Soham Chatterjee
Email: soham.chatterjee@tifr.res.in
Course: Mathematical Foundations for Computer Sciences

Final Exam
Dept: STCS
December 6, 2024

> **Problem 1** Bollobás Inequality
>
> Let $m > 0$ and $A_1, \dots, A_m$ and $B_1, \dots, B_m$ be finite sets. Suppose that for all $i, j \in [m]$, we have $A_i \cap B_j = \emptyset \iff i = j$. Then
> $$\sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} \leq 1$$

Solution: Let $S = \left(\bigcup\limits_{i=1}^{n} A_i\right) \cup \left(\bigcup\limits_{i=1}^{n} B_i\right)$. Assume that $|S| = n$. Now $\binom{|A_i|+|B_i|}{|A_i|} = \frac{(|A_i|+|B_i|)!}{|A_i|!|B_i|!}$. Now $|A_i|!|B_i|!$ be the number of arrangements of the elements of $A_i \cup B_i$ so that first $|A_i|$ elements are form A_i and the last $|B_i|$ elements are from B_i. And $(|A_i| + |B_i|)!$ is the number of arrangements of the elements of $A_i \cup B_i$. Let E_i be the event that in a arrangement of the elements of S has first $|A_i|$ elements from A_i and last $|B_i|$ elements from B_i. Then

$$\mathbb{P}[E_i] = \frac{|A_i|!(n - |A_i| - |B_i|)!|B_i|!}{n!} = \frac{|A_i|!|B_i|!}{(|A_i| + |B_i|)!} = \binom{|A_i| + |B_i|}{|A_i|}^{-1} \qquad \forall\ i \in [m]$$

If we can show that E_i are pairwise disjoint events then we have

$$\sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} = \sum_{i=1}^{n} \mathbb{P}[E_i] = \mathbb{P}\left[\bigcup_{i=1}^{n} E_i\right] \leq 1$$

Hence it suffices to show that the events are pairwise disjoint. Take E_i and E_j where $i \neq j$. Suppose both E_i and E_i events occurred in a arrangement. Then the first $|A_i|$ elements are from A_i and the last $|B_i|$ elements are from B_i. Now $A_i \cap B_j \neq \emptyset$ and $A_j \cap B_i \neq \emptyset$. Then in the first $|A_i|$ elements there is an element from B_j. If $|A_i| \leq |A_j|$ then in the first $|A_j|$ elements there is an element from B_j which is not possible if E_j even occurred. So suppose $|A_i| > |A_j|$. In that case $A_j \cap B_i \neq \emptyset \implies$ there is an element of B_i in the first $|A_i|$ elements which is not possible since the E_i event occurred. Hence contradiction. Both E_i and E_j events cannot occur simultaneously. Therefore E_i and E_j are pairwise disjoint. Therefore $\sum\limits_{i=1}^{m} \binom{|A_i|+|B_i|}{|A_i|}^{-1} \leq 1$. ■

> **Problem 2**
>
> Solve the following parts.
>
> - Let $n > 0$ and S be a set of points on the n-dimensional hypercube such that $\Pr_{x \sim \{0,1\}^n}(x \in S) > \frac{2}{n}$. Show that S contains three points that form an equilateral triangle.
> - Let Sym_n be the symmetric group of permutations over [n]. Let $1 \leq a < b \leq n$. Find the size of the group generated by the permutations $(123 \dots n)$ and (ab).
> - There is a coop run by two brothers, Arjun and Bheem. Initially, the coop has two hens, one owned by Arjun and one owned by Bheem. Every month, a uniformly random hen in the coop lays two eggs that grow up to two more hens. These new hens are owned by the brother who owned the mother hen (and the mother hen can continue to lay eggs in the future). For $n, k \geq 0$, what is the probability that after n months, Arjun owns $2k + 1$ hens in the coop. By considering all possible values of k, derive an identity for 4^n.

Solution:

- Now two points in the hypercube are neighbors if they differ in one coordinate. Let x, y are two distinct points in the boolean hypercube. Suppose they have a common neighbor. Then x, y differ in two coordinates. Hence their distance is $\sqrt{2}$. Now Given that $Pr_{x \sim \{0,1\}^n}[x \in S] > \frac{2}{n}$. Then $|S| > \frac{2}{n}\, 2^n = \frac{2^{n+1}}{n}$.

  Now each point in the hypercube has n neighbors. Consider the set $\{(x,u)\colon x \in S, u \text{ neighbor of } x\}$. We have

  $$|\{(x,u)\colon x \in S, u \text{ neighbor of } x\}| = |S| \times n > \frac{2^{n+1}}{n} \times n = 2^{n+1}$$

  . But the hypercube has 2^n elements. Hence by pigeonhole principle there exists an element z which is neighbor of at least 3 elements of S. Let $x_1, x_2, x_3 \in S$ such that z is neighbor of x_1, x_2, x_3. Then distance between x_i, x_j is $\sqrt{2}$ for all $i, j \in [3]$ with $i \neq j$. Hence x_1, x_2, x_3 forms an equilateral triangle.

- Let $\phi = (12 \cdots n)$ and $\pi = (ab)$. Therefore $\phi^n = 1 \implies \phi^{-1} = \phi^{n-1}$. Since $b > a$ suppose $d = b - a$. Let $G = \langle \phi, \pi \rangle$.

  Lemma 1. *$(1, 1+kd) \in G$ for all $k \in \mathbb{N}$.*

  Proof: We will prove this using induction on k. For $k = 1$, $(1, 1+kd) = (1, 1+b-a)$. Now $(\phi^{-1})^{a-1}\pi\phi^{a-1} \in G$. Now for any $i \in [n]$, $\phi^{a-1}(i) = i + a - 1$ if $i + a - 1 \leq n$ and otherwise $\phi^{a-1}(i) = i + a - 1 - n$. Now if $i + a - 1 \neq a, b$ then $\pi(i+a-1) = i+a-1$ then $\pi(i+a-1-n) = i+a-1-n$. In both of these cases we have $(\phi^{-1})^{a-1}\pi\phi^{a-1}(i) = (\phi^{-1})^{a-1}\phi^{a-1}(i) = i$. Now if $i + a - 1 = a$ then $i = 1$ Then $\pi(a) = b$ Then

  $$(\phi^{-1})^{a-1}(b) = b - (a-1) = 1 + (b-a) = 1 + d$$

  If $i + a - 1 = b$ then $i = 1 + d$. Then $\pi(b) = a$. Then

  $$(\phi^{-1})^{a-1}(a) = a - (a-1) = 1$$

  Hence $(\phi^{-1})^{a-1}\pi\phi^{a-1} = (1, 1+d)$. Hence the base case follows.

