Combinatorics

Instructor: Jie Ma, Scribed by Jun Gao, Jialin He and Tianchi Yang

2020 Fall, USTC

1 Double Counting

The basic setting of the double counting technique is as follows. Suppose that we are given two finite sets A and B, and a subset $S \subseteq A \times B$. If $(a, b) \in S$, then we say that a and b are incident. Let N_a be the number of elements $b \in B$ such that $(a, b) \in S$, and N_b be the number of elements $a \in A$ such that $(a, b) \in S$. Then we have

$$\sum_{a \in A} N_a = |S| = \sum_{b \in B} N_b.$$

Theorem 1.1. *Let $T(j)$ be the number of divisions of a positive integer j. Let $\overline{T(n)} = \frac{1}{n}\sum_{j=1}^{n} T(j)$. Then we have $|\overline{T(n)} - H(n)| < 1$, where $H(n) = \sum_{i=1}^{n} \frac{1}{i}$ is the n^{th} Harmonic number.*

Proof. Define a table $X = (x_{ij})$ where

$$x_{ij} = \begin{cases} 1 & if\ i|j \\ 0 & otherwise. \end{cases}$$

Then

$$\sum_{j=1}^{n} T(j) = \sum_{1 \le i \le j \le n} x_{ij} = \sum_{i=1}^{n} \lfloor \frac{n}{i} \rfloor,$$

which implies that

$$\overline{T(n)} = \frac{1}{n} \sum_{i=1}^{n} \lfloor \frac{n}{i} \rfloor.$$

Then we have

$$|\overline{T(n)} - H(n)| < 1.$$

∎

Exercise 1.2. *Prove that*

$$\left| \frac{1}{n} \sum_{i=1}^{n} \lfloor \frac{n}{i} \rfloor - \sum_{i=1}^{n} \frac{1}{i} \right| < 1.$$

1.1 Sperner's Theorem

Definition 1.3. *Let $\mathcal{F} \subseteq 2^{[n]}$ be a family of subsets of [n]. We say $\mathcal{F}$ is independent (or $\mathcal{F}$ is an independent system), if for any two $A, B \in \mathcal{F}$, we have $A \not\subset B$ and $B \not\subset A$. In other words, $\mathcal{F}$ is independent if and only if there is no "containment" relationship between any two subsets of $\mathcal{F}$.*

Fact 1.4. *For a fixed $k \in [n]$, $\binom{[n]}{k}$ is an independent system.*

Theorem 1.5 (Sperner's Theorem). *For any independent system $\mathcal{F}$ of [n], we have*

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

First proof of Sperner's Theorem (Double-Counting). A *chain* of subsets of $[n]$ is a sequence of distinct subsets

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq ... \subseteq A_k.$$

A *maximal chain* is a chain with the property that no other subsets of [n] can be inserted into it to find a longer chain. We have the following observations.

(1). Any maximal chain looks like:

$$\phi \subseteq \{x_1\} \subseteq \{x_1, x_2\} \subseteq ... \subseteq \{x_1, ..., x_k\} \subseteq ... \subseteq \{x_1, ..., x_n\}.$$

(2). There are exactly $n!$ maximal chains.

This is because any such a maximal chain, say $\mathcal{C} : \phi \subseteq \{x_1\} \subseteq \{x_1, x_2\} \subseteq ... \subseteq \{x_1, x_2, ..., x_n\}$, defines a unique permutation:

$$\pi : [n] \to [n], \pi(i) = x_i, \forall i \in [n].$$

Now we double count the number of pairs $(\mathcal{C}, A)$ satisfying that:

- $\mathcal{C}$ is a maximal chain of [n].
- $A \in \mathcal{C} \cap \mathcal{F}$.

Recall the rule of double counting given at the beginning that

$$\sum_{\mathcal{C}} N_{\mathcal{C}} = \text{the number of pairs } (\mathcal{C}, A) = \sum_{A} N_A,$$

where $N_{\mathcal{C}}$ is the number of subsets $A \in \mathcal{C} \cap \mathcal{F}$ and N_A is the number of maximal chains $\mathcal{C}$ contains A. It is key to observe that

- $N_{\mathcal{C}} \leq 1$,
- $N_A = |A|!(n - |A|)!$

So we have

$$\begin{aligned} n! = \sum_{\mathcal{C}} 1 \geq \sum_{\mathcal{C}} N_{\mathcal{C}} = \sum_{A \in \mathcal{F}} N_A = \sum_{A \in \mathcal{F}} |A|!(n - |A|)! \\ = \sum_{A \in \mathcal{F}} \frac{n!}{\binom{n}{|A|}} \geq \sum_{A \in \mathcal{F}} \frac{n!}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \frac{n!}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} |\mathcal{F}|, \end{aligned}$$

which implies that

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

This finishes the proof. ∎

Now we give another proof of Sperner's Theorem.

