Combinatorics Notes

Leart Ajvazaj

January 2023

Notes from my Combinatorics class at Cambridge with Bela Bollobas. Any mistake is with very high certainty mine.

Chapter 1. Basic Results

1. Chains, Antichains and Scattered Sets of Vectors

In 1943, in the Soviet Journal *Matematicheskiy Sbornik*, Littlewood and Offord published the third part of their series of papers entitled *On the number of real roots of a random algebraic equation*. The first part, in the Journal of the LMS, appeared later, in 1945, and the second part even later, in 1948, in the Annals of Mathematics. With these papers, Littlewood and Offord launched the detailed study of random polynomials as a serious subject, and since then investigations along similar lines have gone from strength to strength.
One of the lemmas Littlewood and Offord needed was the following. Let $z_1, \dots, z_n$ be complex numbers, each of modulus at least 1, and let r be a positive constant. Consider the 2^n sums $\sum_{k=1}^{n} \varepsilon_k z_k$, where ε_k are ± 1. Then the number of these sums that fall into a cirlce of radius r is not greater than $cr2^n(\log n)n^{-1/2}$ for a constant depending on r. Soon after the publication of this paper in 1948, Erdos sharpened this result for real numbers and radius 1 by appealing to Sperner's theorem. In this section we'll prove this fundamental theorem in combinatorics together with some of its extensions and consequences, including an extension of the Littlewood-Offord lemma in a sharp form.
We start by recalling a fundamental theorem of graph theory, the Konig-Egervary-Hall Theorem, which in Cambridge we tend to call Hall's Marriage Theorem, or Hall's Matching Theorem, or simply Hall's Theorem. This theorem should be familiar to every mathematician, just as the Weierstrass Approximation Theorem, the Structure Theorem of Finite Abelian Groups, the PNT and the Law of Large Numbers must be familiar to all.
Given a bipartite graph $G = (U, W; E)$ i.e. a bipartite graph with bipartition (U, W), a **complete matching** (CM) from U into W is a subgraph $H \subset G$ such that $d_H(u) = 1$ for every $u \in U$, and $d_H(w) \leq 1$ for every $w \in W$.

Theorem : 1 (Hall's Marriage Theorem)

A bipartite graph $G = (U, W; E)$ has a complete matching from U into W if, and only if,

$$|\Gamma_G(A)| \geq |A| \tag{1}$$

for every $A \subset U$.

Proof. Left as exercise □

Let's reformulate theorem 1 in terms of set systems. Let $\mathcal{F} = \{A_1, \ldots, A_m\}$ be a set system. A set $\{a_1, \ldots, a_m\}$ of m distinct elements, with $a_i \in A_i$, for every i is a **set of distinct representatives** of $\mathcal{F}$. Trivially, if $\mathcal{F}$ has a set of distinct representatives (SDR), then

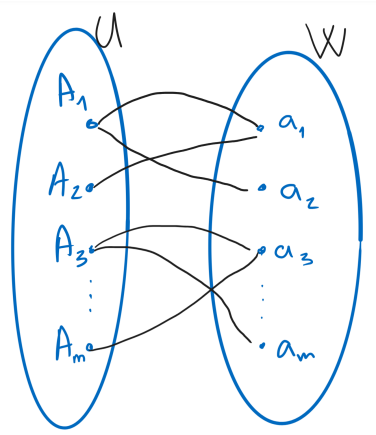
$$|\bigcup_{i \in I} A_i| \geq |I| \tag{2}$$

for every set $I \subset [m] = \{1, \ldots, m\}$.

Theorem : 2

Let $\mathcal{F} = \{A_1, \ldots, A_m\}$ be a set system satisfying (2). Then $\mathcal{F}$ has an SDR.

Proof. This result follows directly from theorem 1. How would we go about putting this problem in a graph theory setting? We could think of a bipartite graph with partition U, W where $U = \{A_1, A_2, \ldots, A_m\}$ and $W = \bigcup A_i$. Let there be an edge between $a \in U$ and $b \in W$ if and only if $b \in a$.



The result now follows from theorem 1. □

We present another proof that doesn't use Theorem 1.

Proof. The necessity is clear. Now proceed by induction on m. For $m = 1$ the assertion is trivial. Suppose that $m > 1$ and that the result holds for smaller values of m. Let $\mathcal{F}$ be a

set system satisfying (2).
Suppose first that for $S \subset [m]$, $S \neq \varnothing$ and $S \neq [m]$, we have strict inequality in (1). Pick an element $x_m \in A_m$ and set $A_i' = A_i \setminus \{x_m\}$ for $i = 1, 2, \ldots, m-1$. Then the system $\{A_1', A_2', \ldots, A_{m-1}'\}$ also satisfies Hall's condition so it has a SDR by the inductive hypothesis. Now suppose that there is a set $S \subset [m]$, $S \neq \varnothing$ and $S \neq [m]$ for which equality holds in (2). Set $\mathcal{F}_1 = \{A_i : i \in S\}$. By our induction hypothesis the system $\mathcal{F}_1$ has a set of distinct representatives, say X_1. Clearly $X_1 = \cup_{i \in S} A_i$, $|X_1| = |S|$. For $i \in [m] \setminus S$ set $A_i' = A_i \setminus X_1$ and $\mathcal{F}_2 = \{A_i' : i \in [m] \setminus S\}$. Does $\mathcal{F}_2$ satisfy Hall's condition? If $T \subset [m] \setminus S$, $T \neq \varnothing$, then

$$|\cup_{i \in T} A_i'| \geq |\cup_{i \in T} A_i \cup X_1| - |X_1| = |\cup_{i \in X \cup T} A_i| - |S| \geq |T|.$$

Hence by the inductive hypothesis $\mathcal{F}_2$ also has a SDR, say X_2. Then $X_1 \cup X_2$ will do the job for $\mathcal{F}$. □

Corollary : 3

Let G be a bipartite graph with bipartition (U, W). Suppose that $d(u) \geq d(w) \geq 1$ for all $u \in U$ and $w \in W$. Then there's a complete matching from U into W.

Proof. We have that there's a $d \in \mathbb{N}$ such that $d(u) \geq d \geq d(w) \geq 1$ for all $u \in U$ and $w \in W$. We have that $d|A| \leq e(A, \Gamma(A)) \leq d|\Gamma(A)|$. Thus $|A| \leq |\Gamma(A)|$. By theorem 1 we're done. □

A bipartite graph $G = (U, W; E)$ is said to be (k, l)-**regular** if $d(u) = k \geq 1$ and $d(w) = l$ for all $u \in U$ and $w \in W$. Also, a bipartite graph is **biregular** if it is (k, l)-regular for some natural numbers k and l. For a bipartite graph with bipartition (U, W) we shall find it useful to define the **weight** or measure of a set $T \subset U$ as $w(T) = |T|/|U|$; similarly, for $T \subset W$, we write $w(T) = |T|/|W|$.

Corollary : 4

Let G be a biregular bipartite graph with bipartition (U, W). Then for $A \subset U$ we have

$$w(\Gamma(A)) \geq w(A).$$

Proof. Say G is (k, l)-regular. We have $|e(A, \Gamma(A))| = k|A|$ and $|e(A, \Gamma(A))| \leq l|\Gamma(A)|$. So $\frac{k}{l}|A| \leq |\Gamma(A)|$ thus

$$\frac{k}{l \cdot |W|}|A| \leq w(\Gamma(A))|.$$

Since $l|W| = k|U|$ we get $w(A) \leq w(\Gamma(A))$. □

Put simply, this "theorem" states the important basic fact that in a biregular graph $(U, W; E)$ the proportion of neighbors of a set $A \cap U$ in W is at least as large as the proportion of A in U.

Corollary : 5

Let G be a (k,l)-regular bipartite graph with bipartition (U,W) and $|U| \leq |W|$ then there is a CM from U into W.

Equivalently, if G is a non-empty biregular bipartite graph with bipartition (U,W) and $|U| \leq |W|$ then there is a CM from U into W.

Proof. Since G is (k,l)-regular we have $k|U| = l|W|$. Since $|U| \leq |W|$, we have $k \geq l$. For $A \subset U$ we have

$$k|A| = e(A, \Gamma(A)) \leq l|\Gamma(A)| \leq k|\Gamma(A)|$$

Thus $|A| \leq |\Gamma(A)|$. □

A standard and, for us, very important, example of a biregular graph has bipartition $X^{(r)}, X^{(s)}$, and the edges are given by containment.

Corollary : 6

As usual, let X be an n-set, and let $0 \leq r < s \leq n$. If $|\frac{n}{2} - r| \geq |\frac{n}{2} - s|$ then there is an injection $f : X^{(r)} \to X^{(s)}$ such that $A \subset f(A)$ for every $A \in X^{(r)}$. If $|\frac{n}{2} - r| \leq |\frac{n}{2} - s|$ then there is an injection $g : X^{(s)} \to X^{(r)}$ such that $A \supset g(A)$ for every $A \in X^{(s)}$.

Our first proof of Sperner's Theorem, to be stated below, will be based on Corollary 6. We shall freely interchange X and $[n]$, and write $\mathcal{P}(n)$ for $\mathcal{P}(X) = \mathcal{P}([n])$. We call a family $\mathcal{A} \subset \mathcal{P}(n)$ an **antichain** or a **Sperner family** if no set in $\mathcal{A}$ is contained in another.

Theorem : 7 (Sperner's Theorem)

Let $\mathcal{A}$ be a Sperner family in $\mathcal{P}(n)$. Then

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

Proof. There are chains from $\lceil \frac{n}{2} \rceil$ level to n level as well as from $\lfloor \frac{n}{2} \rfloor$ level to 1 level. Concatenate these chains and you'll cover the entire $\mathcal{P}(n)$. Since $\mathcal{A}$ is Sperner it can intersect a chain at most at one place so by the fact that the middle layers have $\binom{n}{\lfloor n/2 \rfloor}$ we're done. The even case is handled similarly. □

Let us note the observation of Erdos that Theorem 7 implies the sharp form of the Littlewood-Offord Lemma for real numbers.

Corollary : 8

Let $x_1, x_2, \ldots, x_n$ be n real numbers, each of modulus at least 1, and for $\varepsilon = (\varepsilon_i)_1^n$, $\varepsilon_i = \pm 1$, set $x_\varepsilon = \sum_{i=1}^n \varepsilon_i x_i$. Then at most $\binom{n}{\lfloor n/2 \rfloor}$ of the 2^n sums x_ε fall into the interior of an interval I of length 2. This bound is best possible.

Proof. We may assume that $x_i \geq 1$ for all i since the sums we'll consider are the same if you take $-x_i$ or x_i. Given ε let $F_\varepsilon := \{i : \varepsilon_i = 1\}$ and $\mathcal{F} = \{F_\varepsilon : x_{F_\varepsilon} \in I\}$ (meaning we take $\varepsilon_i = 1$ for $i \in F_\varepsilon$ and -1 for the rest). Note that $\mathcal{F}$ is a Sperner set so you get the upper bound as desired. It's best possible as you can take $x_1 = \cdots = x_n = 1$ and you're done. □

The proof of Theorem 7 we have given in natural, but rather pedestrian. Next, we shall present a substantial extension of Theorem 7 to regular graded posets. Our prime example of a poset is $\mathcal{P}(n)$, with the relation $A < B$ if $A \subset B$. First, a poset is trivial if no two elements are comparable. Needless to say, we are not interested in trivial posets, so all posets we consider are assumed to be nontrivial. In fact, we shall assume more, namely that every poset we consider has a unique maximum and unique minimum.
In a poset $P = (S, <)$ a **chain** is a subset $T \subset S$ such that any two elements of T are comparable, i.e. $T = \{t_0, t_1, \ldots, t_k\}$ with $t_0 < t_1 < \cdots < t_k$. This chain has length $k+1$. An **antichain** is a subset of S in which no two elements are comparable. Trivially, a chain and an antichain meet in at most one element. A poset $P = (S, <)$ is **graded** if it has a partition $S = \cup_0^m S_i$ into antichains S_i such that if $x < y$ then thre is a chain $x = x_i < x_{i+1} < \cdots < x_j = y$ with $x_h \in S_h$ for every h.
A **regular graded poset** is a poset $P = (S, <)$ such that there are natural numbers r_i, s_i with this property: for $i < m$ every $x \in S_i$ is dominated by precisely r_i elements of S_{i+1} and for $i > 0$ it dominates precisely s_i elements of S_{i-1}. The canonical regular graded poset is $\mathcal{P}(X)$, with level sets $X^{(0)}, X^{(1)}, \ldots, X^{(m)}$.
We use the natural extension of the weight we defined earlier: $P = (S, <)$ is a graded poset with level sets $S_0, S_1, \ldots, S_m$ then for $A \subset S$ we define $A_i = A \cap S_i$, $i = 0, \ldots, m$, and $w(A) = \sum_0^m w(A_i) = \sum_0^m w(A \cap S_i) = \sum_0^m \frac{|A \cap S_i|}{|S_i|}$.

Theorem : 9

Let $P = (S, <)$ be a regular graded poset with level sets $S_0, \ldots, S_m$, and let A be an antichain in P. Then $w(A) \leq 1$.

