Производящие функции и их использование в элементарной комбинаторике

1 Основные определения

1. Понятие производящей функции — это одно из основных понятий современной перечислительной комбинаторики. Задача этого пункта — дать вначале неформальное, а затем и вполне строгое определение этой функции.

1.1. Пусть вначале X — это некоторое конечное множество, элементы которого мы собираемся перечислять. Первое, что обычно хочется сделать — это указать мощность этого множества, т.е. количество всех элементов множества X. Однако, как правило, нам этой информации оказывается мало — на следующем шаге мы обычно хотим как-то классифицировать элементы этого множества.

1.2. Например, рассмотрим множество всех товаров на складе. Конечно, полезно знать общее их число. Однако гораздо полезнее разбить эти товары на категории или группы (например, бутылки вина, коробки мыла, пакеты с чаем), и подсчитать количество товаров в каждой группе по отдельности. При этом мы, конечно же, получим и общее число товаров, просто сложив количество товаров каждой категории.

В качестве еще одного примера можно привести гипотетическую задачу подсчета всех собак, живущих на Земле. Если пытаться практически решать эту задачу, то легче, видимо, начинать с подсчета собак данной породы, живущих в данном конкретном городе.

1.3. Рассмотрим теперь счетное множество X, например, множество всех графов. Здесь вообще ставить вопрос о количестве всех объектов данного множества бессмысленно. Однако и в этом случае задачу перечисления можно сделать вполне содержательной, и сделать это можно с помощью того же подхода. Именно, можно разбить множество X на конечные блоки по какому-то разумному признаку, и перечислять количество элементов в каждом блоке.

1.4. Например, бессмысленно перечислять все звезды во Вселенной. Однако перечислить все звезды в данной галактике — вполне разумная задача. Далее, столь же бессмысленно перечислять все графы. Однако перечисление всех различных графов на n вершинах при заданном n есть также вполне содержательная и важная с практической точки зрения задача.

1.5. Как же чисто технически осуществить разбиение множества X на блоки? Наиболее популярный способ сделать это следующий: нужно приписать всем элементам x рассматриваемого множества некоторый вес так, чтобы элементы одного и того же блока имели одинаковый вес. Формально для этого следует ввести отображение

$$w : X \longrightarrow K$$

множества X в некоторое коммутативное кольцо K, при котором всем элементам x_i данного блока сопоставляется одно и то же значение $w(x_i) = k$. При этом все элементы X разбиваются на блоки мощности $c_k := |w^{-1}(k)|$:

$$X = \bigcup_{k\in K} \{x \mid w(x) = k\}.$$

Заметим, что при таком подходе всему множеству X в кольце K отвечает вполне конкретный элемент
$$\sum_{x\in X} w(x) = \sum_{k\in K} c_k \cdot k, \qquad\qquad c_k = |w^{-1}(k)|,$$
называемый энумератором множества X.

1.6. `Пример 1.` Пусть на складе имеются три пакета с чаем, две бутылки вина и четыре коробки с мылом:
$$X = \{\text{чай}_1, \text{чай}_2, \text{чай}_3, \text{вино}_1, \text{вино}_2, \text{мыло}_1, \text{мыло}_2, \text{мыло}_3, \text{мыло}_4\}.$$
Введем теперь отображение $w : X \to \{\text{ч}, \text{в}, \text{м}\}$ следующим образом:
$$w(\text{чай}_i) = \text{ч}, \qquad w(\text{вино}_i) = \text{в}, \qquad w(\text{мыло}_i) = \text{м}.$$
Тогда все множество X разобьется на три блока, а энумератор этого множества примет вид
$$w(X) = 3\text{ч} + 2\text{в} + 4\text{м}.$$

1.7. Как правило, в перечислительной комбинаторике используется отображение $w : X \to K$ весьма специального вида. Именно, в качестве K выбирается множество $\mathbb{C}^{\mathbb{Z}_+}$ всех счетных последовательностей $(a_0, a_1, a_2, \ldots)$ комплексных чисел:
$$\mathbb{C}^{\mathbb{Z}_+} := \{(a_0, a_1, a_2, \ldots) \,|\, a_i \in \mathbb{C} \quad \forall\, i \in \mathbb{Z}_+\}.$$
На этом множестве вводятся две базовые операции — сложения и умножения. Операция сложения числовых последовательностей определяется, как правило, однозначно:
$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots). \tag{1}$$
Операцию же умножения
$$(a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots) = (c_0, c_1, c_2, \ldots)$$
можно вводить по-разному. Чаще всего на практике используются следующие два способа умножения этих последовательностей: свертка
$$c_n = \sum_{i=0}^{n} a_i \cdot b_{n-i} \tag{2}$$
и биномиальная свертка
$$c_n = \sum_{i=0}^{n} \binom{n}{i} \cdot a_i \cdot b_{n-i} \tag{3}$$
числовых последовательностей. В обоих случаях множество $\mathbb{C}^{\mathbb{Z}_+}$ с введенными на нем операциями представляет собой коммутативное кольцо (точнее, коммутативную область целостности).

1.8. На практике с этими кольцами удобно работать как с кольцами $\mathbb{C}[[z]]$ и $\mathbb{C}_e[[z]]$ формальных степенных рядов. В таких кольцах для элемента $(0, 1, 0, 0, \ldots)$ вводится специальное обозначение
$$(0, 1, 0, 0, \ldots) =: z.$$
Рассмотрим вначале случай кольца $\mathbb{C}[[z]]$ с правилом умножения (2). Заметим, что
$$z^2 := z \cdot z = (0, 1, 0, 0, \ldots) \cdot (0, 1, 0, 0, \ldots) = (0 \cdot 0, 0 \cdot 1 + 1 \cdot 0, 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0, \ldots) = (0, 0, 1, 0, \ldots),$$

$$z^3 := z \cdot z \cdot z = (0,1,0,0,\dots)\cdot(0,1,0,0,\dots)\cdot(0,1,0,0,\dots) = (0,1,0,0,\dots)\cdot(0,0,1,0,\dots) =$$
$$= (0\cdot 0, 0\cdot 0 + 1\cdot 0, 0\cdot 1 + 1\cdot 0 + 0\cdot 0, 0\cdot 0 + 1\cdot 1 + 0\cdot 0 + 0\cdot 0, \dots) = (0,0,0,1,0,\dots),$$

В общем случае, как несложно проверить по индукции,

$$z^n := \underbrace{z\cdot z\cdot \ldots \cdot z}_{n} = (0,\dots,0,1,0,\dots),$$

где единица в последнем выражении стоит на n-м месте. Как следствие, любую последовательность чисел $(a_0, a_1, a_2, \dots)$ можно записать в виде формального степенного ряда

$$(a_0, a_1, a_2, \dots) = \sum_{n=0}^{+\infty} a_n z^n =: a(z) \in \mathbb{C}[[z]].$$

Аналогично, в случае кольца $\mathbb{C}_e[[z]]$ с правилом умножения (3) несложно убедиться в том, что

$$z^n := \underbrace{z\cdot z\cdot \ldots \cdot z}_{n} = (0,\dots,0,n!,0,\dots),$$

причем единственное ненулевое число $n!$ стоит в этой последовательности на n-м месте. Поэтому в кольце $\mathbb{C}_e[[z]]$ любая последовательность чисел $(a_0, a_1, a_2, \dots)$ может быть записана так:

$$(a_0, a_1, a_2, \dots) = \sum_{n=0}^{+\infty} a_n \frac{z^n}{n!} =: A(z) \in \mathbb{C}_e[[z]].$$

1.9. Вернемся теперь к перечисляемому нами множеству X и рассмотрим отображение

$$w : X \longrightarrow C[[z]],$$

которое любому элементу $x \in X$ сопоставляет некоторый формальный степенной ряд вида

$$w(x) = 0\cdot 1 + 0\cdot z + \ldots + 0\cdot z^{k-1} + 1\cdot z^k + 0\cdot z^{k+1} + \ldots = z^k.$$

По сути дела, при таком подходе вес любого элемента $x \in X$ определяется степенью $k \in \mathbb{Z}_+$ особого элемента z кольца $C[[z]]$. Тогда суммирование образов $w(x)$ этого отображения по всем $x \in X$ даст нам энумератор вида

$$f(z) := \sum_{x\in X} w(x) = \sum_{n=0}^{+\infty} c_n z^n,$$

то есть формальный степенной ряд, коэффициенты c_n которого дают нам количество элементов $x \in X$, имеющих заданный вес $z^n = (0,0,\dots,0,1,0,\dots)$. Этот энумератор и называют *обыкновенной производящей функцией* множества X.

Аналогично, рассматривая отображение

$$w_e : X \longrightarrow C_e[[z]],$$

сопоставляющее любому элементу $x \in X$ формальный степенной ряд вида

$$w(x) = 0\cdot 1 + 0\cdot \frac{z}{1!} + \ldots + 0\cdot \frac{z^{k-1}}{(k-1)!} + 1\cdot \frac{z^k}{k!} + 0\cdot \frac{z^{k+1}}{(k+1)!} + \ldots = \frac{z^k}{k!},$$

мы получаем в качестве энумератора X формальный степенной ряд

$$F(z) := \sum_{x \in X} w_e(x) = \sum_{n=0}^{+\infty} c_n \frac{z^n}{n!},$$

называемый *экспоненциальной производящей функцией* множества X.

1.10. Согласно определению, каждая производящая функция является элементом кольца формальных степенных рядов. Следовательно, любую пару производящих функций можно складывать и перемножать между собой.

