КОНКРЕТНАЯ МАТЕМАТИКА

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ

ВТОРОЕ ИЗДАНИЕ

ГРЭХЕМ • КНУТ • ПАТАШНИК

ВТОРОЕ ИЗДАНИЕ
КОНКРЕТНАЯ МАТЕМАТИКА
МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ
ГРЭХЕМ
КНУТ
ПАТАШНИК

ВИЛЬЯМС
ADDISON WESLEY

КОНКРЕТНАЯ МАТЕМАТИКА

ВТОРОЕ ИЗДАНИЕ

CONCRETE MATHEMATICS

SECOND EDITION

A Foundation for Computer Science

ADDISON-WESLEY

Upper Saddle River, NJ • Boston • Indianapolis • San Francisco
New York • Toronto • Montrèal • London • Munich • Paris • Madrid
Capetown • Sydney • Tokyo • Singapore • Mexico City

КОНКРЕТНАЯ МАТЕМАТИКА

ВТОРОЕ ИЗДАНИЕ

Математические основы информатики

ВИЛЬЯМС

Издательский дом "Вильямс"
Москва • Санкт-Петербург • Киев
2010

ББК 32.973.26-018.2.75
Г91
УДК 519.688

Издательский дом "Вильямс"

Зав. редакцией *С.Н. Тригуб*

Перевод с английского и редакция канд. техн. наук *И.В. Красикова*

По общим вопросам обращайтесь в Издательский дом "Вильямс" по адресу:
info@williamspublishing.com, http://www.williamspublishing.com

Грэхем, Рональд Л., **Кнут**, Дональд Э., **Паташник**, Орен.

Г91 Конкретная математика. Математические основы информатики, 2-е изд. : Пер. с англ. — М. : ООО "И.Д. Вильямс", 2010. — 784 с. : ил. — Парал. тит. англ.

ISBN 978-5-8459-1588-7 (рус.)

ББК 32.973.26-018.2.75

Научно-популярное издание

Рональд Л. Грэхем, Дональд Э. Кнут, Орен Паташник

Конкретная математика. Математические основы информатики

2-е издание

Литературный редактор *Л. Н. Красножон*
Верстка *А. Н. Полинчик*
Художественный редактор *Е. П. Дынник*
Корректор *И. В. Крассиков*

Подписано в печать 02.10.2009. Формат 70×100/16.
Гарнитура CM-LH. Печать офсетная.
Усл. печ. л. 49,0. Уч.-изд. л. 53,2.
Тираж 1000 экз. Заказ № 19486.

Отпечатано по технологии CtP
в ОАО "Печатный двор" им. А. М. Горького
197110, Санкт-Петербург, Чкаловский пр., 15.

ООО "И. Д. Вильямс", 127055, г. Москва, ул. Лесная, д. 43, стр. 1

ISBN 978-5-8459-1588-7 (рус.)
ISBN 0-201-55802-5 (англ.)

Оглавление

Предисловие

“Аудитория, уровень изложения и трактовка материала — вот для чего предназначены предисловия”.
— П. Р. Халмош (P. R. Halmos) [173]

ЭТА КНИГА ОСНОВАНА на одноименном курсе лекций, который ежегодно читается в Стэнфордском университете начиная с 1970 года. Каждый год его прослушивают около пятидесяти человек — студентов как средних, так и старших курсов, но в первую очередь дипломников (а многие из наших выпускников уже начали вводить такого рода курсы и в других местах). По-видимому, настала пора представить материалы курса более широкой аудитории (включая студентов младших курсов).

“Отдельные лица приобретают дешевый авторитет, оснастив свою речь жаргоном: они могут проповедовать и выставлять напоказ поверхностные суждения. Но от математиков-профессионалов требуются не разглагольствования и даже не степень осведомленности в тех или иных математических вопросах, а готовность применять свои знания и способность реально решать возникающие на практике математические задачи. Короче говоря, мы ждем дел, а не слов”.
— Дж. Хаммерсли (J. Hammersley) [176]

Конкретная математика зарождалась в смутное и неспокойное десятилетие. В те бурные годы подвергалось сомнениям все, включая казавшиеся до этого незыблемыми ценности. Студенческие городки превращались в арены жарких баталий. Оспаривались сами учебные программы, и математика не могла быть исключением. Как раз в те годы Джон Хаммерсли (John Hammersley) написал свою полемическую статью “О снижении уровня математической подготовки в школах и университетах благодаря ‘современной математике’ и подобной ей жидкой интеллектуальной похлебке” [176]; другие обеспокоенные математики [332] даже задавались вопросом “А можно ли спасти математику?” Когда один из авторов этой книги (Д.Э.К.) задумал серию книг под названием *Искусство программирования*, то при написании первого тома он обнаружил, что в его арсенале отсутствуют самые важные инструменты. Математика, которую он изучал в колледже в качестве профильной дисциплины, разительно отличалась от той математики, которая требовалась для досконального, обоснованного понимания компьютерных программ. Поэтому он ввел новый курс, содержащий материал, который хотел бы в свое время прослушать сам.

Новый курс с самого начала получил название “конкретная математика” — сперва как противопоставление “абстрактной математике”, поскольку конкретные классические результаты очень быстро выметались из современного математического образования новой метлой абстрактных идей, популярно именовавшихся “новоматом” (“New Math”). Абстрактная математика — чудесный предмет, в котором нет ничего плохого: она красива, обща и полезна. Однако ее приверженцы впали в опасное заблуждение: они решили, что вся остальная математика ниже ее, занимает подчиненное положение и не заслуживает особого внимания. Погоня за общностью оказалась столь захватывающей, что стала самоцелью для целого поколения математиков, которые потеряли способность находить красоту в частностях, в том числе получать удовольствие от решения численных задач или по достоинству оценивать роль математических методов. Абстрактная математика стала вырождаться и терять связь с действительностью. Математическому образованию срочно потребовался конкретный противовес для восстановления равновесия.

Когда Д.Э.К. приступал к чтению курса конкретной математики в Стэнфордском университете, он пояснял несколько странное название курса тем, что это попытка преподавания в стиле "хард" вместо "софт". Он объявил, что, в отличие от своих коллег, он *не* намерен излагать ни теорию агрегатов, ни теорему вложения Стоуна, ни даже компактификацию Стоуна–Чеха. (Несколько студентов факультета гражданского строительства потихоньку покинули аудиторию.*)

"Сердце математики — в конкретных примерах и конкретных задачах."
— П. Р. Халмош (P. R. Halmos) [172]

Хотя конкретная математика возникла в качестве реакции на другие тенденции в математике, основные причины ее появления на свет скорее позитивны, нежели негативны. По мере того как этот курс завоевывал себе место под солнцем в учебном процессе, содержание предмета "цементировалось" и доказывало свою ценность в целом ряде новых приложений. Тем временем поступило независимое подтверждение уместности подобного наименования, когда З. А. Мелзак (Z. A. Melzak) опубликовал два тома, озаглавленные *Справочник по конкретной математике* [267].

"Непростительный грех — учить абстрактному до изучения конкретного."
— З. А. Мелзак (Z. A. Melzak) [267]

На первый взгляд, материал конкретной математики может показаться беспорядочным нагромождением хитроумных трюков, но на самом деле это упорядоченный набор инструментов. Методы конкретной математики обладают не только внутренним единством, но и внешней привлекательностью для множества людей. Когда другой автор этой книги (Р.Л.Г.) впервые прочел этот курс в 1979 году, восторг студентов был столь велик, что они тут же решили продлить это удовольствие еще на год.

Но что же такое конкретная математика на самом деле? Это смесь КОНтинуальной и дисКРЕТНОЙ математик. Еще более кон-

Конкретная математика — мост к абстрактной математике.

---

* Имеется в виду, что название "конкретная математика" (concrete mathematics) в английском языке, помимо смысла "конкретная", имеет и второй смысл — "бетонная". — *Примеч. пер.*

кретно — это осмысленное оперирование математическими формулами с использованием определенного набора методов решения задач. После того как вы, читатель, изучите материал данной книги, все, что вам потребуется, — это холодная голова, большой лист бумаги и сносный почерк для вычисления ужасно выглядящих сумм, решения запутанных рекуррентных соотношений и выявления коварных закономерностей в данных. Вы овладеете алгебраической техникой в такой степени, что зачастую вам будет проще получать точные результаты, нежели довольствоваться приближенными ответами, справедливыми лишь в ограниченном смысле.

"Более подготовленный читатель, пропустивший кажущиеся ему слишком элементарными части, может потерять больше, чем менее подготовленный читатель, пропустивший части, кажущиеся слишком сложными"
— Г. Пойа (G. Pólya) [297]

Основные темы книги включают исчисление сумм, рекуррентные соотношения, элементарную теорию чисел, биномиальные коэффициенты, производящие функции, дискретную вероятность и асимптотические методы. Основной упор при этом делается на технической стороне дела, а не на теоремах существования или комбинаторных рассуждениях; цель заключается в том, чтобы сделать каждого читателя настолько осведомленным в дискретных операциях (наподобие вычисления функции "наибольшего целого" или конечной суммы), насколько изучающие анализ знакомы с операциями континуальными (наподобие вычисления функции "абсолютной величины" или определенного интеграла).

Заметим, что этот перечень тем кардинально отличается от того, что в наше время обычно читается в качестве спецкурсов под названием "Дискретная математика". Поэтому наш предмет нуждается в отличительном наименовании, и название "Конкретная математика", право, не хуже любого другого.

Мы недостаточно дерзки для названия "Дистинуальная математика".

130 в русском переводе.
— Переводчик

Первоначальным руководством по конкретной математике для стэнфордского курса послужил раздел "Математическое введение" из *Искусства программирования* [207]. Но изложение на этих 110 страницах было слишком сжатым, поэтому еще один автор (О.П.) загорелся желанием внести большое количество дополнений. Настоящая книга выросла на этих примечаниях: она предваряет и расширяет материал "Математического введения". Некоторые вопросы повышенной сложности опущены; в то же время в книгу включено несколько тем, которых не было ранее и без которых материал был бы неполным.

"... конкретный спасательный круг, брошенный студентам, тонущим в море абстракции".
— У. Готтшалк (W. Gottschalk)

Авторы с удовольствием объединили свои усилия для работы над этой книгой, поскольку ее предмет начал зарождаться и обретать собственную жизнь у них на глазах; кажется, что книга написалась как бы сама собой. Более того, несколько различные подходы каждого автора после нескольких лет совместной работы оказались хорошо пригнанными друг к другу, так что возникает ощущение некоторого манифеста выбранного нами способа занятий математикой. Поэтому мы думаем, что эта книга окажется одой красоте и очарованию математики и что наши читатели разделят с нами хотя бы ϵ того удовольствия, которое мы получили при ее написании.

Поскольку книга родилась в университетской среде, мы попытались передать дух аудитории наших дней, выбрав неформальный стиль изложения. Некоторые полагают, что математика — это серьезное дело, холодное и сухое; но мы считаем ее развлечением и не боимся в этом признаться. Так ли уж необходимо проводить четкую грань между работой и игрой? Конкретная математика полна тому примеров — пусть выполняемые действия и не всегда приятны, зато ответы могут принести удивительную радость. Радости и горести математической работы явно присутствуют в этой книге, поскольку представляют собой часть нашей жизни.

Студенты всегда все знают лучше преподавателей, поэтому мы попросили первых студентов, изучавших этот материал, внести свой вклад в виде "граффити" на полях. Некоторые из этих заметок были попросту банальны, некоторые полны смысла; одни из них предупреждали о двусмысленностях или неясностях, другие оказывались типичными комментариями умников с "Камчатки". Часть замечаний положительна, часть отрицательна, ценность еще одной части попросту нулевая. Но все они, несомненно, отражают реальные чувства читателей, что должно облегчить восприятие книги. (Вдохновляющей идеей для таких пометок на полях послужил справочник *Поступающему в Стэнфорд*, в котором официальной линии университета противопоставляются замечания выпускников. Например, в справочнике сказано: "Есть несколько вещей, которые нельзя пропустить в столь аморфном образовании, каковым является Стэнфорд"; в примечании на полях — "Аморфное образование! Ну и выражения! Вокруг одни типичные псевдоинтеллектуалы". Стэнфорд: "Потенциал совместно проживающих студентов безграничен". Граффити: "Здешние общаги — зоопарк без дворника".)

Математическое граффити:

Килрой не был Хааром.
Освободите группу.
Взорвите ядро.
$N = 1 \Rightarrow P = NP$.

Не замечаю за собой интереса к этим замечаниям.

На полях мы также цитируем великих математиков прошлого — подлинные слова, которыми они сообщали о своих фундаментальных открытиях. Представляется уместным привести на одних и тех же страницах слова Лейбница, Эйлера, Гаусса и тех, кому предстоит продолжать их дело. Математика по-прежнему привлекает своих приверженцев, и каждая нить находит свое место в этом богатом полотне.

Это был самый приятный курс из пройденных мною. Но было бы неплохо подытоживать материал по мере продвижения вперед.

Понятно: конкретная математика — это муштра.

В книге имеется более 500 упражнений, разделенных на шесть категорий.

- ***Разминка*** — это упражнения, которые КАЖДЫЙ ЧИТАТЕЛЬ должен попытаться выполнить при первом прочтении материала.
- ***Обязательные упражнения*** предназначены для самостоятельного установления фактов, которые лучше всего усваиваются, если их выводить самому, а не читать о том, как это делали другие.
- ***Домашние задания*** представляют собой задачи для углубленного понимания материала той главы, к которой они относятся.
- ***Контрольные работы*** обычно охватывают материал двух и более глав одновременно; в основном они предназначены для неспешного выполнения дома (а не в цейтноте в аудитории).
- ***Дополнительные задачи*** выходят за рамки ожидаемых возможностей среднего студента, изучающего курс конкретной математики на базе этой книги. Они расширяют текст книги в важных и интересных направлениях.
- ***Исследовательские проблемы*** могут быть разрешимы человеком (а могут и не быть таковыми), но те из них, которые представлены в книге, стоит попытаться решить (без ограничения времени).

Домашнее задание заставило выучить большой объем нового материала, так что каждый затраченный час стоил того.

Домашние контрольные работы вам еще пригодятся — не выбрасывайте их.

Контрольные оказались труднее, чем можно было ожидать по домашним заданиям.

Ответы ко всем этим упражнениям приведены в приложении А, зачастую с дополнительной информацией о родственных результатах. (Конечно же, "ответы" на исследовательские проблемы являются неполными, но даже в этих случаях частичные результаты или указания могут оказаться полезными.) Читателям не возбраняется заглянуть в ответы главным образом разминочных задач, но только ПОСЛЕ серьезных попыток решить задачу самостоятельно, без подсказок.

Лентяи могут сдать курс, просто списывая ответы, но обманут при этом только самих себя.

В приложении Б мы попытались воздать должное первоисточникам каждого упражнения, поскольку составление той или иной поучительной задачи зачастую представляет собой творческий процесс с изрядной долей везения. К сожалению, математики выработали традицию заимствовать упражнения без какой бы то ни было признательности; мы же считаем, что гораздо лучшей является обратная традиция (практикуемая, например, в шахматных книгах и журналах, где принято указывать авторов, дату и место появления оригинальных шахматных задач). Однако мы так и не смогли выявить источники многих задач,

Дополнительные задачи не рассчитаны на студентов, специализирующихся по другим предметам.

ставших частью математического фольклора. Если кому-либо из читателей известно о происхождении того или иного упражнения, ссылка на которое нами пропущена или неточна, мы будем рады получить от него подробную информацию, чтобы исправить упущение в следующих изданиях книги.

Шрифт, которым набраны математические обозначения в книге, — новая разработка Германа Цапфа (Hermann Zapf) [227], заказанная Американским математическим обществом. Она выполнена при содействии комиссии, в состав которой вошли Б. Битон (B. Beeton), Р. Ф. Боас (R. P. Boas), Л. К. Дарст (L. K. Durst), Д. Э. Кнут (D. E. Knuth), Ф. Мердок (P. Murdock), Р. Ш. Пале (R. S. Palais), П. Ренц (P. Renz), Э. Свансон (E. Swanson), С. Б. Уидден (S. B. Whidden) и У. Б. Вульф (W. B. Woolf). Основная идея дизайна Цапфа — отразить особенности написания математических знаков идеальным почерком. Рукописный стиль, в отличие от механического, более естествен, потому что обычно математические формулы выходят из-под пера, куска мела или карандаша. (Примером одного из фирменных признаков этого дизайна является начертание нуля, '0', который слегка заострен сверху, — ведь рукописный нуль редко закругляется в исходной точке.) Буквы расположены прямо, а не наклонно, с тем, чтобы нижние и верхние индексы, а также штрихи, было легче совмещать с обычными символами. Новое семейство шрифтов получило название *AMS Euler*, в честь великого математика Леонарда Эйлера (Leonhard Euler) (1707–1783), открывшего так много в математике из того, что нам известно сегодня. Алфавиты *AMS Euler* включают текстовые ($\mathrm{A}\mathrm{a}\,\mathrm{B}\mathrm{b}\,\mathrm{C}\mathrm{c}\ \ldots\ \mathrm{X}\mathrm{x}\,\mathrm{Y}\mathrm{y}\,\mathrm{Z}\mathrm{z}$), готические ($\mathfrak{A}\mathfrak{a}\,\mathfrak{B}\mathfrak{b}\,\mathfrak{C}\mathfrak{c}\ \ldots\ \mathfrak{X}\mathfrak{x}\,\mathfrak{Y}\mathfrak{y}\,\mathfrak{Z}\mathfrak{z}$) и рукописные прописные ($\mathcal{A}\,\mathcal{B}\,\mathcal{C}\ \ldots\ \mathcal{X}\,\mathcal{Y}\,\mathcal{Z}$) буквы, а также греческие буквы ($\mathrm{A}\alpha\,\mathrm{B}\beta\,\Gamma\gamma\ \ldots\ \mathrm{X}\chi\,\Psi\psi\,\Omega\omega$) и специальные символы наподобие $\wp$ и $\aleph$. Нам особенно приятно торжественно представить это семейство шрифтов в нашей книге, поскольку дух Леонарда Эйлера воистину живет на каждой ее странице: конкретная математика — это эйлерова математика!

Значит, вы не видели мой почерк.

Авторы чрезвычайно признательны Андрею Бродеру (Andrei Broder), Эрнсту Мэйру (Ernst Mayr), Эндрю Яо (Andrew Yao) и Френсис Яо (Frances Yao), которые внесли значительный вклад в книгу в те годы, когда они преподавали конкретную математику в Станфорде. Кроме того, мы выражаем 1024 благодарности ассистентам, творчески подошедшим к записи происходившего в аудитории каждый год и помогавшим составлять экзаменационные вопросы; их имена перечислены в приложении В. Эта книга, которая по сути представляет собой конспект всего ценного, прозвучавшего на лекциях за шестнадцать лет, была бы просто невозможна без их первоклассной работы.

Профессор, спасибо за (1) шутки и каламбуры, (2) серьезность и содержательность предмета.

Не вижу, где бы я мог применить изученный материал...

Стать реальностью данной книге помогало множество других людей. Например, достойны похвалы студенты университетов Брауна и Райса, Колумбийского, Нью-Йоркского, Принстонского и Станфордского университетов, внесшие вклад в виде отобранных для книги граффити и оказавшие помощь в отладке первых черновых версий книги. Наше сотрудничество с издательством Addison-Wesley было особенно эффективным и плодотворным; в частности, мы хотим поблагодарить нашего издателя Питера Гордона (Peter Gordon), технического директора Бет Ааронсон (Bette Aaronson), дизайнера Роя Брауна (Roy Brown) и редактора Лин Дюпре (Lyn Dupré). Неоценимую помощь нам оказали Национальный научный фонд и Отделение военно-морских исследований. При составлении предметного указателя незаменима была Шерил Грэхем (Cheryl Graham). Кроме того, мы хотим поблагодарить наших жен — Фэн, Джилл и Эми — за их терпение, поддержку, ободрение и советы.

С этим предметом было немало хлопот, но я знаю, что он отточил мои математические и умственные способности.

Второе издание книги отличается новым разделом, 5.8, в котором описан ряд важных идей, которые Дорон Зайльбергер (Doron Zeilberger) открыл вскоре после выхода в свет первого издания. Кроме того, исправления к первому изданию встречаются почти на каждой странице.

Мы пытались создать идеальную книгу, но сами мы не идеальны, а потому призываем оказать содействие в исправлении допущенных нами ошибок. Мы с признательностью выплатим премию в сумме 2.56 доллара первому нашедшему любую ошибку в английском оригинальном издании книги, будь то ошибка математическая, историческая или типографская.

Я бы посоветовал случайному студенту держаться от этого курса подальше.

Мюррей-Хилл, Нью-Джерси — Р.Л.Г.
Станфорд, Калифорния Д.Э.К.
Май 1988 и октябрь 1993 года О.П.

От издательства

Вы, читатель этой книги, и есть главный ее критик. Мы ценим ваше мнение и хотим знать, что было сделано нами правильно, что можно было сделать лучше и что еще вы хотели бы увидеть изданным нами. Нам интересно услышать и любые другие замечания, которые вам хотелось бы высказать в наш адрес.

Мы ждем ваших комментариев и надеемся на них. Вы можете прислать нам бумажное или электронное письмо, либо просто посетить наш Web-сервер и оставить свои замечания там. Одним словом, любым удобным для вас способом дайте нам знать, нравится или нет вам эта книга, а также выскажите свое мнение о том, как сделать наши книги более интересными для вас.

Отсылая письмо или сообщение, не забудьте указать название книги и ее авторов, а также свой обратный адрес. Мы внимательно ознакомимся с вашим мнением и обязательно учтем его при отборе и подготовке к изданию последующих книг. Наши координаты:

E-mail: info@williamspublishing.com
WWW: http://www.williamspublishing.com

Адреса для писем из:

России: 127055, г. Москва, ул. Лесная, д. 43, стр. 1
Украины: 03150, Киев, а/я 152

Принятые обозначения

ЧАСТЬ СИМВОЛИКИ, используемой в этой книге, (пока еще?) не стала нормой. Вот список обозначений, которые могут быть незнакомы читателям, изучавшим аналогичный материал по другим книгам. В нем указаны номера страниц, на которых эти обозначения разъясняются.

Обозначение	*Смысл*	*Страница*
$\ln x$	Натуральный логарифм: $\log_e x$	339
$\lg x$	Бинарный логарифм: $\log_2 x$	98
$\log x$	Десятичный логарифм: $\log_{10} x$	541
$\lfloor x \rfloor$	Пол: $\max\{ n \mid n \leqslant x,\ \text{целое } n \}$	95
$\lceil x \rceil$	Потолок: $\min\{ n \mid n \geqslant x,\ \text{целое } n \}$	95
$x \bmod y$	Остаток: $x - y\lfloor x/y \rfloor$	113
$\{x\}$	Дробная часть: $x \bmod 1$	98
$\sum f(x)\,\delta x$	Неопределенная сумма	73
$\sum_a^b f(x)\,\delta x$	Определенная сумма	74
$x^{\underline{n}}$	Убывающая факториальная степень: $x!/(x-n)!$	71, 265
$x^{\overline{n}}$	Возрастающая факториальная степень: $\Gamma(x+n)/\Gamma(x)$	71, 265
$n¡$	Субфакториал: $n!/0! - n!/1! + \cdots + (-1)^n n!/n!$	246
$\Re z$	Действительная часть: x, если $z = x + iy$	91
$\Im z$	Мнимая часть: y, если $z = x + iy$	91
H_n	Гармоническое число: $1/1 + \cdots + 1/n$	50

$H_n^{(x)}$	Обобщенное гармоническое число: $1/1^x + \cdots + 1/n^x$	340
$f^{(m)}(z)$	m-я производная f в точке z	564
$\begin{bmatrix} n \\ m \end{bmatrix}$	Число Стирлинга первого рода (число циклов)	320
$\begin{Bmatrix} n \\ m \end{Bmatrix}$	Число Стирлинга второго рода (число подмножеств)	318
$\left\langle {n \atop m} \right\rangle$	Число Эйлера	328
$\left\langle\!\left\langle {n \atop m} \right\rangle\!\right\rangle$	Число Эйлера второго порядка	331
$(a_m \ldots a_0)_b$	Обозначение для $\sum_{k=0}^{m} a_k b^k$ в системе счисления с основанием b	30
$K(a_1, \ldots, a_n)$	Континуант	367
$F\left({a, b \atop c} \middle\vert z \right)$	Гипергеометрическая функция	258
$\#A$	Мощность: количество элементов в множестве A	62
$[z^n]\, f(z)$	Коэффициент при z^n в разложении $f(z)$	249
$[\alpha \,..\, \beta]$	Замкнутый интервал: множество $\{x \mid \alpha \leqslant x \leqslant \beta\}$	102
$[m = n]$	1, если $m = n$, в противном случае — 0*	45
$[m \backslash n]$	1, если m нацело делит n, в противном случае — 0*	137
$[m \backslash\backslash n]$	1, если m напросто делит n, в противном случае — 0*	189
$[m \perp n]$	1, если m взаимно простое с n, в противном случае — 0*	154

В хорошем учебнике по конкретной математике не обойтись без сбивающего с толку списка обозначений.

*В общем случае, если S — некоторое утверждение, которое может быть истинно или ложно, обозначение в квадратных скобках $[S]$ равно 1, если S истинно, и 0 в противном случае.

Повсюду в книге мы используем одинарные кавычки ('...') для выделения текста, как он *пишется*, а двойные кавычки ("...") — для фразы, как она *произносится*. Так, строка букв 'строка' называется "строка".

'Нестрока' тоже строка.

Выражение вида 'a/bc' означает то же, что и '$a/(bc)$'. Кроме того, $\log x/\log y = (\log x)/(\log y)$ и $2n! = 2(n!)$.

1

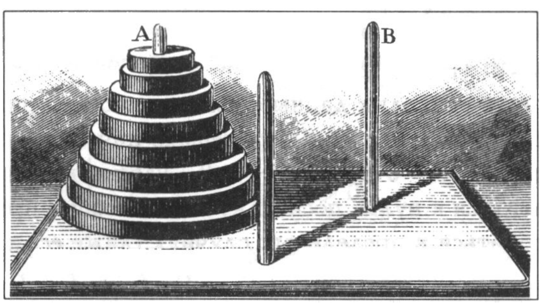
Рекуррентные задачи

В ЭТОЙ ГЛАВЕ в качестве примера рассматриваются три задачи, которые дадут вам понять, что же будет дальше. Эти задачи объединяет то, что их неоднократно изучали математики и их решения основаны на идее *рекуррентности*, согласно которой решение каждой задачи зависит от решений меньших экземпляров той же самой задачи.

1.1 Ханойская башня

Начнем с небольшой изящной головоломки под названием "Ханойская башня", придуманной французским математиком Эдуардом Люка́ (Edouard Lucas) в 1883 году. Имеется башенка, сложенная из восьми дисков, изначально одетых на один из трех колышков в порядке уменьшения размера дисков.

Поднимите руку, если вы никогда этого не видели. Остальные могут сразу перейти к уравнениям (1.1).



Суть головоломки состоит в том, чтобы перенести всю башню на другой колышек (используя третий в качестве вспомогательного) по одному диску за раз, причем больший диск никогда не должен находиться поверх меньшего.

Люка [260] связывал свою головоломку с легендой о гораздо большей "Башне Брахмы", которая, как утверждается, состоит из 64 дисков из чистого золота на трех алмазных шпилях. В начале

А, золото!.. Интересно, а из чего сделаны наши чисто конкретные диски?

времен Бог поместил эти золотые диски на первый шпиль и повелел группе жрецов перенести башню на третий шпиль согласно указанным выше условиям. По слухам жрецы беспрерывно, день и ночь, сменяя друг друга, без устали перекладывают диски, и когда последний диск будет перенесен, башня рассыплется в пыль и настанет конец света.

То, что головоломка имеет решение, на первый взгляд не очевидно. Но если немного подумать (или познакомиться с этой задачей заранее), то можно убедиться, что это так. При этом возникает вопрос: как перенести башню наилучшим образом? То есть какого количества переносов дисков достаточно для решения поставленной задачи?

Лучший способ решения подобного рода вопросов — это обобщить его и решать более общую задачу. Башня Брахмы состоит из 64 дисков, Ханойская башня — из 8. Давайте рассмотрим обобщенную башню из n дисков.

Одно из достоинств такого подхода состоит в том, что можно существенно уменьшить размер задачи. В этой книге мы часто будем видеть, как полезно вначале РАССМОТРЕТЬ МАЛЫЕ СЛУЧАИ. Легко разобраться, как переместить башню, состоящую всего из одного или двух дисков. После небольшого количества экспериментов с парой дисков становится очевидным, как перемещать башню из трех дисков.

Следующий шаг в решении задачи состоит во введении подходящих обозначений: ИМЕНУЙ И ВЛАСТВУЙ. Пусть T_n — минимальное количество перемещений, необходимое для переноса n дисков с одного шпиля на другой в соответствии с правилами Люка. Очевидно, что $T_1 = 1$, $T_2 = 3$.

Можно получить еще одни данные, совершенно не прикладывая усилий: если дисков нет, то и перемещать ничего не надо! Значит, для башни из $n = 0$ дисков $T_0 = 0$. Умные математики не боятся обратиться к нулю, поскольку понять общее проще, если хорошо разобраться в предельных случаях (даже если они тривиальны).

А теперь давайте изменим перспективу и вместо малого подумаем о большом: как переместить большую башню? Эксперименты с тремя дисками показывают, что выигрышная идея состоит в том, чтобы перенести два диска на средний шпиль, переместить большой диск на третий, а затем два диска со среднего шпиля поместить на третий шпиль поверх большого диска. Эта идея применима и для переноса n дисков в общем случае: сначала мы переносим $n-1$ меньших дисков на промежуточный шпиль (для чего требуется T_{n-1} перемещений дисков), затем переносим наибольший диск в его конечное положение (это делается при помощи одного перемещения), а после опять переносим $n-1$ меньших дисков с промежуточного шпиля на наибольший диск (для чего нужны очередные T_{n-1} перемещений). Таким образом, мы можем перенести $n>0$ дисков не более чем за $2T_{n-1}+1$ перемещения:

$$T_n \leqslant 2T_{n-1}+1 \qquad \text{при } n>0.$$

В этой формуле вместо '$=$' используется '$\leqslant$', поскольку наше построение доказывает только то, что $2T_{n-1}+1$ перемещений достаточно; мы еще не показали, что необходимо $2T_{n-1}+1$ перемещений. Вдруг какой-то ученый (или хитрый) муж сумеет отыскать более короткий способ переноса?

Большинство опубликованных "решений" задачи Люка, наподобие раннего решения Аллардиса (Allardice) и Фрейзера (Fraser) [7], не поясняют, почему T_n должно быть $\geqslant 2T_{n-1}+1$.

Но существует ли лучшее решение? На самом деле — нет. На каком-то этапе мы вынуждены переместить самый большой диск. При этом остальные $n-1$ меньших дисков должны находиться на одном шпиле, что требует как минимум T_{n-1} переносов. Если очень хочется, то большой диск можно перекладывать с места на место и несколько раз — например, задумавшись о том, чему равно T_{n-1}. Но когда большой диск, наконец, окажется на своем месте, нужно будет переместить $n-1$ меньших дисков (которые опять должны быть одеты на один шпиль) назад на наибольший диск. Это также потребует выполнения T_{n-1} переносов. Следовательно,

$$T_n \geqslant 2T_{n-1}+1 \qquad \text{при } n>0.$$

Эти два неравенства вместе с тривиальным решением для $n=0$ дают

$$\begin{aligned} T_0 &= 0; \\ T_n &= 2T_{n-1}+1 \qquad \text{при } n>0. \end{aligned} \tag{1.1}$$

(Заметим, что эти формулы дают уже известные нам значения $T_1=1$ и $T_2=3$. Наши эксперименты с малыми количествами дисков не только помогли вывести общую формулу, но и предоставили возможность проверки решения на наличие глупых ошибок. Такие проверки особенно важны при более сложных способах решения в последующих главах.)

Стоп, стоп... кажется, это слово уже встречалось...

Набор равенств типа (1.1) называется *рекуррентностью* (оно же *рекуррентное соотношение* или *рекурсивная зависимость*). Она задается граничным значением и уравнением, которое выражает общий член через предыдущие. Иногда рекуррентностью мы будем называть только это уравнение для общего члена, хотя технически для полного определения рекуррентности требуется еще и начальное значение.

Рекуррентное соотношение позволяет вычислить T_n для произвольного значения n, какого мы пожелаем. Но вряд ли кому-то понравится вычислять значение непосредственно из рекуррентного соотношения при больших значениях n — это потребует слишком много времени. Рекуррентность дает только косвенную, локальную информацию. *Решение рекуррентного соотношения* было бы гораздо более приятным и полезным. То есть было бы здóрово, если бы мы могли получить решение в аналитическом виде, которое позволяло бы быстро вычислять T_n даже для очень больших значений n. Решение в аналитическом виде позволяет легко понять, что же на самом деле представляет собой T_n.

Но как же решить рекуррентное соотношение? Один из способов — угадать правильное решение и затем доказать правильность догадки. Надеяться же при игре в "угадайку" вновь приходится на случаи небольших величин n. Давайте вычислим несколько последовательных значений: $T_3 = 2\cdot3+1 = 7$; $T_4 = 2\cdot7+1 = 15$; $T_5 = 2\cdot15+1 = 31$; $T_6 = 2\cdot31+1 = 63$. Ага! Определенно это что-то напоминает: этот ряд выглядит так, как если бы

$$T_n = 2^n - 1 \qquad \text{при } n \geqslant 0. \tag{1.2}$$

По крайней мере это справедливо при $n \leqslant 6$.

Математическая индукция — это общий способ доказательства истинности некоторого утверждения для целочисленных значений $n \geqslant n_0$. Сначала мы доказываем утверждение для наименьшего значения n, равного n_0. Это называется *базой* (*базисом*, *основанием* (*basis*)) индукции. Затем мы доказываем утверждение для $n > n_0$ в предположении, что оно уже доказано для всех значений от n_0 до $n-1$ включительно; это называется *шагом индукции* или просто *индукцией*. Такое доказательство позволяет получить бесконечное количество результатов при помощи конечного количества работы.

Математическая индукция доказывает, что по лестнице можно взобраться на любую, сколь угодно высокую ступеньку — путем доказательства того, что можно стать на нижнюю ступеньку (база) и того, что с любой ступеньки можно подняться на следующую (шаг индукции).

Рекуррентность и математическая индукция идеально подходят друг другу. Например, в нашем случае (1.2) легко следует из (1.1): основание индукции тривиально, поскольку $T_0 = 2^0-1 = 0$. Шаг индукции выполняется для $n > 0$, если предположить справедливость (1.2), когда n заменяется на $n-1$:

$$T_n = 2T_{n-1}+1 = 2(2^{n-1}-1)+1 = 2^n-1\,.$$

Следовательно, (1.2) выполняется и для n. Отлично! Мы успешно завершили поиск выражения для T_n.

Что же касается задачи браминов, то она далеко не завершена — они продолжают упорно трудиться, и вряд ли им доведется отдохнуть в ближайшие тысячелетия: при $n = 64$ требуется сделать $2^{64} - 1$ перенос дисков, что составляет ни много, ни мало:

$$18\,446\,744\,073\,709\,551\,615,$$

т.е. больше 18 квинтиллионов. Если даже брамины будут перемещать диски с нечеловеческой скоростью, один диск за микросекунду, им потребуется более 5 000 веков для перемещения башни Брахмы. Головоломка Люка практичнее и человечнее — при определенной ловкости рук необходимые $2^8 - 1 = 255$ переносов можно выполнить минуты за четыре.

Рекуррентность Ханойской башни типична для многих задач, возникающих в приложениях разного рода. В процессе поиска аналитического выражения интересующей величины наподобие T_n мы проходим три стадии.

1 Рассмотрение малых значений. Это позволяет понять задачу и помогает нам на стадиях 2 и 3.

Знаете ли вы, что слово "proof" не только переводится как "доказательство", но и означает меру крепости алкогольных напитков (0,5%)?

2 Поиск и доказательство математического выражения интересующей нас величины. В случае Ханойской башни это рекуррентность (1.1), которая позволяет вычислить T_n для любого n.

3 Поиск и доказательство аналитического выражения для полученного рекуррентного соотношения. В случае Ханойской башни это решение рекуррентности (1.2).

В этой книге на первое место выходит третья стадия; именно ей будет уделяться основное внимание. Фактически зачастую стадии 1 и 2 будут полностью опущены, поскольку математическое выражение будет задано в качестве исходной точки. Но даже тогда мы будем встречаться с подзадачами, решение которых проходит через все три стадии.

Наш анализ задачи о Ханойской башне привел к корректному ответу, но слабым его местом является "индуктивный скачок", когда мы полагались на везение при угадывании ответа. Одна из основных целей данной книги — пояснить, каким образом можно решать рекуррентные соотношения, *не* обладая сверхъестественными способностями. Например, рекуррентное соотношение (1.1) можно упростить, прибавляя 1 к обоим сторонам уравнений:

$$\begin{aligned} &T_0 + 1 = 1\,; \\ &T_n + 1 = 2T_{n-1} + 2 \qquad \text{при } n > 0. \end{aligned}$$

Теперь, если положить $U_n = T_n + 1$, получим

Интересно, что от +1 в (1.1) можно избавиться путем прибавления, а не вычитания.

$$\begin{aligned} U_0 &= 1\,; \\ U_n &= 2U_{n-1} \qquad \text{при } n > 0. \end{aligned} \tag{1.3}$$

Не надо быть гением, чтобы обнаружить, что решение *этого* рекуррентного соотношения имеет вид $U_n = 2^n$; следовательно, $T_n = 2^n - 1$. Даже компьютер в состоянии решить такую простую задачу.

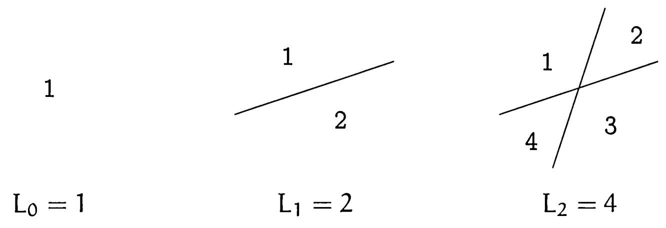
1.2 Прямые на плоскости

Вторая задача явно пахнет геометрией: на сколько кусков можно разрезать блин при помощи n прямолинейных разрезов ножом? Или, говоря более академичным языком: каково максимальное количество L_n областей, на которые можно разделить плоскость n прямыми? Эта задача впервые была решена в 1826 году швейцарским математиком Якобом Штайнером (Jacob Steiner) [338].

Пиццу, о которой идет речь в оригинале, разрезать ровно гораздо сложнее, чем блин...
— Переводчик

Интересно, разреза́л ли он швейцарский сыр?

Мы вновь начнем с рассмотрения малых случаев, помня, что начинать надо с абсолютного нуля. Плоскость без прямых — это единая область; одна прямая делит плоскость на две области, а две прямые — на четыре:

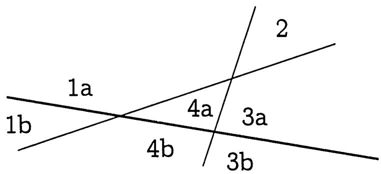


(Каждая прямая неограниченно продолжается в обоих направлениях.)

Этого достаточно, чтобы высказать догадку о том, что $L_n = 2^n$. Похоже, что каждая новая прямая просто удваивает количество областей. Увы, это не так — первый блин оказался комом... Удвоение получалось бы в том случае, если бы новая n-я прямая рассекала каждую без исключения старую область на две; на большее количество частей рассечь старую область невозможно из-за того, что все старые области выпуклы. (Прямая линия может разделить выпуклую область не более чем на две новые области, которые, в свою очередь, также будут выпуклыми.) Но когда мы добавим третью прямую — толстую прямую на приведенном ниже рисунке, — то увидим, что она может рассечь не

Область выпукла, если она включает прямолинейные отрезки между любыми двумя ее точками. (В моем словаре сказано иначе, но давайте поверим математикам.)

более трех старых областей, как бы ни располагались две первые прямые:



Таким образом, $L_3 = 4+3 = 7$ — наилучшее, чего можно достичь.

Если немного подумать, то можно найти и подходящее обобщение. Новая n-я прямая ($n > 0$) увеличивает количество областей на k тогда и только тогда, когда она рассекает k старых областей, а это происходит тогда и только тогда, когда новая прямая пересекает старые прямые в $k-1$ различных местах. Две прямые могут пересекаться не более чем в одной точке. Таким образом, новая прямая не может пересечь $n-1$ старых прямых более чем в $n-1$ разных точках, и должно выполняться соотношение $k \leqslant n$. Итак, мы установили верхнюю границу

Не будем вдаваться в тонкости неевклидовых геометрий...
— Переводчик

$$L_n \leqslant L_{n-1} + n \qquad \text{при } n > 0.$$

Более того, при помощи индукции легко показать, что в этой формуле можно достичь равенства. Разместим n-ю прямую так, чтобы она не была параллельна ни одной из остальных линий (а значит, пересекала каждую из них), причем она не должна проходить ни через одну существующую точку пересечения прямых (а значит, она пересекает все существующие прямые в разных точках). При этом рекуррентное соотношение принимает вид

$$\begin{aligned} &L_0 = 1\,; \\ &L_n = L_{n-1} + n \qquad \text{при } n > 0. \end{aligned} \tag{1.4}$$

Уже известные нам значения L_1, L_2 и L_3 отлично укладываются в это соотношение.

Теперь нужно получить решение в аналитическом виде. Можно опять сыграть в "угадайку", но что-то 1, 2, 4, 7, 11, 16... не вызывает никаких ассоциаций. Пожалуй, самое время попробовать другой подход. Часто можно понять рекуррентность, "разворачивая", или "разматывая", ее до конца следующим образом:

Разворачивая? Скорее уж "подставляя".

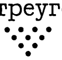
$$\begin{aligned} L_n &= L_{n-1} + n = \\ &= L_{n-2} + (n-1) + n = \\ &= L_{n-3} + (n-2) + (n-1) + n = \end{aligned}$$

$$= L_0 + 1 + 2 + \cdots + (n-2) + (n-1) + n =$$
$$= 1 \ + \ S_n\,, \qquad \text{где } S_n = 1 + 2 + 3 + \cdots + (n-1) + n.$$

Другими словами, L_n на единицу больше суммы S_n первых n натуральных чисел.

Поскольку с величиной S_n нам придется сталкиваться еще много раз, сто́ит составить таблицу ее начальных значений, чтобы легче было ее распознать, встретившись с ней в следующий раз:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
S_n	1	3	6	10	15	21	28	36	45	55	66	78	91	105

Эти числа называются также *треугольными*, поскольку S_n представляет собой количество биллиардных шаров, расположенных в виде треугольника из n рядов. Например, четырехрядная расстановка  состоит из $S_4 = 10$ шаров.

Чтобы вычислить S_n, можно прибегнуть к хитрости, которую, как говорят, Гаусс (Gauss) открыл для себя в 1786 году, когда ему было всего девять лет [88] (и которой воспользовался еще Архимед в предложениях 10 и 11 своего классического трактата о спиралях):

Гауссу приписывают столько, что непонятно, кто именно гений — он или его пресс-секретарь.

А может, он просто был притягательной личностью?

$$\begin{array}{rcccccccccc}
S_n & = & 1 & + & 2 & + & 3 & + \cdots + & (n-1) & + & n \\
+S_n & = & n & + & (n-1) & + & (n-2) & + \cdots + & 2 & + & 1 \\
\hline
2S_n & = & (n+1) & + & (n+1) & + & (n+1) & + \cdots + & (n+1) & + & (n+1)
\end{array}$$

Мы просто складываем S_n с самим собой, но записанным в обратном порядке, так что сумма каждого из n столбцов равна $n+1$. Упрощая, получаем

$$S_n = \frac{n(n+1)}{2} \qquad \text{при } n \geqslant 0. \tag{1.5}$$

Итак, вот окончательное решение:

Приятно, что мы смогли разобраться хотя бы в одном открытии такого выдающегося математика, как Гаусс.

$$L_n = \frac{n(n+1)}{2} + 1 \qquad \text{при } n \geqslant 0. \tag{1.6}$$

Как специалисты высокого класса мы могли бы удовлетвориться этим выводом и считать его доказательством, невзирая на разворачивание и обращение. Но студенты — дело другое, они обязаны отвечать куда более строгим стандартам, поэтому у нас есть отличный повод построить строгое доказательство методом математической индукции. Вот ключевой шаг индукции:

$$L_n = L_{n-1} + n = \left(\tfrac{1}{2}(n-1)n + 1\right) + n = \tfrac{1}{2}n(n+1) + 1\,.$$

Теперь у нас не остается никаких сомнений в справедливости решения в аналитическом виде (1.6).

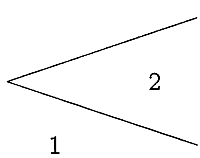
Чем "замкнутое" отличается от "открытого"? Что мы должны представить, услышав эти слова?
Ответ: уравнение "замкнуто", если оно самодостаточно, и не выражается через себя же, т.е. не приводит к рекуррентности. "Вопрос закрыт" означает, что он не возникнет вновь. Иногда метафоры дают нам ключ к пониманию вещей.

Кстати, мы все время говорим о решении в аналитическом виде (или о "замкнутой форме" решения (closed form)), но пока что не дали точного определения этого термина. Обычно все понятно без лишних слов. Рекуррентности наподобие (1.1) и (1.4) не являются замкнутой формой, поскольку выражают некоторую величину через самих себя, в отличие от решений (1.2) и (1.6). Сумма вида $1 + 2 + \cdots + n$ не является закрытой формой из-за наличия '$\cdots$'; но выражение наподобие $n(n+1)/2$ представляет собой замкнутую форму. Можно дать такое грубое определение: выражение для величины $f(n)$ имеет аналитический вид (находится в замкнутой форме), если мы можем вычислить его при помощи фиксированного количества "хорошо известных" стандартных операций, не зависящего от n. Например, выражения $2^n - 1$ и $n(n+1)/2$ представлены в закрытой форме, поскольку включают только сложение, вычитание, умножение, деление и возведение в степень в явном виде.

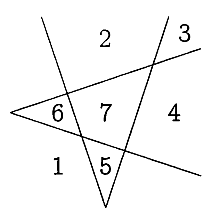
Общее количество простых замкнутых форм ограничено, так что существуют рекуррентные соотношения, не представимые в аналитическом виде. Однако если такие рекуррентные соотношения важны, то можно пополнить множество стандартных операций новыми; это может существенно расширить диапазон задач, решаемых в аналитическом виде. Например, произведение первых n натуральных чисел, $n!$, оказалось настолько важным, что теперь рассматривается как базовая операция. Таким образом, '$n!$' представляет собой замкнутую форму, несмотря на то, что ее эквивалент '$1 \cdot 2 \cdot \ldots \cdot n$' таковой не является.

Достаточно ли научно звучит термин "зиг"?

А теперь бегло рассмотрим вариацию задачи об угощении блином: предположим, что вместо прямых линий используются ломаные, каждая из которых содержит один "зиг". Чему в этом случае равно максимальное количество Z_n областей, на которые разделяют плоскость n таких ломаных? Можно ожидать, что Z_n будет раза в два (ну, или в три) больше L_n. Давайте подсчитаем:



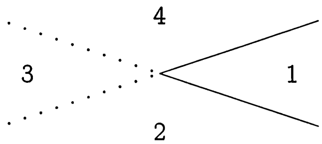
$Z_1 = 2$



$Z_2 = 7$

Из этих малых случаев — после небольшого раздумья — мы приходим к выводу, что ломаная линия похожа на две прямые, с тем отличием, что там, куда после пересечения *не* продолжаются "две" прямые, области сливаются:

...а может, и большого...



Области 2, 3 и 4, которые были бы отдельными областями при наличии двух прямых, становятся одной областью в случае одной ломаной, т.е. мы просто теряем две области. Однако при корректном размещении линий — когда точка излома находится "за" пересечениями с другими линиями — все, что мы теряем, — это две области на одну линию. Таким образом,

Подробности можно найти в упр. 18.

$$
\begin{aligned}
Z_n \;=\; L_{2n} - 2n \;&=\; 2n(2n+1)/2 + 1 - 2n \;= \\
&=\; 2n^2 - n + 1 \qquad \text{при } n \geqslant 0. \qquad (1.7)
\end{aligned}
$$

Сравнивая аналитические решения (1.6) и (1.7), мы находим, что для бо́льших n

$$
\begin{aligned}
L_n &\sim \tfrac{1}{2}n^2\,, \\
Z_n &\sim 2n^2\,;
\end{aligned}
$$

так что ломаные линии дают примерно в четыре раза больше областей, чем прямые. (В последующих главах мы рассмотрим, как анализировать приближенное поведение целочисленных функций при бо́льших значениях аргумента n. Символ '$\sim$' будет определен в разделе 9.1.)

1.3 Задача Иосифа Флавия

Последний из вводных примеров представляет собой вариацию древней задачи*, носящей имя Иосифа Флавия, известного историка I века. Легенда гласит, что Иосиф не был бы знаменитым, историком и вообще живым, если бы не его математические таланты. В ходе Иудейско-римской войны он в соста-

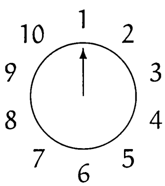
(Увлекательную историю этой задачи обсуждают Аренс (Ahrens) [5, т. 2] и Герштейн (Herstein) и Капланск (Kaplansky) [187]. Сам Иосиф [197] несколько невнятен в описании своих похождений.)

* Описание происхождения этой задачи и множества ее вариантов (средневекового, восточного и др.) можно найти в [20]. — *Примеч. пер.*

ве отряда из 41 иудейского воина был загнан римскими войсками в пещеру, выход из которой был один — в плен. Предпочтя смерть позору плена, воины избрали нетривиальный способ самоубийства — выстроиться в круг и убивать каждого третьего из живых до тех пор, пока не погибнут все. Но Иосиф и его единомышленник сочли такую гибель бессмысленной. Отдавая себе отчет, что в случае прямого выступления против планов воинов первыми падут он со своим единомышленником, Иосиф придумал план получше. Подсчитав спасительные места в круге, он сумел поставить на них себя и своего товарища.

...и только поэтому данная история дошла до наших дней.

Наша версия задачи начинается с выстраивания в круг n человек, пронумерованных от 1 до n, и удалении каждого *второго* из оставшихся в круге до тех пор, пока не останется только один человек. Вот, например, как выглядит начальная конфигурация при $n = 10$:



Порядок исключения — 2, 4, 6, 8, 10, 3, 7, 1, 9, так что остается номер 5. Задача состоит в том, чтобы определить номер уцелевшего $J(n)$.

В этой задаче начать рассмотрение с нуля не получится...

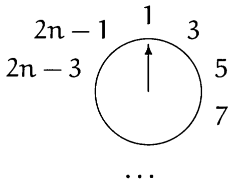
Только что мы видели, что $J(10) = 5$. Можно было бы предположить, что $J(n) = n/2$ при четных n; случай $n = 2$ тоже подтверждает это предположение: $J(2) = 1$. Увы, но другие малые случаи (например, значения для $n = 4$, $n = 6$ и $n = 8$) опровергают это предположение.

n	1	2	3	4	5	6	7	8	9
$J(n)$	1	1	3	1	3	5	7	1	3

Придется вернуться к схеме и попытаться найти решение получше... Ну-ка, ну-ка... Похоже, $J(n)$ всегда нечетное. Кстати, тому есть веская причина: первый проход по кругу исключает все четные номера. Кроме того, если n четное, то мы окажемся в ситуации, подобной начальной, однако людей останется в два раза меньше, и изменятся их номера.

Отрицательный результат — тоже результат, поскольку он помогает глубже проникнуть в задачу.

Итак, предположим, что изначально было $2n$ людей. После первого прохода по кругу в нем останутся номера



и следующим исключаемым номером будет 3. Задача выглядит так, как если бы у нас имелось n человек, но номер каждого при этом был удвоен и уменьшен на 1:

$$J(2n) \;=\; 2J(n) - 1 \qquad \text{при } n \geqslant 1.$$

Теперь можно быстро перейти к большим значениям n. Например, мы знаем, что $J(10) = 5$, так что

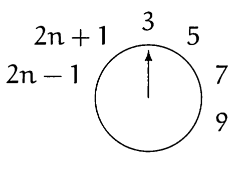
Вот в чем хитрость: $J(2n) = No(J(n))$, *где* $No(k) = 2k - 1$.

$$J(20) \;=\; 2J(10) - 1 \;=\; 2\cdot 5 - 1 \;=\; 9\,.$$

Аналогично $J(40) = 17$, и вообще $J(5\cdot 2^m) = 2^{m+1} + 1$.

Ну, а как же быть с нечетными значениями? При наличии $2n + 1$ человек номер 1 становится жертвой сразу после номера $2n$, и мы оказываемся в ситуации

А может, нечет — не в счет?..



Мы опять получили почти исходную ситуацию с n людьми, но на этот раз номера уцелевших удваиваются и *уменьшаются* на 1. Таким образом,

$$J(2n+1) \;=\; 2J(n) + 1 \qquad \text{при } n \geqslant 1.$$

Объединяя эти уравнения с $J(1) = 1$, мы получим рекуррентные соотношения, определяющие J для всех значений:

$$\begin{aligned} J(1) &= 1\,; \\ J(2n) &= 2J(n) - 1 \qquad \text{при } n \geqslant 1; \\ J(2n+1) &= 2J(n) + 1 \qquad \text{при } n \geqslant 1. \end{aligned} \tag{1.8}$$

Вместо того чтобы получать $J(n)$ из $J(n-1)$, эти соотношения предоставляют в наше распоряжение куда более эффективную вычислительную схему, поскольку при вычислениях n на каждом шаге уменьшается в 2 или более раз. Так, значение $J(1000000)$ можно вычислить при помощи всего лишь 19 применений (1.8).

Но мы все равно ищем решение в аналитическом виде, поскольку оно должно быть не только более быстрым, но и более информативным. В конце концов, не забывайте, что речь идет о жизни и смерти!

Полученное нами рекуррентное соотношение позволяет очень быстро построить таблицу для небольших значений n. Может, мы сумеем увидеть в ней какую-то закономерность и угадать ответ.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$J(n)$	1	1	3	1	3	5	7	1	3	5	7	9	11	13	15	1

Вот! Похоже, можно сгруппировать значения по степеням 2 (что в таблице показано при помощи вертикальных линий); $J(n)$ всегда равно 1 в начале группы и на каждом шаге увеличивается на 2 в пределах группы. Так что если записать n в виде $n = 2^m + l$, где 2^m — наибольшая степень 2, не превосходящая n, а l — разность между n и этой степенью 2, то, похоже, решение нашего рекуррентного соотношения можно записать следующим образом:

$$J(2^m + l) \;=\; 2l + 1 \qquad \text{при } m \geqslant 0 \text{ и } 0 \leqslant l < 2^m. \tag{1.9}$$

(Обратите внимание, что если $2^m \leqslant n < 2^{m+1}$, то остаток $l = n - 2^m$ удовлетворяет неравенству $0 \leqslant l < 2^{m+1} - 2^m = 2^m$.)

Мы должны доказать (1.9). Как и ранее, прибегнем к математической индукции, но в этот раз к индукции по m. При $m = 0$ мы получаем единственное возможное значение $l = 0$; таким образом, база индукции (1.9) сводится к $J(1) = 1$, что, очевидно, верно. Шаг индукции состоит из двух частей, в зависимости от того, четное l или нет. Если $m > 0$ и $2^m + l = 2n$, то l четное и

$$J(2^m + l) \;=\; 2J(2^{m-1} + l/2) - 1 \;=\; 2(2l/2 + 1) - 1 \;=\; 2l + 1$$

в соответствии с (1.8) и гипотезой индукции; это именно то, что мы хотим получить. Аналогичное доказательство работает в нечетном случае, когда $2^m + l = 2n + 1$. Можно также заметить, что из (1.8) вытекает соотношение

$$J(2n + 1) - J(2n) \;=\; 2.$$

В любом случае индукция завершена и (1.9) доказано.

Проиллюстрируем решение (1.9) путем вычисления $J(100)$. В этом случае $100 = 2^6 + 36$, так что $J(100) = 2\cdot 36 + 1 = 73$.

Теперь, когда сделано самое трудное (решена задача), немного задержимся перед тем, как идти дальше: каждое решение

Есть и более простой путь! Ключевой факт заключается в том, что $J(2^m) = 1$ для всех m, а это непосредственно следует из нашего первого уравнения $J(2n) = 2J(n) - 1$. Следовательно, мы знаем, что первый человек спасется, если n представляет собой степень 2. А в общем случае, когда $n = 2^m + l$, количество людей сокращается до степени 2 после l исключений. Первый из оставшихся после этого человек — тот, кто останется в живых — имеет номер $2l+1$.

задачи можно обобщить так, что оно будет применимо к более широкому классу задач. Раз уж мы изучили некоторый метод, будет поучительным познакомиться с ним поближе и посмотреть, как далеко он может нас завести. Поэтому в оставшейся части раздела мы изучим решение (1.9) и исследуем некоторые обобщения рекуррентного соотношения (1.8). Эти исследования выявят структуру, лежащую в основе всех задач такого рода.

Степени 2 играют важную роль в нашем поиске решения, так что вполне естественно рассмотреть двоичные представления n и $J(n)$. Предположим, что двоичное представление n's имеет вид

$$n = (b_m\, b_{m-1} \ldots b_1\, b_0)_2\,,$$

т.е.

$$n = b_m 2^m + b_{m-1} 2^{m-1} + \cdots + b_1 2 + b_0\,,$$

где каждое b_i равно 0 или 1, причем старший бит b_m равен 1. Вспоминая, что $n = 2^m + l$, последовательно получаем

$$\begin{aligned} n &= (1\, b_{m-1}\, b_{m-2} \ldots b_1\, b_0)_2\,,\\ l &= (0\, b_{m-1}\, b_{m-2} \ldots b_1\, b_0)_2\,,\\ 2l &= (b_{m-1}\, b_{m-2} \ldots b_1\, b_0\, 0)_2\,,\\ 2l+1 &= (b_{m-1}\, b_{m-2} \ldots b_1\, b_0\, 1)_2\,,\\ J(n) &= (b_{m-1}\, b_{m-2} \ldots b_1\, b_0\, b_m)_2\,. \end{aligned}$$

(Последний шаг следует из того, что $J(n) = 2l+1$ и $b_m = 1$.) Мы доказали, что

$$J\big((b_m\, b_{m-1} \ldots b_1\, b_0)_2\big) = (b_{m-1} \ldots b_1\, b_0\, b_m)_2\,; \qquad (1.10)$$

т. е., говоря языком компьютерного программирования, $J(n)$ получается из n путем циклического сдвига влево на один бит! Чудеса... Так, если $n = 100 = (1100100)_2$, то $J(n) = J\big((1100100)_2\big) = (1001001)_2$, что равно $64+8+1 = 73$. Наверняка, работая в двоичной системе счисления изначально, мы бы сразу обратили внимание на данную закономерность.

Если начать с n и $m+1$ раз итерировать функцию J, то будет выполнено $m+1$ однобитовых циклических сдвигов, так что, поскольку n представляет собой $(m+1)$-битовое число, казалось бы, мы должны вновь получить число n. Но это не совсем так. Например, если $n = 13$, то $J\big((1101)_2\big) = (1011)_2$, но далее $J\big((1011)_2\big) = (111)_2$ и процесс обрывается: когда 0 становится

“Итерация” в данном случае означает применение функции к себе самой.

ведущим битом, он исчезает и более не рассматривается. По определению $J(n)$ всегда $\leqslant n$, поскольку $J(n)$ — номер уцелевшего человека; следовательно, если $J(n) < n$, мы никогда не вернемся к n в последующих итерациях.

Многократное применение J приводит к последовательности убывающих значений, в конечном счете достигающей "фиксированной точки", в которой $J(n) = n$. Свойство циклического сдвига позволяет легко выяснить, что это за точка: при достаточном количестве итераций всегда будет получаться значение, состоящее из одних единиц в двоичной записи, равное $2^{\nu(n)} - 1$, где $\nu(n)$ — количество битов 1 в двоичном представлении n. Таким образом, поскольку $\nu(13) = 3$, мы получаем

$$\overbrace{J(J(\ldots J}^{2}(13)\ldots)) \;=\; 2^3 - 1 \;=\; 7;$$

Довольно забавно: если M — компактное n-мерное многообразие C^∞ ($n > 1$), то существует дифференцируемое погружение M в $\mathbf{R}^{2n-\nu(n)}$, но необязательно в $\mathbf{R}^{2n-\nu(n)-1}$. Не удивлюсь, если узнаю, что Иосиф — секретный тополог...

аналогично

$$\overbrace{J(J(\ldots J}^{8}((101101101101011)_2)\ldots)) \;=\; 2^{10} - 1 \;=\; 1023.$$

Странно, но факт.

Вернемся к нашему первому предположению, что $J(n) = n/2$ при четных n. Очевидно, что в общем случае это неверно, но теперь мы можем определить, *когда это верно*:

$$\begin{aligned} J(n) &= n/2, \\ 2l+1 &= (2^m + l)/2, \\ l &= \tfrac{1}{3}(2^m - 2). \end{aligned}$$

Если число $l = \frac{1}{3}(2^m - 2)$ целое, то $n = 2^m + l$ является решением, поскольку l меньше 2^m. Несложно проверить, что $2^m - 2$ делится на 3 при нечетном m и не делится, когда m четное. (Такие вещи мы будем изучать в главе 4.) Поэтому имеется бесконечно много решений уравнения $J(n) = n/2$, начиная со следующих:

m	l	$n = 2^m + l$	$J(n) = 2l + 1 = n/2$	n (двоичное)
1	0	2	1	10
3	2	10	5	1010
5	10	42	21	101010
7	42	170	85	10101010

Обратите внимание на вид чисел в крайнем справа столбце. Это двоичные числа, для которых циклический сдвиг на один бит влево дает тот же результат, что и сдвиг на один бит вправо (деление пополам).

Пожалуй, с функцией J мы разобрались в полной мере; следующий шаг — ее обобщение. Что было бы, если бы в нашей задаче получилось рекуррентное соотношение, подобное (1.8), но с другими константами? Вряд ли бы нам повезло угадать решение, которое могло бы быть воистину причудливым. Давайте исследуем случай с константами α, β и γ и попытаемся найти в аналитическом виде решение более общего рекуррентного соотношения

Какая-то Гомеровская грамота...

$$\begin{aligned} f(1) &= \alpha; \\ f(2n) &= 2f(n) + \beta \qquad \text{при } n \geqslant 1; \\ f(2n+1) &= 2f(n) + \gamma \qquad \text{при } n \geqslant 1. \end{aligned} \tag{1.11}$$

(Наше исходное рекуррентное соотношение получается при $\alpha = 1$, $\beta = -1$ и $\gamma = 1$.) Начнем с $f(1) = \alpha$ и, работая так же, как и раньше, построим таблицу для малых значений n:

n	$f(n)$
1	α
2	$2\alpha + \beta$
3	$2\alpha + \gamma$
4	$4\alpha + 3\beta$
5	$4\alpha + 2\beta + \gamma$
6	$4\alpha + \beta + 2\gamma$
7	$4\alpha + 3\gamma$
8	$8\alpha + 7\beta$
9	$8\alpha + 6\beta + \gamma$

(1.12)

Похоже, что коэффициенты при α представляют собой наибольшие степени 2, не превосходящие n. Кроме того, между степенями 2 значения коэффициентов при β уменьшаются на 1 вплоть до 0, а при γ — увеличиваются на 1, начиная с 0. Таким образом, если выразить $f(n)$ в виде

$$f(n) = A(n)\alpha + B(n)\beta + C(n)\gamma, \tag{1.13}$$

отделяя зависимости от α, β и γ друг от друга, похоже, что

$$\begin{aligned} A(n) &= 2^m; \\ B(n) &= 2^m - 1 - l; \\ C(n) &= l. \end{aligned} \tag{1.14}$$

Здесь, как обычно, $n = 2^m + l$ и $0 \leqslant l < 2^m$ при $n \geqslant 1$.

Дальше будет кое-что поновее.

Не так уж трудно доказать по индукции (1.13) и (1.14), но в этом мало как системы, так и информации. К счастью, имеется лучший способ, состоящий в выборе отдельных значений и их комбинации. Проиллюстрируем его на примере частного случая $\alpha = 1$, $\beta = \gamma = 0$, когда $f(n)$ предполагается равной $A(n)$. В этом случае рекуррентные соотношения (1.11) приводятся к виду

$$\begin{aligned} A(1) &= 1\,; \\ A(2n) &= 2A(n) \qquad \text{при } n \geqslant 1; \\ A(2n+1) &= 2A(n) \qquad \text{при } n \geqslant 1. \end{aligned}$$

Достаточно ясно (по индукции по m), что $A(2^m + l) = 2^m$.

Теперь воспользуемся рекуррентным соотношением (1.11) и решением (1.13) *в обратном порядке*, начав с простой функции $f(n)$ и выясняя, нет ли каких-то констант (α, β, γ), которые ее определяют. Подстановка константной функции $f(n) = 1$ в (1.11) говорит, что

Изящная идея!

$$\begin{aligned} 1 &= \alpha, \\ 1 &= 2\cdot 1 + \beta, \\ 1 &= 2\cdot 1 + \gamma, \end{aligned}$$

следовательно, значения $(\alpha, \beta, \gamma) = (1, -1, -1)$, удовлетворяющие этим уравнениям, дают $A(n) - B(n) - C(n) = f(n) = 1$. Аналогично можно подставить $f(n) = n$:

$$\begin{aligned} 1 &= \alpha, \\ 2n &= 2\cdot n + \beta, \\ 2n+1 &= 2\cdot n + \gamma. \end{aligned}$$

Эти уравнения выполняются для всех n, если $\alpha = 1$, $\beta = 0$ и $\gamma = 1$, так что нам не надо доказывать по индукции, что при данных значениях параметров $f(n) = n$. Мы уже *знаем*, что $f(n) = n$ является решением в данном частном случае, поскольку рекуррентное соотношение (1.11) однозначно определяет $f(n)$ для каждого значения n.

По сути, все сделано! Мы показали, что функции $A(n)$, $B(n)$ и $C(n)$ из (1.13), которые решают (1.11) в общем случае, удовлетворяют уравнениям

$$\begin{aligned} A(n) &= 2^m\,, \qquad \text{где } n = 2^m + l \text{ и } 0 \leqslant l < 2^m, \\ A(n) - B(n) - C(n) &= 1\,, \\ A(n) + C(n) &= n\,. \end{aligned}$$

Наши предположения в (1.14) подтверждаются немедленно, поскольку можно решить эти уравнения и получить $C(n) = n - A(n) = l$ И $B(n) = A(n) - 1 - C(n) = 2^m - 1 - l$.

Этот подход иллюстрирует на удивление полезный *метод наборов* (*repertoire method*) для решения рекуррентных соотношений. Сначала находятся значения общих параметров, для которых решения известны. Это обеспечивает нас набором частных случаев, которые мы в состоянии решить. Затем общий случай получается путем комбинирования частных. Нам надо столько независимых частных решений, сколько и независимых параметров (в данном случае их три — α, β и γ). Другие примеры данного подхода вы найдете в упражнениях 16 и 20.

Внимание! Авторы надеются, что читатель поймет идею метода из этих доморощенных примеров, и не излагают его последовательно, "от и до". Лучше всего данный метод работает с "линейными" рекуррентными соотношениями (в том смысле, что их решения могут быть выражены в виде суммы произвольных параметров, умноженных на функции от n, как в (1.13). В этом смысле уравнение (1.13) является ключевым.

Мы знаем, что исходная J-рекуррентность имеет магическое решение в двоичной записи:

$$J\big((b_m\, b_{m-1} \ldots b_1\, b_0)_2\big) \;=\; (b_{m-1} \ldots b_1\, b_0\, b_m)_2\,, \quad \text{где } b_m = 1.$$

Допускает ли подобные чудеса обобщенная рекуррентность Иосифа?

Конечно, почему бы и нет? Можно переписать обобщенную рекуррентность (1.11) в виде

$$\begin{aligned} f(1) &= \alpha\,, \\ f(2n+j) &= 2f(n) + \beta_j \qquad \text{при } j = 0, 1 \quad \text{и} \quad n \geqslant 1\,, \end{aligned} \tag{1.15}$$

если положить $\beta_0 = \beta$ и $\beta_1 = \gamma$. Эта рекуррентность побитово развертывается следующим образом:

$$\begin{aligned} f\big((b_m\, b_{m-1} \ldots b_1\, b_0)_2\big) &= 2f\big((b_m\, b_{m-1} \ldots b_1)_2\big) + \beta_{b_0} \\ &= 4f\big((b_m\, b_{m-1} \ldots b_2)_2\big) + 2\beta_{b_1} + \beta_{b_0} \\ &\;\;\vdots \\ &= 2^m f\big((b_m)_2\big) + 2^{m-1}\beta_{b_{m-1}} + \cdots + 2\beta_{b_1} + \beta_{b_0} \\ &= 2^m \alpha + 2^{m-1}\beta_{b_{m-1}} + \cdots + 2\beta_{b_1} + \beta_{b_0}\,. \end{aligned}$$

Предположим, что теперь мы так расширили двоичную запись, что в ней допустимы любые цифры, а не только 0 и 1. Предыдущий вывод говорит нам, что

(Это 'расширение' ≡ 'уничтожению')

$$f\big((b_m\, b_{m-1} \ldots b_1\, b_0)_2\big) \;=\; (\alpha\, \beta_{b_{m-1}}\, \beta_{b_{m-2}} \ldots \beta_{b_1}\, \beta_{b_0})_2\,. \tag{1.16}$$

Отлично! Это можно было бы увидеть и раньше, если бы мы с самого начала записали таблицу (1.12) иначе:

Кажется, я понял: в двоичных представлениях $A(n)$, $B(n)$ и $C(n)$ единицы находятся в разных позициях.

n	$f(n)$
1	α
2	$2\alpha + \beta$
3	$2\alpha + \gamma$
4	$4\alpha + 2\beta + \beta$
5	$4\alpha + 2\beta + \gamma$
6	$4\alpha + 2\gamma + \beta$
7	$4\alpha + 2\gamma + \gamma$

Например, для $n = 100 = (1100100)_2$ исходная рекуррентность Иосифа $\alpha = 1$, $\beta = -1$ и $\gamma = 1$, как и ранее, дает

$$\begin{array}{rccccccccc}
n = & (\,1 & 1 & 0 & 0 & 1 & 0 & 0\,)_2 & = & 100 \\
\hline
f(n) = & (\,1 & 1 & -1 & -1 & 1 & -1 & -1\,)_2 & & \\
= & +64 & +32 & -16 & -8 & +4 & -2 & -1 & = & 73.
\end{array}$$

Свойство циклического сдвига сохраняется, поскольку каждый блок двоичных цифр $(1\,0\,\ldots\,0\,0)_2$ в представлении n преобразуется в

$$(1\,{-1}\,\ldots\,{-1}\,{-1})_2 = (0\,0\,\ldots\,0\,1)_2\,.$$

“Существует два вида обобщений. Одно — подешевле, другое — подороже. Легко обобщать, разбавляя небольшую идею большими словами. Гораздо сложнее приготовить очищенный концентрат из нескольких хороших ингредиентов.”
— Г. Пойа (G. Pólya) [297]

Наше изменение записи дает нам компактное решение (1.16) обобщенной рекуррентности (1.15). Если очень хочется, можно обобщить ее еще больше. Рекуррентность

$$\begin{aligned}
f(j) &= \alpha_j && \text{при } 1 \leqslant j < d;\\
f(dn + j) &= cf(n) + \beta_j && \text{при } 0 \leqslant j < d \quad \text{и} \quad n \geqslant 1
\end{aligned} \tag{1.17}$$

такая же, как и предыдущая, с тем отличием, что мы начинаем с чисел в системе счисления с основанием d, а получаем значения в системе счисления с основанием c. То есть данная рекуррентность имеет решение с переменным основанием

$$f\big((b_m\,b_{m-1}\ldots b_1\,b_0)_d\big) = (\alpha_{b_m}\,\beta_{b_{m-1}}\,\beta_{b_{m-2}}\ldots\beta_{b_1}\,\beta_{b_0})_c\,. \tag{1.18}$$

Пусть, например, нам несказанно повезло и мы сумели получить рекуррентное соотношение

$$\begin{aligned}
f(1) &= 34\,,\\
f(2) &= 5\,,\\
f(3n) &= 10f(n) + 76 && \text{при } n \geqslant 1;\\
f(3n+1) &= 10f(n) - 2 && \text{при } n \geqslant 1,\\
f(3n+2) &= 10f(n) + 8 && \text{при } n \geqslant 1,
\end{aligned}$$

и предположим, что мы хотим вычислить $f(19)$. Здесь $d = 3$ и $c = 10$. Тогда $19 = (201)_3$, и решение с переменным основанием системы счисления состоит в выполнении перехода от системы счисления с основанием 3 к системе счисления с основанием 10 цифра за цифрой. Старшая цифра 2 становится цифрой 5, а 0 и 1 становятся 76 и -2 соответственно, что дает окончательный ответ

Возможно, это пример невезения.

$$f(19) = f\big((201)_3\big) = (5\ 76\ {-2})_{10} = 1258\,.$$

Так Иосиф и иудейско-римская война привели нас к интересным обобщенным возвратным отношениям.

В общем случае я против возврата войн.

Упражнения

Разминка

Разминка обязательна к выполнению в каждой главе!
— Администрация

1 Все лошади — одной масти. Это утверждение можно доказать по индукции по числу лошадей в данном множестве. Вот как это делается.

"Если имеется только одна лошадь, то очевидно, что она своей собственной масти, так что база индукции тривиальна. Для выполнения шага индукции предположим, что имеется n лошадей с номерами от 1 до n. Согласно гипотезе индукции лошади от первой до $n - 1$ — одинаковой масти. То же самое можно сказать о лошадях со второй по n-ю. Лошади со второй по $n - 1$-ю не в состоянии менять масть, переходя из группы в группу (это же не хамелеоны!), поэтому в силу транзитивности лошади с первой по n-ю должны иметь один и тот же цвет. Следовательно, все n лошадей — одного цвета. QED*". Что в приведенном доказательстве не так?

2 Найдите кратчайшую последовательность переносов для перемещения башни из n дисков со шпиля A на шпиль B, если переносы между шпилями A и B запрещены. (Каждый перенос должен выполняться на средний шпиль или с него. Как обычно, больший диск никогда не должен располагаться поверх меньшего.)

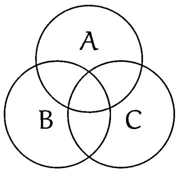
3 Покажите, что в процессе перемещения башни на условиях из предыдущего упражнения будут реализованы все корректные размещения n дисков на трех шпилях.

---

* "Quod erat demonstrandum" — "Что и требовалось доказать" (лат.). — *Примеч. пер.*

4 Имеются ли для трех шпилей некоторые начальное и конечное расположения n дисков, для которых переход из одного в другое по исходным правилам Люка требует более $2^n - 1$ переносов дисков?

5 "Диаграмма Венна" с тремя пересекающимися окружностями часто используется для иллюстрации всех восьми возможных подмножеств трех данных множеств:



Можно ли проиллюстрировать при помощи четырех пересекающихся окружностей все шестнадцать возможностей, возникающих при наличии четырех заданных множеств?

6 Одни области, определяемые n прямыми на плоскости, бесконечны, в то время как другие — конечны. Чему равно максимально возможное число конечных областей?

7 Пусть $H(n) = J(n+1) - J(n)$. Уравнение (1.8) говорит о том, что $H(2n) = 2$, а $H(2n+1) = J(2n+2) - J(2n+1) = \bigl(2J(n+1) - 1\bigr) - \bigl(2J(n) + 1\bigr) = 2H(n) - 2$ при всех $n \geqslant 1$. Поэтому кажется возможным доказательство по индукции по n того, что $H(n) = 2$ при всех n. Что здесь неверно?

Домашние задания

8 Решите рекуррентное соотношение

$$\begin{aligned} Q_0 &= \alpha; \qquad Q_1 = \beta; \\ Q_n &= (1 + Q_{n-1})/Q_{n-2} \qquad \text{при } n > 1. \end{aligned}$$

Считаем, что $Q_n \neq 0$ при всех $n \geqslant 0$. *Указание*: $Q_4 = (1+\alpha)/\beta$.

9 Иногда возможно применение индукции в обратном направлении, т.е. ведущей доказательство от n к $n-1$, а не наоборот! Рассмотрим, например, утверждение

...эта лошадь — другой масти...

$$P(n) \ : \quad x_1 \ldots x_n \leqslant \left(\frac{x_1 + \cdots + x_n}{n}\right)^n, \text{ если } x_1, \ldots, x_n \geqslant 0.$$

Оно верно для $n = 2$, поскольку $(x_1 + x_2)^2 - 4x_1x_2 = (x_1 - x_2)^2 \geqslant 0$.

а Полагая $x_n = (x_1 + \cdots + x_{n-1})/(n-1)$, докажите, что из $P(n)$ вытекает $P(n-1)$ для любого $n > 1$.

б Покажите, что из $P(n)$ и $P(2)$ следует $P(2n)$.

в Поясните, почему отсюда следует справедливость $P(n)$ для всех n.

10 Пусть Q_n — минимальное число переносов, необходимых для переноса башни из n дисков со шпиля A на шпиль B, если все переносы осуществляются *по часовой стрелке*, т.е. с A на B, с B на третий шпиль или с этого третьего шпиля на A. Пусть также R_n — минимальное количество переносов, необходимых для перемещения башни со шпиля B на шпиль A при тех же ограничениях. Докажите, что

$$Q_n = \begin{cases} 0, & \text{если } n = 0; \\ 2R_{n-1}+1, & \text{если } n > 0; \end{cases}$$
$$R_n = \begin{cases} 0, & \text{если } n = 0; \\ Q_n+Q_{n-1}+1, & \text{если } n > 0. \end{cases}$$

(Решать эти рекуррентные соотношения не надо; как это делается, вы узнаете в главе 7.)

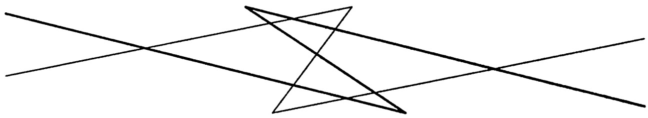
11 Двойная Ханойская башня состоит из $2n$ дисков n разных размеров, по два диска каждого размера. Как обычно, переносить можно только по одному диску за раз, и больший диск нельзя класть поверх меньшего.

а Сколько переносов требуется для перемещения двойной башни с одного шпиля на другой, если диски одного размера неотличимы друг от друга?

б Что если в окончательном расположении дисков требуется воспроизвести исходный порядок дисков в башне? (*Указание*: эту задачу решить не так-то просто — на самом деле она относится к "дополнительным задачам".)

12 Обобщим упражнение 11, а, предполагая, что имеются диски n разных размеров, причем размер k имеют m_k дисков. Определите, чему равно $A(m_1, \ldots, m_n)$ — минимальное количество переносов, необходимое для перемещения башни при условии неразличимости дисков одинакового размера.

13 На какое максимальное количество областей можно разделить плоскость n зигзагообразными линиями,



$ZZ_2 = 12$

каждая из которых состоит из двух полубесконечных прямых, соединенных прямолинейным отрезком?

Подумайте о помощнике, который будет держать сыр.

14 На сколько частей можно разрезать головку сыра пятью плоскими разрезами? (Головка в процессе разрезания должна покоиться, а каждый разрез должен представлять собой плоскость в трехмерном пространстве.) Найдите рекуррентное соотношение для P_n — максимального количества трехмерных областей, на которое может быть разбито трехмерное пространство при помощи n различных плоскостей.

15 У Иосифа был друг, спасая которого, Иосиф поставил на предпоследнее место в списке казнимых. Чему равно $I(n)$, номер предпоследнего выжившего, если казнят каждого второго?

16 Примените метод наборов для решения обобщенного рекуррентного соотношения с четырьмя параметрами:

$$\begin{aligned} g(1) &= \alpha, \\ g(2n+j) &= 3g(n) + \gamma n + \beta_j \qquad \text{при } j = 0, 1 \quad \text{и} \quad n \geqslant 1. \end{aligned}$$

Указание: испытайте функцию $g(n) = n$.

Контрольные работы

17 Пусть W_n — минимальное количество переносов, необходимых для перемещения башни из n дисков с одного шпиля на другой, когда имеется не три, а четыре шпиля. Покажите, что

$$W_{n(n+1)/2} \leqslant 2W_{n(n-1)/2} + T_n \qquad \text{при } n > 0.$$

(Здесь $T_n = 2^n - 1$ — количество переносов в случае трех шпилей.) Воспользуйтесь этим для поиска в аналитическом виде $f(n)$, такой, что $W_{n(n+1)/2} \leqslant f(n)$ при всех $n \geqslant 0$.

18 Покажите, что следующее множество из n ломаных линий делит плоскость на Z_n областей, где Z_n определено в (1.7): j-я ломаная линия ($1 \leqslant j \leqslant n$) имеет излом в точке $(n^{2j}, 0)$ и проходит через точки $(n^{2j} - n^j, 1)$ и $(n^{2j} - n^j - n^{-n}, 1)$.

19 Можно ли получить Z_n областей при использовании n ломаных линий, если угол каждого излома равен 30°?

Это уже пятизвездочная рекуррентность!

20 Воспользуйтесь методом наборов для решения обобщенной рекуррентности с пятью параметрами:

$$h(1) = \alpha,$$

$$h(2n+j) = 4h(n) + \gamma_j n + \beta_j \qquad \text{при } j = 0, 1 \text{ и } n \geqslant 1.$$

Указание: попробуйте применить функции $h(n) = n$ и $h(n) = n^2$.

21 Предположим, что имеется $2n$ человек, построенных в круг, среди которых первые n — "наши", а последние n — "не наши". Покажите, что всегда найдется целое число m (зависящее от n), такое, что если при движении по кругу будет исключаться каждый m-й человек, то сначала будут исключены все "не наши". (Например, при $n = 3$ можно взять $m = 5$; при $n = 4$ можно использовать $m = 30$.)

Дополнительные задачи

22 Покажите, что можно построить диаграмму Венна для всех 2^n возможных подмножеств n заданных множеств при помощи n конгруэнтных выпуклых многоугольников, которые повернуты на разные углы относительно общего центра.

23 Предположим, что Иосифа поставили в круг на конкретное j-е место, но позволили ему назвать параметр q, после чего уничтожается каждый q-й человек. Всегда ли Иосиф сможет спастись?

Исследовательские проблемы

24 Найдите все рекуррентные соотношения вида

$$X_n = \frac{1 + a_1 X_{n-1} + \cdots + a_k X_{n-k}}{b_1 X_{n-1} + \cdots + b_k X_{n-k}},$$

решения которых периодичны, какими бы ни были начальные значения $X_0, \ldots, X_{k-1}$.

25 Решите задачу о Ханойской башне с четырьмя шпилями для бесконечного числа случаев, доказав, что в упр. 17 имеет место равенство.

26 Обобщим упр. 23. Назовем *Иосифовым подмножеством* множества $\{1, 2, \ldots, n\}$ множество из k чисел, такое, что при некотором q первыми будут уничтожены люди с оставшимися $n - k$ номерами. (Эти k мест занимают люди, которых хочет спасти Иосиф.) Оказывается, при $n = 9$ три из 2^9 возможных подмножеств являются неиосифовыми, а именно подмножества $\{1, 2, 5, 8, 9\}$, $\{2, 3, 4, 5, 8\}$ и $\{2, 5, 6, 7, 8\}$. При $n = 12$ имеется 13 неиосифовых подмножеств; при всех прочих $n \leqslant 12$ неиосифовых подмножеств нет. Являются ли неиосифовы подмножества редкостью при больших n?

Да, и неплохо бы их найти.

2

Суммы

СУММЫ В МАТЕМАТИКЕ вездесущи, так что мы обязаны знать основные способы работы с ними. Эта глава посвящена обозначениям и методам, которые помогут сделать суммирование простым и дружественным занятием.

2.1 Обозначения

В главе 1 мы встречались с суммой первых n натуральных чисел, которую записывали как $1 + 2 + 3 + \cdots + (n-1) + n$. Многоточие в таких формулах говорит о необходимости полной записи суммы в соответствии с шаблоном, определенным окружающими членами. Само собой, следует избегать сумм наподобие $1 + 7 + \cdots + 41.7$, которые без соответствующего контекста лишены всякого смысла. С другой стороны, включение членов 3 и $(n-1)$ в определенной степени излишне; шаблон становится совершенно очевидным и при записи $1 + 2 + \cdots + n$. Иногда его можно даже сократить до $1 + \cdots + n$.

Мы будем работать с суммами общего вида

$$a_1 + a_2 + \cdots + a_n\,, \tag{2.1}$$

где каждое a_k представляет собой определенное каким-то образом число. Такая запись обладает тем преимуществом, что при достаточном воображении мы можем “увидеть” всю сумму почти так, как если бы она была записана полностью.

Членом не только почетным, но и по нечетным...

Каждый элемент суммы a_k называется ее *членом*. Зачастую члены определяются неявно, в виде формул, следующих некоторому легко улавливаемому шаблону, и в таких случаях иногда требуется записывать их в развернутом виде, чтобы их смысл был очевиден. Например, если предполагается, что формула

$$1 + 2 + \cdots + 2^{n-1}$$

описывает сумму из n, а не 2^{n-1} членов, ее следует записать более точно как

$$2^0 + 2^1 + \cdots + 2^{n-1}.$$

Запись с тремя точками используется очень часто, но она может оказаться неоднозначной и несколько громоздкой. Имеются и иные способы записи сумм, среди которых особенно выделяется запись с явным указанием пределов

$$\sum_{k=1}^{n} a_k\,, \tag{2.2}$$

которая называется сигма-обозначением, так как использует прописную греческую букву Σ. Такая запись гласит, что в сумму включаются те члены a_k, индекс которых k представляет собой целое число между нижней и верхней границами 1 и n включительно. На обычном языке это звучит как "сумма по k от 1 до n". Жозеф Фурье (Joseph Fourier) ввел это $\sum$-обозначение в 1820 году, и оно очень быстро было принято всем математическим миром.

"Le signe $\sum_{i=1}^{i=\infty}$ indique que l'on doit donner au nombre entier i toutes ses valeurs 1, 2, 3, ..., et prendre la somme des termes."
— Ж. Фурье (J. Fourier) [127]

Не хотите изучить французский язык, чтобы прочесть Фурье (и не только) в подлиннике?
— Переводчик

Кстати, величина после $\sum$ (в данном случае — a_k) называется *общим членом* (*summand*).

Говорится, что индексная переменная k *связана* со знаком $\sum$ в (2.2), поскольку k в a_k не имеет никакого отношения к k вне суммы. k можно было бы заменить любой другой буквой без потери смысла (2.2). В качестве индекса часто используется буква i (вероятно, потому что с нее начинается слово "index"), но мы в общем случае предпочитаем применять суммирование по k, считая разумным позволить i оставаться мнимой единицей $\sqrt{-1}$.

Я бы предпочел не использовать в качестве индекса a или n в (2.2), поскольку они имеют собственное значение за рамками $\sum$.

Еще более полезна обобщенная запись $\sum$ с одним или несколькими условиями под знаком суммы, указывающая множество индексов, по которым выполняется суммирование. Например, суммы в (2.1) и (2.2) могут быть записаны как

$$\sum_{1\leqslant k\leqslant n} a_k\,. \tag{2.3}$$

Этот конкретный пример не сильно отличается от (2.2), но обобщенная форма позволяет брать сумму по множествам значений индексов, не ограниченных последовательными целыми числами. Например, вот как можно выразить сумму квадратов всех нечетных натуральных чисел, меньших 100:

$$\sum_{\substack{1\leqslant k<100\\ k \text{ нечетное}}} k^2\,.$$

Аналог этой суммы с явными пределами,

$$\sum_{k=0}^{49}(2k+1)^2,$$

более громоздкий и менее очевидный. Аналогично сумма обратных значений всех простых чисел между 1 и N записывается как

$$\sum_{\substack{p\leqslant N\\ p\ \text{простое}}}\frac{1}{p};$$

запись с явными пределами имеет при этом вид

$$\sum_{k=1}^{\pi(N)}\frac{1}{p_k},$$

где p_k означает k-е простое число, а $\pi(N)$ — количество простых чисел, не превосходящих N. (Кстати, эта сумма дает приближенное среднее количество различных простых делителей случайного близкого к N целого числа, поскольку около $1/p$ этих целых чисел делятся на p. При большом N это значение приближенно равно $\ln\ln N + M$, где значение

$$M \approx 0.2614972128476427837554268386086958590515666$$

представляет собой константу Мертенса (Mertens) [271], $\ln x$ — натуральный логарифм x, а $\ln\ln x$ — $\ln(\ln x)$.)

Знак суммы похож на сгорбившегося грузчика.

Самым большим преимуществом обобщенной записи суммы является то, что все манипуляции выполняются проще, чем в записи с явно указанными пределами. Предположим, например, что мы хотим изменить индексную переменную k на $k+1$. В случае обобщенной записи мы получим

$$\sum_{1\leqslant k\leqslant n} a_k = \sum_{1\leqslant k+1\leqslant n} a_{k+1}.$$

Здесь очень легко сообразить, что происходит, и подстановка выполняется практически без лишних раздумий. Но в случае явных пределов запись имеет вид

$$\sum_{k=1}^{n} a_k = \sum_{k=0}^{n-1} a_{k+1},$$

и понять, что происходит, несколько сложнее, а совершить ошибку, наоборот, проще.

С другой стороны, запись с явно указанными пределами не является совершенно бесполезной. Она имеет привлекательный округлый вид да и немного короче — так, в (2.2) используется семь символов, а в (2.3) — восемь. Поэтому мы будем часто использовать $\sum$ с верхним и нижним пределами при формулировке задачи или представлении результата, хотя и будем предпочитать работать с указанными под знаком $\sum$ соотношениями при манипуляциях суммой, индексные переменные которой требуется изменять.

Можно сказать — круглая сумма.

Знак $\sum$ в этой книге встречается более тысячи раз, так что надо быть твердо уверенными в точном понимании его смысла. Формально мы записываем

Мелочь по сравнению с количеством Σ в "Илиаде".

$$\sum_{P(k)} a_k \qquad (2.4)$$

как сокращенную запись суммы всех членов a_k, таких, что целое значение k удовлетворяет заданному условию $P(k)$. ("Условие (предикат) $P(k)$" представляет собой некоторое утверждение относительно k, которое может быть ложным либо истинным.) Пока что допустим, что $a_k \neq 0$ только для конечного количества целых значений k, удовлетворяющих условию $P(k)$; в противном случае выполняется суммирование бесконечного числа ненулевых членов, и ситуация несколько усложняется. Другая крайность — когда $P(k)$ ложно для всех целых значений k, мы получаем "пустую" сумму, которая по определению равна нулю.

Если знак суммы находится в тексте абзаца, а не в отдельной формуле, то используется несколько модифицированный вид (2.4): мы записываем '$\sum_{P(k)} a_k$', где $P(k)$ записывается в виде нижнего индекса при $\sum$, чтобы формула не слишком выдавалась за пределы строки. Аналогично '$\sum_{k=1}^{n} a_k$' представляет собой строчную альтернативу записи (2.2), когда соответствующее обозначение следует разместить в строке текста.

Зачастую хочется написать

$$\sum_{k=2}^{n-1} k(k-1)(n-k) \qquad \text{вместо} \qquad \sum_{k=0}^{n} k(k-1)(n-k),$$

поскольку члены $k = 0$, 1 и n в этой сумме равны нулю. Может показаться более эффективным суммировать $n-2$ члена вместо $n+1$. Однако таких соблазнов следует избегать: эффективность вычислительная — совсем не то же, что эффективность понимания! Позже вы убедитесь, что как можно более простые пределы суммирования обладают тем преимуществом, что существенно упрощают работу с суммами. На самом деле запись $\sum_{k=2}^{n-1}$

может быть просто опасно неоднозначной, поскольку ее смысл при $n = 0$ или $n = 1$ не очевиден (см. упр. 1). Нулевые члены безопасны и позволяют избегать многих неприятностей.

Рассматривавшиеся до сих пор обозначения достаточно стандартны, но сейчас мы отступим от сложившихся традиций. Кеннет Айверсон (Kenneth E. Iverson) в своем языке программирования APL [191, с. 11; см. также 220] предложил замечательную идею, которая, как вы увидите, существенно упрощает многое из того, что мы будем делать в книге. Эта идея состоит в том, чтобы заключать логическое утверждение в квадратные скобки и считать, что результат равен 1, если утверждение истинно, и 0, если оно ложно. Например,

Заметим, что "символ Кронекера", с которым часто приходится встречаться в математической литературе (δ_{kn} равно 1 при $k = n$ и 0 — в противном случае), представляет собой всего лишь частный случай записи Айверсона: вместо символа Кронекера можно записать просто $[k = n]$.

$$[p \text{ простое}] = \begin{cases} 1, & \text{если } p \text{ — простое число;} \\ 0, & \text{если } p \text{ — составное число.} \end{cases}$$

Запись Айверсона позволяет выражать суммы без каких бы то ни было ограничений на индекс суммирования, поскольку (2.4) можно переписать в виде

$$\sum_k a_k \big[P(k)\big] . \tag{2.5}$$

Если $P(k)$ ложно, член $a_k\big[P(k)\big]$ равен нулю, так что мы можем спокойно включать его в состав суммируемых членов. Это упрощает работу с индексом суммирования, поскольку позволяет не беспокоиться о граничных условиях.

"Я часто удивляюсь новым, важным приложениям [этого обозначения]".
— Б. де Финетти (B. de Finetti) [123]

Следует упомянуть одну небольшую техническую деталь: иногда a_k не определено для всех целых k. Мы обходим эту трудность при помощи предположения о том, что $\big[P(k)\big]$ имеет "совершенно нулевое значение", когда $P(k)$ ложно. Оно до такой степени нулевое, что делает $a_k\big[P(k)\big]$ равным нулю, даже если a_k не определено. Например, если использовать запись Айверсона для суммы простых чисел, обратных простым, не превышающим N, в виде

$$\sum_p [p \text{ простое}][p \leqslant N]/p ,$$

то не возникнет никаких проблем с делением на нуль при $p = 0$, поскольку наше соглашение гласит, что $[0 \text{ простое}][0 \leqslant N]/0 = 0$.

Давайте теперь просуммируем все, что мы узнали о суммах. Имеется два основных способа записи суммы членов. Первый использует троеточие '$\cdots$', второй — знак суммы '$\sum$'. Запись с троеточием часто подсказывает полезные преобразования (например, группировку смежных членов), поскольку, когда перед глазами вся сумма, легче увидеть определенные упрощающие ее закономерности. Беда только в том, что за лесом можно не увидеть деревьев. Запись с использованием $\sum$ компактна, впечатляюще действует на девушек и часто подсказывает преобразования, которые в записи с троеточием не очевидны. При работе с $\sum$ нулевые члены не мешают, а зачастую и упрощают выполнение манипуляций.

... и уменьшает шансы провалиться на экзамене из-за "недостаточной строгости"

2.2 Суммы и рекуррентности

Отлично, мы разобрались, как записать сумму при помощи причудливых обозначений. Но как найти значение суммы? Один из способов состоит в том, чтобы заметить тесную связь между суммами и рекуррентными соотношениями. Сумма

$$S_n = \sum_{k=0}^{n} a_k$$

эквивалентна рекуррентности

$$\begin{aligned} S_0 &= a_0 ; \\ S_n &= S_{n-1} + a_n \qquad \text{при } n > 0. \end{aligned} \tag{2.6}$$

(Рассматривайте S_n не как отдельное число, а как последовательность, определенную для всех $n \geqslant 0$.)

Таким образом, можно вычислять суммы в аналитическом виде, используя для этого методы решения рекуррентных соотношений в замкнутой форме из главы 1.

Например, если a_n представляет собой сумму константы и кратного n, то рекуррентное соотношение (2.6) приобретает следующий общий вид:

$$\begin{aligned} R_0 &= \alpha ; \\ R_n &= R_{n-1} + \beta + \gamma n \qquad \text{при } n > 0. \end{aligned} \tag{2.7}$$

Прибегая к методам главы 1, мы находим, что $R_1 = \alpha + \beta + \gamma$, $R_2 = \alpha + 2\beta + 3\gamma$ и т.д.; в общем случае решение может быть записано в виде

$$R_n = A(n)\alpha + B(n)\beta + C(n)\gamma , \tag{2.8}$$

где $A(n)$, $B(n)$ и $C(n)$ — коэффициенты при обобщенных параметрах α, β и γ.

Метод наборов советует нам попробовать подставить вместо R_n простые функции от n в надежде найти константные параметры α, β и γ, при которых решение оказывается особенно простым. Подстановка $R_n = 1$ приводит к $\alpha = 1$, $\beta = 0$, $\gamma = 0$; следовательно,

$$A(n) = 1.$$

Подстановка $R_n = n$ приводит к $\alpha = 0$, $\beta = 1$, $\gamma = 0$, что дает нам

$$B(n) = n.$$

Наконец подстановка $R_n = n^2$ дает $\alpha = 0$, $\beta = -1$, $\gamma = 2$, откуда

$$2C(n) - B(n) = n^2$$

Еще проще: $2 \times 2 = \dfrac{\pi}{\sum_{n\geqslant 0} \frac{2}{(4n+1)(4n+3)}}$.

и $C(n) = (n^2 + n)/2$. Просто как дважды два.

Таким образом, если нам надо вычислить сумму

$$\sum_{k=0}^{n} (a + bk),$$

то рекуррентное соотношение (2.6) сведется к (2.7) с $\alpha = \beta = a$, $\gamma = b$ и его решением будет $aA(n) + aB(n) + bC(n) = a(n+1) + b(n+1)n/2$.

И обратно, многие рекуррентные соотношения могут быть сведены к суммам; таким образом, специальные методы вычисления сумм, которые мы изучим позже в этой главе, могут помочь при решении рекуррентностей, справиться с которыми иначе было бы трудно. Примером может служить рекуррентность, связанная с задачей о Ханойской башне:

$$\begin{aligned} T_0 &= 0; \\ T_n &= 2T_{n-1} + 1 \qquad \text{при } n > 0. \end{aligned}$$

Ее можно привести к частному случаю (2.6) путем деления обеих частей на 2^n:

$$\begin{aligned} T_0/2^0 &= 0; \\ T_n/2^n &= T_{n-1}/2^{n-1} + 1/2^n \qquad \text{при } n > 0. \end{aligned}$$

Теперь можно положить $S_n = T_n/2^n$ и получить

$$\begin{aligned} S_0 &= 0; \\ S_n &= S_{n-1} + 2^{-n} \qquad \text{при } n > 0. \end{aligned}$$

Отсюда следует, что

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

(Обратите внимание, что член для $k = 0$ в сумму не включен.) Сумма геометрической прогрессии $2^{-1} + 2^{-2} + \cdots + 2^{-n} = (\frac{1}{2})^1 + (\frac{1}{2})^2 + \cdots + (\frac{1}{2})^n$ будет выведена несколько позже в этой главе — она равна $1 - (\frac{1}{2})^n$. Следовательно, $T_n = 2^n S_n = 2^n - 1$.

Здесь мы преобразовали T_n в S_n, заметив, что исходное рекуррентное соотношение можно разделить на 2^n. Этот трюк — частный случай общего метода, с помощью которого можно почти любую рекуррентность вида

$$a_n T_n = b_n T_{n-1} + c_n \tag{2.9}$$

привести к сумме. Идея состоит в том, что обе стороны умножаются на *суммирующий множитель* s_n:

$$s_n a_n T_n = s_n b_n T_{n-1} + s_n c_n .$$

Этот множитель s_n выбирается так, чтобы сделать

$$s_n b_n = s_{n-1} a_{n-1} .$$

Далее, если записать $S_n = s_n a_n T_n$, то можно получить рекуррентное соотношение

$$S_n = S_{n-1} + s_n c_n .$$

Следовательно,

$$S_n = s_0 a_0 T_0 + \sum_{k=1}^{n} s_k c_k = s_1 b_1 T_0 + \sum_{k=1}^{n} s_k c_k ,$$

и решение исходного рекуррентного соотношения (2.9) представляет собой

$$T_n = \frac{1}{s_n a_n}\left(s_1 b_1 T_0 + \sum_{k=1}^{n} s_k c_k\right) . \tag{2.10}$$

Например, при $n = 1$ мы получаем $T_1 = (s_1 b_1 T_0 + s_1 c_1)/s_1 a_1 = (b_1 T_0 + c_1)/a_1$.

(Так как s_1 сокращается, оно может быть любым, лишь бы не нулем.)

Но достаточно ли мы хитроумны, чтобы найти верное s_n? Запросто: можно развернуть отношение $s_n = s_{n-1} a_{n-1}/b_n$ и выяснить, что подходящим суммирующим множителем является дробь

$$s_n = \frac{a_{n-1} a_{n-2} \ldots a_1}{b_n b_{n-1} \ldots b_2} , \tag{2.11}$$

которая для удобства может быть умножена на любую константу. Например, у рекуррентности в задаче о Ханойской башне

$a_n = 1$ и $b_n = 2$; только что разработанный общий метод гласит, что $s_n = 2^{-n}$ — подходящий множитель, если мы хотим свести рекуррентность к сумме. Чтобы его обнаружить, не требуется особого прозрения.

Как обычно, нужна определенная осторожность, чтобы не разделить что-нибудь на нуль. Метод суммирующего множителя работает в том случае, когда все a и все b — ненулевые.

Быстрая сортировка изобретена в 1962 году Хоаром (Hoare) [189].

Применим эти идеи к рекуррентному соотношению, возникающему при изучении "быстрой сортировки" — одного из наиболее важных методов сортировки данных в компьютере. Среднее количество сравнений в типичной реализации быстрой сортировки при ее применении к n элементам в случайном порядке удовлетворяет рекуррентному соотношению

$$\begin{aligned} C_0 &= C_1 = 0\,; \\ C_n &= n + 1 + \frac{2}{n}\sum_{k=0}^{n-1} C_k \quad \text{при } n > 1. \end{aligned} \tag{2.12}$$

Да, это соотношение выглядит куда страшнее всех встречавшихся нам до сих пор рекуррентностей; мало того что оно включает сумму всех предыдущих значений, так эта сумма еще и делится на n. Малые случаи дают некоторые данные ($C_2 = 3$, $C_3 = 6$, $C_4 = \frac{19}{2}$), которые не могут вселить в нас чувство уверенности.

Однако можно попробовать понемногу снижать сложность (2.12), сначала избавившись от деления, а затем и от знака $\sum$. Идея заключается в умножении обеих частей выражения на n и получении соотношения

$$nC_n \;=\; n^2 + n + 2\sum_{k=0}^{n-1} C_k \quad \text{при } n > 1;$$

если заменить n на $n-1$, можно получить

$$(n-1)C_{n-1} \;=\; (n-1)^2 + (n-1) + 2\sum_{k=0}^{n-2} C_k \quad \text{при } n-1 > 1.$$

Теперь мы можем вычесть это уравнение из первого, и знак $\sum$ исчезнет:

$$nC_n - (n-1)C_{n-1} \;=\; 2n + 2C_{n-1} \quad \text{при } n > 2.$$

Таким образом, исходное рекуррентное соотношение для C_n сводится к гораздо более простому:

$$\begin{aligned} C_0 &= C_1 = 0; \quad C_2 = 3; \\ nC_n &= (n+1)C_{n-1} + 2n \qquad \text{при } n > 2. \end{aligned}$$

Явный прогресс! Теперь можно применить суммирующий множитель, поскольку данное рекуррентное соотношение имеет вид (2.9) с $a_n = n$, $b_n = n+1$ и

$$c_n = 2n - 2[n=1] + 2[n=2].$$

Изложенный выше общий метод говорит о том, что надо умножить все рекуррентное соотношение на некоторое значение, кратное

$$s_n = \frac{a_{n-1}a_{n-2}\ldots a_1}{b_n b_{n-1}\ldots b_2} = \frac{(n-1)\cdot(n-2)\cdot\ldots\cdot 1}{(n+1)\cdot n\cdot\ldots\cdot 3} = \frac{2}{(n+1)n}.$$

В соответствии с (2.10) решение имеет вид

Мы начали с $\sum$ в рекуррентности и долго избавлялись от нее. Затем при помощи суммирующего множителя мы получили новую $\sum$. Так все же суммы — это хорошо или плохо?

$$C_n = 2(n+1)\sum_{k=1}^{n}\frac{1}{k+1} - \frac{2}{3}(n+1) \qquad \text{при } n > 1.$$

Сумма в этом выражении очень похожа на величину, которая часто появляется в разных приложениях. Она возникает так часто, что мы дадим ей специальное название и введем для нее специальное обозначение:

$$H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \sum_{k=1}^{n}\frac{1}{k}. \tag{2.13}$$

Буква H происходит от слова "harmonic", так что H_n — это *гармоническое число*, названное так в связи с тем, что k-я гармоника, извлекаемая из скрипичной струны, представляет собой основной тон, производимый струной, в k раз короче исходной (с длиной, равной $1/k$ от длины исходной струны).

Наше исследование рекуррентности быстрой сортировки (2.12) можно завершить получением аналитической записи для C_n. Это возможно, если считать таковой выражение C_n через H_n. Сумма в формуле для C_n такова:

$$\sum_{k=1}^{n}\frac{1}{k+1} = \sum_{1\leqslant k\leqslant n}\frac{1}{k+1}.$$

Она легко связывается с H_n путем изменения k на $k-1$ и пересмотра граничных условий:

$$\begin{aligned}\sum_{1\leqslant k\leqslant n}\frac{1}{k+1} &= \sum_{1\leqslant k-1\leqslant n}\frac{1}{k} = \sum_{2\leqslant k\leqslant n+1}\frac{1}{k} = \\ &= \Bigl(\sum_{1\leqslant k\leqslant n}\frac{1}{k}\Bigr) - \frac{1}{1} + \frac{1}{n+1} = H_n - \frac{n}{n+1}.\end{aligned}$$

И даже прилично.

Отлично! Мы нашли значение суммы, что требовалось для завершения решения (2.12): среднее количество сравнений при быстрой сортировке n случайно расположенных элементов равно

$$C_n = 2(n+1)H_n - \tfrac{8}{3}n - \tfrac{2}{3} \qquad \text{при } n > 1. \tag{2.14}$$

Как обычно, проверим корректность малых случаев: $C_2 = 3$, $C_3 = 6$.

Не спутайте с незаконными валютными операциями.

2.3 Преобразование сумм

Ключ к успеху при работе с суммами — в умении заменить одну сумму другой, более простой или более близкой к конечной цели. Научиться этому легко, если усвоить несколько фундаментальных правил преобразования и поработать с ними на практике.

Пусть K — некоторое конечное множество целых чисел. Суммы по элементам K могут быть преобразованы с использованием трех простых правил:

В других моих математических книгах эти законы определены по-другому.

$$\sum_{k\in K} ca_k = c\sum_{k\in K} a_k\,; \qquad \text{(дистрибутивность)} \tag{2.15}$$

$$\sum_{k\in K}(a_k+b_k) = \sum_{k\in K} a_k + \sum_{k\in K} b_k\,; \qquad \text{(ассоциативность)} \tag{2.16}$$

$$\sum_{k\in K} a_k = \sum_{p(k)\in K} a_{p(k)}\,. \qquad \text{(коммутативность)} \tag{2.17}$$

Дистрибутивность позволяет вносить константы под знак суммы и выносить их из него. Ассоциативность позволяет разбивать сумму на две или объединять две суммы в одну. Коммутативность позволяет переупорядочивать члены в любом удобном нам порядке; здесь $p(k)$ — произвольная перестановка множества всех целых чисел. Так, если $K = \{-1, 0, +1\}$ и если $p(k) = -k$, то три указанные правила гласят, что

И почему бы не назвать закон пермутативным, а не коммутативным?

$$ca_{-1} + ca_0 + ca_1 = c(a_{-1} + a_0 + a_1)\,; \qquad \text{(дистрибутивность)}$$

$$(a_{-1} + b_{-1}) + (a_0 + b_0) + (a_1 + b_1)$$
$$= (a_{-1} + a_0 + a_1) + (b_{-1} + b_0 + b_1)\,; \qquad \text{(ассоциативность)}$$

$$a_{-1} + a_0 + a_1 = a_1 + a_0 + a_{-1}\,. \qquad \text{(коммутативность)}$$

Трюк Гаусса из главы 1 можно рассматривать как применение трех указанных фундаментальных законов. Предположим, что мы хотим вычислить сумму обобщенной *арифметической* *прогрессии*,

$$S = \sum_{0\leqslant k\leqslant n}(a+bk).$$

Коммутативность позволяет нам заменить k на $n-k$ и получить

Это что-то вроде изменения переменных в интеграле, но попроще.

$$S = \sum_{0\leqslant n-k\leqslant n}\bigl(a+b(n-k)\bigr) = \sum_{0\leqslant k\leqslant n}(a+bn-bk).$$

Ассоциативность позволяет сложить эти два уравнения:

$$2S = \sum_{0\leqslant k\leqslant n}\bigl((a+bk)+(a+bn-bk)\bigr) = \sum_{0\leqslant k\leqslant n}(2a+bn).$$

Применение дистрибутивности дает нам тривиальную, легко вычисляемую сумму:

"Чему равно один плюс один плюс один плюс один плюс один плюс один плюс один плюс один плюс один плюс один?" "Не знаю, — сказала Алиса. — Я сбилась со счета." "Она не умеет складывать!"
— Льюис Кэрролл (Lewis Carroll) [50]

$$2S = (2a+bn)\sum_{0\leqslant k\leqslant n}1 = (2a+bn)(n+1).$$

Разделив на 2, мы доказали, что

$$\sum_{k=0}^{n}(a+bk) = (a+\tfrac{1}{2}bn)(n+1). \tag{2.18}$$

Правую часть можно запомнить как среднее первого и последнего членов, а именно $\frac{1}{2}\bigl(a+(a+bn)\bigr)$, умноженное на количество членов $(n+1)$.

Важно не забывать, что функция $p(k)$ в обобщенном коммутативном законе (2.17) представляет собой перестановку всех целых чисел. Иными словами, для каждого целого n должно быть ровно одно целое k, такое, что $p(k)=n$. В противном случае коммутативный закон может не выполняться, как наглядно иллюстрируется в упражнении 3. Преобразования наподобие $p(k)=k+c$ или $p(k)=c-k$, где c — целочисленная константа, всегда являются перестановками, так что они всегда работоспособны.

С другой стороны, мы можем несколько ослабить ограничения на перестановку: достаточно потребовать, чтобы существовало ровно одно целое число k, такое, что $p(k)=n$, когда n — элемент индексного множества K. Если $n\notin K$ (т.е. если n не принадлежит K), то не имеет значения, сколько раз $p(k)=n$, так как такие k не принимают участия в суммировании. Так, например, можно утверждать, что

$$\sum_{\substack{k\in K\\ k \text{ четное}}} a_k = \sum_{\substack{n\in K\\ n \text{ четное}}} a_n = \sum_{\substack{2k\in K\\ 2k \text{ четное}}} a_{2k} = \sum_{2k\in K} a_{2k}, \qquad (2.19)$$

поскольку имеется ровно одно k, такое, что $2k = n$, когда $n \in K$ и n четное.

Не вместо?

Запись Айверсона, которая позволяет нам получать значения 0 или 1 для логических утверждений в формуле, может использоваться вместе с дистрибутивным, ассоциативным и коммутативным законами для вывода дополнительных свойств сумм. Например, вот важное правило для объединения различных множеств индексов: если K и K' являются некоторыми множествами целых чисел, то

$$\sum_{k\in K} a_k + \sum_{k\in K'} a_k = \sum_{k\in K\cap K'} a_k + \sum_{k\in K\cup K'} a_k. \qquad (2.20)$$

Это следует из общих формул

$$\sum_{k\in K} a_k = \sum_k a_k [k\in K] \qquad (2.21)$$

и

$$[k\in K] + [k\in K'] = [k\in K\cap K'] + [k\in K\cup K']. \qquad (2.22)$$

Обычно правило (2.20) используется либо для объединения двух почти не пересекающихся множеств индексов, как в случае

$$\sum_{k=1}^{m} a_k + \sum_{k=m}^{n} a_k = a_m + \sum_{k=1}^{n} a_k \qquad \text{при } 1 \leqslant m \leqslant n,$$

либо для отделения одного члена суммы, как в случае

(Здесь поменялись местами части (2.20).)

$$\sum_{0\leqslant k\leqslant n} a_k = a_0 + \sum_{1\leqslant k\leqslant n} a_k \qquad \text{при } n \geqslant 0. \qquad (2.23)$$

Эта операция отделения члена представляет собой основу *метода перестановок*, который часто позволяет нам вычислять суммы в аналитическом виде. Идея заключается в том, чтобы начать с неизвестной суммы, подлежащей вычислению, и обозначить ее как S_n:

Именуй и властвуй.

$$S_n = \sum_{0\leqslant k\leqslant n} a_k.$$

Затем мы переписываем S_{n+1} двумя способами, выделяя последний и первый члены:

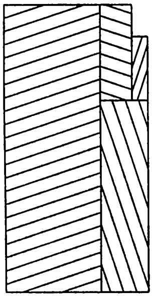
$$\begin{aligned} S_n + a_{n+1} &= \sum_{0\leqslant k\leqslant n+1} a_k = a_0 + \sum_{1\leqslant k\leqslant n+1} a_k = \\ &= a_0 + \sum_{1\leqslant k+1\leqslant n+1} a_{k+1} = \\ &= a_0 + \sum_{0\leqslant k\leqslant n} a_{k+1}\,. \end{aligned} \qquad (2.24)$$

Теперь можно поработать с этой последней суммой и попытаться выразить ее через S_n. Если мы сумеем это сделать, то получим уравнение, решением которого будет искомая сумма.

Воспользуемся этим подходом для вычисления суммы обобщенной *геометрической прогрессии*

Для геометрической прогрессии и доказательство должно быть геометрическим.

$$S_n = \sum_{0\leqslant k\leqslant n} ax^k\,.$$



Общая схема из (2.24) гласит, что

$$S_n + ax^{n+1} = ax^0 + \sum_{0\leqslant k\leqslant n} ax^{k+1}\,,$$

а сумма в правой части равна $x\sum_{0\leqslant k\leqslant n} ax^k = xS_n$ в соответствии с дистрибутивным законом. Таким образом, $S_n+ax^{n+1} = a+xS_n$, и мы можем решить это уравнение относительно S_n и получить

$$\sum_{k=0}^{n} ax^k = \frac{a - ax^{n+1}}{1-x} \qquad \text{при } x \neq 1. \qquad (2.25)$$

(Очевидно, что при $x = 1$ сумма просто равна $(n+1)a$.) Правую часть этой формулы можно запомнить как разность первого члена суммы и первого члена, не входящего в сумму (члена, следующего за последним членом суммы), разделенную на разность 1 и знаменателя прогрессии.

Эту формулу вдолбили в меня еще в школе.

Для настоящих мужчин это слишком просто. Давайте применим данный метод для вычисления несколько более сложной суммы, а именно

$$S_n = \sum_{0\leqslant k\leqslant n} k2^k\,.$$

В этом случае $S_0 = 0$, $S_1 = 2$, $S_2 = 10$, $S_3 = 34$, $S_4 = 98$; какова же общая формула? Согласно (2.24) получим

$$S_n + (n+1)2^{n+1} = \sum_{0 \leqslant k \leqslant n} (k+1)2^{k+1};$$

и теперь нам надо выразить сумму в правой части через S_n. Ассоциативность позволяет разбить ее на две части

$$\sum_{0 \leqslant k \leqslant n} k\,2^{k+1} + \sum_{0 \leqslant k \leqslant n} 2^{k+1},$$

и первая из полученных сумм представляет собой ни что иное, как $2S_n$. Вторая сумма — сумма геометрической прогрессии, которая согласно (2.25) равна $(2-2^{n+2})/(1-2) = 2^{n+2}-2$. Таким образом, $S_n + (n+1)2^{n+1} = 2S_n + 2^{n+2} - 2$, и после применения алгебраических преобразований мы получаем

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

Теперь понятно, почему $S_3 = 34$: это $32 + 2$, а не $2 \cdot 17$.

Аналогичные рассуждения с x вместо 2 привели бы нас к уравнению $S_n + (n+1)x^{n+1} = xS_n + (x - x^{n+2})/(1-x)$; следовательно, мы можем заключить, что

$$\sum_{k=0}^{n} kx^k = \frac{x-(n+1)x^{n+1}+nx^{n+2}}{(1-x)^2} \qquad \text{при } x \neq 1. \tag{2.26}$$

Интересно, что эту же формулу можно легко вывести совершенно иным способом, при помощи элементарных методов дифференциального исчисления. Если начать с уравнения

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$$

и взять производную от обеих частей по x, то можно получить

$$\begin{aligned}\sum_{k=0}^{n} kx^{k-1} &= \frac{(1-x)\big(-(n+1)x^n\big) + 1-x^{n+1}}{(1-x)^2}\\ &= \frac{1-(n+1)x^n + nx^{n+1}}{(1-x)^2},\end{aligned}$$

поскольку производная суммы равна сумме производных слагаемых. В последующих главах мы встретимся с множеством связей между континуальной и дискретной математикой.

2.4 Кратные суммы

Члены суммы могут иметь не один, а два и более индексов. Например, вот как выглядит двойная сумма из девяти членов, управляемая индексами j и k:

Девятеро двоих не ждут?..

$$\sum_{1\leqslant j,k\leqslant 3} a_j b_k = a_1b_1 + a_1b_2 + a_1b_3 + \\ + a_2b_1 + a_2b_2 + a_2b_3 + \\ + a_3b_1 + a_3b_2 + a_3b_3 .$$

Обратите внимание, что $1 \leqslant j,k \leqslant 3$ не означает суммирование по всем $j \geqslant 1$ и всем $k \leqslant 3$.

Для таких сумм используются те же обозначения и методы, что и для сумм с одним индексом. Так, если $P(j,k)$ представляет собой некоторый предикат от j и k, то сумма всех членов $a_{j,k}$, таких, что значение $P(j,k)$ истинно, может быть записана двумя способами. Один из них использует запись Айверсона и суммирование по *всем* парам целых чисел j и k:

$$\sum_{P(j,k)} a_{j,k} = \sum_{j,k} a_{j,k}\,[P(j,k)] .$$

Несмотря на наличие более одного индекса, нам достаточно только одного знака $\sum$, который означает суммирование по всем допустимым комбинациям индексов.

Иногда используются и два знака $\sum$, когда мы говорим о сумме сумм. Например, выражение

$$\sum_j \sum_k a_{j,k}\,[P(j,k)]$$

является ни чем иным, как сокращенной записью суммы

$$\sum_j \Bigl(\sum_k a_{j,k}\,[P(j,k)]\Bigr) ,$$

которая, в свою очередь, представляет собой сумму по всем целым j членов $\sum_k a_{j,k}\,[P(j,k)]$. Эти члены являются суммами по всем целым k всех членов $a_{j,k}$, для которых предикат $P(j,k)$ истинен. В таких случаях говорят, что двойная сумма "сначала суммируется по k". Сумму, которая зависит более чем от одного индекса, можно начинать суммировать с любого из ее индексов.

Кратные Σ вычисляются справа налево (изнутри наружу).

С этим связаны важное правило, именуемое *изменением порядка суммирования*, и обобщающее правило ассоциативности (2.16), с которым мы познакомились ранее:

$$\sum_j \sum_k a_{j,k}[P(j,k)] = \sum_{P(j,k)} a_{j,k} = \sum_k \sum_j a_{j,k}[P(j,k)] . \qquad (2.27)$$

Здесь средний член представляет собой сумму по двум индексам. Слева $\sum_j \sum_k$ означает суммирование сначала по k, а затем по j. Справа $\sum_k \sum_j$ означает суммирование сначала по j, а затем по k. На практике при вычислении двойной суммы в аналитическом виде обычно суммирование по одному индексу выполняется проще, чем по другому, так что мы можем выбирать, какой из них для нас более удобен.

Никто и не боится. Я думаю, это правило по сравнению с некоторыми вещами из главы 1 — сама простота и очевидность.

Суммы сумм — это еще не самое страшное в математике, но начинающему они могут показаться настоящим лабиринтом, так что добавим в качестве путеводной нити еще пару примеров. Сумма из девяти членов, с которой мы встретились в начале раздела, — хорошая иллюстрация преобразования двойных сумм, поскольку она может быть упрощена, и этот процесс упрощения типичен для того, что можно делать с двойными суммами:

$$\begin{aligned}
\sum_{1\leqslant j,k\leqslant 3} a_j b_k &= \sum_{j,k} a_j b_k [1\leqslant j,k\leqslant 3] = \\
&= \sum_{j,k} a_j b_k [1\leqslant j\leqslant 3][1\leqslant k\leqslant 3] = \\
&= \sum_{j}\sum_{k} a_j b_k [1\leqslant j\leqslant 3][1\leqslant k\leqslant 3] = \\
&= \sum_{j} a_j [1\leqslant j\leqslant 3] \sum_{k} b_k [1\leqslant k\leqslant 3] = \\
&= \sum_{j} a_j [1\leqslant j\leqslant 3] \Bigl(\sum_{k} b_k [1\leqslant k\leqslant 3]\Bigr) = \\
&= \Bigl(\sum_{j} a_j [1\leqslant j\leqslant 3]\Bigr)\Bigl(\sum_{k} b_k [1\leqslant k\leqslant 3]\Bigr) = \\
&= \Bigl(\sum_{j=1}^{3} a_j\Bigr)\Bigl(\sum_{k=1}^{3} b_k\Bigr).
\end{aligned}$$

В первой строке записана сумма девяти членов без какого-либо конкретного упорядочения. Во второй строке они сгруппированы в тройки $(a_1b_1+a_1b_2+a_1b_3)+(a_2b_1+a_2b_2+a_2b_3)+(a_3b_1+a_3b_2+a_3b_3)$. В третьей строке для вынесения за скобки всех a-членов использована дистрибутивность суммирования, так как a_j и $[1\leqslant j\leqslant 3]$ не зависят от k; это дает нам $a_1(b_1+b_2+b_3)+a_2(b_1+b_2+b_3)+a_3(b_1+b_2+b_3)$. Четвертая строка аналогична третьей, но в ней имеется лишняя пара скобок, которые требуются для того, чтобы пятая строка не казалась такой загадочной. В ней выносится множитель $(b_1+b_2+b_3)$, который встречается для каждого значения j: $(a_1+a_2+a_3)(b_1+b_2+b_3)$. А последняя строка представляет собой всего лишь иной способ записи предпоследней строки. Этот метод вывода может использоваться для доказательства *обобщенного дистрибутивного закона*

$$\sum_{\substack{j\in J\\ k\in K}} a_j b_k = \Bigl(\sum_{j\in J} a_j\Bigr)\Bigl(\sum_{k\in K} b_k\Bigr), \qquad (2.28)$$

справедливого для любых множеств индексов J и K.

У фундаментального закона изменения порядка суммирования (2.27) имеется ряд вариаций, которые возникают, когда мы хотим ограничить диапазоны значений индексов, а не выполнять суммирование по всем целым j и k. Эти вариации вызывают две ассоциации — ванили и булыжной мостовой. Начнем с ванили:

$$\sum_{j\in J}\sum_{k\in K} a_{j,k} = \sum_{\substack{j\in J\\ k\in K}} a_{j,k} = \sum_{k\in K}\sum_{j\in J} a_{j,k}\,. \qquad (2.29)$$

Это всего лишь иной способ записи (2.27), связанный с тем, что “айверсониан” $[j\in J, k\in K]$ можно разложить на множители: $[j\in J][k\in K]$. “Ванильное правило” применимо в тех случаях, когда диапазоны j и k не зависят друг от друга.

Булыжник не столь гладкий. Эта формула применима тогда, когда диапазон значений индекса внутренней суммы зависит от индексной переменной внешней суммы:

$$\sum_{j\in J}\ \sum_{k\in K(j)} a_{j,k} = \sum_{k\in K'}\ \sum_{j\in J'(k)} a_{j,k}\,. \qquad (2.30)$$

Здесь множества J, $K(j)$, K' и $J'(k)$ должны быть связаны так, чтобы

$$[j\in J]\bigl[k\in K(j)\bigr] = [k\in K']\bigl[j\in J'(k)\bigr]\,.$$

Подобное разложение, в принципе, всегда возможно: можно положить, что $J = K'$ — множество всех целых чисел, а $K(j) = J'(k)$ — множества, соответствующие предикату $P(j,k)$, управляющему двойной суммой. Однако имеются важные частные случаи, когда множества J, $K(j)$, K' и $J'(k)$ приобретают простой вид. Такие случаи — не редкость в приложениях. Например, вот весьма полезное разложение:

$$\begin{aligned}[1\leqslant j\leqslant n][j\leqslant k\leqslant n] &= [1\leqslant j\leqslant k\leqslant n] =\\ &= [1\leqslant k\leqslant n][1\leqslant j\leqslant k]\,. \end{aligned}\qquad (2.31)$$

Это равенство с использованием "айверсонианов" позволяет нам записать следующее:

$$\sum_{j=1}^{n}\sum_{k=j}^{n} a_{j,k} = \sum_{1\leqslant j\leqslant k\leqslant n} a_{j,k} = \sum_{k=1}^{n}\sum_{j=1}^{k} a_{j,k}\,. \tag{2.32}$$

Самое время немного разогреться упражнениями 5 и 6.

(Или забраться в поисках горячительного в бар.)

Одна из приведенных сумм обычно вычисляется проще другой, так что можно воспользоваться (2.32) для того, чтобы переключиться со сложной суммы на простую.

Применим описанные идеи на практике. Рассмотрим массив

$$\begin{bmatrix} a_1a_1 & a_1a_2 & a_1a_3 & \ldots & a_1a_n \\ a_2a_1 & a_2a_2 & a_2a_3 & \ldots & a_2a_n \\ a_3a_1 & a_3a_2 & a_3a_3 & \ldots & a_3a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_na_1 & a_na_2 & a_na_3 & \ldots & a_na_n \end{bmatrix}$$

из n^2 произведений a_ja_k. Наша цель заключается в поиске простой формулы для

$$S_{\triangledown} = \sum_{1\leqslant j\leqslant k\leqslant n} a_ja_k\,,$$

суммы всех элементов на и выше главной диагонали данного массива. Поскольку $a_ja_k = a_ka_j$, массив симметричен относительно главной диагонали; поэтому $S_{\triangledown}$ будет равен примерно половине суммы *всех* элементов (за исключением подгоночного множителя для учета главной диагонали).

Неужто булыжники на дороге бывают подогнанными?

Эти соображения наводят на мысль о следующих преобразованиях:

$$S_{\triangledown} = \sum_{1\leqslant j\leqslant k\leqslant n} a_ja_k = \sum_{1\leqslant k\leqslant j\leqslant n} a_ka_j = \sum_{1\leqslant k\leqslant j\leqslant n} a_ja_k = S_{\triangle}\,,$$

поскольку (j,k) можно переобозначить как (k,j). Кроме того, поскольку

$$[1\leqslant j\leqslant k\leqslant n]+[1\leqslant k\leqslant j\leqslant n] = [1\leqslant j,k\leqslant n]+[1\leqslant j=k\leqslant n]\,,$$

мы получаем

$$2S_{\triangledown} = S_{\triangledown}+S_{\triangle} = \sum_{1\leqslant j,k\leqslant n} a_ja_k + \sum_{1\leqslant j=k\leqslant n} a_ja_k\,.$$

Первая сумма в соответствии с обобщенным дистрибутивным законом (2.28) равна $\left(\sum_{j=1}^{n} a_j\right)\left(\sum_{k=1}^{n} a_k\right) = \left(\sum_{k=1}^{n} a_k\right)^2$, а вторая — $\sum_{k=1}^{n} a_k^2$. Поэтому

$$S_{\triangledown} = \sum_{1\leqslant j\leqslant k\leqslant n} a_j a_k = \frac{1}{2}\left(\left(\sum_{k=1}^{n} a_k\right)^2 + \sum_{k=1}^{n} a_k^2\right). \qquad (2.33)$$

Это сумма элементов верхнетреугольной матрицы, выраженная через более простые однократные суммы.

Воодушевившись таким успехом, рассмотрим другую двойную сумму:

$$S = \sum_{1\leqslant j<k\leqslant n} (a_k - a_j)(b_k - b_j).$$

И вновь мы имеем симметрию при перестановке j и k:

$$S = \sum_{1\leqslant k<j\leqslant n} (a_j - a_k)(b_j - b_k) = \sum_{1\leqslant k<j\leqslant n} (a_k - a_j)(b_k - b_j).$$

Так что, воспользовавшись равенством

$$[1\leqslant j<k\leqslant n] + [1\leqslant k<j\leqslant n] = [1\leqslant j,k\leqslant n] - [1\leqslant j=k\leqslant n],$$

сумму S можно прибавить к самой себе и получить

$$2S = \sum_{1\leqslant j,k\leqslant n} (a_j - a_k)(b_j - b_k) - \sum_{1\leqslant j=k\leqslant n} (a_j - a_k)(b_j - b_k).$$

Вторая сумма равна нулю. А первая? Она распадается на четыре отдельные “ванильные” суммы:

$$\begin{aligned}
&\sum_{1\leqslant j,k\leqslant n} a_j b_j - \sum_{1\leqslant j,k\leqslant n} a_j b_k - \sum_{1\leqslant j,k\leqslant n} a_k b_j + \sum_{1\leqslant j,k\leqslant n} a_k b_k = \\
&\quad = 2\sum_{1\leqslant j,k\leqslant n} a_k b_k - 2\sum_{1\leqslant j,k\leqslant n} a_j b_k = \\
&\quad = 2n\sum_{1\leqslant k\leqslant n} a_k b_k - 2\left(\sum_{k=1}^{n} a_k\right)\left(\sum_{k=1}^{n} b_k\right).
\end{aligned}$$

На последнем шаге обе суммы были упрощены в соответствии с обобщенным дистрибутивным законом (2.28). Если преобразование первой суммы кажется загадкой, повторим его в замедленной съемке:

$$\begin{aligned}
2\sum_{1\leqslant j,k\leqslant n} a_k b_k &= 2\sum_{1\leqslant k\leqslant n}\ \sum_{1\leqslant j\leqslant n} a_k b_k = \\
&= 2\sum_{1\leqslant k\leqslant n} a_k b_k \sum_{1\leqslant j\leqslant n} 1 = \\
&= 2\sum_{1\leqslant k\leqslant n} a_k b_k n = 2n\sum_{1\leqslant k\leqslant n} a_k b_k.
\end{aligned}$$

Индексная переменная, отсутствующая в общем члене (в данном случае это j), может быть легко удалена путем умножения того, что осталось, на размер множества значений этой переменной (в нашем случае — на n).

Вернемся к тому месту, где мы остановились. Теперь мы можем разделить все на 2 и, выполнив перестановку, получить интересную формулу:

$$\Bigl(\sum_{k=1}^{n} a_k\Bigr)\Bigl(\sum_{k=1}^{n} b_k\Bigr) = \\ = n\sum_{k=1}^{n} a_k b_k - \sum_{1\leqslant j<k\leqslant n} (a_k - a_j)(b_k - b_j)\,. \quad (2.34)$$

Это равенство в качестве частных случаев дает *неравенства Чебышева*:

(*Чебышев [58] в действительности доказал аналогичный результат не для сумм, а для интегралов:* $(\int_a^b f(x)dx)\times$ $\times(\int_a^b g(x)dx) \leqslant$ $\leqslant (b-a)\times$ $\times(\int_a^b f(x)g(x)dx)$, *если* $f(x)$ *и* $g(x)$ *— монотонные неубывающие функции.*)

$$\Bigl(\sum_{k=1}^{n} a_k\Bigr)\Bigl(\sum_{k=1}^{n} b_k\Bigr) \leqslant n\sum_{k=1}^{n} a_k b_k\,, \quad \text{если } a_1 \leqslant \cdots \leqslant a_n \\ \text{и } b_1 \leqslant \cdots \leqslant b_n;$$

$$\Bigl(\sum_{k=1}^{n} a_k\Bigr)\Bigl(\sum_{k=1}^{n} b_k\Bigr) \geqslant n\sum_{k=1}^{n} a_k b_k\,, \quad \text{если } a_1 \leqslant \cdots \leqslant a_n \\ \text{и } b_1 \geqslant \cdots \geqslant b_n.$$

(В общем случае, если $a_1 \leqslant \cdots \leqslant a_n$ и если p — перестановка множества $\{1,\ldots,n\}$, нетрудно доказать, что наибольшее значение $\sum_{k=1}^{n} a_k b_{p(k)}$ достигается при $b_{p(1)} \leqslant \cdots \leqslant b_{p(n)}$, а наименьшее значение — при $b_{p(1)} \geqslant \cdots \geqslant b_{p(n)}$.)

Многократное суммирование имеет интересную связь с обобщенной операцией замены индекса суммирования в *однократных* суммах. Нам известно, что в силу коммутативности

$$\sum_{k\in K} a_k = \sum_{p(k)\in K} a_{p(k)}\,,$$

если $p(k)$ представляет собой произвольную перестановку целых чисел. Но что будет, если заменить k на $f(j)$, где f — произвольная функция

$$f\colon J \to K,$$

которая преобразует целое число $j \in J$ в целое число $f(j) \in K$?

Общая формула замены индекса имеет вид

$$\sum_{j\in J} a_{f(j)} = \sum_{k\in K} a_k \,\#f^{-}(k)\,, \qquad (2.35)$$

где $\#f^{-}(k)$ обозначает количество элементов множества

$$f^{-}(k) = \bigl\{ j \bigm| f(j) = k \bigr\},$$

т.е. количество значений $j \in J$, таких, что $f(j)$ равно k.

Формулу (2.35) легко доказать путем изменения порядка суммирования:

$$\sum_{j\in J} a_{f(j)} = \sum_{\substack{j\in J\\ k\in K}} a_k \,\bigl[f(j)=k\bigr] = \sum_{k\in K} a_k \sum_{j\in J} \bigl[f(j)=k\bigr]\,,$$

поскольку $\sum_{j\in J} \bigl[f(j)=k\bigr] = \#f^{-}(k)$. В частном случае, когда f представляет собой взаимно однозначное соответствие между J и K, имеем $\#f^{-}(k) = 1$ при всех k, и общая формула (2.35) сводится к

$$\sum_{j\in J} a_{f(j)} = \sum_{f(j)\in K} a_{f(j)} = \sum_{k\in K} a_k\,.$$

Один мой преподаватель математики именовал это “биекцией”; возможно, в один прекрасный день я все же сумею полюбить это слово...

Это коммутативный закон (2.17), с которым мы уже встречались ранее, но в несколько замаскированном виде.

Пока что наши примеры кратных сумм включали обобщенные члены наподобие a_k или b_k. Но поскольку в названии книги есть слово “конкретная”, давайте рассмотрим кратную сумму с реальными числами:

$$S_n = \sum_{1\leqslant j<k\leqslant n} \frac{1}{k-j}\,.$$

Подумать только: эти авторы считают, что j, k и n — “реальные числа”.

В частности, $S_1 = 0$; $S_2 = 1$; $S_3 = \frac{1}{2-1} + \frac{1}{3-1} + \frac{1}{3-2} = \frac{5}{2}$.

Обычный способ вычисления двойной суммы состоит в суммировании сначала по j или по k, так что рассмотрим оба варианта.

$$\begin{aligned}
S_n &= \sum_{1\leqslant k\leqslant n}\ \sum_{1\leqslant j<k} \frac{1}{k-j} = && \text{(суммируем по } j\text{)}\\
&= \sum_{1\leqslant k\leqslant n}\ \sum_{1\leqslant k-j<k} \frac{1}{j} = && \text{(заменяем } j \text{ на } k-j\text{)}\\
&= \sum_{1\leqslant k\leqslant n}\ \sum_{0<j\leqslant k-1} \frac{1}{j} = && \text{(упрощаем границы для } j\text{)}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{1\leqslant k\leqslant n} H_{k-1} = && \text{(по определению } H_{k-1} \text{ (2.13))}\\
&= \sum_{1\leqslant k+1\leqslant n} H_k = && \text{(заменяем } k \text{ на } k+1)\\
&= \sum_{0\leqslant k<n} H_k\,. && \text{(упрощаем границы для } k)
\end{aligned}
$$

Увы! Мы не знаем, как вычислить сумму гармонических чисел в аналитическом виде.

Займемся самобичеванием?

Если попробовать подойти с другой стороны, получим

$$
\begin{aligned}
S_n &= \sum_{1\leqslant j\leqslant n}\ \sum_{j<k\leqslant n} \frac{1}{k-j} = && \text{(суммируем по } k)\\
&= \sum_{1\leqslant j\leqslant n}\ \sum_{j<k+j\leqslant n} \frac{1}{k} = && \text{(заменяем } k \text{ на } k+j)\\
&= \sum_{1\leqslant j\leqslant n}\ \sum_{0<k\leqslant n-j} \frac{1}{k} = && \text{(упрощаем границы для } k)\\
&= \sum_{1\leqslant j\leqslant n} H_{n-j} = && \text{(по определению } H_{n-j} \text{ (2.13))}\\
&= \sum_{1\leqslant n-j\leqslant n} H_j = && \text{(заменяем } j \text{ на } n-j)\\
&= \sum_{0\leqslant j<n} H_j\,. && \text{(упрощаем границы для } j)
\end{aligned}
$$

И мы опять в том же тупике...

Но есть и *иной* путь, на котором мы заменяем k на $k+j$ *до* приведения S_n к сумме сумм:

$$
\begin{aligned}
S_n &= \sum_{1\leqslant j<k\leqslant n} \frac{1}{k-j} = && \text{(переписываем исходную сумму)}\\
&= \sum_{1\leqslant j<k+j\leqslant n} \frac{1}{k} = && \text{(заменяем } k \text{ на } k+j)\\
&= \sum_{1\leqslant k\leqslant n}\ \sum_{1\leqslant j\leqslant n-k} \frac{1}{k} = && \text{(суммируем сначала по } j)\\
&= \sum_{1\leqslant k\leqslant n} \frac{n-k}{k} = && \text{(сумма по } j \text{ тривиальна)}\\
&= \sum_{1\leqslant k\leqslant n} \frac{n}{k} - \sum_{1\leqslant k\leqslant n} 1 = && \text{(в силу ассоциативности)}\\
&= n\Bigl(\sum_{1\leqslant k\leqslant n} \frac{1}{k}\Bigr) - n = && \text{(надо же!)}\\
&= nH_n - n\,. && \text{(по определению } H_n \text{ (2.13))}
\end{aligned}
$$

Ура! Мы нашли S_n. А комбинируя наши фальшстарты с найденным решением, мы в качестве приза находим сумму гармонических чисел:

$$\sum_{0\leqslant k<n} H_k = nH_n - n. \tag{2.36}$$

Здесь разумно использовать $k \leqslant n$ вместо $k \leqslant n-1$. Простые границы экономят силы.

Понять использованный здесь трюк можно двумя путями — алгебраически и геометрически. (1) Алгебраически, если имеется двойная сумма, члены которой включают $k+f(j)$, где f — произвольная функция, этот пример показывает, что неплохо попробовать заменить k на $k-f(j)$ и просуммировать по j. (2) Геометрически можно рассматривать конкретную сумму S_n в следующем виде (рассмотрен случай $n=4$):

$$\begin{array}{lcccccccc} & k=1 & k=2 & & k=3 & & k=4 \\ j=1 & & \frac{1}{1} & + & \frac{1}{2} & + & \frac{1}{3} \\ j=2 & & & & \frac{1}{1} & + & \frac{1}{2} \\ j=3 & & & & & & \frac{1}{1} \\ j=4 & & & & & & \end{array}$$

Наши первые попытки суммирования сначала по j (по столбцам) или по k (по строкам) давали нам $H_1+H_2+H_3 = H_3+H_2+H_1$. Выигрышная идея, по сути, заключалась в суммировании по диагоналям, что дало $\frac{3}{1}+\frac{2}{2}+\frac{1}{3}$.

2.5 Общие методы

Давайте закрепим изученное, рассмотрев один и тот же пример под разными углами. На следующих страницах мы попытаемся найти аналитический вид суммы первых n квадратов, которую будем обозначать как $\square_n$:

$$\square_n = \sum_{0\leqslant k\leqslant n} k^2 \qquad \text{при } n \geqslant 0. \tag{2.37}$$

Мы увидим, что имеется как минимум восемь разных способов решения этой задачи, и в ходе анализа различных тактических решений мы изучим стратегии успешного наступления на суммы в общем случае.

Сначала, как обычно, рассмотрим малые случаи:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
n^2	0	1	4	9	16	25	36	49	64	81	100	121	144
$\square_n$	0	1	5	14	30	55	91	140	204	285	385	506	650

Никакого очевидного аналитического вида для $\Box_n$ не наблюдается; но если нам удастся его найти, то эти значения могут послужить для проверки.

Метод 0: поиск в справочнике

Задача наподобие суммы первых n квадратов наверняка уже кем-то решена, так что, скорее всего, ее решение можно найти в каком-нибудь математическом справочнике*. И в самом деле, на странице 36 *Стандартных математических таблиц CRC* (*CRC Standard Mathematical Tables*) [28] имеется искомый ответ:

$$\Box_n = \frac{n(n+1)(2n+1)}{6} \qquad \text{при } n \geqslant 0. \tag{2.38}$$

Просто чтобы убедиться, что мы корректно прочли формулу в справочнике и в ней нет опечатки, убедимся, что она совершенно верно дает $\Box_5 = 5\cdot 6\cdot 11/6 = 55$. Кстати, на той же странице 36 *Таблиц CRC* имеется информация о суммах кубов, четвертых, ..., десятых степеней.

Очень мощным справочником по математическим формулам является *Справочник по специальным функциям* (*Handbook of Mathematical Functions*) под редакцией Абрамовица (Abramowitz) и Стиган (Stegun) [2]. На стр. 813–814 этой книги перечислены значения $\Box_n$ для $n \leqslant 100$; а на стр. 804 и 809 приведена формула, эквивалентная (2.38), вместе с аналогичными формулами для сумм кубов, четвертых, ..., пятнадцатых степеней (как для обычных, так и для знакочередующихся сумм).

Более сложные суммы можно найти в таблицах Хансена (Hansen) [178].

Но наилучшим источником ответов на вопросы о последовательностях является небольшая книга *Справочник по целочисленным последовательностям* (*Handbook of Integer Sequences*) Слоана (Sloane) [330], в которой перечислены тысячи последовательностей в соответствии с их числовыми значениями. Если вы сталкиваетесь с некоторой рекуррентностью, которую есть основание считать изученной, все, что нужно сделать, — это вычислить количество членов последовательности, достаточное для распознавания ее среди прочих; после этого у вас появятся шансы обнаружить эту последовательность и ссылки на соответствующую литературу в *Справочнике* Слоана. Например, последовательность 1, 5, 14, 30, ... имеет у Слоана номер 1574 и названа последовательностью "квадратно пирамидальных чисел"

---

* См., например, формулу 8 на стр. 597 в А.П. Прудников, Ю.А. Брычков, О.И. Маричев. *Интегралы и ряды.* — М.: Наука. — 1981. — *Примеч. пер.*

("square pyramidal numbers"), поскольку в пирамиде с квадратным основанием из n шаров имеется $\Box_n$ шаров. Слоан приводит три ссылки, одной из которых является ранее упомянутый справочник Абрамовица и Стиган.

Еще один способ приобщиться к мировой математической мудрости заключается в том, чтобы использовать компьютерную программу (типа Axiom, MACSYMA, Maple или Mathematica), предоставляющую инструментарий для символьных преобразований. Такие программы в особенности необходимы людям, работающим с большими формулами.

Знакомство со стандартными источниками информации крайне желательно, поскольку они могут оказаться исключительно полезными. Но все же Метод 0 не согласуется с духом данной книги, поскольку мы хотим получать ответы самостоятельно. Метод поиска готового решения ограничивается задачами, которые кому-то показались достойными внимания; интересующие нас задачи могут среди них отсутствовать.

Или как минимум задачами с теми же ответами, что и у задач, которые кому-то показались достойными внимания.

Метод 1: угадывание ответа с доказательством по индукции

Принесла ли нам ответ на хвосте сорока или к нему привело получасовое биение в бубен — от нас в любом случае требуется доказать его корректность.

Можно, например, заметить, что значения $\Box_n$ имеют относительно небольшие простые множители, так что можно было бы получить формулу (2.38) для небольших значений n. Можно также предложить эквивалентную формулу

$$\Box_n = \frac{n(n+\frac{1}{2})(n+1)}{3} \qquad \text{при } n \geqslant 0, \tag{2.39}$$

которую несколько легче запомнить. Все говорит в пользу формулы (2.39), но мы обязаны ее доказать так, чтобы в ней не оставалось никаких сомнений. Для этой цели и была изобретена математическая индукция.

"Итак, Ваша Честь, мы знаем, что $\Box_0 = 0 = 0(0+\frac{1}{2})(0+1)/3$, так что с базой все просто. По индукции предположим, что $n > 0$ и что (2.39) выполняется, когда n заменяется на $n-1$. Поскольку

$$\Box_n = \Box_{n-1} + n^2,$$

имеем

$$\begin{aligned} 3\Box_n &= (n-1)(n-\tfrac{1}{2})(n) + 3n^2 = \\ &= (n^3 - \tfrac{3}{2}n^2 + \tfrac{1}{2}n) + 3n^2 = \\ &= (n^3 + \tfrac{3}{2}n^2 + \tfrac{1}{2}n) = \\ &= n(n+\tfrac{1}{2})(n+1). \end{aligned}$$

Следовательно, (2.39) и в самом деле выполняется при всех $n \geqslant 0$". Судья в своей бесконечной мудрости благосклонно согласился с доводами адвоката.

Индукция здесь уместна и несколько более логична, чем поиск ответа. Но все равно это не то, чего мы хотим. Со всеми прочими суммами в этой главе мы справлялись без привлечения индукции, — давайте попробуем точно так же "с нуля" справиться и с $\square_n$. Нам ни к чему озарения, нам нужен метод, который не полагается на везение.

Метод 2: приведение суммы

Вернемся к методу приведения, который так хорошо сработал для геометрической прогрессии (2.25). Выделим первый и последний члены $\square_{n+1}$ для того, чтобы получить уравнение для $\square_n$:

$$\begin{aligned}
\square_n + (n+1)^2 &= \sum_{0\leqslant k\leqslant n} (k+1)^2 = \\
&= \sum_{0\leqslant k\leqslant n} (k^2+2k+1) = \\
&= \sum_{0\leqslant k\leqslant n} k^2 + 2\sum_{0\leqslant k\leqslant n} k + \sum_{0\leqslant k\leqslant n} 1 = \\
&= \square_n \; + \; 2\sum_{0\leqslant k\leqslant n} k \; + \; (n+1)\,.
\end{aligned}$$

Ой!.. Значения $\square_n$ сокращаются... Что ж, бывает и так, что несмотря на все усилия метод приведения приводит к чему-то в духе $\square_n = \square_n$, и в результате мы ничего не добиваемся.

Ничья?.. :)

С другой стороны, мы не зря тратили время и бумагу: мы нашли сумму первых n целых чисел в аналитическом виде

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

хотя на самом деле собирались найти сумму их квадратов. Логично предположить, что в поисках суммы кубов целых чисел, которые можно обозначить как $\boxdot_n$, мы получим выражение для суммы квадратов. Попробуем?

$$\begin{aligned}
\boxdot_n + (n+1)^3 &= \sum_{0\leqslant k\leqslant n} (k+1)^3 = \\
&= \sum_{0\leqslant k\leqslant n} (k^3+3k^2+3k+1) = \\
&= \boxdot_n + 3\square_n + 3\frac{(n+1)n}{2} + (n+1)\,.
\end{aligned}$$

Все в порядке: $\boxdot_n$ сокращаются, и мы получаем достаточно информации для определения $\Box_n$ безо всякой индукции:

Метод 2': приведение в смятение.

$$\begin{aligned} 3\Box_n &= (n+1)^3 - 3(n+1)n/2 - (n+1) = \\ &= (n+1)(n^2+2n+1-\tfrac{3}{2}n-1) = (n+1)(n+\tfrac{1}{2})n\,. \end{aligned}$$

Метод 3: построение набора

Для суммирования квадратов достаточно немного обобщить рекуррентное соотношение (2.7). Решение рекуррентности

$$\begin{aligned} &R_0 = \alpha\,; \\ &R_n = R_{n-1} + \beta + \gamma n + \delta n^2 \qquad \text{при } n > 0 \end{aligned} \tag{2.40}$$

будет иметь общий вид

$$R_n = A(n)\alpha + B(n)\beta + C(n)\gamma + D(n)\delta\,; \tag{2.41}$$

и мы уже определили $A(n)$, $B(n)$ и $C(n)$, поскольку (2.40) — это то же самое, что и (2.7) при $\delta = 0$. Если теперь подставить $R_n = n^3$, то выясняется, что n^3 является решением при $\alpha = 0$, $\beta = 1$, $\gamma = -3$ и $\delta = 3$. Следовательно,

$$3D(n) - 3C(n) + B(n) = n^3\,;$$

это уравнение определяет $D(n)$.

Нас интересует сумма $\Box_n$, равная $\Box_{n-1} + n^2$; если положить в (2.41) $\alpha = \beta = \gamma = 0$ и $\delta = 1$, то мы получим $\Box_n = R_n$. Следовательно, $\Box_n = D(n)$. Нам не нужны алгебраические преобразования для вычисления $D(n)$ из $B(n)$ и $C(n)$, поскольку мы уже знаем, каким будет ответ; самым твердолобым надо самостоятельно выполнить следующие действия:

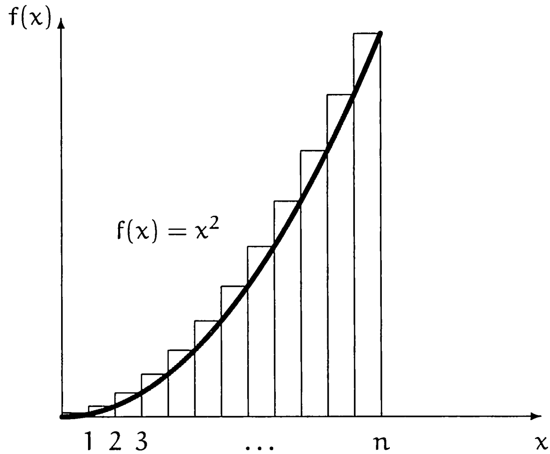
$$\begin{aligned} 3D(n) = n^3 + 3C(n) - B(n) &= n^3 + 3\frac{(n+1)n}{2} - n = \\ &= n(n+\tfrac{1}{2})(n+1)\,. \end{aligned}$$

Метод 4: замена сумм интегралами

Воспитанные на континуальной, а не дискретной математике обычно лучше знакомы с интегралами, чем с суммами, так что для них совершенно естественно заменить $\sum$ на $\int$. Одна из целей этой книги заключается в том, чтобы сделать $\sum$ настолько простой в обращении, чтобы $\int$ казался более сложным, чем $\sum$ (как минимум при точных вычислениях). Тем не менее проследить отношения между $\sum$ и $\int$ — неплохая идея, поскольку суммирование и интегрирование основаны на очень схожих идеях.

В математическом анализе интеграл может рассматриваться как площадь под некоторой кривой, и вычислить эту площадь можно приближенно путем сложения площадей узких прямоугольников, касающихся этой кривой. Если набор узких прямоугольников задан, можно пойти другим путем: поскольку $\square_n$ представляет собой сумму площадей прямоугольников размерами 1×1, 1×4, …, $1 \times n^2$, она приближенно равна площади под кривой $f(x) = x^2$ в интервале от 0 до n.

Здесь масштаб по горизонтали в десять раз превышает масштаб по вертикали.



Площадь под кривой равна $\int_0^n x^2\,dx = n^3/3$; следовательно, значение $\square_n$ приближенно равно $\frac{1}{3}n^3$.

Воспользоваться этим фактом можно для оценки ошибки приближения $E_n = \square_n - \frac{1}{3}n^3$. Поскольку $\square_n$ удовлетворяет рекуррентному соотношению $\square_n = \square_{n-1} + n^2$, мы находим, что E_n удовлетворяет более простому рекуррентному соотношению

$$\begin{aligned} E_n &= \square_n - \frac{1}{3}n^3 = \square_{n-1} + n^2 - \frac{1}{3}n^3 = \\ &= E_{n-1} + \tfrac{1}{3}(n-1)^3 + n^2 - \tfrac{1}{3}n^3 = E_{n-1} + n - \tfrac{1}{3}\,. \end{aligned}$$

Другой способ применения интегрального подхода состоит в поиске формулы для E_n путем суммирования площадей клиновидных фигур, составляющих погрешность. Имеем

Для тех, кто склонен к математическому анализу.

$$\begin{aligned} \square_n - \int_0^n x^2\,dx &= \sum_{k=1}^{n}\left(k^2 - \int_{k-1}^{k} x^2\,dx\right) = \\ &= \sum_{k=1}^{n}\left(k^2 - \frac{k^3 - (k-1)^3}{3}\right) = \sum_{k=1}^{n}\left(k - \tfrac{1}{3}\right). \end{aligned}$$

В любом случае находим сначала E_n, а затем — $\square_n$.

Метод 5: усложнение и упрощение

Еще один способ поиска аналитической записи для $\square_n$ заключается в замене исходной суммы кажущейся более сложной двойной суммой, которая при правильном подходе может быть упрощена:

$$\begin{aligned}
\square_n &= \sum_{1\leqslant k\leqslant n} k^2 = \sum_{1\leqslant j\leqslant k\leqslant n} k = \\
&= \sum_{1\leqslant j\leqslant n}\ \sum_{j\leqslant k\leqslant n} k = \\
&= \sum_{1\leqslant j\leqslant n} \left(\frac{j+n}{2}\right)(n-j+1) = \\
&= \tfrac{1}{2} \sum_{1\leqslant j\leqslant n} \bigl(n(n+1)+j-j^2\bigr) = \\
&= \tfrac{1}{2}n^2(n+1)+\tfrac{1}{4}n(n+1)-\tfrac{1}{2}\square_n = \\
&= \tfrac{1}{2}n(n+\tfrac{1}{2})(n+1)-\tfrac{1}{2}\square_n\,.
\end{aligned}$$

Последний шаг несколько напоминает последний шаг в методе приведения, так как мы получаем уравнение с неизвестной величиной с обеих сторон.

На первый взгляд, переход от одинарной суммы к двойной может показаться шагом назад, но на самом деле это шаг вперед, поскольку он приводит нас к сумме, с которой легче работать. Нельзя ожидать, что любая задача решается только постоянным упрощением: это так же наивно, как считать, что достичь горной вершины можно, только постоянно поднимаясь.

Метод 6: использование конечных разностей

Метод 7: использование производящих функций

Готовьтесь к встрече с еще более интересными способами вычисления $\square_n = \sum_{k=0}^{n} k^2$, которые будут изучаться в следующем разделе и последующих главах.

2.6 Исчисление конечного и бесконечного

Мы изучили ряд способов непосредственной работы с суммами. Теперь настало время расширить кругозор, взглянув на проблему суммирования на более высоком уровне. Математики разработали исчисление "конечных разностей", аналогичное более традиционному исчислению бесконечно малых, которое позволяет работать с суммированием более аккуратно и методично.

Исчисление бесконечно малых основано на свойствах оператора *дифференцирования* D, определяемого как

$$Df(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}\,.$$

Исчисление конечных разностей основано на свойствах *разностного* оператора Δ, определяемого как

$$\Delta f(x) = f(x+1) - f(x). \qquad (2.42)$$

Это — конечный аналог производной, в котором мы ограничиваемся положительными целыми значениями h. Таким образом, $h = 1$ — ближайшее к нулю значение, которое можно получить "в пределе" при $h \to 0$, а $\Delta f(x)$ представляет собой значение $\bigl(f(x+h) - f(x)\bigr)/h$ при $h = 1$.

Символы D и Δ называются *операторами*, поскольку они при работе с функциями дают нам новые функции; это функции от функций, производящие функции. Если f — достаточно гладкая функция, отображающая действительные числа на действительные числа, то Df также является функцией, отображающей действительные числа на действительные числа. И если f — *любая* действительная функция от действительного аргумента, то таковой же будет и функция Δf. Значения функций Df и Δf в точке x задаются приведенными выше определениями.

Из курса дифференциального исчисления известно, как оператор D действует на степенные функции $f(x) = x^m$. В таких случаях $Df(x) = mx^{m-1}$. Неформально это можно записать, опуская f:

$$D(x^m) = mx^{m-1}.$$

Было бы неплохо, если бы оператор Δ мог давать такой же элегантный результат; к сожалению, этого не происходит. Например,

$$\Delta(x^3) = (x+1)^3 - x^3 = 3x^2 + 3x + 1.$$

Математическая степень?

Но имеется разновидность "m-й степени", которая красиво трансформируется оператором Δ, и это обстоятельство делает конечные разности таким интересным предметом. Подобные видоизмененные m-е степени определяются правилом

$$x^{\underline{m}} = \overbrace{x(x-1)\ldots(x-m+1)}^{m\ \text{сомножителей}}, \qquad \text{целое } m \geqslant 0. \qquad (2.43)$$

Обратите внимание на маленький символ подчеркивания под m; он указывает, что m сомножителей пошагово становятся все меньше и меньше. Имеется соответствующее обозначение и для возрастающих сомножителей:

$$x^{\overline{m}} = \overbrace{x(x+1)\ldots(x+m-1)}^{m\ \text{сомножителей}}, \qquad \text{целое } m \geqslant 0. \qquad (2.44)$$

При $m = 0$ получаем $x^{\underline{0}} = x^{\overline{0}} = 1$, поскольку произведение не существующих сомножителей по соглашению полагается равным 1 (точно так же, как сумма отсутствующих слагаемых считается равной 0).

Если вам надо произнести вслух название величины $x^{\underline{m}}$, говорите "x в убывающей степени m"; аналогично $x^{\overline{m}}$ звучит как "x в возрастающей степени m". Эти функции называются также *убывающими факториальными степенями* и *возрастающими факториальными степенями*, поскольку они тесно связаны с факториальной функцией $n! = n(n-1)\ldots(1)$. И действительно, $n! = n^{\underline{n}} = 1^{\overline{n}}$.

В математической литературе встречаются и другие обозначения для факториальных степеней, в частности "символ Похгаммера" $(x)_m$ для $x^{\overline{m}}$ или $x^{\underline{m}}$; для $x^{\underline{m}}$ встречаются также записи наподобие $x^{(m)}$ или $x_{(m)}$. Но лидирующие позиции все же постепенно завоевывает подчеркнуто-надчеркнутое обозначение, поскольку оно легко записывается, запоминается и не требует лишних скобок.

Математическая терминология иногда кажется сущим бредом: на самом деле Похгаммер [293] использовал обозначение $(x)_m$ для биномиального коэффициента $\binom{x}{m}$, а не для факториальных степеней.

Убывающие степени $x^{\underline{m}}$ особенно хороши по отношению к Δ:

$$\begin{aligned}\Delta(x^{\underline{m}}) &= (x+1)^{\underline{m}} - x^{\underline{m}} = \\ &= (x+1)x\ldots(x-m+2) - x\ldots \\ &\qquad (x-m+2)(x-m+1) = \\ &= m\,x(x-1)\ldots(x-m+2)\,,\end{aligned}$$

так что исчисление конечных разностей располагает удобным в обращении правилом, соответствующим правилу $D(x^m) = mx^{m-1}$:

$$\Delta(x^{\underline{m}}) = mx^{\underline{m-1}}. \tag{2.45}$$

Это — фундаментальный факториальный факт.

Оператор D в исчислении бесконечно малых имеет обратный, антидифференциальный (или интегральный), оператор $\int$. Основная теорема дифференциального и интегрального исчислений связывает D с $\int$:

$$\begin{aligned} g(x) &= Df(x) \quad \text{тогда и только тогда,} \\ \text{когда} \quad \int g(x)\,dx &= f(x) + C\,. \end{aligned}$$

Здесь $\int g(x)\,dx$ — неопределенный интеграл $g(x)$ — представляет собой класс функций, производные от которых равны $g(x)$. Аналогично Δ также имеет обратный антиразностный (или сум-

"Quemadmodum ad differentiam denotandam usi sumus signo Δ, ita summam indicabimus signo Σ. ... ex quo æquatio $z = \Delta y$, si invertatur, dabit quoque $y = \Sigma z + C$."
— *Л. Эйлер (L. Euler) [110]*

мирующий) оператор $\sum$; и здесь действует иная основная теорема:

$$g(x) \;=\; \Delta f(x) \quad \text{тогда и только тогда,}$$
$$\text{когда} \quad \sum g(x)\,\delta x \;=\; f(x) + C\,. \tag{2.46}$$

Здесь $\sum g(x)\,\delta x$ — *неопределенная сумма* $g(x)$, класс функций, у которых *разность* равна $g(x)$. (Заметим, что строчная δ соотносится с прописной Δ так же, как строчная d соотносится с прописной D.) Буква "C" в неопределенных интегралах представляет собой произвольную константу; в случае неопределенных сумм "C" — произвольная функция $p(x)$, такая, что $p(x+1) = p(x)$. Например, C может быть периодической функцией $a + b\sin 2\pi x$; такие функции аннулируются при взятии разностей, так же как константы аннулируются при взятии производной. При целочисленных значениях x функция C представляет собой константу.

Теперь мы почти готовы к кульминационному моменту. Исчисление бесконечно малых включает в себя *определенные* интегралы: если $g(x) = Df(x)$, то

$$\int_a^b g(x)\,dx \;=\; f(x)\Big|_a^b \;=\; f(b) - f(a)\,.$$

Исчисление конечных разностей — во всем имитирующее своего более известного собрата — включает определенные *суммы*: если $g(x) = \Delta f(x)$, то

$$\sum\nolimits_a^b g(x)\,\delta x \;=\; f(x)\Big|_a^b \;=\; f(b) - f(a)\,. \tag{2.47}$$

Эта формула поясняет смысл записи $\sum_a^b g(x)\,\delta x$, так же, как предыдущая формула определяет $\int_a^b g(x)\,dx$.

Но что же на самом деле представляет собой $\sum_a^b g(x)\,\delta x$? Мы определили ее по аналогии, но не по существу. Эта аналогия позволяет хорошо запомнить правила исчисления конечных разностей, но она будет совершенно бесполезна, если мы не сможем понять ее значение. Давайте попытаемся вывести ее смысл, рассматривая сначала некоторые частные случаи, полагая, что $g(x) = \Delta f(x) = f(x+1) - f(x)$. Если $b = a$, то

$$\sum\nolimits_a^a g(x)\,\delta x \;=\; f(a) - f(a) \;=\; 0\,.$$

Затем, если $b = a + 1$, мы получим

$$\sum\nolimits_a^{a+1} g(x)\,\delta x \;=\; f(a+1) - f(a) \;=\; g(a)\,.$$

Если увеличить b на 1, то

$$\begin{aligned}\textstyle\sum_a^{b+1} g(x)\,\delta x - \sum_a^{b} g(x)\,\delta x &= \\ &= \big(f(b+1)-f(a)\big) - \big(f(b)-f(a)\big) = \\ &= f(b+1)-f(b) = g(b)\,.\end{aligned}$$

Эти наблюдения и метод математической индукции позволяют нам точно определить, что же означает $\sum_a^b g(x)\,\delta x$ в общем случае, когда a и b представляют собой целые числа, такие, что $b \geqslant a$:

$$\begin{aligned}\textstyle\sum_a^{b} g(x)\,\delta x &= \sum_{k=a}^{b-1} g(k) = \\ &= \sum_{a\leqslant k<b} g(k) \qquad \text{при целых } b \geqslant a. \qquad (2.48)\end{aligned}$$

Другими словами, определенная сумма — это то же самое, что и обычная сумма с пределами суммирования, но с исключенным значением верхнего предела.

Это и есть обещанная кульминация?!

Давайте попробуем повторить вывод несколько иначе. Предположим, что у нас имеется неизвестная сумма, которая, как мы считаем, вычислима в аналитическом виде, причем ее можно записать в следующем виде: $\sum_{a\leqslant k<b} g(k) = \sum_a^b g(x)\,\delta x$. Теория исчисления конечных разностей гласит, что можно выразить ответ как $f(b) - f(a)$, если найти неопределенную сумму или антиразностную функцию f, такую, что $g(x) = f(x+1) - f(x)$. Один из способов понять этот принцип заключается в том, чтобы записать $\sum_{a\leqslant k<b} g(k)$ в развернутом виде, используя запись с троеточием:

$$\begin{aligned}\sum_{a\leqslant k<b} \big(f(k+1)-f(k)\big) &= \\ &= \big(f(a{+}1)-f(a)\big) + \big(f(a{+}2)-f(a{+}1)\big) + \cdots + \\ &\quad + \big(f(b{-}1)-f(b{-}2)\big) + \big(f(b)-f(b{-}1)\big)\,.\end{aligned}$$

Все, что находится в правой части (за исключением $f(b) - f(a)$), сокращается, так что значением суммы является $f(b) - f(a)$. (Суммы вида $\sum_{a\leqslant k<b}\big(f(k+1) - f(k)\big)$ часто называются *телескопическими*, по аналогии со складным телескопом, поскольку толщина стенок сложенного телескопа определяется только радиусами самой большой и самой малой из труб.)

И все время моих размышлений ушло на телескопирование очень длинного выражения в очень короткое.

Правило (2.48) применимо только при $b \geqslant a$; а что же будет в случае $b < a$? Согласно (2.47)

$$\begin{aligned}\sum\nolimits_a^b g(x)\,\delta x \;&=\; f(b)-f(a) \;=\\ &=\; -\bigl(f(a)-f(b)\bigr) \;=\; -\sum\nolimits_b^a g(x)\,\delta x\,,\end{aligned}$$

что аналогично соответствующему уравнению для определенного интеграла. Подобные рассуждения доказывают, что $\sum_a^b + \sum_b^c = \sum_a^c$ — аналог соотношения $\int_a^b + \int_b^c = \int_a^c$ для сумм. Вот его полный вид:

$$\sum\nolimits_a^b g(x)\,\delta x + \sum\nolimits_b^c g(x)\,\delta x \;=\; \sum\nolimits_a^c g(x)\,\delta x \tag{2.49}$$

для всех целых a, b и c.

Вероятно, сейчас некоторые из читателей начнут интересоваться: а что все это нам дает? Начнем с того, что определенные суммы дают нам простой способ вычисления сумм убывающих степеней: основные правила (2.45), (2.47) и (2.48) приводят к общему правилу

Остальных этот вопрос интересует уже давно...

$$\sum_{0\leqslant k<n} k^{\underline{m}} \;=\; \frac{k^{\underline{m+1}}}{m+1}\bigg|_0^n \;=\; \frac{n^{\underline{m+1}}}{m+1} \quad \text{при целых } m, n \geqslant 0. \tag{2.50}$$

Эту формулу легко запомнить, потому что она очень похожа на хорошо известную формулу $\int_0^n x^m\,dx = n^{m+1}/(m+1)$.

В частности, при $m = 1$ мы получаем $k^{\underline{1}} = k$, поэтому исчисление конечных разностей позволяет нам легко запомнить тот факт, что

$$\sum_{0\leqslant k<n} k \;=\; \frac{n^{\underline{2}}}{2} \;=\; n(n-1)/2\,.$$

Метод определенных сумм, помимо прочего, намекает, что суммы в диапазоне $0 \leqslant k < n$ часто оказываются проще сумм в диапазоне $1 \leqslant k \leqslant n$; первые равны $f(n) - f(0)$, а последние требуют вычисления $f(n+1) - f(1)$.

Обычные степени также можно суммировать этим новым способом, если сначала выразить их через убывающие степени. Например,

$$k^2 \;=\; k^{\underline{2}} + k^{\underline{1}},$$

так что

$$\begin{aligned}\sum_{0\leqslant k<n} k^2 \;&=\; \frac{n^{\underline{3}}}{3} + \frac{n^{\underline{2}}}{2} \;=\; \tfrac{1}{3}n(n-1)(n-2+\tfrac{3}{2}) \;=\\ &=\; \frac{1}{3}n(n-\frac{1}{2})(n-1)\,.\end{aligned}$$

Заменяя n на $n+1$, мы получим еще один способ вычисления нашего старого знакомого — $\square_n = \sum_{0\leqslant k\leqslant n} k^2$ — в аналитическом виде.

Старый друг лучше новых двух...

Как все просто! Это и правда проще любого из груды способов, которой мы завалили нашу сумму в предыдущем разделе. Давайте не останавливаться на достигнутом и обратимся к *кубам*. Простые вычисления показывают, что

$$k^3 = k^{\underline{3}} + 3k^{\underline{2}} + k^{\underline{1}}.$$

(Обычные степени всегда можно преобразовать в факториальные и обратно при помощи чисел Стирлинга, с которыми мы познакомимся в главе 6.) Таким образом,

$$\sum_{a\leqslant k<b} k^3 = \frac{k^{\underline{4}}}{4} + k^{\underline{3}} + \frac{k^{\underline{2}}}{2}\bigg|_a^b.$$

Убывающие степени, как видно, отлично подходят для вычисления сумм. Но нет ли у них иных качеств, оправдывающих их введение в математику? Должны ли мы преобразовывать обычные степени в убывающие перед суммированием, а затем выполнять обратное преобразование перед тем, как двигаться дальше, или можно поступать как-то иначе? Да, зачастую можно работать непосредственно с факториальными степенями, обладающими дополнительными свойствами. Например, точно так же, как $(x+y)^2 = x^2 + 2xy + y^2$, справедливо и $(x+y)^{\underline{2}} = x^{\underline{2}} + 2x^{\underline{1}}y^{\underline{1}} + y^{\underline{2}}$, причем та же аналогия имеется и между $(x+y)^m$ и $(x+y)^{\underline{m}}$. (Эта "факториальная биномиальная теорема" доказывается в упр. 5.37.)

До сих пор мы рассматривали убывающие степени только с неотрицательными показателями. Чтобы распространить аналогию с обычными степенями и на отрицательные показатели, необходимо подходящее определение $x^{\underline{m}}$ при $m < 0$. Глядя на последовательность

$$\begin{aligned}
x^{\underline{3}} &= x(x-1)(x-2)\,,\\
x^{\underline{2}} &= x(x-1)\,,\\
x^{\underline{1}} &= x\,,\\
x^{\underline{0}} &= 1\,,
\end{aligned}$$

можно заметить, что для перехода от $x^{\underline{3}}$ через $x^{\underline{2}}$ и $x^{\underline{1}}$ к $x^{\underline{0}}$ выполняется деление на $x-2$, затем на $x-1$, затем на x. Представляется разумным (если не необходимым) очередное деление на $x+1$ для перехода от $x^{\underline{0}}$ к $x^{\underline{-1}}$, получая, таким образом, $x^{\underline{-1}} = 1/(x+1)$.

Продолжая эту последовательность делений, мы получим следующие первые отрицательные убывающие степени:

$$\begin{aligned} x^{\underline{-1}} &= \frac{1}{x+1}, \\ x^{\underline{-2}} &= \frac{1}{(x+1)(x+2)}, \\ x^{\underline{-3}} &= \frac{1}{(x+1)(x+2)(x+3)}, \end{aligned}$$

и в общем случае определение отрицательных убывающих степеней имеет вид

$$x^{\underline{-m}} = \frac{1}{(x+1)(x+2)\ldots(x+m)} \qquad \text{при } m > 0. \qquad (2.51)$$

Но вы не комплексуйте...

(Можно определить убывающие степени и для действительных, и даже для комплексных m, но это мы отложим до главы 5.)

При таком определении убывающие степени обретают дополнительные замечательные свойства. Вероятно, наиболее важным является обобщенный закон степеней, аналогичный закону

$$x^{m+n} = x^m x^n$$

для обычных степеней. Убывающая версия имеет вид

$$x^{\underline{m+n}} = x^{\underline{m}}(x-m)^{\underline{n}}, \qquad \text{целые } m \text{ и } n. \qquad (2.52)$$

Например, $x^{\underline{2+3}} = x^{\underline{2}}(x-2)^{\underline{3}}$; а при отрицательном n

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

Законы обладают и положительными, и отрицательными сторонами.

Если бы мы определили $x^{\underline{-1}}$ как $1/x$, а не как $1/(x+1)$, то закон степеней (2.52) не выполнялся бы в случаях наподобие $m = -1$ и $n = 1$. В действительности можно было бы использовать (2.52) для определения убывающих степеней с отрицательными показателями, положив $m = -n$. Когда существующее обозначение расширяется на большее количество случаев, всегда лучше формулировать определения таким образом, чтобы общие правила оставались справедливыми.

Убедимся теперь, что ключевое разностное свойство справедливо и для вновь определенных убывающих степеней. Действительно ли $\Delta x^{\underline{m}} = m x^{\underline{m-1}}$ при $m < 0$? Если, например, $m = -2$, то разность равна

$$\Delta x^{\underline{-2}} = \frac{1}{(x+2)(x+3)} - \frac{1}{(x+1)(x+2)} = $$
$$= \frac{(x+1)-(x+3)}{(x+1)(x+2)(x+3)} = $$
$$= -2x^{\underline{-3}}.$$

Да, все верно! Аналогичные рассуждения применимы для всех $m < 0$.

Следовательно, правило суммирования (2.50) выполняется как для положительных, так и для отрицательных убывающих степеней, пока отсутствует деление на нуль:

$$\sum\nolimits_a^b x^{\underline{m}}\,\delta x = \frac{x^{\underline{m+1}}}{m+1}\bigg|_a^b \qquad \text{при } m \neq -1.$$

А что делать при $m = -1$? Вспомним, что в случае интегрирования при $m = -1$

$$\int_a^b x^{-1}\,dx = \ln x\,\bigg|_a^b.$$

Неплохо бы получить конечно-разностный аналог $\ln x$; другими словами, мы хотим найти функцию $f(x)$, такую, что

$$x^{\underline{-1}} = \frac{1}{x+1} = \Delta f(x) = f(x+1) - f(x).$$

Несложно увидеть, что такой функцией является

$$f(x) = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{x}$$

при целом x, и эта величина представляет собой гармоническое число H_x из (2.13). Таким образом, H_x — дискретный аналог континуального $\ln x$. (Мы определим H_x для нецелых x в главе 6, а пока что для наших целей достаточно целочисленных значений. В главе 9 мы также увидим, что при больши́х x значение $H_x - \ln x$ приближенно равно $0.577 + 1/(2x)$. Следовательно, H_x и $\ln x$ — не просто аналоги, раз их значения обычно отличаются меньше чем на 1.)

Ровно 0.577? Может, это $1/\sqrt{3}$. А может и нет...

Теперь мы можем дать полное описание сумм убывающих степеней:

$$\sum\nolimits_a^b x^{\underline{m}}\,\delta x \;=\; \begin{cases} \left.\dfrac{x^{\underline{m+1}}}{m+1}\right|_a^b & \text{при } m \neq -1; \\[2ex] \left.H_x\right|_a^b & \text{при } m = -1. \end{cases} \tag{2.53}$$

Эта формула указывает, почему так гармонично при решении дискретных задач наподобие анализа быстрой сортировки возникают гармонические числа — так же, как натуральные логарифмы весьма натуральны для континуальных задач.

Теперь, когда найден аналог $\ln x$, давайте поищем аналог для e^x. Какая функция $f(x)$ обладает тем свойством, что $\Delta f(x) = f(x)$, соответствующая тождеству $De^x = e^x$? Это просто:

$$f(x+1) - f(x) \;=\; f(x) \qquad \Longleftrightarrow \qquad f(x+1) \;=\; 2f(x)\,,$$

так что мы имеем дело с простой рекуррентностью и можем принять $f(x) = 2^x$ в качестве дискретной экспоненты.

Разность функции c^x для произвольного c также достаточно проста, а именно

$$\Delta(c^x) \;=\; c^{x+1} - c^x \;=\; (c-1)c^x\,.$$

Следовательно, антиразностью c^x является $c^x/(c-1)$, если $c \neq 1$. Этот факт наряду с фундаментальными законами (2.47) и (2.48) дает нам компактный способ получения общей формулы для суммы геометрической прогрессии:

$$\sum_{a\leqslant k<b} c^k \;=\; \sum\nolimits_a^b c^x\,\delta x \;=\; \left.\frac{c^x}{c-1}\right|_a^b \;=\; \frac{c^b - c^a}{c-1} \qquad \text{при } c \neq 1.$$

Всякий раз, когда мы сталкиваемся с функцией f, которая может быть полезна в аналитическом виде, мы можем вычислить ее разность $\Delta f = g$; так мы получаем функцию g, неопределенная сумма которой $\sum g(x)\,\delta x$ нам известна. Таблица 80 представляет собой начало большой таблицы перечня пар "разность/антиразность" полезных для вычисления сумм.

'Таблица 80' — это та, которая на странице 80. Уловили?

Несмотря на все параллели между континуальной и дискретной математикой некоторые континуальные понятия не имеют дискретных аналогов. Например, "правило цепочек" в исчислении бесконечно малых — удобное правило для вычисления производной функции от функции; однако соответствующего правила в исчислении конечных разностей нет, так как нет красивой формулы для $\Delta f\bigl(g(x)\bigr)$. В дискретном случае затруднена и замена переменных, за исключением некоторых случаев наподобие замены x на $c \pm x$.

Таблица 80. Суммы и разности

$f = \Sigma g$	$\Delta f = g$	$f = \Sigma g$	$\Delta f = g$
$x^{\underline{0}} = 1$	0	2^x	2^x
$x^{\underline{1}} = x$	1	c^x	$(c-1)c^x$
$x^{\underline{2}} = x(x-1)$	$2x$	$c^x/(c-1)$	c^x
$x^{\underline{m}}$	$mx^{\underline{m-1}}$	cf	$c\Delta f$
$x^{\underline{m+1}}/(m+1)$	$x^{\underline{m}}$	$f+g$	$\Delta f + \Delta g$
H_x	$x^{\underline{-1}} = 1/(x+1)$	fg	$f\Delta g + Eg\Delta f$

Однако $\Delta\bigl(f(x)\,g(x)\bigr)$ *имеет* красивое выражение, которое предоставляет нам правило для *суммирования по частям* — конечно-разностный аналог того, что в исчислении бесконечно малых именуется интегрированием по частям. Вспомним, что в исчислении бесконечно малых формула

$$D(uv) = u\,Dv + v\,Du$$

приводит к правилу интегрирования по частям

$$\int u\,Dv = uv - \int v\,Du$$

после интегрирования и подстановки членов; то же самое можно сделать в исчислении конечных разностей.

Начнем с применения разностного оператора к произведению функций $u(x)$ и $v(x)$:

$$\begin{aligned}\Delta\bigl(u(x)\,v(x)\bigr) &= u(x{+}1)\,v(x{+}1) - u(x)\,v(x) = \\ &= u(x{+}1)\,v(x{+}1) - u(x)\,v(x{+}1) + \\ &\qquad + u(x)\,v(x{+}1) - u(x)\,v(x) = \\ &= u(x)\,\Delta v(x) + v(x{+}1)\,\Delta u(x)\,. \end{aligned} \tag{2.54}$$

Эту формулу можно привести к более удобному виду, воспользовавшись *оператором сдвига* E, определяемым как

$$Ef(x) = f(x+1)\,.$$

Подставляя $Ev(x)$ вместо $v(x{+}1)$, получим компактное правило для разности произведения:

$$\Delta(uv) = u\,\Delta v + Ev\,\Delta u\,. \tag{2.55}$$

(Наличие E представляет определенное неудобство, но при этом делает уравнение корректным.) Беря неопределенную сумму от обеих частей этого уравнения и переставляя члены, мы получаем

В исчислении бесконечно малых $1 \to 0$, и E попросту исчезает.

обещанное правило суммирования по частям:

$$\sum u\,\Delta v \;=\; uv - \sum Ev\,\Delta u\,. \tag{2.56}$$

Как и в исчислении бесконечно малых, ко всем трем членам можно добавить пределы суммирования, делая неопределенные суммы определенными.

Я все понял: $e^x = 2^x$ при малых значениях 1...

Это правило полезно, когда сумму слева вычислить сложнее, чем сумму справа. Рассмотрим конкретный пример. Интеграл $\int xe^x\,dx$ обычно вычисляется путем интегрирования по частям; его дискретным аналогом является сумма $\sum x2^x\,\delta x$, с которой мы уже сталкивались в этой главе. Тогда она была записана в виде $\sum_{k=0}^{n} k\,2^k$. Для суммирования по частям положим $u(x) = x$ и $\Delta v(x) = 2^x$; отсюда $\Delta u(x) = 1$, $v(x) = 2^x$ и $Ev(x) = 2^{x+1}$. Подстановка в (2.56) дает

$$\sum x2^x\,\delta x \;=\; x2^x \;-\; \sum 2^{x+1}\,\delta x \;=\; x2^x \;-\; 2^{x+1} \;+\; C\,.$$

Теперь можно использовать эту формулу для вычисления ранее упоминавшейся суммы, добавив пределы суммирования:

$$\begin{aligned}\sum_{k=0}^{n} k2^k \;&=\; \sum\nolimits_0^{n+1} x2^x\,\delta x \;=\\ &=\; x2^x - 2^{x+1}\,\Big|_0^{n+1} \;=\\ &=\; \big((n+1)2^{n+1} - 2^{n+2}\big) - (0\cdot 2^0 - 2^1) \;=\\ &=\; (n-1)2^{n+1} + 2\,.\end{aligned}$$

Найти сумму таким способом проще, чем при помощи метода приведения, поскольку при этом не надо ни о чем задумываться.

Быть может, сверхцель математики — избавить человечество от размышлений?

Ранее в главе мы наткнулись на формулу для $\sum_{0\leqslant k<n} H_k$ и прыгали от радости. Но формулу (2.36) можно было бы обнаружить закономерно, зная о суммировании по частям. Продемонстрируем это утверждение, взявшись за вычисление еще более сложной на вид суммы $\sum_{0\leqslant k<n} kH_k$. Решение оказывается несложным, если исходить из аналогии с интегралом $\int x\ln x\,dx$: положим $u(x) = H_x$ и $\Delta v(x) = x = x^{\underline{1}}$, так что $\Delta u(x) = x^{\underline{-1}}$, $v(x) = x^{\underline{2}}/2$, $Ev(x) = (x+1)^{\underline{2}}/2$ и мы получаем

$$\begin{aligned}\sum xH_x\,\delta x \;&=\; \frac{x^{\underline{2}}}{2}H_x \;-\; \sum \frac{(x+1)^{\underline{2}}}{2}\,x^{\underline{-1}}\,\delta x \;=\\ &=\; \frac{x^{\underline{2}}}{2}H_x \;-\; \frac{1}{2}\sum x^{\underline{1}}\,\delta x \;=\\ &=\; \frac{x^{\underline{2}}}{2}H_x \;-\; \frac{x^{\underline{2}}}{4} \;+\; C\,.\end{aligned}$$

(При переходе от первой строки ко второй мы объединили две убывающие степени $(x+1)^{\underline{2}}x^{\underline{-1}}$, используя правило показателей (2.52) с $m=-1$ и $n=2$.) Теперь можно добавить пределы и заключить, что

$$\sum_{0\leqslant k<n} kH_k \;=\; \sum\nolimits_0^n xH_x\,\delta x \;=\; \frac{n^{\underline{2}}}{2}\Bigl(H_n-\tfrac{1}{2}\Bigr)\,. \qquad (2.57)$$

2.7 Бесконечные суммы

Определяя обозначение $\sum$ в начале этой главы, мы схитрили в вопросе о бесконечных суммах, по сути, заявив: “Отложим этот вопрос на потом, а пока будем считать, что все суммы состоят только из конечного количества ненулевых членов”. Пришло время расплаты — мы обязаны признать тот факт, что суммы могут быть и бесконечными. Истина заключается в том, что бесконечные суммы — как и все в этом мире — имеют свои хорошие и плохие стороны.

И это — хитрость?

Начнем с плохих новостей: оказывается, что методы, которыми мы пользовались при работе с суммами, *не* всегда применимы к бесконечным суммам. Хорошая же новость заключается в том, что имеется большой, простой в понимании класс бесконечных сумм, для которых все рассмотренные операции вполне применимы. Причины, лежащие в основе обеих новостей, станут понятны после того, как мы более близко познакомимся с суммированием.

Все знают, что такое конечная сумма: мы добавляем к общему результату все новые и новые слагаемые, пока они не иссякнут. Но определять бесконечную сумму следует более аккуратно, чтобы не столкнуться с парадоксами.

Например, представляется естественным считать, что бесконечная сумма

$$S \;=\; 1+\tfrac{1}{2}+\tfrac{1}{4}+\tfrac{1}{8}+\tfrac{1}{16}+\tfrac{1}{32}+\cdots$$

равна 2, поскольку при ее удвоении мы получим

$$2S \;=\; 2+1+\tfrac{1}{2}+\tfrac{1}{4}+\tfrac{1}{8}+\tfrac{1}{16}+\cdots \;=\; 2+S\,.$$

С другой стороны, применяя те же рассуждения к сумме

$$T \;=\; 1+2+4+8+16+32+\cdots,$$

мы получаем, что она равна -1, поскольку при ее удвоении мы получаем

$$2T \;=\; 2+4+8+16+32+64+\cdots \;=\; T-1.$$

Вполне логично: $1+2+4+8+\cdots$ — не что иное, как “бесконечно точное” представление числа -1 в двоичном компьютере с неограниченным размером слова.

Забавно, не правда ли, получить отрицательное число, суммируя положительные значения? Пожалуй, лучше оставить сумму T неопределенной, или, быть может, следует считать, что $T = \infty$, поскольку слагаемые суммы становятся больше любого заранее заданного фиксированного конечного числа. (Заметим, что ∞ представляет собой второе "решение" уравнения $2T = T - 1$; она же "решает" и уравнение $2S = 2 + S$.)

Давайте попробуем сформулировать хорошее определение значения обобщенной суммы $\sum_{k \in K} a_k$, где множество K может быть бесконечным. Для начала предположим, что все члены a_k *неотрицательны*. В таком случае подходящее определение найти нетрудно: если для любого *конечного* подмножества $F \subset K$ существует ограничивающая постоянная A, такая, что

$$\sum_{k \in F} a_k \leqslant A,$$

то мы определяем сумму $\sum_{k \in K} a_k$ как равную *наименьшему* такому A. (Как следует из хорошо известных свойств действительных чисел, множество всех таких A всегда содержит наименьший элемент.) Но если такой константы A не существует, мы говорим, что $\sum_{k \in K} a_k = \infty$; это означает, что если A — некоторое действительное число, то имеется множество из конечного количества членов a_k, сумма которых превосходит A.

Определение в предыдущем абзаце сформулировано достаточно аккуратно, чтобы не зависеть от порядка, который может иметься у множества индексов K. Поэтому аргументация, которая будет приведена далее, будет справедливой не только для сумм по множеству целых чисел, но и для кратных сумм со многими индексами $k_1, k_2, \ldots$.

Множество K может даже быть несчетным. Но если существует ограничивающая константа A, то только счетное количество членов может быть отлично от нуля, поскольку самое большее nA членов $\geqslant 1/n$.

В частном случае, когда K — множество неотрицательных целых чисел, наше определение для неотрицательных членов a_k приводит к

$$\sum_{k \geqslant 0} a_k = \lim_{n \to \infty} \sum_{k=0}^{n} a_k .$$

И вот почему: любая неубывающая последовательность действительных чисел имеет предел (возможно, равный ∞). Если этот предел равен A и если F — некоторое конечное множество неотрицательных чисел, все из которых $\leqslant n$, то $\sum_{k \in F} a_k \leqslant \sum_{k=0}^{n} a_k \leqslant A$; следовательно, либо $A = \infty$, либо A — ограничивающая постоянная. И если A' — некоторое число, меньшее установленной границы A, то существует такое n, что $\sum_{k=0}^{n} a_k >$

A'; следовательно, конечное множество $F = \{0, 1, \ldots, n\}$ свидетельствует о том, что A' ограничивающей постоянной не является.

Теперь можно легко вычислить величины конкретных бесконечных сумм в соответствии с только что данным определением. Например, если $a_k = x^k$, то

$$\sum_{k\geqslant 0} x^k = \lim_{n\to\infty} \frac{1-x^{n+1}}{1-x} = \begin{cases} 1/(1-x), & \text{если } 0 \leqslant x < 1; \\ \infty, & \text{если } x \geqslant 1. \end{cases}$$

В частности, бесконечные суммы S и T, которые рассматривались минуту назад, имеют значения соответственно 2 и ∞, как мы и предполагали. Другой интересный пример:

$$\sum_{k\geqslant 0} \frac{1}{(k+1)(k+2)} = \sum_{k\geqslant 0} k^{\underline{-2}} =$$
$$= \lim_{n\to\infty} \sum_{k=0}^{n-1} k^{\underline{-2}} = \lim_{n\to\infty} \frac{k^{\underline{-1}}}{-1}\bigg|_0^n = 1\,.$$

Рассмотрим теперь случай, когда сумма наряду с неотрицательными членами может содержать и отрицательные. Чему, например, равна сумма

$$\sum_{k\geqslant 0} (-1)^k = 1 - 1 + 1 - 1 + 1 - 1 + \cdots ?$$

Если сгруппировать члены попарно, можно получить

$$(1-1) + (1-1) + (1-1) + \cdots = 0 + 0 + 0 + \cdots,$$

Таким образом, сумма равна нулю; но если начать группирование по парам на шаг позже, то можно получить

$$1 - (1-1) - (1-1) - (1-1) - \cdots = 1 - 0 - 0 - 0 - \cdots,$$

т.е. сумма равна 1.

“Aggregatum quantitatum $a - a + a - a + a - a$ etc. nunc est $= a$, nunc $= 0$, adeoque continuata in infinitum serie ponendus $= a/2$, fateor acumen et veritatem animadversionis tuæ.”

— Г. Гранди (G. Grandi) [163]

Можно, конечно, попробовать подставить значение $x = -1$ в формулу $\sum_{k\geqslant 0} x^k = 1/(1-x)$, которая, как мы доказали, справедлива при $0 \leqslant x < 1$; но тогда придется заключить, что сумма (пусть и бесконечная) *целых чисел* равна $\frac{1}{2}$!

Еще один занимательный пример — “дважды бесконечная” сумма $\sum_k a_k$, в которой $a_k = 1/(k+1)$ при $k \geqslant 0$ и $a_k = 1/(k-1)$ при $k < 0$. Ее можно записать как

$$\cdots + (-\tfrac{1}{4}) + (-\tfrac{1}{3}) + (-\tfrac{1}{2}) + 1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \cdots . \qquad (2.58)$$

Если вычислить эту сумму, начиная с "центрального" элемента и двигаясь наружу,

$$\cdots + \Bigl(-\tfrac{1}{4} + \bigl(-\tfrac{1}{3} + (-\tfrac{1}{2} + (1) + \tfrac{1}{2}) + \tfrac{1}{3}\bigr) + \tfrac{1}{4}\Bigr) + \cdots,$$

то мы получим значение 1; то же значение 1 будет получено, если мы сдвинем все скобки на один элемент влево,

$$\cdots + \Bigl(-\tfrac{1}{5} + \bigl(-\tfrac{1}{4} + (-\tfrac{1}{3} + (-\tfrac{1}{2}) + 1) + \tfrac{1}{2}\bigr) + \tfrac{1}{3}\Bigr) + \cdots,$$

так как сумма всех чисел в n внутренних скобках равна

$$-\frac{1}{n+1} - \frac{1}{n} - \cdots - \frac{1}{2} + 1 + \frac{1}{2} + \cdots + \frac{1}{n-1} = \\ = 1 - \frac{1}{n} - \frac{1}{n+1}.$$

Аналогичные рассуждения показывают, что величина суммы остается равной 1 и при сдвиге на любое фиксированное число элементов влево или вправо. Это заставляет нас решить, что данная сумма действительно равна 1. Но, с другой стороны, если сгруппировать члены следующим образом:

$$\cdots + \Bigl(-\tfrac{1}{4} + \bigl(-\tfrac{1}{3} + (-\tfrac{1}{2} + 1 + \tfrac{1}{2}) + \tfrac{1}{3} + \tfrac{1}{4}\bigr) + \tfrac{1}{5} + \tfrac{1}{6}\Bigr) + \cdots,$$

то n-я пара внутренних скобок будет содержать числа

$$-\frac{1}{n+1} - \frac{1}{n} - \cdots - \frac{1}{2} + 1 + \frac{1}{2} + \cdots + \frac{1}{2n-1} + \frac{1}{2n} = \\ = 1 + H_{2n} - H_{n+1}.$$

В главе 9 будет доказано, что $\lim_{n\to\infty}(H_{2n} - H_{n+1}) = \ln 2$; следовательно, такое группирование предполагает, что дважды бесконечная сумма должна быть равна $1 + \ln 2$.

Есть нечто бессмысленное в сумме, которая дает разные значения при сложении ее членов разными способами. В современных руководствах по анализу имеется целый ряд определений, с помощью которых таким патологическим суммам приписываются осмысленные значения; но если мы позаимствуем эти определения, то не сможем оперировать $\sum$-обозначениями так же свободно, как делали это до сих пор. Цели этой книги таковы, что нам не нужны рафинированные уточнения понятия "условной сходимости" — мы будем придерживаться такого определения бесконечных сумм, которое оставляет в силе все использованные нами в настоящей главе операции.

