

第三章 普通型生成函数计数

(杜玉洁 翻译)

本章引入计数工具中最重要的技巧：生成函数。Wilf 关于生成函数的性质写了一整本书。现有数种不同类型的生成函数，我们从最简单的普通型生成函数开始。在第四、七、八章我们将处理其他类型的生成函数。对于所有类型生成函数的研究，基本的观点都是拿出一列我们感兴趣的数，然后用代数对象，也即多项式或形式幂级数来替代它。这样做的好处是，我们可以用代数对象所具有的性质来研究初始序列。这使得序列的结果的证明有如下优点：

- (1) 证明会非常短。
- (2) 许多证明可以直接计算而不需要其他的方法。
- (3) 有时没有其他已知方法来获取给定的结果。

3.1 生成多项式

令 x 为变量。复数序列

$$a_0, a_1, a_2, \dots, a_n \tag{3.1}$$

有普通生成多项式

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{k=0}^n a_kx^k.$$

在这里，“普通”是为了区别于其他类型的生成多项式。因为我们只在这一章中处理普通生成多项式的情况，我们经常去掉形容词。注意 $f(x)$ 是复数域上一元多项式环 $\mathbb{C}[x]$ 中的元素， x 的系数为复数。我们也称 $f(x)$ 为序列 (3.1) 的生成函数，因为它是可数项序列（但是可能非有限项）生成函数的特殊情况。这种更一般的设定将在 3.3 节中讨论。

我们从一个简单的例子开始讨论，考虑帕斯卡三角形的一行中的二项式系数

$$\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \dots, \binom{n}{n}.$$

相应的生成函数为

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

特别地, 当 $n = 4$ 时, 我们有

$$f(x) = 1 + 4x + 6x^2 + 4x^3 + x^4 = (1 + x)^4.$$

这就是众所周知的二项式定理. 我们给出这一结果的两种证明, 一种是组合方法, 另一种是代数运算.

定理 3.1.1 (二项式定理). 对于 $n \in \mathbb{N}$, 有

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

证明. (组合证明) 用乘法分配律展开乘积

$$(1 + x)^n = \overbrace{(1 + x)(1 + x) \cdots (1 + x)}^n.$$

我们得到 x^k 项是通过选取 k 项 x 和 $n - k$ 项 1 . 但是从 n 个对象中选 k 个的方法数为 $\binom{n}{k}$, 这就是乘积中 x^k 项系数. \square

证明. (代数证明) 我们对 n 做归纳. 当 $n = 0$ 时等式显然成立. 因此假设 $n \geq 1$. 注意, 根据二项式系数的性质, 我们可以将生成函数写作

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

这样做的好处在于不用处理求和的界. 现在利用定理 1.3.3(a) 中关于二项式递归, 改变符号, 归纳得

$$\begin{aligned} \sum_k \binom{n}{k} x^k &= \sum_k \binom{n-1}{k-1} x^k + \sum_k \binom{n-1}{k} x^k \\ &= x \sum_k \binom{n-1}{k-1} x^{k-1} + \sum_k \binom{n-1}{k} x^k \\ &= x \sum_k \binom{n-1}{k} x^k + \sum_k \binom{n-1}{k} x^k \\ &= x(1 + x)^{n-1} + (1 + x)^{n-1} \\ &= (1 + x)^n \end{aligned}$$

\square

第一个证明展示了权重生成函数的乘法规则的使用, 我们将在 3.4 节讨论. 第二个证明表明有关生成函数的计算可以基于传统运算法则. 并且, 扩展求和域的技巧是我们经常使用以简便运算的方法之一. 我们现在想给出一个例子来展示一个生成函数是如何产生的, 到如何使用生成函数来给出其他结果的简单证明. 特别地, 在二项式定理中令 $x = 1$ 就得到

$$\sum_{k=0}^n \binom{n}{k} = (1 + 1)^n = 2^n,$$

这就是定理 1.3.3 的 (c). 类似地, 我们令 $x = -1$, 就有

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

这就是定理 1.3.3 的 (d).

我们通过介绍第一类斯特林数的生成函数来结束本节. 这一结果的证明与二项式定理的代数证明类似, 留作习题. 第二类斯特林数的生成函数将在 3.3 节介绍形式幂级数之后讨论.

定理 3.1.2. 对于 $n \in \mathbb{N}$, 我们有

$$\sum_{k=0}^n c(n, k) x^k = x(x+1)(x+2) \dots (x+n-1).$$

在等式左边, 令 $x = 1$ 就得到特殊情况

$$\#P([n]) = \sum_k c(n, k) = n!.$$

因此该命题可以看做是定理 1.2.1 的一般化. 这样的扩展被称为 q -模拟, 将在下一节讨论.

3.2 统计量与 q -模拟

一种构造生成函数的方法是使用统计量与 q -模拟. 因为超几何级数理论与变量 q 的关系通常被用于这些生成函数. 这是一个助记符选择, 因为有时, 正如我们将在下面看到的, q 代表素数 p 的幂. 目前还没有对 q -模拟的正式定义, 因此我们将用一个例子开始, 将说明我们最终给出的元定义.

集合 S 上的统计量就是 $S \rightarrow \mathbb{N}$ 的函数 st . 因为统计量的值域为 \mathbb{N} , 因此我们可以为有限集合 S 定义相应的生成多项式

$$f(q) = \sum_{s \in S} q^{\text{st}s}.$$

这一生成函数有时被称为 st 在 S 上的分布, 因为它也可以被写作

$$f(q) = \sum_{k \geq 0} a_k q^k$$

其中 a_k 为 $s \in S$ 中满足 $\text{st } s = k$ 的元素个数, 这与概率论中随机变量的分布相似. 排列中最著名的统计量之一是逆序数. 排列 $\pi = \pi_1 \dots \pi_n \in P([n])$ 有逆序集

$$\text{Inv } \pi = \{(i, j) \mid i < j \text{ and } \pi_i > \pi_j\}.$$

我们可以认为该集合是指数对的集合, 其中 π 的相应元素并非它们的自然增长序. 注意, 我们使用指数对而非 π 的元素, 因为这可以更容易地将这个概念推广到允许重复的词. 比如说, 如果 $\pi = \pi_1\pi_2\pi_3\pi_4\pi_5 = 41532$, 那么

$$\text{Inv } \pi = \{(1, 2), (1, 4), (1, 5), (3, 4), (3, 5), (4, 5)\}.$$

π 的逆序数就是

$$\text{inv } \pi = \# \text{Inv } \pi.$$

我们经常使用最开始函数的约定, 首字母大写代表集合, 小写代表集合的基数. 继续我们的例子, $\text{inv } 41532 = 6$. 显然 $\text{inv}: P([n]) \rightarrow \mathbb{N}$ 是一个统计量, 并且它有一个非常有趣的生成多项式.

定理 3.2.1. 对于 $n \geq 0$, 我们有

$$\sum_{\pi \in P([n])} q^{\text{inv } \pi} = (1)(1+q)(1+q+q^2) \cdots (1+q+q^2+\cdots+q^{n-1}).$$

证明. 我们对 n 作归纳, 并且省去琐碎的基本情况. 对每一个 $\pi \in P([n])$, 可以通过 $\sigma \in P([n-1])$ 获得, 只需在 σ 的元素之间形成的 n 个空 (包括 σ_1 的前面和 σ_{n-1} 的后面) 的空) 中插入 n 即可. 令 σ^i 表示把 n 放在距离右边的第 i 个空, 其中 σ_{n-1} 后面的空记为 0. 于是显然有

$$\text{inv } \sigma^i = i + \text{inv } \sigma.$$

用该等式然后进行归纳得到

$$\begin{aligned} \sum_{\pi \in P([n])} q^{\text{inv } \pi} &= \sum_{\sigma \in P([n-1])} \sum_{i=0}^{n-1} q^{\text{inv } \sigma^i} \\ &= \sum_{\sigma \in P([n-1])} q^{\text{inv } \sigma} \cdot \sum_{i=0}^{n-1} q^i \\ &= (1+q)(1+q+q^2) \cdots (1+q+q^2+\cdots+q^{n-1}). \end{aligned}$$

□

将 $q = 1$ 代入此结果得到

$$\#P([n]) = \sum_{\pi \in P([n])} 1 = n!$$

这就是定理 1.2.1 的第二条内容.

现在我们已经遇到了一些 q -模拟 (尽管我们还没有这样命名), 它们的元定义会更有意义. 一个组合对象 \mathcal{O} 的 q -模拟就是 $\mathcal{O}(q)$ 满足

$$\lim_{q \rightarrow 1} \mathcal{O}(q) = \mathcal{O}.$$

注意 \mathcal{O} 可以是许多东西：一个数，一个定义或者一个定理. 比如 $n \in \mathbb{N}$ 的标准 q -模拟是多项式

$$[n]_q = 1 + q + q^2 + \cdots + q^{n-1} \quad (3.2)$$

显然 $[n]_1 = n$. n 的另一可能的 q -模拟是有理函数 $(1 - q^n)/(1 - q)$. 在这种情况下我们不能直接代入 $q = 1$, 而是必须要取极限. 当然, 商和 $[n]_q$ 在 $q \neq 1$ 时是相等的. 另一 q -模拟是 q -阶乘

$$[n]_q! = [1]_q [2]_q \cdots [n]_q.$$

因此定理 3.2.1 可被重新表述为

$$\sum_{\pi \in P([n])} q^{\text{inv} \pi} = [n]_q!.$$

注意, 有时我们把 $[n]_q$ 写作 $[n]$. 这可能与 $[n]$ 作为集合引起混淆, 因此我们只在清楚知道到底代表着什么时才使用此简便记号. 类似地, 在不致引起混淆时我们也可删除下标 q .

还有另一著名的统计量, 用 $[n]_q!$ 作为其分布. $\pi \in P([n])$ 的降序集为

$$\text{Des } \pi = \{i \mid \pi_i > \pi_{i+1}\}. \quad (3.3)$$

相应的降序数为 $\text{des } \pi = \#\text{Des } \pi$. 并且如果 $i \in \text{Des } \pi$ 当且仅当 $(i, i+1) \in \text{Inv } \pi$. 我们也可以定义升序集 $\text{Asc } \pi$ 和升序数 $\text{asc } \pi$, 只需调转 (3.3) 中不等号的方向. 继续我们之前的例子, 有 $\text{Des } 41532 = \{1, 3, 4\}$ 并且 $\text{des } 41532 = 3$. π 的主要指标为

$$\text{maj } \pi = \sum_{i \in \text{Des } \pi} i.$$

因此 $\text{maj } 41532 = 1 + 3 + 4 = 8$. “主要指数”一词是由多米尼克·福阿塔 [26] 创造的, 以纪念英国军队的一名少校珀西·麦克马洪 (Percy MacMahon), 他首先研究了这一统计数据 [61].

定理 3.2.2. 对于 $n \geq 0$, 我们有

$$\sum_{\pi \in P([n])} q^{\text{maj } \pi} = [n]_q!.$$

证明. 我们像定理 3.2.1 一样证明, 但是现在 σ 的空隙编号方式不同. 我们先给我们给 σ_i 和 σ_{i+1} (当 i 为降序) 之间的空以及 σ_{n-1} 之后的空编号, 从右到左从 0 开始编号. 然后给剩下的空编号, 从左到右用 $\text{des } \sigma + 1, \text{des } \sigma + 2, \dots, n-1$ 进行编号. 例子见证明后.

令 $\sigma^{(j)}$ 表示把 n 放在第 j 个位置后的结果. 我们称

$$\text{maj } \sigma^{(j)} = j + \text{maj } \sigma \quad (3.4)$$

事实上, 如果空 j 是下降位或者在 σ 的最后, 那么插入 n 只会移动 j 下降到右侧并包括给定下降一个位置的右侧. 通过定义主指标, $\text{maj } \sigma$ 总共增加了 j . 如果空 j 是在上升位

或者 σ 的开始, 那么插入 n 增加了一个新的下降位, 并且把下降位移到空的后面某一位置的右边. 容易验证, 如果在空 j 插入 n 导致 $\text{maj } \sigma$ 增加了 j , 那么在空 $j+1$ 插入 n 导致 $\text{maj } \sigma$ 增加 $j+1$. 由归纳法知等式 (3.4) 对每个 j 都成立. 证毕. \square

继续前面的例子 $\sigma = 41532$ 且 $\text{maj } \sigma = 8$, 空的编号顺序如下:

$$_4 4_3 1_5 5_2 3_1 2_0.$$

在每一个空依次插入 6 得到

j	0	1	2	3	4	5
$\sigma^{(j)}$	415326	415362	415632	461532	641532	416532
$\text{maj } \sigma^{(j)}$	8	9	10	11	12	13

有很多排列统计量的分布为 $[n]_q!$, 这些统计数据被 Foata 称为 Mahonian. 可以参考文章 Babson 和 Steingrímsson [3] 的 Mahonian 统计列表.

找到了涉及置换的 q -模拟后, 读者可能会怀疑它们也存在于组合中. 对于整数 $0 \leq k \leq n$, 定义 q -二项式系数或高斯系数为

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_q!}{[k]_q! [n-k]_q!}$$

和往常一样, 如果 $k < 0$ 或 $k > n$, 令这一函数为 0. 比如

$$\begin{aligned} \begin{bmatrix} 4 \\ 2 \end{bmatrix} &= \frac{[4]!}{[2]![2]!} \\ &= \frac{[4][3]}{[2][1]} \\ &= \frac{(1+q+q^2+q^3)(1+q+q^2)}{(1+q)} \\ &= 1+q+2q^2+q^3+q^4 \end{aligned} \tag{3.5}$$

只给出定义并不清楚这是一个 q -多项式还是有理函数. 但是利用归纳法及我们下一个结果就很清楚了. 注意, 这一定理给出了普通二项式系数递归的两个 q -模拟. 这说明了一个一般原则, 即 q -模拟不一定是唯一的, 正如我们在 $[n]_q!$ 的 inv 和 maj 解释中也看到的.

定理 3.2.3. 我们有

$$\begin{bmatrix} 0 \\ k \end{bmatrix}_q = \delta_{0,k}$$

并且对 $n \geq 1$, 有

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

证明. 初始条件是平凡的. 我们将证明 q -二项式的第一个递归, 剩下的留作练习. 使用 q -阶乘的定义并找到一个公分母得到

$$\begin{aligned} q^k \begin{bmatrix} n-1 \\ k \end{bmatrix} + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} &= \frac{[n-1]!}{[k]![n-k]!} (q^k [n-k] + [k]) \\ &= \frac{[n-1]!}{[k]![n-k]!} \cdot [n] \\ &= \begin{bmatrix} n \\ k \end{bmatrix}. \end{aligned}$$

证毕. □

下面我们给出二项式定理 (定理 3.1.1) 的 q -模拟. 令 q, t 为两个变量.

