TRACES OF FINITE SETS: EXTREMAL PROBLEMS AND GEOMETRIC APPLICATIONS

BY

ZOLTÁN FÜREDI
Dept. Math., Univ. Illinois, Urbana, IL 61801–2917, USA
E–mail: zoltan@math.uiuc.edu
and
Math. Inst. of the Hungarian Academy
P.O.B. 127, 1364 Budapest, Hungary

and

JÁNOS PACH*
Math. Inst. of the Hungarian Academy
P.O.B. 127, 1364 Budapest, Hungary
and
Courant Institute, New York Univ.,
251 Mercer St., New York, NY 10012
E–mail: pach@cims6.nyu.edu

February, 1992

Abstract. Given a hypergraph $\mathcal{H}$ and a subset S of its vertices, the *trace* of $\mathcal{H}$ on S is defined as $\mathcal{H}|S = \{E \cap S : E \in \mathcal{H}\}$. The *Vapnik–Chervonenkis dimension* (VC-dimension) of $\mathcal{H}$ is the size of the largest subset S for which $\mathcal{H}|S$ has $2^{|S|}$ edges. Hypergraphs of small VC-dimension play a central role in many areas of statistics, discrete and computational geometry, and learning theory. We survey some of the most important results related to this concept with special emphasis on (a) hypergraph theoretic methods and (b) geometric applications.

1. Introduction

A basic theme of mathematics is the characterization of certain structures in terms of their local properties. Many deep results are concerned with global consequences of some local assumptions. Their applicability is explained by the fact that it is often fairly easy to check if the local conditions are satisfied.

Research supported in part by the Hungarian National Science Foundation, Grant Nr. 1812.
* Supported in part by NSF Grant CCR–8901484 and Hungarian Science Foundation Grant OTKA–1814
This copy was printed on January 28, 2009

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

Let $\mathcal{H}$ be a hypergraph whose vertex set and edge set are denoted by $V(\mathcal{H})$ and $E(\mathcal{H})$, respectively. We frequently call (the edge set of) a hypergraph *family* or *set system.* Often, if it does not cause any misunderstanding, we identify the hypergraph $\mathcal{H}$ by its edge set $E(\mathcal{H})$. The edges of $\mathcal{H}$ are called *members* of the set system, the vertices are also called *points.* We call a hypergraph *multihypergraph* if it has multiple edges. If we want to emphasize the lack of multiple edges in $\mathcal{H}$ we call it *simple.* $\mathcal{G}$ is called a *subhypergraph* of $\mathcal{H}$ if $V(\mathcal{G}) \subseteq V(\mathcal{H})$ and $E(\mathcal{G}) \subseteq E(\mathcal{H})$. A fundamental problem of extremal set theory is to estimate the maximum number of edges a hypergraph can have without containing a subhypergraph isomorphic to a given $\mathcal{G}$. (For a recent survey, see [F91].)

In this paper we shall address a very similar type of question. For any $S \subseteq V(\mathcal{H})$, define $\mathcal{H}|S$, the *restriction* of $\mathcal{H}$ to S (or the *trace* of $\mathcal{H}$ on S), to be the hypergraph on the vertex set S with $E(\mathcal{H}|S) = \{E \cap S : E \in E(\mathcal{H})\}$. Let 2^S, $\binom{S}{k}$ and $\binom{S}{\le k}$ denote the hypergraphs on S, whose edges are all subsets, all k element subsets and all at most k element subsets of S, respectively. We shall study the basic properties of those hypergraphs whose restrictions to the 'small' subsets of their vertex sets satisfy certain conditions. For example, we may require that $\mathcal{H}|S \neq 2^S$ for any s element subset $S \subseteq V(\mathcal{H})$. We shall see e.g. that these hypergraphs have relatively few edges and admit small transversals.

Throughout this paper $[n]$ will denote the set of the first n positive integers, and $\deg_{\mathcal{H}}(x)$ (or, simply, $\deg(x)$) will stand for the *degree* of a vertex x in $\mathcal{H}$, i.e., for the number of edges $E \in E(\mathcal{H})$ for which $x \in E$.

2. VC-dimension and arrow relations

The *Vapnik–Chervonenkis dimension* (or VC-*dimension*, for short) of a hypergraph $\mathcal{H}$ is the maximum size of a subset $A \subseteq V(\mathcal{H})$ with the property that for every $B \subseteq A$ there exists $E_B \in E(\mathcal{H})$ with $E_B \cap A = B$. The most significant property of hypergraphs of small VC-dimension is that their number of edges is bounded by a polynomial of $|V(\mathcal{H})|$.

Theorem 2.1. ([S], [P], [Sh], [VC]) *For any hypergraph $\mathcal{H}$ with n vertices and* VC-*dimension d,*

$$|E(\mathcal{H})| \le \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d}.$$

Because of its importance, this theorem was discovered independently by several authors (Sauer [S] 1973, Perles [P] and Shelah [Sh] 1972, Vapnik–Chervonenkis [VC] in an implicit form, 1971), and it was also conjectured by P. Erdős.

The tightness of the upper bound follows from the fact that the hypergraph formed by all at most d-element subsets of an n-set has VC-dimension d. However, there are many extremal families (e.g., the hypergraph consisting of all the sets with at least $n-d$ elements). We return to this problem in Section 5.

The proofs in [S] and [Sh] apply induction on n. Here we present a slightly longer proof, found independently by Frankl [F83] and Alon [A], which gives more insight.

Given a hypergraph $\mathcal{H}$ and a subset $S \subseteq V(\mathcal{H})$, let $\mathcal{H}|S$ denote (as before) the trace of $\mathcal{H}$ on S. Note that in $\mathcal{H}|S$ we take every set only once (it is a *simple* hypergraph). The *arrow* relation $\mathcal{H} \to (s,t)$ means that there is an s-element subset $S \subseteq V(\mathcal{H})$ with at least t traces, i.e., $|E(\mathcal{H}|S)| \ge t$. The relation $(n,m) \to (s,t)$

means that whenever $|E(\mathcal{H})| \geq m$ one can find a subset $S \subset V(\mathcal{H})$ with at least t traces, i.e., $\mathcal{H} \to (s,t)$. Theorem 2.1 states that

$$(n,m) \to (s, 2^s) \quad \text{if } m > \sum_{0 \leq i \leq s-1} \binom{n}{i}.$$

Recall that a hypergraph is called *hereditary* (or *monotone*, or it is said to form an *ideal*) if $F \subset E \in E(\mathcal{H})$ implies $F \in E(\mathcal{H})$.

Theorem 2.2. ([A], [F83]) *For every hypergraph $\mathcal{H}$ there exists a hereditary hypergraph $\mathcal{M}$ on the same vertex set such that $|E(\mathcal{H})| = |E(\mathcal{M})|$ and $|E(\mathcal{H}|S)| \geq |E(\mathcal{M}|S)|$ for all $S \subset V(\mathcal{H})$.*

Actually, Alon [A] proved a little bit more (see later in Section 11), while Frankl [F83] stated Theorem 2.2 in a slightly weaker form, but the their proofs are identical.
Proof: Suppose that $\mathcal{E} = E(\mathcal{H})$ is not hereditary. One can find an $E_0 \in \mathcal{E}$ and an element $i \in E_0$ such that $E_0 \setminus \{i\} \notin \mathcal{E}$. Define the following *push down* operation $P : \mathcal{E} \to 2^{V(\mathcal{H})}$.

$$P(E) = \begin{cases} E \setminus \{i\} \text{ if } i \in E, \text{ and } E \setminus \{i\} \notin \mathcal{E}, \\ E \quad \text{otherwise} \end{cases}$$

Let $P(\mathcal{E}) = \{P(E) : E \in \mathcal{E}\}$. P is an injection, so $|P(\mathcal{E})| = |\mathcal{E}|$. Moreover $P(E_0) = E_0 \setminus \{i\}$, so $\sum_{E \in P(\mathcal{E})} |E| < \sum_{E \in \mathcal{E}} |E|$. We claim, that $|\mathcal{E}|S| \geq |P(\mathcal{E})|S|$ for all $S \subset V$. This is obvious if $i \notin S$. Otherwise, one can split the subsets of S into $2^{|S|-1}$ pairs, $\{Z, Z \cup \{i\}\}$ ($Z \subset S \setminus \{i\}$). It is easy to see, that $|P(\mathcal{E}) \cap \{Z, Z \cup \{i\}\}| \leq |\mathcal{E} \cap \{Z, Z \cup \{i\}\}|$. Clearly, using a series of the above operations one can transform $\mathcal{E}$ in finitely many steps into a hereditary hypergraph possessing the desired properties. $\square$
Proof of Theorem 2.1: Consider a hypergraph $\mathcal{H}$ on n vertices with VC-dimension d. By Theorem 2.2 there exists an ideal $\mathcal{I}$ of the same size on the same vertex set such that $\mathcal{I} \not\to (d+1, 2^{(d+1)})$. All edges of $\mathcal{I}$ have at most d elements, so $|E(\mathcal{H})| = |E(\mathcal{I})| \leq \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d}$. $\square$

3. Collapsing one element

Bondy [Bon] observed that a hypergraph $\mathcal{H}$ on n points and with at most n edges has an element $i \in V$ such that the number of edges in $\mathcal{H}|(V \setminus \{i\})$ is the same as in $\mathcal{H}$. With arrow relations,

$$(n,m) \to (n-1,m) \quad \text{if} \quad m \leq n. \tag{3.1}$$

Bollobás [Bol] proved that

$$(n,m) \to (n-1,m-1) \quad \text{if} \quad m \leq 3n/2. \tag{3.2}$$

Notice that both of these theorems are easy corollaries to Theorem 2.2. E.g., to show (3.2) consider a hypergraph $\mathcal{E}$ with $|\mathcal{E}| = m \leq 3n/2$ on an underlying set V, $|V| = n$. There exists a hereditary hypergraph $\mathcal{M}$ on V with $|\mathcal{M}| = |\mathcal{E}|$ such that $|\mathcal{M}|V_i| \leq |\mathcal{E}|V_i|$ for all $V_i := V \setminus \{i\}$. If $\{i\} \notin \mathcal{M}$, then $|\mathcal{M}|V_i| = m$. Hence $\mathcal{M} \to (n-1,m)$ and so does $\mathcal{E}$. If $\mathcal{M}$ contains all the singletons, then there must be an element, say i which is not covered by any 2-element set in $\mathcal{M}$, otherwise $|\mathcal{M}| \geq 1+n+\lceil n/2 \rceil$. Deleting i from V, the only collapse of the traces on V_i comes from the pair $\{\emptyset, \{i\}\}$, so $\mathcal{M} \to (n-1,m-1)$. $\square$

Theorem 3.1. (Frankl [F83]) $(n, m) \to (n-1, m-(2^{t-1}-1))$ *for* $m \le n(2^t - 1)/t$.