Proof. Assume $A \neq \varnothing$. The span of A is the maximal t such that $A_i \neq \varnothing$ and $A_{i+t} \neq \varnothing$. If $t = 0$ we're done. Assume $t \geq 1$ and that the assertion holds for smaller values of the span. Suppose $A_k \neq \varnothing$ but $A_l = 0$ for $l > k$. Let $A'_{k-1} \subset S_{k-1}$ be the set of elements of A_{k-1} dominated by some elements of A_k. Then $A'_{k-1} \cap (A \setminus A_k) = \varnothing$ (WHY?) By corollary 4 $w(A'_{k-1}) \geq w(A_k)$. Therefore $A' = (A \setminus A_k) \cup A'_{k-1}$ is an anticahin of span $t-1$ and weight $w(A') \geq w(A)$ so we're done by the inductive hypothesis. □

Keeping the notation of Theorem 9, let r_i, s_i be the parameters of P so that $r_0 = 0$ and $s_m = 0$. Let M be the total number of maximal chains and, for $x \in S$, let $m(x)$ be the number of maximal chains containing x. Note that if $x \in S_h$ then $m(x) = \left(\prod_{i=1}^{h} r_i\right)\left(\prod_{i=h}^{m-1} s_i\right)$ so $m(x)$ depends only on the level of x. Also every maximal chain meets S_h in precisely one element, so if $x \in S_h$ then $m(x) = \frac{M}{|S_h|}$.
We can now prove theorem 9 differently.

Proof.

$$M = \sum_{\text{max chain}} 1 \geq \sum_{x \in A} m(x) = \sum_{h=0}^{m} |A_h| \frac{M}{|S_h|}$$

therefore $1 \geq \sum_{h=0}^{m} \frac{|A_h|}{|S_h|} = w(A)$. □

In $\mathcal{P}(X) = \mathcal{P}(n)$, our quintessential regular graded poset with level sets $X^{(h)}$, an antichain is also called a **Sperner family** or a **Sperner system**. In $\mathcal{P}(n)$ both N and $m(x)$ are trivially computed, so we obtain the following extension of Sperner's Theorem, noted by Meshalkin (1963), Bollobas (1965), and Lubell (1966). The elegant proof bellow, based on maximal chains, was found by Lubell. The previous proofs arose from this proof.

Theorem : 10 (LYM Inequality)

Let $\mathcal{F}$ be a Sperner family in $\mathcal{P}(X) = \mathcal{P}(n)$ and for $0 \leq k \leq n$ set $f_k = |\mathcal{F} \cap X^{(k)}|$. Then

$$\sum_{k=0}^{n} f_k \binom{n}{k}^{-1} \leq 1.$$

Proof. There are $n!$ maximal chains in $\mathcal{P}(n)$ since maximal chains correspond to permutations of $[n]$. Also every set in $X^{(k)}$ is contained in $k!(n-k)!$ maximal chains. Since every maximal chain contains at most one set from $\mathcal{F}$, we have

$$n! \geq \sum_{k=0}^{n} f_k k!(n-k)!.$$

□

It is easy to reformulate this proof of Theorem 10 without mentioning maximal chains in $\mathcal{P}$. We say that a set $A \in \mathcal{P}(n)$ is contained in a permutation $x_1 x_2 \dots x_n$ of $[n]$ if $A = \{x_1, \dots, x_k\}$, where $k = |A|$, so that every set of order k is contained in $k!(n-k)!$ permutations. Since every permutation contains at most one set in $\mathcal{F}$, we find that

$$\sum_{F \in \mathcal{F}} |F|!(n - |F|)!,$$

as required.
The partition of $\mathcal{P}(X)$ into chains used in the proof of Theorem 7 is rather haphazard:

essentially all we proved is that it exists. Next, following Kleitman, we shall show that there is a partition into so-called symmetric chains. A **symmetric chain** in $\mathcal{P}(X)$ is a nested sequence $(C_j)_i^{n-i}$ of subsets of X such that $|C_j| = j$ for every j. This symmetric chain has length $n+1-2i$: it can be obtained from a maximal chain by deleting its first i and last i sets.

Theorem : 11

The power set $\mathcal{P}(n)$ can be decomposed into symmetric chains. Every symmetric chain decomposition (SCD) consists of $\binom{n}{\lfloor n/2 \rfloor}$ chains.

Proof. We already have justification for the second claim. To prove the first, induct on n. $n = 1$ is clear. Suppose $n \geq 2$ and that the result holds for smaller values. Let $\mathcal{P}(n-1) = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \cdots \cup \mathcal{C}_s$ be a partition of $\mathcal{P}(n-1)$ into symmetric chains. From each chain $\mathcal{C}_i = \{A_1, \ldots, A_k\}$ where $A_1 \subset \cdots \subset A_k$ we create two new chains:

$$\mathcal{C}_i' = \{A_1, A_2, \ldots, A_k, A_k \cup \{n\}\}$$

and

$$\mathcal{C}_i'' = \{A_1 \cup \{n\}, A_2 \cup \{n\}, \ldots, A_{k-1} \cup \{n\}\}.$$

These are clearly symmetric and they partition $\mathcal{P}(n)$. □

Strictly speaking, the proof above is not quite correct: we have deliberately failed to emphasize the obvious, that a chain $\mathcal{C}_i''$ may be empty, in which case we just discard it. Indeed, if $\mathcal{C}_i$ has length 1, i.e. consists of a set of size $\frac{n-1}{2}$, then $\mathcal{C}_i''$ is defined to be the empty set. If each chain $\mathcal{C}_i$ in $\mathcal{P}(n-1)$ gave rise to two chains in $\mathcal{P}(n)$, then the number of chains would grow as 2^{n-1} (as $\mathcal{P}(1)$ is partitoned into one symmetrical chain), which is impossible.

In a SCD of $\mathcal{P}(n)$ the number of chains that start at level $i < \frac{n}{2}$ is clearly $\binom{n}{i} - \binom{n}{i-1}$, since for every j, $1 \leq j \leq \frac{n}{2}$, every symmetric chain containing a set at level $j-1$ must have a set at level j as well.
In order to extend Corollary 8 from real numbers to vectors in normed spaces, Kleitman used a partition of $\mathcal{P}(n)$ into families whose *profile* is the profile of a partition into symmetric chains, i.e. it is such that for $0 \leq i \leq \frac{n}{2}$ it contains exactly

$$l(n,i) = \binom{n}{i} - \binom{n}{i-1}$$

families with $n+1-2i$ with sets, where $\binom{n}{-1}$ is defined to be 0.

Theorem : 12

Let $x_1, x_2, \ldots$ be vectors of norm at least 1 in a normed space, and for a finite set A of natural numbers set $x_A = \sum_{i \in A} x_i$. Let $\mathcal{A} \subset \mathcal{P}(n)$ be such that if $A, B \in \mathcal{A}$ then $\|x_A - x_B\| < 1$. Then

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

Proof. Call a family $\mathcal{F}$ of finite subsets of natural numbers **sparse** or **scattered** if $\|x_A - x_B\| \geq 1$ for all $A, B \in \mathcal{F}, A \neq B$, and call a decomposition of $\mathcal{P}(n)$ into some families a **symmetric decomposition into sparse sets** (SDSS) if each family is sparse and for every i, $0 \leq i \leq \frac{n}{2}$, the decomposition $l(n,i)$ families with $n+1-2i$ sets.
Since $\mathcal{A}$ has at most one element in a sparse set and $\sum_{i=0}^{\lfloor n/2 \rfloor} l(n,i) = \binom{n}{\lfloor n/2 \rfloor}$, our theorem follows if we prove the following claim.
Claim. The power set $\mathcal{P}(n)$ has an SDSS.
To prove this Claim, we apply induction on n. Clearly, $\{\varnothing, \{1\}\}$ will do for $n = 1$. Assuming that our claim holds for $n - 1 \geq 1$, let us prove it for n. Let Δ_{n-1} be an SDSS of $\mathcal{P}(n-1)$. To enable us to construct an SDSS Δ_n of $\mathcal{P}(n)$ from Δ_{n-1}, we consider a support linear functional f at x_n, so that

$$\|f\| = 1 \text{ and } f(x_n) = \|x_n\| \geq 1.$$

Let $\mathcal{D} \in \Delta_{n-1}$ and pick $D \in \mathcal{D}$ such that $f(x_D) \geq f(x_E)$ for all $E \in \mathcal{D}$.
Define

$$\mathcal{D}' = \mathcal{D} \cup \{D \cup \{n\}\}$$

and

$$\mathcal{D}'' = \{E \cup \{n\} : E \in \mathcal{D}, E \neq D\},$$

and let Δ_n be the collection of non-empty families $\mathcal{D}'$ and $\mathcal{D}''$ we have just constructed. Clearly, Δ_n is a decomposition of $\mathcal{P}(n)$. Also, each set of the form $\mathcal{D}''$ is trivially sparse. Furthermore, $\mathcal{D}'$ is also sparse since if $E \in \mathcal{D}$ then for $D' = D \cup \{n\}$ we have

$$\|x_{D'} - x_E\| = \|x_D + x_n - x_E\| \geq f(x_D - x_E) + f(x_n) \geq f(x_n) \geq 1.$$

Finally, is Δ_n a symmetric decomposition of $\mathcal{P}(n)$ into sparse sets? Since a family $\mathcal{D}$ with $n-2i$ sets gives rise to two families, one with $n+1-2i$ sets and another with $n-1-2i$ sets, for $0 \leq i \leq \frac{n+1}{2}$ the number of families in Δ_n with $n+1-2i$ sets is $l(n-1,i)+l(n-1,i-1)$, i.e.

$$\begin{aligned} l(n-1,i) + l(n-1,i-1) &= \binom{n-1}{i} - \binom{n-1}{i-1} + \binom{n-1}{i-1} - \binom{n-1}{i-2} \\ &= \binom{n-1}{i} + \binom{n-1}{i-1} - \binom{n-1}{i-1} - \binom{n-1}{i-2} \\ &= \binom{n}{i} - \binom{n}{i-1} = l(n,i) \end{aligned}$$

sets, so Δ_n is indeed an SDSS of $\mathcal{P}(n)$, completing our proof of the claim, so our theorem is proved. □

In the proof of the calculations above involving $l(n,i)$ and various binomial coefficients are not needed, using them was an overkill. All we have to note is that the way the profile of a decomposition of $\mathcal{P}(n)$ is the same for an SCD as it was for an SDSS.

2. The EKR Theorem and Katona's Circle Method

The Erdos-Ko-Rado Theorem is a basic result in the combinatorics of finite sets. Although it was proved already in 1938, it was published only in 1961. Strangely, it went into the mathematical consciousness in the following slightly corrupted form:

Calling a family **intersecting** if any two of its sets intersect,, if $1 \leq r \leq \frac{n}{2}$ then an intersecting family of r-subsets of $[n]$ consists of at most $\binom{n-1}{r-1}$ sets.

The EKR Theorem led to numerous developments, including the simple but brilliant circle method of Katona (1968), and many extensions of it due to Frankl, Furedi, Kleitman, Pyber, and others.
In this brief section we start with an extension of EKR due to Bollobas (1973), and deduce the EKR inequality itself, together with other corollaries.
As usual, we take X to be a set with n elements. In a **cyclic order** on X, the elements of X are in an order $x_1, x_2, \ldots, x_n$ with x_{i+1} the successor of x_i and x_1 the successor of x_n. Note that there are $(n-1)!$ cyclic orders on X. An **arc** in a cyclic order on X is a set consisting of successive elements: $x_i, x_{i+1}, \ldots, x_j$. Needless to say, we may always take $i \leq n$ and $j \geq 1$. For example, for $n = 7$ the set $\{x_1, x_2, x_6, x_7\}$ is an arc. In a particular, in a cyclic order on n for $1 \leq k < n$ there are n arcs of length k (i.e. with k elements).
What Katona realized was that the number of times a set $A \in X^{(k)}$ appears as an arc in one of the $(n-1)!$ cyclic orders on X, namely $k!(n-k)!$, is independent of the way A is arranged in X, so whatever we can conclude about the number of arcs obtained from a family $\mathcal{A} \subset \mathcal{P}(X)$ can be pulled back to give information about $\mathcal{A}$. Most importantly, usually it is much easier to study a system of arcs that arise than the original set system. The proof below is an example of this.

Lemma : 1

Let $\mathcal{C} \subset X^{(\leq n/2)}$ be an intersecting Sperner family of arcs in a cyclic order $x_1, x_2, \ldots, x_n$ on X, and let r be the minimal length of an arc in $\mathcal{C}$. Then $\mathcal{C}$ consists of at most r arcs. In particular,

$$\sum_{C \in \mathcal{C}} \frac{1}{|C|} \leq 1. \tag{3}$$

Furthermore, equality holds in (3) if and only if for some point $x \in X$ and integer r, $1 \leq r \leq \frac{n}{2}$, the family $\mathcal{C}$ is the set of r arcs of length r that contain x.

Proof. If $r = 1$, the result is obvious. Assume $r > 1$ and $C_0 = \{x_1, x_2, \ldots, x_r\}$ be an arc of minimal length. Take $C \in \mathcal{C}$ with $C \neq C_0$. Define $f : \mathcal{C} \setminus C_0 \to C_0$ as follows: $f(C) = x_i$ if C

contains exactly one of x_i and x_{i+1}. Since $r \leq \frac{n}{2}$ we know that this function is well-defined. For $C_1, C_2 \in \mathcal{C}$ if $f(C_1) = f(C_2)$, then $C_1 = C_2$ since otherwise we'd get that one is contained in the other which is a contradiction. Thus f is injective. So $|\mathcal{C}| \leq r$.
Note that

$$\sum_{C \in \mathcal{C}} \frac{1}{|C|} \leq \sum_{C \in \mathcal{C}} \frac{1}{r} = |\mathcal{C}| \cdot \frac{1}{r} \leq r \cdot \frac{1}{r} = 1.$$

The third assertion follows. □

Theorem : 2

Let $\mathcal{A} \subset X^{(\leq n/2)}$ be an intersecting Sperner family. Then

$$\sum_{A \in \mathcal{A}} \binom{n-1}{|A|-1}^{-1} \leq 1. \tag{4}$$

If we have equality in (4) then $\mathcal{A} \subset X^{(r)}$ for some r, $1 \leq r \leq n/2$.