  Suppose this is true for $1, \ldots, k-1$. Then $(1, 1+(k-1)d) \in G$. Then using Lemma 2 we have $(1 + (k-1)d, 1+kd) \in G$. Hence

  $$(1, 1+kd) = (1, 1+(k-1)d)(1+(k-1)d, 1+kd)(1, 1+(k-1)d)$$

  Hence $(1, 1+kd) \in G$. By mathematical induction $(1, 1+kd) \in G$ for all $k \in \mathbb{N}$ □

  Lemma 2. *For any $k \in \mathbb{N}$,*

  $$(1+kd, 1+(k+1)d) = \phi^{kd}(1, 1+d)(\phi^{-1})^{kd}$$

  Proof: For any $i \in [n]$, $(\phi^{-1})^{kd}(i) = i - kd$. Now if $i - kd \neq 1, 1+d$ then $\pi(i - kd) = i - kd$. Then $\phi^{kd}\pi(\phi^{-1})^{kd}(i) = \phi^{kd}(\phi^{-1})^{kd}(i) = i$. Now suppose $i - kd = 1$. Then $i = 1 + kd$. Then $\phi(i - kd) = 1 + kd$. Then $\phi^{kd}(1+kd) = 1 + kd - kd = 1$. If $i - kd = 1 + d$ then $i = 1 + (k+1)d$. Then $\pi(1+d) = 1$. Then $\phi^{kd}(1) = 1 + kd$. Therefore $\phi^{kd}(1, 1+d)(\phi^{-1})^{kd} = (1+kd, 1+(k+1)d)$ □

  Now if $gcd(n,d) = 1$ then there exists $a, b \in \mathbb{Z}$ where $a > 0$ such that $ad + bn = 1$. Therefore $bn = 1 - ad$. Therefore $(1, 1+ad) = (1, 2)$. And for any $k \in [n]$, $(1, k) = (1, 1+(k-1)ad)$. Hence for all $k \in [n]$, $(1, k) \in G$. Hence for any $i, j \in [n]$, $i \neq j$, $(i, j) = (1, i)(1, j)(1, i)$. Hence all transpositions are in G. Hence G contains all the permutations of Sym_n. Hence Sym_n is generated by $(12 \cdots n)$ and (ab).

  Now if $gcd(n,d) \neq 1$. Suppose $gcd(n,d) = l$. Then there exists $a, b \in \mathbb{Z}$ with $a > 0$ such that $ad + bn = l$. Then for any $r \in [l]$, $(r, r+kl) \in G$ for any $k \in \mathbb{N}$. Therefore any permutation of the elements

■

[I discussed the solution with Aakash and Shubham]

> **Problem 3**
>
> In this question, we will show that $(n-1)! \equiv -1 \pmod{n}$ if and only if n is prime
>
> - Show that, if G is a finite Abelian group, then the order of any element in G divides the maximal order of all elements.
> - For all $n > 0$, define the set $G_n = \{i \in [n] \mid GCD(i, n) = 1\}$. Show that G_n together with the multiplication operator modulo n forms a finite Abelian Group.
> - If n is prime, then show that $\prod_{0<i<j<n}(j-i) \not\equiv 0 \pmod{n}$. Using a Vandermonde matrix, conclude that G_n must be cyclic.
> - Derive that $(n-1)! \equiv -1 \pmod{n}$ if and only if n is prime.

Solution:

- Let g be the element of G which has the maximum order of all elements. Let r_g be the order of g. Now let h be any element. Assume r_h be the order of h. Suppose $r_h \nmid r_g$. Let r is the lcm of r_g and r_h. Then consider the element gh. Now for any $n \in \mathbb{N}$, $(gh)^n = g^n h^n$. For any $n < r$, either $n \not\equiv 0 \pmod{r_g}$ or $n \not\equiv 0 \pmod{r_h}$. Hence in both cases $g^n h^n \neq e$ where e is the identity element of G. But $(gh)^r = g^r h^r = e \cdot e = e$. Therefore the order of the element is r. Since $r_h \nmid r_g$, $r > r_g, r_h$. So we get a new element which has order more than r_g which contradicts that g has the maximum order. Hence contradiction. $r_h \mid r_g$. Since h is an arbitrary element of G, the order of any element of G divides the maximal order of all elements in G.

-
  - **Closure:** Let $a, b \in G_n$. Then $gcd(a, n) = 1$ and $gcd(b, n) = 1$. Let $gcd(ab, n) = d$. If $d \neq 1$ then there exists d_a and d_b such that $d_a \mid a$ and $d_b \mid b$ nand $d_a d_b = d$. Since $d \neq 1$ both d_a and d_b can not be equal to 1. Hence one of them is at least not equal to 1. WLOG suppose $d_a \neq 1$. Then $d_a \mid a$ and $d_a \mid n$ which contradicts $gcd(a, n) = 1$. Hence $gcd(ab, n) = 1$. So $ab \in G_n$.
  - **Associativity:** Let $a, b, c \in G_n$. Now multiplication of integers follows associativity. Therefore $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. Therefore then going modulo n we get the same values i.e. $(a \cdot b) \cdot c \pmod{n} = a \cdot (b \cdot c) \pmod{n}$. Therefore $(a \cdot b) \cdot c \equiv a \cdot (b \cdot c) \pmod{n}$. Hence the group operation follows associativity.
  - **Identity:** Consider the element 1. Now for any $a \in [n]$ we have $1 \cdot a = a \cdot 1 = a$. Therefore $1 \cdot a \equiv a \cdot 1 \equiv a \pmod{n}$.
  - **Inverse:** Let $a \in G_n$. Now consider the set elements $S = \{k \cdot a \mid k \in G_n\}$. Now in this set S there are $\phi(n)$ elements where ϕ is the Euler Totient Function. Now consider the set $S_{\text{mod } n} = \{x \pmod{n} \mid x \in S\}$. Now for any $n \in \mathbb{N}$, $n \cdot a = \sum\limits_{i=1}^{n} a$. Since G_n is closed under operation, $S_{\text{mod } n} \subseteq G_n$. Now $S_{\text{mod } n}$ has at most $\phi(n)$ elements. Now let there exists two $i, j \in [n-1]$ where $i \neq j$ and $j > i$ such that $i \cdot a \equiv j \cdot a \pmod{n} \implies (i-j) \cdot a \equiv 0 \pmod{n}$. Therefore $(i-j) \cdot a \mid n$. Hence $a \mid n$. But since $a \in G_n$ we know $gcd(a, n) = 1$. Hence contradiction. Therefore such i, j does not exists. Hence $|S_{\text{mod } n}| = \phi(n)$. Therefore $S_{\text{mod } n} = G_n$. Therefore there exists $k \in G_n$ such that $k \cdot a \equiv 1 \pmod{n}$. Therefore there exists inverse of a.
  - **Commutativity:** Let $a, b \in G_n$. Now we know multiplication of integers follows commutativity. Therefore $a \cdot b = b \cdot a$. Therefore going modulo n we have $a \cdot b \bmod n = b \cdot a \bmod n \implies a \cdot b \equiv b \cdot a \pmod{n}$.