Proof: Let $\mathcal{M}$ be a hereditary hypergraph on V such that $\mathcal{M} \not\to (n-1, m-(2^{t-1}-1))$. Then the degree of every point in $\mathcal{M}$ is at least 2^{t-1}. Using a weighted version of the Kruskal–Katona theorem ([Kr], [Ka], see Theorem 7.1) due to Katona [Ka78], one can see that $\sum_{i\in E\in\mathcal{M}} 1/|E| \ge (2^t-1)/t$ for every $i \in V$. Thus,

$$|\mathcal{M}| = 1 + \sum_{E\in\mathcal{M}} \sum_{i\in E} 1/|E| = 1 + \sum_{i\in V} \sum_{i\in E\in\mathcal{M}} 1/|E| \ge 1 + n(2^t-1)/t. \quad \square$$

If t is a divisor of n, the extremal families are obtained by taking n/t disjoint sets of size t and all of their subsets. This weight function method was extended by Watanabe [W]. She proved

$$\begin{aligned}(n,m) &\to (n-1, m-4) \text{ for } m \le 17n/6,\\ (n,m) &\to (n-1, m-5) \text{ for } m \le 13n/4,\\ (n,m) &\to (n-1, m-6) \text{ for } m \le 7n/2.\end{aligned}$$

The coefficients 17/6, 13/4 and 7/2 are best possible. The arrow relation $(n,m) \to (n-1, m-2)$ follows immediately from Theorem 2.2 if $m \le 2n$.

Problem 3.2. *Find the largest* $c = c(k)$ *such that* $(n, m) \to (n-1, m-k)$ *for all* $m \le (c(k) - o(1))n$.

Frankl and Watanabe have recently informed us that they determined $c(k)$ for all $k \le 10$. Bondy and Hajnal conjectured that $(n, 1 + \sum_{0\le i\le d} \binom{n}{i}) \to (n-1, 1 + \sum_{0\le i\le d} \binom{n-1}{i})$. Frankl [F78] disproved this conjecture for $n = 6r+3$, $d = 3r+1$. He asked asked whether it holds for large values of n, $n > n_0(d)$. His example consists of 2^{n-1} sets, namely $\mathcal{M} = \{E \subset V : |E \cup Y_i| \le r$ for at least 2 of the Y_i's$\}$, where Y_1, Y_2, Y_3 is a partition of V into $(2r+1)$-element parts.

Problem 3.3. *Is it true for* $n > n_0(k, d)$ *that* $(n, 1 + \sum_{0\le i\le d} \binom{n}{i}) \to (k, 1 + \sum_{0\le i\le d} \binom{k}{i})$?

4. Arrow relations and Turán numbers

A hypergraph is called *d-uniform* (or a *d-graph*) if each of its edges consists of d points. Let $T(n, k, d)$ denote the maximum number of edges of a d-uniform hypergraph on n points without a complete subhypergraph of k points. These numbers are called the *Turán numbers* and, except for some trivial cases, their exact values are known only for $d = 2$ (Turán's theorem [T]), e.g. $T(n, 3, 2) = \lfloor n^2/4 \rfloor$. A recent survey can be found in this volume [dC]. The following theorem is an easy consequence of Theorem 2.2. For $d = 2$ it was conjectured by Lovász and proved independently by Alon [A] and Frankl [F83].

Theorem 4.1. $(n, 1 + \sum_{0\le i\le d-1} \binom{n}{i} + T(n, d+1, d)) \to (d+1, 2^{d+1}-1)$.

Proof: Let $\mathcal{M}$ be a hereditary hypergraph with $\mathcal{M} \not\to (d+1, 2^{d+1}-1)$. It is easy to check that (1) $\mathcal{M}$ does not have an edge of size at least $d+1$, and (2) it cannot contain any complete d-graph on $d+1$ vertices, as a subhypergraph. $\square$

The sets $S_1, S_2, \ldots, S_{k+1}$ are called *disjointly representable* if there exist $x_1, \ldots, x_{k+1}$ such that $x_i \in E_j$ if and only if $i = j$. Erdős and Gyárfás observed that any large set system of r-sets should contain $k+1$ disjointly representable members.

Conjecture 4.2. (Frankl, Pach [FP84]) *Let $\mathcal{E}$ be an r-uniform hypergraph without $k+1$ disjointly representable edges. Then $|\mathcal{E}| \leq T(r+k, k+1, k)$.*

This upper bound, if true, is best possible, as can be seen by the following example. Take a k-uniform hypergraph $\mathcal{F}$ on the $(r+k)$-element underlying set V without a complete k-graph on $k+1$ elements. Then the system of complements, $\mathcal{E} = \{V \setminus F : F \in \mathcal{F}\}$, is obviously an r-uniform hypergraph without more than k disjointly representable edges.

Frankl and Pach proved their conjecture in the case $r = 2$ (and k is arbitrary), and in the case $k = 2$ (and $r \geq 1$ is arbitrary). Here we reformulate the latter one.

Theorem 4.3. (Frankl, Pach [FP84]) *Let $\mathcal{E} = \{E_1, E_2, \ldots, E_m\}$ be an r-uniform hypergraph such that from any three of its edges one is contained in the union of the other two. Then $m \leq 1 + r + \lfloor r^2/4 \rfloor$.*

Moreover, the only r-uniform hypergraph $\mathcal{E}$ without 3 disjointly representable edges achieving this upper bound can be obtained as follows. Let A and B be two disjoint sets with $|A| = \lfloor (r+2)/2 \rfloor$, $|B| = \lceil (r+2)/2 \rceil$, and let $\mathcal{E}$ be the family of all r-subsets $E \subset A \cup B$ with $|A \setminus E| = 1$ and $|B \setminus E| = 1$.

Proof: Choose a minimal set F_i, such that $F_i \cap E_i = \emptyset$, but $F_i \cap E_j \neq \emptyset$ for all $j \neq i$. Observe that $|F_i| \leq 2$, because, by minimality, F_i represents $|F_i|$ edges of $\mathcal{E}$. Apply induction on r. If there exists an $|F_i| < 2$, then $m \leq 1 + f(r-1) < f(r)$, where $f(x) = 1 + x + \lfloor x^2/4 \rfloor$, and we are done. If $|F_i| = 2$ for all i, then define $\mathcal{G} = \{F_i : F_i \subset E_1\}$. It is easy to see that $\mathcal{G}$ is a triangle-free graph, so Turán's theorem implies $|\mathcal{G}| \leq \lfloor r^2/4 \rfloor$. Moreover, all other F_i's with $i \geq 2$ meet E_1 in distinct points. So the rest contains at most $1 + r$ pairs. $\square$

We are going to return to this problem in Section 8.

5. Extremal families

We say that $\mathcal{F} \to \mathcal{H}$ if $\mathcal{H}$ is contained in a trace of $\mathcal{F}$, i.e., there exists $F_1, \ldots, F_t \in \mathcal{F}$ and $S \subseteq V(\mathcal{F})$, $|S| = |V(\mathcal{H})|$ such that $F_1 \cap S, \ldots, F_t \cap S$ form a hypergraph on S isomorphic to $\mathcal{H}$. Let $f(n, \mathcal{H})$ be the largest m such that there exists a hypergraph $\mathcal{F}$ on n vertices and m edges with $\mathcal{F} \not\to \mathcal{H}$. If $\mathcal{F} \not\to \mathcal{H}$ and $|\mathcal{F}| = f(n, \mathcal{H})$, then $\mathcal{F}$ is called *extremal* (or *$\mathcal{H}$-extremal*).

By this terminology, Theorem 2.1 states that $f(n, 2^S) = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d}$, where $d = |S| - 1$. We shall see that there are many 2^S-extremal hypergraphs. Let $n \geq d \geq \ell > 0$. One might think that $f(n, \binom{S}{\ell})$ is much smaller than $f(n, 2^S)$. However, the following example shows, that they coincide.

Example 5.1. (Füredi, Quinn [FQ]) *Let $\mathcal{F} = \mathcal{F}(n, d, \ell)$ consist of those $E \subset [n]$ for which there exists a j such that $|[j] \cap E| = \ell$ and $|\{j, j+1, \ldots, n\} \setminus E| \leq d - \ell$. Moreover, let $\binom{[n]}{i} \subset \mathcal{F}$ for all $i < \ell$. Then $|\mathcal{F}| = \sum_{0 \leq i \leq d} \binom{n}{i}$, but $\mathcal{F} \not\to \binom{S}{\ell}$. That is, $\mathcal{F}$ is an $\binom{S}{\ell}$-extremal hypergraph.*

Proof: For any set $Y = \{y_1 < y_2 < \cdots < y_{d+1}\} \subset [n]$, $Y \cap E = \{y_1, y_2, \ldots, y_\ell\}$ is impossible. $\square$

The characterization of $\binom{S}{\ell}$-extremal hypergraphs (or more generally, $\mathcal{H}$-extremal hypergraphs) is a challenging open problem.

Conjecture 5.2. (Füredi, Quinn [FQ]) *If $\mathcal{F}$ is an $\binom{S}{\ell}$-extremal hypergraph, then $|\mathcal{F} \cap \binom{[n]}{i}| = |\mathcal{F}(n, d, \ell) \cap \binom{[n]}{i}|$, i.e.,*

$$|\mathcal{F}\cap\binom{[n]}{i}| = \begin{cases} \sum_a \binom{i+d+1-2\ell}{i-\ell+1+2a}\binom{n-i-d-1+2\ell}{\ell-1-a} & \text{for } \ell\le i\le n-d-1+\ell \\ \binom{n}{i} & \text{otherwise.}\end{cases}$$

The cases $\ell = 0$, $\ell = d+1$ are trivial, the case $\ell = 1$ (and hence, by taking complements, the case $\ell = d$) was proved in [FQ]. The case $|S| = 3$, $\ell = 2$ was confirmed by Anstee [An80]. He has also established the following structure theorem. If $\mathcal{F}$ is a $\binom{[3]}{2}$-extremal hypergraph, then $\mathcal{F} \not\to \mathcal{C}_k$ for all k (here $\mathcal{C}_k$ denotes the cycle of length k). Furthermore, if $\mathcal{F}_0 \not\to \mathcal{C}_k$ for all k, then it can be extended to a $\binom{[3]}{2}$-extremal hypergraph. Consequently, the number of $\mathbf{C}_3$-extremal hypergraphs is at least as large as the number of trees, i.e., its order of magnitude is exponential ([An83]).

6. Linear algebraic proofs

The investigation of $\mathcal{C}_3$-free hypergraphs by linear algebraic techniques was initiated by Ryser [R72]. Let $\mathcal{F}$ be a hypegraph with vertex set $[n]$, r a natural number. The *generalized incidence matrix* of $\mathcal{F}$, $M = M(\mathcal{F}, \le r)$, is defined as an $|\mathcal{F}| \times (\sum_{0\le i\le r}\binom{n}{i})$ matrix where the rows are labeled by the edges of $\mathcal{F}$, the columns are labeled by the at most r-element subsets of $[n]$, and

$$M_{F,Y} = \begin{cases} 1 & \text{if } \ Y\subset F, \\ 0 & \text{if } \ Y\not\subset F.\end{cases}$$

The definition of $M(\mathcal{F}, r)$ is similar, except that now the columns are labeled by the r-subsets only. Note that $M(\mathcal{F}, 1)$ is the usual incidence matrix of $\mathcal{F}$.

Theorem 6.1. (Frankl, Pach [FP83]) *Let $\mathcal{F}$ be a hypergraph on n points. If the rows of the generalized incidence matrix $M(\mathcal{F}, \le r)$ are linearly dependent, then $\mathcal{F} \to 2^S$ for some $S \subset [n]$, $|S| \ge r+1$.*

This immediately implies Theorem 2.1. Here we present the proof of a related result of Frankl and Pach, but the proof of Theorem 6.1 is almost identical.

