Week 6-8: The Inclusion-Exclusion Principle

March 27, 2019

1 The Inclusion-Exclusion Principle

Let S be a finite set. Given subsets A, B, C of S, we have

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

$$\begin{aligned} |A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| \\ &\quad + |A \cap B \cap C|. \end{aligned}$$

Let $P_1, P_2, \ldots, P_n$ be properties referring to the objects in S. Let A_i denote the subset of S whose elements satisfy the property P_i, i.e.,

$$A_i = \{x \in S : x \text{ satisfies property } P_i\}, \quad 1 \le i \le n.$$

The elements of A_i may possibly satisfy some properties other than P_i. In many occasions we need to find the number of objects satisfying none of the properties $P_1, P_2, \ldots, P_n$.

Theorem 1.1. *The number of objects of S which satisfy none of the properties $P_1, P_2, \ldots, P_n$ is given by*

$$\begin{aligned} |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n| &= |S| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \sum_{i<j<k} |A_i \cap A_j \cap A_k| \\ &\quad + \cdots + (-1)^n |A_1 \cap A_2 \cap \cdots \cap A_n|. \end{aligned} \tag{1}$$

Proof. The left side of (1) counts the number of objects of S with none of the properties. We establish the identity (1) by showing that an object with none of the properties makes a net contribution of 1 to the right side of (1), and for an object with at least one of the properties makes a net contribution of 0.

Recall the indicator function 1_A of a subset $A \subseteq S$ is defined by $1_A(x) = 1$ if $x \in A$ and $1_A(x) = 0$ if $x \notin A$. We actually prove the following function identity:

$$1_{\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n} = 1_S - \sum_{k=1}^{n} (-1)^k \sum_{i_1 < \cdots < i_k} 1_{A_{i_1} \cap \cdots \cap A_{i_k}}.$$

Let x be an object satisfying none of the properties. Then the net contribution of x to the right side of (1) is

$$1 - 0 + 0 - 0 + \cdots + (-1)^n 0 = 1.$$

Let x be an object of S satisfying exactly r properties of $P_1, P_2, \ldots, P_n$, where $r > 0$. The net contribution of x to the right side of (1) is

$$\binom{r}{0} - \binom{r}{1} + \binom{r}{2} - \binom{r}{3} + \cdots + (-1)^r \binom{r}{r} = (1-1)^r = 0.$$

□

Corollary 1.2. *The number of objects of S which satisfy at least one of the properties $P_1, P_2, \ldots, P_n$ is given by*

$$\begin{aligned} |A_1 \cup A_2 \cup \cdots \cup A_n| &= \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| \\ &\quad - \cdots + (-1)^{n+1} |A_1 \cap A_2 \cap \cdots \cap A_n|. \end{aligned} \tag{2}$$

Proof. Note that the set $A_1 \cup A_2 \cup \cdots \cup A_n$ consists of all those objects in S which possess at least one of the properties, and

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = |S| - |\overline{A_1 \cup A_2 \cup \cdots \cup A_n}|.$$

Then by the DeMorgan law we have

$$\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n.$$

Thus

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = |S| - |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n|.$$

Putting this into the identity (1), the identity (2) follows immediately. □

2 Combinations with Repetition

Given a multiset M and fix an object x, whose repetition number is larger than r. Let M' be the multiset whose objects have the same repetition numbers as those objects in M, except that x repeats exactly r times. Then

$$\#\{r\text{-combinations of } M\} = \#\{r\text{-combinations of } M'\}.$$

Example 2.1. Determine the number of 10-combinations of the multiset

$$M' = \{3a, 4b, 5c\}.$$

Let S be the set of 10-combinations of the multiset $M = \{\infty a, \infty b, \infty c\}$. Let P_1, P_2, andP_3 be the properties that a 10-combination of M' has more than 3 a's, 4 b's, and 5 c's, respectively. Then the number of 10-combinations of M' is the number of 10-combinations of M which have none of the properties P_1, P_2, and P_3. Let A_i denote the sets consisting of the 10-combinations of M which have the property P_i, $1 \leq i \leq 3$. By the Inclusion-Exclusion Principle, the number to be determined is

$$\begin{aligned}|\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3| &= |S| - \big(|A_1| + |A_2| + |A_3|\big) + \big(|A_1 \cap A_2| + |A_1 \cap A_3| \\ &\quad + |A_2 \cap A_3|\big) - |A_1 \cap A_2 \cap A_3|.\end{aligned}$$

Note that

$$\begin{aligned}
|S| &= \left\langle {3 \atop 10} \right\rangle = \binom{3+10-1}{10} = \binom{12}{10} = 66,\\
|A_1| &= \left\langle {3 \atop 6} \right\rangle = \binom{3+6-1}{6} = \binom{8}{6} = 28,\\
|A_2| &= \left\langle {3 \atop 5} \right\rangle = \binom{3+5-1}{5} = \binom{7}{5} = 21,\\
|A_3| &= \left\langle {3 \atop 4} \right\rangle = \binom{3+4-1}{4} = \binom{6}{4} = 15,\\
|A_1 \cap A_2| &= \left\langle {3 \atop 1} \right\rangle = \binom{3+1-1}{1} = \binom{3}{1} = 3,\\
|A_1 \cap A_3| &= \left\langle {3 \atop 0} \right\rangle = \binom{3+0-1}{0} = \binom{2}{0} = 1,\\
|A_2 \cap A_3| &= 0,\\
|A_1 \cap A_2 \cap A_3| &= 0.
\end{aligned}$$

Putting all these results into the inclusion-exclusion formula, we have

$$|\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3| = 66 - (28 + 21 + 15) + (3 + 1 + 0) - 0 = 6.$$

The six 10-combinations are

$$\{3a, 4b, 3c\},\ \{3a, 3b, 4c\},\ \{3a, 2b, 5c\},\ \{2a, 4b, 4c\},\ \{2a, 3b, 5c\},\ \{a, 4b, 5c\}.$$

Example 2.2. Find the number of integral solutions of the equation

$$x_1 + x_2 + x_3 + x_4 = 15$$

which satisfy the conditions

$$2 \le x_1 \le 6, \quad -2 \le x_2 \le 1, \quad 0 \le x_3 \le 6, \quad 3 \le x_4 \le 8.$$

Let $y_1 = x_1 - 2$, $y_2 = x_2 + 2$, $y_3 = x_3$, and $y_4 = x_4 - 3$. Then the problem becomes to find the number of nonnegative integral solutions of the equation

$$y_1 + y_2 + y_3 + y_4 = 12$$

subject to

$$0 \le y_1 \le 4, \quad 0 \le y_2 \le 3, \quad 0 \le y_3 \le 6, \quad 0 \le y_4 \le 5.$$

Let S be the set of all nonnegative integral solutions of the equation $y_1 + y_2 + y_3 + y_4 = 12$. Let P_1 be the property that $y_1 \ge 5$, P_2 the property that $y_2 \ge 4$, P_3 the property that $y_3 \ge 7$, and P_4 the property that $y_4 \ge 6$. Let A_i denote the subset of S consisting of the solutions satisfying the property P_i, $1 \le i \le 4$. Then the problem is to find the cardinality $|\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3 \cap \bar{A}_4|$ by the inclusion-exclusion principle. In fact,

$$|S| = \left\langle {4 \atop 12} \right\rangle = \binom{4 + 12 - 1}{12} = \binom{15}{12} = 455.$$

Similarly,

$$\begin{aligned}
|A_1| &= \left\langle {4 \atop 7} \right\rangle = \binom{4 + 7 - 1}{7} = \binom{10}{7} = 120, \\
|A_2| &= \left\langle {4 \atop 8} \right\rangle = \binom{4 + 8 - 1}{8} = \binom{11}{8} = 165, \\
|A_3| &= \left\langle {4 \atop 5} \right\rangle = \binom{4 + 5 - 1}{5} = \binom{8}{5} = 56, \\
|A_4| &= \left\langle {4 \atop 6} \right\rangle = \binom{4 + 6 - 1}{6} = \binom{9}{6} = 84.
\end{aligned}$$

For the intersections of two sets, we have

$$|A_1 \cap A_2| = \left\langle {4 \atop 3} \right\rangle = \binom{4+3-1}{3} = \binom{6}{3} = 20,$$

$$|A_1 \cap A_3| = 1, \quad |A_1 \cap A_4| = |A_2 \cap A_3| = 4, \quad |A_2 \cap A_4| = 10, \quad |A_3 \cap A_4| = 0.$$

For the intersections of more sets,

$$\begin{aligned} |A_1 \cap A_2 \cap A_3| &= |A_1 \cap A_2 \cap A_4| = |A_1 \cap A_3 \cap A_4| \\ &= |A_2 \cap A_3 \cap A_4| = |A_1 \cap A_2 \cap A_3 \cap A_4| = 0. \end{aligned}$$

Thus the number required is given by

$$|\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3 \cap \bar{A}_4| = 455 - (120 + 165 + 56 + 84) + (20 + 1 + 4 + 4 + 10) = 69.$$

3 Derangements

A permutation of $\{1, 2, \ldots, n\}$ is called a **derangement** if every integer i $(1 \leq i \leq n)$ is not placed at the ith position. We denote by D_n the number of derangements of $\{1, 2, \ldots, n\}$.

