

Contents

1	Pigeonhole Principle	3
2	Binomial Coefficient	4
2.1	Basic Definitions	4
2.2	Binomial Coefficient Definition	4
2.3	Binomial Theorem	4
2.4	Multichoose	5
2.5	Other Identities	6
2.6	Properties of Sequence of Binomial Coefficients	7
2.7	Multinomial Theorem	8
2.8	Generalized Binomial Theorem	8
2.9	Subscripted Binomial Coefficients	9
3	Permutations	10
3.1	Inversions	10
3.2	Cycles	12
3.3	Involutions	15
3.4	Derangements	16
4	Principal of Inclusion/Exclusion	18
5	Partitions	20
5.1	Compositions	20
5.2	Set Partitions	21
5.3	Integer Partitions	26
5.4	Lah Numbers	29
6	Graphs	31
6.1	Chromatic Polynomials	31
6.2	Trees	33
7	Generating Functions	36
7.1	Definitions	36
7.2	Common Generating Functions	37
7.3	Wilf Rules	39
7.4	Recurrences	40
7.5	EGF Examples	42
7.6	Fibonacci Numbers	44
7.7	Catalan Numbers	46
7.8	Counting Using Products and Compositions	50
7.9	Snake Oil Method	51
8	Combinatorial Classes	54
9	Labeled Structures	59
10	Inversion Formulas	63
10.1	Binomial Inversion	63
10.2	Stirling Inversion	65
10.3	Lah Inversion	67
11	Lagrange Inversion	68
11.1	Motzkin Numbers	73

12 Posets	76
12.1 Introduction	76
12.2 Incidence Algebra	79
12.3 Möbius Inversion	83
12.4 Number Theory	84
13 Combinatorial Reciprocity	88

1. Pigeonhole Principle

Theorem 1.1 (Pigeonhole Principle - PHP). If n balls are placed into k bins for $0 < k < n$, then there exists a bin with at least two balls.

Proof. Suppose not. Let $n_i \leq 1$ be the number of balls in bin i . Then $n = \sum_{i=1}^k n_i \leq k$, contradiction. \square

Example 1.2. Let $a_1, \dots, a_n \in \mathbb{N}$. Then there exists a consecutive subsequence summing to a multiple of n .

Form $n + 1$ partial sums (including the empty one). At least two share the same value modulo n by PHP. Thus, their difference (which is also a sum of consecutive a_i) is a multiple of n .

Example 1.3. Any graph with $|V| \geq 2$ has at least two vertices of equal degree.

There are n vertices and n possible degrees. However, we cannot simultaneously have a vertex of degree 0 (isolated) and degree $n - 1$ (fully connected). Thus, apply PHP in each case.

Example 1.4. 6 distinct integers have no prime divisor greater than 10. Show that two share a common prime divisor.

Let A be the set of 6 numbers, and define $f : A \rightarrow \{1, 2, 3, 5, 7\}$ be the largest prime factor dividing the number (or 1 if no prime divides it). By PHP, f maps at least two integers in A to the same value, which cannot be 1 since elements of A are distinct. Thus, the two elements share a common prime divisor.

Example 1.5. 200 balls are distributed in 100 bins, so that no bin has more than 100 balls. Show that some set of bins contains exactly 100 balls.

If all bins contain two balls, pick any 50 bins. Otherwise, there exists a bin with 1 ball. Label the bins so $1 = n_{99} \leq n_1 \leq n_2 \leq \dots \leq n_{98} \leq n_{100} \leq 100$. Form a new sequence m_i by decrementing n_{100} , so that $\sum m_i = 199$. Some consecutive subsequence is a multiple of 100 by Example 1.2, and thus equals 100. If it does not include m_{100} , then the corresponding sequence in n_i sums to 100. Otherwise, the corresponding sequence in n_i with n_{99} removed sums to 100.

Example 1.6. Let $A \in M_n(\mathbb{Z}_2)$ with at least $2n$ entries equal to 1. Prove that A contains two entries equal to 1 so that one of them is strictly above and right of the other.

There are $2n - 1$ antidiagonals of the matrix, so apply PHP.

Theorem 1.7 (Generalized Pigeonhole Principle). If n balls are distributed into k bins for $n > kr > 0$, then there exists a bin with more than r balls.

Proof. Suppose not. Let $n_i \leq r$ be the number of balls in bin i . Then $n = \sum_{i=1}^k n_i \leq kr$, contradiction. \square

Example 1.8. How many people must be in a group to guarantee a cluster of 3 friends or 3 strangers?

5 is not sufficient, since a cyclic graph may be formed (where edges represent friendships), but 6 is. Fix a person, who is either friends or strangers with 3 others by PHP; without loss of generality let it be friends. If any of them are friends with each other, there is a cluster of 3 friends; otherwise, there is a cluster of 3 strangers.

Example 1.9. Let 9 points be distributed in a unit square. Show that 3 points can be covered with a disk of radius $1/2$.

Divide the square into four triangles using the diagonals. By PHP, one of these triangles contains at least 3 points, and a disk can cover this triangle. One can also divide the square into four squares instead.

2. Binomial Coefficient

2.1 Basic Definitions

Definition 2.1. A permutation of a set S is a linear ordering of its elements. Let S_n denote the set of permutations of $[n]$.

↳ S_n is also called the symmetric group

↳ We can also define S_n to be the set of bijections $\pi : S \rightarrow S$

Theorem 2.2. The number of permutations of n elements is $|S_n| = n!$.

Proof. By induction, the number of choices for the first element is n , and the number of options for the remainder is $(n-1)!$. \square

Theorem 2.3. The number of distinguishable permutations of a multiset $S = \{1^{n_1}, 2^{n_2}, \dots, k^{n_k}\}$ is $\binom{n}{n_1 n_2 \dots n_k} = \frac{n!}{n_1! n_2! \dots n_k!}$ (this is called the multinomial coefficient).

Proof. Temporarily label each object to obtain $n!$ permutations. Then, for each object type i , divide by the number of permutations of these objects $n_i!$. \square

Definition 2.4. A k -string (or k -word) over a set (alphabet) A is a sequence $a_1 a_2 \dots a_k \in A$.

Theorem 2.5. If A is an alphabet with $|A| = n$, then the number of k -words is n^k .

Definition 2.6. A k -permutation of S is a k -word on S in which all elements are distinct.

Theorem 2.7. The number of k -permutations of S , for $|S| = n$, is $(n)_k = n(n-1) \dots (n-k+1) = \frac{n!}{(n-k)!}$.

2.2 Binomial Coefficient Definition

Definition 2.8. $\binom{S}{k}$ denotes the collection of subsets of S of size k . The binomial coefficient is $\binom{n}{k} = \left| \binom{S}{k} \right|$ for $|S| = n$.

Theorem 2.9. $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{(n)_k}{k!}$

Proof. There are $(n)_k$ k -permutations, then divide by $k!$ to remove the order. \square

Theorem 2.10. $\binom{n}{k} = \binom{n}{n-k}$

Proof. LHS selects elements in the set S , while RHS selects elements in the set S^c . \square

Theorem 2.11. $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

Proof. RHS conditions on whether n is in the subset $S \subset [n]$. \square

Example 2.12. Show $\sum_{i=1}^n i(n-i+1) = \binom{n+2}{3}$.

LHS conditions on the middle element chosen. \square

2.3 Binomial Theorem

Theorem 2.13 (Binomial Theorem). $\sum_{k=0}^n \binom{n}{k} x^k y^{n-k} = (x+y)^n$

Proof. Consider expanding the RHS. Each term $x^k y^{n-k}$ arises by choosing k factors to include x and $n - k$ to include y , which can be done $\binom{n}{k}$ ways. \square

Theorem 2.14. $\sum_{i=0}^n (-1)^i \binom{n}{i} = \delta_{n0}$

Proof. Take $x = -1$ and $y = 1$ in the Binomial Theorem. Or, for $n > 0$, form the involution mapping $A \subset [n]$ to the symmetric set difference $A \triangle n$ (i.e., add the element n if $n \notin A$ or remove n if $n \in A$). This reverses the parity, so the number of odd/even subsets are the same. \square

Theorem 2.15. $2^n = \sum_{k=0}^n \binom{n}{k}$

Proof. Apply Binomial Theorem with $x = y = 1$. Or, note both sides count subsets of $[n]$, with RHS conditioned by size k . \square

Example 2.16. Prove that there are 3^n ways to select two disjoint subsets from $[n]$.

Proof. Each element may be placed in either of the two subsets or neither, so apply product rule. \square

2.4 Multichoose

Definition 2.17. The collection of multisubsets of S of size k is $\left(\binom{S}{k}\right)$. Define $\left(\binom{n}{k}\right) = \left|\left(\binom{S}{k}\right)\right|$ for $|S| = n$.

- ↳ Each element of $\left(\binom{S}{k}\right)$ can be represented as a n -tuple of whole numbers summing to k (where each value denotes how many times the corresponding element is included in the multisubset)
- ↳ This can be used to count how many ways k votes can be distributed among n candidates (the analog of the binomial coefficient with replacement)

Theorem 2.18. $\left(\binom{n}{k}\right) = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}$

Proof. Every element of $\left(\binom{S}{k}\right)$ can be represented by a string of $n - 1$ bars and k stars. The bars divide the string into n bins, which represent each element of S . The number of stars in each bin represent how often that element is chosen. To count these string, simply choose k stars from $n + k - 1$ symbols. \square

Theorem 2.19. $\sum_{k=0}^n \binom{n}{k} \left(\binom{k}{m}\right) (-1)^k = (-1)^n \delta_{nm}$

Proof. LHS counts the number of ways to select an even number of candidates and then cast m votes for them, minus the corresponding quantity for an odd number of candidates. Group the terms of the sum by the vote distribution. When $n > m$, any person receiving a vote must be a candidate, and the remaining $p \geq 1$ people may either be a candidate or not a candidate. Thus, the terms for some fixed vote distribution are $\sum_{k=0}^p \binom{p}{k} (-1)^k = 0$ by Theorem 2.14. Thus, we have our claim for $n > m$, and it is trivial when $n = m$. \square

Theorem 2.20. $\left(\binom{n}{k}\right) = \left(\binom{n}{k-1}\right) + \left(\binom{n-1}{k}\right)$

Proof. LHS counts number of ways to distribute k votes to n candidates. RHS conditions on whether the last candidate receives a vote. \square

Lemma 2.21. The number of weakly increasing sequences $1 \leq a_1 \leq \dots \leq a_n \leq n$ is $\binom{2n-1}{n-1} = \left(\binom{n}{n}\right)$.

Proof. Choose a multisubset of size n from $[n]$ in $\binom{n}{n}$ ways. Or, use stars and bars. From left to right, read the $n - 1$ stars as incrementing a counter, and the n bars as recording the current value of the counter into the sequence. \square

2.5 Other Identities

Theorem 2.22 (Hockey-Stick Identity). $\sum_{j=0}^{n-k} \binom{k+j}{k} = \binom{n+1}{k+1}$

Proof. RHS counts the number of ways to choose a subset from $[n+1]$ of size $k+1$. LHS conditions on the largest element in the set $k+j+1$. \square

Theorem 2.23 (Extraction/Absorption). $k \binom{n}{k} = n \binom{n-1}{k-1}$ for $0 < k < n$, or $\binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1}$

Proof. LHS selects committee of size n then selects chair, while RHS selects chair then the rest of the committee. \square

Theorem 2.24. $\binom{n}{k} \binom{k}{m} = \binom{n}{m} \binom{n-m}{k-m}$

Proof. LHS selects committee of size k from n people, then selects m chairs. RHS selects chairs, then the rest of the committee. \square

Theorem 2.25. $\binom{n}{m} \binom{n-m}{k} = \binom{n}{k} \binom{n-k}{m}$

Proof. LHS selects committee of size m then a different committee of size k , while RHS selects the committees in the reverse order. \square

Theorem 2.26. $\sum_{k=1}^n k \binom{n}{k} = n2^{n-1}$

Proof. Apply extraction/absorption then the binomial theorem. Or, LHS selects committee of any size then selects chairs, while RHS selects chair then committee of any size. \square

Theorem 2.27. $\sum_{k=2}^n k(k-1) \binom{n}{k} = n(n-1)2^{n-2}$

Proof. LHS constructs committee of any size from n people and then selects a president and vice president, while RHS selects the president, vice president, and then fills the rest of the committee. \square

Example 2.28. $\sum_{k=0}^n \binom{n}{k} (-1)^k \frac{2^k}{k+1} = \frac{1}{n+1} I[n \text{ is even}]$

Proof. Using Extraction/Absorption and the Binomial Theorem:

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} (-1)^k \frac{2^k}{k+1} &= \sum_{k=0}^n \binom{n+1}{k+1} (-1)^k \frac{2^k}{n+1} = \frac{1}{n+1} \sum_{k=0}^n \binom{n+1}{k+1} (-2)^k \\ &= -\frac{1}{2(n+1)} \sum_{k=1}^{n+1} \binom{n+1}{k} (-2)^k = -\frac{1}{2(n+1)} \left[\sum_{k=0}^{n+1} \binom{n+1}{k} (-2)^k - 1 \right] \\ &= \frac{1}{n+1} \frac{1 - (-1)^{n+1}}{2} = \frac{1}{n+1} \frac{1 + (-1)^n}{2} = \frac{1}{n+1} I[n \text{ is even}] \end{aligned} \quad \square$$

Theorem 2.29 (Vandermonde's Identity). $\sum_{j=0}^k \binom{n}{j} \binom{m}{k-j} = \binom{n+m}{k}$

Proof. RHS selects committee of size k from n women and m men, while LHS conditions on the number of women j .

Or, consider a lattice formed by k rows of squares and $n + m - k$ columns. RHS counts paths from the lower left $(0, 0)$ to the upper right $(k, n + m - k)$ which only move up or right, by picking among the $m + n$ moves which k are vertical. LHS conditions on which point $(j, n - j)$ along the diagonal is passed through; there are $\binom{n}{j}$ ways to reach this point and $\binom{m}{k-j}$ ways to complete the path. \square

Theorem 2.30. $\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}$

Proof. Apply Vandermonde's Identity to obtain:

$$\sum_{k=0}^n \binom{n}{k}^2 = \sum_{k=0}^n \binom{n}{k} \binom{n}{n-k} = \binom{2n}{n} \quad \square$$

Theorem 2.31. $\sum_{k=1}^n k \binom{n}{k}^2 = n \binom{2n-1}{n-1}$

Proof. Apply Absorption/Extraction and Vandermonde's Identity to obtain:

$$\begin{aligned} \sum_{k=1}^n k \binom{n}{k}^2 &= \sum_{k=1}^n n \binom{n-1}{k-1} \binom{n}{k} = n \sum_{k=0}^{n-1} \binom{n-1}{k} \binom{n}{k+1} \\ &= n \sum_{k=0}^{n-1} \binom{n-1}{k} \binom{n}{n-k-1} = n \binom{2n-1}{n-1} \quad \square \end{aligned}$$

Theorem 2.32. $\sum_{i=0}^k (-1)^i \binom{n}{i} = (-1)^k \binom{n-1}{k}$

Proof. Using generating functions with Wilf V, as well as the Generalized Binomial Theorem:

$$\begin{aligned} \sum_{i=0}^k (-1)^i \binom{n}{i} &= [x^k] \frac{1}{1-x} \sum_{i=0}^{\infty} (-1)^i \binom{n}{i} x^i = [x^k] \frac{1}{1-x} (1-x)^n = [x^k] (1-x)^{n-1} \\ &= [x^k] \sum_{i=0}^{\infty} (-1)^i \binom{n-1}{i} x^i = (-1)^k \binom{n-1}{k} \quad \square \end{aligned}$$

Example 2.33. For $2 \leq k < n$, we have $k^n < \binom{kn}{n}$.

Proof. Consider a $n \times k$ grid. k^n chooses one entry from each column, while $\binom{kn}{n}$ may choose any n entries in the grid. \square

2.6 Properties of Sequence of Binomial Coefficients

Definition 2.34. A finite sequence $\{a_k\}_{k=1}^n$ is unimodal if there exists m such that $a_1 \leq \dots \leq a_m$ and $a_m \geq \dots \geq a_n$. The sequence is furthermore log-concave if $\frac{a_k}{a_{k-1}} \frac{a_{k+1}}{a_k} = a_k^2 - a_{k-1}a_{k+1} \geq 0$ for all k .

↳ Log-concave sequences have $\frac{a_{k+1}}{a_k} \leq \frac{a_k}{a_{k-1}}$, so that the ratios are monotone decreasing

↳ Using the above interpretation, it is easy to show that any log-concave sequence is indeed unimodal

Theorem 2.35. $\binom{n}{k}$ is unimodal in k .

Proof. We prove that it is monotone increasing when $k \leq \lfloor \frac{n}{2} \rfloor$, then the remainder follows by symmetry. Consider associating elements in S_{k-1} with those in S_k if the former is a subset of the latter. Then every element of S_{k-1} is associated to $n - k + 1$ elements in S_k , and every element in S_k is associated to $k - 1$ elements in S_{k-1} . Thus, $|S_{k-1}|(n - k + 1) = |S_k|(k - 1)$. We have

$n - k + 1 \geq k - 1$, so that $|S_{k-1}| \leq |S_k|$. □

Theorem 2.36. $\binom{n}{k}$ is log-concave in k (this follows algebraically).

2.7 Multinomial Theorem

Theorem 2.37 (Multinomial Theorem). $(x_1 + \cdots + x_k)^n = \sum \binom{n}{n_1 \cdots n_k} x_1^{n_1} \cdots x_k^{n_k}$ where the summation is taken over all nonnegative $n_1 + \cdots + n_k = n$

↳ This follows similarly to the Binomial Theorem

Theorem 2.38. $\binom{n}{n_1 \cdots n_k} = \binom{n}{n_1} \binom{n-n_1}{n_2} \cdots \binom{n-n_1-\cdots-n_{k-1}}{n_k}$

Proof. LHS is distinguishable permutations of $\{1^{n_1}, \dots, k^{n_k}\}$, while RHS successively picks the locations for all 1s, then all 2s, and so on. □

Theorem 2.39. $\sum_{a+b+c=n} a \binom{n}{a \ b \ c} = n3^{n-1}$

Proof. LHS counts number of ways to form three groups of size a, b, c from n people and then select a chair in the first group. RHS selects the chair first then distributes the remaining people. □

2.8 Generalized Binomial Theorem

Definition 2.40. The generalized binomial coefficient $\binom{\alpha}{k}$ for $\alpha \in \mathbb{R}$ and $k \in \mathbb{Z}$ is 0 for $k < 0$, 1 for $k = 1$, and $\frac{(\alpha)_k}{k!}$ for $k > 0$.

Theorem 2.41 (Generalized Binomial Theorem). $(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$ as a formal power series

↳ This follows by computing the Taylor series

↳ This is often useful when working with generating functions

Example 2.42.

$$\begin{aligned} \sqrt{1-4x} &= \sum_{n=0}^{\infty} \binom{1/2}{n} (-4x)^n = 1 - 2x + \sum_{n=2}^{\infty} \frac{1}{2n!} \left(-\frac{1}{2}\right)_{n-1} (-4x)^n = 1 - 2x - \sum_{n=2}^{\infty} \frac{1}{2n!} \left(\frac{1}{2}\right)^{\overline{n-1}} (4x)^n \\ &= 1 - 2x - \sum_{n=2}^{\infty} \frac{4^n}{2n!} \frac{(2n-3)!!}{2^{n-1}} x^n = 1 - 2x - \sum_{n=2}^{\infty} \frac{2^n (2n-3)!!}{n!} x^n \\ &= 1 - 2x - \sum_{n=2}^{\infty} \frac{2^n (2n-3)!! (2n-2)!!}{n! 2^{n-1} (n-1)!} x^n = 1 - 2x - \sum_{n=2}^{\infty} \frac{2(2n-2)!}{n!(n-1)!} x^n \end{aligned}$$

Example 2.43. We can prove Vandermonde's identity using the Generalized Binomial Theorem.

$$\begin{aligned} \sum_{k=0}^{m+n} \binom{m+n}{k} x^k &= (1+x)^{m+n} = (1+x)^m (1+x)^n \\ &= \left[\sum_{k=0}^m \binom{m}{k} x^k \right] \left[\sum_{j=0}^n \binom{n}{j} x^j \right] = \sum_{k=0}^{\infty} \sum_{j=0}^k \binom{m}{j} \binom{n}{k-j} x^k \end{aligned}$$

Theorem 2.44. $\binom{\alpha}{n} = (-1)^n \binom{n-\alpha-1}{n}$

Proof.

$$\binom{\alpha}{n} = \frac{(\alpha)_n}{n!} = \frac{(-\alpha)\bar{n}(-1)^n}{n!} = \frac{(n-\alpha-1)_n(-1)^n}{n!} = (-1)^n \binom{n-\alpha-1}{n} \quad \square$$

Theorem 2.45. $\sum_{n=0}^{\infty} \binom{n}{k} x^n = \frac{x^k}{(1-x)^{k+1}}$

Proof. Using Theorem 2.44 and the Generalized Binomial Theorem:

$$\begin{aligned} \sum_{n=0}^{\infty} \binom{n}{k} x^n &= \sum_{n=k}^{\infty} \binom{n}{k} x^n = \sum_{n=k}^{\infty} \binom{n}{n-k} x^n = \sum_{n=k}^{\infty} (-1)^{n-k} \binom{-k-1}{n-k} x^n \\ &= x^k \sum_{n=k}^{\infty} (-1)^{n-k} \binom{-k-1}{n-k} x^{n-k} = x^k \sum_{n=0}^{\infty} \binom{-k-1}{n} (-x)^n = x^k (1-x)^{-k-1} \\ &= \frac{x^k}{(1-x)^{k+1}} \quad \square \end{aligned}$$

Theorem 2.46. $(1-x)^{-m} = \sum_{n=0}^{\infty} \binom{m+n-1}{n} x^n$

Proof. By the Generalized Binomial Theorem and Theorem 2.44:

$$(1-x)^{-m} = \sum_{n=0}^{\infty} \binom{-m}{n} (-x)^n = \sum_{n=0}^{\infty} \binom{m+n-1}{n} x^n \quad \square$$

2.9 Subscripted Binomial Coefficients

Definition 2.47. Define $\binom{m}{n}_p = \sum_{k=0}^m \binom{k}{p} \binom{k}{n}$.

↳ Note that this is not a standard definition/construct

Lemma 2.48. $\binom{m}{n}_p$ counts the number of ordered pairs (S, T) where $S \in \binom{[m+1]}{n+1}$, $T \in \binom{[m+1]}{p}$, and $\max T < \max S$.

Proof. Condition on $\max S = k+1$. Select the remaining elements in S $\binom{k}{n}$ ways, and select the elements of T $\binom{k}{p}$ ways. $\max S \in [m+1]$, so that $0 \leq k \leq m$. □

Lemma 2.49. $\binom{m+1}{n}_p = \binom{m}{n}_p + \binom{m+1}{p} \binom{m+1}{n}$

Proof. Use the definition and pull out the term for $k = m+1$. □

Lemma 2.50. $\binom{m}{n}_p = \sum_{j=0}^p \binom{m+1}{p-j} \binom{m+1+j}{n+1+j} (-1)^j$

↳ Substituting $p = 0$ gives $\sum_{k=n}^m \binom{k}{n} = \binom{m+1}{n+1}$, which is the Hockey-Stick Identity

Lemma 2.51. The generating function for $\binom{m}{n}_p$ with respect to n is $A_p(m, x) = \sum_{k=0}^m \binom{k}{p} (1+x)^k$.

Proof.

$$A_p(m, x) = \sum_{n=0}^{\infty} \sum_{k=0}^m \binom{k}{p} \binom{k}{n} = \sum_{k=0}^m \binom{k}{p} \sum_{n=0}^{\infty} \binom{k}{n} = \sum_{k=0}^m \binom{k}{p} (1+x)^k \quad \square$$

3. Permutations

3.1 Inversions

Definition 3.1. If $\pi \in S_n$, let $\pi_j = \pi(j)$. Then π can be represented as:

- 2-line representation: $\begin{pmatrix} 1 & 2 & \cdots & n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix}$
- 1-line representation: $\pi_1 \pi_2 \cdots \pi_n$

Concept 3.2. The inverse of a permutation π can be found with the 2-line representation by exchanging the two rows and then reordering the columns so that the first row is in order.

Definition 3.3. Let $\pi = \pi_1 \cdots \pi_n \in S_n$. A pair (π_i, π_j) for $i < j$ is called an inversion of π if $\pi_i > \pi_j$. Let $E(\pi)$ denote the set of all inversions of π . Then the inversion table π_I of π is constructed as $\pi_I = b_1 b_2 \cdots b_n$ by letting b_k count the number of times that k is the second component in $E(\pi)$ (i.e., how many times k appears as the smaller number in an inversion). The inversion number of π is defined $i(\pi) = |E(\pi)| = \sum_{k=1}^n b_k$.

↳ $0 \leq b_j < n$ for all j , so $b_n = 0$

↳ The inversion table provides a unique representation of permutations. The 1-line representation can be reconstructed as follows. First place n . Then, work through π_I in reverse. For each b_j , insert j so that b_j elements are to the left of it.

Example 3.4. Consider $\pi = 463512 \in S_6$. Then:

$$E(\pi) = \{(4, 3), (4, 1), (4, 2), (6, 3), (6, 5), (6, 1), (6, 2), (3, 1), (3, 2), (5, 1), (5, 2)\}$$

We thus have $\pi_I = 442010$. Now suppose we are given $\pi_I = 442010$. To reconstruct π , first place 6. Since $b_5 = 1$, we obtain 65. Since $b_4 = 0$, place it at the left to obtain 465. Continue to get 4635, 46352, and finally 463512.

Theorem 3.5.

- $E(\pi)$ is transitive, so that if (a, b) and (b, c) are in $E(\pi)$, then $(a, c) \in E(\pi)$.
- Let E be any transitive subset of $T = \{(x, y) \mid 1 \leq y < x \leq n, x, y \in [n]\}$ whose complement $E^c = T \setminus E$ is also transitive. Then there is a permutation π such that $E(\pi) = E$.

Proof.

- $(a, b), (b, c) \in E(\pi)$ implies a comes before b and b comes before c in the one-line representation. This implies a comes before c in the one-line representation, so that $(a, c) \in E(\pi)$.
- Consider a directed graph with vertices $\{1, \dots, n\}$. Add an edge $j \rightarrow i$ if $(j, i) \in E$. Add an edge $i \rightarrow j$ if $(j, i) \in E^c$. This graph represents the order in which the elements $\{1, \dots, n\}$ must appear in the one-line representation of π , where edges go from earlier elements to later elements. A permutation realizing this graph exists iff the graph is acyclic (indeed, perform topological sorting). Suppose for contradiction that it was cyclic. By transitivity of red edges and green edges, we can find a sub-cycle such that the edges alternate color. If $v_1 \rightarrow v_2 \rightarrow v_3$ or $v_1 \rightarrow v_2 \leftarrow v_3$, then we must have $v_1 \rightarrow v_3$ for some color, otherwise $v_1 \leftarrow v_3$ for any color yields a contradiction with transitivity. Thus, we can find a smaller sub-cycle as long as the number of vertices in the cycle is at least 3. Continuing, we obtain a cycle with fewer than 3 elements, which is a contradiction, since no self-edges or bidirectional edges are possible. Now we can construct π via topological sorting, and $E(\pi) = E$ by definition of E . \square

Theorem 3.6. The inversion number of π equals the inversion number of π^{-1} , so that $i(\pi) = i(\pi^{-1})$.

Proof. $(\pi_i, \pi_j) \in E(\pi) \iff (\pi_i, \pi_j)$ appears in an opposite order as (i, j) in the 2-line representation $\iff (i, j) \in E(\pi^{-1})$. Thus, we have a bijection from $E(\pi)$ to $E(\pi^{-1})$, so that $i(\pi) = i(\pi^{-1})$. \square

Definition 3.7. A permutation $\pi \in S_n$ can be represented via an inversion grid. On an $n \times n$ grid, place \bullet in cell (j, k) whenever $\pi_j = k$ (so that each element of the permutation is represented by a row). Place \times in any cell where there exists a \bullet directly below it and another \bullet directly right of it, so that \times in cell (i, j) means that (π_i, π_j) is an inversion. The number of \times in column j is equal to b_j , and $|\pi_I|$ is the total number of \times . The transpose of an inversion grid is formed by transposing the grid, and is the inversion grid of π^{-1} .

For example, the inversion grid (and its transpose) of $\pi = 531264$ is:

\times	\times	\times	\times	\bullet	
\times	\times	\bullet			
\bullet					
	\bullet				
			\times		\bullet
			\bullet		

Inversion Grid for π

\times	\times	\bullet			
\times	\times		\bullet		
\times	\bullet				
\times				\times	\bullet
\bullet					
				\bullet	

Inversion Grid for π^{-1} (transpose)

↳ Counting the number of \times symbols directly gives $i(\pi) = i(\pi^{-1})$

Definition 3.8. Let $I_n(k)$ count the number of permutations in S_n with exactly k inversions. Define $I_0(0) = 1$.

Theorem 3.9.

- (a) $I_n(j) > 0$ iff $0 \leq j \leq \binom{n}{2}$
- (b) $I_n\left(\binom{n}{2}\right) = 1$
- (c) $I_n(1) = n - 1$
- (d) $I_n(k) = I_n\left(\binom{n}{2} - k\right)$, so that $I_n(k)$ is symmetric

Proof.

- (a) A permutation in S_n with exactly $0 \leq j \leq \binom{n}{2}$ inversions can be inductively constructed as follows. If $j \leq \binom{n-1}{2}$, then choose a permutation in S_{n-1} with j inversions and add n at the end. Otherwise, pick $(n-1 \cdots 1) \in S_{n-1}$ and then add n so that $j - \binom{n-1}{2} \leq n-1$ elements are to its right.
- (b) A permutation $\pi \in S_n$ has $\binom{n}{2}$ inversions if every (π_i, π_j) is an inversion, which occurs iff the permutation is $\pi = (n \ n-1 \cdots 2 \ 1)$.
- (c) There is one inversion in $\pi \in S^n$ iff precisely one pair of adjacent entries in $(1 \ 2 \cdots n)$ is swapped, which can be done $n-1$ ways.
- (d) Compose the permutation $\pi \in S_n$ with $\sigma : (n \ n-1 \cdots 2 \ 1)$ to obtain an involution from permutations with k inversions to permutations with $\binom{n}{2} - k$ inversions. \square

Theorem 3.10. $I_n(k) = \sum_{j=0}^{n-1} I_{n-1}(k-j)$

↳ Given a permutation in S_{n-1} with k inversions, inserting n to form a permutation in S_n can produce any of $\{k, k+1, \dots, k+n-1\}$ inversions depending on the placement of n

Proof. We form the bijection from permutations in S_n with k inversions and permutations in S_{n-1} with between $k - n + 1$ and k inversions. In the forward direction, remove n from the permutation. In the reverse, add n to the permutation in the location that causes k inversions to be present. Specifically, if the permutation has $k - j$ inversions, insert n so that j values are right of it in the one-line representation. \square

Theorem 3.11. Let $J_n(x) = \sum_{k=0}^{\infty} I_n(k)x^k$ be the generating function for $\{I_n(k)\}_{k=0}^{\infty}$. Then $J_n(x) = (1 + x + \cdots + x^{n-1})J_{n-1}(x)$.

Proof. Using the recurrence in Theorem 3.10:

$$\begin{aligned} J_n(x) &= \sum_{k=0}^{\infty} I_n(k)x^k = \sum_{k=0}^{\infty} \sum_{j=0}^{n-1} I_{n-1}(k-j)x^k = \sum_{j=0}^{n-1} \sum_{k=0}^{\infty} I_{n-1}(k-j)x^k \\ &= \sum_{j=0}^{n-1} x^j \sum_{k=0}^{\infty} I_{n-1}(k-j)x^{k-j} = \sum_{j=0}^{n-1} x^j J_{n-1}(x) = \left(\sum_{j=0}^{n-1} x^j \right) J_{n-1}(x) \end{aligned} \quad \square$$

Corollary 3.12. The generating function for $\{I_n(k)\}_{k=0}^{\infty}$ is:

$$J_n(x) = \frac{(1-x^n)(1-x^{n-1}) \cdots (1-x^2)(1-x)}{(1-x)^n}$$

Proof. Write $(1 + x + \cdots + x^{n-1}) = \frac{1-x^n}{1-x}$ since it is the partial sum of a geometric series, and then apply the recursion in Theorem 3.11. \square

Definition 3.13. The sign of a permutation π is $\text{sgn}(\pi) = (-1)^{i(\pi)}$. If the sign is 1 (so that there are an even number of inversions), then the permutation is called even. If the sign is -1 (so that there are an odd number of inversions), then the permutation is called odd.

- ↳ Composing a simple transposition $\tau = (k, k+1)$ with π causes it to switch sign. This is shown by forming a bijection $E(\pi) \setminus (k+1, k) \leftrightarrow E(\tau\pi) \setminus (k+1, k)$, then noting that precisely one of $E(\pi)$ or $E(\tau\pi)$ contains $(k+1, k)$, so that $E(\tau\pi) = E(\pi) \pm 1$.
- ↳ Any transposition is the composition of an odd number of simple transpositions. This is shown by induction since $(i, i+k+1) = (i, i+k)(i+k, i+k+1)(i, i+k)(i+k, i+k+1)$.
- ↳ Thus, applying any transposition switches the sign of a permutation.
- ↳ Any cycle of length j can be written as a product of $j-1$ transpositions. This is shown by induction since $(a_1, \dots, a_j) = (a_1, a_j)(a_1, \dots, a_{j-1})$.
- ↳ Thus, a cycle of length j is even iff j is odd.
- ↳ sgn is a group homomorphism, so that $\text{sgn}(\pi\delta) = \text{sgn}(\pi)\text{sgn}(\delta)$. This is shown by writing the cycle decomposition of each and then writing each cycle as transpositions.
- ↳ Using the cycle decomposition for a permutation, we have that a permutation is even if it can be written as an even number of transpositions, and odd if it can be written as an odd number of transpositions.

3.2 Cycles

Lemma 3.14. Let $\pi \in S_n$ and $x \in [n]$. Then there exists a positive $1 \leq k \leq n$ such that $\pi^k(x) = x$.

Proof. Consider $\{\pi(x), \pi^2(x), \dots, \pi^n(x)\}$. Either one is equal to x (in which case we are done), or by PHP $\pi^i(x) = \pi^j(x)$ for $i < j$. Apply π^{-i} to each side. \square

Definition 3.15. Let $\pi = (\pi_1 \ \pi_2 \ \dots \ \pi_n) \in S_n$. For some $j \in [n]$, suppose $\pi(j), \pi^2(j), \dots, \pi^k(j)$ are distinct and $\pi^k(j) = j$. This sequence is called a cycle. A permutation can be written in cycle notation as a product of distinct cycles. For example, $\pi = 145263$ can be written as $(1)(24)(356)$. Each cycle should start with its smallest element, and the collection of cycles is ordered by the smallest element in each cycle.

↳ Every element belongs to a cycle by Lemma 3.14

↳ The singletons are sometimes dropped from the notation

↳ There is one permutation in S_n with n cycles (the identity) and $(n-1)!$ with one cycle

Theorem 3.16. Let c_1, \dots, c_n be nonnegative integers with $\sum_{j=1}^n jc_j = n$. Then the number of permutations from S_n with c_j cycles of length j is:

$$\frac{n!}{c_1!c_2! \dots c_n! 1^{c_1} 2^{c_2} \dots n^{c_n}}$$

Proof. Let C be the set of permutations with the given number of cycles. We map $S_n \rightarrow C$ by inserting parentheses from left to right to form all one-cycles, then all two-cycles, and so on. Each permutation in C has $c_1!c_2! \dots c_n! 1^{c_1} 2^{c_2} \dots n^{c_n}$ preimages, since all cycles of length c_i may be permuted $c_i!$ ways, and each cycle of length j has j equivalent ways to write it. \square

Example 3.17.

(a) The number of permutations with exactly one 3-cycle and two 4-cycles is $\frac{11!}{1!2!3!4^2}$

(b) The number of permutations with exactly one cycle in S_n is $\frac{n!}{1!n!} = (n-1)!$

Example 3.18. We can also show that the number of permutations in S_n with exactly one cycle is $(n-1)!$ via a combinatorial argument. Write the cycle in cycle notation such that n is the leftmost element. Map this bijectively to a permutation in S_{n-1} (written in one-line notation) by removing the character n and the parentheses.

Definition 3.19. Let $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ denote the set of all n -permutations with exactly k cycles. Let $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right] = 1$ and $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] = \left| \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] \right|$ otherwise. These are signless Stirling number of the first kind. These are also denoted $c(n, k)$.

↳ This counts the number of ways to seat n knights at k unlabeled tables, so that no table is empty

Definition 3.20. The signed Stirling numbers of the first kind are given by $s(n, k) = (-1)^{n-k} \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$.

↳ This is useful when looking at Stirling inversion

Example 3.21. $\left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right] = \{(1)(234), (1)(243), (12)(34), (13)(24), (14)(23), (123)(4), (132)(4), (124)(3), (142)(3), (134)(2), (143)(2)\}$, so that $\left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right] = 11$.

Proposition 3.22. $\sum_{k=0}^n \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] = n!$

Proof. RHS counts permutations of $[n]$, while LHS conditions on the number of cycles k . \square

Proposition 3.23. For $n \geq k \geq l$, $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] = \left[\begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right] + (n-1) \left[\begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right]$.

Proof. RHS conditions on whether n is a fixed point (its own cycle). If so, permute the remaining $n-1$ elements in $k-1$ cycles. If not, permute the $n-1$ elements in k cycles and then pick which of the $n-1$ elements n should go after (indeed, for each cycle of length p , there are p elements n could go after, so sum over all cycles). \square

Proposition 3.24. The generating function for $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ with respect to k is $g_n(x) = \sum_{k=0}^{\infty} \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] x^k = x^{\overline{n}}$.

Proof. When $n = 0$, we have $g_0(x) = 1$ since $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1$ and $\begin{bmatrix} 0 \\ k \end{bmatrix} = 0$ for all $k \geq 1$. Otherwise:

$$\begin{aligned} g_n(x) &= \sum_k \begin{bmatrix} n \\ k \end{bmatrix} x^k = \sum_k \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} x^k + (n-1) \sum_k \begin{bmatrix} n-1 \\ k \end{bmatrix} x^k = xg_{n-1}(x) + (n-1)g_{n-1}(x) \\ &= g_{n-1}(x) \cdot (x + n - 1) \end{aligned}$$

Thus, by induction, we have $g_n(x) = x \cdot (x+1) \cdots (x+n-1) = x^{\overline{n}}$. \square

Corollary 3.25. $\sum_k (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} x^k = (x)_n$.

Proof. Begin with the generating function, then substitute $-x$ for x and rearrange. \square

Corollary 3.26. $\begin{bmatrix} n \\ k \end{bmatrix} = \sum_{0 < j_1 < \cdots < j_{n-k} < n} j_1 j_2 \cdots j_{n-k}$

Proof. By Proposition 3.24, we have $\sum_{k=0}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} x^k = x^{\overline{n}}$. Apply $[x^k]$ to each side to get $\begin{bmatrix} n \\ k \end{bmatrix} = [x^k]x^{\overline{n}} = [x^k](x(x+1)(x+2)\cdots(x+n-1)) = [x^{k-1}]((x+1)(x+2)\cdots(x+n-1))$. Consider distributing the above factors. For any resulting multiple of x^{k-1} , it must have originated from $k-1$ of the factors being x and $n-k$ of the factors being distinct integers in $[n-1]$. Sum over all such products of $n-k$ integers to obtain the result. \square

Proposition 3.27. The exponential generating function for $\begin{bmatrix} n \\ k \end{bmatrix}$ with respect to n is given by $C_k(x) = \sum_{n=0}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} \frac{x^n}{n!} = \frac{1}{k!} \left(\log \frac{1}{1-x} \right)^k$.

Proof. Consider the labeled structure $[\cdot]_k$. This is equivalent to partitioning the labels L into n nonempty blocks, building a cycle on each block, and then building the structure of sets of size k on the blocks. Within each block, a cycle is formed $(n-1)!$ ways with EGF $\log \frac{1}{1-x}$, and the blocks are formed into a k -set $I[n=k]$ ways, which has EGF $x^k/k!$. Compose these to obtain the result by Theorem 9.10.

Alternatively, we can count the number of ways to form ordered cycles via $[\cdot]_1^k$, then remove the order by dividing by $k!$. The exponential generating function for $[\cdot]_1$ is $\log \frac{1}{1-x}$, so that we obtain $\frac{1}{k!} \left(\log \frac{1}{1-x} \right)^k$ as claimed.

We can also prove this using the recurrence in Proposition 3.23. Rewrite this as $\begin{bmatrix} n+1 \\ k \end{bmatrix} = \begin{bmatrix} n \\ k-1 \end{bmatrix} + n \begin{bmatrix} n \\ k \end{bmatrix}$. Then using Wilf I and II, $C'_k(x) = C_{k-1}(x) + xC'_k(x)$. Rearrange to obtain $C'_k(x) = \frac{1}{1-x} C_{k-1}(x)$, so by induction $C'_k(x) = \frac{1}{1-x} \frac{1}{(k-1)!} (\log(\frac{1}{1-x}))^{k-1}$. Integrate to obtain $C_k(x) = \frac{1}{(k-1)!} \int \frac{1}{1-x} (\log(\frac{1}{1-x}))^{k-1} dx$. Let $u = \log \frac{1}{1-x}$, so that $du = \frac{1}{(1-x)^2} dx = \frac{1}{1-x} dx$. Then $C_k(x) = \frac{1}{(k-1)!} \int u^{k-1} du = u^k/k! + c$.