Definition 1.6. *A chain is* <u>*symmetric*</u> *if it consists of subsets of sizes* $k, k+1, ..., \lfloor \frac{n}{2} \rfloor, ..., n-k-1, n-k$ *for some* $k \geq 0$.

For example, when $n = 3$, $\{\{2\}, \{2,3\}, \{1,2,3\}$ is not symmetric. And when $n = 4$, $\{\phi, \{1,2,3\}\}$ is not symmetric.

Theorem 1.7. *The family* $2^{[n]}$ *can be partitioned into a disjoint union of symmetric chains.*

Proof of Theorem 1.7. For each $A \in 2^{[n]}$, we define a sequence "$a_1 a_2 ... a_n$" consisting of left and right parentheses by defining

$$a_i = \begin{cases} \text{"(", if} \quad i \in A \\ \text{")", otherwise} \end{cases}$$

We then define the <u>"partial pairing of parentheses"</u> as following:

(1). First, we pair up all pairs "()" of adjoint parentheses.

(2). Then, we delete these already paired parentheses.

(3). Repeat the above process until nothing can be done.

Note that when this process stops, the remaining unpaired parentheses must look like this:

$$))))((((($$

We say two subsets $A, B \in 2^{[n]}$ have <u>the same partial pairing</u>, if the paired parentheses are the same (even in the same positions).

We can define an equivalence " $\sim$ " on $2^{[n]}$ by letting $A \sim B$ if and only if A, B have the same partial pairing.

Exercise 1.8. *Each equivalence class indeed forms a symmetric chain.*

Using this fact, now we see that $2^{[n]}$ can be partitioned into disjoint equivalence classes, which are disjoint symmetric chains. This fishes the proof. ∎

Theorem 1.7 can rapidly imply Sperner's Theorem.

Second proof of Sperner's Theorem. Note that by definition, any symmetric chain contains exactly one subset of size $\lfloor \frac{n}{2} \rfloor$. Since there are $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ many subsets of size $\lfloor \frac{n}{2} \rfloor$, by Theorem 1.7, we see that any partition of $2^{[n]}$ into symmetric chains has to consist of exactly $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ symmetric chains. Each symmetric chain can contain at most one subset from $|\mathcal{F}$ and thus we see $|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. ∎

1.2 Littlewood-Offord Problem

Theorem 1.9. *Fix a vector* $\vec{a} = (a_1, a_2, ..., a_n)$ *with each* $|a_i| \geq 1$*. Let* $S = \{\vec{\epsilon} = (\epsilon_1, \epsilon_2, ..., \epsilon_n) : \epsilon_i \in \{1, -1\}$ *and* $\vec{\epsilon} \cdot \vec{a} \in (-1, 1)\}$*, then* $|S| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$*.*

Remark: Note that this is tight for many vectors $\vec{a}$.

Proof. For any $\vec{\epsilon} \in S$, define $A_{\vec{\epsilon}} = \{i \in [n] : a_i\epsilon_i > 0\}$. Let $\mathcal{F} = \{A_{\vec{\epsilon}} : \vec{\epsilon} \in S\}$. Then we have

$$|S| = |\mathcal{F}|.$$

Now we claim that $\mathcal{F}$ is an independent system. Suppose for a contradiction that there exist $A_{\vec{\epsilon}_1}, A_{\vec{\epsilon}_2} \in \mathcal{F}$ with $A_{\vec{\epsilon}_1} \subseteq A_{\vec{\epsilon}_2}$. That also says,

$$\begin{cases} \vec{\epsilon}_1 \cdot \vec{a} \in (-1, 1) \\ \vec{\epsilon}_2 \cdot \vec{a} \in (-1, 1) \end{cases}$$

which imply that

$$|\epsilon_1 \cdot \vec{a} - \epsilon_2 \cdot \vec{a}| < 2.$$

By definition, we have

$$\vec{\epsilon}_1 \cdot \vec{a} = \sum_{i \in A_{\vec{\epsilon}_1}} |a_i| - \sum_{i \notin A_{\vec{\epsilon}_1}} |a_i| = 2 \sum_{i \in A_{\vec{\epsilon}_1}} |a_i| - \sum_{i=1}^{n} |a_i|.$$

Since $A_{\vec{\epsilon}_1} \subseteq A_{\vec{\epsilon}_2}$, we also have that

$$\vec{\epsilon}_2 \cdot \vec{a} - \vec{\epsilon}_1 \cdot \vec{a} = 2(\sum_{i \in A_{\vec{\epsilon}_2}} |a_i| - \sum_{j \in A_{\vec{\epsilon}_1}} |a_j|) \geq 2|a_j| \geq 2, \text{ for some } j \in A_{\vec{\epsilon}_2} \setminus A_{\vec{\epsilon}_1},$$

a contradiction. By Sperner's Theorem, we have $|S| = |\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. This finishes the proof ∎

1.3 Turán Type Problem

Definition 1.10. *A graph* G *is bipartite if its vertex set can be partitioned into two parts (say* A *and* B*) such that each edge joints one vertex in* A *and another in* B*.*

This is equivalent to say that $V(G)$ can be partitioned into two independent subsets. And we say (A, B) is a bipartition of G. For example, all even cycles C_{2k} are bipartite, while all odd cycles C_{2k+1} are not.

