Дискретная математика

1 Основные правила перечислительной комбинаторики

1.1. Начнем с очень краткого напоминания основных понятий теории множеств.

Определение 1.1. Множеством $X = \{x_1, x_2, \dots, x_n\}$ называется совокупность различимых объектов x_i, $i = 1, \dots, n$, объединенных по некоторому признаку.

В качестве характерного примера можно рассмотреть, например, множество X всех студентов, находящихся в данной аудитории. Действительно, все студенты различимы, отличны друг от друга и объединены по признаку "собрались в данной аудитории".

Определение 1.2. Мощностью $|X|$ множества X называется количество элементов в нем. Как правило, мы будем рассматривать конечные множества, в которых $|X| = n$, $n \in \mathbb{N}$, и называть их n-множествами.

1.1.1. Основные операции над множествами — это объединение $A \cup B$ (рис.1,a), пересечение $A \cap B$ (рис.1,b), разность $A \setminus B$ (рис.2,a) и симметрическая разность (рис.2,b) двух множеств A и B. В случае, если множество A является подмножеством некоторого более широкого множества X, удобно также рассматривать операцию дополнения $A' := X \setminus A$ множества A (рис.2,c).

Рис. 1

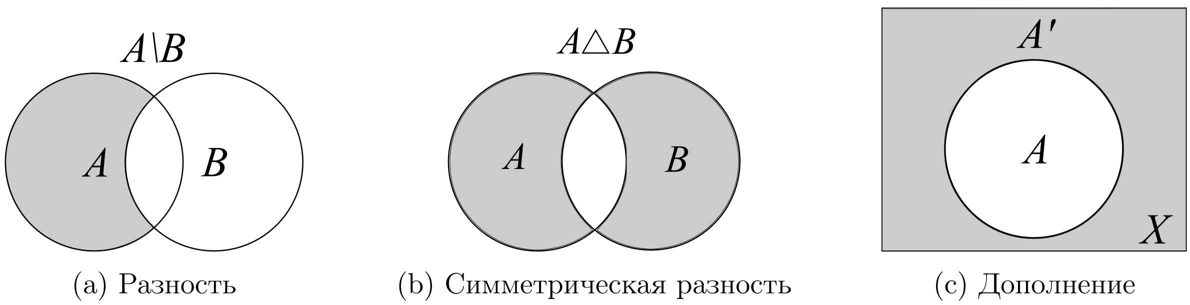

Рис. 2

Свойства операций над множествами удобно изучать графически, с использованием так называемых *диаграмм Эйлера-Венна* (смотри рисунки 1 и 2). Например, с их помощью достаточно просто проиллюстрировать справедливость законов де Моргана

$$A' \cap B' = (A \cup B)', \qquad A' \cup B' = (A \cap B)' \tag{1}$$

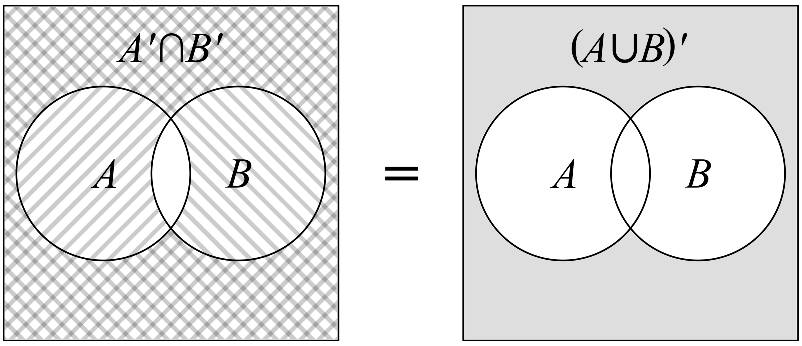

Рис. 3: Графическое доказательство закона де Моргана

(смотри рис.3).

1.1.2. В дальнейшем мы достаточно часто будем использовать понятие покрытия множества X семейством $\{X_1, X_2, \ldots, X_k\}$ множеств, а также всевозможные частные случаи этого понятия.

Определение 1.3. Семейство множеств $\{X_1, X_2, \ldots, X_k\}$ называется *покрытием* множества X, если их объединение дает нам все множество X:

$$X_1 \cup X_2 \cup \ldots \cup X_k = X.$$

Важным частным случаем покрытия является понятие разбиения множества.

Определение 1.4. Говорят, что семейство множеств $\{X_1, X_2, \ldots, X_k\}$ образует *разбиение* множества X, если

1. множества $X_i \neq \varnothing$, $i = 1, \ldots, n$;
2. $X_i \cap X_j = \varnothing \ \forall i \neq j$;
3. $X_1 \cup X_2 \cup \ldots \cup X_k = X$.

Элементы X_i этого семейства называются *блоками* разбиения.

В качестве характерного примера можно рассмотреть разбиение студентов данного курса на группы. Студенческие группы являются при этом блоками данного разбиения.

Если по каким-то причинам оказывается важным порядок следования блоков, то говорят об *упорядоченном разбиении* $(X_1, X_2, \ldots, X_k)$ множества X. Например, если мы выводим группы на сцену для вручения им дипломов, то важен порядок, в котором они туда выходят. Следовательно, в данном случае мы получаем упорядоченное разбиение студентов данного курса.

Наконец, еще одним частным случаем покрытия X семейством множеств $\{X_1, X_2, \ldots, X_k\}$ является понятие *разделения* множества X. Разделение есть аналог упорядоченного разбиения, в котором допускаются пустые блоки. Точное определение таково:

Определение 1.5. Разделением множества X называется упорядоченная последовательность $(X_1, X_2, \ldots, X_k)$ возможно пустых, попарно непересекающихся множеств, объединение которых дает все множество X.

1.1.3. Еще одной часто использующейся в комбинаторике операцией над множествами является операция декартова произведения множеств.

Определение 1.6. Декартовым произведением множеств A и B называется множество

$$A \times B := \{(a, b) \,|\, a \in A,\ b \in B\}$$

всех упорядоченных пар (a, b), таких, что $a \in A$, $b \in B$.

В качестве простейшего примера декартова произведения множеств обычно приводят шахматную доску. Любая клетка шахматной доски имеет координаты "буква-цифра", например, $e5$ или $h4$. Иными словами, координаты клеток шахматной доски являются элементами декартова произведения множеств $A = \{a, b, \ldots, h\}$ и $B = \{1, 2, \ldots, 8\}$.

В частном случае множества A и B могут совпадать. В этом случае декартово произведение $A \times A$ обозначается через $A^{(2)}$.

Определение 1.7. Декартовым произведением k множеств $X_1, X_2, \ldots, X_k$ называется множество

$$X_1 \times X_2 \times \ldots \times X_k := \{(x_1, x_2, \ldots, x_k) \,|\, x_i \in X_i,\ \forall\ i = 1, \ldots, k\}$$

всевозможных упорядоченных k-элементных последовательностей вида $(x_1, x_2, \ldots, x_k)$.

В частном случае $X_1 = X_2 = \ldots = X_k = X$ имеем декартову степень $X \times X \times \ldots \times X =: X^{(k)}$.

1.1.4. Любой элемент $X^{(k)}$ есть упорядоченный набор из k элементов множества X, в котором некоторые элементы могут повторяться. Если же в таком k-множестве порядок элементов не важен, говорят о *k-мультимножестве над множеством X*. Формальное определение k-мультимножества таково.

Определение 1.8. k-мультимножеством над n-элементным множеством X называется пара (X, φ), где $\varphi\colon X \to \mathbb{Z}_+$ есть функция, сопоставляющая любому элементу $x \in X$ количество $\varphi(x)$ его вхождений в k-мультимножество.

Любую функцию φ такого рода можно определить с помощью множества Ξ упорядоченных пар

$$\Xi := \{(x, \varphi(x)) \mid x \in X\}.$$

Поэтому, например, 3-мультимножество над множеством $X = \{x, y\}$, состоящее из двух элементов x и одного элемента y, можно формально записывать в виде $(X, \{(x, 2), (y, 1)\})$. Однако чаще вместо такой формальной записи используют более наглядную форму записи вида $\{x, x, y\}$.

Самый простой и понятный пример мультимножества — это монеты в кошельке. В этом примере в качестве множества X выступает множество из девяти монет разного достоинства:

$$X = \{1 \text{ копейка},\ 5 \text{ копеек},\ 10 \text{ копеек},\ 50 \text{ копеек},\ 1 \text{ рубль},\ 2 \text{ рубля},\ 5 \text{ рублей},\ 10 \text{ рублей}\}.$$

Любой набор из этих монет в количестве k штук образует k-мультимножество над множеством X.

1.1.5. Теория множеств как раздел математики создавалась значительно позже комбинаторики. Поэтому некоторые наиболее важные понятия теории множеств исторически получили в комбинаторике свои, специфические названия. Именно,

1. k-сочетанием без повторений называется любое k-элементное подмножество n-элементного множества;

2. k-сочетанием с повторениями называется любое k-мультимножество над n-множеством;

3. k-перестановкой без повторений называется упорядоченное k-подмножество n-элементного множества;

4. k-перестановкой с повторениями называется любой элемент декартовой степени $X^{(k)}$.

1.2. Теперь перейдем к двум самым простым, но в то же время достаточно важным правилам перечислительной комбинаторики — правилу суммы и правилу произведения.

1.2.1. Начнем с простейшего примера: пусть на одном блюде лежат три яблока, а на втором — две груши; сколькими способами можно выбрать один фрукт? Ответ очевиден: пятью способами.

Обобщающее этот пример простейшее *правило суммы* можно сформулировать так: если некоторый объект из множества A можно выбрать k способами, и, вне зависимости от выбора этого объекта, можно n способами выбрать некоторый элемент множества B, то выбор объекта из множества A *или* из множества B можно осуществить $k+n$ способами.

Очевидна переформулировка этого правила на языке теории множеств: пусть пересечение двух множеств A и B пусто; тогда

$$|A \cup B| = |A| + |B|.$$

В частности, если $A \subset X$ и A' — дополнение множества A, то

$$|A| + |A'| = |X|. \tag{2}$$

В более общем случае, рассматривая произвольное разбиение множества X на блоки, имеем равенство вида

$$|X| = |X_1| + |X_2| + \ldots + |X_k|,$$

которое также называется правилом суммы в комбинаторике.

1.2.2. Под правилом произведения в комбинаторике понимается равенство

$$|X_1 \times X_2 \times \ldots \times X_k| = |X_1| \cdot |X_2| \cdot \ldots \cdot |X_k|.$$

Приведем простейший пример на применение этого правила в комбинаторике. Пусть в аудитории находятся 32 студента одной группы, 24 студента другой группы и 17 студентов третьей группы. В этом случае тройку, состоящую из представителей каждой группы, можно выбрать $32 \cdot 24 \cdot 17$ способами.

1.3. Наряду с правилом суммы, в элементарной комбинаторике также достаточно часто используется и несложное его обобщение — так называемый *принцип включения-исключения*. Если правило суммы связано с разбиением множества X, то принцип включения-исключения связан с некоторым произвольным покрытием этого n-множества семейством $\{X_1, X_2, \ldots, X_k\}$. Сформулируем его для самого простого случая двух множеств.

1.3.1. Рассмотрим два конечных множества A и B, пересечение которых может быть и непусто. Тогда количество элементов в объединении этих множеств, очевидно, равно

$$|A \cup B| = |A| + |B| - |A \cap B|. \tag{3}$$

Действительно, когда мы считаем количество $|A|$ элементов в множестве A и складываем его с количеством $|B|$ элементов в множестве B, мы любой элемент, принадлежащий как множеству A, так и множеству B, считаем дважды. Чтобы этот избыток убрать, нам нужно один раз вычесть количество элементов, содержащихся в пересечении этих двух множеств.

Равенство (3) можно называть обобщенным правилом суммы — оно обобщает правило суммы на случай, когда пересечение двух множеств не пусто.

1.3.2. Предположим теперь, что A и B являются подмножествами некоторого более широкого множества X. В этом случае у множества $A \subset X$ и множества $B \subset X$ имеются дополнения к ним — множества A' и B', причем $A \cup A' = B \cup B' = X$.

Рассмотрим теперь пересечение $A' \cap B'$ дополнений множеств A и B. Согласно одной из теорем де Моргана (1), $A' \cap B' = (A \cup B)'$. Следовательно, количество элементов в этом пересечении с учетом равенства (2) и обобщенного правила суммы (3) можно сосчитать так:

$$|A' \cap B'| = |(A \cup B)'| = |X| - |A \cup B| = |X| - |A| - |B| + |A \cap B|.$$

Равенство

$$|A' \cap B'| = |X| - |A| - |B| + |A \cap B|, \tag{4}$$

и называется в комбинаторике принципом включения-исключения.

1.3.3. Приведем простейший пример его использования. Пусть в аудитории находятся 30 человек, 20 человек из которых знают английский, 12 — французский, а 6 человек знают оба языка. Сколько человек не знает ни один из этих иностранных языков? Ответ, согласно принципу включения-исключения (4), следующий:

$$N = 30 - 20 - 12 + 6 = 4.$$

1.3.4. Несложно обобщить равенства (3) и (4) на случай большего количества множеств. Например, для трех множеств A, B, C соответствующие формулы выглядят так:

а) обобщенное правило суммы:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|; \tag{5}$$

б) принцип включения-исключения:

$$|A' \cap B' \cap C'| = |X| - |A| - |B| - |C| + |A \cap B| + |A \cap C| + |B \cap C| - |A \cap B \cap C|. \tag{6}$$

Действительно, рассмотрим, к примеру, левую часть равенства (5). Она подсчитывает количество элементов, принадлежащих объединению трех множеств (смотри рис.4). Если элемент x_1, принадлежащий этому объединению, содержится в множестве A, но не содержится в множествах B и C, то он один раз подсчитывается в правой части равенства (5) (слагаемое $|A|$, отвечающее зеленой подобласти на рис.4). Если элемент x_2 принадлежит множествам A и B, но не принадлежит C (красная подобласть на рис.4), то в правой части (5) этот элемент входит в слагаемые $|A|$, $|B|$ и $-|A \cap B|$, то есть также подсчитывается ровно один раз. Наконец, если x_3 принадлежит пересечению всех трех множеств (зеленая подобласть на рис.4), то за этот элемент отвечают все слагаемые в правой части (5). Так как четыре из них входят со знаком плюс, а три — со знаком минус, то этот элемент также считается в правой части (5) лишь однажды.

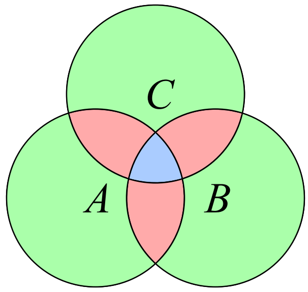

Рис. 4: Диаграмма Эйлера-Венна для трёх множеств

2 Подсчет k-сочетаний из n элементов. Биномиальные коэффициенты.

2.1. Основная задача данного параграфа состоит в подсчете количества всех k-сочетаний из n элементов. Начнем мы с подсчета количества k-сочетаний из n элементов без повторений.

2.1.1. Число k-сочетаний без повторений известно в литературе под названием биномиальных коэффициентов. Ранее в советской литературе такие числа обозначались через C_n^k. В настоящее время для этих коэффициентов используется обозначение $\binom{n}{k}$ (читается "из n по k").

2.1.2. Обычно на вопрос, чему равны биномиальные коэффициенты, вспоминают формулу

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

Эта формула не очень удачна с двух точек зрения — с вычислительной и с идейной. С вычислительной точки зрения ее затруднительно использовать для достаточно больших значений n и k. Например, при $n = 38$ и $k = 19$ числитель и знаменатель могут просто не уместиться в диапазон изменения целых чисел для того или иного языка программирования. С идейной же точки зрения эта формула неудобна потому, что она не позволяет обобщить понятие биномиальных коэффициентов на случай целых, вещественных или комплексных значений n. В следующем параграфе мы получим более удобное выражение для этих коэффициентов, допускающее такое обобщение. Здесь же мы с помощью правила суммы выведем рекуррентное соотношение для биномиальных коэффициентов, позволяющее эти биномиальные коэффициенты эффективно вычислять для достаточно больших значений n и k.

2.1.3. Для получения рекуррентного соотношения введем множество Σ_k всех k-элементных подмножеств n-элементного множества X. Например, для $X = \{x_1, x_2, x_3\}$ множество $\Sigma_2 = \{\{x_1, x_2\}, \{x_1, x_3\}, \{x_2, x_3\}\}$. Заметим, что биномиальный коэффициент $\binom{n}{k}$ как раз и описывает мощность такого множества. Разобьем теперь множество Σ_k на два блока — блок $\Sigma_k^{(1)}$, k-элементные подмножества которого содержат элемент x_1, и блок $\Sigma_k^{(2)}$, подмножества которого этот элемент не содержат. Понятно, что это — непустые, непересекающиеся подмножества, объединение которых дает нам все множество Σ_k. Поэтому по правилу суммы мы получаем равенство вида

$$\binom{n}{k} = |\Sigma_k| = |\Sigma_k^{(1)}| + |\Sigma_k^{(2)}|.$$

Осталось выразить через биномиальные коэффициенты количество элементов в каждом из блоков $\Sigma_k^{(1)}$, $\Sigma_k^{(2)}$. А это делается довольно легко.

2.1.4. Действительно, во всех подмножествах первого блока элемент x_1 уже выбран, и нам из оставшегося $(n-1)$-элементного множества $X \setminus x_1$ остается выбрать $(k-1)$-элементные подмножества . Сделать это можно $\binom{n-1}{k-1}$ способами. Во втором блоке содержатся k-элементные подмножества множества $X \setminus x_1$, состоящего из $(n-1)$-го элемента. Их количество, очевидно, равно $\binom{n-1}{k}$. Таким образом окончательно имеем следующее рекуррентное соотношение:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \qquad k \geqslant 1, \quad n \geqslant 1. \tag{7}$$

2.1.5. Соотношение (7) следует дополнить начальными и граничными условиями. Так как k-элементных подмножеств n-элементного множества в случае $k > n$ не существует, то

$$\binom{n}{k} = 0 \qquad \text{при} \quad k > n.$$

Далее, пустое подмножество можно выбрать всегда и только одним способом; поэтому

$$\binom{n}{0} = 1 \qquad \forall\, n \geqslant 0.$$

Используя эти условия, можно шаг за шагом вычислить коэффициенты $\binom{n}{k}$. Часто их записывают в виде так называемого треугольника Паскаля:

$$\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
 & & & & \dots & & & &
\end{array}$$

2.1.6. Как видно, треугольник Паскаля симметричен, т.е. $\binom{n}{k} = \binom{n}{n-k}$. Комбинаторное доказательство этого факта очевидно. Действительно, выбирая любое k-элементное множество, мы тем самым однозначно выбираем и дополнение к нему, т.е. $(n-k)$-элементное множество. Следовательно, количество k-элементных и $(n-k)$-элементных подмножеств совпадает.

2.2. Наряду с треугольником Паскаля биномиальные коэффициенты допускают и еще одно очень удобное графическое представление — представление на координатной плоскости (n, k)

2.2.1. Данное представление связано со следующей довольно интересной комбинаторной задачей. Рассмотрим плоскость (n, k) и нарисуем на этой плоскости пути, исходящие из начала координат, приходящие в точку с координатами (n, k), $n \geqslant 0$, $k = 0, \dots, n$, и состоящие из диагональных и вертикальных отрезков (рис.5). В самих точках (n, k) отметим количество таких путей, приходящих в эту точку из начала координат. Заметим теперь, что попасть в точку с координатами (n, k) мы можем, пройдя только лишь через точку с координатами $(n-1, k)$ или через точку с координатами $(n-1, k-1)$. С комбинаторной точки зрения это означает, что количество путей $\binom{n}{k}$, приходящих в точку с координатами (n, k), равняется количеству $\binom{n-1}{k}$ путей, приходящих в точку с координатами $(n-1, k)$, плюс количество путей $\binom{n-1}{k-1}$, приходящих в точку с координатами $(n-1, k-1)$. Но равенство

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

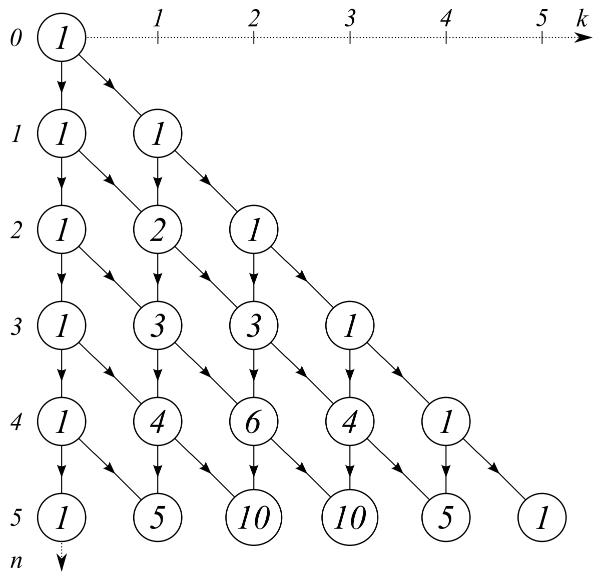

Рис. 5: Графическое представление чисел $\binom{n}{k}$ на координатной плоскости (n, k)

есть основное тождество для биномиальных коэффициентов (7). А это означает, что мы получили новую комбинаторную интерпретацию таких коэффициентов. Именно, числа $\binom{n}{k}$ описывают количество путей, состоящих из диагональных $(1, 1)$ и вертикальных $(1, 0)$ отрезков, выходящих из начала координат — точки $(0, 0)$, и оканчивающихся в точке с координатами (n, k).

2.2.2. Графическое представление чисел $\binom{n}{k}$ на плоскости (n, k) очень удобно для получения разного рода тождеств с биномиальными коэффициентами. В качестве первого примера зафиксируем какое-то конкретное значение параметра k и просуммируем биномиальные коэффициенты $\binom{m}{k}$ по m от k до некоторого фиксированного значения n. Например, выберем $k = 1$, $n = 4$. Складывая числа $1, 2, 3, 4$, мы получим число 10, то есть биномиальный коэффициент, стоящий правее и ниже рассмотренной цепочки биномиальных коэффициентов. Есть подозрение, что данный факт, а именно, равенство вида

$$\binom{k}{k} + \binom{k+1}{k} + \ldots + \binom{n}{k} = \binom{n+1}{k+1} \tag{8}$$

выполняется для любых значений параметров n и k.

2.2.3. Для формального доказательства тождества

$$\sum_{m=0}^{n} \binom{m}{k} = \underbrace{\binom{0}{k} + \binom{1}{k} + \ldots + \binom{k-1}{k}}_{=0} + \binom{k}{k} + \binom{k+1}{k} + \ldots + \binom{n}{k} = \sum_{m=k}^{n} \binom{m}{k} = \binom{n+1}{k+1},$$

называемого формулой суммирования биномиальных коэффициентов по верхнему индексу, применим рекуррентное соотношение (7) к коэффициенту $\binom{m+1}{k+1}$:

$$\binom{m+1}{k+1} = \binom{m}{k} + \binom{m}{k+1} \qquad \Longrightarrow \qquad \binom{m}{k} = \binom{m+1}{k+1} - \binom{m}{k+1}.$$

Просуммируем теперь полученное равенство по m от k до n:

$$\sum_{m=k}^{n}\binom{m}{k}=\binom{n+1}{k+1}+\sum_{m=k}^{n-1}\binom{m+1}{k+1}-\sum_{m=k}^{n}\binom{m}{k+1}=$$

$$=\binom{n+1}{k+1}+\sum_{m=k}^{n-1}\binom{m+1}{k+1}-\sum_{m=k+1}^{n}\binom{m}{k+1}=$$

$$=\binom{n+1}{k+1}+\sum_{m'=k+1}^{n}\binom{m'}{k+1}-\sum_{m=k+1}^{n}\binom{m}{k+1}=\binom{n+1}{k+1}.$$

2.2.4. Комбинаторное доказательство тождества (8) основано на следующем общем подходе: мы разбиваем множество Σ_{k+1} всех $(k+1)$-элементных подмножеств $(n+1)$-элементного множества X на блоки, подсчитываем количество элементов в каждом блоке, а затем пользуемся правилом суммы для подсчета числа $|\Sigma_{k+1}|=\binom{n+1}{k+1}$.

Для реализации этого подхода возьмем $(n+1)$-элементное множество

$$X=\{x_1,x_2,\ldots,x_{k-1},x_k,x_{k+1},\ldots,x_{n-1},x_n,x_{n+1}\}$$

и будем разбивать множество всех $(k+1)$-элементных подмножеств множества X следующим образом. В первый блок поместим все $(k+1)$-элементные подмножества, содержащие элемент x_{n+1}. Такие подмножества элемент x_{n+1} гарантированно содержат. Нам из оставшихся элементов $x_1,\ldots,x_n$ множества $X\setminus x_{n+1}$ нужно выбрать недостающие k элементов. По определению биномиального коэффициента, сделать мы это можем $\binom{n}{k}$ способами.

Теперь мы рассмотрим все $(k+1)$-элементные подмножества, которые не содержат элемент x_{n+1}, но обязательно содержат элемент x_n. Для того, чтобы любое такое подмножество сформировать, нам нужно из множества элементов $\{x_1,\ldots,x_{n-1}\}$ выбрать недостающие k элементов. Это можно сделать $\binom{n-1}{k}$ количеством способов.

Затем рассмотрим все $(k+1)$-элементные подмножества, которые содержат в обязательном порядке элемент x_{n-1} и не содержат элементов x_{n+1} и x_n. Эти подмножества получаются выбором недостающих k элементов из множества элементов $\{x_1,\ldots,x_{n-2}\}$, и выбрать такие подмножества можно $\binom{n-2}{k}$ способами.

Продолжая далее, мы дойдем когда-то до ситуации, в которой нам нужно выбрать все $(k+$ $+\,1)$-элементные подмножества, которые содержат элемент x_{k+1} и не содержат элементы со старшими индексами. Так как элемент x_{k+1} у нас уже выбран, то нам остается из множества $\{x_1,\ldots,x_k\}$ выбрать k-элементное подмножество. Сделать мы это можем, очевидно, $\binom{k}{k}=1$ способом.

Складывая теперь количество элементов в каждом блоке, мы и получаем тождество (8).

Пример 2.1. Пусть $n=4$, $X=\{x_1,x_2,x_3,x_4,x_5\}$, $k=2$, $k+1=3$. Приведем список всех трехэлементных подмножеств этого множества:

$$\{x_1,x_2,x_3\},\quad \{x_1,x_2,x_4\},\quad \{x_1,x_2,x_5\},\quad \{x_1,x_3,x_4\},\quad \{x_1,x_3,x_5\},$$
$$\{x_1,x_4,x_5\},\quad \{x_2,x_3,x_4\},\quad \{x_2,x_3,x_5\},\quad \{x_2,x_4,x_5\},\quad \{x_3,x_4,x_5\}.$$

В первый блок разбиения этого множества подмножеств включим подмножества, содержащие элемент x_5; таковых имеется $\binom{4}{2}=6$ штук. Из *оставшегося* списка выберем все подмножества, содержащие x_4; их $\binom{3}{2}=3$ штуки. Наконец, у нас остается единственное подмножество элементов, не содержащих ни x_4, ни x_5, т.е. подмножество $\{x_1,x_2,x_3,\}$.

2.3. Название “биномиальные коэффициенты” связано с тем, что они, помимо всего прочего, встречаются в формуле бинома Ньютона

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k\, y^{n-k}. \tag{9}$$

2.3.1. Комбинаторное доказательство этой формулы довольно элементарно: нужно просто расписать $(x+y)^n$ в виде произведения n сомножителей

$$(x+y)^n = \underbrace{(x+y)}_{1} \cdot \underbrace{(x+y)}_{2} \cdot \ldots \cdot \underbrace{(x+y)}_{n}$$

и пометить каждый из таких сомножителей числом в диапазоне от единицы до n. В результате мы имеем множество X, состоящее из n *различных* экземпляров сомножителей вида $(x+y)$.

После перемножения этих n скобок получается определенный набор слагаемых вида $x^k y^{n-k}$, $k = 0, 1, \ldots, n$. Для подсчета количества этих слагаемых при фиксированном значении параметра k заметим, что любое слагаемое $x^k y^{n-k}$ можно получить так: выбрать в n-элементном множестве X k-элементное подмножество, взять в этом подмножестве в качестве сомножителей переменные x, а в оставшемся $(n-k)$-элементном подмножестве выбрать в качестве сомножителей переменные y. Как следствие, количество слагаемых $x^k y^{n-k}$ совпадает с количеством способов выбрать k-элементное подмножество n-элементного множества X и равно $\binom{n}{k}$.