Формальные правила сложения и умножения этих функций, естественно, совпадают с описанными выше правилами (1)–(3) сложения и умножения отвечающих этим рядам числовых последовательностей. Именно, пусть

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots \qquad \text{и} \qquad g(z) = b_0 + b_1 z + b_2 z^2 + \ldots$$

есть пара обыкновенных производящих функций для множеств X и Y соответственно, а

$$F(z) = a_0 + a_1 \frac{z}{1!} + a_2 \frac{z^2}{2!} + \ldots \qquad \text{и} \qquad G(z) = b_0 + b_1 \frac{z}{1!} + b_2 \frac{z^2}{2!}$$

есть пара экспоненциальных производящих функций для этой пары множеств. Тогда суммой этих производящих функций называются формальные степенные ряды

$$f(z) + g(z) := (a_0 + b_0) + (a_1 + b_1) z + (a_2 + b_2) z^2 + \ldots$$

и

$$F(z) + G(z) := (a_0 + b_0) + (a_1 + b_1) \frac{z}{1!} + (a_2 + b_2) \frac{z^2}{2!} + \ldots$$

Произведением производящих функций называются формальные степенные ряды

$$h(z) = f(z) \cdot g(z) := c_0 + c_1 z + c_2 z^2 + \ldots \qquad \text{и} \qquad H(z) = F(z) \cdot G(z) := c_0 + c_1 \frac{z}{1!} + c_2 \frac{z^2}{2!} + \ldots,$$

коэффициенты c_n которых вычисляются по формулам (2) и (3) соответственно.

Важно, что сложение и умножение любых двух производящих функций имеет и вполне определенный комбинаторный смысл. Выяснение этого смысла мы отложим до одного из следующих параграфов. Сейчас же, на время забыв о перечисляемом множестве X, займемся изучением свойств производящих функций с точки зрения теории формальных степенных рядов.

2. Итак, по определению, обыкновенная и экспоненциальная производящие функции являются элементами формальных степенных рядов $C[[z]]$ и $C_e[[z]]$. Поговорим немного подробнее о таких формальных степенных рядах.

2.1. Прежде всего, подчеркнем отличие формальных степенных рядов от рядов, встречающихся в математическом анализе. По сути дела, любой формальный степенной ряд — это некоторая картинка, удобная для изображения заданной числовой последовательности $(a_0, a_1, a_2, \ldots)$. Например, рассмотрим формальный степенной ряд

$$f(z) = \sum_{n=0}^{+\infty} n!\, z^n = 1 + z + 2z^2 + 6z^3 + 24z^4 + \ldots, \tag{4}$$

отвечающий числовой последовательности $(1, 1, 2, 4, 24, \ldots)$. Символы $1, z, z^2, z^3, \ldots$ в такой форме записи играют, по сути, роль индексов — они нужны лишь для того, чтобы указывать позиции, на которых стоят элементы этой числовой последовательности. Например, если коэффициент при z^4 равен 24, то это означает лишь, что число 24 есть четвертый элемент рассматриваемой числовой последовательности.

Как следствие, в теории формальных степенных рядов символ z не надо рассматривать как комплексную переменную, которая может принимать какие-то конкретные значения. Нет смысла вычислять значения таких рядов как предел последовательности частичных сумм ни при каких конкретных значениях z. Наконец, в этой теории можно обойтись без понятия сходящихся или расходящихся рядов.

Так, например, с точки зрения математического анализа рассматривать ряд (4) смысла не имеет — он расходится при всех значениях $z > 0$. В комбинаторике же этот ряд, понимаемый как обыкновенная производящая функция, имеет вполне конкрентый и весьма важный комбинаторный смысл: коэффициенты этого ряда задают количество всех перестановок заданного n-элементного множества.

2.2. Несмотря на все вышесказанное, опыт, накопленный при работе с рядами из математического анализа, часто бывает полезен и при анализе формальных степенных рядов. Связано это, прежде всего, с тем, что многие базовые операции над рядами, такие, как сложение или умножение, вводятся одинаково как для обычных, так и для формальных степенных рядов.

`Пример 2.` Перемножим два формальных степенных ряда из $C[[z]]$:

$$f(z) = \sum_{n=0}^{+\infty} \alpha^n z^n \qquad \text{и} \qquad g(z) = 1 - \alpha z.$$

Коэффициент c_0 в произведении $h(z) = f(z) \cdot g(z)$, очевидно, равен $\alpha^0 \cdot 1 = 1$. Несложно убедиться, что остальные коэффициенты c_k равны нулю:

$$c_k = \alpha^k \cdot 1 - \alpha \cdot \alpha^{k-1} \equiv 0.$$

Следовательно,

$$f(z)\cdot g(z) = \Bigl(\sum_{n=0}^{+\infty} \alpha^n z^n\Bigr)\cdot(1-\alpha z) = 1 \qquad \Longrightarrow \qquad f(z) = \sum_{n=0}^{+\infty} \alpha^n z^n = g^{-1}(z) = (1-\alpha z)^{-1} =: \frac{1}{1-\alpha z}.$$

Заметим теперь, что формула

$$1 + \alpha z + \alpha^2 z^2 + \ldots = \frac{1}{1 - \alpha z} \qquad \forall\, \alpha \in \mathbb{C} \tag{5}$$

есть хорошо известная формула для суммы геометрической прогрессии. В математическом анализе смысл этой формулы состоит в следующем: ряд, стоящий в левой ее части, сходится к функции, стоящей в ее правой части, при всех z, таких, что $|\alpha z| < 1$. В теории формальных степенных рядов ее смысл другой: формула означает, что функции $f(z)$ и $g(z)$ являются взаимно-обратными по отношению к операции умножения в кольце $C[[z]]$ формальных степенных рядов. Однако, несмотря на смысловые отличия, вид самого тождества и в том, и в другом случае одинаков.

2.3. Отмеченная в примере 2 закономерность справедлива, как правило, и в общем случае. Именно, пусть имеется некоторое тождество со степенными рядами, которое справедливо в случае, когда эти ряды рассматриваются как некоторые аналитические функции комплексного аргумента z, заданные в некоторой общей для них области определения (например, в некоторой малой окрестности нуля). Тогда высока вероятность того, что это же тождество остается верным как некоторое соотношение между формальными степенными рядами.

С этой точки зрения доказанное в примере 2 тождество (5) может быть получено так. Рассмотрим функции комплексного аргумента

$$f(z) = \frac{1}{1-\alpha z} \qquad \text{и} \qquad g(z) = 1 - \alpha z,$$

аналитические в области $|\alpha z| < 1$. Последнее, в частности, означает, что эти функции в области $|\alpha z| < 1$ единственным образом представляются своими рядами Тейлора

$$f(z) = 1 + \alpha z + \alpha^2 z^2 + \ldots \qquad \text{и} \qquad g(z) = 1 - \alpha z + 0z^2 + \ldots.$$

Кроме того, в этой области для них справедливо очевидное равенство $f(z) \cdot g(z) = 1$, которое в терминах соответствующих им рядов записывается так:

$$f(z) \cdot g(z) = 1 \qquad \Longleftrightarrow \qquad \left(\sum_{n=0}^{+\infty} \alpha^n z^n\right) \cdot (1 - \alpha z) = 1. \tag{6}$$

Все, что теперь остается заметить — это то, что данное равенство останется справедливым, если его рассматривать и как некоторое тождество между формальными степенными рядами в $C[[z]]$ в силу того, что операции умножения для формальных степенных рядов и рядов из математического анализа совпадают.

Далее мы рассмотрим еще несколько примеров из этой же серии.

`Пример 3.` Рассмотрим функции $F(z) = \exp(z)$ и $G(z) = \exp(-z)$, аналитические на всей комплексной плоскости. Очевидное равенство $F(z) \cdot G(z) = 1$ в терминах соответствующих им рядов Тейлора выглядит так:

$$\exp(z) \cdot \exp(-z) = 1 \qquad \Longleftrightarrow \qquad \left(\sum_{n=0}^{+\infty} 1 \cdot \frac{z^n}{n!}\right) \cdot \left(\sum_{n=0}^{+\infty} (-1) \cdot \frac{z^n}{n!}\right) = 1.$$

Раскрывая в этом равенстве скобки и приравнивая коэффициенты при $z^n/n!$, получаем хорошо известное и очень полезное тождество для биномиальных коэффициентов

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = \begin{cases} 0, & \text{если } n > 0, \\ 1, & \text{если } n = 0. \end{cases}$$

`Пример 4.` Пусть $F(z)$, $G(z)$ – пара аналитических функций, связанных в некоторой общей области их определения тождеством вида

$$F(z) = G(z) \cdot e^z.$$

Очевидно, что тогда

$$G(z) = F(z) \cdot e^{-z}.$$

На языке формальных степенных рядов эти равенства отвечают формулам обращения

$$f_n = \sum_{k=0}^{n} \binom{n}{k} g_k \qquad \Longleftrightarrow \qquad g_n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} f_k.$$

3. Как уже неоднократно отмечалось, формула (6) указывает на то, что функции $f(z)$ и $g(z)$ являются взаимно-обратными по отношению к операции умножения в кольце формальных степенных рядов $\mathbb{C}[[z]]$. Возникает вопрос: какими свойствами должна обладать функция $g(z) \in \mathbb{C}[[z]]$ для того, чтобы она имела обратный по отношению к операции умножения элемент в кольце формальных степенных рядов $\mathbb{C}[[z]]$? Оказывается, необходимым и достаточным условием для этого является отличие коэффициента b_0 при $z^0 = 1$ от нуля.

3.1. Действительно, пусть $g(z) = b_0 + b_1 z + b_2 z^2 + \ldots$ и коэффициент $b_0 \neq 0$. Покажем, что в этом случае обязательно существует функция $f(z) = a_0 + a_1 z + a_2 z^2 + \ldots$, такая, что $f(z) \cdot g(z) = 1$.