В действительности наше определение бесконечных сумм достаточно простое. Пусть K — некоторое множество и пусть a_k —

действительный член суммы, определенный для каждого $k \in K$. (Здесь 'k' может на самом деле означать несколько индексов k_1, k_2, ..., так что множество K может быть многомерным.) Любое действительное число x можно записать как разность его положительной и отрицательной частей,

$$x = x^+ - x^-, \qquad \text{где } x^+ = x \cdot [x > 0] \text{ и } x^- = -x \cdot [x < 0].$$

(Либо $x^+ = 0$, либо $x^- = 0$, либо оба.) Мы уже объясняли, как определять величины бесконечных сумм $\sum_{k \in K} a_k^+$ и $\sum_{k \in K} a_k^-$, поскольку a_k^+ и a_k^- неотрицательны. Поэтому наше общее определение таково:

$$\sum_{k \in K} a_k = \sum_{k \in K} a_k^+ - \sum_{k \in K} a_k^-, \tag{2.59}$$

если только обе суммы в правой части не равны ∞. В последнем случае сумма $\sum_{k \in K} a_k$ остается неопределенной.

Пусть $A^+ = \sum_{k \in K} a_k^+$ и $A^- = \sum_{k \in K} a_k^-$. Если обе суммы, A^+ и A^-, конечны, то говорят, что сумма $\sum_{k \in K} a_k$ *абсолютно сходится* к значению $A = A^+ - A^-$. Если $A^+ = \infty$, а A^- конечна, то говорят, что сумма $\sum_{k \in K} a_k$ *расходится* к $+\infty$. Аналогично, если $A^- = \infty$, а A^+ конечна, то говорят, что $\sum_{k \in K} a_k$ расходится к $-\infty$. Если $A^+ = A^- = \infty$, то не говорят ничего.

Другими словами, абсолютная сходимость означает сходимость суммы абсолютных величин.

Мы начали с определения, которое работало для неотрицательных членов, затем распространили его на действительные члены с любыми значениями. Если члены суммы a_k являются комплексными числами, наше определение можно очевидным способом расширить еще раз: сумма $\sum_{k \in K} a_k$ определяется как $\sum_{k \in K} \Re a_k + i \sum_{k \in K} \Im a_k$, где $\Re a_k$ и $\Im a_k$ — действительная и мнимая части a_k при условии, что обе эти суммы определены. В противном случае сумма $\sum_{k \in K} a_k$ не определена (см. упр. 18).

Плохие новости, как упоминалось ранее, заключаются в том, что некоторые бесконечные суммы приходится оставлять неопределенными, поскольку операции, которые мы выполняем с ними, могут приводить к несуразностям (см. упр. 34). Хорошие же новости заключаются в том, что что все операции из настоящей главы совершенно справедливы всякий раз, когда мы имеем дело с абсолютно сходящимися в только что установленном смысле суммами.

Мы можем убедиться в последнем, продемонстрировав, что каждое из наших правил преобразования сохраняет величину любой абсолютно сходящейся суммы неизменной. Более точно это означает, что мы должны доказать дистрибутивный, ассоциативный и коммутативный законы, а также правило, согласно которому можно начинать суммирование по любой индексной переменной. Все остальное, что делалось в данной главе, может быть выведено из этих четырех фундаментальных операций с суммами.

Дистрибутивный закон (2.15) можно сформулировать более строго следующим образом: если сумма $\sum_{k\in K} a_k$ абсолютно сходится к A и если c — некоторое комплексное число, то сумма $\sum_{k\in K} ca_k$ абсолютно сходится к cA. Это можно доказать, разбив сумму сначала на действительную и мнимую, а затем на положительную и отрицательную части, как делалось выше, и доказывая частный случай, когда $c > 0$ и каждый член a_k неотрицателен. Доказательство в этом частном случае работает в силу того, что $\sum_{k\in F} ca_k = c\sum_{k\in F} a_k$ для всех конечных множеств F; последний же факт доказывается индукцией по размеру F.

Ассоциативный закон (2.16) можно сформулировать следующим образом: если $\sum_{k\in K} a_k$ и $\sum_{k\in K} b_k$ абсолютно сходятся соответственно к A и B, то $\sum_{k\in K}(a_k + b_k)$ абсолютно сходится к $A + B$. Оказывается, это является частным случаем более общей теоремы, которую мы вскоре докажем.

Коммутативный закон (2.17) в действительности доказывать не обязательно, так как при рассмотрении (2.35) было показано, как вывести его в качестве частного случая более общего правила изменения порядка суммирования.

Когда вы окажетесь здесь впервые, лучше просто бегло просмотрите эту страницу.
— Сочувствующий ассистент

Главное, что требуется доказать, — это фундаментальный принцип кратных сумм: *абсолютно сходящиеся суммы с двумя или более индексами всегда можно начинать суммировать по любому из этих индексов.* Формально мы должны доказать, что если J и элементы $\{K_j \mid j \in J\}$ — некоторые множества индексов, такие, что

$$\sum_{\substack{j\in J\\ k\in K_j}} a_{j,k} \quad \text{абсолютно сходится к } A\,,$$

то для каждого $j \in J$ существуют комплексные числа A_j, такие, что

$$\sum_{k\in K_j} a_{j,k} \quad \text{абсолютно сходится к } A_j \quad \text{и}$$

$$\sum_{j\in J} A_j \quad \text{абсолютно сходится к } A\,.$$

Достаточно доказать это утверждение для случая, когда все члены суммы неотрицательны, так как в общем случае все можно разбить на действительную и мнимую, положительную и отрицательную части, как это делалось прежде. Поэтому предположим, что $a_{j,k} \geqslant 0$ для всех пар $(j,k) \in M$, где M — основное индексное множество $\{(j,k) \mid j \in J, k \in K_j\}$.

Дано, что сумма $\sum_{(j,k)\in M} a_{j,k}$ конечна, а именно, что

$$\sum_{(j,k)\in F} a_{j,k} \leqslant A$$

для всех конечных подмножеств $F \subseteq M$, и что A является наименьшей из таких верхних границ. Если j — некоторый элемент множества J, то каждая сумма вида $\sum_{k\in F_j} a_{j,k}$, где F_j — конечное подмножество множества K_j, ограничена сверху числом A. Следовательно, эти конечные суммы имеют наименьшую верхнюю границу $A_j \geqslant 0$ и по определению $\sum_{k\in K_j} a_{j,k} = A_j$.

Нам все еще надо доказать, что A является наименьшей верхней границей $\sum_{j\in G} A_j$ для всех конечных подмножеств $G \subseteq J$. Предположим, что G — конечное подмножество множества J, такое, что $\sum_{j\in G} A_j = A' > A$. Мы можем найти конечные подмножества $F_j \subseteq K_j$, такие, что $\sum_{k\in F_j} a_{j,k} > (A/A')A_j$ для каждого $j \in G$, для которого $A_j > 0$. Имеется как минимум одно такое j. Но тогда $\sum_{j\in G, k\in F_j} a_{j,k} > (A/A')\sum_{j\in G} A_j = A$, что противоречит тому факту, что $\sum_{(j,k)\in F} a_{j,k} \leqslant A$ для всех конечных подмножеств $F \subseteq M$. Следовательно, для всех конечных подмножеств $G \subseteq J$ справедливо $\sum_{j\in G} A_j \leqslant A$.

Наконец, пусть A' — некоторое действительное число, меньшее A. Наше доказательство будет завершено, если мы сможем найти конечное множество $G \subseteq J$, такое, что $\sum_{j\in G} A_j > A'$. Мы знаем, что существует конечное множество $F \subseteq M$, такое, что $\sum_{(j,k)\in F} a_{j,k} > A'$; пусть G — множество индексов j в этом F и пусть $F_j = \{k \mid (j,k) \in F\}$. Тогда $\sum_{j\in G} A_j \geqslant \sum_{j\in G}\sum_{k\in F_j} a_{j,k} = \sum_{(j,k)\in F} a_{j,k} > A'$. QED

Итак, теперь все законно. Все, что мы делали с бесконечными суммами, оправданно постольку, поскольку для любой конечной суммы абсолютных величин ее членов существует конечная граница. Так как “дважды бесконечная” сумма (2.58) давала нам два разных ответа при вычислении ее двумя разными способами, ее положительные члены $1+\frac{1}{2}+\frac{1}{3}+\cdots$ должны расходиться к ∞; в противном случае нами был бы получен один и тот же ответ, не зависящий от способа группирования ее членов.

Тогда почему я в последнее время только и слышу о “гармоничной сходимости”?

Упражнения

Разминка

1 Объясните, что означает следующая запись:

$$\sum_{k=4}^{0} q_k$$

2 Упростите выражение $x \cdot \big([x>0]-[x<0]\big)$.

3 Продемонстрируйте свое понимание $\sum$-обозначения, записав суммы

$$\sum_{0\leqslant k\leqslant 5} a_k \qquad \text{и} \qquad \sum_{0\leqslant k^2\leqslant 5} a_{k^2}$$

в развернутом виде. (Осторожно: вторая сумма — с подвохом.)

4 Где ошибка в следующем выводе?

$$\Big(\sum_{j=1}^{n} a_j\Big)\Big(\sum_{k=1}^{n} \frac{1}{a_k}\Big) = \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{a_j}{a_k} = \sum_{k=1}^{n}\sum_{k=1}^{n} \frac{a_k}{a_k} = \\ = \sum_{k=1}^{n} n = n^2 .$$

5 Выразите тройную сумму

$$\sum_{1\leqslant i<j<k\leqslant 4} a_{ijk}$$

в виде троекратной суммы (с тремя $\sum$):

а суммируя сначала по k, затем по j, затем по i;

б суммируя сначала по i, затем по j, затем по k.

Запишите также свои троекратные суммы в полном виде, без применения $\sum$-обозначений, с использованием скобок для того, чтобы показать, что́ именно суммируется вначале.

6 Выразите значение суммы $\sum_k [1\leqslant j\leqslant k\leqslant n]$ в виде функции от j и n.

Уделим внимание возрастающим степеням.

7 Пусть $\nabla f(x) = f(x) - f(x-1)$. Чему равно $\nabla(x^{\overline{m}})$?

8 Чему равно значение $0^{\underline{m}}$, где m — некоторое заданное целое число?

9 Как выглядит правило показателей для возрастающих факториальных степеней, аналогичное (2.52)? Воспользуйтесь им для определения $x^{\overline{-n}}$.

10 В тексте была выведена следующая формула для разности произведения:

$$\Delta(uv) \;=\; u\Delta v + \mathrm{E}v\,\Delta u\,.$$

Как может быть верной эта формула, если ее левая часть симметрична относительно u и v, а правая — нет?

Обязательные упражнения

11 Общее правило суммирования по частям (2.56) эквивалентно следующему:

$$\sum_{0\leqslant k<n}(a_{k+1}-a_k)b_k = a_n b_n - a_0 b_0 - \\ - \sum_{0\leqslant k<n} a_{k+1}(b_{k+1}-b_k) \quad \text{при } n\geqslant 0.$$

Докажите эту формулу непосредственно с применением дистрибутивного, ассоциативного и коммутативного законов.

12 Покажите, что функция $p(k)=k+(-1)^k c$ является перестановкой множества всех целых чисел, когда c — целое число.

13 Воспользуйтесь методом наборов для поиска суммы

$$\sum_{k=0}^{n}(-1)^k k^2$$

в аналитическом виде.

14 Вычислите сумму $\sum_{k=1}^{n} k2^k$, переписав ее в виде кратной суммы $\sum_{1\leqslant j\leqslant k\leqslant n} 2^k$.

15 Вычислите $⧉_n = \sum_{k=1}^{n} k^3$ методом 5 из текста главы: сначала запишите $⧉_n + \square_n = 2\sum_{1\leqslant j\leqslant k\leqslant n} jk$; затем примените (2.33).

16 Докажите, что $x^{\underline{m}}/(x-n)^{\underline{m}} = x^{\underline{n}}/(x-m)^{\underline{n}}$, если только ни один из знаменателей не равен нулю.

17 Покажите, что следующие формулы могут быть использованы для преобразования убывающих степеней в возрастающие и наоборот для всех целых m:

$$x^{\overline{m}} \;=\; (-1)^m(-x)^{\underline{m}} \;=\; (x+m-1)^{\underline{m}} \;=\; 1/(x-1)^{\underline{-m}};$$
$$x^{\underline{m}} \;=\; (-1)^m(-x)^{\overline{m}} \;=\; (x-m+1)^{\overline{m}} \;=\; 1/(x+1)^{\overline{-m}}.$$

(Степень $x^{\overline{-m}}$ определена в ответе к упр. 9.)

18 Пусть $\Re z$ и $\Im z$ — действительная и мнимая части комплексного числа z. Модуль $|z|$ равен $\sqrt{(\Re z)^2 + (\Im z)^2}$. Говорят, что сумма $\sum_{k \in K} a_k$ комплексных членов a_k сходится абсолютно, если сходятся абсолютно обе действительные суммы $\sum_{k \in K} \Re a_k$ и $\sum_{k \in K} \Im a_k$. Докажите, что $\sum_{k \in K} a_k$ сходится абсолютно тогда и только тогда, когда существует ограничивающая константа B, такая, что $\sum_{k \in F} |a_k| \leqslant B$ для всех конечных подмножеств $F \subseteq K$.

Домашние задания

19 Воспользуйтесь суммирующим множителем для решения рекуррентности

$$\begin{aligned} T_0 &= 5\,; \\ 2T_n &= nT_{n-1} + 3 \cdot n! \qquad \text{при } n > 0. \end{aligned}$$

20 Попробуйте, вычисляя сумму $\sum_{k=0}^{n} kH_k$ методом приведения, получить вместо этого значение суммы $\sum_{k=0}^{n} H_k$.

21 Вычислите суммы $S_n = \sum_{k=0}^{n}(-1)^{n-k}$, $T_n = \sum_{k=0}^{n}(-1)^{n-k}k$ и $U_n = \sum_{k=0}^{n}(-1)^{n-k}k^2$ методом приведения, считая, что $n \geqslant 0$.

Трудно доказать тождественность умершего около двух веков тому назад.

22 Докажите *тождество Лагранжа* (не прибегая к индукции):

$$\sum_{1 \leqslant j < k \leqslant n} (a_j b_k - a_k b_j)^2 = \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right) - \left(\sum_{k=1}^{n} a_k b_k\right)^2 .$$

Докажите тождество для более общего случая:

$$\sum_{1 \leqslant j < k \leqslant n} (a_j b_k - a_k b_j)(A_j B_k - A_k B_j)\,.$$

23 Вычислите сумму $\sum_{k=1}^{n}(2k+1)/k(k+1)$ двумя способами:

а заменяя $1/k(k+1)$ "элементарными дробями" $1/k - 1/(k+1)$;

б суммируя по частям.

24 Чему равна сумма $\sum_{0 \leqslant k < n} H_k/(k+1)(k+2)$? *Указание*: обобщите вывод (2.57).

Это обозначение введено Якоби (Jacobi) в 1829 году [192].

25 Обозначение $\prod_{k \in K} a_k$ означает произведение чисел a_k для всех $k \in K$. Предположим для простоты, что $a_k \neq 1$ лишь для конечного числа k, так что нам не надо определять бесконечные произведения. Каким законам, аналогичным дистрибутивному, ассоциативному и коммутативному законам для $\sum$-операции, удовлетворяет подобная $\prod$-операция?

26 Выразите двойное произведение $\prod_{1\leqslant j\leqslant k\leqslant n} a_j a_k$ при помощи манипуляций с $\prod$-обозначением через одинарное произведение $\prod_{k=1}^{n} a_k$. (Это упражнение дает нам произведение элементов верхней треугольной матрицы по аналогии с их суммой (2.33).)

27 Вычислите величину $\Delta(c^{\underline{x}})$ и используйте ее для получения значения $\sum_{k=1}^{n}(-2)^{\underline{k}}/k$.

28 Где следующий вывод свернул с пути истинного?

$$\begin{aligned}
1 &= \sum_{k\geqslant 1}\frac{1}{k(k+1)} = \sum_{k\geqslant 1}\left(\frac{k}{k+1}-\frac{k-1}{k}\right) = \\
&= \sum_{k\geqslant 1}\sum_{j\geqslant 1}\left(\frac{k}{j}[j=k+1]-\frac{j}{k}[j=k-1]\right) = \\
&= \sum_{j\geqslant 1}\sum_{k\geqslant 1}\left(\frac{k}{j}[j=k+1]-\frac{j}{k}[j=k-1]\right) = \\
&= \sum_{j\geqslant 1}\sum_{k\geqslant 1}\left(\frac{k}{j}[k=j-1]-\frac{j}{k}[k=j+1]\right) = \\
&= \sum_{j\geqslant 1}\left(\frac{j-1}{j}-\frac{j}{j+1}\right) = \sum_{j\geqslant 1}\frac{-1}{j(j+1)} = -1\,.
\end{aligned}$$

Контрольные работы

29 Вычислите сумму $\sum_{k=1}^{n}(-1)^k k/(4k^2-1)$.

30 Игроки в криббедж хорошо знают, что $15 = 7+8 = 4+5+6 = 1+2+3+4+5$. Найдите количество способов представления числа 1050 в виде суммы последовательных положительных целых чисел. (Тривиальное представление 1050 в виде самого себя считается одним из таких способов; следовательно, имеется четыре, а не три способа представления числа 15 в виде суммы последовательных положительных целых чисел. Кстати, для решения этой задачи знание правил игры в криббедж не требуется.)

31 Дзета-функция Римана $\zeta(k)$ определяется как бесконечная сумма

$$1+\frac{1}{2^k}+\frac{1}{3^k}+\cdots = \sum_{j\geqslant 1}\frac{1}{j^k}\,.$$

Докажите, что $\sum_{k\geqslant 2}\bigl(\zeta(k)-1\bigr) = 1$. Найдите, чему равна сумма $\sum_{k\geqslant 1}\bigl(\zeta(2k)-1\bigr)$.

32 Пусть $a \dot{-} b = \max(0, a-b)$. Докажите, что

$$\sum_{k\geqslant 0}\min(k, x \dot{-} k) \;=\; \sum_{k\geqslant 0}\bigl(x \dot{-} (2k+1)\bigr)$$

при любом действительном $x \geqslant 0$, и вычислите эту сумму в аналитическом виде.

Дополнительные задачи

33 Пусть $\bigwedge_{k\in K} a_k$ обозначает минимальное из чисел a_k (или их наибольшую нижнюю границу, если K бесконечно) в предположении, что каждое a_k — либо действительное число, либо $\pm\infty$. Какие законы справедливы для $\bigwedge$-операции, аналогичные тем, которые имеют место для $\sum$ и $\prod$ (см. упр. 25)?

Закон джунглей.

34 Докажите, что если сумма $\sum_{k\in K} a_k$ не определена в соответствии с определением (2.59), то она расслаивается в следующем смысле: если A^- и A^+ — любые заданные действительные числа, то можно найти последовательность конечных подмножеств $F_1 \subset F_2 \subset F_3 \subset \cdots$ множества K, такую, что

$$\sum_{k\in F_n} a_k \;\leqslant\; A^-, \text{ если } n \text{ нечетное,}$$

$$\sum_{k\in F_n} a_k \;\geqslant\; A^+, \text{ если } n \text{ четное.}$$

35 Докажите теорему Гольдбаха

$$1 \;=\; \frac{1}{3}+\frac{1}{7}+\frac{1}{8}+\frac{1}{15}+\frac{1}{24}+\frac{1}{26}+\frac{1}{31}+\frac{1}{35}+\cdots \;=\; \sum_{k\in P}\frac{1}{k-1},$$

где P — множество "совершенных степеней", рекурсивно определяемое как

$$P \;=\; \{\, m^n \mid m \geqslant 2, n \geqslant 2, m \notin P \,\}.$$

Совершенная власть развращает совершенно.

(Игра слов, основанная на том, что слово "power" переводится и как "степень", и как "власть". — Переводчик)

36 "Самоописывающая последовательность" Соломона Голомба (Solomon Golomb) $\langle f(1), f(2), f(3), \ldots \rangle$ — единственная неубывающая последовательность положительных целых чисел, основным свойством которой является то, что она содержит ровно $f(k)$ вхождений каждого k. Немного подумав, можно прийти к выводу, что эта последовательность должна начинаться следующим образом:

n	1	2	3	4	5	6	7	8	9	10	11	12
$f(n)$	1	2	2	3	3	4	4	4	5	5	5	6

Пусть $g(n)$ — наибольшее целое m, такое, что $f(m) = n$. Покажите, что

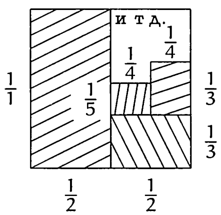
а $g(n) = \sum_{k=1}^{n} f(k)$;

б $g(g(n)) = \sum_{k=1}^{n} kf(k)$;

в $g(g(g(n))) = \frac{1}{2}ng(n)(g(n)+1) - \frac{1}{2}\sum_{k=1}^{n-1} g(k)(g(k)+1)$.

Исследовательская проблема

37 Можно ли уложить все прямоугольники размером $1/k \times 1/(k+1)$ при $k \geqslant 1$ в квадрат размером 1×1? (Вспомните, что сумма их площадей равна 1.)



3

Целочисленные функции

ЦЕЛЫЕ ЧИСЛА составляют костяк дискретной математики, и нам часто требуется преобразовывать дробные или произвольные действительные числа в целые. В этой главе наша цель — познакомиться с такого рода преобразованиями и приобрести навыки работы с ними, а также немного изучить их замечательные свойства.

3.1 Полы и потолки

Начнем с настила пола (floor) и перекрытия потолка (ceil), а именно — с функций "пол" (наибольшее целое) и "потолок" (наименьшее целое), которые определены для всех действительных чисел x следующим образом:

$$\begin{aligned} \lfloor x \rfloor &= \text{наибольшее целое, не превышающее } x\,; \\ \lceil x \rceil &= \text{наименьшее целое, не меньшее } x\,. \end{aligned} \tag{3.1}$$

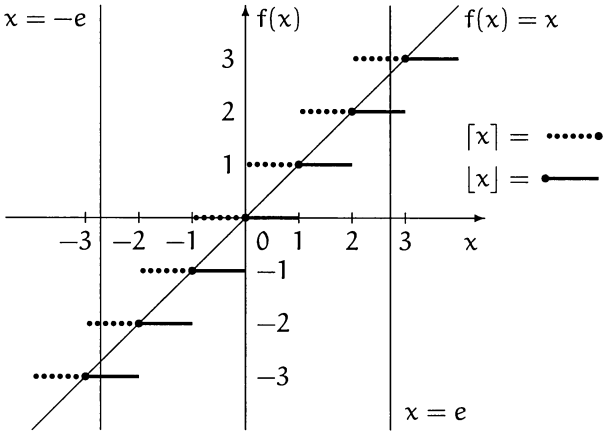
Эти обозначения (и названия) были введены Кеннетом Айверсоном (Kenneth E. Iverson) в начале 1960-х годов [191, с. 12]. Он обнаружил, что наборщики вполне могли бы справиться с этими символами, срезав верхушки и основания символов '[' и ']'. Его обозначения стали настолько распространенными, что сейчас они используются в технических статьях без каких-либо пояснений. До недавнего времени для наибольшего целого $\leqslant x$ часто использовалось обозначение '[x]', при этом эквивалентное обозначение для наименьшего целого отсутствовало. Некоторые авторы пытались использовать ']x[' (понятно, что из этого ничего не могло получиться).

)Бppp...(

Помимо разных вариантов обозначений, имеются разные варианты самих функций. Например, некоторые калькуляторы имеют функцию INT, определенную как $\lfloor x \rfloor$ при положительном

x и $\lceil x \rceil$ при отрицательном x. Вероятно, разработчики этих калькуляторов хотели, чтобы их функция INT удовлетворяла тождеству $\mathrm{INT}(-x) = -\mathrm{INT}(x)$. Но мы остаемся приверженцами наших пола и потолка, поскольку у них имеются свойства гораздо более привлекательные.

Один из способов поближе познакомиться с функциями пола и потолка — взглянуть на их графики, напоминающие ступеньки над и под прямой $f(x) = x$:



Из этого графика, например, видно, что

$$\lfloor e \rfloor = 2, \qquad \lfloor -e \rfloor = -3,$$
$$\lceil e \rceil = 3, \qquad \lceil -e \rceil = -2,$$

поскольку $e = 2.71828\ldots$.

Взглянув на эту иллюстрацию, можно заметить ряд фактов относительно полов и потолков. Начнем с того, что, поскольку функция пола лежит под диагональю $f(x) = x$, справедливо соотношение $\lfloor x \rfloor \leqslant x$; аналогично $\lceil x \rceil \geqslant x$. (Конечно, эти факты следуют непосредственно из определения функций.) В целых точках эти две функции совпадают:

$$\lfloor x \rfloor = x \quad \Longleftrightarrow \quad x \text{ целое} \quad \Longleftrightarrow \quad \lceil x \rceil = x.$$

(Здесь символ '$\Longleftrightarrow$' означает "тогда и только тогда".) Если же они не совпадают, то потолок ровно на 1 выше пола:

$$\lceil x \rceil - \lfloor x \rfloor = [x \text{ не целое}]. \tag{3.2}$$

Красивое применение обозначений, введенных Айверсоном!

Если сдвинуть диагональ вниз на 1, то она окажется полностью под функцией "пол", так что $x - 1 < \lfloor x \rfloor$; аналогично $x + 1 > \lceil x \rceil$. Объединив эти наблюдения, получаем

$$x - 1 < \lfloor x \rfloor \leqslant x \leqslant \lceil x \rceil < x + 1. \tag{3.3}$$

Наконец, эти функции являются отражениями друг друга относительно обеих осей:

$$\lfloor -x \rfloor = -\lceil x \rceil ; \qquad \lceil -x \rceil = -\lfloor x \rfloor . \tag{3.4}$$

Таким образом, каждая функция легко выражается через другую, что помогает объяснить, почему функция "потолок" раньше не имела собственного обозначения. Тем не менее потолки встречаются так часто, что заслуживают отдельного обозначения, так же как отдельные обозначения были введены для возрастающих и убывающих степеней. Математики издавна имеют синус и косинус, тангенс и котангенс, секанс и косеканс, максимум и минимум, а теперь у нас есть пол и потолок.

На следующей неделе обещали подвезти стены...

Для настоящего доказательства свойств этих функций (а не просто разглядывания картинок) особенно полезны следующие правила:

$$\begin{aligned}
\lfloor x \rfloor = n \quad &\iff \quad n \leqslant x < n+1 , && \text{(а)}\\
\lfloor x \rfloor = n \quad &\iff \quad x-1 < n \leqslant x , && \text{(б)}\\
\lceil x \rceil = n \quad &\iff \quad n-1 < x \leqslant n , && \text{(в)}\\
\lceil x \rceil = n \quad &\iff \quad x \leqslant n < x+1 . && \text{(г)}
\end{aligned} \tag{3.5}$$

(Во всех случаях считается, что n — целое, а x — действительное.) Правила (а) и (в) следуют непосредственно из определения (3.1); правила (б) и (г) — те же самые, просто неравенства преобразованы так, что в средине оказывается n.

Целочисленное слагаемое можно свободно вносить и выносить в/за скобки пола/потолка:

На радость ослику Иа...
— Переводчик

$$\lfloor x+n \rfloor = \lfloor x \rfloor + n , \qquad n \text{ — целое.} \tag{3.6}$$

(Правило (3.5, а) гласит, что это утверждение эквивалентно неравенствам $\lfloor x \rfloor + n \leqslant x+n < \lfloor x \rfloor + n + 1$.) Но аналоги этой операции типа вынесения за скобки постоянного множителя в общем случае недопустимы. Например, $\lfloor nx \rfloor \neq n\lfloor x \rfloor$ при $n = 2$ и $x = 1/2$. Это означает, что скобки пола и потолка сравнительно негибки. Обычно мы счастливы, если удается от них избавиться или что-то доказать при их наличии.

Оказывается, что имеется масса ситуаций, когда скобки пола и потолка избыточны, так что их можно удалять или вставлять, как вам угодно. Например, любое неравенство между действительным и целым числами эквивалентно неравенству с полом или потолком между целыми числами:

$$\begin{aligned} x < n &\iff \lfloor x \rfloor < n, &\text{(а)}\\ n < x &\iff n < \lceil x \rceil, &\text{(б)}\\ x \leqslant n &\iff \lceil x \rceil \leqslant n, &\text{(в)}\\ n \leqslant x &\iff n \leqslant \lfloor x \rfloor. &\text{(г)} \end{aligned} \qquad (3.7)$$

Эти правила легко доказать. Например, если $x < n$, то, конечно же, $\lfloor x \rfloor < n$, поскольку $\lfloor x \rfloor \leqslant x$. И наоборот, если $\lfloor x \rfloor < n$, то должно выполняться $x < n$, поскольку $x < \lfloor x \rfloor + 1$ и $\lfloor x \rfloor + 1 \leqslant n$.

Было бы хорошо, если бы эти четыре правила в (3.7) так же легко запоминались, как и доказывались. Каждое неравенство без пола или потолка соответствует такому же неравенству с полом или с потолком; но нужно дважды подумать, прежде чем решить, какое из них подходит.

Разность между x и $\lfloor x \rfloor$ называется *дробной частью* x и встречается в приложениях достаточно часто, чтобы заслужить собственное обозначение:

$$\{x\} = x - \lfloor x \rfloor. \qquad (3.8)$$

Лучше все же не записывать дробную часть как $\{x\}$ в случаях, когда есть шанс спутать ее с множеством, состоящим из единственного элемента x.

Иногда $\lfloor x \rfloor$ называют *целой частью* x, поскольку $x = \lfloor x \rfloor + \{x\}$. Если действительное число x можно записать в виде $x = n + \theta$, где n — целое число, а $0 \leqslant \theta < 1$, то согласно (3.5, а) можно заключить, что $n = \lfloor x \rfloor$ и $\theta = \{x\}$.

Равенство (3.6) не выполняется, если n — произвольное действительное число. Но в общем случае для $\lfloor x + y \rfloor$ имеются две возможности. Если записать $x = \lfloor x \rfloor + \{x\}$ и $y = \lfloor y \rfloor + \{y\}$, то получим $\lfloor x+y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor + \lfloor \{x\}+\{y\} \rfloor$. А поскольку $0 \leqslant \{x\}+\{y\} < 2$, оказывается, что иногда $\lfloor x + y \rfloor$ равно $\lfloor x \rfloor + \lfloor y \rfloor$, а в остальных случаях это $\lfloor x \rfloor + \lfloor y \rfloor + 1$.

Второй случай реализуется тогда и только тогда, когда выполняется "перенос" за десятичную точку при сложении дробных частей $\{x\}$ и $\{y\}$.

3.2 Применения пола и потолка

Итак, базовый инструментарий для работы с полами и потолками у нас есть. Опробуем их на деле, начав с простой задачки: что такое $\lceil \lg 35 \rceil$? (Следуя предложению Эдварда М. Рейнгольда (Edward M. Reingold), мы используем 'lg' для обозначения логарифма по основанию 2.) Ну а поскольку $2^5 < 35 \leqslant 2^6$, взятие логарифма дает нам $5 < \lg 35 \leqslant 6$, откуда согласно (3.5, в) $\lceil \lg 35 \rceil = 6$.

Заметим, что число 35 при записи в двоичной системе счисления состоит из шести битов: $35 = (100011)_2$. Всегда ли верно, что $\lceil \lg n \rceil$ представляет собой длину n в двоичном виде? Не всегда. Нам нужны те же 6 бит и для записи $32 = (100000)_2$, так что $\lceil \lg n \rceil$ — неверный ответ на поставленный вопрос. (Он неверен, только когда n представляет собой степень 2, но это означает, что он неверен в бесконечном числе случаев.) Правильный ответ можно найти, заметив, что для записи числа n, такого, что $2^{m-1} \leqslant n < 2^m$, требуется m бит; (3.5, a) говорит нам, что $m-1 = \lfloor \lg n \rfloor$, так что $m = \lfloor \lg n \rfloor + 1$. То есть для записи любого $n > 0$ требуется $\lfloor \lg n \rfloor + 1$ бит. Аналогичные рассуждения дают другой вариант ответа $\lceil \lg(n+1) \rceil$; эта формула справедлива и для $n = 0$, если считать, что для записи $n = 0$ в двоичном виде требуется 0 бит.

Давайте теперь рассмотрим выражения с несколькими полами или потолками. Что такое $\lceil \lfloor x \rfloor \rceil$? Это просто: поскольку $\lfloor x \rfloor$ — целое число, $\lceil \lfloor x \rfloor \rceil$ — это просто $\lfloor x \rfloor$. Этому же равно любое другое выражение, в котором наиболее глубоко вложенный $\lfloor x \rfloor$ окружен любым количеством полов и потолков.

Вот более сложная задача: докажите или опровергните утверждение

$$\left\lfloor \sqrt{\lfloor x \rfloor} \right\rfloor = \lfloor \sqrt{x} \rfloor \qquad \text{для действительного } x \geqslant 0. \tag{3.9}$$

Очевидно, что равенство достигается, когда x — целое число, так как $x = \lfloor x \rfloor$. Равенство выполняется и для частных случаев $\pi = 3.14159\ldots$, $e = 2.71828\ldots$ и $\phi = (1+\sqrt{5})/2 = 1.61803\ldots$, поскольку во всех этих случаях мы получаем $1 = 1$. Безуспешные попытки найти контрпример наводят на мысль, что равенство справедливо в общем случае, так что давайте попытаемся его доказать.

Ну конечно, первые действительные числа, приходящие на ум, — это π, e и ϕ, разве не так?

Скептицизм полезен для здоровья только до определенной степени. Скептическое отношение к намерениям и планам (особенно своим), вероятно, избавит вас от волнений и обеспечит спокойную жизнь. Но проявление слишком скептического отношения наверняка заставит вас все время быть прикованным к работе, вместо того чтобы позволить себе выкроить время для физических упражнений

Кстати, когда мы сталкиваемся с задачей "доказать или опровергнуть", то обычно начинаем с попыток опровержения при помощи контрпримера по двум причинам: опровержение потенциально проще (достаточно одного опровергающего контрпримера), а его поиски будят спящие творческие силы. Даже если утверждение истинно, поиски контрпримеров зачастую приводят к доказательству — как только скептик понимает, почему контрпример невозможен. И вообще, скептицизм полезен для здоровья.

Если бы мы пытались доказать, что $\left\lfloor \sqrt{\lfloor x \rfloor} \right\rfloor = \lfloor \sqrt{x} \rfloor$ при помощи анализа, то для этого можно было бы начать с разбиения x на целую и дробную части $\lfloor x \rfloor + \{x\} = n + \theta$ с последующим разложением квадратного корня, применяя биномиальную теорему: $(n+\theta)^{1/2} = n^{1/2} + n^{-1/2}\theta/2 - n^{-3/2}\theta^2/8 + \cdots$. Но это слишком громоздкий, непрактичный подход.

Гораздо проще воспользоваться имеющимся инструментарием. Вот возможная стратегия: разобрать наружный пол и квадратный корень в $\lfloor\sqrt{\lfloor x\rfloor}\rfloor$, затем удалить внутренний пол, после чего собрать разобранное, чтобы получить $\lfloor\sqrt{x}\rfloor$. Положим $m = \lfloor\sqrt{\lfloor x\rfloor}\rfloor$ и воспользуемся правилом (3.5, a), получая $m \leqslant \sqrt{\lfloor x\rfloor} < m+1$. Тем самым мы избавляемся от скобок наружного пола, ничего при этом не теряя. Применим возведение в квадрат, поскольку все три выражения неотрицательны. При этом мы получим $m^2 \leqslant \lfloor x\rfloor < (m+1)^2$. Так мы избавляемся от квадратного корня. Затем мы удаляем пол, воспользовавшись правилом (3.7, г) для левого неравенства и (3.7, а) для правого: $m^2 \leqslant x < (m+1)^2$. Теперь восстановить проделанное не составляет особого труда: извлечь квадратные корни, получив $m \leqslant \sqrt{x} < m+1$, и воспользоваться правилом (3.5, а), чтобы получить $m = \lfloor\sqrt{x}\rfloor$. Таким образом, $\lfloor\sqrt{\lfloor x\rfloor}\rfloor = m = \lfloor\sqrt{x}\rfloor$, т.е. утверждение оказалось истинным. Аналогично доказывается, что

и восстановления сил. Чрезмерный скептицизм — это прямая дорога к состоянию "конченного человека", когда вы становитесь настолько озабочены соблюдением строгости и законченности, что уже никогда не можете ничего довести до конца.
— Скептик

$$\lceil\sqrt{\lceil x\rceil}\rceil = \lceil\sqrt{x}\rceil \qquad \text{для действительного } x \geqslant 0.$$

Только что найденное доказательство не так сильно полагается на свойства квадратных корней. Более пристальный взгляд показывает, что можно обобщить лежащие в основе доказательства идеи и доказать более общее утверждение. Пусть $f(x)$ — некоторая непрерывная монотонно возрастающая функция, обладающая тем свойством, что

$$f(x) = \text{целое число} \qquad \Longrightarrow \qquad x = \text{целое число}.$$

(Символ '$\Longrightarrow$' означает "отсюда следует".) Тогда

$$\lfloor f(x)\rfloor = \lfloor f(\lfloor x\rfloor)\rfloor \qquad \text{и} \qquad \lceil f(x)\rceil = \lceil f(\lceil x\rceil)\rceil \tag{3.10}$$

Это наблюдение было сделано Р.Д. Мак-Элисом (R. J. McEliece), когда он заканчивал университет.

всякий раз, когда определены функции $f(x)$, $f(\lfloor x\rfloor)$ и $f(\lceil x\rceil)$. Поскольку до этого мы занимались полами, давайте докажем это общее свойство для потолков, тем более что для полов доказательство почти такое же. Если $x = \lceil x\rceil$, то доказывать нечего. В противном случае $x < \lceil x\rceil$, а $f(x) < f(\lceil x\rceil)$, поскольку f — возрастающая функция. Следовательно, $\lceil f(x)\rceil \leqslant \lceil f(\lceil x\rceil)\rceil$, так как $\lceil\,\rceil$ — функция неубывающая. Если $\lceil f(x)\rceil < \lceil f(\lceil x\rceil)\rceil$, то, поскольку функция f непрерывна, должно существовать число y, такое, что $x \leqslant y < \lceil x\rceil$ и $f(y) = \lceil f(x)\rceil$. Это y является целым числом в силу специального свойства f. Но между $\lfloor x\rfloor$ и $\lceil x\rceil$ не может быть никакого целого числа. Это противоречие означает, что должно выполняться $\lceil f(x)\rceil = \lceil f(\lceil x\rceil)\rceil$.

Важный частный случай этой теоремы заслуживает того, чтобы быть явно упомянутым:

$$\left\lfloor\frac{x+m}{n}\right\rfloor = \left\lfloor\frac{\lfloor x\rfloor+m}{n}\right\rfloor \quad\text{и}\quad \left\lceil\frac{x+m}{n}\right\rceil = \left\lceil\frac{\lceil x\rceil+m}{n}\right\rceil, \tag{3.11}$$

если m и n — целые числа, а знаменатель n положителен. Например, пусть $m = 0$; тогда $\lfloor\lfloor\lfloor x/10\rfloor/10\rfloor/10\rfloor = \lfloor x/1000\rfloor$. Троекратное деление на 10 и отбрасывание цифр остатка — это то же, что и деление на 1000 с последующим отбрасыванием всего остатка.

Попробуем теперь доказать или опровергнуть следующее утверждение:

$$\lceil\sqrt{\lfloor x\rfloor}\,\rceil \stackrel{?}{=} \lceil\sqrt{x}\,\rceil \qquad \text{при действительном } x \geqslant 0.$$

Оно справедливо при $x = \pi$ и $x = e$, но не выполняется при $x = \phi$. Этого достаточно, чтобы утверждать, что в общем случае оно неверно.

ОТСТУПЛЕНИЕ Перед тем как идти дальше, давайте на минуту прервемся и рассмотрим различные уровни задач, которые могут предлагаться в математических книгах.

Уровень 1. Для явно указанных объекта x и свойства $P(x)$ надо *доказать* истинность $P(x)$. Например, "докажите, что $\lfloor\pi\rfloor = 3$". Задача требует доказательства некоторого предполагаемого факта.

Уровень 2. Для явно указанных множества X и свойства $P(x)$ надо доказать истинность $P(x)$ *для всех* $x \in X$. Например, "докажите, что $\lfloor x\rfloor \leqslant x$ для всех действительных x". Задача вновь состоит в поиске доказательства, которое в этот раз имеет общий характер. Мы работаем уже не с арифметикой, а с алгеброй.

Уровень 3. Для явно указанных множества X и свойства $P(x)$ надо доказать *или опровергнуть* истинность $P(x)$ для всех $x \in X$. Например, "докажите или опровергните, что $\lceil\sqrt{\lfloor x\rfloor}\,\rceil = \lceil\sqrt{x}\,\rceil$ для всех действительных $x \geqslant 0$". Здесь в задачу добавляется дополнительный уровень неопределенности — результат может оказаться как тем, так и другим. Это ближе к реальным ситуациям, с которыми постоянно сталкивается математик: утверждения в книгах обычно истинны, но в реальной исследовательской работе к новому надо относиться непредвзято. Если утверждение ложно, задача состоит в том, чтобы найти контрпример. Если утверждение истинно, следует доказать его, как и в случае задачи второго уровня.

В других моих книгах "доказать или опровергнуть" в 99.44% случаев означает то же самое, что и "доказать". К данной книге это утверждение не относится.

Уровень 4. Для явно указанных множества X и свойства $P(x)$ надо найти *необходимое и достаточное условие* $Q(x)$ того, что $P(x)$ истинно. Например, "найдите необходимое и достаточное условие того, что $\lfloor x \rfloor \geqslant \lceil x \rceil$". Задача состоит в том, чтобы найти Q, такое, что $P(x) \iff Q(x)$. Конечно, всегда имеется тривиальный ответ $Q(x) = P(x)$, но задача подразумевает поиск настолько простого условия, насколько это возможно. Для поиска работающего простого условия необходимо творчество. (Например, в нашем случае ответ — "$\lfloor x \rfloor \geqslant \lceil x \rceil \iff x$ является целым числом".) Дополнительный элемент исследования, необходимый для поиска $Q(x)$, делает эту разновидность задач более сложной, но и более типичной для математика, работающего в "реальной жизни". Ну и, конечно же, должно быть доказано, что $P(x)$ истинно тогда и только тогда, когда истинно $Q(x)$.

Но не проще.
— А. Эйнштейн (A. Einstein)

Уровень 5. Для явно указанного множества X найти *интересное общее свойство* $P(x)$ его элементов. Это область чистой науки, где, по мнению студентов, царит полный хаос. Это настоящая математика, которую авторы учебников крайне редко допускают в свои книги. КОНЕЦ ОТСТУПЛЕНИЯ

Давайте вернемся к нашему вопросу и переведем его с уровня 3 на уровень 4: какое условие является необходимым и достаточным для того, чтобы $\lceil \sqrt{\lfloor x \rfloor} \rceil = \lceil \sqrt{x} \rceil$? Мы уже выяснили, что равенство выполняется при $x = 3.142$, но не при $x = 1.618$; дальнейшие эксперименты показывают нам, что оно не выполняется и при x, лежащем между 9 и 10. Более того, можно обнаружить, что условие не выполняется при $m^2 < x < m^2 + 1$, поскольку при этом в левой части мы получаем m, а в правой — $m+1$. Во всех прочих случаях, где определен $\sqrt{x}$, а именно при $x = 0$ или $m^2 + 1 \leqslant x \leqslant (m+1)^2$, мы получаем равенство. Таким образом, для выполнения равенства необходимым и достаточным является следующее условие: либо x — целое число, либо $\sqrt{\lfloor x \rfloor}$ — не целое число.

Для следующей задачи рассмотрим сначала новое удобное обозначение, предложенное Ч. Э. Р. Хоаром (C. A. R. Hoare) и Лайлом Рэмшоу (Lyle Ramshaw), для интервалов действительной прямой: $[\alpha \,..\, \beta]$ означает множество действительных чисел x, таких, что $\alpha \leqslant x \leqslant \beta$. Это множество называется *замкнутым интервалом*, поскольку он включает обе конечные точки α и β. Интервал $(\alpha \,..\, \beta)$, не включающий ни одной из конечных точек, состоит из всех x, таких, что $\alpha < x < \beta$; такой интервал называется *открытым интервалом*. Интервалы $[\alpha \,..\, \beta)$ и $(\alpha \,..\, \beta]$, содержащие только по одной конечной точке, определяются аналогично и называются *полуоткрытыми*.

Пессимисты называют их полузамкнутыми.

Сколько целых чисел содержится в таких интервалах? Полуоткрытые интервалы проще, так что начнем с них. Фактически полуоткрытые интервалы почти всегда удобнее открытых или закрытых. Например, они аддитивны — при объединении полуоткрытых интервалов $[\alpha \mathinner{.\,.} \beta)$ и $[\beta \mathinner{.\,.} \gamma)$ получается полуоткрытый интервал $[\alpha \mathinner{.\,.} \gamma)$. В случае открытых интервалов это не получается, так как точка β оказывается исключенной; для замкнутых интервалов проблема иная — точка β оказывается включенной дважды.

Вернемся к нашей задаче. Если α и β — целые числа, ответ прост: интервал $[\alpha \mathinner{.\,.} \beta)$ содержит $\beta - \alpha$ целых чисел α, $\alpha + 1$, $\ldots$, $\beta - 1$ при условии, что $\alpha \leqslant \beta$. Аналогично интервал $(\alpha \mathinner{.\,.} \beta]$ содержит $\beta - \alpha$ целых чисел при том же условии. Но наша задача более трудная, так как по условию α и β — произвольные действительные числа. Но эту задачу можно свести к более простой, поскольку для целых n согласно (3.7)

$$\begin{aligned} \alpha \leqslant n < \beta \qquad &\iff \qquad \lceil \alpha \rceil \leqslant n < \lceil \beta \rceil, \\ \alpha < n \leqslant \beta \qquad &\iff \qquad \lfloor \alpha \rfloor < n \leqslant \lfloor \beta \rfloor. \end{aligned}$$

Интервалы справа имеют целые конечные точки и содержат столько же целых чисел, что и интервалы слева с действительными конечными точками. Поэтому интервал $[\alpha \mathinner{.\,.} \beta)$ содержит ровно $\lceil \beta \rceil - \lceil \alpha \rceil$ целых чисел, а $(\alpha \mathinner{.\,.} \beta]$ — $\lfloor \beta \rfloor - \lfloor \alpha \rfloor$. Это один из случаев, когда нам надо не избавляться от скобок пола или потолка, а, напротив, добавить их.

О мнемонических правилах: знаете, как запомнить первые знаки числа π? Они соответствуют количеству букв в словах стишка "это я знаю и помню прекрасно, пи многие знаки мне лишни, напрасны".

Кстати, имеется мнемоническое правило для запоминания, когда следует применять полы, а когда — потолки: полуоткрытые интервалы, которые включают не правую, а левую конечную точку (такие, как $0 \leqslant \theta < 1$), используются несколько чаще тех, которые включают правую точку, а не левую, так же как полы используются чаще потолков. Но по закону Мэрфи правило противоположно ожидаемому — потолки соответствуют интервалам $[\alpha \mathinner{.\,.} \beta)$, а полы — интервалам $(\alpha \mathinner{.\,.} \beta]$.

Аналогичный анализ показывает, что закрытый интервал $[\alpha \mathinner{.\,.} \beta]$ содержит ровно $\lfloor \beta \rfloor - \lceil \alpha \rceil + 1$ целых чисел, а открытый интервал $(\alpha \mathinner{.\,.} \beta)$ — ровно $\lceil \beta \rceil - \lfloor \alpha \rfloor - 1$, но в последнем случае мы налагаем дополнительное ограничение $\alpha \neq \beta$, чтобы эта формула не смущала нас заявлением о том, что пустой интервал $(\alpha \mathinner{.\,.} \alpha)$ содержит -1 целых чисел. Подытожим найденные факты:

$$\begin{array}{lccr}
\text{интервал} & \text{количество целых чисел} & \text{ограничение} & \\
[\alpha\,..\,\beta] & \lfloor\beta\rfloor-\lceil\alpha\rceil+1 & \alpha\leqslant\beta\,, & \\
[\alpha\,..\,\beta) & \lceil\beta\rceil-\lceil\alpha\rceil & \alpha\leqslant\beta\,, & (3.12)\\
(\alpha\,..\,\beta] & \lfloor\beta\rfloor-\lfloor\alpha\rfloor & \alpha\leqslant\beta\,, & \\
(\alpha\,..\,\beta) & \lceil\beta\rceil-\lfloor\alpha\rfloor-1 & \alpha<\beta\,. &
\end{array}$$

А вот задача, от которой мы не сможем отказаться. В "Клубе конкретной математики" есть казино (только для покупателей этой книги), в котором имеется рулетка с колесом на тысячу гнезд, пронумерованных от 1 до 1000. Если при вращении колеса выпадает номер n, который делится на пол своего кубического корня, т.е. если

$$\lfloor\sqrt[3]{n}\rfloor \setminus n\,,$$

то это номер выигрышный, и казино платит нам 5 долларов; в противном случае это проигрышный номер, и мы платим 1 доллар казино. (Обозначение '$a\backslash b$', которое читается как "a делит b", означает, что b в точности кратно a; это отношение детально рассматривается в главе 4.) Можно ли "сделать деньги", играя в казино?

Опрос в аудитории дал следующие результаты: 28 студентов отказались играть, 13 захотели рискнуть, остальные решили не отвечать на вопрос.

Можно вычислить средний выигрыш — т.е. сумму, которую мы выигрываем (или теряем) за одну игру, — подсчитав сперва количество W выигрышных номеров и $L = 1000 - W$ проигрышных. Если каждый номер выпадет по одному разу в серии из 1000 игр, мы выиграем $5W$ долларов и проиграем L долларов, так что средний выигрыш составляет

$$\frac{5W-L}{1000} = \frac{5W-(1000-W)}{1000} = \frac{6W-1000}{1000}\,.$$

Если имеется не менее 167 выигрышных номеров, мы получаем прибыль, в противном случае выигрывает казино.

Но как же нам подсчитать количество выигрышных номеров среди 1000? Несложно обнаружить закономерность. Номера от 1 до $2^3-1=7$ выигрышные, так как для каждого из них $\lfloor\sqrt[3]{n}\rfloor=1$. Среди номеров от $2^3=8$ до $3^3-1=26$ выигрышными являются только четные номера. Среди номеров от $3^3=27$ до $4^3-1=63$ — только те, которые делятся на 3, и т.д.

Систематический анализ задачи возможен при помощи методов суммирования из главы 2 с привлечением записи Айверсона для логических выражений:

$$\begin{aligned}
W &= \sum_{n=1}^{1000} [n\text{— выигрышный номер}] = \\
&= \sum_{1\leqslant n\leqslant 1000} \bigl[\lfloor\sqrt[3]{n}\rfloor \backslash n\bigr] = \sum_{k,n} \bigl[k=\lfloor\sqrt[3]{n}\rfloor\bigr][k\backslash n][1\leqslant n\leqslant 1000] = \\
&= \sum_{k,m,n} \bigl[k^3\leqslant n<(k+1)^3\bigr][n=km][1\leqslant n\leqslant 1000] = \\
&= 1+\sum_{k,m} \bigl[k^3\leqslant km<(k+1)^3\bigr][1\leqslant k<10] = \\
&= 1+\sum_{k,m} \bigl[m\in \bigl[k^2\,..\,(k+1)^3/k\bigr)\bigr][1\leqslant k<10] = \\
&= 1+\sum_{1\leqslant k<10} \bigl(\lceil k^2+3k+3+1/k\rceil - \lceil k^2\rceil\bigr) = \\
&= 1+\sum_{1\leqslant k<10} (3k+4) = 1+\frac{7+31}{2}\cdot 9 = 172.
\end{aligned}$$

Этот вывод заслуживает внимательного изучения. Обратите внимание, что в шестой строке использована наша формула (3.12) для количества целых чисел в полуоткрытом интервале. Единственная "сложность" состоит в рассмотрении при переходе от третьей строки к четвертой $n = 1000$ как отдельного случая. (Неравенство $k^3 \leqslant n < (k+1)^3$ нелегко объединить с неравенством $1 \leqslant n \leqslant 1000$ при $k = 10$.) В общем случае граничные условия — одно из наиболее критичных мест при работе с суммами.

Вот именно!

В последней строке выясняется, что $W = 172$; следовательно, наш средний выигрыш в одной игре равен $(6\cdot 172-1000)/1000$ долларам, т.е. 3.2 цента. Можно ожидать, что, сделав 100 ставок по доллару, вы станете богаче примерно на 3.20 доллара (если, конечно, казино не сделает так, что одни числа будут более равны, чем другие).

Где, вы говорите, находится это казино?

Задача о казино, которую мы только что так благополучно решили, представляет собой приукрашенную версию сухого вопроса "сколько целых чисел n, где $1 \leqslant n \leqslant 1000$, удовлетворяют отношению $\lfloor\sqrt[3]{n}\rfloor \backslash n$?" Математически это два одинаковых вопроса. Однако иногда сто́ит приукрасить задачу. Мы воспользовались более богатым словарем ("выигрышные номера" и "проигрышные номера"), что помогло нам лучше понять происходящее.

Обобщим задачу. Предположим, мы изменили 1000 на 1000000 или на еще большее число N (считаем, что у казино большие связи, и достать колесо побольше для него не проблема). Сколько выигрышных номеров имеется теперь?

В этом случае вполне приемлемы те же рассуждения, что и ранее, но надо быть поосторожнее с самым большим значением k, которое мы для удобства обозначим как K:

$$K = \lfloor \sqrt[3]{N} \rfloor .$$

(Ранее K было равно 10.) Общее количество выигрышных номеров для произвольного N становится равным

$$\begin{aligned} W &= \sum_{1 \leqslant k < K} (3k+4) + \sum_m [K^3 \leqslant Km \leqslant N] = \\ &= \tfrac{1}{2}(7+3K+1)(K-1) + \sum_m \bigl[m \in [K^2 \,..\, N/K]\bigr] = \\ &= \tfrac{3}{2}K^2 + \tfrac{5}{2}K - 4 + \sum_m \bigl[m \in [K^2 \,..\, N/K]\bigr] . \end{aligned}$$

Мы знаем, что оставшаяся сумма равна $\lfloor N/K \rfloor - \lceil K^2 \rceil + 1 = \lfloor N/K \rfloor - K^2 + 1$; следовательно, формула

$$W = \lfloor N/K \rfloor + \tfrac{1}{2}K^2 + \tfrac{5}{2}K - 3 , \qquad K = \lfloor \sqrt[3]{N} \rfloor \tag{3.13}$$

дает общий ответ для колеса с N гнездами.

Первые два члена этой формулы приближенно равны $N^{2/3} + \frac{1}{2}N^{2/3} = \frac{3}{2}N^{2/3}$, а остальные члены при большом N оказываются гораздо меньшими. В главе 9 мы научимся выводить выражения наподобие

$$W = \tfrac{3}{2}N^{2/3} + O(N^{1/3}) ,$$

где $O(N^{1/3})$ означает величину, не превосходящую некоторой константы, умноженной на $N^{1/3}$. Какой бы ни была константа, мы знаем, что она не зависит от N; так что при большом N вклад O-члена в W будет мал по сравнению с $\frac{3}{2}N^{2/3}$. Например, в следующей таблице показано, насколько значения $\frac{3}{2}N^{2/3}$ близки к значениям W:

N	$\frac{3}{2}N^{2/3}$	W	Погрешность, %
1 000	150.0	172	12.791
10 000	696.2	746	6.670
100 000	3231.7	3343	3.331
1 000 000	15000.0	15247	1.620
10 000 000	69623.8	70158	0.761
100 000 000	323165.2	324322	0.357
1 000 000 000	1500000.0	1502497	0.166

Вполне нормальное приближение.

Приближенные формулы полезны постольку, поскольку они проще формул с полами и потолками. Однако зачастую важна абсолютная точность, в особенности при малых значениях N, с которыми обычно приходится иметь дело на практике. Например, владелец казино может опрометчиво предположить, что при $N = 1000$ имеется только $\frac{3}{2}N^{2/3} = 150$ выигрышных номеров (в этом случае доход казино составил бы 10 центов).

Завершим раздел рассмотрением так называемых спектров. Определим *спектр* действительного числа α как бесконечное мультимножество целых чисел

$$\mathrm{Spec}(\alpha) = \{\lfloor\alpha\rfloor, \lfloor 2\alpha\rfloor, \lfloor 3\alpha\rfloor, \ldots\}.$$

(Мультимножество — это то же самое, что и множество, но оно может содержать повторяющиеся элементы.) Например, спектр $1/2$ начинается следующим образом: $\{0, 1, 1, 2, 2, 3, 3, \ldots\}$.

Легко доказать, что двух одинаковых спектров не существует, т.е. из $\alpha \neq \beta$ следует $\mathrm{Spec}(\alpha) \neq \mathrm{Spec}(\beta)$. Предположим без потери общности, что $\alpha < \beta$. Существует такое положительное целое число m, что $m(\beta - \alpha) \geqslant 1$. (В действительности таковым является любое целое $m \geqslant \lceil 1/(\beta - \alpha)\rceil$; но не будем постоянно хвастать нашим знанием полов и потолков.) Следовательно, $m\beta - m\alpha \geqslant 1$ и $\lfloor m\beta\rfloor > \lfloor m\alpha\rfloor$. Таким образом, $\mathrm{Spec}(\beta)$ содержит менее m элементов $\leqslant \lfloor m\alpha\rfloor$, в то время как $\mathrm{Spec}(\alpha)$ содержит их как минимум m.

...без нахождения общности...

Спектры обладают рядом красивых свойств. Рассмотрим, например, мультимножества

$$\begin{aligned}&\mathrm{Spec}(\sqrt{2}\,) = \\ &\quad = \{1, 2, 4, 5, 7, 8, 9, 11, 12, 14, 15, 16, 18, 19, 21, 22, 24, \ldots\}, \\ &\mathrm{Spec}(2 + \sqrt{2}\,) = \\ &\quad = \{3, 6, 10, 13, 17, 20, 23, 27, 30, 34, 37, 40, 44, 47, 51, \ldots\}.\end{aligned}$$

“Если x есть некоторое иррациональное число, меньшее единицы, то можно взять один из двух рядов величин m/x, $m/(1 - x)$, где m — целое число; каждое число, принадлежащее к тому или иному ряду, и только оно одно, будет заключено между любыми двумя заданными последовательными числами”.

— Рэлей (Rayleigh) [304]

Спектр $\mathrm{Spec}(\sqrt{2}\,)$ легко вычислить на калькуляторе, а n-й элемент спектра $\mathrm{Spec}(2 + \sqrt{2}\,)$ согласно (3.6) ровно на $2n$ больше n-го элемента спектра $\mathrm{Spec}(\sqrt{2}\,)$. Более внимательное рассмотрение показывает, что эти два спектра связаны гораздо более удивительным образом: похоже, что любое число, которого нет в одном спектре, имеется в другом, но никакое число не содержится одновременно в обоих спектрах! И это на самом деле так: положительные целые числа представляют собой объединение непересекающихся множеств $\mathrm{Spec}(\sqrt{2}\,)$ и $\mathrm{Spec}(2 + \sqrt{2}\,)$. Говорят, что эти спектры образуют *разбиение* (*partition*) положительных целых чисел.

Для доказательства этого факта подсчитаем, сколько элементов в множестве $\mathrm{Spec}(\sqrt{2}\,)$ не превышают n и сколько таких же элементов в множестве $\mathrm{Spec}(2+\sqrt{2}\,)$. Если для любого n в сумме их окажется n, то эти два спектра и в самом деле образуют разбиение всех целых положительных чисел.

Верно: при возрастании n на 1 должно возрастать только одно из этих чисел.

Пусть α положительно. Количество элементов в $\mathrm{Spec}(\alpha)$, таких, что они $\leqslant n$, равно

$$\begin{aligned} N(\alpha,n) &= \sum_{k>0}\bigl[\lfloor k\alpha\rfloor \leqslant n\bigr] = \\ &= \sum_{k>0}\bigl[\lfloor k\alpha\rfloor < n+1\bigr] = \\ &= \sum_{k>0}[k\alpha < n+1] = \\ &= \sum_{k}\bigl[0<k<(n+1)/\alpha\bigr] = \\ &= \lceil (n+1)/\alpha\rceil - 1\,. \end{aligned} \tag{3.14}$$

Два момента в этом выводе особенно интересны. Во-первых, в нем для замены '$\leqslant$' на '$<$' используется правило

$$m \leqslant n \quad\Longleftrightarrow\quad m < n+1\,, \qquad m \text{ и } n \text{— целые}, \tag{3.15}$$

так что в соответствии с (3.7) можно убрать скобки пола. Во-вторых, и это более тонкий момент, суммирование выполняется по всем $k>0$, а не $k\geqslant 1$, поскольку величина $(n+1)/\alpha$ может быть меньше 1 при определенных n и α. Если попытаться применить (3.12) для определения количества целых чисел в интервале $[1\,..\,(n+1)/\alpha)$, в отличие от количества целых чисел в интервале $(0\,..\,(n+1)/\alpha)$, то получится верный ответ, но вывод при этом окажется неверен в силу несоблюдения условий его применимости.

Итак, у нас есть формула для $N(\alpha,n)$. Теперь можно проверить, образуют ли $\mathrm{Spec}(\sqrt{2}\,)$ и $\mathrm{Spec}(2+\sqrt{2}\,)$ разбиение положительных целых чисел, путем проверки, выполняется ли соотношение $N(\sqrt{2},n)+N(2+\sqrt{2},n)=n$ для всех целых чисел $n>0$, используя для этого (3.14):

$$\begin{aligned} &\left\lceil\frac{n+1}{\sqrt{2}}\right\rceil - 1 + \left\lceil\frac{n+1}{2+\sqrt{2}}\right\rceil - 1 = n \\ \Longleftrightarrow\quad &\left\lfloor\frac{n+1}{\sqrt{2}}\right\rfloor + \left\lfloor\frac{n+1}{2+\sqrt{2}}\right\rfloor = n \qquad \text{согласно (3.2)}, \end{aligned}$$

$$\iff \quad \frac{n+1}{\sqrt{2}} - \left\{\frac{n+1}{\sqrt{2}}\right\} + \\ + \frac{n+1}{2+\sqrt{2}} - \left\{\frac{n+1}{2+\sqrt{2}}\right\} = n \quad \text{согласно (3.8)}.$$

Все упрощается благодаря изящному равенству

$$\frac{1}{\sqrt{2}} + \frac{1}{2+\sqrt{2}} = 1,$$

так что наше условие сводится к проверке, верно ли, что

$$\left\{\frac{n+1}{\sqrt{2}}\right\} + \left\{\frac{n+1}{2+\sqrt{2}}\right\} = 1$$

для всех $n > 0$. А это верно, потому что перед нами дробные части нецелых чисел, которые в сумме дают целое число $n+1$. Значит, мы действительно имеем дело с разбиением.

3.3 Рекуррентности с полом и потолком

Полы и потолки добавляют новое интересное измерение в изучение рекуррентных соотношений. Давайте для начала рассмотрим рекуррентность

$$\begin{aligned} K_0 &= 1; \\ K_{n+1} &= 1 + \min(2K_{\lfloor n/2 \rfloor}, 3K_{\lfloor n/3 \rfloor}) \qquad \text{при } n \geqslant 0. \end{aligned} \tag{3.16}$$

Так, например, K_1 равно $1+\min(2K_0, 3K_0) = 3$, а начало последовательности имеет вид 1, 3, 3, 4, 7, 7, 7, 9, 9, 10, 13, … . Один из авторов этой книги с присущей ему скромностью решил назвать эти числа числами Кнута.

В упр. 25 требуется доказать или опровергнуть, что $K_n \geqslant n$ при всех $n \geqslant 0$. Первые несколько перечисленных чисел K удовлетворяют этому неравенству, так что, похоже, оно может выполняться и в общем случае. Давайте попробуем доказать его по индукции. База индукции при $n = 0$ непосредственно следует из определения рекуррентности. Для шага индукции предположим, что неравенство выполняется для всех значений до некоторого фиксированного неотрицательного n, и попытаемся показать, что $K_{n+1} \geqslant n+1$. Из рекуррентного соотношения нам известно, что $K_{n+1} = 1+\min(2K_{\lfloor n/2 \rfloor}, 3K_{\lfloor n/3 \rfloor})$. Гипотеза индукции гласит, что $2K_{\lfloor n/2 \rfloor} \geqslant 2\lfloor n/2 \rfloor$ и $3K_{\lfloor n/3 \rfloor} \geqslant 3\lfloor n/3 \rfloor$. Однако $2\lfloor n/2 \rfloor$ может быть равным самое меньшее $n-1$, а $3\lfloor n/3 \rfloor$ может быть равным самое меньшее $n-2$. Самое большее, что нам удается заключить из гипотезы индукции, — это то, что $K_{n+1} \geqslant 1+(n-2)$, но этого чуть-чуть не хватает до $K_{n+1} \geqslant n+1$.

Теперь у нас есть причины усомниться в верности неравенства $K_n \geqslant n$, так что давайте попробуем его опровергнуть. Если нам удастся найти такое n, что либо $2K_{\lfloor n/2 \rfloor} < n$, либо $3K_{\lfloor n/3 \rfloor} < n$, — другими словами, такое, что

$$K_{\lfloor n/2 \rfloor} < n/2 \qquad \text{или} \qquad K_{\lfloor n/3 \rfloor} < n/3\,,$$

то мы получим $K_{n+1} < n+1$. Возможно ли это? Ответ мы пока что не дадим, чтобы не портить впечатление от упр. 25.

Рекуррентные соотношения с полами и/или потолками часто возникают в информатике; это связано с тем, что масса алгоритмов основана на принципе "разделяй и властвуй", которые часто сводят задачу размера n к решению аналогичных задач меньших размеров, являющихся долями n. Например, один из способов отсортировать n записей (где $n > 1$) состоит в разделении их на две примерно равные части, одну — размером $\lceil n/2 \rceil$ и вторую — размером $\lfloor n/2 \rfloor$. (Между прочим, формула

$$n = \lceil n/2 \rceil + \lfloor n/2 \rfloor \tag{3.17}$$

часто оказывается очень кстати при решении такого рода задач.) После того, как каждая из частей отсортирована по отдельности (тем же рекурсивно примененным методом), их можно объединить в требуемом порядке, выполняя не более $n-1$ сравнений. Таким образом, общее количество сравнений не превышает $f(n)$, где

$$\begin{aligned} &f(1) = 0\,, \\ &f(n) = f(\lceil n/2 \rceil) + f(\lfloor n/2 \rfloor) + n - 1 \qquad \text{при } n > 1. \end{aligned} \tag{3.18}$$

Решение этой рекуррентности приводится в упр. 34.

Задача Иосифа из главы 1 содержит подобную рекуррентность, которую можно привести к виду

$$\begin{aligned} J(1) &= 1\,; \\ J(n) &= 2J(\lfloor n/2 \rfloor) - (-1)^n \qquad \text{при } n > 1. \end{aligned}$$

Теперь мы уже оснащены гораздо бо́льшим количеством методов по сравнению с главой 1, так что давайте рассмотрим более приближенную к реальности задачу Иосифа, в которой убивают каждого третьего, а не второго. Если применить методы из главы 1 к этой более сложной задаче, то мы придем к рекуррентности вида

$$J_3(n) = \left(\left\lceil \tfrac{3}{2} J_3\left(\lfloor \tfrac{2}{3} n \rfloor\right) + a_n \right\rceil \bmod n\right) + 1\,,$$

где 'mod' — это функция, с которой мы вскоре познакомимся, а $a_n = -2$, $+1$ или $-\frac{1}{2}$ соответственно при $n \bmod 3 = 0$, 1 или 2. Но на эту рекуррентность страшно даже смотреть, не то что ее решать.

Есть и другой подход к задаче Иосифа, который приводит к более благоприятной ситуации. Всякий раз, когда счет пропускает человека, оставляя его на этом круге в живых, ему можно присваивать новый номер. Таким образом, 1-й и 2-й человек становятся $n+1$-м и $n+2$-м, затем 3-го казнят; 4-й и 5-й становятся $n+3$-м и $n+4$-м, затем 6-го казнят; $\ldots$; $3k+1$-й и $3k+2$-й становятся $n+2k+1$-м и $n+2k+2$-м, затем $3k+3$-го казнят; $\ldots$; затем $3n$-го казнят (или он остается в живых). Например, при $n = 10$ перенумерация имеет следующий вид:

1	2	3	4	5	6	7	8	9	10
11	12		13	14		15	16		17
18			19	20			21		22
			23	24					25
			26						27
			28						
			29						
			30						

Казнимый k-м человек прекращает существование вместе с номером $3k$. Так что можно вывести, кто спасется, если определить исходный номер человека с номером $3n$.

Если $N > n$, то жертва с номером N должна была иметь некоторый предыдущий номер, найти который можно следующим образом: поскольку $N = n + 2k + 1$ или $N = n + 2k + 2$, то $k = \lfloor (N-n-1)/2 \rfloor$. Предыдущий номер был соответственно $3k+1$ или $3k+2$. То есть он был равен $3k + (N - n - 2k) = k + N - n$. Следовательно, номер спасшегося $J_3(n)$ можно вычислить следующим образом:

$$N := 3n;$$

$$\textbf{пока } N > n\textbf{, выполнять } N := \left\lfloor \frac{N-n-1}{2} \right\rfloor + N - n;$$

$$J_3(n) := N.$$

"Not too slow, not too fast".
— Л. Армстронг (L. Armstrong)

Это не аналитическая запись $J_3(n)$ и даже не рекуррентное соотношение. Но как минимум это алгоритм, который при большом n позволяет достаточно быстро получить интересующий нас ответ.

К счастью, имеется возможность упрощения этого алгоритма путем использования переменной $D = 3n + 1 - N$ вместо N. (Это изменение соответствует присвоению номеров, уменьшающихся от $3n$ до 1 вместо увеличивающихся от 1 до $3n$.) Тогда усложненное правило присвоения значений N приобретает вид

$$D := 3n+1-\left(\left\lfloor\frac{(3n+1-D)-n-1}{2}\right\rfloor+(3n+1-D)-n\right) =$$
$$= n+D-\left\lfloor\frac{2n-D}{2}\right\rfloor = D-\left\lfloor\frac{-D}{2}\right\rfloor = D+\left\lceil\frac{D}{2}\right\rceil = \left\lceil\tfrac{3}{2}D\right\rceil,$$

и алгоритм можно переписать следующим образом:

$D := 1$;
пока $D \leqslant 2n$, **выполнять** $D := \lceil\frac{3}{2}D\rceil$;
$J_3(n) := 3n + 1 - D$.

Ага! Это уже гораздо лучше, так как n очень просто входит в вычисления. Фактически, при помощи аналогичных рассуждений можно показать, что если уничтожается каждый q-й человек, то номер уцелевшего $J_q(n)$ можно вычислить следующим образом:

$D := 1$;
пока $D \leqslant (q-1)n$, **выполнять** $D := \lceil\frac{q}{q-1}D\rceil$; (3.19)
$J_q(n) := qn + 1 - D$.

В случае $q = 2$, который хорошо нам знаком, D растет до 2^{m+1} при $n = 2^m + l$; следовательно, $J_2(n) = 2(2^m+l)+1-2^{m+1} = 2l+1$. Отлично.

Знакома ли вам так же хорошо оценка 2?

В (3.19) вычисляется последовательность целых чисел, которая может быть определена при помощи следующего рекуррентного соотношения:

$$D_0^{(q)} = 1;$$
$$D_n^{(q)} = \left\lceil\frac{q}{q-1}D_{n-1}^{(q)}\right\rceil \qquad \text{при } n > 0. \tag{3.20}$$

Непохоже, чтобы эти числа были как-то связаны с известными функциями простым образом, за исключением $q = 2$; следовательно, вряд ли удастся найти для них приемлемую аналитическую запись. Но если принять последовательность $D_n^{(q)}$ как "известную", то решение обобщенной задачи Иосифа очень легко описать: номер спасшегося $J_q(n)$ равен $qn + 1 - D_k^{(q)}$, где k — наименьшее возможное число, такое, что $D_k^{(q)} > (q-1)n$.

"Известную", как, например, гармонические числа. Э. М. Одлыжко (A. M. Odlyzko) и Г. С. Вильф (H. S. Wilf) показали [283], что $D_n^{(3)} = \lfloor(\frac{3}{2})^n C\rfloor$, где $C \approx 1.622270503$.

3.4 'MOD': бинарная операция

Если m и n — целые положительные числа, то частное от деления n на m равно $\lfloor n/m \rfloor$. Полезно ввести простое обозначение и для остатка от этого деления. Обозначим его '$n \bmod m$'. Основная формула

$$n = \underbrace{m\lfloor n/m \rfloor}_{\text{частное}} + \underbrace{n \bmod m}_{\text{остаток}}$$

гласит, что можно выразить $n \bmod m$ как $n - m\lfloor n/m \rfloor$. Можно обобщить эту формулу и на отрицательные целые числа, а фактически и на произвольные действительные числа:

$$x \bmod y = x - y\lfloor x/y \rfloor \qquad \text{при } y \neq 0. \tag{3.21}$$

Откуда взялось название «'mod': бинарная операция»? Вы узнаете это в следующей главе.

Тем самым 'mod' определяется как бинарная операция, такая же, как сложение или вычитание. Неформально математики давно пользуются mod, вычисляя различные величины по модулю 10, по модулю 2π и т.д., но формально это обозначение утвердилось только в последние десятилетия. Так что это — новое обозначение для старого понятия.

Можно легко уловить смысл $x \bmod y$ для положительных действительных чисел x и y интуитивно. Представим себя бегуном по окружности длиной y, точкам которой назначены действительные значения из интервала $[0\,..\,y)$. Если пройти по окружности расстояние x, начиная с точки 0, то мы окажемся в точке $x \bmod y$. (При этом мы $\lfloor x/y \rfloor$ раз пройдем через точку 0.)

Остерегайтесь языков программирования, в которых используется иное определение. (В частности, это относится к языкам программирования C и C++. — Переводчик)

При отрицательных x или y следует внимательно рассмотреть определение операции, чтобы точно понять, что же в этом случае означает операция 'mod'. Вот некоторые примеры с целыми значениями:

$$\begin{aligned}
5 \bmod 3 &= 5 - 3\lfloor 5/3 \rfloor &&= 2\,,\\
5 \bmod -3 &= 5 - (-3)\lfloor 5/(-3) \rfloor &&= -1\,,\\
-5 \bmod 3 &= -5 - 3\lfloor -5/3 \rfloor &&= 1\,,\\
-5 \bmod -3 &= -5 - (-3)\lfloor -5/(-3) \rfloor &&= -2\,.
\end{aligned}$$

*Может, назвать его мофигом?**

Число после 'mod' называется *модулем*; как называть число перед 'mod', пока что не придумали. В приложениях модуль

* В оригинале — "modumor". Непереводимая игра слов, основанная на созвучии второй части слова "модуль" (modulus) слову "less" (меньше), а второй части слова "modumor" — слову "more" (больше). — Переводчик

обычно положителен, но данное выше определение сохраняет смысл и при отрицательных значениях. В обоих случаях значение $x \bmod y$ находится между 0 и модулем:

$$\begin{aligned} 0 \leqslant\ & x \bmod y < y \qquad \text{при } y > 0, \\ 0 \geqslant\ & x \bmod y > y \qquad \text{при } y < 0. \end{aligned}$$

Но что делать с $y = 0$? Определение (3.21), чтобы избежать деления на 0, оставляет данный случай неопределенным, но для полноты можно определить

$$x \bmod 0 = x. \qquad (3.22)$$

Нет модуля — нет и операции!
— Переводчик

Это соглашение сохраняет то свойство, что $x \bmod y$ всегда отличается от x на величину, кратную y. (Может показаться более естественным сделать эту функцию непрерывной в 0 при помощи следующего определения: $x \bmod 0 = \lim_{y\to 0} x \bmod y = 0$. Но в главе 4 мы увидим, что это сделало бы ее менее полезной. Непрерывность — не самый важный аспект операции mod.)

Мы уже сталкивались с одним замаскированным частным случаем операции mod, когда записывали число x в виде суммы его целой и дробной частей: $x = \lfloor x \rfloor + \{x\}$. Дробная часть может быть записана также как $x \bmod 1$, потому что

$$x = \lfloor x \rfloor + x \bmod 1.$$

Обратите внимание, что в этой формуле круглые скобки не нужны: считается, что знак связывает операнды сильнее, чем знаки сложения и вычитания (т.е. *приоритет* этого оператора выше приоритетов операторов сложения или вычитания. — *Примеч. пер.*).

При определении операции mod использовалась функция "пол", в то время как функция "потолок" такого внимания не была удостоена. Возможно, стоило бы воспользоваться функцией "потолок" для определения аналога операции mod наподобие

$$x \operatorname{mumble} y = y\lceil x/y \rceil - x.$$

В 1970-е годы был популярен стиль "mod". Может, стоит назвать новую функцию не "mumble", а "punk" (панк)?

Нет — я автор, как хочу, так и называю! Мне нравится "mumble".

("Mumble" — нечленораздельное бормотание (англ.). — *Примеч. пер.*) В случае нашей аналогии с бегом по кругу это соответствует расстоянию, которое осталось пробежать бегуну, преодолевшему расстояние x, чтобы попасть в исходную точку 0. Но, конечно же, требуется более подходящее название, чем 'mumble'. Впрочем, если бы нашлось применение этой операции, за названием дело бы не стало...

Дистрибутивный закон представляет собой наиболее важное алгебраическое свойство операции mod:

$$c(x \bmod y) \;=\; (cx) \bmod (cy) \tag{3.23}$$

Заметим, что $x \operatorname{mumble} y = (-x) \bmod y$.

для всех действительных c, x и y. (Если кто-то предпочитает считать, что оператор mod связывает операнды слабее умножения, он может убрать скобки из правой части равенства.) Этот закон легко доказать, исходя из определения (3.21), так как

$$c(x \bmod y) = c(x - y\lfloor x/y \rfloor) = cx - cy\lfloor cx/cy \rfloor = cx \bmod cy\,,$$

если $cy \neq 0$; случаи с нулевыми модулями тривиальны. Приведенные ранее четыре примера с ± 5 и ± 3 дважды иллюстрируют этот закон при $c = -1$. Тождество типа (3.23) действует успокаивающе, ибо оно позволяет надеяться, что операция 'mod' определена надлежащим образом.

А попросту — в остатке?

В оставшейся части этого раздела мы рассмотрим приложение операции 'mod', в котором она оказывается весьма полезной, хотя и не играет центральной роли. Рассматриваемая задача часто возникает в многочисленных ситуациях, когда надо максимально равномерно разбить n предметов на m групп.

Предположим, например, что у нас есть n коротких строк текста, которые надо разместить в m столбцах. По эстетическим соображениям желательно, чтобы столбцы располагались в убывающем по высоте порядке (фактически — в неубывающем порядке) и чтобы высота столбцов была примерно одинакова — никакие два из них не должны отличаться по высоте более чем на одну строку текста. Если 37 строк текста разбиваются на пять столбцов, то предпочтительно такое их размещение, как показано справа:

8	8	8	8	5
line 1	line 9	line 17	line 25	line 33
line 2	line 10	line 18	line 26	line 34
line 3	line 11	line 19	line 27	line 35
line 4	line 12	line 20	line 28	line 36
line 5	line 13	line 21	line 29	line 37
line 6	line 14	line 22	line 30	
line 7	line 15	line 23	line 31	
line 8	line 16	line 24	line 32	

8	8	7	7	7
line 1	line 9	line 17	line 24	line 31
line 2	line 10	line 18	line 25	line 32
line 3	line 11	line 19	line 26	line 33
line 4	line 12	line 20	line 27	line 34
line 5	line 13	line 21	line 28	line 35
line 6	line 14	line 22	line 29	line 36
line 7	line 15	line 23	line 30	line 37
line 8	line 16			

Кроме того, хотелось бы распределить строки текста "постолбцово", т.е. вначале решается, сколько строк будет в первом столбце, затем — во втором, третьем и так далее в соответствии с порядком чтения текста человеком. Построчное распределение автоматически приведет к требуемому размещению, но порядок чтения строк при этом окажется нарушенным (мы получили бы нечто подобное размещению справа, но в первом столбце были бы строки 1, 6, 11, ..., 36, а не 1, 2, 3, ..., 8, как должно быть).

Построчную стратегию распределения использовать нельзя, но она позволяет определить,сколько строк разместится в каждом столбце. Если n не кратно m, то в процессе построчного размещения выясняется, что длинные столбцы должны содержать по $\lceil n/m \rceil$ строк, а короткие — по $\lfloor n/m \rfloor$. Длинных столбцов будет ровно $n \bmod m$ (и, как выясняется, ровно $n \operatorname{mumble} m$ коротких).

Давайте обобщим терминологию и поговорим о 'предметах' и 'группах' вместо 'строк' и 'столбцов'. Только что мы выяснили, что первая группа должна содержать $\lceil n/m \rceil$ предметов; следовательно, можно попробовать такую схему последовательного распределения: для распределения n предметов по m группам при $m > 0$ поместим $\lceil n/m \rceil$ предметов в одну группу, а затем рекурсивно применим ту же процедуру к остальным $n' = n - \lceil n/m \rceil$ оставшимся предметам для их размещения в $m' = m - 1$ прочих группах.

Например, если $n = 314$, а $m = 6$, то распределение выполняется следующим образом:

Оставшиеся предметы	Оставшиеся группы	⌈ Предметов в группе ⌉
314	6	53
261	5	53
208	4	52
156	3	52
104	2	52
52	1	52

Как видите, схема работает — мы получаем группы практически неизменного размера, несмотря на то, что делитель изменяется.

Почему она работает? В общем случае можно предположить, что $n = qm + r$, где $q = \lfloor n/m \rfloor$ и $r = n \bmod m$. Если $r = 0$, процесс размещения прост: мы помещаем $\lceil n/m \rceil = q$ предметов в первую группу и заменяем n на $n' = n - q$, оставляя $n' = qm'$ предметов для размещения в остальных $m' = m - 1$ группах. Если же $r > 0$, мы помещаем $\lceil n/m \rceil = q + 1$ предметов в первую группу и заменяем n на $n' = n - q - 1$, оставляя $n' = qm' + r - 1$ предметов для последующих групп. Новый остаток $r' = r - 1$, а q остается тем же самым. Отсюда следует, что получится r групп с $q + 1$ предметами, за которыми следуют $m - r$ групп с q предметами.

Сколько предметов окажется в k-й группе? Для ответа на этот вопрос нам нужна формула, которая давала бы $\lceil n/m \rceil$ при

$k \leqslant n \bmod m$ и $\lfloor n/m \rfloor$ в противном случае. Нетрудно убедиться, что требуемым условиям отвечает формула

$$\left\lceil \frac{n-k+1}{m} \right\rceil,$$

поскольку она сводится к $q + \lceil (r-k+1)/m \rceil$, если, как и в предыдущем абзаце, записать $n = qm + r$ (здесь $q = \lfloor n/m \rfloor$). Получаем, что $\lceil (r-k+1)/m \rceil = [k \leqslant r]$, если $1 \leqslant k \leqslant m$ и $0 \leqslant r < m$. Следовательно, можно записать следующее тождество, которое выражает разбиение n на m как можно более равных частей в невозрастающем порядке:

$$n = \left\lceil \frac{n}{m} \right\rceil + \left\lceil \frac{n-1}{m} \right\rceil + \cdots + \left\lceil \frac{n-m+1}{m} \right\rceil . \tag{3.24}$$

Это тождество справедливо при любом целом положительном m и при любом целом n (положительном, отрицательном или нуле). Мы уже сталкивались со случаем $m = 2$ в (3.17), хотя и записанном в несколько ином виде: $n = \lceil n/2 \rceil + \lfloor n/2 \rfloor$.

Если бы мы пожелали, чтобы все части располагались в неубывающем порядке — когда меньшие группы предшествуют бóльшим, — можно было бы действовать тем же способом, но с $\lfloor n/m \rfloor$ предметами в первой группе, получая соответствующее тождество

$$n = \left\lfloor \frac{n}{m} \right\rfloor + \left\lfloor \frac{n+1}{m} \right\rfloor + \cdots + \left\lfloor \frac{n+m-1}{m} \right\rfloor . \tag{3.25}$$

Тождества (3.25) и (3.24) можно преобразовывать друг в друга, пользуясь либо соотношением (3.4), либо тождеством из упр. 12.

Некоторые утверждают, что слишком опасно заменять что-либо на mx.
(Имеются в виду ракеты MX. — Переводчик)

Теперь, если в (3.25) заменить n на $\lfloor mx \rfloor$ и применить правило (3.11) для удаления полов внутри полов, то получится тождество, выполняющееся при всех действительных x:

$$\lfloor mx \rfloor = \lfloor x \rfloor + \left\lfloor x + \frac{1}{m} \right\rfloor + \cdots + \left\lfloor x + \frac{m-1}{m} \right\rfloor . \tag{3.26}$$

Это в некоторой степени удивительный результат, так как функция "пол" является целочисленной аппроксимацией действительной величины, и тем не менее единственное приближение слева оказывается равным сумме целого набора приближений справа. Если считать, что $\lfloor x \rfloor$ — это, грубо говоря, в среднем $x - \frac{1}{2}$, то левая часть в грубом приближении равна $mx - \frac{1}{2}$, а правая оказывается равной $(x - \frac{1}{2}) + (x - \frac{1}{2} + \frac{1}{m}) + \cdots + (x - \frac{1}{2} + \frac{m-1}{m}) = mx - \frac{1}{2}$. И вот — сумма этих грубых приближений оказывается точной величиной!

3.5 Суммы с полами и потолками

Уравнение (3.26) демонстрирует, что в аналитическом виде можно получить по крайней мере один вид сумм со скобками $\lfloor\ \rfloor$. Есть ли другие виды? Есть. Обычно применяемая в таких случаях уловка состоит в том, чтобы избавиться от пола или потолка, введя новую переменную.

Рассмотрим, например, можно ли получить сумму

$$\sum_{0\leqslant k<n} \lfloor\sqrt{k}\rfloor$$

в аналитическом виде. Идея заключается во введении переменной $m = \lfloor\sqrt{k}\rfloor$; это можно сделать чисто "механически", действуя так же, как в задаче с рулеткой:

$$\begin{aligned}\sum_{0\leqslant k<n} \lfloor\sqrt{k}\rfloor &= \sum_{k,m\geqslant 0} m[k<n]\big[m=\lfloor\sqrt{k}\rfloor\big] = \\ &= \sum_{k,m\geqslant 0} m[k<n]\big[m\leqslant\sqrt{k}<m+1\big] = \\ &= \sum_{k,m\geqslant 0} m[k<n]\big[m^2\leqslant k<(m+1)^2\big] = \\ &= \sum_{k,m\geqslant 0} m\big[m^2\leqslant k<(m+1)^2\leqslant n\big] + \\ &\qquad + \sum_{k,m\geqslant 0} m\big[m^2\leqslant k<n<(m+1)^2\big]\,.\end{aligned}$$

И вновь определенные трудности вызывают граничные условия. Допустим сначала, что $n = a^2$ — точный квадрат. Тогда вторая сумма равна нулю, а первая может быть вычислена по обычному правилу:

$$\begin{aligned}&\sum_{k,m\geqslant 0} m\big[m^2\leqslant k<(m+1)^2\leqslant a^2\big] = \\ &\quad= \sum_{m\geqslant 0} m\big((m+1)^2-m^2\big)[m+1\leqslant a] = \\ &\quad= \sum_{m\geqslant 0} m(2m+1)[m<a] = \\ &\quad= \sum_{m\geqslant 0} (2m^{\underline{2}}+3m^{\underline{1}})[m<a] = \\ &\quad= \sum\nolimits_0^a (2m^{\underline{2}}+3m^{\underline{1}})\,\delta m = \\ &\quad= \tfrac{2}{3}a(a-1)(a-2)+\tfrac{3}{2}a(a-1) = \tfrac{1}{6}(4a+1)a(a-1)\,.\end{aligned}$$

Убывающие степени убивают сумму.

В общем случае можно положить $a = \lfloor\sqrt{n}\rfloor$; тогда нужно всего лишь сложить члены при $a^2 \leqslant k < n$, причем все они равны a, так что сумма оказывается равной $(n-a^2)a$. Это дает нам искомый аналитический вид

$$\sum_{0\leqslant k<n} \lfloor\sqrt{k}\rfloor = na - \tfrac{1}{3}a^3 - \tfrac{1}{2}a^2 - \tfrac{1}{6}a, \qquad a = \lfloor\sqrt{n}\rfloor. \tag{3.27}$$

Другой подход к вычислению таких сумм состоит в замене выражения вида $\lfloor x\rfloor$ на $\sum_j [1\leqslant j\leqslant x]$; это всегда законно при $x \geqslant 0$. Вот как работает этот метод при вычислении суммы $\lfloor$квадратных корней$\rfloor$, если для удобства считать, что $n = a^2$:

$$\begin{aligned}\sum_{0\leqslant k<n} \lfloor\sqrt{k}\rfloor &= \sum_{j,k} [1\leqslant j\leqslant \sqrt{k}\,][0\leqslant k<a^2] = \\ &= \sum_{1\leqslant j<a}\ \sum_{k} [j^2\leqslant k<a^2] = \\ &= \sum_{1\leqslant j<a} (a^2-j^2) = a^3 - \tfrac{1}{3}a(a+\tfrac{1}{2})(a+1)\,.\end{aligned}$$

Вот другой пример, в котором замена переменной приводит к преобразованной сумме. Примерно в одно и то же время в 1909 году независимо друг от друга три математика — Боль (Bohl) [34], Серпиньски (Sierpiński) [326] и Вейль (Weyl) [368] — обнаружили замечательный факт: если α иррационально, то при $n \to \infty$ числа $\{n\alpha\}$ в высшей степени равномерно распределены между 0 и 1. Одна из формулировок звучит следующим образом:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{0\leqslant k<n} f\bigl(\{k\alpha\}\bigr) = \int_0^1 f(x)\,dx \tag{3.28}$$

при любом иррациональном α и любой ограниченной функции f, которая непрерывна почти везде. Например, *среднее* значение $\{n\alpha\}$ можно найти, положив $f(x) = x$. Мы получим значение $\frac{1}{2}$. (Именно этого и следовало ожидать; но приятно знать, что это верно независимо от того, чему именно равно иррациональное число α.)

Внимание: это достаточно сложный материал. При первом чтении лучше бегло проскочить следующие пару страниц: они не так существенны.
— Сочувствующий ассистент

Теорема Боля, Серпиньски и Вейля доказывается путем аппроксимации $f(x)$ сверху и снизу "ступенчатыми функциями", которые являются линейными комбинациями простых функций

$$f_\nu(x) = [0\leqslant x<\nu]$$

На взлет!

при $0 \leqslant \nu \leqslant 1$. Наша цель состоит не в том, чтобы доказать теорему, — это дело других книг по матанализу. Мы же попробуем выяснить фундаментальные причины, по которым она справедлива, рассмотрев частный случай $f(x) = f_\nu(x)$. Другими словами, давайте попытаемся выяснить, насколько близка сумма

$$\sum_{0\leqslant k<n} \big[\{k\alpha\}<\nu\big]$$

к "идеальному" значению $n\nu$ при большом n и иррациональном α.

Для этой цели мы определим *отклонение* $D(\alpha, n)$ как максимальное абсолютное значение по всем $0 \leqslant \nu \leqslant 1$ суммы

$$s(\alpha, n, \nu) \;=\; \sum_{0\leqslant k<n} \Big(\big[\{k\alpha\}<\nu\big] - \nu\Big)\,. \qquad (3.29)$$

Наша задача — показать, что $D(\alpha, n)$ "не слишком велико" по сравнению с n, показав, что величина $|s(\alpha, n, \nu)|$ всегда достаточно мала при иррациональном α. Без потери общности можно считать, что $0 < \alpha < 1$.

Сначала можно переписать $s(\alpha, n, \nu)$ в более простом виде, а затем ввести новую индексную переменную j:

$$\begin{aligned}\sum_{0\leqslant k<n} \Big(\big[\{k\alpha\}<\nu\big] - \nu\Big) \;&=\; \sum_{0\leqslant k<n} \big(\lfloor k\alpha\rfloor - \lfloor k\alpha-\nu\rfloor - \nu\big) \;= \\ &= -n\nu + \sum_{0\leqslant k<n}\,\sum_{j} \big[k\alpha-\nu<j\leqslant k\alpha\big] \;= \\ &= -n\nu + \sum_{0\leqslant j<\lceil n\alpha\rceil}\,\sum_{k<n} \big[j\alpha^{-1}\leqslant k<(j+\nu)\alpha^{-1}\big]\,.\end{aligned}$$

Если нам повезет, мы сможем просуммировать по k. Но прежде сто́ит ввести некоторые новые переменные, чтобы привести эту формулу к приличному виду. Положим

Знакомо: именуй и властвуй. Главное здесь — замена переменной k на j.
— Сочувствующий ассистент

$$\begin{aligned} a &= \lfloor\alpha^{-1}\rfloor\,, &\qquad \alpha^{-1} &= a+\alpha'\,, \\ b &= \lceil\nu\alpha^{-1}\rceil\,, &\qquad \nu\alpha^{-1} &= b-\nu'\,. \end{aligned}$$

Таким образом, $\alpha' = \{\alpha^{-1}\}$ — это дробная часть α^{-1}, а ν' — mumble-дробная часть $\nu\alpha^{-1}$.

И вновь единственный источник неприятностей — это граничные условия. Давайте пока забудем об ограничении '$k < n$' и вычислим сумму по k без него:

$$\begin{aligned}\sum_{k}\Big[k\in\big[j\alpha^{-1}\,..\,(j+\nu)\alpha^{-1}\big)\Big] \;&=\; \big\lceil(j+\nu)(a+\alpha')\big\rceil - \big\lceil j(a+\alpha')\big\rceil \;= \\ &= b + \lceil j\alpha'-\nu'\rceil - \lceil j\alpha'\rceil\,.\end{aligned}$$

Все просто; теперь выполняем подстановку

$$s(\alpha, n, \nu) = \\ = -n\nu + \lceil n\alpha \rceil b + \sum_{0 \leqslant j < \lceil n\alpha \rceil} \left(\lceil j\alpha' - \nu' \rceil - \lceil j\alpha' \rceil \right) - S, \quad (3.30)$$

где S — поправка для случаев $k \geqslant n$, которые мы не смогли исключить. Величина $j\alpha'$ будет целой только при $j = 0$, поскольку α (а значит, и α') иррационально; величина $j\alpha' - \nu'$ будет целой не более чем для одного значения j. Так что можно заменить "потолочные" члены "половыми":

$$s(\alpha, n, \nu) = \\ = -n\nu + \lceil n\alpha \rceil b - \sum_{0 \leqslant j < \lceil n\alpha \rceil} \left(\lfloor j\alpha' \rfloor - \lfloor j\alpha' - \nu' \rfloor \right) - S + \{0 \text{ или } 1\}.$$

Запись $\{0$ или $1\}$ означает нечто, равное 0 либо 1; нет смысла связывать себя деталями, не имеющими никакого значения.

Интересно. Вместо аналитического вида мы получаем сумму, которая имеет вид, похожий на $s(\alpha, n, \nu)$, но с иными параметрами: α' вместо α, $\lceil n\alpha \rceil$ вместо n и ν' вместо ν. А что если получить рекуррентное соотношение для $s(\alpha, n, \nu)$, которое, как мы надеемся, приведет к рекуррентному соотношению для отклонения $D(\alpha, n)$? Это означает, что мы хотим ввести в действие сумму

$$s(\alpha', \lceil n\alpha \rceil, \nu') = \sum_{0 \leqslant j < \lceil n\alpha \rceil} \left(\lfloor j\alpha' \rfloor - \lfloor j\alpha' - \nu' \rfloor - \nu' \right),$$

получая рекуррентность

$$s(\alpha, n, \nu) = \\ = -n\nu + \lceil n\alpha \rceil b - \lceil n\alpha \rceil \nu' - s(\alpha', \lceil n\alpha \rceil, \nu') - S + \{0 \text{ или } 1\}.$$

Вспоминая, что $b - \nu' = \nu\alpha^{-1}$, мы обнаруживаем, что все прекрасно упростится, если мы заменим $\lceil n\alpha \rceil (b - \nu')$ на $n\alpha(b - \nu') = n\nu$:

$$s(\alpha, n, \nu) = -s(\alpha', \lceil n\alpha \rceil, \nu') - S + \epsilon + \{0 \text{ или } 1\}.$$

Здесь ϵ — положительная ошибка, меньшая $\nu\alpha^{-1}$. В упр. 18 доказывается, что подобно ей S лежит в пределах от 0 до $\lceil \nu\alpha^{-1} \rceil$. Можно также удалить из суммы член $j = \lceil n\alpha \rceil - 1 = \lfloor n\alpha \rfloor$, поскольку его внесение в сумму — либо ν', либо $\nu' - 1$. Следовательно, если мы берем максимум абсолютных значений по всем ν, то получаем

$$D(\alpha, n) \leqslant D(\alpha', \lfloor \alpha n \rfloor) + \alpha^{-1} + 2. \quad (3.31)$$

Методы, которые мы будем изучать в последующих главах, позволяют сделать из этого рекуррентного соотношения вывод о том, что $D(\alpha, n)$ всегда гораздо меньше n, если n достаточно велико. Следовательно, теорема (3.28) справедлива; однако сходимость к пределу не всегда очень быстрая (см. упр. 9.45 и 9.61).

Посадка

Ну вот, это было всего лишь небольшое упражнение по работе с суммами, полами и потолками. Читателям, не приученным "доказывать, что погрешности невелики", трудно поверить в то, что у кого-то хватает смелости не отступать перед такими страшными суммами. На самом деле второй взгляд показывает, что все это вычисление пронизывает одна простая мысль. Основная идея в том, что некоторая сумма $s(\alpha, n, \nu)$ из n членов может быть сведена к аналогичной сумме из не более чем $\lceil \alpha n \rceil$ членов. Все остальное сводится на нет, за исключением небольшого остатка, образованного из близких к граничным членов.

Вдохните поглубже — сейчас мы вычислим еще одну сумму, которая, несмотря на свою нетривиальность, обладает тем преимуществом (по сравнению с только что вычисленной), что выражается в аналитическом виде, так что можно легко проверить полученный ответ. Теперь наша задача будет состоять в обобщении суммы в (3.26). Нам требуется найти выражение для

$$\sum_{0 \leqslant k < m} \left\lfloor \frac{nk + x}{m} \right\rfloor, \qquad m, n \text{— целые}, \quad m > 0.$$

Поиск этой суммы в аналитическом виде — крепкий орешек, разгрызть который потруднее, чем те, с которыми мы имели дело до сих пор (за исключением, возможно, задачи об отклонении, с которой только что расправились). Но эта задача настолько поучительна, что мы провозимся с ней до конца главы.

Это более твердая сумма полов или сумма более твердых полов?

Как обычно, особенно при решении трудных задач, начнем с рассмотрения простых крайних случаев. Частный случай $n = 1$ представляет собой тождество (3.26), в котором x заменено на x/m:

$$\left\lfloor \frac{x}{m} \right\rfloor + \left\lfloor \frac{1 + x}{m} \right\rfloor + \cdots + \left\lfloor \frac{m - 1 + x}{m} \right\rfloor = \lfloor x \rfloor .$$

Предупреждаем: это только начало тянущегося до конца главы решения длинной и сложной задачи, для решения которой нет других поводов, кроме любопытства.

— Студенты

И, как в главе 1, имеет смысл получить дополнительную информацию, обратившись к случаю $n = 0$:

$$\left\lfloor \frac{x}{m} \right\rfloor + \left\lfloor \frac{x}{m} \right\rfloor + \cdots + \left\lfloor \frac{x}{m} \right\rfloor = m \left\lfloor \frac{x}{m} \right\rfloor .$$

Можно согласиться. Но что это за поколение, которому всегда требуются практическая польза и прибыль? Неужели в вас умерла любознательность?

У задачи два параметра — m и n; давайте рассмотрим малые случаи для параметра m. При $m = 1$ сумма состоит из единственного члена $\lfloor x \rfloor$. Если же $m = 2$, сумма равна $\lfloor x/2 \rfloor + \lfloor (x+n)/2 \rfloor$. Можно избавиться от взаимодействия x и n, если вынести n за

Эта сумма возникает, например, при изучении и тестировании генерации случайных чисел. Но математики рассматривали эту сумму задолго до появления компьютеров, поскольку считали совершенно естественным поинтересоваться, нет ли способа просуммировать арифметическую прогрессию, "поставленную на пол".
— Преподаватель

знак функции "пол", но, чтобы сделать это, надо по отдельности рассмотреть случаи четного и нечетного n. Если n четное, $n/2$ — целое число, так что можно вынести его за знак пола:

$$\left\lfloor\frac{x}{2}\right\rfloor + \left(\left\lfloor\frac{x}{2}\right\rfloor + \frac{n}{2}\right) = 2\left\lfloor\frac{x}{2}\right\rfloor + \frac{n}{2}.$$

Если же n нечетное, то $(n-1)/2$ — целое число, и мы получаем

$$\left\lfloor\frac{x}{2}\right\rfloor + \left(\left\lfloor\frac{x+1}{2}\right\rfloor + \frac{n-1}{2}\right) = \lfloor x\rfloor + \frac{n-1}{2}.$$

Последний шаг следует из (3.26) при $m = 2$.

Эти формулы для четного и нечетного n немного напоминают формулы для $n = 0$ и 1, но ясной закономерности пока что не видно. Поэтому давайте продолжим рассмотрение малых случаев. При $m = 3$ сумма равна

$$\left\lfloor\frac{x}{3}\right\rfloor + \left\lfloor\frac{x+n}{3}\right\rfloor + \left\lfloor\frac{x+2n}{3}\right\rfloor,$$

и возможны три варианта для n: либо оно кратно 3, либо оно на 1 или на 2 больше кратного, т.е. $n \bmod 3 = 0$, 1 или 2. Если $n \bmod 3 = 0$, то $n/3$ и $2n/3$ — целые числа, так что сумма равна

$$\left\lfloor\frac{x}{3}\right\rfloor + \left(\left\lfloor\frac{x}{3}\right\rfloor + \frac{n}{3}\right) + \left(\left\lfloor\frac{x}{3}\right\rfloor + \frac{2n}{3}\right) = 3\left\lfloor\frac{x}{3}\right\rfloor + n.$$

"Пытливый ум требует приносящей удовлетворение умственной деятельности как необходимого условия его совершенствования. "Необходимость — источник открытий" — бесхитростная поговорка. "Необходимость — источник бесплодных усилий" — это гораздо ближе к истине. Основой роста современных открытий является наука, а наука почти целиком произрастает на почве удовлетворения собственного любопытства".
— А. Н. Уайтхед (A. N. Whitehead) [371]

Если $n \bmod 3 = 1$, то $(n-1)/3$ и $(2n-2)/3$ — целые числа, так что

$$\left\lfloor\frac{x}{3}\right\rfloor + \left(\left\lfloor\frac{x+1}{3}\right\rfloor + \frac{n-1}{3}\right) + \left(\left\lfloor\frac{x+2}{3}\right\rfloor + \frac{2n-2}{3}\right) = \lfloor x\rfloor + n - 1.$$

Последний шаг опять следует из (3.26), на этот раз при $m = 3$. И наконец, если $n \bmod 3 = 2$, то

$$\left\lfloor\frac{x}{3}\right\rfloor + \left(\left\lfloor\frac{x+2}{3}\right\rfloor + \frac{n-2}{3}\right) + \left(\left\lfloor\frac{x+1}{3}\right\rfloor + \frac{2n-1}{3}\right) = \lfloor x\rfloor + n - 1.$$

Наши левые полушария мозга уже разобрались с $m = 3$, но правые пока отстают, так что рассмотрим случай $m = 4$:

$$\left\lfloor\frac{x}{4}\right\rfloor + \left\lfloor\frac{x+n}{4}\right\rfloor + \left\lfloor\frac{x+2n}{4}\right\rfloor + \left\lfloor\frac{x+3n}{4}\right\rfloor.$$

Как минимум теперь мы знаем достаточно, чтобы рассмотреть случаи, основываясь на значении $n \bmod m$. Если $n \bmod 4 = 0$, то

$$\left\lfloor\frac{x}{4}\right\rfloor + \left(\left\lfloor\frac{x}{4}\right\rfloor + \frac{n}{4}\right) + \left(\left\lfloor\frac{x}{4}\right\rfloor + \frac{2n}{4}\right) + \left(\left\lfloor\frac{x}{4}\right\rfloor + \frac{3n}{4}\right) = 4\left\lfloor\frac{x}{4}\right\rfloor + \frac{3n}{2}.$$

А если $n \bmod 4 = 1$, то

$$\left\lfloor\frac{x}{4}\right\rfloor + \left(\left\lfloor\frac{x+1}{4}\right\rfloor + \frac{n-1}{4}\right) + \left(\left\lfloor\frac{x+2}{4}\right\rfloor + \frac{2n-2}{4}\right) + \\ + \left(\left\lfloor\frac{x+3}{4}\right\rfloor + \frac{3n-3}{4}\right) = \lfloor x\rfloor + \frac{3n}{2} - \frac{3}{2}.$$

Случай $n \bmod 4 = 3$, оказывается, дает тот же ответ. Наконец, в случае $n \bmod 4 = 2$ мы получаем нечто несколько иное, оказывающееся важным ключом к установлению поведения в общем случае:

$$\left\lfloor\frac{x}{4}\right\rfloor + \left(\left\lfloor\frac{x+2}{4}\right\rfloor + \frac{n-2}{4}\right) + \left(\left\lfloor\frac{x}{4}\right\rfloor + \frac{2n}{4}\right) + \left(\left\lfloor\frac{x+2}{4}\right\rfloor + \frac{3n-2}{4}\right) = \\ = 2\left(\left\lfloor\frac{x}{4}\right\rfloor + \left\lfloor\frac{x+2}{4}\right\rfloor\right) + \frac{3n}{2} - 1 = 2\left\lfloor\frac{x}{2}\right\rfloor + \frac{3n}{2} - 1.$$

На последнем шаге выполняется упрощение выражения, имеющего вид $\lfloor y/2\rfloor + \lfloor(y+1)/2\rfloor$, что опять представляет собой частный случай (3.26).

Вот сводка значений нашей суммы при малых значениях m:

m	$n \bmod m = 0$	$n \bmod m = 1$	$n \bmod m = 2$	$n \bmod m = 3$
1	$\lfloor x\rfloor$			
2	$2\left\lfloor\frac{x}{2}\right\rfloor + \frac{n}{2}$	$\lfloor x\rfloor + \frac{n}{2} - \frac{1}{2}$		
3	$3\left\lfloor\frac{x}{3}\right\rfloor + n$	$\lfloor x\rfloor + n - 1$	$\lfloor x\rfloor + n - 1$	
4	$4\left\lfloor\frac{x}{4}\right\rfloor + \frac{3n}{2}$	$\lfloor x\rfloor + \frac{3n}{2} - \frac{3}{2}$	$2\left\lfloor\frac{x}{2}\right\rfloor + \frac{3n}{2} - 1$	$\lfloor x\rfloor + \frac{3n}{2} - \frac{3}{2}$

Это выглядит так, как если бы у нас было нечто, имеющее вид

$$a\left\lfloor\frac{x}{a}\right\rfloor + bn + c\,,$$

где a, b и c некоторым образом зависят от m и n. Даже близорукие могут увидеть, что b, вероятно, равно $(m-1)/2$. Труднее разглядеть выражение для a; но случай $n \bmod 4 = 2$ подсказывает нам, что, вероятно, a представляет собой НОД(m,n) — наибольший общий делитель m и n. Это не лишено смысла, так как НОД(m,n) — это тот множитель, который мы убираем из m и n при приведении дроби n/m к несократимой, а наша сумма включает в себя дробь n/m. (Операции с наибольшим общим делителем будут детально рассматриваться в главе 4.) Значение c кажется более загадочным, но будем надеяться, что оно получится при доказательстве наших предположений об a и b.

Вычисляя сумму при малых m, мы, по сути, переписали каждый ее член в виде

$$\left\lfloor\frac{x+kn}{m}\right\rfloor = \left\lfloor\frac{x+kn \bmod m}{m}\right\rfloor + \frac{kn}{m} - \frac{kn \bmod m}{m},$$

поскольку $(kn - kn \bmod m)/m$ — целое число, которое может быть вынесено за скобки пола. Таким образом, исходную сумму можно развернуть в следующую таблицу:

$$\begin{array}{lcccc}
 & \left\lfloor\dfrac{x}{m}\right\rfloor & + \quad \dfrac{0}{m} & - & \dfrac{0 \bmod m}{m} \\
+ & \left\lfloor\dfrac{x+n \bmod m}{m}\right\rfloor & + \quad \dfrac{n}{m} & - & \dfrac{n \bmod m}{m} \\
+ & \left\lfloor\dfrac{x+2n \bmod m}{m}\right\rfloor & + \quad \dfrac{2n}{m} & - & \dfrac{2n \bmod m}{m} \\
 & \vdots & \vdots & & \vdots \\
+ & \left\lfloor\dfrac{x+(m-1)n \bmod m}{m}\right\rfloor & + \quad \dfrac{(m-1)n}{m} & - & \dfrac{(m-1)n \bmod m}{m}.
\end{array}$$

При экспериментах с малыми значениями m эти столбцы приводили нас соответственно к $a\lfloor x/a\rfloor$, bn и c.

В частности, теперь можно понять, откуда берется b. Второй столбец представляет собой арифметическую прогрессию с известной нам суммой, которая равна среднему первого и последнего членов, помноженному на число членов:

$$\frac{1}{2}\left(0+\frac{(m-1)n}{m}\right)\cdot m = \frac{(m-1)n}{2}.$$

Так что наша догадка о том, что $b = (m-1)/2$, проверена и подтверждена.

Первый и третий столбцы так легко не сдаются; чтобы определить a и c, надо повнимательнее рассмотреть последовательность чисел

$$0 \bmod m,\quad n \bmod m,\quad 2n \bmod m,\quad \ldots,\quad (m-1)n \bmod m.$$

Предположим, например, что $m = 12$ и $n = 5$. Если представить эту последовательность в виде циферблата, то эти числа представляют собой 0 часов (примем 12 часов за 0), затем — 5 часов, 10 часов, 3 часа (= 15 часов), 8 часов и т.д. Оказывается, что каждый час встречается только один раз.

Предположим теперь, что $m = 12$ и $n = 8$. Тогда данные числа представляют собой 0 часов, 8 часов, 4 часа (= 16 часов), но затем 0, 8 и 4 повторяются. Поскольку 8 и 12 кратны 4 и числа начинаются с 0 (тоже кратного 4), вырваться из этого замкнутого круга невозможно — все числа должны быть кратны 4.

В этих двух случаях мы имеем НОД$(12,5) = 1$ и НОД$(12,8) = 4$. Общее правило, которое будет доказано в следующей главе, гласит, что если $d = $ НОД(m,n), то мы получаем числа 0, d, $2d$, $\ldots$, $m-d$ в некотором порядке, за которыми следует еще $d-1$ копий той же последовательности. Например, при $m = 12$ и $n = 8$ набор 0, 8, 4 встречается четыре раза.

Сначала лемма, затем дилемма.

Первый столбец нашей суммы приобретает смысл. Он содержит d копий членов $\lfloor x/m \rfloor$, $\lfloor (x+d)/m \rfloor$, $\ldots$, $\lfloor (x+m-d)/m \rfloor$ в некотором порядке, так что их сумма равна

$$
\begin{aligned}
d\left(\left\lfloor\frac{x}{m}\right\rfloor + \left\lfloor\frac{x+d}{m}\right\rfloor + \cdots + \left\lfloor\frac{x+m-d}{m}\right\rfloor\right) &= \\
= d\left(\left\lfloor\frac{x/d}{m/d}\right\rfloor + \left\lfloor\frac{x/d+1}{m/d}\right\rfloor + \cdots + \left\lfloor\frac{x/d+m/d-1}{m/d}\right\rfloor\right) &= \\
= d\left\lfloor\frac{x}{d}\right\rfloor .
\end{aligned}
$$

Здесь последний шаг заключается в еще одном применении (3.26). Наша догадка относительно a также проверена и подтверждена:

$$a = d = \text{НОД}(m,n).$$

Как мы и предполагали, теперь мы можем вычислить c, поскольку третий столбец становится понятен — он содержит d копий арифметической прогрессии $0/m$, d/m, $2d/m$, $\ldots$, $(m-d)/m$, так что их сумма равна

$$d\left(\frac{1}{2}\left(0 + \frac{m-d}{m}\right)\cdot\frac{m}{d}\right) = \frac{m-d}{2}.$$

Поскольку в третьем столбце члены не складываются, а вычитаются, получаем

$$c = \frac{d-m}{2}.$$

Чудеса и приключения закончены. Вот искомый аналитический вид:

$$\sum_{0\leqslant k<m}\left\lfloor\frac{nk+x}{m}\right\rfloor = d\left\lfloor\frac{x}{d}\right\rfloor + \frac{m-1}{2}n + \frac{d-m}{2},$$

где $d = \text{НОД}(m, n)$. Для проверки можно убедиться, что эта формула работает для частных случаев $n = 0$ и $n = 1$, с которыми мы встречались ранее. При $n = 0$ мы получаем $d = \text{НОД}(m, 0) = m$; два последних члена обращаются в нуль, так что формула дает верный ответ $m\lfloor x/m \rfloor$. При $n = 1$ мы получаем $d = \text{НОД}(m, 1) = 1$; два последних члена сокращаются, и сумма равна просто $\lfloor x \rfloor$.

Немного повозившись с аналитической записью, ее можно даже сделать симметричной относительно m и n:

$$\begin{aligned}\sum_{0\leqslant k<m}\left\lfloor\frac{nk+x}{m}\right\rfloor &= d\left\lfloor\frac{x}{d}\right\rfloor + \frac{m-1}{2}n + \frac{d-m}{2} = \\ &= d\left\lfloor\frac{x}{d}\right\rfloor + \frac{(m-1)(n-1)}{2} + \frac{m-1}{2} + \frac{d-m}{2} = \\ &= d\left\lfloor\frac{x}{d}\right\rfloor + \frac{(m-1)(n-1)}{2} + \frac{d-1}{2}. \end{aligned} \tag{3.32}$$

Все, я повержен на пол...

Это немного неожиданно, поскольку нет никаких алгебраических причин подозревать такую симметрию. Мы доказали "закон взаимности"

$$\sum_{0\leqslant k<m}\left\lfloor\frac{nk+x}{m}\right\rfloor = \sum_{0\leqslant k<n}\left\lfloor\frac{mk+x}{n}\right\rfloor, \qquad \text{целые } m, n > 0.$$

Например, если $m = 41$ и $n = 127$, левая сумма состоит из 41 члена, а правая — из 127; но они остаются равными при любом действительном x.

Упражнения

Разминка

1 Анализируя задачу Иосифа в главе 1, мы представляли произвольное положительное целое число n в виде $n = 2^m + l$, где $0 \leqslant l < 2^m$. Приведите явные формулы для l и m как функций от n с использованием скобок пола и/или потолка.

2 Какой вид имеет формула для ближайшего целого к данному действительному числу x? Приведите формулы, которые в случае "равновесия", когда x находится ровно посредине между двумя целыми числами, округляют результат (а) в сторону увеличения, до $\lceil x \rceil$; (б) в сторону уменьшения, до $\lfloor x \rfloor$.

3 Вычислите $\lfloor \lfloor m\alpha \rfloor n/\alpha \rfloor$, где m и n — положительные целые числа, а α — иррациональное число, большее n.

4 В главе описаны задачи с первого по пятый уровень. А что собой представляет задача нулевого уровня? (Кстати, это упражнение — задача *не* нулевого уровня.)

5 Найдите необходимое и достаточное условие того, что $\lfloor nx \rfloor = n\lfloor x \rfloor$, где n — положительное целое число. (Ваше условие должно включать $\{x\}$.)

6 Что интересного можно сказать об $\lfloor f(x) \rfloor$, где $f(x)$ — непрерывная монотонно *убывающая* функция, которая принимает целые значения, только когда x — целое число?

7 Решите рекуррентное соотношение

$$\begin{aligned} X_n &= n && \text{при } 0 \leqslant n < m; \\ X_n &= X_{n-m} + 1 && \text{при } n \geqslant m. \end{aligned}$$

8 Докажите *принцип ящиков Дирихле*: если n предметов размещаются в m ящиках, то некоторый ящик должен содержать $\geqslant \lceil n/m \rceil$ предметов, а другой — $\leqslant \lfloor n/m \rfloor$ предметов.

9 Египетские математики в 1800 г. до н.э. представляли рациональные числа между 0 и 1 как сумму "единичных дробей" $1/x_1 + \cdots + 1/x_k$, где все x — различные положительные целые числа. Например, они записывали $\frac{1}{3} + \frac{1}{15}$ вместо $\frac{2}{5}$. Докажите, что всегда имеется возможность систематической записи следующим образом: если $0 < m/n < 1$, то

$$\frac{m}{n} = \frac{1}{q} + \left\{\text{представление } \frac{m}{n} - \frac{1}{q}\right\}, \qquad q = \left\lceil \frac{n}{m} \right\rceil .$$

(Это *алгоритм Фибоначчи*, открытый Леонардо Фибоначчи (Leonardo Fibonacci) в 1202 году.)

Обязательные упражнения

10 Покажите, что выражение

$$\left\lceil \frac{2x+1}{2} \right\rceil - \left\lceil \frac{2x+1}{4} \right\rceil + \left\lfloor \frac{2x+1}{4} \right\rfloor$$

всегда равно либо $\lfloor x \rfloor$, либо $\lceil x \rceil$. При каких условиях получается тот или иной случай?

11 Приведите детали представленного в тексте доказательства того, что открытый интервал $(\alpha \mathinner{..} \beta)$ содержит ровно $\lceil \beta \rceil - \lfloor \alpha \rfloor - 1$ целых чисел при $\alpha < \beta$. Почему для корректности доказательства нужно исключить случай $\alpha = \beta$?

12 Докажите, что

$$\left\lceil \frac{n}{m} \right\rceil = \left\lfloor \frac{n+m-1}{m} \right\rfloor$$

при любом целом n и любом положительном целом m. (Это тождество дает другой способ превращения потолков в полы и обратно вместо применения рефлексивного закона (3.4).)

13 Пусть α и β — положительные действительные числа. Докажите, что $\mathrm{Spec}(\alpha)$ и $\mathrm{Spec}(\beta)$ образуют разбиение всех положительных целых чисел тогда и только тогда, когда α и β иррациональны и $1/\alpha + 1/\beta = 1$.

14 Докажите или опровергните:

$$(x \bmod ny) \bmod y = x \bmod y \qquad \text{при целом } n.$$

15 Имеется ли тождество, аналогичное (3.26), в котором вместо полов используются потолки?

16 Докажите, что $n \bmod 2 = \bigl(1-(-1)^n\bigr)/2$. Найдите и докажите аналогичное выражение для $n \bmod 3$ вида $a + b\omega^n + c\omega^{2n}$, где ω — комплексное число $(-1+i\sqrt{3}\,)/2$. *Указание*: $\omega^3 = 1$ и $1 + \omega + \omega^2 = 0$.

17 Вычислите сумму $\sum_{0 \leqslant k < m} \lfloor x + k/m \rfloor$ в случае $x \geqslant 0$ путем подстановки $\sum_j [1 \leqslant j \leqslant x + k/m]$ вместо $\lfloor x + k/m \rfloor$ с последующим суммированием сначала по k. Согласуется ли ваш ответ с (3.26)?

18 Докажите, что остаточный член S в (3.30) не превышает $\lceil \alpha^{-1} v \rceil$. *Указание*: покажите, что малые значения j не существенны.

Домашние задания

19 Найдите необходимое и достаточное условие для действительного числа $b > 1$, такое, что

$$\lfloor \log_b x \rfloor = \lfloor \log_b \lfloor x \rfloor \rfloor$$

справедливо для всех $x \geqslant 1$.

20 Найдите сумму всех кратных x чисел в замкнутом интервале $[\alpha \,..\, \beta]$ при $x > 0$.

21 Сколько чисел вида 2^m, где $0 \leqslant m \leqslant M$, начинаются в десятичной записи с цифры 1?

22 Вычислите суммы $S_n = \sum_{k \geqslant 1} \lfloor n/2^k + \frac{1}{2} \rfloor$ и $T_n = \sum_{k \geqslant 1} 2^k \lfloor n/2^k + \frac{1}{2} \rfloor^2$.

23 Покажите, что n-й элемент последовательности

$$1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, \ldots$$

равен $\lfloor\sqrt{2n}+\frac{1}{2}\rfloor$. (Последовательность включает каждое число m ровно m раз.)

24 В упр. 13 установлено интересное соотношение между мультимножествами $\mathrm{Spec}(\alpha)$ и $\mathrm{Spec}(\alpha/(\alpha-1))$, когда α— иррациональное число > 1, поскольку $1/\alpha + (\alpha-1)/\alpha = 1$. Найдите (и докажите) другое интересное соотношение между мультимножествами $\mathrm{Spec}(\alpha)$ и $\mathrm{Spec}(\alpha/(\alpha+1))$, где α— произвольное положительное целое число.

25 Докажите или опровергните, что числа Кнута (3.16) удовлетворяют условию $K_n \geqslant n$ для всех неотрицательных n.

26 Покажите, что вспомогательные числа Иосифа (3.20) удовлетворяют неравенству

$$\left(\frac{q}{q-1}\right)^n \leqslant D_n^{(q)} \leqslant q\left(\frac{q}{q-1}\right)^n \qquad \text{при } n \geqslant 0.$$

27 Докажите, что среди чисел $D_n^{(3)}$, определенных в (3.20), бесконечно много как четных, так и нечетных.

28 Решите рекуррентное соотношение

$$\begin{aligned} a_0 &= 1; \\ a_n &= a_{n-1} + \lfloor\sqrt{a_{n-1}}\rfloor \qquad \text{при } n > 0. \end{aligned}$$

29 Покажите в дополнение к (3.31), что

Эта формула отклоняется от (3.31).

$$D(\alpha, n) \geqslant D(\alpha', \lfloor\alpha n\rfloor) - \alpha^{-1} - 2.$$

30 Покажите, что рекуррентное соотношение

$$\begin{aligned} X_0 &= m, \\ X_n &= X_{n-1}^2 - 2 \qquad \text{при } n > 0 \end{aligned}$$

имеет решение $X_n = \lceil\alpha^{2^n}\rceil$, если m— целое число, большее 2, где $\alpha + \alpha^{-1} = m$ и $\alpha > 1$. Например, если $m = 3$, то решением является

$$X_n = \lceil\phi^{2^{n+1}}\rceil, \qquad \phi = \frac{1+\sqrt{5}}{2}, \qquad \alpha = \phi^2.$$

31 Докажите или опровергните, что $\lfloor x\rfloor + \lfloor y\rfloor + \lfloor x+y\rfloor \leqslant \lfloor 2x\rfloor + \lfloor 2y\rfloor$.

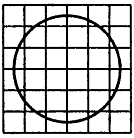
32 Обозначим расстояние от x до ближайшего целого как $\|x\| = \min(x - \lfloor x \rfloor, \lceil x \rceil - x)$. Чему равно значение

$$\sum_k 2^k \|x/2^k\|^2\,?$$

(Обратите внимание, что эта сумма может быть дважды бесконечной (в обе стороны). Например, при $x = 1/3$ члены ненулевые как при $k \to -\infty$, так и при $k \to +\infty$.)

Контрольные работы

33 На шахматной доске размером $2n \times 2n$ симметрично начерчена окружность диаметром $2n-1$; вот какой вид это имеет при $n = 3$:



а Через сколько клеток проходит такая окружность?

б Найдите функцию $f(n,k)$, такую, что внутри данной окружности полностью помещается $\sum_{k=1}^{n-1} f(n,k)$ клеток доски.

34 Пусть $f(n) = \sum_{k=1}^{n} \lceil \lg k \rceil$.

а Найдите аналитический вид $f(n)$ для $n \geqslant 1$.

б Докажите, что $f(n) = n - 1 + f(\lceil n/2 \rceil) + f(\lfloor n/2 \rfloor)$ для всех $n \geqslant 1$.

Упростите, но при этом не измените ее значение.

35 Упростите формулу $\lfloor (n+1)^2 n!\, e \rfloor \bmod n$.

36 В предположении, что n — неотрицательное целое число, найдите аналитическое выражение суммы

$$\sum_{1<k<2^{2^n}} \frac{1}{2^{\lfloor \lg k \rfloor} 4^{\lfloor \lg \lg k \rfloor}}\,.$$

37 Докажите тождество

$$\sum_{0 \leqslant k < m} \left(\left\lfloor \frac{m+k}{n} \right\rfloor - \left\lfloor \frac{k}{n} \right\rfloor \right) =$$
$$= \left\lfloor \frac{m^2}{n} \right\rfloor - \left\lfloor \frac{\min(m \bmod n, (-m) \bmod n)^2}{n} \right\rfloor$$

для всех положительных целых чисел m и n.

38 Пусть $x_1, \ldots, x_n$ — действительные числа, такие, что тождество

$$\sum_{k=1}^{n} \lfloor m x_k \rfloor = \Big\lfloor m \sum_{1\leqslant k\leqslant n} x_k \Big\rfloor$$

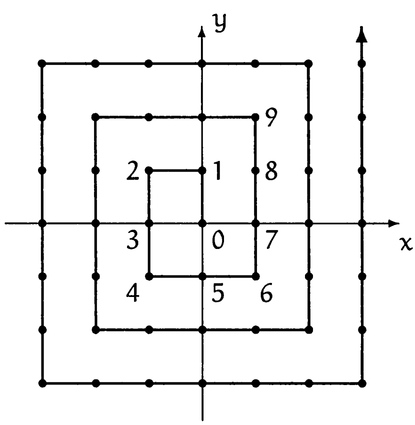
выполняется для всех положительных целых чисел m. Докажите какое-нибудь интересное свойство $x_1, \ldots, x_n$.

39 Докажите, что двойная сумма

$$\sum_{0\leqslant k\leqslant \log_b x} \; \sum_{0<j<b} \lceil (x + jb^k)/b^{k+1} \rceil$$

равна $(b-1)\big(\lfloor \log_b x \rfloor + 1\big) + \lceil x \rceil - 1$, для всех действительных чисел $x \geqslant 1$ при любом целом $b > 1$.

40 Спиральная функция $\sigma(n)$, схематически показанная ниже, отображает целое неотрицательное число n на упорядоченную пару целых чисел $\big(x(n), y(n)\big)$. Например, она отображает $n = 9$ на упорядоченную пару $(1, 2)$.



В южном полушарии люди используют спираль, закрученную в обратную сторону.

а Докажите, что если $m = \lfloor \sqrt{n} \rfloor$, то

$$x(n) = \\ = (-1)^m \Big(\big(n - m(m+1)\big) \cdot \big[\lfloor 2\sqrt{n} \rfloor \text{ четное}\big] + \lceil \tfrac{1}{2} m \rceil \Big),$$

и найдите аналогичную формулу для $y(n)$. *Указание*: разбейте спираль на сегменты W_k, S_k, E_k, N_k в зависимости от значения $\lfloor 2\sqrt{n} \rfloor = 4k-2$, $4k-1$, $4k$, $4k+1$.

б Докажите, что, наоборот, можно установить значение n из значения $\sigma(n)$ при помощи формулы вида

$$\begin{aligned} n &= (2k)^2 \pm \bigl(2k + x(n) + y(n)\bigr), \\ k &= \max\bigl(|x(n)|, |y(n)|\bigr). \end{aligned}$$

Укажите правило, когда вместо $\pm$ должен использоваться знак $+$, а когда — знак $-$.

Дополнительные задачи

41 Пусть f и g — возрастающие функции, обладающие тем свойством, что множества $\{f(1), f(2), \ldots\}$ и $\{g(1), g(2), \ldots\}$ образуют разбиение всех положительных целых чисел. Предположим, что f и g связаны условием $g(n) = f\bigl(f(n)\bigr) + 1$ при всех $n > 0$. Докажите, что $f(n) = \lfloor n\phi \rfloor$ и $g(n) = \lfloor n\phi^2 \rfloor$, где $\phi = (1 + \sqrt{5}\,)/2$.

42 Существуют ли действительные числа α, β и γ, такие, что $\mathrm{Spec}(\alpha)$, $\mathrm{Spec}(\beta)$ и $\mathrm{Spec}(\gamma)$ совместно образуют разбиение множества положительных целых чисел?

43 Найдите интересную интерпретацию чисел Кнута, разворачивая рекуррентное соотношение (3.16).

44 Покажите наличие целых чисел $a_n^{(q)}$ и $d_n^{(q)}$, таких, что

$$a_n^{(q)} = \frac{D_{n-1}^{(q)} + d_n^{(q)}}{q - 1} = \frac{D_n^{(q)} + d_n^{(q)}}{q} \qquad \text{при } n > 0,$$

где $D_n^{(q)}$ является решением рекуррентного соотношения (3.20). Воспользуйтесь этим фактом для решения обобщенной задачи Иосифа:

$$J_q(n) = 1 + d_k^{(q)} + q(n - a_k^{(q)}) \quad \text{для } a_k^{(q)} \leqslant n < a_{k+1}^{(q)}.$$

45 Примените прием из упр. 30 для нахождения решения рекуррентного соотношения

$$\begin{aligned} Y_0 &= m, \\ Y_n &= 2Y_{n-1}^2 - 1 \qquad \text{при } n > 0 \end{aligned}$$

в аналитическом виде, если m — положительное целое число.

46 Докажите, что если $n = \bigl\lfloor (\sqrt{2}^{\,l} + \sqrt{2}^{\,l-1})m \bigr\rfloor$, где m и l — неотрицательные целые числа, то $\bigl\lfloor \sqrt{2n(n+1)} \bigr\rfloor = \bigl\lfloor (\sqrt{2}^{\,l+1} + \sqrt{2}^{\,l})m \bigr\rfloor$. Используйте это замечательное свойство для поиска решения рекуррентного соотношения

$$\begin{aligned} L_0 &= a, && \text{целое } a > 0; \\ L_n &= \left\lfloor \sqrt{2L_{n-1}(L_{n-1}+1)} \right\rfloor && \text{для } n > 0 \end{aligned}$$

в аналитическом виде. *Указание*:

$$\left\lfloor \sqrt{2n(n+1)} \right\rfloor = \left\lfloor \sqrt{2}\left(n+\frac{1}{2}\right) \right\rfloor.$$

47 Говорят, что функция $f(x)$ *репликативна*, если она удовлетворяет условию

$$f(mx) = f(x) + f\left(x+\frac{1}{m}\right) + \cdots + f\left(x+\frac{m-1}{m}\right)$$

для любого положительного целого числа m. Найдите необходимые и достаточные условия, которым должно удовлетворять действительное число c, чтобы следующие функции были репликативны:

а $f(x) = x + c$,

б $f(x) = [x + c \text{ целое}]$,

в $f(x) = \max(\lfloor x \rfloor, c)$,

г $f(x) = x + c\lfloor x \rfloor - \frac{1}{2}[x \text{ не целое}]$.

48 Докажите тождество

$$x^3 = 3x\lfloor x\lfloor x \rfloor \rfloor + 3\{x\}\{x\lfloor x \rfloor\} + \{x\}^3 - 3\lfloor x \rfloor \lfloor x\lfloor x \rfloor \rfloor + \lfloor x \rfloor^3$$

и покажите, как можно получить подобные формулы для x^n при $n > 3$.

49 Найдите необходимое и достаточное условие, которому должны удовлетворять действительные числа $0 \leqslant \alpha < 1$ и $\beta \geqslant 0$, так чтобы можно было определить α и β из бесконечного мультимножества значений

$$\{\, \lfloor n\alpha \rfloor + \lfloor n\beta \rfloor \mid n > 0 \,\}.$$

Исследовательские проблемы

50 Найдите необходимое и достаточное условие, которому должны удовлетворять действительные числа α и β, чтобы α и β можно было определить из бесконечного мультимножества величин

$$\{\, \lfloor \lfloor n\alpha \rfloor \beta \rfloor \mid n > 0 \,\}.$$

51 Пусть x — действительное число $\geqslant \phi = \frac{1}{2}(1+\sqrt{5})$. Тогда решение рекуррентного соотношения

$$\begin{aligned} Z_0(x) &= x, \\ Z_n(x) &= Z_{n-1}(x)^2 - 1 \quad \text{при } n > 0, \end{aligned}$$

может быть записано в виде $Z_n(x) = \lceil f(x)^{2^n} \rceil$, если x — целое число, где

$$f(x) = \lim_{n\to\infty} Z_n(x)^{1/2^n},$$

так как в этом случае $Z_n(x) - 1 < f(x)^{2^n} < Z_n(x)$. Какими другими интересными свойствами обладает функция $f(x)$?

52 Для заданных неотрицательных действительных чисел α и β положим

$$\mathrm{Spec}(\alpha;\beta) = \{\lfloor\alpha+\beta\rfloor, \lfloor 2\alpha+\beta\rfloor, \lfloor 3\alpha+\beta\rfloor, \ldots\}.$$

Специфицировать Spec — специфичная задача.

$\mathrm{Spec}(\alpha;\beta)$ — мультимножество, которое обобщает мультимножество $\mathrm{Spec}(\alpha) = \mathrm{Spec}(\alpha;0)$. Докажите или опровергните, что если $m \geqslant 3$ мультимножеств $\mathrm{Spec}(\alpha_1;\beta_1)$, $\mathrm{Spec}(\alpha_2;\beta_2)$, …, $\mathrm{Spec}(\alpha_m;\beta_m)$ образуют разбиение всех положительных целых чисел и при этом параметры $\alpha_1 < \alpha_2 < \cdots < \alpha_m$ представляют собой действительные числа, то

$$\alpha_k = \frac{2^m - 1}{2^{k-1}} \qquad \text{при } 1 \leqslant k \leqslant m.$$

53 Алгоритм Фибоначчи (см. упр. 9) является “жадным” в том смысле, что на каждом шаге он выбирает как можно более малое q. Известен более сложный алгоритм, с помощью которого любая дробь m/n с нечетным n может быть представлена как сумма различных базовых дробей $1/q_1 + \cdots + 1/q_k$ с *нечетными* знаменателями. Всегда ли заканчивается жадный алгоритм при поиске такого представления?

4

Теория чисел

ЦЕЛЫЕ ЧИСЛА ЗАНИМАЮТ ЦЕНТРАЛЬНОЕ МЕСТО в дискретной математике, на которую делается особый упор в данной книге. Поэтому мы хотим рассмотреть *теорию чисел* — важный раздел математики, работающий со свойствами целых чисел.

Мы немного заплывали в воды теории чисел в предыдущей главе, когда имели дело с бинарными операциями 'mod' и 'НОД'. Теперь пришло время окунуться в них с головой...

Другими словами, "не тратьте, куме, сили, спускайтеся на дно!"

4.1 Делимость

Говорят, что m делит n (или n делится на m), если $m > 0$ и отношение n/m представляет собой целое число. Это свойство лежит в основе всей теории чисел, так что сто́ит обзавестись для него специальным обозначением. Поэтому мы будем записывать

$$m\backslash n \iff m > 0 \quad \text{и } n = mk \text{ при некотором целом } k\,. \qquad (4.1)$$

(В настоящее время в математической литературе более распространено обозначение '$m|n$', а не '$m\backslash n$'. Но вертикальная черта и так употребляется без меры — и для абсолютного значения, и для разделителей множеств, и для условных вероятностей и т.д., в то время как обратная косая черта явно недооценена. Помимо прочего, '$m\backslash n$' создает впечатление, что m представляет собой знаменатель некоторого подразумеваемого отношения. Так что наклоним символ делимости немного влево.)

Если m не делит n, мы пишем '$m\nmid n$'.

Имеется аналогичное отношение "n кратно m", которое означает почти то же самое, за исключением того, что m не обязано быть положительным. Это просто означает, что $n = mk$ для некоторого целого k. Так, например, существует только одно число, кратное 0 (а именно — 0), но ничто не делится на 0. Любое целое число кратно -1, но, строго говоря, ни одно целое число не делится на -1. Эти определения применимы, когда m и n — произвольные действительные числа; например, 2π делится на π. Но мы почти всегда будем использовать их, когда m и n представляют собой целые числа. В конце концов, мы занимаемся теорией чисел.

Наибольший общий делитель двух целых чисел m и n — наибольшее целое, которое делит их оба:

$$\text{НОД}(m,n) \;=\; \max\{\,k \mid k\backslash m \quad\text{и}\quad k\backslash n\,\}. \tag{4.2}$$

В Великобритании он называется 'hcf' (highest common factor), в США — 'gcd' (greatest common divisor), в отечественной литературе — 'НОД'.

Например, $\text{НОД}(12,18) = 6$. Это хорошо знакомое всем понятие, поскольку оно представляет собой не что иное, как общий множитель, который еще в младших классах учат выделять из дроби m/n, приводя ее к несократимому виду: $12/18 = (12/6)/(18/6) = 2/3$. Обратите внимание, что если $n > 0$, то мы имеем $\text{НОД}(0,n) = n$, так как любое положительное число делит 0, а n является наибольшим делителем самого себя. Значение $\text{НОД}(0,0)$ не определено.

Еще одно хорошо знакомое понятие — *наименьшее общее кратное*

Не пытайтесь искать наибольшее общее кратное.

$$\text{НОК}(m,n) \;=\; \min\{\,k \mid k>0, \quad m\backslash k \quad\text{и}\quad n\backslash k\,\}; \tag{4.3}$$

при $m \leqslant 0$ или $n \leqslant 0$ оно не определено. Изучающие арифметику знают эту величину как наименьший общий знаменатель, к которому приводятся при сложении дроби со знаменателями m и n. Например, $\text{НОК}(12,18) = 36$, так что даже ученики начальных классов знают, что $\frac{7}{12} + \frac{1}{18} = \frac{21}{36} + \frac{2}{36} = \frac{23}{36}$. Понятие НОК в чем-то похоже на понятие НОД, но мы не будем уделять ему такое же внимание, как и НОД, так как последний обладает рядом красивых свойств.

Одним из наиболее привлекательных свойств НОД является простота его вычисления при помощи известного более 2300 лет метода под названием *алгоритм Евклида*. Для вычисления $\text{НОД}(m,n)$ для заданных значений $0 \leqslant m < n$ в алгоритме Евклида используется рекуррентность

$$\begin{aligned} \text{НОД}(0,n) &= n\,; \\ \text{НОД}(m,n) &= \text{НОД}(n \bmod m,\, m) \qquad \text{при } m > 0. \end{aligned} \tag{4.4}$$

Так, например, $\text{НОД}(12,18) = \text{НОД}(6,12) = \text{НОД}(0,6) = 6$. Указанная рекуррентность корректна, поскольку любой общий делитель m и n должен быть общим делителем как m, так и $n \bmod m$, которое представляет собой $n - \lfloor n/m \rfloor m$. Вряд ли имеется иная рекуррентность для $\text{НОК}(m,n)$, которая хотя бы отчасти была столь же простой, как эта. (См. упр. 2.)

Алгоритм Евклида дает нам больше, чем просто НОД: его можно расширить так, что он будет вычислять целые числа m' и n', удовлетворяющие уравнению

$$m'm + n'n = \text{НОД}(m, n). \tag{4.5}$$

(Не забывайте, что m' и n' могут быть отрицательными.)

Вот как это делается. Если $m = 0$, мы просто выбираем $m' = 0$ и $n' = 1$. В противном случае мы полагаем $r = n \bmod m$ и применяем метод рекурсивно — подставляя r и m вместо m и n и вычисляя $\bar{r}$ и $\bar{m}$, такие, что

$$\bar{r}r + \bar{m}m = \text{НОД}(r, m).$$

Поскольку $r = n - \lfloor n/m \rfloor m$ и $\text{НОД}(r, m) = \text{НОД}(m, n)$, это уравнение гласит, что

$$\bar{r}\left(n - \lfloor n/m \rfloor m\right) + \bar{m}m = \text{НОД}(m, n).$$

Левую часть можно переписать, чтобы показать ее зависимость от m и n:

$$\left(\bar{m} - \lfloor n/m \rfloor \bar{r}\right) m + \bar{r}n = \text{НОД}(m, n);$$

следовательно, $m' = \bar{m} - \lfloor n/m \rfloor \bar{r}$ и $n' = \bar{r}$ — те целые числа, которые нужны в (4.5). Например, в нашем любимом случае $m = 12$, $n = 18$ этот метод дает $6 = 0 \cdot 0 + 1 \cdot 6 = 1 \cdot 6 + 0 \cdot 12 = (-1) \cdot 12 + 1 \cdot 18$.

Но чем же так привлекательно уравнение (4.5)? Главным образом тем, что числа m' и n' фактически *доказывают*, что алгоритм Евклида дает правильный результат в каждом конкретном случае. Предположим, компьютер сообщил нам, что $\text{НОД}(m, n) = d$ и что $m'm + n'n = d$, но мы относимся к этому с изрядной долей скепсиса и полагаем, что на самом деле имеется больший общий делитель, пропущенный машиной. Однако этого не может быть, так как любой общий делитель m и n должен делить $m'm + n'n$, а раз так, он должен делить и d; следовательно, он должен быть $\leqslant d$. Кроме того, легко убедиться, что d делит как m, так и n. (Алгоритмы, предоставляющие собственные доказательства верности результата, называются *самоподтверждающими* (*self-certifying*).)

Мы будем много раз использовать (4.5) в оставшейся части этой главы. Одним из важных следствий этого уравнения является следующая мини-теорема:

$$k \backslash m \quad \text{и} \quad k \backslash n \qquad \iff \qquad k \backslash \text{НОД}(m, n). \tag{4.6}$$

(Доказательство: если k делит и m, и n, то оно делит и $m'm + n'n$, а значит, оно делит и НОД(m,n). Обратно, если k делит НОД(m,n), то оно делит делитель m и делитель n, так что оно делит и m, и n.) Мы всегда знали, что всякий общий делитель чисел m и n должен быть *меньше или равен* их НОД — по определению наибольшего общего делителя. Теперь мы знаем, что любой общий делитель является *делителем* их НОД.

Иногда требуется просуммировать по всем делителям числа n. В этом случае зачастую полезно воспользоваться удобным правилом

$$\sum_{m\backslash n} a_m = \sum_{m\backslash n} a_{n/m}\,, \qquad \text{целое } n > 0, \tag{4.7}$$

которое выполняется, потому что n/m пробегает те же делители числа n, что и m. Например, при $n = 12$ это правило гласит, что $a_1 + a_2 + a_3 + a_4 + a_6 + a_{12} = a_{12} + a_6 + a_4 + a_3 + a_2 + a_1$.

Имеется также несколько более общее равенство

$$\sum_{m\backslash n} a_m = \sum_{k} \sum_{m>0} a_m [n = mk]\,, \tag{4.8}$$

которое является непосредственным следствием определения (4.1). Если n — положительное число, то правая часть (4.8) представляет собой $\sum_{k\backslash n} a_{n/k}$; следовательно, из (4.8) вытекает (4.7). Равенство (4.8) работает и при отрицательных n. (В таких случаях ненулевые члены справа появляются только тогда, когда k является отрицательным делителем n.)

Кроме того, двойная сумма по делителям может быть "переставлена" по правилу

$$\sum_{m\backslash n} \sum_{k\backslash m} a_{k,m} = \sum_{k\backslash n} \sum_{l\backslash(n/k)} a_{k,kl}\,. \tag{4.9}$$

Например, это правило при $n = 12$ приобретает следующий вид:

$$\begin{aligned} & a_{1,1} + (a_{1,2} + a_{2,2}) + (a_{1,3} + a_{3,3}) + \\ &\qquad + (a_{1,4} + a_{2,4} + a_{4,4}) + (a_{1,6} + a_{2,6} + a_{3,6} + a_{6,6}) + \\ &\qquad + (a_{1,12} + a_{2,12} + a_{3,12} + a_{4,12} + a_{6,12} + a_{12,12}) = \\ &= (a_{1,1} + a_{1,2} + a_{1,3} + a_{1,4} + a_{1,6} + a_{1,12}) + \\ &\qquad + (a_{2,2} + a_{2,4} + a_{2,6} + a_{2,12}) + (a_{3,3} + a_{3,6} + a_{3,12}) + \\ &\qquad + (a_{4,4} + a_{4,12}) + (a_{6,6} + a_{6,12}) + a_{12,12}\,. \end{aligned}$$

Можно доказать (4.9) при помощи обозначения Айверсона.

Левая сторона равенства представляет собой

$$\sum_{j,l}\sum_{k,m>0} a_{k,m}[n=jm][m=kl] = \sum_{j}\sum_{k,l>0} a_{k,kl}[n=jkl],$$

а правая —

$$\sum_{j,m}\sum_{k,l>0} a_{k,kl}[n=jk][n/k=ml] = \sum_{m}\sum_{k,l>0} a_{k,kl}[n=mlk],$$

что то же самое, если не считать переименования индексов. Этот пример указывает, что изученные в главе 2 методы пригодны и при изучении теории чисел.

4.2 Простые числа

Положительное целое число p называется *простым*, если оно имеет только два делителя, а именно — 1 и p. *Везде до конца данной главы буква* p *будет обозначать простое число, даже если это не оговаривается явно.* По соглашению 1 не является простым, так что последовательность простых чисел начинается так:

А 'p' в слове 'простое'?

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, \ldots .$$

Некоторые числа кажутся простыми, но таковыми не являются, наподобие 91 $(= 7\cdot 13)$ и 161 $(= 7\cdot 23)$. Эти и другие числа, которые имеют нетривиальные делители, называются *составными*. Каждое целое число, большее 1, является либо простым, либо составным, но не тем и другим одновременно.

Простые числа очень важны, так как они представляют собой "кирпичи", из которых строятся все положительные целые числа. Любое положительное целое число n может быть записано как произведение простых чисел:

$$n = p_1 \ldots p_m = \prod_{k=1}^{m} p_k, \qquad p_1 \leqslant \cdots \leqslant p_m. \tag{4.10}$$

Например, $12 = 2\cdot 2\cdot 3$; $11011 = 7\cdot 11\cdot 11\cdot 13$; $11111 = 41\cdot 271$. (Произведения, обозначаемые при помощи $\prod$, аналогичны суммам, обозначаемым при помощи $\sum$, в том смысле, как это поясняется в упр. 2.25. Если $m = 0$, то такое произведение считается пустым, а его значение равным 1 по определению; именно таким образом число $n = 1$ оказывается представленным при помощи (4.10).) Такое разложение на множители всегда возможно, так как если $n > 1$ — не простое число, то оно имеет делитель n_1, такой, что $1 < n_1 < n$; таким образом, можно записать $n = n_1 \cdot n_2$, а (по индукции) мы знаем, что n_1 и n_2 могут быть записаны как произведения простых чисел.

Кроме того, разложение (4.10) *единственно*: возможен только один способ записи n в виде произведения простых чисел в неубывающем порядке. Это утверждение называется ОСНОВНОЙ ТЕОРЕМОЙ АРИФМЕТИКИ, и оно представляется настолько очевидным, что возникают сомнения — а надо ли его доказывать? Как два разных множества простых чисел могут дать одно и то же произведение? Да, не могут, но причина — *не просто* "по определению простых чисел". Например, если рассмотреть множество всех действительных чисел вида $m + n\sqrt{10}$, где m и n — целые числа, то произведение любых двух таких чисел будет числом того же вида, и можно назвать число "простым", если оно не может быть разложено на множители нетривиальным способом. Число 6 имеет два представления, $2 \cdot 3 = (4 + \sqrt{10}\,)(4 - \sqrt{10}\,)$, и тем не менее в упр. 36 показано, что 2, 3, $4+\sqrt{10}$ и $4 - \sqrt{10}$ являются в этой системе "простыми".

Следовательно, требуется строго доказать, что разложение (4.10) единственное. Определенно, имеется только один способ при $n = 1$, поскольку в этом случае произведение должно быть пустым. Так что предположим, что $n > 1$ и что все меньшие числа разлагаются на множители единственным образом. Предположим, что у нас есть разложения

$$n = p_1 \ldots p_m = q_1 \ldots q_k, \qquad p_1 \leqslant \cdots \leqslant p_m \quad \text{и} \quad q_1 \leqslant \cdots \leqslant q_k,$$

где все числа p и q — простые. Докажем, что $p_1 = q_1$. Если это не так, то можно считать, что $p_1 < q_1$, т.е. p_1 меньше любого из q. Так как p_1 и q_1 простые, их НОД должен быть равен 1; следовательно, самоподтверждающий алгоритм Евклида дает нам целые числа a и b, такие, что $ap_1 + bq_1 = 1$. Тогда

$$ap_1 q_2 \ldots q_k + bq_1 q_2 \ldots q_k = q_2 \ldots q_k.$$

Итак, p_1 делит оба члена в левой части, поскольку $q_1 q_2 \ldots q_k = n$; следовательно, p_1 делит правую часть $q_2 \ldots q_k$. Тогда число $q_2 \ldots q_k / p_1$ целое, а $q_2 \ldots q_k$ — разложение на простые множители, в котором присутствует p_1. Но $q_2 \ldots q_k < n$, так что оно допускает единственное разложение на множители (по индукции). Это противоречие показывает, что в конце концов p_1 должно быть равно q_1. Таким образом, можно разделить оба разложения n на p_1, получая $p_2 \ldots p_m = q_2 \ldots q_k < n$. По индукции должны быть равны и остальные множители, так что наше доказательство единственности разложения завершено.

Единственное — разложение, а не теорема.

Иногда более полезно сформулировать основную теорему иначе: *любое положительное целое число может быть единственным образом записано в виде*

$$n = \prod_p p^{n_p}, \quad \text{где каждое } n_p \geqslant 0. \tag{4.11}$$

Правая часть представляет собой произведение по бесконечно многим простым числам; но для конкретного n все показатели степени, за исключением небольшого количества, равны нулю, так что соответствующие множители равны 1. Следовательно, на самом деле это конечное произведение, так же как многие "бесконечные" суммы на самом деле являются конечными, поскольку в основном их члены нулевые.

Формула (4.11) однозначно представляет n, так что последовательность $\langle n_2, n_3, n_5, \ldots \rangle$ можно рассматривать как *систему счисления* положительных целых чисел. Например, представление в виде степеней простых чисел числа 12 имеет вид $\langle 2, 1, 0, 0, \ldots \rangle$, а представление в виде степеней простых чисел числа 18 — $\langle 1, 2, 0, 0, \ldots \rangle$. Чтобы перемножить два числа, надо просто сложить их представления. Другими словами,

$$k = mn \qquad \iff \qquad k_p = m_p + n_p \quad \text{при всех } p. \tag{4.12}$$

Отсюда вытекает, что

$$m \backslash n \qquad \iff \qquad m_p \leqslant n_p \quad \text{при всех } p, \tag{4.13}$$

а отсюда непосредственно следует, что

$$k = \text{НОД}(m, n) \iff k_p = \min(m_p, n_p) \quad \text{при всех } p, \tag{4.14}$$
$$k = \text{НОК}(m, n) \iff k_p = \max(m_p, n_p) \quad \text{при всех } p. \tag{4.15}$$

Например, поскольку $12 = 2^2 \cdot 3^1$ и $18 = 2^1 \cdot 3^2$, можно получить их НОД и НОК, беря минимумы и максимумы соответствующих степеней:

$$\text{НОД}(12, 18) = 2^{\min(2,1)} \cdot 3^{\min(1,2)} = 2^1 \cdot 3^1 = 6,$$
$$\text{НОК}(12, 18) = 2^{\max(2,1)} \cdot 3^{\max(1,2)} = 2^2 \cdot 3^2 = 36.$$

Если простое число p делит произведение mn, то оно делит либо число m, либо n, а возможно, и оба в силу теоремы о единственности разложения на простые множители. Но составные числа таким свойством не обладают. Например, составное число 4 делит $60 = 6 \cdot 10$, но не делит ни 6, ни 10. Причина проста: в разложении $60 = 6 \cdot 10 = (2 \cdot 3)(2 \cdot 5)$ два простых сомножителя числа $4 = 2 \cdot 2$ оказываются разделенными по разным частям, так что 4 не делит ни одну из них. Но простое число неразделимо, так что оно должно делить один из исходных сомножителей.

4.3 Простые примеры простых чисел

Сколько всего простых чисел? Много. Очень много. Бесконечно много, как давным-давно доказал Евклид в своей теореме 9:20. И вот как он это сделал. Предположим, что имеется конечное количество простых чисел, скажем, k чисел 2, 3, 5, ..., P_k. Тогда, как говорит Евклид, мы должны рассмотреть число

$$M = 2 \cdot 3 \cdot 5 \cdot \ldots \cdot P_k + 1.$$

Ни одно из k простых чисел не может делить M, поскольку каждое из них делит $M - 1$. Таким образом, должно иметься некоторое другое простое число, которое делит M; возможно, число M само является простым. Это противоречит нашему предположению о том, что только 2, 3, ..., P_k являются простыми числами, так что на самом деле их должно быть бесконечно много.

"Οἱ πρῶτοι ἀριθμοὶ πλείους εἰσὶ παντὸς τοῦ προτεθέντος πλήθους πρώτων ἀριθ μῶν."
— Евклид [98]
(Перевод: "Простых чисел больше, чем в любом заданном множестве простых чисел.")

Доказательство Евклида наводит на мысль об определении *числа Евклида* при помощи рекуррентного соотношения

$$e_n = e_1 e_2 \ldots e_{n-1} + 1 \qquad \text{при } n \geqslant 1. \tag{4.16}$$

Последовательность этих чисел начинается следующим образом:

$$\begin{aligned} e_1 &= 1 + 1 = 2, \\ e_2 &= 2 + 1 = 3, \\ e_3 &= 2 \cdot 3 + 1 = 7, \\ e_4 &= 2 \cdot 3 \cdot 7 + 1 = 43, \end{aligned}$$

причем все приведенные числа — простые. Но уже следующее число, e_5, представляет собой $1807 = 13 \cdot 139$. Далее, число $e_6 = 3263443$ простое, в то время как

$$\begin{aligned} e_7 &= 547 \cdot 607 \cdot 1033 \cdot 31051, \\ e_8 &= 29881 \cdot 67003 \cdot 9119521 \cdot 6212157481. \end{aligned}$$

Известно, что e_9, ..., e_{17} — составные числа, так же, как, вероятно, и все остальные числа e_n. Однако числа Евклида являются *взаимно простыми* друг по отношению к другу, т.е.

$$\text{НОД}(e_m, e_n) = 1, \qquad \text{когда } m \neq n.$$

Алгоритм Евклида (какой же еще?) подтверждает это за три коротких шага: поскольку $e_n \bmod e_m = 1$ при $n > m$, то

$$\text{НОД}(e_m, e_n) = \text{НОД}(1, e_m) = \text{НОД}(0, 1) = 1.$$

Следовательно, если положить q_j равным наименьшему множителю e_j для каждого $j \geqslant 1$, то все простые числа $q_1, q_2, q_3, \ldots$ будут различны — они образуют бесконечную последовательность простых чисел.

Давайте сделаем небольшое отступление и рассмотрим числа Евклида с точки зрения главы 1. Нельзя ли записать e_n в аналитическом виде? Рекуррентное соотношение (4.16) можно упростить, удалив троеточие: если $n > 1$, то мы имеем

$$\begin{aligned} e_n &= e_1 \ldots e_{n-2}e_{n-1} + 1 = \\ &= (e_{n-1} - 1)e_{n-1} + 1 = e_{n-1}^2 - e_{n-1} + 1. \end{aligned}$$

Таким образом, в e_n примерно вдвое больше десятичных цифр, чем в e_{n-1}. В упр. 37 доказывается существование константы $E \approx 1.264$, такой, что

$$e_n = \left\lfloor E^{2^n} + \tfrac{1}{2} \right\rfloor. \tag{4.17}$$

А в упр. 60 доказывается подобная формула,

$$p_n = \left\lfloor P^{3^n} \right\rfloor, \tag{4.18}$$

которая дает только простые числа при некоторой константе P. Однако уравнения наподобие (4.17) и (4.18) не могут в действительности рассматриваться как аналитическая запись, поскольку константы E и P вычисляются по числам e_n и p_n, образуя замкнутый круг. Неизвестно (и его существование маловероятно) соотношение, которое бы связывало эти константы с другими независимыми математическими значениями.

На самом деле никто не знает *никакой* полезной формулы, которая давала бы произвольно большие простые и только простые числа.* Однако несмотря на это, кибернетики из компании Chevron Geosciences в 1984 году вместо нефтяного открыли

---

* Однако *бесполезные* формулы существуют. Так, например, в книге Г. Уоррен, *Алгоритмические трюки для программистов.*— М.: Издательский дом "Вильямс", 2003, приводится несколько формул для n-го простого числа наподобие $p_n = 1 + \sum_{m=1}^{2^n} \left\lfloor \sqrt[n]{n} \left(\sum_{x=1}^{m} \left\lfloor \cos^2 \pi \frac{(x-1)!+1}{x} \right\rfloor \right)^{-1/n} \right\rfloor$.— *Переводчик.*

"математический источник". Используя программу Дэвида Словински (D. Slowinski), они при тестировании нового суперкомпьютера Cray X-MP обнаружили наибольшее известное на тот момент времени простое число

$$2^{216091} - 1\,.$$

Это число легко вычислить на персональном компьютере за несколько миллисекунд, поскольку современные компьютеры работают в двоичной системе счисления, а это число — просто $(11\ldots1)_2$. Все 216 091 его битов равны 1. Но гораздо труднее доказать его простоту. Практически любая операция с этим числом выполняется очень долго из-за его размера. Например, даже интеллектуальный алгоритм требует на персональном компьютере нескольких минут только для перевода $2^{216091} - 1$ в десятичную запись* Пересылка распечатки такого числа, состоящего из 65 050 цифр, требует 78 центов.

Кстати, $2^{216091} - 1$ — это количество переносов, необходимых для решения задачи о Ханойской башне из 216 091 диска. Числа вида

Ко времени, когда вы это прочтете, — наверняка еще больше.

$$2^p - 1$$

(где p, как и везде в этой главе, — простое число) называются *числами Мерсенна*, в честь преподобного Марина Мерсенна (Marin Mersenne), который изучал некоторые их свойства в XVII веке [269]. Известные к настоящему времени простые числа Мерсенна получаются при $p = 2, 3, 5, 7, 13, 17, 19, 31, 61$, 89, 107, 127, 521, 607, 1 279, 2 203, 2 281, 3 217, 4 253, 4 423, 9 689, 9 941, 11 213, 19 937, 21 701, 23 209, 44 497, 86 243, 110 503, 132 049, 216 091, 756 839, 859 433, 1 257 787, 1 398 269, 2 976 221, 3 021 377, 6 972 593, 13 466 917, 20 996 011, 24 036 583, 25 964 951, 30 402 457, 32 582 657, 37 156 667 и 43 112 609**

---

* Это во многом зависит от представления числа в компьютере. Так, при использовании библиотеки для работы с большими числами MAPM Майкла Ринга (Michael C. Ring) на компьютере с процессором Pentium 4 с тактовой частотой 3 ГГц для вычисления этого числа требуется около 15 мс, а для его превращения в десятичную запись — менее миллисекунды (для наибольшего известного на момент перевода книги простого числа $2^{43112609} - 1$ соответствующие значения равны соответственно около 15 с и 100 мс). — *Переводчик.*

** (В оригинальном издании книги были приведены известные простые числа Мерсенна по состоянию на начало 1998 года. На

Число $2^n - 1$ не может быть простым при составном n, так как $2^{km} - 1$ имеет одним из сомножителей $2^m - 1$:

$$2^{km} - 1 = (2^m - 1)(2^{m(k-1)} + 2^{m(k-2)} + \cdots + 1).$$

Однако $2^p - 1$ — не всегда простое при простом p; так, $2^{11} - 1 = 2047 = 23 \cdot 89$ — первое из составных чисел Мерсенна (и Мерсенн знал о нем).

В настоящее время разложение на множители и проверка простоты больших чисел — тема многочисленных исследовательских работ. Сводка того, что было известно об этом вплоть до 1981 года, приведена в разделе 4.5.4 книги [208], и с того времени появилось много новых результатов. На с. 391–394 этой же книги (с. 459–461 русского перевода) поясняется метод проверки чисел Мерсенна на простоту.

На протяжении последних пяти столетий наибольшим известным простым числом оказывалось, по большей части, простое число Мерсенна, хотя известно лишь несколько десятков этих чисел. Многие пытаются найти большие простые числа, но это не так просто. Те, кто хотят оказаться в *Книге рекордов Гиннесса* заслуженно, а не просто из-за везения, могут попытать счастья и поискать простые числа вида $2^n k + 1$ при малых значениях k, например, 3 или 5. Проверка простоты этих чисел почти такая же быстрая, как и чисел Мерсенна (детали см. в упр. 4.5.4–27 в книге [208]).

Но мы пока что так и не ответили полностью на первоначальный вопрос о том, сколько же всего простых чисел. Их бесконечно много, но одни бесконечные множества "плотнее" других. Например, среди положительных целых чисел имеется бесконечно много четных чисел и бесконечно много полных квадратов, но при этом в определенных важных смыслах четных чисел больше,

---

момент перевода этой книги известно 46 простых чисел Мерсенна; последние два открыты в 2008 году, причем проведена исчерпывающая проверка до 39-го числа Мерсенна; так что возможно наличие других простых чисел Мерсенна, кроме известных, между $2^{13466917} - 1$ и $2^{43112609} - 1$ (эвристические оценки показывают, что в интервале для p от 10 000 000 до 80 000 000 ждут своего открытия еще два неизвестных простых числа Мерсенна). Последние 12 простых чисел Мерсенна найдены в рамках проекта GIMPS (Great Internet Mersenne Prime Search) — широкомасштабного Интернет-проекта распределенных вычислений по поиску простых чисел Мерсенна (`http://www.mersenne.org/`). — *Переводчик.*)

чем полных квадратов. Один из таких "смыслов" — сравнивать n-е значения: n-е четное число равно $2n$, а n-й полный квадрат равен n^2; так как $2n$ гораздо меньше n^2 при больших n, n-е четное целое число встречается гораздо раньше n-го полного квадрата, так что можно сказать, что имеется гораздо больше четных целых чисел, чем полных квадратов. Похожая точка зрения заключается в том, чтобы сравнить количество искомых величин, не превосходящих x. Имеется $\lfloor x/2 \rfloor$ таких целых четных чисел и $\lfloor \sqrt{x} \rfloor$ таких полных квадратов; поскольку $x/2$ гораздо больше, чем $\sqrt{x}$ при больших x, снова можно говорить о том, что целых четных чисел гораздо больше, чем полных квадратов.

Странно. Поскольку существует взаимно однозначное соответствие четных чисел и квадратов, я думал, что и тех, и других — одинаковое количество...

Что же можно сказать о простых числах в этих двух смыслах? Оказывается, что n-е простое число P_n примерно в n раз больше натурального логарифма n:

$$P_n \sim n \ln n .$$

(Символ '$\sim$' можно читать как "асимптотически равно", что означает, что предел отношения $P_n / n \ln n$ равен 1 при n, стремящемся к бесконечности.) Аналогично для количества простых чисел $\pi(x)$, не превышающих x, имеется "теорема о распределении простых чисел":

$$\pi(x) \sim \frac{x}{\ln x} .$$

Доказательство этих двух фактов выходит за рамки нашей книги, хотя можно легко показать, что каждый из них влечет за собой другой. В главе 9 мы рассмотрим скорости роста данных функций и увидим, что функция $n \ln n$, наше приближение для P_n, асимптотически лежит между $2n$ и n^2. Следовательно, простых чисел меньше, чем четных, но больше, чем полных квадратов.

Эти формулы, справедливые только в пределе при n или $x \to \infty$, можно заменить более точными оценками. Так, Россер (Rosser) и Шёнфельдь (Schoenfeld) [312] установили следующие удобные границы

$$\ln x - \tfrac{3}{2} < \tfrac{x}{\pi(x)} < \ln x - \tfrac{1}{2} \quad \text{при } x \geqslant 67; \tag{4.19}$$

$$\begin{aligned} n\bigl(\ln n + \ln\ln n - \tfrac{3}{2}\bigr) &< P_n < \\ &< n\bigl(\ln n + \ln\ln n - \tfrac{1}{2}\bigr) \quad \text{при } n \geqslant 20. \end{aligned} \tag{4.20}$$

Если взять "случайное" целое число n, то шансы, что оно окажется простым, составляют примерно один к $\ln n$. Например, если выбирать числа вблизи 10^{16}, то придется проверить примерно $16 \ln 10 \approx 36.8$ из них, прежде чем будет найдено простое число. (Оказывается, что между $10^{16} - 370$ и $10^{16} - 1$ имеется ровно 10 простых чисел.) Тем не менее в распределении простых чисел много нерегулярностей. Например, все числа между $P_1P_2 \ldots P_n+2$ и $P_1P_2 \ldots P_n+P_{n+1}-1$ включительно — составные. Известно много примеров "простых близнецов" p и $p+2$ (5 и 7, 11 и 13, 17 и 19, 29 и 31, ..., 9999999999999641 и 9999999999999643, ...), но никому не известно, бесконечно ли много пар таких близнецов. (См. Харди (Hardy) и Райт (Wright) [181, §1.4 и §2.8].)

Один из простейших способов вычислить все $\pi(x)$ простых чисел $\leqslant x$ — так называемое "решето Эратосфена". Сначала выписываются все целые числа от 2 до x. Затем обводится кружок вокруг числа 2, помечающий его как простое число, и вычеркиваются все числа, кратные 2. Затем последовательно обводятся наименьшие из невычеркнутых и необведенных чисел и вычеркиваются все кратные им. Когда останутся только обведенные или вычеркнутые числа, все обведенные числа — суть простые. Например, при $x = 10$ выписываются числа от 2 до 10, обводится 2, затем вычеркиваются кратные 2 числа 4, 6, 8 и 10. Очередное наименьшее невычеркнутое необведенное число — 3, так что оно обводится, и вычеркиваются кратные ему числа 6 и 9. Теперь наименьшим необведенным становится 5, так что мы обводим его и вычеркиваем 10. И наконец мы обводим 7. В результате обведенными оказываются числа 2, 3, 5 и 7, так что имеется $\pi(10) = 4$ простых числа, не превосходящих 10.

"Je me sers de la notation très simple $n!$ pour désigner le produit de nombres décroissans depuis n jusqu'à l'unité, savoir $n(n-1)$ $(n-2)\ldots.$ 3. 2. 1. L'emploi continuel de l'analyse combinatoire que je fais dans la plupart de mes démonstrations, a rendu cette notation indispensable".
— Х. Крамп (Ch. Kramp) [228]

4.4 Факториальные факты

Давайте теперь посмотрим, как выглядит разложение на простые множители самых составных чисел — факториалов:

$$n! = 1\cdot 2\cdot \ldots \cdot n = \prod_{k=1}^{n} k, \qquad \text{целое } n \geqslant 0. \tag{4.21}$$

В соответствии с нашим соглашением о пустом произведении факториал $0!$ определяется как равный 1. Таким образом, для каждого положительного целого числа n справедливо соотношение $n! = (n-1)!\,n$. Это число равно количеству перестановок n различных объектов, т.е. количеству способов разместить в ряд n предметов: первый предмет можно разместить n способами; для каждого размещения первого предмета второй предмет можно разместить $n-1$ способом; для каждого из $n(n-1)$ вариантов размещения первых двух предметов третий можно разместить $n-2$ способами; и т.д., что дает нам $n(n-1)(n-2)\ldots(1)$ вариантов размещения в целом. Вот несколько первых значений функции "факториал":

n	0	1	2	3	4	5	6	7	8	9	10
$n!$	1	1	2	6	24	120	720	5040	40320	362880	3628800

Полезно знать несколько фактов о факториалах — например, первые шесть или около того значений, что $10!$ равно $3\frac{1}{2}$ миллиона с небольшим, или что количество цифр в $n!$ превышает n при $n \geqslant 25$.

Можно доказать, что $n!$ очень велико, воспользовавшись уловкой наподобие гауссовой из главы 1:

$$n!^2 = (1\cdot 2\cdot \ldots \cdot n)(n\cdot \ldots \cdot 2\cdot 1) = \prod_{k=1}^{n} k(n+1-k).$$

Имеем $n \leqslant k(n+1-k) \leqslant \frac{1}{4}(n+1)^2$, поскольку квадратный полином $k(n+1-k) = \frac{1}{4}(n+1)^2 - \left(k-\frac{1}{2}(n+1)\right)^2$ принимает наименьшее значение при $k = 1$ и при $k = n$, а наибольшее значение — при $k = \frac{1}{2}(n+1)$. Поэтому

$$\prod_{k=1}^{n} n \leqslant n!^2 \leqslant \prod_{k=1}^{n} \frac{(n+1)^2}{4},$$

т.е.

$$n^{n/2} \leqslant n! \leqslant \frac{(n+1)^n}{2^n}. \tag{4.22}$$

Это соотношение указывает, что факториальная функция обладает экспоненциальным ростом!!

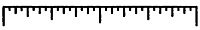
Более точное приближение $n!$ при больших n дает формула Стирлинга, которая будет выведена в главе 9:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{4.23}$$

Еще более точное приближение говорит нам об асимптотической погрешности формулы Стирлинга, которая дает заниженные значения $n!$ на множитель, примерно равный $1/(12n)$. Это более точное приближение вполне неплохо работает даже при достаточно малых n. Например, приближение Стирлинга (4.23) при $n = 10$ дает значение, близкое к 3598696, что примерно на $0.83\% \approx 1/120$ меньше точного значения. Замечательная вещь — асимптотика!

Но вернемся к простым числам. Хотелось бы для каждого заданного простого числа p определить наибольшую степень этого числа, которая делит $n!$, т.е. показатель степени при p в единственном разложении $n!$ на простые множители. Обозначим это число как $\epsilon_p(n!)$ и начнем наше исследование с малых случаев $p = 2$ и $n = 10$. Поскольку $10!$ — произведение десяти чисел, величину $\epsilon_2(10!)$ можно найти, суммируя вхождения степеней 2 в эти числа; это вычисление соответствует суммированию столбцов в следующем массиве:

	1	2	3	4	5	6	7	8	9	10	Степени 2
Делится на 2		x		x		x		x		x	$5 = \lfloor 10/2 \rfloor$
Делится на 4				x				x			$2 = \lfloor 10/4 \rfloor$
Делится на 8								x			$1 = \lfloor 10/8 \rfloor$
Степени 2	0	1	0	2	0	1	0	3	0	1	8

(Суммарные значения столбцов образуют то, что иногда называют *линеечной функцией* (*ruler function*) $\rho(k)$ из-за ее сходства с линейкой , деления которой отмечают доли дюйма.) Сумма этих десяти значений равна 8; следовательно, 2^8 делит $10!$, а 2^9 — нет.

А может быть, сантиметра или парсека.

Есть и другой способ: можно сосчитать вклады строк. В первой строке отмечены числа, которые вносят степень 2 (и, таким образом, делятся на 2): таких чисел $\lfloor 10/2 \rfloor = 5$. Во второй строке отмечены числа, которые вносят дополнительную степень 2; таковых $\lfloor 10/4 \rfloor = 2$. В третьей строке помечены числа, которые вносят еще кое-что; их $\lfloor 10/8 \rfloor = 1$. Тем самым подсчитываются все вклады, так что мы получаем $\epsilon_2(10!) = 5 + 2 + 1 = 8$.

Для произвольного n данный метод дает

$$\epsilon_2(n!) = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n}{4} \right\rfloor + \left\lfloor \frac{n}{8} \right\rfloor + \cdots = \sum_{k \geqslant 1} \left\lfloor \frac{n}{2^k} \right\rfloor .$$

Эта сумма в действительности конечна, поскольку при $2^k > n$ общий член суммы превращается в нуль. Следовательно, в ней только $\lfloor \lg n \rfloor$ ненулевых членов, и ее очень легко вычислить. Например, при $n = 100$ мы получаем

$$\epsilon_2(100!) = 50 + 25 + 12 + 6 + 3 + 1 = 97 .$$

Каждый член этой суммы представляет собой пол от половины предыдущего члена. Это справедливо для всех n, так как в качестве частного случая (3.11) мы имеем $\lfloor n/2^{k+1} \rfloor = \lfloor \lfloor n/2^k \rfloor / 2 \rfloor$. Все особенно просто при записи чисел в двоичной системе счисления:

$$\begin{aligned} 100 &= (1100100)_2 = 100 \\ \lfloor 100/2 \rfloor &= \ (110010)_2 = \ 50 \end{aligned}$$

$$\begin{aligned}
\lfloor 100/4 \rfloor &= (11001)_2 = 25\\
\lfloor 100/8 \rfloor &= (1100)_2 = 12\\
\lfloor 100/16 \rfloor &= (110)_2 = 6\\
\lfloor 100/32 \rfloor &= (11)_2 = 3\\
\lfloor 100/64 \rfloor &= (1)_2 = 1
\end{aligned}$$

Мы просто отбрасываем младший бит члена для получения следующего.

Двоичное представление одновременно показывает нам, как вывести еще одну формулу:

$$\epsilon_2(n!) \;=\; n - \nu_2(n)\,, \tag{4.24}$$

где $\nu_2(n)$ — количество единиц в двоичном представлении n. Такое упрощение работоспособно, поскольку каждая 1, которая дает вклад 2^m в значение n, обеспечивает вклад $2^{m-1}+2^{m-2}+\cdots+2^0 = 2^m - 1$ в значение $\epsilon_2(n!)$.

Обобщая наши рассуждения для произвольного простого числа p, мы получаем

$$\epsilon_p(n!) \;=\; \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots \;=\; \sum_{k\geqslant 1} \left\lfloor \frac{n}{p^k} \right\rfloor . \tag{4.25}$$

Насколько велико значение $\epsilon_p(n!)$? Можно получить простую (но хорошую) верхнюю границу, просто удалив пол в общем члене суммы и просуммировав бесконечную геометрическую прогрессию:

$$\begin{aligned}
\epsilon_p(n!) \;&<\; \frac{n}{p} + \frac{n}{p^2} + \frac{n}{p^3} + \cdots \;=\\
&=\; \frac{n}{p}\left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) \;=\\
&=\; \frac{n}{p}\left(\frac{p}{p-1}\right) \;=\\
&=\; \frac{n}{p-1}\,.
\end{aligned}$$

При $p = 2$ и $n = 100$ это неравенство дает оценку 100 (точное значение — $\epsilon_2(100!) = 97 < 100$). Таким образом, верхняя граница 100 не только верна, но и достаточно близка к точному значению 97. Фактически истинное значение $n - \nu_2(n)$ в общем случае $\sim n$, так как $\nu_2(n) \leqslant \lceil \lg n \rceil$ асимптотически гораздо меньше n.

При $p = 2$ и 3 наши формулы дают $\epsilon_2(n!) \sim n$ и $\epsilon_3(n!) \sim n/2$, так что представляется правдоподобным, что время от времени значение $\epsilon_3(n!)$ должно быть ровно в два раза меньше $\epsilon_2(n!)$. Например, так и есть при $n = 6$ и $n = 7$, поскольку $6! = 2^4 \cdot 3^2 \cdot 5 = 7!/7$. Но еще никто не доказал, что такое совпадение встречается бесконечно часто.

Граница для $\epsilon_p(n!)$, в свою очередь, дает границу для $p^{\epsilon_p(n!)}$, которая представляет собой вклад p в $n!$:

$$p^{\epsilon_p(n!)} < p^{n/(p-1)} .$$

Можно упростить эту формулу (рискуя сильно ослабить верхнюю границу), заметив, что $p \leqslant 2^{p-1}$; следовательно, $p^{n/(p-1)} \leqslant (2^{p-1})^{n/(p-1)} = 2^n$. Другими словами, вклад любого простого числа в $n!$ меньше 2^n.

Можно использовать это наблюдение для другого доказательства существования бесконечного количества простых чисел. Действительно, если бы имелось только k простых чисел 2, 3, $\ldots$, P_k, то мы имели бы $n! < (2^n)^k = 2^{nk}$ для всех $n > 1$, поскольку каждое простое число может внести в разложение $n!$ множитель, не больший $2^n - 1$. Но получить противоречие с неравенством $n! < 2^{nk}$ очень легко, выбрав n достаточно большим, например $n = 2^{2k}$. Тогда

$$n! < 2^{nk} = 2^{2^{2k}k} = n^{n/2} ,$$

что противоречит неравенству $n! \geqslant n^{n/2}$, полученному в (4.22). Значит, имеется бесконечно много простых чисел.

Можно еще немного повозиться с этим рассуждением и получить грубую оценку для $\pi(n)$ — количества простых чисел, не превышающих n. Каждое такое простое число вносит в разложение $n!$ множитель, меньший 2^n; поэтому, как и ранее,

$$n! < 2^{n\pi(n)} .$$

Если заменить $n!$ приближением Стирлинга (4.23), которое представляет собой нижнюю границу, и прологарифмировать, то можно получить

$$n\pi(n) > n \lg(n/e) + \tfrac{1}{2} \lg(2\pi n) ,$$

а следовательно,

$$\pi(n) > \lg(n/e) .$$

Эта нижняя граница очень слабая по сравнению с действительным значением $\pi(n) \sim n/\ln n$, так как $\log n$ гораздо меньше $n/\log n$ при больших n. Но зато нам почти не пришлось трудиться при ее получении — в конце концов, граница есть граница.

4.5 Взаимная простота

Если НОД$(m, n) = 1$, то целые числа m и n не имеют общих простых сомножителей, и мы говорим, что они *взаимно простые*. Эта концепция настолько важна на практике, что следовало бы ввести для нее специальное обозначение. Увы, специалисты в области теории чисел никак не могут прийти к согласию. Поэтому мы умоляем: ВЫСЛУШАЙТЕ НАС, О МАТЕМАТИКИ ВСЕГО МИРА! ЖДАТЬ ВОЛЬШЕ НЕЛЬЗЯ! МНОЖЕСТВО ФОРМУЛ НЕМЕДЛЕННО СТАНЕТ ПОНЯТНЕЕ ПРИ ВВЕДЕНИИ ТАКОГО ОБОЗНАЧЕНИЯ! ДАВАЙТЕ ОСТАНОВИМСЯ НА ЗАПИСИ '$m \perp n$' И БУДЕМ ГОВОРИТЬ, ЧТО "m ПРОСТО С n" (ПО ОТНОШЕНИЮ К n), ЕСЛИ m И n ВЗАИМНО ПРОСТЫ. Другими словами, объявим, что

Подобно тому, как перпендикулярные линии не имеют общего направления, так и перпендикулярные числа не имеют общих множителей.

$$m \perp n \quad \Longleftrightarrow \quad m, n \text{— целые и НОД}(m, n) = 1. \qquad (4.26)$$

Дробь m/n несократима тогда и только тогда, когда $m \perp n$. Поскольку мы сводим дроби к несократимым, сокращая на наибольший общий множитель числитель и знаменатель, можно предположить, что в общем случае

$$m/\text{НОД}(m, n) \quad \perp \quad n/\text{НОД}(m, n), \qquad (4.27)$$

и это действительно так — это вытекает из доказанного в упр. 14 более общего правила НОД$(km, kn) = k\,$НОД(m, n).

Отношение $\perp$ имеет простую формулировку при работе с представлением чисел в виде степеней простых чисел, так как в силу правила (4.14) для НОД

$$m \perp n \quad \Longleftrightarrow \quad \min(m_p, n_p) = 0 \quad \text{при всех } p. \qquad (4.28)$$

Более того, поскольку m_p и n_p неотрицательны, это можно переписать следующим образом:

$$m \perp n \quad \Longleftrightarrow \quad m_p n_p = 0 \quad \text{при всех } p. \qquad (4.29)$$

Нуль, как и при скалярном произведении ортогональных векторов.

Теперь можно доказать важный закон, согласно которому можно объединять и разъединять два отношения $\perp$ с одной и той же левой частью:

$$k \perp m \ \text{ и } \ k \perp n \quad \Longleftrightarrow \quad k \perp mn. \qquad (4.30)$$

Ввиду соотношения (4.29) это правило представляет собой другой вариант следующего утверждения: $k_p m_p = 0$ и $k_p n_p = 0$ тогда и только тогда, когда $k_p(m_p + n_p) = 0$ при неотрицательных m_p и n_p.

Интересно, что математики говорят "открыт" там, где все остальные сказали бы "придуман".

Существует красивый способ построения множества всех неотрицательных дробей m/n с $m \perp n$, называемый *деревом Штерна–Броко*, потому что он был открыт независимо друг от друга немецким математиком Морицем Штерном (Moritz Stern) [339] и французским часовщиком Ахиллом Броко (Achille Brocot) [40]. Суть этого способа в том, чтобы начать с двух дробей $(\frac{0}{1}, \frac{1}{0})$, а затем повторять необходимое количество раз следующую операцию:

$$\text{вставить } \frac{m+m'}{n+n'} \text{ между двумя соседними дробями } \frac{m}{n} \text{ и } \frac{m'}{n'}\,.$$

Новая дробь $(m+m')/(n+n')$ называется *медиантой* (*mediant*) дробей m/n и m'/n'. Например, первый шаг дает одну новую запись между $\frac{0}{1}$ и $\frac{1}{0}$:

$$\frac{0}{1},\ \frac{1}{1},\ \frac{1}{0};$$

следующий шаг добавляет две записи:

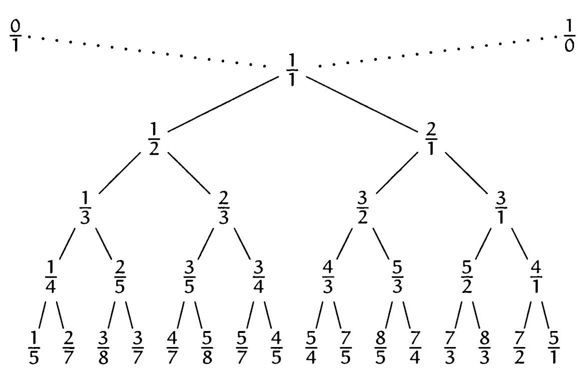
$$\frac{0}{1},\ \frac{1}{2},\ \frac{1}{1},\ \frac{2}{1},\ \frac{1}{0}\,.$$

Очередной шаг добавляет четыре дроби:

$$\frac{0}{1},\ \frac{1}{3},\ \frac{1}{2},\ \frac{2}{3},\ \frac{1}{1},\ \frac{3}{2},\ \frac{2}{1},\ \frac{3}{1},\ \frac{1}{0};$$

Я думаю, что 1/0 — это бесконечность "в несократимом виде".

затем получаются 8, 16 и т.д. новых дробей. Весь массив можно представить в виде бесконечного бинарного дерева, верхние уровни которого имеют следующий вид:



Каждая дробь имеет вид $\frac{m+m'}{n+n'}$, где $\frac{m}{n}$ — ближайший предок вверху слева, а $\frac{m'}{n'}$ — вверху справа. ("Предок" представляет собой дробь, достижимую при следовании по ветвям вверх.) В таком дереве можно найти массу интересных закономерностей.

Но почему такое дерево работает? Почему, например, каждая медианта $(m+m')/(n+n')$ на таком дереве оказывается несократимой? (Если бы все числа m, m', n и n' были нечетными, то мы получили бы дробь вида "четное/четное", но построение дерева гарантирует, что дроби с нечетными числителями и знаменателями никогда не появятся рядом.) И почему все возможные дроби m/n встречаются ровно по одному разу? Почему какая-нибудь дробь не может встретиться дважды (или не встретиться вообще)?

Все эти вопросы имеют на удивление простые ответы, основанные на следующем фундаментальном факте: *если* m/n *и* m'/n' — *последовательные дроби на любом шаге их построения, то*

$$m'n - mn' = 1. \tag{4.31}$$

Это отношение изначально справедливо ($1\cdot 1 - 0\cdot 0 = 1$); а когда мы вставляем новую медианту $(m+m')/(n+n')$, то можем проверить новые соотношения:

$$\begin{aligned} (m+m')n - m(n+n') &= 1;\\ m'(n+n') - (m+m')n' &= 1. \end{aligned}$$

Оба эти уравнения эквивалентны исходному условию (4.31), которое они заменяют. Поэтому условие (4.31) представляет собой инвариант на всех этапах построения.

Кроме того, если $m/n < m'/n'$ и если все эти величины неотрицательны, то легко убедиться в том, что

$$m/n < (m+m')/(n+n') < m'/n'.$$

Медианта лежит где-то между своими предшественниками, хотя и не точно посредине между ними. Следовательно, наше построение сохраняет порядок, и мы не можем получить одну и ту же дробь в разных местах.

Бесплатный совет: если в каком-то месте у вас застряла дробь, покажитесь врачу!

Остается еще один вопрос. Может ли какая-нибудь положительная дробь a/b с $a \perp b$ оказаться пропущенной? Ответ — нет, поскольку можно ограничить наше построение непосредственными соседями a/b, а в этой области поведение дерева легко поддается анализу. Изначально мы имеем

$$\frac{m}{n} = \frac{0}{1} < \left(\frac{a}{b}\right) < \frac{1}{0} = \frac{m'}{n'},$$

где скобки у дроби $\frac{a}{b}$ указывают, что на самом деле ее здесь нет. Затем, если на некотором этапе

$$\frac{m}{n} < \left(\frac{a}{b}\right) < \frac{m'}{n'},$$

построение образует дробь $(m+m')/(n+n')$ и имеются три разные ситуации. Либо $(m+m')/(n+n') = a/b$ и искомая дробь имеется в дереве; либо $(m+m')/(n+n') < a/b$ и мы можем положить $m \leftarrow m+m'$, $n \leftarrow n+n'$; либо $(m+m')/(n+n') > a/b$ и мы можем положить $m' \leftarrow m+m'$, $n' \leftarrow n+n'$. Этот процесс не может продолжаться до бесконечности, поскольку условия

$$\frac{a}{b} - \frac{m}{n} > 0 \qquad \text{и} \qquad \frac{m'}{n'} - \frac{a}{b} > 0$$

означают, что

$$an - bm \geqslant 1 \qquad \text{и} \qquad bm' - an' \geqslant 1;$$

следовательно,

$$(m'+n')(an-bm) + (m+n)(bm'-an') \geqslant \\ \geqslant m'+n'+m+n;$$

а это то же самое, что и $a+b \geqslant m'+n'+m+n$ согласно (4.31). На каждом этапе возрастает либо m, либо n, либо m', либо n', так что мы должны получить искомую дробь не более чем через $a+b$ шагов.

Последовательность Фарея (Farey) порядка N (обозначается $\mathcal{F}_N$) представляет собой множество всех несократимых дробей между 0 и 1, знаменатели которых не превосходят N и которые расположены в возрастающем порядке. Например, при $N = 6$ мы имеем

$$\mathcal{F}_6 = \frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1}.$$

В общем случае можно получить $\mathcal{F}_N$, начав с $\mathcal{F}_1 = \frac{0}{1}, \frac{1}{1}$ и вставляя медианты везде, где это возможно сделать, не нарушив условия, что знаменатель дроби не должен превосходить N. При этом ни одна дробь не будет пропущена, поскольку известно, что построение Штерна–Броко этого не допускает, а медианта со знаменателем $\leqslant N$ не может получиться из дроби с знаменателем $> N$. (Другими словами, $\mathcal{F}_N$ определяет в дереве Штерна–Броко *поддерево*, получающееся путем обрезки ненужных ветвей.) Отсюда следует, что $m'n - mn' = 1$ всегда, когда m/n и m'/n' представляют собой последовательные элементы последовательности Фарея.

Этот метод построения показывает, что последовательность $\mathcal{F}_N$ может быть легко получена из последовательности $\mathcal{F}_{N-1}$: мы просто вставляем дробь $(m+m')/N$ между последовательными дробями m/n и m'/n' последовательности $\mathcal{F}_{N-1}$, сумма знаменателей которых равна N. Например, легко получить $\mathcal{F}_7$ из элементов $\mathcal{F}_6$ путем вставки $\frac{1}{7}$, $\frac{2}{7}$, ..., $\frac{6}{7}$ в соответствии с указанным правилом:

$$\mathcal{F}_7 = \tfrac{0}{1}, \tfrac{1}{7}, \tfrac{1}{6}, \tfrac{1}{5}, \tfrac{1}{4}, \tfrac{2}{7}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{3}{7}, \tfrac{1}{2}, \tfrac{4}{7}, \tfrac{3}{5}, \tfrac{2}{3}, \tfrac{5}{7}, \tfrac{3}{4}, \tfrac{4}{5}, \tfrac{5}{6}, \tfrac{6}{7}, \tfrac{1}{1}.$$

Если N — простое число, то появится $N-1$ новых дробей; в противном случае их будет меньше $N-1$, так как этот процесс генерирует только дроби, числители которых взаимно просты с N.

В (4.5) мы доказали (правда, другими словами), что если $m \perp n$ и $0 < m \leqslant n$, то можно найти целые числа a и b, такие, что

$$ma - nb = 1. \tag{4.32}$$

(На самом деле мы говорили, что $m'm + n'n = \text{НОД}(m, n)$, но вместо $\text{НОД}(m, n)$ можно подставить 1, вместо m' — a, и вместо $-n'$ — b.) Последовательность Фарея дает нам другое доказательство (4.32), поскольку мы можем положить b/a равной дроби, предшествующей m/n в последовательности $\mathcal{F}_n$. Таким образом, (4.5) снова есть просто (4.31). Например, одним из решений уравнения $3a - 7b = 1$ является $a = 5$, $b = 2$, поскольку дробь $\frac{2}{5}$ предшествует дроби $\frac{3}{7}$ в последовательности $\mathcal{F}_7$. Из построения последовательности следует, что всегда можно найти решение (4.32) с $0 \leqslant b < a \leqslant n$, если $0 < m \leqslant n$. Аналогично, если $0 \leqslant n < m$ и $m \perp n$, то уравнение (4.32) с $0 < a \leqslant b \leqslant m$ можно решить, полагая a/b равной дроби, *следующей за* n/m в $\mathcal{F}_m$.

Последовательные тройки дробей в последовательности Фарея обладают удивительным свойством, которое доказывается в упр. 61. Но лучше прекратить обсуждение последовательности Фарея, поскольку дерево Штерна–Броко намного интереснее.

В действительности дерево Штерна–Броко можно рассматривать как *систему счисления* для представления рациональных чисел, поскольку каждая положительная несократимая дробь встречается в нем только один раз. Воспользуемся символами L и R для обозначения левой и правой ветвей при продвижении от корня вниз к определенной дроби. Тогда строка символов L и R однозначно определяет местоположение дроби в дереве. Например, строка LRRL означает, что мы идем влево от $\frac{1}{1}$ к $\frac{1}{2}$, затем — вправо к $\frac{2}{3}$, снова вправо к $\frac{3}{4}$ и влево к $\frac{5}{7}$. Мы можем рассматривать LRRL как представление дроби $\frac{5}{7}$. Таким образом можно однозначно представить любую положительную дробь в виде строки из символов L и R.

Имеется и небольшая проблема: дробь $\frac{1}{1}$ соответствует *пустой* строке, так что для нее требуется особое обозначение. Давайте обозначать ее I, поскольку этот символ напоминает цифру 1 и с него начинается слово "identity" (единица).

Такое представление вызывает два естественных вопроса. (1) Если заданы два положительных целых числа m и n, $m \perp n$, то какой будет строка символов L и R, соответствующая m/n? (2) Если задана некоторая строка символов L и R, то какая дробь ей соответствует? Вопрос 2 кажется более простым, так что начнем с него. Пусть

$$f(S) \;=\; \text{дробь, соответствующая } S,$$

где S — строка символов L и R. Например, $f(LRRL) = \frac{5}{7}$.

В соответствии с построением $f(S) = (m+m')/(n+n')$, если m/n и m'/n' являются ближайшими предшествующими и следующими за S дробями на верхних уровнях нашего дерева. Изначально $m/n = 0/1$ и $m'/n' = 1/0$; затем по мере продвижения по дереву соответственно влево или вправо мы последовательно заменяем медиантой $(m+m')/(n+n')$ либо m/n, либо m'/n'.

Как выразить это поведение при помощи математических формул? Несколько проведенных экспериментов показывают, что наилучший способ — поддержка матрицы 2×2

$$M(S) \;=\; \begin{pmatrix} n & n' \\ m & m' \end{pmatrix},$$

которая содержит все четыре величины, входящие в дроби m/n и m'/n', охватывающие $f(S)$. Подобно записи дробей можно было бы поместить величины m вверху, а n — внизу, но расположение "кверху ногами" оказывается более подходящим, так как изначально $M(I) = \binom{1\,0}{0\,1}$, а матрица $\binom{1\,0}{0\,1}$ — ни что иное, как единичная матрица I.

Шаг влево заменяет n' на $n+n'$ и m' на $m+m'$; следовательно,

$$\begin{aligned} M(SL) \;&=\; \begin{pmatrix} n & n+n' \\ m & m+m' \end{pmatrix} \;= \\ &=\; \begin{pmatrix} n & n' \\ m & m' \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \;=\; M(S) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

(Это частный случай более общего правила,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} aw+by & ax+bz \\ cw+dy & cx+dz \end{pmatrix},$$

умножения матриц 2×2.) Аналогично оказывается, что

Если вы не знакомы с матрицами, не волнуйтесь: больше в этой книге их нет.

$$M(SR) = \begin{pmatrix} n+n' & n' \\ m+m' & m' \end{pmatrix} = M(S)\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Поэтому, если определить L и R как матрицы 2×2,

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \tag{4.33}$$

то индукция по длине S дает простую формулу $M(S) = S$. Правда красиво? (Буквы L и R исполняют две роли: матриц и символов в строковом представлении.) Например,

$$M(LRRL) = LRRL = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3 & 4 \\ 2 & 3 \end{smallmatrix}\right);$$

так что дроби-предки, которые охватывают дробь $LRRL = \frac{5}{7}$, представляют собой $\frac{2}{3}$ и $\frac{3}{4}$. Это построение дает нам ответ на вопрос 2:

$$f(S) = f\left(\begin{pmatrix} n & n' \\ m & m' \end{pmatrix}\right) = \frac{m+m'}{n+n'}. \tag{4.34}$$

А как насчет вопроса 1? Теперь, когда мы понимаем фундаментальную связь между узлами дерева и матрицами 2×2, все просто. При заданной паре положительных целых чисел m и n, с $m \perp n$, положение дроби m/n в дереве Штерна–Броко можно найти "бинарным поиском":

$S := I$;
пока $m/n \neq f(S)$ **выполнять**
  если $m/n < f(S)$ **то** (вывод(L); $S := SL$)
  иначе (вывод(R); $S := SR$).

