MTH 318: COMBINATORICS

KRISHNA KAIPA

Abstract. Class notes for the course MTH318 Combinatorics at IISER-Pune during August semester of 2019.

1. Introduction

The course policy can be found on the course webpage

`www.iiserpune.ac.in/~kaipa/teaching/MTH318.`

In this introductory lecture, we gave an overview of the contents of the course. Then we discussed two problems without going into details, in order to motivate the subject and its methods.

(1) Let $b_{n,k}$ denote the number of ways to pick k objects out of a set of n-objects with repetition (i.e. with replacement). This is also the number of 'monomials' $X_1^{a_1}X_2^{a_2}\dots X_n^{a_n}$ with a_i non-negative integers satisfying $a_1+\dots+a_n=k$. We talked about the *generating functions* method to obtain this number $b_{n,k}$.
Consider the 'formal power series' $B_n(x)=\sum_{k\geq 0} b_{n,k}x^k$.
Clearly $b_{1,k}=1$ for all $k\geq 0$ and hence $B_1(x)=1+x+x^2+\dots$. Then we showed that $b_{n,k}=\sum_j b_{n-1,j}$ by considering the possible values for a_n in the monomials interpretation above. This gives $B_n(x)=(1+x+x^2+\dots)B_{n-1}(x)$. Thus $B_n(x)=(1+x+x^2+\dots)^n$. Since $(1+x+x^2+\dots)(1-x)=1$, it follows that $B_n(x)(1-x)^n=1$. From this, it can be shown (see §2 below) that $b_{n,k}=\binom{n+k-1}{k}$.

(2) The number of *derangements* of a set S of size n. A derangement is a permutation (bijective function) $f:S\to S$ such that $f(s)\neq s$ for all $s\in S$. The main technique for determining the number of derangements is the *inclusion-exclusion principle*.

2. Permutations, Combinations, and Formal Power Series

The number of r-permutations of a set of size n.
In other words the number of ordered arrangements of r objects taken from a set of n objects is

$$ {}_nP_r = n(n-1)\dots(n-r+1) = \frac{n!}{(n-r)!}. $$

This is because there are n ways to pick the first object, $n-1$ ways to pick the second object, and so on until we get to $(n-r+1)$ ways to pick the r-th object.

The number of r-permutations of a set of size n with repetition or replacement.
In other words the number of r-permutations of a multiset $\{\infty \cdot X_1, \infty \cdot X_2, \dots, \infty \cdot X_n\}$ consisting of an infinite supply of n distinct objects is simply

$$ n^r. $$

This is because there are n ways to pick the i-th objects for all $1 \leq i \leq r$.

The number of r-combinations of a set of size n.
In other words the number of size r subsets a size n set is:

$$ \binom{n}{r} = n(n-1)\dots(n-r+1)/r! = \frac{n!}{r!(n-r)!}. $$

To see this we note that ${}_nP_r$ is also equal to the number of ways to pick a subset of size r from a set of size n, and then to order the chosen r elements. In other words ${}_nP_r = r!\binom{n}{r}$ which gives the above formula for $\binom{n}{r}$. It is useful to note that $\binom{n}{r} = \binom{n}{n-r}$.

The number of r-combinations of a set of size n with repetition/replacement.
In other words the number of r-combinations of the multiset $\{\infty \cdot X_1, \infty \cdot X_2, \dots, \infty \cdot X_n\}$ (as defined above). Let us call this number $b_{n,r}$. This is also the number of monomials $X_1^{a_1} X_2^{a_2} \dots X_n^{a_n}$ with a_i non-negative integers satisfying $a_1 + \dots + a_n = k$. By setting $b_i = a_i + 1$ we can say that

$$ b_{n,r} = \#\{b_1, \dots, b_n \in \mathbb{N} : b_1 + b_2 + \dots + b_n = n + r\}. $$

Consider a string of $n+r$ dots. If we mark $n-1$ of the $n+r-1$ spaces between the dots, then the marks separate the $n+r$ dots into n nonempty collection of dots. Conversely, given $b_1, \dots, b_n \in \mathbb{N}$ satisfying $b_1 + \dots + b_n = n + r$, we mark the $n-1$ spaces after b_1 dots, $b_1 + b_2$ dots, and so on upto $b_1 + \dots + b_{n-1}$ dots. This establishes that

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

We now elaborate on the generating functions approach to obtaining $b_{n,r}$.

The ring of formal power series
Let $\mathbb{R}[[X]]$ denote the set of formal power series of the form $a_0 + a_1 X + a_2 X^2 + \dots$ where $a_i \in \mathbb{R}$. The word 'formal' means that we do not need to think of the convergence questions of the series. Given series $\sum_{i \geq 0} a_i X^i$ and $\sum_{i \geq 0} b_i X^i$ we can add them to get $\sum_{i \geq 0} (a_i + b_i) X^i$, and we can also

multiply them to get

$$(\sum_{i\geq 0} a_i X^i)(\sum_{i\geq 0} b_i X^i) = \sum_{i\geq 0}\sum_{j\geq 0} a_i b_j X^{i+j} = \sum_{k\geq 0} X^k(a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0)$$

As an example of multiplication, check that the product of the formal geometric series $1 + X + X^2 + \ldots$ with the formal series $1 - X$ is just the series 1. More generally the formal series $\sum_{i\geq 0} a_i X^i$ has a multiplicative inverse if and only if $a_0 \neq 0$. When $a_0 \neq 0$ the multiplicative inverse is unique. (See Assignment 1). The formal derivative of the monomial X^i is iX^{i-1}. We define the derivative of $\sum_{i\geq 0} a_i X^i$ to be $\sum_{i\geq 1} i a_i X^{i-1}$. Given formal series $a(X)$ and $b(X)$, the product rule

$$\tfrac{d}{dX} a(X)b(X) = (\tfrac{d}{dX} a(X))b(X) + a(X)(\tfrac{d}{dX} b(X))$$

holds (See Assignment 1). As in calculus, there is also a higher-order derivative version of product rule (See Assignment 1)

$$\tfrac{d^n}{dX^n} a(X)b(X) = \sum_{i=0}^{n} \tbinom{n}{i} (\tfrac{d^i}{dX^i} a(X))(\tfrac{d^{n-i}}{dX^{n-i}} b(X))$$

Generating function approach to obtaining $b_{n,r}$ As discussed in §1, in terms of $B_n(X) = \sum_{r\geq 0} b_{n,r} X^r$, we have $B_n(X) = (1 + X + X^2 + \ldots)^n$. Thus $B_n(X)$ is the multiplicative inverse of $(1 - X)^n$. Using the higher-derivative product rule on the identity $B_1(X)(1 - X) = 1$, we get

$$(\tfrac{d^m}{dX^m} B_1(X))(1 - X) - m(\tfrac{d^{m-1}}{dX^{m-1}} B_1(X) = 0.$$

This gives

$$(\tfrac{d^{n-1}}{dX^{n-1}} B_1(X))(1 - X)^{n-1} = (n - 1)! B_1(X).$$

Multiplying by $(1 - X)$ we get

$$(\tfrac{d^{n-1}}{dX^{n-1}} B_1(X))(1 - X)^n = (n - 1)!,$$

which shows that the unique multiplicative inverse of $(1 - X)^n$ is

$$B_n(X) = \frac{1}{n-1!}\frac{d^{n-1}}{dX^{n-1}}(1 + X + X^2 + \ldots) = \sum_{r\geq 0} \binom{n + r - 1}{n - 1} X^r.$$

Therefore $b_{n,r} = \binom{n+r-1}{r}$.

Some practice problems.

P1) (from [Brualdi]) How many seven-digit numbers are there such that the digits are distinct integers taken from $\{1, 2, \ldots, 9\}$ and such that the digits 5 and 6 do not appear consecutively in either order?

Ans: The desired number is ${}_9P_7$ minus the number of seven-digit numbers with the last condition replaced by the condition that 5 and 6 appear together. The latter number is $({}_7P_5) \times (6) \times 2$, where ${}_7P_5$ is the number of five-digit numbers with distinct digits drawn from $\{1, 2, 3, 4, 7, 9\}$, and the factor 6 is for the 6 places in this five-digit number where the symbol 56 or 65 can be inserted, and the last factor of 2 is for the two choices $56, 65$.

3. Permutations of Multisets, Equivalence Relations

Practice problems continued

P2) (from [Brualdi]) Ten people, including two who do not wish to sit next to one another, are to be seated at a round table. How many circular seating arrangements are there?
Ans: Without the restriction of the two people (call them A and B) not wanting to sit next to each other, the answer would be $10!/10 = 9!$. The number arrangements which we do not want, are $(9!/9) \times 2$ where $(9!/9)$ is the number of circular permutations of nine objects (by treating A and B as one bundle), and the factor 2 is for the ordering AB, BA. So the answer is $9! - 2 \cdot 8!$.

P3) (from [Brualdi]) How many eight-letter words can be constructed by using the 26 letters $\{A, B, \ldots, Z\}$ if each word contains j letters from the set $\{A, E, I, O, U\}$? It is understood that there is no restriction on the number of times a letter can be used in a word.
Ans: $\binom{8}{j} \cdot 5^j \cdot 21^{8-j}$.

P4) (from [Brualdi]) A bakery has eight varieties of biscuits. If a box of biscuits contains 12 biscuits, how many different options are there for a box of biscuits?
Ans: $\binom{19}{12} = \binom{n+k-1}{k}$ with $n = 8$ and $k = 12$.

P4) (from [Brualdi]) What is the number of integral solutions of the equation $x_1+x_2+x_3+x_4 = 20$ satisfying the constraints $x_1 \geq 3, x_2 \geq 1, x_3 \geq 0, x_4 \geq 5$?
Ans: same as the number of solutions in non-negative integers of the equation $y_1 + y_2 + y_3 + y_4 = 11$ where $y_1 = x_1 - 3, y_2 = x_2 - 1, y_3 = x_3, y_4 = x_4 - 5$. Therefore, the answer is $\binom{14}{11} = \binom{n+k-1}{k}$ with $n = 4$ and $k = 11$.

Permutations of Multisets

Consider a multiset $\{a_1 \cdot X_1, a_2 \cdot X_2, \ldots, a_n \cdot X_n\}$ of n distinct objects X_i appearing with multiplicity $a_i \in \mathbb{N}$ for $1 \leq i \leq n$. (Here all a_i's are finite. The number of permutations of this multiset is:

$$\frac{(a_1 + \cdots + a_n)!}{a_1!a_2! \ldots a_n!}.$$

Let m be the desired number. If the multiple copies of each symbol X_i are treated as being distinct, then the number of permutations is clearly $(a_1 + \cdots + a_n)!$. But this also equals the number m in question multiplied by the number $(a_1!a_2! \ldots a_n!)$ of permutations of the a_i copies of X_i for each i. Thus we get the asserted formula for m.

It is complicated to write down a general formula for the number of r-permutations of this multiset. But for given n, r, a_i's we can always determine the number

P6) (from [Brualdi]) Consider the multiset $S = \{3 \cdot x, 2 \cdot y, 4 \cdot z\}$ of nine objects of three types. Find the number of 8-permutations of S.
Ans: The desired number is the number of 8-permutations of $\{2 \cdot x, 2 \cdot y, 4 \cdot z\}$ plus the number of 8-permutations $\{3 \cdot x, 1 \cdot y, 4 \cdot z\}$ plus the number of 8-permutations of $\{3 \cdot x, 2 \cdot y, 3 \cdot z\}$. Hence the answer is $8!/(2!2!4!) + 8!/(3!1!4!) + 8!/(3!2!3!)$. See also Assignment.

Equivalence Relations on a set
Given a set S, a relation on S is a subset R of $S \times S$. A relation is called an *equivalence relation* if i) $(a, a) \in R$ for all $a \in S$, ii) $(a, b) \in R \Leftrightarrow (b, a) \in R$, iii) $(a, b), (c, b) \in R \Rightarrow (a, c) \in R$. The conditions i), ii),iii) are known as reflexivity, symmetry, and transitivity. Instead of the notation $R \subset S \times S$, we often say $a \sim b$ whenever $(a, b) \in R$. For $a \in S$, the *equivalence class* of an element a is $[a] = \{b \in S : b \sim a\}$.

Theorem. *The equivalence classes of an equivalence relation $\sim$ on a set S partition S (into non-empty parts). Conversely, every partition of S arises from a unique equivalence relation on S.*

Proof. Given an equivalence relation $\sim$ on S, suppose $c \in [a] \cap [b]$ then the fact that $c \sim a, c \sim b$ implies $a \sim b$ by transitivity. Therefore $x \sim a$ if and only if $x \sim b$, and hence $[a] = [b]$. Thus, we have shown that distinct equivalence classes are disjoint. In order to prove that the equivalence classes partition S, it remains to show that every $a \in S$ is in some class. This is true because $a \in [a]$.
Conversely, given a partition $S = \coprod_{i \in I} S_i$ into disjoint parts S_i indexed by an indexing set I, we define $a \sim b$ if and only if there exists some $i \in I$ with $a, b \in S_i$. Clearly i) $a \sim a$, ii) $a \sim b \Leftrightarrow b \sim a$, iii) $a \sim b, b \sim c \Rightarrow a \sim c$. Therefore $\sim$ is an equivalence relation, and clearly the equivalence classes are just the S_i for $i \in I$.

We note that the equivalence relation constructed (above) from a partition of S, coincides with $\sim$ when the partition is by equivalence classes of an equivalence relation $\sim$. Similarly, the partition into equivalence classes of an equivalence relation $\sim$ on S constructed (above) from a partition of S coincides with the original partition. This shows that the assignment $\sim \mapsto$ partition by eq. classes of $\sim$ gives a bijective correspondence between the set of equivalence relations on S, and the set of partitions of S. $\square$

4. Assignment 1

(1) Given a multiset $S = \{a_1 \cdot x_1, \ldots, a_n \cdot x_n\}$ of size $m = a_1 + \cdots + a_n$. Show that the number of $(m-1)$ permutations of S is the number of permutations of S.

(2) In how many ways can four men and eight women be seated at a round table if there are to be two women between consecutive men around the table?

(3) A group of mn people are to be arranged into m teams each with n players.
(a) Determine the number of ways if each team has a different name.
(b) Determine the number of ways if the teams don't have names.

(4) Prove that a power series $\sum_{i\geq 0} a_i x^i$ has a multiplicative inverse if and only if $a_0 \neq 0$. Moreover, the multiplicative inverse is unique when it exists.

(5) Prove the product rule and the higher-order derivative product rule for formal power series.

(6) Let n be a positive integer. Suppose we choose a sequence $i_1, i_2, \ldots, i_n$ of integers between 1 and n at random.
(a) What is the probability that the sequence contains exactly $n-2$ different integers? [Ans: $n(n-1)(n-2)(3n-5)n!/(48n^n)$]
(b) What is the probability that the sequence contains exactly $n-3$ different integers? [Ans: $\binom{n}{4}n!(n-2)(n-3)/(12n^n)$]

(7) Consider the multiset $\{n \cdot a, n \cdot b, 1, 2, 3, ..., n+1\}$ of size $3n+1$. Determine the number of its n-combinations. [Ans: $(n+1)2^n$.]

(8) Let $\mathcal{P}$ be the set of r-permutations of a set S of size m. Put an equivalence relation on $\mathcal{P}$, such that the set of equivalence classes is
   (a) the set of r-combinations of S
   (b) the set of circular r-permutations of S
   (c) if $r = m$, the set of permutations of a multiset $S' = \{a_1 \cdot X_1, \ldots, a_n \cdot X_n\}$ of size m. (Given an eq. reln on a set A, we get an equivalence relation on any subset $B \subset A$. We can take $B = \mathcal{P}$ and $A = S^m$)

   Show that each equivalence class has the same size, and hence obtain the known formulas for the sizes of a), b), c).

5. Pigeonhole principle

Pigeonhole principle states that if $n > k$ identical balls are distributed among k boxes, then there exists a box with at least two balls.
A proof by contradiction is as follows: let n_i be the number of balls placed in the i-th box for $1 \leq i \leq k$. Suppose $n_i \leq 1$ for all i, then

$$n - k = \sum_{i=1}^{k} (n_i - 1) \leq 0,$$

contradicting $n > k$.

Some sample applications of the pigeonhole principle.

1) (from [Bona]) A chess tournament has n participants, and any two players play exactly one game against each other. Then it is true that in any given point of time, there are two players who have finished the same number of games.

For $0 \leq k \leq n-1$, let b_k be the number of players who have played k matches at the given point in time. We have $b_0 + b_1 + \cdots + b_{n-1} = n$. If $b_0 = 0$ or if $b_{n-1} = 0$, then by pigeonhole principle (with $k = n-1$) we get that some $b_i \geq 2$. The case in which both $b_0, b_{n-1} > 0$ is impossible because $b_0 > 0 \Rightarrow b_{n-1} = 0$.

2) (from [Bona]) There is an element in the sequence $\{7, 77, 777, 7777, \ldots, \}$ that is divisible by 2019.

Let a_i demote the i-th element of the sequence. We claim there exists $1 \leq i \leq 2019$ for which a_i is divisible by 2019. Suppose not, then the list of 2019 remainders $a_i \text{mod} 2019$ take values in $1, 2, \ldots, 2018$. By pigeonhole principle, there exist $1 \leq i < j \leq 2019$ with $a_i = a_j \text{ mod} 2019$, i.e. 2019 divides the number $77\ldots7 \times 10^i = a_{j-i} \cdot 10^i$. Since 10^i is relatively prime to 2019, it follows that 2019 divides a_{j-i} contradicting the hypothesis. Therefore, there exists $1 \leq i \leq 2019$ for which a_i is divisible by 2019.
3) (from [Brualdi]) A chess master who has 11 weeks to prepare for a tournament decides to play at least one game every day but, to avoid tiring himself, he decides not to play more than 12 games during any calendar week. Show that there exists a succession of (consecutive) days during which the chess master will have played exactly 21 games.

Let a_i be the number of games played at the end of day i. We have $1 < a_1 < a_2 < \cdots < a_{77} \leq 12 \cdot 11 - 132$. Consider the list of natural numbers:

$$a_1, a_2, \ldots, a_{77}, a_1 + 21, a_2 + 21, \ldots, a_{77} + 21$$

If the list has repetition, the only possibility is that $a_i + 21 = a_j$ for some $i < j$, which means that on the days $i + 1, \ldots j$ exactly 21 games have been played. Therefore, it suffices to show that the list has repetition. Suppose not, then the list has 154 entries. This is impossible because the largest entry of the list $a_{77} + 21 \leq 132 + 21 = 153$.

4) (from[Brualdi]) Chinese Remainder Theorem: Given relatively prime natural numbers m and n, and integers $0 \leq a \leq m-1$ and $0 \leq b \leq n$. There exists a natural number x such that $x \equiv a \bmod m$ and $x \equiv b \bmod n$.

Consider the list of n remainders $a+jm \bmod n$ for $0 \leq j \leq n-1$. In case $a+im \equiv a+jm \bmod n$ for some $0 \leq i < j \leq n-1$, then n divides $(j-i)m$. Since n is relatively prine to m, it follows that n divides $j-i$ contradicting the fact that $0 < j-i < n$. Thus $j \mapsto a+jm \bmod n$ is a one-one function from $\{0, 1, \ldots, n-1\}$ to itself, and therefore is onto. In particular some $a+im \equiv b \bmod n$. The principle used here is analogous to the pigeonhole principle which can be restated as: if A and B are finite sets with $|A| > |B|$ then a function $f : A \to B$ cannot be one-one.

6. Assignment 2

(1) (from [Bona]) (a) The set M consists of nine positive integers, none of which has a prime divisor larger than six. Prove that M has two elements whose product is the square of an integer.

(b)* The set A consists of $n+1$ positive integers, none of which have a prime divisor that is larger than the n-th smallest prime number. Prove that there exists a non-empty subset $B \subset A$ so that the product of the elements of B is a perfect square. (Hint: rank nullity theorem holds over any field including the binary field $\{0,1\}$ with $1+1=0, 1\cdot 1=1$.)

(2) (from [Bona]) Prove that among 502 positive integers, there are always two integers so that either their sum or their difference is divisible by 1000.

(3) (from [Bona]) We select $n+1$ different integers from the set $\{1,2,\ldots,2n\}$. Prove that there will always be two relatively prime integers among the selected integers.

7. Pigeon-hole Principle, general version

Pigeon-hole Principle, general version Let n, m and r be positive integers so that $n > rm$. Let us distribute n identical balls into m identical boxes. Then there will be at least one box into which we place at least $r + 1$ balls.

Example (from [Bona]): Ten points are given within a square of unit size. Then there are two of them that are closer to each other than 0.48, and there are three of them that can be covered by a disk of radius 0.5.
We divide the unit square into 9 equal squares. Each sub-square has diameter $\sqrt{2}/3 < 0.48$. By pigeonhole principle, there exists a sub-square with at least 2 points.
Similarly if we divide the unit square into four equal sub-squares, then each sub-square has a circumcircle of radius $1/\sqrt{8} < 0.5$.

Problem (1) from Assignment 2: a) Label the numbers $n_1, \ldots, n_9$ and write $n_j = 2^{a_j}3^{b_j}5^{c_j}$. Consider the 3×9 matrix A with entries over $\{0, 1\}$ whose j-th column is $(a_j \bmod 2, b_j \bmod 2, c_j \bmod 2)^T$. There are in total 8 elements of $\{0, 1\}^3$ and hence two of the nine column are identical, say columns i and j. Then $n_i n_j$ is a square.