Proof. Suppose $\mathcal{A} \neq \varnothing$ and $n \geq 4$. Let $A \in \mathcal{A}$. Assign a weight of $\frac{1}{|A|}$ to every cyclic order π on X under which A is an arc. There are $|A|!(n-|A|)!$ such permutations. So the total weight due to A across all permutations is $\frac{1}{|A|} \cdot |A|!(n-|A|)! = (|A|-1)!(n-|A|)!$. We know that for each permutation (there's $(n-1)!$ of them) the sum of weights from $\mathcal{A}$ is not bigger than 1 by Lemma 1. Hence

$$\sum_{A \in \mathcal{A}} (|A|-1)!(n-|A|)! \leq (n-1)!$$

which is equivalent to what we set out to prove. □

The EKR Theorem below is an immediate consequence of Theorem 1.

Theorem : 3 (Erdos-Ko-Rado)

Let $1 \leq r \leq n/2$ and let $\mathcal{A} \subset X^{(\leq r)}$ be an intersecting Sperner family. Then $|\mathcal{A}| \leq \binom{n-1}{r-1}$.

Proof. Recall that from the previous result we have

$$\sum_{A \in \mathcal{A}} \binom{n-1}{|A|-1}^{-1} \leq 1.$$

Hence we can get bounds for $|\mathcal{A}|$ by considering $\min \binom{n-1}{|A|-1}^{-1}$ and $\max \binom{n-1}{|A|-1}^{-1}$. Since $r \leq n/2$, we have that $\min \binom{n-1}{|A|-1}^{-1} = \binom{n-1}{r-1}^{-1}$ (the further from the middle, the longer the

binomial coefficients). Putting everything together we have:

$$1 \geq \sum_{A \in \mathcal{A}} \binom{n-1}{|A|-1}^{-1} \geq |\mathcal{A}| \binom{n-1}{r-1}^{-1}.$$

Rearrange and we get the desired result. □

In their original paper, Erdös, Ko and Rado wanted to extend their theorem to larger intersections. They hoped to show that if $\mathcal{A} \subset X^{(r)}$ in such that $|A \cap B| \leq l$ for all $A, B \in \mathcal{A}$ and $n \geq n_0(r,l)$ then $|\mathcal{A}| \leq \binom{n-l}{r-l}$. [As the family $\{A \in X^{(r)} : A \supset [l]\}$ shows, this bound would be best possible]. Unsurprisingly, EKR hoped to prove this with a not-too-large (ideally, smallest-possible) threshold function $n_0(k,l)$. Unfortunately they never managed to prove this extension, so eventually (over twenty years later) they published this result with a rather large threshold $n_0(r,l)$. Here we shall give a trivial result about l-intersections with an even worse threshold $n_0(r,l)$. The proof below uses a (perhaps "the") mindless head-on attack.

Theorem : 4

Let $2 \leq l < r$ be fixed. There is a constant $n_0(r,l)$ such that if $n \geq n_0(r,l)$ and any two sets in $\mathcal{A} \subset X^{(r)}$ intersect in at least l elements, then $|\mathcal{A}| \leq \binom{n-l}{r-l}$.

Proof. If some l-set is contained in all $A \in \mathcal{A}$, we're done. Otherwise, assume $\mathcal{A}$ is maximal. So $|A_1 \cap A_2| = l$ and $|A_1 \cap A_2 \cap A_3| < l$ for some A_1, A_2, A_3. This means that for all $A \in \mathcal{A}$ we have that A meets $U = A_1 \cup A_2 \cup A_3$ in at least $l+1$ places. Since $|U| \leq 3r$ we have that

$$|\mathcal{A}| \leq \binom{3r}{l+1}\binom{n}{r-l-1} < \binom{n-l}{r-l}$$

when n is large enough. □

This result tells us that, for $1 \leq l < r$ fixed, $\mathcal{A}_0 = \{A \in X^{(r)} : [l] \subset A\}$ is an extremal family whenever $n \geq n_0(r,l)$. But what about the smaller values of n? It is easily seen that of n? It is easily seen that if n if hardly larger than r then $\mathcal{A}_0$ cannot be extremal. Indeed, there are several other possibilities. For $0 \leq t \leq r-l$ set

$$\mathcal{A}_t = \{A \in X^{(r)} : |\mathcal{A} \cap [l_2 t]| \geq l + t\}.$$

Then each $\mathcal{A}_t$ is an l-intersecting family, so a natural candidate for an extremal l-intersecting family of r-sets. As it happens, $\mathcal{A}_0$ is the largest of these families if, and only if, $n > (l+1)(r-l+1)$, and Frankl proved in 1978 that in this case $\mathcal{A}_0$ is the unique extremal family.
Concerning the general case, Frankl conjectured that the maximal size of an l-intersecting family $\mathcal{A} \subset X^{(r)}$ is

$$\max_{0 \leq t \leq r-l} |\mathcal{A}_t|,$$

where the family $\mathcal{A}_t$ is given above. This daring conjecture was proven by Ahlswede and Khachatrian in 1997.

Changing intersections to unions, given $r \geq \frac{n}{2}$, at most how large is $\mathcal{F} \subset X^{(r)}$ if X is not the union of any two members of $\mathcal{F}$? Taking complements, we see that we have already answered this question. Indeed, setting $\mathcal{A} = \{X \setminus F : F \in \mathcal{F}\}$, the family $\mathcal{F}$ is intersecting if and only if no two members of $\mathcal{F}$ have X as their union. Therefore the EKR theorem tells us that $|\mathcal{F}| \leq \binom{n-1}{n-r-1} = \binom{n-1}{r}$. But at most how large is $\mathcal{F} \subset X^{(r)}$ if $k \geq 3$, $kr \geq n$ and no k members of $\mathcal{F}$ have X as their union? It is rather surprising that, as shown by Frankl, this question is easy to answer.

Theorem : 5

Let $2 \leq k$, $r < n \leq kr$, and let $\mathcal{F} \subset X^{(r)}$ be such that X is not the union of k sets in $\mathcal{F}$. Then

$$|\mathcal{F}| \leq \binom{n-1}{r}.$$

Proof. We may assume that $k = \lceil n/r \rceil$, i.e. k is as small as possible, subject to $kr \geq n$. Equivalently, $(k-1)r < n \leq kr$ - otherwise we may decrease k to obtain a stronger assertion. Let Π be the set of all $(n-1)!$ cyclic permutations of X. For $\pi \in \Pi$, let $\mathcal{F}_\pi$ be the set of π-arcs obtained from the sets in $\mathcal{F}$. If $F \in \mathcal{F}$ is a π-arc then we denote this π-arc by F_π.
Claim. $|\mathcal{F}_\pi| \leq n - r$ for every π.
We may assume that π is the standard (identity) order $1, 2, \ldots, n$ on X, the family $\mathcal{F}_\pi$ is not empty, and $A_0 = \{n-r+1, n-r+2, \ldots, n\} \in \mathcal{F}$.
Set $|\mathcal{F}_\pi| = l$. Assign to each π-arc F_π its last element, except assign the set $\{n, n+1, \ldots, kr\}$ to A_0. Thus in total we have assigned a set K of $l + kr - n$ elements to $\mathcal{F}_\pi$. Note that the set $[n]$ is a subset of $[kr]$; also having fixed the arcs F_π, we no longer care about the cyclic order on $[n]$.
As $X = [n]$ is not the union of k arcs, if $1 \leq j \leq r$ then

$$|K \cap \{j, j+r, \ldots, j+(k-1)r\}| \leq k-1.$$

Indeed, otherwise the intersection would give us k arcs of F_π, which would cover X. Since the sets $\{j, j+r, \ldots, j+(k-1)r\}$ partition $[kr]$ and $[kr]$ contains K, this tells us that $|K| = l + kr - n \leq (k-1)r$, i.e. $l \leq n - r$.
From here the theorem follows by a simple double counting. Write P for the set of pairs $\{(F, \pi) : \pi \in \Pi, F_\pi \in \mathcal{F}\}$. Then by the claim above,

$$|P| = |\mathcal{F}| r!(n-r)! = \sum_{\pi \in \Pi} |\mathcal{F}_\pi| \leq (n-1)!(n-r),$$

so $|\mathcal{F}| \leq \binom{n-1}{r}$ as required □

This theorem is another extension of the EKR theorem, let's state this extension explicitly.

Corollary : 6

Let $k \geq 2$, $2 \leq r \leq \frac{k-1}{k}n$ and let $\mathcal{A} \subset X^{(r)}$ be such that $\cap_{i=1}^{k} A_i \neq \varnothing$ whenever $A_1, \ldots, A_k \in \mathcal{A}$. Then $|\mathcal{A}| \leq \binom{n-1}{r-1}$.

Proof. Set $\mathcal{F} = \{X \setminus A : A \in \mathcal{A}\}$. Then $\cup_{i=1}^{k} F_i \neq X$ whenever $F_1, \ldots, F_k \in \mathcal{F}$ so $|\mathcal{A}| = |\mathcal{F}| \leq \binom{n-1}{n-r} = \binom{n-1}{r-1}$. □

3. The Cube and the Kruskal-Katona Theorem

Our aim in this section is to prove two more fundamental results in the combinatorics of set systems: the Vertex Isoperimetric Theorem in the Cube and the Kruskal-Katona Theorem about shadows.

An isoperimetric inequality concerns the relationship between the size of a "body" and the size of its "boundary", no matter how the body and boundary are defined. If the original body is large, at least how large is the boundary?

Our main concern will be the discrete case. Given a graph $G = (V, E)$, the **vertex isoperimetric problem** in G is as follows: given a set $A \subset V$ of a vertices, at least how large is its **closed neighborhood**

$$N(A) = \{x \in V : d(x, A) \leq 1\}$$

in terms of $a = |A|$?

In a variant of this problem, we start with a bipartite or r-partite graph, and study the neighborhood of a set A contained in one of the parts. Thus, let G be a bipartite graph with bipartition (V_1, V_2) and edge set E. Given a set $A \subset V_1$ with a vertices, we should like to know at least how large the open neighborhood

$$\Gamma(A) = \{y \in V_2 : xy \in E \text{ for some vertex } x \in A\}$$

is in terms of a. (Here it is nicer to take the open neighborhood since the closed neighborhood of a set $A \subset V_1$ is just $N(A) = A \cup \Gamma(A)$.) In the classical case, when we are interested in what happens in the plane, the 'real' solution of the isoperimetric problem is that the perimeter length of an open set in the plane with area a is at least as large as the perimeter of a circular disc of area a. Similarly, in the n-dimensional cube Q_n, we wish to find a set of a given size whose neighborhood is as small as possible. Thus our vertex isoperimetric problem in the cube Q_n, considered as a graph is as follows: given an integer a, $1 \leq a \leq 2^n - 1$, find a set $I_a \subset Q_n$ with $|I_a| = a$ such that

$$|N(A)| \geq |N(I_a)|$$

for all $A \subset Q_n$ with $|A| = a$.

Harper proved that the sets I_a can be chosen to be nested, so that there is an order on the set of vertices in which (our chosen) I_a is precisely the initial segment of size a in this order. If $a \geq 2^n - n$ then $N(A) = Q_n$ for every set A with a vertices. If $a = 2^n - n - 1$ then there's

only one extremal set (up to isomorphism).
The canonical order solving the vertex-isoperimetric problem in the cube Q_n is the **simplicial order**, in which for $x, y \in Q_n$ we have $x < y$ if $|x| < |y|$ or $|x| = |y| = r$ and $x < y$ in the lex order on $X^{(r)}$.
In the **lex order** on $X^{(r)}$ we have $x < y$ if $\min x \Delta y \in x$, i.e. in the first place where x and y disagree, x has the smaller value. $x < y$ in the **colex order** if $\max x \Delta y \in y$.
So, for example, the simplicial order on Q_4 is: $\varnothing, 1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234, 1234$.
It's easy to see that the neighborhood of an initial segment is an initial segment.
It's clear from above that we view $Q_n \simeq \{0,1\}^n \simeq \mathcal{P}(n)$ as the power set $\mathcal{P}(n)$ and the vertices of Q_n are subsets of $X = [n]$. Unlike previous sections, we use single capital letters for subsets of the cube. To reduce the clutter, we write 1234 for $\{1, 2, 3, 4\}$, $x - i$ for $x \setminus \{i\}$, etc.
We shall prove Harper's theorem by induction so in our lemmas we'll assume that in Q_{n-1} the initial segments of the simplicial order are best for the vertex isoperimetric inequality.
To define a **1-codimensional i-compression** or simply **i-compression**, we partition the cube into two faces in every direction. For $A \subset Q_n \simeq \mathcal{P}(n)$ and $1 \le i \le n$, define

$$A_-^{(i)} = \{x : x \in A, i \notin x\} \subset \mathcal{P}(x \setminus \{i\})$$

$$A_+^{(i)} = \{x - i : x \in A, i \in x\} \subset \mathcal{P}(x \setminus \{i\})$$

Here $A_-^{(i)}$ is the part of A on the **lower face** $(Q_n)_-^{(i)}$ of Q_n in direction i, and $A_+^{(i)}$ is the **upper face** $(Q_n)_+^{(i)}$. Most importantly, $A_+^{(i)}$ and $A_-^{(i)}$ are sets in the cube $\mathcal{P}(x \setminus \{i\})$. We emphasize that the neighborhood of $A_\pm^{(i)}$ is taken in $\mathcal{P}(x \setminus \{i\})$.
A point $x \in (Q_n)_+^{(i)}$ belongs to $N(A)_+^{(i)}$ if it's in $N(A_+^{(i)})$ or $A_-^{(i)}$. The same holds with $+$ and $-$ interchanged.
For $1 \le i \le n$ we define the **i-compression** $C_i(A)$ of a set $A \subset Q_n$ to be the set $B \subset Q_n$ such that $B_\pm^{(i)}$ is the initial segment of length $|A_\pm^{(i)}|$ in the simplicial order on $\mathcal{P}(x \setminus \{i\})$.