定理 3.2.4. 对于 $n \geq 0$ 我们有

$$(1+t)(1+qt)(1+q^2t) \cdots (1+q^{n-1}t) = \sum_{k=0}^n q^{\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q t^k. \quad (3.6)$$

证明. 我们对 n 作归纳, $n=0$ 的情况很容易验证. 对于 $n > 0$ 的情形, 我们用前面结果的第二个递推关系以及归纳假设有

$$\begin{aligned} \sum_k q^{\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q t^k &= \sum_k q^{\binom{k}{2}} \begin{bmatrix} n-1 \\ k \end{bmatrix}_q t^k + \sum_k q^{\binom{k}{2}+n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q t^k \\ &= (1+t)(1+qt) \cdots (1+q^{n-2}t) + q^{n-1}t \sum_k q^{\binom{k-1}{2}} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q t^{k-1} \\ &= (1+t)(1+qt) \cdots (1+q^{n-2}t) + q^{n-1}t(1+t)(1+qt) \cdots (1+q^{n-2}t) \\ &= (1+t)(1+qt) \cdots (1+q^{n-1}t), \end{aligned}$$

这就是我们想要的结果. □

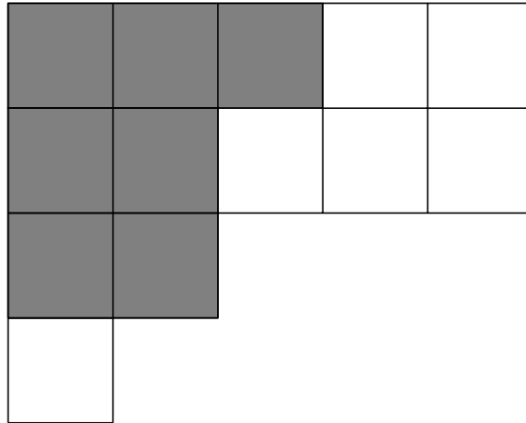
q -二项式系数有许多种组合解释, 我们在这里展示其中两种. 假设 $\lambda = (\lambda_1, \dots, \lambda_k)$ 何 $\mu = (\mu_1, \dots, \mu_l)$ 是整数分拆, 如果 $k \geq l$ 且当 $i \leq l$ 时有 $\lambda_i \geq \mu_i$. 则称 λ 包含 μ , 记作 $\lambda \supseteq \mu$. 换句话说, λ 的杨图包含 μ 的杨图, 如果适当放置使它们的西北角对齐. 比如说, $(5, 5, 2, 1) \supseteq (3, 2, 2)$, 见图 3.1, 阴影部分是 μ 的杨图, 包含在 λ 的杨图中. 符号 $\mu \subseteq \lambda$ 应该是不言自明的. 给定 $\mu \subseteq \lambda$, 也有相应的倾斜分拆

$$\lambda/\mu = \{(i, j) \in \lambda \mid (i, j) \notin \mu\}. \quad (3.7)$$

图 3.1 中倾斜分拆的单元格是白色的.

$k \times l$ 矩形是整数分拆, 其重数表示法为 (k^l) . 考虑包含在这一矩形的分拆集合

$$\mathcal{R}(k, l) = \{\lambda \mid \lambda \subseteq (k^l)\}$$

图 3.1: $(5, 5, 2, 1) \supseteq (3, 2, 2)$ 的杨图

$|\lambda|$ 是分拆 λ 的部分数, 我们考虑生成函数 $\sum_{\lambda \in \mathcal{R}(k,l)} q^{|\lambda|}$. 比如, 如果 $k = l = 2$, 那么我们有

$$\frac{\lambda \subseteq (2^2)}{q^{|\lambda|}} \parallel \begin{array}{c|c|c|c|c|c} \emptyset & (1) & (2) & (1^2) & (2, 1) & (2^2) \\ \hline 1 & q & q^2 & q^2 & q^3 & q^4 \end{array},$$

因而

$$\sum_{\lambda \in \mathcal{R}(2,2)} q^{|\lambda|} = 1 + q + 2q^2 + q^3 + q^4.$$

读者会注意到这一结果与 (3.5) 类似, 这并非巧合.

定理 3.2.5. 对于 $k, l \geq 0$, 我们有

$$\sum_{\lambda \in \mathcal{R}(k,l)} q^{|\lambda|} = \begin{bmatrix} k+l \\ k \end{bmatrix}_q.$$

证明. 对 k 作归纳, $k = 0$ 的情况留给读者. 如果 $k > 0$ 且 $\lambda \subseteq (k^l)$, 则有两种可能. 一种情况是 $\lambda_1 < k$ 并且 $\lambda \subseteq ((k-1)^l)$. 另一种情况是 $\lambda_1 = k$, 此时 λ 可以写作 $\lambda = (k, \lambda')$, 其中 λ' 是 λ 的不包含 λ_1 的分拆. 因此 $\lambda' \subseteq (k^{l-1})$. 注意在这种情况下 $|\lambda| = |\lambda'| + k$. 我们用归纳法及定理 3.2.3 得到

$$\begin{aligned} \sum_{\lambda \in \mathcal{R}(k,l)} q^{|\lambda|} &= \sum_{\lambda \in \mathcal{R}(k-1,l)} q^{|\lambda|} + \sum_{\lambda' \in \mathcal{R}(k,l-1)} q^{|\lambda'|+k} \\ &= \begin{bmatrix} k+l-1 \\ k-1 \end{bmatrix} + q^k \begin{bmatrix} k+l-1 \\ k \end{bmatrix} \\ &= \begin{bmatrix} k+l \\ k \end{bmatrix}, \end{aligned}$$

证毕. □

对于第二种高斯多项式的组合解释, 需要一些线性代数的工具. 令 q 为素幂, 令 \mathbb{F}_q 为具有 q 个元素的伽罗瓦域. 令 V 为 \mathbb{F}_q 上维数 $\dim V = n$ 的向量空间. 元素. 我们用 $W \leq V$ 表示 W 是 V 的子空间. 令

$$\left[\begin{array}{c} V \\ k \end{array} \right] = \{W \leq V \mid \dim W = k\}.$$

维数为 k 的子空间与 $k \times n$ 的满秩的行最简形矩阵成一一对应. 比如说, 如果 $n = 4$, $k = 2$, 那么可能的矩阵为

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & * & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{bmatrix},$$

$$\begin{bmatrix} 1 & * & * & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * & 0 & * \\ 0 & 0 & 1 & * \end{bmatrix}, \begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \end{bmatrix},$$

其中 $*$ 表示 \mathbb{F}_q 的任意元素. 所以相应星图之一的子空间的数目是 q^s , 其中 s 是星星的数量. 因此

$$\# \left[\begin{array}{c} \mathbb{F}_q^4 \\ 2 \end{array} \right] = 1 + q + 2q^2 + q^3 + q^4,$$

在这一点上看来非常相似. 但是请注意, 与之前的情况相反, 这实际上代表一个整数而不是多项式, 因为 q 是素幂. 当然, 这个例子是概括性的. 因为这个结果, 人们有时半开玩笑地称伽罗瓦域上只有一个元素的集合成为向量空间是不存在的.

定理 3.2.6. 如果 V 是 \mathbb{F}_q 维数为 n 的向量空间, 那么

$$\# \left[\begin{array}{c} V \\ k \end{array} \right] = \left[\begin{array}{c} n \\ k \end{array} \right]_q.$$

证明. 给定 $W \leq V$ 使得 $\dim W = k$, 我们首先确定 W 的可能有序基 $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ 的数目. 注意, 因为 $\dim V = n$, 我们有 $\#V = \#\mathbb{F}_q^n = q^n$. 我们可以为 \mathbf{v}_1 选出任意非零向量, 因此可以选择的数目为 $q^n - 1$. 对于 \mathbf{v}_2 , 我们可以在 V 中选择不由 \mathbf{v}_1 生成的任意向量, 有 $q^n - q$ 种可能选择. 用这种方式继续, 计数得

$$(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{k-1}).$$

通过类似的论证, 生成维数为 k 的空间 W 的不同有序基的数目为

$$(q^k - 1)(q^k - q)(q^k - q^2) \dots (q^k - q^{k-1}).$$

因此可能的子空间 W 的数目为

$$\begin{aligned} \frac{(q^n - 1)(q^n - q) \cdots (q^n - q^{k-1})}{(q^k - 1)(q^k - q) \cdots (q^k - q^{k-1})} &= \frac{q^{\binom{k}{2}}(q^n - 1)(q^{n-1} - 1) \cdots (q^{n-k+1} - 1)}{q^{\binom{k}{2}}(q^k - 1)(q^{k-1} - 1) \cdots (q - 1)} \\ &= \frac{(q - 1)^k [n][n-1] \cdots [n-k+1]}{(q - 1)^k [k][k-1] \cdots [1]} \\ &= \begin{bmatrix} n \\ k \end{bmatrix}, \end{aligned}$$

证毕. □

由于 Knuth [50] 使用了前一个示例中的行缩减梯形矩阵来证明这一结果. 读者在练习中给出详细计算证明.

3.3 形式幂级数的代数

我们现在希望将生成函数的概念从有限序列推广到可数无限序列. 为此, 我们将不得不使用幂级数. 但我们希望避免使用解析幂级数时出现的收敛问题. 因此, 我们将研究形式幂级数的代数. 这将意味着我们必须小心, 因为在代数中, 只允许有限次数的加法或乘法. 但是还有收敛的另一个概念会考虑这个问题. 我们应该注意到有一个组合学的分支, 它使用分析技术从相应幂级数的序列提取有用的信息, 例如其增长率. 有关此方法的信息, 请参阅 Flajolet 和 Sedgewick 的书 [25].

形式幂级数形如

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots = \sum_{n=0}^{\infty} a_nx^n,$$

其中 a_n 是复数. 我们也称 $f(x)$ 是序列 $a_n, n \geq 0$ 的普通型生成函数. 在本章中我们通常省去形容词“普通”, 因为我们没有遇到其他类型的生成函数.

注意, 这些级数被认为是形式的, 因为 x 的幂只是占位符, 不允许用值替换 x . 在这样定义下, 解析收敛不是问题, 对于 $\sum_{n \geq 0} n!x^n$, 除了在 $x = 0$ 处外没有其他点收敛, 我们也可讨论形式幂级数. 我们用符号

$$\mathbb{C}[[x]] = \left\{ \sum_{n \geq 0} a_nx^n \mid a_n \in \mathbb{C} \text{ for all } n \geq 0 \right\}$$

表示形式幂级数的集合. 这个集合是代数, 称为形式幂级数的代数, 加法、数乘、乘法规则定义如下

$$\begin{aligned} \sum_{n \geq 0} a_nx^n + \sum_{n \geq 0} b_nx^n &= \sum_{n \geq 0} (a_n + b_n)x^n, \\ c \sum_{n \geq 0} a_nx^n &= \sum_{n \geq 0} (ca_n)x^n, \\ \sum_{n \geq 0} a_nx^n \cdot \sum_{n \geq 0} b_nx^n &= \sum_{n \geq 0} c_nx^n, \end{aligned}$$

其中 $c \in \mathbb{C}$ 且

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

读者可能会反对, 正如前面提到的, 在代数中只允许有限数量的加法, 而 $C[[x]]$ 的元素似乎涉及无限多个. 但这是一种错觉. 请记住, x 只是一个形式参数, 因此表达式 $\sum_n a_n x^n$ 只是一个助记符, 帮助定义三种代数运算, 特别是乘法. 我们可以简单地把 $\mathbb{C}[[x]]$ 定义为复数向量 (a_0, a_1, a_2, \dots) 的集合, 服从向量加法运算规则

$$(a_0, a_1, a_2, \dots) + (b_0, b_1, b_2, \dots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots),$$

数乘和乘法规则也可类似定义. 值得注意的是, $C[[x]]$ 中的元素只允许加或乘有限次. 因此只能进行改变 x 的给定幂的系数有限次的运算.

给定一个复数序列 a_0, a_1, a_2, \dots , 我们将它与普通型生成函数联系起来

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots \in \mathbb{C}[[x]].$$

我们有时称, 如果合适的话, 这一级数对由 a_n 枚举的对象计数. 就像对生成多项式做的一样, 这样做的原因在于挖掘 $\mathbb{C}[[x]]$ 的性质, 以获取关于原始序列的相关信息. 我们经常把这一生成函数写作 $\sum_n a_n x^n$, 假定指数的范围为 $n \geq 0$.

我们用一个简单的例子开始, 考虑序列 $1, 1, 1, \dots$, 其生成函数为 $\sum_n x^n$. 我们想将其简化为几何级数

$$1 + x + x^2 + \dots = \frac{1}{1-x}. \quad (3.8)$$

但是右边的等式意味着什么? 因为 $1/(1-x)$ 作为有理函数出现, 因此它不属于 $C[[x]]$ 吗? 解决这个难题的方法是记住, 给定代数 A 中的元素 a , a 有可能有一个逆元素, 即元素 a^{-1} 使得 $a \cdot a^{-1} = 1$, 其中 1 是 A 的单位元素. 因此要证 (3.8) 式, 即证 $\sum_n x^n$ 与 $1-x$ 互逆. 由分配律很易证明:

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

这一例子表明, 关于分析幂级数的结果同样适用于形式幂级数, 尽管某些地方需要验证是否成立. 在大多数情况下, 我们都可以这样做. 但明智的做法再举几个例子来说明. 另一例子是 $1/x$ 在 $C[[x]]$ 中没有意义, 因为 x 没有逆. 假设对形式幂级数 $f(x)$, 有 $xf(x) = 1$. 那么左边常数项系数为 0, 右边为 1, 产生矛盾.

另一例子, 对 $n \geq 0$ 考虑序列 $1/n!$. 相应的生成函数为

$$e^x = \sum_{n \geq 0} \frac{x^n}{n!},$$

但是, 问题是 e^x 不是 $C[[x]]$ 的先验元素. 解决问题的方法是把 e^x 定义为这一幂级数的形式代表. 然后, 为了完整, 我们会验证所有通常的指数规则都成立, 例如 $e^{2x} = (e^x)^2$.

在这我们不再验证. 但是我们要指明这一规则不成立的情况. 特别地, 在 $\mathbb{C}[[x]]$ 中, 不能写作

$$e^{1+x} = ee^x.$$

这是因为左边没有定义良好. 确实, 在展开 $\sum_n (1+x)^n/n!$ 时, 计算 x 的给定幂的系数需要无限次做加法, 然而我们已经指出的, 这是不允许的.

尽管我们不会证明本文中形式幂级数所需的每一个分析幂级数等式, 在 $\mathbb{C}[[x]]$ 中有一些关于运算的一般结果也是好的. 首先我们要处理形式幂级数可逆的情形.

定理 3.3.1. 如果 $f(x) = \sum_n a_n x^n$, 那么在 $\mathbb{C}[[x]]$ 中 $f(x)^{-1}$ 存在, 当且仅当 $a_0 \neq 0$.

证明. 一方面, 假设 $f(x)g(x) = 1$, 其中 $g(x) = \sum_n b_n x^n$. 比较两边的常数项得到 $a_0 b_0 = 1$, 因而 $a_0 \neq 0$.

另一方面, 假设 $a_0 \neq 0$. 我们构造逆 $g(x) = \sum_n b_n x^n$, 满足 $f(x)g(x) = 1$. 比较两边 x^n 系数, 有 $a_0 b_0 = 1$, 且当 $n \geq 1$ 时,

$$a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 = 0.$$

因为 $a_0 \neq 0$, 所以 $b_0 = 1/a_0$. 类似地, 当 $n \geq 1$ 时, 我们可以在上式中解出 b_n , 通过给出其递推关系. 这样我们就建立了 $g(x)$.