  Therefore G_n is a finite Abelian group.

- Since n is prime all the positive integers less than n are in G_n. Hence for all $i, j \in [n-1]$ where $i < j$ and $i \neq j - 1$ we have $j - i \nmid n$ otherwise $j - i$ will be a nontrivial factor of n. Therefore we have $\prod\limits_{0<i<j<n} (j - i) = \prod\limits_{0<i<j<n, i<j-1} (j - i)(j - i) \not\equiv 0 \pmod{n}$.

Since n is prime all the positive integers less than n are in G_n So consider the matrix $M = (m_{i,j})_{1 \le i,j \le n-1}$ where $m_{i,j} = i^{j-1}$.

Lemma 3. *If*

$$M = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1^{n-3} & a_2^{n-3} & \cdots & a_{n-2}^{n-3} & a_{n-1}^{n-3} \\ a_1^{n-2} & a_2^{n-2} & \cdots & a_{n-2}^{n-2} & a_{n-1}^{n-2} \end{bmatrix} \implies \det M = \prod_{0<i<j<n} (a_j - a_i)$$

Proof: We will prove this using induction on n. For $n = 1$ this is true. So the base case follows. Suppose this is true for $n - 1$. We will show for n. Now consider the matrix M' where er subtracted the $i - 1$ times the j^{th} row from $(j+1)^{th}$ row and we

$$M' = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & a_2 - a_1 & \cdots & a_{n-2} - a_1 & a_{n-1} - a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_2^{n-2} - a_1 a_2^{n-3} & \cdots & a_{n-2}^{n} - a_1 a_{n-1}^{n-3} & a_{n-1}^{n-2} - a_1 a_{n-1}^{n-3} \end{bmatrix}$$

Therefore

$$\det M = \det M' = \prod_{j=2}^{n-1} (a_j - a_1) \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_2 & a_3 & \cdots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_2^{n-2} & a_3^{n-2} & \cdots & a_{n-1}^{n-2} \end{bmatrix}$$

Hence by inductive hypothesis we have

$$\det M = \prod_{j=2}^{n-1} (a_j - a_1) \prod_{1<i<j<n} (a_j - a_i) = \prod_{0<i<j<n} (a_j - a_i)$$

Hence by mathematical induction this is true for all n. □

Therefore using the lemma we have $\det M = \prod\limits_{0<i<j<n} (j - i)$. Now we know $\det M \neq 0 \pmod{n}$ if n is prime. Therefore M is invertible.

By the first part we know order of any element of G_n divides the maximal order of all elements. Now let g be the element in G_n which has the maximal order. Consider the vector $\hat{g} = \begin{bmatrix} 1 & g & g^2 & \cdots & g^{n-2} \end{bmatrix}$. Now $\hat{g}$ is a column of M. Since M is invertible $1, g, \ldots, g^{n-2}$ are linearly independent. Therefore order of g is $n - 1$. Hence $1, g, \ldots, g^{n-2}$ are all elements of G_n. But G_n has $n - 1$ elements. Therefore g is the generator of G_n. Hence G_n is cyclic.

- Let n is a prime. Then $(n-1)! = \prod\limits_{g \in G_n} g$. Now for all elements $g \in G_n$ there exists $g^{-1} \in G_n$. If $g \neq g^{-1}$ then their product in $\prod\limits_{g \in G_n} g$ vanishes. Therefore only the elements whose product actually contributes to the value of $\prod\limits_{g \in G_n} g$ are g's in G_n such that $g = g^{-1} \iff g^2 = 1$. Therefore

$$\prod_{g \in G_n} g = \prod_{\substack{g \in G_n \\ g^2 = 1}} g$$

Now for any element $g \in G_n$

$$g^2 = 1 \implies g^2 - 1 \equiv 0 \pmod{n} \implies (g+1)(g-1) \equiv 0 \pmod{n}$$

Therefore either $g + 1 \equiv 0 \pmod{n} \implies g = n - 1$ or $g - 1 \equiv 0 \pmod{n} \implies g = 1$. Therefore

$$\prod_{g \in G_n} g = (n-1) \cdot 1 = n - 1 \equiv -1 \pmod{n} \implies (n-1)! \equiv -1 \pmod{n}$$

Now suppose n is not prime. Then there exists $1 < a, b < n$ such that $a \cdot b = n$. Then

$$(n-1)! = a \cdot b \cdot \prod_{i=1, i \neq a,b}^{n-1} i = n \prod_{i=1, i \neq a,b}^{n-1} i$$

Hence $n \mid (n-1)! \implies (n-1)! \equiv 0 \pmod{n}$.

Therefore $(n-1)! \equiv -1 \pmod{n}$ if and only if n is a prime.

■

Problem 5

Let Ω be a set and Sym_Ω be the set of permutations on Ω. We say that G acts on Ω if there is a homomorphism θ from G to Sym_Ω. For all $x \in \Omega$ and an action θ, we define the orbit of x under θ to be the set $\text{orb}(x) = \{\theta(g)(x) \mid g \in G\}$ and the stabilizer of x under θ to be $\text{stab}(x) = \{g \in G \mid \theta(g)(x) = x\}$. In the following, assume that G and Ω are finite.

- Show that $|\text{orb}(x)| \cdot |\text{stab}(x)| = |G|$.
- Let $x, y \in \Omega$ be related if $y \in \text{orb}(x)$. Show that this is an equivalence relation.
- Let $\text{fix}(g) = |\{x \in \Omega \mid g \in \text{stab}(x)\}|$. Show that the number of different orbits is $\mathbb{E}_{g \sim G}[|\text{fix}(g)|]$.
- Using the above, compute the number of ways to color the vertices of a regular tetrahedron using k colors.