Theorem 6.2. (Frankl, Pach [FP84]) *Let $\mathcal{F}$ be an k-uniform hypergraph on $[n]$, $(n \ge k > r \ge 1)$. If the rows of the matrix $M(\mathcal{F}, r)$ are linearly dependent, then* VC-dim$(\mathcal{F}) \ge r+1$.

Proof: Consider a linear dependence with coefficients $\alpha(E)$, i.e., $\sum_{E\in\mathcal{F}}\alpha(E)M_{E,Y} = 0$ for all $Y \subset [n]$, $|Y| = r$. It follows that $\sum_{E\in\mathcal{F},Y\subset E}\alpha(E) = 0$ holds for all $|Y| \le r$. Let S be a minimal subset of $[n]$ with $\sum_{E\in\mathcal{F},S\subset E}\alpha(E) \ne 0$. Clearly, $|S| \ge r+1$. Using a backward induction it can be shown that for all $Z \subset S$

$$\sum_{E\in\mathcal{F},E\cap S=Z}\alpha(E) = (-1)^{|S\setminus Z|}\sum_{E\in\mathcal{F},S\subset E}\alpha(E).$$

This yields, in particular, that for every $Z \subset S$ there exists $E \in \mathcal{F}$ with $E \cap S = Z$. $\square$

Let $f(n,k,\mathcal{H})$ be the largest m such that there exists a (simple) k-uniform hypergraph $\mathcal{F}$ on n vertices with $\mathcal{F} \not\to \mathcal{H}$. The upper bound in the following inequality is an immediate corollary of Theorem 6.2. For every $n \geq k \geq i \geq 0$ one has

$$\binom{n-1}{k-1} \leq f(n,k,\binom{[k]}{\geq i}) \leq \binom{n}{k-1} \tag{6.1}$$

The lower bound follows by considering all k-element sets containing a common element. The upper bound cannot be improved, for example in the case $n = 2k-1$, $i = 0$. Applying Theorem 8.1 to the system of complements, we obtain

$$f(n,k,\binom{[k]}{\geq k-1}) = \binom{n-1}{k-1}. \tag{6.2}$$

Conjecture 6.3. (Frankl, Pach [FP84]) *For n sufficiently large, ($n > n_0(k)$), the lower bound in (6.1) is tight.*

This conjecture is particularly interesting in the case $i = 0$, for it would yield a new generalization of the Erdős–Ko–Rado theorem [EKR].

7. The number of dense subsets

We shall demonstrate the power of classical hypergraph theory by answering a question of M. Karchmer [Kar] about the minimum number of dense subsets.

Let $\mathcal{H}$ be a hypergraph with vertex set $[n]$. A subset $S \subset [n]$ is called *d-dense* if $\mathcal{H}|S = 2^S$ and $|S| = d$. Let $\mathcal{D}_d(\mathcal{H})$ is the family of d-dense sets. Karchmer posed the following question. Given $|\mathcal{H}|$, what is the minimum number of d-dense sets? By Theorem 2.1, $\mathcal{D}_d(\mathcal{H}) \neq \emptyset$ for $|\mathcal{H}| > \sum_{0\leq i\leq d-1}\binom{n}{i}$, but it can be $\emptyset$ below that threshold.

The *antilexicographic* ordering of the (finite) subsets of $\mathbf{N} = \{1,2,3,\dots\}$ is defined by $A < B$ if and only if $\max(A \triangle B) \in B$. This is a total (linear) order, starts with $\emptyset$, 1, 2, 21, 3, 31, 32, 321, 4, 41, 42, 421, 43, 431, … Let $\mathcal{F}(f,k)$ ($\mathcal{F}(f,\geq k)$) denote the first f k-sets (subsets of size at least k, respectively) according to this ordering.

Theorem 7.1. (Kruskal [Kr], Katona [Ka]) *Let $\mathcal{F}$ be a k-uniform hypergraph, $|\mathcal{F}| = f$. For any $0 < \ell < k$ the number of ℓ-element sets contained in at least one edge of $\mathcal{F}$ is at least as large as the corresponding number for $\mathcal{F}(f,k)$.*

Denote the set of ℓ-subsets covered by some edge of $\mathcal{F}$ by $\partial_\ell \mathcal{F}$. It is easy to see that any positive integer f can be uniquely written in the following *cascade* form:

$$f = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \cdots + \binom{a_t}{t}, \tag{7.1}$$

where $a_k > a_{k-1} > \cdots > a_t \geq t \geq 1$ are integers. Then Theorem 7.1 states that $|\partial_\ell \mathcal{F}| \geq \binom{a_k}{\ell} + \binom{a_{k-1}}{\ell-1} + \cdots + \binom{a_t}{t-(k-\ell)}$. A short proof of the Kruskal–Katona theorem was found by Frankl [F84]. Now we are in a position to answer Karchmer's question.

Theorem 7.2. *Let $\mathcal{H}$ be a hypergraph with vertex set $[n]$, and write $|\mathcal{H}|$ in the form*

$$|\mathcal{H}| = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d-1} + m. \tag{7.2}$$

Then the number of d-dense sets in $\mathcal{H}$, $|\mathcal{D}_d(\mathcal{H})|$, is minimal if $\mathcal{H} = \mathcal{F}(n,d,m) := \binom{[n]}{\le d-1} \cup \mathcal{F}(m, \ge d)$.

For example, if $m = 2^k - \sum_{0 \le i \le d-1} \binom{k}{i}$, then $|\mathcal{D}_d(\mathcal{H})| \ge \binom{k}{d}$.
Proof: Apply Theorem 2.2 to an optimal hypergraph $\mathcal{H}$. We obtain an ideal $\mathcal{I}$ with $|\mathcal{I}| = |\mathcal{H}|$ and $\mathcal{D}_d(\mathcal{I}) = \mathcal{D}_d(\mathcal{H})$. Let $\mathcal{I}^{(i)}$ denote the set of i-element edges of $\mathcal{I}$. Let f be the largest integer such that for the initial segment $\mathcal{F} := \mathcal{F}(f, \ge d)$ one has $|\mathcal{F}^{(d)}| = |\mathcal{I}^{(d)}|$. Then, $\mathcal{F}^{(i)}$ is an initial segment in $\binom{\mathbf{N}}{i}$, with $\partial_d \mathcal{F}^{(i)} \subseteq \mathcal{F}^{(d)}$ $(i \ge d)$. The hypergraph $\mathcal{F}$ (together with the sets of size less than d) is again an ideal. The Kruskal-Katona theorem implies that $\mathcal{F}^{(i)}$ is a largest i-uniform hypergraph $\mathcal{J}$ with $|\partial_d(\mathcal{J})| \le |\mathcal{F}^{(d)}|$. That is, $|\mathcal{F}^{(i)}| \ge |\mathcal{I}^{(i)}|$ holds for every i. As $\mathcal{F}(m, \ge d)$ is a subfamily of $\mathcal{F}$, and $|\mathcal{D}_d(\mathcal{I})| = |\mathcal{F}^{(d)}|$, the proof is complete. □

Using (7.1) one can also obtain an exact formula.

8. More about disjointly representable sets

Let $m(r,k)$ be the largest m such that there exists an r-uniform hypergraph of size m without $k+1$ disjointly representable edges. This function was investigated in Section 4. The definition of $m(\le r, k)$ is analogous.

Theorem 8.1. (Füredi, Tuza [FT]) *Let $\mathcal{F}$ be a hypergraph of rank r and $|\mathcal{F}| > \binom{r+k}{r}$. Then $\mathcal{F} \to \binom{[k+1]}{\le 1}$, i.e., there exists $F_0, F_1, \ldots, F_{k+1} \in \mathcal{F}$ and a set $Y = \{y_1, \ldots, y_{k+1}\}$, such that $Y \cap F_0 = \emptyset$ and $Y \cap F_i = \{y_i\}$ for all $1 \le i \le k+1$.*

In the proof we are going to use the following important result.

Theorem 8.2. (Frankl [F82], Kalai [K]) *Let $A_1, A_2, \ldots, A_m$ be at most r-element, $B_1, B_2, \ldots, B_m$ be at most k-element sets with $A_i \cap B_i = \emptyset$. Suppose that $A_i \cap B_j \neq \emptyset$ for $i > j$. Then $m \le \binom{r+k}{r}$.*

The same assertion was proved by Bollobás [B] in 1965, under the stronger assumption that $A_i \cap B_j = \emptyset$ iff $i \neq j$.
Proof of Theorem 8.1: Let $\mathcal{F} = \{F_1, F_2, \ldots, F_m\}$ be an ordering of the sets such that $|F_i| \le |F_j|$ for $i < j$. For every set $F \in \mathcal{F}$ one can find an at most k-element set $B = B(F)$ such that $F \cap B(F) = \emptyset$ and $B(F) \cap F' \neq \emptyset$ whenever $F' \in \mathcal{F}$ and $F' \not\subset F$. Then, Theorem 8.2 can be applied to $\{F_j, B(F_j)\}_{1 \le j \le m}$. □

Together with the results of Section 4, this yields

$$T(r+k, k+1, k) \le m(r,k) \le m(\le r, k) = \binom{r+k}{k}. \tag{8.1}$$

Note that here the lower and upper bounds are very close to each other. If k is fixed and r tends to infinity then the ratio of the lower and upper bounds above tends to a limit $c_k > 0$. Moreover, $\lim_{k\to\infty} c_k = 1$. Using Example 5.1, one can obtain several extremal families, e.g. $\mathcal{F} := \cup_{0 \le i \le r} \binom{[i+k-1]}{i}$. The description of all extremal families does not seem to be hopeless. In [FT] it was proved that for any hypergraph $\mathcal{F}$ of rank r with $|\mathcal{F}| = \binom{r+k}{r}$ and with $\mathcal{F} \not\to \binom{[k+1]}{\le 1}$ one has

$$|\mathcal{F}_i| = \binom{i+k-1}{i}, \tag{8.2}$$

where $\mathcal{F}_i$ denotes the set of i-element edges of $\mathcal{F}$. The structure of $\binom{[k+1]}{1}$-extremal families is probably more complicated, (8.2) does not hold. In the case $k = 2$, other extremal families can be exhibited using the result of Anstee [An83] mentioned after Conjecture 5.2.

The sets $E_1, E_2, \dots, E_\ell$ form a Δ-*system* of size ℓ with kernel K if $E_i \cap E_j = K$ for every pair $1 \le i < j \le \ell$. The sets $F_1, \dots, F_{k+1}$ are called *disjointly t-representable* if one can choose disjoint t-element subsets $Y_i \subset F_i$ such that $Y_i \cap F_j = \emptyset$ whenever $i \ne j$.