□

我们关于 e^x 的例子表明我们在替换时要特别注意. 定义 $g(x)$ 在 $f(x) = \sum_n a_n x^n$ 中的替换为

$$f(g(x)) = \sum_{n \geq 0} a_n g(x)^n.$$

上式右边是关于形式幂级数的无限求和, 而不仅仅是形式变量. 为讨论这样的和, 需要在 $\mathbb{C}[[x]]$ 中引入收敛的概念.

下面这一记号对形式幂级数 $f(x)$ 来说非常方便:

$$[x^n] f(x) = \text{在 } f(x) \text{ 中 } x^n \text{ 项系数},$$

也即 a_n . 假设有形式幂级数序列 $f_0(x), f_1(x), f_2(x), \dots$. 如果, 对于任意 n , 幂级数序列中 x^n 项系数趋近于常数, 且该常数等于 $f(x)$ 中 x^n 项系数, 则称这一序列收敛到 $f(x) \in \mathbb{C}[[x]]$, 记作

$$\lim_{k \rightarrow \infty} f_k(x) = f(x).$$

一般地, 给定 n , 存在相应的 K 使得 $[x^n] f_k(x) = [x^n] f(x)$ 对于 $k \geq K$ 都成立. 否则, 我们称序列不收敛或极限不存在.

考虑序列

$$f_0(x) = 1, f_1(x) = 1 + x, f_2(x) = 1 + x + x^2, \dots,$$

因而 $f_k(x) = 1 + x + \cdots + x^k$. 那么该序列有极限, 即

$$\lim_{k \rightarrow \infty} f_k(x) = \sum_{n \geq 0} x^n = \frac{1}{1-x}.$$

对于给定 n , 令 $K = n$. 因此对 $k \geq n$ 有 $[x^n] f_k = [x^n] f_n = 1$. 另一方面, 考虑序列

$$f_0(x) = 1 + x, f_1(x) = 1/2 + x/2, f_2(x) = 1/4 + x/4, \dots,$$

其通项为 $f_k(x) = 1/2^k + x/2^k$. 这一序列在 $\mathbb{C}[[x]]$ 上不收敛, 因为对于任意 n , $[x]f_k(x)$ 对不同的 k 是不等的. 这与该序列收敛到 0 的分析情形矛盾.

像在分析幂级数中一样, 我们用序列的收敛定义级数的收敛. 给定 $f_0(x), f_1(x), f_2(x), \dots$, 如果

$$\lim_{k \rightarrow \infty} s_k(x) = f(x),$$

我们称其和存在且收敛到 $f(x)$, 记作 $\sum_{k \geq 0} f_k(x) = f(x)$, 其中

$$s_k(x) = f_0(x) + f_1(x) + \cdots + f_k(x)$$

为其前 k 项和. 注意, 这一定义域形式幂级数定义一致, 因为给定序列 a_0, a_1, a_2, \dots , 可以令 $f_k(x) = a_k x^k$, 然后证明 $\sum_{k \geq 0} f_k(x) = f(x)$, 其中 $f(x) = \sum_{k \geq 0} a_k x^k$.

为陈述级数收敛准则, 需要定义 $f(x) = \sum_n a_n x^n$ 的最小度. 如果 $f(x) \neq 0$,

$$\text{mdeg } f(x) = \text{满足 } a_n \neq 0 \text{ 的最小 } n.$$

如果 $f(x) = 0$ 则 $\text{mdeg } f(x) = \infty$. 事实证明, 要说明幂级数和收敛, 取整数的极限就足够了.

定理 3.3.2. 给定 $f_0(x), f_1(x), f_2(x), \dots \in \mathbb{C}[[x]]$, 那么 $\sum_{k \geq 0} f_k(x)$ 存在当且仅当

$$\lim_{k \rightarrow \infty} (\text{mdeg } f_k(x)) = \infty.$$

证明. 我们先证必要性, 充分性的证明留作练习. 我们按照 (3.9) 式定义序列 $s_k(x)$. 对于给定的 n , 存在 K 使得

$$[x^n] s_K(x) = [x^n] s_{K+1}(x) = [x^n] s_{K+2}(x) = \cdots.$$

但是对于 $j \geq 0$, 有

$$s_{K+j}(x) = s_K(x) + f_{K+1}(x) + f_{K+2}(x) + \cdots + f_{K+j}(x).$$

因而当 $k > K$ 时有 $[x^n] f_k(x) = 0$. 对于给定的 n , 令 N 为所有 K -值的最大值, 其中 K 是小于或等于 n 的整数. 因而, 当 $n > N$ 时, $\text{mdeg } f_k(x) > n$. 根据实数极限的定义, $\lim_{k \rightarrow \infty} (\text{mdeg } f_k(x)) = \infty$. \square

我们现在对替代的定义很清楚了, 这意味着幂级数的总收敛. 我们可以使用前面的结果给出将一个生成函数替换为另一个生成函数时的简单收敛标准.

定理 3.3.3. 给定 $f(x), g(x) \in \mathbb{C}[[x]]$, 复合 $f(g(x))$ 存在当且仅当

- (1) $f(x)$ 为多项式或者
- (2) $g(x)$ 常数项为 0.

证明. 如果 $f(x)$ 为多项式, 那么 $f(g(x))$ 为有限和, 显然收敛. 因此假设 $f(x) = \sum_n a_n x^n$ 不是多项式. 如果 $g(x)$ 没有常数项, 则 $\text{mdeg } a_n g(x)^n \geq n$. 由前一定理, 收敛的极限为无穷, $f(g(x))$ 是定义良好的.

若 $[x^0]g(x) \neq 0$, 因为 $f(x)$ 不是多项式, 存在无穷多个 n 使得 $a_n \neq 0$. 但是对于这些 n , $\text{mdeg } a_n g(x)^n = 0$. 因此极限不是无穷, 并且 $f(g(x))$ 在 $\mathbb{C}[[x]]$ 中不存在. \square

我们还会发现考虑某些无限乘积很有用. 他们的收敛性和无限和一样. 给定序列 $f_0(x), f_1(x), f_2(x), \dots$ 如果

$$\lim_{k \rightarrow \infty} p_k(x) = f(x),$$

我们称他们的积存在并且收敛到 $f(x)$ 记作 $\prod_{k \geq 0} f_k(x) = f(x)$, 其中

$$p_k(x) = f_0(x)f_1(x) \dots f_k(x).$$

下面结果的证明与定理 3.3.2 非常类似, 将其留给读者.

定理 3.3.4. 令 $f_0(x), f_1(x), f_2(x), \dots$ 为常数项为 0 的幂级数. 则 $\prod_{k \geq 0} (1 + f_k(x))$ 存在当且仅当

$$\lim_{k \rightarrow \infty} (\text{mdeg } f_k(x)) = \infty.$$

\square

让我们通过展示前面的结果如何简单地验证积是否存在来结束本节. 考虑 $\prod_{k \geq 1} (1 + x^k)$. 在这种情况下, $f_k(x) = x^k$ 且 $\text{mdeg } x^k = k$. 因此极限为无穷, 积存在. 我们会在 3.5 节中看到, 我们将在 3.5 节中看到, 它计算具有不同部分的整数分区. 相比之下, $\prod_{k \geq 0} (1 + x/2^k)$ 不收敛因为 $\text{mdeg } x/2^k = 1$.

3.4 普通型生成函数的加法与乘法规则

就像集合一样, 普通型生成函数也有加法规则和乘积规则. 为了说明这些结果, 我们需要权重生成函数的概念. 这种方法使得用组合方式为各种序列构建生成函数成为可能. 我们将其首先应用于更深入探索二项式定理.

令 S 为集合. S 的权重为函数 $\text{wt} : S \rightarrow \mathbb{C}[[x]]$. 大多数情况下, 如果 $s \in S$, 那么 $\text{wt } s$ 将只是反映 s 某些性质的单项式. 比如说, 如果 st 是 S 上的任意统计量, 那么我们可以取 $\text{wt } s = x^{\text{st } s}$. 选一个贯穿此节的更具体的例子, 令 $S = 2^{[n]}$, 对于 $T \in S$, 定义

$$\text{wt } T = x^{|T|} \quad (3.10)$$

给定权重集 S , 我们可以形成相应的权重生成函数

$$f(x) = f_S(x) = \sum_{s \in S} \text{wts}.$$

值得注意的是, 上述求和属于 $\mathbb{C}[[x]]$, 若若此, 称集合 S 为可求和集. 当然, 如果 S 是有限集, 自然是可求和的. 以 $S = 2^{[3]}$ 为例, 有

T	\emptyset	$\{1\}$	$\{2\}$	$\{3\}$	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$	$\{1, 2, 3\}$
wt T	1	x	x	x	x^2	x^2	x^2	x^3

因此

$$f_S(x) = 1 + 3x + 3x^2 + x^3.$$

更一般地, 对于 $S = 2^{[n]}$ 有

$$\begin{aligned} f_S(x) &= \sum_{T \in 2^{[n]}} x^{|T|} \\ &= \sum_{k=0}^n \sum_{T \in \binom{[n]}{k}} x^k \\ &= \sum_{k=0}^n \binom{n}{k} x^k. \end{aligned}$$

我们得出了一行帕斯卡三角形的生成函数.

以下定理使得使用权重生成函数更轻松. 对于加法规则, 如果 S, T 是互不相交的权重集, 那么我们求 $u \in S \uplus T$ 的权重就需要在求 u 在 S 或 T 中的权重, 这分别取决于 $u \in S$ 或 $u \in T$. 对于任意的 S, T , 我们计算 $S \times T$ 的权重即令

$$\text{wt}(s, t) = \text{wts} \cdot \text{wtt}.$$

引理 3.4.1. 令 S, T 为可求和的集合.

(1) (加法规则) 集合 $S \cup T$ 是可求和的. 如果 $S \cap T = \emptyset$, 则

$$f_{S \uplus T}(x) = f_S(x) + f_T(x).$$

(2) (乘法规则) $S \times T$ 是可求和的, 且

$$f_{S \times T}(x) = f_S(x) \cdot f_T(x).$$

证明. (1) 因为 S 是可求和的, 对于给定的 $n \in \mathbb{N}$, 只存在有限个 $s \in S$ 使得 wts 中 x^n 项的系数非零. 这对 T 同样成立. 这就说明 $S \cup T$ 中只有有限个元素满足 x^n 项的系数非零, 意味着 $S \cup T$ 是可求和的. 且

$$f_{S \uplus T}(x) = \sum_{u \in S \uplus T} \text{wt } u = \sum_{u \in S} \text{wt } u + \sum_{u \in T} \text{wt } u = f_S(x) + f_T(x).$$

(2) $S \times T$ 的可求和性留作练习.

$$f_{S \times T}(x) = \sum_{(s,t) \in S \times T} \text{wt}(s, t) = \sum_{s \in S} \text{wt } s \cdot \sum_{t \in T} \text{wt } t = f_S(x) \cdot f_T(x)$$

□

我们可以使用这些规则直接计算各种生成函数. 首先用这种方式重新证明二项式定理. 我们发现求和一边是 $S = 2^{[n]}$ 的权重生成函数. 对于乘积一边, 我们用多重符号重新表示 S . 特别地, 考虑

$$S' = \{T' = (1^{m_1}, 2^{m_2}, \dots, n^{m_n}) \mid m_i = 0 \text{ 或 } 1, \forall i \in [n]\},$$

权重为

$$\text{wt } T' = x^{\sum_i m_i}.$$

很显然我们有双射 $f: S \rightarrow S'$, $f(T) = (1^{m_1}, 2^{m_2}, \dots, n^{m_n})$, 其中

$$m_i = \begin{cases} 0 & \text{若 } i \notin T, \\ 1 & \text{若 } i \in T. \end{cases}$$

(事实上, 这就是定理 1.3.1 的证明中用到的映射.) 此外, 这一双射是保持权重不变的: $\text{wt } f(T) = \text{wt } T$. 考虑一个具体的例子, 如果 $n = 5$ 且 $T = \{2, 4, 5\}$, 那么 $f(T) = (1^0, 2^1, 3^0, 4^1, 5^1)$, $\text{wt } f(T) = x^3 = \text{wt } T$. 使用 S' 的好处在于, 它明显是集合 $\{i^0, i^1\}$ ($i \in [n]$) 的权重乘积, 其中 $\text{wt } i^0 = 1$, $\text{wt } i^1 = x$. 因此我们可以写作 S

$$S' = \{1^0, 1^1\} \times \{2^0, 2^1\} \times \dots \times \{n^0, n^1\} = (1^0 \uplus 1^1) \times (2^0 \uplus 2^1) \times \dots \times (n^0 \uplus n^1),$$

其中对于不同的 a 和 b , 我们用 $a \uplus b$ 表示 $\{a\} \uplus \{b\}$. 将此表达式分别应用引理 3.4.1 中加法和乘法规则, 得

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} x^k &= f_S(x) \\ &= f_{S'}(x) \\ &= (\text{wt } 1^0 + \text{wt } 1^1) (\text{wt } 2^0 + \text{wt } 2^1) \dots (\text{wt } n^0 + \text{wt } n^1) \\ &= (1 + x)^n. \end{aligned}$$

这是应用生成函数的第一个例子, 所以写得比较详细. 然而, 在实践中, 通常比较简洁, 例如, 不区分 S 与 S' 的区别, 因为它们产生相同的权重生成函数, 所以可以认为它们是相同的集合. 我们通常忽略可求和性的验证. 下面考虑一个更实质性的例子, 即负指数的二项式定理.

定理 3.4.2. 如果 $n \in \mathbb{N}$, 那么

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

证明. 求和一边表明我们应该考虑

$$S = \{T \mid T \text{ 是 } [n] \text{ 上的多重集}\},$$

权重函数见 (3.10) 式. 于是

$$f_S(x) = \sum_{T \in S} \text{wt } T = \sum_{k \geq 0} \left(\sum_{T \in \left(\binom{[n]}{k} \right)} x^k \right) = \sum_{k \geq 0} \left(\binom{n}{k} \right) x^k.$$

将集合 S 写作

$$\begin{aligned} S &= \{(1^{m_1}, 2^{m_2}, \dots, n^{m_n}) \mid m_i \geq 0, \forall i \in [n]\} \\ &= (1^0 \uplus 1^1 \uplus 1^2 \uplus \dots) \times (2^0 \uplus 2^1 \uplus 2^2 \uplus \dots) \times \dots \times (n^0 \uplus n^1 \uplus n^2 \uplus \dots), \end{aligned}$$

其权重函数为 $\text{wt } i^k = x^k$. 由引理 3.4.1 得到

$$\begin{aligned} f_S(x) &= (\text{wt } 1^0 + \text{wt } 1^1 + \text{wt } 1^2 + \dots) \cdots (\text{wt } n^0 + \text{wt } n^1 + \text{wt } n^2 + \dots) \\ &= (1 + x + x^2 + \dots)^n \\ &= \frac{1}{(1-x)^n}. \end{aligned}$$

证毕. □

关于这个结果, 有几点需要说明. 首先, 将其与我们的第一个版本的二项式定理进行对比. 在定理 3.1.1 中, 我们对 $[n]$ 的子集计数, 其中不允许重复元素, 且生成函数为 $(1+x)^n$. 在定理 3.4.2 中我们对 $[n]$ 上的多重集计数, 其中元素允许重复, 其生成函数为 $1/(1-x)^n$. 在下节中我们用另一例子继续讨论.