Для этого перемножим два формальных степенных ряда $f(z)$ и $g(z)$, а затем приравняем в равенстве $f(z) \cdot g(z) = 1$ коэффициенты при одинаковых степенях z:

$$\begin{array}{lll}
z^0: & a_0 \cdot b_0 = 1 & \Longrightarrow \quad a_0 = \dfrac{1}{b_0} \neq 0; \\
z^1: & a_0 \cdot b_1 + a_1 \cdot b_0 = 0 & \Longrightarrow \quad a_1 = -\dfrac{1}{b_0} \cdot (a_0\, b_1); \\
z^2: & a_0 \cdot b_2 + a_1 \cdot b_1 + a_2 \cdot b_0 = 0 & \Longrightarrow \quad a_2 = -\dfrac{1}{b_0} \cdot (a_0\, b_2 + a_1\, b_1); \\
 & & \qquad\quad \ldots
\end{array}$$

Видно, что таким образом шаг за шагом можно восстановить все коэффициенты a_n формального степенного ряда $f(z)$.

Обратно, пусть $f(z) = a_0 + a_1 z + a_2 z^2 + \ldots$ и $g(z) = b_0 + b_1 z + b_2 z^2 + \ldots$ — обыкновенные производящие функции, такие, что $f(z) \cdot g(z) = 1$. Тогда, согласно формуле умножения обыкновенных производящих функций, $a_0 \cdot b_0 = 1$. Отсюда, в частности, следует, что $b_0 \neq 0$.

3.2. Пусть теперь $h(z) = c_0 + c_1 z + c_2 z^2 + \ldots$ и $g(z) = b_0 + b_1 z + b_2 z^2 + \ldots$ — два произвольных элемента кольца $\mathbb{C}[[z]]$, $b_0 \neq 0$. Пусть для определенности $b_0 = 1$. В этом случае можно рассматривать формальные дроби вида $h(z)/g(z)$ как элементы $f(z) \in \mathbb{C}[[z]]$, такие, что $f(z) \cdot g(z) = h(z)$. Говорят, что все эти элементы $f(z)$ образуют *поле частных* кольца $\mathbb{C}[[z]]$. По сути же дела, их можно рассматривать как результат *деления* формального степенного ряда $h(z)$ на формальный степенной ряд $g(z)$. Так, в рассмотренном в пункте 2.4 примере ряд

$$f(z) = 1 + \alpha z + \alpha^2 z^2 + \ldots$$

можно рассматривать как результат деления формальных степенных рядов $h(z) = 1$ и $g(z) = 1 - \alpha z$.

3.3. В качестве еще одного чрезвычайно полезного для дальнейшего изложения примера приведем следующее обобщение формулы (5):

$$\frac{1}{(1-\alpha z)^k} = \sum_{n=0}^{+\infty} \left(\!\!\binom{k}{n}\!\!\right) (\alpha\, z)^n. \qquad (7)$$

Для проверки этой формулы следует, вообще говоря, показать, что умножение стоящего в правой части равенства (7) ряда $f(z)$ на ряд $g(z) = (1-\alpha z)^k$ дает единицу для любого $k \in \mathbb{N}$. Однако проще доказать это равенство, воспользовавшись приемом, описанным в пункте 2.3 настоящего параграфа.

Действительно, рассмотрим пару функций

$$f(z) = \frac{1}{(1-\alpha z)^k} \qquad \text{и} \qquad g(z) = (1-\alpha z)^k$$

комплексного аргумента z, аналитических в общей для них области $|\alpha z| < 1$. Ясно, что в этой области справедливо равенство

$$f(z) \cdot g(z) = 1.$$

Все, что теперь осталось для доказательства тождества (7) — это разложить функцию $f(z)$ в ряд Тейлора в области $|\alpha z| < 1$.

Для этого вспомним хорошо известную в классическом анализе формулу бинома Ньютона

$$(1+z)^q = \sum_{n=0}^{+\infty} \frac{(q)_n}{n!} z^n = 1 + \sum_{n=1}^{+\infty} \frac{q\,(q-1)\,\ldots\,(q-n+1)}{n\,(n-1)\,\ldots\,1} z^n, \tag{8}$$

справедливую для любого $q \in \mathbb{C}$ и $|z| < 1$. Полагая в (8) в качестве $q = -k$, $k \in \mathbb{N}$, и заменяя в ней z на αz, получаем следующую цепочку равенств, приводящую к формуле (7):

$$\frac{1}{(1-\alpha z)^k} = (1-\alpha z)^{-k} = \sum_{n=0}^{+\infty} \frac{(-k)_n}{n!} (-\alpha\, z)^n =$$

$$= 1 + \sum_{n=1}^{+\infty} \frac{(-k)\,(-k-1)\,\ldots\,(-k-(n-1))}{n\,(n-1)\,\ldots\,1} (-1)^n (\alpha\, z)^n =$$

$$= \sum_{n=0}^{+\infty} \binom{k+n-1}{n} (\alpha\, z)^n = \sum_{n=0}^{+\infty} \left(\!\binom{k}{n}\!\right) (\alpha\, z)^n.$$

В частном случае $k = 2$ и $k = 3$ из (7) имеем следующие полезные формулы:

$$\frac{1}{(1-\alpha z)^2} = 1 + 2\,\alpha\, z + 3\,\alpha^2 z^2 + 4\,\alpha^3 z^3 + \ldots + (n+1)\alpha^n z^n + \ldots, \tag{9}$$

$$\frac{1}{(1-\alpha z)^3} = 1 + 3\,\alpha\, z + 6\,\alpha^2 z^2 + 10\,\alpha^3 z^3 + \ldots + \frac{(n+1)(n+2)}{2} \alpha^n z^n + \ldots. \tag{10}$$

3.4. *Замечание.* Все рассуждения, связанные с делением обыкновенных производящих функций как элементов кольца $C[[z]]$, легко переносятся и на элементы кольца $C_e[[z]]$. В частности, формальный степенной ряд $G(z) = b_0 + b_1 z/1! + b_2 z^2/2! + \ldots$ имеет обратный по отношению к операции умножения в $C_e[[z]]$, если $b_0 \neq 0$.

4. Еще раз подчеркнем, что построение производящей функции для заданного множества X представляет собой вполне разумный и полный ответ на некоторую комбинаторную задачу, связанную с перечислением элементов множества X. Именно, построение такой функции означает, что мы разбили X на блоки и подсчитали количество элементов в каждом блоке — это количество равно коэффициентам a_n при z^n или при $z^n/n!$.

Однако в комбинаторных задачах наряду с такими формами ответа встречаются и другие. Например, ответ часто можно получить в виде некоторого рекуррентного выражения, связывающего коэффициенты a_n между собой. В связи с этим возникает проблема перехода от одной формы ответа к другой. Способы такого перехода и будут предметом изучения следующих трех параграфов. Параллельно с этим мы изучим и еще некоторые весьма полезные операции над формальными степенными рядами.

2 Рекуррентные соотношения

1. Начнем с простого примера. Популяция лягушек в озере увеличивается в четыре раза каждый год. В последний день каждого года 100 лягушек отлавливают и переправляют на другие озера. Предполагая, что в начале первого года в озере было 50 лягушек, найти количество лягушек в начале любого последующего года.

1.1. Обозначим через a_n количество лягушек в начале $(n+1)$-го года; $a_0 = 50$. Тогда, очевидно,

$$a_1 = 4 \cdot 50 - 100 = 100, \qquad a_2 = 4 \cdot 100 - 100 = 300,$$

а в общем случае

$$a_{n+1} = 4 \cdot a_n - 100, \qquad n = 0, 1, 2, \ldots$$

Полученное равенство является простейшим примером *рекуррентного соотношения.*

1.2. *Определение.* Пусть $a_0, a_1, a_2, \ldots$ — произвольная числовая последовательность. Если для любого $n \geqslant m$ число a_{n+m} является некоторой функцией от m предыдущих членов последовательности, т.е.

$$a_{n+m} = f(a_n, a_{n+1}, \ldots, a_{n+m-1}), \tag{11}$$

то такая последовательность называется рекуррентной последовательностью, а соотношение (11) — рекуррентным соотношением m-го порядка.

1.3. В частном случае линейной функции f имеем так называемое линейное рекуррентное соотношение

$$a_{n+m} = b_1(n)\, a_{n+m-1} + b_2(n)\, a_{n+m-2} + \ldots + b_{m-1}(n)\, a_{n+1} + b_m(n)\, a_n + u_n. \tag{12}$$

В случае $u_n = 0$ оно называется однородным, в противном случае — неоднородным.

1.4. Самый простой случай рекуррентного соотношения – это линейное однородное рекуррентное соотношение с постоянными коэффициентами

$$a_{n+m} = b_1\, a_{n+m-1} + b_2\, a_{n+m-2} + \ldots + b_{m-1}\, a_{n+1} + b_m\, a_n. \tag{13}$$

1.5. Очевидно, что для однозначного определения всех a_n необходимо наряду с самим рекуррентным соотношением задать и первые m членов $a_0, a_1, \ldots, a_{m-1}$ данной последовательности, т.е., как говорят, задать начальные условия для рекуррентного соотношения.

1.6. Итак, наличие рекуррентного соотношения и начальных условий позволяет нам последовательно, шаг за шагом, определить любое наперед заданное количество n членов рекуррентной числовой последовательности. Зачастую это все, что нам требуется от задачи; иными словами, получение рекуррентного соотношения для искомой числовой последовательности a_n является вполне приемлемым, а зачастую — и наиболее удобным с вычислительной точки зрения ответом поставленной комбинаторной задачи.

