LECTURES ON EXTREMAL SET SYSTEMS AND TWO-COLORINGS OF HYPERGRAPHS

Gy. Károlyi

Department of Algebra, Eötvös University, Budapest
Institute of Discrete Mathematics, University of Bonn

Contents

1	**Introduction**	
3	**Part A**	**Extremal Set Systems**
3	**Lecture 1**	**Intersection Conditions**
7	**Lecture 2**	**Sperner Systems**
11	**Lecture 3**	**The Littlewood–Offord Problem**
14	**Lecture 4**	**The Kruskal–Katona Theorem**
18	**Part B**	**Two-colorings of Hypergraphs**
18	**Lecture 5**	**Blocking Sets I**
21	**Lecture 6**	**Blocking Sets II**
25	**Lecture 7**	**Discrepancy I**
30	**Lecture 8**	**Discrepancy II**

Introduction

The aim of these lectures is to give a short introduction to two developing areas of combinatorics concerning hypergraphs.

In the first part we present the main classical results in the theory of extremal set systems. The basic question is that how large can a family $\mathcal{H}$ of subsets of an underlying set X with n elements be if it satisfies a given intersection, union or inclusion property. We are also interested in what the set systems having maximum size look like. We try to illustrate the most fruitful methods of the theory. Furthermore, in Section 3, we also deal with a nice application to a geometric problem.

In the second part we consider two approaches for 2-colorings of hypergraphs. On the one hand we look for conditions that guarantee the existence of blocking sets — 2-colorings of X without monochromatic members of $\mathcal{H}$. On the other hand we investigate the possibility of a balanced 2-coloring: we would like to divide X into two parts such that these parts divide the elements of $\mathcal{H}$ as evenly as possible. Of course we cannot expect that every member of $\mathcal{H}$ contains the same number of elements of the two color classes. Instead of this we investigate the discrepancy of the hypergraph — the measure of inevitable irregularities. Combining various ideas we will demonstrate the strength of the so-called probabilistic method, a powerful tool in combinatorics.

Throughout these notes by a hypergraph or set system we always mean a family of subsets of the fixed underlying set $X = \{1, 2, \ldots, n\}$. We will denote by 2^X the family consisting of all subsets of X. We will denote hypergraphs by script letters $\mathcal{F}, \mathcal{G}, \mathcal{H}$ and so on. Block capitals $A, B, C \ldots$ will denote subsets of X. The elements of X are often called points or vertices, the members of $\mathcal{H}$ are the edges of the hypergraph $\mathcal{H}$. $\mathcal{H}$ is r-uniform if all of its edges contain exactly r vertices. Thus hypergraphs may be considered as a generalization of graphs: the simple graphs are the 2-uniform hypergraphs.

References

Here we give only the four main sources of the material we treated. The reader can find particular references and further readings at the end of each Lecture.

Part A

I. Anderson: *Combinatorics of Finite Sets, Chapters 1–3,5,7,11.*
(The Clarendon Press, Oxford Univ. Press, N.Y., 1987)

L. Lovász: *Combinatorial Problems and Exercises, Chapter 13.*
(Akadémiai Kiadó, Budapest, 1979)

Part B

N. Alon and J.H. Spencer: *The Probabilistic Method, Chapters 1–3,5,12.*
(Wiley-Interscience, 1991)

L. Lovász: *Combinatorial Problems and Exercises, Chapters 2,13.*
(Akadémiai Kiadó, Budapest, 1979)

J. Spencer: *Ten Lectures on the Probabilistic Method, Chapters 4,5,8,9.*
(SIAM, 1987)

Part A
Extremal Set Systems

1 Intersection Conditions

As warming-up we investigate the following simple problem which allows us to get an insight to the nature of our subject.

Question. *A hypergraph $\mathcal{H}$ has no disjoint edges. How big can $|\mathcal{H}|$ be?*

Answer. $|\mathcal{H}| \leq 2^{n-1}$. Indeed, if $A \subseteq X$ is an edge of $\mathcal{H}$, then $X \setminus A \notin \mathcal{H}$. Therefore the elements of 2^X form 2^{n-1} pairs and $\mathcal{H}$ may contain at most one member of each pair. On the other hand, there are 2^{n-1} subsets of X containing a fixed element of X, and these subsets clearly satisfy the intersection condition.

Exercise. *Suppose we are given a hypergraph $\mathcal{H}$ with the intersection property, and $|\mathcal{H}| < 2^{n-1}$. Prove that $\mathcal{H}$ can be extended to a collection of 2^{n-1} subsets of X also satisfying this property.*

Next we consider r-uniform hypergraphs with the intersection property.

If $r > \frac{n}{2}$, then the question is not interesting because there are no disjoint r-element subsets of X.

Theorem 1.1 (Erdős–Ko–Rado). *Let $\mathcal{H}$ be an r-uniform hypergraph $(n \geq 2r)$ the edges of which are pairwise intersecting. Then $|\mathcal{H}| \leq \binom{n-1}{r-1}$.*

Proof. Katona proved this theorem with the help of "cyclic permutations". If we count the ways the numbers $1, 2, \ldots, n$ can be arranged along a circle we find $(n-1)!$ different ways, these are called the cyclic permutations of the elements $1, 2, \ldots, n$. Indeed, we can identify n different permutations to obtain the same arrangement (e.g. $2, 1, 3, \ldots, n$ with $1, 3, \ldots, n, 2$). A member of $\mathcal{H}$ may be considered in a cyclic permutation, as r consecutive elements. More presicely, an r-element subset of X is contained in exactly $r!(n-r)!$ cyclic permutations. On the other hand, a cyclic permutation can contain at most r different members of $\mathcal{H}$, this is the point where we use the intersection condition, noting that $r \leq \frac{n}{2}$. (See the exercise below.) Summarizing these observations we obtain the estimate

$$|\mathcal{H}|r!(n-r)! \leq r(n-1)! \ ,$$

$$|\mathcal{H}| \leq \frac{r(n-1)!}{r!(n-r)!} = \binom{n-1}{r-1}.$$

To see that this inequality is sharp, consider the r-element subsets of $\mathcal{H}$ containing a fixed element of X.

Exercise. *Prove that a cyclic permutation contains at most r distinct edges of the hypergraph $\mathcal{H}$.*

In the remaining part of this section we consider set systems with stronger intersection properties. A basic fact which plays a role for example in the theory of block designs is the well-known Fisher's inequality, that we state in a more general setting.

Theorem 1.2. *Suppose that $\emptyset \neq \mathcal{H}$ and the intersection of any two distinct edges of $\mathcal{H}$ has the same cardinality λ. Then $|\mathcal{H}| \leq n$.*

Proof. It is obvious if $\lambda = 0$, hence we may suppose $\lambda \geq 1$. Let $\mathcal{H} = \{A_1, \ldots, A_m\}$, we may suppose $|A_i| > \lambda$ for $i = 1, 2, \ldots, m$. We will carry out the proof using elementary linear algebra. We may associate to each A_i a $0-1$ vector $\mathbf{v}_i$ of length n: the j-th coordinate $v_{i,j}$ of $\mathbf{v}_i$ is 1 if and only if $j \in A_i$. We will call $\mathbf{v}_i$ the characteristic vector of A_i. If we can show that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent over the field $\mathbb{R}$ of real numbers, then we are done: m cannot be greater than n, the dimension of the vector space of real vectors of length n. First of all we examine the scalar product of the vectors $\mathbf{v}_i$. Let $\alpha_i = |A_i| - \lambda \geq 1$. Then

$$\mathbf{v}_i \mathbf{v}_j = \sum_{k=1}^{n} v_{i,k} v_{j,k} = \begin{cases} \lambda & \text{if } i \neq j \\ \lambda + \alpha_i & \text{if } i = j. \end{cases}$$

Suppose that there is a linear combination of the vectors $\sum_{i=1}^{m} c_i \mathbf{v}_i = \mathbf{0}$, $c_i \in \mathbb{R}$. Our aim is to show that then $c_i = 0$ for $i = 1, 2, \ldots, m$. Scalar multiplying both sides by $\mathbf{v}_j$ we obtain

$$\lambda \sum_{i=1}^{m} c_i + c_j \alpha_j = 0 \quad , \quad c_j = -\frac{\lambda}{\alpha_j} \sum_{i=1}^{m} c_i \qquad (j = 1, \ldots, m).$$

If

$$\sum_{i=1}^{m} c_i = 0 \ ,$$

then we are done. If not, then summarizing these equalities we get

$$\sum_{j=1}^{m} c_j = -\lambda(\sum_{j=1}^{m} \frac{1}{\alpha_j}) \sum_{i=1}^{m} c_i,$$

a contradiction, because

$$1 + \lambda \sum_{j=1}^{m} \frac{1}{\alpha_j} > 0 \ .$$

Note that equality holds in Fisher's inequality for instance in the case of projective planes.

Although perhaps this is not the simplest way to prove Fisher's inequality, it has the advantage that we can adopt the idea to prove the following nonuniform version of the Ray-Chaudhuri–Wilson Theorem.

Theorem 1.3 (Frankl–Wilson). *Let $\lambda_1, \ldots, \lambda_s < n$ be nonnegative integers. If $|A \cap B| \in \{\lambda_1, \ldots, \lambda_s\}$ for every $A, B \in \mathcal{H}$, then $|\mathcal{H}| \leq \sum_{k=0}^{s} \binom{n}{k}$. (We use the convention that A and B denote different edges of $\mathcal{H}$.)*

Note. In the special case when the hypergraph $\mathcal{H}$ is r-uniform ($r \geq s$), the original Ray-Chaudhuri–Wilson Theorem states that $|\mathcal{H}| \leq \binom{n}{s}$.

