# Combinatorics - Lecture Note

Lectures by professor Catherine Yan Notes by Byeongsu Yu

December 26, 2018

#### Abstract

This note is based on the course, Combinatorics given by professor Catherine Yan on Fall 2018 at Texas A&M University. Much part of this note was  $T_EX$ -ed after class. Every blemish on this note is Byeongsu's own.

# Contents

1	Bas	ics - Fundamental Coefficients	1
	1.1	Basic Number and Rules	1
	1.2	Basic identities	3
	1.3	Lattice Paths and Stirling number	6
	1.4	Short introduction to Generating function and Exponential Generating Function	7
	1.5	Permutation Statistics	12
		1.5.1 Inversion	14
		1.5.2 Descents	16
		1.5.3 Eulerian Polynomial	17
	1.6	Composition	22
	1.7	12fold way	24
	1.8	q-Analog and Gaussian Coefficients	25
<b>2</b>	For	mal Power Series and Infinite Matrices	29
	2.1	Ordinary and Exponential generating function.	29
	2.2	q-analog of generating function	30
	2.3	Probabilistic Generating Function	35
3	Ger	nerating Function	40
	3.1	Evaluating sums	43
	3.2	Enumeration on trees	49
	3.3	Lagrange Inversion Formula	52
	3.4	Generating function for Integer Partition	56
	-	3.4.1 Euler's Pentagonal Theorem	58
		3.4.2 Jacobi Triple Identity	60
4	Ger	atle introduction to WZ method	61
<b>5</b>	Siev	ve Method	63
	5.1	Principle of Inclusion-Exclusion	63
	5.2	Möbius inversion	70
	5.3	Involution Principle	81
	5.4	The Lemma of Gessel - Viennot	86

6	Enu	imeration of Patterns	90
	6.1	Symmetries and Patterns	90
	6.2	Cycle index	93
	6.3	The Theorem of PólyaRedfield	94
7	Cat	alan numbers and structures	95
	7.1	Example of catalan structures.	95
	7.2	NC matching of [n]	98

# **1** Basics - Fundamental Coefficients

#### 1.1 Basic Number and Rules

Let  $\mathbb{N} = \{0, 1, 2, \dots\}, [n] := \{1, 2, \dots, n\}.$ 

• [Rule of sum]: If  $A \cap B = \emptyset$ , then

$$|A \cup B| = |A| + |B|.$$

• [Rule of product]: If A, B are finite set, then

$$|A \times B| = |A| \cdot |B|.$$

Application: A task has k steps. There are  $a_i \in \mathbb{N}$  ways to do *i*-th step. Then, # of ways to do the task is  $a_1 \cdots a_k = \prod_{i=1}^k a_i$ .

• [Rule of bijection]: If  $\exists$  a bijection from A onto B, then

$$|A| = |B|.$$

For example,

# of subsets of 
$$[n] = 2^n$$
  
# of words of length  $n = 2^n$ 

We can see this by encoding the information of subset of [n] using word of length n. For example,

$$\{2, 3, 4\} \mapsto 01110000.$$

• [Rule of double counting]: If 2 formulas count the same set then they are equal. This leads to how we can provide a combinatorial proof. For example,

$$C(n,k) = \binom{n}{k} = \#$$
 of subsets of  $[n]$  of size  $k$ .

Since complementary of size k subset is a subset of n - k subset and this relationship is bijective, we can say that

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

without calculation. Another example is Pascal's identity, which is,

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

To see this, # of choosing k element from the n element is the same as fixing one element, and choosing k-1 elements from the other n-1 elements (so that we can choose k elements containing such fixed one) plus choosing k elements from the other n-1 elements (so that we can choose k elements without fixed one.) In summary, the number of choosing k elements from n elements is the number of choosing k elements containing some special element plus the number of choosing k elements without such special one.

• [Rule of subtraction]:

$$|A| = |U| - |\bar{A}|.$$

This leads to the principle of inclusion and exclusion, i.e.,

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

**Definition 1.1.1** (Permutation). A k-permutation is an ordered list of k-distinct elements from n, where  $0 \le k \le n$ . The # of k-permutation is denoted as P(n, k), and

$$P(n,k) = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1).$$

Definition 1.1.2 (Falling and Rising factorial). We call

$$n^{\underline{k}} = n(n-1)\cdots(n-k+1)$$

as k-falling factorial of n. Similarly, we call

$$n^{\overline{k}} = n(n+1)\cdots(n+k-1)$$

as k-rising factorial of n.

Thus,

$$\binom{n}{k} = \frac{P(n,k)}{k!} = \frac{n^{\underline{k}}}{k!}$$

Actually k-falling (resp. rising) factorial can be generalized to  $\mathbb{R}$ ; for any  $x \in \mathbb{R}$ ,

$$x^{\underline{k}} := x(x-1)\cdots(x-k+1)$$
$$x^{\overline{k}} := x(x+1)\cdots(x+k-1)$$

Then we have identity that

$$(-x)^{\underline{k}} = (-x)(-x-1)\cdots(-x-k+1) = (-1)^{k}(x)(x+1)\cdots(x+k-1) = (-1)^{k}x^{\overline{k}}$$
$$(-x)^{\overline{k}} = (-x)(-x+1)\cdots(-x+k-1) = (-1)^{k}(x)(x-1)\cdots(x-k+1) = (-1)^{k}x^{\underline{k}}.$$

Now we can define that

$$\binom{x}{k} := \frac{x^{\underline{k}}}{k!}.$$

Actually,  $\binom{-1}{n}, \binom{1}{2}$  occurs often in combinatorics. From this definition, for any  $c \in \mathbb{R}, k \in \mathbb{N}$ ,

$$\binom{-c}{k} = \frac{(-c)(-c-1)\cdots(-c-k+1)}{k!} = \frac{(-1)^k c(c+1)\cdots(c+k-1)}{k!} = (-1)^k \binom{c+k-1}{k}.$$

Similarly,

$$\binom{c}{k} = \frac{c(c-1)\cdots(c-k+1)}{k!} = \frac{(-1)^k(-c)(-c+1)\cdots(-c+k-1)}{k!} = (-1)^k \frac{(-c+k-1)(-c+k-2)\cdots(c)}{k!} = (-1)^k \binom{k-c-1}{k}.$$

Theorem 1.1.3 (Reciprocity Law).

$$\binom{-c}{k} = (-1)^k \binom{c+k-1}{k} \text{ and } \binom{c}{k} = (-1)^k \binom{k-c-1}{k}.$$

From this, we know that

$$\begin{pmatrix} -1\\n \end{pmatrix} = (-1)^n \begin{pmatrix} 1+n-1\\n \end{pmatrix} = (-1)^n \\ \begin{pmatrix} \frac{1}{2}\\n \end{pmatrix} = (-1)^n \begin{pmatrix} n-\frac{3}{2}\\n \end{pmatrix} \\ \begin{pmatrix} -\frac{1}{2}\\n \end{pmatrix} = (-1)^n \begin{pmatrix} n-\frac{1}{2}\\n \end{pmatrix}$$

#### 1.2 Basic identities

Theorem 1.2.1 (Binomial Theorem).

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}.$$

Maybe induction can be used for proving this, but we can see other ways.

*Proof 1.* For polynomial identities, verify it for sufficiently many values. In practices, we can deal with infinitely many values, such as  $\mathbb{P}$ . Now let X, Y be 2 disjoint sets with |X| = x, |Y| = y. Our goal is to count # of words of length n with alphabet set  $X \cup Y$ .

- Counting 1 : It is just  $(x+y)^n$ .
- Counting 2 : In a words on  $X \cup Y$  of length n, let

$$k = \#$$
 of letters in  $X, 0 \le k \le n$ .

For a fixed k, # of words of length n with k letters in X and (n-k) letters in Y is

$$\binom{n}{k} x^k y^{n-k}$$

since first of all, we should choose k positions from the words for letters from X, and the number of choosing such positions is  $\binom{n}{k}$ . From each chosen positions, there are x possible letters from X, and for each unchosen position, y possible letters. Thus, the number of possible words with chosen k position is  $x^k y^{n-k}$ .

Now summing over k, we get

# of words 
$$=\sum_{k=0}^{n} \binom{n}{k} x^{k} y^{n-k}$$
.

Remark 1.2.2. By convention, we define

$$\sum_{\emptyset} = 0, \prod_{\emptyset} = 1.$$

Theorem 1.2.3 (Newton Binomial Theorem).

$$(1+x)^{\alpha} = \sum_{k=0}^{\infty} {\alpha \choose k} x^k.$$

for |x| < 1.

Useful cases of Newton's binomial theorem is

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots$$

Actually,

$$(1-x)^{-1} = \sum_{k \ge 0} \frac{-1}{k} (-x)^k = \sum_{k \ge 0} (-1)^k \binom{k}{k} (-x)^k = \sum_{k \ge 0} x^k.$$

You can find useful identity here. Similarly,

$$(x+y)^{\underline{n}} = \sum_{k=0}^{n} \binom{n}{k} x^{\underline{k}} y^{\underline{n-k}}$$
$$(x+y)^{\overline{n}} = \sum_{k=0}^{n} \binom{n}{k} x^{\overline{k}} y^{\overline{n-k}}$$

We can prove first one using combinatorial structure. If we assume x = |X| and y = |Y| for some set X, Y, and assume n < x + y, then we can count the number of length n words from characteriset  $X \cup Y$  without repeating character in two ways;

- Counting 1: Just  $(x+y)^{\underline{n}}$ .
- Counting 2: In a words on  $X \cup Y$  of length n, let

$$k = \#$$
 of letters in  $X, 0 \le k \le n$ .

For a fixed k, # of words of length n with k letters in X and (n-k) letters in Y is

$$\binom{n}{k} x^{\underline{k}} y^{\underline{n-k}}$$

since first of all, we should choose k positions of characters of fixed word for letters from X, and the number of choosing such positions is  $\binom{n}{k}$ . For k chosen positions, there are  $x^{\underline{k}}$  possible ways to have letters from X, and for n - k unchosen positions, there are  $y^{\underline{k}}$  possible ways to have letters from Y. Thus, the number of possible words for fixed k chosen position is  $x^{\underline{k}}y^{\underline{n-k}}$ .

Now summing over k, we get

# of words 
$$=\sum_{k=0}^{n} {n \choose k} x^{\underline{k}} y^{\underline{n-k}}$$
.

For the second one, think about it...

**Definition 1.2.4** (Polynomial sequence of binomial type). A polynomial sequence of binomial type is a sequence of polynomials  $\{p_n(x)\}_{n=0}^{\infty}$  such that  $\deg(p_n(x)) = n$  and for any  $n \in \mathbb{N}$ ,

$$p_n(x+y) = \sum_{k=0}^n \binom{n}{k} p_k(x) p_{n-k}(x).$$

The second condition is called **binomial type**.

Theorem 1.2.5 (Hockey Stick Identity).

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

In this case, we define  $\binom{i}{k} = 0$  if i < k.

*Proof.* Choosing k + 1 elements from n + 1 elements,

- Counting 1:  $\binom{n+1}{k+1}$ .
- Counting 2: We can divide case where a special person is chosen or not; i.e,

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

Now  $\binom{n}{k+1}$  also can be represented by a second special person is chosen or not, i.e.,

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

By repeating this, we get

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n-1}{k} + \dots + \binom{k+1}{k} + \binom{k+1}{k+1}.$$

and since  $\binom{k+1}{k+1} = \binom{k}{k}$ , we are done.

Theorem 1.2.6 (Vandermonde Identity).

$$\sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k} = \binom{x+y}{n}.$$
$$\sum_{a+b=n} \binom{x}{a} \binom{y}{b} = \binom{x+y}{n}.$$

or

*Proof.* Let X, Y be sets such that |X| = x, |Y| = y. Then how many ways of choosing n elements from  $X \cup Y$ ?

- Counting 1:  $\binom{x+y}{n}$ .
- Counting 2: Let k be a number of chosen elements from X. Then,  $k = 0, 1, \dots, n$ . Then, for each fixed k, there are  $\binom{x}{k}\binom{y}{n-k}$  ways to choosing k elements from X and choosing n-k elements from Y. By summing these for all possible k, we get

$$\sum_{k=0}^{n} \binom{x}{k} \binom{y}{n-k}.$$

Theorem 1.2.7 (Multinomial Theorem).

$$(x_1 + \dots + x_k)^n = \sum \binom{n}{a_1, \dots, a_k} x_1^{a_1} \cdots x_k^{a_k}$$

where  $\binom{n}{a_1, \cdots, a_k} := \frac{n!}{a_1! \cdots a_k!}$ .

Note that

$$\binom{n}{a} = \binom{n}{a, n-a}.$$

#### **1.3** Lattice Paths and Stirling number

- 1.  $\binom{m+n}{n} \to \#$  of lattice paths from  $\overline{0} = (0,0)$  to (m,n). To see this, note that each lattice path consisting of *m*-step in East and *n*-step in North. Thus, a path corresponds to a sequence of  $\{E, N\}$  with exactly m E-step and n N-step.
- 2.  $\binom{n}{k}$ : Choose k items for a set of n without repetition.
- 3.  $\binom{n}{k} = \binom{n+k-1}{k}$ : Choose k items for a set of n allowing repetition. Why? assume the items are  $S_1, \dots, S_n$ . A k-multiset is  $(x_1, \cdot, sx_n)$  where  $x_i \in \mathbb{N}$  is the multiplicity of  $S_i$  and  $x_1 + \dots + x_n = k$ . Hence,

$$\binom{n}{k} = \#$$
 of solutions of  $x_i \in \mathbb{N}$  such that  $x_1 + \dots + x_n = k$ 

And to see this, use bar-ball trick. For example, if n = 4, k = 10, then

$$3+2+5+0=10 \quad \leftrightarrow \quad ooo|oo|ooooo|$$

a sequence of 10 balls and 3 bars. Hence,  $x_i$  denotes a horizontal steps and + denote a vertical steps. So we can denote it as a lattice path from (0,0) to (10,3) by



4. Set Partition:

**Definition 1.3.1.** Given  $S \neq \emptyset$ , |S| = n, a partition of S is a collection of subsets of S, say  $\{B_1, \dots, B_k\}$  such that  $S = \bigcup_{i=1}^k B_i$  and  $B_i \neq \emptyset$ ,  $B_i \cap B_j = \emptyset$  if  $i \neq j$ . Each  $B_i$  is called a block.

We use notation as

$$\Pi(S) = \{ \text{ partitions of } S \}, \Pi_n = \Pi([n]).$$

**Definition 1.3.2** (Stirling number of the 2nd kind). Let  $S_{n,k}$  be the Stirling number of the 2nd kind, denoting the number of partitions of [n] with k blocks.

For example,  $S_{n,0} = 0$ ,  $S_{0,0} = 1$  by convention.

$$S_{n,1} = 1, S_{n,n} = 1.S_{n,n-1} = \binom{n}{2}, S_{n,2} = \frac{2^n - 2}{2}$$

To see  $S_{n,2}$ , note that each nonempty proper subset A gives natural partition  $\{A, A^c\}$ . To see  $S_{n,n-1}$  note that only one block should have 2 elements and the other blocks are singletons.

**Definition 1.3.3** (Bell number).

$$B(n) := \sum_{k} S_{n,k} = \# \text{ partitions of } [n].$$

**Properties:** 

- (a)  $S_{n,k} = kS_{n-1,k} + S_{n-1,k-1}$ . To see this, consider the blocks containing n. Then,
  - It is in singleton, and in this case,  $S_{n-1,k-1}$  cases occur.
  - It is not a singleton, then choosing k partitions, and assign n to one of given partitions. This gives  $k \cdot S_{n-1,k}$  ways.

$$B(n+1) = \sum_{i=0}^{n} \binom{n}{i} B(i) \quad (n \ge 0)$$

*Proof.* Consider the size of blocks containing elements n + 1. If size is i + 1, then there are  $\binom{n}{i}$  ways to making a block with size i + 1, and the number of partitioning remaining elements are B(i).

# 1.4 Short introduction to Generating function and Exponential Generating Function

- To describe sequence  $f_0, f_1, f_2, \cdots$ ,
- (a) Closed formular: For example,

$$2^n, n^2 + n, \text{ or } \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^n + \left( \frac{1 - \sqrt{5}}{2} \right)^n \right).$$

(b) Summation: For example,

$$\sum_{i=1}^n \binom{n}{i}.$$

To use this, you need usefulness of this notation.

(c) Recurrence Relation: For example,

$$F_n = F_{n-1} + F_{n-2}, F_0 = 0, F_1 = 1.$$

(d) Generating function: We can represent the sequence as a formal power series, by

$$\sum_{i\geq 0} f_i x^i \quad \text{or } \sum_{i\geq 0} f_i \frac{x^i}{i!}.$$

These are from [Hai10], not dealt on the lecture. Let *structure* be a mathematical object constructed on each  $[n], n \in \mathbb{N}_0$ . Let f(n) be the number of structures on each [n]. Then, we can define

Definition 1.4.1 (Exponential Generating Function).

$$F(x) = \sum_{n} f(n) \frac{x^{n}}{n!}.$$

Example 1.4.2 (Trivial structures).

• If a structure is just "being a set," then

$$f(n) = 1, \forall n \in \mathbb{N}_0 \implies F(x) = e^x$$

• Structure = "1-element set" then

$$f(n) = \begin{cases} 1 & n = 1 \\ 0 & n \neq 1 \end{cases} \implies F(x) = x.$$

• Structure = "empty set" then

$$f(n) = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} \implies F(x) = 1.$$

• Structure = "non-empty set" then

$$f(n) = \begin{cases} 1 & n \neq 0 \\ 0 & n = 0 \end{cases} \implies F(x) = e^x - 1.$$

• Structure = "even size set" then

$$f(n) = \begin{cases} 1 & n \equiv 0 \mod 2 \\ 0 & n \equiv 1 \mod 2 \end{cases} \implies F(x) = \sum_{n \ge 0} \frac{x^{2n}}{(2n)!} = \cosh(x)$$

• Structure = "odd size set" then

$$f(n) = \begin{cases} 0 & n \equiv 0 \mod 2\\ 1 & n \equiv 1 \mod 2 \end{cases} \implies F(x) = \sum_{n \ge 0} \frac{x^{2n+1}}{(2n+1)!} = \sinh(x)$$

Now we have addition and multiplication principles for exponential generating function. Suppose f-structure be a structure which gives  $\{f(n)\}$ . Then,

**Principle 1.4.3** (Addition Principle for exponential generating functions). Suppose that the set of f-structures on each set is the disjoint union of the set of g-structures and the set of h-structures. Then

$$F(x) = G(x) + H(x)$$

where  $F(x) = \sum_{n} f(n) \frac{x^{n}}{n!}, G(x) = \sum_{n} g(n) \frac{x^{n}}{n!}, H(x) = \sum_{n} h(n) \frac{x^{n}}{n!}$ 

**Example 1.4.4.** If we can consider f = "trivial" structure, g = "empty" structure, h = "non-empty" structure. Every set is either empty or non-empty, and these possibilities are mutually exclusive, so we have

$$F(x) = G(x) + H(x) = 1 + (e^x - 1) = e^x.$$

To establish multiplication principle, we need a definition for multiplicative structure.

**Definition 1.4.5.** Let g and h denote two types of structures on finite sets. A  $g \times h$  structure on a set A consists of

- 1. an ordered partition of A into disjoint subsets  $A = A_1 \cup A_2$
- 2. a g-structure on  $A_1$ ,
- 3. a h-structure on  $A_2$ ,
- 4. where the structures in  $A_1, A_2$  are chosen independently.

**Principle 1.4.6** (Multiplication principle for exponential generating functions). If G(x) and H(x) are the exponential generating functions for g-structures and h-structures, respectively, then the exponential generating function for  $g \times h$  structures is

$$F(x) = G(x)H(x).$$

In general, let  $g_1 \times g_2 \times \cdots \otimes g_r$  structure on A consists of an ordered partition of A into r disjoint subsets  $A = A_1 \cup \cdots \cup A_r$ , and independently chosen  $g_i$  structures on each subset  $A_i$ . Then, by following induction on  $g_1 \times (g_2 \times \cdots \otimes g_r)$  and so on, the generating function of  $g_1 \times g_2 \times \cdots \otimes g_r$  structure is

$$G(x) = G_1(x)G_2(x)\cdots G_r(x).$$

[Hai10] gives a proof for this principle.

#### Example 1.4.7.

• To counting subsets of [n], note that choosing a subset is choosing a partition of [n] by 2, and imposing "being a set" structure for both. Thus by the multiplication principle,

$$F(x) = e^x \cdot e^x = e^{2x} = \sum_n 2^n \frac{x^n}{n!}$$

and  $f(n) = 2^n$  is consistent with our prior knowledge.

• To counting a non-empty subsets, this is equivalent to choosing a partition of [n] by 2 and imposing "non-empty set" structure on first one and trivial structure for second one (since we only count in case of chosen subset is nonempty; if complement of that set is empty, this implies chosen set is nonproper, thus nonempty.) Hence,

$$F(x) = (e^x - 1) \cdot e^x = e^{2x} - e^x \sum_n (2^n - 1) \frac{x^n}{n!}.$$

Also,  $f(n) = 2^n - 1$  is consistent with our prior knowledge.

• To counting number of functions from [n] to [k], note that this is equivalent of partition of [n] by k (allowing empty set) and giving a "being a set" structure. Hence,

$$F(x) = (e^x)^k = e^{kx} = \sum_n k^n \frac{x^n}{n!}.$$

Also,  $f(n) = k^n$  is consistent with our prior knowledge.

• To counting number of "surjective" functions from [n] to [k], we should give a "nonempty set" structure for every partition of [n] by k. This gives

$$F(x) = (e^x - 1)^k.$$

And note that number of surjective functions can be counted by making k nonempty partition of [n]and assign each element of [k], which can be represented by  $f(n) = k!S_{n,k}$  way. Thus,

$$F(x) = \sum_{n \ge k} k! S_{n,k} \frac{x^n}{n!}$$

Therefore,

$$\sum_{n \ge k} S_{n,k} \frac{x^n}{n!} = \frac{(e^x - 1)^k}{k!}.$$

There is another way of seeing above equation. Let

$$F_k(x) := \sum_{n \ge k} S_{n,k} \frac{x^n}{n!}$$

Then,

$$kF_k(x) + F_{k-1}(x) = \sum_{n \ge k} (kS_{n,k} + S_{n,k-1}) \frac{x^n}{n!} + S_{k-1,k-1} \frac{x^{k-1}}{(k-1)!} = F'_k(x) - S_{k,k} \frac{x^{k-1}}{(k-1)!} + S_{k-1,k-1} \frac{x^{k-1}}{(k-1)!} = F'_k(x) - S_{k-1,k-1} \frac{x^{k-1}}{(k-1)!} + S_{k-1,k-1} \frac$$

since  $S_{k,k} = 1 = S_{k-1,k-1}$ . Now note that for k = 0, 1  $F_k(x) = \frac{(e^x - 1)^k}{k!}$  holds. So, by inductive hypothesis, we can replace  $F_{k-1}(x)$  with  $\frac{(e^x - 1)^{k-1}}{(k-1)!}$ . Then,

$$F'_k(x) = kF_k(x) + \frac{(e^x - 1)^{k-1}}{(k-1)!}$$

Let  $y(x) = F_k(x), l(x) = \frac{(e^x - 1)^{k-1}}{(k-1)!}$ . Then,

$$y' - ky = l(x).$$

By multiplying  $e^{-kx}$  both sides, we get

$$y'e^{-kx} + y(-ke^{-kx}) = \frac{e^{-kx}(e^x - 1)^{k-1}}{(k-1)!}$$

By product rule,

$$ye^{-kx} = \int \frac{e^{-kx}(e^x - 1)^{k-1}}{(k-1)!} dx = \frac{e^{-kx}(e^x - 1)^k}{k} + C$$

So

$$y = \frac{(e^x - 1)^k}{k} + Ce^{kx}.$$

and from the inductive hypothesis C = 0, done.

**Remark 1.4.8** (Small contradiction). Note that the exercise 2.23 gives an equation of Stirling number of first kind, not the second kind.

• For the Bell number, note that B(n) is just consisting of all possible (nonempty) partitions of n, which implies just summing all possible partitions of n by k with respect to k. Hence, by addition principle,

$$\sum_{n} B(n) \frac{x^{n}}{n!} = \sum_{k} \sum_{n \ge k} S_{n,k} \frac{x^{n}}{n!} = \sum_{k} \frac{(e^{x} - 1)^{k}}{k!} = e^{(e^{x} - 1)^{k}}$$

See another identities about  $S_{n,k}$ .

5.

$$\sum_{k\geq k} S_{n,k} x^n = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}.$$

To see this [Aig07][p.63-64], let  $S_k(x) = \sum_{n \ge k} S_{n,k} x^n$ . Then,

$$S_{k-1}(x) + kS_k(x) = \sum_n (S_{n,k-1} + kS_{n,k})x^n = \sum_n S_{n+1,k}x^n = xS_k(x).$$

Thus

$$S_k(x) = \frac{xS_{k-1}(x)}{1-kx}$$

Since  $S_0(x) = 1$  by definition, we get

$$S_k(x) = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}$$

by induction.

6. 
$$x^n = \sum_{k \ge 1} S_{n,k} x^{\underline{k}}$$
.

*Proof.* Assume  $x \in \mathbb{P}$ . Then

$$x^n = \#$$
 of maps from  $[n]$  to X where  $|X| = x$   
=  $\sum_{K \subseteq X} \#$  of surjections from  $[n]$  to K

And if we fix K, then the number of surjection from [n] to K is  $S_{n,t}t!$ . Thus,

$$x^{n} = \sum_{t=1}^{n} \sum_{|K|=t, K \subseteq X} S_{n,t} t! = \sum_{t=1}^{n} \binom{x}{t} S_{n,t} t! = \sum_{t=1}^{n} S_{n,t} x^{\underline{t}} = \sum_{k \ge 1} S_{n,t} x^{\underline{k}}$$

OEIS: Online encyclopedia of interger sequence.

Note that table of  $S_{n,k}$  is

$n \setminus k$	0	1	2	3	4	5	6
0	1						
1	0	1					
2	0	1	1				
3	0	1	3	1			
4	0	1	7	6	1		
5	0	1	15	25	10	1	
6	0	1	31	90	65	15	1

You can see that it gives infinite dimensional matrix. In general, if S is infinite dimensional matrix then

$$a_n \sum_k S_{n,k} b_k \iff b_n \sum_k (S^{-1})_{n,k} a_k$$

where  $a_n, b_n$  are two integer sequences. From this inversion, we can get

$$x^{\underline{n}} \sum_{i=1}^{n} (S^{-1})_{n,k} x^k$$

Also,

$$(1+x)^n = \sum_k \binom{n}{k} x^k \implies x^n \sum (-1)^{n-k} \binom{n}{k} (1+x)^k$$

from inversion. You can check it by expanding  $((1 + x) - 1))^n$ .

7.

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

*Proof.* LHS is the number of surjections from [n] to [k]. Let A, B a set with |A| = a, |B| = b. Then, number of functions  $f : A \to B$  is  $b^a$ . The number of one-to-one functions is  $b^a$ . To get the number of surjections, let U be set of all functions from [n] to [k]. Take

$$A_i = \{ f \in U : i \notin \operatorname{Im}(f) \}$$

Then the number of surjections is

$$\left|\overline{A_1 \cup \cdots A_k}\right| = \left|\bigcap_{i=1}^k \overline{A_i}\right|$$

By inclusion-exclusin formular,

$$\left|\bigcap_{i=1}^{k} \overline{A_i}\right| = |U| - \sum |A_i| + \sum_{i < j} |A_i \cap A_j| + \dots + (-1)^k |A_1 \cap \dots \cap A_k|.$$

Since

$$|A_i| = (k-1)^n, |A_i \cap A_j| = (k-2)^n, \cdots, |A_{i_1} \cap \cdots \cap A_{i_r}|(k-r)^n,$$

for  $r = 1, 2, \cdots, k$ , we get

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

for RHS.

### 1.5 Permutation Statistics

There are several notation to denote permuations.

1.

$$\pi = a_1 \cdots a_n,$$

rearrangement of  $1, 2, \dots, n$ .

2. Bijection: for example,

$$f = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$
 or  $g = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ .

The first row denotes domain, and second row denotes range. So,

$$f \circ g = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

3. Picture: for f,

Note that permutation is union of cycles.

4. Notation for cycles. If

 $\pi = \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 1 & 4 & 6 & 8 & 2 & 7 \end{matrix}$ 

then we can rewrite it as

$$\pi = (13)(25687)$$

And standard is to write down each cycle starting with the largest entry, and list cycles in increasing order according to their initial entries. So, in this case,

$$\pi = (31)(4)(87256)$$

Now let  $\phi$  be a map from standard cycle notation to word of length *n* consisting of characters in [*n*] by removing all parentheses. It is bijection, since from any length *n*-word, we can recover it by find all left-to-right maximals and make a parenthese. For example, from 31487256 we can find it by

 $(31487256 \rightarrow (31)(487256) \rightarrow (31)(4)(87256 \rightarrow (31)(4)(87256)))$ 

**Theorem 1.5.1.** # of permutation of length n with k cycles = # of permutation of length n with k left-toright maximum

Definition 1.5.2 (Stirling number of first kind). The Stirling number of first kind is defined as

 $\mathscr{S}_{n,k} := \# of permutations of [n] with k cycles.$ 

Set  $\mathscr{S}_{0,0} = 1, \mathscr{S}_{0,k} = 0$  if  $k \ge 1$ , and  $\mathscr{S}_{n,0} = 0$ . Another easy example is

$$\mathscr{S}_{n,1} = (n-1)!, \mathscr{S}_{n,n} = 1, \mathscr{S}_{n,n-1} = \binom{n}{i}.$$

Lemma 1.5.3.

$$\mathscr{S}_{n,k} = \mathscr{S}_{n-1,k-1} + (n-1)\mathscr{S}_{n-1,k}.$$

*Proof.* Note that we can divides case of permutation of [n] with k cycles by

- A case that n is in the singleton cycle, which can be generated by  $\mathscr{S}_{n-1,k-1}$  way.
- n is not in the singleton cycle, which can be generated by  $(n-1)\mathscr{S}_{n-1,k}$  way. (Making a k cycle and assingn n, which gives (n-1) possibility, since assigning n in the word of length n-1 seems to give us n possibility, but the first position, which is the most left and a position between first cycle and second cycle gives the same cycle, if we think that assigned n is contained in the cycle which its left element contained.

Now see generating function. Fix n. Then,

Theorem 1.5.4.

$$\sum_{k\geq 0}\mathscr{S}_{n,k}x^k = x^{\overline{n}} = x(x+1)\cdots(x+n-1).$$

*Proof.*  $\sum_{\pi \in \mathfrak{S}n} x^{\# \operatorname{cycle}(\pi)}$ 

If  $\pi = (13)(25)(476)$ , denoting cycle by  $x_1 = (13), x_2 = (25), x_3 = (476)$  for example, then let X be a finite set with x elements in X. Then, Label each cycle by an element in X (allowing repeatition). Then,

#### LHS = # X-labeled permutations

Alternatively, we can construct X-labeled permutation by following step. Introduce entries one by one.

- 1. Integer 1: It has to appear a cycle, make a cycle (1) and label it in x ways.
- 2. Integer 2: either start a new cycle (x ways) or join the existing cycle (1 ways) so x + 1 ways occur.
- 3. Integer 3: either start a new cycle (x ways) or join the existing cycle (2 ways)
- 4. and so on.

From this,

$$RHS = x^{\overline{n}}.$$

Also note that

$$x^{\underline{n}} = (-1)^{n} (-x)^{\overline{n}} = \sum_{k=0}^{n} (-1)^{n-k} \mathscr{S}_{n,k} x^{k}.$$

From this equation, some authors define Stirling number of first kind with sign  $(-1)^{n-k} \mathscr{S}_{n,k}$  and call  $\mathscr{S}_{n,k}$  as signless Stirling number of first kind. From the equation

$$x^n = \sum_{k=0}^n S_{n,k} x^{\underline{k}},$$

We know that  $\{S_{n,k}\}$  and  $\{(-1)^{n-k}\mathscr{S}_{n,k}\}$  is inverse to each other.  $(S_{n,k}$  is the Stirling number of second kind.)

#### 1.5.1 Inversion

Let  $a_1 \cdots a_n \in \mathfrak{S}n$ , where  $\mathfrak{S}n$  is the symmetric group of n.

**Definition 1.5.5** (Inversion). An inversion is a pair  $(i, j) \in [n] \times [n]$  such that i < j and  $a_i > a_j$  from the given permutation  $a_1 \cdots a_n$ .

For example, if  $\pi = 35146827$ , then inversions are (13) or 31, and (17) or 32. First (13) denotes it by position, and second 31 denotes it by its entries. Now we use entry notation for this. So, the example has

as inversions. Define

$$INV(\pi) = \{ \text{ set of all inversions of } \pi \}, inv(\pi) = |INV(\pi)|.$$

Usually, capital letter denotes set and lower case letters denotes cardinality of the set. So, for above  $\pi$ ,  $inv(\pi) = 9$ . Actually you can check it by counting how many crossings occur on below function graph;



Given  $\pi$ , let

 $a_i := |\{j : j > i \text{ and } j \text{ appears before } i\}|.$ 

Then,

 $a_k = |\{ \text{ inversion with a form } *k \text{ in } \pi \}|$ 

for any k. From this know that

$$\sum_{i=1}^{n} a_i = inv(\pi) \tag{1}$$

since each set inducing  $a_i$  gives a partition of  $INV(\pi)$ . Note that

 $a_n = 0, 0 \le a_{n-1} \le 1, \cdots, 0 \le a_{n-k} \le k$ for  $j \in [n]$ .

**Definition 1.5.6** (Inversion Table). Inversion Table is a map  $\mathfrak{S}n \to A_{n-1} \times \cdots \times A_0$  where  $A_i = [i] \cup \{0\}$ , defined by  $\pi \to (a_1, \cdots, a_n)$ .

#### Claim 1.5.7. Inversion table is a bijection.

Note that if we know that domain and codomain has the same (finite) cardinality, then one-to-one or onto implies bijection. And in this case, we know

$$|\mathfrak{S}n| = n! = |A_{n-1}| \cdots |A_0| = |A_{n-1} \times \cdots \times A_0|.$$

*Proof.* It suffices to show that there exists a inverse map. Note that given  $(a_1, \dots, a_n)$ , we can reconstruct  $\pi$  by inserting integer from i = n to 1. For each i, insert i so that there are  $a_i$  terms in front of i.

For example, suppose (1, 5, 2, 0, 4, 2, 0, 1, 0)

$a_9 = 0 \implies$	9	
$a_8 = 1 \implies$	98	1 numeric should be in front of $8$
$a_7 = 0 \implies$	798	0 numeric should be in front of $7$
$a_6 = 2 \implies$	7968	2 numeric should be in front of $6$
$a_5 = 4 \implies$	79685	4 numeric should be in front of $5$
$a_4 = 0 \implies$	479685	0 numeric should be in front of $4$
$a_3 = 2 \implies$	4739685	2 numeric should be in front of $3$
$a_2 = 5 \implies$	47396285	5 numeric should be in front of 2 $$
$a_1 = 1 \implies$	417396285	1 numeric should be in front of 1

Say  $f: S_n \to \prod_{i=n-1}^0 A_i$  be inversion table and g be process of making sequence to permutation. Then, we claim  $f \circ g = id_{\prod_{i=n-1}^0 A_i}$ . Note that g is one-to-one since it is deterministic process. And f is onto since any element in codomain can be rechable by f using construction of g. Done.

Generating function is

$$\sum_{\pi \in \mathfrak{S}n} q^{inv(\pi)} = \sum_{(a_1, \cdots, a_n) \in A_{n-1} \times \cdots \times A_0} q^{\sum a_i} = \sum_{(a_1, \cdots, a_n) \in A_{n-1} \times \cdots \times A_0} q^{a_1} \cdots q^{a_n} = \left(\sum_{a_1 \in A_{n-1}} q^{a_1}\right) \cdots \left(\sum_{a_n \in A_0} q^{a_n}\right) = (1 + q + \cdots + q^{n-1}) \cdots (1 + q)(1).$$

By letting q-integer  $[n]_q$  as

$$[n]_q := (1 + q + \cdots q^{n-1}),$$

we can denote it as

$$\left(\sum_{a_1 \in A_{n-1}} q^{a_1}\right) \cdots \left(\sum_{a_n \in A_0} q^{a_n}\right) = (1+q+\dots+q^{n-1})\cdots(1+q)(1) = [1]_q[2]_q\cdots[n]_q = [n]_q!$$

Any permutation statistics  $\alpha$  with the same distribution as inversion is called *Mahonian Statistics*, i.e.,  $\sum q^{\alpha(\pi)} = [n]_q!$ 

#### 1.5.2 Descents

Let  $\pi = \pi_1 \cdots \pi_n$  be a permutation. (Each  $\pi_i$  denotes permuting *i* such that  $\pi_i = \pi(i)$ ). Then,

#### Definition 1.5.8.

$$DES(\pi) := \{i : \pi_i > \pi_{i+1}\} \subseteq [n-1], des(\pi) = |DES(\pi)|.$$

For example, if  $\pi = 231564$ , then since 3 > 1, 6 > 4,

$$DES(\pi) = \{2, 5\}, des(\pi) = 2.$$

Given  $D \subseteq [n-1]$  let

$$\beta(D) = |\{\pi \in S_n : DES(\pi) = D\}|, \alpha(D) = |\{\pi \in S_n : DES(\pi) \subseteq D\}|$$

If we counting this using multinomial by length of each cycles, this disregards information between cycles.

**Proposition 1.5.9.** If  $D = \{d_1, \dots, d_k\}$  with increasing order, then

$$\alpha(D) = \begin{pmatrix} n \\ d_1, d_2 - d_1, \cdots, d_k - d_{k-1}, n - d_k \end{pmatrix} = \begin{pmatrix} n \\ \Delta d \end{pmatrix}$$

where  $\Delta d$  be a sequence  $d_1, d_2 - d_1, \dots, d_k - d_{k-1}, n - d_k$ , called difference sequence.

