

UNIVERSITY OF
CAMBRIDGE

MATHEMATICS TRIPOS

Part III

Combinatorics

Michaelmas, 2018

Lectures by

I. B. LEADER

Notes by

QIANGRU KUANG

Contents

1	Set systems	2
1.1	Chains & Antichains	2
1.2	Shadows	5
1.2.1	Two total orderings on $X^{(r)}$	6
1.2.2	Compressions	6
1.3	Intersecting families	11
2	Isoperimetric inequalities	14
	Index	19

1 Set systems

Definition (set system). Let X be a set. A *set system* on X (or *family of subsets*) of X is a family $\mathcal{A} \subseteq \mathcal{P}(X)$.

Example. We write $X^{(r)} = \{A \subseteq X : |A| = r\}$

In this course we almost deal exclusively with finite sets so unless otherwise stated, in this course we assume all sets are finite and $X = [n] = \{1, 2, \dots, n\}$.

For example, $|X^{(r)}| = \binom{n}{r}$. More concretely, for example,

$$[4]^{[2]} = \{12, 13, 14, 23, 24, 34\}$$

where 12 denotes $\{1, 2\}$ to avoid heavy notation. Therefore $|[4]^{[2]}| = 6$.

What is the mental picture for power set? Often we make $\mathcal{P}(X)$ into a graph, called Q_n , by joining A to B if $|A \Delta B| = 1$ where Δ is the symmetric difference, i.e. if $A = B \cup \{i\}$ for some $i \notin B$ (or vice versa).

Example. Q_3

Example. General picture for Q_n where n is even:
and when n is odd:

If we identify a set $A \subseteq X$ with a $\{0, 1\}$ sequence of length n (e.g. $134 \leftrightarrow 1011000 \dots 0$), via $A \leftrightarrow \mathbf{1}_A$ or χ_A . In this way, we can represent Q_n as a n -dimensional cube:

For this reason, Q_n is often called the *hypercube* or *discrete cube* or *n-cube*. In this way, the study of a set system become the study of a graph.

1.1 Chains & Antichains

Definition (chain). A family $\mathcal{A} \subseteq \mathcal{P}(X)$ is a *chain* if for all $A, B \in \mathcal{A}$, $A \subseteq B$ or $B \subseteq A$.

Example. $\{12, 125, 123589\}$.

On the other hand, we have

Definition (antichain). A family $\mathcal{A} \subseteq \mathcal{P}(X)$ is an *antichain* if for all $A, B \in \mathcal{A}$ with $A \neq B$, $A \not\subseteq B$.

Example. $\{1, 467, 2456\}$.

A natural question is: how large can a chain be? Obviously we can achieve $|\mathcal{A}| = n + 1$. We also cannot exceed $n + 1$ since a chain must meet each “level” $X^{(r)}$ for $0 \leq r \leq n$ in at most one place.

A less trivial question is how large an antichain can be. We could achieve $|\mathcal{A}| = n$ by enumerating all the singletons. It is maximal since adjoining any nonempty set to the family will result in an inclusion. However, it is not the largest antichain. Indeed, we could take $\mathcal{A} = X^{(r)}$ for any r . Thus we can achieve $|\mathcal{A}| = \binom{n}{\lfloor n/2 \rfloor}$. Can we beat it?

Pause for a moment and consider the longest chain problem. Why can a chain meet each level at only one place? One way to see this is that each level is an antichain: as we can decompose Q_n into $n + 1$ antichains, we cannot have a chain longer than that. Inspired by this, we try to decompose Q_n into chains.

Theorem 1.1 (Sperner's lemma). *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ where $|X| = n$ be an antichain. Then*

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

Proof. Sufficient to partition $\mathcal{P}(X)$ into $\binom{n}{\lfloor n/2 \rfloor}$ chains. For this sufficient to show:

1. for all $r < \frac{n}{2}$, there exist matchings from $X^{(r)}$ to $X^{(r+1)}$ where a matching is just a set of non-adjacent edges,
2. for all $r > \frac{n}{2}$, there exist matchings from $X^{(r)}$ to $X^{(r-1)}$.

Then put these matchings together to form chains. Each passes through $X^{(n/2)}$ so there are $\binom{n}{\lfloor n/2 \rfloor}$ of them.

By taking complements, sufficient to prove 1. Consider the subgraph of Q_n spanned by $X^{(r)} \cup X^{(r+1)}$ which is bipartite. For any $\mathcal{B} \subseteq X^{(r)}$, let $\Gamma(\mathcal{B})$ be the neighbourhood (in $X^{(r+1)}$) of \mathcal{B} . Then we have

$$\#(\mathcal{B} - \Gamma(\mathcal{B}) \text{ edges}) = |\mathcal{B}|(n - r)$$

as each point in $X^{(r)}$ has degree $n - r$. Meanwhile

$$\#(\mathcal{B} - \Gamma(\mathcal{B}) \text{ edges}) \leq |\Gamma(\mathcal{B})|(r + 1)$$

as each point in $X^{(r+1)}$ has degree $r + 1$. Thus

$$|\Gamma(\mathcal{B})| \geq |\mathcal{B}| \frac{n - r}{r + 1} \geq |\mathcal{B}|$$

as $r < \frac{n}{2}$. Hence by Hall's theorem there exist matchings. \square

Remark.

1. $\binom{n}{\lfloor n/2 \rfloor}$ is achievable, e.g. $\mathcal{A} = X^{(\lfloor n/2 \rfloor)}$.
2. Note that the theorem says nothing about extremal cases — which antichain have this size?

The aim is to show that for \mathcal{A} an antichain,

$$\sum_{r=0}^n \frac{|\mathcal{A} \cap X^{(r)}|}{\binom{n}{r}} \leq 1.$$

Note that this trivially implies Sperner's lemma. We will use the same setup but bound the numbers more carefully. To justify the following definition, we will write $X^{(r)}$ above $X^{(r-1)}$ in Q_n .