Let S be the set of all permutations of $\{1, 2, \ldots, n\}$. Then $|S| = n!$. Let P_i be the property that a permutation of $\{1, 2, \ldots, n\}$ has the integer i in its ith position, and let A_i be the set of all permutations satisfying the property P_i, where $1 \leq i \leq n$. Then

$$D_n = |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n|.$$

For each $(i_1, i_2, \ldots, i_k)$ such that $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, a permutation of $\{1, 2, \ldots, n\}$ with $i_1, i_2, \ldots, i_k$ fixed at the i_1th, i_2th, ..., i_kth position respectively can be identified as a permutation of the set $\{1, 2, \ldots, n\} - \{i_1, i_2, \ldots, i_k\}$ of $n - k$ objects. Thus

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = (n-k)!.$$

By the inclusion-exclusion principle, we have

$$
\begin{aligned}
|\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n| &= |S| + \sum_{k=1}^{n} (-1)^k \sum_{i_1<i_2<\cdots<i_k} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| \\
&= n! + \sum_{k=1}^{n} (-1)^k \sum_{i_1<i_2<\cdots<i_k} (n-k)! \\
&= \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)! \\
&= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \simeq \frac{n!}{e} \quad \text{(when } n \text{ is large.)}
\end{aligned}
$$

Theorem 3.1. *For $n \geq 1$, the number D_n of derangements of $\{1, 2, \ldots, n\}$ is*

$$D_n = n! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} \right). \tag{3}$$

Here are a few derangement numbers:

$$D_0 \equiv 1, \quad D_1 = 0, \quad D_2 = 1, \quad D_3 = 2, \quad D_4 = 9, \quad D_5 = 44.$$

Corollary 3.2. *The number of permutations of $\{1, 2, \ldots, n\}$ with exactly k numbers displaced is*

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

Proposition 3.3. *The derangement sequence D_n satisfies the recurrence relation*

$$D_n = (n-1)(D_{n-1} + D_{n-2}), \quad n \geq 3$$

with the initial condition $D_1 = 0, D_2 = 1$. The sequence D_n satisfies the recurrence relation

$$D_n = nD_{n-1} + (-1)^n, \quad n \geq 2.$$

Proof. The recurrence relations can be proved without using the formula (3). Let S_k denote the set of derangements of $\{1, 2, \ldots, n\}$ having the pattern $ka_2a_3 \cdots a_n$, where $k = 2, 3, \ldots, n$. We may think of $a_2a_3 \ldots a_n$ as a permutation of $\{2, \ldots, k-1, 1, k+1, \ldots, n\}$ with respect to the order

$$23 \cdots (k-1)1(k+1) \cdots n.$$

The derangements of S_k can be partitioned into two types:

$$ka_2a_3\cdots a_k\cdots a_n\ (a_k\neq 1)\quad\text{and}\quad ka_2a_3\cdots a_{k-1}1a_{k+1}\cdots a_n.$$

The first type can be considered as permutations of $k23\ldots(k-1)1(k+1)\ldots n$ such that the first member is fixed and no one is placed in its original place for other members. The number of such permutations is D_{n-1}. The second type can be considered as permutations of $k23\ldots(k-1)1(k+1)\ldots n$ such that the first and the kth members are fixed, and no one is placed in its original place for other members. The number of such permutations is D_{n-2}. We thus obtain the recurrence relation

$$D_n=(n-1)(D_{n-1}+D_{n-2}),\quad n\geq 3.$$

Let us rewrite the recurrence relation as

$$D_n-nD_{n-1}=-\big(D_{n-1}-(n-1)D_{n-2}\big),\quad n\geq 3.$$

Applying this recurrence relation continuously, we have

$$D_n-nD_{n-1}=(-1)^i\big(D_{n-i}-(n-i)D_{n-i-1}\big),\quad 1\leq i\leq n-2.$$

Thus $D_n-nD_{n-1}=(-1)^{n-2}(D_2-D_1)=(-1)^n$. Hence $D_n=nD_{n-1}+(-1)^n$. □

4 Surjective Functions

Let X be a set of m objects and Y a set of n objects. Then the number of functions of X to Y is n^m. The number of injective functions from X to Y is

$$\binom{n}{m}m!=P(n,m).$$

Let $C(m,n)$ denote the number of surjective functions from X to Y. What is $C(m,n)$?

Theorem 4.1. *The number $C(m,n)$ of surjective functions from a set of m objects to a set of n objects is given by*

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

Proof. Let S be the set of all functions of X to Y. Write $Y = \{y_1, y_2, \ldots, y_n\}$. Let A_i be the set of all functions f such that y_i is not assigned to any element of X by f, i.e., $y_i \notin f(X)$, where $1 \le i \le n$. Then

$$C(m,n) = |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n|.$$

For each $(i_1, i_2, \ldots, i_k)$ such that $1 \le i_1 < i_2 < \cdots < i_k \le n$, the intersection

$$A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}$$

can be identified to the set of functions f from X to the set $Y \setminus \{y_{i_1}, y_{i_2}, \ldots, y_{i_k}\}$. Thus

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = (n-k)^m.$$

By the Inclusion-Exclusion Principle, we have

$$\begin{aligned} |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n| &= |S| + \sum_{k=1}^{n} (-1)^k \sum_{i_1 < i_2 < \cdots < i_k} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| \\ &= n^m + \sum_{k=1}^{n} (-1)^k \sum_{i_1 < i_2 < \cdots < i_k} (n-k)^m \\ &= \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)^m. \end{aligned}$$

□

Note that $C(m,n) = 0$ for $m < n$; we have

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)^m = 0 \quad \text{if} \quad m < n.$$

Corollary 4.2. *For integers $m, n \ge 1$,*

$$\sum_{\substack{i_1 + \cdots + i_n = m \\ i_1, \ldots, i_n \ge 1}} \binom{m}{i_1, \ldots, i_n} = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)^m.$$

Proof. The integer $C(m,n)$ can be interpreted as the number of ways to place objects of X into n distinct boxes so that no box is empty. Let the 1st box be placed i_1 objects, …, the nth box be placed i_n objects; then $i_1 + \cdots + i_n = m$.

The number of placements of X into n distinct boxes, such that the 1st box contains exactly i_1 objects, ..., the nth box contains exactly i_n objects, is $\frac{m!}{i_1!\cdots i_n!}$, which is the multinomial coefficient $\binom{m}{i_1,\ldots,i_n}$. We thus have

$$C(m,n) = \sum_{\substack{i_1+\cdots+i_n=m \\ i_1,\ldots,i_n\geq 1}} \binom{m}{i_1,\ldots,i_n}.$$

□

5 Euler Totient Function

Let n be a positive integer. We denote by $\phi(n)$ the number of integers of $[1,n]$ which are coprime to n, i.e., $\phi(n) = |\{k \in [1,n] : \gcd(k,n) = 1\}|$. For example,

$$\phi(1) = 1, \quad \phi(2) = 1, \quad \phi(3) = 2, \quad \phi(4) = 2, \quad \phi(5) = 4, \quad \phi(6) = 2.$$

The integer-valued function ϕ is defined on the set of positive integers, called the **Euler phi (totient) function**.

Theorem 5.1. *Let n be a positive integer factorized into the form*

$$n = p_1^{e_1} p_2^{e_r} \cdots p_r^{e_r},$$

where $p_1, p_2, \ldots, p_r$ are distinct primes and $e_1, e_2, \ldots, e_r \geq 1$. Then

$$\phi(n) = n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right).$$

Proof. Let $S = \{1, 2, \ldots, n\}$. Let P_i be the property of integers in S having factor p_i, and let A_i be the set of integers in S that satisfy the property P_i, where $1 \leq i \leq r$. Then $\phi(n)$ is the number of integers satisfying none of the properties $P_1, P_2, \ldots, P_r$, i.e.,

$$\phi(n) = |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_r|.$$

Note that

$$A_i = \left\{1p_i, 2p_i, \ldots, \left(\frac{n}{p_i}\right) p_i\right\}, \quad 1 \leq i \leq r.$$

Likewise, for $q = p_{i_1} p_{i_2} \cdots p_{i_k}$ with $1 \le i_1 < i_2 < \cdots < i_k \le r$,

$$A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k} = \left\{ 1q, 2q, \ldots, \left(\frac{n}{q}\right) q \right\}.$$

Thus

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = \frac{n}{q} = \frac{n}{p_{i_1} p_{i_2} \cdots p_{i_k}}.$$

By the Inclusion-Exclusion Principle, we have

$$\begin{aligned}
|\bar{A}_1 \cap \cdots \cap \bar{A}_r| &= |S| + \sum_{k=1}^{r} (-1)^k \sum_{i_1 < \cdots < i_k} |A_{i_1} \cap \cdots \cap A_{i_k}| \\
&= n + \sum_{k=1}^{r} (-1)^k \sum_{i_1 < i_2 < \cdots < i_k} \frac{n}{p_{i_1} p_{i_2} \cdots p_{i_k}} \\
&= n \left[1 - \left(\frac{1}{p_1} + \cdots + \frac{1}{p_r} \right) \right. \\
&\quad + \left(\frac{1}{p_1 p_2} + \frac{1}{p_1 p_3} + \cdots + \frac{1}{p_{r-1} p_r} \right) \\
&\quad - \left(\frac{1}{p_1 p_2 p_3} + \frac{1}{p_1 p_2 p_4} + \cdots + \frac{1}{p_{r-2} p_{r-1} p_r} \right) \\
&\quad \left. + \cdots + (-1)^r \frac{1}{p_1 p_2 \cdots p_r} \right] \\
&= n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i} \right).
\end{aligned}$$

□

Example 5.1. For the integer 36 $(= 2^2 3^2)$, we have

$$\phi(36) = 36 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) = 12.$$

The following are the twelve specific integers of $[1, 36]$ that are coprime to 36:

$$1,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 25,\ 29,\ 31,\ 35.$$

Corollary 5.2. *For any prime number p,*

$$\phi(p^k) = p^k - p^{k-1}.$$

Proof. The result can be directly proved without Theorem 5.1. The set $[1, p^k]$ has p^{k-1} integers $1p, 2p, \ldots p^{k-1}p$ not coprime to p^k. Thus $\phi(p^k) = p^k - p^{k-1}$. □

Lemma 5.3. *Let* $m = m_1m_2$. *If* $\gcd(m_1, m_2) = 1$, *then we have*

(i) *The function* $f : [m] \to [m_1] \times [m_2]$ *defined by* $f(a) = (r_1, r_2)$, *where*

$$a = q_1m_1 + r_1 = q_2m_2 + r_2 \in [m], \quad 1 \leq r_1 \leq m_1, \quad 1 \leq r_2 \leq m_2,$$

is a bijection.

(ii) *The restriction of* f *to* $\{a \in [m] : \gcd(a, m) = 1\}$ *is a map to the product set*

$$\big\{a \in [m_1] : \gcd(a, m_1) = 1\big\} \times \big\{a \in [m_2] : \gcd(a, m_2) = 1\big\},$$

and is also a bijection.

