

Combinatorics

Adam Kelly (ak2316@cam.ac.uk)

October 14, 2021

This article constitutes my notes for the ‘Combinatorics’ course, held in Michaelmas 2021 at Cambridge. These notes are *not a transcription of the lectures*, and differ significantly in quite a few areas. Still, all lectured material should be covered.

Contents

1	Set Systems	1
1.1	Chains and Antichains	1
1.2	Shadows	4

§1 Set Systems

We will begin our study of combinatorics by considering *set systems* – collections of subsets of a set (which will typically be $X = [n] = \{1, 2, \dots, n\}$).

Definition 1.1 (Set Systems)

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

It’s often useful to think about the power set of a set X , $\mathcal{P}(X)$, as a graph. We can do this by joining two elements A and B if $|A \Delta B| = 1$, where Δ is the symmetric difference. This graph is the **discrete cube** Q_n ¹.

§1.1 Chains and Antichains

We are first going to look at what happens when sets are contained or not contained in each-other. If you know anything about posets, this will likely be familiar.

Definition 1.2 (Chain and Antichain)

We say that $\mathcal{A} \subset \mathcal{P}(X)$ is a **chain** if for all $A, B \in \mathcal{A}$ we have $A \subset B$ or $B \subset A$.

We say that \mathcal{A} is a **antichain** if for all $A, B \in \mathcal{A}$ with $A \neq B$ we have $A \not\subset B$ and $B \not\subset A$.

¹This is the same graph as the boolean hypercube, in the obvious way.

Example 1.3

$\{\{1, 4\}, \{1, 4, 7, 8\}, \{1, 2, 4, 7, 8\}\}$ is a chain, and $\{\{1, 4\}, \{1, 7, 8\}, \{32, 3, 8\}\}$ is an antichain.

A natural first question is how big can a chain be? We can easily get $|\mathcal{A}| = n - 1$ by taking

$$\mathcal{A} = \{\emptyset, \{1\}, \{1, 2\}, \dots, [n]\}.$$

Can we beat this? No, since \mathcal{A} must meet $X^{(r)}$ (the set of r element subsets of X) at at most one point.

How about antichains? We can achieve $|\mathcal{A}| = n$ by taking all singleton sets, but can we do any better? Well with the same idea we can take each $\lfloor n/2 \rfloor$ -element subset of $[n]$, giving $|\mathcal{A}| = \binom{n}{\lfloor n/2 \rfloor}$. Can we do better than *this*? It's not quite obvious (and it's this type of question that we will come across frequently in this course...).

Theorem 1.4 (Sperner's Lemma)

Let $\mathcal{A} \in \mathcal{P}(X)$ be an antichain. Then

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

and this is sharp.

Inspired by what made chains have size at most $n + 1$, we can try and decompose $\mathcal{P}(X)$ into $\binom{n}{\lfloor n/2 \rfloor}$ chains. If we can do this, we will be done as \mathcal{A} meets each in at most one point. We can try and do this decomposition explicitly, but it turns out that it's incredibly hard to do. Instead, it's sufficient to find:

1. For each $r < n/2$, a matching (set of disjoint edges) from $X^{(r)}$ to $X^{(r+1)}$,
2. For each $r > n/2$, a matching from $X^{(r)}$ to $X^{(r-1)}$.

We'd then be done as we can just put these matchings together to form chains, and we end up with the right number. By taking complements, it's enough to prove the first, and to prove there's a matching, we just need to check Hall's condition.

Proof of Sperner's Lemma. Consider the subgraph of Q_n spanned by $X^{(r)} \cup X^{(r+1)}$. This graph is bipartite, and $d(A) = n - r$ for all $A \in X^{(r)}$ and $d(A) = r + 1$ for all $A \in X^{(r+1)}$.