*Proof.* Proof is from [Sta11][p.38]. To obtain such a permutation, for the first cycle, take  $d_1$  elements and arrange it with ascending order. It gives  $\binom{n}{d_1}$  way. Secondly, take  $d_2 - d_1$  elements from  $n - d_1$  remainings and arrange it with ascending order. It gives  $\binom{n-d_1}{d_2-d_1}$  way to choose. In this manner there are

$$\binom{n}{d_1}\binom{n-d_1}{d_2-d_1}\cdots\binom{n-d_k}{n-d_k} = \binom{n}{d_1,d_2-d_1,\cdots,d_k-d_{k-1},n-d_k}.$$

Claim 1.5.10.

$$\alpha(D) = \sum_{T \subseteq D} \beta(T)$$

and

$$\beta(D) = \sum_{T \subseteq D} (-1)^{|D-T|} \alpha(T).$$

*Proof.* First equation is just derived from definition of  $\alpha(D)$ . To see the second one, we should use inclusion exclusion principle. To see this,

$$A(D) := \{ \pi \in S_n : DES(\pi) \subseteq D \}, B(D) := \{ \pi \in S_n : DES(\pi) = D \}$$

Then,

$$B(D) = A(D) \setminus \bigcup_{T \subsetneq D} A(T)$$

Thus

$$\beta(D) = |B(D)| = |A(D) \setminus \bigcup_{T \subsetneq D} A(T)| = |A(D)| - \left| \bigcup_{T \subsetneq D} A(T) \right|.$$

Now to see  $\left|\bigcup_{T\subsetneq D} A(T)\right|$ , note that from the inclusion,

$$\bigcup_{T\subsetneq D} A(T) = \bigcup_{T\subsetneq D, |T| = |D| - 1} A(T)$$

since for any proper subset of D should be a subset of element |D| - 1. Now call a proper subset of D with cardinality |D| - 1 as **facet**. Letting  $D = \{d_1, \dots, d_k\}$  and giving index for facets, such as

$$T_i = D \setminus \{d_i\}, A_i := A(T_i)$$

Then,

$$\bigcup_{T \subsetneq D, |T| = |D| - 1} A(T) = \bigcup_{i=1}^{k} A_i.$$

By the inclusion exclusion principle,

$$\left|\bigcup_{i=1}^{k} A_{i}\right| = \sum_{j=1}^{k} (-1)^{j+1} \left(\sum_{1 \le i_{1} < \dots < i_{j} \le n} |A_{i_{1}} \cap \dots \cap A_{i_{j}}|\right)$$

And it is clear that

$$A_{i_1} \cap \dots \cap A_{i_j} = A(T_{i_1,\dots,i_j}) \text{ where } T_{i_1,\dots,i_j} = D \setminus \{d_1,\dots,d_j\}.$$

Hence,  $j + 1 = |D - T_{i_1, \dots, i_j}| + 1$ . Since it visits all case of proper subsets of D, we can say

$$\sum_{j=1}^{k} (-1)^{j+1} \left( \sum_{1 \le i_1 < \dots < i_j \le n} |A_{i_1} \cap \dots \cap A_{i_j}| \right) = \sum_{T \subsetneq D} (-1)^{|D-T|+1} \alpha(T).$$

Hence,

$$\beta(D) = \alpha(D) - \sum_{T \subsetneq D} (-1)^{|D-T|+1} \alpha(T) = \alpha(D) + \sum_{T \subsetneq D} (-1)^{|D-T|} \alpha(T) = \sum_{T \subseteq D} (-1)^{|D-T|} \alpha(T).$$

Given a set S, we can linearly list all subsets of S, with bigger ones first. For example, if  $S = \{a, b\}$ , then we can make a matrix with entries  $S_{T_1,T_2} = \begin{cases} 1 & \text{if } T_1 \subseteq T_2 \\ 0 & \text{otherwise} \end{cases}$ 

Then, inver se of this is given by

$$S_{T_1,T_2}^{-1} = \begin{cases} (-1)^{|T_2 - T_1|} & \text{if } T_1 \subseteq T_2 \\ 0 & \text{otherwise} \end{cases}$$

#### 1.5.3 Eulerian Polynomial

**Definition 1.5.11** (Eulerian Polynomial).

$$A_n(x) = \sum_{\pi \in \mathfrak{S}n} x^{1 + des(\pi)}.$$

+1 is to assigning ascending permutation; which gives  $des(\pi) = 0$ . Data are below;

$$A_1(x) = x, A_2(x) = x + x^2, A_3(x) = x + 4x^2 + x^3, A_4(x) = x + 11x^2 + 11x^3 + x^4.$$

Now let

$$A_{n,k} := |\{\pi \in \mathfrak{S}n : des(\pi) = k-1\}|$$

Then,

$$A_n(x) = \sum_k A_{n,k} x^k.$$

Basic properties are below.

•

$$A_{n,k} = A_{n,n-k+1}.$$

*Proof.* It suffices to show that

$$\{\pi\in\mathfrak{S}n:des(\pi)=k-1\}\cong\{\pi\in\mathfrak{S}n:des(\pi)=(n-k+1)-1=n-k\}$$

has a bijection. Suppose  $\pi = a_1 \cdots a_n \in \mathfrak{S}n$  has  $des(\pi) = k - 1$ . Then, let  $\pi_{rev}$  be a reversion of  $\pi$ , i.e.,

$$\pi_{rev} = a_n a_{n-1} \cdots a_1.$$

Then it has  $des(\pi) = n - 1 - (k - 1) = n - k$ , and  $(\pi_{rev})_{rev} = \pi$  for any  $\pi$ , hence this reversing map has a bijection with image. Note that this map gives bijection with image if we think it as a map from  $\{\pi \in \mathfrak{S}n : des(\pi) = (n - k + 1) - 1 = n - k\}$ , thus it gives a bijection.

$$A_{n,k} = kA_{n-1,k} + (n-k+1)A_{n-1,k-1}$$

To see this, note that we can generate  $des(\pi) = k - 1$  permutation by

- 1. Making descent k-1 sequence first using [n-1] and distribute it on the rightmost end (which doesn't add descent) or inside of each descent (for example, if  $a_i a_{i+1}$  form a descent, then  $a_i n a_{i+1}$  form a new descent  $a_i n$  and delete original one, hence number of descent is preserved.) This gives k possible way of assigning n.
- 2. Making descent k-2 sequence and generating new descent by assigning leftmost end or inside of  $a_i a_j$  which are not descent. This gives 1 + (n-k) ways.

From this property, we get

Proposition 1.5.12 (Proposition 1.4.4 in [Sta11]).

$$\sum_{m \ge 0} m^d x^m = \frac{A_d(x)}{(1-x)^{d+1}}.$$

*Proof.* Proof is induction on d. If d = 0, then it holds since  $A_0(x) = 1$  and  $\sum_m x^m = 1/(1-x)$ . Now suppose it holds for some d. Then by differenting and multiplying x for both sides we get

$$\sum_{m \ge 0} m^{d+1} x^m = x \cdot \frac{(1-x)A'_d(x) + (d+1)A_d(x)}{(1-x)^{d+2}}$$

So it suffices to show that

$$A_{d+1}(x) = (1-x)A'_d(x) + (d+1)A_d(x).$$

By seeing it coefficientwise, it suffices to show that

$$A_{d+1,k} = kA_{d,k} + (d-k+2)A_{d,k-1}$$

which is holds from the above property.

Theorem 1.5.13 (Worpitzky's Identity).

$$x^{n} = \sum_{k \ge 1} A_{n,k} \binom{x+n-k}{n} = \sum_{k \ge 1} A_{n,k} \binom{x+k-1}{n}$$

*Proof.* To see the second equality, use  $A_{n,k} = A_{n,n+1-k}$  and change index of sum from k to n + 1 - k. For the first equality, note that let

$$S = \{ (a_1, \cdots, a_n) : 0 \le a_i < x, a_i \in \mathbb{N} \}$$

Then we can count |S| by two ways.

- $|S| = x^n$ .
- From given sequence, we can generate a sequence of order by this; for any sequence  $(a_1, \dots, a_n)$ , we get terms and  $\sigma \in \mathfrak{S}n$

$$1 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq x, \sigma = \sigma = \sigma_1 \cdots \sigma_n$$

where  $\sigma_i$  is position of  $t_i$  in original sequence. For example, if  $\sigma = 213$  and  $t_1 = 1, t_2 = 2, t_3 = 2$  then the original sequence is 212. However, this sequence of  $t_i$  should compatible with  $\sigma$ , which is

compatibility: if 
$$i \in DES(\sigma)$$
 then  $t_i < t_{i+1}$ .

So if we give  $\sigma = 231$  and  $t_1 = 1, t_2 = 2, t_3 = 2$ , still this gives sequence 212, but it is not compatible case since the condition that second minimal one and third minimal one is the same and has position 3 and 1 respectively implies position 1 should be written first and position 3 secondly.

•

Thus,

$$\#\operatorname{seq} = \sum_{\sigma \in \mathfrak{S}n} \#\{(t_1, \cdots, t_n) : (t_1, \cdots, t_n) \text{ is compatible with } \sigma\}$$

Now we should ask that, given  $\sigma \in \mathfrak{S}n$ ,  $DES(\sigma) = \{d_1, \dots, d_{k-1}\}$ , how many such compatible  $(t_1, \dots, t_n)$  exists?

Note that picking 5 nondecreasing sequence from [100] is

$$\left( \begin{pmatrix} 100\\5 \end{pmatrix} \right) = \begin{pmatrix} 100+5-1\\5 \end{pmatrix}$$

If we think about  $t = (t_1, \dots, t_n)$ , then it is nondecreasing sequence. And note that, for example, picking a sequence

$$1 \le t_1 < t_2 \le t_3 \le x$$

is the same as picking

$$1 \le t_1 \le t_2 - 1 \le t_3 - 1 \le t_3$$

a nondecreasing sequence. Hence, by the standard treatment for all positions in  $DES(\sigma)$ , it is equivalent to picking nondecreasing sequence

$$1 \le t_1' \le \dots \le t_n' \le x - (k-1)$$

since  $des(\sigma) = k - 1$ . Hence, for given  $\sigma$ , we had

$$\left(\binom{x-(k-1)}{n}\right) = \binom{x-(k-1)+n-1}{n} = \binom{x+n-k}{n}.$$

Hence,

$$x^n = \sum_{\sigma \in \mathfrak{S}n} {\binom{x+n-k}{n}}$$
 where  $des(\sigma) = k-1$ 

and since number of  $\sigma \in \mathfrak{S}n$  with  $des(\sigma) = k - 1$  is  $A_{n,k}$ , thus

$$x^{n} = \sum_{k=1}^{n-1} A_{n,k} \binom{x+n-k}{n}$$

Any permutation statistics with the same distribution as  $des(\sigma)$  is called *Eulerian Statistics*.

Claim 1.5.14. For  $\pi = \pi_1 \pi_2 \cdots \pi_n \in \mathfrak{S}n$ , let

$$EXC(\pi) = \{i : \pi_i = \pi(i) > i\}, weak EXC(\pi)\{i : \pi_i = \pi(i) \ge i\}$$

which is called **exceedances**. It is Eulerian, i.e., the number of permutations in  $\mathfrak{S}n$  with k exceedances is  $A_{n,k+1}$  (so is that with k+1 weak exceedances.)

*Proof.* See [Sta11][p.39-40]. Let  $\pi \in \mathfrak{S}n$  and represent it as standard notation; then

$$\pi = (a_1, \cdots, a_i)(a_{i+1}, \cdots, a_{i_1})(a_{i_1+1}, \cdots, a_{i_2})\cdots(a_{i_{k-1}+1}, \cdots, a_{i_k}).$$

Then, if we permuting  $a_i$  using  $\pi$ , there are two possible results;  $a_i$  goes to  $a_{i+1}$  or  $a_i$  goes to back, for example, if  $i_l \leq i \leq i_{l+1}$  then it goes to  $a_{i_l}$ . And from the standard notation, this means  $a_{i_l} = \pi(a_i) > a_i$  since each cycle should start with the maximal element. And since

$$a_1 < a_{i_1} < \cdots < a_{i_k},$$

we get

$$a_i < a_{i_l} < a_{i_{l+1}} = a_{i+1}.$$

Thus we can summarize this result as

If 
$$\pi(a_i) \neq a_{i+1}$$
, then  $a_i < a_{i+1}$ .

So if  $a_i < a_{i+1}$  or i = d, then permuting  $a_i$  by  $\pi$  gives  $a_{i_l}$ , which is greater than or equal to  $a_i$  or gives  $a_{i+1}$ , which is also greater than or equal to  $a_i$ , thus

$$\pi(a_i) \ge a_i$$

, where equality comes from the case when cycle containing  $a_i$  is singleton.

Conversely, if  $w(a_i) \ge a_i$ , if equal it implies cycle containing  $a_i$  is singleton, hence  $a_{i+1} > a_i$  by standard notation or i = d. If  $w(a_i) > a_i$  then in case of  $w(a_i) = a_{i+1}$  gives  $a_{i+1} > a_i$  or in case of  $w(a_i) = a_{i_l}$  gives  $a_i < a_{i_l} < a_{i_{l+1}} = a_{i+1}$ , thus we can say that

$$a_i < a_{i+1}$$
 or  $i = d$  iff  $w(a_i) \ge a_i$ .

Hence,

$$d - des(\pi) = |\{i \in [d] : \pi(i) \ge i\}|$$

which is called a set of *weak exceedance of*  $\pi$ . Now to get a relationship between weak exceedance and exceedance of  $\pi$ , we should represent  $\pi$  in usual notation, i.e.,

$$\pi = \pi_1 \cdots \pi_d$$
 where  $\pi_i = \pi(i)$ .

permuting i. Then let

$$u_i = d + i - w_{d-i+1}.$$

Note that

$$\begin{array}{l} u_{d-i+1} < d-i+1 \iff d+1 - w_{d-(d-i+1)+1} < d+1 - i \iff i < w_i \\ u_{d-i+1} > d-i+1 \iff d+1 - w_{d-(d-i+1)+1} > d+1 - i \iff i > w_i \\ u_{d-i+1} < d-i+1 \iff d+1 - w_{d-(d-i+1)+1} < d+1 - i \iff i = w_i \end{array}$$

Thus, if w has k weak excedance, then corresponding terms in u does not have any excedance, which implies d - k characters of u which is from no weak excedance turns out to be excedance.

Also note that  $w \to u$  is bijection since applying this transformation twice we can always get the same permutations, and it is definitely one-to one by construction.

Now, if w has k excedances, then its corresponding u has d - k weak excedances, so u has k descents. Since it is the same sum of counting excedances of w or counting descents of u by permuting all possible elements in  $\mathfrak{S}n$ , so  $A_{n,k+1}$  is the number of all permutations in  $\mathfrak{S}n$  with k excedances. Similarly, if w has k + 1 weak excedances, then w has d - k - 1 descents, so the reversing of w has k descents. Since reversing is also a bijection, it is equivalent to counting permutations with k + 1 weak excedances is to counting permutations with k descents. Done.

**Definition 1.5.15** (Major index). If  $\pi = \pi_1 \cdots \pi_n \in \mathfrak{S}n$ , then

$$maj(\pi) = \sum_{i \in DES(\pi)} i$$

For example,

$\sigma \in S_3$	$DES(\sigma)$	$maj(\sigma)$	$inv(\sigma)$
123	Ø	0	0
132	$\{2\}$	2	1
213	$\{1\}$	1	1
231	$\{2\}$	2	2
312	$\{1\}$	1	2
321	$\{1,2\}$	3	3

Claim 1.5.16 (Proposition 1.4.6 in [Sta11]).

$$\sum_{\sigma \in S_n} q^{maj(\sigma)} = \sum_{\sigma \in S_n} q^{inv(\sigma)} = [n]_q!$$

*Proof.* We already saw that  $\sum_{\sigma \in S_n} q^{inv(\sigma)} = [n]_q!$ . So it suffices to show that

$$\sum_{\sigma \in S_n} q^{maj(\sigma)} = [n]_q!$$

Use induction. Given  $\sigma = \sigma_1 \cdots \sigma_{n-1} \in \mathfrak{S}n - 1$ , insert *n* in *n* possible ways and check the major index on the *n* new permutations. For example, let  $\pi = 356214$ . Then

$$DES(\pi) = \{3, 4\} \implies maj(\pi) = 3 + 4 = 7.$$

By adding 7 on  $\pi$  we get

new $\sigma$	DES(new $\sigma$ )	$  maj(\sigma)  $
356214 <mark>7</mark>	$\{3, 4\}$	7+0
3562174	$\{3, 4, 6\}$	7+6
3562 <b>7</b> 14	$\{3, 5\}$	7+1
356 <mark>7</mark> 214	$\{4, 5\}$	7+2
35 <b>7</b> 6214	$\{3, 4, 5\}$	7 + 5
3 <b>7</b> 56214	$\{2, 4, 5\}$	7+4
<b>7</b> 356214	$\{1, 4, 5\}$	7+3

To generalize this results, let  $\pi = \pi_1 \cdots \pi_{n-1} \in \mathfrak{S}n - 1$ , and think about the case of adding n. Let  $A_{\pi} = \{\pi' \in \mathfrak{S}n : \pi' \text{ is generated by adding } n \text{ to fixed } \pi \in \mathfrak{S}n - 1\}$ . Let

$$\psi: A_{\pi} \to \mathbb{N}$$
 by  $\psi(\pi') = maj(\pi')$ 

What we want to show is that

$$\psi(A_{\pi}) = maj(\pi) + [n-1] \cup \{0\}.$$

If this is true, then

$$\sum_{\sigma \in \mathfrak{S}n} q^{maj(\sigma)} = \sum_{\sigma \in \mathfrak{S}n-1} (q^{maj(\sigma) + \sum_{i=0}^{n-1} i} = [n]_q \sum_{\sigma \in \mathfrak{S}n-1} = [n]_q ([n-1]_q!) = [n]_q!$$

where last equality comes from inductive hypothesis. (Also note that the first step holds when n = 1.)

To see this, all we left is just counting; let  $DES(\pi) = \{d_1, \dots, d_k\}$  where

$$d_0 = 0 < d_1 < \dots < d_k < n = d_{k+1}.$$

If n is added on  $(d_j + 1)$ -th position, then it gives

$$\pi' = \cdots \pi_{d_i} n \pi_{d_i+1}$$

since  $\pi_{d_j} < n > \pi_{d_j+1}$ ,

$$DES(\pi') = (DES(\pi) \setminus \{d_j, d_{j+1}, \cdots, d_k\}) \cup \{d_j + 1, d_{j+1} + 1 \cdots, d_k + 1\} \implies maj(\pi') = maj(\pi) + k - j + 1.$$

$$\psi(DES(\pi)) = maj(\pi) + \{1, 2, \cdots, k\}.$$

If n is added on  $d_{k-j} + 1 < i < d_{k-(j-1)} + 1$  position for  $j = 1, 2, \dots, k$ , then, it moves descents  $d_{k-(j-1)}, \dots, d_k$  to  $d_{k-(j-1)} + 1, \dots, d_k + 1$  respectively, and make new descent *i*. Thus,

$$DES(\pi') = \left(DES(\pi) \setminus \{d_{k-(j-1)}, \cdots, d_k\}\right) \cup \{i\} \cup \{d_{k-(j-1)}+1, \cdots, d_k+1\} \implies maj(\pi') = maj(\pi) + i + j$$

Hence,

$$\psi(\{i: d_{k-j}+1 < i < d_{k-(j-1)}+1\}) = maj(\pi) + j + (d_{k-j}+1, d_{k-(j-1)}+1) \cap \mathbb{N}, \text{ for } j = 1, 2, \cdots, k$$

For j = 0 case, for adding n on the last position, which is n, then it gives  $maj(\pi') = maj(\pi) + 0$  since no new descents occur; otherwise, if  $\{i : d_k + 1 < i < d_{k+1} = n\}$  is not empty in  $\pi$ , then adding n on such is does not change previous descents, but gives new descent i. Hence,

$$\psi(\{i: d_{k-j} + 1 < i < d_{k-(j-1)} + 1\}) = \begin{cases} \{0\} & \text{if given set is empty} \\ \{d_k + 2, \cdots, n-1\} \cup \{0\} \end{cases}$$

And if given set is empty, then  $d_k = n - 2$  so that  $maj(\pi') = maj(\pi) + (n - 2) + 1 = maj(\pi) + n - 1$  occur in j = 1 case with adding n in  $d_k$  position. Now by counting with j, we get

j	$\psi(\{i: d_{k-j} + 1 < i < d_{k-(j-1)} + 1\})$
0	$\int maj(\pi) + \{0\}$ if given set is empty,
0	$maj(\pi) + \{d_k + 2, \cdots, n-1\} \cup \{0\}$ nonempty.
1	$\int maj(\pi) + \{d_{k-1} + 3, d_{k+1} = n - 1\}  \text{if } j = 0 \text{ case is } \{0\},$
1	$maj(\pi) + \{d_{k-1} + 3, \cdots, d_k + 1\}$ otherwise
2	$maj(\pi) + \{d_{k-1} + 4, \cdots, d_{k-1} + 2\}$
3	$maj(\pi) + \{d_{k-1} + 5, \cdots, d_{k-1} + 3\}$
k-1	$maj(\pi) + \{d_1 + k + 1 \cdots, d_2 + k - 1\}$
k	$maj(\pi) + \{k+1, \cdots, d_1 + k\}$

Thus,

$$\psi([n-1] \setminus DES(\pi)) = maj(\pi) + ([n-1] \setminus [k]) \cup \{0\}$$

and from the above result  $\psi(DES(\pi)) = maj(\pi) + [k]$ , done.

#### 1.6 Composition

Let  $x_1 + \cdots + x_k = n$ . Then the number of  $\mathbb{N} \cup \{0\}$ -solutions is  $\binom{n+k-1}{n}$  by bar-ball tricks. And its  $\mathbb{P}$ -solution is  $\binom{n-1}{k-1}$ , since now we should assign + (bar) only n-1 position without repeat.

**Definition 1.6.1** (Composition). Any  $(x_1, \dots, x_k), x_i \in \mathbb{P}, i \in [k]$  with sum  $\sum_{i=1}^k x_i = n$  is called a composition of n.

For example, 2 + 3 + 2 = 7 is a composition of 7.

Number of total composition of n is  $2^{n-1}$ , since it is just number of (nonempty) subsets of [n-1], by thinking fixing n-1 positions among n balls which used for putting + sign. Then the number of balls in each section divided by + sign represents composition.

For notation, if we let  $\lambda$  be an (unordered) composition of n by  $\lambda_1, \dots, \lambda_k$ , then we denote

$$\lambda = (\lambda_1, \cdots, \lambda_k)$$
 where  $\lambda_1 \ge \cdots \ge \lambda_k \ge 1$ .

and say that

$$\lambda \vdash n$$

Note that k is nontrivial number, and  $\lambda$  is integer partition. For example,

$$4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1.$$

Let

$$P(n,k) := \{\lambda : |\lambda| = n, \lambda \text{ has } k \text{ parts.}\}, \quad p(n,k) = |P(n,k)|, \quad p(n) := \sum_{k \ge 0} p(n,k)$$

Now, *Ferrer's diagram* (or *Young's diagram*) is a way of representing (unordered) partition. This diagram is a set of blocks constructed as follow; Fix starting point (or block). Then, make  $\lambda_1$  many squares (young diagram) or dots (Ferre diagram) on a first row, from the starting block to right. So the first row has  $\lambda_1$  blocks including the block at starting point. Now, make  $\lambda_j$  many blocks, which starts at the below of starting block of (j-1)-th row to the right. For example, if  $\lambda = (6, 4, 2, 2, 1)$ , then



If we read it vertically i.e., counting number of blocks in a row from left to right, we can get 5, 4, 2, 2, 1, 1 which is another partition of 15. So we call this as *conjugate* of  $\lambda$ , and denote it as

$$\lambda^* = (5, 4, 2, 2, 1, 1)$$

Note that

• p(n,k) = p(n-1,k-1) + p(n-k,k). To see this we can decompose P(n,k) as 1) a partition containing 1 and 2) a partition not containing 1. To generate a partition containing 1, then make partition of n-1 with k-1 parts, and just add 1 for k-th part. To generate a partition not containing 1, then to make partition of n-k with k parts, then add 1 for each part. This gives a partition of n having no 1.

Note that this relation yields definition of p(n, 0) = 0, p(0, 0) = 1  $(n \ge 1)$ , p(n, k) = 0 if k > n.

٠

 $P(n, \leq k) \stackrel{\lambda^*}{\longleftrightarrow} P(n, \text{ the largest part is of size } \leq k).$ 

To see this, just see Young diagram of  $\lambda$ . Since  $\lambda$  has at most k parts, the columns has at most k, and they are bijection. So,

$$\sum_{n \ge 0} p(n, \le k) x^n = \sum_{\lambda \text{ at most } k \text{ part}} x^{|\lambda|} = \sum_{\lambda^* \text{ largest part is of size } \le k} x^{|\lambda^*|}$$

If  $\mu$  has the largest part  $\leq k$ , then  $\mu$  is a multiset such that

$$\mu = \{k^{n_k}, (k-1)^{n_{k-1}}, \cdots, 1^{n_1}\}$$

which means  $n_k$  many k,  $n_{k-1}$  many  $k-1, \dots, n_1$  many 1, where  $n_i \in \mathbb{N}$ . Then,

$$\sum_{\mu = \{k^{n_k}, (k-1)^{n_{k-1}}, \dots, 1^{n_1}\}} x^{|\mu|} = \sum_{\mu = \{k^{n_k}, (k-1)^{n_{k-1}}, \dots, 1^{n_1}\}} x^{n_1 + 2n_2 + \dots + kn_k} = \prod_{i=1}^k \sum_{n_i=0}^\infty x^{in_i} = \prod_{i=1}^k \frac{1}{1-x^i}.$$

Thus,

$$\sum_{n \ge 0} p(n)x^n = \prod_{i=1}^{\infty} \frac{1}{1 - x^i}.$$

Now denote

 $P(*, \leq m, \leq n) := \{\lambda : \lambda \text{ has no more than } m \text{ parts and each part is } \leq n.\}$ 

Actually,  $P(*, \leq m, \leq n)$  has bijection with a set of all lattice paths from (0,0) to (m,n). To see this, think a young diagram and its boundary; if we assume that each block' edges has distance 1 and put the left upper vertice of blocks on (0, n) points in the coordinate space, then the rightmost vertice has (m, n) in

the coordinate space. And its path from (0,0) to (m,n) along the right side of the diagram gives a lattice path. It's lattice path is unique up to young diagram, thus those sets have a bijection between them.

Since we know that the number of lattice paths from (0,0) to (m,n) is  $\binom{n+m}{n}$ ,

$$p(*, \le m, \le n) = \binom{m+n}{n}.$$

And you can show it in the homework that

$$\sum_{\lambda \in P(*, \leq m, \leq n)} q^{|\lambda|} = \binom{m+n}{n}_q$$

#### 1.7 12fold way

Let  $f: N \to R$  where |N| = n, |R| = m. There are 12 cases, like this;

$(N,R) \setminus f$	all function	1-1	onto
N:distinguishable R:distinguishable	$r^n$	$r^{\underline{n}}$	r!S(n,r)
N:indistinguishable R:distinguishable	$\binom{n+r-1}{n}$	$\binom{r}{n}$	$\binom{n-1}{r-1}$
N:distinguishable R:indistinguishable	$S(n, \le r) = \sum_{i=1}^{r} S(n, i)$	$\chi(n \le r)$	S(n,r)
N:indistinguishable R:indistinguishable	$p(n, \leq r)$	$\chi(n \le r)$	p(n,r)

To see this..

- Case 1: N, R are distinguishable. Then all possible function is just assigning r to one of element in N hence each elements of R has n possibility, thus  $r^n$ . For 1-1 functions, first element in R has n choices, but second one has n-1 choices since first element takes at least one elements in N, and so on. Hence,  $r^{\underline{n}}$  possible ways. For onto case, first of all, we should partitioning N by r nonzero subsets. This gives S(n,r) possibilities, where S(n,r) is the Stirling number of second kind. Then there are r! ways to assigning elements in R to given r partitions of N.
- Case 2: N is indistinguishable, and R is distinguishable. Then, actually,  $f: N \to R$  and  $g: N \to R$  is distinct if and only if  $|f^{-1}(r)| = |g^{-1}(r)|$  for all  $r \in R$ . Thus, it is equivalent to just partitioning the same balls into r parts, allowing empty part. Hence it is the same as bar-ball problem with r bar and n ball, therefore  $\binom{n+r-1}{n}$ . For 1-1 correspondence, all preimage of element in R should be singleton or empty set. And there should be n singleton, since f is a function. Thus, it is equivalent to choosing n elements from R whose preimage is singleton, and assign emptyset for the others. So,  $\binom{r}{n}$ . For onto case, note that no preimage of element in R is empty; and from the condition

$$|f^{-1}(r_1)| + \dots + |f^{-1}(r_{|R|})| = n,$$

it is the same as counting all positive solutions of above equation, which is just the number of ordered k-partition of n, which is  $\binom{n-1}{r-1}$ .

- Case 3: N is distinguishable, and R is indistinguishable. Note that f gives a set of r disjoint subsets of N (containing multiple emptysets in general) whose union is N. Since we do not distinguish elements in R, two functions differ if and only if their corresponding a set of r sets are differ. And by removing empty sets from the set, we can regard it as a partition of n with  $\leq r$  parts. So  $S(n, \leq r)$  comes up. For 1-1, the function get r set which contains n singleton and emptysets (if r > n.) This is unique, since we only think f and g differ if they have two distinct sets of r sets. So  $\chi(n \leq r)$  which is characteristic function, is the answer. For the onto function, it gives a set of r sets consisting of nonempty disjoint subset of N, which is just a partition of N. Hence, S(n, r).
- Case 3: N is indistinguishable, and R is indistinguishable. Then all function is just partitioning, since now two functions are differ if their given r set's cardinalities (without ordering) are differ; and this cardinality form a partition number, thus  $p(n, \leq r)$ . 1-1 is just equivalent, since cardinality of singleton is just 1. For onto case, now we even distinguish two sets in r sets if they have the same cardinality; thus it is just p(n, r).

#### **1.8** *q*-Analog and Gaussian Coefficients

**Definition 1.8.1** (q-something). *q*-integer:  $[q]_q = 1 + q + q^2 + \dots + q^{k-1} = \frac{1-q^k}{1-q}$ . *q*-factorial:  $[n]_q = [1]_q[2]_q \dots [n]_q$ . *q*-binomial coefficient:  ${n \brack k}_q = \frac{[n]_q!}{[k]_q![n-k]_q!}$ 

Actually we can derive those from a vector spaces over  $\mathbb{F}_q$ . Use notation that

 $V_q(n) :=$  vector space of dimension n over finite field  $\mathbb{F}_q$ .

Then

$$V_q(n) = \{(\alpha_1, \cdots, \alpha_n) : \alpha_i \in \mathbb{F}_q\}$$

**Proposition 1.8.2.**  $\begin{bmatrix} n \\ k \end{bmatrix}_{q} = the number of k-dimensional subspace of <math>V_q(n), 0 \le k \le n$ .

Proof. Let U be a k-dimensional subspace of  $V_q(n)$ . So U has some basis  $\langle e_1, \cdots, e_k \rangle$  where  $e_i \in U, \{e_i\}_{i=1}^k$ is linearly independent from  $V_q(n)$ . Now counting the number of ordered list. For  $e_1$ , there are  $q^n - 1$  choices, since any vector except 0 is linearly independent. If we choose  $e_1$ , then for  $e_2$ , there are  $q^n - q$  choices since we should get rid of any element which is linear combination of  $e_1$ , and the number of all such combination is q. Similarly, for  $e_j$ , then there are  $q^n - q^{j-1}$  choices where  $q^{j-1}$  stands for the number of all linear combination of  $e_1, \cdots, e_{j-1}$ . Hence, for  $e_k$ , there are  $(q^n - q^{k-1})$  choices. Thus, the number of ordered basis in  $V_q(n)$  for k-dimensional subspace is

$$(q^n-1)(q^n-q)\cdots(q^n-q^{k-1}).$$

Now fix U. Then, how many ordered basis exists for such fixed U? We can do the same argument replacing  $V_q(n)$  by U. Thus it gives

$$(q^k-1)\cdots(q^k-q^{k-1}).$$

Thus, its quotient is

$$\frac{(q^n-1)(q^n-q)\cdots(q^n-q^{k-1})}{(q^k-1)\cdots(q^k-q^{k-1})} = \frac{\frac{(q^n-1)(q^n-q)\cdots(q^n-q^{k-1})}{(1-q)^k}}{\frac{(q^k-1)\cdots(q^k-q^{k-1})}{(1-q)^k}} = \frac{[n]_q\cdots[n-k+1]_q}{[k]_q!} = \frac{[n]_q!}{[k]_q![n-k]_q!}.$$

#### Corollary 1.8.3.

1. 
$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n \\ n-k \end{bmatrix}_q$$
.  
2.  $\begin{bmatrix} n \\ k \end{bmatrix}_q = q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k \end{bmatrix}_q$ 

*Proof.* For 1, Any k-dimensional vector space is complement to some n-k dimensional dimensional vector space.

For 2, we can count k-dimensional space of  $V_q(n)$  in another way; first of all, fix a basis  $\langle e_1, \cdots, e_n \rangle$  of  $V_q(n)$ , and let W be a subspace generated by a basis  $\langle e_1, \cdots, e_{n-1} \rangle$ . Then

1. To count the number of k dimensional subspace which is contained in W, it is just  $\binom{n-1}{k}_{a}$ .

To count the number of k-dimensional subspace U which is not contained in W, let's say U be such space. Then, since dim W = n - 1, dim $(U \cap W) = k - 1$ . Thus, U is spanned by basis of  $U \cap W$  and f where  $f \notin W$ , so first of all, count all possible basis for  $U \cap W$ , which gives  $\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q$ . Now for f, if we assume  $f = e_n + \sum_{i=1}^{n-1} c_i e_i$  by scaling any  $f \notin W$ , then there exists  $q^{n-1}$  possible f, by choosing  $c_1, \dots, c_{n-1} \in \mathbb{F}_q$ . And, if

$$\operatorname{span} U \cap W \cup \{f_1\} = \operatorname{span} U \cap W \cup \{f_2\}$$

then  $f_1 - f_2 \in U \cap W$ . And note that  $U \cap W$  consists of  $q^{k-1}$  vectors. Thus, for any possible f up to scaling, f + v where  $v \in U \cap W$  will results in the same vectorspace. Hence,

Number of different 
$$U = \begin{bmatrix} n-1\\ k-1 \end{bmatrix}_q \frac{q^{n-1}}{q^{k-1}} = \begin{bmatrix} n-1\\ k-1 \end{bmatrix}_q q^{n-k}.$$

There are two more combinatorial interpretation of  $[ ]_q$ .

1. Lattice path: We already know that the number of lattice paths from (0,0) to (m,n) is  $\binom{n+m}{n}$ .

Claim 1.8.4.

$$\sum_{L:(0,0)\to(m,n)} q^{area(L)} = \begin{bmatrix} m+n\\n \end{bmatrix}_q,$$

where area(L) means the area generated by the lattice path L and line segments [(0,0), (0,n)] and [(0,n), (m,n)].

*Proof.* Let N = m + n. We want to show surjection and some kind of functional from *m*-dimensional subspace of  $V_q(N)$  to the *q*-lattice path. Now start from a *m*-dimensional subspace *K*. Fix a basis in  $V_q(N)$ . Then, *K* is represented by  $m \times N$  matrix consisting of basis on a row. (So row is the coordinate of basis with respect to fixed basis of  $V_q(N)$ .) Now we can transform this matrix to the reduced row echelon form. Then, since rows are linearly independent,

- (a) Each row starts with leading 1.
- (b) First nonzero entry in row i + 1 is on the right of first nonzero entry in row i.
- (c) For each column containing leading 1, 1 is the only nonzero entry.

Thus, by get rid of columns containing 1, we get a new  $m \times n$  matrix. And rewrite \* for entries which may not be zero. For example, if m = 4, n = 3, N = 7,

$$n \times N \text{ matrix} \to \begin{pmatrix} 1 & * & 0 & 0 & * & 0 & * \\ 0 & 0 & 1 & 0 & * & 0 & * \\ 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \end{pmatrix} \to \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix}$$

If we reflect the matrix about y-axis, we get a Ferre-diagram consisting of \*, hence it gives partition in  $P(*, \leq m, \leq n)$ . i.e., partition has no more than m parts and each part is less than or equal to n. In the example, we get  $\lambda = (3, 2, 2, 1)$ .

If the pivot positions are  $1 = a_1 < a_2 < \cdots < a_n \leq N$ , then

$$\lambda_i = n - a_i + i = n - (a_i - i).$$

To see this, from N columns, get rid of m columns containing leading 1, we get n. Now, we claim that  $a_i = i$  gives n stars; otherwise, by row echelon form, one column containing star which is left to the  $a_i$ -th column exists; done. Now, to see original statement, note that the number of star is decreasing by 1 when  $a_i$  changed to  $a_i + 1$  by the definition of row-echelon form. (Just see the picture.) Thus the equation holds.

Thus, each element in  $P(*, \leq m, \leq n)$  corresponds to some set of *m*-dimensional subspace of  $V_q(N)$ . Note that for each given  $\lambda \in P(*, \leq m, \leq n)$ ,  $\lambda$  can generates  $q^{area(\lambda)}$  reduced row echelon form matrix, and each matrix of these corresponds to the *m*-dimensional subspace. Hence,

$$\begin{bmatrix} n+m\\n \end{bmatrix}_q = \sum_{U \le V_q(N)} 1 = \sum_{\text{partition of } U \text{ which gives the same } \lambda} q^{area(\lambda)} = \sum_{\lambda \in P(*, \le m, \le n)} q^{area(\lambda)} = \sum_{L:(0,0) \to (m,n)} q^{area(L)} = \sum_{L$$

Corollary 1.8.5.

$$\sum_{\lambda \in P(*, \leq m, \leq n)} q^{|\lambda|} = {n+m \choose n}_q$$

*Proof.* Note that  $area(\lambda) = |\lambda|$  since number of points in Ferrer diagram of  $\lambda$  is just area of  $\lambda$ . **Theorem 1.8.6** (*q*-binomial Theorem).

$$(1+xq)(1+xq^1)\cdots(1+xq^n) = \sum_{k=0}^n {n \brack k}_q q^{\binom{k+1}{2}} x^k.$$

*Proof.* We already know that

$$\sum_{n \ge 0} p(n) x^n = \prod_{i=1}^{\infty} \frac{1}{1 - x^i} = \sum_{\lambda \text{ has } n_i \text{ parts of size i}} x^{|\lambda|},$$

and

$$\sum_{\lambda \text{ has distinct } \leq k \text{ parts with each part is less than } k} = \prod_{i=1}^k (1+x^i).$$

(We already showed first one; for the second one, note that RHS generates all possible distinct  $\leq k$  partitions where each part is less than k.)