Proof. (i) It suffices to show that f is surjective. Since $\gcd(m_1, m_2) = 1$, by the Euclidean Algorithm there exist integers x and y such that $xm_1 + ym_2 = 1$.

For each $(r_1, r_2) \in [m_1] \times [m_2]$, the integer $r := r_2xm_1 + r_1ym_2$ can be written as

$$r = (r_2 - r_1)xm_1 + r_1(xm_1 + ym_2) = (r_1 - r_2)ym_2 + r_2(xm_1 + ym_2).$$

Since $xm_1 + ym_2 = 1$, we have

$$r = (r_2 - r_1)xm_1 + r_1 = (r_1 - r_2)ym_2 + r_2.$$

We modify r by adding an appropriate multiple qm of m to obtain

$$a := qm + r \quad \text{such that} \quad 1 \leq a \leq m.$$

Then $a = q_1m_1 + r_1 = q_2m_2 + r_2 \in [m]$ for some integers q_1 and q_2. We thus have $f(a) = (r_1, r_2)$. This shows that f is surjective. Since both $[m]$ and $[m_1] \times [m_2]$ have the same cardinality m_1m_2, it follows that f must be a bijection.

(ii) It follows from the fact that an integer $a \in [m_1m_2]$ is coprime to m_1m_2 iff a is coprime to m_1 and coprime to m_2. □

Theorem 5.4. *For positive integers* m *and* n *such that* $\gcd(m, n) = 1$,

$$\phi(mn) = \phi(m)\phi(n).$$

If $n = p_1^{e_1} \cdots p_r^{e_r}$ *with* $e_1, \ldots, e_r \geq 1$*, where* $p_1, \ldots, p_r$ *are distinct primes, then*

$$\phi(n) = n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right).$$

Proof. The first part follows from Lemma 5.3. Note that $[p_i^{e_i}]$ has $p_i^{e_i-1}$ integers $1p_i, 2p_i, \ldots, p_i^{e_i-1}p_i$ not coprime to $p_i^{e_i}$. So $\phi(p^{e_i}) = p^{e_i} - p^{e_i-1}$. The second part follows from the first part, i.e.,

$$\begin{aligned} \phi(n) &= \prod_{i=1}^{r} \phi(p_i^{e_i}) = \prod_{i=1}^{r} \left(p_i^{e_i} - p_i^{e_i-1}\right) \\ &= \prod_{i=1}^{r} p_i^{e_i} \left(1 - \frac{1}{p_i}\right) = n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right). \end{aligned}$$

□

6 Permutations with Forbidden Positions

Let $X_1, X_2, \ldots, X_n$ be subsets (possibly empty) of $\{1, 2, \ldots, n\}$. We denote by $P(X_1, X_2, \ldots, X_n)$ the set of all permutations $a_1 a_2 \cdots a_n$ of $\{1, 2, \ldots, n\}$ such that

$$a_1 \notin X_1, \quad a_2 \notin X_2, \quad \ldots, \quad a_n \notin X_n.$$

In other words, a permutation of S belongs to $P(X_1, X_2, \ldots, X_n)$ provided that no members of X_1 occupy the first place, no members of X_2 occupy the second place, ..., and no members of X_n occupy the nth place. Let

$$p(X_1, X_2, \ldots, X_n) = |P(X_1, X_2, \ldots, X_n)|.$$

It is known that there is a one-to-one correspondence between permutations of $\{1, 2, \ldots, n\}$ and the placement of n non-attacking indistinguishable rooks on an n-by-n board. The permutation $a_1 a_2 \cdots a_n$ of $\{1, 2, \ldots, n\}$ corresponds to the placement of n rooks on the board in the squares with coordinates

$$(1, a_1), \quad (2, a_2), \quad \ldots, \quad (n, a_n).$$

The permutations in $P(X_1, X_2, \ldots, X_n)$ corresponds to placements of n non-attacking rooks on an n-by-n board in which certain squares are not allowed to be put a rook.

Let S be the set of all placements of n non-attacking rooks on an $n \times n$-board. A rook placement in S is said to satisfy the **property** P_i provided that the rook in the ith row having column index in X_i, where $1 \le i \le n$. Let A_i be the set of rook placements satisfying the property P_i. Then by the Inclusion-Exclusion Principle,

$$\begin{aligned} p(X_1, X_2, \ldots, X_n) &= |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n| \\ &= |S| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \cdots \\ &\quad \cdots + (-1)^n |A_1 \cap A_2 \cap \cdots \cap A_n|. \end{aligned}$$

Proposition 6.1. *Let r_k $(1 \le k \le n)$ denote the number of ways to place k non-attacking rooks on an $n \times n$-board where each of the k rooks is in a forbidden position. Then*

$$r_k = \frac{1}{(n-k)!} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|. \tag{4}$$

Proof. Fix $(i_1, i_2, \ldots, i_k)$ with $1 \le i_1 < i_2 < \cdots < i_k \le n$. Let $r(i_1, i_2, \ldots, i_k)$ denote the number of ways to place k non-attacking rooks such that

- the rook on the i_1th row has column index in X_{i_1},
- the rook on the i_2th row has column index in X_{i_2}, ..., and
- the rook on the i_kth row has column index in X_{i_k}.

For each such k rook arrangement, delete the i_1th row, i_2th row, ..., i_kth row, and delete the columns where the i_1th, or i_2th, ..., or i_kth position is arranged a rook; the other $n-k$ rooks cannot be arranged in the deleted rows and columns. The leftover is an $(n-k) \times (n-k)$-board, and the other $n-k$ rooks can be arranged in $(n-k)!$ ways. So

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = r(i_1, i_2, \ldots, i_k)\,(n-k)!.$$

Since $r_k = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} r(i_1, i_2, \ldots, i_k)$, it follows that

$$r_k(n-k)! = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|.$$

□

Theorem 6.2. *The number of ways to place n non-attacking rooks on an $n \times n$-board with forbidden positions is given by*

$$p(X_1, X_2, \ldots, X_n) = \sum_{k=0}^{n} (-1)^k r_k (n-k)!,$$

where r_k is the number of ways to place k non-attacking rooks on an $n \times n$-board where each of the k rooks is in a forbidden position.

Example 6.1. Let $n = 5$ and $X_1 = \{1, 2\}$, $X_2 = \{3, 4\}$, $X_3 = \{1, 5\}$, $X_4 = \{2, 3\}$, and $X_5 = \{4, 5\}$.

×	×			
		×	×	
×				×
	×	×		
			×	×

Find the number of rook placements with the given forbidden positions.

Solution. Note that $r_0 = 1$. It is easy to see that

$$r_1 = 5 \times 2 = 10.$$

Since $r_1 = \frac{1}{4!} \sum_i |A_i|$, we have

$$\sum_i |A_i| = r_1 4! = 10 \cdot 4!. \quad \text{(This is not needed.)}$$

Since

$$|A_1 \cap A_2| = |A_2 \cap A_3| = |A_3 \cap A_4| = |A_4 \cap A_5| = |A_1 \cap A_5| = 4 \cdot 3!,$$

$$|A_1 \cap A_3| = |A_1 \cap A_4| = |A_2 \cap A_4| = |A_2 \cap A_5| = |A_3 \cap A_5| = 3 \cdot 3!,$$

we see that

$$r_2 = \frac{1}{3!}\sum_{i<j} |A_i \cap A_j| = 5 \times 4 + 5 \times 3 = 35.$$

Using the symmetry between A_1, A_2, A_3, A_4, A_5 and A_5, A_4, A_3, A_2, A_1 respectively, we see that

$$\begin{aligned} |A_1 \cap A_2 \cap A_3| &= |A_1 \cap A_2 \cap A_5| = |A_1 \cap A_4 \cap A_5| \\ &= |A_2 \cap A_3 \cap A_4| = |A_3 \cap A_4 \cap A_5| \\ &= 6 \cdot 2!, \\ |A_1 \cap A_2 \cap A_4| &= |A_1 \cap A_3 \cap A_4| = |A_1 \cap A_3 \cap A_5| \\ &= |A_2 \cap A_3 \cap A_5| = |A_2 \cap A_4 \cap A_5| \\ &= 4 \cdot 2!. \end{aligned}$$

These can be obtained by considering the following six patterns:

$\times$	$\times$			
		$\times$	$\times$	
$\times$				$\times$

$\times$	$\times$			
		$\times$	$\times$	
	$\times$	$\times$		

$\times$	$\times$			
		$\times$	$\times$	
			$\times$	$\times$

$\times$	$\times$			
$\times$				$\times$
	$\times$	$\times$		

$\times$	$\times$			
$\times$				$\times$
			$\times$	$\times$

		$\times$	$\times$	
$\times$				$\times$
	$\times$	$\times$		

We then have

$$r_3 = 5 \cdot 6 + 5 \cdot 4 = 50.$$

Using the symmetric position again, we see that

$$\begin{aligned} |A_1 \cap A_2 \cap A_3 \cap A_4| &= |A_1 \cap A_2 \cap A_3 \cap A_5| = |A_1 \cap A_2 \cap A_4 \cap A_5| \\ &= |A_1 \cap A_3 \cap A_4 \cap A_5| = |A_2 \cap A_3 \cap A_4 \cap A_4| \\ &= 5 \cdot 1!. \end{aligned}$$

Thus

$$r_4 = 5 \times 5 = 25.$$

Finally,

$$r_5 = |A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5| = 2.$$

The answer $\sum_{k=0}^{5}(-1)^k r_k(5-k)!$ is

$$5! - 10 \times 4! + 35 \times 3! - 50 \times 2! + 25 \times 1! - 2 = 13.$$

A permutation of $\{1, 2, \ldots, n\}$ is **nonconsecutive** if 12, 23, ..., $(n-1)n$ do not occur. We denote by Q_n the number of nonconsecutive permutations of $\{1, 2, \ldots, n\}$. We have $Q_1 = 1$, $Q_2 = 1$, $Q_3 = 3$, $Q_4 = 13$.