Given a set $S \subset X^{(r)}$, the number of edges between S and $\Gamma(S)$, the neighborhood of S , is $|S|(n - r)$ counting from below and at most $|\Gamma(S)|(r + 1)$ counting from above. So

$$|\Gamma(S)| \geq |S| \frac{n - r}{r + 1} \geq |S|,$$

since $r \leq (n - 1)/2$. Thus there is a matching by Hall's theorem from $X^{(r)}$ to $X^{(r+1)}$ which is sufficient. \square

Remark. While this proof does tell us about size, it doesn't tell us about uniqueness – which antichains have size $\binom{n}{\lfloor n/2 \rfloor}$?

It is possible to strengthen Sperner's lemma significantly.

We know that if we have an antichain, then if we add up all of the numbers of sets of each of $X^{(1)}, \dots, X^{(n)}$, we get at most $|X^{\lfloor n/2 \rfloor}|$. We can show if you add up the fraction of each of $X^{(1)}, \dots, X^{(n)}$ that the antichain occupies, you end up with at most 1. To prove this though, we will need to do a little bit more work.

We are going to introduce the concept of the *shadow* of a set system, and this will reoccur throughout the course.

Definition 1.5 (Shadow)

The **shadow** of $\mathcal{A} \subset X^{(r)}$ is the set system $\partial\mathcal{A} \subset X^{(r-1)}$ given by

$$\partial\mathcal{A} = \bigcup_{S \in \mathcal{A}} S^{(r-1)}.$$

For example if $\mathcal{A} = \{\{1, 2, 3\}, \{2, 3, 4\}, \{1, 4, 5\}, \{1, 2, 7\}\}$ then $\partial\mathcal{A} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{2, 4\}, \{3, 4\}, \{1, 5\}, \{4, 5\}, \{1, 7\}, \{2, 7\}\}$.

We can then prove that the shadow of \mathcal{A} occupies at least as great a fraction of it's 'layer' (with respect to fixed-sized subsets of X) as \mathcal{A} does.

Lemma 1.6 (Local LYM (Lubell, Yamamoto, Meshalyin) Inequality)

Let $1 \leq r \leq n$, and let $\mathcal{A} \subset X^{(r)}$, then

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

Proof. The number of edges between \mathcal{A} and $\partial\mathcal{A}$ is given by $|\mathcal{A}|r$, counting from above, and from below is at most $|\partial\mathcal{A}|(n-r+1)$. Thus

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

□

To have equality, we would need every $A \in \mathcal{A}$, $i \in A$ and $j \notin A$ to have $(A - \{i\}) \cup \{j\} \in \mathcal{A}$. Thus equality only happens when $\mathcal{A} = \emptyset$ or $\mathcal{A} = X^{(r)}$.

We can now use this lemma to prove our strengthening of Sperner's lemma.

Theorem 1.7 (LYM Inequality)

Let $\mathcal{A} \subset \mathcal{P}(X)$ be an antichain. Then

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

Proof 1 ('Bubble down by Local LYM'). For each r , write \mathcal{A}_r for $\mathcal{A} \cap X^{(r)}$.

First, $|\mathcal{A}_n|/\binom{n}{n} \leq 1$ trivially. Then $\partial\mathcal{A}_n$ and \mathcal{A}_n are disjoint, since \mathcal{A} is an antichain.

Thus

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

Thus by Local LYM, we get

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

Then $\partial(\partial\mathcal{A}_n \cup \mathcal{A}_{n-1})$ and \mathcal{A}_{n-2} are disjoint, so

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

and thus by Local LYM

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

so

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

We can continue on like this until we end up with the desired result. \square

To get equality, we would need equality to hold for each application of Local LYM, so for the greatest r with $\mathcal{A}_r \neq \emptyset$, we must have $\mathcal{A}_r = X^{(r)}$.

Another proof can be given that's a little bit more magic.