Solution:

- We will first prove that $\text{stab}(x)$ is a subgroup of G for any $x \in \Omega$. Now first of all $\text{stab}(x)$ contains the identity of element of G. Therefore $\text{stab}(x) \neq \emptyset$. Now let $g, h \in \text{stab}(x)$. Then $\theta(g)(x) = x$ and $\theta(h)(x) = x$. Then

$$\theta(gh)(x) = \theta(g) \cdot \theta(h)(x) = \theta(g)\left(\theta(h)(x)\right) = \theta(g)(x) = x$$

Therefore $gh \in \text{stab}(x)$. Hence $\text{stab}(x)$ is a subgroup of G. Now consider the set of cosets of $\text{stab}(x)$ in G be $G/\text{stab}(x)$. By Lagrange Theorem $|G/\text{stab}(x)| = \frac{|G|}{|\text{stab}(x)|}$. Hence it is enough to show that $|G/\text{stab}(x)| = |\text{orb}(x)|$. Now let $g \cdot \text{stab}(x)$ be a coset of $G/\text{stab}(x)$. Then any element of $g \cdot \text{stab}(x)$ is of the form gh where $h \in \text{stab}(x)$. Now consider the function $f : G/\text{stab}(x) \to \text{orb}(x)$ where $f(g \cdot \text{stab}(x)) = \theta(g)(x)$. Now for all $\theta(gh)(x) = \theta(g)\theta(h)(x) = \theta(g)(x)$. Therefore for any two elements $h, h' \in g \cdot \text{stab}(x)$, $g(g \cdot \text{stab}(x)) = \theta(g)(x) = \theta(h)(x) = \theta(h')(x)$. So f is well defined. Now the for all $y \in \text{orb}(x)$, $\exists\, g \in G$ such that $\theta(g)(x) = y$. So in $G/\text{stab}(x)$ consider the coset $g \cdot \text{stab}(x)$. Then $f(g \cdot \text{stab}(x)) = y$. Therefore f is surjective. Now let there exists $g \cdot \text{stab}(x), g' \cdot \text{stab}(x) \in G/\text{stab}(x)$ such that $f(g \cdot \text{stab}(x)) = f(g' \cdot \text{stab}(x)) = y$ for some $y \in \text{orb}(x)$. In that case we have

$$\theta(g^{-1} \cdot g')(x) = \theta(g^{-1})\theta(g')(x) = \theta(g^{-1})(y) = x$$

Hence $g^{-1} \cdot g \in \text{stab}(x)$. Therefore $g \cdot (g^{-1} \cdot g') = g'$. Hence $g' \in g \cdot \text{stab}(x)$. Therefore all the elements of G for which sends x to same destination are in same equivalence class in $G/\text{stab}(x)$. Therefore f is injective. Hence f is a bijection. Therefore $|G/\text{stab}(x)| = |\text{orb}(x)|$. Hence we get $|\text{orb}(x)| \cdot |\text{stab}(x)| = |G|$.

- $x, y \in \Omega$ are related if $y \in \operatorname{orb}(x)$ i.e. $\exists\ g \in G$ such that $\theta(g)(x) = y$. Now for any $x \in \Omega$, x is related to x since the identity element of G, $\theta(1)(x) = x$. Hence the relation is reflexive. Now let $x.y \in Om$ are related. then $\exists\ \ g \in G$ such that $\theta(g)(x) = y$. Then $\theta(g^{-1})(y) = x$. Therefore the relation is symmetric. Now suppose $x, y, z \in \Omega$ and x, y are related, y, z are related. Hence there exists $g, h \in G$ such that $\theta(g)(x) = y$ and $\theta(h)(y) = z$. Then $\theta(hg)(x) = \theta(h)\theta(g)(x) = \theta(h)(y) = z$. Hence x, z are also related. Therefore the relation is transitive. Hence the relation is an equivalence relation.

- Let Γ denote the set of different orbits.

$$\begin{aligned}
\frac{1}{|G|}\sum_{g\in G}|\operatorname{fix}(g)| &= \frac{1}{|G|}\sum_{g\in G}|\{x \in \Omega : gx = x\}| \\
&= \frac{1}{|G|}|\{(g, x)\colon g \in G, x \in \Omega, g \cdot x = x\}| \\
&= \frac{1}{|G|}\sum_{x\in\Omega}|\{g \in G : gx = x\}| \\
&= \frac{1}{|G|}\sum_{x\in\Omega}|\operatorname{stab}(x)| \\
&= \frac{1}{|G|}\sum_{x\in\Omega}\frac{|G|}{|\operatorname{orb}(x)|} \\
&= \sum_{x\in\Omega}\frac{1}{|\operatorname{orb}(x)|} \\
&= \sum_{\operatorname{orb}(x)\in\Gamma}\left(\sum_{y\in\operatorname{orb}(y)}\frac{1}{|\operatorname{orb}(x)|}\right) \\
&= \sum_{\operatorname{orb}(x)\in\Gamma}\left(\sum_{y\in\operatorname{orb}(x)}\frac{1}{|\operatorname{orb}(x)|}\right) \\
&= \sum_{\operatorname{orb}(x)\in\Gamma} 1 \\
&= |\Gamma|
\end{aligned}$$

- The rotational symmetries of a regular tetrahedrons are:
  - Identity Rotation: It leaves all the vertices fixed. So there is only one such symmetry.
  - 120° or 240° rotation about an axis which passes through a vertex and the center of the opposite face. Each axis has 2 rotation choices. And there are 4 choices for the vertex. Hence there are 8 such symmetries.
  - 180° rotation about an axis through midpoints of opposite edges. For each axis there is only one rotation. There are 3 such axes. Hence there are 3 such symmetries.

  Hence total number of symmetries $1+8+3 = 12$. A_4 be the set of permutations of S_4 which contains even permutations. Since half of the permutations of S_4 are even permutations we have $|A_4| = 12$. Let $g, h \in A_4$. Then g and h can be broken into even number product of transpositions. Let g can be written as product of $2k$ transpositions and h can be written as product of $2l$ transpositions. Therefore gh is also even permutation since gh can be written as product $2(k + l)$ transpositions where the first $2k$ are transpositions from g and the last $2l$ are transpositions from h. Hence $gh \in A_4$. Therefore A_4 is a subgroup of S_4.

  Lemma 4. *Symmetry group of a regular tetrahedron is A_4*

  Proof: We first enumerate the vertices of the regular tetrahedron by [4]. The identity rotation is basically fixing all the vertices. Therefore the identity element of A_4 corresponds to the identity rotation.

Now consider the 120° or 240° rotations about an axis which passes through a vertex and the center of the opposite face. Now in A_4 all the 3-length cycles corresponds to such rotations. Because for any 3-cycle we first fix an element. WLOG suppose we fix 4. Then we permute the rest of the elements. So the only options are (123) and (132). Now the permutation (123) corresponds to 120° rotation about the axis passing through the vertex 4 and middle point of the triangle generated by the vertices 1,2,3. Similarly (123) corresponds to 240° rotation about the axis passing through the vertex 4 and middle point of the triangle generated by the vertices 1,2,3. Hence each of the 120° or 240° rotations symmetries corresponds to each of the 3-cycles of A_4.

For 180° rotation about an axis through midpoints of opposite edges if one of the edge is between $a, b \in [4]$, $a \neq b$ then the other edge is between $[4] - \{a, b\}$ since they are opposite edge. Let $\{c, d\} = [4] - \{a, b\}$. Now rotating $180-^\circ$ about the axis through the middle points of the edges ab and cd is basically permuting $\{a, b\}$ and $\{c, d\}$. Therefore $180-^\circ$ about the axis through the middle points of the edges ab and cd corresponds to the permutation $(ab)(cd)$. There are only one such symmetry for the axis through the middle point of edges ab and cd. Now there are 3 such permutations in A_4. And each of them corresponds to 180° rotation about an axis through midpoints of opposite edges.

