Lecture 3

Combinatorial Methods (Winter 2023)
University of Toronto
Swastik Kopparty
Scribe: Yuchong Zhang and Amy Mann

1 Extremal Combinatorics: Intersecting Family

Extremal combinatorics studies how large (or small) a collection of finite objects can be, if it has to satisfy certain restrictions. Here we consider the example of intersecting families of sets. Recall that $[n] := \{1, 2, \ldots, n\}$. A family $\mathcal{F} \subset \mathcal{P}([n])$ is *intersecting* if for all $A, B \in \mathcal{F}$, $A \cap B \neq \emptyset$.

Examples

1. $\{1\}, \{1, 2\}, \{1, 2, 3\}, \ldots, \{1, 2, 3, \ldots n\}$ form an intersecting family of size n.
2. Let $\binom{[n]}{n-1}$ be the collection of all subsets of $[n]$ of size $n-1$. This is an intersecting family of size n.
3. Let $\binom{[n]}{>n/2}$ be the collection of all subsets of size $> n/2$. This is an intersecting family of size 2^{n-1} if n is odd, and size $\frac{1}{2}(2^n - \binom{n}{n/2})$ if n is even.
4. All subsets of $[n]$ that contain 1 form an intersecting family of size 2^{n-1}

In fact, the largest intersecting family $\mathcal{F} \subset \mathcal{P}([n])$ has size 2^{n-1}. To see this, pair up each set $A \subset [n]$ with its complement A^c. Clearly there are 2^{n-1} such pairs. By the pigeon hole principle, if $|\mathcal{F}| > 2^{n-1}$, then there must exist $A \subset [n]$ such that $A, A^c \in \mathcal{F}$. But then $\mathcal{F}$ would not be intersecting.

A related problem is the Erdös-Ko-Rado problem: given $n, k \in \mathbb{N}$, what is the size of the largest intersecting family $\mathcal{F} \subset \binom{[n]}{k}$?

If $k > \frac{n}{2}$, then clearly $\binom{[n]}{k}$ itself is an intersecting family, so $\mathcal{F}$ can have size $\binom{n}{k}$.

If $k \leq \frac{n}{2}$, then we can consider the collection of all size k subsets that contain a particular element, same as before in example 4. Such collection has size $\binom{n-1}{k-1}$. It turns out that in this case, $\binom{n-1}{k-1}$ is the largest size an intersecting family can have.

Theorem 1. *(EKR Theorem): For $k < \frac{n}{2}$, $\binom{n-1}{k-1}$ is the largest size for any intersecting family $\mathcal{F} \subset \binom{[n]}{k}$.*

A necklace made out of the letters $1, \ldots, n$ is a string that contains each letter exactly once and written along a circle in the clockwise direction. Two strings represent the same necklace if one can be obtained by rotating the other clockwise for some integer number of letters.

More formally, each such string can be identified with a permutation on $[n]$. From this perspective, we can define an equivalence relation on the permutation group S_n where $\sigma, \tau \in S_n$ are equivalent if there exists some $k \in \mathbb{N}$ such that

$$(1\ 2 \dots n)^k \sigma = \tau$$

and a necklace is simply an equivalence class. Since $(1\ 2 \dots n)$ has order n, it follows that the set of all necklaces has size $(n-1)!$. With this definition, we now prove the EKR theorem.

Proof. Let $\mathcal{F} \subset \binom{[n]}{k}$ be an intersecting family. Pick a necklace $a_1 \dots a_n$ uniformly at random. Consider the n length-k substrings of this necklace. View these substrings as sets $S_1, \dots, S_n$: $S_1 = \{a_1, a_2, \dots a_k\}$, $S_2 = \{a_2, a_3, \dots, a_{k+1}\}$, …, $S_n = \{a_n, a_1, \dots, a_{k-1}\}$. Let $Z = \{S_1, \dots, S_n\}$. Since $k < n/2$, for any $1 \le m \le n$, there can be at most k sets in Z (including S_m itself) that intersect S_m. Thus at most k of S_i's can be in $\mathcal{F}$. Consider the random variable $X = \sum_{A \in \mathcal{F}} \mathbb{1}[A \in Z]$. On one hand, clearly $X = |\mathcal{F} \cap Z|$, which we just showed is at most than k. So $\mathbb{E}X \le k$. On the other hand:

$$\begin{aligned} \mathbb{E}X &= \sum_{A \in \mathcal{F}} \mathbb{E}\mathbb{1}[A \in Z] \\ &= \sum_{A \in \mathcal{F}} \Pr(A \in Z) \quad \star \end{aligned}$$

