Combinatorics, 2015 Fall, USTC
Week 3

An easy application of the method of generating functions

- Recall: Given subsets $I_1, I_2, ..., I_k$ of non-negative integers, let $f_j(x) := \sum_{l \in I_j} x^l$ for all $1 \leq j \leq k$, and let $f(x) := \prod_{j=1}^{k} f_j(x)$. Then the number a_n of integer solutions to $i_1 + i_2 + ... + i_k = n$, where $i_j \in I_j$, is equal to $[x^n]f$. In other words, $f(x)$ is the generating function of $\{a_n\}$.

- The following problem is a warm-up of the method of generating functions.

- Let a_n be the number of ways to pay n Chinese Yuan using 1-Yuan bills, 2-Yuan bills and 5-Yuan bills (assume there exist such bills). What is the generating function of this sequence $\{a_n\}$?

- Observe that a_n corresponds to the number of integer solutions (i_1, i_2, i_3) to

  $$i_1+i_2+i_3 = n, \quad \text{where } i_1 \in I_1 := \{0,1,2,...\}, i_2 \in I_2 := \{0,2,4,...\} \text{ and } i_3 \in I_3 := \{0,5,10,...\}.$$

  Let $f_j(x) := \sum_{m \in I_j} x^m$ for $j = 1,2,3$. Then $f(x) := \prod_{1 \leq j \leq 3} f_j(x)$ is such that $[x^n]f = a_n$. That is, the generating function of $\{a_n\}$ is $f(x) = \frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^5}$.

Integer partition

- How many ways are there to write a natural number n as a sum of several natural numbers?

- The answer is not too difficult if we count *ordered partitions* of n. Here "ordered partition" means that we will view $1+1+2, 1+2+1$ as two different partitions of 4.

  For $1 \leq k \leq n$, let a_k be the number of ordered partitions of n such that n is partitioned into k natural numbers. Then this counts the number of integer solutions to

  $$i_1 + i_2 + ... + i_k = n, \quad \text{where each } i_j \geq 1.$$

  So $a_k = \binom{n-1}{k-1}$.
  Therefore the total number of ordered partitions of n is $\sum_{1 \leq k \leq n} \binom{n-1}{k-1} = 2^{n-1}$.

- From now on, we consider the *unordered partitions*. For instance, we will view $1+2+3$ and $3+2+1$ as the same one.

  Let p_n be the number of partitions of n in this sense.

- Let n_j be the number of the j's in such a partition of n. Then it holds that

  $$\sum_{j \geq 1} j \cdot n_j = n.$$

  If we use i_j to express the contribution of the addends equal to j in a partition of n (i.e., $i_j = j \cdot n_j$), then

  $$\sum_{j \geq 1} i_j = n, \quad \text{where } i_j \in \{0, j, 2j, 3j, ...\}.$$

Note that in the above summation, j can run from 1 to infinity, or run from 1 to n.

So p_n is the coefficient of x_n in the product

$$P_n(x) := (1 + x + x^2 + ...)(1 + x^2 + x^4 + ...)...(1 + x^n + x^{2n} + ...) = \prod_{k=1}^{n} \frac{1}{1 - x^k}.$$

- What is the generating function $P(x)$ of $\{p_n\}$ then?

  As the index j in the summation can be viewed from 1 to $+\infty$, the generating function $P(x)$ is an infinite product of polynomials

$$P(x) = \prod_{k=1}^{+\infty} \frac{1}{1 - x^k}.$$

The Catalan number

- First let us recall the definition of $\binom{r}{k}$ for real number r and positive integer k, and the Newton's binomial Theorem. We obtained that

$$\binom{\frac{1}{2}}{k} = \frac{(-1)^{k-1}2}{4^k} \cdot \frac{(2k-2)!}{k!(k-1)!}.$$

- Let n-gon be a polygon with n corners, labelled as corner 1, corner 2,..., corner n.
- **Definition.** A triangulation of the n-gon is a way to add lines between corners to make triangles such that these lines do not cross inside of the polygon.
- Let b_{n-1} be the number of triangulations of the n-gon, for $n \geq 3$. It is not hard to see that $b_2 = 1, b_3 = 2, b_4 = 5$.
- We want to find the general formula of b_n.
- We consider the triangle T in a triangulation of n-gon which contains corners 1 and 2. The triangle T should contain a third corner, say i. Since $3 \leq i \leq n$, we can divide the set of triangulations of n-gon into cases.