Now from given equation,

$$(1+xq)(1+xq^1)\cdots(1+xq^n)$$

each term generated by multiplying all chosen term from the box  $(1 + xq^j)$  is like

$$q^l x^k$$

and we can get combinatorial meaning from this; if we choose  $xq^{j_1}, \dots, xq^{j_k}$ , then  $(j_k, \dots, j_l)$  form a partition having distinct k parts and each part is less than n. Hence,

$$(1+xq)(1+xq^{1})\cdots(1+xq^{n}) = \sum_{\substack{\lambda \text{ has distinct part}\\ \text{ each part} \leq n}} x^{\# \text{ of parts}} q^{|\lambda|} = \sum_{k=0}^{n} x^{k} \left(\sum_{\substack{\lambda \text{ has } k \text{ distinct part}\\ \text{ each part} \leq n}} q^{|\lambda|}\right)$$

To counting such  $\lambda = (\lambda_k > \cdots > \lambda_1 > 0)$ , let  $\mu_i = \lambda_i - i$ . Then,

$$n-k>\mu_k\geq\cdots\geq\mu_1\geq 0).$$

Hence,  $\mu$  is a partition having less than or equal to k parts and each part is less than or equal to n-k. Thus,

$$\lambda = \mu + (k, k - 1, \cdots, 1).$$

Thus,

$$|\lambda| = |\mu| + \binom{k+1}{2}.$$

So,

$$\sum_{k=0}^{n} x^{k} \left( \sum_{\substack{\lambda \text{ has } k \text{ distinct part} \\ \text{ each part} \leq n}} q^{|\lambda|} \right) = \sum_{k=0}^{n} x^{k} \left( \sum_{\mu \in P(*, \leq k, \leq n-k} q^{|\mu| + \binom{k+1}{2}} \right) = \sum_{k=0}^{n} x^{k} q^{\binom{k+1}{2}} \sum_{\mu \in P(*, \leq k, \leq n-k} q^{|\mu|}$$
$$= \sum_{k=0}^{n} x^{k} q^{\binom{k+1}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_{q}.$$

where last equality comes from the above corollary.

2. Inversion of words. Let M be a multiset

$$M:=\{1^{a_i},\cdots,k^{a_k}\},\$$

and  $n = \sum_{i=1}^{k} a_i$ . Then, let

 $S(M) := \{ \text{ all distinct permutation of } M \}.$ 

For example, if  $M = \{1^2, 2^2\}$ 

$$S(M) = \{1122, 1212, 1221, 1212, 2121, 2211\}.$$

Then their inversion is

Note that

$$|S(M)| = \binom{n}{a_1, \cdots, a_k}.$$

Claim 1.8.7.

$$\sum_{w \in S(M)} q^{inv(v)} = \begin{bmatrix} n \\ a_1, \cdots, a_k \end{bmatrix}_q := \frac{[n]_q!}{\prod_{i=1}^k [a_i]_q}.$$

*Proof.* We'll show that

$$\prod_{i=1}^{k} [a_i]_q \sum_{w \in S(M)} q^{inv(v)} = [n]_q.$$

Now construct a map

$$S(a_1) \times S(a_2) \times \cdots \times S(a_k) \times S(M) \to S(n)$$

such that for  $\pi_{a_1} \times \cdots \times \pi_{a_k} \times b \in S(a_1) \times S(a_2) \times \cdots \times S(a_k) \times S(M)$ ,

replacing js in b with  $\pi_{a_j} + a_{j-1}$  from left to right,  $j \in \{1, \dots, k\}, a_0 = 0$ .

For example, if  $M = \{1^3, 2^4\}$ , then

$$S_3 \times S_4 \times S(M) \to S(7)$$
 by  $b_1 b_2 b_3 \times c_1 c_2 c_3 c_4 \times 2121212 \mapsto \begin{array}{cccc} 1 & 3 & 2 \\ 7 & 4 & 5 & 6 \end{array} = 7143526.$ 

Note that it gives bijection; for example, from S(n), we can have a map to  $S(a_1) \times S(a_2) \times \cdots \times S(a_k) \times S(M)$  by, for  $b \in S(n)$ 

change elements of b between  $a_{i-1}$  and  $a_i$  to i, for  $i \in [k]$  to get  $\pi \in S(M)$  and collecting all elements of b corresponding to i in  $\pi$  and let it be  $\pi_i \in S(a_i)$ .

This recovers the original sequence always, and we already know that

$$|S(a_1) \times S(a_2) \times \dots \times S(a_k) \times S(M)| = \left(\prod_{i=1}^k a_i!\right) \binom{n}{a_1, \cdots, a_k} = n! = |S(M)|$$

so it is bijection done.

Also, observe that inversion of corresponding sequence is preserved; i.e., if  $\pi_1 \times \pi_2 \times \cdots \times \pi_k \times w \mapsto \pi$ , then inversion occur in  $\pi_j$  also occus in  $\pi$  with same position, and vice versa because order of elements is preserved. Also, if inversion occur in w, then inversion occur in  $\pi$  and vice versa, because the

corresponding elements in  $\pi$  of the bigger element in w have been added by  $a_i$  which assures that the corresponding element in  $\pi$  bigger than the other corresponding one. Hence,

$$inv(\pi) = \sum_{i=1}^{k} inv(\pi_i) + inv(w).$$

Thus, by formular from inversion,

$$[n]_{q}! = \sum_{\pi \in S_{n}} q^{inv(\pi)} = \sum_{(\pi_{1}, \cdots, \pi_{n}, w)} q^{\sum_{i=1}^{k} inv(\pi_{i}) + inv(w)} = \left(\prod_{i=1}^{k} \sum_{\pi_{i} \in S(a_{i})} q^{inv(\pi_{i})}\right) \sum_{w \in S(M)} q^{inv(w)} = \prod_{i=1}^{k} [a_{i}]_{q} \sum_{w \in S(M)} q^{inv(w)}$$

1

gives

$$\sum_{w \in S(M)} q^{inv(w)} = \frac{[n]_q!}{\prod_{i=1}^k [a_i]_q} = \begin{bmatrix} n\\a_1, \cdots, a_k \end{bmatrix}_q.$$

Also note that given  $M = \{1^k, 2^{n-k}\}$ , we can make a lattice paths by relating 1 be horizontal and 2 be vertical; so we can use k horizontal steps and n-k vertical steps to go (k, n-k) from (0, 0). And this lattice path bijectively corresponds to permutation of M.

# 2 Formal Power Series and Infinite Matrices

#### 2.1 Ordinary and Exponential generating function.

Let  $f: \mathbb{N} \to \mathbb{N}$  be a function for counting. Then, we can represent it as  $\{f(n)\}_{n=0}^{\infty}$ . Now let

$$F(x) := \sum_{n \ge 0} f(n) x^n$$

It is called *ordinary generating function*. Since we don't care about convergence,  $F(x) \in \mathbb{C}[[x]]$ , which is a space of formal power series. Hence, F + G and  $\alpha F$  is well defined for  $G \in \mathbb{C}[[x]], \alpha \in \mathbb{C}$ .

If  $A(x) = \sum a_n x^n$ ,  $B(x) = \sum b_n x^n$ , then let

$$c_n = \sum_{k \ge 0} a_k b_{n-k}.$$

Then,

$$C(x) = \sum_{n \ge 0} c_n x^n = A(x)B(x).$$

This comes from just counting coefficient of A(x)B(x).

Also, we can differentiate and integrate the function, and some formal power series contains inverse. Actually, there are two kind of inverse. One is  $\frac{1}{A(x)}$  and the other is, given A(x), there exists B(x) such that A(x)B(x) = 1. Note that B(x) exists if and only if  $A(0) = a_0 \neq 0$ . To see this, if  $a_0 \neq 0$  then take  $b_0 = \frac{1}{a_0}$  and we can solve system of linear equation for coefficient of  $x^n$  in A(x)B(x) by induction. Conversely, if such B(x) exists, then  $b_0a_0 = 1$  since it is only constant term of product, which implies  $a_0 \neq 0$ .

We can also define exponential generating function; for given  $\{f_n\}_{n\geq 0}$  we get

$$\hat{F}(x) := \sum_{n \ge 0} f_n \frac{x^n}{n!}.$$

Then,

$$\hat{C}(x) = \hat{A}(x)\hat{B}(x) \iff c_n = \sum_{k\geq 0} \binom{n}{k} a_k b_{n-k}.$$

since n! term is canceled out by  $\frac{x^n}{n!}$  part in  $\hat{C}(x)$ . This means choose something by A structure and put B structure on the remaining.

### 2.2 *q*-analog of generating function.

We can also define

$$F_q(x) := \sum_{n \ge 0} f(n) \frac{x^n}{[n]_q!}.$$

Hence, by the same argument,

$$C_q(x) = A_q(x)B_q(x) \iff c_n = \sum_{k\geq 0} \begin{bmatrix} n \\ k \end{bmatrix}_q a_k b_{n-k}.$$

And note that if we set P be the lattice of all finite-dimensional subspaces of a vector space of infinite dimension over  $\mathbb{F}_q$ , ordered by inclusion. Then, **factorial function** of P is  $[n]_q$ , and this poset is called **binomial poset**. Intuitively, this is a poset such that for any  $x, y \in P$  with  $x \leq y$ , the interval [x, y] only depends on rank(y). See more detail on [Sta11][p.364].

Also there is a sequence of nonzero number  $\{q_i\}$ . Then, let  $Q_n = \prod_{i=1}^n q_i$ . We can also make generating function

$$\sum_{n\geq 0} f_n \frac{x^n}{Q_n}.$$

Note that  $Q_n$  can be regarded as the number of maximum choice in  $\{x_1, \dots, x_n\}$  where  $x_i$  should be chosen only  $q_i$  numbers.

**Remark 2.2.1** (Coefficient). If  $F(x) = \sum_{n>0} a_n x^n$ , then we denote  $a_n$  by

$$a_n =: [x^n]F(X).$$

**Definition 2.2.2** (Convergence). Let  $F_1(x), \dots, be$  sequence of formal power series. We can define

$$F_i(x) \to F(x) \iff \forall n \in \mathbb{N}, [x^n]F_i \text{ eventually } [x^n]F.$$

Note that this implies the sequence  $[x^n]F_i$  converges to constant  $[x^n]F$  in usual sense, i.e.,  $\exists M > 0$  such that  $\forall m > M, [x^n]F_m = [x^n]F$ .

Or we can make equivalent definition.

**Definition 2.2.3** (Degree). Let  $\deg(F) = \min\{i : [x^i] F \neq 0\}$ .

**Claim 2.2.4.**  $F_i \to F$  if and only if  $\lim_{i\to\infty} deg(F(x) - F_i(x)) = \infty$ .

Proof. Suppose  $F_i \to F$ . For j = 1, there exists M > 0 such that  $[x](F(x) - F_i(x)) = 0$  for all i > M, hence  $1 \notin \{j : [x^j](F - F_i) \neq 0\}$  for i > M. Hence  $\lim_{i\to\infty} deg(F(x) - F_i(x)) > 1$ . Now suppose  $\lim_{i\to\infty} deg(F(x) - F_i(x)) > j - 1$  for some j. Then, from convergence, we can take  $M \in \mathbb{N}$  so that  $[x^j](F - F_i) = 0$  for any i > M. This implies  $j \notin \{j : [x^j](F - F_i) \neq 0\}$  for any i > M. Hence,  $\lim_{i\to\infty} deg(F(x) - F_i(x)) > j$ . Thus by induction,  $\lim_{i\to\infty} deg(F(x) - F_i(x)) = \infty$ .

Conversely, if  $\lim_{i\to\infty} deg(F(x) - F_i(x)) = \infty$ , then for any fixed j, there exists M > 0 such that  $[x^j](F - F_i) = 0$  for i > M. Since j was arbitrarily chosen it implies  $F_i \to F$ .

For example, note that

$$\prod_{i\geq 1} \frac{1}{1-x^i} = \prod_{i\geq 1} (1+x^i+x^{2i}+\cdots) = \lim_{k\to\infty} \prod_{i=1}^k (1+x^i+x^{2i}+\cdots) = 1+p_1x^i+p_2x^{2i}+\cdots$$

for some sequence  $p_1$ . Hence to make infinite product make sense, for all N,  $P_N$  is stabilized by finitely many products. Namely, to compute coefficient, we really care about finitely many factors. So

$$\prod_{i=1}^{\infty} (1 + x + \dots + x^i)$$

make sense, and

$$\prod_{i=1}^{\infty} (1 + F_i(x))$$

is well-defined if and only if  $deg(F_i) \to \infty$  as  $i \to \infty$ . To see this, if  $deg(F_i) \to \infty$ , then for any  $n \in \mathbb{N}, \exists M > 0$ such that  $[x^n]F_i = 0$  for all i > M, hence, to calculate coefficient of  $x^n$  we only care about  $F_1, \dots, F_M$ , so we can define the coefficient in finitely many operation. Conversely, if it is well-defined, then we can compute coefficient by finitely many  $F_i$ s, this implies  $[x^n]F_i = 0$  for i > M with fixed M > 0, done.

Also, composition of 2 formal power series

$$A(B(x)) = \sum_{n=0}^{\infty} a_n (B(x))^n$$

makes sense if and only if B(0) = 0. To see this, if  $B(0) \neq 0$ , then any  $x^n$  term should consider all possible  $B(x)^i$ ,  $i \in \mathbb{N}$  since  $b_0$  generates  $x^n$  for each power. Otherwise, if B(0) = 0, then  $x^n$  term only care about  $B(x), \dots, B^n(x)$ .

Belows are useful Formal Power Series.

• 
$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

• 
$$\sum_{n=0}^{\infty} {m \choose n} x^n = (1+x)^m.$$

•  $\sum_{n=0}^{\infty} {m+n-1 \choose n} x^n = \frac{1}{(1-x)^m}, m \ge 2.$ 

• 
$$\sum_{n=0}^{\infty} {n \choose m} z^n = \frac{z^m}{(1-z)^{m+1}}$$

• 
$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$
.

•  $\sum_{n=0}^{\infty} \frac{x^n}{n} = -\log(1-x).$ 

*Proof.* First one is trivial. For second one, it is just from binomial theorem. If m < 0, then since  $(1 + z)^m$  is the inverse of  $(1 + z)^{-m}$ , we should show that  $\sum_{n\geq 0} {m \choose n} z^m$  is the inverse of  $\sum_{n\geq 0} {-m \choose n} z^n$ . However, by producting two sums, its coefficient is just from convolution that

$$\sum_{k=0}^{n} \binom{-m}{k} \binom{m}{n-k} = \binom{0}{n}$$

by Vandermonde identity, hence just 1. For third one, it is actually the same as bar-ball problem with m bars and n balls. To see the fourth one, it is index shift; namely,

$$\frac{z^m}{(1-z)^{m+1}} = z^m \sum_{n \ge 0} \binom{n+m}{m} z^n = \sum_{n \ge 0} \binom{n+m}{m} z^{n+m} = \sum_{n \ge 0} \binom{n}{m} z^n.$$

Fifth one is also trivial and the last one is from calculus.

For example,

Claim 2.2.5.

$$F(z) = \sum_{n \ge 0} \binom{2n}{n} z^n.$$

*Proof.* Let  $f_n = \binom{2n}{n}$ . Then,

$$f_n = \frac{2n(2n-1)}{n^2} \binom{2n-2}{n-1} = \frac{4n-2}{n} f_{n-1}.$$

Thus,

$$nf_n = 4nf_{n-1} - 2f_{n-1}.$$

Let  $F(z) = \sum f_n z^n$ . Then,

$$F'(z) = \left(\sum f_n z^n\right)' = \sum n f_n z^{n-1}.$$

Thus, by multiplying  $z^{n-1}$  for both sides, we get

$$nf_n z^{n-1} = 4nf_{n-1}z^{n-1} - 2f_{n-1}z^{n-1}$$

Now note that

$$(zF(z))' = (\sum f_n z^{n+1})' = \sum (n+1)f_n z^n = \sum nf_{n-1}z^{n-1}$$

Thus,

$$F'(z) = 4(zF(z))' - 2F(z) = 4zF'(z) + 2F(z)$$

By ODE calculation,

$$\frac{F'}{F} = \frac{2}{1 - 4z} \implies \log F = -\frac{1}{2}\log(1 - 4z) + C.$$

For c, we have initial value for z = 0, since F(0) = 1. Thus, C = 0. Therefore,

$$F(z) = \frac{1}{\sqrt{1-4z}} = \sum_{n\geq 0} {\binom{2n}{n}} z^n.$$

- 6		
		L
		L

Also check that

$$\frac{1}{\sqrt{1-4z}} = \sum_{n \ge 0} \binom{-\frac{1}{2}}{n} (-4z)^n.$$

To see this, note that

$$\binom{1/2}{n} = (-1)^n \frac{1 \cdot 3 \cdots (2n-1)}{n! 2^n}$$

Thus,

$$\binom{-\frac{1}{2}}{n}(-4z)^n = \frac{1 \cdot 3 \cdots (2n-1)2^n}{n!} = \binom{2n}{n}$$

since

$$\binom{2n}{n} = \frac{(2n(2n-2)\cdots 2)((2n-1)\cdots 3\cdot 1)}{n!n!} = \frac{1\cdot 3\cdots (2n-1)2^n}{n!}.$$

Corollary 2.2.6.

$$\sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k} = 4^{n}$$

*Proof.* Each coefficient of  $F^2$  is

$$F^{2} = \frac{1}{1 - 4z} = \sum_{n \ge 0} 4^{n} z^{n},$$

now apply convolution on  $F^2$ , done.

Theorem 2.2.7 (Binomial Inverse).

$$a_n = \sum_{k \ge 0} \binom{n}{k} b_k \iff b_n = \sum_{k \ge 0} (-1)^{n-k} \binom{n}{k} a_k.$$

*Proof.* Let  $\hat{A}, \hat{B}$  are exponential series such that

$$\hat{B} = \hat{A}e^z.$$

Then, from  $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$  and convolution,

$$\frac{b_n}{n!} = \sum_{k=0}^n \frac{1}{(n-k)!} \frac{a_k}{k!} \iff b_n = \sum_{k=0}^n \binom{n}{k} a_k.$$

Also, this holds if and only if  $\hat{B}e^{-z} = \hat{A}$ , and we know

$$\frac{a_n}{n!} = \sum_{k=0}^n \frac{(-1)^{n-k}}{(n-k)!} \frac{b_k}{k!} \iff a_n = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} b_k.$$

Also,

Proposition 2.2.8 (Inversion formula from Stirling).

$$a_n = \sum_k S_{n,k} b_k \iff b_n = \sum_k (-1)^{n-k} s_{n,k}$$

where  $S_{n,k}$  is the Stirling number of second kind (which count the number of ways to partition a set of n elements into k nonempty subsets) and  $s_{n,k}$  is the Stirling numbers of the first kind (which count the number of permutations of n elements with k disjoint cycles).

*Proof.* We already prove that

$$x^n = \sum_k S_{n,k} x^{\underline{k}}$$
 and  $x^{\underline{n}} = \sum_k (-1)^{n-k} s_{n,k} x^k$ .

And  $\{x^n\}_{n>0}, \{x^{\underline{k}}\}_{k\geq 0}$  are two bases of the polynomial. Hence, if we let  $A = (S_{n,k})_{nk}, B = (s_{n,k})_{nk}$ , then A, B are linear endomorphism of the vector space of polynomial which are inverse, i.e., AB = I. (Note that this infinite matrix and its product is well-defined since for each element of product, only finitely many nonzero terms occur.)  $\square$ 

Actually, this proof can be done for any two polynomial sequences which are bases of polynomial. Below are examples;

(1) Recall

$$x^{m} = \sum_{k \ge 0} k! S_{m,k} x^{\underline{k}} = \sum_{k \ge 0} k! S_{m,k} \binom{x}{k}.$$

Replace x by n,

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

Now let  $a_n = n^m, b_k = k! S_{m,k}$ .

You can make m = n since even if k is big enough, then  $S_{m,k}$  or  $\binom{n}{k}$  will be zero. So

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

Thus,

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

where last inequality comes from the formular.

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

(2) f(x) is polynomial of degree n.  $\{x^{\underline{k}}\}_{k=0}^{\infty}$  is also a basis.

$$f(x) = \sum_{k=0}^{n} a_k x^k$$

where  $a_n \neq 0$  since  $\deg(f) = n$ . Let x = n.

$$f(n) = \sum_{k=0}^{n} \binom{n}{k} k! a_k \iff n! a_n = \sum_{k \ge 0} (-1)^{n-k} \binom{n}{k} f(k).$$

by binomial Inversion.

If deg  $f < n \implies a_n = 0$ . So  $\sum_k (-1)^{n-k} {n \choose k} f(k) = 0$ . Similarly,  $f(x+1), f(x+2), \cdots$  all have degree < n. which implies

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

(To see this, just apply binomial inversion for  $f(x+n) = \sum_{k=0}^{n} {n \choose k} k! a'_k x^{\underline{k}}$ . Note that such  $a'_k$ s exist since  $\{x^{\underline{k}}\}$  is a basis.) Let  $\Delta$  difference operator

$$\Delta f(k) := f(x+1) - f(x) = (E - I)f(x).$$

where E(f(x)) = f(x+1). Note that E is linear; just check. Claim 2.2.9.  $\deg(f) < n \iff \Delta^n f = 0$ .

Proof.

$$\Delta^{n}(f) = (E - I)^{n}(f) = \sum_{k} (-1)^{n-k} \binom{n}{k} E^{k} f = \sum_{k} (-1)^{n-k} \binom{n}{k} f(x+k) = 0.$$

Thus  $\deg(f) < n \implies \Delta^n(f) = 0$  by above argument. Convesely, if  $\Delta^n(f) = 0$  for all x, then each f(x+n) has degree less than n, which implies f(x) has degree less than n, by inversion.

(3) Evaluation.

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

It is related with Vandermonde. However, in this time, we use

$$n!a_n = \sum_k (-1)^{n-k} \binom{n}{k} f(k).$$

Let

$$f(k) := \binom{n+k}{n} \implies f(x) := \binom{x+n}{n}.$$

Note that this also can be represented using basis  $\{x^{\underline{k}}\}$ , i.e.,

$$f(x) = \sum_{k \ge 0} a_k x^k$$

Then, by above formular,

$$\sum_{k} (-1)^{k} \binom{n}{k} f(k) = (-1)^{n} n! a_{n}.$$

Since  $f(x) = \sum_{k \ge 0} a_k x^{\underline{k}}$ , and  $\binom{x+n}{n}$  has leading term 1/n!,

$$a_n = \frac{1}{n!}$$

This implies

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

## 2.3 Probabilistic Generating Function

Let

- $\Omega$  be a probability space
- $X: \Omega \to \mathbb{N}$  be a random (discrete) variable.
- $p_n: P_r(X = n)$ , the probability that X takes the value n, where  $p_n \in [0, 1]/$ , i.e.,

$$p_n = Pr(X = n) = \frac{\# \text{ structure with } x = n}{-\text{total structure}}$$

**Definition 2.3.1.** the probability generating function is denoted by  $P_X(z)$ , and defined as

$$P_X(z) := \sum_{n \ge 0} p_n z^n,$$

where  $P_X(1) = 1$ . And the expected value of X is

$$\mathbb{E}X := \sum_{n \ge 0} nP_n = P'_X(1).$$

And the variance of X is

$$VarX = P_X''(1) + P_X'(1) - (P_X'(1))^2.$$

**Example 2.3.2.**  $\Omega = \mathfrak{S}n$ , let  $X = inv(\pi)$  for  $\pi \in S_n$ . Then,

$$P_X(z) = \sum n \ge 0 p_n z^n$$

and  $p_k = Pr(inv(\pi) = k) = \frac{I_{n,k}}{n!}$  where  $I_{n,k} = \#\{\pi \in \mathfrak{S}n : inv(\pi) = k\}$ . Thus,

$$P_X(z) = \sum n \ge 0 p_n z^n = \frac{1}{n!} \sum_{k \ge 0} I_{n,k} z^k = \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} z^{inv(\pi)} = \frac{1}{n!} \prod_{i=1}^n (1 + z + \dots + z^{i-1}).$$

whee last equality comes from Mahonian Statistics. (See inversion part.)

$$\mathbb{E}X = P'_X(1) = \left. \frac{1}{n!} \left( \prod_{i=1}^n (1+z+\dots+z^{i-1}) \right)' \right|_{z=1} = \sum_{i=1}^n \frac{i-1}{2} = \frac{1}{2} \binom{n}{2}.$$

Actually,

$$P'_X(z) = \sum_{i=1}^n \left( \prod_{\substack{j\neq i}}^n \frac{1+z+\dots+z^{j-1}}{j} \right) \left( \frac{1+2z+\dots+(i-1)z^{i-2}}{i} \right)$$
$$P''_X(z) = \sum_{i=1}^n \left( \prod_{\substack{j\neq i}}^n \frac{1+z+\dots+z^{j-1}}{j} \right)' \left( \frac{1+2z+\dots+(i-1)z^{i-2}}{i} \right)$$
$$+ \sum_{i=1}^n \left( \prod_{\substack{j\neq i}}^n \frac{1+z+\dots+z^{j-1}}{j} \right) \left( \frac{1+2z+\dots+(i-1)z^{i-2}}{i} \right)'$$

and by letting z = 1, we get  $\left(\prod_{j \neq i}^{n} \frac{1+z+\dots+z^{j-1}}{j}\right) = 1$ , and  $\left(\frac{1+2z+\dots+(i-1)z^{i-2}}{i}\right) = \frac{1}{i}\sum_{j=1}^{i-1} j = \frac{i-1}{2}$ , thus

$$P_X''(z)|_{z=1} = \sum_{i=1}^n \frac{i-1}{2} \left( \prod_{j \neq i}^n \frac{1+z+\dots+z^{j-1}}{j} \right)' + \sum_{i=1}^n \left( \frac{1+2z+\dots+(i-1)z^{i-2}}{i} \right)'$$
And

$$\left(\prod_{j\neq i}^{n} \frac{1+z+\dots+z^{j-1}}{j}\right)' = \sum_{j\neq i}^{n} \left(\prod_{k\neq i,j}^{n} \frac{1+z+\dots+z^{k-1}}{k}\right) \left(\frac{1+2z+\dots+(j-1)z^{j-2}}{j}\right)$$

by letting z = 1, we get

$$\left(\prod_{j\neq i}^{n} \frac{1+z+\dots+z^{j-1}}{j}\right)' \bigg|_{z=1} = \sum_{j\neq i}^{n} 1 \cdot \frac{j-1}{2} = \frac{1}{2} \binom{n}{2} - \frac{i-1}{2}.$$

Thus

$$P_X''(z)|_{z=1} = \sum_{i=1}^n \frac{i-1}{2} \cdot \left(\frac{1}{2}\binom{n}{2} - \frac{i-1}{2}\right) + \sum_{i=1}^n \left(\frac{1+2z+\dots+(i-1)z^{i-2}}{i}\right)'$$

and first term is just

$$\sum_{i=1}^{n} \frac{i-1}{2} \cdot \left(\frac{1}{2}\binom{n}{2} - \frac{i-1}{2}\right) = \frac{n(n-1)}{8} \sum_{i=1}^{n} (i-1) - \sum_{i=1}^{n} \frac{(i-1)^2}{4} = \frac{n^2(n-1)^2}{16} - \frac{n(2n^2 - 3n + 1)}{24}$$

hence

$$P_X''(z)|_{z=1} = \frac{n^2(n-1)^2}{16} - \frac{n(2n^2 - 3n + 1)}{24} + \sum_{i=1}^n \left(\frac{1 + 2z + \dots + (i-1)z^{i-2}}{i}\right)'$$

And

$$\begin{split} \sum_{i=1}^{n} \left( \frac{1+2z+\dots+(i-1)z^{i-2}}{i} \right)' \bigg|_{z=1} &= \sum_{i=1}^{n} \frac{1\cdot 2+2\cdot 3+\dots+(i-2)(i-1)}{i} = \sum_{i=1}^{n} \frac{1}{i} \sum_{j=1}^{i-2} j(j+1) \\ &= \sum_{i=1}^{n} \frac{1}{i} \left( \frac{(i-2)(i-1)(2i-3)}{6} + \frac{(i-2)(i-1)}{2} \right) \\ &= \sum_{i=1}^{n} \frac{(i-1)(i-2)}{2i} \left( \frac{(2i-3)}{3} + 1 \right) = \sum_{i=1}^{n} \frac{(i-1)(i-2)}{3} \\ &= \frac{1}{3} \left( \frac{n(n+1)(2n+1)}{6} - 3\frac{n(n+1)}{2} + 2n \right) \\ &= \frac{n}{9} (n^2 - 3n + 2). \end{split}$$

Thus,

$$P_X''(z)|_{z=1} = \frac{n^2(n-1)^2}{16} - \frac{n(2n^2 - 3n + 1)}{24} + \frac{n}{9}(n^2 - 3n + 2)$$
$$= \frac{n(9n^3 - 14n^2 - 21n + 26)}{144}$$

Thus variance is

$$VarX = P_X''(1) + P_X'(1) - (P_X'(1))^2 = \frac{(n-1)n(2n-5)}{72}.$$

**Remark 2.3.3.**  $\mathbb{E}X$  can be computed easier. Let  $x_{ij} = \begin{cases} 1 & \text{if } i, j \text{ form inversion} \\ 0 & o.w. \end{cases}$ , and

$$X = \sum_{i < j}^{n} x_{ij}.$$

For example, if n = 2, then,

$$\mathbb{E}X = \sum i < j\mathbb{E}(x_{ij}) = \mathbb{E}x_{12} = 1 \cdot Pr(21) + 0 \cdot Pr(12) = \frac{1}{2}$$

For general n,

$$\mathbb{E}X = \frac{1}{2} \binom{n}{2}$$

by choosing 2 numbers then there are two possible way of arranging and only one arrange gives inversion.

Now think about lattice walk. Suppose we walk on  $\mathbb{Z}$ . Let W denotes walk. Start at 0 at t = 0, and each time go left or right with evenly distributed probability, i.e.,

$$Pr(Lef) = Pr(Right) = 1/2.$$

Basic problem is

What is 
$$Pr(W \text{ reaches } m \text{ or } -m)$$

where  $m \in \mathbb{P} = \mathbb{N} \setminus \{0\}$ . Now make a condition on W hits  $\pm m$ , such as, How many steps (X) does it need? The answer is

$$Pr(W \text{ eaches } m \text{ or } -m.) = 1.$$

To see this we need some setting. Let  $G_n^{(m)}$  be the number of walks reach  $\pm m$  for the first time at time n. Then let

$$G^{(m)}(z) = \sum_{n \ge 0} G_n^{(m)} z^n.$$

Note that X is the number of steps used to hit  $\pm m$ , so

$$Pr(X=n) = \frac{G_n^{(m)}}{2^n}$$

since each walks in  $G^{(m)}$  has probability  $2^n$ . Thus,

$$P_X(z) = \sum_{n \ge 0} Pr(X=n)z^n = G^{(m)}(\frac{z}{2}).$$

Thus we get

1. 
$$P_X(1) = G^{(m)}(1/2)$$
  
2.  $\mathbb{E}X = P'_X(1) = \frac{1}{2}G^{(m)}(z/2)'|_{z=\frac{1}{2}} = \frac{1}{2}G^{(m)}(1/2)'.$ 

Hence it suffices to find  $G^{(m)}(z)$ . Let  $E_n^{(m)}$  be the number of positive walk from 0 to 0 in n step, i.e., return to the 0 for the first time at n step, not reaching m. So walk counted by  $E_n^{(m)}$  has always positive in the middle. Then, for any walk reaching m we can decompose it as last partial walk from 0 to m which contains all positive walk in the middle, and other walks from 0 to 0. Now let  $\tilde{E}_n^{(m)}$  be the number of positive walk from 0 to 0 in n steps not reaching m. Note that their difference is  $E_n^{(m)}$  cannot touch 0 in the middle but  $\tilde{E}_n^{(m)}$  does. Then,

$$\tilde{E}_{n}^{(m)} = \sum_{k \ge 0} E_{k}^{(m)} \tilde{E}_{n-k}^{(m)}$$

Thus, using convolution we can conclude that

$$\tilde{E}^{(m)}(z) = E^{(m)}(z)\tilde{E}^{(m)(z)} + 1 \implies \tilde{E}^{(m)}(z) = \frac{1}{1 - E^{(m)}(z)}.$$
(2)

with the initial condition  $E_0^{(m)} = 0$ ,  $\tilde{E}_0^{(m)}$  (Or you can think that actually, if we set whole structure  $A(z) := \tilde{E}_n^{(m)}$ , then it can be decomposed each walks following  $B(z) := E_n^{(m)}$ . Hence,

$$A(z) = 1 + B(z) + B(z)^{2} + \cdots$$

where  $B(z)^k$  stands for in case of partitioning the walk with k part following  $E_*^{(m)}$ , which implies A(z) is the composition of 1/(1-x) and B(z), so we get the above equality.

Also we claim  $E_n^{(m)} = \tilde{E}_{n-2}^{(m-1)}$ . Why? If we remove first two step of any walk counted by  $E_n^{(m)}$ , then it gives a walk counted by  $\tilde{E}_{n-2}^{(m-1)}$ . This gives

$$E^{(m)}(z) = z^2 \tilde{E}^{(m-1)} \tag{3}$$

Now by applying (2), we get

$$E^{(m)}(z) = z^2 \tilde{E}^{(m-1)} = \frac{z^2}{1 - E^{(m-1)}(z)}, E^{(1)}(z) = 0.$$

Now set  $A_m = E^{(m)}(1/2)$ . Then,  $A_m = \frac{1}{4(1-A_{m-1})}$ ,  $A_1 = 0$ , and induction on m yields

$$A_m = \frac{m-1}{2m}, (m \ge 1).$$

Thus, if we let  $B_m = E^{(m)}(1/2)'$ , then using above two equations, we get

$$B_m = \frac{2(m^2 - 1)}{3m}, (m \ge 1).$$

Now let  $F_n^{(m)}$  be the number of positive walk from 0, and hitting *m* for the first time at time *n*. Then,  $F_0^{(m)} = 0$ , and

$$F_n^{(m)} = \sum_{k=1}^n E_k^{(m)} F_{n-k}^{(m)} + F_{n-1}^{(m-1)}$$

where first term denotes the case when walk come back to 0 at k-step, and the last one denotes the case when a walk hits m without touching 0 in the middle of steps. Thus, from the convolution, we get

$$F^{(m)}(z) = E^{(m)}F^{(m)}(z) + zF^{(m-1)}.$$

Hence, by putting  $z = \frac{1}{2}$ , you can get

$$F^{(m)}(1/2) = \frac{1}{m+1}, F^{(m)}(1/2)' = \frac{2m(m+2)}{3(m+1)}.$$

#### **Remark 2.3.4.** Also, how can you get such an induction?

*Proof.* This is called Ricatti recurrence; I will give just way of general exact form. If the recurrence has a form

$$x_{n+1} = \frac{ax_n + b}{cx_n + d}$$

then let  $y_n = cx_n + d$ , so we can turn this equation into

$$y_{n+1} = \alpha + \frac{\beta}{y_n}$$

and by letting  $y_n := \frac{w_{n+1}}{w_n}$  we get

$$w_{n+2} - \alpha w_{n+1} + \beta w_n = 0$$

And we can solve this recurrence relation using characteristic equation, so done.

Thus, we can calculate  $G^{(m)}(1/2)$  using above information. Let  $G^{(m)}_{+,n}$  is the number of walks counted by  $G^{(m)}_n$  having positive first step. Then, these walks can be divided by two cases; One case is that it come back to origin. It can be counted by  $E^{(m)}(z)G^{(m)}(z)$ . The other case is not coming back to 0, in this case we can count it using  $zF^{(m-1)}(z)$ . Hence,

$$G_{+}^{(m)}(z) = E^{(m)}(z)G^{(m)}(z) + zF^{(m-1)}(z)$$

And by symmetry,

$$G^{(m)}(z) = 2E^{(m)}G^{(m)} + 2zF^{(m-1)}$$

Hence,

$$G^{(m)}(z) = \frac{2zF^{(m-1)}(z)}{1 - 2E^{(m)}(z)},$$

and  $G_0^{(m)} = 0$ . Thus,

 $G^{(m)}(1/2) = 1$ 

and

$$\mathbb{E}X = \frac{1}{2}G^{(m)}(1/2)' = m^2.$$

These are easy calculation if we get above inductions. The problem is how to get such formular for above.

# **3** Generating Function

Ordinary one is

$$F(x) = \sum_{n \ge 0} f_n x^n.$$

For example, let k be an integer  $\geq 1$ . Let  $h_n$  be the number of integer solution of  $x_1 + \cdots + x_k = n$ . Then, since it is just bar-ball problem, we know

$$h_n = \binom{n+k-1}{n} = \binom{k}{n}.$$

So, its generating function is

$$H_k(x) = \sum_{n \ge 0} \binom{n+k-1}{n} x^n = \frac{1}{(1-x)^k}$$

where last inequality comes from notes on useful identities. If we see the combinatorial meaning of  $H_k(x)$ , then

$$H_k(x) = \sum_{n \ge 0} h_n x^n = \sum_{n \ge 0} \left( \sum_{\substack{(x_1, \cdots, x_k) \\ \sum_i x_i \in \mathbb{N}}} 1 \right) x^n = \sum_{\substack{(x_1, \cdots, x_k) \\ x_i \in \mathbb{N}}} x^{x_1 + \dots + x_k} = \prod_{i=1}^k \left( \sum_{x_i \in \mathbb{N}} x^{x_i} \right) = \frac{1}{(1-x)^k}$$

Similarly, fix k. Then, let  $a_n$  be the number of integer solution of  $x_1 + \cdots + x_k = n$  where  $x_1 \in S_1, \cdots, x_k \in S_k$ and each  $S_i$  is a subset of  $\mathbb{N}$ . Then by the same argument above,

$$A_k(x) = \prod_{i=1}^k \left( \sum_{x_i \in S_i} x^{x_i} \right).$$

Let  $b_n$  be the number of non-negative integer solution of

,

$$3x_1 + 4x_2 + 2x_3 + 5x_4 = n.$$

Then let  $y_i = c_i x_i$  where  $c_i$  denotes coefficient of each term in the above equation, then it turns out to be counting integer solution of  $y_1 + \cdots + y_4 = n$  such that  $y_i \in c_i \mathbb{N}$ . Hence,

$$B(x) = \sum_{n \ge 0} b_n x^n = \frac{1}{1 - x^3} \cdot \frac{1}{1 - x^4} \cdot \frac{1}{1 - x^2} \cdot \frac{1}{1 - x^5}$$

Also, we can calculate the generating function with linear recurrence of degree k with constant coefficient. (Homogeneous case).