我们还可以使定理 3.4.2 看起来几乎完全像定理 3.1.1. 事实上, 如果 $n \leq 0$, 那么由定理 3.4.2 及等式 (1.6) (用 $t-n$ 替代 n), 得

$$(1+x)^n = \frac{1}{(1-(-x))^{-n}} = \sum_{k \geq 0} \left(\binom{-n}{k} \right) (-x)^k = \sum_{k \geq 0} \binom{n}{k} x^k.$$

这就很像定理 3.1.1, 除了我们有一个无限级数, 而对于正的 n 我们有一个多项式.

二项式定理对任何 $n \in \mathbb{C}$ 都是有意义的, 只要 $|x| < 1$ 使级数收敛. 在 $\mathbb{C}[[x]]$ 中, 可以使 $(1+x)^n$ 对任意有理数 $n \in \mathbb{Q}$ 有意义. 见本章练习 12, 并证明下式:

定理 3.4.3. 对于任意 $n \in \mathbb{Q}$, 有

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

□

3.5 整数分拆再讨论

整数分拆理论是普通型生成函数发挥了核心作用的一部分内容. 在这种情况下和其他情况下, 有必要考虑无限乘积. 但是, 正如我们在上一节中看到的, 我们必须注意这些乘积收敛. 我们对集合有相应的限制可以用来构造权重生成函数. 我们从讨论下一问题开始.

令 S 为权重集. 我们称 S 是有根的, 如果存在元素 $r \in S$ 满足

(1) $\text{wt } r = 1$ 且

(2) 如果 $s \in S - \{r\}$ 则 $\text{wt } s$ 常数项为 0,

此时 r 被称为根. 比如, 定理 3.4.2 的证明中用到的集合 (n^0, n^1, n^2, \dots) , 其根为 $r = n^0$, 因为 $\text{wt } n^0 = 1$ 且当 $k \geq 1$ 时 $\text{wt } n^k = x^k$. 给定有根集的序列 S_1, S_2, S_3, \dots , 其中 S_i 的根为 r_i , 它们的直和定义为

$$\begin{aligned} S_1 \oplus S_2 \oplus S_3 \oplus \dots \\ = \{(s_1, s_2, s_3, \dots) \mid s_i \in S_i (\forall i \in \{1, 2, \dots\}) \text{ 且 } s_i \neq r_i \text{ 只对有限个 } i \text{ 成立}\}. \end{aligned}$$

注意, 如果 S_i 的个数是有限的, 那么它们的直和等于它们的积. 但是当其数量无限时, 根的条件就开始起作用了. 因为这一条件, 我们在 $\bigoplus_{i \geq 1} S_i$ 上有定义良好的权重, 即

$$\text{wt}(s_1, s_2, s_3, \dots) = \prod_{i \geq 1} \text{wt } s_i,$$

因为该乘积只有有限多个不等于 1 的因子. 此外, 引理 3.4.1 中的乘法规则必须适当修改以满足收敛. 但是证明与前面的结果相似, 因而留作练习.

定理 3.5.1. 令 S_1, S_2, S_3, \dots 是可求和的有根集序列, 满足

$$\lim_{i \rightarrow \infty} [\text{mdeg}(f_{S_i}(x) - 1)] = \infty.$$

那么直和 $S_1 \oplus S_2 \oplus S_3 \oplus \dots$ 也是可求和的, 且

$$f_{S_1 \oplus S_2 \oplus S_3 \oplus \dots}(x) = \prod_{i \geq 1} f_{S_i}(x).$$

□

我们现在将证明一个欧拉定理, 给出 $p(n)$ 的生成函数, 即 n 的整数分拆数. 读者应将证明与定理 3.4.2 中给出的计数多重集的证明进行对比, 明显相似.

定理 3.5.2.

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

证明. 根据上式求和一边, 我们考虑 $n \geq 0$ 的整数分拆集合 λ 组成的集合 S , 权重为

$$\text{wt } \lambda = x^{|\lambda|}, \quad (3.11)$$

$|\lambda|$ 为 λ 的部分数. 因而

$$f_S(x) = \sum_{\lambda \in S} \text{wt} \lambda = \sum_{n \geq 0} \sum_{|\lambda|=n} x^n = \sum_{n \geq 0} p(n) x^n.$$

我们用多重集符号将 S 表示成直和的形式, 即

$$\begin{aligned} S &= \{(1^{m_1}, 2^{m_2}, 3^{m_3}, \dots) \mid m_i \geq 0 \forall i \in [n], \text{且只有有限多个 } m_i \neq 0\} \\ &= (1^0 \uplus 1^1 \uplus 1^2 \uplus \dots) \oplus (2^0 \uplus 2^1 \uplus 2^2 \uplus \dots) \oplus (3^0 \uplus 3^1 \uplus 3^2 \uplus \dots) \oplus \dots \end{aligned}$$

请注意, 由于我们希望 $\text{wt} \lambda$ 上的指数是其各部分的总和, 且 i^k 表示部分 i 重复 k 次, 我们取

$$\text{wt} i^k = x^{ik},$$

与定理 3.4.2 证明中使用的权重相反. 由前面的定理将其用生成函数表示得

$$\begin{aligned} f_S(x) &= \prod_{i \geq 1} (\text{wt} i^0 + \text{wt} i^1 + \text{wt} i^2 + \text{wt} i^3 + \dots) \\ &= \prod_{i \geq 1} (1 + x^i + x^{2i} + x^{3i} + \dots) \\ &= \prod_{i \geq 1} \frac{1}{1 - x^i}. \end{aligned}$$

证毕. □

读者会注意到以前的证明实际上说明更多. 特别地, 因子 $1/(1 - x^i)$ 表示 λ 中等于 i 的部分. 下面详细说明.

命题 3.5.3. 给定 $n \in \mathbb{N}$ 且 $P \subseteq \mathbb{P}$, 令 $p_P(n)$ 表示 n 的满足分拆部分属于 P 的分拆个数. 则

(1)

$$\sum_{n \geq 0} p_P(n) x^n = \prod_{i \in P} \frac{1}{1 - x^i}.$$

(2) 特别地, 对 $k \in \mathbb{P}$,

$$\sum_{n \geq 0} p_{[k]}(n) x^n = \frac{1}{(1 - x)(1 - x^2) \cdots (1 - x^k)}.$$

证明. 对 (1), 使用定理 3.5.2 证明中的想法, 除了 S 的元素只包含 r^{m_r} 形式的分量, 其中 $r \in P$. (2) 是 (1) 的直接结果. □

我们可以限制分拆的部分数, 而不是限制分拆的部分的集合. 回顾 $p(n, k)$ 是长度为 $\ell(\lambda) \leq k$ 的 $\lambda \vdash n$ 分拆数.

推论 3.5.4. 对 $k \geq 0$, 有

$$\sum_{n \geq 0} p(n, k) x^n = \frac{1}{(1 - x)(1 - x^2) \cdots (1 - x^k)}.$$

证明. 从前面的结果可以看出, 由 $p_{[k]}(n)$ 计算的分拆与由 $p(n, k)$ 计算的分拆之间存在保持大小的双射. $\lambda \rightarrow \lambda^t$ 就是满足条件的一个映射. 事实上, λ 只使用 $[k]$ 中的部分当且仅当 $\lambda_1 \leq k$. 用杨图来说, 这意味着 λ 的第一行的长度最多为 k . 随之有 λ^t 的第一列长度最多为 k , 相当于 λ^t 有最多 k 部分. \square

在上一节中, 我们指出了集合和多重集合的生成函数之间的关系. 这一关系对整数分拆同样成立. 就像在 2.3 节定义的, $p_d(n)$ 表示将 n 分拆成不同部分的分拆个数.

定理 3.5.5.

$$\sum_{n \geq 0} p_d(n) x^n = \prod_{i \geq 1} (1 + x^i)$$

证明. 到目前为止, 我们一直在写出用权重生成函数证明的大部分细节, 来让读者熟悉该方法. 但是到现在, 仅仅写出重点就足够了. 我们用 S 表示将 $n \in \mathbb{N}$ 分拆成不同部分的所有分拆集合, 并且使用 (3.11) 式中的权重. $f_S(x)$ 是定理的和侧. 为得到乘积一边, 我们写作

$$\begin{aligned} S &= \{(1^{m_1}, 2^{m_2}, 3^{m_3}, \dots) \mid m_i = 0 \text{ 或 } 1, \text{ 只有有限多个 } m_i \neq 0\} \\ &= \bigoplus_{i \geq 1} (i^0 \uplus i^1). \end{aligned}$$

因而生成函数即为

$$f_S(x) = \prod_{i \geq 1} (1 + x^i).$$

证毕. \square

定理 3.5.6 (欧拉). 对所有的 $n \geq 0$,

$$p_o(n) = p_d(n).$$

证明. 只需证这两个序列有相同的生成函数. 由定理 3.5.3 (1) 和定理 3.5.5, 我们有

$$\begin{aligned} \sum_{n \geq 0} p_d(n) x^n &= (1 + x) (1 + x^2) (1 + x^3) \cdots \\ &= (1 + x) (1 + x^2) (1 + x^3) \cdots \frac{(1 - x) (1 - x^2) (1 - x^3) \cdots}{(1 - x) (1 - x^2) (1 - x^3) \cdots} \\ &= \frac{(1 - x^2) (1 - x^4) (1 - x^6) \cdots}{(1 - x) (1 - x^2) (1 - x^3) \cdots} \\ &= \frac{1}{(1 - x) (1 - x^3) (1 - x^5) \cdots} \\ &= \sum_{n \geq 0} p_o(n) x^n. \end{aligned}$$

证毕. \square

3.6 递推关系与生成函数

读者可能已经注意到, 第 1 章中描述的许多组合序列满足递推关系. 如果有一个由递归定义的序列, 则可以用生成函数找到该序列的显式表达式. 对于那些难以从递归本身提取信息的序列, 可通过生成函数收集信息. 本节就来解决这些问题.

我们从一个简单的例子入手, 利用生成函数求递推关系. 给定递归序列 a_0, a_1, a_2, \dots 及初始条件, 我们希望找到第 n 项的表达式.

(1) 将递推同乘以 x^n ; 一般地, n 的最佳选择是递推中所有项的最大指标. 对所有 $n \geq d$ 求和, 其中 d 是使得递推为空的最小指标.

(2) 令

$$f(x) = \sum_{n \geq 0} a_n x^n,$$

使用边界条件用 $f(x)$ 表示步骤 (1) 中的等式.

(3) 解出 $f(x)$.

(4) 求 a_n 作为 x^n 在 $f(x)$ 中的系数.

我们注意到部分分式展开有助于完成步骤 (4).

考虑一个简单的例子, 假设序列定义为 $a_0 = 2$, 且当 $n \geq 1$ 时, $a_n = 3a_{n-1}$. 计算该序列一些前面项的值, 得 $a_1 = 2 \cdot 3, a_2 = 2 \cdot 3^2, a_3 = 2 \cdot 3^3$. 因此, 很容易猜出然后归纳证明得 $a_n = 2 \cdot 3^n$. 我们现在想要用生成函数来解决. 步骤 (1) 很容易, 就是写作 $a_n x^n = 3a_{n-1} x^n$, 然后对其求和得

$$\sum_{n \geq 1} a_n x^n = \sum_{n \geq 1} 3a_{n-1} x^n.$$

由步骤 (2), 令 $f(x) = \sum_{n \geq 0} a_n x^n$, 即得

$$\sum_{n \geq 1} a_n x^n = f(x) - a_0 = f(x) - 2,$$

且

$$\sum_{n \geq 1} 3a_{n-1} x^n = 3x \sum_{n \geq 1} a_{n-1} x^{n-1} = 3xf(x),$$

其中最后一个等式由在和中将 n 替换为 $n-1$ 得到. 由步骤 (3) 得

$$f(x) - 2 = 3xf(x) \implies f(x) - 3xf(x) = 2 \implies f(x) = \frac{2}{1-3x} \quad (3.12)$$

对步骤 (4), 我们可将 $1/(1-3x)$ 展开为几何级数, 也就是利用 (3.8) 式, 将 $3x$ 替换为 x , 我们得到

$$f(x) = 2 \sum_{n \geq 0} 3^n x^n = \sum_{n \geq 0} 2 \cdot 3^n x^n.$$

提取出 x^n 的系数即得 $a_n = 2 \cdot 3^n$.

在前面的例子中, 猜出 a_n 的表达式然后用归纳法证明比生成函数的方法更简单. 但是, 有时候不能猜出表达式, 但仍可用生成函数求递推关系. 我们以斐波那契数列为例.

如果我们使用由 (1.1) 给出的此序列的定义, 我们的结果将略微更好. 遵循算法, 我们写作

$$\sum_{n \geq 2} F_n x^n = \sum_{n \geq 2} (F_{n-1} + F_{n-2}) x^n.$$

令 $f(x) = \sum_{n \geq 0} F_n x^n$ 得

$$\sum_{n \geq 2} F_n x^n = f(x) - F_0 - F_1 x = f(x) - x$$

且

$$\sum_{n \geq 2} (F_{n-1} + F_{n-2}) x^n = x(f(x) - F_0) + x^2 f(x) = (x + x^2) f(x).$$

令左右两边的表达式相等并求解 $f(x)$ 得到

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

对于最后一步, 我们希望使用部分分式, 即因子 $1 - x - x^2$. 使用二次公式, 我们看到分母有根

$$r_1 = \frac{-1 + \sqrt{5}}{2} \quad \text{且} \quad r_2 = \frac{-1 - \sqrt{5}}{2}.$$

因而

$$1 - x - x^2 = \left(1 - \frac{x}{r_1}\right) \left(1 - \frac{x}{r_2}\right).$$

因为两边都在 $x = r_1, r_2$ 处无意义, 并且两边都有常数项 1. 所以我们有下面部分分式分解

$$f(x) = \frac{x}{\left(1 - \frac{x}{r_1}\right) \left(1 - \frac{x}{r_2}\right)} = \frac{A}{\left(1 - \frac{x}{r_1}\right)} + \frac{B}{\left(1 - \frac{x}{r_2}\right)}.$$

A, B 为常数. 约分得

$$x = A \left(1 - \frac{x}{r_2}\right) + B \left(1 - \frac{x}{r_1}\right).$$

将 $x = r_1$ 代入此等式化简得 $r_1 = A(1 - r_1/r_2)$, 解出 A 得 $A = 1/\sqrt{5}$. 类似地, 将 $x = r_2$ 代入得 $B = -1/\sqrt{5}$. 将这些值重新代入 (3.13) 式并展开级数, 得

$$f(x) = \frac{1}{\sqrt{5}} \cdot \sum_{n \geq 0} \frac{x^n}{r_1^n} - \frac{1}{\sqrt{5}} \cdot \sum_{n \geq 0} \frac{x^n}{r_2^n}.$$

通过分母有理化可以检验 $1/r_1 = (1 + \sqrt{5})/2$ 且 $1/r_2 = (1 - \sqrt{5})/2$. 取两边 x^n 系数得

$$F_n = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n \quad (3.14)$$

这个例子展示了生成函数方法的真正威力. 仅通过计算 F_n 的值是不可能猜出 (3.14) 中的表达式的. 实际上, 右边是一个整数甚至都不明显!