Выводом этого алгоритма является искомая строка символов L и R.

Имеется и другой способ добиться того же результата — изменяя m и n вместо использования S. Если S — некоторая матрица 2×2, то мы имеем

$$f(RS) = f(S) + 1,$$

поскольку матрица RS подобна матрице S с тем отличием, что в матрице RS верхняя строка добавлена к нижней. (Немного медленнее:

$$S = \begin{pmatrix} n & n' \\ m & m' \end{pmatrix}; \qquad RS = \begin{pmatrix} n & n' \\ m+n & m'+n' \end{pmatrix};$$

следовательно, $f(S) = (m+m')/(n+n')$ и $f(RS) = \big((m+n) + (m'+n')\big)/(n+n')$.) Если алгоритм бинарного поиска применяется к некоторой дроби m/n с $m > n$, то первый вывод этого алгоритма — R; следовательно, дальнейшая работа алгоритма даст $f(S)$ ровно на 1 больше, чем если бы мы начали с дроби $(m-n)/n$, а не с m/n. Аналогичное свойство выполняется и для L, так что мы имеем

$$\begin{aligned} \frac{m}{n} &= f(RS) &\iff&& \frac{m-n}{n} &= f(S) && \text{при } m > n, \\ \frac{m}{n} &= f(LS) &\iff&& \frac{m}{n-m} &= f(S) && \text{при } m < n. \end{aligned}$$

Это означает, что алгоритм бинарного поиска можно преобразовать в следующую процедуру без использования матриц:

$$\begin{array}{l} \textbf{пока}\ \ m \ne n\ \ \textbf{выполнять} \\ \qquad \textbf{если}\ \ m < n\ \ \textbf{то} \qquad \big(\text{вывод}(L);\ \ n := n - m\big) \\ \qquad\qquad\qquad\qquad\ \ \textbf{иначе}\ \big(\text{вывод}(R);\ \ m := m - n\big)\,. \end{array}$$

Например, если дано $m/n = 5/7$, то при помощи этого алгоритма мы получаем

$$\begin{array}{lccccc} m = & 5 & 5 & 3 & 1 & 1 \\ n = & 7 & 2 & 2 & 2 & 1 \\ \text{Вывод} & L & R & R & L & \end{array}$$

В дереве Штерна–Броко нет иррациональных чисел, но зато имеются "близкие" к ним рациональные. Например, если попытаться выполнить бинарный поиск для числа $e = 2.71828\ldots$ вместо дроби m/n, мы получим бесконечную строку из символов L и R, которая начинается следующим образом:

$$RRLRRLRLLLLRLRRRRRRLRLLLLLLLLRLR\ldots.$$

Эту бесконечную строку можно рассматривать как представление числа e в системе счисления Штерна–Броко, так же как число e можно представить в виде бесконечной десятичной дроби $2.718281828459045\ldots$ или в виде бесконечной двоичной дроби

$(10.101101111110\ldots)_2$. Кстати, оказывается, что представление числа e в системе счисления Штерна–Броко обнаруживает определенную закономерность:

Это замечательное двоичное представление было продемонстрировано Германом Минковским (Hermann Minkowski) на Международном конгрессе математиков в Гейдельберге в 1904 году.

$$e = RL^0RLR^2LRL^4RLR^6LRL^8RLR^{10}LRL^{12}RL\ldots,$$

что эквивалентно частному случаю одного открытия, сделанного Эйлером (Euler) [105] в 24 года.

Из этого представления можно вывести, что дроби

$$\begin{array}{ccccccccccccccccc} R & R & L & R & R & L & R & L & L & L & L & R & L & R & R & R & R \\ \frac{1}{1}, & \frac{2}{1}, & \frac{3}{1}, & \frac{5}{2}, & \frac{8}{3}, & \frac{11}{4}, & \frac{19}{7}, & \frac{30}{11}, & \frac{49}{18}, & \frac{68}{25}, & \frac{87}{32}, & \frac{106}{39}, & \frac{193}{71}, & \frac{299}{110}, & \frac{492}{181}, & \frac{685}{252}, & \frac{878}{323}, \ldots \end{array}$$

служат простейшими рациональными верхним и нижним приближениями к e. Действительно, если дробь m/n отсутствует в списке, то между m/n и e находится некоторая дробь из списка, числитель которой $\leqslant m$, а знаменатель $\leqslant n$. Например, $\frac{27}{10}$ нет в списке, так что эта дробь не является таким же простым приближением, как дробь $\frac{19}{7} = 2.714\ldots$, которая имеется в списке и которая ближе к e. Можно убедиться в этом на основании того, что дерево Штерна–Броко не просто включает все рациональные числа, но включает их в упорядоченном виде, а также на основании того, что все дроби с малыми числителем и знаменателем располагаются выше всех менее простых дробей. Так, например, дробь $\frac{27}{10} = RRLRRLL$ меньше дроби $\frac{19}{7} = RRLRRL$, которая, в свою очередь, меньше $e = RRLRRLR\ldots$. Этот способ позволяет находить превосходные приближения. Например, $\frac{878}{323} \approx 2.718266 \approx .999994e$; эта дробь получается из первых 16 символов представления e в системе счисления Штерна–Броко, а точность при этом примерно такая же, как и при 16-битовом двоичном представлении числа e.

Бесконечное представление произвольного иррационального числа α можно получить путем простой модификации безматричной процедуры бинарного поиска:

$$\begin{aligned} \textbf{если } \alpha < 1 \textbf{ то } & \bigl(\text{вывод}(L);\ \alpha := \alpha/(1-\alpha)\bigr) \\ \textbf{иначе } & \bigl(\text{вывод}(R);\ \alpha := \alpha - 1\bigr). \end{aligned}$$

(Эти действия должны быть повторены бесконечное количество раз — ну, или пока вам не надоест...) Если число α рациональное, бесконечное представление, получающееся таким образом, оказывается тем же, что и ранее, но к (конечному) представлению α справа добавляется строка RL^∞. Например, если $\alpha = 1$, мы получим представление $RLLL\ldots$, соответствующее бесконечной последовательности дробей $\frac{1}{1}$, $\frac{2}{1}$, $\frac{3}{2}$, $\frac{4}{3}$, $\frac{5}{4}$, ..., стремящейся в пределе к 1. Эта ситуация аналогична обычному двоичному представлению, если рассматривать L как 0 и R как 1: точно так же, как каждое действительное число x из $[0\,..\,1)$ имеет бесконечное двоичное представление $(.b_1b_2b_3\ldots)_2$, не завершающееся одними цифрами 1, так и каждое действительное число α из $[0\,..\,\infty)$ имеет бесконечное представление Штерна–Броко $B_1B_2B_3\ldots$, не завершающееся одними символами R. Таким образом, имеется взаимно однозначное, сохраняющее порядок соответствие между $[0\,..\,1)$ и $[0\,..\,\infty)$, если исходить из соответствия $0 \leftrightarrow \mathrm{L}$ и $1 \leftrightarrow \mathrm{R}$.

Существует тесная связь между алгоритмом Евклида и представлением Штерна–Броко рациональных чисел. Для данного $\alpha = m/n$ мы получаем $\lfloor m/n \rfloor$ символов R, затем — $\lfloor n/(m \bmod n) \rfloor$ символов L, затем — $\lfloor (m \bmod n)/\bigl(n \bmod (m \bmod n)\bigr) \rfloor$ символов R и т.д. Приведенные выше числа $m \bmod n$, $n \bmod (m \bmod n)$, ... представляют собой значения, фигурировавшие в алгоритме Евклида. (Требуется небольшая изобретательность, чтобы обеспечить отсутствие в конце бесконечного числа символов R.) Эту связь мы будем рассматривать в главе 6.

4.6 'MOD': отношение сравнимости по модулю

"Numerorum congruentiam hoc signo, $\equiv$, in posterum denotabimus, modulum ubi opus erit in clausulis adiungentes, $-16 \equiv 9$ (mod. 5), $-7 \equiv 15$ (mod. 11)."
— К. Ф. Гаусс (C. F. Gauss) [142]

Модулярная арифметика представляет собой один из главных инструментов, которым располагает теория чисел. Мы бегло коснулись ее в главе 3 при использовании бинарной операции 'mod' — как правило, в качестве одной из операций наряду с другими в составе некоторого выражения. В этой главе 'mod' будет использоваться в качестве самостоятельного выражения, в связи с чем более удобна несколько иная форма записи:

$$a \equiv b \pmod m \quad\iff\quad a \bmod m = b \bmod m\,. \tag{4.35}$$

Например, $9 \equiv -16 \pmod 5$, так как $9 \bmod 5 = 4 = (-16) \bmod 5$. Формула '$a \equiv b \pmod m$' читается как "$a$ сравнимо с b по модулю m". Это определение имеет смысл, когда a, b и m — произвольные действительные числа, но мы почти всегда будем использовать его только с целыми числами.

Поскольку $x \bmod m$ отличается от x на величину, кратную m, можно истолковать понятие сравнимости иначе:

$$a \equiv b \pmod m \quad\iff\quad a - b \text{ кратно } m. \tag{4.36}$$

Действительно, если $a \bmod m = b \bmod m$, то определение 'mod' в (3.21) говорит нам, что $a-b = a \bmod m+km-(b \bmod m+lm) = (k-l)m$ при некоторых целых k и l. И наоборот, если $a-b = km$, то $a = b$ при $m = 0$; в противном случае

$$\begin{aligned} a \bmod m \;=\; a - \lfloor a/m \rfloor m \;&=\; b + km - \lfloor (b+km)/m \rfloor m \;= \\ &=\; b - \lfloor b/m \rfloor m \;=\; b \bmod m\,. \end{aligned}$$

Зачастую проще применять толкование $\equiv$ из (4.36), чем из (4.35). Например, $8 \equiv 23 \pmod 5$, так как $8 - 23 = -15$ кратно 5; при этом нам не надо вычислять ни $8 \bmod 5$, ни $23 \bmod 5$.

Знак сравнения по модулю '$\equiv$' выглядит очень похоже на '$=$', что удобно, поскольку сравнения очень похожи на равенства. Например, сравнение по модулю является *отношением эквивалентности*; т.е. оно обладает свойствами рефлексивности '$a \equiv a$', симметричности '$a \equiv b \;\Rightarrow\; b \equiv a$' и транзитивности '$a \equiv b \equiv c \;\Rightarrow\; a \equiv c$'. Все эти свойства легко доказать, поскольку любое отношение '$\equiv$', удовлетворяющее соотношению '$a \equiv b \iff f(a) = f(b)$' для некоторой функции f, является отношением эквивалентности (в нашем случае $f(x) = x \bmod m$). Кроме того, сравнимые элементы можно складывать и вычитать без потери их сравнимости:

"Сегодня я чувствую себя отлично по модулю небольшой головной боли".
— Словарь хакера [337]

$$\begin{aligned} a \equiv b \quad\text{и}\quad c \equiv d \quad&\Longrightarrow\quad a + c \equiv b + d \pmod m\,; \\ a \equiv b \quad\text{и}\quad c \equiv d \quad&\Longrightarrow\quad a - c \equiv b - d \pmod m\,. \end{aligned}$$

Действительно, если $a - b$ и $c - d$ оба кратны m, то же можно сказать и о $(a+c)-(b+d) = (a-b)+(c-d)$ и $(a-c)-(b-d) = (a-b)-(c-d)$. Кстати, не обязательно писать '(mod m)' после каждого знака '$\equiv$'. Если модуль представляет собой константу, достаточно упомянуть его однократно для указания контекста. В этой возможности состоит большое удобство понятия сравнения по модулю.

То же самое справедливо и для умножения, если только мы работаем с целыми числами:

$$a \equiv b \quad\text{и}\quad c \equiv d \quad\Longrightarrow\quad ac \equiv bd \pmod m\,, \quad \text{целые } b, c.$$

Доказательство: $ac - bd = (a-b)c + b(c-d)$. Многократное применение этого свойства умножения позволяет нам получить степени:

$$a \equiv b \quad\Longrightarrow\quad a^n \equiv b^n \pmod m\,, \quad \text{целые } a, b; \quad \text{целое } n \geqslant 0.$$

Например, поскольку $2 \equiv -1 \pmod 3$, имеем $2^n \equiv (-1)^n \pmod 3$; это означает, что $2^n - 1$ кратно 3 тогда и только тогда, когда n четное.

Таким образом, большинство обыкновенно выполняемых с равенствами алгебраических операций можно выполнять и со сравнениями по модулю. Большинство, но не все. Операция деления отсутствует как класс. Если $ad \equiv bd \pmod m$, это не означает, что всегда $a \equiv b$. Например, $3\cdot 2 \equiv 5\cdot 2 \pmod 4$, но $3 \not\equiv 5$.

Однако свойство сократимости сравнений по модулю можно сохранить в часто встречающемся случае, когда d и m взаимно просты:

$$ad \equiv bd \iff a \equiv b \pmod m, \quad \text{целые } a, b, d, m \text{ и } d \perp m. \tag{4.37}$$

Например, из того, что $15 \equiv 35 \pmod m$ вполне закономерно следует вывод, что $3 \equiv 7 \pmod m$, если только модуль m не кратен 5.

Для доказательства этого свойства снова воспользуемся расширением правила (4.5) для НОД, находя d' и m', такие, что $d'd + m'm = 1$. Тогда, если $ad \equiv bd$, можно умножить обе части сравнения по модулю на d', получая $ad'd \equiv bd'd$. Поскольку $d'd \equiv 1$, получаем $ad'd \equiv a$ и $bd'd \equiv b$; следовательно, $a \equiv b$. Это доказательство показывает, что при рассмотрении сравнений $\pmod m$ число d' действует наподобие числа $1/d$; так что будем называть его "обратным к d по модулю m".

Другое применение операции деления к сравнению по модулю заключается в делении модуля наряду с другими числами:

$$\begin{aligned} ad &\equiv bd \pmod{md} \\ &\iff a \equiv b \pmod m \quad \text{при } d \ne 0. \end{aligned} \tag{4.38}$$

Это правило справедливо для всех действительных чисел a, b, d и m, поскольку оно зависит только от дистрибутивного закона $(a \bmod m)d = ad \bmod md$: имеем $a \bmod m = b \bmod m \iff (a \bmod m)d = (b \bmod m)d \iff ad \bmod md = bd \bmod md$. Так, например, из $3\cdot 2 \equiv 5\cdot 2 \pmod 4$ заключаем, что $3 \equiv 5 \pmod 2$.

Можно объединить (4.37) и (4.38), чтобы получить общее правило, изменяющее значение модуля как можно в меньшей степени:

$$\begin{aligned} ad &\equiv bd \pmod m \\ &\iff a \equiv b \left(\bmod \frac{m}{\text{НОД}(d, m)}\right), \quad \text{целые } a, b, d, m. \end{aligned} \tag{4.39}$$

Чтобы получить это правило, можно умножить $ad \equiv bd$ на d', где $d'd + m'm = \text{НОД}(d, m)$; это даст нам сравнение по модулю $a \cdot \text{НОД}(d, m) \equiv b \cdot \text{НОД}(d, m) \pmod m$, которое можно разделить на $\text{НОД}(d, m)$.

Давайте немного разовьем идею изменения величины модуля. Если известно, что $a \equiv b \pmod{100}$, то должно выполняться и $a \equiv b \pmod{10}$ или по модулю любого делителя 100. Утверждение, что $a - b$ кратно 100, более сильное, чем утверждение, что оно кратно 10. В общем случае

$$a \equiv b \pmod{md} \quad \Longrightarrow \quad a \equiv b \pmod m, \quad \text{целое } d, \qquad (4.40)$$

поскольку величина, кратная md, кратна и m.

И наоборот, если мы знаем, что $a \equiv b$ по некоторым двум малым модулям, можно ли на этом основании сделать вывод, что $a \equiv b$ и по некоторому большему модулю? Да, можно: соответствующее правило гласит, что

Модуляшкам?

$$a \equiv b \pmod m \quad \text{и} \quad a \equiv b \pmod n \quad \Longleftrightarrow \quad a \equiv b \pmod{\text{НОК}(m, n)}, \quad \text{целые } m, n > 0. \qquad (4.41)$$

Например, если мы знаем, что $a \equiv b$ по модулю 12 и 18, мы можем с уверенностью заключить, что $a \equiv b \pmod{36}$. Основанием для такого заключения служит то, что если разность $a - b$ кратна как m, так и n, то она кратна и $\text{НОК}(m, n)$. Это следует из принципа единственности разложения на простые множители.

Частный случай этого правила — $m \perp n$ — чрезвычайно важен, поскольку $\text{НОК}(m, n) = mn$, если m и n взаимно просты. Сформулируем его в явном виде:

$$a \equiv b \pmod{mn} \quad \Longleftrightarrow \quad a \equiv b \pmod m \quad \text{и} \quad a \equiv b \pmod n, \quad \text{если } m \perp n. \qquad (4.42)$$

Например, $a \equiv b \pmod{100}$ тогда и только тогда, когда $a \equiv b \pmod{25}$ и $a \equiv b \pmod 4$. Иначе говоря, если мы знаем, что $x \bmod 25$ и $x \bmod 4$, то имеем достаточно информации для того, чтобы определить, что $x \bmod 100$. Это — частный случай *китайской теоремы об остатках* (см. упр. 30), которая называется так потому, что была открыта в Китае Сунь Цзы (Sun Tsŭ) примерно в 350 г. н. э.

Модули m и n в (4.42) могут быть далее разложены на взаимно простые множители, пока не будут выделены все различные простые числа. Следовательно,

$$a \equiv b \pmod m \quad \Longleftrightarrow \quad a \equiv b \pmod{p^{m_p}} \ \text{при всех } p,$$

если разложение (4.11) числа m на простые множители представляет собой $\prod_p p^{m_p}$. Сравнения по модулю степеней простых чисел — это те строительные блоки, из которых строятся все сравнения по модулю целых чисел.

4.7 Независимые остатки

Одно из важнейших приложений сравнений по модулю — *система счисления в остатках* (*residue number system*), в которой целое число x представляется в виде последовательности его остатков (или вычетов) по взаимно простым модулям:

$$\mathrm{Res}(x) = (x \bmod m_1, \ldots, x \bmod m_r), \quad \text{если } m_j \perp m_k \text{ при } 1 \leqslant j < k \leqslant r.$$

Знание $x \bmod m_1$, ..., $x \bmod m_r$ ничего не говорит о x, но позволяет определить $x \bmod m$, где m — произведение $m_1 \ldots m_r$. Поскольку в практических приложениях часто известно, что x находится в определенном диапазоне, то можно узнать об x все, если известно, что $x \bmod m$ и если m достаточно велико.

Рассмотрим, например, малый случай системы счисления в остатках, когда имеются только два модуля, 3 и 5:

$x \bmod 15$	$x \bmod 3$	$x \bmod 5$
0	0	0
1	1	1
2	2	2
3	0	3
4	1	4
5	2	0
6	0	1
7	1	2
8	2	3
9	0	4
10	1	0
11	2	1
12	0	2
13	1	3
14	2	4

Все упорядоченные пары $(x \bmod 3, x \bmod 5)$ различны, поскольку $x \bmod 3 = y \bmod 3$ и $x \bmod 5 = y \bmod 5$ тогда и только тогда, когда $x \bmod 15 = y \bmod 15$.

Согласно правилам сравнений обе компоненты можно складывать, вычитать и умножать *независимо*. Например, если мы хотим умножить $7 = (1, 2)$ на $13 = (1, 3)$ по модулю 15, то вычисляем $1 \cdot 1 \bmod 3 = 1$ и $2 \cdot 3 \bmod 5 = 1$ и получаем окончательный ответ $(1, 1) = 1$; следовательно, $7 \cdot 13 \bmod 15$ должно быть равно 1. Само собой, так оно и есть.

Этот принцип независимости полезен в компьютерных приложениях, поскольку разные компоненты могут обрабатываться раздельно (например, на разных компьютерах). Если каждый модуль m_k представляет собой простое число p_k, несколько меньшее 2^{31}, то компьютер, выполняющий базовые арифметические операции с целыми числами в диапазоне $[-2^{31} \,..\, 2^{31})$, может легко вычислять суммы, разности и произведения по модулю p_k. Набор из r таких простых чисел позволяет складывать, вычитать и умножать "числа многократной точности", состоящие почти из $31r$ бит, а система остатков дает возможность сделать это быстрее, чем при выполнении тех же действий другими способами.

Например, хорошо подходит простое число Мерсенна $2^{31} - 1$.

В определенных ситуациях можно выполнять и деление. Предположим, например, что мы хотим вычислить точное значение большого определителя, состоящего из целых чисел. Результатом будет целое число D, а границы для $|D|$ можно получить на основе размера его элементов. Все известные быстрые способы вычисления определителей требуют выполнения деления, что приводит к дробям (и потере точности, если довольствоваться двоичными приближениями). Выход заключается в вычислении $D \bmod p_k = D_k$ для разных больших простых чисел p_k. Мы можем безопасно выполнять деление по модулю p_k, если только делитель не окажется кратным p_k. Это очень маловероятно, но если такое и случится, то можно выбрать другое простое число. Наконец, зная D_k для достаточно большого количества простых чисел, мы можем определить и само значение D.

Мы пока что не объяснили, как от заданного набора остатков $(x \bmod m_1, \ldots, x \bmod m_r)$ вернуться к $x \bmod m$. Мы показали, что такой переход, в принципе, может быть выполнен, но вычисления могут оказаться такими громоздкими, что на практике окажутся неприемлемыми. К счастью, имеется относительно простой способ выполнить эту работу, который мы можем проиллюстрировать на примере набора $(x \bmod 3, x \bmod 5)$ из нашей таблицы. Ключевая идея состоит в разрешении проблемы для случаев $(1, 0)$ и $(0, 1)$; действительно, если $(1, 0) = a$ и $(0, 1) = b$, то $(x, y) = (ax + by) \bmod 15$, поскольку сравнения могут умножаться и суммироваться.

В нашем случае таблица дает $a = 10$ и $b = 6$; но как найти a и b, когда модули огромны? Другими словами, если $m \perp n$, то где тот путь, который позволит найти числа a и b, такие, что будут выполняться уравнения

$$a \bmod m = 1, \quad a \bmod n = 0,$$
$$b \bmod m = 0, \quad b \bmod n = 1\ ?$$

И вновь нас спасает (4.5): с помощью алгоритма Евклида можно найти такие числа m' и n', что

$$m'm + n'n = 1\,.$$

Поэтому можно взять $a = n'n$ и $b = m'm$, при необходимости приводя их по mod mn.

При больших модулях для уменьшения количества вычислений могут понадобиться и другие хитрости; и хотя подробности выходят за рамки данной книги, их можно найти в [208, с. 274]. Переход от остатков к соответствующим исходным числам — операция, в принципе, выполнимая, но достаточно медленная, так что выигрыш во времени выполнения можно получить, только если можно выполнить сразу некоторую последовательность операций в системе счисления в остатках перед обратным преобразованием.

Попробуем подкрепить эти идеи путем решения небольшой задачи: сколько решений имеет сравнение по модулю

$$x^2 \equiv 1 \pmod m\,, \tag{4.43}$$

если решения x и x' считаются одним и тем же при $x \equiv x'$?

В соответствии с поясненными ранее общими принципами сначала мы должны рассмотреть случай, когда m представляет собой степень простого числа p^k при $k > 0$. Тогда сравнение $x^2 \equiv 1$ может быть записано в виде

$$(x-1)(x+1) \equiv 0 \pmod{p^k}\,,$$

так что p должно делить либо $x-1$, либо $x+1$, либо оба. Но p не может делить и $x-1$, и $x+1$, если только оно не равно 2 (оставим этот случай на потом). Если $p > 2$, то $p^k\backslash(x-1)(x+1) \iff p^k\backslash(x-1)$ или $p^k\backslash(x+1)$; поэтому имеется ровно два решения: $x \equiv +1$ и $x \equiv -1$.

Случай $p = 2$ немного отличается. Если $2^k\backslash(x-1)(x+1)$, то либо $x-1$, либо $x+1$ делится на 2, но не на 4, так что другое число должно делиться на 2^{k-1}. Это означает, что при $k \geqslant 3$

имеется четыре решения, а именно $x \equiv \pm 1$ и $x \equiv 2^{k-1} \pm 1$. (Например, если $p^k = 8$, то эти четыре решения представляют собой $x \equiv 1, 3, 5, 7 \pmod 8$; кстати, часто полезно знать, что *квадрат любого нечетного числа имеет вид* $8n + 1$.)

Итак, $x^2 \equiv 1 \pmod m$ тогда и только тогда, когда $x^2 \equiv 1 \pmod{p^{m_p}}$ для всех простых чисел p с $m_p > 0$ в каноническом разложении m на простые множители. Каждое простое число не зависит от других, и для $x \bmod p^{m_p}$ (за исключением случая $p = 2$) имеется ровно две возможности. Поэтому, если m имеет ровно r различных простых делителей, общее количество решений сравнения по модулю $x^2 \equiv 1$ равно 2^r (за исключением поправки на четное m). В общем случае точное количество решений равно

Все простые числа нечетны, кроме 2, которое нечетнее всех.

$$2^{r+[8\backslash m]+[4\backslash m]-[2\backslash m]}\,. \tag{4.44}$$

Так, имеется четыре "квадратных корня единицы по модулю 12", а именно 1, 5, 7 и 11. При $m = 15$ данная четверка — это числа, которые дают остатки ± 1 по mod 3 и mod 5, а именно $(1,1)$, $(1,4)$, $(2,1)$ и $(2,4)$ в системе счисления в остатках. В обычной (десятичной) системе счисления это числа 1, 4, 11 и 14.

4.8 Дополнительные приложения

С главы 3 за нами небольшой должок: мы обещали доказать, что m чисел

$$0 \bmod m, n \bmod m, 2n \bmod m, \ldots, (m-1)n \bmod m \tag{4.45}$$

состоят в точности из d наборов m/d чисел

$$0, d, 2d, \ldots, m-d$$

в некотором порядке, где $d = \text{НОД}(m,n)$. Например, при $m = 12$ и $n = 8$ имеем $d = 4$, и эти числа представляют собой 0, 8, 4, 0, 8, 4, 0, 8, 4, 0, 8, 4.

Первая часть доказательства — показать, что мы получаем d наборов первых m/d значений — теперь тривиальна. Согласно (4.38)

Любят математики говорить, что то или иное тривиально...

$$jn \equiv kn \pmod m \qquad \iff \qquad j(n/d) \equiv k(n/d) \pmod{m/d},$$

откуда мы получаем d наборов значений, полученных при $0 \leqslant k < m/d$.

Теперь мы должны показать, что эти m/d чисел представляют собой $\{0, d, 2d, \ldots, m-d\}$ в некотором порядке. Положим

$m = m'd$ и $n = n'd$. Тогда по дистрибутивному закону (3.23) $kn \bmod m = d(kn' \bmod m')$, так что величины, которые получаются при $0 \leqslant k < m'$,— это взятые по d раз числа

$$0 \bmod m', n' \bmod m', 2n' \bmod m', \ldots, (m'-1)n' \bmod m'.$$

Но на основании (4.27) известно, что $m' \perp n'$,— они уже поделены на их НОД. Поэтому надо рассмотреть только случай $d = 1$, а именно— когда m и n взаимно просты.

Итак, допустим, что $m \perp n$. В этом случае при помощи "принципа голубиных гнезд" легко видеть, что числа (4.45)— это просто $\{0, 1, \ldots, m-1\}$ в некотором порядке. Этот принцип гласит, что если m голубей рассаживаются по m гнездам, то пустое гнездо имеется тогда и только тогда, когда в каком-то другом гнезде будет находиться больше одного голубя. (Это аналог принципа ящиков Дирихле, доказанного в упр. 3.8.) Известно, что числа (4.45) различны, поскольку

$$jn \equiv kn \pmod m \quad\Longleftrightarrow\quad j \equiv k \pmod m,$$

если $m \perp n$ (это свойство (4.37)). Следовательно, все гнезда с номерами $0, 1, \ldots, m-1$ должны заполниться m различными числами. На этом доказательство завершается и наши долги из главы 3 оплачиваются.

Доказательство завершено, но можно доказать большее, если использовать прямой метод, а не полагаться на косвенную "птичью" аргументацию. Если $m \perp n$ и если задано значение $j \in [0\,..\,m)$, можно явно вычислить число $k \in [0\,..\,m)$, такое, что $kn \bmod m = j$, решая сравнение по модулю

$$kn \equiv j \pmod m$$

относительно k. Мы просто умножаем обе части на n', где $m'm + n'n = 1$, и получаем

$$k \equiv jn' \pmod m;$$

следовательно, $k = jn' \bmod m$.

Только что доказанные факты можно использовать для получения одного важного результата, открытого Пьером де Ферма (Pierre de Fermat) в 1640 году. Ферма был великим математиком, внесшим огромный вклад в создание дифференциального исчисления и многие другие разделы математики. После него остались рукописи, содержащие множество теорем без доказательств, и все они были впоследствии подтверждены. Лишь одна из них противостояла усилиям лучших математиков мира более 350 лет и стала потому самой знаменитой теоремой. Эта "большая теорема Ферма" гласит, что

$$a^n + b^n \neq c^n \qquad (4.46)$$

для всех положительных целых чисел a, b, c и n, если $n > 2$. (Само собой, уравнения $a + b = c$ и $a^2 + b^2 = c^2$ имеют бесконечное количество решений.) Эндрю Уайльс (Andrew Wiles) окончательно разрешил вопрос, опубликовав эпохальное доказательство (4.46) в *Annals of Mathematics* **141** (1995), с. 443–551.

СЕНСАЦИЯ

Эйлер [115] предположил, что $a^4 + b^4 + c^4 \neq d^4$, *но Наум Элкис (Noam Elkies) [92] в августе 1987 года обнаружил, что существует бесконечно много решений. Роджер Фрай (Roger Frye) выполнил исчерпывающий компьютерный поиск и выяснил, затратив около 110 часов работы суперкомпьютера Connection Machine, что единственным решением при* $d < 1000000$ *является*

$$95800^4 + 217519^4 + 414560^4 = 422481^4.$$

Однако проверить теорему Ферма выпуска 1640 года гораздо проще. Сейчас она называется "малая теорема Ферма" (или для краткости просто "теорема Ферма") и гласит, что

$$n^{p-1} \equiv 1 \pmod p, \qquad \text{если } n \perp p. \qquad (4.47)$$

Доказательство. Как обычно, полагаем, что p — простое число. Мы знаем, что $p-1$ чисел $n \bmod p$, $2n \bmod p$, ..., $(p-1)n \bmod p$ представляют собой числа $1, 2, \ldots, p-1$ в некотором порядке. Таким образом, перемножив их все, получим

$$\begin{aligned} n \cdot (2n) \cdot \ldots \cdot ((p-1)n) &\equiv \\ &\equiv (n \bmod p) \cdot (2n \bmod p) \cdot \ldots \cdot ((p-1)n \bmod p) \\ &\equiv (p-1)!, \end{aligned}$$

где выполняется сравнение по модулю p. Это означает, что

$$(p-1)!\, n^{p-1} \equiv (p-1)! \pmod p,$$

и можно выполнить сокращение на множитель $(p-1)!$, поскольку он не делится на p. QED

Иногда удобнее работать с альтернативной формулировкой теоремы:

$$n^p \equiv n \pmod p, \qquad \text{целое } n. \qquad (4.48)$$

Это сравнение по модулю имеет место при любом целом n. Доказательство этого факта очень простое: если $n \perp p$, мы просто умножаем (4.47) на n. Если нет, то $p \backslash n$, так что $n^p \equiv 0 \equiv n$.

В том же году, когда была открыта теорема (4.47), Ферма написал письмо Мерсенну, в котором высказал предположение о том, что числа

$$f_n = 2^{2^n} + 1$$

"...laquelle proposition, si elle est vraie, est de très grand usage".
— П. де Ферма
P. de Fermat [121]

должны быть простыми при всех $n \geqslant 0$. Он знал, что первые пять таких чисел — простые:

$$2^1+1 = 3; \qquad 2^2+1 = 5; \qquad 2^4+1 = 17;$$
$$2^8+1 = 257; \qquad 2^{16}+1 = 65537;$$

однако он не мог сообразить, как доказать простоту следующего числа $2^{32}+1 = 4294967297$.

Интересно, что Ферма смог бы доказать, что $2^{32}+1$ *не* является простым, если бы воспользовался собственной только что открытой теоремой и потратил немного времени для нескольких десятков умножений. Подставим $n = 3$ в (4.47), что дает нам

$$3^{2^{32}} \equiv 1 \pmod{2^{32}+1}, \qquad \text{если } 2^{32}+1 \text{ — простое число.}$$

И это — всего лишь <u>малая</u> теорема Ферма!

Это соотношение можно проверить вручную, начиная с 3 и возводя его в квадрат 32 раза, сохраняя при этом только остатки по модулю $2^{32}+1$. Сначала мы получим $3^2 = 9$, затем — $3^{2^2} = 81$, затем — $3^{2^3} = 6561$, и так далее, пока не доберемся до

$$3^{2^{32}} \equiv 3029026160 \pmod{2^{32}+1}.$$

Результат не равен 1, так что $2^{32}+1$ — составное число. Этот метод опровержения не дает нам никакой информации о сомножителях составного числа, но доказывает, что они существуют. (Это 641 и 6700417, найденные Эйлером в 1732 году [102].)

Если бы число $3^{2^{32}}$ оказалось равным 1 по модулю $2^{32}+1$, то вычисления не доказывали бы, что $2^{32}+1$ — простое; они всего лишь не опровергали бы это. Но в упр. 47 рассматривается обращение теоремы Ферма, с помощью которого *можно* доказать простоту больших чисел, не прибегая к громоздким арифметическим вычислениям.

Мы доказали теорему Ферма, сокращая обе части уравнения на $(p-1)!$. Оказывается, величина $(p-1)!$ всегда сравнима с -1 по модулю p; это — часть классического результата, известного как теорема Вильсона:

$$(n-1)! \equiv -1 \pmod n \iff n \text{ простое,} \quad \text{если } n > 1. \tag{4.49}$$

Половина этой теоремы тривиальна: если $n > 1$ — не простое число, то у него есть простой делитель p, который появляется в качестве множителя в разложении $(n-1)!$, так что $(n-1)!$ не может быть сравнимо с -1. (Если бы $(n-1)!$ было сравнимо с -1

по модулю n, то оно было бы сравнимо и с -1 по модулю p, но это не так.)

Вторая половина теоремы Вильсона гласит, что $(p-1)! \equiv -1 \pmod{p}$. Ее можно доказать, объединяя попарно числа с обратными к ним по модулю p. Если $n \perp p$, то известно, что существует n', такое, что

$$n'n \equiv 1 \pmod{p};$$

здесь n' — обратное к n, а n, в свою очередь, обратное к n'. Любые два обратных к n числа должны быть сравнимы друг с другом, поскольку из $nn' \equiv nn''$ вытекает, что $n' \equiv n''$.

Если p — простое, то p' — производно простое или просто производное?

Теперь предположим, что каждое число между 1 и $p-1$ образует пару со своим обратным. Поскольку произведение некоторого числа и обратного к нему сравнимо с 1, произведение всех пар обратных чисел также сравнимо с 1; по-видимому, $(p-1)!$ также сравнимо с 1. Проверим это, например, для $p = 5$. Получаем $4! = 24$, но это число сравнимо с 4, а не с 1, по модулю 5. Ой! Значит, мы ошиблись, но где? Давайте посмотрим на обратные числа повнимательнее:

$$1' = 1, \qquad 2' = 3, \qquad 3' = 2, \qquad 4' = 4.$$

Вот оно в чем дело: 2 и 3 образуют пару, но 1 и 4 — нет (они являются обратными сами себе).

Чтобы исправить наш анализ, надо определить, какие числа являются обратными сами себе. Если x обратно самому себе, то $x^2 \equiv 1 \pmod{p}$. Мы уже доказали, что это сравнение имеет ровно два корня при $p > 2$. (Если $p = 2$, то очевидно, что $(p-1)! \equiv -1$, так что беспокоиться об этом случае не надо.) Эти корни равны 1 и $p-1$, а все остальные числа (между 1 и $p-1$) благополучно разбиваются на пары; следовательно,

$$(p-1)! \equiv 1 \cdot (p-1) \equiv -1,$$

что и требовалось.

К сожалению, мы не умеем эффективно вычислять факториалы, так что в качестве практического инструмента проверки теорема Вильсона бесполезна — это всего лишь теория.

4.9 Фи и мю

Сколько среди целых чисел $\{0, 1, \ldots, m-1\}$ взаимно простых с m? Эта важная величина называется $\varphi(m)$, тотиентой m (это название придумано британским математиком

Дж. Дж. Сильвестром (J. J. Sylvester) [347], которому нравилось вводить новые термины). Итак, имеем $\varphi(1) = 1$, $\varphi(p) = p - 1$, и $\varphi(m) < m - 1$ для всех составных чисел m.

Функция φ называется *фи-функцией Эйлера*, так как Эйлер был первым, кто ее изучал. Эйлер открыл, например, что теорема Ферма (4.47) может быть обобщена на составные модули следующим образом:

"Si fuerit N ad x numerus primus et n numerus partium ad N primarum, tum potestas x^n unitate minuta semper per numerum N erit divisibilis."
— Л. Эйлер [111]

$$n^{\varphi(m)} \equiv 1 \pmod m, \qquad \text{если } n \perp m. \tag{4.50}$$

(Теорему Эйлера предлагается доказать в упр. 32.)

Если m — степень простого числа p^k, то вычислить $\varphi(m)$ легко, поскольку $n \perp p^k \iff p\nmid n$. Кратными p среди чисел $\{0, 1, \ldots, p^k - 1\}$ являются числа $\{0, p, 2p, \ldots, p^k - p\}$; следовательно, всего их p^{k-1}, а $\varphi(p^k)$ подсчитывает то, что остается:

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

Обратите внимание, что эта формула при $k = 1$ корректно дает $\varphi(p) = p - 1$.

Если $m > 1$ не является степенью простого числа, то можно записать $m = m_1 m_2$, где $m_1 \perp m_2$. Тогда числа $0 \leqslant n < m$ могут быть представлены в системе остатков как $(n \bmod m_1, n \bmod m_2)$. Согласно (4.30) и (4.4)

$$n \perp m \iff n \bmod m_1 \perp m_1 \quad \text{и} \quad n \bmod m_2 \perp m_2.$$

Следовательно, $n \bmod m$ является "хорошей" тогда и только тогда, когда "хорошими" являются и $n \bmod m_1$, и $n \bmod m_2$ (если считать хорошим качеством взаимную простоту). Общее количество хороших величин по модулю m теперь может быть вычислено рекурсивно как $\varphi(m_1)\varphi(m_2)$, поскольку имеется $\varphi(m_1)$ хороших способов выбора первой компоненты $n \bmod m_1$ и $\varphi(m_2)$ хороших способов выбора второй компоненты $n \bmod m_2$ в модульном представлении.

Например, $\varphi(12) = \varphi(4)\varphi(3) = 2 \cdot 2 = 4$, поскольку n взаимно простое с 12 тогда и только тогда, когда $n \bmod 4 = (1 \text{ или } 3)$ и $n \bmod 3 = (1 \text{ или } 2)$. Вот четыре величины в системе остатков, взаимно простых с 12: $(1,1)$, $(1,2)$, $(3,1)$, $(3,2)$; в обычной десятичной записи это 1, 5, 7, 11. Теорема Эйлера гласит, что $n^4 \equiv 1 \pmod{12}$ при $n \perp 12$.

"Si sint A et B numeri inter se primi et numerus partium ad A primarum sit $= a$, numerus vero partium ad B primarum sit $= b$, tum numerus partium ad productum AB primarum erit $= ab$."
— Л. Эйлер [111]

Функция $f(m)$ положительных целых чисел называется *мультипликативной*, если $f(1) = 1$ и

$$f(m_1 m_2) = f(m_1) f(m_2) \qquad \text{для любых } m_1 \perp m_2. \tag{4.51}$$

Мы только что доказали, что $\varphi(m)$ — мультипликативная функция. Мы уже встречались с примером такой функции в этой главе: количество несравнимых решений сравнения по модулю $x^2 \equiv 1 \pmod m$ мультипликативно. Еще одним примером мультипликативной функции служит функция $f(m) = m^\alpha$ при любом показателе степени α.

Мультипликативная функция полностью определяется ее значениями на множестве степеней простых чисел, поскольку любое положительное целое число m может быть разложено на степени простых сомножителей, которые взаимно просты друг с другом. Общая формула

$$f(m) = \prod_p f(p^{m_p}), \qquad \text{если } m = \prod_p p^{m_p}, \tag{4.52}$$

выполняется тогда и только тогда, когда f — мультипликативная функция.

В частности, эта формула дает нам значение фи-функции Эйлера для произвольного числа m:

$$\varphi(m) = \prod_{p\backslash m}(p^{m_p} - p^{m_p-1}) = m\prod_{p\backslash m}\left(1-\frac{1}{p}\right). \tag{4.53}$$

Например, $\varphi(12) = (4-2)(3-1) = 12(1-\frac{1}{2})(1-\frac{1}{3})$.

Рассмотрим теперь применение φ-функции для изучения рациональных чисел по модулю 1. Говорят, что дробь m/n — *правильная*, если $0 \leqslant m < n$. Следовательно, $\varphi(n)$ — это количество несократимых правильных дробей со знаменателем n, а все несократимые правильные дроби со знаменателем n или меньшим и неправильная дробь $\frac{1}{1}$ входят в состав последовательности Фарея.

Множество *всех* правильных дробей со знаменателем 12 до их приведения к несократимым имеет вид

$$\frac{0}{12}, \frac{1}{12}, \frac{2}{12}, \frac{3}{12}, \frac{4}{12}, \frac{5}{12}, \frac{6}{12}, \frac{7}{12}, \frac{8}{12}, \frac{9}{12}, \frac{10}{12}, \frac{11}{12}.$$

Приведение дает

$$\frac{0}{1}, \frac{1}{12}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{5}{12}, \frac{1}{2}, \frac{7}{12}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, \frac{11}{12},$$

и эти дроби можно сгруппировать по знаменателям:

$$\frac{0}{1}; \quad \frac{1}{2}; \quad \frac{1}{3}, \frac{2}{3}; \quad \frac{1}{4}, \frac{3}{4}; \quad \frac{1}{6}, \frac{5}{6}; \quad \frac{1}{12}, \frac{5}{12}, \frac{7}{12}, \frac{11}{12}.$$

И что нам это дает? Ну, например, то, что среди знаменателей встречаются все делители d числа 12 вместе со всеми $\varphi(d)$ своими числителями. Все знаменатели являются делителями 12.

Таким образом,

$$\varphi(1)+\varphi(2)+\varphi(3)+\varphi(4)+\varphi(6)+\varphi(12) = 12.$$

Очевидно, что то же самое получится, если начать с несокращенных дробей $\frac{0}{m}$, $\frac{1}{m}$, ..., $\frac{m-1}{m}$ для любого m, следовательно,

$$\sum_{d\backslash m}\varphi(d) = m. \tag{4.54}$$

В начале этой главы говорилось, что в задачах теории чисел часто приходится вычислять суммы по делителям некоторого числа. Формула (4.54) — пример такой суммы, а далее мы встретимся и с другими примерами.

Вот один забавный факт: если f — некоторая функция, такая, что сумма

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

представляет собой мультипликативную функцию, то и сама f является мультипликативной функцией. (Этот результат вместе с (4.54) и тем фактом, что $g(m) = m$ — очевидно мультипликативная функция, дает другое объяснение тому, почему $\varphi(m)$ — функция мультипликативная.) Этот забавный факт можно доказать по индукции по m: база индукции устанавливается просто, так как $f(1) = g(1) = 1$. Пусть $m > 1$, и допустим, что $f(m_1m_2) = f(m_1)f(m_2)$ для любых $m_1 \perp m_2$ и $m_1m_2 < m$. Если $m = m_1m_2$ и $m_1 \perp m_2$, то

$$g(m_1m_2) = \sum_{d\backslash m_1m_2} f(d) = \sum_{d_1\backslash m_1}\sum_{d_2\backslash m_2} f(d_1d_2)$$

и $d_1 \perp d_2$, так как все делители m_1 взаимно просты со всеми делителями m_2. Согласно гипотезе индукции $f(d_1d_2) = f(d_1)f(d_2)$, за исключением, возможно, случая, когда $d_1 = m_1$ и $d_2 = m_2$; следовательно, мы получаем

$$\begin{aligned}\Bigl(\sum_{d_1\backslash m_1} f(d_1)\sum_{d_2\backslash m_2} f(d_2)\Bigr) - f(m_1)f(m_2) + f(m_1m_2) &=\\ &= g(m_1)g(m_2) - f(m_1)f(m_2) + f(m_1m_2).\end{aligned}$$

Но это равно $g(m_1m_2) = g(m_1)g(m_2)$, так что $f(m_1m_2) = f(m_1)f(m_2)$.

Обратно, если $f(m)$ является мультипликативной функцией, то соответствующая ей функция суммы по делителям $g(m) = \sum_{d\backslash m} f(d)$ всегда мультипликативна. На самом деле, как показано в упр. 33, справедливо даже большее. Следовательно, наш забавный факт и обратный ему — непреложные факты.

Функция Мёбиуса $\mu(m)$, названная в честь математика XIX века Августа Мёбиуса (August Möbius), известного также своей лентой, определяется для всех целых $m \geqslant 1$ уравнением

$$\sum_{d\backslash m} \mu(d) = [m=1]. \qquad (4.55)$$

Это уравнение на самом деле представляет собой рекуррентное соотношение, поскольку его левая часть является суммой из $\mu(m)$ и некоторых значений $\mu(d)$ при $d < m$. Например, если последовательно подставлять $m = 1, 2, \ldots, 12$, то можно вычислить первые двенадцать значений данной функции:

m	1	2	3	4	5	6	7	8	9	10	11	12
$\mu(m)$	1	−1	−1	0	−1	1	−1	0	0	1	−1	0

Ричард Дедекинд (Richard Dedekind) [77] и Жозеф Лиувилль (Joseph Liouville) [251] в 1857 году заметили следующий важный "принцип обращения":

$$g(m) = \sum_{d\backslash m} f(d) \iff f(m) = \sum_{d\backslash m} \mu(d) g\left(\frac{m}{d}\right). \qquad (4.56)$$

Согласно этому принципу функция μ обеспечивает нас новым способом понять, что же собой представляет функция $f(m)$, для которой известна сумма $\sum_{d\backslash m} f(d)$.

Сейчас самое время размяться с помощью упр. 11.

Для доказательства (4.56) воспользуемся приемами (4.7) и (4.9), которые описанны в начале этой главы. Если $g(m) = \sum_{d\backslash m} f(d)$, то

$$\begin{aligned}
\sum_{d\backslash m} \mu(d) g\left(\frac{m}{d}\right) &= \sum_{d\backslash m} \mu\left(\frac{m}{d}\right) g(d) = \\
&= \sum_{d\backslash m} \mu\left(\frac{m}{d}\right) \sum_{k\backslash d} f(k) = \\
&= \sum_{k\backslash m} \sum_{d\backslash (m/k)} \mu\left(\frac{m}{kd}\right) f(k) = \\
&= \sum_{k\backslash m} \sum_{d\backslash (m/k)} \mu(d) f(k) = \\
&= \sum_{k\backslash m} [m/k=1] f(k) = f(m).
\end{aligned}$$

Вторая половина (4.56) доказывается аналогично (см. упр. 12).

Соотношение (4.56) дает нам полезное свойство функции Мёбиуса, и мы уже протабулировали первые двенадцать ее значений. Но что собой представляет $\mu(m)$ при больших m? Как разрешить рекуррентное соотношение (4.55)? Очевидно, что функция $g(m) = [m=1]$ мультипликативна — в конце концов, она равна нулю при всех m, за исключением $m = 1$. Так что функция Мёбиуса, определенная уравнением (4.55), должна быть мультипликативной в соответствии с забавным фактом, доказанным пару минут назад. Таким образом, можно выяснить, что из себя представляет $\mu(m)$, если вычислить $\mu(p^k)$.

Точное время зависит от скорости чтения книги.

Когда $m = p^k$, (4.55) гласит, что

$$\mu(1) + \mu(p) + \mu(p^2) + \cdots + \mu(p^k) \;=\; 0$$

при всех $k \geqslant 1$, поскольку делителями p^k являются $1, \ldots, p^k$. Отсюда следует, что

$$\mu(p) \;=\; -1\,; \qquad \mu(p^k) \;=\; 0 \quad \text{при } k > 1.$$

Поэтому в силу (4.52) получаем общую формулу

$$\mu(m) \;=\; \prod_{p\backslash m} \mu(p^{m_p}) \;=\; \begin{cases} (-1)^r, & \text{если } m = p_1 p_2 \ldots p_r; \\ 0, & \text{если } m \text{ делится на квадрат} \\ & \text{некоторого простого } p^2. \end{cases} \tag{4.57}$$

Вот что такое μ.

Если рассматривать (4.54) как рекуррентное соотношение для функции $\varphi(m)$, то его можно разрешить при помощи правила Дедекинда–Лиувилля (4.56). Мы получим

$$\varphi(m) \;=\; \sum_{d\backslash m} \mu(d)\,\frac{m}{d}\,. \tag{4.58}$$

Например,

$$\begin{aligned} \varphi(12) &= \mu(1)\cdot 12 + \mu(2)\cdot 6 + \mu(3)\cdot 4 + \mu(4)\cdot 3 + \mu(6)\cdot 2 + \mu(12)\cdot 1 \\ &= 12-6-4+0+2+0 = 4\,. \end{aligned}$$

Если m делится на r различных простых чисел, скажем, на $\{p_1, \ldots, p_r\}$, то сумма (4.58) имеет только 2^r ненулевых членов, так как функция μ часто равна нулю. Так, можно убедиться, что (4.58) согласуется с формулой (4.53), записанной в виде

$$\varphi(m) \;=\; m\Bigl(1-\frac{1}{p_1}\Bigr)\ldots\Bigl(1-\frac{1}{p_r}\Bigr)\,;$$

если перемножить все r сомножителей $(1 - 1/p_j)$, мы получим в точности 2^r ненулевых членов (4.58). Достоинство функции Мёбиуса заключается в том, что она применима и во многих иных ситуациях, а не только в рассмотренной.

Например, попробуем выяснить, сколько дробей насчитывается в последовательности Фарея $\mathcal{F}_n$. Это количество равно числу несократимых дробей из $[0\,..\,1]$, знаменатели которых не превосходят n, так что оно на 1 больше, чем величина $\Phi(n)$, которая определяется как

$$\Phi(x) = \sum_{1 \leqslant k \leqslant x} \varphi(k). \qquad (4.59)$$

(Мы вынуждены прибавить 1 к $\Phi(n)$ из-за последней дроби $\frac{1}{1}$.) Сумма (4.59) выглядит сложной, но $\Phi(x)$ можно определить косвенно, заметив, что

$$\sum_{d \geqslant 1} \Phi\left(\frac{x}{d}\right) = \frac{1}{2}\lfloor x\rfloor\lfloor 1+x\rfloor \qquad (4.60)$$

при всех действительных $x \geqslant 0$. Почему выполняется это тождество? Несмотря на его несколько пугающий вид, на самом деле его доказательство не выходит за рамки наших возможностей. Всего имеется $\frac{1}{2}\lfloor x\rfloor\lfloor 1+x\rfloor$ правильных дробей m/n с $0 \leqslant m < n \leqslant x$, с учетом как сокращенных, так и не сокращенных дробей; это дает нам правую часть тождества. Число таких дробей с НОД$(m, n) = d$ равно $\Phi(x/d)$, потому что такие дроби суть m'/n' с $0 \leqslant m' < n' \leqslant x/d$ после замены m на $m'd$ и n на $n'd$. Таким образом, левая часть подсчитывает те же дроби иным способом, и это тождество должно быть истинным.

Давайте более внимательно рассмотрим ситуацию, чтобы уравнения (4.59) и (4.60) стали более понятными. Из определения $\Phi(x)$ вытекает, что $\Phi(x) = \Phi(\lfloor x\rfloor)$; однако удобнее считать функцию $\Phi(x)$ определенной для произвольных действительных чисел, а не только для целых. Для целых значений мы получаем таблицу:

Такое распространение на действительные числа — полезный прием для решения многих рекуррентных соотношений, возникающих в ходе анализа алгоритмов.

n	0	1	2	3	4	5	6	7	8	9	10	11	12
$\varphi(n)$	–	1	1	2	2	4	2	6	4	6	4	10	4
$\Phi(n)$	0	1	2	4	6	10	12	18	22	28	32	42	46

и можем проверить тождество (4.60) при $x = 12$:

$$\Phi(12) + \Phi(6) + \Phi(4) + \Phi(3) + \Phi(2) + \Phi(2) + 6\cdot\Phi(1)$$
$$= 46 + 12 + 6 + 4 + 2 + 2 + 6 = 78 = \tfrac{1}{2}\cdot 12\cdot 13.$$

Удивительно!

Тождество (4.60) можно рассматривать как неявную рекуррентность относительно $\Phi(x)$; например, мы только что видели, как при вычислении $\Phi(12)$ использовались некоторые значения $\Phi(m)$ при $m < 12$. Разрешать такие рекуррентные соотношения можно при помощи еще одного красивого свойства функции Мёбиуса:

В действительности Мёбиус [273] изобрел свою функцию из-за свойства (4.61), а не (4.56).

$$g(x) = \sum_{d\geqslant 1} f(x/d) \iff f(x) = \sum_{d\geqslant 1} \mu(d)\, g(x/d). \qquad (4.61)$$

Это правило обращения выполняется для всех функций f, таких, что $\sum_{k,d\geqslant 1} |f(x/kd)| < \infty$. Это можно доказать следующим образом. Предположим, что $g(x) = \sum_{d\geqslant 1} f(x/d)$. Тогда

$$\begin{aligned}\sum_{d\geqslant 1} \mu(d)\, g(x/d) &= \sum_{d\geqslant 1} \mu(d) \sum_{k\geqslant 1} f(x/kd) = \\ &= \sum_{m\geqslant 1} f(x/m) \sum_{d,k\geqslant 1} \mu(d)[m = kd] = \\ &= \sum_{m\geqslant 1} f(x/m) \sum_{d\backslash m} \mu(d) = \\ &= \sum_{m\geqslant 1} f(x/m)[m = 1] = f(x).\end{aligned}$$

Доказательство в обратном направлении по сути такое же.

Теперь мы в состоянии разрешить рекуррентное соотношение (4.60) относительно $\Phi(x)$:

$$\Phi(x) = \frac{1}{2} \sum_{d\geqslant 1} \mu(d) \lfloor x/d \rfloor \lfloor 1 + x/d \rfloor. \qquad (4.62)$$

Это всегда конечная сумма, например

$$\begin{aligned}\Phi(12) &= \tfrac{1}{2}(12\cdot 13 - 6\cdot 7 - 4\cdot 5 + 0 - 2\cdot 3 + 2\cdot 3 - \\ &\qquad - 1\cdot 2 + 0 + 0 + 1\cdot 2 - 1\cdot 2 + 0) = \\ &= 78 - 21 - 10 - 3 + 3 - 1 + 1 - 1 = 46.\end{aligned}$$

Из главы 9 вы узнаете, как использовать (4.62) для получения хорошего приближения $\Phi(x)$; фактически мы докажем результат, полученный в 1874 году Мертенсом (Mertens) [270]:

$$\Phi(x) = \frac{3}{\pi^2} x^2 + O(x \log x).$$

Следовательно, функция $\Phi(x)$ растет “плавно”, усредняя изменчивое поведение $\varphi(k)$.

Придерживаясь начатой в предыдущей главе традиции, завершим данную главу задачей, которая иллюстрирует многое из того, с чем вы уже познакомились, и что, кроме того, имеет отношение к следующей главе. Предположим, что у нас имеется n бусинок разных цветов; наша цель — подсчитать, сколькими способами из них можно составить ожерелье длиной m. Можно попробовать применить принцип “именуй и властвуй”, обозначив количество возможных ожерелий через $N(m,n)$.

Например, из бусинок цветов R и B ожерелье длиной 4 можно собрать $N(4,2) = 6$ различными способами:

$$\begin{array}{ccc} & R & \\ R & & R \\ & R & \end{array}\qquad \begin{array}{ccc} & R & \\ R & & R \\ & B & \end{array}\qquad \begin{array}{ccc} & R & \\ R & & B \\ & B & \end{array}\qquad \begin{array}{ccc} & R & \\ B & & B \\ & R & \end{array}\qquad \begin{array}{ccc} & R & \\ B & & B \\ & B & \end{array}\qquad \begin{array}{ccc} & B & \\ B & & B \\ & B & \end{array}$$

Все остальные способы эквивалентны одному из приведенных, так как вращение ожерелья новое ожерелье не дает. Однако отражения считаются различными; так, например, для случая $m = 6$

$$\begin{array}{ccc} & B & \\ R & & R \\ | & & | \\ R & & B \\ & B & \end{array}\qquad \text{отличается от} \qquad \begin{array}{ccc} & B & \\ R & & R \\ | & & | \\ B & & R \\ & B & \end{array}.$$

Задача подсчета таких конфигураций была впервые решена П. А. Мак-Мэханом (P. A. MacMahon) [264] в 1892 году.

Явного рекуррентного соотношения для $N(m,n)$ нет, но количество ожерелий можно подсчитать, разрывая и вытягивая в нить m способами каждое из них и рассматривая получающиеся обрывки. Например, при $m = 4$ и $n = 2$ мы получим следующие наборы:

RRRR	RRRR	RRRR	RRRR
RRBR	RRRB	BRRR	RBRR
RBBR	RRBB	BRRB	BBRR
RBRB	BRBR	RBRB	BRBR
RBBB	BRBB	BBRB	BBBR
BBBB	BBBB	BBBB	BBBB

Каждый из n^m возможных шаблонов встречается в этом массиве из $mN(m,n)$ строк по крайней мере один раз, а некоторые — и более того. Сколько раз встречается конкретный набор $a_0 \ldots a_{m-1}$? Все просто: это количество циклических сдвигов строки $a_k \ldots a_{m-1} a_0 \ldots a_{k-1}$, приводящих к исходной строке

$a_0 \ldots a_{m-1}$. Например, строка BRBR встречается дважды, так как четыре способа разрезать ожерелье BRBR дают четыре циклических сдвига (BRBR, RBRB, BRBR, RBRB), два из которых совпадают с самим BRBR. Это рассуждение показывает, что

$$\begin{aligned} mN(m,n) &= \\ &= \sum_{a_0,\ldots,a_{m-1}\in S_n} \sum_{0\leqslant k<m} \bigl[a_0\ldots a_{m-1} = a_k \ldots a_{m-1} a_0 \ldots a_{k-1}\bigr] = \\ &= \sum_{0\leqslant k<m} \sum_{a_0,\ldots,a_{m-1}\in S_n} \bigl[a_0\ldots a_{m-1} = a_k \ldots a_{m-1} a_0 \ldots a_{k-1}\bigr]. \end{aligned}$$

Здесь S_n — множество из n различных цветов.

Выясним, сколько шаблонов удовлетворяют условию $a_0 \ldots a_{m-1} = a_k \ldots a_{m-1} a_0 \ldots a_{k-1}$ при заданном k. Например, если $m = 12$ и $k = 8$, нужно подсчитать количество решений вида

$$\begin{aligned} &a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 a_{10} a_{11} = \\ &\qquad\qquad = a_8 a_9 a_{10} a_{11} a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 . \end{aligned}$$

Это означает, что $a_0 = a_8 = a_4$; $a_1 = a_9 = a_5$; $a_2 = a_{10} = a_6$ и $a_3 = a_{11} = a_7$. Таким образом, a_0, a_1, a_2 и a_3 могут быть выбраны n^4 способами, а остальные a зависят от выбранных. Вам это ничего не напоминает? В общем случае решение

$$a_j = a_{(j+k) \bmod m} \qquad \text{при } 0 \leqslant j < m$$

заставляет нас уравнять a_j с $a_{(j+kl) \bmod m}$ при $l = 1, 2, \ldots$; нам известно, что кратными k по модулю m являются числа $\{0, d, 2d, \ldots, m-d\}$, где $d = \text{НОД}(k,m)$. Поэтому общее решение состоит в независимом выборе $a_0, \ldots, a_{d-1}$ и последующем присвоении $a_j = a_{j-d}$ при $d \leqslant j < m$. Всего решений n^d.

Мы доказали, что

$$mN(m,n) = \sum_{0\leqslant k<m} n^{\text{НОД}(k,m)} .$$

Эта сумма может быть упрощена, поскольку она включает только члены вида n^d, где $d\backslash m$. Подстановка $d = \text{НОД}(k,m)$ дает

$$\begin{aligned} N(m,n) &= \frac{1}{m}\sum_{d\backslash m} n^d \sum_{0\leqslant k<m} \bigl[d = \text{НОД}(k,m)\bigr] = \\ &= \frac{1}{m}\sum_{d\backslash m} n^d \sum_{0\leqslant k<m} \bigl[k/d \perp m/d\bigr] = \\ &= \frac{1}{m}\sum_{d\backslash m} n^d \sum_{0\leqslant k<m/d} \bigl[k \perp m/d\bigr]. \end{aligned}$$

(Мы можем заменить k/d на k, так как k должно быть кратно d.) Наконец, по определению $\sum_{0\leqslant k<m/d}[k\perp m/d] = \varphi(m/d)$, так что мы получаем формулу Мак-Мэхана:

$$N(m,n) = \frac{1}{m}\sum_{d\backslash m} n^d\,\varphi\Bigl(\frac{m}{d}\Bigr) = \frac{1}{m}\sum_{d\backslash m}\varphi(d)\,n^{m/d}. \qquad (4.63)$$

Например, при $m = 4$ и $n = 2$ количество ожерелий, как и следовало ожидать, равно $\frac{1}{4}(1\cdot 2^4 + 1\cdot 2^2 + 2\cdot 2^1) = 6$.

Не очевидно, что значение $N(m,n)$, определяемое суммой Мак-Мэхана, представляет собой целое число! Давайте попробуем доказать непосредственно, что

$$\sum_{d\backslash m}\varphi(d)\,n^{m/d} \equiv 0 \pmod m, \qquad (4.64)$$

не связывая наши рассуждения с ожерельями. В частном случае, когда m — простое число, это сравнение по модулю сводится к $n^p + (p-1)n \equiv 0 \pmod p$, т.е. к сравнению $n^p \equiv n$. Мы уже видели в (4.48), что это сравнение представляет собой иную формулировку теоремы Ферма. Поэтому сравнение (4.64) справедливо при $m = p$; мы можем рассматривать его как обобщение теоремы Ферма на случай, когда модуль не является простым. (Обобщение Эйлера (4.50) — это другое обобщение.)

Мы доказали (4.64) для всех простых модулей, так что рассмотрим наименьший из оставшихся случаев — $m = 4$. Мы должны доказать, что

$$n^4 + n^2 + 2n \equiv 0 \pmod 4.$$

Доказательство будет проще, если рассматривать четные и нечетные случаи по отдельности. Если n четно, то все три члена в левой части сравнимы с 0 по модулю 4, так что их сумма также сравнима с 0. Если же n нечетно, то n^4 и n^2 сравнимы с 1, а $2n$ сравнимо с 2; следовательно, левая часть сравнима с $1 + 1 + 2$, а значит, сравнима с 0 по модулю 4, и мы получили то, что хотели.

Давайте теперь замахнемся на $m = 12$. Это значение m должно быть интересным, так как у него масса множителей, включая квадрат простого числа (при том что само значение достаточно мало́). (Кроме того, мы надеемся на возможность обобщить доказательство для 12 на произвольное значение m.) Сравнение, которое мы должны доказать, —

$$n^{12} + n^6 + 2n^4 + 2n^3 + 2n^2 + 4n \equiv 0 \pmod{12}.$$

И что дальше? Согласно (4.42) это сравнение выполняется тогда и только тогда, когда оно выполняется по модулю 3 и по модулю 4. Поэтому вначале докажем, что оно выполняется по модулю 3. Наше сравнение по модулю (4.64) выполняется для простых чисел, так что $n^3 + 2n \equiv 0 \pmod 3$. Подумав, этот факт можно использовать для группирования членов большей суммы:

$$\begin{aligned} n^{12} + n^6 + 2n^4 + 2n^3 + 2n^2 + 4n &= \\ &= (n^{12} + 2n^4) + (n^6 + 2n^2) + 2(n^3 + 2n) \equiv \\ &\equiv 0 + 0 + 2 \cdot 0 \equiv 0 \qquad \pmod 3\,. \end{aligned}$$

Так что по модулю 3 все получается.

Полдела сделано. Для доказательства сравнения по модулю 4 воспользуемся тем же трюком. Мы уже доказали, что $n^4 + n^2 + 2n \equiv 0 \pmod 4$, так что воспользуемся этим при группировании:

$$\begin{aligned} n^{12} + n^6 + 2n^4 + 2n^3 + 2n^2 + 4n &= \\ &= (n^{12} + n^6 + 2n^3) + 2(n^4 + n^2 + 2n) = \\ &\equiv 0 + 2 \cdot 0 \equiv 0 \qquad \pmod 4\,. \end{aligned}$$

QED для случая $m = 12$.

Пока что наше сравнение по модулю доказано для простого m, $m = 4$ и $m = 12$. Теперь попробуем доказать его для степеней простого числа. Для конкретности можно предположить, что $m = p^3$ при некотором простом p. Тогда левая часть (4.64) представляет собой

$$\begin{aligned} n^{p^3} + \varphi(p)n^{p^2} + \varphi(p^2)n^p + \varphi(p^3)n &= \\ &= n^{p^3} + (p-1)n^{p^2} + (p^2-p)n^p + (p^3-p^2)n = \\ &= (n^{p^3} - n^{p^2}) + p(n^{p^2} - n^p) + p^2(n^p - n) + p^3 n\,. \end{aligned}$$

Можно показать, что это выражение сравнимо с 0 по модулю p^3, если мы сумеем доказать, что $n^{p^3} - n^{p^2}$ делится на p^3, что $n^{p^2} - n^p$ делится на p^2 и что $n^p - n$ делится на p, потому что тогда все это вместе будет делиться на p^3. Согласно альтернативной формулировке теоремы Ферма $n^p \equiv n \pmod p$, так что p делит $n^p - n$; следовательно, существует целое q, такое, что

$$n^p = n + pq\,.$$

Теперь возведем обе части в p-ю степень, разложим правую часть в соответствии с биномиальной теоремой (с которой мы познакомимся в главе 5) и после перегруппирования получим

$$\begin{aligned} n^{p^2} &= (n+pq)^p = \\ &= n^p + (pq)^1 n^{p-1}\binom{p}{1} + (pq)^2 n^{p-2}\binom{p}{2} + \cdots = \\ &= n^p + p^2 Q \end{aligned}$$

при некотором другом целом Q. Мы в состоянии выделить множитель p^2, потому что во втором слагаемом $\binom{p}{1} = p$ и потому что множитель $(pq)^2$ присутствует во всех последующих членах. Итак, мы нашли, что p^2 делит $n^{p^2} - n^p$.

Снова возведем обе стороны уравнения в p-ю степень, разложим в ряд и перегруппируем, получив

$$\begin{aligned} n^{p^3} &= (n^p + p^2 Q)^p = \\ &= n^{p^2} + (p^2 Q)^1 n^{p(p-1)}\binom{p}{1} + (p^2 Q)^2 n^{p(p-2)}\binom{p}{2} + \cdots = \\ &= n^{p^2} + p^3 \mathbf{Q} \end{aligned}$$

при еще одном целом $\mathbf{Q}$. Итак, p^3 делит $n^{p^3} - n^{p^2}$. Этим завершается доказательство для $m = p^3$, так как мы показали, что p^3 делит левую часть (4.64).

Более того, мы можем доказать по индукции, что

$$n^{p^k} = n^{p^{k-1}} + p^k \mathfrak{Q}$$

при некотором последнем целом $\mathfrak{Q}$ (последнем потому, что мы исчерпали все варианты шрифта для вывода этой буквы); следовательно,

$$n^{p^k} \equiv n^{p^{k-1}} \pmod{p^k} \qquad \text{при } k > 0. \tag{4.65}$$

Таким образом, левая часть (4.64), которая представляет собой

$$(n^{p^k} - n^{p^{k-1}}) + p(n^{p^{k-1}} - n^{p^{k-2}}) + \cdots + p^{k-1}(n^p - n) + p^k n,$$

делится на p^k и потому сравнима с 0 по модулю p^k.

Почти все завершено. Теперь, когда мы доказали (4.64) для степеней простых чисел, все, что осталось сделать,— это доказать его при $m = m_1 m_2$, где $m_1 \perp m_2$, в предположении, что данное сравнение истинно при m_1 и m_2. Наше рассмотрение случая $m = 12$, который распадался на отдельные случаи $m = 3$ и $m = 4$, дает основание считать, что такой подход сработает.

Мы знаем, что функция φ мультипликативна, так что можно записать

$$\sum_{d\backslash m} \varphi(d)\, n^{m/d} \;=\; \sum_{d_1\backslash m_1,\, d_2\backslash m_2} \varphi(d_1 d_2)\, n^{m_1 m_2/d_1 d_2} \;=$$
$$= \sum_{d_1\backslash m_1} \varphi(d_1)\Bigl(\sum_{d_2\backslash m_2} \varphi(d_2)(n^{m_1/d_1})^{m_2/d_2}\Bigr).$$

Но внутренняя сумма сравнима с 0 по модулю m_2, так как мы полагаем, что (4.64) выполняется при m_2; так что вся сумма сравнима с 0 по модулю m_2. Из соображений симметричности находим, что вся сумма сравнима с 0 по модулю m_1. Таким образом, в силу (4.42) она сравнима с 0 по модулю m. QED

Упражнения

Разминка

1 Чему равно наименьшее положительное целое число, имеющее ровно k делителей ($1 \leqslant k \leqslant 6$)?

2 Докажите, что $\text{НОД}(m,n)\cdot\text{НОК}(m,n) = m\cdot n$, и воспользуйтесь этим тождеством для того, чтобы выразить $\text{НОК}(m,n)$ через $\text{НОК}(n \bmod m, m)$ при $n \bmod m \neq 0$. *Указание*: воспользуйтесь правилами (4.12), (4.14) и (4.15).

3 Пусть $\pi(x)$ — количество простых чисел, не превосходящих x. Докажите или опровергните, что

$$\pi(x) - \pi(x-1) \;=\; [x \text{— простое число}]\,.$$

4 Что произошло бы, если бы построение Штерна–Броко начиналось с дробей $\left(\frac{0}{1}, \frac{1}{0}, \frac{0}{-1}, \frac{-1}{0}, \frac{0}{1}\right)$, а не с $\left(\frac{0}{1}, \frac{1}{0}\right)$?

5 Найдите простые формулы для L^k и R^k, когда L и R представляют собой матрицы 2×2 из (4.33).

6 Что означает запись '$a \equiv b \pmod 0$'?

7 Десять человек, пронумерованных от 1 до 10, выстроены в круг, как в задаче Иосифа, и казнят каждого m-го человека (значение m может быть гораздо больше 10). Докажите, что первые три казненных не могут быть 10, k и $k+1$ (в указанном порядке) при любом k.

8 Система счисления в остатках $(x \bmod 3, x \bmod 5)$, рассмотренная в тексте, обладает тем забавным свойством, что числу 13 соответствует пара $(1,3)$, которая выглядит почти так же.

Объясните, как найти все такие совпадения без вычисления всех пятнадцати пар остатков. Другими словами, найдите все соответствующие решения сравнений

$$10x+y \equiv x \pmod 3, \qquad 10x+y \equiv y \pmod 5.$$

Указание: воспользуйтесь тем, что $10u+6v \equiv u \pmod 3$ и $10u+6v \equiv v \pmod 5$.

9 Покажите, что $(3^{77}-1)/2$ — нечетное составное число. *Указание*: чему равно $3^{77} \bmod 4$?

10 Вычислите $\varphi(999)$.

11 Найдите функцию $\sigma(n)$, обладающую тем свойством, что

$$g(n) = \sum_{0\leqslant k\leqslant n} f(k)$$
$$\iff \quad f(n) = \sum_{0\leqslant k\leqslant n} \sigma(k)\, g(n-k).$$

(Это аналог функции Мёбиуса; см. (4.56).)

12 Упростите формулу $\sum_{d\backslash m}\sum_{k\backslash d}\mu(k)\, g(d/k)$.

13 Положительное целое число n называется *свободным от квадратов*, если оно не делится на m^2 ни при каком $m>1$. Найдите необходимое и достаточное условие того, что n свободно от квадратов:

а через представление (4.11) числа n в виде показателей простых чисел;

б через функцию $\mu(n)$.

Обязательные упражнения

14 Докажите или опровергните:

а $\text{НОД}(km,kn) = k\,\text{НОД}(m,n)$;

б $\text{НОК}(km,kn) = k\,\text{НОК}(m,n)$.

15 Встречается ли каждое простое число в качестве сомножителя некоторого числа Евклида e_n?

16 Чему равна сумма величин, обратных первым n числам Евклида?

17 Пусть f_n — "число Ферма" $2^{2^n}+1$. Докажите, что $f_m \perp f_n$, если $m<n$.

18 Покажите, что если 2^n+1 — простое число, то n представляет собой степень 2.

19 Докажите следующие тождества при целом положительном n:

$$\sum_{1\leqslant k<n}\left\lfloor\frac{\varphi(k+1)}{k}\right\rfloor=\sum_{1<m\leqslant n}\left\lfloor\left(\sum_{1\leqslant k<m}\lfloor(m/k)/\lceil m/k\rceil\rfloor\right)^{-1}\right\rfloor=$$
$$=n-1-\sum_{k=1}^{n}\left\lceil\left\{\frac{(k-1)!+1}{k}\right\}\right\rceil.$$

Указание: это хитрый вопрос, но ответ на него достаточно прост.

20 Для каждого положительного целого числа n найдется такое простое число p, что $n < p \leqslant 2n$. (Это, по сути, "постулат Бертрана", который Жозеф Бертран (Joseph Bertrand) проверил в 1845 году для значений $n < 3\,000\,000$, а Чебышев доказал для всех n в 1850 году.) Воспользуйтесь постулатом Бертрана для доказательства существования константы $b \approx 1.25$, такой, что все числа

$$\lfloor 2^b\rfloor,\quad \lfloor 2^{2^b}\rfloor,\quad \lfloor 2^{2^{2^b}}\rfloor,\quad \ldots$$

являются простыми.

21 Пусть P_n — n-е простое число. Найдите константу K, такую, что

$$\lfloor(10^{n^2}K)\bmod 10^n\rfloor = P_n\,.$$

Похоже, это проверка на косоглазие.

22 Число 1111111111111111111 — простое. Докажите, что при любом основании системы счисления b число $(11\ldots1)_b$ может быть простым, только если количество единиц — простое.

23 Укажите рекуррентное соотношение для $\rho(k)$ — линеечной функции из обсуждения $\epsilon_2(n!)$ в этой главе. Покажите существование связи между $\rho(k)$ и диском, перемещаемым на шаге k при перемещении Ханойской башни из n дисков за 2^n-1 переносов $(1 \leqslant k \leqslant 2^n-1)$.

Никто не увлекается нумерологией?..

24 Выразите $\epsilon_p(n!)$ через функцию $\nu_p(n)$, которая представляет собой сумму цифр представления числа n в системе счисления с основанием p, обобщая тем самым формулу (4.24).

25 Будем говорить, что m *напросто делит* (*exactly divides*) n, записывая это как $m\backslash\backslash n$, если $m\backslash n$ и $m \perp n/m$ (например, для рассмотренных в тексте факториальных делителей $p^{\epsilon_p(n!)}\backslash\backslash n!$). Докажите или опровергните следующее:

а $k\backslash\backslash n$ и $m\backslash\backslash n \iff km\backslash\backslash n$, если $k \perp m$;

б для всех $m, n > 0$ либо $\text{НОД}(m, n)\backslash\backslash m$, либо $\text{НОД}(m, n)\backslash\backslash n$.

26 Рассмотрим последовательность $\mathcal{G}_N$ всех неотрицательных несократимых дробей m/n, таких, что $mn \leqslant N$, например

$$\mathcal{G}_{10} = \tfrac{0}{1}, \tfrac{1}{10}, \tfrac{1}{9}, \tfrac{1}{8}, \tfrac{1}{7}, \tfrac{1}{6}, \tfrac{1}{5}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{1}{2}, \tfrac{2}{3}, \tfrac{1}{1}, \tfrac{3}{2}, \tfrac{2}{1}, \tfrac{5}{2}, \tfrac{3}{1}, \tfrac{4}{1}, \tfrac{5}{1}, \tfrac{6}{1}, \tfrac{7}{1}, \tfrac{8}{1}, \tfrac{9}{1}, \tfrac{10}{1}.$$

Верно ли, что $m'n - mn' = 1$ всегда, когда m/n непосредственно предшествует дроби m'/n' в $\mathcal{G}_N$?

27 Приведите простое правило сравнения рациональных чисел, основанное на их представлениях с использованием символов L и R в системе счисления Штерна–Броко.

28 Представление Штерна–Броко числа π имеет вид

$$\pi = R^3L^7R^{15}LR^{292}LRLR^2LR^3LR^{14}L^2R\ldots.$$

Используйте его для поиска всех простейших рациональных приближений π, знаменатели которых меньше 50. Входит ли $\frac{22}{7}$ в их число?

29 В этой главе описано соответствие между двоичными действительными числами $x = (.b_1b_2b_3\ldots)_2$ из $[0\,..\,1)$ и действительными числами Штерна–Броко $\alpha = B_1B_2B_3\ldots$ из $[0\,..\,\infty)$. Если x соответствует α и $x \neq 0$, то какое число соответствует $1 - x$?

30 Докажите следующее утверждение (китайскую теорему об остатках): пусть $m_1, \ldots, m_r$ — положительные целые числа с $m_j \perp m_k$ при $1 \leqslant j < k \leqslant r$; пусть $m = m_1 \ldots m_r$; и пусть $a_1, \ldots, a_r, A$ — целые числа. Тогда существует ровно одно целое число a, такое, что

$$a \equiv a_k \pmod{m_k} \text{ при } 1 \leqslant k \leqslant r \quad \text{и} \quad A \leqslant a < A + m.$$

31 Число в десятичной записи делится на 3 тогда и только тогда, когда сумма его цифр делится на 3. Докажите это общеизвестное правило и обобщите его.

32 Докажите теорему Эйлера (4.50), обобщив доказательство теоремы Ферма (4.47).

Почему “Euler” произносится как “Эйлер”, а “Euclid” — как “Евклид”?

33 Покажите, что если $f(m)$ и $g(m)$ — мультипликативные функции, то таковой же будет и функция

$$h(m) = \sum_{d\backslash m} f(d)\, g(m/d).$$

34 Докажите, что (4.56) представляет собой частный случай (4.61).

Домашние задания

35 Пусть $I(m,n)$ — функция, удовлетворяющая соотношению

$$I(m,n)m + I(n,m)n = \text{НОД}(m,n),$$

когда m и n — неотрицательные целые числа и $m \neq n$. Таким образом, в соотношении (4.5) $I(m,n) = m'$ и $I(n,m) = n'$; значение $I(m,n)$ — *обращение* m относительно n. Найдите рекуррентное соотношение, определяющее $I(m,n)$.

36 Рассмотрим множество $Z(\sqrt{10}) = \{\, m + n\sqrt{10} \mid \text{целые } m,n \,\}$. Число $m + n\sqrt{10}$ называется *обратимым*, или *единицей*, если $m^2 - 10n^2 = \pm 1$, поскольку оно имеет обратное (т.е. поскольку $(m+n\sqrt{10})\cdot\pm(m-n\sqrt{10}) = 1$). Например, $3+\sqrt{10}$ — обратимое, как и $19-6\sqrt{10}$. Пары взаимно сокращаемых обратимых чисел могут быть вставлены в любое разложение на множители, так что будем их игнорировать. Необратимые числа множества $Z(\sqrt{10})$ называются простыми, если они не могут быть записаны как произведение двух необратимых чисел. Покажите, что 2, 3 и $4 \pm \sqrt{10}$ являются простыми числами множества $Z(\sqrt{10})$. *Указание*: если $2 = (k + l\sqrt{10}) \times (m + n\sqrt{10})$, то $4 = (k^2 - 10l^2)(m^2 - 10n^2)$. Кроме того, квадрат любого целого числа по модулю 10 равен 0, 1, 4, 5, 6 или 9.

37 Докажите формулу (4.17). *Указание*: покажите, что $e_n - \frac{1}{2} = (e_{n-1} - \frac{1}{2})^2 + \frac{1}{4}$, и рассмотрите числа $2^{-n}\log(e_n - \frac{1}{2})$.

38 Докажите, что если $a \perp b$ и $a > b > 0$, то

$$\text{НОД}(a^m - b^m,\, a^n - b^n) = a^{\text{НОД}(m,n)} - b^{\text{НОД}(m,n)}, \quad 0 \leqslant m < n.$$

(Здесь все переменные — целые числа.) *Указание*: воспользуйтесь алгоритмом Евклида.

39 Пусть $S(m)$ — наименьшее положительное целое число n, для которого существует возрастающая последовательность целых чисел

$$m = a_1 < a_2 < \cdots < a_t = n,$$

такая, что $a_1 a_2 \ldots a_t$ — полный квадрат. (Если m само является полным квадратом, можно положить $t = 1$ и $n = m$.)

Например, $S(2) = 6$, поскольку наилучшей из таких последовательностей является $a_1 = 2$, $a_2 = 3$, $a_3 = 6$. Имеем:

n	1	2	3	4	5	6	7	8	9	10	11	12
$S(n)$	1	6	8	4	10	12	14	15	9	18	22	20

Докажите, что $S(m) \neq S(m')$ для любых $0 < m < m'$.

40 Докажите, что если представление числа n в системе счисления с основанием p имеет вид $(a_m \ldots a_1 a_0)_p$, то

$$n!/p^{\epsilon_p(n!)} \equiv (-1)^{\epsilon_p(n!)} a_m! \ldots a_1!\, a_0! \qquad (\mathrm{mod}\ p)\,.$$

(Левая часть представляет собой просто $n!$, из которого удалены все множители p. При $n = p$ данное утверждение сводится к теореме Вильсона.)

41 **а** Покажите, что если $p \bmod 4 = 3$, то не существует целого числа n, такого, что p делит $n^2 + 1$. *Указание*: воспользуйтесь теоремой Ферма.

б Покажите, что, с другой стороны, если $p \bmod 4 = 1$, то такое число существует. *Указание*: попробуйте записать $(p-1)!$ как $\prod_{k=1}^{(p-1)/2} k(p-k)$ и вспомните теорему Вильсона.

42 Рассмотрим несократимые дроби m/n и m'/n'. Докажите, что если сумма $m/n + m'/n'$ приведена к несократимому виду, то ее знаменатель будет равен nn' тогда и только тогда, когда $n \perp n'$. (Другими словами, $(mn' + m'n)/nn'$ будет сразу несократимой тогда и только тогда, когда n и n' не имеют общего множителя.)

43 На k-м уровне дерева Штерна–Броко имеется 2^k узлов, соответствующих последовательности матриц L^k, $L^{k-1}R$, …, R^k. Покажите, что эта последовательность может быть получена, начиная с матрицы L^k, путем последовательного умножения на

$$\begin{pmatrix} 0 & -1 \\ 1 & 2\rho(n)+1 \end{pmatrix}$$

при $1 \leqslant n < 2^k$, где $\rho(n)$ — линеечная функция.

44 Докажите, что бейсболист с коэффициентом бэттера .316 должен отбить мяч как минимум 19 раз. (Если у него m удачных ударов за n выходов на биту, то $m/n \in [0.3155\,..\,0.3165)$.)

Радиокомментатор: "…питчер Марк Лешиффр дошел до второй базы!

У Марка, коэффициент бэттера которого .080, это второй удачный удар в этом сезоне". Что-то не так?

45 Число 9376 обладает интересным свойством самовоспроизводимости:

$$9376^2 = 87909376.$$

Сколько четырехзначных чисел x удовлетворяют уравнению $x^2 \bmod 10000 = x$? Сколько n-значных чисел x удовлетворяют уравнению $x^2 \bmod 10^n = x$?

46 **а** Докажите, что если $n^j \equiv 1$ и $n^k \equiv 1 \pmod m$, то $n^{\text{НОД}(j,k)} \equiv 1$.

б Покажите, что $2^n \not\equiv 1 \pmod n$, если $n > 1$. *Указание*: рассмотрите наименьший простой множитель n.

47 Покажите, что если $n^{m-1} \equiv 1 \pmod m$ и если $n^{(m-1)/p} \not\equiv 1 \pmod m$ для любого простого числа p, такого, что $p \backslash (m-1)$, то m — простое число. *Указание*: покажите, что если это условие выполняется, то при $1 \leqslant k < m$ все числа $n^k \bmod m$ различны.

48 Обобщите теорему Вильсона (4.49), установив величину выражения $\left(\prod_{1\leqslant n<m,\, n\perp m} n\right) \bmod m$ при $m > 1$.

49 Пусть $R(N)$ — количество пар (m, n) целых чисел, таких, что $1 \leqslant m \leqslant N$, $1 \leqslant n \leqslant N$ и $m \perp n$.

а Выразите $R(N)$ через Φ-функцию.

б Докажите, что $R(N) = \sum_{d\geqslant 1} \lfloor N/d \rfloor^2 \mu(d)$.

50 Пусть m — положительное целое число и пусть

$$\omega = e^{2\pi i/m} = \cos(2\pi/m) + i\sin(2\pi/m).$$

А что такое корень из раз-ницы?

Говорят, что ω представляет собой *корень из единицы* степени m, поскольку $\omega^m = e^{2\pi i} = 1$. В действительности каждое из m комплексных чисел $\omega^0, \omega^1, \ldots, \omega^{m-1}$ является корнем m-й степени из единицы, так как $(\omega^k)^m = e^{2\pi k i} = 1$; таким образом $z - \omega^k$ — один из сомножителей полинома $z^m - 1$ при $0 \leqslant k < m$. Поскольку эти сомножители различны, полное разложение $z^m - 1$ над полем комплексных чисел должно быть таким:

$$z^m - 1 = \prod_{0\leqslant k<m} (z - \omega^k).$$

а Пусть $\Psi_m(z) = \prod_{0\leqslant k<m,\, k\perp m}(z - \omega^k)$. (Этот полином степени $\varphi(m)$ называется *круговым полиномом порядка* m.) Докажите, что

$$z^m - 1 = \prod_{d\backslash m} \Psi_d(z).$$

б Докажите, что $\Psi_m(z) = \prod_{d\backslash m}(z^d - 1)^{\mu(m/d)}$.

Контрольные работы

51 Докажите теорему Ферма (4.48) путем разложения $(1 + 1 + \cdots + 1)^p$ посредством мультиномиальной теоремы.

52 Пусть n и x — положительные целые числа, такие, что x не имеет делителей $\leqslant n$ (за исключением 1), и пусть p — некоторое простое число. Докажите, что по меньшей мере $\lfloor n/p \rfloor$ чисел $\{x - 1, x^2 - 1, \ldots, x^{n-1} - 1\}$ кратны p.

53 Найдите все положительные целые числа n, такие, что $n \backslash \lceil (n-1)!/(n+1) \rceil$.

54 Вычислите значение $1000! \bmod 10^{250}$ вручную.

55 Пусть P_n — произведение первых n факториалов, $\prod_{k=1}^{n} k!$. Докажите, что при любом положительном целом числе n отношение P_{2n}/P_n^4 является целым числом.

56 Покажите, что выражение

$$\left(\prod_{k=1}^{2n-1} k^{\min(k,\, 2n-k)}\right) \Bigg/ \left(\prod_{k=1}^{n-1} (2k+1)^{2n-2k-1}\right)$$

представляет собой степень 2.

57 Пусть $S(m, n)$ — множество всех целых чисел k, таких, что

$$m \bmod k + n \bmod k \;\geqslant\; k\,.$$

Например, $S(7, 9) = \{2, 4, 5, 8, 10, 11, 12, 13, 14, 15, 16\}$. Докажите, что

$$\sum_{k \in S(m,n)} \varphi(k) \;=\; mn\,.$$

Указание: сначала докажите равенство сумм

$$\sum_{1 \leqslant m \leqslant n} \sum_{d \backslash m} \varphi(d) = \sum_{d \geqslant 1} \varphi(d) \lfloor n/d \rfloor .$$

Затем рассмотрите разность $\lfloor (m+n)/d \rfloor - \lfloor m/d \rfloor - \lfloor n/d \rfloor$.

58 Пусть $f(m) = \sum_{d \backslash m} d$. Найдите необходимое и достаточное условие того, что $f(m)$ является степенью 2.

Дополнительные задачи

59 Докажите, что если $x_1, \ldots, x_n$ — положительные целые числа и $1/x_1 + \cdots + 1/x_n = 1$, то $\max(x_1, \ldots, x_n) < e_n$.

Указание: докажите по индукции следующий более сильный результат: "если $1/x_1 + \cdots + 1/x_n + 1/\alpha = 1$, где x_1, $\ldots$, x_n — положительные целые числа, а α — рациональное число $\geqslant \max(x_1, \ldots, x_n)$, то $\alpha + 1 \leqslant e_{n+1}$ и $x_1 \ldots x_n(\alpha+1) \leqslant e_1 \ldots e_n e_{n+1}$". (Доказательство этого факта нетривиально.)

60 Докажите, что существует константа P, такая, что (4.18) дает только простые числа. Можно воспользоваться следующим (в высшей степени нетривиальным) фактом: при любом достаточно большом p между p и $p + p^\theta$ имеется простое число, если $\theta > \frac{6}{11}$.

61 Докажите, что если m/n, m'/n' и m''/n'' — последовательные элементы $\mathcal{F}_N$, то

$$\begin{aligned} m'' &= \lfloor (n+N)/n' \rfloor m' - m, \\ n'' &= \lfloor (n+N)/n' \rfloor n' - n. \end{aligned}$$

(Это рекуррентное соотношение позволяет вычислить все элементы $\mathcal{F}_N$ по порядку, начиная с $\frac{0}{1}$ и $\frac{1}{N}$.)

62 Какое двоичное число соответствует числу e, исходя из соответствия "двоичная система $\leftrightarrow$ система Штерна–Броко"? (Выразите свой ответ в виде бесконечной суммы; вычислять ее в аналитическом виде не требуется.)

63 Покажите, используя только методы из данной главы, что если большая теорема Ферма (4.46) неверна, то наименьшее n, которое ее опровергает, должно быть простым числом. (Можно считать, что (4.46) выполняется при $n = 4$.) Кроме того, если $a^p + b^p = c^p$ — наименьший контрпример, покажите, что для некоторого целого m

$$a + b = \begin{cases} m^p, & \text{если } p \nmid c, \\ p^{p-1} m^p, & \text{если } p \backslash c. \end{cases}$$

Таким образом, число $c \geqslant m^p/2$ должно быть действительно огромным. *Указание*: положите $x = a + b$ и заметьте, что $\text{НОД}(x, (a^p + (x-a)^p)/x) = \text{НОД}(x, pa^{p-1})$.

64 *Последовательность Пирса* $\mathcal{P}_N$ *порядка* N представляет собой бесконечную строку дробей, разделенных знаками '$<$' или '$=$', которая содержит все неотрицательные дроби m/n с $m \geqslant 0$ и $n \leqslant N$ (включая несокращенные дроби). Она определяется рекурсивно, начиная с последовательности

$$\mathcal{P}_1 = \tfrac{0}{1} < \tfrac{1}{1} < \tfrac{2}{1} < \tfrac{3}{1} < \tfrac{4}{1} < \tfrac{5}{1} < \tfrac{6}{1} < \tfrac{7}{1} < \tfrac{8}{1} < \tfrac{9}{1} < \tfrac{10}{1} < \cdots .$$

При $N \geqslant 1$ последовательность $\mathcal{P}_{N+1}$ образуется путем вставки двух символов непосредственно перед kN-м символом $\mathcal{P}_N$ при всех $k > 0$. Эти вставляемые символы представляют собой

$$\begin{array}{lll} \dfrac{k-1}{N+1} & =, & \text{если } kN \text{ нечетное;} \\ \mathcal{P}_{N,kN} & \dfrac{k-1}{N+1}, & \text{если } kN \text{ четное.} \end{array}$$

Здесь $\mathcal{P}_{N,j}$ обозначает j-й символ последовательности $\mathcal{P}_N$, который представляет собой знак '<' либо '=' при четном j и дробь — при j нечетном, например

$$\mathcal{P}_2=\tfrac{0}{2}=\tfrac{0}{1}<\tfrac{1}{2}<\tfrac{2}{2}=\tfrac{1}{1}<\tfrac{3}{2}<\tfrac{4}{2}=\tfrac{2}{1}<\tfrac{5}{2}<\tfrac{6}{2}=\tfrac{3}{1}<\tfrac{7}{2}<\tfrac{8}{2}=\tfrac{4}{1}<\tfrac{9}{2}<\tfrac{10}{2}=\tfrac{5}{1}<\cdots;$$
$$\mathcal{P}_3=\tfrac{0}{2}=\tfrac{0}{3}=\tfrac{0}{1}<\tfrac{1}{3}<\tfrac{1}{2}<\tfrac{2}{3}<\tfrac{2}{2}=\tfrac{3}{3}=\tfrac{1}{1}<\tfrac{4}{3}<\tfrac{3}{2}<\tfrac{5}{3}<\tfrac{4}{2}=\tfrac{6}{3}=\tfrac{2}{1}<\tfrac{7}{3}<\tfrac{5}{2}<\cdots;$$
$$\mathcal{P}_4=\tfrac{0}{2}=\tfrac{0}{4}=\tfrac{0}{3}=\tfrac{0}{1}<\tfrac{1}{4}<\tfrac{1}{3}<\tfrac{2}{4}=\tfrac{1}{2}<\tfrac{2}{3}<\tfrac{3}{4}<\tfrac{2}{2}=\tfrac{4}{4}=\tfrac{3}{3}=\tfrac{1}{1}<\tfrac{5}{4}<\tfrac{4}{3}<\tfrac{6}{4}=\cdots;$$
$$\mathcal{P}_5=\tfrac{0}{2}=\tfrac{0}{4}=\tfrac{0}{5}=\tfrac{0}{3}=\tfrac{0}{1}<\tfrac{1}{5}<\tfrac{1}{4}<\tfrac{1}{3}<\tfrac{2}{5}<\tfrac{2}{4}=\tfrac{1}{2}<\tfrac{2}{5}<\tfrac{2}{3}<\tfrac{3}{4}<\tfrac{4}{5}<\tfrac{2}{2}=\tfrac{4}{4}=\cdots;$$
$$\mathcal{P}_6=\tfrac{0}{2}=\tfrac{0}{4}=\tfrac{0}{6}=\tfrac{0}{5}=\tfrac{0}{3}=\tfrac{0}{1}<\tfrac{1}{6}<\tfrac{1}{5}<\tfrac{1}{4}<\tfrac{2}{6}=\tfrac{1}{3}<\tfrac{2}{5}<\tfrac{2}{4}=\tfrac{3}{6}=\tfrac{1}{2}<\tfrac{3}{5}<\tfrac{4}{6}=\cdots.$$

(Одинаковые элементы выстраиваются в довольно своеобразном порядке.) Докажите, что знаки '<' и '=', определенные приведенными выше правилами, корректно описывают отношения соседних дробей в последовательности Пирса.

Исследовательские проблемы

65 Являются ли все числа Евклида e_n свободными от квадратов?

66 Являются ли все числа Мерсенна $2^p - 1$ свободными от квадратов?

67 Докажите или опровергните, что $\max_{1 \leqslant j < k \leqslant n} a_k / \text{НОД}(a_j, a_k) \geqslant n$ для любой последовательности целых чисел $0 < a_1 < \cdots < a_n$.

68 Существует ли константа Q, такая, что число $\lfloor Q^{2^n} \rfloor$ простое для всех $n \geqslant 0$?

69 Пусть P_n обозначает n-е простое число. Докажите или опровергните, что $P_{n+1} - P_n = O(\log P_n)^2$.

70 Справедливо ли соотношение $\epsilon_3(n!) = \epsilon_2(n!)/2$ для бесконечно большого количества n?

71 Докажите или опровергните следующее утверждение: если $k \neq 1$, то существует $n > 1$, такое, что $2^n \equiv k \pmod{n}$. Существует ли бесконечно много таких n?

72 Докажите или опровергните следующее утверждение: для любого целого a существует бесконечно много таких n, что $\varphi(n)\backslash(n+a)$.

73 Если бы $\Phi(n)+1$ членов последовательности Фарея

$$\mathcal{F}_n = \left\langle \mathcal{F}_n(0), \mathcal{F}_n(1), \ldots, \mathcal{F}_n\bigl(\Phi(n)\bigr) \right\rangle$$

были распределены достаточно равномерно, то следовало бы ожидать, что $\mathcal{F}_n(k) \approx k/\Phi(n)$. Таким образом, мерой "отклонения $\mathcal{F}_n$ от равномерности" служит сумма $D(n) = \sum_{k=0}^{\Phi(n)} \bigl|\mathcal{F}_n(k) - k/\Phi(n)\bigr|$. Верно ли, что $D(n) = O(n^{1/2+\epsilon})$ при любом $\epsilon > 0$?

74 Оцените, сколько в множестве $\{0! \bmod p, 1! \bmod p, \ldots, (p-1)! \bmod p\}$ содержится различных значений при $p \to \infty$?

5

Биномиальные коэффициенты

ПЕРЕВЕДЕМ ДЫХАНИЕ. В предыдущих главах нам пришлось нелегко со всеми этими суммами с полами, потолками, модулями, фи- и мю-функциями. Сейчас мы перейдем к изучению биномиальных коэффициентов, которые оказываются (а) более важными в приложениях и (б) более простыми в работе, чем все прочие величины.

Удачи!

5.1 Основные тождества

Символ $\binom{n}{k}$ используется для обозначения биномиального коэффициента, названного так из-за важного свойства, которое будет рассмотрено нами в следующем разделе, — биномиальной теоремы. А читается этот символ как "выбор k из n". Это заклинание берет свое начало из комбинаторной интерпретации — это количество способов выбора k-элементного подмножества из n-элементного множества. Например, выбрать два элемента из множества $\{1, 2, 3, 4\}$ можно шестью способами, а именно как

Известные также как k-элементные сочетания из n элементов.

$$\{1,2\}, \quad \{1,3\}, \quad \{1,4\}, \quad \{2,3\}, \quad \{2,4\}, \quad \{3,4\};$$

так что $\binom{4}{2} = 6$.

Чтобы выразить значение $\binom{n}{k}$ в более привычных терминах, проще начать с определения количества k-элементных *последовательностей*, а не подмножеств, выбираемых из n-элементного множества; в последовательностях учитывается порядок следования их элементов. Воспользуемся теми же соображениями, которые применялись в главе 4, чтобы показать, что $n!$ — количество перестановок n объектов. Имеется n вариантов выбора первого элемента последовательности; для каждого из них имеется $n-1$ вариантов выбора второго элемента последовательности;

и так до $n-k+1$ вариантов выбора k-го элемента последовательности. Это дает нам общее количество последовательностей $n(n-1)\ldots(n-k+1) = n^{\underline{k}}$. Поскольку каждое k-элементное подмножество может быть упорядочено $k!$ разными способами, данное число *последовательностей* учитывает каждое *подмножество* ровно $k!$ раз. Для получения окончательного ответа надо просто разделить полученное количество последовательностей на $k!$:

$$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots(1)}.$$

Например,

$$\binom{4}{2} = \frac{4\cdot 3}{2\cdot 1} = 6,$$

что согласуется с нашим предыдущим перечислением.

Назовем n *верхним индексом*, а k — *нижним индексом*. Комбинаторная интерпретация накладывает на индексы требование быть неотрицательными целыми числами, поскольку множества не могут иметь отрицательного или дробного количества элементов. Однако биномиальные коэффициенты применяются не только в комбинаторике, так что мы уберем некоторые из ограничений. Оказывается, что полезнее всего, когда верхний индекс может быть произвольным действительным (или даже комплексным) числом, а нижний — произвольным целым числом. Таким образом, наше формальное определение принимает следующий вид:

$$\binom{r}{k} = \begin{cases} \dfrac{r(r-1)\ldots(r-k+1)}{k(k-1)\ldots(1)} = \dfrac{r^{\underline{k}}}{k!}, & \text{целое } k \geqslant 0; \\ 0, & \text{целое } k < 0. \end{cases} \qquad (5.1)$$

Это определение обладает рядом заслуживающих внимания свойств. Во-первых, верхний индекс обозначен как r, а не n; буква r указывает на тот факт, что биномиальный коэффициент имеет смысл и тогда, когда здесь оказывается любое действительное число. Так, например, $\binom{-1}{3} = (-1)(-2)(-3)/(3\cdot 2\cdot 1) = -1$. Здесь нет никакого комбинаторного смысла, но $r=-1$ является одним из важных частных случаев. Полезны и нецелые индексы наподобие $r=-1/2$.

Во-вторых, $\binom{r}{k}$ можно рассматривать как полином k-й степени от r. Как мы увидим, такая точка зрения часто оказывается полезной.

В-третьих, мы оставили не определенными биномиальные коэффициенты для нецелых нижних индексов. Можно дать приемлемое определение и в этом случае, но на практике такие приложения встречаются редко, так что мы отложим данное обобщение и займемся им немного позже в данной главе.

Последнее замечание. В правой части определения перечислены ограничения 'целое $k \geqslant 0$' и 'целое $k < 0$'. Такие ограничения будут приводиться во всех рассматриваемых нами соотношениях, чтобы яснее представлять область их применения. В общем случае чем меньше ограничений, тем лучше, так как соотношения без ограничений более широко применимы; имеющиеся ограничения представляют собой важную часть соотношения. При работе с биномиальными коэффициентами проще на некоторое время забыть о сложных для запоминания ограничениях, а позже проверить, не нарушаются ли какие-либо из них. Такая проверка должна быть выполнена в обязательном порядке.

Например, почти всегда, когда мы встречаем коэффициент $\binom{n}{n}$, он равен 1, так что легко впасть в заблуждение, что он всегда равен 1. Но внимательное рассмотрение определения (5.1) показывает, что $\binom{n}{n}$ равно 1, только когда $n \geqslant 0$ (в предположении, что n целое); при $n < 0$ мы получаем $\binom{n}{n} = 0$. Такие ловушки могут (и будут) делать жизнь интереснее.

Прежде чем приступить к соотношениям, которыми мы будем пользоваться для укрощения биномиальных коэффициентов, давайте познакомимся с некоторыми их малыми значениями. Числа в табл. 201 образуют начало *треугольника Паскаля*, названного так по имени Блеза Паскаля (Blaise Pascal) (1623–1662), написавшего о них основополагающий трактат [285]. Пустые ме-

Биномиальные коэффициенты были хорошо известны в Азии за много веков до рождения Паскаля [90], но он никоим образом не мог об этом знать.

Таблица 201. Треугольник Паскаля

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$	$\binom{n}{10}$
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

ста в таблице на самом деле представляют собой 0, поскольку в числителе в (5.1) имеется нулевой множитель; так, например, $\binom{1}{2} = (1\cdot 0)/(2\cdot 1) = 0$. Эти места оставлены пустыми просто для того, чтобы выделить остальную часть таблицы.

Сто́ит запомнить формулы для первых трех столбцов:

$$\binom{r}{0} = 1, \qquad \binom{r}{1} = r, \qquad \binom{r}{2} = \frac{r(r-1)}{2}. \tag{5.2}$$

Они справедливы для произвольных действительных чисел. (Вспомним, что $\binom{n+1}{2} = \frac{1}{2}n(n+1)$ — это формула для треугольных чисел, выведенная в главе 1; эти числа сразу же бросаются в глаза в столбце $\binom{n}{2}$ табл. 201.) Неплохо также запомнить первые строк пять треугольника Паскаля: наличие набора чисел 1, 4, 6, 4, 1 при решении какой-нибудь задачи должно наводить на мысль о том, что дело тут не обойдется без биномиальных коэффициентов.

Числа в треугольнике Паскаля удовлетворяют практически бесконечному множеству соотношений, так что не удивляйтесь тому, что при более близком знакомстве с ним мы найдем кое-что интересненькое. Например, интересное "свойство шестиугольника", иллюстрируемое шестью числами, 56, 28, 36, 120, 210, 126, которые окружают число 84 в правой нижней части табл. 201. Два варианта чередующегося перемножения чисел из этого шестиугольника дают одно и то же произведение: $56\cdot 36\cdot 210 = 28\cdot 120\cdot 126 = 423360$. Это свойство выполняется для любого такого шестиугольника из любой части треугольника Паскаля.

Итальянцы называют его треугольником Тартальи (Tartaglia).

А теперь — о различных соотношениях. В этом разделе наша цель заключается в том, чтобы изучить несколько простых правил, которые позволяют решать подавляющее большинство практических задач с биномиальными коэффициентами.

"C'est une chose estrange combien il est fertile en proprietez".
— Б. Паскаль (B. Pascal) [285]

Определение (5.1) можно переписать с применением факториалов в обычном случае, когда верхний индекс r представляет собой некоторое целое число n, большее или равное нижнему индексу k:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}, \qquad \text{целые } n \geqslant k \geqslant 0. \tag{5.3}$$

Чтобы получить эту формулу, мы просто умножаем числитель и знаменатель (5.1) на $(n-k)!$. Иногда полезно выразить биномиальный коэффициент в таком факториальном виде (например, при доказательстве свойства шестиугольника). Но зачастую приходится работать в обратном направлении, заменяя факториалы биномиальными коэффициентами.

Факториальное представление указывает на симметрию треугольника Паскаля: каждый ряд читается одинаково как слева направо, так и справа налево. Соотношение, отображающее это свойство *симметрии*, получается путем замены k на $n-k$:

$$\binom{n}{k} = \binom{n}{n-k}, \qquad \text{целое } n \geqslant 0, \text{ целое } k. \tag{5.4}$$

Эта формула имеет комбинаторный смысл, поскольку, указывая k выбранных предметов из n, мы тем самым указываем $n-k$ невыбранных предметов.

Ограничения, гласящие, что n и k в (5.4) должны быть целыми, очевидны, поскольку каждый нижний индекс должен быть целым числом. Но почему n не может быть отрицательным? Предположим, например, что $n=-1$. Верно ли равенство

$$\binom{-1}{k} \stackrel{?}{=} \binom{-1}{-1-k}?$$

Нет. В частности, при $k=0$ мы получим 1 слева и 0 справа. В действительности при всяком целом $k \geqslant 0$ левая часть равна

$$\binom{-1}{k} = \frac{(-1)(-2)\ldots(-k)}{k!} = (-1)^k,$$

т.е. либо 1, либо -1; правая же часть равна 0, так как нижний индекс отрицателен. При отрицательном k левая часть равна 0, а правая представляет собой

$$\binom{-1}{-1-k} = (-1)^{-1-k},$$

т.е. равна 1 или -1. Таким образом, равенство '$\binom{-1}{k} = \binom{-1}{-1-k}$' всегда неверно!

Соотношение симметрии не выполняется и при всех прочих отрицательных n. Но, к сожалению, это ограничение легко забывается, поскольку выражение в верхнем индексе бывает отрицательным при некоторых малозаметных (но вполне законных) значениях его переменных. Всякий, кто много раз имел дело с биномиальными коэффициентами, попадался на эту удочку не менее трех раз.

Надеюсь, что я не попадусь на эту удочку на коллоквиуме.

Этот недостаток соотношения симметрии с избытком компенсируется существенным достоинством: оно выполняется при всех значениях k, даже при $k < 0$ или $k > n$ (поскольку в таких случаях обе стороны равны нулю). В противном случае

$0 \leqslant k \leqslant n$, и симметрия следует непосредственно из (5.3):

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

Следующее важное тождество позволяет внесение под знак биномиального коэффициента и вынесение из-под него:

$$\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}, \qquad \text{целое } k \neq 0. \tag{5.5}$$

Данное ограничение на k предохраняет нас от деления на 0. Назовем (5.5) тождеством *внесения* (absorption), поскольку мы часто будем использовать его для внесения переменной в биномиальные коэффициенты, если снаружи она мешает. Это равенство следует из определения (5.1), так как $r^{\underline{k}} = r(r-1)^{\underline{k-1}}$ и $k! = k(k-1)!$ при $k > 0$; обе части равны нулю при $k < 0$.

Если умножить обе стороны (5.5) на k, то получим правило внесения, которое выполняется даже при $k = 0$:

$$k\binom{r}{k} = r\binom{r-1}{k-1}, \qquad \text{целое } k. \tag{5.6}$$

У этого соотношения есть компаньон, который сохраняет нижний индекс неизменным:

$$(r-k)\binom{r}{k} = r\binom{r-1}{k}, \qquad \text{целое } k. \tag{5.7}$$

Вывести (5.7) можно, сделав "бутерброд" из (5.6) между двумя правилами симметрии:

$$\begin{aligned}
(r-k)\binom{r}{k} &= (r-k)\binom{r}{r-k} = && \text{(в силу симметрии)}\\
&= r\binom{r-1}{r-k-1} = && \text{(в силу (5.6))}\\
&= r\binom{r-1}{k}. && \text{(в силу симметрии)}
\end{aligned}$$

Но погодите минутку! Мы сказали, что данное тождество справедливо для *всех* действительных r, но наш вывод верен, только когда r — положительное целое число. (Если мы не хотим использовать свойство симметрии (5.4) себе во вред, то верхний индекс $r-1$ должен быть целым неотрицательным числом.) Может, мы "схимичили"? Нет. Да, данный вывод справедлив только для положительных целых чисел r; но можно утверждать, что

Ну, по крайней мере не в этот раз.

это тождество выполняется при всех значениях r, поскольку обе стороны (5.7) представляют собой полиномы от r степени $k+1$. Ненулевой полином степени d или меньшей может иметь не более d различных нулей; таким образом, разность двух таких полиномов, которая также имеет степень d или меньшую, не может быть нулем более чем в d точках, если только она не тождественно равна нулю. Другими словами, если два полинома степени d или меньшей совпадают более чем в d точках, они должны совпадать везде. Мы показали, что $(r-k)\binom{r}{k} = r\binom{r-1}{k}$, когда r — положительное целое число, так что эти два полинома совпадают в бесконечном числе точек, а значит, тождественно равны.

Метод доказательства из предыдущего абзаца, который мы будем называть *полиномиальным доказательством*, часто полезен при распространении многих соотношений, справедливых для целых чисел, на действительные числа, — мы будем встречаться с ним снова и снова. Некоторые равенства наподобие соотношения симметрии (5.4) не являются тождествами для полиномов, так что этот метод пригоден не всегда. Тем не менее многие соотношения имеют нужный вид.

Например, вот еще одно полиномиальное тождество — быть может, самое важное среди биномиальных тождеств, — известное как *формула сложения*:

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}, \qquad \text{целое } k. \tag{5.8}$$

Если r — положительное целое число, то формула сложения показывает, что каждое число в треугольнике Паскаля представляет собой сумму двух чисел из предыдущего ряда — расположенного непосредственно над ним и того, что слева от него. Эта формула применима при отрицательных, действительных и даже комплексных r; единственное налагаемое ограничение — k должно быть целым (чтобы были определены биномиальные коэффициенты).

Один из способов доказать формулу сложения состоит в том, чтобы предположить, что r — положительное целое число, и воспользоваться комбинаторной интерпретацией. Вспомним, что $\binom{r}{k}$ — это количество возможных k-элементных подмножеств r-элементного множества. Если у нас есть множество из r яиц, среди которых одно тухлое давно пора выбросить, то имеется $\binom{r}{k}$ способов выбрать из них k яиц. $\binom{r-1}{k}$ из этих способов выбирают только хорошие яйца; а $\binom{r-1}{k-1}$ из них дают подмножество, содержащее тухлое яйцо, так как каждое такое подмножество выбирается как тухлое яйцо и $k-1$-элементное подмножество из

$r-1$ хороших яиц. Суммирование этих двух величин дает нам формулу (5.8). Этот вывод предполагает, что r — положительное целое число, а $k \geqslant 0$. Но обе стороны тождества равны нулю при $k < 0$, и полиномиальное доказательство говорит о том, что соотношение (5.8) справедливо и во всех остальных случаях.

Можно вывести (5.8) и путем сложения двух тождеств внесения-вынесения (5.7) и (5.6):

$$(r-k)\binom{r}{k} + k\binom{r}{k} = r\binom{r-1}{k} + r\binom{r-1}{k-1};$$

здесь левая часть равна $r\binom{r}{k}$, так что можно разделить обе части равенства на r. Этот вывод справедлив для всех значений, кроме $r=0$, но этот случай легко проверить отдельно.

Те, кому не по душе все эти хитрости и кто предпочитает идти прямым курсом, могут предпочесть вывод (5.8) путем работы непосредственно с определением. Если $k > 0$,

$$\begin{aligned}\binom{r-1}{k} + \binom{r-1}{k-1} &= \frac{(r-1)^{\underline{k}}}{k!} + \frac{(r-1)^{\underline{k-1}}}{(k-1)!} = \\ &= \frac{(r-1)^{\underline{k-1}}(r-k)}{k!} + \frac{(r-1)^{\underline{k-1}}k}{k!} = \\ &= \frac{(r-1)^{\underline{k-1}}r}{k!} = \frac{r^{\underline{k}}}{k!} = \binom{r}{k}.\end{aligned}$$

И вновь случаи с $k \leqslant 0$ можно легко проверить.

Мы только что видели три разных способа доказательства формулы сложения. В этом нет ничего удивительного; биномиальные коэффициенты обладают массой полезных свойств, которые могут приводить к различным способам доказательства интересующего нас соотношения.

Формула сложения по сути представляет собой рекуррентное соотношение для чисел из треугольника Паскаля, и мы еще увидим, насколько она полезна при доказательстве других тождеств по индукции. Новое тождество можно получить немедленно, просто разворачивая рекуррентное соотношение, например,

$$\begin{aligned}\binom{5}{3} &= \binom{4}{3} + \binom{4}{2} = \\ &= \binom{4}{3} + \binom{3}{2} + \binom{3}{1} = \\ &= \binom{4}{3} + \binom{3}{2} + \binom{2}{1} + \binom{2}{0} = \\ &= \binom{4}{3} + \binom{3}{2} + \binom{2}{1} + \binom{1}{0} + \binom{1}{-1}.\end{aligned}$$

Поскольку $\binom{1}{-1}=0$, последний член исчезает, и мы можем остановиться. Этот метод приводит нас к общей формуле

$$\sum_{k\leqslant n}\binom{r+k}{k} = \binom{r}{0}+\binom{r+1}{1}+\cdots+\binom{r+n}{n} =$$
$$= \binom{r+n+1}{n}, \qquad \text{целое } n. \tag{5.9}$$

Обратите внимание, что нет необходимости в нижней границе $k \geqslant 0$ для индекса суммирования, так как все члены с $k < 0$ равны нулю.

Эта формула выражает один биномиальный коэффициент в виде суммы других, верхний и нижний индексы которых остаются "равноудаленными". Мы нашли ее путем многократного разложения биномиального коэффициента с наименьшим нижним индексом: сначала — $\binom{5}{3}$, затем — $\binom{4}{2}$, $\binom{3}{1}$, $\binom{2}{0}$. Что получится, если выполнить разворачивание другим образом — разлагая коэффициенты с наибольшим нижним индексом? Мы получим

$$\begin{aligned}
\binom{5}{3} &= \binom{4}{3}+\binom{4}{2} = \\
&= \binom{3}{3}+\binom{3}{2}+\binom{4}{2} = \\
&= \binom{2}{3}+\binom{2}{2}+\binom{3}{2}+\binom{4}{2} = \\
&= \binom{1}{3}+\binom{1}{2}+\binom{2}{2}+\binom{3}{2}+\binom{4}{2} = \\
&= \binom{0}{3}+\binom{0}{2}+\binom{1}{2}+\binom{2}{2}+\binom{3}{2}+\binom{4}{2}.
\end{aligned}$$

Теперь нулю равен коэффициент $\binom{0}{3}$ (так же, как и $\binom{0}{2}$ и $\binom{1}{2}$, но они придают соотношению определенную привлекательность). Мы можем заметить определенную закономерность:

$$\sum_{0\leqslant k\leqslant n}\binom{k}{m} = \binom{0}{m}+\binom{1}{m}+\cdots+\binom{n}{m} =$$
$$= \binom{n+1}{m+1}, \qquad \text{целые } m,n \geqslant 0. \tag{5.10}$$

Это тождество, которое мы назовем *суммированием по верхнему индексу*, выражает один биномиальный коэффициент в виде суммы других, нижние индексы которых представляют собой константу. В данном случае сумма нуждается в нижней границе $k \geqslant 0$, поскольку члены суммы с $k < 0$ *не нулевые*. Кроме того, m и n в общем случае не могут быть отрицательными.

Тождество (5.10) имеет интересную комбинаторную интерпретацию. Если мы хотим выбрать $m+1$ билетов из множества $n+1$ билетов, пронумерованных от 0 до n, то это можно сделать $\binom{k}{m}$ способами, если наибольший номер выбранного билета равен k.

Доказать соотношения (5.9) и (5.10) можно по индукции, используя формулу сложения, но можно также получить одно соотношение из другого. Например, докажем, что (5.9), на основании (5.10); наше доказательство проиллюстрирует некоторые распространенные методы работы с биномиальными коэффициентами. Общий план действий таков: преобразовать левую часть $\sum\binom{r+k}{k}$ соотношения (5.9) так, чтобы она приняла вид левой части $\sum\binom{k}{m}$ соотношения (5.10); затем воспользоваться этим соотношением для замены суммы одним биномиальным коэффициентом; наконец, привести этот коэффициент к виду правой части соотношения (5.9).

Для удобства будем считать, что r и n — неотрицательные целые числа; общее соотношение (5.9) будет получено из этого частного случая при помощи полиномиального доказательства. Вместо r будем писать m, чтобы эта переменная напоминала нам о том, что мы имеем дело с неотрицательным целым числом. Теперь систематическое выполнение нашего плана протекает следующим образом:

$$\begin{aligned}
\sum_{k\leqslant n}\binom{m+k}{k} &= \sum_{-m\leqslant k\leqslant n}\binom{m+k}{k} = \\
&= \sum_{-m\leqslant k\leqslant n}\binom{m+k}{m} = \\
&= \sum_{0\leqslant k\leqslant m+n}\binom{k}{m} = \\
&= \binom{m+n+1}{m+1} = \binom{m+n+1}{n}.
\end{aligned}$$

Давайте проанализируем вывод пошагово. Ключевой является вторая строка, в которой мы применяем правило симметрии (5.4) для замены $\binom{m+k}{k}$ на $\binom{m+k}{m}$. Это можно сделать только при $m+k\geqslant 0$, так что наш первый шаг ограничивает область изменения k, отбрасывая члены с $k<-m$ (что законно, поскольку отброшенные члены равны нулю). Теперь мы почти готовы применить (5.10); третья строка готовит нас к этому, заменяя k на $k-m$ и приводя в порядок диапазон суммирования. Этот шаг, подобно первому, — просто игра с $\sum$-обозначениями. Теперь k появляется в верхнем индексе, и пределы суммирования приведены в надлежащий вид, так что в четвертой строке применяется (5.10). Завершает работу еще одно применение правила симметрии.

Некоторые суммы, с которыми мы имели дело в главах 1 и 2, на самом деле являются либо частными случаями соотношения (5.10), либо скрытыми вариантами этого соотношения. Например, случай $m = 1$ дает сумму неотрицательных целых чисел до n включительно:

$$\binom{0}{1} + \binom{1}{1} + \cdots + \binom{n}{1} = 0 + 1 + \cdots + n = \\ = \frac{(n+1)n}{2} = \binom{n+1}{2}.$$

Общий случай эквивалентен правилу

$$\sum_{0 \leqslant k \leqslant n} k^{\underline{m}} = \frac{(n+1)^{\underline{m+1}}}{m+1}, \qquad \text{целые } m, n \geqslant 0$$

из главы 2, если мы разделим обе части этой формулы на $m!$. Фактически формула сложения (5.8) говорит нам, что

$$\Delta\left(\binom{x}{m}\right) = \binom{x+1}{m} - \binom{x}{m} = \binom{x}{m-1},$$

если заменить r и k соответственно на $x+1$ и m. Следовательно, методы главы 2 дают нам удобную форму неопределенного суммирования

$$\sum \binom{x}{m} \delta x = \binom{x}{m+1} + C. \tag{5.11}$$

Биномиальные коэффициенты получили свое название от *биномиальной теоремы*, которая имеет дело со степенями биномиального выражения $x+y$. Вот малые случаи данной теоремы:

“Когда ему [Мориарти] исполнился двадцать один год, он написал трактат о биномиальной теореме, завоевавший ему европейскую известность. После этого он получил кафедру математики в одном из наших провинциальных университетов...”
— Ш. Холмс (S. Holmes) [84]

$$\begin{aligned}
(x+y)^0 &= 1x^0y^0 \\
(x+y)^1 &= 1x^1y^0 + 1x^0y^1 \\
(x+y)^2 &= 1x^2y^0 + 2x^1y^1 + 1x^0y^2 \\
(x+y)^3 &= 1x^3y^0 + 3x^2y^1 + 3x^1y^2 + 1x^0y^3 \\
(x+y)^4 &= 1x^4y^0 + 4x^3y^1 + 6x^2y^2 + 4x^1y^3 + 1x^0y^4 .
\end{aligned}$$

Нетрудно понять, почему эти коэффициенты совпадают с числами из треугольника Паскаля. Когда мы расписываем произведение

$$(x+y)^n = \overbrace{(x+y)(x+y)\ldots(x+y)}^{n \text{ сомножителей}},$$

то каждый его член представляет собой произведение n сомножителей, каждый из которых может быть либо x, либо y. Количество таких членов с k сомножителями x и $n-k$ сомножителями y представляет собой коэффициент при $x^k y^{n-k}$ после приведения подобных членов. Эта величина в точности равна количеству способов выбора k из n биномов, из которых в произведение войдет x, т.е. $\binom{n}{k}$.

В некоторых учебниках величина 0^0 считается неопределенной, поскольку функции x^0 и 0^x имеют разные пределы при x, стремящемся к 0. Но это ошибка. Мы должны определить

$$x^0 = 1 \qquad \text{при любом } x,$$

чтобы биномиальная теорема была верна при $x = 0$, $y = 0$, и/или $x = -y$. Эта теорема слишком важна, чтобы произвольно ее ограничивать! Функция же 0^x, напротив, особой ценности не представляет (см. дальнейшее обсуждение этого вопроса в [220]).

Но что именно представляет собой биномиальная теорема? Во всей красе она предстает, будучи записанной в следующем виде:

$$(x+y)^r = \sum_k \binom{r}{k} x^k y^{r-k}, \qquad \begin{array}{c}\text{целое } r \geqslant 0\\ \text{или } |x/y| < 1.\end{array} \tag{5.12}$$

Это сумма по всем целым k; но в действительности при целом неотрицательном r это конечная сумма, поскольку все ее члены равны нулю, за исключением тех, у которых $0 \leqslant k \leqslant r$. С другой стороны, биномиальная теорема верна и при отрицательном r, и даже когда r — произвольное действительное или комплексное число. В этих случаях данная сумма действительно бесконечна и для гарантии абсолютной сходимости суммы требуется выполнение условия $|x/y| < 1$.

Два частных случая биномиальной теоремы достойны особого внимания, несмотря на их исключительную простоту. Если $x = y = 1$ и $r = n$ — неотрицательное, то

$$2^n = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n}, \qquad \text{целое } n \geqslant 0.$$

Это равенство говорит о том, что сумма чисел n-й строки треугольника Паскаля равна 2^n. Если x вместо $+1$ равен -1, то мы получим

$$0^n = \binom{n}{0} - \binom{n}{1} + \cdots + (-1)^n \binom{n}{n}, \qquad \text{целое } n \geqslant 0.$$

Например, $1-4+6-4+1=0$; т.е. если дать элементам n-й строки чередующиеся знаки, то их сумма равна нулю (за исключением верхней строки, когда $n=0$ и $0^0=1$).

Если r не является неотрицательным целым числом, то биномиальная теорема чаще всего используется в частном случае $y=1$. Рассмотрим этот случай явно, записав z вместо x, чтобы подчеркнуть тот факт, что здесь может использоваться произвольное комплексное число:

$$(1+z)^r = \sum_k \binom{r}{k} z^k, \qquad |z| < 1. \tag{5.13}$$

При подстановке $z = x/y$ и умножении обеих частей на y^r получается общая формула (5.12).

Мы доказали биномиальную теорему при помощи комбинаторной интерпретации только для неотрицательных целых r. Мы не можем вывести общую формулу из случая для неотрицательных целых чисел при помощи полиномиального доказательства, так как в общем случае сумма бесконечна. Но при произвольном r можно воспользоваться рядом Тейлора и теорией функций комплексного переменного:

$$\begin{aligned} f(z) &= \frac{f(0)}{0!}z^0 + \frac{f'(0)}{1!}z^1 + \frac{f''(0)}{2!}z^2 + \cdots = \\ &= \sum_{k \geqslant 0} \frac{f^{(k)}(0)}{k!} z^k. \end{aligned}$$

Производные функции $f(z) = (1+z)^r$ легко вычисляются: $f^{(k)}(z) = r^{\underline{k}}(1+z)^{r-k}$. Подстановка $z=0$ дает нам (5.13).

Кроме того, нам надо доказать, что бесконечная сумма сходится при $|z| < 1$. Это действительно так, потому что $\binom{r}{k} = O(k^{-1-r})$ согласно равенству (5.83), приведенному ниже.

Смысл O-обозначения раскрывается в главе 9.

Теперь рассмотрим более пристально значения $\binom{n}{k}$ при отрицательных целых n. Один из подходов к этим величинам состоит в использовании правила сложения (5.8) для заполнения табл. 201 числами, получая таким образом табл. 212. Например, $\binom{-1}{0} = 1$, поскольку $\binom{0}{0} = \binom{-1}{0} + \binom{-1}{-1}$ и $\binom{-1}{-1} = 0$; затем $\binom{-1}{1} = -1$, поскольку $\binom{0}{1} = \binom{-1}{1} + \binom{-1}{0}$; и т.д.

Все эти числа нам знакомы. Строки и столбцы табл. 212 оказываются столбцами табл. 201 (минус знак "минус"). Так что

Таблица 212. Треугольник Паскаля: продолжение вверх

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$	$\binom{n}{10}$
−4	1	−4	10	−20	35	−56	84	−120	165	−220	286
−3	1	−3	6	−10	15	−21	28	−36	45	−55	66
−2	1	−2	3	−4	5	−6	7	−8	9	−10	11
−1	1	−1	1	−1	1	−1	1	−1	1	−1	1
0	1	0	0	0	0	0	0	0	0	0	0

между величинами $\binom{n}{k}$ при отрицательных и положительных n должна иметься некоторая связь. Общее правило таково:

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}, \qquad \text{целое } k, \tag{5.14}$$

что легко доказать, поскольку

$$\begin{aligned} r^{\underline{k}} &= r(r-1)\ldots(r-k+1) = \\ &= (-1)^k(-r)(1-r)\ldots(k-1-r) = (-1)^k(k-r-1)^{\underline{k}} \end{aligned}$$

при $k \geqslant 0$, а при $k < 0$ обе части равны нулю.

Особенно ценно то, что соотношение (5.14) выполняется безо всяких ограничений (конечно, нижний индекс должен быть целым, чтобы были определены биномиальные коэффициенты). Преобразование (5.14) называется *обращением верхнего индекса* или просто верхним обращением.

Но как запомнить эту важную формулу? Прочие соотношения, с которыми мы имели дело — симметрии, внесения, сложения и т.д., — были достаточно простыми, то это выглядит весьма замысловато. Тем не менее имеется мнемоническое правило, что уже не так плохо. Для обращения верхнего индекса начнем с записи $(-1)^k$, где k — нижний индекс (который не должен изменяться). Затем мы сразу же записываем k дважды: на места нижнего и верхнего индексов. Затем мы обращаем исходный верхний индекс путем его *вычитания* из нового верхнего индекса. Завершается работа еще одним *вычитанием* 1 (всегда выполняется вычитание, а не сложение, — в конце-концов, ведь это процесс обращения).

Мнемоническое? Скорее, пневматическое — просто сотрясение воздуха. Тем не менее мне оно помогло.

Давайте для практики обратим два раза подряд верхний индекс. Мы получим

Можно размяться и при помощи упр. 4.

$$\begin{aligned} \binom{r}{k} &= (-1)^k \binom{k-r-1}{k} = \\ &= (-1)^{2k} \binom{k-(k-r-1)-1}{k} = \binom{r}{k}, \end{aligned}$$

так что мы вернулись к тому, с чего начинали. Вряд ли мы стремились именно к этому, но зато это доказывает, что мы все сделали верно, что приятно!

И грустно, если мы хотели добиться чего-то другого.

Конечно, есть и более полезные приложения (5.14), чем показанное. Можно использовать верхнее обращение, например, для перемещения величин между верхним и нижним индексами. Это соотношение имеет симметричный вид:

$$(-1)^m\binom{-n-1}{m} = \\ = (-1)^n\binom{-m-1}{n}, \quad \text{целые } m,n \geqslant 0. \qquad (5.15)$$

Оно справедливо, поскольку обе его части равны $\binom{m+n}{n}$.

Верхнее обращение может также использоваться для вывода интересной формулы суммирования:

$$\sum_{k\leqslant m}\binom{r}{k}(-1)^k = \binom{r}{0}-\binom{r}{1}+\cdots+(-1)^m\binom{r}{m} = \\ = (-1)^m\binom{r-1}{m}, \qquad \text{целое } m. \qquad (5.16)$$

Идея заключается в том, чтобы обратить верхний индекс, применить (5.9) и вновь выполнить верхнее обращение:

(Здесь двойное обращение помогает, поскольку между обращениями выполняется еще одна операция.)

$$\sum_{k\leqslant m}\binom{r}{k}(-1)^k = \sum_{k\leqslant m}\binom{k-r-1}{k} = \binom{-r+m}{m} = \\ = (-1)^m\binom{r-1}{m}.$$

Эта формула дает частичную сумму r-й строки треугольника Паскаля в случае, когда элементам строки приписаны чередующиеся знаки. Например, при $r=5$ и $m=2$ формула дает $1-5+10=6=(-1)^2\binom{4}{2}$.

Заметим, что, если $m \geqslant r$, (5.16) дает знакочередующуюся сумму всей строки, а она равна нулю, если r представляет собой положительное целое число. Мы уже доказывали это ранее, когда раскладывали $(1-1)^r$ в соответствии с биномиальной теоремой. Интересно, что частичные суммы такого разложения вычисляются в аналитическом виде.

А что можно сказать о более простой частичной сумме

$$\sum_{k\leqslant m}\binom{n}{k} = \binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{m}? \qquad (5.17)$$

Наверняка, если мы в состоянии вычислить знакочередующуюся сумму, то справимся и с этой? Увы, это не так: для частичной суммы строки треугольника Паскаля нет соответствующего аналитического выражения. Можно посчитать суммы элементов столбцов (формула (5.10)), но не строк. Забавно, однако, что если умножить элементы на их расстояние до центра строки, то частичные суммы можно вычислить аналитически:

$$\sum_{k\leqslant m}\binom{r}{k}\left(\frac{r}{2}-k\right) = \frac{m+1}{2}\binom{r}{m+1}, \qquad \text{целое } m. \tag{5.18}$$

(Эта формула легко проверяется при помощи индукции по m.) Взаимосвязь между этими частичными суммами — со множителем $(r/2-k)$ в общем члене и без него — напоминает взаимосвязь между интегралами

$$\int_{-\infty}^{\alpha} xe^{-x^2}\,dx = -\tfrac{1}{2}e^{-\alpha^2} \qquad \text{и} \qquad \int_{-\infty}^{\alpha} e^{-x^2}\,dx\,.$$

Кажущийся более сложным интеграл слева, с множителем x, выражается в аналитическом виде, в то время как выглядящий более просто интеграл справа аналитического вида не имеет. Внешность обманчива!

На самом деле он равен $\frac{1}{2}\sqrt{\pi}(1+\operatorname{erf}\alpha)$ — константе плюс величине, кратной "функции ошибок" от α — если это можно считать аналитическим видом.

Ближе к концу этой главы мы изучим метод, позволяющий определить, выражаются ли в аналитическом виде частичные суммы некоторого ряда с биномиальными коэффициентами, в достаточно общей постановке задачи. Этот метод в состоянии обнаружить формулы (5.16) и (5.18), а также позволяет выяснить, что нет смысла тратить время на поиск аналитической записи (5.17).

Частичные суммы биномиального ряда приводят к еще одному занимательному соотношению:

$$\sum_{k\leqslant m}\binom{m+r}{k}x^k y^{m-k} = \sum_{k\leqslant m}\binom{-r}{k}(-x)^k(x+y)^{m-k}, \quad \text{целое } m. \tag{5.19}$$

Это тождество нетрудно доказать по индукции: обе части равны нулю при $m<0$ и 1 — при $m=0$. Если обозначить сумму в левой части через S_m, то можно применить формулу сложения (5.8), а это позволяет легко показать, что

$$S_m = \sum_{k\leqslant m}\binom{m-1+r}{k}x^k y^{m-k} + \sum_{k\leqslant m}\binom{m-1+r}{k-1}x^k y^{m-k}$$

и

$$\sum_{k\leqslant m}\binom{m-1+r}{k}x^k y^{m-k} = yS_{m-1}+\binom{m-1+r}{m}x^m,$$

$$\sum_{k\leqslant m}\binom{m-1+r}{k-1}x^k y^{m-k} = xS_{m-1}$$

при $m > 0$. Следовательно,

$$S_m = (x+y)S_{m-1}+\binom{-r}{m}(-x)^m,$$

а этой рекуррентности удовлетворяет и правая часть (5.19). По индукции обе части должны быть равны. QED

Но есть и более тонкое доказательство. Если r — целое число в диапазоне $0 \geqslant r \geqslant -m$, то биномиальная теорема гласит, что обе части (5.19) равны $(x+y)^{m+r}y^{-r}$. А поскольку обе части представляют собой полиномы от r степени m или меньшей, то их совпадения в $m+1$ разных точках достаточно (но едва-едва!) для доказательства равенства в общем случае.

Может показаться глупым получать соотношение, в котором одна сумма равна другой, причем эти суммы к тому же не выражаются в аналитическом виде. Но иногда вычислить одну сумму проще, чем другую. Например, если положить $x=-1$ и $y=1$, то можно получить альтернативную форму соотношения (5.16):

$$\sum_{k\leqslant m}\binom{m+r}{k}(-1)^k = \binom{-r}{m}, \qquad \text{целое } m \geqslant 0.$$

А если положить $x=y=1$ и $r=m+1$, то мы получим

$$\sum_{k\leqslant m}\binom{2m+1}{k} = \sum_{k\leqslant m}\binom{m+k}{k}2^{m-k}.$$

В левой части просуммирована только половина биномиальных коэффициентов с верхним индексом $2m+1$, а эти коэффициенты равны своим двойникам в правой половине, поскольку треугольник Паскаля обладает лево-правосторонней симметрией. Следовательно, левая часть — это просто $\frac{1}{2}2^{2m+1}=2^{2m}$. Это дает нам формулу, которую мы никак не ожидали получить:

Существует красивое комбинаторное доказательство этой формулы [247].

$$\sum_{k\leqslant m}\binom{m+k}{k}2^{-k} = 2^m, \qquad \text{целое } m \geqslant 0. \tag{5.20}$$

Давайте проверим ее при $m=2$: $\binom{2}{0}+\frac{1}{2}\binom{3}{1}+\frac{1}{4}\binom{4}{2} = 1+\frac{3}{2}+\frac{6}{4}=4$. Поразительно!

До сих пор мы рассматривали либо отдельные биномиальные коэффициенты, либо суммы, в члены которых входило только по одному биномиальному коэффициенту. Но многие задачи, с которыми приходится сталкиваться, включают в себя произведения двух или более биномиальных коэффициентов, так что мы посвятим оставшуюся часть главы рассмотрению случаев такого рода.

Вот удобное правило, которое зачастую помогает упростить произведение двух биномиальных коэффициентов:

$$\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}, \qquad \text{целые } m,\, k. \tag{5.21}$$

Мы уже знакомы с частным случаем этого правила $k = 1$; это — правило внесения (5.6). Хотя обе части (5.21) представляют собой произведения биномиальных коэффициентов, одна часть зачастую суммируется значительно проще другой из-за взаимодействия с остальной частью формулы. Например, в левой части дважды используется m, в то время как в правой оно встречается только один раз. Поэтому при суммировании по m мы обычно стремимся заменить $\binom{r}{m}\binom{m}{k}$ на $\binom{r}{k}\binom{r-k}{m-k}$.

Равенство (5.21) выполняется, в первую очередь, из-за сокращения $m!$ в факториальных представлениях $\binom{r}{m}$ и $\binom{m}{k}$. Если все переменные — целые числа и $r \geqslant m \geqslant k \geqslant 0$, то

$$\begin{aligned}\binom{r}{m}\binom{m}{k} &= \frac{r!}{m!\,(r-m)!}\,\frac{m!}{k!\,(m-k)!} = \\ &= \frac{r!}{k!\,(m-k)!\,(r-m)!} = \\ &= \frac{r!}{k!\,(r-k)!}\,\frac{(r-k)!}{(m-k)!\,(r-m)!} = \binom{r}{k}\binom{r-k}{m-k}.\end{aligned}$$

Это просто. Кроме того, если $m < k$ или $k < 0$, то обе части (5.21) равны нулю; так что это соотношение справедливо для всех целых m и k. Наконец, полиномиальное доказательство распространяет его на все действительные r.

Даже очень.

Биномиальный коэффициент $\binom{r}{k} = r!/(r-k)!\,k!$ после соответствующего переобозначения переменных можно записать в виде $(a+b)!/a!\,b!$. Аналогично величину $r!/k!\,(m-k)!\,(r-m)!$ в средине приведенного выше вывода можно записать как $(a+b+c)!/a!\,b!\,c!$. Это — "триномиальный коэффициент", появляющийся в "триномиальной теореме":

$$(x+y+z)^n = \sum_{\substack{0\leqslant a,b,c\leqslant n\\ a+b+c=n}} \frac{(a+b+c)!}{a!\,b!\,c!}x^a y^b z^c = $$
$$= \sum_{\substack{0\leqslant a,b,c\leqslant n\\ a+b+c=n}} \binom{a+b+c}{b+c}\binom{b+c}{c} x^a y^b z^c .$$

Так что $\binom{r}{m}\binom{m}{k}$ на самом деле представляет собой триномиальный коэффициент в ином обличьи. Триномиальные коэффициенты время от времени возникают в разных приложениях и обычно записываются как

$$\binom{a+b+c}{a,\ b,\ c} = \frac{(a+b+c)!}{a!\,b!\,c!},$$

чтобы подчеркнуть наличие симметрии.

"Excogitavi autem olim mirabilem regulam pro numeris coefficientibus potestatum, non tantum a binomio $x+y$, sed et a trinomio $x+y+z$, imo a polynomio quocunque, ut data potentia gradus cujuscunque v. gr. decimi, et potentia in ejus valore comprehensa, ut $x^5y^3z^2$, possim statim assignare numerum coefficientem, quem habere debet, sine ulla Tabula jam calculata".

— Г. Лейбниц (G. W. Leibniz) [245]

Обобщением биномиальных и триномиальных коэффициентов являются *мультиномиальные коэффициенты*, которые всегда выражаются как произведения биномиальных коэффициентов:

$$\binom{a_1+a_2+\cdots+a_m}{a_1, a_2, \ldots, a_m} = \frac{(a_1+a_2+\cdots+a_m)!}{a_1!\,a_2!\,\ldots\,a_m!} =$$
$$= \binom{a_1+a_2+\cdots+a_m}{a_2+\cdots+a_m}\cdots\binom{a_{m-1}+a_m}{a_m}.$$

Таким образом, при встрече с таким ужасом нам помогут наши стандартные методы.

Так мы добрались до табл. 218, в которой перечислены наиболее важные из рассмотренных стандартных методов. Это те методы, на которые мы опираемся при работе с суммами, включающими произведение биномиальных коэффициентов. Каждое из приведенных равенств представляет собой сумму по k, где k входит по одному разу в каждый биномиальный коэффициент; кроме того, имеются четыре почти независимых параметра, обозначаемые как m, n, r и т.д., по одному в каждом индексе. В зависимости от того, входит ли k в верхний или нижний индекс и с каким знаком, мы получаем различные случаи. Местами имеется дополнительный множитель $(-1)^k$, необходимый для вычислимости суммы в аналитическом виде.

Положите здесь закладку или загните уголок страницы, чтобы быстро найти эту таблицу. Она еще не раз вам понадобится!

Табл. 218 слишком сложна, чтобы всю ее запомнить; ее назначение — служить для справок. Первое соотношение в таблице запоминается гораздо легче других, и его стоит запомнить. Оно

Таблица 218. Суммы произведений биномиальных коэффициентов

$$\sum_k \binom{r}{m+k}\binom{s}{n-k} = \binom{r+s}{m+n}, \quad \text{целые } m, n \tag{5.22}$$

$$\sum_k \binom{l}{m+k}\binom{s}{n+k} = \binom{l+s}{l-m+n}, \quad \begin{array}{l}\text{целое } l \geqslant 0,\\ \text{целые } m, n\end{array} \tag{5.23}$$

$$\sum_k \binom{l}{m+k}\binom{s+k}{n}(-1)^k = (-1)^{l+m}\binom{s-m}{n-l}, \quad \begin{array}{l}\text{целое } l \geqslant 0,\\ \text{целые } m, n\end{array} \tag{5.24}$$

$$\sum_{k\leqslant l} \binom{l-k}{m}\binom{s}{k-n}(-1)^k = (-1)^{l+m}\binom{s-m-1}{l-m-n}, \quad \begin{array}{c}\text{целые}\\ l, m, n \geqslant 0\end{array} \tag{5.25}$$

$$\sum_{-q\leqslant k\leqslant l} \binom{l-k}{m}\binom{q+k}{n} = \binom{l+q+1}{m+n+1}, \quad \begin{array}{l}\text{целые } m, n \geqslant 0,\\ \text{целое } l+q \geqslant 0\end{array} \tag{5.26}$$

гласит, что сумма (по всем целым k) произведения двух биномиальных коэффициентов, у которых верхние индексы постоянны, а нижние при любом k имеют одну и ту же сумму, представляет собой биномиальный коэффициент, получающийся суммированием нижних и верхних индексов. Это тождество известно как *свертка Вандермонда*, поскольку Александр Вандермонд (Alexandre Vandermonde) написал о нем важную статью в конце XVIII века [357]; однако оно было известно Чу Ши-цзе в Китае в 1303 году. Все остальные соотношения в табл. 218 могут быть получены из свертки Вандермонда аккуратным обращением верхних индексов или применением правил симметрии и т.п.; таким образом, свертка Вандермонда — наиболее фундаментальное соотношение среди приведенных.

Доказать свертку Вандермонда можно исходя из красивой комбинаторной интерпретации. Если заменить k на $k-m$ и n на $n-m$, можно считать, что $m=0$; следовательно, требуется доказать

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, \qquad \text{целое } n. \tag{5.27}$$

Пусть r и s — неотрицательные целые числа (общий случай получается при помощи полиномиального доказательства). В правой части $\binom{r+s}{n}$ — количество способов выбрать n человек среди r

Феминисткам может не понравиться то, что мужчины упомянуты первыми...

Им точно так же может не понравиться, если первыми будут упомянуты женщины...

мужчин и s женщин. В левой части каждый член суммы представляет собой количество способов выбрать k мужчин и $n-k$ женщин. Суммирование по всем k учитывает каждую возможность ровно по одному разу.

Чаще всего эти правила применяются слева направо, поскольку упрощение выполняется именно в этом направлении. Однако бывает и так, что оправданно движение и в обратном направлении, даже если приходится временно идти на усложнение. При этом обычно образуется двойная сумма, в которой можно изменить порядок суммирования, а после выполнить упрощение.

Перед тем как идти дальше, посмотрим, как доказываются еще два соотношения из табл. 218. Легко доказать соотношение (5.23); все, что нам надо для этого, — заменить первый биномиальный коэффициент на $\binom{l}{l-m-k}$, а затем применить свертку Вандермонда (5.22).

Следующее соотношение (5.24) несколько сложнее. Путем последовательных преобразований его можно свести к свертке Вандермонда, но можно доказать его столь же просто, прибегнув к методу математической индукции. Зачастую индукция — первое, к чему мы прибегаем, если не можем придумать ничего лучшего; однако здесь индукция по l подходит как нельзя лучше.

В качестве базиса индукции берем случай $l=0$, когда все члены равны нулю, за исключением $k=-m$; таким образом, обе части равенства равны $(-1)^m\binom{s-m}{n}$. Теперь предположим, что соотношение выполняется при всех значениях, меньших некоторого фиксированного $l>0$. Мы можем воспользоваться формулой сложения, чтобы заменить $\binom{l}{m+k}$ на $\binom{l-1}{m+k}+\binom{l-1}{m+k-1}$; тогда исходная сумма разбивается на две, каждая из которых может быть вычислена на основе гипотезы индукции:

$$\sum_k \binom{l-1}{m+k}\binom{s+k}{n}(-1)^k+\sum_k \binom{l-1}{m+k-1}\binom{s+k}{n}(-1)^k =$$
$$= (-1)^{l-1+m}\binom{s-m}{n-l+1}+(-1)^{l+m}\binom{s-m+1}{n-l+1}.$$

Если применить формулу сложения еще раз, то выражение упрощается и получается правая часть (5.24).

В этом выводе заслуживают упоминания две вещи. Во-первых, мы снова убедились в удобстве суммирования по всем целым k, а не только в определенном диапазоне, поскольку при этом нам не надо беспокоиться о граничных условиях. Во-вторых, формула сложения хорошо работает вместе с математической индукцией, поскольку она представляет собой рекуррентное соотношение для биномиальных коэффициентов. Биномиальный коэффициент, верхний индекс которого l, выражается через два биномиальных коэффициента, верхние индексы которых равны $l-1$, и это именно то, что нам надо для применения математической индукции.

Пожалуй, на этом оставим табл. 218 в покое. А что можно сказать о суммах с тремя и более биномиальными коэффициентами? Если индекс суммирования имеется во всех коэффициентах, то шансы получить решение в аналитическом виде малы: для сумм такого рода известны лишь несколько аналитических выражений, так что интересующая нас сумма вполне может не соответствовать предъявленным требованиям. Один из таких раритетов (о котором говорится в упр. 43) перед вами:

$$\sum_k \binom{m-r+s}{k}\binom{n+r-s}{n-k}\binom{r+k}{m+n} = \binom{r}{m}\binom{s}{n}, \qquad \text{целые } m, n. \tag{5.28}$$

Вот еще один, несколько более симметричный, пример:

$$\sum_k \binom{a+b}{a+k}\binom{b+c}{b+k}\binom{c+a}{c+k}(-1)^k = \frac{(a+b+c)!}{a!\,b!\,c!}, \qquad \text{целые } a, b, c \geqslant 0. \tag{5.29}$$

У него есть двойник с двумя коэффициентами:

$$\sum_k \binom{a+b}{a+k}\binom{b+a}{b+k}(-1)^k = \frac{(a+b)!}{a!\,b!}, \qquad \text{целые } a, b \geqslant 0, \tag{5.30}$$

которого, кстати, нет в табл. 218. Аналогичная сумма с четырьмя коэффициентами в аналитическом виде не выражается, но зато можно выразить другую подобную ей сумму:

$$\sum_k (-1)^k \binom{a+b}{a+k}\binom{b+c}{b+k}\binom{c+d}{c+k}\binom{d+a}{d+k}\Big/\binom{2a+2b+2c+2d}{a+b+c+d+k} = \frac{(a+b+c+d)!\,(a+b+c)!\,(a+b+d)!\,(a+c+d)!\,(b+c+d)!}{(2a+2b+2c+2d)!\,(a+c)!\,(b+d)!\,a!\,b!\,c!\,d!},$$
$$\text{целые } a, b, c, d \geqslant 0.$$

Это выражение следует из соотношения с пятью параметрами, открытого Джоном Дугаллом (John Dougall) [82] в начале XX века.

Но и тождество Дугалла — не самое пышное из известных сумм биномиальных коэффициентов. Пальму первенства пока что удерживает сумма

$$\sum_{k_{ij}} (-1)^{\sum_{i<j} k_{ij}} \left(\prod_{1 \leqslant i<j<n} \binom{a_i+a_j}{a_j+k_{ij}} \right) \left(\prod_{1 \leqslant j<n} \binom{a_j+a_n}{a_n+\sum_{i<j} k_{ij} - \sum_{i>j} k_{ji}} \right) =$$

$$= \binom{a_1+\cdots+a_n}{a_1, a_2, \ldots, a_n}, \quad \text{целые } a_1, a_2, \ldots, a_n \geqslant 0. \qquad (5.31)$$

Это сумма по $\binom{n-1}{2}$ индексным переменным k_{ij} при $1 \leqslant i < j < n$. Уравнение (5.29) представляет собой частный случай данного равенства при $n = 3$; случай $n = 4$ может быть записан при подстановке (a, b, c, d) вместо (a_1, a_2, a_3, a_4) и (i, j, k) вместо (k_{12}, k_{13}, k_{23}) в следующем виде:

$$\sum_{i,j,k} (-1)^{i+j+k} \binom{a+b}{b+i} \binom{a+c}{c+j} \binom{b+c}{c+k} \binom{a+d}{d-i-j} \binom{b+d}{d+i-k} \binom{c+d}{d+j+k} =$$

$$= \frac{(a+b+c+d)!}{a!\, b!\, c!\, d!}, \qquad \text{целые } a, b, c, d \geqslant 0.$$

Левая часть (5.31) представляет собой коэффициент при $z_1^0 z_2^0 \ldots z_n^0$ после полного разложения произведения $n(n-1)$ дробей

$$\prod_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} \left(1 - \frac{z_i}{z_j}\right)^{a_i}$$

по положительным и отрицательным степеням z. Предположение о правой части (5.31) было высказано Фриманом Дайсоном (Freeman Dyson) в 1962 году и вскоре доказано сразу несколькими исследователями. В упр. 86 приведено "простое" доказательство (5.31).

Заслуживает внимания еще одно тождество, включающее массу биномиальных коэффициентов:

$$\sum_{j,k} (-1)^{j+k} \binom{j+k}{k+l} \binom{r}{j} \binom{n}{k} \binom{s+n-j-k}{m-j} =$$

$$= (-1)^l \binom{n+r}{n+l} \binom{s-r}{m-n-l}, \quad \text{целые } l, m, n;\ n \geqslant 0. \qquad (5.32)$$

У данного тождества, доказанного в упр. 83, есть шанс встретиться в практическом приложении. Но мы отклоняемся от нашей темы "основные тождества", так что лучше остановиться и подвести итоги.

Таблица 222. Десять главных тождеств с биномиальными коэффициентами

$\dbinom{n}{k} = \dfrac{n!}{k!\,(n-k)!}\,,$	целые $n \geqslant k \geqslant 0$	*Факториальное представление*
$\dbinom{n}{k} = \dbinom{n}{n-k}\,,$	целое $n \geqslant 0$, целое k	*Симметрия*
$\dbinom{r}{k} = \dfrac{r}{k}\dbinom{r-1}{k-1}\,,$	целое $k \neq 0$	*Внесение/вынесение*
$\dbinom{r}{k} = \dbinom{r-1}{k} + \dbinom{r-1}{k-1}\,,$	целое k	*Сложение/разложение*
$\dbinom{r}{k} = (-1)^k\dbinom{k-r-1}{k}\,,$	целое k	*Верхнее обращение*
$\dbinom{r}{m}\dbinom{m}{k} = \dbinom{r}{k}\dbinom{r-k}{m-k}\,,$	целые m, k	*Триномиальный вариант*
$\sum\limits_k \dbinom{r}{k} x^k y^{r-k} = (x+y)^r\,,$	целое $r \geqslant 0$, или $\lvert x/y\rvert < 1$	*Биномиальная теорема*
$\sum\limits_{k\leqslant n} \dbinom{r+k}{k} = \dbinom{r+n+1}{n}\,,$	целое n	*Параллельное суммирование*
$\sum\limits_{0\leqslant k\leqslant n} \dbinom{k}{m} = \dbinom{n+1}{m+1}\,,$	целые $m, n \geqslant 0$	*Верхнее суммирование*
$\sum\limits_k \dbinom{r}{k}\dbinom{s}{n-k} = \dbinom{r+s}{n}\,,$	целое n	*Свертка Вандермонда*

Мы убедились, что биномиальные коэффициенты удовлетворяют такому разнообразию тождеств, которое способно сбить с толку и привести в изумление. К счастью, некоторые из них несложно запомнить, и этим можно воспользоваться, чтобы выводить остальные за несколько шагов. В табл. 222 приведены десять наиболее полезных формул; это именно те формулы, которые сто́ит запомнить.

5.2 Необходимые навыки

В предыдущем разделе мы вывели ряд соотношений, работая с суммами и подставляя их в другие соотношения. Это было не слишком сложное занятие, так как мы знали, что именно хотим доказать, так что можно было сформулировать общий план и наполнить его конкретным содержимым без особых сложностей. Однако обычно в реальном мире мы сталкиваемся не с необходимостью доказать некоторое соотношение, а с необходимостью упростить некоторую сумму. При этом мы не знаем, как должна выглядеть эта упрощенная сумма (и даже существует ли она). В этом и следующем разделах мы поработаем с такими суммами и тем самым отточим свое мастерство в обращении с биномиальными коэффициентами.

Для начала попытаемся поработать с несколькими суммами, содержащими по одному биномиальному коэффициенту.

Алгоритм <u>самообучения</u>:
1 прочесть задачу
2 попытаться решить
3 взглянуть на решение в книге
4 <u>if</u> попытка неудачна <u>goto</u> 1
<u>else</u> <u>goto</u> следующая задача

Задача 1: сумма отношений

Мы хотим найти аналитический вид суммы

$$\sum_{k=0}^{m} \binom{m}{k} \Big/ \binom{n}{k}, \qquad \text{целые } n \geqslant m \geqslant 0.$$

Первый взгляд, брошенный на эту сумму, вызывает панику, потому что до сих пор мы не работали с соотношениями, в которых присутствовало бы отношение биномиальных коэффициентов. (Более того, сумма включает два биномиальных коэффициента, что вроде бы противоречит утверждению, предшествующему этой задаче.) Однако точно так же, как мы использовали факториальные представления, чтобы представить одно произведение биномиальных коэффициентов в виде другого произведения (именно так было получено соотношение (5.21)), можно поступить и в случае отношения биномиальных коэффициентов. На самом деле мы можем избежать факториальных представлений, положив $r = n$ и разделив обе части равенства (5.21) на $\binom{n}{k}\binom{n}{m}$; это дает нам

К сожалению, этот алгоритм может привести к зацикливанию.

Предлагаемое исправление:
0 <u>set</u> $c \leftarrow 0$
3a <u>set</u> $c \leftarrow c + 1$
3b <u>if</u> $c = N$ <u>goto</u> преподаватель

$$\binom{m}{k} \Big/ \binom{n}{k} = \binom{n-k}{m-k} \Big/ \binom{n}{m}.$$

— Э. Дейкстра (E. W. Dijkstra)

Таким образом, мы заменяем отношение слева отношением справа, и сумма приобретает вид

$$\sum_{k=0}^{m} \binom{n-k}{m-k} \Big/ \binom{n}{m}.$$

У нас все еще остается отношение биномиальных коэффициентов, но в знаменателе уже нет индекса суммирования k, так что его можно удалить из суммы (мы вернем его позже).

... но этот подраздел называется “НЕ ОБХОДИМЫЕ навыки”.

Можно упростить и граничные условия, суммируя по всем $k \geqslant 0$; члены с $k > m$ равны нулю. Оставшаяся сумма уже не так страшна:

$$\sum_{k\geqslant 0} \binom{n-k}{m-k}.$$

Она аналогична сумме из тождества (5.9), поскольку индекс k присутствует в двух местах с одним и тем же знаком. Но здесь он $-k$, а в (5.9) — нет. Поэтому следующий шаг очевиден, и другого пути у нас нет:

$$\sum_{k\geqslant 0} \binom{n-k}{m-k} = \sum_{m-k\geqslant 0} \binom{n-(m-k)}{m-(m-k)} = \\ = \sum_{k\leqslant m} \binom{n-m+k}{k}.$$

Теперь можно применить соотношение параллельного суммирования по обоим индексам (5.9):

$$\sum_{k\leqslant m} \binom{n-m+k}{k} = \binom{(n-m)+m+1}{m} = \binom{n+1}{m}.$$

В заключение мы восстанавливаем знаменатель $\binom{n}{m}$, который был удален из суммы ранее, и затем применяем (5.7), чтобы получить требуемый аналитический вид:

$$\binom{n+1}{m} \Big/ \binom{n}{m} = \frac{n+1}{n+1-m}.$$

Этот вывод на самом деле справедлив для любого действительного n, если только не выполняется деление на нуль, т.е. если n не является одним из целых чисел $0, 1, \ldots, m-1$.

Чем сложнее выкладки, тем важнее проверка получившегося результата. Хотя наш вывод был не слишком сложен, все же проверим его. При малых значениях $m = 2$ и $n = 4$ мы получаем

$$\binom{2}{0} \Big/ \binom{4}{0} + \binom{2}{1} \Big/ \binom{4}{1} + \binom{2}{2} \Big/ \binom{4}{2} = 1 + \frac{1}{2} + \frac{1}{6} = \frac{5}{3},$$

что совпадает с величиной $(4+1)/(4+1-2)$, вычисленной по полученной нами формуле.

Задача 2: из литературы по методам сортировки

Следующая сумма относится к седой древности — прошлому тысячелетию (а точнее, к началу 1970-х годов), когда навыки работы с биномиальными коэффициентами не были столь распространены среди исследователей. Одна статья, посвященная улучшенному методу слияния [196], завершалась следующим примечанием: "можно показать, что ожидаемое количество сэкономленных пересылок данных... дается выражением

$$T = \sum_{r=0}^{n} r \, \frac{{}_{m-r-1}C_{m-n-1}}{{}_{m}C_{n}}.$$

Величины m и n определены выше, а ${}_mC_n$ представляет собой обозначение количества сочетаний из m элементов по n... Автор благодарит рецензента за приведение более сложного выражения для среднего количества пересылок к приведенному здесь виду".

Ох, не напоминайте о коллоквиуме, пожалуйста...

Как мы увидим, это, определенно, не окончательный ответ на поставленную задачу, и не годится даже для коллоквиума.

Сначала следует привести эту сумму к виду, с которым можно работать. Одного обозначения ${}_{m-r-1}C_{m-n-1}$ достаточно для того, чтобы остановить любого, кроме разве что энтузиаста-рецензента (не стоит благодарности). В наших обозначениях сумма имеет вид

$$T = \sum_{k=0}^{n} k\binom{m-k-1}{m-n-1} \Big/ \binom{m}{n}, \qquad \text{целые } m > n \geqslant 0.$$

Биномиальный коэффициент в знаменателе не содержит индекс суммирования, так что можно его удалить и рассматривать новую сумму

$$S = \sum_{k=0}^{n} k\binom{m-k-1}{m-n-1}.$$

Что же дальше? Индекс суммирования входит только в верхний индекс биномиального коэффициента. Так что, если бы в формуле у нас не было еще одного k, мы могли бы упростить данную сумму и применить суммирование по верхнему индексу (5.10). Но дополнительное k не позволяет нам так поступить. Если бы мы могли каким-то образом внести k в биномиальный коэффициент, воспользовавшись одним из наших соотношений внесения, то затем можно было бы выполнить суммирование по верхнему индексу. К сожалению, имеющиеся у нас правила здесь не работают. Но если бы вместо k было $m-k$, то можно было бы воспользоваться соотношением внесения (5.6):

$$(m-k)\binom{m-k-1}{m-n-1} = (m-n)\binom{m-k}{m-n}.$$

Вот он, ключ: перепишем k как $m-(m-k)$ и разделим сумму S на две:

$$\begin{aligned}\sum_{k=0}^{n} k\binom{m-k-1}{m-n-1} &= \sum_{k=0}^{n}(m-(m-k))\binom{m-k-1}{m-n-1} = \\ &= \sum_{k=0}^{n} m\binom{m-k-1}{m-n-1} - \sum_{k=0}^{n}(m-k)\binom{m-k-1}{m-n-1} = \\ &= m\sum_{k=0}^{n}\binom{m-k-1}{m-n-1} - \sum_{k=0}^{n}(m-n)\binom{m-k}{m-n} = \\ &= mA-(m-n)B,\end{aligned}$$

где

$$A = \sum_{k=0}^{n}\binom{m-k-1}{m-n-1}, \qquad B = \sum_{k=0}^{n}\binom{m-k}{m-n}.$$

Суммы A и B — наши старые знакомые, в которых изменяется верхний индекс, в то время как нижний остается неизменным. Давайте разберемся сначала с суммой B, которая имеет более простой вид. Для того чтобы привести общий член данной суммы к виду левой части (5.10), надо немного ее преобразовать:

$$\begin{aligned}\sum_{0\leqslant k\leqslant n}\binom{m-k}{m-n} &= \sum_{0\leqslant m-k\leqslant n}\binom{m-(m-k)}{m-n} = \\ &= \sum_{m-n\leqslant k\leqslant m}\binom{k}{m-n} = \\ &= \sum_{0\leqslant k\leqslant m}\binom{k}{m-n}.\end{aligned}$$

На последнем шаге в сумму включаются члены с $0\leqslant k<m-n$, которые все равны нулю, поскольку верхний индекс у них оказывается меньше нижнего. Теперь, используя (5.10), просуммируем по верхнему индексу и получим

$$B = \sum_{0\leqslant k\leqslant m}\binom{k}{m-n} = \binom{m+1}{m-n+1}.$$

Другая сумма A почти такая же, с тем отличием, что m в ней заменено на $m-1$. Так мы получаем аналитическое выражение для суммы S, которое можно еще немного упростить:

$$\begin{aligned} S &= mA - (m-n)B = \\ &= m\binom{m}{m-n} - (m-n)\binom{m+1}{m-n+1} = \\ &= \left(m - (m-n)\frac{m+1}{m-n+1}\right)\binom{m}{m-n} = \\ &= \left(\frac{n}{m-n+1}\right)\binom{m}{m-n}. \end{aligned}$$

Это дает нам аналитическое выражение для исходной суммы:

$$\begin{aligned} T &= S\Big/\binom{m}{n} = \frac{n}{m-n+1}\binom{m}{m-n}\Big/\binom{m}{n} = \\ &= \frac{n}{m-n+1}. \end{aligned}$$

Даже рецензент не смог бы его упростить!

Для проверки полученного результата вновь обратимся к малому случаю. При $m = 4$ и $n = 2$ получаем

$$T = 0\cdot\tbinom{3}{1}/\tbinom{4}{2} + 1\cdot\tbinom{2}{1}/\tbinom{4}{2} + 2\cdot\tbinom{1}{1}/\tbinom{4}{2} = 0 + \tfrac{2}{6} + \tfrac{2}{6} = \tfrac{2}{3},$$

что вполне согласуется с вычисленной по нашей формуле величиной $2/(4-2+1)$.

Задача 3: из старого домашнего задания

Да смертны ли вообще старые задачи?!

Вот еще одна сумма с единственным биномиальным коэффициентом. Данная сумма, в отличие от последней, родилась в университетских стенах: это задача из одного из домашних заданий. В ней предлагалось вычислить величину $Q_{1000000}$, где

$$Q_n = \sum_{k\leqslant 2^n}\binom{2^n-k}{k}(-1)^k, \qquad \text{целое } n \geqslant 0.$$

Эта задача потруднее прочих: к ней неприменимо *ни одно* из известных нам к этому времени соотношений. К тому же мы имеем дело с суммой из $2^{1000000}+1$ членов, так что просто просуммировать их нельзя. Индекс суммирования k имеется в обоих индексах биномиального коэффициента, и в верхнем, и в нижнем, но с противоположными знаками. Обращение верхнего индекса также ничего нам не дает — оно удаляет множитель $(-1)^k$, но вносит в верхний индекс $2k$.

Как всегда, когда ничего не получается, лучшее, что можно сделать, — это рассмотреть малые случаи. Даже если мы и не найдем закономерность, которую можно будет доказать по индукции, то хотя бы получим данные для проверки результатов нашей работы. Вот все ненулевые члены и их суммы для первых четырех значений n.

n		Q_n
0	$\binom{1}{0} \qquad = 1$	$= 1$
1	$\binom{2}{0} - \binom{1}{1} \qquad = 1 - 1$	$= 0$
2	$\binom{4}{0} - \binom{3}{1} + \binom{2}{2} \qquad = 1 - 3 + 1$	$= -1$
3	$\binom{8}{0} - \binom{7}{1} + \binom{6}{2} - \binom{5}{3} + \binom{4}{4} = 1 - 7 + 15 - 10 + 1$	$= 0$

За следующий случай, $n = 4$, лучше не браться, чтобы не запутаться и не допустить арифметическую ошибку. (Вычислением членов типа $\binom{12}{4}$ и $\binom{11}{5}$ вручную, не объединяя их с другими, сто́ит заниматься лишь в безвыходной ситуации.)

Итак, значения Q_n начинаются с 1, 0, −1, 0. Даже если бы мы посчитали еще одно-два значения, вряд ли бы аналитический вид Q_n оказался очевидным. Но если бы мы обнаружили и доказали рекуррентное соотношение для Q_n, то тогда, вероятно, мы бы могли предложить и доказать аналитическую формулу. Для поиска рекуррентного соотношения надо связать Q_n с Q_{n-1} (или с $Q_{\text{меньшие значения}}$); но для этого нам надо связать член наподобие $\binom{128-13}{13}$, который получается при $n = 7$ и $k = 13$, с членом типа $\binom{64-13}{13}$. Многообещающим этот путь не назовешь — мы не знаем никаких более-менее простых соотношений между элементами треугольника Паскаля, разделенными 64 строками. Формула сложения — наш основной инструмент при доказательствах по индукции — связывает элементы, отделенные друг от друга одной строкой.

Но это приводит нас к ключевому наблюдению: на самом деле нет необходимости работать с элементами, разделенными 2^{n-1} строками. Переменная n никогда не фигурирует сама по себе, а только в контексте 2^n. Вот оно: 2^n — это просто ловушка для отвода глаз! Если мы заменим 2^n на m, то все, что нам надо будет сделать, — это найти аналитическое выражение для более общей (но и более простой) суммы

О преподаватель, придумавший эту задачу! Коварство — вот имя твое!

$$R_m = \sum_{k \leqslant m} \binom{m-k}{k} (-1)^k, \qquad \text{целое } m \geqslant 0;$$

тем самым мы получим аналитическое выражение и для $Q_n = R_{2^n}$. И у нас неплохие шансы на то, что формула сложения приведет к рекуррентному соотношению для R_m.

Значения R_m для небольших m можно найти в табл. 201, если поочередно складывать и вычитать величины, располагающиеся на диагонали, идущей с юго-запада на северо-восток. Вот что мы получаем:

m	0	1	2	3	4	5	6	7	8	9	10
R_m	1	1	0	−1	−1	0	1	1	0	−1	−1

Похоже, дальше будет масса сокращений.

Давайте теперь обратимся к формуле для R_m и посмотрим, нельзя ли записать для нее рекуррентное соотношение. Наша стратегия состоит в применении формулы сложения (5.8) и поиске в результирующем выражении сумм, имеющих вид R_k, — похоже на то, что мы делали в методе приведения в главе 2:

$$
\begin{aligned}
R_m &= \sum_{k\leqslant m}\binom{m-k}{k}(-1)^k = \\
&= \sum_{k\leqslant m}\binom{m-1-k}{k}(-1)^k+\sum_{k\leqslant m}\binom{m-1-k}{k-1}(-1)^k = \\
&= \sum_{k\leqslant m}\binom{m-1-k}{k}(-1)^k+\sum_{k+1\leqslant m}\binom{m-2-k}{k}(-1)^{k+1} = \\
&= \sum_{k\leqslant m-1}\binom{m-1-k}{k}(-1)^k+\binom{-1}{m}(-1)^m- \\
&\qquad -\sum_{k\leqslant m-2}\binom{m-2-k}{k}(-1)^k-\binom{-1}{m-1}(-1)^{m-1} = \\
&= R_{m-1}+(-1)^{2m}-R_{m-2}-(-1)^{2(m-1)} = R_{m-1}-R_{m-2}\,.
\end{aligned}
$$

(На предпоследнем шаге использована формула $\binom{-1}{m} = (-1)^m$, которая, как мы знаем, справедлива при $m \geqslant 0$.) Сам вывод корректен при $m \geqslant 2$.

Это знает каждый выполнивший разминочное упр. 4.

Из этого рекуррентного соотношения можно быстро генерировать значения R_m и не менее быстро убедиться в периодичности этой последовательности. В самом деле,

$$
R_m = \begin{cases} 1 \\ 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{cases}, \qquad \text{если } m \bmod 6 = \begin{cases} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{cases}.
$$

Доказательство по индукции выполняется путем простой проверки. Но если требуется более академичное доказательство, то можно развернуть данное рекуррентное соотношение на один шаг, получив

$$R_m = (R_{m-2} - R_{m-3}) - R_{m-2} = -R_{m-3},$$

при $m \geqslant 3$. Следовательно, $R_m = R_{m-6}$ при $m \geqslant 6$.

Наконец, поскольку $Q_n = R_{2^n}$, можно найти Q_n, определяя $2^n \bmod 6$ и используя выражение для R_m. При $n = 0$ мы получаем $2^0 \bmod 6 = 1$; умножая на 2 (mod 6), мы получим повторения чисел 2 и 4. Таким образом,

$$Q_n = R_{2^n} = \begin{cases} R_1 = 1, & \text{если } n = 0, \\ R_2 = 0, & \text{если } n > 0 \text{ нечетное}, \\ R_4 = -1, & \text{если } n > 0 \text{ четное}. \end{cases}$$

Это выражение для Q_n согласуется с первыми четырьмя значениями, вычисленными в начале решения данной задачи. Поэтому мы можем заключить, что $Q_{1000000} = R_4 = -1$.

Задача 4: сумма с двумя биномиальными коэффициентами

Наша следующая задача заключается в поиске аналитического выражения для суммы

$$\sum_{k=0}^{n} k\binom{m-k-1}{m-n-1}, \qquad \text{целые } m > n \geqslant 0.$$

Стоп! А где же второй обещанный биномиальный коэффициент? И почему мы должны упрощать сумму, которую мы уже упрощали? (Это же сумма S из задачи 2.)

Это сумма, которую проще упростить, простите за каламбур, если рассматривать ее общий член как произведение двух биномиальных коэффициентов, а затем воспользоваться одним из общих соотношений из табл. 218. Второй биномиальный коэффициент материализуется, если переписать k как $\binom{k}{1}$:

$$\sum_{k=0}^{n} k\binom{m-k-1}{m-n-1} = \sum_{0 \leqslant k \leqslant n} \binom{k}{1}\binom{m-k-1}{m-n-1}.$$

Далее надо воспользоваться соотношением (5.26), поскольку в нем переменная суммирования входит в оба верхних индекса, и при этом с разными знаками.

Но наша сумма еще не приведена к нужному виду. Чтобы достичь полного соответствия формуле (5.26), верхний предел суммирования должен быть $m - 1$. Без проблем: при $n < k \leqslant$

$m-1$ члены равны нулю, поэтому можно выполнить подстановку $(l, m, n, q) \leftarrow (m-1, m-n-1, 1, 0)$; ответ равен

$$S = \binom{m}{m-n+1}.$$

Это несколько аккуратнее, чем полученная ранее формула. Мы можем преобразовать эту формулу в выведенную ранее, воспользовавшись (5.7):

$$\binom{m}{m-n+1} = \frac{n}{m-n+1}\binom{m}{m-n}.$$

Аналогично можно получить интересные результаты, подставляя в общее соотношение определенные значения. Предположим, например, что мы подставляем в (5.26) $m = n = 1$ и $q = 0$. Тогда это соотношение принимает вид

$$\sum_{0 \leqslant k \leqslant l} (l-k)k = \binom{l+1}{3}.$$

Левая часть представляет собой $l\bigl((l+1)l/2\bigr) - (1^2 + 2^2 + \cdots + l^2)$, так что мы получили еще один способ решения задачи о сумме последовательных квадратов, которую до смерти замучили в главе 2.

Мораль данной истории такова: частные случаи важных сумм иногда лучше приводить к общему виду. После изучения сумм в общем виде разумно рассмотреть их простые специализации.

Задача 5: сумма с тремя сомножителями

Вот еще одна неплохая сумма. Мы хотим упростить сумму

$$\sum_k \binom{n}{k}\binom{s}{k}k, \qquad \text{целое } n \geqslant 0.$$

Индекс суммирования k имеется в обоих нижних индексах с одним и тем же знаком; таким образом, тождество (5.23) из табл. 218 очень похоже на то, в чем мы нуждаемся. После небольших манипуляций мы должны суметь им воспользоваться.

Наибольшее отличие между (5.23) и тем, что у нас есть, заключается в лишнем k в нашей сумме. Но можно внести k в один из биномиальных коэффициентов при помощи соотношения внесения:

$$\sum_k \binom{n}{k}\binom{s}{k}k = \sum_k \binom{n}{k}\binom{s-1}{k-1}s = s\sum_k \binom{n}{k}\binom{s-1}{k-1}.$$

Нам не надо беспокоиться об s, которое появляется вместо k, поскольку это константа. Теперь мы готовы применить тождество и получить аналитический вид

$$s\sum_k \binom{n}{k}\binom{s-1}{k-1} = s\binom{n+s-1}{n-1}.$$

Если бы на первом шаге мы внесли k в $\binom{n}{k}$, а не в $\binom{s}{k}$, то не могли бы применить (5.23) непосредственно, так как величина $n-1$ могла бы оказаться отрицательной, а данное тождество требует неотрицательного значения как минимум в одном из верхних индексов.

Задача 6: очень страшная сумма

Следующая сумма выглядит еще более вызывающе. Требуется найти аналитическую запись для суммы

$$\sum_{k\geqslant 0}\binom{n+k}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1}, \qquad \text{целое } n \geqslant 0.$$

Одной из полезных мер сложности суммы служит количество вхождений в нее индекса суммирования. Здесь эта мера повергает в уныние — k встречается шесть раз. Кроме того, ключевой прием, сработавший в предыдущей задаче — внесение чего-то за пределами биномиальных коэффициентов в один из них, — в данном случае не сработает. Если мы внесем $k+1$, то получим вместо него новое вхождение k. И дело не только в этом: наш индекс k дважды входит в биномиальные коэффициенты с коэффициентом 2. Обычно мультипликативные константы удалять гораздо труднее, чем аддитивные.

Так что, надо шесть раз отказаться от решения этой задачи?

Но в этот раз нам повезло. Двойки находятся именно там, где следует для применения тождества (5.21), так что мы получаем

$$\sum_{k\geqslant 0}\binom{n+k}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1} = \sum_{k\geqslant 0}\binom{n+k}{k}\binom{n}{k}\frac{(-1)^k}{k+1}.$$

Двойки исчезли, прихватив с собой одно вхождение k. Итак, мера сложности снизилась до пяти.

Самая большая из оставшихся неприятностей — наличие $k+1$ в знаменателе, но теперь эту величину можно внести в $\binom{n}{k}$ при помощи тождества (5.6):

$$\sum_{k\geqslant 0}\binom{n+k}{k}\binom{n}{k}\frac{(-1)^k}{k+1} = \sum_k \binom{n+k}{k}\binom{n+1}{k+1}\frac{(-1)^k}{n+1} =$$
$$= \frac{1}{n+1}\sum_k \binom{n+k}{k}\binom{n+1}{k+1}(-1)^k.$$

(Вспомним, что $n \geqslant 0$.) Итак, осталось четыре k.

Для удаления прочих k у нас есть два перспективных пути. Можно воспользоваться симметрией $\binom{n+k}{k}$, а можно обратить верхний индекс $n+k$, тем самым удалив и k, и множитель $(-1)^k$. Давайте испытаем оба варианта, начав с симметрии:

$$\frac{1}{n+1}\sum_k \binom{n+k}{k}\binom{n+1}{k+1}(-1)^k = \frac{1}{n+1}\sum_k \binom{n+k}{n}\binom{n+1}{k+1}(-1)^k .$$

Пора делать ставки!

Остается только три k. Теперь можно продолжить с использованием (5.24): заменив (l, m, n, s) на $(n+1, 1, n, n)$, получаем

$$\frac{1}{n+1}\sum_k \binom{n+k}{n}\binom{n+1}{k+1}(-1)^k = \frac{1}{n+1}(-1)^n\binom{n-1}{-1} = 0 .$$

Как — нуль?! После стольких трудов?! Давайте проверим при $n=2$: $\binom{2}{0}\binom{0}{0}\frac{1}{1} - \binom{3}{2}\binom{2}{1}\frac{1}{2} + \binom{4}{4}\binom{4}{2}\frac{1}{3} = 1 - \frac{6}{2} + \frac{6}{3} = 0$. Таки нуль...

Давайте для интереса испробуем второй вариант, с обращением верхнего индекса $\binom{n+k}{k}$:

$$\frac{1}{n+1}\sum_k \binom{n+k}{k}\binom{n+1}{k+1}(-1)^k = \frac{1}{n+1}\sum_k \binom{-n-1}{k}\binom{n+1}{k+1} .$$

Теперь применим (5.23) с заменой $(l, m, n, s) \leftarrow (n+1, 1, 0, -n-1)$:

$$\frac{1}{n+1}\sum_k \binom{-n-1}{k}\binom{n+1}{k+1} = \frac{1}{n+1}\binom{0}{n} .$$

Стоп-стоп-стоп! Да, это нуль при $n > 0$, но при $n = 0$ это 1! А предыдущее решение давало нам нуль для всех случаев! Что же произошло? Ведь сумма на самом деле равна 1 при $n = 0$, так что верный ответ — '$[n=0]$'. Мы где-то ошиблись в нашем первом выводе.

Попробуйте применить бинарный поиск: вначале повторите среднюю формулу, чтобы понять, произошла ошибка раньше или позже.

Вернемся и повторим наш вывод при $n = 0$, чтобы увидеть, где же возникло такое расхождение. Ну вот — оказывается, мы попались в ловушку, о которой говорилось ранее: мы попытались применить симметрию там, где верхний индекс может оказаться отрицательным! Мы не имели права заменять $\binom{n+k}{k}$ на $\binom{n+k}{n}$, когда диапазон k представляет собой все целые числа, поскольку это превращает нулевую величину в ненулевую при $k < -n$. (Простите за это.)

Другой множитель суммы, $\binom{n+1}{k+1}$, превращается в нуль при $k < -n$, за исключением $n = 0$ и $k = -1$. Следовательно, наша ошибка при $n = 2$ остается невыявленной. В упр. 6 поясняется, как следует поступить.

Задача 7: новое затруднение

Примемся за еще более трудную задачу — найти в аналитическом виде сумму

$$\sum_{k\geqslant 0}\binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1}, \qquad \text{целые } m,n>0.$$

При $m=0$ получается сумма из предыдущей задачи, но в этой задаче m может быть любым, и мы не в состоянии применить здесь что-либо из использовавшегося при решении задачи 6. (В частности, решающий первый шаг.)

Однако если бы удалось избавиться от m, то мы бы могли воспользоваться только что полученным результатом. Итак, наша стратегия состоит в следующем: заменить $\binom{n+k}{m+2k}$ суммой членов вида $\binom{l+k}{2k}$ для некоторого неотрицательного целого l; тогда общий член суммы будет выглядеть аналогично общему члену суммы из задачи 6, и мы сможем изменить порядок суммирования.

Но чем заменить $\binom{n+k}{m+2k}$? Тщательное изучение ранее выведенных соотношений выявляет только одну подходящую кандидатуру, а именно — выведенное ранее соотношение (5.26) из табл. 218. Единственный способ его использования — заменить параметры (l,m,n,q,k) соответственно на $(n+k-1, 2k, m-1, 0, j)$:

$$\sum_{k\geqslant 0}\binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1} =$$

$$= \sum_{k\geqslant 0}\ \sum_{0\leqslant j\leqslant n+k-1}\binom{n+k-1-j}{2k}\binom{j}{m-1}\binom{2k}{k}\frac{(-1)^k}{k+1} =$$

$$= \sum_{j\geqslant 0}\binom{j}{m-1}\sum_{\substack{k\geqslant j-n+1\\ k\geqslant 0}}\binom{n+k-1-j}{2k}\binom{2k}{k}\frac{(-1)^k}{k+1}.$$

На последнем шаге мы изменяем порядок суммирования, работая с граничными условиями под знаками $\sum$ в соответствии с правилами главы 2.

Нельзя просто так заменить полученную внутреннюю сумму результатом решения задачи 6, поскольку в ней имеется дополнительное условие $k\geqslant j-n+1$. Но это дополнительное условие не является излишним, только если $j-n+1>0$, т.е. если $j\geqslant n$. А если $j\geqslant n$, то первый биномиальный коэффициент внутренней суммы равен нулю, поскольку верхний индекс находится между 0 и $k-1$ и, таким образом, строго меньше нижнего индекса $2k$.

Следовательно, на внешнюю сумму может быть наложено дополнительное ограничение $j < n$, которое не оказывает влияния на то, какие ненулевые члены в нее включены. Это делает ограничение $k \geqslant j - n + 1$ излишним, и можно воспользоваться результатом задачи 6. Двойная сумма распадается на две:

$$\sum_{j \geqslant 0} \binom{j}{m-1} \sum_{\substack{k \geqslant j-n+1 \\ k \geqslant 0}} \binom{n+k-1-j}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1} =$$

$$= \sum_{0 \leqslant j < n} \binom{j}{m-1} \sum_{k \geqslant 0} \binom{n+k-1-j}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1} =$$

$$= \sum_{0 \leqslant j < n} \binom{j}{m-1} [n-1-j=0] = \binom{n-1}{m-1}.$$

Внутренняя сумма обращается в нуль при всех j, кроме $j = n-1$, так что мы получаем в качестве ответа простое аналитическое выражение.

Задача 8: еще одно затруднение

Давайте расширим задачу по-другому, рассмотрев сумму

$$S_m = \sum_{k \geqslant 0} \binom{n+k}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1+m}, \qquad \text{целые } m, n \geqslant 0.$$

Если $m = 0$, мы снова получим сумму, с которой имели дело ранее, но теперь m находится в другом месте. Эта задача немного труднее задачи 7, но (к счастью) у нас теперь больший опыт в поиске решений, чем раньше. Начнем так же, как и в задаче 6:

$$S_m = \sum_{k \geqslant 0} \binom{n+k}{k} \binom{n}{k} \frac{(-1)^k}{k+1+m}.$$

Теперь (как и в задаче 7) попытаемся разложить ту часть, которая зависит от m, на члены, с которыми мы уже умеем обращаться. Когда m было равно нулю, мы вносили $k+1$ в $\binom{n}{k}$; при $m > 0$ можно сделать то же самое, если разложить $1/(k+1+m)$ на допускающие такое внесение члены. Нам все еще везет: в задаче 1 нами было доказано подходящее тождество

$$\sum_{j=0}^{m} \binom{m}{j} \binom{r}{j}^{-1} = \frac{r+1}{r+1-m}, \qquad \begin{array}{c} \text{целое } m \geqslant 0, \\ r \notin \{0, 1, \ldots, m-1\}. \end{array} \tag{5.33}$$

Замена r на $-k-2$ дает нам желаемое разложение

$$S_m = \sum_{k \geqslant 0} \binom{n+k}{k} \binom{n}{k} \frac{(-1)^k}{k+1} \sum_{j \geqslant 0} \binom{m}{j} \binom{-k-2}{j}^{-1}.$$

Теперь $(k+1)^{-1}$ можно внести в $\binom{n}{k}$, как и планировалось. В действительности эту величину можно внести и в $\binom{-k-2}{j}^{-1}$. Наличие двух вариантов упрощения наводит на мысль о существовании за сценой еще бо́льших возможностей по сокращению. И это так — представление всего, что входит в новый общий член, с использованием факториалов и возвращение к биномиальным коэффициентам дает нам формулу, которую можно просуммировать по k:

$$\begin{aligned} S_m &= \frac{m!\,n!}{(m+n+1)!}\sum_{j\geqslant 0}(-1)^j\binom{m+n+1}{n+1+j}\sum_k\binom{n+1+j}{k+j+1}\binom{-n-1}{k} = \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{j\geqslant 0}(-1)^j\binom{m+n+1}{n+1+j}\binom{j}{n}. \end{aligned}$$

Они думают, что это можно проверить на клочке бумаги на коленке?

Сумма по всем j равна нулю согласно (5.24). Следовательно, $-S_m$ представляет собой сумму по $j < 0$.

Для вычисления $-S_m$ при $j < 0$ заменим j на $-k-1$ и просуммируем по $k \geqslant 0$:

$$\begin{aligned} S_m &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\geqslant 0}(-1)^k\binom{m+n+1}{n-k}\binom{-k-1}{n} = \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\leqslant n}(-1)^{n-k}\binom{m+n+1}{k}\binom{k-n-1}{n} = \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\leqslant n}(-1)^k\binom{m+n+1}{k}\binom{2n-k}{n} = \\ &= \frac{m!\,n!}{(m+n+1)!}\sum_{k\leqslant 2n}(-1)^k\binom{m+n+1}{k}\binom{2n-k}{n}. \end{aligned}$$

Наконец, применяя (5.25), мы получим окончательный ответ:

$$S_m = (-1)^n\frac{m!\,n!}{(m+n+1)!}\binom{m}{n} = (-1)^n m^{\underline{n}} m^{\underline{-n-1}}.$$

Такое просто необходимо проверить. При $n = 2$ находим, что

$$S_m = \frac{1}{m+1} - \frac{6}{m+2} + \frac{6}{m+3} = \frac{m(m-1)}{(m+1)(m+2)(m+3)}.$$

Наш вывод требует, чтобы m было целым числом, но результат справедлив для всех действительных m, поскольку $(m+1)^{\overline{n+1}}\,S_m$ представляет собой полином от m степени $\leqslant n$.

5.3 Специальные приемы

Давайте теперь рассмотрим три приема, которые значительно усиливают ранее изученные нами методы работы с биномиальными коэффициентами.

Этот прием следовало бы пронумеровать как "прием $\frac{1}{2}$".

Прием 1: уполовинивание

Многие из наших соотношений включают произвольное действительное число r. Когда r имеет специальный вид "целое минус $\frac{1}{2}$", биномиальный коэффициент $\binom{r}{k}$ можно записать в виде совершенно непохожего произведения биномиальных коэффициентов. Это приводит к новому семейству тождеств, работать с которыми можно с удивительной легкостью.

Один из способов увидеть, как это работает, — начать с *формулы удвоения*

$$r^{\underline{k}}(r-\tfrac{1}{2})^{\underline{k}} = (2r)^{\underline{2k}}/2^{2k}, \qquad \text{целое } k \geqslant 0. \tag{5.34}$$

Это тождество становится очевидным, если разложить убывающие степени и расположить множители в левой части в чередующемся порядке:

$$r(r-\tfrac{1}{2})(r-1)(r-\tfrac{3}{2})\ldots(r-k+1)(r-k+\tfrac{1}{2}) = \\ = \frac{(2r)(2r-1)\ldots(2r-2k+1)}{2\cdot 2\cdot \ldots \cdot 2}.$$

Теперь можно разделить обе части на $k!^2$ и получить

$$\binom{r}{k}\binom{r-1/2}{k} = \binom{2r}{2k}\binom{2k}{k}\Big/2^{2k}, \qquad \text{целое } k. \tag{5.35}$$

Если положить $k=r=n$, где n — целое число, это даст

$$\binom{n-1/2}{n} = \binom{2n}{n}\Big/2^{2n}, \qquad \text{целое } n. \tag{5.36}$$

Обращение верхнего индекса приводит к еще одной полезной формуле:

$$\binom{-1/2}{n} = \left(\frac{-1}{4}\right)^n \binom{2n}{n}, \qquad \text{целое } n. \tag{5.37}$$

... половиним ...

Например, при $n = 4$ получаем

$$\binom{-1/2}{4} = \frac{(-1/2)(-3/2)(-5/2)(-7/2)}{4!} = \\ = \left(\frac{-1}{2}\right)^4 \frac{1\cdot3\cdot5\cdot7}{1\cdot2\cdot3\cdot4} = \\ = \left(\frac{-1}{4}\right)^4 \frac{1\cdot3\cdot5\cdot7\cdot2\cdot4\cdot6\cdot8}{1\cdot2\cdot3\cdot4\cdot1\cdot2\cdot3\cdot4} = \left(\frac{-1}{4}\right)^4\binom{8}{4}.$$

Обратите внимание на замену произведения нечетных чисел факториалом.

Тождество (5.35) имеет забавное следствие. Положим $r = \frac{1}{2}n$ и возьмем сумму по всем целым k. Результат согласно (5.23) будет следующим:

$$\sum_k \binom{n}{2k}\binom{2k}{k}2^{-2k} = \sum_k \binom{n/2}{k}\binom{(n-1)/2}{k} =$$
$$= \binom{n-1/2}{\lfloor n/2 \rfloor}, \qquad \text{целое } n \geqslant 0, \qquad (5.38)$$

поскольку либо $n/2$, либо $(n-1)/2$ равно $\lfloor n/2 \rfloor$ — целому неотрицательному числу!

Можно воспользоваться сверткой Вандермонда (5.27) и вывести

$$\sum_k \binom{-1/2}{k}\binom{-1/2}{n-k} = \binom{-1}{n} = (-1)^n, \qquad \text{целое } n \geqslant 0.$$

Подстановка значений из (5.37) дает

$$\binom{-1/2}{k}\binom{-1/2}{n-k} = \left(\frac{-1}{4}\right)^k\binom{2k}{k}\left(\frac{-1}{4}\right)^{n-k}\binom{2(n-k)}{n-k} =$$
$$= \frac{(-1)^n}{4^n}\binom{2k}{k}\binom{2n-2k}{n-k},$$

что в сумме дает $(-1)^n$. Следовательно, мы получили замечательное свойство "средних" элементов треугольника Паскаля:

$$\sum_k \binom{2k}{k}\binom{2n-2k}{n-k} = 4^n, \qquad \text{целое } n \geqslant 0. \qquad (5.39)$$

Например, $\binom{0}{0}\binom{6}{3}+\binom{2}{1}\binom{4}{2}+\binom{4}{2}\binom{2}{1}+\binom{6}{3}\binom{0}{0} = 1\cdot20+2\cdot6+6\cdot2+20\cdot1 = 64 = 4^3$.

Эти иллюстрации к нашему первому приему показывают, что разумно попытаться заменить биномиальные коэффициенты вида $\binom{2k}{k}$ биномиальными коэффициентами вида $\binom{n-1/2}{k}$, где n — некоторое подходящее целое число (обычно 0, 1 или k); получающаяся при этом формула может оказаться существенно более простой.

Прием 2: разности высших порядков

Ранее мы видели, что частичные суммы ряда $\binom{n}{k}(-1)^k$ вычислимы, в то время как для ряда с общим членом $\binom{n}{k}$ — нет.

Оказывается, что имеется множество важных приложений биномиальных коэффициентов с чередующимися знаками, $\binom{n}{k}(-1)^k$. Одной из причин этого является то, что такие коэффициенты тесно связаны с разностным оператором Δ, определенным в разделе 2.6.

Разность Δf функции f в точке x есть

$$\Delta f(x) \;=\; f(x+1)-f(x)\,;$$

если применить оператор Δ еще раз, получим вторую разность,

$$\begin{aligned}\Delta^2 f(x) \;&=\; \Delta f(x+1)-\Delta f(x) \;=\\ &=\; \bigl(f(x{+}2)-f(x{+}1)\bigr)-\bigl(f(x{+}1)-f(x)\bigr) \;=\\ &=\; f(x+2)-2f(x+1)+f(x)\,,\end{aligned}$$

подобную второй производной. Аналогично получаются

$$\begin{aligned}\Delta^3 f(x) \;&=\; f(x+3)-3f(x+2)+3f(x+1)-f(x)\,,\\ \Delta^4 f(x) \;&=\; f(x+4)-4f(x+3)+6f(x+2)-4f(x+1)+f(x)\,,\end{aligned}$$

и т.д. Биномиальные коэффициенты в эти формулы входят с чередующимися знаками.

В общем случае n-я разность представляет собой

$$\Delta^n f(x) \;=\; \sum_k \binom{n}{k}(-1)^{n-k} f(x+k)\,, \qquad \text{целое } n \geqslant 0. \tag{5.40}$$

Эта формула легко доказывается по индукции, но есть красивое доказательство, опирающееся на элементарную теорию операторов. Вспомним, что в разделе 2.6 был определен оператор сдвига E:

$$Ef(x) \;=\; f(x+1)\,;$$

следовательно, оператор Δ представляет собой $E-1$, где 1 — тождественный оператор, определяемый правилом $1f(x)=f(x)$. По биномиальной теореме

$$\Delta^n \;=\; (E-1)^n \;=\; \sum_k \binom{n}{k} E^k (-1)^{n-k}\,.$$

Это равенство, элементами которого выступают операторы; оно эквивалентно (5.40), поскольку E^k — это оператор, который переводит $f(x)$ в $f(x+k)$.