Theorem 6.3. *For $n \geq 1$,*

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

Proof. Let S be the set of permutations of $\{1, 2, \ldots, n\}$. Let P_i be the property that in a permutation the pattern $i(i+1)$ does occur, where $1 \leq i \leq n-1$. Let A_i be the set of all permutations satisfying the property P_i. Then Q_n is the number of permutations satisfying none of the properties $P_1, \ldots, P_{n-1}$, i.e.,

$$Q_n = |\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_{n-1}|.$$

Note that

$$|A_i| = (n-1)!, \quad 1 \leq i \leq n-1.$$

Similarly,

$$|A_i \cap A_j| = (n-2)!, \quad 1 \leq i < j \leq n-1.$$

More generally,

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

Thus by the Inclusion-Exclusion Principle,

$$\begin{aligned} Q_n &= |S| + \sum_{k=1}^{n-1}(-1)^k \sum_{1 \leq i_1 < \cdots < i_k \leq n-1} |A_{i_1} \cap \cdots \cap A_{i_k}| \\ &= \sum_{k=0}^{n-1}(-1)^k \binom{n-1}{k}(n-k)!. \end{aligned}$$

□

Example 6.2. Eight persons line up in one column in such a way that every person except the first one has a person in front. What is the chance when the eight persons reline up after a break so that everyone has a different person in his/her front?

We assign numbers $1, 2, \ldots, 8$ to the eight persons so that the number i is assigned to the ith person (counted from the front). The problem is then to find the number of permutations of $\{1, 2, \ldots, 8\}$ in which the patterns 12, 23, ..., 78 do not occur. For instance, 31542876 is an allowed permutation, while 83475126 is not. The answer is given by

$$P = \frac{Q_8}{8!} = \sum_{k=0}^{7}(-1)^k \binom{7}{k} \frac{(8-k)!}{8!} \approx 0.413864.$$

Example 6.3. There are n persons seated at a round table. The n persons left the table and reseat after a break. How many seating plans can be made in the second time so that each person has a different person seating on his/her left comparing to the person before the break?

This is equivalent to finding the number of circular nonconsecutive permutations of $\{1, 2, \ldots, n\}$. A **circular nonconsecutive** permutation of $\{1, 2, \ldots, n\}$ is a circular permutation of $\{1, 2, \ldots, n\}$ such that 12, 23, ..., $(n-1)n$, $n1$ do not occur in the counterclockwise direction.

Let S be the set of all circular permutations of $\{1, 2, \ldots, n\}$. Let A_i denote the subset of all circular permutations of $\{1, 2, \ldots, n\}$ such that $i(i+1)$ does not occur, $1 \le i \le n$. We understand that A_n is the subset of all circular permutations that $n1$ does not occur. The answer is

$$|\bar{A}_1 \cap \bar{A}_1 \cap \cdots \cap \bar{A}_n|.$$

Note that $|S| = (n-1)!$, and

$$|A_i| = (n-1)!/(n-1) = (n-2)!.$$

More generally,

$$|A_{i_1} \cap \cdots \cap A_{i_k}| = (n-k)!/(n-k) = (n-k-1)!, \quad 1 \le k \le n-1;$$

$$|A_1 \cap A_2 \cap \cdots \cap A_n| = 1.$$

We thus have

$$|\bar{A}_1 \cap \cdots \cap \bar{A}_n| = \sum_{k=0}^{n-1} (-1)^k \binom{n}{k} (n-k-1)! + (-1)^n.$$

Theorem 6.4.

$$Q_n = D_n + D_{n-1}, \quad n \geq 2.$$

Proof.

$$\begin{aligned}
D_n + D_{n-1} &= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} + (n-1)! \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} \\
&= (n-1)! \left(n + n \sum_{k=1}^{n} \frac{(-1)^k}{k!} + \sum_{k=1}^{n} \frac{(-1)^{k-1}}{(k-1)!} \right) \\
&= n! + (n-1)! \sum_{k=1}^{n} \frac{(-1)^k}{k!} (n-k) \\
&= n! + \sum_{k=1}^{n-1} (-1)^k \binom{n-1}{k} (n-k)! \\
&= \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (n-k)! = Q_n.
\end{aligned}$$

□

7 Rook Polynomials

Definition 7.1. Let C be a board; each square of C is referred as a **cell**. Let $r_k(C)$ denote the number of ways to arrange k rooks on the board C so that no one can take another. We assume $r_0(C) = 1$. The **rook polynomial** of C is

$$R(C, x) = \sum_{k=0}^{\infty} r_k(C) x^k.$$

A k**-rook arrangement** on the board C is an arrangement of k rooks on C.

Proposition 7.2. *Given a board C. For each cell σ of C, let $C-\sigma$ denote the board obtained from C by deleting the cell σ, and let C_σ denote the board obtained from C by deleting all cells on the row and column that contains the cell σ. Then*

$$r_k(C) = r_k(C-\sigma) + r_{k-1}(C_\sigma).$$

Equivalently,

$$R(C,x) = R(C-\sigma, x) + xR(C_\sigma, x).$$

Proof. The k-rook arrangements on the board C can be divided into two kinds: the rook arrangements that the square σ is occupied and the rook arrangements that the square is not occupied, i.e., the k-rook arrangements on the board $C-\sigma$ and the $(k-1)$-rook arrangements on the board C_σ. We thus have $r_k(C) = r_k(C-\sigma) + r_{k-1}(C_\sigma)$. □

Two boards C_1 and C_2 are said to be **independent** if they have no common rows and common columns. Independent boards must be disjoint. If C_1 and C_2 are independent boards, we denote by $C_1 + C_2$ the board that consists of the cells either in C_1 or in C_2, i.e., the union of cells.

Proposition 7.3. *Let C_1 and C_2 be independent boards. Then*

$$r_k(C_1 + C_2) = \sum_{i=0}^{k} r_i(C_1)\, r_{k-i}(C_2),$$

where $C_1 + C_2 = C_1 \cup C_2$. Equivalently,

$$R(C_1 + C_2, x) = R(C_1, x)R(C_2, x).$$

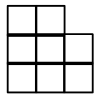
Proof. Since C_1 and C_2 have disjoint rows and columns, each i-rook arrangement of C_1 and each j-rook arrangement of C_2 will constitute a $(i+j)$-rook arrangement of $C_1 + C_2$, and vice versa. Thus

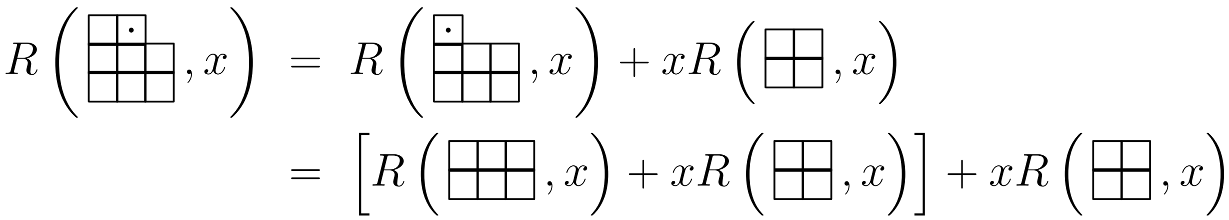
$$r_k(C_1 + C_2) = \sum_{\substack{i+j=k \\ i,j\geq 0}} r_i(C_1) r_j(C_2).$$

□

Example 7.1. The rook polynomial of an m-by-n board C with $m \leq n$,

$$R(C,x) = \sum_{k=0}^{m} \binom{m}{k} \binom{n}{k} k! x^k.$$

Example 7.2. Find the rook polynomial of the board . We use $\boxdot$ (a square with a dot) to denote a selected square when applying the recurrence formula of rook polynomial.



$$= \left(1 + 6x + 3 \cdot 2x^2\right) + 2x\left(1 + 4x + 2x^2\right)$$
$$= 1 + 8x + 14x^2 + 4x^3.$$

8 Weighted Version of Inclusion-Exclusion Principle

Let X be a set, either finite or infinite. The **indicator function** of a subset A of X is a real-valued function 1_A on X, defined by

$$1_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

For real-valued functions f, g, and a real number c, we define functions $f+g$, cf, and fg on X as follows:

$$(f+g)(x) = f(x) + g(x),$$

$$(cf)(x) = cf(x),$$

$$(fg)(x) = f(x)g(x).$$

For subsets $A, B \subseteq X$ and arbitrary function f on X, it is easy to verify the following properties:

(i) $1_{A\cap B} = 1_A 1_B$,

(ii) $1_{\bar{A}} = 1_X - 1_A$,

(iii) $1_{A\cup B} = 1_A + 1_B - 1_{A\cap B}$,

(iv) $1_X f = f$.

The set of all real-valued functions on X is a vector space over $\mathbb{R}$, and is further a commutative algebra with identity 1_X.

Given a function $w : X \to \mathbb{R}$, usually referred to a **weight function** on X, such that w is nonzero at only finitely many elements of X; the value $w(x)$ is called the **weight** of x. For each subset $A \subseteq X$, the **weight** of A is

$$w(A) = \sum_{x\in A} w(x).$$

If $A = \emptyset$, we assume $w(\emptyset) = 0$. For each function $f : X \to \mathbb{R}$, the **weight** of f is

$$w(f) = \sum_{x\in X} w(x)f(x) = \langle w, f\rangle.$$

Clearly, $w(1_A) = w(A)$. For functions f_i and constants c_i ($1 \le i \le m$), we have

$$w\left(\sum_{i=1}^{m} c_i f_i\right) = \sum_{i=1}^{m} c_i\, w(f_i).$$

This means that w is a linear functional on the vector space of all real-valued functions on X.

