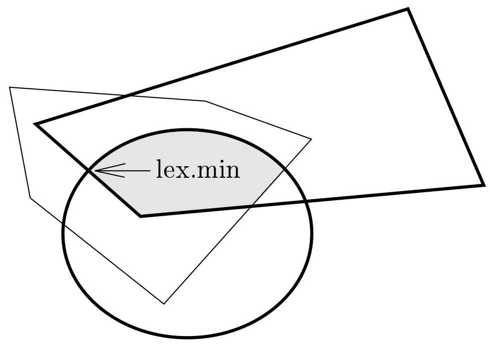


Figure 1: The lexicographically smallest point in the intersection of convex sets, and 2 sets determining it.

fraction of $(d+1)$-tuples of sets intersect? The following theorem says that then a large fraction of the sets must have a point in common.

1.1 Theorem (Fractional Helly Theorem). *For any $\alpha > 0$ there exists a $\beta = \beta(d, \alpha) > 0$ with the following property. Let $F_1, \ldots, F_n$ be convex sets in $\mathbf{R}^d$, $n \geq d+1$, and suppose that for at least $\alpha\binom{n}{d+1}$ of the $(d+1)$-point index sets $I \subseteq \{1, 2, \ldots, n\}$, we have $\bigcap_{i \in I} F_i \neq \emptyset$. Then there exists a point contained in at least βn sets among the F_i.*

This result holds with $\beta = \frac{\alpha}{d+1}$.

Proof. For a subset $I \subseteq \{1, 2, \ldots, n\}$, let us write F_I for the intersection $\bigcap_{i \in I} F_i$.

First we observe that it is enough to prove this theorem for the case when the F_i are closed and bounded (in fact, even polytopes). Indeed, given some arbitrary $F_1, \ldots, F_n$, we take all $(d+1)$-tuples I of indices for which the intersection F_I is nonempty, and we fix some point $p_I \in F_I$. Define $F_i' = \operatorname{conv}\{p_I \colon i \in I\}$; this is a polytope contained in F_i, and it is easy to check that if the theorem holds for these F_i' then it also holds for the original F_i. So, in the rest of the proof, we assume the F_i, and hence also all the nonempty F_I, are compact.

Let $\leq_{lex}$ denote the *lexicographic ordering* of points of $\mathbf{R}^d$ (by their coordinate vectors). It is easy to show that any compact subset of $\mathbf{R}^d$ has a unique lexicographically minimum point. We need the following consequence of Helly's Theorem.

1.2 Lemma. *Let $I \subseteq \{1, 2, \ldots, n\}$ be an index set with $F_I \neq \emptyset$, and let v be the (unique) lexicographically minimum point of F_I. Then there is an at most d-element subset $J \subseteq I$ such that v is the lexicographically minimum point of F_J as well. (In other words, the minimum is always determined by at most d "constraints" F_i; see Fig. 1.)*

Proof. Define a set $C \subset \mathbf{R}^d$, $C = \{x \in \mathbf{R}^d \colon x <_{lex} v\}$. It is easy to check that C is convex. Since v is the lexicographic minimum of F_I, we have $C \cap F_I = \emptyset$. We have a family of convex sets, consisting of C plus the sets F_i with $i \in I$, with empty intersection. Hence by Helly's theorem there are at most $d+1$ sets in this family whose intersection is empty as well. The set C must be one of these sets since all the others contain v. The remaining at most d sets yield the desired index set J. □

We can now finish the proof of the Fractional Helly Theorem. For each of the $\alpha\binom{n}{d+1}$ index sets I of cardinality $d+1$ with $F_I \neq \emptyset$, we fix a d-element set $J = J(I) \subset I$ such that F_J has the same lexicographic minimum as F_I.

The theorem now follows by a double counting. Since we have at most $\binom{n}{d}$ d-tuples J, one of them, J_0, appears as $J(I)$ for at least $\alpha\binom{n}{d+1}/\binom{n}{d} = \alpha(n-d)/(d+1)$ distinct index $(d+1)$-tuples I. Each such I has the form $J_0 \cup \{i\}$ for some $i \in \{1, 2, \ldots, n\}$. The lexicographic minimum of F_{J_0} is thus contained in at least $d + \alpha(n-d)/(d+1) \geq \alpha n/(d+1)$ sets among the F_i. Hence we may set $\beta = \frac{\alpha}{d+1}$. □

Bibliography and Remarks. The Fractional Helly Theorem is due to Katchalski and Liu [KL79].

A quantitatively sharper version was proved by Kalai [Kal84], in which $\beta \to 1$ for $\alpha \to 1$. The best possible value is
$$\beta = 1 - (1-\alpha)^{1/(d+1)}.$$

Exercises

1. Prove the following *Colored Helly Theorem* (due to Lovász): *Let $\mathcal{C}_1, \ldots, \mathcal{C}_{d+1}$ be finite families of convex sets in $\mathbf{R}^d$ such that for any choice of sets $C_1 \in \mathcal{C}_1, \ldots, C_{d+1} \in \mathcal{C}_{d+1}$, the intersection $C_1 \cap \ldots \cap C_{d+1} \neq \emptyset$. Then for some i, all the sets of $\mathcal{C}_i$ have a nonempty intersection.* Apply a method similar to the proof of Fractional Helly, i.e. consider the lexicographic minima of the intersections of suitable collections of the sets.

2. Show that the number β in Theorem 1.1 cannot in general be taken smaller than $1-(1-\alpha)^{1/(d+1)}$.

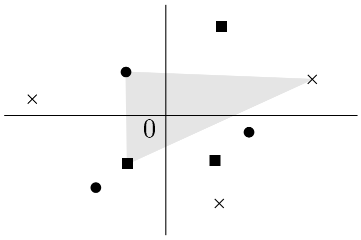
3. Show that any compact set in $\mathbf{R}^d$ has a unique point with lexicographically smallest coordinate vector.

2 Colored Carathéodory Theorem and Tverberg's Theorem

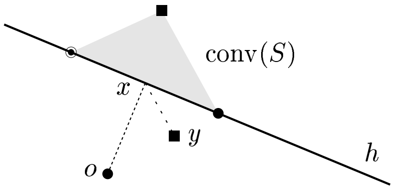
Here we present a "colored version" of Carathéodory's Theorem. In order to have a unified notation with the proof of Tverberg's Theorem below, we will work in space of dimension N (instead of the usual d) for a while.

2.1 Theorem (Colored Carathéodory Theorem). *Consider $N+1$ finite point sets $M_1, \ldots, M_{N+1}$ in $\mathbf{R}^N$ such that the convex hull of each M_i contains the point 0 (the origin). Then there exists an $(N+1)$-point set $S \subseteq M_1 \cup \ldots \cup M_{N+1}$ with $|M_i \cap S| = 1$ for each i and such that $0 \in \operatorname{conv}(S)$. (If we imagine that the points of M_i have "color" i, then we look for a "rainbow" $(N+1)$-point S with $0 \in \operatorname{conv}(S)$, where "rainbow"="containing all colors.)*

Here is a planar illustration:



Proof. Call the convex hull of an $(N+1)$-point rainbow set a *rainbow simplex*. We proceed by contradiction; suppose that no rainbow simplex contains the origin, and choose a rainbow set S such that the distance of $\mathrm{conv}(S)$ to the origin is the smallest possible. Let x be the point of $\mathrm{conv}(S)$ closest to 0. Consider the hyperplane h containing x and perpendicular to the segment $0x$, as in the picture:



Then all of S lies in the closed halfspace h^- bounded by h and not containing 0. We have $\mathrm{conv}(S) \cap h = \mathrm{conv}(h \cap S)$, any by Carathéodory's Theorem, there exists an at most N-point subset $T \subseteq S \cap h$ such that $x \in \mathrm{conv}(T)$.

Let i be a color not occurring in T (i.e. $M_i \cap T = \emptyset$). If all the points of M_i lied in the halfspace h^-, then 0 wouldn't lie in $\mathrm{conv}(M_i)$ (which we assume). Thus, there exists a point $y \in M_i$ lying in the complement of h^- (strictly, i.e. $y \notin h$).

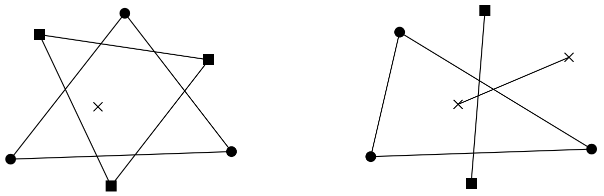
Let us form a new rainbow set S' from S, by replacing the (unique) point of $M_i \cap S$ by y. We have $T \subset S'$, and so $x \in \mathrm{conv}(S')$. Hence the segment xy is contained in $\mathrm{conv}(S')$, and so $\mathrm{conv}(S')$ also lies closer to 0 than $\mathrm{conv}(S)$ — a contradiction. $\square$

This proof suggests an algorithm for finding the rainbow simplex as in the theorem. Namely, start with an arbitrary rainbow simplex, and if it doesn't contain 0, switch one vertex as in the proof. It is not known whether the number of steps of this algorithm can be bounded by a polynomial function of N. It would be very interesting to construct configurations where the number of steps is very large.

Tverberg's Theorem generalizes Radon's Theorem. It is a nice example of a nontrivial result about combinatorial properties of finite sets in $\mathbf{R}^d$, and a building block for other results.

2.2 Theorem (Tverberg's Theorem). *Let d, r be given natural numbers. Let us put $N = (d+1)(r-1)$. Then for any $(N+1)$-point set $A \subset \mathbf{R}^d$ there exist r disjoint subsets $A_1, A_2, \ldots, A_r \subseteq A$ such that $\bigcap_{i=1}^r \mathrm{conv}(A_i) \neq \emptyset$.*

The sets $A_1, A_2, \ldots, A_r$ as in the theorem are called a *Tverberg partition* of A (we may assume that they form a partition of A), and a point in the intersection of their convex hulls is called a *Tverberg point*. The following illustration shows what can such a partition look like for $d = 2$, $r = 3$.



(Are these all Tverberg partitions for this set, or are there more?)

Again, a very interesting open problem is the existence of an efficient algorithm for finding a Tverberg partition for a given set. There is a polynomial-time algorithm if the dimension is fixed, but some NP-hardness results for closely related problems indicate that perhaps this problem might be algorithmically difficult if d is part of input.

Proof of Tverberg's Theorem. We begin with a re-formulation of Tverberg's Theorem, which will be technically easier to handle. For a set $X \subseteq \mathbf{R}^d$, the *convex cone generated by* X is defined as the set of all linear combinations of points of X with nonnegative coefficients; that is, we set

$$\operatorname{cone}(X) = \left\{ \sum_{i=1}^{n} \alpha_i x_i \colon x_1, \ldots, x_n \in X,\ \alpha_1, \ldots, \alpha_n \in \mathbf{R},\ \alpha_i \geq 0 \right\}.$$

Geometrically, $\operatorname{cone}(X)$ is the union of all semilines starting in the origin and passing through a point of $\operatorname{conv}(X)$. The following statement is easily checked to be equivalent to Tverberg's Theorem:

2.3 Proposition (Tverberg's Theorem — "cone" version). *Let $A \subset \mathbf{R}^{d+1}$ be an $N+1$-point set such that $0 \notin \operatorname{conv}(A)$ (where d, r, and N are as in Tverberg's Theorem). Then there exist r disjoint subsets $A_1, A_2, \ldots, A_r \subseteq A$ such that $\bigcap_{i=1}^{r} \operatorname{cone}(A_i) \neq \{0\}$.*

Let us verify that this proposition implies Tverberg's Theorem. Embed $\mathbf{R}^d$ into $\mathbf{R}^{d+1}$ as the hyperplane $x_{d+1} = 1$ (this embedding is frequently used if one needs to pass from an "affine" formulation to a "linear" formulation of a problem). An $(N+1)$-point set $A \subset \mathbf{R}^d$ thus becomes a subset of $\mathbf{R}^{d+1}$; moreover, its convex hull lies in the $x_{d+1} = 1$ hyperplane and thus it doesn't contain 0. By Proposition 2.3, the set A can be partitioned into groups $A_1, \ldots, A_r$ with $\bigcap_{i=1}^{r} \operatorname{cone}(A_i) \neq \{0\}$. The intersection of these cones thus contains a semiline originating in 0. It is easily checked that such a semiline intersects the hyperplane $x_{d+1} = 1$ and that the intersection point is a Tverberg point for A. Hence it suffices to prove Proposition 2.3.