Lemma : 1

Let $A \subset Q_n$ and $1 \le i \le n$. Then $|N(C_i(A))| \le |N(A)|$. The i-compression does not increase the size of the neighborhood.

Proof. Set $B = C_i(A)$. Since $B_\pm^{(i)}$ and $N(B_\pm^{(i)})$ are initial segments, they are nested. By our induction hypothesis and the characterization of $N(A)_\pm^{(i)}$ we have:

$$\begin{aligned}
|N(B)| &= |N(B_-^{(i)}) \cup B_+^{(i)}| + |N(B_+^{(i)}) \cup B_-^{(i)}| \\
&= \max\{|N(B_-^{(i)}|, |B_+^{(i)}|\} + \max\{|N(B_+^{(i)}|, |B_-^{(i)}|\} \\
&\le \max\{|N(A_-^{(i)}|, |A_+^{(i)}|\} + \max\{|N(A_+^{(i)}|, |A_-^{(i)}|\} \qquad (5) \\
&\le |N(A_-^{(i)}) \cup A_+^{(i)}| + |N(A_+^{(i)}) \cup A_-^{(i)}| \\
&= |N(A)_-^{(i)}| + |N(A)_+^{(i)}| = |N(A)|
\end{aligned}$$

as claimed. □

Call a set $A \subset Q_n$ **i-compressed** if $C_i(A)$, and call it **compressed** if it is i-compressed for every i. It is just about trivial that after a sequence of compressions a set becomes compressed, so if we want to determine the minimal size of a boundary of a set with m elements then we may assume that our set is compressed. In an ideal world, only the simplicial initial segments would be compressed, so the lemma below would solve our isoperimetric problem.

Lemma : 2

For every set $A \subset Q_n$ there is a compressed set $B \subset Q_n$ with $|B| = |A|$ and $|N(B)| \leq |N(A)|$.

Proof. Let $A_0 = A, A_1, \ldots, A_t$ be a sequence such that for every s, $0 \leq s < t$ there is a direction i such that $A_{s+1} = C_i(A_s)$ and $A_{s+1} \neq A_s$. This sequence must end, since a compression moving a set A_s takes it closer to the beginning of the simplicial order. At each stage we have $|A_{s+1}| = |A_s|$ and $|N(A_{s+1})| \leq |N(A_s)|$ so $B = C_t$ has the required properties. □

To formalize the proof above, write $n(a)$ for the order of a in the simplicial order. Thus, in $\mathcal{P}(n)$ we have $n(\varnothing) = 1$, $n(1) = 2, \ldots, n(n) = n+1, n(12) = n+2, \ldots, n(1n) = 2n, \ldots$. Also write $n(A) = \sum_{a \in A} n(a)$ for a set $A \subset Q_n$. Note that A is an initial segment if $n(A) = \min\{n(B) : B \subset Q_n, |B| = |A|\}$. This tells us that if $A_{s+1} = C_i(A_s)$ and $A_{s+1} \neq A_s$ then $n(A_{s+1}) < n(A_s)$ so the sequence $A_0, A_1, \ldots$ must end.
Does Lemma 2 complete our proof? Are the initial segments in the simplicial order the only compressed sets? Not quite: In Q_3, say, $\{\varnothing, 1, 2, 12\}$ is compressed and in Q_4 $\{\varnothing, 1, 2, 3, 4, 12, 13, 23\}$ is compressed. And these examples can be extended to the following exceptional sets: for n odd, $n = 2k+1 \geq 3$ the bottom half-cube, i.e. the initial segment of length 2^{n-1} in the simplicial order on Q_n is $H_n = X^{(\leq k)}$, which ends in $x = ((k+2)\ldots(2k)(2k+1))$ and is followed by $y = (12\ldots k(k+1))$ [the set x is the last set with k elements and y is the first with $k+1$.] Exchanging x and y we obtain

$$E_n = (H_n \cup \{y\}) \setminus \{x\}.$$

Then E_n is compressed but not initial.
Similarly, for n even, $n = 2k \geq 4$, the set

$$H_n = X^{(\leq k-1)} \cup ((X \setminus \{1\})^{(k-1)} + 1)$$

is the half-cube in the simplicial order on Q_n. The last element of H_n is $x = (1(k+2)(k+3)\ldots(2k))$ and the successor of x is $y = 23\ldots(k+1)$. Exchange x and y to obtain

$$E_n(H_n \cup \{y\}) \setminus \{x\}.$$

Then E_n is compressed but not initial.

Lemma : 3

Let $B \subset Q_n$ be compressed but not an initial segment in the simplicial order on Q_n. Then B is exactly the exceptional set E_n defined above.

Proof. As B is not an initial segment, there are points $x, y \in Q_n$ such that $x < y$, $x \notin B$ and $y \in B$. For $1 \leq i \leq n$ we cannot have $i \in x$ and $i \in y$, since B is compressed. Similarly, we cannot have $i \notin x$ and $i \notin y$. Hence $y = x^c$ (complement). We have that y is the successor of x in the simplicial order and since $y = x^c$ we must have that $B = E_n$. □

Luckily for us, the neighborhood of E_n is $\lfloor \frac{n-1}{2} \rfloor$ larger than the neighborhood of the half-size initial segment of Q_n, so putting together the lemmas above we find the simplicial initial segments are indeed the solution of the vertex isoperimetric problem in the cube.

Theorem : 4 (Harper)

Let $A \subset Q_n$ and let B be the initial segment of length $|A|$ in the simplicial order on Q_n. Then

$$|N(B)| \leq |N(A)|.$$

In particular, if

$$|A| = \sum_{i=0}^{r} \binom{n}{i} \text{ then } |N(A)| \geq \sum_{i=0}^{r+1} \binom{n}{i}.$$

Proof. For $n \leq 2$ the assertion is trivial, so we take $n \geq 3$. By lemma 2 we may assume that A is compressed. If A is an initial segment, we're done. If not, by lemma 3 A is the exceptional set E_n whose vertex neighborhood is $\lfloor \frac{n-1}{2} \rfloor$ larger than that of the corresponding initial segment. □

Now we turn to the other main result of this section, the Kruskal-Katona theorem. This result concerns set systems and level sets, so we change our notation and terminology: a vertex of the cube becomes a set, and a subset of the cube a family of sets. As usual, we take $X = [n]$, although all we need is that X is a linearly ordered finite set. What Corollary 4 of Section 1 tells us is that if the level sets $X^{(r)}$ and $X^{(s)}$ are endowed with the uniform probability measure then the probability that a random s-set is contained in an r-set belonging to $\mathcal{A}$ is at least the probability that a random r-set belongs to $\mathcal{A}$. Our aim here is to prove a much sharper inequality - in fact, a best possible inequality. As we shall see, it suffices to jump down by only one level, from $X^{(r)}$ to $X^{(r-1)}$, because the general case, jumping from $X^{(r)}$ to $X^{(s)}$ follows from this in $r - s$ steps. When considering $X^{(r)}$ and $X^{(r-1)}$, it is customary to talk about "shadows". As these are our main concern below we define them precisely.
For $1 \leq r \leq n$, the **lower shadow** or simply **shadow** of a family of sets $\mathcal{A} \subset X^{(r)}$ is

$$\underline{\partial}^{(r)}\mathcal{A} = \{B \in X^{(r-1)} : B \subset A, A \in \mathcal{A}\},$$

so by Corollary 4 of Section 1,

$$|\underline{\partial}^{(r)}\mathcal{A}| \geq \frac{r}{n-r+1}|\mathcal{A}|.$$

Our aim is to do better, to prove a best possible isoperimetric inequality. In other words, we should like to determine the "shadow function" $\partial^{(r)}(m)$, i.e. the function $\partial^{(r)}(m)$ such that if $\mathcal{A} \subset X^{(r)}$ and $|\mathcal{A}| = m$ then

$$|\underline{\partial}^{(r)}\mathcal{A}| \geq \partial^{(r)}(m),$$

with equality for some family $\mathcal{A} \subset X^{(r)}$ where $|\mathcal{A}| = m$. A priori it is not clear that there is such a function: after all, this function could easily depend on n as well. However, as we shell see, there is such a shadow function $\partial^{(r)}(m)$: if $1 \leq m \leq \binom{n}{r}$ then

$$\partial^{(r)}(m) = \min\{|\underline{\partial}^{(r)}\mathcal{A}| : \mathcal{A} \subset [n]^{(r)}, |\mathcal{A}| = m\},$$

independently of n.

This problem of bounding $|\underline{\partial}^{(r)}\mathcal{A}|$ for a family $\mathcal{A} \subset X^{(r)}$ of m r-sets is perhaps the most natural isoperimetric problem for families of sets. Much of the time, the notation $\underline{\partial}^{(r)}$ will be abbreviated to ∂ since it is usually clear we are considering the lower shadow of a collection of r-sets. We shall use $\underline{\partial}^{(r)}$ a little longer, since in a moment we shall use the upper shadow operator $\overline{\partial}^{(n-r)}$ as well, where, for $\mathcal{H} \subset X^{(n-r)}$ we have $\overline{\partial}\mathcal{H} = \{K \in X^{(n-r+1)} : K \supset H, H \in \mathcal{H}\}$.

Note that for $r = 2$ our shadow problem is trivial; in that case the question is simply the following: if a graph has m edges, at least how many vertices does it have? Turning it around, if a graph has k vertices, at most how many edges does it have? Clearly, the answer is "at most $\binom{k}{2}$ edges" so the shadow function $\partial^{(2)}(m)$ is as follows:

$$\partial^{(2)}(m) = k \text{ if } \binom{k-1}{2} < m \leq \binom{k}{2}.$$

For $r \geq 3$ the problem is not nearly as easy. As we shall see, the Kruskal-Katona theorem determines the exact shadow function $\partial^{(r)}(m)$ for all r and m.

We have already encountered the two most frequently used linear orders on $X^{(r)}$: the lex (lexicographic) order and the colex (colexicographic) order. $A < B$ in lex if $\sum_{i \in A} 2^{-i} > \sum_{i \in B} 2^{-i}$, and $A < B$ in colex if $\sum_{i \in A} 2^i < \sum_{i \in B} 2^i$. So in lex the slogan is "the small elements should be as small as possible", while in colex "the large elements should be as small as possible".

As we shall see, for every m the initial segment of $X^{(r)}$ of length m in the colex order has as small a lower shadow as any family $\mathcal{A}$ in $X^{(r)}$ with m elements - this is the content of the Kruskal-Katona theorem. That the families with smallest shadows, the initial segments in the colex order, are nested is very fortunate: it helps greatly in our solution of our isoperimetric problem - in fact, it makes the proofs unreasonably easy. In fact, as we shall see now, the Kruskal-Katona theorem is an immediate consequence of the Vertex Isoperimetric Theorem.

Theorem : 5 (Kruskal-Katona)

Let $\mathcal{A} \subset X^{(r)}$, $1 \leq r \leq n$, and let $I_a^{(r)}$ be the initial segment of length $a = |\mathcal{A}|$ in the colex order on $X^{(r)}$. Then

$$|\underline{\partial}^{(r)}\mathcal{A}| \geq |\underline{\partial}^{(r)} I_a^{(r)}|.$$

Proof. Let $J_a^{(n-r)}$ be the initial segment of length a in the lex order on $X^{(n-r)}$, and set $\mathcal{C} = X^{(\geq r+1)} \cup \mathcal{A}$ and $\mathcal{D} = X^{(\leq n-r-1)} \cup J_a^{(n-r)}$. Clearly, $|\mathcal{C}| = |\mathcal{D}|$, so the vertex isoperimetric inequality in the cube tells us that $|N(\mathcal{C})| \geq |N(\mathcal{D})|$. As it is easy to check that $|\underline{\partial} I_a^{(r)}| = |\overline{\partial} J_a^{(n-r)}|$, we find that

$$|N(\mathcal{C})| = |X^{(\geq r)}| + |\underline{\partial}\mathcal{A}| \geq |N(\mathcal{D})| = |X^{(\leq n-r)}| + |\overline{\partial} J_a^{(n-r)}|,$$

so

$$|\underline{\partial}\mathcal{A}| \geq |\overline{\partial} J_a^{(n-r)}| = |\underline{\partial} I_a^{(r)}|,$$

as claimed. □

4. Sumsets

In the classical examples of isoperimetric inequalities in analysis and geometry often the most elegant way of proving them is by the use of sumsets. Given sets $A, B \subset \mathbb{R}^n$, the **Minkowski sum** of A and B is the set $\{x+y : x \in A, y \in B\}$. Needless to say, $\mathbb{R}^n$ can be replaced by any Abelian group. The fundamental inequality about sums is the Brunn-Minkowski inequality.

