Combinatorics

December 22, 2008

The feel of this course is similar to graph theory; if you liked that you will like this. The course is about subsets of finite sets, which might seem the thing least likely to have interesting things to say about them, but in fact we'll find a lot of interesting results. The definitions here are simple; most of the ideas are in the proofs themselves, so you cannot ignore the proofs as in other courses. There are many nice, short proofs, even of very strong results - sometimes results so strong that they seem like they couldn't possibly be true.

An outline: there three chapters. The main one is on set systems; then we have one on isoperimetric inequalities, i.e. minimising the perimiter of a shape with given area, but working with finite rather than continuous sets. We are mostly interested in the behaviour as the number n of dimensions gets large, rather than finding the precisely best solution when $n = 2$ or 3 (though we will do that as well). Finally we have a chapter on intersecting families of sets, where we can obtain some surprisingly strong results with only basic algebra.

As usual, the content of the course is what is on the board rather than anything else. There will be three examples sheets and three examples classes for these; it is important to prepare work rather than just attending to listen to these.

Books should be unnecessary for the course, but some of the audience may want one for reference or to give a second perspective on things. The recommended one is Bollboas' "Combinatorics"; this is very gentle, almost "bedside reading"; it gives a good feeling of what's going on, and shows some things that go beyond this course. Anderson's "Combinatorics of Finite Sets" is simple and clear, but less good in distinguishing between proofs that really are trivial and proofs which merely appear trivial because they've used a very clever idea. Both these books only really cover the first chapter; there are no good published books for the second and third chapters.

The only prerequisites are the basic concepts of graph theory (the notions of a graph and a path, and Hall's theorem), and knowing what the integers mod p are, and what a vector space is.

1 Set Systems

Let X be a set. A set system on X (or family of subsets of X) is a family $\mathcal{A} \subset \mathbb{P}(X)$, e.g. $X^{(r)} = \{A \subset X : |A| = r\}$. Unless otherwise stated we shall take $X = [n] = \{1, \ldots, n\}$. So e.g. $|X^{(r)}| = \binom{n}{r}$; so e.g. $[4]^{(2)} = \{12, 13, 14, 23, 24, 34\}$ (where of course 12 denotes $\{1, 2\}$ and so on). Often we make $\mathcal{P}(X)$ into a graph by joining A to B if $|A \triangle B| = 1$ (where $\triangle$ is the symmetric difference). This graph is called the discrete cube Q_n. There are two important ways to visualize this: firstly, as a set of "levels" $X^{(r)}$, starting with the one-element $X^{(n)}$, then the larger $X^{(n-1)}$, and so on down to the maximally-sized $X^{(\frac{n}{2})}$ for n even (or equally sized $X^{(\lceil\frac{n}{2}\rceil)}$ and $X^{(\lfloor\frac{n}{2}\rfloor)}$ for n odd), then reducing in size to $X^{(0)}$. This is a useful visualization in many ways, but doesn't show the graph interconnections well - there's simply an awful mess of edges between any two adjacent levels. The other visualization, which gives the graph its name, is as follows: identify a point $A \in \mathbb{P}(X)$ with a 0-1 sequence of length n, e.g. $\{1,3\} \to 10100\ldots$. There is then an obvious identification with the vertices of the unit cube in $\mathbb{R}^n$.

Chains and antichains

$\mathcal{A} \subset \mathbb{P}(X)$ is a chain if $\forall A, B \in \mathcal{A}, A \subset B$ or $B \subset A$, e.g. $\{12, 1257, 12357\}$. It is an antichain if $\forall A \neq B \in \mathcal{A}, A \not\subset B$, e.g. $\{1, 346, 2489\}$. How large can a chain be? We can have $|\mathcal{A}| = n+1$ by e.g. $\mathcal{A} = \{\emptyset, 1, 12, \ldots, [n]\}$; clearly we cannot beat this , as a chain can meet a "level" $X^{(r)}$ in at most one point.

How large can an antichain be? We can have $|\mathcal{A}| = n$, e.g. $A = \{1, 2, \ldots, n\}$; in fact $X^{(r)}$ for any r is always an antichain. So we can achieve $\binom{n}{\frac{n}{2}}$ for n even or $\binom{n}{\lfloor\frac{n}{2}\rfloor}$ for n odd; in fact this is optimal, but how do we prove this?

1.1 Theorem (Sperner's Lemma)

(Note that there are two important results called "Sperner's Lemma", the other being in algebraic topology)

Let $\mathcal{A} \subset \mathbb{P}(X)$ be an antichain. Then $|\mathcal{A}| \leq \binom{n}{\lfloor\frac{n}{2}\rfloor}$: we'll decompose $\mathbb{P}(X)$ into $\binom{n}{\lfloor\frac{n}{2}\rfloor}$ chains, then we're done. (This is a very clever idea - it mirrors the chain case, where we proved the maximum size of a chain by decomposing the cube into $n+1$ antichains, the layers). To do this, it is sufficient to show that: i) for each $r < \frac{n}{2}$ there is a matching (recall this means a set of disjoint edges) from $X^{(r)}$ to $X^{(r+1)}$, and ii) for each $r > \frac{n}{2}$ there is a matching from $X^{(r)}$ to $X^{(r-1)}$; then we can just put together these matchings to form our chains. By taking complements, it is sufficient to prove only the first of these.

As an exercise the reader should try to find an explicit matching; this is actually very hard. However, we can and shall of course just use Hall: consider the subgraph G of Q spanned by $X^{(r)} \cup X^{(r+1)}$; it is bipartite. We have $d(A) = n - r \forall A \in X^{(r)}$ and $d(A) = r + 1 \forall A \in X^{(r+1)}$. Given $S \subset X^{(r)}$, the number of $S - \Gamma(S)$ edges is $|S|(n-r)$ (counting from below), but is also $\leq |\Gamma(S)|(r+1)$ (counting from above). Thus $|\Gamma(S)| \geq |S|\frac{n-r}{r+1} \geq |S|$ (as $r < \frac{n}{2}$), so by Hall there is a matching.

Remarks: 1. We can achieve $|\mathcal{A}| = \binom{n}{\lfloor\frac{n}{2}\rfloor}$, by e.g. $A = X^{\lfloor\frac{n}{2}\rfloor}$. 2. We have no result on uniqueness; this proof tells us nothing about when $|\mathcal{A}|$ may $= \binom{n}{\lfloor\frac{n}{2}\rfloor}$.

Aim: If $\mathcal{A}$ is an antichain then $\sum_{r=0}^{n} \frac{|A \cap X^{(r)}|}{\binom{n}{r}} \leq 1$, i.e. the sum of proportions of levels filled is ≤ 1. This trivially implies Sperner.

For $\mathcal{A} \subset X^{(r)}$ the shadow or lower shadow of $\mathcal{A}$ is $\partial\mathcal{A} = \partial^{-}\mathcal{A} = \{B \in X^{(r-1)} : B \cup i \in \mathcal{A} \text{ some } i \in X\} \subset X^{(r-1)}$. E.g. if $\mathcal{A} = \{123, 124, 234, 135\}$ we have $\partial\mathcal{A} = \{12, 13, 23, 14, 24, 34, 15, 35\}$.

1.2 Local Lym

("Lym" here is the initials of the mathematicians who first proved the result, Lubal, Meshachin and Amamoto)

Let $1 \leq r \leq n$ and $\mathcal{A} \in X^{(r)}$. Then $\frac{|\partial\mathcal{A}|}{\binom{n}{r}-1} \geq \frac{|\mathcal{A}|}{\binom{n}{r}}$, i.e. "the fraction occupied by $\partial\mathcal{A}$ is $\geq$ that for $\mathcal{A}$".

Proof: the number of edges from $\mathcal{A}$ to $\partial\mathcal{A}$ is $|\mathcal{A}|^r$ (counting from $\mathcal{A}$) and is $\leq |\partial\mathcal{A}|(n-r+1)$ (counting from $\partial\mathcal{A}$). Thus $|\mathcal{A}|r \leq |\partial\mathcal{A}|(n-r+1)$, so $\frac{|\partial\mathcal{A}|}{|\partial A|} \geq \frac{r}{n-r+1}$. But $\frac{\binom{n}{r}-1}{\binom{n}{r}} = \frac{r}{n-r+1}$, so we have the result (or, for the reader who finds this last numerical part ugly (which it is), we can argue from equality in the case where $\mathcal{A}$ and $\partial\mathcal{A}$ are the whole layer).

When can there be equality in locality? We must have that $\forall A \in \mathcal{A} \forall i \in A \forall j \notin A$, $(A-i) \cup j \in \mathcal{A}$; thus $\mathcal{A}$ must $= \emptyset$ or $X^{(r)}$.

1.3 Theorem (Lym inequality)

Let $\mathcal{A} \subset \mathbb{P}(X)$ be an antichain. Then $\sum_{r=0}^{n} \frac{|\mathcal{A} \cap X^{(r)}|}{\binom{n}{r}} \leq 1$.

Proof 1: "Bubble down using local lym". Write $\mathcal{A}_r$ for $\mathcal{A} \cap X^{(r)}$. Firstly, we have $\frac{|\mathcal{A}_n|}{\binom{n}{n}} \leq 1$. Now, $\partial\mathcal{A}_n$ and $\mathcal{A}_{n-1}$ are disjoint subsets of $X^{(n-1)}$ (since $\mathcal{A}$ is an antichain), so $\frac{|\partial\mathcal{A}_n|}{\binom{n}{n}-1} + \frac{|\mathcal{A}_{n-1}|}{\binom{n}{n}-1} = \frac{|\partial\mathcal{A}_n \cup \mathcal{A}_{n-1}}{\binom{n}{n}-1} \leq 1$, so (by local lym) $\frac{|\mathcal{A}_n|}{\binom{n}{n}} + \frac{|\mathcal{A}_{n-1}}{\binom{n}{n}-1} \leq 1$. Also, $\partial(\partial\mathcal{A}_n \cup \mathcal{A}_{n-1})$ and $\mathcal{A}_{n-2}$ are disjoint, so $\frac{|\partial(\partial\mathcal{A}_n \cup \mathcal{A}_{n-1}|}{\binom{n}{n}-2} + \frac{|\mathcal{A}_{n-2}|}{\binom{n}{n}-2} \leq 1$; thus (by local lym again) $\frac{|\mathcal{A}_n|}{\binom{n}{n}} + \frac{|\mathcal{A}_{n-1}|}{\binom{n}{n}-1} + \frac{|\mathcal{A}_{n-2}|}{\binom{n}{n}-2} \leq 1$. Keep going; we get $\frac{|\mathcal{A}_n|}{\binom{n}{n}} + \cdots + \frac{|\mathcal{A}_0|}{\binom{n}{0}} \leq 1$.

When can we have equality in lym? We must have had equality in each application of local lym; thus, for the greatest r with $\mathcal{A}_r \neq \emptyset$, we must have $\mathcal{A}_r = X^{(r)}$, but then $\mathcal{A} = X^{(r)}$ since $\mathcal{A}$ is an antichain. So we have equality in lym iff $\mathcal{A} = X^{(r)}$, and so in particularly equality in Sperner iff $\mathcal{A} = X^{\frac{n}{2}}$ (for n even) or $\mathcal{A} = X^{\lfloor \frac{n}{2} \rfloor}$ or $\mathcal{A} = X^{\lceil \frac{n}{2} \rceil}$ (for n odd).