Интересный и важный случай возникает при рассмотрении отрицательных убывающих степеней. Пусть $f(x)=(x-1)^{\underline{-1}} =$

$1/x$. Тогда по правилу (2.45) имеем $\Delta f(x) = (-1)(x-1)^{\underline{-2}}$, $\Delta^2 f(x) = (-1)(-2)(x-1)^{\underline{-3}}$, и в общем случае

$$\Delta^n\bigl((x-1)^{\underline{-1}}\bigr) = (-1)^{\underline{n}}(x-1)^{\underline{-n-1}} = (-1)^n \frac{n!}{x(x+1)\ldots(x+n)}\,.$$

Равенство (5.40) дает нам

$$\sum_k \binom{n}{k}\frac{(-1)^k}{x+k} = \frac{n!}{x(x+1)\ldots(x+n)} = \\ = x^{-1}\binom{x+n}{n}^{-1}, \quad x \notin \{0,-1,\ldots,-n\}. \quad (5.41)$$

Например,

$$\frac{1}{x} - \frac{4}{x+1} + \frac{6}{x+2} - \frac{4}{x+3} + \frac{1}{x+4} = \\ = \frac{4!}{x(x+1)(x+2)(x+3)(x+4)} = 1\Big/x\binom{x+4}{4}\,.$$

Сумма в (5.41) представляет собой разложение $n!/\bigl(x(x+1)\ldots(x+n)\bigr)$ на простейшие дроби.

Важные результаты могут быть получены и из положительных убывающих степеней. Если $f(x)$ представляет собой полином степени d, разность $\Delta f(x)$ является полиномом степени $d-1$; таким образом, величина $\Delta^d f(x)$ представляет собой константу, а $\Delta^n f(x) = 0$ при $n > d$. Этот исключительно важный факт упрощает многие формулы.

Более пристальный взгляд дает дальнейшую информацию. Пусть $f(x)$ представляет собой некоторый полином степени d:

$$f(x) = a_d x^d + a_{d-1}x^{d-1} + \cdots + a_1 x^1 + a_0 x^0.$$

В главе 6 мы увидим, что обычные степени можно выразить в виде суммы убывающих степеней (например, $x^2 = x^{\underline{2}} + x^{\underline{1}}$); следовательно, существуют коэффициенты $b_d, b_{d-1}, \ldots, b_1, b_0$, такие, что

$$f(x) = b_d x^{\underline{d}} + b_{d-1}x^{\underline{d-1}} + \cdots + b_1 x^{\underline{1}} + b_0 x^{\underline{0}}.$$

(Оказывается, что $b_d = a_d$ и $b_0 = a_0$, но коэффициенты между ними связаны более сложным образом.) Пусть $c_k = k!\,b_k$ при $0 \leqslant k \leqslant d$. Тогда

$$f(x) = c_d\binom{x}{d} + c_{d-1}\binom{x}{d-1} + \cdots + c_1\binom{x}{1} + c_0\binom{x}{0};$$

таким образом, любой полином может быть представлен в виде суммы величин, кратных биномиальным коэффициентам. Такое разложение называется *рядом Ньютона* для $f(x)$, поскольку оно широко использовалось Исааком Ньютоном (Isaac Newton).

Ранее в этой главе мы заметили, что формула сложения приводит к формуле

$$\Delta\left(\binom{x}{k}\right) = \binom{x}{k-1}.$$

Таким образом, по индукции n-я разность ряда Ньютона определяется очень просто:

$$\begin{aligned}\Delta^n f(x) &= c_d\binom{x}{d-n} + c_{d-1}\binom{x}{d-1-n} + \cdots + \\ &\quad + c_1\binom{x}{1-n} + c_0\binom{x}{-n}.\end{aligned}$$

Если положить $x = 0$, то все члены $c_k\binom{x}{k-n}$ в правой части будут равны нулю, за исключением члена с $k - n = 0$; следовательно,

$$\Delta^n f(0) = \begin{cases} c_n, & \text{если } n \leqslant d; \\ 0, & \text{если } n > d. \end{cases}$$

Ряд Ньютона для $f(x)$, таким образом, представляет собой

$$\begin{aligned}f(x) &= \Delta^d f(0)\binom{x}{d} + \Delta^{d-1} f(0)\binom{x}{d-1} + \cdots + \\ &\quad + \Delta f(0)\binom{x}{1} + f(0)\binom{x}{0}.\end{aligned}$$

Предположим, например, что $f(x) = x^3$. Легко вычислить, что

$$\begin{gathered} f(0) = 0, \qquad f(1) = 1, \qquad f(2) = 8, \qquad f(3) = 27; \\ \Delta f(0) = 1, \qquad \Delta f(1) = 7, \qquad \Delta f(2) = 19; \\ \Delta^2 f(0) = 6, \qquad \Delta^2 f(1) = 12; \\ \Delta^3 f(0) = 6. \end{gathered}$$

Так что ряд Ньютона имеет вид $x^3 = 6\binom{x}{3} + 6\binom{x}{2} + 1\binom{x}{1} + 0\binom{x}{0}$.

Формулу $\Delta^n f(0) = c_n$ можно переписать следующим образом с применением (5.40) при $x = 0$:

$$\sum_k \binom{n}{k}(-1)^k\left(c_0\binom{k}{0} + c_1\binom{k}{1} + c_2\binom{k}{2} + \cdots\right) = (-1)^n c_n, \qquad \text{целое } n \geqslant 0.$$

Здесь $\langle c_0, c_1, c_2, \ldots \rangle$ — произвольная последовательность коэффициентов; бесконечная сумма $c_0\binom{k}{0} + c_1\binom{k}{1} + c_2\binom{k}{2} + \cdots$ в действительности оказывается конечной при любом $k \geqslant 0$, так что вопрос о сходимости не стойт. В частности, можно доказать важное тождество

$$\sum_k \binom{n}{k}(-1)^k(a_0+a_1k+\cdots+a_nk^n) \;=\; (-1)^n n!\, a_n\,, \qquad \text{целое } n \geqslant 0, \tag{5.42}$$

поскольку полином $a_0 + a_1k + \cdots + a_nk^n$ всегда можно записать в виде ряда Ньютона $c_0\binom{k}{0} + c_1\binom{k}{1} + \cdots + c_n\binom{k}{n}$ с $c_n = n!\,a_n$.

Многие суммы, которые, на первый взгляд, представляются безнадежными, в действительности могут быть просуммированы почти тривиально при помощи идеи n-х разностей. Например, рассмотрим соотношение

$$\sum_k \binom{n}{k}\binom{r-sk}{n}(-1)^k \;=\; s^n\,, \qquad \text{целое } n \geqslant 0. \tag{5.43}$$

Оно выглядит очень впечатляюще, потому что совершенно отличается от всего, с чем мы до сих пор имели дело. Но в действительности его легко понять; стóит только обратить внимание на проясняющий дело множитель $\binom{n}{k}(-1)^k$ в общем члене, так как функция

$$f(k) \;=\; \binom{r-sk}{n} \;=\; \frac{1}{n!}(-1)^n s^n k^n+\cdots \;=\; (-1)^n s^n\binom{k}{n}+\cdots$$

представляет собой полином от k степени n, старший коэффициент которого равен $(-1)^n s^n/n!$. Таким образом, (5.43) — это не более чем приложение (5.42).

Ряд Ньютона рассматривался в предположении, что $f(x)$ — полином. Но мы видели, что бесконечный ряд Ньютона

$$f(x) \;=\; c_0\binom{x}{0} + c_1\binom{x}{1} + c_2\binom{x}{2} + \cdots$$

также имеет смысл, поскольку такие суммы всегда конечны, если x — неотрицательное целое число. Наш вывод формулы $\Delta^n f(0) = c_n$ работает в бесконечном случае точно так же, как и в случае полинома; так что общее соотношение таково:

$$f(x) \;=\; f(0)\binom{x}{0}+\Delta f(0)\binom{x}{1}+\Delta^2 f(0)\binom{x}{2}+\Delta^3 f(0)\binom{x}{3}+\cdots\,, \qquad \text{целое } x \geqslant 0. \tag{5.44}$$

Эта формула справедлива для любой функции $f(x)$, которая определена при целых неотрицательных x. Более того, если правая часть этого соотношения сходится при других значениях x, то она определяет некоторую функцию, которая “интерполирует” $f(x)$ в некотором естественном смысле. (Существует бесконечное количество способов интерполяции значений функции, так что нельзя утверждать, что соотношение (5.44) истинно для всех x, которые делают бесконечный ряд сходящимся. Например, если положить $f(x) = \sin(\pi x)$, то получим $f(x) = 0$ во всех целых точках, так что правая часть (5.44) тождественно равна нулю; но левая часть отлична от нуля при всех нецелых x.)

Ряд Ньютона представляет собой конечно-разностный аналог ряда Тейлора из исчисления бесконечно малых. Так же, как ряд Тейлора можно записать в виде

$$g(a+x) = \frac{g(a)}{0!}x^0 + \frac{g'(a)}{1!}x^1 + \frac{g''(a)}{2!}x^2 + \frac{g'''(a)}{3!}x^3 + \cdots,$$

Так как $E = 1+\Delta$, $E^x = \sum_k \binom{x}{k}\Delta^k$; *и* $E^x g(a) = g(a+x)$.

ряд Ньютона для $f(x) = g(a+x)$ может быть записан как

$$g(a+x) = \frac{g(a)}{0!}x^{\underline{0}} + \frac{\Delta g(a)}{1!}x^{\underline{1}} + \frac{\Delta^2 g(a)}{2!}x^{\underline{2}} + \frac{\Delta^3 g(a)}{3!}x^{\underline{3}} + \cdots. \tag{5.45}$$

(Это то же, что и (5.44), так как $\Delta^n f(0) = \Delta^n g(a)$ для всех $n \geqslant 0$, если $f(x) = g(a+x)$.) И ряд Тейлора, и ряд Ньютона конечны, если g — полином или если $x = 0$; кроме того, ряд Ньютона конечен, если x — положительное целое число. В противном случае суммы могут как сходиться, так и расходиться при определенных значениях x. Если ряд Ньютона сходится, когда x не является целым неотрицательным числом, то он может в действительности сходиться к величине, *отличной* от $g(a+x)$, потому что ряд Ньютона (5.45) зависит только от дискретных значений функции $g(a)$, $g(a+1)$, $g(a+2)$, ...

Один из примеров сходящегося ряда Ньютона предоставляет биномиальная теорема. Пусть $g(x) = (1+z)^x$, где z — фиксированное комплексное число, такое, что $|z| < 1$. Тогда $\Delta g(x) = (1+z)^{x+1} - (1+z)^x = z(1+z)^x$, следовательно, $\Delta^n g(x) = z^n(1+z)^x$. В этом случае бесконечный ряд Ньютона

$$g(a+x) = \sum_n \Delta^n g(a)\binom{x}{n} = (1+z)^a \sum_n \binom{x}{n} z^n$$

сходится к “корректному” значению $(1+z)^{a+x}$ при всех x.

Джеймс Стирлинг (James Stirling) пытался использовать ряд Ньютона для обобщения факториальной функции на нецелые числа. Сначала он нашел коэффициенты S_n, такие, что

$$x! = \sum_n S_n\binom{x}{n} = S_0\binom{x}{0} + S_1\binom{x}{1} + S_2\binom{x}{2} + \cdots \qquad (5.46)$$

представляет собой тождество при $x = 0$, $x = 1$, $x = 2$ и т.д. Однако он обнаружил, что получающийся в результате ряд не сходится при целых неотрицательных x. Он предпринял новую попытку, записав на этот раз

$$\ln x! = \sum_n s_n\binom{x}{n} = s_0\binom{x}{0} + s_1\binom{x}{1} + s_2\binom{x}{2} + \cdots. \qquad (5.47)$$

Так как $\Delta(\ln x!) = \ln(x+1)! - \ln x! = \ln(x+1)$, в соответствии с (5.40) получается

$$\begin{aligned} s_n &= \Delta^n(\ln x!)\big|_{x=0} = \Delta^{n-1}\bigl(\ln(x+1)\bigr)\big|_{x=0} = \\ &= \sum_k \binom{n-1}{k}(-1)^{n-1-k}\ln(k+1). \end{aligned}$$

Искомыми коэффициентами являются $s_0 = s_1 = 0$; $s_2 = \ln 2$; $s_3 = \ln 3 - 2\ln 2 = \ln\frac{3}{4}$; $s_4 = \ln 4 - 3\ln 3 + 3\ln 2 = \ln\frac{32}{27}$ и т.д. Так Стирлинг получил сходящийся ряд (хотя и не доказал этого). В действительности найденный им ряд сходится при всех $x > -1$. Таким образом, он оказался в состоянии вычислить $\frac{1}{2}!$. Остальная часть истории — в упр. 88.

"Постольку, поскольку эти члены возрастают очень быстро, их разности будут образовывать расходящуюся прогрессию, что препятствует приближению ординаты параболы, поэтому в данном и подобных ему случаях я интерполирую логарифмы этих членов, разности которых составляют быстро сходящийся ряд."
— Д. Стирлинг (J. Stirling) [343]

(До XIX века доказывать сходимость было необязательно.)

Прием 3: обращение

Частный случай только что выведенного правила (5.45) для ряда Ньютона можно переписать следующим образом:

$$g(n) = \sum_k \binom{n}{k}(-1)^k f(k) \iff f(n) = \sum_k \binom{n}{k}(-1)^k g(k). \qquad (5.48)$$

Это дуальное соотношение между f и g называется *формулой обращения*, которая напоминает формулы обращения Мёбиуса (4.56) и (4.61), с которыми мы встречались в главе 4. Формулы обращения говорят нам, как разрешить "неявные рекуррентные соотношения", в которых неизвестная последовательность встроена в сумму.

Например, пусть $g(n)$ — известная функция, $f(n)$ — неизвестная и при этом можно показать, что $g(n) = \sum_k \binom{n}{k}(-1)^k f(k)$.

Обратите: 'zınb ppo'. (См. ответ на следующей странице.)

Тогда (5.48) позволяет нам выразить $f(n)$ в виде суммы известных значений.

Мы можем доказать (5.48) непосредственно, с применением основных методов из начала данной главы. Если $g(n) = \sum_k \binom{n}{k}(-1)^k f(k)$ при всех $n \geqslant 0$, то

$$\begin{aligned}
\sum_k \binom{n}{k}(-1)^k g(k) &= \sum_k \binom{n}{k}(-1)^k \sum_j \binom{k}{j}(-1)^j f(j) = \\
&= \sum_j f(j) \sum_k \binom{n}{k}(-1)^{k+j}\binom{k}{j} = \\
&= \sum_j f(j) \sum_k \binom{n}{j}(-1)^{k+j}\binom{n-j}{k-j} = \\
&= \sum_j f(j)\binom{n}{j} \sum_k (-1)^k \binom{n-j}{k} = \\
&= \sum_j f(j)\binom{n}{j}[n-j=0] \quad = \quad f(n)\,.
\end{aligned}$$

Доказательство в обратном направлении, разумеется, такое же, поскольку связь между f и g симметрична.

Давайте проиллюстрируем (5.48), применив ее к "задаче о победе футбольной команды". Группа из n фанатов выигрывающей футбольной команды на радостях подбрасывает свои шляпы в воздух. Шляпы падают в случайном порядке — по одной к каждому болельщику. Сколькими способами $h(n, k)$ ровно k болельщиков получат назад собственные шляпы?

Есть и другой вариант постановки задачи — когда n изрядно подвыпивших на вечеринке гостей берут шляпы с вешалки в случайном порядке. Выбирайте, что кому больше по душе.
— Переводчик

Например, если $n = 4$ и и болельщики обозначены буквами A, B, C, D, то $4! = 24$ возможных способа приземления шляп дают следующие количества попаданий в руки законных владельцев:

ABCD	4	BACD	2	CABD	1	DABC	0
ABDC	2	BADC	0	CADB	0	DACB	1
ACBD	2	BCAD	1	CBAD	2	DBAC	1
ACDB	1	BCDA	0	CBDA	1	DBCA	2
ADBC	1	BDAC	0	CDAB	0	DCAB	0
ADCB	2	BDCA	1	CDBA	0	DCBA	0

Таким образом, $h(4,4) = 1$, $h(4,3) = 0$, $h(4,2) = 6$, $h(4,1) = 8$ и $h(4,0) = 9$.

Можно определить $h(n, k)$, заметив, что это количество способов выбора k счастливых обладателей шляп, а именно $\binom{n}{k}$, умноженное на количество способов упорядочения оставшихся $n-k$

шляп так, что ни одна из них не попадет к владельцу, а именно на $h(n-k,0)$. Перестановка называется *беспорядком* (derangement), если в ней перемещен каждый предмет, а само число беспорядков из n объектов иногда обозначается как '$n¡$' (читается как "n субфакториал"). Таким образом, $h(n-k,0) = (n-k)¡$, а общая формула имеет вид

Повернув "zınb ppo" на 180°, мы получим "odd quiz", т.е. "странную шутку".
— Переводчик

$$h(n,k) \;=\; \binom{n}{k} h(n-k,0) \;=\; \binom{n}{k}(n-k)¡\,.$$

(Обозначение для субфакториала не является стандартным, и не ясно, насколько удачно наше; тем не менее попробуем пока что использовать его и посмотрим, насколько оно нам понравится. В конце концов, всегда можно воспользоваться чем-то наподобие 'D_n', если '$n¡$' не будет нас устраивать.)

Наша задача была бы решенной, если бы мы получили аналитическое выражение для $n¡$, так что попробуем его найти. Поскольку сумма $h(n,k)$ по всем k — это общее количество перестановок n шляп, можно легко записать рекуррентное соотношение

$$\begin{aligned} n! \;=\; \sum_k h(n,k) \;&=\; \sum_k \binom{n}{k}(n-k)¡ \;= \\ &=\; \sum_k \binom{n}{k} k¡\,, \qquad \text{целое } n \geqslant 0. \end{aligned} \tag{5.49}$$

(На последнем шаге мы заменяем k на $n-k$ и $\binom{n}{n-k}$ на $\binom{n}{k}$.) С помощью данного неявного рекуррентного соотношения можно вычислить любые требуемые значения $h(n,k)$:

n	$h(n,0)$	$h(n,1)$	$h(n,2)$	$h(n,3)$	$h(n,4)$	$h(n,5)$	$h(n,6)$
0	1						
1	0	1					
2	1	0	1				
3	2	3	0	1			
4	9	8	6	0	1		
5	44	45	20	10	0	1	
6	265	264	135	40	15	0	1

Например, вот как можно вычислить строку для $n = 4$. Два крайних справа элемента очевидны — имеется только один вариант того, что все шляпы попадут к своим хозяевам, и ни одного варианта того, что ровно три болельщика получат свои шляпы (чью шляпу тогда поймал бы четвертый болельщик?). При $k = 2$ и $k = 1$ можно воспользоваться нашим уравнением для $h(n,k)$ и получить $h(4,2) = \binom{4}{2}h(2,0) = 6\cdot 1 = 6$ и $h(4,1) = \binom{4}{1}h(3,0) = 4\cdot 2 = 8$. Это уравнение нельзя использовать для вычисления

Искусство математики, как и жизни, состоит в выяснении, какие из истин бесполезны.

(Вспоминается анекдот о том, как заблудившийся в тумане на воздушном шаре пилот опускается к земле и спрашивает прохожего:
— Где я?!
— На воздушном шаре! — слышит он в ответ и решает, что перед ним математик, ибо кто еще в состоянии дать такой совершенно точный и совершенно бесполезный ответ?
— Переводчик)

$h(4,0)$. Вернее, можно, но оно даст нам $h(4,0) = \binom{4}{0}h(4,0)$, что, конечно, верно, но бесполезно. Пойдем другим путем и используем для вычисления $h(4,0) = 9$ соотношение $h(4,0)+8+6+0+1 = 4!$. Найденное $h(4,0) = 9$ и есть значение $4¡$. Аналогичным образом $n¡$ зависит от других значений $k¡$ при $k < n$.

Но как разрешить рекуррентность вида (5.49)? Просто: она имеет вид (5.48), с $g(n) = n!$ и $f(k) = (-1)^k k¡$. Следовательно, ее решением является

$$n¡ = (-1)^n \sum_k \binom{n}{k}(-1)^k k!\,.$$

Ну, на самом деле это не решение, а сумма, которую по возможности следует записать в аналитическом виде. Но это уже лучше, чем рекуррентное соотношение. Сумму можно упростить, поскольку $k!$ сокращается со скрытым $k!$ в $\binom{n}{k}$, так что попробуем это сделать и получить

$$n¡ = \sum_{0\leqslant k\leqslant n} \frac{n!}{(n-k)!}(-1)^{n+k} = n! \sum_{0\leqslant k\leqslant n} \frac{(-1)^k}{k!}\,. \qquad (5.50)$$

Оставшаяся часть суммы быстро сходится к числу $\sum_{k\geqslant 0}(-1)^k/k! = e^{-1}$. В действительности члены, исключенные из суммы, дают

$$n! \sum_{k>n} \frac{(-1)^k}{k!} = \frac{(-1)^{n+1}}{n+1} \sum_{k\geqslant 0}(-1)^k \frac{(n+1)!}{(k+n+1)!} =$$
$$= \frac{(-1)^{n+1}}{n+1}\left(1-\frac{1}{n+2}+\frac{1}{(n+2)(n+3)}-\cdots\right),$$

где заключенная в скобки величи на лежит между 1 и $1 - \frac{1}{n+2} = \frac{n+1}{n+2}$. Следовательно, абсолютная величина разности $n¡$ и $n!/e$ грубо равна $1/n$; точнее, она лежит между $1/(n+1)$ и $1/(n+2)$. Но $n¡$ — целое число, поэтому оно должно быть равно числу, которое мы получаем при округлении $n!/e$ до ближайшего целого при $n > 0$. Итак, вот решение в аналитическом виде, которое мы искали:

$$n¡ = \left\lfloor \frac{n!}{e} + \frac{1}{2} \right\rfloor + [n=0]\,. \qquad (5.51)$$

Это число вариантов того, что ни один из болельщиков не получит назад свою шляпу. При большом n более важно знать *вероятность*, с которой это происходит. Если предположить, что каждый из $n!$ исходов равновероятен — так как шляпы подброшены очень высоко,— то данная вероятность равна

$$\frac{n¡}{n!} = \frac{n!/e + O(1)}{n!} \sim \frac{1}{e} = .367\ldots.$$

Так что, когда n становится достаточно велико, вероятность, что все шляпы будут перепутаны, составляет почти 37%.

Бейсбольные болельщики: .367 — это еще и коэффициент бэттера Тая Кобба (Ty Cobb) за всю его карьеру — никем не превзойденный рекорд. Совпадение?

Кстати, рекуррентное соотношение (5.49) для субфакториалов точно такое же, как и (5.46), первое рекуррентное соотношение, рассмотренное Стирлингом, когда он пытался обобщить факториальную функцию. Следовательно, $S_k = k¡$. Эти коэффициенты настолько велики, что неудивительно, что бесконечный ряд (5.46) расходится при нецелых x.

(Ну, коэффициент Кобба был 4191/11429 ≈ .366699, в то время как 1/e ≈ .367879. Но если бы Вэйд Боггс (Wade Boggs) провел несколько удачных сезонов, то...)

Перед тем как покончить с данной задачей, рассмотрим вкратце две интересные закономерности, которые так и выпирают из таблицы начальных значений $h(n,k)$. Во-первых, числа 1, 3, 6, 10, 15, ..., находящиеся под диагональю из нулей, представляют собой треугольные числа. Это наблюдение легко доказать, поскольку данные элементы таблицы являются числами $h(n, n-2)$, а

$$h(n, n-2) = \binom{n}{n-2} 2¡ = \binom{n}{2}.$$

Кроме того, похоже, что числа в первых двух столбцах отличаются друг от друга на ± 1. Всегда ли это так? Да, поскольку

$$\begin{aligned} h(n,0)-h(n,1) &= n¡-n(n-1)¡ = \\ &= \Biggl(n! \sum_{0 \leqslant k \leqslant n} \frac{(-1)^k}{k!}\Biggr) - \Biggl(n(n-1)! \sum_{0 \leqslant k \leqslant n-1} \frac{(-1)^k}{k!}\Biggr) = \\ &= n! \frac{(-1)^n}{n!} = (-1)^n. \end{aligned}$$

Другими словами, $n¡ = n(n-1)¡ + (-1)^n$. Это гораздо более простое рекуррентное соотношение для числа беспорядков, чем то, которое мы рассматривали раньше.

Теперь попробуем обратить что-нибудь еще. Если применить инверсию к формуле

Но инверсия — источник смога.

$$\sum_k \binom{n}{k} \frac{(-1)^k}{x+k} = \frac{1}{x} \binom{x+n}{n}^{-1},$$

выведенной в (5.41), то мы найдем, что

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

Это интересно, но в действительности не ново: если обратить верхний индекс в $\binom{x+k}{k}$, то мы просто вновь получим уже открытое соотношение (5.33).

5.4 Производящие функции

Вот мы и добрались до самой важной идеи всей книги — понятия *производящей функции* (generating function). Представляющую интерес бесконечную последовательность $\langle a_0, a_1, a_2, \ldots \rangle$ удобно представить в виде *степенного ряда* относительно вспомогательной переменной z,

$$A(z) = a_0 + a_1 z + a_2 z^2 + \cdots = \sum_{k\geqslant 0} a_k z^k. \tag{5.52}$$

Использование для обозначения переменной буквы z связано с тем, что зачастую под z подразумевается комплексное число. Теория функций комплексного переменного традиционно использует в своих формулах 'z', а степенные ряды (известные также как аналитические или голоморфные функции) занимают в этой теории одно из центральных мест.

В следующих главах мы будем много раз сталкиваться с производящими функциями; так, глава 7 будет полностью посвящена им. В настоящий момент наша цель состоит в том, чтобы ввести основные понятия и продемонстрировать уместность использования производящих функций при изучении биномиальных коэффициентов.

Основная польза производящей функции в том, что одной ею можно представить целую бесконечную последовательность. Зачастую при решении тех или иных задач можно сначала обратиться к производящим функциям, а затем, поработав с ними как следует, вновь вернуться к их коэффициентам. При определенном везении можно узнать о функции достаточно, чтобы выяснить все, что надо знать о ее коэффициентах.

Если $A(z)$ — некоторый степенной ряд $\sum_{k\geqslant 0} a_k z^k$, то его удобно записать в виде

(См. в [223] рассмотрение истории и пользы данного понятия.)

$$[z^n]\,A(z) = a_n\,; \tag{5.53}$$

другими словами, $[z^n]\,A(z)$ означает коэффициент при z^n в $A(z)$.

Пусть $A(z)$ — производящая функция для последовательности $\langle a_0, a_1, a_2, \ldots \rangle$, как в (5.52), и пусть $B(z)$ — производящая функция для последовательности $\langle b_0, b_1, b_2, \ldots \rangle$. Тогда произведение $A(z)B(z)$ представляет собой степенной ряд

$$(a_0+a_1z+a_2z^2+\cdots)(b_0+b_1z+b_2z^2+\cdots) = \\ = a_0b_0+(a_0b_1+a_1b_0)z+(a_0b_2+a_1b_1+a_2b_0)z^2+\cdots;$$

коэффициент при z^n в этом произведении равен

$$a_0b_n + a_1b_{n-1} + \cdots + a_nb_0 = \sum_{k=0}^{n} a_kb_{n-k}.$$

Таким образом, если мы хотим вычислить сумму, которая имеет общий вид

$$c_n = \sum_{k=0}^{n} a_kb_{n-k}, \qquad (5.54)$$

и при этом известны производящие функции $A(z)$ и $B(z)$, то

$$c_n = [z^n]\, A(z)B(z).$$

Последовательность $\langle c_n \rangle$, определяемая правилом (5.54), называется *сверткой* последовательностей $\langle a_n \rangle$ и $\langle b_n \rangle$; эти две последовательности "свертываются", образуя суммы произведений всех тех элементов последовательностей, нижние индексы которых при сложении дают некоторую определенную величину. Суть предыдущего абзаца состоит в том, что свертка последовательностей соответствует умножению их производящих функций.

Производящие функции обеспечивают нас мощными средствами обнаружения и/или доказательства тождеств. Например, биномиальная теорема гласит, что $(1+z)^r$ — это производящая функция для последовательности $\langle \binom{r}{0}, \binom{r}{1}, \binom{r}{2}, \ldots \rangle$:

$$(1+z)^r = \sum_{k\geqslant 0} \binom{r}{k} z^k.$$

Аналогично

$$(1+z)^s = \sum_{k\geqslant 0} \binom{s}{k} z^k.$$

Если перемножить эти производящие функции, то получится другая производящая функция:

$$(1+z)^r(1+z)^s = (1+z)^{r+s}.$$

И вот самый важный момент: приравнивание коэффициентов при z^n в обеих частях этого равенства дает

$$\sum_{k=0}^{n} \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}.$$

(5.27)! = (5.27)(4.27) (3.27)(2.27) (1.27)(0.27)!.

Мы заново открыли свертку Вандермонда (5.27)!

Это было так просто и красиво, что очень хочется попробовать еще. В этот раз мы воспользуемся выражением $(1-z)^r$, которое представляет собой производящую функцию для последовательности $\langle(-1)^n\binom{r}{n}\rangle = \langle\binom{r}{0}, -\binom{r}{1}, \binom{r}{2}, \ldots\rangle$. Умножение на $(1+z)^r$ дает другую производящую функцию, коэффициенты которой нам известны:

$$(1-z)^r(1+z)^r = (1-z^2)^r.$$

Приравнивание коэффициентов при z^n приводит к равенству

$$\sum_{k=0}^{n} \binom{r}{k}\binom{r}{n-k}(-1)^k = (-1)^{n/2}\binom{r}{n/2}[n \text{ четное}]. \qquad (5.55)$$

Следует проверить это соотношение при одном-двух малых случаях. При $n=3$, например, получается

$$\binom{r}{0}\binom{r}{3} - \binom{r}{1}\binom{r}{2} + \binom{r}{2}\binom{r}{1} - \binom{r}{3}\binom{r}{0} = 0.$$

Каждый положительный член сокращается с таким же отрицательным. То же самое получается при любом нечетном n, так что в этих случаях сумма не представляет собой особого интереса. Но если n четное, например когда $n=2$, мы получаем нетривиальную сумму, отличную от свертки Вандермонда:

$$\binom{r}{0}\binom{r}{2} - \binom{r}{1}\binom{r}{1} + \binom{r}{2}\binom{r}{0} = 2\binom{r}{2} - r^2 = -r.$$

Таким образом, при $n=2$ равенство (5.55) подтверждено. Оказывается, что (5.30) представляет собой частный случай нашего нового тождества (5.55).

Биномиальные коэффициенты появляются и в некоторых других производящих функциях. Особенно примечательны следующие важные соотношения, в которых нижний индекс фиксирован, а верхний — изменяется:

Если у вас есть цветной маркер, можете подчеркнуть эти два равенства — они того заслуживают.

$$\frac{1}{(1-z)^{n+1}} = \sum_{k\geqslant 0}\binom{n+k}{n}z^k, \qquad \text{целое } n \geqslant 0 \qquad (5.56)$$

$$\frac{z^n}{(1-z)^{n+1}} = \sum_{k\geqslant 0}\binom{k}{n}z^k, \qquad \text{целое } n \geqslant 0. \qquad (5.57)$$

Здесь второе тождество — это всего лишь первое тождество, умноженное на z^n, т.е. "сдвинутое вправо" на n позиций. Первое тождество представляет собой всего лишь немного замаскированный частный случай биномиальной теоремы: если разложить $(1-z)^{-n-1}$ согласно (5.13), то коэффициент при z^k равен $\binom{-n-1}{k}(-1)^k$, что можно переписать как $\binom{k+n}{k}$ или $\binom{n+k}{n}$ путем обращения верхнего индекса. Эти частные случаи заслуживают явного упоминания, поскольку очень часто возникают в приложениях.

При $n = 0$ получаем частный случай частного случая — геометрический ряд:

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \cdots = \sum_{k\geqslant 0} z^k .$$

Это производящая функция последовательности $\langle 1, 1, 1, \ldots \rangle$, и она особенно полезна потому, что свертка любой другой последовательности с ней представляет собой последовательность сумм: если $b_k = 1$ при всех k, то (5.54) сводится к

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

Таким образом, если $A(z)$ — производящая функция для членов суммы $\langle a_0, a_1, a_2, \ldots \rangle$, то $A(z)/(1-z)$ — производящая функция для сумм $\langle c_0, c_1, c_2, \ldots \rangle$.

Задача о беспорядках, связанная со шляпами и болельщиками, которую мы решили путем обращения, может быть решена при помощи производящих функций еще одним интересным способом. Основное рекуррентное соотношение этой задачи

$$n! = \sum_{k} \binom{n}{k} (n-k)\text{¡}$$

может быть представлено в виде свертки, если выразить $\binom{n}{k}$ через факториалы и разделить обе части на $n!$:

$$1 = \sum_{k=0}^{n} \frac{1}{k!} \, \frac{(n-k)\text{¡}}{(n-k)!} .$$

Производящей функцией для последовательности $\langle \frac{1}{0!}, \frac{1}{1!}, \frac{1}{2!}, \ldots \rangle$ является e^z; следовательно, если положить

$$D(z) = \sum_{k\geqslant 0} \frac{k\text{¡}}{k!} z^k ,$$

то данная свертка/рекуррентность указывает на то, что

$$\frac{1}{1-z} = e^z D(z).$$

Решение относительно $D(z)$ дает

$$D(z) = \frac{1}{1-z}e^{-z} = \frac{1}{1-z}\left(\frac{1}{0!}z^0 - \frac{1}{1!}z^1 + \frac{1}{2!}z^2 + \cdots\right).$$

Приравнивание коэффициентов при z^n показывает, что

$$\frac{n¡}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!};$$

это та же формула, которая была получена ранее путем обращения.

До сих пор исследование производящих функций давало нам доказательства того, что мы знали и так, благодаря иным методам. Мы не использовали производящих функций для получения новых результатов, за исключением (5.55). Теперь мы готовы к поиску нового и неожиданного. Существуют два семейства степенных рядов, которые порождают особенно широкий класс тождеств с биномиальными коэффициентами. Определим *обобщенный биномиальный ряд* $\mathcal{B}_t(z)$ и *обобщенный экспоненциальный ряд* $\mathcal{E}_t(z)$ следующим образом:

$$\mathcal{B}_t(z) = \sum_{k\geqslant 0}(tk)^{\underline{k-1}}\frac{z^k}{k!}; \quad \mathcal{E}_t(z) = \sum_{k\geqslant 0}(tk+1)^{k-1}\frac{z^k}{k!}. \qquad (5.58)$$

В разделе 7.5 мы покажем, что данные функции удовлетворяют тождествам

$$\mathcal{B}_t(z)^{1-t} - \mathcal{B}_t(z)^{-t} = z; \ \mathcal{E}_t(z)^{-t}\ln\mathcal{E}_t(z) = z. \qquad (5.59)$$

В частном случае $t = 0$ имеем

$$\mathcal{B}_0(z) = 1 + z; \qquad \mathcal{E}_0(z) = e^z;$$

это объясняет, почему ряды с произвольным параметром t называются "обобщенными" биномиальными и экспоненциальными рядами.

Следующие пары соотношений справедливы при любом действительном r:

$$\mathcal{B}_t(z)^r = \sum_{k\geqslant 0}\binom{tk+r}{k}\frac{r}{tk+r}z^k;$$

$$\mathcal{E}_t(z)^r = \sum_{k\geqslant 0} r\frac{(tk+r)^{k-1}}{k!}z^k; \qquad (5.60)$$

Обобщенный биномиальный ряд $\mathcal{B}_t(z)$ открыт в 1750-х годах Дж. Х. Ламбертом (J. H. Lambert) [236, §38], который несколькими годами позже заметил [237], что его степени удовлетворяют первому тождеству в (5.60). В упр. 84 поясняется, как вывести (5.61) из (5.60).

$$\frac{\mathcal{B}_t(z)^r}{1-t+t\mathcal{B}_t(z)^{-1}} = \sum_{k\geqslant 0}\binom{tk+r}{k}z^k;$$

$$\frac{\mathcal{E}_t(z)^r}{1-zt\mathcal{E}_t(z)^t} = \sum_{k\geqslant 0}\frac{(tk+r)^k}{k!}z^k. \qquad (5.61)$$

(Если $tk+r=0$, следует проявить осторожность относительно того, как интерпретировать коэффициент при z^k; каждый такой коэффициент представляет собой полином относительно r. Например, постоянным членом ряда $\mathcal{E}_t(z)^r$ является $r(0+r)^{-1}$, а это равно 1 даже при $r=0$.)

Поскольку уравнения (5.60) и (5.61) справедливы при любом r, мы получаем очень общие соотношения путем перемножения рядов, соответствующих различным степеням r и s, например

$$\begin{aligned}\mathcal{B}_t(z)^r\frac{\mathcal{B}_t(z)^s}{1-t+t\mathcal{B}_t(z)^{-1}} &= \sum_{k\geqslant 0}\binom{tk+r}{k}\frac{r}{tk+r}z^k\sum_{j\geqslant 0}\binom{tj+s}{j}z^j = \\ &= \sum_{n\geqslant 0}z^n\sum_{k\geqslant 0}\binom{tk+r}{k}\frac{r}{tk+r}\binom{t(n-k)+s}{n-k}.\end{aligned}$$

Этот степенной ряд должен быть равен

$$\frac{\mathcal{B}_t(z)^{r+s}}{1-t+t\mathcal{B}_t(z)^{-1}} = \sum_{n\geqslant 0}\binom{tn+r+s}{n}z^n;$$

следовательно, можно приравнять коэффициенты при z^n и получить тождество

$$\sum_k\binom{tk+r}{k}\binom{t(n-k)+s}{n-k}\frac{r}{tk+r} = \binom{tn+r+s}{n}, \qquad \text{целое } n,$$

корректное при всех действительных r, s и t. При $t=0$ это тождество сводится к свертке Вандермонда. (Если же в этой формуле $tk+r$ случайно равен нулю, то множитель $tk+r$ в знаменателе надо рассматривать как сокращающийся с $tk+r$ в числителе биномиального коэффициента. Обе части соотношения представляют собой полиномы от r, s и t.) Аналогичные соотношения выполняются при умножении $\mathcal{B}_t(z)^r$ на $\mathcal{B}_t(z)^s$ и т.д. (табл. 255).

Мы уже знаем, что рассмотрение частных случаев общих результатов — идея обычно неплохая. Например, что получится, если положить $t=1$? Обобщенный бином $\mathcal{B}_1(z)$ очень прост — это

$$\mathcal{B}_1(z) = \sum_{k\geqslant 0}z^k = \frac{1}{1-z};$$

Таблица 255. Общие соотношения свертки, справедливые при целых $n \geqslant 0$

$$\sum_k \binom{tk+r}{k}\binom{tn-tk+s}{n-k}\frac{r}{tk+r} = \binom{tn+r+s}{n} \qquad (5.62)$$

$$\sum_k \binom{tk+r}{k}\binom{tn-tk+s}{n-k}\frac{r}{tk+r}\cdot\frac{s}{tn-tk+s} =$$

$$= \binom{tn+r+s}{n}\frac{r+s}{tn+r+s} \qquad (5.63)$$

$$\sum_k \binom{n}{k}(tk+r)^k(tn-tk+s)^{n-k}\frac{r}{tk+r} = (tn+r+s)^n \qquad (5.64)$$

$$\sum_k \binom{n}{k}(tk+r)^k(tn-tk+s)^{n-k}\frac{r}{tk+r}\cdot\frac{s}{tn-tk+s} =$$

$$= (tn+r+s)^n\frac{r+s}{tn+r+s} \qquad (5.65)$$

таким образом, $\mathcal{B}_1(z)$ не дает нам ничего, что мы бы не получили из свертки Вандермонда. Но $\mathcal{E}_1(z)$ — важная функция

$$\mathcal{E}(z) = \sum_{k\geqslant 0}(k+1)^{k-1}\frac{z^k}{k!} = 1+z+\frac{3}{2}z^2+\frac{8}{3}z^3+\frac{125}{24}z^4+\cdots, \qquad (5.66)$$

которую мы пока что не видели. Она удовлетворяет базовому соотношению

$$\mathcal{E}(z) = e^{z\mathcal{E}(z)}. \qquad (5.67)$$

Ха! Это итерированная степенная функция $\mathcal{E}(\ln z) = z^{z^{z^{\cdot^{\cdot^{\cdot}}}}}$, которая всегда приводила меня в изумление.

Zzzzzz...

Эта функция, впервые изученная Эйлером (Euler) [117] и Эйзенштейном (Eisenstein) [91], возникает во многих приложениях [193, 204].

Частные случаи $t = 2$ и $t = -1$ обобщенного бинома представляют особый интерес, поскольку их коэффициенты все время возникают в задачах с рекурсивной структурой. Таким образом, имеет смысл явно указать эти ряды — для возможности последующих ссылок:

Степенной ряд для $\mathcal{B}_{1/2}(z)^r = (\sqrt{z^2+4}+z)^{2r}/4^r$ тоже весьма поучителен.

$$\mathcal{B}_2(z) = \sum_k \binom{2k}{k}\frac{z^k}{1+k} =$$

$$= \sum_k \binom{2k+1}{k}\frac{z^k}{1+2k} = \frac{1-\sqrt{1-4z}}{2z}, \qquad (5.68)$$

$$\mathcal{B}_{-1}(z) = \sum_k \binom{1-k}{k}\frac{z^k}{1-k} =$$

$$= \sum_k \binom{2k-1}{k}\frac{(-z)^k}{1-2k} = \frac{1+\sqrt{1+4z}}{2}, \qquad (5.69)$$

$$\mathcal{B}_2(z)^r = \sum_k \binom{2k+r}{k}\frac{r}{2k+r}z^k, \qquad (5.70)$$

$$\mathcal{B}_{-1}(z)^r = \sum_k \binom{r-k}{k}\frac{r}{r-k}z^k, \qquad (5.71)$$

$$\frac{\mathcal{B}_2(z)^r}{\sqrt{1-4z}} = \sum_k \binom{2k+r}{k}z^k, \qquad (5.72)$$

$$\frac{\mathcal{B}_{-1}(z)^{r+1}}{\sqrt{1+4z}} = \sum_k \binom{r-k}{k}z^k. \qquad (5.73)$$

Коэффициенты $\binom{2n}{n}\frac{1}{n+1}$ из $\mathcal{B}_2(z)$ называются *числами Каталана* C_n, по имени Эжена Каталана (Eugène Catalan), который написал о них основополагающую работу в 1830-х годах [52]. Последовательность этих чисел начинается следующим образом:

n	0	1	2	3	4	5	6	7	8	9	10
C_n	1	1	2	5	14	42	132	429	1430	4862	16796

Коэффициенты $\mathcal{B}_{-1}(z)$ по сути те же самые, за исключением дополнительной единицы в начале и чередования знаков у других чисел: $\langle 1, 1, -1, 2, -5, 14, \ldots\rangle$. Таким образом, $\mathcal{B}_{-1}(z) = 1 + z\mathcal{B}_2(-z)$. Кроме того, $\mathcal{B}_{-1}(z) = \mathcal{B}_2(-z)^{-1}$.

Завершим этот раздел выводом важного следствия из (5.72) и (5.73), а именно — соотношения, которое показывает дальнейшие связи между функциями $\mathcal{B}_{-1}(z)$ и $\mathcal{B}_2(-z)$:

$$\frac{\mathcal{B}_{-1}(z)^{n+1} - (-z)^{n+1}\mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} = \sum_{k\leqslant n}\binom{n-k}{k}z^k.$$

Данное соотношение выполняется потому, что коэффициенты при z^k в $(-z)^{n+1}\mathcal{B}_2(-z)^{n+1}/\sqrt{1+4z}$ имеют вид

$$[z^k]\,\frac{(-z)^{n+1}\mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} = (-1)^{n+1}[z^{k-n-1}]\,\frac{\mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} =$$

$$= (-1)^{n+1}(-1)^{k-n-1}[z^{k-n-1}]\,\frac{\mathcal{B}_2(z)^{n+1}}{\sqrt{1-4z}} =$$

$$= (-1)^k \binom{2(k-n-1)+n+1}{k-n-1} =$$

$$= (-1)^k \binom{2k-n-1}{k-n-1} =$$

$$= (-1)^k \binom{2k-n-1}{k} =$$

$$= \binom{n-k}{k} = [z^k] \frac{\mathcal{B}_{-1}(z)^{n+1}}{\sqrt{1+4z}}$$

при $k > n$. Эти члены отлично сокращаются друг с другом. Теперь можно воспользоваться (5.68) и (5.69) для получения аналитического выражения

$$\sum_{k \leqslant n} \binom{n-k}{k} z^k =$$
$$= \frac{1}{\sqrt{1+4z}} \left(\left(\frac{1+\sqrt{1+4z}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{1+4z}}{2} \right)^{n+1} \right),$$
$$\text{целое } n \geqslant 0. \qquad (5.74)$$

(Частный случай $z = -1$ возникал в задаче 3 раздела 5.2. Поскольку числа $\frac{1}{2}(1 \pm \sqrt{-3}\,)$ являются корнями шестой степени из единицы, суммы $\sum_{k \leqslant n} \binom{n-k}{k} (-1)^k$ имеют периодическое поведение, что и наблюдалось в указанной задаче.) Аналогично можно объединить (5.70) с (5.71) для сокращения больших коэффициентов и получить

$$\sum_{k<n} \binom{n-k}{k} \frac{n}{n-k} z^k = \left(\frac{1+\sqrt{1+4z}}{2} \right)^n + \left(\frac{1-\sqrt{1+4z}}{2} \right)^n,$$
$$\text{целое } n > 0. \qquad (5.75)$$

5.5 Гипергеометрические функции

Методы, которые мы применяли к биномиальным коэффициентам, очень эффективны в тех случаях, когда они работают, но следует иметь в виду, что зачастую они оказываются слишком специфичными, предназначенными для конкретных задач, т.е. по сути они представляют собой приемы, а не технологии. При работе над той или иной задачей мы часто движемся одновременно в нескольких направлениях, и иногда выясняется, что это — движение по кругу. Биномиальные коэффициенты подобны хамелеонам: они очень легко меняют свою внешность. Таким образом, представляется естественным вопрос — нет ли

Они даже изменчивее хамелеонов: их можно расчленить, а затем сочленить разными способами.

некоего унифицирующего принципа, который работал бы со всем многообразием сумм биномиальных коэффициентов единым систематическим способом? К счастью, ответ на этот вопрос положительный. Такой единый принцип основан на теории определенных бесконечных сумм, именуемых *гипергеометрическими рядами*.

Изучать гипергеометрические ряды много лет назад начали Эйлер (Euler), Гаусс (Gauss) и Риман (Riemann); эти ряды и по сей день остаются предметом обширных исследований. Однако гипергеометрическая форма записи выглядит несколько устрашающе, и к ней надо привыкнуть.

То, что пронесло свою форму записи через века, должно быть и вправду полезным.

Обобщенный гипергеометрический ряд — это степенной ряд от z с $m+n$ параметрами, который определяется через возрастающие факториальные степени следующим образом:

$$F\left(\begin{matrix}a_1, \ldots, a_m \\ b_1, \ldots, b_n\end{matrix}\middle| z\right) = \sum_{k\geqslant 0} \frac{a_1^{\overline{k}} \ldots a_m^{\overline{k}}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}}} \frac{z^k}{k!}. \tag{5.76}$$

Чтобы избежать деления на нуль, ни одно b не может быть нулем или целым отрицательным. В остальном все a и b могут быть любыми. В качестве альтернативы двустрочной записи (5.76) может использоваться запись '$F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)$', которая более удобна как минимум с типографской точки зрения. Величины a называются *верхними параметрами* и входят в числитель членов суммы F. Величины b называются *нижними параметрами* и входят в знаменатель. Последняя величина, z, называется *аргументом*.

В стандартных справочниках для обозначения гипергеометрического ряда с m верхними параметрами и n нижними вместо 'F' часто используется '${}_mF_n$'. Однако эти лишние нижние индексы обычно только загромождают формулу и отнимают время на их переписывание. Всегда можно просто пересчитать параметры, так что эти индексы являются избыточными.

Многие важные функции представляют собой просто частные случаи гипергеометрической функции — и в этом их сильная сторона. Например, простейший случай получается при $m = n = 0$: здесь параметров нет вовсе, и мы получаем знакомый ряд

$$F\left(\middle| z\right) = \sum_{k\geqslant 0} \frac{z^k}{k!} = e^z.$$

Вообще-то, когда m или n равно нулю, запись имеет несколько обескураживающий вид. Чтобы этого избежать, можно добавить дополнительные единицы вверху и внизу:

$$F\left({1 \atop 1}\middle| z\right) = e^z .$$

В общем случае функция не меняется при сокращении параметра из числителя и знаменателя, как и при добавлении двух одинаковых параметров в числитель и знаменатель.

Очередной простейший случай получается при $m = 1$, $a_1 = 1$ и $n = 0$; однако мы заменим параметры на $m = 2$, $a_1 = a_2 = 1$, $n = 1$ и $b_1 = 1$, чтобы было $n > 0$. Этот ряд также нам знаком, поскольку $1^{\overline{k}} = k!$:

$$F\left({1,1 \atop 1}\middle| z\right) = \sum_{k\geqslant 0} z^k = \frac{1}{1-z} .$$

Это наш старый знакомый — геометрический ряд. Гипергеометрическая функция $F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)$ называется так потому, что включает геометрический ряд $F(1, 1; 1; z)$ в качестве весьма частного случая.

Общий случай $m = 1$ и $n = 0$ в действительности легко вычисляется в аналитическом виде при использовании (5.13) и (5.14):

$$F\left({a, 1 \atop 1}\middle| z\right) = \sum_{k\geqslant 0} a^{\overline{k}} \frac{z^k}{k!} = \sum_k \binom{a+k-1}{k} z^k = \frac{1}{(1-z)^a} . \quad (5.77)$$

Если заменить a на $-a$ и z на $-z$, то мы получим биномиальную теорему

$$F\left({-a, 1 \atop 1}\middle| -z\right) = (1+z)^a .$$

Отрицательное целое число в качестве верхнего параметра превращает бесконечный ряд в конечный, поскольку$(-a)^{\overline{k}} = 0$ при $k > a \geqslant 0$ и целом a.

Общий случай $m = 0$, $n = 1$ представляет собой еще один знаменитый ряд, хотя он и не столь известен в литературе по дискретной математике:

$$F\left({1 \atop b, 1}\middle| z\right) = \sum_{k\geqslant 0} \frac{(b-1)!}{(b-1+k)!} \frac{z^k}{k!} = I_{b-1}(2\sqrt{z}) \frac{(b-1)!}{z^{(b-1)/2}} . \quad (5.78)$$

Эта функция I_{b-1} называется "модифицированной функцией Бесселя" порядка $b - 1$. Частный случай $b = 1$ дает нам функцию $F\left({1 \atop 1,1}\middle| z\right) = I_0(2\sqrt{z})$, которая представляет собой интересный ряд $\sum_{k\geqslant 0} z^k/k!^2$.

Частный случай $m = n = 1$ называется конфлюэнтным (вырожденным) гипергеометрическим рядом и часто обозначается буквой M:

$$F\left({a \atop b}\middle| z\right) = \sum_{k\geqslant 0} \frac{a^{\overline{k}}}{b^{\overline{k}}} \frac{z^k}{k!} = M(a, b, z). \tag{5.79}$$

Эта функция, имеющая важные применения в технике, введена Эрнстом Куммером (Ernst Kummer).

Возможно, некоторые из читателей удивлены, почему мы не рассматриваем вопрос о сходимости бесконечного ряда (5.76). Дело в том, что сходимость можно игнорировать, если использовать z просто как формальный символ. Несложно убедиться, что формальные бесконечные суммы вида $\sum_{k\geqslant n} \alpha_k z^k$ с $-\infty < n < \infty$ образуют поле, если их коэффициенты α_k лежат в некотором поле. Такие формальные суммы можно суммировать, вычитать, умножать, делить, дифференцировать и выполнять их композицию, не беспокоясь о сходимости; все выведенные тождества формально будут справедливыми. Так, гипергеометрический ряд $F\left({1,1,1 \atop 1}\middle| z\right) = \sum_{k\geqslant 0} k!\, z^k$ не сходится ни при каком ненулевом z; но, как мы увидим в главе 7, его можно использовать для решения задач. С другой стороны, заменяя z конкретным числовым значением, мы должны быть уверены в том, что бесконечная сумма хорошо определена.

Впрочем, вопрос о сходимости (5.56), (5.57), (5.58), ... тоже не рассматривался.

Очередное усложнение дает нам самый знаменитый гипергеометрический ряд. В действительности именно он был *собственно* гипергеометрическим рядом примерно до 1870 года, когда все было обобщено для произвольных m и n. У этого ряда два верхних параметра и один нижний:

$$F\left({a, b \atop c}\middle| z\right) = \sum_{k\geqslant 0} \frac{a^{\overline{k}} b^{\overline{k}} z^k}{c^{\overline{k}}\, k!}. \tag{5.80}$$

Его часто называют гауссовым гипергеометрическим рядом, поскольку многие его тонкие свойства были впервые доказаны Гауссом (Gauss) в его докторской диссертации в 1812 году [143], хотя Эйлер (Euler) [118] и Пфафф (Pfaff) [292] уже открыли к тому времени некоторые замечательные свойства этого ряда. Одним из важных частных случаев является ряд

"Пожалуй, 95% (если не все 100%) функций, изучаемых в настоящее время в университетах студентами-физиками, инженерами и даже математиками, могут быть охвачены одним-единственным символом $F(a, b; c; x)$."
— *У. У. Сойер (W. W. Sawyer) [318]*

$$\ln(1+z) = zF\left({1,1 \atop 2}\middle| -z\right) = z\sum_{k\geqslant 0} \frac{k!\,k!}{(k+1)!} \frac{(-z)^k}{k!} =$$
$$= z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots.$$

Обратите внимание, что $z^{-1}\ln(1+z)$ представляет собой гипергеометрическую функцию, на сама функция $\ln(1+z)$ гипергеометрической не является, так как гипергеометрический ряд всегда равен 1 при $z=0$.

До сих пор гипергеометрические ряды не дали нам ничего, кроме повода упомянуть пару громких имен. Но мы увидели, что несколько очень различных функций могут рассматриваться как гипергеометрические ряды; именно это и будет основным предметом нашего интереса в дальнейшем. Мы увидим, что большой класс сумм может быть записан как гипергеометрический ряд "каноническим" способом; следовательно, мы получим удобную систему учета фактов о биномиальных коэффициентах.

Какие же ряды являются гипергеометрическими? На этот вопрос легко ответить, если рассмотреть отношение последовательных членов гипергеометрического ряда:

$$F\left({a_1,\ldots,a_m \atop b_1,\ldots,b_n}\middle|\, z\right) = \sum_{k\geqslant 0} t_k, \qquad t_k = \frac{a_1^{\overline{k}}\ldots a_m^{\overline{k}}\, z^k}{b_1^{\overline{k}}\ldots b_n^{\overline{k}}\, k!}.$$

Первый член — $t_0=1$, и прочие члены дают отношения

$$\begin{aligned}\frac{t_{k+1}}{t_k} &= \frac{a_1^{\overline{k+1}}\ldots a_m^{\overline{k+1}}}{a_1^{\overline{k}}\ldots a_m^{\overline{k}}}\,\frac{b_1^{\overline{k}}\ldots b_n^{\overline{k}}}{b_1^{\overline{k+1}}\ldots b_n^{\overline{k+1}}}\,\frac{k!}{(k+1)!}\,\frac{z^{k+1}}{z^k} = \\ &= \frac{(k+a_1)\ldots(k+a_m)\,z}{(k+b_1)\ldots(k+b_n)(k+1)}.\end{aligned} \tag{5.81}$$

Это *рациональная функция* от k, т.е. отношение двух полиномов от k. Согласно основной теореме алгебры любая рациональная функция от k может быть разложена на множители над полем комплексных чисел и приведена к указанному виду. Величины a в числителе представляют собой корни полинома в числителе, взятые с обратным знаком, а b — взятые с обратным знаком корни полинома в знаменателе. Если в знаменатель не входит отдельный множитель $(k+1)$, то его можно включить и в числитель, и в знаменатель. Остается постоянный множитель, который можно обозначить через z. Таким образом, гипергеометрические ряды — это в точности те ряды, первый член которых равен 1, а отношение членов t_{k+1}/t_k представляет собой рациональную функцию от k.

Предположим, например, что у нас есть бесконечный ряд с отношением членов

$$\frac{t_{k+1}}{t_k} = \frac{k^2+7k+10}{4k^2+1},$$

которое является рациональной функцией от k. Полином в числителе разделяется на множители $(k+2)(k+5)$, а знаменатель представляет собой произведение $4(k+i/2)(k-i/2)$. Поскольку в знаменателе отсутствует необходимый множитель $(k+1)$, запишем отношение членов в виде

$$\frac{t_{k+1}}{t_k} = \frac{(k+2)(k+5)(k+1)(1/4)}{(k+i/2)(k-i/2)(k+1)}$$

и получим окончательный результат: заданный ряд — это

$$\sum_{k\geqslant 0} t_k = t_0\, F\left(\begin{matrix} 2, 5, 1 \\ i/2, -i/2 \end{matrix}\,\middle|\, 1/4\right).$$

Таким образом, установлен общий метод нахождения гипергеометрического представления некоторой заданной величины S, если такое представление возможно. Вначале S записывается в виде бесконечного ряда, первый член которого отличен от нуля. Выберем такие обозначения, чтобы это был ряд $\sum_{k\geqslant 0} t_k$ с $t_0 \neq 0$. Затем вычисляется t_{k+1}/t_k. Если отношение членов не является рациональной функцией от k, значит, нам не повезло. В противном случае мы выражаем его в виде (5.81); это дает нам параметры $a_1, \ldots, a_m$, $b_1, \ldots, b_n$ и аргумент z, такой, что $S = t_0\, F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)$.

(Сейчас самое время размяться с помощью упр. 11.)

Гауссов гипергеометрический ряд при необходимости подчеркнуть важность отношения членов может быть записан и в рекурсивном виде:

$$F\left(\begin{matrix} a, b \\ c \end{matrix}\,\middle|\, z\right) = \\ = 1 + \frac{a}{1}\frac{b}{c}z\left(1 + \frac{a+1}{2}\frac{b+1}{c+1}z\left(1 + \frac{a+2}{3}\frac{b+2}{c+2}z\,(1+\cdots)\right)\right).$$

Давайте попробуем переформулировать тождества с биномиальными коэффициентами, выведенные ранее в этой главе, в гипергеометрическом виде. Например, давайте выясним, как выглядит правило параллельного суммирования (сложения по диагонали)

$$\sum_{k\leqslant n} \binom{r+k}{k} = \binom{r+n+1}{n}, \qquad \text{целое } n,$$

в гипергеометрической записи. Для этого надо записать данную сумму в виде бесконечного ряда, начинающегося с $k=0$, так что заменим k на $n-k$:

$$\sum_{k\geqslant 0} \binom{r+n-k}{n-k} = \sum_{k\geqslant 0} \frac{(r+n-k)!}{r!\,(n-k)!} = \sum_{k\geqslant 0} t_k\,.$$

Формально данный ряд бесконечный, но на самом деле он конечен, потому что $(n-k)!$ в знаменателе делает $t_k = 0$ при $k > n$. (Позже мы увидим, что $1/x!$ определено при всех x и что $1/x! = 0$ при целом отрицательном x. Ну а пока махнем рукой на все формальности до тех пор, пока не наберемся побольше гипергеометрического опыта.) Отношение членов в данном случае равно

$$\frac{t_{k+1}}{t_k} = \frac{(r+n-k-1)!\,r!\,(n-k)!}{r!\,(n-k-1)!\,(r+n-k)!} = \frac{n-k}{r+n-k} = \\ = \frac{(k+1)(k-n)(1)}{(k-n-r)(k+1)}.$$

Кроме того, $t_0 = \binom{r+n}{n}$. Следовательно, правило параллельного суммирования эквивалентно гипергеометрическому тождеству

$$\binom{r+n}{n} F\left({1,\,-n \atop -n-r}\middle|\,1\right) = \binom{r+n+1}{n}.$$

Деление на $\binom{r+n}{n}$ дает несколько более простую версию

$$F\left({1,\,-n \atop -n-r}\middle|\,1\right) = \frac{r+n+1}{r+1}, \qquad \text{если } \binom{r+n}{n} \ne 0. \tag{5.82}$$

Давайте попробуем еще. В тождестве (5.16),

$$\sum_{k\leqslant m} \binom{r}{k}(-1)^k = (-1)^m\binom{r-1}{m}, \qquad \text{целое } m,$$

после замены k на $m-k$ отношение членов принимает вид $(k-m)/(r-m+k+1) = (k+1)(k-m)(1)/(k-m+r+1)(k+1)$; следовательно, (5.16) выражает в аналитическом виде функцию

$$F\left({1,\,-m \atop -m+r+1}\middle|\,1\right).$$

По сути это то же, что и гипергеометрическая функция в левой части (5.82), но с m вместо n и $r+1$ вместо $-r$. Поэтому тождество (5.16) могло бы быть выведено из (5.82) — гипергеометрической версии (5.9). (Не удивительно, что (5.16) оказалось легко доказать, воспользовавшись (5.9).)

Мало нам беспорядков — теперь еще и вырождение!

Перед тем как идти дальше, мы должны подумать о вырожденных случаях, поскольку гипергеометрические функции не определены, когда нижний параметр равен нулю или представляет собой отрицательное целое число. Тождество, соответствующее правилу параллельного суммирования, обычно применяется при целых положительных r и n; но тогда $-n-r$ — отрицательное целое число, и гипергеометрический ряд (5.76) не определен. Как тогда можно считать тождество (5.82) законным? Ответ в том, что мы можем перейти к пределу $F\left({1,-n \atop -n-r+\epsilon}\middle|1\right)$ при $\epsilon \to 0$.

Позднее в этой главе подобные вещи будут рассмотрены более обстоятельно, а пока просто отдадим себе отчет в том, что некоторые знаменатели могут быть "взрывоопасными". Интересно, однако, что первая сумма, которую мы пытались выразить гипергеометрически, оказалась вырожденной.

Сначала мы доказывали тождества для целого r и использовали полиномиальную аргументацию, чтобы показать их справедливость в общем случае. А теперь мы сначала доказываем их для иррационального r, а затем используем предельную аргументацию, чтобы показать их справедливость для целых чисел!

Возможно, другим больным местом в выводе (5.82) является представление $\binom{r+n-k}{n-k}$ в виде $(r+n-k)!/r!\,(n-k)!$. Такое представление некорректно при отрицательном целом r, поскольку $(-m)!$ должно быть равным ∞, если следовать правилу

$$0! = 0\cdot(-1)\cdot(-2)\cdot\ldots\cdot(-m+1)\cdot(-m)!.$$

И вновь для получения целочисленных значений нам надо рассматривать предел $r+\epsilon$ при $\epsilon \to 0$.

Но факториальное представление $\binom{r}{k} = r!/k!\,(r-k)!$ определено только при целом r. Если же мы хотим работать с гипергеометрическими представлениями, то нам нужна такая факториальная функция, которая была бы определена при всех комплексных числах. К счастью, такая функция имеется и может быть определена несколькими способами. Вот одно из наиболее полезных определений $z!$ (точнее — $1/z!$):

$$\frac{1}{z!} = \lim_{n\to\infty}\binom{n+z}{n}n^{-z}. \tag{5.83}$$

(См. упр. 21. Эйлер [99, 100, 72] открыл это определение в 22 года.) Можно показать, что данный предел существует для всех комплексных z, и что он равен нулю тогда, когда z — отрицательное целое число. Другое важное определение —

$$z! = \int_0^\infty t^z e^{-t}\,dt, \qquad \text{если } \Re z > -1. \tag{5.84}$$

Этот интеграл существует только тогда, когда действительная часть z больше -1, но можно воспользоваться формулой

$$z! = z(z-1)!, \tag{5.85}$$

чтобы распространить определение на все комплексные z (за исключением отрицательных целых чисел). Еще одно определение следует из приближения Стирлинга для $\ln z!$ из (5.47). Все эти подходы приводят к одной и той же обобщенной факториальной функции.

Имеется еще одна очень похожая функция, именуемая *гамма-функцией*, которая связана с обычными факториалами примерно так же, как возрастающие степени связаны с убывающими. В стандартных справочниках факториалы и гамма-функции часто используются совместно, и при необходимости удобно переходить от одних к другим, используя следующие формулы:

$$\Gamma(z+1) = z!\,, \tag{5.86}$$

$$(-z)!\,\Gamma(z) = \frac{\pi}{\sin \pi z}\,. \tag{5.87}$$

Как вы записываете z в степени $\overline{w}$, когда $\overline{w}$ — комплексно сопряженное с w число?

Как $z^{(\overline{w})}$.

Эти обобщенные факториалы можно использовать для определения обобщенных факториальных степеней, когда z и w представляют собой произвольные комплексные числа:

$$z^{\underline{w}} = \frac{z!}{(z-w)!}\,, \tag{5.88}$$

$$z^{\overline{w}} = \frac{\Gamma(z+w)}{\Gamma(z)}\,. \tag{5.89}$$

Единственная оговорка состоит в том, что если эти формулы дают неопределенность типа ∞/∞, то надо использовать соответствующие пределы. (Эти формулы никогда не приводят к неопределенности вида $0/0$, так как факториалы и гамма-функции никогда не обращаются в нуль.) Биномиальный коэффициент можно записать как

$$\binom{z}{w} = \lim_{\zeta\to z}\lim_{\omega\to w}\frac{\zeta!}{\omega!\,(\zeta-\omega)!} \tag{5.90}$$

где z и w — произвольные комплексные числа.

Понятно, что нижний индекс достигает своего предела первым. Вот почему $\binom{z}{w}$ обращается в нуль при отрицательном целом w.

Это значение бесконечно, когда z — отрицательное целое, а w не является целым.

Вооружившись такими инструментами, как обобщенные факториалы, мы можем вернуться к нашей цели — выявить гипергеометрическую сущность выведенных ранее тождеств. Биномиальная теорема (5.13), как и следовало ожидать, оказывается ни чем иным, как суммой (5.77). Так что следующее наиболее интересное для исследования соотношение — это свертка Вандермонда (5.27):

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, \qquad \text{целое } n.$$

Здесь k-й член представляет собой

$$t_k = \frac{r!}{(r-k)!\,k!}\,\frac{s!}{(s-n+k)!\,(n-k)!}\,,$$

и больше не надо избегать использования в этих выражениях обобщенных факториалов. Всякий раз, когда t_k содержит множитель наподобие $(\alpha+k)!$ со знаком 'плюс' перед k, в соответствии с (5.85) мы получаем $(\alpha+k+1)!/(\alpha+k)! = k+\alpha+1$ в отношении t_{k+1}/t_k; это дает '$\alpha+1$' в соответствующем гипергеометрическом представлении — в качестве верхнего параметра, если $(\alpha+k)!$ был в числителе t_k, или в качестве нижнего параметра в противном случае. Аналогично множитель типа $(\alpha-k)!$ приводит к $(\alpha-k-1)!/(\alpha-k)! = (-1)/(k-\alpha)$; это дает '$-\alpha$' в другой совокупности параметров (с измененными ролями верхнего и нижнего параметров) и меняет знак аргумента гипергеометрической функции. Множители наподобие $r!$, не зависящие от k, входят в t_0, но исчезают из отношений членов. Используя такие трюки, можно без дальнейших вычислений показать, что отношение членов в (5.27) представляет собой

$$\frac{t_{k+1}}{t_k} = \frac{k-r}{k+1}\,\frac{k-n}{k+s-n+1},$$

умноженное на $(-1)^2 = 1$, и свертка Вандермонда приобретает вид

$$\binom{s}{n} F\left(\begin{matrix}-r,\,-n\\ s-n+1\end{matrix}\,\middle|\,1\right) = \binom{r+s}{n}. \qquad (5.91)$$

Это равенство можно использовать для определения $F(a,b;c;z)$ в общем случае, когда $z=1$, а b — отрицательное целое число.

Перепишем (5.91) в таком виде, который облегчит нам просмотр таблицы, если потребуется вычислить некоторую новую сумму. В результате оказывается, что

$$F\left(\begin{matrix}a,b\\ c\end{matrix}\,\middle|\,1\right) = \frac{\Gamma(c-a-b)\,\Gamma(c)}{\Gamma(c-a)\,\Gamma(c-b)}; \quad \begin{matrix}\text{целое } b \leqslant 0\\ \text{или } \Re c > \Re a + \Re b.\end{matrix} \qquad (5.92)$$

Свертка Вандермонда (5.27) охватывает только тот случай, когда один из верхних параметров, скажем, b, представляет собой целое не положительное число; но Гаусс (Gauss) доказал [143], что (5.92) справедливо и тогда, когда a, b и c — комплексные числа, действительные части которых удовлетворяют условию $\Re c > \Re a + \Re b$. В остальных случаях бесконечный ряд $F\left({a,b \atop c}\middle|1\right)$ не сходится. При $b=-n$ соотношение можно переписать в более удобном виде с факториальными степенями вместо гамма-функций:

Несколько недель назад мы изучали то, что Гаусс сделал в детсадовском возрасте. Теперь мы изучаем материал, выходящий за рамки его диссертации. Что же с нами делают?!

$$F\left(\begin{matrix}a,\,-n\\ c\end{matrix}\,\middle|\,1\right) = \frac{(c-a)^{\overline{n}}}{c^{\overline{n}}} = \frac{(a-c)^{\underline{n}}}{(-c)^{\underline{n}}}, \quad \text{целое } n \geqslant 0. \qquad (5.93)$$

Оказывается, все пять соотношений из табл. 218 являются частными случаями свертки Вандермонда; все они охватываются формулой (5.93), если уделить должное внимание вырожденным случаям.

Обратите внимание, что (5.82) представляет собой просто частный случай соотношения (5.93) при $a = 1$. Таким образом, на самом деле незачем запоминать соотношение (5.82); более того, на самом деле нам не надо тождество (5.9), которое привело нас к (5.82), несмотря на то что в табл. 222 его предлагалось запомнить. Компьютерная программа для формульных преобразований, столкнувшись с задачей вычисления суммы $\sum_{k\leqslant n}\binom{r+k}{k}$, могла бы преобразовать ее в гипергеометрическую форму и подставить в общее выражение для свертки Вандермонда.

В задаче 1 из раздела 5.2 требовалось найти значение суммы

$$\sum_{k\geqslant 0}\binom{m}{k}\Big/\binom{n}{k}.$$

Для этой задачи естественным является гипергеометрическое представление, и, немного попрактиковавшись, любой гипергеометр сможет тут же записать эту сумму как $F(1,-m;-n;1)$. Да, эта задача — еще одно жалкое подобие Вандермонда!

Сумма в задачах 2 и 4 дает функцию $F(2, 1-n; 2-m; 1)$. (Сначала следует заменить k на $k+1$.) А “страшная” сумма из задачи 6 оказывается всего лишь функцией $F(n+1,-n;2;1)$. Получается, что, кроме скрытых версий свертки Вандермонда, суммировать больше нечего?

Ну зачем же так? Задача 3 несколько иная. В ней мы имеем дело с частным случаем суммы общего вида $\sum_k\binom{n-k}{k}z^k$, рассмотренной в (5.74), и она приводит нас к аналитическому выражению для

$$F\left({1+2\lceil n/2\rceil,\,-n \atop 1/2}\middle|\,-z/4\right).$$

Мы также доказали нечто новое в (5.55), когда рассматривали коэффициенты функции $(1-z)^r(1+z)^r$:

$$F\left({1-c-2n,\,-2n \atop c}\middle|\,-1\right) = (-1)^n\frac{(2n)!}{n!}\frac{(c-1)!}{(c+n-1)!}, \quad \text{целое } n \geqslant 0.$$

При обобщении на комплексные числа эта формула называется *формулой Куммера*:

$$F\left({a,\,b \atop 1+b-a}\middle|\,-1\right) = \frac{(b/2)!}{b!}(b-a)^{\underline{b/2}}. \tag{5.94}$$

(Эрнст Куммер (Ernst Kummer) [229] доказал ее в 1836 году.)

Интересно сравнить две эти формулы. Заменяя c на $1-2n-a$, находим, что результаты согласуются тогда и только тогда, когда

$$(-1)^n \frac{(2n)!}{n!} = \lim_{b\to-2n} \frac{(b/2)!}{b!} = \lim_{x\to-n} \frac{x!}{(2x)!} \tag{5.95}$$

при целом положительном n. Предположим, например, что $n = 3$; тогда должно быть $-6!/3! = \lim_{x\to-3} x!/(2x)!$. Мы знаем, что и $(-3)!$, и $(-6)!$ равны бесконечности. Мы могли бы игнорировать это затруднение и сделать вид, что $(-3)! = (-3)(-4)(-5)(-6)!$, так что два имеющихся $(-6)!$ сократятся. Но на этот соблазн следует ответить решительным отказом, так как он приводит нас к неверному результату! Пределом $x!/(2x)!$ при $x \to -3$ согласно (5.95) является не $(-3)(-4)(-5)$, а $-6!/3! = (-4)(-5)(-6)$.

Правильный способ вычисления предела в (5.95) состоит в применении уравнения (5.87), которое связывает факториалы с отрицательным аргументом и гамма-функции с положительным аргументом. Если заменить x на $-n-\epsilon$ и положить $\epsilon \to 0$, то двойное применение (5.87) дает

$$\frac{(-n-\epsilon)!}{(-2n-2\epsilon)!}\,\frac{\Gamma(n+\epsilon)}{\Gamma(2n+2\epsilon)} = \frac{\sin(2n+2\epsilon)\pi}{\sin(n+\epsilon)\pi}\,.$$

Поскольку $\sin(x+y) = \sin x\cos y + \cos x \sin y$, получившееся отношение синусов приводится к виду

$$\frac{\cos 2n\pi\,\sin 2\epsilon\pi}{\cos n\pi\,\sin \epsilon\pi} = (-1)^n\bigl(2+O(\epsilon)\bigr)$$

при помощи методов из главы 9. Таким образом, согласно (5.86) мы, как и требовалось, имеем

$$\lim_{\epsilon\to0} \frac{(-n-\epsilon)!}{(-2n-2\epsilon)!} = 2(-1)^n \frac{\Gamma(2n)}{\Gamma(n)} = \\ = 2(-1)^n \frac{(2n-1)!}{(n-1)!} = (-1)^n \frac{(2n)!}{n!}\,.$$

Завершим наш обзор тем, что установим заново все прочие соотношения, с которыми мы имели дело в этой главе, нарядив их в гипергеометрическое платье. Сумма с тремя биномиальными коэффициентами в (5.29) может быть записана как

$$F\left({1-a-2n,\ 1-b-2n,\ -2n \atop a,\ b}\,\middle|\,1\right) = \\ = (-1)^n \frac{(2n)!}{n!}\,\frac{(a+b+2n-1)^{\overline{n}}}{a^{\overline{n}}\,b^{\overline{n}}}\,, \qquad \text{целое } n \geqslant 0.$$

Если обобщить эту формулу на комплексные числа, то получится *формула Диксона*:

$$F\left({a,\ b,\ c \atop 1+c-a,\ 1+c-b}\middle|\,1\right) = \frac{(c/2)!}{c!}\,\frac{(c-a)^{\underline{c/2}}\,(c-b)^{\underline{c/2}}}{(c-a-b)^{\underline{c/2}}}\,, \qquad (5.96)$$

$$\Re a+\Re b < 1+\Re c/2.$$

Одной из наиболее общих формул из числа встречавшихся является сумма с тремя биномиальными коэффициентами (5.28), которая приводит к *тождеству Заальшютца*:

Историческая справка: Заальшютц (Saalschütz) [315] независимо открыл эту формулу почти через 100 лет после первой ее публикации Пфаффом (Pfaff) [292]. Взятие предела при $n \to \infty$ дает уравнение (5.92).

$$F\left({a,\ b,\ -n \atop c,\ a+b-c-n+1}\middle|\,1\right) = \frac{(c-a)^{\overline{n}}\,(c-b)^{\overline{n}}}{c^{\overline{n}}\,(c-a-b)^{\overline{n}}} = \qquad (5.97)$$

$$= \frac{(a-c)^{\underline{n}}\,(b-c)^{\underline{n}}}{(-c)^{\underline{n}}\,(a+b-c)^{\underline{n}}}\,, \qquad \text{целое } n \geqslant 0.$$

Эта формула дает при $z = 1$ значение обобщенного гипергеометрического ряда с тремя верхними и двумя нижними параметрами при условии, что один из верхних параметров является целым неположительным числом и что $b_1 + b_2 = a_1 + a_2 + a_3 + 1$. (Если сумма нижних параметров больше суммы верхних параметров не на 1, а на 2, то формула из упр. 25 может быть использована для выражения $F(a_1, a_2, a_3; b_1, b_2; 1)$ через две гипергеометрические функции, удовлетворяющие тождеству Заальшютца.)

Доставшееся нам ценой большого труда соотношение из задачи 8 раздела 5.2 сводится к

$$\frac{1}{1+x}\,F\left({x+1,\ n+1,\ -n \atop 1,\ x+2}\middle|\,1\right) = (-1)^n x^{\underline{n}}\, x^{\underline{-n-1}}.$$

Эх... это всего лишь частный случай тождества Заальшютца (5.97), когда $c = 1$, так что можно было бы сберечь массу усилий, перейдя к непосредственному гипергеометрическому представлению!

А как насчет задачи 7? Та сверхтрудная сумма дает формулу

$$F\left({n+1,\ m-n,\ 1,\ \frac12 \atop \frac12 m+1,\ \frac12 m+\frac12,\ 2}\middle|\,1\right) = \frac{m}{n}\,, \qquad \text{целое } n \geqslant m > 0,$$

которая представляет собой первый случай, когда нам встретились три нижних параметра. Так что выглядит это довольно ново. Но на самом деле ничего нового здесь нет: если воспользоваться упр. 26, то левая часть может быть заменена на

$$F\left({n,\ m-n-1,\ -\frac12 \atop \frac12 m,\ \frac12 m-\frac12}\middle|\,1\right) - 1\,,$$

и мы снова получим тождество Заальшютца.

Что ж, еще один опыт дефляции, но одновременно еще один повод оценить силу гипергеометрических методов.

Историческая справка: исключительная уместность применения гипергеометрических рядов к тождествам с биномиальными коэффициентами впервые была отмечена Джорджем Эндрюсом (George Andrews) в 1974 году [9, разд. 5].

Соотношения свертки из табл. 255 не обладают гипергеометрическими эквивалентами, ибо их отношения членов являются рациональными функциями от k только тогда, когда t — целое число. Равенства (5.64) и (5.65) не являются гипергеометрическими, даже когда $t = 1$. Но можно принять во внимание (5.62) для случая, когда t принимает небольшие целочисленные значения:

$$F\left({\frac{1}{2}r,\ \frac{1}{2}r+\frac{1}{2},\ -n,\ -n-s \atop r+1,\ -n-\frac{1}{2}s,\ -n-\frac{1}{2}s+\frac{1}{2}}\,\middle|\,1\right) = \binom{r+s+2n}{n}\Big/\binom{s+2n}{n};$$

$$F\left({\frac{1}{3}r,\ \frac{1}{3}r+\frac{1}{3},\ \frac{1}{3}r+\frac{2}{3},\ -n,\ -n-\frac{1}{2}s,\ -n-\frac{1}{2}s+\frac{1}{2} \atop \frac{1}{2}r+\frac{1}{2},\ \frac{1}{2}r+1,\ -n-\frac{1}{3}s,\ -n-\frac{1}{3}s+\frac{1}{3},\ -n-\frac{1}{3}s+\frac{2}{3}}\,\middle|\,1\right) = $$

$$= \binom{r+s+3n}{n}\Big/\binom{s+3n}{n}.$$

Первая из этих формул вновь дает решение задачи 7, если заменить величины (r, s, n) на $(1, m-2n-1, n-m)$ соответственно.

Наконец, "неожиданная" сумма (5.20) дает нам неожиданное гипергеометрическое соотношение, которое оказывается весьма поучительным. Рассмотрим это не спеша. Сначала превратим ее в бесконечную сумму:

$$\sum_{k\leqslant m}\binom{m+k}{k}2^{-k} = 2^m \qquad\Longleftrightarrow\qquad \sum_{k\geqslant 0}\binom{2m-k}{m-k}2^k = 2^{2m}.$$

Отношение членов $(2m-k)!\,2^k/m!\,(m-k)!$ равно $2(k-m)/(k-2m)$, так что получается гипергеометрическое тождество с $z = 2$:

$$\binom{2m}{m}F\left({1,\ -m \atop -2m}\,\middle|\,2\right) = 2^{2m}, \qquad \text{целое } m \geqslant 0. \tag{5.98}$$

Но взгляните на нижний параметр '$-2m$'. Целые отрицательные числа запрещены, так что данное тождество не определено!

Теперь самое время тщательно рассмотреть подобные предельные случаи, как и было обещано ранее, поскольку гипергеометрические функции в точках "вырождения" зачастую могут быть вычислены путем приближения к ним из ближайших "невырожденных" точек. Поступая так, следует быть осторожным, поскольку если пределы берутся разными способами, то и результаты могут получиться разными. Например, вот два предела, которые оказываются совершенно разными, если один из параметров увеличен на ϵ:

$$
\begin{aligned}
\lim_{\epsilon\to 0} F\left(\begin{matrix}-1+\epsilon,\ -3\\ -2+\epsilon\end{matrix}\,\middle|\,1\right) &= \lim_{\epsilon\to 0}\Bigl(1+\frac{(-1+\epsilon)(-3)}{(-2+\epsilon)\,1!}+\frac{(-1+\epsilon)(\epsilon)(-3)(-2)}{(-2+\epsilon)(-1+\epsilon)\,2!}\\
&\qquad +\frac{(-1+\epsilon)(\epsilon)(1+\epsilon)(-3)(-2)(-1)}{(-2+\epsilon)(-1+\epsilon)(\epsilon)\,3!}\Bigr) =\\
&= 1-\tfrac{3}{2}+0+\tfrac{1}{2} = 0\,;\\
\lim_{\epsilon\to 0} F\left(\begin{matrix}-1,-3\\ -2+\epsilon\end{matrix}\,\middle|\,1\right) &= \lim_{\epsilon\to 0}\Bigl(1+\frac{(-1)(-3)}{(-2+\epsilon)\,1!}+0+0\Bigr) =\\
&= 1-\tfrac{3}{2}+0+0 = -\tfrac{1}{2}\,.
\end{aligned}
$$

Аналогично мы определили $\binom{-1}{-1} = 0 = \lim_{\epsilon\to 0}\binom{-1+\epsilon}{-1}$; это не то же самое, что $\lim_{\epsilon\to 0}\binom{-1+\epsilon}{-1+\epsilon} = 1$. Чтобы правильно трактовать (5.98) как предел, следует понимать, что верхний параметр $-m$ использован для того, чтобы обратить в нуль все члены ряда $\sum_{k\geqslant 0}\binom{2m-k}{m-k}2^k$ при $k > m$; это значит, что данному соотношению надлежит придать следующую более точную формулировку:

$$\binom{2m}{m}\lim_{\epsilon\to 0} F\left(\begin{matrix}1,\ -m\\ -2m+\epsilon\end{matrix}\,\middle|\,2\right) = 2^{2m}\,, \qquad \text{целое } m \geqslant 0. \tag{5.99}$$

Каждый член этого предельного соотношения вполне определен, так как множитель $(-2m)^{\overline{k}}$ в знаменателе не обращается в нуль до тех пор, пока $k > 2m$. Поэтому данный предел дает именно ту сумму (5.20), с которой мы начали.

5.6 Гипергеометрические преобразования

Теперь должно быть ясно, что база данных известных гипергеометрических представлений в аналитическом виде представляет собой полезный инструмент для суммирования биномиальных коэффициентов. Мы просто преобразуем заданную сумму к ее каноническому гипергеометрическому виду и ищем ее в таблице. Если она там есть — прекрасно, решение найдено. Если нет — ее можно добавить в базу данных, если она окажется выражаемой в аналитическом виде. Можно также включить в таблицу записи, которые гласят "эта сумма в аналитическом виде в общем случае не выражается". Например, сумма $\sum_{k\leqslant m}\binom{n}{k}$ соответствует гипергеометрической функции

$$\binom{n}{m} F\left(\begin{matrix}1,\ -m\\ n-m+1\end{matrix}\,\middle|\,-1\right), \qquad \text{целые } n \geqslant m \geqslant 0, \tag{5.100}$$

которая выражается в аналитическом виде тогда и только тогда, когда m близко к 0, $\frac{1}{2}n$ или n.

Но это не конец истории, поскольку гипергеометрические функции подчиняются и собственным тождествам. Это означает, что каждая аналитическая запись гипергеометрической функции приводит к дополнительным аналитическим формулам и новым записям в базе данных. Например, тождества из упр. 25 и 26 говорят о том, как преобразовать одну гипергеометрическую функцию в две другие с похожими, но отличными параметрами. В свою очередь, эти функции также могут быть преобразованы.

На самом деле это не база данных, а "база знаний".

В 1797 году И. Ф. Пфафф (J. F. Pfaff) [292] открыл удивительный *закон отражения*,

$$\frac{1}{(1-z)^a} F\left({a, b \atop c} \middle| \frac{-z}{1-z}\right) = F\left({a,\ c-b \atop c} \middle| z\right), \qquad (5.101)$$

который служит примером преобразования другого типа. Это формальное тождество относительно степенных рядов, если величина $(-z)^k/(1-z)^{k+a}$ заменена бесконечным рядом $(-z)^k\bigl(1+\binom{k+a}{1}z+\binom{k+a+1}{2}z^2+\cdots\bigr)$, полученным при разложении в ряд левой части (см. упр. 50). Этот закон можно использовать для вывода новых формул из уже известных при $z \neq 1$.

Например, формула Куммера (5.94) может быть объединена с правилом (5.101), если параметры выбраны так, что применимы оба тождества:

$$\begin{aligned} 2^{-a} F\left({a,\ 1-a \atop 1+b-a} \middle| \frac{1}{2}\right) &= F\left({a,\ b \atop 1+b-a} \middle| -1\right) = \\ &= \frac{(b/2)!}{b!}(b-a)^{\underline{b/2}}. \end{aligned} \qquad (5.102)$$

Теперь можно положить $a = -n$ и перейти от данного равенства к новому тождеству относительно биномиальных коэффициентов, которое может понадобиться в один прекрасный день:

$$\begin{aligned} \sum_{k\geqslant 0} \frac{(-n)^{\overline{k}}\,(1+n)^{\overline{k}}}{(1+b+n)^{\overline{k}}}\,\frac{2^{-k}}{k!} &= \sum_k \binom{n}{k}\left(\frac{-1}{2}\right)^k \binom{n+k}{k} \Big/ \binom{n+b+k}{k} = \\ &= 2^{-n}\,\frac{(b/2)!\,(b+n)!}{b!\,(b/2+n)!}, \quad \text{целое } n \geqslant 0. \end{aligned} \qquad (5.103)$$

Например, при $n = 3$ данное тождество гласит, что

$$\begin{aligned} 1 - 3\,\frac{4}{2(4+b)} + 3\,\frac{4\cdot 5}{4(4+b)(5+b)} - \frac{4\cdot 5\cdot 6}{8(4+b)(5+b)(6+b)} &= \\ = \frac{(b+3)(b+2)(b+1)}{(b+6)(b+4)(b+2)}&. \end{aligned}$$

В это почти невозможно поверить, но это так при всех b (за исключением случаев, когда какой-нибудь из сомножителей в знаменателе обращается в нуль).

Неплохо. Давайте попробуем еще? Может, нам удастся найти формулу, которой и в самом деле можно будет похвастать перед друзьями? Что даст нам закон Пфаффа, если применить его к выражению несколько странного вида (5.99), где $z = 2$? В этом случае мы полагаем $a = -m$, $b = 1$ и $c = -2m + \epsilon$ и получаем

$$(-1)^m \lim_{\epsilon\to 0} F\left({-m, 1 \atop -2m+\epsilon}\middle| 2\right) = \lim_{\epsilon\to 0} F\left({-m, -2m-1+\epsilon \atop -2m+\epsilon}\middle| 2\right) =$$
$$= \lim_{\epsilon\to 0} \sum_{k\geqslant 0} \frac{(-m)^{\overline{k}}\,(-2m-1+\epsilon)^{\overline{k}}}{(-2m+\epsilon)^{\overline{k}}} \frac{2^k}{k!} =$$
$$= \sum_{k\leqslant m} \binom{m}{k} \frac{(2m+1)^{\underline{k}}}{(2m)^{\underline{k}}} (-2)^k,$$

поскольку ни один из предельных членов не близок к нулю. Это приводит к другой удивительной формуле,

Истерическое примечание: если у вас получается другой результат, см. упр. 51.

$$\sum_{k\leqslant m} \binom{m}{k} \frac{2m+1}{2m+1-k} (-2)^k = (-1)^m 2^{2m} \Big/ \binom{2m}{m} =$$
$$= 1 \Big/ \binom{-1/2}{m}, \quad \text{целое } m \geqslant 0. \qquad (5.104)$$

При $m = 3$, например, сумма равна

$$1 - 7 + \frac{84}{5} - 14 = -\frac{16}{5},$$

а $\binom{-1/2}{3}$ действительно равно $-\frac{5}{16}$.

Когда мы рассматривали тождества с биномиальными коэффициентами и приводили их к гипергеометрическому виду, мы не придали значения соотношению (5.19), поскольку это было соотношение между двумя суммами, а не аналитическое выражение. Однако теперь (5.19) можно рассматривать как соотношение между гипергеометрическими рядами. Если продифференцировать его n раз по y, а затем заменить k на $m - n - k$, то получим

$$\sum_{k\geqslant 0} \binom{m+r}{m-n-k} \binom{n+k}{n} x^{m-n-k} y^k =$$
$$= \sum_{k\geqslant 0} \binom{-r}{m-n-k} \binom{n+k}{n} (-x)^{m-n-k} (x+y)^k.$$

Это приводит к следующему гипергеометрическому преобразованию:

$$F\left({a,\,-n \atop c}\,\middle|\,z\right) = \frac{(a-c)^{\underline{n}}}{(-c)^{\underline{n}}}\,F\left({a,\,-n \atop 1-n+a-c}\,\middle|\,1-z\right), \quad \begin{matrix}\text{целое}\\ n\geqslant 0.\end{matrix} \tag{5.105}$$

Обратите внимание, что при $z = 1$ это сводится к свертке Вандермонда (5.93).

По-видимому, в дифференцировании что-то есть, если данный пример сколь-нибудь показателен, — оно оказалось полезным и в главе 2 при суммировании $x + 2x^2 + \cdots + nx^n$. Посмотрим, что получится при дифференцировании обобщенного гипергеометрического ряда по z:

$$\begin{aligned}\frac{d}{dz}F\left({a_1,\ldots,a_m \atop b_1,\ldots,b_n}\,\middle|\,z\right) &= \sum_{k\geqslant 1}\frac{a_1^{\overline{k}}\ldots a_m^{\overline{k}}\,z^{k-1}}{b_1^{\overline{k}}\ldots b_n^{\overline{k}}\,(k-1)!} = \\ &= \sum_{k+1\geqslant 1}\frac{a_1^{\overline{k+1}}\ldots a_m^{\overline{k+1}}\,z^{k}}{b_1^{\overline{k+1}}\ldots b_n^{\overline{k+1}}\,k!} = \\ &= \sum_{k\geqslant 0}\frac{a_1(a_1+1)^{\overline{k}}\ldots a_m(a_m+1)^{\overline{k}}\,z^{k}}{b_1(b_1+1)^{\overline{k}}\ldots b_n(b_n+1)^{\overline{k}}\,k!} = \\ &= \frac{a_1\ldots a_m}{b_1\ldots b_n}\,F\left({a_1+1,\ldots,a_m+1 \atop b_1+1,\ldots,b_n+1}\,\middle|\,z\right).\end{aligned} \tag{5.106}$$

Параметры выносятся и увеличиваются на 1.

Дифференцирование можно также применить для того, чтобы "отщипнуть" только один параметр, не трогая остальные. Для этого воспользуемся оператором

$$\vartheta = z\frac{d}{dz},$$

Как вы произносите ϑ?

(Не знаю, как вы, а TeX называет ее 'vartheta'.)

который дифференцирует функцию и умножает полученный результат на z. Этот оператор дает

$$\begin{aligned}\vartheta\,F\left({a_1,\ldots,a_m \atop b_1,\ldots,b_n}\,\middle|\,z\right) &= z\sum_{k\geqslant 1}\frac{a_1^{\overline{k}}\ldots a_m^{\overline{k}}\,z^{k-1}}{b_1^{\overline{k}}\ldots b_n^{\overline{k}}\,(k-1)!} = \\ &= \sum_{k\geqslant 0}\frac{k\,a_1^{\overline{k}}\ldots a_m^{\overline{k}}\,z^{k}}{b_1^{\overline{k}}\ldots b_n^{\overline{k}}\,k!},\end{aligned}$$

что само по себе не представляет особой ценности. Но если умножить F на один из верхних параметров, скажем, на a_1, и добавить

ϑF, то получится

$$(\vartheta + a_1)\, F\left({a_1, \dots, a_m \atop b_1, \dots, b_n} \middle| z\right) = \sum_{k \geqslant 0} \frac{(k + a_1) a_1^{\overline{k}} \dots a_m^{\overline{k}}\, z^k}{b_1^{\overline{k}} \dots b_n^{\overline{k}}\, k!} =$$
$$= \sum_{k \geqslant 0} \frac{a_1 (a_1{+}1)^{\overline{k}}\, a_2^{\overline{k}} \dots a_m^{\overline{k}}\, z^k}{b_1^{\overline{k}} \dots b_n^{\overline{k}}\, k!} =$$
$$= a_1\, F\left({a_1{+}1,\ a_2,\ \dots,\ a_m \atop b_1,\ \dots,\ b_n} \middle| z\right).$$

При этом на 1 увеличивается только один параметр.

Аналогичный трюк срабатывает и с нижними параметрами, но в этом случае они уменьшаются, а не увеличиваются:

$$(\vartheta{+}b_1{-}1)\, F\left({a_1, \dots, a_m \atop b_1, \dots, b_n} \middle| z\right) = \sum_{k \geqslant 0} \frac{(k{+}b_1{-}1) a_1^{\overline{k}} \dots a_m^{\overline{k}}\, z^k}{b_1^{\overline{k}} \dots b_n^{\overline{k}}\, k!} =$$
$$= \sum_{k \geqslant 0} \frac{(b_1{-}1)\, a_1^{\overline{k}} \dots a_m^{\overline{k}}\, z^k}{(b_1{-}1)^{\overline{k}}\, b_2^{\overline{k}} \dots b_n^{\overline{k}}\, k!} =$$
$$= (b_1{-}1)\, F\left({a_1,\ \dots,\ a_m \atop b_1{-}1,\ b_2,\ \dots,\ b_n} \middle| z\right).$$

Можно объединить все эти операции и выразить одну и ту же величину двумя разными способами, сыграв математическую “шутку”:

$$(\vartheta + a_1) \dots (\vartheta + a_m) F = a_1 \dots a_m\, F\left({a_1{+}1,\ \dots,\ a_m{+}1 \atop b_1,\ \dots,\ b_n} \middle| z\right),$$

и

$$(\vartheta + b_1 - 1) \dots (\vartheta + b_n - 1) F =$$
$$= (b_1{-}1) \dots (b_n{-}1)\, F\left({a_1,\ \dots,\ a_m \atop b_1{-}1,\ \dots,\ b_n{-}1} \middle| z\right),$$

где $F = F(a_1, \dots, a_m; b_1, \dots, b_n; z)$. Но (5.106) гласит, что верхняя строка является производной от нижней. Следовательно, обобщенная гипергеометрическая функция F удовлетворяет дифференциальному уравнению

$$D(\vartheta{+}b_1{-}1) \dots (\vartheta{+}b_n{-}1) F = (\vartheta{+}a_1) \dots (\vartheta{+}a_m) F, \qquad (5.107)$$

где D представляет собой оператор $\frac{d}{dz}$.

Это следует подтвердить примером. Давайте найдем дифференциальное уравнение, которому удовлетворяет стандартная гипергеометрическая функция с двумя верхними и одним нижним параметрами $F(z) = F(a, b; c; z)$. В соответствии с (5.107) имеем

$$D(\vartheta + c - 1)F = (\vartheta + a)(\vartheta + b)F.$$

Что это означает в обычной записи? Ну, $(\vartheta+c-1)F$ — это $zF'(z) + (c-1)F(z)$, а производная от этой величины дает левую часть,

$$F'(z) + zF''(z) + (c-1)F'(z).$$

В правой части имеем

$$\begin{aligned}(\vartheta+a)\bigl(zF'(z)+bF(z)\bigr) &= z\frac{d}{dz}\bigl(zF'(z)+bF(z)\bigr) + a\bigl(zF'(z)+bF(z)\bigr) = \\ &= zF'(z)+z^2F''(z)+bzF'(z)+azF'(z)+abF(z).\end{aligned}$$

Приравнивание обеих частей дает

$$z(1-z)F''(z) + \bigl(c - z(a+b+1)\bigr)F'(z) - abF(z) = 0. \qquad (5.108)$$

Это уравнение эквивалентно (5.107), записанному в форме произведения.

Обратно, от дифференциального уравнения можно вернуться к степенному ряду. Допустим, что $F(z) = \sum_{k\geqslant 0} t_k z^k$ — степенной ряд, удовлетворяющий уравнению (5.107). Непосредственное вычисление показывает, что

$$\frac{t_{k+1}}{t_k} = \frac{(k+a_1)\ldots(k+a_m)}{(k+b_1)\ldots(k+b_n)(k+1)};$$

следовательно, функцией $F(z)$ должна быть $t_0\,F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)$. Мы доказали, что гипергеометрический ряд (5.76) представляет собой единственный формальный степенной ряд, который удовлетворяет дифференциальному уравнению (5.107) и имеет постоянный член 1.

Было бы неплохо, если бы гипергеометрические функции позволяли решать все на свете дифференциальные уравнения, но это далеко не так. Правая часть уравнения (5.107) всегда разлагается на сумму членов вида $\alpha_k z^k F^{(k)}(z)$, где $F^{(k)}(z)$ — k-я производная $D^kF(z)$; левая часть всегда разлагается на сумму членов вида $\beta_k z^{k-1}F^{(k)}(z)$ при $k > 0$. Так что дифференциальное уравнение (5.107) всегда принимает специальный вид

$$z^{n-1}(\beta_n - z\alpha_n)F^{(n)}(z) + \cdots + (\beta_1 - z\alpha_1)F'(z) - \alpha_0 F(z) = 0.$$

Уравнение (5.108) иллюстрирует это в случае $n = 2$. И обратно, как будет показано в упр. 6.13, любое дифференциальное уравнение данного вида может быть разложено относительно ϑ-оператора так, чтобы получить уравнение типа (5.107). Таким образом, это дифференциальные уравнения, решениями которых являются степенные ряды с рациональными отношениями членов.

Функция $F(z) = (1-z)^r$ удовлетворяет $\vartheta F = z(\vartheta - r)F$. Это дает еще одно доказательство биномиальной теоремы.

Умножение обеих частей уравнения (5.107) на z позволяет обойтись без оператора D и дает примечательное выражение с ϑ во всех сомножителях:

$$\vartheta(\vartheta+b_1-1)\ldots(\vartheta+b_n-1)F = z(\vartheta+a_1)\ldots(\vartheta+a_m)F. \qquad (5.109)$$

Первый сомножитель $\vartheta = (\vartheta+1-1)$ слева соответствует сомножителю $(k+1)$ в отношении членов (5.81), который, в свою очередь, соответствует $k!$ в знаменателе k-го члена обобщенного гипергеометрического ряда. Остальные сомножители $(\vartheta + b_j - 1)$ соответствуют сомножителям $(k + b_j)$ в знаменателе, которые, в свою очередь, соответствуют $b_j^{\overline{k}}$ в (5.76). Справа же z соответствует z^k, а $(\vartheta + a_j)$ соответствует $a_j^{\overline{k}}$.

Одно из приложений данной дифференциальной теории состоит в поиске и доказательстве новых преобразований. Например, легко убедиться в том, что гипергеометрические функции

$$F\left({2a, 2b \atop a+b+\frac{1}{2}} \middle| z\right) \quad \text{и} \quad F\left({a, b \atop a+b+\frac{1}{2}} \middle| 4z(1-z)\right)$$

удовлетворяют дифференциальному уравнению

$$z(1-z)F''(z) + (a+b+\tfrac{1}{2})(1-2z)F'(z) - 4abF(z) = 0;$$

следовательно, должно выполняться *тождество Гаусса* [143, уравнение 102]

$$F\left({2a, 2b \atop a+b+\frac{1}{2}} \middle| z\right) = F\left({a, b \atop a+b+\frac{1}{2}} \middle| 4z(1-z)\right). \qquad (5.110)$$

Внимание: нельзя безопасно использовать (5.110) в случае $|z| > 1/2$, если только обе части соотношения не полиномы; см. упр. 53.

В частности,

$$F\left({2a, 2b \atop a+b+\frac{1}{2}} \middle| \frac{1}{2}\right) = F\left({a, b \atop a+b+\frac{1}{2}} \middle| 1\right) \qquad (5.111)$$

всякий раз, когда обе бесконечные суммы сходятся. И в самом деле, эти суммы всегда сходятся, за исключением вырожденного случая, когда $a + b + \frac{1}{2}$ представляет собой неположительное целое число.

Каждое новое тождество для гипергеометрических функций влечет за собой новое тождество для биномиальных коэффициентов, и последнее соотношение не является исключением. Рассмотрим сумму

$$\sum_{k\leqslant m}\binom{m-k}{n}\binom{m+n+1}{k}\left(\frac{-1}{2}\right)^k,\qquad \text{целые } m\geqslant n\geqslant 0.$$

Ее члены отличны от нуля при $0\leqslant k\leqslant m-n$, и, соблюдая, как и ранее, известную осторожность при переходе к пределу, эту сумму можно выразить в гипергеометрическом виде:

$$\lim_{\epsilon\to 0}\binom{m}{n}F\left({n-m,\ -n-m-1+\alpha\epsilon \atop -m+\epsilon}\,\bigg|\,\frac{1}{2}\right).$$

Величина α не влияет на предел, поскольку неположительный верхний параметр $n-m$ обрывает сумму раньше. Можно положить $\alpha=2$, так что становится применимым (5.111). Теперь данный предел можно вычислить, поскольку правая часть этого соотношения представляет собой частный случай (5.92). Результат может быть выражен в упрощенном виде

$$\sum_{k\leqslant m}\binom{m-k}{n}\binom{m+n+1}{k}\left(\frac{-1}{2}\right)^k=$$
$$=\binom{(m+n)/2}{n}2^{n-m}[m+n\text{ четное}],\qquad \text{целые } m\geqslant n\geqslant 0,\qquad (5.112)$$

как показано в упр. 54. Например, если $m=5$ и $n=2$, получаем $\binom{5}{2}\binom{8}{0}-\binom{4}{2}\binom{8}{1}/2+\binom{3}{2}\binom{8}{2}/4-\binom{2}{2}\binom{8}{3}/8=10-24+21-7=0$; если же $m=4$ и $n=2$, то обе части равны $\frac{3}{4}$.

Можно также указать случаи, когда соотношение (5.110) при $z=-1$ приводит к суммам биномиальных коэффициентов, но эти суммы, честно говоря, жутковаты. Если положить $a=\frac{1}{6}-\frac{n}{3}$ и $b=-n$, то получается монстроидальная формула

$$F\left({\frac{1}{3}-\frac{2}{3}n,\ -2n \atop \frac{2}{3}-\frac{4}{3}n}\,\bigg|\,-1\right)=F\left({\frac{1}{6}-\frac{1}{3}n,\ -n \atop \frac{2}{3}-\frac{4}{3}n}\,\bigg|\,-8\right).$$

Эти гипергеометрические функции являются невырожденными полиномами, если $n\not\equiv 2\pmod 3$, а их параметры подбираются так искусно, что левую часть выражения можно вычислить по формуле (5.94). Так мы приходим к умопомрачительному результату

$$\sum_k\binom{n}{k}\binom{\frac{1}{3}n-\frac{1}{6}}{k}8^k\bigg/\binom{\frac{4}{3}n-\frac{2}{3}}{k}=$$
$$=\binom{2n}{n}\bigg/\binom{\frac{4}{3}n-\frac{2}{3}}{n},\qquad \text{целое } n\geqslant 0,\ \ n\not\equiv 2\ (\mathrm{mod}\ 3).\qquad (5.113)$$

Это самое поразительное тождество с биномиальными коэффициентами, которое мы видели. Даже малые случаи такого тождества недостаточно просты для проверки вручную. (Оказывается, обе части в действительности дают $\frac{81}{7}$ при $n = 3$.) Это тождество совершенно бесполезно, поскольку наверняка не встретится ни в одной практической задаче.

Единственная польза (5.113) — демонстрация существования бесполезных тождеств.

Таково наше представление о гипергеометрических представлениях. Мы убедились в том, что гипергеометрические ряды обеспечивают нас высоконаучным способом обращения с суммами биномиальных коэффициентов. Масса дополнительной информации содержится в классической книге Бейли (Bailey) [18] и ее продолжении Гаспера (Gasper) и Рахмана (Rahman) [141].

5.7 Частичные гипергеометрические суммы

Большинство сумм в этой главе вычислялось по всему промежутку изменения индекса $k \geqslant 0$, но иногда нам удавалось найти аналитическое выражение, работающее и в случае интервала общего вида $a \leqslant k < b$. Например, из (5.16) известно, что

$$\sum_{k<m} \binom{n}{k} (-1)^k = (-1)^{m-1} \binom{n-1}{m-1}, \qquad \text{целое } m. \tag{5.114}$$

Теория из главы 2 обеспечивает нас замечательным способом трактовки подобных формул: если $f(k) = \Delta g(k) = g(k+1) - g(k)$, то мы условились писать, что $\sum f(k)\,\delta k = g(k) + C$, и

$$\sum\nolimits_a^b f(k)\,\delta k = g(k)\Big|_a^b = g(b) - g(a).$$

Кроме того, если a и b — целые числа и $a \leqslant b$, то

$$\sum\nolimits_a^b f(k)\,\delta k = \sum_{a \leqslant k < b} f(k) = g(b) - g(a).$$

Поэтому тождество (5.114) соответствует формуле вычисления неопределенных сумм

$$\sum \binom{n}{k} (-1)^k \,\delta k = (-1)^{k-1} \binom{n-1}{k-1} + C$$

и формуле вычисления разностей

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

Начав с функции $g(k)$, можно легко вычислить $\Delta g(k) = f(k)$ — функцию, суммой которой будет $g(k) + C$. Но гораздо труднее, начав с $f(k)$, найти ее неопределенную сумму

$\sum f(k)\,\delta k = g(k) + C$; функция g может вообще не выражаться в аналитическом виде. Например, похоже, не существует простой формы записи для $\sum \binom{n}{k}\,\delta k$; в противном случае мы могли бы вычислять суммы типа $\sum_{k\leqslant n/3}\binom{n}{k}$, перед которыми мы беспомощны. Но, может быть, все же существует простая формула для суммы $\sum \binom{n}{k}\,\delta k$, которую мы просто пока что не нашли? Как мы можем быть уверены в том, что ее не существует?

В 1977 году Р. У. Госпер (R. W. Gosper) [154] открыл замечательный способ вычисления неопределенных сумм $\sum f(k)\,\delta k = g(k) + C$, когда f и g принадлежат обширному классу функций, именуемых гипергеометрическими членами. Обозначим через

$$F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n}\,\middle|\, z\right)_k = \frac{a_1^{\overline{k}} \ldots a_m^{\overline{k}}}{b_1^{\overline{k}} \ldots b_n^{\overline{k}}}\,\frac{z^k}{k!} \qquad (5.115)$$

k-й член гипергеометрического ряда $F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)$. Будем рассматривать $F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)_k$ как функцию от k, а не от z. Во многих случаях оказывается, что существуют параметры c, A_1, …, A_M, B_1, …, B_N и Z, такие, что

$$\sum F\left({a_1, \ldots, a_m \atop b_1, \ldots, b_n}\,\middle|\, z\right)_k \delta k = c\,F\left({A_1, \ldots, A_M \atop B_1, \ldots, B_N}\,\middle|\, Z\right)_k + C \qquad (5.116)$$

при заданных a_1, …, a_m, b_1, …, b_n и z. Будем говорить, что заданная функция $F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)_k$ *суммируема в гипергеометрических членах*, если такие константы c, A_1, …, A_M, B_1, …, B_N, Z существуют. Алгоритм Госпера либо находит такие константы, либо доказывает, что они не существуют.

В общем случае мы говорим, что $t(k)$ представляет собой *гипергеометрический член*, если $t(k+1)/t(k)$ представляет собой рациональную функцию от k, не тождественную нулю. Это по сути означает, что $t(k)$ кратно с постоянным множителем члену вида (5.115). (Здесь возникают некоторые технические детали, связанные с нулями, поскольку мы бы хотели, чтобы $t(k)$ имело смысл при отрицательных k, а также когда одно или несколько значений b в (5.115) были нулевыми или отрицательными целыми числами. Строго говоря, наиболее общий гипергеометрический член получается путем умножения (5.115) на ненулевую константу и некоторую степень 0 с последующим сокращением нулей в числителе и знаменателе. Это общее правило поясняется примерами из упр. 12.)

Предположим, мы хотим найти $\sum t(k)\,\delta k$, где $t(k)$ — гипергеометрический член. Алгоритм Госпера состоит из двух шагов, каждый из которых достаточно прямолинеен. Шаг 1 состоит в том, чтобы выразить отношение членов в специальной форме

$$\frac{t(k+1)}{t(k)} = \frac{p(k+1)}{p(k)}\frac{q(k)}{r(k+1)}, \tag{5.117}$$

где p, q и r — полиномы, подчиняющиеся следующему условию:

Делимость полиномов аналогична делимости целых чисел. Например, $(k+\alpha)\backslash q(k)$ означает, что отношение $q(k)/(k+\alpha)$ представляет собой полином. Легко видеть, что $(k+\alpha)\backslash q(k)$ тогда и только тогда, когда $q(-\alpha)=0$.

$$(k+\alpha)\backslash q(k) \quad\text{и}\quad (k+\beta)\backslash r(k) \quad\Longrightarrow$$
$$\Longrightarrow\ \alpha-\beta \text{ не положительное целое число.} \tag{5.118}$$

Это условие легко достижимо: начнем с того, что предварительно положим $p(k)=1$, а $q(k)$ и $r(k+1)$ положим равными числителю и знаменателю отношения членов, и разложим их на линейные множители. Например, если $t(k)$ имеет вид (5.115), то мы начинаем с разложения $q(k) = (k+a_1)\ldots(k+a_m)z$ и $r(k) = (k+b_1-1)\ldots(k+b_n-1)k$. Затем мы проверяем, не нарушено ли условие (5.118). Если q и r имеют множители $(k+\alpha)$ и $(k+\beta)$, где $\alpha-\beta = N > 0$, мы выделяем их из q и r и заменяем $p(k)$ на

$$p(k)(k+\alpha-1)^{\underline{N-1}} = p(k)(k+\alpha-1)(k+\alpha-2)\ldots(k+\beta+1). \tag{5.119}$$

Новые p, q и r по-прежнему удовлетворяют (5.117), и этот процесс можно повторять до тех пор, пока не будет выполнено условие (5.118). Вскоре мы увидим, почему (5.118) так важно.

Шаг 2 алгоритма Госпера завершает работу — он состоит в поиске гипергеометрического члена $T(k)$, такого, что

$$t(k) = T(k+1) - T(k), \tag{5.120}$$

если это возможно. Однако совсем не очевидно, как это сделать. Перед тем как двигаться дальше, требуется разработать какую-то теорию. Госпер путем анализа множества частных случаев заметил, что разумно записать неизвестную функцию $T(k)$ в виде

$$T(k) = \frac{r(k)\,s(k)\,t(k)}{p(k)}, \tag{5.121}$$

где $s(k)$ — некая секретная функция, которую надо каким-то образом найти. Подставив (5.121) в (5.120) и применив (5.117), получаем

В упр. 55 поясняется, почему нам хотелось бы найти эту загадочную подстановку.

$$t(k) = \frac{r(k+1)s(k+1)t(k+1)}{p(k+1)} - \frac{r(k)s(k)t(k)}{p(k)} =$$
$$= \frac{q(k)s(k+1)t(k)}{p(k)} - \frac{r(k)s(k)t(k)}{p(k)};$$

так что должно выполняться

$$p(k) = q(k)s(k+1) - r(k)s(k). \qquad (5.122)$$

Если мы сможем найти функцию $s(k)$, удовлетворяющую этому фундаментальному рекуррентному соотношению, то мы найдем и $\sum t(k)\,\delta k$. Если же не сможем, то T не существует.

Мы предполагаем, что $T(k)$ — гипергеометрический член, а это означает, что $T(k+1)/T(k)$ представляет собой рациональную функцию от k. Таким образом, согласно (5.121) и (5.120), $r(k)s(k)/p(k) = T(k)/\bigl(T(k+1) - T(k)\bigr)$ — рациональная функция от k, а сама функция $s(k)$ должна быть отношением полиномов:

$$s(k) = f(k)/g(k).$$

Но на самом деле можно доказать, что функция $s(k)$ сама является полиномом. Действительно, если $g(k)$ — не константа и если $f(k)$ и $g(k)$ не имеют общих множителей, положим N равным наибольшему целому числу, такому, что и $(k+\beta)$, и $(k+\beta+N-1)$ входят как сомножители в $g(k)$ при некотором комплексном числе β. Величина N положительна, поскольку $N = 1$ всегда удовлетворяет данному условию. Уравнение (5.122) можно переписать как

$$p(k)g(k+1)g(k) = q(k)f(k+1)g(k) - r(k)g(k+1)f(k),$$

и если мы положим $k = -\beta$ и $k = -\beta - N$, то получим

$$r(-\beta)g(1-\beta)f(-\beta) = 0 = q(-\beta-N)f(1-\beta-N)g(-\beta-N).$$

Теперь $f(-\beta) \neq 0$ и $f(1-\beta-N) \neq 0$, потому что f и g не имеют общих корней. К тому же $g(1-\beta) \neq 0$ и $g(-\beta-N) \neq 0$, так как в противном случае функция $g(k)$ содержала бы множитель $(k+\beta-1)$ или $(k+\beta+N)$, что противоречит предположению о максимальности N. Поэтому

$$r(-\beta) = q(-\beta-N) = 0.$$

Но это противоречит условию (5.118). Следовательно, функция $s(k)$ должна быть полиномом.

Теперь понятно: Госперу потребовалось условие (5.118), чтобы это доказательство работало.

Наша задача в настоящее время — найти полином $s(k)$, удовлетворяющий (5.122), если $p(k)$, $q(k)$ и $r(k)$ — заданные полиномы, или доказать, что такой полином не существует. Это легко сделать, если $s(k)$ имеет какую-либо определенную степень d, поскольку можно записать

$$s(k) = \alpha_d k^d + \alpha_{d-1}k^{d-1} + \cdots + \alpha_0, \qquad \alpha_d \neq 0, \qquad (5.123)$$

с неизвестными коэффициентами $(\alpha_d, \ldots, \alpha_0)$ и подставить данное выражение в базовое рекуррентное соотношение (5.122). Полином $s(k)$ будет удовлетворять этому рекуррентному соотношению тогда и только тогда, когда все α удовлетворяют линейным уравнениям, получаемым в результате приравнивания коэффициентов при каждой степени k в (5.122).

Но как определить степень s? Оказывается, что в действительности имеется не более двух возможностей. Можно переписать (5.122) в виде

$$\begin{aligned} 2p(k) &= Q(k)\bigl(s(k+1)+s(k)\bigr)+R(k)\bigl(s(k+1)-s(k)\bigr), \qquad (5.124)\\ \text{где}\quad Q(k) &= q(k)-r(k) \quad\text{и}\quad R(k) = q(k)+r(k). \end{aligned}$$

Если $s(k)$ имеет степень d, то сумма $s(k+1)+s(k) = 2\alpha_d k^d + \cdots$ также имеет степень d, в то время как разность $s(k+1)-s(k) = \Delta s(k) = d\alpha_d k^{d-1} + \cdots$ имеет степень $d-1$. (Можно считать, что нулевой полином имеет степень -1.) Обозначим через $\deg(P)$ степень полинома P. Если $\deg(Q) \geqslant \deg(R)$, то степень правой части равенства (5.124) — $\deg(Q)+d$, так что мы должны получить $d = \deg(p)-\deg(Q)$. С другой стороны, если $\deg(Q) < \deg(R) = d'$, то можно записать $Q(k) = \lambda' k^{d'-1} + \cdots$ и $R(k) = \lambda k^{d'} + \cdots$, где $\lambda \neq 0$. Правая часть уравнения (5.124) имеет вид

$$(2\lambda'\alpha_d + \lambda d\,\alpha_d)k^{d+d'-1} + \cdots .$$

Итак, вот каковы две возможности: либо $2\lambda' + \lambda d \neq 0$ и $d = \deg(p)-\deg(R)+1$, либо $2\lambda'+\lambda d = 0$ и $d > \deg(p)-\deg(R)+1$. Второй случай требует проверки только тогда, когда $-2\lambda'/\lambda$ является целым числом d, бо́льшим, чем $\deg(p)-\deg(R)+1$.

Теперь у нас достаточно фактов для того, чтобы выполнить второй шаг алгоритма Госпера: испытав не более двух значений d, мы можем найти $s(k)$, если только уравнение (5.122) имеет полиномиальное решение. Если $s(k)$ существует, то можно подставить его в (5.121) и получить T. Если же нет, то тем самым доказано, что $t(k)$ не суммируется в гипергеометрических членах.

Пора рассмотреть конкретный пример: попробуем вычислить частичную сумму (5.114). Метод Госпера должен быть в состоянии вывести величину

$$\sum \binom{n}{k}(-1)^k\,\delta k$$

для любого фиксированного n, так что мы будем искать сумму

$$t(k) = \binom{n}{k}(-1)^k = \frac{n!\,(-1)^k}{k!\,(n-k)!}.$$

Шаг 1 состоит в представлении отношения членов в требуемом виде (5.117);

$$\frac{t(k+1)}{t(k)} = \frac{k-n}{k+1} = \frac{p(k+1)\,q(k)}{p(k)\,r(k+1)},$$

т.е. мы просто выбираем $p(k) = 1$, $q(k) = k - n$ и $r(k) = k$. Этот выбор p, q и r удовлетворяет (5.118), если только n не является отрицательным целым числом; предположим, что n таковым не является.

Почему не $r(k) = k+1$? А, понятно...

Теперь перейдем к шагу 2. Согласно (5.124) мы должны рассмотреть полиномы $Q(k) = -n$ и $R(k) = 2k - n$. Поскольку R имеет бо́льшую степень, чем Q, нам надо рассмотреть два случая. Либо $d = \deg(p) - \deg(R) + 1$, что равно 0, либо $d = -2\lambda'/\lambda$, где $\lambda' = -n$ и $\lambda = 2$; следовательно, $d = n$. Первый случай лучше, поскольку не требует, чтобы n было положительным целым числом, так что займемся сначала именно им. Нам надо будет рассмотреть вторую возможность для d, только если первый случай не приведет к успеху. В предположении, что $d = 0$, значение $s(k)$ равно просто α_0, и уравнение (5.122) сводится к

$$1 = (k-n)\alpha_0 - k\alpha_0\,.$$

Следовательно, мы выбираем $\alpha_0 = -1/n$. Это удовлетворяет требуемым условиям и дает точно тот результат, который мы рассчитывали подтвердить:

$$\begin{aligned} T(k) &= \frac{r(k)\,s(k)\,t(k)}{p(k)} = \\ &= k\cdot\left(\frac{-1}{n}\right)\cdot\binom{n}{k}(-1)^k = \\ &= \binom{n-1}{k-1}(-1)^{k-1}, \qquad \text{если } n \ne 0. \end{aligned}$$

Если применить тот же метод к вычислению неопределенной суммы $\sum\binom{n}{k}\,\delta k$, без $(-1)^k$, то все будет почти таким же, однако функция $q(k)$ будет равна $n - k$; следовательно, $Q(k) = n - 2k$ будет иметь степень, бо́льшую, чем степень $R(k) = n$, и мы придем к заключению, что d имеет невозможное значение $\deg(p) - \deg(Q) = -1$. (Полином $s(k)$ не может иметь отрицательную степень, так как не может быть нулевым.) Таким образом, функция $\binom{n}{k}$ не суммируется в гипергеометрических членах.

Однако раз уж мы избавились от невозможного, все оставшееся — каким бы невероятным оно ни казалось — должно быть истинным (как утверждал Ш. Холмс (S. Holmes) [83]). Когда мы определяли p, q и r на шаге 1, то решили игнорировать возможность того, что n может быть отрицательным целым числом. А что если это так? Давайте положим $n = -N$, где N положительно. Тогда отношение членов для $\sum \binom{n}{k} \delta k$ есть

$$\frac{t(k+1)}{t(k)} = \frac{-(k+N)}{(k+1)} = \frac{p(k+1)}{p(k)} \frac{q(k)}{r(k+1)}$$

и согласно (5.119) его следует представлять при помощи $p(k) = (k+1)^{\overline{N-1}}$, $q(k) = -1$, $r(k) = 1$. Шаг 2 алгоритма Госпера говорит нам о том, что мы должны поискать полином $s(k)$ степени $d = N - 1$ (а вдруг нам повезет?). Например, при $N = 2$ рекуррентное соотношение (5.122) говорит о том, что нам надо решить уравнение

$$k+1 = -\big((k+1)\alpha_1 + \alpha_0\big) - (k\alpha_1 + \alpha_0)\,.$$

Приравнивание коэффициентов при k и 1 показывает, что

$$1 = -\alpha_1 - \alpha_1; \qquad 1 = -\alpha_1 - \alpha_0 - \alpha_0;$$

следовательно, $s(k) = -\frac{1}{2}k - \frac{1}{4}$ является решением, а

$$T(k) = \frac{1 \cdot \left(-\frac{1}{2}k - \frac{1}{4}\right) \cdot \binom{-2}{k}}{k+1} = (-1)^{k-1}\frac{2k+1}{4}\,.$$

"Превосходно, Холмс!"
"Элементарно, мой дорогой Ватсон!"

Может ли это быть искомой суммой? Да, все получается:

$$(-1)^k\frac{2k+3}{4} - (-1)^{k-1}\frac{2k+1}{4} = (-1)^k(k+1) = \binom{-2}{k}\,.$$

Кстати, можно записать формулу суммирования и в другом виде, с верхним пределом:

$$\begin{aligned}\sum_{k<m}\binom{-2}{k} &= (-1)^{k-1}\frac{2k+1}{4}\bigg|_0^m = \\ &= \frac{(-1)^{m-1}}{2}\left(m + \frac{1-(-1)^m}{2}\right) = \\ &= (-1)^{m-1}\left\lceil\frac{m}{2}\right\rceil, \qquad \text{целое } m \geqslant 0.\end{aligned}$$

Такое представление скрывает тот факт, что биномиальный коэффициент $\binom{-2}{k}$ является суммируемым в гипергеометрических членах, так как $\lceil m/2 \rceil$ не является гипергеометрическим членом. (См. упр. 12.)

Со знаменателем (5.121) могут возникнуть проблемы, если $p(k) = 0$ для некоторого целого k. Упр. 97 помогает понять, что можно сделать в такой ситуации.

Заметим, что нет никакой необходимости собирать каталог гипергеометрических членов, суммируемых в неопределенном виде, аналогичный базе данных определенных гипергеометрических сумм, о которой мы говорили ранее в этой главе, поскольку алгоритм Госпера предоставляет быстрый регулярный метод, работающий во всех суммируемых случаях.

Марко Петковшек (Marko Petkovšek) [291] нашел красивый способ обобщения алгоритма Госпера на более сложные задачи обращения. Он показал, как найти все гипергеометрические члены $T(k)$, удовлетворяющие рекуррентному соотношению l-го порядка

$$t(k) = p_l(k)T(k+l)+\cdots+p_1(k)T(k+1)+p_0(k)T(k), \qquad (5.125)$$

если даны гипергеометрический член $t(k)$ и полиномы $p_l(k), \ldots, p_1(k), p_0(k)$.

5.8 Механическое суммирование

Алгоритм Госпера, замечательный сам по себе, находит аналитический вид только для некоторых биномиальных сумм из числа встречающихся на практике. Но мы не должны останавливаться на достигнутом. Дорон Зайльбергер (Doron Zeilberger) [384] показал, как расширить алгоритм Госпера так, чтобы он стал еще замечательнее и позволял справиться в еще большем количестве случаев. Расширение Зайльбергера позволяет не только находить частичные суммы, но и выполнять суммирование по всем k, так что у нас появляется альтернатива гипергеометрическим методам из разделов 5.5 и 5.6. Более того, как и в оригинальном методе Госпера, вычисления могут выполняться компьютером практически механически — не требуются ни рассуждения, ни удача.

Основная идея заключается в том, чтобы рассматривать наше слагаемое как функцию $t(n, k)$ от двух переменных n и k. (В алгоритме Госпера мы писали просто $t(k)$.) Если неопределенная сумма $t(n, k)$ по k не выражается в гипергеометрических членах — и давайте считать, что так оно и есть, исключение составляют лишь немногие суммы, — то, как заметил Зайльбергер, часто удается так видоизменить $t(n, k)$ (используя пионерскую идею 1940-х годов сестры Селины Фейсенмайер (Celine Fasenmyer) [382]), чтобы получить новые слагаемые, которые уже *суммируются* в неопределенном виде. Например, на практике зачастую $\beta_0(n)t(n,k) + \beta_1(n)t(n+1,k)$ может быть неопределенно просуммировано по k для подходящих полиномов $\beta_0(n)$ и $\beta_1(n)$. А вычислив сумму по k, мы получаем рекуррентное соотношение по n, которое и решает нашу задачу.

А страница 222 вырвана из книги.

Чтобы познакомиться с общим подходом, начнем с рассмотрения простого случая. Предположим, что мы забыли биномиальную теорему и хотим вычислить $\sum_k \binom{n}{k} z^k$. Как получить ответ, если мы не ясновидящие и никаких догадок у нас нет? Ранее в этой главе, например, в задаче 3 из раздела 5.2, мы научились заменять $\binom{n}{k}$ на $\binom{n-1}{k} + \binom{n-1}{k-1}$ и легко получать требуемый результат. Но есть и более систематический путь.

Пусть $t(n,k) = \binom{n}{k} z^k$ — величина, которую мы хотим просуммировать. Алгоритм Госпера говорит нам, что мы не можем вычислить частные суммы $\sum_{k\leqslant m} t(n,k)$ в гипергеометрических членах для произвольного n, за исключением случая $z = -1$. Так что давайте вместо этого рассмотрим более общее слагаемое

$$\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k). \tag{5.126}$$

Попробуем найти значения $\beta_0(n)$ и $\beta_1(n)$, при которых алгоритм Госпера успешно сработает. Во-первых, мы хотели бы упростить (5.126) при помощи соотношения между $t(n+1,k)$ и $t(n,k)$ для исключения $t(n+1,k)$ из выражения. Поскольку

$$\frac{t(n+1,k)}{t(n,k)} = \frac{(n+1)!\,z^k}{(n+1-k)!\,k!}\frac{(n-k)!\,k!}{n!\,z^k} = \frac{n+1}{n+1-k},$$

имеем

$$\hat{t}(n,k) = p(n,k)\frac{t(n,k)}{n+1-k},$$

где

$$p(n,k) = (n+1-k)\beta_0(n) + (n+1)\beta_1(n).$$

Теперь применим алгоритм Госпера к $\hat{t}(n,k)$ с фиксированным n. Сначала запишем

$$\frac{\hat{t}(n,k+1)}{\hat{t}(n,k)} = \frac{\hat{p}(n,k+1)}{\hat{p}(n,k)}\frac{q(n,k)}{r(n,k+1)}, \tag{5.127}$$

как в (5.117). Метод Госпера нашел бы такое представление, начав с $\hat{p}(n,k) = 1$, но в расширении Зайльбергера лучше начать с $\hat{p}(n,k) = p(n,k)$. Заметим, что если положить $\bar{t}(n,k) =$

$\hat{t}(n,k)/p(n,k)$ и $\bar{p}(n,k) = \hat{p}(n,k)/p(n,k)$, то уравнение (5.127) будет эквивалентно

$$\frac{\bar{t}(n,k+1)}{\bar{t}(n,k)} = \frac{\bar{p}(n,k+1)}{\bar{p}(n,k)}\frac{q(n,k)}{r(n,k+1)}. \tag{5.128}$$

Таким образом, мы можем найти $\hat{p}$, q и r, удовлетворяющие (5.127), если найдем $\bar{p}$, q и r, удовлетворяющие (5.128), начав с $\bar{p}(n,k) = 1$. Это упрощает жизнь, потому что $\bar{t}(n,k)$ не включает неизвестные величины $\beta_0(n)$ и $\beta_1(n)$, которые входят в $\hat{t}(n,k)$. В нашем случае $\bar{t}(n,k) = t(n,k)/(n+1-k) = n!\,z^k/(n+1-k)!\,k!$, так что

$$\frac{\bar{t}(n,k+1)}{\bar{t}(n,k)} = \frac{(n+1-k)z}{k+1}$$

и мы можем взять $q(n,k) = (n+1-k)z$ и $r(n,k) = k$. Предполагается, что эти полиномы от k удовлетворяют условию (5.118). Если же нет, то мы должны удалить из q и r некоторые множители и включить соответствующие множители (5.119) в $\bar{p}(n,k)$; но это надо делать, только когда величина $\alpha-\beta$ в (5.118) представляет собой положительную целую константу, не зависящую от n, поскольку мы хотим, чтобы наши вычисления были справедливы для произвольного n. (Формулы, которые мы выведем, в действительности будут справедливы, даже когда n и k не являются целыми (с применением обобщенных факториалов (5.83)).)

Теперь я уже помню, почему $r(n,k)$ не равно $k+1$.

Наш начальный выбор q и r удовлетворяет в указанном смысле условию (5.118), так что можно переходить прямо к шагу 2 алгоритма Госпера. Мы хотим решить уравнение, аналогичное (5.122), с использованием (5.127) вместо (5.117). Таким образом, мы хотим решить

$$\hat{p}(n,k) = q(n,k)s(n,k+1) - r(n,k)s(n,k) \tag{5.129}$$

относительно секретного полинома

$$s(n,k) = \alpha_d(n)k^d + \alpha_{d-1}(n)k^{d-1} + \cdots + \alpha_0(n). \tag{5.130}$$

(Коэффициенты s рассматриваются не как константы, а как функции от n.) В нашем случае уравнение (5.129) имеет вид

$$\begin{aligned}(n+1-k)\beta_0(n) + (n+1)\beta_1(n) &= \\ &= (n+1-k)zs(n,k+1) - ks(n,k),\end{aligned}$$

и мы рассматриваем его как полиномиальное уравнение по k с коэффициентами, которые представляют собой функции от n. Как

Функция $\deg(Q)$ *означает степень по* k; n *считается константой.*

и ранее, находим степень d полинома s, рассматривая $Q(n,k) = q(n,k) - r(n,k)$ и $R(n,k) = q(n,k) + r(n,k)$. Поскольку $\deg(Q) = \deg(R) = 1$ (в предположении $z \neq \pm 1$), имеем $d = \deg(\hat{p}) - \deg(Q) = 0$ и $s(n,k) = \alpha_0(n)$ не зависит от k. Наше уравнение превращается в

$$(n+1-k)\beta_0(n)+(n+1)\beta_1(n) \;=\; (n+1-k)z\alpha_0(n)-k\alpha_0(n);$$

приравнивая коэффициенты при равных степенях k, мы получаем эквивалентную систему уравнений, не содержащую k:

$$\begin{aligned}(n+1)\beta_0(n) + (n+1)\beta_1(n) - (n+1)z\alpha_0(n) &= 0,\\ -\beta_0(n) \qquad\qquad\qquad\qquad + (z+1)\alpha_0(n) &= 0.\end{aligned}$$

Следовательно, мы имеем решение (5.129) с

$$\beta_0(n) \;=\; z+1, \qquad \beta_1(n) \;=\; -1, \qquad \alpha_0(n) \;=\; s(n,k) \;=\; 1.$$

(По случайному совпадению n исчезло.)

Чисто механическим методом мы установили, что $\hat{t}(n,k) = (z+1)t(n,k) - t(n+1,k)$ суммируется в гипергеометрических членах. Другими словами,

$$\hat{t}(n,k) \;=\; T(n,k+1) - T(n,k), \tag{5.131}$$

где $T(n,k)$ представляет собой гипергеометрический член относительно k. Чему равно $T(n,k)$? В соответствии с (5.121) и (5.128) имеем

$$T(n,k) \;=\; \frac{r(n,k)s(n,k)\hat{t}(n,k)}{\hat{p}(n,k)} \;=\; r(n,k)s(n,k)\bar{t}(n,k), \tag{5.132}$$

поскольку $\bar{p}(n,k) = 1$. (На практике $\bar{p}(n,k)$ почти всегда оказывается равным 1.) Следовательно,

$$T(n,k) \;=\; \frac{k}{n+1-k}\,t(n,k) \;=\; \frac{k}{n+1-k}\binom{n}{k}z^k \;=\; \binom{n}{k-1}z^k.$$

Все получается — равенство (5.131) истинно:

$$(z+1)\binom{n}{k}z^k - \binom{n+1}{k}z^k \;=\; \binom{n}{k}z^{k+1} - \binom{n}{k-1}z^k.$$

Но в действительности нам не требуется точно знать $T(n,k)$, поскольку мы собираемся выполнять суммирование $t(n,k)$ по всем целым k. Все, что нам надо знать, — это то, что $T(n,k)$ не равно нулю только для конечного множества значений k, когда

n — произвольное заданное неотрицательное целое число. Тогда сумма $T(n,k+1) - T(n,k)$ по всем k должна телескопически свернуться в 0.

Пусть $S_n = \sum_k t(n,k) = \sum_k \binom{n}{k} z^k$; это сумма, с которой мы начинали и теперь готовы вычислить ее, поскольку уже много знаем о $t(n,k)$. При помощи процедуры Госпера–Зайльбергера установлено, что

$$\sum_k \bigl((z+1)t(n,k) - t(n+1,k)\bigr) = 0 .$$

Но эта сумма равна $(z+1)\sum_k t(n,k) - \sum_k t(n+1,k) = (z+1)S_n - S_{n+1}$. Таким образом, имеем

$$S_{n+1} = (z+1)S_n . \tag{5.133}$$

Ага! Мы знаем, как решать это рекуррентное соотношение, если нам известно S_0. Очевидно, однако, что $S_0 = 1$. Следовательно, мы заключаем, что $S_n = (z+1)^n$ для всех целых $n \geqslant 0$. QED

На самом деле $\lim_{k\to\infty} T(n,k) = 0$ при $|z| < 1$ и произвольном комплексном n. Поэтому (5.133) справедливо для всех n, и, в частности, $S_n = (z+1)^n$, когда n — отрицательное целое число.

Посмотрим еще раз на проделанные вычисления и подытожим наши действия в виде, применимом и для других слагаемых $t(n,k)$. Алгоритм Госпера–Зайльбергера (считая, что $t(n,k)$ задано) можно сформулировать следующим образом.

0 Положим $l := 0$. (Мы будем искать рекуррентное соотношение по n порядка l.)

1 Пусть $\hat{t}(n,k) = \beta_0(n)t(n,k) + \cdots + \beta_l(n)t(n+l,k)$, где $\beta_0(n)$, …, $\beta_l(n)$ — неизвестные функции. Используя свойства $t(n,k)$, найдем $p(n,k)$ — линейную комбинацию $\beta_0(n)$, …, $\beta_l(n)$, коэффициентами которой являются полиномы от n и k, такую, что $\hat{t}(n,k)$ может быть записана в виде $p(n,k)\bar{t}(n,k)$, где $\bar{t}(n,k)$ представляет собой гипергеометрический член от k. Найдем полиномы $\bar{p}(n,k)$, $q(n,k)$, $r(n,k)$, так чтобы отношение членов $\bar{t}(n,k)$ выражалось в виде (5.128) и при этом $q(n,k)$ и $r(n,k)$ удовлетворяли условию Госпера (5.118). Положим $\hat{p}(n,k) = p(n,k)\bar{p}(n,k)$.

2, а Положим $d_Q := \deg(q-r)$, $d_R := \deg(q+r)$ и

$$d := \begin{cases} \deg(\hat{p}) - d_Q, & \text{если } d_Q \geqslant d_R; \\ \deg(\hat{p}) - d_R + 1, & \text{если } d_Q < d_R. \end{cases}$$

2, б Если $d \geqslant 0$, то определим $s(n,k)$, как в (5.130), и рассмотрим линейные уравнения относительно α_0, …, α_d, β_0, … β_l, получаемые приравниванием коэффициентов при степенях k в основном уравнении (5.129). Если эти уравнения имеют решение, в котором все $\beta_0, \ldots, \beta_l$ не равны одновременно нулю, то переходим к шагу 4. В противном случае, если $d_Q < d_R$ и если $-2\lambda'/\lambda$ представляет собой целое число, большее d (где λ — коэффициент при k^{d_R} в $q+r$, а λ' — коэффициент при k^{d_R-1} в $q-r$), положим $d := -2\lambda'/\lambda$ и повторим шаг 2, б.

3 (Слагаемое $\hat{t}(n,k)$ не суммируется в гипергеометрических членах.) Увеличим l на 1 и вернемся к шагу 1.

4 (Успех.) Положим $T(n,k) := r(n,k)s(n,k)\bar{t}(n,k)/\bar{p}(n,k)$. Алгоритм нашел, что $\hat{t}(n,k) = T(n,k+1) - T(n,k)$.

Ниже мы докажем, что этот алгоритм успешно завершается для всех $t(n,k)$, принадлежащих большому классу выражений, именуемых подходящими членами.

Биномиальную теорему можно вывести многими способами, так что наш первый пример применения алгоритма Госпера–Зайльбергера в большей степени дидактичен, чем впечатляющ. Давайте теперь попробуем поработать со сверткой Вандермонда. Смогут ли Госпер и Зайльбергер алгоритмически вывести, что $\sum_k \binom{a}{k}\binom{b}{n-k}$ имеет простой вид? Алгоритм начинается с $l = 0$, что по сути представляет собой повторение исходного алгоритма Госпера. Мы пытаемся проверить, суммируется ли $\binom{a}{k}\binom{b}{n-k}$ в гипергеометрических членах. Небольшой сюрприз: это слагаемое суммируется, если $a+b$ — некоторое специальное неотрицательное целое число (см. упр. 94). Но нас интересуют произвольные a и b, и алгоритм быстро обнаруживает, что неопределенная сумма в общем случае гипергеометрическим членом не является. Так что l увеличивается с 0 до 1, и алгоритм теперь пробует $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k)$. На следующем шаге, как и при выводе биномиальной теоремы, запишем $\hat{t}(n,k) = p(n,k)\bar{t}(n,k)$, где $p(n,k)$ получается путем сокращения дробей в $t(n+1,k)/t(n,k)$. В этом случае — мы настоятельно рекомендуем читателю повторить все вычисления на листе бумаги (они не такие страшные, как кажется на первый взгляд) — все происходит почти так же, как и ранее, но теперь с

$$\begin{aligned}
p(n,k) &= (n+1-k)\beta_0(n)+(b-n+k)\beta_1(n) \;=\; \hat{p}(n,k)\,,\\
\bar{t}(n,k) &= t(n,k)/(n+1-k) \;=\; a!\,b!/(a-k)!\,k!\,(b-n+k)!\,(n+1-k)!\,,\\
q(n,k) &= (n+1-k)(a-k)\,,\\
r(n,k) &= (b-n+k)k\,.
\end{aligned}$$

Шаг 2, а находит, что $\deg(q-r) < \deg(q+r)$ и $d = \deg(\hat{p}) - \deg(q+r)+1 = 0$, так что $s(n,k)$ опять не зависит от k. Фундаментальное уравнение Госпера (5.129) эквивалентно системе двух уравнений от трех неизвестных

$$\begin{aligned}(n+1)\beta_0(n) + (b-n)\beta_1(n) \quad &- (n+1)a\alpha_0(n) = 0,\\ -\beta_0(n) \qquad\qquad + \beta_1(n) &+ (a+b+1)\alpha_0(n) = 0,\end{aligned}$$

которая имеет решение

$$\beta_0(n) = a+b-n, \qquad \beta_1(n) = -n-1, \qquad \alpha_0(n) = 1.$$

Принципиально важно то, что метод Госпера–Зайльбергера всегда приводит к линейным уравнениям относительно неизвестных α и β, так как левая часть (5.129) линейна относительно β, а правая часть — относительно α.

Мы заключаем, что $(a+b-n)t(n,k) - (n+1)t(n+1,k)$ суммируется по k; следовательно, если $S_n = \sum_k \binom{a}{k}\binom{b}{n-k}$, то имеет место рекуррентное соотношение

$$S_{n+1} = \frac{a+b-n}{n+1} S_n.$$

Таким образом, $S_n = \binom{a+b}{n}$, поскольку $S_0 = 1$. Замечательно!

А что можно сказать о тождестве Заальшютца (5.28) с тремя биномиальными коэффициентами? Доказательство (5.28) в упр. 43 интересно, но требует определенного озарения. Превращая искусство в науку, мы хотим заменить озарение освещением, т.е. методичной работой в поте лица. Давайте посмотрим, сможет ли метод Госпера–Зайльбергера обнаружить и доказать (5.28) чисто механическим путем. Для удобства выполним подстановки $m = b + d$, $n = a$, $r = a + b + c + d$, $s = a + b + c$, так что (5.28) принимает более симметричный вид

$$\sum_k \frac{(a+b+c+d+k)!}{(a-k)!\,(b-k)!\,(c+k)!\,(d+k)!\,k!} = \\ = \frac{(a+b+c+d)!\,(a+b+c)!\,(a+b+d)!}{a!\,b!\,(a+c)!\,(a+d)!\,(b+c)!\,(b+d)!}. \qquad (5.134)$$

Чтобы сумма была конечной, будем полагать, что либо a, либо b — неотрицательное целое число.

Положим $t(n,k) = (n+b+c+d+k)!/(n-k)!\,(b-k)!\,(c+k)!\,(d+k)!\,k!$ и $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k)$. Двигаясь по хорошо известному пути, положим

Единственная немеханическая часть работы — выбрать, какой параметр назвать n.

$$\begin{aligned} p(n,k) &= (n{+}1{-}k)\beta_0(n) + (n{+}1{+}b{+}c{+}d{+}k)\beta_1(n) = \hat{p}(n,k),\\ \bar{t}(n,k) &= \frac{t(n,k)}{n{+}1{-}k} = \frac{(n{+}b{+}c{+}d{+}k)!}{(n{+}1{-}k)!\,(b{-}k)!\,(c{+}k)!\,(d{+}k)!\,k!},\\ q(n,k) &= (n{+}b{+}c{+}d{+}k{+}1)(n{+}1{-}k)(b{-}k),\\ r(n,k) &= (c{+}k)(d{+}k)k \end{aligned}$$

и попытаемся решить (5.129) относительно $s(n,k)$. Снова оказывается, что $\deg(q-r) < \deg(q+r)$, но в этот раз $\deg(\hat{p})$ —

$\deg(q + r) + 1 = -1$, так что здесь мы, кажется, застопорились. Однако шаг 2,б предоставляет еще один важный выбор для степени s — $d = -2\lambda'/\lambda$; давайте испробуем его. Здесь $R(n, k) = q(n, k)+r(n, k) = 2k^3+\cdots$, так что $\lambda = 2$, в то время как полином $Q(n, k) = q(n, k)-r(n, k)$ каким-то чудом имеет степень 1 относительно k — коэффициенты при k^2 сокращаются! Таким образом, $\lambda' = 0$; метод Госпера позволяет взять $d = 0$ и $s(n, k) = \alpha_0(n)$.

Обратите внимание, что λ' не является старшим коэффициентом Q несмотря на то, что λ — старший коэффициент R. Число λ' представляет собой коэффициент при $k^{\deg(R)-1}$ в Q.

Уравнения, которые предстоит решить, имеют вид

$$\begin{aligned}
&(n+1)\beta_0(n)+(n+1+b+c+d)\beta_1(n)-\\
&\qquad -(n+1)(n+1+b+c+d)b\alpha_0(n) \;=\; 0,\\
&-\beta_0(n)+\beta_1(n)-\\
&\qquad -\big((n+1)b-(n+1+b)(n+1+b+c+d)-cd\big)\alpha_0(n) \;=\; 0;
\end{aligned}$$

и мы находим

$$\begin{aligned}
\beta_0(n) &= (n+1+b+c)(n+1+b+d)(n+1+b+c+d),\\
\beta_1(n) &= -(n+1)(n+1+c)(n+1+d),\\
\alpha_0(n) &= 2n+2+b+c+d,
\end{aligned}$$

пролив совсем немного пота. Отсюда немедленно вытекает тождество (5.134).

Пот течет, а тождество вытекает.

Аналогичное доказательство (5.134) можно получить при работе с $n = d$ вместо $n = a$. (См. упр. 99.)

Подход Госпера–Зайльбергера помогает вычислять не только неопределенные суммы по всем k, но и определенные суммы по ограниченному диапазону. Например, рассмотрим сумму

$$S_n(z) \;=\; \sum_{k=0}^{n}\binom{n+k}{k}z^k. \tag{5.135}$$

При $z = \frac{1}{2}$ мы получили “неожиданный” результат (5.20); может быть, для Госпера и Зайльбергера здесь нет ничего неожиданного? Обозначив $t(n, k) = \binom{n+k}{k}z^k$, получим

$$\begin{aligned}
p(n, k) &= (n+1)\beta_0(n) + (n+1+k)\beta_1(n) \;=\; \hat{p}(n, k),\\
\bar{t}(n, k) &= t(n, k)/(n+1) \;=\; (n+k)!\,z^k/k!\,(n+1)!,\\
q(n, k) &= (n+1+k)z,\\
r(n, k) &= k
\end{aligned}$$

и $\deg(s) = \deg(\hat{p}) - \deg(q - r) = 0$. Решением уравнения (5.129) будет $\beta_0(n) = 1$, $\beta_1(n) = z - 1$, $s(n, k) = 1$. Таким образом мы находим

$$t(n, k) + (z-1)t(n+1, k) \;=\; T(n, k+1) - T(n, k), \tag{5.136}$$

где $T(n,k) = r(n,k)s(n,k)\hat{t}(n,k)/\hat{p}(n,k) = \binom{n+k}{k-1}z^k$. Теперь можно просуммировать (5.136) для $0 \leqslant k \leqslant n+1$ и получить

$$\begin{aligned} S_n(z)+t(n,n+1)+(z-1)S_{n+1}(z) &= T(n,n+2)-T(n,0) = \\ &= \binom{2n+2}{n+1}z^{n+2} = \\ &= 2\binom{2n+1}{n}z^{n+2}. \end{aligned}$$

Но $t(n,n+1) = \binom{2n+1}{n+1}z^{n+1} = \binom{2n+1}{n}z^{n+1}$, так что

$$S_{n+1}(z) = \frac{1}{1-z}\left(S_n(z) + (1-2z)\binom{2n+1}{n}z^{n+1}\right). \qquad (5.137)$$

Сразу ясно, что случай $z = \frac{1}{2}$ — частный и что $S_{n+1}(\frac{1}{2}) = 2S_n(\frac{1}{2})$. Более того, можно упростить рекуррентное соотношение (5.137) путем умножения обеих частей на суммирующий множитель $(1-z)^{n+1}$; это приводит к общему тождеству

$$(1-z)^n \sum_{k=0}^{n}\binom{n+k}{k}z^k = 1+\frac{1-2z}{2-2z}\sum_{k=1}^{n}\binom{2k}{k}\bigl(z(1-z)\bigr)^k, \qquad (5.138)$$

которое мало кто мог ожидать до появления Госпера с Зайльбергером. Сейчас же получение таких тождеств — чисто рутинная работа.

А что можно сказать насчет аналогичной суммы

$$S_n(z) = \sum_{k=0}^{n}\binom{n-k}{k}z^k, \qquad (5.139)$$

которая встретилась нам в (5.74)? Полные уверенности, мы полагаем $t(n,k) = \binom{n-k}{k}z^k$ и вычисляем

$$\begin{aligned} p(n,k) &= (n+1-2k)\beta_0(n)+(n+1-k)\beta_1(n) = \hat{p}(n,k), \\ \bar{t}(n,k) &= t(n,k)/(n+1-2k) = (n-k)!\,z^k/k!\,(n+1-2k)!, \\ q(n,k) &= (n+1-2k)(n-2k)z, \\ r(n,k) &= (n+1-k)k. \end{aligned}$$

Но увы — уравнение (5.129) не решается, если $z \neq -\frac{1}{4}$, так как степень s должна быть равна $\deg(\hat{p}) - \deg(q-r) = -1$.

$S_n(-\frac{1}{4})$ *равно* $(n+1)/2^n$.

Никаких проблем. Мы просто добавляем еще один параметр $\beta_2(n)$ и испытываем $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k) + \beta_2(n)t(n+2,k)$:

$$\begin{aligned} p(n,k) = {} & (n+1-2k)(n+2-2k)\beta_0(n)+ \\ & +(n+1-k)(n+2-2k)\beta_1(n)+ \end{aligned}$$

$$
\begin{aligned}
&\qquad +(n+1-k)(n+2-k)\beta_2(n) \;=\; \hat{p}(n,k)\,,\\
\bar{t}(n,k) \;&=\; t(n,k)/(n+1-2k)(n+2-2k) \;=\; (n-k)!\,z^k/k!\,(n+2-2k)!\,,\\
q(n,k) \;&=\; (n+2-2k)(n+1-2k)z\,,\\
r(n,k) \;&=\; (n+1-k)k\,.
\end{aligned}
$$

Теперь можно взять $s(n,k) = \alpha_0(n)$, и (5.129) имеет решение

$$\beta_0(n) \;=\; z\,, \quad \beta_1(n) \;=\; 1\,, \quad \beta_2(n) \;=\; -1\,, \quad \alpha_0(n) \;=\; 1\,.$$

Мы обнаружили, что

$$zt(n,k) + t(n+1,k) - t(n+2,k) \;=\; T(n,k+1) - T(n,k)\,,$$

где

$$
\begin{aligned}
T(n,k) \;&=\; r(n,k)s(n,k)\hat{t}(n,k)/\hat{p}(n,k) \;=\\
&=\; (n+1-k)k\bar{t}(n,k) \;=\; \binom{n+1-k}{k-1}z^k.
\end{aligned}
$$

Суммирование от $k=0$ до $k=n$ дает

$$
\begin{aligned}
&zS_n(z) + \bigl(S_{n+1}(z) - \tbinom{0}{n+1}z^{n+1}\bigr) -\\
&\qquad - \left(S_{n+2}(z) - \binom{0}{n+2}z^{n+2} - \binom{1}{n+1}z^{n+1}\right) \;=\\
&\qquad = \; T(n,n+1) - T(n,0)\,.
\end{aligned}
$$

Но $\binom{1}{n+1}z^{n+1} = \binom{0}{n}z^{n+1} = T(n,n+1)$ для всех $n \geqslant 0$, так что мы получаем

$$S_{n+2}(z) \;=\; S_{n+1}(z) + zS_n(z)\,, \qquad n \geqslant 0. \tag{5.140}$$

Решения таких рекуррентных соотношений мы будем изучать в главах 6 и 7; методы этих глав позволяют немедленно перейти от (5.140) к аналитическому виду (5.74) при условии $S_0(z) = S_1(z) = 1$.

Еще один знаменитый пример завершает картину. В 1978 году французский математик Роже Апери (Roger Apéry) решил задачу, которая долго не поддавалась усилиям математиков, доказав, что число $\zeta(3) = 1 + 2^{-3} + 3^{-3} + 4^{-3} + \cdots$ иррационально [14]. Одна из важнейших частей его работы включает вычисление биномиальной суммы

$$A_n \;=\; \sum_k \binom{n}{k}^2 \binom{n+k}{k}^2. \tag{5.141}$$

Объявленное им рекуррентное соотношение для этой последовательности в то время другие математики проверить не могли. (А числа A_n с тех пор стали называться числами Апери; $A_0 = 1$, $A_1 = 5$, $A_2 = 73$, $A_3 = 1445$, $A_4 = 33001$.) Наконец, Дон Загиер (Don Zagier) и Генри Коэн (Henri Cohen) [356] нашли доказательство предположения Апери, и это их доказательство для данной частной (но сложной) суммы стало одним из ключевых моментов, которые привели Зайльбергера к открытию общего метода, который мы сейчас рассматриваем.

В действительности мы уже увидели достаточное количество примеров, чтобы сумма (5.141) стала почти тривиальной. Полагая $t(n,k) = \binom{n}{k}^2\binom{n+k}{k}^2$ и $\hat{t}(n,k) = \beta_0(n)t(n,k) + \beta_1(n)t(n+1,k) + \beta_2(n)t(n+2,k)$, попытаемся решить (5.129) с

Сначала мы попробовали обойтись без β_2, но эта попытка провалилась.

$$\begin{aligned}
p(n,k) &= (n{+}1{-}k)^2(n{+}2{-}k)^2\beta_0(n) \; = \\
&\quad +(n{+}1{+}k)^2(n{+}2{-}k)^2\beta_1(n) \; = \\
&\quad +(n{+}1{+}k)^2(n{+}2{+}k)^2\beta_2(n) \; = \; \hat{p}(n,k)\,, \\
\bar{t}(n,k) &= t(n,k)/(n{+}1{-}k)^2(n{+}2{-}k)^2 = (n{+}k)!^2/k!^4(n{+}2{-}k)!^2\,, \\
q(n,k) &= (n{+}1{+}k)^2(n{+}2{-}k)^2\,, \\
r(n,k) &= k^4\,.
\end{aligned}$$

(Можно не беспокоиться о том факте, что q имеет множитель $(k+n+1)$, а r — множитель k; это не нарушает (5.118), потому что мы рассматриваем n как переменную, а не как фиксированное целое число.) Поскольку $q(n,k)-r(n,k) = -2k^3+\cdots$, можно положить $\deg(s) = -2\lambda'/\lambda = 2$, так что получим

$$s(n,k) \;=\; \alpha_2(n)k^2 + \alpha_1(n)k + \alpha_0(n)\,.$$

При таком выборе s рекуррентное соотношение (5.129) превращается в пять уравнений относительно шести неизвестных $\beta_0(n)$, $\beta_1(n)$, $\beta_2(n)$, $\alpha_0(n)$, $\alpha_1(n)$, $\alpha_2(n)$. Например, приравнивание коэффициентов при k^0 приводит к уравнению, которое упрощается до

$$\beta_0 + \beta_1 + \beta_2 - \alpha_0 - \alpha_1 - \alpha_2 \;=\; 0\,;$$

уравнение, получающееся для коэффициентов при k^4, —

$$\beta_0 + \beta_1 + \beta_2 + \alpha_1 + (6+6n+2n^2)\alpha_2 \;=\; 0\,.$$

Три других уравнения более сложны. Но главное здесь то, что эти линейные уравнения — как все уравнения, появляющиеся на этой стадии алгоритма Госпера–Зайльбергера, *однородны* (их правые части равны нулю). Поэтому они всегда имеют ненулевое решение, если только число неизвестных превышает число уравнений. В нашем случае решением оказывается

$$\begin{aligned}
\beta_0(n) &= (n+1)^3\,,\\
\beta_1(n) &= -(2n+3)(17n^2+51n+39)\,,\\
\beta_2(n) &= (n+2)^3\,,\\
\alpha_0(n) &= -16(n+1)(n+2)(2n+3)\,,\\
\alpha_1(n) &= -12(2n+3)\,,\\
\alpha_2(n) &= 8(2n+3)\,.
\end{aligned}$$

Следовательно,

$$\begin{aligned}
(n+1)^3t(n,k) - (2n+3)(17n^2+51n+39)t(n+1,k)&+\\
+\,(n+2)^3t(n+2,k) &= T(n,k+1) - T(n,k)\,,
\end{aligned}$$

где $T(n,k) = k^4s(n,k)\bar{t}(n,k) = (2n+3)(8k^2-12k-16(n+1)(n+2))(n+k)!^2/(k-1)!^4(n+2-k)!^2$. Суммирование по k дает немыслимое рекуррентное соотношение Апери

“Профессор Литтлвуд (Littlewood), когда ему приходилось использовать какое-либо алгебраическое тождество, никогда не утруждал себя его доказательством; он утверждал, что алгебраическое тождество, если оно справедливо, может быть проверено в несколько строк любым сомневающимся, который достаточно туп, чтобы ощущать необходимость в такой проверке. Моя цель на протяжении следующих страниц — опровергнуть это утверждение”.
— Ф. Дайсон (F. J. Dyson) [89]

$$\begin{aligned}
(n+1)^3A_n + (n+2)^3A_{n+2} &=\\
&= (2n+3)(17n^2+51n+39)A_{n+1}\,. \qquad (5.142)
\end{aligned}$$

Но работает ли метод Госпера–Зайльбергера со всеми суммами, которые встречались нам в этой главе? Нет. Он непригоден в случае, когда $t(n,k)$ равно $\binom{n}{k}(k+1)^{k-1}(n-k+1)^{n-k-1}$ из (5.65), поскольку отношение членов $t(n,k+1)/t(n,k)$ не является рациональной функцией от k. Он также не в состоянии справиться со случаем наподобие $t(n,k) = \binom{n}{k}n^k$, поскольку другое отношение членов $t(n+1,k)/t(n,k)$ не является рациональной функцией от k. (Мы можем, однако, получить решение в этом случае, просуммировав $\binom{n}{k}z^k$ и затем положив $z = n$.) И наконец, этот метод терпит неудачу для удивительно простого слагаемого, наподобие $t(n,k) = 1/(nk+1)$, несмотря на то что оба отношения, $t(n,k+1)/t(n,k)$ и $t(n+1,k)/t(n,k)$, представляют собой рациональные функции от n и k.

Вместе с тем алгоритм Госпера–Зайльбергера гарантированно приводит к успеху в огромном числе случаев, а именно — когда слагаемое $t(n,k)$ является так называемым *подходящим членом*, т.е. членом, который может быть записан в виде

$$\begin{aligned}
t(n,k) &=\\
&= f(n,k)\frac{(a_1n+a_1'k+a_1'')!\,\ldots(a_pn+a_p'k+a_p'')!}{(b_1n+b_1'k+b_1'')!\,\ldots(b_qn+b_q'k+b_q'')!}\,w^nz^k\,. \qquad (5.143)
\end{aligned}$$

Здесь $f(n,k)$ — полиномы от n и k; коэффициенты a_1, a_1', …, a_p, a_p', b_1, b_1', …, b_q, b_q' — определенные целочисленные константы; параметры w и z ненулевые; остальные величины a_1'', …, a_p'', b_1'', …, b_q'' — произвольные комплексные числа. Мы докажем, что если $t(n,k)$ является подходящим членом, то найдутся такие полиномы $\beta_0(n)$, …, $\beta_l(n)$, не все равные нулю, и подходящий член $T(n,k)$, такой, что

Что случится, если $t(n,k)$ не зависит от n?

$$\beta_0(n)t(n,k)+\cdots+\beta_l(n)t(n+l,k) \;=\; T(n,k+1)-T(n,k)\,. \tag{5.144}$$

Приведенное далее доказательство принадлежит Вильфу и Зайльбергеру [374].

Обозначим через N оператор, который увеличивает n на 1, а через K — оператор, увеличивающий k на 1, так что, например, $N^2K^3t(n,k) = t(n+2,k+3)$. Мы будем изучать линейные разностные операторы от N, K и n, а именно — операторные полиномы вида

$$H(N,K,n) \;=\; \sum_{i=0}^{I}\sum_{j=0}^{J}\alpha_{i,j}(n)N^iK^j\,, \tag{5.145}$$

где каждый коэффициент $\alpha_{i,j}(n)$ представляет собой полином от n. Наше первое наблюдение состоит в том, что если $t(n,k)$ — произвольный подходящий член и $H(N,K,n)$ — произвольный линейный разностный оператор, то $H(N,K,n)t(n,k)$ является подходящим членом. Предположим, что t и H задаются соответственно равенствами (5.143) и (5.145); тогда мы определим "базовый член"

$$\bar t(n,k)_{I,J} \;=\; \frac{\prod_{i=1}^{p}\bigl(a_in+a_i'k+a_iI[a_i<0]+a_i'J[a_i'<0]+a_i''\bigr)!}{\prod_{i=1}^{q}\bigl(b_in+b_i'k+b_iI[b_i>0]+b_i'J[b_i'>0]+b_i''\bigr)!}\,w^nz^k\,.$$

Например, если $t(n,k)$ есть $\binom{n-2k}{k} = (n-2k)!/k!\,(n-3k)!$, то базовым членом, соответствующим линейному разностному оператору степеней I и J, будет $\bar t(n,k)_{I,J} = (n-2k-2J)!/(k+J)!\,(n-3k+I)!$. Главное то, что $\alpha_{i,j}(n)N^iK^jt(n,k)$ равно $\bar t(n,k)_{I,J}$, умноженному на некоторый полином от n и k, если $0 \leqslant i \leqslant I$ и $0 \leqslant j \leqslant J$. Конечные суммы полиномов представляют собой полиномы, так что $H(N,K,n)t(n,k)$ имеет требуемый вид (5.143).

На следующем шаге требуется показать, что если $t(n,k)$ — подходящий член, то всегда найдется такой линейный разностный оператор $H(N,K,n)$, что

$$H(N,K,n)t(n,k) \;=\; 0\,.$$

Если $0 \leqslant i \leqslant I$ и $0 \leqslant j \leqslant J$, то сдвинутый член $N^i K^j t(n,k)$ есть произведение $\bar{t}(n,k)_{I,J}$ на полином от n и k, степень которого по k не превышает

$$\begin{aligned} D_{I,J} &= \deg(f) + |a_1|I + |a_1'|J + \cdots + |a_p|I + |a_p'|J \\ &\qquad + |b_1|I + |b_1'|J + \cdots + |b_q|I + |b_q'|J. \end{aligned}$$

Следовательно, требуемый оператор H существует, если мы сможем решить систему из $D_{I,J}+1$ однородных линейных уравнений с $(I+1)(J+1)$ переменными $\alpha_{i,j}(n)$, коэффициентами которых являются полиномы от n. Все, что нам надо, — это выбрать I и J достаточно большими, чтобы выполнялось неравенство $(I+1)(J+1) > D_{I,J}+1$. Например, можно взять $I = 2A'+1$ и $J = 2A + \deg(f)$, где

$$\begin{aligned} A &= |a_1| + \cdots + |a_p| + |b_1| + \cdots + |b_q|; \\ A' &= |a_1'| + \cdots + |a_p'| + |b_1'| + \cdots + |b_q'|. \end{aligned}$$

Последний шаг доказательства заключается в переходе от уравнения $H(N,K,n)t(n,k) = 0$ к решению (5.144). Пусть H выбран таким образом, чтобы минимизировать J, т.е. так, что H имеет наименьшую возможную степень по K. Можно записать для некоторого линейного разностного оператора $G(N,K,n)$

Примененный здесь трюк основан на представлении H в виде полинома от K с последующей заменой K на $\Delta+1$.

$$H(N,K,n) = H(N,1,n) - (K-1)G(N,K,n).$$

Пусть $H(N,1,n) = \beta_0(n) + \beta_1(n)N + \cdots + \beta_l(n)N^l$ и $T(n,k) = G(N,K,n)t(n,k)$. Тогда $T(n,k)$ является подходящим членом и (5.144) выполняется.

Доказательство почти завершено. Нам осталось убедиться, что оператор $H(N,1,n)$ — не просто нулевой. Если это так, то $T(n,k)$ не зависит от k. Поэтому найдутся такие полиномы $\beta_0(n)$ и $\beta_1(n)$, что $\bigl(\beta_0(n) + \beta_1(n)N\bigr)T(n,k) = 0$. Но тогда $\bigl(\beta_0(n) + \beta_1(n)N\bigr)G(N,K,n)$ — ненулевой разностный оператор степени $J-1$, который аннулирует $t(n,k)$; это противоречит предположению о минимальности J, и наше доказательство (5.144) на этом завершено.

Теперь, зная, что (5.144) выполняется для некоторого подходящего члена T, мы можем быть уверены, что алгоритм Госпера успешно найдет T (или T плюс константу). Хотя мы доказали корректность алгоритма Госпера только для случая гипергеометрических членов $t(k)$, зависящих от одной переменной k, наше доказательство может быть распространено на случай двух переменных следующим образом. Существует бесконечно много комплексных чисел n, для которых выполняется условие (5.118) при разложении $q(n,k)$ и $r(n,k)$ как полиномов от k и для которых вычисление степени d на шаге 2 дает тот же результат, что и в алгоритме Госпера для одной переменной. Для всех таких n наше предшествующее доказательство устанавливает существование подходящего полинома $s(n,k)$ от k; таким образом, существует и требуемый полином $s(n,k)$ от n и k. QED

Мы доказали, что алгоритм Госпера–Зайльбергера сможет найти решение (5.144) для некоторого (как можно меньшего) l. Это решение дает нам рекуррентное соотношение по n для вычисления суммы по k любого подходящего члена $t(n,k)$ при условии, что $t(n,k)$ отлично от нуля только для конечного множества значений k. И, разумеется, можно поменять ролями n и k, поскольку определение (5.143) подходящего члена симметрично относительно n и k.

Упр. 98–108 дают еще несколько дополнительных примеров использования алгоритма Госпера–Зайльбергера, иллюстрируя его разносторонность. Вильф и Зайльбергер [374] существенно расширили эти результаты и получили метод, который справляется с обобщенными биномиальными коэффициентами и кратными суммами.

Упражнения

Разминка

1 Чему равно 11^4? Почему это число легко вычислить тому, кто знаком с биномиальными коэффициентами?

2 При каком значении (значениях) k величина $\binom{n}{k}$ максимальна, если n — заданное положительное целое число? Обоснуйте свой ответ.

3 Докажите “свойство шестиугольника”

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

4 Вычислите $\binom{-1}{k}$, обратив (т.е. сделав положительным) верхний индекс.

5 Пусть p — простое число. Покажите, что $\binom{p}{k} \bmod p = 0$ при $0 < k < p$. Что следует отсюда относительно биномиальных коэффициентов $\binom{p-1}{k}$?

6 Исправьте решение задачи 6 в разделе 5.2, правильно применив симметрию.

Случай ошибочности тождества.

7 Справедлива ли формула (5.34) при $k < 0$?

8 Вычислите

$$\sum_k \binom{n}{k}(-1)^k(1-k/n)^n .$$

Чему приближенно равно значение этой суммы при очень больших значениях n? *Указание*: сумма равна $\Delta^n f(0)$ для некоторой функции f.

9 Покажите, что обобщенные экспоненциальные ряды из (5.58) подчиняются правилу

$$\mathcal{E}_t(z) = \mathcal{E}(tz)^{1/t}, \qquad \text{если } t \neq 0,$$

где $\mathcal{E}(z)$ представляет собой сокращение для $\mathcal{E}_1(z)$.

10 Покажите, что $-2\big(\ln(1-z)+z\big)/z^2$ является гипергеометрической функцией.

11 Выразите функции

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots \qquad \text{и}$$

$$\arcsin z = z + \frac{1\cdot z^3}{2\cdot 3} + \frac{1\cdot 3\cdot z^5}{2\cdot 4\cdot 5} + \frac{1\cdot 3\cdot 5\cdot z^7}{2\cdot 4\cdot 6\cdot 7} + \cdots$$

через гипергеометрические ряды.

12 Какая из приведенных далее функций от k является гипергеометрическим членом в смысле раздела 5.7? Обоснуйте свой ответ.

а n^k.
б k^n.
в $\big(k!+(k+1)!\big)/2$.
г H_k, т.е. $\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{k}$.
д $1/\binom{n}{k}$.
е $t(k)T(k)$, где t и T — гипергеометрические члены.
ж $t(k)+T(k)$, где t и T — гипергеометрические члены.
з $t(n-k)$, где t — гипергеометрический член.
и $a\,t(k)+b\,t(k+1)+c\,t(k+2)$, где t — гипергеометрический член.
й $\lceil k/2 \rceil$.
к $k\,[k>0]$.

Здесь t и T не обязательно связаны между собой, как в (5.120).

Обязательные упражнения

13 Найдите связь между суперфакториальной функцией $P_n = \prod_{k=1}^{n} k!$ из упр. 4.55, гиперфакториальной функцией $Q_n = \prod_{k=1}^{n} k^k$ и произведением $R_n = \prod_{k=0}^{n} \binom{n}{k}$.

14 Докажите тождество (5.25), обратив верхний индекс в правиле свертки Вандермонда (5.22). Затем покажите, что еще одно обращение приводит к (5.26).

15 Чему равна сумма $\sum_k \binom{n}{k}^3(-1)^k$? *Указание*: см. (5.29).

16 Вычислите сумму

$$\sum_k \binom{2a}{a+k}\binom{2b}{b+k}\binom{2c}{c+k}(-1)^k$$

при неотрицательных целых a, b и c.

17 Найдите простое соотношение между $\binom{2n-1/2}{n}$ и $\binom{2n-1/2}{2n}$.

18 Найдите альтернативный вид выражения, аналогичный (5.35), для произведения

$$\binom{r}{k}\binom{r-1/3}{k}\binom{r-2/3}{k}.$$

19 Покажите, что обобщенные биномиальные ряды из (5.58) подчиняются правилу

$$\mathcal{B}_t(z) = \mathcal{B}_{1-t}(-z)^{-1}.$$

20 Определим "обобщенный инфрагеометрический ряд" формулой

$$G\left(\begin{matrix}a_1, \ldots, a_m\\ b_1, \ldots, b_n\end{matrix}\middle| z\right) = \sum_{k\geqslant 0} \frac{a_1^{\underline{k}} \ldots a_m^{\underline{k}}}{b_1^{\underline{k}} \ldots b_n^{\underline{k}}} \frac{z^k}{k!},$$

используя в определении (5.76) убывающие степени вместо возрастающих. Объясните, как ряд G связан с рядом F.

21 Покажите, что определение факториалов Эйлера согласуется с обычным определением, показав, что предел в определении (5.83) равен $1/m!$, если $z = m$ — положительное целое число.

22 Воспользуйтесь определением (5.83) для доказательства *факториальной формулы удвоения*:

$$x!\,(x-\tfrac{1}{2})! = (2x)!\,(-\tfrac{1}{2})!/2^{2x}.$$

Кстати, $(-\frac{1}{2})! = \sqrt{\pi}$.

23 Чему равно значение $F(-n, 1; ; 1)$?

24 Найдите величину $\sum_k \binom{n}{m+k}\binom{m+k}{2k}4^k$, воспользовавшись гипергеометрическим рядом.

25 Покажите, что

$$(a_1 - b_1)\,F\left({a_1,\, a_2,\, \ldots,\, a_m \atop b_1{+}1,\, b_2,\, \ldots,\, b_n}\,\middle|\, z\right) =$$

$$= a_1\,F\left({a_1{+}1,\, a_2,\, \ldots,\, a_m \atop b_1{+}1,\, b_2,\, \ldots,\, b_n}\,\middle|\, z\right) -$$

$$- b_1\,F\left({a_1,\, a_2,\, \ldots,\, a_m \atop b_1,\, b_2,\, \ldots,\, b_n}\,\middle|\, z\right).$$

Найдите подобную зависимость между гипергеометрическими функциями

$$F\left({a_1,\, a_2,\, a_3,\, \ldots,\, a_m \atop b_1,\, \ldots,\, b_n}\,\middle|\, z\right),$$

$$F\left({a_1{+}1,\, a_2,\, a_3,\, \ldots,\, a_m \atop b_1,\, \ldots,\, b_n}\,\middle|\, z\right) \qquad \text{и}$$

$$F\left({a_1,\, a_2{+}1,\, a_3,\, \ldots,\, a_m \atop b_1,\, \ldots,\, b_n}\,\middle|\, z\right).$$

26 Выразите функцию $G(z)$ из формулы

$$F\left({a_1,\, \ldots,\, a_m \atop b_1,\, \ldots,\, b_n}\,\middle|\, z\right) = 1 + G(z)$$

в виде кратного некоторому гипергеометрическому ряду.

27 Докажите, что

$$F\left({a_1,\, a_1{+}\frac{1}{2},\, \ldots,\, a_m,\, a_m{+}\frac{1}{2} \atop b_1,\, b_1{+}\frac{1}{2},\, \ldots,\, b_n,\, b_n{+}\frac{1}{2},\, \frac{1}{2}}\,\middle|\, (2^{m-n-1}z)^2\right) =$$

$$= \frac{1}{2}\left(F\left({2a_1,\ldots,2a_m \atop 2b_1,\ldots,2b_n}\,\middle|\, z\right) + F\left({2a_1,\ldots,2a_m \atop 2b_1,\ldots,2b_n}\,\middle|\, -z\right)\right).$$

28 Докажите *тождество Эйлера*

$$F\left({a,\, b \atop c}\,\middle|\, z\right) = (1-z)^{c-a-b}\,F\left({c{-}a,\, c{-}b \atop c}\,\middle|\, z\right)$$

путем двойного применения закона Пфаффа (5.101).

29 Покажите, что вырожденные гипергеометрические функции удовлетворяют соотношению

$$e^z\,F\left({a \atop b}\,\middle|\, -z\right) = F\left({b{-}a \atop b}\,\middle|\, z\right).$$

30 Какой гипергеометрический ряд F удовлетворяет уравнению $zF'(z) + F(z) = 1/(1-z)$?

31 Покажите, что если $f(k)$ — любая функция, суммируемая в гипергеометрических членах, то f сама по себе является гипергеометрическим членом. Например, если $\sum f(k)\,\delta k = cF(A_1, \ldots, A_M; B_1, \ldots, B_N; Z)_k + C$, то существуют константы $a_1, \ldots, a_m, b_1, \ldots, b_n$ и z, такие, что $f(k)$ кратно (5.115).

32 Найдите величину $\sum k^2\,\delta k$ методом Госпера.

33 Воспользуйтесь методом Госпера, чтобы найти $\sum \delta k/(k^2-1)$.

34 Покажите, что частичная гипергеометрическая сумма всегда может быть представлена в виде предела обычных гипергеометрических функций:

$$\sum_{k\leqslant c} F\left(\begin{matrix} a_1, \ldots, a_m \\ b_1, \ldots, b_n \end{matrix}\middle| z\right)_k = \lim_{\epsilon\to 0} F\left(\begin{matrix} -c, a_1, \ldots, a_m \\ \epsilon-c, b_1, \ldots, b_n \end{matrix}\middle| z\right),$$

если c — неотрицательное целое число. (См. (5.115).) Воспользуйтесь этой идеей для вычисления $\sum_{k\leqslant m}\binom{n}{k}(-1)^k$.

Домашние задания

35 При отсутствии контекста запись $\sum_{k\leqslant n}\binom{n}{k}2^{k-n}$ неоднозначна. Вычислите ее

а как сумму по k;

б как сумму по n.

36 Пусть p^k — наибольшая степень простого числа p, которая делит $\binom{m+n}{m}$, если m и n — неотрицательные целые числа. Докажите, что k представляет собой количество переносов, которые происходят при сложении m с n в системе счисления по основанию p. *Указание*: вам поможет упр. 4.24.

37 Покажите, что для факториальных степеней справедлив аналог биномиальной теоремы, т.е. докажите тождества

$$(x+y)^{\underline{n}} = \sum_k \binom{n}{k} x^{\underline{k}}\, y^{\underline{n-k}} \qquad \text{и}$$

$$(x+y)^{\overline{n}} = \sum_k \binom{n}{k} x^{\overline{k}}\, y^{\overline{n-k}}$$

для всех неотрицательных целых чисел n.

38 Покажите, что все неотрицательные целые числа n могут быть представлены единственным образом в виде $n = \binom{a}{1} + \binom{b}{2} + \binom{c}{3}$, где a, b и c — целые числа $0 \leqslant a < b < c$. (Это называется *комбинаторной системой счисления*.)

39 Покажите, что если $xy = ax + by$, то

$$x^n y^n = \sum_{k=1}^{n} \binom{2n-1-k}{n-1} (a^n b^{n-k} x^k + a^{n-k} b^n y^k)$$

при всех $n > 0$. Найдите подобную формулу для более общего произведения $x^m y^n$. (Эти формулы полезны при нахождении представления в простых дробях, например при $x = 1/(z-c)$ и $y = 1/(z-d)$.)

40 Запишите сумму

$$\sum_{j=1}^{m} (-1)^{j+1} \binom{r}{j} \sum_{k=1}^{n} \binom{-j+rk+s}{m-j}, \qquad \text{целые } m, n \geqslant 0,$$

в аналитическом виде.

41 Вычислите $\sum_k \binom{n}{k} k!/(n+1+k)!$ при целом неотрицательном n.

42 Найдите неопределенную сумму $\sum \big((-1)^x / \binom{n}{x}\big)\,\delta x$ и воспользуйтесь ею для записи суммы $\sum_{k=0}^{n} (-1)^k / \binom{n}{k}$ в аналитическом виде.

43 Докажите трехчленное биномиальное тождество (5.28). *Указание*: сначала замените $\binom{r+k}{m+n}$ на $\sum_j \binom{r}{m+n-j} \binom{k}{j}$.

44 Воспользуйтесь тождеством (5.32) для записи в аналитическом виде двойных сумм

$$\sum_{j,k} (-1)^{j+k} \binom{j+k}{j} \binom{a}{j} \binom{b}{k} \binom{m+n-j-k}{m-j} \quad \text{и}$$

$$\sum_{j,k \geqslant 0} (-1)^{j+k} \binom{a}{j} \binom{m}{j} \binom{b}{k} \binom{n}{k} \Big/ \binom{m+n}{j+k},$$

при заданных целых $m \geqslant a \geqslant 0$ и $n \geqslant b \geqslant 0$.

45 Запишите в аналитическом виде сумму $\sum_{k \leqslant n} \binom{2k}{k} 4^{-k}$.

46 Выразите в аналитическом виде при положительном целом n сумму

$$\sum_k \binom{2k-1}{k} \binom{4n-2k-1}{2n-k} \frac{(-1)^{k-1}}{(2k-1)(4n-2k-1)}.$$

Указание: вам помогут производящие функции.

47 Сумма

$$\sum_k \binom{rk+s}{k}\binom{rn-rk-s}{n-k}$$

представляет собой полином от r и s. Покажите, что она не зависит от s.

48 Тождество $\sum_{k=0}^{n}\binom{n+k}{n}2^{-k} = 2^n$ может быть объединено с формулой $\sum_{k\geqslant 0}\binom{n+k}{n}z^k = 1/(1-z)^{n+1}$, что приводит к

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

Какой вид имеет гипергеометрическая форма последнего соотношения?

49 Воспользуйтесь методом гипергеометрических функций для вычисления

$$\sum_k (-1)^k \binom{x}{k}\binom{x+n-k}{n-k}\frac{y}{y+n-k}\,.$$

50 Докажите закон отражения Пфаффа (5.101), сравнивая коэффициенты при z^n в обеих частях данного равенства.

51 Вывод формулы (5.104) показывает, что

$$\lim_{\epsilon\to 0} F(-m, -2m-1+\epsilon; -2m+\epsilon; 2) = 1\Big/\binom{-1/2}{m}\,.$$

В этом упражнении мы увидим, что немного отличающиеся предельные переходы приводят к существенно разным ответам для вырожденного гипергеометрического ряда $F(-m, -2m-1; -2m; 2)$.

а Покажите, что $\lim_{\epsilon\to 0} F(-m+\epsilon, -2m-1; -2m+2\epsilon; 2) = 0$, воспользовавшись законом отражения Пфаффа для доказательства тождества $F(a, -2m-1; 2a; 2) = 0$ при всех целых $m \geqslant 0$.

б Чему равен $\lim_{\epsilon\to 0} F(-m+\epsilon, -2m-1; -2m+\epsilon; 2)$?

52 Докажите, что если N — неотрицательное целое число, то

$$b_1^{\overline{N}}\ldots b_n^{\overline{N}}\, F\left(\left.{a_1,\ldots,a_m,-N \atop b_1,\ldots,b_n}\right| z\right) =$$
$$= a_1^{\overline{N}}\ldots a_m^{\overline{N}}(-z)^N\, F\left(\left.{1-b_1-N,\ldots,1-b_n-N,-N \atop 1-a_1-N,\ldots,1-a_m-N}\right| \frac{(-1)^{m+n}}{z}\right).$$

53 Если в тождестве Гаусса (5.110) положить $b = -\frac{1}{2}$ и $z = 1$, то левая часть сведется к -1, в то время как правая часть равна $+1$. Почему это не доказывает, что $-1 = +1$?

54 Объясните, как была получена правая часть соотношения (5.112).

55 Покажите, что если гипергеометрические члены $t(k) = F(a_1, \ldots, a_m; b_1, \ldots, b_n; z)_k$ и $T(k) = F(A_1, \ldots, A_M; B_1, \ldots, B_N; Z)_k$ удовлетворяют соотношению $t(k) = c\big(T(k+1) - T(k)\big)$ при всех $k \geqslant 0$, то $z = Z$ и $m - n = M - N$.

56 Используя метод Госпера, найдите общую формулу для $\sum \binom{-3}{k}\,\delta k$. Покажите, что $(-1)^{k-1}\lfloor \frac{k+1}{2} \rfloor \lfloor \frac{k+2}{2} \rfloor$ также является решением.

57 Используя метод Госпера при заданных n и z, найдите константу θ, такую, что

$$\sum \binom{n}{k} z^k (k+\theta)\,\delta k$$

суммируется в гипергеометрических членах.

58 Если m и n — целые числа, такие, что $0 \leqslant m \leqslant n$, то положим

$$T_{m,n} = \sum_{0 \leqslant k < n} \binom{k}{m} \frac{1}{n-k}\,.$$

Найдите соотношение между $T_{m,n}$ и $T_{m-1,n-1}$, а затем решите полученное рекуррентное соотношение, применив суммирующий множитель.

Контрольные работы

59 Найдите аналитический вид

$$\sum_{k \geqslant 1} \binom{n}{\lfloor \log_m k \rfloor}$$

при положительных целых m и n.

60 Воспользуйтесь для вычисления $\binom{m+n}{n}$ при одновременно больших m и n приближением Стирлинга (4.23). К чему сводится ваша формула при $m = n$?

61 Докажите, что если p — простое число, то

$$\binom{n}{m} \equiv \binom{\lfloor n/p \rfloor}{\lfloor m/p \rfloor} \binom{n \bmod p}{m \bmod p} \pmod{p}$$

при любых неотрицательных целых числах m и n.

62 В предположении, что p — простое число, а m и n — положительные целые числа, определите величину $\binom{np}{mp} \bmod p^2$. *Указание*: при желании можете воспользоваться следующим обобщением свертки Вандермонда:

$$\sum_{k_1+k_2+\cdots+k_m=n} \binom{r_1}{k_1}\binom{r_2}{k_2}\ldots\binom{r_m}{k_m} = \binom{r_1+r_2+\cdots+r_m}{n}.$$

63 Найдите аналитический вид суммы

$$\sum_{k=0}^{n} (-4)^k \binom{n+k}{2k}$$

при заданном целом $n \geqslant 0$.

64 Вычислите $\sum_{k=0}^{n} \binom{n}{k} / \lceil \frac{k+1}{2} \rceil$ при заданном целом $n \geqslant 0$.

65 Докажите, что

$$\sum_k \binom{n-1}{k} n^{-k}(k+1)! = n.$$

66 Вычислите “двойную сумму Гарри”

$$\sum_{0\leqslant j\leqslant k} \binom{-1}{j - \lfloor\sqrt{k-j}\rfloor}\binom{j}{m}\frac{1}{2^j}, \qquad \text{целое } m \geqslant 0,$$

как функцию от m. (Это сумма как по j, так и по k.)

67 Запишите в аналитическом виде

$$\sum_{k=0}^{n} \binom{\binom{k}{2}}{2}\binom{2n-k}{n}, \qquad \text{целое } n \geqslant 0.$$

68 Запишите в аналитическом виде

$$\sum_k \binom{n}{k} \min(k, n-k), \qquad \text{целое } n \geqslant 0.$$

69 Запишите в аналитическом виде

$$\min_{\substack{k_1,\ldots,k_m \geqslant 0 \\ k_1+\cdots+k_m=n}} \sum_{j=1}^{m} \binom{k_j}{2}$$

как функцию от m и n.

70 Запишите в аналитическом виде

$$\sum_k \binom{n}{k}\binom{2k}{k}\left(\frac{-1}{2}\right)^k, \qquad \text{целое } n \geqslant 0.$$

71 Пусть

$$S_n = \sum_{k\geqslant 0}\binom{n+k}{m+2k}a_k\,,$$

где m и n — неотрицательные целые числа, а $A(z) = \sum_{k\geqslant 0} a_k z^k$ — производящая функция для последовательности $\langle a_0, a_1, a_2, \ldots\rangle$.

а Выразите производящую функцию $S(z) = \sum_{n\geqslant 0} S_n z^n$ через $A(z)$.

б Воспользуйтесь этим приемом для решения задачи 7 из раздела 5.2.

72 Докажите, что если m, n, и k — целые числа и $n > 0$, то

$$\binom{m/n}{k} n^{2k-\nu(k)} \quad \text{целое},$$

где $\nu(k)$ — количество единиц в бинарном представлении k.

73 Воспользуйтесь методом наборов для решения рекуррентного соотношения

$$\begin{aligned} X_0 &= \alpha; \qquad X_1 = \beta;\\ X_n &= (n-1)(X_{n-1}+X_{n-2}) \qquad \text{при } n > 1. \end{aligned}$$

Указание: этому рекуррентному соотношению удовлетворяют как $n!$, так и $n¡$.

74 Эта задача связана с "нестандартным вариантом" треугольника Паскаля, стороны которого состоят из чисел 1, 2, 3, 4, ..., а не из единиц, но при этом числа внутри треугольника по-прежнему удовлетворяют формуле сложения:

$$\begin{array}{ccccccccccc} &&&&&1&&&&& \\ &&&&2&&2&&&& \\ &&&3&&4&&3&&& \\ &&4&&7&&7&&4&& \\ &5&&11&&14&&11&&5& \\ \cdot&&\cdot&&\cdot&&\cdot&&\cdot&&\cdot \end{array}$$

Если $\left(\!\binom{n}{k}\!\right)$ обозначает k-е число в n-й строке при $1 \leqslant k \leqslant n$, то $\left(\!\binom{n}{1}\!\right) = \left(\!\binom{n}{n}\!\right) = n$ и $\left(\!\binom{n}{k}\!\right) = \left(\!\binom{n-1}{k}\!\right) + \left(\!\binom{n-1}{k-1}\!\right)$ при $1 < k < n$. Выразите величину $\left(\!\binom{n}{k}\!\right)$ в аналитическом виде.

75 Найдите соотношение между функциями

$$S_0(n) = \sum_k \binom{n}{3k},$$

$$S_1(n) = \sum_k \binom{n}{3k+1},$$

$$S_2(n) = \sum_k \binom{n}{3k+2}$$

и величинами $\lfloor 2^n/3 \rfloor$ и $\lceil 2^n/3 \rceil$.

76 Решите следующее рекуррентное соотношение для $n, k \geqslant 0$:

$$Q_{n,0} = 1; \qquad Q_{0,k} = [k=0];$$

$$Q_{n,k} = Q_{n-1,k} + Q_{n-1,k-1} + \binom{n}{k} \qquad \text{при } n, k > 0.$$

77 Чему равно значение

$$\sum_{0 \leqslant k_1, \ldots, k_m \leqslant n} \ \prod_{1 \leqslant j < m} \binom{k_{j+1}}{k_j} \qquad \text{при } m > 1?$$

78 Выразите в аналитическом виде

$$\sum_{k=0}^{2m^2} \binom{k \bmod m}{(2k+1) \bmod (2m+1)}$$

в предположении, что m — положительное целое число.

79 **а** Чему равен наибольший общий делитель чисел $\binom{2n}{1}$, $\binom{2n}{3}$, …, $\binom{2n}{2n-1}$? *Указание*: рассмотрите сумму всех этих n чисел.

б Покажите, что наименьшее общее кратное чисел $\binom{n}{0}$, $\binom{n}{1}$, …, $\binom{n}{n}$ равно $L(n+1)/(n+1)$, где $L(n) = \text{НОК}(1, 2, \ldots, n)$.

80 Докажите, что $\binom{n}{k} \leqslant (en/k)^k$ для всех целых $k, n \geqslant 0$.

Это стоит запомнить.

81 Докажите неравенство

$$(-1)^{n-m-1} \sum_k \binom{l}{k} \binom{m+\theta}{n+k} x^k > 0$$

при $0 < \theta < 1$ и $0 \leqslant x \leqslant 1$ и если l, m, n — неотрицательные целые числа, такие, что $m < n$. *Указание*: попробуйте взять производную по x.

Дополнительные задачи

82 Докажите, что треугольник Паскаля обладает при $0 < k < n$ еще более удивительным свойством шестиугольника, чем упоминавшееся в тексте главы:

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

Например, $\text{НОД}(56, 36, 210) = \text{НОД}(28, 120, 126) = 2$.

83 Докажите удивительное соотношение с пятипараметрической двойной суммой (5.32).

84 Покажите, что вторая пара формул свертки (5.61) следует из первой пары (5.60). *Указание*: продифференцируйте по z.

85 Докажите, что

$$\sum_{m=1}^{n}(-1)^m \sum_{1\leqslant k_1<k_2<\cdot<k_m\leqslant n} \binom{k_1^3+k_2^3+\cdots+k_m^3+2^n}{n} = \\ = (-1)^n n!^3 - \binom{2^n}{n}.$$

(Левая часть представляет собой сумму $2^n - 1$ членов.) *Указание*: на самом деле справедливо гораздо большее.

86 Пусть $a_1, \ldots, a_n$ — неотрицательные целые числа и $C(a_1, \ldots, a_n)$ — коэффициент при постоянном члене $z_1^0 \ldots z_n^0$, когда $n(n-1)$ множителей произведения

$$\prod_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}} \left(1 - \frac{z_i}{z_j}\right)^{a_i}$$

полностью разложены на положительные и отрицательные степени комплексных переменных $z_1, \ldots, z_n$.

а Докажите, что $C(a_1, \ldots, a_n)$ равно левой части (5.31).

б Докажите, что если $z_1, \ldots, z_n$ — различные комплексные числа, то полином

$$f(z) = \sum_{k=1}^{n} \prod_{\substack{1\leqslant j\leqslant n\\ j\neq k}} \frac{z - z_j}{z_k - z_j}$$

тождественно равен 1.

в Умножьте исходное произведение $n(n-1)$ сомножителей на $f(0)$ и выведите, что $C(a_1, a_2, \ldots, a_n)$ равно

$$C(a_1 - 1, a_2, \ldots, a_n) + C(a_1, a_2 - 1, \ldots, a_n) + \\ + \cdots + C(a_1, a_2, \ldots, a_n - 1).$$

(Это рекуррентное соотношение определяет мультиномиальные коэффициенты, так что $C(a_1,\ldots,a_n)$ должно совпадать с правой частью (5.31).)

87 Пусть m — положительное целое число и пусть $\zeta = e^{\pi i/m}$. Покажите, что

$$\sum_{k\leqslant n/m}\binom{n-mk}{k}z^{mk} =$$
$$= \frac{\mathcal{B}_{-m}(z^m)^{n+1}}{(1+m)\mathcal{B}_{-m}(z^m)-m} -$$
$$- \sum_{0\leqslant j<m}\frac{\bigl(\zeta^{2j+1}z\,\mathcal{B}_{1+1/m}(\zeta^{2j+1}z)^{1/m}\bigr)^{n+1}}{(m+1)\mathcal{B}_{1+1/m}(\zeta^{2j+1}z)^{-1}-1}\,.$$

(В частном случае $m = 1$ это сводится к (5.74).)

88 Докажите, что коэффициенты s_k в (5.47) равны

$$(-1)^k\int_0^\infty e^{-t}(1-e^{-t})^{k-1}\,\frac{dt}{t}$$

при всех $k > 1$; следовательно, $|s_k| < 1/(k-1)$.

89 Докажите, что соотношение (5.19) имеет бесконечный аналог

$$\sum_{k>m}\binom{m+r}{k}x^k y^{m-k} = \sum_{k>m}\binom{-r}{k}(-x)^k(x+y)^{m-k}, \quad \text{целое } m,$$

если $|x| < |y|$ и $|x| < |x+y|$. Продифференцируйте это соотношение n раз по y и выразите в гипергеометрическом виде. Какое соотношение вы при этом получите?

90 В задаче 1 раздела 5.2 рассматривается сумма $\sum_{k\geqslant 0}\binom{r}{k}/\binom{s}{k}$ при целых r и s, таких, что $s \geqslant r \geqslant 0$. Чему равно значение этой суммы при нецелых r и s?

91 Докажите *тождество Уиппла* (*Whipple*),

$$F\left({\tfrac12 a,\ \tfrac12 a+\tfrac12,\ 1+a-b-c \atop 1+a-b,\ 1+a-c}\,\Big|\,\frac{-4z}{(1-z)^2}\right) =$$
$$= (1-z)^a\,F\left({a,\ b,\ c \atop 1+a-b,\ 1+a-c}\,\Big|\,z\right),$$

показав, что обе его части удовлетворяют одному и тому же дифференциальному уравнению.

92 Докажите *тождества произведений Клаузена* (*Clausen*)

$$F\left({a,\,b \atop a+b+\frac12}\,\Big|\,z\right)^2 \;=\; F\left({2a,\,a+b,\,2b \atop 2a+2b,\,a+b+\frac12}\,\Big|\,z\right);$$

$$F\left({\frac14+a,\,\frac14+b \atop 1+a+b}\,\Big|\,z\right) F\left({\frac14-a,\,\frac14-b \atop 1-a-b}\,\Big|\,z\right)$$

$$= F\left({\frac12,\,\frac12+a-b,\,\frac12-a+b \atop 1+a+b,\,1-a-b}\,\Big|\,z\right).$$

Какие тождества получатся, если приравнять коэффициенты при z^n в обеих частях этих формул?