Definition. Let $\mathcal{A} \subseteq X^{(r)}$ for some $1 \leq r \leq n$. The *shadow* or *lower shadow* of \mathcal{A} is

$$\partial A = \partial^- A = \{A - \{i\} : A \in \mathcal{A}, i \in A\}.$$

so $\partial A \subseteq X^{(r-1)}$.

Example. Let $\mathcal{A} = \{123, 124, 134, 135\} \subseteq X^{(3)}$. Then

$$\partial A = \{12, 13, 23, 14, 24, 34, 15, 35\} \subseteq X^{(2)}.$$

Lemma 1.2 (Local LYM). *Let $\mathcal{A} \subseteq X^{(r)}$ where $1 \leq r \leq n$. Then*

$$\frac{|\partial\mathcal{A}|}{\binom{n}{r-1}} \geq \frac{|\mathcal{A}|}{\binom{n}{r}}.$$

Informally, the fraction of the later occupied increases when we take the shadow.

Proof.

$$\#(\mathcal{A} - \partial\mathcal{A} \text{ edges in } Q_n) = r|\mathcal{A}|$$

by counting from above and

$$\#(\mathcal{A} - \partial\mathcal{A} \text{ edges in } Q_n) \leq (n-r+1)|\partial\mathcal{A}|$$

counting from below so

$$\frac{|\partial\mathcal{A}|}{|\mathcal{A}|} \geq \frac{r}{n-r+1}$$

but

$$\frac{\binom{r}{r-1}}{\binom{n}{r}} = \frac{r}{n-r+1}.$$

□

When does equality hold in local LYM? We'll need

$$(A - \{i\}) \cup \{j\} \in \mathcal{A}$$

for all $a \in \mathcal{A}, i \in A, j \notin A$. Hence $\mathcal{A} = X^{(r)}$ or \emptyset .

Theorem 1.3 (LYM inequality). *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an antichain. Then*

$$\sum_{r=0}^n \frac{|\mathcal{A} \cap X^{(r)}|}{\binom{n}{r}} \leq 1.$$

Proof. The whole idea of the proof can be summarised by “bubble down with local LYM”. Let $\mathcal{A}_r = \mathcal{A} \cap X^{(r)}$. Obviously

$$\frac{|\mathcal{A}_n|}{\binom{n}{n}} \leq 1.$$

Also $\partial\mathcal{A}_n$ and \mathcal{A}_{n-1} are distinct as \mathcal{A} is an antichain. Thus

$$\frac{|\partial\mathcal{A}_n|}{\binom{n}{n-1}} + \frac{|\mathcal{A}_{n-1}|}{\binom{n}{n-1}} = \frac{|\partial\mathcal{A}_n \cup \mathcal{A}_{n-1}|}{\binom{n}{n-1}} \leq 1$$

so

$$\frac{|\mathcal{A}_n|}{\binom{n}{n}} + \frac{|\mathcal{A}_{n-1}|}{\binom{n}{n-1}} \leq 1$$

by local LYM.

Also $\partial(\partial\mathcal{A}_n \cup \mathcal{A}_{n-1})$ is distinct from \mathcal{A}_{n-2} , again since \mathcal{A} is an antichain. Thus

$$\frac{|\partial(\partial\mathcal{A}_n \cup \mathcal{A}_{n-1})|}{\binom{n}{n-2}} + \frac{|\mathcal{A}_{n-2}|}{\binom{n}{n-2}} \leq 1,$$

so

$$\frac{|\partial\mathcal{A}_n \cup \mathcal{A}_{n-1}|}{\binom{n}{n-1}} + \frac{|\mathcal{A}_{n-1}|}{\binom{n}{n-2}} \leq 1,$$

so

$$\frac{|\mathcal{A}_n|}{\binom{n}{n}} + \frac{|\mathcal{A}_{n-1}|}{\binom{n}{n-1}} + \frac{|\mathcal{A}_{n-2}|}{\binom{n}{n-2}} \leq 1$$

Keep going we have the result desired. \square

Again we can ask when we have equality in LYM inequality. This happens if and only if we have equality in each use of local LYM, so the “first” (greatest) r with $\mathcal{A}_r \neq \emptyset$ must have $\mathcal{A}_r = X^{(r)}$ so $\mathcal{A} = X^{(r)}$. Thus we know equality in Sperner’s lemma ($\mathcal{A} = X^{(n/2)}$ for n even, similar for n odd)

Proof 2. Choose uniformly at random a maximal chain \mathcal{C} (i.e. $C_0 \subseteq C_1 \subseteq \dots \subseteq C_n$ with $|C_i| = i$). For a given r -set A , $\mathbb{P}(A \in \mathcal{C}) = \binom{n}{r}^{-1}$ as all r -sets are equally likely. Therefore

$$\mathbb{P}(\mathcal{A}_r \text{ meets } \mathcal{C}) = \frac{|\mathcal{A}_r|}{\binom{n}{r}}$$

since the events are disjoint. Furthermore for different r the events that \mathcal{A} meets X^r are also disjoint so

$$\mathbb{P}(\mathcal{A} \text{ meets } \mathcal{C}) = \sum_{r=0}^n \frac{|\mathcal{A}_r|}{\binom{n}{r}}$$

and of course it is less than 1. \square

Remark. Equivalently, the number of maximal chains is $n!$ and the number of them containing a given r -set is $r!(n-r)!$, so

$$\sum_{r=0}^n |\mathcal{A}_r| r!(n-r)! \leq n!$$

so this is probability in disguise

1.2 Shadows

For $\mathcal{A} \subseteq X^{(r)}$, we know $|\partial\mathcal{A}| \geq |\mathcal{A}| \frac{r}{n-r+1}$ — but equality is rare (only for $\mathcal{A} = \emptyset$ or $\mathcal{A} = X^{(r)}$). The natural question is: given $|\mathcal{A}|$, how should we choose $\mathcal{A} \subseteq X^{(r)}$ to minimise $|\partial\mathcal{A}|$? Informally, this asks how “tightly” can we pack some r -sets.