2.3.2. Формула (9) оказывается чрезвычайно полезной для вывода разного рода соотношений, связанных с биномиальными коэффициентами. Например, полагая в ней $x = y = 1$, получаем тождество

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n. \tag{10}$$

Заметим теперь, что, суммируя биномиальные коэффициенты $\binom{n}{k}$ по k, мы подсчитываем *все* подмножества n-элементного множества. Иными словами, мы формально доказали тот факт, что количество всех подмножеств данного n-множества равно 2^n. Комбинаторное доказательство этого факта будет дано в следующем параграфе.

2.3.3. Далее, полагая в (9) $x = -1$, $y = 1$, получаем важное тождество

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0, \qquad n > 0. \tag{11}$$

Заметим, что в случае $n = 0$ эта сумма оказывается равной единице. Для того, чтобы понять комбинаторный смысл равенства (11), перепишем его в следующем виде:

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \ldots = 0 \qquad \Longleftrightarrow \qquad \binom{n}{0} + \binom{n}{2} + \ldots = \binom{n}{1} + \binom{n}{3} + \ldots$$

В левой части последнего равенства стоит количество всех подмножеств, в которых содержится четное количество элементов, а в правой — число подмножеств, содержащих нечетное количество элементов. Иными словами, следствием равенства (11) является тот факт, что количество четных подмножеств любого множества, отличного от $\varnothing$, равняется количеству его нечетных подмножеств.

Наконец, отсюда же мы можем заключить, что количество, например, всех нечетных подмножеств ровно в два раза меньше общего количества всех подмножеств. С учетом (10) это означает, что количество всех нечетных подмножеств равно 2^{n-1}.

2.3.4. Наконец, продифференцируем (9) по x:

$$n\,(x+y)^{n-1} = \sum_{k=1}^{n} k\binom{n}{k}\,x^{k-1}\,y^{n-k}. \tag{12}$$

Подставляя в это равенство $x = y = 1$, получим еще одно полезное равенство:

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

2.4. Перейдем теперь к задачам, связанным с подсчетом k-сочетаний с повторениями.

2.4.1. Начнем с примера. Пусть множество X состоит из двух чисел 1 и 2. Выпишем все 3-сочетания с повторениями из 2-множества X:

$$\{1,1,1\}, \qquad \{1,1,2\}, \qquad \{1,2,2\}, \qquad \{2,2,2\}.$$

Как видно, таковых оказалось 4 штуки. Как подсчитать это количество в общем случае?

2.4.2. Для решения данной задачи воспользуемся чрезвычайно полезным и часто использующимся в комбинаторике *принципом биекции*. Формально этот принцип можно сформулировать следующим образом: пусть X, Y — пара конечных множеств, и пусть существует биекция $f\colon X \to Y$, т.е. такое отображение, что

$$\forall\, y \in Y \qquad \exists!\, x \in X : \qquad y = f(x).$$

Тогда количество элементов в множествах X и Y совпадают: $|X| = |Y| = n$.

2.4.3. Неформально использование принципа биекции можно проиллюстрировать на следующем примере. Предположим, что вы устраиваете вечеринку и приглашаете на нее довольно много друзей. Как гостеприимный хозяин, вы встречаете всех своих друзей на входе в дом, но запоминаете только пришедших к вам девушек. В какой-то момент вы решаете подсчитать, сколько парней пришло к вам на вечеринку. Вы знаете количество пришедших к вам девушек, и вам кажется, что количество девушек и парней одинаково. Как вам быстро проверить это предположение? Ответ достаточно очевиден: попросить каждую девушку взять ровно одного парня за руку. Если в результате этой процедуры все множество гостей разбилось на пары, то ваше предположение окажется верным. Тем самым вы сильно упростили себе жизнь — вам не пришлось проделывать довольно утомительную работу по пересчету пришедших к вам парней; вы просто воспользовались для их подсчета результатом уже проделанной работы по пересчету пришедших к вам девушек.

2.4.4. Вернемся теперь к задаче подсчета всех k-сочетаний с повторениями, т.е. всех k-мультимножеств над n-множеством X. Для подсчета количества $\left(\!\binom{n}{k}\!\right)$ таких мультимножеств нам будет удобнее вначале конкретизировать n-множество X, а именно, взять в качестве X множество $[n] := \{1, 2, \ldots, n\}$ первых n натуральных чисел. Любое k-мультимножество такого множества можно записать, очевидно, в следующем виде:

$$1 \leqslant a_1 \leqslant a_2 \leqslant \ldots \leqslant a_k \leqslant n.$$

Например, 3-мультимножество $\{1,1,2\}$ над 2-множеством $X=\{1,2\}$ можно записать так:

$$1 \leqslant (a_1=1) \leqslant (a_2=1) \leqslant (a_3=2) \leqslant (n=2).$$

Теперь превратим в этой цепочке все нестрогие неравенства в строгие. Для этого мы к a_2 прибавим единицу, к a_3 — двойку, к a_4 — тройку, и так далее. К последнему числу a_n мы, таким образом, добавим число $(k-1)$. В результате получим цепочку строгих неравенств вида

$$1 \leqslant a_1 < a_2+1 < a_3+2 < a_4+3 < \ldots < a_k+(k-1) \leqslant n+(k-1).$$

В нашем примере

$$1 \leqslant (a_1=1) < (a_2+1=2) < (a_3+2=4) \leqslant (n+(3-1)=4).$$

Заметим, что в результате этой операции мы получили некоторое k-элементное подмножество *различных* чисел вида $a_i+(i-1)$ множества $\widetilde{X}=[n+k-1]$ всех чисел от единицы до $n+k-1$. Иными словами, мы сопоставили любому k-мультимножеству над множеством $X=[n]$ вполне определенное k-подмножество множества $\widetilde{X}=[n+k-1]$. Очевидно, что это сопоставление взаимно-однозначно. Но количество всех k-подмножеств данного множества мы знаем — оно равно $\binom{n+k-1}{k}$. Следовательно, этому числу равно, по принципу биекции, и количество $\left(\!\binom{n}{k}\!\right)$ всех k-мультимножеств над множеством $X=[n]$:

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

Справедливость этого равенства для произвольного n-элементного множества X вновь следует из принципа биекции.

3 k-перестановки из n элементов. Урновые схемы и схемы раскладки предметов по ящикам.

3.1. Перейдем теперь к подсчету количества k-перестановок из n элементов.

3.1.1. Напомним, что k-перестановкой из n элементов называется *упорядоченный* набор

$$(a_1, a_2, \ldots, a_k)$$

элементов, в котором все a_i принадлежат одному и тому же n-элементному множеству X.

Элементы a_i в наборе $(a_1, a_2, \ldots, a_k)$ могут как повторяться, так и не повторяться. В первом случае говорят о k-перестановках с повторениями, во втором — о k-перестановках без повторений.

Номер паспорта — это типичный пример k-перестановки с повторениями над множеством из десяти цифр $X=\{0,1,\ldots,9\}$. Классическим примером 3-перестановки без повторений является упорядоченный список спортсменов, занявших призовые места в любых спортивных соревнованиях.

3.1.2. В литературе встречается довольно много альтернативных названий для данного объекта. Именно, упорядоченный набор $(a_1, a_2, \ldots, a_k)$, $a_i \in X$, также иногда называется

– k-размещением из n элементов;

– кортежем из k элементов множества X;

– упорядоченной k-выборкой из n элементов;

– k-мерным вектором над множеством X;

– k-элементным словом над n-элементным алфавитом.

3.1.3. Сосчитаем количество k-перестановок с повторениями.

Утверждение 3.1. *Количество k-перестановок с повторениями из n элементов равно n^k.*

Для доказательства можно либо просто сослаться на правило произведения, либо рассмотреть $(a_1, a_2, \ldots, a_k)$, $a_i \in X$ как слово из k элементов над алфавитом из $n = |X|$ букв. На первое место в слове мы можем поставить любую из n букв, на второе — также любую из n букв и так далее. Всего же получаем n^k вариантов записать данное слово.

3.1.4. В качестве важного приложения доказанного выше результата сосчитаем еще раз, на этот раз комбинаторно, количество подмножеств данного множества X. Для этого воспользуемся принципом биекции. Именно, закодируем любое подмножество A множества X бинарной строкой $f(A)$ длины n, то есть строкой над алфавитом $\{0, 1\}$. Единицу на i-м месте поставим в случае, если элемент $x_i \in A$. В противном случае на i-е место поставим ноль.

Например, пусть $X = \{x_1, x_2, x_3, x_4\}$, $A = \{x_2, x_4\}$. Тогда соответствующая подмножеству A строка длины 4 записывается следующим образом:

$$f(A) = (0, 1, 0, 1).$$

Очевидно, что построенное отображение f взаимно-однозначно. Следовательно, количество подмножеств данного n-множества X совпадает с количеством бинарных строк длины n, которое, согласно доказанному выше утверждению 3.1, равно 2^n.

3.1.5. Перейдем теперь к подсчету количества перестановок без повторений.

Утверждение 3.2. *Количество $P(n, k)$ k-перестановок из n элементов без повторений равно*

$$P(n, k) = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) =: (n)_k.$$

Доказательство очевидно — на первое место в строке длины k я могу поставить любой из n элементов, на второе — любой из оставшихся $(n-1)$ элементов и так далее.

3.1.6. В частном случае $k = n$ k-перестановки из n элементов без повторений называются просто перестановками n-элементного множества X. Их количество равно

$$P_n \equiv P(n) = n!, \qquad P(0) = 0! = 1.$$

3.1.7. Любую k-перестановку из n элементов без повторений можно рассматривать и как упорядоченное k-подмножество n-множества. Количество таких подмножеств мы можем сосчитать следующим образом: мы можем $\binom{n}{k}$ способами выбрать k-подмножество n-элементного множества, а затем $k!$ способами его упорядочить. Действительно, на первое место мы можем поставить любой из k элементов подмножества, на второе — любой из оставшихся $k-1$ элементов и так далее. Таким образом, мы получаем некоторое новое выражение для количества всех упорядоченных k-элементных подмножеств n-множества, то есть количества k-перестановок из n элементов без повторений, равно

$$(n)_k = k! \cdot \binom{n}{k} \qquad \Longrightarrow \qquad \binom{n}{k} = \frac{(n)_k}{k!}.$$

Последняя формула часто используется как некомбинаторное определение биномиальных коэффициентов $\binom{n}{k}$ в случае, когда $k \in \mathbb{Z}$, а n принадлежит $\mathbb{Z}$, $\mathbb{R}$ или даже $\mathbb{C}$. Именно, по определению,

$$\binom{q}{k} := \begin{cases} \dfrac{q\,(q-1)\,\ldots\,(q-k+1)}{k!} =: \dfrac{(q)_k}{k!}, & \text{если } k \in \mathbb{N}, \\ 1, & \text{если } k = 0, \\ 0, & \text{если } k < 0, \end{cases} \qquad \forall\, q \in \mathbb{C}.$$

Например,

$$\binom{-1}{3} = \frac{(-1)\cdot(-2)\cdot(-3)}{3\cdot 2\cdot 1} = -1.$$

3.1.8. Функцию $(q)_k$ часто также обозначают через $q^{\underline{k}}$ и называют *убывающей факториальной степенью* [1]. Наряду с убывающей можно ввести и так называемую *возрастающую факториальную степень*

$$q^{(k)} \equiv q^{\overline{k}} := q \cdot (q+1) \cdot \ldots \cdot (q+k-1).$$

В частности, с ее помощью получается удобное для вычислений выражение для количества $\left(\!\binom{n}{k}\!\right)$ сочетаний с повторениями:

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k} = \frac{n^{(k)}}{k!}. \tag{13}$$

3.2. Итак, мы получили простые соотношения для подсчета количества четырех основных объектов элементарной комбинаторики — k-сочетаний и k-перестановок из n элементов с повторениями и без повторений. Эти объекты встречаются в огромном количестве внешне не очень похожих друг на друга задач элементарной комбинаторики. Оказывается, однако, что большинство этих задач можно свести к одной из двух простейших схем — либо к так называемой урновой схеме, либо к схеме раскладки предметов по ящикам.

3.2.1. В урновой схеме имеется урна, в которой находятся n *различимых* предметов. Из урны последовательно вытаскивается k предметов. Задача состоит в подсчете количества различных способов выбора этих предметов, или, как еще говорят, в подсчете различных k-элементных выборок из n предметов, находящихся в урне.

На практике наиболее часто встречаются четыре модификации этой задачи, различающиеся способами формирования k-элементной выборки. Прежде всего, мы можем возвращать или не возвращать вытаскиваемые предметы обратно в урну. В первом случае говорят о *выборке с повторениями*, во втором — о *выборке без повторений*. Далее, в некоторых задачах нам важен порядок вытаскиваемых предметов. В этом случае имеем так называемые *упорядоченные* выборки. В противном случае выборки называются *неупорядоченными*.

Нетрудно понять, что задачи о подсчете k-элементных выборок представляют собой, по сути, те же самые задачи о подсчете k-перестановок или k-сочетаний из n элементов. Действительно, любая неупорядоченная k-элементная выборка представляет собой либо k-элементное подмножество n-множества, либо k-мультимножество над n-элементным множеством в зависимости от того, возвращаем мы вытаскиваемые предметы обратно в урну или не возвращаем. Следовательно, количество таких неупорядоченных выборок совпадает с коэффициентами $\binom{n}{k}$ или $\left(\!\binom{n}{k}\!\right)$. Очевидно также, что любая упорядоченная k-элементная выборка есть просто некоторая k-перестановка n-элементного множества. Поэтому количество таких выборок равно n^k или $(n)_k$ в зависимости от того, говорим ли мы о выборке с повторениями или без повторений.

В результате получаем следующую таблицу решений задач, связанных с урновыми схемами:

Предметы на выходе	с возвращением	без возвращения
упорядоченные	n^k	$(n)_k$
неупорядоченные	$\left(\!\binom{n}{k}\!\right)$	$\binom{n}{k}$

3.2.2. Приведем несколько характерных примеров, достаточно естественно сводящихся к одной из описанных выше урновых схем.

Пример 3.3. Предположим, что у нас имеется некоторое общество, состоящее из двадцати членов. Сколькими способами можно выбрать президента, вице-президента, секретаря и казначея этого общества?

Решение. Очевидно, что любой способ выбора представляет собой упорядоченное 4-элементное подмножество 20-элементного множества. Следовательно, существует ровно $(20)_4$ различных способов выбора членов общества на эти должности.

Пример 3.4. Для того, чтобы открыть сейф, нужно набрать код из пяти символов с помощью вращающихся дисков. На каждом из этих дисков нанесено 12 символов, одинаковых для каждого из дисков. Сколько вариантов различных кодов существует?

Решение. Любой код представляет собой упорядоченную 5-элементную выборку с повторениями или, иначе, строку из пяти символов над алфавитом из 12 букв. Следовательно, имеется 12^5 вариантов различных кодов.

Пример 3.5. На почте продаются открытки десяти различных видов. Сколькими способами можно купить восемь открыток? А восемь открыток разных видов?

Решение. Понятно, что в первом случае любой набор из восьми открыток представляет собой неупорядоченную выборку с повторениями, а во втором — выборку без повторений из 10 элементов. Следовательно, в первом случае имеем $\left(\!\binom{10}{8}\!\right)$, а во втором — $\binom{10}{8}$ способов покупки восьми открыток.

3.2.3. Второй, не менее популярной в элементарной комбинаторике схемой, связанной с подсчетом количества k-перестановок и k-сочетаний, является схема раскладки предметов по ящикам. В этой схеме имеется n *различимых* ящиков, по которым нужно разложить k *различимых* или *неразличимых* предметов. При этом мы можем накладывать определенные ограничения на количество предметов в каждом ящике.

Рассмотрим, к примеру, задачу о подсчете количества способов раскладки k различимых предметов по n различимым ящикам при условии, что в любой ящик можно класть любое количество предметов. Количество способов совершить эти действия равно, очевидно, n^k. Действительно, любой предмет мы можем положить в любой из n ящиков вне зависимости от того, куда мы положили оставшиеся предметы. Поэтому, согласно правилу произведения, это количество равно n^k. Иными словами, данная задача представляет собой переформулировку задачи о подсчете количества k-перестановок из n элементов с повторениями.

Теперь предположим, что в той же схеме мы не имеем права класть более одного предмета в один ящик. Тогда первый предмет мы можем поместить в любой из n ящиков, второй — в любой из оставшихся свободными $(n-1)$ ящиков и так далее. Всего же получаем $(n)_k$ способов раскладки. Следовательно, данная схема соответствует подсчету k-перестановок без повторений.

3.2.4. Пусть теперь у нас имеются n различимых ящиков и k неразличимых предметов. Тогда подсчет количества различных способов раскладки этих предметов по ящикам сводится к задаче о подсчете количества k-сочетаний из n элементов.

Действительно, в данной схеме в качестве n-элементного множества выступает множество, состоящее из n различимых ящиков. В случае, когда в каждый ящик можно класть не более одного предмета, мы, раскладывая предметы по ящикам, выделяем в этом множестве некоторое k-элементное подмножество. Следовательно, количество таких раскладок совпадает с количеством различных k-элементных подмножеств n-множества и равно $\binom{n}{k}$.

В случае же, когда никаких ограничений на количество предметов в ящике не накладывается, мы, раскладывая по ящикам k неразличимых предметов, задаем тем самым некоторое k-мультимножество n-множества X. Поэтому количество различных способов такой раскладки равно количеству $\left(\!\binom{n}{k}\!\right)$ k-сочетаний из n элементов с повторениями.

Подводя итоги, построим таблицу рассмотренных схем раскладок n предметов по k ящикам:

Предметы	Ящики	Произвольное количество предметов в ящике	Не более одного предмета в ящике
различимые	различимые	n^k	$(n)_k$
неразличимые	различимые	$\left(\!\binom{n}{k}\!\right)$	$\binom{n}{k}$

3.2.5. Проиллюстрируем некоторые характерные примеры задач, которые довольно естественно сводятся к схеме схем раскладки предметов по ящикам.

Пример 3.6. Сколькими способами можно разложить по двум *различимым* карманам (например, левому и правому) девять монет *различного* достоинства?

Решение. Рассматриваемый пример является типичной задачей, которая естественным образом сводится к схеме раскладки предметов по ящикам. Действительно, в роли ящиков здесь выступают левый и правый карман, а в роли предметов — девять различных монет. Поэтому ответ в этой задаче — 2^9 способов.

Замечание 3.7. Рассмотренная задача, однако, не всегда решается верно: в качестве ответа иногда выдают 9^2 способов. Путаница, как правило, происходит потому, что эту задачу пытаются свести к урновой схеме, считая, что имеются урна, в которой расположены 9 различных предметов, а также 2 различимые позиции на выходе.

Для того, чтобы этой путаницы избежать, полезно сформулировать следующие основные отличия схемы раскладки по ящикам от соответствующей ей урновой схемы. Во-первых, в схеме раскладки предметов по ящикам предметы обратно не возвращаются, они остаются в ящике. Во-вторых, в этой схеме в любой ящик можно класть любое количество предметов. В аналогичной урновой схеме на любую позицию помещается ровно один предмет.

Приведем теперь два характерных примера, связанных с раскладкой неразличимых предметов по различимым ящикам.

Пример 3.8. У отца имеется 5 (неразличимых) апельсинов, которые он может раздать восьми своим сыновьям. Если его задача состоит в том, чтобы раздать их максимальному количеству сыновей, то он должен поставить дополнительное условие — любой из его сыновей не должен получить более одного апельсина. В этом случае количество способов, которыми он может раздать своим сыновьям эти пять апельсинов, равно, очевидно, $\binom{8}{5}$. Если же он раздает их по каким-то заслугам, и может, таким образом, любому сыну отдать любое количество апельсинов, то количество способов это сделать равно $\left(\!\!\binom{8}{5}\!\!\right) = \binom{12}{5}$.

Пример 3.9. В физике встречаются задачи, в которых имеются n различных уровней энергии и k неразличимых элементарных частиц. Если эти частицы — фермионы, то для них действует так называемый принцип запрета Паули, согласно которому на любом энергетическом уровне может находиться не более одной элементарной частицы. Как следствие, количество различных распределений k фермионов по n энергетическим уровням равно $\binom{n}{k}$. Наряду с фермионами существуют и частицы иного сорта — бозоны, для которых не существует ограничений на количество частиц, занимающих один и тот же уровень энергии. Для бозонов количество таких распределений равно, очевидно, $\left(\!\!\binom{n}{k}\!\!\right)$.

3.2.6. К задачам раскладки неразличимых предметов по различимым ящикам, связанным с подсчетом количества k-сочетаний, сводятся также задачи о так называемом *разбиении* натурального числа k на n слагаемых. Данная задача формулируется следующим образом: сколькими способами можно представить натуральное число k в виде суммы n слагаемых вида

$$a_1 + a_2 + \ldots + a_n = k$$

при условии, что порядок слагаемых важен, то есть при условии, что разбиения вида

$$1 + 3 + 3 + 3 = 10 \qquad \text{и} \qquad 3 + 1 + 3 + 3 = 10$$

считаются различными?

Если на числа a_i накладывается единственное условие вида $a_i \geqslant 0$, то количество разбиений равно количеству $\left(\!\!\binom{n}{k}\!\!\right)$ k-мультимножеств над n-элементным множеством. Действительно, в упорядоченной сумме $a_1 + a_2 + \ldots + a_n$ любой индекс i слагаемого a_i можно рассматривать как i-й ящик, в который мы складываем a_i единиц. Следовательно, эту задачу можно трактовать как задачу о раскладке k “неразличимых” единиц по n различимым ящикам.

Пример 3.10. Подсчитать количество разбиений числа $k = 4$ на два слагаемых:

$$a_1 + a_2 = 4, \qquad a_1, a_2 \geqslant 0.$$

Ответ: $\left(\!\binom{2}{4}\!\right) = \binom{5}{4} = 5$ разбиений: $0+4 = 1+3 = 2+2 = 3+1 = 4+0 = 4$.

К подсчету числа k-сочетаний из n элементов без повторений задача о разбиении числа k сводится в случае, когда на числа a_i накладываются следующие условия:

$$a_i = 0 \qquad \text{или} \qquad a_i = 1.$$

В этом случае индекс i также можно трактовать как i-й ящик; его можно выбрать (положив $a_i = 1$) или не выбрать (положив $a_i = 0$). Всего же нужно выбрать k таких ящиков. Это можно сделать $\binom{n}{k}$ способами.

Пример 3.11. Подсчитать количество разбиений числа 2 на три слагаемых при условии, что любое слагаемое может принимать значения 0 или 1.

Ответ: $\binom{3}{2} = 3$ разбиения: $1+1+0 = 1+0+1 = 0+1+1 = 2$.

3.2.7. Достаточно часто на практике встречаются ситуации, когда одну и ту же задачу можно свести и к урновой схеме, и к схеме раскладки предметов по ящикам.

Пример 3.12. В кондитерском магазине продаются пирожные трех разных видов. Сколькими различными способами можно купить семь пирожных?

Решение. Ответ в этой задаче, очевидно, равен $\left(\!\binom{3}{7}\!\right) = \binom{9}{7}$. Этот ответ можно, например, получить, представляя себе коробку с тремя отделениями, в каждое из которых кладется пирожное только одного вида; в этом случае мы сводим задачу к подсчету количества раскладок семи неразличимых предметов по трем различимым ящикам. Другой способ получить тот же ответ — это представлять себе урну, в которой находятся три разных пирожных, и считать количество способов выбора из урны семи пирожных с возвращениями любого выбранного пирожного обратно в урну. Наконец, можно вообще забыть о любых схемах, если понимать, что любые семь купленных пирожных трех различных видов представляют собой 7-мультимножество над 3-элементным множеством различных видов пирожных.

4 Подсчет количества отображений конечных множеств. Числа Стирлинга второго рода

4.1. Оказывается, задачи о раскладке различимых предметов по различимым же ящикам имеют и еще одну, чрезвычайно важную комбинаторную интерпретацию — они эквивалентны задачам о подсчете количества отображений конечных множеств.

4.1.1. Напомним определение произвольного отображения $f\colon X \to Y$.

Определение 4.1. Пусть X, Y — пара конечных множеств. Отображением f из X в Y называется правило, согласно которому любому элементу $x \in X$ ставится в соответствие единственный элемент $y \in Y$:

$$\forall\, x \in X \qquad \exists!\, y \in Y\colon \qquad y = f(x).$$

С комбинаторной точки зрения любое отображение f из n-элементного множества X в k-элементное множество Y можно рассматривать как некоторый вариант раскладки n различимых предметов по k различимым ящиками при отсутствии каких-либо ограничений на количество предметов в каждом ящике.

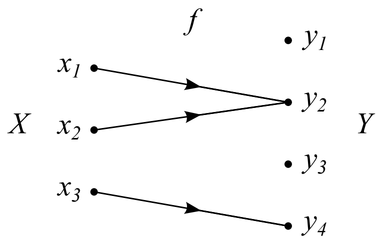

Рис. 6

Пример 4.2. Рассмотрим отображение f из трехэлементного множества X в четырехэлементное множество Y вида

$$f(x_1) = y_2, \qquad f(x_2) = y_2, \qquad f(x_3) = y_4$$

(смотри рис.6). Этому отображению отвечает раскладка трех различимых предметов по четырем различимым ящикам, при которой первые два предмета размещаются во втором ящике, а третий предмет — в четвертом ящике.

Как следствие, общее количество всех отображений n-элементного множества X в k-элементное множество Y равно k^n.

4.1.2. Напомним теперь определение *инъективного* отображения.

Определение 4.3. Отображение $f\colon X \to Y$ называется инъективным, если из условия

$$f(x_1) = f(x_2) \qquad \Longrightarrow \qquad x_1 = x_2.$$

Иными словами, отображение называется инъективным, если у любого элемента $y \in Y$ имеется не более одного прообраза, т.е. элемента $x \in X$, такого, что $y = f(x)$.

Понятно, что любому инъективному отображению $f\colon X \to Y$ отвечает такая раскладка n элементов множества X, при которой в каждом из k ящиков находится не более одного элемента. Как следствие, количество всевозможных инъективных отображений равно $(k)_n$.

4.1.3. Наконец, рассмотрим случай биективного и сюръективного отображений.

Определение 4.4. Отображение $f\colon X \to Y$ называется *биективным,* если

$$\forall\, y \in Y \qquad \exists!\, x \in X\colon \qquad y = f(x).$$

Количество таких отображений равно, очевидно, $n!$, где $n = |X| = |Y|$.

Определение 4.5. Отображение $f\colon X \to Y$ называется сюръективным, если

$$\forall\, y \in Y \qquad \exists\, x \in X\colon \qquad y = f(x).$$

Другими словами, отображение f сюръективно, если у любого элемента $y \in Y$ найдется хотя бы один прообраз $x \in X$, то есть такой x, что $f(x) = y$.

Комбинаторная интерпретация сюръективного отображения такова: это есть некоторая раскладка n различимых предметов по k различимым ящикам при условии, что в каждом ящике находится хотя бы один предмет. Количество таких раскладок при $k > n$ равно, очевидно, нулю. Задача следующего пункта данного параграфа — сосчитать количество этих раскладок для случая $0 \leqslant k \leqslant n$.

4.2. Обозначим через $\widehat{S}(n,k)$ количество всех сюръективных отображений n-элементного множества X в k-элементное множество Y. Сосчитаем $\widehat{S}(n,k)$ для случая $n \geqslant 0$, $0 \leqslant k \leqslant n$.

4.2.1. Рассмотрим множество *всех* отображений из n-множества X в k-множество Y. Как мы знаем, количество таких отображений равно k^n. Наша задача состоит в том, чтобы подсчитать это количество по-другому, выразив k^n через числа $\widehat{S}(n,k)$.

4.2.2. Заметим, что *любое* отображение $f\colon X \to Y$ можно рассматривать как *сюръективное* отображение множества X на множество

$$\operatorname{Im}(f) = \{y \in Y \mid \exists x\colon\ y = f(x)\},$$

являющееся образом множества X при отображении f.

Так, для отображения f из примера 4.2 образ $\operatorname{Im}(f) = \{y_2, y_4\}$, а отображение $f\colon X \to Y$ является сюръективным отображением множества X на подмножество $\operatorname{Im}(f) \subset Y$.