When $k = 0$ and $x = 0$, we have $u = 0$ so that $u^k/k! = 1$. Since $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1$, this means $c = 0$. When $k > 0$ and $x = 0$, we have $u = 0$ and $u^k/k! = 0$. Since $\begin{bmatrix} 0 \\ k \end{bmatrix} = 0$, this also means $c = 0$. Thus, $C_k(x) = u^k/k!$, as claimed. \square

Proposition 3.28. $\begin{bmatrix} n \\ 2 \end{bmatrix} = \frac{n!}{2} \sum_{m=1}^{n-1} \frac{1}{m(n-m)}$

Proof. Rewrite this as $\begin{bmatrix} n \\ 2 \end{bmatrix} = \sum_{m=1}^{n-1} \frac{n!}{2} \frac{1}{m(n-m)}$. LHS counts the number of permutations in S_n with two cycles. RHS counts the same thing via over-counting. First count all permutations of $[n]$, $n!$ ways. Now group the string by placing the first m into a group and the remaining $n-m$ into another group. Of all such ways to do this, precisely $\frac{1}{2m(n-m)}$ ways yield a valid permutation in canonical form. Indeed, there are m ways to rotate the first group, $n-m$ ways to rotate the second, and 2 ways to order the cycles, and all together only one combination of these is a valid permutation in

canonical form. Thus, by summing over all m , we have our result. \square

Proposition 3.29. $\sum_k \binom{n}{k} (-1)^k = 0$ for $n \geq 2$

Proof. We form an involution between permutations with an odd number of cycles and those with an even number of cycles. If 1 and 2 belong to different cycles, then concatenate these cycles; for example, $(134)(25)$ becomes (13425) . If 1 and 2 belong to the same cycle, then split the cycle at 2; for example, (13425) becomes $(134)(25)$. This will always result in a valid permutation in canonical form. This is clearly an involution, and it changes the parity of the number of cycles, hence we have the result. \square

3.3 Involutions

Lemma 3.30. Let i_n count the number of involutions in S_n (i.e., permutations with $\pi^2 = \text{id}$). Then $i_{n+2} = i_{n+1} + (n+1)i_n$.

Proof. The RHS conditions on whether $n+2$ is a fixed point. If it is, then there are i_{n+1} ways to define the involution on the remaining $n+1$ numbers. If it is not, then it must be in a 2-cycle with any of the other $n+1$ numbers, then the remaining n numbers can form an involution i_n ways. \square

Lemma 3.31. The exponential generating function for $\{i_n\}$ is $I(x) = e^{x+x^2/2}$.

Proof. Use Lemma 3.30 and Wilf I-II for exponential generating functions to obtain $I''(x) = I'(x) + xI'(x) + I(x)$. Simplify to obtain $I''(x) = (1+x)I'(x) + I(x)$. It is not particularly clear how to solve this, but it is easy to see that the given $I(x)$ satisfies it.

Alternatively, use labeled structures. Involutions can be counted by partitioning the labels, then forming a 1- or 2-cycle on each block. The number of ways to form the cycles is represented by the EGF $x + x^2/2$, so that the exponential formula gives $e^{x+x^2/2}$ as claimed. \square

Lemma 3.32. Let j_n count the number of permutations of $[n]$ that are both involutions and derangements. Then $j_{2n+1} = 0$ and $j_{2n} = (2n-1)!!$. The exponential generating function is $J(x) = e^{x^2/2}$.

Proof. All involutions are composed of 1-cycles and 2-cycles, but a derangement has no 1-cycles. Thus, we must only have 2-cycles, so that n must be even. In other words, $j_{2n+1} = 0$. By Theorem 3.16, $j_{2n} = \frac{(2n)!}{n!2^n} = (2n-1)!!$.

We can use labeled structures to find the generating function of the j_n . Consider the labeled structure $\mathcal{S}(L)$ which form a permutation consisting solely of 2-cycles on the labels L . This can equivalently be formed by partitioning the labels then forming a 2-cycle on each partition. The exponential generating function for the number of ways to form a 2-cycle is $x^2/2$, so that the exponential generating function $J(x)$ is $e^{x^2/2}$.

Using this generating function, we can re-derive the values of j_n . Indeed:

$$j_n = n![x^n]J(x) = n![x^n] \sum_{k=0}^{\infty} \frac{(x^2/2)^k}{k!} = n![x^n] \sum_{k=0}^{\infty} \frac{x^{2k}}{2^k k!}$$

When n is odd, this is clearly 0. Otherwise:

$$j_{2n} = (2n)![x^{2n}] \sum_{k=0}^{\infty} \frac{x^{2k}}{2^k k!} = \frac{(2n)!}{2^n n!} = (2n-1)!! \quad \square$$

Lemma 3.33. The number of permutations which are derangements and whose cube is the identity is given by the EGF $e^{2x^3/6}$. If we relax this to allow fixed points, the EGF is $e^{x+2x^3/6}$.

Proof. Consider the labeled structure associated with the first problem. It can be formed by partitioning the labels, then forming a 3-cycle on each block. The EGF for forming a 3-cycle is $2x^3/6$, so

that the exponential formula gives the EGF $e^{2x^3/6}$. If we allow for fixed points, the EGF counting the number of ways to form the cycles is $x + 2x^3/6$, so that the exponential formula yields $e^{x+2x^3/6}$. \square

3.4 Derangements

Definition 3.34. A permutation is called a derangement if it has no fixed points.

Theorem 3.35. The number of derangements in S_n is $D_n = \sum_{k=0}^n \binom{n}{k} k! (-1)^{n-k}$.

$\hookrightarrow D_n$ is often denoted $!n$, the subfactorial

Proof. To count the number of derangements D_n in S_n , define $D_{n,k}$ to equal the number of permutations with k fixed points. $D_{n,k} = \binom{n}{k} D_{n-k}$ by the product rule. Conditioning permutations based on their number of fixed points:

$$n! = \sum_{k=0}^n D_{n,k} = \sum_{k=0}^n \binom{n}{k} D_{n-k} = \sum_{k=0}^n \binom{n}{n-k} D_{n-k} = \sum_{k=0}^n \binom{n}{k} D_k$$

Thus, by Binomial Inversion, we have $D_n = \sum_{k=0}^n \binom{n}{k} k! (-1)^{n-k}$.

Or, we can use Inclusion/Exclusion. Letting $N = n!$ and p_j be the property that the permutation fixes j , then $N(p_j) = (n-1)!$, $N(p_j, p_k) = (n-2)!$, and so on. Thus:

$$\begin{aligned} D_n &= n! - \sum_j (n-1)! + \sum_{j < k} (n-2)! - \sum_{j < k < \ell} (n-3)! + \cdots + (-1)^r (n-r)! \\ &= n! - \binom{n}{1} (n-1)! + \binom{n}{2} (n-2)! - \binom{n}{3} (n-3)! + \cdots + (-1)^r \binom{n}{r} (n-r)! \\ &= \sum_{k=0}^n \binom{n}{k} (n-k)! (-1)^k = n! \sum_{k=0}^n \frac{(-1)^k}{k!} = \sum_{k=0}^n \binom{n}{k} k! (-1)^{n-k} \end{aligned} \quad \square$$

Corollary 3.36. The asymptotic fraction of permutations which are derangements is $\lim_{n \rightarrow \infty} \frac{!n}{n!} = e^{-1}$.

Proof.

$$\lim_{n \rightarrow \infty} \frac{!n}{n!} = \lim_{n \rightarrow \infty} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} k! (-1)^{n-k} = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{(n-k)!} (-1)^{n-k} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1} \quad \square$$

Lemma 3.37. $D_{n+1} = n(D_n + D_{n-1})$ for $n \geq 1$.

Proof. Write derangement $\pi \in S_{n+1}$ in cycle notation. There are n options for $\pi_{n+1} \neq n+1$. If $(n+1, \pi_{n+1})$ is a transposition, then there are D_{n-1} ways to permute the remaining elements. Otherwise, removing π_{n+1} is a bijection yielding a derangement in S_n , so that there are D_n options. \square

Lemma 3.38. The number of n -permutations with exactly one cycle of length one is $nD_{n-1} = D_n - (-1)^n$.

Proof. Pick the fixed point n ways, and permute the remaining via a derangement D_{n-1} ways to get $nD_{n-1} = n(n-1)! \sum_{j=0}^{n-1} \frac{(-1)^j}{j!} = n! \sum_{j=0}^{n-1} \frac{(-1)^j}{j!} = D_n - (-1)^n$. \square

Example 3.39. We claim that the number of permutations $p = p_1 \cdots p_n$ so that $p_i \neq i+1$ for all $i \in [n-1]$ is $2D_n - (-1)^n$. Compose any of these permutations p on the left with $\rho = (n, n-1, \dots, 1)$ to obtain a permutation $q = \rho p$ such that $q_i \neq i$ for all $i \in [n-1]$. Either the resulting permutation is a derangement, D_n ways, or it has exactly one fixed point of n , D_{n-1} ways. Since ρ is a bijection, we obtain $D_n + D_{n-1}$ permutations p .

Proposition 3.40. The exponential generating function of the number of derangements D_n is $D(x) = \frac{e^{-x}}{1-x}$. This yields $D_n = n! \sum_{k=0}^n \frac{(-1)^k}{k!}$ as before.

↳ This can be used to obtain the $D_n = n! \sum_{k=0}^n \frac{(-1)^k}{k!}$ as before, since:

$$D_n = n! [x^n] \frac{e^{-x}}{1-x} = n! \sum_{k=0}^n [x^k] e^{-x} = n! \sum_{k=0}^n \frac{(-1)^k}{k!}$$

Proof. $n! = \sum_{k=0}^n \binom{n}{k} D_{n-k}$ by Theorem 3.35. LHS has the EGF $\frac{1}{1-x}$ and by Wilf III, RHS has the EGF $D(x)e^x$. So, $D(x) = \frac{e^{-x}}{1-x}$.

Alternatively, we can use labeled structures and the partition structure. A derangement can be formed uniquely by partitioning the elements in $[n]$, and then forming a one-cycle of length at least 2 on each partition. The labeled structure for forming a cycle is $\begin{bmatrix} \cdot \\ 1 \end{bmatrix}$, which has exponential generating function $\log \frac{1}{1-x}$. Now excluding 1-cycles, we obtain the exponential generating function $\log \frac{1}{1-x} - x$. Applying the permutation structure, we obtain $D(x) = \exp \left(\log \frac{1}{1-x} - x \right) = \frac{e^{-x}}{1-x}$. \square

Proposition 3.41. Let $D_e(n)$ and $D_o(n)$ be the number of even and odd derangements of length n . Then $D_e(n) - D_o(n) = (-1)^{n-1}(n-1)$.

Proof. To count $D_e(n)$, condition on whether n appears in a 2-cycle. If so, then the other element in that cycle can be chosen $n-1$ ways, and the remaining elements are permuted $D_o(n-2)$ ways. Otherwise, we form an odd derangement $D_o(n-1)$ ways and append n in one of the cycles $n-1$ ways. Altogether, $D_e(n) = (n-1)[D_o(n-1) + D_o(n-2)]$, and similarly $D_o(n) = (n-1)[D_e(n-1) + D_e(n-2)]$. Letting $a_n = D_e(n) - D_o(n)$, subtracting yields $a_n = -(n-1)[a_{n-1} + a_{n-2}]$. Rewrite as $a_{n+2} = -(n+1)[a_{n+1} + a_n]$.

$a_1 = 0$ and $a_2 = -1$, which gives us $a_0 = 1$ after solving. Form the ordinary generating function $A(x)$, which satisfies $\frac{A(x)-1}{x^2} = -(xD+1)\left[\frac{A(x)-1}{x} + A(x)\right]$ by Wilf I and II. Simplify to obtain $A'(x) + \frac{x^2+1}{x^3+x^2}A(x) = \frac{1}{x^3+x^2}$. To solve, use integrating factors. Use partial fraction decomposition to write $\frac{x^2+1}{x^3+x^2} = \frac{x^2+1}{x^2(x+1)}$ as $-\frac{1}{x} + \frac{1}{x^2} + \frac{2}{x+1}$, then integrate to obtain $-\ln|x| - \frac{1}{x} + 2\ln|x+1|$. The integrating factor is thus $-x(x+1)^2e^{-1/x}$, so multiply to obtain $D\left[\frac{(x+1)^2}{x}e^{-1/x}A(x)\right] = \frac{x+1}{x^3}e^{-1/x}$. Integrate to obtain the solution $A(x) = \frac{1}{(x+1)^2}[2x+1+cx e^{1/x}]$.

$\frac{1}{x}$ does not admit a power series, so assume $c = 0$ (we verify this yields the true answer at the end). Then $A(x) = \frac{2x+1}{(x+1)^2}$, or $A(x) = \frac{2}{x+1} - \frac{1}{(x+1)^2}$ using partial fraction decomposition. Thus:

$$a_n = [x^n]A(x) = 2[x^n]\frac{1}{x+1} - [x^n]\frac{1}{(x+1)^2} = 2(-1)^n - (-1)^n(n+1) = (-1)^{n-1}(n-1)$$

We show $c = 0$ is correct by verifying that the above form satisfies the recurrence:

$$a_{n+2} = (-1)^{n+1}(n+1) = -(n+1)[(-1)^n] = -(n+1)[(-1)^n n + (-1)^{n-1}(n-1)] = -(n+1)[a_{n+1} + a_n]$$

\square

4. Principal of Inclusion/Exclusion

Theorem 4.1 (Principal of Inclusion/Exclusion). Let S be a collection of N objects and let $P = \{p_1, \dots, p_r\}$ be a set of properties that one or more of the elements in S possess. Let $N(p_{i_1}, \dots, p_{i_n})$ count the number of objects that satisfy at least the properties p_{i_1}, \dots, p_{i_n} , and let P_i be the set of elements satisfying p_i . Let e_j count the number of objects with exactly j properties. Then:

$$e_0 = \underbrace{N}_{N_0} - \underbrace{\sum_{j=1}^r N(p_j)}_{N_1} + \underbrace{\sum_{j < k} N(p_j, p_k)}_{N_2} - \underbrace{\sum_{j < k < \ell} N(p_j, p_k, p_\ell)}_{N_3} + \dots + (-1)^r \underbrace{N(p_1, \dots, p_r)}_{N_r}$$

Alternatively, the number of objects satisfying at least one property is:

$$\left| \bigcup_{j=1}^r P_j \right| = N_1 - N_2 + \dots + (-1)^{r+1} N_r$$

The number satisfying at least j properties is similarly computed, and by taking differences, the number satisfying exactly j properties can be computed. Or, using generating functions $E(x) = \sum_k e_k x^k$ and $N(x) = \sum_k N_k x^k$, we have $E(x) = N(x-1)$, so that expanding $N(x-1)$ gives us all values of e_j .

Proof. We have $N_k = \sum_{j=k}^r \binom{j}{k} e_j$ by conditioning on the number of properties j and then picking k to use. Then the generating functions satisfy:

$$N(x) = \sum_{k=0}^r N_k x^k = \sum_{k=0}^r \sum_{j=k}^r \binom{j}{k} e_j x^k = \sum_{j=0}^r e_j \sum_{k=0}^j \binom{j}{k} x^k = \sum_j e_j (1+x)^j = E(1+x)$$

Then $E(x) = N(x-1)$. Substituting $x=0$ yields $e_0 = N(-1) = \sum_k N_k (-1)^k$. \square

Example 4.2. The number of nonnegative integer solutions to $x_1 + x_2 + x_3 = 17$ is given by the number of weak compositions as $\binom{3}{17} = \binom{19}{2} = 171$. If we add the restriction that $x_1 \leq 6$, $x_2 \leq 10$, and $x_3 \leq 8$, then this can be counted using Inclusion/Exclusion with properties $p_1 = (x_1 > 6)$, $p_2 = (x_2 > 10)$, and $p_3 = (x_3 > 8)$. Allocating the 7 guaranteed parts to x_1 , we obtain that $N(p_1) = \binom{3}{17-7} = \binom{3}{10} = 66$, and similarly $N(p_2) = \binom{3}{6} = 28$, $N(p_3) = \binom{3}{8} = 45$, $N(p_1 p_2) = 0$, $N(p_1 p_3) = \binom{3}{1} = 3$, $N(p_2 p_3) = 0$, and $N(p_1 p_2 p_3) = 0$. Then the total number of given by $171 - 66 - 28 - 45 + 3 = 35$.

Concept 4.3. An equivalent formulation that uses sets is as follows. Let S be a finite set and $S_1, \dots, S_n \subset S$. Then:

$$\left| S \setminus \bigcup_{i=1}^n S_i \right| = |S| - \sum_{i=1}^n |S_i| + \sum_{1 \leq i < j \leq n} |S_i \cap S_j| - \dots + (-1)^n \left| \bigcap_{i=1}^n S_i \right|$$

Proof. Rather than viewing properties on a collection of N objects, this form enumerates the elements of S satisfying property i in the subset S_i . With this realization, the equivalence is clear.

Using this form (and the language from the first form), we now reprove the claim using Möbius Inversion. Define $f, g : B_n \rightarrow \mathbb{R}$ where B_n is the Boolean algebra on $[n]$ and:

$$f(I) = \left| \bigcap_{i \in I} S_i \right| \quad \text{and} \quad g(I) = \left| \bigcap_{i \in I} S_i \setminus \bigcup_{i \notin I} S_i \right|$$

$f(I)$ counts the number of elements in S satisfying at least the properties in I . $g(I)$ counts the number of elements in S satisfying exactly the properties in I . Therefore, we have $f(I) = \sum_{J \supset I} g(J)$. We note that $\mathcal{P}(S)$ is isomorphic to the poset B_N , which is finite. Applying Möbius Inversion and Lemma 12.36 gives:

$$g(I) = \sum_{J \supset I} f(J) \mu(I, J) = \sum_{J \supset I} f(J) (-1)^{|J| - |I|}$$

In particular, substituting $I = \emptyset$ gives:

$$\left| S \setminus \bigcup_{i=1}^n S_i \right| = g(\emptyset) = \sum_{J \in B_n} (-1)^J \left| \bigcap_{j \in J} S_j \right| \quad \square$$

5. Partitions

Concept 5.1. Imagine placing n objects into k bins so that no bin is empty. We can count:

- Surjections/ordered set partitions, where the objects and bins are distinct
- Compositions, where the objects are identical but the bins are distinct
- Set partitions, where the objects are distinct but the bins are identical
- Integer partitions, where the objects and bins are identical

5.1 Compositions

Definition 5.2. A sequence of k nonnegative integers summing to n is called a weak composition of n into k parts. If all elements are nonzero, then it is called a composition.

Definition 5.3. Let $Q([n], k)$ denote the collection of compositions of n into k parts. Let $q(n, k) = |Q([n], k)|$ denote the number of such compositions. Let $Q([n])$ denote the set of compositions of n into any number of parts. Let $q(n) = |Q([n])|$ denote the number of such compositions.

Theorem 5.4. The number of weak compositions of n into k parts is $\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$.

Proof. Weak compositions represent ways of dividing the summands of $n = 1 + \dots + 1$ into k distinct bins, which can be done $\binom{n+k-1}{k-1}$ ways.

Alternatively, we can use combinatorial classes. Weak compositions into k parts are represented by $(\text{SEQ}(\mathcal{Z}))^k$, so that the ordinary generating function is $\frac{1}{(1-x)^k}$. We now use Lemma 7.19 to obtain:

$$\begin{aligned} f_n &= [x^n]F(x) = [x^n]\frac{1}{(1-x)^k} = [x^{n+k}]\frac{x^k}{(1-x)^{k+1}}(1-x) = [x^{n+k}]\frac{x^k}{(1-x)^{k+1}} - [x^{n+k-1}]\frac{x^k}{(1-x)^{k+1}} \\ &= \binom{n+k}{k} - \binom{n+k-1}{k} = \binom{n+k-1}{k-1} \quad \square \end{aligned}$$

Theorem 5.5. The number of compositions of n into k parts is $q(n, k) = \binom{n-1}{k-1}$.

Proof. Put one object in each bin, then distribute the remainder via a weak composition $\binom{n-1}{k-1} = \binom{n-1}{n-k}$ ways. Or, group the summands in $n = 1 + \dots + 1$ into k groups by choosing $k-1$ dividers from $n-1$ gaps.

Alternatively, we can use combinatorial classes. Compositions into k parts are represented by $(\text{SEQ}(\mathcal{Z}) \setminus \{\square\})^k$, where $\text{SEQ}(\mathcal{Z}) \setminus \{\square\}$ represents a natural number. Then the ordinary generating function is $F(x) = \left(\frac{1}{1-x} - 1\right)^k = \frac{x^k}{(1-x)^k}$. We now use Lemma 7.19 to obtain:

$$\begin{aligned} f_n &= [x^n]F(x) = [x^n]\frac{x^k}{(1-x)^k} = [x^n]\frac{x^k}{(1-x)^{k+1}}(1-x) = [x^n]\frac{x^k}{(1-x)^{k+1}} - [x^{n-1}]\frac{x^k}{(1-x)^{k+1}} \\ &= \binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1} \quad \square \end{aligned}$$

Theorem 5.6. The number of compositions of n is $q(n) = 2^{n-1}$.

Proof. Apply binomial theorem, or select any subset of the $n-1$ gaps to be dividers. □

Theorem 5.7. The number of compositions of $n \geq 2$ into an even (odd) number of parts is 2^{n-2} .

Proof. The number of compositions into k parts is $\binom{n-1}{k-1}$, so we can compute (noting that an alternating sum of binomial coefficients is 0):

$$\begin{aligned} \sum_{\substack{k=1 \\ k \text{ even}}}^n \binom{n-1}{k-1} &= \frac{1}{2} \left[\sum_{k=1}^n \binom{n-1}{k-1} (-1)^k + \sum_{k=1}^n \binom{n-1}{k-1} \right] \\ &= \frac{1}{2} \left[- \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k + \sum_{k=0}^{n-1} \binom{n-1}{k} \right] = \frac{1}{2} \sum_{k=0}^{n-1} \binom{n-1}{k} = 2^{n-2} \end{aligned}$$

Or, form an involution from compositions with an even number of parts to those with an odd number. If the composition begins with 1, combine this with the value to its right; otherwise, subtract one from the first value and place 1 as the first value. \square

5.2 Set Partitions

Definition 5.8. Let S be a set. A collection of non-empty disjoint subsets whose union is S is a set partition. Let $\left\{ \begin{smallmatrix} S \\ k \end{smallmatrix} \right\}$ denote set partitions of S into k blocks, and Π_n be the set partitions of $[n]$. Let $S(n, k) = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = \left| \left\{ \begin{smallmatrix} [n] \\ k \end{smallmatrix} \right\} \right|$ be the number of such set partitions (called Stirling set numbers of the second kind).

- ↳ The Bell numbers count all partitions of n : $b_n = \sum_{k=1}^n \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$
- ↳ The number of ordered set partitions is $k! \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$
- ↳ For $S = [n]$, we represent $\{1, 2, 4\}, \{6\}, \{5, 3\}$ by 124/35/6, so that each block is ascending and blocks are ordered by their first element
- ↳ One can also reverse the order of each block (e.g., 421/53/6) so that the dividers may be removed
- ↳ The canonical form of set partition $\sigma \in \Pi_n$ is $w(\sigma) = w_1 w_2 \cdots w_n \in [k]^n$ where $w_i = j$ if $i \in B_j$, for $\sigma = B_1 / \cdots / B_n$
- ↳ For example, $\sigma = 13/267/45$ implies $w(\sigma) = 1213322$

Theorem 5.9. $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = 0$ if $k > n$, $n < 0$, or $k < 0$. $\left\{ \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right\} = 1$ and $\left\{ \begin{smallmatrix} n \\ 0 \end{smallmatrix} \right\} = 0$ for $n > 0$, and otherwise $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\} + k \left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\}$.

Proof. Suppose $n, k \geq 1$. RHS conditions on whether $\{n\}$ is its own block. If so, then the remainder of the partition can be performed $\left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\}$ ways. Otherwise, partition all other elements $\left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\}$ ways then choose which of the k blocks contains n . \square

Theorem 5.10. $k! \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = \sum_{j=0}^k (-1)^j \binom{k}{j} (k-j)^n = \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} j^n$

- ↳ As written, this counts the number of ordered set partitions
- ↳ Dividing by $k!$ yields $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = \sum_{j=0}^k (-1)^j \frac{(k-j)^n}{j!(k-j)!} = \sum_{j=1}^k (-1)^{k-j} \frac{j^{n-1}}{(j-1)!(k-j)!}$

Proof. LHS counts surjections $[n] \rightarrow [k]$. RHS uses the Inclusion-Exclusion principle. There are k^n total functions, and let p_j be the property that j is not in the image of the function. Then we have:

$$\begin{aligned} k! \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} &= k^n - \sum_j N(p_j) + \sum_{j \leq k} N(p_j, p_k) - \cdots + (-1)^n N(p_1, \dots, p_k) \\ &= k^n - \sum_j (k-1)^n + \sum_{j \leq k} (k-2)^n - \cdots + (-1)^n (k-k)^k \\ &= k^n - \binom{k}{1} (k-1)^n + \binom{k}{2} (k-2)^n - \cdots + \binom{k}{k} (-1)^k (k-k)^n \end{aligned}$$

$$= \sum_{j=0}^k (-1)^j \binom{k}{j} (k-j)^n \quad \square$$

Alternatively, we can use generating functions. We could consider $A_n(y) = \sum_k \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} y^k$, $B_k(x) = \sum_n \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} x^n$, or $C(x, y) = \sum_{n,k} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} x^n y^k$, but it turns out the second is the most useful. Using Theorem 5.9, we have:

$$\begin{aligned} B_k(x) &= \sum_n \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} x^n = \sum_n \left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\} x^n + k \sum_n \left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\} x^n \\ &= x \sum_n \left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\} x^{n-1} + kx \sum_n \left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\} x^{n-1} = xB_{k-1}(x) + kxB_k(x) \end{aligned}$$

Solving gives $\frac{B_k(x)}{B_{k-1}(x)} = \frac{x}{1-kx}$. Applying k times gives $B_k(x) = \frac{B_0(x)}{B_0(x)} = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}$. We apply partial fraction decomposition to the denominator. Write $\frac{1}{(1-x)(1-2x)\cdots(1-kx)} = \sum_{j=1}^k \frac{\alpha_j}{1-jx}$. To solve this, first multiply each side by $1-rx$ for some fixed $1 \leq r \leq k$, then substitute $x = \frac{1}{r}$. All of the summands on the right vanish except for α_r , so that:

$$\begin{aligned} \alpha_r &= \frac{1}{(1-\frac{1}{r})(1-\frac{2}{r})\cdots(1-\frac{r-1}{r})(1-\frac{r+1}{r})\cdots(1-\frac{k}{r})} = \frac{r^{k-1}}{(r-1)(r-2)\cdots(1)(-1)\cdots(r-k)} \\ &= (-1)^{(k-r)} \frac{r^{k-1}}{(r-1)!(k-r)!} \end{aligned}$$

Thus, we have:

$$\begin{aligned} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} &= [x^n]B_k(x) = [x^n] \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)} = [x^{n-k}] \frac{1}{(1-x)(1-2x)\cdots(1-kx)} \\ &= \sum_{j=1}^k \alpha_j [x^{n-k}] \frac{1}{1-jx} = \sum_{j=1}^k \alpha_j j^{n-k} = \sum_{j=1}^k (-1)^{(k-j)} \frac{j^{k-1}}{(j-1)!(k-j)!} j^{n-k} \\ &= \sum_{j=1}^k (-1)^{(k-j)} \frac{j^{n-1}}{(j-1)!(k-j)!} \end{aligned}$$

Corollary 5.11. The bell numbers can be written as $b_n = \frac{1}{e} \sum_{j=0}^{\infty} \frac{j^n}{j!}$

↳ The presence of the factorial suggests that exponential generating functions may be useful for representing the Bell numbers

Proof. First, note that $b_n = \sum_{k=1}^M \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ for all $M \geq n$, since $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = 0$ when $k > n$. Fix one such M , then write the following using Theorem 5.10:

$$b_n = \sum_{k=1}^M \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = \sum_{k=1}^M \sum_{j=1}^k (-1)^{k-j} \frac{j^{n-1}}{(j-1)!(k-j)!} = \sum_{j=1}^M \frac{j^{n-1}}{(j-1)!} \sum_{k=j}^M \frac{(-1)^{k-j}}{(k-j)!} = \sum_{j=1}^M \frac{j^n}{j!} \sum_{k=0}^{M-j} \frac{(-1)^k}{k!}$$

It remains to take the limit as $M \rightarrow \infty$. To apply the product rule for limits, we must remove the dependence of j from the limits in the second summation, which is done by showing it is negligible. In particular, write:

$$b_n = \sum_{j=1}^M \frac{j^n}{j!} \sum_{k=0}^M \frac{(-1)^k}{k!} - \sum_{j=1}^M \frac{j^n}{j!} \sum_{k=M-j+1}^M \frac{(-1)^k}{k!}$$

$\sum_{k=M-j+1}^M \frac{(-1)^k}{k!}$ is an alternating series with terms decreasing in magnitude, and so is bounded above in absolute value by $\frac{1}{(M-j+1)!}$. Thus, the second term is bounded above in magnitude by

$\sum_{j=1}^M \frac{j^n}{j!} \cdot \frac{1}{(M-j+1)!}$. Taking the limit as $M \rightarrow \infty$, we have that an arbitrary number of the first terms approach 0 for sufficiently large M . Thus, it suffices to evaluate the limit $\lim_{M \rightarrow \infty} \sum_{j=N}^M \frac{j^n}{j!} \cdot \frac{1}{(M-j+1)!}$ for any N . This is less than $\lim_{M \rightarrow \infty} \sum_{j=N}^M \frac{j^n}{j!}$, but since $\sum_{j=1}^{\infty} \frac{j^n}{j!}$ converges by the ratio test, this means that $\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \sum_{j=N}^M \frac{j^n}{j!} = 0$. Hence, $\lim_{M \rightarrow \infty} \sum_{j=N}^M \frac{j^n}{j!} = 0$ for all N , and we obtain that the second term vanishes.

To simplify the first term, use the product rule for limits. We have $\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1}$, so that the claim holds. \square

Theorem 5.12. $x^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (x)_k$ for $n \geq 0$ and $x \in \mathbb{R}$.

Proof. Each side is a polynomial, so it suffices to prove it for all $x \in \mathbb{N}$. LHS is number of ways to distribute n identical balls into x distinct bins. RHS conditions on the number of nonempty bins k , grouping the balls $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ ways then picking the bins $(x)_k$ ways. \square

Corollary 5.13. $x^n = \sum_{k=0}^n (-1)^{n-k} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^{\bar{k}}$ for $n \geq 0$ and $x \in \mathbb{R}$.

\hookrightarrow Note that $(-1)^{n-k} \left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ are the signed Stirling numbers of the second kind

Proof. Replace $x \leftarrow -x$ in Theorem 5.12 to obtain $(-x)^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (-x)_k$. Now simplify to get $(-1)^n x^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^{\bar{k}} (-1)^k$. Rearrange to obtain the result. \square

Theorem 5.14. $\sum_{k=0}^n (-1)^k k! \left\{ \begin{matrix} n \\ k \end{matrix} \right\} = (-1)^n$

Proof. $(-1)^k k! = (-1)_k$, so apply Theorem 5.12. Or, form an involution between the ordered set partitions with an even number of blocks and those with an odd number of blocks, excluding the set partition $1/2/\dots/n$. If $\{1\}$ is not the first block, then either combine 1 with the block to its right (if $\{1\}$ is a block) or remove 1 from its block and place it in its own block to the left. If $\{1\}$ is the first block, apply the same process with 2, then 3, and so on.

This can also be proved using Theorem 5.12 and Lemma 10.1:

$$\begin{aligned} \sum_{k=0}^m \binom{m}{k} k^n (-1)^k &= \sum_{k=0}^m \binom{m}{k} (-1)^k \sum_{j=0}^n \left\{ \begin{matrix} n \\ j \end{matrix} \right\} (k)_j = \sum_{j=0}^n \left\{ \begin{matrix} n \\ j \end{matrix} \right\} \sum_{k=0}^m \binom{m}{k} (-1)^k (k)_j \\ &= \sum_{j=0}^n j! \left\{ \begin{matrix} n \\ j \end{matrix} \right\} \sum_{k=0}^m \binom{m}{k} \binom{k}{j} (-1)^k = \sum_{j=0}^n j! \left\{ \begin{matrix} n \\ j \end{matrix} \right\} (-1)^m \delta_{mj} = (-1)^m m! \left\{ \begin{matrix} n \\ m \end{matrix} \right\} \square \end{aligned}$$

Theorem 5.15. $(-x)^n = \sum_{k=0}^n (-1)^k k! \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \binom{x+k-1}{k}$

Proof. $(-x)_k = (x+k-1)_k = k! \binom{x+k-1}{k}$, now apply Theorem 5.12. \square

Theorem 5.16. $\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 2^{n-1} - 1$

Proof. $w : \left\{ \begin{matrix} [n] \\ 2 \end{matrix} \right\} \rightarrow [2]^n$ is injective, and the image is all words beginning with 1 and containing both 1 and 2. \square

Theorem 5.17. $\left\{ \begin{matrix} n+1 \\ m+1 \end{matrix} \right\} = \sum_{k=m}^n \binom{n}{k} \left\{ \begin{matrix} k \\ m \end{matrix} \right\}$

Proof. LHS counts set partitions of $[n+1]$ into $m+1$ blocks. RHS conditions on the number of elements k in a block different from $n+1$; the elements are chosen $\binom{n}{k}$ ways and distributed $\left\{ \begin{smallmatrix} k \\ m \end{smallmatrix} \right\}$ ways. \square

Theorem 5.18. $b_{n+1} = \sum_{k=0}^n \binom{n}{k} b_k$

Proof. Take sum of Theorem 5.17. Or, LHS counts ways to partition $[n+1]$, while RHS conditions on number of elements k in a block different from $n+1$. \square

Proposition 5.19. The exponential generating function of the Bell numbers b_n is $B(x) = e^{e^x-1}$.

Proof. $b_{n+1} = \sum_{k=0}^n \binom{n}{k} b_k$ by Theorem 5.18 with $b_0 = 1$. The generating function for b_{n+1} is $B'(x)$ by Wilf I, and the generating function for the right side is $B(x) \sum_{n=0}^{\infty} \frac{x^n}{n!} = B(x)e^x$ by Wilf III. Then $B'(x) = B(x)e^x$, or $\frac{B'(x)}{B(x)} = e^x$. The integral of LHS is $\log(B(x))$, and the integral of RHS is e^x , up to a constant. Thus, $\log(B(x)) = e^x + c$, or $B(x) = e^{e^x+c}$. $B(0) = b_0 = 1$ yields $c = -1$.

Alternatively, we can use Corollary 5.11 to directly calculate this generating function. We have:

$$\begin{aligned} B(x) &= \sum_{n=0}^{\infty} b_n \frac{x^n}{n!} = \frac{1}{e} \sum_{n=1}^{\infty} \frac{x^n}{n!} \sum_{j=1}^{\infty} \frac{j^n}{j!} + 1 = \frac{1}{e} \sum_{j=1}^{\infty} \frac{1}{j!} \sum_{n=1}^{\infty} \frac{(xj)^n}{n!} + 1 = \frac{1}{e} \sum_{j=1}^{\infty} \frac{e^{xj} - 1}{j!} + 1 \\ &= \frac{1}{e} \left[\sum_{j=1}^{\infty} \frac{e^{xj}}{j!} - \sum_{j=1}^{\infty} \frac{1}{j!} \right] + 1 = \frac{1}{e} \left[(e^{e^x} - 1) - (e - 1) \right] + 1 = e^{e^x-1} \end{aligned} \quad \square$$

As another method, we can use the permutation structure to derive this. Let $\mathcal{B}(L) = \{L_1/L_2/\cdots \vdash L\}$ be the labeled structure of set partitions. We can manipulate this:

$$\mathcal{B}(L) = \{\{L_1, L_2, \dots\} \mid L_1/L_2/\cdots \vdash L\} = \{\{S_1, S_2, \dots\} \mid L_1/L_2/\cdots \vdash L \text{ and } S_j \in \overline{E}(L_j)\} = \Pi(\overline{E})(L)$$

We have $F_{\overline{E}}(x) = e^x - 1$, so that the exponential generating function for the Bell numbers is e^{e^x-1} .

Proposition 5.20. The exponential generating function for the set partition numbers is given by

$$\sum_{n=0}^{\infty} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} \frac{x^n}{n!} = \frac{(e^x - 1)^k}{k!}.$$

Proof. When $k = 0$, the statement is obvious. Otherwise, we use induction; assume $k > 0$. Using Theorem 5.9:

$$B'_k(x) = \sum_{n=k}^{\infty} \left\{ \begin{smallmatrix} n+1 \\ k \end{smallmatrix} \right\} \frac{x^n}{n!} = \sum_{n=k}^{\infty} \left\{ \begin{smallmatrix} n \\ k-1 \end{smallmatrix} \right\} \frac{x^n}{n!} + k \sum_{n=k}^{\infty} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} \frac{x^n}{n!} = B_{k-1}(x) + kB_k(x)$$

Use induction to write $B_{k-1}(x) = \frac{(e^x-1)^{k-1}}{(k-1)!}$. Use the integrating factor e^{-kx} to obtain:

$$e^{-kx} B_k(x) = \int \frac{(e^x-1)^{k-1}}{(k-1)!} e^{-kx} dx = \int \frac{((e^x-1)/e^x)^{k-1}}{(k-1)!} e^{-x} dx$$

Let $u = \frac{e^x-1}{e^x} = 1 - e^{-x}$, so that $du = e^{-x} dx$. We obtain $e^{-kx} B_k(x) = u^k/k! + c$. Evaluating at $x = 0$, we should have $B_k(x) = 0$ since $k > 0$ and $u = 0$. Thus, $c = 0$. Solving:

$$B_k(x) = \frac{(1 - e^{-x})^k}{k!} \cdot e^{kx} = \frac{(e^x - 1)^k}{k!} \quad \square$$

Theorem 5.21. $\sum_{k=0}^m k \left\{ \begin{smallmatrix} n+k \\ k \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} n+m+1 \\ m \end{smallmatrix} \right\}$

Proof. RHS counts set partitions of $[n+m+1]$ into m blocks. LHS conditions on the largest k such that $n+k+1$ is not a singleton. This guarantees $(n+m+1) - (n+k+1) = m-k$ singletons, so the number of remaining blocks is k . The number of remaining elements is $n+k+1$. Distribute the first $n+k$, then pick which block contains $n+k+1$.

Or, use a telescoping sum. Note that by Theorem 5.9:

$$k \begin{Bmatrix} n+k \\ k \end{Bmatrix} = \left[\begin{Bmatrix} n+k \\ k-1 \end{Bmatrix} + k \begin{Bmatrix} n+k \\ k \end{Bmatrix} \right] - \begin{Bmatrix} n+k \\ k-1 \end{Bmatrix} = \begin{Bmatrix} n+k+1 \\ k \end{Bmatrix} - \begin{Bmatrix} n+k \\ k-1 \end{Bmatrix}$$

Using this, we get a telescoping sum. \square

Theorem 5.22. For $n \geq 3$, we have $b_n < n!$.

Use the representation where each block is listed in reverse order (so that dividers are not needed). Then every set partition of $[n]$ can be represented uniquely by a permutation in S_n . For $n \geq 3$, not every permutation is a valid set partition, such as any beginning with 231 (since 1 needs to appear in the first block).

Lemma 5.23. Let q_n be the number of all partitions of $[n]$ with no singleton blocks. Then $b_n = q_n + q_{n+1}$.

Proof. We prove $b_n - q_n = q_{n+1}$. LHS counts partitions with a singleton block, and RHS counts partitions with a new element $n+1$ with no singletons. We can form the bijection which combines all singleton elements with $n+1$ into a new block (and the reverse takes all elements in the same block as $n+1$ and places them in singleton blocks). \square

Lemma 5.24. $q_n = (-1)^n + \sum_{j=1}^n (-1)^{j-1} b_{n-j}$

Proof. $q_0 = 1$, so for induction, suppose the result holds for n . Then:

$$\begin{aligned} q_{n+1} &= b_n - q_n = b_n + (-1)^{n+1} + \sum_{j=1}^n (-1)^j b_{n-j} = (-1)^{n+1} + \sum_{j=0}^n (-1)^j b_{n-j} \\ &= (-1)^{n+1} + \sum_{j=1}^{n+1} (-1)^{j-1} b_{n+1-j} \end{aligned} \quad \square$$

Lemma 5.25. $q_n = \sum_{i=0}^n (-1)^i \binom{n}{i} b_{n-i}$

Proof. Let p_i be the property that i is a singleton. Then $N(p_i) = b_{n-1}$. Similarly, $N(p_{i_1}, \dots, p_{i_k}) = b_{n-k}$. Then $N_k = \binom{n}{k} b_{n-k}$, so that Inclusion/Exclusion yields $q_n = \sum_{i=0}^n (-1)^i N_i = \sum_{i=0}^n (-1)^i \binom{n}{i} b_{n-i}$.

Or, use binomial inversion. Let $q_{n,k} = \binom{n}{k} q_{n-k}$ be the number with exactly k singletons. By conditioning on the number of singletons, $b_n = \sum_{k=0}^n q_{n,k} = \sum_{k=0}^n \binom{n}{k} q_{n-k}$. By binomial inversion, $q_n = \sum_{k=0}^n \binom{n}{k} b_k (-1)^{n-k}$. Perform a change of variables to obtain the desired formula. \square

Lemma 5.26. The exponential generating function for q_n is $Q(x) = e^{e^x - x - 1}$.

Proof. Start with Lemma 5.25. By Wilf III, $Q(x) = B(x)e^{-x}$. By Proposition 5.19, we have $Q(x) = e^{e^x - x - 1}$. \square

Lemma 5.27. The exponential generating function for the number of ordered set partitions is $A(x) = \frac{1}{2 - e^x}$.

Proof. This uses compositions of exponential generating functions. An ordered set partition can be viewed as partitioning $[n]$ into k parts, building a structure on each partition $I[n > 0]$ ways, and then ordering the sets $k!$ ways. Building the structures within each partition corresponds to the EGF $e^x - 1$, while ordering the partitions corresponds to $\frac{1}{1-x}$. Composing these, we obtain $A(x) = \frac{1}{1 - (e^x - 1)} = \frac{1}{2 - e^x}$. \square

5.3 Integer Partitions

Definition 5.28. An integer partition of $n \in \mathbb{N}$ is a multiset λ whose elements in \mathbb{N} sum to n . This is notated $(\lambda_1, \dots, \lambda_k) \vdash n$, and λ is often written in descending order. $P([n])$ is the set of integer partitions of n , and $p(n) = |P([n])|$. $P_k([n])$ and $p_k(n)$ are defined similarly for integer partitions into k parts. $P_{\leq k}([n])$ and $p_{\leq k}(n)$ are used for integer partitions into at most k parts.

↳ Integer partitions are similar to compositions, but where order does not matter

Concept 5.29. A Young Diagram can be used to represent an integer partition, such as:

$$\lambda = (5, 3, 2, 2) \vdash 12 \quad \longleftrightarrow \quad \begin{array}{cccc} \square & \square & \square & \square \\ \square & \square & \square & \\ \square & \square & \square & \\ \square & \square & & \end{array}$$

Definition 5.30. The transpose λ^t of an integer partition λ is formed by taking the transpose of the Young Diagram, such as:

$$\lambda^t = (4, 4, 2, 1, 1) \vdash 12 \quad \longleftrightarrow \quad \begin{array}{ccccc} \square & \square & \square & \square & \\ \square & \square & \square & \square & \\ \square & \square & \square & & \\ \square & \square & & & \\ \square & & & & \end{array}$$

If $\lambda^t = (\lambda_1^t, \dots, \lambda_m^t)$, then λ_j^t counts the number of $\lambda_i \geq j$. This is a bijection.