Proof of Proposition 2.3. First we define r linear maps $\varphi_j : \mathbf{R}^{d+1} \to \mathbf{R}^N$. The image space has $N = (d+1)(r-1)$ coordinates; we group them into $r-1$ groups by $d+1$ coordinates each. For $j = 1, 2, \ldots, r-1$, $\varphi_j(x)$ is the vector having the coordinates of x in the jth group and zeros in the other groups; symbolically

$$\varphi_j(x) = (\underbrace{0 \,|\, 0 \,|\, \cdots \,|\, 0}_{(j-1)\times} \,|\, x \,|\, 0 \,|\, \cdots \,|\, 0\,).$$

The last mapping, φ_r, has $-x$ in all groups: $\varphi_r(x) = (-x \,|\, -x \,|\, \cdots \,|\, -x\,)$.

We note the following property of these maps: For any r vectors $u_1, \ldots, u_r \in \mathbf{R}^{d+1}$ we have

$$\sum_{j=1}^{r} \varphi_j(u_j) = 0 \text{ holds if and only if } u_1 = u_2 = \cdots = u_r\,. \tag{1}$$

Indeed, this can be easily seen by expressing

$$\sum_{j=1}^{r} \varphi_j(u_j) = (u_1 - u_r \,|\, u_2 - u_r \,|\, \cdots \,|\, u_{r-1} - u_r)\,.$$

Next, let $A = \{a_1, \ldots, a_{N+1}\} \subset \mathbf{R}^{d+1}$ be a set as in Proposition 2.3. We consider the set $M = \varphi_1(A) \cup \varphi_2(A) \cup \ldots \cup \varphi_r(A)$ in $\mathbf{R}^N$ consisting of r copies of A. The first $r-1$ copies are placed into mutually orthogonal subspaces of $\mathbf{R}^N$, and the last copy of each a_i sums up to 0 with the other $r-1$ copies of a_i. Then we color the points of M by $N+1$ colors; all copies of the same a_i get the color i. In other words, we set $M_i = \{\varphi_1(a_i), \varphi_2(a_i), \ldots, \varphi_r(a_i)\}$. As we have noted, the points in each M_i sum up to 0, which means that $0 \in \operatorname{conv}(M_i)$ and thus the assumptions of the Colored Carathéodory Theorem 2.1 hold for $M_1, \ldots, M_{N+1}$.

Let $S \subseteq M$ be a rainbow set (containing one point of each M_i) with $0 \in \operatorname{conv}(S)$. For each i, let $j(i)$ be the index of the point of M_i contained in S; that is, we have $S = \big\{\varphi_{j(1)}(a_1), \varphi_{j(2)}(a_2), \ldots, \varphi_{j(N+1)}(a_{N+1})\big\}$. Then $0 \in \operatorname{conv}(S)$ means that

$$\sum_{i=1}^{N} \alpha_i \varphi_{j(i)}(a_i) = 0$$

for some nonnegative reals $\alpha_1, \ldots, \alpha_{N+1}$ summing up to 1. Let I_j be the set of indices i with $j(i) = j$, and set $A_j = \{a_i \colon i \in I_j\}$. The above sum can be rearranged to

$$\sum_{j=1}^{r} \sum_{i \in I_j} \alpha_i \varphi_j(a_i) = \sum_{j=1}^{r} \varphi_j\Big(\sum_{i \in I_j} \alpha_i a_i\Big)$$

(the last equality follows from the linearity of each φ_j). Write $u_j = \sum_{i \in I_j} \alpha_i a_i$. This is a linear combination of points of A_j with nonnegative coefficients, hence $u_j \in \operatorname{cone}(A_j)$. Above we have derived $\sum_{j=1}^{r} \varphi_j(u_j) = 0$, and so by the condition (1) we get $u_1 = u_2 = \ldots = u_r$. Hence the common value of all the u_j's is a common point of all the cones $\operatorname{cone}(A_j)$.

It remains to check that $u_j \neq 0$. Since we assume $0 \notin \operatorname{conv}(A)$, the only nonnegative linear combination of points of A equal to 0 is the trivial one, with all coefficients 0. On the other hand, since not all α_i are 0, at least one u_j is expressed as a nontrivial linear combination of points of A. This proves Proposition 2.3. $\square$

Bibliography and Remarks. The Colored Carathéodory Theorem is due to Bárány [Bár82]. Tverberg's Theorem was proved by Tverberg (really!) [Tve66]. His original proof was quite complicated. The idea is simple, though: start with some point configuration for which the theorems is valid, and convert it to a given configuration by moving one point at a time. During the movement, the current partition may stop working at some point, and it must be shown that it can be replaced by another suitable partition.

Later on, Tverberg found a simpler proof [Tve81]. The main idea of the proof presented in the text above is due to Sarkaria [Sar92], and our presentation is based on a simplification by Onn (see [BO97]). The perhaps simplest proof, also due to Tverberg (published in a paper by

Tverberg and Vrećica [TV93]) and inspired by the proof of the Colored Carathéodory theorem, goes roughly as follows. Let $\pi = (A_1, A_2, \ldots, A_r)$ be some partition of the given $N+1$ points into r disjoint nonempty subsets. Consider a ball intersecting all the sets $\mathrm{conv}(A_j)$, $j = 1, 2, \ldots, r$, whose radius $\rho = \rho(\pi)$ is the smallest possible. By a suitable general position assumption, it can be assured that the smallest ball is always unique for any partition. If $\rho(\pi) = 0$ then π is a Tverberg partition. Supposing that $\rho(\pi) > 0$, it can be shown that π can be locally changed (by moving one point) to another partition π' with $\rho(\pi') < \rho(\pi)$. The reader may want to try to complete this proof.

Algorithmic aspects of the Colored Carathéodory Theorem were investigated by Bárány and Onn [BO97].

A conjecture of Sierksma says that the number of Tverberg partitions for a set of $(r-1)(d+1)+1$ points in $\mathbf{R}^d$ in general position is at least $((r-1)!)^d$. A lower bound of $\frac{1}{(r-1)!}\left(\frac{r}{2}\right)^{((r-1)(d+1))/2}$, provided that $r \geq 3$ is a prime number, was proved by Vućić and Živaljević [VŽ93] by an ingenious topological argument.

Exercises

1. Prove that Tverberg's Theorem 2.2 implies Proposition 2.3. Why is the assumption $0 \notin \mathrm{conv}(A)$ necessary in Proposition 2.3?
2. Show that for any $d, r \geq 1$ there is an $(N+1)$-point set in $\mathbf{R}^d$ in general position, $N = (d+1)(r-1)$, having no more than $((r-1)!)^d$ Tverberg partitions.

3 Selection lemma and weak epsilon-nets

Consider arbitrary n points in the plane (in general position), and draw all the $\binom{n}{3}$ triangles with vertices at the given points. Then there exists a point of the plane common to at least $\frac{2}{9}\binom{n}{3}$ of these triangles. In this section, we prove this result (with a worse value of the constant) and its d-dimensional generalization.

For easier formulations, we introduce the following terminology: If $X \subset \mathbf{R}^d$ is a finite set, an *X-simplex* is the convex hull of some $(d+1)$-tuple of points of X. We make the convention that X-simplices are in bijective correspondence with their vertex sets; this means that two X-simplices determined by two distinct $(d+1)$-point subsets of X are considered different even if they coincide as point sets in $\mathbf{R}^d$. Thus, the X-simplices form a multiset in general (this concerns only sets X in degenerate positions; if X is in general position then distinct $(d+1)$-point sets have distinct convex hulls).

3.1 Theorem (Selection Lemma for All Simplices). *Let X be an n-point set in $\mathbf{R}^d$. Then there exists a point $a \in \mathbf{R}^d$ (not necessarily belonging to X) contained in at least $c_d\binom{n}{d+1}$ X-simplices, where $c_d > 0$ is a constant depending on the dimension d only.*

Proof. We may suppose that n is sufficiently large ($n \geq n_0$ for a given constant n_0); otherwise choosing a point contained in a single X-simplex would do with c_d small enough.

Set $r = \lceil n/(d+1) \rceil$. By Tverberg's Theorem 2.2, there exist r pairwise disjoint sets $M_1, \ldots, M_r \subseteq X$ whose convex hulls all have a point in common; call this point a. (A

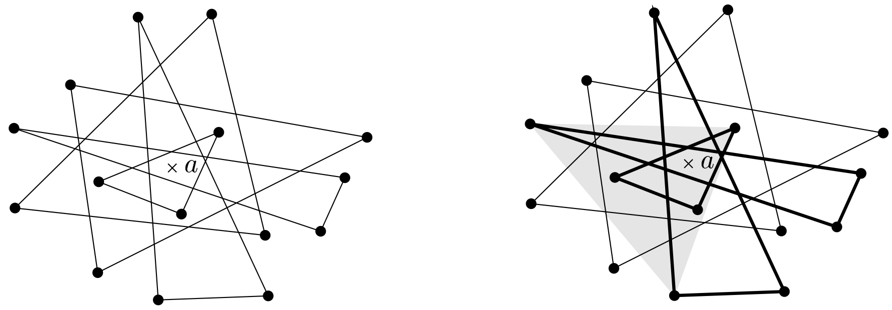


Figure 2: The sets M_i (a), and the selection of an X-simplex containing a using $d+1$ of the M_i (b).

typical M_i has $d+1$ points but some of them may be smaller.) We want show that the point a is contained in many X-simplices; see Fig. 2.

Let $J = \{j_0, \ldots, j_d\} \subseteq \{1, 2, \ldots, r\}$ be a $(d+1)$-point index set. We apply the Colored Carathéodory's Theorem 2.1 for the $(d+1)$ "color" sets $M_{j_0}, \ldots, M_{j_d}$, which all contain a in their convex hull. This yields a rainbow X-simplex S_J containing a and having one vertex from each of the M_{j_i}.

Clearly if $J' \neq J$ are two $(d+1)$-tuples of indices, we have $S_J \neq S_{J'}$. Hence the number of X-simplices containing the point a is at least

$$\binom{r}{d+1} = \binom{\lceil n/(d+1) \rceil}{d+1} \geq \frac{1}{(d+1)^{d+1}} \frac{n(n-(d+1))\ldots(n-d(d+1))}{(d+1)!}.$$

For n sufficiently large, say $n \geq 2d(d+1)$, this is at least $(d+1)^{-(d+1)} 2^{-d} \binom{n}{d+1}$. □

Remark. The best possible value of c_d is not known. The proof just given shows that for n very large, we may take $c_d \approx (d+1)^{-(d+1)}$. In the plane, it is known that $c_2 \geq \frac{2}{9}$; the proof we gave leads to a much worse value, however.

We now pass to the second topic of this section.