  (1). If $i = 3$ or n, the triangle T divides the n-gon into triangle T itself plus a $(n-1)$-gon, which results in b_{n-2} triangulations of n-gon.

  (2). For $4 \leq i \leq n-1$, the triangle T divides the n-gon into three regions: a $(n-i+2)$-gon, triangle T and a $(i-1)$-gon, therefore it results in $b_{i-2} \times b_{n-i+1}$ many triangulations of n-gon. Therefore, combining (1) and (2), we get that

$$b_{n-1} = b_{n-2} + \sum_{i=4}^{n-1} b_{i-2}b_{n-i+1} + b_{n-2} = b_{n-2} + \sum_{j=2}^{n-3} b_j b_{n-j-1} + b_{n-2}$$

  By letting $b_0 = 0$ and $b_1 = 1$, we get

$$b_{n-1} = \sum_{j=0}^{n-1} b_j b_{n-1-j} \quad \text{or} \quad b_k = \sum_{j=0}^{k} b_j b_{k-j} \quad \text{for} \quad k \geq 2.$$

- Let $f(x) = \sum_{k\geq 0} b_k x^k$. Note that $f^2(x) = \sum_{k\geq 0}\left(\sum_{j=0}^{k} b_j b_{k-j}\right) x^k$. Therefore

$$f(x) = x + \sum_{k\geq 2} b_k x^k = x + \sum_{k\geq 2}\left(\sum_{j=0}^{k} b_j b_{k-j}\right) x^k = x + \sum_{k\geq 0}\left(\sum_{j=0}^{k} b_j b_{k-j}\right) x^k = x + f^2(x).$$

- Solving $f^2(x) - f(x) + x = 0$, we get that $f(x) = \frac{1+\sqrt{1-4x}}{2}$ or $\frac{1-\sqrt{1-4x}}{2}$. But notice that $f(0) = 0$, so it has to be the case that

$$f(x) = \frac{1-\sqrt{1-4x}}{2}.$$

- Next, we apply the Newton's binomial theorem to get that

$$f(x) = \frac{1}{2} - \frac{1}{2}\sum_{k\geq 0}\binom{\frac{1}{2}}{k}(-4x)^k = \sum_{k\geq 1}\frac{(-1)^{k+1}4^k}{2}\binom{\frac{1}{2}}{k}x^k.$$

  After plugging the obtained expression of $\binom{\frac{1}{2}}{k} = \frac{(-1)^{k-1}2}{4^k}\cdot\frac{(2k-2)!}{k!(k-1)!}$, we get that

$$f(x) = \sum_{k\geq 1}\frac{(2k-2)!}{k!(k-1)!}x^k = \sum_{k\geq 1}\frac{1}{k}\binom{2k-2}{k-1}x^k.$$

  Note that $f(x)$ is the generating function of $\{b_k\}$, therefore

$$b_k = \frac{1}{k}\binom{2k-2}{k-1}.$$

- **Theorem.** The total number of triangulations of the $(k+2)$-gon is $\frac{1}{k+1}\binom{2k}{k}$, which is also called **the k^{th} Catalan number**.

- **Definition.** A *binary tree* is a tree which can be recursively defined as follows (we will show more properties about *trees* when we talk about Graph Theory):

  A binary tree either is empty (having no vertex), or consists of one distinguished vertex (called the *root*), and an ordered pair of binary trees (called the *left subtree* and *right subtree*).

- Exercise. Show that the number of binary trees with n vertices is equal to the n^{th} Catalan number.

Random walk

- Imagine the real axis drawn in the plane with integer points marked. A frog leaps among integer points according to the following rules of random walk:

  (1). Initially, the frog sits at 1.

  (2). In each coming step, the frog leaps either by distance 2 to the right (from i to $i+2$), or by distance 1 to the left (from i to $i-1$). It decides one of the two actions independently at random, and each action is of the same probability $\frac{1}{2}$.