(b) Label the numbers $m_1, \ldots, m_{n+1}$, and let the first n primes be denoted $p_1, \ldots, p_n$. Write

$$m_j = \prod_{i=1}^{n} p_i^{a_{ij}}.$$

Let M be the $n \times (n + 1)$ matrix with entries in $\{0, 1\}$ defined by $M_{ij} = a_{ij} \bmod 2$. Treating M as a matrix over the binary field $\mathbb{F}_2$, and using the rank-nullity theorem, the matrix M has nullity at least 1. So there exists a non-zero vector $(c_1, \ldots, c_n)^T \in \mathbb{F}_2^n$ such that $Mc = \vec{0}$. Let B be the subset of A consisting of those m_j for which $c_j \neq 0$. It follows that:

$$\sum_{\{j:c_j\neq 0\}} a_{ij} = 0 \bmod 2, \quad \forall\, 1 \leq i \leq n.$$

In particular the product of the numbers in B is a square.

We also discussed Problems (4)-(5) from Assignment 1.

8. Binomial Coefficients generalized

For $x \in \mathbb{R}$ and $k \in \mathbb{Z}$ we define

$$\binom{x}{k} = \begin{cases} 0 & \text{if } k < 0 \\ 1 & \text{if } k = 0 \\ \frac{x(x-1)\ldots(x-k+1)}{k!} & \text{if } k > 0 \end{cases} \tag{1}$$

For example, if n is a positive integer then:

$$\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}.$$

We recall that for a non-negative integer n we have the identity of polynomials:

$$(1+X)^n = \sum_{k=0}^{n} \binom{n}{k} X^k.$$

We also recall *Newton's binomial theorem* from calculus which states that for an arbitrary real number $n \in \mathbb{R}$ and for $-1 < x < 1$, the infinite series

$$\sum_{k \geq 0} \binom{n}{k} x^k,$$

converges to the number $(1+x)^n$.

We recall the identity

$$\binom{x+1}{k} = \binom{x}{k} + \binom{x}{k-1}. \tag{2}$$

To see this, we first note that if $k < 0$ then both sides are 0. If $k = 0$, then the left side is 1 and the right side is $1 + 0 = 1$. If $k > 0$, it suffices to prove the polynomial identity

$$0 = \binom{X+1}{k} - \binom{X}{k} - \binom{X}{k-1}. \tag{3}$$

where $\binom{X}{k} := \frac{X(X-1)\ldots(X+1-k)}{k!}$. This identity is easily checked directly.

There is also a combinatorial proof of the polynomial identity (3) : Since the polynomial on the right has degree at most k, it suffices to show that this polynomial has more than k roots. We will show that every natural number is a root. If $x = n$ with $n \in \mathbb{N}$, then $\binom{x+1}{k}$ is the number of size k subsets of $\{1, \ldots, n+1\}$. The number of such subsets which contain $n+1$ is $\binom{n}{k-1}$ and the number of such subsets which do not contain $(n+1)$ is $\binom{n}{k+1}$. Thus each $x \in \mathbb{N}$ is a root of (3).

The identity (2) can be generalized as below

Lemma 1. *For $x, y \in \mathbb{R}$ and $k \in \mathbb{Z}$, we have:*

$$\binom{x+y}{k} = \sum_{i=0}^{k} \binom{x}{k-i} \binom{y}{i}, \qquad x, y \in \mathbb{R}, k \in \mathbb{Z}. \tag{4}$$

Proof. If $k < 0$ both sides of the identity are 0. If $k = 0$, then the left side is 1 and the right-side is $\binom{x}{0}\binom{y}{0} = 1$. If $k > 0$, then it suffices to show the following identity in the polynomial ring $\mathbb{R}[X, Y]$:

$$0 = h(X, Y) := \binom{X+Y}{k} - \sum_{i=0}^{k} \binom{X}{k-i}\binom{Y}{i}.$$

Note that the (total) degree of $h(X, Y)$ is at-most k. (This means that any monomial $X^\mu Y^\nu$ that appears in $h(X, Y)$ must satisfy $\mu + \nu \leq k$). Therefore, we can write

$$h(X, Y) = \sum_{i=0}^{k} h_i(X)Y^{k-i},$$

for some polynomials $h_i(X) \in \mathbb{R}[X]$ of degree at most i. We note that $h(m, n) = 0$ for $m, n \in \mathbb{N}$: indeed any size-k subset of the set $\{1, \ldots, m+n\}$ is the disjoint union of a size i subset of $\{1, \ldots, m\}$ and a size $(k - i)$ subset of $\{m + 1, \ldots, m + n\}$. Therefore $\binom{m+n}{k} = \sum_{i=0}^{k} \binom{m}{k-i}\binom{n}{i}$. In particular for $m \in \mathbb{N}$, the polynomial $h(m, Y)$ in $\mathbb{R}[Y]$ of degree at most k has infinitely many roots (namely $Y = n$ for $n \in \mathbb{N}$), and is hence the zero polynomial. Therefore

$$h(m, Y) = \sum_{i=0}^{k} h_i(m)Y^{k-i} = 0,$$

(here 0 is the zero polynomial in $\mathbb{R}[Y]$). This, in turn shows that for each $0 \leq i \leq k$, the polynomial $h_i(X) \in \mathbb{R}[X]$ of degree at most i has infinitely many roots (namely $X = m$ for $m \in \mathbb{N}$), and is therefore the zero polynomial. Returning to the expression $h(X, Y) = \sum_{i=0}^{k} h_i(X)Y^{k-i}$, we conclude that $h(X, Y)$ is the zero polynomial in $\mathbb{R}[X, Y]$. □

Example: Setting $y = -1$ and $x = n$ in (4) and using the fact that $\binom{-1}{i} = (-1)^i$ we get the identity

$$\binom{n-1}{k} = \sum_{i=0}^{k} (-1)^{k-i} \binom{n}{i}.$$

Another generalization of (2) is as follows. Repeatedly using (2) $(n + 1)$-times (where n is a non-negative integer), we get:

$$\binom{x+1}{k+1} = \left[\binom{x}{k} + \binom{x-1}{k} + \cdots + \binom{x-n}{k}\right] + \binom{x-n}{k+1}. \tag{5}$$

In particular for $x = n$ where n is a non-negative integer, $k \geq 0$ (or more generally $k \neq -1$) we have:

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n-1}{k} + \cdots + \binom{0}{k} = \sum_{j=k}^{n} \binom{j}{k}. \tag{6}$$

A combinatorial interpretation of (6) is as follows. Given an integer j, there are $\binom{j}{i}\binom{n-j}{k-i}$ subsets of $\{1, \ldots, n + 1\}$ of size $(k + 1)$ with the property that the $(i + 1)$-th largest element of the subset is $j + 1$. Therefore we get the identity:

$$\binom{n+1}{k+1} = \sum_{j} \binom{j}{i}\binom{n-j}{k-i}. \tag{7}$$

We note that (6) is a special case of (7) for $i = k$ (or $i = 0$).

The most well known identity about binomial coefficients is the duality

$$\binom{n}{k} = \binom{n}{n-k} \qquad \text{for } n \in \{0, 1, 2, \dots\}, k \in \mathbb{Z}. \tag{8}$$

This follows from the symmetric expression $n!/(k!(n-k)!)$ for either side of the identity.

Lemma 2. *(Unimodality of binomial coefficients) Let $0 \le k \le n$ be integers. If n is even then the maximum value of $\binom{n}{k}$ as a function of k occurs for $k = n/2$, and:*

$$\binom{n}{0} < \binom{n}{1} < \cdots < \binom{n}{n/2} > \binom{n}{n/2-1} > \cdots > \binom{n}{n}.$$

If n is odd, then the maximum value of $\binom{n}{k}$ as a function of k occurs at $k = (n-1)/2$ and $k = (n+1)/2$, and:

$$\binom{n}{0} < \binom{n}{1} < \cdots < \binom{n}{(n-1)/2} = \binom{n}{(n+1)/2} > \cdots > \binom{n}{n}.$$

Proof. This follows from the fact that

$$\frac{\binom{n}{k}}{\binom{n}{k+1}} - 1 = \frac{k+1}{n-k} - 1 = 2\left(\frac{k - \frac{n-1}{2}}{n-k}\right),$$

is negative if $k < (n-1)/2$, positive if $k > (n-1)/2$ and zero if $k = (n-1)/2$. □

For $n \in \{0, 1, 2, \dots\}$, the n-th Pascal matrix A_n of size $(n+1) \times (n+1)$ matrix defined by $A_{ij} = \binom{j-1}{i-1}$. The matrix A_n is upper triangular and has ones on the diagonal, and is hence invertible.

Lemma 3. *The ij-th entry of A_n^{-1} is $(-1)^{i+j}$ times the ij-th entry of A_n. In other words,*

$$\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & (-1)^n \end{pmatrix} A_n \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & (-1)^n \end{pmatrix}.$$

In long form:

$$\sum_i (-1)^{i+q} \binom{p}{i}\binom{i}{q} = \delta_{pq}, \quad p, q \in \{0, 1, 2, \dots\}$$

Proof. Let V be the $(n+1)$-dimensional vector space of polynomials in one variable X of degree at most n and real coefficients:

$$V = \{a_0 + a_1 X + \cdots + a_n X^n : a_0, \dots, a_n \in \mathbb{R}\}.$$

The function $T : V \to V$ given by $T(f(X)) = f(X+1)$ is a linear transformation. The matrix of T with respect to the basis $\{1, X, X^2, \dots, X^n\}$ is clearly A_n. The inverse transformation T^{-1} is given by $T^{-1}(f(X)) = f(X-1)$. Since

$$(X-1)^{j-1} = X^{j-1} - (j-1)X^{j-2} + \cdots + (-1)^{j-1}X^0,$$

it follows that the ij-th entry of the matrix representing T^{-1} is $(-1)^{i+j}\binom{j-1}{i-1}$. □

Two other proofs of this result are in the Assignment.

9. Assignment 3

(1) Prove Lemma 3 by considering

$$\frac{1}{r!}\frac{d^r}{dX^r}_{|X=-1}(1+X)^m.$$

(2) Show the long form of Lemma 3 by first proving the identity:

$$\binom{n}{k}\binom{k}{m} = \binom{n}{m}\binom{n-m}{k-m}.$$

(3) Give two proofs –one combinatorial and one algebraic– of the following identities:
   (a) $\binom{2n}{n} = \sum_k \binom{n}{k}^2$.
   (b) $\sum_k \binom{n}{k} = 2^n$.
   (c) $\sum_i \binom{p}{i}\binom{i}{q} = \binom{p}{q}2^{p-q}, \quad p, q \in \{0, 1, 2, \dots\}$.

(4) Let n and k be integers with $1 \leq k \leq n$. Prove that

$$\sum_{k=1}^{n}\binom{n}{k}\binom{n}{k-1} = \frac{1}{2}\binom{2n+2}{n+1} - \binom{2n}{n}$$

(5) (optional) Let V_n be the vector space of polynomials in one variable of degree at most n and real coefficients. Consider the basis $\{1, X, X^2/2!, \dots, X^n/n!\}$ for V_n.
   (a) Find the inverse of the $(n+1)\times(n+1)$ upper triangular matrix whose ij-th entry (for $j \geq i$) is $1/(j-i)!$.
   (b) Let $a \in \mathbb{R}$. Find the matrix of the linear transformation $T_a : V_n \to V_n$ given by $T_a(f(X)) = f(X+a)$.
   (c) Find the matrix of the linear transformation $D : V_n \to V_n$ given by $Df(X) = f'(X)$ (the derivative of $f(X)$).
   (d) Verify that the linear transformation $\exp(aD)$ (using the usual series for exp) equals T_a and interpret this.
   (e) If $V = \mathbb{R}[X]$, show that $T_a = \exp(aD)$ by showing that $T_a(f) = \exp(aD)(f)$ for any $f \in \mathbb{R}[X]$.

10. Sperner's theorem

Sperner's theorem Let $S = \{1, \ldots, n\}$. Let 2^S denote the set of all subsets of S. A *chain* of S is a subset $\mathcal{C}$ of 2^S such that if $A, B \in \mathcal{C}$ then either $A \subset B$ or $B \subset A$. The largest possible size of a chain of S is clearly n, and any such chain is of the form

$$\{x_1\} \subset \{x_1, x_2\} \subset \cdots \subset \{x_1, \ldots, x_i\} \subset \cdots \subset \{x_1, \ldots, x_n\},$$

where $(x_1, \ldots, x_n)$ is a permutation of $(1, \ldots, n)$. It follows that there are $n!$ maximal chains of S. Similarly there are $k!(n-k)!$ maximal chains of S which contain a fixed subset B of S of size k.

An *antichain* of S is a subset $\mathcal{A}$ of 2^S such that if $A \not\subset B$ for any pair of elements $A, B \in \mathcal{A}$. For any $1 \leq k \leq n$, the set of all k-element subsets of S is an antichain of size $\binom{n}{k}$.

Theorem 4. *(Sperner's Theorem) The maximum possible size of an antichain of* $\{1, \ldots, n\}$ *is* $\binom{n}{\lfloor n/2 \rfloor}$.

Proof. Let β denote the maximum possible size of an antichain of $S = \{1, \ldots, n\}$. As noted above, there do exist antichains of size $\binom{n}{k}$ for any $1 \leq k \leq n$. It follows from Lemma 2 that there exist antichains of size $\binom{n}{\lfloor n/2 \rfloor}$. Therefore, $\beta \geq \binom{n}{\lfloor n/2 \rfloor}$. It remains to show that $\beta \leq \binom{n}{\lfloor n/2 \rfloor}$.

Suppose $\mathcal{C}$ is a chain of S and $\mathcal{A}$ is an antichain of S. If $A, B \in \mathcal{C}$ then we have $A \subset B$ or $B \subset A$, and hence at most one of A and B can be an element of $\mathcal{A}$ (from the definition of antichain). This shows that $0 \leq |\mathcal{C} \cap \mathcal{A}| \leq 1$.

Given an antichain $\mathcal{A}$, for each $A \in \mathcal{A}$ let $\mathcal{C}_A$ denote the set of all maximal chains of S which contain A. If $A, B \in \mathcal{A}$, then we note that $\mathcal{C}_A$ and $\mathcal{C}_B$ are disjoint because, as noted above, a chain can contain at most one of A and B. Therefore, the quantity

$$\sum_{A \in \mathcal{A}} |\mathcal{C}_A|,$$

is the number of those maximal chains of S which have at least one element in common with $\mathcal{A}$. This number is clearly less than the number of all maximal chains of S, which as observed above is $n!$. Hence, we get

$$n! \geq \sum_{A \in \mathcal{A}} |\mathcal{C}_A| \tag{9}$$

For each $1 \leq k \leq n$, let α_k denote the number of elements of $\mathcal{A}$ of size k. We must show that

$$\binom{n}{\lfloor n/2 \rfloor} \geq |\mathcal{A}| = \alpha_1 + \alpha_2 + \cdots + \alpha_n.$$

We also know that if $|A| = k$, then $|\mathcal{C}_A| = k!(n-k)!$. Therefore:

$$\sum_{A \in \mathcal{A}} |\mathcal{C}_A| = \sum_{k=1}^{n} \alpha_k k!(n-k)! = n! \sum_{k=1}^{n} \frac{\alpha_k}{\binom{n}{k}}$$

Using Lemma 2 we get

$$\sum_{A \in \mathcal{A}} |\mathcal{C}_A| \geq \frac{n!}{\binom{n}{\lfloor n/2 \rfloor}} (\alpha_1 + \alpha_2 + \cdots + \alpha_n). \tag{10}$$

Combining inequalities (9) and (10), we conclude that:

$$\binom{n}{\lfloor n/2 \rfloor} \geq |\mathcal{A}| \geq \beta. \tag{11}$$

This completes the proof that $\beta = \binom{n}{\lfloor n/2 \rfloor}$. □

Remark: It can be shown that the only antichains of size $\binom{n}{\lfloor n/2 \rfloor}$ are: i) the collection of all size $n/2$ subsets of S if n is even, ii) the collection of all size $(n-1)/2$ subsets of S, and the collection of all size $(n+1)/2$ subsets of S if n is odd. Equality holds in (11) if and only if equality holds in the inequalities (9) and (10). Equality in (10) holds if and only if all elements of $\mathcal{A}$ have size either $(n-1)/2$ or $(n+1)/2$ if n is odd, and $n/2$ if n is even. Equality in (9) holds if and only if each of the $n!$ maximal chains of S contain a member of $\mathcal{A}$. If n is even, it follows that equality holds in (11) if and only if $\mathcal{A}$ is the set of all $n/2$-size subsets of S. If n is odd, with a little more work, it can be shown that either all the sets in $\mathcal{A}$ have size $\binom{n}{(n-1)/2}$ or all of them have size $\binom{n}{(n+1)/2}$. (See for example §7.2 in the book Combinatorics: Topics, Techniques, and Algorithms by Peter J. Cameron)

11. Practice Problems for Quiz 1

(1) How many permutations are there of the letters of the word ADDRESSES? How many 8-permutations are there of these nine letters?

(2) How many sets of three integers between 1 and 20 are possible if no two consecutive integers are to be in a set?

(3) Use the pigeonhole principle to prove that the decimal expansion of a rational number m/n eventually is repeating. For example,

$$34478/99900 = 0.34512512512512512\ldots$$

(4) (a) Show that if $n+1$ integers are chosen from the set $\{1, 2, \ldots, 2n\}$, then there are always two which differ by 1.
(b) Show that if $n+1$ distinct integers are chosen from the set $\{1, 2, \ldots, 3n\}$, then there are always two which differ by at most 2.
(c) Generalize the previous two parts of this problem.

(5) A collection of subsets of $\{1, 2, \ldots, n\}$ has the property that each pair of subsets has at least one element in common. Prove that there are at most 2^{n-1} subsets in the collection.

(6) Evaluate $\sum_{k=0}^{n} (-1)^k \binom{n}{k} 10^k$.

(7) Find one binomial coefficient equal to the following expression:

$$\binom{n}{k} + 3\binom{n}{k-1} + 3\binom{n}{k-2} + \binom{n}{k-3}$$

(8) Determine the number of 12-combinations of the multiset $S = \{4 \cdot a, 3 \cdot b, 4 \cdot c, 5 \cdot d\}$.

(9) Determine the number of solutions of the equation $X_1 + X_2 + X_3 + X_4 = 14$ in nonnegative integers X_1, X_2, X_3, and X_4 not exceeding 8.

12. Quiz-1

(1) ([5 points]) Consider the multiset $\{n \cdot a, 1, 2, 3, ..., n\}$ of size $2n$. Determine the number of its n-combinations.

(2) ([5 points]) Find and prove a formula for

$$\sum_{r,s,t\geq 0,\, r+s+t=n} \binom{m_1}{r}\binom{m_2}{s}\binom{m_3}{t}$$

All quantities involved are non-negative integers.

(3) ([5 points]) How many permutations of

$$\texttt{FLOCCINAUCINIHILIPILIFICATION} = F^2L^3O^2C^4I^9N^3A^2U^1H^1P^1T^1$$

are there with no adjacent `I`'s.

(4) ([5 points]) There are 100 people at a party. Each person has an even number (possibly zero) of acquaintances . Prove that there are three people at the party with the same number of acquaintances (i.e. who know each other).

Answers.

(1) An n-combination of the given multiset is the same as the multiset formed by taking a with multiplicity μ together with a size $(n-\mu)$ subset of $\{1, 2, \ldots, n\}$ for some $0 \leq \mu \leq n$. Thus the desired answer is :

$$\sum_{\mu=0}^{n} \binom{n}{\mu} = 2^n.$$

Alternative Solution: We give another solution which is more complicated and hence not preferable. But it does illustrate some binomial coefficient identities and generating function techniques. Let X be the original multiset and let Y be the multiset $\{\infty\cdot a, \infty\cdot 1, \ldots, \infty\cdot n\}$ of $n+1$ distinct objects each having infinite multiplicity. Note that for treating n-combinations of X, we may assume that the multiplicity of a in X is ∞ instead of n. The number of n-combinations of Y is $\binom{(n+1)+n-1}{n} = \binom{2n}{n}$. Let $B_i \subset Y$ denote the set of those n-combinations of Y in which the element i of Y is used more than once. Clearly, the set of n-combinations of X consists of the complement of $\cup_{i=1}^{n} B_i$ in the set of n-combinations of Y. Also for $I = (i_1, \ldots, i_k)$ with $1 \leq i_1 < \cdots < i_k \leq n$, we have

$$|\cap_{i\in I} B_i| = \binom{(n+1)+(n-2k)-1}{n-2k} = \binom{2(n-k)}{n},$$

as it is equal to the number of solutions of $y_0 + y_1 + \cdots + y_n = n - 2k$ with all y_i being non-negative integers. By the inclusion-exclusion principle, the number of n-combinations of X is:

$\binom{2n}{n} - \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}\binom{2(n-k)}{n} = \sum_{k=0}^{n}(-1)^k\binom{n}{n-k}\binom{2(n-k)}{n} = \sum_{j\geq 0}(-1)^{n-j}\binom{n}{j}\binom{2j}{n}$.

The last expression is the constant term in the power series

$$(X^{-2}-1)^n(\sum_{i\geq 0}\binom{i}{n}X^i) = \frac{(1+X)^n(1-X)^n}{X^{2n}}\frac{X^n}{n!}D^n(1-x)^{-1},$$

where $D = d/dX$. Since $D^n(1-X)^{-1} = (1-X)^{-(n+1)}/n!$, we need the constant term in the Laurent-expansion of

$$(1+X)^n/(X^n(1-X))$$

which is the same as the coefficient of X^n in the power series:

$$(1+X)^n(1+X+X^2+\dots).$$

We conclude that the answer is $\sum_{i=0}^{n}\binom{n}{i} = 2^n$.

(2) The desired number is the coefficient of X^n in

$$(1+X)^{m_1}(1+X)^{m_2}(1+X)^{m_3} = (1+X)^{m_1+m_2+m_3}.$$

Therefore, the answer is $\binom{m_1+m_2+m_3}{n}$.