Sketch of the proof (Alon–Babai). With the same notations as in the previous proof define the functions $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, 2, \ldots, m$ by

$$f_i(\mathbf{x}) = \prod_{\lambda_j < |A_i|} \Big(\sum_{k=1}^{n} v_{i,k} x_k - \lambda_j \Big) \ .$$

Then $f_i(\mathbf{v}_i) \neq 0$ and $f_i(\mathbf{v}_j) = 0$ for $i > j$ if the edges are ordered as $|A_1| \leq |A_2| \leq \ldots \leq |A_m|$. As the $\mathbf{v}_i$ are $0-1$ vectors, we have the same result for the functions $g_i : \mathbb{R}^n \to \mathbb{R}$ obtained from f_i by substituting x_k for the powers of $x_k (k = 1, \ldots, n)$. Now it is easy to see that the functions $g_1, \ldots, g_m$ are linearly independent over $\mathbb{R}$. In fact, suppose that

$$\sum_{i=1}^{m} c_i g_i = 0$$

and there exists a nonzero coefficient c_i. Let j be the smallest index i such that $c_i \neq 0$, then

$$0 = \sum_{i=1}^{m} c_i g_i(\mathbf{v}_j) = c_j g_j(\mathbf{v}_j) \neq 0 \ ,$$

a contradiction. The g_i are polynomials linear in each of the variables $x_1, \ldots, x_n$ and deg $g_i \leq s$. Thus they are contained in a subspace of $\mathbb{R}[x_1, \ldots, x_n]$ having dimension

$$\sum_{k=0}^{s} \binom{n}{k} ,$$

and this completes the proof.

References

L. Babai and P. Frankl, *Linear Algebra Methods in Combinatorics, Part I, Preliminary Version,* Chicago, 1988.

R.C. Bose, *A note on Fisher's inequality for balanced incomplete block designs,* Ann. Math. Stat. **20** (1949) 619–620.

P. Erdős, C. Ko and R. Rado, *Intersection theorems for systems of finite sets,* Quart. J. Math. Oxford **12** (1961) 313–320.

R.A. Fisher, *An examination of the different possible solutions of a problem in incomplete blocks,* Annals of Eugenics (London) **10** (1940) 52–75.

P. Frankl and R.M. Wilson, *Intersection theorems with geometric consequences,* Combinatorica **1** (1981) 357–368.

M. Hall, *Combinatorial Theory,* Blaisdell, 1967.

G. Katona, *A simple proof of the Erdős–Chao Ko–Rado theorem,* J. Comb. Th. B **13** (1972) 183–184.

D.K. Ray-Chaudhuri and R.M. Wilson, *On t-designs,* Osaka J. Math. **12** (1975) 737–744.

R.M. Wilson, *The exact bound in the Erdős–Ko–Rado theorem,* Combinatorica **4** (1984) 247–257.

2 Sperner Systems

In this section we investigate set systems satisfying the following inclusion property: if $A, B \in \mathcal{H}$, then $A \not\subseteq B$. Such hypergraphs are called Sperner systems.

Theorem 2.1 (Sperner). *If $\mathcal{H}$ is a Sperner system, then*

$$|\mathcal{H}| \leq \binom{n}{\lfloor n/2 \rfloor} .$$

We will present here three proofs of this theorem demonstrating three different methods of the theory. The first is Sperner's original proof.

First proof: "Push to the middle." It would be illustrative to imagine the subsets of X on $n+1$ levels: the i-th level contains the subsets of X of i elements, their number is $\binom{n}{i}$. We will denote this set by $\binom{X}{i}$. We can connect two elements of two neighbouring levels if one of them contains the other. (This is the usual way to represent partially ordered sets (posets); here the ordering is defined by inclusion.) Which level does have the largest size? The middle level, of course. More precisely, the middle ($\frac{n}{2}$-th) level, if n is even, and the two middle levels, if n is odd; in both cases the $\lfloor \frac{n}{2} \rfloor$-th level has maximum size. Our aim is to push the elements of $\mathcal{H}$ to the $\lfloor \frac{n}{2} \rfloor$-th level, where the upper bound follows trivially. Let $\mathcal{H}_k = \mathcal{H} \cap \binom{X}{k}$ – the k-element edges of $\mathcal{H}$ –, where $0 < k \leq n$. Define the shadow of $\mathcal{H}_k$ by

$$\triangle \mathcal{H}_k = \{A \subseteq X \ : \ |A| = k-1, \ A \subset B \text{ for some } B \in \mathcal{H}_k\}.$$

Claim. $| \triangle \mathcal{H}_k| \geq \frac{k}{n-k+1} |\mathcal{H}_k|$.

Proof. Count the pairs (A, B) with $A \in \Delta\mathcal{H}_k, B \in \mathcal{H}_k, A \subset B$ in two different ways. Each $B \in \mathcal{H}_k$ contains k members of $\Delta\mathcal{H}_k$, hence the number of the pairs is $k|\mathcal{H}_k|$. On the other hand, each $A \in \Delta\mathcal{H}_k$ is contained in $n-(k-1)$ members of $\binom{X}{k}$. Since $\mathcal{H}_k \subseteq \binom{X}{k}$, we obtain

$$k|\mathcal{H}_k| \leq (n-k+1)|\Delta\mathcal{H}_k|,$$

proving the claim.

Returning to the proof suppose that the largest edge of $\mathcal{H}$ has k vertices, $k > \lfloor \frac{n}{2} \rfloor$. Then $|\Delta\mathcal{H}_k| \geq \frac{k}{n-k+1}|\mathcal{H}_k| \geq |\mathcal{H}_k|$, and being $\mathcal{H} \cap \Delta\mathcal{H}_k = \emptyset$, we may replace the k-element edges by distint members of $\Delta\mathcal{H}_k$. It is easy to see that the new hypergraph is a Sperner system again, so we can iterate the process.

By a similar argument we also can push up the small edges to the $\lfloor \frac{n}{2} \rfloor$-th level; we have to consider

$$\nabla\mathcal{H}_k = \{A \subseteq X \ : \ |A| = k+1, A \supset B \text{ for some } B \in \mathcal{H}_k\},$$

and see that $|\nabla\mathcal{H}_k| \geq |\mathcal{H}_k|$ if $0 \leq k < \lfloor \frac{n}{2} \rfloor$.

Exercise. *Prove* $|\nabla\mathcal{H}_k| \geq \frac{n-k}{k+1}|\mathcal{H}_k|$ *for* $0 \leq k < n$.

Exercise. *Let* $\mathcal{H}$ *be a Sperner system. Show that there exists a Sperner system* $\mathcal{G} \subseteq \binom{X}{\lfloor n/2 \rfloor}$ *with* $|\mathcal{G}| = |\mathcal{H}|$ *such that for every* $A \in \mathcal{H}$ *there is a* $B \in \mathcal{G}$ *with* $A \subseteq B$ *or* $B \subseteq A$. (Hint: repeat the previous proof and use Hall's Theorem on matchings in bipartite graphs.)

Second proof: LYM inequality. The idea of Lubell was to prove the theorem by the following inequality (Lubell, Yamamoto, Meschalkin):

$$\sum_{k=0}^{n} \frac{|\mathcal{H}_k|}{\binom{n}{k}} \leq 1.$$

Observing that $\binom{n}{k} \leq \binom{n}{\lfloor n/2 \rfloor}$ for every $k = 0, 1, \ldots, n$, Sperner's Theorem follows immediately. It remains only to prove the LYM inequality.

There are $n!$ permutations of the elements of X. Let $\emptyset \neq A \subseteq X$. We shall say that a permutation π of $X = \{1, 2, \ldots, n\}$ begins with A if the first $|A|$ members of π are the elements of A in some order; the number of such permutations is $|A|!(n-|A|)!$. No permutation can begin with two different sets in $\mathcal{H}$ without violating Sperner's property. Thus

$$\sum_{A \in \mathcal{H}} |A|!(n-|A|)! \leq n! \ ,$$

or equivalently

$$\sum_{k=0}^{n} |\mathcal{H}_k| \frac{k!(n-k)!}{n!} \leq 1 \ ,$$

as was to be proved.

We mention here without a proof the following result:

Theorem 2.2 (Bollobás). Let $A_1 \dots, A_m, B_1, \dots, B_m \subseteq X$ such that $A_i \cap B_j = \emptyset$ if and only if $i = j$. Then

$$\sum_{i=1}^{m} \frac{1}{\binom{|A_i|+|B_i|}{|A_i|}} \leq 1.$$

(Here the A_i's are different subsets and so are the B_i's but $A_i = B_j$ is allowed.)

Exercise. *Derive LYM inequality from the above theorem.*

The method of Lubell's proof provides us to describe the extremal Sperner systems. We present here Lovász's argument.

Theorem 2.3. *Let $\mathcal{H}$ be a Sperner system. If $|\mathcal{H}| = \binom{n}{\lfloor n/2 \rfloor}$, then $\mathcal{H} = \binom{X}{n/2}$ if n is even, and $\mathcal{H} = \binom{X}{\frac{n-1}{2}}$ or $\mathcal{H} = \binom{X}{\frac{n+1}{2}}$ if n is odd.*

Proof. Suppose that $\mathcal{H}$ is an extremal Sperner system. If $\mathcal{H}_k \neq \emptyset$, then $\binom{n}{k} = \binom{n}{\lfloor n/2 \rfloor}$ from Lubell's proof. Thus, if n is even, then we are done. If n is odd, then $\mathcal{H} = \mathcal{H}_m \cup \mathcal{H}_{m+1}$ where $m = \frac{n-1}{2}$. Moreover, to reach equality in the LYM inequality, every permutation of $\{1, 2, \dots, n\}$ must begin with a suitable edge of $\mathcal{H}$. Hence if $A = \{x_1, \dots, x_m\}$ and $B = \{x_1, \dots, x_{m+1}\}$, then A or B must be in $\mathcal{H}$.

Suppose, by way of contradiction, that $\mathcal{H}$ contains some but not all of the m-element subsets of X. Then we can find sets E, F such that $E = F = m + 1, E \in \mathcal{H}, F \notin \mathcal{H}$. By relabelling the elements of X if necessary, we can suppose that $E = \{x_1, \dots, x_{m+1}\}$ and $F = \{x_i \dots, x_{m+i}\}$ for some i. Since $E \in \mathcal{H}$ and $F \notin \mathcal{H}$ there must be a largest integer $j < i$ with $E^* = \{x_j, \dots, x_{m+j}\} \in \mathcal{H}$, then $F^* = \{x_{j+1}, \dots, x_{m+j+1}\} \notin \mathcal{H}$. Since $E^* \cap F^* \subset E^* \in \mathcal{H}$, we must have $E^* \cap F^* \notin \mathcal{H}$. However, we have already seen that $E^* \cap F^*$ or F^* must be in $\mathcal{H}$, a contradiction.