93 Покажите, что при произвольной заданной функции f и произвольной константе $\alpha \neq 0$ бесконечная сумма

$$\sum \left(\prod_{j=1}^{k-1} \bigl(f(j)+\alpha\bigr) \Big/ \prod_{j=1}^{k} f(j) \right) \delta k$$

имеет (достаточно) простой вид.

94 Найдите $\sum \binom{a}{k}\binom{-a}{n-k}\,\delta k$ при положительном целом числе n.

95 Какие условия должны быть наложены в дополнение к (5.118), чтобы полиномы p, q, r из (5.117) стали однозначно определенными?

96 Докажите, что если алгоритм Госпера не находит решения (5.120) при данном гипергеометрическом члене $t(k)$, то не имеет решения и более общее уравнение

$$t(k) \;=\; \bigl(T_1(k{+}1)+\cdots+T_m(k{+}1)\bigr)-\bigl(T_1(k)+\cdots+T_m(k)\bigr)\,,$$

где $T_1(k), \ldots, T_m(k)$ — гипергеометрические члены.

97 Найдите все комплексные числа z, для которых $k!^2\big/\prod_{j=1}^{k}(j^2+jz+1)$ суммируется в гипергеометрических членах.

98 Какое рекуррентное соотношение дает метод Госпера–Зайльбергера для суммы $S_n = \sum_k \binom{n}{2k}$?

99 Воспользуйтесь методом Госпера–Зайльбергера для поиска аналитической записи $\sum_k t(n,k)$, где $t(n,k) = (n+a+b+c+k)!/(n+k)!\,(c+k)!\,(b-k)!\,(a-k)!\,k!$, в предположении, что a — неотрицательное целое число.

100 Найдите рекуррентное соотношение для суммы

$$S_n = \sum_{k=0}^{n} \frac{1}{\binom{n}{k}}$$

и воспользуйтесь им для того, чтобы найти другую формулу для S_n.

101 Найдите рекуррентное соотношение, которому удовлетворяют суммы

Для этого и нескольких последующих упражнений имеет смысл воспользоваться компьютерной алгеброй.

а $$S_{m,n}(z) = \sum_k \binom{m}{k}\binom{n}{k} z^k\,;$$

б $$S_n(z) = S_{n,n}(z) = \sum_k \binom{n}{k}^2 z^k\,.$$

102 Используйте процедуру Госпера–Зайльбергера для обобщения "бесполезного" тождества (5.113): найдите еще какие-нибудь значения a, b и z, такие, что

$$\sum_k \binom{n}{k}\binom{\frac{1}{3}n-a}{k} z^k \Big/ \binom{\frac{4}{3}n-b}{k}$$

имеет простую аналитическую запись.

103 Пусть $t(n,k)$ — подходящий член (5.143). Какие степени будут иметь полиномы $\hat{p}(n,k)$, $q(n,k)$ и $r(n,k)$ относительно переменной k, когда процедура Госпера–Зайльбергера применяется к $\hat{t}(n,k) = \beta_0(n)t(n,k) + \cdots + \beta_l(n)t(n+l,k)$? (Редкие, исключительные случаи можно игнорировать.)

104 Воспользуйтесь процедурой Госпера–Зайльбергера для проверки замечательного тождества

$$\sum_k (-1)^k \binom{r-s-k}{k}\binom{r-2k}{n-k}\frac{1}{r-n-k+1} = \binom{s}{n}\frac{1}{r-2n+1}\,.$$

Поясните, почему для этой суммы не найдено простейшее рекуррентное соотношение.

105 Покажите, что если $\omega = e^{2\pi i/3}$, то

$$\sum_{k+l+m=3n} \binom{3n}{k,l,m}^2 \omega^{l-m} = \binom{4n}{n,n,2n}, \quad \text{целое } n \geqslant 0.$$

106 Докажите удивительное тождество (5.32), положив $t(r,j,k)$ равным результату деления слагаемого на правую часть и показав затем, что найдутся функции $T(r,j,k)$ и $U(r,j,k)$, для которых

$$t(r+1,j,k)-t(r,j,k) = T(r,j+1,k)-T(r,j,k)+ \\ +U(r,j,k+1)-U(r,j,k)\,.$$

107 Докажите, что $1/(nk+1)$ не является подходящим членом.

108 Покажите, что числа Апери A_n из (5.141) представляют собой диагональные элементы $A_{n,n}$ числовой матрицы, определенной следующим образом:

$$A_{m,n} = \sum_{j,k}\binom{m}{j}^2\binom{m}{k}^2\binom{2m+n-j-k}{2m}.$$

Докажите, что эта матрица симметрична и что

$$A_{m,n} = \sum_{k}\binom{m+n-k}{k}^2\binom{m+n-2k}{m-k}^2 = \\ = \sum_{k}\binom{m}{k}\binom{n}{k}\binom{m+k}{k}\binom{n+k}{k}.$$

109 Докажите, что числа Апери (5.141) удовлетворяют сравнению

$$A_n \equiv A_{\lfloor n/p\rfloor}A_{n \bmod p} \pmod p$$

для всех простых p и всех целых $n \geqslant 0$.

Исследовательские проблемы

110 При каких n справедливо сравнение $\binom{2n}{n} \equiv (-1)^n \pmod{(2n+1)}$?

111 Пусть $q(n)$ — наименьший нечетный простой множитель среднего биномиального коэффициента $\binom{2n}{n}$. Согласно упр. 36 нечетными простыми числами p, которые *не* делят $\binom{2n}{n}$, являются те, все цифры которых в представлении n в системе счисления с основанием p не превышают $(p-1)/2$. Вычислительные эксперименты показали, что $q(n) \leqslant 11$ для $1 < n < 10^{10000}$, за исключением $q(3160) = 13$.

а Верно ли, что $q(n) \leqslant 11$ для всех $n > 3160$?

б Достигается ли $q(n) = 11$ для бесконечного количества n?

За решение любой из частей данного упражнения предлагается вознаграждение в $7 \cdot 11 \cdot 13$ долларов.

112 Верно ли, что $\binom{2n}{n}$ делится на 4 или 9 при всех $n > 4$ за исключением $n = 64$ и $n = 256$?

113 Если $t(n+1,k)/t(n,k)$ и $t(n,k+1)/t(n,k)$ — рациональные функции от n и k и если существует линейный разностный оператор $H(N,K,n)$, такой, что $H(N,K,n)t(n,k) = 0$, то следует ли отсюда, что $t(n,k)$ — подходящий член?

114 Пусть m — положительное целое число; определим последовательность $c_n^{(m)}$ при помощи рекуррентного соотношения

$$\sum_k \binom{n}{k}^m \binom{n+k}{k}^m = \sum_k \binom{n}{k}\binom{n+k}{k} c_k^{(m)} .$$

Являются ли числа $c_n^{(m)}$ целыми?

6

Специальные числа

НЕКОТОРЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ чисел возникают в математике так часто, что узнаются с первого взгляда и получают собственные имена. Например, каждый, кто изучал арифметику, знаком с последовательностью квадратов $\langle 1, 4, 9, 16, \ldots \rangle$. В главе 1 мы сталкивались с треугольными числами $\langle 1, 3, 6, 10, \ldots \rangle$; в главе 4 изучали простые числа $\langle 2, 3, 5, 7, \ldots \rangle$; в главе 5 бегло рассмотрели числа Каталана $\langle 1, 2, 5, 14, \ldots \rangle$.

В настоящей главе мы познакомимся с другими важными последовательностями. Первым пунктом повестки дня будут числа Стирлинга ${n \brace k}$ и ${n \brack k}$ и числа Эйлера $\left\langle {n \atop k} \right\rangle$. Они образуют треугольники коэффициентов аналогично биномиальным коэффициентам $\binom{n}{k}$ из треугольника Паскаля. Затем мы тщательно изучим гармонические числа H_n и бросим беглый взгляд на числа Бернулли B_n; они отличаются от других ранее рассматривавшихся последовательностей чисел тем, что являются дробными, а не целыми числами. Наконец, мы исследуем очаровывающие числа Фибоначчи F_n и некоторые из их важных обобщений.

6.1 Числа Стирлинга

Начнем с достаточно близких родственников биномиальных коэффициентов — чисел Стирлинга, которые названы так в честь Джеймса Стирлинга (James Stirling) (1692–1770). Эти числа имеют две разновидности, традиционно именуемые "числами Стирлинга первого и второго рода". Несмотря на долгую историю и многочисленные применения для этих чисел до сих пор нет общепринятого обозначения. Следуя Йовану Карамате (Jovan Karamata), мы будем записывать числа Стирлинга второго рода как ${n \brace k}$, а числа Стирлинга первого рода — как ${n \brack k}$; эти обозначения являются более дружественными и удобными по сравнению со многими другими предлагавшимися обозначениями.

"...par cette notation, les formules deviennent plus symétriques."
— *Й. Карамата [199]*

Таблица 318. Треугольник Стирлинга для числа подмножеств

n	${n \brace 0}$	${n \brace 1}$	${n \brace 2}$	${n \brace 3}$	${n \brace 4}$	${n \brace 5}$	${n \brace 6}$	${n \brace 7}$	${n \brace 8}$	${n \brace 9}$
0	1									
1	0	1								
2	0	1	1							
3	0	1	3	1						
4	0	1	7	6	1					
5	0	1	15	25	10	1				
6	0	1	31	90	65	15	1			
7	0	1	63	301	350	140	21	1		
8	0	1	127	966	1701	1050	266	28	1	
9	0	1	255	3025	7770	6951	2646	462	36	1

В табл. 318 и 319 показано, как выглядят числа ${n \brace k}$ и ${n \brack k}$ при малых n и k. Задача, которая включает в себя числа “6, 11, 6, 1”, скорее всего, должна быть связана с числами ${n \brack k}$, так же как задача, в которой встречаются числа “1, 4, 6, 4, 1”, скорее всего, имеет отношение к $\binom{n}{k}$; это “фирменные знаки” последовательностей при $n = 4$.

Числа Стирлинга второго рода встречаются гораздо чаще, чем числа Стирлинга первого рода, так что первыми будут рассмотрены числа второго рода. Обозначение ${n \brace k}$ используется для количества способов разбиения множества из n элементов на k непустых подмножеств. Например, имеется семь способов разбиения четырехэлементного множества на две части:

Кстати, сам Стирлинг тоже рассматривал их первыми в своей книге [343].

$$\begin{aligned}&\{1,2,3\}\cup\{4\},\quad \{1,2,4\}\cup\{3\},\quad \{1,3,4\}\cup\{2\},\quad \{2,3,4\}\cup\{1\},\\ &\{1,2\}\cup\{3,4\},\quad \{1,3\}\cup\{2,4\},\quad \{1,4\}\cup\{2,3\};\end{aligned} \tag{6.1}$$

таким образом, ${4 \brace 2} = 7$. Обратите внимание, что фигурные скобки используются для обозначения как множеств, так и чисел ${n \brace k}$. Это сходство помогает запомнить смысл записи ${n \brace k}$, которую можно прочесть как “k подмножеств из n”.

Давайте рассмотрим малые значения k. Имеется только один способ разместить n элементов в единственном непустом множестве; следовательно, ${n \brace 1} = 1$ для всех $n > 0$. С другой стороны, ${0 \brace 1} = 0$, потому что 0-элементное множество пусто.

Случай $k = 0$ немного хитрее. Все получается как нельзя лучше, если согласиться с тем, что существует только один способ разбиения пустого множества на нулевое количество непустых частей; следовательно, ${0 \brace 0} = 1$. Но непустое множество

Таблица 319. Треугольник Стирлинга для числа циклов

n	$\left[{n \atop 0}\right]$	$\left[{n \atop 1}\right]$	$\left[{n \atop 2}\right]$	$\left[{n \atop 3}\right]$	$\left[{n \atop 4}\right]$	$\left[{n \atop 5}\right]$	$\left[{n \atop 6}\right]$	$\left[{n \atop 7}\right]$	$\left[{n \atop 8}\right]$	$\left[{n \atop 9}\right]$
0	1									
1	0	1								
2	0	1	1							
3	0	2	3	1						
4	0	6	11	6	1					
5	0	24	50	35	10	1				
6	0	120	274	225	85	15	1			
7	0	720	1764	1624	735	175	21	1		
8	0	5040	13068	13132	6769	1960	322	28	1	
9	0	40320	109584	118124	67284	22449	4536	546	36	1

требует наличия по крайней мере одного подмножества, так что $\left\{{n \atop 0}\right\} = 0$ при $n > 0$.

А что же будет при $k = 2$? Конечно же, $\left\{{0 \atop 2}\right\} = 0$. Если множество из $n > 0$ объектов разделяется на две непустые части, то одна из них содержит последний объект и некоторое подмножество из первых $n-1$ объектов. Существует 2^{n-1} способов выбрать последнее подмножество, поскольку каждый из первых $n-1$ объектов либо входит в него, либо не входит. Но мы не должны помещать в него все эти объекты, потому что мы хотим получить две непустые части. Поэтому вычтем 1:

$$\left\{{n \atop 2}\right\} = 2^{n-1} - 1, \qquad \text{целое } n > 0. \tag{6.2}$$

(Это согласуется с приведенным выше перечислением способов разбиения: $\left\{{4 \atop 2}\right\} = 7 = 2^3 - 1$.)

Модификация данных рассуждений приводит к рекуррентному соотношению, с помощью которого можно вычислить $\left\{{n \atop k}\right\}$ для любого k. Если задано множество из $n > 0$ объектов, которое должно быть разбито на k непустых частей, то мы помещаем последний объект либо в отдельный класс ($\left\{{n-1 \atop k-1}\right\}$ способами), либо в некоторое непустое множество из первых $n-1$ объектов. В последнем случае имеется $k\left\{{n-1 \atop k}\right\}$ возможных вариантов, потому что каждый из $\left\{{n-1 \atop k}\right\}$ способов распределения первых $n-1$ объектов по k непустым частям дает k подмножеств, с которыми можно объединить n-й объект. Следовательно,

$$\left\{{n \atop k}\right\} = k\left\{{n-1 \atop k}\right\} + \left\{{n-1 \atop k-1}\right\}, \qquad \text{целое } n > 0. \tag{6.3}$$

Это именно то правило, в соответствии с которым строится табл. 318; без множителя k оно сводится к формуле сложения (5.8), которая генерирует треугольник Паскаля.

А теперь перейдем к числам Стирлинга первого рода. Они отчасти похожи на числа Стирлинга второго рода, но ${n \brack k}$ подсчитывает количество способов представления n объектов виде k *циклов* вместо подмножеств. Произносить обозначение '${n \brack k}$' можно как "k циклов из n".

Циклы — это циклические представления, аналогичные ожерельям, рассматривавшимся в главе 4. Цикл

$$\begin{array}{ccc} & A & \\ D & & B \\ & C & \end{array}$$

можно записать более компактно как '$[A, B, C, D]$', понимая при этом, что

$$[A, B, C, D] = [B, C, D, A] = [C, D, A, B] = [D, A, B, C];$$

цикл "зацикливается", т.е. его конец соединяется с его началом. С другой стороны, цикл $[A, B, C, D]$ — не одно и то же, что и цикл $[A, B, D, C]$ или $[D, C, B, A]$.

Существует одиннадцать различных способов составления двух циклов из четырех элементов:

$$\begin{array}{llll} [1,2,3]\,[4], & [1,2,4]\,[3], & [1,3,4]\,[2], & [2,3,4]\,[1], \\ [1,3,2]\,[4], & [1,4,2]\,[3], & [1,4,3]\,[2], & [2,4,3]\,[1], \\ [1,2]\,[3,4], & [1,3]\,[2,4], & [1,4]\,[2,3]; & \end{array} \qquad (6.4)$$

следовательно, ${4 \brack 2} = 11$.

"Есть девять и еще шестьдесят способов сложить песнь племени, и-каждый-из-них-по-отдельности-правильный".
— Р. Киплинг (R. Kipling)

Единичный цикл (т.е. цикл, состоящий из одного элемента) — это по сути то же самое, что и единичное множество (т.е. множество, состоящее из одного элемента). Аналогично 2-цикл подобен 2-множеству, потому что $[A, B] = [B, A]$, так же как и $\{A, B\} = \{B, A\}$. Однако имеется два *разных* 3-цикла, $[A, B, C]$ и $[A, C, B]$. Заметим, например, что одиннадцать пар циклов в (6.4) можно получить из семи пар множеств в (6.1), составив из каждого 3-элементного множества по два цикла.

В общем случае из любого n-элементного множества можно составить $n!/n = (n-1)!$ различных n-циклов, если только $n > 0$. (Имеется $n!$ перестановок, и каждый n-цикл соответствует n из них, потому что цикл может начинаться с любого из своих элементов.) Таким образом,

$${n \brack 1} = (n-1)!, \qquad \text{целое } n > 0. \qquad (6.5)$$

Это значительно больше величины ${n \brace 1} = 1$, которая была получена для количества подмножеств Стирлинга. В самом деле, легко убедиться в том, что число циклов должно быть по меньшей мере таким же, как и число подмножеств:

$$\left[{n \atop k}\right] \geqslant \left\{{n \atop k}\right\}, \qquad \text{целые } n, k \geqslant 0, \tag{6.6}$$

потому что каждое разбиение на непустые подмножества приводит как минимум к одному представлению в виде циклов.

Равенство в (6.6) имеет место тогда, когда все циклы с необходимостью являются либо единичными, либо двойными — в силу того, что в таких случаях циклы эквивалентны подмножествам. Это происходит при $k = n$ и $k = n-1$; следовательно,

$$\left[{n \atop n}\right] = \left\{{n \atop n}\right\}; \qquad \left[{n \atop n-1}\right] = \left\{{n \atop n-1}\right\}.$$

В самом деле, легко увидеть, что

$$\left[{n \atop n}\right] = \left\{{n \atop n}\right\} = 1; \qquad \left[{n \atop n-1}\right] = \left\{{n \atop n-1}\right\} = \binom{n}{2}. \tag{6.7}$$

(Количество способов представления n объектов в виде $n-1$ циклов или подмножеств равно количеству способов выбора двух объектов, которые окажутся в одном и том же цикле или подмножестве.) Треугольные числа $\binom{n}{2} = 1, 3, 6, 10, \ldots$ имеются как в табл. 318, так и в табл. 319.

Рекуррентное соотношение для $\left[{n \atop k}\right]$ можно вывести путем видоизменения рассуждений, использовавшихся при выводе $\left\{{n \atop k}\right\}$. Каждое представление n объектов в виде k циклов либо помещает последний объект в отдельный цикл ($\left[{n-1 \atop k-1}\right]$ способами), либо вставляет этот объект в одно из $\left[{n-1 \atop k}\right]$ циклических представлений первых $n-1$ объектов. В последнем случае существует $n-1$ различных способов выполнения такой вставки. (Это требует определенной сообразительности, однако нетрудно убедиться, что имеется j способов размещения нового элемента в j-цикл для получения $(j+1)$-цикла. Так, например, когда $j = 3$, цикл $[A, B, C]$ приводит к

$$[A, B, C, D], \qquad [A, B, D, C] \qquad \text{или} \qquad [A, D, B, C]$$

при вставке нового элемента D, и никаких других возможностей не существует. Суммирование по всем j дает в итоге $n-1$ способов вставки n-го объекта в циклическое разбиение $n-1$ объектов.) Таким образом, искомое рекуррентное соотношение имеет вид

$$\left[{n \atop k}\right] = (n-1)\left[{n-1 \atop k}\right] + \left[{n-1 \atop k-1}\right], \qquad \text{целое } n > 0. \tag{6.8}$$

Это аналог формулы сложения, с помощью которой строится табл. 319.

Сравнение (6.8) и (6.3) показывает, что первый член в их правых частях умножается на его верхний индекс $(n-1)$ в случае чисел Стирлинга первого рода и на нижний индекс k в случае чисел Стирлинга второго рода. Таким образом, при выполнении доказательства при помощи математической индукции можно осуществить "внесение" в членах типа $n{n \brack k}$ и $k{n \brace k}$.

Каждая перестановка эквивалентна некоторому множеству циклов. Рассмотрим, например, перестановку, которая переводит строку цифр 123456789 в 384729156. Для наглядности ее можно представить в виде двух строк,

$$\begin{matrix}1&2&3&4&5&6&7&8&9\\ 3&8&4&7&2&9&1&5&6,\end{matrix}$$

ясно показывающих, что 1 переходит в 3, 2 переходит в 8 и т.д. Возникает циклическая структура, потому что число 1 переходит в 3, которое переходит в 4, которое переходит в 7, которое переходит обратно в 1; таким образом, мы имеем дело с циклом $[1,3,4,7]$. Другим циклом в данной перестановке является $[2,8,5]$; еще один цикл — $[6,9]$. Таким образом, перестановка 384729156 эквивалентна циклическому представлению

$$[1,3,4,7]\,[2,8,5]\,[6,9]\,.$$

Если имеется некоторая перестановка $\pi_1\pi_2\ldots\pi_n$ множества $\{1,2,\ldots,n\}$, то каждый элемент входит только в один цикл. Действительно, если начать с $m_0=m$ и рассматривать $m_1=\pi_{m_0}$, $m_2=\pi_{m_1}$ и т.д., то в конце концов мы будем должны вернуться к $m_k=m_0$. (Числа раньше или позже должны будут повториться, и первым числом, которое появится вновь, будет m_0, потому что мы знаем, что у других чисел m_1, m_2, …, m_{k-1} имеются собственные уникальные предшественники.) Таким образом, каждая перестановка определяет некоторое циклическое представление. И обратно, каждое циклическое представление очевидным образом определяет перестановку (для этого достаточно лишь обратить построение), и это взаимно однозначное соответствие показывает, что перестановки и циклические представления по сути являются одним и тем же.

Следовательно, ${n \brack k}$ — количество перестановок n объектов, содержащих ровно по k циклов. Если просуммировать ${n \brack k}$ по всем k, то должно получиться общее количество перестановок:

$$\sum_{k=0}^{n} {n \brack k} = n!\,, \qquad \text{целое } n \geqslant 0. \tag{6.9}$$

Например, $6+11+6+1=24=4!$.

Числа Стирлинга полезны потому, что рекуррентные соотношения (6.3) и (6.8) возникают в множестве разных задач. Например, если мы решим представить обычные степени x^n через убывающие степени $x^{\underline{n}}$, то обнаружим, что несколько первых значений дают

$$\begin{aligned}
x^0 &= x^{\underline{0}}\,;\\
x^1 &= x^{\underline{1}}\,;\\
x^2 &= x^{\underline{2}} + x^{\underline{1}}\,;\\
x^3 &= x^{\underline{3}} + 3x^{\underline{2}} + x^{\underline{1}}\,;\\
x^4 &= x^{\underline{4}} + 6x^{\underline{3}} + 7x^{\underline{2}} + x^{\underline{1}}\,.
\end{aligned}$$

Эти коэффициенты подозрительно похожи на числа из табл. 318, только прочитанные не слева направо, а справа налево; таким образом, можно быть почти уверенным, что общая формула имеет вид

Лучше было бы определить $\left\{{n \atop k}\right\} = \left[{n \atop k}\right] = 0$ *при* $k < 0$ *и* $n \geqslant 0$.

$$x^n = \sum_k \left\{{n \atop k}\right\} x^{\underline{k}}\,, \qquad \text{целое } n \geqslant 0. \tag{6.10}$$

Действительно, простое доказательство по индукции разрешает все сомнения: $x \cdot x^{\underline{k}} = x^{\underline{k+1}} + kx^{\underline{k}}$, потому что $x^{\underline{k+1}} = x^{\underline{k}}(x-k)$; следовательно, $x \cdot x^{n-1}$ есть

$$\begin{aligned}
x\sum_k \left\{{n-1 \atop k}\right\} x^{\underline{k}} &= \sum_k \left\{{n-1 \atop k}\right\} x^{\underline{k+1}} + \sum_k \left\{{n-1 \atop k}\right\} kx^{\underline{k}} =\\
&= \sum_k \left\{{n-1 \atop k-1}\right\} x^{\underline{k}} + \sum_k \left\{{n-1 \atop k}\right\} kx^{\underline{k}} =\\
&= \sum_k \left(k\left\{{n-1 \atop k}\right\} + \left\{{n-1 \atop k-1}\right\}\right) x^{\underline{k}} = \sum_k \left\{{n \atop k}\right\} x^{\underline{k}}.
\end{aligned}$$

Другими словами, числа подмножеств Стирлинга — это коэффициенты при факториальных степенях, которые дают обычные степени.

Можно пойти и по другому пути, потому что числа циклов Стирлинга представляют собой коэффициенты при обычных степенях, которые дают факториальные степени:

$$\begin{aligned}
x^{\overline{0}} &= x^0\,;\\
x^{\overline{1}} &= x^1\,;\\
x^{\overline{2}} &= x^2 + x^1\,;\\
x^{\overline{3}} &= x^3 + 3x^2 + 2x^1\,;\\
x^{\overline{4}} &= x^4 + 6x^3 + 11x^2 + 6x^1\,.
\end{aligned}$$

Мы имеем $(x+n-1)\cdot x^{\overline{k}} = x^{\overline{k+1}}+(n-1)x^{\overline{k}}$, так что доказательство, подобное только что проведенному, показывает, что

$$(x+n-1)x^{\overline{n-1}} = (x+n-1)\sum_k \begin{bmatrix} n-1 \\ k \end{bmatrix} x^k = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} x^k .$$

Это приводит к доказательству по индукции общей формулы

$$x^{\overline{n}} = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} x^k , \qquad \text{целое } n \geqslant 0. \tag{6.11}$$

(Подстановка $x = 1$ вновь дает (6.9).)

"Минутку! — можете сказать вы. — Это равенство включает возрастающие факториальные степени $x^{\overline{n}}$, в то время как (6.10) включает убывающие факториальные степени $x^{\underline{n}}$. Что если нам потребуется выразить $x^{\underline{n}}$ через обычные степени или x^n через возрастающие степени?" Легко: достаточно добавить несколько знаков "минус" и получить

$$x^n = \sum_k \begin{Bmatrix} n \\ k \end{Bmatrix} (-1)^{n-k} x^{\overline{k}} , \qquad \text{целое } n \geqslant 0; \tag{6.12}$$

$$x^{\underline{n}} = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} x^k , \qquad \text{целое } n \geqslant 0. \tag{6.13}$$

Это срабатывает; например, формула

$$x^{\underline{4}} = x(x-1)(x-2)(x-3) = x^4 - 6x^3 + 11x^2 - 6x$$

почти такая же, как и

$$x^{\overline{4}} = x(x+1)(x+2)(x+3) = x^4 + 6x^3 + 11x^2 + 6x ,$$

за исключением чередующихся знаков. Общее тождество

$$x^{\underline{n}} = (-1)^n (-x)^{\overline{n}} \tag{6.14}$$

из упр. 2.17 превращает (6.10) в (6.12), а (6.11) в (6.13), если изменить знак x.

Достаточно легко запомнить, когда следует вставлять множитель $(-1)^{n-k}$ в формулу наподобие (6.12), поскольку при большом x имеет место естественное упорядочение степеней:

$$x^{\overline{n}} > x^n > x^{\underline{n}} , \qquad \text{для всех } x > n > 1. \tag{6.15}$$

Числа Стирлинга $\left[{n \atop k}\right]$ и $\left\{{n \atop k}\right\}$ неотрицательны, так что знаки "минус" следует использовать при выражении "малых" степеней через "большие".

Таблица 325. Основные тождества для чисел Стирлинга при целом $n \geqslant 0$

Рекуррентные соотношения

$$\left\{{n \atop k}\right\} = k\left\{{n-1 \atop k}\right\} + \left\{{n-1 \atop k-1}\right\}$$

$$\left[{n \atop k}\right] = (n-1)\left[{n-1 \atop k}\right] + \left[{n-1 \atop k-1}\right]$$

Частные случаи

$$\left\{{n \atop 0}\right\} = \left[{n \atop 0}\right] = [n=0]$$

$$\left\{{n \atop 1}\right\} = [n>0] \qquad \left[{n \atop 1}\right] = (n-1)!\,[n>0]$$

$$\left\{{n \atop 2}\right\} = (2^{n-1}-1)[n>0] \qquad \left[{n \atop 2}\right] = (n-1)!\,H_{n-1}\,[n>0]$$

$$\left\{{n \atop n-1}\right\} = \left[{n \atop n-1}\right] = \binom{n}{2}$$

$$\left\{{n \atop n}\right\} = \left[{n \atop n}\right] = \binom{n}{n} = 1$$

$$\left\{{n \atop k}\right\} = \left[{n \atop k}\right] = \binom{n}{k} = 0, \qquad \text{если } k>n$$

Преобразования степеней

$$x^n = \sum_k \left\{{n \atop k}\right\} x^{\underline{k}} = \sum_k \left\{{n \atop k}\right\} (-1)^{n-k} x^{\overline{k}}$$

$$x^{\underline{n}} = \sum_k \left[{n \atop k}\right] (-1)^{n-k} x^k$$

$$x^{\overline{n}} = \sum_k \left[{n \atop k}\right] x^k$$

Формулы обращения

$$\sum_k \left[{n \atop k}\right]\left\{{k \atop m}\right\}(-1)^{n-k} = [m=n]$$

$$\sum_k \left\{{n \atop k}\right\}\left[{k \atop m}\right](-1)^{n-k} = [m=n]$$

Таблица 326. Дополнительные тождества для чисел Стирлинга при целых $l, m, n \geqslant 0$

$$\left\{ {n+1 \atop m+1} \right\} = \sum_k \binom{n}{k} \left\{ {k \atop m} \right\} \tag{6.16}$$

$$\left[{n+1 \atop m+1} \right] = \sum_k \left[{n \atop k} \right] \binom{k}{m} \tag{6.17}$$

$$\left\{ {n \atop m} \right\} = \sum_k \binom{n}{k} \left\{ {k+1 \atop m+1} \right\} (-1)^{n-k} \tag{6.18}$$

$$\left[{n \atop m} \right] = \sum_k \left[{n+1 \atop k+1} \right] \binom{k}{m} (-1)^{m-k} \tag{6.19}$$

$n^m(-1)^{n-m}\left[{n \atop m} \right] = \sum_k \left[{n \atop k} \right] \binom{-m}{k-m} n^k$

$$m! \left\{ {n \atop m} \right\} = \sum_k \binom{m}{k} k^n (-1)^{m-k} \tag{6.20}$$

$$\left\{ {n+1 \atop m+1} \right\} = \sum_{k=0}^{n} \left\{ {k \atop m} \right\} (m+1)^{n-k} \tag{6.21}$$

$$\left[{n+1 \atop m+1} \right] = \sum_{k=0}^{n} \left[{k \atop m} \right] n^{\underline{n-k}} = n! \sum_{k=0}^{n} \left[{k \atop m} \right] / k! \tag{6.22}$$

$$\left\{ {m+n+1 \atop m} \right\} = \sum_{k=0}^{m} k \left\{ {n+k \atop k} \right\} \tag{6.23}$$

$$\left[{m+n+1 \atop m} \right] = \sum_{k=0}^{m} (n+k) \left[{n+k \atop k} \right] \tag{6.24}$$

$$\binom{n}{m} = \sum_k \left\{ {n+1 \atop k+1} \right\} \left[{k \atop m} \right] (-1)^{m-k} \tag{6.25}$$

$$n^{\underline{n-m}} [n \geqslant m] = \sum_k \left[{n+1 \atop k+1} \right] \left\{ {k \atop m} \right\} (-1)^{m-k} \tag{6.26}$$

Также $\binom{n}{m}(n-1)^{\underline{n-m}} = \sum_k \left[{n \atop k} \right] \left\{ {k \atop m} \right\}$, *обобщение (6.9).*

$$\left\{ {n \atop n-m} \right\} = \sum_k \binom{m-n}{m+k} \binom{m+n}{n+k} \left[{m+k \atop k} \right] \tag{6.27}$$

$$\left[{n \atop n-m} \right] = \sum_k \binom{m-n}{m+k} \binom{m+n}{n+k} \left\{ {m+k \atop k} \right\} \tag{6.28}$$

$$\left\{ {n \atop l+m} \right\} \binom{l+m}{l} = \sum_k \left\{ {k \atop l} \right\} \left\{ {n-k \atop m} \right\} \binom{n}{k} \tag{6.29}$$

$$\left[{n \atop l+m} \right] \binom{l+m}{l} = \sum_k \left[{k \atop l} \right] \left[{n-k \atop m} \right] \binom{n}{k} \tag{6.30}$$

Формулу (6.11) можно подставить в (6.12) и получить двойную сумму:

$$x^n = \sum_k \left\{ {n \atop k} \right\} (-1)^{n-k} x^{\overline{k}} = \sum_{k,m} \left\{ {n \atop k} \right\} \left[{k \atop m} \right] (-1)^{n-k} x^m .$$

Это справедливо при всех x, так что коэффициенты при x^0, x^1, ..., x^{n-1}, x^{n+1}, x^{n+2}, ... в правой части должны быть равны нулю, и мы получаем тождество

$$\sum_k \left\{ {n \atop k} \right\} \left[{k \atop m} \right] (-1)^{n-k} = [m=n] , \qquad \text{целые } m, n \geqslant 0. \quad (6.31)$$

Числа Стирлинга, подобно биномиальным коэффициентам, удовлетворяют многим удивительным тождествам. Однако эти тождества не столь гибки, как тождества из главы 5, а потому применяются не столь часто. Таким образом, лучше всего просто перечислить простейшие из них для будущих ссылок в случаях, когда придется колоть какой-нибудь не в меру крепкий стирлингов орешек. В табл. 325 и 326 содержатся наиболее часто применяющиеся формулы; уже выведенные нами тождества также повторены здесь.

При изучении биномиальных коэффициентов в главе 5 мы обнаружили, что выгодно определить $\binom{n}{k}$ для отрицательных n таким образом, чтобы тождество $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ оставалось корректным без ограничений. Используя это тождество для распространения $\binom{n}{k}$ за пределы комбинаторики, мы обнаружили (в табл. 212), что треугольник Паскаля при его продолжении вверх, в сущности, воспроизводит себя в перевернутом виде. Давайте попробуем сделать то же и с треугольниками Стирлинга. Что произойдет, если принять, что базовые рекуррентные соотношения

$$\left\{ {n \atop k} \right\} = k \left\{ {n-1 \atop k} \right\} + \left\{ {n-1 \atop k-1} \right\} \qquad \text{и}$$

$$\left[{n \atop k} \right] = (n-1) \left[{n-1 \atop k} \right] + \left[{n-1 \atop k-1} \right]$$

справедливы при любых целых n и k? Решение становится однозначным, если добавить дополнительные разумные соглашения, что

$$\left\{ {0 \atop k} \right\} = \left[{0 \atop k} \right] = [k=0] \quad \text{и} \quad \left\{ {n \atop 0} \right\} = \left[{n \atop 0} \right] = [n=0] . \qquad (6.32)$$

В самом деле, вырисовывается удивительная и красивая картина: треугольник Стирлинга для циклов оказывается над треугольником Стирлинга для подмножеств, и наоборот! Оба рода чисел

Таблица 328. Треугольники Стирлинга в тандеме

n	${n \brace -5}$	${n \brace -4}$	${n \brace -3}$	${n \brace -2}$	${n \brace -1}$	${n \brace 0}$	${n \brace 1}$	${n \brace 2}$	${n \brace 3}$	${n \brace 4}$	${n \brace 5}$
−5	1										
−4	10	1									
−3	35	6	1								
−2	50	11	3	1							
−1	24	6	2	1	1						
0	0	0	0	0	0	1					
1	0	0	0	0	0	0	1				
2	0	0	0	0	0	0	1	1			
3	0	0	0	0	0	0	1	3	1		
4	0	0	0	0	0	0	1	7	6	1	
5	0	0	0	0	0	0	1	15	25	10	1

Стирлинга связаны исключительно простым законом [220, 221]:

$$\left[{n \atop k}\right] = \left\{{-k \atop -n}\right\}, \qquad \text{целые } k, n. \tag{6.33}$$

Мы получили "дуальность", напоминающую взаимоотношения между min и max, между $\lfloor x \rfloor$ и $\lceil x \rceil$, между $x^{\underline{n}}$ и $x^{\overline{n}}$, между НОД и НОК. Легко проверить, что при таком соответствии рекуррентные соотношения $\left[{n \atop k}\right] = (n-1)\left[{n-1 \atop k}\right] + \left[{n-1 \atop k-1}\right]$ и $\left\{{n \atop k}\right\} = k\left\{{n-1 \atop k}\right\} + \left\{{n-1 \atop k-1}\right\}$ означают одно и то же.

6.2 Числа Эйлера

Время от времени возникает еще один треугольник чисел, которым мы обязаны Эйлеру [104, §13; 110, с. 485] и элементы которого мы обозначаем как '$\left\langle{n \atop k}\right\rangle$'. Угловые скобки в данном обозначении напоминают знаки "меньше" и "больше", ведь $\left\langle{n \atop k}\right\rangle$ — это количество перестановок $\pi_1\pi_2\ldots\pi_n$ множества $\{1, 2, \ldots, n\}$, имеющих k *подъемов*, т.е. k мест, где $\pi_j < \pi_{j+1}$. (*Предупреждение.* Это обозначение еще менее стандартизовано, чем наши обозначения $\left[{n \atop k}\right]$, $\left\{{n \atop k}\right\}$ для чисел Стирлинга. Но мы увидим, что в этом обозначении есть свой резон.)

Кнут [209, первое издание] вместо $\left\langle{n \atop k}\right\rangle$ использовал $\left\langle{n \atop k+1}\right\rangle$.

Например, одиннадцать перестановок множества $\{1, 2, 3, 4\}$ содержат по два участка подъема:

$$1324, \quad 1423, \quad 2314, \quad 2413, \quad 3412;$$
$$1243, \quad 1342, \quad 2341; \qquad 2134, \quad 3124, \quad 4123.$$

Таблица 329. Треугольник Эйлера

n	$\left\langle{n \atop 0}\right\rangle$	$\left\langle{n \atop 1}\right\rangle$	$\left\langle{n \atop 2}\right\rangle$	$\left\langle{n \atop 3}\right\rangle$	$\left\langle{n \atop 4}\right\rangle$	$\left\langle{n \atop 5}\right\rangle$	$\left\langle{n \atop 6}\right\rangle$	$\left\langle{n \atop 7}\right\rangle$	$\left\langle{n \atop 8}\right\rangle$	$\left\langle{n \atop 9}\right\rangle$
0	1									
1	1	0								
2	1	1	0							
3	1	4	1	0						
4	1	11	11	1	0					
5	1	26	66	26	1	0				
6	1	57	302	302	57	1	0			
7	1	120	1191	2416	1191	120	1	0		
8	1	247	4293	15619	15619	4293	247	1	0	
9	1	502	14608	88234	156190	88234	14608	502	1	0

(В первой строке перечислены перестановки с $\pi_1 < \pi_2 > \pi_3 < \pi_4$; во второй — перестановки с $\pi_1 < \pi_2 < \pi_3 > \pi_4$ и $\pi_1 > \pi_2 < \pi_3 < \pi_4$.) Следовательно, $\left\langle{4 \atop 2}\right\rangle = 11$. В табл. 329 перечислены небольшие числа Эйлера; на этот раз "фирменным знаком" служит последовательность 1, 11, 11, 1. При $n > 0$ в наличии может быть не более $n-1$ подъемов, так что на диагонали треугольника мы имеем $\left\langle{n \atop n}\right\rangle = [n=0]$.

Треугольник Эйлера, подобно треугольнику Паскаля, симметричен слева направо, но в данном случае закон симметрии несколько иной:

$$\left\langle{n \atop k}\right\rangle = \left\langle{n \atop n-1-k}\right\rangle, \qquad \text{целое } n > 0. \tag{6.34}$$

Перестановка $\pi_1\pi_2\ldots\pi_n$ имеет $n-1-k$ подъемов тогда и только тогда, когда ее "отражение" $\pi_n\ldots\pi_2\pi_1$ имеет k подъемов.

Давайте попробуем найти рекуррентное соотношение для $\left\langle{n \atop k}\right\rangle$. Каждая перестановка $\rho = \rho_1\ldots\rho_{n-1}$ множества $\{1,\ldots,n-1\}$ приводит к n перестановкам множества $\{1,2,\ldots,n\}$, если вставлять новый элемент n во все возможные места. Предположим, что мы вставляем n в позиции j, получая перестановку $\pi = \rho_1\ldots\rho_{j-1}\,n\,\rho_j\ldots\rho_{n-1}$. Количество подъемов в π такое же, как и в ρ, если $j=1$ или если $\rho_{j-1} < \rho_j$; если же $\rho_{j-1} > \rho_j$ или если $j = n$, то оно на единицу больше, чем в ρ. Таким образом, перестановка π с k подъемами получается $(k+1)\left\langle{n-1 \atop k}\right\rangle$ способами из перестановок ρ, которые содержат k подъемов, плюс $\big((n-2)-(k-1)+1\big)\left\langle{n-1 \atop k-1}\right\rangle$ способами из перестановок ρ, которые имеют $k-1$ подъемов. Искомое рекуррентное соотношение имеет вид

$$\left\langle {n \atop k} \right\rangle = (k+1)\left\langle {n-1 \atop k} \right\rangle + (n-k)\left\langle {n-1 \atop k-1} \right\rangle, \quad \text{целое } n > 0. \tag{6.35}$$

Начать вычисления по рекуррентному соотношению можно, положив

$$\left\langle {0 \atop k} \right\rangle = [k=0], \qquad \text{целое } k, \tag{6.36}$$

и считая, что $\left\langle {n \atop k} \right\rangle = 0$ при $k < 0$.

Числа Эйлера полезны, в первую очередь, потому, что они предоставляют необычные связи между обычными степенями и последовательными биномиальными коэффициентами:

$$x^n = \sum_k \left\langle {n \atop k} \right\rangle \binom{x+k}{n}, \qquad \text{целое } n \geqslant 0. \tag{6.37}$$

(Это так называемое тождество Воропицкого (Worpitzky) [378].) Например,

$$\begin{aligned} x^2 &= \binom{x}{2} + \binom{x+1}{2}, \\ x^3 &= \binom{x}{3} + 4\binom{x+1}{3} + \binom{x+2}{3}, \\ x^4 &= \binom{x}{4} + 11\binom{x+1}{4} + 11\binom{x+2}{4} + \binom{x+3}{4} \end{aligned}$$

Западные ученые недавно узнали о выдающейся китайской книге Ли Сянь-Ляня (Li Shan-Lan) [249; 265, с.320–325], опубликованной в 1867 году, в которой впервые упоминается формула (6.37).

и т.д. Тождество (6.37) легко доказать по индукции (упр. 14).

Кстати, (6.37) дает еще один способ вычисления суммы первых n квадратов: $k^2 = \left\langle {2 \atop 0} \right\rangle \binom{k}{2} + \left\langle {2 \atop 1} \right\rangle \binom{k+1}{2} = \binom{k}{2} + \binom{k+1}{2}$, следовательно,

$$\begin{aligned} 1^2+2^2+\cdots+n^2 &= \left(\binom{1}{2}+\binom{2}{2}+\cdots+\binom{n}{2}\right) + \left(\binom{2}{2}+\binom{3}{2}+\cdots+\binom{n+1}{2}\right) \\ &= \binom{n+1}{3}+\binom{n+2}{3} = \tfrac{1}{6}(n+1)n\big((n-1)+(n+2)\big). \end{aligned}$$

Рекуррентное соотношение Эйлера (6.35) несколько сложнее рекуррентных соотношений Стирлинга (6.3) и (6.8), так что не приходится ожидать, что числа $\left\langle {n \atop k} \right\rangle$ будут удовлетворять такому же количеству простых тождеств. Тем не менее кое-что имеется:

$$\left\langle {n \atop m} \right\rangle = \sum_{k=0}^{m} \binom{n+1}{k} (m+1-k)^n (-1)^k; \tag{6.38}$$

Таблица 331. Треугольник Эйлера второго порядка

n	$\left\langle\!\left\langle {n \atop 0} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 1} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 2} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 3} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 4} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 5} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 6} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 7} \right\rangle\!\right\rangle$	$\left\langle\!\left\langle {n \atop 8} \right\rangle\!\right\rangle$
0	1								
1	1	0							
2	1	2	0						
3	1	8	6	0					
4	1	22	58	24	0				
5	1	52	328	444	120	0			
6	1	114	1452	4400	3708	720	0		
7	1	240	5610	32120	58140	33984	5040	0	
8	1	494	19950	195800	644020	785304	341136	40320	0

$$m!\left\{ {n \atop m} \right\} = \sum_k \left\langle {n \atop k} \right\rangle \binom{k}{n-m}; \qquad (6.39)$$

$$\left\langle {n \atop m} \right\rangle = \sum_k \left\{ {n \atop k} \right\} \binom{n-k}{m} (-1)^{n-k-m} k!. \qquad (6.40)$$

Если умножить (6.39) на z^{n-m} и просуммировать по m, то получим $\sum_m \left\{ {n \atop m} \right\} m!\, z^{n-m} = \sum_k \left\langle {n \atop k} \right\rangle (z+1)^k$. Заменяя z на $z-1$ и приравнивая коэффициенты при z^k, получим (6.40). Таким образом, два последних тождества по сути эквивалентны. Первое тождество, (6.38), дает нам частные случаи при малых m:

$$\left\langle {n \atop 0} \right\rangle = 1;$$

$$\left\langle {n \atop 1} \right\rangle = 2^n - n - 1;$$

$$\left\langle {n \atop 2} \right\rangle = 3^n - (n+1)2^n + \binom{n+1}{2}.$$

Больше нам незачем задерживаться на числах Эйлера — достаточно просто знать о том, что они существуют, и располагать основными тождествами, чтобы иметь возможность при необходимости прибегнуть к ним. Однако, прежде чем оставить эту тему, следует упомянуть еще об одном треугольнике коэффициентов, показанном в табл. 331. Мы назовем эти числа $\left\langle\!\left\langle {n \atop k} \right\rangle\!\right\rangle$ "числами Эйлера второго порядка", потому что они удовлетворяют рекуррентному соотношению, подобному (6.35), но с заменой n на $2n-1$ в одном месте:

$$\left\langle\!\left\langle {n \atop k} \right\rangle\!\right\rangle = (k+1)\left\langle\!\left\langle {n-1 \atop k} \right\rangle\!\right\rangle + (2n-1-k)\left\langle\!\left\langle {n-1 \atop k-1} \right\rangle\!\right\rangle. \qquad (6.41)$$

Эти числа имеют одну интересную комбинаторную интерпретацию, впервые замеченную Гессель (Gessel) и Стенли (Stanley) [147]: если образовать перестановки мультимножества $\{1,1,2,2,\ldots,n,n\}$ с тем особым свойством, что все числа между двумя встречающимися m больше этого m, где $1 \leqslant m \leqslant n$, то $\left\langle\!\left\langle {n \atop k} \right\rangle\!\right\rangle$ представляет собой количество таких перестановок, которые содержат k подъемов. Например, имеется восемь подходящих "одноподъемных" перестановок множества $\{1,1,2,2,3,3\}$:

$$\begin{array}{llll} 113322, & 133221, & 221331, & 221133, \\ 223311, & 233211, & 331122, & 331221. \end{array}$$

Таким образом, $\left\langle\!\left\langle {3 \atop 1} \right\rangle\!\right\rangle = 8$. Мультимножество $\{1,1,2,2,\ldots,n,n\}$ имеет всего

$$\sum_k \left\langle\!\!\left\langle {n \atop k} \right\rangle\!\!\right\rangle = (2n-1)(2n-3)\ldots(1) = \frac{(2n)^{\underline{n}}}{2^n} \qquad (6.42)$$

соответствующих перестановок, потому что оба появления n должны быть смежны, и имеется $2n-1$ мест их вставки в перестановку мультимножества $\{1,1,2,2,\ldots,n-1,n-1\}$. Например, при $n=3$ перестановка 1221 имеет пять точек вставки, что дает 331221, 133221, 123321, 122331 и 122133. Рекуррентное соотношение (6.41) может быть доказано путем распространения на него аргументации, использовавшейся для обычных чисел Эйлера.

Числа Эйлера второго порядка важны в основном в силу своей связи с числами Стирлинга [148]: индукцией по n получаем, что

$$\left\{ {x \atop x-n} \right\} = \sum_k \left\langle\!\!\left\langle {n \atop k} \right\rangle\!\!\right\rangle \binom{x+n-1-k}{2n}, \quad \text{целое } n \geqslant 0; \qquad (6.43)$$

$$\left[{x \atop x-n} \right] = \sum_k \left\langle\!\!\left\langle {n \atop k} \right\rangle\!\!\right\rangle \binom{x+k}{2n}, \quad \text{целое } n \geqslant 0. \qquad (6.44)$$

Например,

$$\left\{ {x \atop x-1} \right\} = \binom{x}{2}, \qquad \left[{x \atop x-1} \right] = \binom{x}{2};$$

$$\left\{ {x \atop x-2} \right\} = \binom{x+1}{4} + 2\binom{x}{4}, \qquad \left[{x \atop x-2} \right] = \binom{x}{4} + 2\binom{x+1}{4};$$

$$\left\{ {x \atop x-3} \right\} = \binom{x+2}{6} + 8\binom{x+1}{6} + 6\binom{x}{6},$$

$$\left[{x \atop x-3} \right] = \binom{x}{6} + 8\binom{x+1}{6} + 6\binom{x+2}{6}.$$

(Мы уже сталкивались со случаем $n = 1$ в (6.7).) Эти тождества выполняются, когда x — целое и n — неотрицательное целое число. Поскольку правые части этих тождеств являются полиномами от x, можно воспользоваться (6.43) и (6.44) для определения чисел Стирлинга $\left\{ {x \atop x-n} \right\}$ и $\left[{x \atop x-n} \right]$ при произвольных действительных (или комплексных) значениях x.

Если $n > 0$, полиномы $\left\{ {x \atop x-n} \right\}$ и $\left[{x \atop x-n} \right]$ равны нулю при $x = 0$, $x = 1$, ... и $x = n$; следовательно, они делятся на $(x - 0)$, $(x - 1)$, ... и $(x - n)$. Интересно взглянуть на то, что остается после деления на эти множители. Определим *полиномы Стирлинга* $\sigma_n(x)$ при помощи правила

$$\sigma_n(x) = \left[{x \atop x-n} \right] \Big/ \big(x(x-1)\ldots(x-n)\big) . \qquad (6.45)$$

(Полином $\sigma_n(x)$ имеет степень $n-1$.) Несколько первых случаев таковы:

Полином ли $1/x$?

(А жаль...)

$$\begin{aligned} \sigma_0(x) &= 1/x ; \\ \sigma_1(x) &= 1/2 ; \\ \sigma_2(x) &= (3x-1)/24 ; \\ \sigma_3(x) &= (x^2-x)/48 ; \\ \sigma_4(x) &= (15x^3-30x^2+5x+2)/5760 . \end{aligned}$$

Они могут быть вычислены через числа Эйлера, например

$$\sigma_3(x) = \big((x-4)(x-5) + 8(x-4)(x+1) + 6(x+2)(x+1)\big)/6!$$

Оказывается, что эти полиномы удовлетворяют двум весьма привлекательным тождествам:

$$\left(\frac{ze^z}{e^z-1}\right)^x = x \sum_n \sigma_n(x)\, z^n ; \qquad (6.46)$$

$$\left(\frac{1}{z} \ln \frac{1}{1-z}\right)^x = x \sum_n \sigma_n(x+n)\, z^n . \qquad (6.47)$$

В общем случае, если степенной ряд $\mathcal{S}_t(z)$ удовлетворяет тождеству

$$\ln\big(1 - z\mathcal{S}_t(z)^{t-1}\big) = -z\mathcal{S}_t(z)^t , \qquad (6.48)$$

то

$$\mathcal{S}_t(z)^x = x \sum_n \sigma_n(x+tn)\, z^n . \qquad (6.49)$$

Таблица 334. Формулы сверток Стирлинга

$$rs\sum_{k=0}^{n} \sigma_k(r+tk)\,\sigma_{n-k}(s+t(n-k)) \;=\; (r+s)\,\sigma_n(r+s+tn) \qquad (6.50)$$

$$s\sum_{k=0}^{n} k\sigma_k(r+tk)\,\sigma_{n-k}(s+t(n-k)) \;=\; n\sigma_n(r+s+tn) \qquad (6.51)$$

$$\left\{ {n \atop m} \right\} \;=\; (-1)^{n-m+1}\frac{n!}{(m-1)!}\sigma_{n-m}(-m) \qquad (6.52)$$

$$\left[{n \atop m} \right] \;=\; \frac{n!}{(m-1)!}\sigma_{n-m}(n) \qquad (6.53)$$

Таким образом, можно вывести общие формулы сверток для чисел Стирлинга, как это делалось для биномиальных коэффициентов в табл. 255; результаты приведены в табл. 334. Если сумма чисел Стирлинга не соответствует тождествам из табл. 325 или 326, то табл. 334 может оказаться как раз тем, что нужно. (Такой пример приводится позже в этой главе, после уравнения (6.100). В упр. 7.19 обсуждаются общие принципы сверток, основанных на тождествах типа (6.46) и (6.49).)

6.3 Гармонические числа

Теперь пришла пора вплотную заняться гармоническими числами, с которыми мы впервые встретились еще в главе 2:

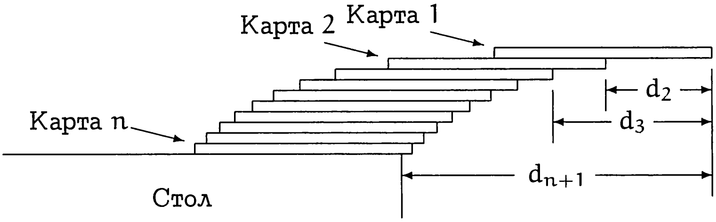
$$H_n \;=\; 1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n} \;=\; \sum_{k=1}^{n}\frac{1}{k}, \qquad \text{целое } n \geqslant 0. \qquad (6.54)$$

Эти числа возникают в процессе анализа алгоритмов так часто, что специалистам потребовалось для них специальное обозначение. Мы воспользуемся обозначением H_n, где 'H' происходит от слова "harmonic" (гармонический), при этом тон с длиной волны $1/n$ называется n-й гармоникой тона, длина волны которого равна 1. Вот несколько первых значений H_n:

n	0	1	2	3	4	5	6	7	8	9	10
H_n	0	1	$\frac{3}{2}$	$\frac{11}{6}$	$\frac{25}{12}$	$\frac{137}{60}$	$\frac{49}{20}$	$\frac{363}{140}$	$\frac{761}{280}$	$\frac{7129}{2520}$	$\frac{7381}{2520}$

В упр. 21 показано, что при $n > 1$ числа H_n никогда не являются целыми.

Ниже описан один карточный фокус, основанный на идее Р. Т. Шарпа (R. T. Sharp) [325], который иллюстрирует, насколько естественно гармонические числа возникают даже в простых ситуациях. Пусть у нас имеется n игральных карт и стол, на который мы хотим выложить эти карты стопкой с максимально возможным выступом над краем стола с учетом действия гравитации:



Для большей определенности задачи будем считать, что карта k располагается на карте $k+1$ ($1 \leqslant k < n$), и потребуем, чтобы края карт были параллельны краю стола; в противном случае величину выступа можно было бы увеличить, разворачивая карты так, чтобы их уголки выступали немного дальше. Для упрощения ответа будем считать, что длина каждой карты равна 2 единицам.

В случае единственной карты выступ максимален, когда ее центр тяжести находится ровно над краем стола. Поскольку центр тяжести однородной карты находится в ее центре, мы получаем выступ длиной в половину карты, т.е. длиной в 1 единицу.

В случае двух карт нетрудно убедиться, что максимальная величина выступа получается тогда, когда центр тяжести верхней карты находится ровно над краем второй карты, а общий центр тяжести обеих карт — ровно над краем стола. А поскольку общий центр тяжести двух карт находится посредине их совмещенных частей, то мы в состоянии увеличить величину выступа еще на пол-единицы.

Рассмотренные случаи подсказывают общий метод, в соответствии с которым карты размещаются так, чтобы центр тяжести k верхних карт располагался над краем $k+1$-й карты (которая находится под этими k верхними картами). Стол играет роль $n+1$-й карты. Чтобы выразить это условие алгебраически, можно обозначить через d_k расстояние от выступающего края самой верхней карты до соответствующего края k-й карты сверху. Тогда $d_1 = 0$, а d_{k+1} следует положить равным центру тяжести k первых карт:

$$d_{k+1} = \frac{(d_1+1)+(d_2+1)+\cdots+(d_k+1)}{k} \quad \text{при } 1 \leqslant k \leqslant n. \qquad (6.55)$$

(Центр тяжести k объектов, которые имеют веса $w_1, \ldots, w_k$ и центры тяжести которых находятся соответственно в точках p_1, $\ldots$ p_k, располагается в точке $(w_1p_1+\cdots+w_kp_k)/(w_1+\cdots+w_k)$.)

Это рекуррентное соотношение можно переписать двумя эквивалентными способами:

$$\begin{aligned} kd_{k+1} &= k + d_1 + \cdots + d_{k-1} + d_k\,, \qquad && k \geqslant 0;\\ (k-1)d_k &= k - 1 + d_1 + \cdots + d_{k-1}\,, && k \geqslant 1. \end{aligned}$$

Вычитание одного уравнения из другого показывает, что

$$kd_{k+1} - (k-1)d_k = 1 + d_k\,, \qquad k \geqslant 1;$$

следовательно, $d_{k+1} = d_k + 1/k$. Вторая карта будет сдвинута на половину единицы длины относительно третьей карты, которая сдвинута на треть единицы длины относительно четвертой карты, которая сдвинута на... Словом, по индукции получается общая формула

$$d_{k+1} = H_k\,, \tag{6.56}$$

и, положив $k = n$, мы получим полную величину выступа при укладке n карт описанным способом, равную $d_{n+1} = H_n$.

Нельзя ли достичь большего, воздерживаясь сначала от сдвига каждой карты на предельно возможное расстояние и накапливая "потенциальную энергию гравитации" для последующего решающего сдвига? Нет: любое устойчивое расположение карт должно удовлетворять неравенству

$$d_{k+1} \leqslant \frac{(1+d_1) + (1+d_2) + \cdots + (1+d_k)}{k}\,, \qquad 1 \leqslant k \leqslant n.$$

Кроме того, $d_1 = 0$. Отсюда по индукции следует, что $d_{k+1} \leqslant H_k$.

Заметим, что для того, чтобы верхняя карта полностью выступала за край стола, вовсе не требуется много карт. Нам нужно столько карт, чтобы величина выступа немного превышала длину одной карты, которая равна 2 единицам. Первым гармоническим числом, превышающим 2, является $H_4 = \frac{25}{12}$, так что нам достаточно всего лишь четырех карт.

Если у нас полная колода из 52 карт, то мы получаем выступ величиной H_{52} единиц, что равно $H_{52}/2 \approx 2.27$ длин карт. (Вскоре мы изучим формулу, которая показывает, как вычислять приближенное значение H_n при больших n, не складывая кучу дробей.)

Всякий, кто действительно пытается выложить 52 карты, похоже, имеет дело с неполной колодой...

Еще одна занятная задачка — "о червяке на резинке" — демонстрирует еще один облик гармонических чисел. Медленный, но чрезвычайно упорный червяк W начинает движение от одного конца метровой полоски резины к другому и ползет со скоростью 1 сантиметр в минуту. В конце каждой минуты столь же упорный держащий резинку некто K, единственная цель жизни которого состоит в доставлении неприятностей W, растягивает ее на метр. Таким образом, после минуты ползания W находится в 1 сантиметре от начала и в 99 сантиметрах от конца, после чего K растягивает ее на метр. В процессе растягивания W сохраняет свое относительное положение на резинке (в 1% от ее начала и в 99% от конца), так что теперь W находится в 2 сантиметрах от начала резинки и в 198 сантиметрах от цели. После еще одной минуты ползания на счету W 3 пройденных сантиметра и 197 сантиметрах впереди; но K в очередной раз растягивает резинку, и они превращаются в 4.5 и 295.5. И так далее... Доползет ли когда-нибудь червяк до финиша? Чем дальше он движется, тем больше удаляется от него цель. (Считается, что терпение (и время жизни) K и W, растяжимость резинки и крохотность червяка беспредельны.)

Напишем несколько формул. Когда K растягивает резинку, та ее часть, которую прополз W, остается неизменной. Таким образом, за первую минуту он проползает 1/100 ее часть, за вторую — 1/200, за третью — 1/300 и т.д. После n минут часть резинки, которую проползет червяк, составляет

$$\frac{1}{100}\left(\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}\right) = \frac{H_n}{100}. \tag{6.57}$$

Итак, червяк достигнет цели, когда H_n превысит 100.

Вскоре мы узнаем, как оценивать значение H_n при больших n, а пока что просто проверим наш анализ, рассмот в той же ситуации "суперчервяка", который в отличие от W в состоянии проползти за одну минуту 50 сантиметров. В соответствии с только что рассмотренным рассуждением за n минут он проползет $H_n/2$ длины резинки. Если наши рассуждения верны, то суперчервяк доберется до цели, прежде чем n достигнет 4, поскольку $H_4 > 2$. И это так: простой подсчет показывает, что после трех минут путешествия суперчервяку останется проползти всего $33\frac{1}{3}$ сантиметров. Он закончит свое путешествие через 3 минуты и 40 секунд.

Ценители виниловых дисков должны помнить это число (оборотов в минуту).

Гармонические числа появляются и в треугольнике Стирлинга. Попробуем найти аналитический вид для $\left[{n \atop 2}\right]$ — количества перестановок n объектов, содержащих ровно два цикла. Рекуррентное соотношение (6.8) гласит, что

$$\left[{n+1 \atop 2}\right] = n\left[{n \atop 2}\right] + \left[{n \atop 1}\right] = n\left[{n \atop 2}\right] + (n-1)! \qquad \text{при } n > 0;$$

и это рекуррентное соотношение является естественным кандидатом для метода суммирующего множителя из главы 2:

$$\frac{1}{n!}\left[{n+1 \atop 2}\right] = \frac{1}{(n-1)!}\left[{n \atop 2}\right] + \frac{1}{n}.$$

Развертывание этого рекуррентного соотношения приводит к $\frac{1}{n!}{n+1 \brack 2} = H_n$; следовательно,

$$ {n+1 \brack 2} = n!H_n \,. \qquad (6.58) $$

В главе 2 мы доказали, что гармонический ряд $\sum_k 1/k$ расходится, что означает, что H_n неограниченно возрастает при $n \to \infty$. Наше доказательство носило косвенный характер — мы нашли, что некоторая бесконечная сумма (2.58) дает разные результаты при перестановке ее членов, так что сумма $\sum_k 1/k$ не может быть ограниченной. Тот факт, что $H_n \to \infty$, представляется противоречащим здравому смыслу, поскольку, помимо прочего, это означает, что достаточно большая стопка карт может образовывать километровый выступ над краем стола и что в конце концов червяк W сможет отдохнуть, добравшись до цели своего путешествия. Давайте внимательнее посмотрим на величину H_n при больших n.

Простейший путь увидеть, что $H_n \to \infty$, вероятно, состоит в группировании членов ряда в соответствии со степенями 2. Поместим один член в группу 1, два члена в группу 2, четыре члена в группу 3, восемь членов в группу 4 и т.д.:

$$ \underbrace{\frac{1}{1}}_{\text{группа 1}} + \underbrace{\frac{1}{2}+\frac{1}{3}}_{\text{группа 2}} + \underbrace{\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}}_{\text{группа 3}} + $$
$$ + \underbrace{\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{11}+\frac{1}{12}+\frac{1}{13}+\frac{1}{14}+\frac{1}{15}}_{\text{группа 4}} + \cdots . $$

Оба члена группы 2 находятся между $\frac{1}{4}$ и $\frac{1}{2}$, так что сумма группы лежит между $2\cdot\frac{1}{4} = \frac{1}{2}$ и $2\cdot\frac{1}{2} = 1$. Все четыре члена группы 3 лежат между $\frac{1}{8}$ и $\frac{1}{4}$, так что их сумма также находится между $\frac{1}{2}$ и 1. И вообще, каждый из 2^{k-1} членов группы k находится между 2^{-k} и 2^{1-k}; следовательно, сумма каждой отдельной группы находится между $\frac{1}{2}$ и 1.

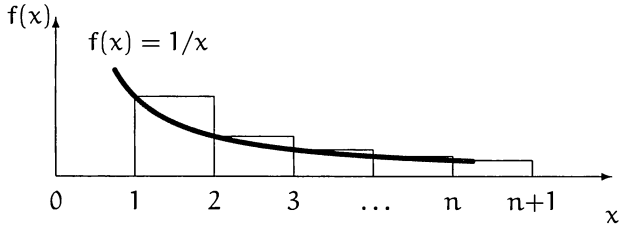
Такая процедура группирования показывает, что если n-й член находится в группе k, то $H_n > k/2$ и $H_n \leqslant k$ (по индукции по k). Таким образом, $H_n \to \infty$, а в действительности

$$ \frac{\lfloor \lg n \rfloor + 1}{2} < H_n \leqslant \lfloor \lg n \rfloor + 1 \,. \qquad (6.59) $$

Так и хочется назвать их "червячными" в честь никуда не торопящегося червяка из рассматривавшейся задачи.

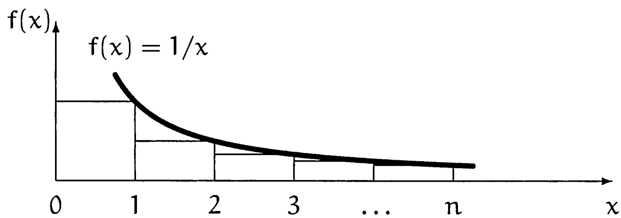
Итак, мы выяснили значение H_n с точностью до множителя 2. Хотя гармонические числа растут до бесконечности, делают они это всего лишь логарифмически, т.е. крайне медленно.

Более точные границы можно получить, если приложить немного больше усилий. Как мы выяснили в главе 2, H_n представляет собой дискретный аналог непрерывной функции $\ln n$. Натуральный логарифм определяется как площадь области под некоторой кривой, так что напрашивается следующее геометрическое сравнение:



Площадь области под данной кривой от 1 и n, равная $\int_1^n dx/x = \ln n$, меньше площади указанных n прямоугольников, которая, в свою очередь, равна $\sum_{k=1}^{n} 1/k = H_n$. Таким образом, $\ln n < H_n$; это более точный результат по сравнению с тем, что мы имели в (6.59). Разместив те же прямоугольники немного по-другому, мы получим аналогично и верхнюю границу:

"Теперь я вижу также способ найти... сумму из y^e членов гармонического ряда... с помощью логарифмов, но [проделать] y^e вычислений для нахождения этих правил было бы слишком затруднительно".
— И. Ньютон (I. Newton) [280]



В этот раз площадь n прямоугольников, равная H_n, меньше площади первого прямоугольника плюс площадь области под данной кривой. Тем самым доказано, что

$$\ln n < H_n < \ln n + 1 \qquad \text{при } n > 1. \tag{6.60}$$

Итак, теперь мы знаем величину H_n с ошибкой, не превосходящей 1.

Гармонические числа "второго порядка" $H_n^{(2)}$ возникают при суммировании квадратов чисел, обратных натуральным, вместо суммирования просто обратных натуральных чисел:

$$H_n^{(2)} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2} = \sum_{k=1}^{n} \frac{1}{k^2}.$$

Аналогично определяются гармонические числа r-го порядка — путем суммирования $(-r)$-х степеней:

$$H_n^{(r)} = \sum_{k=1}^{n} \frac{1}{k^r}. \tag{6.61}$$

Если $r > 1$, то при $n \to \infty$ эти числа стремятся к некоторым соответствующим пределам; в упр. 2.31 говорилось, что этот предел обычно называют дзета-функцией Римана (Riemann):

$$\zeta(r) = H_\infty^{(r)} = \sum_{k\geqslant 1} \frac{1}{k^r}. \tag{6.62}$$

Эйлер [103] открыл изящный способ использования обобщенных гармонических чисел для приближения ими обычных гармонических чисел $H_n^{(1)}$. Рассмотрим бесконечный ряд

$$\ln\left(\frac{k}{k-1}\right) = \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{3k^3} + \frac{1}{4k^4} + \cdots, \tag{6.63}$$

который сходится при $k > 1$. Левая его часть равна $\ln k - \ln(k-1)$; таким образом, при суммировании обеих частей по $2 \leqslant k \leqslant n$ левая сумма телескопируется, и мы получаем

$$\begin{aligned} \ln n - \ln 1 &= \sum_{k=2}^{n}\left(\frac{1}{k}+\frac{1}{2k^2}+\frac{1}{3k^3}+\frac{1}{4k^4}+\cdots\right) = \\ &= (H_n-1)+\tfrac{1}{2}(H_n^{(2)}-1)+\tfrac{1}{3}(H_n^{(3)}-1)+\tfrac{1}{4}(H_n^{(4)}-1)+\cdots. \end{aligned}$$

После перестановки получается выражение для разности между H_n и $\ln n$:

$$H_n - \ln n = 1-\tfrac{1}{2}(H_n^{(2)}-1)-\tfrac{1}{3}(H_n^{(3)}-1)-\tfrac{1}{4}(H_n^{(4)}-1)-\cdots.$$

При $n \to \infty$ правая часть стремится к предельному значению

$$1 - \tfrac{1}{2}(\zeta(2)-1) - \tfrac{1}{3}(\zeta(3)-1) - \tfrac{1}{4}(\zeta(4)-1) - \cdots,$$

известному как *константа Эйлера* и обычно обозначаемому греческой буквой γ. Поскольку $\zeta(r)-1$ приближенно равно $1/2^r$, этот бесконечный ряд сходится достаточно быстро, и мы можем вычислить десятичное значение

"Huius igitur quantitatis constantis C valorem deteximus, quippe est C = 0,577218."
— Л. Эйлер [103]

$$\gamma = 0.5772156649\ldots. \tag{6.64}$$

Рассуждения Эйлера приводят к предельному соотношению

$$\lim_{n\to\infty}(H_n - \ln n) = \gamma; \tag{6.65}$$

таким образом, на H_n приходится примерно 58% длины между обеими границами из (6.60) — так мы постепенно приблизились к его истинному значению.

Как будет видно из главы 9, можно продолжить уточнения. Например, будет доказано, что

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{\epsilon_n}{120n^4}, \quad 0 < \epsilon_n < 1. \qquad (6.66)$$

Эта формула позволяет, не прибегая к сложению миллиона дробей, заключить, что миллионное гармоническое число есть

$$H_{1000000} \approx 14.3927267228657236313811275.$$

Кроме прочего, это означает, что стопка из миллиона карт может выступать над краем стола на длину более чем семи карт.

А что формула (6.66) позволяет сказать о судьбе червяка на резинке? Поскольку значение H_n не ограничено, червяк обязательно доберется до конца ленты — как только H_n превысит 100. Наше приближение H_n утверждает, что это произойдет при n, приближенно равном

$$e^{100-\gamma} \approx e^{99.423}.$$

В действительности в упр. 9.49 доказывается, что переломное значение n есть либо $\lfloor e^{100-\gamma} \rfloor$, либо $\lceil e^{100-\gamma} \rceil$. Можно представить, с каким триумфом червяк W пересечет финишную черту к большому неудовольствию K, который поймет, что 287 дециллионов (287 с 33 нулями) веков растягивания резинки до длины более 10^{27} световых лет (при этом соседние атомы резинки окажутся на межзвездных расстояниях друг от друга) потрачены впустую...

Увы: червяк не доберется до конца резинки, так как гораздо раньше — как только будет полностью перемещена башня Брахмы — наступит конец света...

6.4 Гармоническое суммирование

Рассмотрим теперь некоторые суммы, в которых содержатся гармонические числа. Начнем с небольшого повторения того, о чем мы узнали в главе 2. В (2.36) и (2.57) мы доказали, что

$$\sum_{0 \leqslant k < n} H_k = nH_n - n; \qquad (6.67)$$

$$\sum_{0 \leqslant k < n} kH_k = \frac{n(n-1)}{2} H_n - \frac{n(n-1)}{4}. \qquad (6.68)$$

Давайте рискнем и возьмемся за общую сумму, которая включает обе указанные суммы в качестве частных случаев: чему равна сумма

$$\sum_{0 \leqslant k < n} \binom{k}{m} H_k,$$

если m — отрицательное целое число?

Метод, который лучше всего работал для вычисления (6.67) и (6.68) в главе 2, назывался *суммированием по частям.* Мы записывали общий член суммы в виде $u(k)\Delta v(k)$ и применяли общее правило

$$\sum\nolimits_a^b u(x)\Delta v(x)\,\delta x \;=\; u(x)v(x)\Big|_a^b - \sum\nolimits_a^b v(x+1)\Delta u(x)\,\delta x\,. \qquad (6.69)$$

Вспомнили? Сумма $\sum_{0\leqslant k<n}\binom{k}{m}H_k$, которая теперь перед нами, тоже вполне подходит для этого метода, поскольку можно положить

$$u(k) \;=\; H_k\,, \qquad\qquad \Delta u(k) \;=\; H_{k+1}-H_k \;=\; \frac{1}{k+1}\,;$$
$$v(k) \;=\; \binom{k}{m+1}, \qquad \Delta v(k) \;=\; \binom{k+1}{m+1}-\binom{k}{m+1} \;=\; \binom{k}{m}.$$

(Другими словами, гармонические числа имеют простую Δ, а биномиальные коэффициенты — простую Δ^{-1}, так что мы на верном пути.) Подстановка в (6.69) дает

$$\begin{aligned}\sum_{0\leqslant k<n}\binom{k}{m}H_k \;&=\; \sum\nolimits_0^n \binom{x}{m}H_x\,\delta x\\ &=\; \binom{x}{m+1}H_x\Big|_0^n - \sum\nolimits_0^n \binom{x+1}{m+1}\frac{\delta x}{x+1} \;=\\ &=\; \binom{n}{m+1}H_n - \sum_{0\leqslant k<n}\binom{k+1}{m+1}\frac{1}{k+1}\,.\end{aligned}$$

Оставшаяся сумма проста, поскольку можно внести $(k+1)^{-1}$ под знак биномиального коэффициента, пользуясь нашим старым проверенным равенством (5.5):

$$\begin{aligned}\sum_{0\leqslant k<n}\binom{k+1}{m+1}\frac{1}{k+1} \;&=\; \sum_{0\leqslant k<n}\binom{k}{m}\frac{1}{m+1} \;=\\ &=\; \binom{n}{m+1}\frac{1}{m+1}\,.\end{aligned}$$

Таким образом, вот искомый ответ:

$$\sum_{0\leqslant k<n}\binom{k}{m}H_k \;=\; \binom{n}{m+1}\left(H_n-\frac{1}{m+1}\right). \qquad (6.70)$$

(Что чудесно согласуется с (6.67) и (6.68) при $m=0$ и $m=1$.)

В следующем примере вместо умножения используется деление: давайте вычислим сумму

$$S_n \;=\; \sum_{k=1}^{n}\frac{H_k}{k}\,.$$

Если раскрыть H_k согласно его определению, можно получить двойную сумму

$$S_n = \sum_{1\leqslant j\leqslant k\leqslant n} \frac{1}{j\cdot k}.$$

Теперь нам на помощь приходит еще один метод из главы 2: равенство (2.33) говорит нам, что

$$S_n = \frac{1}{2}\left(\left(\sum_{k=1}^{n}\frac{1}{k}\right)^2 + \sum_{k=1}^{n}\frac{1}{k^2}\right) = \frac{1}{2}\left(H_n^2 + H_n^{(2)}\right). \qquad (6.71)$$

Оказывается, что этот ответ можно получить иначе, если прибегнуть к суммированию по частям (см. упр. 26).

Теперь испытаем свои силы на более трудной задаче [354], которая суммированию по частям не поддается:

$$U_n = \sum_{k\geqslant 1}\binom{n}{k}\frac{(-1)^{k-1}}{k}(n-k)^n, \qquad \text{целое } n \geqslant 1.$$

Чтобы не было и намека на ответ.

(В этой сумме нет и намека на гармонические числа; но как знать — вдруг они появятся позже?)

Мы будем решать эту задачу двумя способами: в одном случае трудом и потом вымучивая ответ, а во втором — полагаясь на смекалку и/или удачу. Сначала помучаемся. Разложим $(n-k)^n$ согласно биномиальной теореме, так чтобы объединить мешающее нам k в знаменателе с числителем:

$$\begin{aligned} U_n &= \sum_{k\geqslant 1}\binom{n}{k}\frac{(-1)^{k-1}}{k}\sum_j\binom{n}{j}(-k)^j n^{n-j} = \\ &= \sum_j\binom{n}{j}(-1)^{j-1}n^{n-j}\sum_{k\geqslant 1}\binom{n}{k}(-1)^k k^{j-1}. \end{aligned}$$

Это не такая путаница, как может показаться на первый взгляд, потому что k^{j-1} во внутренней сумме является полиномом от k, а равенство (5.40) подсказывает, что мы просто берем n-ю разность этого полинома. Оно-то так, но не совсем: сначала нам надо внести ясность в некоторые детали. Прежде всего, k^{j-1} не является полиномом, если $j = 0$; так что нам надо отделить этот член и работать с ним отдельно. Кроме того, нам не хватает члена с $k = 0$ из формулы для n-й разности. Этот член отличен от нуля при $j = 1$, так что лучше восстановить его (и тут же снова вычесть). В результате получается

$$\begin{aligned} U_n &= \sum_{j\geqslant 1}\binom{n}{j}(-1)^{j-1}n^{n-j}\sum_{k\geqslant 0}\binom{n}{k}(-1)^k k^{j-1}- \\ &\quad -\sum_{j\geqslant 1}\binom{n}{j}(-1)^{j-1}n^{n-j}\binom{n}{0}0^{j-1}- \\ &\quad -\binom{n}{0}n^n\sum_{k\geqslant 1}\binom{n}{k}(-1)^k k^{-1}\,. \end{aligned}$$

Отлично! Теперь верхняя строка (единственная, в которой осталась двойная сумма) равна нулю: это сумма величин, кратных n-м разностям полиномов степени, меньшей n, а такие n-е разности равны нулю. Вторая строка тоже равна нулю за исключением случая $j = 1$, когда она равна $-n^n$. Таким образом, единственное оставшееся затруднение — это третья строка. Мы свели исходную задачу к гораздо более простой сумме:

$$U_n = n^n(T_n - 1)\,, \qquad \text{где } T_n = \sum_{k\geqslant 1}\binom{n}{k}\frac{(-1)^{k-1}}{k}\,. \tag{6.72}$$

Например, $U_3 = \binom{3}{1}\frac{8}{1} - \binom{3}{2}\frac{1}{2} = \frac{45}{2}$; $T_3 = \binom{3}{1}\frac{1}{1} - \binom{3}{2}\frac{1}{2} + \binom{3}{3}\frac{1}{3} = \frac{11}{6}$; следовательно, $U_3 = 27(T_3 - 1)$, как и требовалось.

Как вычислить T_n? Один способ состоит в том, чтобы заменить $\binom{n}{k}$ на $\binom{n-1}{k} + \binom{n-1}{k-1}$, получив простую рекуррентную зависимость T_n от T_{n-1}. Однако имеется более поучительный способ: мы уже сталкивались с аналогичной формулой в (5.41), а именно

$$\sum_k \binom{n}{k}\frac{(-1)^k}{x+k} = \frac{n!}{x(x+1)\ldots(x+n)}\,.$$

Если вычесть член при $k = 0$ и положить $x = 0$, то можно получить $-T_n$. Давайте так и сделаем:

$$\begin{aligned} T_n &= \left(\frac{1}{x}-\frac{n!}{x(x+1)\ldots(x+n)}\right)\bigg|_{x=0} = \\ &= \left(\frac{(x+1)\ldots(x+n)-n!}{x(x+1)\ldots(x+n)}\right)\bigg|_{x=0} = \\ &= \left(\frac{x^n{n+1 \brack n+1}+\cdots+x{n+1 \brack 2}+{n+1 \brack 1}-n!}{x(x+1)\ldots(x+n)}\right)\bigg|_{x=0} = \frac{1}{n!}{n+1 \brack 2}\,. \end{aligned}$$

(Мы использовали разложение (6.11) для выражения $(x+1)\ldots(x+n) = x^{\overline{n+1}}/x$; а в числителе можно выделить x потому, что

$\left[{n+1 \atop 1}\right] = n!$.) Но из (6.58) нам известно, что $\left[{n+1 \atop 2}\right] = n!\,H_n$; следовательно, $T_n = H_n$, и мы получаем ответ:

$$U_n \;=\; n^n(H_n - 1)\,. \qquad (6.73)$$

Это один подход. Другой подход — попробовать вычислить значительно более общую сумму

$$U_n(x,y) \;=\; \sum_{k\geqslant 1}\binom{n}{k}\frac{(-1)^{k-1}}{k}(x+ky)^n\,, \quad \text{целое } n \geqslant 0; \qquad (6.74)$$

величина исходной суммы U_n будет тогда представлять ее частный случай $U_n(n,-1)$. (К приданию большей общности нас подталкивает то, что предыдущий вывод “отбросил” бóльшую часть деталей данной задачи; так или иначе те детали не должны иметь отношения к делу, ибо n-я разность их ликвидирует.)

Можно было бы воспроизвести предыдущий вывод с небольшими изменениями и выяснить величину $U_n(x,y)$. Можно также заменить $(x+ky)^n$ на $(x+ky)^{n-1}(x+ky)$, а затем заменить $\binom{n}{k}$ на $\binom{n-1}{k} + \binom{n-1}{k-1}$, что приводит к рекуррентному соотношению

$$U_n(x,y) \;=\; xU_{n-1}(x,y) + x^n/n + yx^{n-1}\,; \qquad (6.75)$$

которое благополучно решается с помощью суммирующего множителя (упр. 5).

Но проще воспользоваться другим приемом, который сослужил нам добрую службу в главе 2: дифференцированием. Производная $U_n(x,y)$ по y привносит k, которое сокращается с k в знаменателе, и мы получаем тривиальную сумму:

$$\begin{aligned}\frac{\partial}{\partial y}U_n(x,y) \;&=\; \sum_{k\geqslant 1}\binom{n}{k}(-1)^{k-1}n(x+ky)^{n-1} \;=\\ &=\; \binom{n}{0}nx^{n-1} - \sum_{k\geqslant 0}\binom{n}{k}(-1)^k n(x+ky)^{n-1} \;=\; nx^{n-1}\,.\end{aligned}$$

(И снова n-я разность полинома степени $< n$ обращается в нуль.)

Мы доказали, что производная $U_n(x,y)$ по y равна nx^{n-1} независимо от y. В общем случае, если $f'(y) = c$, то $f(y) = f(0) + cy$; таким образом, мы должны получить $U_n(x,y) = U_n(x,0) + nx^{n-1}y$.

Остается последнее — найти $U_n(x,0)$. Но $U_n(x,0)$ — это просто x^n, умноженное на сумму $T_n = H_n$, которую мы уже рассматривали в (6.72); таким образом, общая сумма (6.74) имеет аналитический вид

$$U_n(x,y) \;=\; x^n H_n + n x^{n-1} y\,. \tag{6.76}$$

В частности, решением исходной задачи является $U_n(n,-1) = n^n(H_n - 1)$.

6.5 Числа Бернулли

Очередная важная последовательность чисел в нашей повестке дня носит имя Якоба Бернулли (Jakob Bernoulli) (1654–1705), открывшего любопытные соотношения при работе с формулами для сумм m-х степеней [26]. Положим

$$S_m(n) \;=\; 0^m + 1^m + \cdots + (n-1)^m \;=\; \sum_{k=0}^{n-1} k^m \;=\; \sum\nolimits_0^n x^m\,\delta x\,. \tag{6.77}$$

(Таким образом, $S_m(n) = H_{n-1}^{(-m)}$ при $m > 0$ при использовании обозначений для обобщенных гармонических чисел.) Рассматривая ряд приведенных ниже формул, Бернулли обнаружил в них некоторую закономерность:

$$\begin{aligned}
S_0(n) &= n\\
S_1(n) &= \tfrac12 n^2 - \tfrac12 n\\
S_2(n) &= \tfrac13 n^3 - \tfrac12 n^2 + \tfrac16 n\\
S_3(n) &= \tfrac14 n^4 - \tfrac12 n^3 + \tfrac14 n^2\\
S_4(n) &= \tfrac15 n^5 - \tfrac12 n^4 + \tfrac13 n^3 - \tfrac1{30} n\\
S_5(n) &= \tfrac16 n^6 - \tfrac12 n^5 + \tfrac5{12} n^4 - \tfrac1{12} n^2\\
S_6(n) &= \tfrac17 n^7 - \tfrac12 n^6 + \tfrac12 n^5 - \tfrac16 n^3 + \tfrac1{42} n\\
S_7(n) &= \tfrac18 n^8 - \tfrac12 n^7 + \tfrac7{12} n^6 - \tfrac7{24} n^4 + \tfrac1{12} n^2\\
S_8(n) &= \tfrac19 n^9 - \tfrac12 n^8 + \tfrac23 n^7 - \tfrac7{15} n^5 + \tfrac29 n^3 - \tfrac1{30} n\\
S_9(n) &= \tfrac1{10} n^{10} - \tfrac12 n^9 + \tfrac34 n^8 - \tfrac7{10} n^6 + \tfrac12 n^4 - \tfrac3{20} n^2\\
S_{10}(n) &= \tfrac1{11} n^{11} - \tfrac12 n^{10} + \tfrac56 n^9 - n^7 + n^5 - \tfrac12 n^3 + \tfrac5{66} n
\end{aligned}$$

А вы ее не видите? Коэффициент при n^{m+1} в $S_m(n)$ всегда равен $1/(m+1)$. Коэффициент при n^m всегда равен $-1/2$. Коэффициент при n^{m-1} всегда равен… посмотрим… да, $m/12$. Коэффициент при n^{m-2} — всегда нуль. Коэффициент при n^{m-3} всегда… посмотрим… ну-ка, ну-ка… да, конечно, это $-m(m-1)\times(m-2)/720$. Коэффициент при n^{m-4} всегда равен нулю. Похоже, что если эту закономерность продолжить, то коэффициент при n^{m-k} всегда будет представлять собой некоторую константу, умноженную на $m^{\underline{k}}$.

Именно это эмпирически нашел Бернулли. (Доказательства он не дал.) В современной записи мы записываем эти коэффициенты в виде

$$\begin{aligned} S_m(n) &= \frac{1}{m+1}\left(B_0 n^{m+1} + \binom{m+1}{1} B_1 n^m + \cdots + \binom{m+1}{m} B_m n\right) = \\ &= \frac{1}{m+1}\sum_{k=0}^{m}\binom{m+1}{k} B_k n^{m+1-k}. \end{aligned} \tag{6.78}$$

Числа Бернулли определяются неявным рекуррентным соотношением

$$\sum_{j=0}^{m}\binom{m+1}{j} B_j = [m=0] \qquad \text{для всех } m \geqslant 0. \tag{6.79}$$

Например, $\binom{2}{0}B_0 + \binom{2}{1}B_1 = 0$. Несколькими первыми числами Бернулли являются

n	0	1	2	3	4	5	6	7	8	9	10	11	12
B_n	1	$-\frac{1}{2}$	$\frac{1}{6}$	0	$\frac{-1}{30}$	0	$\frac{1}{42}$	0	$\frac{-1}{30}$	0	$\frac{5}{66}$	0	$\frac{-691}{2730}$

(Все надежды на простой аналитический вид B_n улетучиваются при появлении странной дроби $-691/2730$.)

Формулу Бернулли (6.78) можно доказать по индукции по m, с помощью метода приведения (один из способов, применяя который, мы нашли в главе 2 сумму $S_2(n) = \Box_n$):

$$\begin{aligned} S_{m+1}(n) + n^{m+1} &= \sum_{k=0}^{n-1}(k+1)^{m+1} = \\ &= \sum_{k=0}^{n-1}\sum_{j=0}^{m+1}\binom{m+1}{j}k^j = \\ &= \sum_{j=0}^{m+1}\binom{m+1}{j} S_j(n). \end{aligned} \tag{6.80}$$

Обозначим через $\widehat{S}_m(n)$ правую часть (6.78); мы хотим показать, что $S_m(n) = \widehat{S}_m(n)$ в предположении, что $S_j(n) = \widehat{S}_j(n)$ при $0 \leqslant j < m$. Начнем так, как мы делали в главе 2 для $m = 2$, вычитая $S_{m+1}(n)$ из обеих частей (6.80). Затем распишем каждое $S_j(n)$ с использованием (6.78) и перегруппируем члены так, чтобы коэффициенты при степенях n в правой части оказались рядом и упростились:

$$\begin{aligned}
n^{m+1} &= \sum_{j=0}^{m}\binom{m+1}{j}S_j(n) = \sum_{j=0}^{m}\binom{m+1}{j}\widehat{S}_j(n)+\binom{m+1}{m}\Delta = \\
&= \sum_{j=0}^{m}\binom{m+1}{j}\frac{1}{j+1}\sum_{k=0}^{j}\binom{j+1}{k}B_k n^{j+1-k}+(m+1)\Delta = \\
&= \sum_{0\leqslant k\leqslant j\leqslant m}\binom{m+1}{j}\binom{j+1}{k}\frac{B_k}{j+1}n^{j+1-k}+(m+1)\Delta = \\
&= \sum_{0\leqslant k\leqslant j\leqslant m}\binom{m+1}{j}\binom{j+1}{j-k}\frac{B_{j-k}}{j+1}n^{k+1}+(m+1)\Delta = \\
&= \sum_{0\leqslant k\leqslant j\leqslant m}\binom{m+1}{j}\binom{j+1}{k+1}\frac{B_{j-k}}{j+1}n^{k+1}+(m+1)\Delta = \\
&= \sum_{0\leqslant k\leqslant m}\frac{n^{k+1}}{k+1}\sum_{k\leqslant j\leqslant m}\binom{m+1}{j}\binom{j}{k}B_{j-k}+(m+1)\Delta = \\
&= \sum_{0\leqslant k\leqslant m}\frac{n^{k+1}}{k+1}\binom{m+1}{k}\sum_{k\leqslant j\leqslant m}\binom{m+1-k}{j-k}B_{j-k}+(m+1)\Delta = \\
&= \sum_{0\leqslant k\leqslant m}\frac{n^{k+1}}{k+1}\binom{m+1}{k}\sum_{0\leqslant j\leqslant m-k}\binom{m+1-k}{j}B_j+(m+1)\Delta = \\
&= \sum_{0\leqslant k\leqslant m}\frac{n^{k+1}}{k+1}\binom{m+1}{k}[m-k=0]+(m+1)\Delta = \\
&= \frac{n^{m+1}}{m+1}\binom{m+1}{m}+(m+1)\Delta = \\
&= n^{m+1}+(m+1)\Delta, \qquad \text{где } \Delta = S_m(n)-\widehat{S}_m(n).
\end{aligned}$$

(Этот вывод представляет собой хороший обзор стандартных методов, изученных в главе 5.) Таким образом, $\Delta = 0$ и $S_m(n) = \widehat{S}_m(n)$. QED

В главе 7, применяя производящие функции, мы докажем (6.78) гораздо проще. Ключевая идея этого доказательства в том, чтобы показать, что числа Бернулли представляют собой коэффициенты степенного ряда

$$\frac{z}{e^z-1} = \sum_{n\geqslant 0} B_n\frac{z^n}{n!}. \tag{6.81}$$

Далее речь пойдет о несколько более тонких материях, с которыми вы, вероятно, на первый раз предпочтете просто бегло ознакомиться.
— Сочувствующий ассистент

Начало ознакомления

Пока же давайте просто считать, что (6.81) действительно выполняется, так что мы можем вывести из него некоторые удивительные следствия. Если к обеим частям прибавить $\frac{1}{2}z$, избавляясь тем самым от члена $B_1z/1! = -\frac{1}{2}z$ в правой части, можно получить

$$\frac{z}{e^z-1}+\frac{z}{2} = \frac{z}{2}\,\frac{e^z+1}{e^z-1} = \frac{z}{2}\,\frac{e^{z/2}+e^{-z/2}}{e^{z/2}-e^{-z/2}} = \frac{z}{2}\operatorname{cth}\frac{z}{2}. \qquad (6.82)$$

Здесь cth означает "гиперболический котангенс", функцию, известную также как отношение $\operatorname{ch} z/\operatorname{sh} z$, где

$$\operatorname{sh} z = \frac{e^z-e^{-z}}{2}; \qquad \operatorname{ch} z = \frac{e^z+e^{-z}}{2}. \qquad (6.83)$$

Замена z на $-z$ дает $\left(\frac{-z}{2}\right)\operatorname{cth}\left(\frac{-z}{2}\right) = \frac{z}{2}\operatorname{cth}\frac{z}{2}$; следовательно, каждый коэффициент с нечетным номером в разложении $\frac{z}{2}\operatorname{cth}\frac{z}{2}$ должен быть нулевым, и мы имеем

$$B_3 = B_5 = B_7 = B_9 = B_{11} = B_{13} = \cdots = 0. \qquad (6.84)$$

Кроме того, равенство (6.82) приводит к аналитическому виду для коэффициентов разложения cth:

$$z\operatorname{cth} z = \frac{2z}{e^{2z}-1}+\frac{2z}{2} = \sum_{n\geqslant 0} B_{2n}\frac{(2z)^{2n}}{(2n)!} = \\ = \sum_{n\geqslant 0} 4^n B_{2n}\frac{z^{2n}}{(2n)!}. \qquad (6.85)$$

Однако особым спросом на рынке функций гиперболические функции не пользуются; более популярны "реальные" тригонометрические функции. Обычные тригонометрические функции можно выразить через их гиперболические аналоги при помощи правил

$$\sin z = -i\operatorname{sh} iz, \qquad \cos z = \operatorname{ch} iz; \qquad (6.86)$$

соответствующие степенные ряды представляют собой

$$\sin z = \frac{z^1}{1!}-\frac{z^3}{3!}+\frac{z^5}{5!}-\cdots, \qquad \operatorname{sh} z = \frac{z^1}{1!}+\frac{z^3}{3!}+\frac{z^5}{5!}+\cdots;$$

$$\cos z = \frac{z^0}{0!}-\frac{z^2}{2!}+\frac{z^4}{4!}-\cdots, \qquad \operatorname{ch} z = \frac{z^0}{0!}+\frac{z^2}{2!}+\frac{z^4}{4!}+\cdots.$$

"Действительные" функции получаются при использовании мнимых чисел...

Следовательно, $\operatorname{ctg} z = \cos z/\sin z = i\operatorname{ch} iz/\operatorname{sh} iz = i\operatorname{cth} iz$, и мы получаем

$$z\operatorname{ctg} z = \sum_{n\geqslant 0} B_{2n}\frac{(2iz)^{2n}}{(2n)!} = \sum_{n\geqslant 0}(-4)^n B_{2n}\frac{z^{2n}}{(2n)!}. \qquad (6.87)$$

Другая замечательная формула для $z\operatorname{ctg} z$ была найдена Эйлером (см. упр. 73):

$$z\operatorname{ctg} z = 1-2\sum_{k\geqslant 1}\frac{z^2}{k^2\pi^2-z^2}. \qquad (6.88)$$

Формулу Эйлера можно разложить по степеням z^2, получив

$$z\operatorname{ctg} z = 1 - 2\sum_{k\geqslant 1}\left(\frac{z^2}{k^2\pi^2} + \frac{z^4}{k^4\pi^4} + \frac{z^6}{k^6\pi^6} + \cdots\right) =$$
$$= 1 - 2\left(\frac{z^2}{\pi^2}H_\infty^{(2)} + \frac{z^4}{\pi^4}H_\infty^{(4)} + \frac{z^6}{\pi^6}H_\infty^{(6)} + \cdots\right).$$

Приравнивая коэффициенты при z^{2n} с соответствующими коэффициентами в другой формуле, (6.87), получаем почти волшебное аналитическое выражение для бесконечного количества бесконечных сумм:

$$\zeta(2n) = H_\infty^{(2n)} = (-1)^{n-1}\frac{2^{2n-1}\pi^{2n}B_{2n}}{(2n)!}, \quad \text{целое } n > 0. \tag{6.89}$$

Например,

$$\zeta(2) = H_\infty^{(2)} = 1+\tfrac{1}{4}+\tfrac{1}{9}+\cdots = \pi^2 B_2 = \pi^2/6; \tag{6.90}$$
$$\zeta(4) = H_\infty^{(4)} = 1+\tfrac{1}{16}+\tfrac{1}{81}+\cdots = -\pi^4 B_4/3 = \pi^4/90. \tag{6.91}$$

Формула (6.89) не только представляет собой аналитический вид для $H_\infty^{(2n)}$, но и дает информацию о приблизительной величине чисел B_{2n}, поскольку $H_\infty^{(2n)}$ очень близко к 1 при больших n. А еще она говорит о том, что $(-1)^{n-1}B_{2n} > 0$ для всех $n > 0$; таким образом, отличные от нуля числа Бернулли являются знакочередующимися.

И это еще не все. Числа Бернулли появляются и в коэффициентах разложения тангенса,

$$\operatorname{tg} z = \frac{\sin z}{\cos z} = \sum_{n\geqslant 0}(-1)^{n-1}4^n(4^n-1)B_{2n}\frac{z^{2n-1}}{(2n)!}, \tag{6.92}$$

Отсюда можно вообще не читать

а также других тригонометрических функций (упр. 72). Формула (6.92) приводит к другому важному факту о числах Бернулли, а именно, что

$$T_{2n-1} = (-1)^{n-1}\frac{4^n(4^n-1)}{2n}B_{2n} \quad \text{— положительное целое.} \tag{6.93}$$

Например,

n	1	3	5	7	9	11	13
T_n	1	2	16	272	7936	353792	22368256

(Числа T называются *тангенциальными числами.*)

Один из способов доказательства утверждения (6.93), следуя идее Б. Ф. Логана (B. F. Logan), состоит в рассмотрении степенного ряда

$$\frac{\sin z+x\cos z}{\cos z-x\sin z} = x+(1+x^2)z+(2x^3+2x)\frac{z^2}{2}+(6x^4+8x^2+2)\frac{z^3}{6}+\cdots = \\ = \sum_{n\geqslant 0} T_n(x)\frac{z^n}{n!}, \qquad (6.94)$$

Когда $x = \operatorname{tg} w$, это дает $\operatorname{tg}(z+w)$. Следовательно, по теореме Тейлора, n-я производная $\operatorname{tg} w$ равна $T_n(\operatorname{tg} w)$.

где $T_n(x)$ — полином от x. Подстановка $x = 0$ дает $T_n(0) = T_n$, n-е тангенциальное число. Если продифференцировать (6.94) по x, можно получить

$$\frac{1}{(\cos z - x\sin z)^2} = \sum_{n\geqslant 0} T_n'(x)\frac{z^n}{n!};$$

но если продифференцировать по z, то будет получено

$$\frac{1+x^2}{(\cos z - x\sin z)^2} = \sum_{n\geqslant 1} T_n(x)\frac{z^{n-1}}{(n-1)!} = \sum_{n\geqslant 0} T_{n+1}(x)\frac{z^n}{n!}.$$

(Попробуйте сами — здесь получается очень красивое сокращение.) Таким образом, мы имеем

$$T_{n+1}(x) = (1+x^2)T_n'(x), \qquad T_0(x) = x, \qquad (6.95)$$

простое рекуррентное соотношение, из которого следует, что коэффициенты $T_n(x)$ — неотрицательные целые числа. Кроме того, можно легко доказать, что полином $T_n(x)$ имеет степень $n+1$ и что его коэффициенты попеременно нулевые и положительные. Поэтому $T_{2n+1}(0) = T_{2n+1}$ являются положительными целыми числами, что и утверждалось ранее в (6.93).

Рекуррентное соотношение (6.95) предоставляет нам простой способ вычисления чисел Бернулли через тангенциальные числа с использованием только простых операций над целыми числами; определяющее же числа Бернулли рекуррентное соотношение (6.79), напротив, требует сложных вычислений с дробями.

Вычислить сумму n-х степеней от a до $b-1$, а не от 0 до $n-1$, как гласит теория из главы 2, можно следующим образом:

$$\sum_{k=a}^{b-1} k^m = \sum\nolimits_a^b x^m\,\delta x = S_m(b) - S_m(a). \qquad (6.96)$$

Это тождество имеет интересные следствия, если рассматривать отрицательные значения k. Имеем

$$\sum_{k=-n+1}^{-1} k^m = (-1)^m \sum_{k=0}^{n-1} k^m, \qquad \text{когда } m > 0,$$

следовательно,

$$S_m(0) - S_m(-n+1) \;=\; (-1)^m \bigl(S_m(n) - S_m(0)\bigr)\,.$$

Но поскольку $S_m(0) = 0$, мы получаем соотношение

$$S_m(1-n) \;=\; (-1)^{m+1} S_m(n)\,, \qquad m > 0. \tag{6.97}$$

Иоганн Фаулхабер (Johann Faulhaber) неявно использовал (6.97) в 1631 году [119] при поиске простых формул, выражающих $S_m(n)$ в виде полиномов от $n(n+1)/2$ при $m \leqslant 17$; см. [222].

Таким образом, $S_m(1) = 0$. Если мы разложим полином $S_m(n)$ на множители, то среди них всегда будут множители n и $(n-1)$, потому что этот полином имеет корни 0 и 1. В общем случае $S_m(n)$ представляет собой полином степени $m+1$ со старшим членом $\frac{1}{m+1}n^{m+1}$. Более того, можно подставить $n = \frac{1}{2}$ в (6.97) и получить, что $S_m(\frac{1}{2}) = (-1)^{m+1} S_m(\frac{1}{2})$; если m четное, то $S_m(\frac{1}{2}) = 0$, так что еще одним множителем будет $(n - \frac{1}{2})$. Эти наблюдения поясняют, почему в главе 2 нам удалось найти такое простое разложение, как

$$S_2(n) \;=\; \tfrac{1}{3} n(n-\tfrac{1}{2})(n-1);$$

там можно было воспользоваться подобным рассуждением для выяснения величины $S_2(n)$ без ее вычисления! Кроме того, из (6.97) вытекает, что полином с остальными множителями, $\hat{S}_m(n) = S_m(n)/(n-\frac{1}{2})$, всегда удовлетворяет соотношению

$$\hat{S}_m(1-n) \;=\; \hat{S}_m(n)\,, \qquad m \text{ четное}, \quad m > 0.$$

Отсюда следует, что $S_m(n)$ всегда можно записать в виде произведения

$$S_m(n) \;=\; \begin{cases} \dfrac{1}{m+1} \displaystyle\prod_{k=1}^{\lceil m/2 \rceil} (n-\tfrac{1}{2}-\alpha_k)(n-\tfrac{1}{2}+\alpha_k)\,, & m \text{ нечетное;} \\[2ex] \dfrac{(n-\frac{1}{2})}{m+1} \displaystyle\prod_{k=1}^{m/2} (n-\tfrac{1}{2}-\alpha_k)(n-\tfrac{1}{2}+\alpha_k)\,, & m \text{ четное.} \end{cases} \tag{6.98}$$

Здесь $\alpha_1 = \frac{1}{2}$, а $\alpha_2, \ldots, \alpha_{\lceil m/2 \rceil}$ — соответствующие комплексные числа, значения которых зависят от m. Например,

$$\begin{aligned} S_3(n) &= n^2(n-1)^2/4\,; \\ S_4(n) &= n(n-\tfrac{1}{2})(n-1)(n-\tfrac{1}{2}+\sqrt{7/12}\,)(n-\tfrac{1}{2}-\sqrt{7/12}\,)/5\,; \\ S_5(n) &= n^2(n-1)^2(n-\tfrac{1}{2}+\sqrt{3/4}\,)(n-\tfrac{1}{2}-\sqrt{3/4}\,)/6\,; \\ S_6(n) &= n(n-\tfrac{1}{2})(n-1)(n-\tfrac{1}{2}+\alpha)(n-\tfrac{1}{2}-\alpha)(n-\tfrac{1}{2}+\overline{\alpha})(n-\tfrac{1}{2}-\overline{\alpha})/7\,, \\ &\qquad \text{где } \alpha = 2^{-3/2}\,3^{-1/4}\bigl(\sqrt{\sqrt{31}+\sqrt{27}} + i\,\sqrt{\sqrt{31}-\sqrt{27}}\,\bigr). \end{aligned}$$

Конец пропускаемого текста

Если m нечетное и больше, чем 1, то $B_m = 0$; следовательно, $S_m(n)$ делится на n^2 (и на $(n-1)^2$). В противном случае корни полинома $S_m(n)$, по-видимому, не должны подчиняться какому-то простому правилу.

Завершим изучение чисел Бернулли выяснением того, как они связаны с числами Стирлинга. Один из способов вычисления суммы $S_m(n)$ состоит в замене обычных степеней убывающими, поскольку убывающие степени просто суммируются. После вычисления подобных простых сумм можно опять вернуться к обычным степеням:

$$\begin{aligned} S_m(n) = \sum_{k=0}^{n-1} k^m &= \sum_{k=0}^{n-1} \sum_{j\geqslant 0} {m \brace j} k^{\underline{j}} = \sum_{j\geqslant 0} {m \brace j} \sum_{k=0}^{n-1} k^{\underline{j}} = \\ &= \sum_{j\geqslant 0} {m \brace j} \frac{n^{\underline{j+1}}}{j+1} = \\ &= \sum_{j\geqslant 0} {m \brace j} \frac{1}{j+1} \sum_{k\geqslant 0} (-1)^{j+1-k} {j+1 \brack k} n^k . \end{aligned}$$

Таким образом, приравнивая коэффициенты к соответствующим коэффициентам в (6.78), мы должны получить тождество

$$\sum_{j\geqslant 0} {m \brace j} {j+1 \brack k} \frac{(-1)^{j+1-k}}{j+1} = \frac{1}{m+1} \binom{m+1}{k} B_{m+1-k}, \qquad k > 0. \tag{6.99}$$

Было бы неплохо доказать это соотношение непосредственно, получая тем самым другой способ определения чисел Бернулли. Однако ни правила табл. 325, ни правила табл. 326 не дают нам сколь-нибудь явного повода рассчитывать на доказательство по индукции, что сумма слева в (6.99) представляет собой константу, умноженную на $m^{\underline{k-1}}$. Если $k = m+1$, то сумма в левой части равна ${m \brace m}{m+1 \brack m+1}/(m+1) = 1/(m+1)$, так что в этом случае все просто. А если $k = m$, то сумма в левой части сводится к ${m \brace m-1}{m \brack m}m^{-1} - {m \brace m}{m+1 \brack m}(m+1)^{-1} = \frac{1}{2}(m-1) - \frac{1}{2}m = -\frac{1}{2}$, так что и этот случай достаточно прост. Но при $k < m$ сумма в левой части выглядит ужасно. Вряд ли Бернулли открыл бы свои числа, пойдя таким путем.

Единственное, что можно сделать,— это заменить ${m \brace j}$ на ${m+1 \brace j+1} - (j+1){m \brace j+1}$. Тогда $(j+1)$ отлично сокращается с мешающим нам знаменателем и левая часть превращается в

$$\sum_{j\geqslant 0} {m+1 \brace j+1} {j+1 \brack k} \frac{(-1)^{j+1-k}}{j+1} - \sum_{j\geqslant 0} {m \brace j+1} {j+1 \brack k} (-1)^{j+1-k} .$$

При $k < m$ вторая сумма равна нулю в соответствии с (6.31). Это оставляет нас наедине с первой суммой, которая требует смены декораций: переименуем все переменные так, чтобы индексом суммирования было k и чтобы остальными параметрами были m и n. Тогда тождество (6.99) будет эквивалентно тождеству

$$\sum_k \left\{ n \atop k \right\} \left[k \atop m \right] \frac{(-1)^{k-m}}{k} = \frac{1}{n} \binom{n}{m} B_{n-m} + [m=n-1], \quad m > 0. \tag{6.100}$$

Неплохо — мы получили нечто более привлекательное, хотя табл. 326 все еще ничего не подсказывает насчет очередного шага.

Теперь на помощь приходят формулы сверток из табл. 334: можно воспользоваться формулами (6.53) и (6.52) для того, чтобы выразить общий член суммы через полиномы Стирлинга:

$$\left\{ n \atop k \right\} \left[k \atop m \right] = (-1)^{n-k+1} \frac{n!}{(k-1)!} \sigma_{n-k}(-k) \cdot \frac{k!}{(m-1)!} \sigma_{k-m}(k);$$

$$\left\{ n \atop k \right\} \left[k \atop m \right] \frac{(-1)^{k-m}}{k} = (-1)^{n+1-m} \frac{n!}{(m-1)!} \sigma_{n-k}(-k)\, \sigma_{k-m}(k).$$

Ситуация улучшается; свертка в (6.50) при $t = 1$ дает

$$\sum_{k=0}^{n} \sigma_{n-k}(-k)\, \sigma_{k-m}(k) = \sum_{k=0}^{n-m} \sigma_{n-m-k}\big(-n+(n-m-k)\big)\, \sigma_k(m+k) = \\ = \frac{m-n}{(m)(-n)} \sigma_{n-m}\big(m-n+(n-m)\big).$$

Формула (6.100) проверена, и мы установили, что числа Бернулли связаны с постоянными членами полиномов Стирлинга:

Конец беглого ознакомления

$$\frac{B_m}{m!} = -m\sigma_m(0). \tag{6.101}$$

6.6 Числа Фибоначчи

Теперь перейдем к особой последовательности чисел — быть может, самой приятной из всех — последовательности чисел Фибоначчи $\langle F_n \rangle$:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
F_n	0	1	1	2	3	5	8	13	21	34	55	89	144	233	377

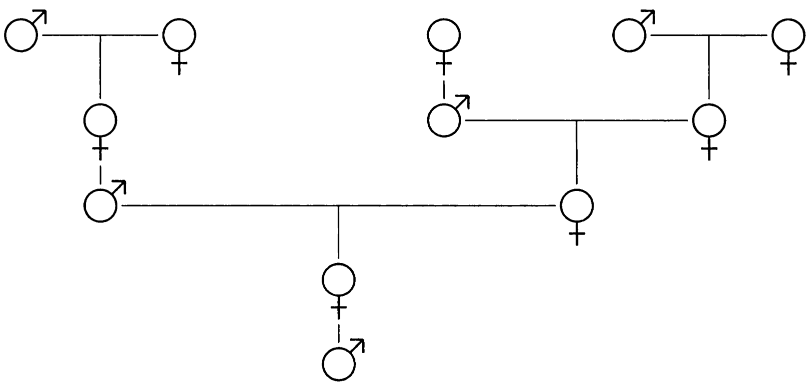
В отличие от гармонических чисел и чисел Бернулли, числа Фибоначчи представляют собой обычные бесхитростные целые числа, определяемые рекуррентным соотношением

$$\begin{aligned} F_0 &= 0; \\ F_1 &= 1; \\ F_n &= F_{n-1} + F_{n-2} \qquad \text{при } n > 1. \end{aligned} \tag{6.102}$$

Простота этого правила — простейшего из всех рекуррентных соотношений, в которых каждое число зависит от предыдущих двух — служит пояснением того, почему числа Фибоначчи так часто встречаются в самых разных ситуациях.

Противоестественная природа этого примера возмутительна! Комиссии по защите общественной морали следует немедленно запретить эту книгу!

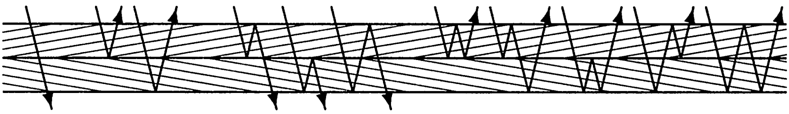
Наглядным примером естественного возникновения чисел Фибоначчи являются "пчелиные деревья". Давайте рассмотрим родословную самца пчелы. Каждый самец пчелы (именуемый трутнем) появляется на свет непарным путем от самки (именуемой маткой); но каждая самка имеет двух родителей — самца и самку. Вот несколько первых уровней такого дерева:



У трутня один дед и одна бабка, один прадед и две прабабки, два прапрадеда и три прапрабабки. В общем случае, как легко установить по индукции, у него ровно F_{n+1} пра^n-дедушек и F_{n+2} пра^n-бабушек.

Числа Фибоначчи часто обнаруживаются в природе — возможно, по причинам, аналогичным закону образования родословного дерева пчел. Например, в цветке типичного подсолнуха плотно упакованные семечки располагаются по спиралям; обычно это 34 спирали, закручивающиеся в одном направлении, и 55 — в другом. Меньшие цветки содержат 21 и 34 спирали, или 13 и 21 спираль; в Англии однажды демонстрировался гигантский подсолнух с 89 и 144 спиралями. Аналогичные закономерности обнаруживаются и в некоторых видах сосновых шишек.

Вот пример другого рода [277]. Предположим, что одна на другую наложены две стеклянные пластинки. Сколько существует способов a_n прохождения луча света через пластинки или отражения от них после изменения его направления n раз? Вот несколько первых случаев:



$a_0 = 1$ $a_1 = 2$ $a_2 = 3$ $a_3 = 5$

Когда n четное, получается четное число преломлений и луч проходит сквозь пластинки; когда n нечетное, луч отражается и выходит с той же стороны, с которой вошел. Похоже, что a_n представляют собой числа Фибоначчи, и непродолжительное разглядывание рисунка показывает, почему: при $n \geqslant 2$ преломляющиеся n раз лучи либо претерпевают первое отражение от внешней поверхности и продолжают прохождение a_{n-1} способами, либо начинают с отражения от внутренней поверхности и затем снова отражаются в обратном направлении, чтобы закончить a_{n-2} способами. Таким образом, мы получаем рекуррентное соотношение Фибоначчи $a_n = a_{n-1} + a_{n-2}$. Начальные условия отличаются, но не слишком сильно, потому что $a_0 = 1 = F_2$ и $a_1 = 2 = F_3$; так что все просто сдвигается на две позиции, и $a_n = F_{n+2}$.

Эти числа были введены Леонардо Фибоначчи (Leonardo Fibonacci) в 1202 году, и математики стали выявлять все новые и новые интересные факты об этих числах. Эдуард Люка (Edouard Lucas), причастный к головоломке о Ханойской башне (которая рассматривалась нами в главе 1), активно работал над ними во второй половине XIX века (кстати, именно с легкой руки Люка эти числа стали именоваться "числами Фибоначчи"). Одним из удивительных результатов, полученных Люка, стало применение свойств чисел Фибоначчи для доказательства того, что 39-значное число Мерсенна $2^{127} - 1$ является простым.

"La suite de Fibonacci possède des propriétés nombreuses fort intéressantes".
— Э. Люка [259]

Одной их старейших теорем, связанных с числами Фибоначчи, было тождество, впервые опубликованное итальянским астрономом Джованни Доменико Кассини (Gian Domenico Cassini) в 1680 году [51]:

$$F_{n+1} F_{n-1} - F_n^2 = (-1)^n \qquad \text{при } n > 0. \tag{6.103}$$

При $n = 6$, например, тождество Кассини совершенно справедливо утверждает, что $13 \cdot 5 - 8^2$ равно 1. (Иоганну Кеплеру (Johannes Kepler) это тождество было знакомо еще с 1608 года [202].)

Полиномиальная формула, которая включает числа Фибоначчи вида $F_{n \pm k}$ при малых значениях k, может быть преобразована в формулу, которая включает в себя только F_n и F_{n+1}, потому что можно воспользоваться правилом

$$F_m = F_{m+2} - F_{m+1} \tag{6.104}$$

для выражения F_m через бо́льшие числа Фибоначчи при $m < n$ и формулой

$$F_m = F_{m-2} + F_{m-1} \tag{6.105}$$

для замены F_m меньшими числами Фибоначчи при $m > n+1$. Так, например, можно заменить F_{n-1} на $F_{n+1} - F_n$ в (6.103), чтобы получить тождество Кассини в виде

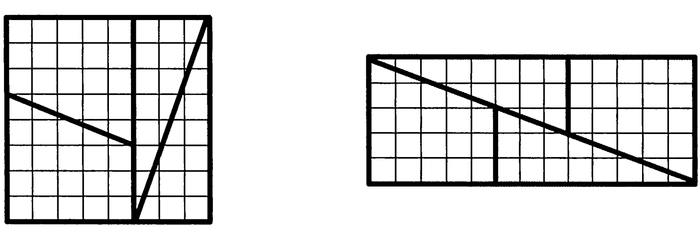
$$F_{n+1}^2 - F_{n+1}F_n - F_n^2 = (-1)^n . \tag{6.106}$$

Кроме того, если заменить n на $n+1$, тождество Кассини примет вид

$$F_{n+2}F_n - F_{n+1}^2 = (-1)^{n+1};$$

это то же самое, что и $(F_{n+1}+F_n)F_n - F_{n+1}^2 = (-1)^{n+1}$, что, в свою очередь, совпадает с (6.106). Таким образом, "Кассини(n)" истинно тогда и только тогда, когда истинно "Кассини($n+1$)", так что по индукции (6.103) выполняется для всех n.

Тождество Кассини лежит в основе геометрического парадокса, который был одной из любимых головоломок Льюиса Кэррола (Lewis Carroll) ([63], [319], [364]). Суть его в том, чтобы разрезать шахматную доску на части, из которых затем составить прямоугольник:



Исходный квадрат из $8\times 8 = 64$ клеток превращается в прямоугольник из $5 \times 13 = 65$ клеток! Аналогичное построение разбивает любой квадрат размером $F_n \times F_n$ на четыре части с размерами сторон F_{n+1}, F_n, F_{n-1} и F_{n-2} вместо соответственно 13, 8, 5 и 3 в нашем примере. В результате получается прямоугольник размером $F_{n-1} \times F_{n+1}$; в соответствии с (6.103) одна клетка либо прибавляется, либо утрачивается — в зависимости от того, четное или нечетное n.

Парадокс объясняется тем, что... впрочем, кто же будет объяснять волшебство?

Строго говоря, правило (6.105) можно применять только при $m \geqslant 2$, потому что мы не определили F_n для отрицательных n.

Бо́льшую свободу действий можно получить, если избавиться от этого условия и использовать (6.104) и (6.105) для определения чисел Фибоначчи с отрицательными индексами. Например, F_{-1} оказывается равным $F_1 - F_0 = 1$, а F_{-2} — равным $F_0 - F_{-1} = -1$. Таким способом мы выводим значения

n	0	−1	−2	−3	−4	−5	−6	−7	−8	−9	−10	−11
F_n	0	1	−1	2	−3	5	−8	13	−21	34	−55	89

и вскоре становится ясно (по индукции), что

$$F_{-n} = (-1)^{n-1}F_n\,, \qquad \text{целое } n. \tag{6.107}$$

Тождество Кассини (6.103) при таком обобщении последовательности Фибоначчи становится справедливым при *любых* целых n, а не только при $n > 0$.

Процесс сведения $F_{n\pm k}$ к комбинации F_n и F_{n+1} с применением правил (6.105) и (6.104) приводит к последовательности формул

$$\begin{aligned}
F_{n+2} &= F_{n+1} + F_n & \qquad F_{n-1} &= F_{n+1} - F_n\\
F_{n+3} &= 2F_{n+1} + F_n & F_{n-2} &= -F_{n+1} + 2F_n\\
F_{n+4} &= 3F_{n+1} + 2F_n & F_{n-3} &= 2F_{n+1} - 3F_n\\
F_{n+5} &= 5F_{n+1} + 3F_n & F_{n-4} &= -3F_{n+1} + 5F_n
\end{aligned}$$

из которой становится очевидной еще одна закономерность:

$$F_{n+k} = F_k F_{n+1} + F_{k-1} F_n\,. \tag{6.108}$$

Это легко доказываемое по индукции тождество справедливо при любых целых k и n (положительных, отрицательных или равных нулю).

Если в (6.108) положить $k = n$, то можно получить, что

$$F_{2n} = F_n F_{n+1} + F_{n-1} F_n\,; \tag{6.109}$$

следовательно, F_{2n} кратно F_n. Аналогично

$$F_{3n} = F_{2n} F_{n+1} + F_{2n-1} F_n\,,$$

и можно заключить, что F_{3n} также кратно F_n. По индукции

$$F_{kn} \text{ кратно } F_n\,, \tag{6.110}$$

при любых целых k и n. Это поясняет, например, почему F_{15} (равное 610) кратно как F_3, так и F_5 (которые равны соответственно 2 и 5). На самом деле истинно даже большее — в упр. 27 доказывается, что

$$\text{НОД}(F_m, F_n) = F_{\text{НОД}(m,n)}\,. \tag{6.111}$$

Например, НОД(F_{12}, F_{18}) = НОД$(144, 2584) = 8 = F_6$.

Теперь можно доказать обращение свойства (6.110): если $n > 2$ и если F_m кратно F_n, то m кратно n. Действительно, если $F_n \backslash F_m$, то $F_n \backslash$ НОД$(F_m, F_n) = F_{\text{НОД}(m,n)} \leqslant F_n$. Это возможно только тогда, когда $F_{\text{НОД}(m,n)} = F_n$, и наше допущение, что $n > 2$, приводит к необходимости того, что НОД$(m, n) = n$. Следовательно, $n \backslash m$.

Обобщение этих идей было использовано Юрием Матиясевичем в его знаменитом доказательстве [266] того, что не существует алгоритма, позволяющего выяснить, разрешимо ли в целых числах заданное полиномиальное уравнение относительно многих неизвестных с целыми коэффициентами. Лемма Матиясевича утверждает, что при $n > 2$ число Фибоначчи F_m кратно F_n^2 тогда и только тогда, когда m кратно nF_n.

Докажем это путем рассмотрения последовательности $\langle F_{kn} \bmod F_n^2 \rangle$ при $k = 1, 2, 3, \ldots$, и выяснения, когда $F_{kn} \bmod F_n^2 = 0$. (Мы знаем, что m должно иметь вид kn, если $F_m \bmod F_n = 0$.) Вначале мы имеем $F_n \bmod F_n^2 = F_n$, что не равно нулю. Затем согласно (6.108) получаем

$$F_{2n} = F_n F_{n+1} + F_{n-1} F_n \equiv 2F_n F_{n+1} \pmod{F_n^2},$$

поскольку $F_{n+1} \equiv F_{n-1} \pmod{F_n}$. Аналогично

$$F_{2n+1} = F_{n+1}^2 + F_n^2 \equiv F_{n+1}^2 \pmod{F_n^2}.$$

Это сравнение позволяет вычислить

$$\begin{aligned} F_{3n} &= F_{2n+1} F_n + F_{2n} F_{n-1} \equiv \\ &\equiv F_{n+1}^2 F_n + (2F_n F_{n+1}) F_{n+1} = 3F_{n+1}^2 F_n \pmod{F_n^2}; \\ F_{3n+1} &= F_{2n+1} F_{n+1} + F_{2n} F_n \equiv \\ &\equiv F_{n+1}^3 + (2F_n F_{n+1}) F_n \equiv F_{n+1}^3 \pmod{F_n^2}. \end{aligned}$$

В общем случае по индукции по k находим, что

$$F_{kn} \equiv kF_n F_{n+1}^{k-1} \quad \text{и} \quad F_{kn+1} \equiv F_{n+1}^k \pmod{F_n^2}.$$

Поскольку F_{n+1} взаимно простое с F_n, то

$$\begin{aligned} F_{kn} \equiv 0 \pmod{F_n^2} \iff kF_n &\equiv 0 \pmod{F_n^2} \iff \\ \iff k &\equiv 0 \pmod{F_n}. \end{aligned}$$

Лемма Матиясевича доказана.

Одно из наиболее важных свойств чисел Фибоначчи — особый способ представления целых чисел с их использованием. Будем писать

$$j \gg k \iff j \geqslant k+2. \tag{6.112}$$

Тогда *каждое положительное целое число имеет единственное представление вида*

$$n = F_{k_1} + F_{k_2} + \cdots + F_{k_r}, \quad k_1 \gg k_2 \gg \cdots \gg k_r \gg 0. \qquad (6.113)$$

(Это "теорема Цеккендорфа (Zeckendorf)" [246], [381].) Например, один миллион можно представить следующим образом:

$$\begin{aligned} 1000000 &= 832040 + 121393 + 46368 + 144 + 55 = \\ &= F_{30} + F_{26} + F_{24} + F_{12} + F_{10}. \end{aligned}$$

Такое представление всегда может быть найдено с применением "жадного" алгоритма, когда в качестве F_{k_1} выбирается наибольшее возможное число Фибоначчи $\leqslant n$, затем в качестве F_{k_2} выбирается наибольшее число $\leqslant n - F_{k_1}$, и т.д. (Более строго, предположим, что $F_k \leqslant n < F_{k+1}$; тогда $0 \leqslant n - F_k < F_{k+1} - F_k = F_{k-1}$. Если n — число Фибоначчи, то представление (6.113) справедливо при $r = 1$ и $k_1 = k$. В противном случае индукция по n показывает, что $n - F_k$ имеет фибоначчиево представление $F_{k_2} + \cdots + F_{k_r}$, и представление (6.113) справедливо, если положить $k_1 = k$, потому что из неравенства $F_{k_2} \leqslant n - F_k < F_{k-1}$ вытекает, что $k \gg k_2$.) И обратно, всякое представление вида (6.113) означает, что

$$F_{k_1} \leqslant n < F_{k_1+1},$$

потому что наибольшим возможным значением $F_{k_2} + \cdots + F_{k_r}$ при $k \gg k_2 \gg \cdots \gg k_r \gg 0$ является

$$F_{k-2} + F_{k-4} + \cdots + F_{k \bmod 2 + 2} = F_{k-1} - 1, \quad \text{если } k \geqslant 2. \qquad (6.114)$$

(Эта формула легко доказывается индукцией по k; ее левая часть обращается в нуль при k, равном 2 или 3.) Поэтому k_1 представляет собой описанную выше "жадно" выбранную величину, и это представление должно быть единственным.

Любая система однозначного представления чисел является системой счисления; следовательно, теорема Цеккендорфа приводит к *фибоначчиевой системе счисления*. Любое неотрицательное целое число n может быть представлено в виде последовательностей нулей и единиц следующим образом:

$$n = (b_m b_{m-1} \ldots b_2)_F \quad \Longleftrightarrow \quad n = \sum_{k=2}^{m} b_k F_k. \qquad (6.115)$$

Эта система счисления отчасти похожа на двоичную (с основанием 2), с тем отличием, что в ней никогда не встречаются две единицы подряд. Например, вот как выглядят числа от 1 до 20, представленные по Фибоначчи:

$$\begin{array}{llll}
1 = (000001)_F & 6 = (001001)_F & 11 = (010100)_F & 16 = (100100)_F \\
2 = (000010)_F & 7 = (001010)_F & 12 = (010101)_F & 17 = (100101)_F \\
3 = (000100)_F & 8 = (010000)_F & 13 = (100000)_F & 18 = (101000)_F \\
4 = (000101)_F & 9 = (010001)_F & 14 = (100001)_F & 19 = (101001)_F \\
5 = (001000)_F & 10 = (010010)_F & 15 = (100010)_F & 20 = (101010)_F
\end{array}$$

Фибоначчиево представление миллиона, показанное ранее, может быть сопоставлено с его двоичным представлением $2^{19}+2^{18}+2^{17}+2^{16}+2^{14}+2^{9}+2^{6}$:

$$\begin{aligned}
(1000000)_{10} &= (10001010000000000010100000000)_F \\
&= (11110100001001000000)_2 .
\end{aligned}$$

Фибоначчиево представление требует несколько большего количества битов, потому что соседние единицы в нем не допускаются; в остальном эти представления аналогичны.

При добавлении 1 в фибоначчиевой системе счисления возникают два случая. Если "разряд единиц" равен 0, мы заменяем его на 1, т.е. добавляем $F_2 = 1$, поскольку разряд единиц связан с F_2. В противном случае двумя младшими цифрами оказываются 01, и мы заменяем их на 10 (тем самым добавляя $F_3 - F_2 = 1$). Наконец, мы должны выполнить необходимое количество "переносов", заменяя набор цифр '011' на '100', пока в строке не будут отсутствовать две подряд идущие цифры 1. (Это правило переноса эквивалентно замене $F_{m+1} + F_m$ на F_{m+2}.) Например, для перехода от $5 = (1000)_F$ к $6 = (1001)_F$ или от $6 = (1001)_F$ к $7 = (1010)_F$ переносы не требуются; но для перехода от $7 = (1010)_F$ к $8 = (10000)_F$ следует выполнить два переноса.

Несмотря на обилие рассмотренных нами свойств чисел Фибоначчи, мы пока не сталкивались с их аналитической записью. И хотя выражения в аналитическом виде не были найдены ни для чисел Стирлинга, ни для чисел Эйлера или Бернулли, для гармонических чисел нам все же удалось получить аналитическое выражение $H_n = \left[{n+1 \atop 2}\right]/n!$. А нет ли какой-нибудь связи между F_n и другими известными нам величинами? Можно ли "разрешить" рекуррентное соотношение, которое определяет F_n?

Ответ утвердительный: действительно, существует простой способ решения этого рекуррентного соотношения, если воспользоваться понятием *производящей функции*, вкратце рассматривавшейся в главе 5. Рассмотрим бесконечный ряд

$$F(z) = F_0 + F_1 z + F_2 z^2 + \cdots = \sum_{n\geqslant 0} F_n z^n. \qquad (6.116)$$

Если мы найдем простую формулу для $F(z)$, то появятся шансы найти простую формулу и для его коэффициентов F_n.

В главе 7 мы полностью сосредоточимся на производящих функциях, но к тому времени, когда мы до них доберемся, полезно будет иметь в запасе этот пример. Степенной ряд $F(z)$ имеет одно замечательное свойство, которое обнаруживается при его умножении на z и z^2:

$$\begin{aligned} F(z) &= F_0 + F_1 z + F_2 z^2 + F_3 z^3 + F_4 z^4 + F_5 z^5 + \cdots, \\ zF(z) &= \qquad F_0 z + F_1 z^2 + F_2 z^3 + F_3 z^4 + F_4 z^5 + \cdots, \\ z^2F(z) &= \qquad\qquad F_0 z^2 + F_1 z^3 + F_2 z^4 + F_3 z^5 + \cdots. \end{aligned}$$

Если теперь вычесть два последних равенства из первого, то члены, включающие z^2, z^3 и более высокие степени z, исчезнут в силу рекуррентного соотношения Фибоначчи. Кроме того, постоянный член F_0 на самом деле никогда не оказывается первым, потому что $F_0 = 0$. Таким образом, все, что остается после вычитания,— это $(F_1 - F_0)z$, т.е. просто z. Другими словами,

$$F(z) - zF(z) - z^2F(z) = z,$$

и решение относительно $F(z)$ дает нам компактную формулу

$$F(z) = \frac{z}{1 - z - z^2}. \qquad (6.117)$$

Итак, вся информация, содержащаяся в последовательности Фибоначчи, свернута в несложное (хотя и непонятное) выражение $z/(1 - z - z^2)$. Поверите вы в это или нет, но это прогресс, потому что теперь мы можем разложить знаменатель на множители и воспользоваться элементарными дробями для получения формулы, которую легко разложить в степенной ряд. Коэффициенты этого степенного ряда дадут нам выражение для чисел Фибоначчи в аналитическом виде.

Набросанный план может оказаться понятнее, если подойти к нему с другого конца. Если имеется какая-нибудь более простая производящая функция (скажем, $1/(1 - \alpha z)$, где α — константа), то нам известны коэффициенты при всех степенях z, потому что

$$\frac{1}{1 - \alpha z} = 1 + \alpha z + \alpha^2 z^2 + \alpha^3 z^3 + \cdots.$$

“Sit $1 + x + 2xx + 3x^3 + 5x^4 + 8x^5 + 13x^6 + 21x^7 + 34x^8$ &c Series nata ex divisione Unitatis per Trinomium $1 - x - xx$.”

— *А. де Муавр (A. de Moivre) [76]*

“Величины r, s, t, показывающие отношения членов, совпадают с аналогичными величинами в знаменателе дроби. Г-н де Муавр был первым, кто применил это свойство, каким бы очевидным оно ни казалось, к решению задач о бесконечных рядах, решить которые по-другому было бы весьма затруднительно.”

— *Дж. Стирлинг (J. Stirling) [343]*

Аналогично, если у нас есть некоторая производящая функция $A/(1-\alpha z)+B/(1-\beta z)$, то коэффициенты также легко вычисляются, потому что

$$\begin{aligned}\frac{A}{1-\alpha z}+\frac{B}{1-\beta z} &= A\sum_{n\geqslant 0}(\alpha z)^n + B\sum_{n\geqslant 0}(\beta z)^n\\ &= \sum_{n\geqslant 0}(A\alpha^n + B\beta^n)z^n\,.\end{aligned}\tag{6.118}$$

Таким образом, все, что от нас требуется, — это найти константы A, B, α и β, такие, что

$$\frac{A}{1-\alpha z}+\frac{B}{1-\beta z} = \frac{z}{1-z-z^2}\,,$$

и тогда мы получим выражение в аналитическом виде $A\alpha^n+B\beta^n$ для коэффициента F_n при z^n в $F(z)$. Левую часть можно переписать как

$$\frac{A}{1-\alpha z}+\frac{B}{1-\beta z} = \frac{A-A\beta z+B-B\alpha z}{(1-\alpha z)(1-\beta z)}\,,$$

так что интересующие нас четыре константы являются решениями двух полиномиальных уравнений:

$$(1-\alpha z)(1-\beta z) = 1-z-z^2\,;\tag{6.119}$$

$$(A+B)-(A\beta+B\alpha)z = z\,.\tag{6.120}$$

Мы хотим представить знаменатель $F(z)$ в виде $(1-\alpha z)(1-\beta z)$; тогда мы сможем выразить $F(z)$ в виде суммы двух дробей, в которых сомножители $(1-\alpha z)$ и $(1-\beta z)$ удачно отделены друг от друга.

Обратите внимание, что в знаменателе уравнения (6.119) сомножители записаны в виде $(1-\alpha z)(1-\beta z)$, а не в более привычной форме $c(z-\rho_1)(z-\rho_2)$, где ρ_1 и ρ_2 — корни уравнения. Причина этого в том, что запись $(1-\alpha z)(1-\beta z)$ приводит к более привлекательным разложениям в степенные ряды.

Как обычно, авторы не могут устоять перед хитростями...

Найти α и β можно разными способами, в одном из которых используется тонкая уловка. Введем новую переменную w и поищем разложение

$$w^2-wz-z^2 = (w-\alpha z)(w-\beta z)\,.$$

Тогда мы сможем просто положить $w=1$ и получить разложение $1-z-z^2$. Корни уравнения $w^2-wz-z^2=0$ можно найти при помощи обычной формулы для корней квадратного уравнения — они равны

$$\frac{z\pm\sqrt{z^2+4z^2}}{2} = \frac{1\pm\sqrt{5}}{2}z\,.$$

Таким образом,

$$w^2 - wz - z^2 \;=\; \left(w - \frac{1+\sqrt{5}}{2}\,z\right)\left(w - \frac{1-\sqrt{5}}{2}\,z\right)$$

и искомые константы α и β найдены.

Число $(1+\sqrt{5})/2 \approx 1.61803$ играет важную роль во многих разделах математики, а равно и в мире искусства, где с античных времен оно считается эстетически наиболее благоприятным. Оно даже получило собственное имя, *отношение золотого сечения*, и обозначается греческой буквой ϕ в честь Фидия, который, как утверждается, сознательно использовал его в своих скульптурах. Второй корень, $(1-\sqrt{5})/2 = -1/\phi \approx -.61803$, обладает многими свойствами числа ϕ, так что у него тоже есть собственное имя — $\hat{\phi}$, "фи с крышей". Оба эти числа являются корнями уравнения $w^2 - w - 1 = 0$, так что

В соответствии с обширными эмпирическими наблюдениями европейских исследователей [136] отношение роста человека к высоте расположения пупка составляет примерно 1.618.

$$\phi^2 \;=\; \phi + 1\,; \qquad \hat{\phi}^2 \;=\; \hat{\phi} + 1\,. \tag{6.121}$$

(Ниже вы узнаете о числах ϕ и $\hat{\phi}$ больше.)

Итак, константы $\alpha = \phi$ и $\beta = \hat{\phi}$, требующиеся в (6.119), найдены; теперь надо лишь установить A и B в (6.120). Приравнивая $z = 0$ в этом уравнении, получим $B = -A$, так что (6.120) сводится к уравнению

$$-\hat{\phi}A + \phi A \;=\; 1\,,$$

решением которого является $A = 1/(\phi - \hat{\phi}) = 1/\sqrt{5}$. Таким образом, разложение (6.117) на элементарные дроби представляет собой

$$F(z) \;=\; \frac{1}{\sqrt{5}}\left(\frac{1}{1-\phi z} - \frac{1}{1-\hat{\phi} z}\right). \tag{6.122}$$

Неплохо — $F(z)$ получено в том виде, в котором мы и хотели его получить. Разложение дробей в степенные ряды, как в (6.118), дает нам выражение для коэффициента при z^n в аналитическом виде:

$$F_n \;=\; \frac{1}{\sqrt{5}}(\phi^n - \hat{\phi}^n)\,. \tag{6.123}$$

(Эта формула впервые была опубликована Даниэлем Бернулли (Daniel Bernoulli) в 1728 году, после чего о ней благополучно забыли до 1843 года, когда ее вновь открыл Жак Бине (Jacques Binet) [31].)

Прежде чем замереть в восторге от нашего вывода, следует убедиться в его правильности. При $n = 0$ формула совершенно верно дает $F_0 = 0$; при $n = 1$ она дает $F_1 = (\phi - \hat{\phi})/\sqrt{5}$, что, конечно же, равно 1. При бо́льших степенях уравнения (6.121) показывают, что числа, определенные формулой (6.123), удовлетворяют фибоначчиевому рекуррентному соотношению, так что по индукции они обязаны быть числами Фибоначчи. (Можно также разложить ϕ^n и $\hat{\phi}^n$ в соответствии с биномиальной теоремой и избавиться от различных степеней $\sqrt{5}$, но это громоздко и приводит к путанице. Смысл выражения в аналитическом виде не в том, чтобы обеспечить способ быстрого вычисления, а в том, чтобы указать связь F_n с другими математическими величинами.)

Проявив некоторую смекалку, можно было бы просто угадать формулу (6.123) и доказать ее по индукции. Однако метод производящих функций является действенным способом именно ее нахождения — в главе 7 мы увидим, что тот же самый метод приводит к решению куда более трудных рекуррентных соотношений. Между прочим, мы нисколько не беспокоились, сходятся ли бесконечные суммы в нашем выводе формулы (6.123); оказывается, что большинство операций с коэффициентами степенного ряда может быть строго обоснованно независимо от того, сходятся ли данные суммы в действительности [182]. Тем не менее скептически настроенные читатели, подозрительно относящиеся к действиям с бесконечными суммами, могут найти успокоение в том факте, что после того, как уравнение (6.123) найдено с использованием бесконечного ряда, оно может быть проверено путем надежного доказательства по индукции.

Одним из интересных следствий формулы (6.123) является то, что целое число F_n чрезвычайно близко к иррациональному числу $\phi^n/\sqrt{5}$ при большом n. (Поскольку $\hat{\phi}$ по абсолютной величине меньше 1, величина $\hat{\phi}^n$ экспоненциально убывает и ее влияние становится почти несущественным.) Например, числа $F_{10} = 55$ и $F_{11} = 89$ очень близки к

$$\frac{\phi^{10}}{\sqrt{5}} \approx 55.00364 \quad \text{и} \quad \frac{\phi^{11}}{\sqrt{5}} \approx 88.99775\,.$$

Этим можно воспользоваться для другого выражения в аналитическом виде,

$$F_n = \left\lfloor \frac{\phi^n}{\sqrt{5}} + \frac{1}{2} \right\rfloor = \frac{\phi^n}{\sqrt{5}}, \quad \text{округленное до ближайшего целого,} \tag{6.124}$$

потому что $\left|\hat{\phi}^n/\sqrt{5}\right| < \frac{1}{2}$ при всех $n \geqslant 0$. Когда n четное, F_n немного меньше $\phi^n/\sqrt{5}$; в противном случае оно немного больше.

Тождество Кассини (6.103) можно переписать как

$$\frac{F_{n+1}}{F_n} - \frac{F_n}{F_{n-1}} = \frac{(-1)^n}{F_{n-1}\,F_n}\,.$$

При большом n величина $1/F_{n-1}F_n$ весьма мала, так что отношение F_{n+1}/F_n должно быть очень близким к F_n/F_{n-1}, а (6.124) указывает, что это отношение приближает ϕ. На самом деле

$$F_{n+1} = \phi F_n + \hat{\phi}^n . \tag{6.125}$$

(Это тождество проверяется при $n = 0$ и $n = 1$ непосредственно, а для $n > 1$ оно доказывается по индукции; оно может быть также доказано непосредственно подстановкой в формулу (6.123).) Отношение F_{n+1}/F_n очень близко к числу ϕ, которое оно приближает то с избытком, то с недостатком.

В силу простого совпадения ϕ достаточно близко к количеству километров в одной миле. (Точное значение равно 1.609344, поскольку 1 дюйм равен в точности 2.54 сантиметра.) Это дает удобный способ перевода в уме километров в мили и обратно, поскольку расстояние в F_{n+1} километров равно (довольно близко) расстоянию в F_n миль.

Если США когда-нибудь все же перейдут на метрическую систему, то американские знаки ограничения скорости изменятся с 55 миль/ч на 89 км/ч. А может, полиция будет столь великодушна, что разрешит ездить со скоростью 90 миль/ч?

Предположим, что мы хотим преобразовать некоторое число (не являющееся числом Фибоначчи) километров в мили. Сколько будет 30 км по-американски? Это легко: мы просто прибегаем к фибоначчиевой системе счисления и переводим в уме число 30 в его фибоначчиевое представление $21+8+1$ при помощи описанного выше жадного алгоритма. Теперь можно сдвинуть каждое число на одну позицию вправо и получить $13 + 5 + 1$. (Первоначальная единица была числом F_2, поскольку $k_r \gg 0$ в (6.113); новая единица соответствует F_1.) Сдвиг вправо практически равносилен делению на ϕ. Следовательно, наша оценка — 19 миль. (Что достаточно точно; правильный ответ — около 18.64 миль.) Аналогично для перехода от миль к километрам можно осуществить сдвиг на одну позицию влево; 30 миль — это примерно $34+13+2 = 49$ километров. (Эта оценка менее точная; правильное значение — около 48.28.)

Оказывается, что подобное правило "сдвига вправо" дает правильно округленное число миль в n километрах для всех $n \leqslant 100$, за исключением случаев $n = 4$, 12, 54, 62, 75, 83, 91, 96 и 99, когда оно отличается меньше чем на 2/3 мили. Правило "сдвига влево" дает либо правильно округленное число километров в n милях, либо на 1 км больше при всех $n \leqslant 113$. (Единственный действительно нарушающий данное правило случай — это $n = 4$, когда отдельные ошибки округления для $n = 3+1$ накладываются друг на друга, вместо того, чтобы компенсировать друг друга.)

Правило сдвига вправо заменяет n на $f(n/\phi)$, а правило сдвига влево заменяет n на $f(n\phi)$, где $f(x) = \lfloor x + \phi^{-1} \rfloor$.

6.7 Континуанты

Числа Фибоначчи обнаруживают важные связи с деревом Штерна–Броко, которое мы рассматривали в главе 4, и допускают важные обобщения на последовательность полиномов, которую обстоятельно изучал Эйлер. Эти полиномы называются *континуантами*, или K-*полиномами*, потому что представляют собой ключ к изучению непрерывных дробей вида

$$a_0+\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+\cfrac{1}{a_4+\cfrac{1}{a_5+\cfrac{1}{a_6+\cfrac{1}{a_7}}}}}}}\,. \tag{6.126}$$

Континуантный полином $K_n(x_1,x_2,\ldots,x_n)$ имеет n параметров и определяется следующим рекуррентным соотношением:

$$\begin{aligned} K_0() &= 1\,;\\ K_1(x_1) &= x_1\,;\\ K_n(x_1,\ldots,x_n) &= K_{n-1}(x_1,\ldots,x_{n-1})x_n+\\ &\quad+K_{n-2}(x_1,\ldots,x_{n-2})\,. \end{aligned} \tag{6.127}$$

Например, три следующих за $K_1(x_1)$ случая таковы:

$$\begin{aligned} K_2(x_1,x_2) &= x_1x_2+1\,;\\ K_3(x_1,x_2,x_3) &= x_1x_2x_3+x_1+x_3\,;\\ K_4(x_1,x_2,x_3,x_4) &= x_1x_2x_3x_4+x_1x_2+x_1x_4+x_3x_4+1\,. \end{aligned}$$

По индукции легко убедиться, что количество членов в K_n равно числу Фибоначчи:

$$K_n(1,1,\ldots,1) = F_{n+1}\,. \tag{6.128}$$

Если число переменных ясно из контекста, можно писать вместо 'K_n' просто 'K', так же, как можно было опускать число параметров, имея дело с гипергеометрическими функциями F из главы 5. Например, $K(x_1,x_2)=K_2(x_1,x_2)=x_1x_2+1$. Конечно же, в формулах наподобие (6.128) нижний индекс n необходим.

Эйлер [112] заметил, что полином $K(x_1,x_2,\ldots,x_n)$ может быть получен, если начать с произведения $x_1x_2\ldots x_n$ и затем вычеркивать соседние пары $x_k x_{k+1}$ всеми возможными способами. Правило Эйлера можно представить графически, образуя всевозможные слова длиной n из точек и тире в "азбуке Морзе", где каждая точка дает длину 1, а каждое тире — длину 2; вот слова длиной 4 из азбуки Морзе:

••••   ••–   •–•   –••   ––

Эти слова из точек и тире соответствуют членам $K(x_1, x_2, x_3, x_4)$; точка означает переменную, которая включена, а тире означает пару переменных, которая исключена. Например, •–• соответствует $x_1 x_4$.

Слово длиной n в азбуке Морзе, которое содержит k тире, содержит $n-2k$ точек, а всего $n-k$ знаков. Эти точки и тире могут быть расставлены $\binom{n-k}{k}$ способами; поэтому, если заменить каждую точку на z, а каждое тире — на 1, то можно получить

$$K_n(z, z, \ldots, z) \;=\; \sum_{k=0}^{n} \binom{n-k}{k} z^{n-2k}\,. \tag{6.129}$$

Кроме того, известно, что общее количество членов в континуанте равно некоторому числу Фибоначчи. Следовательно, установлено тождество

$$F_{n+1} \;=\; \sum_{k=0}^{n} \binom{n-k}{k}\,. \tag{6.130}$$

(Аналитическая запись для (6.129), обобщающая формулу Эйлера–Бине (6.123) для чисел Фибоначчи, приведена в (5.74).)

Связь между континуантами и последовательностями азбуки Морзе показывает, что континуанты обладают зеркальной симметрией:

$$K(x_n, \ldots, x_2, x_1) \;=\; K(x_1, x_2, \ldots, x_n)\,. \tag{6.131}$$

Поэтому в дополнение к правосторонней рекуррентности из определения (6.127) они подчиняются рекуррентности, которая пристраивает переменные слева:

$$K_n(x_1, \ldots, x_n) \;=\; x_1 K_{n-1}(x_2, \ldots, x_n) + K_{n-2}(x_3, \ldots, x_n)\,. \tag{6.132}$$

Обе эти рекуррентности представляют собой частные случаи более общего правила:

$$\begin{aligned} K_{m+n}(x_1, \ldots, x_m, x_{m+1}, \ldots, x_{m+n}) \;&=\\ = K_m(x_1, \ldots, x_m)\, K_n(x_{m+1}, \ldots, x_{m+n}) &+\\ + K_{m-1}(x_1, \ldots, x_{m-1})\, K_{n-1}(x_{m+2}, \ldots, x_{m+n})&\,. \end{aligned} \tag{6.133}$$

Это правило легко истолковать, исходя из аналогии с азбукой Морзе: первое произведение $K_m K_n$ дает те члены K_{m+n}, в которых нет тире в позиции $[m, m+1]$, в то время как второе произведение дает те члены, в которых в этой позиции тире присутствует. Если же положить все x равными 1, то данное тождество показывает, что $F_{m+n+1} = F_{m+1}F_{n+1} + F_m F_n$; таким образом, (6.108) является частным случаем (6.133).

Эйлер [112] обнаружил, что континуанты подчиняются еще более замечательному правилу, которое является обобщением тождества Кассини:

$$\begin{aligned}
&K_{m+n}(x_1, \ldots, x_{m+n})\, K_k(x_{m+1}, \ldots, x_{m+k}) \;= \\
&\quad = K_{m+k}(x_1, \ldots, x_{m+k})\, K_n(x_{m+1}, \ldots, x_{m+n}) + \\
&\qquad + (-1)^k K_{m-1}(x_1, \ldots, x_{m-1}) \times \\
&\qquad \times K_{n-k-1}(x_{m+k+2}, \ldots, x_{m+n}) \,.
\end{aligned} \tag{6.134}$$

Это правило (доказанное в упр. 29) справедливо тогда, когда индексы при всех K неотрицательны. Например, при $k = 2$, $m = 1$ и $n = 3$ имеем

$$K(x_1, x_2, x_3, x_4)\, K(x_2, x_3) \;=\; K(x_1, x_2, x_3)\, K(x_2, x_3, x_4) + 1 \,.$$

Континуанты тесно связаны с алгоритмом Евклида. Предположим, например, что вычисление НОД(m, n) завершается за четыре шага:

$$\begin{aligned}
&\text{НОД}(m, n) \;= \\
&\quad = \text{НОД}(n_0, n_1) \;= && n_0 = m, \quad && n_1 = n\,; \\
&\quad = \text{НОД}(n_1, n_2) \;= && n_2 = n_0 \bmod n_1 && = n_0 - q_1 n_1\,; \\
&\quad = \text{НОД}(n_2, n_3) \;= && n_3 = n_1 \bmod n_2 && = n_1 - q_2 n_2\,; \\
&\quad = \text{НОД}(n_3, n_4) \;= && n_4 = n_2 \bmod n_3 && = n_2 - q_3 n_3\,; \\
&\quad = \text{НОД}(n_4, 0) \;=\; n_4\,. && 0 = n_3 \bmod n_4 && = n_3 - q_4 n_4\,.
\end{aligned}$$

Тогда

$$\begin{aligned}
n_4 &= n_4 &&= K()n_4\,; \\
n_3 &= q_4 n_4 &&= K(q_4)n_4\,; \\
n_2 &= q_3 n_3 + n_4 &&= K(q_3, q_4)n_4\,; \\
n_1 &= q_2 n_2 + n_3 &&= K(q_2, q_3, q_4)n_4\,; \\
n_0 &= q_1 n_1 + n_2 &&= K(q_1, q_2, q_3, q_4)n_4\,.
\end{aligned}$$

В общем случае, если алгоритм Евклида находит наибольший общий делитель d за k шагов, после вычисления последовательности частных q_1, …, q_k, то начальными числами являются $K(q_1, q_2, \ldots, q_k)d$ и $K(q_2, \ldots, q_k)d$. (Этот факт был замечен еще в начале XVIII века Тома Фанте де Ланьи (Thomas Fantet de Lagny) [232], который, по-видимому, был первым, кто непосредственно рассматривал континуанты. При этом Ланьи отмечал, что последовательные числа Фибоначчи, которые фигурируют в качестве континуантов, когда q принимают свои минимальные значения, являются теми наименьшими числами на входе, при которых алгоритму Евклида требуется выполнить некоторое заданное число шагов.)

Континуанты тесно связаны также с непрерывными дробями, от которых они получили свое название. Например,

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3}}} = \frac{K(a_0, a_1, a_2, a_3)}{K(a_1, a_2, a_3)}. \tag{6.135}$$

Та же картина наблюдается для дробей любой глубины. Это легко доказать по индукции; например,

$$\frac{K(a_0, a_1, a_2, a_3 + 1/a_4)}{K(a_1, a_2, a_3 + 1/a_4)} = \frac{K(a_0, a_1, a_2, a_3, a_4)}{K(a_1, a_2, a_3, a_4)}$$

в силу тождества

$$\begin{aligned} K_n(x_1, \ldots, x_{n-1}, x_n + y) &= \\ &= K_n(x_1, \ldots, x_{n-1}, x_n) + K_{n-1}(x_1, \ldots, x_{n-1})y. \end{aligned} \tag{6.136}$$

(Это тождество доказано и обобщено в упр. 30.)

Кроме того, континуанты тесно связаны и с деревом Штерна–Броко, которое обсуждалось в главе 4. Каждый узел в этом дереве может быть представлен в виде последовательности L и R, скажем, в виде

$$R^{a_0} L^{a_1} R^{a_2} L^{a_3} \ldots R^{a_{n-2}} L^{a_{n-1}}, \tag{6.137}$$

где $a_0 \geqslant 0$, $a_1 \geqslant 1$, $a_2 \geqslant 1$, $a_3 \geqslant 1$, $\ldots$, $a_{n-2} \geqslant 1$, $a_{n-1} \geqslant 0$, а n четное. Используя 2×2-матрицы L и R из (4.33), нетрудно доказать по индукции, что матричным эквивалентом (6.137) является

$$\begin{pmatrix} K_{n-2}(a_1, \ldots, a_{n-2}) & K_{n-1}(a_1, \ldots, a_{n-2}, a_{n-1}) \\ K_{n-1}(a_0, a_1, \ldots, a_{n-2}) & K_n(a_0, a_1, \ldots, a_{n-2}, a_{n-1}) \end{pmatrix}. \tag{6.138}$$

(Доказательство этого является частью упр. 87.) Например,

$$R^a L^b R^c L^d = \begin{pmatrix} bc + 1 & bcd + b + d \\ abc + a + c & abcd + ab + ad + cd + 1 \end{pmatrix}.$$

В силу этого можно воспользоваться (4.34), чтобы записать, наконец, в аналитическом виде дробь из дерева Штерна–Броко, L-R-представлением которой является (6.137):

$$f(R^{a_0}\ldots L^{a_{n-1}}) = \frac{K_{n+1}(a_0, a_1, \ldots, a_{n-1}, 1)}{K_n(a_1, \ldots, a_{n-1}, 1)}. \qquad (6.139)$$

(Это теорема Халфена (Halphen) [174].) Например, чтобы найти дробь для представления LRRL, полагаем $a_0 = 0$, $a_1 = 1$, $a_2 = 2$, $a_3 = 1$ и $n = 4$; тогда равенство (6.139) дает

$$\frac{K(0,1,2,1,1)}{K(1,2,1,1)} = \frac{K(2,1,1)}{K(1,2,1,1)} = \frac{K(2,2)}{K(3,2)} = \frac{5}{7}.$$

(Для поглощения единиц спереди и сзади в списках параметров мы воспользовались правилом $K_n(x_1, \ldots, x_{n-1}, x_n + 1) = K_{n+1} \times (x_1, \ldots, x_{n-1}, x_n, 1)$, которое получается, если подставить $y = 1$ в (6.136).)

Сравнение (6.135) и (6.139) показывает, что дробь, соответствующая произвольному узлу (6.137) в дереве Штерна–Броко, допускает представление в виде непрерывной дроби

$$f(R^{a_0}\ldots L^{a_{n-1}}) = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ldots + \cfrac{1}{a_{n-1} + \cfrac{1}{1}}}}}. \qquad (6.140)$$

Таким образом, можно без проволочек переходить от непрерывных дробей к соответствующим узлам дерева Штерна–Броко. Например,

$$f(LRRL) = 0 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1}}}}.$$

В главе 4 мы отмечали, что иррациональные числа определяют бесконечные пути в дереве Штерна–Броко и могут быть представлены в виде бесконечной строки букв “L” и “R”. Если такой бесконечной строкой для числа α является $R^{a_0}L^{a_1}R^{a_2}L^{a_3}\ldots$, то ей соответствует непрерывная дробь

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cfrac{1}{a_5 + \cfrac{1}{\ddots}}}}}}. \tag{6.141}$$

Эта бесконечная непрерывная дробь может быть получена и непосредственно. Пусть $\alpha_0 = \alpha$, и при $k \geqslant 0$ положим

$$a_k = \lfloor \alpha_k \rfloor; \qquad \alpha_k = a_k + \frac{1}{\alpha_{k+1}}. \tag{6.142}$$

Числа a_k называются неполными частными числа α. Если α рациональное, скажем, m/n, то этот процесс перебирает обнаруживаемые с помощью алгоритма Евклида частные и затем останавливается (при $\alpha_{k+1} = \infty$).

Рациональна или иррациональна константа Эйлера γ? Никто не знает. Некоторую информацию об этой нерешенной задаче можно получить, поискав число γ в дереве Штерна–Броко: если оно рационально, мы найдем его, а если иррационально, то мы найдем все наилучшие рациональные приближения к нему. Непрерывная дробь для числа γ начинается со следующих неполных частных:

Или знает, но не говорит об этом.

k	0	1	2	3	4	5	6	7	8
a_k	0	1	1	2	1	2	1	4	3

Поэтому представление Штерна–Броко для этой дроби начинается со строки LRLLRLLRLLLLRRRL...; никакой очевидной закономерности здесь не наблюдается. Вычисления Ричарда Брента (Richard Brent) [38] показывают, что если γ — рациональное число, то его знаменатель должен насчитывать более 10 тысяч десятичных цифр. Поэтому никто не верит в то, что число γ рационально; тем не менее никто не смог доказать, что оно таковым не является.

Ну, γ должно быть иррациональным в силу малоизвестного изречения Эйнштейна: "Бог не загромождает Вселенную большими знаменателями".

Завершим эту главу доказательством одного замечательного тождества, которое связывает воедино многие из подобных понятий. В главе 3 было введено понятие спектра — спектром числа α называется мультимножество чисел $\lfloor n\alpha \rfloor$, где α — заданная константа. Поэтому бесконечный ряд

$$\sum_{n \geqslant 1} z^{\lfloor n\phi \rfloor} = z + z^3 + z^4 + z^6 + z^8 + z^9 + \cdots$$

может быть назван производящей функцией для спектра числа ϕ, где $\phi = (1+\sqrt{5})/2$ — отношение золотого сечения. Тождество, которое мы докажем (открытое в 1976 году Дж. Л. Дейвисоном (J. L. Davison) [73]), представляет собой бесконечную непрерывную дробь, которая связывает эту производящую функцию с последовательностью Фибоначчи:

$$\cfrac{z^{F_1}}{1+\cfrac{z^{F_2}}{1+\cfrac{z^{F_3}}{1+\cfrac{z^{F_4}}{\ddots}}}} = (1-z)\sum_{n\geqslant 1} z^{\lfloor n\phi \rfloor}. \tag{6.143}$$

Интересны обе части (6.143), но вначале рассмотрим числа $\lfloor n\phi \rfloor$. Если представление Фибоначчи (6.113) числа n есть $F_{k_1} + \cdots + F_{k_r}$, то ожидается, что $n\phi$ будет приблизительно равным $F_{k_1+1} + \cdots + F_{k_r+1}$ — числу, которое получается, если сдвинуть фибоначчиево представление влево (как при переводе миль в километры). В действительности из (6.125) известно, что

$$n\phi = F_{k_1+1} + \cdots + F_{k_r+1} - \left(\hat{\phi}^{k_1} + \cdots + \hat{\phi}^{k_r}\right).$$

Поскольку $\hat{\phi} = -1/\phi$ и $k_1 \gg \cdots \gg k_r \gg 0$, то

$$\begin{aligned} \left|\hat{\phi}^{k_1} + \cdots + \hat{\phi}^{k_r}\right| &< \phi^{-k_r} + \phi^{-k_r-2} + \phi^{-k_r-4} + \cdots = \\ &= \frac{\phi^{-k_r}}{1-\phi^{-2}} = \phi^{1-k_r} \leqslant \phi^{-1} < 1; \end{aligned}$$

и, рассуждая аналогично, $\hat{\phi}^{k_1} + \cdots + \hat{\phi}^{k_r}$ имеет тот же знак, что и $(-1)^{k_r}$. Следовательно,

$$\lfloor n\phi \rfloor = F_{k_1+1} + \cdots + F_{k_r+1} - \left[k_r(n) \text{ четное}\right]. \tag{6.144}$$

Условимся называть число n *фибоначчиево-нечетным* (или, для краткости, F-нечетным), если младший бит фибоначчиевого представления равен 1 (что то же самое, что и $k_r(n) = 2$). В противном случае n *фибоначчиево-четное* (F-четное). Например, наименьшими F-нечетными числами являются 1, 4, 6, 9, 12, 14, 17 и 19. Если $k_r(n)$ четное, то число $n-1$ является F-четным в силу (6.114); аналогично, если $k_r(n)$ нечетное, то $n-1$ является F-нечетным. Поэтому

$$k_r(n) \text{ четное} \quad \Longleftrightarrow \quad n-1 \text{ F-четное.}$$

Кроме того, если $k_r(n)$ четное, то из (6.144) следует, что $k_r(\lfloor n\phi \rfloor) = 2$; если $k_r(n)$ нечетное, то (6.144) утверждает, что

$k_r(\lfloor n\phi \rfloor) = k_r(n) + 1$. Таким образом, $k_r(\lfloor n\phi \rfloor)$ всегда четное, и доказано, что

$$\lfloor n\phi \rfloor - 1 \text{ всегда F-четное.}$$

Обратно, если m — некоторое F-четное число, то можно провести это вычисление в обратном порядке и найти n, такое, что $m + 1 = \lfloor n\phi \rfloor$. (Сперва прибавляется 1 в F-записи, как объяснялось выше. Если переносов не происходит, то число n оказывается равным $(m+2)$, сдвинутому вправо; в противном случае число n равно сдвинутому вправо $(m + 1)$.) Поэтому сумма в правой части (6.143) может быть записана в виде

$$\sum_{n\geqslant 1} z^{\lfloor n\phi \rfloor} = z \sum_{m\geqslant 0} z^m [m \text{— F-четное}]. \tag{6.145}$$

А как насчет дроби слева? Перепишем (6.143) так, чтобы данная непрерывная дробь выглядела как (6.141), с 1 во всех числителях:

$$\cfrac{1}{z^{-F_0} + \cfrac{1}{z^{-F_1} + \cfrac{1}{z^{-F_2} + \cfrac{1}{\ddots}}}} = \frac{1-z}{z} \sum_{n\geqslant 1} z^{\lfloor n\phi \rfloor}. \tag{6.146}$$

(Это весьма тонкое преобразование! Числитель и знаменатель исходной дроби, имеющей в числителе z^{F_n}, нужно поделить на $z^{F_{n-1}}$.) Если оборвать эту новую непрерывную дробь на элементе $1/z^{-F_n}$, то ее величина будет равна отношению континуантов

$$\frac{K_{n+2}(0, z^{-F_0}, z^{-F_1}, \ldots, z^{-F_n})}{K_{n+1}(z^{-F_0}, z^{-F_1}, \ldots, z^{-F_n})} = \frac{K_n(z^{-F_1}, \ldots, z^{-F_n})}{K_{n+1}(z^{-F_0}, z^{-F_1}, \ldots, z^{-F_n})},$$

как и в (6.135). Вначале рассмотрим знаменатель, в надежде на его податливость. Полагая $Q_n = K_{n+1}(z^{-F_0}, \ldots, z^{-F_n})$, находим, что $Q_0 = 1$, $Q_1 = 1 + z^{-1}$, $Q_2 = 1 + z^{-1} + z^{-2}$, $Q_3 = 1 + z^{-1} + z^{-2} + z^{-3} + z^{-4}$ и вообще все продолжается как нельзя лучше и получается геометрический ряд

$$Q_n = 1 + z^{-1} + z^{-2} + \cdots + z^{-(F_{n+2}-1)}.$$

Соответствующий этому знаменателю числитель $P_n = K_n(z^{-F_1}, \ldots, z^{-F_n})$ оказывается подобным Q_n, но с меньшим числом членов. Например, имеем

$$P_5 = z^{-1} + z^{-2} + z^{-4} + z^{-5} + z^{-7} + z^{-9} + z^{-10} + z^{-12},$$

в сравнении с $Q_5 = 1 + z^{-1} + \cdots + z^{-12}$. Более внимательное рассмотрение выявляет закономерность, определяющую, какие члены присутствуют. Вот она:

$$
\begin{aligned}
P_5 &= \frac{1 + z^2 + z^3 + z^5 + z^7 + z^8 + z^{10} + z^{11}}{z^{12}} = \\
&= z^{-12} \sum_{m=0}^{12} z^m \,[m \text{— F-четное}]\,;
\end{aligned}
$$

и вообще, можно доказать по индукции, что

$$
P_n = z^{1-F_{n+2}} \sum_{m=0}^{F_{n+2}-1} z^m \,[m \text{— F-четное}]\,.
$$

Поэтому

$$
\frac{P_n}{Q_n} = \frac{\sum_{m=0}^{F_{n+2}-1} z^m \,[m \text{— F-четное}]}{\sum_{m=0}^{F_{n+2}-1} z^m}\,.
$$

Взяв предел при $n \to \infty$, получаем (6.146) в силу (6.145).

Упражнения

Разминка

1 Каковы $\left[{4 \atop 2}\right] = 11$ перестановок множества $\{1, 2, 3, 4\}$, содержащих ровно два цикла? (Используйте обычную запись перестановки наподобие 2314, а не разложение на циклы, как в (6.4).)

2 Всего имеется m^n функций, отображающих множество из n элементов на множество из m элементов. Сколько из них принимают ровно k различных значений?

3 Те, кто складывали стопку карт в реальной жизни, знают, что это следует делать с небольшим запасом прочности, чтобы стопка не опрокинулась при малейшем колебании воздуха. Предположим, что центр тяжести верхних k карт должен находиться по меньшей мере в ϵ единицах от края $(k + 1)$-й карты. (Таким образом, первая карта, например, может выступать над второй самое большее на $1 - \epsilon$ единиц.) Можно ли при этом получить сколь угодно большой выступ, располагая достаточным количеством карт?

4 Выразите $1/1 + 1/3 + \cdots + 1/(2n+1)$ через гармонические числа.

5 Поясните, как получить рекуррентное соотношение (6.75) из определения (6.74) величины $U_n(x, y)$, и решите это рекуррентное соотношение.

6 Некий исследователь оставил на острове пару кроликов. Если они становятся способны к размножению через месяц и если каждая пара половозрелых кроликов способна производить новую пару крольчат ежемесячно, то сколько пар кроликов будет в наличии через n месяцев при полном отсутствии смертности среди кроликов? (Через два месяца будет две пары, одна из которых — новорожденная.) Установите связь между этой задачей и "пчелиным деревом" из настоящей главы.

Если гармонические числа — червячные, то числа Фибоначчи — точно кроличьи!

7 Покажите, что тождество Кассини (6.103) представляет собой как частный случай соотношения (6.108), так и частный случай соотношения (6.134).

8 Воспользуйтесь фибоначчиевой системой счисления, чтобы перевести скорость 65 миль/ч в приблизительное число км/ч.

9 Сколько примерно квадратных километров в 8 квадратных милях?

10 Каково представление ϕ в виде непрерывной дроби?

Обязательные упражнения

11 Чему равна сумма $\sum_k (-1)^k \left[{n \atop k}\right]$ — сумма знакочередующихся чисел ряда треугольника Стирлинга для числа циклов, если n — неотрицательное целое число?

12 Докажите, что числа Стирлинга обладают правилом обращения, которое аналогично правилу (5.48):

$$g(n) = \sum_k \left\{{n \atop k}\right\}(-1)^k f(k) \iff f(n) = \sum_k \left[{n \atop k}\right](-1)^k g(k).$$

13 Дифференциальные операторы $D = \frac{d}{dz}$ и $\vartheta = zD$ упоминались в главах 2 и 5. Имеем

$$\vartheta^2 = z^2D^2 + zD,$$

потому что $\vartheta^2 f(z) = \vartheta z f'(z) = z\frac{d}{dz} z f'(z) = z^2 f''(z) + z f'(z)$, что равносильно $(z^2D^2 + zD)f(z)$. Аналогично можно показать, что $\vartheta^3 = z^3D^3 + 3z^2D^2 + zD$. Докажите общие формулы

$$\vartheta^n = \sum_k \left\{{n \atop k}\right\} z^k D^k,$$

$$z^n D^n = \sum_k \left[{n \atop k}\right] (-1)^{n-k} \vartheta^k$$

для всех $n \geqslant 0$. (Эти формулы могут быть использованы для перехода от дифференциального выражения вида $\sum_k \alpha_k z^k f^{(k)}(z)$ к выражению вида $\sum_k \beta_k \vartheta^k f(z)$, как это делалось в (5.109).)

14 Докажите степенное тождество (6.37) для чисел Эйлера.

15 Докажите тождество Эйлера (6.39), взяв m-ю разность (6.37).

16 Каково общее решение двойного рекуррентного соотношения

$$\begin{aligned} A_{n,0} &= a_n\,[n \geqslant 0]\,; \qquad A_{0,k} = 0\,, \qquad \text{если } k > 0; \\ A_{n,k} &= kA_{n-1,k} + A_{n-1,k-1}\,, \qquad \text{целые } k, n, \end{aligned}$$

когда k и n пробегают множество *всех* целых чисел?

17 Решите следующие рекуррентные соотношения, считая, что $\left|{n \atop k}\right|$ есть нуль при $n < 0$ или $k < 0$:

а $$\left|{n \atop k}\right| = \left|{n-1 \atop k}\right| + n\left|{n-1 \atop k-1}\right| + [n=k=0] \qquad \text{при } n, k \geqslant 0.$$

б $$\left|{n \atop k}\right| = (n-k)\left|{n-1 \atop k}\right| + \left|{n-1 \atop k-1}\right| + [n=k=0] \qquad \text{при } n, k \geqslant 0.$$

в $$\left|{n \atop k}\right| = k\left|{n-1 \atop k}\right| + k\left|{n-1 \atop k-1}\right| + [n=k=0] \qquad \text{при } n, k \geqslant 0.$$

18 Докажите, что полиномы Стирлинга удовлетворяют соотношениям

$$(x+1)\,\sigma_n(x+1) = (x-n)\,\sigma_n(x) + x\sigma_{n-1}(x)\,.$$

19 Докажите, что обобщенные числа Стирлинга удовлетворяют соотношениям

$$\sum_{k=0}^{n} \left\{{x+k \atop x}\right\} \left[{x \atop x-n+k}\right] (-1)^k \Big/ \binom{x+k}{n+1} = 0\,, \quad \text{целое } n > 0.$$

$$\sum_{k=0}^{n} \left[{x+k \atop x}\right] \left\{{x \atop x-n+k}\right\} (-1)^k \Big/ \binom{x+k}{n+1} = 0\,, \quad \text{целое } n > 0.$$

20 Выразите $\sum_{k=1}^{n} H_k^{(2)}$ в аналитическом виде.

21 Покажите, что если $H_n = a_n/b_n$, где a_n и b_n — целые, то знаменатель b_n кратен $2^{\lfloor \lg n \rfloor}$. *Указание*: рассмотрите число $2^{\lfloor \lg n \rfloor - 1} H_n - \frac{1}{2}$.

22 Докажите, что бесконечная сумма

$$\sum_{k \geqslant 1} \left(\frac{1}{k} - \frac{1}{k+z} \right)$$

сходится при любом комплексном числе z, за исключением случаев, когда z — отрицательное целое число, и покажите, что она равна H_z, когда z — неотрицательное целое число. (Таким образом, эту формулу можно использовать для определения гармонических чисел H_z при комплексных z.)

23 Коэффициенты разложения $z/(e^z-1)$ по степеням z задаются равенством (6.81). А каковы коэффициенты разложения $z/(e^z+1)$? *Указание*: рассмотрите тождество $(e^z+1)(e^z-1)=e^{2z}-1$.

24 Докажите, что тангенциальное число T_{2n+1} кратно 2^n. *Указание*: докажите, что все коэффициенты $T_{2n}(x)$ и $T_{2n+1}(x)$ кратны 2^n.

25 Равенство (6.57) доказывает, что в конце концов в некоторый момент времени N червяк достигнет конца резинки. Поэтому сначала должен наступить такой момент времени n, когда он будет ближе к концу резинки по истечении n минут, чем он был по истечении $n-1$ минут. Покажите, что $n < \frac{1}{2}N$.

26 Воспользуйтесь правилами суммирования по частям для вычисления суммы $S_n = \sum_{k=1}^{n} H_k/k$. *Указание*: рассмотрите, кроме того, похожую сумму $\sum_{k=1}^{n} H_{k-1}/k$.

27 Докажите НОД-правило (6.111) для чисел Фибоначчи.

28 *Числа Люка* L_n определяются как $F_{n+1}+F_{n-1}$. Таким образом, согласно (6.109) имеем $F_{2n}=F_nL_n$. Вот таблица нескольких их первых значений:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
L_n	2	1	3	4	7	11	18	29	47	76	123	199	322	521

а Воспользуйтесь методом наборов, чтобы показать, что решение Q_n обобщенного рекуррентного соотношения

$$Q_0 = \alpha; \quad Q_1 = \beta; \quad Q_n = Q_{n-1}+Q_{n-2}, \quad n>1$$

может быть выражено через F_n и L_n.

б Выразите L_n через ϕ и $\hat{\phi}$ в аналитическом виде.

29 Докажите тождество Эйлера для континуантов — равенство (6.134).

30 Обобщите (6.136) с тем, чтобы найти выражение для континуанта с приращением $K(x_1,\ldots,x_{m-1},x_m+y,x_{m+1},\ldots,x_n)$ при $1 \leqslant m \leqslant n$.

Домашние задания

31 Найдите аналитический вид коэффициентов $\left|{n \atop k}\right|$ в представлении возрастающих степеней через убывающие:

$$x^{\overline{n}} = \sum_k \left|{n \atop k}\right| x^{\underline{k}}, \qquad \text{целое } n \geqslant 0.$$

(Например, $x^{\overline{4}} = x^{\underline{4}} + 12x^{\underline{3}} + 36x^{\underline{2}} + 24x^{\underline{1}}$, следовательно, $\left|{4 \atop 2}\right| = 36$.)

32 В главе 5 мы получили формулы

$$\sum_{k\leqslant m} \binom{n+k}{k} = \binom{n+m+1}{m} \quad \text{и} \quad \sum_{0\leqslant k\leqslant m} \binom{k}{n} = \binom{m+1}{n+1},$$

развертывая рекуррентное соотношение $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ двумя способами. А какие получатся тождества при развертывании аналогичного рекуррентного соотношения $\left\{{n \atop k}\right\} = k\left\{{n-1 \atop k}\right\} + \left\{{n-1 \atop k-1}\right\}$?

33 В табл. 325 приведены значения $\left[{n \atop 2}\right]$ и $\left\{{n \atop 2}\right\}$. Как выглядят выражения в аналитическом виде (не включающие чисел Стирлинга) для последующих величин $\left[{n \atop 3}\right]$ и $\left\{{n \atop 3}\right\}$?

34 Чему равны $\left\langle{-1 \atop k}\right\rangle$ и $\left\langle{-2 \atop k}\right\rangle$, если считать, что основное рекуррентное соотношение (6.35) справедливо при любых целых k и n, и если $\left\langle{n \atop k}\right\rangle = 0$ при всех $k < 0$?

35 Докажите, что для любого $\epsilon > 0$ существует целое $n > 1$ (зависящее от ϵ), такое, что $H_n \bmod 1 < \epsilon$.

36 Можно ли сложить стопку из n кирпичей так, чтобы самый верхний кирпич полностью выступал над самым нижним кирпичом, и при этом человек, вес которого равен весу 100 кирпичей, мог находиться на средине верхнего кирпича, не обрушивая всю стопку?

37 Выразите $\sum_{k=1}^{mn} (k \bmod m)/k(k+1)$ через гармонические числа, считая, что m и n — положительные целые числа. Какова предельная величина данной суммы при $n \to \infty$?

38 Вычислите неопределенную сумму $\sum \binom{r}{k}(-1)^k H_k\, \delta k$.

39 Выразите сумму $\sum_{k=1}^{n} H_k^2$ через n и H_n.

40 Докажите, что 1979 делит числитель суммы $\sum_{k=1}^{1319}(-1)^{k-1}/k$, и укажите аналогичную сумму для 1987 (и 2011). *Указание*: воспользуйтесь уловкой Гаусса для получения суммы дробей, числители которых равны 1979. (См. также упр. 4.)

Ага! Это все — простые года!

41 Вычислите сумму

$$\sum_k \binom{\lfloor (n+k)/2 \rfloor}{k}$$

в аналитическом виде, если n — целое (возможно, отрицательное) число.

42 Если S — некоторое множество целых чисел, то пусть $S+1$ представляет собой "сдвинутое" множество $\{x+1 \mid x \in S\}$. Сколько подмножеств множества $\{1, 2, \ldots, n\}$ обладают тем свойством, что $S \cup (S+1) = \{1, 2, \ldots, n+1\}$?

43 Докажите, что бесконечная сумма

$$\begin{aligned} &.1 \\ +&.01 \\ +&.002 \\ +&.0003 \\ +&.00005 \\ +&.000008 \\ +&.0000013 \\ &\ \ \vdots \end{aligned}$$

сходится к рациональному числу.

44 Докажите обращение тождества Кассини (6.106): если k и m — целые числа, такие, что $|m^2 - km - k^2| = 1$, то существует целое n, такое, что $k = \pm F_n$ и $m = \pm F_{n+1}$.

45 Воспользуйтесь методом наборов для решения обобщенного рекуррентного соотношения

$$X_0 = \alpha; \qquad X_1 = \beta; \qquad X_n = X_{n-1} + X_{n-2} + \gamma n + \delta.$$

46 Чему равны $\cos 36°$ и $\cos 72°$?

47 Покажите, что

$$2^{n-1} F_n = \sum_k \binom{n}{2k+1} 5^k,$$

и воспользуйтесь данным соотношением для установления величин $F_p \bmod p$ и $F_{p+1} \bmod p$ при простом p.

48 Докажите, что нулевые параметры континуантов могут быть удалены путем стягивания их соседей:

$$\begin{aligned} &K_n(x_1, \ldots, x_{m-1}, 0, x_{m+1}, \ldots, x_n) = \\ &\quad = K_{n-2}(x_1, \ldots, x_{m-2}, x_{m-1}+x_{m+1}, x_{m+2}, \ldots, x_n), \\ &\qquad 1 < m < n. \end{aligned}$$

49 Найдите представление числа $\sum_{n\geqslant 1} 2^{-\lfloor n\phi \rfloor}$ в виде непрерывной дроби.

50 Определим функцию $f(n)$ для всех положительных целых чисел n при помощи рекуррентного соотношения

$$\begin{aligned} f(1) &= 1\,; \\ f(2n) &= f(n)\,; \\ f(2n+1) &= f(n) + f(n+1)\,. \end{aligned}$$

а При каких n значение $f(n)$ четное?

б Покажите, что функция $f(n)$ может быть выражена через континуанты.

Контрольные работы

51 Пусть p — простое число.

а Докажите, что $\left\{ {p \atop k} \right\} \equiv \left[{p \atop k} \right] \equiv 0 \pmod p$ при $1 < k < p$.

б Докажите, что $\left[{p-1 \atop k} \right] \equiv 1 \pmod p$ при $1 \leqslant k < p$.

в Докажите, что $\left\{ {2p-2 \atop p} \right\} \equiv \left[{2p-2 \atop p} \right] \equiv 0 \pmod p$, если $p > 2$.

г Докажите, что если $p > 3$, то $\left[{p \atop 2} \right] \equiv 0 \pmod{p^2}$. *Указание*: рассмотрите $p^{\underline{p}}$.

52 Пусть H_n записано в виде несократимой дроби a_n/b_n.

а Докажите, что $p \backslash b_n \iff p \nmid a_{\lfloor n/p \rfloor}$, если p — простое.

б Найдите все $n > 0$, такие, что a_n делится на 5.

53 Найдите аналитический вид суммы $\sum_{k=0}^{m} \binom{n}{k}^{-1} (-1)^k H_k$ при $0 \leqslant m \leqslant n$. *Указание*: в упр. 5.42 имеется сумма без множителя H_k.

54 Пусть $n > 0$. Цель данного упражнения — показать, что знаменатель числа B_{2n} представляет собой произведение всех простых чисел p, таких, что $(p-1)\backslash(2n)$.

а Покажите, что $S_m(p) + \bigl[(p-1)\backslash m\bigr]$ кратно p, когда p — простое и $m > 0$.

б Используйте результат части (а) для того, чтобы показать, что

$$B_{2n} + \sum_p \frac{\bigl[(p-1)\backslash(2n)\bigr]}{p} = I_{2n} \quad \text{— целое число.}$$

Указание: достаточно доказать, что если p — некоторое простое число, то знаменатель дроби $B_{2n} + \bigl[(p-1)\backslash(2n)\bigr]/p$ не делится на p.

в Докажите, что знаменатель числа B_{2n} всегда является нечетным кратным 6 и что он равен 6 для бесконечного количества n.

55 Докажите (6.70) как следствие некоторого более общего тождества, суммируя

$$\sum_{0\leqslant k<n} \binom{k}{m}\binom{x+k}{k}$$

и дифференцируя по x.

56 Вычислите $\sum_{k\neq m} \binom{n}{k}(-1)^k k^{n+1}/(k-m)$ в аналитическом виде как функцию целых чисел m и n. (Это сумма по всем целым k, за исключением $k=m$.)

57 "Обернутые биномиальные коэффициенты порядка 5" определяются как

$$\left(\!\binom{n}{k}\!\right) = \left(\!\binom{n-1}{k}\!\right) + \left(\!\binom{n-1}{(k-1) \bmod 5}\!\right), \qquad n>0,$$

и $\left(\!\binom{0}{k}\!\right) = [k=0]$. Пусть Q_n — разность между наибольшим и наименьшим из этих чисел в n-й строке:

$$Q_n = \max_{0\leqslant k<5}\left(\!\binom{n}{k}\!\right) - \min_{0\leqslant k<5}\left(\!\binom{n}{k}\!\right).$$

Найдите и докажите соотношения между числами Q_n и числами Фибоначчи.

58 Запишите суммы $\sum_{n\geqslant 0} F_n^2 z^n$ и $\sum_{n\geqslant 0} F_n^3 z^n$ в аналитическом виде. Что можно сказать о величине $F_{n+1}^3 - 4F_n^3 - F_{n-1}^3$?

59 Докажите, что если m и n — положительные целые числа, то существует целое x, такое, что $F_x \equiv m \pmod{3^n}$.

60 Найдите все положительные целые числа n, такие, что либо F_n+1, либо F_n-1 являются простыми числами.

61 Докажите тождество

$$\sum_{k=0}^{n} \frac{1}{F_{2^k}} = 3 - \frac{F_{2^n-1}}{F_{2^n}}, \qquad \text{целое } n\geqslant 1.$$

Чему равна сумма $\sum_{k=0}^{n} 1/F_{3\cdot 2^k}$?

62 Пусть $A_n = \phi^n + \phi^{-n}$ и $B_n = \phi^n - \phi^{-n}$.

а Найдите константы α и β, такие, что $A_n = \alpha A_{n-1} + \beta A_{n-2}$ и $B_n = \alpha B_{n-1} + \beta B_{n-2}$ для всех $n \geqslant 0$.

б Выразите A_n и B_n через F_n и L_n (см. упр. 28).

в Докажите, что $\sum_{k=1}^{n} 1/(F_{2k+1}+1) = B_n/A_{n+1}$.

г Выразите в аналитическом виде сумму $\sum_{k=1}^{n} 1/(F_{2k+1}-1)$.

Лишние задачи? Запасные задачи?

Дополнительные задачи

63 Сколько перестановок $\pi_1\pi_2\ldots\pi_n$ множества $\{1,2,\ldots,n\}$ содержат ровно k индексов j, таких, что

а $\pi_i < \pi_j$ для всех $i < j$? (Такие j называются "левосторонними максимумами".)

б $\pi_j > j$? (Такие j называются "превышениями".)

64 Чему равен знаменатель дроби $\left[{1/2 \atop 1/2-n}\right]$, если привести ее к несократимому виду?

65 Докажите тождество

$$\int_0^1 \cdots \int_0^1 f\big(\lfloor x_1 + \cdots + x_n \rfloor\big)\, dx_1 \ldots dx_n \;=\; \sum_k \left\langle {n \atop k} \right\rangle \frac{f(k)}{n!}\,.$$

66 Чему равна сумма чисел n-й строки треугольника Эйлера с чередующимися знаками $\sum_k (-1)^k \left\langle {n \atop k} \right\rangle$?

67 Докажите, что

$$\sum_k \left\{ {n+1 \atop k+1} \right\} \binom{n-k}{m-k} (-1)^{m-k} k! \;=\; \left\langle {n \atop m} \right\rangle.$$

68 Покажите, что $\left\langle\!\left\langle {n \atop 1} \right\rangle\!\right\rangle = 2\left\langle {n \atop 1} \right\rangle$, и найдите аналитический вид $\left\langle\!\left\langle {n \atop 2} \right\rangle\!\right\rangle$.

69 Запишите $\sum_{k=1}^{n} k^2 H_{n+k}$ в аналитическом виде.

70 Покажите, что комплексные гармонические числа из упр. 22 раскладываются в степенной ряд $H_z = \sum_{n\geqslant 2} (-1)^n H_\infty^{(n)} z^{n-1}$.

71 Докажите, что обобщенный факториал из уравнения (5.83) может быть записан как

$$\prod_{k\geqslant 1} \left(1 + \frac{z}{k}\right) e^{-z/k} \;=\; \frac{e^{-\gamma z}}{z!}\,,$$

рассмотрев предел при $n \to \infty$ первых n сомножителей данного бесконечного произведения. Покажите, что $\frac{d}{dz}(z!)$ связано с обобщенными гармоническими числами из упр. 22.

72 Докажите, что функция "тангенс" раскладывается в степенной ряд (6.92), и найдите соответствующие ряды для $z/\sin z$ и $\ln\big((\operatorname{tg} z)/z\big)$.

73 Докажите, что $z \operatorname{ctg} z$ равно

$$\frac{z}{2^n} \operatorname{ctg} \frac{z}{2^n} - \frac{z}{2^n} \operatorname{tg} \frac{z}{2^n} + \sum_{k=1}^{2^{n-1}-1} \frac{z}{2^n} \left(\operatorname{ctg} \frac{z+k\pi}{2^n} + \operatorname{ctg} \frac{z-k\pi}{2^n} \right)$$

при любом целом $n \geqslant 1$, и покажите, что при фиксированном k и $n \to \infty$ предел k-го общего члена равен $2z^2/(z^2 - k^2\pi^2)$.

74 Найдите соотношение между числами $T_n(1)$ и коэффициентами разложения $1/\cos z$.

75 Докажите, что тангенциальные числа и коэффициенты разложения $1/\cos z$ располагаются по сторонам бесконечного треугольника, который начинается следующим образом:

$$\begin{array}{ccccccccccccc}
&&&&&&1&&&&&&\\
&&&&&0&&1&&&&&\\
&&&&1&&1&&0&&&&\\
&&&0&&1&&2&&2&&&\\
&&5&&5&&4&&2&&0&&\\
&0&&5&&10&&14&&16&&16&\\
61&&61&&56&&46&&32&&16&&0
\end{array}$$

Каждая строка содержит частичные суммы предыдущей строки, поочередно слева направо и справа налево. *Указание*: рассмотрите коэффициенты степенного ряда для $(\sin z + \cos z)/\cos(w + z)$.

76 Выразите сумму

$$\sum_k (-1)^k \left\{ {n \atop k} \right\} 2^{n-k} k!$$

в аналитическом виде и покажите, что она равна нулю при четном n.

77 При целых m и n, $n \geqslant 0$, формула (6.52) дает значение $\sigma_n(m)$ при $m < 0$, формула (6.53) — при $m > n$, а формула (6.101) — при $m = 0$. Покажите, что в остальных случаях имеет место равенство

$$\sigma_n(m) = \frac{(-1)^{m+n-1}}{m!\,(n-m)!} \sum_{k=0}^{m-1} \left[{m \atop m-k} \right] \frac{B_{n-k}}{n-k}, \qquad \text{целые } n \geqslant m > 0.$$

78 Докажите следующее соотношение, которое связывает числа Стирлинга, Бернулли и Каталана:

$$\sum_{k=0}^{n} \left\{ {n+k \atop k} \right\} \binom{2n}{n+k} \frac{(-1)^k}{k+1} = B_n \binom{2n}{n} \frac{1}{n+1}.$$

79 Покажите, что части шахматной доски из парадокса “64 = 65” можно разместить так, что получится парадокс “64 = 63”.

80 Последовательность чисел, определяемая рекуррентным соотношением $A_1 = x$, $A_2 = y$ и $A_n = A_{n-1} + A_{n-2}$, содержит $A_m = 1000000$ при некотором m. Какие положительные целые числа x и y приводят к максимально возможному m?

81 В этой главе описан способ сведения формулы, содержащей числа $F_{n\pm k}$, к формуле, содержащей только числа F_n и F_{n+1}. Естественно заинтересоваться, могут ли такие "приведенные" формулы быть равными, если они отличаются внешним видом. Пусть $P(x, y)$ — полином от x и y с целочисленными коэффициентами. Найдите необходимое и достаточное условие того, что $P(F_{n+1}, F_n) = 0$ для всех $n \geqslant 0$.

82 Объясните, как выполняется сложение положительных целых чисел в фибоначчиевой системе счисления.

83 Возможно ли, чтобы последовательность $\langle A_n \rangle$, удовлетворяющая рекуррентному соотношению Фибоначчи $A_n = A_{n-1} + A_{n-2}$, не содержала простых чисел, если A_0 и A_1 — взаимно простые числа?

84 Пусть m и n — нечетные положительные целые числа. Выразите в аналитическом виде

$$S^+_{m,n} = \sum_{k\geqslant 0} \frac{1}{F_{2mk+n}+F_m}\,; \quad S^-_{m,n} = \sum_{k\geqslant 0} \frac{1}{F_{2mk+n}-F_m}\,.$$

Указание: суммы из упр. 62 — это $S^+_{1,3} - S^+_{1,2n+3}$ и $S^-_{1,3} - S^-_{1,2n+3}$.

85 Охарактеризуйте все N, такие, что вычеты чисел Фибоначчи $F_n \bmod N$ для $n \geqslant 0$ образуют полное множество $\{0, 1, \ldots, N-1\}$. (См. упр. 59.)

86 Пусть C_1, C_2, … — последовательность ненулевых целых чисел, такая, что

$$\text{НОД}(C_m, C_n) = C_{\text{НОД}(m,n)}$$

для всех положительных целых чисел m и n. Докажите, что все обобщенные биномиальные коэффициенты

$$\binom{n}{k}_{\mathcal{C}} = \frac{C_n C_{n-1} \ldots C_{n-k+1}}{C_k C_{k-1} \ldots C_1}$$

являются целыми числами. (В частности, "коэффициенты Фибоначчи", образованные таким способом из чисел Фибоначчи, целые ввиду (6.111).)

87 Покажите, что K-полиномы представимы в виде произведения матриц

$$\begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_n \end{pmatrix}$$

и в виде определителя

$$\det \begin{pmatrix} x_1 & 1 & 0 & 0 & \dots & 0 \\ -1 & x_2 & 1 & 0 & & 0 \\ 0 & -1 & x_3 & 1 & & \vdots \\ \vdots & & -1 & & & \\ & & & & \ddots & 1 \\ 0 & 0 & \dots & & -1 & x_n \end{pmatrix}.$$

88 Обобщая (6.146), найдите непрерывную дробь, связанную с производящей функцией $\sum_{n\geqslant 1} z^{\lfloor n\alpha \rfloor}$, где α — некоторое положительное иррациональное число.

89 Пусть α — некоторое иррациональное число из интервала $(0\,..\,1)$ и пусть a_1, a_2, a_3, ... — неполные частные в его представлении в виде непрерывной дроби. Покажите, что $|D(\alpha, n)| < 2$, где $n = K(a_1, \dots, a_m)$, а D — отклонение, определенное в главе 3.

90 Пусть Q_n — наибольший знаменатель на n-м уровне дерева Штерна–Броко. (Таким образом, в соответствии со схемой из главы 4 $\langle Q_0, Q_1, Q_2, Q_3, Q_4, \dots \rangle = \langle 1, 2, 3, 5, 8, \dots \rangle$.) Докажите, что $Q_n = F_{n+2}$.

Исследовательские проблемы

91 Как лучше всего распространить определение $\left\{ {n \atop k} \right\}$ на произвольные действительные значения n и k?

92 Пусть число H_n записано в виде несократимой дроби a_n/b_n, как в упр. 52.

а Бесконечно ли много таких n, что $p \backslash a_n$ при некотором фиксированном простом p?

б Бесконечно ли много таких n, что $b_n = \text{НОК}(1, 2, \dots, n)$? (Двумя подобными n являются $n = 250$ и $n = 1000$.)

93 Докажите, что числа γ и e^γ иррациональны.

94 Разработайте общую теорию решения двухпараметрического рекуррентного соотношения

$$\left| {n \atop k} \right| = (\alpha n + \beta k + \gamma) \left| {n-1 \atop k} \right| + \\ + (\alpha' n + \beta' k + \gamma') \left| {n-1 \atop k-1} \right| + [n = k = 0] \quad \text{при } n, k \geqslant 0$$

в предположении, что $\left|{n \atop k}\right| = 0$ при $n < 0$ или $k < 0$. (Биномиальные коэффициенты, числа Стирлинга, числа Эйлера и последовательности чисел из упр. 17 и 31 — это все частные случаи данного рекуррентного соотношения.) Какие специальные значения $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$ дают "фундаментальные решения", через которые может быть выражено общее решение?

95 Найдите эффективный способ распространения алгоритма Госпера–Зайльбергера с гипергеометрических членов на члены, содержащие числа Стирлинга.

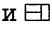
7

Производящие функции

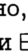
НАИБОЛЕЕ МОЩНЫЙ МЕТОД работы с числовыми последовательностями, как известно, заключается в преобразовании бесконечных рядов, "генерирующих" эти последовательности. Мы уже изучили большое количество последовательностей и встречались с некоторыми производящими функциями; теперь мы готовы к более глубокому исследованию производящих функций, которое покажет нам их исключительную полезность.