Proof 2: Choose, uniformly at random, a maximal chain C (i.e. $A_0 \subset A_1 \subset \cdots \subset A_n$, with $|A_i| = i \forall i$. For a fixed r-set A, we have $P(A \in C) = \frac{1}{\binom{n}{r}}$, so $P(C \text{ meets } \mathcal{A}_r) = \frac{|\mathcal{A}_r|}{\binom{n}{r}}$ (since the events are disjoint). Thus $\sum_{r=0}^{n} \frac{|\mathcal{A}_r|}{\binom{n}{r}} \leq 1$.

Since we didn't use any actual probability here, only the language of probability, we can rewrite the proof using only counting: the number of maximal chains is $n!$ and the number of chains through a given r-set in $r!(n-r)!$, so $\sum_{r=0}^{n} |\mathcal{A}_r| r!(n-r)! \leq n!$ and we have the result. But this only serves to obscure things.

Shadows

If $\mathcal{A} \subset X^{(r)}$, we know $|\partial\mathcal{A}| \geq |\mathcal{A}|\frac{r}{n-r+1}$. But equality is very rare (it only occurs for $\mathcal{A} = \emptyset$ or $\mathcal{A} = X^{(r)}$. What happens in between? Specifically, given $|\mathcal{A}|$ fixed, how would we choose $\mathcal{A} \subset X^{(r)}$ to minimise $|\partial\mathcal{A}|$? It's believable that for $|\mathcal{A}| = \binom{k}{r}$, we should choose $\mathcal{A} = [k]^{(r)}$ yielding $\partial\mathcal{A} = [k]^{(r-1)}$.

What about for $\binom{k}{r} < |\mathcal{A}| < \binom{k+1}{r}$ for some k? It's similarly believable that we'd take $[k]^{(r)}$ and some extra sets from $[k+1]^{(r)}$, e.g. for $\binom{7}{3} + \binom{4}{2}$ 3-sets, we'd try $\mathcal{A} = [7]^{(3)} \cup \{A \cup \{8\} : A \in [4]^2\}$.

Two total orderings on $X^{(r)}$

Given $A, B \in X^{(r)}$, write these as $A = \{a_1 \dots a_r\}, B = \{b_1 \dots b_r\}$ with $a_1 < \dots < a_r, b_1 < \dots < b_r$. We say that $A < B$ in the lexicographic or lex order if $\exists i$ with $a_i < b_i$ and $a_j = b_j \forall j < i$; equivalently $A < B$ if $a_i < b_i$ where $i = \min\{j : a_j \neq b_j\}$. Informally, for a set to be small in this ordering we "use small elements if possible". For example, the lexicographic ordering on $[4]^{(2)}$ gives 12, 13, 14, 23, 24, 34; on $[6]^{(3)}$ it is 123, 124, 125, 126, 134, 135, 136, 145, 146, 156, 234, 235, 236, 245, 246, 256, 345, 346, 456.

We say $A < B$ in the colexicographic or colex order if $\exists i$ with $a_i < b_i$ and $a_j = b_j \forall j > i$; equivalently, $A < B$ if $a_i < b_i$ where $i = \max\{j : a_j \neq b_j\}$; informally, "don't use large elements". Finally, equivalently $A < B$ if $\sum_{i<A} 2^i < \sum_{i \in B} 2^i$, i.e. if we write our sets as binary strings then this is just the usual order on such. E.g. the colexicographic ordering on $[4]^{(2)}$ is 12, 13, 23, 14, 24, 34; on $[6]^{(3)}$ it is 123, 124, 134, 234, 125, 135, 235, 145, 245, 345, 126, 136, 236, 146, 246, 346, 156, 256, 356, 456. Note that $[m]^{(r)}$ is an initial segment of $[m+1]^{(r)}$; thus we could view the colexicographic order as an enumeration of $\mathbb{N}^{(r)}$, which is not the case for the lexicographic order.

Our aim is to show that initial segments of colex have the smallest lower shadow, i.e. if $\mathcal{A} \subset X^{(r)}$ and $C \subset X^{(r)}$ is the initial segment of colex with $|C| = |\mathcal{A}|$, then $|\partial C| \leq |\partial \mathcal{A}|$. In particular, this gives us that $|\mathcal{A}| = \binom{k}{r} \Rightarrow |\partial \mathcal{A}| \geq \binom{k}{r-1}$.

Compressions

Given $\mathcal{A} \subset X^{(r)}$, we'd like a way to replace $\mathcal{A}$ by some $\mathcal{A}' \subset X^{(r)}$ with i) $|\mathcal{A}'| = |\mathcal{A}|$ ii) $|\partial \mathcal{A}'| \leq |\partial \mathcal{A}|$ and iii) $\mathcal{A}'$ "looks more like" C than $\mathcal{A}$ did. (Of course, strictly speaking replacing $\mathcal{A}$ with C immediately is a suitable operation; we also want to be able to easily prove ii)). We'd like to find several such "compression" operations, so that for any possible $\mathcal{A}$ we have $\mathcal{A} \to \mathcal{A}' \to \mathcal{A}'' \dots \to \mathcal{B}$ such that either $\mathcal{B} = C$, or $\mathcal{B}$ is so similar to C that we can see directly that $|\partial \mathcal{B}| \geq |\partial C|$.

ij-compressions

"Colex prefers 1 to 2" inspires: fix $1 \leq i < j \leq n$. Then the ij-compression is to "replace j by i if possible": for $A \subset X^{(r)}$, let $C_{ij}(A) = A \cup i - j$ (the precedence of operations here is obvious; in general we always read from left to right) if $j \in A, i \notin A$ and A otherwise, and for $\mathcal{A} \subset X^{(r)}$ let $C_{ij}(\mathcal{A}) = \{C_{ij}(A) : A \in \mathcal{A}\} \cup \{A \in \mathcal{A} : C_{ij}(A) \in \mathcal{A}\}$ (i.e. we replace A by $C_{ij}(A)$ unless this is another set already in $\mathcal{A}$). So e.g. if $\mathcal{A} = \{123, 124, 135, 235, 245, 367\}$ then $C_{12}(\mathcal{A}) = \{123, 124, 135, 235, 145, 367\}$. We say $\mathcal{A}$ is ij-compressed if $C_{ij}(\mathcal{A}) = \mathcal{A}$.

We clearly have $|C_{ij}(\mathcal{A}) = |\mathcal{A}|$.

1.4 Lemma

Let $1 \leq r \leq n, \mathcal{A} \subset X^{(r)}$. Then for any $1 \leq i < j \leq n$, $|\partial C_{ij}(\mathcal{A}) \leq |\partial \mathcal{A}|$: write $\mathcal{A}'$ for $C_{ij}(\mathcal{A})$. We'll show that for each $B \in \partial \mathcal{A}' - \partial \mathcal{A}$ we have $j \notin B, i \in B$, and $B \cup j - i \in \partial \mathcal{A} - \partial \mathcal{A}'$, then we're clearly done.

We have $B \cup x \in \mathcal{A}'$ for some x (and $B \cup x \notin \mathcal{A}$). So $i \in B \cup x, j \notin B \cup x$. We cannot have $i = x$ as then $B \cup i \in \mathcal{A}'$ so $B \cup i$ or $B \cup j \in \mathcal{A}$, so $B \in \partial\mathcal{A}$, a contradiction. So $i \in B, j \notin B$. Since $B \cup x \cup j - i \in \mathcal{A}$, we certainly have $B \cup j - i \in \partial\mathcal{A}$. We claim $B \cup j - i \notin \partial\mathcal{A}'$: suppose $B \cup j - i \cup y \in \mathcal{A}'$ for some y. Then we cannot have $y = i$ (as then $B \cup j \in \mathcal{A}'$, whence $B \in \partial\mathcal{A}$). So $i \notin B \cup y \cup j - i, j \in B \cup y \cup j - i$, and so $B \cup y \cup j - i$ and $B \cup y$ are both $\in \mathcal{A}$ (by definition of $\mathcal{A}'$). So $B \in \partial\mathcal{A}$, a contradiction.

The above proof is not one to memorize line by line; on the contrary, if you understand what's going on, you should be easily able to reproduce it from scratch.

Remark: We actually showed that $\partial(C_{ij}(\mathcal{A})) \subset C_{ij}(\partial\mathcal{A})$.

We say $\mathcal{A} \subset X^{(r)}$ is left-compressed if $C_{ij}(\mathcal{A}) = \mathcal{A} \forall i < j$.

1.5 Corollary

For $\mathcal{A} \subset X^{(r)} \exists \mathcal{B} \subset X^{(r)}$ with $|\mathcal{B}| = |\mathcal{A}|, |\partial\mathcal{B}| \leq |\partial\mathcal{A}|$ and $\mathcal{B}$ left-compressed: define a sequence $\mathcal{A}_0, \mathcal{A}_1, \dots \subset X^{(r)}$ as follows: set $\mathcal{A}_0 = \mathcal{A}$. Having chosen $\mathcal{A}_0, \dots, \mathcal{A}_k$: if $\mathcal{A}_k$ is left-compressed, stop the sequence with it. Otherwise, choose $i < j$ with $\mathcal{A}_k$ not ij-compressed, and set $\mathcal{A}_{k+1} = C_{ij}(\mathcal{A}_k)$. This must terminate (as e.g. $\sum_{A \in \mathcal{A}_k} \sum_{x \in A} x$ is decreasing in k); the system $\mathcal{B} = \mathcal{A}_k$ has $|\mathcal{B}| = |\mathcal{A}|$ and $|\partial\mathcal{B}| \leq |\partial\mathcal{A}|$ by Lemma 4.

Remarks: 1. Alternatively: among all $\mathcal{B} \subset X^{(r)}$ with $|\mathcal{B}| = |\mathcal{A}|$ and $|\partial\mathcal{B}| \leq |\partial\mathcal{A}|$, choose one with minimal $\sum_{A \in \mathcal{A}_k} \sum_{x \in A} x$. 2. It is possible to apply each C_{ij} at most once, if we choose the order sensibly. But this is not really relevant.

Any initial segment of colex is left-compressed. But the converse is easily false, e.g. $\mathcal{A} = \{123, 124, 125, 126, 127\}$. It is possible to prove the result we want using only left-compression and some clever counting, but this is ugly and painful.

"Colex prefers 23 to 14" inspires: for $U, V \subset X$ with $|U| = |V|$ and $U \cap V = \emptyset$, define the UV-compression by: for $A \in X^{(r)}$, let $C_{UV}(A) = A \cup U - V$ if $V \subset A, A \cap U = \emptyset$, A otherwise, and for $\mathcal{A} \subset X^{(r)}$, set $C_{UV}(\mathcal{A}) = \{C_{UV}(A) : A \in \mathcal{A}\} \cup \{A \in \mathcal{A} : C_{UV}(A) \in \mathcal{A}\}$. Note also that $C_{\{i\},\{j\}}(\mathcal{A}) = C_{ij}(\mathcal{A})$.

Unfortunately, we can have $|\partial C_{UV}(\mathcal{A})| > |\partial\mathcal{A}|$, e.g. let $\mathcal{A} = \{147, 478\}$; $|\partial A| = 5$. But then $C_{23,14}(\mathcal{A}) = \{237, 478\}$ with $|\partial C_{23,14}\mathcal{A}| = 6$.