## **Example 3.0.1.** Consider $\{a_n\}_{n=0}^{\infty}$ satisfies

$$a_n = 5a_{n-1} - 6a_{n-2}$$

with  $a_0 = 1, a_2 = -2$ .

Usual way of solving recurrence relation is

- Step 1: Find the characteristic equation, in this case,  $q^2 = 5q 6$ .
- Step 2: Find the roots of the equation. In this case, q = 2, 3.
- Step 3: Find general solution of the Recurrence relation; if you study ordinary difference equation, then its solution is

$$a_n = c_1 2^n + c_2 3^n, \quad c_1, c_2 \in \mathbb{R}$$

• Step 4: Find  $c_1, c_2$  with  $a_0 = 1, a_1 = -2$ . Using this information, you can get

$$c_1 + c_2 = 1, 2c_1 + 3c_2 = -2 \implies c_1 = 5, c_2 = -4.$$

Hence,

$$a_n = 5 \cdot 2^n - 4 \cdot 3^n.$$

However, if we think about its generating function  $A(x) = \sum_{i \ge 0} a_i x^i$ ,

$$A(x) = a_0 + a_1 x + \sum_{n \ge 2} a_n x^n = 1 - 2x + \sum_{n \ge 2} a_n x^n$$

From the recurrence relation,

$$A(x) = 1 - 2x + \sum_{n \ge 0} 5a_{n-1} - 6a_{n-2})x^n$$
  
= 1 - 2x + 5x  $\left(\sum_{n \ge 1} a_n x^n\right) - 6x^2 \left(\sum_{n \ge 0} a_n x^n\right)$   
= 1 - 2x + 5x (A(x) - 1) - 6x^2 A(x)

implies

$$A(x) = \frac{1 - 7x}{1 - 5x + 6x^2} = \frac{5}{1 - 2x} - \frac{4}{1 - 3x} = 5\sum_{n \ge 0} 2^n x^n - 4\sum_{n \ge 0} 3^n x^n = \sum_{n \ge 0} (5 \cdot 2^n - 4 \cdot 3^b) x^n$$

which gives the same result.

In general, given recurrence relation

$$a_n = r_1 a_{n-1} + \dots + r_k a_{n-k}$$

with initial values, we get

$$A(x) = a_0 + a_1 x + \dots + a_{k-1} x^{k-1} + \sum_{n \ge k} a_n x^n$$

so by applying the recurrence relation on the sum, we can get

$$A(x) = \frac{P(x)}{Q(x)}$$

and using partial fraction, we can get the general solution of the recurrence relation.

Theorem 3.0.2 (Theorem 3.1 in [Aig07] p.95). The followings ar equivalent.

- 1.  $\{h_n\}_{n=0}^{\infty}$  a solution of a linear recurrence relation of degree k with constant coefficient.
- 2. H(x), the ordinary generating function of  $h_n$ , has a form  $H(x) = \frac{R(x)}{Q(x)}$  where R(x), Q(x) are polynomial and deg  $R < \deg Q = k$
- 3.

$$H(x) = \sum_{i=1}^{l} \frac{P_i(x)}{(1 - \alpha_i x)^{d_i}}$$

where  $\deg(P_i) < d_i$  and  $\sum_{i=1}^l d_i = k$ 

4.  $h_n = \sum_{i=1}^k P_i(n)\alpha_i^n$ , where  $P_i(n)$  is a polynomial of n with degree less than  $d_i$ .

*Proof.* From 1 to 2 is just showed over. From 2 to 3 is just partial fraction. From 3 to 4 is just calculating the fractions as formal power series, and from 4 to 1 is just definition of recurrence relation, done.  $\Box$ 

**Example 3.0.3.** Write  $(\sqrt{2} + \sqrt{3})^{1980}$  in decimal form. What is the last digit before and the 1st digit after decimal point?

The idea of solving this example is that we can see  $(\cdot)^n$  as a solution of the recurrence relation. Also, we can use  $(5+2\sqrt{6})^{990}$  instand of  $\sqrt{2}+\sqrt{3}$ . The reason of change is to get integer coefficient on the recurrence relation.

So consider

$$h_n := \left(5 + 2\sqrt{6}\right)^n + \left(5 - 2\sqrt{6}\right)^n$$

and let  $\alpha_1 = 5 + 2\sqrt{6}$ ,  $\alpha_2 = 5 - 2\sqrt{6}$ . This should be a solution of some quadratic equation, and we can get the quadratic equation using Vieta's formula. (Or in usual word, using the relationship between coefficients of a polynomial and zeros of the polynomial.)

$$\alpha_1 + \alpha_2 = 10, \alpha_1 \alpha_2 = 1 \implies q^2 - 10q + 1 = 0$$

So

$$h_n = 10h_{n-1} - h_{n-2}.$$

Thus, by letting n = 0 and 1 on the equation  $h_n := (5 + 2\sqrt{6})^n + (5 - 2\sqrt{6})^n$ , we get

$$h_0 = 2, h_1 = 10.$$

Note that

$$0 < \alpha_2 = 5 - 2\sqrt{6} = \sqrt{25} - \sqrt{24} < \frac{1}{2}$$

Thus,  $\alpha_2^n \to 0$  as  $n \to \infty$ . Therefore,  $\alpha_1^n = h_n - (\text{tiny number})$ , thus, since *n* is integer, then  $\alpha_1 = *.9 * * * \cdots$  for sufficiently large *n*. Thus if we know  $h_n$ 's last digit, then since  $h_n$  is integer, we are done. So let *y* be the last digit of  $h_{990} - 1$ . Then yb mod 10,

$$h_n \cong -h_{n-2} \mod 10$$

it gives sequence  $2, 8, 2, 8, \cdots$  thus

$$h_{990} \cong 8 \mod 10$$

which gives the last digit y = 7.

**Example 3.0.4** (Catalan Number). Let  $C_0 = 1, c_{n+1} = C_0C_n + C_1C_{n-1} + \dots + C_nC_0 = \sum_{i=0}^n C_iC_{n-i}$ . So,

 $1, 1, 2, 5, 14, 42, 132, 429, \cdots$ 

we call this sequence as Catalan Number.

Its combinatorial meaning can be found on the triangulariton of (n + 2) polygon. For example, if n = 3, think about the 5-gon, say P with vertices  $\{0, 1, 2, 3, 4\}$ . Then it has five triangulation, where

$$P \cup \{02, 03\}, P \cup \{13, 14\}, P \cup \{02, 04\}, P \cup \{13, 03\}, P \cup \{14, 24\}$$

Thus it has  $C_3 = 5$  triangulation. To see this for *n*, thinking about the triangulation of (n + 2)-gon with triangle 0, k + 1, n + 2.



Then the left remaining part is (k + 2)-gon and the right remaining part is (n - k + 2)-gon (note that the number of elements from 0 to k + 1 is k + 2, and that of from k + 1(-k + 1) to n + 2 - (k + 1) is n - k + 2.), thus, the triangulation containing triangle of 0, k + 1, n + 2 is  $C_k \cdot C_{n-k}$ . And k can permute  $\{1, 2, \dots, n\}$ . Thus,

$$C_0C_n + \dots + C_nC_0 = C_n$$

triangulations are possible.

Let C(z) be the ordinary generating function of  $C_k$ . Then,

$$C(z) = \sum_{k=0}^{\infty} C_k z^k = 1 + \sum_{k=1}^{\infty} C_k z^k = 1 + \sum_{k\geq 0}^{k-1} \sum_{i=0}^{k-1} C_i C_{k-i-1} z^k = 1 + z \sum_{k\geq 0}^{k} \sum_{i=0}^{k} C_i C_{k-i} z^k = 1 + z C^2(z).$$

where the last equality comes from the convolution, C(z)C(z). Hence, C(z) is a solution of  $y = 1 + zy^2$ . Since

$$y = \frac{1 \pm \sqrt{1 - 4z}}{2z}$$

and since C(0) = 1, we should choose  $\lim_{z\to 0} y = 1$ , which is

$$C(z) = y = \frac{1 - \sqrt{1 - 4z}}{2z}.$$

Numerically,

$$(1-4z)^{1/2} = \sum_{n\geq 0} \binom{1/2}{n} (-4z)^n = 1 + \sum_{n\geq 1} \frac{1}{2n} \binom{-\frac{1}{2}}{n-1} (-4)^n z^n = 1 - 2\sum_{n\geq 1} \frac{1}{n} \binom{2n-2}{n-1} z^n = 1 - 2\sum_{n\geq 0} \frac{1}{n+1} \binom{2n}{n} z^{n+1} = 1 - 2\sum_{n\geq 0} \frac{1}{n-1} \binom{2n}{n-1} z^n = 1 - 2\sum_{n\geq 0} \frac{1}{n-1} \binom{2n}{n-1} \binom$$

Thus,

$$1 - (1 - 4z)^{1/2} = 2z \sum_{n \ge 0} \frac{1}{n+1} \binom{2n}{n} z^n.$$

Hence,

$$C(z) = \frac{1 - (1 - 4z)^{1/2}}{2z} = \sum_{n \ge 0} \frac{1}{n+1} \binom{2n}{n} z^n$$

which implies

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

## 3.1 Evaluating sums

1. Let  $A(x) = \sum_{k \ge 0} a_k x^k$ . Let

$$s_n = \sum_{k=0}^n a_k$$

Then, for  $b_i = 1$  for all i,

$$S(x) = \sum_{n \ge 0} \left( \sum_{k=0}^{n} a_k b_{n-k} \right) x^n = A(x)B(x) = \frac{A(x)}{1-x}$$

since  $B(x) = \sum_{n \ge 0} x^n = \frac{1}{1-x}$ .

2. Let  $S_n = \sum_{k=0}^n \binom{2k}{k} 4^{-k}$ , Then,

$$4^n S_n = \sum_{k=0}^n \binom{2k}{k} 4^{n-k}$$

so using convolution,

$$\sum_{n\geq 0} 4^n S_n x^n = \left(\sum_{k\geq 0} \binom{2k}{k} x^k\right) \left(\sum_{k\geq 0} 4^k x^k\right) = (1-4x)^{-1/2} \frac{1}{1-4x} = (1-4x)^{\frac{3}{2}}$$

where the second equality comes from the note of Useful Identities. Thus,

$$4^{n}s_{n} = \binom{-3/2}{n}(-4)^{n} = (2n+1)\binom{2n}{n}.$$

3. Let  $s_n = \sum_{k=1}^n (-1)^{k-1} k \binom{n}{k} r^{n-k}$ . By letting  $a_k = (-1)^{k-1} k$ ,  $b_{n-k} = r^{n-k}$ ,

$$s_n = \sum_{k=0}^n \binom{n}{k} a_k b_{n-k}$$

since  $k = 0 \implies a_0 = 0$ . Thus,

$$\sum_{n\geq 0} s_n \frac{x^n}{n!} = \left(\sum_{k\geq 0} a_k \frac{x^k}{k!}\right) \left(\sum_{k\geq 0} b_k \frac{x^k}{k!}\right)$$
$$= \left(\sum_{k\geq 0} (-1)^{k-1} k \frac{x^k}{k!}\right) \left(\sum_{k\geq 0} r^k \frac{x^k}{k!}\right)$$
$$= x e^{-x} \cdot e^{rx} = x e^{(r-1)x}.$$

Thus,  $s_n = n(r-1)^{n-1}$ .

4. "External approach." Let  $s_n = \sum_{k=0}^n$  something. Then,

$$\sum_{n \ge 0} s_n x^n = \sum_{n \ge 0} x^n \sum_{k \ge 0} \text{something} = \sum_{k \ge 0} \sum_{n \ge 0} \text{something} \cdot x^n.$$

5.  $s_n = \sum_{k=\lfloor \frac{n}{2} \rfloor}^n {k \choose n-k}$ . Then the binomial coefficient is nonzero iff k > n-k and  $2k \ge n$ . Thus  $s_n = \sum_{k\ge 0} {k \choose n-k}$ . Then,

$$F(x) = \sum_{n \ge 0} s_n x^n = \sum_{n \ge 0} \sum_{k \ge 0} \binom{k}{n-k} x^n = \sum_{k \ge 0} x^k \sum_{n \ge 0} \binom{k}{n-k} x^{n-k} = \sum_{k \ge 0} x^k \sum_{j \ge 0} \binom{k}{j} x^j$$

since j = n - k. But this j visits  $0, 1, \dots, k$ , we can apply the binomial theorem; so

$$F(x) = \sum_{k \ge 0} x^k (1+x)^k = \frac{1}{1 - x(1+x)}$$

Note that  $1 - x - x^2$  is the characteristic equation of  $s_{n+1} = s_n + s_{n-1}$ . Thus, from  $s_0 = 1, s_1 = 1$ , we can conclude that

$$s_n = F_{n-1}$$

where  $F_n$  is the Fibonacci sequence.

6. Some tricks using convolution. Let

$$s_n = \sum_{k \ge 0} \binom{n+k}{m+2k} \binom{2k}{k} \frac{(-1)^k}{(k+1)}.$$

Let

$$S(x) = \sum_{n \ge 0} s_n x^n = \sum_{n \ge 0} x^n \sum_{k \ge 0} \binom{n+k}{m+2k} \binom{2k}{k} \frac{(-1)^k}{(k+1)}$$
$$= \sum_{k \ge 0} \binom{2k}{k} \frac{(-1)^k}{(k+1)} \sum_{n \ge 0} \binom{n+k}{m+2k} x^n$$

Note that

$$\sum_{n\geq 0} \binom{n+k}{m+2k} x^n = x^{-k} \sum_{n\geq 0} \binom{n+k}{A} x^{n+k},$$

where A = m + 2k. Thus by the useful identity 5,

$$\sum_{n \ge 0} \binom{n+k}{m+2k} x^n = x^{-k} \sum_{n \ge 0} \binom{n+k}{A} x^{n+k} = x^{-k} \frac{x^A}{(1-x)^{A+1}}.$$

Thus,

$$S(x) = \sum_{k \ge 0} \binom{2k}{k} \frac{(-1)^k}{(k+1)} \sum_{n \ge 0} \binom{n+k}{m+2k} x^n$$
$$= \sum_{k \ge 0} \binom{2k}{k} \frac{(-1)^k}{k+1} x^{-k} \frac{x^{m+2k}}{(1-x)^{m+2k+1}}$$
$$= \frac{x^m}{(1-x)^{m+1}} \sum_{k \ge 0} (-1)^k \frac{1}{k+1} \binom{2k}{k} \frac{x^k}{(1-x)^{2k}}$$
$$= \frac{x^m}{(1-x)^{m+1}} \sum_{k \ge 0} \frac{1}{k+1} \binom{2k}{k} \left(-\frac{x}{(1-x)^2}\right)^k$$

And we know that

$$\sum_{k \ge 0} \frac{1}{k+1} \binom{2k}{k} z^k = C(z)$$

the generating function of Catalan Series. Hence, from  $C(z)=\frac{1-\sqrt{1-4z}}{2z},$  we can get

$$C(-\frac{x}{(1-x)^2}) = \frac{1 - \frac{1+x}{1-x}}{\frac{-2x}{(1-x)^2}} = (1-x),$$

thus

$$S(x) = \frac{x^m}{(1-x)^{m+1}} C\left(-\frac{x}{(1-x)^2}\right) = \left(\frac{x}{1-x}\right)^m$$

Thus,

$$s_n = [x^n]S(x) = [x^{n-m}]\frac{1}{(1-x)^m} = \binom{n-1}{m-1}$$

from the useful identity 3.

7. "Exponential Convolution" Note that

$$h_n = \sum_{k=0}^n \binom{n}{k} f_k g_{n-k} \iff \sum_{n \ge 0} h_n \frac{x^n}{n!} = \left(\sum_{n \ge 0} f_n \frac{x^n}{n!}\right) \left(\sum_{n \ge 0} g_n \frac{x^n}{n!}\right)$$

For example, let  $\sigma \in \mathfrak{S}n$ , then derangement of  $\sigma = \sigma_1 \cdots \sigma_n$  is element in  $\sigma$  such that  $\sigma_i \neq i, \forall i$ . So any permutation is disjoint union of fixed part and derangement part. For example,

 $\sigma = 3214763,$ 

derangement part is 32175 (position: 12357) and fixed part is 46. So, if we let  $d_n$  be the number of derangement in permutation, then

$$\sum_{k=0}^{n} \binom{n}{k} d_k = n!$$

Thus, by convolution with  $b_{n-k} = 1$ ,  $a_k = d_k$  we can get

$$E_a(x)E_b(x) = \sum_{n\geq 0} n! \frac{x^n}{n!}$$

and  $E_b(x) = e^x$ . Hence,

$$E_a(x) = \frac{e^{-x}}{1-x}.$$

where  $E_a(x)$  denotes the exponential generating function of  $a_n$ .

In general, if h-structure on set is the disjoint union of a set with f-structure and the other with g-structure, then

$$E_h(x) = E_f(x)E_g(x).$$

If g = f, for example, f = g is the number of complete graph, i.e.,  $f_0 = 0, f_n = 1$ ? In this case, we should remove repeated counting case, where f(S)g(T) and f(T)g(S) for any partition  $\{S, T\}$  of [n]

$$h_n = \sum_{S \sqcup T = [n]} f(S)f(T) = \frac{1}{2} \sum_{S \sqcup T = [n]} f(S)f(T) = \frac{1}{2} (E_f)^2.$$

Thus, if h is disjoint k-union with the same f-structure, then

$$E_h(x) = \frac{1}{k!} (E_f)^k.$$

For example,  $S_{n,k}$  is the number of partitions of [n] into k blocks. So,

$$\sum_{n \ge 0} S_{n,k} \frac{x^n}{n!} = \frac{1}{k!} \left( -1 + \sum_{n \ge 0} \frac{x^n}{n!} \right)^k = \frac{(e^x - 1)^k}{k!}$$

since underlying f structure is just nonempty set, i.e.  $f_n = 1$  if n > 0,  $f_0 = 0$ .

**Theorem 3.1.1** (Compositional Formula). Let  $h_n$  counts the partition of [n] into kblocks, and put f-structure on each blocks, and put a g-structure on blocks themselves. Then,

$$h_n = \sum_{k \ge 1} g_k \sum_{\{B_1, \cdots, B_k\} \pi P(n,k)} f(|B_1|) \cdots f(|B_k|)$$

Hence, if  $h_0 = 1, f_0 = 0$ ,

$$E_h = E_g(E_f(x)).$$

*Proof.* For fixed k, we have  $g_k \cdot \frac{1}{k!} (E_f)^k$  structure. So,

$$\sum_{n\geq 0} h_n \frac{x^n}{n!} = \sum_{n\geq 0} \sum_{k\geq 1} g_k \sum_{\{B_1,\cdots,B_k\} \pi P(n,k)} f(|B_1|) \cdots f(|B_k|) x^n = \sum_{n\geq 0} \sum_{k\geq 1} g_k \frac{E_f^k}{k!} = E_g(E_f(x))$$

where second equality comes from the trick that distributing  $x^n$  to each  $x^{|B_1|}, \dots, x^{|B_n|}$  and use the above formular on k.

There are some special cases.

• If  $g_k = 1$  for all k, then it is just exponential composition formular, i.e.,

$$E_h = e^{E_f}$$

• if  $g_k = t^k$ , then

$$h_n = \sum_{\pi \in P(n)} t^{|\#blocks|} \prod_{i=1}^{|\#blocks|} f(|B_i|) = \sum_{h\text{-structure on [n]}} t^{\#\text{components}} f(|B_i|) = \sum_{h\text{-structure on$$

Thus,

$$E_h(x) = \sum_{h\text{-structure on [n]}} t^{\text{#components}} x^{\text{#elements}} / (\text{#elements})!$$

and the composition formular gives

$$E_h(x) = e^{tE_f(x)}$$

since  $e^{tx} = \sum_{k \ge 0} (tx)^k$ . From this, we can get

 $\mathbf{p}$ 

$$\sum_{\substack{n \ge 0, k \ge 0\\ \text{atition on } n \text{ with } k \text{ blocks}}} \frac{x^n t^k}{n!} = e^{t(e^x - 1)}.$$

using the composition formular, with  $E_f(x)$  is just (nonempty) structure for each k block. Similarly,

$$\sum_{n\geq 0} B(n)\frac{x^n}{n!} = e^{e^x - 1}$$

where B(n) is the Bell number, the number of partition of n with any disjoint sets. This also can be derived from giving h structure, i.e., disjoint k sets and giving f-structure, just (nonempty).

• Let  $h_n$  be the number of decomposition of [n] into nonempty blocks then linearly order each block. Then the *f*-structure is  $f_n = n!$  or  $f_0 = 0$ .  $g_n = 1, g_0 = 0$ , since the block has only (nonempty) structure. Thus,

$$E_h(x) = e^{\sum_{n \ge 1} n! \frac{x^n}{n!}} = e^{\frac{x}{1-x}}.$$

• Let  $h_n$  be the number of graphs on [n], i.e.,  $e^{\binom{n}{2}}$ . Then, let  $f_n$  is the number of connected graph on [n]. Then, by setting  $g_n = 1, g_0 = 0$  (since a graph consisting of forests of connected graphs, so no additional structure on connected components themselves.) Then,

$$E_h(x) = e^{E_f(x)}$$

Hence,

$$E_f(x) = \log E_h(x) = \log \left(1 + \sum_{n \ge 1} 2^{\binom{n}{2}} \frac{x^n}{n!}\right)$$

• Let  $h_n$  be the number of decomposition of [n] into nonempty blocks then cyclic order each block. Then the *f*-structure is  $f_n = (n-1)!$  or  $f_0 = 0$ .  $g_n = 1, g_0 = 0$ , since the block has only (nonempty) structure. Thus,

$$E_h(x) = e^{\sum_{n \ge 1} (n-1)! \frac{x^n}{n!}} = e^{\sum_{n \ge 1} \frac{x^n}{n}} = e^{-\log(1-x)} = \frac{1}{1-x}$$

where the last equality comes from the useful identity 9.

**Theorem 3.1.2** (Theorem 3.5 in [Aig07]: A permutation version of the compositional formula). Let f, g, h be defined by

$$h(|X|) = \sum_{\sigma \in S_{|X|}} g(k)f(|C_1|) \cdots f(|C_n|)$$

with  $h_0 = g(0)$ , where  $C_1, \dots, C_n$  are cycle of  $\sigma$ . Then,  $E_h(x) = E_g\left(\sum_{n\geq 1} f(n) \frac{x^n}{n}\right)$ .

*Proof.* There are (j-1)! ways of to make *j*-set into a cyclic permutation. Hence,

$$h(|X|) = \sum_{\{B_1, \dots, B_k\} \in \prod(X)} g(k)(|B_1| - 1)! f(|B_1|) \cdots (|B_n| - 1)! f(|B_n|)$$

thus by the compositional formula with new  $f'_n = (n-1)!f_n$ , we get

$$E_{h}(x) = E_{g}(\sum_{n \ge 1} (n-1)! f_{n} \frac{x^{n}}{n!}) = E_{g}\left(\sum_{n \ge 1} f_{n} \frac{x^{n}}{n}\right)$$

For example, let  $g_k = 1$  for all k, and  $f_n = 1$  only for n = 1, 2, then we are counting permutations with no cycles of length greater than or equal to 3, in other words involutions, i.e.,  $\sigma^2 = e$ . Then,

$$E_h(x) = e^{x + x^2/2}.$$

**Proposition 3.1.3.** Let S be a finite set. Given  $f, g : \mathbb{N} \to K$ , define  $h_1, \dots, h_4$  as follow;

1.  $h_1(\#S) = f(\#S) + g(\#S)$ 2.  $h_2(\#S) = \#S \cdot f(\#S - 1)$ 3.  $h_3(\#S) = f(\#S + 1)$ 4.  $h_4(\#S) = \#Sf(\#S)$ 

Then,

1.  $E_{h_1}(x) = E_f + E_g$ 2.  $E_{h_2}(x) = xE_f$ 3.  $E_{h_3}(x) = E'_f$  4.  $E_{h_4}(x) = x E'_f$ 

Proof. First one is trivial; it comes from the linearlity of the formal sum. For the second,

$$\sum_{n \ge 0} h_n \frac{x^n}{n!} = \sum_{n \ge 1} h_n \frac{x^n}{n!} = \sum_{n \ge 1} f_{n-1} \frac{x^n}{(n-1)!} = x \sum_{n \ge 0} f_n \frac{x^n}{n!} = x E_f(x).$$

where the fist equality comes from  $h_2(0) = 0$ . Fo the third one,

$$\sum_{n\geq 0} h_n \frac{x^n}{n!} = \sum_{n\geq 0} f_{n+1} \frac{x^n}{n!} = \sum_{n\geq 0} (f_{n+1} \frac{x^{n+1}}{(n+1)!})' = E_f(x)'$$

where the second equality comes from derivative of constant is zero. For the fourth one,

$$\sum_{n \ge 0} h_n \frac{x^n}{n!} = \sum_{n \ge 1} f_n \frac{x^n}{(n-1)!} = x \sum_{n \ge 1} f_n \frac{x^{n-1}}{(n-1)!} = x E_f(x)$$

where the last equality comes from (3).

### 3.2 Enumeration on trees

Tree has the labeled vertices V = [n], and a graph G = (V, E), where  $E \subseteq {\binom{[n]}{2}}$ . Then **tree** is a connected graph without cycle. For example,



above trees are all distinct tree. Let  $T_n = \#$  trees on [n]. Note that  $T_3 = 3$ , as seen above. Let  $r_n = \#$  rooted trees on [n]. A **rooted tree** is a tree with a marked vertex, called root. For example,



those are example of rooted tree, where the root is black dots. Thus, for each tree, we can get three distinct rooted tree. In general, for any tree with vertices n, we can get n distinct rooted tree by permuting root. Hence,

$$r_n = nT_n.$$

Let  $f_n = \#$  rooted forests on [n]. A **rooted forest** is a graph without cycle and every connected component is rooted tree. We claim that

$$f_n = T_{n+1}.$$

To see this, note that for any forests with rooted tree, we can make tree with n + 1 vertices by adding 0 vertex and edges between root to 0. Also, we can get the rooted forest by deleting n + 1 labeled vertex. (Note that labeling gives distinct tree even if their shape is the same, so this procedure is inverse to each other.) Thus,

$$r_{n+1} = (n+1)T_{n+1} = (n+1)f_n$$

Let y be the exponential generating function for  $r_n$ , i.e.,

$$y = \sum_{n \ge 1} r_n \frac{x^n}{n!}.$$

and  $F(z) = \sum_{n \ge 0} f_n \frac{z^n}{n!}$ . Then,

$$zF(z) = z \sum_{n \ge 0} f_n \frac{z^n}{n!} = z \sum_{n \ge 0} r_{n+1} \frac{z^n}{(n+1)!} = \sum_{n \ge 0} r_{n+1} \frac{z^{n+1}}{(n+1)!} = y.$$

Also, the number of rooted forests is just having disjoint unions and gives a rooted tree structures, hence by the Exponential formular, (since blocks have no structure)

$$F(z) = e^y.$$

Thus,

$$y = ze^y$$
.

Let  $P_k(n) = \#$  rooted forests on [n] with exactly k trees, i.e.,

$$f_n = \sum_{k \ge 0}^n P_k(n).$$

Then, by the exponential formula,

$$\sum_{n\geq 0} P_k(n) \frac{x^n}{n!} = \frac{1}{k!} y^k$$

since it is just make partition of n into k blocks and gives rooted tree structure. Now fix a subset  $S \subseteq [n]$ , and let |S| =: k. Then,

 $P_S(n) = \#$  of rooted forest on [n] whose roots are exactly S.

So,

$$P_{S_1}(X) = P_{S_2}(X) \iff |S_1| = |S_2|$$

Thus,

$$P_k(n) = \binom{n}{k} P_S(n)$$

for  $|S| = k, S \subseteq [n]$ .

Claim 3.2.1. 
$$P_S(n) = kn^{n-k-1}$$

*Proof.* If |S| = n, then  $P_n(n) = 1 = n \cdot n^{n-n-1}$ , so the equation holds. If |S| = 1, we know that  $T_n = n^{n-2}$ ,  $r_n = n^{n-1}$  using Prüfer code. Prüfer code is a bijection between rooted forests and  $[n]^{n-k-1} \times S$ .  $\Box$ 

Prüfer code is an algorithm such that it sequentially remove the largest leave and write down its neighbor. Then, for n vertices, with k trees, it remains n - k vertices, hence it write down neighbors n - k time. And note that the last neighbor which the algorithm scribes is always in S So the possible number of Prüfer code is  $kn^{n-k-1}$ , where k is the cardinality of S.

*Bijectivity of the Prüfer code.* The map from Prüfer code to rooted forest is injective, since if there is two Prüfer code in the preimage, then the same rooted forests has the distinct neighborhood of leaves structure, contradiction. Also, the other map is unique, since leaves-neighborhood structure of labeled tree is unique.

**Corollary 3.2.2.** 
$$P_k(n) = \binom{n}{k} k n^{n-k-1} = \binom{n-1}{k-1} n^{n-k}$$

Hence,

$$T_n = P_{\{1\}}(n) = n^{n-2}, r_n = nT_n = n^{n-1}, f_n = T_{n+1} = (n+1)^{n-1}.$$

In a rooted forest, for any vertex v,

$$\deg(v) = \# \text{ children of } v = \begin{cases} \# \text{ edge incident to } v & \text{if } v \text{ is a root} \\ \# \text{ edge incident to } v - 1 & \text{if } v \text{ is not a root} \end{cases}$$

*leaf* is a vertex which is not a root and degree 0. So, for each rooted forests, we can represent it as *ordered degree sequence* defined as below.

$$\Delta(F) := (\delta_1, \cdots, \delta_n)$$
 where  $\delta_i = \deg(i), \sum_i \delta_i = n - k$ .

**Theorem 3.2.3.** For  $\Delta = (\delta_1, \dots, \delta_n)$  where  $\delta_i \ge 0, \sum_i \delta_i = n - k$ , let  $N(\Delta)$  be the number of rooted forests on [n] whose ordered degree sequence is  $\Delta$ . Then,

$$N(\Delta) = \binom{n-1}{k-1} \binom{n-k}{\delta_1, \cdots, \delta_n}.$$

Equivalently,

$$\sum_{F \text{ on } [n]} x_1^{\deg(1)} \cdots x_n^{\deg(n)} = \binom{n-1}{k-1} (x_1 + \dots + x_n)^{n-k}$$

where F is the rooted forests on [n].

*Proof.* LHS is

$$\sum_{\substack{(a_1,\cdots,a_{n-k})'s\\\text{Prifer code}\\=|n|^{n-k-1}\times S}} x_{a_1}\cdots x_{a_{n-k}} = \left(\sum_{a_1\in[n]} x_{a_1}\right)\cdots \left(\sum_{a_{n-1}\in[n]} x_{a_{n-1}}\right) \left(\sum_{a_n\in S} x_{a_n}\right) = (x_1+\cdots+x_n)^{n-k-1} \left(\sum_{i\in S} x_i\right)$$

Summing over all  $S \in {\binom{[n]}{k}}$ , we get

$$\sum_{\substack{F \text{ on } [n]\\F \text{ has } k \text{-components}}} x_1^{\deg(1)} \cdots x_n^{\deg(n)} = (x_1 + \cdots + x_n)^{n-k-1} \left( \sum_{|S|=k} \sum_{i \in S} x_i \right).$$

where deg(l) is the number of children of vertex  $l \in [n]$ . Then, note that  $\left(\sum_{|S|=k}\sum_{i\in S} x_i\right) = \binom{n-1}{k-1}(x_1 + \cdots x_n)$ . To see this, note that the way of choosing S with |S| = k with  $x_i$  is  $\binom{n-1}{k-1}$ , since choose  $x_i$  first and choose the rest k-1 elements from n-1 elements. Since i was arbitrary, it gives the desired result. Hence,

$$\sum_{\substack{F \text{ on } [n]\\F \text{ has } k \text{-components}}} x_1^{\deg(1)} \cdots x_n^{\deg(n)} = (x_1 + \cdots + x_n)^{n-k} \binom{n-1}{k-1}.$$

Now, given a rooted forest F, let

$$type(F) := (r_0, \cdots)$$

where  $r_i$  is the number of vertices of degree  $i, \sum r_i = n$ . For example, the forest F below

has type (2, 0, 1), and  $\Delta(F) = (0, 2, 0)$ .

**Claim 3.2.4.** The number of forests with a given type  $(r_0, r_1, \cdots)$  is

$$\binom{n}{r_0, \cdots, r_n} \binom{n-1}{k-1} \binom{n-k}{\delta_1, \cdots, \delta_n}$$

where  $\delta_1, \dots, \delta_n$  is any arrangement of  $r_0$  many 0 on ordered sequence  $\Delta$ ,  $r_1$  many 1 on  $\Delta$ , etc. And this is equivalent to

$$\binom{n}{r_0, \cdots, r_n} \binom{n-1}{k-1} \frac{n-k}{(0!)^{r_0} (1!)^{r_1} (2!)^{r_2} \cdots}$$

where k is given by  $\sum_{i=0}^{n} ir_i = n - k$ .

_	
	_

*Proof.* Note that any arrangement of given type gives the same multinomials; i.e., if type is given by (2, 0, 1), then there are three possible arrangement, (2, 0, 0), (0, 2, 0), (0, 0, 2) and all gives the same multinomial  $\binom{2}{(0!)^2(1!)^0(2!)^1}$ . Hence, done.

So, when k = 1, # of rooted trees of type  $\bar{r} = (r_0, \cdots, r_n)$  is

$$\binom{n}{r_0,\cdots,r_n}\frac{(n-1)!}{(0!)^{r_0}\cdots(n!)^{r_n}}$$

## 3.3 Lagrange Inversion Formula

Laurent Powe series

$$f(x) = \sum_{i \ge n_0} f_i x^i$$

**Lemma 3.3.1.** 1.  $[x^{-1}]f'(x) = 0$ 

2. If f(x) is a formal power series with  $f_0 = 0, f_1 \neq 0$ , then

$$[x^{-1}]f(x)^{i}f'(x) = \begin{cases} 1 & \text{if } i = -1\\ 0 & o.w. \end{cases}$$

*Proof.* First one is obvious since  $\log x$  term is not included in f(x). To see the second one, note that

$$f^{i}f'(x) = \frac{1}{i+1}(f^{i+1}(x))'$$

Thus, if  $i \neq -1$ , then it is just another power series, so by the first result, it gives 0. If i = -1, then from the assumption that  $f_0 = 0$ ,  $f_1 \neq 0$ , we can say f(x) = xg(x) for some g(x). This implies

$$\frac{1}{f(x)} = \frac{1}{x} \cdot \frac{1}{g_0 + g_1 x + \dots} = \frac{1}{x} \cdot \frac{1}{g_0} \left( \frac{1}{1 + \frac{g_1}{g_0} x + \dots} \right) = \frac{1}{g_0 x} \left( 1 - p + p^2 - p^3 + \dots \right)$$

where  $p = \frac{g_1}{g_0}x + \cdots$ , and use the geometric series for the last equality. Thus,

$$\frac{f'(x)}{f(x)} = \frac{f_1 + 2f_2x + 3f_3x^3 + \dots}{f_1x + f_2x^2 + f_3x^3 + \dots} = \frac{\left(f_1 + 2f_2x + 3f_3x^3 + \dots\right)}{f_1x} \cdot \frac{1}{1 + \frac{f_2}{f_1}x + \frac{f_3}{f_1}x^2 + \dots} = \frac{\left(f_1 + 2f_2x + 3f_3x^3 + \dots\right)}{f_1x} \frac{1}{1 + p}$$
$$= \frac{1 + 2\frac{f_2}{f_1}x + 3\frac{f_3}{f_1}x^2 + \dots}{x} \left(1 - p + p^2 - p^3 + \dots\right)$$

So  $[x^{-1}]\frac{f'(x)}{f(x)} = 1$ , since other terms containing at least x except 1 in numerator.

**Theorem 3.3.2** (Lagrange Inversion Formula). Let f(x) be a formal power series, with  $f_0 = 0, f_1 \neq 0$ . Let g(x) be the compositional inverse of f(x), i.e.,

$$g(f(x)) = f(g(x)) = x.$$

Then,

$$[x^{n}]g(x) = \frac{1}{n} \cdot [x^{-1}](f(x))^{-n}.$$

*Proof.* Let  $g(x) = \sum b_i x^i$ . From g(f(x)) = x, we know that

$$x = \sum_{i \ge 1} b_i f(x)^i.$$

By taking derivative, we get

$$1 = \sum_{i \ge 1} i b_i f^{i-1}(x) f'(x).$$

Now divide both sides by  $f^n(x)$ , so we can get

$$\frac{1}{f^n(x)} = \sum_{i \ge 1} i b_i f^{i-1-n}(x) f'(x).$$

Then,

$$[x^{-1}]\frac{1}{f^n(x)} = \sum_{i \ge 1} ib_i [x^{-1}]f^{i-1-n}(x)f'(x) = nb_n = n[x^n]g(x).$$

by the above lemma.