Theorem : 1 (Brunn-Minkowski)

Let A and B be nonempty compact subsets of $\mathbb{R}^n$. Then

$$|A + B|^{1/n} \geq |A|^{1/n} + |B|^{1/n}.$$

Here $|C|$ stands for the Lebesgue measure or volume of a set C. Furthermore, equality holds if and only if A and B are homothetic convex sets.

We leave the proof as an exercise. In the proof it may be assumed that the sets A and B are closures of their interiors, or even that they are unions of rectangular parallelopipeds.
In the discrete case, which is our interest, such inequalities can be trivial, although not as clean as BM. However, most of the time they lead to deep questions. Here we shall hardly scrape the surface by proving some of the simplest inequalities. We start with the fundamental (and terribly easy) Cauchy-Davenport theorem.
We shall consider additively written Abelian groups: we call them **additive groups**. Also, an **additive set** is a finite nonempty subset of an additive group.
Given additive sets A, B, their sum (also called 'Minkowski sum') is

$$A + B = \{a + b : a \in A, b \in B\}.$$

In this very brief chapter we are interested in lower bounds on the number of elements $|A+B|$ in the sumset $A + B$ in terms of $|A|$ and $|B|$. Let us mention now that the case when the ambient groups is cyclic is by no means trivial, and indeed this is often the most interesting case for the results in this chapter.
If A, say, consists of one element then, trivially, $|A+B| = |B|$, so in studying this problem, we care only about the case when both A and B have at least two elements. Since translations

$A \to A+s\{a+s : a \in A\}$ and $B \to B+t$, and so $A+B \to A+B+s+t$ do not change the cardinalities of these sets, we may translate these A and B to our heart's content. Thus, if $A, B \subset \mathbb{Z}$ are non-empty finite sets then the cardinality of $A+B$ remains unchanged if we translate these sets so that they satisfy $\max A = \min B = 0$. In that case $A+B \supset A \cup B$ and $A \cap B = \varnothing$, so $|A+B| \geq |A \cup B| = |A| + |B| - 1$. Another trivial way of seeing this inequality is to enumerate the elements of A and B as $A = \{a_1 < a_2 < \cdots < a_k\}$ and $B = \{b_1 < b_2 < \cdots < b_l\}$ and then note that

$$A + B \supset \{a_1 + b_1, a_1 + b_2, \ldots, a_1 + b_l, a_2 + b_l, \ldots, a_k + b_l\}.$$

Also, for $|A|, |B| \geq 2$ (and still in $\mathbb{Z}$), we have $|A+B| = |A| + |B| - 1$ if and only if A and B are arithmetic progressions with the same common difference.
If G is finite, of order n, say, then $A+B$ need not to be nearly as large. However, if $|A| + |B| \geq n+1$ then $A+B$ is as large as possible: $A+B = G$. Indeed, $g \in A+B$ if and only if the sets A and $B_g = g - B = \{g - b : b \in B\}$ intersect. We have

$$|A \cap B_g| = |A| + |B_g| - |A \cup B_g| = |A| + |B| - |A \cup B_g| \geq n + 1 - |A \cup B_g| \geq 1.$$

So A and B_g do intersect.
On the other hand, if $|A| + |B| \leq n$, then $A+B$ can be quite small. For example, if $n = 2m$ and $A = B = \{2k : 0 \leq k < m\}$ then $A+B = A = B$. As we shall see now, the fundamental Cauchy-Davenport theorem tells us that if p is a prime then in $\mathbb{Z}_p$ a sumset $A+B$ has to be as large as in $\mathbb{Z}$, provided the trivial condition $|A| + |B| \leq p$ is satisfied. This theorem was proved by Davenport in 1935, and was discovered only some years later that it had been proved by Cauchy over one hundred years earlier, in 1813.

Theorem : 2 (Cauchy-Davenport)

Let p be a prime and A and B two non-empty subsets of $\mathbb{Z}_p$ such that $|A|+|B| \leq p+1$. Then $|A+B| \geq |A| + |B| - 1$.

Proof. Without loss of generality, we may assume that $1 \leq |A| \leq |B|$. We will induct on $|A|$.
If $|A| = 1$, then $|A+B| = |a+B| = |B| \geq |A| + |B| - 1 = |B|$.
Now assume that the theorem holds for cardinalities up to (but not including) $|A|$. Since we can translate sets without losing anything, we may assume that $0, a \in A$. Since $\mathbb{Z}_p = \{a, 2a, \ldots, (p-1)a, pa\}$ we have that there exists k such that $ka \in B$ and $(k+1)a \notin B$ (otherwise $B = \mathbb{Z}_p$). We may translate B to $B - ka$, so we will assume that $0 \in B$ and $a \notin B$. Then $1 \leq |A \cap B| < |A|$. So we can use the inductive hypothesis for $1 \leq |A \cap B| \leq |A \cup B|$. We get that $|(A \cap B) + (A \cup B)| \geq |A \cap B| + |A \cup B| - 1 = |A| + |B| - 1$. Since

$$(A \cap B) + (A \cup B) \subset A + B$$

we get that

$$|A+B| \geq |(A \cap B) + (A \cup B)| \geq |A| + |B| - 1.$$

□

Minkowski addition is trivially associative and commutative: for subsets A, B, C of an Abelian group we have $A + B = B + A$, $(A + B) + C = A + (B + C)$, etc. Thus Theorem 2 has the following immediate consequence.

Corollary : 3

Let p be a prime and let $A_1, \ldots, A_k$ be non-empty subsets of $\mathbb{Z}_p$ with $|A_1|+\cdots+|A_k| \leq p + k - 1$. Then

$$|A_1 + \cdots + A_k| \geq |A_1| + \cdots + |A_k| - k + 1.$$

Proof. Since $|A_1| + |A_2| + \cdots + |A_k| \leq p + k - 1$ we get $|A_1| + |A_2| \leq p + 1$ so $|A_1 + A_2| \geq |A_1| + |A_2| - 1$.
We induct. Assume we have the result for $k - 1$. If $|A_1 + A_2 + \cdots + A_{k-1}| + |A_k| \geq p + 1$, then $A_1 + \cdots + A_k = \mathbb{Z}_p$ so we're done. Otherwise, we can apply Cauchy-Davenport to get

$$|A_1 + \cdots + A_k| \leq |A_1 + \cdots + A_{k-1}| + |A_k| - 1 \leq |A_1| + \cdots + |A_k| - k + 2 - 1.$$

□

The Cauchy-Davenport theorem fails for general Abelian groups, even for cyclic groups. For example, as mentioned abouve, if A and B consists of all even numbers in $\mathbb{Z}_n$, with n even, then $A + B$ also consists of all even numbers, so $|A + B| = |A| = |B| = \frac{n}{2}$. More generally, if A and B are nonempty subsets of an Abelian group G and each of them is contained in a coset of the same subgroup H, then $|A + B| \leq |H|$. In 1953, Kneser proved a considerable extension of the Cauchy-Davenport theorem. Here we prove a slightly weaker version of this result.

Theorem : 4

Let A and B be additive subsets of an Abelian group G such that $|A| + |B| \leq |G|$. Then G has a proper subgroup H, $1 \leq |H| < |G|$, such that

$$|A + B| \geq |A| + |B| - |H|. \tag{6}$$

Proof. We may assume that $0 \in A \cap B$. We induct on $|B|$. If $|B| = 1$, we get the result with $H = \{0\}$.
Let $|B| \geq 2$ and suppose that (6) holds for smaller sets B. We have two cases:

- First case: $A + B - B = A$. We have $A + B \subset A + B - B$ and $A + B - B = A = A + 0 \subset A + B$. So $A + B = A + B - B = A$. Then (6) is trivial by picking H to be the subgroup generated by B. Since $H = \{\sum_{i\in I} b_i - \sum_{j\in J} b_j : b_i, b_j \in B\}$ we get that $A + H = A$. We get $2 \leq |B| \leq |H| \leq |A|$ hence H is a proper subgroup of G and $|A + B| = |A| \geq |A| + |B| - |H|$.

- Second case: $A + B - B \neq A$. We have that there exist $a \in A$ and $b, c \in B$ such that $a + b - c \notin A$. Let $d = a - c$, $A' = A \cup (B + d)$ and $B' = A \cap (B + d)$. Then

$c+d=a\in B'$, so $B'\neq\varnothing$ and $B+d\not\subset A$ since $b+d\notin A$. So we have $|B'|<|B|$. We use the inductive hypothesis to get $|A'+B'|\geq|A'|+|B'|-|H|$ for some proper subgroup $H\leq G$. $|A'|+|B'|=|A|+|B+d|=|A|+|B|$ so we have the desired result.

□

In fact, Kneser proved more, namely that if H is chosen to be the stabilizer of $A+B$, then

$$|A+B|\geq|A+H|+|B+H|-|H|\geq|A|+|B|-|H|.$$

Note that theorem 4 is stronger than Cauchy-Davenport since if our group is $\mathbb{Z}_p$ then the only choice for H is the subgroup consisting of only the identity. A little later Kneser extended his theorem to locally compact Abelian groups with the Haar measure.
Exercise 29 is another extension of the Cauchy-Davenport theorem.
Let us turn to another basic theorem about sumsets. Given a sequence of elements $(a_i)_1^m$ of an Abelian group of order n, if $m\geq n$ then some of these a_i add up to 0. Indeed, it is an easy exercise to show that there are $1\leq i\leq j\leq m$ such that $a_i+a_{i+1}+\cdots+a_j=0$. Erdos, Ginzburg and Ziv showed in 1961 that if m is large enough then we can specify the number of summands: some n of the a_i add up to 0. Here we state this theorem in its original form, for the cyclic group $\mathbb{Z}_n$.

Theorem : 5

Every sequence of $2n-1$ natural numbers has n terms whose sum is a multiple of n. Putting it slightly differently, let $n\geq 2$ be a natural number, and let $a_1,a_2,\ldots,a_{2n-1}\in\mathbb{Z}_n$. Then

$$\sum_{i\in I}a_i=0$$

for some set $I\subset[2n-1]^{(n)}$.

Proof. (i) First we prove this result when n is a prime p. We may assume that $0\leq a_1\leq\cdots\leq a_{2p-1}\leq p-1$. If any p of these terms are equal then we're done. Hence assume that $a_i<a_{i+p-1}$ for $i=1,2,\ldots,p$. For $i=1,2,\ldots,p-1$ set $A_i=\{a_i,a_{i+p-1}\}$ and $A_p=\{a_{2p-1}\}$. By Cauchy-Davenport we have

$$|A_1+\cdots+A_p|\geq|A_1|+\cdots+|A_p|-p+1=p.$$

So $A_1+\cdots+A_p=Z_p$ hence we're done.

(ii) Turning to the general case, induct on the number of prime factors of n. We'll work in $\mathbb{Z}$ for convenience. Suppose that $n=pm$ where p is a prime and $m>1$. Say we have $a_1,a_2,\ldots,a_{2n-1}$. Taking these numbers mod m the inductive hypothesis tells us that there are $2p-1$ disjoint m-subsets $S_1,S_2,\ldots,S_{2p-1}$ with sums divisible by m (say $b_1m,b_2m,\ldots,b_{2p-1}m$). Indeed, these m-subsets can be chosen since even after selecting $2p-2$ of them you still have $2n-1-(2p-2)m=2m-1$ additional terms to choose

from. Having found these sets S_i, some p of the $2p-1$ numbers b_i add up to a multiple of p. So $pm = n$ elements add up to a multiple of $pm = n$.

□

Although the proof above is very simple, later we shall give an algebraic proof as well.
There are a great many problems to do with sumsets that are still open. An important example concerns the **Davenport constant** of a (finite abelian) group G, denoted by $D(G)$. This is defined to be the least m such that any sequence of m elements from G has a non-empty subset summing to zero. Thus for example it is easy to see that the Davenport constant of the cyclic group $\mathbb{Z}_n$ is exactly n. What abut the Davenport constant of the group $\mathbb{Z}_n^k$? A moment's thought shows that this is at least $k(n-1)+1$, and this is conjectured to be the correct value. This has been proved for $k = 2$ by Olson, whos also showed it for general k in the case when n is prime. However, the conjecture is open in general. It is remarkable that the Davenport constants of these natural and uncomplicated groups are unknown.
The problem of finding $D(G)$ was proposed by Harold Davenport at a conference at Ohio State University in 1966. As remarked by Davenport on that occasion, if G is the class group of an algebraic number field $\mathbb{F}$, then $D(G)$ is the maximal number of prime ideals (counted with multiplicity) in the decomposition of an irreducible integer in $\mathbb{F}$.
Another constant defined by sumsets is the **Erdos-Ginzburg-Zif constant** $z(G)$ of a finite abelian group G: the minimal integer z such that every sequence of length z in G has a subsequence of length $\exp(G)$ summing to 0. As usual, $\exp(G)$ is the exponent of G, the lest common multiple of the orders of the elements of G. The Fundamental Theorem of Finite Abelian Groups tells us that if G is a nontrivial finite abelian group then $G = \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$ for some unique integers $1 < n_1 < n_2 < \cdots < n_r$. The integer $r = r(G)$ is the **rank** of G, and we set $d^*(G) = \sum_{i=1}^{r}(n_i - 1)$. Olson proved that for p-groups (i.e for groups in which each n_i is a power of the same prime p) the trivial lower bound we have noted for $D(G)$ is, in fact, its value.