(3) The desired number is the same as the the number of permutations of the multiset

$$F^2L^3O^2C^4N^3A^2U^1H^1P^1T^1$$

of size 20 (we just removed the 9 I's) multiplied by the number of ways to pick 9 spots from the 21 spots consisting of the positions preceding, following and intermediate to the 20 letters created in step 1. This works out to be:

$$\frac{20!}{2!3!2!4!3!2!1!1!1!1!} \times \binom{21}{9}$$

(4) Let $n_0, n_2, \dots, n_{98}$ denote the number of persons in the party who have $0, 2, \dots, 98$ acquaintances respectively. Suppose $n_i < 3$ for all i. Since $n_0 + \dots + n_{98} = 100$, we must have $n_i = 2$ for all i. Let A and B denote the 2 persons having zero acquaintances. The set of acquaintances of any person C excludes at least 3 persons namely A, B, C. Thus $n_{98} = 0$, contradicting the fact that $n_{98} = 2$. So the assumption that $n_i < 3$ for all i is false.

13. Inclusion-Exclusion Principle

Inclusion-Exclusion Principle
Let $A_1, A_2, \ldots, A_n$ be finite subsets of a set S. Then

$$|\cup_{i=1}^{n} A_i| = E_1 - E_2 + E_3 + \cdots + (-1)^{n-1} E_n,$$

where

$$E_k = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|.$$

Proof. The proof is by induction on n. The base case $n = 2$ is well known:

$$|A_1 \cup A_2| = |A_2| + |A_2| - |A_1 \cap A_2|.$$

Assume the result is true for $n-1$ sets $A_1, \ldots, A_{n-1}$. Let

$$\begin{aligned} E'_k &= \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n-1} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| \\ E''_k &= \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n-1} |A_n \cap (A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k})| \end{aligned}$$

We also define $E''_0 = |A_n|$. It is understood that the quantities are zero when the summation runs over an empty set: for example $E'_n = E''_n = 0$. We note that :

$$E_k = E'_k + E''_{k-1}, \quad 1 \le k \le n$$

For $1 \le k \le n-1$ define $B_i = A_i \cap A_n$. We can rewrite the expression for E''_k as:

$$E''_k = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n-1} |B_{i_1} \cap B_{i_2} \cap \cdots \cap B_{i_k}|.$$

By the inductive hypothesis, we get:

$$\begin{aligned} \sum_{i=1}^{n-1} (-1)^{i-1} E''_i &= |\cup_{i=1}^{n-1} B_i| = |A_n \cap (\cup_{i=1}^{n-1} A_i)|. \\ \sum_{i=1}^{n-1} (-1)^{i-1} E'_i &= |\cup_{i=1}^{n-1} A_i|. \end{aligned}$$

Therefore

$$\sum_{i=1}^{n} (-1)^{i-1} E_i = \sum_{i=1}^{n-1} (-1)^{i-1} E'_i \quad + \quad E''_0 \quad - \sum_{i=1}^{n-1} (-1)^{i-1} E''_i.$$

This can be written as:

$$|\cup_{i=1}^{n-1} A_i| + |A_n| - |A_n \cap (\cup_{i=1}^{n-1} A_i)|.$$

The last quantity is $|\cup_{i=1}^{n} A_i|$ using the base case $n = 2$. □

Derangements of a finite set.
Let X be a finite set say $X = \{1, \ldots, n\}$. The number of derangements of X is

$$D_n = \#\{f : X \to X \mid f \text{ is bijective and } f(i) \neq i \ \forall i\}.$$

For example $D_1 = 0, D_2 = 1$ and $D_3 = 2$.

Lemma 5. *D_n is the closest integer to $n!/e$.*

Proof. Define

$$A_i = \{f : X \to X \mid f(i) = i, \ f \text{ is bijective}\}.$$

Clearly $D_n = n! - |\cup_{i=1}^n A_i|$. Note that $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}$ is the set of permutations of X which fix $i_1, i_2, \ldots, i_k$ and hence:

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = (n-k)! \quad \text{for } 1 \leq i_1 < i_2 < i_k \leq n.$$

Therefore, the terms E_k in the application of inc.-exc. principle $|\cup_{i=1}^n A_i| = \sum_{i=1}^n (-1)^{i-1} E_i$ equal

$$E_k = (n-k)!\binom{n}{k} = \frac{n!}{k!}.$$

Using this we get

$$D_n = n! - n!(\tfrac{1}{1!} - \tfrac{1}{2!} + \tfrac{1}{3!} + \cdots + (-1)^{n-1}\tfrac{1}{n!}).$$

In other words $D_n/n!$ is the sum of the first $(n+1)$ terms (i.e. the $(n+1)$-th partial sum) of the alternating series

$$e^{-1} = \sum_{i=0}^{\infty} (-1)^i \tfrac{1}{i!}.$$

As the sequence $\frac{1}{i!}$ is monotonically decreasing, we have

$$|e^{-1} - D_n/n!| < \tfrac{1}{n+1!}.$$

Consequently $|n!/e - D_n| < 1/(n+1) \leq 1/2$. This proves that D_n is the closest integer to $n!/e$. □

As another application of the inclusion-exclusion principle, we discussed the problem of determining the number of r-combinations of a multiset $X = \{n_1 \cdot a_1, \ldots, n_k \cdot a_k\}$. Let S be the multiset $\{\infty \cdot a_1, \ldots, \infty \cdot a_k\}$. The number of r-combinations of X is the number of r-combinations of S minus $|A_1 \cup \ldots A_k|$ where A_i is the the number of r-combinations of S which uses a_i more than n_i times. We discussed a specific example from [Brualdi].

14. Möbius inversion on the power-set of a finite set

Möbius inversion on the power-set of a finite set
Let X be a finite set, and 2^X the power set of X. To each function $f : 2^X \to \mathbb{R}$, we define another function $f^\dagger : 2^X \to \mathbb{R}$ by

$$f^\dagger(K) = \sum_{L \subset K} f(L).$$

The following lemma tells us how to recover f from $f^\dagger$:

Lemma 6.

$$f(K) = \sum_{L \subset K} (-1)^{|K \setminus L|} f^\dagger(L)$$

Proof. Using the definition of $f^\dagger$ the right hand side is:

$$\sum_{L \subset K} (-1)^{|K \setminus L|} f^\dagger(L) = \sum_{L \subset K} (-1)^{|K \setminus L|} \sum_{J \subset L} f(J) = \sum_{J \subset K} f(J) \sum_{\{L : K \supset L \supset J\}} (-1)^{|K \setminus L|}$$

Let $\bar{K} = K \setminus J$ and let $\bar{L} = L \setminus J$. We have

$$\sum_{\{L : K \supset L \supset J\}} (-1)^{|K \setminus L|} = \sum_{\bar{L} \subset \bar{K}} (-1)^{|\bar{K} \setminus \bar{L}|} = \sum_{i=0}^{|\bar{K}|} \binom{|\bar{K}|}{i} (-1)^i = \delta_{0,|\bar{K}|}.$$

Since $\delta_{0,|\bar{K}|} = \delta_{J,K}$, we get:

$$\sum_{L \subset K} (-1)^{|K \setminus L|} f^\dagger(L) = \sum_{J \subset K} f(J) \delta_{J,K} = f(K)$$

□

We will now show how this result implies the inclusion-exclusion principle. First we make an observation:
Let $B_1, \ldots, B_n$ be subsets of a set S. Let $X = \{1, \ldots, n\}$ and 2^X the power set of X. For each element I of 2^X, i.e. for each subset of X, let S_I denote the subset of S consisting of those elements which belong to B_i for each $i \in I$ but do not belong to B_j for $j \notin I$. In this definition, note that $S_\emptyset = S \setminus (\cup_{i=1}^n B_i)$ and $S_X = (\cap_{i=1}^n B_i)$. Clearly each element of S is in exactly one of the S_I's. Thus

$$S = \coprod_{I \subset X} S_I,$$

where

$$S_I = (\cap_{i \in I} B_i) \cap (\cap_{j \notin I} (S \setminus B_j)).$$

At this stage we note that

$$(\cap_{i \in I} B_i) = \coprod_{J \supset I} S_J$$

Consider the function $f : 2^X \to \mathbb{R}$ given by $f(I) = |S_{X \setminus I}|$. Using the previous equality, we have

$$f^\dagger(I) = \sum_{J \subset I} |S_{X \setminus J}| = |\cap_{i \notin I} B_i|.$$

In particular $f^\dagger(X) = S$. We note that $S_\emptyset = \cap_{i=1}^n (S \setminus B_i)$, and hence:

$$|\cup_{i=1}^n B_i| = |S| - |S_\emptyset| = |S| - f(X)$$

Therefore,

$$|\cup_{i=1}^n B_i| = |S| - \sum_{I \subset X} (-1)^{|X \setminus I|} f^\dagger(I) = |S| - \sum_{J \subset X} (-1)^{|J|} |\cap_{j \in J} B_j| = \sum_{J \subset X, J \neq \emptyset} (-1)^{|J|-1} |\cap_{j \in J} B_j|$$

The equality between the first and last terms of this chain of equalities is the inclusion-exclusion principle.

15. Assignment 4

(1) Let n be a positive integer and let $p_1, p_2, \dots p_k$ be all the different prime numbers that divide n. Let

$$\phi(n) = \#\{1 \leq j \leq n : j \text{ is relatively prime to } n\}.$$

Use the inclusion-exclusion principle to show that

$$\phi(n) = n \prod_{i=1}^{k} (1 - p_i^{-1})$$

(2) A metro train has six stops on its route from its starting station. There are 10 people on the metro train as it departs its starting station. Each person exits the metro train at one of its six stops, and at each stop at least one person exits. In how many ways can this happen?

(3) Prove that D_n is an even number if and only if n is an odd number.

(4) Determine the number of 10-combinations of the multiset $S = \{\infty \cdot a, 4 \cdot b, 5 \cdot c, 7 \cdot d\}$.

16. Möbius Inversion on Posets

Möbius Inversion on Posets A relation R on a set X is just a subset of $X \times X$. The relation R is called *reflexive* if $(x, x) \in R$ for all $x \in X$. The relation R is called *transitive* if

$$(x, y), (y, z) \in R \Rightarrow (x, z) \in R.$$

The relation R is called *symmetric* if

$$(x, y) \in R \Leftrightarrow (y, x) \in R.$$

The relation R is called *antisymmetric* if

$$(x, y), (y, x) \in R \Leftrightarrow x = y.$$

We recall that an equivalence relation on a set X is a relation that is reflexive, transitive and symmetric.

A *partial order* on a set X is a relation R on X which is reflexive, transitive and *anti-symmetric*. For a partial order R, it is common to denote $(a, b) \in R$, as $a \leq b$. The pair $(X, \leq)$ is called a partially-ordered set or in short a *poset*.

Examples of posets are

(1) $X = 2^S$ the power set of a set S. Here $A \leq B$ if $A \subset B$.
(2) A variant of (1) is to take X as the set of all linear subspaces of a vector space S. Here $A \leq B$ is A is a linear subspace of B.
(3) $X = \mathbb{R}$ with usual meaning of $\leq$
(4) $X = \mathbb{N}$ with $m \leq n$ if m divides n.

A poset $(X, \leq)$ is called a *totally ordered* set, if for any a, b either $a \leq b$ or $b \leq a$. In other words any two elements of X are comparable. In the examples above (3) is totally ordered. The example (1) is totally ordered if S has atmost one element. Similarly the example (2) is totally ordered if $\dim(S) \leq 1$. All other cases of the examples above are not totally ordered.

Given a poset $(X, \leq)$ let

$$\mathcal{F}(X) = \{f : X \times X \to \mathbb{R} : x \not\leq y \Rightarrow f(x, y) = 0\}.$$

Note that $\mathcal{F}(X)$ is a linear subspace of the vector space of all functions $\{f : X \times X \to \mathbb{R}\}$. Under the condition that for each pair of elements $x, y \in X$ the set (sometimes called *interval*) $\{z \in X : x \leq z \leq y\}$ is finite, we define a multiplication (called *convolution*) on $\mathcal{F}$ by defining $f(x, y) = 0$ if $x \not\leq y$ and in case $x \leq y$ we define:

$$(f * g)(x, y) = \sum_{\{z : x \leq z \leq y\}} f(x, z) g(z, y).$$

This multiplication is not necessarily commutative, i.e. $f * g$ need not equal $g * f$. The multiplication is however associative

$$(f * g * h)(x, y) = \sum_{\{u, v : x \leq u \leq v \leq y\}} f(x, u) g(u, v) h(v, y) = (f * g) * h = f * (g * h)$$

There is a unique multiplicative identity g with the property that $f*g = g*f = f$ for all $f \in \mathcal{F}(X)$. The Kronecker delta function $g(x,y) = \delta_{x,y}$ is a multiplicative identity because

$$(f * \delta)(x,y) = f(x,y) = (\delta * f).$$

The uniqueness is left as an exercise (Assignment 5). Note that if $X = \{1, \ldots, n\}$ with the usual meaning of $\leq$, then $\mathcal{F}(X)$ is isomorphic to the algebra U_n of $n \times n$ real upper triangular matrices by the map $f \mapsto A_f$ where the ij-th entry of A_f is $f(i,j)$.

Returning to a general poset X, suppose $f \in \mathcal{F}(X)$ has a left inverse g i.e. $g * f = \delta$. Then, the identity $1 = \delta(x,x) = g(x,x)f(x,x)$ shows that $f(x,x) \neq 0$. Similarly if f has a right inverse h, i.e. $f * h = \delta$, then again the condition $f(x,x)h(x,x) = 1$ forces $f(x,x) \neq 0$.

Lemma 7. *Any $f \in \mathcal{F}(X)$ satisfying $f(x,x) \neq 0$ for all $x \in X$ has a unique two-sided multiplicative inverse.*

Proof. First we note that if g is a left inverse of f and h is a right inverse of f then

$$h = \delta * h = (g * f) * h = g * (f * h) = g * \delta = g.$$

Therefore, if f has a left inverse and a right inverse, then every left inverse equals a right inverse. This shows that there is a unique h such that $f * h = h * f = \delta$. We have already seen that $g(x,x) = h(x,x) = 1/f(x,x)$ for a left inverse g or a right inverse h.

We will now determine $g(x,y)$ for $x < y$. We recall that the interval $\{z : x \leq z \leq y\}$ is finite. We partition this interval into parts $M_0, M_1, \ldots, M_n$ where $M_0 = \{x\}$ and for $i > 0$, M_i consists of those z in this interval for which $\#\{t : x < t \leq z\} = i$. If $z \in M_i$ where $i > 0$ and u satisfies $x \leq u < z$ then

$$\{t : x < t \leq z\} \supset \{t : x < t \leq u\} \amalg \{t : u < t \leq z\}.$$

The sets in the disjoint union are nonempty because u belongs to the first set and z belongs to the second set. Therefore, it follows that

$$\#\{t : x < t \leq u\} < \#\{t : x < t \leq z\} = i,$$

and hence $u \in M_j$ for some $j < i$. Assume inductively that $g(x,z)$ has been defined for $z \in M_0, \ldots, M_{i-1}$. For $z \in M_i$, we have

$$0 = \delta(x,z) = g(x,z)f(z,z) + \sum_{\{u : x \leq u < z\}} g(x,u)f(u,z).$$

As observed above, each u satisfying $x \leq u < z$ also satisfies $u \in M_j$ for some $j < i$. Therefore $g(x,u)$ is known by the inductive hypothesis. This allows us to determine

$$g(x,z) = \frac{-1}{f(z,z)} \sum_{\{u : x \leq u < z\}} g(x,u)f(u,z).$$

Next, we determine the right inverse $h(x,y)$ for $x < y$. We partition the interval $\{z : x \leq z \leq y\}$ into parts $M_0', M_1', \ldots, M_n'$ where $M_0' = \{y\}$ and for $i > 0$, M_i' consists of those z in this interval for which $\#\{t : z \leq t < y\} = i$. If $z \in M_i'$ where $i > 0$ and v satisfies $z < v \leq y$ then

$$\{t : z \leq t < y\} \supset \{t : z \leq t < v\} \amalg \{t : v \leq t < y\}.$$

The sets in the disjoint union are nonempty because v belongs to the second set and z belongs to the first set. Therefore, it follows that

$$\#\{t : v \leq t < y\} < \#\{t : z \leq t < y\} = i,$$

and hence $v \in M'_j$ for some $j < i$. Assume inductively that $h(z, y)$ has been defined for $z \in M'_0, \ldots, M'_{i-1}$. For $z \in M'_i$, we have

$$0 = \delta(z, y) = f(z, z)h(z, y) + \sum_{v:z<v\leq y} f(z, v)h(v, y).$$

As observed above, each v satisfying $z < v \leq y$ also satisfies $v \in M'_j$ for some $j < i$. Therefore $h(v, y)$ is known by the inductive hypothesis. This allows us to determine

$$h(z, y) = \frac{-1}{f(z, z)} \sum_{\{v:z<v\leq y\}} f(z, v)h(v, y).$$

□

17. Möbius Inversion on Posets, continued...

The function $\zeta(x,y)$ defined by $\zeta(x,y) = 1$ if $x \leq y$ and zero otherwise, has an inverse by the previous lemma. The Möbius function of X is the multiplicative inverse of ζ. It is denoted $\mu(x,y)$. We have $\mu(x,x) = 1$ and for $x < y$:

$$\mu(x,y) = -\sum_{\{z:x\leq z<y\}} \mu(x,z).$$

Example 1) If X is a totally ordered set, then

$$\mu(x,y) = \begin{cases} 1 & \text{if } \#\{z : x \leq z < y\} = 0 \\ -1 & \text{if } \#\{z : x \leq z < y\} = 1 \\ 0 & \text{if } \#\{z : x \leq z < y\} > 1 \end{cases}$$

Example 2 Let $(X, \leq)$ is the power set of a finite set S. Let $\mathcal{F}'(X)$ denote the vector space of all $\mathbb{R}$-valued functions on X. Consider the associative multiplication on $\mathcal{F}'(X)$ defined by $(f * g)(B) = \sum_{A\subset B} f(A)g(B \setminus A)$. There is an injective linear map $\imath : \mathcal{F}'(X) \hookrightarrow \mathcal{F}(X)$ given by $\imath(f)(A,B) = 0$ if $A \not\subset B$ and $\imath(f)(A,B) = f(B \setminus A)$ if $A \subset B$. The map $\imath$ also respects multiplication:

$$(\imath(f) * \imath(g))(A,B) = \sum_{C:A\subset C\subset B} \imath(f)(A,C)\imath(g)(C,B) = \sum_{C:A\subset C\subset B} f(C \setminus A)g(B \setminus C) = \imath(f * g).$$

Note that $\zeta = \imath(\zeta')$ and $\delta = \imath(\delta')$ where $\zeta'(A) = 1$ for all $A \subset X$ and $\delta'(A) = \delta_{\emptyset,A}$. Any $f \in \mathcal{F}'(X)$ satisfying $f(\emptyset) \neq 0$ has a multiplicative inverse given by

$$g(Y) = \tfrac{-1}{f(\emptyset)} \sum_{Z\subsetneq Y} g(Z)f(Y \setminus Z).$$

It follows that $\imath(g)$ is the multiplicative inverse of $\imath(f)$. In particular $\mu(A,B) = \mu'(B \setminus A)$ where $\mu'(\emptyset) = 1$ and for non-empty A we have

$$\mu'(A) = -\sum_{B\subsetneq A} \mu'(B).$$

It is clear that $\mu'(A)$ depends only on the size of A, say $\mu'(A) = \varphi(|A|)$ for some function $\varphi : \{0, 1, \dots\} \to \mathbb{R}$ satisfying $\varphi(0) = 1$ and for $n > 0$:

$$\varphi(n) = -\sum_{i=0}^{n-1} \binom{n}{i} \varphi(i).$$

This can be expressed as saying that $A_n(\varphi(0), \dots, \varphi(n))^T = (1, 0, \dots, 0)^T$ where A_n is the $(n+1) \times (n+1)$ lower triangular Pascal matrix. Since we know that the (ij)-th entry of A_n^{-1} is $(-1)^{i+j}\binom{i-1}{j-1}$, we get for $m > 0$:

$$\begin{pmatrix} \varphi(0) \\ \varphi(1) \\ \vdots \\ \varphi(n) \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ \vdots \\ (-1)^n \end{pmatrix}.$$

Therefore $\mu'(A) = (-1)^{|A|}$ and returning to $\mathcal{F}(X)$, we conclude $\mu(K,L) = (-1)^{|L\setminus K|}$ for $K \subset L$.