Hence there is a one-one correspondence between rotational symmetries of regular tetrahedron and the group A_4. Hence the symmetry group of a regular tetrahedron is A_4. □

Now we will calculate $|\text{fix}(g)|$ for each $g \in A_4$. There are k colors.

- Identity: Since this fixes all the vertices all k^4 coloring are fixed under the identity.
- 120° or 240° rotation about an axis which passes through a vertex and the center of the opposite face: One vertex is fixed and the other 3 vertices are cycled. A coloring is invariant if the 3 cycled vertices are of same color and the 4^{th} vertex can be any color. Hence the number of fixed colorings is k^2.
- 180° rotation about an axis through midpoints of opposite edges: Here it swaps two disjoint pairs of vertices. Hence a coloring is invariant if each pair has the same color. Therefore the number of fixed colorings is k^2¿.

Now applying result in previous part we have the number of distinct colorings is $\frac{1}{12}(k^4 + 8k^2 + 3k^2) = \frac{k^4+11k^2}{12}$

■

Problem 7

Let $n > 0$ and M be an $n \times n$ symmetric matrix with positive entries. Let λ be the largest eigenvalue of M. Show that

- $\lambda > 0$.
- There exists an eigenvector for λ with positive entries.
- λ has multiplicity 1.
- For all other eigenvalues λ', we have $\lambda > |\lambda'|$.

Solution:

- Entries of M are positive. Hence the trace of M, $tr(M) > 0$. Now $tr(M)$ is the sum of eigenvalues of M. Hence there must exists an eigenvalue of M which is positive since $tr(M) > 0$. Hence the $\lambda > 0$ since λ is the largest eigenvalue.

- Since M is a real symmetric matrix. It is diagonalizable over $\mathbb{R}$. Therefore all the eigenvalues of M are real and the corresponding eigenvectors of M are real. Let x, y are eigenvectors for distinct eigenvalues λ_1 and λ_2. Then

  $$\lambda_2 y^T x = (\lambda_2 y)^T x = (My)^T x = y^T Mx = y^T(\lambda_1 x) = \lambda_1 y^T x \implies (\lambda_1 - \lambda_2) y^T x = 0 \implies y^T x = 0$$

  Hence eigenvectors for distinct eigenvalues are orthogonal to each other. Therefore using Gram Schmidt we obtain orthonormal eigenbasis within same eigenspace. Therefore we get a real orthornormal eigenbasis for real symmetric matrix.

  Therefore there is an real orthonormal eigenbasis $\{v_i : i \in [n]\}$ with real eigenvalues $\{\lambda_i : i \in [n]\}$ such that $M = \sum\limits_{i=1}^{n} \lambda_i v_i v_i^T$. Let $v \in \mathbb{R}^n$ where all the entries of v are nonnegative and $\|v\| = 1$. Then there exists unique $b_i \in \mathbb{R}$ such that $v = \sum_{i=1}^n b_i v_i$. Then we have

  $$v^T M v = v^T \left(\sum_{i=1}^{n} \lambda_i v_i v_i^T\right)\left(\sum_{i=1}^{n} b_i v_i\right) = \left(\sum_{i=1}^{n} b_i^* v_i^T\right) \sum_{i=1}^{n} \lambda_i b_i v_i = \sum_{i=1}^{n} \lambda_i |b_i|^2$$

  Therefore

  $$|v^T M v| \leq \sum_{i=1}^{n} |\lambda_i||b_i|^2 \leq \lambda \sum_{i=1}^{n} |b_i|^2 = \lambda$$

  Now if v is the eigenvector with eigenvalue λ then $v^T M = \lambda v^T v = \lambda$. Hence the equality holds. Let v is not a eigenvector with eigenvalue λ. Suppose the first k eigenvectors are of eigenvalue λ. Then $|b_i|^2 < 1$ for all $i \in [k]$. Since each $|b_i|^2 \in [0, 1]$, $\sum_{i=1}^n \lambda_i |b_i|^2$ is a convex combination of the eigenvalues of M. Since λ is the largest eigenvalue the largest value the convex combinations of the eigenvalues can take is λ. Since $|b_1|^2 < 1$ the value taken by convex combination is less than λ. Hence $v^T M v < \lambda$. Therefore $v^T M v = \lambda$ if and only if v is an eigenvector for eigenvalue λ.

  Since λ is an eigenvalue of M let w be the normalized real eigenvector of M with corresponding eigen value λ. Let $M = (m_{i,j})_{1 \leq i,j \leq n}$. Now define the vector x, where $x_i = |w_i|$. Now

  $$0 < \lambda = \sum_{i,j} m_{i,j} w_i w_j = \left|\sum_{i,j} m_{i,j} w_i w_j\right| \leq \sum_{i,j} m_{i,j} |w_i||w_j| = \sum_{i,j} m_{i,j} x_i x_j \leq \lambda$$

  Therefore x is an eigenvector with eigenvalue λ. Hence we have $Mx = \lambda x$ and x has all entries nonnegative. Now if for any $i \in [n]$, $x_i \neq 0$ since $\lambda x_i = \sum_{j=1}^n a_{i,j} x_j$ and at least one $x_j \neq 0$. Hence all the entries of x are positive.

- Suppose λ has multiplicity more than 1. Then there exists two real orthonormal eigenvectors vectors u, v for eigenvalue λ. Let $\exists\ i \in [n]$ such that $u_i < 0$. Then

  $$0 = \lambda(u_i + |u_i|) = \sum_{j=1}^{n} a_{i,j}(u_j + |u_j|)$$

  Since $u_j + |u_j| \geq 0$ for all $j \in [n]$ and $a_{i,j} > 0$ for all $i, j \in [n]$ we conclude $u_j + |u_j| = 0$ for all $j \in [n]$. Therefore wither $u_j = |u_j|$ for all $j \in [n]$ or $u_j = -|u_j|$ for all $j \in [n]$. Similarly we obtain the same for v and get $v_j = |v_j|$ for all $j \in [n]$ or $v_j = -|v_j|$ for all $j \in [n]$. Then

  $$v^T u = \sum_{j=1}^{n} v_j u_j = \pm \sum_{j=1}^{n} |v_j u_j| \neq 0$$

  Therefore u, v are not orthogonal. Hence contradiction. Therefore λ has multiplicity 1.

- Let λ' be any other eigenvalue. So $\lambda > \lambda'$. Let u be the normalized eigenvector of λ'. Then we have $\lambda > |u^T M u| = |\lambda' u^T u| = |\lambda'|$. Hence $\lambda > |\lambda'|$ for any other eigenvalue.