Однако иногда нам хочется получить явное аналитическое выражение для общего члена a_n этой последовательности, или, как говорят, *решить* данное рекуррентное соотношение. Мы в данном параграфе покажем, как построить такое решение в случае линейного однородного рекуррентного соотношения с постоянными коэффициентами (13).

2. Прежде чем рассматривать общий случай уравнения (13), рассмотрим наиболее простой его вариант — линейное однородное рекуррентное соотношение первого порядка

$$a_{n+1} = b_1 \cdot a_n, \qquad n = 0, 1, 2, \ldots; \qquad a_0 \text{ — заданное число.} \tag{14}$$

2.1. Решение этого уравнения построить легко. Действительно,

$$a_1 = b_1 \cdot a_0; \qquad a_2 = b_1 \cdot a_1 = b_1^2 \cdot a_0; \qquad \ldots \qquad a_n = b_1^n \cdot a_0.$$

2.2. Предположим теперь, что нам кто-то сразу подсказал вид решения, а именно, что решение нашего уравнения (14) степенным образом зависит от n:

$$a_n = r^n \qquad \text{для некоторого } r\,.$$

Воспользовавшись этой подсказкой, подставим это выражение в исходное рекуррентное соотношение (14). В результате получается равенство вида

$$r^{n+1} = b_1 \cdot r^n,$$

из которого сразу же следует, что $r = b_1$; при этом a_n оказывается равным b_1^n. Это означает, что при $n = 0$ число $a_0 = b_1^0 = 1$. Иными словами, $a_n = b_1^n$ есть решение исходной задачи (14) в частном случае $a_0 = 1$. Поэтому такое решение называется *частным решением* уравнения (14).

2.3. Теперь попытаемся, зная частное решение, построить общее решение задачи (14). Для этого заметим, что любое частное решение уравнения (14), умноженное на произвольную постоянную c, по-прежнему этому уравнению удовлетворяет:

$$c \cdot b_1^{n+1} \equiv c \cdot b_1 \cdot b_1^n.$$

Решение же вида $c \cdot b_1^n$ позволяет удовлетворить любому начальному условию. Действительно, подставляя его в начальное условие для уравнения (14), получим:

$$c \cdot b_1^0 = a_0 \qquad \Longrightarrow \qquad c = a_0 \qquad \Longrightarrow \qquad a_n = a_0 \cdot b_1^n.$$

По этой причине решение вида $c \cdot b_1^n$ называется *общим решением* уравнения (14).

2.4. Подведем предварительные итоги. Предположив, что уравнение (14) степенным образом зависит от n, мы получили частное его решение вида $a_n = b_1^n$. Домножив это решение на произвольную постоянную, мы построили общее решение уравнения (14). Начальное условие позволило нам определить точное значение этой постоянной.

3. Рассмотрим теперь линейное однородное рекуррентное соотношение второго порядка

$$a_{n+2} = b_1 \cdot a_{n+1} + b_2 \cdot a_n, \qquad n = 0, 1, 2, \ldots, \qquad a_0\,,\ a_1 \text{ — заданные числа,} \tag{15}$$

и попытаемся применить к этому уравнению алгоритм, описанный в конце предыдущего пункта.

3.1. Для этого вновь предположим, что частное решение уравнения (15) имеет вид $a_n = r^n$. Подставляя это выражение в рекуррентное соотношение, получим

$$r^{n+2} = b_1 \cdot r^{n+1} + b_2 \cdot r^n \qquad \Longrightarrow \qquad r^2 - b_1\, r - b_2 = 0,$$

т.е. квадратное уравнение на r, любое решение r_0 которого есть некоторое частное решение уравнения (15). По этой причине данное уравнение называется характеристическим уравнением для рекуррентного соотношения (15).

3.2. Рассмотрим вначале случай, когда характеристическое уравнение имеет два различных вещественных корня r_1 и r_2. Покажем, что в таком случае выражение

$$a_n = c_1 r_1^n + c_2 r_2^n, \tag{16}$$

где c_1 и c_2 — произвольные постоянные, является *общим решением* соотношения (15) в том смысле, что любое решение (15) с заданными начальными условиями в нем содержится.

Действительно, покажем вначале, что выражение (16) действительно удовлетворяет рекуррентному соотношению (15):

$$c_1 r_1^{n+2} + c_2 r_2^{n+2} = b_1 c_1 r_1^{n+1} + b_1 c_2 r_2^{n+1} + b_2 c_1 r_1^n + b_2 c_2 r_2^n =$$
$$= c_1 (r_1^2 - b_1 r_1 - b_2) + c_2 (r_1^2 - b_1 r_1 - b_2) \equiv 0.$$

Покажем теперь, что это — действительно общее решение, т.е. что мы всегда можем подобрать константы c_1 и c_2 так, чтобы решение вида (16) удовлетворяло любым заданным начальным условиям. Для этого рассмотрим это выражение при $n = 0$ и $n = 1$:

$$a_0 = c_1 r_1^0 + c_2 r_2^0 = c_1 + c_2,$$
$$a_1 = c_1 r_1^1 + c_2 r_2^1 = c_1 r_1 + c_2 r_2.$$

Эти выражения следует рассматривать как систему линейных уравнений для определения неизвестных постоянных c_1 и c_2. Определитель этой системы

$$\begin{vmatrix} 1 & 1 \\ r_1 & r_2 \end{vmatrix} = r_2 - r_1 \neq 0,$$

поэтому система всегда имеет единственное решение.

3.3. `Упражнение.` Показать, что в случае равных корней характеристического уравнения общее решение уравнения (15) имеет вид

$$a_n = c_1 r^n + c_2 n r^n,$$

а в случае комплексных корней это решение записывается в виде

$$a_n = c_1 \rho^n \cos(n \vartheta) + c_2 \rho^n \sin(n \vartheta),$$

где ρ и ϑ — модуль и аргумент одного из двух комплексных корней характеристического уравнения.

3.4. Пожалуй, самым известным и важным примером рекуррентного соотношения (15) является соотношение вида

$$F_{n+2} = F_{n+1} + F_n, \qquad n = 0, 1, 2, \ldots, \qquad F_0 = 0, \;\; F_1 = 1,$$

определяющее так называемые числа Фибоначчи F_n. Характеристическое уравнение для этого рекуррентного уравнения имеет вид

$$r^2 = r + 1 \qquad \Longrightarrow \qquad r_{1,2} = \frac{1 \pm \sqrt{5}}{2} \qquad \Longrightarrow \qquad F_n = c_1 \left(\frac{1 + \sqrt{5}}{2}\right)^n + c_2 \left(\frac{1 - \sqrt{5}}{2}\right)^n.$$

Константы c_1 и c_2 определим из начальных условий:

$$F_0 = 0 = c_1 + c_2$$
$$F_1 = 1 = c_1 \frac{1+\sqrt{5}}{2} + c_2 \frac{1-\sqrt{5}}{2} = c_1 \left[\frac{1+\sqrt{5}-1+\sqrt{5}}{2}\right] = c_1 \sqrt{5} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad c_1 = +\frac{1}{\sqrt{5}}, \qquad c_2 = -\frac{1}{\sqrt{5}}.$$

Таким образом, окончательно получаем следующее явное выражение для чисел Фибоначчи:

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

4. В принципе, описанный выше способ решения рекуррентных соотношений можно пытаться распространять и на соотношения более высокого порядка, и на неоднородные линейные рекуррентные соотношения. Вместо этого, однако, мы в следующем параграфе продемонстрируем более универсальный и удобный способ решения таких соотношений, использующий аппарат производящих функций.

3 Линейные рекуррентные соотношения с постоянными коэффициентами и обыкновенные производящие функции

1. Итак, в предыдущем параграфе мы показали, как по имеющемуся линейному однородному рекуррентному соотношению с постоянными коэффициентами, связывающему члены a_n числовой последовательности, найти явный их вид как функций параметра n. Как следствие, мы получаем и представление решения в виде обыкновенной производящей функции

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots .$$

Однако в большинстве задач такой порядок построения решения не является оптимальным. На практике, как правило, легче по имеющемуся рекуррентному соотношению построить производящую функцию, используя описанные в первом параграфе данной главы операции над формальными степенными рядами. Проиллюстрируем вышесказанное на простом примере — сформулированной в начале предыдущего параграфа задаче о лягушках.

1.1. Напомним, что для этой задачи нами было получено следующее линейное неоднородное рекуррентное соотношение первого порядка, описывающее изменение популяции лягушек в озере:

$$a_{n+1} = 4\,a_n - 100, \qquad n = 0, 1, 2, \ldots, \qquad a_0 = 50. \tag{17}$$

Покажем, как восстановить для этой рекуррентной числовой последовательности отвечающую ей обыкновенную производящую функцию.