Since the necklace is chosen uniformly at random, any set $A \in \binom{[n]}{k}$ has the same probability of being in Z. Since $|Z| = n$, we have:

$$\sum_{A \in \binom{[n]}{k}} \mathbb{1}[A \in Z] = n$$

and hence

$$\begin{aligned} \mathbb{E}\Big(\sum_{A \in \binom{[n]}{k}} \mathbb{1}[A \in Z] \Big) &= n \\ \sum_{A \in \binom{[n]}{k}} \mathbb{E}\mathbb{1}[A \in Z] &= n \\ \sum_{A \in \binom{[n]}{k}} \Pr(A \in Z) &= n \end{aligned}$$

which implies for any $A \in \binom{[n]}{k}$,

$$\Pr(A \in Z) = \frac{n}{\binom{n}{k}}$$

apply this to $\star$:

$$\begin{aligned} k \ge \mathbb{E}X = \sum_{A \in \mathcal{F}} \Pr(A \in Z) &= |\mathcal{F}| \cdot \frac{n}{\binom{n}{k}} \\ |\mathcal{F}| &\le \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1} \end{aligned}$$

□

2 The Ballot Theorem and Catalan numbers

The necklaces defined in the previous section has surprising applications to other combinatorial problems. Here we consider the Catalan numbers in the following setup: given the $n \times n$ grid in $\mathbb{R}^2$ with corners $(0,0),(0,n),(n,n),(n,0)$, we are interested in paths that go from $(0,0)$ to (n,n) under the restriction that in each step, we are only allowed to go right or up by 1 unit. Clearly there are $\binom{2n}{n}$ such paths in total because we have to make $2n$ steps, n of which must go up. But how many paths are there that do not go above the diagonal?

Fact 2. *If we take n steps where each step must go either up or right, then the probability of landing on the diagonal is*

$$\binom{n}{n/2} \approx \Theta(\frac{1}{\sqrt{n}2^n})$$

Fact 3. *The number of paths that do not go above the diagonal is the nth Catalan number*

$$\frac{1}{n+1}\binom{2n}{n}$$

We will show a proof of Fact 3 using the Ballot theorem, which involves necklaces made out of $+1$'s and -1's. For any such necklace, we say that a gem (meaning a letter on the circle) is *special* if the n partial sums that start there and goes in the clockwise direction are all positive.

Theorem 4. *(Ballot Theorem): Any necklace made of a +1's and b -1's has $a - b$ special gems.*

Proof. Clearly any -1 or any +1 that is immediately followed by a -1 is not a special gem. Moreover, removing a consecutive (+1,-1) pair does not change the set of special gems. This is because for any special gem, any of its partial sum that ends before the (+1, -1) pair is positive, and it stays positive after the removal. The other partial sums are also positive because they are positive prior to the removal and the pair sums to 0. We can then keep removing pairs of the form (+1, -1) and stop when either a (the number of 1's) or b (the number of -1's) reaches zero. If a reaches zero first, then in the original necklace, each +1 falls in some (+1, -1) pair, so there are no special gems. If b reaches zero first, then the number of 1's remaining is $a - b$ and they are all special gems since any -1 is cancelled out by its preceding +1 and thus has no effect on the partial sums. □

Catalan Numbers Via Ballot Theorem

Consider length 2n+1 necklaces with n+1 1's and n -1's. How many special gems are there? Every such necklace has one special gem. Let's view this as a walk. Write $a_1, a_2, \cdots, a_{2n}$. +1 corresponds to the right and -1 corresponds to up. Now we have a Catalan walk. We conclude that for every necklace there is a Catalan walk. So the number of Catalan walks $= \frac{1}{2n+1}\binom{2n+1}{n+1}$.

Remark: If the second last one is +1, we derive a contradiction (it must be special as well). Thus, we can say that the last step has to be +1, and the second last has to be -1.