Alternative Proof of the LYM Inequality. Choose uniformly at random a maximal chain \mathcal{C} , that is, made up of sets $C_0 \subset C_1 \subset \dots \subset C_n$ where $|C_i| = i$ for each i . For a fixed r -set A , $\mathbb{P}(A \in \mathcal{C}) = 1/\binom{n}{r}$, so the chance that \mathcal{C} meets $\mathcal{A}_r = |\mathcal{A}_r|/\binom{n}{r}$, as the events are disjoint. So $\mathbb{P}(\mathcal{C} \text{ meets } \mathcal{A}) = \sum_{r=0}^n |\mathcal{A}_r|/\binom{n}{r}$, and $\mathbb{P}(\mathcal{C} \text{ meets } \mathcal{A}) = \sum_{r=0}^n |\mathcal{A}_r|/\binom{n}{r} \leq 1$. \square

§1.2 Shadows

Let $\mathcal{A} \subset X^{(r)}$ for $1 \leq r \leq n$. We know that $|\partial\mathcal{A}| \geq |\mathcal{A}| \frac{r}{n-r+1}$, but equality occurs only for $\mathcal{A} = \emptyset$ or $X^{(r)}$. What happens for \mathcal{A} not of those sizes? That is, given $|\mathcal{A}|$, how should we choose \mathcal{A} to minimize $|\partial\mathcal{A}|$?

It seems believable that if $|\mathcal{A}| = \binom{k}{r}$ then we should choose $\mathcal{A} = [k]^{(r)}$. This would give us $|\partial\mathcal{A}| = \binom{k}{r-1}$. Now what if $\binom{k}{r} < |\mathcal{A}| < \binom{k+1}{r}$? In this case, we might take \mathcal{A} with $[k]^{(r)} \subset \mathcal{A} \subset [k+1]^{(r)}$. For example, with $\mathcal{A} \subset X^{(3)}$ with $|\mathcal{A}| = \binom{7}{3} + \binom{4}{2}$, we could take $\mathcal{A} = [7]^{(3)} \cup \{A \cup 8 \mid A \in [4]^{(2)}\}$.

This suggests some kind of ordering on $X^{(r)}$...

Definition 1.8 (Ordering $X^{(r)}$)

Let A and B be distinct elements of $X^{(r)}$. Let $A = \{a_1, \dots, a_r\}$ and $B = \{b_1, \dots, b_r\}$, where $a_1 < \dots < a_r$ and $b_1 < \dots < b_r$.

We say that $A < B$ in the **lexicographic** or **lex** ordering if $a_i < b_i$ where i is the first position in which a_i and b_i differ.

We say that $A < B$ in the **colexicographic** or **colex** ordering if $a_i < b_i$ where i is the last position in which a_i and b_i differ.

Example 1.9 (Lexicographic Order)

We can list $^a [6]^{(3)}$ in lexicographic order as $\{123, 124, 125, 126, 134, 135, 136, 145, 146, 156, 234, 235, 236, 245, 246, 256, 345, 346, 356, 456\}$.

^aWe use a shorthand of $123 = \{1, 2, 3\}$.

Example 1.10 (Colexicographic Order)

We can list $[6]^{(3)}$ in colexicographic order as $\{123, 124, 134, 234, 125, 135, 235, 145, 245, 345, 126, 136, 236, 145, 245, 346, 156, 256, 356, 456\}$.

We note that in colex, $[k]^{(r)}$ is the initial segment of $[k+1]^{(r)}$. Indeed, one can view colex as giving an enumeration of $\mathbb{N}^{(r)}$ (which isn't true for lex).

We want to show that initial segments of colex minimize the shadow. That is, if $\mathcal{A} \subset X^{(r)}$, and \mathcal{C} is the initial segment of colex on $X^{(r)}$ with $|\mathcal{C}| = |\mathcal{A}|$, then $|\partial\mathcal{C}| \leq |\partial\mathcal{A}|$, and in particular $|\mathcal{A}| = \binom{k}{r}$ implies that $|\partial\mathcal{A}| \geq \binom{k}{r-1}$.