Theorem : 6

If G is a p-group then $D(G) = d^*(G) + 1$.

Also, Caro and Gao proved independently that for an abelian group G the constants $D(G)$ and $z(G)$ are closely connected.

Theorem : 7

If G is an abelian group of order n then $z(G) = n + D(G) - 1$.

In a later chapter we shall note more recent results about the Davenport and Erdos-Ginzburg-Zif constants.

Chapter 2. Polynomials in Combinatorics

1. Alon's Combinatorial Nullstellensatz

In this section we shall study polynomials in n variables $X_1, \ldots, X_n$ over a field $\mathbb{F}$ or ring R, concentrating on the set of common zeros of our polynomials. Occasionally we shall assume that $\mathbb{F}$ is algebraically closed. We frequently write X for the n-touple $(X_1, \ldots, X_n)$, and $\mathbb{F}[X] = \mathbb{F}[X_1, \ldots, X_n]$ for the ring of polynomials. In our main applications, later in this section, $\mathbb{F}$ will be a finite field. Our notation is self-explanatory; e.g. $\deg f$ is the total degree of $f \in \mathbb{F}[X]$, and $\deg_{X_i} f$ is its degree in the varaible X_i. The degree of the zero polynomial is taken to be $-\infty$.
We start with one of the great theorems in all of mathematics, Hilbert's Nullstellensatz. This theorem is a corner-stone of Algebraic Geometry: it expresses a deep connection between geometry and algebra. This fundamental theorem, proved by David Hilbert in 1893, concerns the common zeros of polynomials over an algebraically closed field.

Let $\mathbb{F}$ be an algebraically closed field, and let I be an ideal in $\mathbb{F}[X]$. Denote by $V(I) \subset K^n$ the set of common zeros of the polynomials in I. If $f \in \mathbb{F}[X_1, \ldots, X_n]$ vanishes at every point of $V(I)$ then some power of f is in I.

The converse assertion is trivial: if $f^k \in I$ for some natural number k then f vanishes at every point of $V(I)$, since it vanishes at every zero of f^k. Thus these two properties are equivalent: f vanishes at every point $V(I)$ if and only if some power of f belongs to I.
Over the years, many proofs of Hilbert's Nullstellensatz have been given; the best known is probably the proof Zariski gave in 1947.
There is a 'weak form' of Hilbert's Nullstellensatz, which says the following:

Let $f_1, \ldots, f_m \in \mathbb{F}[X]$ have no common zero in K^n, where K is an algebraically closed extension of $\mathbb{F}$. Then the ideal these polynomials generate is the entire polynomial ring $\mathbb{F}[X]$, i.e. there are polynomials $g_1, \ldots, g_m \in \mathbb{F}[X]$ that satisfy the Bezout equation

$$1 = g_1 f_1 + \cdots + g_m f_m. \tag{7}$$

An important question concerning Bezout's equation is the following. If each polynomial f_i has degree at most d, what is the minimal δ such that we can choose polynomials g_i of degree at most δ? After numerous early results, starting with Hermann in 1926, this question was taken up by Brownawell in 1987, and a little later Kollar proved the best possible bound $\delta \leq d^{\min(m,n)} - d$ when $d \geq 3$. See also Philippon 1987 and Shiffman 1989.
In fact, the 'weak form' is not that weak since, as shown by Rabinowitsch in 1930, it easily implies the following 'strong' form of the Nullstellensatz.

Let $f, f_1, \ldots, f_m \in \mathbb{F}[X_1, \ldots, X_n]$ be such that f vanishes at all the common zeros of $f_1, \ldots, f_m \in \mathbb{F}[X]$. Then some power of f is in the ideal generated by $f_1, \ldots, f_m$.

To see that the 'weak form' implies this 'strong form', let X_0 be an additional variable, and consider the $m+1$ polynomials $f_1, \ldots, f_m$ and $X_0 f - 1$. These $m+1$ polynomials in our $n+1$ variables $X_0, X_1, \ldots, X_n$ have no common zero, since if $f_i(x) = f_i(x_0, \ldots, x_n) = 0$ for

$i = 1, \ldots, m$ and some $x \in \mathbb{F}^{n+1}$, then $f(x) = 0$, so $x_0 f(x) - 1 = -1 \neq 0$. The 'strong form' tells us that Bezout's equation (7) is solvable, i.e. there are polynomials $g_0, \ldots, g_m$ in $\tilde{X} = (X_0, \ldots, X_n)$ such that

$$1 = \sum_{i=1}^{m} g_i(\tilde{X}) f_i(X) + g_0(\tilde{X}) \cdot (X_0 f(X) - 1).$$

Setting $X_0 = 1/f(X)$, this identity becomes

$$1 = \sum_{i=1}^{m} g_i \left(\frac{1}{f(X)}, X_1, \ldots, X_n \right) f_i = \frac{\sum_{i=1}^{n} h_i f_i(X)}{f(X)^k},$$

where k is the maximal power of X_0 that occurs in the monomials of the g_i, so that $h_i \in \mathbb{F}[X]$. Hence $f^k = \sum_{i=1}^{m} h_i f_i$, as required.
The beautiful and very simple proof we have just given is usually referred to, rather dismissively, as the 'Rabinowitsch Trick'.
Let us turn to the main topic of this section, Alon's Combinatorial Nullstellensatz, which first appeared in print in 1999, although Alon had presented it at a conference four years earlier.

Theorem : 1 (Alon's Combinatorial Nullstellensatz)

For $i = 1, \ldots, n$, let S_i be a non-empty finite subset of a field $\mathbb{F}$ with $|S_i| = d_i + 1$, and let $f \in \mathbb{F}[X] = \mathbb{F}[X_1, \ldots, X_n]$ be a polynomial of degree $d = \sum_{i=1}^{n} d_i$ in whose expansion the coefficient of the monomial $X_1^{d_1} \cdots X_n^{d_n}$ is non-zero. Then f is not identically zero $S_1 \times \cdots \times S_n$.

In 2010, Michalek surprised the world of combinatorics by giving a very simple and extremely elegant proof of Alon's Combinatorial Nullstellensatz. As we have already remarked, this is the second proof we shall present. In this proof we use division by a monic linear polynomial $X - t$ in the ring $R[X, Y, \ldots]$ of polynomials of several variables over a ring R. All we should note is that for $k \geq 1$ and $t \in R$ we have

$$X^k = (X - t)(X^{k-1} + tX^{k-2} + \cdots + t^{k-1}) + t^k = (X - t)q + r$$

where q is a monic polynomial of degree $k - 1$ in $R[X]$ and the remainder r is in R.

Proof. We apply induction on d, the degree of f, starting with $d = 1$ which is trivial. Assume $\deg f = d > 1$ and the assertion holds whenever $\deg f < d$. We may assume that the expansion of f contains a monomial $X_1^{d_1} \ldots X_n^{d_n}$ with coefficient 1, with $d = \sum_{i=1}^{n} d_i$ and $d_1 \geq 1$. Finally, aiming for a contradiction, let's assume f vanishes on $S_1 \times S_2 \times \cdots \times S_n$. Let $s_1 \in S_1$ and divide f by $X_1 - s_1$:

$$f = (X_1 - s_1)q + r \tag{8}$$

where $q \in \mathbb{F}[X_1, \ldots, X_n]$, $r \in \mathbb{F}[X_2, \ldots, X_n]$ and $\deg q = d - 1$. Furthermore the coefficient of $X_1^{d_1 - 1} X_2^{d_2} \ldots X_n^{d_n}$ in q is also 1.

Pick a point $x \in \{s_1\} \times S_2 \times \cdots \times S_n$ and substitute it into (8). Since $f(x) = 0$ we get $r(x) = 0$ and since r doesn't depend on X_1 we have that r vanishes on the entire set $S_1 \times S_2 \times \cdots \times S_n$. As $s - s_1 \neq 0$ whenever $s \in S_1 \setminus \{s_1\}$, the polynomial q vanishes on $(S_1 \setminus \{s_1\}) \times S_2 \times \cdots \times S_n$ contradicting the inductive hypothesis thus we hav e a contradiction. □

As one of the many applications of his Combinatorial Nullstellensatz to Algebra, Alon gave a short proof of the classical theorem of Chevalley. The result was conjuctered by Artin in 1934 and proved by Chevalley and Warning a year later, the former in its simple form below and the latter in a sharper form, giving the best bound on the number of common zeros.

Theorem : 2

Let $\mathbb{F}$ be a finite field of order q and characteristic p, and let $f_1, \ldots, f_m \in \mathbb{F}[X] = \mathbb{F}[X_1, \ldots, X_n]$ be polynomials with a common zero such that $\sum_{i=1}^{m} \deg f_i < n$. Then these polynomials $f_1, f_2, \ldots, f_m$ have another common zero.

Proof. Suppose, by way of contradiction, that they have only one common zero. Assume this common zero is $0 = (0)_1^n \in \mathbb{F}^n$. Take

$$f(X) = \prod_{i=1}^{m}(1 - f_i(X)^{q-1}) + \gamma \prod_{j=1}^{n} \prod_{s \in \mathbb{F}^\times} (X_j - s)$$

with γ such that $f(0) = 0$. Thus f is identically zero on $\mathbb{F} \times \cdots \times \mathbb{F}$. We have $\deg(\prod_{j=1}^{m}(1 - f_i(X)^{q-1})) < n(q-1)$ and $\deg(\gamma \prod_{j=1}^{n} \prod_{s \in \mathbb{F}^\times}(x_j - s)) = (q-1)n$. So the degree of f is $(q-1)n$. The coefficient of $X_1^{q-1} \ldots X_n^{q-1}$ is $\gamma \neq 0$. By ACNS we have a contradiction. Thus there's another zero. □

In 1964 James Ax gave a very simple proof of Warning's first theorem for one polynomial, sharpening Chevalley's theorem above for $m = 1$.

Theorem : 3

Let $\mathbb{F}$ be a finite field of characteristic p and order $q = p^l$, and let $f \in \mathbb{F}[X_1, \ldots, X_n]$ have degree $d < n$. Then $N(f)$, the number of zeros of f is a multiple of p.

Proof. We want to show that $N(f) = 0$ in $\mathbb{F}$. We have $1 - f(x)^{q-1} = \begin{cases} 1 \text{ if } f(x) = 0 \\ 0 \text{ else} \end{cases}$ So $N(f) = \sum_{x \in \mathbb{F}^n}(1 - f(x)^{q-1}) = -\sum_{x \in \mathbb{F}^n} f(x)^{q-1}$. Now $f(X)^{q-1}$ has degree $d(q-1) < n(q-1)$ so it's an $\mathbb{F}$-linear combination of monomials of degree less than $n(q-1)$. If $X^u = X_1^{u_1} \ldots X_n^{u_n}$ is such a monomial then one of u_i is less than $q-1$. Assume $u_1 \leq q-2$. We have $S(u) = \sum_{x \in \mathbb{F}^n} x^u = \prod_{i=1}^{n} \sum_{x_i \in \mathbb{F}} x_i^{u_i} = \sum_{x \in \mathbb{F}} x^{u_1} \prod_{i=2}^{n} \sum_{x_i \in \mathbb{F}} x_i^{u_i} = 0$. Therefore $N(f) = 0$ in $\mathbb{F}$. □

Let us end this section with the Chevalley-Warning theorem about the number of common zeros of several polynomials over the field $\mathbb{F}_p$ of prime order p. We shall prove this result by applying ACNS.

Theorem : 4

Let p be a prime and let $f_1, \ldots, f_m \in \mathbb{F}_p[X_1, \ldots, X_n]$ be polynomials such that $\sum_{i=1}^m \deg f_i < n$. Then the number of common zeros of these polynomials $f_1, \ldots, f_m$ is a multiple of p. In particular, if they have a common zero then they have another common zero.

Proof. Let $z_1, \ldots, z_k \in (\mathbb{F}_p)^n$ be the common zeros with $z_i = (z_1^{(i)}, \ldots, z_n^{(i)})$. Let us assume that k is not a multiple of p, so that our aim is to arrive at a contradiction. Define a polynomial $f \in \mathbb{F}_p[X]$ by

$$f = \prod_{i=1}^{m}(1 - f_i^{p-1}) - \sum_{j=1}^{k}\prod_{i=1}^{n}(1 - (X_i - z_i^{(j)})^{p-1}) = g - h. \tag{9}$$

The coefficient of $X_1^{p-1}X_2^{p-1}\ldots X_n^{p-1}$ in the second summand h in (9) is $k(-1)^n \neq 0 \in \mathbb{F}_p$, and h has no other monomials of degree at least $n(p-1)$ so $\deg h = n(p-1)$. Furthermore, since

$$(p-1)\sum_{i=1}^{m}\deg f_i \leq (p-1)(n-1) < n(p-1),$$

the first summand g in (9) has no monomials of degree at least $n(p-1)$. Thus $\deg f = n(p-1)$ and the coefficient of the monomial $X_1^{p-1}X_2^{p-1}\ldots X_n^{p-1}$ in f is not zero. By ACNS, Theorem 1, f is not identically 0 on $(\mathbb{F}_p)^n$.
To arrive at a contradiction and so to complete our proof, we shall prove that every evaluation $f(x)$ of our polynomial f is 0. In showing this, we consider two cases, whether x is a common zero of our polynomials, or not.
First, what is $f(z_j)$? The first summand in (9) is 1, trivially. Also, the second is -1 since the j-th product is 1, while the other $k-1$ products are 0, because $0 \neq y \in \mathbb{F}_p$ we have $y^{p-1} = 1$. Hence $f(z_j) = 0$.
Second, what is $f(x)$ if $x \notin \{z_1, \ldots, z_k\}$? Then $f_i(x) \neq 0$ for at least one i, so $1 - f_i(x)^{p-1} = 0$, implying that $g(x) = 0$. Also, $h(x) = 0$ as well, since in each product

$$\prod_{i=1}^{n}(1 - (x_i - z_i^{(j)})^{p-1})$$

at least one of the factors is 0, because $x_i \neq z_i^{(j)}$ for at least one i. Hence $f(x) = 0$, and so f is identically zero on $(\mathbb{F}_p)^n$. This contradiction completes our proof. □

2. Applications of ACNS

Let us start with the algebraic proof of the (very very simple) Cauchy-Davenport theorem we promised in section I.4. Here and in the subsequent results in this section we keep the

notation $\mathbb{Z}_p$ for the set of integers modulo a prime p, although we make use of the fact that $\mathbb{Z}_p$ is a field.