Say $\mathcal{A}$ is UV-compressed if $C_{UV}(\mathcal{A}) = \mathcal{A}$.

1.6 Lemma

Let $U, V \subset X$ be disjoint with $|U| = |V|$. Let $\mathcal{A} \subset X^{(r)}$. Suppose that $(\star)\forall u \in U \exists v \in V$ such that $\mathcal{A}$ is $(U - u, V - v)$-compressed; this condition seems artificial at this stage, but comes out of the proof. Then $|\partial C_{UV}(\mathcal{A})| \leq |\partial\mathcal{A}|$. Proof: write $\mathcal{A}'$ for $C_{UV}(\mathcal{A})$. For any $B \in \partial\mathcal{A}' - \partial\mathcal{A}$, we'll show that $U \subset B, V \cap B = \emptyset$ and $B \cup V - U \in \partial\mathcal{A} - \partial\mathcal{A}'$, then done. We have $B \cup x \in \mathcal{A}'$ for some x (and $B \cup x \notin \mathcal{A}$). Thus $U \subset B \cup x, V \cap (B \cup x) = \emptyset$, and $B \cup x \cup V - U \in \mathcal{A}$, so certainly $V \cap B = \emptyset$. Also $x \notin U$, because: if $x \in U$, then $\exists y \in V$ with $\mathcal{A}$ being $(U - x, V - y)$-compressed, but $B \cup x \cup V - U \in \mathcal{A}$, so $B \cup y \in \mathcal{A}$, contradicting $B \notin \partial\mathcal{A}$. Thus $U \subset B$. We have $B \cup V - U \in \partial\mathcal{A}$, because $B \cup x \cup V - U \in \mathcal{A}$. Suppose $B \cup V - U \in \partial\mathcal{A}'$: so $w \cup (B \cup V - U) \in \mathcal{A}^{|}prime$. If $w \notin U$, because $w \cup (B \cup V - U) \in \mathcal{A}'$, we must have $w \cup (B \cup V - U)$ and $W \cup B$ both in $\mathcal{A}$ (by the definitions of C_{UV}), but then

$B \in \partial\mathcal{A}$, a contradiction. If $w \in U$, we have $\mathcal{A}$ $(U - w, V - z)$-compressed for some $z \in V$. So from $w \cup (B \cup V - U) \in \mathcal{A}$ (true as this is a set which contains V so cannot have been moved from another set into $\mathcal{A}'$) we obtain $B \cup z \in \mathcal{A}$, a contradiction.

Remark: We actually showed $\partial C_{UV}(\mathcal{A}) \subset C_{UV}(\partial\mathcal{A})$.

1.7 Theorem (Kruskal-Katona Theorem)

Let $\mathcal{A} \subset X^{(r)}$ $(1 \le r \le n)$ and let C be the initial segment of colex on $X^{(r)}$ with $|C| = |\mathcal{A}|$. Then $|\partial\mathcal{A}| \ge |\partial C|$; in particular if $|\mathcal{A}| = \binom{k}{r}$ then $|\partial\mathcal{A}| \ge \binom{k}{r-1}$. Let $\Gamma = \{(U, V) : U, V \subset X, |U| = |V| > 0, \max V > \max U\}$; this last condition lets us "head towards" colex. Define a sequence $\mathcal{A}_0, \mathcal{A}_1, \dots$ by: set $\mathcal{A}_0 = \mathcal{A}$. Having chosen $\mathcal{A}_0, \dots, \mathcal{A}_k$, if $\mathcal{A}_k$ is UV-compressed for each $(U, V) \in \Gamma$ then stop the sequence with $\mathcal{A}_k$. Otherwise, choose $(U, V) \in \Gamma$ with $|U|$ minimal such that $\mathcal{A}_k$ is not (U, V)-compressed, and set $\mathcal{A}_{k+1} = C_{UV}(\mathcal{A}_k)$. For each $u \in U$, setting $v = \min V$ we have $(U - u, V - v) \in \Gamma \cup \{\emptyset, \emptyset\}$, so $\mathcal{A}_k$ is $(U - u, V - v)$-compressed. Thus by Lemma 6, $|\partial\mathcal{A}_{k+1}| \le |\partial\mathcal{A}_k|$. This sequence must terminate, as $\sum_{A \in \mathcal{A}_k} \sum_{i \in A} 2^i$ is decreasing. The final term $\mathcal{B} = \mathcal{A}_k$ satisfies: $|\mathcal{B}| = |\mathcal{A}|$ and $|\partial\mathcal{B}| \le |\partial\mathcal{A}|$, and $\mathcal{B}$ is (U, V)-compressed $\forall (U, V) \in \Gamma$. We claim $\mathcal{B} = C$: suppose not, i.e. suppose we have $A, B \in X^{(r)}$ with $A < B$ in colex and $A \notin \mathcal{B}, B \in \mathcal{B}$. Set $U = A - B$, $V = B - A$. We have $\max V > \max U$ (as $A < B$ in colex), so $(U, V) \in \Gamma$, so $B \in \mathcal{B} \Rightarrow A \in \mathcal{B}$, a contradiction.

It should not be surprising that this claim should hold, since we invented (U, V)-compressions for use with colex; they are useless for anything else.

Remarks: 1) Equivalently, if $|\mathcal{A}| = \binom{k_r}{r} + \binom{k_{r-1}}{r-1} + \cdots + \binom{k_s}{s}$ where $k_r > k_{r-1} > \cdots > k_s$ and $s > 0$, then $|\partial\mathcal{A}| \ge \binom{k_r}{r-1} + \binom{k_{r-1}}{r-2} + \cdots + \binom{k_s}{s-1}$. 2) The proof only actually used Lemma 6, not Lemma 4 or Corollary 5. But Lemma 4 is very useful in understanding Lemma 6. 3) When do we have equality? The reader can check that if $|\mathcal{A}| = \binom{k}{r}$ and there is equality in Kruskal-Katona (i.e. $|\partial\mathcal{A}| = \binom{k}{r-1}$) then $\mathcal{A} = Y^{(k)}$ for some $Y \subset X$ with $|Y| = k$. But it is <u>false</u> in general that if equality holds ($|\partial\mathcal{A}| = |\partial C|$) then $\mathcal{A}$ is isomorphic to C (where $\mathcal{A}, \mathcal{B}$ are <u>isomorphic</u> if there is a permutation $f : X \to X$ sending $\mathcal{A}$ to $\mathcal{B}$).

For $\mathcal{A} \subset X^{(r)}$ $(0 \le r \le n-1)$, the <u>upper shadow</u> of $\mathcal{A}$ is $\partial^+\mathcal{A} \subset X^{(r+1)}$ given by $\partial^+\mathcal{A} = \{A \cup x : A \in \mathcal{A}, x \in X, x \notin A\}$. Now, $A < B$ in colex on $X^{(r)}$ if and only if $A^c < B^c$ in lex on $X^{(n-r)}$ (with the ground-set order reversed - a condition which is irrelevant up to isomorphism). So:

1.8 Corollary

Let $\mathcal{A} \subset X^{(r)}$ $(0 \le r \le n-1)$ and let C be the initial segment of lex on $X^{(r)}$ with $|C| = |\mathcal{A}|$. Then $|\partial^+ C| \le |\partial^+\mathcal{A}|$.

Since the shadow of an initial segment of colex on $X^{(r)}$ is an initial segment of colex on $X^{(r-1)}$ (if $C = \{A \in X^{(r)} : A \le a_1 \dots a_r\}$ then ∂C is precisely $\{B \subset X^{(r-1)} : X \le a_2 a_3 \dots a_r\}$), we have:

1.9 Corollary

Let $\mathcal{A} \subset X^{(r)}$ and let C be the initial segment of colex on $X^{(r)}$ with $|C| = |\mathcal{A}|$. Then $|\partial^t \mathcal{A}| \geq |\partial^t C| \forall 1 \leq t \leq r$; in particular if $|\mathcal{A}| = \binom{k}{r}$ then $|\partial^t \mathcal{A}| \geq \binom{k}{r-t}$: if $|\partial^t \mathcal{A}| \geq |\partial^t C|$ then $\partial^{r+1} \mathcal{A}| \geq |\partial^{t+1} C|$ by KK.

Intersecting Families

We say $\mathcal{A} \subset \mathbb{P}(X)$ is intersecting if $A \cap B \neq \emptyset \forall A, B \in \mathcal{A}$ (we shan't even bother defining disjoint families, since they are entirely uninteresting). How large can an intersecting family be? We can have $|\mathcal{A}| = 2^{n-1}$ by e.g. $\mathcal{A} = \{A \subset X : 1 \in A\}$.

1.10 Proposition

Let $\mathcal{A} \subset \mathbb{P}(X)$ be intersecting, then $|\mathcal{A}| \leq 2^{n-1}$: for any $A \subset \mathbb{P}(X)$, at most one of A, A^c can belong to $\mathcal{A}$.

Remark: there are many extremal systems, e.g. $\{A \in \mathbb{P}(X) : |A| > \frac{n}{2}\}$ works for odd n.

What if we restrict to $\mathcal{A} \subset X^{(r)}$? If $r > \frac{n}{2}$ this is silly, we can take $\mathcal{A} = X^{(r)}$. If $r = \frac{n}{2}$ this is also a silly case: choosing one set from each complimentary pair A, A^c gives us $\frac{1}{2}\binom{n}{r}$ and this is optimal as we can never have both A and A^c. So we study the case $r < \frac{n}{2}$.

The obvious guess is to set $\mathcal{A} = \{A \in X^{(r)} : 1 \in A\}$; this has $|\mathcal{A}| = \binom{n-1}{r-1} = \frac{r}{n}\binom{n}{r}$. We could also try e.g. $\mathcal{B} = \{A \in X^{(r)} : |A \cap \{1, 2, 3\}| \geq 2\}$. Observe that e.g. in $[8]^{(3)}$, $|\mathcal{A}| = 21$ and $|\mathcal{B}| = 1 + 3 \times 5 = 16 < 21$ (there is one set $\{1, 2, 3\}$ and 15 sets meeting it in two elements each).

1.11 Theorem (Erdős-Ko-Rado)

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

For one way of proving this, "bubble down with KK". For any $A, B \in \mathcal{A}$, we make the key observation that $A \cap B \neq \emptyset$ is equivalent to $A \not\subset B^c$ - this allows us to use shadows etc. Write $\overline{\mathcal{A}} = \{B^c : B \in \mathcal{A}\} \subset X^{(n-r)}$. We have that $\mathcal{A}$ and $\partial^{n-2r}\overline{\mathcal{A}}$ are disjoint subsets of $X^{(r)}$. Suppose that $|\mathcal{A}| > \binom{n-1}{r-1}$, then $|\overline{\mathcal{A}}| > \binom{n-1}{r-1} = \binom{n-1}{n-r}$. So $|\partial^{n-2r}\overline{\mathcal{A}}| \geq \binom{n-1}{r}$ by Corollary 9; thus $|\mathcal{A}| + |\partial^{n-2r}\overline{\mathcal{A}}| > \binom{n-1}{r-1} + \binom{n-1}{r} = \binom{n}{r}$, a contradiction. Note that the way this adds up neatly is no happy coincidence; the numbers "had" to work out, since if $\mathcal{A} = \{A \in X^{(r)} : 1 \in A\}$ then we have equality and $\mathcal{A}, \partial^{n-2r}\overline{\mathcal{A}}$ partition $X^{(r)}$.