我们的算法可用于在以下情况下推导生成函数：有一个三角形数列，而不仅仅是一个序列。在这我们用斯特林数来说明。回顾定理 1.5.2 中第一类无符号斯特林数满足的递推关系及初始条件，将其转化成有符号版本得到 $s(0, k) = \delta_{0,k}$ 且当 $n \geq 1$ 时，

$$s(n, k) = s(n-1, k-1) - (n-1)s(n-1, k).$$

我们希望找到生成函数 $f_n(x) = \sum_k s(n, k)x^k$ 。应用算法得

$$\begin{aligned} f_n(x) &= \sum_k s(n, k)x^k \\ &= \sum_k [s(n-1, k-1) - (n-1)s(n-1, k)]x^k \\ &= xf_{n-1}(x) - (n-1)f_{n-1}(x) \\ &= (x - n + 1)f_{n-1}(x), \end{aligned}$$

这就得到生成函数序列 $f_n(x)$ 的递推关系。由初始条件 $s(0, k)$ 得 $f_0(x) = 1$ 。现在容易猜出 $f_n(x)$ 的表达式，只需写出前面一些值然后用归纳法证明，并得到如下定理，这由定理 3.1.2 很容易得出。

定理 3.6.1. 对于 $n \geq 0$ ，有

$$\sum_k s(n, k)x^k = x(x-1)\cdots(x-n+1).$$

□

以完全类似的方式，可以得到第二类斯特林数的生成函数。证明留给读者。

定理 3.6.2. 对于 $k \geq 0$ ，有

$$\sum_n S(n, k)x^n = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)}.$$

□

比较前两个结果，读者会注意到无重复对象（集合、不同分拆）的生成函数与那些允许重复（多重集、普通分拆）的生成函数比较相似。前面提过了，这将在 3.9 节解释。

到目前为止，我们从递归中导出的所有生成函数都是有理函数。这是因为递归是线性的，我们将在下一节证明一般结果。我们将通过举一些更复杂的例子来结束本节，例如代数函数。我们考虑卡特兰数 $C(n)$ ，其生成函数为 $c(x) = \sum_{n \geq 0} C(n)x^n$ 。利用定理 1.11.2 中的递推关系和初始条件，并用我们习惯的方法计算得

$$c(x) = 1 + \sum_{n \geq 1} C(n)x^n = 1 + \sum_{n \geq 1} \left(\sum_{i+j=n-1} C(i)C(j) \right) x^n = 1 + xc(x)^2.$$

将其写作 $xc(x)^2 - c(x) + 1 = 0$ 并用二次公式解出 $c(x)$ 得

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

$c(x)$ 的这个公式似乎有两个问题. 首先, 我们不知道正号与负号哪个是对的. 其次, 我们似乎遗漏了形式幂级数的环, 因为我们除以 x , 它没有逆. 可以通过选择符号同时解决这两个问题, 使得分子没有常数项. 然后可以简单地通过降低幂次来除以 x . 由定理 3.4.3 得到, $\sqrt{1-4x} = (1-4x)^{1/2}$ 的生成函数有常数项 $\binom{1/2}{0} = 1$. 因此正确的符号应该是负号, 然后我们得到

定理 3.6.3.

$$\sum_{n \geq 0} C(n)x^n = \frac{1 - \sqrt{1-4x}}{2x}$$

□

可以使用这个生成函数来重新导出定理 1.11.3 中 $C(n)$ 的显式表达式, 读者在练习中详细证明.

3.7 有理生成函数与线性递归

读者可能已经注意到, 在前面的章节中, 不管是初始例子还是斐波那契数列, a_n 的递推关系都是形式为 r^n 的函数的线性组合, 其中 r 在相应生成函数分母根的倒数上变化. 这对我们本节研究的递推关系来说都很普遍. 在给出一个刻画这种情况的定理之前, 我们将再研究一个例子来说明会发生什么.

考虑如下定义的序列, $a_0 = 1, a_1 = -4$, 且

$$a_n = 4a_{n-1} - 4a_{n-2} \quad \text{对于 } n \geq 2. \quad (3.15)$$

根据前面四步算法, 对 $f(x) = \sum_{n \geq 0} a_n x^n$, 有

$$\begin{aligned} f(x) - 1 + 4x &= \sum_{n \geq 2} a_n x^n \\ &= 4x \sum_{n \geq 2} a_{n-1} x^{n-1} - 4x^2 \sum_{n \geq 2} a_{n-2} x^{n-2} \\ &= 4x(f(x) - 1) - 4x^2 f(x) \end{aligned}$$

解出 $f(x)$ 并计算部分分式展开的常数得到

$$f(x) = \frac{1-8x}{1-4x+4x^2} = \frac{1-8x}{(1-2x)^2} = \frac{4}{1-2x} - \frac{3}{(1-2x)^2}.$$

两边拿出 x^n 项系数, 利用定理 3.4.2 (互换 n 和 k 的位置), 以及如下等式

$$\left(\binom{k}{n} \right) = \binom{n+k-1}{n} = \binom{n+k-1}{k-1} \quad (3.16)$$

得到

$$a_n = 4 \cdot 2^n - 3 \binom{n+1}{1} 2^n = (1-3n)2^n \quad (3.17)$$

现在, 我们的用 n 的多项式而非常数幂 r^n 作为系数的多项式. 并且多项式的次数小于作为分母根的 $1/r$ 的重数.

考虑复数序列 a_n , 其中 $n \geq 0$. 我们称序列满足带有常数项的度为 d 的齐次线性递归, 即存在 $d \in \mathbb{P}$ 及常数 $c_1, \dots, c_d \in \mathbb{C}$, $c_d \neq 0$ 使得

$$a_{n+d} + c_1 a_{n+d-1} + c_2 a_{n+d-2} + \dots + c_d a_n = 0. \quad (3.18)$$

为了稍后简化, 我们将递归的所有项放在左侧使 a_{n+d} 而非 a_n 成为指数最高的项. 我们也可考虑非齐次的情况, 其中有一个总和 c_{d+1} , 它没有乘以序列的任何项, 但我们在这里暂不考虑此情形. 满足递推关系 (3.18) 的序列就是有理生成函数的序列.

定理 3.7.1. 给定序列 a_n , 其中 $n \geq 0$ 及 $d \in \mathbb{P}$, 则下列命题等价.

(1) 序列满足 (3.18) 式.

(2) 生成函数 $f(x) = \sum_{n \geq 0} a_n x^n$ 具有形式

$$f(x) = \frac{p(x)}{q(x)}, \quad (3.19)$$

其中

$$q(x) = 1 + c_1 x + c_2 x^2 + \dots + c_d x^d \quad (3.20)$$

且 $\deg p(x) < d$.

(3) 可写作

$$a_n = \sum_{i=1}^k p_i(n) r_i^n,$$

其中 r_i 为互不相同的非负复数, 满足

$$1 + c_1 x + c_2 x^2 + \dots + c_d x^d = \prod_{i=1}^k (1 - r_i x)^{d_i}, \quad (3.21)$$

对每个 i , $p_i(n)$ 是度 $\deg p_i(n) < d_i$ 的多项式.

证明. 我们首先证明 (1) 与 (2) 等价. 证明 (1) 意味着 (2) 本质上是我们算法的应用. (3.18) 式乘以 x^{n+d} 并对 $n \geq 0$ 求和可得

$$\begin{aligned} 0 &= \sum_{n \geq 0} a_{n+d} x^{n+d} + c_1 x \sum_{n \geq 0} a_{n+d-1} x^{n+d-1} + \dots + c_d x^d \sum_{n \geq 0} a_n x^n \\ &= \left[f(x) - \sum_{n=0}^{d-1} a_n x^n \right] + c_1 x \left[f(x) - \sum_{n=0}^{d-2} a_n x^n \right] + \dots + c_d x^d f(x) \\ &= q(x) f(x) - p(x), \end{aligned}$$

其中 $q(x)$ 由 (3.20) 给出, $p(x)$ 是剩余项的求和, 意味着 $\deg p(x) < d$. 解出 $f(x)$ 即证出这一方向.

为证 (2) 包含 (1), 对 (3.19) 做交叉乘法并使用 (3.20) 写作

$$p(x) = q(x)f(x) = (1 + c_1x + c_2x^2 + \cdots + c_dx^d)f(x).$$

因为 $\deg p(x) < d$, 我们有 $[x^{n+d}]p(x) = 0$, 对任意 $n \geq 0$ 成立. 拿出上式两边 x^{n+d} 项系数即得递推关系 (3.18).

我们现在证明 (2) 和 (3) 等价. (2) 包含 (3) 由算法得出. 特别地, 使用 (3.19), (3.20) 及 (3.21), 以及部分分式展开, 可得

$$f(x) = \frac{p(x)}{\prod_{i=1}^k (1 - r_ix)^{d_i}} = \sum_{i=1}^k \sum_{j=1}^{d_i} \frac{A_{i,j}}{(1 - r_ix)^j}, \quad (3.22)$$

其中 $A_{i,j}$ 是确定常数. 但是由定理 3.4.2 及等式 (3.16), 我们有

$$[x^n] \frac{1}{(1 - r_ix)^j} = \left(\binom{j}{n} \right) r_i^n = \binom{n+j-1}{j-1} r_i^n,$$

其中

$$\binom{n+j-1}{j-1} = \frac{(n+j-1)(n+j-2)\cdots(n+1)}{(j-1)!}$$

是关于 n 的度为 $j-1$ 的多项式, 现在取 (3.22) 式两边 x^n 系数得

$$a_n = \sum_{i=1}^k \left[\sum_{j=1}^{d_i} A_{i,j} \binom{n+j-1}{j-1} \right] r_i^n.$$

称方括号内的多项式为 $p_i(n)$, 就得到我们想要的展开式.

(3) 包含 (2) 的证明留作练习. □

我们注意到, 上述定理不仅具有理论意义, 而且在计算上非常有用. 特别是, 由于 (1) 和 (3) 的等价性, 可以以更直接的方式解决线性、常数系数递归, 而无需必须使用生成函数. 假设我们有一个满足 (3.18) 式的序列. 然后我们知道 (3) 中的解是关于 r_i , 它们是由 (3.20) 给出的 $q(x)$ 的根的倒数. 为简化, 我们考虑多项式

$$r(x) = x^d q(1/x) = x^d + c_1x^{d-1} + c_2x^{d-2} + \cdots + c_d.$$

与 (3.21) 比较表明 r_i 是 $r(x)$ 的根. 我们现在通过使用初始条件求解这些多项式的系数来找到 $p_i(n)$. 具体来说, 再次考虑我们开始本节的例子 (3.15). 因为当 $n \geq 2$ 时, $a_n - 4a_{n-1} + 4a_{n-2} = 0$, 我们考虑 $r(x) = x^2 - 4x + 4 = (x-2)^2$. 因此 $a_n = p(n)2^n$, 其中 $\deg p(n) < 2$. 因而 $p(n) = A + Bn$, A, B 是常数. 将 $n = 0$ 代入得 $1 = a_0 = A2^0 = A$. 再令 $n = 1$ 得 $-4 = a_1 = (1+B)2^1$ 或 $B = -3$. 因此 a_n 又如 (3.17) 给出. 但这个解显然比第一个解更简单. 这种方法称为待定系数法. 当然, 使用生成函数的最大好处是它们可以用来解决递归, 即使在它们不是线性常系数.

在本节的理论与待定系数法求解常系数线性微分方程之间存在着惊人的相似之处. 这不是偶然, 本节中的材料可以被认为是有限差异理论的一部分, 这与离散的微分方程理论类似. 我们在 5.5 节学习莫比乌斯反演时会详细讨论.

3.8 色多项式

有时生成函数或多项式会以意想不到的方式出现. 我们现在用图的色多项式说明这一现象.

令 $G = (V, E)$ 为图. G 的用颜色集合 S 对 (顶点) 着色就是函数 $c: V \rightarrow S$. 图 3.2 包含一个图表, 它用集合 $S = \{\text{白, 灰, 黑}\}$ 进行两种着色, 我们用该图表继续后面的讨论. 我们称 c 是正常的, 如果对所有的边 $uv \in E$, 有 $c(u) \neq c(v)$. 图 3.2 中第一个着色是正常的, 然而第二个不是, 因为边 vx 的两个端点都染灰色. G 的色数, 用 $\chi(G)$ 表示, 是集合 S 的满足存在一个正常着色 $c: V \rightarrow S$ 的最小基数. 在我们的例子中 $\chi(G) = 3$, 因为我们在图 3.2 中展示了用 3 种颜色的正常着色, 我们不能用少于 3 种颜色进行正常着色, 因为存在三角形 uvx .

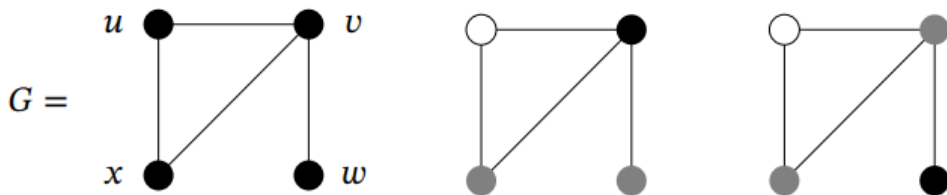


图 3.2: 一个图与两种着色

色数是图论中一个重要的不变量. 但是由它的定义, 它更多地属于极值组合学 (研究最小化或最大化约束的结构), 而不是枚举方面. 虽然在这关于 $\chi(G)$ 我们没什么可多说的, 但是我们得说这是最著名的数学定理之一. 称一个图为平面图, 如果它可以画在平面上, 没有任何两条边交叉.

定理 3.8.1 (四色定理). 如果 G 是平面图, 则

$$\chi(G) \leq 4.$$

□

请注意, 此结果与可以具有任意大色数的普通图形成鲜明对比. 例如, 对完全图有 $\chi(K_n) = n$. 四色定理在 1977 年被 Appel 和 Haken 证明时 (在科赫的帮助下) 引起了不小的轰动 [1,2]. 一方面, 据四色猜想被提出已经过去了 100 多年. 他们的证明也是第一个大量使用计算机来完成, 所有各种情况的计算和演示不能完全由人工检查.