If $|\mathcal{A}| = \binom{k}{r}$, it is believable that we would take $\mathcal{A} = [k]^{(r)}$, which gives $[k]^{(r-1)}$. What if $\binom{k}{r} < |\mathcal{A}| < \binom{k+1}{r}$, it is believable that we’d take $[k]^{(r)}$ and some other r -sets from $[k+1]^{(r)}$. For example if $\mathcal{A} \subseteq X^{(3)}$ with $|\mathcal{A}| = \binom{7}{3} + \binom{4}{2}$, we would take

$$\mathcal{A} = [7]^{(3)} \cup \{A \cup \{8\} : A \in [4]^{(2)}\},$$

i.e. take those in $[7]$ of size 3 so that they are as tightly packed as possible and then choose some other stuff.

If we increment the size of \mathcal{A} by 1, it is believable that we should take the above \mathcal{A} and adjoin another element from $[4]$. Thus it seems that there is a total order on subsets of X of a given size, and we just take the first $|\mathcal{A}|$.

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

Definition (lexicographic order). Given $A, B \in X^{(r)}$, say $A = a_1, \dots, a_r, B = b_1, \dots, b_r$, where we use the notation to mean $a_1 < \dots < a_r$, say $A < B$ in the *lexicographic* or *lex* order if for some i have $a_i < b_i$ and $a_j = b_j$ for all $j < i$.

Equivalently, $a_i < b_i$, where $i = \min\{j : a_j \neq b_j\}$. The slogan is “use small number”.

Example. Lex on $[4]^{(2)}$:

12, 13, 14, 23, 24, 34.

Lex on $[6]^{(3)}$:

123, 124, 125, 126, 134, 135, 136, 145, 146, 156,
234, 235, 236, 245, 246, 256, 345, 346, 356, 456.

However, in the shadow minimisation problem we want to avoid large numbers as much as possible, i.e. keep the largest number as small as possible.

Definition (colexicographic order). Say $A < B$ in the *colexicographic* or *colex* order if for some i have $a_i < b_i$ and $a_j = b_j$ for all $j > i$.

Equivalently, $a_i < b_i$ where $i = \max\{j : a_j \neq b_j\}$. The slogan is “avoid large number”. Equivalently, $A < B$ if $\sum_{i \in A} 2^i < \sum_{i \in B} 2^i$.

Example. Colex on $[4]^{(2)}$:

12, 13, 23, 14, 24, 34.

Colex on $[6]^{(2)}$:

123, 124, 134, 234, 125, 135, 235, 145, 245, 345,
126, 136, 236, 146, 246, 346, 156, 256, 356, 456.

Note. In colex, $[k]^{(r)}$ is an initial segment of $[k+1]^{(r)}$, meaning that it is the first t elements for some t . Therefore we could view colex as an enumeration of $\mathbb{N}^{(r)}$. Try this with lex and see what happens!

Following our heuristics just now, the aim is to show initial segments of colex minimise ∂ , i.e. if $\mathcal{A} \subseteq X^{(r)}$ and $\mathcal{C} \subseteq X^{(r)}$ is the first $|\mathcal{A}|$ r -sets in colex then $|\partial\mathcal{A}| \geq |\partial\mathcal{C}|$, This is Kruskal-Katona theorem, the first theorem in combinatorics. In particular, $|\mathcal{A}| = \binom{k}{r}$ implies $|\partial\mathcal{A}| \geq \binom{k}{r-1}$. However, unless \mathcal{A} is written in very nice form, it is very difficult to estimate $\partial\mathcal{A}$.

1.2.2 Compressions

The idea is to “replace” $\mathcal{A} \subseteq X^{(r)}$ with some $\mathcal{A}' \subseteq X^{(r)}$ such that

1. $|\mathcal{A}'| = |\mathcal{A}|$,
2. $|\partial\mathcal{A}'| \leq |\partial\mathcal{A}|$,

3. \mathcal{A}' “looks more like \mathcal{C} ” than \mathcal{A} did.

Ideally, we would compress

$$\mathcal{A} \rightarrow \mathcal{A}' \rightarrow \mathcal{A}'' \rightarrow \dots \rightarrow \mathcal{B}$$

where either $\mathcal{B} = \mathcal{C}$, or \mathcal{B} is so similar to \mathcal{C} that we can see directly that $|\partial\mathcal{B}| \geq |\partial\mathcal{C}|$.

Colex prefers 1 to 2

Definition (*ij*-compression). For $1 \leq i < j \leq n$, the *ij*-compression C_{ij} is defined by for $A \subseteq X$,

$$C_{ij}(A) = \begin{cases} A - j + i & \text{if } j \in A, i \notin A \\ A & \text{otherwise} \end{cases}$$

and for $\mathcal{A} \subseteq \mathcal{P}(X)$,

$$C_{ij}(\mathcal{A}) = \{C_{ij}(A) : A \in \mathcal{A}\} \cup \{A \in \mathcal{A} : C_{ij}(A) \in \mathcal{A}\}.$$

Example. If $\mathcal{A} = \{123, 134, 234, 235, 247\}$, then

$$C_{12}(\mathcal{A}) = \{123, 134, 234, 135, 147\}.$$

$|C_{ij}(\mathcal{A})| = |\mathcal{A}|$ and after the compression it looks “more” like colex than lex. Say \mathcal{A} is *ij*-compressed if $C_{ij}(\mathcal{A}) = \mathcal{A}$. It is also intuitively obvious that the operation is indeed a compression in the sense that it decreases the shadow.

Proposition 1.4. Let $\mathcal{A} \subseteq X^{(r)}$, $1 \leq i < j \leq n$ then

$$|\partial C_{ij}(\mathcal{A})| \leq |\partial\mathcal{A}|.$$

Proof. Write \mathcal{A}' for $C_{ij}(\mathcal{A})$. We’ll show that if $B \in \partial\mathcal{A}' - \partial\mathcal{A}$ then $i \in B, j \notin B$ and $B \cup j - i \in \partial\mathcal{A} - \partial\mathcal{A}'$ (so B has a preimage under C_{ij}). Then done.