Proposition 5.31. $p_k(n) = p_{\leq k}(n) - p_{\leq k-1}(n)$

Theorem 5.32. $p_k(n) = p_{\leq k}(n - k)$

Proof. For $\lambda \in P_k(n)$, taking the transpose yields a Young Diagram where the first row has precisely k blocks. Removing this first row is a bijection into $P_{\leq k}(n - k)$. Thus, composing these, we have the result. \square

Corollary 5.33. $p_{\leq k}(n) = p_{\leq k-1}(n) + p_{\leq k}(n - k)$

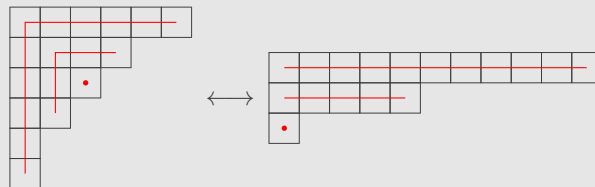
Proof. Consider $\lambda \in P_{\leq k}(n)$, and condition on whether λ is a composition of n into exactly k parts. $p_{\leq k-1}(n)$ counts the ways for this to be false, and $p_k(n) = p_{\leq k}(n - k)$ counts the ways for it to be true by Theorem 5.32. \square

Theorem 5.34. The number of partitions of n into at most k parts is equal to the number of partitions of n into parts not larger than k .

Proof. The transpose of the corresponding Young diagrams forms a bijection between these. \square

Proposition 5.35. The number of self-conjugate integer partitions (where $\lambda = \lambda^t$) is equal to the number of integer partitions with distinct parts all of which are odd.

Proof. Consider a bijection as illustrated below, where L-shaped blocks of cells in the self-conjugate Young diagram are flattened into rows in a new Young diagram. An example is illustrated below:



\square

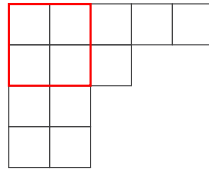
Proposition 5.36. For all $n \geq 2$, $p(n) - p(n - 1)$ counts the number of partitions of n in which the two largest parts are equal.

Proof. The number of partitions of n in which the two largest parts are equal, call this $\tilde{p}(n)$, can be put in bijective correspondence via the transpose with partitions in $P([n])$ all of whose parts are greater than 1. Through negative counting, this equals $p(n)$ minus the number of partitions containing 1. The partitions containing 1 can be bijected to partitions of $n - 1$ by removing a 1, so that $\tilde{p}(n) = p(n) - p(n - 1)$. \square

Proposition 5.37. For all positive integers n , $p(1) + p(2) + \cdots + p(n) < p(2n)$.

Proof. We claim that the LHS counts $\lambda \vdash 2n$ such that $n \leq \lambda_1 < 2n$, so that clearly the inequality holds. Specifically, LHS conditions on λ_1 , so that $p(1)$ counts the case when $\lambda_1 = 2n - 1$, $p(2)$ counts the case when $\lambda_1 = 2n - 2$, and more generally $p(k)$ counts the case when $\lambda_1 = 2n - k$. Indeed, any $\tilde{\lambda} \in P([k])$ will have $\tilde{\lambda}_1 \leq k \leq n$, so that inserting a row of $2n - k \geq n$ at the top of the Young diagram for $\tilde{\lambda}$ yields a valid Young diagram for some $\lambda \in P([2n])$. \square

Example 5.38. The Durfee square of a partition $\lambda \vdash n$ is the largest square fitting in the upper left corner of the Young diagram, such as in the following:



The width of this square can be computed without drawing the Young diagram. In particular, list $\lambda = (\lambda_1, \dots, \lambda_k) \vdash n$. Identify the largest j such that $\lambda_j \geq j$. The width of the Durfee square will then be j .

To prove this, note that the Durfee square has width j iff the Young diagram contains square (j, j) but not $(j + 1, j + 1)$, where we count from the top left. (j, j) is in the Young diagram iff $\lambda_j \geq j$, and $(j + 1, j + 1)$ is not in the Young diagram iff $\lambda_{j+1} < j + 1$. λ is monotone decreasing, so that $\lambda_{j+1} < j + 1$ implies $\lambda_i < i$ for all $i \geq j + 1$. Hence, the width of the Durfee square can be identified by finding the largest j such that $\lambda_j \geq j$.

Example 5.39. For all $k \leq n$, we have $p_k(n) \leq (n - k + 1)^{k-1}$.

↳ In particular, this shows $p_k(n)$ cannot be a polynomial in n . If it were, it would have to be of degree at most $k - 1$, but it would have k zeros for $0 \leq n \leq k - 1$

Proof. Consider the Young diagram for some $\lambda \in P_k([n])$. The maximal width of the Young diagram is $n - k + 1$, which occurs if all parts other than the first are 1. Further note that the last part is entirely determined by the previous $k - 1$ parts, since they must sum to n . Hence, the number of Young diagrams is upper-bounded by $(n - k + 1)^{k-1}$, where we pick each of the first $k - 1$ rows to be between 1 and $n - k + 1$ in length. \square

Theorem 5.40. The generating function for the number of integer partitions $p(n)$ is:

$$\mathcal{E}(x) = \sum_{n=0}^{\infty} p(n)x^n = \frac{1}{1-x} \frac{1}{1-x^2} \cdots = \prod_{j=1}^{\infty} \frac{1}{1-x^j}$$

↳ Although this appears to be an infinite product, we note that only finitely many factors contribute to the coefficient for any x^n . Indeed, $\frac{1}{1-x^j} = 1 + x^j + x^{2j} + \cdots$, so that $\frac{1}{1-x^j}$ has no contribution when $j > n$.

Proof. This is an immediate application of counting using the products of generating functions. Note that an integer partition of n can be represented as a tuple (a_1, a_2, \dots, a_n) . a_j is a multiple of j , and represents the contribution of all j terms in the integer partition. $\mathcal{E}(x)$ counts the number of ways to divide $n = 1 + \cdots + 1$ into the different components of the tuple under this restriction, since $\frac{1}{1-x^j} = 1 + x^j + x^{2j} + \cdots$.

More formally, we note that the combinatorial class of integer partitions is isomorphic to $\text{MSET}(\text{SEQ}(\mathcal{Z}) \setminus \{1\})$.

$\{\square\}$ where \mathcal{Z} is the atomic class. Indeed, $\text{SEQ}(\mathcal{Z}) \setminus \{\square\}$ is isomorphic to \mathbb{N} , so that taking the multiset yields (unordered) integer partitions. Now apply the rules for the generating functions. \square

Theorem 5.41 (Hardy/Ramanujan). $p(n) \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{2n/3}}$.

$\hookrightarrow a_n \sim b_n$ means $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$.

Example 5.42. Let $f(n)$ count the number of partitions of n such that no part is equal to 2. Let $F(x) = \sum_{n \geq 0} f(n)x^n = \frac{1}{1-x} \cdot 1 \cdot \frac{1}{1-x^3} \cdots = (1-x^2)\mathcal{E}(x)$. Then for $n \geq 2$:

$$f_n = [x^n](1-x^2)\mathcal{E}(x) = [x^n]\mathcal{E}(x) - [x^{n-2}]\mathcal{E}(x) = p(n) - p(n-2)$$

Proposition 5.43.

- (a) $d(x) = \prod_{j=1}^{\infty} (1+x^j)$ is the generating function counting the number of integer partitions with distinct parts.
- (b) $o(x) = \prod_{j=1}^{\infty} \frac{1}{1-x^{2j+1}}$ is the generating function counting the number of integer partitions using only odd parts.
- (c) Let $p_d(n)$ be the number of ways to partition n into distinct parts. Let $p_o(n)$ be the number of ways to partition n into parts all of which are odd. Then $p_d(n) = p_o(n)$.

Proof. The first two parts follow by counting with products of generating functions. Now perform the following manipulations, noting that even power of j cancel:

$$d(x) = \prod_{j=1}^{\infty} (1+x^j) = \prod_{j=1}^{\infty} \frac{1-x^{2j}}{1-x^j} = \prod_{\substack{j=1 \\ j \text{ odd}}}^{\infty} \frac{1}{1-x^j} = o(x)$$

We can also form a bijection. Beginning with a partition into distinct parts, iteratively replace each even number $2n$ with $n+n$ until all are odd. Beginning with a partition into odd parts, iteratively combine pairs of identical parts $n+n$ into $2n$. Since order is irrelevant in integer partitions, it is clear that both these procedures are order independent and terminate at a single possible partition. Furthermore, these procedures are inverses, making it a bijection. \square

Proposition 5.44. The number of partitions of n with no parts equal to 1 or 2 is $p(n) - p(n-1) - p(n-2) + p(n-3)$.

Proof. The number of partitions with at least one part equal to 1 is $p(n-1)$; a bijection is formed by adding/removing a 1. Similarly, the number of partitions with at least one 2 is $p(n-2)$, and the number of partitions with both is $p(n-3)$. Now apply inclusion/exclusion. \square

Theorem 5.45. $np(n) = \sum_{k=1}^n \sigma(k)p(n-k)$, where $\sigma(n)$ is the sum of all divisors of n . For example, $\sigma(6) = 1 + 2 + 3 + 6 = 12$.

Proof. We calculate $\mathcal{E}'(x)$ via logarithmic differentiation. In particular, $\log \mathcal{E}(x) = -\sum_{j=1}^{\infty} \log(1-x^j)$, so that taking the derivative yields:

$$\frac{\mathcal{E}'(x)}{\mathcal{E}(x)} = \sum_{j=1}^{\infty} \frac{jx^{j-1}}{1-x^j}$$

Rearrange to obtain $x\mathcal{E}(x) = \mathcal{E}(x) \sum_{j=1}^{\infty} \frac{jx^j}{1-x^j}$. The LHS is the generating function for $np(n)$ by

Wilf II, and by Wilf III, we have:

$$\begin{aligned}
 [x^n] \left(\mathcal{E}(x) \sum_{j=1}^{\infty} \frac{jx^j}{1-x^j} \right) &= \sum_{k=0}^n \left([x^k] \sum_{j=0}^{\infty} \frac{jx^j}{1-x^j} \right) p(n-k) = \sum_{k=0}^n \left([x^k] \sum_{j=1}^k \frac{jx^j}{1-x^j} \right) p(n-k) \\
 &= \sum_{k=0}^n \left(\sum_{j=1}^k [x^{k-j}] \frac{j}{1-x^j} \right) p(n-k) = \sum_{k=1}^n \left(\sum_{j=1}^k j I[j | k] \right) p(n-k) \\
 &= \sum_{k=1}^n \sigma(k) p(n-k) \quad \square
 \end{aligned}$$

5.4 Lah Numbers

Definition 5.46. The unsigned Lah numbers $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ count the number of ways to partition $[n]$ into k blocks, where a linear order is imposed within each block. The set of these structures is denoted $L([n], k)$. When considering all numbers of blocks, we denote the value by ℓ_n , and the associated set $L([n])$.

Lemma 5.47. The exponential generating function for $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ is $F_k(x) = \frac{(\frac{x}{1-x})^k}{k!}$, and the exponential generating function for ℓ_n is $F(x) = e^{\frac{x}{1-x}}$.

Proof. The exponential generating function for forming permutations is $\sum_{n=0}^{\infty} n! \frac{x^n}{n!} = \frac{1}{1-x}$. Thus, the exponential generating function for non-empty permutations is $\frac{1}{1-x} - 1 = \frac{x}{1-x}$. The second claim follows from forming the permutation structure on the corresponding labeled structure. For the first claim, use composition of exponential generating functions. We partition L into blocks, form a nonempty permutation on each with exponential generating function $\frac{x}{1-x}$, and form a k -set from all the blocks with exponential generating function $x^k/k!$. Hence, we obtain $F_k(x) = (\frac{x}{1-x})^k/k!$. \square

Proposition 5.48. $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] = \binom{n-1}{k-1} \frac{n!}{k!} = \binom{n}{k} \frac{(n-1)!}{(k-1)!}$ for $n \geq k \geq 1$.

Proof. Using Lemma 5.47 and Theorem 2.45:

$$\begin{aligned}
 \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] &= n! [x^n] \frac{(\frac{x}{1-x})^k}{k!} = \frac{n!}{k!} [x^n] \frac{x^k}{(1-x)^{k+1}} (1-x) = \frac{n!}{k!} \left([x^n] \frac{x^k}{(1-x)^{k+1}} - [x^{n-1}] \frac{x^k}{(1-x)^{k+1}} \right) \\
 &= \frac{n!}{k!} \left(\binom{n}{k} - \binom{n-1}{k} \right) = \frac{n!}{k!} \binom{n-1}{k-1}
 \end{aligned}$$

The second form then follows from Extraction/Absorption. We can also prove the second form combinatorially. Select the elements to start each block $\binom{n}{k}$ ways (this also ensures each block is nonempty). Sort the blocks so that the starting elements are in increasing order. We now must partition the remaining $n-k$ elements into k blocks so that there is order within each partition, the partitions themselves are ordered, and any partition may be empty. This can be done with stars and bars: we order a set of $n-k$ elements and $k-1$ distinct bars $\binom{n-1}{k-1}$ ways, then we remove the distinctness of the bars by dividing by $(k-1)!$. Hence, $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] = \binom{n}{k} \frac{(n-1)!}{(k-1)!}$. \square

Proposition 5.49. $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] = \sum_{j=k}^n \left[\begin{smallmatrix} n \\ j \end{smallmatrix} \right] \left\{ \begin{smallmatrix} j \\ k \end{smallmatrix} \right\}$

Proof. Consider the summation. It counts how to partition $[n]$ into any number of cycles, and then form a set partition of the set of cycles into k blocks. We can use labeled structures and the composition of exponential generating functions. The EGF for forming a permutation is $\log \frac{1}{1-x}$,

and the EGF for forming set partitions into k blocks is $\frac{(e^x - 1)^k}{k!}$. Composing these yields:

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} = \frac{(e^{\log(1/(1-x))} - 1)^k}{k!} = \frac{(\frac{1}{1-x} - 1)^k}{k!} = \frac{(\frac{x}{1-x})^k}{k!}$$

This is the EGF for $\begin{bmatrix} n \\ k \end{bmatrix}$ by Lemma 5.47.

Alternatively, we can use a combinatorial proof. LHS counts the number of ways to partition $[n]$ into k nonempty blocks with internal linear order, and the RHS counts the number of ways to permute $[n]$, then distribute its cycles into k nonempty blocks. The RHS can be equivalently carried out by partitioning $[n]$ into k blocks and then forming a permutation within each block. So, it remains to show that we can biject a linear ordering to a permutation. But this is already clear via the one-line representation of permutations. \square

Lemma 5.50. $x^{\bar{n}} = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} (x)_k$ and $(x)_n = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} x^{\bar{k}}$

Proof. For the first claim, we use Proposition 3.24, Theorem 5.12, and Proposition 5.49 to write:

$$x^{\bar{n}} = \sum_{j=0}^n \begin{bmatrix} n \\ j \end{bmatrix} x^j = \sum_{j=0}^n \begin{bmatrix} n \\ j \end{bmatrix} \sum_{k=0}^j \left\{ \begin{matrix} j \\ k \end{matrix} \right\} (x)_k = \sum_{k=0}^n (x)_k \sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} (x)_k$$

For the second claim, we use Corollary 3.25, Corollary 5.13, and Proposition 5.49 to write:

$$\begin{aligned} (x)_n &= \sum_{j=0}^n (-1)^{n-j} \begin{bmatrix} n \\ j \end{bmatrix} x^j = \sum_{j=0}^n (-1)^{n-j} \begin{bmatrix} n \\ j \end{bmatrix} \sum_{k=0}^j (-1)^{j-k} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} x^{\bar{k}} = \sum_{k=0}^n (-1)^{n-k} x^{\bar{k}} \sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} \\ &= \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} x^{\bar{k}} \end{aligned} \quad \square$$

6. Graphs

6.1 Chromatic Polynomials

Definition 6.1. Let G be a simple graph. An n -coloring of G is a map $c : V \rightarrow [n]$. It is a proper n -coloring if adjacent vertices in G has different colors, i.e., $c(u) \neq c(v)$ for all edges uv . The number of proper n -colorings is denoted $p_G(n)$, which is called the chromatic polynomial of G .

Lemma 6.2. Consider a graph G with d vertices. The chromatic polynomials for some common graphs include:

- If $|E| = 0$, then $p(x) = x^d$
- If G is a tree, then $p(x) = x(x-1)^{d-1}$
- If $G = K_d$ is a complete graph, then $p(x) = (x)_d$

Theorem 6.3 (4-Color Theorem). Every planar graph has a proper 4-coloring. Or, $p_G(4) > 0$ for every planar graph.

↳ This was first introduced in 1852, but it was not proved until 1976. The proof relied on extensive computer computation and casework, and a simplified proof has been elusive.

↳ The concept of a chromatic polynomial was introduced to help prove this. Despite it not yielding a proof, chromatic polynomials have received extensive research in combinatorics and graph theory.

Definition 6.4. Let $G = (V, E)$ be a graph with an edge $e = uv \in E$. Then the deletion operator gives $G \setminus e = (V, E \setminus \{e\})$. The contraction operator G/e gives the graph obtained by identifying u with v (i.e., merging the two vertices via a quotient).

Lemma 6.5. For a simple graph $G = (V, E)$ and an edge $e \in E$, we have $p_G(n) = p_{G \setminus e}(n) - p_{G/e}(n)$.

↳ This provides a basis for proof by strong induction on $|E|$, since $G \setminus e$ has one fewer edge and G/e has at least one fewer edge

↳ It also provides a way to develop computational algorithms

Proof. Consider the proper n -colorings of $G \setminus e$. Each coloring either has $c(u) = c(v)$, in which case it is a coloring of $p_{G/e}$, or it has $c(u) \neq c(v)$, in which case it is a coloring of G . Each of these relationships are also bijective, so rearrange to obtain the result. \square

Proposition 6.6. If $G = (V, E)$ is a simple graph, then $p_G(n)$ is a polynomial of degree $|V|$ with integer coefficients.

Proof. This is proved by induction on $|E|$. The base case of $|E| = 0$ gives $p_G(n) = n^{|V|}$, which satisfies the claim. By induction for $|E| > 0$, using Lemma 6.5, we can write $p_G(n)$ as a difference of a $|V|$ -degree and a $(|V| - 1)$ -degree polynomial each with integer coefficients. The result is that $p_G(n)$ has integer coefficients and degree $|V|$. \square

Corollary 6.7. Write the chromatic polynomial of G as $p_G(n) = c_d n^d + c_{d-1} n^{d-1} + \dots + c_0$. Then:

- $c_d = 1$
- $c_0 = 0$
- $c_{d-1} = -|E|$
- $(-1)^d p_G(-n) > 0$ for all $n \geq 1$

Proof. This follows via induction on $|E|$ using Lemma 6.5. The base case holds for all the claims when $|E| = 0$, since the chromatic polynomial is n^d which has $c_d = 1$, $c_0 = 0$, $c_{d-1} = 0$, and $(-1)^d (-n)^d = n^d > 0$. For $|E| > 0$, by Lemma 6.5, $p_G(n)$ is the difference of a $|V|$ -degree and a $(|V| - 1)$ -degree polynomial each with fewer edges. By induction, we thus have:

- The leading coefficient of $p_G(n)$ is the same as the leading coefficient of $p_{G \setminus e}(n)$, which is 1
- The constant term of $p_G(n)$ is the difference of the constant terms of the two polynomials,

each of which have constant term 0. Hence, $p_G(n)$ also has constant term zero

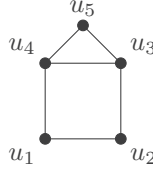
- The term c_{d-1} in $p_G(n)$ is the difference of the term c_{d-1} in $p_{G \setminus e}(n)$ and the leading coefficient of $p_{G/e}(n)$. By induction, this is $-(|E| - 1) - 1 = -|E|$
- We have $(-1)^d p_G(n) = (-1)^d (p_{G \setminus e}(n) - p_{G/e}(n)) = (-1)^d p_{G \setminus e}(n) + (-1)^{d-1} p_{G/e}(n)$, where each summand is positive by induction

□

Theorem 6.8. Let H and L be graphs such that $H \cap L = K_d$. Let $G = H \cup L$. Then $p_G(x) = \frac{p_H(x) \cdot p_L(x)}{(x)_d}$.

Proof. The intersection K_d contains d distinct colors. Consider equivalence classes of colorings on H , L , and G where two colorings are equivalent if the d colors in K_d are the same. By symmetry (namely, picking a permutation of the colors that switches the colors in K_d to any other set), we have that all equivalence classes for a given graph are the same size. There are $(x)_d$ such equivalence classes, so that the size of each equivalence class is $p(x)/(x)_d$. Pick the equivalence class $(x)_d$ ways, pick the remaining coloring on H $p_H(x)/(x)_d$ ways, and pick the remaining coloring on L $p_L(x)/(x)_d$ ways. Multiply these together to obtain the form for $p_G(x)$. □

Example 6.9. Consider the graph given below:



Let $H = \{u_1, u_2, u_3, u_4\}$ and $L = \{u_3, u_4, u_5\}$. Then they intersect on $\{u_3, u_4\}$, which is K_2 . Using Lemma 6.5, we find that $p_H(x) = x(x-1)(x^2 - 3x + 3)$, and we have that $p_L(x) = x(x-1)(x-2)$. Hence, Theorem 6.8, we have $p_G(x) = x(x-1)(x-2)(x^2 - 3x + 3)$.

Definition 6.10. Let G be a graph with vertices $V(G) = \{u_1, \dots, u_d\}$. We define an orientation on G , denoted θ , which specifies an order for each element in $E(G)$. The unordered pairs in $E(G)$ become ordered pairs in θ , so that if $uv \in \theta$, then $u \rightarrow v$ in G .

Definition 6.11. A directed path in G with orientation θ is a sequence of distinct vertices $x_{j-1} \rightarrow x_j$ for $j = 1, \dots, s$ directed according to the given θ . In addition, if $x_s \rightarrow x_0$, then $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_s \rightarrow x_0$ is called a directed cycle.

Definition 6.12. An orientation θ is called acyclic if there are no directed cycles.

Definition 6.13. For a proper n -coloring $c : V \rightarrow [n]$, the orientation induced by c is $\theta = \{u_i u_j \in E \mid c(u_i) > c(u_j)\}$. In other words, all edges point from lower colors to higher colors.

↳ The induced orientation will clearly be acyclic.

Definition 6.14. Consider (θ, c) where θ is an orientation and c is an n -coloring, not necessarily proper. (θ, c) is called compatible if for every oriented edge $u \rightarrow v$ we have $c(u) \leq c(v)$. It's called strictly compatible if $c(u) < c(v)$.

Proposition 6.15. If (θ, c) is strictly compatible, then c is a proper n -coloring and θ is acyclic.

Proof. If (θ, c) is strictly compatible, then since each edge uv is directed, we have either $c(u) > c(v)$ or $c(u) < c(v)$. In particular, $c(u) \neq c(v)$, so that the coloring is proper. To prove that θ is acyclic, suppose for contradiction there existed a directed cycle $u_{i_1} \rightarrow u_{i_2} \rightarrow \dots \rightarrow u_{i_k} \rightarrow u_{i_1}$. Then $c(u_{i_1}) < c(u_{i_2}) < \dots < c(u_{i_k}) < c(u_{i_1})$, which yields a contradiction. □

Theorem 6.16. Let G be a finite simple graph, and consider its chromatic polynomial $p_G(x)$. Then $(-1)^{|V|} p_G(-n)$ equals the number of compatible pairs (θ, c) , where θ is acyclic and c is an n -coloring.

Proof. Let $\lambda_G(n)$ count the number of compatible pairs. We claim $\lambda_G(n) = \lambda_{G \setminus e}(n) + \lambda_{G/e}(n)$ for any fixed $e = uv$. Fix any compatible orientation θ and coloring c on G , and consider the following cases regarding e :

- $c(u) \neq c(v)$. These pairs are in bijective correspondence with the pairs on $G \setminus e$ for which $c(u) \neq c(v)$, with the bijection formed by adding/removing $u \rightarrow v$ in the case $c(u) < c(v)$, or adding/removing $v \rightarrow u$ in the case $c(v) < c(u)$.
- $c(u) = c(v)$, and switching the orientation of uv yields a cyclic orientation. These pairs are in bijective correspondence with the pairs on $G \setminus e$ for which $c(u) \neq c(v)$ and there exists a directed path connecting u and v in some order. The inverse is uniquely defined since only one of $u \rightarrow v$ or $v \rightarrow u$ yields an acyclic orientation.
- $c(u) = c(v)$, $u \rightarrow v$, and switching the orientation of uv yields an acyclic orientation. These pairs are in bijective correspondence with the pairs on $G \setminus e$ for which $c(u) \neq c(v)$ and there do not exist any paths connecting u and v .
- $c(u) = c(v)$, $v \rightarrow u$, and switching the orientation of uv yields an acyclic orientation. These pairs are in bijective correspondence with the pairs on G/e via the natural bijection. The orientation on G/e will be acyclic since $u \rightarrow v$ and $v \rightarrow u$ are acyclic.

Altogether, we have $\lambda_G(n) = \lambda_{G \setminus e}(n) + \lambda_{G/e}(n)$. Now we use this to prove our claim by induction on $|E|$. When $|E| = 0$, there is one orientation with $n^{|V|}$ compatible colorings, so $\lambda_G(n) = n$. We have $p_G(x) = x^{|V|}$, hence $(-1)^{|V|}p_G(-n) = (-1)^{|V|}(-n)^{|V|} = n^{|V|} = \lambda_G(n)$. Now consider $|E| > 0$:

$$\begin{aligned} \lambda_G(n) &= \lambda_{G \setminus e}(n) + \lambda_{G/e}(n) = (-1)^{|V|}p_{G \setminus e}(n) + (-1)^{|V|-1}p_{G/e}(n) \\ &= (-1)^{|V|}[p_{G \setminus e}(n) - p_{G/e}(n)] = (-1)^{|V|}p_G(n) \end{aligned} \quad \square$$

Example 6.17. Consider a triangular graph. It is a complete graph, so that $p_G(x) = x(x-1)(x-2)$. Using $n = 1$ gives the number of acyclic orientations, which is $(-1)^3p_G(-1) = 6$. Substituting $n = 2$ gives 24, which is the number of ways to color the graph with 2 colors and then create an acyclic compatible orientation.

6.2 Trees

Definition 6.18. A tree is a simple, undirected, connected, acyclic graph.

Proposition 6.19. Every tree contains at least two leaf nodes.

Proof. Suppose for contradiction a tree exists that violates this. Start at a leaf node (if exactly one exists) or an arbitrary node (if no leaf nodes exist). Move along an arbitrary edge, and remove that edge. Note that this prevents us from visiting the starting node if it was a leaf node. Since any nodes we visit in this manner cannot be leaf nodes, we will either always find another edge to traverse, or we will visit the same node multiple times. The former contradicts the fact that the tree has finitely many edges, while the latter contradicts the fact that the tree is acyclic. \square

Lemma 6.20. Every tree with n vertices has $n - 1$ edges.

Proof. When $n = 1$, this is trivial. For higher n , we can find a leaf node by Proposition 6.19. Both this leaf node and its outgoing edge can be removed to yield a valid tree with one fewer edge and one fewer node. Continue until the trivial case is reached. \square

Lemma 6.21. Any simple, undirected, acyclic graph with n vertices and $n - 1$ edges must be connected, and is thus a tree.

Proof. If it is not connected, then it can be represented as a union of several disconnected trees. Let the size of these trees be n_i , so that $\sum_{i=1}^k n_i = n$. The number of edges is thus $\sum_{i=1}^k (n_i - 1) = n - k < n - 1$, a contradiction. \square

Lemma 6.22. The number of labeled trees (those with nodes labeled in $[n]$) is n^{n-2} , or 0 when $n = 0$.

Proof. When $n = 0$, we define the number of labeled trees to be 0. $n = 1$ is trivial.

Otherwise, we form a bijection from labeled trees to $[n]^{n-2}$. To convert a tree into a sequence, iteratively remove the leaf with the smallest label, and append to the sequence its unique neighbor. Repeat until two nodes remain, at which point the process is terminated.

To invert this, first note that all labels not appearing in the sequence must be leaf nodes. Indeed, if they were internal nodes, one of their two neighbors would have to be removed in order to reach a tree with two nodes. Conversely, these are the only leaf nodes, since a label occurring in the sequence must have at least two neighbors in order for one to be removed.

Consider the label set, which contains all eligible labels. Identify all eligible labels not present in the code. There will be at least two, since the label set is two larger in size than the sequence. Identify the smallest, and connect it to the first element in the code. Now remove the first element in the code and remove the smallest element from the label set. Now recursively apply this procedure until the code is empty, at which point the remaining two nodes are connected.

In order to show that the inverse is well-defined on all of $[n]^{n-2}$, we must prove that the result is simple, connected, and acyclic. It cannot add the same edge twice, since each recursive step removes a node from consideration. It cannot add a loop, since a node cannot both be in the code and not in the code. For each edge added, one of its endpoints is permanently removed from the label set and thus cannot have additional edges added to it, so that the resulting graph is acyclic. The resulting acyclic graph has $n - 1$ edges, so that it must be a tree by Lemma 6.21.

We have an inverse for our forward map on the image of that map, but it remains to show the forward map is surjective or that the inverse is injective on its entire domain. We prove the latter. Suppose two sequences produces the same tree. The leaf nodes of that tree are determined by the elements not in the sequence, so that both sequences must use the same subset of $[n]$. The lowest leaf node is always connected solely to the first element in the sequence, so that both codes must have the same first element. Now, recursively, continue this argument by looking at the subtree without this node. \square

Definition 6.23. A plane tree is a rooted tree in which an ordering is specified for the children of each vertex.

Definition 6.24. A k -ary tree is a plane tree such that the number of children of any node is either 0 and k . Let $\mathcal{T}^{\{0,k\}}$ denote the class of k -ary trees. The size of a k -ary tree is defined as its number of nodes.

Definition 6.25. For $\Omega \subset \mathbb{W}$ and $0 \in \Omega$, let \mathcal{T}^Ω be the class of plane trees such that the number of children at each node lies in Ω . These trees are called Ω -restricted trees. The characteristic function encapsulating Ω is $\phi(z) = \sum_{k \in \Omega} z^k$.

↳ As an example, the class of k -ary trees has characteristic $\phi(z) = 1 + z^k$ and the class of unrestricted plane trees has characteristic $\phi(z) = (1 - z)^{-1}$.

Proposition 6.26. The ordinary generating function $T^\Omega(x)$ for Ω -restricted trees satisfies the functional equation $T^\Omega(x) = x\phi(T^\Omega(x))$. Therefore, by Lagrange Inversion:

$$[x^n]T^\Omega(x) = \frac{1}{n}[z^{n-1}]\phi(z)^n$$

Proof. This follows from using combinatorial classes. $\mathcal{T}^\Omega = \mathcal{Z} \times \text{SEQ}_\Omega(\mathcal{T}^\Omega)$, so that $T^\Omega(x) = x \sum_{k \in \Omega} T^\Omega(x)^k = x\phi(T^\Omega(x))$. \square

Example 6.27. The number of size- n trees in $\mathcal{T}^{\{0,1,3\}}$ is $\frac{1}{n} \sum_{\substack{k=1 \\ k \text{ odd}}}^n \binom{n}{k} \binom{n-k}{(k-1)/2}$.

Proof.

$$[x^n]T(x) = \frac{1}{n}[z^{n-1}](1 + z + z^3)^n = \frac{1}{n}[z^{n-1}] \sum_{k=0}^n \binom{n}{k} (z + z^3)^{n-k}$$

$$\begin{aligned}
&= \frac{1}{n} \sum_{k=0}^n \binom{n}{k} [z^{n-1}] z^{n-k} (1+z^2)^{n-k} = \frac{1}{n} \sum_{k=1}^n \binom{n}{k} [z^{k-1}] (1+z^2)^{n-k} \\
&= \frac{1}{n} \sum_{\substack{k=1 \\ k \text{ odd}}}^n \binom{n}{k} [z^{(k-1)/2}] (1+z)^{n-k} = \frac{1}{n} \sum_{\substack{k=1 \\ k \text{ odd}}}^n \binom{n}{k} \binom{n-k}{(k-1)/2}
\end{aligned}$$

In the first binomial expansion, we used the power $n-k$ instead of k to cancel out the n in $[z^{n-1}]$. If we used k instead, we would have $[z^{n-k-1}](1+z^2)^k$, which is harder to deal with since now the parity of k for which this is nonzero depends also on the parity of n . \square

Example 6.28. The number of size- n trees in $\mathcal{T}^{\{0,2,3\}}$ is $\frac{1}{n} \sum_{k=0}^n \binom{n}{k} \binom{k}{n-2k-1}$.

Proof.

$$\begin{aligned}
[x^n]T(x) &= \frac{1}{n} [z^{n-1}] (1+z^2+z^3)^n = \frac{1}{n} \sum_{k=0}^n \binom{n}{k} [z^{n-1}] (z^2+z^3)^k \\
&= \frac{1}{n} \sum_{k=0}^n \binom{n}{k} [z^{n-2k-1}] (1+z)^k = \frac{1}{n} \sum_{k=0}^n \binom{n}{k} \binom{k}{n-2k-1}
\end{aligned}$$

\square

Example 6.29. Suppose we change the size of a tree in $\mathcal{T}^{\{0,2,3\}}$ to be the number of non-leaf nodes. Denote this class by $\overline{\mathcal{T}}$. The number of trees of size $n = 1$ is 1, and when $n > 1$, the number of trees of size n is $\frac{1}{n} \sum_{k=0}^{n-1} \binom{2n}{k} \binom{n}{n-k-1} 2^{k+1}$.

Proof. Any tree in $\overline{\mathcal{T}}$ is either a lone root or a root with 2 or 3 children. Then $\overline{\mathcal{T}} = \{\square\} + \mathcal{Z} \times \text{SEQ}_{\{2,3\}}(\overline{\mathcal{T}})$. This yields $\overline{T}(x) = 1 + x(\overline{T}(x)^2 + \overline{T}(x)^3)$. Let $z(x) = \overline{T}(x) - 1$ and $\phi(z) = (z+1)^2(z+2)$. Then $z(x) = x(\overline{T}(x)^2 + \overline{T}(x)^3) = x((z+1)^2 + (z+1)^3) = x(z+1)^2(z+2) = x\phi(z)$. Apply Lagrange Inversion for $n \geq 1$ to obtain:

$$\begin{aligned}
[x^n]z(x) &= \frac{1}{n} [z^{n-1}] \phi(z)^n = \frac{1}{n} [z^{n-1}] (z+1)^{2n} (z+2)^n = \frac{1}{n} \sum_{k=0}^{n-1} [z^k] (z+1)^{2n} \cdot [z^{n-k-1}] (z+2)^n \\
&= \frac{1}{n} \sum_{k=0}^{n-1} \binom{2n}{k} \binom{n}{n-k-1} 2^{k+1}
\end{aligned}$$

\square

7. Generating Functions

7.1 Definitions

Definition 7.1. A formal power series is a summation $\sum_{n=0}^{\infty} a_n x^n$ for a sequence of coefficients $\{a_n\}$

↳ Two series are equal if the coefficients are equal

↳ The formal power series corresponding to $\{a_n\}$ is called its ordinary generating function, and the relationship is denoted $\sum_{n=0}^{\infty} a_n x^n \xleftrightarrow{\text{ogf}} \{a_n\}$

Proposition 7.2. Let $\mathbb{F}[[x]]$ denote the space of formal power series over a field \mathbb{F} , such as $\mathbb{C}[[x]]$. This forms an algebra with the operations:

(i) Addition: $\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$

(ii) Scalar Multiplication: $c \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} c a_n x^n$

(iii) Multiplication: $\left(\sum_{n=0}^{\infty} a_n x^n \right) \left(\sum_{n=0}^{\infty} b_n x^n \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) x^n$

Definition 7.3. $[x^k] : \mathbb{F}[[x]] \rightarrow \mathbb{F}$ defined by $[x^k] \sum_{n=0}^{\infty} a_n x^n = a_k$ is a linear functional. In $\mathbb{F}[[x]]$, we have $[x^k] \sum_{n=0}^{\infty} a_n x^n = 0$ for all $k < 0$.

Proposition 7.4. $[x^n]$ satisfies the following properties:

- Linearity: $[x^n](\alpha A(x) + \beta B(x)) = \alpha [x^n]A(x) + \beta [x^n]B(x)$
- Shifting: $[x^n]x^k A(x) = [x^{n-k}]A(x)$ for $n \geq k$
- Differentiation: $[x^n]A'(x) = (n+1)[x^{n+1}]A(x)$ (using formal differentiation in Proposition 7.13)
- Convolution: $[x^n]A(x)B(x) = \sum_{k=0}^n ([x^k]A(x))([x^{n-k}]B(x))$
- Composition: $[x^n]A(B(x)) = \sum_{k=0}^{\infty} [x^k]A(x) \cdot [x^n](B(x))^k$ (assuming this is well-defined in Definition 7.8)

Proof. Linearity follows trivially, since $[x^n]$ is a linear functional. Shifting follows by definition of multiplication, and differentiation by definition in Proposition 7.13. Convolution follows by definition of multiplication. Composition is proved as follows (where the infinite sum is finite for any given n , so that linearity holds):

$$[x^n]A(B(x)) = [x^n] \sum_{k=0}^{\infty} a_k (B(x))^k = \sum_{k=0}^{\infty} a_k [x^n](B(x))^k = \sum_{k=0}^{\infty} [x^k]A(x) \cdot [x^n](B(x))^k \quad \square$$

Definition 7.5. G is a map from sequences in \mathbb{F} to formal power series, so that $G(\{f_n\}) = \sum_{n=0}^{\infty} f_n x^n$. For simplicity, we often denote $G(f_n) = G(\{f_n\})$.

Proposition 7.6. G satisfies the following properties:

- Linearity: $G(\alpha a_n + \beta b_n) = \alpha G(a_n) + \beta G(b_n)$
- Shifting: $G(a_{n+1}) = \frac{G(a_n) - a_0}{x}$
- Differentiation: $G(na_n) = x D_x G(a_n)$

- Convolution: $G\left(\sum_{k=0}^n a_k b_{n-k}\right) = G(a_n)G(b_n)$
- Composition: $\sum_{k=0}^{\infty} a_k G(b_n)^k = G(a_n) \circ G(b_n)$ (assuming this is well-defined in Definition 7.8)

Proof. Linearity holds by the definition of addition for formal power series. Shifting follows trivially. Differentiation holds as follows:

$$G(na_n) = \sum_{n=0}^{\infty} na_n x^n = x \sum_{n=0}^{\infty} na_n x^{n-1} = x D_x G(a_n)$$

Convolution holds by the definition of multiplication, and composition trivially holds. \square

Theorem 7.7. A formal power series has a reciprocal iff $a_0 \neq 0$, and it is unique if it exists.

Proof. Uniqueness follows from algebra (uniqueness of multiplicative inverses). For $\sum_{n=0}^{\infty} a_n x^n$, suppose $[\sum_{n=0}^{\infty} a_n x^n][\sum_{n=0}^{\infty} b_n x^n] = 1$. For $n = 0$, this implies $a_0 b_0 = 1$, so $a_0 \neq 0$ and $b_0 = -a_0^{-1}$. For $n = 1$, we have $a_0 b_1 + a_1 b_0 = 0$, so $b_1 = -a_0^{-1} a_1 b_0$. This can be continued recursively to obtain a reciprocal. \square

Definition 7.8. The composition of two formal power series $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $g(x)$ is $f(g(x)) = \sum_{n=0}^{\infty} a_n g(x)^n$. However, this may not be well-defined. If $[x^0]g(x) \neq 0$, then infinitely many terms in the above definition are needed to calculate each coefficient of the resulting power series, so since infinite series are not defined, neither is the power series. However, if $[x^0]g(x) = 0$, then $a_n(g(x))^n = a_n(g_1 x + g_2 x^2 + \dots)^n = a_n x^n (g_1 + g_2 x + \dots)^n$, so that finitely many terms are needed to compute each coefficient. So, the composition is well-defined if $[x^0]g(x) = 0$ or if f is a polynomial.

Proposition 7.9. Suppose $[x^0]f(x) = 0$. Then f has an inverse g , so that $f(g(x)) = x = g(f(x))$, iff $[x^1]f(x) \neq 0$.

Definition 7.10. Let $\{f_k(x)\}_{k \geq 0}$ be a sequence of formal power series. $f_k(x)$ converges to $f(x)$ if $[x^n]f_k(x)$ is eventually constant and equal to $[x^n]f(x)$ as $k \rightarrow \infty$ for all n .

↳ We require each $[x^n]f_k(x)$ to become constant since there is not necessarily a notion of convergence in the field \mathbb{F}

Example 7.11. For a power series $A(x) = \sum_{n=0}^{\infty} a_n x^n$, the sequence of partial sums $s_k(x) = \sum_{n=0}^k a_n x^n$ converges to $A(x)$.

Proposition 7.12. Define the degree of a formal power series $f(x)$ to be the smallest n such that $a_n \neq 0$, denoted $\deg(f(x))$. If $f(x) = 0$, then let $\deg(f(x)) = \infty$. $\sum_{k=0}^{\infty} f_k(x)$ exists iff $\lim_{k \rightarrow \infty} \deg(f_k(x)) \rightarrow \infty$.

Proposition 7.13. Define differentiation as $f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}$ for $f(x) = \sum_{n=0}^{\infty} a_n x^n$. All normal derivative rules such as the sum, product, and quotient rule hold for formal power series.