我们现在转向枚举图着色. 令 $t \in \mathbb{N}$. G 的色多项式定义为

$$P(G; t) = \text{正常着色数 } c: V \rightarrow [t].$$

这个概念是由 George Birkhoff [13] 提出的. 目前尚不清楚为什么 $P(G; t)$ 应该称为多项式, 让我们用图 3.2 中的图形计算它. 按顺序 u, v, w, x 考虑顶点着色. 顶点 u 的着色有 t 种选择. 因为 v 不能与 u 染相同的颜色, 于是 v 有 $t - 1$ 种选择. 同理, w 的选择数为 $t - 1$. 最后, x 的选择方法数为 $t - 2$, 因为它不能和 u 与 v 染相同的颜色. 因此, 最终的计数为

$$P(G) = P(G; t) = t(t - 1)(t - 1)(t - 2) = t^4 - 4t^3 + 5t^2 - 2t. \quad (3.23)$$

这是一个关于 t 的多项式, 是染色的方法数! 在证明之前, 我们有一些说明. 首先 $P(G; t)$ 和 $\chi(G)$ 关系密切, 即当 $0 \leq t < \chi(G)$, $P(G; t) = 0$, 但是 $P(G; \chi(G)) > 0$. 这是由 P 和 χ 的定义得来的, 因为后者是 G 的正常着色存在的最小的非负整数, 前者是对这些着色方法数的计数. 其次, 我们并不总能够用上面方法计算出 $P(G; t)$, 并对整数 k 将其表达为因子 $t - k$ 的乘积. 例如, 考虑圈 C_4 , 顶点按顺时针标记为 u, v, w, x . 如果我们试着用这种方式计算 $P(C_4; t)$, 除了顶点 x 的着色外, 其他都很简单, 因为 x 与 u 和 w 都相邻. 但是我们不能确定是否 u 和 w 着色相同, 因为它们本身不相邻.

事实证明, 相同的想法可以用于证明 $P(G; t)$, 并降低计算 $P(C_4; t)$ 的难度. 考虑图 $G = (V, E)$ 及一条边 $e \in E$. 这个图是由 G 删去边 e 得到的, 记作 $G \setminus e$, 它的顶点为 V 且边集为 $E - \{e\}$. 图 3.3 中中间的图就是例子中删去边 $e = vx$ 得到的. 在 G 中收缩一条边 e 表示为 G/e , 它是通过收缩 e 到一个新的顶点 v_e , 使得 v_e 与 e 的端点相邻, 其他顶点和边不变. 在我们的示例图中收缩 vx 得到图 3.3 右侧的图. 下一个引理在 $P(G; t)$ 的研究中至关重要. 理想情况下, 对 $\#E$ 作归纳, 因为 $G \setminus e$ 和 G/e 的边都比 G 的边少.

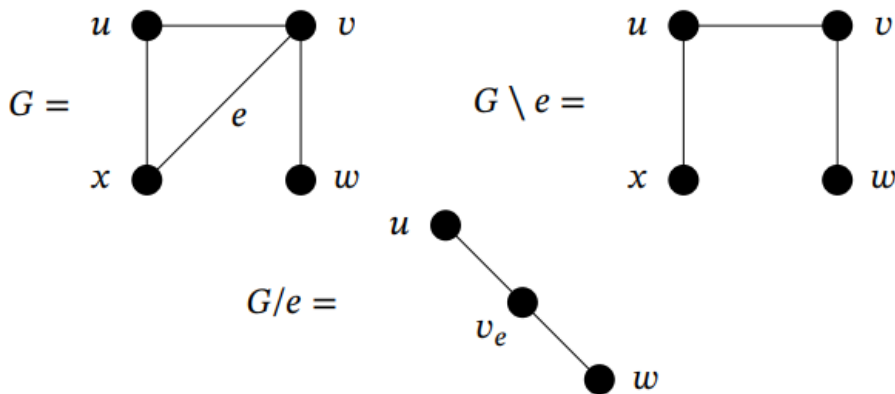


图 3.3: 删除与收缩

引理 3.8.2 (删除-收缩引理). 如果 G 是一个图, $e \in E$, 有

$$P(G; t) = P(G \setminus e; t) - P(G/e; t).$$

证明. 我们证明形式 $P(G \setminus e) = P(G) + P(G/e)$. 假定 $e = uv$. 因为 e 不出现在 $G \setminus e$ 中, 其正常着色数有两种类型: 一种是 $c(u) \neq c(v)$, 另一种是 $c(u) = c(v)$. 如果

$c(u) \neq c(v)$, 那么 $G \setminus e$ 的正常着色数等于 G 的正常着色数, 因而有 $P(G)$ 种着色. 当 $c(u) = c(v)$ 时, $G \setminus e$ 的正常着色与 G/e 之间也有一个双射, 即对 v_e 着色为和 u 与 v 相同的颜色, 其他点与 G 着色相同. 这种情况下有 $P(G/e)$ 种着色方式. 利用加法原则, 证毕. \square

定理 3.8.3. 对任意图 G , $P(G; t)$ 是关于 t 的多项式.

证明. 我们对 $|E|$ 作归纳. 如果 G 没有边, 那么 $P(G; t) = t^{|V|}$, 显然是关于 t 的多项式. 如果 $|E| \geq 1$, 则取边 $e \in E$. 由删除-收缩定理, $P(G; t) = P(G \setminus e; t) - P(G/e; t)$. 由归纳假设, $P(G \setminus e; t)$ 与 $P(G/e; t)$ 都是关于 t 的多项式. 因而它们的差也是关于 t 的多项式. \square

我们现在可以利用引理 3.8.2 来计算 C_4 的色多项式. 回顾一下符号表示, P_n 和 K_n 分别表示 n 个顶点的路和完全图. 现在选出任一边 $e \in E(C_4)$, 我们可以运用删除-收缩定理, 然后确定通过逐个顶点着色得到的图的多项式

$$P(C_4) = P(P_4) - P(K_3) = t(t-1)^3 - t(t-1)(t-2) = t(t-1)(t^2 - 3t + 3).$$

请注意, 二次因子具有复数根, 因此证实了 $P(G)$ 并不总是有整数的根.

可以使用归纳法和引理 3.8.2 来证明关于 $P(G; t)$ 的一系列结果. 由于这些证明都是相似的, 我们将它们留给读者. 我们会使用一种非标准的方式写下这个多项式的系数, 这在后面使用比较简单.

定理 3.8.4. 令 $G = (V, E)$, 且

$$P(G; t) = a_0 t^n - a_1 t^{n-1} + a_2 t^{n-2} - \cdots + (-1)^n a_n.$$

(1) $n = |V|$.

(2) $mdeg P(G; t) = G$ 的分支数.

(3) 对所有 i 有 $a_i \geq 0$, 且对 $0 \leq i \leq n - mdeg P(G; t)$ 有 $a_i > 0$.

(4) $a_0 = 1$ 且 $a_1 = |E|$.

现在我们知道 $P(G; t)$ 是一个多项式, 我们想要知道其系数是否存在组合解释, 我们方法的逆可以用来处理序列, 然后找到其生成函数. 对边集 E 设置一个总序, 如果边 e 在这一序下是小于边 f 的, 记作 $e < f$, 其他符号类似. 如果 C 是 G 中一个圈, 那么相应的破圈 B 就是在 $E(C)$ 移除总序中最小的边得到的边集. 回到图 3.2 中的例子, 令 $b = uv, c = ux, d = vw, e = vx$, 并设置序 $b < c < d < e$. 唯一的圈有边 b, c, e , 且相应的破圈 $B = \{c, e\}$, 就是一条路的边. 也就是说边集 $A \subseteq E$ 不包含破圈或者说是一个 NBC 集合, 如果 $A \supseteq B$ 对任意破圈 B . 令

$$NBC_k = NBC_k(G) = \{A \subseteq E \mid \#A = k \text{ 且 } A \text{ 是一个 NBC 集合}\}$$

且 $nbc_k = nbc_k(G) = \#NBC_k(G)$. 在我们的例图中, 比较上表的最后一列与 $P(G; t)$ 的系数得出我们下面的结果, 这就是 Whitney[99] 的结果. 令人惊讶的是结论确实不取决于赋予边的总序. 我们给出的证明是基于 Blass 和 Sagan [16] 的结果.

k	$\text{NBC}_k(G)$	$\text{nbc}_k(G)$
0	$\{\emptyset\}$	1
1	$\{\{b\}, \{c\}, \{d\}, \{e\}\}$	4
2	$\{\{b, c\}, \{b, d\}, \{b, e\}, \{c, d\}, \{d, e\}\}$	5
3	$\{\{b, c, d\}, \{b, d, e\}\}$	2
4	\emptyset	0

定理 3.8.5. 如果 $\#V = n$, 那么边集 E 的任意序,

$$P(G; t) = \sum_{k=0}^n (-1)^k \text{nbc}_k(G) t^{n-k}.$$

证明. 将每一个 $A \in \text{NBC}_k(G)$ 与生成子森林联系起来. 那么 A 是无圈的因为任何圈都包含一个破圈. 由定理 1.10.2 得 A 是一个有 $n - k$ 个分支树的森林. 因此 $\text{nbc}_k(G)t^{n-k}$ 是数对 (A, c) 的数目, 其中 $A \in \text{NBC}_k(G)$ 且 $c: V \rightarrow [t]$ 是 A 的每一个分支上的染色常数. 我们称这样的染色为 A -不正常. 通过令 $\text{sgn}(A, c) = (-1)^{\#A}$ 使这样的数对集合成为符号集. 因此, 如果我们找到一个在这些数对上的符号反转对合 ι , 其固定点集是正的符号, 并且和 G 的正常着色之间存在双射, 即可证明这一定理.

定义 ι 的固定点集为 (A, c) 使得 $A = \emptyset$ 且 c 是正常的. 这些数对显然具有所需的特性. 对于任何其他数对, c 不是正常的着色, 因此一定有边 $e = uv$ 满足 $c(u) = c(v)$. 令 e 是总序中最小的这样的边. 现在定义 $\iota(A, c) = (A \Delta \{e\}, c) := (A', c)$. 很明显 ι 是符号反转的. 并且它是一个对合, 因为 c 没有改变, 因此最小的单色边在一个对及其图像中是相同的. 我们只需要检验 ι 是定义良好的. 如果 $A' = A - \{e\}$, 那么很明显 A 仍然是一个 NBC 集合, 并且 c 是 A' -不正常的. 如果 $A' = A \cup \{e\}$, 由于 e 连接两个相同颜色的顶点, 所以 c 仍然是 A' -不正常的. 但是利用反证法, 假设 A' 不是 NBC 集. 那么 $A' \supseteq B$, 其中 B 是一个破圈, 且 $e \in B$, 因为 A 是 NBC 的. 因为 c 是 A' -不正常, B 中所有的边有顶点着色 $c(u)$. 但是 e 是有那样性质的最小的边, 因此通过移除这一条边来得到 B 是不存在的. 因此 A' 是 NBC, ι 是定义良好的, 证毕. \square

关于色多项式的一个令人惊奇的事情是, 它经常出现在先验没有业务的地方, 因为不涉及图形着色. 我们现在给出两个例子来说明这一点. 回想一下第 2.6 节, 图 G 中的一个方向图 O , 它是和 G 有相同顶点集的有向图, 通过将 G 的每条边 uv 替换为可能的弧之一 \overrightarrow{uv} 或者 \overrightarrow{vu} . 图 3.4 就是图 G 的两个方向图. 称 O 是无圈的, 如果它不包含任何有向圈, 令 $\mathcal{A}(G)$ 和 $a(G)$ 分别表示图 G 的无圈方向图的集合和计数. 图 3.4 中第一个方向图是无圈的, 而第二个是有圈的. 圈 u, v, x, u 的方向图的总数目为 2^3 , 那些形成一个圈的数目是 2 (顺时针和逆时针). 因为 vw 的任何方向都不可能产生圈, 我们有 $a(G) = 2(2^3 - 2) = 12$. 我们现在做一些非常奇怪的事情, 将 $t = -1$ 代入色多项式 (3.23) 并得到 $P(G; -1) = (-1)(-2)(-2)(-3) = 12$. 虽然完全不清楚用 -1 颜色为图形着色意味着什么, 我们刚刚看到的是下面的著名的斯坦利定理的例子 [85].

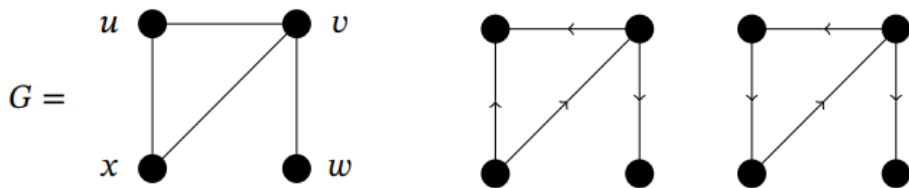


图 3.4: 一个图的方向

定理 3.8.6. 对任意图 G , 顶点数 $\#V = n$, 我们有

$$P(G; -1) = (-1)^n a(G).$$

证明. 我们对 $\#E$ 作归纳, 初始的情况容易验证. 只需验证 $(-1)^n P(G; -1)$ 和 $a(G)$ 满足相同的递推关系. 利用删除-收缩引理, 我们发现只需证明对固定边 $e = uv \in E$ 有 $a(G) = a(G \setminus e) + a(G/e)$. 考虑映射

$$\phi : A(G) \rightarrow A(G \setminus e),$$

它将 $O \mapsto O'$, 其中 O' 通过 O 移除 e 对应的弧得到. 很明显 O' 仍然是无圈的, 因此函数是定义良好的.

我们称 ϕ 是满射. 假设相反, 存在一些 $O' \in A(G \setminus e)$ 使得添加回 \vec{uv} 得到一个有向圈 C , 类似地添加回 \vec{vu} 得到一个有向圈 C' . 那么 $(C - \vec{uv}) \cup (C' - \vec{vu})$ 是闭的, 由第一章习题 14 (2) 得到, 有向途径一定包含一条有向圈. 但是第三条有向圈包含在 O' 中, 产生矛盾.

如果 $O' \in A(G \setminus e)$, 那么由映射的定义, $\#\phi^{-1}(O') \leq 2$. 从前一段我们知道 $\#\phi^{-1}(O') \geq 1$. 因此 $a(G) = x + 2y$, 其中 $x = \#\{O' \mid \phi^{-1}(O') = 1\}$, $y = \#\{O' \mid \phi^{-1}(O') = 2\}$. 因为 $a(G \setminus e) = x + y$, 只需证明 $a(G/e) = y$. 我们建立如下双射

$$\psi : \{O' \in A(G \setminus e) \mid \phi^{-1}(O') = 2\} \rightarrow A(G/e).$$

令 Y 为 ψ 的定义域. 如果存在一对边 $wu, vw \in E(G)$, 那么任意 $O' \in Y$ 包含 \vec{wu} 和 \vec{vw} , 或者包含 \vec{uw} 和 \vec{vw} . 这是因为在所有其他情形中, 添加回 e 的一个方向将形成 G 的带圈的方向图, 与 $\phi^{-1}(O') = 2$ 矛盾. 因此定义 $O'' = \psi(O')$ 为 G/e 的方向图, 它和 O' 上不包含新顶点 v_e 的那些弧是一致的, 在形式为 wv_e 的那些边和 wu 或 wv 方向相同. (如图所示, 这两个方向要么都朝向 e , 要么都远离 e). 证明 ψ 是定义良好的双射留作练习. \square

我们应该提到的是, Stanley 实际上证明了对所有负整数 $-t$ 的 $P(G; -t)$ 进行组合解释的更强结果. 具体细节见练习 28 (2). 所以, 正如我们在 (1.6) 中看到的二项式系数, 我们有其组合解释的另一个实例. 我们将在下一节更广泛地研究这种现象.