Problem 8.3. *Determine max $|\mathcal{F}|$ over all hypergraphs $\mathcal{F}$ of rank r without containing a Δ-system of size ℓ whose kernel has at least $r-t$ points and without having $k+1$ disjointly t-representable edges.*

Frankl and Pach [FP84] proved that there are constants c_1, c_2 depending only from k, ℓ and t such that

$$c_1 r^{k(t-1)} < \max |\mathcal{F}| < c_2 r^{(k+1)(t-1)-1}.$$

9. Forcing an ordered substructure

Given a natural number n and a class $\mathcal{L}$ of $(0,1)$-matrices (so called *forbidden submatrices*), determine the maximum integer m such that there exists an $m \times n$ $(0,1)$-matrix M without repeated rows and containing no element of $\mathcal{L}$ as a submatrix. Let us denote this maximum by $\text{ex}(n, \mathcal{L})$. We write $A \to \mathcal{L}$ if A has a submatrix that belongs to $\mathcal{L}$. So $\text{ex}(n, \mathcal{L}) = \max\{m : M$ is an $m \times n$ simple (i.e, no multiple rows) $(0,1)$-matrix such that $M \not\to \mathcal{L}\}$. If A is the incidence matrix of the hypergraph $\mathcal{F}$, and $\mathcal{L}(\mathcal{H})$ is the family of matrices obtained from the incidence matrix of $\mathcal{H}$ by row and column permutations, then $A \to \mathcal{L}(\mathcal{H})$ if and only if $\mathcal{F} \to \mathcal{H}$. Moreover, using the notation in Section 5, $\text{ex}(n, \mathcal{L}(\mathcal{H})) = f(n, \mathcal{H})$. For example, if $\mathcal{L}_s$ is the family of all $2^s \times s$ matrices containing every $(0,1)$-vector of length s (as a row) exactly once, then Theorem 2.1 states that $\text{ex}(n, \mathcal{L}_s) = \sum_{0 \le i \le s-1} \binom{n}{i}$.

Theorem 9.1. (Frankl, Füredi, Pach [FFP]) *Let $\mathcal{L}$ be any family of forbidden $(0,1)$-matrices, and suppose that there is a $t \times s$-matrix $L \in \mathcal{L}$. Then,*

$$\text{ex}(n, \mathcal{L}) \le \left((t-1)\binom{n}{s} + 1\right)\left(\sum_{0 \le i \le s-1} \binom{n}{i} + 1\right) - 1 \le t n^{2s-1}. \tag{9.1}$$

Proof: Let M be a simple $m \times n$ matrix. If m exceeds the bound of (9.1), then M contains $(t-1)\binom{n}{s} + 1$ copies of $\mathcal{L}_s$, each of them lying entirely below its predecessors. But then, by the pigeonhole principle, there are at least t such copies of $\mathcal{L}_s$ having the same set of s columns. ¿From the ith copy, we can select the ith row of L and so produce a forbidden submatrix. □

Conjecture 9.2. ([AnF], [FFP]) *For any $t \times s$ matrix L, there is a constant c_L so that (9.1) may be replaced by* $\mathrm{ex}(n, \mathcal{L}) \le c_L n^s$.

In [AnF] it was proved that for any $1 \times s$ matrix L

$$\mathrm{ex}(n, L) = \sum_{0 \le i \le s-1} \binom{n}{i}. \tag{9.2}$$

This was given by a generalization of Example 5.1. A number of other cases were also considered:
(9.3) where L is a $2 \times s$ matrix,
(9.4) matrices with repeated rows,
(9.5) matrices having only one column.

As an application of the results, mainly constructions, in [AnF] the case of 2×3 matrices has been studied in detail. There are 14 essentially different 2×3 $(0,1)$-matrices (up to taking $(0,1)$-complements and reversing row and/or column order):

$$F_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad F_2 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad F_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$F_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad F_5 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad F_6 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$F_7 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad F_8 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad F_9 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$F_{10} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad F_{11} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$F_{12} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad F_{13} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad F_{14} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

FIGURE 1. The 14 essentially distinct 2×3 (0,1) matrices

Theorem 9.3. (Anstee, Füredi [AnF])

$$\mathrm{ex}(n, F_i) = \binom{n}{3} + \binom{n}{2} + \binom{n}{1} + \binom{n}{0}, \ \textit{for } 1 \le i \le 7,$$

$$\mathrm{ex}(n, F_i) = \binom{n}{2} + \binom{n}{1} + \binom{n}{0} + \binom{n-1}{2} \ \textit{for } i = 8, 9$$

$$\mathrm{ex}(n, F_{10}) = \binom{n}{3} + O(n^2),$$

$$\mathrm{ex}(n, F_{11}) = \binom{n}{3} + \binom{n}{2} + \binom{n}{1} + \binom{n}{0} + \binom{n-2}{2},$$

$$\binom{n}{3} + \binom{n}{2} + \binom{n}{1} + 1 \le \mathrm{ex}(n, F_i) \le \binom{n}{4} + \binom{n}{3} + \binom{n}{2} + \binom{n}{1} + 1 \ \textit{for } 12 \le i \le 14.$$

Theorem 9.4. (Frankl, Füredi, Pach [FFP]) *Let D_s be a $2\times s$ $(0,1)$-matrix, a row of 1's followed by a row of 0's. Then, for $s=2$ (and $n>1$) $\mathrm{ex}(n,D_2)=\binom{n}{2}+2n-1$. For $s>2$, $n\geq 2s$,*

$$\binom{n}{s}+2\left(\binom{n}{s-1}+\binom{n}{s-2}+\cdots+\binom{n}{0}\right)-\binom{2s-1}{s}\leq$$
$$\mathrm{ex}(n,D_s)\leq\binom{n}{s}+5s^2\binom{n}{s-1}. \quad (9.6)$$

Let $\mathcal{F}$ be a simple hypergraph with vertex set $[n]$, and let $s\geq 2$ a fixed natural number. It is well–known, and follows easily from Theorem 2.1, that, if $|F\setminus F'|<s$ for all $F,F'\in\mathcal{F}$, then $|\mathcal{F}|\leq\sum_{0\leq i\leq s-1}\binom{n}{i}$, and this bound cannot be improved. The problem of determining $\mathrm{ex}(n,D_s)$ can be reformulated as follows. What is the maximum length of a sequence $\{F_1,F_2,\ldots,F_m\}$ of distinct subsets of $[n]$ with the property that

$$|F_i\setminus F_j|<s \text{ for all } i<j? \quad (9.7)$$

Without loss of generality we may suppose that the F_i's are listed in increasing order of their cardinalities, i.e., $|F_i|\leq|F_j|$. The lower bound in (9.6) is shown by the following construction. Fix a chain of subsets $E_1\subset E_2\subset\cdots\subset E_n=[n]$ with $|E_i|=i$ and let $\mathcal{F}_i=\{F\subset[n]:|F|=i,\ F\supset E_{i-k+1}\}$. Finally, define

$$\mathcal{F}=:\{F\subset[n]:|F|<s\}\cup(\cup_{s\leq i\leq n-s}\mathcal{F}_i)\cup\{F\subset[n]:|F|>n-s\}.$$

Conjecture 9.5. (Frankl, Füredi, Pach [FFP]) *There exists a sufficiently large $n_0(k)\geq 2k$ such that if $n\geq n_0(k)$, then $\mathrm{ex}(n,D_s)$ equals the lower bound in (9.6).*

It might be difficult to confirm this conjecture, because if $n=2s+\sqrt{s}/10$, then $\mathrm{ex}(n,D_s)$ significantly exceeds the left hand side of (9.6); and in the case $s=2$ there are exactly 2^{n-2} essentially different extremal families [FFP]. A hypergraph $\mathcal{F}$ is called a *Sperner family* if $F\not\subset G$ holds for every pair $F,G\in\mathcal{F}$.

Conjecture 9.6. (Frankl, Füredi, Pach [FFP]) *Let $\mathcal{F}=\{F_1,F_2,\ldots,F_m\}$ be a Sperner family of the subsets of $[n]$ satisfying condition (9.7). Then $|\mathcal{F}|\leq\binom{n}{s-1}$ holds for $n\geq 2s-3$.*

The cases $2s-3\leq n\leq 2s$ have been settled.

10. The density of matrices without a forbidden submatrix

A *configuration* $C=(c_{ij})$ $(1\leq i\leq t,\ 1\leq i\leq s)$ is a partial matrix with 1's and 'blanks' at the entries. We say that a $(0,1)$-matrix M contains the configuration C if one can find t rows $i_1,i_2,\ldots,i_t$, $i_1<i_2<\cdots<i_t$ and s columns $j_1,j_2,\ldots,j_s$, $j_1<j_2<\cdots<j_s$ in M such that the corresponding submatrix contains C, i.e. $M_{i_\alpha,j_\beta}=1$ whenever $c_{\alpha,\beta}=1$. Let $g(m,n;C)$ denote the maximum number of 1's in an $m\times n$ matrix M not containing C. In the case $n=m$ we write $g(n;C)$. For a collection $\mathcal{C}$ of forbidden configurations, the corresponding threshold function is denoted by $g(n;\mathcal{C})$.

This problem is closely related to the Turán–type questions of extremal graph theory. Our matrices can be considered as bipartite graphs. The important difference is, however, that in our case the vertices (i.e., the rows and columns) are ordered.

In [FH] it was proved that for all configurations C with at most 3 entries, and for 28 of the 37 configurations of 4 entries the threshold function is linear, $g(n;C) = O(n)$. A result on the Zarankiewicz problem (due to [EKST]) immediately gives $g(n; \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}) = \Theta(n^{3/2})$. For all other matrices with four entries $g(n;C) = O(n \log n)$. For example,

$$\Omega(n \frac{\log n}{\log\log n}) \le g(n; \begin{pmatrix} 1 & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}) \le O(n \log n). \tag{10.1}$$

(Bienstock and Győri [BG].)

Conjecture 10.1. (Bienstock, Győri [BG]) *In (10.1) the lower bound gives the correct order of magnitude.*

The result $g(n; \begin{pmatrix} 1 & 1 & \\ 1 & & 1 \end{pmatrix}) = O(n \log n)$ was used in [F90] to prove that the number of unit distances determined by n points in the plane in convex position is at most $12 n \log n$. The best lower bound, $2n - 7$, for this distance problem is due to Edelsbrunner and Hajnal [EH].

More results, problems, connections to the Davenport–Schinzel theory can be found in [FH].

11. Further generalizations

Alon generalized Theorem 2.2 to $[k_1] \times [k_2] \times \dots [k_n]$, instead of the Boolean lattice $2^{[n]}$. The corresponding analogue of Theorem 2.1 was first proved by Karpovsky and Milman [KM], and independently by Anstee and Murty [AM]. Many similar questions can be asked for other classes of matrices, when we allow, say, $0, 1, 2$ entries (instead of 0's and 1's only).

Another direction of research is to investigate the *minimum degree*, instead of the size of the hypergraph. This problem was proposed by Cunningham and Frankl. Let $\mathcal{H}$ be a hypergraph and let $d(n, \mathcal{H}) := \max_{\mathcal{F}} \{\min_{x \in [n]} \deg_{\mathcal{F}}(x)\}$ where $\mathcal{F} \subset 2^{[n]}$ and $\mathcal{F} \not\to \mathcal{H}$. It is easy to see, that $d(n, 2^{[s]}) = \sum_{0 \le i \le s-1} \binom{n-1}{i}$. A hypergraph $\mathcal{H}$ is k-partite if there exists a partition $V_1 \cup \dots \cup V_k = V(\mathcal{H})$ such that $|E \cap V_i| \le 1$ for every edge E, $1 \le i \le k$. In [FQ] it was observed that if $\mathcal{H}$ is not k-partite, then $d(n, \mathcal{H}) \ge \Omega(n^k)$. For further problems consult [FQ].

A hypergraph $\mathcal{F}$ is called *t-wise ℓ-intersecting* if $|F_1 \cap \dots \cap F_t| \ge \ell$ holds for every $F_1, \dots, F_t \in \mathcal{F}$.