3.2 Definition (Weak Epsilon-Net). *Let X be a finite point set in $\mathbf{R}^d$ and $\varepsilon > 0$ a real number. A set $N \subseteq \mathbf{R}^d$ is called* a weak ε-net *for X (with respect to convex sets) if any convex set containing at least $\varepsilon|X|$ points of X contains a point of N. Thus, N has to intersect all "big" convex sets, where the size is measured by the number of points of X.*

A key result about weak ε-nets with respect to convex sets is the following:

3.3 Theorem (Weak ε-Net Theorem). *For any $d \geq 1$, $\varepsilon > 0$, and finite $P \subset \mathbf{R}^d$, there exists a weak ε-net for P with respect to convex sets of size at most $f(d, \varepsilon)$, where $f(d, \varepsilon)$ depends on d and ε but not on P.*

Proof. A simple proof can be given using the Selection Lemma 3.1. Let an $P \subset \mathbf{R}^d$ be an n-point set. The required N is constructed by a greedy algorithm. Set $N_0 = \emptyset$. If N_i

has already been constructed, we look whether there is a convex set C containing at least εn points of P and no point of N_i. If not, N_i is a weak ε-net by definition. If yes, we set $X_i = P \cap C$, and we apply the Selection Lemma 3.1 on X_i. This gives us a point a_i contained in at least $c_d\binom{|X_i|}{d+1} = \Omega(\varepsilon^{d+1}n^{d+1})$ X_i-simplices. We set $N_{i+1} = N_i \cup \{a_i\}$ and continue by the next step of the algorithm.

Altogether there are $\binom{n}{d+1}$ P-simplices. In each step of the algorithm, at least $\Omega(\varepsilon^{d+1}n^{d+1})$ of them are "killed", meaning that they were not intersected by N_i but are intersected by N_{i+1}. Hence the algorithm makes at most $O(\varepsilon^{-(d+1)})$ steps. $\square$

Bibliography and Remarks. The planar Selection Lemma for all simplices, with the best possible constant $\frac{2}{9}$, was proved by Boros and Füredi [BF84]. A generalization to an arbitrary dimension was found by Bárány [Bár82].

Weak ε-nets were introduced by Haussler and Welzl [HW87] in the following more general context. If Y is a (possibly infinite) ground set, $\mathcal{S}$ is a family of subsets of Y, and X is a finite subset of Y, then a set $N \subseteq Y$ is called a *weak ε-net for X* with respect to $\mathcal{S}$ if for any $S \in \mathcal{S}$ such that $|S \cap X| \geq \varepsilon|X|$ we have $S \cap N \neq \emptyset$ (above we had $Y = \mathbf{R}^d$ and $\mathcal{S}$ = all convex sets in $\mathbf{R}^d$). The adjective "weak" refers to the fact that we do not require $N \subseteq X$; if we add this requirement to the definition, we get the notion of the so-called *ε-net* (which we will not need here).

The existence of weak ε-nets with respect to convex sets was proved by Alon, Bárány, Füredi, and Kleitman [ABFK92] by the method shown in the text but with a slight quantitative improvement, achieved by using the Selection Lemma for many simplices 5.1 instead of the Selection Lemma for all simplices.

The minimum worst-case size, in terms of ε, of a weak ε-net with respect to convex sets in $\mathbf{R}^d$ is not known. The best known bound in the plane is $O(\varepsilon^{-2})$ (this was also shown in [ABFK92]; see Exercise 2), and in $\mathbf{R}^d$ it is $O\left(\varepsilon^{-d}\left(\log\frac{1}{\varepsilon}\right)^{b_d}\right)$ for a suitable constant b_d (Chazelle et al. [CEG+95]). But the truth is probably much smaller, maybe around $\varepsilon^{-\lfloor d/2\rfloor}$. Finding good bounds is another very interesting problem in this area.

Exercises

1. Prove the Selection Lemma 3.1 using Tverberg's Theorem 2.2 and the Fractional Helly Theorem 1.1 (instead of the Colored Carathéodory's Theorem). Which proof gives a larger value of c_d?

2. Complete the following sketch of an alternative proof of Theorem 3.3.

   (a) Let X be an n-point set in the plane (assume general position if convenient). Let h be a vertical line with half of the points of X on each side, and let X_1, X_2 be these halves. Let M be the set of all intersections of segments of the form x_1x_2 with h, where $x_1 \in X_1$ and $x_2 \in X_2$. Let N_0 be a weak ε'-net for M (this is a one-dimensional situation!). Recursively construct weak ε''-nets N_1, N_2 for X_1 and X_2, respectively, and set $N = N_0 \cup N_1 \cup N_2$. Show that with a suitable choice of ε' and ε'', N is a weak ε-net for X of size $O(\varepsilon^{-2})$.

   (b) Generalize the proof from (a) to $\mathbf{R}^d$ (use induction on d). Estimate the exponent of ε in the resulting bound on the size of the constructed weak ε-net.

3. The aim of this exercise is to show that if X is a finite set in the plane in convex position, then for any $\varepsilon > 0$ there exists a weak ε-net for X with respect to convex sets of size nearly linear in $\frac{1}{\varepsilon}$.

   (a) Let an n-point convex independent set $X \subset \mathbf{R}^2$ be given, and let $\ell \leq n$ be a parameter. Choose points $p_0, p_1, \ldots, p_{\ell-1}$ of X, appearing in this order around the circumference of $\mathrm{conv}(X)$, in such a way that the set X_i of points of X lying (strictly) between p_{i-1} and p_i has at most n/ℓ points for each i. Construct a weak ε'-net N_i for each X_i (recursively) with $\varepsilon' = \ell\varepsilon/3$, and let M be the set containing the intersection of the segment p_0p_{j-1} with p_jp_i, for all pairs i, j, $1 \leq i < j-1 \leq \ell-2$. Show that the set $N = \{p_0, \ldots, p_{\ell-1}\} \cup N_1 \cup \ldots \cup N_\ell \cup M$ is a weak ε-net for X.

   (b) If $f(\varepsilon)$ denotes the minimum necessary size of a weak ε-net for a finite convex independent point set in the plane, derive a recurrence for $f(\varepsilon)$ using (a) with a suitably chosen ℓ, and prove the bound for $f(\varepsilon) = O\left(\frac{1}{\varepsilon}\left(\log\frac{1}{\varepsilon}\right)^c\right)$. What is the smallest c you can get?

4. In this exercise, we want to show that if X is the vertex set of a regular convex n-gon in the plane, then there exists a weak ε-net for X (with respect to convex sets) of size $O(\frac{1}{\varepsilon})$.

   Suppose X lies on the unit circle u centered at o. For an arc length $\alpha \leq \pi$ radians, let $r(\alpha)$ be the radius of the circle centered at o and touching a chord of u connecting two points on u at arc distance α. For $i = 0, 1, 2, \ldots$, let N_i be a set of $\lfloor\frac{100}{\varepsilon(1.01)^i}\rfloor$ points placed at regular intervals on the circle of radius $r(\varepsilon(1.01)^i/10)$ centered at o (we take only those i for which this is well-defined). Show that o plus the union of the N_i, $i = 0, 1, 2, \ldots$, form a weak ε-net of size $O(\frac{1}{\varepsilon})$ for X (the constants 1.01 etc. are rather arbitrary and can be greatly improved).

4 The Hadwiger–Debrunner (p, q)-problem

Let $\mathcal{F}$ be a finite family of convex sets in the plane. We know that if every 3 sets from $\mathcal{F}$ intersect, then all sets of $\mathcal{F}$ intersect. What if we only know that out of every 4 sets of $\mathcal{F}$, there are some 3 that intersect? Let us say that $\mathcal{F}$ satisfies the $(4,3)$-*condition*. In such case, $\mathcal{F}$ may consist, for instance, of $n-1$ sets sharing a common point and one extra set lying somewhere far away from the others. So we cannot hope for a common intersection of all sets. But can all the sets of $\mathcal{F}$ be pierced by a bounded number of points? That is, does there exist a constant C such that for any family $\mathcal{F}$ of convex sets in $\mathbf{R}^2$ satisfying the $(4,3)$-condition there are at most C points such that each set of $\mathcal{F}$ contains at least one of them?

This is the simplest nontrivial case of the so-called (p,q)-problem raised by Hadwiger and Debrunner. This problem was open for a very long time, despite of a considerable effort and many results for special families of sets, and it has been solved quite recently by Alon and Kleitman.

4.1 Theorem (The (p,q)-Theorem). *Let p, q, d be integers with $p \geq q \geq d+1 > 2$. Then there exists a constant $C = C(d,p,q)$ such that the following is true: Let $\mathcal{F}$ be a finite family of convex sets in $\mathbf{R}^d$ satisfying the (p,q)-condition; that is, among any p sets of $\mathcal{F}$, there are q sets with a common intersection. Then $\mathcal{F}$ can be pierced by at most C points.*

Remarks. Clearly, the condition $q \geq d+1$ is necessary, since n hyperplanes in general position in $\mathbf{R}^d$ satisfy the (d,d)-condition but cannot be pierced by any bounded number of points independent of n.

It has been known for a long time that if $p(d-1) < (q-1)d$, then $C(d,p,q)$ exists and equals $p-q+1$. This is the only case when exact values, or even good estimates, of $C(d,p,q)$ are known.

The Alon–Kleitman proof consists of three main steps, and it combines an amazing number of tools. By now, we have met them all: the Fractional Helly Theorem, the Separation Theorem, and weak ε-nets. A unsatisfactory feature of this method is that the resulting estimates for $C(d,p,q)$ are enormously large, while the truth seems to be much smaller.

Step I: Fractional Helly. We may assume that $q = d+1$ (since we cannot get good bounds for the $C(d,p,q)$ anyway and so we are only interested in their existence, this is sufficient for the proof). A first observation is that if $\mathcal{F}$ satisfies the (p,q)-condition, then many q-tuples of sets of $\mathcal{F}$ intersect. This can be seen by a simple double counting. Every p-tuple of sets of $\mathcal{F}$ contains (at least) one intersecting q-tuple, and a single q-tuple is contained in $\binom{n-q}{p-q}$ p-tuples (where $n = |\mathcal{F}|$), therefore there are at least

$$\frac{\binom{n}{p}}{\binom{n-q}{p-q}} \geq \alpha \binom{n}{q}$$

intersecting q-tuples, with $\alpha > 0$ depending on p, q only. Using this with $q = d+1$, the Fractional Helly Theorem 1.1 implies that at least βn sets of $\mathcal{F}$ have a common point, with $\beta = \beta(\alpha) > 0$ a constant.[1]

For reasons which become apparent later, we will need a *weighted version* of this result:

4.2 Lemma. *Let $\mathcal{F} = \{F_1, \ldots, F_n\}$ be a family of convex sets in $\mathbf{R}^d$ satisfying the $(p, d+1)$-condition, and let $w_1, \ldots, w_n$ be nonnegative real numbers summing up to 1, the* weights *of the sets F_i. Then there exists an index set $I \subseteq \{1, 2, \ldots, n\}$ such that $\bigcap_I F_i \neq \emptyset$, and the total weight of sets F_i with $i \in I$ is at least β, where $\beta > 0$ depends on d and p only.*

Proof. First assume that all the w_i are rational numbers with a common denominator D, so that $n_i = Dw_i$ are integers. Define a new family $\mathcal{F}_w$, which contains n_i copies of F_i, $i = 1, 2, \ldots, n$ (so it is a multiset of sets).

This new family need not satisfy the $(p, d+1)$-condition, since the $(p, d+1)$-condition for $\mathcal{F}$ only speaks of p-tuples of *distinct* sets from $\mathcal{F}$, while a p-tuple of sets from $\mathcal{F}_w$ may contain multiple copies of the same set. But $\mathcal{F}_w$ does satisfy the $(p', d+1)$-condition with $p' = d(p-1)+1$; indeed, a p'-tuple of sets of $\mathcal{F}_w$ contains at least $d+1$ copies of the same set or it contains p distinct sets, an in the latter case the $(p, d+1)$-condition for $\mathcal{F}$ applies. Using the Fractional Helly Theorem (which does not require the sets in the considered family to be distinct), we get that there exists a point a common to at least βD sets of $\mathcal{F}_w$, i.e. the total weight of the sets of $\mathcal{F}$ containing this point is at least β, for some $\beta = \beta(p, d)$.

[1] By removing these βn sets and iterating, we would get that $\mathcal{F}$ can be pierced by $O(\log n)$ points. The whole point is to get rid of this $\log n$ factor.