Third proof: Symmetric chain decomposition. The idea of the proof of de Bruijn, Tengbergen and Kruyswijk is to decompose the partially ordered set 2^X to the disjoint union of symmetric chains. A symmetric chain is a chain $A_k \subset A_{k+1} \subset \dots \subset A_{n-k}$ of subsets of X, where $|A_i| = i$ for $k \leq i \leq n - k$. Clearly a symmetric chain can contain at most one member of a Sperner system. On the other hand, every symmetric chain contains an element of the middle level(s), implying Sperner's Theorem. The existence of a symmetric chain decomposition can be proved by induction on n. The initial step is trivial. Suppose that $n > 1$ and we have already found a symmetric chain decomposition of $2^{\{1,2,\dots,n-1\}}$. If $A_k \subset A_{k+1} \subset \dots \subset A_{n-1-k}$ is a symmetric chain in that decomposition, then we can construct two symmetric chains of 2^X of the form $A_k \subset A_{k+1} \subset \dots \subset A_{n-1-k} \subset A_{n-k-1} \cup \{n\}$ and $A_k \cup \{n\} \subset A_{k+1} \cup \{n\} \subset \dots \subset$

$A_{n-2-k} \cup \{n\}$. To be more precise, if n is odd and $k = \frac{n-1}{2}$, then the second chain does not exist and we have only one new symmetric chain. It is easy to see that 2^X is the disjoint union of these new chains, and the proof is ready.

In a general partially ordered set there is no way to define symmetric chains. But one can generalise the notion of a Sperner system easily to posets, defining antichains: sets without comparable elements. Let us state here the following minimax result:

Theorem 2.4 (Dilworth). *In any poset P, the maximum size of an antichain is equal to the minimum number of chains in a chain decomposition of P.*

References

B. Bollobás, *On generalised graphs,* Acta Math. Acad. Sci. Hung. **16** (1965) 447–452.

N.G. de Bruijn, C. Tengbergen and D. Kruyswijk, *On the set of divisors of a number,* Nieuw Arch. Widsk. **23** (1951) 191–193.

R.P. Dilworth, *A decomposition theorem for partially ordered sets,* Ann. Math. **51** (1950) 161–165.

D. Lubell, *A short proof of Sperner's lemma,* J. Comb. Th. **1** (1966) 299.

L.D. Meschalkin, *A generalization of Sperner's theorem on the number of subsets of a finite set,* Theor. Probl. Appl. **8** (1963) 203–204.

E. Sperner, *Ein Satz über Untermengen einer endlichen Menge,* Math. Z. **27** (1928) 544–548.

K. Yamamoto, *Logarithmic order of free distributive lattices,* J. Math. Soc. Japan **6** (1954) 343–353.

3 The Littlewood–Offord Problem

In this section we apply Sperner-type theorems for the following geometric problem. Given vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ of length ≥ 1 in the d-dimensional Euclidean space, how many of the 2^n sums $\sum_{i=1}^{n} \varepsilon_i \mathbf{x}_i \quad (\varepsilon_i = \pm 1)$ can lie in a ball of radius r? The pioneering result in this direction was that of Littlewood and Offord for complex numbers (i.e. 2-dimensional vectors). Erdős improved their upper bound by an argument based on Sperner's Theorem. Here we follow his idea to prove a more general theorem. First we deal with the simplest case $d = 1$.

Theorem 3.1. *Let $x_1, \ldots, x_n$ be real numbers such that $|x_i| \geq 1$ for each i, and let J be any interval of length 2 open at at least one end. Then the number of sums $\sum_{i=1}^{n} \varepsilon_i x_i \quad (\varepsilon_i = \pm 1)$ lying in J is at most $\binom{n}{\lfloor n/2 \rfloor}$.*

Proof. Without loss of generality we can assume that all of the x_i's are positive; for any negative x_i can be replaced by $-x_i$. We now associate to each sum $\sum_{i=1}^{n} \varepsilon_i x_i$ the set $A = \{i \mid \varepsilon_i = +1\}$. If A_1 and A_2 were two such sets, and $A_1 \subset A_2$, then the corresponding sums would differ by at least two and so they could not both be in J. It follows that the sets A corresponding to sums lying in J must form a Sperner system and the result is therefore an immediate consequence of Theorem 2.1.

Corollary. *If the x_i are as in the Theorem and J is any interval of length $2r$ open at at least one end, then the number of sums $\sum_{i=1}^{n} \varepsilon_i x_i$ $(\varepsilon_i = \pm 1)$ lying in J is at most $\lceil r \rceil \binom{n}{\lfloor n/2 \rfloor}$.*

Now we are ready to state the general result.

Theorem 3.2. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be vectors in the d-dimensional space with $|\mathbf{x}_i| \geq 1$ for each i, and let B_r be an open ball of radius r. Then the number of sums $\sum_{i=1}^{n} \varepsilon_i x_i$ $(\varepsilon_i = \pm 1)$ lying in B_r is at most $(1+\delta) r d \binom{n}{\lfloor n/2 \rfloor}$ for an arbitrarily small positive δ, if r and n are big enough.*

Proof. First observe that each vector has a coordinate of absolute value $\geq \frac{1}{\sqrt{d}}$. By the Pigeonhole Principle there exists an $1 \leq i \leq d$ such that there are at least n/d vectors with i-th coordinate $\geq 1/\sqrt{d}$ or $\leq -1/\sqrt{d}$. Relabelling the vectors and the coordinates and replacing $\mathbf{x}_j$ by $-\mathbf{x}_j$ if necessary, we may assume that the first coordinate of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ is at least $\frac{1}{\sqrt{d}}$, where $t \geq \frac{n}{d}$. Fix the signs $\varepsilon_{t+1}, \ldots, \varepsilon_n$. If $\sum_{i=1}^{n} \varepsilon_i \mathbf{x}_i \in B_r$, then $\sum_{i=1}^{t} \varepsilon_i \mathbf{x}_i \in B'_r = B_r - \sum_{i=t+1}^{n} \varepsilon_i \mathbf{x}_i$. Therefore the first coordinate $\sum_{i=1}^{t} \varepsilon_i x_{i,1}$ of the vector $\sum_{i=1}^{t} \varepsilon_i \mathbf{x}_i$ must lie in an open interval of length $2r$. Applying the Corollary to the real numbers $\sqrt{d} x_{i,1}$, we can see that the number of sums $\sum_{i=1}^{n} \varepsilon_i \mathbf{x}_i$ lying in B_r is at most $\lceil \sqrt{d} r \rceil \binom{t}{\lfloor t/2 \rfloor}$.

There are 2^{n-t} possibilities to choose the signs $\varepsilon_{t+1}, \ldots, \varepsilon_n$. Thus we have an upper bound $2^{n-t} \lceil \sqrt{d} r \rceil \binom{t}{\lfloor t/2 \rfloor}$ for the number of sums $\sum_{i=1}^{n} \varepsilon_i \mathbf{x}_i$ lying in B_r. Using Stirling's well-known formula $n! \approx n^n e^{-n} \sqrt{2\pi n}$ we can estimate $\binom{n}{\lfloor n/2 \rfloor} \approx 2^n \sqrt{\frac{2}{\pi n}}$ and $\binom{t}{\lfloor t/2 \rfloor} \approx 2^t \sqrt{\frac{2}{\pi t}}$. Therefore, if r and n are large enough, we have

$$2^{n-t} \lceil \sqrt{d} r \rceil \binom{t}{\lfloor t/2 \rfloor} < (1+\delta) \sqrt{d} r \binom{n}{\lfloor n/2 \rfloor} \sqrt{\frac{n}{t}} \leq (1+\delta) d r \binom{n}{\lfloor n/2 \rfloor}.$$

Unfortunately, this upper bound is not the best possible. Recently, Frankl and Füredi succeeded in omitting the term d from the estimate.

However, for unit balls (balls of radius 1) the exact upper bound is known. Katona and Kleitman first proved it for $d = 2$.

Theorem 3.3. *Let $z_1, \ldots, z_n$ be complex numbers with $|z_i| \geq 1$ for each i. Then the number of sums $\sum_{i=1}^{n} \varepsilon_i z_i$ $(\varepsilon_i = \pm 1)$ lying inside a circle of unit radius is at most $\binom{n}{\lfloor n/2 \rfloor}$.*

Their proof depends on the following so-called Two-part Sperner Theorem that can be derived by the symmetric chain method.

Theorem 3.4. *Let X be the disjoint union of two proper subsets X_1 and X_2. If there are no $A, B \in \mathcal{H}$ with $A \subset B$ and $B \setminus A \subseteq X_1$ or $B \setminus A \subseteq X_2$, then $|\mathcal{H}| \leq \binom{n}{\lfloor n/2 \rfloor}$.*

Proof of Theorem 3.3. We may assume that each z_i has real part ≥ 0. To each of the sums $\sum_{i=1}^{n} \varepsilon_i z_i$ we can associate two sets, namely

$$A = \{i \mid \varepsilon_i = 1 \text{ and } z_i \text{ is in the first quadrant}\}$$

and

$$B = \{i \mid \varepsilon_i = 1 \text{ and } z_i \text{ is in the fourth quadrant}\} .$$

If two sums give rise to the same A but two different sets B_1, B_2 such that $B_1 \subset B_2$, then the sums differ by a sum of complex numbers all in the fourth quadrant and all of length ≥ 2; the sums therefore cannot both lie inside a unit circle. The same is true if we interchange the role of the two quadrants. So the sets $C = A \cup B$ corresponding to sums in a unit circle must satisfy the conditions of the Two-part Sperner Theorem, and the result follows.

To close this section we mention that Kleitman generalized Theorem 3.3 to arbitrary dimensions: he defined a partition of the set of sums $\sum_{i=1}^{n} \varepsilon_i \mathbf{x}_i$ into $\binom{n}{\lfloor n/2 \rfloor}$ blocks so that no two sums in the same block can lie in the same unit ball. Unfortunately, Theorem 3.4 cannot be generalized to d-partitions of the underlying set if $d \geq 3$.

References

P. Erdős, *On a lemma of Littlewood and Offord,* Bull. Amer. Math. Soc. **51** (1945) 898–902.

P. Frankl and Z. Füredi, *Solution of the Littlewood–Offord problem in high dimensions,* Ann. of Math. (2) **128** (1988) 259–270.

G. Katona, *On a conjecture of Erdős and a stronger form of Sperner's theorem,* Stud. Sci. Math. Hung. **1** (1966) 59–63.

D.J. Kleitman, *On a lemma of Littlewood and Offord on the distribution of certain sums,* Math. Z. **90** (1965) 251–259.

D.J. Kleitman, *On a lemma of Littlewood and Offord on the distribution of linear combination of vectors,* Adv. Math. **5** (1970) 1–3.