Definition 1.11. *Let* $K_{a,b}$ *be the complete bipartite graph with two parts of sizes* a *and* b*. This is a bipartite graph with edge set* $\{(i, j) : i \in A,\ j \in B\}$ *where* $|A| = a$ *and* $|B| = b$*.*

Definition 1.12. *Given a graph* H*, we say a graph* G *is* H*-free is* G *dose not contain a copy of* H *as its subgraph.*

For example, $K_{a,b}$ is K_3-free.

Definition 1.13. *For fixed graph* H*, let the Turán number of* H*, denoted by* $\text{ex}(n, H)$*, be the maximum number of edges in an* n*-vertex* H*-free graph* G*.*

Theorem 1.14. $\mathrm{ex}(n, C_4) \leqslant \frac{n}{4}(1+\sqrt{4n-3})$.

Proof. Let G be a C_4-free graph with n vertices. We need to show that $e(G) \leqslant \frac{n}{4}(1+\sqrt{4n-3})$. Consider $S = \{(\{u_1, u_2\}, w) : u_1wu_2 \text{ is a path of length 2 in } G\}$. Since G is C_4-free, for fixed $\{u_1, u_2\}$, there is at most one vertex w such that $(\{u_1, u_2\}, w) \in S$. So we have

$$|S| = \sum_{\{u_1,u_2\}} \text{the number of } (\{u_1, u_2\}, w) \in S \leqslant \sum_{\{u_1,u_2\}} 1 = \binom{n}{2}.$$

On the other hand, fixed a vertex w, the number of $\{u_1, u_2\}$ such that $(\{u_1, u_2\}, w) \in S$ exactly equals $\binom{d(w)}{2}$, which implies that

$$|S| = \sum_{w\in V(G)} \binom{d(w)}{2} = \frac{1}{2}\sum_{w\in V(G)} d^2(w) - e(G).$$

Putting the above together, we have

$$\binom{n}{2} \geq |S| = \frac{1}{2}\sum_{w\in V(G)} d^2(w) - e(G).$$

Using Cauchy-Schwarz inequality, we have

$$\frac{n^2-n}{2} \geq \frac{n}{2}\sum_{w\in V(G)} \frac{d^2(w)}{n} - e(G) \geq \frac{n}{2}\sum_{w\in V(G)} \left(\frac{d(w)}{n}\right)^2 - e(G),$$

which implies that

$$\frac{2e^2(G)}{n} - e(G) \leq \frac{n^2-n}{2}.$$

Solving it, we can derive easily that $e(G) \leq \frac{n}{4}(1+\sqrt{4n-3})$. ∎

Exercise 1.15. *Prove that* $\mathrm{ex}(n, C_4) < \frac{n}{4}(1+\sqrt{4n-3})$.

Corollary 1.16. *We have* $\mathrm{ex}(n, C_4) \leqslant (\frac{1}{2}+o(n))n^{\frac{3}{2}}$*, where* $o(n) \to 0$ *as* $n \to \infty$.

Theorem 1.17 (Mantal's Thm)**.** $\mathrm{ex}(n, K_3) = \lfloor \frac{n^2}{4} \rfloor$.

Proof. We first consider the lower bound $\mathrm{ex}(n, K_3) \geqslant \lfloor \frac{n^2}{4} \rfloor$ as the complete bipartite graph $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ is K_3-free and has $\lfloor \frac{n^2}{4} \rfloor$ edges.

Next, we show $\mathrm{ex}(n, K_3) \leqslant \lfloor \frac{n^2}{4} \rfloor$. We prove by induction on n that any n-vertex K_3-free graph G has at most $\frac{n^2}{4}$ edges. First it holds trivially when $n \in \{1, 2\}$. Now we assume that any K_3-free graph H with less than n vertices has at most $|V(H)|^2/4$ edges. Let G be K_3-free with n vertices. Take any edge of G, say $xy \in E(G)$. Since G is K_3-free, we say $N_G(x) \cap N_G(y) = \varnothing$, implies that $|d(x)| + |d(y)| \leqslant n$.

Let H be a graph obtained from G by deleting vertex x and y. Note that H is also K_3-free and has $n-2$ vertices. By induction, $e(H) \leqslant \frac{(n-2)^2}{4}$. Thus we have that

$$e(G) = e(H) + |d(x)| + |d(y)| - 1 \leqslant \frac{(n-2)^2}{4} + n - 1 = \frac{n^2}{4}.$$

This finishes the proof.

∎

Exercise 1.18. *The unique n-vertex K_3-graph which attains the maximum number of edges* $\mathrm{ex}(n, K_3)$ *is the complete bipartite graph* $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$.