Definition 7.14. Other types of generating functions are also possible, such as:

Type	Abbreviation	Form	Notation
Ordinary Generating Function	ogf	$f(x) = \sum_{n=0}^{\infty} a_n x^n$	$f(x) \xleftrightarrow{\text{ogf}} \{a_n\}$
Exponential Generating Function	egf	$g(x) = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!}$	$g(x) \xleftrightarrow{\text{egf}} \{b_n\}$
Dirichlet Generating Function	dgf	$h(s) = \sum_{n=1}^{\infty} \frac{c_n}{n^s}$	$h(s) \xleftrightarrow{\text{dgf}} \{c_n\}$

7.2 Common Generating Functions

Example 7.15. $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$, so that $[x^k] \frac{1}{1-x} = 1$

Proof. This can be shown with the Generalized Binomial Theorem:

$$(1-x)^{-1} = \sum_{n=0}^{\infty} \binom{-1}{n} (-x)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} (-x)^n = \sum_{n=0}^{\infty} x^n$$

Or, carry out the multiplication:

$$(1-x) \sum_{n=0}^{\infty} x^n = \sum_{n=0}^{\infty} x^n - \sum_{n=0}^{\infty} x^{n+1} = x^0 = 1 \quad \square$$

Lemma 7.16. $[x^n] \frac{1}{1-ax} = a^n$

Proof. Substitute ax for x in Example 7.15. \square

Lemma 7.17. $[x^n] \frac{1}{x-a} = -\frac{1}{a^{n+1}}$

Proof. Using Lemma 7.16, $[x^n] \frac{1}{x-a} = -\frac{1}{a} [x^n] \frac{1}{1-\frac{1}{a}x} = -\frac{1}{a^{n+1}}$. \square

Lemma 7.18. $[x^n] \frac{A(x)}{1-x} = \sum_{k=0}^n a_k$ from Example 7.15 and definition of multiplication

↳ In other words, dividing by $1-x$ gives the partial sums

Lemma 7.19. $[x^n] \frac{x^k}{(1-x)^{k+1}} = \binom{n}{k}$ by Theorem 2.45.

Lemma 7.20. $[x^n] \frac{1}{(x-a)^k} = (-1)^k \left(\frac{1}{a}\right)^{n+k} \binom{n+k-1}{k-1}$

Proof.

$$\begin{aligned} [x^n] \frac{1}{(x-a)^k} &= [x^{n+k-1}] \frac{x^{k-1}}{(x-a)^k} = \frac{1}{a} [x^{n+k-1}] \frac{(x/a)^{k-1}}{(x/a-1)^k} = (-1)^k \frac{1}{a} [x^{n+k-1}] \frac{(x/a)^{k-1}}{(1-x/a)^k} \\ &\stackrel{7.19}{=} (-1)^k \frac{1}{a} \cdot \left(\frac{1}{a}\right)^{n+k-1} \binom{n+k-1}{k-1} = (-1)^k \left(\frac{1}{a}\right)^{n+k} \binom{n+k-1}{k-1} \quad \square \end{aligned}$$

Theorem 7.21. $\sum_{n=0}^{\infty} \binom{2n}{n} x^n = \frac{1}{\sqrt{1-4x}}$ is the generating function for the central binomial coefficients.

Proof. This can be solved using the Generalized Binomial Theorem:

$$\begin{aligned} \frac{1}{\sqrt{1-4x}} &= \sum_{n=0}^{\infty} \binom{-1/2}{n} (-4x)^n = 1 + \sum_{n=1}^{\infty} \frac{(-1/2)_n}{n!} (-4x)^n \\ &= 1 + \sum_{n=1}^{\infty} \frac{(-1)^n (2n-1)!!}{2^n n!} (-4x)^n = 1 + \sum_{n=1}^{\infty} \frac{2^n (2n-1)!!}{n!} x^n = 1 + \sum_{n=1}^{\infty} \frac{2^n n! (2n-1)!!}{n! n!} x^n \\ &= 1 + \sum_{n=1}^{\infty} \frac{(2n)!! (2n-1)!!}{n! n!} x^n = 1 + \sum_{n=1}^{\infty} \frac{(2n)!}{n! n!} x^n = 1 + \sum_{n=1}^{\infty} \binom{2n}{n} x^n \end{aligned}$$

Alternatively, we can use Lagrange Inversion. Note that $\binom{2n}{n} = [z^n](1+z)^{2n}$. We can thus use Theorem 11.10 with $W(z) = 1$ and $\phi(z) = (1+z)^2$. To find z , we solve $z = x\phi(z) = x(1+z)^2$. We obtain $xz^2 + (2x-1)z + x = 0$, so that using the quadratic formula and picking the root which is well-defined at $x = 0$ gives $z = \frac{-2x+1-\sqrt{1-4x}}{2x}$. Using this:

$$\begin{aligned} \sum_{n=0}^{\infty} \binom{2n}{n} x^n &= \sum_{n=0}^{\infty} x^n [z^n] (1+z)^{2n} = \frac{1}{1-x\phi'(z)} = \frac{1}{1-x \cdot 2(z+1)} = \frac{1}{1-x \cdot 2\left(\frac{-2x+1-\sqrt{1-4x}}{2x} + 1\right)} \\ &= \frac{1}{1-x \cdot 2\left(\frac{1-\sqrt{1-4x}}{2x}\right)} = \frac{1}{\sqrt{1-4x}} \quad \square \end{aligned}$$

Theorem 7.22. $\sum_{n=0}^{\infty} \binom{2n+k}{n} x^n = \frac{1}{\sqrt{1-4x}} \left(\frac{1-\sqrt{1-4x}}{2x} \right)^k$

Proof. This is a more general version of Theorem 7.21, but it can be proved similarly. Note that $\binom{2n+k}{n} = [z^n](1+z)^{2n+k}$. We can thus use Theorem 11.10 with $W(z) = (1+z)^k$ and $\phi(z) = (1+z)^2$. From before, we have $z = \frac{-2x+1-\sqrt{1-4x}}{2x}$. Using this:

$$\begin{aligned} \sum_{n=0}^{\infty} \binom{2n+k}{n} x^n &= \sum_{n=0}^{\infty} x^n [z^n] (1+z)^{2n+k} = \sum_{n=0}^{\infty} x^n [z^n] (1+z)^k (1+z)^{2n} \\ &= \frac{(1+z)^k}{1-x\phi'(z)} = \frac{(1+z)^k}{1-x \cdot 2(z+1)} = \frac{(1+\frac{-2x+1-\sqrt{1-4x}}{2x})^k}{1-x \cdot 2\left(\frac{-2x+1-\sqrt{1-4x}}{2x} + 1\right)} \\ &= \frac{\left(\frac{1-\sqrt{1-4x}}{2x}\right)^k}{1-x \cdot 2\left(\frac{1-\sqrt{1-4x}}{2x}\right)} = \frac{1}{\sqrt{1-4x}} \left(\frac{1-\sqrt{1-4x}}{2x} \right)^k \quad \square \end{aligned}$$

Theorem 7.23. $\sum_{n=0}^{\infty} \frac{k(2n+k-1)!}{n!(n+k)!} x^n = \left(\frac{1-\sqrt{1-4x}}{2x} \right)^k$

Proof. We use Theorem 7.22. We calculate:

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

Taking the derivative, we have:

$$\begin{aligned} A'(x) &= k \sum_{n=1}^{\infty} \binom{2(n-1)+k+1}{n-1} x^{n-1} = k \sum_{n=0}^{\infty} \binom{2n+k+1}{n} x^n = \frac{k}{\sqrt{1-4x}} \left(\frac{1-\sqrt{1-4x}}{2x} \right)^{k+1} \\ &= \frac{k}{\sqrt{1-4x}} \left(\frac{1-\sqrt{1-4x}}{2x} \right)^{k-1} \frac{2-4x-2\sqrt{1-4x}}{4x^2} \end{aligned}$$

Letting $u = \frac{1-\sqrt{1-4x}}{2x}$, we have $\frac{du}{dx} = \frac{4x/\sqrt{1-4x}-2+2\sqrt{1-4x}}{4x^2} = \frac{2-4x-2\sqrt{1-4x}}{4x^2\sqrt{1-4x}}$. Hence, $A'(x) = k(u)^{k-1} \frac{du}{dx}$. Therefore, $A(x) = u^k$, giving us the claim. \square

7.3 Wilf Rules

Theorem 7.24 (Wilf Rules for OGFs).

- I. If $F(x) \xleftrightarrow{\text{ogf}} \{f_n\}$, then for $k \geq 0$, $\{f_{n+k}\} \xleftrightarrow{\text{ogf}} \frac{F(x) - f_0 - f_1x - \dots - f_{k-1}x^{k-1}}{x^k}$
- II. If $F(x) \xleftrightarrow{\text{ogf}} \{f_n\}$ and P is a polynomial, then $\{P(n)f_n\} \xleftrightarrow{\text{ogf}} P(xD)F(x)$, where D is the derivative operator with respect to x
 - \hookrightarrow For example, $(xD)^2 F(x) = x \frac{d}{dx} \left[x \frac{d}{dx} F(x) \right]$
- III. If $F(x) \xleftrightarrow{\text{ogf}} \{f_n\}$ and $G(x) \xleftrightarrow{\text{ogf}} \{g_n\}$, then $F(x)G(x) \xleftrightarrow{\text{ogf}} \left\{ \sum_{k=0}^n a_k b_{n-k} \right\}$

IV. If $F(x) \xleftrightarrow{\text{ogf}} \{f_n\}$ and $k \in \mathbb{N}$, then $F^k(x) \xleftrightarrow{\text{ogf}} \left\{ \sum_{n_1+\dots+n_k=n} a_{n_1} \cdots a_{n_k} \right\}$

V. If $F(x) \xleftrightarrow{\text{ogf}} \{f_n\}$, then $\frac{1}{1-x}F(x) \xleftrightarrow{\text{ogf}} \left\{ \sum_{k=0}^n f_k \right\}$, i.e., $\frac{1}{1-x}F(x)$ is the generating function for the partial sums

VI. If $F(x) \xleftrightarrow{\text{ogf}} \{f_n\}$, then $x^k F(x) \xleftrightarrow{\text{ogf}} \{a_{n-k}\}$, where $a_{n-k} = 0$ when $n - k < 0$

Example 7.25. Suppose $a_{n+2} = 3a_{n+1} - a_n$ with $a_0 = 1$ and $a_1 = -3$. Then by Wilf I, $\frac{A(x)-a_0-a_1x^1}{x^2} = -3\frac{A(x)-a_0}{x} - A(x)$. Rearrange to obtain $A(x) - 1 + 3x = -3xA(x) + 3x - A(x)x^2$, or $A(x) = \frac{1}{1+3x+x^2}$.

Example 7.26. Suppose $a_n = n^2$, as in Example 7.32. Then by Wilf II and Lemma 7.16:

$$A(x) = (xD)^2 \frac{1}{1-x} = x \frac{d}{dx} \left[x \frac{d}{dx} \frac{1}{1-x} \right] = x \frac{d}{dx} \frac{x}{(1-x)^2} = \frac{x(1+x)}{(1-x)^3}$$

Theorem 7.27 (Wilf Rules for EGFs).

I. If $F(x) \xleftrightarrow{\text{egf}} \{f_n\}$, then for $k \geq 0$, $\{f_{n+k}\} \xleftrightarrow{\text{egf}} D^k F(x)$

↳ Shifting in the other direction is harder. Generally, just rewrite the recurrence so that it is a shift in the direction above (e.g., $a_{n+2} = a_{n+1} + 2a_n$ rather than $a_n = a_{n-1} + 2a_{n-2}$)

II. If $F(x) \xleftrightarrow{\text{egf}} \{f_n\}$ and P is a polynomial, then $\{P(n)f_n\} \xleftrightarrow{\text{egf}} P(xD)F(x)$, where D is the derivative operator with respect to x

↳ This is the same as for OGFs

III. If $F(x) \xleftrightarrow{\text{egf}} \{f_n\}$ and $G(x) \xleftrightarrow{\text{egf}} \{g_n\}$, then $F(x)G(x) \xleftrightarrow{\text{egf}} \left\{ \sum_{k=0}^n \binom{n}{k} f_k g_{n-k} \right\}$

Proof. We prove Wilf III using the corresponding rule for ordinary generating functions:

$$F(x)G(x) = \left(\sum_{n=0}^{\infty} \frac{f_n}{n!} x^n \right) \left(\sum_{n=0}^{\infty} \frac{g_n}{n!} x^n \right) \stackrel{\text{III}}{=} \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{f_k g_{n-k}}{k!(n-k)!} \right) x^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \binom{n}{k} f_k g_{n-k} \right) \frac{x^n}{n!}$$

□

Theorem 7.28 (Wilf Rules for DGFs).

I. If $F(s) \xleftrightarrow{\text{dgf}} \{f_n\}$ and $G(s) \xleftrightarrow{\text{dgf}} \{g_n\}$, then $F(s)G(s) \xleftrightarrow{\text{dgf}} \left\{ \sum_{j+k=n} f_j g_k \right\} = \left\{ \sum_{d|n} f_d g_{n/d} \right\}$

↳ For example, $[4^{-s}]F(s)G(s) = f_1 g_4 + f_2 g_2 + f_4 g_1$

II. If $F(s) \xleftrightarrow{\text{dgf}} \{f_n\}$ then $F(s)^k \xleftrightarrow{\text{dgf}} \left\{ \sum_{\substack{d_1, \dots, d_k \in [n] \\ d_1 \cdots d_k = n}} f_{d_1} \cdots f_{d_k} \right\}$

7.4 Recurrences

Concept 7.29. Given a recurrence relation defining a sequence $\{a_n\}$, a closed form for the generating function $A(x)$ can then be found. To find a closed form for the original sequence, the typical procedure (for linear recurrences) is to apply partial fraction decomposition to obtain $A(x)$ as a sum of factors of the form $\frac{c}{1-ax}$. Then, applying $[x^n]$, linearity, and Lemma 7.16 yields a_n . For non-linear recurrences, the procedure may be more complicated (or exponential generating functions may be required).

Example 7.30. Let $a_n = 2a_{n-1}$ with a_0 given. Let $\{a_n\} \longleftrightarrow A(x)$. Then $\sum_{n=0}^{\infty} a_{n+1}x^{n+1} = 2\sum_{n=0}^{\infty} a_nx^{n+1}$, so that $A(x) - a_0 = 2xA(x)$. Solving, $A(x) = \frac{a_0}{1-2x}$. Then $a_n = [x^n] \frac{a_0}{1-2x} = a_0 2^n$ by Lemma 7.16.

Example 7.31. Let $b_{n+1} = 2b_n + 1$ and $b_0 = 1$. Then $\sum_{n=0}^{\infty} b_{n+1}x^{n+1} = 2\sum_{n=0}^{\infty} b_nx^{n+1} + \sum_{n=0}^{\infty} x^{n+1}$, so that $B(x) - b_0 = 2xB(x) + \frac{x}{1-x}$. Solving, $B(x) = \frac{1}{(1-x)(1-2x)} = -\frac{1}{1-x} + \frac{2}{1-2x}$. Thus, $b_n = [x^n] \left(-\frac{1}{1-x} + \frac{2}{1-2x} \right) = -1 + 2(2)^n = 2^{n+1} - 1$.

Example 7.32. Let $a_n = n^2$. We find a closed form for the generating function $A(x)$. Note that $a_{n+2} = 2a_{n+1} - a_n + 2$. Thus, $\sum_{n=0}^{\infty} a_{n+2}x^{n+2} = 2\sum_{n=0}^{\infty} a_{n+1}x^{n+2} - \sum_{n=0}^{\infty} a_nx^{n+2} + 2\sum_{n=0}^{\infty} x^{n+2}$. So, $A(x) - x = 2xA(x) - x^2A(x) + \frac{2x^2}{1-x}$. Solving, $(x^2 - 2x + 1)A(x) = x + \frac{2x^2}{1-x}$, or $A(x) = \frac{x^2+x}{(1-x)^3}$.

Theorem 7.33. Let p, q be coprime polynomials with $0 \leq k = \deg p < \deg q = m$. Write $q(x) = \sum_{j=0}^m q_j x^j$ and suppose $q_0 \neq 0$. Let $f(x) = \frac{p(x)}{q(x)} \xrightarrow{\text{ogf}} \{f_n\}_{n \geq 0}$. Then for all $n \geq 0$, f_n satisfies the recurrence $q_0 f_{n+m} + q_1 f_{n+m-1} + \cdots + q_m f_n = 0$ with initial conditions $f_j = f^{(j)}(0)/j!$ for $0 \leq j \leq m-1$.

↳ This implies that given a linear recurrence with finitely many terms, and a set of initial conditions, the resulting solution as a power series can be represented as a rational function

Proof. The initial conditions follow by writing the Taylor series of f . The recurrence follows by rewriting $q(x)f(x) = p(x)$, then applying $[x^{n+m}]$ to each side. The LHS simplifies by Wilf III, and RHS is 0 since $\deg p < m \leq m+n$. \square

Example 7.34. Let $G(x) = \frac{1}{1-ax-bx^2-cx^3}$ for $c \neq 0$. g_n satisfies the recurrence $g_{n+3} = ag_{n+2} + bg_{n+1} + cg_n$ by Theorem 7.33. The initial conditions are $g_0 = G(0) = 1$, $g_1 = G'(0) = a$, and $g_2 = G''(0)/2 = a^2 + b$.

Example 7.35. Suppose $a_{n+1} = 2a_n + n$ with $a_0 = 1$. Let $b_n = n$, so that $B(x) = xD \frac{1}{1-x} = \frac{x}{(1-x)^2}$ by Wilf II. By Wilf I, $\frac{A(x)-a_0}{x} = 2A(x) + B(x)$. Rearrange and substitute to obtain $A(x) = \frac{1-2x+2x^2}{(1-2x)(1-x)^2}$.

Example 7.36. We find a closed form for $a_{n+1} = (n+1)a_n + n!$ with $a_0 = 0$. Using exponential generating function, Wilf I on LHS and Wilf II on RHS gives $A'(x) = (xD + 1)A(x) + \frac{1}{1-x}$, so that $A'(x) = xA'(x) + A(x) + \frac{1}{1-x}$. This can be solved using integrating factors. Indeed, $(1-x)A'(x) - A(x) = \frac{1}{1-x}$, so that $D[(1-x)A(x)] = \frac{1}{1-x}$. Then $A(x) = -\frac{\ln(1-x)}{1-x}$. Note that this is of the form $A(x) = [\ln(1-x)]D[\ln(1-x)]$, so that by the chain rule $A(x) = \frac{1}{2}D[\ln(1-x)^2]$. Then letting $h_n = \sum_{k=1}^n 1/k$ be the n th harmonic number, we have:

$$\begin{aligned} a_n &= n![x^n]A(x) = \frac{n!}{2}[x^n]D(\ln(1-x)^2) = \frac{n!}{2}[x^{n+1}]xD(\ln(1-x)^2) = \frac{(n+1)!}{2}[x^{n+1}]\ln(1-x)^2 \\ &= \frac{(n+1)!}{2} \sum_{k=0}^{n+1} ([x^k]\ln(1-x))([x^{n-k}]\ln(1-x)) = \frac{(n+1)!}{2} \sum_{k=1}^n \frac{1}{k} \cdot \frac{1}{n+1-k} \\ &= \frac{(n+1)!}{2} \sum_{k=1}^n \frac{1}{n+1} \left[\frac{1}{k} + \frac{1}{n+1-k} \right] = n!h_n \end{aligned}$$

Example 7.37. We find a closed form for the sequence $a_{n+1} = (n+1)a_n + 2(n+1)!$ with $a_0 = 0$. Using Wilf I and Wilf II, we have $A'(x) = (xD + 1)A(x) + 2D[\frac{1}{1-x}]$. Simplifying, $(1-x)A'(x) = A(x) + \frac{2}{(1-x)^2}$. Using integrating factors, we can rewrite it as $D[(1-x)A(x)] = \frac{2}{(1-x)^2}$, so that integrating yields $(1-x)A(x) = \frac{2}{1-x} + c$. Substitute $x = 0$ and $A(0) = a_0 = 0$ to obtain $c = -2$. Then $A(x) = \frac{2}{(1-x)^2} - \frac{2}{(1-x)}$. Hence:

$$\begin{aligned} a_n &= n![x^n]A(x) = 2n! \left([x^n] \frac{1}{(1-x)^2} - [x^n] \frac{1}{1-x} \right) = 2n! \left([x^n]D \left[\frac{1}{(1-x)} \right] - 1 \right) \\ &= 2n! \left([x^{n+1}]xD \left[\frac{1}{(1-x)} \right] - 1 \right) = 2n! \left((n+1)[x^{n+1}] \frac{1}{(1-x)} - 1 \right) = 2n!((n+1) - 1) \\ &= 2n \cdot n! \end{aligned}$$

Example 7.38. Consider counting the number of n -words q_n over alphabet $[k]$ that do not contain a particular substring of length m . Assume this substring has no proper substrings that appear at both the beginning and end of the word (for example, 111211 is invalid since it contains 11 at the beginning and end). The generating function is $Q(x) = \frac{1}{x^m - kx + 1}$.

Proof. $q_n = k^n$ when $n < m$, since the substring cannot be present. Otherwise, proceed with negative counting. Condition on the length r of the word before the rightmost instance of the substring. This word has no restrictions, so there are k^r options. The remaining $n - r - m$ letters cannot contain the substring, so are counted by q_{n-r-m} . Then, we have $q_n = k^n - \sum_{r=0}^{n-m} k^r q_{n-r-m}$. The generating function thus satisfies:

$$\begin{aligned} Q(x) &= \sum_{n=0}^{\infty} \left[k^n - \sum_{r=0}^{n-m} k^r q_{n-r-m} \right] x^n = \frac{1}{1-kx} - \sum_{n=0}^{\infty} \sum_{r=0}^{n-m} k^r q_{n-r-m} x^n \\ &= \frac{1}{1-kx} - \sum_{n=m}^{\infty} \sum_{r=0}^{n-m} k^r q_{n-r-m} x^n = \frac{1}{1-kx} - \sum_{n=0}^{\infty} \sum_{r=0}^n k^r q_{n-r} x^{n+m} \\ &= \frac{1}{1-kx} - x^m \sum_{n=0}^{\infty} \sum_{r=0}^n k^r q_{n-r} x^n = \frac{1}{1-kx} - x^m \frac{1}{1-kx} Q(x) = \frac{1 - x^m Q(x)}{1 - kx} \end{aligned}$$

Therefore, $Q(x) = \frac{1}{x^m - kx + 1}$. Alternatively, we can proceed using Inclusion/Exclusion. Let p_k be the property an instance of the substring begins at position k . Fix j of the properties; the number of words satisfying these j are k^{n-mj} , since this is the number of letters remaining. The number of ways to choose j of the properties is $\binom{n-mj+j}{j}$; viewing each instance of the substring as a single token, then the n -word becomes a sequence of $n - mj + j$ tokens, j of which must be chosen to be the substring in question. Thus, $q_n = \sum_{j=0}^{\infty} (-1)^j \binom{n-mj+j}{j} k^{n-mj}$. The generating function is thus:

$$\begin{aligned} Q(x) &= \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} (-1)^j \binom{n-mj+j}{j} k^{n-mj} x^n = \sum_{j=0}^{\infty} (-1)^j k^{-mj} \sum_{n=0}^{\infty} \binom{n-mj+j}{j} (kx)^n \\ &= \sum_{j=0}^{\infty} (-1)^j k^{-mj} \sum_{n=0}^{\infty} \binom{n}{j} (kx)^{n+mj-j} = \sum_{j=0}^{\infty} (-1)^j k^{-mj} (kx)^{mj-j} \sum_{n=0}^{\infty} \binom{n}{j} (kx)^n \\ &= \sum_{j=0}^{\infty} (-1)^j k^{-mj} (kx)^{mj-j} \frac{(kx)^j}{(1-kx)^{j+1}} = \sum_{j=0}^{\infty} (-1)^j \frac{x^{mj}}{(1-kx)^{j+1}} = \frac{1}{1-kx} \sum_{j=0}^{\infty} \left(-\frac{x^m}{1-kx} \right)^j \\ &= \frac{1}{1-kx} \frac{1}{1 + \frac{x^m}{1-kx}} = \frac{1}{x^m - kx + 1} \quad \square \end{aligned}$$

7.5 EGF Examples

Example 7.39. Suppose n people are split into two (possibly empty) groups, and each person in the first group is assigned one of three colors. There are $c_n = 4^n$ ways to form this, since we can simply assign each person one of the properties {group 2, color 1, color 2, color 3}.

Instead, we could have conditioned on the size of the first group k to get $\sum_{k=0}^n \binom{n}{k} 3^k$. Using the Binomial Theorem yields the results. Or, use Wilf III to write the exponential generating function $C(x) = \left(\sum_{n=0}^{\infty} 3^n \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) = e^{3x} e^x = e^{4x}$. The final answer is then $c_n = n![x^n]C(x) = n![x^n] \sum_{n=0}^{\infty} \frac{(4x)^n}{n!} = 4^n$.

Example 7.40. Now suppose we also must assign a leader to the second group. There are $c_n = n4^{n-1}$ ways to do this, since we can first choose the leader and then distribute the remaining people as before.

Or, we can once again condition on the size of the first group k to get $\sum_{k=0}^n \binom{n}{k} 3^k (n-k)$. Use Wilf III to write the exponential generating function:

$$C(x) = \left(\sum_{n=0}^{\infty} 3^n \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} n \frac{x^n}{n!} \right) = e^{3x} x \left(\sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} \right) = e^{3x} x \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) = x e^{4x}$$

Now we can obtain the solution $c_n = n![x^n]x e^{4x} = n![x^{n-1}]e^{4x} = n4^{n-1}$.

Example 7.41. Now suppose instead of assigning a leader to the second group, we must order them. The closed form is not obvious in this case. Conditioning on the size of the first group k , we get $c_n = \sum_{k=0}^n \binom{n}{k} 3^k (n-k)!$. Use Wilf III to write the exponential generating function:

$$C(x) = \left(\sum_{n=0}^{\infty} 3^n \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} n! \frac{x^n}{n!} \right) = \frac{e^{3x}}{1-x}$$

A closed form for c_n can be found through analytic combinatorics. Namely, $c_n = e^3 \Gamma(1+n, 3)$ where $\Gamma(x, s) = \int_x^{\infty} t^{s-1} e^{-t} dt$ for $s \in \mathbb{C}$.

Example 7.42. Suppose n people are split into 3 groups. Each person in the first group is assigned one of two colors and each person in the second group is assigned one of two colors. There are $c_n = 5^n$ ways to form this, since we are effectively forming 5 groups based on the original groupings and the colors.

Instead, we could have conditioned on the size of the groups to obtain $c_n = \sum_{i=0}^n \sum_{j=0}^{n-i} \binom{n}{i} \binom{n-i}{j} 2^i 2^j$. Use the Binomial Theorem twice to obtain the results. Or, use Wilf III twice to write the exponential generating function $C(x) = e^{5x}$, so that $c_n = [x^n]C(x) = 5^n$.

Example 7.43. Now suppose that the second group has a leader and the third group is ordered. Then we obtain:

$$c_n = \sum_{i=0}^n \sum_{j=0}^{n-i} \binom{n}{i} \binom{n-i}{j} 2^i 2^j j (n-i-j)! = \sum_{i=0}^n \binom{n}{i} 2^i \sum_{j=0}^{n-i} \binom{n-i}{j} 2^j j (n-i-j)!$$

Using Wilf II, the EGF corresponding to $2^j j$ is $xD[e^{2x}] = 2xe^{2x}$. Thus, using Wilf III twice, we have $C(x) = 2xe^{4x}/(1-x)$.

Example 7.44. Let $c_0 = 1$ and for $n > 0$ let c_n count the number of n -permutations in which each cycle is colored red, green, or blue. Instead of forming the permutation then coloring the elements, we can color the elements and then form permutations within each color. Conditioning on the number of red and green elements, we have:

$$c_n = \sum_{i=0}^n \sum_{j=0}^{n-i} \binom{n}{i} \binom{n-i}{j} i! j! (n-i-j)! = n! \sum_{i=0}^n \sum_{j=0}^{n-i} 1 = n! \sum_{i=0}^n (n-i+1) = \frac{n!(n+1)(n+2)}{2} = \frac{(n+2)!}{2}$$

The exponential generating function for c_n is thus:

$$C(x) = \frac{1}{2} D^2 \left[\frac{1}{1-x} \right] = \frac{1}{2} \left[\frac{2}{(1-x)^3} \right] = \frac{1}{(1-x)^3}$$

If instead we let $a_{n+2} = c_n$ and $a_0 = a_1 = 1$, then we have $A(x) = \frac{1}{1-x}$.

Proposition 7.45. The exponential generating function for $\binom{n}{k}$ with respect to n is $B_k(x) = \frac{x^k}{k!} e^x$.

Proof.

$$B_k(x) = \sum_{n=0}^{\infty} \binom{n}{k} \frac{x^n}{n!} = \sum_{n=k}^{\infty} \frac{1}{k!(n-k)!} x^n = \frac{x^k}{k!} \sum_{n=k}^{\infty} \frac{x^{n-k}}{(n-k)!} = \frac{x^k}{k!} \sum_{n=0}^{\infty} \frac{x^n}{n!} = \frac{x^k}{k!} e^x$$

Alternatively, we can prove this via induction. $B_0(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$, so that the base case holds. Then:

$$\begin{aligned} B_{k+1}(x) &= \sum_{n=k+1}^{\infty} \binom{n}{k+1} \frac{x^n}{n!} = \sum_{n=k+1}^{\infty} \frac{n}{k+1} \binom{n-1}{k} \frac{x^n}{n!} = \frac{x}{k+1} \sum_{n=k+1}^{\infty} \binom{n-1}{k} \frac{x^{n-1}}{(n-1)!} \\ &= \frac{x}{k+1} \sum_{n=k}^{\infty} \binom{n}{k} \frac{x^n}{n!} = \frac{x}{k+1} \cdot \frac{x^k}{k!} e^x = \frac{x^{k+1}}{(k+1)!} e^x \end{aligned} \quad \square$$

Example 7.46. Consider the sequence with $a_0 = 1$, $a_1 = 1$, $a_2 = 2$, and the recurrence $a_{n+1} = (n+1)a_n - \binom{n}{2}a_{n-2}$. To find the exponential generating function, we could shift the recurrence to $a_{n+3} = (n+3)a_{n+2} - \binom{n+2}{2}a_n$. Then, we expand the binomial as a polynomial and use Wilf I and II to obtain the ordinary differential equation:

$$A'''(x) = xA'''(x) + 3A''(x) - \frac{1}{2}x^2A''(x) - 2xA'(x) - A(x)$$

However, we can actually derive a simpler differential equation through a clever application of Wilf III. Note that:

$$\sum_{n=2}^{\infty} \binom{n}{2} a_{n-2} = \sum_{n=2}^{\infty} \sum_{k=0}^n \binom{n}{2} I[k=2] a_{n-k} = \frac{x^2}{2} A(x)$$

Using this, we obtain the differential equation $A'(x) = xA'(x) + A(x) - \frac{x^2}{2}A(x)$, or $(1-x)A'(x) = (1 - \frac{x^2}{2})A(x)$. This is separable, so that we need to evaluate the integral of $\frac{1-x^2/2}{1-x} = \frac{1}{2} + \frac{x}{2} + \frac{1}{2} \frac{1}{1-x}$. This yields $\log A(x) = c + \frac{1}{2}x + \frac{x^2}{4} - \frac{1}{2} \log(1-x)$. Since $a_0 = 1$, substituting $x = 0$ gives $c = 0$. Thus, $A(x) = \frac{e^{x/2+x^2/4}}{\sqrt{1-x}}$.

7.6 Fibonacci Numbers

Definition 7.47. The Fibonacci Numbers are those with the recurrence $f_{n+2} = f_{n+1} + f_n$ and initial values $f_0 = f_1 = 1$. They count the number of ways to tile a $1 \times n$ grid with monominoes and dominoes.

Proof. The base case of $f_0 = 1$ and $f_1 = 1$ holds. Now, for arbitrary $n > 1$, condition on whether the first tile is a monomino or domino. In the former, there are f_{n+1} ways to cover the remaining squares. In the latter, there are f_n ways to cover the remaining squares. \square

Lemma 7.48. The generating function of the Fibonacci Numbers is $F(x) = \frac{1}{1-x-x^2}$.

Proof. Using the recurrence and Wilf I, $\frac{F(x)-f_0-f_1x}{x^2} = \frac{F(x)-f_0}{x} + F(x)$. Rearrange to obtain the desired result. Or, using the combinatorial interpretation, let $a_{n,j}$ be the number of ways to tile a $1 \times n$ board with j tiles. Let $A_j(x)$ be the associated generating function in n . Choosing any sequence of monominoes and dominoes yields $A_j(x) = (x^1 + x^2)^j$. By composition, $F(x) = \sum_{j=0}^{\infty} A_j(x) = \sum_{j=0}^{\infty} (x+x^2)^j = \frac{1}{1-(x+x^2)} = \frac{1}{1-x-x^2}$. \square

Example 7.49. $\sum_{n=0}^{\infty} f_{2n+1}x^n = \frac{1}{1-3x+x^2}$

Proof. Let $y = \sqrt{x}$. Then:

$$\begin{aligned} \sum_{n=0}^{\infty} f_{2n+1}x^n &= \sum_{n=0}^{\infty} f_{2n+1}y^{2n} = \frac{1}{y} \sum_{n=0}^{\infty} f_{2n+1}y^{2n+1} = \frac{1}{2y} \left[\sum_{n=0}^{\infty} f_n y^n - \sum_{n=0}^{\infty} f_n (-y)^n \right] \\ &= \frac{1}{2y} \left[\frac{1}{1-y-y^2} - \frac{1}{1+y-y^2} \right] = \frac{1}{1-3y^2+y^4} = \frac{1}{1-3x+x^2} \end{aligned}$$

Or, it suffices to show that $a_n = f_{2n+1}$ satisfy the recurrence $a_{n+2} = 3a_{n+1} - a_n$ and that $a_0 = \frac{1}{1-3(0)+0^2} = 1$ and $a_1 = -(1-3(0)+0^2)^{-2} \cdot (2(0)-3) = 3$ by Theorem 7.33. $a_0 = f_1 = 1$ and $a_1 = f_3 = 3$, so the initial conditions are satisfied. The recurrence holds since:

$$a_{n+2} = f_{2n+5} = f_{2n+4} + f_{2n+3} = 2f_{2n+3} + f_{2n+2} = 3f_{2n+3} - f_{2n+1} = 3a_{n+1} - a_n \quad \square$$

Example 7.50. We find a closed form for $a_n = na_{n-1} + n(n-1)a_{n-2}$ with $a_0 = a_1 = 1$. Rewrite this as $a_{n+2} = (n+2)a_{n+1} + (n+2)(n+1)a_n$. By Wilf I and II, $A''(x) = (xD+2)A'(x) + (xD+2)(xD+1)A(x)$. Simplify to obtain $(x^2+x-1)A''(x) + (4x+2)A'(x) + 2A(x) = 0$. Note that this is of the form $D^2[p(x)A(x)] = 0$ for $p(x) = x^2+x-1$, so that $A(x) = \frac{c_1x+c_0}{x^2+x-1}$. $A(0) = 1$ implies $c_0 = -1$. Taking the derivative, $a_1 = A'(0) = \frac{(0^2+0-1)(c_1)-(c_1(0)-1)(2(0)-1)}{(0^2+0-1)^2} = 1 - c_1$, so that $c_1 = 0$. Thus, $A(x) = \frac{1}{1-x-x^2}$. This is the generating function for the Fibonacci Numbers, so that $a_n = n! \cdot f_n$.

Lemma 7.51.

- (a) $f_{n+2} - 1 = f_n + \cdots + f_0$
- (b) $f_{2n} - 1 = f_1 + f_3 + \cdots + f_{2n-1}$
- (c) $f_{2n+1} = f_0 + f_2 + \cdots + f_{2n}$

Proof.

- (a) LHS counts number of ways to tile a $1 \times (n+2)$ grid with at least one domino. RHS conditions on the number of tiles to the left of the rightmost domino. Or, we can use generating functions. By Wilf I, the generating function for RHS is:

$$\begin{aligned} \frac{F(x) - 1 - x}{x^2} - \frac{1}{1-x} &= \frac{\frac{1}{1-x-x^2} - 1 - x}{x^2} - \frac{1}{1-x} = \frac{\frac{2x^2+x^3}{1-x-x^2}}{x^2} - \frac{1}{1-x} = \frac{2+x}{1-x-x^2} - \frac{1}{1-x} \\ &= \frac{1}{(1-x-x^2)(1-x)} = \frac{F(x)}{1-x} \end{aligned}$$

But this is the generating function for the partial sums of f_n by Lemma 7.18.

- (b) LHS counts number of ways to tile a $1 \times 2n$ grid with at least one monomino. RHS conditions on the number of tiles to the left of the rightmost monomino.
- (c) LHS counts number of ways to tile a $1 \times (2n+1)$ grid, which necessarily has at least one monomino. RHS conditions on the number of tiles to the left of the rightmost monomino. \square

Lemma 7.52. $f_j = \sum_{k=0}^{\infty} \binom{j-k}{k}$

Proof. Condition on the number of dominoes k . Then precisely $j-k$ tiles are needed; choose k of them to be the dominoes. Or, prove it with generating functions:

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

Then:

$$f_j = [x^j] \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \binom{n}{k} x^{n+k} = \sum_{k=0}^{\infty} [x^j] \sum_{n=k}^{\infty} \binom{n}{k} x^{n+k} = \sum_{k=0}^{\infty} [x^{j-k}] \sum_{n=k}^{\infty} \binom{n}{k} x^n = \sum_{k=0}^{\infty} \binom{j-k}{k} \quad \square$$

Lemma 7.53. $\sum_{k=1}^n \binom{n}{k} f_{k-1} = f_{2n-1}$

Proof. RHS counts the number of tilings of a $1 \times (2n-1)$ grid into monominoes and dominoes. LHS conditions on the number of monominoes k among the first n tiles. There is at least one since otherwise the n dominoes would fill a $1 \times 2n$ grid. Choose which are monominoes $\binom{n}{k}$ ways. The monominoes take k squares, and the $n-k$ dominoes take $2(n-k)$ squares, so that $2n-k$ squares are taken by the first n tiles. The number of remaining squares is $2n-1-(2n-k) = k-1$; cover these f_{k-1} ways. \square

Lemma 7.54. For $n \geq 0$, $f_{2n+1} = f_n f_{n+1} + f_{n-1} f_n$.

Proof. LHS counts number of ways to cover a $1 \times (2n+1)$ board with monominoes and dominoes. Split this into a $1 \times n$ board on the left and a $1 \times (n+1)$ board on the right. RHS conditions on whether a domino overlaps the split between these. If not, cover the left f_n ways and the right f_{n+1} ways. If so, cover the rest of the left f_{n-1} ways and the rest of the right f_n ways. \square

Lemma 7.55. $\sum_{k=0}^n f_{2k+1} (-1)^{n-k} = f_n f_{n+1}$

Proof. Let $H(x) = \sum_{n=0}^{\infty} f_n f_{n+1} x^n$. By 7.54, $H(x) + xH(x) = \sum_{n=0}^{\infty} f_{2n+1} x^n$, so that $H(x) = \frac{1}{1+x} \sum_{n=0}^{\infty} f_{2n+1} x^n$. Then:

$$\begin{aligned} f_n f_{n+1} &= [x^n] \frac{1}{1+x} \sum_{k=0}^{\infty} f_{2k+1} x^k = [(-x)^n] \frac{1}{1-x} \sum_{k=0}^{\infty} f_{2k+1} (-x)^k = (-1)^n [x^n] \frac{1}{1-x} \sum_{k=0}^{\infty} f_{2k+1} (-x)^k \\ &= (-1)^n \sum_{k=0}^n f_{2k+1} (-1)^k = \sum_{k=0}^n f_{2k+1} (-1)^{n-k} \quad \square \end{aligned}$$

Lemma 7.56. The closed form for the Fibonacci Numbers is $f_n = \frac{\varphi^{n+1} - \psi^{n+1}}{\sqrt{5}}$ where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio and $\psi = \frac{1-\sqrt{5}}{2}$ is its conjugate.

Proof. First note $\varphi\psi = -1$ and $\varphi + \psi = 1$. Perform partial fraction decomposition to obtain $F(x) = \frac{1}{1-x-x^2} = \frac{-1}{(x+\varphi)(x+\psi)} = \frac{1/\sqrt{5}}{(x+\varphi)} - \frac{1/\sqrt{5}}{(x+\psi)}$. Then using Lemma 7.17:

$$f_n = [x^n] F(x) = \frac{1}{\sqrt{5}} \left([x^n] \frac{1}{x+\varphi} - [x^n] \frac{1}{x+\psi} \right) = \frac{-(-\varphi)^{-n-1} + (-\psi)^{-n-1}}{\sqrt{5}} = \frac{\varphi^{n+1} - \psi^{n+1}}{\sqrt{5}}$$

This could also be solved using exponential generating functions. Let $F(x) \xleftrightarrow{\text{egf}} \{f_n\}$. Then $F''(x) = F'(x) + F(x)$ by Wilf I with $F(0) = f_0 = 1$ and $F'(0) = f_1 = 1$. This is a second order linear homogeneous differential equation. The characteristic polynomial is $r^2 - r - 1$ with roots φ, ψ , so that $G(x) = c_1 e^{\varphi x} + c_2 e^{\psi x}$. Applying the initial conditions, we obtain $c_1 = \frac{1-\psi}{\varphi-\psi} = \frac{\varphi}{\sqrt{5}}$ and $c_2 = \frac{\varphi-1}{\varphi-\psi} = -\frac{\psi}{\sqrt{5}}$, so that $F(x) = \frac{1}{\sqrt{5}} [\varphi e^{\varphi x} - \psi e^{\psi x}]$. Thus:

$$f_n = n! [x^n] \left(\frac{1}{\sqrt{5}} [\varphi e^{\varphi x} - \psi e^{\psi x}] \right) = \frac{\varphi}{\sqrt{5}} n! [x^n] e^{\varphi x} - \frac{\psi}{\sqrt{5}} n! [x^n] e^{\psi x} = \frac{\varphi^{n+1} - \psi^{n+1}}{\sqrt{5}} \quad \square$$

Corollary 7.57. f_n can be computed by rounding $\frac{\varphi^{n+1}}{\sqrt{5}}$ to the nearest integer.

Proof. $\psi/\sqrt{5} = -0.28$ which is less than 0.5 in absolute value, and since $|\psi| < 1$, we have that $|\psi^{n+1}/\sqrt{5}| < 0.5$. Thus, this factor is negligible under rounding. \square

Example 7.58. The number of binary strings with n digits and no consecutive 1s is given by f_{n+1} .

Proof. Let this be t_n . There is 1 0-digit string and 2 1-digit strings, so that $t_0 = f_1$ and $t_1 = f_2$. A string counted by t_n either begins with 0 and is followed by a string in t_{n-1} , or begins with 10 and is followed by a string in t_{n-2} . Then $t_n = t_{n-1} + t_{n-2}$, so that $t_n = f_{n+1}$.