To handle arbitrary real weights w_i, we first construct a finite set $A = A(\mathcal{F})$ as follows. For every subfamily $\mathcal{G} \subseteq \mathcal{F}$ with a nonempty intersection, we choose a point $a_{\mathcal{G}}$ in that intersection, and we let A be the set of all these points. Clearly, for any collection of rational weights w_i, we may require that the "good" point a common to sets of weight at least β belongs to A. Then if the given vector $w = (w_1, \ldots, w_n)$ of weights has arbitrary real components, we choose a sequence of rational vectors $w^{(j)}$ converging to w; clearly, a point $a \in A$ which is good for infinitely many of the weights $w^{(j)}$ is good for w as well. □

Step II: Duality. Let $A = A(\mathcal{F})$ be the finite set as at the end of the proof of Lemma 4.2 (containing one point from each nonempty intersection).

4.3 Lemma. *Each point $a \in A$ can be assigned a nonnegative real weight u_a, in such a way that $\sum_{a \in A} u_a = 1$ and that for any $F \in \mathcal{F}$, the total weight of points of A lying in F is at least β, where β is the positive constant from Lemma 4.2.*

Proof. We use the Separation Theorem. Set $m = |A|$, and let the coordinates in $\mathbf{R}^m$ be indexed by the points of A. The idea is to consider u as the coefficient vector in the equation of a hyperplane h in $\mathbf{R}^m$, and express the conditions on u as requirement on h to separate suitable convex sets.

For each $F_i \in \mathcal{F}$, let $c^{(i)} \in \mathbf{R}^m$ be the characteristic vector of F_i; i.e. $c_a^{(i)} = 1$ for $a \in F_i$ and 0 otherwise. We are looking for a vector $u \in \mathbf{R}^m$ satisfying $\langle c^{(i)}, u \rangle \geq \beta$ for all $i = 1, \ldots, n$. Moreover, the components of u should be nonnegative and sum up to (at most) 1. This can be rephrased by requiring that $\langle u, x \rangle \leq 1$ for all $x \in (-\infty, 1]^m$, or equivalently $\langle u, x \rangle \leq \beta$ for all $x \in (-\infty, \beta]^m$.

Considering the hyperplane $h = \{x \in \mathbf{R}^m \colon \langle u, x \rangle = \beta\}$, we see that it should separate the convex sets $C = \operatorname{conv}\{c^{(1)}, \ldots, c^{(n)}\}$ and $D = (-\infty, \beta)^m$. So for proving that a suitable h exists, it suffices to show that $C \cap D = \emptyset$, or in other words, that any point $x \in C$ has at least one coordinate at least β.

We can write an $x \in C$ as a convex combination $x = \sum_{i=1}^n w_i c^{(i)}$, for some $w_i \geq 0$ summing up to 1. Lemma 4.2 applied with the weights w_i tells us there exists a point $a \in A$ contained in sets of $\mathcal{F}$ of total weight at least β, i.e. $\sum_{i \colon a \in F_i} w_i \geq \beta$. But the sum on the left-hand side of this inequality is just x_a, the coordinate of x indexed by a. Lemma 4.3 is proved.

There are many ways to phrase this proof; for instance, we could use a suitable version of Farkas' Lemma. □

Let us summarize what we have proved so far. Given a family $\mathcal{F}$ with the $(p, d+1)$-property, we have obtained a finite set A with nonnegative point weights, such that the total weight of A is 1 and the weight of $F \cap A$ for any $F \in \mathcal{F}$ is at least a positive constant β. In fact, A with the weight function defines a probability measure in the plane (concentrated on the finite set A) under which each of the sets of $\mathcal{F}$ is "big".

The reader may want to think of another, somewhat similar case: suppose that each of the sets of $\mathcal{F}$ were contained in the unit square and had area at least $\beta > 0$ (so the probability measure would be the usual Lebesgue measure on the unit square). Can you find a simple construction of a set N that pierces all sets of such an $\mathcal{F}$, and whose size only depends on β?

In the proof, we need to deal with a more or less arbitrary probability measure (with a finite support but this doesn't seem to help much). An appropriate tool here are the weak ε-nets for convex sets.

Step III: Weak epsilon-nets. Having Lemma 4.3 from Step II at our disposal, the rest of the proof is a straightforward application of the existence theorem for weak ε-nets (Theorem 3.3). First, we may assume that the weights u_a of points of A as in Lemma 4.3 are all rational numbers (this is immediate if we relax the requirement in the lemma to $\sum_{a\in A\cap F} u_a \geq \beta/2$ for all $F \in \mathcal{F}$, say, which we can afford). Let D be the common denominator of the rational numbers u_a; we construct a multiset $\tilde{A}$ by taking Du_a copies of each $a \in A$. Set $\varepsilon = \beta/2$, and apply Theorem 3.3 to get a set N, of size depending on ε and d only and thus constant, which is a weak ε-net for $\tilde{A}$ with respect to convex sets (one can easily check that the proof of that theorem goes through for multisets as well). Since each set $F \in \mathcal{F}$ contains at least εD points of $\tilde{A}$, we get that N pierces $\mathcal{F}$, and the proof of the (p,q)-Theorem is finished. □

Bibliography and Remarks. The Hadwiger–Debrunner (p,q)-problem was solved by Alon and Kleitman [AK92]; we also refer to that paper for the history of the problem. Several other related results were proved by Alon and Kalai [AK95] by similar techniques. The application of the Alon–Kleitman technique in Exercise 2 below is due to Alon [Alo98]. Earlier, a similar result with a slightly stronger bound was proved by Kaiser [Kai97] by a topological method (following an initial breakthrough by Tardós [Tar95], who dealt with the case $d = 2$).

Exercises

1. For which values of p and r does the following hold? Let $\mathcal{F}$ be a finite family of convex sets in $\mathbf{R}^d$, and suppose that any subfamily consisting of at most p sets can be pierced by at most r points. Then $\mathcal{F}$ can be pierced by at most C' points, for some $C' = C'(d,p,r)$.

2. Let $\ell_1, \ldots, \ell_d$ be distinct parallel lines in the plane. A *d-interval* is a set of the form $I_1 \cup I_2 \cup \ldots \cup I_d)$, where $I_j \subset \ell_j$ is a closed interval on the jth line. Let $\mathcal{H}$ be a family of n d-intervals containing no $k+1$ pairwise disjoint members.

   (a)* Show that there is a point contained in at least $\frac{n}{2dk}$ members of $\mathcal{H}$.

   (b) Use (a) and an analogue of Step II in the text to show that there exists a finite collection of m points such that any member of $\mathcal{H}$ contains at least $\frac{m}{2d^2k}$ of them.

   (c) Show that $\mathcal{H}$ can be pierced by at most $2d^2k$ points.

5 Selection Lemma for many simplices

In this section, we continue using the term X-simplex in the sense of Section 3, i.e. these are the convex hulls of $(d+1)$-point subsets of X. In that section, we saw that if X is a set in $\mathbf{R}^d$ and we consider *all* the X-simplices, then at least a fixed fraction of them have a point in common (this was what we called the Selection Lemma for all simplices). What if we do not have all, but many X-simplices, some α-fraction of all? It turns out that still many of them must have a point in common; this is the contents of the Selection Lemma below. But the known quantitative bounds are quite weak, and it seems that a radically different proof method would be needed to get a substantially better estimate.

5.1 Theorem (Selection Lemma). *Let X be an n-point set in $\mathbf{R}^d$, and let $\mathcal{S}$ be a family of $\alpha\binom{n}{d+1}$ X-simplices, where $\alpha \in (0,1]$ is a parameter. Then there exists a point contained in at least*

$$c\alpha^{s_d}\binom{n}{d+1}$$

X-simplices from $\mathcal{S}$, where $c = c(d) > 0$ and s_d are constants.

Remarks. For $d = 1$, it is not too difficult to obtain asymptotically sharp quantitative bounds (see Exercise 1). Already for $d = 2$, the result is quite nontrivial, and the best known bound (probably still not sharp) is as follows: If $|\mathcal{S}| = n^{3-\nu}$, then there is a point contained in at least $O(n^{3-3\nu}/\log^5 n)$ X-triangles from $\mathcal{S}$, which in the parameterization as in Theorem 5.1 means $s_2 \approx 3$ provided that α is small enough, say $\alpha \le n^{-\delta}$ for some $\delta > 0$. For higher dimensions, the best known proof gives $s_d \approx (4d+1)^{d+1}$.

Proof of the Selection Lemma. We give a proof using the Fractional Helly Theorem 1.1. We need to show that many $(d+1)$-tuples of X-simplices of $\mathcal{S}$ have nonempty intersections. For brevity, let us call a $(d+1)$-tuple with a nonempty intersection *good*. First, let us concentrate on the simpler task of exhibiting at least one good $(d+1)$-tuple; even this seems quite nontrivial. The only known proof (for an arbitrary dimension) is based on the following theorem:

5.2 Theorem (Colored Tverberg Theorem). *For any integers $r, d > 1$ there exists an integer t, such that given any $t(d+1)$-point set $Y \subset \mathbf{R}^d$ partitioned into $d+1$ t-point classes $Y_1, \ldots, Y_{d+1}$, there exist r vertex-disjoint Y-simplices $S_1, \ldots, S_r$, such that each S_i has exactly one vertex of each Y_j, $j = 1, 2, \ldots, d+1$ (that is, the S_i are rainbow), and all of $S_1, \ldots, S_r$ have a point in common.*

In this result, one can take $t \le 4r - 3$.

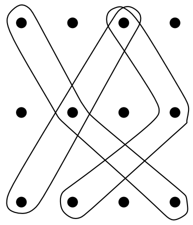
The following drawing shows a 3-colored set in the plane, and the 3 rainbow triangles as in the Colored Tverberg Theorem:

(In the plane, it is known that $t = r$ works.)

The only known proof of the Colored Tverberg Theorem is topological, and we will not discuss it here.

Assuming this result, let us continue the proof of the Selection Lemma 5.1. We can view $\mathcal{S}$ as a $(d+1)$-uniform hypergraph. That is, we regard X as a vertex set, and each X-simplex corresponds to a hyperedge, i.e. a subset of X of size $d+1$. This hypergraph captures the "combinatorial type" of the family $\mathcal{S}$, and a specific placement of the points of X in $\mathbf{R}^d$ then gives a concrete "geometric realization" of $\mathcal{S}$.

In the hypergraph language, the Colored Tverberg Theorem can be rephrased as follows. Let $K_{d+1}(t,t,\ldots,t)$ denote the complete $(d+1)$-partite $(d+1)$-uniform hypergraph with t vertices in each of its $d+1$ vertex classes (the illustration shows a $K_3(4,4,4)$; only three edges are shown as a sample although, of course, all triples connecting vertices at different levels are present).



Then, whenever the vertex set Y of $K_{d+1}(t,t,\ldots,t)$ is placed in $\mathbf{R}^d$, there exists a r-tuple of vertex-disjoint Y-simplices sharing a common point, and all the vertex sets of the Y-simplices in this r-tuple are edges of $K_{d+1}(t,t,\ldots,t)$.

A well-known result of Erdős says that if we have a $(d+1)$-uniform hypergraph on n vertices with sufficiently many edges, then it has to contain a (non-induced) copy of $K_{d+1}(t,t,\ldots,t)$. (We do not formulate this result precisely as we will need a stronger one later. The simplest case of this statement is the well-known fact that any graph on n vertices with more than roughly $n^{3/2}$ edges contains a $K_{2,2}$.) In our situation, we choose $r=d+1$ (then $t=4d+1$ suffices in the Colored Tverberg Theorem). If the given family $\mathcal{S}$ of X-simplices is sufficiently large, the corresponding $(d+1)$-uniform hypergraph contains a copy of $K_{d+1}(t,t,\ldots,t)$, and by the Colored Tverberg Theorem, such a copy contributes a good $(d+1)$-tuple of X-simplices of $\mathcal{S}$.