We have $B \cup x \in \mathcal{A}'$ for some $x \notin B$ and $B \cup x \notin \mathcal{A}$. Hence $i \in B \cup x, j \notin B \cup x$ and $(B \cup x) \cup j - i \in \mathcal{A}$ (so $B \cup x$ is the set that has been compressed). Note that $x \neq i$ else $B \cup j \in \mathcal{A}$, contradicting $B \notin \partial\mathcal{A}$.

Certainly $B \cup j - i \in \partial\mathcal{A}$. Claim that $B \cup j - i \notin \partial\mathcal{A}'$, thereby completing the proof: suppose $(B \cup j - i) \cup y \in \mathcal{A}'$. We cannot have $y = i$, else $B \cup j \in \mathcal{A}'$, whence $B \cup j \in \mathcal{A}$, contradiction. Thus $j \in (B \cup j - i) \cup y$ and $i \notin (B \cup j - i) \cup y$, so $(B \cup j - i) \cup y \in \mathcal{A}$ and $B \cup y \in \mathcal{A}$ by definition of C_{ij} . Contradiction. \square

Remark. We have actually showed that

$$\partial C_{ij}(\mathcal{A}) \subseteq C_{ij}(\partial\mathcal{A}),$$

“shadow of compression lives inside compression of shadow”.

Definition (left-compressed). Say $\mathcal{A} \subseteq X^{(r)}$ is *left-compressed* if $C_{ij}(\mathcal{A}) = \mathcal{A}$ for all $i < j$.

Proposition 1.5. Let $\mathcal{A} \subseteq X^{(r)}$. Then there exists left-compressed $\mathcal{B} \subseteq X^{(r)}$ with $|\mathcal{B}| = |\mathcal{A}|$ and $|\partial\mathcal{B}| \leq |\partial\mathcal{A}|$.

Proof. Morally we only have to show that any sequence of compression terminate. Among all $\mathcal{B} \subseteq X^{(r)}$ with $|\mathcal{B}| = |\mathcal{A}|$ and $|\partial\mathcal{B}| \leq |\partial\mathcal{A}|$, choose one with

$$\sum_{A \in \mathcal{B}} \sum_{x \in A} x$$

minimal. Then \mathcal{B} is left compressed, as if $C_{ij}(\mathcal{B}) \neq \mathcal{B}$ then we contradict minimality. \square

Note.

1. Alternatively, we may apply one C_{ij} then another and so on — it must terminate.
2. In fact we can apply each C_{ij} at most once if we choose a sensible order.

Certainly initial segments of colex are left-compressed. The converse is blatantly false, e.g. $\mathcal{A} = \{123, 124, 125, 126, 127\}$. We need to do more.

colex prefers 23 to 14 We can compress not only singletons but also sets of larger sizes:

Definition (*UV-compression*). For $U, V \subseteq X$ with $|U| = |V|$ and $U \cap V = \emptyset$, the *UV-compression* C_{UV} is defined by for $A \subseteq X$,

$$C_{UV}(A) = \begin{cases} A \cup U - V & \text{if } V \subseteq A, U \cap A = \emptyset \\ A & \text{otherwise} \end{cases}$$

and for $\mathcal{A} \subseteq X^{(r)}$,

$$C_{UV}(\mathcal{A}) = \{C_{UV}(A) : A \in \mathcal{A}\} \cup \{A \in \mathcal{A} : C_{UV}(A) \in \mathcal{A}\}.$$

Example. If $\mathcal{A} = \{123, 134, 235, 145, 146, 157\}$, then

$$C_{23,14}(\mathcal{A}) = \{123, 134, 235, 145, 236, 157\}.$$

Similar observation: $|C_{UV}(\mathcal{A})| = |\mathcal{A}|$. Say \mathcal{A} is *UV-compressed* if $C_{UV}(\mathcal{A}) = \mathcal{A}$. Sadly, C_{UV} need *not* decrease shadow — e.g. $\mathcal{A} = \{146, 467\}$ then $C_{23,14}(\mathcal{A}) = \{236, 467\}$ so $|\partial\mathcal{A}| = 5$, $|\partial C_{23,14}(\mathcal{A})| = 6$. Intuitively, removing one elements and adding another sends a set to a “close neighbour” which shares largely the same shadow, but once we start doing *UV-compression*, some things are “moved a long way”.

However, note that \mathcal{A} in the above example is not left-compressed. It turns out once you have done the “smaller” compressions, doing a larger compression always decrease the shadow. Formally,

Proposition 1.6. *Let $\mathcal{A} \subseteq X^{(r)}$ and $U, V \subseteq X$ with $|U| = |V|$ and $U \cap V = \emptyset$. Suppose that for all $x \in U$ exists $y \in V$ such that \mathcal{A} is $(U-x, V-y)$ -compressed then*

$$|\partial C_{UV}(\mathcal{A})| \leq |\partial \mathcal{A}|.$$

Proof. Write \mathcal{A}' for $C_{UV}(\mathcal{A})$. Given $B \in \partial \mathcal{A}' - \partial \mathcal{A}$, we'll show that $U \subseteq B, V \cap B = \emptyset$ and $B \cup V - U \in \partial \mathcal{A} - \partial \mathcal{A}'$.

We have $B \cup x \in \mathcal{A}'$ for some $x \notin B$, with $B \cup x \notin \mathcal{A}$. So $U \subseteq B \cup x, V \cap (B \cup x) = \emptyset$ and $(B \cup x) \cup V - U \in \mathcal{A}$. Thus certainly $V \cap B = \emptyset$.

If $x \in U$, have \mathcal{A} is $(U-x, V-y)$ -compressed for some $y \in V$ so from $(B \cup x) \cup V - U \in \mathcal{A}$ we obtain $B \cup y \in \mathcal{A}$, contradicting $B \notin \partial \mathcal{A}$. Hence $x \notin U$ and so $U \subseteq B$. Also $B \cup V - U \in \partial \mathcal{A}$ (as $(B \cup x) \cup V - U \in \mathcal{A}$).