J. Littlewood and C. Offord, *On the number of real roots of a random algebraic equation III,* Mat. Sb. **12** (1943) 277–285.

4 The Kruskal–Katona Theorem

Our last theme in the theory of extremal set systems is a deeper analysis of the shadow of a k-uniform hypergraph $\mathcal{H}$. In order to prove Sperner's Theorem it was satisfactory to see that

$$|\Delta\mathcal{H}| \geq \frac{k}{n-k+1}|\mathcal{H}|.$$

The Kruskal–Katona Theorem gives the exact lower bound. To understand the theorem we have to recall the so-called l-binomial representation of positive integers.

Theorem 4.1. *Given positive integers m and l, there exists a unique representation of m in the form*

$$m = \binom{a_l}{l} + \binom{a_{l-1}}{l-1} + \ldots + \binom{a_t}{t} \quad \textit{where} \quad a_l > a_{l-1} > \ldots > a_t \geq t \geq 1.$$

Proof. We prove only the existence of such a representation. First choose a_l as the largest integer for which $\binom{a_l}{l} \leq m$. Then choose a_{l-1} as the largest integer for which $\binom{a_{l-1}}{l-1} \leq m - \binom{a_l}{l}$. Were $a_{l-1} \geq a_l$, we would have $m \geq \binom{a_l}{l} + \binom{a_l}{l-1} = \binom{a_l+1}{l}$, contradicting the maximality of a_l. Therefore $a_{l-1} < a_l$. Continuing this process we eventually reach a stage where the choice of a_t for some $t \geq 2$ actually gives equality:

$$\binom{a_t}{t} = m - \binom{a_l}{l} - \binom{a_{l-1}}{l-1} - \ldots - \binom{a_{t+1}}{t+1};$$

or we get right down to choosing a_1 as the integer such that

$$\binom{a_1}{1} \leq m - \binom{a_l}{l} - \ldots - \binom{a_2}{2} < \binom{a_1+1}{1},$$

$$m = \binom{a_l}{l} + \ldots + \binom{a_1}{1}.$$

Exercise. *Show that the l-binomial representation of the positive integer m is unique.*

To find set systems with the smallest possible shadows it will be useful to define the following ordering on $\binom{X}{k}$. Let $1 \leq a_1 < a_2 < \ldots < a_k \leq n$ and $1 \leq b_1 < b_2 < \ldots < b_k \leq n$ are integers. Define $\{a_1, \ldots, a_k\} \prec \{b_1, \ldots, b_k\}$ if $a_k < b_k$ or if there exists an integer $l < k$ such that $a_k = b_k, \ldots, a_{l+1} = b_{l+1}$ and $a_l < b_l$. It is easy to check that $\prec$ is a linear ordering on the k-element subsets of X. Let $1 \leq m \leq$

$\binom{n}{k}$ and describe the first m elements of $\binom{X}{k}$ according to $\prec$. Let the k-binomial representation of m be $m = \binom{a_k}{k} + \ldots + \binom{a_t}{t}$, then the sets we are looking for are the k-element subsets of $\{1, \ldots, a_k\}$; the $(k-1)$-element subsets of $\{1, \ldots, a_{k-1}\}$ unioned with $\{a_k+1\}, \ldots$, and the t-element subsets of $\{1, \ldots, a_t\}$, unioned with $\{a_k+1, \ldots, a_{t+1}+1\}$. What is the shadow of this set-system? It consists of the $(k-1)$-element subsets of $\{1, \ldots, a_k\}$; the $(k-2)$-element subsets of $\{1, \ldots, a_{k-1}\}$, unioned with $\{a_k+1\}; \ldots$; and the $(t-1)$-element subsets of $\{1, \ldots, a_t\}$, unioned with $\{a_k+1, \ldots, a_{t+1}+1\}$. Thus the shadow of the first $m = \binom{a_k}{k} + \ldots + \binom{a_t}{t}$ elements of $\binom{X}{k}$ is the first $\binom{a_k}{k-1} + \ldots + \binom{a_t}{t-1}$ elements of $\binom{X}{k-1}$. This observation motivates the following

Theorem 4.2 (Kruskal–Katona). *Let $\mathcal{H}$ be a k-uniform hypergraph. If the k-binomial representation of $|\mathcal{H}|$ is*

$$|\mathcal{H}| = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \ldots + \binom{a_t}{t},$$

where $a_k > a_{k-1} > \ldots > a_t \geq t \geq 1$, then

$$|\Delta\mathcal{H}| \geq \binom{a_k}{k-1} + \binom{a_{k-1}}{k-2} + \ldots + \binom{a_t}{t-1}.$$

Sketch of the proof. The idea is "shift to the left": a shifting process that does not increase the shadow of $\mathcal{H}$. For $1 \leq i < j \leq n$ define the shift operator S_i^j by

$$S_i^j(A) = \begin{cases} A \setminus \{j\} \cup \{i\} & \text{if } j \in A, i \notin A \text{ and } A \setminus \{j\} \cup \{i\} \notin \mathcal{H} \\ A & \text{otherwise.} \end{cases}$$

If we define $S_i^j(\mathcal{H}) = \{S_i^j(A) : A \in \mathcal{H}\}$, then the basic observation is that $\Delta S_i^j(\mathcal{H}) \subseteq S_i^j(\Delta\mathcal{H})$.

If we apply the shift operators $S_1^j (j = 2, 3, \ldots, n)$ repeatedly, then the number of sets containing 1 increases, hence after a finite number of applications we must end with a hypergraph $\mathcal{H}^*$ of the same size as $\mathcal{H}$, satisfying $S_1^j(\mathcal{H}^*) = \mathcal{H}^*$ for $j = 2, \ldots, n$ and $|\Delta\mathcal{H}^*| \leq |\Delta\mathcal{H}|$. Let $\mathcal{H}_1^* = \{A \in \mathcal{H}^* : 1 \notin A\}$ and $\mathcal{H}_2^* = \{A \setminus \{1\} : 1 \in A \in \mathcal{H}^*\}$. Then $\Delta\mathcal{H}_1^* \subseteq \mathcal{H}_2^*, |\mathcal{H}^*| = |\mathcal{H}_1^*| + |\mathcal{H}_2^*|$.

The theorem can be proved by a double induction on n and k, it is obvious for arbitrary n if $k = 1$. Using the induction hypothesis it is easy to show, by way of contradiction, that

$$|\mathcal{H}_2^*| \geq \binom{a_k - 1}{k-1} + \ldots + \binom{a_t - 1}{t-1}.$$

If we use the induction hypothesis again we obtain that

$$|\Delta\mathcal{H}| \geq |\Delta\mathcal{H}^*| \geq |\mathcal{H}_2^*| + |\Delta\mathcal{H}_2^*| \geq$$

$$\geq \binom{a_k-1}{k-1} + \ldots + \binom{a_t-1}{t-1} + \binom{a_k-1}{k-2} + \ldots + \binom{a_t-1}{t-2} = \binom{a_k}{k-1} + \ldots + \binom{a_t}{t-1},$$

and this is what we wanted to prove.

It is worth mentioning that this theorem proved to be useful in the investigation of combinatorial properties of simplicial polytopes and related problems.

Let $\mathcal{H}$ be a k-uniform hypergraph, $0 < r < k$, and define

$$\Delta^{(r)}\mathcal{H} = \{A \subset X : |A| = r \text{ and } A \subset B \text{ for some } B \in \mathcal{H}\} \ .$$

Applying the theorem repeatedly we obtain the following

Corollary. *If the k-binomial representation of $|\mathcal{H}|$ is*

$$|\mathcal{H}| = \binom{a_k}{k} + \ldots + \binom{a_t}{t},$$

then

$$|\Delta^{(r)}\mathcal{H}| \geq \binom{a_k}{r} + \ldots + \binom{a_t}{t-k+r}.$$

As another corollary we deduce the Erdős–Ko–Rado Theorem (1.1) from the Kruskal–Katona Theorem. Suppose indirectly that $|\mathcal{H}| > \binom{n-1}{r-1}$. Define the $(n-r)$-uniform hypergraph $\mathcal{H}' = \{X \setminus A : A \in \mathcal{H}\}$. We are in the situation that $A \not\subseteq A'$ if $A \in \mathcal{H}, A' \in \mathcal{H}'$, so the sets in $\Delta^{(r)}\mathcal{H}'$ must be distinct form the sets in $\mathcal{H}$, and therefore $|\mathcal{H}| + |\Delta^{(r)}\mathcal{H}'| \leq \binom{n}{r}$. On the other hand, by the Corollary we have $|\Delta^{(r)}\mathcal{H}'| \geq \binom{n-1}{r}$, so $|\mathcal{H}| + |\Delta^{(r)}\mathcal{H}'| > \binom{n-1}{r-1} + \binom{n-1}{r} = \binom{n}{r}$, a contradiction.

Finally we mention a related result which is not sharp, but more suitable for further applications. If we extend the definition of the binomial coefficient $\binom{n}{k}$ to $\binom{x}{k}$, where x is a positive real number by

$$\binom{x}{k} = \frac{x(x-1)\ldots(x-k+1)}{k!},$$

then every integer $\binom{n-1}{k} < m < \binom{n}{k}$ can be written in the form $m = \binom{x}{k}$ with a suitable $n-1 < x < n$.

Theorem 4.3 (Lovász). *If $\mathcal{H}$ is a k-uniform hypergraph and $|\mathcal{H}| = \binom{x}{k}$ $(x \geq k)$, then $|\Delta\mathcal{H}| \geq \binom{x}{k-1}$.*

References

G.F. Clements, *Intersection theorems for sets of subsets of a finite set,* Quart. J. Math. Oxford **27** (1976) 325–337.

D.E. Daykin, *Erdős–Ko–Rado from Kruskal–Katona,* J. Comb. Th. A **17** (1974) 254–255.

P. Frankl, *A new short proof for the Kruskal–Katona theorem,* Discr. Math. **48** (1984) 327–329.

G. Katona, *A theorem on finite sets,* in Theory of Graphs, Proc. Colloq. Tihany, Hungary, 1966 (P. Erdős and G. Katona, eds.), Akadémiai Kiadó, Budapest, 1968, pp. 187–207.

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

P. McMullen and G.C. Shephard, *Convex Polytopes and the Upper Bound Conjecture,* London Math. Soc. Lecture Note Series **3**, Cambridge University Press, 1971.