For Fractional Helly, we need not only one but many good $(d+1)$-tuples. We use an appropriate stronger hypergraph result, saying that if a hypergraph has enough edges then it contains many copies of $K_{d+1}(t,t,\ldots,t)$:

5.3 Theorem (Erdős–Simonovits Theorem). *Let d,t be positive integers. Let $\mathcal{H}$ be a $(d+1)$-uniform hypergraph on n vertices and with $\alpha\binom{n}{d+1}$ edges, where $\alpha \geq Cn^{-1/t^d}$ for a certain sufficiently large constant C. Then $\mathcal{H}$ contains at least*

$$c\alpha^{t^{d+1}} n^{(d+1)t}$$

copies of $K_{d+1}(t,t,\ldots,t)$ (where $c=c(d,t)>0$ is a constant).

For completeness, a proof is given at the end of this section.

Note that, in particular, the theorem implies that a $(d+1)$-uniform hypergraph having at least a constant fraction of all possible edges contains at least a constant fraction of all possible copies of $K_{d+1}(t,t,\ldots,t)$.

We can now finish the proof of the Selection Lemma 5.1 by a simple double counting. The given family $\mathcal{S}$, viewed as a $(d+1)$-uniform hypergraph, has $\alpha\binom{n}{d+1}$ edges, and thus it contains at least $c\alpha^{t^{d+1}}n^{(d+1)t}$ copies of $K_{d+1}(t,t,\ldots,t)$. Each such copy contributes one good $(d+1)$-tuple of vertex-disjoint X-simplices of $\mathcal{S}$. On the other hand, $d+1$ vertex-disjoint X-simplices have together $(d+1)^2$ vertices, and hence their vertex set can be extended to a vertex set of some $K_{d+1}(t,t,\ldots,t)$ (which has $t(d+1)$ vertices) in at most $n^{(t-d-1)(d+1)}$ ways. This is also the maximum number of copies of $K_{d+1}(t,t,\ldots,t)$ giving rise to the same good $(d+1)$-tuple. Hence there are at least $c\alpha^{t^{d+1}}n^{(d+1)^2}$ good $(d+1)$-tuples of X-simplices of $\mathcal{S}$. Applying the Fractional Helly Theorem 1.1 on $\mathcal{S}$, we get that at least $c'\alpha^{t^{d+1}}n^{d+1}$ X-simplices of $\mathcal{S}$ share a common point, with $c'=c'(d)>0$. This proves the Selection Lemma 5.1, with the exponent $s=(4d+1)^{d+1}$. □

The relation of the Colored Tverberg Theorem to the Selection Lemma in this proof somewhat resembles deriving macroscopic properties in physics (pressure, temperature, etc.) from microscopic properties (laws of motion applied on the molecules, say). From the information about small (microscopic) configurations, we obtained a global (macroscopic) result, saying that a significant portion of the X-simplices of $\mathcal{S}$ have a common point.

Proof of Theorem 5.3. First we recall that for any convex function $f:\mathbf{R}\to\mathbf{R}$ and any reals $x_1,\ldots,x_n$, we have

$$\sum_{i=1}^{n} f(x_i) \geq n\, f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right). \tag{2}$$

By induction on d, we are going to show that a k-uniform hypergraph on n vertices and with m edges contains at least $f_k(n,m)$ copies of $K_k(t,\ldots,t)$, where

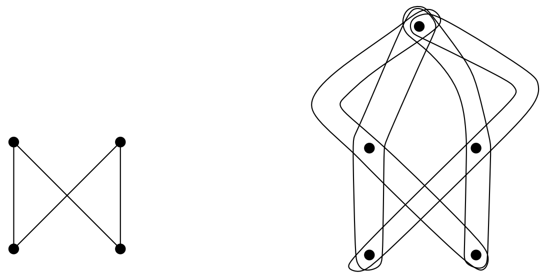
$$f_k(n,m) = c_k n^{tk}\left(\frac{m}{n^k}\right)^{t^k} - C_k n^{t(k-1)},$$

with $c_k>0$, C_k some suitable constants (depending on k and also on t; t is not shown in the notation since it remains fixed). This claim implies the theorem.

For $k=1$, this claim obviously holds.

So let $\mathcal{H}$ be k-uniform with vertex set V, $|V|=n$, and edge set E, $|E|=m$. For a vertex $v\in V$, define a $(k-1)$-uniform hypergraph $\mathcal{H}_v$ on V, whose edges are all edges of $\mathcal{H}$ that contain v, but with v deleted; that is, $\mathcal{H}_v=(V,\{e\setminus\{v\}\colon e\in E, v\in e\})$. Further let $\mathcal{H}'$ be the $(k-1)$-uniform hypergraph whose edge set is the union of edge sets of all $\mathcal{H}_v$.

Let $\mathcal{K}$ denote the set of all copies of the complete $(k-1)$-partite hypergraph $K_{k-1}(t,\ldots,t)$ in $\mathcal{H}'$. The key notion in the proof is that of an *extending vertex* for a copy $K\in\mathcal{K}$: a vertex $v\in V$ is extending for a $K\in\mathcal{K}$ if K is contained in $\mathcal{H}_v$, or in other words, if for each edge e of K, $e\dot{\cup}\{v\}$ is an edge in $\mathcal{H}$. The picture below shows a $K_2(2,2)$ and an extending vertex for it (in a 3-regular hypergraph).



The idea is to count the number of all pairs (K, v), where v is an extending vertex of K, in two ways.

On the one hand, if a fixed copy $K \in \mathcal{K}$ has q_K extending vertices, then it contributes $\binom{q_K}{t}$ distinct copies of $K_k(t, \ldots, t)$ in $\mathcal{H}$. We note that one copy of $K_k(t, \ldots, t)$ comes from at most $O(1)$ distinct $K \in \mathcal{K}$ in this way, and therefore it suffices to lower-bound $\sum_{K\in\mathcal{K}} \binom{q_K}{t}$.

On the other hand, for a fixed vertex v, the hypergraph $\mathcal{H}_v$ contains at least $f_{k-1}(n, m_v)$ copies $K \in \mathcal{K}$ by the inductive assumption, where m_v is the number of edges of $\mathcal{H}_v$. Hence

$$\sum_{K\in\mathcal{K}} q_K \geq \sum_{v\in V} f_{k-1}(n, m_v)\,,$$

and using $\sum_{v\in V} m_v = km$, the convexity of f_{k-1} in the second variable, and (2), we get

$$\sum_{K\in\mathcal{K}} q_K \geq n\, f_{k-1}(n, km/n)\,. \tag{3}$$

To conclude the proof, define a convex function extending the binomial coefficient $\binom{x}{t}$ to the whole domain $\mathbf{R}$:

$$g(x) = \begin{cases} 0 & \text{for } x \leq t-1 \\ \frac{x(x-1)\ldots(x-t+1)}{t!} & \text{for } x > t-1\,. \end{cases}$$

We want to lower-bound $\sum_{K\in\mathcal{K}} g(q_K)$, and we have the bound (3) for $\sum_{K\in\mathcal{K}} q_K$. Using the bound $|\mathcal{K}| \leq n^{t(k-1)}$ (clear since $K_{k-1}(t, \ldots, t)$ has $t(k-1)$ vertices) and (2), we get that the number of copies of $K_k(t, \ldots, t)$ in $\mathcal{H}$ is at least

$$cn^{t(k-1)} g\left(\frac{n\, f_{k-1}(n, km/n)}{n^{t(k-1)}}\right)\,.$$

Calculation finishes the induction step; we omit the details. $\square$

Bibliography and Remarks. The Selection Lemma for many simplices was conjectured, and proved in the planar case, by Bárány, Füredi, and Lovász [BFL90]. They were motivated by the *k-set problem* (how many k-element subsets can be cut off by a halfspace for an n-point set in $\mathbf{R}^d$?). They also proved the planar case of the Colored Tverberg Theorem and conjectured the general case. This was established by Živaljević and Vrećica [ŽV92]; simplified proofs were given later by Björner, Lovász, Živaljević, and Vrećica [BLŽV94] and by Matoušek [Mat96] (using a method of Sarkaria). All these proofs are topological.

The best possible value of t in the Colored Tverberg Theorem is not known. The topological proofs show that if r is a prime, then $t = 2r - 1$ suffices. Recently, this was extended to all prime powers r by Živaljević [Živ] (a similar approach in a different problem was used earlier by Özaydin, by Sarkaria, and by Volovikov). For $d = 2$, Bárány and Larman [BL92] proved that $t = r$ suffices for any r.

We outline a beautiful topological proof, due to Lovász, showing that for $r = 2$, one can take $t = 2$ for all d (this proof is reproduced in [BL92]). Let X be the surface of the $(d+1)$-dimensional *cross-polytope*, i.e. of the convex hull of $V = \{e_1, -e_1, e_2, -e_2, \ldots, e_{d+1}, -e_{d+1}\}$, where $e_1, e_2, \ldots, e_{d+1}$ is the standard orthonormal basis in $\mathbf{R}^{d+1}$. Note that X consists of 2^{d+1} simplices of dimension d, each of them being the convex hull of $d+1$ points of V. Let $Y_i = \{u_i, v_i\} \subset \mathbf{R}^d$, $i = 1, 2, \ldots, d+1$, be the given two-point color classes. Define the mapping $f\colon V \to \mathbf{R}^d$ by setting $f(e_i) = u_i$, $f(-e_i) = v_i$. This mapping has a unique extension $\bar{f}\colon X \to \mathbf{R}^d$

such that $\bar{f}$ is linear on each of the d-dimensional simplices mentioned above. This $\bar{f}$ is a continuous mapping of X into $\mathbf{R}^d$. Since X is homeomorphic to the d-dimensional sprere S^d, a version of the *Borsuk–Ulam Theorem* guarantees that there is $x \in X$ such that $\bar{f}(x) = \bar{f}(-x)$. If $V_1 \subset V$ is the vertex set of a d-dimensional simplex containing x, then $V_1 \cap (-V_1) = \emptyset$, $-x \in \mathrm{conv}(-V_1)$, and as is easy to check, $S_1 = f(V_1)$ and $S_2 = f(-V_1)$ are vertex sets of intersecting rainbow simplices ($\bar{f}(x) = \bar{f}(-x)$ is a common point).

Theorem 5.3 is from Erdős and Simonovits [ES83].

Exercises

1. (a) Prove a one-dimensional Selection lemma: Given an n-point set $X \subset \mathbf{R}$ and a family $\mathcal{S}$ of $\alpha\binom{n}{2}$ X-intervals, there exists a point common to $\Omega(\alpha^2\binom{n}{2})$ intervals of $\mathcal{S}$. What is the best value of the constant of proportionality you can get?

   (b) Show that this result is sharp (up to the value of the multiplicative constant) in the full range of α.

2. (a) Show that the exponent s_2 in the Selection Lemma 5.1 in the plane cannot be smaller than 2.

   (b) Show that also $s_3 \geq 2$. Can you also show $s_d \geq 2$?

   (c) Show that the proof method via the Fractional Helly Theorem cannot give a better value of s_2 than 3 in the Selection Lemma 5.1. That is, construct an n point set and $\alpha\binom{n}{3}$ triangles on it, in such a way that no more than $O(\alpha^5 n^9)$ triples of these triangles have a point in common.

6 A hypergraph regularity lemma

Here we consider a tool from the theory of hypergraphs. It is a result inspired by the famous *Szemerédi Regularity Lemma* for graphs. Very roughly speaking, the Szemerédi Regularity Lemma says that for given $\varepsilon > 0$, the vertex set of any sufficiently large graph G can be partitioned into some number, not too small and not too large, of parts in such a way that the bipartite graphs between "most" pairs of the parts look like random bipartite graphs, up to an "error" bounded by ε. An exact formulation is rather complicated. The result discussed here is a hypergraph analogue of a simple consequence of the Szemerédi Regularity Lemma. It is considerably easier to prove than the Szemerédi Regularity Lemma.