Theorem : 1

Let p be a prime and A and B two non-empty subsets of $\mathbb{Z}_p$ such that $|A|+|B| \leq p+1$. Then $|A+B| \geq |A|+|B|-1$.

Proof. Suppose, for a contradiction, that $A+B \subset C \subset \mathbb{Z}_p$ with $|C| \leq |A|+|B|-2 \leq p-1$. Define $f \in \mathbb{Z}_p[X,Y]$ by $f(X,Y) = \prod_{c\in C}(X+Y-c)$. $f(x,y) = 0$ for $(x,y) \in A \times B$. $\deg f = |A|+|B|-2$ and the coefficient of $X^{|A|-1}Y^{|B|-1}$ in f is $\binom{|A|+|B|-2}{|A|-1} \neq 0 \in \mathbb{Z}_p$ contradicting ACNS. □

Similarly, the prime order case of the Erdos-Ginzburg-Zib theorem, Theorem 5 in I.4., is an easy consequence of Chevalley's theorem, Theorem 2 in II.1.

Theorem : 2

For a prime p, every sequence $a_1, a_2, \ldots, a_{2p-1} \in \mathbb{Z}_p$ has p terms whose sum is 0.

Proof. Let $f_1(X) = \sum_{i=1}^{2p-1} a_i X_i^{p-1} \in \mathbb{Z}_p[X_1, \ldots, X_{2p-1}]$ and $f_2(X) = \sum_{i=1}^{2p-1} X_i^{p-1} \in \mathbb{Z}_p[X_1, \ldots, X_{2p-1}]$. Note that $\deg f_1 + \deg f_2 = 2(p-1) < 2p-1$. f_1 and f_2 have 0 as a common zero. From Theorem 2 of the previous section, we have that there's another common zero $\overline{x} = (x_1, \ldots, x_{2p-1})$. $f_2(\overline{x}) = 0$ implies that none or p of x_i are nonzero. Since $\overline{x} \neq 0$, p of the x_i are nonzero. Label them as $x_{f(1)}, \ldots, x_{f(p)}$. Note that $f_1(\overline{x}) = 0$, thus $\sum_{i=1}^{p} a_{f(i)} = 0$. □

Returning to the topic of sumsets, this time we shall investigate restricted sums of sets. Given sets $A, B \subset \mathbb{Z}_p$, at least how large is their sum if we demand that the summands are different. Thus, at least how large is

$$A \oplus B = \{a+b : a \in A, b \in B, a \neq b\}?$$

Trivially, if $|A| = 1$ then $|A \oplus B| = |B|$ if $A \not\subset B$ and $|A \oplus B| = |B| - 1$ if $A \subset B$. Also, $[n] \oplus [n] = [3, 2n-1]$, so for $n \geq 2$ we have $|[n] \oplus [n]| \geq 2n-3$ provided $2n-3 \leq p$. In 1964, Erdos and Heilbronn conjectured that for $A = B$ the restricted sum is at least as large as in the case when A is an interval:

$$|A \oplus A| \geq 2|A| - 3,$$

if $2|A| - 3 \leq p$. This conjectured inequality seems to be a short step from the Cauchy-Davenport theorem but, amazingly, in spite of a number of serious attacks by many people, the Erdos-Heilbronn conjecture remained unsolved for thirty years: eventully, in 1994, Dias da Silva and Hamidoune managed to prove it. Soon after this breakthrough result, Alon, Nathanson and Ruzsa realized that in fact this conjecture is a very simple consequence of

ACNS. Thus, given the right tool, the following theorem of Alon, Nathanson and Ruzsa, implying the Dias da Silva-Hamidoune theorem, is hardly more than not a too difficult exercise. Note that if $A = [n]$ and $B = [m]$ for $n \neq m$ then $A \oplus B = [3, n+m]$ so $|A \oplus B| = n + m - 2$. The result below claims that this is an extremal example.

Theorem : 3

Let p be a prime, and A and B non-empty subset of $\mathbb{Z}_p$ with $|A| \neq |B|$ and $|A|+|B| \leq p+2$. Then

$$|A \oplus B| \geq |A| + |B| - 2.$$

Proof. Set $a = |A|$ and $b = |B|$ and assume without loss of generality that $a < b$. Suppose, by way of contradiction, that $A \oplus B \subset C \subset \mathbb{Z}_p$ such that $|C| = a+b-3$. Consider $f(X,Y) = (X-Y)\prod_{c \in C}(X+Y-c)$. This vanishes on $A \times B$. Also $\deg f = 1+a+b-3 = (|A|-1)(|B|-1)$. Need to check the coefficient of $X^{|A|-1}Y^{|B|-1}$. It is $\binom{a+b-3}{a-2} - \binom{a+b-3}{b-2} = \frac{b-a}{a-1}\binom{a+b-3}{b-2} \neq 0 \in \mathbb{Z}_p$ since $a+b-3 < p$. But this contradicts ACNS, therefore $|A \oplus B| \geq |A| + |B| - 2$. □

The following immediate consequence of this theorem is slightly stronger than the Erdos-Heilbronn conjecture, as stated above.

Corollary : 4

Let p be a prime and let A and B be non-empty subsets of $\mathbb{Z}_p$ with $|A| + |B| \leq p+3$. Then $|A \oplus B| \geq |A| + |B| - 3$.

Proof. If $|A| \neq |B|$ it's clear from theorem 3. If $|A| = |B|$ pick $B' \subset B$ such that $|B'| = |B|-1$. Then $|A \oplus B'| \geq |A|+|B'|-2 = |A|+|B|-3$. So $|A \oplus B| \geq |A \oplus B'| \geq |A|+|B|-3$. □

Our next application is very different. In 1973, Berge and Sauer made the fascinating conjecture that every 4-regular graph contains a 3-regular subgraph. (Not a spanning subgraph: it is trivial that not every 4-regular graph has a 1-factor.) Although this conjecture was proved by Tashkinov in 1982, we shall not present that results, but a close relative of it, proved by Alon, Friedland and Kalai two years later in a very simple way.

Theorem : 5

Let p be a prime, and $G = (V, E)$ a loopoless multigraph of average degree greater than $2p-2$ and maximum degree $2p-1$. Then G contains a p-regular subgraph.

Proof. Let $A = (a_{v,e})$ be the incidence matrix of G, i.e. for $v \in V$ and $e \in E$ set

$$a_{v,e} = \begin{cases} 1 & \text{if } e \text{ is incident with } v, \\ 0 & \text{otherwise} \end{cases}$$

Let us define a polynomial $f(X) = f(X_e : e \in E)$ in variables X_e indexed by the edges $e \in E$ as follows:

$$f(X) = \prod_{v \in V} \left(1 - \left(\sum_{e \in E} a_{v,e} X_e\right)^{p-1}\right) - \prod_{e \in E}(1 - X_e).$$

Writing n for the number of vertices and m for the number of edges of G, the degree of the first term in f is $(p-1)n$, and that of the second term is $m > (p-1)n$, so f has degree m. Also the coefficient of the monomial $\prod_{e \in E} X_e$ in f is $(-1)^{m+1} \neq 0$, so by ACNS f is not identically 0 on $\{0,1\}^E$; say, for $x = (x_e)_{e \in E}$, with x_e is 0 or 1 for each edge e, we have $f(x) \neq 0$. Since $f(0) = 0$, this tells us that $x \neq 0$, i.e.

$$F = \{e : x_e = 1\} \neq \varnothing.$$

Finally, all we have to observe is that in the greaph $H = (V, F)$ the degree of every vertex is a multiple of p. Indeed, then every vertex has degree 0 or p, as no vertex has degree $2p$ or more. In the evaluation of $f(X)$ at x the second term is 0, since not every x_e is 0. Hence if $f(x) \neq 0$ then in the evaluation of $f(X)$ at x the first term is non-zero. But again by Lagrange's theorem this implies that

$$\sum_{e \in E} a_{v,e} x_e = 0$$

for every vertex v, i.e. the degree of every vertex is 0 in $\mathbb{Z}_p$. □

The last result we shall present is an extension of the Erdos-Ginzburg-Ziv theorem to two dimensions. For a prime p, let $s(p,2)$ be the minimal m such that every sequence of m vectors in the two-dimensional $\mathbb{Z}_p$-vector space $\mathbb{Z}_p \oplus \mathbb{Z}_p$ (here $\oplus$ doesn't mean restricted sumset) contains p terms summing to 0. Clearly $s(p,2) \geq 4p-3$, since the sequence of length $4p-4$ in which each of the vectors $(0,0), (0,1), (1,0)$, and $(1,1)$ occurs $p-1$ times does not have p terms whose sum is 0.
In 1983, Kemnitz conjectured that this is an extremal example, so $s(p,2) = 4p-3$ for every prime p. This conjecture is trivial for $p = 2$ but the case $p = 3$ already needs some effort. In 2000, Ronyai came close to proving this conjecture: he proved that $s(p,2) \leq 4p-2$. As a preparation, we prove the following result.

Theorem : 6

Let $p \geq 2$ be a prime, and let $v_1, v_2, \ldots, v_{3p}$ be $3p$ vectors in $V = \mathbb{Z}_p \oplus \mathbb{Z}_p$ with 0 sum: $\sum_{i=1}^{3p} v_i = 0 \in V$. Show that some p of these vectors also sum to 0.

Proof. With $v_i = (a_1, b_1)$ define three polynomials over $\mathbb{Z}_p$ in $3p-1$ variables $X_1, X_2, \ldots, X_{3p-1}$:

$$f_1(X) = \sum_{i=1}^{3p-1} a_i X_i^{p-1}, \quad f_2(X) = \sum_{i=1}^{3p-1} b_i X_i^{p-1}, \quad f_3(X) = \sum_{i=1}^{3p-1} X_i^{p-1}.$$

We have $\sum \deg f_i = 3(p-1) < 3p-1$ so the number of common zeros is a multiple of p. Let $0 \neq x = (x_i)$ be a common zero. Let $J = \{j : x_j \neq 0\}$ so that $J \neq \varnothing$. Then $x_j^{p-1} = 1$ for $j \in J$. Hence $\sum_{j\in J} a_j = \sum_{j\in J} b_j = 0 \in V$ and $|J| = p$ or $2p$. In the first case simply take $I = J$. In the second case take $I = [3p] \setminus J$. □

Here is then Ronyai's theorem

Theorem : 7 (Ronyai)

Let p be a prime and let $v_1, v_2, \ldots, v_m$ be $m = 4p-2$ vectors in the vector space $V = \mathbb{Z}_p \oplus \mathbb{Z}_p$ over $\mathbb{Z}_p$. Then some p of the vectors v_i add up to 0.