There exists an alternate form; let  $f(x) = \frac{x}{\phi(x)} \iff g(x) = x\phi(g(x))$ . Then,

$$[x^{n}]g(x) = \frac{1}{n}[x^{n-1}]\phi^{n}(x).$$

To see this, note that f(g(x)) = x = g(f(x)). Hence by the above theorem, we can get

$$[x^{n}]g(x) = \frac{1}{n} \cdot [x^{-1}] (f(x))^{-n} = \frac{1}{n} \cdot [x^{-1}] \frac{\phi^{n}(x)}{x^{n}} = \frac{1}{n} [x^{n-1}] \phi^{n}(x).$$

In the textbook, it says that if  $f(x) = xG(x) \iff g(x) = \frac{x}{G(g(x))}$ , then

$$[x^{n}]g(x) = \frac{1}{n}[x^{n-1}]G^{-n}(x).$$

We can derive this just thinking  $[x^{-1}]\frac{1}{f^n(x)} = [x^{-1}]x^{-n}G^{-n}(x) = [x^{n-1}]G^{-n}(x)$ . Corollary 3.3.3 (Textbook version). Let  $F(z) = zG(F(z)), G(0) \neq 0$ , then

$$[x^{n}]F(z) = \frac{1}{n}[z^{n-1}]G^{n}(z).$$

We already show this using above theorem, but we can show it using combinatorial argument. *Proof.* Assume

$$G(z) = g_0 + \frac{g_1}{1!}z + \frac{g_2}{2!}z^2 + \cdots$$

Then,

$$\begin{aligned} G^{n}(z) &= \left(g_{0} + \frac{g_{1}}{1!}z + \frac{g_{2}}{2!}z^{2} + \cdots\right)^{n} \\ &= \sum_{r_{0}, \cdots, r_{n}} \binom{n}{r_{0}, r_{1}, \cdots, r_{n}} g_{0}^{r_{0}} \left(\frac{g_{1}}{1!}z\right)^{r_{1}} \left(\frac{g_{2}}{2!}z^{2}\right)^{r_{2}} \cdots \left(\frac{g_{n}}{n!}z^{n}\right)^{r_{n}} \\ &= \sum_{r_{0}, r_{1}, \cdots, r_{n}} \binom{n}{r_{0}, r_{1}, \cdots, r_{n}} g_{0}^{r_{0}} \cdots g_{n}^{r_{n}} \frac{z^{r_{1}+2r_{2}+\cdots+nr_{n}}}{(0!)^{r_{0}}(1!)^{r_{1}}(2!)^{r_{2}}\cdots(n!)^{r_{n}}} \end{aligned}$$

Thus,

$$[z^{n-1}]G^n(z) = \sum_{\substack{r_0, \cdots, r_n \\ \sum_{i=0}^n ir_i = n-1}} \binom{n}{r_0, r_1, \cdots, r_n} g_0^{r_0} \cdots g_n^{r_n} \frac{1}{(0!)^{r_0} (1!)^{r_1} (2!)^{r_2} \cdots (n!)^{r_n}}$$

Then if we just thinking type  $(r_0, \dots, r_n)$ , then  $g_0^{r_0} \cdots g_n^{r_n} = \prod_{i \in [n]} g_{\deg(i)}$ , so

$$(n-1)![z^{n-1}]G^{n}(z) = \sum_{\substack{r_{0}, \cdots, r_{n} \\ \sum_{i=0}^{n} ir_{i}=n-1}} \binom{n}{r_{0}, r_{1}, \cdots, r_{n}} g_{0}^{r_{0}} \cdots g_{n}^{r_{n}} \frac{(n-1)!}{(0!)^{r_{0}} (1!)^{r_{1}} (2!)^{r_{2}} \cdots (n!)^{r_{n}}}$$
$$= \sum_{T: \text{ rooted tree on } n \text{ vertices}} g_{0}^{r_{0}} g_{1}^{r_{1}} \cdots g_{n}^{r_{n}}$$

since  $\sum_{\substack{r_0,\cdots,r_n\\\sum_{i=0}^n ir_i=n-1}}^{r_0,\cdots,r_n} {n \choose r_0,r_1,\cdots,r_n} \frac{(n-1)!}{(0!)^{r_0}(1!)^{r_1}(2!)^{r_2}\cdots(n!)^{r_n}}$  counts the number of rooted trees on n vertices, and  $g_0^{r_0}g_1^{r_1}\cdots g_n^{r_n}$  counts how many vertices with specific degree occurs. So we can think  $g_0^{r_0}g_1^{r_1}\cdots g_n^{r_n}$  as a degree counter for each rooted tree T the summation visit.

Now, note that the desired equation

$$[x^{n}]F(z) = \frac{1}{n}[z^{n-1}]G^{n}(z)$$

is equivalent to say that for the representation  $F(z) = \sum_{n \ge 1} \frac{f_n}{n!} z^n$ ,

$$\frac{f_n}{n!} = \frac{1}{n} \cdot \frac{1}{(n-1)!} \sum_{T: \text{ rooted tree on } n \text{ vertices}} g_0^{r_0} g_1^{r_1} \cdots g_n^{r_n}$$

which is equivalent to say that

$$f_n = \sum_{T: \text{ rooted tree on } n \text{ vertices}} g_0^{r_0} g_1^{r_1} \cdots g_n^{r_n}$$

Thus, it suffices to show that if  $f_n = \sum_{T: \text{ rooted tree on } n \text{ vertices }} g_0^{r_0} g_1^{r_1} \cdots g_n^{r_n}$ , then  $y = E_f$  satisfy y = zG(y), where  $E_f(z) = \sum_{n \ge 1} f_n \frac{z^n}{n!}$ . This is because such F satisfying the condition F(z) = zG(F(z)) is unique; to see this, if H also satisfies the equation H(z) = zG(H(z)), then

$$(F - H)(z) = z(G(F(z)) - G(H(z))) = z\left(\sum_{k \ge 1} g_k(F(z)^k - H(z)^k)\right)$$

and by comparing each coefficients of  $x^k$  using induction, we can conclude F = H, thus done.

To see this, we need a combinatorial picture; suppose  $f_n = \sum_{T: \text{ rooted tree on } n \text{ vertices }} g_0^{r_0} g_1^{r_1} \cdots g_n^{r_n}$ . And note that any rooted tree on [n+1] can be regarded as a rooted forest on [n] by deleting its root. So,

$$f_{n+1} = (n+1) \cdot \sum_{T_1 \cup \dots \cup T_k \text{ on } [n+1] \setminus \{\text{root}\}} \left( \bar{g}^{T_1} \cdots \bar{g}^{T_k} \right) \cdot g_k$$

which means, (n + 1) is just the number of ways of choosing a root, and the summation just permutes all possible rooted tree with component k, where  $1 \le k \le n$ , and then each component of the forest should have particular degree counts  $\bar{g}^T = g_0^{r_0} \cdots g_n^{r_n}$  where  $(r_0, \cdots, r_n)$  is type of the tree T, and the last  $g_k$  term counts the root. Thus,

$$\frac{f_{n+1}}{n+1} = \sum_{T_1 \cup \dots \cup T_k \text{ on } [n+1] \setminus \{\text{root}\}} \left(\bar{g}^{T_1} \cdots \bar{g}^{T_k}\right) \cdot g_k.$$

Also, if we let  $\bar{g}^k$  as sum of all  $\bar{g}^T$  with |T| = k, then  $\bar{g}^k = f_{|T_k|}$ , so we can apply the compositional formular (theorem 3.3 in [Aig07]) to get

$$\frac{f_{n+1}}{n+1} = [x^n]G(F(z)).$$

(Note that this can be derived by representing the sum as a sum of partitions of  $[n+1] \setminus \{k\}$ .) Hence,

$$F(z) = \sum_{n \ge 1} f_n \frac{z^n}{n!} = z \sum_{n \ge 0} \frac{f_{n+1}}{n+1} \frac{z^n}{n!} = z \sum_{n \ge 0} \left( [x^n] G(F(z)) \right) \frac{z^n}{n!} = z G(F(z))$$

as desired.

#### Corollary 3.3.4 (Lagrange Inversion formula).

1. Suppose  $f(x) \in x\mathbb{C}[[x]]$  and  $g(x) = f^{\langle -1 \rangle}(x)$ , i.e., compositional inverse. Then,

$$\begin{split} [x^{n}]g(x) &= \frac{1}{n} [x^{-1}] f^{-n}(x) \\ [x^{n}]g^{k}(x) &= \frac{k}{n} [x^{-k}] f^{-n}(x) \\ &= \frac{k}{n} [x^{-1}] x^{k-1} f^{-n}(x) \\ [x^{n}]H(g(x)) &= [x^{n}] \frac{H'(x)}{f^{n}(x)} \text{ for any } H. \end{split}$$

2. Let  $f(x) = \frac{x}{\phi(x)}$  where  $\phi(x) \in \mathbb{C}[[x]], \ \phi(0) \neq 0$ , and let  $g(x) = x\phi(g(x))$ . Then,

$$[x^{n}]g(x) = \frac{1}{n}[x^{n-1}]\phi^{n}(x)$$
$$[x^{n}]g^{k}(x) = \frac{k}{n}[x^{n-k}]\phi^{n}(x)$$
$$[x^{n}]H(g(x)) = [x^{n}]H'(x)\phi^{n}(x) \text{ for any } H.$$

3. Let  $f(x) = xG(x) \iff g(x) = \frac{x}{G(g(x))}$ . Then,  $[x^n]g(x) = \frac{1}{n}[x^{n-1}]G^{-n}(x)$   $[x^n]g^k(x) = \frac{k}{n}[x^{n-k}]G^{-n}(x)$  $n[x^n]H(g(x)) = [x^{n-1}]H'(x)G^{-n}(x)$  for any H.

*Proof.* For (i), the proof of second inequality follows almost same procedure for proving the first equality, which we proved earlier. Let  $g^k(x) = \sum_{i \ge k} b_i x^i$ . Then,

$$x^k = \sum_{i \ge k} p_i f(x)^i.$$

By differentiation and dividing both sides by  $f^n(x)$ , we get

$$\frac{kx^{k-1}}{f^n(x)} = \sum_{i \ge k} ip_i f(x)^{i-n-1} f'(x).$$

Then,

$$[x^{-1}]\sum_{i\geq k} ip_i f(x)^{i-n-1} f'(x) = np_n$$

since  $f(x)^{i-n-1}f'(x) = \frac{1}{i-n}\frac{d}{dx}(f^{i-n}(x))$  for  $i \neq n$ , thus any other values except i = n gives zero, and i = n value is  $np_n$  as we shown above. And, for LHS,

$$\frac{kx^{k-1}}{f^n(x)} = \frac{kx^{k-1}}{(f_1x + f_2x^2 + \dots)^n} = \frac{kx^{k-n-1}}{(f_1 + f_2x + \dots)^n}$$

Thus,

$$[x^{-1}]\frac{kx^{k-1}}{f^n(x)} = np_n = n[x^n]g(x)$$

Now note that

$$[x^{-1}]\frac{kx^{k-1}}{f^n(x)} = k[x^{-1}]\frac{x^{k-1}}{f^n(x)} = k[x^{-k}]f^{-n}(x)$$

For the third inequality, note that H consists of linear sum of  $x^k$ 's hence it suffices to show that  $H = x^k$ , and this case is what we proved earlier, done.

For other cases, just put such functions on the first case, and thinking the index number.

There are several examples of applications of Lagrange Inversion Formula.

1. Let  $r_n$  be the number of rooted tree. Then,  $y = \sum_{n \ge 1} r_m \frac{x^n}{n!}$  satisfies

$$y = xe^y$$

as we shown above, (introducing the rooted forest). Then, by the version 2 of Lagrange Inversion Formula, we have  $a_{1} = a_{1} (a_{1}) \text{ where } f(a_{2}) = c^{2}$ 

$$y = x\phi(y)$$
 where  $\phi(x) = e^{\frac{1}{2}}$ 

 $\mathbf{SO}$ 

$$\frac{r_n}{n!} = [x^n]y = \frac{1}{n}[x^{n-1}]e^{nx} = \frac{1}{n}[x^{n-1}]\sum_{k\ge 0}\frac{n^k}{k!}x^k = \frac{1}{n}\cdot\frac{n^{n-1}}{(n-1)!}$$

Thus,

$$r_n = n^{n-1}.$$

Also,

$$[x^{n}]y^{k} = \frac{k}{n}[x^{n-k}]e^{nx} = \frac{k}{n}\frac{n^{n-k}}{(n-k)!}$$

Since

$$\sum_{k\geq 0} P_k(n) \frac{x^k}{n!} = \frac{y^k}{k!}$$

where  $P_k(n)$  is the number of rooted forests on [n] with exactly k trees, and this equation is also shown in above. Thus,

$$P_k(n)\frac{k!}{n!} = \frac{k}{n} \cdot \frac{n^{n-k}}{(n-k)!} \implies P_k(n) = \binom{n-1}{k-1}n^{n-k}$$

2. For catalan number, let

$$y = \sum_{n \ge 0} C_n z^n$$

By the Catalan number's generating function, we get

$$y = zy^2 + 1.$$

To apply LIF,

(a) Let g = zy and  $g = g^2 + z \implies g(1 - g) = z$ . So, g is compositional inverse of f(z) = z(1 - z). Hence by applying version 3, on G(z) := 1 - z, we get

$$[z^{n}]g(z) = \frac{1}{n}[z^{n-1}](1-z)^{-n} = \frac{1}{n}\binom{-n}{n-1}(-1)^{n-1} = \frac{1}{n}\binom{n+(n-1)-1}{n-1} = \frac{1}{n}\binom{2n-2}{n-1}.$$

Thus,

$$c_n = [z^{n+1}]g = \frac{1}{n+1} \binom{2n}{n}.$$

(b) Let g = y - 1. Then,  $1 + g = z(1 + g)^2 + 1$  and  $g = z(1 + g)^2$ . Then to use the version 2 of LIF, let  $\phi(z) = (1 + z)^2$ . Then,

$$[z^{n}]g(z) = \frac{1}{n}[z^{n-1}]\phi^{n}(z) = \frac{1}{n}[z^{n-1}](1+z)^{2n} = \frac{1}{n}\binom{2n}{n} = \frac{1}{n}\binom{2n-2}{n-1}$$

## 3.4 Generating function for Integer Partition

Let  $\lambda = (\lambda_1, \dots, \lambda_k) \ge 0$  be a k-tuple such that  $\lambda_1 + \dots + \lambda_k$ . Then denote  $\lambda \vdash n$  and call that  $\lambda$  partitions n. And size of  $\lambda$ , say  $|\lambda|$  is n. Now denote

$$\begin{split} p(n) &:= \#\{\lambda \vdash n\}\\ p(n, \leq k) &:= \#\{\lambda \vdash n : \text{ the largest part of } \lambda \text{ is less than or equal to } k \; \}. \end{split}$$

Then we have

$$\sum_{n \ge 0} p(n, \le k) z^n = \frac{1}{(1-z)(1-z^2)\cdots(1-z^k)} = \prod_{i=1}^k \frac{1}{1-z^i}$$

since

$$\frac{1}{(1-z)(1-z^2)\cdots(1-z^k)} = \underbrace{(1+z+z^2+\cdots)}_{a_1 \text{ parts of size } 1} \underbrace{(1+z^2+z^4+\cdots)}_{a_2 \text{ parts of size } 2} \cdots \underbrace{(1+z^k+z^{2k}+\cdots)}_{a_k \text{ parts of size } k} = \sum_{\substack{\lambda = (\lambda_1, \cdots, \lambda_l) \\ \lambda_i \le k}} z^{|\lambda|}$$

And the last sum is just the same one as  $\sum_{n\geq 0} p(n, \leq k) z^n$ . Now, below one counts how many parts the  $\lambda$  has using generating function;

$$\sum_{\substack{\lambda = (\lambda_1, \cdots, \lambda_l) \\ \lambda_i \le k}} t^{\# \text{part}(\lambda)} z^{|\lambda|} = (1 + tz + t^2 z^2 + \cdots)(1 + tz^2 + t^2 z^4 + \cdots) \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k + t^2 z^{2k} + \cdots) = \prod_{i=1}^{\kappa} \frac{1}{1 - tz^i} \cdots (1 + tz^k +$$

where each exponent of t in paranthese of  $1/(1-z^l)$  counts how many parts with size l occur on the partition. If  $k \to \infty$ , then

$$\lim_{k \to \infty} \sum_{n \ge 0} p(n, \le k) z^n = \lim_{k \to \infty} \prod_{i=1}^k \frac{1}{1 - z^i} = \prod_{i=1}^\infty \frac{1}{1 - z^i} = \sum_{n \ge 0} p(n) z^n.$$

Similarly,

$$\lim_{k \to \infty} \sum_{\substack{\lambda = (\lambda_1, \cdots, \lambda_l) \\ \lambda_i \le k}} t^{\# \operatorname{part}(\lambda)} z^{|\lambda|} = \prod_{i=1}^{\infty} \frac{1}{1 - t z^i}.$$

Now define

$$p(n,k) = #\{\lambda \vdash n : \text{ with exactly } k \text{ parts}\} = #\{\lambda \vdash n, \lambda_1 = k\}$$

since every  $\lambda \vdash n$  with exactly k parts has conjugate  $\lambda^*$  having  $\lambda_1^* = k$  and  $|\lambda^*| = n$ . So counting such conjugates is the same as counting those with exactly k parts. Thus,

$$\sum_{n \ge 0} p(n,k) z^n = \frac{z^k}{(1-z)\cdots(1-z^k)}$$

since size k parts occur at least once in every  $\lambda$  counted by p(n,k) and vice versa.

Now let  $\lambda$  is a partition having  $\leq k$  parts and the highest summand is  $\leq m$ . Then, in the section of Gaussian Coefficient, we proved that

$$\sum_{\lambda \in Par(;\leq k;\leq m)} q^{|\lambda|} = \binom{m+k}{k}_q$$

and

$$\sum_{\lambda \in Par(k; k \leq m)} q^{|\lambda|} = \binom{m+k-1}{k}_q.$$

The second one can be derived by counting any path with at most k parts and the highest summand at most m-1, then add size k block on the first line of the Ferrer diagram, and conjugate it.



Then,

$$p(n, \text{distinct}) = \#\{\lambda \vdash n; \lambda_1 > \lambda_2 > \dots\} = \sum_{n \ge 0} p(n, \text{distinct}) z^n t^{\# \text{part}(\lambda)} = \prod_{i=1}^{\infty} (1 + z^i t)$$

since exactly 1 case of size i occur.

From  $\sum_{n\geq 0} p(n) z^n = \prod_{i=1}^{\infty} \frac{1}{1-z^i}$ , we can derive the number of  $\lambda$  with exactly k distinct parts as

$$\sum_{\lambda = (\lambda_1 > \dots > \lambda_k > 0)} z^{|\lambda|} = \sum_{\mu: \text{ partition}} z^{|\mu| + \binom{k+1}{2}} = z^{\binom{k+1}{2}} \prod_{i=1}^k \frac{1}{1 - z^i}$$

since any such partition can be decomposed with

$$\lambda = \mu + (k, k - 1, \cdots, 1)$$

Thus by counting all partition with at most k parts and adding  $(k, k-1, \dots, 1)$  we can construct every such partition.

### 3.4.1 Euler's Pentagonal Theorem

 $\sim$ 

Theorem 3.4.1.

$$\prod_{i=1}^{\infty} (1-z^i) = \sum_{n=-\infty}^{\infty} (-1)^n z^{\frac{3n^2+n}{2}}$$

Motivation of the theorem is from computing p(n); Note that

$$\left(\sum_{n\geq 0} p(n)z^n\right)\prod_{i=1}^{\infty} (1-z^i) = \prod_{i=1}^{\infty} \frac{1}{1-z^i}\prod_{i=1}^{\infty} (1-z^i) = 1$$

And we can observe that

$$\prod_{i=1}^{\infty} (1-z^i) = 1 - z - z^2 + z^5 + z^7 - z^{12} - z^{15} + z^{22} + z^{26} + \cdots$$

To see this, combinatorially,

$$[z^n]LHS = E(n) - O(n)$$

where E(n) is the number of partitions of n with an even number of distinct parts and O(n) is the number of partitions of n with a odd number of distinct parts. This is because

$$\prod_{i=1}^{\infty} (1-z^i) = \sum_{\substack{k_i \in \{1,0\}\\i \in \mathbb{N}}} (-1)^{\sum_{j=1}^{\infty} k_i} z^{\sum_{j=1}^{\infty} j k_j}.$$

From choosing each box of  $(1 - z^i)$  we have  $k_i = 1$  or 0, and sign is determined by evenness or odddness of the number of chosen box.

Now we claim that

$$E(n) - O(n) = \begin{cases} (-1)^m & n = \frac{3m^2 \pm m}{2} \\ 0 & \text{o.w.} \end{cases}$$

And actually, this claim gives the name pentagonal, since the nonzero index  $1, 5, 12, 22, \cdots$  with changed sign compared with preceding term is just number of points in the expanding pentagon with length  $n \in \mathbb{N}$ ;



For example, above black dots gives 1, and black and blue dots gives 5, and black, blue and red dots gives 1 + 4 + 7 = 11, and so on.

Now to prove the claim, we need to investigate the Ferrer diagram; think below diagram for example.



a is the number of points in the last line of the diagram (blue region), and b is the number of points in the farthest diagonal line (red region). Then, if a > b, move part of red region to bottom of the a parts. If  $a \le b$  then move part of blue region to outside of red region. So in the above example,



below is the result of the operation.



This gives a bijection between underlying sets of E(n) and O(n) generically. "Generically" means that, actually this bijection cannot holds when blue region and red region contains nonempty intersection, and a = b or a = b + 1 holds. (Note that any other difference, the above operation still works since each operation requires  $\min(a, b)$  amounts of position for  $\max(a, b)$  region. Only nonempty intersection with a = b or a = b + 1 cannot satisfy this requirement.)

So from this we can count such bad cases; as a table it occur when

And note that a, b in the "bad" cases determines whole diagram, since it determines all boundary of the diagram. Actually, in any bad case,

$$ab + \frac{b(b-1)}{2} = \frac{3b^2 \pm b}{2}$$

points occur. Thus for each n, if  $n = \frac{3b^2 \pm b}{2}$ , then only one case occur, and since it has b parts, so sign is determined by b, i.e.,  $(-1)^b$ , done.

For more information, see [DK16] by google.

#### 3.4.2 Jacobi Triple Identity

Theorem 3.4.2 (Jacobi Triple Identity).

$$\prod_{k\geq 1} (1+zq^k)(1+z^{-1}q^{k-1})(1-q^k) = \sum_{n=-\infty}^{\infty} q^{\frac{n(n+1)}{2}} z^n$$

Remark that if we put  $q^3 \rightarrow q$  and  $z = -q^{-1}$ , then the above equation gives

$$\prod_{k\geq 1} (1-q^{3k-1})(1-q^{3k-2})(1-q^{3k}) = \prod_{k\geq 1} (1-q^k) \underbrace{=}_{\text{Euler's Pentagonal Thm}} \sum_{n=-\infty}^{\infty} (-1)^n q^{\frac{3n^2+n}{2}}$$

In fact, it has original form, which is

$$\prod_{m \ge 1} (1 - x^{2m})(1 + x^{2m-1}y^2) \left(1 + \frac{x^{2m-1}}{y}\right) = \sum_{n = -\infty}^{\infty} x^{n^2} y^{2n}$$

You can see it by putting  $x = \sqrt{q}$  and  $y^2 = \sqrt{q}z$ .

Also, this is equivalent to saying that

$$\prod_{k \ge 1} (1 + zq^k)(1 + z^{-1}q^{k-1}) = \left(\sum_{n = -\infty}^{\infty} q^{\frac{n(n+1)}{2}} z^n\right) \cdot \left(\prod_{k=1}^{\infty} \frac{1}{1 - q^k}\right).$$

Note that  $\frac{1}{1-q^k} = \sum_{n \ge 0} p(n) q^n$ .

To prove the theorem, let  $F(z) := \prod_{k \ge 1} (1 + zq^k)(1 + z^{-1}q^{k-1})$ . First step is to find out combinatorial interpretation of F(z). So if we choose n+iz-terms from  $\prod_{k\ge 1} (1+zq^k)$  and iz-terms from  $\prod_{k\ge 1} (1+z^{-1}q^{k-1})$  then it gives  $z^n$ . Thus,

$$[z^{n}]F(z) = \sum_{\substack{\lambda:\lambda_{1} > \dots > \lambda_{n+i} \\ \mu:\mu_{1} > \dots > \mu_{i} > 0}} q^{\sum_{j=1}^{n+i}\lambda_{i} - \sum_{l=1}^{i}(\mu_{l}-1)} = \sum_{\substack{\lambda:\lambda_{1} > \dots > \lambda_{n+i} > 0 \\ \mu:\mu_{1} > \dots > \lambda_{i} > 0}} q^{|\lambda| + |\tilde{\mu}|} = \sum_{\substack{\lambda:\lambda_{1} > \dots > \lambda_{n+i} > 0 \\ \tilde{\mu}:\tilde{\mu}_{1} > \dots > \tilde{\mu}_{i} \ge 0}} q^{|\lambda| + |\tilde{\mu}|}$$

where  $\tilde{\mu} = (\mu_1 - 1, \cdots, \mu_i - 1)$ . Second step is let

$$F(z) = \sum_{n = -\infty}^{\infty} a_n(q) z^n.$$

Then, by replacing z with qz, we get

$$F(qz) = \prod_{k \ge 1} (1 + (zq)q^k)(1 + z^{-1}q^{-1}q^{k-1}) = F(z)\frac{1 + z^{-1}q^{-1}}{1 + zq} = \frac{1}{zq}F(z).$$

Note that second step can be attained from comparing each terms in F(z) and F(qz); F(qz) has  $(1+(zq)^{-1})$  term but has no (1+zq) term. This implies

$$\sum_{n=-\infty}^{\infty} a_n(q)q^n z^n = F(qz) = \frac{1}{qz}F(z) = \sum_{n=-\infty}^{\infty} a_n(q)q^{-1}z^{n-1}$$

By equating coefficient of  $z^{n-1}$ , we get

 $a_{n-1}(q)q^{n-1} = a_n(q)q^{-1}, \forall n \in \mathbb{Z} \implies a_n(q) = q^n a_{n-1}(q), \forall n \in \mathbb{Z}.$ 

So, if  $n \ge 0$ , then

$$a_n(q) = q^{\binom{n+1}{2}} a_0(q)$$

and for n < 0, we have

 $a_n(q)q^{n+1} = a_{n+1}(q), \forall n \in \mathbb{Z} \implies a_n(q) = q^{-(n+1)-(n+2)-\dots-1-0}a_0(q) = q^{\frac{-n(-n+1)}{2}}a_0(q) = q^{\binom{n+1}{2}}a_0(q), \forall n \in \mathbb{Z} \setminus \mathbb{N}.$ since  $-(n+1) - (n+2) - \dots - 1 - 0$  is just sum from 0 to |n| - 1. Hence, in any case,

$$a_n(q) = q^{\binom{n+1}{n}} a_0(q).$$

Thus,

$$F(z) = a_0(q) \sum_{n \in \mathbb{Z}} q^{\binom{n+1}{2}} z^n$$

Step three is just show what  $a_0(q)$  is. It suffices to show that

$$a_0(q) = \prod_{k=1}^{\infty} \frac{1}{(1-q^k)}$$

Then, from step 1, we know

$$a_0(q) = [z^0]F(z) = \sum_{i=1}^{\infty} \sum_{\substack{\lambda:\lambda_1 > \dots > \lambda_i > 0\\ \tilde{\mu}: \tilde{\mu_1} > \dots > \tilde{\mu}_i \ge 0}} q^{|\lambda| + |\tilde{\mu}|}$$

and actually, any partition can be decomposed by two distinct partitions, one has nonzero elements; to see this, just think the Ferrer diagram, and cut the diagram by diagonal line, and the upper part with diagonal one gives  $\lambda$  and the lower part without diagonal gives  $\tilde{\mu}$ . Thus,  $a_0(q)$  counts all possible partitions, so

$$a_0(q) = \sum_{n \ge 0} p(n)q^n = \prod_{k=1}^{\infty} \frac{1}{(1-q^k)},$$

done.

# 4 Gentle introduction to WZ method

We want to prove or disprove the given equation

$$\sum_{k \ge 0} f(n,k) = F(n).$$

Then, divide both sids to get 1, in this case,

$$\sum_{k\ge 0}\frac{f(n,k)}{F(n)} = 1.$$

Let  $F(n,k) = \frac{f(n,k)}{F(n)}$ . Then our goal is to check that

$$\sum_{k \ge 0} F(n,k) = 1,$$

independent of n. And it suffices to check that

- 1.  $\sum_{k\geq 0} (F(n+1,k) F(n,k)) = 0$  for all n.
- 2. Check the value  $\sum_{k\geq 0} F(n,k)$  in case of n=0.

To deal with the first condition, our dream is as below;

1. Dream1: If  $\exists G(n,k)$  such that

$$F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k)$$

then

$$\sum_{k=B}^{A} F(n+1,k) - F(n,k) = \sum_{k=B}^{A} G(n,k+1) - G(n,k) = G(n,A+1) - G(n,B).$$

2. Dream2: For A sufficiently large and B sufficiently small,

$$G(n, A) = 0 = G(n, B).$$

Then, G(n, A + 1) - G(n, B) = 0.

For example, think about  $\sum_{k=0}^{n} {n \choose k} = 2^{n}$ . Then, it is equivalent to say

$$\sum_{k=0}^{n} \frac{\binom{n}{k}}{2^n} = 1.$$

Let  $F(n,k) = \frac{\binom{n}{k}}{2^n}$ . Then,

$$G(n,k) = -\frac{\binom{n}{k-1}}{2^{n+1}}.$$

Then we can check that

$$G(n, k+1) - G(n, k) = F(n+1, k) - F(n, k)$$

Actually if we compare F(n+1,k)+G(n,k) and F(n,k)+G(n,k) using Pascal's identity, you can check that it is true. Thus, for any sum between A, B with respect to k the Dream1 comes true. Also, for k < 1, k > n, G(n,k) = 0. Hence our dreams come true, and by checking the case n = 1, actually  $\sum_{k=0}^{n} F(0,k) = 1 = 2^{0}$ , done.

The pair (F, G) is called **WZ-pair**. W denotes WilfZeilberger pair. How to find G(n, k)? There is a algorithmic way of finding it. It is known as **Gosper's algorithm**. Our question can be refined as follow; given

$$D(n,k) = F(n+1,k) - F(n,k)$$

we want to find G(n,k) such that G(n,k+1) - G(n,k) = D(n,k), i.e.,

$$G(n, k+1) = \sum_{i=-\infty}^{k} D(n, i).$$

So, first we want to ask is that G(n,k) has whether closed form or not. Now suppose that D(n,k) is hypergeometric function. Then, Gosper's algorithm says that there exists an algorithm which will tell you in finite time either G(n,k) with closed form exists or not. And meaning of hypergeometric function is that

$$\frac{D(n+1,k)}{D(n,k)}, \frac{D(n,k+1)}{D(n,k)}$$

are rational function of n and k.

# 5 Sieve Method

### 5.1 Principle of Inclusion-Exclusion

We just abbreviate this principle as PIE. Let  $A_1, A_2, \dots, A_n$  be subset of X. Then,

$$\left|\overline{\bigcup_{i} A_{i}}\right| = \left|\bigcap_{i} \overline{A_{i}}\right| = |X| - \sum_{i} |A_{i}| + (-1)^{2} \sum_{i \leq j} |A_{i} \bigcap A_{j}| + \dots + (-1)^{n} |A_{1} \cap A_{2} \cap \dots \cap A_{n}|.$$

For special case, if i = 1, then

$$|A| = |X| - |A|.$$
  
 $|A \cup B| = |A| + |B| - |A \cap B|.$ 

In application, let X be universe. Then there are n properties  $p_1, \dots, p_n$ . For  $T \subseteq [n]$ , let

 $N_{=T} = \#\{x \in X : x \text{ has exactly all properties } p_i \text{ in } T\}$ 

and

 $N_{\supseteq T} = \#\{x \in X : x \text{ has at least all properties } p_i \text{ in } T\}$ 

Now define

 $A_i := \{ x \in X : x \text{ has property } p_i \}.$ 

So if  $T = \{1, 2, 3\}$ , then

$$N_{\supset T} = |A_1 \cap A_2 \cap A_3|.$$

Now, PIE can be restated as

Theorem 5.1.1 (Principle of Inclusion-Exclusion).

$$N_{=\emptyset} = \sum_{T \subseteq [n]} (-1)^{|T|} N_{\supseteq T}.$$

If  $N_{=T}, N_{\supseteq T}$  depend on |T| but not content of T, which we say **homogeneous case**, then let  $f_k := N_{=T}$ and  $g_k = N_{\supseteq T}$  for |T| = k. Then, PIE implies

$$f_0 = \sum_{k=0}^n (-1)^k \binom{n}{k} g_k.$$

For example,

1. Think the problem of finding the number of integer solutions of

$$x_1 + \dots + x_n = k.$$

Then by bar-ball technique, we know the number of nonnegative integer solution is  $\binom{n+k-1}{k}$ . Also, the number of integer solutions such that  $x_i \ge a_i$  for any  $a_i \in \mathbb{Z}$  is the same as nonnegative integer solution of

$$\sum_{i=1}^{n} y_i = k - \sum a_i.$$

since any solution  $(y_i)_{i=1}^n$  can be translated to  $(x_i)_{i=1}^n = (y_i)_{i=1}^n + (a_i)_{i=1}^n$ .

Now, using PIE, we can count the number of integer solution when each  $x_i$  has an upper bound. For example, let  $x_1 + x_2 + x_3 = 200$  and  $x_1 \le 100, x_2 \le 101, x_3 \le 102$ . Then let

$$X = \{ \text{ nonnegative integer solution of } x_1 + x_2 + x_3 = 200 \}$$
  

$$A_1 = \{ \text{ solutions where } x_1 \ge 101 \}$$
  

$$A_2 = \{ \text{ solutions where } x_1 \ge 102 \}$$
  

$$A_3 = \{ \text{ solutions where } x_1 \ge 103 \}.$$

Then,  $|\overline{A_1 \cup A_2 \cup A_3}|$  is the desired answer, and we can calculate it via PIE. For special case, where

$$x_1 + \dots + x_n = k, 0 \le x_i < s, x_i \in \mathbb{N},$$

we can get

$$|A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_j}| = \#\{(y_1, \dots, y_n) \in \mathbb{N}^n : y_1 + \dots + y_n = k - js\} \\ = \binom{k - js + n - 1}{n - 1}.$$

Thus, we can define  $g_j$ , so it gives

$$N_{=\emptyset} = \sum_{j=0}^{n} (-1)^{j} \binom{n}{j} \binom{k-js+n-1}{n-1}.$$

If s = 1, then the number of solution is  $\delta_{0,k}$ , since if  $k \neq 0$  then there is no solution. Thus

$$N_{=\emptyset} = \sum_{j=0}^{n} (-1)^{j} \binom{n}{j} \binom{k-j+n-1}{n-1} = \delta_{0,k}.$$

### 2. Evaluate

$$\sum_{k=0}^{m} (-1)^k \binom{m}{k} \binom{n-k}{r} \text{ where } m \le r \le n.$$

Idea is to regard  $\binom{n-k}{r}$  as some kind of  $N_{\geq T}$  for |T| = k. We can design a combinatorial situation; fix m elements, say  $\{1, 2, \dots, m\}$  from [n] and X be set of all r-subsets of [n]. Let

$$p_i := \{T \subseteq [n] : i \notin T\}, \forall i \in [m].$$

Then,

$$N_{\supseteq T} = \# r$$
-subsets of  $[n]$  not containing  $|T|$  elements  $= \binom{n-k}{r}$ 

And,

$$N_{=\emptyset} = \# r$$
-subsets of  $[n]$  containing  $1, 2, \cdots, m$ . =  $\binom{n-m}{r-m}$ 

Thus, by PIE and above equation, we get

$$\binom{n-m}{r-m} = N_{=\emptyset} = \sum_{k=0}^{m} (-1)^k \binom{m}{k} \binom{n-k}{r} \text{ where } m \le r \le n.$$

Also note that

 $N_{=T} = \# r$ -subsets of [n] not containing elements in T but containing elements in [m] - T which is

$$\binom{n-m}{r-m+k}$$
 where  $k = |T|$ .