Let $\mathcal{H} = (X, E)$ be a k-partite hypergraph whose vertex set is the union of k pairwise disjoint n-element sets $X_1, X_2, \ldots, X_k$, and whose edges are k-tuples containing precisely one element from each X_i. For subsets $Y_i \subseteq X_i$, $i = 1, 2, , \ldots, k$, let $e(Y_1, \ldots, Y_k)$ denote the number of edges of $\mathcal{H}$ contained in $Y_1 \cup \ldots \cup Y_k$. In this notation, the total number of edges of $\mathcal{H}$ is equal to $e(X_1, \ldots, X_k)$. Further, let

$$\rho(Y_1, \ldots, Y_k) = \frac{e(Y_1, \ldots, Y_k)}{|Y_1| \cdot |Y_2| \cdot \ldots \cdot |Y_k|}$$

denote the *density* of the subhypergraph induced by the Y_i.

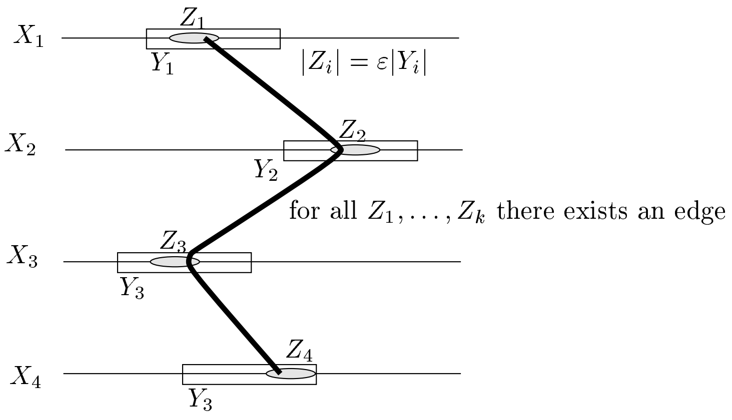
6.1 Theorem (Weak Regularity Lemma (for hypergraphs)). *Let $\mathcal{H}$ be a k-partite hypergraph as above, and suppose that $\rho(\mathcal{H}) \geq \beta$ for some $\beta > 0$. Let $0 < \varepsilon < 1/2$. Suppose that n is sufficiently large in terms of β and ε.*

Then there exist subsets $Y_i \subseteq X_i$ of equal size $|Y_i| = s \geq \beta^{1/\varepsilon^k} n$, $i = 1, 2, \ldots, k$, such that

(i) (High density) $\rho(Y_1, \ldots, Y_k) \geq \beta$, and

(ii) (Edges on all large subsets) $e(Z_1, \ldots, Z_k) > 0$ for any $Z_i \subseteq Y_i$ with $|Z_i| \geq \varepsilon s$, $i = 1, 2, \ldots, k$.

The following scheme illustrates the situation (but, of course, the vertices of the Y_i and Z_i need not be contiguous).



Proof. Intuitively, the sets Y_i should be selected in such a way that the subhypergraph induced by them is as dense as possible. We then want to show that if there were $Z_1, \ldots, Z_k$ of size at least εs with no edges on them, we could replace the Y_i by sets with a still larger density. But if we looked at the usual density $\rho(Y_1, \ldots, Y_k)$, we would typically get too small sets. The trick is to look at a modified density parameter which slightly favors larger sets. Thus, we define the *magical density* $\mu(Y_1, \ldots, Y_k)$ by

$$\mu(Y_1, \ldots, Y_k) = \frac{e(Y_1, \ldots, Y_k)}{(|Y_1| \cdot |Y_2| \cdot \ldots \cdot |Y_k|)^{1-\varepsilon^k/k}}.$$

We choose $Y_1, \ldots, Y_k$, $Y_i \subseteq X_i$, as sets of *equal size* that have the maximum possible magical density $\mu(Y_1, \ldots, Y_k)$. We denote the common size $|Y_1| = \cdots = |Y_k|$ by s.

First we derive the condition (i) in the theorem for this choice of the Y_i. We have

$$\frac{e(Y_1, \ldots, Y_k)}{s^{k-\varepsilon^k}} = \mu(Y_1, \ldots, Y_k) \geq \mu(X_1, \ldots, X_k) = \beta n^{\varepsilon^k} \geq \beta s^{\varepsilon^k},$$

and so $e(Y_1, \ldots, Y_k) \geq \beta s^k$, which verifies (i). Since obviously $e(Y_1, \ldots, Y_k) \leq s^k$, we have $\mu(Y_1, \ldots, Y_k) \leq s^{\varepsilon^k}$. Combining with $\mu(Y_1, \ldots, Y_k) \geq \beta n^{\varepsilon^k}$ derived above, we also obtain that $s \geq \beta^{1/\varepsilon^k} n$.

It remains to prove (ii). Since εs is a large number by the assumptions, rounding it up to an integer doesn't matter in the subsequent calculations (as can be checked by a simple but somewhat tedious analysis). In order to simplify matters, we will thus assume that

εs is an integer, and we let $Z_1 \subseteq Y_1, \ldots, Z_k \subseteq Y_k$ be εs-element sets. We want to prove $e(Z_1, \ldots, Z_k) > 0$. We have

$$\begin{aligned} e(Z_1, \ldots, Z_k) = e(Y_1, \ldots, Y_k) & \\ - e(Y_1 \setminus Z_1, Y_2, Y_3, \ldots, Y_k) & \\ - e(Z_1, Y_2 \setminus Z_2, Y_3, \ldots, Y_k) & \\ - e(Z_1, Z_2, Y_3 \setminus Z_3, \ldots, Y_k) & \\ \vdots \quad & \\ - e(Z_1, Z_2, Z_3, \ldots, Y_k \setminus Z_k). & \end{aligned} \tag{4}$$

We want to show that the negative terms are not too large, using the assumption that the magical density of $Y_1, \ldots, Y_k$ is maximum. The problem is that $Y_1, \ldots, Y_k$ only maximize the magical density among the sets of equal size, while we have sets of different sizes in the terms. To get back to sets of equal size, we use the following observation. If, say, R_1 is a randomly chosen subset of Y_1 of some given size r, we have

$$\mathbf{E}\left[\rho(R_1, Y_2, \ldots, Y_k)\right] = \rho(Y_1, \ldots, Y_k),$$

where $\mathbf{E}[.]$ denotes the expectation with respect to the random choice of r-element $R_1 \subseteq Y_1$. This preservation of density by choosing a random subset is quite intuitive, and it is not difficult to verify it by counting (Exercise 1). For estimating the term $e(Y_1 \setminus Z_1, Y_2, \ldots, Y_k)$, we use random subsets $R_2, \ldots, R_k$ of size $(1-\varepsilon)s$ of $Y_2, \ldots, Y_k$, respectively. Thus,

$$e(Y_1 \setminus Z_1, Y_2, \ldots, Y_k) = (1-\varepsilon)s^k \mathbf{E}\left[\rho(Y_1 \setminus Z_1, R_2, \ldots, R_k)\right].$$

Now for any choice of $R_2, \ldots, R_k$, we have

$$\begin{aligned} \rho(Y_1 \setminus Z_1, R_2, \ldots, R_k) &= ((1-\varepsilon)s)^{-\varepsilon^k} \mu(Y_1 \setminus Z_1, R_2, \ldots, R_k) \\ &\leq ((1-\varepsilon)s)^{-\varepsilon^k} \mu(Y_1, Y_2, \ldots, Y_k) \\ &= (1-\varepsilon)^{-\varepsilon^k} \rho(Y_1, \ldots, Y_k). \end{aligned}$$

Therefore, we obtain

$$e(Y_1 \setminus Z_1, Y_2, \ldots, Y_k) \leq (1-\varepsilon)^{1-\varepsilon^k} e(Y_1, \ldots, Y_k) \leq (1-\varepsilon) e(Y_1, \ldots, Y_k).$$

To estimate the term $e(Z_1, Z_2, \ldots, Z_{i-1}, Y_i \setminus Z_i, Y_{i+1}, \ldots, Y_k)$, we use random subsets $R_i \subset Y_i \setminus Z_i$ and $R_{i+1} \subset Y_{i+1}, \ldots, R_k \subset Y_k$, this time all of size εs. A similar calculation as before yields

$$e(Z_1, Z_2, \ldots, Z_{i-1}, Y_i \setminus Z_i, Y_{i+1}, \ldots, Y_k) \leq \varepsilon^{i-1-\varepsilon^k}(1-\varepsilon) e(Y_1, \ldots, Y_k).$$

(This estimate is also valid for $i = 1$ but it is worse than the one derived above and it would not suffice in the subsequent calculation.) From (4) we get that $e(Z_1, \ldots, Z_k)$ is at least $e(Y_1, \ldots, Y_k)$ multiplied by the factor

$$1 - (1-\varepsilon) - (1-\varepsilon)\varepsilon^{-\varepsilon^k} \sum_{i=2}^{k} \varepsilon^{i-1} = \varepsilon - \varepsilon^{1-\varepsilon^k}\left(1 - \varepsilon^{k-1}\right)$$

$$
\begin{aligned}
&= \varepsilon \left(1 + \varepsilon^{-\varepsilon^k}(\varepsilon^{k-1} - 1)\right) \\
&= \varepsilon \left(1 + e^{\varepsilon^k \ln(1/\varepsilon)}(\varepsilon^{k-1} - 1)\right) \\
&\geq \varepsilon \left(1 + (1 + \varepsilon^k \ln \tfrac{1}{\varepsilon})(\varepsilon^{k-1} - 1)\right) \\
&= \varepsilon^{k+1} \left(\tfrac{1}{\varepsilon} - \ln \tfrac{1}{\varepsilon} + \varepsilon^k \ln \tfrac{1}{\varepsilon}\right) \\
&\geq \varepsilon^{k+1} \left(\tfrac{1}{\varepsilon} - \ln \tfrac{1}{\varepsilon}\right) \\
&> 0.
\end{aligned}
$$

Theorem 6.1 is proved. □

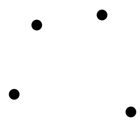
Bibliography and Remarks. The Szemerédi Regularity Lemma is from [Sze78]. Our presentation of Theorem 6.1 essentially follows Pach [Pac98], whose treatment is an adaptation of an approach of Komlós and Sós. A survey of applications and variations of the Szemerédi Regularity Lemma can be found in [KS96].

Exercises

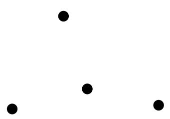
1. Verify the equality $\mathbf{E}\left[\rho(R_1, Y_2, \ldots, Y_k)\right] = \rho(Y_1, \ldots, Y_k)$, where the expectation is with respect to a random choice of an r-element $R_1 \subseteq Y_1$. Also derive the other similar equalities used in the proof.

7 Same Type Lemma with applications

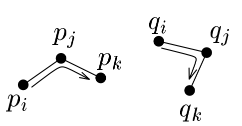
Order type of a set. There are infinitely many 4-point sets in the plane in general position, but there are only two "combinatorially distinct" types of such sets:



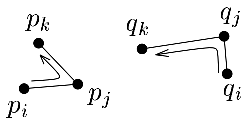
and



What is an appropriate equivalence relation that would capture the intuitive notion of two finite point sets in $\mathbf{R}^d$ being "combinatorially the same"? One suitable notion is based on the *order type* of a configuration. First we explain this notion for planar configurations in general position, where it is quite simple. Let $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ be two sequences of points in $\mathbf{R}^2$, both in general position (no two points coincide and no three are collinear). The n-tuples P and Q have the same order type if, for any three indices $i < j < k$, we turn in the same direction (right or left) when going from p_i to p_k via p_j and when going from q_i to q_k via q_j:



or



We say that both the triples (p_i, p_j, p_k) and (q_i, q_j, q_k) have the same orientation.