■

[I discussed with Shubham and Soumyadeep and Aakash]

Problem 8

Let $0 < n \ll k$ and $x_1, \ldots, x_k$ be points in $\mathbb{R}^n$ and $v_1, \ldots, v_k \in \mathbb{R}$ be corresponding values. In many learning settings, the goal is to find the smallest simple function, say a low degree polynomial, that best approximates the given values. For example, one can try to find the degree 3 polynomial $p(\cdot)$ (over n variables) that minimizes the "loss function":

$$\sum_{i=1}^{k} \left(p\left(x_i\right) - v_i\right)^2$$

- Show that the set of all polynomials with degree at most 3 with standard operations is a vector space V and, if all the values are 0 , the set of polynomials for which the loss function is 0 is a subspace of V.
- Show that $x_1, \ldots, x_k$ define a linear transformation θ from V to $\mathbb{R}^k$. Use this to show an inner product function for which the above loss minimization problem is just a problem of computing the projection of a given vector.
- Solve the learning problem using the adjoint transformation $\theta^\dagger$. You may assume that θ is one-to-one.

Solution:

- Consider the set V of all polynomials with degree 3 at most.
  - **Closure:** For all $p, q \in S$, $p + q$ has degree at most 3. Therefore $p + q \in V$. So V is closed under addition
  - **Associativity** For all $p, q, r \in V$, $\forall\ x \in \mathbb{R}$,

    $$\begin{aligned}((p+q)+r)(x) = (p+q)(x) + r(x) = (p(x)+q(x)) + r(x) = p(x) + (q(x)+r(x)) \\ = p(x) + (q+r)(x) = (p+(q+r))(x)\end{aligned}$$

    Hence $(p+q)+r = p+(q+r)$. So V follows associativity.
  - **Identity:** The identity element of V is the 0 polynomial. It has degree 0 so $0 \in V$. And for any $x \in \mathbb{R}$, and $p \in V$, $p(x) + 0(x) = 0(x) + p(x) + p(x)$. Hence $p + 0 + 0 + p = p$.
  - **Inverse:** For any $p \in V$, $p \neq 0$, the inverse of p is $-p$ where the coefficients of $-p$ are the negative of the coefficients of p. Then for all $x \in \mathbb{R}$,

    $$(p+(-p))(x) = p(x)+(-p)(x) = p(x)-p(x) = 0 \quad ((-p)+p)(x) = (-p)(x)+p(x) = -p(x)+p(x) = 0$$

    Therefore $p + (-p) = (-p) + p = 0$.
  - **Commutativity:** Let $p, q \in V$. Then for any $x \in \mathbb{R}$, $(p+q)(x) = p(x) + q(x) = q(x) + p(x) = (q+p)(x)$. Therefore V is commutative.
  - **Scaler Multiplication:** For any $a, b \in \mathbb{R}$, and $\forall\ p \in V$

    $$\forall\ x \in \mathbb{R}\ (1 \cdot p)(x) = p(x) \implies 1 \cdot p = p$$

    $$\forall\ x \in \mathbb{R}\ (a \cdot (b \cdot p))(x) = a \cdot (b \cdot p)(x) = a \cdot (b \cdot p(x)) = (a \cdot b)p(x) \implies a \cdot (b \cdot p) = (a \cdot b)p$$

    So V is closed under scaler multiplication.
  - **Distributivity:** For all $p, q \in V$ and $a, b \in \mathbb{R}$

    $$\begin{aligned}\forall\ x \in \mathbb{R}\ (a \cdot (p+q))(x) = a \cdot (p+q)(x) = a \cdot (p(x)+q(x)) = a \cdot p(x) + a \cdot q(x) = (a \cdot p + a \cdot q)(x) \\ \implies a \cdot (p+q) = a \cdot p + a \cdot q\end{aligned}$$

And

$$\forall\, x \in \mathbb{R}\; ((a+b)\cdot p)(x) = (a+b)\cdot p(x) = a\cdot p(x)+b\cdot p(x) = (a\cdot p+b\cdot p)(x) \implies (a+b)\cdot p = a\cdot p+b\cdot p$$

Hence V follows distributivity property

Therefore V is a vector space over $\mathbb{R}$

Now for any $a \in \mathbb{R}$, consider the map $v_a : V \to \mathbb{R}$ where $v_a(p) = p(a)$. Now for any $p, q \in V$, $v_a(p+q) = (p+q)(a) = p(a) + q(a) = v_a(p) + v_a(q)$. Therefore v_a is a linear map. Now the consider the map $\theta : V \to \mathbb{R}^k$ where

$$\theta(p) = (v_{x_i}(p) \colon i \in [k])$$

Therefore θ is also a linear map between the vector spaces V and $\mathbb{R}^k$.

Suppose all the values are 0. Then the set of polynomials of V for which the loss function is 0 is exactly the polynomials whose evaluations at x_i is 0 for all $i \in [k]$. Hence the set of polynomials in V for which the loss function is 0 is $\ker \theta$. Since $\ker \theta$ is a subspace of V

- As we explained before for any $a \in \mathbb{R}$, $v_a : V \to \mathbb{R}$ where $v_a(p) = p(a)$ is a linear transformation. Therefore the map $\theta = (v_{x_i} \colon i \in [k])$ where for any $p \in V$, $\theta(p) = (p(x_1), p(x_2), \dots, p(x_k))$ is also linear map. Let $W = \operatorname{Im} \theta$.

  We will take the standard inner product on $\mathbb{R}^k$. Then $\mathbb{R}^k = W \oplus W^\perp$ where $W^\perp$ is orthogonal to W. Let $v \in \mathbb{R}$. Then v can be written uniquely as $v = w + w^\perp$ where $w \in W$ and $w^\perp \in W^\perp$. Then we have

  $$\|v - \theta(p)\|^2 = \|w - \theta(p)\|^2 + \|w^\perp\|^2$$

  Hence $p \in V$ be such that $\|v - \theta(p)\|$ is minimum then $\|w - \theta(p)\|^2$ is minimum which can be zero. Therefore $p \in \theta^{-1}(w)$. Therefore the loss is minimum when we pick a polynomial from the $\theta^{-1}(v - w^\perp)$ where $w^\perp$ is the projection or v orthogonal to the space $\operatorname{Im} \theta$.

■

Problem 9

Let $n > 0$. Call an operator $\theta \in L(\mathbb{R}^n)$ a "rotation" if for all $v \in \mathbb{R}^n$, the (Euclidean) norm of v equals the (Euclidean) norm of $\theta(v)$. As an example, the transformation $\theta \in L\left(\mathbb{R}^2\right)$ mapping $(x, y) \to (y, x)$ is a rotation. For what values of $n > 0$ does there exists a rotation that does not have any non-zero fixed points, i.e., for all $v \neq 0 \in \mathbb{R}^n$, we have $\theta(v) \neq v$.