PART B
Two-colorings of Hypergraphs

5 Blocking Sets I

The existence problem and the construction of blocking sets are widely studied in the theory of symmetric structures such as block designs or finite geometries. In our language of hypergraphs, a subset of X is a blocking set in $\mathcal{H}$ if it intersects every edge of $\mathcal{H}$ but contains no one of them. The investigation of blocking sets is also of central importance at position games: if the hypergraph formed of the winning sets contains no blocking sets, then the first player has a winning strategy.

An other terminology is accepted in combinatorial theory. A hypergraph $\mathcal{H}$ is called 2-colorable (or having Property B) if there is a 2-coloring of the points without a monochromatic edge. Clearly 2-colorability is equivalent to the existence of blocking sets: each color class in a good 2-coloration is just a blocking set in the hypergraph.

Exercise. *Suppose that no two edges of $\mathcal{H}$ have exactly one point in common. Prove that $\mathcal{H}$ is 2-colorable.*

In this section we deal with the following problem. At least how many edges must an r-uniform hypergraph have if it contains no blocking sets? The following result of Erdős gives an exponential lower bound independent of the size of the underlying set.

Theorem 5.1. *If an r-uniform hypergraph has at most 2^{r-1} edges, then it is 2-colorable.*

Proof. The theorem can be proved easily by counting the bad 2-colorings. Nevertheless we take the opportunity to introduce the probabilistic method through this simple example. Let $\mathcal{H} = \{A_1, \ldots, A_m\}$. Color the points of $\mathcal{H}$ by red and blue randomly, independently of each other with probability 1/2. Denote by E_i the event that A_i is monochromatic. The probability of this event is $\mathbf{P}(E_i) = \frac{1}{2^{r-1}}$. Therefore the probability of the bad 2-colorings is

$$\mathbf{P}(\bigcup_{i=1}^{m} E_i) \leq \sum_{i=1}^{m} \mathbf{P}(E_i) = \frac{m}{2^{r-1}} \leq 1.$$

If $m < 2^{r-1}$, then we are done. If $m = 2^{r-1}$, we need a tiny refinement. Consider the event, when all points are red: it has a positive probability. Hence $\mathbf{P}(\bigcap_{i=1}^{m} E_i) > 0$ and

$$\mathbf{P}(\bigcup_{i=1}^{m} E_i) \le \sum_{i=1}^{m} \mathbf{P}(E_i) - (m-1)\mathbf{P}(\bigcap_{i=1}^{m} E_i) < 1.$$

Therefore the probability of a good 2-coloring is positive and this means that there must be a good 2-coloring.

With a clever refinement of this method Beck improved this result. His idea was to color the points at random and then – by a suitably chosen probability – to change the color of bad points: points contained in monochromatic edges. This argument leads to the following

Theorem 5.2 (Beck). *Let ε be an arbitrarily small positive constant, and suppose that r is large enough. If an r-uniform hypergraph has at most $2^r r^{1/3-\varepsilon}$ edges, then it is 2-colorable.*

On the other hand, these theorems are quite good:

Theorem 5.3 (Erdős). *There exist non-2-colorable r-uniform hypergraphs with $c \cdot r^2 \cdot 2^r$ edges.*

Proof. Choose m r-element subsets of X, there are $\binom{\binom{n}{r}}{m}$ ways to do this. It can be shown that for a suitable choice of n and m, most of these hypergraphs are non-2-colourable. Count the number N of the pairs $(\mathcal{H}, f)$ where $\mathcal{H} \subseteq \binom{X}{r}, |\mathcal{H}| = m$ and $f : X \to \{red,\ blue\}$ is a good 2-coloring of $\mathcal{H}$. Fix a 2-coloring f and denote by X_r and X_b the set of red and blue points, respectively, and let n_r and n_b be their cardinalities, $n_r + n_b = n$. If it contains no monochromatic edges, then each edge of $\mathcal{H}$ intersects both X_r and X_b. The number of the r-element subsets of X having this property is clearly $\binom{n}{r} - \binom{n_r}{r} - \binom{n_b}{r}$. Since there are 2^n 2-colorings,

$$N = \sum_f \binom{\binom{n}{r} - \binom{n_r}{r} - \binom{n_b}{r}}{m} \le 2^n \binom{\binom{n}{r} - 2\binom{n/2}{r}}{m}.$$

$$\text{If} \quad 2^n \binom{\binom{n}{r} - 2\binom{n/2}{r}}{m} < \binom{\binom{n}{r}}{m}, \text{ then } N < \binom{\binom{n}{r}}{m},$$

and therefore there must be an $\mathcal{H}$ without a good 2-coloration. This is the case when $n = r^2$ and $m > cr^2 2^r$ with a suitable constant c. We omit the detailed calculation which is a bit tiresome.

References

J. Beck, *On 3-chromatic hypergraphs,* Discr. Math. **24** (1978) 127–137.

P. Erdős, *On a combinatorial problem I,* Nordisk Mat. Tidskrift **11** (1963) 5–10.

P. Erdős, *On a combinatorial problem II,* Acta Math. Acad. Sci. Hung. **15** (1964) 445–447.

P. Erdős and J.L. Selfridge, *On a combinatorial game,* J. Comb. Th. A **14** (1973) 298–301.

P. Erdős and J. Spencer, *Probabilistic Methods in Combinatorics,* Akadémiai Kiadó, Budapest, 1974.

A.W. Hales and R.I. Jewett, *Regularity and positional games,* Trans. Amer. Math. Soc. **106** (1963) 222–229.

6 Blocking Sets II

In the previous section we saw edge-number conditions leading to the existence of good 2-colorings in r-uniform hypergraphs. In the present section we consider degree conditions and their consequences. The degree of a point is the number of edges that contain it. The degree $\deg \mathcal{H}$ of the hypergraph $\mathcal{H}$ is the maximum degree of its points. An easy observation is the following.

Theorem 6.1. *If an r-uniform hypergraph $\mathcal{H}$ has degree $\deg \mathcal{H} \leq \frac{r}{2}$, then $\mathcal{H}$ is 2-colorable.*

Sketch of the proof. It follows from the degree condition that the union of any k edges has at least $2k$ points. Hall's theorem on bipartite graphs guarantee the existence of two disjoint sets of distinct representatives, i.e. two blocking sets. (The vertex classes of the bipartite graph are the points and the edges of $\mathcal{H}$, respectively; the point i is connected to the edge A if and only if $i \in A$.)

Lovász succeeded in weakening the linear degree condition to an exponential one, namely $\deg \mathcal{H} \leq 2^r/8r$. In fact, he proved a bit more:

Theorem 6.2. *If every edge of an r-uniform hypergraph $\mathcal{H}$ meets at most 2^{r-3} other edges, then $\mathcal{H}$ is 2-colorable.*

The main point of the proof is a refinement of counting the "bad events" if they are more or less independent of each other. Let $E_1, E_2, \ldots, E_m$ be events (the "bad" events) and consider them as the vertices of a graph $\mathcal{G}$. Connect two events – to vertices of $\mathcal{G}$ – with an edge if and only if they are not independent. $\mathcal{G}$ is called the dependency graph of the events $E_1, \ldots, E_m$. More precisely, we assume that every event is indpendent from the sigma-algebra generated by its neighbours in $\mathcal{G}$.

Theorem 6.3 (Lovász Local Lemma). *Let $E_1, \ldots, E_m$ be events and denote their dependency graph by $\mathcal{G}$, $d = \deg \mathcal{G}$. If $\mathbf{P}(E_i) \leq 1/4d$ for every $i = 1, \ldots, m$, then $\mathbf{P}(\bigcup_{i=1}^{m} E_i) < 1$.*

Proof. Denote by $\overline{E}$ the complementary event of E: $\overline{E}$ does happen if and only if E fails; of course $\mathbf{P}(\overline{E}) = 1 - \mathbf{P}(E)$. Now $\overline{\cup E_i} = \cap \overline{E_i}$, thus $\mathbf{P}(\cup E_i) < 1$ means that the good event happens with positive probability. Recall the definition of the conditional probability $\mathbf{P}(E|F) = \mathbf{P}(E \cap F)/\mathbf{P}(F)$; if E and F are independent, then $\mathbf{P}(E \cap F) = \mathbf{P}(E)\mathbf{P}(F)$ and $\mathbf{P}(E|F) = \mathbf{P}(E)$.

Our aim is to prove by induction on k that

$$\mathbf{P}(F_{k+1} \mid \overline{F_1} \cap \ldots \cap \overline{F_k}) \le \frac{1}{2d}$$

whenever $\{F_1, \ldots, F_{k+1}\} \subseteq \{E_1, \ldots, E_m\}$; then

$$\mathbf{P}(\bigcap_{i=1}^{m} \overline{E_i}) = \prod_{i=1}^{m} \mathbf{P}(\overline{E_i} \mid \overline{E_1} \cap \ldots \cap \overline{E_{i-1}}) \ge \left(1 - \frac{1}{2d}\right)^m > 0.$$

The initial step is clear: $\mathbf{P}(E_i) \le \frac{1}{4d} < \frac{1}{2d}$. We may suppose that the events $F_{d+1}, \ldots, F_k$ are independent of F_{k+1}. Then

$$\mathbf{P}(F_{k+1} \mid \overline{F_1} \cap \ldots \cap \overline{F_k}) = \frac{\mathbf{P}(F_{k+1} \cap \overline{F_1} \cap \ldots \cap \overline{F_k})}{\mathbf{P}(\overline{F_1} \cap \ldots \cap \overline{F_k})} =$$

$$= \frac{\mathbf{P}(F_{k+1} \cap \overline{F_1} \cap \ldots \cap \overline{F_d} \mid \overline{F_{d+1}} \cap \ldots \cap \overline{F_k})}{\mathbf{P}(\overline{F_1} \cap \ldots \cap \overline{F_d} \mid \overline{F_{d+1}} \cap \ldots \cap \overline{F_k})} \le \frac{\mathbf{P}(F_{k+1} \mid \overline{F_{d+1}} \cap \ldots \cap \overline{F_k})}{1 - \sum_{i=1}^{d} \mathbf{P}(F_i \mid \overline{F_{d+1}} \cap \ldots \cap \overline{F_k})}.$$

Therefore, using the induction hypothesis for the denominator

$$\mathbf{P}(F_{k+1} \mid \overline{F_1} \cap \ldots \cap \overline{F_k}) \le \frac{\mathbf{P}(F_{k+1})}{1 - \sum_{i=1}^{d}(1/2d)} \le 2p \le \frac{1}{2d},$$

as was to be proved.