An alternative proof, which is shorter and from first principles but requires one idea "out of a hat": consider cyclic orderings of X, i.e. bijections $\mathbb{Z}_n \hookrightarrow X$. Ask: how many of the $A \in \mathcal{A}$ are intervals (i.e. blocks of r consecutive elements) in this ordering. The answer is at most r: suppose $\{c_1, c_2, \ldots, c_r\} \in \mathcal{A}$¿ Then for each $1 \leq i \leq r - 1$, at most one of $\{c_{i-r+1} \ldots c_i\}$ and $\{c_{i+1} \ldots c_{i+r}\}$ belongs to $\mathcal{A}$ (and no other sets to). Also, each r-set is (an interval) in precisely $nr!(n - r)!$ of the $n!$ possible cyclic orderings - it could be any of the n intervals around the circle, and there are $r!$ ways to order the elements within the interval and $(n - r)!$ ways to order those without. So $|\mathcal{A}| \leq \frac{r}{n}\binom{n}{r}$ as required.

Remarks: 1. Formally, we are double-counting the edges in the bipartite graph with vertex classes $\mathcal{A}$ and the set of cyclic orderings, where we join $A \in \mathcal{A}$

to a cyclic ordering C if A is an interval in C. 2. This method is called averaging or Katona's method.

Equality in Erdős-Ito-Rado

We want that if $\mathcal{A} \subset X^{(r)}$ is intersecting ($r < \frac{n}{2}$) and $|\mathcal{A}| = \binom{n-1}{r-1}$ then $\mathcal{A} = \{A \in X^{(r)} : i \in A\}$ for some i. From the second proof: for each cyclic ordering C we have r intervals in $\mathcal{A}$; these must be all the r intervals containing some point $x(C)$. Our task is to show that $x(C) = x(C') \forall C, C'$; sufficient to do this for C and C' differing by a single transposition of adjacent elements. Fix a cyclic ordering C, and wlong take $x(C)$ to be the point C_0. Thus $\{C_0, \ldots, C_{r-1}\}, \{C_{-r+1}, \ldots, C_0\} \in \mathcal{A}, \{C_1, \ldots, C_r\}, \{C_{-r}, \ldots, C_{-1}\} \notin \mathcal{A}$.

Let C' be obtained from C by swapping two adjacent elements $\neq C_0$, say C_i and C_{i+1}. Wlog take $i \geq \frac{n-1}{2}$ (otherwise we can just reflect the cyclic ordering). Then $\{C_0, \ldots, C_{r-1}\}$ is an interval of C' and $\{C_1, \ldots, C_r\}$ an interval of C' (unless $r = \frac{n-1}{2}$ and $i = \frac{n-1}{2}$, in which case $\{C_1 \ldots C_{r-1}, C_{r+1}\}$ is an interval of C'. But $\{C_0 \ldots C_{r-1}\} \in \mathcal{A}, \{C_1 \ldots C_r\} \notin \mathcal{A}$ (and $\{C_1 \ldots C_{r-1}, C_{r+1}\} \notin \mathcal{A}$ as it is disjoint from $\{C_{-r+1}, \ldots, C_0\} \in \mathcal{A}$). Thus $x(C') = C_0$.

2 Isoperimetric Inequalities

We ask: for a set of given size, how small can the boundary be? E.g. (some continuous examples, which we will not cover except in passing) among subsets of $\mathbb{R}^2$ of given area, the disc has the smallest perimiter. Amongst subsets of $\mathbb{R}^3$ of given volume, the solid sphere has the smallest surface area. Among subsets of S^2 of a given area, a "circular cap" has the smallest perimiter.

For a graph G and $A \subset V(G)$, the boundary of A is $b(A) = \{x \in G : x \notin A, xy \in E \text{ some } y \in A\}$. An isoperimetric inequality on G is an inequality of the form $|b(A)| \geq f(|A|) \forall A \subset V(G)$.

Based on the above continuous examples, often a good guess is $B(x, r) = \{y : d(x, y) \leq r\}$ (where d is the usual graph distance, length of shortest path). A sometimes useful trick is to minimise the neighbourhood of A, $N(A) = \{x : d(x, A) \leq 1\}$.

What happens in Q_n? For e.g. $|A| = 4$ in Q_3, the obvious things to try are $B(0, 1)$, a corner and its three neighbours, which has $|b(A)| = 3$, or one square face of the cube, which has $|b(A)| = 4$. (Note we are emphasising that we are seeing Q_n only as a graph now; thus sets of points in it are A rather than $\mathcal{A}$). We quess that for general Q_n, $B(\emptyset, r) = X^{(\leq r)} = X^{(0)} \cup \cdots \cup X^{(r)}$ are best. What if $|X^{(\leq r)}| < |A| < |X^{(\leq r+1)}|$? We guess we should take $X^{(\leq r)} \cup B$ for some $B \subset X^{(r+1)}$. Then $b(A) = (X^{(r+1)} \setminus B) \cup \partial^+ B$, so we'll take B an initial segment of lex (by KK). Which suggests:

The simplicial ordering on $\mathbb{P}(X)$ is defined by: $x < y$ if either $|x| < |y|$ or $|x| = |y|, x < y$ in lex.

Aim: initial segments of the simplicial order are best.

Let $A \subset Q_n$ and $1 \leq i \leq n$. The i-sections of A are the sets $A_+^{(i)}, A_-^{(i)} \subset \mathbb{P}(X - i)$ given by $A_-^{(i)} = \{x \in A : i \notin x\}, A_+^{(i)} = \{x - i : x \in A, i \in x\}$. The i-compression $C_i(A)$ of A is defined by giving its i-sections: $(C_i(A))_+^{(i)}$ is the initial segment of simplicial on $\mathbb{P}(x-i)$ of size $|A_+^{(i)}|$ and $(C_i(A))_-^{(i)}$ is the inital segment of simplicial on $\mathbb{P}(x-i)$ of size $|A_-^{(i)}|$. Note that $|C_i(A)| = |A|$; note also that $C_i(A)$ "looks more like" a Hamming ball than A did (A Hamming ball is a set $A \subset Q_n$ with $X^{(r)} \subset A \subset X^{(r+1)}$ for some r). We say A is i-compressed if $C_i(A) = A$.

2.1 Theorem (Harper's Theorem)

(This theorem obeys Gauss' law; it was first proven (correctly) by Katona).

Let $A \subset Q_n$ and let C be the initial segment of the simplicial order with $|C| = |A|$. Then $|N(A)| \geq |N(C)|$; in particular $|A| \geq \sum_{i=0}^{r} \binom{n}{i} \Rightarrow |N(A)| \geq \sum_{i=0}^{r+1} \binom{n}{i}$.

Remarks: 1. If we knew A was a Hamming ball, we would be done by KK. 2. Conversely, theorem 1 implies KK; given $B \subset X^{(r)}$, apply theorem 1 to $A = X^{(<r)} \cup B$.

Proof: we induct on n, the $n = 1$ case is done. Given $A \subset Q_n$ for $n > 1$ and $1 \leq i \leq n$, we claim: $N(C_i(A)) \leq N(A)$: write B for $C_i(A)$. Note $|N(A) = |N(A_-) \cup A_+| + |N(A_+) \cup A_-|$ (the two sets being $N(A)_-, N(A)_+$ respectively). And $|N(B)| = |N(B_-) \cup B_+| + |N(B_+) \cup B_-|$. Now $|B_+| = |A_+|$ and $|N(B_-)| \leq |N(A_-)|$ by the inductive hypothesis. But B_+ and $N(B_-)$ are both initial segments of the simplicial ordering (as the neighbourhood of such an initial segment is an initial segment), so they are nested (one is a subset of the other), and so certainly $|N(B_-) \cup B_+|$ ($= \max(|N(B_-)|, |B_+|)$) is $\leq |N(A_-) \cup A_+|$. (Note that this proof only works because our "winning" sets are "nice"; they are initial segments of a fixed ordering). Similarly we have $|N(B_+) \cup B_-| \leq |N(A_+) \cup A_-|$ and we have the claim.

Define $A_0, A_1, \ldots$ as follows: set $A_0 = A$. Having chosen $A_0, \ldots, A_k$, if A_k is i-compressed $\forall i$, then stop the sequence with A_k. If not, choose i with $C_i(A_k) \neq A_k$, and set $A_{k+1} = C_i(A_k)$; continue. This must terminate; using the "most stupid" reason possible, $\sum_{x \in A_k} f(x)$ is decreasing where $f(x)$ denotes the position of x in the simplicial ordering. Then $B = A_k$ satisfies: $|B| = |A|, |N(B)| = |N(A)|$, B is i-compressed $\forall i$.

Does B i-compressed $\forall i$ imply B is an initial segment of the simplicial order? (If so, then we are done, for $B = C$. The answer is no, e.g. the bottom face of Q_3. However, we have:

2.2 Lemma

Let $B \subset Q_n$ be i-compressed $\forall i$, but not an initial segment of the simplicial ordering. Then for n odd, $= 2k+1$, we have $B = X^{(\leq k)} - \{(k+2)(k+3)\ldots(2k+1)\} \cup \{123\ldots(k+1)\}$, and for n even, $= 2k$, we have $B = X^{(\leq k-1)} \cup \{x \in X^{(k)} : 1 \in x\} - \{1(k+2)(k+3)\ldots 2k\} \cup \{234\ldots(k+1)\}$ (then we are done, as in each of these cases we have $|N(B)| \geq |N(c)|$ (in fact it is generally much larger)). We have some $x < y$ with $x \notin B, y \in B$. For each i, we cannot have $i \in x, y$ (as B is i-compressed) and we cannot have $i \notin x, y$ (as again B is i-compressed). Thus $x = y^c$. So for each $x \notin B$ we have at most 1 later point $y \in B$ (namely x^c) and for each $y \in B$ we have at most 1 earlier point $x \notin B$ (nomely y^c). So $B = \{z : z \leq y\} - \{x\}$, where x is the predecessor of y and $x = y^c$. So for n odd we must have x the last $\frac{n-1}{2}$-set, as required, and for n even we must have x the last $\frac{n}{2}$-set containing 1, again as required.

Notes: 1. We can also prove Harper's Theorem by UV-compressions; the proof is quite nice, but harder. 2. We can also use these "codimension-1" compressions to prove KK; this gives a proof which is short, but bad for the soul as a first proof of KK to see.

For $A \subset Q_n$, the <u>t-neighbourhood</u> of A is $N^t(A)$; equivalently this is $\{x \in Q_n : d(x, A) \leq t\}$.

2.3 Corollary

Let $A \subset Q_n$ with $|A| \geq \sum_{i=0}^{r} \binom{n}{r}$. Then for $1 \leq t \leq n-r$, we have $|N^t(A)| \geq \sum_{i=0}^{r+t} \binom{n}{i}$, by Theorem 1 and induction.

To get a feel for the strength of Corollary 3, we'll need some numerical estimates on $\sum_{i=0}^{r} \binom{n}{i}$ and similar quantities.