对于色多项式的多变性质的第二个示例, 我们需要假设我们的图 G 有顶点集 $[n]$ 以便顶点有一个总序. 令 F 为 G 的生成森林, 并将 F 的每个分支树 T 的最小顶点 r 作为根. 称 F 是增的, 如果对于所有的根 r , 以 r 为起点的路的顶点编号形成增序. 在图 3.5 中, 是一个图 G , 其顶点用 $[4]$ 编号, 与两个生成树. 我们称 F_1 是增序, 因为任意单元节点都是增的, 并且对非平凡树, 从根开始的唯一的最大路为 $1, 2, 4$, 显然是一个增序. 另一方面, F_2 非增, 因为有路 $1, 4, 2$.

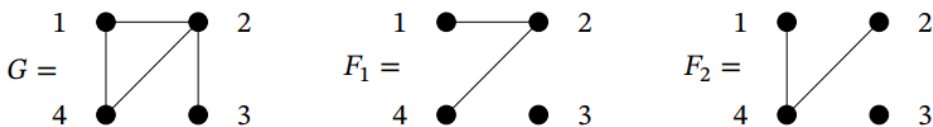


图 3.5: 一个图与两种生成森林

对一个图 $G = (V, E)$, 我们定义

$$\text{ISF}_m(G) = \{F \mid F \text{ 是 } G \text{ 的有 } m \text{ 条边的生成森林}\},$$

且 $\text{isf}_m(G) = \# \text{ISF}_m(G)$. 如果 $\#V = n$, 那么考虑相应的生成多项式

$$\text{isf}(G) = \text{isf}(G; t) = \sum_{m=0}^n (-1)^m \text{isf}_m(G) t^{n-m}.$$

我们用这一公式计算我们的例图. 有零条边或一条边的树都是增的, 因此 $\text{isf}_0(G) = 1$ 且 $\text{isf}_1(G) = \#E = 4$. G 的任意的这样的边对形成一个增森林, 除了图 3.5 中 F_2 那样的边对. 因此 $\text{isf}_2(G) = \binom{4}{2} - 1 = 5$. 类似地我们可以验证 $\text{isf}_3(G) = 2$. 并且 $\text{isf}_4(G) = 0$, 因为 G 本身不是森林. 因此

$$\text{isf}(G; t) = t^4 - 4t^3 + 5t^2 - 2t = t(t-1)^2(t-2) = P(G; t).$$

我们并不总是有 $\text{isf}(G) = P(G)$, 因为前者依赖于顶点的编号 (即使我们的符号掩盖了事实), 然而后者并不依赖. 因此, 我们将暂时不考虑它们何时相等, 现在只关注它们在 \mathbb{Z} 上的因式分解, 正如我们将看到的, 它不是巧合. 事实上, 根将是由如下定义的边集的基数

$$E_k = \{ik \in E \mid i < k\}, \quad (3.12)$$

其中 $1 \leq k \leq n$. 在我们的例子中 $E_1 = \emptyset$ (因为没有顶点比 1 小), $E_2 = \{12\}$, $E_3 = \{23\}$, $E_4 = \{14, 24\}$.

引理 3.8.7. 如果 G 有顶点 $V = [n]$, 那么一个生成子图 F 是一个增森林当且仅当对于 $k \in [n]$ 它是通过从每个 E_k 中选择最多一条边获得的.

证明. 对于前一个方向, 利用反证法, 假设 F 包含 ik 和 jk , 其中 $i, j < k$. 因此如果 r 是包含 i, j, k 的树的根, 那么, 由增序条件, i 必须是唯一的 $r-k$ 路径上 k 之前的顶点. 但是 j 也必须如此, 这就产生矛盾.

对于另一个方向, 我们必须首先验证 F 是无圈的. 但是如果 F 包含一个圈 C , 那么令 k 为其最大顶点. 因此有 $ik, jk \in E(C)$, 并且根据最大要求, 有 $i, j < k$. 这与假设相矛盾. 类似地, 可以证明如果 F 不是增序, 则可以从相同的 E_k 产生两个边, 证毕. \square

现在证明 Hallam 和 Sagan [41] 的以下结果就很简单了. 这里给出的证明是由 Hallam、Martin 和 Sagan [40] 得到的.

定理 3.8.8. 如果 G 有顶点 $V = [n]$, 那么

$$\text{isf}(G; t) = \prod_{k=1}^n (t - |E_k|)$$

证明. 乘积中 t^{n-m} 的系数是形式 $|E_{i_1}| |E_{i_2}| \cdots |E_{i_m}|$ 的所有项的和, 其中 i_j 是不同的指标. 但是这个乘积是从每个集合 $E_{i_1}, E_{i_2}, \dots, E_{i_m}$ 中选出一条边的方法数. 所以, 根据前面引理, 总和是具有 m 条边的增森林的数量, 证毕. \square

回到色多项式与增生成森林多项式何时相等的问题, 我们需要下面的定义. 如果 V 的总序 v_1, v_2, \dots, v_n 使得对所有的 k , 在总序中在为 v_k 之前并与 v_k 相邻的顶点集合形成一个团 (G 的完整子图), 则图 G 具有完美消除序 (peo). 乍一看, 这个定义似乎很奇怪, 但是它在各种图理论环境中很有用. 回到我们的例子, 我们看到序 $1, 2, 3, 4$ 是一个完美消除序, 因为 1 与之前的顶点不相邻, 顶点 2 和 3 都与单个先前的顶点相邻, 即 K_1 , 并且 4 与 1 和 2 相邻, 这形成一条边, 也即 K_2 . 我们可以证明来自 [40] 的另一个结果.

引理 3.8.9. 令 G 顶点 $V = [n]$. 将 G 的边写作 ij 其中 $i < j$, 并将它们按字典序排列. 对所有的 $m \geq 0$, 我们有

$$\text{ISF}_m(G) \subseteq \text{NBC}_m(G).$$

此外, 对所有的 m 上面两个集合相等当且仅当 $[n]$ 上的自然顺序是完美消除序.

证明. 为了证明包含关系, 我们假设 F 是一个增的生成森林, 它包含破圈 B , 并导出矛盾. 按边的字典序, B 必须是形式为 v_1, v_2, \dots, v_l 的路, 其 $v_1 = \min \{v_1, \dots, v_l\}$, 且 $v_2 > v_l$. 所以一定存在一个最小指标 $i \geq 2$ 使得 $v_i > v_{i+1}$. 随之有 $v_{i-1}, v_{i+1} < v_i$, 所以 B 的两条相应的边与引理 3.8.7 矛盾.

对于第二个陈述的前一个方向, 我们必须证明如果 $i, j < k$ 并且 $ik, jk \in E(G)$, 那么 $ij \in E(G)$. 再利用引理 3.8.7, $\{ik, jk\}$ 不是增生成森林的边集. 因此, 通过假定的相等性, 这个集合必须包含一个破圈. 由于只有两条边, 这一边集一定是破圈, 并且 $ij \in E(G)$ 一定是用来形成圈的边. 另一方向的证明留作练习. \square

从这个结果, 我们立即得出以下结论.

定理 3.8.10. 令 G 顶点 $V = [n]$. 那么 $\text{isf}(G; t) = P(G; t)$ 当且仅当 $[n]$ 上的自然序是一个完美消除序. \square

3.9 组合互反性

将一个负参数代入一个计数函数得到一个符号乘另一个计数函数, 那么这就称为组合互反. 这一概念是由 Stanley[86] 引入和研究的. 我们已经在等式 (1.6) 与定理 3.8.6 中看到了两个例子 (更一般地, 见本章练习 28). 在这里, 我们将递归与有理生成函数联系起来. 请参阅 Beck 和 Sanyal [5] 的关于该问题的论述.

在陈述一个一般定理之前, 我们回到第 3.6 节开始的例子. 也即如下定义的序列 $a_0 = 2$, 且当 $n \geq 1$ 时 $a_n = 3a_{n-1}$. 我们可以将这一递推关系的定义域扩展到所有的整数 n , 在这种情况下可以得到 $2 = a_0 = 3a_{-1}$, 因此 $a_{-1} = 2/3$. 那么 $2/3 = a_{-1} = 3a_{-2}$, 于是得到 $a_{-2} = 2/3^2$, 以此类推. 作简要的归纳得到当 $n \leq 0$ 有 $a_n = 2 \cdot 3^n$, 与 $n \geq 0$ 时一样. 我们也可对这一序列的负指标部分计算生成函数, 从 a_{-1} 开始计算, 得到几何级数

$$\sum_{n \geq 1} a_{-n} x^n = \sum_{n \geq 1} \frac{2x^n}{3^n} = \frac{2x/3}{1 - x/3} = \frac{2x}{3 - x}.$$

将其与 (3.12) 中的式子 $f(x) = \sum_{n \geq 0} a_n x^n$ 比较得到

$$-f(1/x) = \frac{-2}{1 - 3/x} = \frac{2x}{3 - x} = \sum_{n \geq 1} a_{-n} x^n.$$

读者应该对写 $f(1/x)$ 有些疑虑, 因为我们已经指出 x 在 $\mathbb{C}[[x]]$ 中没有逆. 事实上, 如果我们使用定义 $f(x) = \sum_{n \geq 0} a_n x^n$, 那么 $f(1/x) = \sum_{n \geq 0} a_n / x^n$, 这不是一个形式幂级数. 但是如果 $f(x)$ 可以被表达成有理函数的形式, 即 $f(x) = p(x)/q(x)$, 其中 $\deg p(x) \leq \deg q(x) := d$, 那么作如下替换就有意义. 因为 $q(x)$ 度为 d , 我们知道 $x^d q(1/x)$ 也是一个多项式, 并且是可逆的, 因为它的常数项不为零 (定理 3.3.1). 此外, $x^d p(1/x)$ 也是一个多项式, 因为 $d \geq \deg p(x)$. 所以我们可以定义

$$f(1/x) = \frac{x^d p(1/x)}{x^d q(1/x)}, \quad (3.25)$$

它仍在形式幂级数环中. 在这样定义下, 下面的结果就有意义.

定理 3.9.1. 假设 a_n 是一个序列, $n \in \mathbb{Z}$ 满足 (3.18) 中常系数的线性递推关系. 令 $f(x) = \sum_{n \geq 0} a_n x^n$, 我们有

$$\sum_{n \geq 1} a_{-n} x^n = -f(1/x)$$

证明. 由 (3.25), 要证明该定理只需证明

$$x^d q(1/x) \sum_{n \geq 1} a_{-n} x^n = -x^d p(1/x).$$

注意由 (3.20) 式, 我们有

$$x^d q(1/x) = x^d + c_1 x^{d-1} + c_2 x^{d-2} + \cdots + c_d.$$

因此, 如果 $m \geq 1$, 那么利用 (3.18) 及 $x^d p(1/x)$ 的度至多为 d , 得

$$\begin{aligned} [x^{m+d}] x^d q(1/x) \sum_{n \geq 1} a_{-n} x^n &= a_{-m} + c_1 a_{-m-1} + c_2 a_{-m-2} + \cdots + c_d a_{-m-d} \\ &= 0 \\ &= [x^{m+d}] (-x^d p(1/x)). \end{aligned}$$

类似地, 我们可以证明该不等式对 $-d \leq m \leq 0$ 仍成立. 证毕. \square

为了说明这一定理, 我们考虑负的二项式展开. 如果固定 $n \geq 1$, 应用定理 3.4.2 及等式 (3.16), 得

$$f(x) := \frac{1}{(1-x)^n} = \sum_{k \geq 0} \left(\binom{n}{k} \right) x^k = \sum_{k \geq 0} \binom{n+k-1}{n-1} x^k.$$

注意因为 n 是固定的, 我们将 $\binom{n+k-1}{n-1}$ 看作关于 k 的函数. 在二项式系数中将 k 替换为 $-k$, 我们考虑相应的生成函数

$$g(x) = \sum_{k \geq 1} \binom{n-k-1}{n-1} x^k.$$

注意到当 $1 \leq k < n$ 时 $\binom{n-k-1}{n-1} = 0$, 因为 $0 \leq n-k-1 < n-1$. 所以 x^n 可以从 $g(x)$ 分解出来, 使用 (1.6) 和上面的负二项式展开, 得到

$$\begin{aligned} g(x) &= x^n \sum_{k \geq n} \binom{n-k-1}{n-1} x^{k-n} \\ &= x^n \sum_{j \geq 0} \binom{-j-1}{n-1} x^j \\ &= (-1)^{n-1} x^n \sum_{j \geq 0} \left(\binom{j+1}{n-1} \right) x^j \\ &= (-1)^{n-1} x^n \sum_{j \geq 0} \binom{n+j-1}{n-1} x^j \\ &= \frac{(-1)^{n-1} x^n}{(1-x)^n}. \end{aligned}$$

另一方面, 我们可以应用定理 3.9.1 将 $g(x)$ 写作

$$g(x) = \frac{-1}{(1-1/x)^n} = \frac{-x^n}{(x-1)^n} = \frac{(-1)^{n-1} x^n}{(1-x)^n},$$

得到同样的结果, 但是计算更少.