1.2. Пусть

$$f(z) = a_0 + a_1\, z + a_2\, z^2 + \ldots + a_n z^n + \ldots$$

есть искомая обыкновенная производящая функция для числовой последовательности a_n, удовлетворяющей уравнению (17). Домножим (17) на z^{n+1} и просуммируем полученное уравнение по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} a_{n+1}\, z^{n+1} = 4 \sum_{n=0}^{+\infty} a_n\, z^{n+1} - 100 \sum_{n=0}^{+\infty} z^{n+1}. \tag{18}$$

Левая часть этого равенства — это "почти" $f(z)$; переходя в этой сумме к новому индексу суммирования $k = n+1$, $k = 1, \ldots, +\infty$, получаем

$$\sum_{k=1}^{+\infty} a_k\, z^k = f(z) - a_0.$$

Первая сумма в правой части равенства (18), очевидно, равна

$$4 \sum_{n=0}^{+\infty} a_n\, z^{n+1} = 4\, z \sum_{n=0}^{+\infty} a_n\, z^n = 4\, z\, f(z).$$

Наконец, последнюю сумму в (18) можно свернуть и записать в виде дроби

$$\sum_{n=0}^{+\infty} z^{n+1} = z \sum_{n=0}^{+\infty} z^n = \frac{z}{1-z}.$$

Поэтому окончательно равенство (18) переписывается в виде

$$f(z) - a_0 = 4\, z\, f(z) - 100 \frac{z}{1-z} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad f(z) = \frac{a_0}{1-4z} - \frac{100\, z}{(1-z)(1-4z)}. \tag{19}$$

1.3. Итак, мы построили производящую функцию для числовой последовательности a_n. Наша же исходная задача заключалась в отыскании явного выражения для этих чисел. Оказывается, что теперь это сделать довольно просто — достаточно разложить правую часть (19) в ряд по степеням z.

С первым слагаемым в правой части (19) справиться легко — мы знаем, что

$$g(z) = \frac{1}{1-4z} = 1 + 4z + (4z)^2 + \ldots.$$

Для того, чтобы проделать ту же операцию со вторым слагаемым, нам предварительно необходимо разложить эту дробь на простейшие:

$$\frac{z}{(1-z)(1-4z)} = \frac{A}{1-z} + \frac{B}{1-4z} = \frac{A - 4Az + B - Bz}{(1-z)(1-4z)} \Longrightarrow \begin{cases} A + B = 0 \\ 4A + B = -1 \end{cases} \Longrightarrow \begin{cases} A = -1/3 \\ B = 1/3 \end{cases}$$

Как следствие,

$$-\frac{100\, z}{(1-z)(1-4z)} = -\frac{100}{3} \left[\sum_{n=0}^{+\infty} (4^n - 1)\, z^n \right],$$

и мы окончательно для a_n получаем следующее явное аналитическое выражение:

$$a_n = a_0\, 4^n - \frac{100}{3} \left[4^n - 1\right] = \frac{50}{3} \left[4^n + 2\right], \qquad n = 0, 1, 2, \ldots.$$

2. Итак, опишем общий алгоритм использования аппарата обыкновенных производящих функций для решения линейных рекуррентных соотношений m-го порядка с постоянными коэффициентами.

1) Для числовой последовательности $\{a_n\}$ вводится обыкновенная производящая функция $f(z)$.
2) Рекуррентное соотношение трансформируется в уравнение для $f(z)$ следующим образом: это соотношение домножается на z^{n+m}, суммируется по n от 0 до $+\infty$, а затем каждая из сумм выражается через $f(z)$.
3) Полученное уравнение формально разрешается относительно $f(z)$.
4) Числа a_n находятся как коэффициенты при z^n в разложении $f(z)$ по степеням z.

Применим этот алгоритм к линейному неоднородному рекуррентному соотношению m-го порядка с постоянными коэффициентами, записанному в следующем виде:

$$b_0 a_{n+m} + b_1 a_{n+m-1} + \ldots + b_m a_n = u_n, \quad n = 0, 1, 2, \ldots, \quad b_0 \neq 0, \;\; a_0, a_1, \ldots, a_{m-1} \text{ – заданы.} \tag{20}$$

2.1. Введем для последовательности $a_0, a_1, a_2, \ldots$ обыкновенную производящую функцию

$$f(z) = a_0 + a_1\, z + a_2\, z^2 + \ldots = \sum_{n=0}^{+\infty} a_n z^n.$$

2.2. Домножим (20) на z^{n+m} и просуммируем полученное уравнение по n от 0 до $+\infty$. В полученном соотношении разберем каждое слагаемое отдельно:

$$\begin{aligned}
b_0 \sum_{n=0}^{+\infty} a_{n+m} z^{n+m} &= b_0\, [f(z) - a_0 - a_1 z - \ldots - a_{m-1} z^{m-1}],\\
b_1 z \sum_{n=0}^{+\infty} a_{n+m-1} z^{n+m-1} &= b_1 z\, [f(z) - a_0 - a_1 z - \ldots - a_{m-2} z^{m-2}],\\
&\ldots\\
b_{m-2} z^{m-2} \sum_{n=0}^{+\infty} a_{n+2} z^{n+2} &= b_{m-2} z^{m-2}\, [f(z) - a_0 - a_1 z],\\
b_{m-1} z^{m-1} \sum_{n=0}^{+\infty} a_{n+1} z^{n+1} &= b_{m-1} z^{m-1}\, [f(z) - a_0],\\
b_m z^m \sum_{n=0}^{+\infty} a_n z^n &= b_m z^m f(z),\\
\sum_{n=0}^{+\infty} u_n z^n &=: z^m u(z).
\end{aligned}$$

Введем также следующие обозначения:

$$g(z) := b_0 + b_1 z + b_2 z^2 + \ldots + b_{m-1} z^{m-1} + b_m z^m,$$

$$c_n = \sum_{i=0}^{n} a_i b_{n-i}, \qquad h(z) := \sum_{n=0}^{m-1} c_n z^n.$$

С учетом этих обозначений и сделанных выше преобразований получается следующее уравнение на производящую функцию $f(z)$:

$$f(z)\cdot g(z) = b_0a_0+(b_0a_1+b_1a_0)z+\ldots+(b_0a_{n+m-1}+b_1a_{n+m-2}+\ldots+b_{m-2}a_1+b_{m-1}a_0)z^{m-1}+z^m u(z) =$$

$$= \sum_{n=0}^{m-1} c_n z^n + z^m u(z) = h(z) + z^m u(z) \qquad \Longrightarrow$$

$$\Longrightarrow \qquad f(z) = \frac{h(z) + z^m u(z)}{g(z)}.$$

Заметим, что деление на $g(z) = b_0 + b_1 z + \ldots + b_m z^m$ законно — коэффициент b_0 отличен от нуля.

2.3. В случае *однородного* линейного рекуррентного соотношения с постоянными коэффициентами m-го порядка производящая функция для рекуррентной последовательности $a_0, a_1, a_2, \ldots$ представляет собой рациональную функцию

$$f(z) = \frac{h(z)}{g(z)} = \frac{\sum\limits_{n=0}^{m-1} c_n z^n}{\sum\limits_{n=0}^{m} b_n z^n}.$$

По сути, этим устанавливается взаимно-однозначное соответствие между линейными однородными рекуррентными соотношениями на коэффициенты a_n и рациональными производящими функциями.

2.4. Как правило, при вычислении явного вида коэффициентов a_n вместо формального деления степенных рядов целесообразно разложить дробь на простейшие — элементарные дроби вида

$$\frac{A}{1-\alpha\, z}, \qquad \frac{B}{(1-\alpha\, z)^2}, \qquad \ldots \qquad \frac{D}{(1-\alpha\, z)^m},$$

а затем воспользоваться готовыми формулами вида (7).

4 Линейные рекуррентные соотношения с переменными коэффициентами

1. Перейдем теперь к более общему виду линейного рекуррентного соотношения (12) с переменными коэффициентами. Как вскоре мы увидим, при решении таких уравнений с помощью производящих функций естественным образом возникают производные таких функций. Однако производящая функция — это формальный степенной ряд, поэтому понятие производной для такой функции требует специального определения.

1.1. *Определение.* Пусть $(a_0, a_1, a_2, \ldots)$ — некоторая числовая последовательность, и пусть

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots$$

есть формальный степенной ряд для этой последовательности, понимаемый как элемент кольца $\mathbb{C}[[z]]$. Производной ряда $f(z)$ называется формальный степенной ряд вида

$$a_1 + 2a_2 z + 3a_3 z^2 + \ldots = \sum_{n=1}^{+\infty} n\, a_n z^{n-1} := f'(z).$$

Иными словами, производной числовой последовательности $(a_0, a_1, a_2, \ldots)$ в кольце с операцией умножения (2) называется числовая последовательность

$$(0 \cdot a_0, 1 \cdot a_1, 2 \cdot a_2, 3 \cdot a_3, \ldots).$$

1.2. *Определение.* Пусть

$$F(z) = a_0 + a_1 \frac{z}{1!} + a_2 \frac{z^2}{2!} + \ldots + a_n \frac{z^n}{n!} + \ldots \in \mathbb{C}_e[[z]]$$

есть формальный степенной ряд для числовой последовательности $(a_0, a_1, a_2, \ldots)$. Производной этого ряда называется формальный степенной ряд вида

$$F'(z) = a_1 + a_2 \frac{z}{1!} + a_3 \frac{z^2}{2!} + \ldots + a_{n+1} \frac{z^n}{n!} + \ldots$$

Иными словами, "экспоненциальной" производной числовой последовательности $(a_0, a_1, a_2, \ldots)$ является сдвинутая на одну позицию влево числовая последовательность $(a_1, a_2, a_3, \ldots)$.

1.3. Для операции взятия производной в кольцах $\mathbb{C}[[z]]$ и $\mathbb{C}_e[[z]]$ формальных степенных рядов можно выводить свойства, аналогичные привычным нам свойствам производной из классического анализа. При этом вывод этих свойств зачастую оказывается даже более простым по сравнению с аналогичным выводом в курсе математического анализа.

Пример 1. Докажем, что если $f(z) \in \mathbb{C}[[z]]$ и $f'(z) = 0$, то $f(z) = a_0$, т.е. $f(z)$ отвечает числовая последовательность вида $(a_0, 0, 0, \ldots)$.