Proof of Theorem 6.2. Color the points of X randomly with red and blue, independently and with the same probability 1/2. If $\mathcal{H} = \{A_1, \ldots, A_m\}$, let E_i be the event that A_i is monochromatic, then $\mathbf{P}(E_i) = \frac{1}{2^{r-1}}$. If $A_i \cap A_j = \emptyset$, then the events E_i and E_j are independent of each other, hence the degree of the dependency graph of our events is $d = \deg \mathcal{G} \le 2^{r-3}$, $\mathbf{P}(E_i) \le \frac{1}{4d}$. Thus Theorem 6.3 can be applied, the probability of a good 2-coloration is positive – there exists a good 2-coloring of $\mathcal{H}$.

Sometimes it is worth using the following asymmetric version of Lovász's Local Lemma, which was first used by Spencer in order to prove lower bounds for Ramsey-numbers.

Theorem 6.4. *Let $E_1, \ldots, E_m$ be events with dependency graph $\mathcal{G}$ and suppose there exist $x_1, \ldots, x_m \in [0, 1)$ with*

$$\mathbf{P}(E_i) < x_i \prod_{\{E_i, E_j\} \text{is an edge of } \mathcal{G}} (1 - x_j)$$

for $i = 1, \ldots, m$. *Then*

$$\mathbf{P}(\bigcap_{i=1}^{m} \overline{E_i}) > \prod_{i=1}^{m}(1 - x_i) > 0.$$

Exercise. *Deduce Lovász's Local Lemma from Theorem 6.4.*

Exercise. *Prove Theorem 6.5.* (Hint: interpret the proof of Theorem 6.4.)

Let us see some applications of Theorem 6.2 to almost disjoint set-systems, due to Erdős and Lovász.

Theorem 6.5. *Let $\mathcal{H}$ be an r-uniform hypergraph such that any two edges of $\mathcal{H}$ have at most one point in common.*

A) If $n \le 2^{r-4}$, then $\mathcal{H}$ is 2-colorable.

B) If $\mathcal{H}$ is not 2-colorable, then it contains at least $2^{r-4}/r - 1$ points of degree at least $2^{r-4}/r - 1$.

C) If $|\mathcal{H}| \le 4^{r-4}/r^3$, then $\mathcal{H}$ is 2-colorable.

Proof. A) Suppose that $\mathcal{H}$ is not 2-colorable, then by Theorem 6.2 there exists an edge A which intersects more than 2^{r-3} other edges. Then one of the points of A belongs to more than $2^{r-3}/r$ of these. Let x be a point with maximum degree $d(x) = \deg \mathcal{H} > 2^{r-3}/r$. Then the edges containing x have no other point in common by the assumption, and so they cover at least $1 + \frac{2^{r-3}}{r}(r-1) > 2^{r-4}$ points, a contradiction.

B) For every $A \in \mathcal{H}$ let $\varphi(A)$ denote a vertex of A with largest degree and set $A' = A \backslash \{\varphi(A)\}$. The sets A' form an $(r-1)$-uniform hypergraph $\mathcal{H}'$. If $\mathcal{H}'$ is 2-colorable, then so is $\mathcal{H}$, thus we may suppose that there is a point x of $\mathcal{H}'$ with degree $t = d(x) > \frac{2^{r-4}}{r-1}$ (in $\mathcal{H}'$, of course). Let $A'_1, \ldots, A'_t$ be the edges of $\mathcal{H}'$ incident to x. Now the degree of the point $\varphi(A_i)$ in $\mathcal{H}$ is at least t. Moreover, the points $\varphi(A_1), \ldots, \varphi(A_t)$ are distinct because of the intersection condition.

C) Suppose indirectly, that $\mathcal{H}$ is not 2-colorable. Then, by the previous statement, there exist $2^{r-4}/r$ points with degree $> 2^{r-4}/r$. If we count the edges incident to these points we get more than $4^{r-4}/r^2$, and each of them is counted at most r times. Hence $|\mathcal{H}| > 4^{r-4}/r^3$.

We mention that result C) is essentially sharp: there exist r-uniform hypergraphs with less than $cr^4 4^r$ edges which are not 2-colorable and any two edges intersect in at most one point.

We close this section with an application to projective spaces. In the d-dimensional projective space of order q, $PG(d,q)$, there are $\approx q^d$ points and $\approx q^{2d-2}$ lines. Therefore if $q^{2d-2} \lesssim 4^{q-3}/q^3$, then there exist blocking sets in $PG(d,q)$ as Theorem 6.5 states it. Taking logarithms of base 2 on both sides we get

$$(2d-2)\log q \lesssim 2(q-3) - 3\log q \ .$$

So for $d \lesssim q/\log q$ there exist blocking sets in $PG(d,q)$, and this bound is much better than the bounds obtained from the various constructions.

References

P. Erdős and L. Lovász, *Problems and results on 3-chromatic hypergraphs and some related questions*, in Infinite and Finite Sets II (A. Hajnal, R. Radó and V.T. Sós, eds.), Colloq. Math. Series Soc. J. Bolyai **10**, North Holland, Amsterdam, 1975, pp. 609–627.

J. Spencer, *Asymptotic lower bounds for Ramsey functions,* Discr. Math. **20** (1977) 69–76.

T. Szőnyi, *Blocking sets in projective spaces,* manuscript.

7 Discrepancy I

In this section we define various concepts of discrepancy and give an upper bound for the discrepancy of a hypergraph in a function of the number of its edges. As we are looking for balanced 2-colorings, the most common algebraic formulation of discrepancy is

$$\operatorname{disc}(\mathcal{H}) = \min_{f:X\to\{+1,-1\}} \quad \max_{A\in\mathcal{H}} |\sum_{x\in A} f(x)|.$$

Dealing with the special case first when the size of the hypergraph is equal to that of the underlying set it is easy to prove the following result:

Theorem 7.1 (Erdős). *If* $|\mathcal{H}| = n = |X|$, *then* $\operatorname{disc}(\mathcal{H}) \le \sqrt{2n\ln(2n)}$.

The proof is based on a probabilistic lemma which has its own interest: it is a strong variant of Chebyshew's famous inequality for random variables having binomial distribution, it gives an efficient estimate for the sum of the small binomial coefficients.

Chernoff's Inequality. *Let* $\xi_1, \ldots, \xi_n$ *be independent random variables with common distribution function* $\mathbf{P}(\xi_i = 1) = \mathbf{P}(\xi_i = -1) = \frac{1}{2}$ *, and let* $\eta_n = \xi_1 + \ldots + \xi_n$ *. Then* $\mathbf{P}(\eta_n > \lambda) < e^{\frac{-\lambda^2}{2n}}$ *for any* $\lambda > 0$ *.*

Proof. The expectation $\mathbf{E}(\xi)$ of the random variable ξ (with discrete distribution) is $\mathbf{E}(\xi) = \sum_{\alpha} \alpha \mathbf{P}(\xi = \alpha)$. In particular, for $y > 0$

$$\mathbf{E}(e^{y\xi_i}) = \frac{1}{2}(e^y + e^{-y}) \le e^{y^2/2},$$

the inequality can be shown e.g. by comparing Taylor series. The expectation of the product of independent random variables is equal to the product of the expectations, hence

$$\mathbf{E}(e^{y\eta_n}) = \mathbf{E}(\prod_{i=1}^{n} e^{y\xi_i}) = \prod_{i=1}^{n} \mathbf{E}(e^{y\xi_i}) < (e^{y^2/2})^n \ .$$

Therefore we obtain the estimate

$$\mathbf{P}(\eta_n > \lambda) = \mathbf{P}(e^{y\eta_n} > e^{y\lambda}) \le \frac{1}{e^{y\lambda}} \mathbf{E}(e^{y\eta_n}) < e^{\frac{y^2 n}{2} - y\lambda} \ .$$

If we choose $y = \lambda/n$, optimising the inequality, we get

$$\mathbf{P}(\eta_n > \lambda) < e^{-\lambda^2/2n} .$$

Proof of Theorem 7.1. Let $f : X \to \{+1, -1\}$ be a random 2-coloring.

If $A \in \mathcal{H}, |A| = r$, then $f(A) = \sum_{x \in A} f(x) = \eta_r$. Thus

$$\mathbf{P}(|f(A)| > \lambda) = 2\mathbf{P}(\eta_r > \lambda) < 2e^{-\lambda^2/2r} \leq 2e^{-\lambda^2/2n} .$$

Therefore, defining

$$\mathrm{disc}(f) = \max_{A \in \mathcal{H}} |f(A)| ,$$

$$\mathbf{P}(\mathrm{disc}(f) > \lambda) < \sum_{A \in \mathcal{H}} \mathbf{P}(|f(A)| > \lambda) < |\mathcal{H}| \cdot 2e^{-\lambda^2/2n}.$$

Taking $\lambda = \sqrt{2n \ln(2n)}$ we obtain

$$\mathbf{P}(\mathrm{disc}(f) > \sqrt{2n \ln(2n)}\ ) < 1,$$

or equivalently, there exists a 2-coloring f with

$$\mathrm{disc}(f) \leq \sqrt{2n \ln(2n)}.$$

As a matter of fact, this proof gives $\mathrm{disc}(f) \leq \sqrt{2n \ln(2|\mathcal{H}|)}$ in the general case. Our next aim is to show, how one can reduce the term $\sqrt{n}$ to $2\sqrt{|\mathcal{H}|}$ if $|\mathcal{H}| < n/4$. The heart of this is a linear algebraic method worked out by Beck and Spencer.

First of all we introduce two other concepts of discrepancy. The hereditary discrepancy of $\mathcal{H}$ is

$$\mathrm{herdisc}(\mathcal{H}) = \max_{Y \subseteq X} \mathrm{disc}(\mathcal{H}|Y)$$

where the restriction of $\mathcal{H}$ to Y is $\mathcal{H}|Y = \{A \cap Y \mid A \in \mathcal{H}\}$.