If the point sequences P and Q are in $\mathbf{R}^d$, we will require that all the corresponding $(d+1)$-element subsequences have the same orientation. The notion of orientation is best explained for d-tuples of vectors in $\mathbf{R}^d$. If $v_1, \ldots, v_d$ are vectors in $\mathbf{R}^d$, there is a unique linear mapping sending the vector e_i of the standard basis of $\mathbf{R}^d$ to v_i, $i = 1, 2, \ldots, d$. The matrix A of this mapping has the vectors $v_1, \ldots, v_d$ as the columns. The orientation of $(v_1, \ldots, v_d)$ is defined as the sign of $\det(A)$; so it can be $+1$ (positive orientation), -1 (negative orientation), or 0 (the vectors are linearly dependent and lie in a $(d-1)$-dimensional linear subspace). For a $(d+1)$-tuple of points $(p_1, p_2, \ldots, p_{d+1})$, we define the orientation to be the orientation of the d vectors $p_2 - p_1, p_3 - p_1, \ldots, p_{d+1} - p_1$.

Returning to the order type, let $P = (p_1, p_2, \ldots, p_n)$ be a point sequence in $\mathbf{R}^d$. The *order type* of P is defined as the mapping assigning to each $(d+1)$-tuple $(i_1, i_2, \ldots, i_{d+1})$ of indices, $1 \leq i_1 < i_2 < \cdots < i_{d+1} \leq n$, the orientation of the $(d+1)$-tuple $(p_{i_1}, p_{i_2}, \ldots, p_{i_{d+1}})$. Thus, the order type of P can be described by a sequence of $+1$s, -1s, and 0s with $\binom{n}{d+1}$ terms.

From the order type of a sequence, one can read off various properties of the sequence, such as general position, convex position, etc.—see Exercise 1.

For point sequences in general position, having the same order type is a good notion of combinatorial equivalence. For degenerate configurations, the order type doesn't take into account the lower-dimensional structure.

Same-type transversals. Let $(Y_1, Y_2, \ldots, Y_m)$ be an m-tuple of finite sets in $\mathbf{R}^d$. By a *transversal* of this m-tuple we mean any m-tuple $(y_1, y_2, \ldots, y_m)$ such that $y_i \in Y_i$ for all i. We say that $(Y_1, Y_2, \ldots, Y_m)$ *has same-type transversals* if all of its transversals have the same order type.

If $(X_1, X_2, \ldots, X_m)$ are very large finite sets such that $X_1 \cup \cdots \cup X_m$ is in general position, we can find not too small subsets $Y_1 \subseteq X_1, \ldots, Y_m \subseteq X_m$ such that $(Y_1, \ldots, Y_m)$ has same-type transversals. To see this, color each transversal of $(X_1, X_2, \ldots, X_m)$ by its order type. This is a coloring of the edges of the complete m-partite hypergraph on $(X_1, \ldots, X_m)$ by no more than 2^m colors. By the appropriate Ramsey-type theorem for multipartite hypergraphs, there are sets $Y_i \subseteq X_i$, not too small, such that all the edges induced by $Y_1 \cup \cdots \cup Y_m$ have the same color, i.e. $(Y_1, \ldots, Y_m)$ has same-type transversals.

As is the case for many other geometric applications of Ramsey-type theorems, this result can be quantitatively improved tremendously by a geometric argument: for m and d fixed, the size of the sets Y_i can be made a constant fraction of $|X_i|$.

7.1 Theorem (Same Type Lemma). *For any natural numbers d and m, there exists $c = c(d, m) > 0$ such that the following holds. Let $X_1, X_2, \ldots, X_m$ be finite sets in $\mathbf{R}^d$ such that $X_1 \cup \cdots \cup X_m$ is in general position. Then there are $Y_1 \subseteq X_1, \ldots, Y_m \subseteq X_m$ such that the m-tuple $(Y_1, Y_2, \ldots, Y_m)$ has same-type transversals and $|Y_i| \geq c|X_i|$ for all $i = 1, 2, \ldots, m$.*

Proof. First we observe that it is sufficient to prove the Same Type Lemma for $m = d+1$. For larger m, we begin with $(X_1, X_2, \ldots, X_m)$ as the current m-tuple of sets. Then we go through all $(d+1)$-tuples $(i_1, i_2, \ldots, i_{d+1})$ of indices, and if $(Z_1, \ldots, Z_m)$ is the current m-tuple of sets, we apply the Same Type Lemma to the $(d+1)$-tuple $(Z_{i_1}, \ldots, Z_{i_{d+1}})$. These sets are replaced by smaller sets $(Z'_1, \ldots, Z'_{d+1})$ such that this $(d+1)$-tuple has same-type transversals. After this step is executed for all $(d+1)$-tuples of indices, the resulting current m-tuple of sets has same-type transversals.

Note that this method gives the rather small lower bound $c(d,m) \geq c(d,d+1)^{\binom{m-1}{d}}$.

To handle the crucial case $m = d+1$, we will use the following criterion for a $(d+1)$-tuple of sets having same-type transversals.

7.2 Lemma. *Let $C_1, C_2, \ldots, C_{d+1} \subseteq \mathbf{R}^d$ be convex sets. The following two conditions are equivalent:*

(i) There is no hyperplane simultaneously intersecting all of $C_1, C_2, \ldots, C_{d+1}$.

(ii) For each nonempty index set $I \subset \{1,2,\ldots,d+1\}$, the sets $\bigcup_{i\in I} C_i$ and $\bigcup_{j\notin I} C_j$ can be strictly separated by a hyperplane.

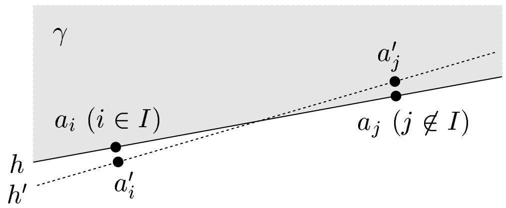
Moreover, if $X_1, X_2, \ldots, X_{d+1} \subset \mathbf{R}^d$ are finite sets such that the sets $C_i = \operatorname{conv}(X_i)$ have the property (i) (and (ii)), then $(X_1, \ldots, X_{d+1})$ has same-type transversals.

The proof of this neat result is left to Exercise 2. We will not need the assertion that (i) implies (ii).

Same Type Lemma for $d+1$ sets. To prove the Same Type Lemma for the case $m = d+1$, it now suffices to choose the sets $Y_i \subseteq X_i$ in such a way that their convex hulls are separated in the sense of (ii) in Lemma 7.2. This can be done by an iterative application of the Ham-Sandwich Theorem.

Suppose that for some nonempty index set $I \subset \{1,2,\ldots,d+1\}$, the sets $\operatorname{conv}(\bigcup_{i\in I} X_i)$ and $\operatorname{conv}(\bigcup_{j\notin I} X_j)$ cannot be separated by a hyperplane. We may assume that $d+1 \in I$. Let h be a hyperplane simultaneously bisecting $X_1, X_2, \ldots, X_d$, whose existence is guaranteed by the Ham-Sandwich Theorem. Let γ be a closed halfspace bounded by h and containing at least half of the points of X_{d+1}. For all $i \in I$, including $i = d+1$, we discard the points of X_i not lying in γ, and for $j \notin I$, we throw away the points of X_j that lie in the interior of γ (note that points on h are never discarded); see Fig. 3.

We claim that union of the resulting sets with indices in I is now strictly separated from the union of the remaining sets. If h contains no points of the sets then it is a separating hyperplane. Otherwise, let the points contained in h be $a_1, \ldots, a_t$; we have $t \leq d$ by the general position assumption. For each a_j, choose a point a_j' very near to a_j. If a_j lies in some X_i with $i \in I$, a_j' is chosen in the complement of γ, and otherwise it is chosen in the interior of γ. We let h' be a hyperplane passing through $a_1', \ldots, a_t'$ and lying very close to h. Then h' is the desired separating hyperplane, provided that the a_j' are sufficiently close to the corresponding a_j, as in the picture below:



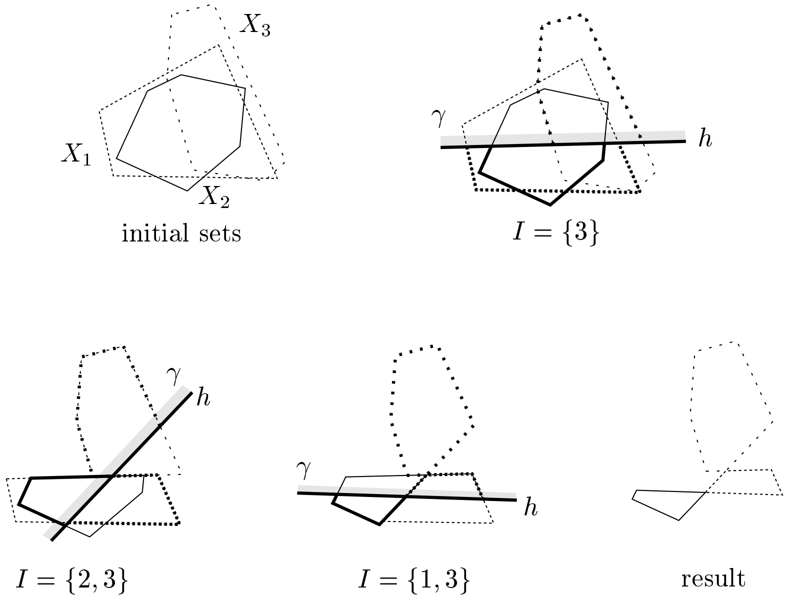


Figure 3: Proof of the Same Type Lemma for $d = 2$, $m = 3$.

Thus, we have "killed" the index set I, at the price of halving the sizes of the current sets; more precisely, the size of a set X_i is reduced from $|X_i|$ to $\lceil |X_i|/2 \rceil$ (or larger). We can continue with other index sets in the same manner. After no more than 2^d halvings, we obtain sets satisfying the separation condition and thus having same-type transversals. The Same Type Lemma is proved. Note that the bound for $c(d, d+1)$ is doubly exponential, roughly 2^{-2^d}. $\square$

A simple application. We recall that by the Erdős–Szekeres Theorem, for any natural number k, there is a natural number $n = n(k)$ such that any n-point set in the plane in general position contains a subset of k points in convex position (forming the vertices of a convex k-gon). The Same Type Lemma immediately gives the following

7.3 Theorem (Positive Fraction Erdős–Szekeres Theorem). *For every integer $k \geq 4$ there is a constant $c_k > 0$ such that every sufficiently large finite set $X \subset \mathbf{R}^2$ contains k disjoint subsets $Y_1, \ldots, Y_k$, of size at least $c|X|$ each, such that each transversal of $(Y_1, \ldots, Y_k)$ is in convex position.*

Proof. Let $n = n(k)$ be the number as in the Erdős–Szekeres Theorem. We partition X into n sets $X_1, \ldots, X_n$ of almost equal sizes, and we apply the Same Type Lemma on them, obtaining sets $Y_1, \ldots, Y_n$, $Y_i \subseteq X_i$, with same-type transversals. Let $(y_1, \ldots, y_n)$ be a transversal of $(Y_1, \ldots, Y_n)$. By the Erdős–Szekeres Theorem, there are $i_1 < i_2 < \cdots < i_k$ such that $y_{i_1}, \ldots, y_{i_k}$ are in convex position. Then $Y_{i_1}, \ldots, Y_{i_k}$ are as required in the theorem. $\square$

Another selection lemma: all simplices on a positive fraction of points. An ingenious application of the Same Type Lemma is another, quite strong "selection lemma." In the plane, given n red points, n white points, and n blue points, we can select $\frac{n}{12}$ red, $\frac{n}{12}$ white, and $\frac{n}{12}$ blue points in such a way that *all* the red-white-blue triangles for the resulting sets have a point in common. Here is the d-dimensional generalization.