Solution: For all $n > 0$, consider the operator $\theta \in L(\mathbb{R}^n)$ where for any $v \in \mathbb{R}^n$, $\theta(v) = -v$. Now $\|v\| = \|-v\|$. Hence θ is also a rotation operator. Then this operator has only one fixed point which is the 0 vector. Therefore for all $n > 0$ there exists a rotation that does not have any non-zero fixed point. ■
[Me and Soumyadeep solved this together.]

Problem 10

Fields with only 1 element do not exist. Nonetheless, we can still use them in problem sets. Let $q > 0$ and V be a vector space over a field $\mathbb{F}$ of size q. Define $n = \dim(V)$.

- For $0 \leq k \leq n$, compute the number of subspaces of V of dimension k. Call this number $\binom{n}{k}_q$.
- Prove the identities:

$$\binom{n}{k}_q = q^k \cdot \binom{n-1}{k}_q + \binom{n-1}{k-1}_q = \binom{n-1}{k}_q + q^{n-k} \cdot \binom{n-1}{k-1}_q$$

$$\binom{2n}{n}_q = \sum_{k=0}^{n} q^{k^2} \cdot \binom{n}{k}_q^2$$

Setting $q = 1$, observe how a vector space over a field of size 1 behaves like a set.

Solution:

- For $k = 0$ we have $\binom{n}{0}_q = 1$. So now we will assume $k \geq 1$. We will first calculate the number of ways to choose k linearly independent vectors in V which forms as a basis of a k-dimensional vector space.

  Lemma 5. *Let W be an n-dimensional vector space over $\mathbb{F}_q$, the finite field with q elements, and let $0 \leq k \leq n$. Then there exist*

  $$\frac{(q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{k-1})}{k!}$$

  many linearly independent subsets of W consisting of k elements.

  Proof: W has q^n elements. So number of ways to select a nonzero vector from W is $q^n - 1$. Now we have chosen the first linearly independent vector v_1. Now the second linearly independent vector must not by in the space $\langle v_1 \rangle$. So number ways to choose the second vector is $q^n - q$. So now we have chosen v_2. Continuing like this suppose we have chosen v_i. Then the next vector should not belong to the subspace $\langle v_1, \ldots, v_i \rangle$. Therefore there are $q^n - q^i$ choices for the next vector so that it is linearly independent of all the other vectors chosen before. So there are $q^n - q^i$ choices for v_{i+1}. Therefore for $i = k-1$ there are $q^n - q^{k-1}$ choices for choosing v_k. Hence there are $\prod\limits_{i=0}^{k}(q^n - q^i)$ many set of linearly independent k-vector tuples $(v_1, \ldots, v_k)$. Now we can pick the same set of linearly independent vectors $\{v_1, \ldots, v_k\}$ in any order. So the number of ways to choose k linearly independent vectors from W is $\frac{1}{k!}\prod\limits_{i=0}^{k-1}(q^n - q^i)$. □

  Now for any k-dimensional subspace of V is spanned by k linearly independent vectors of V. But same k-dimensional subspace can be generated by multiple set of k linearly independent vectors. So we have to divide the total number of ways to chose k linearly independent subsets from V by total number of ways to chose k linearly independent subsets from a k-dimensional vector space. So number of different k-dimensional subspaces of V is $\frac{\frac{1}{k!}\prod_{i=0}^{k-1}(q^n-q^i)}{\frac{1}{k!}\prod_{i=0}^{k-1}(q^k-q^i)} = \frac{\prod_{i=0}^{k-1}(q^n-q^i)}{\prod_{i=0}^{k-1}(q^k-q^i)}$. Therefore $\binom{n}{k}_q = \prod\limits_{i=0}^{k-1}\frac{q^n-q^i}{q^k-q^i} = \prod\limits_{i=0}^{k-1}\frac{q^{n-i}-1}{q^{k-i}-1}$.

- –

  $$\begin{aligned}
  q^k\binom{n-1}{k}_q + \binom{n-1}{k-1}_q &= q^k\prod_{i=0}^{k-1}\frac{q^{n-1-i}-1}{q^{k-i}-1} + \prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\
  &= \left[\frac{q^k(q^{n-k}-1)}{q^k-1} + 1\right]\prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\
  &= \frac{q^k(q^{n-k}-1)-(q^k-1)}{q^k-1}\prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\
  &= \frac{q^n-1}{q^k-1}\prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\
  &= \frac{q^n-1}{q^k-1}\prod_{i=1}^{k-1}\frac{q^{n-i}-1}{q^{k-i}-1} \\
  &= \prod_{i=0}^{k-1}\frac{q^{n-i}-1}{q^{k-i}-1} = \binom{n}{k}_q
  \end{aligned}$$

–

$$\begin{aligned}\binom{n-1}{k}_q + q^{n-k}\binom{n-1}{k-1}_q &= \prod_{i=0}^{k-1}\frac{q^{n-1-i}-1}{q^{k-i}-1} + q^{n-k}\prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\ &= \left[\frac{q^{n-k}-1}{q^k-1} + q^{n-k}\right]\prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\ &= \frac{q^{n-k}-1+(q^k-1)q^{n-k}}{q^k-1}\prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\ &= \frac{q^n-1}{q^k-1}\prod_{i=0}^{k-2}\frac{q^{n-1-i}-1}{q^{k-1-i}-1} \\ &= \frac{q^n-1}{q^k-1}\prod_{i=1}^{k-1}\frac{q^{n-i}-1}{q^{k-i}-1} \\ &= \prod_{i=0}^{k-1}\frac{q^{n-i}-1}{q^{k-i}-1} = \binom{n}{k}_q\end{aligned}$$

– Now we have

$$\begin{aligned}\binom{n}{k}_q &= \prod_{i=0}^{k-1}\frac{q^{n-i}-1}{q^{k-i}-1} \\ &= \prod_{i=0}^{k-1}\frac{q^{n-i}-1}{q^{k-i}-1}\prod_{i=0}^{n-k-1}\frac{q^{n-k-i}-1}{q^{n-k-i}-1} \\ &= \frac{\prod\limits_{i=0}^{n-1}(q^{n-i}-1)}{\prod\limits_{i=0}^{k-1}(q^{k-i}-1)\prod\limits_{i=0}^{n-k-1}(q^{n-k-i}-1)} \\ &= \prod_{i=0}^{n-k-1}\frac{q^{n-i}-1}{q^{n-k-i}-1} = \binom{n}{n-k}_q\end{aligned}$$

First we will prove an identity on the generating function of $\binom{n}{k}_q$.

Lemma 6. $\prod_{k=0}^{n-1}(1+q^kx) = \sum_{k=0}^{n} q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^k$ *for any* $n \in \mathbb{N}$.