Example 3) If $X_1, \dots, X_k$ are posets, then the Cartesian product $X = X_1 \times \cdots \times X_k$ is a poset by defining $(a_1, \dots, a_k) \leq (b_1, \dots b_k)$ if $a_i \leq b_i$ for each i. (Check reflexive, transitive, and antisymmetric properties Assignment 5). Let $\mathcal{F}'(X) = \prod_{i=1}^k \mathcal{F}(X_i)$ be the linear subspace of $\mathcal{F}(X)$ consisting of functions of the form

$$(\Pi_{i=1}^k f_i)\left((a_1, \dots, a_k), (b_1, \dots b_k)\right) = f_1(a_1, b_1) f_2(a_2, b_2) \dots f_k(a_k, b_k).$$

Let $\vec{a}$ denote $(a_1, \dots, a_k)$ and let $\vec{c} \leq \vec{a}$ denote $c_i \leq a_i$ for each $1 \leq i \leq k$. We note that $\mathcal{F}'(X)$ is also a sub-algebra of $\mathcal{F}(X)$ because

$$((\Pi_{i=1}^k f_i) * (\Pi_{i=1}^k g_i))(\vec{a}, \vec{b}) = \textstyle\sum_{\vec{a} \leq \vec{c} \leq \vec{b}} (\Pi_{i=1}^k f_i)(\vec{a}, \vec{c})\, (\Pi_{i=1}^k g_i)(\vec{c}, \vec{b}) = \Pi_{i=1}^k \sum_{a_i \leq c_i \leq b_i} f_i(a_i, c_i) g_i(c_i, b_i) =$$
$$= \Pi_{i=1}^k (f_i * g_i)(a_i, b_i) = (\Pi_{i=1}^k (f_i * g_i))(\vec{a}, \vec{b}).$$

As an application of this, suppose $f_i \in \mathcal{F}(X_i)$ have multiplicative inverses denoted f_i^{-1}. We get

$$(\Pi_{i=1}^k f_i) * (\Pi_{i=1}^k f_i^{-1}) = \Pi_{i=1}^k (f_i * f_i^{-1}) = \Pi_{i=1}^k \delta_{X_i} = \Pi_{i=1}^k (f_i^{-1} * f_i) = (\Pi_{i=1}^k f_i^{-1}) * (\Pi_{i=1}^k f_i)$$

Since $\delta_X = \Pi_{i=1}^k \delta_{X_i}$, we conclude that $\Pi_{i=1}^k f_i^{-1}$ is the multiplicative inverse of $\Pi_{i=1}^k f_i$. In particular

$$\mu_X = \Pi_{i=1}^k \mu_{X_i}$$

For a natural number n with k prime factors $n = p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}$, let X_i be the totally ordered set

$$X_i : p_i^0 < p_i^1 < \cdots < .$$

The product $X = X_1 \times \cdots \times X_k$ (the set of natural numbers with no prime factors other than $p_1, \dots, p_k$) has the product partial order $t \leq s$ if $t|s$. The Möbius function

$$\mu_X(\Pi_{i=1}^k p_i^{a_i}, \Pi_{i=1}^k p_i^{b_i}) = \Pi_{i=1}^k \mu_{X_i}(p_i^{a_i}, p_i^{b_i}) = \begin{cases} 0 & \text{if some } |b_i - a_i| > 1 \\ (-1)^{\Pi_{i=1}^k (b_i - a_i)} & \text{if } 0 \leq b_i - a_i \leq 1 \text{ for all } i \end{cases}$$

We note that for $m|n$, we have $\mu(m, n) = \mu(1, n/m)$. In particular $\mu_X(1, n)$ is 0 if n is not square-free (i.e. some $m_i > 1$), and it is equal to $(-1)^k$ if n is square free (i.e. $m_1 = \cdots = m_k = 1$).

Theorem 8. *(Möbius Inversion on posets) Let $(X, \leq)$ be a poset with the finite intervals property. Suppose that there is a least element 0 in X (i.e. $0 \leq x$ for all $x \in X$). Given $f : X \to \mathbb{R}$ define $f^\dagger(x) = \sum_{y \leq x} f(y)$. We can recover f from $f^\dagger$ by the formula*

$$f(x) = \sum_{y \leq x} f^\dagger(y) \mu_X(y, x).$$

Proof. Let $F \in \mathcal{F}(X)$ be defined by $F(z, x) = \delta_{0,z} f(x)$. Let $G = F * \zeta$. We note that $G(z, x) = \delta_{0,z} f^\dagger(x)$. Since $G * \mu_X = F * \zeta * \mu = F$, it follows that

$$f(x) = F(0, x) = \sum_{y \leq x} G(0, y) \mu_X(y, x) = \sum_{y \leq x} f^\dagger(y) \mu_X(y, x).$$

□

Example 1) If X is a totally ordered set with a least element then Möbius Inversion states that for any $f : X \to \mathbb{R}$:

$$f^\dagger(x) = \sum_{y \leq x} f(y) \quad \Rightarrow \quad f(x) = f^\dagger(x) - f^\dagger(x - 1)$$

where $x-1$ denotes the predecessor of x in the total order X.

Example 2) For $X = 2^S$ with S finite, the least element is the empty set $\emptyset$ and Möbius Inversion states that for any $f : X \to \mathbb{R}$:

$$f^\dagger(K) = \sum_{L \subset K} f(L) \quad \Rightarrow \quad f(K) = \sum_{L \subset K} (-1)^{|K \setminus L|} f^\dagger(L),$$

which is Lemma 6.

Example 3) (Möbius Inversion in elementary number theory) Given a function $f : \mathbb{N} \to \mathbb{R}$, let $f^\dagger : \mathbb{N} \to \mathbb{R}$ be defined by $f^\dagger(m) = \sum_{d|m} f(d)$, then

$$f(n) = \sum_{d|n} \mu(1, n/d) f^\dagger(d).$$

To see this, fix $n = p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}$, and let $X = \{1, \dots, n\}$ with $i \leq j$ if $i|j$. The least element of this poset is 1. By Möbius inversion in X, we have $f(n) = \sum_{d|n} f^\dagger(d)\mu(d, n) = \sum_{d|n} f^\dagger(d)\mu(1, n/d)$.

Example 4) (Euler ϕ-function in elementary number theory) Let

$$\phi(n) = \#\{1 \leq j \leq n : j \text{ is relatively prime to } n\}.$$

Let $\xi = \exp(2\pi\sqrt{-1}/n)$. For $1 \leq m \leq n$, the least natural number j for which $(\xi^m)^j = 1$ is n/d where $d = \gcd(m, n)$. For a given divisor d of n, clearly there are $\phi(n/d)$ such values of m. This shows that

$$\phi^\dagger(n) = \sum_{d|n} \phi(d) = \sum_{d|n} \phi(n/d) = n.$$

Therefore, by inversion formula we have:

$$\phi(n) = \sum_{d|n} d\, \mu(1, n/d) = n \sum_{d|n} \mu(1, d)/d.$$

If $p_1, \dots, p_k$ are the prime divisors of n, then $\mu(1, d)$ is zero unless $d|p_1 p_2 \dots p_k$. In case $d|p_1 p_2 \dots p_k$, we have $\mu(1, d)$ equals $+1$ if d has an even number of prime factors and -1 is d has an odd number of prime factors. It follows that

$$\sum_{d|n} \mu(1, d)/d = (1 - p_1^{-1})(1 - p_2^{-1}) \dots (1 - p_k^{-1}),$$

and hence

$$\phi(n) = n \prod_{i=1}^{k} (1 - p_i^{-1}).$$

18. Assignment 5

(1) Check that if X is poset, then $\delta(x, y)$ is the unique multiplicative identity element of $\mathcal{F}(X)$.

(2) If $X_1, \ldots, X_k$ are posets, consider the relation on on the Cartesian product $X = X_1 \times \cdots \times X_k$ defined by $(a_1, \ldots, a_k) \leq (b_1, \ldots b_k)$ if $a_i \leq b_i$ for each i. Check that $\leq$ is a partial order.

(3) Consider the power set of $\{1, 2, 3\}$ denoted $X = 2^{\{1,2,3\}}$ partially ordered by set containment. Let a function $f \in \mathcal{F}(X)$ be defined by

$$f(A, B) = \begin{cases} 1, & \text{if } A = B \\ 2, & \text{if } A \subset B \text{ and } |B| - |A| = 1 \\ 1, & \text{if } A \subset B \text{ and } |B| - |A| = 2 \\ -1, & \text{if } A \subset B \text{ and } |B| - |A| = 3 \end{cases}$$

Find the inverse of f with respect to the convolution product.

(4) Consider the multiset $X = \{n_1 \cdot a_1, n_2 \cdot a_2, \ldots, n_k \cdot a_k\}$ of k distinct elements with positive repetition numbers $n_1, n_2, \cdots, n_k$. We introduce a partial order on the combinations of X by stating the following relationship: If $A = \{p_1 \cdot a_1, p_2 \cdot a_2, \cdots, p_k \cdot a_k\}$ and $B = \{q_1 \cdot a_1, q_2 \cdot a_2, \ldots, q_k \cdot a_k\}$ are combinations of X, then $A \leq B$ provided that $p_i \leq q_i$ for $i = 1, 2, \ldots, k$. Prove that this statement defines a partial order on X and then compute its Mobius function.

19. Quiz-2

(1) ([4 points]) Let X be a poset with finite intervals property. Let $\mu(\cdot,\cdot)$ denote the Möbius function of X. Write down a recursive/inductive formula for $\mu(x,y)$ where $x \leq y$.

Ans: $\mu(x,x) = 1$ and $\mu(x,y) = -\sum_{\{z:x\leq z<y\}} \mu(x,z)$.

(2) ([8 points]) The function $\lambda(n)$ is defined for positive integers n by the rule

$$\sum_{d|n} \lambda(d) = \log(n).$$

Prove that $\lambda(n) = \log(p)$ if $n = p^e$ where p is a prime, and $\lambda(n) = 0$ otherwise.

Ans: By inversion we have

$$\lambda(n) = \sum_{d|n} \log(n/d)\mu(d),$$

where $\mu(d)$ is 0 if d is not square-free, and in case d is square free then $\mu(d) = (-1)^{\#\text{prime factors of } d}$. If $n = 1$, we get

$$\lambda(1) = \log(1)\mu(1) = 0 \times 1 = 0.$$

If $n > 1$, let $p_1, \ldots, p_k$ denote the prime divisors of n. We can write:

$$\lambda(n) = \sum_{I\subset\{1,\ldots,k\}} \left(\log(n) - \log(\Pi_{i\in I} p_i)\right)(-1)^{\#I} = \log(n)(1-1)^k - \sum_{I\subset\{1,\ldots,k\}} (-1)^{\#I}\log(\Pi_{i\in I} p_i))$$

We recall our convention that $0^k = \delta_{0,k}$. Since $k \geq 1$, the first term is 0. The other term is

$$-\sum_{i=1}^{k} \log(p_i) \sum_{\{i\}\subset I\subset\{1,\ldots,k\}} (-1)^{\#I} = \sum_{i=1}^{k} \log(p_i)\,(1-1)^{k-1}.$$

Since $(1-1)^{k-1} = \delta_{1,k}$, we conclude that $\lambda(n) = 0$ if n has more than one prime factor, and in case $n = p^e$ for some prime p and $e \in \mathbb{N}$, we have $\log(n) = \log(p)$.

(3) ([8 points]) Let $a_{k,n}$ denote the number of surjective functions (also called onto functions) from $\{1,2,\ldots,k\}$ to $\{1,2,\ldots,n\}$. Using inclusion-exclusion principle, find an expression for $a_{k,n}$. (Remark: The case $k = 10$ and $n = 6$ should be familiar !)

Ans: For $I \subset \{1,\ldots,n\}$ let B_I denote the set of functions from $\{1,\ldots,k\}$ to $\{1,\ldots,n\}$ whose image does not contain the members of I. For example if $I = \emptyset$ then B_I consists of all functions from $\{1,\ldots,k\}$ to $\{1,\ldots,n\}$. Clearly $\#B_I = (n - \#I)^k$. By inclusion-exclusion principle

$$a_{k,n} = \sum_{I\subset\{1,\ldots,n\}} (-1)^{\#I}(n-\#I)^k = \sum_{j=0}^{n} (-1)^{n-j}\binom{n}{j} j^k.$$

20. Mid-Term Examination

(1) (a) (3 points) Give a combinatorial proof of the identity $\binom{2n}{n} = \sum_k \binom{n}{k}^2$.
(b) (4 points) Prove that among 502 positive integers, there are always two integers so that either their sum or their difference is divisible by 1000.

solution

a) The left side is the number of ways of picking a size-n subset of $\{1, \ldots, 2n\}$. Partitioning $\{1, \ldots, 2n\}$ as $X \amalg Y$ where $X = \{1, \ldots, n\}$ and $Y = \{n+1, \ldots, 2n\}$, each size-n subset of $\{1, \ldots, 2n\}$ is obtained by picking a size-k subset of X and a size $(n-k)$ subset of Y for some $k \in \{0, \ldots, n\}$. Therefore $\binom{2n}{n}$ equals

$$\sum_k \binom{n}{k}\binom{n}{n-k} = \sum_k \binom{n}{k}^2.$$

b) Suppose we have a set of 502 positive integers $x_1, \ldots, x_{502}$ such that the difference of any two of these integers is not divisible by 1000. The remainders $y_i = x_i \bmod 1000$ form a size-502 subset of $\{0, \ldots, 999\}$. The latter set can be partitioned into 501 boxes:

$$\{0\}, \{500\}, \{1, 999\}, \{2, 998\}, \{499, 501\}.$$

Any subset of $\{0, \ldots, 999\}$ which does not contain $\{j, 1000-j\}$ for any $j \in \{1, \ldots, 499\}$, has size at most 501. Therefore, every size 502-subset of $\{0, \ldots, 999\}$ must contain $\{j, 1000-j\}$ for some $j \in \{1, \ldots, 499\}$, in other words there exist $i \neq j$ such that $x_i + x_j$ is divisible by 1000.

(2) (6 points) How many permutations are there of the digits $1, 2, \ldots, 8$ in which none of the patterns 12, 34, 56, 78 appears?

solution Let $j \in \{1, 2, 3, 4\}$. The number of permutations of the digits $1, 2, \ldots, 8$ in which j of the patterns 12, 34, 56, 78 appear, is $(8-j)!$. Therefore, by inclusion-exclusion principle the answer to the problem is

$$8! - 4 \times 7! + 6 \times 6! - 4 \times 5! + 4! = 24024.$$

(3) (7 points) Let n be a positive integer. Suppose we choose a sequence $i_1, i_2, \ldots, i_n$ of integers between 1 and n at random. Note that order of the sequence matters, and repetition of digits is allowed. What is the probability that the sequence contains exactly $n-3$ different integers?

solution The probability is #favourable outcomes/n^n. A favourable item is obtained by the following steps

(a) Pick a size $n-3$ subset of $\{1, \ldots, n\}$,
(b) Form a multiset of size n whose underlying set is the set picked in the previous step
(c) Choose a permutation of the multiset in the previous step,

There are $\binom{n}{3}$ ways to pick the set in step 1. For step 2, there are 3 different kinds of multisets depending on the set of multiplicities: a) $4, 1, \ldots, 1$, b) $3, 2, 1, \ldots, 1$, c) $2, 2, 2, 1, \ldots, 1$. The number of corresponding multisets is a) $\binom{n-3}{1}$, b) $(n-3)(n-4)$, c) $\binom{n-3}{3}$. For step 3, the number of permutations depends on the types a)-c) in step 2): it is a) $n!/4!$, b) $n!/(3!2!)$, c) $n!/(2!2!2!)$.

The total number of favourable outcomes is therefore:

$$\binom{n}{3} \times \left(\binom{n-3}{1} \cdot \tfrac{n!}{4!} + (n-3)(n-4) \cdot \tfrac{n!}{3!2!} + \binom{n-3}{3} \cdot \tfrac{n!}{2!2!2!} \right),$$

Dividing this by n^n and simplifying the desired probability is:

$$\tbinom{n}{4}n!(n-2)(n-3)/(12n^n).$$

(4) Let d_n be the number of derangements of $\{1,2,\ldots,n\}$. Take $d_0 = 1$. Let $g(x)$ be the formal power series

$$g(X) = \sum_{j=0}^{\infty} \frac{d_j X^j}{j!}.$$

(a) (4 points) Assuming the recurrence in part b), derive a differential equation for $g(X)$ and solve it to get a closed form solution for $g(X)$.

(b) (4 points) Using the expression for d_n or some other method prove that

$$d_{n+1} = n(d_n + d_{n-1}), \qquad n \in \mathbb{N}$$

solution

(a)

$$g(X) = 1 + \sum_{n=1}^{\infty} \tfrac{d_{n+1}X^{n+1}}{n+1!} = 1 + \sum_{n=1}^{\infty} \tfrac{n(d_n+d_{n-1})X^{n+1}}{n+1!}.$$

Differentiating, we get:

$$g'(X) = \sum_{n=1}^{\infty} \tfrac{n(d_n+d_{n-1})X^{n}}{n!} = Xg(X) + Xg'(X).$$

Thus, we get the differential equation $g'(X)(1-X) = Xg(X)$. We can write this as

$$\frac{d}{dX}(g(X)(1-X)) + g(X)(1-X) = 0.$$

The unique solution of this differential equation is $g(X) = ce^{-X}/(1-X)$ where $c = g(0) = 1$. Thus $g(X) = e^{-X}/(1-X)$.

(b) We know $d_n/n! = \sum_{i=0}^{n}(-1)^i/i!$. Therefore

$$\tfrac{n}{n+1!}(d_n + d_{n-1}) = \tfrac{n}{n+1}(\sum_{i=0}^{n} \tfrac{(-1)^i}{i!}) + \tfrac{1}{n+1}(\sum_{i=0}^{n-1} \tfrac{(-1)^i}{i!}) = (\sum_{i=0}^{n-1} \tfrac{(-1)^i}{i!}) + \tfrac{(-1)^n}{n!} + \tfrac{(-1)^{n+1}}{(n+1)!}.$$

The last expression above is also equal to $d_{n+1}/(n+1)!$. This shows that $d_{n+1} = n(d_n+d_{n-1})$.

Alternative solution of (b): Let Δ_n denote the set of derangements of $\{1,\ldots,n\}$. For $1 \le i \le n$, let F_i be the set of derangements of $\{1,\ldots,n+1\}$ which map $n+1$ to i. Clearly

$$\Delta_{n+1} = \amalg_{i=1}^{n} F_i.$$

It suffices to show that $\#F_i = (d_n+d_{n-1})$. We can partition each F_i as $\amalg_{j=1}^{n} F_{ij}$ where $F_{ij} \subset F_i$ consists of those derangements which map j to $n+1$. Clearly $\#F_{ii} = d_{n-1}$. For $j \neq i$ any element of F_{ij} is a bijective function $g : \{1,\ldots,j-1,j+1,\ldots,n\} \to \{1,\ldots,i-1,i+1,\ldots,n\}$ satisfying $g(x) \neq x$ for all x. For each $j \neq i$, there is a bijection from F_{ij} to the set of derangements of $\{1,\ldots,n\}$ which map j to i: This bijection is given by $g \mapsto \hat{g}$ where

$$\hat{g}(x) = \begin{cases} g(x) & \text{if } x \neq j \\ i & \text{if } x = j. \end{cases}$$

Thus

$$\#\Delta_n = \#(\amalg_{j\in\{1,\dots,n\}\setminus i} F_{ij}.)$$

This shows that

$$\#F_i = \sum_{i=1}^{n} \#F_{ij} = d_{n-1} + d_n$$

(5) (7 points) Let $a \leq b$ be elements of a poset X with finite intervals property. By a chain of length i between a and b we mean a sequence $a = x_0 < \cdots < x_i = b$. Let μ_X denote the Möbius function of X. Prove that :

$$\mu_X(a,b) = \sum_C (-1)^{\text{length}(C)},$$

where the sum runs over all C chains between a and b.

solution We prove the result by induction on the size n of the interval $\{c : a \leq c \leq b\}$. In order to prove that the result holds for the base case $n = 1$ we must show $\sum_C (-1)^{\text{length}(C)} = \mu(a,a) = 1$. This follows because there is only one chain $a = x_0 = a$ between a and $b = a$. For $n > 1$, assume inductively that the result holds if the size of the interval $\{c : a \leq c \leq b\}$ is at most $n-1$. Next, suppose that the interval $\{c : a \leq c \leq b\}$ has size n. By the inductive formula for μ_X we have

$$\mu(a,b) = -\sum_{\{c:a\leq c<b\}} \mu(a,c).$$

Any chain $C : a = x_0 < x_1 < \cdots < x_i = b$ between a and b, can be written as $C' < b$ where C' is the chain $C' : a = x_0 < \cdots < x_{i-1}$. Thus $C \mapsto C'$ gives a bijection from the set of chains between a and b to the set of chains which begin at a and end at some $c < b$. In this notation:

$$\sum_C (-1)^{\text{length}(C)} = -\sum_{C'} (-1)^{\text{length}(C')}.$$

By the inductive hypothesis, the right side of this equation is

$$-\sum_{\{c:a\leq c<b\}} \mu(a,c).$$

21. Recurrence relations and generating functions

Recurrence relations and generating functions
The Hemachandra/Fibonacci numbers f_n: Let f_n denote the number of rhythms of (time)-duration $n-1$ consisting of two kinds of beats. A short beat and a long beat, and assume that short beat is one time unit, and the long beat is two time units.

In terms of Sanskrit poetry, we want a *vritta* (poetry meter) of $n-1$ *matras* (time units), using two types of *varnas* (syllables): a *laghu varna* having one matra (short syllable of duration one) and a *guru varna* having two matras (long syllable of duration 2). This goes back to Indian mathematician Pingala (before Christ), or in less ancient work of Hemachandra (around 1150 AD but before Fibonacci). The number of meters with i long syllables, is the number of permutations of the multiset $\{(n-1-2i)\cdot a, i\cdot b\}$ which is $\binom{n-1-i}{i}$. Thus

$$f_n = \sum_i \binom{n-1-i}{i}.$$

Also $f_n = f'_n + f''_n$ where f'_n, f''_n are the number of meters in which the last syllable is short, long respectively. Clearly $f'_n = f_{n-1}$ and $f''_n = f_{n-2}$, and hence

$$f_n = f_{n-1} + f_{n-2}.$$

We assume that $f_0 = 0$ and $f_1 = 1$. Let $H(x) = \sum_{i=0}^{\infty} f_i X^i$ be the generating function for the Hemachandra numbers. We have

$$H(X) = X + \sum_{i\geq 2}(f_{i-1} + f_{i-2})X^i = X + XH(X) + X^2H(X),$$

which gives

$$H(X) = X/(1 - X - X^2).$$

Using the fact that $X^2 + X - 1 = (X - \phi^{-1})(X + \phi)$ where $\phi = \frac{1+\sqrt{5}}{2}$ and using partial fractions we get:

$$H(X) = \tfrac{1}{\sqrt{5}}\,(\tfrac{1}{1-X\phi} - \tfrac{1}{1+X/\phi})$$

Using geometric series expansion we get:

$$f_n = \tfrac{1}{\sqrt{5}}\,(\phi^n - (-\phi^{-1})^n) = \tfrac{1}{\sqrt{5}}\,((\tfrac{1+\sqrt{5}}{2})^n - (\tfrac{1-\sqrt{5}}{2})^n).$$

As another example of generating functions we recall that:

$$(1-X)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} X^k.$$

The Stirling number of the second kind $S_{n,k}$ is the number of equivalence relations on $\{1,\ldots,n\}$ having exactly k equivalence classes. Clearly $k!S_{n,k}$ is the number of surjective functions from $\{1,\ldots,n\}$ to $\{1,\ldots,k\}$ which we have worked out to be $\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\,j^n$. Therefore

$$S_{n,k} = \tfrac{1}{k!}\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\,j^n.$$

For fixed k, the generating function $S_k(X) = \sum_{n\geq k} S_{n,k}X^n = \sum_{n\geq 0} S_{n,k}X^n$ can be expressed as

$$S_k(X) = \sum_{n\geq 0} X^n \tfrac{1}{k!} \sum_{j=0}^{k} (-1)^{k-j}\tbinom{k}{j}\, j^n = \tfrac{1}{k!}\textstyle\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\,\frac{1}{1-jX}.$$

(Note that we have used the convention $0^0 = 1$ here for the term $j = 0$.)