3.10 习题

1. 用三种方式证明对 $n \in \mathbb{N}$

$$\sum_{k=0}^n 2^k \binom{n}{k} = 3^n.$$

(1) 应用二项式定理。(2) 应用组合解释。(3) 对任意 $c \in \mathbb{N}$, 通过把 2^k 替换为 c^k 将等式一般化, 用二项式定理及组合解释给出证明。

2. 对 $m, n, k \in \mathbb{N}$, 证明

$$\binom{m+n}{k} = \sum_{l \geq 0} \binom{m}{l} \binom{n}{k-l}$$

in three ways: by induction, using the Binomial Theorem, and using a combinatorial argument. (3) Let x_1, \dots, x_m be variables. Prove the multinomial coefficient identity

$$\sum_{n_1 + \dots + n_m = n} \binom{n}{n_1, \dots, n_m} x_1^{n_1} \cdots x_m^{n_m} = (x_1 + \dots + x_m)^n$$

in three ways: (a) by inducting on n , (b) by using the Binomial Theorem and inducting on m , (c) by a combinatorial argument. (4) (a) Prove Theorem 3.1.2. (b) Use this generating function to rederive Corollary 1.5.3. (5) (a) Recall that an inversion of $\pi \in P([n])$ is a pair (i, j) with $i < j$ and $\pi_i > \pi_j$ and we call π_i the maximum of the inversion. The inversion table of π is $I(\pi) = (a_1, a_2, \dots, a_n)$ where a_k is the number of inversions with maximum k . Show that $0 \leq a_k < k$ for all k and that

$$\text{inv } \pi = \sum_{k=1}^n a_k$$

(b) Let

$$\mathcal{I}_n = \{(a_1, a_2, \dots, a_n) \mid 0 \leq a_k < k \text{ for all } k\}.$$

Show that the map $\pi \mapsto I(\pi)$ is a bijection $P([n]) \rightarrow \mathcal{I}_n$. (c) Use part (b) and weight-generating functions to rederive Theorem 3.2.1. (6) Let $\text{st}: \mathfrak{S}_n \rightarrow \mathbb{N}$ be a statistic on permutations. Recall that for a permutation π we let $\text{Av}_n(\pi)$ denote the set of permutations in \mathfrak{S}_n avoiding π . Say that permutations π, σ are st-Wilf equivalent, written $\pi \equiv^{\text{st}} \sigma$, if for all $n \geq 0$ we have equality of the

generating functions

$$\sum_{\tau \in \text{Av}_n(\pi)} q^{\text{st}\tau} = \sum_{\tau \in \text{Av}_n(\sigma)} q^{\text{st}\tau}$$

(a) Show that if π and σ are st-Wilf equivalent, then they are Wilf equivalent. (b) Show that

$$132 \stackrel{\text{inv}}{\equiv} 13$$

and

$$231 \stackrel{\text{inv}}{\equiv} 312$$

and that there are no other inv-Wilf equivalences between two permutations in \mathfrak{S}_3 . (c) Show that

$$\begin{aligned} 132 &\stackrel{\text{maj}}{\equiv} 231 \\ 213 &\stackrel{\text{maj}}{\equiv} 312 \end{aligned}$$

and that there are no other maj-Wilf equivalences between two permutations in \mathfrak{S}_3 . (7) (a) Prove that

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n \\ n-k \end{bmatrix}$$

in three ways: using the q -factorial definition, using the interpretation in terms of integer partitions, and using the interpretation in terms of subspaces. (b) Prove the second recursion in Theorem 3.2.3 in two ways: by mimicking the proof of the first recursion and by using the first recursion in combination with part (a). (c) If $S \subseteq [n]$, then let ΣS be the sum of the elements of S . Give two proofs of the following q -analogue of the fact that $\# \begin{pmatrix} [n] \\ k \end{pmatrix} = \begin{pmatrix} n \\ k \end{pmatrix}$:

$$\sum_{S \in \begin{pmatrix} (n) \\ k \end{pmatrix}} q^{\Sigma S} = q^{\begin{pmatrix} k+1 \\ 2 \end{pmatrix}} \begin{bmatrix} n \\ k \end{bmatrix}_q,$$

one proof by induction and the other using Theorem 3.2.5. (d) Give two other proofs of Theorem 3.2.6: one by inducting on n and one by using Theorem 3.2.5. (8) (a) Reprove Theorem 3.2.4 in two ways: using integer partitions and using subspaces. (b) The Negative q -Binomial Theorem states that

$$\frac{1}{(1-t)(1-qt)(1-q^2t)\cdots(1-q^{n-1}t)} = \sum_{k \geq 0} \begin{bmatrix} n+k-1 \\ k \end{bmatrix} t^k$$

Give three proofs of this result: inductive, using integer partitions, and using subspaces.

(9) (a) Given $n_1 + n_2 + \cdots + n_m = n$, define the corresponding q -multinomial coefficient to be

$$[n_1, n_2, \dots, n_m]_q = \frac{[n]_q}{[n_1]_q! [n_2]_q! \cdots [n_m]_q!}$$

if all $n_i \geq 0$ or zero otherwise. Prove that

$$\begin{aligned} &[n_1, n_2, \dots, n_m]_q \\ &= \sum_{i=1}^m q^{n_1+n_2+\cdots+n_{i-1}} \begin{bmatrix} n-1 \\ n_1, \dots, n_{i-1}, n_i-1, n_{i+1}, \dots, n_m \end{bmatrix}_q \end{aligned}$$

(b) Define inversions, descents, and the major index for permutations (linear orderings) of the multiset $M = \{1^{n_1}, 2^{n_2}, \dots, m^{n_m}\}$ exactly the way as was done for permutations without repetition. Let $P(M)$ be the set of permutations of M . Prove that

$$\sum_{\pi \in P(M)} q^{\text{inv}\pi} = \sum_{\pi \in P(M)} q^{\text{maj}\pi} = \left[\begin{matrix} n \\ n_1, n_2, \dots, n_m \end{matrix} \right]_q$$

(c) Let V be a vector space over \mathbb{F}_q of dimension n . Assume $S = \{s_1 < \dots < s_m\} \subseteq \{0, 1, \dots, n\}$. Then a flag of type S is a chain of subspaces

$$F : W_1 < W_2 < \dots < W_m \leq V$$

such that $\dim W_i = s_i$ for all i . The reason for this terminology is that when $n = 2$ and $S = \{0, 1, 2\}$, then F consists of a point (the origin) contained in a line contained in a plane which could be viewed as a drawing of a physical flag with the point being the hole in the ground, the line being the flag pole, and the plane being the cloth flag itself. Give two proofs that

$$\#\{F \text{ of type } S\} = [s_1, s_2 - s_1, s_3 - s_2, \dots, s_m - s_{m-1}, n - s_m]_q,$$

one by mimicking the proof of Theorem 3.2.6 and one by induction on m . (10) (a) Prove that in $C[[x]]$ we have $e^{kx} = (e^x)^k$ for any $k \in \mathbb{N}$. (b) Define formal power series for the trigonometric functions using their usual Taylor expansions. Prove that in $C[[x]]$ we have $\sin^2 x + \cos^2 x = 1$ and $\sin 2x = 2 \sin x \cos x$. (c) If one is given a sequence a_0, a_1, a_2, \dots and defines

$$f_k(x) = a_k x^k,$$

then show that $\sum_{k \geq 0} f_k(x) = f(x)$ where $f(x) = \sum_{k \geq 0} a_k x^k$. (d) Prove the backwards direction of Theorem 3.3.2. (e) Use Theorems 3.3.1 and 3.3.3 to reprove that $1/x$ and e^{1+x} are not well-defined in $\mathbb{C}[[x]]$. (f) Prove Theorem 3.3.4 (11) Prove that if S, T are summable sets, then so is $S \times T$.

(12) Say that $f(x) \in \mathbb{C}[[x]]$ has a square root if there is $g(x) \in \mathbb{C}[[x]]$ such that $f(x) = g(x)^2$ (a) Prove that $f(x)$ has a square root if and only if $\text{mdeg } f(x)$ is even. (b) Show that as formal power series

$$(1+x)^{1/2} = \sum_{k \geq 0} \binom{1/2}{k} x^k$$

(c) Show that as formal power series

$$(e^x)^{1/2} = \sum_{k \geq 0} \frac{x^k}{2^k k!}$$

(d) Generalize the previous parts of this exercise to m th roots for $m \in \mathbb{P}$. (13) Give a second proof of Theorem 3.4.2 by using induction. (14) Prove Theorem 3.4.3. (15) Prove Theorem 3.5.1. (16) (a) Finish the proof of Theorem 3.5.3 (a). (b) Give a second proof of Theorem 3.5.3 (b) by using Theorem 3.2.5. (c) Show that the generating function for the number of partitions of n with largest part k equals the generating function for the number of partitions of n with exactly k parts and that both are equal to the product

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

(d) The Durfee square of a Young diagram λ is the largest square partition (d^d) such that $(d^d) \subseteq \lambda$. Use this concept to prove that

$$\sum_{n \geq 0} p(n)x^n = \sum_{d \geq 0} \frac{x^{d^2}}{(1-x)^2(1-x^2)^2\cdots(1-x^d)^2}$$

(17) Let a_n be the number of integer partitions of n such that any part i is repeated fewer than i times and let b_n be the number of integer partitions of n such that no part is a square. Use generating functions to show that $a_n = b_n$ for all n . (18) Given $m \geq 2$, use generating functions to show that the number of partitions of n where each part is repeated fewer than m times equals the number of partitions of n into parts not divisible by m . Note that bijective proofs of this result were given in Exercise 15 of Chapter 2. (19) (a) Show that F_n is the closest integer to

$$\frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n$$

for $n \geq 1$. (b) Prove that

$$F_n^2 = \begin{cases} F_{n-1}F_{n+1} + 1 & \text{if } n \text{ is odd,} \\ F_{n-1}F_{n+1} - 1 & \text{if } n \text{ is even} \end{cases}$$

in two ways: using equation (3.14) and by a combinatorial argument.

(20) (a) Use the algorithm in Section 2.6 to rederive Theorem 3.1.1. (b) Complete the proof of Theorem 3.6.1. (c) Give a second proof of Theorem 3.6.1 using Theorem 3.1.2 (d) Prove Theorem 3.6.2. (e) Let s be the infinite matrix with rows and columns indexed by \mathbb{N} and with $s(n, k)$ being the entry in row n and column k . Similarly define S with entries $S(n, k)$. Show that $Ss = sS = I$ where I is the $\mathbb{N} \times \mathbb{N}$ identity matrix. Hint: Use Theorems 3.6.1 and 3.6.2. (21) Given $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, define the corresponding q -Stirling number of the second kind by $S[0, k] = \delta_{0,k}$ and, for $n \geq 1$,

$$S[n, k] = S[n-1, k-1] + [k]_q S[n-1, k].$$

(a) Show that

$$\sum_{n \geq 0} S[n, k] x^n = \frac{x^k}{(1 - [1]_q x)(1 - [2]_q x) \cdots (1 - [k]_q x)}.$$

(b) All set partitions $\rho = B_1/B_2/\dots/B_k$ in the rest of this problem will be written in standard form, which means that

$$1 = \min B_1 < \min B_2 < \cdots < \min B_k.$$

An inversion of ρ is a pair (b, B_j) where $b \in B_i$ for some $i < j$ and $b > \min B_j$. We let $\text{inv } \rho$ be the number of inversions of ρ . For example, $\rho = B_1/B_2/B_3 = 136/25/4$ has inversions $(3, B_2), (6, B_2), (6, B_3)$, and $(5, B_3)$ so that $\text{inv } \rho = 4$. Show that

$$S[n, k] = \sum_{\rho \in S([n], k)} q^{\text{inv } \rho}.$$

(c) A descent of a set partition ρ is a pair (b, B_{i+1}) where $b \in B_i$ and $b > \min B_j$. We let $\text{des } \rho$ be the number of descents of ρ . In the previous example, ρ has descents $(3, B_2), (6, B_2)$, and $(5, B_3)$ so that $\text{des } \rho = 3$. The descent multiset of ρ is denoted $\text{Des } \rho$ and is the multiset

$$\{ \{1^{d_1}, 2^{d_2}, \dots, k^{d_k} \mid \text{for all } i, d_i = \text{number of descents } (b, B_{i+1})\} \}.$$

The major index of ρ is

$$\text{maj } \rho = \sum_{i \in \text{Des } \rho} i = d_1 + 2d_2 + \cdots + kd_k.$$

In our running example $\text{Des } \rho = \{ \{1^2, 2^1\} \}$ so that $\text{maj } \rho = 1 + 1 + 2 = 4$. Show that

$$S[n, k] = \sum_{\rho \in S([n], k)} q^{\text{maj } \rho}.$$

(22) Given $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, define the corresponding signless q -Stirling number of the first kind by $c[0, k] = \delta_{0, k}$ and, for $n \geq 1$,

$$c[n, k] = c[n-1, k-1] + [n-1]_q c[n-1, k]$$

(a) Show that

$$\sum_{k \geq 0} c[n, k] x^k = x(x + [1]_q)(x + [2]_q) \cdots (x + [n-1]_q).$$

(b) The standard form of $\pi \in \mathfrak{S}_n$ is $\pi = \kappa_1 \kappa_2 \cdots \kappa_k$ where the κ_i are the cycles of π ,

$$\min \kappa_1 < \min \kappa_2 < \cdots < \min \kappa_k,$$

and each κ_i is written beginning with $\min \kappa_i$. Define the cycle major index of π to be $\text{maj}_c \pi = \text{maj } \pi'$ where π' is the permutation in one-line form obtained by removing the cycle parentheses in the standard form of π . If, for example, $\pi = (1, 7, 2)(3, 6, 8)(4, 5)$, then $\pi' = 17236845$ so that $\text{maj}_c \pi = 2 + 6 = 8$. Show that

$$c[n, k] = \sum_{\pi \in c([n], k)} q^{\text{maj}_c \pi}.$$

(23) Reprove the formula

$$C(n) = \frac{1}{n+1} \binom{2n}{n}$$

by using Theorem 2.6.3. (24) (a) Show that if $k, l \in \mathbb{N}$ are constants, then $\binom{n+l}{k}$ is a polynomial in n of degree k . (b) Show that the polynomials

$$\binom{n}{0}, \binom{n+1}{1}, \binom{n+2}{2}, \dots$$

form a basis for the algebra of polynomials in n . (c) Use part (a) to complete the proof of Theorem 3.7.1. (25) Redo the solution for the first recursion in Section 3.6 as well as the one for F_n using the method of undetermined coefficients. (26) Prove for the n -cycle that

$$P(C_n; t) = (t-1)^n + (-1)^n(t-1)$$

in two ways: using deletion-contraction and using NBC sets. (27) Prove Theorem 3.8.4 using induction and give a second proof of parts (b)-(d) using NBC sets. (28) (a) Complete the proof of Theorem (b) Let $G = (V, E)$ be a graph and $t \in \mathbb{P}$. Call an acyclic orientation O and a (not necessarily proper) coloring $c : V \rightarrow [t]$ compatible if, for all arcs \vec{uv} of O , we have $c(u) \leq c(v)$. Show that if $\#V = n$, then

$$P(G; -t) = (-1)^n (\text{number of compatible pairs } (O, c)).$$

(c) Show that Theorem 3.8.6 is a special case of part (b). (29) Finish the proofs of Lemma 3.8.7 and Lemma 3.8.9. (30) (a) Call a permutation σ which avoids $\Pi = \{231, 312, 321\}$ tight. Show that σ is tight if and only if σ is an involution having only 2-cycles of the form $(i, i+1)$ for some i . (b) Let G be a graph with $V = [n]$. Call a spanning forest F of G tight if the sequence of labels on any path starting at a root of F avoids Π as in part (a). Let $\text{TSF}_m(G) = \{F \mid F \text{ is a tight spanning forest of } G \text{ with } m \text{ edges}\}$.

Show that if G has no 3-cycles, then for all $m \geq 0$

$$\text{TSF}_m(G) \subseteq \text{NBC}_m(G).$$

(c) A candidate path in a graph G with no 3-cycles is a path of the form

$$a, c, b, v_1, v_2, \dots, v_m = d$$

such that $a < b < c, m \geq 1$, and v_m is the only v_i smaller than c . A total order on $V(G)$ is called a quasiperfect ordering (qpo) if every candidate path satisfies the following condition: either $ad \in E(G)$, or $d < b$ and $cd \in E(G)$. Consider the generating function

$$\text{tsf}(G; t) = \sum_{m \geq 0} (-1)^m \text{tsf}_m(G) t^{n-m}.$$

Show that $\text{tsf}(G; t) = P(G; t)$ if and only if the natural order on $[n]$ is a qpo. (31) Fill in the details of the case $-d \leq m \leq 0$ in the proof of Theorem 3.9.1. (32) (a) Extend the Fibonacci numbers F_n to all $n \in \mathbb{Z}$ by insisting that their recursion continue to hold for $n < 0$. Show that if $n \geq 0$, then

$$F_{-n} = (-1)^{n-1} F_n.$$

(b) Find $\sum_{n \geq 1} F_{-n} x^n$ in two ways: by using part (a) and by using Theorem 3.9.1.