Or, use generating functions. Build a string by joining k entries from $\{0, 10\}$ and then optionally appending a 1. The generating function holding the number of ways to do this grouped by the length of the string is $A_k(x) = (x + x^2)^k (1 + x)$. Sum over all k to obtain $A(x) = \frac{1+x}{1-x-x^2}$. Then $t_n = [x^n] \frac{1+x}{1-x-x^2} = [x^n] \frac{1}{1-x-x^2} + [x^n] \frac{x}{1-x-x^2} = f_n + f_{n-1} = f_{n+1}$. \square

7.7 Catalan Numbers

Definition 7.59. The Catalan Numbers c_n count the number of sequences of $2n$ values in $\{\pm 1\}$ that add to 0 and have no negative partial sums. For example, all of the following are realizations of the Catalan Numbers:

- ↳ A person begins with an empty jar, and each day adds a penny or removes a penny. After $2n$ days it is empty again. How many ways could this have occurred?
- ↳ Two candidates received an equal number of votes. How many ways can the votes be tallied so that candidate A never trails candidate B ?
- ↳ Consider a string of matched parentheses such that all left parentheses are eventually closed by right parentheses. How many such strings are there?

We will use the last interpretation to derive more about the Catalan Numbers.

Definition 7.60. A string of parentheses is called legal if there are an equal number of left and right parentheses, and while reading from left to right, the number of right parentheses never exceeds the number of left parentheses. Let C_n denote the set of all legal strings of $2n$ parentheses, and let $c_n = |C_n|$. Note $C_1 = \{()\}$, $C_2 = \{()(), (())\}$, and $C_3 = \{()()(), (())(), ()()(), ((())), (())()()\}$, so that $c_1 = 1$, $c_2 = 2$, and $c_3 = 5$.

For $w \in C_n$, let the index of w be the number of parenthesis pairs in the first legal substring (and let this substring be the prefix). For example, the indices in C_3 above are 1, 2, 1, 3, 3 respectively. w is primitive if its index is n . The number of primitive strings is c_{n-1} (indeed, form a bijection by removing/adding the outermost set of parentheses).

Theorem 7.61. The Catalan Numbers satisfy the recurrence $c_n = \sum_{k=1}^n c_{k-1}c_{n-k}$.

Proof. The number of strings in C_n with index k is $c_{k-1}c_{n-k}$, since the prefix is primitive and thus can be chosen c_{k-1} ways, the the remainder can be chosen c_{n-k} ways. Sum over all k to get $c_n = \sum_{k=1}^n c_{k-1}c_{n-k}$. \square

Lemma 7.62. The ordinary generating function for the Catalan numbers satisfies the equation $C(x) = 1 + xC(x)^2$.

Proof. The generating function for c_n is:

$$C(x) = \sum_{n=0}^{\infty} c_n x^n = c_0 + \sum_{n=1}^{\infty} \sum_{k=1}^n c_{k-1}c_{n-k}x^n = 1 + \sum_{k=1}^{\infty} c_{k-1} \sum_{n=k}^{\infty} c_{n-k}x^n = 1 + xC(x)^2 \quad \square$$

Lemma 7.63. The ordinary generating function for the Catalan numbers is $C(x) = \frac{1 - \sqrt{1 - 4x}}{2x}$.

Proof. This follows by applying the quadratic equation to Lemma 7.62 and picking the root that is well-defined at $x = 0$. \square

Theorem 7.64. The Catalan Numbers are given by $c_n = \frac{1}{n+1} \binom{2n}{n}$.

Proof. Using Example 2.42 (or the Generalized Binomial Theorem) and the generating function given in Lemma 7.63:

$$c_n = [x^n] \frac{1 - \sqrt{1 - 4x}}{2x} = [x^n] \frac{1}{2x} \left[2x + \sum_{n=2}^{\infty} \frac{2(2n-2)!}{n!(n-1)!} x^n \right] = [x^n] \sum_{n=0}^{\infty} \frac{(2n)!}{(n+1)!n!} x^n = \frac{1}{n+1} \binom{2n}{n}$$

Alternatively, we can use Lagrange Inversion. From Lemma 7.62, we have $C(x) = 1 + xC(x)^2$, or $C(x) - 1 = xC(x)^2$. Thus, take $z = C(x) - 1$ and $\phi(z) = (1+z)^2$. We note that $\phi(0) = 1 \neq 0$ and $z = x\phi(z)$ as required. Thus:

$$[x^n]z(x) = \frac{1}{n} [z^{n-1}](1+z)^{2n} = \frac{1}{n} [z^{n-1}] \sum_{k=0}^{2n} \binom{2n}{k} z^k = \frac{1}{n} \binom{2n}{n-1} = \frac{1}{n+1} \binom{2n}{n} \quad \square$$

Theorem 7.65. c_n count the number of ways to parenthesize a product of $n+1$ objects.

Proof. Write the multiplication scheme with outermost enclosing parentheses. We form a bijection to strings on the alphabet $\{A, B\}$ such that reading left to right, the number of A s never trails the number of B s. In particular, form a bijection by associating all multiplication symbols with A and all right parentheses with B . This results in a valid string since a right parenthesis must have an associated multiplication operation to its left. \square

Corollary 7.66. c_n counts the number of full binary trees with $n+1$ leaves.

Proof. There is a bijection between multiplication schemes given in Theorem 7.65 and full binary trees representing the order of multiplication.

Alternatively, we can use recursive structures and generating functions. Consider the combinatorial class \mathcal{G} of full binary trees, where the size of the tree is given by its number of leaf nodes. Any tree of size at least one contains a root followed optionally by two children, each of which are members of \mathcal{G} . Hence $\mathcal{G} = \{\cdot\} + \text{SEQ}_2(\mathcal{G})$. Hence, the ordinary generating function satisfies $G(x) = x + G(x)^2$. By the quadratic formula, $G(x) = \frac{1 \pm \sqrt{1-4x}}{2}$. In order to achieve nonnegative coefficients, we must have $G(x) = \frac{1 - \sqrt{1-4x}}{2}$. Thus, $G(x) = xC(x)$ by Lemma 7.63, so that $G_{n+1} = c_n$. \square

Theorem 7.67. c_n counts the number of plane/ordered trees with $n + 1$ total nodes.

Proof. Consider the combinatorial class \mathcal{G} of trees where the size is given by the number of nodes. Any tree of size at least 1 can be represented as a root followed by any number of subtrees as children. Then $\mathcal{G} = \{\cdot\} \times \text{SEQ}(\mathcal{G})$. The generating function thus satisfies $G(x) = \frac{x}{1-G(x)}$. Solving gives $G(x) - G^2(x) = x$. The same procedure as in Corollary 7.66 now gives the result.

Instead of solving the functional equation, we could do this more easily using Lagrange Inversion. Let $\phi(z) = \frac{1}{1-z}$ so that $G(x) = x\phi(G(x))$. Then:

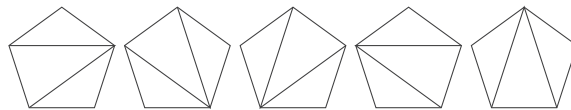
$$[x^n]G(x) = \frac{1}{n}[z^{n-1}]\phi(z)^n = \frac{1}{n}[z^{n-1}]\frac{1}{(1-z)^n} = \frac{1}{n}\binom{(n-1)+n-1}{n-1} = \frac{1}{n}\binom{2n-2}{n-1} = c_{n-1}$$

Alternatively, biject rooted trees to legal strings of parentheses. Given a string of parentheses, we convert it into a tree as follows:

- Create a root node.
- While the string is nonempty, remove its prefix. From this prefix, remove the outer parentheses and use the same process to create a tree for this shorter string. Attach it as a subtree to the root node.
- Once the string is empty, terminate the process of adding subtrees.

For the reverse process, we create a string of parentheses from a tree by representing each leaf node as $()$, and then grouping together subtrees with parentheses. The outermost parentheses corresponding to the root node are then removed. \square

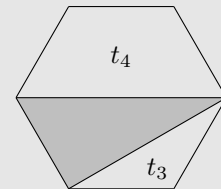
Theorem 7.68. c_n counts the number of non-overlapping triangulations of a convex $(n + 2)$ -gon.



5 Non-Overlapping Triangulations of a Pentagon

Proof.

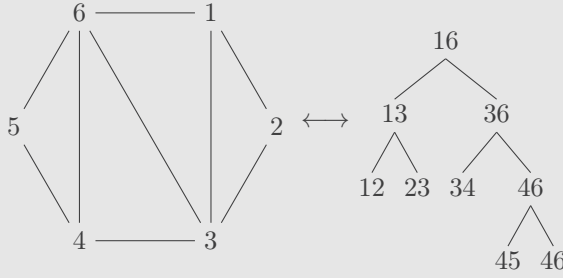
Let t_n count the number of triangulations of an n -gon. Pick an edge of the polygon. There are $n-2$ triangles it could belong to, since there are $n-2$ other vertices. In each case, it splits the polygon into two smaller ones with number of sides k and $n+3-k$ for $2 \leq k \leq n+1$. The number of ways to finish the triangulation is thus $t_{k-2}t_{n+1-k}$. Summing over all k , we have $\sum_{k=2}^{n+1} t_{k-2}t_{n+1-k} = \sum_{k=1}^n t_{k-1}t_{n-k}$. But this is the same recurrence as the Catalan Numbers in Theorem 7.61. Since $t_2 = 1 = c_0$ and $t_3 = 1 = c_1$, we have that $t_{n+2} = c_n$.



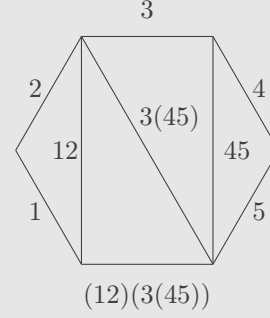
Another proof bijects triangulations to full binary trees whose vertices are the $n + 2$ exterior edges and the $n - 1$ internal edges (resulting in $n + 1$ leaf nodes). Fix an edge to be the root of the tree, and identify the triangle it belongs to. Let the other two edges be the children of the root. Now continue recursively. Any exterior edge encountered is a leaf node. For any internal edge, it borders

two triangles: the one used to reach that edge, and a triangle not yet counted. Let the children of the edge be the other two edges of the uncounted triangle.

A final proof bijects triangulations to parenthesizations. Label all except one of the edges, so that there are $n + 1$ labeled edges. Now, iteratively label the interior edges by combining labels, using parentheses to specify the order in which they are combined. The final parenthesization is the label of the last exterior edge.



Method 2



Method 3

□

Theorem 7.69. Consider a string of parentheses that satisfies the criteria to be legal, except that not all left parentheses need to be closed. Let q_n count the number of possible strings of length n . Then $q_n = \binom{n}{\lfloor n/2 \rfloor}$, the Central Binomial Coefficients.

Proof. If no prefix exists, then removing the leading left parenthesis is a bijection to q_{n-1} . Otherwise, condition on the index, so that there are $\sum_{k=1}^{\lfloor n/2 \rfloor} c_{k-1} q_{n-2k}$ options. Together, we have $q_n = q_{n-1} + \sum_{k=1}^{\lfloor n/2 \rfloor} c_{k-1} q_{n-2k}$ for $n > 0$ and $q_0 = 1$. The corresponding generating function is:

$$\begin{aligned} Q(x) &= \sum_{n=0}^{\infty} q_n x^n = q_0 + \sum_{n=1}^{\infty} \left[q_{n-1} + \sum_{k=1}^{\lfloor n/2 \rfloor} c_{k-1} q_{n-2k} \right] x^n = 1 + \sum_{n=1}^{\infty} q_{n-1} x^n + \sum_{n=1}^{\infty} \sum_{k=1}^{\lfloor n/2 \rfloor} c_{k-1} q_{n-2k} x^n \\ &= 1 + xQ(x) + \sum_{k=1}^{\infty} \sum_{n=2k}^{\infty} c_{k-1} q_{n-2k} x^n = 1 + xQ(x) + \sum_{k=1}^{\infty} c_{k-1} x^{2k} \sum_{n=2k}^{\infty} q_{n-2k} x^{n-2k} \\ &= 1 + xQ(x) + Q(x)x^2 \sum_{k=0}^{\infty} c_k x^{2k} = 1 + xQ(x) + Q(x)x^2 C(x^2) \end{aligned}$$

Solving:

$$\begin{aligned} Q(x) &= -\frac{1}{x^2 C(x^2) + x - 1} = -\frac{1}{\frac{1-\sqrt{1-4x^2}}{2} + x - 1} = \frac{2}{\sqrt{1-4x^2} - 2x + 1} \\ &= \frac{2(\sqrt{1-4x^2} + 2x - 1)}{(\sqrt{1-4x^2} - 2x + 1)(\sqrt{1-4x^2} + 2x - 1)} = \frac{2\sqrt{1-4x^2} + 4x - 2}{1 - 4x^2 - (2x - 1)^2} = \frac{2\sqrt{1-4x^2} + 4x - 2}{-8x^2 + 4x} \\ &= \frac{\sqrt{1-4x^2} + 2x - 1}{2x(1-2x)} \end{aligned}$$

Thus, using Theorem 2.32:

$$\begin{aligned} q_n &= [x^n]Q(x) = \frac{1}{2}[x^n] \frac{\sqrt{1-4x^2} + 2x - 1}{x(1-2x)} = \frac{1}{2}[x^{n+1}] \left(\frac{\sqrt{1-4x^2}}{1-2x} - 1 \right) = \frac{1}{2}[x^{n+1}] \frac{\sqrt{1-4x^2}}{1-2x} \\ &= 2^n [x^{n+1}] \frac{\sqrt{1-x^2}}{1-x} = 2^n \sum_{k=0}^{n+1} [x^k] (1-x^2)^{1/2} = 2^n \sum_{k=0}^{n+1} [x^k] \sum_{k=0}^{\infty} \binom{1/2}{k} (-1)^k x^{2k} \end{aligned}$$

$$\begin{aligned}
&= 2^n \sum_{k=0}^{n+1} [x^k] \sum_{k=0}^{\infty} [k \text{ even}] \binom{1/2}{k/2} (-1)^{k/2} x^k = 2^n \sum_{\substack{k=0 \\ k \text{ even}}}^{n+1} \binom{1/2}{k/2} (-1)^{k/2} \\
&= 2^n \sum_{k=0}^{\lfloor \frac{n+1}{2} \rfloor} \binom{1/2}{k} (-1)^k = 2^n \binom{-1/2}{\lfloor (n+1)/2 \rfloor} (-1)^{\lfloor (n+1)/2 \rfloor} = 2^{n-2\lfloor \frac{n+1}{2} \rfloor} \binom{2\lfloor \frac{n+1}{2} \rfloor}{\lfloor \frac{n+1}{2} \rfloor}
\end{aligned}$$

When n is even, we obtain $\binom{n}{n/2}$. When n is odd, we obtain $\frac{1}{2} \binom{n+1}{(n+1)/2} = \frac{1}{2} \frac{n+1}{(n+1)/2} \binom{n}{(n-1)/2} = \binom{n}{(n-1)/2}$. In either case, $q_n = \binom{n}{\lfloor n/2 \rfloor}$. \square

7.8 Counting Using Products and Compositions

Example 7.70. Suppose a semester of n days is divided into two parts (of arbitrary length), and there is 1 holiday during the first part and 2 holidays during the second part. Let s_n be the number of ways this can be accomplished. If the first part has k days, then there are $a_k = k$ ways to select the holiday. If the second part has k days, then there are $b_k = \binom{k}{2}$ ways to select the two holidays. By Lemma 7.19:

$$S(x) = A(x)B(x) = \frac{x}{(1-x)^2} \cdot \frac{x^2}{(1-x)^3} = \frac{x^3}{(1-x)^5}$$

Thus, we compute:

$$s_n = [x^n] \frac{x^3}{(1-x)^5} = [x^n] \frac{1}{x} \frac{x^4}{(1-x)^5} = [x^{n+1}] \frac{x^4}{(1-x)^5} = \binom{n+1}{4}$$

Example 7.71. Consider dividing a group of people into subgroups A , B , and C , and ask each subgroup to form a line. We also require that A have an odd number of people, and that B have an even number of people. We count the number of ways to do this. The exponential generating function representing the number of ways to arrange n people in group A is $A(x) = \sum_{n=0}^{\infty} \frac{1-(-1)^n}{2} n! \frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{1-(-1)^n}{2} x^n = \frac{1}{2} [\sum_{n=0}^{\infty} x^n - \sum_{n=0}^{\infty} (-x)^n] = \frac{1}{2} \left[\frac{1}{1-x} - \frac{1}{1+x} \right]$. The exponential generating function representing the number of ways to arrange n people in group B is $B(x) = \sum_{n=0}^{\infty} \frac{1+(-1)^n}{2} n! \frac{x^n}{n!} = \frac{1}{2} \left[\frac{1}{1-x} + \frac{1}{1+x} \right]$. The remaining generating function is $C(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. The total number of ways is thus encoded in the exponential generating function:

$$\begin{aligned}
A(x)B(x)C(x) &= \frac{1}{2} \left[\frac{1}{1-x} - \frac{1}{1+x} \right] \frac{1}{2} \left[\frac{1}{1-x} + \frac{1}{1+x} \right] \frac{1}{1-x} = \frac{1}{4} \left[-\frac{2x}{x^2-1} \right] \left[-\frac{2}{x^2-1} \right] \frac{1}{1-x} \\
&= \frac{-x}{(x+1)^2(x-1)^3} = \frac{-1/16}{x+1} + \frac{-1/8}{(x+1)^2} + \frac{1/16}{x-1} + \frac{-1/4}{(x-1)^3}
\end{aligned}$$

Then, we have:

$$\begin{aligned}
[x^n]A(x)B(x)C(x) &= -\frac{1}{16}[x^n] \frac{1}{x+1} - \frac{1}{8}[x^n] \frac{1}{(x+1)^2} + \frac{1}{16}[x^n] \frac{1}{x-1} - \frac{1}{4}[x^n] \frac{1}{(x-1)^3} \\
&= -\frac{1}{16}(-1)^{n+1} - \frac{1}{8} \left((-1)^{n+2} \binom{n+1}{1} \right) + \frac{1}{16}(-1)^{n+1} - \frac{1}{4} \left((-1)^{n+3} \binom{n+2}{2} \right) \\
&= \frac{1}{16}(-1)^{n+1} + \frac{1}{8}(-1)^{n+1}(n+1) - \frac{1}{16} + \frac{1}{8}(n+2)(n+1) \\
&= \frac{1}{16}[(-1)^{n+1} - 1] + \frac{1}{8}(n+1)[n+2 + (-1)^{n+1}]
\end{aligned}$$

Since this is an exponential generating function, the total number of ways is:

$$n! \left[\frac{1}{16}[(-1)^{n+1} - 1] + \frac{1}{8}(n+1)[n+2 + (-1)^{n+1}] \right] = \begin{cases} \frac{n! \cdot n(n+2)}{8} & \text{if } n \text{ is even} \\ \frac{(n+1)! \cdot (n+3)}{8} & \text{if } n \text{ is odd} \end{cases}$$

Example 7.72. Consider counting the number of ways to make change for a given amount n . The number of ways with solely pennies is $[x^n] \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, with solely nickels is $[x^n] \sum_{k=0}^{\infty} x^{5k} = \frac{1}{1-x^5}$,

and so on. Thus, the product $\frac{1}{1-x} \frac{1}{1-x^5} \frac{1}{1-x^{10}} \cdots$ counts the number of ways to make change for n with any combination of coins.

This can also work with a finite bank of coins by truncating each power series (instead of using an infinite series).

7.9 Snake Oil Method

Concept 7.73. The Snake Oil Method uses generating functions to compute sums containing a free variable. First, a free variable n is identified and the sum is named g_n . The generating function $G(x)$ is written, which now contains two summations. Interchange the order of summation, then try to identify the coefficients in this new form.

↳ The name points to the fact that this method may or may not work, and when it does work, it may seem miraculous or not easily explainable

Example 7.74. Consider the summation $a_n = \sum_{k=0}^{\infty} \binom{n+k}{mk} \alpha^{n-k}$ for some $n, m \in \mathbb{N}$ and $\alpha \in \mathbb{R}$. Its generating function with respect to n is $A(x) = \frac{\alpha(1-\alpha x)^{m-1}}{(1-\alpha x)^m - (\alpha x)^{m-1}}$, which can be used to find the closed form of a_n depending on the value of m and α . For example, when $m = 2$ and $\alpha = 2$, we obtain $a_n = \frac{1}{3}(1 + 2 \cdot 4^n)$.

Proof. Using the free variable n , we have:

$$\begin{aligned} A(x) &= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \binom{n+k}{mk} \alpha^{n-k} x^n = \sum_{k=0}^{\infty} \alpha^{-k} \sum_{n=0}^{\infty} \binom{n+k}{mk} (\alpha x)^n \\ &= \sum_{k=0}^{\infty} \alpha^{-k} (\alpha x)^{-k} \sum_{n=0}^{\infty} \binom{n+k}{mk} (\alpha x)^{n+k} = \sum_{k=0}^{\infty} \alpha^{-k} (\alpha x)^{-k} \sum_{n=k}^{\infty} \binom{n}{mk} (\alpha x)^n \\ &\stackrel{2.45}{=} \sum_{k=0}^{\infty} \alpha^{-k} (\alpha x)^{-k} \frac{(\alpha x)^{mk}}{(1-\alpha x)^{mk+1}} = \frac{1}{1-\alpha x} \sum_{k=0}^{\infty} \left(\frac{(\alpha x)^{m-1}}{\alpha(1-\alpha x)^m} \right)^k = \frac{1}{1-\alpha x} \frac{1}{1 - \frac{(\alpha x)^{m-1}}{\alpha(1-\alpha x)^m}} \\ &= \frac{1}{1-\alpha x} \frac{\alpha(1-\alpha x)^m}{\alpha(1-\alpha x)^m - (\alpha x)^{m-1}} = \frac{\alpha(1-\alpha x)^{m-1}}{\alpha(1-\alpha x)^m - (\alpha x)^{m-1}} \end{aligned}$$

Consider the case where $m = 2$ and $\alpha = 2$. Then $A(x) = \frac{1-2x}{(1-2x)^2-x} = \frac{1-2x}{1-5x+4x^2}$. Then:

$$a_n = [x^n] \frac{1-2x}{1-5x+4x^2} = [x^n] \left(\frac{1/3}{1-x} + \frac{2/3}{1-4x} \right) = \frac{1}{3}(1 + 2 \cdot 4^n) \quad \square$$

Example 7.75. Consider the summation $\sum_{k=0}^{\infty} \binom{n}{k} \binom{2n}{n-k}$. This is equal to $\binom{3n}{n}$ by Vandermonde's Identity, but we can also use the Snake Oil Method. Using n as the free variable results in the summation $\sum_{k=0}^{\infty} \binom{2n}{n-k} \frac{x^k}{(1-x)^{k+1}}$, which is not particularly nice to evaluate. Instead, if we add a free variable m so that $a_n = \sum_{k=0}^{\infty} \binom{n}{k} \binom{m}{n-k}$, then we can proceed with the generating function with respect to m :

$$\begin{aligned} A(x) &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \binom{n}{k} \binom{m}{n-k} x^m = \sum_{k=0}^{\infty} \binom{n}{k} \sum_{m=0}^{\infty} \binom{m}{n-k} x^m = \sum_{k=0}^{\infty} \binom{n}{k} \frac{x^{n-k}}{(1-x)^{n-k+1}} \\ &= \frac{x^n}{(1-x)^{n+1}} \sum_{k=0}^{\infty} \binom{n}{k} \left(\frac{1-x}{x} \right)^k = \frac{x^n}{(1-x)^{n+1}} \left(1 + \frac{1-x}{x} \right)^n = \frac{1}{(1-x)^{n+1}} \end{aligned}$$

Thus, $a_n = [x^{2n}] \frac{1}{(1-x)^{n+1}} = \binom{3n}{n}$.

Example 7.76. $\sum_{k=0}^{\infty} \binom{2n+1}{2p+2k+1} \binom{p+k}{k} = 4^{n-p} \binom{2n-p}{p}$

Proof. Let $s = 2n + 1$, and consider the generating function with respect to s :

$$\begin{aligned}
 S(x) &= \sum_{s=0}^{\infty} \sum_{k=0}^{\infty} \binom{s}{2p+2k+1} \binom{p+k}{k} x^s = \sum_{k=0}^{\infty} \binom{p+k}{p} \sum_{s=0}^{\infty} \binom{s}{2p+2k+1} x^s \\
 &= \sum_{k=0}^{\infty} \binom{p+k}{p} \frac{x^{2p+2k+1}}{(1-x)^{2p+2k+2}} = \sum_{k=p}^{\infty} \binom{k}{p} \frac{x^{2k+1}}{(1-x)^{2k+2}} = \frac{x}{(1-x)^2} \sum_{k=p}^{\infty} \binom{k}{p} \left(\frac{x^2}{(1-x)^2} \right)^k \\
 &= \frac{x}{(1-x)^2} \frac{\left(\frac{x^2}{(1-x)^2} \right)^p}{\left(1 - \frac{x^2}{(1-x)^2} \right)^{p+1}} = \frac{x}{(1-x)^2} \left(\frac{x^2}{(1-x)^2} \right)^p \left(\frac{(1-x)^2}{1-2x} \right)^{p+1} = \frac{x^{2p+1}}{(1-2x)^{p+1}}
 \end{aligned}$$

Extracting the coefficients:

$$\begin{aligned}
 [x^s] \frac{x^{2p+1}}{(1-2x)^{p+1}} &= [x^{s-2p-1}] \frac{1}{(1-2x)^{p+1}} = (-2)^{-p-1} [x^{s-2p-1}] \frac{1}{(x-1/2)^{p+1}} \\
 &= (-2)^{-p-1} \left(\frac{1}{1/2} \right)^{(s-2p-1)+(p+1)} \binom{(s-2p-1)+(p+1)-1}{(p+1)-1} (-1)^{p+1} \\
 &= 2^{s-2p-1} \binom{s-p-1}{p} = 2^{2n-2p} \binom{2n-p}{p} = 4^{n-p} \binom{2n-p}{p} \quad \square
 \end{aligned}$$

Example 7.77. $\sum_{k=0}^{\infty} \binom{2n+1}{2k} \binom{m+k}{2n} = \binom{2m+1}{2n}$

Proof. Consider the generating function in m :

$$\begin{aligned}
 A(x) &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \binom{2n+1}{2k} \binom{m+k}{2n} x^m = \sum_{k=0}^{\infty} \binom{2n+1}{2k} \sum_{m=0}^{\infty} \binom{m+k}{2n} x^m \\
 &= \sum_{k=0}^{\infty} \binom{2n+1}{2k} x^{-k} \sum_{m=0}^{\infty} \binom{m+k}{2n} x^{m+k} = \sum_{k=0}^{\infty} \binom{2n+1}{2k} x^{-k} \sum_{m=k}^{\infty} \binom{m}{2n} x^m
 \end{aligned}$$

In order for a term in the original sum to be nonzero, we must have $k \leq n$, so that $k < 2k \leq 2n$. Thus, we may change the index of the above summation to be $m = 0$, since all terms introduced will have $\binom{m}{2n} = 0$. Then:

$$\begin{aligned}
 A(x) &= \sum_{k=0}^{\infty} \binom{2n+1}{2k} x^{-k} \sum_{m=0}^{\infty} \binom{m}{2n} x^m = \sum_{k=0}^{\infty} \binom{2n+1}{2k} x^{-k} \frac{x^{2n}}{(1-x)^{2n+1}} \\
 &= \sum_{k=0}^n \binom{2n+1}{2n+1-2k} \frac{x^{2n-k}}{(1-x)^{2n+1}} = \sum_{k=0}^n \binom{2n+1}{2(n-k)+1} \frac{x^{n+(n-k)}}{(1-x)^{2n+1}} \\
 &= \sum_{k=0}^n \binom{2n+1}{2k+1} \frac{x^{n+k}}{(1-x)^{2n+1}} = \frac{x^n}{(1-x)^{2n+1}} \sum_{k=0}^n \binom{2n+1}{2k+1} x^k
 \end{aligned}$$

Now consider the generating function in $y = \sqrt{x}$:

$$\begin{aligned}
 B(x) &= \frac{y^{2n}}{(1-y^2)^{2n+1}} \sum_{k=0}^n \binom{2n+1}{2k+1} y^{2k} = \frac{y^{2n-1}}{(1-y^2)^{2n+1}} \sum_{k=0}^n \binom{2n+1}{2k+1} y^{2k+1} \\
 &= \frac{y^{2n-1}}{(1-y)^{2n+1}(1+y)^{2n+1}} \frac{(1+y)^{2n+1} - (1-y)^{2n+1}}{2} = \frac{y^{2n-1}}{2(1-y)^{2n+1}} - \frac{y^{2n-1}}{2(1+y)^{2n+1}} \\
 &= \frac{1}{2y} \left[\frac{y^{2n}}{(1-y)^{2n+1}} - \frac{y^{2n}}{(1+y)^{2n+1}} \right]
 \end{aligned}$$

Thus:

$$\begin{aligned}
 [x^m]A(x) &= [y^{2m}]B(x) = [y^{2m}]\frac{1}{2y} \left(\frac{y^{2n}}{(1-y)^{2n+1}} - \frac{y^{2n}}{(1+y)^{2n+1}} \right) \\
 &= \frac{1}{2} \left([y^{2m+1}]\frac{y^{2n}}{(1-y)^{2n+1}} - [y^{2m+1}]\frac{y^{2n}}{(1+y)^{2n+1}} \right) \\
 &= \frac{1}{2} \left([y^{2m+1}]\frac{y^{2n}}{(1-y)^{2n+1}} - [(-y)^{2m+1}]\frac{(-y)^{2n}}{(1-y)^{2n+1}} \right) \\
 &= \frac{1}{2} \left([y^{2m+1}]\frac{y^{2n}}{(1-y)^{2n+1}} + [y^{2m+1}]\frac{y^{2n}}{(1-y)^{2n+1}} \right) = [y^{2m+1}]\frac{y^{2n}}{(1-y)^{2n+1}} = \binom{2m+1}{2n} \square
 \end{aligned}$$

8. Combinatorial Classes

Definition 8.1. A combinatorial class is a finite or countable set of objects with a size defined for each object in the set satisfying:

- The size of each object is a nonnegative integer
- The number of objects of a given size is finite

If \mathcal{A} is such a class, then the size of $a \in \mathcal{A}$ is denoted $|a|$.

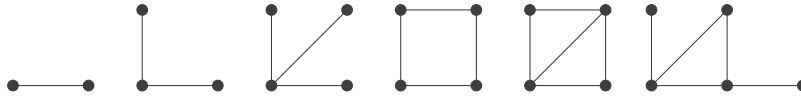
Definition 8.2. We can associate an ordinary generating function with a combinatorial class \mathcal{A} by:

$$\sum_{a \in \mathcal{A}} x^{|a|} = \sum_{n \geq 0} A_n x^n \text{ where } A_n \text{ is the subset of } \mathcal{A} \text{ containing objects of size } n$$

$\{A_n\}$ is called the counting sequence.

Example 8.3. Consider permutations on $[5]$, and let the size of a permutation $\pi \in S_5$ be given by its inversion number. Consider the combinatorial class $\mathcal{B} = \{21345, 21354, 12453, 14523, 41325, 51243\}$. The inversion numbers are $(1, 2, 4, 4, 5)$ respectively, so that the ogf is $B(x) = x^1 + 2x^2 + 2x^4 + x^5$ and the counting sequence is $\{0, 1, 2, 0, 2, 1\}$.

Example 8.4. Consider the following collection of finite graphs, and let the size of a graph be given by its order (number of vertices):



Then the ordinary generating function is $C(x) = x^2 + x^3 + 3x^4 + x^5$. We could also use a bivariate generating function, with the exponent on x representing the number of vertices and the exponent on y representing the number of edges. This gives $K(x, y) = x^2y + x^3y^2 + x^4y^3 + x^4y^4 + x^4y^5 + x^5y^5$.

Definition 8.5. Two classes \mathcal{A} and \mathcal{B} are called equivalent if the counting sequences are identical.

Definition 8.6. The power of combinatorial classes is that we construct more complicated structures from simpler classes iteratively, and using generating functions, obtain the sizes of these new classes. A construction is called admissible if the counting sequence $\{A_n\}$ of \mathcal{A} only depends on the counting sequence of B_n^1, B_n^2, \dots and not on any internal structure of any B_n^k , and also yields a valid combinatorial class satisfying the required properties.

- ↳ For example, $\mathcal{A} = \mathcal{B} \cup \mathcal{C}$ is not an admissible construction because it relies on knowledge of the elements of \mathcal{B} and \mathcal{C} , rather than just their counting sequences

Definition 8.7. Define two basic classes as follows:

- The neutral class $\mathcal{E} = \{\square\}$, where $|\square| = 0$. The corresponding generating function is $E(x) = 1$.
- The atomic class $\mathcal{Z} = \{\cdot\}$, where $|\cdot| = 1$. For example, $Z_a = \{a\}$ with $|a| = 1$. The corresponding generating function is $Z(x) = x$.

Definition 8.8. Define the following constructions:

- Cartesian Product: $\mathcal{A} = \mathcal{B} \times \mathcal{C} = \{(b, c) | b \in \mathcal{B}, c \in \mathcal{C}\}$ using the size $|a|_{\mathcal{A}} = |b|_{\mathcal{B}} + |c|_{\mathcal{C}}$ (i.e., add the component sizes). The generating function is $A(x) = B(x)C(x)$.
 - ↳ We allow for infinite Cartesian products as well, but only when any term of the resulting counting sequence depends on finitely many terms in the product. With this constraint, it is clear that the construction is admissible.
 - ↳ This Cartesian product ultimately represents a Cauchy product. It counts the number of ways to build two combinatorial structures whose sizes add to a particular value. In other words, we split $[n]$ into two intervals, build one structure on the left half, and build a second structure on the right half.

- **Combinatorial Sum:** $\mathcal{A} = \mathcal{B} + \mathcal{C} = (\{\triangle\} \times B) \cup (\{\diamond\} \times C)$ using the size $|(\triangle, a)|_{\mathcal{A}} = |a|_{\mathcal{B}}$ and $|(\diamond, a)|_{\mathcal{A}} = |a|_{\mathcal{C}}$. The generating function is $A(x) = B(x) + C(x)$.
 - ↳ This is effectively a union, but we ensure the two sets are disjoint by multiplying with arbitrary \triangle and \diamond .
- **Sequence Construction:** Assume $\square \notin \mathcal{B}$. Then $\text{SEQ}(\mathcal{B}) = \mathcal{E} + B + B \times B + B^3 + \dots$. The ordinary generating function for the $\mathcal{A} = \text{SEQ}(\mathcal{B})$ is $A(x) = 1 + B(x) + B^2(x) + \dots = \frac{1}{1-B(x)}$.
 - ↳ \mathcal{E} allows selection of the empty sequence and B^k allows for sequences of length k .
 - ↳ If \mathcal{B} contained \square , then there would be an infinite number of objects of any given size, generated by inserting arbitrary many \square elements.
 - ↳ We also define $\text{SEQ}_k(\mathcal{B}) = B^k$, which are sequences of fixed length. The ordinary generating function is $B(x)^k$.
 - ↳ $\text{SEQ}_{\leq k}(\mathcal{B}) = \sum_{j=0}^k B_j$ has the generating function $\frac{1-B(x)^{k+1}}{1-B(x)}$.
 - ↳ $\text{SEQ}_{\geq k}(\mathcal{B}) = \sum_{j=k}^{\infty} B_j$ has the generating function $\frac{B(x)^k}{1-B(x)}$.
 - ↳ For $\Omega \subset \mathbb{W}$, let $\text{SEQ}_{\Omega}(\mathcal{B})$ represent sequences whose lengths are in Ω . This has generating function $\sum_{k \in \Omega} B(x)^k$.
- **Cycle Construction:** Assume $\square \notin \mathcal{B}$. Define an equivalence relation S between sequences, where two sequences are equivalent if there exists a cyclic permutation (circular shift) mapping one to the other (ignoring the initial arbitrary element \triangle , \diamond , etc.). Then $\text{CYC}(\mathcal{B}) = (\text{SEQ}(\mathcal{B}) \setminus \{\square\})/S$. The ordinary generating function for $\mathcal{A} = \text{CYC}(\mathcal{B})$ is $A(x) = \sum_{k=1}^{\infty} \frac{\phi(k)}{k} \log\left(\frac{1}{1-B(x^k)}\right)$, where ϕ is the Euler totient function.
 - ↳ The empty sequence is excluded by convention.
 - ↳ We also define $\text{CYC}_k(\mathcal{B}) = (\text{SEQ}_k(\mathcal{B}) \setminus \{\square\})/S$, which are cycles of fixed length.
- **Multiset Construction:** Assume $\square \notin \mathcal{B}$. Define an equivalence relation Π between sequences, where two sequences are equivalent if there exists an arbitrary permutation mapping one to the other (ignoring the initial arbitrary element \triangle , \diamond , etc.). Then $\text{MSET}(\mathcal{B}) = \text{SEQ}(\mathcal{B})/\Pi$. The ordinary generating function for $\mathcal{A} = \text{MSET}(\mathcal{B})$ is $A(x) = \prod_{n=1}^{\infty} (1 - x^n)^{-B_n} = \exp \sum_{k=1}^{\infty} \frac{B(x^k)}{k}$.
 - ↳ We also define $\text{MSET}_k(\mathcal{B}) = \text{SEQ}_k(\mathcal{B})/\Pi$, which are multisets of fixed size.
- **Powerset Construction:** Assume $\square \notin \mathcal{B}$. $\text{PSET}(\mathcal{B}) \subset \text{MSET}(\mathcal{B})$ with the restriction that all elements are distinct (ignoring the initial arbitrary element \triangle , \diamond , etc.). Define this similarly for $\text{PSET}_k(\mathcal{B})$. The ordinary generating function for $\mathcal{A} = \text{PSET}(\mathcal{B})$ is $A(x) = \prod_{n=1}^{\infty} (1 + x^n)^{B_n} = \exp \sum_{k=1}^{\infty} \frac{(-1)^{k-1} B(x^k)}{k}$.
 - ↳ We also define $\text{PSET}_k(\mathcal{B})$ similarly, which are subsets of a fixed size.

Proof. For the Cartesian product, we have:

$$A(x) = \sum_{n=0}^{\infty} A_n x^n = \sum_{n=0}^{\infty} \sum_{k=0}^n B_k C_{n-k} x^n \stackrel{\text{III}}{=} B(x)C(x)$$

For an infinite Cartesian product, we note that any initial finite number of terms can be correctly captured with a finite product. Thus, the infinite product is a valid formal power series, and is the limit of the partial products.

For the sum construction, we have:

$$A(x) = \sum_{a \in \mathcal{A}} x^{|a|_{\mathcal{A}}} = \sum_{a \in (\{\triangle\} \times B) \cup (\{\diamond\} \times C)} x^{|a|_{\mathcal{A}}} = \sum_{b \in \mathcal{B}} x^{|b|_{\mathcal{B}}} + \sum_{c \in \mathcal{C}} x^{|c|_{\mathcal{C}}} = B(x) + C(x)$$

For the sequence construction, start by noting that $\frac{1}{1-B(x)}$ is a valid composition of power series since $[x^0]B(x) = 0$. Formally, we have an infinite combinatorial sum, but we can show similarly to

the above proof that the resulting generating function is the infinite sum of the generating functions. Hence, the result follows immediately.

For the cycle construction, consider any nonempty sequence in $\text{SEQ}(\mathcal{B})$, and note that it has a unique root which is primitive. For example, if $\mathcal{B} = \{\alpha, \beta\}$, the sequence $\alpha\beta\alpha\beta\beta\alpha\beta\alpha\beta\beta\alpha\beta\beta$ decomposes uniquely into $\alpha\beta\alpha\beta\beta|\alpha\beta\alpha\beta\beta|\alpha\beta\alpha\beta\beta$ where $\alpha\beta\alpha\beta\beta$ cannot be decomposed into a word repeated multiple times (and hence is called primitive). We use a bivariate generating function to describe $\text{SEQ}_{\geq 1}(\mathcal{B})$, with z tracking the size as normal, and u tracking the length of the sequence. The resulting generating function is thus $\frac{uB(z)}{1-uB(z)}$. Letting $PS(z, u)$ be the generating function representing primitive sequences over \mathcal{B} only, we have $\frac{uB(z)}{1-uB(z)} = \sum_{k=1}^{\infty} PS(z^k, u^k)$, where k conditions on the number of times the root is repeated.

Apply Möbius Inversion to obtain $PS(z, u) = \sum_{k=1}^{\infty} \mu(k) \frac{u^k B(z^k)}{1-u^k B(z^k)}$. More formally, consider any term of the power series, written as $[z^{in} u^{jn}]$ for coprime i, j . Let $a_n = [z^{in} u^{jn}] \frac{uB(z)}{1-uB(z)}$ and $b_n = [z^{in} u^{jn}] PS(z, u)$. Then the previous equality gives us the following, from which we can apply Möbius Inversion and obtain equality for all i, j (so that the bivariate power series are equal):

$$\begin{aligned} a_n &= \sum_{k=1}^{\infty} [z^{in} u^{jn}] PS(z^k, u^k) = \sum_{k|in \text{ and } k|jn} [z^{in} u^{jn}] PS(z^k, u^k) = \sum_{k|n} [z^{in} u^{jn}] PS(z^k, u^k) \\ &= \sum_{k|n} [z^{in/k} u^{jn/k}] PS(z, u) = \sum_{k|n} [z^{ik} u^{jk}] PS(z, u) = \sum_{k|n} b_k \end{aligned}$$

Next, notice that each primitive k -cycle has k distinct primitive word representations. So, the generating function $PC(z, u)$ for the number of primitive k -cycles has coefficients $[z^n u^k] PS(z, u)/k$, which can be obtained through the integral $\int_0^u PS(z, t)/t dt$. Substituting and using a u -substitution gives $PC(z, u) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \log \frac{1}{1-u^k B(z^k)}$.

To count the number of primitive cycles of given size without conditioning on their length, substitute $u \leftarrow 1$ to obtain $PC(z) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \log \frac{1}{1-B(z^k)}$. Every cycle has a unique root which is a primitive cycle, so that the generating function for $\text{CYC}(\mathcal{B})$ satisfies $C(z) = \sum_{k=1}^{\infty} PC(z^k)$, where again k conditions on the number of times the root is repeated. So, we obtain:

$$C(z) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \frac{\mu(j)}{j} \log \frac{1}{1-B(z^{jk})} = \sum_{n=1}^{\infty} \sum_{j|n} \frac{\mu(j)}{j} \log \frac{1}{1-B(z^n)} = \sum_{n=1}^{\infty} \frac{\phi(n)}{n} \log \frac{1}{1-B(z^n)}$$

We used Corollary 12.63 to simplify $\sum_{j|n} \frac{\mu(j)}{j}$ and obtain the result.