Proposition 8.1. *Let $P_1, \ldots, P_n$ be some properties about the elements of a set X. Let A_i denote the set of elements of X that satisfy the property P_i, $1 \le i \le n$. Given a weight function w on X. Then the Inclusion-Exclusion Principle can be stated as*

$$1_{\bar{A}_1\cap\bar{A}_2\cap\cdots\cap\bar{A}_n} = 1_X + \sum_{k=1}^{n}(-1)^k \sum_{1\le i_1<i_2<\cdots<i_k\le n} 1_{A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}}; \tag{5}$$

$$w\left(\bar{A}_1\cap\cdots\cap\bar{A}_n\right) = w(X) + \sum_{k=1}^{n}(-1)^k \sum_{i_1<\cdots<i_k} w\big(A_{i_1}\cap\cdots\cap A_{i_k}\big). \tag{6}$$

Proof. Applying properties about indicator functions,

$$\begin{aligned}
1_{\bar{A}_1\cap\cdots\cap\bar{A}_n} &= 1_{\bar{A}_1}\cdots 1_{\bar{A}_n} = (1_X - 1_{A_1})\cdots(1_X - 1_{A_n}) \\
&= \sum f_1\cdots f_n \quad (f_i = 1_X \text{ or } f_i = -1_{A_i}, 1 \le i \le n) \\
&= \underbrace{1_X\cdots 1_X}_{n} + \sum_{k=1}^{n}\sum_{i_1<\cdots<i_k}\underbrace{1_X\cdots 1_X}_{n-k}(-1_{A_{i_1}})\cdots(-1_{A_{i_k}}) \\
&= 1_X + \sum_{k=1}^{n}(-1)^k \sum_{1\le i_1<\cdots<i_k\le n} 1_{A_{i_1}\cap\cdots\cap A_{i_k}}.
\end{aligned}$$

Applying weight w to both sides, we obtain

$$w\left(\bar{A}_1\cap\cdots\cap\bar{A}_n\right) = w(X) + \sum_{k=1}^{n}(-1)^k\sum_{i_1<\cdots<i_k} w\big(A_{i_1}\cap\cdots\cap A_{i_k}\big).$$

□

Let X be a finite set and $A_1,\ldots,A_n$ be subsets of X. Let $[n]=\{1,2,\ldots,n\}$. We introduce two functions α and β on the power set $\mathcal{P}([n])$ of $[n]$ as follows: For each subset $I\subseteq[n]$,

$$\alpha(I) = \begin{cases} w\left(\bigcap_{i\in I} A_i\right) & \text{if } I\neq\emptyset, \\ 0 & \text{if } I=\emptyset; \end{cases}$$

$$\beta(I) = \begin{cases} w\left(\bigcup_{i\in I} A_i\right) & \text{if } I\neq\emptyset, \\ 0 & \text{if } I=\emptyset. \end{cases}$$

By Inclusion-Exclusion,

$$1_{\bigcup_{i=1}^n A_i} = \sum_{k=1}^{n}(-1)^{k-1}\sum_{1\le i_1<\cdots<i_k\le n} 1_{A_{i_1}\cap\cdots\cap A_{i_k}} = \sum_{I\subseteq[n],\, I\neq\emptyset}(-1)^{|I|-1}1_{\bigcap_{i\in I}A_i}.$$

Taking weight w on both sides, we obtain

$$\beta([n]) = \sum_{I\subseteq[n],\, I\neq\emptyset}(-1)^{|I|-1}\alpha(I) = \sum_{I\subseteq[n]}(-1)^{|I|+1}\alpha(I).$$

If one replace $\bar{A}_i$ with A_i in (5), we have

$$\begin{aligned}
1_{\bigcap_{i=1}^n A_i} &= 1_X + \sum_{k=1}^n (-1)^k \sum_{i_1<\cdots<i_k} 1_{\bar{A}_{i_1}\cap\cdots\cap\bar{A}_{i_k}} \quad \left(\textstyle\sum_{k=0}^n \binom{n}{k} (-1)^k = 0\right) \\
&= \sum_{k=1}^n (-1)^{k+1} \sum_{i_1<\cdots<i_k} \left(1_X - 1_{\bar{A}_{i_1}\cap\cdots\cap\bar{A}_{i_k}}\right) \\
&= \sum_{k=1}^n (-1)^{k+1} \sum_{i_1<\cdots<i_k} 1_{A_{i_1}\cup\cdots\cup A_{i_k}} \\
&= \sum_{I\subseteq[n],\, I\neq\emptyset} (-1)^{|I|+1} 1_{\bigcup_{i\in I} A_i}.
\end{aligned}$$

Taking the weight w on both sides, we obtain

$$\alpha([n]) = \sum_{I\subseteq[n],\, I\neq\emptyset} (-1)^{|I|+1}\beta(I) = \sum_{I\subseteq[n]} (-1)^{|I|+1}\beta(I).$$

Theorem 8.2. *We have the identities*

$$\beta(J) = \sum_{I\subseteq J} (-1)^{|I|+1}\alpha(I), \quad \forall J \subseteq [n]; \tag{7}$$

$$\alpha(J) = \sum_{I\subseteq J} (-1)^{|I|+1}\beta(I), \quad \forall J \subseteq [n]. \tag{8}$$

9 Möbius Inversion

Let $(X, \leq)$ be a **locally finite** poset, i.e., for each $x \leq y$ in X the interval $[x, y] = \{z \in X : x \leq z \leq y\}$ is a finite set. Let $\mathcal{I}(X)$ be the set of all functions $f : X \times X \to \mathbb{R}$ such that

$$f(x, y) = 0 \quad \text{if } x \not\leq y;$$

such functions are called **incidence functions** on the poset X. For an incidence function f, we only specify the values $f(x, y)$ for the pairs (x, y) such that $x \leq y$, since $f(x, y) = 0$ for all pairs (x, y) such that $x \not\leq y$.

The **convolution product** of two incidence functions $f, g \in \mathcal{I}(X)$ is an incidence function $f * g : X \times X \to \mathbb{R}$, defined by

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

In fact, $(f * g)(x, y) = 0$ if $x \not\leq y$ (since either $x \not\leq z$ or $z \not\leq y$ for each z) and

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

The convolution product satisfies the associative law:

$$f * (g * h) = (f * g) * h,$$

where $f, g, h \in \mathcal{I}(X)$. Indeed, for $x \leq y$, we have

$$\begin{aligned} (f * (g * h))(x, y) &= \sum_{x \leq z_1 \leq y} f(x, z_1)(g * h)(z_1, y) \\ &= \sum_{x \leq z_1 \leq y} f(x, z_1) \sum_{z_1 \leq z_2 \leq y} g(z_1, z_2) h(z_2, , y) \\ &= \sum_{x \leq z_1 \leq z_2 \leq y} f(x, z_1) g(z_1, z_2) h(z_2, y). \end{aligned}$$

Likewise, for $x \leq y$, we have

$$((f * g) * h))(x, y) = \sum_{x \leq z_1 \leq z_2 \leq y} f(x, z_1) g(z_1, z_2) h(z_2, y).$$

For $x \not\leq y$, we automatically have $(f * (g * h))(x, y) = ((f * g) * h))(x, y) = 0$. The vector space $\mathcal{I}(X)$ together with the convolution $*$ is called the **incidence algebra** of X.

We may think of that incidence functions f are only defined on the set $\{(x, y) \in X \times X : x \leq y\}$, and the convolution is defined as

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

Example 9.1. Let $[n] = \{1, 2, \ldots, n\}$ be the poset with the natrual order of natural numbers. An incidence function $f : [n] \times [n] \to \mathbb{R}$ can be viewed as a upper triangular $n \times n$ matrix $A = [a_{ij}]$ given by $a_{ij} = f(i, j)$. The convolution is just the multiplication of upper triangular matrices.

There is a special function $\delta \in \mathcal{I}(X)$, called the **delta function** of the poset $(X, \leq)$, defined by

$$\delta(x,y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases}$$

The delta function δ is the **identity** of the algebra $\mathcal{I}(X)$, i.e., for all $f \in \mathcal{I}(X)$,

$$\delta * f = f = f * \delta.$$

Indeed, for $x \leq y$,

$$(\delta * f)(x,y) = \sum_{x \leq z \leq y} \delta(x,z) f(z,y) = f(x,y);$$

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

Given an incidence function $f \in \mathcal{I}(X)$. A **left inverse** of f is a function $g \in \mathcal{I}(X)$ such that

$$g * f = \delta.$$

A **right inverse** of f is a function $h \in \mathcal{I}(X)$ such that

$$f * h = \delta.$$

If f has a left inverse g and a right inverse h, then $g = h$. In fact,

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

If f has both a left and right inverse, we say that f is **invertible**; the left inverse and right inverse of f must be same and unique, and it is just called the **inverse** of f.

Note that

$$g * f = \delta \quad \Leftrightarrow \quad \sum_{x \leq z \leq y} g(x,z) f(z,y) = \delta(x,y), \quad \forall\, x \leq y.$$

When $x = y$, we have $g(x,x)f(x,x) = 1$, i.e., $g(x,x) = \frac{1}{f(x,x)}$; so $f(x,x) \neq 0$. We can obtain $g \in \mathcal{I}(X)$ inductively as follows:

$$g(x,x) = \frac{1}{f(x,x)}, \quad \forall\, x \in X, \tag{9}$$

$$g(x,y) = \frac{-1}{f(y,y)} \sum_{x \leq z < y} g(x,z) f(z,y), \quad \forall\, x < y. \tag{10}$$

This means that f is invertible iff $f(x,x) \neq 0$ for all $x \in X$.

Likewise,

$$f * g = \delta \quad \Leftrightarrow \quad \sum_{x \leq z \leq y} f(x,z)g(z,y) = \delta(x,y), \quad \forall\, x \leq y.$$

We can obtain $g \in \mathcal{I}(X)$ inductively as follows:

$$g(x,x) \;=\; \frac{1}{f(x,x)}, \quad \forall\, x \in X, \tag{11}$$

$$g(x,y) \;=\; \frac{-1}{f(x,x)} \sum_{x < z \leq y} f(x,z)g(z,y), \quad \forall\, x < y. \tag{12}$$

The **zeta function** ζ of the poset $(X, \leq)$ is an incidence function such that $\zeta(x,y) = 1$ for all (x,y) with $x \leq y$. Clearly, ζ is invertible. The **Möbius function** μ of the poset $(X, \leq)$ is the inverse of the zeta function ζ in the incidence algebra $\mathcal{I}(X)$, i.e.,

$$\mu = \zeta^{-1}.$$

The Möbius function μ can be inductively defined by

$$\mu(x,x) \;=\; 1, \quad \forall x \in X, \tag{13}$$

$$\mu(x,y) \;=\; -\sum_{x \leq z < y} \mu(x,z) = -\sum_{x < z \leq y} \mu(z,y), \quad \forall\, x < y. \tag{14}$$

Example 9.2. Let $X = \{1, 2, \ldots, n\}$ and consider the linearly ordered set $(X, \leq)$, where $1 < 2 < \cdots < n$. Then for $(k,l) \in X \times X$ with $k \leq l$, the Möbius function is given by

$$\mu(k,l) = \begin{cases} 1 & \text{if } l = k, \\ -1 & \text{if } l = k+1, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that $\mu(k,k) = 1$ and $\mu(k,k+1) = -1$. It follows that $\mu(k,k+2) = 0$ and subsequently, $\mu(k,k+i) = 0$ for all $i \geq 2$.