4.2.3. Разобъем теперь все множество отображений $f\colon X \to Y$ на блоки, включив в i-й блок все отображения, образ $\operatorname{Im}(f)$ которых содержит ровно i элементов: $|\operatorname{Im}(f)| = i$, $i = 1, \ldots, k$. Все, что нам остается — это сосчитать количество элементов в каждом блоке, а затем воспользоваться правилом суммы для того, чтобы получить общее количество k^n всех отображений.

4.2.4. Заметим, что существует $\binom{k}{i}$ способов выбрать i-элементное подмножество k-множества Y. Для каждого из этих подмножеств имеется $\widehat{S}(n,i)$ различных сюръективных отображений из n-элементного множества X в выбранное i-элементное подмножество множества Y. Таким образом, по правилу произведения, общее количество элементов в i-м блоке равно

$$\binom{k}{i} \cdot \widehat{S}(n,i).$$

Тогда по правилу суммы можем записать, что

$$k^n = \sum_{i=0}^{k} \binom{k}{i} \cdot \widehat{S}(n,i). \tag{14}$$

При этом мы суммируем не от 1 до k, а от 0 до k, учитывая, что $\widehat{S}(n,0) = 0$ для всех $n > 0$.

Замечание 4.6. Формулу (14) полезно иногда записывать в виде

$$k^n = \sum_{i=0}^{n} \binom{k}{i} \cdot \widehat{S}(n,i). \tag{15}$$

Несложно убедиться в том, что эта формула непосредственно следует из (14), а также в том, что она оказывается справедливой как для случая $n \geqslant k$, так и для случая $n < k$.

4.2.5. Мы выразили количество всех отображений n-элементного множества X в k-элементное множество Y через количество $\widehat{S}(n,i)$ сюръективных отображений. Нам же нужна обратная формула, выражающая количество $\widehat{S}(n,k)$ сюръективных отображений через число i^n всех отображений. Для ее получения воспользуемся так называемыми *формулами обращения.*

Утверждение 4.7. *Пусть* $(f_0, f_1, f_2, \ldots)$ *и* $(g_0, g_1, g_2, \ldots)$ — *две числовые последовательности, и пусть одна из них выражается через вторую по формулам*

$$f_k = \sum_{i=0}^{k} \binom{k}{i} g_i, \qquad k \geqslant 0. \tag{16}$$

Тогда справедлива следующая формула обращения:

$$g_k = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} f_i, \qquad k \geqslant 0. \tag{17}$$

С учетом этих формул обращения можно, считая n параметром, из соотношения (14) получить следующую явную формулу для вычисления чисел $\widehat{S}(n,k)$:

$$\widehat{S}(n,k) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} \cdot i^n.$$

4.3. Задачи подсчета количества отображений n-элементного множества X в k-элементное множество Y имеют еще одну важную комбинаторную интерпретацию.

4.3.1. Начнем с простого примера.

Пример 4.8. Для трехэлементного множества $X = \{x_1, x_2, x_3\}$ и двухэлементного множества $Y = \{y_1, y_2\}$ имеется, как мы знаем, $2^3 = 8$ различных отображений множества X в множество Y. Запишем все эти отображения как упорядоченные пары подмножеств множества X:

$$(\{x_1, x_2, x_3\}, \varnothing), \qquad (\{x_1, x_2\}, \{x_3\}), \qquad (\{x_1, x_3\}, \{x_2\}), \qquad (\{x_2, x_3\}, \{x_1\}),$$

$$(\{x_1\}, \{x_2, x_3\}), \qquad (\{x_2\}, \{x_1, x_3\}), \qquad (\{x_3\}, \{x_1, x_2\}), \qquad (\varnothing, \{x_1, x_2, x_3\}).$$

Видно, что записанное в таком виде решение представляет собой не что иное, как список всех возможных *разделений* множества X, то есть упорядоченных разбиений X на два блока, один из которых может быть и пустым.

4.3.2. Очевидно, что данный результат справедлив и в общем случае. Именно, любое отображение $f\colon X \to Y$ задает нам некоторое разделение множества X, то есть разбиение этого множества на k упорядоченных блоков, часть из которых могут быть пустыми. Как следствие, количество таких разделений совпадает с количеством всех отображений f и равно k^n.

Аналогичные рассуждения показывают, что любое сюръективное отображение $f\colon X \to Y$ задает нам некоторое *упорядоченное разбиение* множества X на блоки. Поэтому количество всех упорядоченных разбиений n-элементного множества X на k блоков равно числу $\widehat{S}(n,k)$.

4.3.3. Рассмотрим теперь некоторый специальный вид k-разделений множества X, а именно, такие k-разделения, в которых в первом блоке содержится a_1 элемент, во втором блоке — a_2 элемента, в k-м блоке — a_k элементов. Очевидно, что при этом общая сумма всех элементов должна быть равна мощности $|X| = n$:

$$a_1 + a_2 + \ldots + a_k = n, \qquad a_i \geqslant 0.$$

Утверждение 4.9. *Количество всех таких k-разделений n-множества X равно*

$$P(n; a_1, a_2, \ldots, a_k) := \binom{n}{a_1} \cdot \binom{n-a_1}{a_2} \cdot \ldots \cdot \binom{n-a_1-a_2-\ldots-a_{k-1}}{a_k} = \frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}. \quad (18)$$

Доказательство. Действительно, из любого n-элементного множества X мы $\binom{n}{a_1}$ способами можем выбрать a_1 элементов и положить их в первый ящик (отнести к первому блоку разбиения). Затем для каждого такого выбора мы $\binom{n-a_1}{a_2}$ способами можем из оставшегося $(n - a_1)$-элементного множества выбрать a_2 элементов и положить их во второй ящик (отнести ко второму блоку разбиения), и так далее. Формула (18), описывающая общее количество способов совершить все эти действия, следует теперь из правила произведения. □

Следствие 4.10. *Общее количество k^n всех k-разделений n-множества X выражается через числа $P(n; a_1, a_2, \ldots, a_k)$ по формуле*

$$k^n = \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ a_i \geqslant 0}} P(n; a_1, a_2, \ldots, a_k) = \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ a_i \geqslant 0}} \frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}. \quad (19)$$

Замечание 4.11. Если в условии рассматриваемой задачи заменить нестрогие неравенства $a_i \geqslant 0$ на строгие, то есть на неравенства $a_i > 0$, то вместо разделения мы получим упорядоченное разбиение специального вида. Количество таких упорядоченных разбиений также описывается формулой (18), а вместо формулы (19) получается не менее полезное соотношение вида

$$\widehat{S}(n, k) = \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ a_i > 0}} \frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}. \quad (20)$$

4.4. Числа $P(n; a_1, a_2, \ldots, a_k)$ имеют и еще один важный комбинаторный смысл. Именно, они перечисляют так называемые *перестановки n-множества X с повторениями.*

4.4.1. Рассмотрим в качестве элементарного примера следующую задачу: на полке имеются 15 различных книг по математике, 16 по информатике и 12 по физике; каково количество способов перестановки этих книг на полке? Ответ очевиден: $(15 + 16 + 12)! = 43!$ способов.

Предположим теперь, что мы перестали различать книги, посвященные одному и тому же предмету. В этом случае количество различных способов перестановки таких книг уменьшится. Обозначим это количество через λ_n. Так как существует 15! способов упорядочить книги по математике, 16! – по информатике и 12! – по физике, то по правилу произведения мы можем записать, что

$$43! = \lambda_n \cdot 16! \cdot 15! \cdot 12! \qquad \Longrightarrow \qquad \lambda_n = \frac{43!}{16! \cdot 15! \cdot 12!} = P(43; 15, 16, 12).$$

4.4.2. Аналогичные рассуждения справедливы и в общем случае. Именно, пусть среди n переставляемых предметов имеется a_1 неразличимых предметов первого сорта, a_2 неразличимых предметов второго сорта и так далее, причем $a_1 + a_2 + \ldots + a_k = n$. Тогда для количества перестановок таких предметов с повторениями получаем уже знакомую нам формулу

$$\frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!} = P(n; a_1, a_2, \ldots, a_k).$$

4.4.3. В частном случае перестановки n предметов двух различных сортов получаем

$$P(n; k, n-k) = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{k}.$$

Иными словами, количество таких перестановок совпадает с количеством различных k-элементных подмножеств n-элементного множества. Для комбинаторного доказательства данного факта можно воспользоваться рассуждениями, которые мы проводили при подсчете количества всех подмножеств данного множества. Напомним, что там мы любое подмножество кодировали упорядоченной битовой строкой длины n, состоящей из k единиц и $(n-k)$ нулей. Осталось заметить, что любая такая строка представляет собой некоторую перестановку k элементов первого сорта (единиц) и $(n-k)$ элементов второго сорта (нулей).

4.4.4. В заключение отметим еще одну полезную биекцию, позволяющую несколько по-другому сосчитать количество k-мультимножеств n-элементного множества. Мы знаем, что любому k-мультимножеству над n-множеством отвечает некоторая раскладка k неразличимых предметов по n различимым ящикам. В свою очередь, любую такую раскладку можно рассматривать как упорядоченный набор, состоящий из k неразличимых предметов одного сорта (например, k неразличимых шаров), и $(n-1)$-го предмета второго сорта ($(n-1)$-й неразличимой перегородки между этими шарами). Как следствие, количество всех k-мультимножеств

$$\left(\!\binom{n}{k}\!\right) = P(k+n-1; k, n-1) = \frac{(n+k-1)!}{(n-1)! \cdot k!} = \binom{n+k-1}{k}.$$

4.5. Вернемся к числам $\widehat{S}(n,k)$, описывающим, в частности, количество всех упорядоченных разбиений n-множества на k блоков. Обозначим теперь через $S(n,k)$ количество обычных, неупорядоченных разбиений n-множества на k блоков. Так как для любого неупорядоченного разбиения существует $k!$ способов упорядочить k его блоков, то

$$\widehat{S}(n,k) = k!\, S(n,k).$$

Отсюда с учетом (14) и (20) получаем следующие две явные формулы, позволяющие вычислять числа $S(n,k)$:

$$S(n,k) = \frac{1}{k!}\sum_{i=0}^{k}(-1)^k\binom{k}{i}(k-i)^n = \frac{1}{k!}\sum_{\substack{a_1+a_2+\ldots+a_k=n\\ a_i>0}} \frac{n!}{a_1!\cdot a_2!\cdot\ldots\cdot a_k!}. \tag{21}$$

Числа $S(n,k)$ называются *числами Стирлинга второго рода* и встречаются в большом количестве комбинаторных приложений. Исследуем эти числа поподробнее.

4.5.1. Для практического расчета чисел $S(n,k)$ удобно использовать рекуррентные соотношения.

Утверждение 4.12. *Числа Стирлинга второго рода удовлетворяют следующим рекуррентным соотношениям:*

$$S(n,k) = S(n-1,k-1) + k\cdot S(n-1,k), \qquad k = 1,\ldots,n; \tag{22}$$

$$S(0,0) = 1; \qquad S(n,0) = 0 \quad \forall\, n > 0; \qquad S(n,k) = 0 \quad \forall\, k > n.$$

Доказательство. Граничные условия $S(n, 0) = 0$ для всех $n > 0$ и $S(n, k) = 0$ для всех $k > n$ очевидны — n-элементное множество в случае $n > 0$ нельзя разбить на 0 блоков, а также на k блоков в случае, когда $k > n$. Равенство $S(0, 0) = 1$ введено просто для удобства.

Докажем теперь соотношение (22). Разобьем для этого множество всех k-разбиений на два блока. К первому блоку Σ_1 отнесем все разбиения, содержащие одноэлементное подмножество $\{x_1\}$. Ко второму блоку Σ_2 отнесем все оставшиеся k-разбиения, т.е. разбиения, в которых элемент x_1 входит в подмножества, содержащие как минимум два элемента.

Рассмотрим, к примеру, все 2-разбиения множества $X = \{x_1, x_2, x_3, x_4\}$:

$$\{\{x_1, x_2, x_3\}, \{x_4\}\}, \quad \{\{x_1, x_2, x_4\}, \{x_3\}\}, \quad \{\{x_1, x_3, x_4\}, \{x_2\}\}, \quad \{\{x_2, x_3, x_4\}, \{x_1\}\},$$

$$\{\{x_1, x_2\}, \{x_3, x_4\}\}, \quad \{\{x_1, x_3\}, \{x_2, x_4\}\}, \quad \{\{x_1, x_4\}, \{x_2, x_3\}\}.$$

Для этого примера к первому блоку относится единственное разбиение вида $\{\{x_2, x_3, x_4\}, \{x_1\}\}$. Остальные шесть разбиений относятся в данном случае ко второму блоку.

Довольно очевидно, что количество элементов в первом блоке равно количеству $S(n-1, k-1)$ всех $(k-1)$-разбиений оставшегося $(n-1)$-элементного множества. В примере это число равно $S(3, 1) = 1$ — любое множество можно только одним способом разбить на один блок.

Для подсчета количества элементов во втором блоке заметим, что, по сути дела, это число равно количеству всевозможных разбиений $(n-1)$-элементного множества на k непустых подмножеств с поочередным добавлением элемента x_1 в каждое из этих подмножеств. Действительно, для разобранного выше примера имеется три разбиения трехэлементного множества $\{x_2, x_3, x_4\}$ на два непустых подмножества, а именно, разбиения вида

$$\{\{x_2\}, \{x_3, x_4\}\}, \quad \{\{x_3\}, \{x_2, x_4\}\}, \quad \{\{x_4\}, \{x_2, x_3\}\}.$$

Добавляя к каждому из этих подмножеств элемент x_1, получим $2 \cdot 3 = 6$ выписанных выше разбиений четырехэлементного множества X на блоки требуемого вида. Количество таких разбиений в общем случае равно, очевидно, $k \cdot S(n-1, k)$. □

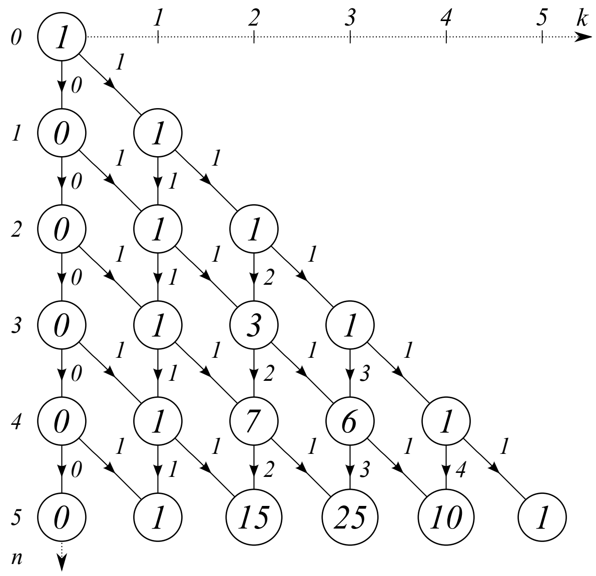

Рис. 7: Графическое представление чисел $S(n, k)$

4.5.2. Как и биномиальные коэффициенты, числа Стирлинга второго рода удобно представлять в виде треугольного массива на плоскости (n, k). На рис.7 показаны первые несколько строк такого массива, вычисленных с помощью рекуррентных соотношений (22). Такой рисунок дает нам еще одну комбинаторную интерпретацию чисел $S(n, k)$ — это есть количество *взвешенных* путей, идущих из начала координат в точку с координатами (n, k).

4.5.3. Числа Стирлинга второго рода встречается в очень большом количестве самых разнообразных задач. В качестве примера вернемся к формуле (15) и перепишем ее через числа Стирлинга второго рода:

$$k^n = \sum_{i=0}^{n} \binom{k}{i} \cdot i! \cdot S(n, i) = \sum_{i=0}^{n} (k)_i \cdot S(n, i). \tag{23}$$

Оказывается, эта формула справедлива для любых вещественных и даже комплексных значений k. Именно, справедливо равенство

$$x^n = \sum_{i=0}^{n} (x)_i \cdot S(n, i).$$

Эта формула активно используется в теории конечных операторов — она позволяет перейти от базиса x^n к базису $(x)_n$.

4.5.4. С точки зрения схемы раскладки предметов по ящикам числа Стирлинга второго рода описывают количество способов разложить n различимых предметов по k неразличимым ящикам так, чтобы в любом ящике содержался хотя бы один предмет. Сосчитаем, зная $S(n, k)$, количество $B(n, k)$ различных раскладок n различимых предметов по k неразличимым ящикам в случае, когда ограничения на количество предметов в любом ящике отсутствуют.

Для этого вновь разобьем множество всех таких раскладок на блоки, поместив в i-й блок все раскладки, в которых занято ровно i ящиков. В случае различимых ящиков в процессе аналогичного разбиения множества на блоки мы в i-м блоке имели $\binom{k}{i} \cdot \widehat{S}(n, i)$ элементов. В случае неразличимых ящиков мы i из k ящиков выбираем ровно одним способом, а затем $S(n, i)$ способами заполняем эти ящики n предметами. Используя правило суммы, получаем отсюда следующее выражение для чисел $B(n, k)$:

$$B(n, k) = \sum_{i=1}^{k} S(n, i).$$

4.5.5. Числа $B(n, k)$ в случае $n = k$ называются числами Белла $B(n)$. Эти числа перечисляют количество *всех* возможных разбиений n-элементного множества X.

Заметим, что любое разбиение множества X можно получить, введя на этом множестве некоторое отношение эквивалентности. Как следствие, количество всех возможных отношений эквивалентности на n-элементном множестве X описывается числами Белла B_n.

4.5.6. Покажем, что для чисел Белла $B(n)$ справедливо следующее рекуррентное соотношение:

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k. \tag{24}$$

Действительно, рассмотрим блок разбиения $(n+1)$-элементного множества, содержащий число $n+1$. Все возможные разбиения мы можем разбить на $n+1$ группу (блок) в зависимости от того, сколько чисел содержится вместе с $n+1$. Предоположим, что вместе с $n+1$ находятся i чисел, $i = 0, \dots, n$. Эти числа мы можем выбрать $\binom{n}{i} = \binom{n}{n-i}$ способами. Оставшиеся $k = n - i$ чисел, $k = 0, \dots, n$, мы можем $B(k)$ способами разбить на блоки. Суммируя по всем k, мы и получим формулу (24).

4.6. Подведем итоги, построив таблицу различных схем раскладок n предметов по k ящикам:

Элементы множества X (предметы)	Элементы множества Y (ящики)	Произвольное количество предметов в ящике	Не более 1 предмета в ящике	Как минимум 1 предмет в ящике
различимые	различимые	k^n	$(k)_n$	$\widehat{S}(n,k)$
неразличимые	различимые	$\left(\!\binom{k}{n}\!\right)$	$\binom{k}{n}$	$\binom{n-1}{k-1}$
различимые	неразличимые	$B(n,k)$	$0, \quad n > k$ $1, \quad n \leqslant k$	$S(n,k)$

Как видно, остался еще один неразобранный вариант — схема размещения n неразличимых предметов по k неразличимым ящикам. Для количества таких размещений явных аналитических формул не существует. Для того, чтобы их перечислить, необходимо использовать аппарат производящих функций, к изучению которого мы приступим в одном из следующих параграфов.

5 Рекуррентные соотношения

5.1. Мы уже несколько раз получали решения комбинаторных задач, записанные в виде тех или иных рекуррентных соотношений. Настало время поговорить о них немного поподробнее.

5.1.1. Начнем с простого примера [2].

Пример 5.1. 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.

Решение. Обозначим через a_n количество лягушек в начале $(n+1)$-го года. По условию задачи, $a_0 = 50$. Тогда, очевидно,

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

а в общем случае

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

Полученное равенство является простейшим примером *рекуррентного соотношения.*

5.1.2. Перейдем теперь к формальным определениям.

Определение 5.2. 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.

$$a_{n+m} = f(a_n, a_{n+1}, \ldots, a_{n+m-1}), \tag{25}$$

то такая последовательность называется рекуррентной последовательностью, а соотношение (25) — рекуррентным соотношением m-го порядка.

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

$$a_{n+m} = b_1(n)\, a_{n+m-1} + b_2(n)\, a_{n+m-2} + \ldots + b_{m-1}(n)\, a_{n+1} + b_m(n)\, a_n + u_n. \tag{26}$$

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

Самый простой случай рекуррентного соотношения (26) — это *линейное однородное рекуррентное соотношение с постоянными коэффициентами*

$$a_{n+m} = b_1\, a_{n+m-1} + b_2\, a_{n+m-2} + \ldots + b_{m-1}\, a_{n+1} + b_m\, a_n. \tag{27}$$

Очевидно, что для однозначного определения всех a_n необходимо наряду с самим рекуррентным соотношением (25) задать и первые m членов $a_0, a_1, \ldots, a_{m-1}$ данной последовательности, то есть, как говорят, *задать начальные условия* для рекуррентного соотношения.

5.1.3. 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 для поставленной комбинаторной задачи.

Однако иногда нам полезно иметь явную формулу, которая для любого n позволяет вычислять значение a_n. Такую формулу называют *решением* рассматриваемого рекуррентного соотношения. Мы в данном параграфе покажем, как строить такие решения для случая линейного однородного рекуррентного соотношения с постоянными коэффициентами (27).

5.2. Прежде чем рассматривать общий случай уравнения (27), рассмотрим наиболее простой его вариант — линейное однородное рекуррентное соотношение первого порядка

$$a_{n+1} = b_1 \cdot a_n, \qquad n = 0, 1, 2, \ldots; \qquad a_0 \text{ — заданное число.} \tag{28}$$

5.2.1. Решение этого уравнения построить легко. Действительно,

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

5.2.2. Предположим теперь, что нам кто-то сразу подсказал вид решения, а именно, что решение нашего уравнения (28) степенным образом зависит от n:

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

Воспользовавшись этой подсказкой, подставим это выражение в исходное рекуррентное соотношение (28). 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 (28) в частном случае $a_0 = 1$. Поэтому такое решение называется *частным решением* уравнения (28).

5.2.3. Теперь попытаемся, зная частное решение, построить общее решение задачи (28). Для этого заметим, что в силу однородности уравнения (28) любое его частное решение, 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 (28), 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 (28).

5.3. Подведем предварительные итоги. Так как исходное уравнение (28) было очень простым, то нам удалось сразу построить его решение и выяснить, что оно степенным образом зависит от параметра n. Затем мы заметили, что если бы нам кто-то заранее подсказал степенной характер решения уравнения (28), то нам хватило бы этой информации для построения как общего, так и частного решения нашего рекуррентного уравнения.

Возникает вопрос: зачем же нам нужен для столь простого уравнения столь сложный алгоритм построения его решения? Оказывается, что этот алгоритм практически без изменений работает и для построения решения линейного однородного уравнения произвольного порядка.

5.3.1. Рассмотрим в качестве примера линейное однородное рекуррентное соотношение второго порядка

$$a_{n+2} = b_1 \cdot a_{n+1} + b_2 \cdot a_n, \qquad n = 0, 1, 2, \ldots, \qquad a_0,\, a_1 \text{ — заданные числа,} \tag{29}$$

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

Для этого предположим, что частное решение уравнения (29) по прежнему степенным образом зависит от параметра n, то есть предположим, что существуют такие значения параметров a_0 и a_1, при которых $a_n = r^n$. Подставляя это выражение в рекуррентное соотношение, получим

$$r^{n+2} = b_1 \cdot r^{n+1} + b_2 \cdot r^n \qquad \Longrightarrow \qquad r^2 - b_1\, r - b_2 = 0, \tag{30}$$

т.е. квадратное уравнение на r. Любое его решение r_0 дает нам некоторое частное решение уравнения (29). По этой причине данное уравнение называется *характеристическим уравнением* для рекуррентного соотношения (29).

5.3.2. Предположим, что характеристическое уравнение (30) имеет два различных вещественных корня r_1 и r_2. Покажем, что в таком случае выражение

$$a_n = c_1\, r_1^n + c_2\, r_2^n, \tag{31}$$

где c_1 и c_2 — произвольные постоянные, является *общим решением* соотношения (29) в том смысле, что любое решение (29) с заданными начальными условиями в нем содержится.

Действительно, покажем вначале, что выражение (31) действительно удовлетворяет рекуррентному соотношению (29):

$$c_1\, r_1^{n+2} + c_2\, r_2^{n+2} = b_1\, c_1\, r_1^{n+1} + b_1\, c_2\, r_2^{n+1} + b_2\, c_1\, r_1^n + b_2\, c_2\, r_2^n =$$

$$= c_1\,(r_1^2 - b_1\,r_1 - b_2) + c_2\,(r_1^2 - b_1\,r_1 - b_2) \equiv 0.$$

Покажем теперь, что это — действительно общее решение, т.е. что мы всегда можем подобрать константы c_1 и c_2 так, чтобы решение вида (31) удовлетворяло любым заданным начальным условиям. Для этого рассмотрим это выражение при $n = 0$ и $n = 1$:

$$\begin{aligned} a_0 &= c_1\,r_1^0 + c_2\,r_2^0 = c_1 + c_2, \\ a_1 &= c_1\,r_1^1 + c_2\,r_2^1 = c_1\,r_1 + c_2\,r_2. \end{aligned}$$

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.

5.3.3. Пожалуй, самым известным и важным примером рекуррентного соотношения (29) 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:

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

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

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

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n. \tag{32}$$

5.3.4. Предположим теперь, что характеристическое уравнение (30) имеет ровно один кратный корень $r_1 = r_2 =: \rho$. Покажем, что в этом случае общее решение уравнения (29) имеет вид

$$a_n = c_1\,\rho^n + c_2\,n\,\rho^n.$$

Мы уже показали, что частное решение вида ρ^n удовлетворяет уравнению (29). С учетом линейности и однородности этого уравнения нам осталось показать, что этому уравнению удовлетворяет и частное решение вида $n\,\rho^n$. Подставляя в уравнение (29) выражение $n\,\rho^n$, получим

$$(n+2)\,\rho^{n+2} = b_1\,(n+1)\,\rho^{n+1} + b_2\,n\,\rho^n \qquad \Longrightarrow \qquad n\,(\rho^2 - b_1\,\rho - b_2) + \rho\,(2\,\rho - b_1) = 0.$$

Первое слагаемое в левой части этого выражения равно нулю в силу характеристического уравнения. Для того, чтобы понять, что равно нулю и последнее слагаемое, заметим, что в случае совпадающих корней $r_1 = r_2 = \rho$ имеем

$$\rho = \frac{b_1 \pm 0}{2} \qquad \Longrightarrow \qquad 2\rho - b_1 = 0.$$

Осталось показать, что такой вид решения может удовлетворить любым начальным условиям. Подставив начальные условия в выражение для a_n, получим следующую систему двух линейных алгебраических уравнений относительно параметров c_1 и c_2:

$$c_1 = a_0,$$

$$\rho\,(c_1 + c_2) = a_1.$$

Понятно, что такая система имеет решение при любых a_0, a_1 и $\rho \neq 0$. Случай же $\rho = 0$ тривиален — в этом случае $b_1 = b_2 = 0$, и потому $a_n = 0$ для любых n.

5.3.5. Наконец, давайте рассмотрим случай, когда уравнение (30) имеет два комплексно сопряжённых корня

$$r_1 = x + iy = \rho\, e^{i\vartheta}, \qquad r_2 = x - iy = \rho\, e^{-i\vartheta}, \qquad \rho, \vartheta \neq 0.$$

Покажем, что в таком случае общее решение линейного однородного рекуррентного соотношения с постоянными коэффициентами второго порядка (29) можно записать так:

$$a_n = \widetilde{c}_1\, \rho^n\, \cos(n\,\vartheta) + \widetilde{c}_2\, \rho^n\, \sin(n\,\vartheta).$$

Действительно, рассуждения, аналогичные проведенным для случая двух различных вещественных корней, показывают, что выражение вида

$$a_n = c_1\, r_1^n + c_2\, r_2^n$$

удовлетворяет рекуррентному соотношению (29) при любых c_1 и c_2. Рассмотрим теперь систему линейных уравнений

$$\begin{aligned} &a_0 = c_1 + c_2, \\ &a_1 = c_1\, \rho\, e^{i\vartheta} + c_2\, \rho\, e^{-i\vartheta} = \rho\,(c_1 + c_2)\cos\vartheta + \rho\, i\,(c_1 - c_2)\sin\vartheta. \end{aligned} \tag{33}$$

Определитель этой системы

$$\begin{vmatrix} 1 & 1 \\ \rho\, e^{i\vartheta} & \rho\, e^{-i\vartheta} \end{vmatrix} = \rho(e^{-i\vartheta} - e^{i\vartheta}) = -\rho \cdot 2i \cdot \sin(\vartheta) \neq 0,$$

поэтому мы с помощью общего решения действительно сможем удовлетворить любым начальным условиям.