Равенство двух формальных степенных рядов означает равенство коэффициентов при соответствующих степенях z. Поэтому равенство $f'(z) = 0$ означает, что все коэффициенты левого формального степенного ряда равны нулю:

$$n \cdot a_n = 0 \quad \forall\, n = 1, 2, \ldots \qquad \Longrightarrow \qquad a_1 = a_2 = \ldots = 0 \qquad \Longrightarrow \qquad f(z) = a_0 + 0 \cdot z + 0 \cdot z^2 + \ldots = a_0.$$

Очевидно, что это же свойство выполняется и для функции $F(z) \in \mathbb{C}_e[[z]]$.

Пример 2. Докажем, что если $f(z) \in \mathbb{C}[[z]]$ и $f'(z) = f(z)$, то

$$f(z) = c \cdot \left[1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots\right] =: c \cdot e^z.$$

Как и в предыдущем примере, приравняем коэффициенты рядов $f(z)$ и $f'(z)$ при одинаковых степенях z:

$$n \cdot a_n = a_{n-1} \quad \forall\, n = 1, 2, \ldots \quad \Longrightarrow \quad a_n = \frac{a_{n-1}}{n} = \frac{a_{n-2}}{n\,(n-1)} = \ldots = \frac{a_0}{n!} \quad \Longrightarrow \quad f(z) = a_0 \cdot e^z.$$

1.4. *Замечание.* Видно, что и определение, и свойства операции дифференцирования формальных степенных рядов совпадают с определением и свойствами операции дифференцирования функций $f(z)$ комплексного аргумента, аналитических в некоторой малой окрестности точки $z = 0$. Этим обстоятельством активно пользуются при решении конкретных рекуррентных соотношений.

2. Перейдем теперь к анализу линейных рекуррентных соотношений с переменными коэффициентами. Начнем, как всегда, с простого примера.

2.1. Рассмотрим числовую последовательность, записанную в следующем явном виде:

$$a_n = \binom{2n}{n}, \qquad n = 0, 1, 2, \ldots$$

Так как

$$a_{n+1} = \binom{2n+2}{n+1} = \frac{(2n+2)!}{(n+1)! \cdot (n+1)!} = \frac{2(2n+1)(2n)!}{(n+1) \cdot n! \cdot n!} = \frac{2(2n+1)}{n+1}\, a_n,$$

то числовая последовательность a_n удовлетворяет рекуррентному соотношению

$$(n+1)a_{n+1} = 4na_n + 2a_n, \qquad n = 0, 1, 2, \ldots, \qquad a_0 = 1.$$

2.2. Итак, мы решили в каком-то смысле обратную задачу — мы из явной формулы для коэффициентов a_n вывели рекуррентное соотношение, которому эти коэффициенты удовлетворяют. Постараемся теперь с использованием обыкновенных производящих функций решить прямую задачу, а именно, по заданному рекуррентному соотношению определить явный вид чисел a_n.

Для этого домножим наше рекуррентное соотношение на z^n и просуммируем его по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} (n+1)a_{n+1}z^n = 4\sum_{n=0}^{+\infty} n\, a_n z^n + 2\sum_{n=0}^{+\infty} a_n z^n = 4z\sum_{n=1}^{+\infty} n\, a_n z^{n-1} + 2\sum_{n=0}^{+\infty} a_n z^n.$$

По определению производной обыкновенной производящей функции,

$$\sum_{n=1}^{+\infty} n\, a_n z^{n-1} = \sum_{n=0}^{+\infty} (n+1)a_{n+1}z^n = f'(z).$$

Поэтому предыдущее равенство можно записать в следующем компактном виде:

$$(1-4z)\, f'(z) = 2f(z), \qquad f(0) := a_0 = 1. \tag{21}$$

2.3. Возникает вопрос, как из этого уравнения определить обыкновенную производящую функцию $f(z)$. Стандартный прием здесь состоит в следующем. Соотношение вида (21) рассматривается как задача Коши для обыкновенного дифференциального уравнения первого порядка. Пусть функция $f(z)$, представляющая собой решение этой задачи, является аналитической функцией комплексного аргумента в некоторой малой окрестности начала координат. В этом случае она единственным образом раскладывается в этой окрестности в степенной ряд вида

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots .$$

При этом, так как операции дифференцирования и умножения таких рядов и формальных степенных рядов $f(z) \in \mathbb{C}[[z]]$ совпадают, то коэффициенты a_n этого ряда удовлетворяют исходному рекуррентному соотношению.

В рассматриваемом примере задача (21) легко решается методом разделения переменных:

$$\frac{df}{f} = \frac{2\,dz}{1-4z} = -\frac{1}{2}\frac{d(1-4z)}{(1-4z)} \quad \Longleftrightarrow \quad d\ln(f) = -\frac{1}{2}d\ln(1-4z) \quad \Longrightarrow \quad f(z) = \frac{1}{\sqrt{1-4z}}.$$

Полученная функция $f(z)$ комплексного аргумента z является аналитической в окрестности точки $z=0$ — она раскладывается в степенной ряд в этой окрестности по формуле бинома Ньютона:

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

$$= 1+\sum_{n=1}^{+\infty}\frac{1\cdot 3\cdot 5\cdot\ldots\cdot(2n-1)}{n!}2^n z^n = \left|_{n!2^n=(1\cdot 2)(2\cdot 2)(3\cdot 2)\ldots(n\cdot 2)=2\cdot 4\cdot 6\cdot\ldots\cdot 2n}\right| =$$

$$= 1+\sum_{n=1}^{+\infty}\frac{1\cdot 2\cdot 3\cdot\ldots\cdot(2n-1)(2n)}{(n!)^2}z^n = \sum_{n=0}^{+\infty}\binom{2n}{n}z^n.$$

2.4. Заметим, что условие аналитичности функции $f(z)$ в окрестности нуля является в данном алгоритме существенным. Оно обычно выполняется лишь в том случае, если соответствующие коэффициенты a_n растут не слишком быстро. В противном случае описанный выше алгоритм решения рекуррентных соотношений с переменными коэффициентами может перестать работать.

В качестве характерного примера рассмотрим следующее несложное линейное рекуррентное соотношение с переменными коэффициентами:

$$a_{n+1} = (n+1)\,a_n, \qquad n = 0,1,2,\ldots, \qquad a_0 = 1. \tag{22}$$

Попытаемся решить его с помощью обыкновенных производящих функций. Домножая рекуррентное соотношение на z^{n+1} и суммируя по n от 0 до $+\infty$, имеем

$$\sum_{n=0}^{+\infty}a_{n+1}z^{n+1} = \sum_{n=0}^{+\infty}n\,a_n z^{n+1} + \sum_{n=0}^{+\infty}a_n z^{n+1} = z^2\sum_{n=0}^{+\infty}n\,a_n z^{n-1} + z\sum_{n=0}^{+\infty}a_n z^n \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \qquad f(z)-1 = z^2 f'(z) + zf(z), \qquad f(0)=1.$$

Полученная задача Коши не имеет решения, аналитического в окрестности начала координат. Этот результат в данном случае легко объясним. Действительно, исходное рекуррентное соотношение (22) настолько простое, что мы легко можем получить явное выражение для коэффициентов a_n:

$$a_{n+1} = (n+1)\,a_n = (n+1)\,n\,a_{n-1} = \ldots = (n+1)!\,a_0 = (n+1)!$$

Отвечающий этой числовой последовательности степенной ряд

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots + n! z^n + \ldots = 1 + z + 2z^2 + 6z^3 + \ldots + n!\,z^n + \ldots,$$

как уже отмечалось в параграфе 1, расходится при любых $z > 0$, если рассматривать его с точки зрения обычного математического анализа.

2.5. Полученный выше не слишком утешительный результат не означает, однако, что рекуррентное соотношение (22) невозможно решить с помощью аппарата формальных степенных рядов. Для того, чтобы понять, как подправить описанный выше алгоритм, заметим следующее: числовой последовательности $a_n = n!$ в кольце $\mathbb{C}_e[[z]]$ формальных степенных рядов отвечает ряд

$$F(z) = 1 + 1! \cdot \frac{z^1}{1!} + 2! \cdot \frac{z^2}{2!} + \ldots + n! \cdot \frac{z^n}{n!} + \ldots = 1 + z + z^2 + \ldots + z^n + \ldots$$

Такому степенному ряду в математическом анализе отвечает функция $1/(1-z)$, аналитическая в области $|z| < 1$. Следовательно, есть надежда, что заменяя в алгоритме обыкновенную производящую функцию на экспоненциальную, мы сможем добиться успеха.

Действительно, введем для числовой последовательности a_n, описываемой рекуррентным соотношением (22), экспоненциальную производящую функцию

$$F(z) = a_0 + a_1 \frac{z}{1!} + a_2 \frac{z^2}{2!} + \ldots + a_n \frac{z^n}{n!} + \ldots$$

Домножим (22) на $z^{n+1}/(n+1)!$ и просуммируем его по n:

$$\sum_{n=0}^{+\infty} a_{n+1} \frac{z^{n+1}}{(n+1)!} = \sum_{n=0}^{+\infty} a_n \frac{z^{n+1}}{n!} \qquad \Longleftrightarrow \qquad F(z) - 1 = z\,F(z) \qquad \Longrightarrow \qquad F(z) = \frac{1}{1-z}.$$

Итак, в случае, когда коэффициенты a_n, отвечающие линейному рекуррентному соотношению с переменными коэффициентами, растут слишком быстро, разумно для решения этого соотношения использовать экспоненциальные производящие функции.