7.1 Теория домино и размен

Производящие функции достаточно важны, а для многих из нас достаточно новы, чтобы оправдать небольшой отдых перед тем, как заняться ими вплотную. Поэтому попытаемся в начале главы развить интуицию в отношении производящих функций при помощи игр и забав. Рассмотрим два приложения производящих функций — одно, связанное с домино, и второе, связанное с монетами.

Каково количество T_n способов покрытия прямоугольника размером $2 \times n$ костями домино размером 2×1? Считаем, что все кости домино идентичны (например, потому что они лежат лицевой стороной вниз или потому что кто-то сделал их неразличимыми, выкрасив в красный цвет); следовательно, имеет значение только ориентация костяшки — вертикальная или горизонтальная; мы можем просто представить себе, что работаем с одинаковыми плитками в форме домино. Например, существует три покрытия прямоугольника размером 2×3, а именно , , и ; так что $T_3 = 3$.

"Сочти пути".
— Е. Б. Браунинг
(E. B. Browning)

Чтобы найти общее выражение для T_n в аналитическом виде, как обычно, начнем с рассмотрения малых значений. При $n = 1$ существует, очевидно, только одно покрытие, ; при $n = 2$ имеется два покрытия,  и .

А что можно сказать о случае $n = 0$? Сколько существует покрытий прямоугольника размером 2×0? Сразу не ясно, как понимать этот вопрос, но раньше мы уже встречались с подобными ситуациями. Так, существует одна перестановка нуля объектов (а именно — пустая перестановка), так что $0! = 1$. Существует единственный способ выбрать нуль предметов из n (а именно — не выбрать ничего), поэтому $\binom{n}{0} = 1$. Существует единственный способ разбить пустое множество на нуль непустых подмножеств, но нет ни одного такого способа для разбиения непустого множества; так что $\left\{ {n \atop 0} \right\} = [n = 0]$. Рассуждая аналогично, можно сделать заключение, что существует единственный способ покрыть костями домино прямоугольник размером 2×0, а именно — не использовать домино вообще; таким образом, $T_0 = 1$. (Это противоречит простой формуле $T_n = n$, которая выполняется при $n = 1, 2$ и 3; по-видимому, придется эту простую формулу отвергнуть, поскольку T_0 должно быть равно 1 в соответствии с логикой ситуации.) Правильное понимание нулевого случая — залог верного решения задачи перечисления.

Давайте рассмотрим еще один малый случай, $n = 4$. Имеются две возможности покрытия левого конца прямоугольника — там можно положить либо одну кость домино вертикально, либо две горизонтально. В случае вертикальной кости домино получается частичное решение , и остающийся прямоугольник размером 2×3 может быть покрыт T_3 способами. Если выбрать две горизонтальные кости домино, то частичное решение  можно завершить T_2 способами. Таким образом, $T_4 = T_3 + T_2 = 5$. (Вот эти пять покрытий: , , ,  и .)

Теперь нам известны пять первых значений T_n:

n	0	1	2	3	4
T_n	1	1	2	3	5

Все это подозрительно похоже на числа Фибоначчи, и нетрудно сообразить, почему. Рассуждения, с помощью которых мы доказали, что $T_4 = T_3 + T_2$, легко обобщаются на $T_n = T_{n-1} + T_{n-2}$ для $n \geqslant 2$. Таким образом, здесь получается то же самое рекуррентное соотношение, что и для чисел Фибоначчи, но с немного иными начальными значениями $T_0 = 1$ и $T_1 = 1$. Но эти начальные значения представляют собой последовательные числа Фибоначчи F_1 и F_2, так что T являются числами Фибоначчи, сдвинутыми на одну позицию:

$$T_n = F_{n+1} \qquad \text{при } n \geqslant 0.$$

(Это выражение можно рассматривать как аналитический вид T_n, потому что числа Фибоначчи достаточно важны для того, чтобы рассматривать их как "известные". Кроме того, сами F_n имеют выражение в аналитическом виде (6.123), использующее алгебраические операции.) Заметим, что это равенство подтверждает мудрость принятия $T_0 = 1$.

Смело вперед по дороге, вымощенной желтыми домино!..

Но какое отношение имеет все это к производящим функциям? Мы как раз приближаемся к тому, чтобы ими воспользоваться (и это будет еще один способ вывода T_n). Этот новый способ основан на одной весьма смелой идее. Рассмотрим "сумму" всех возможных расположений костей домино в $(2 \times n)$-прямоугольнике при всех $n \geqslant 0$ и назовем ее T:

$$T = \text{|} + \text{▯} + \text{▯▯} + \text{⊟} + \text{▯▯▯} + \text{▯⊟} + \text{⊟▯} + \cdots . \tag{7.1}$$

(Первый член, '|', в правой части означает пустое покрытие прямоугольника размером 2×0.) Сумма T содержит массу информации. Ее польза в том, что можно доказать некоторое свойство T в целом, вместо того, чтобы выполнять доказательство (по индукции) для отдельных слагаемых.

Членами данной суммы являются покрытия, которые представляют собой комбинаторные объекты. Не будем задумываться над вопросом, насколько допустимо суммирование бесконечного числа покрытий. Все можно сделать совершенно строго, но в настоящий момент цель состоит в том, чтобы выйти в своем сознании за рамки привычных алгебраических формул.

Мы сложили покрытия, но их можно и умножать, приписывая одно за другим. Например, можно умножить покрытие ▯ на покрытие ⊟ и получить покрытие ▯⊟. Заметьте, однако, что умножение некоммутативно, т.е. следует учитывать порядок умножения: ▯⊟ отличается от ⊟▯.

При использовании таких обозначений для умножения нетрудно увидеть, что пустое покрытие играет особую роль — оно представляет собой мультипликативную единицу, например $\text{|} \times \text{⊟} = \text{⊟} \times \text{|} = \text{⊟}$.

Теперь можно использовать "доминошную арифметику" для работы с бесконечной суммой T:

$$\begin{aligned} T &= \text{|} + \text{▯} + \text{▯▯} + \text{⊟} + \text{▯▯▯} + \text{▯⊟} + \text{⊟▯} + \cdots = \\ &= \text{|} + \text{▯}(\text{|} + \text{▯} + \text{▯▯} + \text{⊟} + \cdots) + \text{⊟}(\text{|} + \text{▯} + \text{▯▯} + \text{⊟} + \cdots) = \\ &= \text{|} + \text{▯}T + \text{⊟}T . \end{aligned} \tag{7.2}$$

В правых частях этих равенств каждое правильное покрытие содержится ровно по одному разу, так что выполненные преобразования вполне осмысленны, несмотря на предостережение из главы 2 об "абсолютной сходимости". Последняя строка уравнения гласит, что любое покрытие в T представляет собой либо пустое покрытие, либо вертикальную кость домино, за которой следует что-то из T, либо две горизонтальные кости домино, за которыми следует что-то из T.

Нутром чую, что эти суммы должны сходиться, если кости домино будут достаточно малых размеров!

Давайте теперь попытаемся разрешить это уравнение относительно T. Заменим T в левой части на IT и вычтем два последних члена справа, мы получим

(I − ▯ − ⊟)T = I. (7.3)

Для проверки раскроем скобки:

  I + ▯ + ▯▯ + ⊟ + ▯▯▯ + ▯⊟ + ⊟▯ + ⋯
− ▯ − ▯▯ − ▯▯▯ − ▯⊟ − ▯▯▯▯ − ▯▯⊟ − ▯⊟▯ − ⋯
− ⊟ − ⊟▯ − ⊟▯▯ − ⊟⊟ − ⊟▯▯▯ − ⊟▯⊟ − ⊟⊟▯ − ⋯
———————————————————————
  I

Каждый член верхней строки, за исключением первого, сокращается с каким-либо членом второй или третьей строки, так что наше уравнение вполне корректно.

До сих пор нам было сравнительно легко придавать комбинаторный смысл уравнениям, с которыми мы работали. Однако теперь, чтобы получить компактное выражение для T, выполним комбинаторное деление. Если слепо положиться на алгебраические преобразования, можно поделить обе части уравнения (7.3) на I − ▯ − ⊟ и получить

T = I / (I − ▯ − ⊟). (7.4)

(Умножение, как уже говорилось, некоммутативно, так что, не различая правое и левое деления, мы практически жульничаем. Но в данном случае это неважно, поскольку I коммутирует с чем угодно. Не будем слишком придирчивы, пока наши заумные идеи не привели к парадоксу.)

Следующий шаг состоит в развертывании этой дроби в степенной ряд с использованием правила

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \cdots .$$

Пустое покрытие I, являющееся мультипликативной единицей в нашей комбинаторной арифметике, играет роль обычной мультипликативной единицы 1; а ▯ + ⊟ играет роль z. Так что мы получаем разложение

$$\frac{\text{|}}{\text{|}-\text{▯}-\text{⊟}} = \text{|}+(\text{▯}+\text{⊟})+(\text{▯}+\text{⊟})^2+(\text{▯}+\text{⊟})^3+\cdots =$$
$$= \text{|}+(\text{▯}+\text{⊟})+(\text{▯▯}+\text{▯⊟}+\text{⊟▯}+\text{⊟⊟})+$$
$$+(\text{▯▯▯}+\text{▯▯⊟}+\text{▯⊟▯}+\text{▯⊟⊟}+\text{⊟▯▯}+\text{⊟▯⊟}+\text{⊟⊟▯}+\text{⊟⊟⊟})+\cdots.$$

Это T, но теперь покрытия расположены в ином порядке, не так, как ранее. Каждое покрытие встречается в этой сумме ровно один раз; так, ▯⊟⊟▯▯⊟▯ содержится в $(\text{▯}+\text{⊟})^7$.

Можно извлечь из этой суммы полезную информацию, если сжать ее, игнорируя детали, не представляющие интереса. Например, можно представить себе, что отдельные кости домино не склеены друг с другом и коммутируют между собой; тогда покрытие ▯⊟⊟▯▯⊟▯ превращается в $\text{▯}^4\text{▭}^6$, потому что оно содержит четыре вертикальные и шесть горизонтальных костей домино. Приведение подобных членов дает ряд

$$\text{T} = \text{|}+\text{▯}+\text{▯}^2+\text{▭}^2+\text{▯}^3+2\text{▯}\text{▭}^2+\text{▯}^4+3\text{▯}^2\text{▭}^2+\text{▭}^4+\cdots.$$

Здесь слагаемое $2\text{▯}\text{▭}^2$ представляет два члена старого разложения, ▯⊟ и ⊟▯, имеющих по одной вертикальной и по две горизонтальные кости домино; аналогично $3\text{▯}^2\text{▭}^2$ представляет три члена, ▯▯⊟, ▯⊟▯ и ⊟▯▯. По сути, мы рассматриваем ▯ и ▭ как обычные (коммутативные) переменные.

Выразить коэффициенты в коммутативной версии T в аналитическом виде можно, воспользовавшись биномиальной теоремой:

$$\frac{\text{|}}{\text{|}-(\text{▯}+\text{▭}^2)} = \text{|}+(\text{▯}+\text{▭}^2)+(\text{▯}+\text{▭}^2)^2+(\text{▯}+\text{▭}^2)^3+\cdots =$$
$$= \sum_{k\geqslant 0}(\text{▯}+\text{▭}^2)^k =$$
$$= \sum_{j,k\geqslant 0}\binom{k}{j}\text{▯}^j\text{▭}^{2k-2j} =$$
$$= \sum_{j,m\geqslant 0}\binom{j+m}{j}\text{▯}^j\text{▭}^{2m}. \qquad (7.5)$$

(На последнем шаге $k-j$ заменяется на m; это корректная замена, потому что $\binom{k}{j}=0$ при $0\leqslant k<j$.) Таким образом, мы заключаем, что $\binom{j+m}{j}$ есть число способов покрытия прямоугольника размером $2\times(j+2m)$ с помощью j вертикальных и $2m$ горизонтальных костей домино. Например, мы только что рассматривали покрытие 2×10 прямоугольника размером ▯⊟⊟▯▯⊟▯, которое включает четыре вертикальные и шесть горизонтальных

костей домино; всего имеется $\binom{4+3}{4} = 35$ таких покрытий, следовательно, одним из членов в коммутативном варианте T является $35▯^4▭^6$.

Можно отбросить еще некоторые детали, если не обращать внимания на ориентацию костей домино. Предположим, что нас не интересуют вертикальные и горизонтальные кости домино и мы хотим знать только общее количество покрытий $2 \times n$. (Это как раз и есть то самое число T_n, с которого мы начинали наше исследование.) Необходимую информацию можно получить путем простой подстановки одной величины z вместо ▯ и ▭. Можно также заменить I на 1 и получить

Теперь, когда нас не интересует ориентация, я дезориентирован...

$$T = \frac{1}{1 - z - z^2}. \tag{7.6}$$

Это — производящая функция (6.117) для чисел Фибоначчи, за исключением того, что в числителе отсутствует множитель z; поэтому мы заключаем, что коэффициент при z^n в T равен F_{n+1}.

Компактные представления $\mathrm{I}/(\mathrm{I}-▯-⊟)$, $\mathrm{I}/(\mathrm{I}-▯-▭^2)$ и $1/(1-z-z^2)$, выведенные нами для T, называются *производящими функциями*, потому что они генерируют интересующие нас коэффициенты.

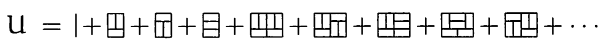
Кстати, наш вывод позволяет заключить, что количество покрытий прямоугольника размером $2 \times n$ ровно m парами горизонтальных костей домино равно $\binom{n-m}{m}$. (Действительно, в покрытии будет $j = n - 2m$ вертикальных костей домино, следовательно, в соответствии с нашей формулой найдется $\binom{j+m}{j} = \binom{j+m}{m} = \binom{n-m}{m}$ способов осуществить покрытие.) В главе 6 мы видели, что $\binom{n-m}{m}$ представляет собой число букв азбуки Морзе длиной n, содержащих m тире; действительно, нетрудно увидеть прямое соответствие между $(2\times n)$-покрытиями и буквами азбуки Морзе. (Покрытие ▯⊟⊟▯▯⊟ соответствует •−−••−•.) Таким образом, покрытия костями домино тесно связаны с континуантами, которые мы изучали в главе 6. Мир тесен...

Задачу вычисления T_n мы решили двумя способами. Первый способ, заключающийся в том, чтобы угадать ответ и доказать его по индукции, проще; второй способ — использовать бесконечную сумму покрытий и извлечь из нее интересующие нас коэффициенты — более изощренный. Почему же нам понадобился второй способ? Неужели только потому, что с костями домино очень интересно играть, представляя их алгебраическими переменными? Конечно, нет; истинная причина использования второго способа в том, что метод бесконечных сумм обладает существенно бо́льшими возможностями. Он примени́м к гораздо большему количеству задач, потому что не требует от нас сверхъестественных догадок.

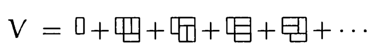
Давайте поднимемся еще на одну ступень обобщения, занявшись задачей, угадать ответ которой будет выше наших сил. Каково количество U_n способов покрытия костями домино прямоугольника размером $3 \times n$?

Первые несколько случаев дают нам мало пищи для размышлений. Пустое покрытие дает $U_0 = 1$. При $n = 1$ допустимых покрытий не существует, поскольку одна кость домино размером 2×1 не заполнит прямоугольник размером 3×1, а для двух в нем не хватит места. Следующий случай, $n = 2$, легко разрешить вручную: имеется три покрытия, ,  и , так что $U_2 = 3$. (Кстати, из решения предыдущей задачи мы уже знаем, что $T_3 = 3$, а количество способов покрытия прямоугольника размером 3×2 такое же, как и количество покрытий прямоугольника размером 2×3.) При $n = 3$, как и при $n = 1$, покрытий не существует. В этом можно убедиться, перебрав все возможные варианты... или перейдя на более высокий уровень и просто сообразив, что площадь прямоугольника размером 3×3 нечетная, и никак не может быть покрыта костями домино с четной площадью. (Очевидно, что те же соображения применимы в случае любого нечетного n.) Наконец, когда $n = 4$, имеется около дюжины покрытий; чтобы найти точное число, надо потратить изрядное количество времени, проверяя полноту списка.

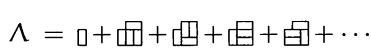
Поэтому попробуем подход с бесконечной суммой, который успешно сработал в прошлый раз:

 (7.7)

Каждое непустое покрытие начинается либо с , либо с  или ; но, к сожалению, в первых двух из них нельзя просто вынести множитель и снова прийти к U. Сумма всех членов в U, начинающихся с , может быть записана в виде V, где



есть сумма всех покрытий "выщербленного" прямоугольника размером $3 \times n$, в котором отсутствует левый нижний квадратик. Аналогично члены U, начинающиеся с , можно записать как Λ, где



состоит из всех покрытий прямоугольника без верхнего левого угла. Ряд Λ представляет собой зеркальное отражение V. Эти разложения позволяют записать

$$\mathrm{U} = | + \text{▯}_{\text{▭}}\mathrm{V} + \text{▯}^{\text{▭}}\Lambda + \text{☰}\mathrm{U}.$$

Точно так же разлагаются сами V и Λ, потому что входящие в них покрытия могут начинаться лишь двумя способами:

$$\begin{aligned}
\mathrm{V} &= \text{▯}\mathrm{U} + \genfrac{}{}{0pt}{}{\text{▭}\,}{\,\text{▭}}\mathrm{V},\\
\Lambda &= \text{▯}\mathrm{U} + \genfrac{}{}{0pt}{}{\,\text{▭}}{\text{▭}\,}\Lambda.
\end{aligned}$$

Теперь у нас есть три уравнения с тремя неизвестными (U, V и Λ). Их можно решить, выразив сначала V и Λ через U, а затем подставив результаты в уравнение для U:

$$\begin{aligned}
\mathrm{V} &= \Bigl(|-\genfrac{}{}{0pt}{}{\text{▭}\,}{\,\text{▭}}\Bigr)^{-1}\text{▯}\mathrm{U}, \qquad \Lambda = \Bigl(|-\genfrac{}{}{0pt}{}{\,\text{▭}}{\text{▭}\,}\Bigr)^{-1}\text{▯}\mathrm{U};\\
\mathrm{U} &= | + \text{▯}_{\text{▭}}\Bigl(|-\genfrac{}{}{0pt}{}{\text{▭}\,}{\,\text{▭}}\Bigr)^{-1}\text{▯}\mathrm{U} + \text{▯}^{\text{▭}}\Bigl(|-\genfrac{}{}{0pt}{}{\,\text{▭}}{\text{▭}\,}\Bigr)^{-1}\text{▯}\mathrm{U} + \text{☰}\mathrm{U}.
\end{aligned}$$

Последнее уравнение можно разрешить относительно U, получив компактную формулу

$$\mathrm{U} = \frac{|}{| - \text{▯}_{\text{▭}}\Bigl(|-\genfrac{}{}{0pt}{}{\text{▭}\,}{\,\text{▭}}\Bigr)^{-1}\text{▯} - \text{▯}^{\text{▭}}\Bigl(|-\genfrac{}{}{0pt}{}{\,\text{▭}}{\text{▭}\,}\Bigr)^{-1}\text{▯} - \text{☰}}. \tag{7.8}$$

Это выражение определяет бесконечную сумму U, так же как (7.4) определяет T.

Следующий шаг состоит в использовании коммутативности. Все прекрасно упрощается, если рассоединить все кости домино и использовать только степени ▯ и ▭:

$$\begin{aligned}
\mathrm{U} &= \frac{1}{1 - \text{▯}^2\text{▭}(1-\text{▭}^3)^{-1} - \text{▯}^2\text{▭}(1-\text{▭}^3)^{-1} - \text{▭}^3} =\\
&= \frac{1-\text{▭}^3}{(1-\text{▭}^3)^2 - 2\text{▯}^2\text{▭}} =\\
&= \frac{(1-\text{▭}^3)^{-1}}{1-2\text{▯}^2\text{▭}(1-\text{▭}^3)^{-2}} =\\
&= \frac{1}{1-\text{▭}^3} + \frac{2\text{▯}^2\text{▭}}{(1-\text{▭}^3)^3} + \frac{4\text{▯}^4\text{▭}^2}{(1-\text{▭}^3)^5} + \frac{8\text{▯}^6\text{▭}^3}{(1-\text{▭}^3)^7} + \cdots =\\
&= \sum_{k\geqslant 0}\frac{2^k\text{▯}^{2k}\text{▭}^k}{(1-\text{▭}^3)^{2k+1}} =\\
&= \sum_{k,m\geqslant 0}\binom{m+2k}{m}2^k\text{▯}^{2k}\text{▭}^{k+3m}.
\end{aligned}$$

(Этот вывод заслуживает тщательной проверки. На последнем шаге использована формула $(1-w)^{-2k-1} = \sum_m \binom{m+2k}{m}w^m$, тож-

Из другого курса я узнал о "регулярных выражениях". Если я не ошибаюсь, то на языке регулярных выражений можно записать

$$\mathrm{U} = \Bigl(\text{▯}_{\text{▭}}\genfrac{}{}{0pt}{}{\text{▭}\,}{\,\text{▭}}^{*}\text{▯} + \text{▯}^{\text{▭}}\genfrac{}{}{0pt}{}{\,\text{▭}}{\text{▭}\,}^{*}\text{▯} + \text{☰}\Bigr)^{*};$$

поэтому должна существовать связь между регулярными выражениями и производящими функциями.

дество (5.56).) Посмотрим внимательно на нижнюю строку и попытаемся понять, что она может нам сказать. Она говорит, что любое покрытие размером $3 \times n$ использует четное количество вертикальных костей домино. Кроме того, если имеется $2k$ вертикальных костей домино, то горизонтальных костей домино должно быть не менее k, и общее количество горизонтальных костей домино должно равняться $k+3m$ для некоторого $m \geqslant 0$. Наконец, количество возможных покрытий с $2k$ вертикальными и $k+3m$ горизонтальными костями домино равно $\binom{m+2k}{m}2^k$.

Теперь мы можем проанализировать 3×4-покрытия, которые вызвали затруднения в начале рассмотрения задачи о покрытии прямоугольника размером $3\times n$. При $n = 4$ общая площадь равна 12, так что всего нам потребуется шесть костей домино. Из них $2k$ вертикальных и $k+3m$ горизонтальных для некоторых k и m; следовательно, $2k+k+3m = 6$. Другими словами, $k+m = 2$. Если мы не используем вертикальные кости домино, то $k = 0$ и $m = 2$; количество вариантов равно $\binom{2+0}{2}2^0 = 1$. (Здесь учтено покрытие .) Если использовать две вертикальные кости домино, то $k = 1$ и $m = 1$; всего имеется $\binom{1+2}{1}2^1 = 6$ таких покрытий. А если использовать четыре вертикальные кости домино, то $k = 2$ и $m = 0$; таких покрытий имеется $\binom{0+4}{0}2^2 = 4$, так что общее количество покрытий — $U_4 = 11$. В общем случае при четном n те же рассуждения показывают, что $k+m = \frac{1}{2}n$, следовательно, $\binom{m+2k}{m} = \binom{n/2+k}{n/2-k}$ и общее количество $(3 \times n)$-покрытий равно

$$U_n = \sum_k \binom{n/2+k}{n/2-k} 2^k = \sum_m \binom{n-m}{m} 2^{n/2-m}. \qquad (7.9)$$

Как и ранее, можно заменить символы ▯ и ▭ на z, получив при этом производящую функцию без дискриминации костей домино за их убеждения. Результатом будет функция

$$U = \frac{1}{1-z^3(1-z^3)^{-1}-z^3(1-z^3)^{-1}-z^3} = \frac{1-z^3}{1-4z^3+z^6}. \qquad (7.10)$$

Если разложить эту дробь в степенной ряд, получится

$$U = 1 + U_2 z^3 + U_4 z^6 + U_6 z^9 + U_8 z^{12} + \cdots,$$

производящая функция для чисел U_n. (Любопытное несовпадение индексов и показателей степеней в формуле легко объяснимо. Коэффициент при z^9, например, есть U_6, отвечающий за покрытия прямоугольника размером 3×6. Это именно то, чего мы хотели, потому что каждое такое покрытие содержит девять костей домино.)

Можно продолжить анализ (7.10) и получить коэффициенты в аналитическом виде, но лучше отложить это занятие, пока мы не наберемся опыта. Так что временно отложим домино и возьмемся за следующую задачу, посвященную "размену".

Сколькими способами можно заплатить 50 копеек, если платить копейками $\textcircled{1}$, пятаками $\textcircled{5}$, гривенниками $\textcircled{10}$, четвертаками $\textcircled{25}$ и полтинниками $\textcircled{50}$. Пойа (Pólya) [298] популяризовал эту задачу, продемонстрировав поучительный способ ее решения с применением производящих функций.

А ведь я помню времена, когда на копейку можно было купить коробку спичек или стакан газировки, а за пятак проехаться в метро...
— Переводчик

Запишем бесконечную сумму, которая представляет все возможные способы размена, так же, как ранее в задаче о домино бесконечные суммы представляли все возможные покрытия. Начать проще всего со случая, когда имеется меньше разновидностей монет, поэтому положим для начала, что у нас нет никаких монет, кроме копеек. Сумма всевозможных способов уплаты некоторого количества копеек (и только копеек) записывается в виде

$$\begin{aligned} \mathrm{P} &= \cancel{1} + \textcircled{1} + \textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{1}\textcircled{1}\textcircled{1} + \cdots \\ &= \cancel{1} + \textcircled{1} + \textcircled{1}^2 + \textcircled{1}^3 + \textcircled{1}^4 + \cdots . \end{aligned}$$

Первый член означает не платить ничего, второй — заплатить одну копейку, третий — две копейки и т.д. Если, кроме копейки, допускается плата еще и пятаками, то сумма всех возможных способов уплаты принимает вид

$$\begin{aligned} \mathrm{N} &= \mathrm{P} + \textcircled{5}\,\mathrm{P} + \textcircled{5}\textcircled{5}\,\mathrm{P} + \textcircled{5}\textcircled{5}\textcircled{5}\,\mathrm{P} + \textcircled{5}\textcircled{5}\textcircled{5}\textcircled{5}\,\mathrm{P} + \cdots \\ &= (\cancel{5} + \textcircled{5} + \textcircled{5}^2 + \textcircled{5}^3 + \textcircled{5}^4 + \cdots)\,\mathrm{P}, \end{aligned}$$

поскольку каждый вариант выплаты включает некоторое количество пятаков, выбираемых из первого множителя, и некоторого количества копеек, выбираемого из P. (Обратите внимание, что N *не равно* сумме $\cancel{1} + \textcircled{1} + \textcircled{5} + (\textcircled{1} + \textcircled{5})^2 + (\textcircled{1} + \textcircled{5})^3 + \cdots$, потому что такая сумма включает многие варианты выплат более чем по одному разу. Например, член $(\textcircled{1} + \textcircled{5})^2 = \textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{5} + \textcircled{5}\textcircled{1} + \textcircled{5}\textcircled{5}$ рассматривает $\textcircled{1}\textcircled{5}$ и $\textcircled{5}\textcircled{1}$ как разные варианты, но мы хотим перечислить все множества монет по одному разу безотносительно к их порядку.)

Аналогично, если допустить к оплате еще и гривенники, то получится бесконечная сумма

$$\mathrm{D} = (\cancel{10} + \textcircled{10} + \textcircled{10}^2 + \textcircled{10}^3 + \textcircled{10}^4 + \cdots)\,\mathrm{N},$$

которая при полном раскрытии будет содержать слагаемые вида $\textcircled{10}^3\textcircled{5}^3\textcircled{1}^5 = \textcircled{10}\textcircled{10}\textcircled{10}\textcircled{5}\textcircled{5}\textcircled{5}\textcircled{1}\textcircled{1}\textcircled{1}\textcircled{1}\textcircled{1}$. Каждый из этих членов представляет собой какой-то из способов размена. Добавление к допустимым монетам четвертаков и полтинников дает

А многие ли помнят, что трехкопеечная монета называлась “алтын”?
— Переводчик

$$
\begin{aligned}
Q &= (\cancel{25} + \textcircled{25} + \textcircled{25}^2 + \textcircled{25}^3 + \textcircled{25}^4 + \cdots)\,D\,;\\
C &= (\cancel{50} + \textcircled{50} + \textcircled{50}^2 + \textcircled{50}^3 + \textcircled{50}^4 + \cdots)\,Q\,.
\end{aligned}
$$

Наша задача заключается в поиске количества членов C, которые сто́ят ровно 50 копеек.

Эта задача красиво решается при помощи следующего трюка. Заменим $\textcircled{1}$ на z, $\textcircled{5}$ на z^5, $\textcircled{10}$ на z^{10}, $\textcircled{25}$ на z^{25} и $\textcircled{50}$ на z^{50}. Тогда каждый член заменяется на z^n, где n — стоимость каждого исходного члена в копейках. Например, член $\textcircled{50}\textcircled{10}\textcircled{5}\textcircled{5}\textcircled{1}$ превращается в $z^{50+10+5+5+1} = z^{71}$. Четыре способа выплаты 13 копеек, а именно $\textcircled{10}\textcircled{1}^3$, $\textcircled{5}\textcircled{1}^8$, $\textcircled{5}^2\textcircled{1}^3$ и $\textcircled{1}^{13}$, сводятся к z^{13}; следовательно, коэффициентом при z^{13} после z-подстановки будет 4.

Пусть P_n, N_n, D_n, Q_n и C_n обозначают количество способов заплатить n копеек при использовании монет не старше соответственно 1, 5, 10, 25 и 50 копеек. Наш анализ говорит нам, что эти числа представляют собой коэффициенты при z^n в соответствующих степенных рядах

$$
\begin{aligned}
P &= 1 + z + z^2 + z^3 + z^4 + \cdots\,,\\
N &= (1 + z^5 + z^{10} + z^{15} + z^{20} + \cdots)\,P\,,\\
D &= (1 + z^{10} + z^{20} + z^{30} + z^{40} + \cdots)\,N\,,\\
Q &= (1 + z^{25} + z^{50} + z^{75} + z^{100} + \cdots)\,D\,,\\
C &= (1 + z^{50} + z^{100} + z^{150} + z^{200} + \cdots)\,Q\,.
\end{aligned}
$$

Интересно, сколько всего существует копеек в реальности? Если n больше, скажем, 10^{10}, то готов держать пари, что “в реальном мире” $P_n = 0$.

Очевидно, что $P_n = 1$ для всех $n \geqslant 0$. Небольшое размышление позволяет убедиться, что $N_n = \lfloor n/5 \rfloor + 1$: чтобы заплатить n копеек копейками и пятаками, мы должны взять 0 или 1, или …, или $\lfloor n/5 \rfloor$ пятаков, после чего останется единственный способ выбрать необходимое количество копеек. Таким образом, P_n и N_n очень просты, чего не скажешь о значениях D_n, Q_n и C_n.

Один из подходов к этим формулам состоит в том, чтобы заметить, что $1 + z^m + z^{2m} + \cdots$ есть просто $1/(1 - z^m)$. Следовательно, можно записать

$$
\begin{aligned}
P &= 1/(1-z)\,,\\
N &= P/(1-z^5)\,,\\
D &= N/(1-z^{10})\,,\\
Q &= D/(1-z^{25})\,,\\
C &= Q/(1-z^{50})\,.
\end{aligned}
$$

Умножая на знаменатель, получаем

$$\begin{aligned}(1-z)\,P &= 1\,,\\ (1-z^5)\,N &= P\,,\\ (1-z^{10})\,D &= N\,,\\ (1-z^{25})\,Q &= D\,,\\ (1-z^{50})\,C &= Q\,.\end{aligned}$$

Теперь, приравнивая коэффициенты при z^n в этих уравнениях, получаем рекуррентные соотношения, из которых легко вычисляются интересующие нас коэффициенты:

$$\begin{aligned}P_n &= P_{n-1}+[n=0]\,,\\ N_n &= N_{n-5}+P_n\,,\\ D_n &= D_{n-10}+N_n\,,\\ Q_n &= Q_{n-25}+D_n\,,\\ C_n &= C_{n-50}+Q_n\,.\end{aligned}$$

Например, коэффициент при z^n в $D=(1-z^{25})Q$ равен $Q_n - Q_{n-25}$; так что мы должны получить $Q_n - Q_{n-25} = D_n$, как и ожидалось.

Можно раскрыть эти рекуррентные соотношения и найти, например, что $Q_n = D_n + D_{n-25} + D_{n-50} + D_{n-75} + \cdots$, где сумма обрывается при отрицательных индексах. Однако неитеративная форма удобнее, потому что каждый коэффициент вычисляется при помощи единственного сложения, как в треугольнике Паскаля.

Давайте воспользуемся полученными рекуррентного соотношениями для поиска C_{50}. Начнем с того, что $C_{50} = C_0 + Q_{50}$, так что нам надо найти Q_{50}. Далее, $Q_{50} = Q_{25} + D_{50}$ и $Q_{25} = Q_0 + D_{25}$; так что нам требуется также знать D_{50} и D_{25}. Эти D_n, в свою очередь, зависят от D_{40}, D_{30}, D_{20}, D_{15}, D_{10} и D_5 и от N_{50}, N_{45}, …, N_5. Таким образом, для определения всех требуемых коэффициентов достаточно выполнить простые вычисления:

n	0	5	10	15	20	25	30	35	40	45	50
P_n	1	1	1	1	1	1	1	1	1	1	1
N_n	1	2	3	4	5	6	7	8	9	10	11
D_n	1	2	4	6	9	12	16		25		36
Q_n	1					13					49
C_n	1										50

(Не считая оплаты кредитной карточкой.)

Последнее значение таблицы и есть искомый ответ C_{50}: имеется ровно 50 способов заплатить 50 копеек. (Ответ верен для набора монет США; нетрудно подсчитать, что российский набор монет достоинством 1, 5, 10 и 50 копеек позволяет выплатить ту же сумму 37 способами, украинский (1, 2, 5, 10, 25 и 50 копеек) — 407 способами, а советский (1, 2, 3, 5, 10, 15, 20 и 50 копеек) — 2956 способами. — *Примеч. пер.*)

А что можно сказать об аналитическом виде C_n? Перемножение всех уравнений дает компактное выражение

$$C = \frac{1}{1-z}\,\frac{1}{1-z^5}\,\frac{1}{1-z^{10}}\,\frac{1}{1-z^{25}}\,\frac{1}{1-z^{50}}, \qquad (7.11)$$

но как извлечь из него коэффициенты при z^n, сразу не понятно. К счастью, это возможно, и позже в этой главе мы еще вернемся к данной задаче.

Более элегантные формулы получаются, если размен выполняется в стране, где чеканятся монеты любого положительного достоинства (①, ②, ③, ...), а не только пять рассмотренных ранее. Соответствующая производящая функция представляет собой бесконечное произведение дробей

$$\frac{1}{(1-z)(1-z^2)(1-z^3)\ldots},$$

и коэффициент при z^n в этом выражении после перемножения и раскрытия скобок называется $p(n)$, количество *разбиений* n. Разбиение n — это представление n в виде суммы положительных целых чисел без учета их порядка. Например, имеется семь различных разбиений 5, а именно

$$5 = 4{+}1 = 3{+}2 = 3{+}1{+}1 = 2{+}2{+}1 = 2{+}1{+}1{+}1 = 1{+}1{+}1{+}1{+}1\,;$$

следовательно, $p(5) = 7$. (Кроме того, $p(2) = 2$, $p(3) = 3$, $p(4) = 5$ и $p(6) = 11$; начало выглядит так, как будто $p(n)$ — всегда простые числа. Увы, $p(7) = 15$ не оставляет от этой гипотезы камня на камне.) Выражение в аналитическом виде для $p(n)$ неизвестно (хотя имеются приближенные формулы, например

$$p(n) = \frac{e^{\pi\sqrt{2n'/3}}}{4n'\sqrt{3}}\left(1 - \frac{1}{\pi}\sqrt{\frac{3}{2n'}}\right)\left(1 + O\left(e^{-\pi\sqrt{n/6}}\right)\right),$$
$$n' = n - \frac{1}{24}.$$

— *Примеч. пер.*), но теория разбиений представляет собой захватывающую область математики, в которой было сделано немало замечательных открытий. Например, Рамануджан (Ramanujan) доказал, что $p(5n+4) \equiv 0 \pmod 5$, $p(7n+5) \equiv 0 \pmod 7$ и $p(11n+6) \equiv 0 \pmod{11}$, выполнив остроумные преобразования производящих функций (см. Эндрюс (Andrews) [11, глава 10]).

7.2 Основные манипуляции

Рассмотрим теперь более внимательно некоторые методы, благодаря которым о степенных рядах можно говорить в превосходной степени.

Сначала — несколько слов о терминологии и обозначениях. Обобщенная производящая функция имеет вид

$$G(z) \;=\; g_0 + g_1 z + g_2 z^2 + \cdots \;=\; \sum_{n \geqslant 0} g_n z^n \,, \tag{7.12}$$

и мы говорим, что $G(z)$ (или, для краткости, просто G) является производящей функцией для последовательности $\langle g_0, g_1, g_2, \ldots \rangle$, которую мы также записываем как $\langle g_n \rangle$. Коэффициент g_n при z^n в $G(z)$ часто обозначается $[z^n]\,G(z)$, как в разделе 5.4.

Сумма в (7.12) берется по всем $n \geqslant 0$, но часто более удобно распространить ее на все целые n. Это можно сделать, просто положив $g_{-1} = g_{-2} = \cdots = 0$. В таких случаях можно по-прежнему говорить о последовательности $\langle g_0, g_1, g_2, \ldots \rangle$, как если бы g_n не существовали для отрицательных n.

При работе с производящими функциями можно говорить о двух типах выражений в "аналитическом виде". Возможно выражение $G(z)$ в аналитическом виде как функции от z, а возможно выражение в аналитическом виде g_n как функции от n. Например, производящая функция для чисел Фибоначчи имеет аналитический вид $z/(1-z-z^2)$; сами числа Фибоначчи в аналитическом виде записываются как $(\phi^n - \hat{\phi}^n)/\sqrt{5}$. Какой именно аналитический вид имеется в виду, будет понятно из контекста.

Теперь несколько слов о перспективах. Производящая функция $G(z)$ может являться в двух ипостасях в зависимости от точки зрения. Иногда это функция комплексной переменной z, обладающая всеми стандартными свойствами, доказываемыми в учебниках по анализу. В других случаях производящая функция представляет собой формальный степенной ряд, в котором z — не более чем формальная переменная. В предыдущем разделе, например, мы воспользовались второй интерпретацией; мы встречались с несколькими примерами, в которых z подставлялось вместо некоторого свойства комбинаторного объекта в "сумму" таких объектов. Коэффициент при z^n после этого от-

Если физики вовсю используют корпускулярно-волновой дуализм, то и математикам не пристало от них отставать.

вечал количеству комбинаторных объектов, имеющих это свойство в количестве n.

При рассмотрении $G(z)$ в качестве функции комплексного переменного встает вопрос о сходимости ряда. В главе 2 мы говорили, что бесконечная сумма $\sum_{n\geqslant 0} g_n z^n$ сходится (абсолютно) тогда и только тогда, когда существует ограничивающая константа A, такая, что конечные суммы $\sum_{0\leqslant n\leqslant N} |g_n z^n|$ никогда не превосходят A, ни при каком N. Таким образом, легко увидеть, что если сумма $\sum_{n\geqslant 0} g_n z^n$ сходится для некоторого значения $z = z_0$, то она сходится и для всех z, для которых $|z| < |z_0|$. Кроме того, должно выполняться $\lim_{n\to\infty} |g_n z_0^n| = 0$; следовательно, в обозначениях главы 9, $g_n = O(|1/z_0|^n)$, если имеется сходимость в точке z_0. И обратно, если $g_n = O(M^n)$, то ряд $\sum_{n\geqslant 0} g_n z^n$ сходится при всех $|z| < 1/M$. Таковы основные факты о сходимости степенных рядов.

Однако для наших целей сходимость — скорее отвлекающий фактор, не играющий особой роли, если только мы не изучаем асимптотическое поведение коэффициентов. Почти каждая операция над производящими функциями может быть строго обоснована как операция над формальными степенными рядами, и такие операции являются корректными, даже если ряды расходятся. (Соответствующую теорию можно найти, например, у Белла (Bell) [23], Найвена (Niven) [282] и Хенричи (Henrici) [182, глава 1].)

Даже если превысим скорость, забыв пристегнуться ремнем безопасности из-за выпитого...

Кроме того, даже если мы отбросим все предосторожности и выведем формулы без строгого обоснования, то в общем случае окажется возможным доказать результаты вывода по индукции. Например, производящая функция для чисел Фибоначчи сходится только при $|z| < 1/\phi \approx 0.618$, но нам не требуется это знать при доказательстве формулы $F_n = (\phi^n - \hat{\phi}^n)/\sqrt{5}$. Будучи найденной, эта формула может быть доказана непосредственно, если не доверять теории формальных степенных рядов. Таким образом, в этой главе мы будем игнорировать вопросы сходимости; сейчас они для нас послужат не помощью, а помехой.

До сих пор речь шла о перспективах. Теперь же рассмотрим инструментарий, т.е. набор приемов преобразования производящих функций — сложения, сдвига, замены переменных, дифференцирования, интегрирования и умножения. В последующем изложении мы примем, если не оговорено противное, что $F(z)$ и $G(z)$ — производящие функции последовательностей $\langle f_n\rangle$ и $\langle g_n\rangle$. Будем также считать, что f_n и g_n нулевые для отрицательных n, поскольку такое предположение позволит в некоторых случаях не заботиться о пределах суммирования.

Совершенно очевидно, что получится, если умножить каждую из производящих функций F и G на константу и сложить получившиеся произведения:

$$\alpha F(z) + \beta G(z) = \alpha \sum_n f_n z^n + \beta \sum_n g_n z^n = \\ = \sum_n (\alpha f_n + \beta g_n) z^n . \qquad (7.13)$$

Мы получаем производящую функцию для последовательности $\langle \alpha f_n + \beta g_n \rangle$.

Сдвиг производящей функции ненамного сложнее. Чтобы сдвинуть $G(z)$ вправо на m позиций, т.е. для образования производящей функции для последовательности $\langle 0, \ldots, 0, g_0, g_1, \ldots \rangle = \langle g_{n-m} \rangle$, начинающейся с m нулей, ее просто надо умножить на z^m:

$$z^m G(z) = \sum_n g_n z^{n+m} = \sum_n g_{n-m} z^n , \quad \text{целое } m \geqslant 0. \qquad (7.14)$$

Это та операция, которую мы дважды использовали вместе со сложением для вывода уравнения $(1 - z - z^2)F(z) = z$ при поиске чисел Фибоначчи в аналитическом виде в главе 6.

Чтобы сдвинуть $G(z)$ на m позиций влево, т.е. для получения производящей функции для последовательности с отброшенными первыми m элементами $\langle g_m, g_{m+1}, g_{m+2}, \ldots \rangle = \langle g_{n+m} \rangle$, вычтем из $G(z)$ первые m членов и разделим результат на z^m:

$$\frac{G(z) - g_0 - g_1 z - \cdots - g_{m-1} z^{m-1}}{z^m} = \sum_{n \geqslant m} g_n z^{n-m} = \\ = \sum_{n \geqslant 0} g_{n+m} z^n . \qquad (7.15)$$

(Последнюю сумму нельзя распространять на все n, если только не оказывается, что $g_0 = \cdots = g_{m-1} = 0$.)

Умножение z на константу представляет собой еще один из наших трюков:

$$G(cz) = \sum_n g_n (cz)^n = \sum_n c^n g_n z^n ; \qquad (7.16)$$

это дает нам производящую функцию для последовательности $\langle c^n g_n \rangle$. Особенно полезен частный случай $c = -1$.

Не нравится мне это $\frac{d}{dz}$: так и кажется, что оно не генерирует, а d-генерирует последовательность...

Зачастую к коэффициентам следует добавить множитель n. Сделать это позволяет дифференцирование:

$$G'(z) = g_1 + 2g_2 z + 3g_3 z^2 + \cdots = \sum_n (n+1) g_{n+1} z^n . \qquad (7.17)$$

Выполняя сдвиг на одну позицию вправо, получаем формулу, которая иногда оказывается даже более полезной:

$$zG'(z) = \sum_n n g_n z^n . \qquad (7.18)$$

Это — производящая функция последовательности $\langle n g_n \rangle$. Повторное дифференцирование позволяет умножить g_n на любой требуемый полином от n.

Обратная операция, интегрирование, позволяет разделить элементы последовательности на n:

$$\int_0^z G(t)\,dt = g_0 z + \frac{1}{2} g_1 z^2 + \frac{1}{3} g_2 z^3 + \cdots = \sum_{n \geqslant 1} \frac{1}{n} g_{n-1} z^n . \qquad (7.19)$$

(Обратите внимание, что постоянный член равен нулю.) Если вместо $\langle g_{n-1}/n \rangle$ нам нужна производящая функция для последовательности $\langle g_n/n \rangle$, то сначала надо выполнить сдвиг на одну позицию влево, заменив под знаком интеграла $G(t)$ на $\bigl(G(t) - g_0\bigr)/t$.

Наконец, вот как выполняется умножение производящих функций:

$$\begin{aligned} F(z)G(z) &= \\ &= (f_0 + f_1 z + f_2 z^2 + \cdots)(g_0 + g_1 z + g_2 z^2 + \cdots) = \\ &= (f_0 g_0) + (f_0 g_1 + f_1 g_0) z + (f_0 g_2 + f_1 g_1 + f_2 g_0) z^2 + \cdots = \\ &= \sum_n \Bigl(\sum_k f_k g_{n-k} \Bigr) z^n . \end{aligned} \qquad (7.20)$$

Как мы видели в главе 5, полученная сумма представляет собой производящую функцию последовательности $\langle h_n \rangle$, *свертки* $\langle f_n \rangle$ и $\langle g_n \rangle$. Сумму $h_n = \sum_k f_k g_{n-k}$ можно записать и как $h_n = \sum_{k=0}^n f_k g_{n-k}$, потому что $f_k = 0$ при $k < 0$, а $g_{n-k} = 0$ при $k > n$. Умножение/свертка немного сложнее других операций, но при этом очень полезна — настолько, что мы посвятим примерам ее использования целый раздел 7.5.

Имеется несколько частных случаев умножения, которые заслуживают рассмотрения в качестве отдельных операций. Один из таких случаев мы уже видели: если $F(z) = z^m$, то мы получаем операцию сдвига (7.14). В этом случае сумма h_n превращается в единственный член g_{n-m}, потому что все коэффициенты f_k равны 0, за исключением $f_m = 1$.

Еще один полезный частный случай имеет место, когда $F(z)$ — хорошо известная нам функция $1/(1-z) = 1 + z + z^2 + \cdots$;

Таблица 406. Преобразования производящих функций

$$\alpha F(z) + \beta G(z) = \sum_n (\alpha f_n + \beta g_n) z^n$$

$$z^m G(z) = \sum_n g_{n-m} z^n, \quad \text{целое } m \geqslant 0$$

$$\frac{G(z) - g_0 - g_1 z - \cdots - g_{m-1} z^{m-1}}{z^m} = \sum_{n \geqslant 0} g_{n+m} z^n, \quad \text{целое } m \geqslant 0$$

$$G(cz) = \sum_n c^n g_n z^n$$

$$G'(z) = \sum_n (n+1) g_{n+1} z^n$$

$$zG'(z) = \sum_n n g_n z^n$$

$$\int_0^z G(t)\,dt = \sum_{n \geqslant 1} \frac{1}{n} g_{n-1} z^n$$

$$F(z)G(z) = \sum_n \Bigl(\sum_k f_k g_{n-k}\Bigr) z^n$$

$$\frac{1}{1-z} G(z) = \sum_n \Bigl(\sum_{k \leqslant n} g_k\Bigr) z^n$$

тогда все f_k (при $k \geqslant 0$) равны 1, и мы получаем важную формулу

$$\frac{1}{1-z} G(z) = \sum_n \Bigl(\sum_{k \geqslant 0} g_{n-k}\Bigr) z^n = \sum_n \Bigl(\sum_{k \leqslant n} g_k\Bigr) z^n. \qquad (7.21)$$

Умножение производящей функции на $1/(1-z)$ дает нам производящую функцию для последовательности частичных сумм исходной последовательности.

В табл. 406 собраны все рассмотренные к настоящему моменту операции. Для эффективного их использования полезно иметь в запасе достаточно большой набор производящих функций. В табл. 407 перечислены некоторые простейшие из них; используя приведенную в таблицах информацию, мы уже в состоянии решать некоторые задачи.

Таблица 407. Простые последовательности и их производящие функции

Последовательность	Производящая функция	Аналитический вид
$\langle 1,0,0,0,0,0,\ldots\rangle$	$\sum_{n\geqslant 0}[n=0]\,z^n$	1
$\langle 0,\ldots,0,1,0,0,\ldots\rangle$	$\sum_{n\geqslant 0}[n=m]\,z^n$	z^m
$\langle 1,1,1,1,1,1,\ldots\rangle$	$\sum_{n\geqslant 0} z^n$	$\dfrac{1}{1-z}$
$\langle 1,-1,1,-1,1,-1,\ldots\rangle$	$\sum_{n\geqslant 0}(-1)^n z^n$	$\dfrac{1}{1+z}$
$\langle 1,0,1,0,1,0,\ldots\rangle$	$\sum_{n\geqslant 0}[2\backslash n]\,z^n$	$\dfrac{1}{1-z^2}$
$\langle 1,0,\ldots,0,1,0,\ldots,0,1,0,\ldots\rangle$	$\sum_{n\geqslant 0}[m\backslash n]\,z^n$	$\dfrac{1}{1-z^m}$
$\langle 1,2,3,4,5,6,\ldots\rangle$	$\sum_{n\geqslant 0}(n+1)\,z^n$	$\dfrac{1}{(1-z)^2}$
$\langle 1,2,4,8,16,32,\ldots\rangle$	$\sum_{n\geqslant 0} 2^n z^n$	$\dfrac{1}{1-2z}$
$\langle 1,4,6,4,1,0,0,\ldots\rangle$	$\sum_{n\geqslant 0}\dbinom{4}{n} z^n$	$(1+z)^4$
$\langle 1,c,\binom{c}{2},\binom{c}{3},\ldots\rangle$	$\sum_{n\geqslant 0}\dbinom{c}{n} z^n$	$(1+z)^c$
$\left\langle 1,c,\binom{c+1}{2},\binom{c+2}{3},\ldots\right\rangle$	$\sum_{n\geqslant 0}\dbinom{c+n-1}{n} z^n$	$\dfrac{1}{(1-z)^c}$
$\langle 1,c,c^2,c^3,\ldots\rangle$	$\sum_{n\geqslant 0} c^n z^n$	$\dfrac{1}{1-cz}$
$\left\langle 1,\binom{m+1}{m},\binom{m+2}{m},\binom{m+3}{m},\ldots\right\rangle$	$\sum_{n\geqslant 0}\dbinom{m+n}{m} z^n$	$\dfrac{1}{(1-z)^{m+1}}$
$\langle 0,1,\frac12,\frac13,\frac14,\ldots\rangle$	$\sum_{n\geqslant 1}\dfrac{1}{n} z^n$	$\ln\dfrac{1}{1-z}$
$\langle 0,1,-\frac12,\frac13,-\frac14,\ldots\rangle$	$\sum_{n\geqslant 1}\dfrac{(-1)^{n+1}}{n} z^n$	$\ln(1+z)$
$\langle 1,1,\frac12,\frac16,\frac1{24},\frac1{120},\ldots\rangle$	$\sum_{n\geqslant 0}\dfrac{1}{n!} z^n$	e^z

Каждая производящая функция из табл. 407 сама по себе достаточно важна, чтобы потратить время на ее запоминание. Многие из них представляют собой частные случаи других, а другие могут быть легко получены при помощи основных операций из табл. 406; таким образом, перенапрягать память не придется.

Подсказка: если последовательность состоит из биномиальных коэффициентов, то производящая функция обычно включает биномы $1 \pm z$.

Например, рассмотрим последовательность $\langle 1, 2, 3, 4, \dots \rangle$, производящая функция которой $1/(1-z)^2$ часто оказывается полезной. Эта производящая функция находится примерно в средине табл. 407 и представляет собой частный случай $m = 1$ последовательности $\langle 1, \binom{m+1}{m}, \binom{m+2}{m}, \binom{m+3}{m}, \dots \rangle$, которая располагается ниже в таблице; это также частный случай $c = 2$ родственной последовательности $\langle 1, c, \binom{c+1}{2}, \binom{c+2}{3}, \dots \rangle$. Вывести нужную формулу можно из производящей функции последовательности $\langle 1, 1, 1, 1, \dots \rangle$, взяв ее частичные суммы, как в (7.21), т.е. поделив $1/(1-z)$ на $(1-z)$. Можно также вывести ее из $\langle 1, 1, 1, 1, \dots \rangle$ путем дифференцирования, воспользовавшись (7.17).

Ну все, все — уговорили!

Последовательность $\langle 1, 0, 1, 0, \dots \rangle$ представляет собой еще один пример последовательности, производящая функция которой может быть получена разными способами. Очевидно, что можно просто вывести формулу $\sum_n z^{2n} = 1/(1-z^2)$ путем подстановки z^2 вместо z в тождество $\sum_n z^n = 1/(1-z)$; можно также применить накопительное суммирование к последовательности $\langle 1, -1, 1, -1, \dots \rangle$, производящая функция которой равна $1/(1+z)$, что дает $1/(1+z)(1-z) = 1/(1-z^2)$. Имеется и третий путь, основанный на общем методе извлечения членов с четными номерами $\langle g_0, 0, g_2, 0, g_4, 0, \dots \rangle$ из *любой* последовательности: если сложить $G(-z)$ с $G(+z)$, можно получить

$$G(z) + G(-z) = \sum_n g_n\bigl(1 + (-1)^n\bigr)z^n = 2\sum_n g_n[n \text{ четное}]z^n;$$

следовательно,

$$\frac{G(z) + G(-z)}{2} = \sum_n g_{2n}\, z^{2n}. \tag{7.22}$$

Аналогично можно извлечь члены с нечетными номерами:

$$\frac{G(z) - G(-z)}{2} = \sum_n g_{2n+1}\, z^{2n+1}. \tag{7.23}$$

В частном случае $g_n = 1$ и $G(z) = 1/(1-z)$ получаем, что производящая функция для $\langle 1, 0, 1, 0, \dots \rangle$ есть $\frac{1}{2}\bigl(G(z) + G(-z)\bigr) = \frac{1}{2}\bigl(\frac{1}{1-z} + \frac{1}{1+z}\bigr) = \frac{1}{1-z^2}$.

Попробуем применить этот трюк к производящей функции для чисел Фибоначчи. Мы знаем, что $\sum_n F_n z^n = z/(1-z-z^2)$;

следовательно,

$$\sum_n F_{2n} z^{2n} = \frac{1}{2}\left(\frac{z}{1-z-z^2} + \frac{-z}{1+z-z^2}\right) =$$
$$= \frac{1}{2}\left(\frac{z+z^2-z^3-z+z^2+z^3}{(1-z^2)^2-z^2}\right) = \frac{z^2}{1-3z^2+z^4}.$$

Эта функция генерирует последовательность $\langle F_0, 0, F_2, 0, F_4, \ldots\rangle$; значит, последовательность чисел Фибоначчи с четными индексами $\langle F_0, F_2, F_4, F_6, \ldots\rangle = \langle 0, 1, 3, 8, \ldots\rangle$ имеет простую производящую функцию:

$$\sum_n F_{2n} z^n = \frac{z}{1-3z+z^2}. \tag{7.24}$$

7.3 Решение рекуррентных соотношений

Займемся теперь одним из наиболее важных применений производящих функций — решением рекуррентных соотношений.

Нас интересует поиск выражения в аналитическом виде как функции от n для элементов g_n последовательности $\langle g_n\rangle$, которая удовлетворяет некоторому рекуррентному соотношению. Решение этой задачи при помощи производящих функций выполняется за четыре шага, настолько механических, что их почти можно запрограммировать на компьютере.

1 Запишите одно уравнение, выражающее g_n через другие элементы последовательности. Это уравнение должно быть корректным для всех целых n с учетом того, что $g_{-1} = g_{-2} = \cdots = 0$.

2 Умножьте обе части уравнения на z^n и просуммируйте по всем n. В левой части получится сумма $\sum_n g_n z^n$, которая представляет собой производящую функцию $G(z)$. Правую часть следует преобразовать таким образом, чтобы она превратилась в некоторое другое выражение, включающее $G(z)$.

3 Решите получившееся уравнение, получив для $G(z)$ выражение в аналитическом виде.

4 Разложите $G(z)$ в степенной ряд и возьмите коэффициент при z^n; это и будет аналитический вид g_n.

Этот метод работает, потому что единая функция $G(z)$ представляет всю последовательность $\langle g_n\rangle$ способом, допускающим выполнение большого количества преобразований.

Пример 1: еще раз о числах Фибоначчи

Попробуем в качестве примера повторить вывод формулы для чисел Фибоначчи из главы 6. Там мы нащупывали путь, изучая новый метод; теперь же будем действовать более систематически. Нам дано рекуррентное соотношение

$$\begin{aligned} g_0 &= 0; \qquad g_1 = 1; \\ g_n &= g_{n-1} + g_{n-2} \qquad \text{при } n \geqslant 2. \end{aligned}$$

Найдем выражение для g_n, выполнив приведенные выше четыре шага.

Шаг 1 гласит, что мы должны записать рекуррентное соотношение в виде “одного уравнения” для g_n. Мы можем написать

$$g_n = \begin{cases} 0, & \text{если } n \leqslant 0; \\ 1, & \text{если } n = 1; \\ g_{n-1} + g_{n-2}, & \text{если } n > 1; \end{cases}$$

но это самообман. В действительности на шаге 1 требуется формула, не содержащая конструкций с перечислением случаев. Одно уравнение

$$g_n = g_{n-1} + g_{n-2}$$

имеет место для $n \geqslant 2$ и выполняется для $n \leqslant 0$ (потому что $g_0 = 0$ и $g_{\text{отриц}} = 0$). Но при $n = 1$ мы получим 1 в левой части и 0 в правой. К счастью, ситуацию легко исправить, поскольку можно добавить в правую часть член $[n=1]$; тем самым будет добавлена 1 при $n = 1$, а при $n \neq 1$ уравнение останется неизменным. Итак, имеем

$$g_n = g_{n-1} + g_{n-2} + [n=1];$$

это и есть то уравнение, которое требуется получить на шаге 1.

На шаге 2 мы должны преобразовать уравнение для $\langle g_n \rangle$ в уравнение для $G(z) = \sum_n g_n z^n$. Это нетрудно:

$$\begin{aligned} G(z) = \sum_n g_n z^n &= \sum_n g_{n-1} z^n + \sum_n g_{n-2} z^n + \sum_n [n=1] z^n = \\ &= \sum_n g_n z^{n+1} + \sum_n g_n z^{n+2} + z = \\ &= zG(z) + z^2 G(z) + z. \end{aligned}$$

Шаг 3 в нашем случае тоже прост; имеем

$$G(z) = \frac{z}{1 - z - z^2},$$

что, конечно, неудивительно.

Вся загвоздка в шаге 4. В главе 6 выполнить этот шаг помогло внезапное озарение; здесь же мы будем двигаться медленнее, чтобы затем, в более сложных задачах, уверенно справляться с этим шагом. Итак, что же представляет собой

$$[z^n]\,\frac{z}{1-z-z^2}\,,$$

коэффициент при z^n в разложении $z/(1-z-z^2)$ в степенной ряд? Как в общем случае, имея некоторую рациональную функцию

$$R(z) \;=\; \frac{P(z)}{Q(z)}\,,$$

где P и Q — полиномы, найти коэффициент $[z^n]\,R(z)$?

Имеется один вид рациональных функций с особо хорошими коэффициентами, а именно

$$\frac{a}{(1-\rho z)^{m+1}} \;=\; \sum_{n\geqslant 0}\binom{m+n}{m}a\rho^n z^n\,. \tag{7.25}$$

(В табл. 407 содержится данная формула при $\rho = 1$; общую формулу, показанную здесь, можно получить подстановкой ρz вместо z и умножением на a.) Коэффициенты конечной суммы функций вида (7.25),

$$\begin{aligned} S(z) \;=\; &\frac{a_1}{(1-\rho_1 z)^{m_1+1}} + \frac{a_2}{(1-\rho_2 z)^{m_2+1}} + \\ &+\cdots+\frac{a_l}{(1-\rho_l z)^{m_l+1}}\,, \end{aligned} \tag{7.26}$$

также неплохи:

$$\begin{aligned} [z^n]\,S(z) \;=\; &a_1\binom{m_1+n}{m_1}\rho_1^n + a_2\binom{m_2+n}{m_2}\rho_2^n + \\ &+\cdots+a_l\binom{m_l+n}{m_l}\rho_l^n\,. \end{aligned} \tag{7.27}$$

Мы покажем, что любая рациональная функция $R(z)$, такая, что $R(0) \neq \infty$, может быть выражена в виде

$$R(z) \;=\; S(z) + T(z)\,, \tag{7.28}$$

где $S(z)$ имеет вид (7.26), а $T(z)$ представляет собой полином. Таким образом, для коэффициентов $[z^n]\,R(z)$ имеется выражение в аналитическом виде. Поиск $S(z)$ и $T(z)$ эквивалентен поиску “разложения на элементарные дроби” функции $R(z)$.

Обратите внимание, что $S(z) = \infty$ для значений z, равных $1/\rho_1, \ldots, 1/\rho_l$. Следовательно, числа ρ_k, которые нам надо найти, если мы хотим представить $R(z)$ в виде $S(z) + T(z)$, должны быть обратны числам α_k, для которых $Q(\alpha_k) = 0$. (Вспомните, что $R(z) = P(z)/Q(z)$, где P и Q — полиномы; $R(z) = \infty$, только если $Q(z) = 0$.)

Предположим, что $Q(z)$ имеет вид

$$Q(z) \;=\; q_0 + q_1 z + \cdots + q_m z^m\,, \qquad \text{где } q_0 \neq 0 \text{ и } q_m \neq 0.$$

"Обращенный" полином

$$Q^R(z) \;=\; q_0 z^m + q_1 z^{m-1} + \cdots + q_m$$

имеет важную взаимосвязь с $Q(z)$:

$$\begin{aligned} Q^R(z) \;&=\; q_0(z-\rho_1)\ldots(z-\rho_m) \quad\Longleftrightarrow \\ &\Longleftrightarrow\quad Q(z) \;=\; q_0(1-\rho_1 z)\ldots(1-\rho_m z)\,. \end{aligned}$$

Таким образом, корни Q^R обратны корням Q и наоборот. Следовательно, для нахождения нужных нам чисел ρ_k надо разложить обращенный полином $Q^R(z)$ на множители.

Например, в случае последовательности Фибоначчи имеем

$$Q(z) \;=\; 1 - z - z^2\,; \qquad Q^R(z) \;=\; z^2 - z - 1\,.$$

Корни Q^R можно найти при помощи общей формулы для корней квадратного уравнения $\bigl(-b \pm \sqrt{b^2 - 4ac}\,\bigr)/2a$, положив $(a, b, c) = (1, -1, -1)$; они оказываются равными

$$\phi \;=\; \frac{1+\sqrt{5}}{2} \quad \text{и} \quad \hat\phi \;=\; \frac{1-\sqrt{5}}{2}\,.$$

Следовательно, $Q^R(z) = (z-\phi)(z-\hat\phi)$ и $Q(z) = (1-\phi z)(1-\hat\phi z)$.

После того как найдены числа ρ, можно перейти к построению разложения на элементарные дроби. Сделать это проще всего в случае различных корней, поэтому сначала рассмотрим именно этот частный случай. Мы можем сформулировать и доказать общий результат.

Теорема о разложении рациональных функций в случае различных корней

Если $R(z) = P(z)/Q(z)$, *где* $Q(z) = q_0(1-\rho_1 z)\ldots(1-\rho_l z)$ *и числа* $(\rho_1, \ldots, \rho_l)$ *различны, и если* $P(z)$ *является полиномом степени меньше* l, *то*

$$\begin{aligned} &[z^n]\, R(z) \;=\; a_1 \rho_1^n + \cdots + a_l \rho_l^n\,, \\ &\text{где } a_k = \frac{-\rho_k P(1/\rho_k)}{Q'(1/\rho_k)}. \end{aligned} \tag{7.29}$$

Доказательство. Пусть $a_1, \ldots, a_l$ — константы с указанными выше значениями. Формула (7.29) справедлива, когда $R(z) = P(z)/Q(z)$ равно

$$S(z) = \frac{a_1}{1-\rho_1 z} + \cdots + \frac{a_l}{1-\rho_l z}.$$

Можно доказать, что $R(z) = S(z)$, показав, что функция $T(z) = R(z) - S(z)$ не бесконечна при $z \to 1/\rho_k$. Отсюда будет следовать, что рациональная функция $T(z)$ никогда не обращается в бесконечность; следовательно, функция $T(z)$ должна быть полиномом. Можно также показать, что $T(z) \to 0$ при $z \to \infty$; следовательно, $T(z)$ должна быть нулем.

Сразите своих родителей, оставив книгу открытой на этой странице, когда уйдете гулять.

(Скорее всего, они будут сражены не формулами, а единственным, что им будет понятно на этой странице, — данным советом...
— Переводчик)

Пусть $\alpha_k = 1/\rho_k$. Чтобы доказать, что $\lim_{z\to\alpha_k} T(z) \neq \infty$, достаточно показать, что $\lim_{z\to\alpha_k} (z-\alpha_k)T(z) = 0$, потому что $T(z)$ — рациональная функция от z. Таким образом, мы хотим показать, что

$$\lim_{z\to\alpha_k} (z-\alpha_k)R(z) = \lim_{z\to\alpha_k} (z-\alpha_k)S(z).$$

Предел в правой части равен $\lim_{z\to\alpha_k} a_k(z-\alpha_k)/(1-\rho_k z) = -a_k/\rho_k$, потому что $(1-\rho_k z) = -\rho_k(z-\alpha_k)$ и $(z-\alpha_k)/(1-\rho_j z) \to 0$ для $j \neq k$. Предел в левой части равен

$$\lim_{z\to\alpha_k} (z-\alpha_k)\frac{P(z)}{Q(z)} = P(\alpha_k) \lim_{z\to\alpha_k} \frac{z-\alpha_k}{Q(z)} = \frac{P(\alpha_k)}{Q'(\alpha_k)}$$

по правилу Лопиталя. Таким образом, теорема доказана.

Возвратимся к примеру с числами Фибоначчи. Здесь $P(z) = z$ и $Q(z) = 1 - z - z^2 = (1-\phi z)(1-\hat\phi z)$; следовательно, $Q'(z) = -1 - 2z$ и

$$\frac{-\rho P(1/\rho)}{Q'(1/\rho)} = \frac{-1}{-1-2/\rho} = \frac{\rho}{\rho+2}.$$

Согласно (7.29) коэффициент при ϕ^n в $[z^n]\,R(z)$ равен $\phi/(\phi+2) = 1/\sqrt{5}$; коэффициент при $\hat\phi^n$ равен $\hat\phi/(\hat\phi+2) = -1/\sqrt{5}$. Таким образом, теорема гласит, что $F_n = (\phi^n - \hat\phi^n)/\sqrt{5}$, как в (6.123).

Если $Q(z)$ имеет кратные корни, вычисления становятся сложнее, но мы можем усовершенствовать доказательство и установить справедливость следующего, более общего результата.

Общая теорема о разложении рациональных функций

Если $R(z) = P(z)/Q(z)$, *где* $Q(z) = q_0(1-\rho_1 z)^{d_1} \ldots (1-\rho_l z)^{d_l}$ *и числа* $(\rho_1, \ldots, \rho_l)$ *различны, и если* $P(z)$ *является полиномом степени меньше* $d_1 + \cdots + d_l$, *то*

$$[z^n]\,R(z) = f_1(n)\rho_1^n + \cdots + f_l(n)\rho_l^n \quad \text{для всех } n \geqslant 0, \qquad (7.30)$$

где все $f_k(n)$ *являются полиномами степени* $d_k - 1$ *со старшим коэффициентом*

$$\begin{aligned} a_k &= \frac{(-\rho_k)^{d_k} P(1/\rho_k) d_k}{Q^{(d_k)}(1/\rho_k)} = \\ &= \frac{P(1/\rho_k)}{(d_k - 1)!\, q_0 \prod_{j \ne k} (1 - \rho_j/\rho_k)^{d_j}} . \end{aligned} \qquad (7.31)$$

Это можно доказать по индукции по $\max(d_1, \ldots, d_l)$, используя тот факт, что

$$R(z) - \frac{a_1 (d_1 - 1)!}{(1 - \rho_1 z)^{d_1}} - \cdots - \frac{a_l (d_l - 1)!}{(1 - \rho_l z)^{d_l}}$$

является рациональной функцией, в знаменателе которой находится полином, не делящийся на $(1 - \rho_k z)^{d_k}$ ни при каких k.

Пример 2: достаточно случайное рекуррентное соотношение

Теперь, познакомившись с более общими методами, мы готовы взяться за новые задачи. Попробуем найти аналитическое решение рекуррентного соотношения

$$\begin{aligned} g_0 &= g_1 = 1 ; \\ g_n &= g_{n-1} + 2g_{n-2} + (-1)^n \qquad \text{при } n \geqslant 2. \end{aligned} \qquad (7.32)$$

Всегда неплохо иметь таблицу нескольких малых случаев, и рекуррентное соотношение легко позволяет это сделать:

n	0	1	2	3	4	5	6	7
$(-1)^n$	1	-1	1	-1	1	-1	1	-1
g_n	1	1	4	5	14	23	52	97

Простая формула не просматривается, и этой последовательности нет даже в *Справочнике* [330] Слоана (Sloane); так что нам не избежать четырехходового процесса, если нам действительно надо решить это рекуррентное соотношение.

Шаг 1 прост, поскольку все, что нам надо, — это придумать дополнительные члены для коррекции ситуации при $n < 2$: уравнение

$$g_n = g_{n-1} + 2g_{n-2} + (-1)^n [n \geqslant 0] + [n = 1]$$

выполняется для всех целых n. Теперь можно выполнить шаг 2:

N.B.: верхний индекс в $\sum_{n=1} z^n$ не пропущен!

$$\begin{aligned} G(z) &= \sum_n g_n z^n = \\ &= \sum_n g_{n-1} z^n + 2\sum_n g_{n-2} z^n + \sum_{n\geqslant 0} (-1)^n z^n + \sum_{n=1} z^n = \\ &= zG(z) + 2z^2 G(z) + \frac{1}{1+z} + z\,. \end{aligned}$$

(Кстати, можно было использовать $\binom{-1}{n}$ вместо $(-1)^n[n\geqslant 0]$, тем самым по биномиальной теореме получив $\sum_n \binom{-1}{n} z^n = (1+z)^{-1}$.) Шаг 3 требует лишь элементарной алгебры и дает

$$G(z) = \frac{1+z(1+z)}{(1+z)(1-z-2z^2)} = \frac{1+z+z^2}{(1-2z)(1+z)^2}\,.$$

Теперь остается только шаг 4.

Определенную сложность представляет множитель в квадрате в знаменателе, поскольку, как мы знаем, в случае кратных корней анализ выполняется труднее, чем когда корни различны. Но ничего не поделаешь — нам попался именно этот случай. У знаменателя два корня, $\rho_1 = 2$ и $\rho_2 = -1$; общая теорема о разложении (7.30) гласит, что

$$g_n = a_1 2^n + (a_2 n + c)(-1)^n$$

с некоторой константой c, где

$$a_1 = \frac{1+1/2+1/4}{(1+1/2)^2} = \frac{7}{9}\,; \qquad a_2 = \frac{1-1+1}{1-2/(-1)} = \frac{1}{3}\,.$$

(Когда знаменатель имеет простые множители, удобнее использовать для вычисления a_k вторую формулу в (7.31), а не первую. Нам нужно лишь подставить $z = 1/\rho_k$ везде в $R(z)$, за исключением множителя, обращающегося в нуль, и поделить результат на $(d_k - 1)!$; это даст нам коэффициент при $n^{d_k-1}\rho_k^n$.) Подстановка $n = 0$ показывает, что неизвестная пока константа c должна быть равна $\frac{2}{9}$; следовательно, искомый ответ —

$$g_n = \tfrac{7}{9}2^n + \left(\tfrac{1}{3}n + \tfrac{2}{9}\right)(-1)^n\,. \qquad (7.33)$$

Несложно проверить случаи $n = 1$ и 2, просто чтобы убедиться в том, что мы не ошиблись. Возможно, следует проверить и случай $n = 3$, поскольку формула выглядит несколько странно. Но на самом деле все в порядке.

Можно ли было угадать формулу (7.33)? Ну, протабулировав значения далее, можно было бы заметить, что $g_{n+1} \approx 2g_n$ при больши́х n. При большой наблюдательности и удаче можно было бы додуматься до константы $\frac{7}{9}$. Но, конечно, гораздо проще и надежнее использовать в качестве инструмента производящие функции.

Пример 3: взаимно рекуррентные последовательности

Иногда встречаются два или более рекуррентных соотношений, зависящих одно от другого. В таких случаях можно образовать производящие функции для обеих последовательностей и решить их, немного расширив наш четырехходовый метод.

Например, вернемся к задаче о покрытии костями домино прямоугольника размером $3 \times n$, с которой мы уже работали в этой главе. Если мы хотим узнать только U_n — общее количество вариантов покрытия прямоугольника $3 \times n$ костями домино, не разделяя случаи вертикальных и горизонтальных костей домино, то нет необходимости вдаваться в такие подробности, как раньше. Можно просто записать рекуррентные соотношения

$$\begin{aligned} &U_0 = 1\,, \qquad U_1 = 0\,; \qquad V_0 = 0\,, \qquad V_1 = 1\,; \\ &U_n = 2V_{n-1} + U_{n-2}\,, \qquad V_n = U_{n-1} + V_{n-2} \qquad \text{при } n \geqslant 2. \end{aligned}$$

Здесь V_n — количество покрытий прямоугольника размером $3 \times n$ без угла с использованием $(3n-1)/2$ костей домино. Эти рекуррентные соотношения легко вывести, если рассмотреть возможные расположения костей домино на левой стороне прямоугольника, как мы делали ранее. Вот значения U_n и V_n при малых n:

n	0	1	2	3	4	5	6	7
U_n	1	0	3	0	11	0	41	0
V_n	0	1	0	4	0	15	0	56

(7.34)

Найдем за четыре шага выражение в аналитическом виде. Сначала (шаг 1) запишем

$$U_n \;=\; 2V_{n-1} + U_{n-2} + [n=0]\,, \qquad V_n \;=\; U_{n-1} + V_{n-2}$$

для всех n. Следовательно (шаг 2),

$$U(z) \;=\; 2zV(z) + z^2U(z) + 1\,, \qquad V(z) \;=\; zU(z) + z^2V(z)\,.$$

Теперь (шаг 3) мы должны решить два уравнения с двумя неизвестными; это легко сделать, получив из второго уравнения

$V(z) = zU(z)/(1-z^2)$; далее находим

$$U(z) = \frac{1-z^2}{1-4z^2+z^4}; \qquad V(z) = \frac{z}{1-4z^2+z^4}. \tag{7.35}$$

(Мы уже видели эту формулу для $U(z)$ в (7.10), но тогда в ней вместо z^2 было z^3. В том выводе n представляло собой количество костей домино, а здесь это ширина прямоугольника.)

Знаменатель $1-4z^2+z^4$ является функцией от z^2; это обеспечивает $U_{2n+1} = 0$ и $V_{2n} = 0$, как и должно быть. Можно извлечь выгоду из этого свойства, сохранив z^2 при разложении знаменателя на множители: нам вовсе не обязательно разлагать $1-4z^2+z^4$ в произведение четырех множителей вида $(1-\rho_k z)$, поскольку двух множителей вида $(1-\rho_k z^2)$ будет достаточно для определения коэффициентов. Другими словами, рассмотрев производящую функцию

$$W(z) = \frac{1}{1-4z+z^2} = W_0 + W_1\,z + W_2\,z^2 + \cdots, \tag{7.36}$$

получим $V(z) = zW(z^2)$ и $U(z) = (1-z^2)W(z^2)$; следовательно, $V_{2n+1} = W_n$ и $U_{2n} = W_n - W_{n-1}$. Мы сохраним время и силы, работая с более простой функцией $W(z)$.

Полином $1-4z+z^2$ разлагается на множители $(z-2-\sqrt{3})$ и $(z-2+\sqrt{3})$, их также можно записать в виде $\bigl(1-(2+\sqrt{3})z\bigr)$ и $\bigl(1-(2-\sqrt{3})z\bigr)$, потому что этот полином обратен сам себе. Таким образом, получаем

$$\begin{aligned} V_{2n+1} = W_n &= \frac{3+2\sqrt{3}}{6}(2+\sqrt{3})^n + \frac{3-2\sqrt{3}}{6}(2-\sqrt{3})^n; \\ U_{2n} = W_n - W_{n-1} &= \frac{3+\sqrt{3}}{6}(2+\sqrt{3})^n + \frac{3-\sqrt{3}}{6}(2-\sqrt{3})^n = \\ &= \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} + \frac{(2-\sqrt{3})^n}{3+\sqrt{3}}. \end{aligned} \tag{7.37}$$

Это и есть искомое выражение для количества покрытий прямоугольника размером $3 \times n$ костями домино.

Кстати, формулу для U_{2n} можно упростить, заметив, что второе слагаемое всегда лежит между 0 и 1. Число U_{2n} целое, так что

$$U_{2n} = \left\lceil \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} \right\rceil \qquad \text{при } n \geqslant 0. \tag{7.38}$$

В действительности второй член $(2-\sqrt{3})^n/(3+\sqrt{3})$ исчезающе мал при бóльших n, потому что $2-\sqrt{3} \approx 0.268$. Это следует иметь в виду при использовании формулы (7.38) для расчетов. Например, достаточно дорогой калькулятор, выпускаемый фирмой с хорошей репутацией, при вычислении $(2+\sqrt{3})^{10}/(3-\sqrt{3})$ дает 413403.0005. Этот ответ точен до девятой значащей цифры; однако точное значение немного *меньше* 413403, а не чуть-чуть больше. Таким образом, было бы ошибкой вычислять потолок от 413403.0005; точный ответ, $U_{20} = 413403$, получается *округлением* до ближайшего целого. Потолки могут быть опасными.

Как и полы...

Пример 4: выражение в аналитическом виде для задачи размена

Мы оставили задачу размена, подсчитав только количество способов выплаты 50 копеек. Давайте попробуем найти то же число для одного рубля (или миллиона), при том же способе оплаты — копейками, пятаками, гривенниками, четвертаками и полтинниками.

Производящая функция, которую мы вывели ранее, имеет вид

$$C(z) = \frac{1}{1-z}\,\frac{1}{1-z^5}\,\frac{1}{1-z^{10}}\,\frac{1}{1-z^{25}}\,\frac{1}{1-z^{50}};$$

это рациональная функция от z, знаменатель которой имеет степень 91. Таким образом, можно разложить знаменатель на 91 множитель и выразить C_n — количество способов выплаты n копеек — в "аналитическом виде" из 91 слагаемого. Но это слишком грустное решение. Нельзя ли в данном конкретном случае найти что-то более красивое?

Первый проблеск надежды появляется, когда мы замечаем, что знаменатель почти является функцией от z^5. Тот же трюк, использованный нами для упрощения вычислений в случае знаменателя $1-4z^2+z^4$, который представлял собой функцию от z^2, можно применить и к $C(z)$, заменив $1/(1-z)$ на $(1+z+z^2+z^3+z^4)/(1-z^5)$:

$$\begin{aligned}
C(z) &= \frac{1+z+z^2+z^3+z^4}{1-z^5}\,\frac{1}{1-z^5}\,\frac{1}{1-z^{10}}\,\frac{1}{1-z^{25}}\,\frac{1}{1-z^{50}} = \\
&= (1+z+z^2+z^3+z^4)\check{C}(z^5),\\
\check{C}(z) &= \frac{1}{1-z}\,\frac{1}{1-z}\,\frac{1}{1-z^2}\,\frac{1}{1-z^5}\,\frac{1}{1-z^{10}}.
\end{aligned}$$

Знаменатель "сжатой" функции $\check{C}(z)$ имеет степень, равную всего лишь 19, так что эта функция гораздо лучше исходной. Новое выражение для $C(z)$, кстати, показывает, что $C_{5n} = C_{5n+1} = C_{5n+2} = C_{5n+3} = C_{5n+4}$; и действительно, это соотношение легко объяснить: количество способов выплаты 53 копеек такое же, как и количество способов выплаты 50 копеек, потому что количество копеек по модулю 5 предопределено.

О сжатой функции говорим сжато.

Но даже для $\check{C}(z)$ не существует простого выражения, основанного на корнях знаменателя. Вероятно, простейший способ вычисления коэффициентов $\check{C}(z)$ получится, если заметить, что каждый сомножитель в знаменателе является делителем $1-z^{10}$. Следовательно, можно записать

$$\check{C}(z) \;=\; \frac{A(z)}{(1-z^{10})^5}\,, \quad \text{где } A(z) = A_0+A_1z+\cdots+A_{31}z^{31}. \qquad (7.39)$$

Для полноты картины вот как выглядит $A(z)$:

$$\begin{aligned}
&(1+z+\cdots+z^9)^2(1+z^2+\cdots+z^8)(1+z^5) \;=\\
&\quad=\; 1+2z+4z^2+6z^3+9z^4+13z^5+18z^6+24z^7+31z^8+\\
&\qquad+39z^9+45z^{10}+52z^{11}+57z^{12}+63z^{13}+67z^{14}+\\
&\qquad+69z^{15}+69z^{16}+67z^{17}+63z^{18}+57z^{19}+52z^{20}+\\
&\qquad+45z^{21}+39z^{22}+31z^{23}+24z^{24}+18z^{25}+13z^{26}+\\
&\qquad+9z^{27}+6z^{28}+4z^{29}+2z^{30}+z^{31}\,.
\end{aligned}$$

В завершение, поскольку $1/(1-z^{10})^5 = \sum_{k\geqslant 0}\binom{k+4}{4}z^{10k}$, можно определить коэффициент $\check{C}_n = [z^n]\,\check{C}(z)$ следующим образом:

$$\begin{aligned}
\check{C}_{10q+r} \;&=\; \sum_{j,k} A_j\binom{k+4}{4}[10q+r=10k+j]\\
&=\; A_r\binom{q+4}{4}+A_{r+10}\binom{q+3}{4}+A_{r+20}\binom{q+2}{4}+\\
&\qquad+A_{r+30}\binom{q+1}{4}\,,
\end{aligned} \qquad (7.40)$$

где $n = 10q+r$ и $0 \leqslant r < 10$. Это дает нам десять случаев, по одному для каждого значения r; но это все же неплохая формула по сравнению с другими альтернативами, включающими степени комплексных чисел.

Например, это выражение можно использовать для определения значения $C_{50q} = \check{C}_{10q}$. Тогда $r = 0$, и мы имеем

$$C_{50q} \;=\; \binom{q+4}{4}+45\binom{q+3}{4}+52\binom{q+2}{4}+2\binom{q+1}{4}\,.$$

Количество способов оплатить 50 коп. равно $\binom{5}{4}+45\binom{4}{4} = 50$; один рубль можно заплатить $\binom{6}{4}+45\binom{5}{4}+52\binom{4}{4} = 292$ способами, а для 1 000 000 рублей это количество составляет

$$\binom{2000004}{4}+45\binom{2000003}{4}+52\binom{2000002}{4}+2\binom{2000001}{4} \;=\\
=\; 66666793333412666685000001\,.$$

Пример 5: расходящийся ряд

Попытаемся теперь найти выражение в аналитическом виде для чисел g_n, определяемых соотношениями

$$\begin{aligned} g_0 &= 1; \\ g_n &= ng_{n-1} \qquad \text{при } n > 0. \end{aligned}$$

Поразмышляв несколько наносекунд, легко сообразить, что g_n — это просто $n!$; метод суммирующих множителей из главы 2 даст этот ответ едва ли не еще быстрее. Но попробуем найти решение при помощи производящих функций хотя бы для того, чтобы посмотреть, что же у нас получится. (Действительно мощный метод должен работать и для простых соотношений наподобие приведенного, и в сложных случаях, когда угадать ответ представляется невозможным.)

От нанотехнологий не спрятаться даже в математике.

Уравнение

$$g_n = ng_{n-1} + [n=0]$$

справедливо при всех n и приводит к

$$G(z) = \sum_n g_n z^n = \sum_n ng_{n-1}\, z^n + \sum_{n=0} z^n\,.$$

Для завершения шага 2 надо выразить $\sum_n ng_{n-1}\, z^n$ через $G(z)$; табл. 406 (основных методов преобразования) подсказывает, что здесь должна помочь производная: $G'(z) = \sum_n ng_n z^{n-1}$. Подгоним наше выражение к сумме требуемого вида:

$$\begin{aligned} G(z) &= 1 + \sum_n (n+1) g_n\, z^{n+1} = \\ &= 1 + \sum_n ng_n\, z^{n+1} + \sum_n g_n\, z^{n+1} = \\ &= 1 + z^2G'(z) + zG(z)\,. \end{aligned}$$

Проверим это уравнение, подставляя значения g_n для малых n. Поскольку

$$\begin{aligned} G &= 1 + z + 2z^2 + 6z^3 + 24z^4 + \cdots\,, \\ G' &= \phantom{1 + z + {}} 1 + 4z + 18z^2 + 96z^3 + \cdots\,, \end{aligned}$$

имеем

$$\begin{aligned} z^2G' &= \phantom{z + {}} z^2 + 4z^3 + 18z^4 + 96z^5 + \cdots\,, \\ zG &= z + z^2 + 2z^3 + 6z^4 + 24z^5 + \cdots\,, \\ 1 &= 1\,. \end{aligned}$$

Суммирование этих строк дает нам G, так что пока все в порядке. Кстати, зачастую удобнее писать 'G' вместо '$G(z)$'; лишнее '(z)' только загромождает формулу, если мы не заменяем z чем-нибудь еще.

Теперь идет шаг 3, и он отличается от того, что мы делали раньше, поскольку в данном случае нам предстоит решить дифференциальное уравнение. С этим дифференциальным уравнением мы можем справиться при помощи гипергеометрических рядов из раздела 5.6 — эти методы не так уж плохи. (Если вы не знакомы с гипергеометрическими функциями — не тревожьтесь, все будет сделано очень быстро.)

"Это будет быстро",— утешал меня доктор, втыкая в меня настоящий гипершприц...

Начнем с избавления от константы 1, для чего возьмем производную от обеих частей:

$$\begin{aligned} G' = (z^2G' + zG + 1)' &= (2zG' + z^2G'') + (G + zG') = \\ &= z^2G'' + 3zG' + G\,. \end{aligned}$$

Теория из главы 5 требует от нас переписать все это с применением оператора ϑ, а из упр. 6.13 мы знаем, что

$$\vartheta G = zG'\,, \qquad \vartheta^2 G = z^2G'' + zG'\,.$$

Итак, дифференциальное уравнение в требуемом нам виде выглядит следующим образом:

$$\vartheta G = z\vartheta^2 G + 2z\vartheta G + zG = z(\vartheta + 1)^2 G\,.$$

В соответствии с (5.109) решением при $g_0 = 1$ является гипергеометрический ряд $F(1,1;;z)$.

Шаг 3 оказался сложнее, чем мы ожидали; но теперь, когда мы знаем функцию G, шаг 4 очень прост — определение гипергеометрического ряда (5.76) дает требуемое разложение:

$$G(z) = F\left({1,1 \atop } \middle| z\right) = \sum_{n\geqslant 0} \frac{1^{\overline{n}}\,1^{\overline{n}}\,z^n}{n!} = \sum_{n\geqslant 0} n!\,z^n\,.$$

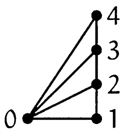
Мы подтвердили выражение, которое знали все время: $g_n = n!$.

Обратите внимание, что метод производящих функций дал правильный ответ несмотря на то, что $G(z)$ расходится при всех ненулевых z. Последовательность $n!$ растет настолько быстро, что член $|n!\,z^n|$ стремится к ∞ при $n \to \infty$, за исключением точки $z = 0$. Это показывает, что с формальными степенными рядами можно выполнять алгебраические действия, не заботясь о сходимости.

Пример 6: рекуррентное соотношение с возвратом

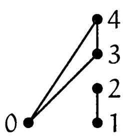
В завершение этого раздела рассмотрим применение производящих функций к задаче из теории графов. *Фан* порядка n —

это граф с вершинами $\{0, 1, \ldots, n\}$ и $2n-1$ ребрами, определяемый следующим образом: вершина 0 соединяется ребрами с каждой из остальных n вершин, а вершина k соединяется с вершиной $k+1$ для $1 \leqslant k < n$. Ниже, например, приведен фан порядка 4, имеющий пять вершин и семь ребер.



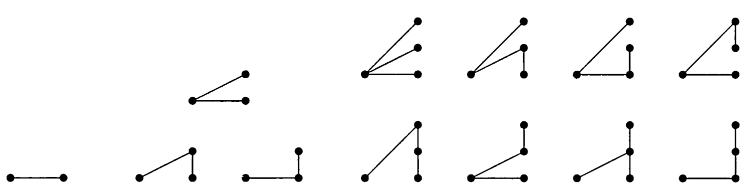
Интересующая нас задача состоит в следующем: сколько в таком графе остовных деревьев f_n? *Остовным деревом* называется подграф, содержащий все вершины и такое количество ребер, чтобы этот подграф был связным, но не настолько много, чтобы в нем появился хоть один цикл. Оказывается, что любое остовное дерево в графе с $n+1$ вершинами имеет ровно n ребер. Подграф менее чем с n ребрами не будет связным, а если число ребер превысит n, в подграфе появится цикл; это доказывается в книгах по теории графов.

Существует $\binom{2n-1}{n}$ способов выбрать n ребер из $2n-1$, имеющихся в фане порядка n, но не любой такой выбор дает остовное дерево. Например, подграф



имеет четыре ребра, но не является остовным деревом; в нем есть цикл $0-4-3-0$ и нет связи между вершинами $\{1,2\}$ и остальными. Мы хотим подсчитать, сколько из $\binom{2n-1}{n}$ вариантов дают остовное дерево.

Рассмотрим несколько начальных случаев. Можно легко перечислить все остовные деревья при $n = 1$, 2 и 3:

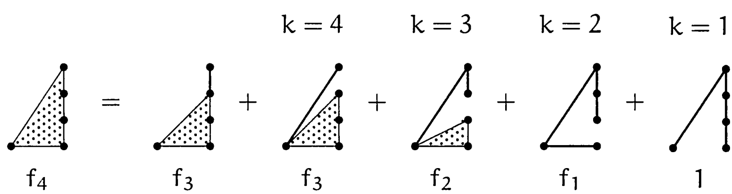


$f_1 = 1$ $\qquad f_2 = 3$ $\qquad f_3 = 8$

(Можно не указывать метки у вершин, если всегда рисовать вершину 0 слева.) Что можно сказать о случае $n = 0$? На первый взгляд, кажется разумным положить $f_0 = 1$; но мы примем

$f_0 = 0$, потому что существование фана порядка 0 (который должен иметь $2n - 1 = -1$ ребро) вызывает серьезные сомнения.

Наша четырехходовая процедура предписывает найти рекуррентное соотношение для f_n, которое выполняется при всех n. Для получения такого соотношения рассмотрим, как самая верхняя вершина (вершина n) соединяется с остальной частью остовного дерева. Если она не соединена с вершиной 0, то она должна быть соединена с вершиной $n - 1$, поскольку эта последняя должна соединяться с остальным графом. В этом случае любое из f_{n-1} остовных деревьев для остающегося графа (фана с вершинами от 0 до $n - 1$) вместе с новым ребром образует остовное дерево для всего графа. В противном случае вершина n *соединена* с вершиной 0, и найдется некоторое число $k \leqslant n$, такое, что вершины $n, n-1, \ldots, k$ непосредственно соединены между собой, но ребро между вершинами k и $k - 1$ не входит в поддерево. Тогда в дереве не может быть ребер между 0 и $\{n - 1, \ldots, k\}$, иначе получится цикл. Следовательно, для $k = 1$ остовное дерево полностью определено. Если же $k > 1$, то для завершения остовного дерева достаточно добавить любое из f_{k-1} остовных деревьев на вершинах $\{0, 1, \ldots, k - 1\}$. Например, вот что дает этот анализ для $n = 4$:



Общее уравнение, справедливое для $n \geqslant 1$, имеет вид

$$f_n = f_{n-1} + f_{n-1} + f_{n-2} + f_{n-3} + \cdots + f_1 + 1 .$$

(Выглядит все так, как будто '1' в конце есть f_0, так что нам следовало бы принять $f_0 = 1$; но мы останемся при своей точке зрения.) Небольшое изменение делает уравнение справедливым при всех целых n:

$$f_n = f_{n-1} + \sum_{k<n} f_k + [n>0] . \tag{7.41}$$

Это рекуррентное соотношение как раз содержит "возврат к началу" от f_{n-1} через все предшествующие значения, и в этом главное отличие данного рекуррентного соотношения от всех рекуррентных соотношений, с которыми мы имели дело ранее в этой главе. Чтобы избавиться от аналогичной суммы в правой части рекуррентного соотношения (2.12), в главе 2 был применен специальный прием, а именно — вычитание одного экземпляра соотношения из другого ($f_{n+1} - f_n$). Здесь этот способ тоже годится; но, как будет показано далее, метод производящих функций позволяет непосредственно работать с суммами такого рода. (И это хорошо, потому что вскоре нам придется иметь дело с гораздо более сложными рекуррентными соотношениями.)

Шаг 1 завершен; на шаге 2 нам придется применить новый прием:

$$\begin{aligned} F(z) &= \sum_n f_n z^n = \sum_n f_{n-1} z^n + \sum_{k,n} f_k z^n [k<n] + \sum_n [n>0] z^n = \\ &= zF(z) + \sum_k f_k z^k \sum_n [n>k] z^{n-k} + \frac{z}{1-z} = \\ &= zF(z) + F(z) \sum_{m>0} z^m + \frac{z}{1-z} = \\ &= zF(z) + F(z)\frac{z}{1-z} + \frac{z}{1-z}. \end{aligned}$$

Ключевой трюк заключается в замене z^n на $z^k z^{n-k}$, что позволяет выразить двойную сумму через $F(z)$, как и требуется на шаге 2.

Шаг 3 сводится к простой алгебре, и мы находим

$$F(z) = \frac{z}{1-3z+z^2}.$$

Обладатели хорошей памяти узнают в этом выражении производящую функцию (7.24) для чисел Фибоначчи с четными номерами. Так что шаг 4 оказывается лишним; мы уже нашли несколько неожиданный ответ задачи о количестве остовных деревьев фанов:

$$f_n = F_{2n} \qquad \text{при } n \geqslant 0. \tag{7.42}$$

7.4 Специальные производящие функции

Пройти шаг 4 нашей четырехходовой процедуры будет проще, если мы будем знать коэффициенты большого количества разных степенных рядов. Разложения в табл. 407 достаточно полезны, если они применимы, однако имеется масса других типов выражений в аналитическом виде. Таким образом, следует дополнить эту таблицу еще одной, в которую будут включены степенные ряды, соответствующие “специальным числам”, рассматривавшимся в главе 6.

Таблица 425. Производящие функции для специальных чисел

$$\frac{1}{(1-z)^{m+1}} \ln \frac{1}{1-z} = \sum_{n\geqslant 0} (H_{m+n}-H_m) \binom{m+n}{n} z^n \qquad (7.43)$$

$$\frac{z}{e^z-1} = \sum_{n\geqslant 0} B_n \frac{z^n}{n!} \qquad (7.44)$$

$$\frac{F_m z}{1-(F_{m-1}+F_{m+1})z+(-1)^m z^2} = \sum_{n\geqslant 0} F_{mn}\, z^n \qquad (7.45)$$

$$\sum_k \left\{ {m \atop k} \right\} \frac{k!\, z^k}{(1-z)^{k+1}} = \sum_{n\geqslant 0} n^m z^n \qquad (7.46)$$

$$(z^{-1})^{\overline{-m}} = \frac{z^m}{(1-z)(1-2z)\ldots(1-mz)} = \sum_{n\geqslant 0} \left\{ {n \atop m} \right\} z^n \qquad (7.47)$$

$$z^{\overline{m}} = z(z+1)\ldots(z+m-1) = \sum_{n\geqslant 0} \left[{m \atop n} \right] z^n \qquad (7.48)$$

$$(e^z-1)^m = m! \sum_{n\geqslant 0} \left\{ {n \atop m} \right\} \frac{z^n}{n!} \qquad (7.49)$$

$$\left(\ln \frac{1}{1-z}\right)^m = m! \sum_{n\geqslant 0} \left[{n \atop m} \right] \frac{z^n}{n!} \qquad (7.50)$$

$$\left(\frac{z}{\ln(1+z)}\right)^m = \sum_{n\geqslant 0} \frac{z^n}{n!} \left\{ {m \atop m-n} \right\} \Big/ \binom{m-1}{n} \qquad (7.51)$$

$$\left(\frac{z}{1-e^{-z}}\right)^m = \sum_{n\geqslant 0} \frac{z^n}{n!} \left[{m \atop m-n} \right] \Big/ \binom{m-1}{n} \qquad (7.52)$$

$$e^{z+wz} = \sum_{m,n\geqslant 0} \binom{n}{m} w^m \frac{z^n}{n!} \qquad (7.53)$$

$$e^{w(e^z-1)} = \sum_{m,n\geqslant 0} \left\{ {n \atop m} \right\} w^m \frac{z^n}{n!} \qquad (7.54)$$

$$\frac{1}{(1-z)^w} = \sum_{m,n\geqslant 0} \left[{n \atop m} \right] w^m \frac{z^n}{n!} \qquad (7.55)$$

$$\frac{1-w}{e^{(w-1)z}-w} = \sum_{m,n\geqslant 0} \left\langle {n \atop m} \right\rangle w^m \frac{z^n}{n!} \qquad (7.56)$$

Табл. 425 представляет собой требующуюся нам базу данных. Тождества из этой таблицы нетрудно доказать, так что не будем задерживаться на этом вопросе; данная таблица предназначена, в первую очередь, служить справочником, когда мы столкнемся с некоторой новой задачей. Однако для первой формулы (7.43) имеется красивое доказательство, на котором хочется остановиться. Начнем с тождества

$$\frac{1}{(1-z)^{x+1}} = \sum_n \binom{x+n}{n} z^n$$

и продифференцируем его по x. В левой части выражение $(1-z)^{-x-1}$ равно $e^{(x+1)\ln(1/(1-z))}$, так что d/dx внесет множитель $\ln\bigl(1/(1-z)\bigr)$. В правой части числитель $\binom{x+n}{n}$ равен $(x+n)\ldots(x+1)$, и d/dx разобьет его на сумму n слагаемых, что эквивалентно умножению $\binom{x+n}{n}$ на

$$\frac{1}{x+n} + \cdots + \frac{1}{x+1} = H_{x+n} - H_x \,.$$

Замена x на m дает (7.43). Заметьте, что $H_{x+n} - H_x$ имеет смысл даже при нецелом x.

Кстати, при дифференцировании сложных произведений, как оказывается, лучше оставлять их в виде произведений, а не записывать производную как сумму. Например, правая часть тождества

$$\frac{d}{dx}\bigl((x+n)^n \ldots (x+1)^1\bigr) = (x+n)^n \ldots (x+1)^1 \left(\frac{n}{x+n} + \cdots + \frac{1}{x+1}\right)$$

была бы менее наглядна, если бы была записана в виде суммы.

У общих тождеств из табл. 425 имеется много важных частных случаев. Например, формула (7.43) при $m = 0$ упрощается до производящей функции для чисел H_n:

$$\frac{1}{1-z} \ln \frac{1}{1-z} = \sum_n H_n z^n \,. \qquad (7.57)$$

Это уравнение можно вывести и иначе; например, можно взять степенной ряд для $\ln\bigl(1/(1-z)\bigr)$ и поделить его на $1-z$, получив тем самым производящую функцию для частичных сумм.

В тождествах (7.51) и (7.52) встречаются отношения $\left\{ {m \atop m-n} \right\} / \binom{m-1}{n}$ и $\left[{m \atop m-n} \right] / \binom{m-1}{n}$ соответственно, которые для $n \geqslant m$ превращаются в неопределенности вида $0/0$. Однако имеется способ придания им смысла с использованием полиномов Стирлинга из (6.45), потому что

$$\left\{ {m \atop m-n} \right\} \Big/ \binom{m-1}{n} = (-1)^{n+1} n!\, m\sigma_n(n-m)\,; \tag{7.58}$$

$$\left[{m \atop m-n} \right] \Big/ \binom{m-1}{n} = n!\, m\sigma_n(m)\,. \tag{7.59}$$

Так, например, при $m = 1$ формулу (7.51) следует записывать не как степенной ряд $\sum_{n\geqslant 0}(z^n/n!)\left\{ {1 \atop 1-n} \right\}/\binom{0}{n}$, а как

$$\frac{z}{\ln(1+z)} = -\sum_{n\geqslant 0}(-z)^n\sigma_n(n-1) = 1 + \tfrac{1}{2}z - \tfrac{1}{12}z^2 + \cdots\,.$$

Тождества (7.53), (7.54), (7.55) и (7.56) представляют собой двойные производящие функции, или суперпроизводящие функции, потому что они имеют вид $G(w,z) = \sum_{m,n} g_{m,n}w^m z^n$. Коэффициент при w^m представляет собой производящую функцию относительно переменной z; коэффициент при z^n представляет собой производящую функцию относительно переменной w. Уравнению (7.56) можно придать более симметричный вид:

$$\frac{e^w - e^z}{we^z - ze^w} = \sum_{m,n\geqslant 0} \left\langle {m+n+1 \atop m} \right\rangle \frac{w^m z^n}{(m+n+1)!}\,. \tag{7.60}$$

7.5 Свертки

А я всегда думал, что свертка — это то, что происходит с моими мозгами при решении задач...

Свертка двух последовательностей, $\langle f_0, f_1, \ldots\rangle = \langle f_n\rangle$ и $\langle g_0, g_1, \ldots\rangle = \langle g_n\rangle$, представляет собой последовательность $\langle f_0g_0,\ f_0g_1 + f_1g_0,\ \ldots\rangle = \langle\sum_k f_k g_{n-k}\rangle$. В разделах 5.4 и 7.2 мы видели, что свертка последовательностей соответствует умножению их производящих функций. Этот факт позволяет легко вычислить многие суммы, с которыми было бы трудно справиться другими методами.

Пример 1: фибоначчиевы свертки

Попробуем, например, найти аналитическое выражение для суммы $\sum_{k=0}^{n} F_k F_{n-k}$. Это — свертка последовательности $\langle F_n\rangle$ с самой собой, поэтому интересующая нас сумма должна быть коэффициентом при z^n в $F(z)^2$, где $F(z)$ — производящая функция для последовательности $\langle F_n\rangle$. Все, что требуется от нас, — это вычислить значение указанного коэффициента.

Производящая функция $F(z)$ есть $z/(1-z-z^2)$ — отношение полиномов; так что, как нам известно из общей теоремы о разложении рациональных функций, ответ можно получить, найдя разложение в правильные дроби. Можно воспользоваться общей теоремой о разложении (7.30) и проделать требуемые вычисления. А можно использовать представление

$$\begin{aligned} F(z)^2 &= \left(\frac{1}{\sqrt{5}}\left(\frac{1}{1-\phi z}-\frac{1}{1-\hat{\phi} z}\right)\right)^2 = \\ &= \frac{1}{5}\left(\frac{1}{(1-\phi z)^2}-\frac{2}{(1-\phi z)(1-\hat{\phi} z)}+\frac{1}{(1-\hat{\phi} z)^2}\right) = \\ &= \frac{1}{5}\sum_{n\geqslant 0}(n+1)\phi^n z^n-\frac{2}{5}\sum_{n\geqslant 0}F_{n+1}z^n+\frac{1}{5}\sum_{n\geqslant 0}(n+1)\hat{\phi}^n z^n . \end{aligned}$$

Вместо того чтобы записать ответ через ϕ и $\hat{\phi}$, давайте попробуем найти аналитический вид в терминах чисел Фибоначчи. Вспомнив, что $\phi+\hat{\phi}=1$, получим

$$\begin{aligned} \phi^n+\hat{\phi}^n &= [z^n]\left(\frac{1}{1-\phi z}+\frac{1}{1-\hat{\phi} z}\right) = \\ &= [z^n]\,\frac{2-(\phi+\hat{\phi})z}{(1-\phi z)(1-\hat{\phi} z)} = [z^n]\,\frac{2-z}{1-z-z^2} = 2F_{n+1}-F_n . \end{aligned}$$

Следовательно,

$$F(z)^2 = \frac{1}{5}\sum_{n\geqslant 0}(n+1)(2F_{n+1}-F_n)z^n-\frac{2}{5}\sum_{n\geqslant 0}F_{n+1}\,z^n ,$$

и мы получаем искомый ответ:

$$\sum_{k=0}^{n}F_kF_{n-k} = \frac{2nF_{n+1}-(n+1)F_n}{5} . \qquad (7.61)$$

Например, при $n=3$ эта формула дает $F_0F_3+F_1F_2+F_2F_1+F_3F_0 = 0+1+1+0=2$ в левой части и $(6F_4-4F_3)/5=(18-8)/5=2$ — в правой.

Пример 2: гармонические свертки

Эффективность работы некоего компьютерного алгоритма "сортировки по образцу" определяется значением суммы

$$T_{m,n} = \sum_{0\leqslant k<n}\binom{k}{m}\frac{1}{n-k}, \qquad \text{целые } m,n\geqslant 0.$$

В упр. 5.58 эта сумма вычисляется при помощи довольно хитроумной двойной индукции с использованием суммирующих множителей. Получить результат можно гораздо проще, если заметить, что $T_{m,n}$ представляет собой n-й член свертки $\langle\binom{0}{m},\binom{1}{m},$

$\binom{2}{m}, \ldots\rangle$ с $\langle 0, \frac{1}{1}, \frac{1}{2}, \ldots\rangle$. Обе последовательности в табл. 407 имеют простые производящие функции:

$$\sum_{n\geqslant 0}\binom{n}{m}z^n = \frac{z^m}{(1-z)^{m+1}}; \qquad \sum_{n>0}\frac{z^n}{n} = \ln\frac{1}{1-z}.$$

Следовательно, согласно (7.43)

$$T_{m,n} = [z^n]\,\frac{z^m}{(1-z)^{m+1}}\ln\frac{1}{1-z} = [z^{n-m}]\,\frac{1}{(1-z)^{m+1}}\ln\frac{1}{1-z} =$$
$$= (H_n-H_m)\binom{n}{n-m}.$$

В действительности имеется гораздо большее количество сумм, сводящихся к сверткам такого вида, поскольку для любых r и s

$$\frac{1}{(1-z)^{r+1}}\ln\frac{1}{1-z}\cdot\frac{1}{(1-z)^{s+1}} = \frac{1}{(1-z)^{r+s+2}}\ln\frac{1}{1-z}.$$

Приравнивание коэффициентов при z^n дает нам общее тождество

$$\sum_k\binom{r+k}{k}\binom{s+n-k}{n-k}(H_{r+k}-H_r)$$
$$= \binom{r+s+n+1}{n}(H_{r+s+n+1}-H_{r+s+1}). \tag{7.62}$$

Оно такое гармоничное...

Оно выглядит слишком хорошо, чтобы быть правдой. Но проверка, по крайней мере при $n=2$, проходит успешно:

$$\binom{r+1}{1}\binom{s+1}{1}\frac{1}{r+1}+\binom{r+2}{2}\binom{s+0}{0}\left(\frac{1}{r+2}+\frac{1}{r+1}\right) =$$
$$= \binom{r+s+3}{2}\left(\frac{1}{r+s+3}+\frac{1}{r+s+2}\right).$$

Частные случаи наподобие $s=0$ столь же замечательны, как и общее тождество.

И это еще не все. Можно воспользоваться тождеством для свертки

$$\sum_k\binom{r+k}{k}\binom{s+n-k}{n-k} = \binom{r+s+n+1}{n}$$

и перенести H_r в другую часть, поскольку H_r не зависит от k:

$$\sum_k\binom{r+k}{k}\binom{s+n-k}{n-k}H_{r+k} =$$
$$= \binom{r+s+n+1}{n}(H_{r+s+n+1}-H_{r+s+1}+H_r). \tag{7.63}$$

Но и это не все. Если r и s представляют собой неотрицательные целые числа l и m, то можно заменить $\binom{r+k}{k}$ на $\binom{l+k}{l}$ и $\binom{s+n-k}{n-k}$ на $\binom{m+n-k}{m}$; затем, заменив k на $k-l$ и n на $n-m-l$, получим

$$\sum_{k=0}^{n}\binom{k}{l}\binom{n-k}{m}H_k = \binom{n+1}{l+m+1}(H_{n+1}-H_{l+m+1}+H_l), \quad \text{целые } l,m,n \geqslant 0. \tag{7.64}$$

Даже частный случай этого тождества $l=m=0$ вызывал у нас трудности в главе 2! (См. (2.36).) Мы проделали немалый путь.

Пример 3: свертки сверток

Если образовать свертку последовательностей $\langle f_n\rangle$ и $\langle g_n\rangle$, а затем свернуть результат с третьей последовательностью $\langle h_n\rangle$, то можно получить последовательность, n-й член которой равен

$$\sum_{j+k+l=n} f_j\, g_k\, h_l\,.$$

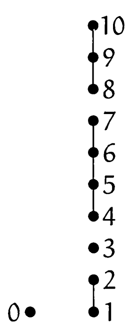
Производящей функцией этой тройной свертки будет, разумеется, произведение $F(z)G(z)H(z)$. Аналогично m-кратная свертка последовательности $\langle g_n\rangle$ с самой собой имеет n-й член вида

$$\sum_{k_1+k_2+\cdot+k_m=n} g_{k_1}\, g_{k_2}\,\cdots\, g_{k_m},$$

а ее производящая функция равна $G(z)^m$.

Эти соображения можно применить к рассмотренной ранее задаче об остовных деревьях фана (пример 6 в разделе 7.3). Оказывается, существует и другой способ вычислить количество остовных деревьев в n-фане f_n, основанный на анализе конфигураций ребер дерева, соединяющих вершины $\{1,2,\ldots,n\}$. Ребро между вершинами k и $k+1$ может быть включено в поддерево или не включено в него; и каждый способ выбора таких ребер задает разбиение множества вершин на блоки соединенных соседних вершин. Например, если $n=10$, можно соединить вершины $\{1,2\}$, $\{3\}$, $\{4,5,6,7\}$ и $\{8,9,10\}$:

Железобетонные блоки.



Сколько же можно получить остовных деревьев, добавляя ребра в вершину 0? Нам нужно соединить вершину 0 с каждым из четырех блоков; существует два способа соединения 0 с $\{1, 2\}$, один способ соединения с $\{3\}$, четыре способа — с $\{4, 5, 6, 7\}$ и три способа — с $\{8, 9, 10\}$, что дает в итоге $2 \cdot 1 \cdot 4 \cdot 3 = 24$ способа. Суммирование по всем способам разбиения на блоки дает следующее выражение для общего числа остовных деревьев:

$$f_n = \sum_{m>0} \sum_{\substack{k_1+k_2+\cdots+k_m=n \\ k_1,k_2,\ldots,k_m>0}} k_1 k_2 \ldots k_m \,. \tag{7.65}$$

Например, $f_4 = 4+3\cdot1+2\cdot2+1\cdot3+2\cdot1\cdot1+1\cdot2\cdot1+1\cdot1\cdot2+1\cdot1\cdot1\cdot1 = 21$.

Это — сумма m-кратных сверток последовательности $\langle 0, 1, 2, 3, \ldots \rangle$ для $m = 1, 2, 3, \ldots$; следовательно, производящей функцией для $\langle f_n \rangle$ будет

$$F(z) = G(z) + G(z)^2 + G(z)^3 + \cdots = \frac{G(z)}{1 - G(z)},$$

где $G(z)$ — производящая функция последовательности $\langle 0, 1, 2, 3, \ldots \rangle$, а именно — $z/(1-z)^2$. Таким образом, имеем

$$F(z) = \frac{z}{(1-z)^2 - z} = \frac{z}{1 - 3z + z^2},$$

как и ранее. Такой подход к вычислению $\langle f_n \rangle$ симметричнее и привлекательнее хитроумного рекуррентного соотношения, с которым мы имели дело ранее.

Пример 4: рекуррентное соотношение со сверткой

Наш следующий пример особенно важен. Это на самом деле классический пример, иллюстрирующий полезность производящих функций при решении рекуррентных соотношений.

Предположим, что у нас есть $n + 1$ переменных $x_0, x_1, \ldots, x_n$ и мы хотим вычислить их произведение при помощи n умножений. Сколько имеется способов (C_n) расставить скобки в произведении $x_0 \cdot x_1 \cdot \ldots \cdot x_n$ так, чтобы порядок умножений был полностью определен? Например, для $n = 2$ имеется два способа: $x_0 \cdot (x_1 \cdot x_2)$ и $(x_0 \cdot x_1) \cdot x_2$. При $n = 3$ имеется пять способов:

$$x_0 \cdot (x_1 \cdot (x_2 \cdot x_3)), \quad x_0 \cdot ((x_1 \cdot x_2) \cdot x_3), \quad (x_0 \cdot x_1) \cdot (x_2 \cdot x_3),$$
$$(x_0 \cdot (x_1 \cdot x_2)) \cdot x_3, \quad ((x_0 \cdot x_1) \cdot x_2) \cdot x_3 \,.$$

Таким образом, $C_2 = 2$, $C_3 = 5$; имеем также $C_1 = 1$ и $C_0 = 1$.

Попробуем применить четырехходовую процедуру из раздела 7.3. Какому же рекуррентному соотношению удовлетворяет последовательность C_i? Ключевым моментом здесь является то наблюдение, что для $n > 0$ ровно одна операция '$\cdot$' лежит вне всех скобок; это последнее умножение, которое связывает все в единое целое. Если эта операция '$\cdot$' располагается между x_k и x_{k+1}, то имеется C_k способов расстановки скобок в произведении $x_0 \cdot \ldots \cdot x_k$ и C_{n-k-1} способов — в $x_{k+1} \cdot \ldots \cdot x_n$; следовательно,

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-1} C_0\,, \qquad \text{если } n > 0.$$

В этом выражении мы легко распознаем свертку, и формулу нетрудно подправить так, чтобы она оставалась справедливой при всех целых n:

$$C_n = \sum_k C_k C_{n-1-k} + [n=0]\,. \tag{7.66}$$

Шаг 1 на этом завершен. Шаг 2 предписывает умножение на z^n и суммирование:

$$\begin{aligned} C(z) &= \sum_n C_n z^n = \\ &= \sum_{k,n} C_k C_{n-1-k} z^n + \sum_{n=0} z^n = \\ &= \sum_k C_k z^k \sum_n C_{n-1-k} z^{n-k} + 1 = \\ &= C(z) \cdot zC(z) + 1\,. \end{aligned}$$

Подумать только — свертка стала произведением! В мире производящих функций бывает и не такое. Жизнь полна сюрпризов.

Авторы шутят.

Шаг 3 тоже прост. Функцию $C(z)$ мы находим по формуле для корней квадратного уравнения:

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

Но какой знак следует выбрать, $+$ или $-$? Обе функции удовлетворяют уравнению $C(z) = zC(z)^2 + 1$, но решение задачи дает только одна из них. Можно выбрать знак $+$ просто на том основании, что всегда следует поступать положительно; но мы быстро выясним, что при этом $C(0) = \infty$, что противоречит фактам. (У правильной функции $C(z)$ должно быть $C(0) = C_0 = 1$.) Таким образом, мы приходим к выводу, что

$$C(z) = \frac{1 - \sqrt{1-4z}}{2z}\,.$$

Наконец, шаг 4. Чему равен коэффициент $[z^n]\,C(z)$? Биномиальная теорема говорит, что

$$\sqrt{1-4z} \;=\; \sum_{k\geqslant 0}\binom{1/2}{k}(-4z)^k \;=\; 1+\sum_{k\geqslant 1}\frac{1}{2k}\binom{-1/2}{k-1}(-4z)^k\,;$$

следовательно, используя (5.37), получаем

$$\frac{1-\sqrt{1-4z}}{2z} \;=\; \sum_{k\geqslant 1}\frac{1}{k}\binom{-1/2}{k-1}(-4z)^{k-1} \;=\; \sum_{n\geqslant 0}\binom{-1/2}{n}\frac{(-4z)^n}{n+1} \;=\; \sum_{n\geqslant 0}\binom{2n}{n}\frac{z^n}{n+1}\,.$$

Количество способов расстановки скобок C_n равно $\binom{2n}{n}\frac{1}{n+1}$.

Мы предвосхитили этот результат в главе 5, введя последовательность *чисел Каталана* $\langle 1,1,2,5,14,\ldots\rangle = \langle C_n\rangle$. Она встречается в десятках задач, на первый взгляд никак не связанных между собой [46], потому что во многих ситуациях рекурсивная природа задачи отвечает рекуррентному соотношению со сверткой (7.66).

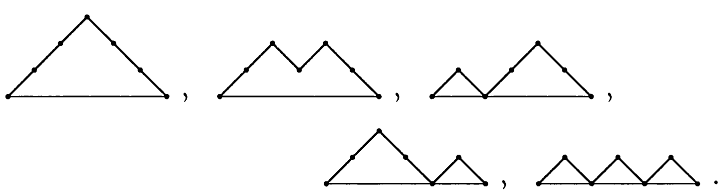
Например, рассмотрим такую задачу: каково количество последовательностей $\langle a_1, a_2, \ldots, a_{2n}\rangle$, состоящих из $+1$ и -1, обладают тем свойством, что

$$a_1 + a_2 + \cdots + a_{2n} \;=\; 0\,,$$

а все их частичные суммы

$$a_1, \quad a_1 + a_2, \quad \ldots, \quad a_1 + a_2 + \cdots + a_{2n}$$

неотрицательны? В последовательности n раз должно встречаться $+1$ и n раз — -1. Графически эту задачу можно представить, нарисовав график последовательности частичных сумм $s_n = \sum_{k=1}^{n} a_k$ как функцию от n. Так, пять возможных решений при $n=3$ есть



Это “горные хребты” шириной $2n$, которые можно изобразить при помощи отрезков двух видов: $\diagup$ и $\diagdown$. Оказывается, имеется ровно C_n способов сделать это, и каждая последовательность может быть связана с некоторой расстановкой скобок следующим образом: добавьте к формуле внешнюю пару скобок, так, чтобы получилось n пар скобок, соответствующих n умножениям. Затем замените каждый знак '$\cdot$' на $+1$ и каждый ')' на -1 и удалите все остальное. Например, по этому правилу формула $x_0\cdot\bigl((x_1\cdot x_2)\cdot(x_3\cdot x_4)\bigr)$ соответствует последовательности $\langle +1,+1,-1,+1,+1,-1,-1,-1\rangle$. Пять способов расстановки скобок в выражении $x_0\cdot x_1\cdot x_2\cdot x_3$ соответствует пяти показанным выше горным хребтам для $n=3$.

Небольшая переформулировка нашей задачи подсчета количества последовательностей позволяет получить удивительно простое комбинаторное решение, не использующее производящих функций. Новая задача заключается в поиске числа последовательностей $\langle a_0, a_1, a_2, \ldots, a_{2n}\rangle$ из $+1$ и -1, обладающих тем свойством, что

$$a_0+a_1+a_2+\cdots+a_{2n} \;=\; 1\,,$$

и при этом все частичные суммы

$$a_0,\quad a_0+a_1,\quad a_0+a_1+a_2,\quad \ldots,\quad a_0+a_1+\cdots+a_{2n}$$

должны быть *положительными*. Очевидно, что это в точности последовательности из предыдущей задачи, к которым спереди приписан элемент $a_0=+1$. Однако последовательности в новой задаче можно просто сосчитать, если воспользоваться замечательным фактом, открытым Джорджем Рени (George Raney) [302] в 1959 году: *если $\langle x_1, x_2, \ldots, x_m\rangle$ — произвольная последовательность целых чисел, сумма которых равна $+1$, то ровно у одного из ее циклических сдвигов*

$$\langle x_1, x_2, \ldots, x_m\rangle,\ \ \langle x_2, \ldots, x_m, x_1\rangle,\ \ \ldots,\ \ \langle x_m, x_1, \ldots, x_{m-1}\rangle$$

все частичные суммы положительны. Например, рассмотрим последовательность $\langle 3,-5,2,-2,3,0\rangle$. Ее циклическими сдвигами являются

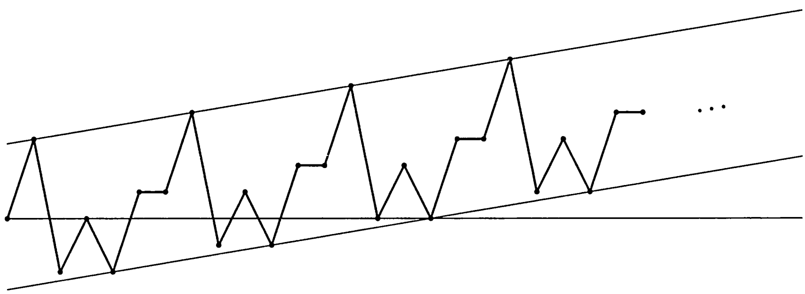
$$\begin{array}{ll}
\langle 3,-5,2,-2,3,0\rangle & \langle -2,3,0,3,-5,2\rangle \\
\langle -5,2,-2,3,0,3\rangle & \langle 3,0,3,-5,2,-2\rangle\ \checkmark \\
\langle 2,-2,3,0,3,-5\rangle & \langle 0,3,-5,2,-2,3\rangle
\end{array}$$

и только отмеченная последовательность имеет лишь положительные частичные суммы.

Лемма Рени может быть доказана при помощи простых геометрических рассуждений. Продолжим нашу последовательность периодически до бесконечной последовательности

$$\langle x_1, x_2, \ldots, x_m, x_1, x_2, \ldots, x_m, x_1, x_2, \ldots \rangle;$$

таким образом, мы полагаем $x_{m+k} = x_k$ для всех $k > 0$. Если теперь изобразить график частичных сумм $s_n = x_1 + \cdots + x_n$ как функции от n, то получившийся график будет иметь "средний наклон" $1/m$, потому что $s_{m+n} = s_n + 1$. Например, график, соответствующий нашей последовательности $\langle 3, -5, 2, -2, 3, 0, 3, -5, 2, \ldots \rangle$, начинается следующим образом:



Эх, если бы цены росли так же медленно... (о снижении не приходится и мечтать).

Весь график может быть заключен между двумя прямыми с наклоном $1/m$, как показано на рисунке, где $m = 6$. В общем случае ограничивающие прямые касаются графика только один раз на каждом периоде из m точек, поскольку прямые с наклоном $1/m$ могут проходить через точки с целыми координатами лишь один раз на m единиц. Единственная нижняя точка касания есть как раз то единственное место в цикле, начиная с которого все частичные суммы будут положительными, поскольку справа от любой другой точки на расстоянии меньше m имеются точки касания.

Вниманию кибернетиков: частичные суммы в этой задаче показывают изменение размера стека как функции от времени при вычислении произведения $n + 1$ сомножителей, поскольку каждая операция внесения операнда в стек увеличивает его размер на $+1$, а каждое умножение — на -1.

С помощью леммы Рени можно легко сосчитать последовательности $\langle a_0, \ldots, a_{2n} \rangle$ из чисел $+1$ и -1, все частичные суммы которых положительны, а общая сумма равна $+1$. Всего имеется $\binom{2n+1}{n}$ последовательностей с n элементами -1 и $n+1$ элементами $+1$, и лемма Рени говорит нам, что ровно у $1/(2n+1)$ части этих последовательностей все частичные суммы положительны. (Выпишите все $N = \binom{2n+1}{n}$ последовательностей и все $2n + 1$ их циклических сдвигов в виде массива $N \times (2n+1)$. Каждая строка содержит ровно одно решение. Каждое решение встречается в каждом столбце ровно один раз. Поэтому во всем массиве содержится $N/(2n+1)$ различных решений, и каждое из них встречается ровно $(2n+1)$ раз.) Общее количество последовательностей с положительными частичными суммами равно

$$\binom{2n+1}{n}\frac{1}{2n+1} = \binom{2n}{n}\frac{1}{n+1} = C_n .$$

Пример 5: рекуррентное соотношение с m-кратной сверткой

Только что рассмотренную задачу можно обобщить, если рассмотреть последовательности $\langle a_0, \ldots, a_{mn}\rangle$ из $+1$ и $(1-m)$, все частичные суммы которых положительны, а общая сумма равна $+1$. Такие последовательности можно назвать m-*последовательностями Рени.* Если в последовательности элемент $(1-m)$ встречается k раз, а $+1$ встречается $mn+1-k$ раз, то

$$k(1-m) + (mn+1-k) = 1$$

и, следовательно, $k = n$. Всего имеется $\binom{mn+1}{n}$ последовательностей с n вхождениями $(1-m)$ и $mn+1-n$ вхождениями $+1$, а лемма Рени говорит нам, что количество таких последовательностей со сплошь положительными частичными суммами составляет ровно

Вниманию кибернетиков: стековая интерпретация применима и в этом случае, но с m-арными, а не бинарными операциями.

$$\binom{mn+1}{n}\frac{1}{mn+1} = \binom{mn}{n}\frac{1}{(m-1)n+1} . \qquad (7.67)$$

Это и есть количество m-последовательностей Рени. Назовем его числом Фусса-Каталана $C_n^{(m)}$, потому что последовательность $\langle C_n^{(m)}\rangle$ была впервые исследована Н. И. Фуссом [135] в 1791 году (за много лет до Каталана). Обычные числа Каталана представляют собой $C_n = C_n^{(2)}$.

Теперь, когда мы знаем ответ (7.67), попробуем сформулировать вопрос к этому ответу. В случае $m = 2$ вопрос был такой: “Какие числа C_n удовлетворяют рекуррентному соотношению $C_n = \sum_k C_k C_{n-1-k} + [n=0]$?” Попробуем найти аналогичный вопрос (аналогичное рекуррентное соотношение) в общем случае.

Тривиальная последовательность $\langle +1\rangle$ длиной 1, очевидно, является m-последовательностью Рени. Если поместить число $(1-m)$ справа от любых m последовательностей, каждая из которых является m-последовательностью Рени, то мы снова получим m-последовательности Рени; частичные суммы остаются положительными — сначала они возрастают до $+2$, $+3$, …, $+m$, а окончательная сумма равна $+1$. И обратно, можно показать, что все m-последовательности Рени $\langle a_0, \ldots, a_{mn}\rangle$ при $n > 0$ получаются этим способом: последний член a_{mn} должен быть равен $(1-m)$. Частичные суммы $s_j = a_0 + \cdots + a_{j-1}$ положительны для $1 \leqslant j \leqslant mn$ и $s_{mn} = m$, потому что $s_{mn} + a_{mn} = 1$. Пусть k_1 означает наибольший индекс $\leqslant mn$, такой, что $s_{k_1} = 1$; пусть k_2 — наибольший индекс, такой, что $s_{k_2} = 2$; и т.д. Таким образом, $s_{k_j} = j$ и $s_k > j$ для $k_j < k \leqslant mn$ и $1 \leqslant j \leqslant m$. Отсюда следует, что $k_m = mn$, и можно легко проверить, что каждая из подпоследовательностей $\langle a_0, \ldots, a_{k_1-1}\rangle$, $\langle a_{k_1}, \ldots, a_{k_2-1}\rangle$, …, $\langle a_{k_{m-1}}, \ldots, a_{k_m-1}\rangle$ является m-последовательностью Рени. При этом должно выполняться $k_1 = mn_1+1$, $k_2-k_1 = mn_2+1$, …, $k_m - k_{m-1} = mn_m + 1$ для некоторых неотрицательных целых чисел n_1, n_2, …, n_m.

Следовательно, $\binom{mn+1}{n}\frac{1}{mn+1}$ является ответом на следующие два интересных вопроса: “Чему равны числа $C_n^{(m)}$, определяемые рекуррентным соотношением

$$C_n^{(m)} = \left(\sum_{n_1+n_2+\quad+n_m=n-1} C_{n_1}^{(m)} C_{n_2}^{(m)} \ldots C_{n_m}^{(m)}\right) + [n=0] \qquad (7.68)$$

при всех целых n?” и “Если $G(z)$ — степенной ряд, удовлетворяющий уравнению

$$G(z) \;=\; zG(z)^m + 1\,, \qquad (7.69)$$

то каковы коэффициенты $[z^n]\,G(z)$?”

Заметим, что это непростые вопросы. В случае обычных чисел Каталана ($m = 2$) мы решили уравнение (7.69) для $G(z)$ и его коэффициентов при помощи формулы для корней квадратного уравнения и биномиальной теоремы; однако для $m = 3$ ни один из стандартных методов не дает ни малейшего намека на то, как решить кубическое уравнение $G = zG^3 + 1$. Таким образом, оказывается проще сначала ответить на вопрос, а затем задать его.

Теперь, однако, мы знаем достаточно, чтобы поставить еще более сложные вопросы и получить на них ответы. Вот пример такого вопроса: “Чему равен коэффициент $[z^n]\,G(z)^l$, если l — положительное целое число, а $G(z)$ — степенной ряд, определяемый соотношением (7.69)?” Рассуждая так, как ранее, получим, что $[z^n]\,G(z)^l$ представляет собой количество последовательностей длиной $mn + l$ со следующими свойствами:

- каждый элемент равен либо $+1$, либо $(1 - m)$;
- все частичные суммы положительны;
- общая сумма равна l.

Действительно, все такие последовательности можно единственным образом получить, приписав друг за другом l последовательностей, обладающих свойствами m-последовательности Рени. Количество способов выполнения этих действий равно

$$\sum_{n_1+n_2+\cdots+n_l=n} C^{(m)}_{n_1} C^{(m)}_{n_2} \ldots C^{(m)}_{n_l} = [z^n]\, G(z)^l .$$

Рени доказал обобщение своей леммы, которое говорит, каким образом подсчитать эти последовательности: *если* $\langle x_1, x_2, \ldots, x_m\rangle$ — *произвольная последовательность целых чисел, такая, что* $x_j \leqslant 1$ *для всех* j, *и* $x_1 + x_2 + \cdots + x_m = l > 0$, *то ровно* l *из циклических сдвигов*

$$\langle x_1, x_2, \ldots, x_m\rangle, \quad \langle x_2, \ldots, x_m, x_1\rangle, \quad \ldots, \quad \langle x_m, x_1, \ldots, x_{m-1}\rangle$$

имеют сплошь положительные частичные суммы.

Например, можно проверить это утверждение на последовательности $\langle -2, 1, -1, 0, 1, 1, -1, 1, 1, 1\rangle$. Вот все ее циклические сдвиги:

$$\begin{array}{ll}
\langle -2, 1, -1, 0, 1, 1, -1, 1, 1, 1\rangle & \langle 1, -1, 1, 1, 1, -2, 1, -1, 0, 1\rangle \\
\langle 1, -1, 0, 1, 1, -1, 1, 1, 1, -2\rangle & \langle -1, 1, 1, 1, -2, 1, -1, 0, 1, 1\rangle \\
\langle -1, 0, 1, 1, -1, 1, 1, 1, -2, 1\rangle & \langle 1, 1, 1, -2, 1, -1, 0, 1, 1, -1\rangle \ \checkmark \\
\langle 0, 1, 1, -1, 1, 1, 1, -2, 1, -1\rangle & \langle 1, 1, -2, 1, -1, 0, 1, 1, -1, 1\rangle \\
\langle 1, 1, -1, 1, 1, 1, -2, 1, -1, 0\rangle \ \checkmark & \langle 1, -2, 1, -1, 0, 1, 1, -1, 1, 1\rangle
\end{array}$$

Только два из них, помеченные '$\checkmark$', обладают тем свойством, что все их частичные суммы положительны. Эта обобщенная лемма доказывается в упр. 13.

Последовательность из элементов $+1$ и $(1-m)$ длиной $mn+l$ с общей суммой l должна включать ровно n элементов $(1-m)$. Обобщенная лемма говорит нам, что $l/(mn+l)$ из этих $\binom{mn+l}{n}$ последовательностей имеют только положительные частичные суммы; следовательно, наш трудный вопрос имеет удивительно простой ответ:

$$[z^n]\, G(z)^l = \binom{mn+l}{n} \frac{l}{mn+l} \tag{7.70}$$

для всех целых $l > 0$.

Читатели, не забывшие главу 5, вполне могут подумать что-нибудь вроде: "У меня *déjà vu* или эта формула встречалась ранее?" Да, она действительно уже встречалась; уравнение Ламберта (5.60) гласит:

$$[z^n]\, \mathcal{B}_t(z)^r = \binom{tn+r}{n} \frac{r}{tn+r} .$$

Следовательно, производящая функция $G(z)$ из (7.69) должна совпадать с обобщенным биномиальным рядом $\mathcal{B}_m(z)$. И действительно, из уравнения (5.59) имеем

$$\mathcal{B}_m(z)^{1-m} - \mathcal{B}_m(z)^{-m} = z,$$

что эквивалентно

$$\mathcal{B}_m(z) - 1 = z\mathcal{B}_m(z)^m.$$

Теперь, когда мы знаем, что имеем дело с обобщенным биномиальным рядом, давайте переключимся на обозначения из главы 5. В ней без доказательства были сформулированы некоторые тождества. Частично восполним этот пробел, доказав, что степенной ряд $\mathcal{B}_t(z)$, определяемый как

$$\mathcal{B}_t(z) = \sum_n \binom{tn+1}{n} \frac{z^n}{tn+1},$$

обладает тем замечательным свойством, что

$$\mathcal{B}_t(z)^r = \sum_n \binom{tn+r}{n} \frac{rz^n}{tn+r},$$

если только t и r — положительные целые числа.

Можно ли распространить эти результаты на *произвольные* значения t и r? Да, можно, потому что коэффициенты $\binom{tn+r}{n}\frac{r}{tn+r}$ представляют собой полиномы от t и r. Общая r-я степень, определяемая как

$$\begin{aligned}
\mathcal{B}_t(z)^r &= e^{r \ln \mathcal{B}_t(z)} = \sum_{n \geqslant 0} \frac{\bigl(r \ln \mathcal{B}_t(z)\bigr)^n}{n!} = \\
&= \sum_{n \geqslant 0} \frac{r^n}{n!} \left(- \sum_{m \geqslant 1} \frac{\bigl(1 - \mathcal{B}_t(z)\bigr)^m}{m} \right)^n,
\end{aligned}$$

имеет в качестве коэффициентов некоторые полиномы от t и r; эти полиномы совпадают с $\binom{tn+r}{n}\frac{r}{tn+r}$ для бесконечно многих значений t и r. Таким образом, эти две последовательности полиномов должны тождественно совпадать.

В главе 5 упоминается также обобщенный экспоненциальный ряд

$$\mathcal{E}_t(z) = \sum_{n \geqslant 0} \frac{(tn+1)^{n-1}}{n!} z^n,$$

который, как утверждается в (5.60), обладает не менее замечательным свойством:

$$[z^n]\,\mathcal{E}_t(z)^r \;=\; \frac{r(tn+r)^{n-1}}{n!}\,. \tag{7.71}$$

Можно доказать это как предельный случай формулы для $\mathcal{B}_t(z)$, потому что, как нетрудно показать,

$$\mathcal{E}_t(z)^r \;=\; \lim_{x\to\infty} \mathcal{B}_{xt}(z/x)^{xr}\,.$$

7.6 Экспоненциальные производящие функции

Иногда производящая функция последовательности $\langle g_n\rangle$ очень сложна, в то время как связанная с ней последовательность $\langle g_n/n!\rangle$ имеет достаточно простую производящую функцию. В таких случаях мы, естественно, предпочтем работать с $\langle g_n/n!\rangle$ и в самом конце выполнить умножение на $n!$. Этот трюк срабатывает столь часто, что степенной ряд

$$\widehat{G}(z) \;=\; \sum_{n\geqslant 0} g_n \frac{z^n}{n!} \tag{7.72}$$

получил собственное название — *экспоненциальная производящая функция* последовательности $\langle g_0, g_1, g_2, \ldots\rangle$. Такое имя возникло из-за экспоненциальной функции e^z, которая представляет собой экспоненциальную производящую функцию последовательности $\langle 1, 1, 1, \ldots\rangle$.

Многие производящие функции в табл. 425 на самом деле являются экспоненциальными производящими функциями. Например, уравнение (7.50) гласит, что $\bigl(\ln\frac{1}{1-z}\bigr)^m/m!$ представляет собой экспоненциальную производящую функцию последовательности $\langle \left[{0 \atop m}\right], \left[{1 \atop m}\right], \left[{2 \atop m}\right], \ldots\rangle$. Обычная производящая функция для этой последовательности существенно сложнее (да еще и расходится).

У экспоненциальных производящих функций имеются собственные приемы работы с ними, аналогичные рассматривавшимся в разделе 7.2 операциям. Например, если умножить экспоненциальную производящую функцию $\langle g_n\rangle$ на z, можно получить

$$\sum_{n\geqslant 0} g_n \frac{z^{n+1}}{n!} \;=\; \sum_{n\geqslant 1} g_{n-1} \frac{z^n}{(n-1)!} \;=\; \sum_{n\geqslant 0} n g_{n-1} \frac{z^n}{n!}\,;$$

а это не что иное, как экспоненциальная производящая функция последовательности $\langle 0, g_0, 2g_1, \ldots \rangle = \langle n g_{n-1} \rangle$.

Здорово, правда?

Дифференцирование экспоненциальной производящей функции последовательности $\langle g_0, g_1, g_2, \ldots \rangle$ по z дает

$$\sum_{n\geqslant 0} n g_n \frac{z^{n-1}}{n!} = \sum_{n\geqslant 1} g_n \frac{z^{n-1}}{(n-1)!} = \sum_{n\geqslant 0} g_{n+1} \frac{z^n}{n!}; \tag{7.73}$$

а это — экспоненциальная производящая функция $\langle g_1, g_2, \ldots \rangle$. Таким образом, дифференцирование экспоненциальной производящей функции соответствует операции сдвига влево $\bigl(G(z) - g_0\bigr)/z$ для обычных производящих функций. (Это свойство экспоненциальных производящих функций мы использовали при изучении гипергеометрических рядов (5.106).) Интегрирование экспоненциальной производящей функции дает

$$\int_0^z \sum_{n\geqslant 0} g_n \frac{t^n}{n!}\,dt = \sum_{n\geqslant 0} g_n \frac{z^{n+1}}{(n+1)!} = \sum_{n\geqslant 1} g_{n-1} \frac{z^n}{n!}; \tag{7.74}$$

т.е. сдвиг вправо, экспоненциальную производящую функцию последовательности $\langle 0, g_0, g_1, \ldots \rangle$.

Наиболее интересной операцией над экспоненциальными производящими функциями, как и в случае обычных производящих функций, является умножение. Если $\widehat{F}(z)$ и $\widehat{G}(z)$ представляют собой экспоненциальные производящие функции последовательностей $\langle f_n \rangle$ и $\langle g_n \rangle$, то $\widehat{F}(z)\widehat{G}(z) = \widehat{H}(z)$ — экспоненциальная производящая функция последовательности $\langle h_n \rangle$, которая называется *биномиальной сверткой* $\langle f_n \rangle$ и $\langle g_n \rangle$:

$$h_n = \sum_k \binom{n}{k} f_k\, g_{n-k}\,. \tag{7.75}$$

Биномиальные коэффициенты появляются здесь из-за того, что $\binom{n}{k} = n!/k!\,(n-k)!$, следовательно,

$$\frac{h_n}{n!} = \sum_{k=0}^{n} \frac{f_k}{k!} \frac{g_{n-k}}{(n-k)!};$$

другими словами, $\langle h_n/n! \rangle$ — это обычная свертка последовательностей $\langle f_n/n! \rangle$ и $\langle g_n/n! \rangle$.

Биномиальные свертки часто встречаются в приложениях. Так, в (6.79) мы определили числа Бернулли при помощи неявного рекуррентного соотношения

$$\sum_{j=0}^{m} \binom{m+1}{j} B_j = [m=0] \qquad \text{при всех } m \geqslant 0;$$

это соотношение можно переписать в виде биномиальной свертки, если заменить $m+1$ на n и добавить к обеим сторонам член B_n:

$$\sum_k \binom{n}{k} B_k = B_n + [n=1] \qquad \text{при всех } n \geqslant 0. \tag{7.76}$$

Теперь можно перевести это соотношение на язык степенных рядов (как было обещано в главе 6), если ввести в рассмотрение экспоненциальную производящую функцию для чисел Бернулли $\widehat{B}(z) = \sum_{n\geqslant 0} B_n z^n/n!$. Левая часть (7.76) представляет собой биномиальную свертку $\langle B_n \rangle$ с константной последовательностью $\langle 1,1,1,\ldots \rangle$; следовательно, экспоненциальная производящая функция левой части равна $\widehat{B}(z)e^z$. Экспоненциальная производящая функция правой части имеет вид

$$\sum_{n\geqslant 0} \bigl(B_n + [n=1]\bigr) z^n/n! = \widehat{B}(z) + z.$$

Таким образом, должно выполняться равенство $\widehat{B}(z) = z/(e^z - 1)$; мы доказали равенство (6.81), которое встречается также в табл. 425 как формула (7.44).

А теперь вернемся к сумме, неоднократно появлявшейся в нашей книге:

$$S_m(n) = 0^m + 1^m + 2^m + \cdots + (n-1)^m = \sum_{0\leqslant k<n} k^m .$$

В этот раз мы попытаемся применить к ней производящие функции: вдруг это упростит дело. Будем считать n фиксированным, а m — переменным. Таким образом, наша задача заключается в том, чтобы понять, как ведут себя коэффициенты степенного ряда

$$S(z) = S_0(n) + S_1(n)\,z + S_2(n)\,z^2 + \cdots = \sum_{m\geqslant 0} S_m(n)\,z^m .$$

Мы знаем, что производящая функция последовательности $\langle 1, k, k^2, \ldots \rangle$ равна

$$\frac{1}{1-kz} = \sum_{m\geqslant 0} k^m z^m ,$$

следовательно,

$$S(z) = \sum_{m\geqslant 0} \sum_{0\leqslant k<n} k^m z^m = \sum_{0\leqslant k<n} \frac{1}{1-kz},$$

если поменять местами порядок суммирования. Эту сумму можно записать в аналитическом виде:

$$\begin{aligned} S(z) &= \frac{1}{z}\left(\frac{1}{z^{-1}-0}+\frac{1}{z^{-1}-1}+\cdots+\frac{1}{z^{-1}-n+1}\right) = \\ &= \frac{1}{z}\left(H_{z^{-1}}-H_{z^{-1}-n}\right); \end{aligned} \qquad (7.77)$$

но мы ничего не знаем о разложении такой суммы по степеням z.

Нам на помощь приходят экспоненциальные производящие функции. Для последовательности $\langle S_0(n), S_1(n), S_2(n), \ldots\rangle$ экспоненциальная производящая функция равна

$$\widehat{S}(z,n) = S_0(n)+S_1(n)\frac{z}{1!}+S_2(n)\frac{z^2}{2!}+\cdots = \sum_{m\geqslant 0} S_m(n)\frac{z^m}{m!}.$$

Для получения коэффициентов $S_m(n)$ можно воспользоваться экспоненциальной производящей функцией последовательности $\langle 1, k, k^2, \ldots\rangle$, а именно

$$e^{kz} = \sum_{m\geqslant 0} k^m \frac{z^m}{m!},$$

что дает

$$\widehat{S}(z,n) = \sum_{m\geqslant 0}\sum_{0\leqslant k<n} k^m\frac{z^m}{m!} = \sum_{0\leqslant k<n} e^{kz}.$$

Последняя же сумма является суммой геометрической прогрессии, которая в аналитическом виде записывается как

$$\widehat{S}(z,n) = \frac{e^{nz}-1}{e^z-1}. \qquad (7.78)$$

Эврика! Все, что нам надо сделать,— выписать коэффициенты этой сравнительно простой функции, и мы получим выражение для $S_m(n)$, потому что $S_m(n) = m!\,[z^m]\widehat{S}(z,n)$.

Здесь в игру вступают числа Бернулли. Совсем недавно мы заметили, что экспоненциальной производящей функцией для чисел Бернулли является

$$\widehat{B}(z) = \sum_{k\geqslant 0} B_k\frac{z^k}{k!} = \frac{z}{e^z-1};$$

следовательно, можно записать

$$\begin{aligned} \widehat{S}(z,n) &= \widehat{B}(z)\,\frac{e^{nz}-1}{z} = \\ &= \left(B_0\frac{z^0}{0!}+B_1\frac{z^1}{1!}+B_2\frac{z^2}{2!}+\cdots\right)\left(n\frac{z^0}{1!}+n^2\frac{z^1}{2!}+n^3\frac{z^2}{3!}+\cdots\right). \end{aligned}$$

Сумма $S_m(n)$ равна коэффициенту при z^m в этом произведении, умноженному на $m!$, например

$$\begin{aligned}
S_0(n) &= 0!\left(B_0\frac{n}{1!\,0!}\right) &&= n;\\
S_1(n) &= 1!\left(B_0\frac{n^2}{2!\,0!}+B_1\frac{n}{1!\,1!}\right) &&= \tfrac12 n^2-\tfrac12 n;\\
S_2(n) &= 2!\left(B_0\frac{n^3}{3!\,0!}+B_1\frac{n^2}{2!\,1!}+B_2\frac{n}{1!\,2!}\right) &&= \tfrac13 n^3-\tfrac12 n^2+\tfrac16 n.
\end{aligned}$$

Таким образом, в очередной раз оказалась выведена хорошо знакомая формула $\Box_n = S_2(n) = \frac13 n(n-\frac12)(n-1)$, и это самый простой вывод из всех: при помощи нескольких строк найдено поведение $S_m(n)$ при любых m.

В общем случае формула имеет вид

$$S_{m-1}(n) = \frac{1}{m}\bigl(B_m(n)-B_m(0)\bigr), \tag{7.79}$$

где $B_m(x)$ — *полином Бернулли*, определяемый как

$$B_m(x) = \sum_k \binom{m}{k} B_k\, x^{m-k}. \tag{7.80}$$

Вот почему это так: полином Бернулли является биномиальной сверткой последовательности $\langle B_0, B_1, B_2, \ldots\rangle$ с $\langle 1, x, x^2, \ldots\rangle$; следовательно, экспоненциальная производящая функция для $\langle B_0(x), B_1(x), B_2(x), \ldots\rangle$ представляет собой произведение экспоненциальных производящих функций этих последовательностей,

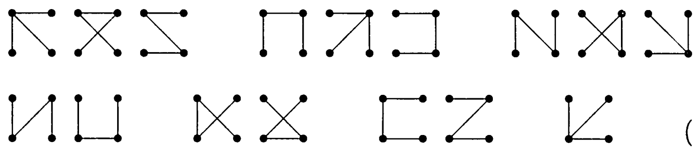
$$\widehat{B}(z,x) = \sum_{m\geqslant 0} B_m(x)\frac{z^m}{m!} = \frac{z}{e^z-1}\sum_{m\geqslant 0} x^m\frac{z^m}{m!} = \frac{ze^{xz}}{e^z-1}. \tag{7.81}$$

Отсюда вытекает равенство (7.79), потому что экспоненциальная производящая функция для $\langle 0, S_0(n), 2S_1(n), \ldots\rangle$ согласно (7.78) представляет собой

$$z\,\frac{e^{nz}-1}{e^z-1} = \widehat{B}(z,n)-\widehat{B}(z,0).$$

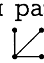
Обратимся еще к одной задаче, в которой экспоненциальная производящая функция — как раз то, что нужно: каково количество остовных деревьев в *полном графе* с n вершинами $\{1, 2, \ldots, n\}$? Обозначим это число как t_n. Полный граф содержит $\frac12 n(n-1)$ ребер, по одному ребру между каждыми двумя вершинами; так что мы, по существу, ищем полное число способов соединить n данных объектов при помощи $n-1$ линий между ними.

Имеем $t_1 = t_2 = 1$. Кроме того, $t_3 = 3$, потому что полный граф с тремя вершинами представляет собой фан порядка 2, а мы знаем, что $f_2 = 3$. В случае $n = 4$ имеется шестнадцать остовных деревьев:



(7.82)

Следовательно, $t_4 = 16$.

Наш опыт решения аналогичной задачи для фанов подсказывает, что наилучший способ справиться с этой задачей состоит в том, чтобы выделить одну вершину и посмотреть на те связные компоненты или блоки, на которые разобьется остовное дерево, если проигнорировать все ребра, проходящие через выделенную вершину. Если невыделенные вершины образуют m компонентов размером $k_1, k_2, \ldots, k_m$, то можно соединить их с выделенной вершиной $k_1 k_2 \ldots k_m$ способами. Например, в случае $n = 4$ можно считать выделенной нижнюю левую вершину. В верхнем ряду (7.82) показано $3t_3$ случаев, когда три остальные вершины соединены между собой одним из t_3 способов и затем соединены с левой нижней вершиной одним из трех способов. В нижнем ряду изображено $2 \cdot 1 \times t_2 t_1 \times \binom{3}{2}$ решений, в которых остальные три вершины разбиты на компоненты размером 2 и 1 одним из $\binom{3}{2}$ способов; там же показан случай , когда три остальные вершины полностью изолированы.

Эти рассуждения приводят к рекуррентному соотношению

$$t_n = \sum_{m>0} \frac{1}{m!} \sum_{k_1+\cdots+k_m=n-1} \binom{n-1}{k_1, k_2, \ldots, k_m} \times \\ \times k_1 k_2 \ldots k_m \, t_{k_1} t_{k_2} \ldots t_{k_m}$$

для всех $n > 1$. Вот почему: имеется $\binom{n-1}{k_1, k_2, \ldots, k_m}$ способов представить $n-1$ элементов в виде последовательности из m компонентов размером соответственно $k_1, k_2, \ldots, k_m$; имеется $t_{k_1} t_{k_2} \ldots t_{k_m}$ способов соединить вершины внутри этих компонентов какими-то остовными деревьями; есть $k_1 k_2 \ldots k_m$ способов соединить вершину n с этими компонентами; и мы должны выполнить деление на $m!$, потому что порядок компонентов не учитывается. Например, для $n = 4$ это рекуррентное соотношение дает

$$\begin{aligned} t_4 &= 3t_3 + \frac{1}{2}\left(\binom{3}{1,2}2t_1t_2 + \binom{3}{2,1}2t_2t_1\right) + \frac{1}{6}\left(\binom{3}{1,1,1}t_1^3\right) = \\ &= 3t_3 + 6t_2t_1 + t_1^3\,. \end{aligned}$$

Рекуррентное соотношение для t_n, на первый взгляд, выглядит устрашающе, но на самом деле оно не так уж плохо, просто слегка свернуто. Обозначим

$$u_n = n\,t_n,$$

и тогда все существенно упростится:

$$\frac{u_n}{n!} = \sum_{m>0} \frac{1}{m!} \sum_{k_1+k_2+\cdots+k_m=n-1} \frac{u_{k_1}}{k_1!}\frac{u_{k_2}}{k_2!}\cdots\frac{u_{k_m}}{k_m!}\,, \quad \text{если } n > 1. \tag{7.83}$$

Внутренняя сумма представляет собой коэффициент при z^{n-1} в экспоненциальной производящей функции $\widehat{U}(z)$, возведенной в m-ю степень; формула будет правильной и для $n = 1$, если добавить слагаемое $\widehat{U}(z)^0$, отвечающее случаю $m = 0$. Итак,

$$\frac{u_n}{n!} = [z^{n-1}] \sum_{m\geqslant 0} \frac{1}{m!}\,\widehat{U}(z)^m = [z^{n-1}]\, e^{\widehat{U}(z)} = [z^n]\, z e^{\widehat{U}(z)}$$

при любом $n > 0$, и мы получаем уравнение

$$\widehat{U}(z) = z\,e^{\widehat{U}(z)}\,. \tag{7.84}$$

Прогресс! Уравнение (7.84) выглядит почти так же, как и уравнение

$$\mathcal{E}(z) = e^{z\,\mathcal{E}(z)}\,,$$

которое определяет обобщенный экспоненциальный ряд $\mathcal{E}(z) = \mathcal{E}_1(z)$ в (5.59) и (7.71); и действительно, между этими функциями имеется связь:

$$\widehat{U}(z) = z\,\mathcal{E}(z)\,.$$

Таким образом, можно просто прочитать ответ к нашей задаче:

$$\begin{aligned} t_n &= \frac{u_n}{n} = \frac{n!}{n}\,[z^n]\,\widehat{U}(z) = \\ &= (n-1)!\,[z^{n-1}]\,\mathcal{E}(z) = n^{n-2}\,. \end{aligned} \tag{7.85}$$

Полный граф с вершинами $\{1, 2, \ldots, n\}$ $(n > 0)$ имеет ровно n^{n-2} остовных деревьев.

7.7 Производящие функции Дирихле

Помимо рассмотренных, возможны и другие способы порождения последовательностей по рядам; здесь можно использовать, по крайней мере в принципе, любую систему "ядер" $K_n(z)$, такую, что

$$\sum_n g_n K_n(z) = 0 \quad \Longrightarrow \quad g_n = 0 \text{ для всех } n.$$

Для обычных производящих функций мы использовали ядра $K_n(z) = z^n$, а для экспоненциальных производящих функций — $K_n(z) = z^n/n!$; можно также попробовать использовать убывающие факториальные степени $z^{\underline{n}}$ или биномиальные коэффициенты $z^{\underline{n}}/n! = \binom{z}{n}$.

Наиболее важная альтернатива производящим функциям и экспоненциальная производящая функциям получается, если использовать ядра $1/n^z$; они рассчитаны на последовательности $\langle g_1, g_2, \ldots \rangle$, начинающиеся с $n = 1$, а не с $n = 0$:

$$\widetilde{G}(z) = \sum_{n \geqslant 1} \frac{g_n}{n^z}. \tag{7.86}$$

Эта функция называется *производящей функцией Дирихле*, поскольку огромный вклад в ее изучение был сделан немецким математиком Густавом Дирихле (Gustav Dirichlet) (1805–1859).

Например, производящая функция Дирихле для константной последовательности $\langle 1, 1, 1, \ldots \rangle$ равна

$$\sum_{n \geqslant 1} \frac{1}{n^z} = \zeta(z). \tag{7.87}$$

Это — *дзета-функция* Римана, которую мы называли также обобщенным гармоническим числом $H_\infty^{(z)}$ для $z > 1$.

Произведение производящих функций Дирихле соответствует специальному виду свертки:

$$\widetilde{F}(z)\widetilde{G}(z) = \sum_{l,m \geqslant 1} \frac{f_l}{l^z} \frac{g_m}{m^z} = \sum_{n \geqslant 1} \frac{1}{n^z} \sum_{l,m \geqslant 1} f_l\, g_m\, [l \cdot m = n].$$

Таким образом, $\widetilde{F}(z)\widetilde{G}(z) = \widetilde{H}(z)$ есть производящая функция Дирихле последовательности

$$h_n = \sum_{d \backslash n} f_d\, g_{n/d}. \tag{7.88}$$

Например, из (4.55) мы знаем, что $\sum_{d\backslash n} \mu(d) = [n=1]$; это свертка Дирихле последовательности Мёбиуса $\langle \mu(1), \mu(2), \mu(3), \ldots \rangle$ с последовательностью $\langle 1, 1, 1, \ldots \rangle$, значит,

$$\widetilde{M}(z)\zeta(z) = \sum_{n\geqslant 1} \frac{[n=1]}{n^z} = 1. \tag{7.89}$$

Другими словами, производящая функция Дирихле последовательности $\langle \mu(1), \mu(2), \mu(3), \ldots \rangle$ равна $\zeta(z)^{-1}$.

Производящие функции Дирихле особенно полезны, когда последовательность $\langle g_1, g_2, \ldots \rangle$ является *мультипликативной функцией*, а именно — когда

$$g_{mn} = g_m g_n \qquad \text{при } m \perp n.$$

В таких случаях значения g_n для всех n определяются значениями g_n для n, являющихся степенями простых чисел, и производящую функцию Дирихле можно разложить в произведение по простым числам:

$$\widetilde{G}(z) = \prod_p \left(1 + \frac{g_p}{p^z} + \frac{g_{p^2}}{p^{2z}} + \frac{g_{p^3}}{p^{3z}} + \cdots\right). \tag{7.90}$$

Если, например, положить $g_n = 1$ для всех n, то получится представление дзета-функции Римана в виде произведения:

$$\zeta(z) = \prod_p \left(\frac{1}{1-p^{-z}}\right). \tag{7.91}$$

Для функции Мёбиуса $\mu(p) = -1$ и $\mu(p^k) = 0$ при $k > 1$, следовательно, ее производящая функция Дирихле равна

$$\widetilde{M}(z) = \prod_p (1-p^{-z}); \tag{7.92}$$

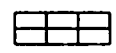
что, само собой, согласуется с (7.89) и (7.91). Для функции Эйлера φ имеем $\varphi(p^k) = p^k - p^{k-1}$, следовательно, ее производящая функция Дирихле в виде произведения записывается как

$$\widetilde{\Phi}(z) = \prod_p \left(1 + \frac{p-1}{p^z - p}\right) = \prod_p \frac{1-p^{-z}}{1-p^{1-z}}. \tag{7.93}$$

Отсюда можно сделать вывод, что $\widetilde{\Phi}(z) = \zeta(z-1)/\zeta(z)$.

Упражнения

Разминка

1 Некий эксцентричный коллекционер покрытий $(2 \times n)$-прямоугольников костями домино платит 4 доллара за каждую вертикальную кость домино и 1 доллар за горизонтальную. Каково количество покрытий, оцененных таким образом в m долларов? Например, для $m = 6$ имеется три покрытия: ,  и .

2 Запишите производящую функцию и экспоненциальную производящую функцию для последовательности $\langle 2, 5, 13, 35, \ldots \rangle = \langle 2^n + 3^n \rangle$ в аналитическом виде.

3 Чему равна сумма $\sum_{n \geqslant 0} H_n / 10^n$?

4 Общая теорема о разложении рациональной функции $P(z)/Q(z)$ не вполне общая, поскольку имеется ограничение, что степень P должна быть меньше степени Q. Что произойдет, если степень P окажется больше?

5 Найдите производящую функцию $S(z)$, такую, что

$$[z^n]\, S(z) \;=\; \sum_k \binom{r}{k} \binom{r}{n-2k}\,.$$

Обязательные упражнения

6 Покажите, что рекуррентное соотношение (7.32) можно решить методом наборов, не прибегая к производящим функциям.

7 Решите рекуррентное соотношение

$$\begin{aligned} g_0 &= 1\,; \\ g_n &= g_{n-1} + 2g_{n-2} + \cdots + ng_0 \qquad \text{при } n > 0. \end{aligned}$$

8 Вычислите $[z^n]\, \big(\ln(1-z)\big)^2/(1-z)^{m+1}$?

9 Воспользуйтесь результатом предыдущего упражнения для вычисления $\sum_{k=0}^{n} H_k H_{n-k}$.

10 Положите в тождестве (7.62) $r = s = -1/2$ и затем избавьтесь от всех вхождений $1/2$ с помощью трюков наподобие (5.36). Какое удивительное тождество вы вывели?

Заблудившееся, должно быть?

11 Эта задача, состоящая из трех независимых частей, позволяет попрактиковаться в работе с производящими функциями. Полагаем, что $A(z) = \sum_n a_n z^n$, $B(z) = \sum_n b_n z^n$, $C(z) =$

$\sum_n c_n z^n$ и что при отрицательных n коэффициенты равны 0.

а Выразите C через A и B, если $c_n = \sum_{j+2k\leqslant n} a_j b_k$.

б Выразите A через B, если $nb_n = \sum_{k=0}^{n} 2^k a_k/(n-k)!$.

в Выразите A через B, если $a_n = \sum_{k=0}^{n} \binom{r+k}{k} b_{n-k}$, где r — действительное число; затем используйте вашу формулу для поиска коэффициентов $f_k(r)$, таких, что $b_n = \sum_{k=0}^{n} f_k(r) a_{n-k}$.

12 Сколько существует способов разместить числа $\{1, 2, \ldots, 2n\}$ в массиве размером $2 \times n$ так, чтобы и строки, и столбцы массива были упорядочены по возрастанию слева направо и сверху вниз? Например, одним из решений при $n = 5$ является

$$\begin{pmatrix} 1 & 2 & 4 & 5 & 8 \\ 3 & 6 & 7 & 9 & 10 \end{pmatrix}.$$

13 Докажите обобщенную лемму Рени, сформулированную непосредственно перед (7.70).

14 Решите рекуррентное соотношение

$$\begin{aligned} g_0 &= 0, \qquad g_1 = 1, \\ g_n &= -2ng_{n-1} + \sum_k \binom{n}{k} g_k g_{n-k} \qquad \text{при } n > 1, \end{aligned}$$

воспользовавшись экспоненциальной производящей функцией.

15 *Число Белла* ϖ_n представляет собой количество способов разбиения множества из n предметов на подмножества. Например, $\varpi_3 = 5$, потому что множество $\{1, 2, 3\}$ можно разбить на следующие подмножества:

$$\{1,2,3\}; \quad \{1,2\}\cup\{3\}; \quad \{1,3\}\cup\{2\}; \quad \{1\}\cup\{2,3\}; \quad \{1\}\cup\{2\}\cup\{3\}.$$

Докажите, что $\varpi_{n+1} = \sum_k \binom{n}{k} \varpi_{n-k}$, и используйте это рекуррентное соотношение для того, чтобы записать экспоненциальную производящую функцию $P(z) = \sum_n \varpi_n z^n/n!$ в аналитическом виде.

16 Последовательности $\langle a_n \rangle$ и $\langle b_n \rangle$ связаны между собой формулой свертки

$$b_n = \sum_{k_1+2k_2+\cdots+nk_n=n} \binom{a_1+k_1-1}{k_1} \binom{a_2+k_2-1}{k_2} \ldots \binom{a_n+k_n-1}{k_n};$$

кроме того, $a_0 = 0$ и $b_0 = 1$. Докажите, что соответствующие производящие функции удовлетворяют уравнению $\ln B(z) = A(z) + \frac{1}{2}A(z^2) + \frac{1}{3}A(z^3) + \cdots$.

17 Покажите, что экспоненциальная производящая функция $\widehat{G}(z)$ связана с обычной производящей функцией $G(z)$ той же последовательности соотношением

$$\int_0^\infty \widehat{G}(zt)e^{-t}\,dt = G(z),$$

если этот интеграл существует.

18 Найдите производящие функции Дирихле для последовательностей

а $g_n = \sqrt{n}$;

б $g_n = \ln n$;

в $g_n = [n$ свободно от квадратов$]$.

Выразите свои ответы через дзета-функции. (Свойство свободы от квадратов определено в упр. 4.13.)

19 Любой степенной ряд $F(z) = \sum_{n\geqslant 0} f_n z^n$ с $f_0 = 1$ определяет последовательность полиномов $f_n(x)$ по правилу

$$F(z)^x = \sum_{n\geqslant 0} f_n(x) z^n,$$

где $f_n(1) = f_n$ и $f_n(0) = [n=0]$. В общем случае $f_n(x)$ имеет степень n. Покажите, что такие полиномы удовлетворяют формулам свертки

Что значит "в общем случае"? Если $f_1 = f_2 = \cdots = f_{m-1} = 0$, степень $f_n(x)$ не превосходит $\lfloor n/m \rfloor$.

$$\sum_{k=0}^{n} f_k(x) f_{n-k}(y) = f_n(x+y);$$

$$(x+y)\sum_{k=0}^{n} k f_k(x) f_{n-k}(y) = x n f_n(x+y).$$

(Тождества в табл. 255 и 334 являются частными случаями этого приема.)

20 Степенной ряд $G(z)$ называется *дифференцируемо конечным*, если найдется конечное число полиномов $P_0(z)$, …, $P_m(z)$, среди которых не все равны нулю, и такие, что

$$P_0(z)G(z) + P_1(z)G'(z) + \cdots + P_m(z)G^{(m)}(z) = 0.$$

Последовательность чисел $\langle g_0, g_1, g_2, \ldots \rangle$ называется *полиномиально рекурсивной*, если найдутся полиномы $p_0(z)$,

…, $p_m(z)$ (в конечном количестве), среди которых не все равны нулю, и такие, что

$$p_0(n)g_n + p_1(n)g_{n+1} + \cdots + p_m(n)g_{n+m} = 0$$

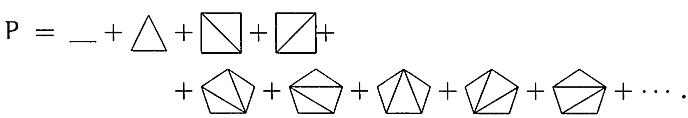
для всех целых $n \geqslant 0$. Докажите, что производящая функция дифференцируемо конечна тогда и только тогда, когда ее последовательность коэффициентов полиномиально рекурсивна.

Домашние задания

21 Грабитель врывается в банк и требует 500 долларов десяти- и двадцатидолларовыми банкнотами. Он также желает знать, сколькими способами кассир может дать ему эти деньги. Найдите производящую функцию $G(z)$, у которой это количество представляет собой коэффициент $[z^{500}]\,G(z)$, а также более компактную производящую функцию $\check{G}(z)$, у которой это число является коэффициентом $[z^{50}]\,\check{G}(z)$. Используйте для решения (а) элементарные дроби; (б) метод, аналогичный (7.39).

Он что, хочет заняться покрытием прямоугольника размером $2 \times n$?

22 Пусть P — сумма всех возможных способов “триангуляции” многоугольников:



(Первый член представляет вырожденный многоугольник с двумя вершинами; все остальные слагаемые представляют собой многоугольники, разбитые на треугольники. Например, пятиугольник можно триангулировать пятью способами.) Определите операцию “умножения” $A\triangle B$ триангулированных многоугольников A и B так, чтобы было справедливо уравнение

$$P = \_\_ + P\triangle P.$$

Затем замените каждый треугольник на ‘z’. Что после этого можно сказать о количестве разбиений n-угольника на треугольники?

23 Сколькими способами можно построить колонну размером $2 \times 2 \times n$ из кирпичей размером $2 \times 1 \times 1$?

Где n определяется ценами на кирпичи и цемент и содержимым вашего кошелька.

24 Сколько остовных деревьев имеется в n-колесе (графе с циклом из n “внешних” вершин, каждая из которых соединена с $(n+1)$-й вершиной “оси”) при $n \geqslant 3$?

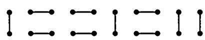
25 Пусть $m \geqslant 2$ — целое число. Выразите в аналитическом виде, как функцию от z и m, производящую функцию последовательности $\langle n \bmod m \rangle$. Используя эту производящую функцию, запишите выражение '$n \bmod m$' в терминах комплексного числа $\omega = e^{2\pi i/m}$. (Например, для $m = 2$ будем иметь $\omega = -1$ и $n \bmod 2 = \frac{1}{2} - \frac{1}{2}(-1)^n$.)

26 Числа Фибоначчи второго порядка $\langle \mathfrak{F}_n \rangle$ определяются рекуррентным соотношением

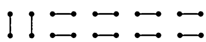
$$\begin{aligned} \mathfrak{F}_0 &= 0; \qquad \mathfrak{F}_1 = 1; \\ \mathfrak{F}_n &= \mathfrak{F}_{n-1} + \mathfrak{F}_{n-2} + F_n \qquad \text{при } n > 1. \end{aligned}$$

Выразите $\mathfrak{F}_n$ через обычные числа Фибоначчи F_n и F_{n+1}.

27 Каждое покрытие прямоугольника размером $2 \times n$ костями домино можно рассматривать как некоторый способ нарисовать n несоприкасающихся отрезков в массиве точек размером $2 \times n$:



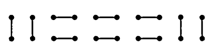
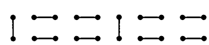
Если наложить друг на друга две подобные схемы, можно получить несколько циклов, поскольку с каждой точкой связаны два отрезка. Если, например, приведенная выше картинка объединяется с

 ,

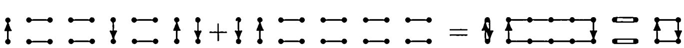
то в результате получается

 .

Такой же набор циклов получается и при объединении

 с  .

Однако можно однозначно восстановить исходные схемы по результату их объединения, если приписать каждому вертикальному отрезку некоторую ориентацию при помощи стрелок. Стрелки расставим поочередно вверх/вниз/вверх/вниз/$\cdots$ в первой схеме и вниз/вверх/вниз/вверх/$\cdots$ во второй, например

 .

Количество таких ориентированных циклических схем должно быть, следовательно, равно $T_n^2 = F_{n+1}^2$, и мы должны быть способны доказать это при помощи алгебры. Пусть Q_n

представляет собой количество ориентированных циклических схем размером $2 \times n$. Найдите рекуррентное соотношение для Q_n, решите его при помощи производящих функций и получите алгебраическое доказательство того, что $Q_n = F_{n+1}^2$.

28 Коэффициенты $A(z)$ в (7.39) удовлетворяют соотношению $A_r + A_{r+10} + A_{r+20} + A_{r+30} = 100$ для $0 \leqslant r < 10$. Найдите "простое" пояснение этого.

29 Чему равна сумма фибоначчиевых произведений

$$\sum_{m>0} \sum_{\substack{k_1+k_2+\quad+k_m=n\\ k_1,k_2,\ldots,k_m>0}} F_{k_1}F_{k_2}\ldots F_{k_m}\,?$$

30 Если производящая функция $G(z) = 1/(1-\alpha z)(1-\beta z)$ имеет разложение на элементарные дроби $a/(1-\alpha z)+b/(1-\beta z)$, то каково разложение на элементарные дроби функции $G(z)^n$?

31 Какая функция $g(n)$, определенная на положительных целых числах n, удовлетворяет рекуррентному соотношению

$$\sum_{d\backslash n} g(d)\,\varphi(n/d) \;=\; 1\,,$$

где φ — функция Эйлера?

32 *Арифметическая прогрессия* представляет собой бесконечное множество целых чисел

$$\{an+b\} \;=\; \{b, a+b, 2a+b, 3a+b, \ldots\}.$$

Множество арифметических прогрессий $\{a_1n + b_1\}$, $\ldots$, $\{a_mn + b_m\}$ называется *точным покрытием*, если всякое неотрицательное целое число встречается в одной и только одной прогрессии. Например, точное покрытие образуют три прогрессии: $\{2n\}$, $\{4n+1\}$, $\{4n+3\}$. Покажите, что если $\{a_1n+b_1\}$, $\ldots$, $\{a_mn+b_m\}$— точное покрытие, такое, что $2 \leqslant a_1 \leqslant \cdots \leqslant a_m$, то $a_{m-1} = a_m$. *Указание*: воспользуйтесь производящими функциями.

Контрольные работы

33 Чему равно $[w^mz^n]\,\big(\ln(1+z)\big)/(1-wz)$?

34 Найдите аналитический вид производящей функции $\sum_{n\geqslant 0} G_n(z)w^n$, если

$$G_n(z) \;=\; \sum_{k\leqslant n/m} \binom{n-mk}{k} z^{mk}.$$

(Здесь m— фиксированное положительное целое число.)

35 Вычислите сумму $\sum_{0<k<n} 1/k(n-k)$ двумя способами:

а разложив слагаемые на простейшие дроби;

б рассмотрев указанную сумму как свертку и использовав производящие функции.

36 Пусть $A(z)$ — производящая функция для $\langle a_0, a_1, a_2, a_3, \ldots \rangle$. Выразите $\sum_n a_{\lfloor n/m \rfloor} z^n$ через A, z и m.

37 Пусть a_n обозначает количество способов записи положительного целого числа n в виде суммы степеней 2 без учета порядка. Например, $a_4 = 4$, поскольку $4 = 2+2 = 2+1+1 = 1+1+1+1$. Положим по определению $a_0 = 1$. Обозначим через $b_n = \sum_{k=0}^{n} a_k$ частичные суммы последовательности a.

а Составьте таблицу чисел a_n и b_n до $n = 10$. Какое замечательное соотношение можно усмотреть в таблице? (Пока что не доказывайте его.)

б Представьте производящую функцию $A(z)$ в виде бесконечного произведения.

в Воспользуйтесь выражением из п. (б), чтобы доказать результат из п. (а).

38 Найдите аналитический вид двойной производящей функции

$$M(w,z) \;=\; \sum_{m,n\geqslant 0} \min(m,n)\, w^m z^n\,.$$

Обобщите свой ответ таким образом, чтобы при фиксированном $m \geqslant 2$ получить аналитическое выражение для

$$M(z_1,\ldots,z_m) \;=\; \sum_{n_1,\ldots,n_m\geqslant 0} \min(n_1,\ldots,n_m)\, z_1^{n_1}\ldots z_m^{n_m}\,.$$

39 Для заданных положительных целых чисел m и n найдите аналитический вид выражений

$$\sum_{1\leqslant k_1<k_2<\ <k_m\leqslant n} k_1k_2\ldots k_m \quad \text{и}$$

$$\sum_{1\leqslant k_1\leqslant k_2\leqslant\,\cdot\,\leqslant k_m\leqslant n} k_1k_2\ldots k_m\,.$$

(Например, при $m = 2$ и $n = 3$ суммы представляют собой соответственно $1\cdot2+1\cdot3+2\cdot3$ и $1\cdot1+1\cdot2+1\cdot3+2\cdot2+2\cdot3+3\cdot3$.) *Указание*: какие коэффициенты при z^m будут в производящих функциях $(1+a_1z)\ldots(1+a_nz)$ и $1/(1-a_1z)\ldots(1-a_nz)$?

40 Выразите $\sum_k \binom{n}{k}(kF_{k-1}-F_k)(n-k)¡$ в аналитическом виде.

41 *Возрастающе-убывающей перестановкой* порядка n называется перестановка $a_1 a_2 \ldots a_n$ чисел $\{1, 2, \ldots, n\}$, которая поочередно возрастает и убывает:

$$a_1 < a_2 > a_3 < a_4 > \cdots .$$

Например, 35142 есть возрастающе-убывающая перестановка порядка 5. Пусть A_n обозначает количество возрастающе-убывающих перестановок порядка n. Покажите, что экспоненциальной производящей функцией $\langle A_n \rangle$ является $(1 + \sin z)/\cos z$.

42 Космический зонд обнаружил, что органика на Марсе имеет ДНК, в состав которой входят пять аминокислот (a, b, c, d, e), а не четыре, как на Земле. В марсианской ДНК никогда не встречаются пары cd, ce, ed и ee; все цепочки, в которых нет запрещенных пар, возможны. (Таким образом, цепочка $bbcda$ запрещена, а $bbdca$ разрешена.) Сколько всего цепочек ДНК длиной n возможно на Марсе? (Для $n = 2$ ответ — 21, потому что мы различаем левый и правый концы цепочки.)

43 *Производящая функция Ньютона* последовательности $\langle g_n \rangle$ определяется как

$$\dot{G}(z) = \sum_n g_n \binom{z}{n} .$$

Найдите формулу свертки, которая устанавливает соотношение между последовательностями $\langle f_n \rangle$, $\langle g_n \rangle$ и $\langle h_n \rangle$, производящие функции Ньютона которых удовлетворяют соотношению $\dot{F}(z)\dot{G}(z) = \dot{H}(z)$. Постарайтесь, чтобы ваша формула была как можно более простой и симметричной.

44 Пусть q_n обозначает количество возможных исходов при попарном сравнении n чисел $\{x_1, \ldots, x_n\}$. Например, $q_3 = 13$, потому что возможны следующие исходы:

$$\begin{array}{llll}
x_1 < x_2 < x_3\,; & x_1 < x_2 = x_3\,; & x_1 < x_3 < x_2\,; & x_1 = x_2 < x_3\,; \\
x_1 = x_2 = x_3\,; & x_1 = x_3 < x_2\,; & x_2 < x_1 < x_3\,; & \\
x_2 < x_1 = x_3\,; & x_2 < x_3 < x_1\,; & x_2 = x_3 < x_1\,; & \\
x_3 < x_1 < x_2\,; & x_3 < x_1 = x_2\,; & x_3 < x_2 < x_1\,. &
\end{array}$$

Найдите аналитический вид экспоненциальной производящей функции $\widehat{Q}(z) = \sum_n q_n z^n/n!$. Найдите также последовательности $\langle a_n \rangle$, $\langle b_n \rangle$, $\langle c_n \rangle$, такие, что

$$q_n = \sum_{k \geqslant 0} k^n a_k = \sum_k \left\{ {n \atop k} \right\} b_k = \sum_k \left\langle {n \atop k} \right\rangle c_k$$

при всех $n > 0$.

45 Вычислите $\sum_{m,n>0}[m \perp n]/m^2n^2$.

46 Вычислите

$$\sum_{0 \leqslant k \leqslant n/2} \binom{n-2k}{k} \left(\frac{-4}{27}\right)^k$$

в аналитическом виде. *Указание*: $z^3 - z^2 + \frac{4}{27} = (z + \frac{1}{3})(z - \frac{2}{3})^2$.

47 Покажите, что приведенные в (7.34) числа U_n и V_n из задачи о покрытии костями домино прямоугольника размером $3 \times n$ тесно связаны с дробями в дереве Штерна–Броко, сходящимися к $\sqrt{3}$.

48 Некоторая определенная последовательность $\langle g_n \rangle$ удовлетворяет рекуррентному соотношению

$$a g_n + b g_{n+1} + c g_{n+2} + d \;=\; 0\,, \qquad \text{целое } n \geqslant 0,$$

для некоторых целых чисел (a, b, c, d) с НОД$(a, b, c, d) = 1$. Она также имеет аналитический вид

$$g_n \;=\; \lfloor \alpha(1+\sqrt{2}\,)^n \rfloor\,, \qquad \text{целое } n \geqslant 0,$$

для некоторого действительного числа α между 0 и 1. Найдите a, b, c, d и α.

О четных, нечетных и почетных степенях.

49 Задача о степенях и четности.

а Рассмотрим последовательность $\langle a_0, a_1, a_2, \ldots \rangle = \langle 2, 2, 6, \ldots \rangle$, определяемую формулой

$$a_n \;=\; (1+\sqrt{2})^n + (1-\sqrt{2})^n\,.$$

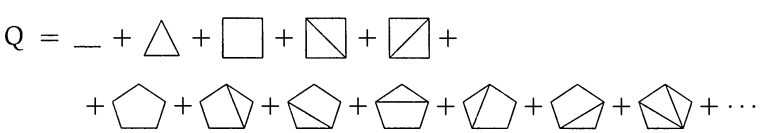
Найдите простое рекуррентное соотношение, которому удовлетворяет эта последовательность.

б Докажите, что $\lceil (1+\sqrt{2})^n \rceil \equiv n \pmod 2$ для всех целых $n > 0$.

в Найдите число α вида $(p + \sqrt{q})/2$, где p и q — положительные целые числа, такие, что $\lfloor \alpha^n \rfloor \equiv n \pmod 2$ для всех целых $n > 0$.

Дополнительные задачи

50 В продолжение упр. 22 рассмотрим сумму всех способов разбиения многоугольников на многоугольники:



Найдите символическое уравнение для Q и используйте его для нахождения производящей функции для количества способов проведения непересекающихся диагоналей в выпуклом n-угольнике. (Требуется найти выражение производящей функции в аналитическом виде как функции от z; находить аналитический вид коэффициентов не нужно.)

51 Докажите, что произведение

$$2^{mn/2} \prod_{\substack{1\leqslant j\leqslant m\\ 1\leqslant k\leqslant n}} \left(\left(\cos^2 \frac{j\pi}{m+1}\right)\text{▯}^2 + \left(\cos^2 \frac{k\pi}{n+1}\right)\text{▭}^2\right)^{1/4}$$

является производящей функцией для покрытий костями домино прямоугольника размером $m \times n$. (Входящие в формулу mn сомножителей можно представлять выписанными в mn клетках прямоугольника. Если mn нечетное, то средний сомножитель нулевой. Коэффициентом при $\text{▯}^j\text{▭}^k$ является количество способов покрытия прямоугольника с помощью j вертикальных и k горизонтальных костей домино.) *Указание*: это сложная задача, выходящая за рамки данной книги, поэтому можно ограничиться простой проверкой справедливости формулы в случае $m = 3$, $n = 4$.

Так это указание или предостережение?

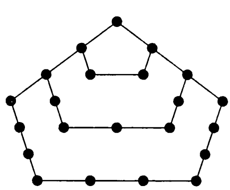
52 Докажите, что полиномы, определяемые рекуррентным соотношением

$$p_n(y) = \left(y - \frac{1}{4}\right)^n - \sum_{k=0}^{n-1} \binom{2n}{2k} \left(\frac{-1}{4}\right)^{n-k} p_k(y),$$

целое $n \geqslant 0$,

имеют вид $p_n(y) = \sum_{m=0}^{n} \left|{n \atop m}\right| y^m$, где $\left|{n \atop m}\right|$ — положительное целое число при $1 \leqslant m \leqslant n$. *Указание*: это упражнение очень поучительное, но не очень простое.

53 Последовательность *пятиугольных чисел* $\langle 1, 5, 12, 22, \ldots \rangle$ является очевидным обобщением треугольных и квадратных чисел:



Пусть $T_n = n(n+1)/2$ — n-е треугольное число; $P_n = n(3n-1)/2$ — n-е пятиугольное число; и, наконец, пусть U_n — количество покрытий прямоугольника размером $3 \times n$ костями домино, определенное в (7.38). Докажите, что треугольное число $T_{(U_{4n+2}-1)/2}$ является одновременно пятиугольным числом. *Указание*: $3U_{2n}^2 = (V_{2n-1} + V_{2n+1})^2 + 2$.

54 Рассмотрим следующую интересную конструкцию:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	...
1	2	3	4		6	7	8	9		11	12	13	14		16	...
1	3	6	10		16	23	31	40		51	63	76	90		106	...
1	3	6			16	23	31			51	63	76			106	...
1	4	10			26	49	80			131	194	270			376	...
1	4				26	49				131	194				376	...
1	5				31	80				211	405				781	...
1					31					211					781	...
1					32					243					1024	...

(Начните со строки, содержащей все положительные целые числа. Затем удалите каждый m-й столбец; здесь $m = 5$. Замените оставшиеся элементы их частичными суммами. Удалите каждый $(m-1)$-й столбец. Вновь замените элементы частичными суммами и т.д.) Используя производящие функции, покажите, что в конце концов получится последовательность m-х степеней. Например, для $m = 5$ получается, как показано на рисунке, последовательность $\langle 1^5, 2^5, 3^5, 4^5, \ldots \rangle$.

55 Докажите, что если степенные ряды $F(z)$ и $G(z)$ обладают свойством дифференцируемой конечности (определенной в упр. 20), то тем же свойством обладают $F(z)+G(z)$ и $F(z)\,G(z)$.

Исследовательские проблемы

56 Докажите, что в некотором большом классе "простых аналитических выражений" не существует "простого аналитического выражения" для коэффициента при z^n в $(1+z+z^2)^n$ в виде функции от n.

57 Докажите или опровергните: если все коэффициенты ряда $G(z)$ равны 0 или 1 и при этом все коэффициенты $G(z)^2$ меньше некоторой константы M, то у $G(z)^2$ бесконечное количество коэффициентов равно нулю.

8

Дискретная вероятность

ЭЛЕМЕНТ СЛУЧАЙНОСТИ привлекается практически всегда, когда мы пытаемся объяснить окружающий нас мир. Математическая *теория вероятностей* позволяет вычислять вероятности сложных событий в предположении, что эти события подчиняются некоторым аксиомам. Она имеет важные приложения во всех областях науки и тесно связана с методами, изучавшимися в предыдущих главах.

Вероятности называются дискретными, если можно вычислить вероятности всех событий путем суммирования, а не интегрирования. Мы уже умеем работать с суммами, так что нет ничего удивительного в применении наших знаний на практике для вычисления ряда интересных вероятностей и средних значений.

8.1 Определения

Теория вероятностей начинается с идеи *пространства вероятностей* (или *вероятностного пространства*), которое представляет собой множество Ω всего, что может случиться в рассматриваемой задаче, в совокупности с правилом, которое назначает каждому элементарному событию $\omega \in \Omega$ вероятность $\Pr(\omega)$. Вероятность $\Pr(\omega)$ должна быть неотрицательным действительным числом, и при этом должно выполняться условие

Читатели, не знакомые с теорией вероятностей, с большой вероятностью получат пользу от изучения классического введения в этот предмет, принадлежащего перу Феллера (Feller) [120].

$$\sum_{\omega \in \Omega} \Pr(\omega) = 1. \tag{8.1}$$

Таким образом, каждое значение $\Pr(\omega)$ должно находиться в диапазоне $[0\,..\,1]$. Мы называем $\Pr$ *распределением вероятностей*, потому что функция $\Pr$ распределяет суммарную вероятность 1 между событиями ω.

Вот пример: в случае пары игральных костей множество элементарных событий Ω представляет собой $D^2 = \{$⚀⚀,

⚀⚁, ..., ⚅⚅}, где

$$D = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$$

представляет собой множество всех шести способов выбрасывания игральной кости. Исходы ⚀⚁ и ⚁⚀ рассматриваются как различные; следовательно, всего данное пространство вероятностей состоит из $6^2 = 36$ элементов.

Учтите — пристрастие к азартным играм может плохо кончиться!

Обычно игральные кости считаются "честными", т.е. все шесть вариантов выпадения одной игральной кости имеют одинаковую вероятность $\frac{1}{6}$, а каждый из 36 вариантов выпадения двух игральных костей из Ω имеет вероятность $\frac{1}{36}$. Но можно рассмотреть и "фальшивые" игральные кости, у которых распределение вероятностей неоднородно. Например, пусть

$$\Pr{}_1(⚀) = \Pr{}_1(⚅) = \tfrac{1}{4};$$
$$\Pr{}_1(⚁) = \Pr{}_1(⚂) = \Pr{}_1(⚃) = \Pr{}_1(⚄) = \tfrac{1}{8}.$$

Тогда $\sum_{d \in D} \Pr_1(d) = 1$, так что $\Pr_1$ является распределением вероятностей на множестве D, и мы можем присвоить вероятности элементам $\Omega = D^2$ посредством правила

$$\Pr{}_{11}(d\,d') = \Pr{}_1(d)\,\Pr{}_1(d'). \tag{8.2}$$

Например, $\Pr_{11}(⚅⚂) = \frac{1}{4} \cdot \frac{1}{8} = \frac{1}{32}$. Это корректное распределение, потому что

$$\sum_{\omega \in \Omega} \Pr{}_{11}(\omega) = \sum_{d\,d' \in D^2} \Pr{}_{11}(d\,d') = \sum_{d,d' \in D} \Pr{}_1(d)\,\Pr{}_1(d') =$$
$$= \sum_{d \in D} \Pr{}_1(d) \sum_{d' \in D} \Pr{}_1(d') = 1 \cdot 1 = 1.$$

Мы можем рассмотреть и случай выбрасывания честной и фальшивой игральных костей,

$$\Pr{}_{01}(d\,d') = \Pr{}_0(d)\,\Pr{}_1(d'), \qquad \text{где } \Pr{}_0(d) = \tfrac{1}{6}; \tag{8.3}$$

в этом случае $\Pr_{01}(⚅⚂) = \frac{1}{6} \cdot \frac{1}{8} = \frac{1}{48}$. В реальном мире грани игральных костей не могут быть совершенно одинаковы, но значение $\frac{1}{6}$ обычно весьма близко к реальности.

Если бы все грани были совершенно одинаковы, как бы мы могли их различить?

Событием называется подмножество Ω. В игре в кости, например, множество

$$\{⚀⚀, ⚁⚁, ⚂⚂, ⚃⚃, ⚄⚄, ⚅⚅\}$$

представляет собой событие "выпадение дубля". Отдельные элементы ω множества Ω называются *элементарными событиями*, потому что они не могут быть разложены на более мелкие подмножества; ω можно рассматривать как одноэлементное событие $\{\omega\}$.

Вероятность события A определяется формулой

$$\Pr(\omega \in A) \;=\; \sum_{\omega \in A} \Pr(\omega)\,; \tag{8.4}$$

и в общем случае, если $R(\omega)$ — некоторое утверждение относительно ω, записываем '$\Pr\bigl(R(\omega)\bigr)$' для суммы всех $\Pr(\omega)$, таких, что $R(\omega)$ истинно. Таким образом, например, вероятность выпадения дубля в случае честных игральных костей равна $\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}=\frac{1}{6}$; но если обе игральные кости несимметричны и имеют распределение вероятностей $\Pr_1$, то вероятность выпадения дубля составит $\frac{1}{16}+\frac{1}{64}+\frac{1}{64}+\frac{1}{64}+\frac{1}{64}+\frac{1}{16}=\frac{3}{16}>\frac{1}{6}$. Смещение центра тяжести игральных костей приводит к повышению вероятности выпадения дубля.

(Использованное здесь обозначение $\sum$ носит более общий характер, чем определенное в главе 2: суммы в (8.1) и (8.4) распространяются на все элементы ω произвольного множества, а не только на целые числа. Однако это расширение в действительности не вызывает никаких проблем; можно условиться использовать под знаком $\sum$ специальное обозначение, когда подразумеваются не целые числа, так что никаких коллизий с обычными обозначениями не возникнет. Остальные определения из главы 2 остаются корректными; в частности, определение бесконечной суммы из главы 2 дает подходящую интерпретацию для наших сумм в случае бесконечного множества Ω. Каждая вероятность неотрицательна, а сумма их всех ограничена, так что вероятность события A в (8.4) корректно определена для всех подмножеств $A \subseteq \Omega$.)

Функция, определенная на элементарных событиях ω из пространства вероятностей, называется *случайной величиной* (*случайной переменной*, *random variable*). Например, если $\Omega = D^2$, можно определить $S(\omega)$ как сумму очков на выпавших гранях для события ω, так что $S(⚅⚂) = 6+3 = 9$. Вероятность того, что сумма очков составит 7, есть вероятность события $S(\omega)=7$, а именно

$$\Pr(⚀⚅)+\Pr(⚁⚄)+\Pr(⚂⚃)$$
$$\qquad+\Pr(⚃⚂)+\Pr(⚄⚁)+\Pr(⚅⚀)\,.$$

В случае честных игральных костей ($\Pr = \Pr_{00}$) это событие происходит с вероятностью $\frac{1}{6}$; но если игральные кости фальшивые ($\Pr = \Pr_{11}$), то вероятность будет $\frac{1}{16} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{16} = \frac{3}{16}$, т.е. такой же, как и найденная ранее для выпадения дубля.

При рассмотрении случайных величин принято опускать обозначение '(ω)', потому что обычно в каждой задаче имеется только одно пространство вероятностей. Таким образом, можно говорить просто '$S = 7$' для обозначения того, что выпало 7 очков, или '$S = 4$' для обозначения события {⚀⚂, ⚁⚁, ⚂⚀}.

Случайная величина характеризуется распределением вероятностей ее значений. Так, например, S принимает одиннадцать возможных значений $\{2, 3, \ldots, 12\}$, и можно протабулировать вероятность каждого события $S = s$ для всех s из этого множества:

s	2	3	4	5	6	7	8	9	10	11	12
$\Pr_{00}(S=s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
$\Pr_{11}(S=s)$	$\frac{4}{64}$	$\frac{4}{64}$	$\frac{5}{64}$	$\frac{6}{64}$	$\frac{7}{64}$	$\frac{12}{64}$	$\frac{7}{64}$	$\frac{6}{64}$	$\frac{5}{64}$	$\frac{4}{64}$	$\frac{4}{64}$

Если мы работаем над задачей, которая включает только случайную величину S, но не другие характеристики игральных костей, то можно получить ответ, основанный только на этих вероятностях, безотносительно к деталям множества $\Omega = D^2$. Фактически можно определить пространство вероятностей как меньшее множество $\Omega = \{2, 3, \ldots, 12\}$ с желаемым распределением вероятностей $\Pr(s)$. Тогда '$S = 4$' будет элементарным событием. Таким образом, зачастую можно игнорировать пространство вероятностей Ω и работать непосредственно со случайными величинами и их распределениями.

Если на одном и том же пространстве вероятностей Ω определены две случайные величины X и Y, то, чтобы охарактеризовать их поведение, не зная ничего об Ω, требуется знать их "совместное распределение"

$$\Pr(X = x \text{ и } Y = y)$$

для каждого x из X и каждого y из Y. Будем говорить, что X и Y — *независимые* случайные величины, если

$$\Pr(X = x \text{ и } Y = y) = \Pr(X = x) \cdot \Pr(Y = y) \tag{8.5}$$

для всех x и y. Интуитивно это означает, что значение X не влияет на значение Y.

Например, если Ω есть множество бросаний игральных костей D^2, то можно определить S_1 как количество очков на первой игральной кости и S_2 — как количество очков на второй игральной кости. Тогда случайные величиныS_1 и S_2 независимы по отношению к каждому из рассмотренных ранее распределений вероятностей $\Pr_{00}$, $\Pr_{11}$ и $\Pr_{01}$, потому что мы определяли вероятность каждого элементарного события dd' как произведение вероятностей событий $S_1 = d$ и $S_2 = d'$. Эти вероятности можно определить и иначе, скажем, так, чтобы

Костлявое неравенство.

Pr(⚀⚄) / Pr(⚀⚅) $\neq$ Pr(⚁⚄) / Pr(⚁⚅) ;

но мы не сделали этого, поскольку обычно игральные кости не влияют друг на друга. В наших обозначениях оба эти отношения равны друг другу и $\Pr(S_2 = 5)/\Pr(S_2 = 6)$.