Перепишем теперь выражение для a_n:

$$a_n = c_1\, r_1^n + c_2\, r_2^n = c_1\, \rho^n\, e^{i\,\vartheta\, n} + c_2\, \rho^n\, e^{-i\,\vartheta\, n} =$$

$$= (c_1 + c_2)\, \rho^n\, \cos(n\,\vartheta) + i\,(c_1 - c_2)\, \rho^n\, \sin(n\,\vartheta).$$

Из (33) следует, что коэффициенты $\widetilde{c}_1 = c_1 + c_2$ и $\widetilde{c}_2 = i\,(c_1 - c_2)$ являются вещественными числами. Следовательно, общее решение в данном случае действительно можно записать в виде

$$a_n = \widetilde{c}_1\, \rho^n\, \cos(n\,\vartheta) + \widetilde{c}_2\, \rho^n\, \sin(n\,\vartheta).$$

5.3.6. Описанная выше техника решения линейных однородных рекуррентных соотношений с постоянными коэффициентами достаточно легко обобщается и на случай соотношений более высокого порядка. Именно, рассмотрим линейное однородное рекуррентное соотношение m-го порядка (27)

$$a_{n+m} = b_1 \cdot a_{n+m-1} + b_2 \cdot a_{n+m-2} + \ldots + b_m \cdot a_n, \qquad a_0, \ldots, a_{m-1} \text{ — заданы.}$$

Предположим, что частное решение этого соотношения имеет вид $a_n = r^n$. Подставляя это выражение в исходное рекуррентное соотношение, получаем для параметра r следующее характеристическое уравнение:

$$r^m = b_1 \cdot r^{m-1} + b_2 \cdot r^{m-2} + \ldots + b_m \cdot r.$$

Пусть $r_1, \ldots, r_k$ есть корни этого уравнения. Каждый такой корень дает свой вклад в общее решение рассматриваемого рекуррентного соотношения. В частности, если r_j есть корень кратности q_j, то вклад этого корня в общее решение задается выражением вида

$$c_{1,j} \cdot r_j^n + c_{2,j} \cdot n \cdot r_j^n + \ldots + c_{q_j,j} \cdot n^{q_j-1} \cdot r_j^n.$$

Так, если характеристическое уравнение для некоторого рекуррентного соотношения шестого порядка имеет вид

$$(r-1)^3 \cdot (r-2)^2 \cdot (r-3) = 0,$$

то общее решение такого соотношения записывается так:

$$a_n = c_1 + c_2 \cdot n + c_3 \cdot n^2 + c_4 \cdot 2^n + c_5 \cdot n \cdot 2^n + c_6 \cdot 3^n.$$

5.4. Читатели, хорошо знакомые с математическим анализом, уже, видимо, заметили, насколько похожи методы решения линейных рекуррентных соотношений и методы решения линейных обыкновенных дифференциальных уравнений с постоянными коэффициентами.

5.4.1. Действительно, рассмотрим линейное однородное дифференциальное уравнение первого порядка вида

$$\frac{dy}{dx} = b_1 \cdot y(x), \qquad y(x_0) = y_0.$$

Такое уравнение легко решается с помощью разделения переменных:

$$\frac{dy}{y} = b_1 \cdot dx \qquad \Longleftrightarrow \qquad d\ln(y) = b_1 \cdot dx \qquad \Longleftrightarrow \qquad \int\limits_{y_0}^{y} d\ln(y) = b_1 \cdot \int\limits_{x_0}^{x} dx \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \qquad \ln(y) - \ln(y_0) = \ln\frac{y}{y_0} = b_1 \cdot (x - x_0) \qquad \Longleftrightarrow \qquad y = y_0\, e^{b_1 \cdot (x - x_0)}.$$

Как и в случае рекуррентных соотношений, мы могли бы решить это уравнение и без использования метода разделения переменных — для его решения достаточно знать характер решения такого дифференциального уравнения. Именно, предположим, что решение $y(x)$ экспоненциально зависит от x:

$$y(x) = e^{\lambda \cdot x}.$$

Подставим это выражение в исходное дифференциальное уравнение:

$$y'(x) = \lambda \cdot e^{\lambda \cdot x} = b_1 \cdot e^{\lambda \cdot x} \qquad \Longleftrightarrow \qquad \lambda = b_1.$$

Заметим теперь, что решение $y(x) = e^{b_1 \cdot x}$ удовлетворяет начальным условиям $y(x_0) = e^{b_1 \cdot x_0}$. Для того, чтобы удовлетворить произвольным начальным условиям, достаточно воспользоваться однородностью данного уравнения и рассмотреть вместо частного общее решение вида $y(x) = c \cdot e^{b_1 \cdot x}$. Константа c легко находится из начальных условий:

$$y(x_0) = y_0 = c \cdot e^{b_1 \cdot x_0} \qquad \Longrightarrow \qquad c = y_0 \cdot e^{-b_1 \cdot x_0} \qquad \Longrightarrow \qquad y = y_0 \cdot e^{b_1 (x - x_0)}.$$

5.4.2. Теперь рассмотрим линейное однородное дифференциальное уравнение второго порядка с постоянными коэффициентами

$$y''(x) = b_1 \cdot y'(x) + b_2 \cdot y(x), \qquad\qquad y(0) = x_0, \quad y'(0) = x_1.$$

Предполагая, что характер решения этого уравнения остался тем же, что и в предыдущем случае, а именно, полагая, что $y(x) = e^{\lambda \cdot x}$, мы в результате подстановки этого выражения в исходное дифференциальное уравнение получаем абсолютно то же самое характеристическое уравнение на параметр λ, что и в случае линейного однородного рекуррентного соотношения второго порядка с постоянными коэффициентами:

$$y''(x) = \lambda^2 \cdot e^{\lambda \cdot x} = b_1 \cdot \lambda \cdot e^{\lambda \cdot x} + b_2 \cdot e^{\lambda \cdot x} \qquad \Longrightarrow \qquad \lambda^2 - b_1 \cdot \lambda - b_2 = 0.$$

Если теперь это уравнение имеет пару различных вещественных корней λ_1 и λ_2, то общее решение обыкновенного дифференциального уравнения в силу его линейности и однородности записывается в виде

$$y(x) = c_1 \cdot e^{\lambda_1 \cdot x} + c_2 \cdot e^{\lambda_2 \cdot x}.$$

В случае единственного вещественного корня λ характеристического уравнения решение попрежнему ищется в виде

$$y(x) = c_1 \cdot e^{\lambda \cdot x} + c_2 \cdot x \cdot e^{\lambda \cdot x}.$$

Наконец, в случае комплексных корней $\lambda_{1,2} = \alpha \pm i \cdot \beta$ решение обыкновенного дифференциального уравнения можно искать в виде

$$y(x) = c_1 \cdot e^{\alpha \cdot x} \cos(\beta \cdot x) + c_2 \cdot e^{\alpha \cdot x} \sin(\beta \cdot x).$$

5.4.3. Отмеченная аналогия наводит нас на мысль о существовании взаимно-однозначного отображения, позволяющего сопоставить линейному однородному рекуррентному соотношению с постоянными коэффициентами аналогичное линейное однородное дифференциальное уравнение с постоянными коэффициентами. Оказывается, что такое отображение построить достаточно просто. Рассмотрим для определенности линейное однородное рекуррентное соотношение с постоянными коэффициентами второго порядка:

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

Домножим это уравнение на $x^n/n!$ и просуммируем по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} a_{n+2} \frac{x^n}{n!} = b_1 \sum_{n=0}^{+\infty} a_{n+1} \frac{x^n}{n!} + b_2 \sum_{n=0}^{+\infty} a_n \frac{x^n}{n!}.$$

Предположим теперь, что ряд

$$a_0 + a_1 \frac{x^1}{1!} + a_2 \frac{x^2}{2!} + \ldots + a_n \frac{x^n}{n!} + \ldots = \sum_{n=0}^{+\infty} a_n \frac{x^n}{n!}$$

представляет собой ряд Тейлора некоторой функции $y(x)$, построенный в окрестности точки $x = 0$:

$$y(x) = \sum_{n=0}^{+\infty} a_n \frac{x^n}{n!}, \qquad a_n = y^{(n)}(0).$$

Тогда

$$y'(x) = a_1 + a_2\frac{x^1}{1!} + a_3\frac{x^2}{2!} + \ldots + a_n\frac{x^{n-1}}{(n-1)!} + a_{n+1}\frac{x^n}{n!} + \ldots = \sum_{n=0}^{+\infty} a_{n+1}\frac{x^n}{n!},$$

$$y''(x) = a_2 + a_3\frac{x^1}{1!} + a_4\frac{x^2}{2!} + \ldots + a_n\frac{x^{n-2}}{(n-2)!} + a_{n+1}\frac{x^{n-1}}{(n-1)!} + +a_{n+2}\frac{x^n}{n!} + \ldots = \sum_{n=0}^{+\infty} a_{n+2}\frac{x^n}{n!},$$

так что исходное рекуррентное соотношение превращается в обыкновенное дифференциальное уравнение вида

$$y''(x) = b_1 \cdot y'(x) + b_2 \cdot y(x)$$

с начальными условиями $y(0) = a_0$, $y'(0) = a_1$. Обратно, раскладывая функцию $y(x)$ в ряд Тейлора в окрестности точки $x = 0$, подставляя это разложение в обыкновенное дифференциальное уравнение и приравнивая коэффициенты при одинаковых степенях $x^n/n!$, мы из исходного обыкновенного дифференциального уравнения получаем рекуррентное соотношение на коэффициенты $a_n = y^{(n)}(0)$.

5.4.4. Полученное нами взаимно-однозначное преобразование полезно как для решения обыкновенных дифференциальных уравнений, так и для решения рекуррентных соотношений. Так, с вычислительной точки зрения иногда имеет смысл переходить от обыкновенного дифференциального уравнения к соответствующему рекуррентному соотношению. Быстрое вычисление чисел a_n позволяет эффективно вычислять коэффициенты в разложении искомой функции в ряд Тейлора, а следовательно, и значение этой функции в произвольной точке x.

Напротив, методы, разработанные для решения обыкновенных дифференциальных уравнений, часто оказываются полезными и для решения аналогичных рекуррентных соотношений. В качестве примера рассмотрим неоднородные линейные дифференциальные уравнения m-го порядка. Общее решение таких уравнений есть общее решение соответствующего однородного уравнения плюс частное решение неоднородного. Для нахождения частного решения неоднородного уравнения часто можно использовать так называемый метод неопределенных коэффициентов. На языке рекуррентных соотношений его можно сформулировать следующим образом.

Рассмотрим неоднородное линейное рекуррентное соотношение

$$a_{n+m} = b_1 \cdot a_{n+m-1} + b_2 \cdot a_{n+m-2} + \ldots + b_m \cdot a_n + u_n, \qquad a_0, \ldots, a_{m-1} \text{ — заданы.}$$

Ранее мы отмечали, что общее решение соответствующего однородного рекуррентного соотношения представляет собой сумму функций $f_j(n)$ вида

$$f_j(n) = (c_{1,j} + c_{2,j} \cdot n + c_{3,j} \cdot n^2 + \ldots + c_{q_j,j} \cdot n^{q_j-1}) \cdot r_j^n.$$

Предположим, что неоднородность $u_n \equiv u(n)$ имеет вид

$$u(n) = (d_0 + d_1 \cdot n + \ldots + d_l \cdot n^l) \cdot r^s =: P_l(n) \cdot r^s.$$

Тогда частное решение исходного неоднородного уравнения следует искать в виде

$$(c_0 + c_1 \cdot n + \ldots + c_l \cdot n^l) \cdot n^q \cdot r^s,$$

где $q = 0$ в случае, если r отличен от корней r_j характеристического уравнения, и равно кратности корня r_j в случае, если r совпадает с одним из корней r_j. В случае, когда неоднородность представляет собой сумму описанных выше функций u_t, частное решение рекуррентного соотношения можно искать в виде суммы частных решений неоднородных рекуррентных соотношений вида

$$a_{n+m} = b_1 \cdot a_{n+m-1} + b_2 \cdot a_{n+m-2} + \ldots + b_m \cdot a_n + u_j(n).$$

5.4.5. Разберем несколько примеров решения неоднородных рекуррентных соотношений.

Пример 5.3. Построить общее решение неоднородного рекуррентного соотношения второго порядка

$$a_{n+2} = 4a_{n+1} - 4a_n + 3 \cdot 2^n.$$

Решение. Характеристическое уравнение для соответствующего однородного рекуррентного соотношения имеет следующий вид:

$$r^2 - 4r + 4 = 0 \qquad \Longleftrightarrow \qquad (r-2)^2 = 0.$$

Так как $r = 2$ есть корень кратности 2, то частное решение исходного неоднородного рекуррентного соотношения следует искать в виде

$$c \cdot n^2 \cdot 2^n.$$

Подставим это выражение в рекуррентное соотношение:

$$c \cdot (n+2)^2 \cdot 2^{n+2} = 4c \cdot (n+1)^2 \cdot 2^{n+1} - 4c \cdot n^2 \cdot 2^n + 3 \cdot 2^n \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \qquad 4c(n^2 + 4n + 4) = 8c(n^2 + 2n + 1) - 4cn^2 + 3 \qquad \Longleftrightarrow \qquad 8c = 3.$$

Следовательно, $c = 3/8$, так что частное решение заданного рекуррентного соотношения имеет вид

$$3 \cdot n^2 \cdot 2^{n-3}.$$

Пример 5.4. Построить общее решение следующего линейного неоднородного рекуррентного соотношения второго порядка:

$$a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot n + 3^n \cdot \cos(\pi\, n/2).$$

Решение. Для построения общего решения соответствующего однородного рекуррентного соотношения запишем его характеристическое уравнение:

$$r^2 = 6r - 9 \qquad \Longleftrightarrow \qquad (r-3)^2 = 0.$$

Как видно, это уравнение имеет корень $r = 3$ кратности два. Следовательно, общее решение однородного рекуррентного соотношения записывается в виде

$$(d_1 + d_2 \cdot n) \cdot 3^n.$$

Неоднородная часть рекуррентного соотношения состоит из двух слагаемых:

$$u_1(n) = 3^n \cdot n \qquad \text{и} \qquad u_2(n) = 3^n \cdot \cos(\pi\, n/2).$$

Так как
$$\cos(\pi\, n/2) = \frac{e^{i\,\pi\, n/2} + e^{-i\,\pi\, n/2}}{2},$$
то выражение для $u_2(n)$ можно переписать так:
$$u_2(n) = \frac{1}{2}\left(r_1^n + r_2^n\right), \qquad \text{где} \qquad r_1 = 3\cdot e^{i\,\pi/2}, \qquad r_2 = 3\cdot e^{-i\,\pi/2}.$$
Иными словами, неоднородная часть рекуррентного соотношения состоит из трех слагаемых вида $r^n \cdot P_i(n)$, и только в первом из них показатель совпадает с корнем характеристического уравнения. Как следствие, мы можем искать отдельно частное решение, связанное с первым слагаемым, и частное решение, связанное с оставшимися слагаемыми.

Рассмотрим вначале рекуррентное соотношение
$$a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot n.$$
Частное его решение следует искать в виде $(c_1 + c_2 \cdot n) \cdot n^2 \cdot 3^n$. Подставляя это выражение в рекуррентное соотношение, получим равенство
$$54n \cdot c_2 + 54 \cdot c_2 + 18 \cdot c_1 = n.$$
Приравнивая коэффициенты при одинаковых степенях n, получаем значения констант $c_2 = 1/54$, $c_1 = -1/18$.

Теперь запишем рекуррентное соотношение
$$a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot \cos(\pi\, n/2) \quad \Longleftrightarrow \quad a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot e^{i\,\pi\, n/2}/2 + 3^n \cdot e^{-i\,\pi\, n/2}/2.$$
Частное его решение можно искать в виде умноженной на 3^n линейной комбинации экспонент с показателями $\pm\, i\,\pi\, n/2$, либо, что более удобно, в виде линейной комбинации синуса и косинуса, умноженной на 3^n:
$$c_3 \cdot 3^n \cdot \cos(\pi\, n/2) + c_4 \cdot 3^n \cdot \sin(\pi\, n/2).$$
Подставляя это выражение в записанное выше рекуррентное соотношение, получаем равенство
$$18 \cdot c_3 \cdot \sin(\pi\, n/2) - (18 \cdot c_4 + 1) \cdot \cos(\pi\, n/2) = 0,$$
из которого следует, что $c_3 = 0$, $c_4 = -1/18$.

Окончательно получаем следующее общее решение исходного рекуррентного соотношения:
$$(d_1 + d_2 \cdot n) \cdot 3^n + (n-3) \cdot n^2 \cdot 3^{n-3}/2 - 3^{n-2} \cdot \sin(\pi\, n/2)/2.$$

6 Производящие функции и формальные степенные ряды

6.1. Понятие производящей функции является, пожалуй, самым важным понятием современной перечислительной комбинаторики. Основная задача данного параграфа состоит в том, чтобы объяснить комбинаторный смысл производящей функции, а также дать по возможности строгое ее определение.

6.1.1. Пусть вначале X есть некоторое *конечное* множество. Первое, что обычно нам хочется сделать — это указать мощность этого множества, то есть подсчитать количество всех элементов в нем. Однако, как правило, нам этой информации оказывается недостаточно — на следующем этапе мы обычно хотим как-то классифицировать элементы этого множества X, то есть разбить множество на блоки, объединенные по тому или иному признаку.

Пример 6.1. Рассмотрим множество всех товаров на складе. Конечно, может быть кому-то и полезно знать общее количество N единиц хранимых на складе товаров. Однако любой кладовщик вам скажет, что значительно полезнее иметь информацию о количестве единиц товара каждой категории в отдельности — например, о количестве бутылок вина, коробок с мылом, пакетов с чаем. При таком подходе мы, конечно же, знаем и общее число N товаров — для его получения нам достаточно сложить количества товаров каждой категории.

Пример 6.2. Как мы уже знаем, общее количество подмножеств n-элементного множества X равно 2^n. Однако, как правило, нас интересует более подробная информация об этих подмножествах, например, сколько существует подмножеств, содержащих ровно k элементов. Число таких подмножеств равно, по определению, биномиальному коэффициенту $\binom{n}{k}$, а суммирование этих коэффициентов по всем $k = 0, \ldots, n$ дает нам 2^n.

6.1.2. 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 тому или иному признаку, и перечислять количество элементов в каждом конкретном блоке.

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.

6.1.3. Заметим также, что именно этот принцип, а именно, принцип разбиения множества перечисляемых нами элементов на блоки с дальнейшим подсчетом элементов в каждом отдельном блоке, мы, по сути, использовали для получения всех основных формул элементарной комбинаторики. Если при этом количество перечисляемых объектов в данном множестве выражалось через количество тех же самых объектов, но в меньших множествах, то мы на выходе получали некоторое рекуррентное соотношение, позволяющее определить количество перечисляемых объектов.

6.1.4. Оказывается, существует очень полезный способ разбиения множества на блоки, который, собственно, и приводит к понятию производящей функции. Именно, припишем всем элементам x рассматриваемого множества X некоторый вес $w(x)$ так, чтобы этот вес совпадал на всех элементах одного и того же блока. Формально сделать это можно с помощью отображения

$$w : X \longrightarrow K$$

множества X в некоторое коммутативное кольцо K, сопоставляющего всем элементам x данного блока B_k одно и то же значение $w(x) = k$, $k \in K$:

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

Заметим, что при таком подходе всему множеству X в кольце K отвечает вполне конкретный элемент

$$\sum_{x\in X} w(x) = \sum_{k\in K} c_k \cdot k, \qquad\qquad c_k = |w^{-1}(k)| = |B_k|\,,$$

называемый *энумератором* множества X.

Пример 6.3. Пусть на складе имеются три пакета с чаем, две бутылки вина и четыре коробки с мылом:

$$X = \{\text{чай}_1, \text{чай}_2, \text{чай}_3, \text{вино}_1, \text{вино}_2, \text{мыло}_1, \text{мыло}_2, \text{мыло}_3, \text{мыло}_4\}.$$

Рассмотрим теперь отображение w этого множества X в коммутативное кольцо K (то есть кольцо, в котором операция умножения коммутативна и ассоциативна; как правило, считают, что это кольцо с единицей, то есть в нем существует нейтральный элемент относительно операции умножения) над множеством, содержащим элементы $\{\text{ч}, \text{в}, \text{м}\}$, задаваемое равенствами

$$w(\text{чай}_i) = \text{ч}, \qquad w(\text{вино}_i) = \text{в}, \qquad w(\text{мыло}_i) = \text{м}.$$

Тогда все множество X разобьется на три блока, а энумератор этого множества примет вид

$$w(X) = 3\text{ч} + 2\text{в} + 4\text{м}.$$

Пример 6.4. Рассмотрим множество 2^X всех подмножеств n-элементного множества X, а также коммутативное кольцо K над множеством, содержащим элементы x и y. Сопоставим любому элементу $Y \subset X$ множества 2^X вес $w = x^k y^{n-k}$, где k есть мощность подмножества Y. Отображение $w : 2^X \to K$ разобъет множество 2^X на блоки B_k мощности $c_k = \binom{n}{k}$. При этом самому множеству 2^X в кольце K будет соответствовать энумератор вида

$$w(2^X) = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n. \tag{34}$$

6.1.5. Как правило, в перечислительной комбинаторике при построении отображения $w : X \to K$ используют несколько стандартных колец. Именно, рассмотрим множество $\mathbb{C}^{\mathbb{Z}_+}$ всех счетных последовательностей $(a_0, a_1, a_2, \ldots)$ комплексных чисел:

$$\mathbb{C}^{\mathbb{Z}_+} := \{(a_0, a_1, a_2, \ldots) \,|\, a_i \in \mathbb{C} \quad \forall\, i \in \mathbb{Z}_+\}.$$

Введем на этом множестве две базовые операции — сложения и умножения. Операция сложения числовых последовательностей определяется обычно как операция поэлементного сложения:

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots). \tag{35}$$

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

$$c_n = \sum_{i=0}^{n} a_i \cdot b_{n-i} \tag{36}$$

и биномиальная свертка

$$c_n = \sum_{i=0}^{n} \binom{n}{i} \cdot a_i \cdot b_{n-i} \tag{37}$$

числовых последовательностей. В обоих случаях множество $\mathbb{C}^{\mathbb{Z}_+}$ с введенными на нем операциями представляет собой коммутативное кольцо (точнее, коммутативную область целостности, то есть кольцо, в котором произведение ненулевых элементов кольца отлично от нуля).

6.1.6. На практике с такими 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 (36). Заметим, что

$$z^2 := z \cdot z = (0, 1, 0, 0, \ldots) \cdot (0, 1, 0, 0, \ldots) = (0 \cdot 0, 0 \cdot 1 + 1 \cdot 0, 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0, \ldots) = (0, 0, 1, 0, \ldots),$$

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

В общем случае, как несложно проверить по индукции,

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

где единица в последнем выражении стоит на n-й позиции, $n = 0, 1, \ldots$. Как следствие, любую последовательность чисел $(a_0, a_1, a_2, \ldots)$ можно записать в виде формального степенного ряда

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

Аналогично, в случае кольца $\mathbb{C}_e[[z]]$ с правилом умножения (37) несложно убедиться в том, что

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

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

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

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

$$w : X \longrightarrow \mathbb{C}[[z]],$$

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 $\mathbb{C}[[z]]$. 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 $z^n = (0, 0, \ldots, 0, 1, 0, \ldots)$. Этот энумератор и называют *обыкновенной производящей функцией* множества X.

Аналогично, рассматривая отображение

$$w_e : X \longrightarrow \mathbb{C}_e[[z]],$$

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.

6.1.8. 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 (35)–(37) 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 (36) и (37) 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.

6.2. Итак, по определению, обыкновенная и экспоненциальная производящие функции являются элементами формальных степенных рядов $\mathbb{C}[[z]]$ и $\mathbb{C}_e[[z]]$. Поговорим немного подробнее о таких формальных степенных рядах.

6.2.1. 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

$$f(z) = \sum_{n=0}^{+\infty} n!\, z^n = 1 + z + 2z^2 + 6z^3 + 24z^4 + \ldots, \tag{38}$$

отвечающий числовой последовательности $(1, 1, 2, 6, 24, \ldots)$. 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 (38) 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.

6.2.2. 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 6.5. Перемножим два формальных степенных ряда из $\mathbb{C}[[z]]$:

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

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,

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

Заметим теперь, что формула

$$1 + \alpha z + \alpha^2 z^2 + \ldots = \frac{1}{1-\alpha z} \qquad \forall\, \alpha \in \mathbb{C} \tag{39}$$

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 $\mathbb{C}[[z]]$ формальных степенных рядов. Однако, несмотря на смысловые отличия, вид самого тождества и в том, и в другом случае одинаков.

6.2.3. Сформулированное выше наблюдение можно использовать для получения различных соотношений, связывающих элементы кольца формальных степенных рядов. Именно, пусть у нас имеются две функции $f(z)$ и $g(z)$ комплексного аргумента, аналитические в некоторой общей для них окрестности точки $z = 0$. Предположим, что в этой окрестности функции удовлетворяют какому-то общему для них соотношению. Представим эти функции в форме сходящихся степенных рядов. Тогда высока вероятность того, что указанное выше соотношение будет справедливым и как некоторое соотношение между формальными степенными рядами, то есть как некоторое соотношение, справедливое для некоторых двух элементов кольца формальных степенных рядов.

С этой точки зрения доказанное в примере 6.5 тождество (39) 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:

$$f(z)\cdot g(z) = 1 \qquad \Longleftrightarrow \qquad \Big(\sum_{n=0}^{+\infty} \alpha^n z^n\Big)\cdot(1-\alpha z) = 1. \tag{40}$$

Все, что теперь остается проверить — это то, что данное равенство останется справедливым, если его рассматривать и как некоторое тождество между формальными степенными рядами в $\mathbb{C}[[z]]$.

Рассмотрим еще несколько примеров из этой же серии.

Пример 6.6. 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:

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

Несложно убедиться, что это равенство оказывается справедливым и как соотношение между двумя элементами кольца $\mathbb{C}_e[[z]]$. Применяя теперь правило (37) умножения в кольце $\mathbb{C}_e[[z]]$, 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 6.7. Пусть $F(z)$, $G(z)$ – пара аналитических функций, связанных в некоторой общей для них окрестности точки $z = 0$ тождеством вида

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

Очевидно, что тогда

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

На языке формальных степенных рядов $F(z), G(z) \in \mathbb{C}_e[[z]]$ эти равенства отвечают формулам обращения

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

Замечание 6.8. Подводя итоги, можно сказать, что описанный выше способ есть хороший способ *получения* разного рода соотношений между формальными степенными рядами. При этом, вообще говоря, это не есть способ их *доказательства* — по идее, мы обязательно должны затем проверить справедливость любого такого тождества, используя только лишь определение операций над формальными степенными рядами в соответствующем кольце формальных степенных рядов. Однако в случае, когда в полученных соотношениях фигурируют только лишь операции сложения и умножения, мы можем сразу сказать, что эти соотношения верны и как соотношения между формальными степенными рядами. Действительно, справедливость таких соотношений сразу следует из того факта, что операции сложения и умножения определены одинаково как в математическом анализе, так и в теории формальных степенных рядов.

6.3. Как уже неоднократно отмечалось, формула (40) 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.

6.3.1. 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:

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

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

Обратно, пусть $f(z)=a_0+a_1z+a_2z^2+\dots$ и $g(z)=b_0+b_1z+b_2z^2+\dots$ — 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 $b_0\neq 0$.

6.3.2. Пусть теперь $h(z)=c_0+c_1z+c_2z^2+\dots$ и $g(z)=b_0+b_1z+b_2z^2+\dots$ — два произвольных элемента кольца $\mathbb{C}[[z]]$, причем $b_0\neq 0$. 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 примере 6.5 ряд

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

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

7 Производящие функции и линейные рекуррентные соотношения

7.1. Ранее мы научились по заданному линейному однородному рекуррентному соотношению с постоянными коэффициентами, связывающему члены a_n числовой последовательности, находить явные выражения для коэффициентов a_n как функций параметра n. Используя эти выражения, мы можем, в принципе, записать решение и в виде формального степенного ряда

$$f(z)=a_0+a_1z+a_2z^2+\dots+a_nz^n+\dots.$$

Однако в большинстве задач такой порядок построения решения не является оптимальным. На практике, как правило, легче по имеющемуся рекуррентному соотношению построить соответствующую данной комбинаторной задаче производящую функцию, используя описанные в предыдущем параграфе операции над формальными степенными рядами. Проиллюстрируем вышесказанное на простом примере, описывающем изменение популяций лягушек в озере.