For the multiset construction, order \mathcal{B} , so that every equivalence class in $\mathcal{A} = \text{MSET}(\mathcal{B})$ can be represented by a sequence whose elements are in order. Thus, $\mathcal{A} = \prod_{b \in \mathcal{B}} \text{SEQ}(\{b\})$. Thus:

$$\begin{aligned} A(x) &= \prod_{b \in \mathcal{B}} \frac{1}{1-x^{|b|}} = \prod_{n=1}^{\infty} \frac{1}{(1-x^n)^{B_n}} = \exp \left(- \sum_{n=1}^{\infty} B_n \log(1-x^n) \right) = \exp \left(\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} B_n \frac{(x^n)^k}{k} \right) \\ &= \exp \left(\sum_{k=1}^{\infty} \frac{1}{k} \sum_{n=1}^{\infty} B_n (x^k)^n \right) = \exp \sum_{k=1}^{\infty} \frac{1}{k} B(x^k) \end{aligned}$$

For the powerset construction, note that $\mathcal{A} = \text{PSET}(\mathcal{B}) \cong \prod_{b \in \mathcal{B}} (\{\square\} + \{b\})$, where \square represents not including an element b in the final set. Thus:

$$\begin{aligned} A(x) &= \prod_{b \in \mathcal{B}} (1+x^{|b|}) = \prod_{n=1}^{\infty} (1+x^n)^{B_n} = \exp \sum_{n=1}^{\infty} B_n \log(1+x^n) = \exp \sum_{n=1}^{\infty} B_n \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(x^n)^k}{k} \\ &= \exp \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sum_{n=1}^{\infty} B_n (x^k)^n = \exp \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} B(x^k) \end{aligned}$$

In the multiset and powerset constructions, since $B_0 = 0$, the infinite product used above is admissible. \square

Theorem 8.9. Consider combinatorial classes \mathcal{B} and \mathcal{C} with corresponding generating functions B and C . Assume $b_0 = 0$. Then $C(B(x))$ counts the number of ways to partition $[n]$ into intervals, build a \mathcal{B} -structure on each interval, and build a \mathcal{C} -structure on the set of intervals.

Proof. Let \mathcal{A} be the combinatorial class describing the number of ways to accomplish this. By the definition of the product, \mathcal{B}^k is the combinatorial class describing the number of ways to divide $[n]$ into k intervals and build a \mathcal{B} -structure on each. The corresponding generating function is $B^k(x)$. The number of ways to build a \mathcal{C} -structure on the intervals is $c_k B^k(x)$. Sum over all k to obtain the result. \square

Example 8.10. Let $\mathcal{Z} = \{\cdot\}$ be an atomic class. Let $\mathcal{B} = \{\mathcal{Z}, \mathcal{Z} \times \mathcal{Z}\} = \{\cdot, (\cdot, \cdot)\}$, so that $B(x) = x + x^2$. Now we can form $\mathcal{A} = \text{SEQ}(\mathcal{B})$ and obtain $A(x) = \frac{1}{1-x-x^2}$. This is the generating function for the Fibonacci numbers, which can be understood by associating \cdot with monominoes and (\cdot, \cdot) with dominoes.

Example 8.11. More generally, consider $\mathcal{B}_{j,k} = \mathcal{Z}^j + \mathcal{Z}^k$ with generating function $x^j + x^k$. Then the generating function for $\mathcal{A} = \text{SEQ}(\mathcal{B}_{j,k})$ is $A(x) = \frac{1}{1-x^j-x^k}$. We could then consider $\mathcal{C} = \text{SEQ}(\mathcal{A} \setminus \mathcal{E})$, which represents sequences of the j -omino and k -omino tilings. We have $C(x) = \frac{1}{1 - (\frac{1}{1-x^j-x^k} - 1)} = \frac{1-x^j-x^k}{1-2x^j-2x^k}$.

Example 8.12. Let $\mathcal{Z} = \{\cdot\}$ be an atomic class. Let $\mathcal{N} = \text{SEQ}(\mathcal{Z}) \setminus \{\square\}$, which is the combinatorial class corresponding to the natural numbers with size function being the identity. The corresponding generating function is $N(x) = \frac{1}{1-x} - 1$. Now consider $\mathcal{A} = \text{SEQ}(\mathcal{N})$, so that $A(x) = \frac{1}{1 - (\frac{1}{1-x} - 1)} = \frac{1-x}{1-2x} = \frac{1}{2} \cdot \frac{1}{1-2x} + \frac{1}{2}$. For $n \geq 1$, we thus have $[x^n]A(x) = 2^{n-1}$, which is the number of compositions of n . For $n = 0$, there is also one composition, reflected by $[x^0]A(x) = 1$.

Example 8.13. Consider $b\text{SEQ}(a)$, which is words starting with a b and followed by an arbitrary number of copies of a . This is a Cartesian product of $\{b\}$ and $\text{SEQ}(a)$, so that the generating function is $x \cdot \frac{1}{1-x} = \frac{x}{1-x}$. Then $\text{SEQ}(b\text{SEQ}(a))$, which is all words over $\{a, b\}$ starting with b , has the generating function $\frac{1}{1 - \frac{x}{1-x}} = \frac{1-x}{1-2x}$. Notice that this is equivalent to $\text{SEQ}(\mathcal{N})$, since we can view the b letters as delimiting a sequence of natural numbers.

To count all words over $\{a, b\}$, we use $\text{SEQ}(a)\text{SEQ}(b\text{SEQ}(a))$, which has the generating function $\frac{1}{1-x} \cdot \frac{1-x}{1-2x} = \frac{1}{1-2x}$. This satisfies $[x^n] \frac{1}{1-2x} = 2^n$ as expected.

Example 8.14. Let $\mathcal{B} = \{\bullet, \bullet\bullet\bullet, \bullet\bullet\bullet\bullet\}$, which is shorthand for $\mathcal{B} = \{\cdot, (\cdot, (\cdot, \cdot)), (\cdot, \cdot, \cdot)\}$, or a combinatorial class with one element of size 1 and two elements of size 3. The five elements of size 4 in $\mathcal{A} = \text{SEQ}(\mathcal{B})$ are:



The generating function for \mathcal{A} is $\frac{1}{1-x-2x^3}$. If we wish to find the generating function of $\text{SEQ}(\bullet\bullet\bullet\mathcal{A})$, then we obtain:

$$\frac{1}{1-x^2 \cdot \frac{1}{1-x-2x^3}} = \frac{1-x-2x^3}{1-x-x^2-2x^3}$$

Example 8.15. Suppose $\mathcal{A} = \mathcal{Z}^k \times \text{SEQ}(\mathcal{A})$. Then the generating function satisfies $A(x) = \frac{x^k}{1-A(x)}$.

Rearranging and solving with the quadratic formula yields $A(x) = \frac{1-\sqrt{1-4x^k}}{2} = x^k C(x^k)$, where $C(x)$ is the generating function for the Catalan numbers given in Lemma 7.63.

Definition 8.16. For an element a of size $|a|$, the weight of a is $x^{|a|}$. More generally, we can consider weights that involve multiple variables, such as assigning a weight of $x^\alpha y^\beta$. Normally, the weight is a multiple of $x^{|a|}$, either a constant multiple or a multiple that is a monomial in other variables. Similar to the restriction on sizes for combinatorial classes, we must have that there are a finite number of elements of any given weight. Given this, we denote the resulting weighted combinatorial class by \mathcal{A}_{wt} .

A weighted generating function A_{wt} is formed by summing the weights of all elements, so that:

$$A_{\text{wt}}(x_1, \dots, x_m) = \sum_{a \in \mathcal{A}_{\text{wt}}} \text{wt}(a)$$

Proposition 8.17. If B_{wt} and C_{wt} are weighted combinatorial classes, then $B_{\text{wt}} + C_{\text{wt}}$ and $B_{\text{wt}} \times C_{\text{wt}}$ are admissible, and $\text{SEQ}(B_{\text{wt}})$ is as well assuming that $\square \notin B_{\text{wt}}$.

Example 8.18. Let \mathcal{S} be the collection of ordered pairs $([n], A)$ where $A \subset [n]$ and $n > 0$. Let the size of $([n], A)$ be n . We have $\mathcal{S} \cong \text{SEQ}(\{0, 1\})$ with $|0| = |1| = 1$. Indeed, the length of the sequence corresponds to n , with 0 representing excluded elements and 1 representing included elements. The ordinary generating function is $S(x) = \frac{1}{1-2x}$.

We now form a weighted generating function where the weight of $([n], A)$ is $x^n y^{|A|}$. $\text{wt}(0) = x$ and $\text{wt}(1) = xy$. Therefore, the weighted generating function becomes:

$$S_{\text{wt}}(x, y) = \frac{1}{1 - (x + xy)} = \sum_{n=0}^{\infty} (x + xy)^n = \sum_{n=0}^{\infty} x^n (1 + y)^n = \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} y^k x^n$$

Example 8.19. We can use the above weighted generating function to derive Pascal's identity (a recurrence for the binomial coefficient). We have $1 = (1 - x - xy) \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} y^k x^n$. Apply $[x^n y^k]$ with $(n, k) \neq (0, 0)$ to each side to obtain:

$$\begin{aligned} 0 &= [x^n y^k] \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} x^n y^k - [x^n y^k] x \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} x^n y^k - [x^n y^k] xy \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} x^n y^k \\ &= [x^n y^k] \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} x^n y^k - [x^{n-1} y^k] \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} x^n y^k - [x^{n-1} y^{k-1}] \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} x^n y^k \\ &= \binom{n}{k} - \binom{n-1}{k} - \binom{n-1}{k-1} \end{aligned}$$

Rearranging then yields the result.

Example 8.20. We can also use the weighted generating function to show the Binomial Theorem, namely $(1 + y)^n = \sum_{k=0}^n \binom{n}{k} y^k$. Rewrite $\frac{1}{1-(x+xy)} = \frac{1}{1-x(1+y)} = \sum_{n=0}^{\infty} (1+y)^n x^n$. Applying $[x^n]$ yields $(1+y)^n$, and applying $[x^n]$ to the equivalent quantity $\sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} x^n y^k$ yields $\sum_{k=0}^n \binom{n}{k} y^k$ as desired.

9. Labeled Structures

Definition 9.1. A labeled structure is a function \mathcal{S} which maps each finite set L to a finite multiset $\mathcal{S}(L)$ such that if $|L| = |M|$, then $|\mathcal{S}(L)| = |\mathcal{S}(M)|$. The set L is called the labels and $\mathcal{S}(L)$ is the structure on L .

Suppose that $|L| = n$, and let $s_n = |\mathcal{S}(L)|$. The corresponding exponential generating function is $F_{\mathcal{S}}(x) = F_{\mathcal{S}(\cdot)}(x) = \sum_{n=0}^{\infty} s_n \frac{x^n}{n!}$.

↳ From another point of view, we can view labeled structures as an endowment of additional information onto combinatorial classes. Consider a combinatorial class \mathcal{A} . However, for each $a \in \mathcal{A}$, we associate to it $|a|$ labels. The number of ways to label the structure a must only depend on $|a|$, and we denote it $s_{|a|}$.

Concept 9.2. The reason for making $\mathcal{S}(L)$ a multiset rather than a set arises due to complications arising with products of labeled structures. See Example 9.20 for an example. When $\mathcal{S}(L) \cap \mathcal{S}(M) = \emptyset$ for distinct $L \neq M$, then we can let $\mathcal{S}(L)$ simply be a set rather than a multiset. If we desire to only work with sets rather than multisets, we can modify the definition of problematic structures that do not satisfy $\mathcal{S}(L) \cap \mathcal{S}(M) = \emptyset$ by endowing each $S \in \mathcal{S}(L)$ with knowledge of its original label set L . In other words, we define the equivalent combinatorial structure $\tilde{\mathcal{S}}(L) = \{(S, L) \mid S \in \mathcal{S}(L)\}$.

Definition 9.3. \mathcal{S} and \mathcal{T} are equivalent, denotes $\mathcal{S} \equiv \mathcal{T}$, if $|\mathcal{S}(L)| = |\mathcal{T}(L)|$ for all finite L .

Definition 9.4. Let \mathcal{S} and \mathcal{T} be labeled structures. If $\mathcal{S}(L)$ and $\mathcal{T}(L)$ are disjoint for any finite L , then \mathcal{S} and \mathcal{T} are called disjoint.

Definition 9.5. For disjoint \mathcal{S} and \mathcal{T} , define their sum as $(\mathcal{S} \cup \mathcal{T})(L) = \mathcal{S}(L) \cup \mathcal{T}(L)$. Since all unions are disjoint, this is a valid labeled structure.

Definition 9.6. Let \mathcal{S}, \mathcal{T} be labeled structures. Then their product is:

$$(\mathcal{S} \times \mathcal{T})(L) = \{(S, T) \mid S \in \mathcal{S}(L_1), T \in \mathcal{T}(L_2) \text{ where } (L_1, L_2) \text{ is a weak partition of } L\}$$

$$= \bigcup_{k=0}^{\infty} \left\{ (S, T) \mid S \in \mathcal{S}(L_1), T \in \mathcal{T}(L \setminus L_1), L_1 \in \binom{L}{k} \right\}$$

Note that the above union is finite for any fixed L . L_1 or L_2 may be empty. We also denote the weak partition by $L_1/L_2 \vdash L$. This counts the number of ways to divide the labels in L between the two structures \mathcal{S} and \mathcal{T} , and then form valid labelings/structures on each structure.

↳ It may be possible to obtain the same tuple with different values of L_1 and L_2 , which is why we must consider the above as a multiset.

Proposition 9.7. Let \mathcal{S} and \mathcal{T} be labeled structures with exponential generating functions $F_{\mathcal{S}}$ and $F_{\mathcal{T}}$. For $|L| = n$, $s_n = |\mathcal{S}(L)|$, and $t_n = |\mathcal{T}(L)|$, we have:

- (a) $|(\mathcal{S} \cup \mathcal{T})(L)| = s_n + t_n$ assuming that \mathcal{S} and \mathcal{T} are disjoint. $F_{\mathcal{S} \cup \mathcal{T}}(x) = F_{\mathcal{S}}(x) + F_{\mathcal{T}}(x)$.
- (b) $|(\mathcal{S} \times \mathcal{T})(L)| = \sum_{k=0}^n \binom{n}{k} s_k t_{n-k}$ and $F_{\mathcal{S} \times \mathcal{T}}(x) = F_{\mathcal{S}}(x) F_{\mathcal{T}}(x)$.

Proof. $|(\mathcal{S} \cup \mathcal{T})(L)| = s_n + t_n$ holds since \mathcal{S} and \mathcal{T} are disjoint. The formula for the exponential generating functions thus follows immediately.

$|(\mathcal{S} \times \mathcal{T})(L)| = \sum_{k=0}^n \binom{n}{k} s_k t_{n-k}$ follows by conditioning on $k = |L_1|$, picking the values of L_1 $\binom{n}{k}$ ways, and then forming the \mathcal{S} -structure s_k ways and the \mathcal{T} -structure t_{n-k} ways. Now apply Wilf III to obtain the corresponding formula for the exponential generating functions. \square

Definition 9.8. Define the partition structure of \mathcal{S} as:

$$\Pi(\mathcal{S})(L) = \{\{S_1, S_2, \dots\} \mid L_1/L_2/\dots \vdash L \text{ and } S_j \in \mathcal{S}(L_j)\}$$

Under this structure, we partition L in all possible ways, and then build an \mathcal{S} -structure on each block in all possible ways. Note that we must have $\mathcal{S}(\emptyset) = \emptyset$, so that $|\mathcal{S}(\emptyset)| = 0$. Otherwise, the partition structure would map finite sets to infinite sets (since partitions could be empty).

↳ This is akin to the definition of a product, but order does not matter (and the number of factors is arbitrary)

Theorem 9.9. Suppose the exponential generating function of a labeled structure \mathcal{S} is $F_{\mathcal{S}}(x) = \sum_{n=1}^{\infty} s_n \frac{x^n}{n!}$. Note that $s_0 = 0$. Then:

$$F_{\Pi(\mathcal{S})}(x) = e^{F_{\mathcal{S}}(x)} = \sum_{n=0}^{\infty} \frac{F_{\mathcal{S}}(x)^n}{n!}$$

Proof. Condition on the number of blocks in the partition:

$$\Pi(\mathcal{S})(L) = \bigcup_{n=0}^{\infty} \{ \{S_1, \dots, S_n\} \mid L_1 / \dots / L_n \vdash L \text{ and } S_j \in \mathcal{S}(L_j) \}$$

Since $s_0 = 0$, the above union is finite for all L . Fix any n :

$$\begin{aligned} |\{ \{S_1, \dots, S_n\} \mid L_1 / \dots / L_n \vdash L \text{ and } S_j \in \mathcal{S}(L_j) \}| &= \frac{1}{n!} |\{(S_1, \dots, S_n) \mid L_1 / \dots / L_n \vdash L \text{ and } S_j \in \mathcal{S}(L_j)\}| \\ &= \frac{1}{n!} |\mathcal{S}^n(L)| \end{aligned}$$

Hence, the exponential generating function for the single term is $F_{\mathcal{S}}(x)^n/n!$. Sum over all n to obtain the result. \square

Theorem 9.10. More generally, suppose we have a labeled structure \mathcal{S} and a labeled structure \mathcal{T} with exponential generating functions $F_{\mathcal{S}}$ and $F_{\mathcal{T}}$. Assume $s_0 = 0$ and $t_0 = 1$. Then $F_{\mathcal{S}}(F_{\mathcal{T}}(x))$ counts the number of ways to partition L into any number of blocks, build a \mathcal{T} -structure within each block, and build a \mathcal{S} -structure on the set of blocks.

↳ This immediately implies the previous result, where the structure we build on each block is merely itself (labeled structure E , with exponential generating function e^x)

Proof. Let $F(x) = F_{\mathcal{S}}(F_{\mathcal{T}}(x))$ be the exponential generating function for $\{f_n\}$, representing the labeled structure:

$$\mathcal{F}(L) = \bigcup \{ \mathcal{T}(\{S_1, S_2, \dots\}) \mid L_1 / L_2 / \dots \vdash L \text{ and } S_j \in \mathcal{S}(L_j) \}$$

Condition on the number of blocks in the partition:

$$\mathcal{F}(L) = \bigcup_{n=0}^{\infty} \bigcup \{ \mathcal{T}(\{S_1, \dots, S_n\}) \mid L_1 / \dots / L_n \vdash L \text{ and } S_j \in \mathcal{S}(L_j) \}$$

Fix any n :

$$\begin{aligned} & \left| \bigcup \{ \mathcal{T}(\{S_1, \dots, S_n\}) \mid L_1 / \dots / L_n \vdash L \text{ and } S_j \in \mathcal{S}(L_j) \} \right| \\ &= t_n |\{ \{S_1, \dots, S_n\} \mid L_1 / \dots / L_n \vdash L \text{ and } S_j \in \mathcal{S}(L_j) \}| \\ &= \frac{t_n}{n!} |\{(S_1, \dots, S_n) \mid L_1 / \dots / L_n \vdash L \text{ and } S_j \in \mathcal{S}(L_j)\}| = \frac{t_n}{n!} |\mathcal{S}^n(L)| \end{aligned}$$

Hence, the exponential generating function for the single term is $F_{\mathcal{S}}(x)^n \cdot t_n/n!$. Sum over all n to obtain the result. \square

Example 9.11. Consider $\mathcal{S}(L) = \binom{L}{k}$, where $k \in \mathbb{W}$. Then $s_n = \binom{n}{k}$, so that:

$$F_{(\cdot)_k}(x) = \sum_{n=0}^{\infty} \binom{n}{k} \frac{x^n}{n!} = \sum_{n=k}^{\infty} \binom{n}{k} \frac{x^n}{n!} = \sum_{n=k}^{\infty} \frac{1}{k!(n-k)!} x^n = \frac{x^k}{k!} \sum_{n=k}^{\infty} \frac{x^{n-k}}{(n-k)!} = \frac{x^k}{k!} e^x$$

Example 9.12. Consider the powerset structure \mathcal{P} . Then $s_n = 2^n$, so that $F_{\mathcal{P}} = \sum_{n=0}^{\infty} 2^n \frac{x^n}{n!} = e^{2x}$. We can write the powerset structure as a sum of disjoint labeled structures:

$$\mathcal{S}(L) = \bigcup_{k=0}^{|L|} \binom{L}{k}$$

By the sum formula and Example 9.11, we similarly derive $F_{\mathcal{P}} = e^{2x}$.

We could also, in theory, coerce this into a form utilizing the partition structure. A powerset can be counted by representing each element of the set as a singleton, then choosing whether this singleton is included or not in two ways. Hence, $\mathcal{P} = \Pi(\mathcal{S})$, where \mathcal{S} is the labeled structure corresponding to assigning a singleton to included or not included. Then $F_{\mathcal{S}}(x) = 2x$, and $F_{\mathcal{P}}(x) = e^{2x}$.

Example 9.13. Consider $E(L) = \{L\}$, so that $s_n = 1$ for all n . Then:

$$F_E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

Example 9.14. Consider \overline{E} such that $\overline{E} : L \mapsto \{L\}$ when $L \neq \emptyset$ and $\overline{E}(\emptyset) = \emptyset$ otherwise. Then $s_n = I[n > 0]$. We thus have $F_{\overline{E}}(x) = e^x - 1$.

Example 9.15. Suppose $T([n])$ counts the number of labeled trees. The number of labeled trees is $s_n = n^{n-2}$ for $n \geq 1$ and $s_0 = 0$ by Lemma 6.22. Hence, $F_T(x) = \sum_{n=1}^{\infty} n^{n-2} \frac{x^n}{n!}$.

Example 9.16. Suppose $L \mapsto \left\{ \begin{smallmatrix} L \\ k \end{smallmatrix} \right\}$ gives set partitions on L of size k . Then $s_n = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$, so that $F_{\left\{ \cdot \atop k \right\}} = \sum_{n=0}^{\infty} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} \frac{x^n}{n!} = \frac{(e^x - 1)^k}{k!}$ by Proposition 5.20. Multiplying by $k!$ gives us the exponential generating function $(e^x - 1)^k$ for ordered set partitions of L into k parts.

Example 9.17. Suppose $L \mapsto \mathcal{S}(L)$ gives ordered set partitions of L . Then $F_{\mathcal{S}} = \frac{1}{2-e^x}$ by Lemma 5.27.

Example 9.18. Suppose $L \mapsto \mathcal{S}(L)$ gives permutations on L . Then $s_n = n!$, so that $F_{\mathcal{S}} = \sum_{n=0}^{\infty} n! \frac{x^n}{n!} = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$.

Example 9.19. Suppose $L \mapsto \left[\begin{smallmatrix} L \\ k \end{smallmatrix} \right]$, which are permutations with exactly k cycles. Then $s_n = \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$, so that $F_{\left[\cdot \atop k \right]} = \frac{1}{k!} \left(\log \frac{1}{1-x} \right)^k$ by Proposition 3.27.

In particular, when $k = 1$, we obtain $F_{\left[\cdot \atop 1 \right]} = \log \frac{1}{1-x}$.

Example 9.20. Consider the labeled structure $(\mathcal{P} \times \mathcal{P})(L)$. The exponential generating function is e^{4x} by Example 9.12 and the product formula, so that $|(\mathcal{P} \times \mathcal{P})(L)| = 4^{|L|}$. Take $L = [2]$, and we obtain $|(\mathcal{P} \times \mathcal{P})([2])| = 16$. Note, however, that the resulting $(\mathcal{P} \times \mathcal{P})([2])$ is a multiset with duplicate entries. (\emptyset, \emptyset) is generated by four different partitions of L : $\emptyset/12$, $1/2$, $2/1$, and $12/\emptyset$.

The cause of this issue is that from any $S \in \mathcal{S}(L)$, it is not possible to recover L . In other words, there exists $L \neq M$ such that $\mathcal{S}(L) = \mathcal{S}(M)$.

Example 9.21. The labeled structure of ordered set partitions into k parts is equivalent to $(\overline{E})^k$.

Proof. Let the former be \mathcal{S} . To show this equivalence, we must prove $|\mathcal{S}(L)| = |(\overline{E})^k(L)|$ for all finite L . In particular, we can show this by forming a bijection between $\mathcal{S}(L)$ and $(\overline{E})^k(L)$. An ordered set partition can be viewed as a k -tuple of disjoint nonempty sets whose union is L . $\overline{E}(L)$ is the labeled structure for nonempty sets containing all elements in L , so that $(\overline{E})^k(L)$ are tuples of nonempty sets containing disjoint elements, whose union is L . Hence, we have an obvious bijection.

We can use this, for example, to show $\left\{ \begin{smallmatrix} n \\ 2 \end{smallmatrix} \right\} = 2^{n-1} - 1$. The labeled structure of ordered set partitions into 2 parts is equivalent to $(\overline{E})^2$, so by the product rule has EGF $F(x) = (e^x - 1)^2 = e^{2x} - 2e^x + 1$. For $n \geq 1$, $[x^n]F(x) = 2^n - 2$. Divide by two to remove the order, yielding $\left\{ \begin{smallmatrix} n \\ 2 \end{smallmatrix} \right\} = 2^{n-1} - 1$. \square

Example 9.22. The labeled structure $\left[\begin{smallmatrix} L \\ k \end{smallmatrix} \right]$ with ordered cycles is equivalent to $\left[\begin{smallmatrix} L \\ 1 \end{smallmatrix} \right]^k$.

Proof. We can view the ordered cycles in a k -tuple, where each element is a nonempty cycle with distinct values in L , such that all elements in L are in at least one cycle. Each element of the tuple is represented by the labeled structure $\left[\begin{smallmatrix} L \\ 1 \end{smallmatrix} \right]$, so that we have an obvious bijection.

We can use this, for example, to show $\left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] = n! \sum_{k=1}^n \frac{1}{k}$. The labeled structure $\left[\begin{smallmatrix} L \\ 2 \end{smallmatrix} \right]$ with ordered cycles is equivalent to $\left[\begin{smallmatrix} L \\ 1 \end{smallmatrix} \right]^2$. $\left[\begin{smallmatrix} L \\ 1 \end{smallmatrix} \right]$ has the exponential generating function $\log \frac{1}{1-x}$ since there are $(n-1)!$ permutations of $[n]$ with one cycle, so that the exponential generating function for $\left[\begin{smallmatrix} L \\ 2 \end{smallmatrix} \right]$ is $\left(\log \frac{1}{1-x} \right)^2$ by the product rule. Now use Wilf III and partial fraction decomposition to obtain:

$$\begin{aligned} \left\{ \begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right\} &= \frac{1}{2} \cdot (n+1)! [x^{n+1}] \left(\log \frac{1}{1-x} \right)^2 = \frac{1}{2} \cdot (n+1)! \sum_{k=1}^n \frac{1}{k} \frac{1}{n+1-k} \\ &= \frac{1}{2} \cdot (n+1)! \frac{1}{n+1} \sum_{k=1}^n \left(\frac{1}{k} + \frac{1}{n+1-k} \right) = n! \sum_{k=1}^n \frac{1}{k} \quad \square \end{aligned}$$

Example 9.23. The labeled structure $\mathcal{P}(\cdot)$ is equivalent to $E \times E$.

Proof. We can biject $A \leftrightarrow (A, L \setminus A)$ for any $A \in \mathcal{P}(L)$. Thus, we have a bijection from $\mathcal{P}(L)$ to $(E \times E)(L)$. \square

Example 9.24. For an example, we enumerate $\Pi(\left\{ \begin{smallmatrix} \cdot \\ 2 \end{smallmatrix} \right\})([5])$:

$$\begin{array}{ccccc} \{1234/5\} & \{123/45\} & \{12/3, 4/5\} & \{14/3, 2/5\} & \{2/34, 1/5\} \\ \{1235/4\} & \{124/35\} & \{13/2, 4/5\} & \{1/34, 2/5\} & \{23/5, 1/4\} \\ \{1245/3\} & \{125/34\} & \{1/23, 4/5\} & \{13/5, 2/4\} & \{25/3, 1/4\} \\ \{1345/2\} & \{134/25\} & \{12/4, 3/5\} & \{15/3, 2/4\} & \{2/35, 1/4\} \\ \{1/2345\} & \{135/24\} & \{14/2, 3/5\} & \{1/35, 2/4\} & \{24/5, 1/3\} \\ & \{145/23\} & \{1/24, 3/5\} & \{14/5, 2/3\} & \{25/4, 1/3\} \\ & \{15/234\} & \{12/5, 3/4\} & \{15/4, 2/3\} & \{2/45, 1/3\} \\ & \{14/235\} & \{15/2, 3/4\} & \{1/45, 2/3\} & \{34/5, 1/2\} \\ & \{13/245\} & \{1/25, 3/4\} & \{23/4, 1/5\} & \{35/4, 1/2\} \\ & \{12/345\} & \{13/4, 2/5\} & \{24/3, 1/5\} & \{3/45, 1/2\} \end{array}$$

Example 9.25. Let \mathcal{S} be the labeled structure of permutations, and let $\mathcal{C}(L) = \left[\begin{smallmatrix} L \\ 1 \end{smallmatrix} \right]$ be the labeled structure of single-cycle permutations. Then writing each permutation as a cycle decomposition, we can see $\mathcal{S} = \Pi(\mathcal{C})$. In other words, we partition the elements in the permutation, then build a cycle on each block.

We have $F_{\mathcal{C}}(x) = \log \frac{1}{1-x}$, so that $F_{\mathcal{S}}(x) = \exp(F_{\mathcal{C}}(x)) = \frac{1}{1-x}$ as expected.

10. Inversion Formulas

10.1 Binomial Inversion

Lemma 10.1. $\sum_{k=0}^n \binom{n}{k} \binom{k}{m} (-1)^k = (-1)^n \delta_{nm}$

↳ This is sometimes referred to as an orthogonality condition, and is used in the development of the corresponding inversion formulas

Proof. By Theorem 2.24 and Theorem 2.14:

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} \binom{k}{m} (-1)^k &= \sum_{k=m}^n \binom{n}{m} \binom{n-m}{k-m} (-1)^k \\ &= \binom{n}{m} (-1)^m \sum_{k=0}^{n-m} \binom{n-m}{k} (-1)^k = \binom{n}{m} (-1)^m \delta_{nm} = (-1)^n \delta_{nm} \end{aligned}$$

Or, note that $\sum_{k=0}^n \binom{n}{k} \binom{k}{m}$ counts the number of ordered pairs (T, S) for $T \subset S \subset [n]$ and $|T| = m$. Let E, \mathcal{O} be the sets for which $|S|$ is even or odd, respectively. If $n \leq m$, the result is trivial. Otherwise, form the involution $(T, S) \mapsto (T, S \triangle \max([n] \setminus T))$, i.e., add/remove to S the largest element in $[n]$ that is not in T . This maps $E \leftrightarrow \mathcal{O}$, so we have the result. \square

Theorem 10.2 (Binomial Inversion). Let $\{a_n\}, \{b_n\}$ be sequences for $n \geq 0$. Then:

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

↳ The symmetric version is $a_n = \sum_{k=0}^n \binom{n}{k} b_k (-1)^k$ iff $b_n = \sum_{k=0}^n \binom{n}{k} a_k (-1)^k$

↳ Binomial inversion can sometimes offer an alternative for proofs other than Inclusion/Exclusion, such as with derangements

Proof. Suppose the RHS. Using Lemma 10.1:

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} b_k &= \sum_{k=0}^n \binom{n}{k} \sum_{j=0}^k \binom{k}{j} a_j (-1)^{k-j} = \sum_{j=0}^n a_j (-1)^{-j} \sum_{k=j}^n \binom{n}{k} \binom{k}{j} (-1)^k \\ &= \sum_{j=0}^n a_j (-1)^{-j} (-1)^n \delta_{nj} = a_n \end{aligned}$$

The converse and symmetric version follow by applying the above to modified versions of a_n and b_n .

Alternatively, we can use generating functions. Suppose RHS, and let A and B be the exponential generating functions of the sequences. Then using Proposition 7.45:

$$\begin{aligned} B(x) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^n \binom{n}{k} a_k (-1)^{n-k} = \sum_{k=0}^{\infty} (-1)^k a_k \sum_{n=k}^{\infty} \binom{n}{k} \frac{(-x)^n}{n!} \\ &= \sum_{k=0}^{\infty} (-1)^k a_k \frac{(-x)^k}{k!} e^{-x} = e^{-x} \sum_{k=0}^{\infty} a_k \frac{x^k}{k!} = e^{-x} A(x) \end{aligned}$$

Then by Wilf III, we have our result:

$$A(x) = e^x B(x) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \binom{n}{k} b_k \right) \frac{x^n}{n!}$$

The converse follows by applying the above to modified versions of a_n and b_n .

Instead of exponential generating function, we can use ordinary generating functions A and B . By Theorem 2.45, we have:

$$\begin{aligned} B(x) &= \sum_{n=0}^{\infty} x^n \sum_{k=0}^n \binom{n}{k} a_k (-1)^{n-k} = \sum_{k=0}^{\infty} (-1)^k a_k \sum_{n=k}^{\infty} \binom{n}{k} (-x)^n = \sum_{k=0}^{\infty} (-1)^k a_k \frac{(-x)^k}{(1+x)^{k+1}} \\ &= \frac{1}{1+x} \sum_{k=0}^{\infty} a_k \left(\frac{x}{1+x} \right)^k = \frac{1}{1+x} A\left(\frac{x}{1+x} \right) \end{aligned}$$

Take $x \leftarrow \frac{x}{1-x}$ to obtain our result:

$$\begin{aligned} A(x) &= \frac{1}{1-x} B\left(\frac{x}{1-x} \right) = \frac{1}{1-x} \sum_{k=0}^{\infty} b_k \left(\frac{x}{1-x} \right)^k = \sum_{k=0}^{\infty} b_k \frac{x^k}{(1-x)^{k+1}} \\ &= \sum_{k=0}^{\infty} b_k \sum_{n=0}^{\infty} \binom{n}{k} x^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \binom{n}{k} b_k \right) x^n \end{aligned}$$

One can also view this as a consequence of Möbius Inversion. Consider the poset P of all subsets of \mathbb{N} ordered by inclusion. Let $f(S) = b_{|S|}$ and $g(S) = a_{|S|}$. Then Möbius inversion gives $g(T) = \sum_{S \subset T} f(S)$ iff $f(T) = \sum_{S \subset T} g(S) \mu(S, T)$. Letting $n = |T|$, conditioning on $k = |S|$, and noting that $\mu(S, T) = (-1)^{|T|-|S|}$ by Lemma 12.36, we have $a_n = \sum_{k=0}^n \binom{n}{k} b_k$ iff $b_n = \sum_{k=0}^n \binom{n}{k} a_k (-1)^{n-k}$ as desired.

As one final alternative, we can use Lagrange Inversion (not a preferred proof, but perhaps informative). We can derive $B(x) = \frac{1}{1+x} A\left(\frac{x}{1+x}\right)$ as above. We let $y = \frac{x}{1+x}$, which can be rearranged as $x = y(1+x) = y\phi(x)$ for $\phi(x) = 1+x$. Using this, $A(y) = (1+x)B(x)$, so taking $W(x) = (1+x)B(x)$, we have:

$$\begin{aligned} a_n &= [x^n]A(x) = \frac{1}{n} [z^{n-1}](W'(z)\phi(z)^n) = \frac{1}{n} [z^{n-1}]((1+z)B'(z) + B(z))(1+z)^n \\ &= \frac{1}{n} \sum_{k=0}^{n-1} [z^k]((1+z)B'(z) + B(z)) [z^{n-k-1}](1+z)^n \\ &= \frac{1}{n} \sum_{k=0}^{n-1} ([z^k]B'(z) + [z^{k-1}]B'(z) + [z^k]B(z)) \binom{n}{n-k-1} \\ &= \frac{1}{n} \sum_{k=0}^{n-1} ((k+1)b_{k+1} + kb_k + b_k) \binom{n}{k+1} = \frac{1}{n} \sum_{k=0}^{n-1} (k+1)(b_{k+1} + b_k) \binom{n}{k+1} \\ &= \sum_{k=0}^{n-1} (b_{k+1} + b_k) \binom{n-1}{k} = \sum_{k=0}^{n-1} b_k \binom{n-1}{k} + \sum_{k=0}^{n-1} b_{k+1} \binom{n-1}{k} \\ &= \sum_{k=0}^{n-1} b_k \binom{n-1}{k} + \sum_{k=1}^n b_k \binom{n-1}{k-1} = \sum_{k=0}^n b_k \left(\binom{n-1}{k} + \binom{n-1}{k-1} \right) = \sum_{k=0}^n b_k \binom{n}{k} \quad \square \end{aligned}$$

Definition 10.3. The binomial transform of $\{a_n\}$ is $b_n = \sum_{k=0}^n \binom{n}{k} a_k (-1)^k$.

↳ This is an involution, corresponding to the symmetric form of binomial inversion

↳ We also have $b_n = (-1)^n (\delta^n a)_0$, the n th forward difference with alternating sign

↳ Sometimes the $(-1)^k$ is not included

Example 10.4. Using Theorem 5.10, the binomial transform of $a_n = n^2$ is:

$$b_n = \sum_{k=0}^n \binom{n}{k} k^2 (-1)^k = (-1)^n n! \left\{ \begin{matrix} 2 \\ n \end{matrix} \right\} = \{0, -1, 2, 0, 0, \dots\}$$

This can also be found using the forward difference interpretation:

$$\begin{array}{ccccccc} \boxed{0} & & 1 & & 4 & & 9 & & 16 \\ & \boxed{1} & & 3 & & 5 & & 7 & \\ & & \boxed{2} & & 2 & & 2 & & \\ & & & \boxed{0} & & 0 & & & \end{array}$$

Example 10.5. The binomial transform of $a_n = 1$ is:

$$b_n = \sum_{k=0}^n \binom{n}{k} a_k (-1)^k = \sum_{k=0}^n \binom{n}{k} (-1)^k = \delta_{n0}$$

Example 10.6. Using Extraction/Absorption, the binomial transform of $a_n = n$ is:

$$b_n = \sum_{k=0}^n \binom{n}{k} k (-1)^k = \sum_{k=1}^n n \binom{n-1}{k-1} (-1)^k = -n \sum_{k=0}^{n-1} \binom{n-1}{k} (-1)^k = -n \delta_{n1} = -\delta_{n1}$$

Example 10.7. The binomial transform of $a_n = a^n$ is:

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

Example 10.8. The binomial transform of $a_n = \delta_{n1}$ is:

$$b_n = \sum_{k=0}^n \binom{n}{k} \delta_{k1} (-1)^k = \binom{n}{1} (-1)^1 = -n$$

10.2 Stirling Inversion

Lemma 10.9. $\sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \left\{ \begin{matrix} k \\ m \end{matrix} \right\} (-1)^{n-k} = \delta_{nm}$ and $\sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^{n-k} = \delta_{nm}$

↳ We use these claims to prove Stirling inversion. However, if the inversion is proved by other means, we could also derive these quickly from the Stirling inversion formulas.

Proof. We use Proposition 3.24 and Theorem 5.12, which state that:

$$x^{\overline{n}} = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} x^k \quad \text{and} \quad x^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (x)_k$$

We combine these to obtain:

$$\begin{aligned} x^n &= \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (x)_k = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (-x)^{\overline{k}} (-1)^k = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \sum_{m=0}^k \begin{bmatrix} k \\ m \end{bmatrix} (-x)^m (-1)^k \\ &= \sum_{m=0}^n \left(\sum_{k=m}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^{k-m} \right) x^m \end{aligned}$$

Hence, we have our second result, after rewriting $(-1)^{k-m} = (-1)^{m-k}$ and then multiplying each side of the claim by $(-1)^{m-n}$. For the first claim:

$$\begin{aligned} (-x)^{\overline{n}} &= \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} (-x)^k = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} x^k (-1)^k = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \sum_{m=0}^k \left\{ \begin{matrix} k \\ m \end{matrix} \right\} (x)_m (-1)^k \\ &= \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \sum_{m=0}^k \left\{ \begin{matrix} k \\ m \end{matrix} \right\} (-x)^{\overline{m}} (-1)^{k-m} = \sum_{m=0}^n \left(\sum_{k=m}^n \begin{bmatrix} n \\ k \end{bmatrix} \left\{ \begin{matrix} k \\ m \end{matrix} \right\} (-1)^{k-m} \right) (-x)^{\overline{m}} \end{aligned}$$

Note that $(-x)^{\overline{n}}$ are linearly independent for different n (indeed, they have different degrees). Hence,

we have our first result, after rewriting $(-1)^{k-m} = (-1)^{m-k}$ and then multiplying each side of the claim by $(-1)^{m-n}$.

Alternatively, we can prove these combinatorially. If $n = m$, then the claims are both trivial. For the first claim, rewrite it as $\sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \left\{ \begin{matrix} k \\ m \end{matrix} \right\} (-1)^k = 0$. $\begin{bmatrix} n \\ k \end{bmatrix} \left\{ \begin{matrix} k \\ m \end{matrix} \right\}$ counts the number of ways to place the elements of $[n]$ into k cycles and then distribute the k cycles into m non-empty sets. Among the sets with multiple elements (which exist by PHP since $m < n$), identify the one with the smallest element. Use Corollary 3.29 to form a bijection between permutations in this set with an even and odd number of cycles. Extend this bijection so that it is the identity on all other sets. This bijection changes the parity of the number of cycles, hence yielding the claim.