2.4 Proposition

Let $0 < \epsilon < \frac{1}{4}$ (of course we are only interested in ϵ small; the uppper bound is just for convenience). Then $\sum_{i=0}^{\lfloor (\frac{1}{2}-\epsilon)n \rfloor} \binom{n}{i} \leq \frac{1}{\epsilon} e^{-\epsilon^2 \frac{n}{2}} 2^n$; note that (for ϵ fixed as $n \to \infty$) this is an <u>exponentially small</u> fraction of 2^n. Some readers may have seen this fact already in the context of binomial or normal distributions; one way to think about it is that our limit is "$\sim \epsilon \sqrt{n}$ standard deviations from the mean $\frac{n}{2}$". This is a crude estimate (more sophisticated ones exist) and our proof will be simple, but it suffices for our purposes. $\binom{n}{i-1} = \binom{n}{i} \frac{i}{n-i+1}$. So for $i \leq \lfloor (\frac{1}{2} - \epsilon) n \rfloor$, have $\frac{\binom{n}{i-1}}{\binom{n}{i}} = \frac{i}{n-i+1} \leq \frac{(\frac{1}{2}-\epsilon)n}{(\frac{1}{2}+\epsilon)n} = 1 - \frac{2\epsilon}{\frac{1}{2}+\epsilon} \leq 1 - 2\epsilon$. So $\sum_{i=0}^{\lfloor (\frac{1}{2}-\epsilon)n \rfloor} \binom{n}{i} \leq \frac{1}{2\epsilon} \binom{n}{\lfloor (\frac{1}{2}-\epsilon)n \rfloor}$ (by the sum of a geometric progression); similarly $\binom{n}{\lfloor (\frac{1}{2}-\epsilon)n \rfloor} \leq \binom{n}{\lfloor (\frac{1}{2}-\frac{\epsilon}{2})n \rfloor} (1 - 2\frac{\epsilon}{2})^{\frac{\epsilon n}{2} - 1}$ (the -1 in the exponent being because of the $\lfloor\rfloor$ rubbish), by the same argument with ϵ replaced by $\frac{\epsilon}{2}$. Thus $\sum_{0}^{\lfloor (\frac{1}{2}-\epsilon)n \rfloor} \binom{n}{i} \leq \frac{1}{2\epsilon} 2 e^{-\frac{\epsilon n^2}{2}} 2^n$ and we have the result.

2.5 Theorem

Let $A \subset Q_n$, $0 < \epsilon < \frac{1}{4}$. Then $\frac{|A|}{2^n} \geq \frac{1}{2} \Rightarrow \frac{|A_{(\epsilon n)}|}{2^n} \geq 1 - \frac{1}{\epsilon} e^{-\frac{\epsilon^2 n}{2}}$ - "$\frac{1}{2}$-sized nets have exponentially large neighbourhoods": we have $|A| \geq \sum_{0}^{\lceil \frac{n}{2} - 1 \rceil} \binom{n}{i}$, so by Harper we have $|A_{(\epsilon n)}| \geq \sum_{0}^{\lceil \frac{n}{2} + \epsilon n - 1 \rceil} \binom{n}{i}$ so $|A^c_{(\epsilon n)}| \leq \sum_{\lceil \frac{n}{2} + \epsilon n \rceil}^{n} \binom{n}{i} = \sum_{0}^{\lfloor \frac{n}{2} - \epsilon n \rfloor} \binom{n}{i} \leq \frac{1}{\epsilon} e^{-\epsilon^2 \frac{n}{2}}$.

Remark: the above is concerned with $\frac{1}{2}$-sized sets, but the same argument would show $\frac{|A|}{2^n} \geq \frac{1}{\epsilon} e^{-\epsilon \frac{n^2}{2}} \Rightarrow \frac{|A_{2\epsilon n}|}{2^n} \geq 1 - \frac{1}{\epsilon} e^{-\frac{\epsilon^2 n}{2}}$.

Concentration of Measure

We say $f : Q_n \to \mathbb{R}$ is <u>Lipschitz</u> if $|f(x) - f(y)| \leq 1 \forall x, y$ adjacent. A real number M is a <u>median</u> or <u>Levy mean</u> for f if $|\{x : f(x) \leq M\}| \geq 2^{n-1}$, $|\{x : f(x) \geq M\}| \geq 2^{n-1}$.

We are now ready to show that "every well-behaved function on Q_n is roughly constant nearly everywhere".

2.6 Theorem

Let f be a Lipschitz function on Q_n with median M. Then $\frac{|\{x:|f(x)-M|\leq\epsilon n\}|}{2^n} \geq 1-\frac{2}{\epsilon}e^{-\frac{\epsilon^2 n}{2}}$ (for $0<\epsilon<\frac{1}{4}$) (This is the "concentration of measure" phenomenon).

Let $A=\{x: f(x)\leq M\}$, then $\frac{|A|}{2^n}\geq\frac{1}{2}$ so $\frac{|A_{(\epsilon n)}|}{2^n}\geq 1-\frac{1}{\epsilon}e^{-\frac{\epsilon^2 n}{2}}$. But $\forall x\in A_{(\epsilon n)}$ we have $f(x)\leq M+\epsilon n$ (as f is Lipschitz) and so $\frac{|\{x:f(x)\leq M+\epsilon n\}|}{2^n}\geq 1-\frac{1}{\epsilon}e^{-\frac{\epsilon^2 n}{2}}$; similarly $\frac{|\{x:f(x)\geq M-\epsilon n\}|}{2^n}\geq 1-\frac{1}{\epsilon}e^{-\frac{\epsilon^2 n}{2}}$ and we have the result.

Let G be a graph of diameter D (where diameter is $\max d(x,y) : x,y\in G$). Define $\alpha(G,\epsilon)=\max\{1-\frac{|A_{(\epsilon D)}|}{|G|} : A\subset G, \frac{|A|}{|G|}\geq\frac{1}{2}\}$. So "$\alpha(G,\epsilon)$ small says: $\frac{1}{2}$-sized sets have big ϵD-neighbourhoods".

A sequence $G_1, G_2,\dots$ of graphs is a Levy family if $\alpha(G_n,\epsilon)\to 0$ as $n\to\infty$ for each $\epsilon>0$. So e.g. theorem 5 tells us that $(Q_n)_{n=1}^{\infty}$ is a Levy family.

So we again have concentration of measure (Lipschitz functions on G_n are almost constant nearly everywhere) for any Levy family.

It turns out that many natural families of graphs - almost anything we could write down, any "natural" family of graphs - are Levy families; some people say this is the "concertration of measure phenomenon". Noone really knows why this is so. For example, the permutation groups S_n (made into a graph by σ adjacent to τ if $\sigma^{-1}\tau$ is a trasposition) is a Levy family.

Similary, we can define $\alpha(S,\epsilon)$ for any metric measure space S (of finite diameter and finite measure), so we can again define Levy families, and it turns out that many natural families of metric spaces form Levy families e.g. the sphere S^n:

Two ingredients: 1) Isoperimetric inequality in $S^n : |A|=|C|\Rightarrow|A_{(\epsilon)}|\geq|C_{(\epsilon)}|$ where C is a circular cap. A sketch of the proof: we can use compressions, e.g. the analogue of ij-compressions are "two-point symmetrisation" where we "stamp on our set" - fix a direction x, then a general line in this direction hits the sphere in two places; if the "upper" (in the x direction) is in our set but the lower is not, replace the upper with the lower. Here we can do this "nicely" in infinitely many directions, unlike Q_n where there were only n directions, so this works; we can form a sequence and use compactness to say that an optimal set exists, and then check it's a circular cap. Or this can also be done by codimension-1 compressions.

2) Estimate: A $\frac{1}{2}$-sized circular cap has angle $\frac{\pi}{2}$, so its ϵ-neighbourhood is a circular cap of angle $\frac{\pi}{2}+\epsilon$. But $\int_{\epsilon}^{1}\cos^n t dt\to 0$ as $n\to\infty$ for any fixed ϵ.

We deduced concentration of measure from isoperimetric estimates. Conversely:

2.7 Proposition

Let G be a graph such that for every Lipschitz $f: G\to\mathbb{R}$ of median M we have $\frac{|\{x\in G:|f(x)-M|>t\}|}{|G|}\leq\alpha$ for some fixed t,α. Then $\frac{|A|}{|G|}\geq\frac{1}{2}\Rightarrow\frac{|A(t)|}{|G|}\geq 1-\alpha$: Let $f(x)=d(x,A)$. Then f is Lipschitz and has 0 as a median (as $|A|\geq\frac{1}{2}|G|$), so we have the result.

Edge-isoperimetric inequalities

For a graph G, $A \subset V(G)$ the edge-boundary of A is $\partial_e A = \partial A = \{xy \in E : x \in A, y \notin A\}$. An edge-isoperimetric inequality on G is an inequality of the form: for $A \subset G$, $|A| = m \Rightarrow |\partial A| \geq f(m)$. E.g. for $|A| = 4$ in Q_3, a vertex and its three neighbours has $|\partial A| = 6$, while a face of the cube has $|\partial A| = 4$. This suggests that subcubes are best.

The binary ordering on Q_n is given by: $x < y$ if $\max(x \Delta y) \in y$ - equivalently $\sum_{i \in X} 2^i < \sum_{i \in y} 2^i$, i.e. we "go up in subcubes". Our aim is to show that initial segments of binary minimise ∂. For $A \subset Q_n$ and $1 \leq i \leq n$, the i-binary-compression $B_i(A)$ is defined by giving its i-sections: $B_i(A)_+^{(i)}$ is the initial segment of $\mathbb{P}(x - i)$ of size $|A_+^{(i)}|$, $B(A)_-^{(i)}$ is the initial segment of binary on $\mathbb{P}(x - i)$ of size $|A_-^{(i)}|$. Clearly $|B_i(A)| = |A|$. We say A is binary-compressed if $B_i(A) = A$ $(\forall i)$.

2.8 Theorem (Edge-isoperimetric inequality in the cube)

(Sometimes called the "theorem of Harper, Lindsey, Bernstein and Hart")

Let $A \subset Q_n$ and let C be the initial segment of binary with $|C| = |A|$. Then $|\partial C| \leq |\partial A|$; in particular, $|A| = 2^k \Rightarrow |\partial A| \geq (n - k)2^k$.

We proove by induction on n; $n = 1$ case done. Given $A \subset Q_n$, $1 \leq i \leq n$, we claim $|\partial B_i(A)| \leq |\partial A|$: write B for $B_i(A)$. We have $|\partial A| = |\partial(A_-)| + |\partial(A_+)| + |A_+ \Delta A_-|$, $|\partial B| = |\partial B_-| + |\partial B_+| + |B_+ \Delta B_-|$ Now $|\partial(B_-)| \leq |\partial(A_-)|, |\partial(B_+)| \leq |\partial(A_+)|$ by the inductive hypothesis; also $|B_+ \Delta B_-| \leq |A_+ \Delta A_-|$ because $|B_+| = |A_+|, |B_-| = |A_-|$ and the sets B_+, B_- are nested (as each is an IS of binary), thus $|\partial B| \leq |\partial A|$ and we have the claim. Define $A_0, A_1, \ldots$ by: set $A_0 = A$. Having chosen $A_0 \ldots A_k$, if A_k is i-binary-compressed $\forall i$ then stop, if not choose i with $B_i(A_k) \neq A_k$ and set $A_{k+1} = B_i(A_k)$. This must terminate, e.g. because $\sum_{x \in A_k}$(position of x in binary) is decreasing.