3 Fermat's Little Theorem

It is a statement in number theory: P is prime, $a \in N$, and $a^p :- a$ (mod p). Consider the set of necklaces of length p whose gems are elements of 1,2,$\cdots$, a (i.e. [a]).

Q*: How many necklaces are there?

Q: How many strings are there? A: a^p.

Q: How many different necklaces does one string give? A: p.

Thus, we can answer Q*. $\sum_{\text{necklaces } b} deg(b) = a^p$.. deg(necklace b) = the number of different strings that we can get when we swap and write out the necklace. If necklace $b = b_1, \cdots, b_n$ has degree d, then $d|n$ and $b = (b_1 * \cdots * b_{n/2})^{n/d}$ (repeated n/d times). This converse is also true; this is an if and only if statement. The orbit of a string, σ, is $\frac{2p}{\text{stabilizer}(\sigma)}$. Some of the orbits are of size p. Some are size one.

Q: How many orbits are there of size 1? A: There are a orbits of size 1. There are m orbits of size m. $m * p + a = a^p \implies a^p = a$ (mod p).

Note: Action of Z_p on $[a]^p$ is $i * (\sigma_1\sigma_2 \cdots \sigma_p) = (\sigma_{i+1}\sigma_{i+2}) \cdots \sigma_{p+i})$.

There is a generalization of Fermat's Little Theorem: We have that the number of necklaces of size $p = a + \frac{a^p - a}{p}$. What is the formula for a general length, n? We will count the number of divisor of the degree n or less. How many necklaces are of size 2n+1 of made of n+1 1's and n -1's? It cannot be periodic (n+1 and 2n+1 are relatively prime). Every necklace has degree 2n+1. Suggestion: We should look at the distinct primes dividing n.

Number of necklaces of size n = $\sum_{d|n}$ number of necklaces of degree exactly d. Let u_d be the number of necklaces of degree exactly d. Note: a is fixed. Let v_d be the number of strings of degree dividing d coming from . $v_d = \sum_{k|d} u_k$ (k—d is all k's that divide d). We have a formula for v_d.

$$v_d = a^{n/d}$$

. This is a special system of equation related to the Mobius function. We will now spend time to understand the Mobius function.

Mobius inversion

For any partial order, there is a Mobius inversion formula. First, we define a function $\mu : \{1, 2, \cdots\} \to \{-1, 0, 1\}$. $\mu(p_1^{e_1} p_2^{e_2} \cdots p_n^{e_k}) = 0$ if any $e_i \geq 2, (-1)^k$ otherwise.

If f,g: $\{1, 2, 3, \cdots\} \to \mathbb{R}$ and g(n) = $\sum_{k|n} f(k)$. Then, f(z) = $\sum_{l|z} g(l) * \mu(\frac{z}{l})$.

Consider

$$\sum_{l|z} g(l)\mu(z/l) = \sum_{l/z} (\sum_{k/l} f(k)) * u(z/l)$$

$$= \sum_{k/l} (\sum_{i \text{ such that } k|l|z} \mu(z/l)) * f(k) = (*) f(z)$$

The last equality is a claim that we will now prove.

Lemma: For all m, $\sum_{l|m} \mu(l) = 1$ if m = 1 and 0 otherwise.

This is the uniquely defining property of μ.

Claim: From the lemma we want to show that for any $k|z$, $\sum_{l \text{ s.t. } k|l|z} \mu(z/l) = 1$ if k =2 and 0 otherwise. The Mobius function is a function that comes out this poset. So there is a general Mobius function for every poset.

This helps us because we now have that $u_z = \sum_{l|z} a^{n/l} \mu(z/l)$.

Thus, total number of necklaces of length n = $\sum_{z|n} u_z$.

Sanity check: Suppose n = p. Total number of necklaces = $u_1 + u_p$.

$$u_1 = 1/1a; u_p = \frac{a * p * 1 + a(-1)}{p} = \frac{a * p - a}{p}$$

This makes sense.

Let n = q*p, q,p prime.

Number of necklaces = $u_1 + u_p + u_q + u_{pq}$. We will get that

$$u_{pq} = 1/pq * (a^{pq} - a^p - a^p + a).$$