2.6. В связи с последним утверждением возникает естественный вопрос: а что, если числовая последовательность a_n будет расти столь быстро, что и отвечающий ей ряд $F(z)$, понимаемый в смысле математического анализа, будет расходиться всюду в окрестности нуля? В принципе такие примеры придумать можно. Однако на практике такие задачи, как правило, все же не встречаются.

2.7. Разумеется, не все линейные рекуррентные соотношения с переменными коэффициентами сводятся к линейным алгебраическим уравнениям на экспоненциальные производящие функции. Часто соответствующие этим соотношениям уравнения содержат производные экспоненциальных производящих функций.

Рассмотрим, к примеру, рекуррентное соотношение

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k, \qquad B_0 = 1 \tag{23}$$

для чисел Белла B_n, описывающих количество всевозможных разбиений n-элементного множества. Домножим это равенство на $z^n/n!$ и просуммируем по n от нуля до $+\infty$:

$$\sum_{n=0}^{+\infty} B_{n+1} \frac{z^n}{n!} = \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \binom{n}{k} B_k \frac{z^n}{n!}.$$

В левой части этого равенства стоит, по определению, производная $B'(z)$ рассматриваемой экспоненциальной производящей функции. Правая его часть представляет собой произведение пары экспоненциальных функций — $B(z)$ и функции

$$1 + 1\cdot\frac{z}{1!} + 1\cdot\frac{z^2}{2!} + \ldots + 1\cdot\frac{z^n}{n!} + \ldots := e^z.$$

Следовательно, в терминах экспоненциальных производящих функций равенство (23) записывается так:

$$B'(z) = e^z B(z), \qquad B(0) := B_0 = 1.$$

Рассмотрим теперь последние равенства как задачу Коши для функции $B(z)$ комплексного аргумента z. Эта задача легко решается:

$$d\ln B(z) = de^z, \qquad B(0) = 1 \qquad \Longrightarrow \qquad B(z) = e^{e^z-1}.$$

Заметим, что получившаяся в результате решения функция $B(z)$ с точки зрения математического анализа представляет собой хоть и быстрорастущую, но аналитическую функцию комплексного аргумента при любом $z \in \mathbb{C}$.

5 Числа Каталана. Нелинейные рекуррентные соотношения

1. Начнем с характерного примера — с подсчета так называемых правильных скобочных последовательностей.

1.1. Как известно, порядок вычислений в любом арифметическом выражении можно однозначно задать расстановкой скобок. Давайте возьмем какое-то достаточно произвольное арифметическое выражение, например,

$$(3-1)\cdot(4+(15-9)\cdot(2+6)),$$

и сотрем в нем все числа и знаки арифметических операций. В результате такого действия мы получим последовательность открывающихся и закрывающихся скобок

$$()(()()),$$

представляющую собой так называемую правильную скобочную последовательность.

1.2. *Определение.* Правильная скобочная последовательность — это строка, состоящая из n открывающихся и n закрывающихся скобок, обладающая следующим свойством: при проходе вдоль этой структуры слева направо количество открывающихся скобок всегда больше или равно количеству закрывающихся скобок.

1.3. Перечислим все правильные скобочные последовательности с числом *пар* скобок $n = 1, 2, 3$:

$n = 1:$	$()$	– 1 последовательность;
$n = 2:$	$()(),\ (())$	– 2 последовательности;
$n = 3:$	$()()(),\ ()(()),\ (())(),\ (()()),\ ((()))$	– 5 последовательностей.

Числа C_n, описывающие количество таких последовательностей, называются *числами Каталана.* Как мы увидим несколько позднее, удобно по определению положить $C_0 = 1$.

2. Последовательность чисел Каталана

$$1, 1, 2, 5, 14, 42, 132, 429, 1430, \ldots$$

(последовательность $A000108$ в `OEIS` (`oeis.org`)) встречается в огромном количестве различных комбинаторных задач. В книге Р.Стенли "Перечислительная комбинаторика", том 2, приведено порядка 100 примеров, в которых эти числа появляются. Приведем лишь несколько наиболее важных из них.

2.1. В качестве первого примера рассмотрим очень простую интерпретацию правильной скобочной последовательности — задачу об очереди в кассу (см. Н.Я.Виленкин). Предположим, что у кассы кинотеатра стоит очередь, состоящая из $2n$ человек. У половины из них имеется по 100 рублей, у второй половины — по 50 рублей. Билет в кино стоит 50 рублей. В начале продажи билетов касса кинотеатра пуста. Спрашивается, сколькими способами можно расставить людей в очереди правильно, т.е. так, чтобы никому не пришлось ждать у кассы сдачу.

Очевидно, что между этой задачей и задачей о количестве правильных скобочных последовательностей имеется биекция: любому человеку, имеющему 50 рублей, отвечает открывающаяся скобка в правильной скобочной последовательности, а человеку со 100 рублями — закрывающаяся скобка.

2.2. Задачу об очереди в кассу часто формулируют более формально, а именно, как задачу о подсчете количества различных *слов Дика* длины $2n$. Словом Дика называют строку длины $2n$ над алфавитом, состоящим из двух символов (например, X и Y), в которой количество символов X и Y совпадает, и в которой никакой начальный сегмент строки не содержит символов Y больше, чем символов X. В этой формулировке биекция с задачей о подсчете правильных скобочных последовательностей имеет, очевидно, вид

$$(\;\longleftrightarrow\; X, \qquad\qquad) \;\longleftrightarrow\; Y.$$

2.3. К задаче о подсчете правильных скобочных последовательностей сводится, очевидно, и задача о перечислении путей на плоскости, выходящих из начала координат, состоящих из отрезков $(1, 1)$ и $(1, -1)$, заканчивающихся в точке $(2n, 0)$ на оси абсцисс и нигде не пересекающих эту ось. Такие пути на плоскости называются *путями Дика.* Биекция с правильными скобочными последовательностями здесь такова:

$$(\;\longleftrightarrow\; \text{вектор } (1,1), \qquad\qquad) \;\longleftrightarrow\; \text{вектор } (1,-1).$$

`Упражнение 1.` Доказать, что числу Каталана также равно количество путей на плоскости, выходящих из начала координат, приходящих в точку с координатами (n, n), состоящих из отрезков $(0, 1)$ и $(1, 0)$, и не поднимающихся выше диагонали $x = y$.

2.4. Одной из основных дискретных структур, использующихся в теории алгоритмов, является плоское корневое бинарное дерево, иногда называющаяся просто бинарным деревом. Рекурсивно такую структуру можно задать следующим образом:

1) пустое плоское корневое бинарное дерево является, по определению, плоским корневым бинарным деревом;

2) пусть r — это некоторый элемент данных, ρ, λ — несвязные (т.е. не имеющие общих элементов) плоские корневые бинарные деревья, не содержащие элемента r; тогда объединение $\beta = (\lambda, r, \rho)$ этих объектов является плоским корневым бинарным деревом, в котором r является корнем, λ — левым, а ρ — правым поддеревом дерева β.

Иными словами, плоское корневое бинарное дерево — это корневое дерево, в котором любая вершина имеет ровно двух (возможно, пустых) потомков — левого и правого.

`Упражнение 2.` Установить биекцию таких структур с правильными скобочными последовательностями, доказав, тем самым, что количество таких деревьев равно числу Каталана C_n.

`Упражнение 3.` Доказать, что количество *полных* плоских корневых бинарных деревьев (т.е. деревьев, у которых любая вершина имеет либо ровно двух потомков, либо ни одного) с $(n+1)$-м листом равно числу Каталана C_n.

2.5. Рассмотрим выражение вида

$$a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_n \cdot a_{n+1},$$

где элементы a_i принадлежат множеству S с введенной на нем неассоциативной бинарной операцией $'\cdot'$. Очевидно, что это выражение не имеет смысла до тех пор, пока мы не расставим скобки так, чтобы указать на последовательность проводимых операций.

`Упражнение 4.` Докажите, что количество расстановки скобок в описанном выше выражении есть число Каталана C_n.

2.6. Еще одна важная комбинаторная интерпретация чисел Каталана появилась впервые в работах Леонарда Эйлера (и, кстати сказать, задолго до работ самого Эжена Каталана). Эйлер рассмотрел количество разбиений выпуклого $(n+2)$-угольника с занумерованными (т.е. различимыми) вершинами на треугольники непересекающимися между собой диагоналями этого $(n+2)$-угольника.

`Упражнение 5.` Доказать, что это количество описывается числами Каталана C_n, установив биекцию между всеми триангуляциями выпуклого $(n+2)$-угольника и плоскими корневыми бинарными деревьями, построенными на n вершинах. Нарисовать эту биекцию для одной из триангуляций выпуклого шестиугольника.

2.7. Рассмотрим теперь все плоские корневые деревья, построенные на $(n+1)$-й вершине, $n = 0, 1, 2, \ldots$. Количество таких деревьев также равно числу Каталана C_n. Для того, чтобы это понять, осуществим в любом таком дереве поиск в глубину. При движении вдоль некоторого ребра вниз, т.е. от корня, сопоставим этому ребру левую открывающуюся скобку. При движении в обратном направлении сопоставим этому же ребру закрывающуюся скобку. Тем самым мы устанавливаем биекцию между всеми такими деревьями и всеми правильными скобочными последовательностями.

2.8. Наконец, перейдем к еще одному полезному понятию — стек-сортируемой последовательности. Согласно Кнутту, стек-сортируемая последовательность — это линейно упорядоченный набор из n чисел, который можно получить из линейно упорядоченной последовательности $(1, 2, \ldots, n)$ с помощью единственного стека.