The final set $B = A_k$ satisfies: $|B| = |A|, |\partial B| \leq |\partial A|$, B is i-binary-compressed $\forall i$. But B need not be an IS, e.g. our corner vertex and three neighbours has this.

2.9 Lemma

Let $B \subset Q_n$ be i-binary compressed $\forall i$, not an initial segment of binary. Then $B = \mathbb{P}(n - 1) \cup \{n\} - \{123 \ldots n - 1\}$ (then we are done, as certainly $|\partial B| \geq |\partial C|$ in this case). We have some $x < y$ with $x \notin B, y \in B$, then for each i we cannot have $i \in X, Y$ or $i \notin X, Y$, as B is i-binary-compressed, so $x = y^c$. So for each $x \notin B$ we have at most one $y > x$ with $y \in B$ (namely $y = x^c$) and for each $y \in B$ there is at most one $x < y$ with $x \notin B$, namely $x = y^c$. Thus $B = \{z : z \leq y\} - \{x\}$ where x is the predecessor of y and $x = y^c$, hence $y = [n]$ and the set must be as described.

Remark: it was vital in the above proof that extremal sets in dimension $n - 1$ were nested, i.e. given by the initial segments of some ordering.

For a graph G, the isoperimetric number of G is $i(G) = \min\{\frac{|\partial A|}{|A|} : A \subset G, \frac{|A|}{|G|} \leq \frac{1}{2}\}$ - represents "how small the average out-degree can be".

2.10 Corollary

$i(Q_n) = 1$: the set $A = \mathbb{P}(n-1)$ shows $i(Q_n) \leq 1$. Let C be any initial segment of binary of size $\leq 2^{n-1}$, then $C \subset \mathbb{P}(n-1)$ so certainly $|\partial C| \geq |C| \Rightarrow i(Q_n) \geq 1$.

We will now study the grit. This is somewhat of a more "natural" space; it more closely resembles $\mathbb{R}^n$ than does the cube.

Inequalities in the Grid

The grid is the graph on $[k]^n = \{(x_1, \ldots, x_n) : x_i \in \{1, \ldots, k\} \forall i\}$, in which $x = (x_1, \ldots, x_n)$ is joined to $y = (y_1, \ldots, y_n)$ if for some i we have $|x_i = y_i| = 1, x_j = y_j \forall j \neq i$ (i.e. x, y are 1 apart in the L_1-norm). Note that for $k = 2$ this is exactly the graph Q_n. Our main questions are: do theorem 1 and theorem 8 extend to the grid?

Vertex-isoperimetric inequality: the first obvious thing to try is a "triangle" in the bottom corner, of side d, which has $b(A) \simeq d \simeq \sqrt{2|A|}$; the other obvious thing is a "square" or cube in the corner of side d, which has $b(A) \sim 2d \sim 2\sqrt{|A|}$. So a "good guess" is that sets of the form $\{x : |x| \leq r\}$ are best, where $|x| = x_1 + \cdots + x_n$ (in fact the triangle should have been obviously the better of the two approaches, from the result in the Q_n special case).

What if $|\{x : |X| \leq r\}| < |A| < |\{x : |x| \leq r+1\}|$? We'd take A of the form $\{x : |x| \leq r\} \cup B$ for some $B \subset \{x : |x| = r+1\}$. We'd want to take the "bottom corner" of the "sloping face" of our triangle; thinking about it, this is the region of $|x| = r+1$ where x_1 is big.

Define the simplicial ordering on $[n]^r$ by: $x < y$ if $|x| < |y|$ or $|x| = |y|$ and $x_i > y_i$ where $i = \min\{j : x_j \neq y_j\}$. E.g. on $[3]^2$ this is $(1,1), (2,1), (1,2), (3,1), (2,2), (1,3), (3,2), (2,3), (3,3)$; on $[4]^3$ it is $(1,1,1), (2,1,1), (1,2,1), (1,1,2), (3,1,1), (2,2,1), (2,1,2), (1,3,1), (1,2,2), (1,1,3), (4,1,1), \ldots$. Note that this agrees with our previous definition of simplicial for $k = 2$.

Our aim is to show that initial segments of simplicial are best for vertex-iso. Let $A \subset [k]^n$. For $1 \leq i \leq n$, the i-sections of A are the sets $A_1, \ldots, A_k$ (or $A_1^{(i)}, \ldots, A_k^{(i)}$) in $[k]^{n-1}$ given by $A_t = \{x = (x_1 \ldots x_{n-1}) \in [k]^{n-1} : (x_1 \ldots x_{i-1} t x_i x_{i+1} \ldots x_{n-1}) \in A\}$. The i-compression of A is the set $C_i(A) \subset [k]^n$ defined by: $(C_i(A))_t$ is the IS of simplicial on $[k]^{n-1}$ of size $|A_t|$. We say A is i-compressed if $C_i(A) = A$.

2.11 Theorem ("Vertex-isoperimetric inequality in the grid")

Let $A \subset [k]^n$ and let C be the initial segment of simplicial with $|C| = |A|$. Then $|N(C)| \leq |N(A)|$. We induct on n; the $n = 1$ case is done (for any $A \subset [k]^1$, $A \neq \emptyset, [k]^1$ we have $|N(A)| \geq |A| + 1 = |N(C)|$). Given $A \subset [k]^n$, fix $1 \leq i \leq n$. We claim $|N(C_i(A))| \leq |N(A)|$: write B for $C_i(A)$. For any $1 \leq t \leq k$, we have $N(A)_t = N(A_t) \cup A_{t-1} \cup A_{t+1}$ (setting $A_0 = A_{k+1} = \emptyset$). So $|N(A)| = \sum_t |N(A_t) \cup A_{t-1} \cup A_{t+1}|$; the same expression is true for $N(B)$. But $|B_{t-1}| = |A_{t-1}|, AB_{t+1}| = |A_{t+1}|, |N(B_t)| \leq |N(A_t)|$ by the induction hypothesis, and the sets $N(B_t), B_{t-1}, B_{t+1}$ are nested (as each is an IS of simplicial on $[k]^{n-1}$). Thus $|N(A_t) \cup A_{t-1} \cup A_{t+1}| \geq |N(B_t) \cup B_{t-1} \cup B_{t+1}|$.

Among all $B \subset [k]^n$ with $|B| = |A|$ and $|N(B)| \leq |N(A)|$, choose one with $\sum_{x \in B}$(position of x in simplicial) minimal. Then B is i-compressed for all i (else $C_i(B)$ contradicts minimality of B). It remains to prove $|N(B)| \geq |N(C)|$. Note that e.g. for $n = 2$ we can have a set B that "looks very different from simplicial": if we draw the space as a square, the plot of any (non-strictly) decreasing function gives us a possible B. But we will see that this is not possible in higher dimensions.

Case 1: $n = 2$. B is i-compressed $\forall i$ iff B is a down-set, i.e. if $x \in B$, y has $y_i \leq x_i \forall i$ then $y \in B$. Let $s = \max\{|x| : x \in B\}, r = \min\{|x| : x \notin B\}$, then $r \leq s$ (or else $B = C$). If $r = s$, then $\{x : |x| < r\} \subset B \subset \{x : |x| \leq r\}$, so certainly $N(B) \geq N(C)$. If $r < s$: we cannot have $\{x : |X| = s\} \subset B$, since B is a down-set (and we have some $x \notin B$ with $|x| = r < s$). So $\exists x, x'$ with $|x| = |x'| = s, x \in B, x' \notin B$, and $x = x' \pm (e_1 - e_2)$ say. Similarly, we cannot have $\{y : |y| = r\} \subset B^c$, so $\exists y, y' : y \in B, y' \notin B, |y| = |y'| = r, y = y' \pm (e_1 - e_2)$. But now let $B' = B - \{x\} \cup \{y'\}$. Then $|N(B')| \leq |N(B)$ (we have lost at least one point from the neighbourhood and gained at most 1 point), contradicting the choice of B.

Case 2: $n \geq 3$. For $x \in B$ with $x_n \geq 2$, we have $x - e_n + e_i \in B \forall i$ with $x_i < k$¡ as B is j-compressed for any $j \neq n, i$ (this is where we need $n \geq 3$). Thus $N(B_t) \subset B_{t-1} \forall t = 2, \ldots, k$.

We had $N(B)_t = N(B_t) \cup B_{t+1} \cup B_{t-1}$, so actually we have $N(B)_t = B_{t-1}$. So $|N(B)| = |B_{k-1}| + |B_{k-2}| + \cdots + |B_1| + |N(B_1)| = |B| - |B_k| + |N(B_1)|$; similarly for C. Thus, to complete the proof, it suffices to show that $|B_k| \leq |C_k|$ and $|B_1| \geq |C_1|$, as then we will have $|N(B_1)| \geq |N(C_1)|$ since these are initial segments of simplicial.

For $|B_k| \leq |C_k|$, define a set $D \subset [k]^n$ by: $D_k = B_k, D_t = N(D_{t+1})$ for $t = k_1, k-2, \ldots, 1$. Then D is clearly an IS of simplicial, and $D \subset B$, so $|D| \leq |B| = |C|$¡ so $D \subset C$ (as each is an IS of simplicial). So $D_k \subset C_k$ and $|D_k| \leq |C_k$ as required.

For $|B_1| \geq |C_1$, define a set $E \subset [k]^n$ by $E_1 = B_1, E_t = \{x \in [k]^{n-1} : N(\{x\}) \subset E_{t-1}\}$ for $t = 2, 3, \ldots 2k$. Then E is an IS of simplicial and $E \supset B$ so $|E| \geq |B| = |C|$, so $E \supset C$; thus $E_1 > C_1$ and $|E_1| \geq |C_1|$ as required.

2.12 Corollary

Let $A \subset [k]^n$ with $|A| \geq |\{x : |x| \leq r\}|$, then $|A_{(t)}| \geq |\{x : |x| \leq r + t\}|$.

Remark: We can check from this that for any fixed k, the sequence $([k]^n)_{n=1}^{\infty}$ is a normal Levy family.

Edge-isoperimetric inequalities in the grid

Unsurprisingly this is our next topic. To minimise $|\partial A|$ for $|A|$ given, if we take a "square" in the bottom left corner [$n = 2$ for now] of side r this has size r^2, while a triangle of side r has size $\sim \frac{1}{2}r^2$, but both these have the same boundary $2r$, suggesting squares are best. But as the size of our set grows, when the square reaches side length $\frac{k}{2}$ it is equalled by a "column" - the left hand $\frac{1}{4}$ of the square has the same edge-perimeter, and for increasing size-of-set the column then wins. This continues until we have a column of width $\frac{3k}{4}$, whose perimiter equals that of the complement of a square of side $\frac{k}{2}$, and after that complements of gradually smaller squares are best.

Sadly, these extremal sets are not nested; there are "phase transitions" at size $\frac{k^2}{4}$ and $\frac{3k^2}{4}$.