Of course, $\mathrm{disc}(\mathcal{H}) \leq \mathrm{herdisc}(\mathcal{H})$. The linear discrepancy of $\mathcal{H}$ is usually defined by

$$\mathrm{lindisc}(\mathcal{H}) = \max_{\alpha \in [0,1]^n} \min_{\varepsilon \in \{0,1\}^n} \max_{A \in \mathcal{H}} |\sum_{i \in A} (\varepsilon_i - \alpha_i)|.$$

It is easy to see that $\mathrm{lindisc}(\mathcal{H}) \geq \frac{1}{2}\mathrm{disc}(\mathcal{H})$; indeed, when all $\alpha_i = \frac{1}{2}$, the possible "errors" $\varepsilon_i - \alpha_i = \pm\frac{1}{2}$, so

$$\min_{\varepsilon \in \{0,1\}^n} \max_{A \in \mathcal{H}} |\sum_{i \in A} (\varepsilon_i - \alpha_i)|$$

is just the discrepancy of the 2-colorings of $\mathcal{H}$ by the colors $\pm\frac{1}{2}$, hence the factor $\frac{1}{2}$. The crucial relation is

Theorem 7.2. $\mathrm{lindisc}(\mathcal{H}) \le \mathrm{herdisc}(\mathcal{H})$.

Proof. For the sake of simplicity let $\mathrm{herdisc}(\mathcal{H}) = K$. We have to show that for arbitrary $\alpha_1, \dots, \alpha_n \in [0,1]$ there exist $\varepsilon_1, \dots, \varepsilon_n \in \{0,1\}$ with $|\sum_{i\in A}(\varepsilon_i - \alpha_i)| \le K$ for all $A \in \mathcal{H}$. Actually it is enough to prove this for $\alpha_1, \dots, \alpha_n$ having finite binary expressions, because of a standard compactness argument: if δ is an arbitrarily small positive number, there exist $\alpha_1(\delta), \dots, \alpha_n(\delta) \in [0,1]$ with finite binary expansions satisfying $|\alpha_i - \alpha_i(\delta)| < \frac{\delta}{n}$. Now if $|\sum_{i\in A}(\varepsilon_i - \alpha_i(\delta)| \le K$, then $|\sum_{i\in A}(\varepsilon_i - \alpha_i)| < K + \delta$.

Suppose therefore that $\alpha_1, \dots, \alpha_n \in [0,1]$ have finite binary expressions, then there is a minimal natural number k such that $2^k\alpha_i \in \mathbb{Z}$ for $i = 1, 2, \dots, n$. Let Y be the set of i such that α_i has 1 as its k-th binary digit. As $\mathrm{disc}(\mathcal{H}|Y) \le \mathrm{herdisc}(\mathcal{H}) = K$, there exist $\delta_i = \pm 1$ so that $|\sum_{i\in A\cap Y}\delta_i| \le K$ for all $A \in \mathcal{H}$. Define approximations $\alpha_{1,1}, \dots, \alpha_{1,n}$ by

$$\alpha_{1,i} = \begin{cases} \alpha_i & \text{if } i \notin Y \\ \alpha_i + \delta_i \cdot 2^{-k} & \text{if } i \in Y\ . \end{cases}$$

Then for any $A \in \mathcal{H}$ we have

$$|\sum_{i\in A}(\alpha_{1,i} - \alpha_i)| = |\sum_{i\in A\cap Y} 2^{-k}\delta_i| \le 2^{-k}K\ .$$

The numbers $\alpha_{1,i} \in [0,1]$ have binary expansions of length at most $k-1$. Hence we can repeat the above procedure to obtain $\alpha_{2,1}, \dots, \alpha_{2,n}$ with

$$|\sum_{i\in A}(\alpha_{2,i} - \alpha_{1,i})| \le 2^{-(k-1)}K\ ,$$

for every $A \in \mathcal{H}$. Iterating this process finally we reach $\varepsilon_i = \alpha_{k,i} \in \{0,1\}$ satisfying

$$|\sum_{i\in A}(\varepsilon_i - \alpha_i)| \le \sum_{j=0}^{k-1} |\sum_{i\in A}(\alpha_{j+1,i} - \alpha_{j,i})| \le \sum_{j=0}^{k-1} 2^{-(k-j)}K < K = \mathrm{herdisc}(\mathcal{H})$$

for every $A \in \mathcal{H}$, and this is what we wanted to prove.

Now we are in the position to formulate the reduction step for hypergraphs having more points than edges.

Theorem 7.3. *Assume that* $\operatorname{lindisc}(\mathcal{H}|Y) \leq K$ *for all* $Y \subseteq X$ *with at most* $|\mathcal{H}|$ *elements,* $|\mathcal{H}| \leq n$. *Then* $\operatorname{lindisc}(\mathcal{H}) \leq K$.

Proof. Let $\mathcal{H} = \{A_1, \dots, A_m\}$, and define the incidence matrix $\mathbf{H} = (h_{ij})$ of $\mathcal{H}$: it is an $m \times n$ matrix with 0 or 1 entries , $h_{ij} = 1$ if and only if $j \in A_i$. For an arbitrary $\mathbf{x} \in [0,1]^n$ we have $\mathbf{Hx} = \mathbf{y} \in \mathbb{R}^m$ where

$$y_i = \sum_{s \in A_i} x_s.$$

The following simple fact from linear algebra will help us.

Lemma. *Let* $\mathbf{A}$ *be an* $m \times n$ *(real) matrix and suppose that the system of linear equations* $\mathbf{Ax} = \mathbf{b}$ *has a solution* $\mathbf{a}$ *with* $0 \leq a_i \leq 1$ *for all* $i = 1, \dots, n$. *Then there exists a solution* $\mathbf{a}'$ *with* $0 \leq a'_i \leq 1$ *for* $i = 1, \dots, n$, *where at least* $n - m$ a'_i *is* 0 *or* 1.

Proof (sketch). Consider the system of equations $\mathbf{Ax} = \mathbf{b}$ in variables x_i where we consider $x_i = a_i$ as a constant if $a_i = 0$ or 1. If less than $n - m$ a_i is 0 or 1, then we have more variables than equations. Since we have a solution located in the unit cube, there exist infinitely many solutions containing a line of solutions which must intersect the boundary of the unit cube. This leads us to a solution with more 0 or 1 entries. Repeating the process, if necessary, we obtain the result.

Now it is easy to finish the proof of Theorem 7.3. Let $\alpha_1, \dots, \alpha_n \in [0,1]$ be given. By the Lemma, there exist $\alpha'_1, \dots, \alpha'_n \in [0,1]$ with $\mathbf{H}\alpha = \mathbf{H}\alpha'$ and at most m α'_i are distinct from 0 or 1. Let $Y = \{i \mid \alpha'_i \neq 0 \text{ or } 1\}$. By the definition of linear discrepancy, there exist $\varepsilon_i \in \{0,1\}$ for every $i \in Y$ satisfying

$$|\sum_{i \in A \cap Y} (\varepsilon_i - \alpha'_i)| \leq \operatorname{lindisc}(\mathcal{H}|Y) \leq K,$$

for all $A \in \mathcal{H}$. $\mathbf{H}\alpha = \mathbf{H}\alpha'$ implies $\sum_{i \in A} (\alpha'_i - \alpha_i) = 0$ for every $A \in \mathcal{H}$. Thus setting $\varepsilon_i = \alpha'_i$ if $i \notin Y$ we obtain the estimate

$$|\sum_{i \in A} (\varepsilon_i - \alpha_i)| = |\sum_{i \in A} (\varepsilon_i - \alpha'_i) + \sum_{i \in A} (\alpha'_i - \alpha_i)| = |\sum_{i \in A} (\varepsilon_i - \alpha'_i)| = |\sum_{i \in A \cap Y} (\varepsilon_i - \alpha'_i)| \leq K$$

for all $A \in \mathcal{H}$, and Theorem 7.3 follows.

Theorem 7.4 (Beck–Fiala). *For an arbitrary hypergraph* $\mathcal{H}$

$$\operatorname{disc}(\mathcal{H}) \leq 2\sqrt{2|\mathcal{H}| \ln(2|\mathcal{H}|)}.$$

Proof. We have already seen it for $|\mathcal{H}| \geq n$. Otherwise we can proceed as

$$\operatorname{disc}(\mathcal{H}) \leq 2\operatorname{lindisc}(\mathcal{H}) \leq 2 \max_{|Y| \leq |\mathcal{H}|} \operatorname{lindisc}(\mathcal{H}|Y) \leq 2 \max_{|Y| \leq |\mathcal{H}|} \operatorname{herdisc}(\mathcal{H}|Y) =$$

$$= 2 \max_{|Y| \leq |\mathcal{H}|} \operatorname{disc}(\mathcal{H}|Y) \leq 2\sqrt{2|\mathcal{H}| \ln(2|\mathcal{H}|)}.$$

We close this section mentioning the beautiful result of Spencer which is best possible up to a constant factor.

Theorem 7.5. *For any hypergraph* $\mathcal{H}$, $\operatorname{disc}(\mathcal{H}) \leq 12\sqrt{|\mathcal{H}|}$.

In the next section we will see hypergraphs with $\operatorname{disc}(\mathcal{H}) \geq \frac{1}{4}\sqrt{|\mathcal{H}|}$. Let us mention that recently Beck found an application of Theorem 7.5 to a problem of Littlewood in harmonic analysis.

References

J. Beck, *Flat polynomials on the unit circle – Note on a problem of Littlewood,* Bull. London Math. Soc. **23** (1991) 269–277.

J. Beck and T. Fiala, *Integer making theorems*, Discr. Appl. Math. **3** (1981) 1–8.

J. Beck and J. Spencer, *Integral approximation sequences,* Math. Programming **30** (1984) 88–98.

H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statist. **23** (1952) 493–507.

L. Lovász, J. Spencer and K. Vesztergombi, *Discrepancy of set-systems and matrices,* Europ. J. Combinatorics **7** (1986) 151–160.

C. McDiarmid, *On the method of bounded differences,* in Surveys in Combinatorics, 1989 (J. Siemons, ed.), London Math. Soc. Lecture Notes Series **141**, Cambridge University Press, 1989.

J.E. Olson and J. Spencer, *Balancing families of sets,* J. Comb. Th. A **25** (1978) 29–37.