7.1.1. Напомним, что для этого примера нами было получено следующее линейное неоднородное рекуррентное соотношение первого порядка, описывающее изменение популяции лягушек в озере:

$$a_{n+1}=4\,a_n-100,\qquad n=0,1,2,\dots,\qquad a_0=50. \tag{41}$$

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

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 (41). Домножим (41) на z^{n+1} и просуммируем полученное уравнение по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} a_{n+1}\, z^{n+1} = 4\sum_{n=0}^{+\infty} a_n\, z^{n+1} - 100\sum_{n=0}^{+\infty} z^{n+1}. \tag{42}$$

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 (42), 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 (42) 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 (42) переписывается в виде

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

$$\Longrightarrow \qquad f(z) = \frac{a_0}{1-4z} - \frac{100\, z}{(1-z)(1-4z)}. \tag{43}$$

7.1.2. 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 (43) в ряд по степеням z.

С первым слагаемым в правой части (43) 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:
$$a_n=a_0\,4^n-\frac{100}{3}\,[4^n-1]=\frac{50}{3}\,[4^n+2]\,,\qquad n=0,1,2,\ldots.$$

7.2. Опишем теперь общий алгоритм использования аппарата обыкновенных производящих функций для решения линейных рекуррентных соотношений m-го порядка с постоянными коэффициентами:

(1) ввести обыкновенную производящую функцию $f(z)$ для числовой последовательности $a_0,a_1,a_2,\ldots$;

(2) трансформировать заданное рекуррентное соотношение в уравнение для $f(z)$, домножив это соотношение на z^{n+m}, просуммировав полученное выражение по n от 0 до $+\infty$ и выразив каждую из полученных таким образом сумм через $f(z)$;

(3) разрешить полученное уравнение относительно $f(z)$;

(4) определить числа a_n 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:
$$b_0a_{n+m}+b_1a_{n+m-1}+\ldots+b_ma_n=u_n,\quad n=0,1,2,\ldots,\quad b_0\neq 0,\ a_0,a_1,\ldots,a_{m-1}\text{ – заданы.}\qquad(44)$$

7.2.1. На первом шаге введем для последовательности $a_0,a_1,a_2,\ldots$ обыкновенную производящую функцию
$$f(z)=a_0+a_1\,z+a_2\,z^2+\ldots=\sum_{n=0}^{+\infty}a_nz^n.$$
Домножим (44) на z^{n+m} и просуммируем полученное уравнение по n от 0 до $+\infty$. В полученном

соотношении разберем каждое слагаемое отдельно:

$$b_0 \sum_{n=0}^{+\infty} a_{n+m} z^{n+m} = b_0 \left[f(z) - a_0 - a_1 z - \ldots - a_{m-1} z^{m-1}\right],$$

$$b_1 z \sum_{n=0}^{+\infty} a_{n+m-1} z^{n+m-1} = b_1 z \left[f(z) - a_0 - a_1 z - \ldots - a_{m-2} z^{m-2}\right],$$

$$\ldots$$

$$b_{m-2} z^{m-2} \sum_{n=0}^{+\infty} a_{n+2} z^{n+2} = b_{m-2} z^{m-2} \left[f(z) - a_0 - a_1 z\right],$$

$$b_{m-1} z^{m-1} \sum_{n=0}^{+\infty} a_{n+1} z^{n+1} = b_{m-1} z^{m-1} \left[f(z) - a_0\right],$$

$$b_m z^m \sum_{n=0}^{+\infty} a_n z^n = b_m z^m f(z),$$

$$z^m \sum_{n=0}^{+\infty} u_n z^n =: z^m u(z).$$

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 $f(z)$:

$$f(z) \cdot g(z) = b_0 a_0 + (b_0 a_1 + b_1 a_0) z + \ldots + (b_0 a_{n+m-1} + b_1 a_{n+m-2} + \ldots + b_{m-2} a_1 + b_{m-1} a_0) z^{m-1} + z^m u(z) =$$

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

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

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

7.2.2. 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.

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

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

а затем воспользоваться готовыми формулами вида (**??**).

7.3. Перейдем теперь к решению линейных рекуррентных соотношений с переменными коэффициентами.

7.3.1. 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.

Определение 7.1. Пусть $(a_0, a_1, a_2, \dots)$ — некоторая числовая последовательность, и пусть

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

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

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

Иными словами, производной числовой последовательности $(a_0, a_1, a_2, \dots)$ в кольце с операцией умножения (36) называется числовая последовательность

$$(1 \cdot a_1, 2 \cdot a_2, 3 \cdot a_3, \dots).$$

Определение 7.2. Пусть

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

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

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

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

7.3.2. 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 7.3. 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 $(a_0, 0, 0, \dots)$.

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 7.4. 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:

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

Замечание 7.5. 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.

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

7.4.1. 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

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

7.4.2. 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:

$$(1-4z)\, f'(z) = 2f(z), \qquad f(0) := a_0 = 1. \tag{45}$$

7.4.3. Возникает вопрос, как из этого уравнения определить обыкновенную производящую функцию $f(z)$. Стандартный прием здесь состоит в следующем. Соотношение вида (45) 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 (45) 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:

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

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

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

7.4.4. 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:

$$a_{n+1} = (n+1)\, a_n, \qquad n = 0, 1, 2, \ldots, \qquad a_0 = 1. \tag{46}$$

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

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

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

Полученная задача Коши не имеет решения, аналитического в окрестности начала координат. Этот результат в данном случае легко объясним. Действительно, исходное рекуррентное соотношение (46) 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 ранее, расходится при любых $z > 0$, если рассматривать его с точки зрения обычного математического анализа.

7.4.5. В принципе, полученное нами дифференциальное уравнение можно преобразовать так, чтобы его решение выражалось через гипергеометрические ряды, и построить искомое решение $a_n = n!$ (см., например, [1]). Однако сама процедура получения такого решения оказывается крайне сложной по сравнению со сложностью исходного рекуррентного соотношения. Оказывается, однако, что ситуацию можно подправить, решая это рекуррентное соотношение не с помощью обыкновенных, а с помощью экспоненциальных производящих функций.

Действительно, заметим, что 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.

Именно, введем для числовой последовательности a_n, описываемой рекуррентным соотношением (46), 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 (46) 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.

7.4.6. 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.

7.4.7. 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

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k, \qquad B_0 = 1 \tag{47}$$

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 (47) 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 $z \in \mathbb{C}$.

8 Производящие функции Дирихле. Формулы обращения Мебиуса

8.1. Как уже отмечалось выше, существует достаточно большое количество различных правил умножения числовых последовательностей, из которых (36) и (37) лишь самые популярные. Так, например, в комбинаторных задачах, связанных с циклическими последовательностями, а также в аналитической теории чисел очень часто используется еще одно правило умножения — так называемая свертка Дирихле.

8.1.1. Именно, пусть

$$(a_1, a_2, \ldots) \qquad \text{и} \qquad (b_1, b_2, \ldots)$$

есть пара числовых последовательностей. Их сверткой Дирихле называется числовая последовательность $(c_1, c_2, \ldots)$, коэффициенты c_n которой рассчитываются по формулам

$$c_n = \sum_{d \backslash n} a_d b_{n/d}. \tag{48}$$

Суммирование в этой формуле происходит по всем делителям d целого положительного числа n, то есть по всем таким положительным числам d, для которых существует $k \in \mathbb{Z}$, такое, что выполняется равенство $n = k \cdot d$. Множество $\mathbb{C}^{\mathbb{N}}$ числовых последовательностей $(a_1, a_2, \ldots)$ с введенными на нем операциями сложения (35) и умножения (48) образует коммутативное кольцо.

8.1.2. Поставим в соответствие числовой последовательности $(a_1, a_2, \ldots)$ формальный степенной ряд вида

$$f(z) = \frac{a_1}{1^z} + \frac{a_2}{2^z} + \frac{a_3}{3^z} + \ldots = \sum_{n=1}^{+\infty} \frac{a_n}{n^z}, \tag{49}$$

называемый *рядом Дирихле* числовой последовательности $(a_1, a_2, \ldots)$. С точки зрения математического анализа этот ряд в области $\Re z > \sigma$ задает некоторую аналитическую функцию $f(z)$ комплексного аргумента $z \in \mathbb{C}$. Несложно показать, что коэффициенты c_n у получающейся в результате перемножения двух таких функций

$$f(z) = \frac{a_1}{1^z} + \frac{a_2}{2^z} + \frac{a_3}{3^z} + \ldots \qquad \text{и} \qquad g(z) = \frac{b_1}{1^z} + \frac{b_2}{2^z} + \frac{b_3}{3^z} + \ldots$$

функции $h(z) = \sum\limits_{n=1}^{+\infty} \frac{c_n}{n^z}$ рассчитываются по формулам (48).

Как следствие, ряд вида (49) можно рассматривать как элемент кольца $\mathbb{C}_d[[z]]$ формальных степенный рядов, операции сложения и умножения в котором определяются формулами (35) и (48) соответственно. Нейтральным элементом по умножению в этом кольце является функция $I(z) = 1/1^z = 1$, отвечающая числовой последовательности

$$I_n = \begin{cases} 1, & \text{если } n = 1, \\ 0, & \text{если } n > 1, \end{cases}$$

а любой элемент кольца $f(z) \in \mathbb{C}_d[[z]]$ с ненулевым свободным членом $a_1 \neq 0$ обратим по умножению.

8.1.3. Рассмотрим числовую последовательность $(1, 1, 1, \ldots)$. В кольце $\mathbb{C}_d[[z]]$ такой последовательности отвечает формальный степенной ряд вида

$$\zeta(z) = \frac{1}{1^z} + \frac{1}{2^z} + \frac{1}{3^z} + \ldots,$$

известный как *ζ-функция Римана*. Ее аналогами в кольцах $\mathbb{C}[[z]]$ и $\mathbb{C}_e[[z]]$ являются функции $f(z) = 1/(1-z)$ и $F(z) = e^z$ соответственно.

8.1.4. Функцией Мебиуса $\mu(z)$ называется функция, обратная к ζ-функции по отношению к операции умножения в кольце формальных степенных рядов $\mathbb{C}_d[[z]]$. Это означает, что

$$\mu(z)\cdot\zeta(z)=\zeta(z)\cdot\mu(z)=I(z) \qquad \Longleftrightarrow \qquad \sum_{d\backslash n}\mu_d=\begin{cases}1, & \text{если } n=1,\\ 0, & \text{если } n>1.\end{cases}$$

В упражнении **??** предлагается доказать, что коэффициенты μ_n этой функции рассчитываются по формулам

$$\mu_n=\begin{cases}1, & \text{если } n=1,\\ (-1)^s, & \begin{array}{l}\text{если в каноническом разложении } n \text{ на простые множители } n=p_1^{\alpha_1}\dots p_s^{\alpha_s}\\ \text{все показатели } \alpha_i=1,\end{array}\\ 0, & \text{если в этом разложении существует хотя бы одно } \alpha_i>1.\end{cases} \tag{50}$$

Так, например,

n	1	2	3	4	5	6	7	8	9	10	11
μ_n	1	-1	-1	0	-1	1	-1	0	0	1	-1

8.1.5. Из определения функций $\zeta(z)$ и $\mu(z)$ следует, что если функции $f(z), g(z)\in\mathbb{C}_d[[z]]$ связаны равенством

$$g(z)=\zeta(z)\cdot f(z), \qquad \text{то} \qquad f(z)=\mu(z)\cdot g(z).$$

Этим равенствам отвечают следующие *формулы обращения Мебиуса:*

$$b_n=\sum_{d\backslash n}a_d \qquad \Longrightarrow \qquad a_n=\sum_{d\backslash n}\mu_d\, b_{n/d}. \tag{51}$$

8.2. Несложно проверить следующие свойства отношения делимости на множестве P всех делителей натурального числа n:

(1) Рефлексивность: $n\backslash n$.

(2) Транзитивность: если $d\backslash n$, а $n\backslash m$, то $d\backslash m$.

(3) Антисимметричность: если $d\backslash n$ и $n\backslash d$, то $d=n$.

Эти три свойства означают, что делимость вводит на множестве P *отношение частичного порядка,* превращая это множество в частично упорядоченное множество. Оказывается, что формулы вида (51) можно обобщить для достаточно широкого класса частично упорядоченных множеств.

8.2.1. Напомним определение частично упорядоченного множества.

Определение 8.1. Частично упорядоченным множеством называется множество P с введенным на нем бинарным отношением $\preccurlyeq$ частичного порядка, удовлетворяющим следующим трем свойствам:

(1) Рефлексивность: $\forall x \in P\ x \preccurlyeq x$.

(2) Транзитивность: $\forall x, y, z \in P$: если $x \preccurlyeq y$, а $y \preccurlyeq z$, то $x \preccurlyeq z$.

(3) Антисимметричность: $\forall x, y \in P$: если $x \preccurlyeq y$ и $y \preccurlyeq x$, то $x = y$.

8.2.2. Приведем еще несколько важных примеров частично упорядоченных множеств. Прежде всего, мы можем рассмотреть множество $P = 2^{[n]}$ всех подмножеств n-множества $[n] = \{1, 2, \ldots, n\}$. В упражнении **??** предлагается доказать, что отношение $\subseteq$ вводит на множестве P отношение частичного порядка, т.е. что множество P вместе с отношением $\subseteq$ образует частично упорядоченное множество.

Еще одним важным примером частично упорядоченного множества является множество неотрицательных целых чисел, принадлежащих замкнутому интервалу $[0, n]$, с отношением $\leqslant$ сравнения этих чисел по величине. Более того, это отношение задает на любом таком множестве $[0, n]$ структуру линейно упорядоченного множества L_n.

8.2.3. Введем некоторые полезные вспомогательные понятия.

Определение 8.2. Два элемента x, y частично упорядоченного множества P называются *сравнимыми,* если *либо* $x \preccurlyeq y$, *либо* $y \preccurlyeq x$.

Определение 8.3. Говорят, что элемент y покрывает элемент x, если $x \prec y$ и не существует элемента $z \in P$, такого, что $x \prec z \prec y$.

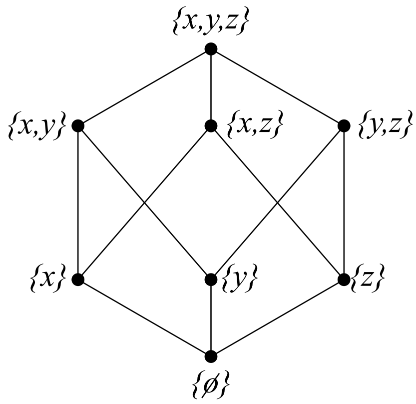

Рис. 8

Определение 8.4. Диаграммой Хассе конечного частично упорядоченного множества P называется граф, вершинам которого отвечают элементы частично упорядоченного множества P. Если $x \prec y$, то вершина, соответствующая y, располагается в диаграмме выше вершины, отвечающей элементу x. Если, кроме того, y покрывает x, то в диаграмме проводится ребро, соединяющее вершины x и y.

На рис.8 в качестве примера показана диаграмма Хассе частично упорядоченного множества B_3.

Рассмотрим, наконец, пару сравнимых между собой элементов $x \preccurlyeq y$ частично упорядоченного множества P.

Определение 8.5. Все элементы z частично упорядоченного множества P, удовлетворяющие неравенствам

$$x \preccurlyeq z \preccurlyeq y,$$

образуют так называемый замкнутый интервал $[x, y]$ между элементами x и y. Если в P любой интервал конечен, то P называется локально конечным частично упорядоченным множеством.

Заметим, что любое конечное частично упорядоченное множество, то есть множество, состоящее из конечного числа элементов, является локально конечным. Еще один пример локально конечного множества — множество всех целых чисел с обычным отношением сравнения $\leqslant$. Далее для простоты мы будем рассматривать конечные частично упорядоченные множества.

8.2.4. Рассмотрим теперь множество $\mathrm{Int}(P)$ всех интервалов произвольного конечного частично упорядоченного множества P. Сопоставив каждому интервалу $I = [x, y]$ некоторое вещественное или комплексное число $f(I) =: f(x, y)$, мы введем на множестве $\mathrm{Int}(P)$ некоторую вещественную или комплексную функцию $f(x, y)$. При этом мы полагаем, что $f(x, y) = 0$ в случае, когда x и y не сравнимы между собой, а также в случае, когда x и y сравнимы, но $y \prec x$. Каждую такую функцию мы можем представить в виде матрицы F, строки и столбцы которой проиндексированы всеми элементами частично упорядоченного множества. Любой элемент этой матрицы, стоящий на пересечении строки x и столбца y, отвечает значению функции $f(x, y)$ на интервале $[x, y]$. Условие $f(x, y) = 0$ в случае, если $y \prec x$, означает, что такая матрица является верхнетреугольной.

В качестве примера рассмотрим частично упорядоченное множество B_2. Пометим строки и столбцы матрицы 4×4 элементами $\varnothing$, $\{1\}$, $\{2\}$ и $\{1, 2\}$. Тогда матрица

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

задает на декартовом произведении $B_2 \times B_2$ функцию $\delta(X, Y)$, такую, что $\delta(X, Y) = 1$ в случае, если $X = Y$, и $\delta(X, Y) = 0$ в противном случае. Еще один пример — это матрица

$$Z = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Она отвечает функции $\zeta(X, Y)$, равной единице в случае, если $X \subseteq Y$.

Верхнетреугольные матрицы мы можем умножать на числа, а также складывать между собой. Кроме того, мы их можем перемножать. Как следствие, мы можем ввести аналогичные операции на множестве всех функций, заданных на множестве $\mathrm{Int}(P)$. В частности, умножение функций f и g, заданных на $\mathrm{Int}(P)$, определяется по формуле

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

Несложно показать, что множество функций на $\mathrm{Int}(P)$ замкнуто относительно этой операции, а само множество вместе с операциями сложения, умножения на число и умножения функций образует алгебру — так называемую *алгебру инцидентности,* изоморфную алгебре верхнетреугольных матриц размерами $n \times n$.

8.2.5. Вернемся к функциям $\delta(x, y)$ и $\zeta(x, y)$, которые мы определили для частично упорядоченного множества B_2. Подобного рода функции можно совершенно аналогичным образом ввести и для любого конечного частично упорядоченного множества, полагая

$$\delta(x,y) = \begin{cases} 1, & \text{если } x = y \\ 0, & \text{если } x \neq y \end{cases}, \qquad \zeta(x,y) = \begin{cases} 1, & \text{если } x \preccurlyeq y \\ 0 & \text{в противном случае} \end{cases}.$$

Функция $\delta(x, y)$ играет роль нейтрального элемента по отношению к операции умножения в алгебре инцидентности и называется *дельта-функцией* частично упорядоченного множества P. Функция $\zeta(x, y)$ называется *дзета-функцией* частично упорядоченного множества и играет важную роль в теории таких множеств. Еще более важную роль играет обратная по отношению к операции умножения функция — так называемая *функция Мёбиуса* $\mu(x, y)$:

$$\mu(x,y) \cdot \zeta(x,y) = \zeta(x,y) \cdot \mu(x,y) = \delta(x,y).$$

Равенство $\mu(x, y) \cdot \zeta(x, y) = \delta(x, y)$ можно переписать так:

$$\sum_{x \preccurlyeq z \preccurlyeq y} \mu(x,z) = \begin{cases} 1, & \text{если } x = y \\ 0 & \text{в противном случае.} \end{cases} \tag{52}$$

Отсюда следует рекуррентная формула, позволяющая последовательно вычислять значения функции Мебиуса:

$$\mu(x,y) = \begin{cases} 1, & \text{если } x = y; \\ -\sum\limits_{x \preccurlyeq z \prec y} \mu(x,z), & \text{если } x \prec y. \end{cases}$$

Заметим, что если мы возьмем за основу равенство $\zeta(x, y) \cdot \mu(x, y) = \delta(x, y)$, то мы получим несколько иную формулу для вычисления функции Мебиуса:

$$\mu(x,y) = \begin{cases} 1, & \text{если } x = y; \\ -\sum\limits_{x \prec z \preccurlyeq y} \mu(z,y), & \text{если } x \prec y. \end{cases}$$

И та, и другая формула позволяют при фиксированном x, стартуя со значения $\mu(x, x) = 1$, последовательно вычислять величины $\mu(x, y)$.

8.2.6. В качестве примера рассмотрим функцию Мебиуса $\mu(x, y)$, обратную к дзета-функции $\zeta(x, y)$ для частично упорядоченного множества $P = B_2$. Значения этой функции можно подсчитать так:

$$\begin{aligned}
&\mu(\varnothing,\varnothing) = 1;\\
&\mu(\varnothing,\{1\}) = -\mu(\varnothing,\varnothing) = -1;\\
&\mu(\varnothing,\{2\}) = -\mu(\varnothing,\varnothing) = -1;\\
&\mu(\varnothing,\{1,2\}) = -\mu(\varnothing,\varnothing) - \mu(\varnothing,\{1\}) - \mu(\varnothing,\{2\}) = -1 - (-1) - (-1) = 1;\\
&\mu(\{1\},\{1\}) = 1;\\
&\mu(\{1\},\{2\}) = 0;\\
&\mu(\{1\},\{1,2\}) = -\mu(\{1\},\{1\}) = -1;\\
&\mu(\{2\},\{2\}) = 1;\\
&\mu(\{2\},\{1,2\}) = -\mu(\{2\},\{2\}) = -1;\\
&\mu(\{1,2\},\{1,2\}) = 1.
\end{aligned}$$

Как следствие, отвечающая этой функции верхнетреугольная матрица M имеет вид

$$Z = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

В упражнении **??** предлагается показать, что в частично упорядоченном множестве B_n для любых двух подмножеств S и T, таких, что $S \subseteq T$, функция Мебиуса

$$\mu(S,T) = (-1)^{|T-S|}.$$

В случае линейно упорядоченного множества L_n функция Мебиуса равна

$$\mu(x,y) = \begin{cases} 1, & \text{если } x = y; \\ -1, & \text{если } x = y - 1; \\ 0 & \text{в противном случае.} \end{cases}$$

Наконец, в случае частично упорядоченного множества всех положительных делителей натурального числа n функция Мебиуса определяется по формуле

$$\mu(x,y) = \begin{cases} 1, & \text{если } x = y; \\ (-1)^s, & \text{если } y/x \text{ есть произведение } s \text{ различных простых чисел;} \\ 0 & \text{в противном случае.} \end{cases}$$

Как следствие, $\mu(x,y) = \mu(y/x)$, где $\mu(z)$ есть классическая функция Мебиуса одной переменной в теории чисел.

8.3. Предположим теперь, что помимо функций $f(x,y)$, заданных на множестве $\mathrm{Int}(P)$, у нас имеются и какие-то функции $f\colon P \to \mathbb{R}$ или $f\colon P \to \mathbb{C}$, заданные на самом частично упорядоченном множестве P. Оказывается, с помощью функций $\zeta(x,y)$ и $\mu(x,y)$ мы для функций $f(x)$ и $g(x)$ можем записать соотношения, аналогичные формулам обращения (51).

8.3.1. Именно, справедлива следующая

Теорема 8.6 (Формула обращения Мебиуса)**.** *Пусть P есть конечное частично упорядоченное множество. Предположим, что на множестве элементов такого частично упорядоченного множества P задана пара вещественных или комплексных функций f и g, значения которых связаны равенством*

$$g(y) = \sum_{x \preccurlyeq y} f(x). \tag{53}$$

Тогда для всех $y \in P$

$$f(y) = \sum_{x \preccurlyeq y} g(x) \cdot \mu(x,y). \tag{54}$$

Доказательство. Введем вектор-строки $\mathbf{f}$ и $\mathbf{g}$, коэффициентами которых являются значения функций f и g. Пусть Z и M есть отвечающие функциям $\zeta(x,y)$ и $\mu(x,y)$ верхнетреугольные матрицы. Тогда первое равенство можно переписать в матричном виде следующим образом:

$$\mathbf{g} = \mathbf{f} \cdot Z.$$

Домножая левую и правую часть этого равенства справа на матрицу M и учитывая, что $Z \times \times M = I$, получаем, что

$$\mathbf{f} = \mathbf{g} \cdot M.$$

8.3.2. Посмотрим, что означают формулы обращения (53)–(54) для конкретных частично упорядоченных множеств. Начнем с линейно упорядоченного множества L целых неотрицательных чисел, принадлежащих отрезку $[0, n]$. Сопоставим каждому значению числа i из этого отрезка некоторое вещественное или комплексное число $f(i) =: f_i$. Иными словами, зададим на множестве целых неотрицательных чисел из этого отрезка некоторую вещественную или комплексную функцию $f(i)$. Предположим, что у нас имеются две такие функции f и g, значения которых связаны равенством

$$g_n = \sum_{i=0}^{n} f_i.$$

Тогда, как мы уже знаем, коэффициенты f_n можно выразить через g_i по формулам

$$f_n = g_n - g_{n-1}.$$

Если теперь вспомнить вид дзета-функции $\zeta(x, y)$ и функции Мебиуса $\mu(x, y)$ на таком частично упорядоченном множестве, то последние формулы мы можем переписать так:

$$g(n) = \sum_{i \leqslant n} f(i) \cdot \zeta(i, n) \qquad \Longleftrightarrow \qquad \mathbf{g} = \mathbf{f} \cdot Z,$$

$$f(i) = \sum_{i \leqslant n} g(i) \cdot \mu(i, n) \qquad \Longleftrightarrow \qquad \mathbf{f} = \mathbf{g} \cdot M,$$

где $\mathbf{f}$, $\mathbf{g}$ — вектор-строки, отвечающие последовательностям f_n и g_n.

Заметим, что формулы, связывающие f_n и g_n, есть не что иное, как конечно-разностный аналог основной теоремы анализа: операторы $\sum$ и Δ взаимно обратны.

8.3.3. Рассмотрим теперь функции $f\colon B_n \to \mathbb{C}$, заданные на всевозможных подмножествах множества $[n]$. Пусть пара таких функций связана соотношением

$$g(T) = \sum_{S \subseteq T} f(S).$$

Тогда значения функции f можно выразить через значения функции g по формуле

$$f(T) = \sum_{S \subseteq T} (-1)^{|T-S|} \cdot g(S).$$

Последняя формула есть не что иное, как обобщенный принцип включения-исключения (см.домашнее задание).

8.3.4. Наконец, вспомним формулы обращения (51). Эти формулы допускают следующую трактовку. Рассмотрим множество всех делителей d числа n. Зададим на этом множестве некоторые вещественные или комплексные функции. Предположим, что значения функций f и g связаны равенством

$$g(n) = \sum_{d \backslash n} f(d).$$

Тогда значения функции $f(n)$ можно вычислить через значения функции $g(d)$ по формулам

$$f(n) = \sum_{d \backslash n} \mu_{n/d} \cdot g(d) = \sum_{d \backslash n} g(d) \cdot \mu(d, n).$$

8.3.5. Заметим теперь, что каждая из описанных выше формул обращения имеет некоторый аналог на языке формальных степенных рядов. Рассмотрим, к примеру, обыкновенные производящие функции $f(z)$ и $g(z)$, связанные равенствами

$$g(z) = \frac{f(z)}{1-z} = f(z) \cdot (1 + z + z^2 + \ldots + z^n + \ldots) \qquad \Longleftrightarrow \qquad f(z) = g(z) \cdot (1 - z).$$

Отсюда с учетом формул (36) умножения таких производящих функций мы получаем, что коэффициенты a_n и b_n функций $f(z)$ и $g(z)$ связаны соотношениями

$$b_n = \sum_{i=0}^{n} a_i \qquad \Longleftrightarrow \qquad a_n = b_n - b_{n-1}.$$

А это есть не что иное, как формулы обращения Мебиуса для линейно упорядоченного множества L_n. При этом функция $1/(1-z)$ представляет собой, по сути, ζ-функцию, а функция $1-z$ — функцию Мебиуса.