Proof: We will prove this using induction on n. For $n = 1$ we have $\prod_{k=0}^{0}(1+q^kx) = 1+x$. Now

$$\sum_{k=0}^{1} q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^k = \binom{n}{0}_q + \binom{1}{1}_q x = 1+x$$

Hence the base case follows. Suppose this is true for $n-1$. Then

$$\begin{aligned}
\prod_{k=0}^{n}(1+q^kx) &= (1+q^n)\sum_{k=0}^{n} q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^k \\
&= \sum_{k=0}^{n} q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^k + \sum_{k=0}^{n} q^n q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^{k+1} \\
&= 1+\sum_{k=1}^{n} q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^k + \sum_{k=0}^{n-1} q^n q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^{k+1} + q^{n+\frac{n(n-1)}{2}}\binom{n}{n}_q x^{n+1} \\
&= 1+\sum_{k=0}^{n-1} q^{\frac{(k+1)k}{2}}\binom{n}{k+1}_q x^{k+1} + \sum_{k=0}^{n-1} q^n q^{\frac{k(k-1)}{2}}\binom{n}{k}_q x^{k+1} + q^{n+\frac{n(n-1)}{2}}\binom{n}{n}_q x^{n+1} \\
&= 1+\sum_{k=0}^{n-1}\left(q^{\frac{(k+1)k}{2}}\binom{n}{k+1}_q + q^{n+\frac{k(k-1)}{2}}\binom{n}{k}_q\right)x^{k+1} + q^{n+\frac{n(n-1)}{2}}\binom{n}{n}_q x^{n+1} \\
&= 1+\sum_{k=0}^{n-1}\left(q^{\frac{(k-1)k}{2}}\binom{n}{k+1}_q + q^{n+\frac{k(k-1)}{2}+k}\binom{n}{k}_q\right)x^{k+1} + q^{n+\frac{n(n-1)}{2}}\binom{n}{n}_q x^{n+1} \\
&= 1+\sum_{k=0}^{n-1} q^k q^{\frac{(k-1)k}{2}}\left(\binom{n}{k+1}_q + q^{n-k}\binom{n}{k}_q\right)x^{k+1} + q^{n+\frac{n(n-1)}{2}}\binom{n}{n}_q x^{n+1} \\
&= 1+\sum_{k=0}^{n-1} q^{\frac{(k+1)k}{2}}\binom{n+1}{k+1}_q x^{k+1} + q^{n+\frac{n(n-1)}{2}}\binom{n}{n}_q x^{n+1} \\
&= 1+\sum_{k=1}^{n} q^{\frac{(k-1)k}{2}}\binom{n+1}{k}_q x^{k} + q^{n+\frac{n(n-1)}{2}}\binom{n}{n}_q x^{n+1} \\
&= \sum_{k=0}^{n+1} q^{\frac{(k-1)k}{2}}\binom{n+1}{k}_q x^{k}
\end{aligned}$$

Hence by mathematical induction this is true for all $n \in \mathbb{N}$. □

Lemma 7. $\displaystyle\prod_{k=0}^{m+n-1}(1+q^kx) = \left(\prod_{k=0}^{m-1}(1+q^kx)\right)\left(\prod_{k=0}^{n-1}(1+q^{k+m}x)\right)$ *for any* $m, n \in \mathbb{N}$.

Proof: First we fix any m. We will show this using induction on n. For $n=1$ we have

$$\left(\prod_{k=0}^{m-1}(1+q^kx)\right)\left(\prod_{k=0}^{1-1}(1+q^{k+m}x)\right) = \left(\prod_{k=0}^{m-1}(1+q^kx)\right)(1+q^m) = \prod_{k=0}^{m}(1+q^kx)$$

Hence the base case follows. Now suppose this is true for n. Then

$$\begin{aligned}
\prod_{k=0}^{m+n}(1+q^kx) = (1+q^{m+n}x)\left(\prod_{k=0}^{m-1}(1+q^kx)\right)\left(\prod_{k=0}^{n-1}(1+q^{k+m}x)\right) \\
= \left(\prod_{k=0}^{m-1}(1+q^kx)\right)\left(\prod_{k=0}^{n}(1+q^{k+m}x)\right)
\end{aligned}$$

Therefore by mathematical induction this is true for n. □

Lemma 8. $\displaystyle\binom{m+n}{k}_q = \sum_{j=0}^{k}\binom{m}{k-j}_q\binom{n}{j}_q q^{j(m-k+j)}$.

Proof: Using the Lemma 6 for any $m, n \in \mathbb{N}$ we have

$$\prod_{k=0}^{m+n-1} (1+q^k x) = \sum_{k=0}^{m+n} q^{\frac{k(k-1)}{2}} \binom{m+n}{k}_q x^k$$

Now we also have using Lemma 7

$$\begin{aligned}
\prod_{k=0}^{m+n-1} (1+q^k x) &= \left(\prod_{k=0}^{m-1} (1+q^k x)\right) \left(\prod_{k=0}^{n-1} (1+q^k (q^m x))\right) \\
&= \left(\sum_{k=0}^{m} q^{\frac{k(k-1)}{2}} \binom{m}{k}_q x^k\right) \left(\sum_{k=0}^{n} q^{\frac{k(k-1)}{2}} \binom{m}{k}_q (q^m x)^k\right) \\
&= \left(\sum_{k=0}^{m} q^{\frac{k(k-1)}{2}} \binom{m}{k}_q x^k\right) \left(\sum_{k=0}^{n} q^{mk+\frac{k(k-1)}{2}} \binom{m}{k}_q x^k\right) \\
&= \sum_{k=0}^{m+n} \left(\sum_{l=0}^{k} q^{\frac{(k-l)(k-l-1)}{2}} \binom{m}{k-l}_q q^{ml+\frac{l(l-1)}{2}} \binom{n}{l}_q \right) x^k \\
&= \sum_{k=0}^{m+n} \left(\sum_{l=0}^{k} q^{\frac{k(k+1)}{2}} q^{l(m+l-k)} \binom{m}{k-l}_q \binom{n}{l}_q \right) x^k
\end{aligned}$$

Hence by equating coefficients we have $\binom{m+n}{k}_q = \sum\limits_{j=0}^{k} \binom{m}{k-j}_q \binom{n}{j}_q q^{j(m-k+j)}$ □

Using Lemma 8 for $m = n = k$ we get

$$\binom{2n}{n}_q = \sum_{j=0}^{n} \binom{n}{n-j}_q \binom{n}{j}_q q^{j(n-n+j)} = \sum_{j=0}^{n} \binom{n}{n-j}_q \binom{n}{j}_q q^{j^2} = \sum_{j=0}^{n} \binom{n}{j}_q^2 q^{j^2}$$

■

[Me and Soumyadeep solved this together. I discussed with Aakash and Shubham]