Suppose $B \cup V - U \in \partial \mathcal{A}'$ then $(B \cup V - U) \cup w \in \mathcal{A}'$ for some w .

1. if $w \notin U$, then $V \subseteq (B \cup V - U) \cup w$ and $U \cap ((B \cup V - U) \cup w) = \emptyset$ so from $(B \cup V - U) \cup w \in \mathcal{A}'$ we conclude that both $(B \cup V - U) \cup w \in \mathcal{A}$ and $B \cup w \in \mathcal{A}$, contradicting $B \notin \partial \mathcal{A}$.
2. if $w \in U$, we have \mathcal{A} is $(U-w, V-z)$ -compressed for some $z \in V$. From $(B \cup V - U) \cup w \in \mathcal{A}$ (as it is in \mathcal{A}' and contains V , so could not have moved), we deduce $B \cup z \in \mathcal{A}$, contradicting $B \notin \partial \mathcal{A}$.

□

Remark. We have actually showed that

$$\partial C_{UV}(\mathcal{A}) \subseteq C_{UV}(\partial \mathcal{A}).$$

Theorem 1.7 (Kruskal-Katona). *Let $\mathcal{A} \subseteq X^{(r)}$ where $1 \leq r \leq n$ and let \mathcal{C} be the initial segment of colex on $X^{(r)}$ with $|\mathcal{C}| = |\mathcal{A}|$, then*

$$|\partial \mathcal{A}| \geq |\partial \mathcal{C}|.$$

In particular if $|\mathcal{A}| = \binom{k}{r}$ then

$$|\partial \mathcal{A}| \geq \binom{k}{r-1}.$$

Proof. Let

$$\Gamma = \{(U, V) : U, V \subseteq X, |U| = |V| > 0, U \cap V = \emptyset, \max U < \max V\}$$

which are the ordered pairs (U, V) with $U < V$ in colex, which are exactly the sensible pairs to do UV compression on in order to decrease shadow. Define a sequence of set systems $\mathcal{A}_0, \mathcal{A}_1, \dots$ in $X^{(r)}$ as follow:

1. $\mathcal{A}_0 = \mathcal{A}$.
2. If \mathcal{A}_k is (U, V) -compressed for all $(U, V) \in \Gamma$, then stop the sequence with \mathcal{A}_k .

3. If not, choose $(U, V) \in \Gamma$ such that \mathcal{A}_k is not (U, V) -compressed with U minimal. Set $\mathcal{A}_{k+1} = C_{UV}(\mathcal{A}_k)$. Note that for all $x \in U$, we have $(U - x, V - y) \in \Gamma \cup \{(\emptyset, \emptyset)\}$ for $y = \min V$ so by Proposition 1.6 have $|\partial\mathcal{A}_{k+1}| \leq |\partial\mathcal{A}_k|$.

The sequence must terminate, for example because as $\sum_{A \in \mathcal{A}_k} \sum_{i \in A} 2^i$ is decreasing in k .

The final system $\mathcal{B} = \mathcal{A}_k$ satisfies $|\mathcal{B}| = |\mathcal{A}|$ and $|\partial\mathcal{B}| \leq |\partial\mathcal{A}|$. Moreover, \mathcal{B} is (U, V) -compressed for all $(U, V) \in \Gamma$. Claim that $\mathcal{B} = \mathcal{C}$:

Proof. Suppose \mathcal{B} is not an initial segment of colex. Then there exists $A < B$ in colex with $A \notin \mathcal{B}$ and $B \in \mathcal{B}$. But then $U = A - B$ and $V = B - A$ have $(U, V) \in \Gamma$ and $C_{UV}(\mathcal{B}) = A$. Absurd. \square

\square

Remark.

1. Equivalently, we may state the theorem in numerical form by translating initial segments of colex into the size of its constituent “lex” parts: if $\mathcal{A} \subseteq X^{(r)}$ with

$$|\mathcal{A}| = \binom{k_r}{r} + \binom{k_{r-1}}{r-1} + \dots + \binom{k_s}{s}$$

where $k_r > k_{r-1} > \dots > k_s$ and $s > 0$, then

$$|\partial\mathcal{A}| \geq \binom{k_r}{k-1} + \binom{k_{r-1}}{r-2} + \dots + \binom{k_s}{s-1}.$$

2. In proof of Krustal-Katona we used only Proposition 1.6, but neither Proposition 1.4 nor Proposition 1.5. However, deriving simpler results for ij -compression provides motivation and intuition for the further full-fledged proof.
3. Can we ask for uniqueness? We can check that if $|\partial\mathcal{A}| = |\partial\mathcal{C}|$ and $|\mathcal{A}| = \binom{k}{r}$ then $\mathcal{A} = Y^{(r)}$ for some k -set Y . Thus this is unique up to isomorphism. But in general it is not true that $|\partial\mathcal{A}| = |\partial\mathcal{C}|$ implies \mathcal{A} is isomorphic to \mathcal{C} ($\mathcal{A} \subseteq \mathcal{P}(X), \mathcal{B} \subseteq \mathcal{P}(Y)$ are isomorphic if there exists a bijection between X and Y sending \mathcal{A} to \mathcal{B}).

What about upper shadow?

Definition (upper shadow). For $\mathcal{A} \subseteq X^{(r)}$ where $0 \leq r \leq n-1$, the *upper shadow* of \mathcal{A} is

$$\partial^+\mathcal{A} = \{A \cup x : A \in \mathcal{A}, x \notin A\}.$$

Note that $A < B$ in colex if and only if $A^c < B^c$ in lex with ground-set order reversed. You can mess around with complement and other set operations, but think about it until it becomes clear!