Theorem 11.1. *There exists a function $v(r,t,\ell)$ such that the following holds. If $\mathcal{F}$ is a t-wise ℓ-intersecting r-uniform hypergraph, then there exists a set S with $|S| \le v(r,t,\ell)$ such that any t edges of $\mathcal{F}$ meet in at least ℓ vertices of S, i.e., $|F_1 \cap \dots \cap F_t \cap S| \ge \ell$ for all $F_1, \dots, F_t \in \mathcal{F}$.*

This theorem was proved by Całczyńska-Karłowitz [C-K] in the most important special case $t = 2, \ell = 1$. His upper bound has been significantly improved by Erdős and Lovász (see [F90]), and Tuza [Tu]:

$$2r - 4 + 2\binom{2r-4}{r-2} \le v(r,2,1) \le \binom{2r-1}{r-1} + \binom{2r-4}{r-2}.$$

The general case has been studied independently by several people. For any fixed t and ℓ, Alon and Füredi [AF] found a lower bound on $v(r,t,\ell)$ exponential in r. For the latest developments and conjectures in this area, consult Kohayakawa [K90].

12. Packings, coverings and fractional matchings

The *packing number* of a hypergraph $\mathcal{H}$ is defined as the largest number $\nu = \nu(\mathcal{H})$ of pairwise disjoint edges in $\mathcal{H}$. The *covering number* (or *transversal number*) of $\mathcal{H}$ is the smallest number $\tau = \tau(\mathcal{H})$ such that one can choose τ vertices of $\mathcal{H}$ with the property that any edge of $\mathcal{H}$ contains at least one of them. Clearly, $\nu(\mathcal{H}) \le \tau(\mathcal{H})$. There are many interesting results in integer programming showing that for some very special classes of hypergraphs equality holds here, but in general $\tau(\mathcal{H})$ is not even bonded from above by any function of $\nu(\mathcal{H})$.

One can define the "fractional" versions of these parameters, as follows. A *fractional packing* of a hypergraph $\mathcal{H} = (V(\mathcal{H}), E(\mathcal{H}))$ is a nonnegative real function p defined on the edge set $E(\mathcal{H})$ such that

$$\sum_{E \ni x} p(E) \le 1 \quad \text{for every vertex } x \in V(\mathcal{H}).$$

The maximum of $\sum_{E\in E(\mathcal{H})} p(E)$ over all fractional packings of $\mathcal{H}$ is called the *fractional packing number* of $\mathcal{H}$, and is denoted by $\nu^*(\mathcal{H})$.

In the same spirit, a *fractional covering* (or *transversal*) of $\mathcal{H}$ is a nonnegative real function t defined on the vertex set $V(\mathcal{H})$ such that

$$\sum_{x\in E} t(x) \ge 1 \quad \text{for every edge } E \in E(\mathcal{H}).$$

The minimum of $\sum_{x\in V(\mathcal{H})} t(x)$ over all fractional transversals is called the *fractional transversal number* of $\mathcal{H}$, and is denoted by $\tau^*(\mathcal{H})$. It immediately follows from the duality theorem of linear programming that

$$\nu(\mathcal{H}) \le \nu^*(\mathcal{H}) = \tau^*(\mathcal{H}) \le \tau(\mathcal{H}) \quad \text{for all } \mathcal{H}.$$

13. Examples with bounded VC-dimension

¿From any hypergraph of small VC-dimension we can construct a large variety of other hypergraphs of bounded dimension, using the following rules.

Lemma 13.1. *Let $\mathcal{H}$ be a hypergraph of* VC*-dimension d, and let $\varphi(E_1, \dots, E_k)$ be a boolean formula of k variables.*

(i) *If every edge E' of a hypergraph $\mathcal{H}'$ can be expressed as $\varphi(E_1, \dots, E_k)$ for some $E_i \in E(\mathcal{H})$, then*

$$\text{VC-dim}(\mathcal{H}') \le 2dk \log(2dk);$$

(ii) *If $\mathcal{H}^*$ denotes the dual hypergraph of $\mathcal{H}$, then*

$$\text{VC-dim}(\mathcal{H}^*) < 2^{d+1}.$$

It may be an interesting problem to decide if the above upper bounds are asymptotically tight in the worst case.

EXAMPLE 13.2 Given an n element point set $S \subseteq \mathbb{R}^d$, let $\mathcal{H}$ be a hypergraph on the vertex set S, whose edges are all possible intersections of S with a closed halfspace. It is easy to see (using Radon's theorem) that VC-dim$(\mathcal{H}) = d+1$. A similar statement is true if we intersect S with closed balls rather than halfspaces.

EXAMPLE 13.3 Given an n element point set $S \subseteq \mathbb{R}^d$, let $\mathcal{H}$ denote the hypergraph with $V(\mathcal{H}) = S$ and

$$E(\mathcal{H}) = \{S \cap P_k \mid \text{for all convex polyhedra } P_k \text{ with} \le k \text{ facets}\}$$

By Lemma 13.1 (i) and the previous example, now we have VC-dim$(\mathcal{H}) \le 2(d+1)k\log(2(d+1)k)$.

EXAMPLE 13.4 Given an n element set L of straight lines in the plane, let $\mathcal{H}$ be defined by $V(\mathcal{H}) = L$ and

$$E(\mathcal{H}) = \{L_s \mid \text{for all segment } s\},$$

where L_s denotes the set of all elements of L that intersect s. It is not difficult to show that there is a small integer $d(\ge 15)$ independent of n such that VC-dim$(\mathcal{H}) \le d$.

EXAMPLE 13.5 Given an n element set L of straight lines in the plane, let $\mathcal{H}$ be defined by $V(\mathcal{H}) = L$ and

$$E(\mathcal{H}) = \{L_\Delta \mid \text{for all triangles } \Delta\},$$

where L_Δ denotes the set of all elements of L that intersect Δ. Using Lemma 13.1 (i) again, we obtain that VC-dim$(\mathcal{H}) \le 6d\log(6d)$, where d denotes the same constant as in 13.4.

It might be interesting to note that Example 13.4 can be regarded as a dual counterpart of (the special case $d = k = 2$ of) Example 13.3, under the standard duality between points and lines in the plane. Examples 13.2 and 13.3 appear to be quite general in another sense, too. A subset $A_{k,j} \subseteq \mathbb{R}^d$ is said to be *semialgebraic* (with parameters k and j), if it consists of all points $x = (x_1, \dots, x_d)$ satisfying k algebraic inequalities of degree j, i.e.,

$$\sum_I a_I^{(t)} x^I \ge 0 \qquad (1 \le t \le k),$$

where the coefficients $a_I^{(t)}$ are real, the $\sum$ is taken over all $I = (i_1, \dots, i_d)$ with nonnegative integer coordinates whose sum is at most j, and

$$x^I = x_1^{i_1} \cdots x_d^{i_d}.$$

Given an n element point set $S \subseteq \mathbb{R}^d$, define a hypergraph $\mathcal{H}$ on the vertex set S by

$$E(\mathcal{H}) = \{S \cap A_{k,j} \mid \text{for all semialgebraic } A_{k,j} \text{ with parameters } k,\ j\}.$$

The mapping $\varphi : \mathbb{R}^d \to \mathbb{R}^{\binom{d+j}{d}-1}$ defined by

$$\varphi(x) = (\dots, x^I, \dots),$$

where we assign a coordinate to each $I = (i_1, \dots, i_d) \neq 0$, takes every $A_{k,j}$ into a set satisfying k linear inequalities

$$\left\langle a^{(t)}, \varphi(x) \right\rangle \geq -a_0^{(t)} \quad (1 \leq t \leq k).$$

Thus, φ defines a bijection between the sets $S \cap A_{k,j}$ and the sets of the form $\varphi(S) \cup P_k$, where P_k is a convex polyhedron in $\mathbb{R}^{\binom{d+j}{d}-j}$ having at most k facets. By Example 13.3, this yields that

$$\text{VC-dim}(\mathcal{H}) \leq 2\binom{d+j}{d} \log\left(2\binom{d+j}{d}k\right).$$

In most applications this upper bound is much too generous, and can be improved substantially by using ad hoc methods. However, as we shall see later, usually it is sufficient for our purposes that there is an upper bound on the VC-dimension of a hypergraph, independent of the number of its vertices. The actual value of VC-dim($\mathcal{H}$) will only effect the constants implied in our results.

The above argument might suggest that there exists a function $f(d)$ such that any hypergraph of VC-dimension d can be embedded into the $f(d)$-dimensional Euclidean space so that every edge can be obtained as the intersection of the vertex set with some convex polyhedron with at most $f(d)$ facets. However, as it was pointed out by Alon, Haussler, Welzl and Wöginger [AHWW], $f(d)$ does not exist even for $d = 2$. This fact also follows from a result of Goodman and Pollack [GP], which implies that the number of hypergraphs embeddable into $\mathbb{R}^{f(d)}$ is much smaller than the total number of hypergraphs of VC-dimension d.

14. Bounds on the transversal number of a hypergraph

V. N. Vapnik and A. Ya. Chervonenkis [VC] have discovered that hypergraphs of small VC-dimension can be statistically very well represented by taking relatively small samples from their vertex sets (see later in Section 15). Their ingenious proof technique was adapted by Haussler and Welzl [HW] to establish the upper bound

$$\tau(\mathcal{H}) \leq \left\lfloor \frac{8d}{\varepsilon} \log_2 \frac{8d}{\varepsilon} \right\rfloor$$

for the transversal number of a hypergraph $\mathcal{H}$ with VC-dimension d, all of whose edges are of size at least $\varepsilon|V(\mathcal{H})|$ for some $0 < \varepsilon < 1$. This bound was improved by Blumer, Ehrenfeucht, Haussler and Warmuth [BEHW], Shawe-Taylor, Anthony and Biggs ([STAB], see also Anthony [Ant]) to $O\left((d/\varepsilon)\log(1/\varepsilon)\right)$. We shall sketch the proof of the best known result of this type, which is also based on the original ideas of Vapnik and Chervonenkis.

Theorem 14.1. (Komlós, Pach and Woeginger [KPW]) *Let $\mathcal{H}$ be a hypergraph of* VC*-dimension d, all of whose edges have measure at least ε with respect to an arbitrary probability distribution μ on $V(\mathcal{H})$. Then*

$$\tau(\mathcal{H}) \leq \frac{d}{\varepsilon}\left(\log\frac{1}{\varepsilon} + 2\log\log\frac{1}{\varepsilon} + 3\right),$$

provided that $\varepsilon > 0$ is sufficiently small.

PROOF. Pick a sequence x of t random points from $V(\mathcal{H})$ with possible repetition, where the selections are done with respect to μ. Then choose another sequence y of $T-t$ points in the same way, and let $z = xy \in [V(\mathcal{H})]^T$. Furthermore, let $\langle z\rangle = \langle xy\rangle$ denote the multiset of all elements occurring in z, i.e., the elements are counted with multiplicities, but their order is irrelevant.