J. Spencer, *Six standard deviations suffice,* Trans. Amer. Math. Soc. **289** (1985) 679–706.

8 Discrepancy II

We begin this last section with a geometric approach to discrepancy which allows us to give a lower estimate for the linear discrepancy of $\mathcal{H}$ in terms of the determinant of its incidence matrix. Let $\mathbf{H}$ denote the incidence matrix of $\mathcal{H}$, and consider the set $U = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{Hx}||_\infty \leq 1\}$; i.e. the points $\mathbf{x}$ in $\mathbb{R}^n$ for which all the coordinates of $\mathbf{Hx}$ have absolute value at most 1. In the special case when $n = m(= |\mathcal{H}|)$ and $\det(\mathbf{H}) \neq 0$, we can define U as $\mathbf{H}^{-1}([-1,1]^n)$. Clearly U is a convex polyhedron centrally symmetric to the origin. We can identify the vertices of the cube $[-1,1]^n$ to the 2-colorings of X. If for some $t > 0$ tU contains a vertex of the cube $[-1,1]^n$, then we have a 2-coloring of X with discrepancy $\leq t$. Hence $\mathrm{disc}(\mathcal{H})$ is just the least number t for which tU covers a vertex of $[-1,1]^n$. We can formulate this observation in the following more standard way. Consider 2^n translated copies U_i of tU, centered at the vertices of the unit cube $[0,1]^n$. Then $\frac{1}{2}\mathrm{disc}(\mathcal{H})$ is the least number t for which U_i contains the points $(\frac{1}{2}, \ldots, \frac{1}{2})$.

Exercise. *Prove that* $\mathrm{lindisc}(\mathcal{H})$ *is the least number* t *for which the sets* U_j *cover the whole unit cube.*

Now we are ready to prove the promised

Theorem 8.1 (Lovász–Spencer–Vesztergombi). *If* $|\mathcal{H}| = n$ *and* $\mathbf{H}$ *denotes the incidence matrix of* $\mathcal{H}$, *then* $\mathrm{lindisc}(\mathcal{H}) \geq \frac{1}{2}\sqrt[n]{\det(\mathbf{H})}$.

Proof. If the sets $tU + \mathbf{v}, \mathbf{v} \in \{0,1\}^n$ cover the unit cube, then the sets $tU + \mathbf{v}, \mathbf{v} \in \mathbb{Z}^n$ must cover the whole space $\mathbb{R}^n$. Therefore it follows from a standard averaging argument that the volume of tU, $\mathrm{Vol}(tU) \geq 1$. As

$$\mathrm{Vol}(tU) = t^n \mathrm{Vol}(U) = t^n \mathrm{Vol}(\mathbf{H}^{-1}[-1,1]) =$$

$$= t^n \det(\mathbf{H}^{-1}) \mathrm{Vol}([-1,1]^n) = 2^n t^n (\det(\mathbf{H}))^{-1},$$

we obtain $t \geq \frac{1}{2}\sqrt[n]{\det(\mathbf{H})}$.

Comparing this with the previous exercise the result follows.

We can construct set systems with large disrepancy by Hadamard matrices. A ± 1 square matrix is called an Hadamard matrix if its row vectors (or equivalently, if its coloumn vectors) are pairwise orthogonal. Therefore the determinant of an $m \times m$ Hadamard matrix is $m^{m/2}$. The spectrum of the applications of the Hadamard matrices is very wide, e.g. in coding theory, therefore it is a very interesting question to generate

Hadamard matrices. However it is a longstanding open problem to determine the possible size of an Hadamard matrix.

Exercise. *Show that if there exist Hadamard matrices of size m and n, respectively, then there exists an Hadamard matrix of size mn.*

Exercise. *Prove that if* $\mathbf{H}$ *is an $m \times m$ Hadamard matrix, then either $m = 2$ or m is divisible by 4.*

Let $\mathbf{H}_0$ be an Hadamard matrix of size n. We may assume that its first row contains $+1$ entries only. Adding the first row to all the other rows we obtain a matrix $\mathbf{H}_1$ with $\det(\mathbf{H}_1) = \det(\mathbf{H}_0)$. If we devide the elements of $\mathbf{H}_1$ (except of the first row) by 2 we obtain a $0-1$ matrix $\mathbf{H}$ with $\det(\mathbf{H}) = \frac{1}{2^{n-1}} \cdot \det(\mathbf{H}_0)$. $\mathbf{H}$ can be considered as the incidence matrix of a set-system $\mathcal{H}$ on vertices $\{1, 2, \ldots, n\}$. Therefore, by Theorems 7.2 and 8.1 we obtain that

$$\operatorname{herdisc}(\mathcal{H}) \geq \operatorname{lindisc}(\mathcal{H}) \geq \frac{1}{2}\sqrt[n]{\det(\mathbf{H})} > \frac{1}{4}\sqrt{n}.$$

Thus, there exists a restriction $\mathcal{G} = \mathcal{H}|Y$ with $\operatorname{disc}(\mathcal{G}) > \frac{1}{4}\sqrt{n} \geq \frac{1}{4}\sqrt{|\mathcal{G}|}$, showing that Theorem 7.5 is sharp up to a constant factor.

The following theorem gives an upper bound to discrepancy by the degree of the hypergraph.

Theorem 8.2 (Beck–Fiala). *If the degree of the hypergraph $\mathcal{H}$ is positive, then*

$$\operatorname{lindisc}(\mathcal{H}) < \deg(\mathcal{H}) \ .$$

(Of course, if $\deg(\mathcal{H}) = 0$*, then* $\operatorname{herdisc}(\mathcal{H}) \leq 1$.)

Proof. Let $\alpha_1, \ldots, \alpha_n \in [0,1]$. We shall define a sequence $\alpha_0 = \alpha,\ \alpha_1, \ldots \in [0,1]^n$ satisfying the following properties. Setting $X_k = \{i \in X : \alpha_{k,i} = 0 \text{ or } 1\}$ we require

(1) $X_{k+1} \subset X_k$ $(X_{k+1} \neq X_k)$,
(2) $\alpha_{k+1,i} = \alpha_{k,i}$ if $\alpha_{k,i} = 0$ or 1, and finally
(3) $\sum_{i \in Y} \alpha_{k+1,i} = \sum_{i \in Y} \alpha_{k,i}$ if $Y \in \mathcal{H}$ with $|Y \cap X_K| > \deg(\mathcal{H})$.

Then Theorem 8.2 follows easily: there exists an s with $X_s = \emptyset$. Choose $\varepsilon_i = \alpha_{s,i}$, then for every $Y \in \mathcal{H}$ there exists a smallest integer k with $|Y \cap X_k| \leq \deg(\mathcal{H})$. In this situation we have

$$|\sum_{i \in Y} (\varepsilon_i - \alpha_i)| \leq |\sum_{i \in Y \cap X_k} (\varepsilon_i - \alpha_i)| \ + \ |\sum_{i \notin Y \cap X_k} (\varepsilon_i - \alpha_i)| =$$

$$= \left| \sum_{i\in Y\cap X_k} (\varepsilon_i - \alpha_i) \right| < |Y \cap X_k| \le \deg(\mathcal{H}).$$

Therefore it remained to define the sequence (α_k) with the desired properties. We define it by induction on k. Suppose that α_k is defined and $X_k \neq \emptyset$. (If $X_k = \emptyset$, then the procedure stops.) Let $\mathcal{H}_k = \{Y \in \mathcal{H} : |Y \cap X_k| > \deg(\mathcal{H})\}$. Counting the pairs $(i,\ Y \cap X_k)$ with $i \in Y \cap X_k$ on two different ways we obtain

$$|\mathcal{H}_k| \cdot \deg(\mathcal{H}) < \sum_{Y\in\mathcal{H}_k} |Y \cap X_k| \le |X_k| \cdot \deg(\mathcal{H}_k) \le |X_k| \cdot \deg(\mathcal{H})$$

yielding to $|\mathcal{H}_k| < |X_k|$. Now associating a real variable x_i to each $i \in X_k$ and defining $x_i = \alpha_{k,i}$ if $i \notin X_k$ we consider the system of linear equations

$$\sum_{i\in Y\cap X_k} x_i = \sum_{i\in Y\cap X_k} \alpha_{k,i} \quad (Y \in \mathcal{H}_k).$$

Applying the lemma of the preceeding section we get a solution $x_i = \alpha_{k+1,i}$ satisfying the desired properties.

The following result of Beck is a consequence of the Beck–Fiala Theorem.

Theorem 8.3. *Let $\mathcal{P} = \{p_1, \dots, p_N\}$ be a finite set of points in the plane. There exists a function $f : \mathcal{P} \longrightarrow \{-1, +1\}$ such that*

$$\left| \sum_{p\in T} f(p) \right| < c(\log N)^4$$

for every rectangle T with sides parallel to the coordinate axes.

Recently Bohus reduced the exponent 4 by 1 in the theorem. This result inspired Rödl and Winkler to investigate representations of grey areas combined of black and white dots. The "smoothness" of the resulting shade of grey color can be modelled as a discrepancy problem and estimated by Bohus's Theorem.

A conjecture of great interest is that $\deg(\mathcal{H})$ in the Beck–Fiala Theorem can be reduced to $c\sqrt{\deg(\mathcal{H})}$. It seems to be very hard to prove – or disprove, but somehow it can be approximated if the number of the vertices is not too huge.

Theorem 8.4 (Beck). *In any hypergraph $\mathcal{H}$ with $\deg(\mathcal{H}) = d$,*

$$\mathrm{disc}(\mathcal{H}) < c\sqrt{d \cdot \ln d \cdot \ln m} \cdot \ln n\ .$$

Note that the upper bound is hereditary to the restrictions of $\mathcal{H}$, and hence, by Theorem 7.2, we have the same upper bound for lindisc$(\mathcal{H})$. We do not prove this theorem here, because the proof is a bit complicated. We refer to the excellent monograph of Beck and Chen. We note that using Lovász's Local Lemma instead of Chernoff's Inequality, one can reduce the term $\sqrt{\ln m}$ to $\sqrt{\ln d}$ in the theorem.

References

J. Beck and W.W.L. Chen, *Irregularities of Distribution*, Cambridge Tracts in Mathematics **89**, Cambridge University Press, 1987.

J. Beck and T. Fiala, *Integer making theorems*, Discr. Appl. Math. **3** (1981) 1–8.

J. Beck and V.T. Sós, *Discrepancy theory,* in Handbook of Combinatorics (R.L. Graham, M. Grötschel and L. Lovász, eds.), to appear.

G. Bohus, *On the discrepancy of 3 permutations,* Random Structures and Algorithms **1** (1990) 215–220.

L. Lovász, J. Spencer and K. Vesztergombi, *Discrepancy of set-systems and matrices,* Europ. J. Combinatorics **7** (1986) 151–160.

V. Rödl and P. Winkler, *Concerning a matrix approximation problem,* preprint