Corollary 1.8. *Let $\mathcal{A} \subseteq X^{(r)}$ where $0 \leq r \leq n-1$ and let \mathcal{C} be the initial segment of lex with $|\partial\mathcal{C}| = |\mathcal{A}|$. Then*

$$|\partial^+ \mathcal{A}| \geq |\partial^+ \mathcal{C}|.$$

Proof. Take complements. □

Also the shadow of an initial segment of colex is again an initial segment of colex. Indeed, if

$$\mathcal{C} = \{A \subseteq X^{(r)} : A \leq a_1, a_2, \dots, a_r\}$$

then

$$\partial\mathcal{C} = \{B \subseteq X^{(r-1)} : B \leq a_2, \dots, a_r\}.$$

Kruskal-Katona also proves a generalised version of itself:

Corollary 1.9. *Let $\mathcal{A} \subseteq X^{(r)}$ and let $\mathcal{C} \subseteq X^{(r)}$ be the initial segment of colex with $|\mathcal{C}| = |\mathcal{A}|$. Then*

$$|\partial^t \mathcal{A}| \geq |\partial^t \mathcal{C}|$$

for all $1 \leq t \leq r$. In particular if $|\mathcal{A}| = \binom{k}{r}$ then

$$|\partial^t \mathcal{A}| \geq \binom{k}{r-t}.$$

Proof. If $|\partial^t \mathcal{A}| \geq |\partial^t \mathcal{C}|$ then $|\partial^{t+1} \mathcal{A}| \geq |\partial^{t+1} \mathcal{C}|$ by **Kruskal-Katona**. □

1.3 Intersecting families

Definition (intersecting family). A family $\mathcal{A} \subseteq \mathcal{P}(X)$ is *intersecting* if $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{A}$.

How large can $|\mathcal{A}|$ be? We can achieve $|\mathcal{A}| = 2^{n-1}$ by fixing an element, e.g. $\mathcal{A} = \{A \subseteq X : 1 \in A\}$. We can't beat it either.

Proposition 1.10. *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be intersecting. Then $|\mathcal{A}| \leq 2^{n-1}$.*

Proof. For each $A \subseteq X$, we have at most one of A or A^c in \mathcal{A} . □

Note. There are many examples with $|\mathcal{A}| = 2^{n-1}$. In the discrete cube, we can take the “top” layer along any element. In the blob graph, we can take $\{A \subseteq X : |A| > \frac{n}{2}\}$.

Now consider the same problem restricted to one level. How larger can $|\mathcal{A}|$ be if $\mathcal{A} \subseteq X^{(r)}$ is intersecting? If $r > \frac{n}{2}$, it is trivial as we can take $\mathcal{A} = X^{(r)}$. If $r = \frac{n}{2}$, the maximum is $\frac{1}{2} \binom{n}{r}$ — just choose one of A, A^c for each $A \in X^{(r)}$. Thus from now on assume $r < \frac{n}{2}$.

Take

$$\mathcal{A} = \{A \in X^{(r)} : 1 \in A\},$$

we have

$$|\mathcal{A}| = \binom{n-1}{r-1} = \frac{r}{n} \binom{n}{r},$$

where the second equality comes from the probability an element of $X^{(r)}$ containing 1. We may also take another family, say

$$\mathcal{B} = \{A \in X^{(r)} : |A \cap \{1, 2, 3\}| \geq 2\}.$$

For example, in case of $[8]^{(3)}$,

$$|\mathcal{A}| = \binom{3}{2} = 2, |\mathcal{B}| = 1 + \frac{3}{2} \binom{5}{1} = 16 < 21$$

so \mathcal{A} beats \mathcal{B} .

Theorem 1.11 (Erdős-Ko-Rado). *Let $r < \frac{n}{2}$ and let $\mathcal{A} \subseteq X^{(r)}$ be intersecting. Then*

$$|\mathcal{A}| \leq \frac{n-1}{r-1}.$$

Proof 1. “Bubble down with Krustal-Katona”

For $A, B \in \mathcal{A}$ have $A \cap B \neq \emptyset$, i.e. $A \not\subseteq B^c$. Write

$$\overline{\mathcal{A}} = \{A^c : A \in \mathcal{A}\} \subseteq X^{(n-r)}.$$

This says that $\partial^{n-2r}\overline{\mathcal{A}}$ is disjoint from \mathcal{A} . Suppose $|\mathcal{A}| > \binom{n-1}{r-1}$. Then $|\overline{\mathcal{A}}| = \binom{n-1}{n-r} = \binom{n-1}{n-r}$ so by iterated Krustal-Katona, have

$$|\partial^{n-2r}\overline{\mathcal{A}}| \geq \binom{n-1}{r}.$$

But $\binom{n-1}{r-1} + \binom{n-1}{r} = \binom{n}{r}$, i.e.

$$|\partial^{n-2r}\overline{\mathcal{A}}| + |\mathcal{A}| > |X^{(r)}|,$$

absurd. □

Remark. Note that in the proof we “happen” to have $|\overline{\mathcal{A}}| = \binom{n-1}{n-r}$. In fact, the numbers have to work as we get equality for $\mathcal{A} = \{A \in X^{(r)} : 1 \in A\}$.

Proof 2. Consider a cyclic ordering c of $[n]$, i.e. a bijection $c : [n] \rightarrow \mathbb{Z}/(n\mathbb{Z})$. How many $A \in \mathcal{A}$ are intervals (sets of r consecutive elements) in our ordering? The answer is at most r . Indeed, suppose $c_1, \dots, c_r \in A$. Then for each $1 \leq i \leq r-1$, at most one of the two intervals $\dots, c_{i-1}c_i$ and $c_{i+1}, c_{i+2} \dots$ can belong to \mathcal{A} so that \mathcal{A} is intersecting. Also a given r -set A is an interval in exactly $n \cdot r! \cdot (n-r)!$ of the $n!$ cyclic orderings. Therefore

$$|\mathcal{A}| n \cdot r! (n-r)! \leq n! r$$

i.e.

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

□

Remark.

1. Equivalently, we are double-counting the edges in the bipartite graph with vertex classes \mathcal{A} and all cyclic orderings, in which $A \in \mathcal{A}$ is joined to c if A is an interval in c .
2. This method, by constructing an clever auxiliary object and counting a parameter of it in two ways, is called *averaging* or *Katona's method*.