- **Problem.** What is the probability that the frog returns to 0?
- In each step, we use "+" and "-" to express the situation that the frog leaps to the right and the left, respectively. Then the probability space Ω can be viewed as the set of infinite vectors, where each coordinate is either + or -.
- Let A be the event that the frog starts at 1 and returns to 0. For $i \geq 1$, let A_i be the event that the frog starts at 1 and reaches 0 at the i^{th} step for the first time.

  Therefore, $A = \cup_{i=1}^{+\infty} A_i$ is a union of disjoint events. Thus, the probability that the frog returns to 0 is $P(A) = \sum_{i=1}^{+\infty} P(A_i)$.
- To computer the value of $P(A_i)$ for $i \geq 1$, we define a_i to be the number of trajectories of the first i steps such that the frog starts at 1 and reaches 0 at the i^{th} step for the first time.

  Note that in the first i steps, there are 2^i trajectories in total (in each step, there are two choices, going left or right); moreover each trajectory occurs with the same probability. Therefore,

$$P(A_i) = \frac{a_i}{2^i}.$$

  Let $f(x) = \sum_{i=0}^{+\infty} a_i x^i$ be the generating function of $\{a_i\}$, where $a_0 := 0$. Then,

$$P(A) = \sum_{i=1}^{+\infty} P(A_i) = \sum_{i=1}^{+\infty} \frac{a_i}{2^i} = f\left(\frac{1}{2}\right).$$

- We now turn to study the generating function $f(x)$.

  Let b_i to be the number of trajectories of the first i steps such that the frog starts at 2 and reaches 0 at the i^{th} step for the first time.

  Let c_i to be the number of trajectories of the first i steps such that the frog starts at 3 and reaches 0 at the i^{th} step for the first time.

  We want to express b_i in terms of $\{a_j\}$. Since the frog only can leap to left by distance 1, if the frog starts at 2 and successfully reaches 0 at the i^{th} step for the first time, then it needs to reach 1 first. Let j be the number of the steps by which the frog reaches 1 for the first time. Then there are a_j trajectories that the frog starts at 2 and reaches 1 at the j^{th} step for the first time. After these j leaps, $i-j$ leaps remain for the frog to move from 1 to 0, and by definition there are a_{i-j} such trajectories that the frog can finish in exactly $i-j$ steps. Observe that j can range from 1 to $i-1$. We derive that

$$b_i = \sum_{j=1}^{i-1} a_j a_{i-j} = \sum_{j=0}^{i} a_j a_{i-j},$$

  where $a_0 = 0$. This implies that $\sum_{i\geq 0} b_i x^i = (\sum_{i\geq 0} a_i x^i)^2 = f^2(x)$.

  Similarly, if we count the number c_i of trajectories from 3 to 0, we can obtain that

$$c_i = \sum_{j=1}^{i-1} a_j b_{i-j} = \sum_{j=0}^{i} a_j b_{i-j},$$

where the second equality is because $a_0 = b_0 = 0$. Therefore,

$$\sum_{i\geq 0} c_i x^i = \left(\sum_{i\geq 0} b_i x^i\right)\left(\sum_{i\geq 0} a_i x^i\right) = f^3(x).$$

- Let us investigate the number a_i of trajectories from 1 to 0 from a different point of view. After the first step, either the frog reaches 0 directly (which shows that $a_1 = 1$), or it leaps to the number 3. In the latter case, it needs to reach 0 in the remaining $i-1$ steps, so it has c_{i-1} trajectories of starting 3 and reaching 0. This shows that for $i \geq 2$, $a_i = c_{i-1}$.

  Now we see that ($c_0 = 0$)

$$f(x) = \sum_{i\geq 0} a_i x^i = x + \sum_{i\geq 2} a_i x^i = x + \sum_{i\geq 2} c_{i-1} x^i = x + x \cdot f^3(x).$$

- Let $a := P(A)$, which is also $f(1/2)$. So $a = \frac{1}{2} + \frac{a^3}{2}$, i.e., $(a-1)(a^2+a-1) = 0$.