For $n = 3$, [as the size of the set grows] our optimal sets are at first cubes $[a]^3$, then square columns $[a]^2 \times [k]$, then half-spaces $[a] \times [k]^2$, then the complements of square columns and finally the complements of cubes - so "things have got worse" in terms of the number of phase transitions.

We will aim to show: the best sets are $[a]^d \times [k]^{n-d}$ or complements of these, for some d depending on the size of the set. It's ok not to calculate d exactly, since given this result, for any given size we can then easily find the "best" set by exhaustively checking the possible values of d.

Observe that if $A = [a]^d \times [k]^{n-d}$ then $|\partial A| = da^{d-1}k^{n-d} = d|A|^{1-\frac{1}{d}}k^{\frac{n}{d}-1}$.

2.13 Theorem (Edge-isoperimetric inequality in the grid)

Let $A \subset [k]^n$ with $|A| \leq \frac{k^n}{2}$. Then $|\partial A| \geq \min\{d|A|^{1-\frac{1}{d}}k^{\frac{n}{d}-1} : d = 1, \ldots, n\}$. (By stating the result in this form, we don't have to "fiddle" with the cases where $|A|$ is not divisible by some power of k).

This is the hardest of all known isoperimetric inequalities; there are some fiddly bits to the proof, so this will only be a non-examinable sketch. Wlog take A a down-set in $[k]^n$ (otherwise "stamp" on A with 1-dimensional compressions). For $1 \leq i \leq n$, define $C_i(A)$ by giving its i-sections: $C_i(A)_t$ is extremal (i.e. "winning" - of the form $[a]^d \times [k]^{n-1-d}$ or the complement of such) in $[k]^{n-1}$ with $|C_i(A)_t| = |A_t|$ (We're already "cheating" here - the size of A_t may not divide neatly to give a set of that precise form. But this can be gotten around with some fiddling).

Write B for $C_i(A)$. Then $|\partial A| = \sum |\partial A_t| + |A_1| - |A_k|$ - the first term gives the horizontal edges out of A, and the other terms are the vertical edges since A is a down-set.

However, what is $|\partial B|$? It is $\sum |\partial B_t|+$ some unknown quantity - B is not a down-set, as the extremal sets in $[k]^{n-1}$ are not nested. We might hope that the reduction in the first term in going from A to B balances any possible increase in the other terms, but this is not so - if A is on the bottom layer a square slightly larger than a threshold value and on the next layer a square slightly smaller than this, B is on the bottom layer a column and the same top layer, and this really does have a larger boundary.

So we try to introduce a "fake boundary" ∂'; we want $\partial' \leq \partial$ with equality for our extremal sets, and $\partial' C_i(A) \leq \partial' A \forall A$.

The obvious guess is $\partial' A = \sum |\partial A_t| + |A_1| - |A_k|$; note $\partial' = \partial$ on all down-sets. We do have $\partial' B \leq \partial' A$ where $B = C_i(A)$, but this fails for C_j for $j \neq i$ (and our next guess, $\partial'' A = \sum_{i=1}^n |A_1^{(i)}| - |A_k^{(i)}|$ cannot work e.g. because the "outside shell" of $[k]^n$ ould have $\partial'' = 0$.

We know: $|\partial A| = \partial' A \geq \partial' B = \sum |\partial B_t| + |B_1| - |B_k| = \sum f(|B_t|) + |B_1| - |B_k|$, where f is the extremal function in dimension $n-1$ [i.e. $f(r)$ is the smallest possible boundary for a set of size r]. Now $f(x)$ is the minimum of some functions of the form $cx^{1-\frac{1}{d}}$ and $c(k^{n-1} - x)^{1-\frac{1}{d}}$, each of which is concanve, so their pointwise minimum f is also concave [and therefore its minima over any given range occur at the extremes of the range]. Consider varying $|B_2|, \ldots, |B_{k-1}|$ keeping $|B_2| + \cdots + |B_{k-1}|$ fixed and $|B_1| \geq |B_2| \geq \cdots \geq |B_k|$. Then by concavity of f we have $\partial' B \geq \partial' C$ where C is given by: for some λ, $C_t = B_1$ if $t \leq \lambda$, B_k if $t > \lambda$ (again we are cheating a little).

We have : $|\partial A| = \partial' A \geq \partial' B \geq \partial' C$, but C is still not a down-set; $\partial' C = \lambda f(|B_1|) + (k - \lambda) f(|B_k|) + |B_1| - |B_k|$. Now consider varying $|B_1|$ and $|B_k|$ (with λ fixed), keeping $\lambda |B_1| + (k - \lambda)|B_k|$ fixed and keeping $|B_1| \geq |B_k$. Again this is a concave function of $|B_1|$ (as it is the sum of two concave functions and a linear one), so we have $\partial' C \geq \partial' D$ where either $D_t = D_1 \forall t$ or $D_t = D_1$ for $t \leq \lambda$, $\emptyset$ for $t > \lambda$ or $D_t = [k]^{n-1}$ if $t \leq \lambda$, D_k if $t > \lambda$.

Thus $|\partial A| = \partial' A \geq \partial' B \geq \partial' C \geq \partial' D = |\partial D|$ - "miraculously", D is a down-set. So we take our i-compression to be $A \mapsto D$, then finish the proof as usual.

Remarks: To make the arguments precise (rather than our "cheats"), work in the continuous cube $[0,1]^n$ instead (using some funny L^1 definition of surface area), and then pass to the discrete cube at the end.

Very few isoperimetric inequalities are known, even approximately. E.g. what about r-sets? Consider the graph on $X^{(r)}$ with x and y joined if $|x \cap y| = r-1$ (i.e. $d(x,y) = 2$ in Q_n); there is no good isoperimetric inequality known. The most interesting case is $r = \frac{n}{2}$ where the conjecture is that balls, i.e. sets of the form $\{x \in X^{(r)} : d(x, x_0) \leq d\}$ are best, i.e. sets of the form $\{x \subset [2r] : x \cap [r] \geq k\}$. But there is no proven result here.

3 Intersecting Families

There are two directions taken in this section; the first is obvious, the second is less natural, but leads to some very elegant mathematics which also has various applications.

t-intersecting families

We say $A \subset \mathbb{P}(X)$ is t-intersecting if $|x \cap y| \geq t \forall x, y \in A$. E.g. for $t = 2$, we could take $\{x : 1, 2 \in x\}$; this has size $\frac{1}{4}2^n$. It is beaten by: $\{x : |x| \geq \frac{n}{2} + 1\}$, of size $\sim \frac{1}{2}2^n$. This suggests:

3.1 Theorem (Katona's t-intersecting theorem)

Let $A \subset \mathbb{P}(X)$ be t-intersecting, taking $n + t$ even for convenience. Then $|A| \leq |X^{(\geq \frac{n+t}{2})}|$: for any $x, y \in A$, $d(x, y^c) \geq t$. Thus, writing $\overline{A} = \{x^c : x \in A\}$, we have $d(A, \overline{A}) \geq t$, i.e. $A_{(t-1)}$ is disjoint from $\overline{A}$¿ Suppose $|A| > |X^{(\geq \frac{n+t}{2})}|$. Then $|A_{t-1}| \geq |X^{(\geq \frac{n+t}{2} - (t-1))}|$ by Harper, but this is $|X^{(\geq \frac{n-t}{2}+1)}|$ and $|\overline{A}| > |X^{(\leq \frac{n-t}{2})}|$, contratictting $A_{(t-1)} \cap \overline{A} = \emptyset$.

What about $A \subset X^{(r)}$ t-intersecting $(1 \leq t \leq r)$? We could alsto try e.g. for $1 \leq \alpha \leq r - t$, $A_\alpha = \{x \in X^{(r)} : |x \cap \{1, 2, \ldots, t + 2\alpha\}| \geq t + \alpha\}$. E.g. $t = 2$ in $[8]^{(4)}$: $|A_0| = \binom{6}{2} = 15, |A_1| = 1 + \binom{4}{3}\binom{4}{1} = 17, |A_2| = \binom{6}{4} = 17$. In $[7]^{(4)}$, $|A_0| = \binom{5}{2} = 10, |A_1| = 1 + \binom{4}{3}\binom{4}{1} = 13, |A_2| = \binom{6}{4} = 15$. In $[9]^{(4)}$, $|A_0| = \binom{7}{2} = 21, |A_1| = 1 + \binom{4}{3}\binom{5}{1} = 21, |A_2| = \binom{6}{4} = 15$. As A_0 is quadratic (in n), $|A_1|$ linear, and $|A_2|$ constant, we will have A_0 winning if n is large.

3.2 Theorem

Let $A \subset X^{(r)}$ be t-intersecting, where r, t fixed, $1 \leq t \leq r$. Then $|A| \leq |A_0| = \binom{n-t}{r-t}$ if n sufficiently large.

Remarks: 1. This is sometimes called the "Second Erdős-Ko-Rado Theorem" 2. The bound on n given by our proof is $(16r)^r$; with more care one can make this $2tr^3$. However, we "don't care" what this bound actually is; only that the result is true for n sufficiently large.

Proof: The idea here is that A_0 has $r-t$ "degrees of freedom". Wlog take A maximal t-intersecting, so $\exists x, y \in A$ with $|x \cap y| = t$ (else for every $x \in A \forall i \in x, j \notin x$ we have $x \cup j - i \in A$, by maximality, so $A = X^{(r)}$, contradiction). We cannot have $x \cap y \subset z \forall z \in A$ (else $|A| \leq |A_0|$ and we are done). So choose $z \in A$ with $z \not\supset x \cap y$. Then for any $w \in A$, we must have $|w \cap (x \cup y \cup z)| \geq t+1$, so $|A| \geq 2^{3r}(\binom{n}{r-t-1} + \binom{n}{r-t-2} + \cdots + \binom{n}{0})$ (the 2^{3r} comes from the part of w on $x \cup y \cup z$, the other terms from the part off $x \cup y \cup z$). This is a polynomial of degree $r-t-1$, so $< |A_0|$ for n sufficiently large.

Remark: The Frankl conjecture was that if $A \subset X^{(r)}$ is t-intersecting then $|A| \leq \max(A_0, A_1, \ldots, A_{r-t})$. This was open for many years, and finally proven by Ahlswede and Khachatrian in 1998.

Modular Intersection

So far, we have "banned" $|x \cap y| = 0$. What if instead we ban $|x \cap y| \equiv 0$, modulo some number? E.g. suppose r odd, and we want $A \subset X^{(r)} : |x \cap y|$ odd $\forall x, y \in A$. We can achieve $|A| = \binom{\lfloor \frac{n-1}{2} \rfloor}{\frac{r-1}{2}}$ by taking sets x of the form: 1, tegether with $\frac{r-1}{2}$ of the pairs: $23, 45, 67, \ldots$.