Given any $E \in E(\mathcal{H})$, let $I(E,x)$ denote the number of bits of x that belong to E. By the independence of x and y, we have

$$\mathrm{Prob}\,[\exists E \in E(\mathcal{H}) \text{ such that } I(E,x)=0] \le \frac{\mathrm{Prob}\,[\exists E \in E(\mathcal{H}) \text{ such that } I(E,x)=0,\ I(E,y)\ge m_E]}{\min_{E\in E(\mathcal{H})}\mathrm{Prob}\,[I(E,y)\ge m_E]}.$$

Choosing m_E to be the median of $I(E,y)$, we get

$$\mathrm{Prob}\,[\exists E \in E(\mathcal{H}) \text{ such that } I(E,x)=0] \le 2\mathrm{Prob}\,[\exists E \in E(\mathcal{H}) \text{ such that } I(E,x)=0,\ I(E,y)\ge m_E]\,.$$

For a fixed $E \in E(\mathcal{H})$, the conditional probability for given $\langle z\rangle = \langle xy\rangle$

$$\begin{aligned}&\mathrm{Prob}\,[I(E,x)=0, I(E,y)\ge m_E \mid \langle z\rangle]\\ &\quad\le \chi\,[I(E,z)\ge m_E]\binom{T-t}{I(E,z)}\Big/\binom{T}{I(E,z)}\\ &\quad\le \chi\,[I(E,z)\ge m_E]\left(1-\frac{t}{T}\right)^{m_E}\end{aligned}$$

By the Sauer–Shelah–VC theorem (Theorem 2.1), a fixed multiset $\langle z\rangle$ has at most $\sum_{i=0}^{d}\binom{T}{i}$ different intersections with the edges of $\mathcal{H}$. Thus, for a given $\langle z\rangle$,

$$\begin{aligned}&\mathrm{Prob}\,[\exists E \in E(\mathcal{H}) \text{ such that } I(E,x)=0, \quad I(E,y)\ge m_E \mid \langle z\rangle]\\ &\quad\le \sum_{i=0}^{d}\binom{T}{i}\left(1-\frac{t}{T}\right)^{m},\end{aligned}$$

where $m = \min_E m_E \ge (T-t)\varepsilon - 1$. Choosing $t = \lfloor (d/\varepsilon)\,(\log(1/\varepsilon) + 2\log\log(1/\varepsilon)+3)\rfloor$, $T = \lfloor (d/\varepsilon)\log^2(1/\varepsilon)\rfloor$, we get

$$\begin{aligned}&\mathrm{Prob}\,[\exists E \in E(\mathcal{H}) \text{ such that } I(E,x)=0]\\ &\quad\le 2\mathrm{Prob}\,[\exists E \in E(\mathcal{H}) \text{ such that } I(E,x)=0, \quad I(E,y)\ge m_E]\\ &\quad\le 2\sum_{i=0}^{d}\binom{T}{i}\left(1-\frac{t}{T}\right)^{(T-t)\varepsilon-1} < 1.\end{aligned}$$

Hence, there exists a sequence x of t points intersecting every edge of $\mathcal{H}$. □

Most often we apply Theorem 14.1 for the uniform distribution on $V(\mathcal{H})$, i.e., when $\mu(E) = |E|/|V(\mathcal{H})|$.

Corollary 14.2. *Let $\mathcal{H}$ be a hypergraph of* VC*-dimension d, and let ε be a fixed small positive number. Then there exists an at most $\frac{d}{\varepsilon}\left(\log\frac{1}{\varepsilon}+2\log\log\frac{1}{\varepsilon}+3\right)$-element subset $N \subseteq V(\mathcal{H})$ which intersects every edge $E \in E(\mathcal{H})$ with $|E| \geq \varepsilon|V(\mathcal{H})|$.*

A subset $N \subseteq V(\mathcal{H})$ with the above property is usually called an *ε-net* for $\mathcal{H}$.

Another interesting special case of Theorem 14.1 is the following. Let t be a fractional transversal of $\mathcal{H}$ with

$$\sum_{x\in V(\mathcal{H})} t(x) = \tau^*(\mathcal{H}).$$

Then $\mu(x) = t(x)/\tau^*(\mathcal{H})$, $x \in V(\mathcal{H})$ is a probability distribution, and

$$\mu(E) = \frac{\sum_{x\in E} t(x)}{\tau^*(\mathcal{H})} \geq \frac{1}{\tau^*(\mathcal{H})} \text{ for every } E \in E(\mathcal{H}).$$

Theorem 14.1 now yields the following result.

Corollary 14.3. *For any hypergraph $\mathcal{H}$ of* VC*-dimension d,*

$$\tau(\mathcal{H}) \leq d\tau^*(\mathcal{H})(\log\tau^*(\mathcal{H}) + 2\log\log\tau^*(\mathcal{H}) + 3),$$

provided that $\tau^*(\mathcal{H})$ is sufficiently large.

Given a hypergraph $\mathcal{H}$, we can define many other interesting parameters similar to the VC-dimension. For instance, for any fixed k, let $d_k(\mathcal{H})$ denote the maximum size of a subset $A \subseteq V(\mathcal{H})$ with the property that for any k element subset $B \subseteq A$ there exists $E_B \in E(\mathcal{H})$ such that $E_B \cap A = B$. Obviously, VC-dim$(\mathcal{H}) \leq d_k(\mathcal{H})$ for every k. Combining Corollary 14.3 with Ramsey's theorem, Ding, Seymour and Winkler [DSW] established the following.

Theorem 14.4. (Ding, Seymour, Winkler [DSW]) *For any hypergraph $\mathcal{H}$,*

$$\tau(\mathcal{H}) \leq 6d_2^2(\mathcal{H}^*)(d_2(\mathcal{H}^*)+\nu(\mathcal{H}))\binom{d_2(\mathcal{H}^*)+\nu(\mathcal{H})}{\nu(\mathcal{H})}^2,$$

where $\mathcal{H}^*$ denotes the dual of $\mathcal{H}$.

Ding et al [DSW] have also established the following geometric corollary to the above bound. Let P and $\mathcal{R}$ be a set of points and a set of axis-parallel rectangles in the plane, respectively. Then, for any natural number k, either there are $k+1$ rectangles in $\mathcal{R}$ so that no point of P belongs to more than one of them, or one can find at most $(k+63)^127$ points in P meeting every member of $\mathcal{R}$. This has been improved recently by Pach and Tőrőcsik [PT].

In many interesting geometric applications of Theorem 14.1 and Corollary 14.2 the logarithmic terms can be replaced by constants. (See Agarwal [Ag], Chazelle and Friedman [CF], Clarkson [C], Matoušek,, Seidel and Welzl [MSW], Pach and Woeginger [PW].) However, in the combinatorial setting these results are not far from being best possible.

Theorem 14.5. (Komlós, Pach, Woeginger [KPW]) *Given any natural number $d \geq 2$ and $\gamma < 2/(d+2)$, for any sufficiently small $\varepsilon > 0$ one can find a hypergraph $\mathcal{H}$, all of whose edges are of size at least $\varepsilon|V(\mathcal{H})|$ and*

$$\tau(\mathcal{H}) \geq \frac{d-2+\gamma}{\varepsilon}\log\frac{1}{\varepsilon}.$$

15. Discrepancies and ε-approximations

Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph with n vertices and m edges, and let $c : V \to \{-1, +1\}$ be a mapping. This defines a partition of the vertices $V = V^+ \cup V^-$. We will call such a partition a *coloring* of $\mathcal{H}$. For a set $S \subset V$ let $c(S) = \frac{1}{2}(\sum_{p\in S} c(p))$, i.e., the difference between the number of elements with the same sign and $|S|/2$. Define the *discrepancy of* c on $\mathcal{H}$ by

$$\operatorname{disc}(\mathcal{H}, c) = \max_{E\in\mathcal{E}} \frac{1}{2}||V^+ \cap E| - |V^- \cap S||,$$

and the *discrepancy of* $\mathcal{H}$ by

$$\operatorname{disc}(\mathcal{H}) = \min_{c:V\to\{-1,+1\}} \operatorname{disc}(\mathcal{H}, c).$$

The discrepancy of a hypergraph $\mathcal{H}$ with n vertices and m edges is at most $\operatorname{disc}(\mathcal{H}) \leq 2\sqrt{n \log m}$ ($m \geq 4$). If $m \sim n^k$, then this gives $O(\sqrt{n \log n})$, and a random construction shows that this bound is best possible. However, much better bounds hold if the hypergraph has bounded VC-dimension. In fact, a somewhat weaker assumption will already enable us to show that $\operatorname{disc}(\mathcal{H}) = o(\sqrt{n})$.

The *primal density function* $\pi_{\mathcal{H}} : [n] \to \mathbb{R}$ of a hypergraph $\mathcal{H}$ is defined by

$$\pi_{\mathcal{H}}(s) = \max_{S\subset V,\, |S|\leq s} |\{E \cap S : E \in \mathcal{E}\}|.$$

The *dual density function* $\pi^*_{\mathcal{H}} : [m] \to \mathbb{R}$ is the primal density function of the dual hypergraph of $\mathcal{H}$ arising by exchanging the role of points and edges. Thus, $\pi^*_{\mathcal{H}}(t)$ is the maximum number of atoms into which the points of V can be partitioned by a collection of t edges of $\mathcal{E}$. By Theorem 2.1, and Lemma 13.1(ii), hypergraphs with bounded VC-dimension have polynomially bounded density (and dual density) functions. For example, consider $(P, \mathcal{B})$ where P is a finite set of points in $\mathbb{R}^d$ and $\mathcal{B}$ is the set of intersections of P with balls. Then the primal density function is of order $O(s^{d+1})$, while the dual density is of order $O(t^d)$ and the VC-dimension is $d + 1$.

A subset S is an *ε-approximation* for $\mathcal{H}$ if

$$\left| \frac{|E \cap S|}{|S|} - \frac{|E|}{|V|} \right| \leq \varepsilon$$

for every set $E \in \mathcal{E}$.

Theorem 15.1. (Vapnik and Chervonenkis [VC]) *Let d be fixed and let $\mathcal{H}$ be a hypergraph of* VC*-dimension (at most) d. Then for every $r > 1$, there exists a $(1/r)$-approximation for $\mathcal{H}$ of size $O(r^2 \log r)$.*

The following theorem greatly expands Theorem 15.1 and the discrepancy bounds for hypergraphs with polynomially bounded (dual) density functions.

Theorem 15.2. (Matoušek, Welzl and Wernisch [MWW]) *Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph on n vertices, d and C constants. Suppose that $\pi_{\mathcal{H}}(s) \leq Cs^d$ for all $s \leq n$. Then the discrepancy* $\operatorname{disc}(\mathcal{H})$ *of $\mathcal{H}$ is bounded by*

$$O(n^{(1/2)-(1/2d)}(\log n)^{1+(1/2d)}),\ \textit{if } d > 1,\ \textit{and } O(\log^{5/2} n),\ \textit{if } d = 1.$$

Moreover, for every $r \leq n$, there exists a $(1/r)$-approximation for $\mathcal{H}$ of size

$$O(r^{2-(2/(d+1))}(\log r)^{2-1/(d+1)}),\ \textit{if } d > 1,\ \textit{and } O(r \log^{5/2} r),\ \textit{if } d = 1.$$

Theorem 15.3. (Matoušek, Welzl and Wernisch [MWW]) *Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph on n vertices, d and C constants. Suppose that $\pi^*_{\mathcal{H}}(t) \leq Ct^d$ for all $t \leq n$. Then the discrepancy* $\operatorname{disc}(\mathcal{H})$ *of $\mathcal{H}$ is bounded by*

$$O(n^{(1/2)-(1/2d)} \log n), \text{ if } d > 1, \text{ and } O(\log^{3/2} n), \text{ if } d = 1.$$

Moreover, for every $r \leq n$, there exists a $(1/r)$-approximation for $\mathcal{H}$ of size

$$O(r^{2-(2/(d+1))}(\log r)^{2-2/(d+1)}), \text{ if } d > 1, \text{ and } O(r \log^{3/2} r), \text{ if } d = 1.$$

These beautiful results allow to rederive and extend many of the upper bounds of Beck [BC] about geometric discrepancies. For example, if $P \subset \mathbb{R}^d$ is an n-set and $\mathcal{B}$ is a family of intersections of P with balls, then

$$\operatorname{disc}(\mathcal{B}) = O(n^{(1/2)-(1/2d)} \sqrt{\log n}).$$

This result holds for every probability measure μ of $\mathbb{R}^d$. This bound can easily be generalized to the case when $\mathcal{B}$ is the set of k-fold boolean combination of balls, too.