**Theorem 5.1.2** (5.1 in [Aig07]). Let E be a finite set. f, g are two functions from  $2^E$  to  $\mathbb{R}$ . Then,

$$f(A) = \sum_{T \supseteq A} g(T) \iff g(A) = \sum_{T \supseteq A} (-1)^{|T| - |A|} f(T), \forall A \in 2^E$$

*Proof.* If we let  $\bar{f}, \bar{g}$  are vectors of dimension  $2^n$  induced by  $2^E$ , then the above statement is equivalent to

$$\bar{f} = M\bar{g} \iff \bar{g} = M^{-1}\bar{f}$$

where  $M, M^{-1}$  are  $2^n \times 2^n$  matrices indexed by  $I = 2^E$  such that

$$M_{S,T} = \chi(S \subseteq T)$$

Then,

$$M_{S,T}^{-1} = (-1)^{|T| - |S|} \chi(S \subseteq T)$$

Actually, to verify this  $M, M^{-1}$  works, we need a proof of the theorem; assume LHS. Then,

$$\sum_{T \supseteq A} (-1)^{|T| - |A|} f(T) = \sum_{T \supseteq A} (-1)^{|T| - |A|} \sum_{U \supseteq T} g(U) = \sum_{U \supseteq A} \left( \sum_{U \supseteq T \supseteq A} (-1)^{|T| - |A|} \right) g(U).$$

,

And if  $|U \setminus A| = m$ , then

$$\sum_{U \supseteq T \supseteq A} (-1)^{|T| - |A|} = \sum_{k=0}^{m} (-1)^k \binom{m}{k} = (1-1)^m = \delta_{m,0}.$$

Thus,

$$\sum_{T \supseteq A} (-1)^{|T| - |A|} f(T) = \sum_{U \supseteq A} \left( \sum_{U \supseteq T \supseteq A} (-1)^{|T| - |A|} \right) g(U) = (-1)^{|A| - |A|} g(A) = g(A).$$

Conversely, if we assume RHS, then

$$\sum_{T \supseteq A} g(T) = \sum_{T \supseteq A} \sum_{U \supseteq T} (-1)^{|U| - |T|} f(U) = \sum_{U \supseteq A} \left( \sum_{U \supseteq T \supseteq A} (-1)^{|U| - |T|} \right) f(U)$$

Then, if  $|U \setminus A| = m$ , then

$$\sum_{U \supseteq T \supseteq A} (-1)^{|U| - |T|} = \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} (1-1)^{m} = \delta_{m,0},$$

 $\mathbf{SO}$ 

$$\sum_{T \supseteq A} g(T) = \sum_{U \supseteq A} \left( \sum_{U \supseteq T \supseteq A} (-1)^{|U| - |T|} \right) f(U) = (-1)^{|A| - |A|} f(A) = f(A).$$

**Corollary 5.1.3.** Let  $f(T) = N_{\supset T}, g(T) = N_{=T}$ . Then,

- 1.  $N_{=A} = \sum_{T \supset A} (-1)^{|T| |A|} N_{\supseteq T}$
- 2. Let  $N_p := \sum_{|A|=p} N_{=A}$ , i.e., # of elements having exactly p of the n properties. Then,

$$N_p = \sum_{|A|=p} \sum_{T \supseteq A} (-1)^{|T|-p} N_{\supseteq T} = \sum_{T:|T|\ge p} (-1)^{|T|-p} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{A:A \subseteq T,|A|=p} 1 = \sum_{k=p}^n (-1)^{k-p} \binom{k}{p} \sum_{T:|T|=k} N_{\supseteq T} \cdot \sum_{T:|T|=k} N_{\Box T} \cdot \sum_{T:|T|=k}$$

If it is homogenoeus case, then

$$N_p = \binom{n}{p} \sum_{k=p}^n (-1)^{k-p} \binom{n-p}{k-p} g_k.$$

3.

$$f(A) = \sum_{T \subseteq A} g(T), \forall A \in 2^E \iff g(A) = \sum_{T \subseteq A} (-1)^{|A| - |T|} f(T).$$

*Proof.* First one is derived from letting  $g(A) = N_{=A}, f(A) = N_{\supseteq A}$ . For the second one, just note that  $\sum_{A:A\subseteq T, |A|=p} 1 = \binom{k}{p}$  if |T| = k, then counting T with respect to |T|. For the third one, do the same thing as we did for proving theorem 5.1.

There are some examples.

1. Let  $D_n = \#$  of derangements  $= \#\{\pi \in \mathfrak{S}n : \pi_i \neq i \text{ for all } i\}$ . Let  $X := \mathfrak{S}n$ , and property i is  $\pi_i = i$ ,  $A_i := \{ \pi \in \mathfrak{S}n : \pi \text{ has property } i \}, \text{ Then,}$ 

$$D_n = \# |\overline{\bigcup_i A_i}|.$$

And we know that

$$|X| = n!, |A_i| = (n-1)!, |A_i \cap A_j| = (n-2)!, \dots |A_{i_1} \cap \dots \cap A_{i_r}| = (n-r)!$$

for generic case, i.e.,  $i_1, \cdots, i_r$  are distinct. Thus, we can define

$$g_k = (n-k)!$$

And by the formula,

$$D_n = \# |\overline{\bigcup_i A_i}| = \sum_{k=0}^n (-1)^k \binom{n}{k} g_k = \sum_{k=0}^n (-1)^k \frac{n!}{k!}$$

Thus,

$$D_n = n! \sum_{k=0}^n \frac{(-1)^k}{k!},$$

and since

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1}$$

we know that  $D_n \sim \frac{n!}{e}$  approximately.

Note that if we fix a permutation  $a_1, \dots, a_n$ , then

$$\#\{\pi \in \mathfrak{S}n : \pi_i \neq a_i\} = D_n.$$

Then,

$$N_p = \#\{\pi \in \mathfrak{S}n : \pi \text{ has exactly } p \text{ fixed pt}\} = \binom{n}{p} \sum_{k=p}^n (-1)^{k-p} \binom{n-p}{k-p} (n-k)!$$
$$= \binom{n}{p} \sum_{k=p}^n (-1)^{k-p} \frac{(n-p)!}{(k-p)!} = \binom{n}{p} (n-p)! \sum_{k=p}^n \frac{(-1)^{k-p}}{(k-p)!}$$

And if we denote k - p = j, then

$$(n-p)!\sum_{k=p}^{n} \frac{(-1)^{k-p}}{(k-p)!} = (n-p)!\sum_{j=0}^{n-p} (-1)^{j} \frac{1}{j!} = D_{n-p}$$

Thus,

$$N_p = \binom{n}{p} D_{n-p}$$

2. Let  $\pi \in \mathfrak{S}n$  with  $\pi = \pi_1 \cdots \pi_n$ . Then

$$DES(\pi) = \{i : \pi_i > \pi_{i+1}\}, \beta(S) = \{\pi \in \mathfrak{S}n : DES(\pi) = S\} \text{ for } S \in 2^{[n-1]}, \alpha(S) = \{\pi \in \mathfrak{S}n : DES(\pi) \subseteq S\}.$$

Then we have

$$\alpha(S) = \sum_{T \subseteq S} \beta(T)$$

and

$$\alpha(S) = \binom{n}{s_1, s_2 - s_1, \cdots, s_k - s_{k-1}, n - s_k} = \binom{n}{\Delta s}$$

which is shown in previous section, where  $s = \{s_1, \dots, s_k\}, s_1 \leq \dots \leq s_k$ . Then, by the corollary version 3 and above, we have

$$\beta(S) = \sum_{T \subseteq S} (-1)^{|S| - |T|} \alpha(T) = \sum_{T = \{s_{i_1}, \cdots, s_{i_j}\} \subseteq S} (-1)^{k-j} \binom{n}{s_1, s_2 - s_1, \cdots, s_k - s_{k-1}, n - s_k}.$$

**Lemma 5.1.4.** Let f be a function defined on  $[0, k+1] \times [0, k+1]$  satisfying  $f_{ij} = f(i, j) = \begin{cases} 1 & i = j \\ 0 & i > j \end{cases}$  and let

$$A_k := \sum_{1 \le i_1 < i_2 < \dots < i_j \le k} (-1)^{k-j} f(0, i_1) f(i_1, i_2) \cdots f(i_j, k+1).$$

Then,

$$\det(F) = A_{I}$$

where F is a matrix with index  $[0, k] \times [0, k]$  such that  $F_{ij} = f(i, j + 1)$ , i.e.,

$$F = \begin{pmatrix} f_{01} & f_{02} & f_{03} & \cdots & f_{0,k+1} \\ f_{11} & f_{12} & f_{13} & \cdots & f_{1,k+1} \\ 0 & f_{22} & f_{23} & \cdots & f_{2,k+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & f_{k,k+1} \end{pmatrix}$$

with  $f_{ii} = 1$ .

*Proof.* Note that empty sequence should be counted. For k = 1,

$$A_k = f(0,1)f(1,2) - f(0,2) = f(0,1)f(1,2) - f(1,1)f(0,2) = \det(F)$$

Now suppose it holds for  $1, 2, \dots, k$ . Then, for k + 1 case, we know that

$$\det F = f_{01} \cdot \det \begin{pmatrix} f_{12} & f_{13} & \cdots & f_{1,k+1} \\ f_{22} & f_{23} & \cdots & f_{2,k+1} \\ 0 & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & f_{k,k+1} \end{pmatrix} - f_{11} \det \begin{pmatrix} f_{02} & f_{03} & \cdots & f_{0,k+1} \\ f_{22} & f_{23} & \cdots & f_{2,k+1} \\ 0 & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & f_{k,k+1} \end{pmatrix}$$
$$= \sum_{\substack{1=i_1 < i_2 < \cdots < i_j \le k+1 \\ 44}} (-1)^{k-j+1} f(0,1) f(1,i_2) \cdots f(i_j,k+1)$$
$$+ \sum_{\substack{2 \le i_1 < i_2 < \cdots < i_j \le k+1 \\ 44}} (-1)^{k-j+1} f(0,i_1) f(i_1,i_2) \cdots f(i_j,k+1)$$

Note that (-1) for first term occurs from determinant since it fixes  $i_1$ , so actually it gives  $(-1)^{k-(j-1)}$ , and for the second term occurs from  $-f_{11}$ .

From this lemma, and by letting  $f(i,j) = \frac{1!}{(s_{j+1}-s_i)!}$  for undefined part, we get

$$\beta(S) = n! \det\left[\frac{1}{(s_{j+1} - s_i)!}\right]$$

with convention  $s_0 = 0, s_{k+1} = n$ .

Also, we can think about q-version of  $\alpha$  and  $\beta$ , i.e.,

$$\beta(S) = \sum_{\substack{\pi \in S_n \\ DES(\pi) = S}} 1 \implies \beta_q(S) = \sum_{\substack{\pi \in S_n \\ DES(\pi) = S}} q^{inv(\pi)}.$$

and

$$\alpha(S) = \sum_{\substack{\pi \in S_n \\ DES(\pi) \subseteq S}} 1 \implies \beta_q(S) = \sum_{\substack{\pi \in S_n \\ DES(\pi) \subseteq S}} q^{inv(\pi)}.$$

Claim 5.1.5.

$$\sum_{\substack{\pi \in S_n \\ DES(\pi) = S}} q^{inv(\pi)} = \binom{n}{s_1, s_2 - s_1, \cdots, s_k - s_{k-1}, n - s_k}_q$$

To see this, we can represent a permutation as a filling of a squares. For example,  $\pi = 4275136$  is represented as below table.

7						*	
6			*				
5	*						
4					*		
3							*
2		*					
1				*			
	1	2	3	4	5	6	7

Now from this we can obtain  $DES(\pi)$ . Note that if  $Des(\pi) \subseteq S$ , then sequence is increase in each interval  $[0, s_1], (s_1, s_2], \cdots, (n - s_k, n]$ .



Note that row denotes index i, and column denotes  $\pi_i$ . Let  $S = \{s_1, \dots, s_k\}$ . Then, from above picture, if  $DES(\pi) \subseteq S$ , then at least  $\pi$  has increasing sequence from i = 1 to  $i = s_1$ . This can be denoted by  $s_1 \times n$  table with  $\pi_1 < \pi_2 < \pi_3$ . For example, if  $S = \{3, 5\}$ , then

3							*
2				*			
1			*				
	1	2	3	4	5	6	7

Now think about inversion occured by this example. For the first column, it has index greater than  $s_1 = 3$ , which gives 3 inversions. By the same argument, the second column gives three inversions, the fifth and sixth column gives one inversions. This can be denoted by gray color below;

3							*
2				*			
1			*				
	1	2	3	4	5	6	7

Note that this inversion must occur regardless of position of  $\pi_i$  with  $i = s_1 + 1$  to i = n. And for given  $\pi$  with fixed  $\pi_1, \dots, \pi_{s_1}$ , the possible permutations can be derived by permutation group isomorphic to  $\mathfrak{S}n - s_1$ , although we need a index coordination. Hence, we can see that

$$\{\pi \in \mathfrak{S}n : DES(\pi) \subseteq S\} \cong \bigcup_{\substack{\pi_f = (\pi_1, \pi_2, \cdots, \pi_{s_1}) \in [n]^{s_1} \\ \pi_1 < \pi_2 < \cdots < \pi_{s_1}}} \pi_f \times T_{\pi_f}$$

where

$$T_{\pi_f} = \{\pi' \in [n]^{n-s_1} : \pi_f + \pi' \text{ gives a permutation in } \mathfrak{S}n\} \cong \{\pi' \in \mathfrak{S}n - s_1\}$$

Note that all isomorphism implies just bijection. This bijection can be proved using the above table; since permutation table has a star at (i, j) if and only if *i*-th row and *j*-th column has unique star at (i, j), so by removing row and columns containing stars corresponds to  $(1, \pi_1), \dots, (s_1, \pi_{s_1})$ , we get a  $(n - s_1) \times (n - s_1)$  table whose row and column has unique star, which gives bijection to  $\mathfrak{S}n - s_1$ .

Note that, as we shown above, for each  $\pi \in \mathfrak{S}n$ , inversion generated first  $s_1$  term depends only on those first  $s_1$  terms, regardless of terms after index  $s_1$ . Hence, we can rewrite the equation as

$$\sum_{\substack{\pi \in \mathfrak{S}n \\ DES(\pi) \subseteq S = \{s_1, \cdots, s_k\}}} q^{inv(\pi)} = \sum_{\substack{\pi \in \mathfrak{S}n \\ \pi_1 < \cdots < \pi_{s_1}}} q^{inv_f(\pi_f) + inv(\pi')}$$

where  $\pi_f, \pi'$  are unique decomposition derived from above bijection, and  $inv_f(\pi_f)$  is the number of inversion of  $\pi$  generated by the first  $s_1$  factors.

Now note that set of possible  $\pi'$  for fixed  $\pi_f$  is isomorphic to  $\mathfrak{S}n - s_1$  as shown above. And this isomophism still preserve  $inv(\pi')$  since this inversion also can be calculated by table which delete  $s_1 \times n$ below part and delete each columns indexed by  $\pi_1, \dots, \pi_{s_1}$ , which actually gives the bijection. Since  $inv(\pi')$ is only determined by factors in  $\pi'$  in  $\mathfrak{S}n - s_1$ , and since this table still give a bijection, we can say that  $inv(\pi')$  is preserved. This argument is clear from picture; for  $\pi$  as below,

7						*	
6			*				
5	*						
4					*		
3							*
2				*			
1		*					
	1	2	3	4	5	6	7

has  $\pi'$ 

7						*	
6			*				
5	*						
4					*		
	1	2	3	4	5	6	7

which is isomorphic to  $3124 \in \mathfrak{S}4$ . So inversion of 3124 is inversion of  $\pi'$ .

7				*
6		*		
5	*			
4			*	
	1	3	5	6

From this, we can say that

$$\sum_{\substack{\pi \in \mathfrak{S}n \\ DES(\pi) \subseteq S = \{s_1, \cdots, s_k\}}} q^{inv(\pi)} = \sum_{\substack{\pi \in \mathfrak{S}n \\ \pi_1 < \cdots < \pi_{s_1}}} q^{inv_f(\pi_f) + inv(\pi')} = \left(\sum_{\substack{\pi \in \mathfrak{S}n \\ \pi_1 < \cdots < \pi_{s_1}}} q^{inv_f(\pi_f)}\right) \left(\sum_{\substack{\pi \in \mathfrak{S}n - s_1 \\ DES(\pi) \subseteq S'}} q^{inv(\pi)}\right)$$

Now it suffices to show that what  $\left(\sum_{\substack{\pi \in \mathfrak{S}n \\ \pi_1 < \cdots < \pi_{s_1}}} q^{inv_f(\pi_f)}\right)$  is. If we calculate it, then by induction, we can apply the same argument for  $s_2 - s_1$  from new S' and  $\mathfrak{S}S_{n-s_1}$ , and inductively apply for  $s_3, \cdots, s_k$ . To see this, I claim that if we project the table

3							*
2				*			
1		*					
	1	2	3	4	5	6	7

onto the x-axis, we get a table

1		*		*			*
	1	2	3	4	5	6	7

This can be turned out into a sequence generated from a multiset  $M = \{1^{s_1}, 2^{n-s_1}\}$  by assigning empty set be 2 and change star to 1. We claim that this is bijective; to see bijection, note that this projection is reversible; since we know  $s_1$ , and \* is ascending order, there exists reverse operation. Also, we claim that inversion of  $\pi$  generated by first  $s_1$  elements is the same as inversion generated by the corresponding sequence; to see this, just observe that each the number of gray cells in a column is exactly same as the number of stars right to the column. And by the proposition 1.7.1 in [Sta11], we have

$$\left(\sum_{\substack{\pi \in \mathfrak{S}n \\ \pi_1 < \dots < \pi_{s_1}}} q^{inv_f(\pi_f)}\right) = \binom{n}{s_1}_q.$$

Now by inductively apply this result, we get

$$\sum_{\substack{\pi \in \mathfrak{S}n \\ DES(\pi) \subseteq S}} q^{inv(\pi)} = \binom{n}{s_1}_q \left( \sum_{\substack{\pi \in \mathfrak{S}n - s_1 \\ DES(\pi) \subseteq S'}} q^{inv(\pi)} \right)$$
$$= \binom{n}{s_1}_q \cdot \binom{n - s_1}{s_2 - s_1}_q \cdots \binom{n - s_1 - s_2 - \dots - s_k}{n - s_k}_q$$
$$= \binom{n}{s_1, s_2 - s_1, \cdots, n - s_k}_q.$$

### 5.2 Möbius inversion

**Definition 5.2.1.** A set P with a relation  $\leq$  on P is called partially ordered set if the relation is

- 1. reflexive, i.e.,  $\forall a \in P, a \leq a$
- 2. antisymmetry, i.e.,  $\forall a, b \in P, a \leq b, b \leq a \implies a = b$
- 3. transitive, i.e.,  $\forall a, b, c \in P, a \leq b, b \leq c \implies a \leq c$ .

Note that this gives us a diagram, called *Hasse diagram*, i.e., a diagram drawn by all pair of covers with line goes upward, i.e., if x < y and there is no  $z \in P$  such that x < z < y, then we call x is covered by y, and draw a line from x vertex to y vertex as upward.

For example,

- 1. Let P be a set of real #,  $\leq$  be a numerical oder. Then,  $[n], \mathbb{Z}, \mathbb{N}$  is called a *chain*, i.e., set with linear ordering. (Every two elements a, b has  $a \leq b$  or  $b \leq a$ .)
- 2. Let P be a power set of a set S, and  $\leq$  be an inclusion. We just call this case as ordered by inclusion. For example, let  $\mathbb{B}_n = \{S \subseteq [n]\}$ . Then  $S \leq_{\mathbb{B}} T$  if and only if  $S \subseteq T$ . If n = 2 we can draw the Hasse diagram



3. Let  $D_n = \{i : i \text{ is positive factor of } n\}$ . And let  $i \leq j \iff i | j$ . Then, if n = 12,  $D_{12} = \{1, 2, 3, 4, 6, 12\}$  and the Hasse diagram of  $D_n$  is



4.



Define  $\leq$  as refinement, i.e.,

 $\pi \leq \sigma \iff$  each block of  $\sigma$  is a union of blocks in  $\pi$ .

For example, if  $\sigma = 1256/347$  and  $\pi = 12/34/56/7$  then  $\pi \leq \sigma$ . If n = 3, it gives



There are several properties.

- (a) y covers x if  $x \leq y$  and there is no z such that  $x \leq z \leq y$ .
- (b) *interval* [x, y] is defined by

$$[x,y] = \{z \in P : x \le z \le y\}$$

Thus,  $x, y \in [x, y]$ . And  $\emptyset$  is not interval.

- (c) P is called *locally finite* if every interval of two elements in P is finite. For example,  $\mathbb{N}, \mathbb{Z}$  are locally finite.
- (d) Hasse diagram is a graph having vertex set as P and edge xy when y covers x.
- (e) minimum  $\hat{0}$  is an element in P s.t.  $\hat{0} \leq x$  for all  $x \in P$ .
- (f) maximum  $\hat{1}$  is an element in P s.t.  $x \leq \hat{1}$  for all  $x \in P$ .

Note that  $\hat{0}$ ,  $\hat{1}$  may not exists; for example, see below Hasse diagram.



In this case, x, y are minimal elements, z, w are maximal elements, but there is no  $\hat{0}, \hat{1}$ .

**Definition 5.2.2** (Incidence algebra). Let  $(P, \leq)$  be a locally finite poset. Then, let Int(P) is defined as set of all intervals of P, i.e.,

$$Int(P) := \{ [x, y] : x \le y, x, y \in P \}.$$

Let k be a field of characteristic 0. Then the incidence algebra I(P,k) consists of all functions  $f: Int(P) \to k$  with multiplication

$$f * g(x, y) = \sum_{z \in [x, y]} f(x, z)g(z, y).$$

Note that

- \* is not commutative in general, but associative.
- $\delta(x,y) := \begin{cases} 1 & x = y \\ 0 & \text{o.w.} \end{cases}$  is the two-sided identity.
- f can be viewed as a formal expression

$$f = \sum_{x \le y} a_{xy}[x, y]$$

so that  $f(x, y) = a_{xy}$  for any  $[x, y] \in Int(P)$ . In this case, we can define f \* g for formal expression by defining multiplication on Int(P) such that

$$[x, y] * [w, z] = \delta_{yw}[x, z].$$

• If P is finite or countable, we can denote f with matrix form; list elements of P as

$$x_1, x_2, \cdots,$$

such that if  $x_i \leq_P x_j$  then i < j. This is called *linear extension of* P. Then, we can make  $M_f$  with size  $|P| \times |P|$  where

$$(M_f)_{x,y} = \begin{cases} f(x,y) & \text{if } x \le y \\ 0 & \text{o.w.} \end{cases}$$

This gives  $f * g = M_f M_g$ , and  $M_f$  is upper triangular. Also

 $M_f$  is invertible  $\iff f(x, x) \neq 0 \forall x \in P.$
• There exists a special element in I(P, k), which is called zeta function  $\zeta$ ,

$$\zeta(x,y) = \begin{cases} 1 & x \le y \\ 0 & x \ne y \end{cases}$$

Note that

$$\begin{split} \zeta^2(x,y) &= \zeta * \zeta(x,y) = \sum_{x \leq t \leq y} 1 = |[x,y]| \\ \zeta^k(x,y) &= \sum_{x = t_0 \leq t_1 \leq \dots \leq t_k = y} 1 \end{split}$$

Thus,

$$(\zeta - \delta)^k(x, y) = \sum_{x=t_0 < t_1 < \dots < t_k = y} 1$$

i.e., the number of chains  $x = t_0 < t_1 < \cdots < t_k = y$  with length k from x to y. Then,

$$\sum_{k=0}^{\infty} (\zeta - \delta)^k = \frac{1}{\delta - (\zeta - \delta)} = \frac{1}{2\delta - \zeta}$$

since  $\delta$  is identity in this algebra. Thus,  $(2\delta - \zeta)^{-1}(x, y)$  counts the total number of chains from x to y.

(Note that to say the above sum rigorously, we should define a topology by saying that  $f_1, f_2, \cdots$  converges to f if for any  $t \in P$  and  $s \leq t$ ,  $\exists N = N(s,t) \in \mathbb{N}$  such that  $f_n(s,t) = f(t)$  for any  $n \geq N$ . This definition of convergence gives a closure of each set, thus we can generate the topology. And in this topology the above infinite series should converges )

**Definition 5.2.3** (Möbius function). Let  $\mu = \zeta^{-1}$ . This is the Möbius function of P, i.e.,

$$\mu * \zeta = \zeta * \mu = \delta.$$

Explicitly, we can define

$$\mu(x,y) := \begin{cases} 1 & x = y \\ -\sum_{x \le t < y} \mu(x,t) & \text{for all } x < y \text{ in } P. \end{cases}$$

To see the definition is true, we need a left inverse;

#### Lemma 5.2.4.

$$f_l^{-1}(x,y) := \begin{cases} \frac{1}{f(x,y)} & \text{if } x = y\\ \frac{1}{f(y,y)} \left( -\sum_{x \le z < y} f^{-1}(x,z) f(z,y) \right) \end{cases}$$

is a left inverse of  $f \in I(P,k)$  having property that  $f(x,x) \neq 0$  for all  $x \in P$ . Similarly,

$$f_r^{-1}(x,y) := \begin{cases} \frac{1}{f(x,y)} & \text{if } x = y\\ \frac{1}{f(x,x)} \left( -\sum_{x < z \le y} f(x,z) f^{-1}(z,y) \right) \end{cases}$$

is the right inverse, and  $f_l^{-1} = f_r^{-1}$ .

Note that this definition is inductively well-defined.

*Proof.* Notes that this is just rearranging the equation  $\sum_{x \le z \le y} f^{-1}(x, z) f(z, y) = 0$ . Similarly,

$$f_r^{-1}(x,y) := \begin{cases} \frac{1}{f(x,y)} & \text{if } x = y\\ \frac{1}{f(y,y)} \left( -\sum_{x \le z < y} f(x,z) f^{-1}(z,y) \right) \end{cases}$$

is just rearranging of the equation.  $\sum_{x \le z \le y} f^{-1}(x, z) f(z, y) = 0$ . Hence  $f * f_r^{-1} = f_l^{-1} * f = \delta$  therefore

$$f_r^{-1} = f_r^{-1} * \delta = f_r^{-1} * f * f_l^{-1} = \delta * f_l^{-1} = f_l^{-1}.$$

Then definition of explicit form of  $\mu$  is derived straightforward; just put  $f = \zeta$  and think  $\zeta(x, y) = 1$  if  $x \leq y$ .

1. For the power set of [3], we have Möbius function with values as below;



$\mu$	1/2/3/	12/3	13/2	23/1	123
1/2/3	1	-1	-1	-1	2
12/3		1			-1
13/2			1		-1
23/1				1	-1
1/2/3					1

2.

3. For simple poset [2] with 1 < 2, we know  $\mu(1,1) = 1 = \mu(2,2)$  But  $\mu * \zeta(1,2) = 0$  implies  $\mu(1,2) = -1$ . For [3], we have

$$\mu(x,x) = 1, \mu(x-1,x) = -1, \mu(x-2,x) = 0$$

since

$$\mu(1,3) = -\mu(1,1) - \mu(1,2)$$

by the explicit definition. Thus, for [n], we know that

$$\mu(x,x) = 1, \mu(x-1,x) = -1, \mu(x,y) = 0$$
 if  $|x-y| > 1$ .

since for the last case, y = k + x for some k > 1, hence

$$\mu(x, y) = -\mu(x, x) - \mu(x, x+1) - \dots - \mu(x, y-1)$$

and inductively  $\mu(x, x + 2) = 0, \mu(x, x + 3) = 0, \dots, \mu(x, y - 1) = 0$  implies  $\mu(x, y) = 0$ .

**Definition 5.2.5.** An order ideal I of a poset P is a subset  $I \subseteq P$  such that if  $t \in I, s \leq t \implies s \in I$ . Similarly, a principal order ideal  $\langle t \rangle$  of a poset P is an ideal generated by t; i.e., t is the unique maximal element of an Ideal.

**Theorem 5.2.6** (Möbius Inversion Formula). Assume P be a poset such that every principal order ideal is finite. Then,  $\forall f, g : P \to k \in$ ,

$$g(x) = \sum_{y \leq x} f(y), \forall x \in P \iff f(x) = \sum_{y \leq x} g(y) \mu(y, x).$$

*Proof.* Note that  $k^P$ , a space of functions from P to k is a vector space. We claim that I(P,k) is a subset of  $End(k^P)$ , a set of all linear transformation from  $k^P$  to itself by defining right action as below; forall  $\pi \in I(P,k)$ ,

$$f.\pi(t) = \sum_{s \le t} f(s)\pi(s,t).$$

To see this, note that standard basis of  $k^P$  is  $\{e_x\}_{x\in P}$  such that  $e_x(y) = \begin{cases} 1 & x = y \\ 0 & \text{o.w.} \end{cases}$ that action of I(P,k) is well-defined and linear. Let  $\pi \in I(P,k)$ . Then, as we defined above,  $f.\pi$  is a still function in  $k^P$ . Also  $f.\delta(t) = f(t)$  and

$$\begin{split} f.\pi * \psi &= \sum_{s \le t} f(s)\pi * \xi(s,t) = \sum_{s \le t} f(s) \sum_{s \le z \le t} \pi(s,z)\xi(z,t) = \sum_{z \le t} \sum_{s \le z} f(s)\pi(s,z)\xi(z,t) \\ &= \sum_{z \le t} \left( \sum_{s \le z} f(s)\pi(s,z) \right) \xi(z,t) = ((f.\pi).\xi)(t). \end{split}$$

where the change of indices of the double sum is well-defined from the finite sum property. And for any  $f, g \in k^P$ ,

$$(f+g).\pi(t) = \sum_{s \le t} (f(s) + g(s))\pi(s,t) = f.\pi(t) + g.\pi(t) \text{ and } cf.\pi(t) = c \sum_{s \le t} f(s)\pi(s,t)$$

from finite sum's linerity.

Then our desired statement is equivalent to saying that

$$f.\zeta = g \iff f = g.\mu.$$

This is easily shown since

$$f.\zeta = g \iff (f.\zeta).\mu = g.\mu \iff f.(\zeta\mu) = g.\mu \iff f.\delta = g.\mu \iff f = g.\mu.$$

See some examples.

1. For [n], we already know that

$$\mu(i,j) = \begin{cases} 1 & i = j \\ -1 & i+1 = j \\ 0 & \text{o.w.} \end{cases}$$

Now the above theorem tells us that

$$g(i) = \sum_{k \le i} f(k) = \sum_{k=0}^{i} f(k) \iff f(i) = g(i) - g(i-1).$$

2. Assume P, Q poset. Then we can define product of two poset as a poset such that

$$P \times Q := \{(x,y) : x \in P, y \in Q\} \text{ and } (x,y) \le (x',y') \iff x \le x' \text{ in } P, y \le y' \text{ in } Q.$$

Now let  $\mu_P, \mu_Q$  be Möbius function of P and Q. Then, we claim that

$$\mu_{P \times Q} = \mu_P \cdot \mu_Q$$
, i.e., $\mu_{P \times Q}((x, y), (x', y')) = \mu_P(x, x')\mu_Q(y, y')$ .

To see this, note that

$$\mu_{P \times Q}((x, y), (x, y)) = 1 = \mu_P(x, x) \cdot \mu_Q(y, y)$$

and by interchanging index of finite double sum we have

$$\sum_{(x,y) \le (u,v) \le (x',y')} \mu_P(x,u) \mu_Q(y,v) = \left(\sum_{x \le u < x'} \mu_P(x,u)\right) \left(\sum_{y \le v < y'} \mu_Q(y,v)\right) = \delta_{x,x'} \delta_{y,y'} = \delta_{(x,y),(x',y')}$$

where  $\delta$  is kronerker delta function. Now, by inductively, if we assume we proved the desired statements for every pair of elements less than (x', y'), then

$$\begin{aligned} \mu_{P \times Q}((x, y), (x', y')) &= -\sum_{(x, y) \le (u, v) < (x', y')} \mu_{P \times Q}((x, y), (u, v)) = -\sum_{(x, y) \le (u, v) < (x', y')} \mu_{P}(x, u) \mu_{Q}(y, v) \\ &= -\delta_{(x, y), (x', y')} + \mu_{P}(x, x') \mu_{Q}(y, y') = \mu_{P}(x, x') \mu_{Q}(y, y') \end{aligned}$$

since (x, y) < (x', y'). Now if  $P \cap \operatorname{pro}[n]$  [m]

Now if P, Q are  $[n], [m] \subset \mathbb{N}$ , then

$$\mu_{P \times Q}((x, y), (x', y')) = \begin{cases} 1 & x = x', y = y' \text{ or } x' = x + 1, y = y + 1 \\ -1 & x = x', y' = y + 1 \text{ or } x' = x + 1, y' = y \\ 0 & \text{ o.w.} \end{cases}$$

These are trivial by above claim.

3. Let  $B_n = 2^{[n]}$ , i.e., power set of [n]. Then each subset S of [n] corresponds to positive indicator  $(\epsilon_1, \dots, \epsilon_n)_S \in \{0, 1\}^n$  where  $\epsilon_i = \begin{cases} 1 & i \in S \\ 0 & i \notin S \end{cases}$  for all  $i \in [n]$ . Also,  $S \subseteq T$  if and only if  $(s_1, \dots, s_n)_S \leq (t_1, \dots, t_n)_T$  coordinatewise. So we just think  $B_n \cong [2] \times \dots \times [2]$  with n copies. Now from the coordinatewise ordering we have Möbius function  $\mu_{B_n}$  and it has a property that

$$\mu_{B_n}(S,T) = \mu_{[2]}(s_1,t_1)\cdots\mu_{[2]}(s_n,t_n)$$

since coordinatewise ordering is just prouduct of posets [2] with natural ordering. And we know that

$$\mu_{[2]}(s,t) = \begin{cases} 1 & s = t \\ -1 & s < t \end{cases}$$

thus,

$$\mu_{B_n}(S,T) = \mu_{[2]}(s_1,t_1)\cdots\mu_{[2]}(s_n,t_n) = (-1)^{|T-S|}.$$

Hence, by the Möbius inversion formula,

$$g(T) = \sum_{S \subseteq T} f(S) \iff f(T) = \sum_{S \subseteq T} g(S)(-1)^{|T-S|}.$$

4. Let  $D_n = \{a : a \in \mathbb{P}, a | n\}$  where  $\mathbb{P} = \mathbb{N}$  without zero. And order the set by divisibility, i.e.,

$$a \leq b \iff a \mid b.$$

Let  $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ . Then, elements in  $D_n$  is

$$a = p_1^{a_1} \cdots p_k^{a_k}, b = p_1^{b_1} \cdots p_k^{b_k}$$

with  $0 \leq a_i, b_i \leq \alpha_i$  for all  $i \in [k]$ . So,

$$a \leq b \iff a \nmid b \iff a_i \leq b_i, \forall i \in [k].$$

So  $D_n \cong [\alpha_1 + 1] \times \cdots \times [\alpha_k + 1]$  with usual ordering. Hence,

$$\mu(a,b) = \begin{cases} 0 & \text{unless } b/a \text{ is sqaurefree} \\ (-1)^r & b/a = p_{i_1} \cdots p_{i_r} \text{ for distinct prime} \\ 1 & b = a \end{cases}$$

In the number theory, just let  $\mu(b/a) := \mu(a, b)$  as abuse of notation. In this notation, for any integer  $m \in D_n$ ,

$$\mu(m) = \begin{cases} 0 & \text{if } p^2 | m \\ (-1)^r & \text{if } m = p_{i_1} \cdots p_{i_r} \text{ for distinct primes} \end{cases}$$

So Möbius inversion formula,

$$g(n) = \sum_{d|n} f(d) \iff f(n) = \sum_{d|n} g(d)\mu(n/d).$$

**Definition 5.2.7** (Lattice). A poset P is a lattice if  $\forall x, y \in P$ ,

- 1. They have the unique least upper bound called  $x \lor y$  (called join.)
- 2. They have the greatest lower bound called  $x \wedge y$  (called meet.)

Note that u is the least upper bound of x and y when every upper bound v of x and y satisfies  $v \ge u$ , and upper bound of x and y is an element u such that  $u \ge x, u \ge y$  simultaneously. (Lower bound is defined in a similar way.)

For example,

1. In [n],  

$$i \lor j = \max(i, j), i \land j = \min(i, j)$$
  
2. In  $B_n$ ,  
 $S \lor T = S \cup T, S \land T = S \cap T$   
3. In  $D_n$ ,  
 $a \lor b = lcm(a, b), a \land b = gcd(a, b)$ 

4. In  $\mathscr{L}(V_n)$ , a set of all subspaces of  $V_n$  ordered by inclusion,

$$S \lor T = \operatorname{span}(S \cup V), S \land T = S \cap T.$$

Counter example is below which we see previously. Note that  $\hat{0}, \hat{1}$  may not exists; for example, see below Hasse diagram.



In this case, x, y are minimal elements, z, w are maximal elements, but there is no  $z \lor w$  or  $x \land y$ .

**Theorem 5.2.8** (Weisner's theorem, 3.9.3 in [Sta11]). For a finite lattice L, note that  $\hat{0}$ ,  $\hat{1}$  exists by taking join or meet of every elements. If L has at least 2 elements, let  $a \neq \hat{1}$ . Then,

$$\sum_{x \wedge a = \hat{0}} \mu(x, \hat{1}) = 0$$

and it gives a recurrence relation

$$\mu(\hat{0},\hat{1}) = -\sum_{\substack{x:x\neq\hat{0}\\x\wedge a=\hat{0}}}\mu(x,\hat{1}).$$

Similarly, it has a dual form; for  $a \neq \hat{1}$ ,

$$\sum_{x\vee a=\hat{1}}\mu(\hat{0},x)=0$$

and it gives a recurrence relation

$$\mu(\hat{0},\hat{1}) = -\sum_{\substack{x:x\neq\hat{1}\\x\vee a=\hat{1}}} \mu(\hat{0},x).$$

To prove the Weisner's theorem, we need a kind of construction in [Sta11][p. 314]. For any lattice L, let A(L,k) be the semigroup algebra of L with the meet operation over k, i.e., A(L,k) is a k-vector space with basis L, with bilinear multiplication  $s \cdot t = s \wedge t$ . We call A(L,k) as *Möbius algebra*.

Now for each  $t \in L$ , define

$$\delta_t = \sum_{s \le t} \mu(s, t) s.$$

If we regard  $\delta_t$  as a function of  $t \in L$ , and s in the above sum as identity function, then by the Möbius inversion theorem,

$$t = \sum_{s \le t} \delta_s.$$

Note that  $\{\delta_t\}_{t \in L}$  is a k-basis of A(L, k) since it is linearly independent (by inductive construction  $\sum_{s < t} \mu(s, t) s$ 

Then, let  $A'(L,k) = \bigoplus_{t \in L} k_t$  where each  $k_t \cong k$  with multiplication defined as follow; for each identity elements  $\delta'_t \in k_t$ ,

$$\delta_t'\delta_s' = \delta_{ts}\delta_t'.$$

where  $\delta_{ts}$  is kronecker delta function. Let  $\theta : A(L,k) \to A'(L,k)$  by  $\theta(\delta_t) = \delta'_t$ 

Claim 5.2.9.  $\theta$  is isomorphism of algebra.

*Proof.* First of all, it is clear that  $\theta$  is a vector space isomorphism, since it maps basis to basis. So it suffices to show that it is multiplicative homomorphism. Let  $t' = \sum_{s \leq t} \delta'_s \in A'(L,k)$ . Then,

$$\theta(s \wedge t) = \theta\left(\sum_{u \le s, u \le t} \delta_u\right) = \sum_{u \le s, u \le t} \delta'_u = \left(\sum_{u \le s} \delta'_u\right) \left(\sum_{u \le t} \delta'_u\right) = \theta(s)\theta(t).$$

Now we can prove the Weisner's theorem.