22. Stirling numbers of Second kind, Bell numbers

Multiplying the expression for $S_k(X)$ by $\Pi_{i=1}^{k}(1-iX)$ we get the expression:

$$f(X) = \prod_{j=1}^{k}(X-1/j) + \sum_{j=1}^{k} \tfrac{1}{j^k} \prod_{\{i:i\neq j\}} \tfrac{X-1/i}{1/j-1/i}.$$

We recall that the Lagrange interpolation polynomial for the data points $(x_1, y_1), \ldots, (x_k, y_k)$ (i.e. the unique polynomial of degree at most $k-1$ whose graph passes through these points) is

$$\sum_{j=1}^{k} y_k \prod_{\{i:i\neq j\}} \tfrac{X-x_i}{x_j-x_i}.$$

Therefore, we conclude $f(X) - \prod_{j=1}^{k}(X-1/j)$ is the desired Lagrange interpolation polynomial for the data points $(1/j, 1/j^k)$ for $1 \leq j \leq k$. However the polynomial $X^k - \prod_{j=1}^{k}(X-1/j)$ is also a polynomial of degree at most $k-1$ which interpolates the same data points. Therefore, we conclude that $f(X) = X^k$. In other words

$$S_k(X) = \frac{X^k}{(1-X)(1-2X)\ldots(1-kX)}.$$

The numbers $S_{n,k}$ obey the recurrence

$$S_{n,k} = kS_{n-1,k} + S_{n-1,k-1}.$$

To see this, we note that given a partition of $\{1, \ldots, n\}$ into k parts, either the subset $\{1, \ldots, n-1\}$ is partitioned into k parts or $k-1$ parts. In the former case, there are k ways to complete this partition of $\{1, \ldots, n-1\}$ to a partition of $\{1, \ldots, n\}$ (by deciding where to send n). In the latter case the only partition of $\{1, \ldots, n\}$ into k parts is to take $\{n\}$ as a part together with the $k-1$ parts of $\{1, \ldots, n-1\}$. Using this recurrence we get

$$S_k(X) = X^k + \sum_{n>k} S_{n,k}X^n = X^k + (X\sum_{n>k} S_{n-1,k-1}X^{n-1}) + kX(\sum_{n>k} S_{n-1,k}X^{n-1}).$$

In other words

$$S_k(X) = XS_{k-1}(X) + kXS_k(X).$$

Since $S_{n,1} = 1$, we have $S_1(X) = X/(1-X)$ and hence we recover the formula:

$$S_k(X) = \frac{X^k}{(1-X)(1-2X)\ldots(1-kX)}.$$

The exponential generating function $S_k^E(X) = \sum_{n\geq 0} S_{n,k}X^n/n!$ can be expressed as

$$S_k^E(X) = \sum_{n\geq 0} X^n \tfrac{1}{n!k!} \sum_{j=0}^{k} (-1)^{k-j}\tbinom{k}{j} j^n = \tfrac{1}{k!}\sum_{j=0}^{k}(-1)^{k-j}\tbinom{k}{j} \exp jX = (e^X-1)^k/k!.$$

The Bell number B_n is the number of equivalence relations on a set a set of size n.

$$B_n = S_{n,0} + \cdots + S_{n,n}.$$

Note that $B_0 = 1$. A recurrence satisfied by the bell numbers is

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k.$$

To see this, we note that given a partition of $\{1, \ldots, n+1\}$, the size of the part containing $n+1$ is an integer $(n-k+1)$ (where $0 \leq k \leq n$) and this part can be chosen in $\binom{n}{k}$ ways. The remaining subset of $\{1, \ldots, n+1\}$ of size k can be partitioned in B_k ways. The exponential generating function

$$B(X) = 1 + \sum_{n\geq 0} B_n X^n/n! = 1 + \sum_{i\geq 0} B_{i+1} X^{i+1}/(i+1!) = 1 + \sum_{i\geq 0}\sum_{k\geq 0} \tbinom{i}{k} B_k X^{i+1}/(i+1)!.$$

Differentiating and setting $j = i - k$, we get:

$$B'(X) = \sum_{k\geq 0} B_k X^k/k! \sum_{j\geq 0} X^j/j! = e^X B(X).$$

If $g(X) \in \mathbb{R}[[X]]$ is a formal power series with zero constant term, then for any $f \in \mathbb{R}[[X]]$ the composition $f(g(X))$ is well defined in $\mathbb{R}[[X]]$. To see this suppose $g(X) = b_1 X + b_2 X^2 + \cdots +$, and $f(X) = a_0 + a_1 X + \ldots$, then the coefficient of X^m in $f(g(X))$ is the coefficient of X^m in the polynomial $\tilde{f}(\tilde{g}(X)$ where:

$$\tilde{f}(X) = a_0 + a_1 X + \ldots a_m X^m, \quad g(X) = b_1 X + b_2 X^2 + \cdots + b_m X^m.$$

In particular if $g(X) \in \mathbb{R}[[X]]$ has zero constant term, then $e^{g(X)}$ is well defined. For example $e^{e^X - 1} \in \mathbb{R}[[X]]$ but e^{e^X} is not defined in $\mathbb{R}[[X]]$.

Returning to the differential equation $B'(X) - e^X B(X) = 0$ in $\mathbb{R}[[X]]$, we can write this within $\mathbb{R}[[X]]$ as

$$e^{e^X - 1} \tfrac{d}{dX}(B(X) e^{1-e^X}) = 0$$

Note that for $f(X), g(X) \in \mathbb{R}[[X]]$, the product $f(X)g(X)$ is the zero element of $\mathbb{R}[[X]]$ if and only if one of $f(X), g(X)$ is the zero element of $\mathbb{R}[[X]]$ (Assignment). Since e^{e^X-1} is not zero, it follows that $\frac{d}{dX}(B(X)e^{1-e^X}) = 0$. A formal power series $f(X)$ has zero derivative if and only if $f(X) = c$ has only the constant term. Therefore, we conclude that $B(X) = ce^{e^X-1}$. Since both of the series $B(X)$ and e^{e^X-1} have constant term 1, it follows that $c = 1$ and hence:

$$B(X) = e^{e^X - 1}.$$

Another way to derive this is to note that $B_n = S_{n,0} + \cdots + S_{n,n} = \sum_{k=0}^{\infty} S_{n,k}$. Thus the exponential generating function $B(X)$ equals $\sum_{k=0}^{\infty} S_k^E(X)$ (the sum of the exponential generating functions of the Stirling numbers of the second kind). Since $S_k^E(X) = (e^X - 1)^k/k!$ we again get $B(X) = e^{e^X-1}$.

Stepping out of $\mathbb{R}[[X]]$ into the analytic setting of convergent real power series, we can write $e^{e^X-1} = \frac{1}{e} e^{e^X}$. The coefficient of $X^n/n!$ in this convergent power series is:

$$e^{-1} \sum_{j\geq 1}^{n} \text{coeff. of } X^n/n! \text{ in } e^{jX}/j! = e^{-1} \sum_{j=1}^{\infty} \tfrac{j^n}{j!}$$

Therefore, we obtain B_n as the limit of a series:

$$B_n = \tfrac{1}{e} \sum_{j=1}^{\infty} \tfrac{j^n}{j!}$$

23. Assignment 6

(1) A Poisson random variable X with parameter λ has probability mass function

$$\Pr(X = k) = \lambda^k e^{-\lambda}/k!, \qquad k = 0, 1, 2, \ldots.$$

The n-th moment of a discrete random variable Y is $\sum_y y^n \Pr(Y = y)$. Determine the n-th moment of the Poisson random variable having unit expectation (mean) in terms of the topics of this section. (Ans: B_n.)

(2) Let V_m be the vector space of polynomials with real coefficients of degree at most m in one variable. Show that $E_k(X) = X(X-1)\ldots(X-k+1)$ for $0 \leq k \leq m$ (with $E_0(X) = 1$) form a basis for V_m. Express the standard basis $e_k(X) = X^k$ for $0 \leq k \leq m$ in terms of the basis $E_0, \ldots, E_m$. (Ans: $e_j = \sum_i S_{j,i} E_i$ where $S_{j,i}$ is the Stirling number of second kind.)
(First method: See quiz 3)
(Second method: Let P be the $m+1 \times m+1$ matrix such that $[e_0, \ldots, e_m] = [E_0, \ldots, E_m]P$. Let $\tilde{P}$ be the matrix defined by $\tilde{P}_{ij} = (i-1)!P_{ij}$. Obtain the relation $B = A\tilde{P}$ where A is the upper triangular Pascal matrix and $B_{ij} = (i-1)^{j-1}$. Use knowledge of A^{-1} to solve the problem.)

(3) Prove that the n-th Fibonacci number f_n is the integer that is closest to the number $\frac{1}{\sqrt{5}}(\frac{1+\sqrt{5}}{2})^n$.

(4) Prove that f_n is divisible by 4 if and only if $6|n$.

(5) By examining the Fibonacci sequence, make a conjecture about when f_n is divisible by 7 and then prove your conjecture. (Ans $8|n$.)

(6) Show that the product of two elements of $\mathbb{R}[[X]]$ is zero if and only if one of them is zero.

(7) (Optional) Let m and n be positive integers. Prove that if m is divisible by n, then f_m is divisible by f_n.

More generally $\gcd(f_m, f_n) = f_{\gcd(m,n)}$.

24. Catalan numbers

Examples of solving recurrence relations.
Towers of Hanoi problem: There are 3 pegs, and on one of the pegs n circular disks of distinct sizes are arranged in increasing order of size from top to bottom. Let h_n be the number of moves required transfer the disks to a different peg with the condition that at no point should a larger disk be placed on a smaller disk. The only way to do this is to i) first move all but the largest disk to a different peg (h_{n-1} moves required for this), then ii) move the largest disk to the remaining vacant peg (1 move for this), and iii) move the other $n-1$ disks to the peg in step ii) (again h_{n-1} moves required). Thus, we have $h_n = 2h_{n-1}+1$ for $n \geq 1$ and $h_0 = 0$. Let $H(X) = \sum_{n\geq 0} h_n X^n$. From the recurrence we obtain

$$H(X) = X(2H(X) + \tfrac{1}{1-X}) \quad \Rightarrow H(X) = \tfrac{X}{(1-X)(1-2X)} = \tfrac{1}{1-2X} - \tfrac{1}{1-X} = \sum_{n\geq 0}(2^n - 1)X^n.$$

Catalan numbers
Ballot problem: A and B are contesting an election. Suppose n people vote for A and n people vote for B. What is the probability that while counting the votes A never trails B?

Lattice Paths: Consider the integer lattice $\mathbb{Z} \times \mathbb{Z}$. A walk consists of one step right $(m, n) \mapsto (m+1, n)$ or one step up $(m, n) \mapsto (m, n+1)$. How many paths are there from $(0, 0)$ to (n, n) which stay below the diagonal (but can touch it).

Triangulating a convex $(n+2)$-gon (Euler): In how many ways can we triangulate a $(n+2)$-gon using $(n-1)$ 'diagonals' which do not cross each other.

The answer to the lattice problem and the polygon problem is C_n and the answer to the ballot problem is $C_n/\binom{2n}{n}$ where C_n is the Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

The first few Catalan numbers are $C_0 = C_1 = 1$ and $C_2 = 2$. In order to get the above formula for C_n we will first derive the recurrence

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-1} C_0, \quad n \geq 1,$$

and then use the recurrence to determine the generating function $C(X) = \sum_{n\geq 0} C_n X^n$. The coefficient of X^n in $C(X)$ will give the formula for C_n. The Ballot problem can be converted to the Lattice problem as follows. To a count of votes, we associate a lattice walk beginning at $(0, 0)$ and ending at (n, n) by taking a vote for A to a step 'right', and a vote for B as a step 'up'. A vote count in which A never trails B is a lattice walk which remains below diagonal. For the lattice problem, given a sub-diagonal walk W let (k, k) – for $0 \leq k \leq n-1$ – be the last point prior to (n, n) which lies on the given walk. Note that the walk can be broken as

$$W' \to (k+1, k) \to W' \to (n, n),$$

where W' is a lattice walk from $(0, 0)$ to (k, k) and W' is a lattice walk from $(k+1, k)$ to $(n, n-1)$. The number of such walks W' is C_k and the number of such walks W'' is the same as the walk from $(0, 0)$ to $(n, n-1) - (k+1, k) = (n-k-1, n-k-1)$ which is C_{n-k-1}. Thus we get the desired

recurrence $C_n = \sum_{k=0}^{n-1} C_k C_{n-k-1}$.

To obtain the same recurrence for the polygon problem, label the vertices as $v_0, \ldots, v_{n+1}$ and treat the edge $v_n v_{n+1}$ as the 'base' of the polygon. Given a triangulation of the polygon, the base is part of a triangle $v_n v_i v_{n+1}$ for some $0 \leq i \leq n-1$. To the left of this triangle is a triangulated $(n-i+1)$-gon (with vertices $v_i, v_{i+1}, \ldots, v_n$) and to the right of this triangle is a triangulated $(i+2)$-gon (with vertices $v_{n+1}, v_0, \ldots, v_i$). Thus we again get $C_n = \sum_{i=0}^{n-1} C_i C_{n-i-1}$ as desired.

In terms of the generating function $C(X)$, the recurrence relation can be simply stated as saying that the coefficient of X^{n-1} in $C(X)^2$ for $n \geq 1$ is the coefficient of X^n in $C(X)$. In other words

$$XC(X)^2 = C(X) - 1.$$

By completing squares, we get

$$(2XC(X) - 1) = \pm(1-4X)^{1/2} = \pm \sum_{n \geq 0} \tbinom{1/2}{n} (-4)^n X^n.$$

Since the constant term on the left side is -1, we must take $\pm$ in the right side to be $-$ rather than $+$. Therefore

$$C(X) = \sum_{n \geq 1} \tbinom{1/2}{n} (-1)^{n-1} 2^{2n-1} X^{n-1} = \textstyle\sum_{j \geq 0} \tbinom{1/2}{j+1} (-1)^j 2^{2j+1} X^j.$$

Since

$$\binom{1/2}{j+1} = (-1)^j \frac{1 \cdot 3 \cdot \ldots (2j-1)}{(j+1)! 2^{j+1}} = (-1)^j \frac{2j!}{j!(j+1)! 2^{2j+1}},$$

we get

$$C_n = \binom{1/2}{n+1} (-1)^j 2^{2n+1} = \frac{1}{n+1} \binom{2n}{n}.$$

25. Assignment 7

(1) Solve the recurrence relation $h_n = 8h_{n-1} - 16h_{n-2}$ for $n \geq 2$. Take initial values $h_0 = -1$ and $h_1 = 0$.
(Ans: $h(x) = (8x-1)/(1-4x)^2 = \sum_{n=0}^{\infty} 4^n(n-1)x^n$)

(2) Solve the non-homogeneous recurrence relation $h_n = 2h_{n-1} + n,\ n \geq 1$ with $h_0 = 1$.
(Ans: $h(X) = \sum_{n=0}^{\infty} h_n X^n = (1 + X(1-X)^{-2})/(1-2X)$ and $h_n = 3 \cdot 2^n - n - 2$)

(3) Let h_n be the size of the set $\mathcal{M}_n$ consisting of $2 \times n$ matrices of the form $\left(\begin{smallmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{smallmatrix}\right)$ with entries from $\{1, 2, \ldots, 2n\}$ (with no repetition) with rows and columns strictly increasing. Prove that $h_n = C_n$ (the n-th Catalan number).

Solution 1: Define $h_0 = 1$. For a matrix $M \in \mathcal{M}_n$, we note that $y_j \geq 2j$ because y_j is the largest member of $\{x_1, \ldots, x_j, y_1, \ldots, y_j\}$. Similarly, $x_j \leq 2j-1$ because x_j is the least member of $\{x_j, \ldots, x_n, y_j, \ldots, y_n\}$. Given $M \in \mathcal{M}_n$, let j be the largest integer less than n such that $y_j = 2j$. We then have for $n \geq 1$, the recurrence $h_n = h_0\tilde{h}_n + h_1\tilde{h}_{n-1} + \cdots + h_n\tilde{h}_0$ where $\tilde{h}_n$ is the size of the subset $\tilde{\mathcal{M}}_n \subset \mathcal{M}_n$ consisting of matrices with the additional property that $y_j > 2j$ for $1 \leq j \leq n-1$. For $\tilde{M} \in \tilde{\mathcal{M}}_n$, we note that the associated matrix $\left(\begin{smallmatrix} x_2-1 & x_3-1 & \cdots & x_n-1 \\ y_1-1 & y_2 & \cdots & y_{n-1} \end{smallmatrix}\right)$ is in $\mathcal{M}_{n-1}$: The columns are increasing because $x_j - 1 \leq 2j-2 < y_{j-1}$. Thus, $\tilde{h}_n = h_{n-1}$. We have shown that h_n's obey the same recurrence as C_n and $h_0 = C_0 = 1$ and hence $h_n = C_n$)

Solution 2: A Dyck path is a path in the first quadrant of the integer lattice from $(0,0)$ to $(2n, 0)$ using n steps of the form $(x, y) \mapsto (x+1, y+1)$ (denoted $\nearrow$) and n steps of the form $(x, y) \mapsto (x+1, y-1)$ (denoted $\searrow$). We have seen that the number of Dyck paths is the Catalan number C_n. For a Dyck path, we associate a matrix $M \in \mathcal{M}$: the first row consists of integers i such that the i-th step is $\nearrow$ and the second row consists of integers i such that the i-th step is $\searrow$. Note that x_i is the step number of the i-th $\nearrow$ step and y_i is the step number of the i-th $\searrow$ step. At the end of x_i steps the location is $(x_i, 2i - x_i)$ which must be strictly above the x-axis, and hence which $x_i \leq 2i-1$. At the end of y_i steps the location is $(y_i, y_i - 2i)$ which must be on or above the x-axis, and hence which $y_i \geq 2i$. This shows that M has increasing columns. Reversing the process we can define a Dyck path to a matrix $M \in \mathcal{M}$: The i-th step is $\nearrow$ or $\searrow$ according as i appears in the first or second row of M). The maps are inverses of each other.