The dual density function seems to be more appropriate to investigate geometric discrepancies, because it depends on the complexity of the figures determining the subsets rather than the dimension of the space.

Finally, we remark that Beck [BC] and, by a different method, Alexander [Ale] gave almost matching lower bounds (up to logarithmic factors) for all the discrepancy bounds of Theorems 15.2 and 15.3.

The proof of Theorem 15.3 is based on the following result about the *crossing number* of hypergraphs. We say that an edge E crosses the pair $\{x, y\}$ if $|E \cap \{x, y\}| = 1$. A *spanning path*, P, of $\mathcal{H}$ is a linear order of the vertices $V = \{v_1, v_2, \ldots, v_n\}$. The *crossing number* of E with respect to P is the number of pairs $\{v_i, v_{i+1}\}$ crossed by E. The crossing number of a hypergraph $\mathcal{H} = (V, \mathcal{E})$ with respect to P is the maximum of the crossing numbers of the edges $E \in \mathcal{E}$.

Theorem 15.4. (Chazelle and Welzl [CW]) *Suppose that the dual density function $\pi^*_{\mathcal{H}}(t)$ is bounded by $\pi^*_{\mathcal{H}}(t) \leq Ct^d$ for some constants C and d. Then there exists a spanning path P with crossing number $O(n^{1-(1/d)} \log n)$, if $d > 1$ and $O(\log^2 n)$, if $d = 1$.*

This result also shows that the use of intervals in constructing extremal families in Section 5 is probably not a coincidence.

The discrepancy of a *matrix* A is

$$\operatorname{disc}(A) = \min_{\mathbf{x} \text{ is a } (-1/2, +1/2) \text{ vector}} ||A\mathbf{x}||_\infty$$

and the hereditary discrepancy of A is

$$\operatorname{herdisc}(A) = \max_{B \text{ is a submatrix of } A} \operatorname{disc}(B).$$

For further definitions, see [LSV]. Note that if A is the incidence matrix of a hypergraph $\mathcal{H}$, then $\operatorname{disc}(A) = \operatorname{disc}(\mathcal{H})$. Furthermore, all entries of an integer matrix A with $\operatorname{herdisc}(A) \leq d$ are between $-2d$ and $2d$. Let $h(n, d)$ denote the maximum number of distinct rows in an $m \times n$ integer matrix A with $\operatorname{herdisc}(A) \leq d$.

Theorem 15.5. (Lovász and Vesztergombi [LV])

$$\sum_{0\le j\le n}\binom{n}{j}\binom{n+j}{j}\binom{2d}{d}\le h(n,d)\le \sum_{0\le j\le n}\binom{n}{j}\frac{(2\pi d)^j}{\left(\Gamma\left(\frac{j}{2}+1\right)\right)^2}$$

where Γ *is the usual gamma function.*

Anstee [An90] showed, using Theorem 2.1, that for a fixed d, $h(n,d)$ is bounded by a polynomial of n. More precisely, $h(n,d)=O(n^{7d^2+3d})$. This is the first step to establish the conjecture of Lovász and Vesztergombi [LV], $h(n,d)=O(n^{4d})$.

16. A problem of Danzer and Rogers

Problem 16.1. *Determine the smallest number* $f(\varepsilon)$ *of points that can be arranged in the unit square* $[0,1]^2$ *so that every convex set* $C\subseteq[0,1]^2$ *with area* $\ge\varepsilon$ *contains at least one of them.*

Let $\varepsilon>0$ be fixed. Then we can choose a sufficiently large m with the property that any convex set $\Delta\subseteq[0,1]^2$ with area $\ge\varepsilon/3$ contains at least $\varepsilon m^2/6$ points belonging to the m by m square grid $S_{m\times m}$ placed on the unit square. Construct a hypergraph $\mathcal{H}$ with $V(\mathcal{H})=S_{m\times m}$, whose edges are all sets of the form $S_{m\times m}\cap\Delta$, where $\Delta\subseteq[0,1]^2$ is a triangle of area at least $\varepsilon/3$. By Example 13.3, VC-dim$(\mathcal{H})\le 6\log 18<18$, and every edge of $\mathcal{H}$ has at least $\varepsilon m^2/6=(\varepsilon/6)|V(\mathcal{H})|$ points. Thus, Corollary 14.2 implies that there exists an at most $(18/(\varepsilon/6))\left(\log(6/\varepsilon)+2\log\log(6/\varepsilon)+3\right)=O\left((1/\varepsilon)\log(1/\varepsilon)\right)$ element set $N\subseteq S_{m\times m}$ intersecting all triangles Δ of area $\varepsilon/3$. Since every convex set $C\subseteq[0,1]^2$ of area at least ε contains such a triangle, we obtain

$$f(\varepsilon)=O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right).$$

Note that this relation can also be shown by a straightforward construction. However, there does not appear to exist an equally easy construction establishing the same bound in higher dimensions, without using Corollary 14.2 or some other probabilistic argument. It is an exciting open problem to decide whether $f(\varepsilon)=O\left(1/\varepsilon\right)$. The answer is probably in the negative.

17. Bounds on incidences

The fact that a large variety of hypergraphs arising in geometric problems has bounded VC-dimension was already noted by A. Rényi, C. Rényi and J. Surányi [RRS] in the 1950's. Although it has been known for more than twenty years that this concept plays a crucial role in mathematical statistics, the relevance of the methods discussed in previous sections to a wide range of problems in discrete and computational geometry was realized only a couple of years ago by Haussler and Welzl.

In what follows, we shall outline an elegant argument due to Clarkson, Edelsbrunner, Guibas, Sharir and Welzl [CEGSW], which provides an elegant alternative proof of a well–known result of Szemerédi and Trotter [ST] on the number of incidences between points and lines in the plane. It is a typical application of the so–called "ε-net technique" based on Corollary 14.2.

Theorem 17.1. (Szemerédi, Trotter [ST]) *There exists a constant c such that the maximum number of incidences, I, between m distinct points and n distinct lines in the plane,*

$$I(m,n) \le c(m^{2/3}n^{2/3} + m + n).$$

Proof. [CEGSW] Let $S = \{p_1, \dots, p_m\}$ and $L = \{\ell_1, \dots, \ell_n\}$ be given sets of points and lines in the plane, respectively. Construct a bipartite graph G with $V(G) = S \cup L$, where p_i and ℓ_j are joined by an edge if and only if p_i is incident to ℓ_j. Obviously, G does not contain a cycle of length 4, so we can apply an old result of Erdős and Kővári, Sós and Turán [EKST] to obtain that

$$I(m,n) = \max_G |E(G)| \le c'(mn^{1/2} + n), \tag{17.1}$$

where c' is a suitable constant.

Define a hypergraph $\mathcal{H}$ on the vertex set $V(\mathcal{H}) = L$ by

$$E(\mathcal{H}) = \{L_\Delta \mid \text{ for all triangles } \Delta\},$$

where L_Δ denotes the set of all elements of L intersecting a triangle Δ. By Example 13.5, VC-dim$(\mathcal{H}) \le 500$ (but, of course, this bound can be substantially lowered).

According to Corollary 14.2, for any sufficiently small $\varepsilon > 0$, $\mathcal{H}$ has an ε-net of size at most $(10^3/\varepsilon)\log(1/\varepsilon)$. That is, there exists an at most $(10^3/\varepsilon)\log(1/\varepsilon)$ element subset $L' \subseteq L$ with the property that any triangle Δ with $|L_\Delta| \ge \varepsilon n$ intersects at least one element of L'. The lines of L' partition the plane into some polygonal regions, which can be further subdivided into at most

$$5|L'|^2 < \frac{10^7}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^2$$

triangles Δ_i $(i = 1, 2, \dots)$. Note that here we also allow Δ_i to be a degenerate triangle having only two sides, i.e., an infinite cone. Using an additional trick, we can spare the logarithmic factor, i.e., the number of these triangles can be reduced to $10^7/\varepsilon^2$. Let n_i denote the number of lines of L intersecting the interior of Δ_i. By our construction, $n_i < \varepsilon n$ for every i. Let m_i denote the number of points of S belonging to Δ_i. (If a point of S is lying on the boundary of several triangles, then it is counted only at one of them.) Thus, $\sum m_i = m$, and we have that the number of incidences between S and L,

$$I(m,n) \le \sum_{i=1}^{10^7/\varepsilon^2} I(m_i, n_i) + \left(\frac{10^3}{\varepsilon}n + m\right),$$

where the second term bounds the number of incidences between S and L'. Using (17.1), we get

$$\begin{aligned} I(m,n) &\le c' \sum_{i=1}^{10^7/\varepsilon^2} (m_i(\varepsilon n)^{1/2} + \varepsilon n) + \frac{10^3}{\varepsilon}n + m \\ &\le c''\left(m(\varepsilon n)^{1/2} + \frac{n}{\varepsilon} + m\right). \end{aligned}$$

If $n \geq m^2$, then (17.1) yields that $I(m,n) \leq 2c'n$. Otherwise, let $\varepsilon = \max\big((n/m^2)^{1/3}, 1/n\big)$, and Theorem 17.1 follows from the last inequality. □

The main advantage of this proof technique lies in its flexibility. In particular, it allows us to generalize the above theorem without too much effort in many different directions.

Let Γ be a class of curves in the x-y plane defined in terms of d real parameters. We say that Γ is a *regular class* of curves with d degrees of freedom if there exists an integer s such that

(a) the dependence of the curves on x, y and the parameters is algebraic of degree at most s;
(b) no two disjoint curves intersect in more than s points;
(c) for any d points in the plane, there are at most s element of Γ passing through all of them.

Theorem 17.2. (Pach and Sharir [PS]) *Let Γ be a regular class of curves with d degrees of freedom. Then there exists a constant $c = c_\Gamma$ such that the number of incidences between m distinct points and n distinct curves in Γ is at most*

$$c(m^{d/(2d-1)}n^{(2d-2)/(2d-1)} + m + n).$$

Since the class of all unit circles is regular with $d = 2$, we immediately obtain the following result.

Corollary 17.3. (Spencer, Szemerédi and Trotter [SST]) *The maximum number of times that the unit distance can occur among n points in the plane is at most $cn^{4/3}$.*

For further proofs and geometric applications see [PA].

References

[Ag] P. K. Agarwal, Partitioning arrangements of lines I–II, *Discr. Comput. Geom.* **5** (1990), 449–483, 553–573.