Мы определили S как сумму двух чисел очков, $S_1 + S_2$. Рассмотрим другую случайную величину P — произведение $S_1 S_2$. Являются ли S и P независимыми? Рассуждая неформально, нет; если нам сказали, что $S = 2$, то мы знаем, что P может быть равным только 1. Но и формально ответ отрицателен, потому что условие независимости (8.5) не выполняется (как минимум для честных игральных костей): для всех допустимых значений s и p имеем $0 < \Pr_{00}(S = s) \cdot \Pr_{00}(P = p) \leqslant \frac{1}{6} \cdot \frac{1}{9}$, что не может равняться значению $\Pr_{00}(S = s \text{ и } P = p)$, которое кратно $\frac{1}{36}$.

Желая понять типичное поведение некоторой случайной величины, мы часто интересуемся ее "средним" значением. Однако понятие среднего неоднозначно; обычно, говоря о последовательности чисел, люди подразумевают под средним три разных значения:

- *среднее арифметическое* (сумма всех значений, деленная на их количество);
- *медиана* (средний элемент упорядоченной последовательности);
- *мода* (значение, встречающееся чаще других).

Например, среднее арифметическое последовательности $\langle 3, 1, 4, 1, 5 \rangle$ равно $\frac{3+1+4+1+5}{5} = 2.8$; медиана равна 3, а мода — 1.

Однако обычно теория вероятности имеет дело не с последовательностями чисел, а со случайными величинами, так что надо определить понятие среднего для случайной величины. Предположим, что мы многократно повторяем эксперимент, производя независимые испытания таким образом, чтобы каждое значение X появлялось с частотой, примерно соответствующей его вероятности. (Например, мы могли бы много раз бросать игральные кости, наблюдая за значениями S и/или P.) Мы хотели бы так определить среднее значение случайной величины, чтобы последовательность чисел, полученная в результате таких экспериментов, имела, как правило, приблизительно те же значения среднего арифметического, медианы и моды, что и соответствующие значения для случайной величины X, согласно нашим определениям.

Вот как это можно сделать: *среднее* случайной величины X с действительными значениями на пространстве вероятности Ω определяется как

$$\sum_{x \in X(\Omega)} x \cdot \Pr(X = x), \tag{8.6}$$

если эта (потенциально бесконечная) сумма существует. (Здесь $X(\Omega)$ означает множество всех значений, которые может принимать X.) *Медиана* X определяется как множество всех таких x, что

$$\Pr(X \leqslant x) \geqslant \frac{1}{2} \quad \text{и} \quad \Pr(X \geqslant x) \geqslant \frac{1}{2}. \tag{8.7}$$

Наконец, *мода* X определяется как множество всех таких x, что

$$\Pr(X = x) \geqslant \Pr(X = x') \qquad \text{для всех } x' \in X(\Omega). \tag{8.8}$$

В нашем примере с бросанием игральных костей среднее значение S оказывается равным $2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + \cdots + 12 \cdot \frac{1}{36} = 7$ для распределения $\Pr_{00}$, для распределения $\Pr_{11}$ среднее значение то же — 7. И медианой, и модой для обоих распределений также оказывается $\{7\}$. Таким образом, S имеет одно и то же среднее значение по всем трем определениям. С другой стороны, случайная величина P при распределении $\Pr_{00}$ имеет среднее арифметическое $\frac{49}{4} = 12.25$; ее медианой является $\{10\}$, а модой — $\{6, 12\}$. Среднее значение P остается тем же и при переходе к игральным костям с распределением $\Pr_{11}$, но медиана при этом уменьшается до $\{8\}$, а модой становится единственный элемент $\{6\}$.

Для среднего арифметического случайной величины в теории вероятностей имеется специальное название и обозначение: оно называется *математическим ожиданием* и записывается как

$$EX = \sum_{\omega \in \Omega} X(\omega) \Pr(\omega). \tag{8.9}$$

В нашем примере с бросанием игральных костей эта сумма содержит 36 членов (по одному для каждого элемента Ω), в то время как (8.6) представляет собой сумму всего лишь 11 чисел. Но обе суммы имеют одно и то же значение, потому что обе равны

$$\sum_{\substack{\omega \in \Omega \\ x \in X(\Omega)}} x \Pr(\omega) \bigl[x = X(\omega)\bigr].$$

Итак, в среднем "среднее" означает "среднее арифметическое".

В приложениях среднее арифметическое случайной величины оказывается полезнее других средних, так что мы практически не будем встречаться с медианами и модами, и термины "ожидаемое значение", "математическое ожидание", "среднее арифметическое" и "среднее" до конца главы являются взаимозаменяемыми.

Если X и Y — две произвольные случайные величины, определенные на одном и том же пространстве вероятностей, то $X+Y$ также является случайной величиной на этом пространстве. По формуле (8.9) средним суммы случайных величин является сумма их средних:

$$E(X+Y) = \sum_{\omega\in\Omega} \bigl(X(\omega)+Y(\omega)\bigr)\Pr(\omega) = EX+EY. \tag{8.10}$$

Аналогично, если α — произвольная константа, то справедливо простое правило

$$E(\alpha X) = \alpha EX. \tag{8.11}$$

Но соответствующее правило для произведения случайных величин оказывается в общем случае более сложным; математическое ожидание определяется как сумма по элементарным событиям, а сумма произведений обычно не имеет простого выражения. Несмотря на эту сложность в частном случае произведения независимых случайных величин работает очень простая и красивая формула:

$$E(XY) = (EX)(EY), \qquad \text{если } X \text{ и } Y \text{ независимы.} \tag{8.12}$$

Это правило можно доказать при помощи дистрибутивного закона для произведений:

$$\begin{aligned} E(XY) &= \sum_{\omega\in\Omega} X(\omega)Y(\omega)\cdot\Pr(\omega) = \\ &= \sum_{\substack{x\in X(\Omega)\\ y\in Y(\Omega)}} xy\cdot\Pr(X=x \text{ и } Y=y) = \\ &= \sum_{\substack{x\in X(\Omega)\\ y\in Y(\Omega)}} xy\cdot\Pr(X=x)\Pr(Y=y) = \\ &= \sum_{x\in X(\Omega)} x\Pr(X=x) \cdot \sum_{y\in Y(\Omega)} y\Pr(Y=y) = (EX)(EY). \end{aligned}$$

Например, мы знаем, что $S = S_1 + S_2$ и $P = S_1S_2$, где S_1 и S_2 — числа очков на первой и второй игральных костях. Имеем $ES_1 = ES_2 = \frac{7}{2}$, следовательно, $ES = 7$; кроме того, S_1 и S_2

независимы, так что $EP = \frac{7}{2} \cdot \frac{7}{2} = \frac{49}{4}$, как и утверждалось ранее. Имеем также $E(S + P) = ES + EP = 7 + \frac{49}{4}$. Но S и P не являются независимыми, так что мы не можем утверждать, что $E(SP) = 7 \cdot \frac{49}{4} = \frac{343}{4}$. Фактически ожидаемое значение SP равно $\frac{637}{6}$ при распределении $\Pr_{00}$ и ровно 112 — при распределении $\Pr_{11}$.

8.2 Математическое ожидание и дисперсия

Следующим по важности свойством случайной величины после математического ожидания является ее *дисперсия*, определяемая как средний квадрат отклонения от среднего:

$$VX = E\big((X - EX)^2\big). \tag{8.13}$$

Если обозначить EX через μ, то дисперсия VX представляет собой ожидаемое значение $(X - \mu)^2$. Это мера "разброса" распределения X.

В качестве простого примера вычисления дисперсии предположим, что нам сделали предложение, от которого невозможно отказаться: некто подарил нам два билета для участия в некоторой лотерее. Организаторы лотереи продают еженедельно по 100 билетов, участвующих в отдельном тираже. В тираже выбирается равномерным случайным образом один из билетов, и обладатель счастливого билета получает сто миллионов долларов. Остальные 99 владельцев билетов не получают ничего.

Подарком можно воспользоваться двумя способами: либо использовать оба билета в одной лотерее, либо по одному — в двух лотереях. Какая стратегия лучше? Попробуем провести анализ. Обозначим через X_1 и X_2 случайные величины, представляющие размер нашего выигрыша по первому и второму билетам. Ожидаемое значение X_1 в миллионах составляет

Тут есть тонкость: в зависимости от выбранной стратегии следует использовать разные пространства вероятностей, но EX_1 и EX_2 в обоих случаях одни и те же.

$$EX_1 = \frac{99}{100} \cdot 0 + \frac{1}{100} \cdot 100 = 1,$$

и то же самое справедливо и для X_2. Ожидаемые значения аддитивны, так что средний суммарный выигрыш составит

$$E(X_1 + X_2) = EX_1 + EX_2 = 2 \text{ миллиона долларов}$$

независимо от выбранной стратегии.

Тем не менее стратегии различны. Выйдя за рамки ожидаемых значений и рассмотрев распределение вероятностей $X_1 + X_2$, мы получим

	Выигрыш (миллионы)		
	0	100	200
Один тираж	.9800	.0200	
Разные тиражи	.9801	.0198	.0001

Если использовать оба билета в одной лотерее, то шансы не выиграть ничего составят 98%, а шансы выиграть 100 миллионов — 2%. Если же билеты играют в разных тиражах, то шансы не выиграть ничего составляет 98.01%, что несколько больше, чем раньше; шансы выиграть 200 миллионов составляют 0.01%, что тоже больше, чем было раньше; шансы же выиграть 100 миллионов равны теперь 1.98%. Таким образом, во втором случае распределение случайной величины $X_1 + X_2$ несколько более разбросано; среднее значение, 100 миллионов, в этом случае менее вероятно, зато немного более вероятны крайние значения.

Дисперсия призвана отразить именно эту "разбросанность" случайной величины. Мерой разброса служит квадрат отклонения случайной величины от ее математического ожидания. В первом случае дисперсия равна

$$.98(0M - 2M)^2 + .02(100M - 2M)^2 = 196M^2 ;$$

во втором —

$$.9801(0M{-}2M)^2 + .0198(100M{-}2M)^2 + .0001(200M{-}2M)^2 = \\ = 198M^2 .$$

Как и ожидалось, последняя величина несколько больше, так как распределение вероятностей во втором случае более разбросано.

Интересно, что дисперсия суммы в долларах выражается в квадратных долларах.

При работе с дисперсиями все числа возводятся в квадрат, так что в результате получаются очень большие значения. (Так, множитель M^2 равен триллиону — впечатляющая величина даже для профессиональных игроков.) Для возврата к более осмысленной шкале часто из дисперсии извлекается квадратный корень, что дает *стандартное отклонение*, которое обычно обозначается греческой буквой σ:

$$\sigma = \sqrt{VX} . \tag{8.14}$$

Стандартные отклонения случайной величины $X_1 + X_2$ для двух рассмотренных стратегий равны $\sqrt{196M^2} = 14.00M$ и $\sqrt{198M^2} \approx 14.071247M$. В некотором смысле вторая стратегия на 71 247 долларов рискованнее.

Чем дисперсия может помочь при выборе стратегии? Ясного ответа на этот вопрос нет. Стратегия с большей дисперсией рискованнее, но что лучше для кошелька — риск или безопасность?

Предположим, что можно купить не два билета, а все сто. Тогда выигрыш в одной лотерее был бы гарантирован, а дисперсия была бы равна нулю. Можно было бы сыграть в сотне тиражей, ничего не получая с вероятностью $.99^{100} \approx .366$, но с ненулевой вероятностью выиграть 10 000 000 000 долларов. Какую из стратегий выбрать — этот вопрос выходит за рамки данной книги, в которой все, что мы можем сделать, — это объяснить, как выполнить необходимые расчеты.

Один из способов снижения риска состоит в подкупе комиссии, но, по моему мнению, при этом вероятность превращается из дискретной в дискредитирующую.

(N.B.: замечания на полях не всегда совпадают с точкой зрения администрации.)

В действительности имеется и более простой способ вычисления дисперсии, чем прямое использование определения (8.13). (Есть все основания предполагать здесь наличие какой-то скрытой математики, ведь дисперсия в лотерейных примерах оказалась кратной M^2.) Имеем

$$\begin{aligned} E\big((X-EX)^2\big) &= E\big(X^2-2X(EX)+(EX)^2\big) = \\ &= E(X^2)-2(EX)(EX)+(EX)^2, \end{aligned}$$

поскольку (EX) — константа; следовательно,

$$VX = E(X^2)-(EX)^2. \tag{8.15}$$

"Дисперсия есть среднее значение квадрата минус квадрат среднего значения".

Например, в "лотерейной" задаче среднее $(X_1+X_2)^2$ равно $.98(0M)^2+.02(100M)^2 = 200M^2$ или $.9801(0M)^2+.0198(100M)^2+.0001(200M)^2 = 202M^2$. Вычтя $4M^2$ (квадрат среднего), мы получим те же результаты, которые ранее заставили нас потрудиться.

Есть еще более простая формула, применимая при вычислении $V(X+Y)$ для независимых X и Y:

$$\begin{aligned} E\big((X+Y)^2\big) &= E(X^2+2XY+Y^2) = \\ &= E(X^2)+2(EX)(EY)+E(Y^2), \end{aligned}$$

поскольку мы знаем, что в случае независимых величин справедливо соотношение $E(XY) = (EX)(EY)$. Следовательно,

$$\begin{aligned} V(X+Y) &= E\big((X+Y)^2\big)-(EX+EY)^2 = \\ &= E(X^2)+2(EX)(EY)+E(Y^2) = \\ &\quad -(EX)^2-2(EX)(EY)-(EY)^2 = \\ &= E(X^2)-(EX)^2+E(Y^2)-(EY)^2 = \\ &= VX+VY. \end{aligned} \tag{8.16}$$

"Дисперсия суммы независимых случайных величин равна сумме их дисперсий".

Например, дисперсия суммы, которую можно выиграть на один лотерейный билет, равна

$$E(X_1^2) - (EX_1)^2 \;=\; .99(0M)^2 + .01(100M)^2 - (1M)^2 \;=\; 99M^2\,.$$

Следовательно, дисперсия суммарного выигрыша по двум лотерейным билетам в двух отдельных (независимых) лотереях равна $2 \times 99M^2 = 198M^2$. Соответствующее значение дисперсии для n независимых лотерейных билетов равно $n \times 99M^2$.

Дисперсия суммы S очков, выпавших на двух игральных костях, может быть получена по той же формуле, поскольку $S = S_1 + S_2$ есть сумма двух независимых случайных величин. Имеем

$$VS_1 \;=\; \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2) - \left(\frac{7}{2}\right)^2 \;=\; \frac{35}{12}$$

для честной игральной кости; следовательно, $VS = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}$. Для фальшивой игральной кости

$$VS_1 \;=\; \frac{1}{8}(2\cdot 1^2+2^2+3^2+4^2+5^2+2\cdot 6^2) - \left(\frac{7}{2}\right)^2 \;=\; \frac{45}{12}\,;$$

следовательно, $VS = \frac{45}{6} = 7.5$, когда обе игральные кости фальшивые. Заметим, что в последнем случае дисперсия S больше, хотя S принимает значение 7 чаще, чем в случае фальшивых игральных костей. Если наша цель — выбросить побольше приносящих удачу семерок, то дисперсия — не лучший показатель успеха.

Мы знаем, как вычислить дисперсию. Но зачем это делать? Основная причина заключается в *неравенстве Чебышева* ([29] и [57]), которое устанавливает важное свойство дисперсии:

$$\Pr\big((X - EX)^2 \geqslant \alpha\big) \;\leqslant\; VX/\alpha \qquad \text{для всех } \alpha > 0. \tag{8.17}$$

(Это неравенство отличается от неравенств Чебышева для сумм, с которыми мы встречались в главе 2.) Грубо (8.17) гласит, что случайная величина X редко принимает значения, далекие от своего среднего EX, если ее дисперсия VX мала. Доказательство неравенства на удивление простое. Имеем

$$\begin{aligned} VX \;&=\; \sum_{\omega\in\Omega} \big(X(\omega) - EX\big)^2 \Pr(\omega) \;\geqslant \\ &\geqslant \sum_{\substack{\omega\in\Omega \\ (X(\omega)-EX)^2 \geqslant \alpha}} \big(X(\omega) - EX\big)^2 \Pr(\omega) \;\geqslant \\ &\geqslant \sum_{\substack{\omega\in\Omega \\ (X(\omega)-EX)^2 \geqslant \alpha}} \alpha \Pr(\omega) \;=\; \alpha\cdot\Pr\big((X - EX)^2 \geqslant \alpha\big)\,; \end{aligned}$$

деление на α завершает доказательство.

Если ввести обозначение μ для математического ожидания и σ — для стандартного отклонения и заменить α на c^2VX в (8.17), то условие $(X - EX)^2 \geqslant c^2VX$ превратится в $(X-\mu)^2 \geqslant (c\sigma)^2$; следовательно, (8.17) гласит, что

$$\Pr\bigl(|X-\mu| \geqslant c\sigma\bigr) \leqslant 1/c^2 . \qquad (8.18)$$

Таким образом, X будет лежать в пределах c-кратного стандартного отклонения от своего среднего значения, за исключением случаев, вероятность которых не превышает $1/c^2$. Случайная величина будет лежать в пределах 2σ от μ как минимум в 75% случаев, а в пределах между $\mu - 10\sigma$ и $\mu + 10\sigma$ — по крайней мере для 99%. Это случаи $\alpha = 4VX$ и $\alpha = 100VX$ неравенства Чебышева.

При бросании пары честных игральных костей n раз общая сумма очков во всех n бросаниях при бо́льших n почти всегда будет близка к $7n$. Вот в чем причина: дисперсия n независимых бросаний равна $\frac{35}{6}n$. Дисперсия $\frac{35}{6}n$ означает, что стандартное отклонение составляет $\sqrt{\frac{35}{6}n}$. Поэтому неравенство Чебышева говорит нам, что сумма очков будет лежать между

$$7n - 10\sqrt{\tfrac{35}{6}n} \quad \text{и} \quad 7n + 10\sqrt{\tfrac{35}{6}n}$$

как минимум в 99% всех экспериментов по бросанию n раз пары честных игральных костей. Например, можно держать пари 99 к 1, что результат одного миллиона бросков пары честных игральных костей окажется между 6.975 миллиона и 7.025 миллиона.

В общем случае пусть X — *любая* случайная величина на пространстве вероятностей Ω, имеющая конечное математическое ожидание μ и конечное стандартное отклонение σ. Тогда можно рассмотреть пространство вероятностей Ω^n, элементарными событиями которого являются n-последовательности $(\omega_1, \omega_2, \ldots, \omega_n)$, где каждое $\omega_k \in \Omega$, а вероятности определяются как

$$\Pr(\omega_1, \omega_2, \ldots, \omega_n) = \Pr(\omega_1)\Pr(\omega_2)\ldots\Pr(\omega_n) .$$

Если теперь определить случайные величины X_k по формуле

$$X_k(\omega_1, \omega_2, \ldots, \omega_n) = X(\omega_k) ,$$

то величина

$$X_1 + X_2 + \cdots + X_n$$

будет суммой n независимых случайных величин, которая соответствует взятию n независимых "образцов" X на Ω и их суммированию. Математическое ожидание $X_1 + X_2 + \cdots + X_n$ равно $n\mu$, а стандартное отклонение равно $\sqrt{n}\,\sigma$; следовательно, среднее значение n образцов,

$$\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\,,$$

будет лежать в пределах от $\mu - 10\sigma/\sqrt{n}$ до $\mu + 10\sigma/\sqrt{n}$ как минимум в 99% случаев. Другими словами, если выбрать достаточно большое n, то среднее арифметическое n независимых испытаний почти всегда окажется очень близким к ожидаемому значению EX. (В учебниках по теории вероятности доказывается более строгая теорема, называющаяся усиленным законом больших чисел; но нам будет достаточно и простого следствия только что выведенного неравенства Чебышева.)

То есть среднее окажется в указанных пределах в как минимум 99% всех случаев, когда мы рассматриваем набор из n независимых образцов для любого фиксированного n. Не распространяйте это утверждение на среднее бесконечной последовательности X_1, X_2, X_3, ... с меняющимся n.

Иногда нам неизвестны характеристики пространства вероятностей, но требуется оценить математическое ожидание случайной величины X при помощи многократных наблюдений ее значения. (Например, нам могла бы понадобиться средняя полуденная температура января в Киеве или средняя продолжительность жизни на Украине.) Если в нашем распоряжении имеются независимые эмпирические наблюдения $X_1, X_2, \ldots, X_n$, то можно предположить, что истинное математическое ожидание приближенно равно

$$\widehat{E}X = \frac{X_1 + X_2 + \cdots + X_n}{n}\,. \tag{8.19}$$

Можно оценить и дисперсию — по формуле

$$\widehat{V}X = \frac{X_1^2 + X_2^2 + \cdots + X_n^2}{n-1} - \frac{(X_1 + X_2 + \cdots + X_n)^2}{n(n-1)}\,. \tag{8.20}$$

Возникает ощущение, что $(n-1)$ в этой формуле — опечатка и здесь должно быть n, как в (8.19), потому что истинное значение дисперсии определяется в (8.15) через ожидаемые значения. Однако использование $n-1$ вместо n дает лучшую оценку, поскольку из определения (8.20) вытекает, что

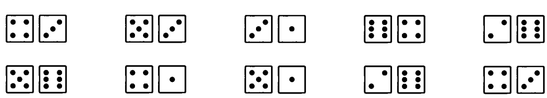
$$E(\widehat{V}X) = VX\,. \tag{8.21}$$

Вот почему:

$$\begin{aligned}
E(\widehat{V}X) &= \frac{1}{n-1}E\Big(\sum_{k=1}^{n} X_k^2 - \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n} X_jX_k\Big) = \\
&= \frac{1}{n-1}\Big(\sum_{k=1}^{n} E(X_k^2) - \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n} E(X_jX_k)\Big) = \\
&= \frac{1}{n-1}\Big(\sum_{k=1}^{n} E(X^2) - \\
&\qquad - \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\big(E(X)^2[j\ne k] + E(X^2)[j=k]\big)\Big) = \\
&= \frac{1}{n-1}\Big(nE(X^2) - \frac{1}{n}\big(nE(X^2) + n(n-1)E(X)^2\big)\Big) = \\
&= E(X^2) - E(X)^2 = VX.
\end{aligned}$$

(Этот вывод при замене $E(X_jX_k)$ на $(EX)^2[j\ne k] + E(X^2)[j=k]$ использует независимость наблюдений.)

На практике для оценки результатов эксперимента со случайной величиной X обычно вычисляются эмпирическое среднее $\hat{\mu} = \widehat{E}X$ и эмпирическое среднее отклонение $\hat{\sigma} = \sqrt{\widehat{V}X}$ и ответ представляется в виде ‘ $\hat{\mu} \pm \hat{\sigma}/\sqrt{n}$ ’. Например, пусть результаты выбрасываний двух (предположительно правильных) игральных костей выглядят следующим образом:



Эмпирическое среднее суммы очков S равно

$$\hat{\mu} = (7+11+8+5+4+6+10+8+8+7)/10 = 7.4;$$

эмпирическая дисперсия равна

$$(7^2+11^2+8^2+5^2+4^2+6^2+10^2+8^2+8^2+7^2-10\hat{\mu}^2)/9 \approx 2.1^2.$$

Таким образом, из этого эксперимента мы получаем следующую оценку суммы очков игральных костей: $7.4 \pm 2.1/\sqrt{10} \approx 7.4 \pm 0.7$.

Давайте рассмотрим еще один пример и разберемся, как вычислять математическое ожидание и дисперсию теоретически, а не эмпирически. Одной из рассмотренных в главе 5 задач была задача о n футбольных болельщиках, бросавших свои шляпы в воздух, в результате чего они случайным образом перепутывались. В (5.51) было показано, что с вероятностью $n¡/n! \approx 1/e$

никто не получит свою шляпу назад. Была также выведена формула

$$P(n,k) = \frac{1}{n!}\binom{n}{k}(n-k)¡ = \frac{1}{k!}\frac{(n-k)¡}{(n-k)!} \qquad (8.22)$$

вероятности того, что ровно k шляп вернутся к своим владельцам.

Для переформулировки этих результатов с использованием только что изученного формализма можно рассмотреть пространство вероятностей Π_n, состоящее из всех $n!$ перестановок π множества $\{1, 2, \ldots, n\}$, где $\Pr(\pi) = 1/n!$ для всех $\pi \in \Pi_n$. Случайная величина

Не путайте с числом Фибоначчи.

$$F_n(\pi) = \text{количество неподвижных точек в перестановке } \pi \in \Pi_n$$

отвечает числу правильных "шляпопадений" в задаче о футбольной победе. Уравнение (8.22) дает $\Pr(F_n = k)$; но мы пока забудем, что нам известны такие формулы. Мы хотим просто найти среднее значение величины F_n и ее стандартное отклонение.

Среднее значение можно легко вычислить, избежав всех сложностей главы 5. Достаточно просто заметить, что

$$\begin{aligned} F_n(\pi) &= F_{n,1}(\pi) + F_{n,2}(\pi) + \cdots + F_{n,n}(\pi)\,, \\ F_{n,k}(\pi) &= [\text{позиция } k \text{— неподвижная точка} \\ &\qquad \text{перестановки } \pi \in \Pi_n]. \end{aligned}$$

Следовательно,

$$EF_n = EF_{n,1} + EF_{n,2} + \cdots + EF_{n,n}\,.$$

А ожидаемое значение $F_{n,k}$ есть просто вероятность события $F_{n,k} = 1$, равная $1/n$, потому что ровно $(n-1)!$ из $n!$ перестановок $\pi = \pi_1\pi_2\ldots\pi_n \in \Pi_n$ имеют $\pi_k = k$. Следовательно,

$$EF_n = n/n = 1 \qquad \text{для } n > 0. \qquad (8.23)$$

Средняя шляпа.

В среднем на свое место попадает одна шляпа. *"Случайная перестановка имеет в среднем одну неподвижную точку".*

А как насчет стандартного отклонения? Этот вопрос труднее, потому что величины $F_{n,k}$ не являются независимыми друг от друга. Но дисперсию можно вычислить, проанализировав их взаимозависимость:

$$
\begin{aligned}
E(F_n^2) &= E\Bigl(\Bigl(\sum_{k=1}^{n} F_{n,k}\Bigr)^2\Bigr) = E\Bigl(\sum_{j=1}^{n}\sum_{k=1}^{n} F_{n,j}\,F_{n,k}\Bigr) = \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} E(F_{n,j}\,F_{n,k}) = \\
&= \sum_{1\leqslant k\leqslant n} E(F_{n,k}^2) + 2\sum_{1\leqslant j<k\leqslant n} E(F_{n,j}\,F_{n,k})\,.
\end{aligned}
$$

(Подобный трюк нами использовался при выводе (2.33) в главе 2.) Далее, $F_{n,k}^2 = F_{n,k}$, поскольку $F_{n,k}$ равен либо 0, либо 1; следовательно, $E(F_{n,k}^2) = EF_{n,k} = 1/n$, как и ранее. Кроме того, если $j < k$, то $E(F_{n,j}\,F_{n,k}) = \Pr(\text{в перестановке } \pi \text{ точки } j \text{ и } k \text{ неподвижны}) = (n-2)!/n! = 1/n(n-1)$. Следовательно,

$$
E(F_n^2) = \frac{n}{n} + \binom{n}{2}\frac{2}{n(n-1)} = 2 \qquad \text{для } n \geqslant 2. \tag{8.24}
$$

(В качестве проверки при $n = 3$ имеем $\frac{2}{6}0^2 + \frac{3}{6}1^2 + \frac{0}{6}2^2 + \frac{1}{6}3^2 = 2$.) Дисперсия равна $E(F_n^2) - (EF_n)^2 = 1$, так что стандартное отклонение (как и математическое ожидание) равно 1. "Случайная перестановка $n \geqslant 2$ элементов имеет 1 ± 1 неподвижных точек".

8.3 Вероятностные производящие функции

Если X — случайная величина, принимающая только целые неотрицательные значения, то ее распределение вероятностей можно компактно представить с применением методов из главы 7. *Производящей функцией случайной величины* X называется ряд

$$
G_X(z) = \sum_{k\geqslant 0} \Pr(X=k)\,z^k\,. \tag{8.25}
$$

Этот степенной ряд относительно z содержит всю информацию о случайной величине X. Его можно записать двумя другими способами:

$$
G_X(z) = \sum_{\omega\in\Omega} \Pr(\omega)\,z^{X(\omega)} = E(z^X)\,. \tag{8.26}
$$

Коэффициенты $G_X(z)$ неотрицательны, а их сумма равна 1; последнее условие можно записать как

$$
G_X(1) = 1\,. \tag{8.27}
$$

И обратно, любой степенной ряд $G(z)$ с неотрицательными коэффициентами и обладающий тем свойством, что $G(1) = 1$, является производящей функцией некоторой случайной величины.

Самое приятное свойство производящих функций случайных величин заключается в том, что обычно они упрощают вычисление средних и дисперсий. Например, вот как просто выражается математическое ожидание:

$$\begin{aligned} EX &= \sum_{k\geqslant 0} k\cdot \Pr(X=k) = \\ &= \sum_{k\geqslant 0} \Pr(X=k)\cdot kz^{k-1}\Big|_{z=1} = \\ &= G'_X(1)\,. \end{aligned} \tag{8.28}$$

Надо просто продифференцировать производящую функцию случайной величины по z и подставить $z = 1$.

Вычислить дисперсию лишь немногим сложнее:

$$\begin{aligned} E(X^2) &= \sum_{k\geqslant 0} k^2\cdot \Pr(X=k) = \\ &= \sum_{k\geqslant 0} \Pr(X=k)\cdot \bigl(k(k-1)z^{k-2} + kz^{k-1}\bigr)\Big|_{z=1} = \\ &= G''_X(1) + G'_X(1)\,. \end{aligned}$$

Следовательно,

$$VX = G''_X(1) + G'_X(1) - G'_X(1)^2\,. \tag{8.29}$$

Уравнения (8.28) и (8.29) гласят, что мы можем вычислить среднее и дисперсию, если сумеем найти две производные, $G'_X(1)$ и $G''_X(1)$. Нет необходимости знать аналитический вид выражений для вероятностей; нам даже не требуется знать аналитический вид самой функции $G_X(z)$.

Удобно ввести обозначения

$$\mathrm{Mean}(G) = G'(1)\,, \tag{8.30}$$

$$\mathrm{Var}(G) = G''(1) + G'(1) - G'(1)^2 \tag{8.31}$$

для произвольной функции G, поскольку нам часто придется вычислять именно указанные комбинации производных.

Еще одно приятное свойство производящих функций случайных величин заключается в том, что во многих важных случаях они являются сравнительно простыми функциями от z. Например, рассмотрим *равномерное распределение* порядка n, когда случайная величина принимает каждое из значений $\{0, 1, \ldots,$

$n-1\}$ с вероятностью $1/n$. В этом случае производящая функция случайной величины равна

$$\begin{aligned} U_n(z) &= \frac{1}{n}(1+z+\cdots+z^{n-1}) = \\ &= \frac{1}{n}\frac{1-z^n}{1-z} \qquad \text{для } n \geqslant 1. \end{aligned} \tag{8.32}$$

Мы легко получили аналитический вид $U_n(z)$, потому что это простая геометрическая прогрессия.

Однако это аналитическое выражение несколько обманчиво: при подстановке $z = 1$ (самого важного значения производящей функции случайной величины) мы получим неопределенное отношение $0/0$, хотя сама функция $U_n(z)$ представляет собой полином, вполне определенный при любых значениях z. Значение $U_n(1) = 1$ очевидно из исходного выражения $(1+z+\cdots+z^{n-1})/n$, хотя, конечно же, можно прибегнуть к правилу Лопиталя и найти с его помощью $\lim_{z\to 1} U_n(z)$. Определение значения $U_n'(1)$ по правилу Лопиталя еще сложнее, потому что в знаменателе окажется множитель $(z-1)^2$; вычисление $U_n''(1)$ создаст еще бо́льшие трудности.

К счастью, есть простой выход из этого положения. Если степенной ряд $G(z) = \sum_{n\geqslant 0} g_n z^n$ сходится по крайней мере в одной точке z, такой, что $|z| > 1$, то степенной ряд $G'(z) = \sum_{n\geqslant 0} n g_n z^{n-1}$ также обладает указанным свойством, а также ряды $G''(z)$, $G'''(z)$ и т.д. Поэтому по формуле Тейлора можно записать

$$G(1+t) = G(1)+\frac{G'(1)}{1!}t+\frac{G''(1)}{2!}t^2+\frac{G'''(1)}{3!}t^3+\cdots; \tag{8.33}$$

и все производные $G(z)$ при $z = 1$ появятся в качестве коэффициентов разложения $G(1+t)$ по степеням t.

Например, таким образом легко вычислить производные для равномерной производящей функции случайной величины $U_n(z)$:

$$\begin{aligned} U_n(1+t) &= \frac{1}{n}\frac{(1+t)^n-1}{t} = \\ &= \frac{1}{n}\binom{n}{1}+\frac{1}{n}\binom{n}{2}t+\frac{1}{n}\binom{n}{3}t^2+\cdots+\frac{1}{n}\binom{n}{n}t^{n-1}. \end{aligned}$$

Сравнение с (8.33) дает

$$U_n(1) = 1; \quad U_n'(1) = \frac{n-1}{2}; \quad U_n''(1) = \frac{(n-1)(n-2)}{3}; \tag{8.34}$$

и в общем случае $U_n^{(m)}(1) = (n-1)^{\underline{m}}/(m+1)$, хотя для вычисления среднего и дисперсии нам достаточно случаев $m = 1$ и $m = 2$.

Среднее равномерного распределения есть

$$U_n'(1) = \frac{n-1}{2}, \qquad (8.35)$$

а дисперсия равна

$$\begin{aligned} U_n''(1) + U_n'(1) - U_n'(1)^2 &= \\ &= 4\,\frac{(n-1)(n-2)}{12} + 6\,\frac{(n-1)}{12} - 3\,\frac{(n-1)^2}{12} = \\ &= \frac{n^2-1}{12}. \end{aligned} \qquad (8.36)$$

Третье место среди свойств производящих функций случайных величин занимает то свойство, что произведение производящих функций случайных величин соответствует сумме независимых случайных величин. Из глав 5 и 7 мы знаем, что произведение производящих функций соответствует свертке последовательностей; однако для приложений даже более важно то, что свертка вероятностей отвечает сумме независимых случайных величин. Действительно, если X и Y представляют собой случайные величины, принимающие только целочисленные значения, то вероятность события $X+Y=n$ равна

$$\Pr(X+Y=n) = \sum_k \Pr(X=k \text{ и } Y=n-k).$$

Если X и Y независимы, получаем

$$\Pr(X+Y=n) = \sum_k \Pr(X=k)\,\Pr(Y=n-k),$$

т.е. свертку. Следовательно, у нас есть эффектная формула

$$G_{X+Y}(z) = G_X(z)\,G_Y(z), \qquad \text{если } X \text{ и } Y \text{ независимы.} \qquad (8.37)$$

Ранее в этой главе мы выяснили, что $V(X+Y) = VX + VY$, если X и Y независимы. Пусть $F(z)$ и $G(z)$ — производящие функции случайных величин X и Y и пусть $H(z)$ — производящая функция случайной величины $X+Y$. Тогда

$$H(z) = F(z)G(z),$$

и из наших формул (8.28)–(8.31) для среднего и дисперсии можно сделать вывод, что

$$\mathrm{Mean}(H) = \mathrm{Mean}(F) + \mathrm{Mean}(G); \qquad (8.38)$$

$$\mathrm{Var}(H) = \mathrm{Var}(F) + \mathrm{Var}(G). \qquad (8.39)$$

Эти формулы, являющиеся свойствами производных, $\mathrm{Mean}(H) = H'(1)$ и $\mathrm{Var}(H) = H''(1)+H'(1)-H'(1)^2$, не выполняются для произведения произвольных функций $H(z) = F(z)G(z)$; вместо этого мы имеем

$$\begin{aligned} H'(z) &= F'(z)G(z) + F(z)G'(z)\,, \\ H''(z) &= F''(z)G(z) + 2F'(z)G'(z) + F(z)G''(z)\,. \end{aligned}$$

Но если подставить $z = 1$, то можно увидеть, что (8.38) и (8.39) будут справедливы в общем случае, если выполняется условие

$$F(1) \;=\; G(1) \;=\; 1 \tag{8.40}$$

и требуемые производные существуют. Для справедливости этих формул не требуется, чтобы "вероятности" лежали в диапазоне $[0\,..\,1]$. Чтобы удовлетворить необходимому условию, можно нормализовать функции $F(z)$ и $G(z)$ путем их деления соответственно на $F(1)$ и $G(1)$, если только $F(1)$ и $G(1)$ не равны нулю.

На среднем и дисперсии история не заканчивается. Эти характеристики — всего лишь первые из бесконечного ряда так называемых *кумулянтов*, введенных датским астрономом Торвальдом Николаи Тиеле (Thorvald Nicolai Thiele) [351] в 1903 году. Первые два кумулянта случайной величины, κ_1 и κ_2, представляют собой то, что до сих пор мы называли "математическое ожидание" и "дисперсия"; вдобавок к ним имеются кумулянты и более высоких порядков, выражающие более тонкие свойства распределений. Если имеется производящая функция случайной величины $G(z)$, то кумулянты всех порядков выражаются с помощью формулы

$$\ln G(e^t) \;=\; \frac{\kappa_1}{1!}t + \frac{\kappa_2}{2!}t^2 + \frac{\kappa_3}{3!}t^3 + \frac{\kappa_4}{4!}t^4 + \cdots\,. \tag{8.41}$$

Давайте познакомимся с кумулянтами поближе. Если $G(z)$ представляет собой производящую функцию случайной величины X, то

$$\begin{aligned} G(e^t) \;=\; \sum_{k\geqslant 0}\Pr(X=k)e^{kt} \;&=\; \sum_{k,m\geqslant 0}\Pr(X=k)\frac{k^m t^m}{m!} \;= \\ &=\; 1+\frac{\mu_1}{1!}t+\frac{\mu_2}{2!}t^2+\frac{\mu_3}{3!}t^3+\cdots\,, \end{aligned} \tag{8.42}$$

где

$$\mu_m \;=\; \sum_{k\geqslant 0} k^m \Pr(X=k) \;=\; E(X^m)\,. \tag{8.43}$$

Эта величина — μ_m — называется m-м моментом X. Можно взять экспоненту от обеих частей (8.41) и получить другую формулу для $G(e^t)$:

$$\begin{aligned} G(e^t) &= 1+\frac{(\kappa_1 t+\frac{1}{2}\kappa_2 t^2+\cdots)}{1!}+\frac{(\kappa_1 t+\frac{1}{2}\kappa_2 t^2+\cdots)^2}{2!}+\cdots = \\ &= 1+\kappa_1 t+\tfrac{1}{2}(\kappa_2+\kappa_1^2)t^2+\cdots . \end{aligned}$$

Приравнивание коэффициентов при одинаковых степенях t приводит к следующему ряду формул

$$\begin{aligned} \kappa_1 &= \mu_1 , && (8.44)\\ \kappa_2 &= \mu_2-\mu_1^2 , && (8.45)\\ \kappa_3 &= \mu_3-3\mu_1\mu_2+2\mu_1^3 , && (8.46)\\ \kappa_4 &= \mu_4-4\mu_1\mu_3+12\mu_1^2\mu_2-3\mu_2^2-6\mu_1^4 , && (8.47)\\ \kappa_5 &= \mu_5-5\mu_1\mu_4+20\mu_1^2\mu_3-10\mu_2\mu_3+ \\ &\qquad +30\mu_1\mu_2^2-60\mu_1^3\mu_2+24\mu_1^5 , && (8.48)\\ &\vdots \end{aligned}$$

которые определяют кумулянты через моменты. Обратите внимание, что κ_2 действительно представляет собой дисперсию $E(X^2)-(EX)^2$, как и утверждалось выше.

Из уравнения (8.41) ясно, что кумулянты, порождаемые произведением $F(z)G(z)$ двух производящих функций случайных величин, будут суммами соответствующих кумулянтов $F(z)$ и $G(z)$, потому что логарифм произведения представляет собой сумму логарифмов. Следовательно, все кумулянты суммы независимых случайных величин аддитивны, так же, как математическое ожидание и дисперсия. Это свойство делает кумулянты более важными, чем моменты.

"Для этих семиинвариантов высших порядков мы не предлагаем никаких специальных имен".
— Т. Н. Тиеле [351]

Если пойти несколько иным путем и записать

$$G(1+t) = 1+\frac{\alpha_1}{1!}t+\frac{\alpha_2}{2!}t^2+\frac{\alpha_3}{3!}t^3+\cdots ,$$

то уравнение (8.33) показывает, что числа α — это "факториальные моменты"

$$\begin{aligned} \alpha_m &= G^{(m)}(1) \\ &= \sum_{k\geqslant 0}\Pr(X=k)k^{\underline{m}}z^{k-m}\Big|_{z=1} = \\ &= \sum_{k\geqslant 0}k^{\underline{m}}\Pr(X=k) = E(X^{\underline{m}}) . \qquad (8.49) \end{aligned}$$

Отсюда вытекает, что

$$\begin{aligned}
G(e^t) &= 1+\frac{\alpha_1}{1!}(e^t-1)+\frac{\alpha_2}{2!}(e^t-1)^2+\cdots = \\
&= 1+\frac{\alpha_1}{1!}(t+\tfrac12 t^2+\cdots)+\frac{\alpha_2}{2!}(t^2+t^3+\cdots)+\cdots = \\
&= 1+\alpha_1 t+\tfrac12(\alpha_2+\alpha_1)t^2+\cdots,
\end{aligned}$$

и мы можем выразить кумулянты через производные $G^{(m)}(1)$:

$$\begin{aligned}
\kappa_1 &= \alpha_1\,, && (8.50)\\
\kappa_2 &= \alpha_2+\alpha_1-\alpha_1^2\,, && (8.51)\\
\kappa_3 &= \alpha_3+3\alpha_2+\alpha_1-3\alpha_2\alpha_1-3\alpha_1^2+2\alpha_1^3\,, && (8.52)\\
&\ \vdots
\end{aligned}$$

Эта последовательность формул порождает "аддитивные" тождества, которые распространяют (8.38) и (8.39) на все кумулянты.

А теперь спустимся с небес на землю и применим рассмотренные идеи к простым примерам. Простейшей случайной величиной является "случайная константа": случайная величина X принимает определенное фиксированное значение x с вероятностью 1. В этом случае $G_X(z) = z^x$ и $\ln G_X(e^t) = xt$; следовательно, математическое ожидание равно x, а все остальные кумулянты равны нулю. Отсюда следует, что операция умножения любой производящей функции случайной величины на z^x увеличивает математическое ожидание на x, но при этом оставляет неизменными дисперсию и другие кумулянты.

Как применить производящие функции к игральным костям? Распределению числа очков на одной игральной кости соответствует производящая функция случайной величины

$$G(z) = \frac{z+z^2+z^3+z^4+z^5+z^6}{6} = zU_6(z)\,,$$

где U_6 представляет собой производящую функцию случайной величины для равномерного распределения порядка 6. Множитель 'z' добавляет к математическому ожиданию 1, так что оно оказывается равным 3.5 вместо $\frac{n-1}{2} = 2.5$, которое дает (8.35); однако это 'z' не влияет на дисперсию (8.36), которая равна $\frac{35}{12}$.

Производящая функция случайной величины суммы очков на двух независимых игральных костях представляет собой квадрат производящей функции случайной величины для одной игральной кости, т.е.

$$\begin{aligned}
G_S(z) &= \frac{z^2+2z^3+3z^4+4z^5+5z^6+6z^7+5z^8+4z^9+3z^{10}+2z^{11}+z^{12}}{36} = \\
&= z^2U_6(z)^2\,.
\end{aligned}$$

При бросании пары игральных костей n раз вероятность получения k очков аналогично равна

$$[z^k]\, G_S(z)^n \;=\; [z^k]\, z^{2n} U_6(z)^{2n} \;=\; [z^{k-2n}]\, U_6(z)^{2n}\,.$$

Распределение шляп представляет собой разновидность равномерного распределения.

В рассмотренной ранее задаче о подбрасывании шляп футбольными болельщиками, известной также как задача подсчета неподвижных точек случайной перестановки, производящая функция случайной величины известна из (5.49):

$$F_n(z) \;=\; \sum_{0\leqslant k\leqslant n} \frac{(n-k)¡}{(n-k)!}\,\frac{z^k}{k!} \qquad \text{для } n \geqslant 0. \tag{8.53}$$

Следовательно,

$$\begin{aligned} F_n'(z) \;&=\; \sum_{1\leqslant k\leqslant n} \frac{(n-k)¡}{(n-k)!}\,\frac{z^{k-1}}{(k-1)!} \;=\\ &=\; \sum_{0\leqslant k\leqslant n-1} \frac{(n-1-k)¡}{(n-1-k)!}\,\frac{z^k}{k!} \;=\; F_{n-1}(z)\,. \end{aligned}$$

Даже не зная никаких подробностей о коэффициентах, из рекуррентного соотношения $F_n'(z) = F_{n-1}(z)$ можно заключить, что $F_n^{(m)}(z) = F_{n-m}(z)$; следовательно,

$$F_n^{(m)}(1) \;=\; F_{n-m}(1) \;=\; [n \geqslant m]\,. \tag{8.54}$$

Эта формула упрощает вычисление математического ожидания и дисперсии; мы находим, как и ранее (но быстрее), что обе эти характеристики равны 1 при $n \geqslant 2$.

На самом деле сейчас мы можем показать, что и m-й кумулянт κ_m этой случайной величины равен 1 при $n \geqslant m$. Действительно, m-й кумулянт зависит только от $F_n'(1)$, $F_n''(1)$, …, $F_n^{(m)}(1)$, а все они равны 1; следовательно, для m-го кумулянта ответ не изменится, если заменить $F_n(z)$ предельной производящей функцией случайной величины

$$F_\infty(z) \;=\; e^{z-1}\,, \tag{8.55}$$

такой, что $F_\infty^{(m)}(1) = 1$ для производных всех порядков. Кумулянты F_∞ тождественно равны 1, потому что

$$\ln F_\infty(e^t) \;=\; \ln e^{e^t-1} \;=\; e^t - 1 \;=\; \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\,.$$

8.4 Бросание монет

Обратимся теперь к процессам, имеющим ровно два исхода. Если мы бросаем монету, то имеется некоторая вероятность p, что монета упадет вверх орлом, и вероятность q, что она упадет вверх решкой, причем

$$p+q = 1.$$

Мошенники знают, что $p \approx 0.1$, если закрутить новое американское пенни на гладком столе (распределение масс в этой монете таково, что голова Линкольна перетягивает и падает вниз.)

(Мы считаем, что монета не останется стоять на ребре, не провалится в щель, не зависнет в воздухе и т.п.) На протяжении этого раздела числа p и q всегда дают в сумме 1. Если монета *правильная*, то $p = q = \frac{1}{2}$; в противном случае монета называется *несимметричной*.

Производящая функция случайной величины для числа выпадений орла в результате одного бросания монеты есть

$$H(z) = q+pz. \tag{8.56}$$

Если монета брошена n раз, в предположении, что все бросания монеты независимы, количество выпавших орлов порождается функцией

$$H(z)^n = (q+pz)^n = \sum_{k\geqslant 0} \binom{n}{k} p^k q^{n-k} z^k, \tag{8.57}$$

в соответствии с биномиальной теоремой. Таким образом, вероятность выбросить ровно k орлов в n бросаниях равна $\binom{n}{k}p^k q^{n-k}$. Эта последовательность вероятностей называется *биномиальным распределением*.

Предположим, что мы подбрасываем монету до тех пор, пока впервые не выпадет орел. Какова вероятность, что нам потребуется ровно k бросаний? С вероятностью p мы получим $k = 1$ (поскольку это вероятность того, что в первом броске выпадет орел); вероятность события $k = 2$ равна qp (поскольку это вероятность того, что в первый раз выпадет решка, а затем орел); а для произвольного k вероятность равна $q^{k-1}p$. Таким образом, производящая функция есть

$$pz + qpz^2 + q^2pz^3 + \cdots = \frac{pz}{1-qz}. \tag{8.58}$$

Повторение этого процесса до выпадения n орлов дает производящую функцию случайной величины

$$\begin{aligned}\left(\frac{pz}{1-qz}\right)^n &= p^n z^n \sum_k \binom{n+k-1}{k} (qz)^k = \\ &= \sum_k \binom{k-1}{k-n} p^n q^{k-n} z^k. \end{aligned} \tag{8.59}$$

Это, кстати, умноженная на z^n функция

$$\left(\frac{p}{1-qz}\right)^n = \sum_k \binom{n+k-1}{k} p^n q^k z^k, \qquad (8.60)$$

которая является производящей для *отрицательного биномиального распределения.*

Пространство вероятностей в примере (8.59), в котором мы бросаем монету до выпадения n орлов, отличается от пространств вероятностей, с которыми мы встречались ранее в этой главе, потому что оно содержит бесконечное количество элементов. Каждый элемент представляет собой конечную последовательность орлов и решек, содержащую ровно n орлов и заканчивающуюся орлом; вероятность такой последовательности равна $p^n q^{k-n}$, где $k-n$ — количество решек. Таким образом, если записывать О при выпадении орла и Р при выпадении решки, то, например, при $n = 3$ последовательность РОРРРОО является элементом рассматриваемого пространства вероятностей, а его вероятность равна $qpqqqpp = p^3q^4$.

Орел — я выигрываю, решка — ты проигрываешь.

Нет? Ладно: решка — ты проигрываешь, орел — я выигрываю.

Опять не так? Ну ладно, орел — ты проигрываешь, решка — я выигрываю.

Пусть X — случайная величина с биномиальным распределением (8.57), а Y — случайная величина с отрицательным биномиальным распределением (8.60). Эти распределения зависят от параметров n и p. Математическое ожидание X равно $nH'(1) = np$, поскольку ее производящая функция случайной величины есть $H(z)^n$; дисперсия же равна

$$n\big(H''(1) + H'(1) - H'(1)^2\big) = n(0 + p - p^2) = npq. \qquad (8.61)$$

Таким образом, стандартное отклонение составляет $\sqrt{npq}$: если подбросить монету n раз, следует ожидать выпадения орла примерно $np \pm \sqrt{npq}$ раз. Математическое ожидание и дисперсию Y можно найти аналогично. Положим

$$G(z) = \frac{p}{1-qz}.$$

Тогда имеем

$$G'(z) = \frac{pq}{(1-qz)^2},$$
$$G''(z) = \frac{2pq^2}{(1-qz)^3};$$

следовательно, $G'(1) = pq/p^2 = q/p$ и $G''(1) = 2pq^2/p^3 = 2q^2/p^2$. Поэтому математическое ожидание Y равно nq/p, а дисперсия — nq/p^2.

Более простой способ получения среднего и дисперсии Y заключается в использовании обратной производящей функции

$$F(z) \;=\; \frac{1-qz}{p} \;=\; \frac{1}{p} - \frac{q}{p}z \qquad (8.62)$$

и записи

$$G(z)^n \;=\; F(z)^{-n}\,. \qquad (8.63)$$

Этот полином $F(z)$ не является производящей функцией, потому что у него есть отрицательный коэффициент. Однако он удовлетворяет главному свойству — $F(1) = 1$. Таким образом, формально $F(z)$ представляет собой бином, отвечающей монете, у которой "вероятность" выпадения орла равна $-q/p$; а $G(z)$ формально соответствует выбрасыванию такой монеты -1 раз(!). Таким образом, отрицательное биномиальное распределение с параметрами (n,p) можно рассматривать как обычное биномиальное распределение с параметрами $(n',p') = (-n,-q/p)$. Действуя формально, получаем, что математическое ожидание должно быть равно $n'p' = (-n)(-q/p) = nq/p$, а дисперсия должна равняться $n'p'q' = (-n)(-q/p)(1+q/p) = nq/p^2$. Этот формальный вывод, включающий отрицательные вероятности, тем не менее, справедлив, поскольку наши предыдущие рассуждения для обычных биномиальных распределений были основаны на тождествах для формальных степенных рядов, в которых нигде не использовалось предположение $0 \leqslant p \leqslant 1$.

Вероятность того, что я становлюсь моложе, явно отрицательна...

Ой! Но тогда вероятность того, что я старею, больше единицы?!

Перейдем к другому примеру: сколько раз надо выбросить монету, чтобы орел выпал два раза подряд? Пространство вероятностей теперь состоит из всех последовательностей О и Р, заканчивающихся на ОО, в которых нет двух последовательных О до последней позиции:

$$\Omega \;=\; \{\text{ОО}, \text{РОО}, \text{РРОО}, \text{ОРОО}, \text{РРРОО}, \text{РОРОО}, \text{ОРРОО}, \ldots\}\,.$$

Вероятность любой заданной последовательности можно получить, заменяя О на p, а Р на q; например, последовательность РОРОО встречается с вероятностью

$$\Pr(\text{РОРОО}) \;=\; qpqpp \;=\; p^3q^2\,.$$

Теперь можно поиграться с производящими функциями, как мы делали это в начале главы 7, обозначив через S бесконечную сумму

$$S \;=\; \text{ОО} + \text{РОО} + \text{РРОО} + \text{ОРОО} + \text{РРРОО} + \text{РОРОО} + \text{ОРРОО} + \cdots$$

всех элементов Ω. Если заменить каждое О на pz, а каждое Р на qz, то можно получить производящую функцию случайной величины для количества бросаний до появления двух последовательных орлов.

Имеется довольно неожиданная связь между S и суммой всех покрытий костями домино

$$\mathsf{T} = \text{|} + ▯ + ▯▯ + ⊟ + ▯▯▯ + ▯⊟ + ⊟▯ + \cdots$$

в уравнении (7.1). Оказывается, что можно получить S из Т, если заменить каждую кость домино ▯ на Р, а каждую кость домино ⊟ на ОР, и затем приписать ОО в конце. Это соответствие легко доказать, потому что каждый элемент Ω имеет вид $(\mathrm{P}+\mathrm{OP})^n\mathrm{OO}$ для некоторого $n \geqslant 0$, а каждый член Т имеет вид $(▯+⊟)^n$. Следовательно, из (7.4) имеем

$$\mathrm{S} = (1-\mathrm{P}-\mathrm{OP})^{-1}\mathrm{OO},$$

а производящая функция случайной величины для нашей задачи равна

$$G(z) = \bigl(1 - qz - (pz)(qz)\bigr)^{-1}(pz)^2 = \\ = \frac{p^2z^2}{1-qz-pqz^2}\,. \qquad (8.64)$$

Наш опыт работы с отрицательным биномиальным распределением подсказывает, что проще всего подсчитать математическое ожидание и дисперсию функции (8.64), если записать

$$G(z) = \frac{z^2}{F(z)},$$

где

$$F(z) = \frac{1-qz-pqz^2}{p^2},$$

и подсчитать "математическое ожидание" и "дисперсию" этой псевдо-производящей функции случайной величины $F(z)$. (Эта вновь введенная функция обладает свойством $F(1) = 1$.) Имеем

$$F'(1) = (-q-2pq)/p^2 = 2 - p^{-1} - p^{-2};$$
$$F''(1) = -2pq/p^2 = 2 - 2p^{-1}.$$

Таким образом, поскольку $z^2 = F(z)G(z)$, $\mathrm{Mean}(z^2) = 2$ и $\mathrm{Var}(z^2) = 0$, математическое ожидание и дисперсия распределения $G(z)$ таковы:

$$\mathrm{Mean}(G) = 2 - \mathrm{Mean}(F) = p^{-2} + p^{-1}; \tag{8.65}$$

$$\mathrm{Var}(G) = -\mathrm{Var}(F) = p^{-4} + 2p^{-3} - 2p^{-2} - p^{-1}. \tag{8.66}$$

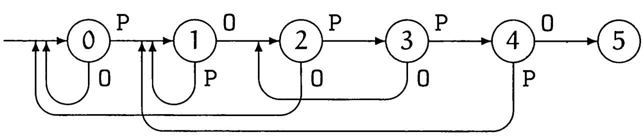
Для $p = \frac{1}{2}$ математическое ожидание и дисперсия составляют соответственно 6 и 22. (В упр. 4 обсуждается вычисление средних и дисперсий путем вычитания.)

Представим себе более хитроумный эксперимент: будем бросать монету до тех пор, пока не встретится последовательность РОРРО. Теперь сумма выигрышных позиций имеет вид

$$\begin{aligned} S = {} & \mathrm{POPPO} + \mathrm{OPOPPO} + \mathrm{PPOPPO} \\ & + \mathrm{OOPOPPO} + \mathrm{OPPOPPO} + \mathrm{POPOPPO} + \mathrm{PPPOPPO} + \cdots; \end{aligned}$$

описать эту сумму сложнее, чем предыдущую. Возвращаясь к методам, использовавшимся при решении задачи о домино в главе 7, можно получить формулу для S, рассматривая множество слагаемых как "язык конечных состояний", определяемый следующим "конечным автоматом":

"Вы просто автомат, вычислительная машина! — воскликнул я. — Временами в вас есть что-то нечеловеческое."
— Дж. Х. Ватсон (J. H. Watson) [83]



Элементарными событиями в пространстве вероятности являются такие последовательности символов О и Р, которые переводят автомат из состояния 0 в состояние 5. Предположим, например, что мы только что выбросили РОР; тогда мы окажемся в состоянии 3. Если теперь выпадет решка, мы перейдем в состояние 4; если выпадет орел — в состояние 2 (но не в начальное состояние 0, поскольку последние РО могут быть продолжены последовательностью РРО).

В этой постановке задачи можно обозначить через S_k сумму всех последовательностей О и Р, которые приводят в состояние k; тогда

$$\begin{aligned} S_0 &= 1 + S_0\,\mathrm{O} + S_2\,\mathrm{O}, \\ S_1 &= S_0\,\mathrm{P} + S_1\,\mathrm{P} + S_4\,\mathrm{P}, \\ S_2 &= S_1\,\mathrm{O} + S_3\,\mathrm{O}, \\ S_3 &= S_2\,\mathrm{P}, \\ S_4 &= S_3\,\mathrm{P}, \\ S_5 &= S_4\,\mathrm{O}. \end{aligned}$$

Теперь искомая сумма S — это сумма S_5; найти ее можно, решая выписанную систему из шести уравнений с шестью неизвестными S_0, S_1, …, S_5. Заменяя O на pz, а P на qz, получим производящую функцию, причем коэффициент при z^n в S_k представляет собой вероятность того, что мы окажемся в состоянии k после n бросков.

Подобным образом любая диаграмма переходов между состояниями, в которой переход из состояния j в состояние k происходит с заданной вероятностью $p_{j,k}$, дает систему линейных уравнений, решениями которой являются производящие функции для вероятностей состояний после n переходов. Объекты такого типа называются *марковскими процессами*, и теория их поведения тесно связана с теорией линейных уравнений.

Однако задачу о бросании монеты можно решить гораздо проще, не прибегая к сложной общей процедуре с построением конечного автомата. Вместо шести уравнений с шестью неизвестными S_0, S_1, …, S_5 для полного описания S достаточно всего двух уравнений с двумя неизвестными. Трюк состоит в том, чтобы ввести в рассмотрение вспомогательную сумму $N = S_0 + S_1 + S_2 + S_3 + S_4$ всех последовательностей результатов бросания, не содержащих заданную последовательность POPPO:

$$N \;=\; 1 + \mathrm{O} + \mathrm{P} + \mathrm{OO} + \cdots + \mathrm{POPOP} + \mathrm{POPPP} + \cdots .$$

Имеем

$$1 + N(\mathrm{O} + \mathrm{P}) \;=\; N + S\,, \qquad (8.67)$$

потому что любой член в левой части либо заканчивается на POPPO (и принадлежит S), либо нет (и принадлежит N); обратно, любое слагаемое из правой части либо пустое, либо принадлежит $N\,\mathrm{O}$ или $N\,\mathrm{P}$. Кроме того, мы имеем еще одно важное уравнение

$$N\,\mathrm{POPPO} \;=\; S + S\,\mathrm{PPO}\,, \qquad (8.68)$$

потому что любое слагаемое из левой части содержит в себе член S, который либо оканчивается первой O из добавленной к N последовательности POPPO (и в этом случае за ним следует PPO), либо заканчивается последней O. Любое слагаемое из правой части, в свою очередь, имеется в левой.

Эта система легко решается. Из соотношения (8.67) получаем $N = (1 - S)(1 - \mathrm{O} - \mathrm{P})^{-1}$, следовательно,

$$(1 - S)(1 - \mathrm{P} - \mathrm{O})^{-1}\,\mathrm{POPPO} \;=\; S(1 + \mathrm{PPO})\,.$$

Как и ранее, для получения производящей функции случайной величины $G(z)$ для количества бросаний монеты следует заменить O на pz и P на qz. После упрощений, вытекающих из тождества $p+q=1$, получаем

$$\frac{(1-G(z))\,p^2q^3z^5}{1-z} = G(z)(1+pq^2z^3);$$

следовательно, решение имеет вид

$$G(z) = \frac{p^2q^3z^5}{p^2q^3z^5+(1+pq^2z^3)(1-z)}. \qquad (8.69)$$

Обратите внимание, что $G(1)=1$, если $pq \neq 0$; мы с вероятностью 1 дождемся последовательности OPPO, если, конечно, монета не настолько несимметричная, что всегда выпадает только одной стороной — только орлом или только решкой.

Для вычисления математического ожидания и дисперсии распределения (8.69) инвертируем $G(z)$, как мы уже делали это раньше, и запишем $G(z)=z^5/F(z)$, где F — полином:

$$F(z) = \frac{p^2q^3z^5+(1+pq^2z^3)(1-z)}{p^2q^3}. \qquad (8.70)$$

Соответствующие производные равны

$$\begin{aligned} F'(1) &= 5-(1+pq^2)/p^2q^3\,, \\ F''(1) &= 20-6pq^2/p^2q^3\,; \end{aligned}$$

и если X — требуемое число бросаний, то мы получаем

$$\begin{aligned} EX = \operatorname{Mean}(G) &= 5\text{—}\operatorname{Mean}(F) = p^{-2}q^{-3}+p^{-1}q^{-1}\,; && (8.71)\\ VX = \operatorname{Var}(G) &= -\operatorname{Var}(F) = \\ &= -25+p^{-2}q^{-3}+7p^{-1}q^{-1}+\operatorname{Mean}(F)^2 = \\ &= (EX)^2-9p^{-2}q^{-3}-3p^{-1}q^{-1}\,. && (8.72) \end{aligned}$$

Для $p=\frac{1}{2}$ математическое ожидание и дисперсия равны 36 и 996.

Попробуем действовать более общо. Уже решенная задача в достаточной степени "случайна", чтобы показать, как следует анализировать случай появления *произвольной* последовательности орлов и решек A. Вновь обозначим через S сумму всех выигрышных последовательностей O и P, а через N — сумму всех последовательностей, в которых подпоследовательность A пока что не встречалась. Уравнение (8.67) остается тем же самым; уравнение (8.68) превращается в

$$\begin{aligned} NA = S\bigl(1 &+ A^{(1)}\,[A^{(m-1)}=A_{(m-1)}] + A^{(2)}\,[A^{(m-2)}=A_{(m-2)}] \\ &+ \cdots + A^{(m-1)}\,[A^{(1)}=A_{(1)}]\bigr)\,, \end{aligned} \qquad (8.73)$$

где m — длина A и где $A^{(k)}$ и $A_{(k)}$ обозначают соответственно последние k и первые k символов A. Например, если A представляет собой уже рассмотренную последовательность РОРРО, то

$$\begin{aligned} A^{(1)} &= \mathrm{O}, & A^{(2)} &= \mathrm{PO}, & A^{(3)} &= \mathrm{PPO}, & A^{(4)} &= \mathrm{OPPO};\\ A_{(1)} &= \mathrm{P}, & A_{(2)} &= \mathrm{PO}, & A_{(3)} &= \mathrm{POP}, & A_{(4)} &= \mathrm{POPP}. \end{aligned}$$

Поскольку единственное точное совпадение $A^{(2)} = A_{(2)}$, уравнение (8.73) сводится к (8.68).

Обозначим через $\widetilde{A}$ результат подстановки p^{-1} вместо О и q^{-1} вместо Р в последовательность A. Тогда нетрудно обобщить наш вывод уравнений (8.71) и (8.72) и получить математическое ожидание и дисперсию в общем случае (упр. 20):

$$EX = \sum_{k=1}^{m} \widetilde{A}_{(k)}\,[A^{(k)} = A_{(k)}]; \tag{8.74}$$

$$VX = (EX)^2 - \sum_{k=1}^{m} (2k-1)\widetilde{A}_{(k)}\,[A^{(k)} = A_{(k)}]. \tag{8.75}$$

В частном случае $p = \frac{1}{2}$ имеется особенно простая интерпретация этих формул. Если дана последовательность A из m орлов и решек, то положим

$$A{:}A = \sum_{k=1}^{m} 2^{k-1}\,[A^{(k)} = A_{(k)}]. \tag{8.76}$$

Очень легко выписать двоичное представление данного числа. Для этого надо записать '1' под каждой позицией, обладающей тем свойством, что данная последовательность, если ее начало сдвинуть в эту позицию, в точности совпадет с первоначальной последовательностью в их общей части:

```
A =   OPOPOOPOPO
A:A = (1000010101)_2 = 512 + 16 + 4 + 1 = 533
      OPOPOOPOPO    √
       OPOPOOPOPO
        OPOPOOPOPO
         OPOPOOPOPO
          OPOPOOPOPO
           OPOPOOPOPO    √
            OPOPOOPOPO
             OPOPOOPOPO    √
              OPOPOOPOPO
               OPOPOOPOPO    √
```

Уравнение (8.74) утверждает, в сущности, что ожидаемое число бросаний до появления подпоследовательности A есть в точности $2(A{:}A)$ для правильной монеты, поскольку $\widetilde{A}_{(k)} = 2^k$ для $p = q = \frac{1}{2}$. Этот результат, найденный советским математиком А. Д. Соловьевым в 1966 году [331], на первый взгляд выглядит парадоксальным: несамосовмещающиеся последовательности появляются раньше, чем самосовмещающиеся! Появления последовательности OOOOO придется ждать почти вдвое дольше, чем последовательности OOOOP или POOOO.

"Чем больше периодов у нашего слова, тем позже оно появляется".
— А. Д. Соловьев

Давайте теперь рассмотрим забавную игру, которую придумал Уолтер Пенни (Walter Penney) [289] в 1969 году. Алиса и Билл бросают монету до тех пор, пока не встретится последовательность OOP или OPP; Алиса выигрывает, если первой встречается последовательность OOP, Билл побеждает, если первой выпадет последовательность OPP. Эта игра — сейчас ее называют "Penney ante", — определенно, выглядит справедливой при использовании правильной монеты: ведь последовательности OOP и OPP имеют одинаковые характеристики, если рассматривать каждую из них по отдельности. Производящая функция случайной величины для времени ожидания последовательности OOP имеет вид

$$G(z) = \frac{z^3}{z^3 - 8(z-1)},$$

и точно такая же производящая функция случайной величины соответствует последовательности OPP. Следовательно, ни Алиса, ни Билл не имеют никакого преимущества при игре в одиночку.

Конечно, нет! Перед кем они могут иметь преимущество, играя в одиночестве?

Но если рассматривать обе подпоследовательности вместе, то между ними возникает крайне интересное взаимодействие. Обозначим через S_A сумму конфигураций, выигрышных для Алисы, а через S_B — сумму конфигураций, выигрышных для Билла:

$$\begin{aligned} S_A &= \mathrm{OOP} + \mathrm{OOOP} + \mathrm{POOP} + \mathrm{OOOOP} + \mathrm{OPOOP} + \mathrm{POOOP} + \cdots; \\ S_B &= \mathrm{OPP} + \mathrm{POPP} + \mathrm{OPOPP} + \mathrm{PPOPP} + \mathrm{POPOPP} + \mathrm{PPPOPP} + \cdots. \end{aligned}$$

Вспоминая наш трюк, использованный в случае одной подпоследовательности, снова обозначим через N сумму всех последовательностей, для которых пока ни один из игроков не выиграл:

$$N = 1+\mathrm{O}+\mathrm{P}+\mathrm{OO}+\mathrm{OP}+\mathrm{PO}+\mathrm{PP}+\mathrm{OOO}+\mathrm{OPO}+\mathrm{POO}+\cdots. \qquad (8.77)$$

Можно легко убедиться в справедливости следующих уравнений:

$$\begin{aligned} 1+N(O+P) &= N+S_A+S_B\,; \\ N\,OOP &= S_A\,; \\ N\,OPP &= S_A\,P+S_B\,. \end{aligned} \tag{8.78}$$

Если теперь положить $O = P = \frac{1}{2}$, то получающееся значение S_A будет вероятностью выигрыша Алисы, а S_B — вероятностью выигрыша Билла. Наши три уравнения сводятся к следующим:

$$1+N \;=\; N+S_A+S_B\,; \quad \tfrac{1}{8}N \;=\; S_A\,; \quad \tfrac{1}{8}N \;=\; \tfrac{1}{2}S_A+S_B\,;$$

и мы находим $S_A = \frac{2}{3}$, $S_B = \frac{1}{3}$. Алиса побеждает в два раза чаще Билла!

В обобщенной игре Алиса и Билл выбирают последовательности орлов и решек A и B и бросают монету до тех пор, пока в последовательности результатов выбрасываний не встретится одна из подпоследовательностей A или B. Эти две подпоследовательности не обязательно одинаковой длины, однако мы будем считать, что A не входит в B и B не входит в A. (В противном случае игра вырождается. Например, если A = OP и B = POPO, то у бедного Билла нет никаких шансов на победу; а если A = OPO и B = PO, то игра может закончиться вничью.) Тогда можно записать три уравнения, аналогичные (8.73) и (8.78):

$$\begin{aligned} 1+N(O+P) &= N+S_A+S_B\,; \\ NA &= S_A\sum_{k=1}^{l} A^{(l-k)}\,[A^{(k)}=A_{(k)}]+ \\ &\qquad + S_B\sum_{k=1}^{\min(l,m)} A^{(l-k)}\,[B^{(k)}=A_{(k)}]\,; \\ NB &= S_A\sum_{k=1}^{\min(l,m)} B^{(m-k)}\,[A^{(k)}=B_{(k)}]+ \\ &\qquad + S_B\sum_{k=1}^{m} B^{(m-k)}\,[B^{(k)}=B_{(k)}]\,. \end{aligned} \tag{8.79}$$

Здесь l — длина A, а m — длина B. Например, если A=OPPOPOPO и B = POPOPPO, то два последних уравнения, зависящие от A и B, будут иметь вид

$$\begin{aligned} N\,OPPOPOPO &= S_A\,PPOPOPO+S_A+S_B\,PPOPOPO+S_B\,POPO\,; \\ N\,POPOPPO &= S_A\,POPPO+S_A\,PPO+S_B\,POPPO+S_B\,. \end{aligned}$$

Для выяснения вероятности победы при игре правильной монетой следует положить $O = P = \frac{1}{2}$; это приведет два решающих уравнения к виду

$$\begin{aligned} N &= S_A \sum_{k=1}^{l} 2^k [A^{(k)} = A_{(k)}] + S_B \sum_{k=1}^{\min(l,m)} 2^k [B^{(k)} = A_{(k)}]\,; \\ N &= S_A \sum_{k=1}^{\min(l,m)} 2^k [A^{(k)} = B_{(k)}] + S_B \sum_{k=1}^{m} 2^k [B^{(k)} = B_{(k)}]\,. \end{aligned} \tag{8.80}$$

Для дальнейшей работы нам надо обобщить операцию $A{:}A$ из (8.76) на случай двух независимых последовательностей A и B:

$$A{:}B = \sum_{k=1}^{\min(l,m)} 2^{k-1} [A^{(k)} = B_{(k)}]\,. \tag{8.81}$$

Уравнения (8.80) упрощаются до

$$S_A(A{:}A) + S_B(B{:}A) = S_A(A{:}B) + S_B(B{:}B)\,;$$

преимущество Алисы выражается отношением

$$\frac{S_A}{S_B} = \frac{B{:}B - B{:}A}{A{:}A - A{:}B}\,. \tag{8.82}$$

(Эту красивую формулу открыл Джон Хортон Конвей (John Horton Conway) [137].)

Например, если, как и ранее, $A = \text{ОРРОРОРО}$ и $B = \text{РОРОРРО}$, то $A{:}A = (10000001)_2 = 129$, $A{:}B = (0001010)_2 = 10$, $B{:}A = (0001001)_2 = 9$ и $B{:}B = (1000010)_2 = 66$; так что отношение S_A/S_B равно $(66-9)/(129-10) = 57/119$. Алиса будет выигрывать в среднем только 57 раз из каждых 176.

В игре Пенни встречаются всякие странности. Например, последовательность ООРО выигрывает у последовательности ОРОО с преимуществом $3/2$, а последовательность ОРОО выигрывает у последовательности РООО с преимуществом $7/5$. Казалось бы, последовательность ООРО должна быть существенно лучше, чем последовательность РООО. Но на самом деле последовательность РООО выигрывает у последовательности ООРО с преимуществом $7/5$! Отношение последовательностей в этой игре нетранзитивно. В упр. 57 доказывается, что, какую бы последовательность $\tau_1\tau_2\ldots\tau_l$ длиной $l \geqslant 3$ ни выбрала Алиса, Билл всегда может повысить свои шансы на победу, выбирая последовательность $\overline{\tau}_2\tau_1\tau_2\ldots\tau_{l-1}$, где $\overline{\tau}_2$ — противоположный к τ_2 символ (орел, если τ_2 — решка, и решка, если τ_2 — орел).

СтРаннО и даже ОРигинальнО...

8.5 Хеширование

"...хотя глагол «хешировать» (to hash) стал стандартным термином в средине 60-х годов, никто не решался использовать в печати столь недостойное слово до 1967 года!"
—Д. Э. Кнут [209]

Завершим эту главу приложением теории вероятностей к программированию. Некоторые важные алгоритмы хранения и выборки данных в компьютерах основаны на методе, носящем название "хеширование". Общая задача заключается в хранении некоторого множества записей, каждая из которых содержит значение "ключа" K и некоторые данные $D(K)$, связанные с этим ключом. Мы хотим иметь возможность быстро находить $D(K)$ по заданному K. Например, каждый ключ может представлять собой имя студента, а связанные с ним данные — оценки за домашние задания.

На практике компьютеры не настолько оснащены памятью, чтобы можно было выделить по отдельной ячейке для каждого возможного ключа. Возможны миллиарды ключей, но в каждом конкретном приложении их присутствует не так уж и много. Одно из решений проблемы заключается в том, чтобы поддерживать две таблицы — KEY[j] (для ключей) и DATA[j] (для данных), $1 \leqslant j \leqslant N$, где N — общее количество записей, которые могут быть размещены в памяти; еще одна переменная n указывает, сколько записей размещено на самом деле. Тогда поиск заданного ключа K можно выполнять при помощи очевидного последовательного сканирования.

S1 Установить $j := 1$. (Уже просмотрены все позиции $< j$.)
S2 Если $j > n$, остановиться. (Неудачный поиск.)
S3 Если KEY[j] $= K$, остановиться. (Поиск успешно завершен.)
S4 Увеличить j на 1 и вернуться к шагу S2. (Пробуем дальше.)

После успешного поиска нужная запись с данными $D(K)$ находится в DATA[j]. После неудачного поиска можно вставить K и $D(K)$ в таблицу путем установок

$$n := j, \quad \mathtt{KEY}[n] := K, \quad \mathtt{DATA}[n] := D(K)$$

в предположении, что таблица еще не заполнена до предела.

Этот метод работает правильно, но может быть ужасающе медленным. Так, при неудачном поиске шаг S2 приходится выполнять $n+1$ раз, а n может быть очень большим.

Хеширование было разработано для того, чтобы ускорить поиск. Основная его идея, в одном из популярных вариантов, заключается в использовании m отдельных списков вместо одного огромного. "Хеш-функция" преобразует каждый возможный ключ K в номер списка $h(K)$ в диапазоне от 1 до m. Вспомогательная таблица FIRST[i] для каждого $1 \leqslant i \leqslant m$ указывает на первый элемент в списке i; еще одна вспомогательная таблица

$\mathtt{NEXT[j]}$, $1 \leqslant j \leqslant N$, указывает на запись, следующую за записью j в том списке, которому эта запись принадлежит. Будем считать, что

$$\begin{aligned} \mathtt{FIRST[i]} &= -1\,, &&\text{если список } i \text{ пуст;}\\ \mathtt{NEXT[j]} &= 0\,, &&\text{если запись } j \text{ последняя в своем списке.}\end{aligned}$$

Как и ранее, имеется переменная n, которая говорит о том, сколько всего записей сохранено.

Например, пусть ключи представляют собой имена и имеются $m = 4$ списка, основанные на первой букве имени:

$$h(\text{name}) = \begin{cases} 1, & \text{для A–F;}\\ 2, & \text{для G–L;}\\ 3, & \text{для M–R;}\\ 4, & \text{для S–Z.}\end{cases}$$

Поначалу у нас имеются четыре пустых списка и $n = 0$. Если, скажем, ключ первой записи — имя Nora, то $h(\text{Nora}) = 3$, так что Nora становится ключом первого элемента списка 3. Если следующими двумя именами будут Glenn и Jim, то оба они попадут в список 2. Теперь таблица в памяти имеет следующий вид:

$$\mathtt{FIRST[1]} = -1,\quad \mathtt{FIRST[2]} = 2,\quad \mathtt{FIRST[3]} = 1,\quad \mathtt{FIRST[4]} = -1.$$
$$\begin{aligned} &\mathtt{KEY[1]} = \text{Nora}, && \mathtt{NEXT[1]} = 0;\\ &\mathtt{KEY[2]} = \text{Glenn}, && \mathtt{NEXT[2]} = 3;\\ &\mathtt{KEY[3]} = \text{Jim}, && \mathtt{NEXT[3]} = 0; \qquad n = 3.\end{aligned}$$

(Значения DATA[1], DATA[2] и DATA[3] содержат конфиденциальную информацию и здесь не показаны.) После вставки 18 записей списки могли бы содержать следующие имена:

Здесь увековечены имена студентов, сидевших в первых рядах на занятиях по конкретной математике и предоставивших свои имена для этого эксперимента.

Список 1	Список 2	Список 3	Список 4
Dianne	Glenn	Nora	Scott
Ari	Jim	Mike	Tina
Brian	Jennifer	Michael	
Fran	Joan	Ray	
Doug	Jerry	Paula	
	Jean		

Эти имена в массиве KEY записаны вперемешку, но записи массива NEXT позволяют разделить списки. Теперь, чтобы найти имя John, нам потребуется просканировать шесть имен второго списка (который оказался самым длинным); но даже этот длинный список все равно существенно короче общего количества 18 имен.

Вот точное описание алгоритма, который ищет ключ K в соответствии с данной схемой:

H1 Установить $i := h(K)$ и $j := \texttt{FIRST}[i]$.

H2 Если $j \leqslant 0$, остановиться. (Поиск неудачен.)

H3 Если $\texttt{KEY}[j] = K$, остановиться. (Поиск успешен.)

H4 Установить $i := j$, затем установить $j := \texttt{NEXT}[i]$ и вернуться к шагу H2. (Очередная попытка.)

Например, для поиска имени Jennifer из приведенного примера на шаге H1 будет установлено $i := 2$ и $j := 2$; на шаге H3 выяснится, что Glenn $\neq$ Jennifer; на шаге H4 будет установлено $j := 3$; на шаге H3 выяснится, что Jim $\neq$ Jennifer. Еще одна итерация шагов H4 и H3 обнаружит Jennifer в таблице.

Готов спорить, что их родители рады этому факту.

После успешного поиска искомые данные $D(K)$ содержатся в $\texttt{DATA}[j]$, как и в предыдущем алгоритме. После неудачного поиска можно внести в таблицу K и $D(K)$ путем следующих операций:

$$\begin{aligned}&n := n+1;\\&\textbf{Если } j<0 \textbf{ то } \texttt{FIRST}[i] := n \textbf{ иначе } \texttt{NEXT}[i] := n;\\&\texttt{KEY}[n] := K;\quad \texttt{DATA}[n] := D(K);\quad \texttt{NEXT}[n] := 0.\end{aligned} \tag{8.83}$$

Теперь таблица соответствует имеющимся данным.

Мы надеемся, что все списки будут примерно одинаковой длины, поскольку это делает поиск примерно в m раз быстрее. Обычно значение m гораздо больше 4, так что множитель $1/m$ — это значительное улучшение.

Заранее неизвестно, какие именно ключи будут храниться в таблице, но обычно можно выбрать хеш-функцию h так, чтобы значения $h(K)$ были случайной величиной, равномерно распределенной между 1 и m и независимой от хеш-кодов других присутствующих ключей. При этом вычисление хеш-функции подобно бросанию игральной кости с m гранями. Может случиться, что все записи попадут в один список, так же, как на игральной кости всегда может выпадать ⚅; но теория вероятностей говорит нам, что списки *почти всегда* будут достаточно хорошо сбалансированы.

Анализ хеширования: введение

"Анализ алгоритмов" — это раздел информатики, занимающийся получением количественной информации об эффективности компьютерных методов. "Вероятностный анализ алгоритмов" изучает время работы алгоритмов, рассматриваемое в качестве случайной величины, зависящей от предполагаемых характеристик исходных данных. Хеширование особенно хорошо подходит для вероятностного анализа, потому что метод хеширования очень эффективен в среднем, при том что наихудший его случай просто ужасен. (В наихудшем случае все ключи имеют один и тот же хеш-код.) Так что программисту, использующему хеширование, лучше бы верить в теорию вероятностей...

Пусть P — количество выполнений шага H3 при работе описанного ранее алгоритма. (Каждое выполнение шага H3 называется "испытанием".) Если мы знаем P, то можем узнать и количество выполнений каждого шага в зависимости от успешности поиска:

Шаг	Неудачный поиск	Успешный поиск
H1	1 раз	1 раз
H2	$P+1$ раз	P раз
H3	P раз	P раз
H4	P раз	$P-1$ раз

Таким образом, главная характеристика, определяющая время работы процедуры поиска, есть число испытаний P.

Чтобы представить себе работу алгоритма, можно вообразить, что у нас имеется адресная книга, организованная некоторым специальным образом, заключающимся в том, что на каждой странице имеется только одна запись. На обложке книги имеется номер страницы первой записи для каждого из m списков; каждое имя адресата K определяет список $h(K)$, которому он принадлежит. На каждой странице книги имеется ссылка на следующую страницу списка, которому принадлежит текущая страница. Число испытаний, требующееся для поиска адреса в такой книге, — это количество страниц, которые нам придется просмотреть.

Если в таблицу внесено n элементов, то их расположение зависит только от соответствующих хеш-кодов $\langle h_1, h_2, \ldots, h_n \rangle$. Все m^n возможных последовательностей $\langle h_1, h_2, \ldots, h_n \rangle$ считаются равновероятными, случайная величина P зависит от этой последовательности.

Случай 1: ключ отсутствует

Загляните под коврик у двери.

Рассмотрим сначала поведение P при неудачном поиске в предположении, что в таблицу внесено n записей. В этом случае пространство вероятностей состоит из m^{n+1} элементарных событий

$$\omega = (h_1, h_2, \ldots, h_n, h_{n+1}),$$

где h_j — хеш-код для j-го ключа в таблице, а h_{n+1} — хеш-код ключа, который не удалось найти. Мы полагаем, что хеш-функция h корректно выбрана, так что $\Pr(\omega) = 1/m^{n+1}$ для любого ω.

Например, для $m = n = 2$ имеется восемь равновероятных возможностей:

h_1	h_2	h_3:	P
1	1	1 :	2
1	1	2 :	0
1	2	1 :	1
1	2	2 :	1
2	1	1 :	1
2	1	2 :	1
2	2	1 :	0
2	2	2 :	2

Если $h_1 = h_2 = h_3$, то мы делаем два неудачных испытания, прежде чем обнаружим, что новый ключ K в таблице отсутствует; если $h_1 = h_2 \neq h_3$, то нам не надо делать ни одного испытания; и т.д. Этот список всех возможностей показывает, что распределение вероятности P описывается при $m = n = 2$ производящей функцией случайной величины $(\frac{2}{8} + \frac{4}{8}z + \frac{2}{8}z^2) = (\frac{1}{2} + \frac{1}{2}z)^2$.

Неудачный поиск требует по одному испытанию для каждого элемента списка h_{n+1}, так что общая формула имеет вид

$$P = [h_1 = h_{n+1}] + [h_2 = h_{n+1}] + \cdots + [h_n = h_{n+1}]. \quad (8.84)$$

Вероятность того, что $h_j = h_{n+1}$, равна $1/m$ для $1 \leqslant j \leqslant n$; поэтому

$$EP = E[h_1 = h_{n+1}] + E[h_2 = h_{n+1}] + \cdots + E[h_n = h_{n+1}] = \frac{n}{m}.$$

Может быть, следует действовать помедленнее. Итак, пусть X_j — случайная величина

$$X_j = X_j(\omega) = [h_j = h_{n+1}].$$

Тогда $P = X_1 + \cdots + X_n$, и $EX_j = 1/m$ для всех $j \leqslant n$; следовательно,

$$EP = EX_1 + \cdots + EX_n = n/m.$$

Хорошо: как мы и надеялись, среднее количество испытаний в m раз меньше, чем без хеширования. Более того, случайные величины X_j независимы и имеют одну и ту же производящую функцию случайной величины

$$X_j(z) = \frac{m-1+z}{m};$$

следовательно, производящая функция случайной величины для общего количества испытаний при неудачном поиске будет

$$P(z) = X_1(z)\ldots X_n(z) = \left(\frac{m-1+z}{m}\right)^n. \qquad (8.85)$$

Это биномиальное распределение с $p = 1/m$ и $q = (m-1)/m$; другими словами, количество испытаний в безуспешном поиске ведет себя точно так же, как и число выпавших орлов при бросании несимметричной монеты с вероятностью выпадания орла в каждом бросании $1/m$. Уравнение (8.61) говорит нам, что дисперсия P равна

$$npq = \frac{n(m-1)}{m^2}.$$

При большом m дисперсия P приближенно равна n/m, так что стандартное отклонение примерно равно $\sqrt{n/m}$.

Случай 2: ключ присутствует

Рассмотрим теперь успешный поиск. Пространство вероятностей для этого случая несколько более сложное, в зависимости от рассматриваемого приложения. Выберем в качестве Ω множество элементарных событий

$$\omega = (h_1, \ldots, h_n; k), \qquad (8.86)$$

где h_j представляет собой хеш-код для j-го ключа, как и ранее, а k есть индекс искомого ключа (хеш-код которого — h_k). Таким образом, $1 \leqslant h_j \leqslant m$ для $1 \leqslant j \leqslant n$ и $1 \leqslant k \leqslant n$; всего имеется $m^n \cdot n$ элементарных событий ω.

Пусть s_j — вероятность того, что выполняется поиск j-го помещенного в таблицу ключа. Тогда

$$\Pr(\omega) = s_k/m^n, \qquad (8.87)$$

если ω — событие (8.86). (В некоторых приложениях чаще осуществляется поиск ключей, вставленных в таблицу первыми (или наоборот, последними), так что мы не будем полагать, что все $s_j = 1/n$.) Заметим, что $\sum_{\omega\in\Omega}\Pr(\omega) = \sum_{k=1}^{n} s_k = 1$, следовательно, (8.87) определяет конкретное распределение вероятностей.

Количество испытаний P при успешном поиске равно p, если ключ K был вставлен в список p-м. Следовательно,

$$P(h_1,\ldots,h_n;k) = [h_1 = h_k] + [h_2 = h_k] + \cdots + [h_k = h_k]; \qquad (8.88)$$

или, если обозначить случайную величину $[h_j = h_k]$ как X_j,

$$P = X_1 + X_2 + \cdots + X_k. \qquad (8.89)$$

Предположим, например, что $m = 10$, $n = 16$ и что хеш-значения образовали такую "случайную" последовательность:

По-моему, где-то я уже видел эту последовательность...

$$(h_1,\ldots,h_{16}) = 3\ 1\ 4\ 1\ 5\ 9\ 2\ 6\ 5\ 3\ 5\ 8\ 9\ 7\ 9\ 3\ ;$$
$$(P_1,\ldots,P_{16}) = 1\ 1\ 1\ 2\ 1\ 1\ 1\ 1\ 2\ 2\ 3\ 1\ 2\ 1\ 3\ 3\ .$$

Количество испытаний P_j, требующихся, чтобы найти j-й ключ, показано под h_j.

Уравнение (8.89) представляет P в виде суммы случайных величин, но мы не можем вычислить EP как $EX_1 + \cdots + EX_k$, потому что значение k само по себе является случайной величиной. Какова же производящая функция случайной величины P? Для ответа на этот вопрос нам придется ненадолго отвлечься от нашей задачи и поговорить об *условной вероятности.*

Уравнение (8.43) — тоже небольшое отступление.

Если A и B — события пространства вероятности, то мы говорим, что условная вероятность A при условии B есть

$$\Pr(\omega \in A \mid \omega \in B) = \frac{\Pr(\omega \in A \cap B)}{\Pr(\omega \in B)}. \qquad (8.90)$$

Например, если X и Y — случайные величины, то условная вероятность события $X = x$ при условии $Y = y$ равна

$$\Pr(X = x \mid Y = y) = \frac{\Pr(X = x \text{ и } Y = y)}{\Pr(Y = y)}. \qquad (8.91)$$

Для любого фиксированного y, принадлежащего диапазону изменения Y, сумма всех этих условных вероятностей по всем x в диапазоне изменений X составит $\Pr(Y = y)/\Pr(Y = y) = 1$; следовательно, (8.91) определяет распределение вероятностей, и можно определить новую случайную величину '$X|y$', такую, что $\Pr\bigl((X|y) = x\bigr) = \Pr(X = x \mid Y = y)$.

Если X и Y независимы, то случайная величина $X|y$ по сути совпадает с X независимо от значения y, потому что $\Pr(X{=}x \mid Y{=}y)$ равно $\Pr(X = x)$ в силу (8.5); это и есть условие независимости. Но если X и Y зависимы, то случайные величины $X|y$ и $X|y'$ могут быть совершенно не похожи одна на другую при $y \ne y'$.

Если X принимает только неотрицательные целые значения, ее производящую функцию случайной величины можно разложить на сумму условных производящих функций случайных величин по отношению к любой другой случайной величине Y:

$$G_X(z) = \sum_{y\in Y(\Omega)} \Pr(Y=y)\,G_{X|y}(z)\,. \tag{8.92}$$

Это соотношение справедливо потому, что коэффициент при z^x в левой части есть $\Pr(X=x)$ для всех $x \in X(\Omega)$, а в правой части он равен

$$\sum_{y\in Y(\Omega)} \Pr(Y=y)\Pr(X=x\,|\,Y=y) = \sum_{y\in Y(\Omega)} \Pr(X=x \text{ и } Y=y) = \\ = \Pr(X=x)\,.$$

Например, если X — произведение очков, выпавших на двух правильных игральных костях, а Y — сумма очков, то производящая функция случайной величины $X|6$ есть

$$G_{X|6}(z) = \tfrac{2}{5}z^5 + \tfrac{2}{5}z^8 + \tfrac{1}{5}z^9,$$

потому что условные вероятности для $Y = 6$ содержат пять равновероятных событий {⚀⚄, ⚁⚃, ⚂⚂, ⚃⚁, ⚄⚀}. Уравнение (8.92) в этом случае сводится к

$$\begin{aligned} G_X(z) = {} & \tfrac{1}{36}G_{X|2}(z)+\tfrac{2}{36}G_{X|3}(z)+\tfrac{3}{36}G_{X|4}(z)+\tfrac{4}{36}G_{X|5}(z)+ \\ & +\tfrac{5}{36}G_{X|6}(z)+\tfrac{6}{36}G_{X|7}(z)+\tfrac{5}{36}G_{X|8}(z)+\tfrac{4}{36}G_{X|9}(z)+ \\ & +\tfrac{3}{36}G_{X|10}(z)+\tfrac{2}{36}G_{X|11}(z)+\tfrac{1}{36}G_{X|12}(z)\,, \end{aligned}$$

Теперь мне ясно и очевидно, что имеют в виду математики, говоря “очевидно”, “ясно” или “тривиально”.

(Аспирант:
— Шеф, я подготовил статью, но из сорока страниц выкладок потерял все, кроме первой и последней...
Руководитель:
— Ерунда! Просто напишите в конце первой: “путем тривиальных преобразований получаем...”
— Переводчик)

формуле, которая, будучи один раз понятой, становится совершенно очевидной. (Конец отвлечения.)

В случае хеширования (8.92) говорит нам о том, как записать производящую функцию случайной величины для количества испытаний при успешном поиске, если положить $X = P$ и $Y = K$. Для любого фиксированного k между 1 и n случайная величина $P|k$ определяется как сумма независимых случайных величин $X_1 + \cdots + X_k$; это в точности равно (8.89). Поэтому ее производящая функция случайной величины равна

$$G_{P|k}(z) = \left(\frac{m-1+z}{m}\right)^{k-1} z\,.$$

"Это очевидно, и я полагаю, что хороший первокурсник должен быть в состоянии это сделать, хотя это и не тривиально".
— Поль Эрдёш (Paul Erdős) [94].

Очевидно, что производящая функция случайной величины P может быть записана как

$$
\begin{aligned}
G_P(z) &= \sum_{k=1}^{n} s_k G_{P|k}(z) = \\
&= \sum_{k=1}^{n} s_k \left(\frac{m-1+z}{m}\right)^{k-1} z = \\
&= z\,S\left(\frac{m-1+z}{m}\right), \qquad (8.93)
\end{aligned}
$$

где

$$S(z) = s_1 + s_2 z + s_3 z^2 + \cdots + s_n z^{n-1} \qquad (8.94)$$

есть производящая функция случайной величины для вероятностей поиска s_k (для удобства разделенная на z).

Хорошо. У нас есть производящая функция случайной величины P; теперь можно найти математическое ожидание и дисперсию путем дифференцирования. Вычисления несколько упрощаются, если убрать множитель z, как мы делали ранее, и найти таким образом математическое ожидание и дисперсию $P-1$:

$$
\begin{aligned}
F(z) &= G_P(z)/z = S\left(\frac{m-1+z}{m}\right); \\
F'(z) &= \frac{1}{m} S'\left(\frac{m-1+z}{m}\right); \\
F''(z) &= \frac{1}{m^2} S''\left(\frac{m-1+z}{m}\right).
\end{aligned}
$$

Следовательно,

$$EP = 1 + \mathrm{Mean}(F) = 1 + F'(1) = 1 + m^{-1}\,\mathrm{Mean}(S); \qquad (8.95)$$

$$
\begin{aligned}
VP = \mathrm{Var}(F) &= F''(1) + F'(1) - F'(1)^2 = \\
&= m^{-2} S''(1) + m^{-1} S'(1) - m^{-2} S'(1)^2 = \\
&= m^{-2}\,\mathrm{Var}(S) + (m^{-1} - m^{-2})\,\mathrm{Mean}(S). \qquad (8.96)
\end{aligned}
$$

Эти общие формулы выражают математическое ожидание и дисперсию количества испытаний P через математическое ожидание и дисперсию заданного распределения искомых ключей S.

Например, предположим, что $s_k = 1/n$ для $1 \leqslant k \leqslant n$. Это означает, что мы выполняем чисто "случайные" поиски с равной вероятностью для всех ключей в таблице. Тогда $S(z)$ представляет собой равномерное распределение вероятности $U_n(z)$ из (8.32), и мы получим $\mathrm{Mean}(S) = (n-1)/2$, $\mathrm{Var}(S) = (n^2-1)/12$. Следовательно,

$$EP = \frac{n-1}{2m}+1\,; \tag{8.97}$$

$$VP = \frac{n^2-1}{12m^2}+\frac{(m-1)(n-1)}{2m^2} = \frac{(n-1)(6m+n-5)}{12m^2}\,. \tag{8.98}$$

Мы вновь получили ускорение в m раз. Если $m \approx n/\ln n$ и $n \to \infty$, среднее количество испытаний при успешном поиске будет примерно равно $\frac{1}{2}\ln n$, а стандартное отклонение асимптотически равно $(\ln n)/\sqrt{12}$.

С другой стороны, можно принять $s_k = (kH_n)^{-1}$ для $1 \leqslant k \leqslant n$; это распределение называется “законом Зипфа”. Тогда $\mathrm{Mean}(S) = n/H_n - 1$, $\mathrm{Var}(S) = \frac{1}{2}n(n+1)/H_n - n^2/H_n^2$. Среднее количество испытаний для $m \approx n/\ln n$ при $n \to \infty$ примерно равно 2 со стандартным отклонением, асимптотически равным $\sqrt{\ln n}/\sqrt{2}$.

В обоих случаях проведенный анализ позволяет спокойно спать пессимистам, опасающимся наихудшего случая: неравенство Чебышева говорит нам о том, что списки будут хорошими и короткими, за исключением крайне редких случаев.

Случай 2, продолжение: разные дисперсии

Мы только что вычислили дисперсию количества испытаний в случае успешного поиска, рассматривая P как случайную величину на пространстве вероятности из $m^n \cdot n$ элементов $(h_1, \ldots, h_n; k)$. Но мы могли бы принять и другую точку зрения. Каждый набор хеш-кодов $(h_1, \ldots, h_n)$ определяет случайную величину $P|(h_1, \ldots, h_n)$, которая представляет собой количество испытаний при успешном поиске в данной конкретной хеш-таблице с n заданными ключами. Среднее значение $P|(h_1, \ldots, h_n)$,

$$A(h_1, \ldots, h_n) = \sum_{p=1}^{n} p \cdot \Pr\bigl((P|(h_1, \ldots, h_n)) = p\bigr)\,, \tag{8.99}$$

как можно считать, представляет время успешного поиска. Величина $A(h_1, \ldots, h_n)$ является случайной величиной, которая зависит только от $(h_1, \ldots, h_n)$, но не от последнего компонента k. Ее можно записать в виде

$$A(h_1, \ldots, h_n) = \sum_{k=1}^{n} s_k\, P(h_1, \ldots, h_n; k)\,,$$

где $P(h_1,\ldots,h_n;k)$ определено в (8.88), поскольку $P|(h_1,\ldots,h_n)=p$ с вероятностью

$$\frac{\sum_{k=1}^{n}\Pr\bigl(P(h_1,\ldots,h_n;k)=p\bigr)}{\sum_{k=1}^{n}\Pr(h_1,\ldots,h_n;k)} = \\ = \frac{\sum_{k=1}^{n} m^{-n}s_k\bigl[P(h_1,\ldots,h_n;k)=p\bigr]}{\sum_{k=1}^{n} m^{-n}s_k} = \\ = \sum_{k=1}^{n} s_k\bigl[P(h_1,\ldots,h_n;k)=p\bigr].$$

Среднее значение $A(h_1,\ldots,h_n)$, полученное суммированием по всем m^n возможным наборам $(h_1,\ldots,h_n)$ и делением на m^n, будет тем же, что и ранее найденное в (8.95). Однако *дисперсия* $A(h_1,\ldots,h_n)$ несколько отличается; это дисперсия m^n средних, а не $m^n\cdot n$ отдельных значений количества испытаний. Например, если $m=1$ (так что у нас имеется только один список), "среднее" значение $A(h_1,\ldots,h_n)=A(1,\ldots,1)$ на самом деле представляет собой константу, так что ее дисперсия VA равна нулю; однако количество испытаний в успешном поиске не постоянно, и его дисперсия VP — не нуль.

VP (Vox Populi, глас народа) не ноль только во время выборов.

"И не безлики вы, и вы — не тени, Коль надо в урны бросить бюллетени!"
— В. С. Высоцкий (Доб. пер.)

Можно проиллюстрировать это различие между дисперсиями, выполнив вычисления для произвольных m и n в простейшем случае, когда $s_k=1/n$ для $1\leqslant k\leqslant n$. Другими словами, мы временно полагаем, что ключи поиска распределены равномерно. Любая данная последовательность хеш-кодов $(h_1,\ldots,h_n)$ определяет m списков, которые содержат соответственно $(n_1,n_2,\ldots,n_m)$ записей для некоторых чисел n_j, где

$$n_1+n_2+\cdots+n_m = n.$$

Успешный поиск, в котором каждый из n ключей в таблице равновероятен, потребует в среднем

$$A(h_1,\ldots,h_n) = \frac{(1+\cdots+n_1)+(1+\cdots+n_2)+\cdots+(1+\cdots+n_m)}{n} = \\ = \frac{n_1(n_1+1)+n_2(n_2+1)+\cdots+n_m(n_m+1)}{2n} = \\ = \frac{n_1^2+n_2^2+\cdots+n_m^2+n}{2n}$$

испытаний. Наша цель заключается в вычислении дисперсии величины $A(h_1,\ldots,h_n)$ на пространстве вероятности, состоящем из всех m^n последовательностей $(h_1,\ldots,h_n)$.

Как оказывается, проще вычислить дисперсию несколько иной величины

$$B(h_1,\ldots,h_n) \;=\; \binom{n_1}{2}+\binom{n_2}{2}+\cdots+\binom{n_m}{2}.$$

Имеем

$$A(h_1,\ldots,h_n) \;=\; 1+B(h_1,\ldots,h_n)/n\,,$$

следовательно, математическое ожидание и дисперсия A удовлетворяют соотношениям

$$EA \;=\; 1+\frac{EB}{n}\,; \qquad VA \;=\; \frac{VB}{n^2}\,. \tag{8.100}$$

Вероятность того, что размеры списков составят n_1, n_2, …, n_m, равна биномиальному коэффициенту

$$\binom{n}{n_1,n_2,\ldots,n_m} \;=\; \frac{n!}{n_1!\,n_2!\,\ldots n_m!}\,,$$

деленному на m^n; следовательно, производящая функция случайной величины $B(h_1,\ldots,h_n)$ представляет собой

$$B_n(z)=\sum_{\substack{n_1,n_2,\ldots,n_m\geqslant 0\\ n_1+n_2+\cdots+n_m=n}}\binom{n}{n_1,n_2,\ldots,n_m}z^{\binom{n_1}{2}+\binom{n_2}{2}+\cdots+\binom{n_m}{2}}\,m^{-n}\,.$$

Для неискушенного глаза эта сумма выглядит жутковато, но наш опыт, приобретенный в главе 7, помогает распознать в ней m-кратную свертку. Действительно, если определить экспоненциальную суперпроизводящую функцию

$$G(w,z) \;=\; \sum_{n\geqslant 0} B_n(z)\frac{m^n w^n}{n!}\,,$$

то можно легко убедиться, что $G(w,z)$ есть просто m-я степень:

$$G(w,z) \;=\; \biggl(\sum_{k\geqslant 0} z^{\binom{k}{2}}\,\frac{w^k}{k!}\biggr)^m\,.$$

Для проверки подставим $z=1$; мы получим $G(w,1)=(e^w)^m$, так что коэффициент при $m^n w^n/n!$ равен $B_n(1)=1$.

Если бы мы нашли значения $B_n'(1)$ и $B_n''(1)$, то могли бы вычислить $\mathrm{Var}(B_n)$. Возьмем поэтому частную производную $G(w,z)$

по z:

$$\frac{\partial}{\partial z}G(w,z) = \sum_{n\geqslant 0} B_n'(z)\frac{m^n w^n}{n!} =$$

$$= m\Bigl(\sum_{k\geqslant 0} z^{\binom{k}{2}}\,\frac{w^k}{k!}\Bigr)^{m-1}\sum_{k\geqslant 0}\binom{k}{2}z^{\binom{k}{2}-1}\,\frac{w^k}{k!}\,;$$

$$\frac{\partial^2}{\partial z^2}G(w,z) = \sum_{n\geqslant 0} B_n''(z)\frac{m^n w^n}{n!} =$$

$$= m(m-1)\Bigl(\sum_{k\geqslant 0} z^{\binom{k}{2}}\,\frac{w^k}{k!}\Bigr)^{m-2}\Bigl(\sum_{k\geqslant 0}\binom{k}{2}z^{\binom{k}{2}-1}\,\frac{w^k}{k!}\Bigr)^2 =$$

$$+m\Bigl(\sum_{k\geqslant 0} z^{\binom{k}{2}}\,\frac{w^k}{k!}\Bigr)^{m-1}\sum_{k\geqslant 0}\binom{k}{2}\Bigl(\binom{k}{2}-1\Bigr)z^{\binom{k}{2}-2}\,\frac{w^k}{k!}\,.$$

Да, это действительно сложные выражения. Но все значительно упрощается, если положить $z = 1$. Например,

$$\sum_{n\geqslant 0} B_n'(1)\frac{m^n w^n}{n!} = me^{(m-1)w}\sum_{k\geqslant 2}\frac{w^k}{2(k-2)!} =$$

$$= me^{(m-1)w}\sum_{k\geqslant 0}\frac{w^{k+2}}{2k!} = \frac{mw^2e^{(m-1)w}}{2}e^w =$$

$$= \sum_{n\geqslant 0}\frac{(mw)^{n+2}}{2m\,n!} = \sum_{n\geqslant 0}\frac{n(n-1)m^n w^n}{2m\,n!}\,,$$

откуда следует, что

$$B_n'(1) = \binom{n}{2}\frac{1}{m}\,. \qquad (8.101)$$

Выражение для EA в (8.100) в полном согласии с (8.97) теперь дает $EA = 1 + (n-1)/2m$.

Формула для $B_n''(1)$ включает похожую сумму

$$\sum_{k\geqslant 0}\binom{k}{2}\Bigl(\binom{k}{2}-1\Bigr)\frac{w^k}{k!} = \frac{1}{4}\sum_{k\geqslant 0}\frac{(k+1)k(k-1)(k-2)w^k}{k!} =$$

$$= \frac{1}{4}\sum_{k\geqslant 3}\frac{(k+1)w^k}{(k-3)!} = \frac{1}{4}\sum_{k\geqslant 0}\frac{(k+4)w^{k+3}}{k!} = (\tfrac{1}{4}w^4+w^3)e^w\,;$$

следовательно, находим, что

$$\sum_{n\geqslant 0} B_n''(1)\frac{m^n w^n}{n!} = m(m-1)e^{w(m-2)}\left(\tfrac{1}{2}w^2e^w\right)^2 + \\ + me^{w(m-1)}\left(\tfrac{1}{4}w^4+w^3\right)e^w = \\ = me^{wm}\left(\tfrac{1}{4}mw^4+w^3\right);$$

$$B_n''(1) = \binom{n}{2}\left(\binom{n}{2}-1\right)\frac{1}{m^2}. \qquad (8.102)$$

Теперь можно собрать все вместе и вычислить искомую дисперсию VA. При этом сокращается масса членов, и результат оказывается на удивление простым:

$$VA = \frac{VB}{n^2} = \frac{B_n''(1)+B_n'(1)-B_n'(1)^2}{n^2} = \\ = \frac{n(n-1)}{m^2n^2}\left(\frac{(n+1)(n-2)}{4}+\frac{m}{2}-\frac{n(n-1)}{4}\right) = \\ = \frac{(m-1)(n-1)}{2m^2n}. \qquad (8.103)$$

Встретив такое "совпадение", мы можем заподозрить какую-то математическую причину этого; может существовать и другой подход к решению поставленной задачи, поясняющий, почему ответ выглядит так просто. И действительно, такой подход имеется (в упр. 61), и он показывает, что в общем случае дисперсия среднего успешного поиска имеет общий вид

$$VA = \frac{m-1}{m^2}\sum_{k=1}^{n} s_k^2(k-1), \qquad (8.104)$$

где s_k — вероятность поиска элемента, вставленного k-м по счету. Уравнение (8.103) представляет собой частный случай $s_k = 1/n$ для $1 \leqslant k \leqslant n$.

Помимо дисперсии среднего, можно рассмотреть и среднее дисперсии. Другими словами, каждая последовательность $(h_1, \ldots, h_n)$, задающая хеш-таблицу, определяет также распределение вероятностей для успешного поиска, и дисперсия этого распределения вероятностей показывает, насколько сильно будет разбросано количество испытаний в различных успешных поисках. Например, вернемся к случаю, когда мы размещали $n = 16$ элементов в $m = 10$ списках:

$$(h_1,\ldots,h_{16}) = 3\ 1\ 4\ 1\ 5\ 9\ 2\ 6\ 5\ 3\ 5\ 8\ 9\ 7\ 9\ 3$$
$$(P_1,\ldots,P_{16}) = 1\ 1\ 1\ 2\ 1\ 1\ 1\ 1\ 2\ 2\ 3\ 1\ 2\ 1\ 3\ 3$$

По-моему, где-то я уже видел эту последовательность...

По-моему, где-то я уже видел это замечание...

"Это я знаю и помню прекрасно, пи многие знаки мне лишни, напрасны"...

$\pi \approx 3.1415926535$
8979323846264338
3279502884197169
3993751058209749
4459230781640628
6208998628034825
3421170679821480
8651328230664709
3844609550582231
7253594081284811
1745028410270193
8521105559644622
9489549303819644
2881097566593344
6128475648233786
7831652712019091
4564856692346034
8610454326648213
3936072602491412
7372458700660631
5588174881520920
9628292540917153
6436789259036001
1330530548820466
5213841469519415
1160943305727036
5759591953092186
1173819326117931
0511854807446237
9962749567351...

Успешный поиск в результирующей таблице имеет производящую функцию случайной величины

$$G(3,1,4,1,\ldots,3) = \sum_{k=1}^{16} s_k z^{P(3,1,4,1,\ldots,3;k)} = s_1 z + s_2 z + s_3 z + s_4 z^2 + \cdots + s_{16} z^3 .$$

Только что мы занимались средним числом проб для успешного поиска в этой таблице, а именно $A(3,1,4,1,\ldots,3) = \mathrm{Mean}\big(G(3,1,4,1,\ldots,3)\big)$. Можно также рассмотреть дисперсию

$$s_1\cdot 1^2 + s_2\cdot 1^2 + s_3\cdot 1^2 + s_4\cdot 2^2 + \cdots + s_{16}\cdot 3^2 - (s_1\cdot 1 + s_2\cdot 1 + s_3\cdot 1 + s_4\cdot 2 + \cdots + s_{16}\cdot 3)^2 .$$

Эта дисперсия является случайной величиной, зависящей от $(h_1,\ldots,h_n)$, так что рассмотрение ее среднего значения вполне закономерно.

Другими словами, имеются три естественных вида дисперсии, которые могут быть полезны при изучении поведения успешного поиска. Это *общая дисперсия* количества испытаний, взятая по всем $(h_1,\ldots,h_n)$ и k; *дисперсия среднего* числа испытаний, где среднее берется по всем k, а затем дисперсия берется по всем $(h_1,\ldots,h_n)$; и *среднее дисперсии* количества испытаний, где дисперсия берется по всем k, а затем среднее — по всем наборам $(h_1,\ldots,h_n)$. В символьном виде общая дисперсия записывается как

$$VP = \sum_{1\leqslant h_1,\ldots,h_n\leqslant m}\ \sum_{k=1}^{n} \frac{s_k}{m^n} P(h_1,\ldots,h_n;k)^2 - \left(\sum_{1\leqslant h_1,\ldots,h_n\leqslant m}\ \sum_{k=1}^{n} \frac{s_k}{m^n} P(h_1,\ldots,h_n;k)\right)^2 ;$$

дисперсия среднего — как

$$VA = \sum_{1\leqslant h_1,\ldots,h_n\leqslant m} \frac{1}{m^n}\left(\sum_{k=1}^{n} s_k P(h_1,\ldots,h_n;k)\right)^2 - \left(\sum_{1\leqslant h_1,\ldots,h_n\leqslant m} \frac{1}{m^n}\sum_{k=1}^{n} s_k P(h_1,\ldots,h_n;k)\right)^2 ;$$

и среднее дисперсии — как

$$AV = \sum_{1\leqslant h_1,\ldots,h_n\leqslant m} \frac{1}{m^n}\left(\sum_{k=1}^{n} s_k P(h_1,\ldots,h_n;k)^2 - \left(\sum_{k=1}^{n} s_k P(h_1,\ldots,h_n;k)\right)^2\right) .$$

Оказывается, что эти три величины связаны между собой простым соотношением:

$$VP = VA + AV. \tag{8.105}$$

В самом деле, условные распределения вероятностей всегда удовлетворяют тождеству

$$VX = V\big(E(X|Y)\big) + E\big(V(X|Y)\big), \tag{8.106}$$

если X и Y — случайные величины в произвольном пространстве вероятностей, причем X принимает вещественные значения. (Это тождество доказывается в упр. 22.) Уравнение (8.105) является частным случаем, когда X — количество испытаний при успешном поиске, а Y — последовательность хеш-кодов $(h_1, \ldots, h_n)$.

Общее уравнение (8.106) требует внимательного рассмотрения, поскольку в нем скрыто несколько различных случайных величин и пространств вероятностей, в которых надлежит вычислять математические ожидания и дисперсии. Для каждого y из диапазона изменения Y в (8.91) определена случайная величина $X|y$, и эта случайная величина имеет ожидаемое значение $E(X|y)$, зависящее от y. Далее, $E(X|Y)$ обозначает случайную величину, принимающую значения $E(X|y)$ при y, принимающем все возможные для Y значения, а $V\big(E(X|Y)\big)$ есть дисперсия этой случайной величины по отношению к распределению вероятностей Y. Аналогично $E\big(V(X|Y)\big)$ представляет собой среднее значение случайной величины $V(X|y)$ при изменении y. В левой части (8.106) находится VX, безусловная дисперсия X. Поскольку дисперсии неотрицательны, мы всегда будем иметь

Сейчас самое время размяться упражнением 6.

$$VX \geqslant V\big(E(X|Y)\big) \quad \text{и} \quad VX \geqslant E\big(V(X|Y)\big). \tag{8.107}$$

Случай 1, еще раз: пересмотр неудачного поиска

В завершение нашего микроскопического изучения хеширования выполним еще одно вычисление, типичное для анализа алгоритмов. На этот раз рассмотрим более подробно *общее время работы* неудачного поиска в предположении, что компьютер вставляет в память до этого отсутствовавший там ключ.

В процессе вставки (8.83) возможны два случая в зависимости от того, какое значение принимает переменная j — отрицательное или нулевое. При этом $j < 0$ тогда и только тогда, когда $P = 0$, поскольку отрицательное значение может появиться только из элемента FIRST для пустого списка. Таким образом, если список был пуст, то $P = 0$ и следует установить $\mathrm{FIRST}[h_{n+1}] := n+1$ (новая запись будет помещена в позицию $n+1$). В противном

P по-прежнему обозначает количество испытаний.

случае $P > 0$ и значение записи NEXT должно быть установлено равным $n+1$. Выполнение алгоритма в этих двух случаях может отнять различное время; поэтому общее время работы для неудачного поиска можно записать в виде

$$T = \alpha + \beta P + \delta[P=0], \tag{8.108}$$

где α, β и δ — некоторые константы, зависящие от компьютера и от того, каким образом алгоритм хеширования закодирован в машинных командах компьютера. Было бы неплохо найти математическое ожидание и дисперсию случайной величины T, поскольку данная информация более важна для практики, чем математическое ожидание и дисперсия P.

До сих пор мы использовали производящие функции случайных величин только для тех случайных величин, которые принимают неотрицательные целые значения. Оказывается, однако, что с выражением

$$G_X(z) = \sum_{\omega\in\Omega} \Pr(\omega) z^{X(\omega)}$$

для произвольной случайной величины X, принимающей действительные значения, можно работать ничуть не хуже, потому что все существенные характеристики X зависят только от поведения G_X вблизи $z=1$, где все степени z вполне определены. Например, время работы неудачного поиска (8.108) является случайной величиной, определенной на пространстве вероятности из равновероятных последовательностей хеш-кодов $(h_1, \ldots, h_n, h_{n+1})$ с $1 \leqslant h_j \leqslant m$; можно считать, что ряд

$$G_T(z) = \\ = \frac{1}{m^{n+1}} \sum_{h_1=1}^{m} \cdots \sum_{h_n=1}^{m} \sum_{h_{n+1}=1}^{m} z^{\alpha+\beta P(h_1,\ldots,h_{n+1})+\delta[P(h_1,\ldots,h_{n+1})=0]}$$

представляет собой производящую функцию случайной величины даже в тех случаях, когда α, β и δ — не целые. (Фактически параметры α, β, δ — физические величины, имеющие размерность времени; это не просто числа! Тем не менее их можно использовать в показателе степени z.) Можно по-прежнему вычислять математическое ожидание и дисперсию T, находя $G_T'(1)$ и $G_T''(1)$ и комбинируя эти значения обычным образом.

Производящая функция для P (а не T) представляет собой

$$P(z) = \left(\frac{m-1+z}{m}\right)^n = \sum_{p\geqslant 0} \Pr(P=p) z^p .$$

Таким образом, имеем

$$\begin{aligned} G_T(z) &= \sum_{p\geqslant 0} \Pr(P=p) z^{\alpha+\beta p+\delta[p=0]} = \\ &= z^\alpha\Bigl((z^\delta-1)\Pr(P=0)+\sum_{p\geqslant 0}\Pr(P=p)z^{\beta p}\Bigr) = \\ &= z^\alpha\Bigl((z^\delta-1)\Bigl(\frac{m-1}{m}\Bigr)^n+\Bigl(\frac{m-1+z^\beta}{m}\Bigr)^n\Bigr). \end{aligned}$$

Подсчет $\mathrm{Mean}(G_T)$ и $\mathrm{Var}(G_T)$ оказывается теперь рутинной операцией:

$$\mathrm{Mean}(G_T) = G_T'(1) = \alpha+\beta\frac{n}{m}+\delta\Bigl(\frac{m-1}{m}\Bigr)^n; \tag{8.109}$$

$$\begin{aligned} G_T''(1) &= \alpha(\alpha-1)+2\alpha\beta\frac{n}{m}+\beta(\beta-1)\frac{n}{m}+\beta^2\frac{n(n-1)}{m^2}+ \\ &\quad +2\alpha\delta\Bigl(\frac{m-1}{m}\Bigr)^n+\delta(\delta-1)\Bigl(\frac{m-1}{m}\Bigr)^n; \\ \mathrm{Var}(G_T) &= G_T''(1)+G_T'(1)-G_T'(1)^2 = \\ &= \beta^2\frac{n(m-1)}{m^2}-2\beta\delta\Bigl(\frac{m-1}{m}\Bigr)^n\frac{n}{m}+ \\ &\quad +\delta^2\Bigl(\Bigl(\frac{m-1}{m}\Bigr)^n-\Bigl(\frac{m-1}{m}\Bigr)^{2n}\Bigr). \end{aligned} \tag{8.110}$$

В главе 9 мы научимся оценивать подобные величины при больших m и n. Если, например, $m=n$ и $n\to\infty$, то с помощью методов главы 9 можно показать, что математическое ожидание и дисперсия T равны соответственно $\alpha+\beta+\delta e^{-1}+O(n^{-1})$ и $\beta^2-2\beta\delta e^{-1}+\delta^2(e^{-1}-e^{-2})+O(n^{-1})$. Если $m=n/\ln n+O(1)$ и $n\to\infty$, то соответствующие величины составят

$$\begin{aligned} \mathrm{Mean}(G_T) &= \beta\ln n+\alpha+O\bigl((\log n)^2/n\bigr); \\ \mathrm{Var}(G_T) &= \beta^2\ln n+O\bigl((\log n)^2/n\bigr). \end{aligned}$$

Упражнения

Разминка

1 Какова вероятность выпадения дубля при распределении вероятностей $\Pr_{01}$ из (8.3), когда одна из игральных костей правильная, а у другой центр масс смещен? Какова вероятность выбросить сумму $S=7$?

2 Какова вероятность того, что в перетасованной колоде карт верхняя и нижняя карты окажутся тузами? (Все 52! перестановки карт имеют вероятность 1/52!.)

3 В 1979 году студентов Станфорда, изучающих конкретную математику, попросили подбрасывать монету до тех пор, пока не выпадут два орла подряд, и сообщить, сколько бросаний потребовалось. В результате получилось

Почему приведено только десять чисел?

Либо остальные студенты чистые теоретики, либо забросили монету подальше...

$$3,\ 2,\ 3,\ 5,\ 10,\ 2,\ 6,\ 6,\ 9,\ 2\,.$$

В 1987 году в Принстоне студентов попросили сделать то же самое, и они получили такие результаты:

$$10,\ 2,\ 10,\ 7,\ 5,\ 2,\ 10,\ 6,\ 10,\ 2\,.$$

Оцените математическое ожидание и дисперсию на основе данных (а) из Станфорда; (б) из Принстона.

4 Пусть $H(z) = F(z)/G(z)$, где $F(1) = G(1) = 1$. Докажите, что аналогично (8.38) и (8.39)

$$\begin{aligned} \mathrm{Mean}(H) &= \mathrm{Mean}(F) - \mathrm{Mean}(G)\,, \\ \mathrm{Var}(H) &= \mathrm{Var}(F) - \mathrm{Var}(G)\,, \end{aligned}$$

если указанные производные существуют при $z = 1$.

5 Предположим, что Алиса и Билл играют в игру (8.78) фальшивой монетой, вероятность выпадения орла у которой равна p. Существует ли значение p, при котором игра станет справедливой?

6 К чему сведется закон дисперсии для условных распределений (8.106), если X и Y представляют собой независимые случайные величины?

Обязательные упражнения

7 Покажите, что если у обеих игральных костей смещен центр масс, причем распределение вероятностей у них одинаковое, то вероятность дубля будет всегда не меньше $\frac{1}{6}$.

8 Пусть A и B — события, такие, что $A \cup B = \Omega$. Докажите, что

$$\Pr(\omega \in A \cap B) = \Pr(\omega \in A)\Pr(\omega \in B) - \Pr(\omega \notin A)\Pr(\omega \notin B)\,.$$

9 Докажите или опровергните: если X и Y — независимые случайные величины, то таковыми же являются и $F(X)$ и $G(Y)$, где F и G — любые функции.

10 Каково максимальное количество элементов, которые могут быть медианами случайной величины X в соответствии с определением (8.7)?

11 Постройте случайную величину, имеющую конечное математическое ожидание и бесконечную дисперсию.

12 **а** Пусть $P(z)$ — производящая функция случайной величины X. Докажите, что

$$\begin{aligned} \Pr(X \leqslant r) &\leqslant x^{-r}P(x) \qquad \text{для } 0 < x \leqslant 1; \\ \Pr(X \geqslant r) &\leqslant x^{-r}P(x) \qquad \text{для } x \geqslant 1. \end{aligned}$$

(Эти важные соотношения называются *неравенствами хвоста*.)

б В частном случае $P(z) = (1+z)^n/2^n$ воспользуйтесь первым неравенством хвоста для доказательства того, что $\sum_{k\leqslant\alpha n}\binom{n}{k} \leqslant 1/\alpha^{\alpha n}(1-\alpha)^{(1-\alpha)n}$ при $0 < \alpha < \frac{1}{2}$.

13 Докажите, что если $X_1, \ldots, X_{2n}$ — независимые случайные величины с одинаковыми распределениями, а α — произвольное действительное число, то

$$\Pr\left(\left|\frac{X_1 + \cdots + X_{2n}}{2n} - \alpha\right| \leqslant \left|\frac{X_1 + \cdots + X_n}{n} - \alpha\right|\right) \geqslant \frac{1}{2}.$$

14 Пусть $F(z)$ и $G(z)$ — производящие функции случайных величин и пусть

$$H(z) = p\,F(z) + q\,G(z),$$

где $p + q = 1$. (Такая функция называется *смесью* функций F и G; соответствующая ей случайная величина получится, если подбросить монету и затем выбрать распределение F или G в зависимости от того, какой стороной упадет монета.) Выразите математическое ожидание и дисперсию H через p, q и математическое ожидание и дисперсию F и G.

15 Если $F(z)$ и $G(z)$ — производящие функции случайных величин, то можно определить новую производящую функцию случайной величины $H(z)$ при помощи их “композиции”:

$$H(z) = F\big(G(z)\big)\,.$$

Выразите $\mathrm{Mean}(H)$ и $\mathrm{Var}(H)$ через $\mathrm{Mean}(F)$, $\mathrm{Var}(F)$, $\mathrm{Mean}(G)$ и $\mathrm{Var}(G)$. (Уравнение (8.93) будет тогда частным случаем.)

16 Найдите аналитическое выражение для суперпроизводящей функции $\sum_{n\geqslant 0} F_n(z)w^n$, где $F_n(z)$ — "футбольная" производящая функция, определенная в (8.53).

17 Пусть $X_{n,p}$ и $Y_{n,p}$ имеют соответственно биномиальное распределение и отрицательное биномиальное распределение с параметрами (n, p). (Эти распределения определены в (8.57) и (8.60).) Докажите, что $\Pr(Y_{n,p} \leqslant m) = \Pr(X_{m+n,p} \geqslant n)$. Какое тождество для биномиальных коэффициентов отсюда вытекает?

Распределение рыб в единице объема океана.

(Учел ли автор, что рыбы часто ходят косяками?... — Переводчик)

18 Говорят, что случайная величина X имеет *распределение Пуассона* с математическим ожиданием μ, если $\Pr(X=k) = e^{-\mu}\mu^k/k!$ для всех $k \geqslant 0$.

а Какова производящая функция этой случайной величины?

б Чему равны ее математическое ожидание, дисперсия и прочие кумулянты?

19 Продолжая предыдущее упражнение, допустим, что случайная величина X_1 имеет распределение Пуассона с математическим ожиданием μ_1, а случайная величина X_2 независима от X_1 и имеет распределение Пуассона с математическим ожиданием μ_2.

а Чему равна вероятность того, что $X_1 + X_2 = n$?

б Чему равны математическое ожидание, дисперсия и прочие кумулянты $2X_1 + 3X_2$?

20 Докажите (8.74) и (8.75) — общие формулы для математического ожидания и дисперсии времени до появления заданной последовательности орлов и решек.

21 Что выражает значение N, если и O, и P в (8.77) равны $\frac{1}{2}$?

22 Докажите (8.106) — закон условного математического ожидания и дисперсии.

Домашние задания

23 Пусть $\Pr_{00}$ — распределение вероятностей двух правильных игральных костей, а $\Pr_{11}$ — распределение вероятностей двух игральных костей со смещенным центром масс, определенное в (8.2). Найдите все такие события A, что $\Pr_{00}(A) = \Pr_{11}(A)$. Какие из этих событий зависят только от случайной величины S? (Пространство вероятностей с $\Omega = D^2$ имеет 2^{36} событий; лишь 2^{11} из них зависят только от S.)

24 Игрок J бросает $2n+1$ правильных игральных костей и убирает те из них, на которых выпало ⚅. Затем игрок K называет число от 1 до 6, бросает оставшиеся игральные кости

и убирает те, на которых выпало названное им число очков. Этот процесс повторяется до тех пор, пока в игре не останется ни одной игральной кости. Выигрывает игрок, убравший в совокупности большее число игральных костей (т.е. $n+1$ или более).

а Чему равно математическое ожидание и дисперсия общего числа игральных костей, убранных игроком J? *Указание*: игральные кости независимы.

б С какой вероятностью выигрывает игрок J при $n=2$?

25 Рассмотрим следующую азартную игру. Вы ставите определенную сумму A и бросаете правильную игральная кость. Если выпало k очков, то ваша ставка умножается на $2(k-1)/5$. (В частности, ставка удваивается, если выпадет ⚅, но вы проигрываете все, если выпадет ⚀.) В любой момент вы можете остановиться и потребовать текущую сумму ставки. Чему равно среднее значение и дисперсия суммы ставки после n бросков? (Пренебрегите эффектами округления суммы до целого числа.)

26 Найдите математическое ожидание и дисперсию количества циклов длиной l в случайной перестановке n элементов. (Задача о победе футбольной команды, ответ к которой дан в (8.23), (8.24) и (8.53), представляет собой частный случай $l=1$.)

27 Пусть X_1, X_2, ..., X_n — независимые значения случайной величины X. Уравнения (8.19) и (8.20) говорят нам, как оценить математическое ожидание и дисперсию X по результатам этих наблюдений. Приведите аналогичную формулу для оценки третьего кумулянта κ_3. (Ваша формула должна давать “несмещенную” оценку в том смысле, что ожидаемое значение оценки должно быть равно κ_3.)

28 Какова средняя продолжительность игры в монету (8.78)

а при условии, что выигрывает Алиса;

б при условии, что выигрывает Билл?

29 Алиса, Билл и Компьютер бросают правильную монету до тех пор, пока не появится одна из последовательностей A = ООРО, B = ОРОО или C = РООО. (Если в игре принимают участие только две из этих последовательностей, то, как мы знаем из (8.82), A, вероятно, выиграет у B, B, вероятно, выиграет у C и C, вероятно, выиграет у A; но в данной игре участвуют все три последовательности одновременно.) Каковы шансы каждого игрока на выигрыш?

30 В тексте главы мы рассматривали три вида дисперсии, связанные с успешным поиском в хеш-таблице. На самом деле есть еще два вида: можно рассмотреть среднее (по k) дисперсий (по $h_1, \ldots, h_n$) количества испытаний $P(h_1, \ldots, h_n; k)$; и можно рассмотреть дисперсию (по k) средних (по $h_1, \ldots, h_n$). Вычислите эти величины.

Червяк Шредингера.

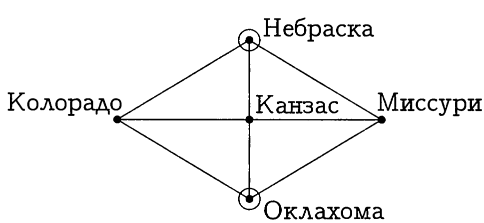
31 В вершине A пятиугольника $ABCDE$ находится яблоко, а на расстоянии двух ребер, в вершине C, находится червяк. Каждый день червяк переползает в одну из соседних вершин с равной вероятностью. Так, через один день червяк окажется в вершине B или вершине D с вероятностью оказаться в каждой из вершин, равной $\frac{1}{2}$. Через два дня червяк может вернуться в C, потому что он не запоминает предыдущего местоположения. Достигнув вершины A, он останавливается на обед.

а Чему равно математическое ожидание и дисперсия количества дней, прошедших до обеда?

б Пусть p — вероятность того, что это количество дней не менее 100. Что говорит о значении p неравенство Чебышева?

в Что позволяет сказать о p неравенство хвоста (упр. 12)?

32 Алиса и Билл в армии, и их части дислоцируются в пяти штатах: Канзасе, Небраске, Миссури, Оклахоме и Колорадо. Изначально Алиса находится в Небраске, а Билл — в Оклахоме. Каждый месяц их переводят в соседний штат, причем любой из соседних штатов выбирается с равной вероятностью). (Вот диаграмма “соседства” штатов:



Нештатная штатная ситуация.

Начальные штаты обведены кружками.) Например, Алиса через один месяц будет переведена с вероятностью $1/3$ в любой из штатов Колорадо, Канзас и Миссури. Найдите математическое ожидание и дисперсию количества месяцев, которое потребуется, чтобы Алиса и Билл встретились. (Возможно, вы захотите прибегнуть к помощи компьютера.)

33 Независимы ли случайные величины X_1 и X_2 из (8.89)?

34 Джина играет в гольф и на каждом ударе с вероятностью $p = .05$ имеет "супершот" и выигрывает один удар, с вероятностью $q = .91$ имеет обычный "шот" и с вероятностью $r = .04$ — "субшот", который стоит ей одного удара в сравнении с расчетным числом. (Для не играющих в гольф: на каждом круге она делает 2, 1 или 0 шагов к цели с вероятностями соответственно p, q и r. В игре с m-ударной лункой ее результат есть минимальное n, такое, что она продвигается на m или более шагов за n кругов. Маленький результат лучше большого.)

Для выполнения расчетов воспользуйтесь калькулятором.

а Покажите, что Джина выигрывает игру с 4-ударной лункой чаще, чем проигрывает, когда она играет с игроком, делающим всегда обычные удары. (Другими словами, вероятность получения результата меньше 4 выше, чем вероятность получения результата больше 4.)

б Покажите, что ее средний результат в игре с 4-ударной лункой больше 4. (Таким образом, в среднем она проигрывает "стабильному" игроку по общей сумме очков, хотя в среднем выигрывает в матче с зачетом по лункам.)

Контрольные работы

35 Центр масс игральной кости смещен так, что получается следующее распределение вероятностей:

$$\Pr(⚀) = p_1\,; \quad \Pr(⚁) = p_2\,; \quad \ldots\,; \quad \Pr(⚅) = p_6\,.$$

Пусть S_n обозначает сумму очков после n бросаний этой игральной кости. Найдите необходимые и достаточные условия, которым должно удовлетворять распределение вероятностей, чтобы две случайные величины, $S_n \bmod 2$ и $S_n \bmod 3$, были независимы друг от друга при всех n.

36 На шести гранях некоторой игральной кости вместо обычных ⚀, ..., ⚅ нанесены следующие метки:

а Покажите, что можно так расставить очки на шести гранях второй игральной кости, что сумма очков при бросании этих двух игральных костей будет иметь то же распределение вероятностей, что и при бросании пары обычных игральных костей. (Все 36 комбинаций граней равновероятны.)

б Обобщая предыдущий пункт, найдите все способы расстановки очков на $6n$ гранях n игральных костей, такие, что распределение суммы очков будет совпадать с распределением суммы очков у n обычных игральных костей. (На каждой грани должно быть положительное целое число очков.)

37 Пусть p_n — вероятность того, что правильную монету придется бросить ровно n раз, прежде чем два раза подряд выпадет орел, и пусть $q_n = \sum_{k\geqslant n} p_k$. Найдите аналитическое выражение p_n и q_n через числа Фибоначчи.

38 Чему равна производящая функция случайной величины для количества бросаний правильной игральной кости, требуемых для выпадения всех шести граней? Рассмотрите общий случай правильной "игральной кости" с m гранями и найдите аналитическое выражение для количества бросаний, требуемых для появления l из m граней. Чему равна вероятность, что потребуется ровно n бросаний?

39 *Производящая функция Дирихле случайной величины* имеет вид

$$P(z) = \sum_{n\geqslant 1} \frac{p_n}{n^z}.$$

Таким образом, $P(0) = 1$. Пусть X— случайная величина, для которой $\Pr(X=n) = p_n$. Выразите $E(X)$, $V(X)$ и $E(\ln X)$ через $P(z)$ и ее производные.

40 Для биномиального распределения (8.57) m-й кумулянт κ_m имеет вид $nf_m(p)$, где f_m— полином степени m. (Например, $f_1(p) = p$ и $f_2(p) = p - p^2$, потому что математическое ожидание и дисперсия равны np и npq.)

а Найдите аналитическое выражение для коэффициента при p^k в $f_m(p)$.

б Докажите, что $f_m(\frac{1}{2}) = (2^m - 1)B_m/m + [m=1]$, где B_m— m-е число Бернулли.

41 Пусть случайная величина X_n представляет собой количество бросаний правильной монеты до появления в совокупности n орлов. Покажите, что $E(X_{n+1}^{-1}) = (-1)^n(\ln 2 + H_{\lfloor n/2\rfloor} - H_n)$. Используя методы главы 9, оцените это значение с абсолютной погрешностью $O(n^{-3})$.

Неужели эта вероятность не зависит и от экономического кризиса?
— Переводчик

42 Некоторый человек испытывает трудности с устройством на работу. Если он является безработным утром любого дня, то с некоторой вероятностью p_h, независимой от предыдущих событий, он нанимается на работу в течение дня; но если он имеет работу в начале дня, то с постоянной вероятностью p_f он будет уволен к вечеру. Найдите среднее число вечеров, когда его ожидает работа на следующий день, в предположении, что первоначально он принят на работу и что этот процесс продолжается n дней. (Например, если $n = 1$, то искомый ответ — $1 - p_f$.)

43 Найдите аналитический вид производящей функции случайной величины $G_n(z) = \sum_{k \geqslant 0} p_{k,n} z^k$, где $p_{k,n}$ есть вероятность того, что случайная перестановка n объектов имеет ровно k циклов. Чему равны математическое ожидание и стандартное отклонение количества циклов?

44 Спортивный клуб проводит среди своих членов турнир по теннису с выбыванием. В турнире участвуют 2^n игроков. В первом круге игроки случайным образом разбиваются на пары, так что любое разбиение равновероятно, и играется 2^{n-1} матчей. Победители переходят на следующий круг, где с помощью такого же процесса выявляется 2^{n-2} победителей. И так далее; на k-м круге играется 2^{n-k} случайно выбираемых матчей между 2^{n-k+1} непобежденными игроками. На n-м круге определяется чемпион. Хотя устроители турнира этого и не знают, все игроки в действительности оказались упорядочены по силе игры, так что x_1 играет лучше всех, x_2 второй по силе игры, ..., x_{2^n} играет хуже всех. Когда игрок x_j играет с игроком x_k и $j < k$, то x_j побеждает с вероятностью p и с вероятностью $1 - p$ побеждает x_k, причем результат их игры не зависит от других игр. Мы считаем, что для всех j и k действует одна и та же вероятность p.

Странные, однако, игроки...

а С какой вероятностью игрок x_1 победит в турнире?

б С какой вероятностью в n-м круге (финале) встретятся два лучших игрока — x_1 и x_2?

в С какой вероятностью в k-м от конца круге ($2^{-(k-1)}$-финале) будут соревноваться 2^k лучших игроков? (Предыдущие вопросы представляют собой частные случаи данного вопроса при $k = 0$ и $k = 1$.)

г Пусть $N(n)$ обозначает количество существенно различных результатов турнира; два турнира считаются по сути совпадающими, если в них все матчи играются между одними и теми же игроками с одними и теми же результатами. Докажите, что $N(n) = 2^n!$.

д Какова вероятность того, что турнир выиграет x_2?

е Докажите, что если $\frac{1}{2} < p < 1$, то вероятность выигрыша турнира игроком x_j строго больше, чем игроком x_{j+1} $(1 \leqslant j < 2^n)$.

"Une rapide opération arithmétique montre que, grâce à cette ingénieuse cascade, les xérès ont toujours au moins trois ans. Pousser plus loin le calcul de leur âge donne le vertige".
— Revue du vin de France (Nov 1984)

45 Настоящий херес производят в Испании в соответствии с многостадийной системой, именуемой "Солера". Для простоты будем считать, что у винодела имеется всего три бочки, A, B и C. Каждый год третья часть вина из бочки C бутилируется и заменяется вином из бочки B; затем бочка B доливается доверху вином из бочки A; и наконец, в бочку A доливается новое вино. Пусть $A(z)$, $B(z)$, $C(z)$ — производящие функции случайных величин, в которых коэффициентом при z^n служит доля вина n-летней выдержки в соответствующей бочке сразу же после переливаний.

а Считаем, что описанные операции производятся с незапамятных времен, так что достигнуто установившееся состояние, и производящие функции случайных величин $A(z)$, $B(z)$ и $C(z)$ одни и те же в начале каждого года. Найдите аналитические выражения для этих производящих функций случайных величин.

б При тех же предположениях найдите математическое ожидание и стандартное отклонение возраста вина в каждой бочке. Каков средний возраст бутилируемого вина? Какая доля вина имеет выдержку 25 лет?

в Теперь примите во внимание конечность времени. Пусть во все три бочки в начале нулевого года залили новое вино. Каким будет средний возраст вина, бутилируемого в начале года n?

46 Стефан Банах (Stefan Banach) имел обыкновение носить с собой две коробки спичек, в каждой из которых изначально находилось по n спичек. Когда ему нужен был огонь, он случайным образом с вероятностью $\frac{1}{2}$ независимо от предыдущего выбора выбирал одну из коробок. Взяв из нее спичку, он клал коробку обратно в карман (даже если она была пустой — именно так и поступают великие математики). Если выбранная коробка оказывалась пустой, Банах выбрасывал ее и доставал другую.

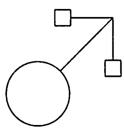
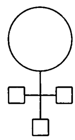
а Однажды он обнаружил, что обе коробки пусты. Какова вероятность такого события? (При $n = 1$ это происходит в половине случаев, а при $n = 2$ — с вероятностью 3/8.) Для ответа на заданный вопрос найдите аналитическое выражение для производящей функции случайной величины $P(w, z) = \sum_{m,n} p_{m,n} w^m z^n$, где $p_{m,n}$ — вероятность, начав с m спичек в одной коробке и n в другой, получить обе пустые коробки, когда первой выбирается пустая коробка. Затем найдите аналитическое выражение для $p_{n,n}$.

б Обобщая ответ к части (a), найдите аналитическое выражение для вероятности того, что в момент выбрасывания первой пустой коробки в другой будет ровно k спичек.

в Найдите выражение для среднего количества спичек в этой другой коробке в аналитическом виде.

А также для количества спичек в пустой коробке.

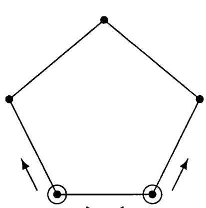
47 Несколько физиологов в сотрудничестве с несколькими физиками открыли недавно две разновидности микробов с очень странным способом размножения. Мужской микроб, *дифаг*, имеет на своей поверхности два рецептора; женский микроб, *трифаг*, имеет три рецептора:

дифаг:  трифаг:  рецептор: □

При облучении культуры дифагов и трифагов ψ-частицами, каждая частица поглощается ровно в одном рецепторе одного из фагов (при этом все рецепторы равновероятны). Если частица попадает в рецептор дифага, то дифаг превращается в трифаг; если же частица попадает в рецептор трифага, то этот трифаг делится на два дифага. Так, если в начале эксперимента имеется один дифаг, то первая ψ-частица вызовет его превращение в трифаг, вторая разделит его на два дифага, третья превратит одного из дифагов в трифаг. Четвертая ψ-частица попадет либо в дифаг, либо в трифаг, и в результате получится либо два трифага (с вероятностью $\frac{2}{5}$), либо три дифага (с вероятностью $\frac{3}{5}$). Найдите аналитическое выражение для среднего числа дифагов, если начать с одного дифага и облучить культуру n одиночными ψ-частицами.

48 Пять человек стоят в вершинах пятиугольника и бросают друг другу летающие диски.

Если этот пятиугольник — Пентагон, они перебрасываются ракетами.



У них имеется два диска, первоначально расположенных в соседних вершинах, как показано на рисунке. В очередной отрезок времени каждый диск бросают либо влево, либо вправо с одинаковой вероятностью. Этот процесс продолжается до тех пор, пока в одного человека не бросят два диска одновременно, и в этом случае игра заканчивается. (Все броски независимы от предыдущих.)

а Найдите математическое ожидание и дисперсию количества пар бросков.

б Найдите аналитическое выражение с использованием чисел Фибоначчи для вероятности того, что игра продлится более 100 шагов.

49 Люк Снегоход проводит зимний отпуск в своей хижине в горах. На передней веранде хижины стоит m пар ботинок, а на задней — n пар. Каждый раз, отправляясь на прогулку, Люк бросает (правильную) монету и решает, идти ли ему с передней или с задней веранды. На выбранной веранде он обувается и выходит из дому. Возвращается он равновероятно к передней или задней веранде с вероятностью $\frac{1}{2}$ независимо от того, с какой веранды началась прогулка. Так, после первой прогулки на передней веранде будет $m + [-1, 0, \text{или} +1]$ пар ботинок, а на задней — $n - [-1, 0, \text{или} +1]$ пар. Если все ботинки окажутся на одной веранде, а Люк соберется выйти с другой, то он пойдет босиком, обморозит ноги и на этом его отпуск заканчивается. Считая, что он продолжает свои прогулки до такого печального конца, обозначим через $P_N(m, n)$ вероятность завершения ровно N прогулок без обморожения, если изначально на передней веранде было m пар, а на задней — n. Тогда, если числа m и n положительны, имеем

$$\begin{aligned} P_N(m,n) \;=\; &\tfrac{1}{4}P_{N-1}(m-1,n+1) + \tfrac{1}{2}P_{N-1}(m,n) + \\ &+ \tfrac{1}{4}P_{N-1}(m+1,n-1)\,, \end{aligned}$$

поскольку первая прогулка является передне-задней, передне-передней, задне-задней или задне-передней с вероятностью $\frac{1}{4}$ в каждом случае, и еще остается $N - 1$ прогулок.

а Дополните рекуррентное соотношение для $P_N(m, n)$ формулами для $m = 0$ или $n = 0$. Используя данное рекуррентное соотношение, найдите уравнения для производящих функций случайных величин

$$g_{m,n}(z) \;=\; \sum_{N \geqslant 0} P_N(m,n) z^N\,.$$

б Продифференцируйте свои уравнения и положите в них $z = 1$, получив тем самым соотношения относительно величин $g'_{m,n}(1)$. Решите полученные уравнения и найдите среднее количество прогулок до обморожения.

в Покажите, что для $g_{m,n}$ имеется выражение в аналитическом виде, которое можно получить, заменяя $z = 1/\cos^2\theta$:

$$g_{m,n}\Big(\frac{1}{\cos^2\theta}\Big) = \frac{\sin(2m+1)\theta + \sin(2n+1)\theta}{\sin(2m+2n+2)\theta}\cos\theta\,.$$

50 Рассмотрим функцию

$$H(z) = 1 + \frac{1-z}{2z}\big(z - 3 + \sqrt{(1-z)(9-z)}\,\big)\,.$$

Цель данного упражнения — доказать, что $H(z)=\sum_{k\geqslant 0} h_k z^k$ является производящей функцией случайной величины, и выяснить некоторые ее фундаментальные свойства.

а Положим $(1-z)^{3/2}(9-z)^{1/2} = \sum_{k\geqslant 0} c_k z^k$. Докажите, что $c_0 = 3$, $c_1 = -14/3$, $c_2 = 37/27$ и $c_{3+l} = 3\sum_k \binom{l}{k} \times \binom{1/2}{3+k}\big(\frac{8}{9}\big)^{k+3}$ для всех $l \geqslant 0$. *Указание*: воспользуйтесь тождеством

$$(9-z)^{1/2} = 3(1-z)^{1/2}\big(1 + \tfrac{8}{9}z/(1-z)\big)^{1/2}$$

и разложите последний множитель по степеням $z/(1-z)$.

б Используя решение п. (а) и упр. 5.81, покажите, что все коэффициенты функции $H(z)$ положительны.

в Докажите удивительное тождество

$$\sqrt{\frac{9-H(z)}{1-H(z)}} = \sqrt{\frac{9-z}{1-z}} + 2\,.$$

г Чему равны математическое ожидание и дисперсия H?

51 В государственной лотерее в Эльдорадо используется описанное в предыдущей задаче распределение выигрышей H. Каждый лотерейный билет стоит 1 дублон, а выплата по этому билету составляет k дублонов с вероятностью h_k. Шансы на выигрыш по каждому билету никоим образом не зависят от выигрышей по другим билетам; иными словами, выигрыш или проигрыш по одному билету не оказывает влияния на вероятность выигрыша по любому другому билету, который вы, возможно, приобрели для участия в том же тираже.

"...у меня сдали нервы, когда я убедился, что Федеральных Штатов Эльдорадо нет. На Земле, я имею в виду."
— *Р. Артур (R. Jay Arthur), "Марки страны Эльдорадо"*
(Добав. пер.)

а Предположим, что вы начинаете игру, имея один дублон. Если вы выигрываете k дублонов, то покупаете k билетов следующего тиража. Затем весь выигрыш во второй игре целиком вкладывается в третий тираж, и т.д. Если не выиграл ни один билет, игра вынужденно прекращается. Докажите, что производящая функция случайной величины вашей суммы после n кругов игры имеет вид

$$1-\frac{4}{\sqrt{(9-z)/(1-z)}+2n-1}+\frac{4}{\sqrt{(9-z)/(1-z)}+2n+1}\,.$$

б Пусть g_n — вероятность того, что вы впервые проигрываете все свои деньги на n-й игре, и пусть $G(z) = g_1z+g_2z^2+\cdots$. Докажите, что $G(1) = 1$. (Это означает, что рано или поздно вы все равно проиграете с вероятностью 1, хотя в процессе игры можете временно стать миллиардером.) Чему равны математическое ожидание и дисперсия G?

в Сколько в среднем билетов вы купите, если продолжите играть до полного разорения?

г Чему равно среднее количество игр до разорения, если вначале у вас был не один дублон, а два?

Дубл 2 он.

Дополнительные задачи

52 Покажите, что определенные в тексте медиана и мода случайных величин соответствуют в некотором разумном смысле медиане и моде последовательностей, если пространство вероятностей конечно.

53 Докажите или опровергните: если X, Y и Z — случайные величины, обладающие тем свойством, что все три пары, (X,Y), (X,Z) и (Y,Z), независимы, то независимы $X+Y$ и Z.

54 Уравнение (8.20) доказывает, что среднее значение $\widehat{V}X$ равно VX. Чему равна *дисперсия* $\widehat{V}X$?

55 В обычной колоде игральных карт 52 карты, по четыре карты каждого из возможных значений $\{\mathtt{A},2,3,4,5,6,7,8,9,10,\mathtt{J},\mathtt{Q},\mathtt{K}\}$. Пусть X и Y обозначают соответственно значения верхней и нижней карт в колоде. Рассмотрим следующий алгоритм тасования.

A, J, Q и K означают туза, валета, даму и короля соответственно. — Переводчик

S1 Перемешать колоду случайным образом, так, чтобы любая перестановка появлялась с вероятностью $1/52!$.

S2 Если $X \neq Y$, подбросить несимметричную монету с вероятностью выпадения орла p, и вернуться к шагу S1, если выпал орел. В противном случае остановиться.

Все бросания монеты и все перемешивания колоды предполагаются независимыми. При каком значении p после остановки описанной процедуры значения X и Y будут независимыми случайными величинами?

56 Обобщите задачу о дисках из упр. 48 на случай m-угольника. Чему в общем случае равны математическое ожидание и дисперсия количества бросаний без коллизий, если первоначально диски находятся в соседних вершинах? Покажите, что если m нечетное, то производящая функция случайной величины для числа бросаний может быть записана в виде произведения распределений для бросания монеты:

$$G_m(z) = \prod_{k=1}^{(m-1)/2} \frac{p_k z}{1 - q_k z},$$

$$\text{где} \;\; p_k = \sin^2 \frac{(2k-1)\pi}{2m}, \quad q_k = \cos^2 \frac{(2k-1)\pi}{2m}.$$

Указание: попробуйте выполнить подстановку $z = 1/\cos^2 \theta$.

57 Докажите, что в игре Пенни последовательность $\tau_1\tau_2 \ldots \tau_{l-1}\tau_l$ всегда уступает последовательности $\bar{\tau}_2\tau_1\tau_2 \ldots \tau_{l-1}$, если $l \geqslant 3$, а монета правильная.

58 Существует ли последовательность $A = \tau_1\tau_2 \ldots \tau_{l-1}\tau_l$ из $l \geqslant 3$ орлов и решек, такая, что и последовательность $O\tau_1\tau_2 \ldots \tau_{l-1}$, и последовательность $P\tau_1\tau_2 \ldots \tau_{l-1}$ дают одинаковые результаты при сравнении с A в игре Пенни?

59 Существуют ли последовательности A и B из орлов и решек, такие, что A длиннее B, но A встречается ранее, чем B, более чем в половине случаев при бросании правильной монеты?

60 Пусть k и n — фиксированные положительные целые числа и $k < n$.

а Найдите аналитический вид производящей функции случайной величины

$$G(w, z) = \frac{1}{m^n} \sum_{h_1=1}^{m} \cdots \sum_{h_n=1}^{m} w^{P(h_1,\ldots,h_n,k)} z^{P(h_1,\ldots,h_n,n}$$

для совместного распределения количества проб, требуемых для поиска k-го и n-го элементов, вставленных в хеш-таблицу с m списками.

б Хотя случайные величины $P(h_1, \ldots, h_n; k)$ и $P(h_1, \ldots, h_n; n)$ зависимы, покажите, что в некотором смысле они независимы:

$$E\big(P(h_1, \ldots, h_n; k) P(h_1, \ldots, h_n; n)\big) = \\ = \big(EP(h_1, \ldots, h_n; k)\big)\big(EP(h_1, \ldots, h_n; n)\big).$$

61 Используя результат предыдущего упражнения, докажите (8.104).

62 Продолжая упр. 47, найдите *дисперсию* количества дифагов после облучения n ψ-частицами.

Исследовательская проблема

63 *Нормальное распределение* представляет собой недискретное распределение вероятностей, характеризуемое тем свойством, что все его кумулянты равны нулю, за исключением математического ожидания и дисперсии. Имеется ли простой способ, позволяющий определить по заданной последовательности кумулянтов $\langle \kappa_1, \kappa_2, \kappa_3, \ldots \rangle$, не проистекает ли она из *дискретного* распределения? (Все вероятности в дискретном распределении должны быть "атомарными".)

9

Асимптотика

ТОЧНЫЕ ОТВЕТЫ, если их удается найти, — это замечательно; они вызывают чувство глубокого удовлетворения и повышают самооценку. Но не относитесь к приближенным значениям свысока — они тоже могут пригодиться. Имея дело с некоторой суммой или рекуррентным соотношением, которое, как нам кажется, невозможно представить в аналитическом виде, мы наверняка предпочтем получить хотя бы приближенный ответ, чем остаться вовсе без информации о поведении функции или суммы. Да и в случае, когда у нас имеется аналитическое решение, наши знания могут оказаться неполными, поскольку не всегда ясно, как сравнить полученный ответ с другими формулами в аналитическом виде.

Например, аналитического выражения для суммы

$$S_n = \sum_{k=0}^{n} \binom{3n}{k},$$

похоже, не существует. Приятно, однако, узнать, что

$$S_n \sim 2\binom{3n}{n} \qquad \text{при } n \to \infty;$$

мы говорим, что сумма “асимптотически стремится к” $2\binom{3n}{n}$. Еще приятнее получить более детальную информацию наподобие

$$S_n = \binom{3n}{n}\left(2 - \frac{4}{n} + O\left(\frac{1}{n^2}\right)\right), \tag{9.1}$$

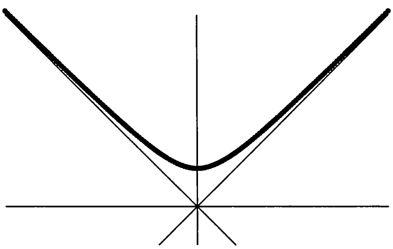
которая дает нам “относительную ошибку порядка $1/n^2$”. Но даже этого недостаточно для того, чтобы сказать, как S_n соотносится с другими величинами. Что больше — S_n или число Фибоначчи F_{4n}? Ответ: при $n = 2$ имеем $S_2 = 22 > F_8 = 21$, но в конечном счете F_{4n} больше, потому что $F_{4n} \sim \phi^{4n}/\sqrt{5}$ и $\phi^4 \approx 6.8541$, в то время как

$$S_n = \sqrt{\frac{3}{\pi n}}(6.75)^n\left(1 - \frac{151}{72n} + O\left(\frac{1}{n^2}\right)\right). \qquad (9.2)$$

Наша цель в данной главе — научиться понимать и получать подобные результаты без чрезмерных усилий.

Слово *асимптотика* имеет греческое происхождение и означает "не сходящиеся вместе". Изучая конические сечения, древнегреческие математики рассматривали, в частности, гиперболы наподобие графика функции $y = \sqrt{1 + x^2}$,

От того же греческого корня происходят такие слова, как "симптом" и "птомаин".



для которого прямые $y = x$ и $y = -x$ являются "асимптотами". При $x \to \infty$ кривая бесконечно приближается к асимптотам, но никогда не соприкасается с ними. В наши дни термин "асимптотика" имеет более широкий смысл и означает любую приближенную величину, которая становится все более и более точной по мере приближения некоторого параметра к предельному значению. Для нас "асимптотика" означает "почти соприкасающиеся".

Вывод некоторых асимптотических формул очень сложен и не укладывается в рамки данной книги. Мы удовольствуемся введением в эту тему, полагая, что на данном базисе можно будет построить более сложные методы. В особенности нас будут интересовать определения символов '$\sim$', 'O' и им подобных; мы также изучим основные способы преобразования асимптотических соотношений.

9.1 Иерархия

Встречающиеся на практике функции от n имеют разные "асимптотические скорости роста"; одни из них стремятся к бесконечности быстрее других. Формализуем это понятие при помощи определения

$$f(n) \prec g(n) \iff \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0. \qquad (9.3)$$

Ловись, функция, большая и маленькая.

Это отношение транзитивно: если $f(n) \prec g(n)$ и $g(n) \prec h(n)$, то $f(n) \prec h(n)$. Можно также записать $g(n) \succ f(n)$, если $f(n) \prec g(n)$. Эти обозначения были введены в 1871 году Полем Дюбуа-Раймоном (Paul du Bois-Reymond) [85].

Например, $n \prec n^2$; неформально мы говорим, что n растет медленнее, чем n^2. Фактически

$$n^\alpha \prec n^\beta \iff \alpha < \beta, \tag{9.4}$$

для произвольных действительных чисел α и β.

Конечно, помимо степенных, имеется множество иных функций. Отношение $\prec$ можно использовать, чтобы упорядочить множество функций. Каждая функция будет "клевать" другую в соответствии с асимптотическим ростом. В этой цепочке могут встретиться, например, такие члены:

$$1 \prec \log\log n \prec \log n \prec n^\epsilon \prec n^c \prec n^{\log n} \prec c^n \prec n^n \prec c^{c^n}.$$

(Здесь ϵ и c — произвольные константы, для которых выполняется условие $0 < \epsilon < 1 < c$.)

Все перечисленные выше функции, за исключением 1, стремятся к бесконечности при $n \to \infty$. Таким образом, пытаясь вставить в данную иерархию новую функцию, мы будем выяснять не то, *стремится ли* она к бесконечности, а *насколько быстро* она это делает.

Занимаясь асимптотическим анализом, не следует мелочиться: работая со стремящейся к бесконечности переменной, следует ДУМАТЬ ПО-ВОЛЬШОМУ. Например, из приведенной иерархии вытекает, что $\log n \prec n^{0.0001}$; может показаться, что это не так, если ограничиться крохотными числами вроде гугола $n = 10^{100}$. В этом случае $\log n = 100$, в то время как $n^{0.0001}$ равно всего лишь $10^{0.01} \approx 1.0233$. Но если рассмотреть гуголплекс $n = 10^{10^{100}}$, то $\log n = 10^{100}$ бледнеет в сравнении с $n^{0.0001} = 10^{10^{96}}$.

Даже если ϵ чрезвычайно мало (меньше, чем, скажем, $1/10^{10^{100}}$), значение $\log n$ будет гораздо меньше значения n^ϵ, если взять n достаточно большим. Действительно, если положить $n = 10^{10^{2k}}$, где k достаточно велико, чтобы выполнялось условие $\epsilon \geqslant 10^{-k}$, то мы получим $\log n = 10^{2k}$, но $n^\epsilon \geqslant 10^{10^k}$. Отношение $(\log n)/n^\epsilon$, таким образом, стремится к нулю при $n \to \infty$.

Малорархию?

Приведенная выше иерархия относится к функциям, стремящимся к бесконечности. Однако зачастую нас интересуют функции, стремящиеся к нулю, так что было бы неплохо иметь аналогичную иерархию и для бесконечно малых функций. Построить ее легко, потому что если $f(n)$ и $g(n)$ никогда не обращаются в нуль, то

$$f(n) \prec g(n) \iff \frac{1}{g(n)} \prec \frac{1}{f(n)}. \qquad (9.5)$$

Так, например, все приведенные далее функции стремятся к нулю (за исключением 1):

$$\frac{1}{c^{c^n}} \prec \frac{1}{n^n} \prec \frac{1}{c^n} \prec \frac{1}{n^{\log n}} \prec \frac{1}{n^c} \prec \frac{1}{n^\epsilon} \prec \frac{1}{\log n} \prec \frac{1}{\log\log n} \prec 1.$$

Давайте рассмотрим еще несколько функций и попробуем вставить их в иерархию. Число $\pi(n)$ простых чисел, не превосходящих n, как известно, приближенно равно $n/\ln n$. Поскольку $1/n^\epsilon \prec 1/\ln n \prec 1$, умножение на n дает

$$n^{1-\epsilon} \prec \pi(n) \prec n.$$

Соотношение (9.4) можно обобщить, если заметить, например, что

$$n^{\alpha_1}(\log n)^{\alpha_2}(\log\log n)^{\alpha_3} \prec n^{\beta_1}(\log n)^{\beta_2}(\log\log n)^{\beta_3} \iff$$
$$\iff (\alpha_1,\alpha_2,\alpha_3) < (\beta_1,\beta_2,\beta_3). \qquad (9.6)$$

Здесь '$(\alpha_1,\alpha_2,\alpha_3) < (\beta_1,\beta_2,\beta_3)$' означает лексикографический порядок (порядок слов в словаре); другими словами, это неравенство означает, что либо $\alpha_1 < \beta_1$, либо $\alpha_1 = \beta_1$ и $\alpha_2 < \beta_2$, либо $\alpha_1 = \beta_1$ и $\alpha_2 = \beta_2$ и $\alpha_3 < \beta_3$.

А что можно сказать о функции $e^{\sqrt{\log n}}$? Где ее место в иерархии? Ответить на подобные вопросы можно с помощью правила

$$e^{f(n)} \prec e^{g(n)} \iff \lim_{n\to\infty}\bigl(f(n)-g(n)\bigr) = -\infty, \qquad (9.7)$$

которое в два шага следует из определения (9.3) путем логарифмирования. Следовательно,

$$1 \prec f(n) \prec g(n) \implies e^{|f(n)|} \prec e^{|g(n)|}.$$

А поскольку $1 \prec \log\log n \prec \sqrt{\log n} \prec \epsilon\log n$, получаем $\log n \prec e^{\sqrt{\log n}} \prec n^\epsilon$.

Если две функции, $f(n)$ и $g(n)$, имеют *одинаковую* скорость роста, мы записываем '$f(n) \asymp g(n)$'. "Официальное" определение в данном случае таково:

$$f(n) \asymp g(n) \iff |f(n)| \leqslant C|g(n)| \text{ и } |g(n)| \leqslant C|f(n)|$$
для некоторого C и всех достаточно больших n. (9.8)

Это имеет место, например, если $f(n) = \cos n + \operatorname{arctg} n$, а $g(n)$ — ненулевая константа. Вскоре мы докажем справедливость данного соотношения для случая, когда $f(n)$ и $g(n)$ представляют собой полиномы одинаковых степеней. Имеется и более сильное соотношение, определяемое правилом

$$f(n) \sim g(n) \iff \lim_{n\to\infty} \frac{f(n)}{g(n)} = 1. \tag{9.9}$$

В этом случае мы говорим, что "$f(n)$ есть асимптотика для $g(n)$".

Г. Г. Харди (G. H. Hardy) [179] ввел интересную и важную концепцию, называемую классом *логарифмически-экспоненциальных функций*, рекурсивно определяемым как наименьшее семейство $\mathfrak{L}$ функций, удовлетворяющих следующим свойствам.

- Постоянная функция $f(n) = \alpha$ принадлежит $\mathfrak{L}$ для всех действительных α.
- Тождественная функция $f(n) = n$ принадлежит $\mathfrak{L}$.
- Если $f(n)$ и $g(n)$ принадлежат $\mathfrak{L}$, то $f(n) - g(n)$ также принадлежит $\mathfrak{L}$.
- Если $f(n)$ принадлежит $\mathfrak{L}$, то $\mathfrak{L}$ принадлежит и $e^{f(n)}$.
- Если $f(n)$ принадлежит $\mathfrak{L}$ и является "существенно положительной", то $\ln f(n)$ принадлежит $\mathfrak{L}$.

Функция $f(n)$ называется существенно положительной, если существует такое число n_0, что $f(n) > 0$ при $n \geqslant n_0$.

Эти правила можно использовать, например, для того, чтобы показать, что $f(n) + g(n)$ принадлежит $\mathfrak{L}$, если $\mathfrak{L}$ принадлежат $f(n)$ и $g(n)$, потому что $f(n) + g(n) = f(n) - \bigl(0 - g(n)\bigr)$. Если $f(n)$ и $g(n)$ — существенно положительные члены $\mathfrak{L}$, то их произведение $f(n)g(n) = e^{\ln f(n) + \ln g(n)}$ и частное $f(n)/g(n) = e^{\ln f(n) - \ln g(n)}$ также принадлежат $\mathfrak{L}$; то же справедливо и для функций наподобие $\sqrt{f(n)} = e^{\frac{1}{2}\ln f(n)}$ и т.п. Харди доказал, что всякая логарифмически-экспоненциальная функция является существенно положительной, существенно отрицательной или равной нулю. Следовательно, произведение и частное любых двух $\mathfrak{L}$-функций принадлежат $\mathfrak{L}$, за исключением случая деления на тождественно нулевую функцию.

Основная теорема Харди о логарифмически-экспоненциальных функциях заключается в том, что эти функции образуют асимптотическую иерархию: *если* $f(n)$ *и* $g(n)$ — *любые функции из* $\mathfrak{L}$, *то либо* $f(n) \prec g(n)$, *либо* $f(n) \succ g(n)$, *либо* $f(n) \asymp g(n)$. *В последнем случае существует константа* α, *такая, что*

$$f(n) \sim \alpha\, g(n).$$

Доказательство теоремы Харди выходит за рамки данной книги; но полезно знать о существовании этой теоремы, потому что почти все функции, с которыми нам придется встречаться, принадлежат $\mathfrak{L}$. На практике в большинстве случаев можно без особого труда вставить заданную функцию в иерархию.

9.2 O-обозначения

В 1894 году Пауль Бахман (Paul Bachmann) придумал замечательное обозначение для асимптотического анализа. В последующие годы его популярности немало поспособствовали Эдмунд Ландау (Edmund Landau) и др. Мы встречали его в формулах наподобие

$$H_n = \ln n + \gamma + O(1/n)\,, \tag{9.10}$$

которая говорит нам, что n-е гармоническое число равно натуральному логарифму n плюс константа Эйлера, плюс некоторая величина, которая составляет "О большое от одной n-й". Эта последняя величина точно не определена, однако какой бы она ни была, обозначение позволяет утверждать, что она не превосходит некоторую константу, умноженную на $1/n$.

"...wenn wir durch das Zeichen $O(n)$ eine Grösse ausdrücken, deren Ordnung in Bezug auf n die Ordnung von n nicht überschreitet; ob sie wirklich Glieder von der Ordnung n in sich enthält, bleibt bei dem bisherigen Schlussverfahren dahingestellt."
— *П. Бахман (P. Bachmann) [17]*

Красота O-обозначений в том, что они скрывают несущественные детали и позволяют сконцентрироваться на главных особенностях поведения: величину $O(1/n)$ можно считать пренебрежимо малой, если только нас не интересуют величины, отличающиеся от $1/n$ постоянным множителем.

Более того, символ 'O' удобно использовать внутри формул. Если мы хотим выразить соотношение (9.10) с применением обозначений из раздела 9.1, то нам надо перенести $\ln n + \gamma$ в левую часть и указать более слабый результат, например

$$H_n - \ln n - \gamma \prec \frac{\log \log n}{n},$$

или более строгий результат

$$H_n - \ln n - \gamma \asymp \frac{1}{n}.$$

При помощи O можно указывать необходимое количество деталей на месте, без перестановки членов.

Идея неточно определенных величин может стать более понятной, если мы рассмотрим некоторые другие примеры. Иногда мы пишем '± 1' для обозначения чего-то, что есть либо $+1$, либо -1; мы не знаем (или нам не важно), чему именно равна эта величина; тем не менее мы можем работать с ней в формулах.

Н. Г. де Брейн (N. G. de Bruijn) начал свою книгу *Асимптотические методы в анализе* [74] с рассмотрения обозначения "большого L", которое помогает понять "большое O". Если записать $L(5)$ для обозначения числа, абсолютное значение которого меньше 5 (не говоря, что именно это за число), можно выполнять определенные вычисления, не обладая полной информацией. Например, можно вывести формулы наподобие $1 + L(5) = L(6)$; $L(2) + L(3) = L(5)$; $L(2)L(3) = L(6)$; $e^{L(5)} = L(e^5)$ и т.д. Но нельзя заключить, например, что $L(5)-L(3) = L(2)$, поскольку левая часть может оказаться $4 - 0$. Фактически лучшее, что можно сказать в этом случае,— это $L(5) - L(3) = L(8)$.

Оно не бессмысленно, а просто бесполезно.

Обозначение 'O' Бахмана похоже на L-обозначение, но оно еще менее точное: $O(\alpha)$ означает число, абсолютное значение которого не превосходит некоторую константу, умноженную на $|\alpha|$. Мы не говорим, что это за число и даже что это за константа. Разумеется, понятие "константы" бессмысленно, если нет чего-либо переменного; так что мы используем O-обозначение только в контексте, где имеется по крайней мере одна величина (скажем, n), значение которой меняется. Формула

$$f(n) = O\bigl(g(n)\bigr) \qquad \text{для всех } n \tag{9.11}$$

в этом контексте означает, что существует такая константа C, что

$$\bigl|f(n)\bigr| \leqslant C\bigl|g(n)\bigr| \qquad \text{для всех } n; \tag{9.12}$$

а если обозначение $O\bigl(g(n)\bigr)$ использовано внутри формулы, то оно означает функцию $f(n)$, удовлетворяющую условию (9.12). Значения $f(n)$ неизвестны, но мы знаем, что они не слишком велики. Аналогично введенная выше величина $L(n)$ представляет неопределенную функцию $f(n)$, значения которой удовлетворяют неравенству $\bigl|f(n)\bigr| < |n|$. Основное отличие между L и O заключается в том, что O-обозначение включает неопределенную константу C; каждое вхождение O может подразумевать свое значение C, но все эти C не зависят от n.

Например, мы знаем, что сумма первых n квадратов равна

$$\Box_n = \tfrac{1}{3}n(n+\tfrac{1}{2})(n+1) = \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n.$$

Можно записать

$$\Box_n = O(n^3),$$

потому что $|\frac{1}{3}n^3+\frac{1}{2}n^2+\frac{1}{6}n| \leqslant \frac{1}{3}|n|^3+\frac{1}{2}|n|^2+\frac{1}{6}|n| \leqslant \frac{1}{3}|n^3|+\frac{1}{2}|n^3|+\frac{1}{6}|n^3| = |n^3|$ для всех целых n. Аналогично можно получить более точную формулу

$$\square_n = \tfrac{1}{3}n^3 + O(n^2);$$

можно также небрежно отбросить часть информации и сказать, что

$$\square_n = O(n^{10}).$$

Ничто в определении O не требует от нас давать наилучшую возможную оценку.

Но минутку! А что если переменная n не целочисленная? Что если у нас имеется формула наподобие $S(x) = \frac{1}{3}x^3+\frac{1}{2}x^2+\frac{1}{6}x$, где x — действительное число? Тогда мы не можем сказать, что $S(x) = O(x^3)$, потому что отношение $S(x)/x^3 = \frac{1}{3} + \frac{1}{2}x^{-1} + \frac{1}{6}x^{-2}$ становится неограниченным при $x \to 0$. Нельзя также сказать, что $S(x) = O(x)$, потому что отношение $S(x)/x = \frac{1}{3}x^2 + \frac{1}{2}x + \frac{1}{6}$ неограниченно растет при $x \to \infty$. Так что мы, судя по всему, не можем использовать O-обозначения для оценки $S(x)$.

Эта дилемма разрешается путем наложения определенных условий на переменные, используемые с O. Например, если поставить условие $|x| \geqslant 1$ или $x \geqslant \epsilon$, где ϵ — произвольная положительная константа, или что x — целое число, то можно записать $S(x) = O(x^3)$. Если же наложить условие $|x| \leqslant 1$ или $|x| \leqslant c$, где c — произвольная положительная константа, то можно записать $S(x) = O(x)$. Обозначение 'O' зависит от контекста, от ограничений, наложенных на используемые переменные.

Эти ограничения часто задаются в виде предельных соотношений. Например, мы могли бы сказать, что

$$f(n) = O\big(g(n)\big) \qquad \text{при } n \to \infty. \tag{9.13}$$

Это означает, что условия, накладываемые обозначением 'O', предполагаются выполненными, когда n "близко" к ∞; нас не волнует, что происходит при не слишком больших n. Более того, мы даже не определяем, что значит "близко"; в такой ситуации каждое вхождение O подразумевает существование *двух* констант C и n_0, таких, что

$$\big|f(n)\big| \leqslant C\big|g(n)\big| \qquad \text{при } n \geqslant n_0. \tag{9.14}$$

Ты лучше всех!
Я так тебя люблю!
Я люблю, как единица на f, при f стремящемся к нулю.
Сверху.

Значения C и n_0 могут быть разными для разных O, но они не зависят от n. Аналогично запись

$$f(x) = O\big(g(x)\big) \qquad \text{при } x \to 0$$

означает, что существуют две константы, C и ϵ, такие, что

$$\bigl|f(x)\bigr| \leqslant C\bigl|g(x)\bigr| \qquad \text{при } |x| \leqslant \epsilon. \tag{9.15}$$

Предельное значение не обязано быть равным ∞ или 0; можно записать

$$\ln z = z - 1 + O\bigl((z-1)^2\bigr) \qquad \text{при } z \to 1,$$

потому что можно доказать, что $|\ln z - z + 1| \leqslant |z-1|^2$, если $|z-1| \leqslant \frac{1}{2}$.

Мы постепенно, на протяжении нескольких страниц развили наше определение O, начав со вполне очевидных вещей, и получили в итоге нечто, представляющееся ужасно сложным; теперь у нас O представляет неопределенную функцию и одну или две неопределенные константы, зависящие от контекста. Все это может показаться слишком сложным для сколь-нибудь разумного обозначения, но это еще не все! Есть еще одна тонкость, пока что ускользнувшая от нас. А именно, следует осознавать, что запись

$$\tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n = O(n^3)$$

вполне корректна, но в этом равенстве *нельзя* менять местами левую и правую части. В противном случае можно прийти к нелепым выводам наподобие $n = n^2$, исходя из верных тождеств $n = O(n^2)$ и $n^2 = O(n^2)$. При работе с O-обозначением и любыми другими формулами с не полностью определенными величинами мы имеем дело с *односторонними равенствами*. Правая часть уравнения содержит не больше информации, чем левая, и фактически может содержать меньше информации; правая часть представляет собой "огрубление" левой.

"And to auoide the tediouse repetition of these woordes: is equalle to: I will sette as I doe often in woorke use, a paire of paralleles, or Gemowe lines of one lengthe, thus: ═══ , bicause noe .2. thynges, can be moare equalle".
— Р. Рекорде (R. Recorde) [305]

Со строго формальной точки зрения запись '$O\bigl(g(n)\bigr)$' означает не единственную функцию $f(n)$, а *множество* всех функций $f(n)$, таких, что $\bigl|f(n)\bigr| \leqslant C\bigl|g(n)\bigr|$ для некоторой константы C. Обычная формула $g(n)$, не включающая символа 'O', обозначает множество, состоящее из одной функции $f(n) = g(n)$. Если S и T представляют собой множества функций от n, то запись '$S+T$' обозначает множество всех функций вида $f(n)+g(n)$, где $f(n) \in S$ и $g(n) \in T$; другие обозначения наподобие $S-T$, ST, S/T, $\sqrt{S}$, e^S, $\ln S$ определяются аналогично. Тогда "равенство" между двумя такими множествами функций есть, строго говоря, *теоретико-множественное включение*; знак '=' на самом деле означает '$\subseteq$'. Эти формальные определения подводят под все наши манипуляции с O твердый логический фундамент.

Например, "уравнение"

$$\tfrac{1}{3}n^3 + O(n^2) \;=\; O(n^3)$$

означает, что $S_1 \subseteq S_2$, где S_1 представляет собой множество всех функций вида $\frac{1}{3}n^3 + f_1(n)$, для которых найдется константа C_1, такая, что $\bigl|f_1(n)\bigr| \leqslant C_1|n^2|$, и где S_2 есть множество всех функций $f_2(n)$, для которых найдется константа C_2, такая, что $\bigl|f_2(n)\bigr| \leqslant C_2|n^3|$. Можно строго доказать это "равенство", взяв произвольный элемент из левой части и показав, что он принадлежит правой части: пусть $\frac{1}{3}n^3 + f_1(n)$ таково, что $\bigl|f_1(n)\bigr| \leqslant C_1|n^2|$; мы должны доказать, что существует такая константа C_2, что $|\frac{1}{3}n^3 + f_1(n)| \leqslant C_2|n^3|$. Константа $C_2 = \frac{1}{3} + C_1$ решает проблему, поскольку $n^2 \leqslant |n^3|$ для всех целых n.

Если '$=$' на самом деле означает '$\subseteq$', то почему мы не используем '$\subseteq$' вместо некорректного знака равенства? Тому есть четыре причины.

Во-первых, традиция. Начали использовать символ 'O' со знаком равенства теоретики-числовики, и эта практика получила распространение. Мы давно убедились, что пытаться изменить привычки математического сообщества — безнадежное занятие.

Во-вторых, традиция. Программисты привыкли к неверному применению знака равенства — многие годы при программировании на FORTRAN и BASIC программисты записывали присваивания вида '$N = N + 1$'. Одной некорректной трактовкой больше, одной меньше — какая разница!

В-третьих, традиция. Мы часто читаем '$=$' как слово 'есть'. Например, формулу $H_n = O(\log n)$ можно произнести как "H_n есть O большое от $\log n$". Но слово 'есть' несимметричное. Мы говорим, что птица есть животное, но не можем сказать, что животное есть птица; понятие "животное" есть огрубление понятия "птица".

В-четвертых, для наших целей такое обозначение весьма естественно. Если ограничиться только ситуациями, в которых символ 'O' полностью занимает правую часть формулы — как в аппроксимации гармонических чисел $H_n = O(\log n)$ или в описании времени работы алгоритма сортировки $T(n) = O(n \log n)$, — то не важно, какой знак используется, '$=$' или какой-то иной. Однако если использовать O внутри выражений, как часто делается в асимптотическом анализе, то нашей интуиции отвечает трактовка знака '$=$' как равенства, а величи́ны наподобие $O(1/n)$ — как чего-то очень маленького.

Поэтому мы и дальше будем использовать '$=$' и считать, что $O\bigl(g(n)\bigr)$ обозначает не полностью определенную функцию, имея

"Очевидно, что знак «$=$» не подходит для таких отношений, поскольку он подразумевает симметрию, которой на самом деле нет. [...] Однако после такого предупреждения использование знака «$=$» не принесет большого вреда, и мы будем использовать его просто потому, что так принято".
— Н. Г. де Брейн (N. G. de Bruijn) [74]

в виду, что в случае необходимости мы всегда сможем вернуться к теоретико-множественным определениям.

Всесторонне рассматривая определение, мы обязаны отметить еще одну техническую деталь. Если в формуле используется несколько переменных, то символ O представляет множество функций от двух или более переменных, а не только от одной. В область определения каждой функции входят все переменные, которые в данном контексте "свободны" для изменений.

Здесь есть определенная тонкость ввиду того, что переменные могут иметь смысл только в части выражения, если они связаны знаком $\sum$ или чем-то подобным. Например, посмотрим внимательно на соотношение

$$\sum_{k=0}^{n}\bigl(k^2+O(k)\bigr) \;=\; \tfrac{1}{3}n^3+O(n^2)\,, \qquad \text{целое } n \geqslant 0. \tag{9.16}$$

Выражение $k^2+O(k)$ в левой части отвечает множеству всех функций от двух переменных вида $k^2+f(k,n)$, для которых найдется константа C, такая, что $\bigl|f(k,n)\bigr| \leqslant Ck$ для $0 \leqslant k \leqslant n$. Сумма таких множеств функций для $0 \leqslant k \leqslant n$ есть множество всех функций $g(n)$ вида

$$\sum_{k=0}^{n}\bigl(k^2+f(k,n)\bigr) \;=\; \tfrac{1}{3}n^3+\tfrac{1}{2}n^2+\tfrac{1}{6}n+f(0,n)+f(1,n)+\cdots+f(n,n)\,,$$

где f обладает указанным свойством. Поскольку

$$\begin{aligned}\bigl|\tfrac{1}{2}n^2+\tfrac{1}{6}n+f(0,n)+f(1,n)+\cdots+f(n,n)\bigr| &\leqslant \\ \leqslant\; \tfrac{1}{2}n^2+\tfrac{1}{6}n^2+C\cdot 0+C\cdot 1+\cdots+C\cdot n \;&<\\ <\; n^2+C(n^2+n)/2 \;<\; (C+1)n^2\,,&\end{aligned}$$

Самое время заняться разминочными упражнениями 3 и 4.

все такие функции $g(n)$ принадлежат правой части (9.16); следовательно, (9.16) справедливо.

Иногда O-обозначения понимают неправильно, считая, что символ 'O' указывает точный порядок роста, и используют его так, как будто наряду с верхней границей он указывает и нижнюю. Так, можно услышать, что некоторый алгоритм сортировки n чисел неэффективен, "потому что время его работы — $O(n^2)$". Но время работы $O(n^2)$ не означает, что это время не есть одновременно $O(n)$. Для указания нижней границы имеется другое обозначение — "большая «омега»":

$$f(n) \;=\; \Omega\bigl(g(n)\bigr) \quad\Longleftrightarrow\quad \bigl|f(n)\bigr| \;\geqslant\; C\bigl|g(n)\bigr| \quad \text{для некоторого } C>0. \tag{9.17}$$

Мы имеем $f(n) = \Omega(g(n))$ тогда и только тогда, когда $g(n) = O(f(n))$. При достаточно больши́х n алгоритм сортировки со временем работы $\Omega(n^2)$ неэффективен по сравнению с алгоритмом, время работы которого составляет $O(n \log n)$.

Наконец, символ "большая «тета»" указывает точный порядок роста:

$$f(n) = \Theta(g(n)) \iff \begin{array}{c} f(n) = O(g(n)) \\ \text{и} \quad f(n) = \Omega(g(n)) . \end{array} \qquad (9.18)$$

Поскольку Ω и Θ — заглавные греческие буквы, O на самом деле должно быть заглавной греческой буквой "омикрон" — в конце концов, именно греки придумали асимптотику.

Соотношение $f(n) = \Theta(g(n))$ справедливо тогда и только тогда, когда $f(n) \asymp g(n)$ в обозначениях, введенных в (9.8).

Эдмунд Ландау (Edmund Landau) [238] изобрел символ "о малое",

$$\begin{aligned} & f(n) = o(g(n)) \\ & \iff \quad |f(n)| \leqslant \epsilon|g(n)| \qquad \begin{array}{l} \text{для всех } n \geqslant n_0(\epsilon) \text{ и} \\ \text{для всех констант } \epsilon > 0. \end{array} \end{aligned} \qquad (9.19)$$

Это есть, по сути, отношение $f(n) \prec g(n)$ из (9.3). Кроме того,

$$f(n) \sim g(n) \iff f(n) = g(n) + o(g(n)) . \qquad (9.20)$$

Многие авторы используют в асимптотических формулах 'o', однако всегда предпочтительнее более явное выражение с 'O'. Например, среднее время работы алгоритма "пузырьковой сортировки" зависит от асимптотического значения суммы $P(n) = \sum_{k=0}^{n} k^{n-k} k!/n!$. Элементарных асимптотических методов достаточно, чтобы доказать формулу $P(n) \sim \sqrt{\pi n/2}$, которая означает, что $P(n)/\sqrt{\pi n/2}$ стремится к 1 при $n \to \infty$. Однако для того, чтобы лучше понять поведение $P(n)$, следует рассмотреть не отношение, а *разность* $P(n) - \sqrt{\pi n/2}$:

n	$P(n)/\sqrt{\pi n/2}$	$P(n) - \sqrt{\pi n/2}$
1	0.798	−0.253
10	0.878	−0.484
20	0.904	−0.538
30	0.918	−0.561
40	0.927	−0.575
50	0.934	−0.585

Числовые значения в среднем столбце не очень убедительны и едва ли могут считаться аргументом в пользу быстрой сходимости $P(n)/\sqrt{\pi n/2}$ к 1, если таковая сходимость вообще имеет место.

Но правый столбец показывает, что значение $P(n)$ действительно весьма близко к $\sqrt{\pi n/2}$. Таким образом, мы гораздо лучше охарактеризуем поведение $P(n)$, если сможем вывести формулы наподобие

$$P(n) \;=\; \sqrt{\pi n/2} + O(1)$$

или даже более точную оценку вида

$$P(n) \;=\; \sqrt{\pi n/2} - \tfrac{2}{3} + O(1/\sqrt{n}\,)\,.$$

Для доказательства результатов с O требуются более сильные методы асимптотического анализа, но затраченные на их изучение усилия сторицей окупаются лучшим пониманием асимптотики, которое дают O-оценки.

Более того, время работы многих алгоритмов сортировки имеет вид

$$T(n) \;=\; A\,n\lg n \;+\; B\,n \;+\; O(\log n)$$

для некоторых констант A и B. Если анализ останавливается на $T(n) \sim A\,n\lg n$, то мы получаем только ограниченную информацию; выбор алгоритма сортировки только на основе значения A — плохая стратегия. Хорошее 'A' в алгоритмах часто достигается ценой ухудшения 'B'. Поскольку $n\lg n$ растет только немного быстрее n, асимптотически более быстрый алгоритм (с несколько меньшим значением A) может оказаться быстрее только для таких значений n, которые никогда не будут достигнуты на практике. Таким образом, для правильного выбора необходим асимптотический метод, позволяющий пойти дальше первого члена и оценить B.

Перед тем как продолжить изучение обозначения 'O', поговорим об одном небольшом аспекте математического стиля. В этой главе для обозначения логарифма используются три разные записи: lg, ln и log. Мы часто используем lg в связи с вычислительными методами, в которых наиболее подходящим оказывается именно двоичный логарифм. В чистой математике чаще используется ln, поскольку формулы с натуральным логарифмом проще и изящнее. Но что такое log? Неужели это "обычный" логарифм по основанию 10, с которым вы сталкивались в старших классах — тот самый "обычный" логарифм, который так необычен для математики и информатики? Да, это он; вместе с тем многие математики вносят путаницу в этот вопрос, используя log для обозначения натурального или двоичного логарифма. Здесь нет единого соглашения, но можно вздохнуть с облегчением, если логарифм встречается внутри O-обозначения, потому что для O мультипликативные константы несущественны. При $n \to \infty$ между $O(\lg n)$, $O(\ln n)$ и $O(\log n)$ нет никакой разницы, как нет никакой разницы и между $O(\lg\lg n)$, $O(\ln\ln n)$ и $O(\log\log n)$. Мы можем выбрать тот вариант, который нам больше нравится; log кажется наиболее подходящим, потому что его легче произносить. Таким образом, мы, как правило, будем использовать log во всех контекстах, где удобочитаемость не привносит неоднозначности.

Заметьте, что $\log\log\log n$ неопределен при $n \leqslant 10$.

9.3 Работа с O

Как и для любого математического формализма, в случае O-обозначений имеются свои правила работы с ними, которые освобождают нас от необходимости прибегать всякий раз к громоздкому определению. Доказав единожды справедливость этих правил с применением определения, мы можем в дальнейшем подняться на более высокий уровень и забыть о проверке включения одного множества функций в другое. Нам даже не придется вычислять константы C, подразумеваемые каждым O, поскольку мы будем применять правила, гарантирующие существование таких констант.

Секрет, как стать занудой, заключается в том, чтобы говорить обо всем.
— Вольтер (Voltaire)

Например, можно раз и навсегда доказать, что

$$n^m = O(n^{m'}) \qquad \text{при } m \leqslant m'; \tag{9.21}$$
$$O\big(f(n)\big) + O\big(g(n)\big) = O\big(|f(n)| + |g(n)|\big). \tag{9.22}$$

Тогда мы сможем сразу сказать, что $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n = O(n^3) + O(n^3) + O(n^3) = O(n^3)$, без трудоемких вычислений из предыдущего раздела.

Вот еще несколько правил, которые точно так же легко можно получить из определений:

$$f(n) = O\big(f(n)\big); \tag{9.23}$$
$$c \cdot O\big(f(n)\big) = O\big(f(n)\big), \qquad \text{если } c \text{ — константа}; \tag{9.24}$$
$$O\big(O\big(f(n)\big)\big) = O\big(f(n)\big); \tag{9.25}$$
$$O\big(f(n)\big)O\big(g(n)\big) = O\big(f(n)g(n)\big); \tag{9.26}$$
$$O\big(f(n)\,g(n)\big) = f(n)O\big(g(n)\big). \tag{9.27}$$

В упр. 9 доказывается (9.22); доказательство остальных правил выполняется аналогично. Всегда можно заменять что-либо, имеющее вид одной из левых частей, тем, что записано справа, безотносительно к ограничениям, накладываемым на переменную n.

Замечание. Формула $O(f(n))^2$ не означает множество всех функций $g(n)^2$, где $g(n)$ принадлежит $O(f(n))$; такие функции $g(n)^2$ не могут принимать отрицательные значения, тогда как множество $O(f(n))^2$ включает отрицательные функции. В общем случае, если S — множество, то S^2 обозначает множество всех произведений $s_1 s_2$, где s_1 и s_2 принадлежат S, а не множество всех квадратов s^2, где $s \in S$.

Уравнения (9.27) и (9.23) позволяют доказать тождество $O\big(f(n)^2\big) = O\big(f(n)\big)^2$. Это свойство позволяет иногда сэкономить скобки, поскольку теперь можно записать

$$O(\log n)^2 \qquad \text{вместо} \qquad O\big((\log n)^2\big)\,.$$

Оба эти варианта предпочтительнее записи $O(\log^2 n)$, которая двусмысленна из-за того, что некоторые авторы используют $O(\log^2 n)$ для обозначения $O(\log\log n)$.

Можно ли также записать

$$O(\log n)^{-1} \qquad \text{вместо} \qquad O\big((\log n)^{-1}\big)\,?$$

Нет! Это некорректно, так как множество функций $1/O(\log n)$ не является ни подмножеством, ни надмножеством для $O(1/\log n)$. Правильно будет заменять $O\big((\log n)^{-1}\big)$ на $\Omega(\log n)^{-1}$, но это выглядит нелепо. Поэтому "возведение O в степень" будем использовать только в тех случаях, когда показатель степени представляет собой положительное целое число.

Степенные ряды дают нам одну из самых полезных операций. Если сумма

$$S(z) \;=\; \sum_{n\geqslant 0} a_n z^n$$

сходится абсолютно для некоторого комплексного числа $z = z_0$, то

$$S(z) \;=\; O(1) \qquad \text{для всех } |z| \leqslant |z_0|.$$

Это очевидно, поскольку

$$|S(z)| \;\leqslant\; \sum_{n\geqslant 0} |a_n|\,|z|^n \;\leqslant\; \sum_{n\geqslant 0} |a_n|\,|z_0|^n \;=\; C \;<\; \infty\,.$$

В частности, $S(z) = O(1)$ при $z \to 0$ и $S(1/n) = O(1)$ при $n \to \infty$, при том только условии, что $S(z)$ сходится хотя бы для одного ненулевого значения z. Этот принцип можно использовать для того, чтобы, отбрасывая хвост степенного ряда с любого удобного места, оценить этот хвост через O. Например, не только $S(z) = O(1)$, но и

$$\begin{aligned} S(z) \;&=\; a_0 + O(z)\,,\\ S(z) \;&=\; a_0 + a_1 z + O(z^2)\,, \end{aligned}$$

и т.д., потому что

$$S(z) \;=\; \sum_{0\leqslant k<m} a_k z^k + z^m \sum_{n\geqslant m} a_n z^{n-m},$$

а последняя сумма, как и сама $S(z)$, абсолютно сходится при $z = z_0$ и есть $O(1)$. В табл. 545 перечислены наиболее полезные асимптотические формулы, половина из которых получена просто путем отбрасывания членов степенного ряда в соответствии с указанным правилом.

Аналогично можно укорачивать ряды Дирихле, представляющие собой суммы вида $\sum_{k\geqslant 1} a_k/k^z$: если ряд Дирихле сходится абсолютно для некоторого $z = z_0$, то можно отбросить его хвост, начав с любого члена, и получить аппроксимацию

$$\sum_{1\leqslant k<m} a_k/k^z + O(m^{-z}),$$

справедливую для $\Re z \geqslant \Re z_0$. В табл. 545 этот принцип иллюстрируется асимптотической формулой для чисел Бернулли B_n.

Напомним, что $\Re$ означает "действительная часть".

С другой стороны, асимптотические формулы для H_n, $n!$ и $\pi(n)$ в табл. 545 не являются начальными отрезками сходящихся рядов; если неограниченно продолжить эти формулы, то полученные ряды будут расходиться при всех n. Особенно легко просматривается это в случае $\pi(n)$, поскольку, как мы уже видели в разделе 7.3, пример 5, степенной ряд $\sum_{k\geqslant 0} k!/(\ln n)^k$ всюду расходится. И тем не менее эти отрезки расходящихся рядов дают полезные аппроксимации.

Говорят, что асимптотическая аппроксимация имеет *абсолютную погрешность* $O\bigl(g(n)\bigr)$, если она имеет вид $f(n) + O\bigl(g(n)\bigr)$, где $f(n)$ не включает O. Приближение вида $f(n)\bigl(1 + O\bigl(g(n)\bigr)\bigr)$ имеет *относительную погрешность* $O\bigl(g(n)\bigr)$, если $f(n)$ не включает O. Например, аппроксимация H_n в табл. 545 имеет абсолютную погрешность $O(n^{-6})$; аппроксимация $n!$ имеет относительную погрешность $O(n^{-4})$. (Правая часть (9.29) не такая, как требуется, $f(n)\bigl(1+O(n^{-4})\bigr)$, но при желании ее можно переписать как

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+\frac{1}{12n}+\frac{1}{288n^2}-\frac{139}{51840n^3}\right)\bigl(1+O(n^{-4})\bigr);$$

аналогичное преобразование рассматривается в упр. 12.) Абсолютная погрешность этой аппроксимации составляет $O(n^{n-3.5}e^{-n})$. Абсолютная погрешность соотносится с числом верных десятичных цифр справа от десятичной точки, которые сохраняются после отбрасывания члена O; относительная погрешность связана с количеством верных "значащих цифр".

Относительная погрешность удобна при вычислении обратных величин, так как $1/(1 + O(\epsilon)) = 1 + O(\epsilon)$.