Example 9.3. Let $X = \{1, 2, \ldots, n\}$. The Möbius function of the poset $(\mathcal{P}(X), \subseteq)$ is given by

$$\mu(A,B) = (-1)^{|B-A|}, \quad \text{where } A \subseteq B.$$

This can be proved by induction on $|B - A|$. For $|B - A| = 0$, i.e., $A = B$, it is obviously true. Consider the case of $|B - A| = m \geq 1$ and assume that it is true when $|B - A| < m$. In fact,

$$\begin{aligned}
\mu(A, B) &= -\sum_{A \subseteq C \subsetneq B} \mu(A, C) = -\sum_{A \subseteq C \subsetneq B} (-1)^{|C-A|} \\
&= -\sum_{D \subsetneq B-A} (-1)^{|D|} = -\sum_{k=0}^{m-1} \binom{m}{k} (-1)^k \\
&= (-1)^m - \sum_{k=0}^{m} \binom{m}{k} (-1)^k = (-1)^{|B-A|}.
\end{aligned}$$

Example 9.4. Consider the poset of 12 members whose Hasse diagram is as follows. Fix an minimal element x, the second of the bottom member from the left blow. If y_1 is the first member of the second bottom layer, then $\mu(x, y_1) = -1$. If y_2 is the second of the second top layer, then $\mu(x, y_2) = 2$. If y_3 is the first of the top layer, then $\mu(x, y_3) = -2$.

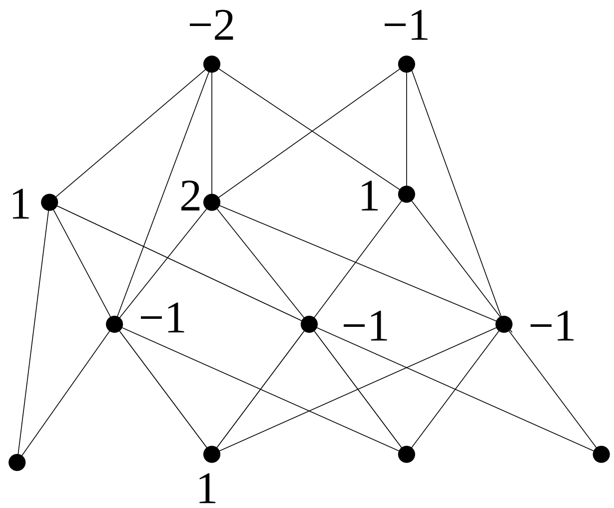


Figure 1: Computing the Möbius function by Hasse diagram

Given a finite poset $(X, \leq)$. For each function $f : X \to \mathbb{R}$, we can multiply an incidence function $\alpha \in \mathcal{I}(X)$ to the left of f and to the right as follows to obtain two functions $\alpha * f$ and $f * \alpha$ on X, defined by

$$(\alpha * f)(x) = \sum_{x \leq y} \alpha(x, y) f(y), \quad \forall x \in X; \tag{15}$$

$$(f * \alpha)(y) = \sum_{x \le y} f(x)\alpha(x, y), \quad \forall y \in X. \tag{16}$$

Theorem 9.1. *Let $(X, \le)$ be a finite poset. Given invertible $\alpha \in \mathcal{I}(X)$, $f, g \in F(X)$. Then $g = \alpha * f$ iff $f = \alpha^{-1} * g$, i.e.,*

$$g(x) = \sum_{x \le y} \alpha(x, y) f(y), \ \forall\, x \in X \ \Leftrightarrow \ f(x) = \sum_{x \le y} \alpha^{-1}(x, y) g(y), \ \forall\, x \in X.$$

Likewise, $g = f * \alpha$ iff $f = g * \alpha^{-1}$, i.e.,

$$g(y) = \sum_{x \le y} f(x)\alpha(x, y), \ \forall\, y \in X \ \Leftrightarrow \ f(y) = \sum_{x \le y} g(x)\alpha^{-1}(x, y), \ \forall\, y \in X. \tag{17}$$

Proof. It follows from the fact $g = \alpha * f$ iff $\alpha^{-1} * g = \alpha^{-1} * (\alpha * f)$, and the fact

$$\alpha^{-1} * (\alpha * f) = (\alpha^{-1} * \alpha) * f = \delta * f = f.$$

Likewise, $g = f * \alpha \Leftrightarrow g * \alpha^{-1} = f * \alpha * \alpha^{-1} = f * \delta = f$. $\square$

Theorem 9.2. *Let $(X, \le)$ be a finite poset. Let f, g be real-valued functions on X. Then*

$$g(x) = \sum_{x \le y} f(y), \ \ \forall\, x \in X \ \Leftrightarrow \ f(x) = \sum_{x \le y} \mu(x, y) g(y), \ \ \forall\, x \in X;$$

$$g(y) = \sum_{x \le y} f(x), \ \ \forall\, y \in X \ \Leftrightarrow \ f(y) = \sum_{x \le y} g(x)\mu(x, y), \ \ \forall\, y \in X. \tag{18}$$

Proof. The first inversion formula follows from the fact that $g = \zeta * f \Leftrightarrow f = \zeta^{-1} * g = \mu * g$. The second inversion formula follows from the fact that $g = f * \zeta \Leftrightarrow f = g * \zeta^{-1} = g * \mu$.

Writing in summations, for each fixed $y \in X$, we have

$$\begin{aligned}
\sum_{x \le y} g(x)\mu(x,y) &= \sum_{x \le y}\sum_{u \le x} f(u)\mu(x,y) \\
&= \sum_{x \le y}\sum_{u \le x} f(u)\zeta(u,x)\mu(x,y) \\
&= \sum_{u \le y} f(u) \sum_{u \le x \le y} \zeta(u,x)\mu(x,y) \\
&= \sum_{u \le y} f(u)\delta(u,y) \\
&= f(y).
\end{aligned}$$

□

Corollary 9.3. *Let* $[n] = \{1, 2, \ldots, n\}$. *Let* $f, g : \mathcal{P}([n]) \to \mathbb{R}$ *be functions such that*

$$g(I) = \sum_{J \subseteq I} f(J), \quad I \subseteq [n].$$

Then

$$f(I) = \sum_{J \subseteq I} (-1)^{|I-J|} g(J), \quad I \subseteq [n].$$

Permanent. Fix a positive integer n. Let $\mathfrak{S}_n$ denote the symmetric group of $[n] = \{1, 2, \ldots, n\}$, i.e., the set of all permutations of $[n]$. Let A be an $n \times n$ real matrix. The **permanent** of A is defined as the number

$$\operatorname{per}(A) = \sum_{\sigma \in \mathfrak{S}_n} \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

For the chessboard C in Example 6.1, we associate a 0-1 matrix $A = [a_{ij}]$ as follows:

$$C = \begin{array}{|c|c|c|c|c|}\hline \times & \times & & & \\ \hline & & \times & \times & \\ \hline \times & & & & \times \\ \hline & \times & \times & & \\ \hline & & & \times & \times \\ \hline \end{array}\,, \qquad A = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

Then the number of ways to put 5 non-attacking indistinguishable rooks on C is the permanent $\mathrm{per}(A)$.

Fix an n-by-n matrix A. For each subset $I \subseteq [n]$, let A_I denote the submatrix of A, whose rows are those of A indexed by members of I. Let $F(I)$ be the set of all functions $\sigma : [n] \to I$, and let $G(I)$ be the set of all surjective functions from $[n]$ onto I. Then

$$F(I) = \bigsqcup_{J \subseteq I} G(J).$$

We introduce a real-valued function f on the power set $\mathcal{P}([n])$ of $[n]$, defined by

$$\begin{aligned} f(\emptyset) &= 0, \\ f(I) &= \sum_{\sigma \in G(I)} \prod_{i=1}^{n} a_{i,\sigma(i)}, \quad \forall I \subseteq [n],\ I \neq \emptyset. \end{aligned}$$

Note that $f([n]) = \mathrm{per}(A)$. Let $g : \mathcal{P}([n]) \to \mathbb{R}$ be defined by

$$g(I) = \sum_{J \subseteq I} f(J), \quad \forall I \subseteq [n].$$

Then

$$\begin{aligned} g(I) &= \sum_{J \subseteq I} \sum_{\sigma \in G(J)} \prod_{i=1}^{n} a_{i,\sigma(i)} \\ &= \sum_{\sigma \in F(I)} \prod_{i=1}^{n} a_{i,\sigma(i)}, \\ &= \prod_{i=1}^{n} \left(\sum_{j \in I} a_{ij} \right), \quad \forall I \subseteq [n]. \end{aligned}$$

Thus by the Möbius inversion, we have

$$f(I) = \sum_{J \subseteq I} (-1)^{|I-J|} g(J), \quad I \subseteq [n].$$

In particular,

$$f([n]) = \sum_{I \subseteq [n]} (-1)^{n-|I|} g(I).$$

Since $f([n]) = \mathrm{per}(A)$, it follows that

$$\mathrm{per}(A) = \sum_{I \subseteq [n]} (-1)^{n-|I|} \prod_{i=1}^{n} \Big(\sum_{j \in I} a_{ij} \Big). \tag{19}$$

However this formula is not much useful because there are 2^n terms in the summation.

Definition 9.4. Let $(X_i, \preceq_i)$ $(i = 1, 2)$ be two posets. The product poset $(X_1 \times X_2, \preceq)$ is given by

$$(x_1, x_2) \preceq (y_1, y_2) \quad \text{iff} \quad x_1 \preceq_1 y_1, \ x_2 \preceq_2 y_2.$$

For the convenience, we write $\preceq_1$ and $\preceq_2$ simply as $\preceq$. Then $(X_1 \times X_2, \preceq)$ is a poset.