[Ale] R. Alexander, Geometric methods in the theory of uniform distribution, *Combinatorica* **10** (1990), 115–136.

[A] N. Alon, On the density of sets of vectors, *Discrete Mathematics* **46** (1983), 199-202.

[AF] N. Alon and Z. Füredi, On the kernel of intersecting families, *Graphs and Combinatorics* **3** (1987), 91–94.

[AHWW] N. Alon, D. Haussler, E. Welzl and G. Wöginger, Partitioning and geometric embedding of range spaces of finite Vapnik–Chervonenkis dimension. *Proc. 3rd ACM Symp. Comput. Geom.* (1987), 331–340.

[An80] R. Anstee, Properties of (0,1)-matrices with no triangles, *Journal of Combinatorial Theory, Ser. A* **29** (1980), 186–198.

[An83] R. Anstee, Hypergraphs with no special edges, *Combinatorica* **3** (1983), 141–146.

[An83] R. Anstee, Forbidden configurations, discrepancy and determinants, *European Journal of Combinatorics* **11** (1990), 15–19.

[AnF] R. Anstee and Z. Füredi, Forbidden submatrices, *Discrete Mathematics* **62** (1986), 225–243.

[AM] R. Anstee and U. S. R. Murty, Matrices with forbidden subconfigurations, *Discrete Mathematics* **54** (1985), 113–116.

[Ant] M. Anthony, Uniform Convergence and Learnability, Ph. D. Thesis, London School of Economics, Reprint Series in Mathematics, No. 11, 1991.

[BC] J. Beck and W. Chen, *Irregularities of distribution*, Cambridge University Press, 1987.

[BG] D. Bienstock and E. Győri, An extremal problem on sparse 0–1 matrices, *SIAM J. Discrete Math.* **4** (1991), 17–27.

[BEHW] A. Blumer, A. Ehrenfeucht, D. Haussler and M. Warmuth, Classifying learnable geometric concepts with the Vapnik–Chervonenkis dimension. *J. of ACM* **36** (1989), 929–965.

[B] B. Bollobás, On generalized graphs, *Acta Math. Hungar.* **16** (1965), 447–452.

[Bol] B. Bollobás, unpublished, see in [L] Problem 13.10.

[Bon] J. A. Bondy, Induced subsets, *Journal of Combinatorial Theory, Ser. B* **12** (1972), 201–202.

[BH] J. A. Bondy and A. Hajnal, Problem 5, *in* Hypergraph Seminar, (C. Berge and D. K. Ray-Chaudhuri, Eds.), Columbus, Ohio, 1972. *Lecture Notes in Mathematics* **411**, Springer-Verlag, 1974, p. 279.

[dC] de Caen, D., The current status of Turán's problem on hypergraphs, see in this volume

[C-K] M. Całczyńska-Karłowitz, Theorem on families of finite sets, *Bull. Acad. Polon. Sci. Ser. Math. Astr. Phys.* **12** (1964), 87–89.

[CF] B. Chazelle and J. Friedman, A deterministic view of random sampling and its use in geometry, *Combinatorica* **10** (1990), 229–249.

[CW] B. Chazelle and E. Welzl, Quasi–optimal range searching in spaces of finite VC-dimension, *Discr. Comput. Geom.* **4** (1989), 467–490.

[C] K. Clarkson, New applications of random sampling in computational geometry, *Discr. Comput. Geom.* **2** (1987), 195–222.

[CEGSW] K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir and E. Welzl, Combinatorial complexity bounds for arrangements of curves and surfaces, *Discr. Compact. Geom.* **2** (1987), 195–222.

[DSW] G. Ding, P. Seymour and P. Winkler, Bounding the vertex cover number of a hypergraph, *Combinatorica*, submitted in 1991.

[EH] H. Edelsbrunner and P. Hajnal, A lower bound on the number of unit distances between the vertices of a convex polygon, *Journal of Combinatorial Theory, Ser. A* **56** (1991), 312–316.

[EKR] P. Erdős, C. Ko, and R. Rado, An intersecting theorem for finite sets, *Quart. J. Math.* **12** (1961), 313–320.

[EKST] P. Erdős and T. Kővári, V. T. Sós and P. Turán, On a problem of Zarankiewicz, *Coll. Math.* **3** (1954), 50–57.

[F78] P. Frankl, On a problem of Bondy and Hajnal, *in* Problèmes Combinatoires et Théorie des Graphes, (Orsay, France, 1976), *Colloq. Internat. C.N.R.S.*, No. **260** (1978), p. 159.

[F82] P. Frankl, An extremal problem for two families of sets, *European Journal of Combinatorics* **3** (1982), 125–127.

[F83] P. Frankl, On the traces of finite sets, Journal of Combinatorial Theory, Ser. A **34** (1983), 41–45.

[F84] P. Frankl, A new short proof for the Kruskal–Katona theorem, *Discrete Mathematics* **48** (1984), 327–329.

[FFP] P. Frankl, Z. Füredi and J. Pach, Bounding one-way differences, *Graphs and Combinatorics* **3** (1987), 341–347.

[FP83] P. Frankl and J. Pach, On the number of sets in a null t-design, *European Journal of Combinatorics* **4** (1983), 21–23.

[FP84] P. Frankl and J. Pach, On disjointly representable sets, *Combinatorica* **4** (1984), 39–45.

[F88] Z. Füredi, Matchings and covers in hypergraphs, *Graphs and Combinatorics* **4** (1988), 115–206.

[F90] Z. Füredi, The maximum number of unit distances in a convex n-gon, Journal of Combinatorial Theory, Ser. A **55** (1990), 316–320.

[F91] Z. Füredi, Turán type problems, *in* Surveys in Combinatorics. 1991, (Proc. of the 13th British Combinatorial Conference, ed. A. D. Keedwell), Cambridge Univ. Press. *London Math. Soc. Lecture Note Series* **166** (1991), 253–300.

[FH] Z. Füredi and P. Hajnal, Davenport-Schinzel theory of matrices, *Discrete Mathematics* **103** (1992), 233–251.

[FQ] Z. Füredi and F. Quinn, Traces of finite sets, *Ars Combinatoria* **18** (1984), 195–200.

[FT] Z. Füredi and Zs. Tuza, Hypergraphs without a large star, *Discrete Mathematics* **55** (1985), 317–321.

[GP] J. E. Goodman and R. Pollack, Upper bounds for configurations and polytopes in $\mathbb{R}^d$, *Discr. Comput. Geom.* **1** (1986), 219–227.

[HW] D. Haussler and E. Welzl, ε-nets and simplex range queries, *Discr. Comput. Geom.* **2** (1987), 127–151.

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

[KM] M. G. Karpovsky and V. D. Milman, Coordinate density of sets of vectors, *Discrete Mathematics* **24** (1978), 177-184.

[Ka] G. O. H. Katona, A theorem on finite sets, *in:* Theory of Graphs, (Proc. Colloq., Tihany, Hungary, 1966, P. Erdős et al., eds.), Akadémiai Kiadó, Budapest and Academic Press, New York, 1968, pp. 187–207.

[Ka] G. O. H. Katona, Optimization for order ideals under weight assignments, *in* Problèmes Combinatoires et Théorie des Graphes, (Orsay, France, 1976), *Colloq. Internat. C.N.R.S.*, No. **260** (1978), pp. 257–258.

[Kar] M. Karchmer, private communication, MIT, 1991.

[Ko] Y. Kohayakawa, A note on kernels of intersecting families, *European J. Combinatorics* **11** (1990), 155–164.

[KPW] J. Komlós, J. Pach and G. Woeginger, Almost tight bounds for epsilon–nets. *Discr. Comput. Geom.* **7** (1992), to appear.

[Kr] J. B. Kruskal, The number of simplices in a complex, *in* Mathematical Optimization Techniques (ed. R. Bellmann), Univ. California Press, Berkeley, 1963, pp. 251–278.

[L] L. Lovász, *Combinatorial Problems and Exercises*, North-Holland, Amsterdam, 1979.

[LSV] L. Lovász, J. Spencer and K. Vesztergombi, Discrepancy of set systems and matrices, *European J. Combinatorics* **7** (1986), 151–160.

[LV] L. Lovász and K. Vesztergombi, Extremal problems for discrepancy, *in* Irregularities of partitions (Proc. Colloq., Fertőd, Hungary, 1986), pp. 107–113, (Algorithms and Combin.: Study Res. Texts) **8** (1989), Springer, Berlin–New York.

[MSW] J. Matoušek, R. Seidel and E. Welzl, How to net a lot with little: Small ε-nets for disks and halfspaces, *Proc. 6th ACM Symp. Comput. Geom.*, 1990.

[MWW] J. Matoušek, E. Welzl and L. Wernisch, Discrepancy and approximations for bounded VC-dimension, *Combinatorica* to appear.

[PA] J. Pach and P. K. Agarwal, *Combinatorial Geometry*, manuscript, 223+*vi* pages, DIMACS Technical Report 91-51, New Brunswick, NJ, July, 1991.

[PS] J. Pach and M. Sharir, Repeated angles in the plane and related problems, *J. Combinat. Th. Ser. A*, **59** (1992), 12-22.

[PT] smc J. Pach and J. Törőcsik, Some geometric applications of Dilworth's theorem, *Discr. Computat. Geom.* to appear.

[PW] J. Pach and G. Woeginger, Some new bounds for ε-nets, *Proc. 6th Symp. on Computational Geometry*, (1990), 10–15.

[P] M. Perles, unpublished, see [Sh]

[RRS] A. Rényi, C. Rényi and J. Surányi, Sur l'independence des domaines simples dans l'espace Euclidean a'n dimensions, *Colloq. Math.* **2** (1951), 130–135. See also in: *Selected Papers of A. Rényi* (edited by P. Turán), Vol. 1., Akadémiai Kiadó, Budapest, 1976.

[R72] H. J. Ryser, A fundamental matrix equation for finite sets, *Proc. Amer. Math. Soc.* **34** (1972), 332–336.

[S] N. Sauer, On the density of families of sets, Journal of Combinatorial Theory, Ser. A **13** (1972), 145–147.

[STAB] J. Shawe-Taylor, M. Anthony and N. L. Biggs, Bounding sample size with the Vapnik–Chervonenkis dimension, Tech. Report CSD-TR-618, Dept. Comp. Sci., Royal Holloway and Bedford New College, London University, 1989.

[Sh] S. Shelah, A combinatorial problem: stability and order for models and theories in infinitary languages, *Pacific J. Math.* **41** (1972), 247–261.

[SST] J. Spencer, E. Szemerédi and W. T. Trotter, Unit distances in the Euclidean plane. In: *Graph Theory and Combinatorics* (B. Bollobás, ed.), Academic Press, New York, 293–303.

[ST] E. Szemerédi and W. T. Trotter, A combinatorial distinction between the Euclidean and projective planes, *Europ. J. Combin.* **4** (1983), 385–394.

[T] P. Turán, On an extremal problem in graph theory, (Hungarian), *Mat. Fiz. Lapok* **48**, 436–452. also see [dC]

[Tu] Zs. Tuza, Critical hypergraphs and intersecting set-pair systems, Journal of Combinatorial Theory, Ser. B **39** (1985), 134–145.

[VC] V. N. Vapnik and A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.* **16** (1971), 264–280.

[W] Mika Watanabe, Arrow relations on families of finite sets, *Discrete Mathematics* **94** (1991), 53–64.