(4) Show that in $\mathbb{R}[[X]]$, we have:

$$(1-4X)^{-1/2} = \sum_{n \geq 0} \binom{2n}{n} X^n \qquad \text{and} \qquad \frac{1 - \sqrt{1-4X}}{2X} = \sum_{n \geq 0} \frac{1}{n+1}\binom{2n}{n} X^n.$$

Using this prove the identity:

$$\binom{2n+1}{n} = \frac{1}{2}\binom{2n+2}{n+1} = \sum_{j=0}^{n} \frac{1}{j+1}\binom{2j}{j}\binom{2n-2j}{n-j}$$

26. Quiz-3 : 14-Oct 2019

(1) (7 points) Prove that two consecutive Fibonacci numbers are relatively co-prime. In other words $\gcd(f_n, f_{n+1}) = 1$ for all $n \geq 0$.
Ans: Clearly $\gcd(f_0, f_1) = \gcd(1, 1) = 1$. Assume $\gcd(f_k, f_{k+1}) = 1$ for $k \leq n$. We have

$$\gcd(f_{n+1}, f_{n+2}) = \gcd(f_{n+1}, f_{n+1} + f_n) = \gcd(f_n, f_{n+1}) = 1.$$

(2) (7 points) Suppose you deposit ₹5000 in a bank account that pays 6% interest at the end of each year (compounded annually). From the second year onwards, at the beginning of each year you deposit ₹1000. Let h_n be the amount in your account after n years (so $h_0 =$ ₹5000). Determine the generating function $h(x) = \sum_{n=0}^{\infty} h_n x^n$ and find a formula for h_n .
Ans: We have $h_0 = 5000$ and $h_n = 1000 + 1.06h_{n-1},\ n \geq 1$. Consider the generating function $h(x) = \sum_{n\geq 0} h_n x^n$. We have:

$$h(x) = 5000 + \sum_{n\geq 1} x^n(1000 + 1.06h_{n-1}) = 4000 + 1000(\sum_{n\geq 0} x^n) + 1.06xh(x)$$

Thus

$$h(x) = 1000(1 - 1.06x)^{-1}[4 + (1 - x)^{-1}] = 1000\left(\sum_{n\geq 0}(1.06x)^n\right) \cdot \left(5 + \sum_{n\geq 1} x^n\right).$$

In particular

$$h_n = 1000\left(5 \cdot (1.06)^n + \sum_{k=1}^{n}(1.06)^{n-k}\right) = 1000\left(5(1.06)^n + \frac{1.06^n - 1}{0.06}\right)$$

(3) (6 points) Let V_m be the vector space of polynomials with real coefficients of degree at most m in one variable. Consider the basis $E_k(X) = X(X-1)\dots(X-k+1)$ for $0 \leq k \leq m$ (with $E_0(X) = 1$) for V_m. Express the standard basis $e_k(X) = X^k$ for $0 \leq k \leq m$ in terms of the basis $E_0, \dots, E_m$ and the Stirling number of second kind. (with proof)
Ans: Given $m \leq n$, the total number m^n of functions from $\{1, \dots, n\}$ to $\{1, \dots, m\}$ equals the sum over k (as $k$ranges from 0 to m) of the number of ways choosing a size k subset of $\{1, \dots, m\}$ and then choosing a surjective map from $\{1, \dots, n\}$ to this subset. Hence:

$$m^n = \sum_{k=0}^{m} \binom{m}{k} k! S_{n,k}.$$

The sum on k in the right side can be taken as going from 0 to n instead of 0 to m (because $\binom{m}{k} = 0$ when $k > m$). In particular the polynomial

$$X^n - \sum_{k=0}^{n} S_{n,k}\, E_k(X),$$

of degree at most n has $n + 1$ roots namely $0, 1, \dots, m$. This shows that

$$e_j(X) = \sum_{i=0}^{j} S_{j,i} E_i(X).$$

27. Stirling numbers of the First kind

Stirling numbers of the first kind $c_{n,k}$ is the number of ways to arrange n people into k non-empty circles. The recurrence obeyed by $c_{n,k}$ is

$$c_{n+1,k} = c_{n,k-1} + nc_{n,k}.$$

This is because the first n of the $n+1$ people can be arranged in either i) $k-1$ circles or ii) k circles. In case i), $n+1$-st person forms a circle all by him/herself. In case ii) the $n+1$-st person has to be inserted into one of the k circles, and there are n ways to insert. Let $C_k(X) = \sum_{n\geq k} c_{n,k}X^n/n!$ be the exponential generating function of the numbers $c_{n,k}$. We note that $c_{n,0} = \delta_{0,n}$ and hence $C_0(X) = 1$. We also consider a generating function $C(X,Y) \in \mathbb{R}[[X,Y]]$ defined by

$$C(X,Y) = \sum_{k\geq 0} C_k(X)Y^k$$

The recurrence above for $c_{n+1,k}$ gives

$$\tfrac{d}{dX}C_k(X) = \tfrac{X^{k-1}}{k-1!} + \sum_{m\geq k} \tfrac{X^m}{m!}(c_{m,k-1} + mc_{m,k}) = C_{k-1}(X) + X\tfrac{d}{dX}C_k(X).$$

Thus

$$\tfrac{d}{dX}C_k(X) = \tfrac{C_{k-1}(X)}{1-X}.$$

This in turn gives the differential equation

$$\tfrac{\partial}{\partial X}C(X,Y) - \tfrac{Y}{1-X}C(X,Y) = 0.$$

As usual, if we have an element of $h(X,Y) = \mathbb{R}[[X,Y]]$ with zero constant term (i.e. $h(0,0) = 0$) and such that $\frac{\partial h}{\partial X} = -\frac{Y}{1-X}$, then we can write the above equation as

$$\exp(-h)\,\tfrac{\partial}{\partial X}(C(X,Y)\exp(h)) = 0.$$

We can take

$$h(X,Y) = Y\log(1-X) = -Y(X + X^2/2 + X^3/3 + \dots).$$

We note that

$$\exp(h) = (1-X)^Y = \sum_{n\geq 0}\binom{Y}{n}(-X)^n = \sum_{n\geq 0} E_n(Y)(-1)^n X^n/n!$$

where $E_n(Y) = Y(Y-1)\dots(Y-n+1)$. We leave it as an exercise to check that the product of two elements of $\mathbb{R}[[X,Y]]$ is zero if and only if one of the factors is zero. Since the element $\exp(h) = \sum_{n\geq 0} E_n(Y)(-1)^n X^n/n!$ is not the zero element of $\mathbb{R}[[X,Y]]$, it follows that $\frac{\partial}{\partial X}(C(X,Y)\exp(h)) = 0$. As an exercise check that $\frac{\partial}{\partial X}f = 0$ for an element $f \in \mathbb{R}[[X,Y]]$ if and only if $f \in \mathbb{R}[[Y]]$. Therefore we conclude that $C(X,Y)\exp(h) \in \mathbb{R}[[Y]]$. Therefore

$$C(X,Y)(1-X)^Y = C(0,Y)(1-0)^Y = \sum_{k\geq 0} C_k(0)Y^k = \sum_{k\geq 0} C_{0,k}Y^k = C_{0,0} = 1.$$

Finally, we get

$$C(X,Y) = (1-X)^{-Y}.$$

We can expand $C(X,Y)$ in two ways

$$\sum_{n\geq 0} E_n(-Y)(-1)^n X^n/n! = C(X,Y) = \sum_{k\geq 0} Y^k(-\log(1-X))^k/k!$$

where the first expansion has been noted above, and the second follows by writing $(1-X)^{-Y}$ as $\exp(-Y\log(1-X))$ This shows that

$$C_k(X) = (-\log(1-X))^k/k!$$

Moreover the expansion

$$\sum_{n\geq 0} E_n(-Y)(-1)^n X^n/n! = C(X,Y) = \sum_{n\geq 0} X^n/n!\,(\sum_{k\leq n} C_{n,k}Y^k)$$

yields the identity (with $Z=-Y$)

$$E_n(Z) = \sum_{k\leq n}(-1)^{n-k}C_{n,k}Z^k$$

Therefore, just as the Stirling numbers $S_{n,k}$ of the second kind were change of basis coefficients expressing e_j's in terms of E_i's, we see that the Stirling numbers $C_{n,k}$ of the first kind (multiplied by $(-1)^{n-k}$) are change of basis coefficients expressing E_j's in terms of e_i's in the vector space of polynomials of degree at most m in one variable Z.

28. Partition numbers

The analogue of Bell numbers for Stirling numbers of the first kind, i.e. $\sum_k C_{n,k}$ works out to be just $n!$. To see this we compare the coefficient of X^n in the first and last power series in the equation below:

$$\sum_{i\geq 0} X^i = (1-X)^{-1} = C(X,1) = \sum_{n\geq 0} X^n/n!\,(\sum_{k\leq n} C_{n,k}).$$

In combinatorial terms, the number of ways of arranging n people into circles is $n!$. We note that $n!$ is also the number of permutations of the same n people, and in fact there is a natural bijection between $\mathrm{Perm}(S)$ the set of permutations of a set S of size n, and the set $\mathrm{Circ}(S)$ of arranging S into circles (where the circles are not labeled). Given $\sigma \in \mathrm{Perm}(S)$ we will associate an element $\sigma^\dagger \in \mathrm{Circ}(S)$. Similarly, given $\tau \in \mathrm{Circ}(S)$ we will associate an element $\tau_\dagger \in \mathrm{Perm}(S)$, and we will show that these maps are inverses of each other, thus defining a bijection between $\mathrm{Perm}(S)$ and $\mathrm{Circ}(S)$.

Let $S = \{1, 2, \ldots, n\}$. Given $\sigma \in \mathrm{Perm}(S)$ and $i \in \{1, \ldots, n\}$ the orbit of i is the subset of S given by $\{i, \sigma(i), \sigma^2(i), \ldots\}$. Since this is a finite set there must be natural numbers k, m such that $\sigma^k(i) = \sigma^{k+m}(i)$. Since σ is a bijection, it follows that $\sigma^m(i) = i$. Let m be the smallest integer with this property. It follows that the orbit of i has size m. Note that if j is in the orbit of i, then the orbit of j is the same as the orbit of i. In other words two distinct orbits do not overlap. Moreover any $j \in S$ is in its own orbit. We write the orbit as a circle with the clockwise traversal of the circle corresponding to the sequence $i, \sigma(i), \sigma^2(i), \ldots$. Thus, we have obtained an arrangement S into circles (namely the orbits), which we define to be $\sigma^\dagger$. Given an element $\tau \in \mathrm{Circ}(S)$, we define a permutation $\tau_\dagger$ of S by defining $\sigma(i)$ (for each $i \in S$) to be the successor of i while traversing the circle clockwise. (in particular, if the circle containing i has no other member except i, then σ fixes i). Starting with $\sigma \in \mathrm{Perm}(S)$, it is clear that the element of $\mathrm{Perm}(S)$ defined by $(\sigma^\dagger)_\dagger$ equals σ. Similarly starting with a $\tau \in \mathrm{Circ}(S)$, the element of $\mathrm{Circ}(S)$ defined by $(\tau_\dagger)^\dagger$ equals τ.

Partitions

We now study the analogue of Bell numbers when the set $S = \{1, \ldots, n\}$ is replaced by the multiset $\{n \cdot a\}$. Let $p_0 = 1$ and let p_n denote the ways to partition the multiset $\{n \cdot a\}$ into any number of parts. In other words the number of ways to write $n = n_1 + n_2 + \cdots + n_k$ where $k \in \{1, \ldots, n\}$ and each $n_i > 0$. Suppose there are a_i parts of size i (where we allow $a_i = 0$, then clearly p_n is the number of solutions $a_1, \ldots, a_n$ in non-negative integers to the equation

$$1a_1 + 2a_2 + \cdots + na_n = n.$$

We note that the coefficient of X^n in the infinite product

$$(\sum_{i\geq 0} X^i)(\sum_{i\geq 0} X^{2i}) \ldots (\sum_{i\geq 0} X^{ji}) \ldots$$

is also the number of solution to the above equation. Thus the generating function

$$p(X) = \sum_{n\geq 0} p_n X^n = \prod_{i=1}^{\infty} \frac{1}{1-X^i}.$$

Note that the infinite product makes sense because p_n is the coefficient of X^n in the finite product

$$\prod_{j=1}^{n}(\sum_{i\geq 0} X^{ji}) = \prod_{i=1}^{n}\frac{1}{1-X^i}.$$

Similarly, the coefficient of X^n in the infinite product $\prod_{i=1}^{\infty}(1-X^i)$ actually equals the coefficient of X^n in the finite product $\prod_{i=1}^{n}(1-X^i)$. Since the coefficient of X^n in the product

$$\prod_{i=1}^{\infty}\frac{1}{1-X^i}\cdot\prod_{i=1}^{\infty}(1-X^i)$$

is the same as the coefficient of X^n in the finite product

$$\prod_{i=1}^{n}\frac{1}{1-X^i}\cdot\prod_{i=1}^{n}(1-X^i) = 1,$$

it follows that the reciprocal of $p(X)$ is the infinite product $\prod_{i=1}^{\infty}(1-X^i)$. Euler calculated the first 50 or more terms in the expansion of this series and observed a pattern which he was able to prove many years later (around 1783 AD). This is Euler's famous Pentagonal number theorem. A pentagonal number is an integer of the form $(3m^2 \pm m)/2$.

Theorem. *(Euler's pentagonal number theorem) The coefficient of X^n in $1/p(X)$ is $(-1)^m$ if $n = (3m^2 \pm m)/2$ and is zero if n is not a pentagonal number:*

$$\prod_{i=1}^{\infty}(1-X^i) = 1/p(X) = \sum_{m=0}^{\infty}(-1)^m(X^{(3m^2-m)/2} + X^{(3m^2+m)/2}).$$

An elementary combinatorial proof is certainly within the scope of this course, but will not give it here. The interested student can see look up 'Franklin's proof of Euler's pentagonal number theorem'.

Another famous partition identity of Euler is:

Theorem. *(Euler) Let $p_o(n)$ be the number of partitions of n into parts of odd size, and let $p_d(n)$ be the number of partitions of n into parts of distinct sizes. Then $p_o(n) = p_d(n)$.*

Proof. We prove this by showing that the generating functions $p_o(X) = \sum_{n\geq 0} p_o(n)X^n$ and $p_d(X) = \sum_{n\geq 0} p_d(n)X^n$ are equal. It is easy to see that

$$p_d(X) = \prod_{i=1}^{\infty}(1+X^i),$$

because the coefficient of X^n in this infinite product is the number of solutions to the above equation $\sum_{i=1}^{n} ia_i = n)$ with $a_i \in \{0,1\}$. Thus

$$p_d(X) = \frac{\prod_{i=1}^{\infty}(1-X^{2i})}{\prod_{i=1}^{\infty}(1-X^i)} = \frac{1}{\prod_{i=1}^{\infty}(1-X^{2i-1})} = p_o(X)$$

□

The subject of partitions is a vast and deep subject into which we do not go further. Ramanujan observed and proved that : i) $5|p(n)$ if $n \equiv 4 \bmod 5$, ii) $7|p(n)$ if $n \equiv 5 \bmod 7$, iii) $11|p(n)$ if $n \equiv 6 \bmod 11$. The pentagonal number's theorem is closely related to the Dedekind Eta function in the

theory of modular forms.

Remark (for those familiar with group theory): Let $S = \{1, \ldots, n\}$. Recall the bijection $\sigma \mapsto \sigma^\dagger$ from $\mathrm{Perm}(S)$ to $\mathrm{Circ}(S)$. Given $\sigma \in \mathrm{Perm}(S)$ let L_σ be the multiset consisting of the lengths of the circles in $\sigma^\dagger$. Note that L_σ is just a partition of n. We introduce an equivalence relation on $\mathrm{Perm}(S)$ be defining σ and τ to be equivalent if $L_\sigma = L_\tau$. It is clear that the number of equivalence classes is p_n the number of partitions of n. The significance of this is that the equivalence relation on $\mathrm{Perm}(S)$ we defined has an interpretation in the theory the symmetric group S_n (i.e. $\mathrm{Perm}(S)$ viewed as a group): two permutations are equivalent if they are *conjugate* in the sense of group theory. The equivalence classes are the *conjugacy classes* in the symmetric group. The conclusion is that p_n is the number of conjugacy classes in the symmetric group S_n.

29. Counting labeled trees

Number of labeled trees on n vertices A graph consists of a set V consisting of n vertices, and a subset E of the $\binom{n}{2}$ possible edges between these vertices. A path $v_1v_2\dots v_k$ in the graph consists of a sequence of distinct vertices $v_1,\dots,v_k$ such that there is an edge between v_i and v_{i+1} for each $1\leq i\leq k-1$. A *tree* is a graph in which there is a unique path between any two vertices. Clearly, a tree with one vertex has no edges. So, we now assume $n>1$. Pick any vertex, say v_n. Let r be the number of edges emanating from v_n. If $r=0$, then there can be no path starting at v_n which is not the case. Therefore $1\leq r\leq n-1$. Let $w_1,\dots,w_r$ be the neighbours of v_n. If we remove v_n and the r edges emanating from v_n, then we are left with a graph G' on $n-1$ vertices. In G' , if a vertex u can be connected to w_i then it cannot be connected to w_j for $j\neq i$: for otherwise, in the original graph, we would have two distinct paths between w_i and w_j, namely $w_i-v_n-w_j$ and the path PP' where P and P' are the paths in the new graph from w_i to u and u to w_j. Moreover, any vertex u is connected to w_i for some i, because otherwise u cannot be connected to v_n in the original graph G. Thus G' decomposes into a disjoint union of graphs $G_1,\dots,G_r$ with the property that if $u,v\in G_i$ then the unique path joining u and v in G actually lies in G_i. In other words, each G_i is a tree. We claim that the number of edges in a tree is one less than the number of vertices ν. This is true if $\nu=1$. Assume, inductively that the result is true for less than ν vertices. In particular , in the graph G' above the number of vertices exceeds the number of edges by r. In G, there is one more vertex v_n and r more edges (joining v_n to $w_1,\dots,w_r$). Therefore the number of vertices of G exceeds the number of edges by $r+1-r=1$.

Let us label the vertices as $v_1,v_2,\dots,v_n$. How many distinct trees are there on this set of vertices? Let t_n denote this number. Clearly $t_1=1$ (no edges), and $t_2=1$ (one edge) . If $n=3$, of the 3 vertices one of them has 2 edges and the other two have 1 edge emanating. The vertex with 2 edges can be chosen in $\binom{3}{1}=3$ ways, and hence $t_3=3$.

Theorem. *(Cayley 1885)* $t_n=n^{n-2}$.

There are several proofs of this theorem. We give a proof using exponential generating functions. A rooted labeled tree is a tree in which a particular vertex has been singled out. Let τ_n denote the number of rooted labeled trees on n vertices. Clearly $\tau_n=nt_n$. Our goal is to show that $\tau_n=n^{n-1}$. This will show $t_n=n^{n-2}$. It is easy to see that $\tau_1=t_1=1$. We define a formal power series:

$$T(x)=\sum_{n=1}^{\infty}\tau_nX^n/n! \tag{12}$$

We now assume $n\geq 2$. As above, let $1\leq r\leq n-1$ be the number of edges emanating from v_n. We recall that the graph G' is an unordered collection of r disjoint trees (on total vertex set $\{v_1,\dots,v_{n-1}\}$) and each of these r trees has a root (the unique element of this tree which is a neigbour of v_n). Let us order these trees as $G_1,\dots,G_r$. There are $r!$ ways to do so. Let w_i denote the root of G_i. Let ν_i denote the number of vertices of G_i. The number of ways of choosing the vertex sets of $G_1,\dots,G_r$ is the multinomial coefficient $\binom{n-1}{\nu_1,\dots,\nu_r}$. Therefore, for $n\geq 2$, we get:

$$\tau_n/n=\sum_{r=1}^{n-1}\tfrac{1}{r!}\sum_{\nu_1+\dots+\nu_r=n-1}\frac{n-1!}{\nu_1!\nu_2!\dots\nu_r!}\,\tau_{\nu_1}\tau_{\nu_2}\dots\tau_{\nu_r}.$$

Thus, for $n \geq 2$, we conclude that $\tau_n/n!$ is the coefficient of X^{n-1} in $\sum_{r=1}^{n-1} T(x)^r/r!$, or also

$$e^{T(X)} - 1 = \sum_{r=1}^{n-1} T(X)^r/r!,$$

because $T(X)^r$ has no terms X^i for $i < r$. Thus, we get:

$$-1 + T(X)/X = \sum_{n=2}^{\infty} \tau_n X^{n-1}/n! = \exp(T(x)) - 1.$$

Thus we have obtain the functional equation:

$$T(x)e^{-T(x)} = X \tag{13}$$

We will now solve this equation, to show $\tau_n = n^{n-1}$. Let $f(X) = Xe^{-X}$. Note that the composition $f \circ T = X$. According to the Lagrange inversion theorem presented below, we have

$$k \cdot \tfrac{\tau_k}{k!} = [X^{k-1}]e^{kX} = \tfrac{k^{k-1}}{(k-1)!}.$$

Therefore, $\tau_k = k^{k-1}$.

Theorem. *(Lagrange $\sim$ 1770). Let $f(X) = \sum_{n\geq 0} f_n X^n \in \mathbb{R}[[X]]$ with $f_0 = 0, f_1 \neq 0$. There is a unique* compositional inverse *$g(X) = \sum_{n\geq 0} g_n X^n \in \mathbb{R}[[X]]$ with $g_0 = 0, g_1 \neq 0$ satisfying $f(g(X)) = X$. The coefficients of g are given by :*

$$k[X^k]g(X) = \mathrm{Res}(1/f^k)$$

We will explain the term Res used above. Before, doing that we note that the existence of $g(X)$ is not difficult: If we plug in an unknown series $g(X) = g_1X + g_2X^2 + \dots$ into the equation $f(g(X)) = X$, we get:

$$f_1(g_1X + g_2X^2 + \dots) + f_2(g_1X + g_2X^2 + \dots)^2 + \cdots = X(f_1g_1) + X^2(f_1g_2 + f_2g_1^2) + \dots$$

This gives $g_1 = 1/f_1$, $g_2 = -f_2g_1^2/f_1$. Similarly, having solved for $g_1, \dots, g_k$ we can determine g_{k+1} in terms of $g_1, \dots, g_k$. Thus the exact formula for the coefficient g_k is the important part of this theorem.

We need to enlarge our ring $\mathbb{R}[[X]]$ of formal power series to include negative powers of X. The ring of *Laurent series* denoted $\mathbb{R}((X))$ consists of the set of formal series

$$\{\sum_{n\geq k} a_n X^n \mid k \in \mathbb{Z}\}$$

with finitely many negative powers of X. The operations of addition and multiplication of Laurent series are clear. Note that any non-zero element $f(X)$ of of $\mathbb{R}((X))$ can be written uniquely as $X^m f^\dagger(X)$ where $m \in \mathbb{Z}$ (possibly negative) and $f^\dagger(X) \in \mathbb{R}[[X]]$ and has nonzero constant term. The integer m is called the order of f and denoted $\mathrm{ord}(f)$. At this point we note that the reciprocal of f already exists in $\mathbb{R}((X))$: it is $X^{-m}(1/f^\dagger(X))$ where the reciprocal $1/f^\dagger(X)$ exists in $\mathbb{R}[[X]]$ because $f^\dagger(X)$ has non-zero constant term. Since every non-zero element of $\mathbb{R}((X))$ has a reciprocal, it follows that $\mathbb{R}((X))$ is a field.

We can extend the differentiation operation D on $\mathbb{R}[[X]]$ to $\mathbb{R}((X))$: we just define $DX^n = nX^{n-1}$ for all integers n and differentiate a Laurent series term-wise. We note that

$$D : \mathbb{R}((X)) \to \mathbb{R}((X))$$

is a linear transformation whose kernel is clearly the constants (i.e. power series with only constant term). Note that the image of D consists of all Laurent series in which the coefficient of X^{-1} is zero. This is because $DX^n = nX^{n-1}$ can never be a nonzero scalar multiple of X^{-1}.

We define the *residue* $\mathrm{Res}(f)$ of a Laurent series f to be the coefficient of X^{-1} in f. Note that

$$\mathrm{Res} : \mathbb{R}((X)) \to \mathbb{R}$$

is a surjective linear transformation (of vector spaces over $\mathbb{R}$). As noted above:

$$\ker(\mathrm{Res}) = \mathrm{Im}(D).$$

In other words $\mathrm{Res}(f') = 0$ where $f' = Df$. As a corollary we note that:

$$\mathrm{Res}(fg') = -\mathrm{Res}(f'g).$$

Another property of Res is:

$$[X^k]f = \mathrm{Res}(X^{-k-1}f),$$

which follows from the definition of Res.