Теперь рассмотрим пару экспоненциальных производящих функций $F(z)$ и $G(z)$, связанных равенствами

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

Отсюда с учетом (37) получаем следующие соотношения на коэффициенты a_n и b_n функций $F(z)$ и $G(z0)$:

$$g_n = \sum_{i=0}^{n} \binom{n}{i} a_i \qquad \Longleftrightarrow \qquad a_n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} a_i.$$

А это есть, в свою очередь, аналог формул обращения для случая частично упорядоченного множества B_n. Иными словами, в рассматриваемом случае аналогами ζ-функции и функции Мебиуса являются экспоненциальные производящие функции e^z и e^{-z}.

Наконец, мы уже отмечали, что классические функции $\zeta(z)$ и $\mu(z)$ в теории чисел являются одномерными аналогами соответствующих функций для частично упорядоченного множества всех делителей натурального числа n.

Таким образом, разные виды сверток числовых последовательностей, а следовательно, и разные виды производящих функций, связаны с разными способами частично упорядочить заданное n-элементное множество. Так, интуитивно кажется, что линейному порядку на $[n]$ отвечает обыкновенная производящая функция, отношению $\subseteq$ соответствует экспоненциальная производящая функция, а отношению делимости $\backslash$ отвечает производящая функция Дирихле. Оказывается, эти интуитивные наблюдения можно формализовать, введя понятие так называемых биномиальных частично упорядоченных множеств и формальных степенных рядов на этих множествах. За подробностями отсылаем заинтересованных читателей к монографии [3].

8.4. В случае частично упорядоченных множеств P общего вида вычисление функции Мебиуса $\mu(x, y)$ представляет собой достаточно тяжелую задачу. Существует, однако, крайне важный с практической точки зрения класс частично упорядоченных множеств, для которых функция

Мебиуса вычисляется достаточно просто. Этот класс частично упорядоченных множеств носит название решеток.

8.4.1. Прежде чем вводить понятие решетки, напомним, что для любого конечного частично упорядоченного множества мы можем ввести два важных понятия — максимального элемента и элемента, наибольшего по включению.

Определение 8.7. Элемент $x \in P$ называется *максимальным элементом* частично упорядоченного множества, если во-первых, этот элемент сравним со всеми остальными элементами частично упорядоченного множества, а во-вторых, для любых элементов $z \in P$ выполняется отношение $z \preccurlyeq x$. Элемент y называется *наибольшим по включению элементом* частично упорядоченного множества, если не существует элементов $z \in P$, таких, что $y \prec z$.

Аналогично вводятся понятия минимального элемента и элемента, наименьшего по включению.

Важно отметить, что далеко не в любом конечном частично упорядоченном множестве P существует максимальный (минимальный) элемент. При этом в случае, когда такой элемент существует, он обязательно является единственным. Наибольшие (наименьшие) по включению элементы в *конечном* частично упорядоченном множестве существуют всегда, и их всегда несколько штук в случае, если минимальный элемент в P отсутствует.

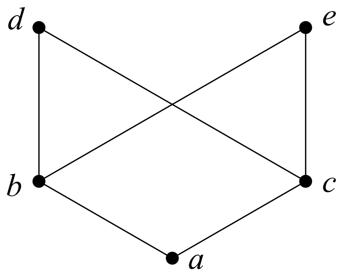

Рис. 9

Так, для частично упорядоченного множества, изображенного на рис.9, элемент a является единственным минимальным элементом. Максимального элемента в данном частично упорядоченном множестве не существует, а наибольших по включению элементов в этом множестве две штуки — это элементы d и e.

8.4.2. Для определения решетки нам понадобятся понятия, в чем-то аналогичные понятиям максимального и минимального элементов. Именно, если у нас имеется какое-то подмножество S частично упорядоченного множества P, то для него мы можем ввести понятие *верхней грани* — элемента $x \in P$, такого, что $z \preccurlyeq x$ для всех $z \in S$. Эта верхняя граница называется *наименьшей верхней гранью* (или супремумом, или точной верхней гранью), если для любого $y \in P$, такого, что $z \preccurlyeq y$ для всех $z \in S$, выполняется условие $x \preccurlyeq y$. Аналогично определяется понятие наибольшей нижней грани (или инфимума, или точной нижней грани) частично упорядоченного множества P.

Определение 8.8. Решеткой L называется частично упорядоченное множество, в котором любое подмножество $S \subseteq L$ имеет как наименьшую верхнюю, так и наибольшую нижнюю грани.

8.4.3. Заметим, что наличие наименьшей верхней (наибольшей нижней) грани достаточно проверять для произвольной пары элементов x, y частично упорядоченного множества. Действительно, пусть любая такая пара будет иметь наибольшую нижнюю и наименьшую верхнюю грани. Тогда по индукции несложно доказать, что и любое подмножество также будет иметь соответствующие грани.

Наименьшая верхняя грань двух элементов x и y обозначается через $x \vee y$ и называется в англоязычной литературе join, а наибольшая нижняя грань обозначается через $x \wedge y$ и называется meet.

В качестве примера решетки рассмотрим частично упорядоченное множество B_n всех подмножеств n-элементного множества $[n]$ (см.рис.8). В таком частично упорядоченном множестве для любых двух подмножеств$S \subseteq [n]$ и $T \subseteq [n]$ наименьшей верхней гранью является подмножество $S \cup T$, а наибольшей нижней гранью — подмножество $S \cap T$. Напротив, частично упорядоченное множество, показанное на рис.9, решеткой не является — элементы b и c имеют две верхние грани, ни одна из которых не является наименьшей.

8.4.4. Для проверки того, что частично упорядоченное множество является решеткой, нам необходимо убедиться в том, что это любые два элемента этого множества имеют как наименьшую верхнюю грань, так и наибольшую нижнюю грань. Существует, однако, один важный частный случай частично упорядоченного множества, для которого нам достаточно проверить существование лишь одной из этих двух граней.

Лемма 8.9. *Конечное частично упорядоченное множество L, в котором существует максимальный элемент $\widehat{1}$, и в котором для любых двух элементов существует наибольшая нижняя грань, является решеткой.*

Доказательство. Пусть x, y есть два произвольных элемента частично упорядоченного множества L. Обозначим через B множество всех верхних граней этих двух элементов. Подмножество $B = \{b_1, \ldots, b_k\} \subseteq L$ конечно и непусто — максимальный элемент $\widehat{1}$ множества L ему гарантированно принадлежит. Лемма будет доказана, если мы докажем, что в B имеется единственный минимальный элемент — в этом случае он и будет единственной наименьшей верхней гранью $x \vee y$ элементов x и y. Рассмотрим для этого элемент

$$b = b_1 \wedge b_2 \wedge \ldots \wedge b_k$$

и покажем, что b принадлежит подмножеству B. Действительно, раз $x \preccurlyeq b_1$ и $x \preccurlyeq b_2$, то x представляет собой какую-то нижнюю грань элементов b_1 и b_2. Но $b_1 \wedge b_2$ представляет собой наибольшую нижнюю грань этих двух элементов, так что $x \preccurlyeq b_1 \wedge b_2$. Аналогично показывается, что и $y \preccurlyeq b_1 \wedge b_2$. Как следствие, $b_1 \wedge b_2 \in B$. Далее по индукции можно показать, что и $b \in B$. Кроме того, b является минимальным элементом подмножества B — просто потому, что $b \preccurlyeq b_i$ для любого $i = 1, \ldots, k$. Следовательно, $b = x \vee y$. □

8.4.5. Доказанная лемма полезна в случае, когда доказать существование наибольшей нижней грани двух элементов x и y частично упорядоченного множества P легче, чем существование наименьшей верхней грани элементов x и y.

В качестве примера рассмотрим множество Π_n всех разбиений множества $[n]$ на блоки. Его можно превратить в частично упорядоченное множество, введя так называемое *отношение измельчения* $\preccurlyeq$. Именно, пусть α и β есть пара произвольных разбиений множества $[n]$ на блоки. Говорят, что $\alpha \preccurlyeq \beta$, если любой блок разбиения β представляет собой объединение нескольких блоков разбиения α. Например,

$$\alpha = \{1, 2\}, \{3\}, \{4\}, \{5\}, \{6\} \preccurlyeq \beta = \{1, 2, 3\}, \{4\}, \{5, 6\}.$$

На рис.10 в качестве примера показана диаграмма Хассе частично упорядоченного множества Π_3.

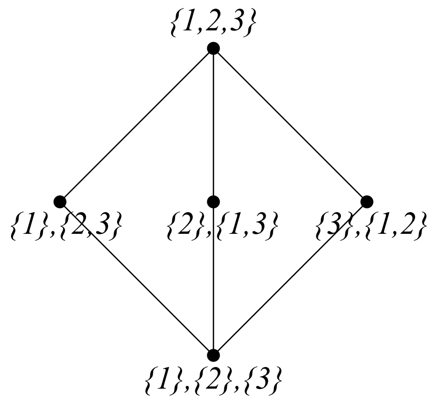

Рис. 10

Покажем, что частично упорядоченное множество Π_n всех разбиений множества $[n]$ на блоки вместе с отношением измельчения является решеткой. Очевидно, что Π_n имеет максимальный элемент — разбиение, содержащее единственный блок $[n]$. Рассмотрим теперь произвольную пару α и β разбиений множества $[n]$. Сопоставим этой паре разбиение γ по следующему принципу: поместим элементы i и j в один и тот же блок разбиения γ тогда и только тогда, когда они являются элементами одного и того же блока как в α, так и в β. В противном случае поместим эти элементы в отдельные блоки $\{i\}$, $\{j\}$. Например, если $\alpha = \{1\}, \{2, 3, 4\}$, а $\beta = \{1, 2\}, \{3, 4\}$, то $\gamma = \{1\}, \{2\}, \{3, 4\}$. Ясно, что $\gamma \preccurlyeq \alpha$ и $\gamma \preccurlyeq \beta$. Кроме того, для любой нижней грани δ разбиений α и β справедливо соотношение $\delta \preccurlyeq \gamma$. Следовательно, $\gamma = \alpha \wedge \beta$, и согласно доказанной лемме, Π_n является решеткой.

8.4.6. Одно из достоинств решеток состоит в том, что для таких частично упорядоченных множеств значительно проще вычисляются функции Мебиуса.

Теорема 8.10 (Weisner, 1935). *В случае решетки L с минимальным элементом $\widehat{0}$ и максимальным элементом $\widehat{1}$ функция Мебиуса*

$$\mu(\widehat{0}, \widehat{1}) = -\sum_{\substack{x:\ x \wedge a = \widehat{0} \\ x \neq \widehat{0}}} \mu(x, \widehat{1}), \tag{55}$$

где a — произвольный элемент множества $L - \{\widehat{1}\}$.

Напомним, что в случае обыкновенного частично упорядоченного множества для подсчета $\mu(\widehat{0}, \widehat{1})$ мы должны сосчитать сумму $n - 1$ значений этой функции в других точках. В случае же решетки L нам достаточно сосчитать сумму значений этой функции в точках, наибольшая нижняя грань с a которых равна $\widehat{0}$. Выбирая в качестве a достаточно большой элемент, мы вероятнее всего получим не так много элементов $x \in L$, для которых $x \wedge a = \widehat{0}$, и тем самым уменьшим количество вычислений.

Пример 8.11. Покажем, что для решетки Π_n функция Мебиуса

$$\mu(\widehat{0}, \widehat{1}) = (-1)^{n-1}(n-1)!$$

Выберем в качестве элемента a разбиение $\alpha = \{1, 2, \ldots, n-1\}, \{n\}$ и найдем все разбиения β, такие, что $\beta \wedge \alpha = \widehat{0}$. В любом таком разбиении β числа i и j не должны входить в блок $\{1, 2, \ldots, n-1\}$. Это возможно тогда и только тогда, когда разбиение β состоит из блока $\{i, n\}$, а также из $(n-2)$-х одноэлементных блоков вида $\{j\}$.

Нам осталось сосчитать для любого такого β функцию Мебиуса $\mu(\beta, \widehat{1})$, $\widehat{1} = \{1, \ldots, n\}$. Заметим для этого, что любой интервал $[\beta, \widehat{1}]$ изоморфен решетке Π_{n-1}. Действительно, если $\{i, n\}$ представляет собой единственный нетривиальный блок в разбиении β, то элементы i и n должны входить в один и тот же блок любого разбиения γ, такого, что $\beta \preccurlyeq \gamma$. Удаляя из всех таких разбиений n, мы получим все возможные разбиения множества $[n-1]$.

Теперь воспользуемся теоремой 8.10:

$$\mu_{\Pi_n}(\widehat{0}, \widehat{1}) = -\sum_{\substack{\beta:\ \beta\wedge\alpha=\widehat{0} \\ \beta\neq\widehat{0}}} \mu(\beta, \widehat{1}) = -(n-1)\cdot\mu_{\Pi_{n-1}}(\widehat{0}, \widehat{1}) = \ldots = (-1)^{n-1}\cdot(n-1)!$$

Замечание 8.12. Мы сказали, что интервал $[\beta, \widehat{1}]$ изоморфен Π_{n-1}. Оказывается, что этот результат можно обобщить. Именно, если β есть разбиение множества $[n]$ на k блоков, то интервал $[\beta, \widehat{1}]$ изоморфен Π_k. Действительно, если два числа i, j входят в один и тот же блок β, то они должны входить в один и тот же блок любого разбиения γ, такого, что $\beta \preccurlyeq \gamma$. Следовательно, в частично упорядоченном множестве $[\beta, \widehat{1}]$ блоки β играют роль отдельных элементов, а само это множество изоморфно Π_k.

Пример 8.13. Пусть разбиение $\beta = \{1\}, \{2, 3\}, \{4\} \in \Pi_4$. Элементы γ, такие, что $\beta \preccurlyeq \gamma$ и $\gamma \preccurlyeq \widehat{1}$, $\widehat{1} = \{1, 2, 3, 4\}$, представляют собой разбиения $\gamma_1 = \{1, 4\}, \{2, 3\}$, $\gamma_2 = \{1\}, \{2, 3, 4\}$ и $\gamma_3 = \{1, 2, 3\}, \{4\}$. Тогда интервал

$$[\beta, \widehat{1}] = (\{1\}, \{2, 3\}, \{4\}), (\{1, 4\}, \{2, 3\}), (\{1\}, \{2, 3, 4\}), (\{1, 2, 3\}, \{4\}), (\{1, 2, 3, 4\})$$

изоморфен частично упорядоченному множеству

$$\Pi_3 = (\{1\}, \{2\}, \{3\}), (\{1, 3\}, \{2\}), (\{1\}, \{2, 3\}), (\{1, 2\}, \{3\}), (\{1, 2, 3\}).$$

8.4.7. Прежде чем доказывать теорему 8.10, приведем крайне важный с практической точки зрения пример использования формулы обращения Мебиуса для решетки Π_n. Мы в свое время достаточно легко подсчитали количество всех простых графов, построенных на множестве $[n]$ вершин — это количество равно $2^{\binom{n}{2}}$. Теперь мы покажем, как с помощью (51) искать количество простых *связных* графов, построенных на множестве $[n]$.

Заметим, прежде всего, что любому простому графу G отвечает некоторое разбиение α множества его вершин на блоки, каждый из которых представляет собой подмножество вершин, отвечающих связной компоненте графа G. Пусть теперь α есть некоторое произвольное разбиение множества $[n]$, блоки которого имеют размеры $c_1, \ldots, c_k$. Ясно, что одному и тому же разбиению соответствует какой-то набор простых графов G. Нам хочется научиться подсчитывать количество $f(\alpha)$ всех таких графов. Однако в лоб сосчитать $f(\alpha)$ достаточно сложно. Значительно проще подсчитать количество $g(\alpha)$ графов G_β, соответствующих всем разбиениям β, таким, что $\beta \preccurlyeq \alpha$. Действительно, у таких графов отсутствуют ребра между вершинами, отвечающими различным блокам разбиения α. Поэтому любой граф G_β представляет собой несвязное объединение k простых, не обязательно связных графов, построенных на соответствующих подмножествах вершин. А количество таких графов сосчитать легко — количество способов построить простой подграф на c_i вершинах равно $2^{\binom{c_i}{2}}$, так что общее количество $g(\alpha)$ графов G_β, отвечающих всевозможным разбиениям $\beta \preccurlyeq \alpha$, равно $2^{c_1} \cdot 2^{c_2} \cdot \ldots \cdot 2^{c_r}$.

Заметим теперь, что $g(\alpha)$ мы можем выразить через функцию $f(\alpha)$ по формуле

$$g(\alpha) = \sum_{\beta \preccurlyeq \alpha} f(\beta).$$

Отсюда по формуле (51) обращения Мебиуса мы получаем, что

$$f(\alpha) = \sum_{\beta \preccurlyeq \alpha} g(\beta) \cdot \mu(\beta, \alpha).$$

Мы хотим подсчитать количество всех связных графов. Все такие графы отвечают разбиению α, состоящему из единственного блока $\widehat{1}$, совпадающего со всем множеством $[n]$. Следовательно, количество связных простых графов рассчитывается по формуле

$$f(\widehat{1}) = \sum_{\beta \in \Pi_n} g(\beta) \cdot \mu(\beta, \widehat{1}),$$

и все, что нам осталось — это научиться вычислять функцию Мебиуса вида $\mu(\beta, \widehat{1})$ для произвольного разбиения $\beta \in \Pi_n$. А это сделать довольно легко — действительно, мы сказали, что любой интервал $[\beta, \widehat{1}]$ изоморфен Π_k, где k — количество блоков в разбиении β. Поэтому $\mu(\beta, \widehat{1}) = (-1)^k \cdot (k-1)!$, и количество $f(\widehat{1})$ всех связных графов на n вершинах рассчитывается по формуле

$$f(\widehat{1}) = \sum_{\beta \in \Pi_n} 2^{\binom{c_1}{2} \cdot \ldots \cdot \binom{c_k}{2}} \cdot (-1)^k \cdot (k-1)!,$$

где k есть количество блоков разбиения β, размеры которых равны $c_1, \ldots, c_k$.

8.4.8. Перейдем теперь к доказательству теоремы 8.10. Заметим, прежде всего, что так как $\widehat{0} \wedge a = \widehat{0}$, то сформулированное в теореме равенство можно переписать так:

$$\sum_{x\colon\, x \wedge a = \widehat{0}} \mu(x, \widehat{1}) = 0.$$

Введем функцию $f(x)$, равную единице в случае, когда $x \wedge a = \widehat{0}$, и нулю во всех остальных случаях (характеристическая функция). Тогда последнее равенство эквивалентно равенству

$$\sum_{x \in L} \mu(x, \widehat{1}) \cdot f(x) = 0.$$

Теперь заметим, что функцию $f(x)$ мы можем переписать в следующем виде:

$$f(x) = \sum_{y \in [\widehat{0}, x \wedge a]} \mu(\widehat{0}, y).$$

Действительно, согласно определению функции Мебиуса (52), стоящая в правой части равенства сумма равна единице в случае, если интервал $[\widehat{0}, x \wedge a]$ тривиален, то есть в случае, когда $x \wedge a = \widehat{0}$, и нулю в противном случае. Подставляя это выражение для функции $f(x)$ в доказываемое равенство, получим, что

$$\sum_{x \in L} \mu(x, \widehat{1}) \sum_{y \in [\widehat{0}, x \wedge a]} \mu(\widehat{0}, y) = 0.$$

Но неравенство $y \preccurlyeq x \wedge a$ выполняется тогда и только тогда, когда $y \preccurlyeq x$ и $y \preccurlyeq a$. Поэтому в последнем равенстве мы можем поменять порядок суммирования, записав его так:

$$\sum_{y \preccurlyeq a} \mu(\widehat{0}, y) \sum_{x \in [y, \widehat{1}]} \mu(x, \widehat{1}) = 0.$$

А такое равенство действительно справедливо — так как $y \preccurlyeq a \prec \widehat{1}$, то интервал $[y, \widehat{1}]$ не пуст, и потому внутренняя сумма в полученном выражении всегда равна нулю. $\square$

9 Числа Каталана. Нелинейные рекуррентные соотношения

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

9.1.1. 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.

Определение 9.1. 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.

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

$$\begin{array}{lll} n=1: & () & \text{– 1 последовательность;} \\ n=2: & ()(),\ (()) & \text{– 2 последовательности;} \\ n=3: & ()()(),\ ()(()),\ (())(),\ (()()),\ ((())) & \text{– 5 последовательностей.} \end{array}$$

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

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 [4] приведено порядка 100 задач, в которых эти числа появляются. Приведем лишь несколько наиболее важных из них.

9.1.2. Рассмотрим вначале очень простую интерпретацию правильной скобочной последовательности — задачу об очереди в кассу (см. [5]). 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.

9.1.3. 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

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

9.1.4. 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* (рис 11). Биекция с правильными скобочными последовательностями здесь такова:

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

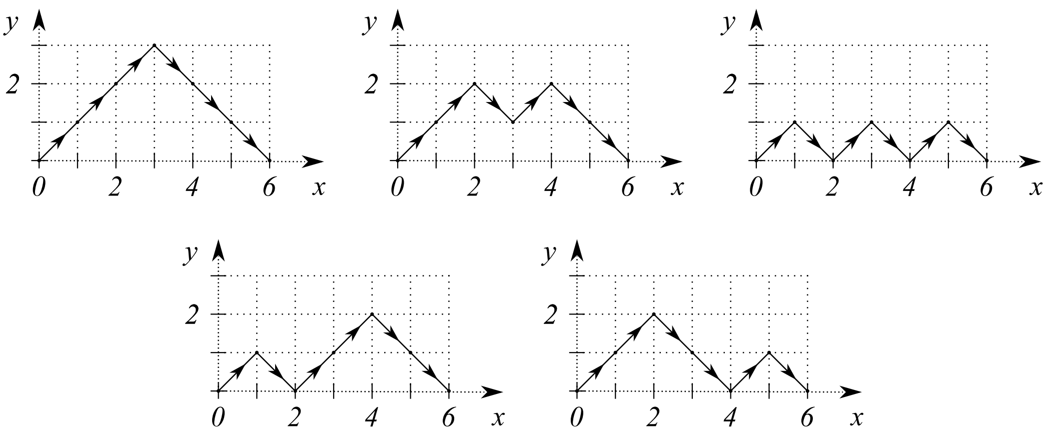

Рис. 11: Все пути Дика из трёх пар шагов.

9.1.5. Одной из основных дискретных структур, использующихся в теории алгоритмов, является плоское корневое бинарное дерево, иногда называющаяся просто бинарным или двоичным деревом. Неформально корневое дерево — это дерево, в котором любая вершина имеет ровно двух (возможно, пустых) потомков — левого и правого (рис. 12). В упражнении **??** предлагается доказать, что количество таких деревьев на n вершинах описывается числами Каталана C_n.

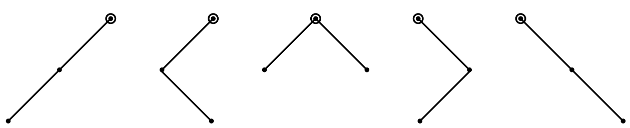

Рис. 12: Все бинарные деревья на трёх вершинах.

9.1.6. Рассмотрим выражение вида

$$a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_n \cdot a_{n+1}, \tag{56}$$

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. Несложно убедиться, что количество расстановки скобок в описанном выше выражении есть число Каталана C_n. Так, для случая $n = 3$ мы имеем следующие 5 различных вариантов расстановки скобок:

$$((a_1 \cdot a_2) \cdot a_3) \cdot a_4, \quad (a_1 \cdot (a_2 \cdot a_3)) \cdot a_4, \quad (a_1 \cdot a_2) \cdot (a_3 \cdot a_4), \quad a_1 \cdot ((a_2 \cdot a_3) \cdot a_4), \quad a_1 \cdot (a_2 \cdot (a_3 \cdot a_4)).$$

9.1.7. 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 (рис. 13).

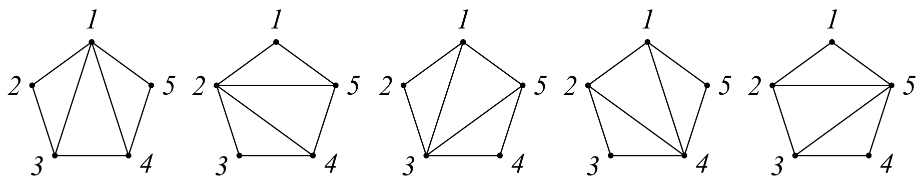

Рис. 13: Все триангуляции пятиугольника.

В упражнении **??** предлагается 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.

9.1.8. Рассмотрим теперь все плоские корневые деревья (рис 14), 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.

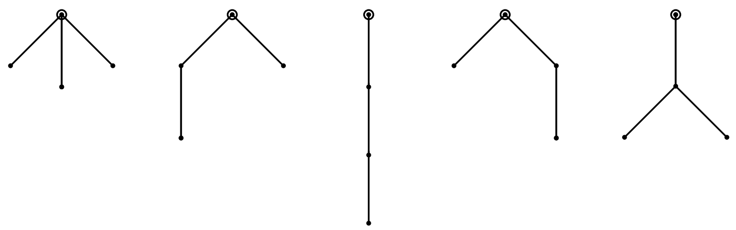

Рис. 14: Все плоские деревья на четырёх вершинах.

9.1.9. Наконец, перейдем к еще одному полезному понятию — стек-сортируемой последовательности. Согласно Кнуту [6], 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:

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

В случае $n = 3$ имеется пять стек-сортируемых последовательностей:

$$(1, 2, 3), \qquad (3, 2, 1), \qquad (2, 1, 3), \qquad (1, 3, 2), \qquad (3, 1, 2).$$

Единственной перестановкой, которую нельзя перевести в последовательность $(1, 2, 3)$ с помощью стека, является перестановка вида $(2, 3, 1)$.

9.2. Теперь, после стольких примеров, пора научиться вычислять числа Каталана. Начнем с вывода рекуррентного соотношения для этих чисел. В качестве основного объекта мы 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.

9.2.1. 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.

9.2.2. 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:

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

9.2.3. 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.

9.2.4. 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

$$C_n = \sum_{k=1}^{n} C_{k-1}\, C_{n-k}, \qquad n = 1, 2, \ldots; \qquad C_0 = 1. \tag{57}$$

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.

9.3. Постараемся решить рекуррентное соотношение (57).

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

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

Нам будет удобно переписать рекуррентное соотношение (57) в следующем виде:

$$C_{n+1} = C_0 \cdot C_n + C_1 \cdot C_{n-1} + \ldots + C_n \cdot C_0 = \sum_{k=0}^{n} C_k C_{n-k}.$$

Домножим это рекуррентное соотношение на z^{n+1} и просуммируем полученное равенство по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} C_{n+1} z^{n+1} = \sum_{n=0}^{+\infty} z^{n+1} \Big(\sum_{k=0}^{n} C_k\, C_{n-k} \Big) \quad \Longleftrightarrow \quad f(z) - 1 = z \Bigg(\sum_{n=0}^{+\infty} z^n \Big(\sum_{k=0}^{n} C_k\, C_{n-k} \Big) \Bigg) = z \cdot f^2(z).$$

Следовательно, функция $f(z)$ определяется из следующего равенства:

$$f(z) = 1 + z\, f^2(z). \tag{58}$$

9.3.2. Как видно, нелинейное рекуррентное соотношение (57) для чисел C_n приводит к нелинейному же уравнению (58) 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 (58). 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 (57).

Итак, все, что нам остается сделать — это найти функцию $f(z)$, аналитическую в окрестности точки $z = 0$ и удовлетворяющую уравнению (58), а затем разложить ее по степеням z^n в окрестности этой точки. Для этого разрешим уравнение (58) 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

$$f(z) = \frac{1 - \sqrt{1-4z}}{2z} = \frac{1}{2z} - \frac{1}{2z}\,(1-4z)^{1/2} \tag{59}$$

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

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \qquad n = 0, 1, 2, \ldots \tag{60}$$

9.3.3. Явная аналитическая формула (60) для чисел Каталана, переписанная в виде

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

допускает следующее элегантное комбинаторное доказательство, принадлежащее французскому математику A.André (1878). Общее количество всех возможных расстановок n открывающихся и n закрывающихся скобок в строке из $2n$ символов равно, очевидно,

$$P(2n; n, n) = \frac{(2n)!}{n! \cdot n!} = \binom{2n}{n},$$

где $P(2n; n, n)$ — количество перестановок с повторениями в строке длины $2n$ n неразличимых символов первого сорта и n неразличимых символов второго сорта (см. параграф 5 главы 1). Осталось из этого количества вычесть "неправильные" скобочные последовательности, т.е. такие, в которых нарушается условие "при проходе строки слева направо количество открывающихся скобок больше или равно количеству закрывающихся скобок".