Theorem 9.5. *Let μ_i be the Möbius functions of posets $(X_i, \preceq_i)$, $i = 1, 2$. Then the Möbius function μ of $X_1 \times X_2$ for $(x_1, x_2) \preceq (y_1, y_2)$ is given by*

$$\mu((x_1, x_2), (y_1, y_2)) = \mu_1(x_1, y_1)\, \mu_2(x_2, y_2).$$

Proof. We proceed by induction on $\ell((x_1, x_2), (y_1, y_2))$, the length of the longest chains in the interval $[(x_1, x_2), (y_1, y_2)]$. It is obviously true when $\ell = 0$. For $\ell \geq 1$, by inductive definition of μ,

$$\begin{aligned}
\mu((x_1, x_2), (y_1, y_2)) &= - \sum_{(x_1,x_2) \preceq (z_1,z_2) \prec (y_1,y_2)} \mu((x_1, x_2), (z_1, z_2)) \\
&= - \sum_{(x_1,x_2) \preceq (z_1,z_2) \prec (y_1,y_2)} \mu_1(x_1, z_1)\, \mu_2(x_2, z_2) \quad \text{(by IH)} \\
&= \mu_1(x_1, y_1)\, \mu_2(x_2, y_2) - \sum_{x_1 \preceq z_1 \preceq y_1} \mu_1(x_1, z_1) \sum_{x_2 \preceq z_2 \preceq y_2} \mu_2(x_2, z_2) \\
&= \mu_1(x_1, y_1)\, \mu_2(x_2, y_2) - \delta_1(x_1, y_1)\, \delta_2(x_2, y_2) \\
&= \mu_1(x_1, y_1)\, \mu_2(x_2, y_2).
\end{aligned}$$

□

Example 9.5. The set $\mathbb{Z}_+ = \{1, 2, \ldots\}$ of positive integers is a poset with the partial order of divisibility. Let $n \in \mathbb{Z}_+$ be factored as

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k},$$

where p_i are distinct primes and e_i are positive integers. Since $\mu(m,m) = 1$ for all $m \in \mathbb{Z}_+$ and $\mu(1,n)$ is inductively given by

$$\mu(1,n) = -\sum_{m\in\mathbb{Z}_+,\, m|n,\, m\neq n} \mu(1,m)$$

We only need to to consider the subposet $(D(n), \text{divisibility})$, where

$$D(n) = \{d \in [n] : d \mid n\}.$$

For $r, s \in D(n)$, they can be written as

$$r = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}, \quad s = p_1^{b_1} p_2^{b_2} \cdots p_k^{b_k},$$

where $0 \le a_i, b_i \le e_i$. Then $r \mid s$ iff $a_i \le b_i$. This means that the poset $D(n)$ is isomorphic to the product poset

$$Q = \big\{(a_1, \ldots, a_k) : a_i \in [0, e_i]\big\} = \prod_{i=1}^{k} [0, e_i],$$

where $[0, e_i] = \{0, 1, \ldots, e_i\}$. Thus $\mu(1,n) = \mu_Q((0,\ldots,0),(e_1,\ldots,e_k))$, where

$$\mu_Q((0,\ldots,0),(e_1,\ldots,e_k)) = \prod_{i=1}^{k} \mu_{[0,e_i]}(0, e_i).$$

Note that

$$\mu_{[0,e_i]}(0,e_i) = \begin{cases} 1 & \text{if } e_i = 0, \\ -1 & \text{if } e_i = 1, \\ 0 & \text{if } e_i \ge 2. \end{cases} = \begin{cases} (-1)^{e_i} & \text{if } e_i \le 1, \\ 0 & \text{if } e_i \ge 2. \end{cases}$$

It follows that

$$\begin{aligned} \mu(1,n) &= \begin{cases} (-1)^{e_1+\cdots+e_k} & \text{if all } e_i \le 1, \\ 0 & \text{otherwise.} \end{cases} \\ &= \begin{cases} 1 & \text{if } n = 1, \\ (-1)^j & \text{if } n \text{ is a product of } j \text{ distinct primes}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Now for arbitrary $m, n \in \mathbb{Z}_+$ such that $m \mid n$, the bijection

$$\{u \in \mathbb{Z}_+ : m \mid u,\ u \mid n\} \xrightarrow{\sim} \left\{v \in \mathbb{Z}_+ : v \mid \frac{n}{m}\right\}, \quad u \mapsto \frac{u}{m}$$

is an isomorphism of posets for the partial order of divisibility. We thus have

$$\mu(m,n) = \mu\left(1, \frac{n}{m}\right).$$

In number theory, we write $\mu(1,n)$ as $\mu(n)$.

Theorem 9.6. *Let $f, g : \mathbb{Z}_+ \to \mathbb{C}$ be two functions. Then*

$$g(n) = \sum_{d|n} f(d), \quad \forall\, n \in \mathbb{Z}_+$$

is equivalent to

$$f(n) = \sum_{d|n} g(d)\mu\left(\frac{n}{d}\right), \quad \forall\, n \in \mathbb{Z}_+.$$

Example 9.6. Let $\Phi_n = \{a \in [n] : \gcd(a,n) = 1\}$. Then $\phi(n) = |\Phi_n|$. Define

$$g(n) = \sum_{d|n} \phi(d), \quad \forall\, n \in \mathbb{Z}_+.$$

Consider the set $\Phi_{n,d} = \{k \in [n] : \gcd(k,n) = d\}$ for each factor d of n. In particular, if $d = 1$, then $\Phi_{n,1} = \Phi_n$. In fact, there is a bijection

$$\Phi_{n,d} \to \Phi_{n/d}, \quad k \mapsto k/d.$$

(Injectivity is trivial. Surjectivity follows from $da \mapsto a$ for $a \in \Phi_{n/d}$.) Then $\phi(n/d) = |\Phi_{n/d}| = |\Phi_{n,d}|$.

Note that for each integer $k \in [n]$, there is a unique integer $d \in [n]$ such that $\gcd(k,n) = d$. We have $[n] = \bigsqcup_{d|n} \Phi_{n,d}$ (disjoint union). Thus

$$n = \sum_{d|n} |\Phi_{n,d}| = \sum_{d|n} \phi\left(\frac{n}{d}\right) = \sum_{k|n} \phi(k) = \sum_{d|n} \phi(d).$$

By the Möbius inversion,

$$\phi(n) = \sum_{k|n} k\mu(k,n) = \sum_{k|n} k\mu\left(\frac{n}{k}\right) = \sum_{dk=n} k\mu(d) = \sum_{d|n} \mu(d) \cdot \frac{n}{d}. \qquad (20)$$

Let $n \geq 2$ and let $p_1, p_2, \ldots, p_r$ be distinct primes dividing n. Then

$$\{d \in [n] : d \mid n,\ \mu(d) \neq 0\} = \left\{\textstyle\prod_{i \in I} p_i : I \subseteq [r]\right\},$$

where $\prod_{i\in\emptyset} p_i = 1$. Since $\mu(1) = 1$, $\mu(d) = (-1)^k$ if $d = p_{i_1}\cdots p_{i_k}$ is a product of k distinct primes, and $\mu(d) = 0$ otherwise, we see that (20) becomes

$$\begin{aligned}\phi(n) &= n - \left(\frac{n}{p_1} + \frac{n}{p_2} + \cdots\right) + \left(\frac{n}{p_1p_2} + \frac{n}{p_1p_3} + \cdots\right) - \cdots \\ &\quad + \cdots + (-1)^r \frac{n}{p_1p_2\cdots p_r} \\ &= n\prod_{i=1}^{r}\left(1 - \frac{1}{p_i}\right) = n \prod_{p\mid n,\ \text{primes}} \left(1 - \frac{1}{p}\right).\end{aligned}$$

Example 9.7. Let $\Sigma = \{a_1, \ldots, a_k\}$ be a set and $M = \{\infty\cdot a_1, \ldots, \infty\cdot a_k\}$ a multiset over Σ. A **circular n-permutation** of M is an arrangement of n elements of M around a circle. Each circular n-permutation of M may be considered as a periodic double-infinite sequence

$$(x_i) = (x_i)_{i\in\mathbb{Z}} = \cdots x_{-2}x_{-1}x_0x_1x_2\cdots$$

of period n, i.e., $x_{i+n} = x_i$ for all $i \in \mathbb{Z}$. The **minimum period** of a circular permutation (x_i) of M is the smallest positive integer among all periods. We shall see below that the minimum period of a double-infinite sequence divides all periods of the sequence.

Let Σ_n denote the set of n-words over Σ, and Σ^* the set of all words over Σ. Then $\Sigma^* = \bigsqcup_{n\geq 0}\Sigma_n$ (disjoint union). Consider the map

$$\sigma : \Sigma_n \to \Sigma_n, \quad \sigma(x_1x_2\cdots x_n) = x_2\cdots x_nx_1, \quad \sigma(\lambda) = \lambda.$$

An n-word w is **primitive** if

$$w, \quad \sigma(w), \quad \sigma^2(w), \quad \ldots, \quad \sigma^{n-1}(w)$$

are distinct. A **period** of an n-word w is a positive integer m such that

$$\sigma^m(w) = w.$$

Every n-word has a trivial period n. The **minimum period** d of an n-word w is the smallest positive integer among all periods of w, which is a common factor of all periods of w; in particular, $d \mid n$. In fact, for a period m of an n-word w, write $m = qd + r$, where $0 \leq r < d$. Suppose $r > 0$. Then

$$\sigma^r(w) = \sigma^r\underbrace{\sigma^d\cdots\sigma^d}_{q}(w) = \sigma^{qd+r}(w) = \sigma^m(w) = w,$$

which means that r is a period of w and is smaller than d, subsequently, contradictory to the minimality of d.