A third property of Res is that for a nonzero f we have:

$$\mathrm{Res}(f'/f) = \mathrm{ord}(f).$$

To see this we write $f = X^{\mathrm{ord}(f)}\, f^\dagger$ and note that

$$f'/f = \mathrm{ord}(f)/X + Df^\dagger/f^\dagger.$$

The first term has residue equal to $\mathrm{ord}(f)$ and the second term is in $\mathbb{R}[[X]]$ and hence has residue zero.

The most important property of Res (for the proof of Lagrange's theorem) is

$$\mathrm{Res}((f \circ g)(X) \cdot g'(X)) = \mathrm{Res}(f)\mathrm{ord}(g)$$

where $g(X) \in \mathbb{R}[[X]]$ with zero constant term (in other words $\mathrm{ord}(g) > 0$). To see this we write $f = \mathrm{Res}(f)X^{-1} + f_1(X)$, where $f_1 \in \mathrm{Ker}(Res)$. Since $\mathrm{Ker}(Res) = \mathrm{Im}(D)$ it follows that $f_1 = f_2'$ for some Laurent series f_2. We have

$$(f \circ g)(X) \cdot g'(X) = \mathrm{Res}(f)\tfrac{g'}{g} + (f_2 \circ g)'.$$

Thus $\mathrm{Res}((f \circ g)(X) \cdot g'(X)) = \mathrm{Res}(f)\mathrm{ord}(g)$.

Returning to the proof of Lagrange's theorem, let $g(X)$ be the compositional inverse of $f(X)$. Note that $\mathrm{ord}(f) = \mathrm{ord}(g) = 1$.

$$k[X^k]g = k\mathrm{Res}(X^{-k-1}g) = -\mathrm{Res}((DX^{-k})g) = \mathrm{Res}(X^{-k}g') = \mathrm{Res}((f^{-k} \circ g)g') = \mathrm{Res}(f^{-k})\mathrm{ord}(g)$$

Since $\mathrm{ord}(g) = 1$, we conclude that

$$k[X^k]g = \mathrm{Res}(f^{-k}).$$

Remarks: Our presentation of Lagrange's theorem is based on https://en.wikipedia.org/wiki/Formal_power_series#The_Lagrange_inversion_formula.

The function $W(x) = -T(-x)$ is the power series at $x = 0$ of the *Lambert W-function* which occurs in many contexts in physics and chemistry. See https://en.wikipedia.org/wiki/Lambert_W_function.

30. Assignment 8

(1) Consider a rooted tree on n vertices. We divide the vertex set as $V_0, V_1, \dots$ where V_0 consists of the root and V_i consists of those vertices such that the unique path from the root to this vertex has i edges. We say that an unlabeled rooted tree is a *binary tree* if for each for each i and for each $v \in V_i$ set of neighbours of v in V_{i+1} –called the children of v– is at most 2 in number. The children are partitioned into two (possibly empty) parts called left children and right children. If a vertex has only one child, then this child can be left or right. If a vertex has two children, then one child is left and the other is right. Show that the number of binary trees on n vertices is the Catalan number C_n. Example: for $n = 3$ the five binary trees are (where r is root):

$$r \to L \to L, \quad r \to L \to R, \quad r \to R \to L, \quad r \to R \to R, \quad r \to LR$$

(Ans: Let a_n be the number of such binary trees. Define $a_0 = 1$. The left side 'branch' of the root is again a binary tree with j vertices, and the right side 'branch' of the root is a binary tree with $n - 1 - j$ vertices. Therefore, the numbers a_n satisfy the same recurrence as that of the Catalan numbers $a_n = a_0 a_{n-1} + a_1 a_{n-2} + \cdots + a_{n-1} a_0$ for $n \geq 1$ and $a_0 = C_0$. Therefore, $a_n = C_n$.)

(2) In this problem we will give a combinatorial proof of the following partition identity in $\mathbb{R}[[X, Y]]$:

$$\prod_{i=1}^{\infty} \frac{1}{1 - YX^i} = \sum_{k=0}^{\infty} Y^k \frac{X^k}{(1-X)(1-X^2)\dots(1-X^k)}$$

(a) We can expand the left side as a power series $\sum_{k,n} p_{n,k} X^n Y^k$. Interpret $p_{n,k}$ as the number of partitions of n into exactly k parts.

(b) We can expand the right side as a power series $\sum_{k,n} q_{n,k} X^n Y^k$. Interpret $q_{n,k}$ as the number of partitions of n whose largest part is k.

(c) Prove the identity by proving that $p_{n,k} = q_{n,k}$. Given a partition $n = n_1 + \cdots + n_k$ with $n_1 \geq n_2 \geq \cdots \geq n_k$, consider a $n \times n$ matrix with entries in $\{0, 1\}$. The i-th row has n_i ones followed by $(n - n_i)$ zeros. For $i > k$ the rows consist of zeros. Conversely to each such $n \times n$ matrix (with $\{0, 1\}$ entries, in which the number of ones in each row is non-increasing, and within each row the zeros trail the ones) we get a partition. Consider the transpose operation on such matrices.

(3) Consider the equation $XC(X)^2 = C(X) - 1$ for the generating function $\sum_{n \geq 0} C_n X^n$ of the Catalan numbers. Let $A(X) = C(X) - 1$. Convert the above equation to a functional equation for $A(X)$ of a form to which Lagrange inversion can be applied. Using Lagrange inversion obtain the formula $C_n = \frac{1}{n+1}\binom{2n}{n}$.

31. Philip Hall marriage theorem/Systems of Distinct Representatives

Systems of Distinct Representatives
Let X be a finite set, and let let $S_1, \ldots, S_m$ be a collection of subsets of X (we allow $S_i = S_j$ even if $i \neq j$). A *system of distinct representatives* (in short SDR) or a *transversal* for the collection $S_1, \ldots, S_m$ is a choice of m distinct elements $x_1, \ldots, x_m$ of X with the property that $x_i \in S_i$.

Examples:

- Let $X = \{1, \ldots, 5\}$ and $S_1 = \{1, 2, 3\}$, $S_2 = \{1, 4, 5\}$ and $S_3 = \{3, 5\}$. Here a transversal is given by $f(1) = 2, f(2) = 4, f(3) = 5$. Another transversal is $f(1) = 2, f(2) = 1, f(3) = 3$.
- Let S_1, S_2, S_3, S_4 be the family of subsets of the set $X = \{a, b, c, d, e\}$, defined by $S_1 = \{a, b, c\}, S_2 = S_4 = \{b, d\}, S_3 = \{a, b, d\}$. Here (c, b, a, d) is an SDR.

Marriage Problem Let X be a set of n men, and $A = \{W_1, \ldots, W_m\}$ a set of m women. Each woman W_i prefers a subset $S_i \subset X$. The marriage problem is to find a SDR, i.e. to find m distinct men $M_1, \ldots M_m$ such that for each $1 \leq i \leq m$, the man M_i is in the preferred list of W_i.

A necessary condition for existence of a SDR:
For a SDR to exist it is necessary that for each $B \subset \{1, \ldots, m\}$ we must have

$$| \cup_{i \in B} S_i| \geq |B|. \tag{14}$$

Indeed, if this condition fails for some $B \subset \{1, \ldots, m\}$ of size say j where $1 \leq j \leq m$, then it is impossible to find j distinct representatives for the $S_i, i \in B$ because the union of these S_i's itself has size less than j. It turns out that there is no other obstruction to finding an SDR. The following theorem is attributed to Philip Hall (1935) but D. Kőnig had also proved it earlier (in context of matching for bipartite graphs)

Theorem. *(Philip Hall marriage theorem) The necessary condition is sufficient.*

Proof. The proof is by induction on m. If $m = 1$ then we just have to pick an element from S_1 and this is possible because the condition (14) says that $|S_1| \geq 1$. Assume inductively that given any finite set Y and $k < m$ subsets of Y, then the condition (14) is also sufficient. Returning to the m subsets $S_1, \ldots, S_m$ of X, we consider two cases: i) either there exists $B \subset \{1, \ldots, m\}$ with $1 \leq |B| \leq m - 1$ for which equality holds in (14), or ii) for every $B \subset \{1, \ldots, m\}$ with $1 \leq |B| \leq m - 1$, inequality holds in (14).

In case i) suppose wlog that $|S_1 \cup \cdots \cup S_j| = j$. Since $j < m$ by inductive hypothesis, we get distinct elements $x_1, \ldots, x_j$ such that $x_i \in S_i$. Moreover $S_1 \cup \cdots \cup S_j$ must equal $\{x_1, \ldots, x_j\}$ as both sets have the same size j. Now let $Y = X \setminus \{x_1, \ldots, x_j\}$ and $A_i = S_{j+i} \setminus \{x_1, \ldots, x_j\}$ for $1 \leq i \leq m-j$. We claim that the condition (14) holds for this new collection: for $B \subset \{1, \ldots, m-j\}$ we have:

$$\cup_{i \in B} A_i = (\cup_{i \in B} S_{j+i}) \setminus \{x_1, \ldots, x_j\} = ((S_1 \cup \cdots \cup S_j) \cup (\cup_{i \in B} S_{j+i})) \setminus \{x_1, \ldots, x_j\}.$$

Since $(S_1 \cup \cdots \cup S_j) \cup (\cup_{i \in B} S_{j+i})$ has size at least $|B| + j$ by (14), it follows that

$$| \cup_{i \in B} A_i| = | ((S_1 \cup \cdots \cup S_j) \cup (\cup_{i \in B} S_{j+i})) \setminus \{x_1, \ldots, x_j\}| \geq |B|.$$

Since $m-j<m$, the inductive hypotheses gives an SDR $x_{j+1},\dots,x_m$ for $A_1,\dots,A_{m-j}$. Combining with $x_1,\dots,x_j$ we get an SDR $x_1,\dots,x_m$ for $S_1,\dots,S_m$.

Case ii): here $|\cup_{i\in B}S_i|>|B|$ for all $B\subset\{1,\dots,m\}$. Since $|S_m|>1$, pick $x_m\in S_m$ and consider $Y=X\setminus\{x_m\}$ and $A_i=S_i\setminus\{x_m\}$ for $1\leq i\leq m-1$. The condition (14) holds for this new collection because for $B\subset\{1,\dots,m-1\}$ we have:

$$|\cup_{i\in B}A_i|=|(\cup_{i\in B}S_i)\setminus\{x_m\}|>|B|-1\geq|B|.$$

Since $m-1<m$, the inductive hypotheses gives an SDR $x_1,\dots,x_{m-1}$ for $A_1,\dots,A_{m-1}$. Combining with x_m we get an SDR $x_1,\dots,x_m$ for $S_1,\dots,S_m$. □

First application of Hall's theorem

Lemma 9. *Suppose we have two partitions $\coprod_{i=1}^m A_i$ and $\coprod_{i=1}^m B_i$ of a set Y such that each of the sets A_i and B_j have the same size n. We claim that there exist $a_i\in A_i$ for $1\leq i\leq m$ and a permutation σ of $\{1,\dots,n\}$ such that $a_i\in B_{\sigma(i)}$ for $1\leq i\leq m$.*

Proof. Let $X=\{1,\dots,m\}$ and consider subset $S_1,\dots,S_m$ of X defined by

$$S_i=\{1\leq j\leq m: A_i\cap B_j\neq\emptyset\}.$$

Given $1\leq j\leq m$ and $B\subset\{1,\dots,m\}$ of size j we claim (14) holds: indeed $|\cup_{i\in B}S_i|$ equals the number of nonempty sets in the right side of

$$\amalg_{i\in B}A_i=\amalg_{i=1}^m(B_i\cap(\amalg_{i\in B}A_i))$$

Therefore the size of the set on the right side in this equation is at-most $n\cdot|\cup_{i\in B}S_i|$ (note each B_i has size n) where as the set on the left side has size $|B|n$. This establishes that $|\cup_{i\in B}S_i|\geq|B|$. By Hall's theorem, we get an SDR $\sigma(1),\dots,\sigma(m)$ of m distinct elements of $\{1,\dots,m\}$ (in other words a permutation of $\{1,\dots,m\}$) such that $\sigma(i)\in S_i$ which means that $A_i\cap B_{\sigma(i)}\neq\emptyset$. Pick $a_i\in A_i\cap B_{\sigma(i)}$. □

As a corollary, we can prove the following theorem related to group theory (this topic is optional):

Theorem. *Let H be a subgroup of a group G such that $[G:H]=m$ is finite. There exist $g_1,\dots,g_m\in G$ such that*

$$G=\cup_{i=1}^m g_iH=\cup_{i=1}^m Hg_i.$$

Proof. We will prove the theorem under the extra assumption that G is finite. In the Assignment, we will drop this assumption. Let $|H|=n$ and $|G|=mn$. We apply the previous lemma with $A_1,\dots,A_m$ and $B_1,\dots,B_m$ being respectively, the set of left cosets of H in G and the set of right cosets of H in G. By the lemma there exist distinct elements $a_1,\dots,a_m$ of G such that:

$$G=\amalg_{i=1}^m a_iH=\amalg_{i=1}^m Ha_i.$$

□

32. Birkhoff- von Neumann theorem on double stochastic matrices

Second application of Philip Hall theorem: The Birkhoff- von Neumann theorem.

We begin with some definitions.

- A vector $v \in \mathbb{R}^n$ is called a probability vector is $v_i \geq 0$ and $\sum v_i = 1$.
- A $n \times n$ real matrix A is called *Markov* if each of the n columns of A is a probability vector.
- A $n \times n$ real matrix A is called *doubly stochastic* if both A and A^T (the transpose of A) are Markov.
- A $n \times n$ permutation matrix A is a matrix with entries in $\{0, 1\}$ such that each row and column has exactly one non-zero entry. There are $n!$ permutation matrices. For each permutation σ of $\{1, \ldots, n\}$ the permutation matrix A_σ is defined by $A_{ij} = \delta_{\sigma(i),j}$. Note that a permutation matrix is doubly stochastic.
- If V is a real vector space and $v_1, \ldots, v_m \in V$ then a convex combination of $v_1, \ldots, v_m$ is a linear combination $t_1 v_1 + \cdots + t_m v_m$ with the property that $t_i \geq 0$ and $t_1 + \cdots + t_m = 1$.
- A convex combination of Markov matrices is Markov, and similarly a convex combination of doubly stochastic matrices is doubly stochastic.

Theorem. *(G. Birkhoff 1946, J. von Neumann 1953) Every $n \times n$ doubly stochastic matrix is a convex combination of $n \times n$ permutation matrices.*

Proof. The proof is by induction on the number m_A of non-zero entries of the doubly stochastic matrix A. Since each column of A must have a non-zero entry, it follows that $m_A \geq n$. In fact, if $m_A = n$ then it is clear that A is obtained by permuting the columns of the $n \times n$ identity matrix, i.e. A is a permutation matrix. Thus the base case of the induction, $m_A = n$ has been established. We now assume inductively that the theorem is true if $m_A < m$. Given a doubly stochastic matrix A with $m_A = m$, let $X = \{1, \ldots, n\}$ and define $S_1, \ldots, S_n \subset X$ by

$$S_i = \{j \in X : A_{ij} \neq 0\}.$$

For $B \subset X$ let $S_B = \cup_{i \in B} S_i$. We claim that $|S_B| \geq |B|$ for all $B \subset X$:

$$|B| = \sum_{i \in B} 1 = \sum_{i \in B} \sum_{j=1}^{n} A_{ij} = \sum_{i \in B} \sum_{j \in S_B} A_{ij} = \sum_{j \in S_B} \sum_{i \in B} A_{ij} \leq \sum_{j \in S_B} \sum_{i=1}^{n} A_{ij} = \sum_{j \in S_B} 1 = |S_B|.$$

Therefore, by Philip Hall theorem there exists a permutation σ of $X = \{1, \ldots, n\}$ such that $\sigma(i) \in S_i$, or in other words $A_{i,\sigma(i)} > 0$. Let $\lambda = \min\{A_{i,\sigma(i)} : 1 \leq i \leq n\}$. Note that $0 < \lambda \leq 1$ and that $\lambda = 1$ if and only if $A = P_\sigma$. Therefore, if $\lambda = 1$ then we have already expressed A as a convex combination of permutation matrices. If $\lambda < 1$, we define

$$A' = \frac{A - \lambda P_\sigma}{1 - \lambda}.$$

Note that the sum of the j-th column of A' equals

$$\frac{(\text{ sum of the } j\text{-th column of A}) - \lambda}{1 - \lambda} = 1$$

The same is true for the row-sums of A' also. Moreover $A'_{ij} \geq 0$, because:

$$A'_{ij} = \begin{cases} A_{ij}/(1-\lambda) & \text{if } j \neq \sigma(i) \\ \frac{A_{i\sigma(i)}-\lambda}{1-\lambda} & \text{if } j = \sigma(i) \end{cases} \geq 0.$$

Thus A' is doubly stochastic. Also, the number of non-zero entries of A exceeds the number of non-zero entries of A' by

$$\#\{i : A_{i,\sigma(i)} = \lambda\} \geq 1.$$

Thus the number of non-zero entries of A' is strictly less than the number of non-zero entries of A. By the inductive hypothesis, it follows that $A' = \sum_\tau t_\tau P_\tau$, $\sum_\tau t_\tau = 1$, where the sum runs over all permutations τ of $\{1, \ldots, n\}$, and where $t_\tau \geq 0$. It follows that

$$A = \lambda P_\sigma + \sum_\tau (1-\lambda) t_\tau P_\tau.$$

We note that the right side is indeed a convex combination of permutation matrices. □

33. Generalized Phillip Hall marriage theorem, Kőnig-Egerváry theorem

Let $S_1, \ldots, S_m$ be subsets of a finite set X. For a nonempty subset B of $\{1, \ldots, m\}$, let $S_B = \cup_{i \in B} S_i$. We recall the Philip Hall theorem: for an SDR of $S_1, \ldots, S_m$ to exist, the necessary condition

$$|S_B| \geq |B| \quad \forall B \subset \{1, \ldots, m\},$$

is also sufficient. There is a generalization of the theorem when the necessary condition does not hold. We define a *deficit* of $S_1, \ldots, S_m$ to capture the worst case of the violation of the necessary condition, i.e. the largest value of $|B| - |S_B|$ for which $|S_B| < |B|$:

$$\Delta = \max\{|B| - |S_B| \, : \, |S_B| < |B|\}$$

If the necessary condition is never violated, we define $\Delta = 0$. In other words

$$\Delta = \max_{B \subset \{1, \ldots, m\}} \quad \max\{0, |B| - |S_B|\}.$$

Theorem. *(Generalized Phillip Hall marriage theorem) The largest sub-collection of $S_1, \ldots, S_m$ for which an SDR exists is $m - \Delta$.*

Proof. Suppose a sub-collection of size $m - t$ of $S_1, \ldots, S_m$ has an SDR. We may assume this sub-collection is $S_1, \ldots, S_{m-t}$ and let $x_i \in S_i$ for $1 \leq i \leq m - t$ be an SDR . Enlarge the set X by adding t extra elements $\tilde{X} = X \amalg \{z_1, \ldots, z_t\}$, and define a collection of m subsets of $\tilde{X}$ by $\tilde{S}_i = S_i \amalg \{z_1, \ldots, z_t\}$. The collection $\tilde{S}_1, \ldots, \tilde{S}_m$ of subsets of $\tilde{X}$ has an SDR: take the SDR $x_i \in S_i \subset \tilde{S}_i$ for $1 \leq i \leq m - t$, and take $z_i \in \tilde{S}_i$ for $m - t + 1 \leq i \leq m$. Since the Hall condition is necessary, we must have for each non-empty $B \subset \{1, \ldots, m\}$:

$$|B| \leq |\cup_{i \in B} \tilde{S}_i| = t + |\cup_{i \in B} S_i|,$$

which shows that $\Delta \leq t$, i.e. $m - \Delta \geq m - t$.

Conversely, let $\tilde{X} = X \amalg \{z_1, \ldots, z_\Delta\}$ and consider the collection of m subsets of $\tilde{X}$ given by $\tilde{S}_i = S_i \amalg \{z_1, \ldots, z_\Delta\}$. For each non-empty $B \subset \{1, \ldots, m\}$, we have:

$$|\cup_{i \in B} \tilde{S}_i| = \Delta + |\cup_{i \in B} S_i| = |B| + [\, \Delta - (\, |B| - |\cup_{i \in B} S_i|\,)\,] \geq |B|,$$

where the last inequality follows from the definition of Δ. Since the Hall condition is satisfied, there exists an SDR $x_i \in \tilde{S}_i$ for $1 \leq i \leq m$. Note that $\tilde{S}_i = S_i \amalg \{z_1, \ldots, z_\Delta\}$ and hence at most Δ of these x_i's can be in $\{z_1, \ldots, z_\Delta\}$. It follows that a sub-collection of size at least $m - \Delta$ of $S_1, \ldots, S_m$ has an SDR.

Combining these two assertions,we conclude that the maximum possible size of a sub-collection of $S_1, \ldots, S_m$ which has an SDR is $m - \Delta$. □

The next theorem ($\sim$ 1914) stated in terms of $0, 1$-matrices is a another version of the generalized Philip Hall theorem.

Theorem 10. *(Kőnig-Egerváry theorem) Let A be a $m \times n$ matrix with $0, 1$ entries. Define a line of a $m \times n$ matrix to be a row or column. The maximum possible size of a subset of the set of 1's of A, no two of which are on the same line is equal to the minimum possible size of a subset of lines of A which together contain all the 1's of A.*

Proof. Let $X = \{1, \ldots, n\}$ and consider the collection of subsets $S_1, \ldots, S_m$ of X given by

$$S_i = \{j \in X : A_{ij} = 1\}.$$

Let B be a non-empty subset of $\{1, \ldots, m\}$ and suppose the sub-collection S_i for $i \in B$ has an SDR, given by $j_i \in S_i$ for $i \in B$. We note that the collection of 1's of A indexed by the entries $\{(i, j_i) : i \in B\}$ is a collection of 1's no two of which are on the same line. Conversely given such a collection of 1's, appearing in places $(i_1, j_1), (i_2, j_2), \ldots, (i_\mu, j_\mu)$ it follows that the $i_1 \ldots, i_\mu$ are distinct $j_1, \ldots, j_\mu$ are distinct. Thus the sub-collection $S_{i_1}, \ldots, S_{i_\mu}$ has an SDR. It follows that the maximum possible size of a collection of such 1's is the maximum possible size of a sub-collection of $S_1, \ldots, S_m$ which has an SDR.