Для этого рассмотрим любую такую "неправильную" скобочную последовательность. Найдем в этой строке первую закрывающуюся скобку, в которой условие "правильности" нарушается. Эта скобка будет стоять на $(2k+1)$-й позиции для некоторого $k = 0, 1, 2, \dots,$ а слева от нее в подстроке длины $2k$ количество открывающихся скобок, равное k, в точности равно количеству закрывающихся скобок. Теперь заменим в получившейся подстроке длины $(2k+1)$ все открывающиеся скобки закрывающимися и наоборот. В результате получим некоторую строку длины $2n$, содержащую ровно $(n+1)$ открывающуюся скобку и ровно $(n-1)$ закрывающуюся.

Теперь возьмем *любую* строку длины $2n$, содержащую $(n+1)$ открывающуюся скобку и $(n - 1)$ закрывающуюся. Оказывается, ее всегда можно превратить в *неправильную* скобочную последовательность обратным преобразованием. Действительно, будем идти вдоль такой строки слева направо и проверять условие "количество закрывающихся скобок больше или равно количеству открывающихся скобок". Так как общее количество открывающихся скобок в такой строке строго больше общего количества закрывающихся скобок, то обязательно найдется открывающаяся скобка, для которой это условие нарушится. Она будет стоять на $(2k+1)$-й позиции, а слева от нее будет стоять подстрока длины $2k$, в которой количество закрывающихся скобок (равное k) равно количеству открывающихся скобок. Меняя в такой подстроке открывающиеся скобки на закрывающиеся и наоборот, мы получаем строку, состоящую из n открывающихся и n закрывающихся скобок, в которой на $(2k+1)$-й позиции нарушается условие правильности скобочной последовательности.

Итак, мы установили взаимно-однозначное соответствие между множеством всех неправильных скобочных последовательностей и множеством всех строк длины $2n$, содержащих $(n+1)$ открывающуюся скобку и $(n-1)$ закрывающуюся скобку. Мощность последнего множества равна, очевидно,

$$P(2n; n+1, n-1) = \frac{(2n)!}{(n+1)! \cdot (n-1)!} = \binom{2n}{n+1}.$$

Согласно принципу биекции, этому же числу равна и мощность множества всех неправильных скобочных последовательностей длины $2n$. Как следствие, количество всех правильных скобочных последовательностей равно

$$P(2n; n, n) - P(2n; n+1, n-1) = \binom{2n}{n} - \binom{2n}{n+1} = \frac{1}{n+1}\binom{2n}{n} = C_n,$$

что и требовалось доказать.

10 Мультииндексные рекуррентные соотношения

10.1. Ранее мы встречали многочисленные примеры рекуррентных соотношений, в которых искомые числовые последовательности зависели более чем от одного индекса. Самый известный и простой пример такого рода — это рекуррентное соотношение для биномиальных коэффициентов

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \qquad k \geqslant 1, \quad n \geqslant 1;$$

$$\binom{n}{k} = 0 \quad \text{при} \quad k > n; \qquad\qquad \binom{n}{0} = 1 \quad \forall\, n \geqslant 0.$$

Аналогичные рекуррентные соотношения встречались у нас и для чисел $\left(\!\binom{n}{k}\!\right)$ с повторениями, и для чисел $S(n,k)$ Стирлинга второго рода, и для других важных числовых последовательностей. Задача этого параграфа — научиться решать такие соотношения, т.е. находить явное выражение искомых чисел как функций параметров n и k, с помощью производящих функций.

10.2. Кажется очевидным, что двухпараметрическим числовым последовательностям должны соответствовать производящие функции, зависящие от двух переменных. Дадим строгие определения.

Определение 10.1. (Обыкновенной) производящей функцией для числовой последовательности $a_{n,k}$ называется формальный степенной ряд $f(z,t) \in \mathbb{C}[[z,t]]$ вида

$$f(z,t) = a_{0,0} + a_{1,0}t + a_{0,1}z + a_{2,0}t^2 + a_{1,1}zt + a_{0,2}z^2 + a_{3,0}t^3 + a_{2,1}zt^2 + a_{1,2}z^2t + a_{0,3}z^3 + \ldots +$$

$$+a_{n,k}z^k t^n + \ldots = \sum_{n,k=0}^{+\infty} a_{n,k} z^k t^n.$$

Определение 10.2. Экспоненциальной производящей функцией для числовой последовательности $a_{n,k}$ называется формальный степенной ряд $F(z,t) \in \mathbb{C}_{ee}[[z,t]]$ вида

$$FF(z,t) = a_{0,0} + a_{1,0}\frac{t}{1!} + a_{0,1}\frac{z}{1!} + a_{2,0}\frac{t^2}{2!} + a_{1,1}\frac{z}{1!}\frac{t}{1!} + a_{0,2}\frac{z^2}{2!} + a_{3,0}\frac{t^3}{3!} + a_{2,1}\frac{z}{1!}\frac{t^2}{2!} + a_{1,2}\frac{z^2}{2!}\frac{t}{1!} + a_{0,3}\frac{z^3}{3!} + \ldots +$$

$$+a_{n,k}\frac{z^k}{k!}\frac{t^n}{n!} + \ldots = \sum_{n,k=0}^{+\infty} a_{n,k}\frac{z^k}{k!}\frac{t^n}{n!}.$$

Наконец, наряду с этими двумя производящими функциями в комбинаторике широко используется и так называемая полуэкспоненциальная производящая функция, а именно, формальный степенной ряд $F(z,t) \in \mathbb{C}_e[[z,t]]$ вида

$$F(z,t) = a_{0,0} + a_{1,0}\frac{t}{1!} + a_{0,1}z + a_{2,0}\frac{t^2}{2!} + a_{1,1}z\frac{t}{1!} + a_{0,2}z^2 + a_{3,0}\frac{t^3}{3!} + a_{2,1}z\frac{t^2}{2!} + a_{1,2}z^2\frac{t}{1!} + a_{0,3}z^3 + \ldots +$$

$$+a_{n,k}z^k\frac{t^n}{n!} + \ldots = \sum_{n,k=0}^{+\infty} a_{n,k}z^k\frac{t^n}{n!}.$$

Очень часто числовые последовательности $a_{n,k}$ являются треугольными, т.е. обладают свойством $a_{n,k} = 0$ при $k > n$. Производящие функции для таких последовательностей можно записать в несколько более простом виде. Так, для полуэкспоненциальных функций имеем

$$F(z,t) = a_{0,0} + (a_{1,0} + a_{1,1}z)\frac{t}{1!} + (a_{2,0} + a_{2,1}z + a_{2,2}z^2)\frac{t^2}{2!} + \ldots +$$

$$+(a_{n,0} + a_{n,1}z + \ldots + a_{n,k}z^k + \ldots + a_{n,n}z^n)\frac{t^n}{n!} + \ldots = \sum_{n=0}^{+\infty} P_n(z)\frac{t^n}{n!}.$$

10.3. Перейдем теперь к алгоритмам поиска решений двухпараметрических рекуррентных соотношений с помощью таких функций. Опишем, как обычно, эти алгоритмы вначале на простейшем примере, а именно, на примере рекуррентного соотношения для биномиальных коэффициентов $\binom{n}{k} \equiv c_{n,k}$.

10.3.1. Числовая последовательность $c_{n,k}$, как мы знаем, удовлетворяет следующему рекуррентному соотношению:

$$c_{n+1,k+1} = c_{n,k} + c_{n,k+1}, \qquad c_{n,0} = 1 \quad \forall\, n \geqslant 0, \qquad c_{n,k} = 0 \quad \forall\, k > n. \tag{61}$$

Будем решать эту задачу поэтапно. Именно, на первом этапе, считая n параметром, мы попытаемся получить рекуррентное соотношение на производящие функции $P_n(z)$. На втором этапе мы введем производящие функции

$$w(z,t) = \sum_{n=0}^{+\infty} P_n(z)\, t^n \qquad \text{или} \qquad W(z,t) = \sum_{n=0}^{+\infty} P_n(z)\frac{t^n}{n!}$$

и постараемся получить для них явные аналитические выражения.

10.3.2. Начнем с получения уравнения на функции $P_n(z)$. В силу условия $c_{n,k} = 0$ при $k > n$ эти функции представляют собой полиномы вида

$$P_n(z) = c_{n,0} + c_{n,1}z + c_{n,2}z^2 + \ldots + c_{n,n}z^n.$$

Для получения уравнения на $P_n(z)$ домножим (61) на z^{k+1} и просуммируем его по k от нуля до $n+1$:

$$\sum_{k=0}^{n+1} c_{n+1,k+1}z^{k+1} = \sum_{k=0}^{n+1} c_{n,k}z^{k+1} + \sum_{k=0}^{n+1} c_{n,k+1}z^{k+1}.$$

В левой части этого равенства имеем полином вида

$$c_{n+1,1}z + c_{n+1,2}z^2 + \ldots + c_{n+1,n+1}z^{n+1} = P_{n+1}(z) - c_{n+1,0} = P_{n+1}(z) - 1.$$

Правую часть того же равенства с учетом граничных условий можно переписать так:

$$\sum_{k=0}^{n} c_{n,k}z^{k+1} + \sum_{k=0}^{n} c_{n,k+1}z^{k+1} = z\, P_n(z) + P_n(z) - 1.$$

Таким образом, полиномы $P_n(z)$ удовлетворяют рекуррентному соотношению

$$P_{n+1}(z) = (1+z)\, P_n(z). \tag{62}$$

В данном частном случае рекуррентное соотношение (62) на $P_n(z)$ решается легко: видно, что $P_n(z) = (1+z)^n$, откуда сразу следует, что $c_{n,k} = \binom{n}{k}$. В более сложных случаях нам следует перейти ко второму этапу.

10.3.3. Именно, домножая в нашем примере уравнение (62) на t^{n+1} и суммируя по n, получаем

$$\sum_{n=0}^{+\infty} P_{n+1}(z)t^{n+1} = (1+z)t\sum_{n=0}^{+\infty} P_n(z)t^n$$

В левой части этого равенства стоит выражение, равное

$$w(z,t) - P_0(z) = w(z,t) - 1.$$

Правая часть того же равенства, очевидно, равна $(1+z)tw(z,t)$. Следовательно, обыкновенная производящая функция для чисел $c_{n,k}$ имеет вид

$$w(z,t) = \frac{1}{1-(1+z)\,t} = \sum_{n=0}^{+\infty}(1+z)^n t^n = \sum_{n=0}^{+\infty}\sum_{k=0}^{n}\binom{n}{k} z^k t^n.$$

10.3.4. Для получения полуэкспоненциальной производящей функции $W(z,t)$ домножим (62) на $t^n/n!$ и просуммируем результат по n:

$$\sum_{n=0}^{+\infty} P_{n+1}(z)\frac{t^n}{n!} = (1+z)\sum_{n=0}^{+\infty} P_n(z)\frac{t^n}{n!}.$$

Левая часть этого равенства представляет собой производную функции $W(z,t)$ по t. Правая же его часть равна, очевидно, $(1+z)\,W(z,t)$. Следовательно, при таком подходе мы получаем дифференциальное уравнение

$$\frac{\partial W}{\partial t} = (1+z)\,W$$

на функцию $W(z,t)$ по переменной t, зависящее от z как от параметра. Добавляя к нему начальное условие $W(z,0) = P_0(z) = c_{0,0} = 1$, получаем задачу Коши, решение которой, очевидно, имеет вид

$$W(z,t) = e^{t(1+z)}.$$

Разложение этого решения в малой окрестности точки $t = 0$ дает нам искомый результат:

$$W(z,t) = \sum_{n=0}^{+\infty}(1+z)^n\frac{t^n}{n!}.$$

10.4. Давайте теперь разберем чуть более содержательную задачу — перечисление так называемых путей Дика (Dyck pahts) на плоскости.

10.4.1. По определению, путь Дика представляет собой путь, выходящий из начала координат, составленный из векторов $(1,1)$ и $(1,-1)$, и заканчивающийся в точке с координатами (n,k). В случае $k = 0$ мы получаем пути Дика, перечисляемые числами Каталана.

Обозначим через $d_{n,k}$ количество различных путей Дика, заканчивающихся в точке с координатами (n,k). Эти числа отличны от нуля только в случае, если сумма $n+k$ является четным числом (см.рис.15). Наша задача — построить производящую функцию для этих чисел.

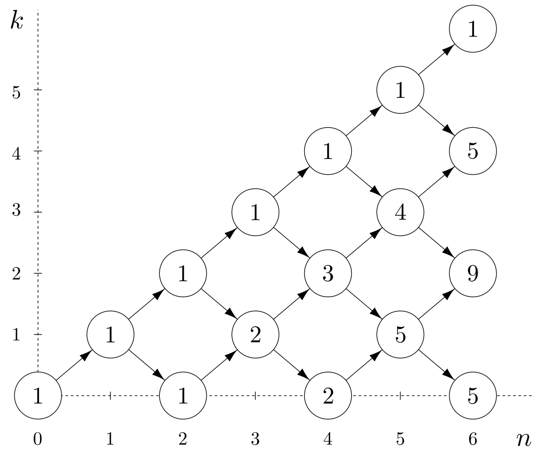

Рис. 15

10.4.2. Запишем для чисел $d_{n,k}$ рекуррентное соотношение:

$$d_{n+1,k} = d_{n,k-1} + d_{n,k+1}, \qquad d_{n,0} = C_n, \quad d_{n,k} = 0, \quad \text{при } k > n \text{ и } k < 0.$$

Здесь C_n — это числа Каталана. Домножая это соотношение на z^k и суммируя по k от нуля до $n+1$, получим

$$\sum_{k=0}^{n+1} d_{n+1,k} z^k = \sum_{k=0}^{n+1} d_{n,k-1} z^k + \sum_{k=0}^{n+1} d_{n,k+1} z^k.$$

Введем полиномы $P_n(z)$ вида

$$P_n(z) = d_{n,0} + d_{n,1} \cdot z + d_{n,2} \cdot z^2 + \ldots + d_{n,n} \cdot d^n, \qquad n = 0, 1, \ldots, +\infty.$$

Тогда в левой части полученного выше равенства будет стоять полином

$$d_{n+1,0} + d_{n+1,1} z + d_{n+1,2} z^2 + \ldots + d_{n+1,n+1} z^{n+1} = P_{n+1}(z).$$

Правую часть того же равенства с учетом граничных условий можно переписать так:

$$\sum_{k=1}^{n+1} d_{n,k-1} z^k + \sum_{k=0}^{n-1} d_{n,k+1} z^k = z\, P_n(z) + \frac{P_n(z) - d_{n,0}}{z} = z\, P_n(z) + \frac{P_n(z) - C_n}{z}.$$

Таким образом, для полиномов $P_n(z)$ мы получили такое уравнение:

$$P_{n+1}(z) = z\, P_n(z) + \frac{P_n(z) - C_n}{z}, \qquad P_0(z) = 1.$$

10.4.3. Для решения уравнения на $P_n(z)$ введем обыкновенную производящую функцию

$$w(z,t) = \sum_{n=0}^{+\infty} P_n(z) \cdot t^n.$$

Домножая уравнение для $P_n(z)$ на t^{n+1} и суммируя по n, получаем

$$w(z,t) - w(z,0) = z\, t\, w(z,t) + \frac{t}{z}\,(w(z,t) - f(t)),$$

где $f(t)$ — обыкновенная производящая функция для чисел Каталана C_n. Так как $w(z,0) = d_{0,0} = 1$, то

$$w(z,t) = \frac{1 - \dfrac{t}{z} f(t)}{1 - t\left(z + \dfrac{1}{z}\right)} = \frac{1 - \dfrac{1 - \sqrt{1-4t^2}}{2zt}}{1 - t\left(z + \dfrac{1}{z}\right)}.$$

11 Принцип инволюции

11.1. Мы уже встречались с достаточно большим количеством задач, решение которых получалось с помощью принципа биекции, а также с помощью принципа, который мы называли принципом "плохой – хороший". Данный параграф посвящен обобщению этих двух принципов — принципу инволюции.

11.1.1. Предположим, что у нас имеется какое-то множество S, которое мы можем разбить на два блока, которые мы обозначим через S^+ и S^-. Допустим также, что нам требуется подсчитать разность $|S^+| - |S^-|$ мощностей этих двух блоков. В случае, когда мощности $|S^+|$ и $|S^-|$ примерно равны друг другу, мы можем эту задачу несколько упростить, введя на множестве S так называемую *альтернированную инволюцию.* Напомним, что инволюцией называется функция φ, обратная самой себе:

$$\varphi(\varphi(x)) = x \qquad\qquad \forall\ x \in S.$$

Такая функция разбивает каждый из блоков S^+, S^- на два подблока — подблоки $\mathrm{Fix}_\varphi \mathrm{S}^+$ и $\mathrm{Fix}_\varphi \mathrm{S}^-$ элементов множества S, остающихся неподвижными под действием инволюции φ, а также подблоки $\mathrm{Var}_\varphi \mathrm{S}^+$ и $\mathrm{Var}_\varphi \mathrm{S}^-$ точек множества S, меняющихся под действием φ.

Определение 11.1. Инволюция $\varphi\colon S \to S$ называется альтернированной, если она переводит любой элемент $x \in \mathrm{Var}_\varphi \mathrm{S}^+$ в какой-то элемент $\varphi(x) \in \mathrm{Var}_\varphi \mathrm{S}^-$ и наоборот.

Заметим теперь, что альтернированная инволюция φ устанавливает взаимно-однозначное соответствие между элементами подмножеств $\mathrm{Var}_\varphi \mathrm{S}^+$ и $\mathrm{Var}_\varphi \mathrm{S}^-$. Как следствие, мощности этих подмножеств совпадают, так что

$$|S^+| - |S^-| = |\mathrm{Fix}_\varphi \mathrm{S}^+| - |\mathrm{Fix}_\varphi \mathrm{S}^-|. \tag{63}$$

Полученное равенство и называется принципом инволюции.

11.1.2. Рассмотрим важный частный случай, при котором $\mathrm{Fix}_\tau \mathrm{S}^- = \varnothing$. В этом случае равенство

$$|S^+| - |S^-| = |\mathrm{Fix}_\varphi \mathrm{S}^+| \tag{64}$$

можно положить в основу перечисления элементов некоторого множества X. Именно, нам следует вложить X в какое-то более широкое множество $S = S^+ \cup S^-$ так, чтобы $X \subseteq S^+$, а затем найти альтернированную инволюцию φ, такую, что

$$\mathrm{Fix}_\varphi \mathrm{S}^- = \varnothing, \qquad\qquad \mathrm{Fix}_\varphi \mathrm{S}^+ = \mathrm{X}.$$

Тогда мы получаем равенство

$$|X| = |S^+| - |S^-|,$$

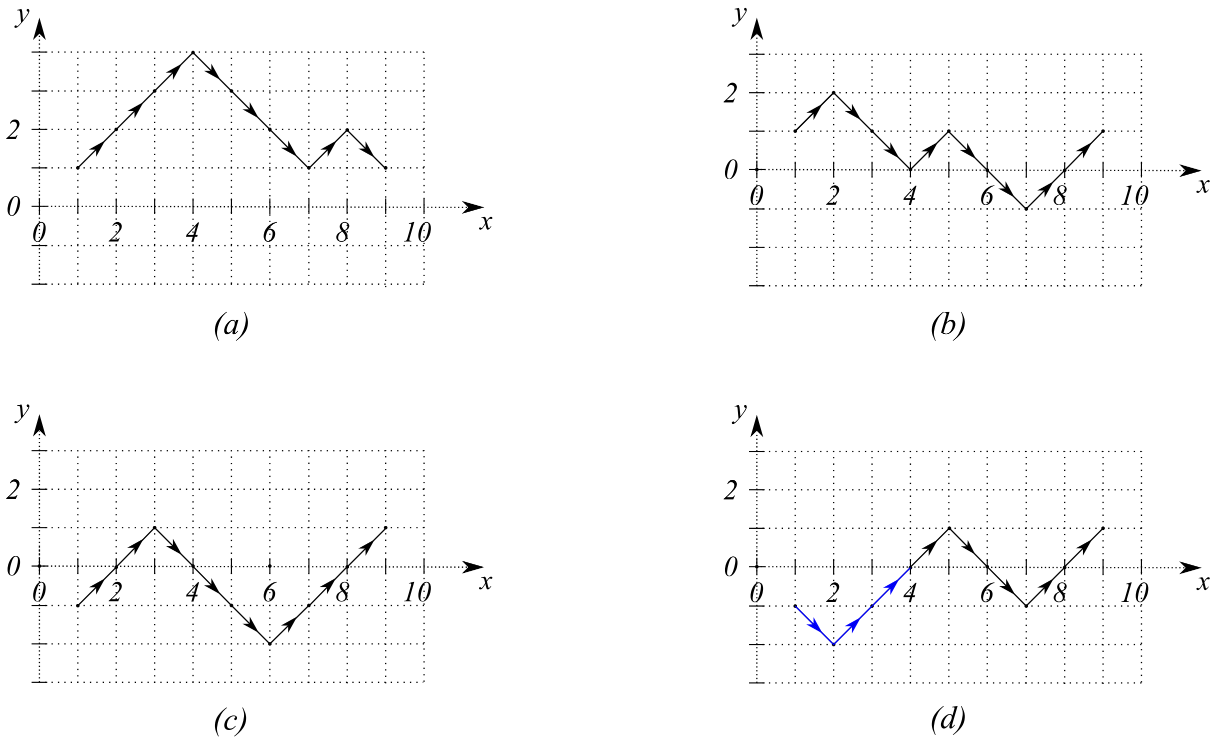

Рис. 16

которое, по сути, и представляет собой формальное описание принципа “плохой-хороший”.

11.1.3. В качестве характерного примера давайте еще раз вернемся к явной формуле (**??**) для чисел Каталана C_n, полученной в конце параграфа, посвященного этим числам. Оказывается, что комбинаторный вывод этой формулы, предложенный André, как раз и использует принцип инволюции.

Давайте для наглядности будем вместо правильных скобочных последовательностей рассматривать пути Дика на плоскости (n, k), начинающиеся в точке с координатами $(1, 1)$, заканчивающиеся в точке с координатами $(2n + 1, 1)$ и не опускающиеся ниже прямой $x = 1$. координат и заканчивающиеся в точках $(2n, 0)$ на оси абсцисс (рис.16,a). Расширим множество X, добавив к нему пути на плоскости с началом в точке $(1, 1)$ и концом в точке $(2n + 1, 1)$, которые могут пересекать прямую $x = 1$ (рис.16,b). Такое множество путей обозначим через S^+. Кроме того, введем в рассмотрение множество S^- путей с началом в точке $(1, -1)$ и с концом в точке $(2n + 1, 1)$, состоящих из отрезков $(1, 1)$ и $(1, -1)$ (рис.16,c). Зададим теперь на множестве $S = S^+ \cup S^-$ инволюцию φ, определив для любого пути $w \in S$ новый путь $\varphi(w)$ как путь, не отличающийся от w на участке от точки $(x, 0)$ первого касания пути w оси абсцисс до последней точки пути, и представляющий собой путь $\varphi(w)$, полученный из w отражением относительно оси абсцисс на участке от точки $(1, 1)$ или $(1, -1)$ до точки $(x, 0)$ первого касания пути w оси абсцисс (см. синий участок пути на рис.16,d, полученный отражением соответствующего участка пути, изображенного на рис.16,b).

Достаточно очевидно, что φ является альтернированной инволюцией. Более того, $\mathrm{Fix}_\tau \mathrm{S}^- = \varnothing$, так как любой путь из точки $(1, -1)$ обязательно пересечет ось абсцисс. Наконец, заметим, что количество путей из $\mathrm{Fix}_\tau \mathrm{S}^+$ в точности совпадает с количеством $|X|$ классических путей Дика.

Тогда, согласно принципу инволюции (64),

$$C_n = |X| = |S^+| - |S^-| = \frac{(2n)!}{n!\,n!} - \frac{(2n)!}{(n-1)!\,(n+1)!} = \binom{2n}{n} - \binom{2n}{n-1}.$$

11.1.4. Принцип инволюции можно обобщить на случай так называемого взвешенного множества S, то есть множества, любому элементу $x \in S$ которого приписан некоторый вес $w(x)$. Самому множеству S также можно приписать вес, равный сумме весов его элементов:

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

Определение 11.2. Говорят, что альтернированная инволюция $\varphi\colon S \to S$ сохраняет вес, если $w(\varphi(x)) = w(x)$ для всех $x \in S$.

Для случая альтернированной инволюции, сохраняющей вес, принцип инволюции (63) можно переписать так:

$$w(S^+) - w(S^-) = w(\mathrm{Fix}_\varphi \mathrm{S}^+) - \mathrm{w}(\mathrm{Fix}_\varphi \mathrm{S}^-). \tag{65}$$

Полагая, в частности, вес $w(x) = 1$ для любого $x \in S$, мы вновь получаем классический принцип инволюции (63).

Покажем, как использовать обобщенный принцип инволюции (65) для доказательства тождества Вандермонта

$$\det \begin{pmatrix} x_1^{n-1} & x_2^{n-1} & \dots & x_{n-1}^{n-1} & x_n^{n-1} \\ x_1^{n-2} & x_2^{n-2} & \dots & x_{n-1}^{n-2} & x_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots & \\ x_1 & x_2 & \dots & x_{n-1} & x_n \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} = \prod_{1 \leqslant i < j \leqslant n} (x_i - x_j), \tag{66}$$

играющего важную роль в линейной алгебре, теории интерполяции и дискретной математике. В частности, с его помощью можно доказать единственность так называемой задачи интерполяции функции $f(x)$ полиномом $P_{n-1}(x)$: если нам известны значения функции $f(x)$ в точках $x_1, \dots, x_n$, то для нахождения полинома $P_{n-1}(x)$, такого, что

$$P_{n-1}(x_i) = a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_{n-1} x_i^{n-1} = f(x_i), \qquad i = 1, \dots, n,$$

мы имеем систему линейных алгебраических уравнений, определителем которой является определитель Вандермонда. Так как он, согласно (66), отличен от нуля при любых попарно различных значениях x_i, то задача интерполяции всегда имеет единственное решение.

11.1.5. В качестве примера посмотрим, чему равны левая и правая часть равенства (66) в случае $n = 3$. В левой его части мы получаем выражение

$$\det \begin{pmatrix} x_1^2 & x_2^2 & x_3^2 \\ x_1 & x_2 & x_3 \\ 1 & 1 & 1 \end{pmatrix} = x_1^2 x_2 - x_2^2 x_1 + x_3^2 x_1 - x_1^2 x_3 + x_2^2 x_3 - x_3^2 x_2,$$

а в правой — полином

$$(x_1 - x_2) \cdot (x_1 - x_3) \cdot (x_2 - x_3) = (x_1^2 - x_1 x_3 - x_1 x_2 + x_2 x_3) \cdot (x_2 - x_3) =$$

$$= x_1^2x_2 - x_2^2x_1 + x_3^2x_1 - x_1^2x_3 + x_2^2x_3 - x_3^2x_2 - x_1x_2x_3 + x_1x_2x_3,$$

который после сокращения двух последних слагаемых в точности совпадает с левой частью равенства (66).

11.1.6. Для доказательства (66) отметим, что левая часть этого равенства содержит $n!$ слагаемых вида

$$\mathrm{sign}(\sigma) \cdot x_{\sigma(1)}^{n-1} \cdot x_{\sigma(2)}^{n-2} \cdot \ldots x_{\sigma(n)}^{0},$$

где σ — одна из $n!$ перестановок множества $[n]$. Подсчитаем количество слагаемых в правой части (66), полученных после перемножения сомножителей $(x_i - x_j)$, а заодно дадим и комбинаторную интерпретацию каждого из таких слагаемых.

Заметим для этого, что в правой части (66) стоит $\binom{n}{2}$ сомножителей — мы берем все двуэлементные подмножества $\{i, j\}$ множества $[n]$ и линейно упорядочиваем индексы в каждом из этих подмножеств для того, чтобы получить выражение вида $x_i - x_j$, $1 \leqslant i < j \leqslant n$. Для получения конкретного слагаемого в правой части (66) мы должны в каждом из сомножителей выбрать либо x_i, либо $-x_j$, а затем перемножить выбранные $\binom{n}{2}$ сомножителей между собой. Сопоставим любому такому слагаемому s турнир, построенный на n вершинах по следующему правилу: если в выражении $x_i - x_j$ мы выбрали в качестве сомножителя x_i, то проведем в турнире ребро из вершины i в вершину j; в противном случае направим ребро из вершины j в вершину i.