For the second claim, rewrite it as $\sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^k = 0$. $\left\{ \begin{matrix} n \\ k \end{matrix} \right\} \begin{bmatrix} k \\ m \end{bmatrix}$ counts the number of ways to set partition $[n]$ into k parts and then create m cycles of the sets. Among the cycles with multiple sets (which exist by PHP since $m < n$), identify the one with the smallest element a . When a is in a set containing other elements, map this to a new structure by removing a from its set and adding a new element $\{a\}$ into the cycle directly after its original location. If a is in its own set, then merge it with the set before it in the cycle. This is an involution that reverses the parity of k , hence yielding the claim. \square

Theorem 10.10 (Stirling Inversion). Let $\{a_n\}, \{b_n\}$ be sequences for $n \geq 0$. Then:

$$a_n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} b_k \text{ for all } n \iff b_n = \sum_{k=0}^n \underbrace{(-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix}}_{s(n,k)} a_k \text{ for all } n$$

Proof. Suppose RHS. Using Lemma 10.9, we have:

$$\begin{aligned} \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} b_k &= \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \sum_{j=0}^k (-1)^{k-j} \begin{bmatrix} k \\ j \end{bmatrix} a_j = \sum_{j=0}^n (-1)^{n-j} a_j \sum_{k=j}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \begin{bmatrix} k \\ j \end{bmatrix} (-1)^{n-k} \\ &= \sum_{j=0}^n (-1)^{n-j} a_j \delta_{nj} = a_n \end{aligned}$$

Now suppose LHS. Also the other claim in Lemma 10.9, we have:

$$\begin{aligned} \sum_{k=0}^n (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} a_k &= \sum_{k=0}^n (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} \sum_{m=0}^k \left\{ \begin{matrix} k \\ m \end{matrix} \right\} b_m = \sum_{m=0}^n b_m \sum_{k=m}^n (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} \left\{ \begin{matrix} k \\ m \end{matrix} \right\} \\ &= \sum_{m=0}^n b_m \delta_{nm} = b_n \end{aligned}$$

Alternatively, we can use generating functions. Let A and B be the exponential generating functions of the sequences. Suppose LHS holds. Then using Proposition 5.20 and Proposition 3.27

$$A(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} b_k = \sum_{k=0}^{\infty} b_k \sum_{n=k}^{\infty} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \frac{x^n}{n!} = \sum_{k=0}^{\infty} b_k \frac{(e^x - 1)^k}{k!} = B(e^x - 1)$$

Take $x \leftarrow \log(1+x)$ to obtain $B(x) = A(\log(1+x))$. Expanding:

$$\begin{aligned} B(x) &= A(\log(1+x)) = \sum_{k=0}^{\infty} a_k \frac{(\log(1+x))^k}{k!} = \sum_{k=0}^{\infty} a_k \frac{(-\log(\frac{1}{1+x}))^k}{k!} = \sum_{k=0}^{\infty} (-1)^k a_k \frac{(\log(\frac{1}{1+x}))^k}{k!} \\ &= \sum_{k=0}^{\infty} (-1)^k a_k \sum_{n=0}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} \frac{(-x)^n}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^n (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} a_k \end{aligned}$$

Applying $[x^n]$ yields the result. The reverse direction is the same but in reverse. \square

10.3 Lah Inversion

Lemma 10.11. $\sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^k = (-1)^n \delta_{nm}$

Proof. We use the closed form for the Lah numbers in Proposition 5.48 and the corresponding orthogonality result for binomial coefficients in Lemma 10.1:

$$\begin{aligned} \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^k &= \sum_{k=0}^n \binom{n}{k} \frac{(n-1)!}{(k-1)!} \binom{k}{m} \frac{(k-1)!}{(m-1)!} (-1)^k = \frac{(n-1)!}{(m-1)!} \sum_{k=0}^n \binom{n}{k} \binom{k}{m} (-1)^k \\ &= \frac{(n-1)!}{(m-1)!} \delta_{nm} (-1)^n = (-1)^n \delta_{nm} \quad \square \end{aligned}$$

Theorem 10.12 (Lah Inversion). Let $\{a_n\}, \{b_n\}$ be sequences for $n \geq 0$. Then:

$$a_n = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} b_k \text{ for all } n \iff b_n = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} a_k \text{ for all } n$$

↳ The symmetric version is $a_n = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} b_k (-1)^k$ iff $b_n = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} a_k (-1)^k$

Proof. Suppose the RHS. Using Lemma 10.11:

$$\begin{aligned} \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} b_k &= \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \sum_{j=0}^k \begin{bmatrix} k \\ j \end{bmatrix} a_j (-1)^{k-j} = \sum_{j=0}^n a_j (-1)^{-j} \sum_{k=j}^n \begin{bmatrix} n \\ k \end{bmatrix} \begin{bmatrix} k \\ j \end{bmatrix} (-1)^k \\ &= \sum_{j=0}^n a_j (-1)^{-j} (-1)^n \delta_{nj} = a_n \end{aligned}$$

For the symmetric version, substitute $b_n \leftarrow (-1)^n b_n$. From the symmetric version, by taking $a_n \leftarrow (-1)^n a_n$, we can obtain the reverse of the original version.

Alternatively, we can use generating functions. Let A and B be the exponential generating functions of the sequences. Suppose LHS holds. Then by Lemma 5.47:

$$A(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} b_k = \sum_{k=0}^{\infty} b_k \sum_{n=k}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} \frac{x^n}{n!} = \sum_{k=0}^{\infty} b_k \frac{\left(\frac{x}{1-x}\right)^k}{k!} = B\left(\frac{x}{1-x}\right)$$

Take $x \leftarrow \frac{x}{1+x}$. Then we obtain our result:

$$\begin{aligned} B(x) &= A\left(\frac{x}{1+x}\right) = \sum_{k=0}^{\infty} a_k \frac{\left(\frac{x}{1+x}\right)^k}{k!} = \sum_{k=0}^{\infty} a_k (-1)^k \frac{\left(\frac{-x}{1-(-x)}\right)^k}{k!} = \sum_{k=0}^{\infty} a_k (-1)^k \sum_{n=0}^{\infty} \begin{bmatrix} n \\ k \end{bmatrix} \frac{(-x)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} a_k \end{aligned}$$

Hence, we have the forward direction. Proceed the same as before to get the reverse and symmetric claims. \square

11. Lagrange Inversion

Theorem 11.1 (Lagrange Inversion). Suppose $W(z)$ and $\phi(z)$ are formal power series in z with $\phi(0) \neq 0$. Then there is a unique formal power series $z(x)$ such that $z = x\phi(z)$. Additionally, the composition $W(z(x))$ as a formal power series in x has:

$$n[x^n]W(z(x)) = [z^{n-1}](W'(z)\phi(z)^n)$$

Taking $W(z) = z$ yields the simplified form:

$$n[x^n]z(x) = [z^{n-1}]\phi(z)^n$$

An equivalent formulation is as follows for any formal power series $W(x)$ and $F(x)$ with $f_0 = 0$ and $f_1 \neq 0$:

$$n[x^n]W(F^{-1}(x)) = [x^{n-1}]W'(x) \left(\frac{x}{F(x)} \right)^n$$

Here, $F^{-1}(x)$ denotes the compositional inverse, which exists by Proposition 7.9. The quantity $\frac{x}{F(x)}$ is formally the reciprocal of $\frac{F(x)}{x}$, which exists since $f_0 = 0$ and $f_1 \neq 0$ by Theorem 7.7.

- ↳ When counting combinatorial objects, oftentimes one will be able to derive a functional equation for the generating functions. Lagrange Inversion can allow us to extract the coefficients even when this functional equation is not solvable.
- ↳ When a formal power series counts certain objects, then suitable functions of it can count other meaningful objects. This is why W is included.

Proof. Start with the second formulation. We show that it is equivalent to the first form, and that it also yields a $z(x)$ satisfying $z = x\phi(z)$. We then argue that this recurrence ensures z is unique. We finish with the main statement of the Theorem, namely the coefficients of $W(z(x))$.

Suppose the second formulation is true, and consider any $W(z)$ and $\phi(z)$ with $\phi(0) \neq 0$. Take $F(z) = z/\phi(z)$, which is well-defined by Theorem 7.7. This satisfies $[z^0]F(z) = 0$ and $[z^1]F(z) = [z^0]1/\phi(z) = -\phi(0)^{-1} \neq 0$ by the proof of Theorem 7.7. We now show that there exists $z(x)$ such that $z = x\phi(z)$. Take $z(x) = F^{-1}(x)$, which exists by Proposition 7.9. Then $x\phi(z) = x \frac{z}{F(z)} = x \frac{z}{F(F^{-1}(z))} = x$ as desired. Substituting $F^{-1}(x) = z(x)$ and $\frac{x}{F(x)} = \phi(x)$ gives $n[x^n]W(z(x)) = [z^{n-1}](W'(z)\phi(z)^n)$, which is the first formulation.

Now suppose the first formulation is true, and consider any $W(x)$ and $F(x)$ with $f_0 = 0$ and $f_1 \neq 0$. Take $\phi(z) = z/F(z)$, which satisfies $\phi(0) \neq 0$ by the proof of Theorem 7.7. As we already showed, $z(x) = F^{-1}(x)$ satisfies $z = x\phi(z)$. Substituting yields $n[x^n]W(F^{-1}(x)) = [x^{n-1}]W'(x) \left(\frac{x}{F(x)} \right)^n$.

Now that we have equivalence of the two formulations, we prove that z is unique. This is done by strong induction, showing that z_n is unique for each n . $z_0 = 0$ since $[x^0]x\phi(z) = [x^{-1}]\phi(z) = 0$, hence establishing the base case. We now have $z_n = [x^n]x\phi(z) = [x^{n-1}]\phi(z)$ which only depends on ϕ and z_0, \dots, z_{n-1} .

Finally, we prove the main statement of the Theorem, namely the coefficients of $W(z(x))$. We proceed by strong induction on n . Working term-by-term in the expansion of W , it suffices to show $n[x^n]z(x)^k = k[z^{n-1}](z^{k-1}\phi(z)^n)$ for all k . The claim is trivial for $n \leq 0$, which we use as our base case. Consider a few special cases for $n > 0$:

- $k = 0$. In this case, RHS is 0, and since $n > 0$, we have LHS is 0.
- $k > n$. Then $n - k < 0$, so that $n[x^n]z(x)^k = n[x^n]x^k\phi(z)^k = n[x^{n-k}]\phi(z)^k = 0$, and thus LHS is 0. RHS is 0 by the same reasoning.
- $k = n$. Then:

$$\begin{aligned} n[x^n]z(x)^n &= n[x^n]x^n\phi(z)^n = n[x^0]\phi(z)^n = n\phi(z(0))^n \\ &= n\phi(0)^n = n[z^0]\phi(z)^n = n[z^{n-1}]z^{n-1}\phi(z)^n \end{aligned}$$

Now suppose $0 < k < n$. Then:

$$\begin{aligned}
 n[x^n]z(x)^k &= n[x^n]x^k\phi(z)^k = n[x^{n-k}]\phi(z)^k \stackrel{\text{I.H.}}{=} \frac{n}{n-k}[z^{n-k-1}](\phi(z)^k)'\phi(z)^{n-k} \\
 &= \frac{n}{n-k}[z^{n-k-1}]k\phi(z)^{k-1}\phi'(z)\phi(z)^{n-k} = \frac{k}{n-k}[z^{n-k-1}]n\phi(z)^{n-1}\phi'(z) \\
 &= \frac{k}{n-k}[z^{n-k-1}](\phi(z)^n)' = \frac{k}{n-k}[z^{n-k}]z(\phi(z)^n)' \stackrel{\text{II}}{=} \frac{k}{n-k}[z^{n-k}](n-k)\phi(z)^n \\
 &= k[z^{n-1}]z^{k-1}\phi(z)^n
 \end{aligned}$$

□

Definition 11.2. The formal Laurent series $\mathbb{F}((x))$ is an extension of the formal power series $\mathbb{F}[[x]]$ given by:

$$\mathbb{F}((x)) = \left\{ \sum_{n \geq -N} c_n x^n \mid N \in \mathbb{Z}, c_n \in \mathbb{F} \right\}$$

↳ Essentially, we allow a finite number of negative powers

Proposition 11.3. Formal Laurent series satisfy the same rules for sums, products, and formal derivatives as formal power series.

Lemma 11.4.

- (a) If $z(x) \in \mathbb{F}((x))$, then $[x^{-1}]z'(x) = 0$.
- (b) If $f(x) \in \mathbb{F}[[x]]$ with $f(0) = 0$ and $f'(0) \neq 0$, then $[x^{-1}]f(x)^k f'(x) = \delta_{-1,k}$ for $k \in \mathbb{Z}$

Proof. The first claim follows since any x^{-1} term would have to originate from the derivative of an x^0 term, but this will always be zero. For the second claim, suppose $k \neq -1$. Then $[x^{-1}]f(x)^k f'(x) = [x^{-1}] \frac{1}{k+1} (f(x)^{k+1})' = 0$ by the first claim. If instead $k = 1$, then we have:

$$\begin{aligned}
 f(x)^{-1}f'(x) &= \frac{f_1 + 2f_2x + 3f_3x^2 + \cdots}{f_1x + f_2x^2 + f_3x^3 + \cdots} = \frac{1}{x} + \frac{f_2x + f_3x^2 + \cdots}{f_1x + f_2x^2 + f_3x^3 + \cdots} \\
 &= \frac{1}{x} + \frac{f_2 + f_3x + \cdots}{f_1 + f_2x + f_3x^2 + \cdots}
 \end{aligned}$$

Since $f'(0) \neq 0$, we have $f_1 \neq 0$. Hence, one divided by the denominator is a formal power series by Theorem 7.7, so that the entire fraction is a formal power series. Thus, $[x^{-1}]f(x)^{-1}f'(x) = 1$. □

Proposition 11.5. Consider formal power series $f(x) \in \mathbb{F}[[x]]$ with $f_0 = 0$ and $f_1 \neq 0$. Using Definition 7.12, $\deg(f(x)^{-k}) = -k$.

Proof.

$$\deg(f(x)^{-k}) = \deg\left(\frac{1}{f(x)^k}\right) = \deg\left(x^{-k} \frac{1}{(f(x)/x)^k}\right) = \deg\left(\frac{1}{(f(x)/x)^k}\right) - k = -k$$

This is justified since $f(x)/x$ is a formal power series with nonzero constant term, so that $(f(x)/x)^k$ is as well. Hence, the reciprocal of $(f(x)/x)^k$ is a formal power series with nonzero constant term by Theorem 7.7, and thus has degree 0. □

Theorem 11.6. We can also use formal Laurent series to offer an alternative proof to Lagrange Inversion, specifically for the second formulation.

Proof. Working term-by-term in the expansion of W , it suffices to show:

$$n[x^n]F^{-1}(x)^k = k[x^{n-1}]\left(x^{k-1}\left(\frac{x}{F(x)}\right)^n\right) \quad \text{or} \quad n[x^n]F^{-1}(x)^k = k[x^{n-k}]\left(\frac{x}{F(x)}\right)^n$$

Write the expansion of $F^{-1}(x)^k$ as $\sum_{j=k}^{\infty} p_j x^j$, where we have noted that the constant term in F^{-1}

is zero. Substituting $x \leftarrow F(x)$ gives:

$$x^k = F^{-1}(F(x))^k = \sum_{j=k}^{\infty} p_j F(x)^j$$

Differentiate each side and then divide by $F(x)^n$:

$$\frac{kx^{k-1}}{F(x)^n} = \sum_{j=k}^{\infty} j p_j F(x)^{j-n-1} F'(x)$$

Thus, we have:

$$\begin{aligned} k[x^{n-k}] \left(\frac{x}{F(x)} \right)^n &= [x^{-1}] \frac{kx^{k-1}}{F(x)^n} = [x^{-1}] \sum_{j=k}^{\infty} j p_j F(x)^{j-n-1} F'(x) \stackrel{11.4}{=} \sum_{j=k}^{\infty} j p_j \delta_{nj} = n p_n \\ &= n[x^n] \sum_{j=k}^{\infty} p_j x^j = n[x^n] F^{-1}(x)^k \end{aligned} \quad \square$$

Example 11.7. $b_n = \sum_{k=0}^n \binom{k}{n-k} a_k$ iff $a_n = \frac{1}{n} \sum_{k=0}^n \binom{2n-k-1}{n-k} k b_k (-1)^{n-k}$ and $a_0 = 0$

Proof. Consider the generating function for $\{b_n\}$:

$$\begin{aligned} B(x) &= \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{k}{n-k} a_k x^n = \sum_{k=0}^{\infty} a_k \sum_{n=k}^{\infty} \binom{k}{n-k} x^n = \sum_{k=0}^{\infty} a_k x^k \sum_{n=k}^{\infty} \binom{k}{n-k} x^{n-k} \\ &= \sum_{k=0}^{\infty} a_k x^k \sum_{n=0}^{\infty} \binom{k}{n} x^n = \sum_{k=0}^{\infty} a_k x^k (1+x)^k = A(x(1+x)) \end{aligned}$$

By the quadratic formula, we could explicitly solve for the compositional inverse of $x(1+x)$, as we did in the proofs of other inversion formulas. However, the power of Lagrange Inversion is that it does not require this. Let $F(x) = x(1+x)$. Then $A(F(x)) = B(x)$. Substituting $x \leftarrow F^{-1}(x)$ gives $A(x) = B(F^{-1}(x))$. By Lagrange Inversion:

$$\begin{aligned} a_n &= [x^n] B(F^{-1}(x)) = \frac{1}{n} [x^{n-1}] B'(x) \left(\frac{x}{F(x)} \right)^n = \frac{1}{n} [x^{n-1}] B'(x) \left(\frac{1}{1+x} \right)^n \\ &= \frac{1}{n} \sum_{k=0}^{n-1} ([x^k] B'(x)) [x^{n-k-1}] \left(\frac{1}{1+x} \right)^n = \frac{1}{n} \sum_{k=0}^{n-1} (k+1) b_{k+1} [x^{n-k-1}] \frac{1}{(1+x)^n} \\ &= \frac{1}{n} \sum_{k=1}^n k b_k [x^{n-k}] \frac{1}{(x+1)^n} \stackrel{7.20}{=} \frac{1}{n} \sum_{k=1}^n k b_k (-1)^n \left(\frac{1}{-1} \right)^{n-k+n} \binom{n-k+n-1}{n-k-1} \\ &= \frac{1}{n} \sum_{k=1}^n k b_k (-1)^{n-k} \binom{2n-k-1}{n-k-1} = \frac{1}{n} \sum_{k=0}^n \binom{2n-k-1}{n-k-1} k b_k (-1)^{n-k} \end{aligned}$$

For the reverse direction, begin with a_n equaling the last expression, and work backwards to obtain $a_n = [x^n] B(F^{-1}(x))$. Hence, $B(x) = A(F(x))$, so work backwards in the first equation to obtain $B(x) = \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{k}{n-k} a_k x^n$. In other words, the above work proves both directions.

If we did not use Lagrange Inversion, we would need to invert $x(1+x)$. Solving with the quadratic formula and picking the root with no constant term gives $F^{-1}(x) = \frac{-1+\sqrt{1+4x}}{2}$. In order to solve this, we require more complicated machinery, such as 2.5.15 in generating functionology:

$$a_n = [x^n] A(x) = [x^n] B \left(\frac{-1+\sqrt{1+4x}}{2} \right) = \sum_{k=0}^{\infty} b_k [x^n] \left(\frac{-1+\sqrt{1+4x}}{2} \right)^k$$

$$\begin{aligned}
&= \sum_{k=0}^{\infty} b_k (-1)^{n-k} [x^n] \left(\frac{1 - \sqrt{1-4x}}{2} \right)^k \\
&= \sum_{k=0}^{\infty} b_k (-1)^{n-k} \frac{1}{n} \cdot ((n-1) + 1) [x^{(n-1)+1}] \left(\frac{1 - \sqrt{1-4x}}{2} \right)^k \\
&= \frac{1}{n} \sum_{k=0}^{\infty} b_k (-1)^{n-k} [x^{n-1}] D \left[\left(\frac{1 - \sqrt{1-4x}}{2} \right)^k \right] \\
&= \frac{1}{n} \sum_{k=0}^{\infty} k b_k (-1)^{n-k} [x^{n-1}] \frac{1}{\sqrt{1-4x}} \left(\frac{1 - \sqrt{1-4x}}{2} \right)^{k-1} \\
&= \frac{1}{n} \sum_{k=0}^{\infty} k b_k (-1)^{n-k} [x^{n-k}] \frac{1}{\sqrt{1-4x}} \left(\frac{1 - \sqrt{1-4x}}{2x} \right)^{k-1} \\
&= \frac{1}{n} \sum_{k=0}^{\infty} k b_k (-1)^{n-k} \binom{2(n-k) + (k-1)}{n-k} = \frac{1}{n} \sum_{k=0}^{\infty} \binom{2n-k-1}{n-k} k b_k (-1)^{n-k} \quad \square
\end{aligned}$$

Example 11.8. Suppose the vertices of a convex polygon P are labeled cyclically $1, 2, 1, 2, 1, 2, \dots$. If the number of vertices is odd, one pair of adjacent vertices will have the same label (both will be 1, and this edge is often drawn on top). Call a triangulation of P proper if no triangle is monochromatic. Let a_N count the number of such triangulations if P has $N + 2$ vertices. Then for $N > 0$:

$$a_N = \begin{cases} \frac{2^n}{2n+1} \binom{3n}{n}, & N = 2n \\ \frac{2^{n+1}}{2n+2} \binom{3n+1}{n}, & N = 2n + 1 \end{cases}$$

↳ This originates from a 2009 paper by Bruce Sagan called “Proper Partitions of a Polygon and k -Catalan Numbers”

Proof. We first show that a_n satisfy:

$$\begin{aligned}
a_{2n+1} &= \sum_{j=0}^n a_{2j} a_{2n-2j} \\
a_{2n} &= \sum_{j=0}^{2n-1} a_j a_{2n-1-j}
\end{aligned}$$

For the first case, there will be an edge in P both of whose vertices are labeled 1. Condition on the other vertex in the triangle containing this edge. It must be one of the $n + 1$ vertices that is a 2, and it will divide P into a $(j + 2)$ -gon and a $(2n + 3 - (j + 2) + 1)$ -gon for any even j . There are thus $a_j a_{2n-j}$ ways to finish the triangulation. Replacing j with $2j$ to ensure that it is even gives the recurrence.

For the second case, fix any edge and condition on the triangle containing this edge. It can contain any of the other $2n$ vertices, and will divide P into a $(j + 2)$ -gon and a $(2n + 2 - (j + 2) + 1)$ -gon. There are thus $a_j a_{2n-1-j}$ ways to finish the triangulation, giving the result.

Let $E(x) = \sum_{n=1}^{\infty} a_{2n} x^n$ and $O(x) = \sum_{n=0}^{\infty} a_{2n+1} x^n$. We calculate:

$$\begin{aligned}
E(x) &= \sum_{n=1}^{\infty} a_{2n} x^n = \sum_{n=1}^{\infty} \sum_{j=0}^{2n-1} a_j a_{2n-1-j} x^n = \sum_{j=0}^{\infty} a_j \sum_{n=\lceil (j+1)/2 \rceil}^{\infty} a_{2n-1-j} x^n \\
&= \sum_{j=0}^{\infty} a_{2j+1} \sum_{n=j+1}^{\infty} a_{2(n-j)-1} x^n + \sum_{j=0}^{\infty} a_{2j} \sum_{n=j+1}^{\infty} a_{2(n-j)-1} x^n \\
&= \sum_{j=0}^{\infty} a_{2j+1} x^{j+1} (1 + E(x)) + \sum_{j=0}^{\infty} a_{2j} x^{j+1} O(x) \\
&= x(1 + E(x))O(x) + xO(x)(1 + E(x)) = 2xO(x)(1 + E(x))
\end{aligned}$$

We also calculate:

$$\begin{aligned} O(x) &= \sum_{n=0}^{\infty} a_{2n+1}x^n = \sum_{n=0}^{\infty} \sum_{j=0}^n a_{2j}a_{2n-2j}x^n = \sum_{j=0}^{\infty} a_{2j} \sum_{n=j}^{\infty} a_{2(n-j)}x^n \\ &= \sum_{j=0}^{\infty} a_{2j}x^j \sum_{n=0}^{\infty} a_{2n}x^n = (1 + E(x))^2 \end{aligned}$$

Substituting, we have $E(x) = 2x(1 + E(x))^3$. Hence, we can let $\phi(z) = 2(1 + z)^3$, so that $E(x) = x\phi(E(x))$. Thus, for $n > 0$, Lagrange Inversion gives:

$$[x^n]E(x) = \frac{1}{n}[z^{n-1}]\phi(z)^n = \frac{2^n}{n}[z^{n-1}](1+z)^{3n} = \frac{2^n}{n} \binom{3n}{n-1} = \frac{2^n}{2n+1} \binom{3n}{n}$$

Since $O(x)$ is a function of $E(x)$, its coefficients can be computed by the same application of Lagrange Inversion, but with $W(z) = (1+z)^2$:

$$\begin{aligned} [x^n]O(x) &= [x^n](1 + E(x))^2 = [x^n]W(E(x)) = \frac{1}{n}[z^{n-1}]W'(z)\phi(z)^n = \frac{1}{n}[z^{n-1}]2(1+z) \cdot 2^n(1+z)^{3n} \\ &= \frac{2^{n+1}}{n}[z^{n-1}](1+z)^{3n+1} = \frac{2^{n+1}}{n} \binom{3n+1}{n-1} = \frac{2^{n+1}}{2n+2} \binom{3n+1}{n} \end{aligned}$$

This highlights the importance of using $W(z)$ when computing functions of other generating functions or counting sequences. Directly expanding $W(z)$ to compute the coefficients of $O(x)$ becomes messy and does not yield a usable form. \square

Example 11.9. Consider counting the number of ways n identical variables can be bracketed, where a bracket is defined recursively as:

- x is a bracket
- For $k \geq 2$, if $\delta_1, \dots, \delta_k$ are bracketed expressions, then so is $(\delta_1\delta_2 \cdots \delta_k)$

Proof. Let \mathcal{S} be the class of all bracketings with the size defined by the number of non-parenthesis characters. Then $\mathcal{S} = \mathcal{Z} + \text{SEQ}_{\geq 2}(\mathcal{S})$. This implies that $S(x) = x + \frac{S(x)^2}{1-S(x)}$. Rearranging gives $2S(x)^2 + (1-x)S(x) + x = 0$. Applying the quadratic formula and picking the root which is well-defined at zero gives $S(x) = \frac{1+x-\sqrt{1-6x+x^2}}{4}$. We can use the Generalized Binomial Theorem to extract the coefficients, but the result is messy:

$$\begin{aligned} [x^n]S(x) &= \frac{\delta_{0n} + \delta_{1n} + \sum_{k=0}^{\infty} \binom{1/2}{k} [x^n](x^2 - 6x)^k}{4} = \frac{\delta_{0n} + \delta_{1n} + \sum_{k=0}^{\infty} \binom{1/2}{k} [x^{n-k}](x-6)^k}{4} \\ &= \frac{\delta_{0n} + \delta_{1n} + \sum_{k=0}^{\infty} \binom{1/2}{k} \binom{k}{n-k} (-6)^{k-(n-k)}}{4} = \frac{\delta_{0n} + \delta_{1n} + \sum_{k=0}^{\infty} \binom{1/2}{k} \binom{k}{2k-n} (-6)^{2k-n}}{4} \end{aligned}$$

Alternatively, we can use Lagrange Inversion. Multiply the recurrence by $1 - S(x)$ to obtain $S(x)(1 - S(x)) = x(1 - S(x)) + S(x)^2$. Rearrange to obtain $S(x) - 2S(x)^2 = x(1 - S(x))$, or $S(x) = x \frac{1-S(x)}{1-2S(x)}$. So, let $\phi(z) = \frac{1-z}{1-2z}$. \square

Theorem 11.10. Suppose $z = z(x)$ satisfies $z = x\phi(z)$ and $\phi(0) \neq 0$. Let $W(x)$ be a formal power series. Then for $n \geq 0$:

$$[x^n]W(z)\phi(z)^n = [x^n] \frac{xW(z)z'(x)}{z(x)} = [x^n] \frac{W(z)}{1 - x\phi'(z)}$$

Proof. By working term-by-term in the expansion of $W(z)$, it suffices to prove the claim when $W(z) = z^k$ for some $k \geq 0$. Choose $H(z)$ so that $H'(z) = \phi(z)^{k-1}$. When $k = 1$, the reciprocal is well-defined since $\phi(0) \neq 0$. Then:

$$[x^n]W(z)\phi(z)^n = [x^{n-k}]\phi(z)^n = [x^{n-k}]H'(z)\phi(z)^{n-k+1} = (n-k+1)[x^{n-k+1}]H(z(x))$$

$$\begin{aligned}
&= [x^{n-k}]D[H(z(x))] = [x^{n-k}]H'(z)z'(x) = [x^{n-k}]\frac{\phi(z)^k z'(x)}{\phi(z)} = [x^{n-k}]\frac{z^k z'(x)}{x^k \phi(z)} \\
&= [x^n]\frac{W(z)z'(x)}{\phi(z)} = [x^n]\frac{xW(z)z'(x)}{z(x)}
\end{aligned}$$

Now, we use $z(x) = x\phi(z)$ to derive $z'(x) = \phi(z) + x\phi'(z)z'(x)$. Solving, we obtain $z'(x) = \frac{\phi(z)}{1-x\phi'(z)} = \frac{z}{x(1-x\phi'(z))}$. Substituting yields the result. \square

Example 11.11. Let $g_n = [x^n](1+x+x^2)^n$. Then we also have $g_n = [x^n]\frac{1}{\sqrt{1-2x-3x^2}}$.

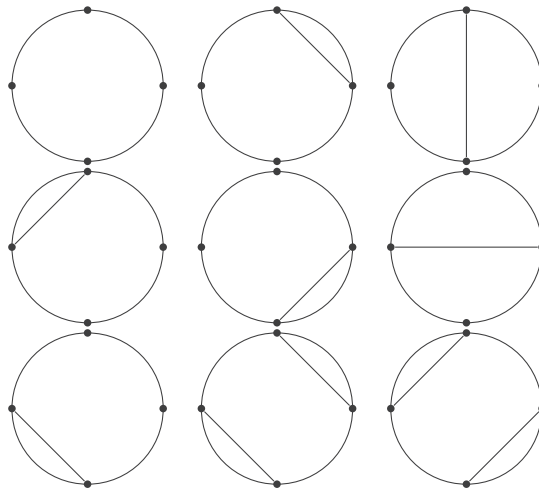
Proof. Let $\phi(z) = 1+z+z^2$. We solve for $z = z(x)$ so that $z(x) = x\phi(z(x))$. We have $z = x(1+z+z^2)$, or $xz^2 + (x-1)z + x = 0$. Applying the quadratic formula and picking the root that is well-defined at 0 gives $z(x) = \frac{1-x-\sqrt{1-2x-3x^2}}{2x}$. Now, we apply Theorem 11.10 to obtain:

$$g_n = [z^n]\phi(z)^n = [x^n]\frac{1}{1-x\phi'(z)} = [x^n]\frac{1}{1-x(2 \cdot \frac{1-x-\sqrt{1-2x-3x^2}}{2x} + 1)} = [x^n]\frac{1}{\sqrt{1-2x-3x^2}} \quad \square$$

11.1 Motzkin Numbers

Definition 11.12. Consider an election with two candidates x and y , and suppose that the candidates receive the same number of votes but voters may also choose “none of the above.” Represent the sequence of votes using a string of x , y , and N . Consider a string to be legal if the number of x and y are the same, and the number of y never exceeds the number of x as we read from left to right. The Motzkin numbers m_n count the number of legal strings of length n defined in this way. Let the set of such strings be $M([n])$.

Proposition 11.13. Consider a circle with n points on its edge. The Motzkin numbers also count the number of ways to add chords between these points such that none of the chords intersect, even at the edge of the circle. Not all points must have a chord attached. Below are examples for $n = 4$, yielding $m_4 = 4$:



Proof. We form a bijection from the circle representation to the string representation. Identify a vertex of the circle to be the starting vertex. If it is part of a chord, begin the string with x . Otherwise, begin the string with N . Now, proceed clockwise. To the string, add an N for each vertex not belonging to a chord, and add either an x or y otherwise. Add an x if the chord has been visited already, and a y otherwise. The string will be legal since each chord will cause both an x and a y to be added to the string, with the x preceding the y .

For the reverse, label the vertices in clockwise order with the integers in $[n]$. From the string, representing each x by a (, each y by), and removing each N yields a legal string of parentheses. Identify corresponding x and y by identifying corresponding parentheses. Draw chords between the

indices of the corresponding x and y . The nested structure of the parentheses (and thus the x and y) ensures that no chords will intersect. \square

Proposition 11.14. $m_0 = 1$ and $m_n = m_{n-1} + \sum_{k=2}^n m_{k-2}m_{n-k}$

Proof. First, we use the string representation. Let $T \in M([n])$. Then T is either of the form $T = NT'$ or $T = xT''yT'''$, where T', T'', T''' are legal strings. The first case has m_{n-1} elements. For the second case, condition on the position $k \neq 2$ of y . Then $T'' \in M([k-2])$ and $T''' \in M([n-k])$, so that there are $m_{k-2}m_{n-k}$ elements.

We can also use the chord representation. Label the vertices in clockwise order with the integers in $[n]$. If vertex 1 has no chord attached, then it can be removed to yield a chord representation on $n-1$ vertices. Otherwise, merge the vertex with its opposing vertex $k \geq 2$ at the other side of the chord to obtain a figure-8, then split this into two circles. This will yield a circle with $k-2$ vertices and one with $n-k$ vertices. Sum over all k . \square

Proposition 11.15. The generating function for the Motzkin numbers satisfies $M(x) = 1 + xM(x) + x^2M(x)^2$.

Proof. Using Proposition 11.14, we have that the ordinary generating function for m_n satisfies:

$$\begin{aligned} M(x) &= 1 + \sum_{n=1}^{\infty} \left(m_{n-1} + \sum_{k=2}^n m_{k-2}m_{n-k} \right) x^n = 1 + \sum_{n=1}^{\infty} m_{n-1}x^n + \sum_{n=1}^{\infty} \sum_{k=2}^n m_{k-2}m_{n-k}x^n \\ &= 1 + xM(x) + \sum_{k=2}^{\infty} m_{k-2} \sum_{n=k}^{\infty} m_{n-k}x^n = 1 + xM(x) + \sum_{k=2}^{\infty} m_{k-2}x^k \sum_{n=0}^{\infty} m_nx^n \\ &= 1 + xM(x) + x^2M(x)^2 \end{aligned} \quad \square$$

Proposition 11.16. $M_n = \frac{1}{n+1} \sum_{k=0}^{n+1} \binom{n+1}{k} \binom{n-k+1}{k-1}$

Proof. We will use Lagrange Inversion. Multiply each side of Proposition 11.15 by x to obtain $xM(x) = x(1+xM(x)+x^2M(x)^2)$, then let $z(x) = xM(x)$ so that $z(x) = x\phi(z)$ with $\phi(z) = 1+z+z^2$. Thus:

$$\begin{aligned} [x^n]M(x) &= [x^{n+1}]z(x) = \frac{1}{n+1} [z^n]\phi(z)^{n+1} = \frac{1}{n+1} [z^n] \sum_{k=0}^{n+1} \binom{n+1}{k} (z+z^2)^{n-k+1} \\ &= \frac{1}{n+1} \sum_{k=0}^{n+1} \binom{n+1}{k} [z^n] z^{n-k+1} (1+z)^{n-k+1} = \frac{1}{n+1} \sum_{k=0}^{n+1} \binom{n+1}{k} [z^{k-1}] (1+z)^{n-k+1} \\ &= \frac{1}{n+1} \sum_{k=0}^{n+1} \binom{n+1}{k} \binom{n-k+1}{k-1} \end{aligned} \quad \square$$

Concept 11.17. It can be of interest to group the elements of $M([n])$ by the number of entries that are N . Calling this number k , we can form a table with columns k and rows n . When $k=0$ and n is even, we obtain the Catalan number $c_{n/2}$.

Lemma 11.18. $m_n = \sum_{k=0}^{\lceil n/2 \rceil} \binom{n}{2k} c_k$

Proof. Condition on the number of letters $2k$ in the string that are not N . Choose their locations in the string $\binom{n}{2k}$ ways, and pick whether they are x or y by picking a legal string counted by c_k . \square

Lemma 11.19. $c_{n+1} = \sum_{k=0}^n \binom{n}{k} m_k$

Proof. Let a_n count the number of ways to write a string of length n over the alphabet $\{X, Y, N, M\}$ such that the tallies for X always are at least the tallies for Y (counting left to right) and the number of X equals the number of Y . Such a string is called legal. We claim that $c_{n+1} = a_n$, which we show by forming a bijection as follows. To convert a legal string of $2n+2$ letters in $\{x, y\}$ to a legal string of n letters over $\{X, Y, N, M\}$, remove the leading x and the trailing y from the string, group the remaining $2n$ letters into n pairs, and make the replacements $xx \rightarrow X$, $yy \rightarrow Y$, $xy \rightarrow N$, and $yx \rightarrow M$. The inverse then makes the opposite replacements and adds a leading x and a trailing y .

Should the map be well-defined, it is clearly a bijection. However, we must show that the forward and inverse directions yield legal strings. It is clear that the requirement of an equal number of x and y (and an equal number of X and Y) are satisfied, so it remains to show that the tally requirements are satisfied. When tallying the first i letters in the string, let $t_{\alpha,i}$ denote the tally for symbol α . We claim that for two corresponding strings, $t_{x,i} \geq t_{y,i}$ for all i if and only if $t_{X,i} \geq t_{Y,i}$ for all i .

By definition of the map, $t_{x,2i+1} = 1 + 2t_{X,i} + t_{N,i} + t_{M,i}$ and $t_{y,2i+1} = 2t_{Y,i} + t_{N,i} + t_{M,i}$. Hence, $t_{x,2i+1} - t_{y,2i+1} = 1 + 2(t_{X,i} - t_{Y,i})$. This shows that $t_{x,2i+1} \geq t_{y,2i+1}$ if and only if $t_{X,i} \geq t_{Y,i}$. Furthermore, $t_{x,2i+1} - t_{y,2i+1}$ must be odd and thus at least 1. $t_{\alpha,i}$ viewed as a sequence in i can only change by at most 1 each time i is incremented, so that assuming $t_{x,2i+1} - t_{y,2i+1} \geq 1$ for all i , we have $t_{x,2i} - t_{y,2i} \geq 0$ for all i . Altogether, we have $t_{x,i} \geq t_{y,i}$ for all i if and only if $t_{X,i} \geq t_{Y,i}$ for all i . Hence, the bijection takes legal strings in one alphabet to legal strings in the other.

We have thus shown that $c_{n+1} = a_n$. It remains to show $a_n = \sum_{k=0}^n \binom{n}{k} m_k$. For a string counted by a_n , condition on the number of letters k that are not M . The $n-k$ letters that are M can appear in any location in the string without restriction, and thus can be picked $\binom{n}{n-k} = \binom{n}{k}$ ways. The remaining letters k can be viewed as a legal string as defined by the Motzkin numbers, so that there are m_k ways to pick these. We therefore have the result. \square

12. Posets

12.1 Introduction

Definition 12.1. A poset is a pair (P, \leq) where P is a set and \leq is a binary relation on P . The relation must be reflexive ($a \leq a$), anti-symmetric ($a \leq b$ and $b \leq a$ implies $a = b$), and transitive ($a \leq b$ and $b \leq c$ implies $a \leq c$). It need not be strongly connected (there can exist two elements where neither $a \leq b$ or $b \leq a$).

↳ We use $a < b$ to mean $a \leq b$ and $a \neq b$

Definition 12.2. $x \in P$ is a maximal element of P if there does not exist a y with $x < y$. If $z \leq x$ for all $z \in P$, then x is the maximum element of P . Similar definitions hold for minimal and minimum elements.

↳ Every finite poset has a maximal and minimal element, but does not necessarily have a maximum or minimum element.

↳ If they exist, the maximum element is denoted with $\hat{1}$, and the minimum element is denoted $\hat{0}$.

Definition 12.3. Suppose $s, t \in P$ are elements in a poset. Then t covers s (or s is covered by t) if $s < t$ and no element $u \in P$ satisfies $s < u < t$. This is denoted by $s < t$.

Definition 12.4. We denote the closed interval $[s, t] = \{u \in P \mid s \leq u \leq t\}$. Using this, t covers s if $[s, t] = \{s, t\}$. $[s, s] = \{s\}$ and the empty set is not considered an interval.

Definition 12.5. Two posets P and Q are isomorphic ($P \cong Q$) if there exists an order-preserving bijection $\phi : P \rightarrow Q$, so that $s \leq_P t$ if and only if $\phi(s) \leq_Q \phi(t)$.

Definition 12.6. A poset can be represented using a Hasse diagram. The nodes of the diagram represent each $x \in P$, and edges are drawn between nodes to represent covering relations. For each $x < y$, we must have that x is vertically below y and an edge connects x and y .

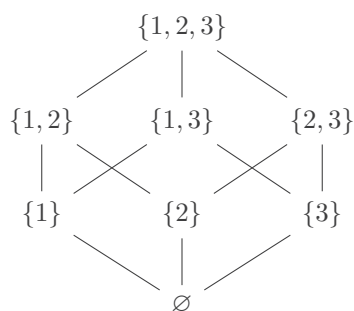
↳ By transitivity, $x < y$ iff x and vertically below y and there is a path of edges from x to y .

Definition 12.7. A chain is a poset (or subset of a poset) in which any two elements are comparable. If no elements in a subset of a poset are comparable, it is called an antichain. A chain or antichain is maximal if no further element can be added to it without breaking the chain/antichain property. It is maximum if P has no longer chain or antichain.

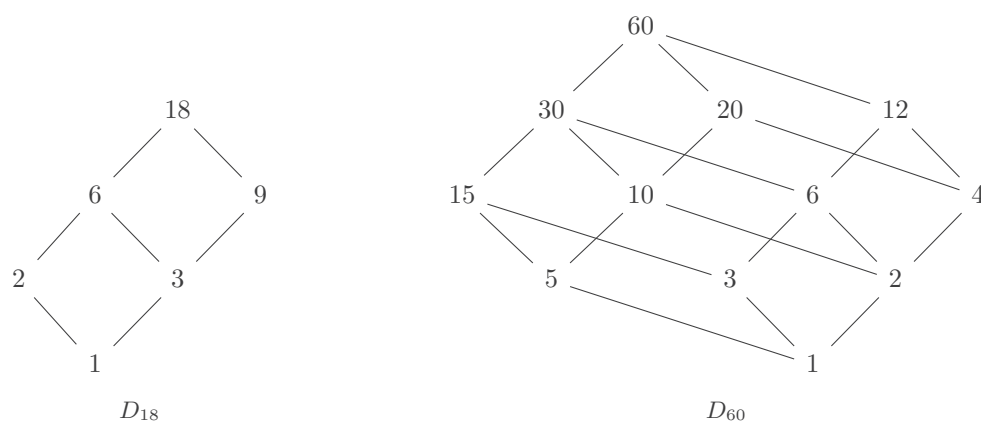
Example 12.8. Consider a chain of length n given by $C_n = \{0, 1, \dots, n\}$ with partial order defined as normal in \mathbb{Z} . The Hasse diagram is:



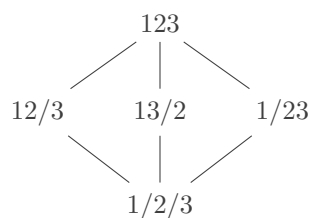
Example 12.9. The Boolean Algebra (B_n, \subset) where $B_n = \mathcal{P}([n])$ is a poset. The Hasse diagram for B_3 is as follows. Note that not all elements are comparable.