  Solving this, we get that $a = 1, \frac{\sqrt{5}-1}{2}$ or $\frac{-\sqrt{5}-1}{1}$. But $P(A) \geq 0$, so $P(A) = 1$ or $\frac{\sqrt{5}-1}{2}$.

  To determine the value of $P(A)$, we consider the inverse function of $f(x)$, that is, $g(x) := \frac{x}{1+x^3}$. Consider the figure of $g(x)$. We find that $g(x)$ is increasing in the interval around $\frac{\sqrt{5}-1}{2}$ but decreasing around 1. Since $f(x) = \sum a_i x^i$ is increasing as x grows, its inverse function $g(x)$ should also be increasing in the region we consider here.

  This explains that the probability $P(A)$ should be $\frac{\sqrt{5}-1}{2}$, which is the golden ratio.

Exponential generating function

Let $\mathbb{N}, \mathbb{N}_e$ and $\mathbb{N}_o$ be the sets of nonnegative integers, nonnegative even integers and nonnegative odd integers, respectively.

- The ordinary generating function of the sequence $\{a_n\}_{n\geq 0}$ is the power series $f(x) = \sum_{n\geq 0} a_n x^n$. We recall the following fact of ordinary generating functions (again!).

  Fact. Let $f_j(x) := \sum_{i\in I_j} x^i$ for $j = 1, 2, ..., n$. And let a_k be the number of integer solutions to $i_1 + i_2 + ... + i_n = k$, where $i_j \in I_j$, that is

$$a_k := \sum_{i_1+i_2+...+i_n=k \text{ for } i_j\in I_j} 1.$$

  Then $\prod_{j=1}^{n} f_j(x)$ is the ordinary generating function of $\{a_k\}$.

- Problem 1. Let S_n be the number of selections of n letters chosen from an unlimited supply of a's, b's and c's such that both of the numbers of a's and b's are even. We can write S_n as

$$S_n = \sum_{e_1+e_2+e_3=n,\ e_1,e_2\in\mathbb{N}_e,\ e_3\in\mathbb{N}} 1.$$

  Using the previous fact, we see that $S_n = [x^n]f$, where

$$f(x) = \left(\sum_{i\in\mathbb{N}_e} x^i\right)^2 \left(\sum_{j\in\mathbb{N}} x^j\right) = \left(\frac{1}{1-x^2}\right)^2 \cdot \frac{1}{1-x}.$$

- We consider a more complicate problem.

  Problem 2. Let T_n be the number of arrangements (or words) of n letters chosen from an unlimited supply of a's, b's and c's such that both of the numbers of a's and b's are even. What is the value of T_n?

- We need the following fact.

  Fact: If we have n letters including x a's, y b's and z c's (i.e. $x+y+z=n$), then we can form $\frac{n!}{x!y!z!}$ distinct words using these n letters.

- Notice that the arrangements of length n can be classified into many cases, desponding on the selection of the n letters, i.e., the numbers of a's, b's and c's used. And each selection of n letters, say x a's, y b's and z c's, should contribute $\frac{n!}{x!y!z!}$ to the count of the total number of arrangements. Therefore, from the above facts, we obtain that

$$T_n = \sum_{e_1+e_2+e_3=n,\ e_1,e_2\in\mathbb{N}_e,\ e_3\in\mathbb{N}} \frac{n!}{e_1!e_2!e_3!}.$$

- **Definition.** The *exponential generating function* for the sequence $\{a_n\}_{n\geq 0}$ is defined to be a power series

$$\sum_{n\geq 0} \frac{a_n}{n!} \cdot x^n.$$

- Similar to the way of defining $f(x)$ (a product of three ordinary generating functions) in Problem 1, we define

$$g(x) := \left(\sum_{i\in\mathbb{N}_e} \frac{x^i}{i!}\right)^2 \left(\sum_{j\in\mathbb{N}} \frac{x^j}{j!}\right)$$

  to be a product of three exponential generating functions, raised from the possible numbers of letters a, b, c.

  We will prove that the so-defined $g(x)$ is the exponential generating function of the sequence $\{T_n\}$. That is:

- **Theorem.**

$$[x^n]g = \frac{T_n}{n!}.$$