(a) Турнир (b) Транзитивный турнир

Рис. 17: Турнир и транзитивный турнир

Пример 11.3. Рассмотрим в качестве примера случай $n = 5$. Выбирая в выражении

$$(x_1 - \underline{x_2}) \cdot (x_1 - \underline{x_3}) \cdot (\underline{x_1} - x_4) \cdot (\underline{x_1} - x_5) \cdot (x_2 - \underline{x_3}) \cdot (x_2 - \underline{x_4}) \cdot (\underline{x_2} - x_5) \cdot (x_3 - \underline{x_4}) \cdot (x_3 - \underline{x_5}) \cdot (x_4 - \underline{x_5})$$

подчеркнутые снизу сомножители, мы получаем слагаемое

$$-x_1^2x_2^2x_3^2x_4^2x_5^2$$

в правой части (66). Ему отвечает турнир, изображенный на рис.17,а.

Так как имеется всего $2^{\binom{n}{2}}$ турниров, построенных на n вершинах, то и в правой части (66) имеется ровно $2^{\binom{n}{2}}$ слагаемых.

11.1.7. Теперь попытаемся дать комбинаторную интерпретацию слагаемых, стоящих в левой части равенства (66). Для этого нам понадобятся некоторые факты из теории графов, касающиеся транзитивных турниров.

Напомним, что турнир называется транзитивным, если из существования ориентированных ребер (i, j) и (j, k) следует существование ребра (i, k). Основное свойство транзитивного турнира состоит в том, что он является ациклическим орграфом.

Утверждение 11.4. *Турнир T является транзитивным тогда и только тогда, когда в T отсутствуют циклы.*

Доказательство. Пусть T является ациклическим графом. Предположим, что $(x, y) \in E(T)$ и $(x, z) \in E(T)$. Так как T ациклический, то $(z, x) \notin E(T)$. Но так как T есть ориентированный полный граф, то в T должно существовать ребро между x и z, а именно, ребро (x, z).

Обратно, пусть в T содержится цикл $C_k = (x_1, \dots, x_k, x_1)$. Так как T транзитивен, то из условий $(x_1, x_2) \in E(T)$ и $(x_2, x_3) \in E(T)$ следует, что $(x_1, x_3) \in E(T)$. Продолжая далее, мы получим, что $(x_1, x_k) \in E(T)$, что невозможно, так как $(x_k, x_1) \in E(T)$. $\square$

Утверждение 11.5. *Турнир T является транзитивным тогда и только тогда, когда все исходящие степени его вершин различны.*

Доказательство. Предположим вначале, что T является транзитивным турниром. Рассмотрим произвольное ребро $(x, y) \in E(T)$. Обозначим через $N(y)$ множество вершин $z \in V(T)$, в которые приходят ребра из вершины y. Так как турнир транзитивен, то из условий $(x, y) \in E(T)$ и $(y, z) \in E(T)$ следует, что $(x, z) \in E(T)$. Следовательно, все вершины из $N(y)$ содержатся в подмножестве вершин $N(x)$. Но в $N(x)$ содержится еще и вершина y, так что $\operatorname{outdeg}(x) = |N(x)| > |N(x)| = \operatorname{outdeg}(y)$.

Обратно, предположим, что последовательность исходящих степеней вершин в турнире T имеет вид

$$(0, 1, 2, \dots, n-1).$$

Перенумеруем вершины турнира так, чтобы $\operatorname{outdeg}(i) = i$ для любой вершины $i \in V(T)$. Тогда из вершины $n-1$ будет исходить $n-1$ ребро в оставшиеся вершины турнира, из вершины $n-2$ — $n-2$ ребер, и так далее. Иными словами, из вершины с номером i идут ребра во все вершины с меньшими номерами. Теперь транзитивность турнира T следует из свойства транзитивности линейного порядка на множестве вершин этого турнира. Действительно, если $i > j$, а $j > k$, то $i > k$, а значит из вершины i обязательно ведер ребро в вершину k. $\square$

11.1.8. Вернемся к равенству (66). Так как транзитивный турнир представляет собой ориентированный ациклический граф, то на множестве его вершин можно задать линейный порядок, упорядочив эти вершины по убыванию исходящих степеней этих вершин. Теперь возьмем произвольную перестановку σ и сопосотавим ей транзитивный турнир, считая, что из вершины $\sigma(1)$ исходит $n-1$ ребро, из вершины $\sigma(2)$ выходят $n-2$ ребра и так далее. Ясно, что такое сопоставление взаимно-однозначно. Следовательно, мы любому слагаемому в левой части (66) можем поставить в соответствие некоторый транзитивный турнир.

Пример 11.6. Перестановке

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 1 & 5 \end{pmatrix}$$

отвечает транзитивный турнир, показанный на рис.17,b.

11.1.9. Для того, чтобы воспользоваться принципом инволюции (65), нам следует приписать вес любому турниру. Для этого любому ориентированному ребру $e = (i, j)$, направленному из i в j, мы припишем вес $w(e)$ и знак $\operatorname{sign}(e)$ по формулам

$$w(e) := x_i, \qquad \operatorname{sign}(e) = \begin{cases} 1, & \text{если } i < j; \\ -1, & \text{если } i > j. \end{cases}$$

Тогда вес $w(T)$ всего турнира T, а также его знак $\operatorname{sign}(T)$ можно определить по формулам

$$w(T) := \prod_{e \in E(T)} w(e) = \prod_{i=1}^{n} x_i^{\operatorname{outdeg}(i)}, \qquad \operatorname{sign}(T) := \prod_{e \in E(T)} \operatorname{sign}(e),$$

а правую часть (66) можно переписать так:

$$\prod_{1 \leqslant i < j \leqslant n} (x_i - x_j) = \sum_{T} \operatorname{sign}(T) \cdot w(T).$$

Левую же часть равенства (66) можно записать в виде

$$\det \begin{pmatrix} x_1^{n-1} & x_2^{n-1} & \dots & x_{n-1}^{n-1} & x_n^{n-1} \\ x_1^{n-2} & x_2^{n-2} & \dots & x_{n-1}^{n-2} & x_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots & \\ x_1 & x_2 & \dots & x_{n-1} & x_n \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot w(T_\sigma),$$

где

$$w(T_\sigma) = x_{\sigma(1)}^{n-1} \cdot x_{\sigma(2)}^{n-2} \cdot \ldots x_{\sigma(n)}^{0}, \qquad \operatorname{sign}(T_\sigma) = (-1)^{\operatorname{inv}(\sigma)} = \operatorname{sign}(\sigma).$$

Сравнивая два полученных выше выражения, мы приходим к заключению, что нам следует отсеить среди всех турниров те, которые не являются транзитивными. Для этого в качестве множества S возьмем множество всех нетранзитивных турниров и положим

$$S^+ := \{T \in S\colon\ \operatorname{sign}(T) = 1\}, \qquad S^- := \{T \in S\colon\ \operatorname{sign}(T) = -1\}.$$

Нам хочется показать, что $|S^+| = |S^-|$. Сделать это проще всего, построив на множестве $S = S^+ \cup S^-$ альтернированную инволюцию φ без неподвижных точек, которая сохраняла бы нам вес турнира, а затем воспользоваться принципом инволюции (65).

11.1.10. Мы ранее доказали, что в турнире T, не являющемся транзитивным, обязательно существует хотя бы одна пара вершин $i < j$, такая, что $\operatorname{outdeg}(i) = \operatorname{outdeg}(j)$. Выберем среди всех таких вершин турнира T лексикографически минимальную пару, то есть такую пару, у которой вершина i имеет минимальный номер среди вершин, имеющих одинаковую исходящую степень, а вершина j имеет минимальный номер среди вершин, исходящая степень которых совпадает с $\operatorname{outdeg}(i)$. Например, для турнира со степенной последовательностью

$$(\operatorname{outdeg}(1), \operatorname{outdeg}(2), \operatorname{outdeg}(3), \operatorname{outdeg}(4), \operatorname{outdeg}(5), \operatorname{outdeg}(6), \operatorname{outdeg}(7)) = (2, 3, 0, 3, 2, 6, 5)$$

нужной нам парой вершин будет пара $(1, 5)$, а для турнира, изображенного на рис.17,a, такой парой является пара вершин $(1, 2)$. Кроме того, мы для определенности будем считать, что ребро, соответствующее этой паре вершин, направлено из вершины i в вершину j.

Рассмотрим теперь все оставшиеся вершины турнира. Имеются четыре варианта ориентации ребер в треугольнике, образованном вершинами i, j, а также произвольной третьей вершиной k турнира — все они показаны на рис.18. Заметим, что вариант, показанный на рис.18,c, не дает никакого вклада в исходящие степени вершин i и j, а вариант, изображенный на рис.18,d, дает равный вклад в исходящие степени вершин i и j. Но так как ребро (i, j) дает вклад, равный единице, в исходящую степень вершины i, и нулевой вклад в исходящую степень вершины j, то с учетом равенства $\operatorname{outdeg}(i) = \operatorname{outdeg}(j)$ мы приходим к выводу о том, что в турнире обязана

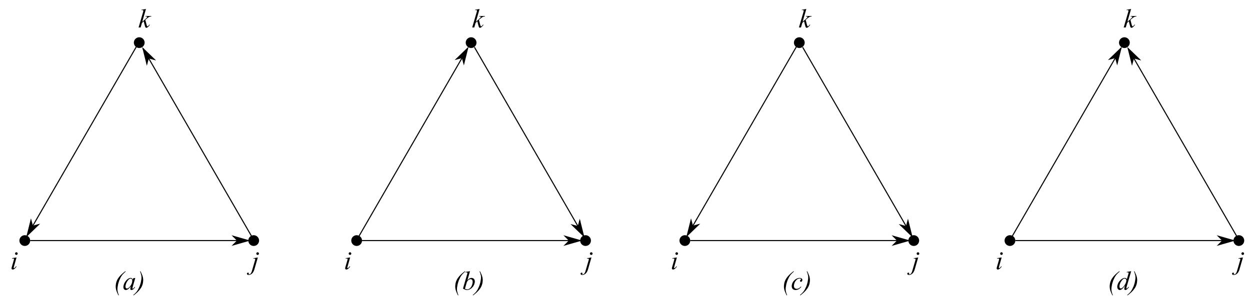

Рис. 18

существовать хотя бы одна вершина k, соответствующая вариантам (a) или (b) расположения ребер в треугольнике, соединяющем вершины i, j, k (см. рис.18,a,b).

Более тщательный анализ вариантов (a) и (b) показывает, что число $|a|$ треугольников типа (a) ровно на единицу больше количества $|b|$ треугольников типа (b). Действительно, вариант (a) дает вклад, равный единице, в исходящую степень вершины i, и нулевой вклад в исходящую степень вершины j, а вариант (b), наоборот, дает вклад, равный единице, в исходящую степень вершины j, и нулевой вклад в исходящую степень вершины i. Но, как мы уже отмечали ранее, само ребро (i, j) дает вклад, равный единице, в исходящую степень вершины i и нулевой вклад в $\operatorname{outdeg}(j)$. Следовательно, степени d_i и d_j вершин i и j равны соответственно

$$d_i = |d| + |b| + 1, \qquad\qquad d_j = |d| + |a|.$$

С учетом условия $d_i = d_j$ мы и получаем равенство

$$|a| = |b| + 1.$$

Отсюда, в частности, следует, что $|a| > 0$ для любого нетранзитивного турнира. Тем самым мы заодно доказали тот факт, что в любом нетранзитивном турнире для любой пары вершин (i, j) обязательно найдется третья вершина k, такая, что через эти три вершины проходит ориентированный цикл длины три.

11.1.11. Теперь мы, наконец, можем определить альтернированную инволюцию φ. Именно, для любого турнира T определим новый турнир $\varphi(T)$, меняя ориентацию ребер на противоположную во всех треугольниках типа (a) и (b), и оставляя ориентацию остальных ребер в турнире неизменной. Ясно, что при такой операции степени вершин, отличных от i и j, останутся неизменными. Покажем, что при этом не поменяются и степени вершин i, j. Обозначим через $\widetilde{d}_i$ и $\widetilde{d}_j$ степени вершин i и j в турнире $\varphi(T)$. Тогда

$$\widetilde{d}_i = |a| + |d| \qquad \Longrightarrow \qquad \widetilde{d}_i - d_i = |a| - |b| - 1 = 0.$$

Аналогично,

$$\widetilde{d}_j = |b| + |d| + 1 \qquad \Longrightarrow \qquad \widetilde{d}_j - d_j = |b| + 1 - |a| = 0.$$

Итак, вес

$$w(T) = \prod_{e \in E(T)} w(e) = \prod_{i=1}^{n} x_i^{\operatorname{outdeg}(i)}$$

всего турнира не меняется при преобразовании φ. Так как $\varphi(\varphi(T))$ дает нам турнир T, то φ действительно представляет собой инволюцию, сохраняющую вес $w(T)$. Наконец, заметим, что φ меняет ориентацию пары ребер (i, k) и (j, k) в любом треугольнике типа (a) или (b), а кроме того, меняет и ориентацию ребра (i, j). Следовательно, $\operatorname{sign}(\varphi(T)) = -\operatorname{sign}(T)$ для любого нетранзитивного турнира. Таким образом, мы доказали, что

$$w(S^+) = w(S^-).$$

Теперь равенство Вандермонда вытекает из принципа инволюции (65).

11.2. В заключение этого параграфа используем принцип инволюции для доказательства очень красивой теоремы, известной как the Lemma of Gessel-Viennot, устанавливающей связь между определителями и решеточными путями.

11.2.1. Займемся вначале комбинаторной интерпретацией определителя

$$\det(M) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot \ldots \cdot a_{n,\sigma(n)} \tag{67}$$

квадратной матрицы M размера $n \times n$. Матрицу M мы можем представить в виде ориентированного взвешенного полного двудольного графа D, в котором вершинам блока X отвечают строки матрицы M, вершинам блока Y — столбцы этой матрицы. При этом любому ребру e, соединяющего вершину $x_i \in X$ с вершиной $y_j \in Y$, приписывается вес $w(e)$, равный элементу $a_{i,j}$ матрицы M.

Для комбинаторной интерпретации равенства (67) рассмотрим в орграфе D систему P_σ вершинно непересекающихся путей вида

$$P_1\colon x_1 \to y_{\sigma(1)}, \qquad P_2\colon x_2 \to y_{\sigma(2)}, \qquad \ldots \qquad P_n\colon x_n \to y_{\sigma(n)},$$

соединяющую вершины множества X с вершинами множества Y. Любой такой системе P_σ мы можем приписать вес $w(P_\sigma)$ как произведение весов $w(P_i)$ отдельных путей этой системы:

$$w(P_\sigma) := w(P_1) \cdot w(P_2) \cdot \ldots \cdot w(P_n) = a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot \ldots \cdot a_{n,\sigma(n)},$$

а также знак $\operatorname{sign}(P_\sigma)$, совпадающий со знаком $\operatorname{sign}(\sigma)$ перестановки σ. Тогда равенство (67) можно переписать в виде

$$\det(M) = \sum_{\sigma \in S_n} \operatorname{sign}(P_\sigma) \cdot w(P_\sigma).$$

и трактовать его следующим образом. В левой части этого равенства стоит определитель матрицы, любой элемент $a_{i,j}$ которой описывает вес единственного в рассматриваемом случае ориентированного пути из вершины x_i в вершину y_j. В правой же части этого равенства стоит знакопеременная сумма весов всевозможных систем P_σ вершинно непересекающихся путей, соединяющих вершины множества X с вершинами множества Y.

11.2.2. Описанная выше комбинаторная интерпретация равенства (67) кажется несколько искусственной. Однако обобщение заложенной в ней идеи на случай произвольных ориентированных ациклических графов оказывается крайне полезной и может быть использовано для решения целого ряда нетривиальных комбинаторных задач.

Пусть D есть ориентированный ациклический граф, в котором любому ребру e приписан какой-то вес $w(e)$. Орграф D имеет конечное количество ориентированных путей P между любой парой вершин (x, y) этого графа. Каждому такому пути P мы можем приписать вес по формуле

$$w(P) = \prod_{e \in P} w(e), \qquad w(P) = 1 \quad \text{в случае, если } P \text{ есть тривиальный путь } x \to x.$$

Предположим теперь, что у нас задано два произвольных, не обязательно не пересекающихся друг с другом n-элементных подмножества вершин $X = \{x_1, \dots, x_n\}$ и $Y = \{y_1, \dots, y_n\}$. Введем для этих подмножеств матрицу M путей, любой элемент $a_{i,j}$ которой равен сумме весов всевозможных путей, соединяющих вершины x_i и y_j:

$$a_{i,j} = \sum_{P \colon x_i \to y_j} w(P).$$

Под системой путей $\mathcal{P}$ из подмножества X в подмножество Y мы будем понимать некоторый набор, состоящий ровно из n путей P_i, соединяющих вершины x_i с вершинами $y_{\sigma(i)}$, где $\sigma \in S_n$ есть какая-то перестановка множества $[n]$. Определим вес $w(\mathcal{P})$ этой системы по формуле

$$w(\mathcal{P}) := \prod_{i=1}^{n} w(P_i),$$

и будем считать, что знак $\operatorname{sign}(\mathcal{P})$ этой системы путей совпадает со знаком $\operatorname{sign}(\sigma)$ перестановки σ. Наконец, мы назовем систему путей вершинно непересекающейся, если любая пара путей, входящих в эту систему, не имеет общих вершин, включая начальные и конечные вершины этих путей. Обозначим через VP семейство вершинно непересекающихся путей.

Лемма 11.7 (Gessel, Viennot, Lindström)**.** *Пусть D есть взвешенный ориентированный ациклический граф, $X = \{x_1, \dots, x_n\} \subseteq V(D)$, $Y = \{y_1, \dots, y_n\} \subseteq V(D)$, а M есть соответствующая этим подмножествам X и Y матрица путей. Тогда определитель этой матрицы равен знакопеременной сумме весов систем вершинно непересекающихся путей, соединяющих вершины множества X с вершинами множества Y:*

$$\det(M) = \sum_{P \in VD} \operatorname{sign}(P) \cdot w(P). \tag{68}$$

11.2.3. Для доказательства равенства (68) заметим, что левая его часть представляет собой сумму слагаемых вида

$$\operatorname{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} = \operatorname{sign}(\sigma) \cdot \left(\sum_{P_1 \colon x_1 \to y_{\sigma(1)}} w(P_1) \right) \cdot \ldots \cdot \left(\sum_{P_1 \colon x_1 \to y_{\sigma(1)}} w(P_1) \right).$$

Суммируя по всем возможным перестановкам $\sigma \in S_n$, мы получаем, что

$$\det(M) = \sum_{P} \operatorname{sign}(P) \cdot w(P),$$

где P пробегает по всем системам путей из X в Y. Для доказательства равенства (68) нам нужно доказать, что

$$\sum_{P \in N} \operatorname{sign}(P) \cdot w(P) = 0,$$

где N — семейство путей, пересекающихся хотя бы по одной вершине. А это мы можем сделать с помощью принципа инволюции (65). Именно, мы докажем лемму, если найдем инволюцию на N без неподвижных точек, такую, что $w(\varphi(P)) = w(P)$ и $\operatorname{sign}(\varphi(P)) = -\operatorname{sign}(P)$. А это сделать в данной задаче достаточно просто.

Действительно, пусть $P \in N$. Среди всех $i = 1, \ldots, n$ выберем наименьший индекс i_0, такой, что путь P_{i_0} пересекается с каким-то из путей P_j, $j > i_0$. Обозначим через z первую вершину, в которой пути P_{i_0} и P_j пересекаются друг с другом. Наконец, среди всех путей P_j, пересекающихся с P_{i_0} в вершине z, выберем путь P_{j_0} с минимальным индексом. Зададим теперь отображение $\varphi(P) = \{P'_1, \ldots, P'_n\}$, так, чтобы P'_k совпадали с P_k для всех $k \neq i_0, j_0$. Кроме того, мы потребуем, чтобы путь P'_{i_0} совпадал с участком пути P_{i_0} от начальной вершины пути до вершины z и с участком пути P_{j_0} от вершины z до конечной вершины $y_{\sigma(j_0)}$ пути P_{j_0}, а путь P'_{j_0}, напротив, совпадал с P_{j_0} на участке от x_{j_0} до z и с путем P_{i_0} на участке от z до y_{i_0}. Так как P и $\varphi(P)$ используют те же самые ребра, то $w(\varphi(P)) = w(P)$. Кроме того, $\operatorname{sign}(\varphi(P)) = -\operatorname{sign}(P)$, так как связанная с системой $\varphi(P)$ путей перестановка σ' представляет собой произведение перестановки σ на транспозицию (i_0, j_0). Построенная таким образом инволюция доказывает равенство (68).

11.2.4. Лемма Гессель-Вьенно позволяет вывести все основные свойства определителей, рассматривая соответствующие графы. В качестве характерного примера покажем, как с помощью равенства (68) доказать формулу Бине-Коши

$$\det(A \cdot B) = \sum_S \det(A[m, S]) \cdot \det(B[S, m]). \tag{69}$$

Здесь A и B — прямоугольные матрицы размерами $m \times n$ и $n \times m$ соответственно, $n \geqslant m$, S — подмножество мощности m индексного множества $[n]$, $A[m, S]$ и $B[S, m]$ — соответствующие этому подмножеству S подматрицы матриц A и B.

Для доказательства формулы (69) представим матрицу A в виде полного двудольного ориентированного графа $K^A_{m,n}$, построенного на множествах вершин $X = \{x_1, \ldots, x_m\}$ и $Y = \{y_1, \ldots, y_n\}$, ребра (x_i, y_j) которого имеют веса $a_{i,j}$. Аналогично, матрицу B представим в виде полного двудольного ориентированного графа $K^B_{n,m}$ на множествах вершин Y и $Z = \{z_1, \ldots, z_m\}$, ребрам (y_j, z_k) которого приписаны веса $b_{j,k}$. Образуем теперь из этих двух двудольных орграфов единый орграф D, отождествляя множества Y вершин этих двух графов. Рассмотрим всевозможные системы P путей из X в Z в таком орграфе. Любой элемент $m_{i,k}$ матрицы M путей для подмножеств X и Z вершин орграфа D записывается в виде

$$m_{i,k} = \sum_{P\colon x_i \to z_k} w(P) = \sum_{j=1}^{n} a_{i,j} \cdot b_{j,k},$$

а значит, $M = A \cdot B$.

Теперь заметим, что произвольную систему, состоящую из n вершинно непересекающихся путей из X в Z, мы можем построить, выбрав в Y подмножество S вершин мощности m и соединив X с S, а S с Z. Количество вершинно непересекающихся путей из X в S описывается определителем $\det(A[m, S])$, а количество вершинно непересекающихся путей из S в Z — определителем $\det(B[S, m])$. Следовательно, согласно комбинаторному правилу произведения,

$$\sum_P \operatorname{sign}(P) \cdot w(P) = \det(A[m, S]) \cdot \det(B[S, m]),$$

и формула Бине-Коши получается как непосредственное следствие леммы Гессель-Вьенно.

11.2.5. Еще одно очень красивое приложение леммы Гессель-Вьенно связано с матрицами Ханкеля чисел Каталана.

Определение 11.8. Рассмотрим произвольную последовательность $(a_0, a_1, a_2, \dots, a_n, \dots)$ комплексных чисел. Матрицами Ханкеля этой последовательности называются матрицы вида

$$H_n := \begin{pmatrix} a_0 & a_1 & \dots & a_n \\ a_1 & a_2 & \dots & a_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n+1} & \dots & a_{2n} \end{pmatrix} \qquad \text{и} \qquad H_n^{(1)} := \begin{pmatrix} a_1 & a_2 & \dots & a_{n+1} \\ a_2 & a_3 & \dots & a_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n+1} & a_{n+1} & \dots & a_{2n+1} \end{pmatrix}.$$

Как следствие, с любой последовательностью $\{a_n\}$ можно связать последовательность чисел

$$(\det(H_0), \det(H_0^{(1)}, \det(H_1), \det(H_1^{(1)}, \dots, \det(H_n), \det(H_n^{(1)}, \dots).$$

Обратно, по любой такой последовательности чисел можно восстановить исходную последовательность $\{a_n\}$ при условии, что все определители этих матриц отличны от нуля. Например, $\det(H_0) = a_0$, $\det(H_1) = a_1$, $\det(H_1) = a_0 a_2 - a_1^2$, так что a_2 определяется отсюда однозначно с учетом условия $a_0 = \det(H_0) \neq 0$, и так далее.

Утверждение 11.9. *Числа Каталана представляют собой единственную числовую последовательность вещественных чисел, для которой*

$$\det(H_n) = \det(H_n^{(1)}) = 1 \qquad \forall \quad n = 0, 1, 2, \dots$$

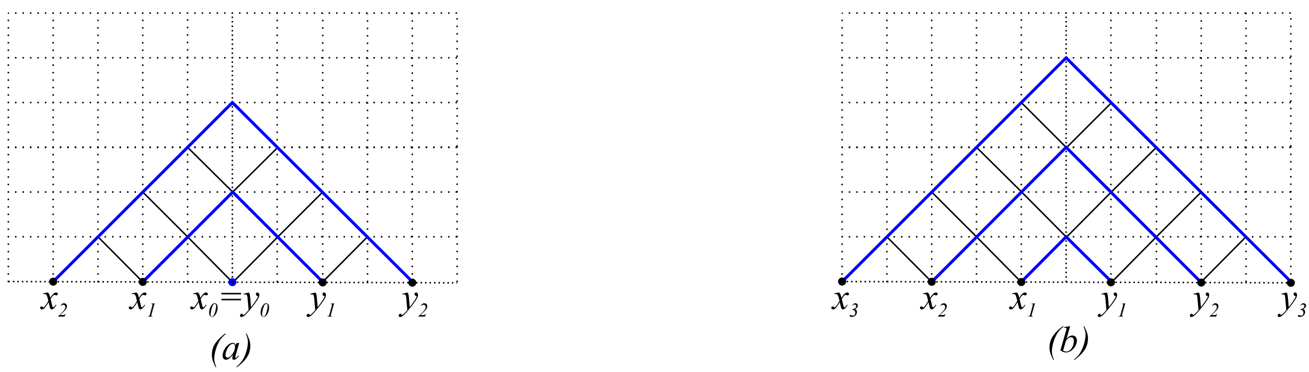

Рис. 19

Покажем с помощью леммы Гессель-Вьенно, что $\det(H_n) = 1$ и $\det(H_n^{(1)}) = 1$. Для этого рассмотрим на плоскости решеточные пути, показанные на рис.19,a и рис.19,b. Первому рисунку отвечают пути Дика, исходящие из точек с координатами $x_i = (-2i, 0)$, и приходящие в точки с координатами $y_i = (2i, 0)$, $i = 0, 1, 2, \dots$. Второму рисунку отвечают пути Дика, исходящие из точек с координатами $x_i = (-2i - 1, 0)$, и приходящие в точки с координатами $y_i = (2i + 1, 0)$, $i = 0, 1, 2, \dots$. Мы знаем, что для первого случая количество путей Дика, соединяющих точки x_i и y_j, равно C_{i+j}. Для второго случая это количество описывается числами Каталана C_{i+j+1}. Теперь заметим, что в обоих случаях существует только одна вершинно непересекающаяся система путей, соединяющая вершины x с вершинами y (см. линии, помеченные синим цветом на рис.19). Следовательно, согласно (68), оба определителя равны единице.

Список литературы

[1] Орен Паташник Дональд Кнут, Роналд Грэхем. *Конкретная математика. Основание информатики.* М.: Мир; Бином. Лаборатория знаний, 2006.

[2] M. Bona. *A Walk Through Combinatorics: an introduction to enumeration and graph theory.* World Sciendific Publishing Co, second edition edition, 2006.

[3] Р. Стенли. *Перечислительная комбинаторика.* М.: Мир, 1990.

[4] Р. Стенли. *Перечислительная комбинаторика. Деревья, производящие функции и симметрические функции.* М.: Мир, 2005.

[5] П.А.Виленкин Н.Я. Виленкин, А.Н.Виленкин. *Комбинаторика.* М.: МЦНМО, 2006.

[6] D.Knut. *Искусство программирования.* Мир, 2000.