We may ask when we have equality in Erdős-Ko-Rado. We can check that equality holds if and only if $\mathcal{A} = \{A \in X^{(r)} : i \in A\}$ for some i . This follows (proof 1) from equality case of Kruskal-Katona (when size is nice), or (proof 2) by considering changing the cyclic ordering bit by bit.

2 Isoperimetric inequalities

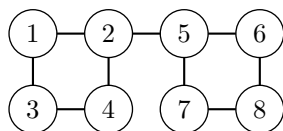
The general theme in this chapter is the question “how tightly can we pack a subset of given size in a space”. Many results in combinatorics follow from this chapter, and Krustal-Katona is almost one of the cases.

To motivate the name “isoperimetric”, observe the basic fact in geometry that among subsets of \mathbb{R}^2 of given area, a disc has smallest perimeter. Similarly among subsets of \mathbb{R}^3 of given volume, a ball has smallest surface area. Among subsets of S^2 of given area, the cap has the smallest perimeter. What are the corresponding notions in discrete settings, say subgraphs of a graph? Obviously the “volume” should be the number of vertices. What about the “surface area” and more generally, its boundary?

Definition (boundary). For a set A of vertices in a graph G , the *boundary* of A is

$$b(A) = \{x \in V(G) : x \notin A, xy \in E \text{ for some } y \in A\}.$$

Example. In the following graph if $A = \{1, 2, 3\}$ then $b(A) = \{4, 5\}$.



An *isoperimetric inequality* of G is an inequality of the form

$$b(A) \geq f(|A|)$$

for $A \subseteq V(G)$. This should be a “good” (at least nontrivial) inequality, which warrants that such an inequality should describe a fixed graph, and not applicable to graphs in general.

To minimise the size $b(A)$ is equivalently to (and usually easier to) the size of

Definition (neighbourhood). A *neighbourhood* of A

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

where d the usual graph distance.

A natural guess for boundary minimising A is often

$$B(x, r) = \{y \in G : d(x, y) \leq r\}.$$

Let’s test it. Take $|A| = 4$ in Q_3 . Then in the discrete A is either a hyperplane or a ball. We have $|b(A)| = 3$ for ball and $|b(A)| = 4$ for hyperplane. This works so we conjecture that “balls are the best”, i.e. if $|A| = |X^{(\leq r)}|$ then $|N(A)| \geq |X^{(\leq r+1)}|$. When $|A|$ is strictly between $\sum_{i=0}^r \binom{n}{i}$ and $\sum_{i=0}^{r+1} \binom{n}{i}$, conjecture that $b(A)$ is minimised if $A = X^{(\leq r)} \cup B$ for some $B \subseteq X^{(r+1)}$. This is a *Hamming ball*. Note that if we assume this result then

$$N(A) = X^{(\leq r+1)} \cup \partial^+ B,$$

so by Krustal-Katona, we'd take B to be an initial segment of lex on $X^{(r+1)}$. This strongly suggests that there is a ordering on $\mathcal{P}(X)$, where subsets are ordered by sizes and lex is used to decide order amongs subsets of equal size, such that the boundary minimising sets are the initial segments of this order.

Definition (simplicial order). The *simplicial order* on Q_n is defined by $x < y$ if either $|x| < |y|$ or $|x| = |y|$ and $x < y$ in lex.

Notation. In this chapter we use roman instead of curly upper case letters such as A, B to denote a family of subsets of X , and lower case letters such as x, y to denote element of X . This is compatible with the vertex convention in graph theory.

Our aim is to show that initial segments of simplicial are the best. As we have Krustal Katona at disposal, we would like to do similar “compressions” to move a set “closer” to Hamming ball. In the proof of Krustal Katona the compressions C_{UV} can be seen as operations of dimension $|U|$ in the discrete cube. Here we consider “codimension 1” compression.

Definition (section). Given $A \subseteq Q_n$ and $1 \leq i \leq n$, the *i -sections* are the set systems $A_+^{(i)}, A_-^{(i)} \subseteq \mathcal{P}(X - i)$, thought as Q_{n-1} , given by

$$A_+^{(i)} = \{x \in A : i \notin x\}$$

$$A_-^{(i)} = \{x - \{i\} : x \in A, i \in x\}$$

If it is clear which i is being considered we may omit the superscripts for simplicity.

Definition (i -compression). Define the *i -compression* $C_i(A)$ of A by compression on its i -sections, i.e. $C_i(A)_+^{(i)}$ is initial segment of Q_{n-1} of size $|A_+^{(i)}|$ and $C_i(A)_-^{(i)}$ is initial segment of Q_{n-1} of size $|A_-^{(i)}|$.

Example. (graph)

Certainly $|C_i(A)| = |A|$. Moreover, $C_i(A)$ “looks more like” the initial segment of simplicial than A did.

Say A is *i -compresses* if $C_i(A) = A$.

Theorem 2.1 (Harper). Let $A \subseteq Q_n$ and C be the initial segment of simplicial order with $|C| = |A|$. Then

$$|N(A)| \geq |N(C)|.$$

Same as in Krustal Katona, what we are really interested in is the “nice” size, i.e. if $|A| \geq \sum_{i=0}^r \binom{n}{i}$ then

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

Remark. If A is a Hamming ball then done by Krustal Katona. Conversely, Harper implies Krustal-Katona. Indeed, given $B \subseteq X^{(r)}$, apply Harper to $A = B \cup X^{(\leq r-1)}$.