Given a collection of lines of A which includes all the 1's of A, let $C \subset \{1, \ldots, m\}$ be the set of rows in this collection (it is possible that $C = \emptyset$). Let $B = \{1, \ldots, m\} \setminus C$. The columns $S_B = \cup_{i \in B} S_i$ must be in the collection of lines. Thus the minimum possible size of a collection of such lines is

$$\min\{m - |B| + |S_B| : B \subset \{1, \ldots, m\}\}.$$

If $B = \emptyset$ i.e. $C = \{1, \ldots, m\}$ then the corresponding collection has size m. So the minimum value above is $m - \Delta$ where

$$\Delta = \max\{|B| - |S_B| \ : \ B \subset \{1, \ldots, m\}, |B| > |S_B|\}.$$

The result now follows from the generalized Philip Hall theorem. Moreover, given a collection $S_1, \ldots, S_m$ of subsets of a finite set X (say of size n), we can form the $0, 1$ -matrix A of size $m \times n$ above, and hence the Kőnig-Egerváry theorem is equivalent to the generalized Philip Hall theorem. □

34. Assignment 9

(1) Let $\mathcal{A} = (A_1, A_2, \ldots, A_n)$ be a family of sets with an SDR. Let x be an element of A_1. Prove that there is an SDR containing x, but show by example that it may not be possible to find an SDR in which x represents A_1.

(2) Suppose $\mathcal{A} = (A_1, A_2, \ldots, A_n)$ is a family of sets that "more than satisfies" the marriage condition in the sense that:

$$|A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_k}| \geq k+1 \quad \forall\, k = 1, 2, \ldots, n,\ 1 \leq i_1 < \cdots < i_k \leq n.$$

Let $x \in A_1$. Prove that $\mathcal{A}$ has an SDR in which x represents A_1 .

(3) Let $n > 1$ and let $\mathcal{A} = (A_1, A_2, \ldots, A_n)$ be the family of subsets of $\{1, 2, \ldots, n\}$, where $A_i = \{1, 2, \ldots, n\} \setminus \{i\}$ for $i = 1, 2, \ldots, n$. Prove that $\mathcal{A}$ has an SDR and that the number of SDRs is the n-th derangement number D_n.

(4) A corporation has seven available positions $Y_1, Y_2, \ldots, Y_7$ and there are ten applicants $X_1, \ldots, X_{10}$. The set of positions each applicant is qualified for is given, respectively, by

$$\{Y_1, Y_2, Y_6\}, \{Y_2, Y_6, Y_7\}, \{Y_3, Y_4\}, \{Y_1, Y_5\}, \{Y_6, Y_7\}, \{Y_3\}, \{Y_2, Y_3\}, \{Y_1, Y_3\}, \{Y_1\}, \{Y5\}.$$

Determine the largest number of positions that can be filled by the qualified applicants and justify your answer.

(5) Let $\mathcal{A} = (A_1, A_2, A_3, A_4, A_5, A_6)$ where

$$A_1 = \{a, b, c\}, A_2 = \{a, b, c, d, e\}, A_3 = \{a, b\}, A_4 = \{b, c\}, A_5 = \{a\}, A_6 = \{a, c, e\}.$$

Does the family $\mathcal{A}$ have an SDR? If not, what is the largest number of sets in the family with an SDR?

(6) (optional) In this problem we drop the assumption that G is finite, in the theorem about group theory in the previous section.

(a) Prove the following group theory lemma:

Lemma. *Let H be a subgroup of a group G such that $[G : H] = m$ is finite. Prove that there is a normal subgroup N of G with N contained in H and such that the quotient group G/N is finite. (Hint: consider the kernel of the homomorphism $G \to S_m$ afforded by the left action of G on G/H.)*

(b) Let N be as in part a) and let $\bar{G}$ denote the finite group G/N. Let $\bar{H} = H/N$. Note that $[\bar{G} : \bar{H}] = m$. (Prove this if it is not clear). Since we have proved the theorem for finite G we obtain elements $x_1, x_2, \ldots, x_m$ of G such that

$$\bar{G} = \cup_{i=1}^{m} x_i \bar{H} = \cup_{i=1}^{m} \bar{H} x_i.$$

Use this to show that:

$$G = \cup_{i=1}^{m} x_i H = \cup_{i=1}^{m} H x_i.$$

35. Quiz-4

(1) (6 points) Consider the equation $XC(X)^2 = C(X)-1$ for the generating function $\sum_{n\geq 0} C_n X^n$ of the Catalan numbers. Let $A(X) = C(X) - 1$. Convert the above equation to a functional equation for $A(X)$ of a form to which Lagrange inversion can be applied. Using Lagrange inversion obtain the formula $C_n = \frac{1}{n+1}\binom{2n}{n}$.

Ans: We have $X = A(X)/(1+A(X))^2$ which shows that $A(X)$ is the compositional inverse of $f(X) = X/(1+X)^2$. By Lagrange inversion, for $n \geq 1$, we have

$$C_n = \tfrac{1}{n}\mathrm{Res}(f^{-k}) = \tfrac{1}{n}\mathrm{Res}(X^{-n}(1+X)^{2n}) = \tfrac{1}{n}[X^{n-1}](1+X)^{2n} = \tfrac{1}{n}\tbinom{2n}{n-1} = \tfrac{1}{n+1}\tbinom{2n}{n}.$$

(2) (7 points) Let $S_1, \ldots, S_n$ be n sets that have an SDR. Suppose that $a_1, \ldots, a_t$ are an SDR for $S_1, \ldots, S_t$ where $t < n$. Prove that $S_1, \ldots, S_n$ have an SDR including the elements $a_1, \ldots, a_t$, but not necessarily as representatives of $S_1, \ldots, S_t$. Give an example of sets $S_1, \ldots, S_n$ and elements $a_1, \ldots, a_t$ that constitute an SDR for $S_1, \ldots, S_t$ but which cannot be representatives of $S_1, \ldots, S_t$ in any SDR of $S_1, \ldots, S_n$.

Ans: Let $x_1, \ldots, x_n$ be an SDR for $S_1, \ldots, S_n$. If $\{a_1, \ldots, a_t\} \subset \{x_1, \ldots, x_n\}$ then we are done. Otherwise, there is an integer $0 \leq \mu \leq t-1$ such that (upto relabeling $a_1, \ldots, a_t$ and $S_1, \ldots, S_t$) that $\{a_1, \ldots, a_\mu\} \subset \{x_1, \ldots, x_n\}$ and $a_{\mu+1}, \ldots, a_t \notin \{x_1, \ldots, x_n\}$. It suffices to show that we can increase the quantity $\{a_1, \ldots, a_t\} \cap \{a_1, \ldots, a_t\}$.

Case 1: If $\mu < t/2$. In this case replacing $x_{\mu+1}, \ldots, x_t$ by $a_{\mu+1}, \ldots, a_t$ respectively, we get an SDR

$$x_1, \ldots, x_\mu, a_{\mu+1}, \ldots, a_t, x_{t+1}, \ldots, x_n$$

which contains $\{a_{\mu+1}, \ldots, a_t\}$. This SDR has at least $t - \mu > \mu$ elements from $\{a_1, \ldots, a_t\}$.

Case 2: $\mu \geq t/2$: Suppose there is an $i \in \{1, \ldots, t-\mu\}$ such that $x_{\mu+i} \notin \{a_1, \ldots, a_\mu\}$ (this is always possible if $\mu > t/2$) then replacing $x_{\mu+i}$ by $a_{\mu+i}$ we get a new SDR with at least $\mu + 1$ elements from $\{a_1, \ldots, a_t\}$.

It remains to consider the case when $\mu = t/2$ and $\{x_{t/2+1}, \ldots, x_t\} = \{a_1, \ldots, a_{t/2}\}$. Here we note that $\{x_1, \ldots, x_{t/2}\}$ and $\{a_1, \ldots, a_{t/2}\}$ are disjoint. Replacing x_i by a_i and $x_{t/2+i}$ by $a_{t/2+i}$ for $1 \leq i \leq t/2$, we get an SDR containing all of $a_1, \ldots, a_t$.

part2) : Let $S_1, \ldots, S_n \subset S$ where $|S| \geq n$ (which is necessary for an SDR of $S_1, \ldots, S_n$ to exist). Let $a_1, \ldots, a_t$ be distinct elements of S, and suppose $S_1 = \cdots = S_{n-1} = S$ and $S_n = \{a_t\}$. The Hall condition holds for $S_1, \ldots, S_n$: For a subset B of $\{1, \ldots, n\}$, if $B = \{n\}$ then $|\cup_{i\in B} S_i| = 1 = |B|$. For any other B, we have $|\cup_{i\in B} S_i| = |S| \geq n \geq |B|$. Thus an SDR for $S_1, \ldots, S_n$ does exist. Clearly in any such SDR a_t represents S_n and hence $a_1, \ldots, a_t$ cannot represent $S_1, \ldots, S_t$ in any SDR for $S_1, \ldots, S_t$. Since $S_1 = \cdots = S_t = S$, clearly $a_1, \ldots, a_t$ is an SDR for $S_1, \ldots, S_t$.

(3) (7 Points) Let $P(X) = \sum_{n=1}^{\infty} p_n X^n$ be the generating function of the partition numbers. Consider the formal power series defined by:

$$\sum_{n=1}^{\infty} a_n X^n = \frac{XP'(X)}{P(X)}.$$

(where $P'(X)$ is the derivative of $P(X)$). Give an interpretation of a_n in terms of elementary number theory.

Ans: Since $P(X) = \Pi_{k=1}^{\infty}(1-X^k)^{-1}$ we get $\frac{XP'(X)}{P(X)} = \sum_{k=1}^{\infty} \frac{kX^k}{1-X^k}$ (see justification below). The contribution to the coefficient of X^n of the k-th summand is 0 if $k \nmid n$ and k if $k \mid n$. Thus $a_n = \sigma(n)$ the sum of divisors of n.

Justification for $\frac{XP'(X)}{P(X)} = \sum_{k=1}^{\infty} \frac{kX^k}{1-X^k}$:
In terms of $Q(X) = 1/P(X) = \prod_{k=1}^{\infty}(1-X^k)$, we must show

$$-XQ \cdot Q' = \sum_{k=1}^{\infty} \frac{kX^k}{1-X^k},$$

which in turn reduces to

$$Q' = \sum_{k=1}^{\infty} (1-X^k)' \prod_{i \in \mathbb{N}\setminus\{k\}} (1-X^i)$$

Let Tr_n denote the polynomial of degree at most n obtained by truncating a power series up-to the X^n-th term. In other words $\mathrm{Tr}_n C(X) = (c_0 + c_1 X + \cdots + c_n X^n)$. Now, two elements $A(X), B(X) \in \mathbb{R}[[X]]$ are equal if and only if $\mathrm{Tr}_n A = \mathrm{Tr}_n B$ for all non-negative integers n. Thus we must show (for each n):

$$\mathrm{Tr}_n(Q') = \mathrm{Tr}_n \left(\sum_{k=1}^{\infty} (1-X^k)' \prod_{i \in \mathbb{N}\setminus\{k\}} (1-X^i) \right)$$

This follows from:

$$\mathrm{Tr}_n(Q') = \mathrm{Tr}_n \left(\prod_{k=1}^{n+1} (1-X^k) \right)' = \mathrm{Tr}_n \left(\sum_{k=1}^{n+1} (1-X^k)' \prod_{i \in \{1,\ldots n+1\}\setminus\{k\}} (1-X^i) \right) = \mathrm{Tr}_n \left(\sum_{k=1}^{\infty} (1-X^k)' \prod_{i \in \mathbb{N}\setminus\{k\}} (1-X^i) \right)$$

36. Additional Practice Problems

(1) Recall that the Stirling number $S(n,k)$ of the second kind, is the number of ways to put n distinct objects into k unlabeled boxes, such that no box is non-empty. Determine the following values (with reasoning), assuming $n \geq 2$. Express your final answer in terms of binomial coefficients.
 (a) $S(n,2)$.
 (b) $S(n,n-1)$.
 (c) $S(n,n-2)$.

(2) a) Show that the number of partitions of n with each part at most 2 is $\lfloor n/2 \rfloor + 1$.
b)* Show that the number of partitions of n with each part at most 3 is the nearest integer to $(n+3)^2/12$.

(3) Determine a recurrence relation for the number a_n of ternary strings (made up of 0's, 1's, and 2's) of length n that do not contain two consecutive 0's or two consecutive 1's. Then find a formula for a_n.

(4) Let S be the multiset $\{\infty \cdot e_1, \infty \cdot e_2, \infty \cdot e_3, \infty \cdot e_4\}$. Determine the generating function for the sequence $h_0, h_1, h_2, \cdots, h_n$ where h_n is the number of n- combinations of S with the following added restrictions:
 (a) Each e_i occurs an odd number of times.
 (b) Each e_i occurs a multiple-of-3 number of times.
 (c) The element e_1 does not occur, and e_2 occurs at most once.
 (d) The element e_1 occurs 1,3, or 11 times, and the element e_2 occurs $2, 4$ or 5 times.
 (e) Each e_i occurs at least 10 times.

(5) Find the exponential generating function for the number of symmetric $n \times n$ permutation matrices.

37. Final Exam

(1) (7 pts) Suppose we have two partitions $\coprod_{i=1}^m A_i$ and $\coprod_{i=1}^m B_i$ of a set Y such that each of the sets A_i and B_j have the same size n. Show that there is a relabeling of the sets $B_1, \ldots, B_m$ such that each $A_i \cap B_i$ is non-empty.

(2) (a) (4 pts) Give a formula (proof not required) for the generating function $C(X) = \sum_{n=0}^{\infty} C_n X^n$ of the Catalan numbers (Recall $C_0 = C_1 = 1$ and $C_2 = 2$)
(b) (4 pts) Using part a) or otherwise, express the following expression as a binomial coefficient (with proof)

$$\sum_{j=0}^{n} \frac{1}{j+1} \binom{2j}{j} \binom{2n-2j}{n-j}$$

(3) Let h_n denote the number of ways of putting a certain structure (let us call it h-structure) on a set of size $n \geq 1$. The exponential generating function for h-structures is $h(X) = \sum_{n=1}^{\infty} h_n X^n/n!$. For example if h is the trivial structure (putting no structure) on a set of size n (so that $h_n = 1$) then $h(X) = e^X - 1$.
(a) (5 pts) Let g and h be two structures. An f-structure on a set S of size n is obtained by
- partitioning S into any number of non-empty blocks. The blocks are not labeled.
- putting an h-structure on each block.
- putting a g-structure on the set of blocks.

Obtain an expression for $f(X)$ in terms of $g(X)$ and $h(X)$.
(b) (5 pts) Recall that the Bell number B_n is the number of equivalence relations on a set of size n. Obtain a formula for $B(X) = \sum_{n=1}^{\infty} B_n X^n/n!$, preferably using part a).
(c) (5 pts) Recall that the Stirling number of the first kind $c_{n,k}$ is the number of ways of arranging n distinct people into k unlabeled circles. Obtain a formula for $C_k(X) = \sum_{n \geq 1} c_{n,k} X^n/n!$, preferably using part a).
(d) (5 pts) A rooted forest on n labeled vertices $\{v_1, \ldots, v_n\}$ is a disjoint union of any number of rooted labeled trees such that union of the vertex set of all the trees is $\{v_1, \ldots, v_n\}$. Recall that a rooted tree is just a tree with one of its vertices singled out. Let τ_n the number of rooted labeled trees on n vertices, and let f_n be the number of rooted forests on n labeled vertices. For example $f_1 = 1, f_2 = 3$, and $\tau_n = n^{n-1}$. Express, preferably using part a), $F(X) = \sum_{n=1}^{\infty} f_n X^n/n!$ in terms of $\tau(X) = \sum_{n=1}^{\infty} \tau_n X^n/n!$. Find a simple relation between the sequence $\{f_n\}_{n=1}^{\infty}$ and the sequence $\{\tau_n\}_{n=2}^{\infty}$, and use it to obtain the functional equation for $\tau(X)$.

Solution to Question 1) done in class

Solution to Question 2)
a) We have

$$C(X) = \sum_{n \geq 0} \tfrac{1}{n+1} \tbinom{2n}{n} X^n = \tfrac{1-\sqrt{1-4X}}{2X}.$$

b) We have

$$\sum_{n\geq 0}\tbinom{2n}{n}X^n = (XC(X))' = (1-4X)^{-1/2}.$$

Thus

$$\sum_{j=0}^{n}\tfrac{1}{j+1}\tbinom{2j}{j}\tbinom{2n-2j}{n-j}$$

is the coefficient of X^n in

$$\tfrac{1-\sqrt{1-4X}}{2X}\cdot\tfrac{1}{\sqrt{1-4X}} = \tfrac{(1-4X)^{-1/2}-1}{2X} = \sum_{n\geq 1}\tbinom{-1/2}{n}(-4)^n X^{n-1}/2 = \textstyle\sum_{n\geq 1}\tbinom{2n-1}{n-1}X^{n-1}.$$

Therefore, the answer is $\binom{2n+1}{n}$.

Solution to Question 3)

a) We claim $f(X) = g(h(X))$. Indeed:

$$f_n = \sum_{k\geq 1} g_k \sum_{\{(i_1,i_2,\ldots,i_k)\in\mathbb{N}^k : i_1+\cdots+i_k=n\}} \frac{1}{k!}\binom{n}{i_1 i_2\ldots i_k} h_{i_1}h_{i_2}\ldots h_{i_k}$$

This gives

$$f(X) = \sum_{n\geq 1}\tfrac{f_n X^n}{n!} = \sum_{k\geq 1}\tfrac{g_k}{k!}\prod_{(i_1,i_2,\ldots,i_k)\in\mathbb{N}^k}\tfrac{h_{i_1}X^{i_1}}{i_1!}\tfrac{h_{i_2}X^{i_2}}{i_2!}\cdots\tfrac{h_{i_k}X^{i_k}}{i_k!} = \sum_{k\geq 1}\tfrac{g_k h(X)^k}{k!} = g(h(X))$$

b) Here $f(X) = B(X)$ and $h(X) = g(X) = e^X - 1$. Thus $B(X) = e^{e^x-1} - 1$.

c) Here $f(X) = C_k(X)$ and $g_n = \delta_{k,n}$ which gives $g(X) = X^k/k!$. Also $h_n = (n-1)!$ as it is the number of circular permutation of n objects, which gives $h(X) = \sum_{n\geq 1} X^n/n = -\log(1-X)$. Thus

$$C_k(X) = (-\log(1-X))^k/k!$$

d) For $F(X)$ we must take $h_n = \tau_n$ and $g_n = 1$. Thus $F(X) = e^{\tau(X)} - 1$.

Next, we note that for $n \geq 2$, we have τ_n/n (the number of unrooted labeled trees on n vertices $\{v_1,\ldots,v_n\}$) equals f_{n-1}: to see this we note that given such a tree, if we remove v_n and all edges emanating from it, we get a rooted forest on $\{v_1,\ldots,v_{n-1}\}$. Thus

$$\tau(X) = X + \sum_{n\geq 2}\tfrac{n f_{n-1}X^n}{n!} = X + XF(X) = Xe^{\tau(X)}.$$

This we get the functional equation $\tau(X)e^{-\tau(X)} = X$.

38. Epilogue

There are some topics which should have been covered to some extant during the course, but which we did not have time for. Here we list some of these topics/techniques.

Introduction to Ramsey Theory. This topic will be covered in the Graph Theory course. A basic example of this theory is the following observation. In a party with at least 6 people, there will either be a group of 3 people who know each other, or a group of 3 people who are strangers to each others. Given positive integers m and n, the Ramsey number, $R(m,n)$ is the minimum number of people that should be in a party so that either there is a group of m people who know each other pairwise, or a group of n people who are pairwise strangers. The technique known as the *probabilistic method* gained prominence after it was used by Erdős (1947) to obtain the lower bound $R(m,m) \geq 2^{m/2}$.

An upper bound for the Ramsey numbers is $R(m,n) \leq \binom{m+n-2}{m-1}$. A nice way to prove this is the technique of double induction below

Technique of Double Induction.

Theorem. *If a statement $P(m,n)$ for $(m,n) \in \mathbb{N} \times \mathbb{N}$ is true for $(m,n) \in \{1\} \times \mathbb{N}$ and for $(m,n) \in \mathbb{N} \times \{1\}$, and if $P(m,n)$ follows from $P(m-1,n)$ and $P(m,n-1)$, then $P(m,n)$ holds for all $(m,n) \in \mathbb{N} \times \mathbb{N}$.*

Proof. This is just a variant of ordinary induction, by inducting on $k = m+n$. Let $R(k)$ be the assertion that $P(m,n)$ is true whenever $m+n=k$. The base case is $k=2$. Assume $R(k-1)$ holds. For any (m,n) with $m+n=k$, we note that $(m-1,n)$ and $(m,n-1)$ are true by inductive hypothesis. We now use the fact $P(m-1,n)$ and $P(m,n-1) \Rightarrow P(m,n)$ to deduce that $P(m,n)$ is true. Thus $R(k)$ holds. □

The marching band Problem. *Exercise 3.4.26 of [Brualdi]*
Suppose that the mn people of a marching band are standing in a rectangular formation of m rows and n columns in such a way that in each row each person is taller than the one to his or her right. Suppose that the leader rearranges the people in each column in decreasing order of height from front to back. Show that the rows are still arranged in decreasing order of height from left to right.