В качестве примера рассмотрим последовательность чисел $(1, 2, 3, 4)$. Разместим ее справа от стека и будем сдвигать ее влево, помещая числа в стек и вынимая их из этого стека с помощью следующей последовательности действий:

– поместить число 1 в стек;

– поместить число 2 в стек;

– извлечь число 2 из стека;

– поместить число 3 в стек;

– поместить число 4 в стек;

– извлечь число 4 из стека;

– извлечь число 3 из стека;

– извлечь число 1 из стека.

В результате этих действий мы получим линейно упорядоченный набор чисел $(2, 4, 3, 1)$.

Возникает вопрос: сколько различных наборов чисел можно получить из последовательности $(1, 2, \ldots, n)$ с помощью подобных операций? Ответ, конечно же, вполне ожидаем — это количество равно числу Каталана C_n.

Для того, чтобы убедиться в этом, достаточно установить следующую биекцию между парой используемых в алгоритме операций и парой скобок:

$$(\longleftrightarrow \text{ поместить число в стек,} \qquad) \longleftrightarrow \text{ извлечь число из стека.}$$

3. Итак, теперь, после стольких примеров, пора научиться вычислять числа Каталана. Начнем с вывода рекуррентного соотношения для этих чисел.

3.1. В качестве основного объекта выберем множество всех правильных скобочных последовательностей. Рекуррентное соотношение для чисел C_n получим, воспользовавшись хорошо нам уже знакомым принципом разбиения множества на блоки и подсчетом количества элементов в каждом из блоков в отдельности.

3.2. Рассмотрим произвольную правильную скобочную последовательность. Для любой открывающейся скобки в такой последовательности можно ввести понятие парной ей закрывающейся скобки. Для этого будем идти от открывающейся скобки вправо и для каждой закрывающейся скобки будем проверять условие "количество закрывающихся скобок равно количеству открывающихся скобок". Первая закрывающаяся скобка, для которой это правило выполнится, и будет парной для нашей открывающейся скобки.

3.3. Разобъем теперь множество всех правильных скобочных последовательностей на блоки. Для этого возьмем крайнюю левую открывающуюся скобку в правильной скобочной последовательности и поместим в k-й блок все правильные скобочные последовательности, для которых парная ей закрывающаяся скобка стоит на $2k$-м месте.

Так, в случае $n = 3$ имеем разбиение множества, состоящего из пяти различных правильных скобочных последовательностей, на три блока:

$$\underbrace{()()(), \qquad ()(())}_{k=1}; \qquad \underbrace{(())()}_{k=2}; \qquad \underbrace{(())(), \qquad ((()))}_{k=3}.$$

3.4. Рассмотрим крайнюю левую открывающуюся и парную ей закрывающуюся скобки — выделенную пару скобок в нашей последовательности. Основное наблюдение здесь состоит в следующем: как внутри, так и снаружи указанной пары скобок стоят правильные скобочные последовательности.

Действительно, рассмотрим вначале последовательность скобок, находящуюся внутри выделенной пары скобок. Мы выбирали выделенную пару из того условия, что в подпоследовательности скобок, начинающейся с крайней левой открывающейся скобки и заканчивающейся парной ей закрывающейся скобкой, количество открывающихся скобок равно количеству закрывающихся скобок. Но, если мы крайнюю пару скобок удалим, то это условие для оставшейся скобочной подпоследовательности сохранится. Условие "количество открывающихся скобок больше или равно количеству закрывающихся скобок" в этой подпоследовательности также выполняется — в противном случае оно бы было нарушено и для всей последовательности скобок в целом.

Проверка того, что и правая подпоследовательность является правильной, проводится с помощью аналогичных рассуждений.

3.5. Подсчитаем теперь количество элементов в k-м блоке. Для этого заметим, что количество способов построить правильную скобочную последовательность внутри выделенной пары скобок равно, очевидно, C_{k-1}. Вне зависимости от выбора этой последовательности мы C_{n-k} способами можем построить правильную скобочную подпоследовательность справа от выделенной пары скобок. Следовательно, по правилу произведения в каждом блоке существует $C_{k-1}\cdot C_{n-k}$ способов построить правильную скобочную последовательность длины $2n$. Общее же число способов получить такую последовательность согласно правилу суммы равно

$$C_n=\sum_{k=1}^{n} C_{k-1}\,C_{n-k}, \qquad n=1,2,\ldots; \qquad C_0=1. \tag{24}$$

3.6. Заметим, что рекуррентное соотношение для чисел Каталана является нелинейным. Кроме того, n-й член последовательности C_n зависит от всех n предыдущих членов этой последовательности.

4. Постараемся решить рекуррентное соотношение (24).

4.1. Для этого введем обыкновенную производящую функцию для последовательности чисел Каталана:

$$f(z)=C_0+C_1z+C_2z^2+\ldots+C_nz^n+\ldots$$

Домножим (24) на z^n и просуммируем полученное равенство по n от единицы до $+\infty$:

$$\sum_{n=1}^{+\infty} C_nz^n=\sum_{n=1}^{+\infty} z^n\Bigl(\sum_{k=1}^{n} C_{k-1}\,C_{n-k}\Bigr) \qquad\Longleftrightarrow\qquad f(z)-1=z\Bigl(\sum_{n=1}^{+\infty} z^{n-1}\Bigl(\sum_{k=1}^{n} C_{k-1}\,C_{n-k}\Bigr)\Bigr).$$

Заметим теперь, что

$$\sum_{n=1}^{+\infty} z^{n-1}\Bigl(\sum_{k=1}^{n} C_{k-1}\,C_{n-k}\Bigr)=\bigl|_{m=n-1;\,n=m+1;\,m=0,1,\ldots}\bigr|=\sum_{m=0}^{+\infty} z^m\Bigl(\sum_{k=1}^{m+1} C_{k-1}\,C_{m+1-k}\Bigr)=$$

$$=\bigl|_{i=k-1;\,k=i+1;\,i=0,\ldots,m}\bigr|=\sum_{m=0}^{+\infty} z^m\Bigl(\sum_{i=0}^{m} C_i\,C_{m-i}\Bigr)=f(z)\cdot f(z)$$

по определению произведения обыкновенных производящих функций. Следовательно, функция $f(z)$ определяется из следующего равенства:

$$f(z)=1+z\,f^2(z). \tag{25}$$

4.2. Как видно, нелинейное рекуррентное соотношение (24) для чисел C_n приводит к нелинейному же уравнению (25) на производящую функцию $f(z)$. Как правило, решение такого рода уравнений строится с помощью формулы обращения Лагранжа, о которой подробно будет рассказано в следующей главе. Мы же сейчас вновь воспользуемся подробно описанным в первом параграфе данной главы подходом, основанным на связи формальных степенных рядов с функциональными рядами из математического анализа.

Именно, предположим, что имеется функция $f(z)$ комплексного аргумента z, аналитическая в окрестности точки $z = 0$ и удовлетворяющая уравнению (25). Тогда эта функция единственным образом раскладывается в окрестности точки $z = 0$ в степенной ряд

$$f(z) = C_0 + C_1 z + C_2 z^2 + \ldots$$

При этом, так как правила сложения и умножения таких рядов и формальных степенных рядов $f(z) \in \mathbb{C}[[z]]$ совпадают, то коэффициенты C_n, полученные в результате такого разложения, будут удовлетворять исходному рекуррентному соотношению (24).

Итак, все, что нам остается сделать — это найти функцию $f(z)$, аналитическую в окрестности точки $z = 0$ и удовлетворяющую уравнению (25), а затем разложить ее по степеням z^n в окрестности этой точки. Для этого разрешим уравнение (25) относительно функции $f(z)$:

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

Решение

$$f(z) = \frac{1 + \sqrt{1-4z}}{2z}$$

расходится в окрестности точки $z = 0$, поэтому его следует исключить из рассмотрения. Второе решение

$$f(z) = \frac{1 - \sqrt{1-4z}}{2z} = \frac{1}{2z} - \frac{1}{2z}\,(1-4z)^{1/2}$$

разложим в ряд в окрестности точки $z = 0$, используя формулу бинома Ньютона:

$$f(z) = \frac{1}{2z} - \frac{1}{2z}\left(1 + \sum_{n=1}^{+\infty} \frac{(1/2)(1/2-1)(1/2-2)\ldots(1/2-n+1)}{n!}(-4z)^n\right) =$$

$$= -\frac{1}{2z}\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{1\cdot 1\cdot 3\cdot 5\cdot\ldots\cdot(2n-3)}{2^n n!}(-4)^n z^n = \sum_{n=1}^{+\infty}\frac{2^{n-1}\cdot 1\cdot 3\cdot 5\cdot\ldots\cdot(2n-3)}{n!}z^{n-1} =$$

$$= \sum_{n=1}^{+\infty}\frac{(2n-2)!}{n\,((n-1)!)^2}\,z^{n-1} = \left|_{n'=n-1;\, n=n'+1;\, n'=0,1,2,\ldots}\right| = \sum_{n'=0}^{+\infty}\frac{(2n')!}{(n'+1)\,(n'!)^2}\,z^{n'}.$$

Следовательно, числа Каталана могут быть выражены через биномиальные коэффициенты по формуле

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \qquad n = 0, 1, 2, \ldots$$

6 Комбинаторный смысл сложения и умножения производящих функций

1. Пусть теперь X, Y — пара конечных или счетных множеств, и пусть каждому из этих множеств поставлены в соответствие обыкновенные ($f(z)$ и $g(z)$) или экспоненциальные ($F(z)$ и $G(z)$) производящие функции. Так как производящие функции являются элементами кольца формальных степенных рядов, то их можно складывать и перемножать между собой. Эти формальные операции имеют вполне определенный комбинаторный смысл, к изучению которого мы и перейдем.