Proof. Induction on n : $n = 1$ is trivial. Given $A \subseteq Q_n$ where $n > 1$, fix $1 \leq i \leq n$. Claim that

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

Proof. Write $B = C_i(A)$. Given an element in the neighbourhood of A , it is either in the downstairs part A_- or the upstairs part A_+ so

$$|N(A)| = |A_+ \cup N(A_-)| + |A_- \cup N(A_+)|$$

Similarly

$$|N(B)| = |B_+ \cup N(B_-)| + |B_- \cup N(B_+)|.$$

Now $|B_+| = |A_+|$ and $|N(B_-)| \leq |N(A_-)|$ by induction. But $N(B_-)$ is an initial segment of simplicial (on Q_{n-1}), as is B_+ , so $N(B_-)$ and B_+ are nested. Hence

$$|B_+ \cup N(B_-)| \leq |A_+ \cup N(A_-)|$$

and similarly

$$|B_- \cup N(B_+)| \leq |A_- \cup N(A_+)|$$

thus

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

□

Among all $B \subseteq Q_n$ with $|B| = |A|$ and $|N(B)| \leq |N(A)|$, choose one with $\sum_{x \in B} f(x)$ minimal, where $f(x)$ is position of x in simplicial ordering of Q_n . Then B is i -compressed for all i .

Must such B be an initial segment of simplicial? Unfortunately, no. For example $A = \{\emptyset, 1, 2, 12\} \subseteq Q_3$. However, we have

Lemma 2.2. *Let $B \subseteq Q_n$ be i -compressed for all i but not an initial segment of simplicial order. Then if n is odd, say $n = 2k + 1$, we have*

$$B = X^{(\leq k)} - \underbrace{\text{last } k\text{-set}}_{(k+2)(k+3)\dots(2k)(2k+1)} \cup \underbrace{\text{first } (k+1)\text{-set}}_{12\dots k(k+1)}$$

whereas if n is even, say $n = 2k$, we have

$$B = X^{(\leq k-1)} \cup \{x \in X^{(k)} : 1 \in x\} - \underbrace{\text{last } k\text{-set with } 1}_{1(k+2)(k+3)\dots(2k)} \cup \underbrace{\text{first } k\text{-set without } 1}_{234\dots k(k+1)}$$

Then done as in each case have $|N(B)| \geq |N(C)|$.

Proof. Have $x \notin B, y \in B$ for some $x < y$ in simplicial. Cannot have $i \in X, i \in y$ as B is i -compressed, and cannot have $i \notin x, i \notin y$, again as B is i -compressed. So for each i , i belongs to exactly one of x or y . Thus $y = x^c$.

Hence for each $y \in B$, at most one of $x < y$ has $x \notin B$ (namely y^c) and for each $x \notin B$, at most one of $y > x$ has $y \in B$ (namely x^c). Hence

$$B = \{z : z \leq y\} - \{x\}$$

where x is the immediate predecessor of y and $x = y^c$.

But then x is the last k -set (if $n = 2k + 1$) or last k -set containing 1 (if $n = 2k$), by definition of simplicial ordering. \square

\square

Remark.

1. It is also possible to prove Harper by UV -compression.
2. We can also use these “codimension 1” compressions to prove Krustal Katona directly.

In analysis, we often use isoperimetric inequality in blow-up form, e.g. give a disk, if we expand it a little bit by $\varepsilon > 0$ at some point then the resulting shape has perimeter at least as large as that of the disk.

Definition (t -neighbourhood). For $A \subseteq Q_n$, the t -neighbourhood of A is

$$A_{(t)} = N^t(A) = \{x \in Q_n : d(x, A) \leq t\}$$

Corollary 2.3. Let $A \subseteq Q_n$ with $|A| = |X^{(\leq r)}|$. Then for $1 \leq t \leq n - r$, have

$$|A_{(t)}| \geq |X^{(\leq r+t)}|.$$

Proof. Harper and induction. \square

To get a better for for what the corollary is saying, we’ll need some estimates and things like $\sum_{i=0}^r \binom{n}{i}$.

Proposition 2.4. Let $0 < \varepsilon < \frac{1}{4}$, then

$$\sum_{i=0}^{\lfloor (\frac{1}{2}-\varepsilon)n \rfloor} \binom{n}{i} < \frac{1}{\varepsilon} \exp(-\frac{\varepsilon^2 n}{2}) 2^n.$$

i.e. for fixed ε , the sum is an exponentially small fraction of 2^n

In standard deviation language, going $\sim \varepsilon$ standard deviates from mean

Proof. For $i \leq (\frac{1}{2} - \varepsilon)n$,

$$\frac{\binom{n}{i-1}}{\binom{n}{i}} = \frac{i}{n-i+1} \leq \frac{\frac{1}{2}-\varepsilon}{\frac{1}{2}+\varepsilon} = 1 - \frac{2\varepsilon}{\frac{1}{2}+\varepsilon} \leq 1 - 2\varepsilon.$$

Hence by comparison with geometric progression

$$\sum_{i=0}^{\lfloor (\frac{1}{2}-\varepsilon)n \rfloor} \binom{n}{i} \leq \frac{1}{2\varepsilon} \binom{n}{\lfloor (\frac{1}{2}-\varepsilon)n \rfloor}.$$

Similarly

$$\binom{n}{\lfloor (\frac{1}{2}-\varepsilon)n \rfloor} \leq (1-\varepsilon)^{\varepsilon n/2-1} \binom{n}{\lfloor (\frac{1}{2}-\frac{\varepsilon}{2})n \rfloor} \leq 2(1-\varepsilon)^{\varepsilon n/2} 2^n \leq 2 \exp -\varepsilon \varepsilon n / 2 \cdot 2^n$$

Therefore

$$\sum_{i=0}^{\lfloor (\frac{1}{2}-\varepsilon)n \rfloor} \leq \frac{1}{2\varepsilon} 2 \exp -\varepsilon^2 n / 22^n.$$

□

Missed a lecture on 06/11/18

Index

- UV*-compression, 8
- i*-compression, 15
- ij*-compression, 7

- antichain, 2

- chain, 2
- colexicographic order, 6

- Erdős-Ko-Rado theorem, 12

- Hamming ball, 14
- Harper's theorem, 15

- intersecting, 11

- isoperimetric inequality, 14

- Kruskal-Katona theorem, 9

- left-compressed, 8
- lexicographic order, 6
- local LYM, 4
- LYM inequality, 4

- section, 15
- set system, 2
- simplicial order, 15
- Sperner's lemma, 3

- upper shadow, 10