Proof of the Weisner's theorem. First of all, note that

$$a'\delta'_{\hat{1}} = \left(\sum_{t \le a} \delta'_t\right)\delta'_{\hat{1}} = \sum_{t \le a} \delta_{t,\hat{1}}\delta'_t = \begin{cases} 0 & \text{if } a' \neq \hat{1}'\\\delta'_{\hat{1}} & a' = \hat{1}' \end{cases}.$$

Thus, by the isomorphism  $\theta$ ,

$$a\delta_{\hat{1}} = \begin{cases} 0 & \text{if } a \neq \hat{1} \\ \delta_{\hat{1}} & \text{if } a = \hat{1} \end{cases}.$$

From this, if we denote  $a\delta_{\hat{1}} = \sum_{t \in L} c_t \cdot t$ , then  $c_{\hat{0}} = 0$ . Also, from the definition of  $\delta_{\hat{1}}$ ,

$$a\delta_{\hat{1}} = a\sum_{t\in L}\mu(t,\hat{1})t = \sum_{t\in L}\mu(t,\hat{1})a\wedge t.$$

Also, if we denote  $a\delta_{\hat{1}} = \sum_{t \in L} c_t \cdot t$ , then  $c_{\hat{0}} = \sum_{t \wedge a = \hat{0}} \mu(t, \hat{1})$ . Dual form is derived by the algebra generated by Möbius inversion formular for dual form, which is also have an isomorphism with A'(L,k).

Recurrence relation is derived by just moving  $\mu(\hat{0}, \hat{1})$  term in the right side.

There are some examples.

1. Let  $V_n$  be a vector space over  $F_n$ ,  $L(V_n) = \{W \subseteq V_n : W \text{ is a subspace of } V_n \text{ over } F_n\}$  ordered by inclusion. Then,

$$[u, w] \cong L(w/u), \text{ by } u \mapsto 0, w \mapsto 1.$$

So to calculate  $\mu(u, w)$ , it suffices to calculate  $\mu_m = \mu(\hat{0}, \hat{1})$  in  $L(V_m)$  for suitable m, where m = $\dim w - \dim u$ . Also note that  $\mu_0 = 1$ . To calculate  $\mu_n$  in general, let a be a 1-dimensional subspace of  $V_n$ . Let  $X \in L(V_n)$  such that  $span(X, a) = V_n$  with  $X \neq V_n$ . Then, X is a (n-1)-dimensional subspace and  $a \notin X$ . So by above argument, for all such X,

$$\mu(0,x) = \mu_{n-1}.$$

And the number of all (whether it contain a or not) (n-1) dimensional subspace of  $V_n$  is  $\binom{n}{n-1}_a$ , which we showed in the previous section. By the same argument, for fixed X, the number of subspaces of dimension n-1 containing a is

$$#\{X = span(x', a) : x' \text{ is } n-2 \text{ dimensional subspaces of } V_n\} = \binom{n-1}{n-2}_q.$$

To see this, we need to recall the argument of proving

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

Let W be n-1 dimensional vector space doesn't containing a. Then, for any U with k = n-2dimensional subspace which doesn't contained in W, U has a basis as union of a basis of  $U \cap W$  and  $\{f\}$  where  $f \notin W$ . Then, our desired number is such U having a as its basis. Such U is just counted by setting f = a and count all possible number of a basis  $U \cap W$  generating distinct vector space. Then, if we just thinking W as a whole space, the number of possible distinct n-2 subspace  $U \cap W$ is  $\binom{n-1}{n-2}_a$ . So

$$#\{X = span(x', a) : x' \text{ is } n-2 \text{ dimensional subspaces of } V_n\} = \binom{n-1}{n-2}_q.$$

Also,  $\binom{n-1}{k}_{q}$  represents (n-2)-dimensional subspaces which is contained in W, therefore it doesn't contain a so we only care about the first case, done.

Hence the number of desired X is

$$\binom{n}{n-1}_q - \binom{n-1}{n-2}_q = \binom{n}{1}_q - \binom{n-1}{1}_q = [n]_q - [n-1]_q = (1+q+\dots+q^{n-1}) - (1+q+\dots+q^{n-2}) = q^{n-1}$$

Thus,

$$\mu_n = -q^{n-1}\mu_{n-1}$$

since  $\mu_n$  counts only case of X having (n-1)-dimension and having no a, and each such space has the same Möbius function value  $\mu_{n-1}$ . Also we know  $\mu_0 = 1$ . Hence by induction,

$$\mu_n = (-1)^n q^{\binom{n}{2}}.$$

2.  $\Pi_n := \{ \text{ all set of partitions of } [n] \}$  ordered by refinement, i.e.  $\pi \leq \sigma$  iff each block of  $\sigma$  is a union of blocks of  $\pi$ . Thus,

$$\hat{0} = 1/2/\cdots/n, (n\text{-blocks})\hat{1} = 123\cdots n, (1\text{-block.})$$

If  $\pi \leq \sigma$ , then  $\sigma$  covers  $\pi$  means  $\pi$  can be created by choosing one block in  $\sigma$  and chop it to make it two blocks. Thus we can make a grade such that

$$rank(\pi) := n - \#$$
 blocks in  $\pi$ .

Note that the number of elements in  $\pi_n$  of rank k is  $S_{n,n-k}$ , by definition of the stirling number of second kind. Now we can find out what the meet and join in this poset. For example, if

$$\sigma = \frac{125}{3}/\frac{47}{68}, \pi = \frac{124}{3}/\frac{57}{6}/\frac{8}{8}$$

then

$$\sigma \wedge \pi = \frac{12}{3} \frac{4}{5} \frac{6}{7} \frac{8}{8}$$

since each blocks are determined by

$$\{\pi_i \cap \sigma_j \neq \emptyset : \text{ each } \pi_i, \sigma_j \text{ is block of } \pi, \sigma \text{ respectively} \}.$$

And,

$$\sigma \lor \pi = 12547/3/68.$$

This is determined by this simple rule; let  $\pi_i, \sigma_j$  be a two block. Then think  $\pi_i \Delta \sigma_j = \pi_i \cup \sigma_j \setminus \pi_i \cap \sigma_j$ , i.e., symmetric difference. If there is another block containing such symmetric difference, then union with those blocks, and find the symmetric difference again, and union the related blocks, and so on. Since each partition is finite, this process must be terminated.

Or geometric way of seeing this  $\pi \lor \sigma$  is represent these as a path on linear graph, i.e. joining two vertices if they are in the same block.



Then  $\sigma \vee \pi$  is just set of all connected component as a block in the above graph.

Now investigate what the interval of this poset is. Let  $\sigma = 1234/5678, \pi = 12/3/4/57/6/8$ . Then,  $\sigma \ge \pi$ , thus

$$[\pi,\sigma] = \left\{ \begin{cases} \{1,2\},\{3\},\{4\} \\ \{1,2,3\},\{4\} \\ \{1,2,4\},\{3\} \\ \{1,2\}\{3,4\} \\ \{1,2,3,4\} \end{cases} \times \left\{ \begin{cases} \{5,7\},\{6\},\{8\} \\ \{5,7,6\},\{8\} \\ \{5,7,8\},\{6\} \\ \{5,7\}\{6,8\} \\ \{5,7,6,8\} \end{cases} \right\} \cong \Pi_3 \times \Pi_3 \end{cases} \right\}$$

where  $\Pi_A$  denotes a set of partitions of A and  $\Pi_n := \Pi_{[n]}$  for all  $n \in \mathbb{N}$ . This is easily generalized; if  $\sigma \ge \pi$  and each block of  $\sigma$ , say  $\sigma_1, \cdots, \sigma_k$  consists of union of  $\pi$ 's block, i.e.,  $\sigma_j = \{\pi_{j1}, \cdots, \pi_{jt_j}\}$ , then

$$[\pi,\sigma] \cong [\hat{0},\sigma/\pi] \cong \Pi_{t_1} \times \cdots \times \Pi_{t_k}$$

From this and the product rule of the Möbius function, we have

$$\mu(\pi,\sigma) = \mu_{t_1}(\hat{0},\hat{1}) \times \cdots \times \mu_{t_k}(\hat{0},\hat{1}).$$

Thus to calculate  $\mu(\pi, \sigma)$  for any  $\pi, \sigma$ , it suffices to calculate  $\pi_n(\hat{0}, \hat{1})$  for any n. First of all,

$$\mu_1(\hat{0},\hat{1}) = 1$$

since  $\Pi_1 = \{\{1\}\}, \text{ thus }$ 

$$\mu_1(\hat{0}, \hat{1}) = \mu(\{1\}, \{1\}) = 1$$

For  $\mu_n$ , we can use the Weisner theorem. First of all, let  $a = 123 \cdots (n-1)/n$ . Then all  $t \in \Pi_n$ such that  $a \wedge t = \hat{0}$  is  $t = \hat{0}$  or t consists of  $\{i, n\}$  and singletons for each element in  $[n] \setminus \{i, n\}$  for  $i = 1, \cdots, n$ . To see this, if  $t \neq \hat{0}$  has  $\{n\}$  as a singleton, then t itself is  $a \wedge t$ . If a block containing nhas more than 1 other element, say i, j are in the block containing n with  $i \neq j$ , then  $a\hat{t}$  contains  $\{i, j\}$ . Thus all possible t having  $a \wedge t = \hat{0}$  should contain two element block  $\{i, j\}$ , and the other elements should form a singleton, otherwise  $a \wedge t$  has nonsingleton block. Thus, such possible number of nonzero t is (n-1). And for each such  $t, \mu(t, \hat{1}) = \mu_{n-1}$  since t has n-1 blocks and  $\hat{1}$  is just one block. Hence, by Weisner's theorem,

$$\sum_{t \wedge a=0} \mu(t, \hat{1}) = 0 \implies \mu_n = \mu(\hat{0}, \hat{1}) = -\sum_{\substack{t \neq \hat{0} \\ t \wedge a=0}} \mu(t, \hat{1}) = -(n-1)\mu_{n-1}.$$

Hence, by considering  $\mu_1 = 1$ , we can get

$$\mu_n = (-1)^{n-1}(n-1)!.$$

### 5.3 Involution Principle

Suppose a set S is disjoint union of  $S^+ \cup S^-$ . Then,  $\phi : S \to S$  is an *involution* if for any  $x \in S$  with  $\phi(x) \neq x$ , x and  $\phi(x)$  has different sign, i.e.,  $x \in S^+, \phi(x) \in S^-$  or vice versa. Then, we can divide

$$\phi(S^+) = \phi(S^+) \cap S^+ \cup \phi(S^+) \cap \S^-$$

We call  $\phi(S^+) \cap S^+$  be fixed points in  $S^+$ . Similarly,  $\phi(S^-) \cap S^-$  be fixed points in  $S^-$ . So,

$$|S^+| - |Fix(S^+)| = |S^-| - |Fix(S^{-1})|.$$

We can think of special case when  $Fix(S^{-}) = \emptyset$ . Then,

1. 
$$|S^{-}| \leq |S^{+}|$$
.

2.  $|Fix(\phi)| = |S^+| - |S^-|$ 

from above equation. For example, which is very artificial, note that

$$N_{\emptyset} = \sum_{T} (-1)^{|T|} N_{\geq T} = \sum_{T: \text{ even}} N_{\geq T} - \sum_{T: \text{ odd}} N_{\geq T}$$

from the Sieve method, where we assume that there exists a universe X and n "bad properties," and

$$\begin{split} N_{=A} &= \#\{x \in X : x \text{ has exactly properties in } A\}\\ N_{\geq A} &= \#\{x \in X : x \text{ has properties at least in } A\}\\ N_{\emptyset} &= \#\{x \in X : x \text{ has no bad properties}\}. \end{split}$$

Now set

 $S = \{(x, Z, T) : \text{ where } x \in X, Z \text{ is a set of properties } x \text{ has }, T \text{ is a subset of } Z\}.$ 

Then,  $S^+ := \{(x, Z, T) \in S : |T| \text{ is even}\}, S^- := \{(x, Z, T) \in S : |T| \text{ is odd}\}$ . From this setting, we can denote the above principle of inclusion-exclusion formula  $N_{\emptyset} = \sum_{T: \text{ even}} N_{\geq T} - \sum_{T: \text{ odd}} N_{\geq T}$  as

$$N_{\emptyset} = |S^+| - |S^-|$$

In this case we can define an involution explicitly; let

$$\phi: (x, Z, T) \mapsto (x, Z, T') \text{ where } T' := \begin{cases} T \cup \{i\} & \text{ if } i \notin T \\ T \setminus \{i\} & \text{ if } i \in T \end{cases}.$$

where *i* is the minimal element of Z. (Note that  $Z \subseteq [n]$  since we index bad properties by [n].) Then, only fixed elements is when  $Z = \emptyset$ . Then,  $(x, \emptyset, T) = (x, \emptyset, \emptyset)$ , and they are in  $S^+$ . Thus

$$N_{\emptyset} = |Fix(S^+)| = |S^+| - |S^-| = \sum_{T: \text{ even}} N_{\ge T} - \sum_{T: \text{ odd}} N_{\ge T}$$

There are examples for involution principle.

For example, if n = 3, then

• Ballot Sequence. Let  $n \in \mathbb{N}$ , then a ballot sequence with length 2n is a sequence  $a_1 \cdots a_{2n}$  with  $a_i \in \{\pm 1\}$  such that

$$\sum_{i=1}^{k} a_i \ge 00, \sum_{i=1}^{2n} = a_i = 0.$$

$$\begin{pmatrix} + & + & + & - & - & - \\ + & + & - & + & - & - \\ + & - & + & - & - & + & - \\ + & - & + & - & - & + & - \\ + & - & + & - & + & - & - \end{pmatrix}.$$

Then, we can represent the ballot sequence as a Dyck path, a path from (0,0) to (2n,0) such that each path comprises of  $\nearrow$ ,  $\searrow$  which stands for +, - respectively, and that the path is always above on y = 0. Also, we can represent it as a Catalan path, a path from (0,0) to (n,n) such that each path comprises of  $\rightarrow$ ,  $\uparrow$  which stands for +, - respectively, and that the path is always below than y = x. This condition above and below are from the condition that partial sum of the sequence is greater than or equal to 0.

**Claim 5.3.1.** The number of ballo sequence with length 2n is the same as n-th catalan number  $\frac{1}{n+1}\binom{2n}{n}$ .

Classical proof. We can calculate it by counting catalan paths using involution principle. Note that the number of catalan path is just the number of paths from (0,0) to (n,n) which never touch y = x + 1. To calculate this, we need to count such paths touching y = x + 1.

Now let  $\phi$  be a map from lattice path from (0,0) to (n,n) to a lattice path (0,0) to (n-1, n+1) such that if a path touches y = x + 1, we reflect the last portion of the original path, i.e., a partial path from the last touched point to (n, n) with respect to y = x + 1. For example,



dashed red path is  $\phi$  of straight blue path. Now we can see that there is 1-1 correspondence between a set of lattice paths from (0,0) to (n,n) touching y = x + 1 and a set of lattice paths from (0,0) to

(n-1, n+1). To see this, note that every path from (0,0) to (n-1, n+1) touch y = x + 1, thus  $\phi$  maps it to a path from (0,0) to (n,n). And note that reflection is injective. Thus, from the formula of counting lattice paths, the number of paths from (0,0) to (n,n) touching y = x + 1 is  $\binom{2n}{n+1}$ , thus the number of catalan paths is

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

• We want to prove that

$$\binom{n}{k} \leq \binom{n}{k+1}$$
 if  $k < \lfloor \frac{n-1}{2} \rfloor$ .

Combinatorial proof. Note that  $\binom{n}{k}$  is the number of lattice paths from (0,0) to (n-k,k). Then we can define  $\phi$  in the sense of above definition; reflect the last portion of lattice path from (0,0) to (n-k,k), i.e., a portion from last point of the path touching y = x + 2k - n + 1. Then, if  $k < \lfloor \frac{n-1}{2} \rfloor$ , then this y = x + 2k - n + 1 touches x-axis when x > 0, otherwise it touches y axis when y > 0. Thus,  $k < \lfloor \frac{n-1}{2} \rfloor$ , then every path from (0,0) to (n-k,k) can be injectively mapped into a path from (0,0) to (n-k-1,k+1), which implies  $\binom{n}{k+1} \ge \binom{n}{k}$ .

We want to show

$$\binom{n}{k-1}\binom{n}{k+1} \le \binom{n}{k}^2$$

in a combinatorial way. Note that a sequence  $\{a_i\}$  has *log-concave* if  $a_{i-1}a_{i+1} \leq a_i^2$ . Let  $L_k$  be a set of lattice paths (n-k) east steps and k north steps. Our goal is to construct an injective map

$$L_{k-1} \times L_{k+1} \to L_k \times L_k.$$

From last example, i.e.,  $\binom{n}{k} \leq \binom{n}{k+1}$  if  $k < \lfloor \frac{n-1}{2} \rfloor$ , we know that a sequence  $|L_k| = \binom{n}{k}$  is unimodal, i.e., has a unique peak. And note that log concave implies unimodality, since the relationship  $\frac{a_i}{a_{i-1}} \geq \frac{a_{i+1}}{a_i}$  on ratio shows that until the ratio decrease below 1, the sequence increase, and after this point, the sequence decrease.

To see the injective map, note that we can draw any pair  $(P_1, P_2)$  in  $L_{k-1} \times L_{k+1}$  in a following way; draw  $P_1$  as usual, and draw translated  $P_2$ , starting at (1, -1) and ending at (n-k, k). For example,



Then take a map  $\phi: L_{k-1} \times L_{k+1}$  to  $L_k \times L_k$  by  $(P_1, P_2) \cong (P_1, P_2 + (1, -1)) \mapsto (P'_1, P'_2)$  where

- $P_1' =$  first part of  $P_1$  before last intersection point of  $P_1 \cap P_2$ 
  - + last part of  $P_2$  before last intersection point of  $P_1 \cap P_2$
- $P'_2$  = first part of  $P_2$  before last intersection point of  $P_1 \cap P_2$ 
  - + last part of  $P_1$  before last intersection point of  $P_1 \cap P_2$

Then  $P'_1 \in L_k$ , and  $P'_2$  is a path from (1, -1) to (n - k + 1, k - 1), which is equivalent to a path from (0, 0) to (n - k, k), so  $P'_2 \in L_k$ . So this is well-defined map. And it is injection since distinct pair in  $L_{k-1}, L_{k+1}$  gives different image in the coordinate space, thus gives different image.

Now note that it may not be a surjection, since there exists a non crossing pair of paths  $(Q_1, Q_2) \in L_k \times (L_k + (1, -1))$ , as below.



• Given points  $A, B, C, D \in \mathbb{Z}^2$ , with property that  $\overline{AD}$  and  $\overline{BC}$  meets at one point, the number of noncrossing paths  $(A \to B, C \to D)$  is

$$\det \begin{pmatrix} \#P(A \to B) & \#P(A \to D) \\ \#P(C \to B) & \#P(C \to D) \end{pmatrix},$$

where  $P(A \to B)$  is a set of lattice paths from A to B. To see this, note that a crossing pairs  $(P_1, P_2)$  in  $P(A \to B) \times P(C \to D)$  has canonically mapped into a set of crossing pairs in  $P(A \to D) \times P(C \to B)$ , as we did in previous example; take first part of  $P_1$  and last part of  $P_2$  with respect to their last intersection, and vice versa. Also, from the condition  $\overline{AD}$  and  $\overline{BC}$ , every pair in  $P(A \to D) \times P(C \to B)$  is crossing pair. Hence the number of crossing pairs in  $P(A \to B) \times P(C \to D)$  is just  $\#P(A \to D) \times \#P(C \to B)$ , so the number of noncrossing pairs in  $P(A \to B) \times P(C \to D)$  is

$$\#P(A \to B) \times \#P(C \to D) - \#P(A \to D) \times \#P(C \to B) = \det \begin{pmatrix} \#P(A \to B) & \#P(A \to D) \\ \#P(C \to B) & \#P(C \to D) \end{pmatrix}.$$

Think about ballot sequence. Let A<sub>1</sub>, ..., A<sub>n</sub> be candidates, and they received m<sub>1</sub> ≥ m<sub>2</sub> ≥ ... ≥ m<sub>n</sub> votes from the ballot. Then, how many ways of sequential voting which reveals voting lead, i.e., A<sub>1</sub> leads A<sub>2</sub>, A<sub>2</sub> leads A<sub>3</sub>, ..., A<sub>n-1</sub> leads A<sub>n</sub> in the any middle step of counting vote?

We can generate a combinatorial model for it. We can think it as a lattice path on  $\mathbb{Z}^n$ , starting at  $\vec{0} = (0, \dots, 0) \in \mathbb{Z}^n$  and ends at  $\vec{m} = (m_1, \dots, m_n)$  with allowed step  $e_i \in \mathbb{Z}^n$ , a standard vector such

that  $(e_i)_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$ . Then for each point  $\vec{x} = (x_1, \dots, x_n)$  in this lattice path should satisfy the inequality  $x_1 \ge \dots \ge x_n$ . This is equivalent to say that the lattice path never cross (n-1) hyperplanes  $x_2 = x_1 + 1 \cdots + x_n = x_{n-1} + 1$ . If we say such lattice path is 'good' lattice path and the other lattice

 $x_2 = x_1 + 1, \dots, x_n = x_{n-1} + 1$ . If we say such lattice path is 'good' lattice path and the other lattice paths as 'bad' lattice path, then

$$\#Good(\vec{0} \to \vec{m}) = \#\mathscr{F}(\vec{0} \to \vec{m}) - \#Bad(\vec{0} \to \vec{m})$$

where  $Good(\vec{0} \to \vec{m})$  denotes a set of good lattice paths,  $\mathscr{F}(\vec{0} \to \vec{m})$  denotes a set of all lattice paths, and  $Bad(\vec{0} \to \vec{m})$  denotes a set of bad lattice paths. Actually we know that for any two points  $\vec{A} = (a_1, \dots, a_n)$  and  $\vec{B} = (b_1, \dots, b_n)$ ,

$$#\mathscr{F}(\vec{A} \to \vec{B}) = \begin{pmatrix} \sum_{i=1}^{n} b_i - \sum_{i=1}^{a_i} \\ b_1 - a_1, \cdots, b_n - a_n \end{pmatrix}.$$

since each lattice path is just choosing each step for n directions.

Now the idea for general n case is just do reflection with respect to those n-1 hyperplanes. To do this, we can expand this problem not only counting path from 0 to  $\vec{m}$ , but also couting path from other points to  $\vec{m}$ . Define

$$e_{\pi} := (1 - \pi(1), 2 - \pi(2), \cdots, n - \pi(n)), \forall \pi \in \mathfrak{S}n.$$

Then,  $e_{id} = \vec{0}$ , and  $e_{\pi}$  satisfies  $x_1 \ge x_2 \ge \cdots \ge x_n$  if and only if  $\pi = id$  since, the inequality condition requires that

$$i - \pi(i) \ge i + 1 - \pi(i+1), \forall i$$

so  $\pi(i+1) - \pi(i) \ge 1$ , which is equivalent to say that  $\pi = id$ .

Now reflect the first segment of a path from  $e_{\pi} \to \vec{m}$  with respect to  $x_i - x_{i+1} = -1$ . This means we change  $e_{\pi}$  to  $e'_{\pi}$ , where

$$e_{\pi} = (1 - \pi(1), 2 - \pi(2), \cdots, i - 1 - \pi(i - 1), i - \pi(i), i + 1 - \pi(i + 1), i + 2 - \pi(i + 2), \cdots, n - \pi(n))$$
  
$$e'_{\pi} = (1 - \pi(1), 2 - \pi(2), \cdots, i - 1 - \pi(i - 1), i - \pi(i + 1), i + 1 - \pi(i), i + 2 - \pi(i + 2), \cdots, n - \pi(n))$$

i.e., take a point which has the distance with  $x_i - x_{i+1} = -1$  the same as  $e_{\pi}$ . (You can check that their middle point is in  $x_i - x_{i+1} = -1$  easily.) Then,

$$i - \pi(i+1) = i - \pi'(i), i + 1 - \pi'(i+1) = i + 1 - \pi(i) \implies \pi' = \pi \cdot (i, i+1) \implies sgn(\pi) = -sgn(\pi').$$

Hence,

$$\bigcup_{\substack{\pi \in \mathfrak{S}n \\ sgn(\pi) = \text{ even}}} Bad(e_{\pi} \to \vec{m}) \xleftarrow{1-1} \bigcup_{\substack{\pi \in \mathfrak{S}n \\ sgn(\pi) = \text{ odd}}} Bad(e_{\pi} \to \vec{m})$$

Thus,

$$\sum_{\pi \in \mathfrak{S}n} sgn(\pi) |Bad(e_{\pi} \to \vec{m})| = 0.$$

This implies that

$$|Bad(\vec{0} \to \vec{m})| + \sum_{\pi \in \mathfrak{S}n, \pi \neq id} sgn(\pi)|Bad(e_{\pi} \to \vec{m})| = 0$$

Then we know that

$$|Bad(\vec{0} \rightarrow \vec{m})| = |F(\vec{0} \rightarrow \vec{m})| - |Good(\vec{0} \rightarrow \vec{m})|$$

and

$$Bad(e_{\pi} \to \vec{m}) = F(e_{\pi} \to \vec{m}), \text{ if } \pi \neq id$$

since  $e_{\pi}$  itself changed by reflection implies every path from  $e_{\pi}$  to  $\vec{m}$  must cross the hyperplane. And also note that

$$sgn(\pi) = (-1)^{length(\pi)} = (-1)^{inv(\pi)}$$

where length of  $\pi$  implies the least number of transpositions representing  $\pi$ , which implies  $inv(\pi)$ . Thus,

$$|Good(\vec{0} \to \vec{m})| = \sum_{\pi \in \mathfrak{S}n} sgn(\pi) F(e_{\pi} \to \vec{m}).$$

And we know that from  $\vec{m} = (m_1, \cdots, m_n)$ ,

$$F(e_{\pi} \to \vec{m}) = \begin{pmatrix} \sum m_i - \sum_{i=1}^n (i - \pi(i)) \\ m_1 - 1 + \pi(1), \cdots, m_n - n + \pi(n) \end{pmatrix} = \begin{pmatrix} \sum m_i \\ m_1 - 1 + \pi(1), \cdots, m_n - n + \pi(n) \end{pmatrix}$$

since  $\sum_{i=1}^{n} (i - \pi(i)) = 0$  from the bijectiveness of  $\pi$ . Thus,

$$|Good(\vec{0} \to \vec{m})| = \sum_{\pi \in \mathfrak{S}n} sgn(\pi)F(e_{\pi} \to \vec{m})$$
$$= \left(\sum_{i=1}^{n} m_{i}\right)! \cdot \sum_{\pi \in \mathfrak{S}n} sgn(\pi) \frac{1}{(m_{1} - 1 + \pi(1))! \cdots (m_{n} - n + \pi_{n})!}$$
$$= \left(\sum_{i=1}^{n} m_{i}\right)! \cdot \det\left[\frac{1}{(m_{i} - i + j)!}\right]$$

where the last equality comes from the definition of determinant using uniqueness of antisymmetric function.

## 5.4 The Lemma of Gessel - Viennot

Think about a complete bipartite graph



each edge between  $A_i$  to  $B_j$  has weight  $m_{ij}$ .

**Definition 5.4.1.** A matching is a set of n-disjoint edges from A's to B's, i.e. a system of path  $(P_1, \dots, P_n)$  from A's to B's such that  $P_i$ 's are disjoint paths starting at a point in A's and ending at a point in B's. We say that such path is vertex disjoint paths.

Now let G be an acyclic (i.e., no directed cycles in a graph) direct weighted (finite) graph having  $V(G) = \{A_i, B_i\}_{i=1}^n$ . Then for two set of vertices  $A = \{A_1, \dots, A_n\}$ , and  $B = \{B_1, \dots, B_n\}$ , for a path from a point  $v \in A, w \in B$ , let

$$w(P) = \pi_{e \in P} w(e).$$

For example, if  $P = e_1 \rightarrow e_2 \rightarrow e_3$  then  $w(p) = \pi_{i=1}^3 w(e_i)$ . Now build a  $n \times n$  matrix m such that

$$M_{ij} = \sum_{P:A_i \to B_j} w(P).$$

Then for a given vertex disjoint path system  $(P_1, \dots, P_n)$ , our goal is calculate

$$\det M = \sum_{\substack{\text{VD path system}\\\sigma \in \mathfrak{S}n}} sgn(\sigma) w(P(A_i \to B_{\sigma(i)}))$$

We can summarize this as follow;

**Lemma 5.4.2** (Gessel-Viennot). Let G = (V, E) is a finite acyclic digraph, and each cycle e in E has a weight w(e). Fix two set of vertices  $A = \{A_1, \dots, A_n\} \subset V$  and  $B = \{B_1, \dots, B_n\} \subset V$ . For any path P in G, define

$$w(P) := \prod_{e \in P} w(e).$$

For a path system  $\mathscr{P} = (P_1, \cdots, P_n)$ , define

$$w(\mathscr{P}) = \prod_{i=1}^{n} w(P_i).$$

And let  $\mathscr{P}_{\sigma}$  be a path system such that

 $\mathscr{P}_{\sigma} := (P_1, \cdots, P_n)$  where  $P_i$  is a path from  $A_i$  to  $B_{\sigma(i)}$ .

Now define a  $n \times n$  matrix M where

$$M_{ij} = \sum_{P:A_i \to B_j} w(P).$$

Then,

$$\begin{split} \det(M) &= \sum_{\sigma \in \mathfrak{S}n} sgn(\sigma) M_{1\sigma(1)} \cdots M_{n\sigma(n)} = \sum_{\sigma \in \mathfrak{S}n} sgn(\sigma) \left( \sum_{P:A_1 \to B_{\sigma_1}} w(P) \right) \cdots \left( \sum_{P:A_n \to B_{\sigma_n}} w(P) \right) \\ &= \sum_{\sigma \in \mathfrak{S}n} sgn(\mathscr{P}_{\sigma}) w(\mathscr{P}_{\sigma}) \end{split}$$

by defining  $sgn(\mathscr{P}_{\sigma}) = sgn(\sigma)$ .

Then the lemma conclude that

$$\det(M) = \sum_{\mathscr{P}_{\sigma}: \ vertex \ disjoint} \ sgn(\mathscr{P}_{\sigma})w(\mathscr{P}_{\sigma}).$$

where vertex disjoint means that if no two path has a vertex in common.

Basic idea of showing this lemma is in below picture;



If the given path system contains a pair of paths crossing to each other, for example,  $(A_1 \rightarrow B_1, A_2 \rightarrow B_2)$ , then this can be canceled by a path system containing a pair of paths  $(A_1 \rightarrow B_2, A_2 \rightarrow B_1)$  since they are transposition relation, so their sign should be negative.

*Proof.* Just need to show that

$$\sum_{\mathscr{P}_{\sigma}: \text{ not vertex disjoint}} sgn(\mathscr{P}_{\sigma})w(\mathscr{P}_{\sigma}) = 0.$$

To achieve this, we define an involution

$$\mathscr{P}_{\sigma} \to \mathscr{P}_{\sigma'}$$

both are not vertex disjoint such that

$$w(\mathscr{P}_{\sigma}) = w(\mathscr{P}_{\sigma'})$$
 and  $sgn(\sigma) = -sgn(\sigma')$ .

Now describe the involution we want to get.

- 1. Find the smallest i such that  $P_i$  intersection other paths.
- 2. Find the 1st intersection point z on  $P_i$ .
- 3. Let j be the smallest index > i such that  $P_i$  passes z.

Now define the operation as swapping the part of  $P_i$  and  $P_j$ , i.e.,

$$P'_i = P_i(A_i \to z) + P_j(z \to B_{\sigma(j)}), P'_j = P_j(A_j \to z) + P_i(z \to \sigma(i))$$

where + is concatenation in the above. Then this operation is clearly involution, since applying twice gives identity, and send every  $P_{\sigma}$  with  $sgn(\sigma) = 1$  to  $\sigma' = \sigma \cdot (i, j)$ , which is odd sign. And this involution only change the sign. Since such involutive image always exists for nonvertex disjoint paths, the sum of weights all nonvertex disjoint paths are zero. done.

There are quick aplications.

1.  $\det(M) = \det(M^t)$ . To prove this, let G be a complete bipartite graph from  $(A_1, \dots, A_n)$  to  $(B_1, \dots, B_n)$  with  $w(A_i, B_j) = m_{ij}$ . Then,

$$\det M = \sum_{\sigma \in \mathfrak{S}n} sgn(\mathscr{P}_{\sigma})w(\mathscr{P}_{\sigma})$$

For each  $\sigma$ , there is only one  $\mathscr{P}_{\sigma}$  consisting edges (so it is only one vertex-disjoint system.)

Similarly, det $(M^t)$  is the same thing for path systems from B to A. And for each path  $P_{\sigma}$  from A to B, there exists a path  $P_{\sigma^{-1}}$  from B to A, such that if  $P_{\sigma}$  send  $A_i$  to  $B_{\sigma(i)}$ , then  $P_{\sigma'}$  send  $B_i$  to  $B_{\sigma^{-1}(i)}$ . And they have the same sign since  $\sigma$  and  $\sigma^{-1}$  has the same sign in  $\mathfrak{S}n$ . So, two determinants are the same.

2.  $\det(MM') = \det(M) \cdot \det(M')$  (Note that M' is not a transpose of M but any matrix having same dimension with M.) To see this let G be a complete bipartite graph with  $V = \{A_i, B_i\}_{i=1}^n$  with resepct to weight  $M_{ij}$  and G' be a complete bipartite graph with  $V = \{B_i, C_i\}_{i=1}^n$  with resepct to weight  $M'_{ij}$ . Then,

$$(MM')_{ij} = \sum_{k} M_{ik}M'_{kj} = \sum_{P:A_i \to C_j} w(P)$$

where  $P: A_i \to C_j$  is a path from  $A_i$  to  $C_j$  by concatenating two graph G, G'. By the lemma, we have

$$\det(MM') = \sum_{\substack{\mathscr{P}_{\sigma} \text{ vertex disjoint}\\\text{from } A \text{ to } C}} sgn(\sigma)w(\mathscr{P}_{\sigma}).$$

And from the complete bipartitiness, each vertex disjoint path system comprises by paths having only two edge;  $A \to B \to C$ , so we can decompose this path by  $P_{\sigma_1} : A \to B$  and  $P_{\sigma_2} : B \to C$ , both are unique vertex disjoint edges of  $\sigma_1, \sigma_2$ , respectively. Thus, if we denote  $\mathscr{P}_{\sigma}$  as a path system consists of  $A \to C$  with  $A_i$  to  $C_i$  (note that we abuse the notation.) then,

$$\det(MM') = \sum_{\sigma \in \mathfrak{S}n} sgn(\sigma)w(\mathscr{P}_{\sigma}) = \sum_{\sigma \in \mathfrak{S}n} \sum_{\substack{(\sigma_1, \sigma_2) \in \mathfrak{S}n^2 \\ \sigma_2 \sigma_1 = \sigma}} sgn(\sigma_1)sgn(\sigma_2)w(\mathscr{P}_{A \to B, \sigma_1})w(\mathscr{P}_{B \to C, \sigma_2})$$
$$= \left(\sum_{\sigma_1 \in \mathfrak{S}n} sgn(\sigma_1)w(\mathscr{P}_{A \to B, \sigma_1})\right) \left(\sum_{\sigma_2 \in \mathfrak{S}n} sgn(\sigma_2)w(\mathscr{P}_{B \to C, \sigma_2})\right)$$

where  $\mathscr{P}_{A\to B,\sigma_1}$  is a path system consists of matching of  $\sigma_1$  from A to B. Note the last inequality comes from the fact that for given  $\sigma$  and  $\sigma_1$ ,  $\sigma_2$  is determined uniquely as  $\sigma_1^{-1}\sigma$ , thus it is just the same as permuting all elements in  $\mathfrak{S}n^2$ , done.

3. More intersecting graphs. Suppose  $A_i = (a_i, 0), B_i = (0, b_i)$  with  $0 \le a_i < \cdots < a_n, 0 \le b_1, \cdots, b_n$ . Now each edge in grid on the 2-dimensional coordinate space has weight 1. We denote path system from A to B be a path on the grid graph from A to B. Then, let

$$M_{ij} = \sum_{P:A_i \to B_j} w(P) = \sum_{p:A_i \to B_j} 1 = \binom{a_i + b_j}{a_j}$$

by counting lattice path from  $A_i$  to  $B_j$  which has equivalent to counting lattice path from 0 to  $(a_i, b_j)$ . Then, the GV lemma says that

$$\det\left(\binom{a_i+b_j}{a_i}_{i,j=1}^n\right) = \sum_{\mathscr{P}_{\sigma}: \text{ vertex disjoint}} sgn(\sigma)w(\mathscr{P}_{\sigma}) = \sum_{\mathscr{P}_{\sigma}: \text{ vertex disjoint}} sgn(\sigma)$$

since weight of any path, or path system should be 1. Now note that a path in  $\mathcal{P}_{\sigma}$  is vertex disjoint if and only if  $\sigma = id$ . To see this note that if there exists a path from  $A_1$  to  $B_j$  with  $j \neq 1$ , then any path from A to  $B_1$  must cross the path  $A_1 \rightarrow B_j$ , so j = 1, Do the same argument on 2, 3,  $\cdots$ , n, done. Thus,

$$\det\left(\binom{a_i+b_j}{a_i}_{i,j=1}^n\right) = \sum_{\mathscr{P}_{\sigma}: \text{ vertex disjoint}} sgn(\sigma) = \sum_{\mathscr{P}_{id}: \text{ vertex disjoint}} 1$$

thus the determinant is just number of noncrossing path systems from  ${\cal A}_i$  to  ${\cal B}_i$  .

4. Similarly, let  $0 \le a_i < \cdots < a_n$ ,  $0 \le b_1, \cdots, b_n$ ). What is the determinant of M where  $M_{ij} = {a_i \choose b_j}$ ? Thinking a picture where

$$A_i = (0, -a_i), B_j = (b_j, -b_j).$$

Then, the number of lattice paths from  $A_i$  to  $B_j$  is  $\binom{b_j + (-b_j) - (-a_i)}{b_j} = \binom{a_i}{b_j}$ . However, in this case, by the same argument above, we can see that the only vertex disjoint path occur when  $\sigma = id$ , thus the determinant is just the number of noncrossing path system from  $A_i$  to  $B_i$ ,  $i = 1, 2, \dots, n$ . Moreover, if  $a_i = m + i - 1, b_j = j - 1$ , then there exists only one noncrossing path system occur; see below picture, which shows when n = 4, m = 3.