Let Σ_d^0 denote the set of primitive d-words over Σ, and $\Sigma_{n,d}$ the subset of Σ_n whose n-words have minimum period d, where $d \mid n$. Clearly,

$$|\Sigma_{n,d}| = |\Sigma_d^0|.$$

Let $\mathbb{Z}(\Sigma)$ denote the set of double-infinite sequences over Σ, and $\mathbb{Z}_n(\Sigma)$ the subset of $\mathbb{Z}(\Sigma)$ whose members have period n. Let $\mathbb{Z}_d^0(\Sigma)$ denote the subset of $\mathbb{Z}_d(\Sigma)$ whose members have minimum period d. Then

$$\mathbb{Z}_n(\Sigma) = \bigsqcup_{d|n} \mathbb{Z}_d^0(\Sigma).$$

Let $C_n(\Sigma)$ denote the set of all circular n-permutations of M. Then $C_n(\Sigma)$ can be identified to the set $\mathbb{Z}_n(\Sigma)$. Thus

$$|C_n(M)| = |\mathbb{Z}_n(\Sigma)| = \sum_{m|n} |\mathbb{Z}_m^0(\Sigma)|.$$

Now we consider the map

$$F : \Sigma_m^0 \longrightarrow \mathbb{Z}_m^0(\Sigma), \quad w = s_1 s_2 \cdots s_m \mapsto (x_i) = \cdots www \cdots ,$$

which is clearly surjective and each member of $\mathbb{Z}_m^0(\Sigma)$ receives exactly m members of Σ_m^0. So $|\mathbb{Z}_m^0(\Sigma)| = |\Sigma_m^0|/m$. Thus

$$|C_n(M)| = \sum_{m|n} |\mathbb{Z}_m^0(\Sigma)| = \sum_{m|n} |\Sigma_m^0|/m.$$

Since $\Sigma_n = \bigsqcup_{d|n} \Sigma_{n,d}$, where $\Sigma_{n,d}$ is the subset of Σ_n whose words have minimum period d, we have

$$|\Sigma_n| = \sum_{d|n} |\Sigma_{n,d}| = \sum_{d|n} |\Sigma_d^0|.$$

By the Möbius inversion,

$$|\Sigma_n^0| = \sum_{d|n} |\Sigma_d| \, \mu(d,n) = \sum_{d|n} |\Sigma_d| \, \mu\left(\frac{n}{d}\right).$$

It follows that

$$\begin{aligned} |C_n(M)| &= \sum_{m|n} \frac{1}{m} \sum_{a|m} |\Sigma_a| \, \mu\left(\frac{m}{a}\right) \\ &= \sum_{a|m,\, m|n} \frac{1}{m} |\Sigma_a| \, \mu\left(\frac{m}{a}\right). \end{aligned}$$

Set $b := m/a$, i.e., $ab = m$. Then $a \mid m$ and $m \mid n$ are equivalent to $a \mid n$ and $b \mid (n/a)$. Thus

$$|C_n(M)| = \sum_{a|n} |\Sigma_a| \sum_{b|(n/a)} \frac{\mu(b)}{ab}.$$

Since $\phi(n) = \sum_{d|n} \mu(d) \cdot \frac{n}{d}$ by (20), we see that

$$\sum_{b|(n/a)} \frac{\mu(b)}{ab} = \frac{1}{n} \sum_{b|(n/a)} \mu(b) \cdot \frac{n/a}{b} = \frac{1}{n} \phi\left(\frac{n}{a}\right).$$

We finally have

$$\begin{aligned} |C_n(M)| &= \frac{1}{n} \sum_{a|n} |\Sigma_a| \, \phi\left(\frac{n}{a}\right) \quad (\text{since } |\Sigma_a| = k^a|) \\ &= \frac{1}{n} \sum_{a|n} k^a \phi\left(\frac{n}{a}\right) \quad (\text{set } d = n/a) \\ &= \frac{1}{n} \sum_{d|n} k^{n/d} \phi(d). \end{aligned}$$

Theorem 9.7. *The number of circular n-permutations of a set of k objects with repetition allowed is*

$$\frac{1}{n} \sum_{d|n} k^{n/d} \phi(d)$$

Theorem 9.8. *The number circular permutations of a multiset M of type $(n_1, \ldots, n_k)$ with $m = \gcd(n_1, \ldots, n_k)$ and $n = n_1 + \cdots + n_k$ is given by*

$$\frac{1}{n} \sum_{a|m} \phi(a) \binom{n/a}{n_1/a, \ldots, n_k/a}. \tag{21}$$

Proof. If a is a period of a permutation of M, it is easy to see that $a \mid m$. Let

$$M_d = \{(dn_1/m) \cdot a_1,\ \ldots,\ (dn_k/m) \cdot a_k\}, \quad d \geq 1.$$

Clearly, $\gcd\{dn_1/m, \ldots, dn_k/m\} = d$ and $M_m = M$.

Let $\mathfrak{S}(M_d)$ denote the set of all permutations of M_d, $\mathfrak{S}_a^0(M_d)$ the set of all permutations of M_d with minimum period a, and $\mathfrak{S}^0(M_d)$ the set of all primitive permutations of M_d, i.e., permutations whose minimum period is the cardinality dn/m of M_d. For each $w \in \mathfrak{S}(M_d)$, let a be the minimum period of w. Then $w = \underbrace{w_1 w_1 \cdots w_1}_{b}$ with a primitive word w_1. Thus w_1 is a word of length a of type

$$\left(\frac{dn_1/m}{b}, \ldots, \frac{dn_k/m}{b}\right).$$

Since $b \mid (dn_i/m)$ for all i, it follows that $b \mid \gcd\{dn_1/m, \ldots, dn_k/m\}$, i.e., $b \mid d$. So $w \in \mathfrak{S}_a^0(M_d)$ with $a = \sum_{i=1}^k (d/b)(n_i/m) = (d/b)(n/m)$. Note that

$$\gcd\{(d/b)(n_1/m), \ldots, (d/b)(n_k/m)\} = d/b.$$

We see that

$$\mathfrak{S}(M_d) = \bigsqcup_{b \mid d} \mathfrak{S}^0_{(d/b)(n/m)}(M_d),$$

$$\mathfrak{S}^0_{(d/b)(n/m)}(M_d) \simeq \mathfrak{S}^0(M_{d/b}) \quad \text{if} \quad b \mid d.$$

We then have

$$|\mathfrak{S}(M_d)| = \sum_{b \mid d} |\mathfrak{S}^0(M_{d/b})| = \sum_{a \mid d} |\mathfrak{S}^0(M_a)|.$$

By the Möbius inversion,

$$|\mathfrak{S}^0(M_d)| = \sum_{a \mid d} |\mathfrak{S}(M_a)|\, \mu\left(\frac{d}{a}\right).$$

Let $C(M_d)$ denote the set of all circular permutations of M_d, $C_a^0(M_d)$ the set of all circular permutations of M_d with minimum period a, and $C^0(M_d)$ the set of all primitive circular permutations of M_d. Likewise,

$$C(M_d) = \bigsqcup_{a \mid d} C_a^0(M_d),$$

$$|C_a^0(M_d)| = |C^0(M_a)| \quad \text{if} \quad a \mid d.$$

Note that $|M_a| = an/m$ and $|\mathfrak{S}^0(M_a)| = |C^0(M_a)| \cdot an/m$. We have

$$\begin{aligned} |C(M_d)| &= \sum_{a|d} |C^0(M_a)| = \sum_{a|d} |\mathfrak{S}^0(M_a)| \cdot \frac{1}{an/m} \\ &= \sum_{a|d} \frac{m}{an} \sum_{b|a} |\mathfrak{S}(M_b)| \, \mu\left(\frac{a}{b}\right). \end{aligned}$$

Set $c = a/b$, i.e., $a = bc$, then $a \mid d$ and $b \mid a$ are equivalent to $b \mid d$ and $c \mid (d/b)$. Thus

$$\begin{aligned} |C(M_d)| &= \frac{1}{n} \sum_{b|d} \frac{m}{d} |\mathfrak{S}(M_b)| \sum_{c|\frac{d}{b}} \frac{d/b}{c} \mu(c) \\ &= \frac{1}{n} \sum_{b|d} \frac{m}{d} |\mathfrak{S}(M_b)| \, \phi\left(\frac{d}{b}\right) \quad (\text{set } a = d/b) \\ &= \frac{1}{n} \sum_{a|d} \frac{m}{d} |\mathfrak{S}(M_{d/a})| \, \phi(a). \end{aligned}$$

Let $d = m$, we have $M = M_m$. Recall $|\mathfrak{S}(M_b)| = \binom{bn/m}{bn_1/m, \ldots, bn_k/m}$, therefore

$$\begin{aligned} |C(M)| &= |C(M_m)| = \frac{1}{n} \sum_{a|m} \phi(a) \, |\mathfrak{S}(M_{m/a})| \\ &= \frac{1}{n} \sum_{a|m} \phi(a) \binom{n/a}{n_1/a, \ldots, n_k/a}. \end{aligned}$$

□

Example 9.8. Consider the multiset $M = \{12a_1, 24a_2, 18a_3\}$ of type $(12, 24, 18)$. Then $m = \gcd(12, 24, 18) = 6$, whose factors are $1, 2, 3, 6$. Recall the values

$$\phi(1) = 1, \quad \phi(2) = 1, \quad \phi(3) = 2, \quad \phi(6) = 2.$$

The number of circular permutations of M is

$$\frac{1}{54}\left[\phi(1)\binom{54}{12, 24, 18} + \phi(2)\binom{27}{6, 12, 9} + \phi(3)\binom{18}{4, 8, 6} + \phi(6)\binom{9}{2, 4, 3}\right]$$

10 Problems

1. Let $(P, \leq)$ be a finite poset. Recall that na incidence function is a function $F : P \times P \to \mathbb{C}$ such that $f(x, y) = 0$ if $x \not\leq y$. The convolution of two incidence functions f, g is a function $f * g : P \to \mathbb{C}$ defined by

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

   (a) Show that $f * g$ is an incidence function, i.e., $(f * g)(x, y) = 0$ for all pairs (x, y) such that $x \not\leq y$.

   (b) If $x \leq y$, show that

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

2. Let P be a finite poset. Think of each incidence function $f : P \times P \to \mathbb{C}$ as a square matrix whose row and column indices are members of P, and whose (x, y)-entry is $f(x, y)$.

   (a) Show that the convolution of incidence functions is just the matrix multiplication.

   (b) Incidence algebra of the poset P is a subalgebra of the algebra of matrices whose rows and columns are indexed by members of P.

3.