Proof. Assume, by way of contradiction, that the assertion is false. Then by the result in Theorem 6, for all $J \subset [m]$ with $|J| = p$ or $|J| = 3p$ we have $\sum_{j\in J} v_j \neq 0 \in V$.
Setting $v_i = (a_i, b_i)$, $i = 1, 2, \ldots, m$ define a polynomial $f(X) \in \mathbb{Z}_p[X] = \mathbb{Z}_p[X_1, \ldots, X_m]$ as follows:

$$f(X) = \left(\left(\sum_{i=1}^{m} a_i X_i\right)^{p-1} - 1\right) \cdot \left(\left(\sum_{i=1}^{m} b_i X_i\right)^{p-1} - 1\right) \cdot \left(\left(\sum_{i=1}^{m} X_i\right)^{p-1} - 1\right) \cdot (r(X) - 2)$$

where

$$r(X) = \sum_{I\in[m]^{(p)}} \prod_{i\in I} X_i.$$

Clearly $\deg f \leq 4p-3 = m-1$. We claim that for $c = (c_i)_1^m \in \{0,1\}^m$ we have

$$f(c) = \begin{cases} 2 & \text{if } c = 0 = (0)_1^m, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Indeed, $f(0) = 2$, trivially. Also, for $c = (c_i)_1^m \in \{0,1\}^m$ the third factor in the definition of $f(c)$ is 0 unless the sum $w(c) = \sum_i c_i$ in $\mathbb{Z}$ is a multiple of p. If $w(c) = 2p$ then the fourth factor vanishes since $\binom{2p}{p} = 2$ in $\mathbb{Z}_p$, and in the other two cases our hypothesis would imply that at least one of the first two factors is 0.
The polynomial $g(X) = 2(X_1 - 1)(X_2 - 1)\ldots(X_m - 1)$ satisfies the same identity (10) as $f(X)$. So the difference $h(X) = g(X) - h(X)$ vanishes on $\{0,1\}^m$. Since g has degree m and $\deg f \leq m-1$ we have that $\deg h = m$. The coefficient of $X_1X_2\ldots X_m$ in h is nonzero so by ACNS h cannot be identically zero on $\{0,1\}^m$ but this is a contradiction. □

Chapter III - Topology in Combinatorics

1. The Kneser Graph and the LSB Theorem

A set of vertices of a graph $G = (V, E)$ is **independent** if no two of its vertices are joined by an edge. The **independence number** $\alpha(G)$ of G is the maximal cardinality of an

independent set of vertices.
A **proper** k-coloring (of the vertices) of a graph $G = (V, E)$ is a map $\varphi : V \to C$, where C is a set of k elements ("colors"), such that $\varphi(a) \neq \varphi(b)$ whenever $ab \in E$. Equivalently, a k-coloring is a partition of the vertex set into k independent sets, with empty sets allowed. The **chromatic number** $\chi(G)$ of G is the minimal k for which G has a (proper) coloring, i.e. the minimal number of independent sets covering the entire vertex set. Trivially, if $H \subset G$ then $\chi(H) \leq \chi(G)$; also, $\chi(G) \geq \dfrac{|G|}{\alpha(G)}$. [Here, as elsewhere, $|G| = |V|$ is the number of vertices.]
In 1955 Martin Kneser defined a graph which by now is called the **Kneser graph** with parameters n and k. The vertex set of this graph $KG(n, k)$ is $[n]^{(k)}$, the set of k-subsets of $[n]$, and two k-subsets are adjacent if and only if they are disjoint. For $n < 2k$ the Kneser graph $KG(n, k)$ has no edges, for $k = 1$ (and any n) it is a complete graph, and for $n = 2k$ it is the complement of a complete matching. However, for $n > 2k$ the structure of $KG(n, k)$ is fascinating. From now on we shall always assume that $n \geq 2k$. Thus the smallest (and simplest) non-trivial Kneser graph is $KG(5, 2)$, which is exactly the Petersen graph.
By definition, $\mathcal{A} \subset [n]^{(k)}$ is an independent set in $KG(n, k)$ if and only if $\mathcal{A}$ is an **intersecting family of sets**, i.e. if any two sets in $\mathcal{A}$ have non-empty intersection. Since there are $\binom{n-1}{k-1}$ k-sets in $[n]^{(k)}$ that contain a fixed element (1, say), $\alpha(KG(n, k)) \geq \binom{n-1}{k-1}$.
Also EKR tells us that

$$\alpha(KG(n, k)) = \binom{n-1}{k-1} \quad \text{if} \quad n \geq 2k \tag{11}$$

Consequently, we find that the chromatic number of $KG(n, k)$ cannot be too small: if $n \geq 2k$ then

$$\chi(KG(n, k)) \geq \frac{\binom{n}{k}}{\alpha(KG(n, k))} = \frac{\binom{n}{k}}{\binom{n-1}{k-1}} = \frac{n}{k}. \tag{12}$$

What about a proper coloring with few colors? For $d = n - 2k \geq 0$ here is a canonical coloring φ of $KG(n, k)$ with $d + 2$ colors: for $A \in [n]^{(k)}$ define

$$\varphi(A) = \begin{cases} \min A & \text{if } \min A \leq d + 1, \\ d + 2 & \text{otherwise.} \end{cases}$$

The trivial lower bound and the coloring φ tell us that

$$\frac{2k + d}{k} \leq \chi(KG(2k + d, k)) \leq d + 2$$

for $k \geq 1$ and $d \geq 0$.
Note that if both k and d are large then these two bounds are far from each other. If $d < k$, say, then there are independent sets containing more than one third of the vertices, but the canonical coloring φ uses $d+2 \gg 3$ colors. Nevertheless, Kneser conjectured that $d+1$ colors are not sufficient. It is trivial that Kneser's conjecture holds when d is 0 or 1, and it is easy to show that it holds for $k = 2$ and all $n \geq 4$, but for over twenty years no real progress was made with the general conjecture. It was quite a surprise when in 1978 Lovasz used cohomology to

prove it in full. Strangely, even before Lovasz's paper appeared Barany had found a beautiful and considerably simpler proof. What is more amazing is that in 2002 Joshua Greene, who at the time was an undergraduate at Harvey Mudd College, further simplified Barany's beautiful and simple proof, turning it into an essentially trivial argument. All one needs is that, like Barany, one should apply Borsuk's theorem. [One should emphasize that the idea that Borsuk's theorem should be applied is more than 99% of the achievement in any case.] The basic theorem goes under several different names: Borsuk's theorem, the Borsuk-Ulam theorem, the Lusternik-Schnirelman theorem, the Lusternik-Schnirelman-Borsuk theorem, or simply the LSB theorem, proved by Lusternik and Schnirelman in 1930 an by Borsuk (proving a conjecture of Ulam) in 1933. These names refer to various equivalent assertions: here we shall state the result in only two incarnations, some others will be given as exercises. So here is what we shall call the LSB theorem.

Theorem : 1 (LSB)

Let $d \geq 1$. The following equivalent assertions hold.

(i) For every continuous mapping $f : S^d \to \mathbb{R}^d$ there is a point $x \in S^d$ such that $f(x) = f(-x)$.

(ii) Let $\{V_1, \ldots, V_{d+1}\}$ be a collection of closed sets covering S^d. Then there is a set V_i that contains a pair of antipodal points, i.e. there is a point $x \in S^d$ such that $x, -x \in V_i$.

Greene noted the following slight extension of part (ii) of this theorem.

Theorem : 2

For every cover of S^d with $d+1$ sets, each of which is either open or closed, one of the sets contains a pair of antipodal points.

Proof. For $0 \leq k \leq d+1$, let $\{V_1, \ldots, V_k\} \cup \{U_{k+1}, \ldots, U_{d+1}\}$ be a cover of S^d, with each V_i closed and each U_j open. Our talk is to show that one of these $d+1$ sets contains a pair of antipodal points.
Case 1: $k = d+1$. This is part (ii) of the LSB Theorem, Theorem 1.
Case 2: $k = 0$. In this case $\{U_1, \ldots, U_{d+1}\}$ is an open cover of S^d. Let $\lambda > 0$ be a Lebesgue number of this cover $S^d = \cup_{i=1}^{d+1} U_i$, so that every open ball of radius λ in S^d is contained in (at least) one of the sets U_i. Set

$$V_i = \{x \in S^d : \text{dist}(x, \overline{U_i}) \geq \lambda/2\},$$

where $\overline{U}_i = S^d \backslash U_i$ is the complement of U_i. Then V_i is a closed subset of U_i and $S^d = \cup_{i=1}^{d+1} V_i$. By Theorem 1(ii), some V_i contains a pair of antipodes $x, -x$ and so $\{x, -x\} \subset V_i \subset U_i$.
Case 3: $1 \leq k \leq d$. Suppose for a contradiction that neither a set V_i nor a set U_j contains

a pair of antipodes. It is easy to see that for every i, $1 \leq i \leq k$, there is an open set U_i containing V_i that still fails to contain a pair of antipodes. But then the open cover $\{U_1, \ldots, U_{d+1}\}$ contradicts Case 2.
To define U_i precisely, for $i = 1, \ldots, k$ and $x \in S^d$, set

$$f_i(x) = \max\{\text{dist}(x, V_i) + \text{dist}(-x, V_i)\}.$$

Then f_i is a continuous and strictly positive function on the entire compact set S^d, so $m_i = \min f_i > 0$. Set

$$U_i = \{x : \text{dist}(x, V_i) < m_i/2\};$$

then U_i is an open set that contains V_i and does not contain a pair of antipodes, as required. □

As shown by Greene, a proof of Lovasz' Theorem on Kneser's conjecture is within half a step of Theorem 2.

Theorem : 3

For $n = 2k + d \geq 2k$, the chromatic number of $KG(n, k)$ is $d + 2$.

Proof. Since we know that $\chi(KG(n, k)) \leq d + 2$, what we need is that if X is an n-set and $X^{(k)} = \cup_{i=1}^{d+1} \mathcal{A}_i$ then some $\mathcal{A}_i$ contains two disjoint k-sets. To show this, let us assume, as we may, that the elements of X are points on $S^{d+1} \subset \mathbb{R}^{d+2}$ in general position, i.e. no hyperplane (or great sphere S^d) contains $d+2$ of these points. Let U_i be the set of points of S^{d+1} whose open hemispheres contain a k-set of "color" i:

$$U_i = \{x \in S^{d+1} : \exists A \in \mathcal{A}_i \text{ such that } A \subset H(x)\},$$

where $H(x)$ is the open hemisphere of S^{d+1} with center x. Finally, set

$$V_{d+2} = S^{d+1} \setminus (U_1 \cup U_2 \cup \cdots \cup U_{d+1}).$$

Then $U_1, \ldots, U_{d+1}$ are open sets, V_{d+2} is closed, and altogether they cover the entire sphere S^{d+1}. Consequently, by Theorem 2, one of these sets contains an antipodal pair $(x, -x)$.
Note that $\{x, -x\} \subset V_{d+2}$ cannot hold since if it did, both $H(x)$ and $H(-x)$ would contain at most $k - 1$ points of X, so the great-sphere of these hemispheres would contain at least $n - 2(k - 1) = d + 2$ points of X, containing that the points of X are in general position. Hence $\{x, -x\} \subset U_i$ for some i, $1 \leq i \leq d + 1$. This means that each of the two disjoint hemispheres $H(x)$ and $H(-x)$ contains a k-subset that belongs to $\mathcal{A}_i$: these k-subsets are disjoint. □

Applications of the LSB Theorem

As I have no time to cover this section in my lectures, I thought you'd enjoy some exercises about this topic.

We say that a Borel measure μ on $[0,1]$ is **continuous** or **non-atomic** if every point has zero measure: equivalently, if $\mu(0,t)$ is a continuous function of t.

Exercise 1. Let $\mu_1, \dots, \mu_n$ be non-atomic signed Borel measures on $[0,1]$ with $\mu_k([0,1]) = 1$ for every k. Then there are $a_0 = 0 \le a_1 \le \dots \le a_n \le a_{n+1} = 1$ such that, for some $I \subset [n+1] = \{1, 2, \dots, n+1\}$ and $A = \cup_{i \in I}[a_{i-1}, a_1]$, we have $\mu_k(A) = \dfrac{1}{2}$ for every k.

Strong Hint. For $x = (x_j)_1^{n+1} \in S^n$, i.e. $\sum_1^{n+1} x_j^2 = 1$, set $y = y(x) = (y_i)_0^{n+1} \in \mathbb{R}^{n+2}$, where $y_i \sum_1^i x_j^2$ for $1 \le i \le n+1$. In particular, $y_0 = 0$ and $y_{n+1} = 1$. Define a function $f : S^n \to \mathbb{R}^n$ by setting $f(x) = (\xi_k)_1^n$, where

$$\xi_k = \sum_{i=1}^{n+1} (\text{sign } x_i)\mu_k([y_{i-1}, y_i]).$$

Let us turn to another colorful result, the Ham Sandwich Theorem. Like the Borsuk part of LSB, the origin of this also goes back to the Scottish Cafe in Lwow (or Lemberg): inspired by Borsuk's Theorem, Hugo Steinhaus asked whether it was always possible to dissect three solids located anywhere in $\mathbb{R}^3$ by an appropriate plane.

Exercise 2. Let $\mu_1, \dots, \mu_d$ be Borel probability measures on $\mathbb{R}^d$ (i.e. $\mu_i(\mathbb{R}^d) = 1$) such that every affine hyperplane has measure 0 in each of the measures. Then there is a half-space of measure $1/2$ in each of the measures.

Strong Hint. Let us define a function $f : S^d \to \mathbb{R}^d$, mapping a point $x \in S^d$, $x = (x_0, x_1, \dots, x_d)$ into $f(x) = (f_1(x), \dots, f_d(x))$ by setting

$$H(x) = \{y \in \mathbb{R}^d : x_1y_1 + x_2y_2 + \dots + x_dy_d \le x_0\}$$

and then

$$f_i(x) = \mu_i(H(x)).$$

Note that if $x_0 = 1$ then $H(x) = \mathbb{R}^d$, so $f_i(x) = 1$ for every i, and if $x_0 = -1$ then $H(x) = \varnothing$, so $f_i(x) = 0$ for every i; in every other case, $H(x)$ is a closed half-space, and $f_i(x) + f_i(-x) = 1$. It is easy to show (see Exercise 3) that f is a continuous map. Consequently, by Theorem 1(i), $f(x) = f(-x)$ for some $x \in S^d$, so $f_i(x) = \frac{1}{2}$ for every i. Hence the half-space $H(x)$ has the desired properties.

Let us note some easy combinatorial consequences of these two exercises.

Exercise 3. Every open necklace with k types of bead, an even number of each kind, can be divided between two thieves with at most k cuts such that each thief gets half of the beads of each kind.

Exercise 4. Let $A_1, \dots, A_d$ be finite sets of points in $\mathbb{R}^d$, with $|A_i| = 2a_i + 1$ odds for each i. Then there is an affine hyperplane such that each of the two open half-spaces it defines contains at most a_i points of A_i for each i.

Exercise 5. Let $A_1, \dots, A_d$ be finite sets of points in $\mathbb{R}^d$, with $\cup_{i=1}^d A_i$ in general position. Then there is an affine hyperplane such that each of the two open half-spaces it defines contains precisely $\lfloor \frac{|A_i|}{2} \rfloor$ points of A_i for each i.