In general, we claim that the only possible noncrossing path has a form  $A_i \to A_i + (i-1,0) \to B_i$ . If i = 1, trivially done, since they're in the same vertical line. For i + 1 case, note that this path should follow a line from  $A_{i+1}$  to  $A_{i+1} + (i+1-1,0)$ , otherwise the path touch  $A_i \to B_i$  constructed before. Then the only one way to construct the remaining path is to follow  $A_{i+1} + (i+1-1,0) \to B_{j+1}$  vertically. Thus, in this case,

$$\det\left(\binom{m+i-1}{j-1}\right) = 1.$$

5. Noncrossing Dyck paths. We defined the Dyck Path before. Now we want to counting noncrossing Dyck paths from a set of points A to B. Note that if n = 1, then the number of Dyck path is 1, when n = 2, 3, and n = 3 then 14. Now define

$$A_i = (-2i, 0), B_i = (2(i - 1 + n), 0).$$

and a path system from  $A = \{A_1, \dots, A_n\}$  to  $B = \{B_1, \dots, B_n\}$  consisting of Dyck path from  $A_i$  to  $B_{\sigma(i)}$  for some  $\sigma$  for  $i \in [n]$  as Dyck path of semi-length n. Then say the Dyck path of semi-length n is noncrossing if no two paths cross each other. Note that from the picture we can easily get only noncrossing Dyck path occur when  $\sigma = id$ , by the same argument above. Now define  $M_{ij}$  be the number of paths in the digraph from  $\{A_i\}_{i=1}^n$  to  $\{B_i\}_{i=1}^n$  consisting of all possible Dyck paths. (Thus, we implicitly assume that w(e) = 1 for each edge in Dyck path.) Then, from the section of Catalan number, we know

$$M = \begin{pmatrix} C_n & C_{n+1} & \cdots & C_{n+k-1} \\ C_{n+1} & C_{n+2} & \cdots & C_{n+k} \\ \cdots & \cdots & \cdots & \cdots \\ C_{n+k-1} & C_{n+k} & \cdots & C_{n+2k-2} \end{pmatrix}$$

By the GV Lemma, we get

$$\det M = \sum_{\sigma \in \mathfrak{S}n} sgn(\sigma)w(\mathscr{P}_{\sigma}) = \# \text{ of vertex disjoint path from } A_i \text{ to } B_i$$

Now, if n = 1, then we can show that # of vertex disjoint path from  $A_i$  to  $B_i = 1$  as follow; let  $A_0 = 0 = B_0$ . Then there exists only a path from  $A_0$  to  $B_0$ , which is just vertex 0. Then by induction, we can see that only path from  $A_i$  to  $A_j$  is just a triangle path, by the similar argument as we did in the previous example. Thus,

$$\det \begin{pmatrix} C_0 & C_1 & \cdots & C_n \\ C_1 & C_2 & \cdots & C_{n+1} \\ \cdots & \cdots & \cdots \\ C_n & C_{n+1} & \cdots & C_{n+2(k-1)} \end{pmatrix} = \sum_{\sigma \in \mathfrak{S}n} = 1$$

6. Let  $A = (0,0), B = (b_1, b_2)$  and x is a lattice point in a rectanular generated by A and B. Then,

$$LP(A \to B; \text{ avoid } x) = LP(A \to B) - LP(A \to x)LP(x \to B).$$

Question is, what is the number of  $LP(A \to B, \text{ avoids } x_1, \dots, x_n)$ ? To see this, let's make a path system from  $(A, x_1, \dots, x_n)$  to  $(B, x_1, \dots, x_n)$  with assumption that for any  $\sigma \neq id$ , any path in  $P_{\sigma}$  is not vertex disjoint. Then, by the lemma,

$$#LP(A \to B; \text{ avoids } x_1, \cdots, x_n) = \det M \text{ where } M = \begin{pmatrix} LP(A \to B) & LP(A \to x_1) & \cdots \\ LP(X_1 \to B) & \cdots & \cdots \\ \cdots & \dots & LP(x_n \to x_n) \end{pmatrix}$$

## 6 Enumeration of Patterns

#### 6.1 Symmetries and Patterns

Think about a geometric figure. For example, regular n gon can have symmetry represented by  $D_{2n}$ , dihedral n group. Now think any graph G = (V, E) and a coloring of vertice  $c : V \to \mathbb{N}$ . Then the symmetry g acting on c, say g \* c is another coloring. For example, consider a 5-gon colored with  $* - \bullet - * - \bullet - \bullet - (*)$ . If we set g be a rotating this 5-gon by  $\frac{2\pi}{5}$ , then we can just let  $g * c := c \circ g$  meaning that

$$g * c(v) = c(g.v), \forall v \in V.$$

So for above example,



if we say c as left, then g \* c is the right. (Note that in other text book, they define  $g * c = c \circ g^{-1}$ .) Now say that 2 coloring c, c' are equivalent if there exists g such that g \* c = c'. Then a pattern denotes an equivalence class. Our goal is just count the number of patterns. For example,

- 1. As we showed above, think n polygon with distinct vertices. Then symmetry can be given by  $C_n$ , a cyclic group with n. Then the number of patterns is just the number of circular permutations, which is (n-1)!.
- 2. Let *n*-polygon with colored vertice be given, such that coloring is k coloring, i.e., range of k has a cardinality k, and symmetric group is just  $C_n$ . Then we call a pattern as *necklace*, since it looks like a pattern in real necklace. And a pattern induced from k-colors for vertices with symmetric group  $D_n$  is called *key-chain*.

We can formalize above argument as below; Given a finite set N, R consider all map  $f : N \to R$ , and say f be coloring of N with color in R. Let G be a finite group acting on N, and each g acts a a permutation of N.

**Definition 6.1.1.** Two maps f and f' are equivalent  $f \sim f'$  whenever  $\exists g \in G$  such that

$$f' = g * f := f \circ g.$$

And we call an equivalence class as a pattern.

Think about a group G acting on a set X, for example, a set of all colorings. Then, for all  $x \in X$ , we can define

- Orbit of x:  $M(x) := \{gx : g \in G\}$
- Stabilizer of x:  $G_x := \{g \in G : gx = x\}$
- Fixed point of  $g: X_q := \{x \in X : gx = x\}.$

For example, let  $G = C_4 = \{e, \rho, \rho^2, \rho^3\}$  and let X be a set of all possible 2 coloring of vertex in a cycle of four vertices. Let x = then,

$$M\left(\stackrel{\bullet}{\bullet}\stackrel{\bullet}{\bullet}\right) = \left\{\stackrel{\bullet}{\bullet}\stackrel{\bullet}{\bullet}, \stackrel{\bullet}{\bullet}\stackrel{\bullet}{\bullet}\right\} = \{x, \rho(x)\}.$$
$$G_x = \{e, \rho^2\}$$
$$X_{\rho^2} = \left\{\stackrel{\bullet}{\bullet}\stackrel{\bullet}{\bullet}, \stackrel{\bullet}{\bullet}\stackrel{\bullet}{\bullet}, \stackrel{\bullet}{\bullet}\stackrel{\bullet}{\bullet}, \stackrel{\bullet}{\bullet}\stackrel{\bullet}{\bullet}\right\}$$

Then the number of orbits of X under G is  $|\{M(x) : x \in X\}|$ . Note that M(x) = M(y) or  $M(x) \cap M(y) = \emptyset$  for any  $x, y \in X$ . Now observe that

$$\sum_{x \in X} |G_x| = \sum_{g \in G} |X_g|.$$

Both just count the number of  $\{(g, x) : gx = x\}$ . Also,

$$|M_x| = |G|/|G_x|.$$

To see this, note that  $M_x = \{gx : g \in G\}$ . So if  $g_1x = g_ix$ , this implies  $x = g_1^{-1}g_ix$ , thus  $g_1^{-1}g_i \in G_x$ , say  $G_x = \{a_1, \dots, a_j\}$ . Then,

$$g_1^{-1} = g_i = a_j$$
 for some  $j \iff g_i = g_1 a_j$ , for some  $j$ .

Exactly,  $|G_x|$  many  $g_{ix}$  are same  $g_1x$ . Since  $g_1$  is arbitrarily chosen, such analysis can be applied for any g. Hence,

$$|M_x| = |G|/|G_x|.$$

Lemma 6.1.2 (Burnside Frobenius Formula).

$$\#orbits = \frac{1}{|G|} \sum_{g \in G} |X_g|.$$

*Proof.* Note that

$$\sum_{g \in G} |X_g| = \sum_{x \in X} |G_x| = \sum_{o: \text{ orbits } x \in o} |G_x|$$

If  $x, y \in o$ , then  $|M_x| = |M_y| \implies |G_x| = |G_y|$  from the formular  $|M_x| = |G|/|G_x|$ . Thus,

$$\sum_{g \in G} |X_g| = \sum_{o: \text{ orbits } x \in o} \sum_{x \in o} |G_x| = \sum_{o: \text{ orbits } x \in o} \sum_{x \in o} |G_x| \cdot |M_x| = |G| \cdot \# \text{ orbits.}$$

There are several applications.

1. Coloring of N = [n] with  $R = \{1, 2, \dots, r\}$ . If  $G = \{id\}$ , then the number of homogeneous coloring is just any function from N to R, so  $r^n$ . By the above formula

$$\#\text{orbits} = \frac{1}{1}(X_{id}) = r^r$$

thus all elements of X is fixed elements.

2. Coloring of N = [n] by R = [r]. Let  $G = S_n$ . Then by the above formula,

$$\#\text{orbits} = \frac{1}{n!} \sum_{\pi \mathfrak{S}n} |X_{\pi}|$$

Now note that a coloring can be fixed by  $\pi$  if and only if all element in the same cycle of  $\pi$  receives the same color. Thus,  $|X_{\pi}| = r^{\# \text{cycle}(\pi)}$ . Hence,

$$\#\text{orbits} = \frac{1}{n!} \sum_{\pi \mathfrak{S}n} r^{\#\text{cycle}(\pi)} = \frac{1}{n!} r(r+1) \cdots (r+n-1) = \binom{r+n-1}{n}$$

This is come from the Stirling number of first kind. See proposition 1.3.7 in [Sta11][p.33]. And recall that the number of nonnegative number solutions of  $x_1 + \cdots + x_r = n$  is  $\binom{r+n-1}{n}$ .

3. For example, think about 2 coloring of a rectangular, with symmetric group  $C_4 = \{e, \rho, \rho^2, \rho^3\}$ . Then

g	$ X_g $
e	$2^{4}$
$\rho$	2
$\rho^2$	$2^{2}$
$\rho^3$	2

By letting r = 2, we get

$$\#orbit = \frac{1}{4}(r^4 + r^2 + 2r)$$

by the above lemma.

4. Let  $G = C_n$ . X be a set of all words of length n on alphabet  $\{1, 2, \dots, r\}$ . Then,  $\rho^k \in G$  acts on X by sending  $a_i$  to  $a_{i+k}$  where i + k denotes  $i + k \mod n$ . Then, by the lemma,

$$\#\text{necklace} = \frac{1}{n!} \sum_{g \in G} |X_g|.$$

If n is a prime, then

$$\begin{array}{|c|c|c|}\hline g & |X_g| \\\hline e & r^n \\\rho & r \\\hline & & r \\\rho^{n-1} & r \\\hline & & r \\\rho^{n-1} & r \\\hline \end{array}$$

(Think that multiplicative subgroup of  $\mathbb{Z}_n$  is cyclic.) Thus,

#necklace = 
$$\frac{1}{n!} \sum_{g \in G} |X_g| = \frac{1}{n} (r^n + (n-1)r).$$

5. Think about pentagon with vertices



Let X be a set of r colorings of vertices,  $G = D_5 = \{\rho^i, \tau \rho^i\}_{i=0}^4$ , where  $\rho$  is rotating once counterclockwise, and  $\tau$  is reflection. Then,

g	Permutation induced from $g$	$ X_g $
e	(1)(2)(3)(4)(5)	$r^n$
$\rho^i$	$(i,i+1,\cdots,i-1)\mathrm{m}$	r
$\tau$	(1)(25)(34)	$r^3$

And  $\tau \rho^i$  has a decomposition with 3 cycles since the same reflection works for different one. Hence, by the formula,

#orbits = 
$$\frac{1}{10}(r^5 + 4r + 5r^3)$$
.

### 6.2 Cycle index

From above applications, we can observe that if g is a permutation of set X, a set of r colorings. Then,

 $|X_q| = |$ Coloring that are fixed by  $g| = r^{\# \text{cycle}(g)}$ .

So we can introduce a generating function called *cycle index*. For each  $g \in G$ , g induces a permutation of X, so it has cycle decomposition  $c_1, \dots, c_k$ . Define that

$$mono(g) := z_{k_1}, \cdots, z_{k_i}$$

where  $k_i$  denotes the length of cycle  $c_i$  of the cycle decomposition of g. For example,

g	Permutation induced from $g$	mono(g)
e	(1)(2)(3)(4)(5)	$z_1^5$
$\rho^i$	$(i,i+1,\cdots,i-1)\mathrm{m}$	$z_5$
$\tau$	(1)(25)(34)	$z_1 z_2^2$

So, if g has  $e_1$  cycles of length 1,  $e_2$  cycles of length 2,  $\cdots$ ,  $e_l$  cycles of length l, and so on, then

$$mono(g) = z_1^{e_1} \cdots z_l^{e_l} \cdots$$

Then define a cycle index; let

$$P_G(z_1, \cdots, z_n) = \frac{1}{|G|} \sum_{g \in G} mono(g).$$

where n = |X|. For example,

$$P_{D_5}(z) = \frac{1}{10}(z_1^5 + 4z_5 + 5z_1z_2^2).$$

Note that if we put  $z_i = r$  for any *i*, then

$$P_G(r, \cdots, r) = \#$$
 of patterns.

New question is, how many 2-coloring of vertices of a pentagon using exactly 3 black and 2 whites? Let X' be a subset of X with such colorings. Then,

g	Permutation induced from $g$	$ X'_g $
e	(1)(2)(3)(4)(5)	$\binom{5}{3}$
$\rho^i$	$(i,i+1,\cdots,i-1)\mathrm{m}$	Ũ
$\tau$	(1)(25)(34)	2

Since  $|X'| = {5 \choose 3}$ , and  $X'_{\tau}$  can be constructed by choosing a vertex in 1 cycle as a black, and choosing one of the 2-cycle as black and the other as white. This can be applied for  $\tau \rho^i$  for any  $i = 1, \dots, 5$ . Hence, the number of orbits are  $\frac{10+5\times 2}{10} = 2$ .

Now think about the general case; 2-coloring with p elements black, and q elements white, with p+q=n. To compute  $X'_g$ , we need some cycles of g are black, and other cycles are white. So the total number of elements in black cycle is p, and that of white cycles is q. Assume there are  $x_i$  many cycles of length i that are black,  $i = 1, 2, \dots, k$  for some k. Then, if we assume g has  $l_i$  cycle of length i, then

$$0 \le x_i \le l_i, \qquad \sum_{i=1}^k x_i \cdot i = p_i$$

Then,

$$mono(G)(x_1^1+1,\cdots,x^n+1) = (x^1+x^0)^{l_1}\cdots(x^k+x^0)^{l_k} = \sum_{\text{all two colorings of cycle } g} x^{\# \text{ black elements}}$$

is just  $mono(G)(x_1^1 + 1, \dots, x^n + 1)$ . Thus,

 $[x^p] \sum_{\text{all two colorings of cycle } g} x^{\# \text{ black elements}} = \# \text{ of patterns with } p \text{ black and } q \text{ white elements generated by } g.$ 

Hence, the number of patterns with exactly p blacks and q white is

$$[x^p]P_G(x+1, x^2+1, \cdots, x^n+1)$$

If we let 1 = y, actually, we can track of white elements, i.e.,

$$[y^{q}]P_{G}(x+y, x^{2}+y^{2}, \cdots, x^{n}+y^{n})$$

So, for general r colors, we can track it by

$$[x_1^{p_1}\cdots x_r^{p_r}]P_G(z_1,\cdots,z_n)$$

where  $z_i = \sum_{j=1}^n x_n^i$ .

## 6.3 The Theorem of PólyaRedfield

**Theorem 6.3.1** (PólyaRedfield). Let X be a set of elements and G be a group of permutations on X. Let  $\{u_1, \dots, u_k\}$  be a set of k colors and C be a set of coloring of X. Then, the generating function of inequivalent coloring in C according to the number of colors of each color k is

$$\sum_{f: \text{ inequivalent coloring of } C} u_1^{\#\{x \text{ with color } 1\}} u_2^{\#\{x \text{ with color } 2\}} \cdots u_k^{\#\{x \text{ with color } 1\}} = P_G(z_1, \cdots, z_k)$$

4 1

where  $z_i = \sum_{j=1}^k u_j^i$ 

*Proof.* We already show in above.

**Example 6.3.2.** Think about  $D_4$  on a rectangular  $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ . Then,

g	Permutation induced from g	mono(g)
e	(1)(2)(3)(4)	$z_1^4$
$\rho^1$	(1234)m	$z_4$
$\rho^2$	(13)(24)m	$z_{2}^{2}$
$\rho^3$	(1432)m	$z_4$
$\tau$	(14)(23)	$z_2^2$
$\tau \rho^1$	(13)	$z_1^2 z_2$
$\tau \rho^2$	(12)(34)	$z_2^2$
$\tau \rho^3$	(24)	$z_1^2 z_2$

Thus,

$$P_G(z_1, \cdots, z_4) = \frac{1}{8}(z_1^4 + 3z_1^2 + 2z_4 + 2z_1^2 z_2).$$

If we want to find 2-colorings, then by the theorem,

$$\sum_{\substack{f: \text{ inequivalent 2-coloring} \\ x^{\#\{x \text{ with coloring 1}\}}y^{\#\{x \text{ with coloring 2}\}} = P_G(x+y, x^2+y^2, x^3+y^3, x^4+y^4) = x^4 + x^3y + 2x^2y^2 + xy^3 + y^4.$$

# 7 Catalan numbers and structures

## 7.1 Example of catalan structures.

Note that  $C_n := \frac{1}{n+1} \binom{2n}{n}$  is a catalan number, which we dealt on previous lecture. For the first values we know that

$$C_0 = 1, 1, 2, 5, 14, 42, 132, \cdots$$

We already know that

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$$
 and  $C(x) := \sum_{n=1}^{\infty} C_n x^n = \frac{1 - \sqrt{1 - 4x}}{2x}.$ 

There are fundamental structures related to the Catalan numbers.

- Triangulation of (n+2)-gon.
- Ballot sequence
- Dyck Path: above three topics were dealt previously.

• Binary trees.

Definition 7.1.1. A tree is called binary tree if it is generated by below recursive process; first of all,

root

 $\emptyset$  be a binary tree. And if  $T_1, T_2$  are binary, then so is  $\begin{array}{cc} & & \\ T_1 & & T_2 \end{array}$ .

For example,



Then,

 $C_n = \#$  of binary tree with *n* vertices.

• Now we can define a *complete binary tree*, which is a binary tree such that every vertex v either have two children or no child. Note that we can convert binary tree to complete binary tree by adding edges which is denoted by dashed line in the above figure. Thus,

 $C_n = \#$  complete binary tree with *n* interval vertices

where interval vertex means a vertex with child. This comes from the bijection from complete binary tree to binary tree; to see the inverse operation, just remove all leaves from the complete tree.

• Plane tree.

**Definition 7.1.2.** A tree is plane tree if it can be generated by following recursive process;

1.  $(\{\bullet\}, \emptyset)$  is a plane tree.

2. If 
$$P_1, \dots, P_m$$
 are plane tree, then so is  $P_1 \cdots P_m$ .

Then we can show that

 $C_n = \#$  plane trees with n + 1 vertices.

For example, when n = 0, then only a verte is plane tree. If n = 1, then  $\bullet - \bullet$  is the only one plane tree. If n = 2,  $\bullet - \operatorname{root} - \bullet$  and  $\operatorname{root} - \bullet - \bullet$  are those. (The latter is just when case m = 1 with  $P_1 = \bullet - \bullet$ . For n = 3, there are five cases, where the top vertex is root of the tree.



• Parenthesization (Bracketing.) Think about the number of ways of apply a binary operation to n + 1 vertices; if n = 0, then no case exists. If n = 1, the only one way is  $(x_1x_2)$ . If n = 2,  $((x_1x_2)x_3), (x_1, (x_2, x_3))$ , and for n = 3,

 $(((x_1x_2)x_3)x_4), ((x_1, x_2)(x_3, x_4)), ((x_1(x_2, x_3))x_4), (x_1((x_2x_3)x_4)), (x_1(x_2(x_3x_4))).$ 

There exists a map from plane tree to binary tree; for example, let r(v), a rank of v be a length from v to the root. Then r(root) = 0. Then, generate binary tree by this process; make a line graph for r(v) = 1. Then, for a set with rank v, we can partition it by its parent. (v must have only one parent by construction of plane tree, and the parent has r(v) = 1.) Then, for each parent v, attach the line graph generated by a partition of  $\{w : r(w) = 2\}$  having v as its parent. Do the same thing for r(v) = 3, 4, and so on. For example,



And bracket can be converted into a complete binary tree by using preorder system. For example,

$$(((x_1x_2)x_3)x_4)$$

can be represented to the binary tree



traveling the most left vertices for each level.

Thus, we can have bijection among plane tree, ballot sequence, parenthesization with complete binary tree.

**Lemma 7.1.3** (Cyclic lemma).  $C_n = \frac{1}{n+1} \binom{2n}{n} = \frac{1}{2n+1} \binom{2n+1}{n}$ .

This is already shown in previous chapter.

**Lemma 7.1.4** (Ramney's lemma). For any sequence  $a_1, \dots, a_{2n+1}$  containing n+1+1's and n-1's, exactly one cyclic shift is a strict ballot sequence, i.e., any partial sum starting at 1 to  $i \in [2n+1]$  is strictly greater than 0, i.e.,

$$\sum_{i=1}^{j} b_i > 0, \forall j \in [2n+1].$$

There are  $\binom{2n+1}{n}$  sequences we can partition it by cyclic shift. Then the lemma says that each orbit and a one "strict ballot sequence." Since each orbit has exactly 2n + 1 elements, hence

# strict ballot sequence 
$$= {\binom{2n+1}{n}} \frac{1}{2n+1} = C_n.$$

*Proof.* Think +1 as  $\nearrow$ , i.e.,  $(0,0) \rightarrow (1,1)$  and -1 be  $\searrow$ , i.e.,  $(0,0) \rightarrow (1,-1)$ . Then, we can make a kind of path from (0,0) to (2n+1,1). If we copy this path from (0,0) to (2n+1) and paste it to the point starting at (2n+1,1), then we have a path from (0,0) to (4n+2,2). Then there are the lowest point. Take the rightmost lowest point. Note that this point is in the region where  $x \in [0, 2n+1]$ , since the pasted path goes

one unit north. Now if we think the rightmost lowest point as (0,0), then by construction, its path from new (0,0) to (2n + 1, 1) doesn't contain a point touching x = 0 except (0,0), so the partial sum is always strict, done.

To see its uniqueness, if there are two strict ballot sequence  $a_1, \dots, a_{2n+1}$  and  $a_i, \dots, a_{2n+1}, a_1, \dots, a_{i-1}$ , note that

$$\sum_{j=1}^{i-1} a_j \ge 1, \sum_{j=i}^{2n+1} a_j \ge 1 \implies 1 = a_1 + \dots + a_{2n+1} \ge 2$$

contradiction.

There is a variation of cycle lemma, called Spitzer's lemma.

**Lemma 7.1.5** (Spitzer's lemma). Let  $a_1, \dots, a_n$  be a real number with property that

- 1.  $\sum_{i=1}^{N} a_i = 0$
- 2. Read cyclically no other nonempty sum of consecutive  $a_i$ 's is 0, i.e. total sum is zero but any consecutive sum is nonzero.

Then, there exists unique cyclic shift  $b_1, \dots, b_n$  such that

$$\sum_{i=1}^{k} b_i \ge 0, \forall k \in [n]$$

*Proof.* Take a path  $(0,0) \to (1,a_1) \to (1,a_1) + (2,a_1+a_2) \to \cdots \to (N,\sum_{i=1}^N a_i) = (N,0)$ . Then we can do the same thing as above, starting at the lowest point.

To see the uniqueness, if it is not unique, then there exists two positive partial sum, but sum of these partial sums are just total sum which is zero, contradiction. This is also just following proof of Ramney's lemma.  $\Box$ 

Application of Spitzer's lemma is here; let gcd(r,s) = 1 for some  $r, s \in \mathbb{N}$ . Then, the number of lattice paths from  $\vec{0}$  to (r,s) stay weakly below the line  $y = \frac{s}{r}x$ , i.e. a diagonal line between  $\vec{0}$  and (r,s), is

$$\frac{1}{r+s}\binom{r+s}{r}.$$

*Proof.* Take a path from  $\vec{0}$  to (r, s). Then we can denote it by a sequence of r east step and s north steps. Then generate a sequence  $a_1, \dots, a_{r+s}$  whre

$$a_i = \begin{cases} s & \text{for east step} \\ -r & \text{for north step} \end{cases}$$

which corresponds to the each step at time *i*. So, the total sum is zero, and consecutive partial sum is nonzero, since gcd(r,s) = 1. Then the lattice path stays below  $y = \frac{s}{r}x$  if and only if any partial sum is greater then 0. To see this, if we had *k* east step and *q* north step such that  $k/q \leq s/r$ , i.e., a point in the path is still stay below then the line  $y = \frac{s}{r}$ , then  $k/q \leq s/r$  is equivalent to ks - qr > 0 (this cannot be equal to 0 from coprime condition) which is equivalent to sum from  $a_1$  to  $a_{k+q}$  is greater than 0, done.  $\Box$ 

### 7.2 NC matching of [n]

This is about parint of 2n points with noncrossing diagonals. i.e., there is no 2 blocks  $B_1, B_2$  such that  $i, j \in B_1, k, l \in B_2$ , and i < k < j < l. From this definition, we just think of case when points are in linear, and block partition by an edge should not cross one another if we restrict edge be upward on the line of points. For example,



the left one is noncrossing but the right one is crossing pair.

We know that  $C_n$  is just number of noncrossing matchings on [2n].

• Pattern avoidance in permutation pattern  $\sigma \in \mathfrak{S}k$ : Define a map from the given sequence of distinct number  $a_1, \dots, a_k$ . let  $std(a_1, \dots, a_k)$  be a permutation obtained from replacing the *i*th minimal entry with *i*. For example,

 $std(35127) \mapsto 34125.$ 

So std only care about the relative order of entry. Given a permutation  $\pi \in \mathfrak{S}n$  with  $n \geq k$ , say  $\pi$  contains  $\sigma$  if there exists a subsequence  $\pi'$  of  $\pi$  such that  $std(\pi') = \sigma$ . (Note that std comes from "standardization.") For example, if  $\sigma = 123$  and  $\pi = 75326148$ , then  $\pi$  has subsequences 148,268 whose standardization is 123. Otherwise, we say that  $\pi$  avoids  $\sigma$ . Let

$$Av_n(\sigma) = \{\pi \in \mathfrak{S}n : \pi \text{ avoids } \sigma\}.$$

Then it is known that for any  $\sigma \in \mathfrak{S}3$ ,

 $|Av_n(\sigma)| = C_n.$ 

For any permutation  $\sigma = \sigma_1 \cdots \sigma_n$ , let  $\sigma^r := \sigma_n \sigma_{n-1} \cdots \sigma_1$ , i.e., reversing the permutation, and let

$$\sigma^c := (n+1-\sigma_1)\cdots, (n+1-\sigma_n).$$

Then,

$$|Av_n(\sigma)| = |Av_n(\sigma^n)| = |Av_n(\sigma^c)|$$

since applying the same map  $(\bullet)^r, (\bullet)^c$  on  $Av_n(\sigma)$  gives a bijection. Also, if

$$|Av_n(\pi)| = |Av_n(\sigma)|$$

for some  $\pi, \sigma$  then we say  $\pi$  and  $\sigma$  is *wilf-equivalent*, i.e.,  $\pi \stackrel{\text{W}}{\sim} \sigma$ .

• Noncrossing partition. Let  $NC_n$  be a set of noncrossing partitions of [n]. Precisely,  $\pi = (B_1, \dots, B_k) \in \Pi_n$  is noncrossing if and only if no  $i, j \in B_{r_1}, k, l \in B_{r_2}$  satisfy i < k < j < l for any  $r_1, r_2 \in [k]$ . We can represent it by convex hull on a points where  $j \in [n]$  is represented by  $e^{\frac{\pi}{2}i - \frac{j2\pi}{n}i}$ . For example, [n] = 9 with noncrossing partition (136)(2)(45)(79)(8) can be represented as



To see why the cardinality of  $NC_n$  is  $C_n$ , think about a convex hull on a circle representing noncrossing partition  $\pi$  which is arbitrarily chosen, and let k be a maximum element in the block containing 1. Then,  $\pi$  is disjoint union of 1) noncrossing partition of  $\{1, 2, \dots, k\}$  such that 1 and k are in the same block, and 2) noncrossing partition of  $\{k + 1, \dots, n\}$ , which gives  $C_{n-k}$  by inductive hypothesis. And to calculate 1) note that it is 1 - 1 relationship with a noncrossing partition of  $\{1, \dots, k-1\}$  by just adding or deleting k from the partitions. Thus it gives  $C_{k-1}$  way, hence

$$\#NC_n = \sum_{k=1}^n C_{n-k}C_{k-1} = C_n$$

Now, we claim that

or

or

**Claim 7.2.1.**  $NC_n$  is a lattice, i.e.,  $\forall \pi, \sigma \in NC_n$ , there exists the unique greatest lower bound and the least upper bound.

Note that  $\Pi_n$ , a lattice of partition is defined from the Poset structure such that  $\pi \leq \sigma$  if  $\pi$  refines  $\sigma$ , i.e., every block in  $\pi$  is subset of block in  $\sigma$ .

*Proof.* Define meet as  $\pi \wedge \sigma = \{\pi_i \cap \sigma_j \neq \emptyset : \pi_i \in \pi, \sigma_j \in \sigma\}$ . Then, since it is just meet in  $\Pi_n$ , and  $NC_n$  is subposed of  $\Pi_n$ , it suffices to show that this meet is also noncrossing. To see this, if we take any two block in  $\pi \wedge \sigma$  and each two elements from those blocks, then we can take two blocks from  $\pi$  or  $\sigma$  distinguishing those two elements (at least one of  $\pi$  or  $\sigma$  distinguish them,) thus it satisfies noncrossing relationship.

However, in case of defining join, thing is not easy. Actually, join in  $\Pi_n$  may not be a noncrossing partition. For example, let  $\pi = 135 - 2 - 4 - 6$  and  $\sigma = 1 - 3 - 5 - 246$ , then  $\pi \vee \sigma = 135 - 246$ , which is crossing partition. To solve this, we need theorem in lattice.

**Theorem 7.2.2.** Let P be a finite poset which is meet-closed and P has  $\hat{1}$ . Then P is a lattice, i.e.,  $\forall x, y \in P, x \lor y$  exists.

Proof of the theorem. For all  $x, y \in P$ , take  $S = \{z : z \ge x, z \ge y\}$ . Then,  $S \ne \emptyset$  since  $\hat{1} \in S$ . Also, S is finite since P is finite. Since P is meet closed, let  $u = \wedge_{z \in S} z$ , i.e., meet of all  $z \in S$ . Then,  $u = x \lor y$ .

In  $NC_n$ ,  $\hat{1} = 123 \cdots n$ , and meet closed as we shown above. Thus, by the theorem,  $NC_n$  is a lattice.  $\Box$ 

What is cover relation in  $NC_n$ ? We call  $\sigma$  is a *cover* of  $\pi \in NC_n$  such that there exists no  $\pi' \in NC_n$  such that  $\pi < \pi' < \sigma$ .

**Lemma 7.2.3.** If  $\pi$  is in  $NC_n$  and  $\pi$  has at least 3 blocks. it always possible to find  $\pi' \in NC_n$  such that  $\pi < \pi' < \hat{1}$ .

*Proof.* Take  $\pi'$  be unioning any two blocks of  $\pi$ .

From the proof of lemma, if  $\sigma$  covers  $\pi$ , then  $\#block(\sigma) = \#block(\pi) - 1$ . (Otherwise we can construct a noncrossing partition between  $\sigma$  and  $\pi$ , contradiction.) Hence,  $NC_n$  can be graded by  $rank(\pi) = n - \#blocks(\pi)$ . From this, we can conclude that  $NC_n$  is a poset having a layer, which is self-symmetric, thus we can calculate a Möbius function of  $NC_n$ . For example,  $NC_4$  looks like this;



1. Self-dual property in  $NC_n$ : there exists an involution  $\tau NC_n(k) \to NC_n(n+1-k)$  where  $NC_n(k)$  means a set of NC partitions having k blocks. For example, let  $\pi = 138 - 2 - 4 - 57 - 6$ . Then we can put it in the circle with points  $1, 2, \dots, 8$  as below.



Now draw new number  $1', \dots, 8'$  in reverse order between  $1, 2, \dots, 8$  and draw a convex hull which can be partitioned by above convex hull. For example,



This red convex hull denote 145 - 23 - 67 - 8. This labeling gives involution. (Note that other labeling of prime points may not be involution.) To see this why it is involution, draw a line between 8 and 8' through 4 and 4' and think about mirror symmetry.

So if n is odd number, then involution gives a bijection for the subset of  $NC_n$  having a middle rank.

How many fixed points here? If n is even, then no fixed points exist. If n = 2m + 1 is odd, then  $C_m$  fixed points exists. Also,  $|NC_n(k)| = \frac{1}{n} {n \choose k} {n \choose k-1}$ , which is called *Narayana number*.

- 2. Now think about interval  $[\pi, \sigma] \subseteq NC_n$ .
  - (a) Case 1: Let  $\pi = \hat{0} = 1 2 \dots n$ , and let  $\sigma = (B_1, \dots, B_k)$ . Then, we can see that

$$[\hat{0},\sigma] = \prod_{i=1}^{k} NC(B_i).$$

To see this, note that any partition  $\tau$  less than  $\sigma$  is refinement of  $\sigma$ , so if we think all blocks in  $\tau$  contained in  $B_i$  for arbitrarily chosen *i*, they form a noncrossing partition of  $B_i$ .

(b) In general, for example, let  $\sigma = 1, 6, 9, 12 - 245 - 3 - 78 - 10, 11, \pi = 19 - 25 - 3 - 4 - 6 - 78 - 10 - 11 - 12$ . Then we can generate this interval by cartesian product of subinterval; namely,

$$[\pi, \sigma] = [(1, 9 - 6 - 12), (1, 6, 9, 12)] \times [(25 - 4), (245)] \times [3, 3] \times [78, 78] \times [10 - 11, (10, 11)]$$

This is because for  $\tau \in [\pi, \sigma]$ , each block of  $\tau$  refining  $\sigma$  but  $\pi$  should be refining of  $\tau$ , so for each block in  $\sigma$ , block of  $\tau$  has also noncrossing property. And to calculate each summand of cartesian product, we can use the involution, since involution preserves cardinality. Note

that each subinterval has  $\hat{1}$  in those points. Thus, by applying involution, it goes to  $\hat{0}$ . For example, in above case, we get

$$\begin{split} [(1,9-6-12),(1,6,9,12)] \mapsto [1-6-9-12,(1,12-6,9)] &\cong NC_2 \times NC_2 \\ [(25-4),(245)] \mapsto [2-4-5,24-5] &\cong NC_2 \times NC_1 \\ [3,3] \mapsto [3,3] &\cong NC_1 \\ [10-11,(10,11)] \mapsto [10-11,(10,11)] &\cong NC_2 \end{split}$$

So,  $[\pi, \sigma] = \prod NC_i$  for some proper index *i* which can be repeated, which implies

$$\mu(\pi,\sigma) = \prod \mu_{NC_i}(\hat{0},\hat{1}).$$

(c) Now calculate  $s_n = \mu_n(\hat{0}, \hat{1})$ . First three values are easy;  $s_1 = 1, s_2 = -1, s_3 = 2$ . Actually,  $s_4 = -5$  by computing directly. Now we claim that

Claim 7.2.4.  $s_n = (-1)^{n-1} C_{n-1}$ .

*Proof.* Recall Weisner's theorem, i.e., if  $a \neq \hat{1}$ , then

$$\mu(\hat{0},\hat{1}) = -\sum_{\substack{x:x \neq \hat{1} \\ x \lor a = \hat{1}}} \mu(\hat{0},x).$$

In  $NC_n$ , let  $a = 1 - 2 - 3 - \dots - (n - 2) - (n - 1, n)$ . Then, if  $x \lor a = \hat{1}$  but  $x \neq \hat{1}$ , then x cannot have a block included in [n - 2]; then we can construct a greater element of x and a by only thinking the remaining elements in x except that block. So every block in x has to intersect with  $\{n - 1, n\}$ . This implies x has only two block, one block containing n - 1 and the other block containing n. To satisfy noncrossing condition, this x looks like  $\{\{1, 2, \dots, k - 1, n\}, \{k, \dots, n - 1\}\}$ . Thus all possible  $x \neq \hat{1}$  is when  $k = 1, \dots, n - 1$ . Hence each block size is k, n - k, so

$$\mu(\hat{0}, x) = \mu(NC_i \times NC_{n-i} = s_i \times s_{n-i})$$

 $\mathbf{SO}$ 

$$\mu_{NC_n}(\hat{0},\hat{1}) = -\sum_{i=1}^{n-1} s_i s_{n-i}.$$

And from the initial condition, we can get

$$s_i = (-1)^{i-1} C_{i-1}.$$

# References

[Aig07] Martin Aigner, A course in Enumeration, Springer, 2007.

- [Hai10] Mark Haiman, Notes on exponential generating functions and structures. (2010), https://math.berkeley.edu/ ~mhaiman/math172-spring10/exponential.pdf. Accessed Sep. 6, 2018.
- [Sta11] Richard P. Stanley, Enumerative Combinatorics. Vol. 1, Second Edition, 2011.
- [DK16] Dick Koch, The Pentagonal Number Theorem and All That (2016).