What if (still for r odd) we wanted $|x \cap y|$ even $\forall x, y \in A$ $(x \neq y)$. We can achieve $|A| = n - r + 1$ by taking all x containing $1, 2, \ldots, r-1$. Amazingly:

3.3 Proposition

Let r be odd and let $A \subset X^{(r)}$ have $|x \cap y|$ even $\forall x, y \in A$ $(x \neq y)$. Then $|A| \leq n$. As a vague approximate motivation for what is mostly an out-of-the-hat proof: what is the only place n (rather than some $\binom{n}{k}$ or similar) appears in Q_n? It's only the dimension. So we will try to find $|A$ linearly independent vectors in an n-dimensional vector space: View each point $x \in Q_n$ as a point $\overline{x}$ in $\mathbb{Z}_2^n$ (e.g. $\{1,3,4\} \leftrightarrow (1,0,1,1,0,\ldots,0)$). Consider $\{\overline{x} : x \in A\}$. We have $\langle \overline{x}, \overline{x} \rangle = 1 \forall x \in A$ (where $\langle , \rangle$ is just the usual dot product). And $\langle \overline{x}, \overline{y} \rangle = 0 \forall x, y \in A, x \neq y$. Hence, $\{\overline{x} : x \in A\}$ are orthonormal, so linearly independent (if $\sum \lambda_i x_i = 0$ then dotting with x_j we obtain $\lambda_j = 0, \forall j$). So $|\{\overline{x} : x \in A\}| \leq n$.

Obviously it is possible to write out the above proof without linear algebra (by expressing the theorem that any linearly independent set can be extended to a basis in the language of sets), but this gives a completely incomprehensible (and several pages long) proof. No good elementary combinatorial proof is known.

For r even: if $|x \cap y|$ even $\forall$ distinct $x, y \in A \subset X^{(r)}$, we can have $|A| \geq \binom{\lfloor \frac{n}{2} \rfloor}{\frac{r}{2}}$ by $A = \{$sets of $\frac{r}{2}$ of the pairs $(12), (34), \ldots\}$. If $|x \cap y|$ odd for all distinct $x, y \in A \subset X^{(r)}$, then we must have $|A| \leq n+1$ - otherwise, add the point $n+1$ to each $x \in A$ and we obtain a contradiction to proposition 3. So banning $|x \cap y| \equiv r (\mod 2)$ forces $|A|$ to be small.

Does this generalise? We shall now show: "s allowed intersections $\mod p \Rightarrow |A| \leq \binom{n}{s}$":

3.4 Theorem (Frankl-Wilson)

(Unusually, particularly in this course, this is a theorem which does not follow Gauss' Law)

Let p be prime, and let $A \subset X^{(r)}$ be such that $\forall$ distinct $x, y \in A$ we have $|x \cap y| \equiv \lambda_i(\mod p)$ for some i, where $\lambda_1, \ldots, \lambda_s$ ($s \leq r$) are integers, none of which is $\equiv r(\mod p)$. Then $|A| \leq \binom{n}{s}$.

Remarks: 1. The bound is a polynomial in n, independent of r. This should be surprising. 2. The boind is essentially the best possible: take $A = \{x \in X^{(r)} : [r-s] \subset x\}$. This has $|A| = \binom{n-r+s}{s} \sim \binom{n}{s}$ (i.e. the ratio tends to 1 as $n \to \infty$ with r, s fixed). 3. There is no polynomial bound if we allow $|x \cap y| \equiv r \mod p$: write $r = a + \lambda p$ with $0 \leq a \leq p-1$, then we have $|x \cap y| \equiv r \mod p \forall x, y \in A$ and $|A| = \binom{\lfloor \frac{n-a}{p} \rfloor}{\lambda}$ by fixing a points, dividing the rest of X into sets of size p, and then taking sets consisting of the a fixed points and λ of the p-sets. This is not a polynomial; the degree grows as λ grows.

Proof: The idea is to try to find $|A|$ linearly independent points in a vector space of dimension $\binom{n}{s}$ by in some sense "applying the polynomial $(t-\lambda_1)\ldots(t-\lambda_s)$ to the values of $|x \cap y|$". For $i \leq j$, let $M(i, j)$ be the $\binom{n}{i} \times \binom{n}{j}$ matrix with rows indexed by $X^{(i)}$, columns indexed by $X^{(j)}$, with $M(i, j)_{xy} = 1$ if $x \subset y$, 0 otherwise.

The matrix $M(s, r)$ has $\binom{n}{s}$ rows; let V be the real vector space (real so that we may divide by natural numbers rather than having to worry whether they are 0 modulo some prime) spanned by these rows; we have $\dim V \leq \binom{n}{s}$.

Consider $M(i, s)M(s, r)$ for $i \leq s$. For any $x \in X^{(i)}$ and $y \in X^{(r)}$, $(M(i, s)M(s, r))_{xy}$ is the number of $z \in X^{(s)}$ with $x \subset z \subset y$; this is $\binom{r-i}{s-i}$ if $x \subset y$, 0 otherwise, i.e. $= M(i, r)_{xy}\binom{r-i}{s-i}$, so $M(i, s)M(s, r) = \binom{r-i}{s-i}M(i, r)$. So the rows of $M(i, r)$ belong to V (since they are just linear combinations of the rows of $M(s, r)$).

Consider $M(i) := M(i, r)^T M(i, r)$. Again, this must have all rows lying in V. For $x, y \in X^{(r)}$, $M(i)_{xy}$ is the number of $z \in X^{(i)}$ with $z \subset x, z \subset y$, i.e. $\binom{|x \cap y|}{i}$. Write the integer polynomial $(t-\lambda_1)\ldots(t-\lambda_s)$ as $\sum_{i=0}^{s} a_i \binom{t}{i}$ for some integers a_i; this is possible since $t(t-1)\ldots(t-i+1) = i!\binom{t}{i}$. Let $M = \sum_{i=0}^{s} a_i M(i)$. All rows of this are in V. Also, $M_{xy} = \sum_{i=0}^{s} a_i \binom{|x \cap y|}{i} = (|x \cap y| - \lambda_1)\ldots(|x \cap y| - \lambda_s)$ (setting $\binom{j}{i} = 0$ for $i > j$. So $M_{xy} \equiv 0 \mod p$ if $|x \cap y| \equiv$ some $\lambda_i(\mod p)$, $\not\equiv 0(\mod p)$ otherwise.

Now look at the submatrix whose rows and columns are indexed by A; it has diagonal entries $\not\equiv 0 \mod p$ and off-diagonal entries $\equiv 0 \mod p$. The rows are integer-valued (as a_i are integers) and linearly independent over $\mathbb{Z}_p$, and therefore over $\mathbb{Z}$, so over $\mathbb{Q}$, so over $\mathbb{R}$. Thus $|A| \leq \dim V = \binom{n}{s}$ and we are done.

Remark: We do need p prime: Grolmusz constructed, for each n, a value of $r \equiv 0(\mod 6)$ (or for a more general composite m in place of 6) and $A \subset X^{(r)}$ with $|x \cap y| \not\equiv 0(\mod 6) \forall$ distinct $x, y \in A$, with $|A| > n^{c \frac{\log n}{\log \log n}}$ for some $c > 0$; this grows faster than any polynomial in n.

3.5 Corollary

Let $A \subset X^{(r)}$ with $|x \cap y| \not\equiv r \mod p \forall$ distinct $x, y \in A$, for some prime $p < r$. Then $|A| \leq \binom{n}{p-1}$: we've allowed $p-1$ values of $|x \cap y| \mod p$.

Two $\frac{n}{2}$ sets from $[n]$ intersect (on average) in about $\frac{n}{4}$ points. Of course (for large n), intersection size exactly $= \frac{n}{4}$ is very unlikely. However, amazingly:

3.6 Corollary

Let p be prime, and let $A \subset [4p]^{(2p)}$ be such that $|x \cap y| \neq p \forall x, y \in A$. Then $|A| \leq 2\binom{4p}{p-1}$; note this is an exponentially small fraction of e.g. $\binom{4p}{2p}$; we have $\binom{n}{\frac{n}{4}} \leq 2e^{-\frac{n}{32}}2^n$ but $\binom{n}{\frac{n}{2}} \sim \frac{c}{\sqrt{2}}2^n$: by halving $|A|$, we may wlog assume we never have $x, x^c \in A$. Hence for distinct $x, y \in A$, we have $|x \cap y| \neq 0, p$ so $|x \cap y| \not\equiv 0$ mod p. Thus $|A| \leq \binom{4p}{p-1}$ by corollary 5. Note the result is true for general n, it's merely simpler to express for $n = 4p$.

Borsuk's Conjecture

Suppose S is a bounded set in $\mathbb{R}^n$. How many pieces do we need to break S into such that each piece has smaller diameter than S? Taking $S \subset \mathbb{R}^n$ to be a regular simplex, it is clear that we may need as many as $n + 1$ pieces. Borsuk's conjecture is that $n + 1$ pieces always suffices. This is known for $n = 1$ (trivial), $n = 2$ (easy), $n = 3$ (fiddly), and also for special cases e.g. $S \subset \mathbb{R}^n$ being a smooth convex body or a symmetric convex body (one for which $x \in S \Rightarrow -x \in S$) - and both these are substantial results. However, in fact, Borsuk's conjecture is false, and massively so:

3.7 Theorem (Kan, Kalai)

$\forall n \exists$ bounded $S \subset \mathbb{R}^n$ such that to break S into pieces of smaller diameter requires $\geq c^{\sqrt{n}}$ pieces, for some constant $c > 1$.

Notes: 1. Our proof will give that Borsuk is false for $n \geq 2000$; with much effort one can show this for $n \geq 581$. It would be nice to have a result for e.g. $n = 4, 5$, which remain open. 2. We will proove this for n of the form $\binom{4p}{2}$, where p prime; then the result follows (with a different c) $\forall n$, e.g. because $\exists p$ with $\frac{n}{2} \leq p \leq n$.

Proof: We consider $S \subset Q_n \subset \mathbb{R}^n$; in fact, $S \subset [n]^{(r)}$ (this is already a nonobvious thing to do). For $x, y \in S$, we have $\|x - y\|^2$ (where $\|\|$ is the usual Euclidean distance) $= 2(r - |x \cap y|)$. Thus we seek S with $|\ |x \cap y| = k$, say, but any subset of S with $\min |x \cap y| > k$ is much smaller than S. Now, the very clever part; let $n = \binom{4p}{2}$ where p prime. Identify $[n]$ with $E(K_{4p})$, the edge-set of the complete graph on $1, \ldots, 4p$. For each $x \in [4p]^{2p}$, let G_x be the complete bipartite graph on vertex classes x, x^c. Let $S = \{G_x : x \in [4p]^{2p}\}$. So $S \subset [n]^{(4p^2)}$, and $|S| = \frac{1}{2}\binom{4p}{2p}$. We have $|G_x \cap G_y| = |x \cap y|^2 + (2p - |x \cap y|)^2 = d^2 + (2p - d)^2$, where $d = |x \cap y|$. This is minimised when $d = p$, i.e. when $|x \cap y| = p$.

Now let $S' \subset S$ have smaller diameter than S: say $S' = \{G_x : x \in A\}$. Then we must have $|x \cap y| \neq p \forall x, y \in A$ (else diameter of S' = diameter of S). So $|A| \leq 2\binom{4p}{p-1}$ by Corollary 5. Thus the number of pieces needed is $\geq \frac{\frac{1}{2}\binom{4p}{2p}}{2\binom{4p}{p-1}} \geq \frac{c\frac{2^{4p}}{\sqrt{p}}}{4e^{-\frac{p}{2}}2^{4p} \geq c'^p}$ for some $c' > 1$, $\geq (c'')^{\sqrt{n}}$ for some $c'' > 1$, as required.

This is the end of the course.