7.4 Theorem (Positive Fraction Selection Lemma). *For all natural numbers d, there exists $c = c(d) > 0$ with the following property. Let $X_1, X_2, \ldots, X_{d+1} \subset \mathbf{R}^d$ be finite sets of equal size, with $X_1 \cup X_2 \cup \cdots \cup X_{d+1}$ in general position. Then there is a point $a \in \mathbf{R}^d$ and subsets $Y_1 \subseteq X_1, \ldots, Y_{d+1} \subseteq X_{d+1}$ with $|Y_i| \geq c|X_i|$ such that the convex hull of any transversal of $(Y_1, \ldots, Y_{d+1})$ contains a.*

Note that this result implies the Selection Lemma for all simplices (Theorem 3.1), although currently this gives a much worse value of the constant.

Proof. Let $X = X_1 \cup \cdots \cup X_{d+1}$. We may suppose that all the X_i are large (for otherwise one-point Y_i will do). Let $\mathcal{S}_0$ be the set of all "rainbow" X-simplices, i.e. of all transversals of $(X_1, \ldots, X_{d+1})$ (where the transversals are formally considered as sets for the moment). The size of $\mathcal{S}_0$ is, for d fixed, at least a constant fraction of $\binom{|X|}{d+1}$ (here we use the assumptions that the X_i are of equal size). Therefore, by the Selection Lemma 5.1, there is a subset $\mathcal{S}_1 \subseteq \mathcal{S}_0$ of at least βn^{d+1} X-simplices containing a common point a, where $\beta = \beta(d) > 0$. (Note that we do not need the full power of the Selection Lemma here, since we deal with the complete $(d+1)$-partite hypergraph.)

Next, consider the $(d+1)$-partite hypergraph $\mathcal{H}$ with vertex set X and edge set $\mathcal{S}_1$. We let $\varepsilon = c(d, d+2)$, where $c(d, m)$ is as in the Same Type Lemma, and we apply the Weak Regularity Lemma (Theorem 6.1) to $\mathcal{H}$. This yields sets $Y_1 \subseteq X_1, \ldots, Y_{d+1} \subseteq X_{d+1}$, whose size is at least a fixed fraction of the size of the X_i, and such that any subsets $Z_1 \subseteq Y_1, \ldots, Z_{d+1} \subseteq Y_{d+1}$ of size at least $\varepsilon|Y_i|$ induce an edge, i.e. there is a rainbow X-simplex with vertices in the Z_i and containing the point a.

Now suppose for a moment that the set $X \cup \{a\}$ is in general position. Then the argument is immediately finished by applying the Same Type Lemma with the $d+2$ sets $Y_1, Y_2, \ldots, Y_{d+1}$ and $Y_{d+2} = \{a\}$. We obtain sets $Z_1 \subseteq Y_1, \ldots, Z_{d+1} \subseteq Y_{d+1}$ and $Z_{d+2} = \{a\}$ with same-type transversals, and with $|Z_i| \geq \varepsilon|Y_i|$ for $i = 1, 2, \ldots, d+1$. (Indeed, the Same Type Lemma guarantees that at least one point is selected even from a small set.) Now, either all transversals of $(Z_1, \ldots, Z_{d+1})$ contain the point a in their convex hull or none does (see Exercise 1(d)). But the latter possibility is excluded by the choice of the Y_i (by the Weak Regularity Lemma).

Unfortunately, we have no guarantee that the point a is in general position with respect to the set X, even if X is in general position (in the usual sense, meaning that no $d+1$ points lie on a common hyperplane). For example, it may happen that the intersection of all the X-simplices in $\mathcal{S}_1$ is a single point, and then the position of a is fixed.

There are several ways around this; we present the following simple argument. It can be shown that if X is in general position, then at most $O(n^d)$ X-simplices can have the point a at the boundary (Exercise 3). If we delete the at most $O(n^d)$ simplices with a on the boundary from $\mathcal{S}_1$, we obtain a collection $\mathcal{S}_1'$ of X-simplices that contain a in the interior, and $|\mathcal{S}_1'| \geq \beta' n^{d+1}$ for a suitable $\beta' = \beta'(d) > 0$. With this collection, the argument is finished as before. □

It is amazing how many quite heavy tools are used in this proof. It would be nice to find a more direct argument.

Bibliography and Remarks. For more information on order types, the reader may consult the survey by Goodman and Pollack [GP93]. The Same Type Lemma is from Bárány and Valtr [BV98], and a very similar idea was used by Pach [Pac98]. Bárány and Valtr proved the Positive

Fraction Erdős–Szekeres Theorem 7.3 (the case $k = 4$ was established earlier by Nielsen), and they gave several more applications of the Same Type Lemma, such as a positive fraction Radon theorem and a positive fraction Tverberg theorem.

The planar case of Theorem 7.4 was proved by Bárany, Füredi, and Lovász [BFL90] (with $c(2) \geq \frac{1}{12}$), and the result for arbitrary dimension is due to Pach [Pac98]. The solution to Exercise 3 was communicated to me by János Pach.

The equivalence of (i) and (ii) in Lemma 7.2 is from Goodman, Pollack, and Wenger [GPW96].

Exercises

1. Let $(x_1, x_2, \ldots, x_n)$ be a finite sequence of points in $\mathbf{R}^d$, and suppose that we know its order type. Which of the following properties of the sequence can be determined from the order type?

   (a) No $d+1$ points among $x_1, x_2, \ldots, x_n$ lie on a common hyperplane.

   (b) For any k points among $x_1, x_2, \ldots, x_n$, $k = 2, 3, \ldots, d+1$, the affine hull has the maximum dimension $k-1$.

   (c) The point x_{d+2} lies in the convex hull of $\{x_1, x_2, \ldots, x_{d+1}\}$ (assuming $n \geq d+2$).

   (d) As in (c), but moreover $x_1, \ldots, x_{d+1}$ do not lie in a common hyperplane.

   (e) The points $x_1, \ldots, x_n$ form the set of vertices of a d-dimensional convex polytope.

   (f) None of the x_i lies in the convex hull of the others.

2. (a) Prove that, in the setting of Lemma 7.2, if the convex hulls of the X_i have the property (i) then $(X_1, \ldots, X_{d+1})$ has same-type transversals. Proceed by contradiction.

   (b) Prove that property (ii) (separation) implies property (i) (no hyperplane transversal). Proceed by contradiction and use Radon's Theorem.

   (c) Prove that (i) implies (ii).

3. Let $X \subset \mathbf{R}^d$ be a set of $n \geq d+1$ points in general position (meaning that no $d+1$ points of X lie on a common hyperplane). Let $a \in \mathbf{R}^d$ be an arbitrary point. Prove that at most $O(n^{d-1})$ of the $\binom{n}{d}$ hyperplanes determined by d-tuples the points of X contain a. (Consequently, at most $O(n^d)$ X-simplices have a on their boundary.)

Ackonowledgment

I would like to thank Gregor Baudis and Andreas Marwinski for pointing out several errors in a previous version of these lecture notes.

References

[ABFK92] N. Alon, I. Bárány, Z. Füredi, and D. Kleitman. Point selections and weak ε-nets for convex hulls. *Combin., Probab. Comput.*, 1(3):189–200, 1992.

[AK92] N. Alon and D. Kleitman. Piercing convex sets and the Hadwiger Debrunner (p, q)-problem. *Adv. Math.*, 96(1):103–112, 1992.

[AK95] N. Alon and G. Kalai. Bounding the piercing number. *Discr. Comput. Geom.*, 1995. Special volume (L. Féjes Tóth Festschrift).

[Alo98] N. Alon. Piercing d-intervals. *Discr. Comput. Geom.*, 19:333–334, 1998.

[Bár82] I. Bárány. A generalization of Carathéodory's theorem. *Discrete Math.*, 40:141–152, 1982.

[BF84] E. Boros and Z. Füredi. The number of triangles covering the center of an n-set. *Geom. Dedicata*, 17:69–77, 1984.

[BFL90] I. Bárány, Z. Füredi, and L. Lovász. On the number of halving planes. *Combinatorica*, 10:175–183, 1990.

[BL92] I. Bárány and D. Larman. A colored version of Tverberg's theorem. *J. London Math. Soc. II. Ser.*, 45:314–320, 19992.

[BLŽV94] A. Björner, L. Lovász, R. Živaljević, and S. Vrećica. Chessboard complexes and matching complexes. *J. London Math. Soc.*, 49:25–39, 1994.

[BO97] I. Bárány and S. Onn. Colourful linear programming and its relatives. *Math. Oper. Res.*, 22:550–567, 1997.

[BV98] I. Bárány and P. Valtr. A positive fraction Erdős–Szekeres theorem. *Discr. Comput. Geom*, 19:335–342, 1998.

[CEG⁺95] B. Chazelle, H. Edelsbrunner, M. Grigni, L. Guibas, M. Sharir, and E. Welzl. Improved bounds on weak ϵ-nets for convex sets. *Discrete Comput. Geom.*, 13:1–15, 1995.

[ES83] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3:181–192, 1983.

[GP93] J. E. Goodman and R. Pollack. Allowable sequences and order types in discrete and computational geometry. In J. Pach, editor, *New Trends in Discrete and Computational Geometry*, volume 10 of *Algorithms and Combinatorics*, pages 103–134. Springer-Verlag, 1993.

[GPW96] J. E. Goodman, R. Pollack, and R. Wenger. Bounding the number of geometric permutations induced by k-transversals. *J. Combin. Theory Ser. A*, 75:187–197, 1996.

[HW87] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.

[Kai97] T. Kaiser. Transversals of d-intervals. *Discr. Comput. Geom.*, 18:195–203, 1997.

[Kal84] G. Kalai. Intersection patterns of convex sets. *Israel J. Math.*, 48:161–174, 1984.

[KL79] M. Katchalski and A. Liu. A problem of geometry in R^n. *Proc. Amer. Math. Soc.*, 75:284–288, 1979.

[KS96] J. Komlós and M. Simonovits. Szemerédi's regularity lemma and its applications in graph theory. In *Combinatorics, Paul Erdős is eighty., Vol. 2 (D. Miklos et al. eds.)*, pages 295–352. János Bólyai Mathematical Society, Budapest, 1996.

[Mat96] J. Matoušek. Note on the colored Tverberg theorem. *J. Combin. Theory Ser. B*, 66:146–151, 1996.

[Pac98] J. Pach. A Tverberg-type result on multicolored simplices. *Comput. Geom.: Theor. Appl.*, 10:71–76, 1998.

[Sar92] K. Sarkaria. Tverberg's theorem via number fields. *Isr. J. Math.*, 79:317, 1992.

[Sze78] E. Szemerédi. Regular partitions of graphs. In *Problemes combinatoires et theorie des graphes, Orsay 1976, Colloq. int. CNRS No.260*, pages 399–401. 1978.

[Tar95] G. Tardos. Transversals of 2-intervals, a topological approach. *Combinatorica*, 15:123–134, 1995.

[TV93] H. Tverberg and S. Vrečica. On generalizations of Radon's theorem and the Ham sandwich theorem. *European J. Comb.*, 14:259–264, 1993.

[Tve66] H. Tverberg. A generalization of Radon's theorem. *J. London Math. Soc.*, 41:123–128, 1966.

[Tve81] H. Tverberg. A generalization of Radon's theorem II. *Bull. Aust. Math. Soc.*, 24:321–325, 1981.

[VŽ93] A. Vućić and R. Živaljević. Note on a conjecture of Sierksma. *Discr. Comput. Geom*, 9:339–349, 1993.

[Živ] R. T. Živaljević. User's guide to equivariant methods in combinatorics II. *Publ. Inst. Math. Beograde.* To appear.

[ŽV92] Živaljević and S. Vrećica. The colored Tverberg's problem and complexes of injective functions. *J. Combin. Theory Ser. A*, 61:309–318, 1992.