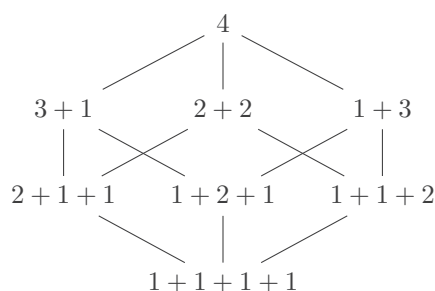
Example 12.10. The Divisor Lattice $(D_n, |)$ consists of all positive integers k that divide n . Example Hasse diagrams include:



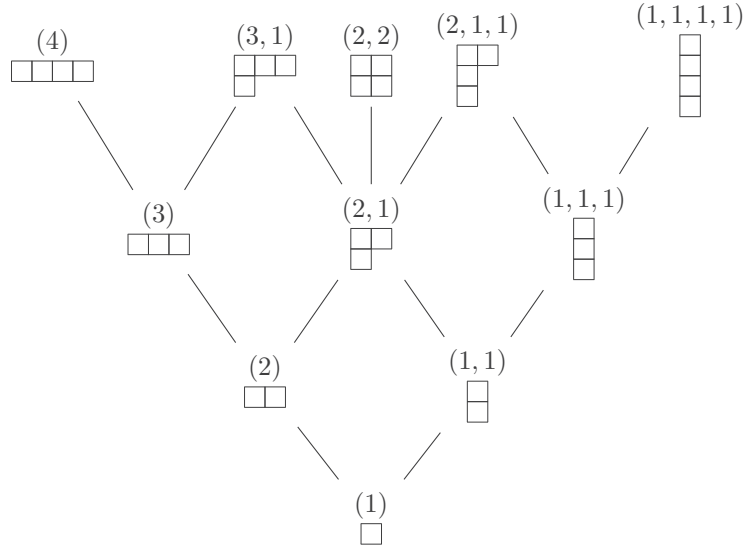
Example 12.11. The lattice of partitions (Π_n, \leq) is the set of all set partitions of $[n]$. For two partitions $\gamma, \delta \in \Pi_n$, $\gamma \leq \delta$ if every block of γ is contained in some block of δ , so that γ is a refinement of δ . The Hasse diagram for Π_3 is:



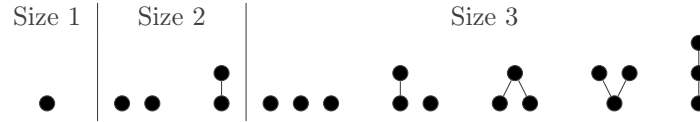
Example 12.12. The lattice of compositions (K_n, \leq) is the set of all compositions of n with order defined by the refinement of compositions. The Hasse diagram for K_4 is:



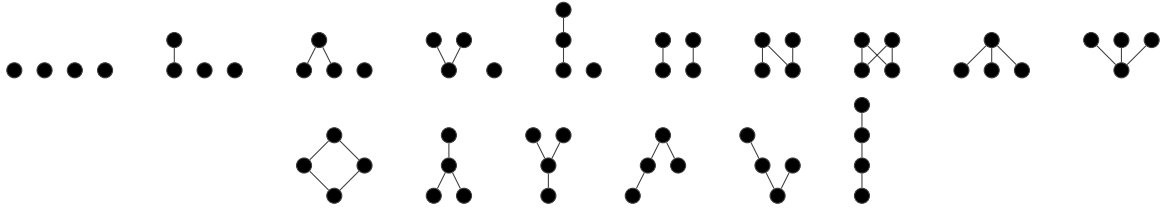
Example 12.13. The Young lattice (Y, \leq) is the set of all integer partitions and $\alpha \leq \beta$ if the Young diagram for α is contained in the Young diagram for β . The Hasse diagram for Y up to integer partitions of 4 is:



Concept 12.14. We can the Hasse diagrams for isomorphism classes of posets of a given size. We have:



Example 12.15. The posets of size 4 are:



Example 12.16. Every finite poset P is isomorphic to a subset of B_n for some n .

Proof. Pick $n = |P|$ and label each element by a unique element in $[n]$ via a linear extension; we use the values in $[n]$ to refer to the elements of P . Form the isomorphism $\phi : P \rightarrow Q$ for $Q \subset B_n$ by $k \mapsto \{k\} \cup \bigcup_{j < k} \phi(j)$. We prove by induction on k that if $q \in \phi(k)$, then $q \leq k$. The base case of $k = 1$ is trivial since 1 is minimal and maps to $\{1\}$. For induction, suppose $q \in \phi(k) = \{k\} \cup \bigcup_{j < k} \phi(j)$. Then either $q = k$ (which satisfies the claim) or $q \in \phi(j)$ for some $j < k$. By induction, $q \leq j < k$, so $q < k$ (which also satisfies the claim).

We show ϕ is injective. If $\phi(j) = \phi(k)$, then $j \in \phi(k)$ which means $j \leq k$. Conversely, $k \in \phi(j)$ which implies $k \leq j$. Hence, $k = j$.

Now we show ϕ is order-preserving. For any $j \leq k$, we have $\phi(j) \leq \phi(k)$ by definition. Furthermore, for any $j \not\leq k$, we have $j \notin \phi(k)$ by the contrapositive of the previous claim. Since $j \in \phi(j)$, we have $\phi(j) \not\leq \phi(k)$. \square

Definition 12.17. A linear extension of a poset with n elements is an order-preserving bijection from P onto $[n]$. That is, for bijection $f : P \rightarrow [n]$, $x \leq y$ implies $f(x) \leq f(y)$.

↳ Linear extensions allow us to code many aspects of a poset using upper triangular matrices, as we will see in later sections.

Example 12.18. We describe the linear extensions of elements of the posets of size 3 (see Concept 12.14). The first has $6 = 3!$ linear extensions, corresponding to any bijection from the three vertices to $[3]$. The second has 3 linear extensions. Indeed, the lone vertex can be mapped to any element in $[3]$,

and then the remaining two are uniquely determined so that order is preserved. The third has 2 linear extensions, where the two non-comparable elements are mapped to $\{1, 2\}$ and the maximum element is mapped to 3. The fourth also has 2 linear extensions, where the minimum element is mapped to 1 and the remaining two are mapped to $\{2, 3\}$. The last is a chain, and thus has 1 linear extension.

Example 12.19. Consider the posets:



Each has precisely 8 linear extensions. Indeed, the element with no relations can be assigned 4 ways. After this, the maximum/minimum element of the remaining elements must map to the maximum/minimum element of the remaining elements in $[4]$. The last two elements can be mapped $2! = 2$ ways.

Definition 12.20. A chain cover of a poset P is a collection of disjoint chains whose union is P . The size of a chain cover is the number of chains in it.

↳ The chains do not need to be maximal. In other words, the relations in the chains do not need to be covering relations.

Theorem 12.21 (Dilworth's Theorem). In a finite poset P , the size of any maximum antichain is equal to the size of the smallest chain cover.

Proof. Let a be the size of a maximum antichain and b be the size of any smallest chain cover. Then $a \leq b$ since each element of the antichain must belong to a separate chain in the cover. For the converse, we induct on the size n of P . The base case $n = 1$ is trivial, so assume $n > 1$. There are two cases:

- P contains an a -element antichain A with at least one element that is not minimal and at least one element that is not maximal. Then A splits P into two sets: U which contains all elements greater than or equal to some element of A , and L which contains all elements less than or equal to some element of A . $L \cap U = A$ since otherwise we would arrive at a contradiction by transitivity. Since A contains both a non-minimal and non-maximal element, U and L are non-empty, and so are each unions of a chains by the inductive hypothesis. Each chain in L has an element of A at its top, and each chain in U has an element of A at its bottom. Join the corresponding chains to obtain a covering of p by a chains.
- P does not have such an antichain. Then either all elements are minimal, or all are maximal. Pick x minimal and y maximal with $x \leq y$. Let P' be the poset with x and y removed. The maximal antichain of P' now has $a - 1$, so by induction, P' can be covered with $k - 1$ chains. Now add the chain $x \leq y$ to obtain that P can be covered by a chains.

□

Example 12.22. The number of 2-element chains in B_n is $3^n - 2^n$. Indeed, for $x \leq y$, there are 3^n ways to mark elements as either in x , in both x and y , or in neither. Then we subtract the 2^n pairs (x, y) such that $x = y$. This results means the number of 2-element antichains in B_n is:

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

12.2 Incidence Algebra

Definition 12.23. A poset is called locally finite if all closed intervals have a finite number of elements.

↳ Not all posets that are locally finite are also finite. For example, the integers are locally finite but not finite.

Definition 12.24. A set of elements $I \subset P$ is called an ideal if $x \in I$ and $y \leq x$ implies $y \in I$. In other words, an ideal is a set closed under choosing smaller elements. A principal ideal is a ideal generated by a single elements, so that $I = \{y \mid y \leq x\}$ for some $x \in P$.

Concept 12.25. In some of our later results, we require that every principal ideal of P is finite. This means that each element is larger than a finite number of elements only. This is a stronger requirement than P being locally finite. For example, the integers are locally finite but has no finite principal ideals.

Definition 12.26. Let P be a locally finite poset. The incidence algebra $I(P)$ is the set of all functions $f : \text{Int}(P) \rightarrow \mathbb{C}$, where $\text{Int}(P)$ is the set of all intervals in P . We denote $f(x, y) = f([x, y])$ for ease of notation. Addition is defined by $(f + g)(x, y) = f(x, y) + g(x, y)$ and multiplication by $(fg)(x, y) = \sum_{x \leq z \leq y} f(x, z)g(z, y)$.

This can be viewed through matrices for a finite poset. Indeed, consider a linear extension x_1, \dots, x_n of P , then for $f \in I(P)$ form the $n \times n$ matrix with ij th entry $f(x_i, x_j)$. Here, we define $f(x_i, x_j)$ with $[x_i, x_j] \notin \text{Int}(P)$. Since the linear extension respects the order in the poset, the matrix will be upper triangular with a certain pattern of zeros corresponding to non-comparable elements. The addition and multiplication operations in $I(P)$ correspond to the usual operations in the matrices. In particular, $I(P)$ is isomorphic to a subalgebra of the $n \times n$ upper triangular matrices.

↳ The proof that $\text{Int}(P)$ is an algebra is straight forward

Definition 12.27. The multiplicative identity in $I(P)$ is $\delta(x, y) = I[x = y]$. This satisfies $\delta f = f\delta = f$, since the corresponding matrix to δ is the identity matrix.

Lemma 12.28. Let $f \in I(P)$ for a locally finite poset P . Then the following conditions are equivalent:

- (i) f has a left inverse
- (ii) f has a right inverse
- (iii) f is invertible
- (iv) $f(x, x) \neq 0$ for all $x \in P$

Proof. (iii) implies (iv) since $fg = \delta$ implies $f(x, x)g(x, x) = 1$ for all $x \in P$. (i) and (ii) together imply (iii). It remains to show that (iv) implies (i) and (ii). We construct the right inverse, and a similar technique can be used for the left inverse. We inductively define:

$$g(x, x) = f(x, x)^{-1} \quad \text{and} \quad g(x, y) = -f(x, x)^{-1} \sum_{x < z \leq y} f(x, z)g(z, y) \quad \text{when } x \neq y$$

Note that each recursive computation of $g(x, y)$ decreases the size of the interval $[x, y]$, which is finite since P is locally finite. Hence, the above is well-defined. It is clearly the inverse when $x = y$, and when $x \neq y$, multiplying by $f(x, x)$ and rearranging yields $\sum_{x \leq z \leq y} f(x, z)g(z, y) = 0$, or $(fg)(x, y) = 0 = \delta(x, y)$ as desired. \square

Definition 12.29. Let P be a locally finite poset. The zeta function of P is defined by $\zeta(x, y) = I[x \leq y] = 1$, where $\zeta \in I(P)$.

Definition 12.30. Lemma 12.28 shows that ζ is invertible. The inverse μ is called the Möbius function of P . Using the proof of Lemma 12.28 and the fact that $\zeta(x, y) = 1$ for all $[x, y] \in I(P)$, we can compute μ for $x \neq y$ as either of:

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

↳ An easy way of viewing this is $\sum_{x \leq z \leq y} \mu(x, z) = \delta(x, y)$. For a Hasse diagram, we can evaluate $\mu(x, y)$ by starting at x and working upwards to label the value of $\mu(x, s)$ for $x \leq s \leq y$. Each value is determined by ensuring the sum of it along with the values below it sum to zero.

Definition 12.31. If P has a maximum and minimum element, then we denote $\mu(P) = \mu(\hat{0}, \hat{1})$.

Lemma 12.32. $\mu(x, y)$ is an integer for all $x, y \in P$.

Proof. Consider the matrix representation Z of ζ , which is upper triangular with diagonal elements 1. Thus, Z^{-1} also has diagonal elements 1, so that the claim follows inductively by the recursive form given in Definition 12.30.

Alternatively, note that $\det(Z) = 1$, since it is upper triangular with diagonal entries 1. Hence, Cramer's rule indicates that the solutions to $Z\mathbf{x} = \mathbf{e}_i$ will be integers for any standard basis vector \mathbf{e}_i . Hence, Z^{-1} has integer entries. \square

Definition 12.33. The Möbius function in number theory is defined as:

$$\mu(n) = \begin{cases} +1, & \text{if } n \text{ is a square-free positive integer with an even number of prime factors} \\ -1, & \text{if } n \text{ is a square-free positive integer with an odd number of prime factors} \\ 0, & \text{if } n \text{ has a squared prime factor} \end{cases}$$

Lemma 12.34. For a divisor lattice P (or more generally the lattice of natural numbers with order given by divisibility), the two different Möbius functions satisfy $\mu(x, y) = \mu(y/x)$ for $x|y$.

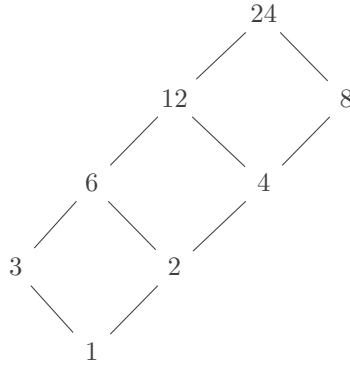
↳ As an alternative partial proof, note that when y/x is a power of k distinct primes, then the interval $[x, y]$ is isomorphic to the Boolean Algebra B_k with $x \mapsto \emptyset$ and $y \mapsto [k]$. This gives $\mu(x, y) = \mu(\emptyset, [k]) = (-1)^k$ by Lemma 12.36.

Proof. We prove this by induction on y/x . When $y/x = 1$, the claim is trivial since $\mu(x, y) = \mu(x, x) = 1 = \mu(1)$. Otherwise, let the distinct prime factors of $n = y/x$ be p_1, \dots, p_m . By the inductive hypothesis:

$$\begin{aligned} \mu(x, y) &= - \sum_{\substack{x|k|y \\ k \neq y}} \mu(x, k) = - \sum_{\substack{x|k|y \\ k \neq y}} \mu(k/x) = - \sum_{\substack{k|n \\ k \neq n}} \mu(k) = - \sum_{\substack{S \subset \{p_1, \dots, p_m\} \\ \prod S \neq n}} \mu\left(\prod S\right) \\ &= - \sum_{i=0}^m \binom{m}{i} (-1)^i + I \left[\prod_{i=1}^m p_i = n \right] (-1)^m = I \left[\prod_{i=1}^m p_i = n \right] (-1)^m = \mu(n) = \mu(y/x) \end{aligned}$$

Here, we used the fact that $\mu(k)$ is zero if it contains a square, so that it suffices to sum over products of distinct primes. We then used the fact that μ only depends on the number of such primes i . When $\prod_{i=1}^m p_i = n$, we have to add back the term corresponding to $i = m$ in the last summation, since we require $\prod S \neq n$. \square

Example 12.35. Consider the divisor lattice D_{24} given below:



For any divisor lattice, the order in \mathbb{N} provides a linear extension, so we use this here (e.g., $8 \in D_{24}$ corresponds to the 6th element in the linear extension). The matrix representation of ζ is:

$$\zeta \mapsto Z = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Inverting the matrix yields:

$$\mu \mapsto M = \begin{pmatrix} 1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Note that the first row matches what we would expect from using the number-theoretic Möbius function. This could also have been determined using the recursive definition of μ .

Lemma 12.36. Consider the Boolean Algebra poset B_n and let $S, T \in B_n$ with $T \leq S$. Then $\mu(T, S) = (-1)^{|S \setminus T|} = (-1)^{|S| - |T|}$.

Proof. We induct on $k = |S| - |T|$. $k = 0$ trivially gives $\mu(T, S) = \mu(T, T) = 1 = (-1)^0$. Otherwise, we have:

$$\begin{aligned} \mu(T, S) &= - \sum_{T \leq R < S} \mu(T, R) = - \sum_{T \leq R < S} (-1)^{|R| - |T|} = - \sum_{j=|T|}^{|S|-1} \binom{k}{j - |T|} (-1)^{j - |T|} \\ &= - \sum_{j=0}^{k-1} \binom{k}{j} (-1)^j = - \sum_{j=0}^k \binom{k}{j} (-1)^j + \binom{k}{k} (-1)^k = (-1)^k = (-1)^{|S| - |T|} \end{aligned}$$

Here, we conditioned in $j = |R|$, then reindexed the summation. \square

Lemma 12.37. The number of chains of length k that start at x and end at y is $(\zeta - \delta)^k(x, y)$. Using a linear extension, the number of chains of length k that start at x_i and end at x_j is given by the entries of $(Z - I)^k$.

↳ The matrix form is analogous to a transition matrix in stochastics, where we subtract I to prevent a chain from containing duplicates.

Proof. We induct on k . The claim is trivial when $k = 1$. For $k > 1$, let the chain be $x = x_1 < \dots < x_k = y$, and condition on x_{k-1} , which we call z . Form a chain of length $k - 1$ from x_1 to z $(\zeta - \delta)^{k-1}(x_1, z)$ ways, and form a chain of length 1 from z to x_k $(\zeta - \delta)(z, x_k)$ ways. The total number of chains is thus:

$$\sum_{z \in [x, y]} (\zeta - \delta)^{k-1}(x_1, x_k)(\zeta - \delta)(z, x_k) = (\zeta - \delta)^k(x_1, x_k) \quad \square$$

Proposition 12.38. Let P be any locally finite poset, let $x_i, x_j \in P$, and assume $x_i < x_j$. Then $\mu(x_i, x_j) = c_0 - c_1 + c_2 - c_3 + \dots$ where c_k is the number of chains of length k from x_i to x_j .

Proof. The number of chains of length k from x_i to x_j is given by the entries of $(Z - I)^k(x_i, x_j)$ by Lemma 12.37. Hence, the given summation corresponds to the matrix $\sum_{k=0}^{\infty} (I - Z)^k$. Use the geometric series formula to obtain $(I - (I - Z))^{-1} = Z^{-1} = M$. \square

Definition 12.39. Consider the covering matrix C of a poset P , where the rows and columns are indexed by the elements of P according to some linear extension. We let $C_{ij} = 1$ when x_j covers x_i and $C_{ij} = 0$ otherwise.

Proposition 12.40. The (i, j) -th entry of $(I - C)^{-1}$ is equal to the number of maximal chains of the interval $[x_i, x_j]$.

Proof. Since C is upper triangular with diagonal entries zero, it is nilpotent. Therefore, $(I - C)^{-1} = I + C + C^2 + \dots$, where the terms in the summation eventually become identically zero. The (i, j) th entry of C^k counts the number of chains from x_i to x_j of length k such that each successive element

covers the previous element. Such a chain is necessarily maximal in $[x_i, x_j]$ by definition of covering, so sum over all k to obtain the result. \square

Definition 12.41. Let P and Q be posets. Then their direct product $P \times Q$ is a poset whose elements (p, q) are ordered by $(p, q) \leq (p', q')$ if $p \leq p'$ and $q \leq q'$.

↳ Reflexivity, anti-symmetry, and transitivity follow immediately from the corresponding properties of P and Q .

Proposition 12.42. $\mu_{P \times Q}((p, q), (p', q')) = \mu_P(p, p')\mu_Q(q, q')$

Proof. If $p < p'$, then $0 = \sum_{p \leq z \leq p'} \mu_P(p, z)$ and $0 = \sum_{q \leq w \leq q'} \mu_Q(q, w) = 0$. We have:

$$\begin{aligned} \sum_{(z, w) \in [(p, q), (p', q')]} \mu_P(p, z)\mu_Q(q, w) &= \sum_{p \leq z \leq p'} \sum_{q \leq w \leq q'} \mu_P(p, z)\mu_Q(q, w) \\ &= \left(\sum_{p \leq z \leq p'} \mu_P(p, z) \right) \left(\sum_{q \leq w \leq q'} \mu_Q(q, w) \right) = 0 \end{aligned}$$

$\mu_P(p, p')\mu_Q(q, q') = 1$ when $p = p'$ and $q = q'$. Together, these two properties uniquely define $\mu_P \times \mu_Q$. Since these are the same definition properties of $\mu_{P \times Q}$, we have our claim. \square

Lemma 12.43. Let $x < y$ in Π_n under the usual refinement order. Suppose y has k blocks and each of those blocks is further partitioned into n_i blocks in x . Then $[x, y] \cong \prod_{i=1}^k \Pi_{n_i}$, so that $\mu(x, y) = \prod_{i=1}^k \mu(\Pi_{n_i})$. To compute this, we have $\mu(\Pi_n) = (n-1)!(-1)^{n-1}$.

Proof. To see that the isomorphism holds, the product holds by separating the elements in each block of y into separate elements in the tuple (in the direct product), then each block in x is treated as a single element. The second claim follows from the first by Proposition 12.42.

It remains to compute the value of $\mu(\Pi_n)$. For the latter, let $x \in \Pi_n$ be a set partition and let $|x|$ denote its number of blocks. The number of ways to color the blocks in x with q colors is $q^{|x|}$. Such a color induces a new partition $y \geq x$ where each block is formed by combining blocks in x with the same color. Given y , the number of colorings of it is $(q)_{|y|}$. Therefore, we have the equality $q^{|x|} = \sum_{y \geq x} (q)_{|y|}$.

Apply Möbius Inversion to obtain $(q)_{|x|} = \sum_{y \geq x} q^{|y|} \mu(x, y)$. Let $x = \hat{0}$ to obtain $(q)_n = \sum_{y \geq \hat{0}} q^{|y|} \mu(\hat{0}, y)$. There is only one partition with $|y| = 1$, which is $y = \hat{1}$. Hence:

$$\mu(\Pi_n) = \mu(\hat{0}, \hat{1}) = [q^1](q)_n = [q^1]q(q-1)_{n-1} = [q^0](q-1)_{n-1} = (0-1)_{n-1} = (n-1)!(-1)^{n-1}$$

\square

Example 12.44. Suppose P is the product of a k -element chain and an n -element chain. The result is isomorphic to a lattice, whose largest chain is length $n+k$. The largest antichain is $\min(n, k)$. Indeed, for a collection of incomparable $(x, y) \in P$, we must that all x and y are distinct. An antichain can then be formed by $(1, n), (2, n-1), \dots$.

12.3 Möbius Inversion

Theorem 12.45 (Möbius Inversion Formula). Let P be a poset such that each principal ideal is finite. Let $f, g : P \rightarrow \mathbb{F}$ be functions. Then:

$$g(y) = \sum_{x \leq y} f(x) \quad \text{if and only if} \quad f(y) = \sum_{x \leq y} g(x)\mu(x, y)$$

An equivalent form when every element in P is smaller than a finite number of elements is:

$$g(x) = \sum_{y \geq x} f(y) \quad \text{if and only if} \quad f(x) = \sum_{y \geq x} g(y)\mu(x, y)$$

Proof. The second form is equivalent to the first by forming a poset where the order is reversed, so it remains to prove the first claim. Assume LHS:

$$\sum_{x \leq y} g(x) \mu(x, y) = \sum_{x \leq y} \mu(x, y) \sum_{z \leq x} f(z) = \sum_{z \leq y} f(z) \sum_{z \leq x \leq y} \mu(x, y) = \sum_{z \leq y} f(z) \delta(x, y) = f(y)$$

The same logic works similarly in the reverse direction:

$$\sum_{x \leq y} f(x) = \sum_{x \leq y} \sum_{z \leq x} g(z) \mu(z, x) = \sum_{z \leq y} g(z) \sum_{z \leq x \leq y} \mu(z, x) = \sum_{z \leq y} g(z) \delta(z, y) = g(y)$$

Alternatively, we can use the matrix representation of the incidence algebra. Note that for any fixed y , it suffices to restrict the poset to the principal ideal of y , making it finite. Thus, we can form a linear extension x_1, x_2, \dots, x_n of P , define the rows vectors \mathbf{f}, \mathbf{g} with $f_i = f(x_i)$ and $g_i = g(x_i)$, and form the matrices $\zeta \mapsto Z$ and $\mu \mapsto M$. The equality $g(y) = \sum_{x \leq y} f(x)$ corresponds to $\mathbf{g} = \mathbf{f}Z$, while the equality $f(y) = \sum_{x \leq y} g(x) \mu(x, y)$ corresponds to $\mathbf{f} = \mathbf{g}M$. Hence, we have that these imply each other, since $M = Z^{-1}$. \square

Concept 12.46. Möbius Inversion is similar to a generalized version of the Principal of Inclusion/Exclusion. When using a poset that is a chain $[n]$, it is comparable to integration or differentiation (partial summation and first differences). Indeed, $g(y) = \sum_{x \leq y} f(x)$ calculates the partial sums, so that $f(y) = \sum_{x \leq y} g(x) \mu(x, y) = g(y) - g(y-1)$. On a Boolean algebra, it is equivalent to the Principal of Inclusion/Exclusion, as shown in Concept 4.3.

12.4 Number Theory

Proposition 12.47. We have $\{1\} \xleftrightarrow{\text{dgf}} \sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s)$ is the Riemann Zeta function.

Proposition 12.48. For any k , $\{n^k\} \xleftrightarrow{\text{dgf}} \sum_{n=1}^{\infty} \frac{1}{n^{s-k}} = \zeta(s-k)$.

Concept 12.49. By Wilf I and Proposition 12.47, $[n^{-s}] \zeta^2(s) = \sum_{d|n} 1 \cdot 1 = d(n)$, where $d(n)$ is the number of divisors of n . Hence, the coefficients in $\zeta^2(s)$ are $1, 2, 2, 3, 2, 4, \dots$. Note that the Riemann Zeta function is related to prime numbers, since $[n^{-s}] \zeta^2(s) = 2$ exactly when n is prime.

Concept 12.50. As an extension to Concept 12.49, Wilf II gives us that $[n^{-s}] \zeta^k(s)$ is the number of ways to express n as a product of k ordered factors. The number of ways to express n as a product of k ordered factors excluding 1 is $[n^{-s}] (\zeta(s) - 1)^k$.

Definition 12.51. A number-theoretic function is a function whose domain is the set of positive integers. Such a function is multiplicative if for any coprime m and n , we have $f(mn) = f(m)f(n)$.

\hookrightarrow A multiplicative function is uniquely determined by its action on primes and their powers

Proposition 12.52. Every multiplicative function f which is not identically zero satisfies $f(1) = 1$.

Proof. $\gcd(1, 1) = 1$, so that $f(1) \cdot f(1) = f(1 \cdot 1) = f(1)$. Hence, $f(1) = 0$ or $f(1) = 1$. If $f(1) = 0$, then $f(a) = f(a \cdot 1) = f(a)f(1) = 0$. \square

Proposition 12.53. If $f(n)$ is a multiplicative function, then so is $g(n) = \sum_{d|n} f(d)$.

Proof. Let $m, n \in \mathbb{N}$ be coprime. Then:

$$\begin{aligned} g(mn) &= \sum_{d|mn} f(d) = \sum_{d_1|m} \sum_{d_2|n} f(d_1 d_2) = \sum_{d_1|m} \sum_{d_2|n} f(d_1) f(d_2) \\ &= \left(\sum_{d_1|m} f(d_1) \right) \left(\sum_{d_2|n} f(d_2) \right) = g(m)g(n) \end{aligned} \quad \square$$

Theorem 12.54. Let f be a multiplicative function. Then:

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_{p \text{ prime}} \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \dots \right)$$

↳ For any integer n , there is a unique way to factor it into powers of primes. Hence, $[n^{-s}]f(s)$ can be uniquely represented as $f(p_1^{a_1})f(p_2^{a_2})\cdots$, so that summing over all n gives the above equation

Proposition 12.55. The constant function equal to 1 (whose Dirichlet series is the Riemann Zeta function) and $\mu(n)$ are multiplicative functions.

Proof. The constant function equal to 1 is trivially multiplicative, and μ being multiplicative follows immediately from definition. Indeed, for coprime m, n , we have:

- $\mu(mn) = 0$ iff mn is divisible by a square, which only occurs if m or n are divisible by a square, so that $\mu(m) = 0$ or $\mu(n) = 0$
- $\mu(mn) \neq 0$ implies $\mu(m) \neq 0$ and $\mu(n) \neq 0$ since otherwise mn would be divisible by a square. Hence, the number of prime factors in mn is the sum of the number of primes factors in m and n , which then gives $\mu(mn) = \mu(m)\mu(n)$

□

Concept 12.56. We can formulate Möbius Inversion for Dirichlet generating functions (effectively just using the divisor lattice as the poset). By Theorem 12.54, we can write $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{\text{prime } p} \frac{1}{1-p^{-s}}$. Similarly, since $\mu(p^a) = 0$ when $a > 1$, $\mu(p^a) = -1$ when $a = 1$, and $\mu(p^a) = 1$ when $a = 0$, we have:

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \prod_{\text{prime } p} (1 - p^{-s}) = \frac{1}{\zeta(s)}$$

Hence, $[n^{-s}] \frac{1}{\zeta(s)} = \mu(n)$. Suppose we have two sequences $\{a_n\}_{n \geq 1}, \{b_n\}_{n \geq 1}$ such that $a_n = \sum_{d|n} b_d$. Then by Wilf I for Dirichlet generating functions, we have $A(s) = B(s)\zeta(s)$, where $A(s)$ and $B(s)$ and the corresponding Dirichlet generating functions. Rearranging, we have $B(s) = A(s) \cdot \frac{1}{\zeta(s)}$, so that Wilf I gives $b_n = \sum_{d|n} a_d \mu(\frac{n}{d})$.

Example 12.57. Let $\sigma(n) = \sum_{d|n} d$ be the sum of the divisors of n . Then σ is multiplicative. It can be calculated from its prime factorization using $\sigma(p^k) = \frac{p^{k+1}-1}{p-1}$. Its Dirichlet generating function can be written as $\zeta(s)\zeta(s-1)$. Furthermore, Möbius Inversion gives $n = \sum_{d|n} \sigma(d)\mu(\frac{n}{d})$.

Proof. It is straightforward to prove that σ is multiplicative by definition, but it also immediately follows from Proposition 12.53 since the identity function is trivially multiplicative. The formula for $\sigma(p^k)$ follows from the geometric series formula, since its divisors are precisely powers of p .

One can prove the generating function using the form for $\sigma(p^k)$ and Theorem 12.54, but it is easier to prove it using Wilf I. Indeed, $\sigma(n) = \sum_{d|n} d \cdot 1$, where the generating function for $\{1\}$ is $\zeta(s)$ and the generating function for $\{n\}$ is $\zeta(s-1)$ by Proposition 12.48. Hence, $\{\sigma(n)\}$ has generating function $\zeta(s)\zeta(s-1)$. □

Example 12.58. Let $\Lambda(p^m) = \log p$ for prime p and $m \geq 1$, and $\Lambda(n) = 0$ otherwise. Then $\log n = \sum_{d|n} \Lambda(d)$, so that Möbius Inversion gives $\Lambda(n) = -\sum_{d|n} \mu(d) \log d$.

Proof. Write $n = p_1^{a_1} \cdots p_k^{a_k}$, so that $\log(n) = a_1 \log(p_1) + \cdots + a_k \log(p_k)$. For prime p_i , the number of its powers dividing n is a_i . Hence, we have $\log n = \sum_{d|n} \Lambda(d)$ as desired. Apply Möbius Inversion on the D_n to obtain:

$$\begin{aligned} \Lambda(n) &= \sum_{d|n} \log(d)\mu(n/d) = \sum_{d|n} \mu(d) \log(n/d) = \sum_{d|n} \mu(d)(\log n - \log d) \\ &= \log(n) \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log d = -\sum_{d|n} \mu(d) \log d \end{aligned}$$

□

Definition 12.59. The Euler Totient function $\phi(n)$ which counts the number of integers at most n which are coprime to n . We use $\phi([n])$ to denote the set of such integers.

Proposition 12.60. $\sum_{d|n} \phi(d) = n$

Proof. We prove this by forming a bijection $\Phi : [n] \rightarrow \underbrace{\{(d, k) \mid d \in [n], d \mid n, k \in \phi([d])\}}_K$. For $m \in [n]$, let $\frac{k}{d}$ be the reduced form of $\frac{m}{n}$. Then $(d, k) \in K$, so that Φ is well-defined. It is injective since $\frac{k_1}{d_1} = \frac{k_2}{d_2}$ implies $\frac{m_1}{n} = \frac{m_2}{n}$, and it is surjective since $kn/d \mapsto (d, k)$, where $kn/d \in [n]$ since $k \leq d$. We see that $|K| = \sum_{d|n} \phi(d)$, so that we have the claim. \square

Proposition 12.61. For a prime p , we have $\phi(p^k) = p^k - p^{k-1}$.

Proof. All integers in $\phi([p^k])^c$ must be divisible by p . Division by p yields a bijection from $\phi([p^k])^c$ to $[p^{k-1}]$, so that $|\phi([p^k])^c| = p^{k-1}$. Hence, $|[p^k]| - \phi(p^k) = p^{k-1}$. Rearrange to obtain $\phi(p^k) = p^k - p^{k-1}$. \square

Theorem 12.62. $\phi(n) = n \prod_{p|n} (1 - p^{-1})$ for primes p dividing n

Proof. $\phi([n])$ contains precisely the integers in $[n]$ which are divisible by none of the primes $p|n$. Hence, we can use the Inclusion/Exclusion. Let the primes dividing n be p_1, \dots, p_r , and let $N(p_{i_1}, \dots, p_{i_k})$ count the number of elements in $[n]$ divisible by at least the primes p_{i_1}, \dots, p_{i_k} . Since these are all distinct primes, $N(p_{i_1}, \dots, p_{i_k})$ is the number of elements in $[n]$ divisible by $p_{i_1} \cdots p_{i_k}$, which is $\frac{n}{p_{i_1} \cdots p_{i_k}}$. Hence, the Inclusion/Exclusion gives:

$$\begin{aligned} \phi(n) &= n - \sum_i \frac{n}{p_i} + \sum_{i < j} \frac{n}{p_i p_j} - \sum_{i < j < k} \frac{n}{p_i p_j p_k} + \cdots + (-1)^r \frac{n}{p_1 \cdots p_r} \\ &= n \left(1 - \sum_i \frac{1}{p_i} + \sum_{i < j} \frac{1}{p_i p_j} - \sum_{i < j < k} \frac{1}{p_i p_j p_k} + \cdots + (-1)^r \frac{1}{p_1 \cdots p_r} \right) \\ &= n \prod_{i=1}^r \left(1 - \frac{1}{p_i} \right) = n \prod_{p|n} (1 - p^{-1}) \end{aligned} \quad \square$$

Corollary 12.63. $\phi(n)$ is a multiplicative function. It can be calculated based on its factorization using Proposition 12.61. By Möbius Inversion, $\phi(n) = \sum_{d|n} d\mu(n/d) = n \sum_{d|n} \mu(d)/d$.

Proof. Theorem 12.62 immediately gives that $\phi(n)$ is multiplicative, since $\phi(nm) = nm \prod_{p|nm} (1 - p^{-1}) = n \prod_{p|n} (1 - p^{-1}) \cdot m \prod_{p|m} (1 - p^{-1})$ for coprime n, m . Concept 12.56 now gives us $\phi(n) = \sum_{d|n} d\mu(n/d)$, or switching $d \leftrightarrow n/d$, $\phi(n) = n \sum_{d|n} \mu(d)/d$. \square

Proposition 12.64. The Dirichlet generating function for $\phi(n)$ is $\zeta(s-1)/\zeta(s)$.

Proof. Corollary 12.63 gives $\phi(n) = \sum_{d|n} d\mu(n/d)$. The generating function for $\{n\}$ is $\zeta(s-1)$, and the generating function for $\{\mu(n)\}$ is $1/\zeta(s)$ by Concept 12.56. Hence, Wilf I gives us the generating function $\zeta(s-1)/\zeta(s)$. \square

Example 12.65. The function $|\mu(n)|$ is multiplicative, which we can view as being 1 when n is not divisible by a square greater than 1 and 0 otherwise. This is because $\mu(n)$ is multiplicative and $|mn| = |m| \cdot |n|$. We have $|\mu(p^k)| = I[k \leq 1]$ for prime p .

We prove that its Dirichlet generating function is $\zeta(s)/\zeta(2s)$. By Theorem 12.54 and Concept 12.56, we have:

$$\sum_{n=1}^{\infty} \frac{|\mu(n)|}{n^s} = \prod_{p \text{ prime}} (1 + p^{-s}) = \prod_{p \text{ prime}} \frac{1 - p^{-2s}}{1 - p^{-s}} = \frac{\prod_{p \text{ prime}} (1 - p^{-2s})}{\prod_{p \text{ prime}} (1 - p^{-s})} = \frac{\zeta(s)}{\zeta(2s)}$$

Theorem 12.66 (Basel Problem). $\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$

Proof. Begin with the geometric series formula to obtain:

$$1 + e^{it} + e^{2it} + \cdots = \frac{1}{1 - e^{it}} = \frac{1}{2} + \frac{1}{2}i \cot(t/2)$$

Take the real part to obtain $\frac{1}{2} + \cos(t) + \cos(2t) + \cdots = 0$ for all $0 < t < 2\pi$. Replace t by $t + \pi$ to obtain $\frac{1}{2} - \cos(t) + \cos(2t) - \cos(2t) + \cdots = 0$. Integrate from $t = 0$ to $t = \alpha$ to obtain $\sin(t) - \frac{1}{2}\sin(2t) + \frac{1}{3}\sin(3t) - \cdots = \frac{t}{2}$. This is conditionally convergent, so integrating again yields $1 - \cos t + \frac{1 - \cos(2t)}{2^2} + \frac{1 - \cos(3t)}{3^2} + \cdots = \frac{t^2}{4}$. Substitute $t = \pi$ to obtain $2 + \frac{2}{3^2} + \frac{2}{5^2} + \cdots = \frac{\pi^2}{4}$. Divide by 2 to obtain $1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}$. Then $\frac{\pi^2}{8} = \zeta(2) - \frac{1}{2^2} - \frac{1}{4^2} - \frac{1}{6^2} - \cdots$. This gives $\frac{\pi^2}{8} = \zeta(2) - \frac{1}{4}\zeta(2)$, so that $\zeta(2) = \frac{4}{3}\frac{\pi^2}{8} = \frac{\pi^2}{6}$. \square

Theorem 12.67 (Prime Number Theorem). Let $\pi(x)$ be the number of primes less than or equal to x . Then $\pi(x) \sim \frac{x}{\ln(x)}$, so that $\lim_{x \rightarrow \infty} \frac{x}{\ln(x)} = 1$.

\hookrightarrow This is equivalent to $\lim_{n \rightarrow \infty} \frac{M(n)}{n} = 0$ for $M(n) = \sum_{k=1}^n \mu(k)$

Definition 12.68. The Riemann Hypothesis is that $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ extended analytically to \mathcal{C} has its nontrivial zeros all on the line $\Re(s) = \frac{1}{2}$. The trivial zeros are $-2, -4, -6, \dots$

\hookrightarrow This is equivalent to $|M(n)| < n^{1/2+\epsilon}$ for all $\epsilon > 0$ and n sufficiently large.

Definition 12.69. Consider the partial sums $s_n = \sum_{k=1}^n a_k$. If $\sigma_n = \frac{1}{n} \sum_{k=1}^n s_k$ converges to L as $n \rightarrow \infty$, then $\{a_n\}$ is Cesaro summable.

Theorem 12.70. If $s_n \rightarrow L$ as $n \rightarrow \infty$, then $\sigma_n \rightarrow L$ as $n \rightarrow \infty$. In other words, convergence implies Cesaro summability, and the values to which s_n and σ_n converge are equivalent. Such summability methods are called regular.

Example 12.71. $\sum_{n=1}^{\infty} (-1)^{n+1}$ does not converge, but $\{(-1)^{n+1}\}$ is Cesaro summable and sums to $1/2$. Indeed, we have:

$$\sigma_n = \frac{1}{n} \sum_{k=1}^n s_k = \frac{1}{n} \sum_{k=1}^n \sum_{i=1}^k (-1)^{i+1} = \frac{1}{n} \sum_{k=1}^n I[k \bmod 2 = 1] = \frac{1}{n} \left\lceil \frac{n}{2} \right\rceil \rightarrow \frac{1}{2}$$

13. Combinatorial Reciprocity

Theorem 13.1. Suppose that $\{a_n\}_{n \in \mathbb{Z}}$ is a sequence satisfying a linear recurrence $a_{n+d} + c_1 a_{n+d-1} + c_2 a_{n+d-2} + \cdots + c_d a_n = 0$ for $c_1, \dots, c_d \in \mathcal{F}$ and $c_d \neq 0$. Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Then $\sum_{n=1}^{\infty} a_{-n} x^n = -f(1/x)$.

Technically, $1/x$ is not a formal power series. However, we can resolve this in a formal way by defining $f(1/x)$ as a symbol. By Theorem 7.33, the generating function $f(x)$ can be expressed as a rational function $\frac{p(x)}{q(x)}$ for coprime polynomials p, q with $\deg(p) < \deg(q) = d$. Then $x^d p(1/x)$ and $x^d q(1/x)$ are also polynomials, and $x^d q(1/x)$ has a nonzero constant term and thus has a reciprocal. Given this, we define $f(x) = \frac{x^d p(1/x)}{x^d q(1/x)}$.

Example 13.2. Consider $\{a_n\}_{n \geq 0}$ given by the recurrence $a_n = 3a_{n-1}$ with $a_0 = 2$. Then $a_n = 2 \cdot 3^n$. The ordinary generating function is $f(x) = \sum_{n=0}^{\infty} a_n x^n = \frac{2}{1-3x}$. Suppose we wished to extend this sequence to all of \mathbb{Z} , so that $a_n = 2 \cdot 3^n$ for all $n \leq 0$ as well. We can introduce a generating function $g(x) = \sum_{n=1}^{\infty} a_{-n} x^n = \frac{2 \cdot x/3}{1-x/3} = \frac{2x}{3-x}$. We observe that $g(x) = -f(1/x)$, as given by Theorem 13.1.

Example 13.3. Consider $f(x) = \frac{1}{(1-x)^n}$. Via Lemma 7.19, this equals $\sum_{k=0}^{\infty} \binom{n}{k} x^k = \sum_{k=0}^n \binom{n+k-1}{n-1} x^k$. By Theorem 13.1, we have $\sum_{k=1}^{\infty} \binom{n-k-1}{n-1} x^k = -f(1/x) = -\frac{1}{(1-1/x)^n} = \frac{(-1)^{n+1} x^n}{(1-x)^n}$.