Воспользовавшись сокращением степенных рядов, можно доказать общие правила

$$\ln\bigl(1+O(f(n))\bigr) = O\bigl(f(n)\bigr), \qquad \text{если } f(n) \prec 1; \tag{9.36}$$

$$e^{O(f(n))} = 1+O\bigl(f(n)\bigr), \qquad \text{если } f(n) = O(1). \tag{9.37}$$

Таблица 545. Асимптотические аппроксимации, справедливые при $n \to \infty$ и $z \to 0$

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} + O\Big(\frac{1}{n^6}\Big) \qquad (9.28)$$

$$n! = \sqrt{2\pi n}\,\Big(\frac{n}{e}\Big)^n \Big(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} + O\Big(\frac{1}{n^4}\Big)\Big) \qquad (9.29)$$

$$B_n = 2[n \text{ четное}](-1)^{n/2-1}\frac{n!}{(2\pi)^n}\big(1 + 2^{-n} + 3^{-n} + O(4^{-n})\big) \qquad (9.30)$$

$$\pi(n) = \frac{n}{\ln n} + \frac{n}{(\ln n)^2} + \frac{2!\,n}{(\ln n)^3} + \frac{3!\,n}{(\ln n)^4} + O\Big(\frac{n}{(\log n)^5}\Big) \qquad (9.31)$$

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + O(z^5) \qquad (9.32)$$

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + O(z^5) \qquad (9.33)$$

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + z^4 + O(z^5) \qquad (9.34)$$

$$(1+z)^\alpha = 1 + \alpha z + \binom{\alpha}{2} z^2 + \binom{\alpha}{3} z^3 + \binom{\alpha}{4} z^4 + O(z^5) \qquad (9.35)$$

(Здесь мы полагаем, что $n \to \infty$; аналогичные формулы справедливы для $\ln\big(1 + O(f(x))\big)$ и $e^{O(f(x))}$ при $x \to 0$.) Например, пусть $\ln\big(1 + g(n)\big)$ — некоторая функция, принадлежащая множеству, описываемому левой частью (9.36). Тогда найдутся такие константы C, n_0 и c, что

$$\big|g(n)\big| \leqslant C\big|f(n)\big| \leqslant c < 1 \qquad \text{для всех } n \geqslant n_0.$$

Отсюда следует, что бесконечная сумма

$$\ln\big(1 + g(n)\big) = g(n) \cdot \big(1 - \tfrac{1}{2}g(n) + \tfrac{1}{3}g(n)^2 - \cdots\big)$$

сходится при всех $n \geqslant n_0$ и ряд в скобках ограничен константой $1 + \frac{1}{2}c + \frac{1}{3}c^2 + \cdots$. Это доказывает (9.36); доказательство (9.37) аналогично. Скомбинировав (9.36) и (9.37), получим полезную формулу

$$\big(1 + O(f(n))\big)^{O(g(n))} = 1 + O\big(f(n)g(n)\big), \qquad \begin{matrix}\text{если } f(n) \prec 1 \text{ и} \\ f(n)g(n) = O(1).\end{matrix} \qquad (9.38)$$

Задача 1: вновь о рулетке

Давайте теперь попытаем счастья в некоторых асимптотических задачах. В главе 3 мы вывели уравнение (3.13) для количества выигрышных позиций в некоторой игре:

$$W = \lfloor N/K \rfloor + \tfrac{1}{2}K^2 + \tfrac{5}{2}K - 3, \qquad K = \lfloor \sqrt[3]{N} \rfloor.$$

Мы обещали, что асимптотическое значение W будет получено в главе 9. Сейчас мы находимся именно в этой главе, так что самое время оценить W при $N \to \infty$.

Основная идея состоит в том, чтобы избавиться от скобок функции "пол", заменив K на $N^{1/3} + O(1)$. Далее можно записать

$$K = N^{1/3}\bigl(1 + O(N^{-1/3})\bigr);$$

эта операция называется "вынесение за скобки главной части". (Мы часто будем пользоваться этим методом.) Теперь из (9.38) и (9.26) мы имеем

$$\begin{aligned} K^2 &= N^{2/3}\bigl(1 + O(N^{-1/3})\bigr)^2 = \\ &= N^{2/3}\bigl(1 + O(N^{-1/3})\bigr) = N^{2/3} + O(N^{1/3}). \end{aligned}$$

Аналогично

$$\begin{aligned} \lfloor N/K \rfloor &= N^{1-1/3}\bigl(1 + O(N^{-1/3})\bigr)^{-1} + O(1) = \\ &= N^{2/3}\bigl(1 + O(N^{-1/3})\bigr) + O(1) = N^{2/3} + O(N^{1/3}). \end{aligned}$$

Отсюда следует, что количество выигрышных позиций равно

$$\begin{aligned} W &= N^{2/3} + O(N^{1/3}) + \tfrac{1}{2}\bigl(N^{2/3} + O(N^{1/3})\bigr) + O(N^{1/3}) + O(1) = \\ &= \tfrac{3}{2}N^{2/3} + O(N^{1/3}). \end{aligned} \qquad (9.39)$$

Торстен Силлке (Torsten Sillke) в 1996 году нашел, что $W = (3/2)N^{2/3} + (5/2)N^{1/3} - R$, где $3 \leqslant R < 121/24$ (причем последняя оценка — наилучшая возможная).

Обратите внимание, как члены с O поглощают друг друга, пока не остается только одно O. Это типичная ситуация и еще одна иллюстрация полезности применения O внутри формул.

Задача 2: метаморфозы формулы Стирлинга

Несомненно, самая знаменитая асимптотическая формула — приближение Стирлинга для $n!$. Мы докажем ее позже в этой главе; пока же попробуем лучше познакомиться с ее свойствами. Одна из версий приближения может быть записана в виде

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + \frac{a}{n} + \frac{b}{n^2} + O(n^{-3})\right) \quad \text{при } n \to \infty, \qquad (9.40)$$

для некоторых констант a и b. Поскольку это приближение справедливо при всех больших n, оно будет справедливо и при замене n на $n-1$:

$$\begin{aligned}(n-1)! \;&=\; \sqrt{2\pi(n-1)}\left(\frac{n-1}{e}\right)^{n-1}\times\\ &\times\left(1+\frac{a}{n-1}+\frac{b}{(n-1)^2}+O\big((n-1)^{-3}\big)\right). \qquad (9.41)\end{aligned}$$

Мы, конечно, знаем, что $(n-1)! = n!/n$; следовательно, правая часть последней формулы должна приводиться к правой части (9.40), деленной на n.

Давайте попробуем упростить (9.41). Из первого сомножителя можно вынести главную часть:

$$\begin{aligned}\sqrt{2\pi(n-1)} \;&=\; \sqrt{2\pi n}\,(1-n^{-1})^{1/2} \;=\\ &=\; \sqrt{2\pi n}\left(1-\frac{1}{2n}-\frac{1}{8n^2}+O(n^{-3})\right).\end{aligned}$$

Здесь использовано соотношение (9.35).

Аналогично имеем

$$\begin{aligned}\frac{a}{n-1} \;&=\; \frac{a}{n}(1-n^{-1})^{-1} \;=\; \frac{a}{n}+\frac{a}{n^2}+O(n^{-3});\\ \frac{b}{(n-1)^2} \;&=\; \frac{b}{n^2}(1-n^{-1})^{-2} \;=\; \frac{b}{n^2}+O(n^{-3});\\ O\big((n-1)^{-3}\big) \;&=\; O\big(n^{-3}(1-n^{-1})^{-3}\big) \;=\; O(n^{-3}).\end{aligned}$$

Единственная часть (9.41), преобразование которой требует некоторой изобретательности,— это множитель $(n-1)^{n-1}$, который равен

$$n^{n-1}(1-n^{-1})^{n-1} \;=\; n^{n-1}(1-n^{-1})^{n}\big(1+n^{-1}+n^{-2}+O(n^{-3})\big).$$

(Мы выписываем члены до тех пор, пока не получим относительную погрешность $O(n^{-3})$. Дело в том, что относительная погрешность произведения равна сумме относительных погрешностей сомножителей и все члены $O(n^{-3})$ объединяются.)

Чтобы раскрыть $(1-n^{-1})^n$, сначала вычислим $\ln(1-n^{-1})$ и затем образуем экспоненту, $e^{n\ln(1-n^{-1})}$:

$$\begin{aligned}(1-n^{-1})^n \;&=\; \exp\big(n\ln(1-n^{-1})\big) \;=\\ &=\; \exp\big(n(-n^{-1}-\tfrac12 n^{-2}-\tfrac13 n^{-3}+O(n^{-4}))\big) \;=\\ &=\; \exp\big(-1-\tfrac12 n^{-1}-\tfrac13 n^{-2}+O(n^{-3})\big) \;=\end{aligned}$$

$$
\begin{aligned}
&= \exp(-1)\cdot\exp(-\tfrac{1}{2}n^{-1})\cdot\exp(-\tfrac{1}{3}n^{-2})\cdot\exp\bigl(O(n^{-3})\bigr) = \\
&= \exp(-1)\cdot\bigl(1-\tfrac{1}{2}n^{-1}+\tfrac{1}{8}n^{-2}+O(n^{-3})\bigr)\times \\
&\qquad \times\bigl(1-\tfrac{1}{3}n^{-2}+O(n^{-4})\bigr)\cdot\bigl(1+O(n^{-3})\bigr) = \\
&= e^{-1}\bigl(1-\tfrac{1}{2}n^{-1}-\tfrac{5}{24}n^{-2}+O(n^{-3})\bigr)\,.
\end{aligned}
$$

Здесь вместо e^z мы использовали обозначение $\exp z$, поскольку это позволяет записывать сложный показатель степени в основной строке, а не в позиции верхнего индекса. Чтобы получить в конечном итоге относительную погрешность $O(n^{-3})$, нам приходится раскрывать $\ln(1-n^{-1})$ с абсолютной погрешностью $O(n^{-4})$, потому что этот логарифм умножается на n.

Итак, правая часть (9.41) приведена к виду "$\sqrt{2\pi n}$ умножить на n^{n-1}/e^n умножить на произведение нескольких членов":

$$
\begin{aligned}
&\bigl(1-\tfrac{1}{2}n^{-1}-\tfrac{1}{8}n^{-2}+O(n^{-3})\bigr)\times \\
&\quad\times\bigl(1+n^{-1}+n^{-2}+O(n^{-3})\bigr) \\
&\quad\times\bigl(1-\tfrac{1}{2}n^{-1}-\tfrac{5}{24}n^{-2}+O(n^{-3})\bigr) \\
&\quad\times\bigl(1+an^{-1}+(a+b)n^{-2}+O(n^{-3})\bigr)\,.
\end{aligned}
$$

Перемножение скобок и включение всех малых асимптотических членов в один $O(n^{-3})$ дает

$$1+an^{-1}+(a+b-\tfrac{1}{12})n^{-2}+O(n^{-3})\,.$$

Гм... Мы надеялись получить $1+an^{-1}+bn^{-2}+O(n^{-3})$, поскольку именно это требуется для соответствия правой части (9.40). Мы где-то ошиблись? Нет, все правильно; следовательно, $a+b-\frac{1}{12}=b$.

Хотя это рассуждение не доказывает справедливость аппроксимации Стирлинга, кое-что все-таки доказано, а именно — то, что формула (9.40) не может быть корректной, если a не равно $\frac{1}{12}$. Если заменить $O(n^{-3})$ в (9.40) на $cn^{-3}+O(n^{-4})$ и выполнить все вычисления с относительной погрешностью $O(n^{-4})$, то можно заключить, что b должно быть равно $\frac{1}{288}$, как и утверждается в табл. 545. (Это не самый простой путь найти a и b, но вполне работоспособный.)

Задача 3: n-е простое число

Уравнение (9.31) представляет собой асимптотическую формулу для $\pi(n)$, количества простых чисел, не превосходящих n. Заменяя n на $p = P_n$, n-е простое число, получим $\pi(p) = n$; следовательно,

$$n = \frac{p}{\ln p}+O\Bigl(\frac{p}{(\log p)^2}\Bigr) \qquad (9.42)$$

при $n \to \infty$. Попробуем "решить" это уравнение относительно p; тогда мы найдем приблизительную величину n-го простого числа.

Первый шаг состоит в упрощении члена O. Поделив обе части на $p/\ln p$, мы обнаружим, что $n \ln p/p \to 1$; следовательно, $p/\ln p = O(n)$ и

$$O\Bigl(\frac{p}{(\log p)^2}\Bigr) = O\Bigl(\frac{n}{\log p}\Bigr) = O\Bigl(\frac{n}{\log n}\Bigr).$$

(Имеем $(\log p)^{-1} \leqslant (\log n)^{-1}$, потому что $p \geqslant n$.)

Второй шаг состоит в перестановке двух частей (9.42), за исключением члена O. Эта процедура корректна ввиду общего правила

$$a_n = b_n + O\bigl(f(n)\bigr) \iff b_n = a_n + O\bigl(f(n)\bigr). \qquad (9.43)$$

(Любое из этих соотношений следует из другого, если умножить обе части на -1 и прибавить $a_n + b_n$.) Следовательно,

$$\frac{p}{\ln p} = n + O\Bigl(\frac{n}{\log n}\Bigr) = n\bigl(1 + O(1/\log n)\bigr),$$

а значит,

$$p = n \ln p\bigl(1 + O(1/\log n)\bigr). \qquad (9.44)$$

Это уравнение— "приближенное рекуррентное соотношение" для $p = P_n$, выражающее его через самого себя. Наша цель заключается в преобразовании его в "приближенный аналитический вид", и мы сможем это сделать, раскрыв рекуррентное соотношение асимптотически. Итак, попытаемся раскрыть (9.44).

Логарифмируя обе части, получаем

$$\ln p = \ln n + \ln\ln p + O(1/\log n). \qquad (9.45)$$

Это значение можно подставить вместо $\ln p$ в (9.44), но хотелось бы, прежде чем выполнять подстановку, полностью избавиться от p в правой части. Где-то в ходе преобразований последнее p должно исчезнуть; но от него нельзя избавиться обычным для рекуррентных соотношений способом, потому что в (9.44) не определены начальные условия для малых p.

Один из подходов к решению состоит в том, чтобы сначала доказать более слабый результат $p = O(n^2)$. В нем можно убедиться, если возвести в квадрат (9.44) и разделить на pn^2,

$$\frac{p}{n^2} = \frac{(\ln p)^2}{p}\bigl(1 + O(1/\log n)\bigr),$$

поскольку правая часть стремится к нулю при $n \to \infty$. Прекрасно, теперь мы знаем, что $p = O(n^2)$; следовательно, $\log p = O(\log n)$ и $\log\log p = O(\log\log n)$. Теперь из (9.45) можно получить

$$\ln p = \ln n + O(\log\log n);$$

эта оценка позволяет заключить, что $\ln\ln p = \ln\ln n + O(\log\log n/\log n)$, и (9.45) теперь дает

$$\ln p = \ln n + \ln\ln n + O(\log\log n/\log n).$$

Это выражение можно подставить в правую часть (9.44) и получить

$$p = n\ln n + n\ln\ln n + O(n).$$

Это и есть приближенная величина n-го простого числа.

Эту оценку можно улучшить, использовав вместо (9.42) более точное приближение $\pi(n)$. Следующий член (9.31) говорит нам, что

$$n = \frac{p}{\ln p} + \frac{p}{(\ln p)^2} + O\left(\frac{p}{(\log p)^3}\right); \tag{9.46}$$

действуя, как и ранее, получим рекуррентное соотношение

А теперь, как и ранее, выбросьте эту исписанную бумажку.

$$p = n\ln p\left(1 + (\ln p)^{-1}\right)^{-1}\left(1 + O(1/\log n)^2\right), \tag{9.47}$$

относительная погрешность которого — $O(1/\log n)^2$ вместо $O(1/\log n)$. Прологарифмировав и сохранив члены до нужной точности (но не чересчур много), получим

Шум, свист.

$$\begin{aligned}\ln p &= \ln n + \ln\ln p + O(1/\log n) = \\ &= \ln n\left(1 + \frac{\ln\ln p}{\ln n} + O(1/\log n)^2\right); \\ \ln\ln p &= \ln\ln n + \frac{\ln\ln n}{\ln n} + O\left(\frac{\log\log n}{\log n}\right)^2.\end{aligned}$$

Наконец, подставив полученные результаты в (9.47), находим ответ:

$$P_n = n\ln n + n\ln\ln n - n + n\frac{\ln\ln n}{\ln n} + O\left(\frac{n}{\log n}\right). \tag{9.48}$$

Например, при $n = 10^6$ эта оценка дает $15\,631\,363.6 + O(n/\log n)$; на самом деле миллионное простое число — $15\,485\,863$. В упр. 21

показано, что можно получить еще более точное приближение к P_n, если начать с более точного приближения $\pi(n)$, чем имеющееся в (9.46).

Задача 4: сумма из старого выпускного экзамена

Когда в Станфордском университете в 1970–1971 учебном году началось преподавание конкретной математики, студентам было предложено найти сумму

$$S_n = \frac{1}{n^2+1} + \frac{1}{n^2+2} + \cdots + \frac{1}{n^2+n} \tag{9.49}$$

с абсолютной погрешностью $O(n^{-7})$. Представим, что мы получили такое задание на выпускном экзамене. Какова наша первая инстинктивная реакция?

Мысль о том, что экзамен завален, будем рассматривать как нулевую инстинктивную реакцию.
— Переводчик

Нет, не паниковать. Первая реакция — думать по-большому. Если положить n равным, скажем, 10^{100} и посмотреть на сумму, то можно увидеть, что она состоит из n слагаемых, каждое из которых немного меньше $1/n^2$; следовательно, сумма будет немного меньше $1/n$. В общем случае обычно можно получить неплохую отправную точку для асимптотического анализа, если рассмотреть ситуацию в целом и получить оценку ответа "на глазок".

Давайте попытаемся улучшить эту грубую оценку, выписав главные члены каждого слагаемого. Имеем

$$\frac{1}{n^2+k} = \frac{1}{n^2(1+k/n^2)} = \frac{1}{n^2}\left(1-\frac{k}{n^2}+\frac{k^2}{n^4}-\frac{k^3}{n^6}+O\left(\frac{k^4}{n^8}\right)\right),$$

так что вполне естественно попробовать просуммировать все эти оценки:

$$\begin{aligned}
\frac{1}{n^2+1} &= \frac{1}{n^2} - \frac{1}{n^4} + \frac{1^2}{n^6} - \frac{1^3}{n^8} + O\left(\frac{1^4}{n^{10}}\right)\\
\frac{1}{n^2+2} &= \frac{1}{n^2} - \frac{2}{n^4} + \frac{2^2}{n^6} - \frac{2^3}{n^8} + O\left(\frac{2^4}{n^{10}}\right)\\
&\ \vdots\\
\frac{1}{n^2+n} &= \frac{1}{n^2} - \frac{n}{n^4} + \frac{n^2}{n^6} - \frac{n^3}{n^8} + O\left(\frac{n^4}{n^{10}}\right)\\
\hline
S_n &= \frac{n}{n^2} - \frac{n(n+1)}{2n^4} + \cdots .
\end{aligned}$$

Выглядит так, как будто мы приходим к результату $S_n = n^{-1} - \frac{1}{2}n^{-2} + O(n^{-3})$, полученному из суммы первых двух столбцов, только вот вычисления что-то очень разрастаются.

Придерживаясь этого подхода, в конце концов мы получим искомый ответ; но мы не будем так поступать по двум причинам. Во-первых, в последнем столбце появятся члены порядка $O(n^{-6})$, когда $n/2 \leqslant k \leqslant n$, так что погрешность вырастет до $O(n^{-5})$; этого слишком много, и нам надо добавить еще несколько столбцов (неужели экзаменатор может быть таким садистом?). Надо думать, есть способ решения и получше. Во-вторых, такой способ получше и в самом деле есть, и он прямо у вас перед глазами.

Может ли птица летать?

А именно, мы знаем выражение для S_n в аналитическом виде: это просто $H_{n^2+n} - H_{n^2}$. И мы знаем хорошее приближение для гармонических чисел, которое и применим дважды:

$$\begin{aligned} H_{n^2+n} &= \ln(n^2+n)+\gamma+\frac{1}{2(n^2+n)}-\frac{1}{12(n^2+n)^2}+O\Big(\frac{1}{n^8}\Big)\,; \\ H_{n^2} &= \ln n^2+\gamma+\frac{1}{2n^2}-\frac{1}{12n^4}+O\Big(\frac{1}{n^8}\Big)\,. \end{aligned}$$

Теперь можно вынести большие члены и выполнить упрощения, как мы делали в процессе анализа приближения Стирлинга. Имеем

$$\begin{aligned} \ln(n^2+n) &= \ln n^2+\ln\Big(1+\frac{1}{n}\Big) = \ln n^2+\frac{1}{n}-\frac{1}{2n^2}+\frac{1}{3n^3}-\cdots\,; \\ \frac{1}{n^2+n} &= \frac{1}{n^2}-\frac{1}{n^3}+\frac{1}{n^4}-\cdots\,; \\ \frac{1}{(n^2+n)^2} &= \frac{1}{n^4}-\frac{2}{n^5}+\frac{3}{n^6}-\cdots\,. \end{aligned}$$

После ряда успешных сокращений получим

$$\begin{aligned} S_n = n^{-1} - \tfrac{1}{2}n^{-2} + \tfrac{1}{3}n^{-3} - \tfrac{1}{4}n^{-4} + \tfrac{1}{5}n^{-5} - \tfrac{1}{6}n^{-6} - \\ - \tfrac{1}{2}n^{-3} + \tfrac{1}{2}n^{-4} - \tfrac{1}{2}n^{-5} + \tfrac{1}{2}n^{-6} + \\ + \tfrac{1}{6}n^{-5} - \tfrac{1}{4}n^{-6} \end{aligned}$$

плюс члены вида $O(n^{-7})$. Немного арифметики — и мы свободны:

$$S_n = n^{-1}-\tfrac{1}{2}n^{-2}-\tfrac{1}{6}n^{-3}+\tfrac{1}{4}n^{-4}-\tfrac{2}{15}n^{-5}+\tfrac{1}{12}n^{-6}+O(n^{-7})\,. \qquad (9.50)$$

Было бы здорово, если бы можно было проверить этот ответ численно, как в предыдущих главах, когда мы получали точные результаты. Асимптотические формулы проверять труднее — в O может скрываться произвольно большая константа, так что числовой тест не дает окончательный ответ. Но на практике у нас нет причин думать, что нам противостоит хитрый и изощренный противник, так что априори можно надеяться, что неизвестная константа в O достаточно мала. Калькулятора достаточно, чтобы быстро вычислить $S_4 = \frac{1}{17} + \frac{1}{18} + \frac{1}{19} + \frac{1}{20} = 0.2170107$; а наша асимптотическая оценка для $n = 4$ дает

$$\tfrac{1}{4}\bigl(1 + \tfrac{1}{4}\bigl(-\tfrac{1}{2} + \tfrac{1}{4}(-\tfrac{1}{6} + \tfrac{1}{4}(\tfrac{1}{4} + \tfrac{1}{4}(-\tfrac{2}{15} + \tfrac{1}{4}\cdot\tfrac{1}{12})))\bigr)\bigr) \;=\; 0.2170125\,.$$

Если мы допустили ошибку, скажем, величиной $\frac{1}{12}$ в члене с n^{-6}, то разность $\frac{1}{12}\frac{1}{4096}$ должна проявиться в пятой десятичной цифре; так что, скорее всего, полученная нами асимптотическая формула верна.

Задача 5: бесконечная сумма

Теперь рассмотрим еще одну асимптотическую задачу, сформулированную Соломоном Голомбом (Solomon Golomb) [152]: чему равно приближенное значение

$$S_n \;=\; \sum_{k\geqslant 1} \frac{1}{kN_n(k)^2}\,, \qquad (9.51)$$

где $N_n(k)$ — количество цифр, требуемых для записи числа k в системе счисления с основанием n?

Начнем с попытки найти грубую оценку. Количество цифр $N_n(k)$ приближенно равно $\log_n k = \log k/\log n$; так что слагаемые грубо равны $(\log n)^2/k(\log k)^2$. Суммирование по k дает $\approx (\log n)^2 \sum_{k\geqslant 2} 1/k(\log k)^2$, а эта сумма сходится к константе, так как ее можно сравнить с интегралом

$$\int_2^\infty \frac{dx}{x(\ln x)^2} \;=\; -\frac{1}{\ln x}\bigg|_2^\infty \;=\; \frac{1}{\ln 2}\,.$$

Таким образом, мы ожидаем, что величина S_n примерно равна $C(\log n)^2$ для некоторой константы C.

Такой беглый анализ полезен для общей ориентации, но не для решения задачи — здесь нужна лучшая оценка. Одна из возможных идей заключается в точном выражении $N_n(k)$:

$$N_n(k) \;=\; \lfloor \log_n k \rfloor + 1\,. \qquad (9.52)$$

Так, например, для записи k в системе счисления по основанию n требуются три цифры, если $n^2 \leqslant k < n^3$, а это выполняется в точности тогда, когда $\lfloor \log_n k \rfloor = 2$. Таким образом, $N_n(k) > \log_n k$, следовательно, $S_n = \sum_{k\geqslant 1} 1/kN_n(k)^2 < 1 + (\log n)^2 \sum_{k\geqslant 2} 1/k(\log k)^2$.

Действуя так же, как в задаче 1, можно записать $N_n(k) = \log_n k + O(1)$ и подставить это выражение в формулу для S_n. Слагаемое, записанное здесь как $O(1)$, всегда находится между

0 и 1 и в среднем близко к $\frac{1}{2}$, так что, вероятно, оно ведет себя достаточно хорошо. Тем не менее это приближение недостаточно точное, чтобы сказать что-нибудь о S_n; оно дает нуль значащих цифр (т.е. большую относительную погрешность) при малых k, а именно эти слагаемые вносят основной вклад в сумму. Нужна другая идея.

Ключ к решению (как и в задаче 4) — применить наше искусство преобразования выражений и, прежде чем переходить к асимптотическим оценкам, привести сумму к более удобному виду. Можно ввести новую переменную суммирования $m = N_n(k)$:

$$\begin{aligned}
S_n &= \sum_{k,m\geqslant 1} \frac{[m=N_n(k)]}{km^2} = \\
&= \sum_{k,m\geqslant 1} \frac{[n^{m-1}\leqslant k<n^m]}{km^2} = \\
&= \sum_{m\geqslant 1} \frac{1}{m^2}\left(H_{n^m-1} - H_{n^{m-1}-1}\right).
\end{aligned}$$

Это выражение может показаться еще худшим, чем исходная сумма, но в действительности сделан шаг в правильном направлении, поскольку мы располагаем очень хорошим приближением для гармонических чисел.

Но пока что придержим его при себе и попытаемся упростить еще кое-что. Не надо спешить к асимптотическим оценкам. Суммирование по частям позволяет сгруппировать члены с одинаковыми значениями H_{n^m-1}, которые нам придется аппроксимировать:

$$S_n = \sum_{k\geqslant 1} H_{n^k-1}\left(\frac{1}{k^2} - \frac{1}{(k+1)^2}\right).$$

Например, H_{n^2-1} сначала умножается на $1/2^2$, а затем — на $-1/3^2$. (Мы использовали тот факт, что $H_{n^0-1} = H_0 = 0$.)

Теперь мы готовы подставить выражение для гармонических чисел. Наш опыт оценки $(n-1)!$ подсказывает, что проще будет аппроксимировать H_{n^k}, а не H_{n^k-1}, поскольку в случае (n^k-1) появится масса лишних членов. Так что мы пишем

$$\begin{aligned}
H_{n^k-1} &= H_{n^k} - \frac{1}{n^k} = \ln n^k + \gamma + \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right) - \frac{1}{n^k} = \\
&= k\ln n + \gamma - \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right).
\end{aligned}$$

Теперь наша сумма сводится к

$$S_n = \sum_{k\geqslant 1}\Bigl(k\ln n+\gamma-\frac{1}{2n^k}+O\Bigl(\frac{1}{n^{2k}}\Bigr)\Bigr)\Bigl(\frac{1}{k^2}-\frac{1}{(k+1)^2}\Bigr) =$$
$$= (\ln n)\Sigma_1+\gamma\Sigma_2-\tfrac{1}{2}\Sigma_3(n)+O\bigl(\Sigma_3(n^2)\bigr)\,. \qquad (9.53)$$

Осталось расписать четыре суммы: Σ_1, Σ_2, $\Sigma_3(n)$ и $\Sigma_3(n^2)$.

Большим-большим O.

Начнем с Σ_3, поскольку $\Sigma_3(n^2)$ входит в O; таким образом, сразу будет видно, какую погрешность мы получим. (Нет смысла выполнять остальные вычисления с высокой точностью, если эти члены все равно будут поглощены O.) Эта сумма — просто степенной ряд

$$\Sigma_3(x) = \sum_{k\geqslant 1}\Bigl(\frac{1}{k^2}-\frac{1}{(k+1)^2}\Bigr)x^{-k}\,,$$

и этот ряд сходится при $x \geqslant 1$, так что мы можем оборвать его в любом желаемом месте. Если оборвать $\Sigma_3(n^2)$ на члене $k=1$, то получим $\Sigma_3(n^2)=O(n^{-2})$; следовательно, (9.53) имеет абсолютную погрешность $O(n^{-2})$. (Для уменьшения этой абсолютной погрешности можно использовать более точное приближение для H_{n^k}; но пока что точности $O(n^{-2})$ нам вполне достаточно.) Если оборвать $\Sigma_3(n)$ на члене $k=2$, можно получить

$$\Sigma_3(n) = \tfrac{3}{4}n^{-1}+O(n^{-2})\,;$$

это как раз та точность, которая нам нужна.

Совсем просто вычислить Σ_2:

$$\Sigma_2 = \sum_{k\geqslant 1}\Bigl(\frac{1}{k^2}-\frac{1}{(k+1)^2}\Bigr)\,.$$

Это телескопический ряд $(1-\frac14)+(\frac14-\frac19)+(\frac19-\frac1{16})+\cdots=1$.

Наконец, Σ_1 дает нам главный член суммы S_n, коэффициент при $\ln n$ в (9.53):

$$\Sigma_1 = \sum_{k\geqslant 1}k\Bigl(\frac{1}{k^2}-\frac{1}{(k+1)^2}\Bigr)\,.$$

Это $(1-\frac14)+(\frac24-\frac29)+(\frac39-\frac3{16})+\cdots=\frac11+\frac14+\frac19+\cdots=H_\infty^{(2)}=\pi^2/6$. (Не примени мы ранее суммирование по частям, мы бы непосредственно увидели, что $S_n \sim \sum_{k\geqslant 1}(\ln n)/k^2$, потому что $H_{n^k-1}-H_{n^{k-1}-1}\sim \ln n$; так что суммирование по частям не помогает в оценке главного члена, хотя и упрощает кое-что другое.)

Итак, мы оценили каждую из сумм в (9.53), так что теперь можно объединить все части и получить ответ к задаче Голомба:

$$S_n = \frac{\pi^2}{6}\ln n+\gamma-\frac{3}{8n}+O\Bigl(\frac{1}{n^2}\Bigr)\,. \qquad (9.54)$$

Обратите внимание, что это выражение растет медленнее, чем наша исходная прикидочная оценка $C(\log n)^2$. Иногда дискретные суммы не отвечают непрерывной интуиции.

Задача 6: Φ ***большое***

В конце главы 4 мы нашли, что количество дробей в последовательности Фарея $\mathcal{F}_n$ равно $1+\Phi(n)$, где

$$\Phi(n) \;=\; \varphi(1)+\varphi(2)+\cdots+\varphi(n)\,;$$

и в (4.62) мы показали, что

$$\Phi(n) \;=\; \frac{1}{2}\sum_{k\geqslant 1}\mu(k)\lfloor n/k\rfloor\lfloor 1+n/k\rfloor\,. \qquad (9.55)$$

Давайте попробуем оценить $\Phi(n)$ при больших n. (В первую очередь, именно суммы такого вида и привели Бахмана (Bachmann) к введению символа 'O'.)

Размышления по-большому говорят нам о том, что $\Phi(n)$, вероятно, пропорционально n^2. Действительно, если бы последний множитель был равен $\lfloor n/k\rfloor$, а не $\lfloor 1+n/k\rfloor$, то мы имели бы верхнюю границу $\bigl|\Phi(n)\bigr|\leqslant\frac{1}{2}\sum_{k\geqslant1}\lfloor n/k\rfloor^2\leqslant\frac{1}{2}\sum_{k\geqslant1}(n/k)^2=\frac{\pi^2}{12}n^2$, потому что функция Мёбиуса $\mu(k)$ принимает значение только -1, 0 или $+1$. Слагаемое '$1+$' в последнем множителе добавляет $\sum_{k\geqslant1}\mu(k)\lfloor n/k\rfloor$; но для $k>n$ это нуль, так что сумма по абсолютной величине не может быть больше, чем $nH_n=O(n\log n)$.

Этот предварительный анализ показывает, что, вероятно, будет удобно записать

$$\begin{aligned}\Phi(n) = \frac{1}{2}\sum_{k=1}^{n}\mu(k)\Bigl(\Bigl(\frac{n}{k}\Bigr)+O(1)\Bigr)^2 &= \frac{1}{2}\sum_{k=1}^{n}\mu(k)\Bigl(\Bigl(\frac{n}{k}\Bigr)^2+O\Bigl(\frac{n}{k}\Bigr)\Bigr) =\\ &= \frac{1}{2}\sum_{k=1}^{n}\mu(k)\Bigl(\frac{n}{k}\Bigr)^2+\sum_{k=1}^{n}O\Bigl(\frac{n}{k}\Bigr) =\\ &= \frac{1}{2}\sum_{k=1}^{n}\mu(k)\Bigl(\frac{n}{k}\Bigr)^2+O(n\log n)\,.\end{aligned}$$

Таким образом мы убрали функцию "пол"; теперь нам остается оценить сумму $\frac{1}{2}\sum_{k=1}^{n}\mu(k)n^2/k^2$ с точностью $O(n\log n)$; другими словами, мы хотим вычислить $\sum_{k=1}^{n}\mu(k)1/k^2$ с точностью $O(n^{-1}\log n)$. Но это уже легко; можно просто распространить сумму на все значения k до ∞, потому что вновь добавляемые члены составляют

$$\sum_{k>n} \frac{\mu(k)}{k^2} = O\Big(\sum_{k>n} \frac{1}{k^2}\Big) = O\Big(\sum_{k>n} \frac{1}{k(k-1)}\Big) =$$
$$= O\bigg(\sum_{k>n}\Big(\frac{1}{k-1} - \frac{1}{k}\Big)\bigg) = O\Big(\frac{1}{n}\Big).$$

Мы доказали в (7.89), что $\sum_{k\geqslant 1} \mu(k)/k^z = 1/\zeta(z)$. Итак, $\sum_{k\geqslant 1} \mu(k)/k^2 = 1/\big(\sum_{k\geqslant 1} 1/k^2\big) = 6/\pi^2$, и мы получаем окончательный ответ:

$$\Phi(n) = \frac{3}{\pi^2}n^2 + O(n\log n). \tag{9.56}$$

Как было показано Салтыковым в 1960 году [316], погрешность этой формулы не превосходит $O(n(\log n)^{2/3} \times (\log\log n)^{1+\epsilon})$. *С другой стороны, она не столь мала, как* $o(n(\log\log n)^{1/2})$, *согласно Монтгомери [275].*

9.4 Два асимптотических приема

Теперь, приобретя некоторую легкость в обращении с O, попробуем взглянуть на сделанное, имея в виду более далекую перспективу. Тогда наш арсенал асимптотических методов пополнится новыми важными боевыми средствами, которые понадобятся нам для борьбы с более трудными задачами.

Прием 1: раскрутка

При оценке n-го простого числа P_n в задаче 3 раздела 9.3 мы решали асимптотическое рекуррентное соотношение вида

$$P_n = n\ln P_n\big(1 + O(1/\log n)\big).$$

Мы доказали, что $P_n = n\ln n + n\ln\ln n + O(n)$, сначала использовав рекуррентное соотношение для получения более слабого результата $O(n^2)$. Это частный случай общего метода, именуемого *раскруткой* (bootstrapping), при котором для нахождения асимптотического решения рекуррентного соотношения мы начинаем с грубой оценки и подставляем ее в рекуррентное соотношение; таким способом часто удается получать все лучшие и лучшие оценки, как бы, подобно барону Мюнхгаузену, "поднимая себя за волосы".

Следующая задача очень хорошо иллюстрирует метод раскрутки: каково асимптотическое значение коэффициента $g_n = [z^n]\,G(z)$ в производящей функции

$$G(z) = \exp\Big(\sum_{k\geqslant 1} \frac{z^k}{k^2}\Big) \tag{9.57}$$

при $n \to \infty$? Продифференцировав это выражение по z, мы найдем

$$G'(z) = \sum_{n=0}^{\infty} n g_n z^{n-1} = \Bigl(\sum_{k\geqslant 1} \frac{z^{k-1}}{k}\Bigr) G(z)\,;$$

приравнивание коэффициентов при z^{n-1} в обеих частях дает рекуррентное соотношение

$$n g_n = \sum_{0\leqslant k<n} \frac{g_k}{n-k}\,. \tag{9.58}$$

Наша задача эквивалентна поиску асимптотической формулы для решения (9.58) с начальным условием $g_0 = 1$. Первые несколько значений

n	0	1	2	3	4	5	6
g_n	1	1	$\frac{3}{4}$	$\frac{19}{36}$	$\frac{107}{288}$	$\frac{641}{2400}$	$\frac{51103}{259200}$

мало что дают для выявления закономерности, а целочисленная последовательность $\langle n!^2 g_n\rangle$ в справочнике Слоана [330] отсутствует; так что об аналитическом виде g_n, похоже, не стоит и мечтать, и лучшее, на что можно надеяться,— вывод асимптотики.

Нашим первым шагом на пути решения задачи будет наблюдение, что $0 < g_n \leqslant 1$ для всех $n \geqslant 0$; это легко доказать по индукции. Итак, для начала имеем

$$g_n = O(1)\,.$$

Это уравнение можно использовать в качестве отправной точки для раскрутки: подставив его в правую часть (9.58), получим

$$n g_n = \sum_{0\leqslant k<n} \frac{O(1)}{n-k} = H_n O(1) = O(\log n)\,;$$

следовательно,

$$g_n = O\Bigl(\frac{\log n}{n}\Bigr) \qquad \text{для } n > 1.$$

Можно сделать еще один шаг раскрутки:

$$\begin{aligned} n g_n &= \frac{1}{n} + \sum_{0<k<n} \frac{O\bigl((1+\log k)/k\bigr)}{n-k} = \\ &= \frac{1}{n} + \sum_{0<k<n} \frac{O(\log n)}{k(n-k)} = \end{aligned}$$

$$= \frac{1}{n} + \sum_{0<k<n} \left(\frac{1}{k} + \frac{1}{n-k}\right) \frac{O(\log n)}{n} =$$

$$= \frac{1}{n} + \frac{2}{n} H_{n-1} O(\log n) = \frac{1}{n} O(\log n)^2 ,$$

что дает

$$g_n = O\left(\frac{\log n}{n}\right)^2 . \qquad (9.59)$$

Продолжится ли это до бесконечности? Может быть, мы получим $g_n = O(n^{-1} \log n)^m$ для всех m?

На самом деле не получим; мы уже достигли поворотного пункта. Попытка продолжить раскрутку приведет к сумме

$$\sum_{0<k<n} \frac{1}{k^2(n-k)} = \sum_{0<k<n} \left(\frac{1}{nk^2} + \frac{1}{n^2 k} + \frac{1}{n^2(n-k)}\right) =$$

$$= \frac{1}{n} H^{(2)}_{n-1} + \frac{2}{n^2} H_{n-1} ,$$

которая представляет собой $\Omega(n^{-1})$; таким образом, мы не сможем получить для g_n оценку, меньшую $\Omega(n^{-2})$.

В действительности сейчас мы уже знаем о g_n достаточно, чтобы применить наш старый прием выделения наибольшей части:

$$n g_n = \sum_{0\leqslant k<n} \frac{g_k}{n} + \sum_{0\leqslant k<n} g_k \left(\frac{1}{n-k} - \frac{1}{n}\right) =$$

$$= \frac{1}{n} \sum_{k\geqslant 0} g_k - \frac{1}{n} \sum_{k\geqslant n} g_k + \frac{1}{n} \sum_{0\leqslant k<n} \frac{k g_k}{n-k} . \qquad (9.60)$$

Первая сумма здесь — $G(1) = \exp(\frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \cdots) = e^{\pi^2/6}$, потому что $G(z)$ сходится для всех $|z| \leqslant 1$. Вторая сумма представляет собой хвост первой; можно получить верхнюю оценку, воспользовавшись (9.59):

$$\sum_{k\geqslant n} g_k = O\left(\sum_{k\geqslant n} \frac{(\log k)^2}{k^2}\right) = O\left(\frac{(\log n)^2}{n}\right) .$$

Эта последняя оценка следует, например, из

$$\sum_{k>n} \frac{(\log k)^2}{k^2} < \sum_{m\geqslant 1} \sum_{n^m<k\leqslant n^{m+1}} \frac{(\log n^{m+1})^2}{k(k-1)} < \sum_{m\geqslant 1} \frac{(m+1)^2 (\log n)^2}{n^m} .$$

(В упр. 54 рассматривается более общий способ оценки остатков подобных рядов.)

Третья сумма в (9.60) равна

$$O\Bigl(\sum_{0<k<n}\frac{(\log n)^2}{k(n-k)}\Bigr) = O\Bigl(\frac{(\log n)^3}{n}\Bigr),$$

что устанавливается при помощи уже знакомых нам аргументов. Таким образом, (9.60) доказывает, что

$$g_n = \frac{e^{\pi^2/6}}{n^2} + O\bigl(\log n/n\bigr)^3. \qquad (9.61)$$

Наконец, мы можем вновь подставить эту формулу в рекуррентное соотношение, выполнив еще один шаг раскрутки; в результате получим

$$g_n = \frac{e^{\pi^2/6}}{n^2} + O(\log n/n^3). \qquad (9.62)$$

(Упр. 23 "заглядывает внутрь" оставшегося O.)

Прием 2: кручение хвостов

Вывод (9.62) в чем-то похож на вывод асимптотического значения (9.56) для $\Phi(n)$: в обоих случаях мы начинали с конечной суммы, но к асимптотическому выражению подходили через рассмотрение бесконечной суммы. При этом нельзя было просто получить бесконечную сумму, добавив O к слагаемым; вместо этого приходилось действовать аккуратно и использовать один подход для малых k и другой — для k бо́льших.

Эти выводы были частными случаями важного трехшагового асимптотического метода суммирования, который мы сейчас обсудим с большей общностью. Чтобы оценить значение $\sum_k a_k(n)$, можно применить следующий прием.

Этот важный метод был впервые применен Лапласом (Laplace) [240].

1 Сначала разбить весь диапазон суммирования на два непересекающихся диапазона, D_n и T_n. Сумма по D_n должна быть "доминантной" частью в том смысле, что она включает достаточно членов, чтобы верно представлять наиболее значащие цифры всей суммы при больших n. Суммирование по диапазону T_n должно представлять собой только "хвост", вносящий в общий итог незначительный вклад.

2 Найти асимптотическую оценку

$$a_k(n) = b_k(n) + O\bigl(c_k(n)\bigr),$$

справедливую при $k \in D_n$. Не требуется, чтобы эта оценка имела место при $k \in T_n$.

3 Наконец, доказать, что все три приведенные ниже суммы малы:

$$\Sigma_a(n) = \sum_{k\in T_n} a_k(n); \qquad \Sigma_b(n) = \sum_{k\in T_n} b_k(n);$$
$$\Sigma_c(n) = \sum_{k\in D_n} \bigl|c_k(n)\bigr|. \tag{9.63}$$

Если все три шага успешно завершаются, то в итоге получается хорошая оценка:

$$\sum_{k\in D_n\cup T_n} a_k(n) = \sum_{k\in D_n\cup T_n} b_k(n) + O\bigl(\Sigma_a(n)\bigr) + O\bigl(\Sigma_b(n)\bigr) + O\bigl(\Sigma_c(n)\bigr).$$

И вот почему. Можно "обрубить" хвост у данной суммы, получив хорошую оценку в диапазоне D_n, там, где такая оценка действительно необходима:

$$\sum_{k\in D_n} a_k(n) = \sum_{k\in D_n} \bigl(b_k(n) + O(c_k(n))\bigr) = \sum_{k\in D_n} b_k(n) + O\bigl(\Sigma_c(n)\bigr).$$

Хвост же можно заменить другим, даже весьма плохо аппроксимирующим первый, поскольку ни один из них не играет заметной роли:

Асимптотический анализ — искусство знать, когда можно быть небрежным, а когда — предельно точным.

$$\sum_{k\in T_n} a_k(n) = \sum_{k\in T_n} \bigl(b_k(n) - b_k(n) + a_k(n)\bigr) =$$
$$= \sum_{k\in T_n} b_k(n) + O\bigl(\Sigma_b(n)\bigr) + O\bigl(\Sigma_a(n)\bigr).$$

Например, вычисляя сумму в (9.60), мы имели

$$\begin{aligned} a_k(n) &= [0\leqslant k<n]g_k/(n-k), \\ b_k(n) &= g_k/n, \\ c_k(n) &= kg_k/n(n-k); \end{aligned}$$

диапазонами суммирования были

$$D_n = \{0, 1, \ldots, n-1\}, \qquad T_n = \{n, n+1, \ldots\};$$

и мы нашли, что

$$\Sigma_a(n) = 0, \quad \Sigma_b(n) = O\bigl((\log n)^2/n^2\bigr), \quad \Sigma_c(n) = O\bigl((\log n)^3/n^2\bigr).$$

Это привело к (9.61).

Аналогично, оценивая $\Phi(n)$ в (9.55), мы имели

$$\begin{aligned} a_k(n) &= \mu(k)\lfloor n/k\rfloor\lfloor 1+n/k\rfloor\,, \quad b_k(n) = \mu(k)n^2/k^2\,, \quad c_k(n) = n/k\,;\\ D_n &= \{1,2,\ldots,n\}, \qquad T_n = \{n+1,n+2,\ldots\}. \end{aligned}$$

Мы вывели (9.56), заметив, что $\Sigma_a(n) = 0$, $\Sigma_b(n) = O(n)$ и $\Sigma_c(n) = O(n\log n)$.

Вот другой пример, где такое "кручение хвостов" оказывается эффективным. (В отличие от предыдущих примеров, этот пример иллюстрирует рассматриваемый прием в его максимальной общности, когда $\Sigma_a(n) \neq 0$.) Мы ищем асимптотическое значение

$$L_n = \sum_{k\geqslant 0} \frac{\ln(n+2^k)}{k!}\,.$$

Наибольший вклад в эту сумму вносят малые k, потому что в знаменателе имеется $k!$. В диапазоне малых k

$$\ln(n+2^k) = \ln n + \frac{2^k}{n} - \frac{2^{2k}}{2n^2} + O\Big(\frac{2^{3k}}{n^3}\Big)\,. \tag{9.64}$$

Можно доказать эту оценку для $0 \leqslant k < \lfloor \lg n\rfloor$, поскольку члены, включенные в O, ограничены сходящимся рядом

$$\sum_{m\geqslant 3}\frac{2^{km}}{mn^m} \leqslant \frac{2^{3k}}{n^3}\sum_{m\geqslant 3}\frac{2^{k(m-3)}}{n^{m-3}} \leqslant \frac{2^{3k}}{n^3}\Big(1+\frac{1}{2}+\frac{1}{4}+\cdots\Big) = \frac{2^{3k}}{n^3}\cdot 2\,.$$

(В этом диапазоне $2^k/n \leqslant 2^{\lfloor\lg n\rfloor-1}/n \leqslant \frac{1}{2}$.)

Следовательно, можно применить описанный трехшаговый метод, положив

$$\begin{aligned} a_k(n) &= \ln(n+2^k)/k!\,,\\ b_k(n) &= (\ln n + 2^k/n - 4^k/2n^2)/k!\,,\\ c_k(n) &= 8^k/n^3k!\,;\\ D_n &= \{0,1,\ldots,\lfloor\lg n\rfloor-1\},\\ T_n &= \{\lfloor\lg n\rfloor, \lfloor\lg n\rfloor+1,\ldots\}. \end{aligned}$$

Все, что остается сделать, — это найти хорошие оценки для трех величин Σ в (9.63), и мы узнаем, что $\sum_{k\geqslant 0} a_k(n) \approx \sum_{k\geqslant 0} b_k(n)$.

Погрешность в доминантной части суммы,

$$\Sigma_c(n) = \sum_{k\in D_n} 8^k/n^3k!,$$

очевидным образом ограничена суммой $\sum_{k\geqslant 0} 8^k/n^3k! = e^8/n^3$, так что ее можно записать как $O(n^{-3})$. Погрешность, вносимая новым хвостом, составляет

$$
\begin{aligned}
\left|\Sigma_b(n)\right| &= \Bigl|\sum_{k\geqslant\lfloor\lg n\rfloor} b_k(n)\Bigr| < \\
&< \sum_{k\geqslant\lfloor\lg n\rfloor} \frac{\ln n + 2^k + 4^k}{k!} < \\
&< \frac{\ln n + 2^{\lfloor\lg n\rfloor} + 4^{\lfloor\lg n\rfloor}}{\lfloor\lg n\rfloor!} \sum_{k\geqslant 0} \frac{4^k}{k!} = O\Bigl(\frac{n^2}{\lfloor\lg n\rfloor!}\Bigr).
\end{aligned}
$$

Поскольку $\lfloor\lg n\rfloor!$ растет быстрее, чем любая степень n, эта маленькая погрешность поглощается слагаемым $\Sigma_c(n) = O(n^{-3})$. Погрешность, вносимая хвостом исходного ряда

$$
\Sigma_a(n) = \sum_{k\geqslant\lfloor\lg n\rfloor} a_k(n) < \sum_{k\geqslant\lfloor\lg n\rfloor} \frac{k+\ln n}{k!},
$$

еще меньше.

Наконец, можно легко вычислить сумму $\sum_{k\geqslant 0} b_k(n)$ в аналитическом виде, и мы получаем желаемую асимптотическую формулу:

$$
\sum_{k\geqslant 0} \frac{\ln(n+2^k)}{k!} = e\ln n + \frac{e^2}{n} - \frac{e^4}{2n^2} + O\Bigl(\frac{1}{n^3}\Bigr). \tag{9.65}
$$

Из использованного метода совершенно ясно, что на самом деле

$$
\sum_{k\geqslant 0} \frac{\ln(n+2^k)}{k!} = e\ln n + \sum_{k=1}^{m-1} (-1)^{k+1} \frac{e^{2^k}}{kn^k} + O\Bigl(\frac{1}{n^m}\Bigr) \tag{9.66}
$$

для любого фиксированного $m > 0$. (Это начальный отрезок ряда, расходящегося для всех фиксированных n при $m \to \infty$.)

В нашем решении имеется только один недостаток: мы были слишком осторожны. Мы вывели (9.64) в предположении $k < \lfloor\lg n\rfloor$, но упр. 53 доказывает, что эта оценка в действительности справедлива для всех значений k. Если бы мы знали этот более сильный результат, нам бы не пришлось использовать трюк с двумя хвостами — мы могли бы сразу прийти к конечному результату! Однако позже мы еще встретимся с задачей, в которой замена хвостов оказывается единственным доступным методом.

9.5 Формула суммирования Эйлера

А теперь в качестве очередного трюка (который фактически является последним важным приемом, рассматривающимся в данной книге) мы обратимся к общему методу аппроксимации сумм, опубликованному Леонардом Эйлером [101] в 1732 году. (Идею метода иногда связывают с именем Колина Маклорена (Colin Maclaurin), профессора математики в Эдинбурге, независимо открывшего этот метод несколько позднее [263, с. 305].)

Вот основная формула:

$$\sum_{a\leqslant k<b} f(k) = \int_a^b f(x)\,dx + \sum_{k=1}^{m} \frac{B_k}{k!} f^{(k-1)}(x)\Big|_a^b + R_m\,, \qquad (9.67)$$

$$\text{где } R_m = (-1)^{m+1}\int_a^b \frac{B_m(\{x\})}{m!} f^{(m)}(x)\,dx\,, \qquad \begin{matrix}\text{целые } a\leqslant b;\\ \text{целое } m\geqslant 1.\end{matrix} \qquad (9.68)$$

Слева находится типичная сумма, оценка которой нам может понадобиться. Справа — другое выражение для той же суммы, включающее интегралы и производные. Если $f(x)$ — достаточно "гладкая" функция, то она будет иметь m производных $f'(x), \ldots, f^{(m)}(x)$, и эта формула окажется тождеством. Выражение в правой части зачастую оказывается превосходной аппроксимацией суммы в левой части, в том смысле, что остаток R_m мал. Например, мы увидим, что приближение Стирлинга для $n!$ является следствием формулы суммирования Эйлера; то же справедливо и для нашей асимптотической аппроксимации для гармонических чисел H_n.

Числа B_k в (9.67) представляют собой числа Бернулли, с которыми мы уже встречались в главе 6; функция $B_m(\{x\})$ в (9.68) представляет собой полином Бернулли из главы 7. Запись $\{x\}$ означает дробную часть, $x - \lfloor x\rfloor$, как в главе 3. Формула суммирования Эйлера сводит все эти понятия вместе.

Вспомним значения нескольких первых чисел Бернулли; всегда удобно иметь этот список недалеко от общей формулы Эйлера:

$$B_0 = 1\,,\ B_1 = -\tfrac{1}{2}\,,\ B_2 = \tfrac{1}{6}\,,\ B_4 = -\tfrac{1}{30}\,,\ B_6 = \tfrac{1}{42}\,,\ B_8 = -\tfrac{1}{30}\,;$$
$$B_3 = B_5 = B_7 = B_9 = B_{11} = \cdots = 0\,.$$

Якоб Бернулли (Jakob Bernoulli) открыл эти числа при изучении сумм степеней целых чисел, и формула Эйлера объясняет, почему: если положить $f(x) = x^{m-1}$, то $f^{(m)}(x) = 0$; следовательно,

$R_m = 0$, и (9.67) сводится к

$$\sum_{a\leqslant k<b} k^{m-1} = \frac{x^m}{m}\bigg|_a^b + \sum_{k=1}^{m}\frac{B_k}{k!}(m-1)^{\underline{k-1}}x^{m-k}\bigg|_a^b =$$
$$= \frac{1}{m}\sum_{k=0}^{m}\binom{m}{k}B_k\cdot(b^{m-k}-a^{m-k}).$$

Например, при $m = 3$ мы получаем наш любимый пример суммирования:

$$\sum_{0\leqslant k<n} k^2 = \frac{1}{3}\left(\binom{3}{0}B_0n^3 + \binom{3}{1}B_1n^2 + \binom{3}{2}B_2n\right) =$$
$$= \frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6}.$$

Хорошего понемножку.

(Здесь мы в последний раз в этой книге вывели эту формулу.)

Перед тем как приступить к доказательству формулы Эйлера, рассмотрим соображения высшего порядка (принадлежащие Лагранжу (Lagrange) [234]) о том, почему такая формула должна иметь место. В главе 2 был определен разностный оператор Δ и объяснялось, что оператор $\sum$ — обратный к Δ, так же как $\int$ обратен оператору дифференцирования D. Можно выразить Δ через D при помощи формулы Тейлора:

$$f(x+\epsilon) = f(x) + \frac{f'(x)}{1!}\epsilon + \frac{f''(x)}{2!}\epsilon^2 + \cdots .$$

Подстановка $\epsilon = 1$ дает

$$\begin{aligned}\Delta f(x) &= f(x+1)-f(x) = \\ &= f'(x)/1!+f''(x)/2!+f'''(x)/3!+\cdots = \\ &= (D/1!+D^2/2!+D^3/3!+\cdots)\,f(x) = (e^D-1)\,f(x).\end{aligned} \tag{9.69}$$

Здесь e^D обозначает дифференциальный оператор $1 + D/1! + D^2/2! + D^3/3! + \cdots$. Поскольку $\Delta = e^D - 1$, обратным оператором должен быть $\Sigma = 1/\Delta\ 1/(e^D-1)$; а из табл. 425 мы знаем, что $z/(e^z-1) = \sum_{k\geqslant 0} B_k z^k/k!$ — степенной ряд, включающий числа Бернулли. Таким образом,

$$\sum = \frac{B_0}{D}+\frac{B_1}{1!}+\frac{B_2}{2!}D+\frac{B_3}{3!}D^2+\cdots = \int+\sum_{k\geqslant 1}\frac{B_k}{k!}D^{k-1}. \tag{9.70}$$

Применив это операторное уравнение к $f(x)$ и добавив пределы, получим

$$\sum\nolimits_a^b f(x)\,\delta x = \int_a^b f(x)\,dx + \sum_{k\geqslant 1}\frac{B_k}{k!}f^{(k-1)}(x)\bigg|_a^b, \tag{9.71}$$

что в точности представляет собой формулу суммирования Эйлера (9.67) без остаточного члена. (Ни Эйлер, ни кто-либо иной на самом деле не рассматривал остаток до тех пор, пока С. Д. Пуассон (S. D. Poisson) [295] не опубликовал в 1823 году важную работу о приближенном интегрировании. Остаточный член играет важную роль, поскольку бесконечная сумма

$$\sum_{k\geqslant 1}(B_k/k!)f^{(k-1)}(x)\Big|_a^b$$

часто оказывается расходящейся. Наш вывод (9.71) был чисто формальным, мы не касались вопросов сходимости.)

Докажем теперь формулу (9.67), включающую остаточный член. Достаточно провести доказательство только для случая $a=0$ и $b=1$, а именно — доказать, что

$$f(0) \;=\; \int_0^1 f(x)\,dx+\sum_{k=1}^{m}\frac{B_k}{k!}\,f^{(k-1)}(x)\Big|_0^1-(-1)^m\int_0^1\frac{B_m(x)}{m!}f^{(m)}(x)\,dx\,,$$

потому что затем можно заменить $f(x)$ на $f(x+l)$ для любого целого l, получив

$$f(l)=\int_l^{l+1} f(x)\,dx+\sum_{k=1}^{m}\frac{B_k}{k!}\,f^{(k-1)}(x)\Big|_l^{l+1}-(-1)^m\int_l^{l+1}\frac{B_m(\{x\})}{m!}f^{(m)}(x)\,dx$$

Общая формула (9.67) представляет собой просто сумму этих тождеств по диапазону $a\leqslant l<b$, потому что промежуточные члены благополучно сокращаются.

Доказательство при $a=0$ и $b=1$ будем проводить индукцией по m, начиная с $m=1$:

$$f(0) \;=\; \int_0^1 f(x)\,dx-\frac{1}{2}\big(f(1)-f(0)\big)+\int_0^1 (x-\tfrac{1}{2})f'(x)\,dx\,.$$

(Полином Бернулли $B_m(x)$ в общем случае определяется уравнением

$$B_m(x) \;=\; \binom{m}{0}B_0x^m+\binom{m}{1}B_1x^{m-1}+\cdots+\binom{m}{m}B_mx^0; \qquad (9.72)$$

в частности $B_1(x)=x-\frac{1}{2}$.) Другими словами, мы хотим доказать, что

$$\frac{f(0)+f(1)}{2} \;=\; \int_0^1 f(x)\,dx+\int_0^1 (x-\tfrac{1}{2})f'(x)\,dx\,.$$

Но это всего лишь частный случай формулы интегрирования по частям

$$u(x)v(x)\Big|_0^1 = \int_0^1 u(x)\,dv(x) + \int_0^1 v(x)\,du(x) \qquad (9.73)$$

для $u(x) = f(x)$ и $v(x) = x - \frac{1}{2}$. Таким образом, случай $m = 1$ оказался простым.

Чтобы перейти от $m-1$ к m и завершить индукцию при $m > 1$, нам надо показать, что $R_{m-1} = (B_m/m!)f^{(m-1)}(x)\big|_0^1 + R_m$, а именно — что выполняется равенство

$$(-1)^m \int_0^1 \frac{B_{m-1}(x)}{(m-1)!} f^{(m-1)}(x)\,dx = \\ = \frac{B_m}{m!} f^{(m-1)}(x)\Big|_0^1 - (-1)^m \int_0^1 \frac{B_m(x)}{m!} f^{(m)}(x)\,dx\,.$$

Оно приводится к уравнению

$$(-1)^m B_m f^{(m-1)}(x)\Big|_0^1 = m\int_0^1 B_{m-1}(x) f^{(m-1)}(x)\,dx + \\ + \int_0^1 B_m(x) f^{(m)}(x)\,dx\,.$$

Авторы хотя бы в конце книги станут серьезными?

И снова к этим двум интегралам применимо соотношение (9.73), с $u(x) = f^{(m-1)}(x)$ и $v(x) = B_m(x)$, потому что производная полинома Бернулли (9.72) есть

$$\frac{d}{dx}\sum_k \binom{m}{k} B_k x^{m-k} = \sum_k \binom{m}{k}(m-k)B_k x^{m-k-1} = \\ = m\sum_k \binom{m-1}{k} B_k x^{m-1-k} = \\ = mB_{m-1}(x)\,. \qquad (9.74)$$

(Здесь полезно тождество поглощения (5.7).) Таким образом, требуемая формула будет иметь место в том и только в том случае, когда

$$(-1)^m B_m f^{(m-1)}(x)\big|_0^1 = B_m(x) f^{(m-1)}(x)\big|_0^1\,.$$

Другими словами, необходимо, чтобы

$$(-1)^m B_m = B_m(1) = B_m(0) \qquad \text{для } m > 1. \qquad (9.75)$$

Это может слегка смутить, потому что ясно, что $B_m(0)$ равно B_m, а не $(-1)^m B_m$. Но на самом деле здесь все в порядке, поскольку $m > 1$; мы знаем, что B_m для четных m равно нулю. (Но опасность была близка.)

Для завершения доказательства формулы суммирования Эйлера осталось показать, что $B_m(1) = B_m(0)$, а это то же самое, что сказать

$$\sum_k \binom{m}{k} B_k = B_m \qquad \text{для } m > 1.$$

Но это просто определение чисел Бернулли, (6.79), так что доказательство завершено.

Тождество $B'_m(x) = mB_{m-1}(x)$ означает, что

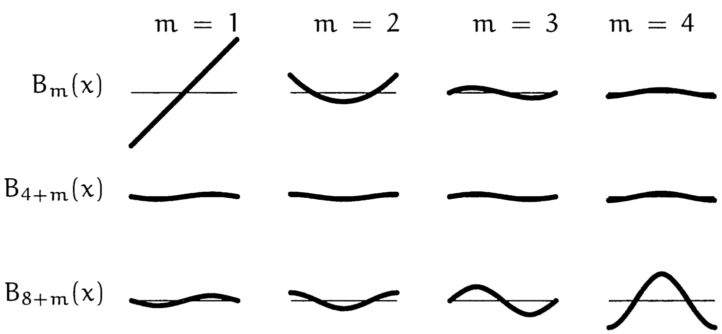
$$\int_0^1 B_m(x)\,dx = \frac{B_{m+1}(1) - B_{m+1}(0)}{m+1},$$

а мы знаем, что этот интеграл равен нулю при $m \geqslant 1$. Следовательно, в остаточном члене формулы Эйлера

$$R_m = \frac{(-1)^{m+1}}{m!} \int_a^b B_m\bigl(\{x\}\bigr) f^{(m)}(x)\,dx$$

множителем при $f^{(m)}(x)$ стоит функция $B_m\bigl(\{x\}\bigr)$, среднее значение которой равно нулю. Это означает, что R_m, скорее всего, окажется малым.

Давайте более внимательно посмотрим на функции $B_m(x)$ при $0 \leqslant x \leqslant 1$, поскольку $B_m(x)$ управляет поведением R_m. Ниже приведены графики $B_m(x)$ для первых двенадцати значений m:



Хотя полиномы от $B_3(x)$ до $B_9(x)$ достаточно малы, в конечном итоге полиномы и числа Бернулли сильно растут. К счастью, R_m содержит компенсирующий множитель $1/m!$, позволяющий усмирить стихию.

График $B_m(x)$ при $m \geqslant 3$ начинает выглядеть очень похоже на волну синуса; в упр. 58 доказывается, что и на самом деле функция $B_m(x)$ может быть аппроксимирована отрицательным кратным $\cos(2\pi x - \frac{1}{2}\pi m)$ с погрешностью $O\bigl(2^{-m} \max_x B_m(\{x\})\bigr)$.

В общем случае функция $B_{4k+1}(x)$ отрицательна при $0 < x < \frac{1}{2}$ и положительна при $\frac{1}{2} < x < 1$. Следовательно, ее интеграл $B_{4k+2}(x)/(4k+2)$ убывает при $0 < x < \frac{1}{2}$ и возрастает при $\frac{1}{2} < x < 1$. Более того, имеет место равенство

$$B_{4k+1}(1-x) \;=\; -B_{4k+1}(x) \qquad \text{для } 0 \leqslant x \leqslant 1,$$

откуда следует

$$B_{4k+2}(1-x) \;=\; B_{4k+2}(x) \qquad \text{для } 0 \leqslant x \leqslant 1.$$

Постоянное слагаемое B_{4k+2} обеспечивает нулевое значение интеграла $\int_0^1 B_{4k+2}(x)\,dx$; следовательно, $B_{4k+2} > 0$. Интеграл от $B_{4k+2}(x)$ представляет собой функцию $B_{4k+3}(x)/(4k+3)$, которая, таким образом, должна быть положительна при $0 < x < \frac{1}{2}$ и отрицательна при $\frac{1}{2} < x < 1$; кроме того, $B_{4k+3}(1-x) = -B_{4k+3}(x)$, так что $B_{4k+3}(x)$ обладает всеми свойствами, указанными для $B_{4k+1}(x)$, но с обратным знаком. Следовательно, $B_{4k+4}(x)$ имеет свойства функции $B_{4k+2}(x)$, но с обратным знаком. Следовательно, $B_{4k+5}(x)$ имеет свойства функции $B_{4k+1}(x)$; на этом цикл завершен, и индуктивно установлена справедливость указанных свойств для всех k.

В соответствии с выполненным выше анализом максимальное значение $B_{2m}(x)$ должно достигаться либо при $x = 0$, либо при $x = \frac{1}{2}$. В упр. 17 доказывается, что

$$B_{2m}(\tfrac{1}{2}) \;=\; (2^{1-2m}-1)B_{2m}\,; \tag{9.76}$$

следовательно, имеем

$$\bigl|B_{2m}(\{x\})\bigr| \;\leqslant\; |B_{2m}|\,. \tag{9.77}$$

Это неравенство можно использовать для получения полезной оценки верхней границы остатка в формуле суммирования Эйлера, потому что из (6.89) нам известно, что

$$\frac{|B_{2m}|}{(2m)!} \;=\; \frac{2}{(2\pi)^{2m}}\sum_{k\geqslant 1}\frac{1}{k^{2m}} \;=\; O\bigl((2\pi)^{-2m}\bigr) \qquad \text{при } m > 0.$$

Следовательно, формулу суммирования Эйлера (9.67) можно переписать следующим образом:

$$\sum_{a\leqslant k<b} f(k) \;=\; \int_a^b f(x)\,dx - \frac{1}{2}f(x)\Big|_a^b + \sum_{k=1}^{m}\frac{B_{2k}}{(2k)!}f^{(2k-1)}(x)\Big|_a^b + \\ + O\bigl((2\pi)^{-2m}\bigr)\int_a^b \bigl|f^{(2m)}(x)\bigr|\,dx\,. \tag{9.78}$$

Например, при $f(x) = e^x$ все производные одинаковы и последняя формула гласит, что $\sum_{a\leqslant k<b} e^k = (e^b - e^a)\bigl(1 - \frac{1}{2} + B_2/2! + B_4/4! + \cdots + B_{2m}/(2m)!\bigr) + O\bigl((2\pi)^{-2m}\bigr)$. Конечно, мы знаем, что на самом деле это сумма геометрической прогрессии, равная $(e^b - e^a)/(e-1) = (e^b - e^a)\sum_{k\geqslant 0} B_k/k!$.

Если $f^{(2m)}(x) \geqslant 0$ при $a \leqslant x \leqslant b$, то интеграл $\int_a^b |f^{(2m)}(x)|\,dx$ представляет собой просто $f^{(2m-1)}(x)\big|_a^b$, так что мы имеем

$$|R_{2m}| \leqslant \left| \frac{B_{2m}}{(2m)!} f^{(2m-1)}(x)\Big|_a^b \right|;$$

другими словами, в этом случае остаточный член ограничен величиной *последнего члена* (находящегося непосредственно перед остаточным). Можно дать еще лучшую оценку, если известно, что

$$f^{(2m+2)}(x) \geqslant 0 \quad \text{и} \quad f^{(2m+4)}(x) \geqslant 0 \quad \text{для } a \leqslant x \leqslant b. \tag{9.79}$$

Оказывается, отсюда вытекает соотношение

$$R_{2m} = \theta_m \frac{B_{2m+2}}{(2m+2)!} f^{(2m+1)}(x)\Big|_a^b \qquad \text{для некоторого } 0 \leqslant \theta_m \leqslant 1; \tag{9.80}$$

т.е. остаточный член лежит между 0 и *первым отброшенным членом* в (9.78) — членом, который мы добавили бы вслед за последним при увеличении m.

Вот доказательство: формула суммирования Эйлера справедлива при всех m, и $B_{2m+1} = 0$ при $m > 0$; следовательно, $R_{2m} = R_{2m+1}$, и первый отброшенный член должен быть

$$R_{2m} - R_{2m+2}.$$

Таким образом, нам надо доказать, что R_{2m} лежит между 0 и $R_{2m} - R_{2m+2}$; а это имеет место тогда и только тогда, когда R_{2m} и R_{2m+2} имеют противоположные знаки. Мы утверждаем, что

$$f^{(2m+2)}(x) \geqslant 0 \quad \text{для } a\leqslant x\leqslant b \quad \text{влечет} \quad (-1)^m R_{2m} \geqslant 0. \tag{9.81}$$

Вместе с (9.79) это доказывает, что R_{2m} и R_{2m+2} имеют противоположные знаки, так что на этом доказательство (9.80) завершается.

Доказать (9.81) нетрудно, если вспомнить определение R_{2m+1} и доказанные ранее факты о графике $B_{2m+1}(x)$. А именно, имеем

$$R_{2m} = R_{2m+1} = \int_a^b \frac{B_{2m+1}\bigl(\{x\}\bigr)}{(2m+1)!} f^{(2m+1)}(x)\,dx,$$

причем $f^{(2m+1)}(x)$ возрастает, потому что ее производная $f^{(2m+2)}(x)$ положительна. (Говоря более точно, $f^{(2m+1)}(x)$ — неубывающая функция, так как ее производная неотрицательна.) График $B_{2m+1}(\{x\})$ выглядит как умноженная на $(-1)^{m+1}$ синусоида, так что из геометрических соображений ясно, что вторая половина каждой волны синусоиды, будучи умноженной на возрастающую функцию, оказывает большее влияние на интеграл, чем первая. Это делает $(-1)^m R_{2m+1} \geqslant 0$, что нам и требовалось. В упр. 16 этот результат доказывается формально.

9.6 Завершающее суммирование

Перед тем как завершить книгу, просуммируем все и подведем итоги. Рассмотрим несколько важных и интересных примеров и применим к ним формулу суммирования Эйлера.

Сумма 1: слишком простая

Но начнем мы с рассмотрения интересного и *не важного* примера, а именно — с суммы, которую мы уже умеем вычислять. Давайте посмотрим, что скажет нам формула суммирования Эйлера при ее применении к телескопической сумме

$$S_n = \sum_{1\leqslant k<n} \frac{1}{k(k+1)} = \sum_{1\leqslant k<n} \left(\frac{1}{k} - \frac{1}{k+1}\right) = 1 - \frac{1}{n}.$$

Наверное, не помешает предварить серьезные применения формулы суммирования Эйлера анализом асимптотического эквивалента матрешки.

Начать можно с записи функции $f(x) = 1/x(x+1)$ через простейшие дроби,

$$f(x) = \frac{1}{x} - \frac{1}{x+1},$$

поскольку это упрощает интегрирование и дифференцирование. Итак, мы имеем $f'(x) = -1/x^2 + 1/(x+1)^2$ и $f''(x) = 2/x^3 - 2/(x+1)^3$; в общем случае

$$f^{(k)}(x) = (-1)^k k!\left(\frac{1}{x^{k+1}} - \frac{1}{(x+1)^{k+1}}\right) \qquad \text{для } k \geqslant 0.$$

Далее,

$$\int_1^n f(x)\,dx = \ln x - \ln(x+1)\Big|_1^n = \ln\frac{2n}{n+1}.$$

Подставляя эти выражения в формулу суммирования (9.67), получаем

$$S_n = \ln\frac{2n}{n+1} - \sum_{k=1}^{m}(-1)^k\frac{B_k}{k}\left(\frac{1}{n^k} - \frac{1}{(n+1)^k} - 1 + \frac{1}{2^k}\right) + R_m(n),$$

где $R_m(n) = -\int_1^n B_m(\{x\})\left(\frac{1}{x^{m+1}} - \frac{1}{(x+1)^{m+1}}\right)dx$.

Например, при $m = 4$ правая часть равна

$$\ln\frac{2n}{n+1} - \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+1} - \frac{1}{2}\right) - \frac{1}{12}\left(\frac{1}{n^2} - \frac{1}{(n+1)^2} - \frac{3}{4}\right)$$
$$+ \frac{1}{120}\left(\frac{1}{n^4} - \frac{1}{(n+1)^4} - \frac{15}{16}\right) + R_4(n).$$

Ох, похоже, тут что-то не так — что-то не видно ничего общего с правильным ответом $1 - n^{-1}$. Но не будем останавливаться; пойдем дальше и посмотрим, что здесь можно сделать. Мы знаем, как расписать слагаемые правой части через отрицательные степени n до, скажем, $O(n^{-5})$:

$$\begin{aligned}
\ln\frac{n}{n+1} &= -n^{-1} + \tfrac{1}{2}n^{-2} - \tfrac{1}{3}n^{-3} + \tfrac{1}{4}n^{-4} + O(n^{-5});\\
\frac{1}{n+1} &= n^{-1} - n^{-2} + n^{-3} - n^{-4} + O(n^{-5});\\
\frac{1}{(n+1)^2} &= n^{-2} - 2n^{-3} + 3n^{-4} + O(n^{-5});\\
\frac{1}{(n+1)^4} &= n^{-4} + O(n^{-5}).
\end{aligned}$$

Следовательно, сумма членов в правой части приближения дает

$$\begin{aligned}
&\ln 2 + \tfrac{1}{4} + \tfrac{1}{16} - \tfrac{1}{128} + \left(-1 - \tfrac{1}{2} + \tfrac{1}{2}\right)n^{-1} + \left(\tfrac{1}{2} - \tfrac{1}{2} - \tfrac{1}{12} + \tfrac{1}{12}\right)n^{-2} +\\
&\quad + \left(-\tfrac{1}{3} + \tfrac{1}{2} - \tfrac{2}{12}\right)n^{-3} + \left(\tfrac{1}{4} - \tfrac{1}{2} + \tfrac{3}{12} + \tfrac{1}{120} - \tfrac{1}{120}\right)n^{-4} + R_4(n) =\\
&\qquad = \ln 2 + \tfrac{39}{128} - n^{-1} + R_4(n) + O(n^{-5}).
\end{aligned}$$

Коэффициенты при n^{-2}, n^{-3} и n^{-4}, как им и положено, прекрасно сократились.

Если бы все в мире шло как надо, то мы должны были бы суметь доказать, что $R_4(n)$ асимптотически мало, может быть, $O(n^{-5})$, и в результате получили бы аппроксимацию нашей суммы. Только вот мы не сможем этого сделать, потому что мы знаем, что правильный постоянный член равен 1, а не $\ln 2 + \frac{39}{128}$ (что, хотя и близко к 1, но все же не равно ей, а примерно равно

0.9978). Таким образом, $R_4(n)$ в действительности равно $\frac{89}{128} - \ln 2 + O(n^{-5})$, только вот формула суммирования Эйлера ничего нам об этом не говорит.

Другими словами, нас постигла неудача.

Один из способов справиться с проблемой заключается в том, чтобы заметить, что постоянное слагаемое в аппроксимации по мере роста m следует определенной схеме:

$$\ln 2 - \tfrac{1}{2}B_1 + \tfrac{1}{2}\cdot\tfrac{3}{4}B_2 - \tfrac{1}{3}\cdot\tfrac{7}{8}B_3 + \tfrac{1}{4}\cdot\tfrac{15}{16}B_4 - \tfrac{1}{5}\cdot\tfrac{31}{32}B_5 + \cdots .$$

Может быть, мы покажем, что этот ряд сходится к 1 при бесконечном количестве слагаемых? Нет — числа Бернулли становятся очень большими. Например, $B_{22} = \frac{854513}{138} > 6192$; следовательно, $|R_{22}(n)|$ будет гораздо больше, чем $|R_4(n)|$. И тут тоже неудача.

Тем не менее выход есть, и этот обходной путь, оказывается, очень полезен и в других приложениях формулы Эйлера. Ключевой момент здесь — заметить, что $R_4(n)$ стремится к некоторому пределу при $n \to \infty$:

$$\lim_{n\to\infty} R_4(n) \;=\; -\int_1^\infty B_4(\{x\})\Bigl(\frac{1}{x^5} - \frac{1}{(x+1)^5}\Bigr)\,dx \;=\; R_4(\infty)\,.$$

Интеграл $\int_1^\infty B_m(\{x\})\, f^{(m)}(x)\,dx$ существует, если $f^{(m)}(x)=O(x^{-2})$ при $x \to \infty$, и это несомненно так в случае $f^{(4)}(x)$. Кроме того, имеем

$$\begin{aligned} R_4(n) \;&=\; R_4(\infty) + \int_n^\infty B_4(\{x\})\Bigl(\frac{1}{x^5} - \frac{1}{(x+1)^5}\Bigr)\,dx \;= \\ &=\; R_4(\infty) + O\Bigl(\int_n^\infty x^{-6}\,dx\Bigr) \;=\; R_4(\infty) + O(n^{-5})\,. \end{aligned}$$

Таким образом, мы доказали с помощью формулы Эйлера, что

$$\begin{aligned} \sum_{1\leqslant k<n} \frac{1}{k(k+1)} \;&=\; \ln 2 + \tfrac{39}{128} - n^{-1} + R_4(\infty) + O(n^{-5}) \;= \\ &=\; C - n^{-1} + O(n^{-5}) \end{aligned}$$

для некоторой константы C. Мы не знаем, какова эта константа — для ее определения требуются другие методы, — но формула суммирования Эйлера позволила нам установить ее существование.

Предположим, что мы выбрали гораздо большее значение m. Тогда те же рассуждения покажут, что

$$R_m(n) \;=\; R_m(\infty) + O(n^{-m-1})\,,$$

и мы получим формулу

$$\sum_{1\leqslant k<n}\frac{1}{k(k+1)} = C-n^{-1}+c_2n^{-2}+c_3n^{-3}+\cdots+c_mn^{-m}+O(n^{-m-1})$$

для некоторых констант c_2, c_3, ... Мы знаем, что в нашем случае все константы c оказываются нулевыми; тем не менее докажем это, хотя бы для того, чтобы восстановить репутацию (если не свою, то по крайней мере формулы Эйлера). Член $\ln\frac{n}{n+1}$ вносит в c_m слагаемое $(-1)^m/m$; член $(-1)^{m+1}(B_m/m)n^{-m}$ вносит $(-1)^{m+1}B_m/m$, а вклад члена $(-1)^k(B_k/k)(n+1)^{-k}$ составляет $(-1)^m\binom{m-1}{k-1}B_k/k$. Следовательно,

$$\begin{aligned}(-1)^m c_m &= \frac{1}{m}-\frac{B_m}{m}+\sum_{k=1}^{m}\binom{m-1}{k-1}\frac{B_k}{k} = \\ &= \frac{1}{m}-\frac{B_m}{m}+\frac{1}{m}\sum_{k=1}^{m}\binom{m}{k}B_k = \frac{1}{m}\bigl(1-B_m+B_m(1)-1\bigr).\end{aligned}$$

Совершенно ясно, что это нуль для $m>1$. Мы доказали, что

$$\sum_{1\leqslant k<n}\frac{1}{k(k+1)} = C-n^{-1}+O(n^{-m-1}) \quad \text{для всех } m\geqslant 1. \qquad (9.82)$$

Этого недостаточно, чтобы доказать, что сумма точно равна $C-n^{-1}$; она может быть равна $C-n^{-1}+2^{-n}$ или чему-то еще в этом роде. Но формула суммирования Эйлера все же дала нам точность $O(n^{-m-1})$ для произвольно большого m, даже несмотря на то, что мы не оценивали явно ни одного остаточного члена.

Опять сумма 1: выводы и обобщения

Перед тем как окончательно разделаться с нашим тренировочным примером, попробуем вновь взглянуть на сделанное в более широком аспекте. Мы начинали с суммы

$$S_n = \sum_{1\leqslant k<n} f(k)$$

и, воспользовавшись формулой суммирования Эйлера, получили

$$S_n = F(n)-F(1)+\sum_{k=1}^{m}\bigl(T_k(n)-T_k(1)\bigr)+R_m(n), \qquad (9.83)$$

где $F(x)$ есть $\int f(x)\,dx$, а $T_k(x)$ — определенное выражение, включающее B_k и $f^{(k-1)}(x)$. Мы также заметили, что существует константа c, такая, что

$$f^{(m)}(x) = O(x^{c-m}) \qquad \text{при } x\to\infty \text{ для всех достаточно больших } m.$$

(А именно, у нас было $f(k) = 1/k(k+1)$; $F(x) = \ln\bigl(x/(x+1)\bigr)$; $c = -2$ и $T_k(x) = (-1)^{k+1}(B_k/k)\bigl(x^{-k} - (x+1)^{-k}\bigr)$.) Это означает, что для всех достаточно больших m остаточные члены имеют маленький хвост,

$$\begin{aligned} R'_m(n) &= R_m(\infty) - R_m(n) = \\ &= (-1)^{m+1} \int_n^\infty \frac{B_m(\{x\})}{m!} f^{(m)}(x)\,dx = O(n^{c+1-m}). \end{aligned} \tag{9.84}$$

Следовательно, можно заключить, что существует константа C, такая, что

$$S_n = F(n) + C + \sum_{k=1}^{m} T_k(n) - R'_m(n). \tag{9.85}$$

(Заметьте, что C, к нашему счастью, вобрала в себя неприятные члены $T_k(1)$.)

Можно сэкономить усилия в будущих задачах, сразу утверждая существование константы C в тех случаях, когда $R_m(\infty)$ существует.

Предположим, что $f^{(2m+2)}(x) \geqslant 0$ и $f^{(2m+4)}(x) \geqslant 0$ для $1 \leqslant x \leqslant n$. Мы доказали, что отсюда следует простая оценка (9.80) остаточного члена:

$$R_{2m}(n) = \theta_{m,n}\bigl(T_{2m+2}(n) - T_{2m+2}(1)\bigr),$$

где $\theta_{m,n}$ лежит где-то между 0 и 1. Но на самом деле нас не интересует оценка, содержащая $R_{2m}(n)$ и $T_{2m+2}(1)$; мы, в конце концов, избавились от $T_k(1)$ путем введения константы C. Что нам действительно нужно, так это оценка вида

$$-R'_{2m}(n) = \phi_{m,n} T_{2m+2}(n),$$

где $0 < \phi_{m,n} < 1$; она позволила бы вывести из (9.85) соотношение

$$S_n = F(n) + C + T_1(n) + \sum_{k=1}^{m} T_{2k}(n) + \phi_{m,n} T_{2m+2}(n), \tag{9.86}$$

и здесь остаточный член действительно будет лежать между нулем и первым отброшенным членом.

Незначительное изменение наших предыдущих рассуждений полностью исправляет положение. Допустим, что

$$f^{(2m+2)}(x) \geqslant 0 \quad \text{и} \quad f^{(2m+4)}(x) \geqslant 0 \qquad \text{при } x \to \infty. \tag{9.87}$$

Правая часть (9.85), если рассматривать только остаточные члены, имеет точно такой же вид, как формула суммирования Эйлера (9.67) для $a = n$ и $b = \infty$, взятая с обратным знаком, и последующие остаточные члены получаются индукцией по m. Следовательно, в этом случае можно применить наши предыдущие рассуждения.

Сумма 2: гармония гармонических чисел

Теперь, научившись столь многому на тривиальном (но надежном) примере, мы легко справимся с нетривиальным. Используем формулу суммирования Эйлера для вывода уже известной нам аппроксимации чисел H_n.

В этом случае $f(x) = 1/x$. Мы уже изучили интеграл и производные f, разбираясь с суммой 1; $f^{(m)}(x) = O(x^{-m-1})$ при $x \to \infty$. Следовательно, можно сразу же использовать формулу (9.85):

$$\sum_{1 \leqslant k < n} \frac{1}{k} = \ln n + C + B_1 n^{-1} - \sum_{k=1}^{m} \frac{B_{2k}}{2kn^{2k}} - R'_{2m}(n)$$

для некоторой константы C. Сумма слева есть H_{n-1}, а не H_n; однако более удобно работать с H_{n-1} и добавить $1/n$ позже, чем возиться с $(n+1)$ в правой части. Тогда $B_1 n^{-1}$ станет $(B_1+1)n^{-1} = 1/(2n)$. Обозначим константу буквой 'γ' вместо 'C', поскольку константа Эйлера γ фактически определяется как $\lim_{n\to\infty}(H_n - \ln n)$.

С помощью только что разработанной теории можно хорошо оценить остаточный член, потому что $f^{(2m)}(x) = (2m)!/x^{2m+1} \geqslant 0$ для всех $x > 0$. Следовательно, из (9.86) получаем

$$H_n = \ln n + \gamma + \frac{1}{2n} - \sum_{k=1}^{m} \frac{B_{2k}}{2kn^{2k}} - \theta_{m,n} \frac{B_{2m+2}}{(2m+2)n^{2m+2}}, \qquad (9.88)$$

где $\theta_{m,n}$ — некоторая дробь между 0 и 1. Первые члены этой общей формулы приведены в табл. 545. Например, для $m = 2$ получаем

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \frac{\theta_{2,n}}{252n^6}. \qquad (9.89)$$

Это уравнение, кстати, дает хорошее приближение γ даже при $n = 2$:

$$\gamma = H_2 - \ln 2 - \tfrac{1}{4} + \tfrac{1}{48} - \tfrac{1}{1920} + \epsilon = 0.577165\ldots + \epsilon,$$

где ϵ лежит между 0 и $\frac{1}{16128}$. Взяв $n = 10^4$ и $m = 250$, получим значение γ с 1271 верным десятичным знаком, начинающееся как

$$\gamma = 0.57721\,56649\,01532\,86060\,65120\,90082\,40243\ldots. \qquad (9.90)$$

Но константа Эйлера встречается и в других формулах, которые позволяют вычислять ее еще эффективнее [345].

Сумма 3: приближение Стирлинга

Если $f(x) = \ln x$, то $f'(x) = 1/x$, так что можно оценить сумму логарифмов, используя практически те же вычисления, что и для суммы обратных величин. Формула суммирования Эйлера дает

$$\sum_{1\leqslant k<n} \ln k = n\ln n - n + \sigma - \frac{\ln n}{2} + \\ + \sum_{k=1}^{m} \frac{B_{2k}}{2k(2k-1)n^{2k-1}} + \varphi_{m,n}\frac{B_{2m+2}}{(2m+2)(2m+1)n^{2m+1}},$$

где σ — некоторая "константа Стирлинга", и $0 < \varphi_{m,n} < 1$. (В этом случае $f^{(2m)}(x)$ не положительна, а отрицательна; но можно по-прежнему утверждать, что остаточный член подчинен первому отброшенному, потому что можно начать с функции $f(x) = -\ln x$ вместо $f(x) = \ln x$.) Добавление к обеим частям $\ln n$ при $m = 2$ дает

$$\ln n! = n\ln n - n + \frac{\ln n}{2} + \sigma + \frac{1}{12n} - \frac{1}{360n^3} + \frac{\varphi_{2,n}}{1260n^5}. \qquad (9.91)$$

Аппроксимацию из табл. 545 можно получить, применив 'exp' к обеим частям. (Значение e^σ оказывается равным $\sqrt{2\pi}$, но пока что мы не готовы вывести эту формулу. В действительности сам Стирлинг не знал точного значения константы еще несколько лет после того, как де Муавр [76] доказал ее существование.)

Если m фиксировано, а $n \to \infty$, то общая формула дает все лучшие и лучшие приближения $\ln n!$ в смысле абсолютной погрешности, а следовательно, эти приближения являются все улучшающимися приближениями $n!$ в смысле погрешности относительной. Но если n фиксировано, а m возрастает, то оценка погрешности $|B_{2m+2}|/(2m+2)(2m+1)n^{2m+1}$ убывает только до некоторой точки, после чего начинает расти. Таким образом, у данной аппроксимации $n!$ имеется некоторый порог точности, переступить который мешает что-то вроде принципа неопределенности.

Не занимался ли Гейзенберг математикой?

В главе 5, в соотношении (5.83), мы обобщили факториалы на произвольные действительные α при помощи определения

$$\frac{1}{\alpha!} = \lim_{n\to\infty}\binom{n+\alpha}{n}n^{-\alpha},$$

предложенного Эйлером. Предположим, что α — большое число; тогда

$$\ln\alpha! = \lim_{n\to\infty}\Bigl(\alpha\ln n + \ln n! - \sum_{k=1}^{n}\ln(\alpha+k)\Bigr),$$

и для оценки этой суммы можно применить формулу суммирования Эйлера с $f(x)=\ln(x+\alpha)$:

$$\begin{aligned}\sum_{k=1}^{n}\ln(k+\alpha) &= F_m(\alpha,n)-F_m(\alpha,0)+R_{2m}(\alpha,n)\,,\\ F_m(\alpha,x) &= (x+\alpha)\ln(x+\alpha)-x+\frac{\ln(x+\alpha)}{2}+\\ &\qquad +\sum_{k=1}^{m}\frac{B_{2k}}{2k(2k-1)(x+\alpha)^{2k-1}}\,,\\ R_{2m}(\alpha,n) &= \int_0^n\frac{B_{2m}(\{x\})}{2m}\frac{dx}{(x+\alpha)^{2m}}\,.\end{aligned}$$

(Здесь мы воспользовались (9.67) с $a=0$ и $b=n$, а затем добавили к обеим частям $\ln(n+\alpha)-\ln\alpha$.) Если вычесть это приближение суммы $\sum_{k=1}^{n}\ln(k+\alpha)$ из приближения Стирлинга для $\ln n!$, добавить $\alpha\ln n$ и перейти к пределу при $n\to\infty$, то можно получить

$$\begin{aligned}\ln\alpha! &= \alpha\ln\alpha-\alpha+\frac{\ln\alpha}{2}+\sigma+\\ &\qquad+\sum_{k=1}^{m}\frac{B_{2k}}{(2k)(2k-1)\alpha^{2k-1}}-\int_0^\infty\frac{B_{2m}(\{x\})}{2m}\frac{dx}{(x+\alpha)^{2m}}\,,\end{aligned}$$

потому что $\alpha\ln n+n\ln n-n+\frac12\ln n-(n+\alpha)\ln(n+\alpha)+n-\frac12\ln(n+\alpha)\to-\alpha$, а остальные, не показанные здесь члены, стремятся к нулю. Таким образом, аппроксимация Стирлинга для обобщенных факториалов (и для гамма-функции $\Gamma(\alpha+1)=\alpha!$) выглядит точно так же, как и для обычных факториалов.

Сумма 4: колоколообразные слагаемые

Обратимся теперь к сумме совершенно иного характера:

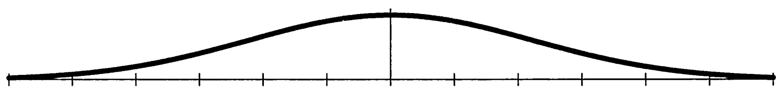
$$\begin{aligned}\Theta_n &= \sum_k e^{-k^2/n} = \qquad\qquad (9.92)\\ &= \cdots+e^{-9/n}+e^{-4/n}+e^{-1/n}+1+e^{-1/n}+e^{-4/n}+e^{-9/n}+\cdots\,.\end{aligned}$$

Это дважды бесконечная сумма, члены которой достигают максимального значения $e^0=1$ при $k=0$. Мы назвали ее Θ_n, потому что это степенной ряд, содержащий величину $e^{-1/n}$ в степени

$p(k)$, где $p(k)$ — полином степени 2; такие степенные ряды традиционно называются тета-функциями. Если $n = 10^{100}$, то

$$e^{-k^2/n} = \begin{cases} e^{-.01} \approx 0.99005 & \text{при } k = 10^{49}; \\ e^{-1} \approx 0.36788 & \text{при } k = 10^{50}; \\ e^{-100} < 10^{-43} & \text{при } k = 10^{51}. \end{cases}$$

Таким образом, слагаемые остаются очень близки к 1, пока k не достигнет значений порядка $\sqrt{n}$, когда слагаемые уменьшаются и становятся очень близки к нулю. Можно предположить, что Θ_n будет пропорционально $\sqrt{n}$. Вот как выглядит график $e^{-k^2/n}$ для $n = 10$:



При бо́льших значениях n график просто растянется по горизонтали в $\sqrt{n}$ раз.

Чтобы оценить Θ_n с помощью формулы суммирования Эйлера, положим $f(x) = e^{-x^2/n}$ и $a = -\infty$, $b = +\infty$. (Если бесконечности вас пугают, то возьмите $a = -A$ и $b = +B$, а затем перейдите к пределу $A, B \to \infty$.) Интеграл от $f(x)$ равен

$$\int_{-\infty}^{+\infty} e^{-x^2/n}\,dx = \sqrt{n}\int_{-\infty}^{+\infty} e^{-u^2}\,du = \sqrt{n}\,C,$$

если заменить x на $u\sqrt{n}$. Значение $\int_{-\infty}^{+\infty} e^{-u^2}\,du$ хорошо известно, но пока что мы обозначим его буквой C с тем, чтобы вернуться к нему после того, как закончим разбирательство с формулой суммирования Эйлера.

Следующее, что нам нужно знать, — это последовательность производных $f'(x)$, $f''(x)$, ... ; для их нахождения удобно воспользоваться заменой

$$f(x) = g(x/\sqrt{n}), \qquad g(x) = e^{-x^2}.$$

Тогда правило замены переменной из дифференциального исчисления говорит нам, что

$$\frac{df(x)}{dx} = \frac{dg(y)}{dy}\frac{dy}{dx}, \qquad y = \frac{x}{\sqrt{n}};$$

а это то же самое, что

$$f'(x) = \frac{1}{\sqrt{n}}\,g'(x/\sqrt{n}).$$

По индукции

$$f^{(k)}(x) \;=\; n^{-k/2}g^{(k)}\big(x/\sqrt{n}\,\big)\,.$$

Например, $g'(x) = -2xe^{-x^2}$ и $g''(x) = (4x^2-2)e^{-x^2}$; следовательно,

$$f'(x) \;=\; \frac{1}{\sqrt{n}}\left(-2\frac{x}{\sqrt{n}}\right)e^{-x^2/n}\,,$$
$$f''(x) \;=\; \frac{1}{n}\left(4\Big(\frac{x}{\sqrt{n}}\Big)^2-2\right)e^{-x^2/n}\,.$$

Чтобы было легче увидеть, что произойдет в дальнейшем, лучше работать с более простой функцией $g(x)$.

Нам не понадобится точно вычислять производные $g(x)$, потому что нас интересуют только предельные значения при $x = \pm\infty$. Нам достаточно заметить, что любая производная функции $g(x)$ есть произведение e^{-x^2} и некоторого полинома от x:

$$g^{(k)}(x) \;=\; P_k(x)e^{-x^2}\,, \qquad \text{где } P_k \text{ — полином степени } k.$$

Это доказывается по индукции.

Экспонента e^{-x^2} стремится к нулю при $x \to \pm\infty$ гораздо быстрее, чем $P_k(x)$ стремится к бесконечности, поэтому мы делаем вывод, что

$$f^{(k)}(+\infty) \;=\; f^{(k)}(-\infty) \;=\; 0$$

для всех $k \geqslant 0$. Следовательно, все члены в

$$\sum_{k=1}^{m}\frac{B_k}{k!}f^{(k-1)}(x)\Big|_{-\infty}^{+\infty}$$

исчезают, и остается лишь слагаемое с $\int f(x)\,dx$ и остаточный член:

$$\begin{aligned}
\Theta_n \;&=\; C\sqrt{n}+(-1)^{m+1}\int_{-\infty}^{+\infty}\frac{B_m\big(\{x\}\big)}{m!}f^{(m)}(x)\,dx \;=\\
&=\; C\sqrt{n}+\frac{(-1)^{m+1}}{n^{m/2}}\int_{-\infty}^{+\infty}\frac{B_m\big(\{x\}\big)}{m!}g^{(m)}\left(\frac{x}{\sqrt{n}}\right)dx \;=\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (x = u\sqrt{n}\,)\\
&=\; C\sqrt{n}+\frac{(-1)^{m+1}}{n^{(m-1)/2}}\int_{-\infty}^{+\infty}\frac{B_m\big(\{u\sqrt{n}\}\big)}{m!}P_m(u)e^{-u^2}\,du \;=\\
&=\; C\sqrt{n}+O(n^{(1-m)/2})\,.
\end{aligned}$$

Выписанная O-оценка остаточного члена вытекает из того факта, что $\left|B_m(\{u\sqrt{n}\})\right|$ ограничено и интеграл $\int_{-\infty}^{+\infty}|P(u)|e^{-u^2}\,du$ существует, если P — полином. (Константа в этом O зависит от m.)

Мы доказали, что $\Theta_n = C\sqrt{n} + O(n^{-M})$ для сколь угодно большого M; разность между Θ_n и $C\sqrt{n}$ "экспоненциально мала". Займемся теперь определением константы C, играющей столь важную роль в значении Θ_n.

Один из способов найти C состоит в поиске значения интеграла в таблице; однако предпочтительно самостоятельно вычислить это значение, чтобы в будущем уметь подсчитать и те интегралы, которых нет в таблицах. Для вычисления C достаточно элементарных знаний, если догадаться рассмотреть не одинарный, а двойной интеграл

$$C^2 = \int_{-\infty}^{+\infty} e^{-x^2}\,dx \int_{-\infty}^{+\infty} e^{-y^2}\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)}\,dx\,dy\,.$$

Переход к полярным координатам дает

$$\begin{aligned} C^2 &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\,r\,dr\,d\theta = \\ &= \frac{1}{2}\int_0^{2\pi} d\theta \int_0^{\infty} e^{-u}\,du = \qquad\qquad (u = r^2)\\ &= \frac{1}{2}\int_0^{2\pi} d\theta = \pi\,. \end{aligned}$$

Следовательно, $C = \sqrt{\pi}$. Тот факт, что $x^2 + y^2 = r^2$ есть уравнение окружности длиной $2\pi r$, в какой-то мере объясняет, откуда здесь берется π.

Еще один способ вычисления C состоит в замене x на $\sqrt{t}$ и dx на $\frac{1}{2}t^{-1/2}\,dt$:

$$C = \int_{-\infty}^{+\infty} e^{-x^2}\,dx = 2\int_0^{\infty} e^{-x^2}\,dx = \int_0^{\infty} t^{-1/2}e^{-t}\,dt\,.$$

Этот интеграл равен $\Gamma\left(\frac{1}{2}\right)$, поскольку в соответствии с (5.84) $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}e^{-t}\,dt$. Таким образом, мы показали, что $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Итак, наша окончательная формула такова:

$$\Theta_n = \sum_k e^{-k^2/n} = \sqrt{\pi n} + O(n^{-M}) \qquad \text{для любого фиксированного } M. \quad (9.93)$$

Константа в O зависит от M; вот почему мы говорим о том, что M “фиксировано”.

Например, при $n=2$ бесконечная сумма Θ_2 равна 2.506628288; это уже отличное приближение к $\sqrt{2\pi} \approx 2.506628275$, несмотря на малость n. Значение же Θ_{100} совпадает с $10\sqrt{\pi}$ до 427-й десятичной цифры! В упр. 59 при помощи более совершенных методов получается быстро сходящийся ряд для Θ_n; оказывается, что

$$\Theta_n/\sqrt{\pi n} \;=\; 1 + 2e^{-n\pi^2} + O(e^{-4n\pi^2})\,. \tag{9.94}$$

Сумма 5: завершающий удар

Теперь займемся последней суммой, которая позволит нам найти значение константы Стирлинга σ. Эта последняя сумма иллюстрирует также многие другие методы из настоящей главы (да и из всей книги), так что это вполне уместный пример в завершение нашего исследования конкретной математики.

Последняя задача выглядит абсурдно простой: мы попытаемся найти асимптотическое значение суммы

$$A_n \;=\; \sum_k \binom{2n}{k}$$

с помощью формулы суммирования Эйлера.

Здесь мы вновь сталкиваемся с ситуацией, когда заранее известен ответ (так ли?); однако всегда интересно применить новые методы в старых задачах, чтобы сравнить результаты и, может быть, открыть что-нибудь новое.

Так что мы ДУМАЕМ ПО-БОЛЬШОМУ и обнаруживаем, что основной вклад в A_n вносят средние члены, вблизи $k=n$. Почти всегда удобно выбрать такие обозначения, чтобы основной вклад в сумму происходил вблизи $k=0$, потому что в таком случае можно воспользоваться приемом замены хвоста, чтобы избавиться от членов с большими $|k|$. Таким образом, заменим k на $n+k$:

$$A_n \;=\; \sum_k \binom{2n}{n+k} \;=\; \sum_k \frac{(2n)!}{(n+k)!\,(n-k)!}\,.$$

Дела идут неплохо, поскольку мы знаем, как аппроксимировать $(n \pm k)!$ при больших n и малых k.

Теперь мы хотим применить трехходовую процедуру, подразумеваемую приемом замены хвоста. А именно, мы хотим записать

$$\frac{(2n)!}{(n{+}k)!\,(n{-}k)!} \;=\; a_k(n) \;=\; b_k(n){+}O\bigl(c_k(n)\bigr) \qquad \text{для } k \in D_n$$

с тем, чтобы получить оценку

$$A_n = \sum_k b_k(n) + O\Bigl(\sum_{k\notin D_n} a_k(n)\Bigr) + \\ + O\Bigl(\sum_{k\notin D_n} b_k(n)\Bigr) + \sum_{k\in D_n} O\bigl(c_k(n)\bigr).$$

Таким образом, попробуем оценить $\binom{2n}{n+k}$ в области малых $|k|$. Можно было бы использовать аппроксимацию Стирлинга, как она представлена в табл. 545, но проще работать с ее логарифмическим эквивалентом в виде (9.91):

$$\begin{aligned}\ln a_k(n) &= \ln(2n)! - \ln(n+k)! - \ln(n-k)! = \\ &= 2n\ln 2n - 2n + \tfrac12\ln 2n + \sigma + O(n^{-1}) - \\ &\quad -(n+k)\ln(n+k) + n + k - \tfrac12\ln(n+k) - \sigma + O\bigl((n+k)^{-1}\bigr) - \\ &\quad -(n-k)\ln(n-k) + n - k - \tfrac12\ln(n-k) - \sigma + O\bigl((n-k)^{-1}\bigr).\end{aligned} \tag{9.95}$$

Хотелось бы преобразовать это в простую и красивую O-оценку.

Я не из тех, кто доминирует...

Метод замены хвоста позволяет использовать оценки, справедливые только для k из "доминирующего" множества D_n. Но как нам определить D_n? Мы должны сделать D_n достаточно малым, чтобы можно было получить хорошую оценку; так, например, не стоит допускать приближения k к n, в этом случае член $O\bigl((n-k)^{-1}\bigr)$ в (9.95) "взрывается". С другой стороны, D_n должно быть достаточно большим, чтобы остаточные члены (члены с $k \notin D_n$) были пренебрежимо малы в сравнении со всей суммой. Обычно для определения подходящего множества D_n требуется действовать методом проб и ошибок; в данной задаче вычисления, которые мы вскоре проделаем, покажут, что разумно определить множества следующим образом:

$$k \in D_n \iff |k| \leqslant n^{1/2+\epsilon}. \tag{9.96}$$

Здесь ϵ — небольшая положительная константа, значение которой будет выбрано позже, после того как мы получше изучим данную местность. (Наши O-оценки будут зависеть от ϵ.) Уравнение (9.95) можно теперь сократить до

$$\begin{aligned}\ln a_k(n) &= (2n+\tfrac12)\ln 2 - \sigma - \tfrac12\ln n + O(n^{-1}) - \\ &\quad -(n+k+\tfrac12)\ln(1+k/n) - (n-k+\tfrac12)\ln(1-k/n).\end{aligned} \tag{9.97}$$

(Мы вынесли основные части логарифмов, записав

$$\ln(n\pm k) = \ln n + \ln(1\pm k/n),$$

в результате чего многие члены, содержащие $\ln n$, сократились.)

Теперь нужно разложить члены $\ln(1 \pm k/n)$ асимптотически, так, чтобы погрешность стремилась к нулю при $n \to \infty$. Мы умножаем $\ln(1 \pm k/n)$ на $(n \pm k + \frac{1}{2})$, так что мы должны разложить логарифм до членов $o(n^{-1})$, с учетом предположения $|k| \leqslant n^{1/2+\epsilon}$:

$$\ln\left(1 \pm \frac{k}{n}\right) \;=\; \pm\frac{k}{n} - \frac{k^2}{2n^2} + O(n^{-3/2+3\epsilon})\,.$$

Умножение на $n \pm k + \frac{1}{2}$ дает

$$\pm k - \frac{k^2}{2n} + \frac{k^2}{n} + O(n^{-1/2+3\epsilon})$$

плюс другие члены, которые поглощаются слагаемым $O(n^{-1/2+3\epsilon})$. Таким образом, (9.97) превращается в

$$\ln a_k(n) \;=\; (2n + \frac{1}{2})\ln 2 - \sigma - \frac{1}{2}\ln n - k^2/n + O(n^{-1/2+3\epsilon})\,.$$

Взяв экспоненту, получим

$$a_k(n) \;=\; \frac{2^{2n+1/2}}{e^\sigma\sqrt{n}}\, e^{-k^2/n}\big(1 + O(n^{-1/2+3\epsilon})\big)\,. \qquad (9.98)$$

Это и есть наша аппроксимация с

$$b_k(n) \;=\; \frac{2^{2n+1/2}}{e^\sigma\sqrt{n}}\, e^{-k^2/n}\,, \qquad c_k(n) \;=\; 2^{2n}\, n^{-1+3\epsilon}\, e^{-k^2/n}\,.$$

Заметьте, что k входит в $b_k(n)$ и $c_k(n)$ очень простым образом. Это явная удача, потому что мы будем выполнять суммирование по k.

Прием замены хвоста говорит нам, что $\sum_k a_k(n)$ приближенно равна $\sum_k b_k(n)$, если грамотно выполнить оценку. Давайте вычислим

$$\sum_k b_k(n) \;=\; \frac{2^{2n+1/2}}{e^\sigma\sqrt{n}} \sum_k e^{-k^2/n} \;=$$
$$=\; \frac{2^{2n+1/2}}{e^\sigma\sqrt{n}}\Theta_n \;=\; \frac{2^{2n}\sqrt{2\pi}}{e^\sigma}\big(1 + O(n^{-M})\big)\,.$$

(И снова удача: можно использовать сумму Θ_n из предыдущего примера.) Это радует, поскольку, как нам известно, исходная сумма в действительности равна

Удивительное совпадение!

$$A_n \;=\; \sum_k \binom{2n}{k} \;=\; (1+1)^{2n} \;=\; 2^{2n}\,.$$

Таким образом, похоже, что $e^\sigma = \sqrt{2\pi}$, как и было объявлено.

Но не торопитесь радоваться. Надо еще доказать, что наши оценки достаточно точны. Посмотрим сначала на погрешность, вносимую $c_k(n)$:

Я очень устал, продираясь через длинные, тяжелые книги. И ни в одной я не встретил ни слова поощрения и благодарности от автора. А как было бы приятно прочесть: "Благодарим за время, затраченное на чтение этой книги, и надеемся, что оно было затрачено не зря", вместо того чтобы после длинного сухого доказательства обнаружить холодную и твердую картонную обложку. Вы согласны?

$$\Sigma_c(n) = \sum_{|k|\leqslant n^{1/2+\epsilon}} 2^{2n} n^{-1+3\epsilon} e^{-k^2/n} \leqslant 2^{2n} n^{-1+3\epsilon} \Theta_n = \\ = O(2^{2n} n^{-\frac{1}{2}+3\epsilon}).$$

Хорошо; эта величина асимптотически меньше предыдущей суммы, если $3\epsilon < \frac{1}{2}$.

Теперь следует проверить хвосты. Имеем

$$\sum_{k>n^{1/2+\epsilon}} e^{-k^2/n} < \exp\bigl(-\lfloor n^{1/2+\epsilon}\rfloor^2/n\bigr)(1+e^{-1/n}+e^{-2/n}+\cdots) = \\ = O(e^{-n^{2\epsilon}}) \cdot O(n),$$

а это есть $O(n^{-M})$ для любого M; так что $\sum_{k\notin D_n} b_k(n)$ асимптотически пренебрежимо мала. (Мы выбрали границу $n^{1/2+\epsilon}$ как раз так, чтобы $e^{-k^2/n}$ было экспоненциально малым вне D_n. Вполне подошли бы и другие выражения, наподобие $n^{1/2}\log n$, и оценки были бы тогда несколько точнее, зато формулы оказались бы сложнее. Нам совсем не обязательно находить наиболее точную оценку, поскольку главная наша цель — найти значение константы σ.) Аналогично другой хвост,

$$\sum_{k>n^{1/2+\epsilon}} \binom{2n}{n+k},$$

ограничен $2n$-кратным максимальным слагаемым, получаемым в разделительной точке $k \approx n^{1/2+\epsilon}$. Это слагаемое, как известно, приближенно равно $b_k(n)$, что, в свою очередь, экспоненциально мало в сравнении с A_n; экспоненциально малый множитель перевешивает множитель $2n$.

Итак, мы успешно применили метод замены хвоста и доказали оценку

$$2^{2n} = \sum_k \binom{2n}{k} = \frac{\sqrt{2\pi}}{e^\sigma} 2^{2n} + O(2^{2n} n^{-\frac{1}{2}+3\epsilon}), \\ \text{если } 0 < \epsilon < \tfrac{1}{6}. \qquad (9.99)$$

Благодарим за время, затраченное на чтение этой книги, и надеемся, что оно было затрачено не зря.
— Авторы

Можно выбрать $\epsilon = \frac{1}{8}$ и сделать вывод, что

$$\sigma = \tfrac{1}{2}\ln 2\pi.$$

QED

Упражнения

Разминка

1 Докажите или опровергните: если $f_1(n) \prec g_1(n)$ и $f_2(n) \prec g_2(n)$, то $f_1(n) + f_2(n) \prec g_1(n) + g_2(n)$.

2 Какая из функций растет быстрее:

а $n^{(\ln n)}$ или $(\ln n)^n$,

б $n^{(\ln \ln \ln n)}$ или $(\ln n)!$,

в $(n!)!$ или $((n-1)!)!\,(n-1)!^{n!}$,

г $F^2_{\lceil H_n \rceil}$ или H_{F_n}?

3 Что неверно в следующих рассуждениях: "Поскольку $n = O(n)$, $2n = O(n)$ и т.д., то $\sum_{k=1}^{n} kn = \sum_{k=1}^{n} O(n) = O(n^2)$"?

4 Приведите пример верного соотношения, содержащего символ 'O' в левой части, но не в правой. (Не используйте умножение на нуль; это слишком просто.) *Указание*: рассмотрите переход к пределу.

5 Докажите или опровергните: $O\bigl(f(n)+g(n)\bigr) = f(n)+O\bigl(g(n)\bigr)$, если $f(n)$ и $g(n)$ положительны для всех n. (Сравните с (9.27).)

6 Умножьте $\bigl(\ln n + \gamma + O(1/n)\bigr)$ на $\bigl(n + O(\sqrt{n}\,)\bigr)$ и выразите ответ в O-обозначениях.

7 Оцените $\sum_{k\geqslant 0} e^{-k/n}$ с абсолютной погрешностью $O(n^{-1})$.

Обязательные упражнения

8 Приведите пример функций $f(n)$ и $g(n)$, таких, что ни одно из отношений $f(n) \prec g(n)$, $f(n) \succ g(n)$ и $f(n) \asymp g(n)$ не выполняется, несмотря на то что обе эти функции монотонно стремятся к ∞.

9 Строго докажите (9.22), показав, что левая часть является подмножеством правой части в соответствии с теоретико-множественным определением O.

10 Докажите или опровергните: $\cos O(x) = 1 + O(x^2)$ для всех действительных x.

11 Докажите или опровергните: $O(x+y)^2 = O(x^2) + O(y^2)$.

12 Докажите, что

$$1 + \frac{2}{n} + O(n^{-2}) \;=\; \Bigl(1 + \frac{2}{n}\Bigr)\bigl(1 + O(n^{-2})\bigr)$$

при $n \to \infty$.

13 Вычислите $(n+2+O(n^{-1}))^n$ с относительной погрешностью $O(n^{-1})$.

14 Докажите, что $(n+\alpha)^{n+\beta} = n^{n+\beta}e^{\alpha}\bigl(1+\alpha(\beta-\frac{1}{2}\alpha)n^{-1}+O(n^{-2})\bigr)$.

15 Укажите асимптотическую формулу для "среднего" триномиального коэффициента $\binom{3n}{n,n,n}$, относительная погрешность которой равна $O(n^{-3})$.

16 Покажите, что если $B(1-x) = -B(x) \geqslant 0$ для $0 < x < \frac{1}{2}$, то

$$\int_a^b B(\{x\})\,f(x)\,dx \;\geqslant\; 0$$

при дополнительном условии $f'(x) \geqslant 0$ для $a \leqslant x \leqslant b$.

17 Используя производящие функции, покажите, что $B_m(\frac{1}{2}) = (2^{1-m}-1)B_m$ для всех $m \geqslant 0$.

18 Найдите значение суммы $\sum_k \binom{2n}{k}^{\alpha}$ при $\alpha > 0$ с относительной погрешностью $O(n^{-1/4})$.

Домашние задания

19 Воспользуйтесь компьютером для сравнения левой и правой частей аппроксимаций из табл. 545, положив $n = 10$, $z = \alpha = 0.1$ и $O(f(n)) = O(f(z)) = 0$.

20 Докажите или опровергните следующие оценки при $n \to \infty$:

а $$O\left(\left(\frac{n^2}{\log\log n}\right)^{1/2}\right) = O(\lfloor\sqrt{n}\rfloor^2)\,;$$

б $$e^{(1+O(1/n))^2} = e + O(1/n)\,;$$

в $$n! = O\Bigl(\bigl((1-1/n)^n n\bigr)^n\Bigr)\,.$$

21 Уравнение (9.48) дает n-е простое число с относительной погрешностью $O(\log n)^{-2}$. Использовав в (9.46) еще один член (9.31), улучшите относительную погрешность до $O(\log n)^{-3}$.

22 Улучшите (9.54) до $O(n^{-3})$.

23 Сделайте еще один шаг в улучшении (9.62), получив абсолютную погрешность $O(n^{-3})$. *Указание*: пусть $g_n = c/(n+1)(n+2) + h_n$; какому рекуррентному соотношению удовлетворяет h_n?

24 Предположим, что $a_n = O(f(n))$ и $b_n = O(f(n))$. Докажите или опровергните то, что свертка $\sum_{k=0}^{n} a_k b_{n-k}$ также есть $O(f(n))$, в следующих случаях:

а $f(n) = n^{-\alpha}$, $\alpha > 1$;

б $f(n) = \alpha^{-n}$, $\alpha > 1$.

25 Докажите формулы (9.1) и (9.2).

26 Уравнение (9.91) показывает, как вычислить $\ln 10!$ с абсолютной погрешностью $< \frac{1}{126000000}$. Следовательно, взяв экспоненту, мы получим $10!$ с относительной погрешностью, меньшей, чем $e^{1/126000000} - 1 < 10^{-8}$. (Фактически эта аппроксимация дает 3628799.9714.) Округлив полученный результат до ближайшего целого (зная, что $10!$ — целое число), мы получим точный ответ.

Всегда ли можно вычислить $n!$ подобным методом, взяв достаточное количество членов приближения Стирлинга? Оцените значение m, дающее наилучшее приближение к $\ln n!$ для фиксированного (большого) целого n. Сравните абсолютную погрешность этого приближения с самим значением $n!$.

27 Воспользуйтесь формулой суммирования Эйлера для поиска асимптотического значения $H_n^{(-\alpha)} = \sum_{k=1}^{n} k^{\alpha}$, где α — любое фиксированное действительное число. (Ваш ответ может включать константу, значение которой в аналитическом виде вам неизвестно.)

28 В упр. 5.13 определена функция "гиперфакториал" $Q_n = 1^1 2^2 \ldots n^n$. Найдите асимптотическое значение Q_n с относительной погрешностью $O(n^{-1})$. (Ваш ответ может включать константу, значение которой в аналитическом виде вам неизвестно.)

29 Оцените функцию $1^{1/1} 2^{1/2} \ldots n^{1/n}$, как в предыдущем упражнении.

30 Найдите асимптотическое значение $\sum_{k\geqslant 0} k^l e^{-k^2/n}$ с абсолютной погрешностью $O(n^{-3})$, если l — фиксированное неотрицательное целое число.

31 Вычислите $\sum_{k\geqslant 0} 1/(c^k + c^m)$ с абсолютной погрешностью $O(c^{-3m})$, если $c > 1$, а m — положительное целое число.

Контрольные работы

32 Вычислите $e^{H_n + H_n^{(2)}}$ с абсолютной погрешностью $O(n^{-1})$.

33 Вычислите $\sum_{k\geqslant 0} \binom{n}{k}/n^{\overline{k}}$ с абсолютной погрешностью $O(n^{-3})$.

34 Определите значения коэффициентов от A до F, такие, что $(1 + 1/n)^{nH_n}$ равно

$$An + B(\ln n)^2 + C\ln n + D + \frac{E(\ln n)^2}{n} + \frac{F\ln n}{n} + O(n^{-1}).$$

35 Вычислите $\sum_{k=1}^{n} 1/kH_k$ с абсолютной погрешностью $O(1)$.

36 Вычислите сумму $S_n = \sum_{k=1}^{n} 1/(n^2+k^2)$ с абсолютной погрешностью $O(n^{-5})$.

37 Вычислите $\sum_{k=1}^{n}(n \bmod k)$ с абсолютной погрешностью $O(n \log n)$.

38 Вычислите $\sum_{k\geqslant 0} k^k \binom{n}{k}$ с относительной погрешностью $O(n^{-1})$.

39 Вычислите $\sum_{0\leqslant k<n} \ln(n-k)(\ln n)^k/k!$ с абсолютной погрешностью $O(n^{-1})$. *Указание*: покажите, что члены при $k \geqslant 10\ln n$ пренебрежимо малы.

40 Пусть m — (фиксированное) положительное целое число. Вычислите $\sum_{k=1}^{n}(-1)^k H_k^m$ с абсолютной погрешностью $O(1)$.

41 Вычислите "факториал Фибоначчи" $\prod_{k=1}^{n} F_k$ с относительной погрешностью $O(n^{-1})$ или меньшей. Ваш ответ может включать константу, значение которой в аналитическом виде вам неизвестно.

42 Пусть α — константа в диапазоне $0<\alpha<\frac{1}{2}$. Из предыдущих глав мы знаем, что общего выражения в аналитическом виде для суммы $\sum_{k\leqslant \alpha n}\binom{n}{k}$ не существует. Покажите, однако, что имеется асимптотическая формула

$$\sum_{k\leqslant \alpha n}\binom{n}{k} = 2^{nH(\alpha)-\frac{1}{2}\lg n+O(1)},$$

где $H(\alpha)=\alpha\lg\frac{1}{\alpha}+(1-\alpha)\lg(\frac{1}{1-\alpha})$. *Указание*: покажите, что $\binom{n}{k-1} < \frac{\alpha}{1-\alpha}\binom{n}{k}$ при $0<k\leqslant \alpha n$.

43 Покажите, что C_n, количество способов размена n копеек (рассмотренное в главе 7) асимптотически равно $cn^4+O(n^3)$ для некоторой константы c. Чему равна эта константа?

44 Докажите, что

$$x^{\underline{1/2}} = x^{1/2}\left[{1/2 \atop 1/2}\right] - x^{-1/2}\left[{1/2 \atop -1/2}\right] + x^{-3/2}\left[{1/2 \atop -3/2}\right] + O(x^{-5/2})$$

при $x\to\infty$. (Вспомните определение $x^{\underline{1/2}} = x!/(x-\frac{1}{2})!$ из (5.88) и определение обобщенных чисел Стирлинга из табл. 334.)

45 Пусть α — иррациональное число между 0 и 1. В главе 3 рассматривалась величина $D(\alpha,n)$, измеряющая максимальное отклонение множества $\{k\alpha\}$ для $0\leqslant k<n$ от равномерного

распределения. В (3.31) было доказано рекуррентное соотношение

$$D(\alpha, n) \leqslant D(\{\alpha^{-1}\}, \lfloor \alpha n \rfloor) + \alpha^{-1} + 2;$$

кроме того, имеются очевидные границы

$$0 \leqslant D(\alpha, n) \leqslant n\,.$$

Докажите, что $\lim_{n\to\infty} D(\alpha, n)/n = 0$. *Указание*: в главе 6 рассматриваются непрерывные дроби.

46 Покажите, что числа Белла $\varpi_n = e^{-1}\sum_{k\geqslant 0} k^n/k!$ из упр. 7.15 асимптотически равны

$$m(n)^n e^{m(n)-n-1/2}/\sqrt{\ln n}\,,$$

где $m(n)\ln m(n) = n - \frac{1}{2}$. Оцените относительную погрешность этой аппроксимации.

47 Пусть m — целое число $\geqslant 2$. Проанализируйте суммы

$$\sum_{k=1}^{n} \lfloor \log_m k \rfloor \qquad \text{и} \qquad \sum_{k=1}^{n} \lceil \log_m k \rceil\,.$$

Какая из них асимптотически ближе к $\log_m n!$?

48 Рассмотрим таблицу гармонических чисел H_k для $1 \leqslant k \leqslant n$ в десятичной записи. В ней k-й элемент $\widehat{H}_k$ корректно округлен до d_k значащих цифр, где d_k выбрано ровно таким, чтобы можно было отличить этот элемент от H_{k-1} и H_{k+1}. Например, вот как выглядят пять строк этой таблицы в месте, где H_k преодолевает рубеж 10:

k	H_k	$\widehat{H}_k$	d_k
12364	9.99980041−	9.9998	5
12365	9.99988128+	9.9999	5
12366	9.99996215−	9.99996	6
12367	10.00004301−	10.0000	6
12368	10.00012386+	10.0001	6

Оцените общее количеств цифр в таблице $\sum_{k=1}^{n} d_k$ с абсолютной погрешностью $O(n)$.

49 В главе 6 мы рассмотрели историю червяка на резинке, который достигнет ее конца через n минут, где $H_{n-1} < 100 \leqslant H_n$.

Докажите, что если n — положительное целое число, такое, что $H_{n-1} \leqslant \alpha \leqslant H_n$, то

$$\lfloor e^{\alpha-\gamma} \rfloor \leqslant n \leqslant \lceil e^{\alpha-\gamma} \rceil .$$

50 Рисковые капиталисты из Кремниевой долины получили предложение, дающее шанс на экспоненциально большую прибыль по их инвестициям: консорциум ГКП обещает в случае вклада в размере n миллионов долларов, где $n \geqslant 2$, выплатить через год сумму в N миллионов долларов, где $N = 10^n$. Конечно, есть определенный риск; на самом деле ГКП платит k миллионов долларов с вероятностью $1/(k^2 H_N^{(2)})$ для целых k в диапазоне $1 \leqslant k \leqslant N$. (Все платежи осуществляются в мегадолларах, т.е. в точных кратных миллиону долларов; выплата определяется истинно случайным процессом.) Обратите внимание, что вкладчик всегда получает назад по крайней мере один миллион долларов.

a Чему равен асимптотический ожидаемый результат через год, если было вложено n миллионов долларов? (Другими словами, какова средняя сумма выплаты?) Ваш ответ должен быть точен в пределах абсолютной погрешности $O(10^{-n})$ долларов.

Как-то раз я заработал $O(10^{-n})$ долларов.

б С какой асимптотической вероятностью вы будете иметь прибыль, вложив n миллионов? (Другими словами, какова вероятность, что вы получите назад больше, чем вложили?) Здесь допускается абсолютная погрешность ответа $O(n^{-3})$.

Дополнительные задачи

51 Докажите или опровергните: $\int_n^\infty O(x^{-2})\,dx = O(n^{-1})$ при $n \to \infty$.

52 Покажите, что существует степенной ряд $A(z) = \sum_{k\geqslant 0} a_n z^n$, сходящийся при всех комплексных z, такой, что

$$A(n) \succ \left. n^{n^{\cdot^{\cdot^{\cdot^{n}}}}} \right\} n .$$

53 Докажите, что если $f(x)$ — функция, производные которой удовлетворяют условиям

$$f'(x) \leqslant 0, \quad -f''(x) \leqslant 0, \quad f'''(x) \leqslant 0, \quad \ldots,$$
$$(-1)^m f^{(m+1)}(x) \leqslant 0$$

для всех $x \geqslant 0$, то

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \cdots + \frac{f^{(m-1)}(0)}{(m-1)!}x^{m-1} + O(x^m)$$

для $x \geqslant 0$.

В частности, в случае $f(x) = -\ln(1+x)$ получается доказательство (9.64) для всех $k, n > 0$.

54 Пусть $f(x)$ — положительная, дифференцируемая функция, такая, что $xf'(x) \prec f(x)$ при $x \to \infty$. Докажите, что

$$\sum_{k \geqslant n} \frac{f(k)}{k^{1+\alpha}} = O\left(\frac{f(n)}{n^\alpha}\right), \qquad \text{если } \alpha > 0.$$

Указание: рассмотрите величину $f(k-\frac{1}{2})/(k-\frac{1}{2})^\alpha - f(k+\frac{1}{2})/(k+\frac{1}{2})^\alpha$.

55 Улучшите (9.99), уменьшив относительную погрешность до $O(n^{-3/2+5\epsilon})$.

56 Величина $Q(n) = 1 + \frac{n-1}{n} + \frac{n-1}{n}\frac{n-2}{n} + \cdots = \sum_{k \geqslant 1} n^{\underline{k}}/n^k$ встречается при анализе многих алгоритмов. Найдите ее асимптотическое значение с абсолютной погрешностью $o(1)$.

57 В (9.54) получена асимптотическая формула для суммы Голомба $\sum_{k \geqslant 1} 1/k\lfloor 1 + \log_n k\rfloor^2$. Найдите асимптотическую формулу для аналогичной суммы без округления для целых $\sum_{k \geqslant 1} 1/k(1+\log_n k)^2$. *Указание*: рассмотрите интеграл $\int_0^\infty ue^{-u}k^{-tu}\,du = 1/(1+t\ln k)^2$.

58 Докажите, что

$$B_m(\{x\}) = -2\frac{m!}{(2\pi)^m}\sum_{k \geqslant 1}\frac{\cos(2\pi kx - \frac{1}{2}\pi m)}{k^m} \qquad \text{для } m \geqslant 2.$$

Для этого вычислите с помощью теории вычетов интеграл

$$\frac{1}{2\pi i}\oint \frac{2\pi i\, e^{2\pi i z\theta}}{e^{2\pi i z} - 1}\frac{dz}{z^m}$$

по квадратному контуру $z = x + iy$, где $\max(|x|, |y|) = M + \frac{1}{2}$, и устремите целое M к ∞.

59 Пусть $\Theta_n(t) = \sum_k e^{-(k+t)^2/n}$ — периодическая функция от t. Покажите, что разложение $\Theta_n(t)$ в ряд Фурье выглядит следующим образом:

$$\Theta_n(t) = \sqrt{\pi n}\big(1 + 2e^{-\pi^2 n}(\cos 2\pi t) + 2e^{-4\pi^2 n}(\cos 4\pi t) + \\ + 2e^{-9\pi^2 n}(\cos 6\pi t) + \cdots\big).$$

(Эта формула дает быстро сходящийся ряд для суммы $\Theta_n = \Theta_n(0)$ из уравнения (9.93).)

60 Объясните, почему у всех коэффициентов в асимптотическом разложении

$$\binom{2n}{n} = \frac{4^n}{\sqrt{\pi n}}\left(1-\frac{1}{8n}+\frac{1}{128n^2}+\frac{5}{1024n^3}-\frac{21}{32768n^4}+O(n^{-5})\right)$$

знаменатели являются степенями 2.

61 В упр. 45 доказывается, что отклонение от равномерного определения $D(\alpha, n)$ есть $o(n)$ для всех иррациональных чисел α. Предъявите иррациональное α, такое, что $D(\alpha, n)$ *не* является $O(n^{1-\epsilon})$ ни для какого $\epsilon > 0$.

62 Для данного n пусть $\left\{{n \atop m(n)}\right\} = \max_k \left\{{n \atop k}\right\}$ — максимальный элемент строки n треугольника Стирлинга (для количества подмножеств). Покажите, что для всех достаточно больших n выполняется одно из равенств $m(n) = \lfloor \overline{m}(n) \rfloor$ или $m(n) = \lceil \overline{m}(n) \rceil$, где

$$\overline{m}(n)(\overline{m}(n)+2)\ln(\overline{m}(n)+2) = n(\overline{m}(n)+1)\,.$$

Указание: это трудно.

63 Докажите, что самоописывающая последовательность Голомба из упр. 2.36 удовлетворяет соотношению

$$f(n) = \phi^{2-\phi} n^{\phi-1} + O(n^{\phi-1}/\log n).$$

64 Найдите доказательство тождества

$$\sum_{n\geqslant 1} \frac{\cos 2n\pi x}{n^2} = \pi^2\left(x^2 - x + \tfrac{1}{6}\right) \qquad \text{для } 0 \leqslant x \leqslant 1,$$

которое использует только "Эйлерову" математику (т.е. не более того, что было известно во времена Эйлера).

65 Чему равны коэффициенты асимптотического ряда

$$1+\frac{1}{n-1}+\frac{1}{(n-1)(n-2)}+\cdots+\frac{1}{(n-1)!} = a_0+\frac{a_1}{n}+\frac{a_2}{n^2}+\cdots\ ?$$

Исследовательские проблемы

66 Найдите "комбинаторное" доказательство аппроксимации Стирлинга. (Обратите внимание на то, что n^n есть число отображений множества $\{1, 2, \ldots, n\}$ в себя, тогда как $n!$ — количество отображений множества $\{1, 2, \ldots, n\}$ на себя.)

67 Рассмотрим решетку точек размером $n \times n$, $n \geqslant 3$, в которой каждая точка имеет четырех соседей. (На краях мы “перескакиваем” на противоположную сторону по модулю n.) Пусть χ_n означает количество способов раскрасить эти точки в три цвета — красный, белый и синий — так, чтобы никакие две соседние точки не были окрашены в один цвет. (Так, $\chi_3 = 12$.) Докажите, что

$$\chi_n \sim \left(\tfrac{4}{3}\right)^{3n^2/2} e^{-\pi/6}.$$

68 Обозначим через Q_n наименьшее целое m, такое, что $H_m > n$. Найдите наименьшее целое n, для которого $Q_n \neq \lfloor e^{n-\gamma} + \frac{1}{2} \rfloor$, или докажите, что такого n не существует.

Ура, это коне-е-е-е-ец!!!

A

Ответы к упражнениям

КАЖДОЕ УПРАЖНЕНИЕ сопровождается (по крайней мере кратким) ответом, причем некоторые ответы выходят за рамки вопроса. Читатели извлекут куда больше пользы из упражнений, если попытаются найти собственные ответы, ПРЕЖДЕ ЧЕМ ЗАГЛЯНУТ в данное приложение.

Авторам будет интересно узнать о любых (даже частичных) решениях исследовательских задач, как и о любых более простых (или более правильных) способах решения других задач.

Первому нашедшему любую ошибку в оригинальном издании книги выплачивается премия в 2.56 доллара.

1.1 Доказательство вполне корректно, если не считать случая $n = 2$. Если все множества из двух лошадей содержат лошадей одной и той же масти, то данное утверждение справедливо для любого количества лошадей.

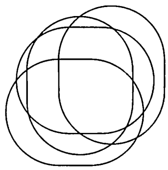
1.2 Если обозначить количество переносов как X_n, то $X_0 = 0$ и $X_n = X_{n-1} + 1 + X_{n-1} + 1 + X_{n-1}$ для $n > 0$. Отсюда следует (например, путем прибавления 1 к обеим частям уравнения), что $X_n = 3^n - 1$. (После $\frac{1}{2}X_n$ перемещений вся башня оказывается на среднем шпиле — полпути пройдено!)

1.3 Всего имеется 3^n разных размещений, поскольку каждый диск может находиться на любом из шпилей. Мы должны повстречаться с каждым из них, поскольку кратчайшее решение требует $3^n - 1$ переносов. (Эта конструкция эквивалентна "тернарному коду Грея", который проходит все числа от $(0\ldots0)_3$ до $(2\ldots2)_3$, изменяя каждый раз только одну цифру.)

1.4 Нет. Если самый большой диск не требуется перекладывать, то достаточно $2^{n-1} - 1$ переносов (доказывается по индукции); в противном случае будет достаточно $(2^{n-1}-1)+1+(2^{n-1}-1)$ переносов (также доказывается по индукции).

1.5 Нет. Различные окружности могут пересечься не более чем в двух точках, так что четвертая окружность может увеличить количество областей не более чем до 14. Однако это вполне возможно, если воспользоваться овалами:

Главное — в количестве точек пересечения; выпуклость играет роль отвлекающего момента.



Венн (Venn) [359] утверждал, что не существует способа проиллюстрировать случай пяти множеств при помощи эллипсов, но Грюнбаум (Grünbaum) [167] сумел найти такое построение.

1.6 Если n-я прямая пересекает предыдущие в $k > 0$ различных точках, мы получим $k-1$ новых ограниченных областей (в предположении, что взаимно параллельных прямых среди уже имеющихся нет) и две новые бесконечные области. Следовательно, максимальное количество конечных областей равно

Предполагается, что $n > 0$.

$$(n-2)+(n-3)+\cdots = S_{n-2} = (n-1)(n-2)/2 = L_n - 2n.$$

1.7 Не доказано основание индукции; и в самом деле, $H(1) \neq 2$.

1.8 $Q_2 = (1+\beta)/\alpha$, $Q_3 = (1+\alpha+\beta)/\alpha\beta$, $Q_4 = (1+\alpha)/\beta$, $Q_5 = \alpha$, $Q_6 = \beta$. Данная последовательность периодична!

1.9 (а) Утверждение $P(n-1)$ получается из неравенства

$$x_1 \ldots x_{n-1}\left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right) \leqslant \left(\frac{x_1+\cdots+x_{n-1}}{n-1}\right)^n .$$

(б) $x_1 \ldots x_n x_{n+1} \ldots x_{2n} \leqslant \Bigl(\bigl((x_1+\cdots+x_n)/n\bigr)\bigl((x_{n+1}+\cdots+x_{2n})/n\bigr)\Bigr)^n$ в силу $P(n)$; произведение внутри $\leqslant \bigl((x_1+\cdots+x_{2n})/2n\bigr)^2$ в силу $P(2)$. (в) Например, $P(5)$ следует из $P(6)$, которое, в свою очередь, следует из $P(3)$, которое следует из $P(4)$, которое следует из $P(2)$.

1.10 Сначала покажите, что $R_n = R_{n-1} + 1 + Q_{n-1} + 1 + R_{n-1}$ при $n > 0$. Кстати, методы из главы 7 покажут нам, что

$$Q_n = \bigl((1+\sqrt{3}\,)^{n+1} - (1-\sqrt{3}\,)^{n+1}\bigr)/(2\sqrt{3}) - 1.$$

1.11 (а) Лучшее, что мы можем сделать, — это переместить сначала двойную $(n-1)$-башню, затем — (с изменением их относительного порядка) два наибольших диска, после чего — опять двойную $(n-1)$-башню. Следовательно, $A_n = 2A_{n-1} + 2$ и $A_n =$

$2T_n = 2^{n+1} - 2$. При таком решении происходит перестановка местами двух наибольших дисков; относительный порядок остальных $2n - 2$ дисков остается тем же, что и ранее.

(б) Обозначим минимальное количество переносов как B_n. Тогда $B_1 = 3$, и может быть показано, что не существует стратегии лучшей, чем $B_n = A_{n-1} + 2 + A_{n-1} + 2 + B_{n-1}$ при $n > 1$. Следовательно, $B_n = 2^{n+2} - 5$ для всех $n > 0$. Любопытно, что это просто $2A_n - 1$, и мы также имеем $B_n = A_{n-1} + 1 + A_{n-1} + 1 + A_{n-1} + 1 + A_{n-1}$.

1.12 Если все $m_k > 0$, то $A(m_1, \ldots, m_n) = 2A(m_1, \ldots, m_{n-1}) + m_n$. Это уравнение "обобщенного Иосифова" типа с решением

$$(m_1 \ldots m_n)_2 = 2^{n-1} m_1 + \cdots + 2m_{n-1} + m_n.$$

Кстати, соответствующее обобщение упр. 11, б удовлетворяет рекуррентному соотношению

$$B(m_1, \ldots, m_n) = \begin{cases} A(m_1, \ldots, m_n), & \text{если } m_n = 1, \\ 2m_n - 1, & \text{если } n = 1, \\ 2A(m_1, \ldots, m_{n-1}) + 2m_n + & \text{если } n > 1 \\ \quad + B(m_1, \ldots, m_{n-1}), & \text{и } m_n > 1. \end{cases}$$

1.13 Поскольку n прямых определяют L_n областей, можно заменить их очень узкими загзагами с отрезками, достаточно длинными для того, чтобы каждая пара зигзагов имела девять пересечений. Это показывает, что $ZZ_n = ZZ_{n-1} + 9n - 8$ при всех $n > 0$; следовательно,

$$ZZ_n = 9S_n - 8n + 1 = \frac{9}{2}n^2 - \frac{7}{2}n + 1.$$

1.14 Количество новых трехмерных областей, определяемых каждым новым разрезом, представляет собой количество двухмерных областей, образованных на новой плоскости линиями ее пересечений с прежними плоскостями. Следовательно, $P_n = P_{n-1} + L_{n-1}$, что дает нам $P_5 = 26$. (Шесть разрезов кубической головки сыра могут дать 27 сырных кубиков или до $P_6 = 42$ кусков неправильной формы.)

Кстати, решение этого рекуррентного соотношения приобретает красивый вид, если выразить его через биномиальные коэффициенты (см. главу 5):

Готов спорить, что я знаю ответ для четырех измерений!

$$X_n = \binom{n}{0} + \binom{n}{1};$$

$$L_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2};$$

$$P_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3}.$$

Здесь X_n — максимальное количество одномерных областей, на которые n точек делят прямую.

1.15 При $n > 1$ функция I удовлетворяет тому же рекуррентному соотношению, что и функция J, но значение $I(1)$ не определено. Поскольку $I(2) = 2$ и $I(3) = 1$, не существует значения $I(1) = \alpha$, такого, что оно позволило бы нам воспользоваться обобщенным методом; "исход" развертывания зависит от двух старших битов двоичного представления n.

Если $n = 2^m + 2^{m-1} + k$, где $0 \leqslant k < 2^{m+1} + 2^m - (2^m + 2^{m-1}) = 2^m + 2^{m-1}$, то решением является $I(n) = 2k+1$ при всех $n > 2$. Другой способ записи решения с использованием представления $n = 2^m + l$ имеет вид

$$I(n) = \begin{cases} J(n) + 2^{m-1}, & \text{если } 0 \leqslant l < 2^{m-1}; \\ J(n) - 2^m, & \text{если } 2^{m-1} \leqslant l < 2^m. \end{cases}$$

1.16 Пусть $g(n) = a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 + d(n)\gamma$. Из (1.18) нам известно, что $a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 = (\alpha\,\beta_{b_{m-1}}\,\beta_{b_{m-2}}\ldots\beta_{b_1}\,\beta_{b_0})_3$ при $n = (1\,b_{m-1}\ldots b_1\,b_0)_2$; это определяет $a(n)$, $b(n)$ и $c(n)$. Подставив $g(n) = n$ в рекуррентное соотношение, получим, что $a(n) + c(n) - d(n) = n$, так что нам известно все, что надо. (Подстановка $g(n) = 1$ дает дополнительное тождество $a(n) - 2b(n) - 2c(n) = 1$, которое можно использовать для определения $b(n)$ через более простые функции $a(n)$ и $a(n) + c(n)$.)

1.17 В общем случае $W_m \leqslant 2W_{m-k} + T_k$ при $0 \leqslant k \leqslant m$. (Это соотношение соответствует перемещению верхних $m-k$ дисков с последующим использованием только трех шпилей для перемещения нижних k дисков и завершающим перемещением верхних $m-k$ дисков.) Данное соотношение приводит к зависимости от единственного значения k, которое минимизирует правую часть общего неравенства при $m = n(n+1)/2$. (Однако сделать вывод о том, что выполняется равенство, нельзя — могут существовать многие другие приемлемые стратегии перемещения башни.) Если принять $Y_n = (W_{n(n+1)/2} - 1)/2^n$, находим, что $Y_n \leqslant Y_{n-1} + 1$; следовательно, $W_{n(n+1)/2} \leqslant 2^n(n-1) + 1$.

1.18 Достаточно показать, что обе линии, выходящие из точки $(n^{2j}, 0)$, пересекают обе линии, выходящие из точки $(n^{2k}, 0)$, и что все точки пересечения различны.

Линия, выходящая из точки $(x_j, 0)$ и проходящая через точку $(x_j - a_j, 1)$, пересекает линию, выходящую из точки $(x_k, 0)$ и проходящую через точку $(x_k - a_k, 1)$, в точке $(x_j - ta_j, t)$, где $t = (x_k - x_j)/(a_k - a_j)$. Пусть $x_j = n^{2j}$ и $a_j = n^j + (0 \text{ или } n^{-n})$. Тогда значение отношения $t = (n^{2k} - n^{2j})/\bigl(n^k - n^j + (-n^{-n}$ или $0,$

или n^{-n})) лежит строго между $n^j + n^k - 1$ и $n^j + n^k + 1$; следовательно, ордината точки пересечения единственным образом определяет j и k. Различны также четыре точки пересечения с одинаковыми j и k.

1.19 Не в случае $n > 5$. Ломаная линия, полупрямые которой образуют при вершине углы θ и $\theta+30°$, может пересечься четыре раза с другой ломаной, полупрямые которой образуют углы ϕ и $\phi + 30°$, только если $30° < |\theta - \phi| < 150°$. Невозможно выбрать более 5 углов, удаленных друг от друга на требуемое расстояние. (5 углов выбрать можно.)

1.20 Пусть $h(n) = a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 + d(n)\gamma_0 + e(n)\gamma_1$. Из (1.18) известно, что $a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 = (\alpha\, \beta_{b_{m-1}}\, \beta_{b_{m-2}} \ldots \beta_{b_1}\, \beta_{b_0})_4$ при $n = (1\, b_{m-1} \ldots b_1\, b_0)_2$; это определяет $a(n)$, $b(n)$ и $c(n)$. Подстановка $h(n) = n$ в рекуррентность дает $a(n)+c(n)-2d(n) - 2e(n) = n$; подстановка $h(n) = n^2$ дает $a(n) + c(n) + 4e(n) = n^2$. Следовательно, $d(n) = \bigl(3a(n) + 3c(n) - n^2 - 2n\bigr)/4$; $e(n) = \bigl(n^2 - a(n) - c(n)\bigr)/4$.

1.21 Можно положить m равным наименьшему (или любому иному) общему кратному чисел $2n$, $2n-1$, $\ldots$, $n+1$. (Нестрогие рассуждения приводят к тому, что "случайное" значение m подойдет с вероятностью

$$\frac{n}{2n}\,\frac{n-1}{2n-1}\cdots\frac{1}{n+1} \;=\; 1\Big/\binom{2n}{n} \;\sim\; \frac{\sqrt{\pi n}}{4^n}\,,$$

так что можно ожидать, что удастся найти m, меньшее чем 4^n.)

1.22 Возьмите правильный многоугольник с 2^n сторонами и пометьте стороны элементами "цикла де Брейна" (de Bruijn) длиной 2^n. (Это циклическая последовательность из 0 и 1, в которой все n-кортежи из смежных элементов различны; см. [207, упр. 2.3.4.2–23] и [208, упр. 3.2.2–17].) Добавьте к каждой стороне, помеченной 1, очень тонкое выпуклое расширение. Требуемые множества представляют собой n таких многоугольников, повернутых "на длину" k сторон при $k = 0, 1, \ldots, n-1$.

Довелось мне покататься на [мото]цикле де Брейна, будучи у него в гостях в голландском городе Нойене.

1.23 Да. (Для решения необходимо знать основные понятия элементарной теории чисел из главы 4.) Пусть $L(n) = \text{НОК}(1, 2, \ldots, n)$. Будем считать, что $n > 2$; тогда в соответствии с постулатом Бертрана между $n/2$ и n имеется простое число p. Будем также считать, что $j > n/2$, поскольку $q' = L(n)+1-q$ оставляет в живых $j' = n+1-j$ тогда и только тогда, когда q оставляет в живых j. Выберем q так, что $q \equiv 1 \pmod{L(n)/p}$ и $q \equiv j+1-n \pmod p$. В этом случае люди выбывают в порядке $1, 2, \ldots, n-p$, $j+1, j+2, \ldots, n, n-p+1, \ldots, j-1$.

1.24 Вот единственные известные примеры: $X_n = 2i\sin\pi r + 1/X_{n-1}$, где r — рациональное число и $0 \leqslant r < \frac{1}{2}$ (при изменении r все периоды имеют длину $\geqslant 2$); рекуррентность Гаусса (Gauss) с периодом 5 из упр. 8; еще более примечательная рекуррентность Тодда (H. Todd) $X_n = (1 + X_{n-1} + X_{n-2})/X_{n-3}$ с периодом 8 (см. [261]); а также рекуррентности, получающиеся из указанных путем замены X_n на X_{mn}, умноженного на константу. Можно считать, что первый ненулевой коэффициент в знаменателе равен 1 и первый ненулевой коэффициент в числителе (при наличии такового коэффициента) имеет неотрицательную вещественную часть. Компьютерная алгебра позволяет легко показать отсутствие иных решений с периодом $\leqslant 5$ при $k = 2$. Частичная теория была разработана Линессом (Lyness) [261, 262] и Куршаном (Kurshan) и Гопинатом (Gopinath) [231].

Интересным примером иного типа с вещественными начальными значениями служит рекуррентность $X_n = |X_{n-1}| - X_{n-2}$ с периодом 9, открытая Мортоном Брауном (Morton Brown) [43]. Получение нелинейных рекуррентностей с любым наперед заданным периодом $\geqslant 5$ может основываться на континуантах [65].

1.25 Если обозначить минимальное количество переносов, необходимых для перемещения n дисков с k вспомогательными колышками как $T^{(k)}(n)$ (так что $T^{(1)}(n) = T_n$ и $T^{(2)}(n) = W_n$), то $T^{(k)}(\binom{n+1}{k}) \leqslant 2T^{(k)}(\binom{n}{k}) + T^{(k-1)}(\binom{n}{k-1})$. Примеры пар (n, k), для которых это неравенство не является равенством, не известны. Если k мало́ по сравнению с n, формула $2^{n+1-k}\binom{n-1}{k-1}$ дает приемлемую верхнюю границу $T^{(k)}(\binom{n}{k})$.

1.26 Перестановка “порядка казни” может быть вычислена за $O(n \log n)$ шагов для всех m и n [209, упр. 5.1.1–2 и 5.1.1–5]. Бьёрн Поонен (Bjorn Poonen) доказал, что неиосифовы множества ровно с четырьмя “не нашими” существуют при любых $n \equiv 0 \pmod 3$ и $n \geqslant 9$; в действительности таких множеств имеется по меньшей мере $\epsilon\binom{n}{4}$ для некоторого $\epsilon > 0$. Он также обнаружил при помощи большого количества вычислений, что единственным среди других $n < 24$ с неиосифовыми множествами является $n = 20$, которое имеет 236 таких множеств при $k = 14$ и два при $k = 13$. (Одно из последних — $\{1, 2, 3, 4, 5, 6, 7, 8, 11, 14, 15, 16, 17\}$; другое — его отражение относительно 21.) При $n = 15$ и $k = 9$ имеется единственное неиосифово множество, а именно — $\{3, 4, 5, 6, 8, 10, 11, 12, 13\}$.

2.1 По этому поводу нет общего согласия — приемлемы три ответа. (1) Можно считать, что $\sum_{k=m}^{n} q_k$ всегда эквивалентна

$\sum_{m\leqslant k\leqslant n} q_k$; тогда указанная сумма равна нулю. (2) Можно было бы сказать, что данная сумма равна $q_4+q_3+q_2+q_1+q_0$, если суммировать по уменьшающимся значениям k. Но это противоречит общепринятому соглашению о том, что $\sum_{k=1}^{n} q_k = 0$ при $n=0$. (3) Можно считать, что $\sum_{k=m}^{n} q_k = \sum_{k\leqslant n} q_k - \sum_{k<m} q_k$; тогда указанная сумма равна $-q_1-q_2-q_3$. Такое соглашение может показаться странным, но оно удовлетворяет полезному правилу $\sum_{k=a}^{b} + \sum_{k=b+1}^{c} = \sum_{k=a}^{c}$ при всех a, b, c.

Лучше всего использовать обозначение $\sum_{k=m}^{n}$, только когда $n-m \geqslant -1$; тогда и (1), и (3) дают одинаковые результаты.

2.2 Это $|x|$. Кстати, величина $\bigl([x>0]-[x<0]\bigr)$ часто называется “знаком x” и обозначается $\mathrm{sign}(x)$. Она равна $+1$ при $x>0$, 0 при $x=0$ и -1 при $x<0$.

2.3 Первая сумма, конечно, представляет собой $a_0+a_1+a_2+a_3+a_4+a_5$; вторая— $a_4+a_1+a_0+a_1+a_4$, так как это сумма по значениям $k\in\{-2,-1,0,+1,+2\}$. Коммутативный закон в данном случае не применим, поскольку функция $p(k)=k^2$ не является перестановкой. Некоторым значениям n (например, $n=3$) не соответствует ни одно k, такое, что $p(k)=n$; другим (например, $n=4$) соответствуют два таких k.

2.4 Один и тот же индекс ‘k’ использован для обозначения двух разных индексных переменных, хотя k связан с внутренней суммой. Это широко распространенная как в математике, так и в программировании ошибка. Результат будет верен, если $a_j=a_k$ при всех j и k, $1\leqslant j,k\leqslant n$.

2.5 (a) $\sum_{i=1}^{4}\sum_{j=i+1}^{4}\sum_{k=j+1}^{4} a_{ijk}=\sum_{i=1}^{2}\sum_{j=i+1}^{3}\sum_{k=j+1}^{4} a_{ijk}=\bigl((a_{123}+a_{124})+a_{134}\bigr)+a_{234}$.

(б) $\sum_{k=1}^{4}\sum_{j=1}^{k-1}\sum_{i=1}^{j-1} a_{ijk}=\sum_{k=3}^{4}\sum_{j=2}^{k-1}\sum_{i=1}^{j-1} a_{ijk}=a_{123}+\bigl(a_{124}+(a_{134}+a_{234})\bigr)$.

2.6 Это $[1\leqslant j\leqslant n](n-j+1)$. Первый множитель необходим здесь потому, что при $j<1$ или $j>n$ должен получаться нуль.

2.7 Это $mx^{\overline{m-1}}$. Версия разностного исчисления на основе ∇, а не Δ, сделала бы факториальные степени неравноправными, особо выделив *возрастающие* факториальные степени.

2.8 0, если $m\geqslant 1$; $1/|m|!$, если $m\leqslant 0$.

2.9 $x^{\overline{m+n}} = x^{\overline{m}}(x+m)^{\overline{n}}$ при целых m и n. Полагая $m=-n$, получим, что $x^{\overline{-n}} = 1/(x-n)^{\overline{n}} = 1/(x-1)^{\underline{n}}$.

2.10 Другой возможный вариант записи правой части: $\mathrm{E}u\,\Delta v + v\,\Delta u$.

2.11 Разбейте левую часть на две суммы и во второй из них замените k на $k+1$.

2.12 Если $p(k) = n$, то $n + c = k + \big((-1)^k + 1\big)c$ и $\big((-1)^k + 1\big)$ четное; следовательно, $(-1)^{n+c} = (-1)^k$ и $k = n - (-1)^{n+c}c$. И обратно, это значение k дает $p(k) = n$.

2.13 Пусть $R_0 = \alpha$ и $R_n = R_{n-1} + (-1)^n(\beta + n\gamma + n^2\delta)$ при $n > 0$. Тогда $R(n) = A(n)\alpha + B(n)\beta + C(n)\gamma + D(n)\delta$. Подстановка $R_n = 1$ дает $A(n) = 1$. Подстановка $R_n = (-1)^n$ дает $A(n) + 2B(n) = (-1)^n$. Подстановка $R_n = (-1)^n n$ дает $-B(n) + 2C(n) = (-1)^n n$. Подстановка $R_n = (-1)^n n^2$ дает $B(n) - 2C(n) + 2D(n) = (-1)^n n^2$. Следовательно, $2D(n) = (-1)^n(n^2 + n)$; искомая сумма равна $D(n)$.

2.14 Предложенная запись корректна, поскольку $k = \sum_{1\leqslant j\leqslant k} 1$ при $1 \leqslant k \leqslant n$. Сначала суммируем по k; в результате кратная сумма сводится к $\sum_{1\leqslant j\leqslant n}(2^{n+1} - 2^j) = n2^{n+1} - (2^{n+1} - 2)$.

2.15 На первом шаге заменяем $k(k+1)$ на $2\sum_{1\leqslant j\leqslant k} j$. Второй шаг дает $⧉_n + \square_n = \big(\sum_{k=1}^{n} k\big)^2 + \square_n$.

2.16 $x^{\underline{m}}(x-m)^{\underline{n}} = x^{\underline{m+n}} = x^{\underline{n}}(x-n)^{\underline{m}}$ согласно (2.52).

2.17 Воспользуйтесь индукцией для доказательства первых двух равенств и правилом (2.52) — для третьего. Вторая формула вытекает из первой.

2.18 Воспользуйтесь тем, что $(\Re z)^+ \leqslant |z|$, $(\Re z)^- \leqslant |z|$, $(\Im z)^+ \leqslant |z|$, $(\Im z)^- \leqslant |z|$ и $|z| \leqslant (\Re z)^+ + (\Re z)^- + (\Im z)^+ + (\Im z)^-$.

2.19 Умножьте обе части на $2^{n-1}/n!$ и положите $S_n = 2^n T_n/n! = S_{n-1} + 3\cdot 2^{n-1} = 3(2^n - 1) + S_0$. Решением будет $T_n = 3\cdot n! + n!/2^{n-1}$. (Как мы увидим в главе 4, T_n является целым числом только тогда, когда n равно 0 или степени 2.)

2.20 Метод приведения дает

$$S_n + (n+1)H_{n+1} = S_n + \Big(\sum_{0\leqslant k\leqslant n} H_k\Big) + n + 1\,.$$

2.21 Выделение последнего члена в S_{n+1} дает $S_{n+1} = 1 - S_n$; выделение первого члена дает

$$\begin{aligned} S_{n+1} &= (-1)^{n+1} + \sum_{1\leqslant k\leqslant n+1} (-1)^{n+1-k} = \\ &= (-1)^{n+1} + \sum_{0\leqslant k\leqslant n} (-1)^{n-k} = \\ &= (-1)^{n+1} + S_n\,. \end{aligned}$$

“Глубоко ошибочный трюизм, который повторяется во всех прописях и речах почтенных ораторов, сводится к тому, что мы должны иметь привычку обдумывать свои поступки. В действительности имеет место как раз обратное. Прогресс цивилизации выражается в умножении числа поведенческих актов, которые мы совершаем необдуманно. Мыслительные же акты подобны кавалерийским атакам, число которых весьма ограниченно, — они требуют свежих лошадей и должны совершаться лишь в решающие моменты битвы”.
— А. Н. Уайтхед (A. N. Whitehead) [370]

Следовательно, $2S_n = 1 + (-1)^n$ и $S_n = [n \text{ четное}]$. Аналогично находим

$$T_{n+1} = n+1-T_n = \sum_{k=0}^{n}(-1)^{n-k}(k+1) = T_n + S_n\,,$$

следовательно, $2T_n = n + 1 - S_n$ и $T_n = \frac{1}{2}(n + [n \text{ нечетное}])$. Наконец, тот же подход приводит к

$$\begin{aligned} U_{n+1} = (n+1)^2 - U_n &= U_n + 2T_n + S_n = \\ &= U_n + n + [n \text{ нечетное}] + [n \text{ четное}] = \\ &= U_n + n + 1\,. \end{aligned}$$

Итак, U_n — треугольное число $\frac{1}{2}(n+1)n$.

2.22 Удвоение общей суммы дает “ванильную” сумму по $1 \leqslant j, k \leqslant n$, которая раскладывается и дает

$$\Bigl(\sum_k a_k A_k\Bigr)\Bigl(\sum_k b_k B_k\Bigr) - \Bigl(\sum_k a_k B_k\Bigr)\Bigl(\sum_k b_k A_k\Bigr).$$

2.23 (a) Этот подход дает четыре суммы, вычисление которых приводит к $2n + H_n - 2n + (H_n + \frac{1}{n+1} - 1)$. (Было бы проще заменить общий член на $1/k + 1/(k+1)$.) (б) Пусть $u(x) = 2x+1$ и $\Delta v(x) = 1/x(x+1) = (x-1)^{\underline{-2}}$; тогда $\Delta u(x) = 2$ и $v(x) = -(x-1)^{\underline{-1}} = -1/x$. Ответ: $2H_n - \frac{n}{n+1}$.

2.24 Суммируя по частям, $\sum x^{\underline{m}} H_x\,\delta x = x^{\underline{m+1}} H_x/(m+1) - x^{\underline{m+1}}/(m+1)^2 + C$; следовательно, $\sum_{0\leqslant k<n} k^{\underline{m}} H_k = n^{\underline{m+1}}(H_n - 1/(m+1))/(m+1) + 0^{\underline{m+1}}/(m+1)^2$. В нашем случае $m = -2$, так что сумма сводится к $1 - (H_n+1)/(n+1)$.

2.25 Вот некоторые основные аналоги.

$$\sum_{k\in K} c a_k = c\sum_{k\in K} a_k \quad\longleftrightarrow\quad \prod_{k\in K} a_k^c = \Bigl(\prod_{k\in K} a_k\Bigr)^c$$

$$\sum_{k\in K}(a_k+b_k) = \sum_{k\in K} a_k + \sum_{k\in K} b_k \quad\longleftrightarrow\quad \prod_{k\in K} a_k b_k = \Bigl(\prod_{k\in K} a_k\Bigr)\Bigl(\prod_{k\in K} b_k\Bigr)$$

$$\sum_{k\in K} a_k = \sum_{p(k)\in K} a_{p(k)} \quad\longleftrightarrow\quad \prod_{k\in K} a_k = \prod_{p(k)\in K} a_{p(k)}$$

$$\sum_{\substack{j\in J\\ k\in K}} a_{j,k} = \sum_{j\in J}\sum_{k\in K} a_{j,k} \quad\longleftrightarrow\quad \prod_{\substack{j\in J\\ k\in K}} a_{j,k} = \prod_{j\in J}\prod_{k\in K} a_{j,k}$$

$$\sum_{k\in K} a_k = \sum_k a_k[k\in K] \quad\longleftrightarrow\quad \prod_{k\in K} a_k = \prod_k a_k^{[k\in K]}$$

$$\sum_{k\in K} 1 = \#K \quad\longleftrightarrow\quad \prod_{k\in K} c = c^{\#K}$$

2.26 $P^2 = \left(\prod_{1\leqslant j,k\leqslant n} a_j a_k\right)\left(\prod_{1\leqslant j=k\leqslant n} a_j a_k\right)$. Первый множитель равен $\left(\prod_{k=1}^{n} a_k^n\right)^2$, а второй — $\prod_{k=1}^{n} a_k^2$. Следовательно, $P = \left(\prod_{k=1}^{n} a_k\right)^{n+1}$.

2.27 $\Delta(c^{\underline{x}}) = c^{\underline{x}}(c-x-1) = c^{\underline{x+2}}/(c-x)$. Подстановка $c=-2$ и уменьшение x на 2 дает $\Delta(-(-2)^{\underline{x-2}}) = (-2)^{\underline{x}}/x$, так что искомая сумма равна $(-2)^{\underline{-1}} - (-2)^{\underline{n-1}} = (-1)^n n! - 1$.

2.28 Незаконно изменение порядка суммирования при переходе от второй строки к третьей — члены этой суммы не сходятся абсолютно. Все остальное абсолютно корректно, за исключением того, что промежуточный результат $\sum_{k\geqslant 1}[k=j-1]k/j$ следовало бы, пожалуй, записать в виде $[j-1\geqslant 1](j-1)/j$ и упростить.

Может ли что-то быть корректным не абсолютно?

2.29 Воспользуйтесь элементарными дробями для того, чтобы получить

$$\frac{k}{4k^2-1} = \frac{1}{4}\left(\frac{1}{2k+1} + \frac{1}{2k-1}\right).$$

Теперь множитель $(-1)^k$ заставляет обе половины каждого члена суммы сократиться с их соседями. Следовательно, искомый ответ равен $-1/4 + (-1)^n/(8n+4)$.

2.30 $\sum_a^b x\,\delta x = \frac{1}{2}(b^{\underline{2}} - a^{\underline{2}}) = \frac{1}{2}(b-a)(b+a-1)$, поэтому

$$(b-a)(b+a-1) = 2100 = 2^2\cdot 3\cdot 5^2\cdot 7.$$

Существует одно решение для каждой записи 2100 в виде $x\cdot y$, где x четное, а y — нечетное: мы полагаем $a = \frac{1}{2}|x-y| + \frac{1}{2}$ и $b = \frac{1}{2}(x+y) + \frac{1}{2}$. Поэтому искомое число решений равно числу делителей $3\cdot 5^2\cdot 7$, а именно — 12. В общем случае имеется $\prod_{p>2}(n_p+1)$ способов представления $\prod_p p^{n_p}$, где произведения берутся по простым числам.

2.31 $\sum_{j,k\geqslant 2} j^{-k} = \sum_{j\geqslant 2} 1/j^2(1-1/j) = \sum_{j\geqslant 2} 1/j(j-1)$. Аналогично находится вторая сумма 3/4.

2.32 Если $2n \leqslant x < 2n+1$, то данные суммы суть $0+\cdots+n+(x-n-1)+\cdots+(x-2n) = n(x-n) = (x-1)+(x-3)+\cdots+(x-2n+1)$. Если же $2n-1 \leqslant x < 2n$, то аналогичным образом обе суммы равны $n(x-n)$. (Забегая вперед, в главу 3, заметим, что оба случая описываются формулой $\lfloor\frac{1}{2}(x+1)\rfloor\left(x - \lfloor\frac{1}{2}(x+1)\rfloor\right)$.)

2.33 Если K — пустое множество, то $\bigwedge_{k\in K} a_k = \infty$. Вот основные аналоги

$$\sum_{k\in K} ca_k = c\sum_{k\in K} a_k \qquad\longleftrightarrow\qquad \bigwedge_{k\in K}(c+a_k) = c+\bigwedge_{k\in K} a_k$$

$$\sum_{k\in K}(a_k+b_k) = \sum_{k\in K} a_k+\sum_{k\in K} b_k \quad\longleftrightarrow\quad \bigwedge_{k\in K}\min(a_k,b_k) = \min\Bigl(\bigwedge_{k\in K} a_k, \bigwedge_{k\in K} b_k\Bigr)$$

$$\sum_{k\in K} a_k = \sum_{p(k)\in K} a_{p(k)} \quad\longleftrightarrow\quad \bigwedge_{k\in K} a_k = \bigwedge_{p(k)\in K} a_{p(k)}$$

$$\sum_{\substack{j\in J\\ k\in K}} a_{j,k} = \sum_{j\in J}\sum_{k\in K} a_{j,k} \quad\longleftrightarrow\quad \bigwedge_{\substack{j\in J\\ k\in K}} a_{j,k} = \bigwedge_{j\in J}\bigwedge_{k\in K} a_{j,k}$$

$$\sum_{k\in K} a_k = \sum_{k} a_k[k\in K] \quad\longleftrightarrow\quad \bigwedge_{k\in K} a_k = \bigwedge_{k} a_k\infty^{[k\notin K]}$$

Перестановка, поглощающая члены одного знака быстрее, чем другого, может сделать сумму равной какому угодно значению.

2.34 Пусть $K^+=\{k \mid a_k \geqslant 0\}$ и $K^-=\{k \mid a_k < 0\}$. Тогда, если, например, n нечетное, мы выбираем F_n равным $F_{n-1} \cup E_n$, где подмножество $E_n \subseteq K^-$ достаточно большое, чтобы $\sum_{k\in(F_{n-1}\cap K^+)} a_k - \sum_{k\in E_n}(-a_k) < A^-$.

2.35 То, что сумма Гольдбаха равна

$$\sum_{m,n\geqslant 2} m^{-n} = \sum_{m\geqslant 2}\frac{1}{m(m-1)} = 1,$$

можно показать следующим образом. Если развернуть каждый ее член в сумму геометрической прогрессии, то она равна $\sum_{k\in P,\, l\geqslant 1} k^{-l}$; таким образом, доказательство будет завершено, если мы сможем найти взаимно однозначное соответствие между упорядоченными парами (m,n) с $m,n \geqslant 2$ и упорядоченными парами (k,l) с $k \in P$ и $l \geqslant 1$, где $m^n = k^l$ в случае соответствия этих пар. Если $m \notin P$, мы полагаем $(m,n) \longleftrightarrow (m^n,1)$, но если $m = a^b \in P$, то полагаем $(m,n) \longleftrightarrow (a^n,b)$.

При таком самоописании вряд ли кому-то захочется ее описывать...

2.36 (а) По определению $g(n) - g(n-1) = f(n)$. (б) Согласно части (а) $g\bigl(g(n)\bigr) - g\bigl(g(n-1)\bigr) = \sum_k f(k)\bigl[g(n-1)<k\leqslant g(n)\bigr] = n\bigl(g(n)-g(n-1)\bigr) = nf(n)$. (в) Вновь в соответствии с частью (а) $g\bigl(g(g(n))\bigr) - g\bigl(g(g(n-1))\bigr)$ представляет собой

$$\begin{aligned}
&\sum_k f(k)\bigl[g(g(n-1))<k\leqslant g(g(n))\bigr] = \\
&\quad= \sum_{j,k} j\,\bigl[j=f(k)\bigr]\bigl[g(g(n-1))<k\leqslant g(g(n))\bigr] = \\
&\quad= \sum_{j,k} j\,\bigl[j=f(k)\bigr]\bigl[g(n-1)<j\leqslant g(n)\bigr] = \\
&\quad= \sum_{j} j\,\bigl(g(j)-g(j-1)\bigr)\bigl[g(n-1)<j\leqslant g(n)\bigr] = \\
&\quad= \sum_{j} jf(j)\,\bigl[g(n-1)<j\leqslant g(n)\bigr] = n\sum_j j\,\bigl[g(n-1)<j\leqslant g(n)\bigr].
\end{aligned}$$

Колин Мэллоуз (Colin Mallows) заметил, что эта последовательность может быть определена при помощи рекуррентности

$$f(1) = 1; \qquad f(n+1) = 1 + f\bigl(n+1-f(f(n))\bigr) \quad \text{при } n \geqslant 0.$$

2.37 (Грэхем полагает, что, скорее всего, нельзя; Кнут считает, что, скорее всего, можно; Паташник благоразумно не компрометирует себя.)

3.1 $m = \lfloor \lg n \rfloor$; $l = n - 2^m = n - 2^{\lfloor \lg n \rfloor}$.

3.2 (а) $\lfloor x + 0.5 \rfloor$. (б) $\lceil x - 0.5 \rceil$.

3.3 Это $\lfloor mn - \{m\alpha\}n/\alpha \rfloor = mn - 1$, поскольку $0 < \{m\alpha\} < 1$.

3.4 Что-то, требующее не доказательства, а удачного предположения (я так предполагаю).

3.5 Имеем $\lfloor nx \rfloor = \lfloor n\lfloor x \rfloor + n\{x\} \rfloor = n\lfloor x \rfloor + \lfloor n\{x\} \rfloor$ согласно (3.8) и (3.6). Следовательно, $\lfloor nx \rfloor = n\lfloor x \rfloor \iff \lfloor n\{x\} \rfloor = 0 \iff 0 \leqslant n\{x\} < 1 \iff \{x\} < 1/n$, в предположении, что n — положительное целое число. (Заметим, что в этом случае $n\lfloor x \rfloor \leqslant \lfloor nx \rfloor$ при всех x.)

3.6 $\lfloor f(x) \rfloor = \lfloor f(\lceil x \rceil) \rfloor$.

3.7 $\lfloor n/m \rfloor + n \bmod m$.

3.8 Если в каждом из ящиков содержится $< \lceil n/m \rceil$ предметов, то $n \leqslant (\lceil n/m \rceil - 1)m$, так что $n/m + 1 \leqslant \lceil n/m \rceil$, что противоречит (3.5). Доказательство второго факта выполняется аналогично.

3.9 Имеем $m/n - 1/q = (n \operatorname{mumble} m)/qn$. Данный процесс должен завершаться, поскольку $0 \leqslant n \operatorname{mumble} m < m$. Знаменатели в таком представлении строго возрастают, а следовательно, они различны, так как $qn/(n \operatorname{mumble} m) > q$.

3.10 $\lceil x + \frac{1}{2} \rceil - \bigl[(2x+1)/4 \text{ — не целое}\bigr]$ — это ближайшее целое к x, если $\{x\} \neq \frac{1}{2}$; в противном случае это ближайшее четное целое число. (См. упр. 2.) Таким образом, данная формула дает "несмещенный" способ округления.

3.11 Если n — целое число, то $\alpha < n < \beta \iff \lfloor \alpha \rfloor < n < \lceil \beta \rceil$. Количество целых чисел, удовлетворяющих условию $a < n < b$ при целых a и b, равно $(b - a - 1)[b > a]$. Поэтому, если бы было $\alpha = \beta =$ целое, мы бы получили неверный ответ.

3.12 Вычтем из обеих частей $\lfloor n/m \rfloor$ в соответствии с (3.6) и получим $\lceil (n \bmod m)/m \rceil = \lfloor (n \bmod m + m - 1)/m \rfloor$. Теперь обе части равны $[n \bmod m > 0]$, поскольку $0 \leqslant n \bmod m < m$.

Более короткое, но менее прямое доказательство состоит в наблюдении, что первый член (3.24) должен быть равен последнему члену в (3.25).

3.13 Если они образуют разбиение, то формула для $N(\alpha, n)$ из текста влечет за собой $1/\alpha + 1/\beta = 1$, поскольку коэффициенты при n в уравнении $N(\alpha, n) + N(\beta, n) = n$ должны быть согласованы так, чтобы данное уравнение выполнялось при больших n. Следовательно, числа α и β либо оба рациональны, либо оба иррациональны. Если они оба иррациональны, то мы получаем разбиение, как показано в тексте. Если же оба они могут быть записаны с числителем m, то значение $m - 1$ не встретится ни в одном из спектров, а m будет находиться в обоих. (Однако Голомб (Golomb) [151] обнаружил, что множества $\{\lfloor n\alpha \rfloor \mid n \geqslant 1\}$ и $\{\lceil n\beta \rceil - 1 \mid n \geqslant 1\}$ всегда образуют разбиение, когда $1/\alpha + 1/\beta = 1$.)

3.14 При $ny = 0$ это очевидно из (3.22); в противном случае это истинно на основании (3.21) и (3.6).

3.15 Подставьте $\lceil mx \rceil$ вместо n в (3.24): $\lceil mx \rceil = \lceil x \rceil + \lceil x - \frac{1}{m} \rceil + \cdots + \lceil x - \frac{m-1}{m} \rceil$.

3.16 Можно убедиться в справедливости формулы $n \bmod 3 = 1 + \frac{1}{3}\big((\omega - 1)\omega^n - (\omega + 2)\omega^{2n}\big)$, проверив ее при $0 \leqslant n < 3$.

Общая формула для $n \bmod m$ при любом целом положительном m приводится в упр. 7.25.

3.17 $\sum_{j,k} [0 \leqslant k < m][1 \leqslant j \leqslant x + k/m] = \sum_{j,k} [0 \leqslant k < m][1 \leqslant j \leqslant \lceil x \rceil]\big[k \geqslant m(j - x)\big] = \sum_{1 \leqslant j \leqslant \lceil x \rceil} \sum_k [0 \leqslant k < m] - \sum_{j = \lceil x \rceil} \sum_k \big[0 \leqslant k < m(j - x)\big] = m\lceil x \rceil - \big\lceil m(\lceil x \rceil - x) \big\rceil = -\lceil -mx \rceil = \lfloor mx \rfloor$.

3.18 Имеем

$$S = \sum_{0 \leqslant j < \lceil n\alpha \rceil} \sum_{k \geqslant n} \big[j\alpha^{-1} \leqslant k < (j + \nu)\alpha^{-1}\big].$$

Если $j \leqslant n\alpha - 1 \leqslant n\alpha - \nu$, вклад отсутствует, потому что $(j + \nu)\alpha^{-1} \leqslant n$. Следовательно, $j = \lfloor n\alpha \rfloor$ — единственный имеющий значение случай, и искомая величина в этом случае равна $\big\lceil (\lfloor n\alpha \rfloor + \nu)\alpha^{-1} \big\rceil - n \leqslant \lceil \nu\alpha^{-1} \rceil$.

3.19 Данное равенство выполняется тогда и только тогда, когда b — целое число. (Если b целое, $\log_b x$ — непрерывная возрастающая функция, которая принимает целочисленные значения только в целых точках. Если b не целое, условие не выполняется при $x = b$.)

3.20 Имеем $\sum_k kx[\alpha \leqslant kx \leqslant \beta] = x \sum_k k\big[\lceil \alpha/x \rceil \leqslant k \leqslant \lfloor \beta/x \rfloor\big]$, что в сумме дает $\frac{1}{2}x\big(\lfloor \beta/x \rfloor \lfloor \beta/x + 1 \rfloor - \lceil \alpha/x \rceil \lceil \alpha/x - 1 \rceil\big)$.

3.21 Если $10^n \leqslant 2^M < 10^{n+1}$, то имеется ровно $n+1$ таких степеней 2, поскольку для каждого k имеется только одна такая степень 2 из k цифр. Так что ответ — $1 + \lfloor M \log 2 \rfloor$.

Примечание: найти количество степеней 2, начинающихся с цифры $l > 1$ более сложно; оно равно $\sum_{0 \leqslant n \leqslant M} \bigl(\lfloor n \log 2 - \log l \rfloor - \lfloor n \log 2 - \log(l+1) \rfloor\bigr)$.

3.22 Все члены при n и $n-1$ одинаковы, за исключением k-го, где $n = 2^{k-1}q$ и q нечетное; имеем $S_n = S_{n-1} + 1$ и $T_n = T_{n-1} + 2^k q$. Следовательно, $S_n = n$ и $T_n = n(n+1)$.

3.23 $X_n = m \iff \frac{1}{2}m(m-1) < n \leqslant \frac{1}{2}m(m+1) \iff m^2 - m + \frac{1}{4} < 2n < m^2 + m + \frac{1}{4} \iff m - \frac{1}{2} < \sqrt{2n} < m + \frac{1}{2}$.

3.24 Пусть $\beta = \alpha/(\alpha+1)$. Тогда число вхождений целого неотрицательного m в $\mathrm{Spec}(\beta)$ ровно на единицу больше количества его вхождений в $\mathrm{Spec}(\alpha)$. Почему? Потому что $N(\beta, n) = N(\alpha, n) + n + 1$.

3.25 Если бы мы могли, продолжая начатое в тексте, найти такое значение m, что $K_m \leqslant m$, то мы бы нарушили указанное неравенство для $n+1$, когда $n = 2m+1$ (а также когда $n = 3m+1$ и $n = 3m+2$). Но для существования такого $m = n'+1$ требуется, чтобы $2K_{\lfloor n'/2 \rfloor} \leqslant n'$ или $3K_{\lfloor n'/3 \rfloor} \leqslant n'$, т.е. чтобы

$$K_{\lfloor n'/2 \rfloor} \leqslant \lfloor n'/2 \rfloor \qquad \text{или} \qquad K_{\lfloor n'/3 \rfloor} \leqslant \lfloor n'/3 \rfloor .$$

Ага! Это уводит нас все дальше и дальше, приводя к тому, что $K_0 \leqslant 0$; но $K_0 = 1$.

Что нам действительно нужно доказать, так это то, что K_n *строго больше* n для всех $n > 0$. Это легко доказывается по индукции, несмотря на то, что это более сильный результат, чем тот, который мы не могли доказать!

(Данное упражнение преподает нам важный урок. Это в большей степени упражнение, посвященное природе математической индукции, чем свойствам функции “пол”.)

“Когда вы пытаетесь найти доказательство при помощи математической индукции, оно может не удаваться вам по двум противоположным причинам. Оно может не удаваться потому, что вы пытаетесь доказать слишком много: ваше $P(n+1)$ — слишком тяжелый груз. Неудача может быть вызвана и тем, что вы пытаетесь доказать слишком мало: ваше $P(n)$ — слишком слабая опора. В общем случае вы должны сбалансировать утверждение своей теоремы так, чтобы опора была как раз достаточной для груза.”
— Г. Пойа (G. Pólya) [297]

3.26 Воспользуйтесь доказательством по индукции с более сильной гипотезой

$$D_n^{(q)} \leqslant (q-1)\left(\left(\frac{q}{q-1}\right)^{n+1} - 1\right) \qquad \text{при } n \geqslant 0.$$

3.27 Если $D_n^{(3)} = 2^m b - a$, где a равно 0 или 1, то $D_{n+m}^{(3)} = 3^m b - a$.

3.28 Ключевое наблюдение заключается в том, что $a_n = m^2$ влечет за собой $a_{n+2k+1} = (m+k)^2 + m - k$ и $a_{n+2k+2} = (m+k)^2 + 2m$ при $0 \leqslant k \leqslant m$; следовательно, $a_{n+2m+1} = (2m)^2$. Решение может быть записано в красивом виде, найденном Карлом Уитти (Carl Witty):

$$a_{n-1} = 2^l + \left\lfloor \left(\frac{n-l}{2}\right)^2 \right\rfloor, \quad \text{где } 2^l + l \leqslant n < 2^{l+1} + l + 1.$$

3.29 Величина $D(\alpha', \lfloor \alpha n \rfloor)$ не превышает максимума абсолютного значения выражения

$$s(\alpha', \lfloor n\alpha \rfloor, \nu') = -s(\alpha, n, \nu) - S + \epsilon + \{0 \text{ или } 1\} + \nu' - \{0 \text{ или } 1\}.$$

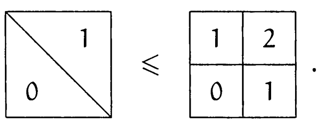
3.30 По индукции $X_n = \alpha^{2^n} + \alpha^{-2^n}$ и X_n — целое число.

Эта логика просто валит с ног на пол!

3.31 Вот "элегантное", "впечатляющее" доказательство (непонятно, каким образом оно было открыто):

$$\begin{aligned}
\lfloor x \rfloor + \lfloor y \rfloor + \lfloor x+y \rfloor &= \lfloor x + \lfloor y \rfloor \rfloor + \lfloor x+y \rfloor \leqslant \\
&\leqslant \lfloor x + \tfrac{1}{2}\lfloor 2y \rfloor \rfloor + \lfloor x + \tfrac{1}{2}\lfloor 2y \rfloor + \tfrac{1}{2} \rfloor = \\
&= \lfloor 2x + \lfloor 2y \rfloor \rfloor = \lfloor 2x \rfloor + \lfloor 2y \rfloor .
\end{aligned}$$

Но есть простое графическое доказательство, основанное на том наблюдении, что требуется рассмотреть только случай $0 \leqslant x$, $y < 1$. Тогда соотношение графически можно изобразить следующим образом:



Возможен и несколько более сильный результат, а именно

$$\lceil x \rceil + \lfloor y \rfloor + \lfloor x+y \rfloor \leqslant \lceil 2x \rceil + \lfloor 2y \rfloor ;$$

но он сильнее только при $\{x\} = \frac{1}{2}$. Если заменить (x, y) на $(-x, x+y)$ и применить рефлексивный закон (3.4), то можно получить

$$\lfloor y \rfloor + \lfloor x+y \rfloor + \lfloor 2x \rfloor \leqslant \lfloor x \rfloor + \lfloor 2x+2y \rfloor .$$

3.32 Обозначим интересующую нас сумму как $f(x)$. Поскольку $f(x) = f(-x)$, можно считать, что $x \geqslant 0$. Члены суммы ограничены величиной 2^k при $k \to -\infty$ и величиной $x^2/2^k$ при $k \to +\infty$, так что сумма существует при любом действительном x.

Имеем $f(2x) = 2\sum_k 2^{k-1}\|x/2^{k-1}\|^2 = 2f(x)$. Пусть $f(x) = l(x) + r(x)$, где $l(x)$ — исходная сумма при $k \leqslant 0$, а $r(x)$ — исходная сумма при $k > 0$. Тогда $l(x+1) = l(x)$ и $l(x) \leqslant 1/2$ при всех x. Если $0 \leqslant x < 1$, то $r(x) = x^2/2 + x^2/4 + \cdots = x^2$ и $r(x+1) = (x-1)^2/2 + (x+1)^2/4 + (x+1)^2/8 + \cdots = x^2 + 1$. Следовательно, $f(x+1) = f(x) + 1$, если $0 \leqslant x < 1$.

Теперь по индукции можно доказать, что $f(x+n) = f(x)+n$ при всех целых $n \geqslant 0$, когда $0 \leqslant x < 1$. В частности, $f(n) = n$. Поэтому в общем случае $f(x) = 2^{-m}f(2^m x) = 2^{-m}\lfloor 2^m x\rfloor + 2^{-m}f(\{2^m x\})$. Но $f(\{2^m x\}) = l(\{2^m x\}) + r(\{2^m x\}) \leqslant \frac{1}{2} + 1$; так что $\bigl|f(x) - x\bigr| \leqslant \bigl|2^{-m}\lfloor 2^m x\rfloor - x\bigr| + 2^{-m}\cdot\frac{3}{2} \leqslant 2^{-m}\cdot\frac{5}{2}$ при любом целом m.

Неизбежный вывод заключается в том, что $f(x) = |x|$ для всех действительных x.

3.33 Пусть $r = n - \frac{1}{2}$ — радиус окружности. (a) Между клетками доски $2n-1$ горизонтальных и $2n-1$ вертикальных линий, причем окружность пересекает каждую из них дважды. Поскольку r^2 — не целое число, теорема Пифагора утверждает, что окружность не проходит ни через один угол клетки. Следовательно, окружность проходит через столько клеток, сколько имеется ее пересечений с прямыми, а именно $8n - 4 = 8r$. (Эта же формула дает количество клеток по краям доски.) (б) $f(n,k) = 4\lfloor\sqrt{r^2-k^2}\rfloor$.

Из (а) и (б) следует, что

$$\tfrac{1}{4}\pi r^2 - 2r \leqslant \sum_{0<k<r}\lfloor\sqrt{r^2-k^2}\rfloor \leqslant \tfrac{1}{4}\pi r^2, \qquad r = n - \tfrac{1}{2}.$$

Задача получения более точных оценок этой суммы представляет собой знаменитую проблему теории чисел, которой занимались Гаусс (Gauss) и многие другие; см. Диксон (Dickson) [78, т. 2, гл. 6].

3.34 (а) Пусть $m = \lceil \lg n\rceil$. Можно добавить еще $2^m - n$ членов для упрощения вычислений в граничных точках:

$$\begin{aligned} f(n) + (2^m - n)m &= \sum_{k=1}^{2^m}\lceil \lg k\rceil = \sum_{j,k} j[j = \lceil \lg k\rceil][1 \leqslant k \leqslant 2^m] = \\ &= \sum_{j,k} j[2^{j-1} < k \leqslant 2^j][1 \leqslant j \leqslant m] = \\ &= \sum_{j=1}^{m} j\,2^{j-1} = 2^m(m-1) + 1. \end{aligned}$$

Следовательно, $f(n) = nm - 2^m + 1$.

(б) Имеем $\lceil n/2 \rceil = \lfloor (n+1)/2 \rfloor$, откуда следует, что решение рекуррентного соотношения $g(n) = a(n) + g(\lceil n/2 \rceil) + g(\lfloor n/2 \rfloor)$ должно удовлетворять соотношению $\Delta g(n) = \Delta a(n) + \Delta g(\lfloor n/2 \rfloor)$. В частности, если $a(n) = n - 1$, то соотношению $\Delta f(n) = 1 + \Delta f(\lfloor n/2 \rfloor)$ удовлетворяет количество битов в двоичном представлении n, а именно $\lceil \lg(n+1) \rceil$. Теперь осталось перейти от Δ к Σ.

Более прямое решение может быть основано на соотношениях
$\lceil \lg 2j \rceil = \lceil \lg j \rceil + 1$ и $\lceil \lg(2j-1) \rceil = \lceil \lg j \rceil + [j > 1]$ для $j \geqslant 1$.

3.35 $(n+1)^2 n!\, e = A_n + (n+1)^2 + (n+1) + B_n$, где

$$A_n = \frac{(n+1)^2 n!}{0!} + \frac{(n+1)^2 n!}{1!} + \cdots + \frac{(n+1)^2 n!}{(n-1)!}$$

кратно n, а

$$\begin{aligned} B_n &= \frac{(n+1)^2 n!}{(n+2)!} + \frac{(n+1)^2 n!}{(n+3)!} + \cdots = \\ &= \frac{n+1}{n+2}\left(1 + \frac{1}{n+3} + \frac{1}{(n+3)(n+4)} + \cdots\right) < \\ &< \frac{n+1}{n+2}\left(1 + \frac{1}{n+3} + \frac{1}{(n+3)(n+3)} + \cdots\right) = \\ &= \frac{(n+1)(n+3)}{(n+2)^2} \end{aligned}$$

меньше 1. Следовательно, искомый ответ — $2 \bmod n$.

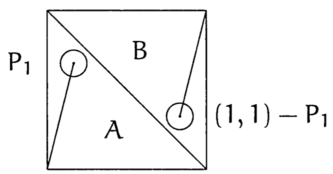
3.36 Сумма равна

$$\begin{aligned} &\sum_{k,l,m} 2^{-l} 4^{-m} \big[m = \lfloor \lg l \rfloor\big] \big[l = \lfloor \lg k \rfloor\big] [1 < k < 2^{2^n}] = \\ &\quad = \sum_{k,l,m} 2^{-l} 4^{-m} [2^m \leqslant l < 2^{m+1}][2^l \leqslant k < 2^{l+1}][0 \leqslant m < n] = \\ &\quad = \sum_{l,m} 4^{-m} [2^m \leqslant l < 2^{m+1}][0 \leqslant m < n] = \\ &\quad = \sum_{m} 2^{-m} [0 \leqslant m < n] = 2(1 - 2^{-n}). \end{aligned}$$

3.37 Сначала рассмотрите случай $m < n$, который распадается на подклассы в зависимости от того, меньше ли m, чем $\frac{1}{2}n$; затем покажите, что обе части изменяются одинаково при увеличении m на n.

3.38 Не более одного x_k может быть нецелым. Отбросим все целые x_k и предположим, что остается n чисел. Если $\{x\} \neq 0$, то среднее значение $\{mx\}$ при $m \to \infty$ лежит между $\frac{1}{4}$ и $\frac{1}{2}$; следовательно, $\{mx_1\} + \cdots + \{mx_n\} - \{mx_1 + \cdots + mx_n\}$ не может иметь нулевого среднего значения при $n > 1$.

Это доказательство основано на сложной теореме о равномерном распределении. Однако возможно и элементарное доказательство, набросок которого для $n = 2$ приведен далее. Пусть P_m — точка $\bigl(\{mx\}, \{my\}\bigr)$. Разделим единичный квадрат $0 \leqslant x, y < 1$ на треугольные области A и B, для которых $x + y < 1$ и $x + y \geqslant 1$. Мы хотим показать, что $P_m \in B$ при некотором m, если $\{x\}$ и $\{y\}$ ненулевые. Если $P_1 \in B$, то показывать больше ничего не надо. В противном случае найдется круг D радиусом $\epsilon > 0$ с центром в P_1, такой, что $D \subseteq A$. В соответствии с принципом Дирихле, если N достаточно велико, последовательность $P_1, \ldots, P_N$ должна содержать две точки с $|P_k - P_j| < \epsilon$ и $k > j$.



Отсюда следует, что P_{k-j-1} лежит в ϵ-окрестности точки $(1,1) - P_1$; следовательно, $P_{k-j-1} \in B$.

3.39 Замените j на $b - j$ и добавьте к сумме член $j = 0$, так что для вычисления суммы по j можно воспользоваться упр. 15. В результате члены

$$\lceil x/b^k \rceil - \lceil x/b^{k+1} \rceil + b - 1$$

телескопируются при суммировании по k.

3.40 Пусть $\lfloor 2\sqrt{n} \rfloor = 4k + r$, где $-2 \leqslant r < 2$, и пусть $m = \lfloor \sqrt{n} \rfloor$. Тогда по индукции можно доказать следующие соотношения.

Сегмент	r	m	x	y	тогда и только тогда, когда
W_k	-2	$2k-1$	$m(m+1)-n-k$	k	$(2k-1)(2k-1) \leqslant n \leqslant (2k-1)(2k)$
S_k	-1	$2k-1$	$-k$	$m(m+1)-n+k$	$(2k-1)(2k) < n < (2k)(2k)$
E_k	0	$2k$	$n-m(m+1)+k$	$-k$	$(2k)(2k) \leqslant n \leqslant (2k)(2k+1)$
N_k	1	$2k$	k	$n-m(m+1)-k$	$(2k)(2k+1) < n < (2k+1)(2k+1)$

Таким образом, когда $k \geqslant 1$, W_k представляет собой сегмент длиной $2k-1$ там, где путь идет на запад и $y(n) = k$; S_k представляет собой внутренность сегмента длиной $2k$ там, где путь идет на юг и $x(n) = -k$; и т.д. (а) Таким образом, искомая формула имеет вид

$$y(n) = (-1)^m\Big(\big(n - m(m+1)\big) \cdot \big[\lfloor 2\sqrt{n}\rfloor \text{ нечетное}\big] - \lceil \tfrac{1}{2}m\rceil\Big).$$

(б) На всех сегментах $k = \max\big(|x(n)|, |y(n)|\big)$. На сегментах W_k и S_k имеем $x < y$ и $n + x + y = m(m+1) = (2k)^2 - 2k$; на сегментах E_k и N_k имеем $x \geqslant y$ и $n - x - y = m(m+1) = (2k)^2 + 2k$. Следовательно, знак определяется по формуле $(-1)^{[x(n)<y(n)]}$.

3.41 Поскольку $1/\phi + 1/\phi^2 = 1$, указанные последовательности действительно разбивают все положительные целые числа. Поскольку условие $g(n) = f\big(f(n)\big) + 1$ определяет f и g единственным образом, нам надо только показать, что $\big\lfloor\lfloor n\phi\rfloor\phi\big\rfloor + 1 = \lfloor n\phi^2\rfloor$ для всех $n > 0$. Это следует из упр. 3 при $\alpha = \phi$ и $n = 1$.

3.42 Нет. Рассуждение, подобное приведенному в тексте и в упр. 13, показывает, что разбиение на три части имеет место тогда и только тогда, когда $1/\alpha + 1/\beta + 1/\gamma = 1$ и

$$\left\{\frac{n+1}{\alpha}\right\} + \left\{\frac{n+1}{\beta}\right\} + \left\{\frac{n+1}{\gamma}\right\} = 1$$

для всех $n > 0$. Но по теореме о равномерном распределении среднее значение последовательности $\big\{(n+1)/\alpha\big\}$ равно $1/2$, если α иррационально. Все параметры не могут быть рациональными, а если $\gamma = m/n$, то среднее значение равно $3/2 - 1/(2n)$. Следовательно, γ должно быть целым числом, но это тоже не годится. (Имеется также доказательство невозможности, использующее только простые принципы, без привлечения теоремы о равномерном распределении; см. [155].)

3.43 Один шаг развертывания рекуррентности для чисел K_n дает минимум из четырех чисел $1 + a + a \cdot b \cdot K_{\lfloor (n-1-a)/(a \cdot b)\rfloor}$, где a и b могут быть равны либо 2, либо 3 каждое. (Это упрощение требует применения правила (3.11) для удаления полов внутри полов вместе с тождеством $x + \min(y, z) = \min(x+y, x+z)$. Следует опустить члены с отрицательными индексами, т.е. те, у которых $n - 1 - a < 0$.)

Продолжая в том же духе, мы приходим к следующей интерпретации: K_n — наименьшее число $> n$ в мультимножестве S всех чисел вида

$$1 + a_1 + a_1a_2 + a_1a_2a_3 + \cdots + a_1a_2a_3\ldots a_m,$$

где $m \geqslant 0$ и каждое a_k равно 2 либо 3. Так, в мультимножестве

$$S = \{1,3,4,7,9,10,13,15,19,21,22,27,28,31,31,\ldots\}$$

число 31 встречается дважды, поскольку оно имеет два представления $1+2+4+8+16 = 1+3+9+18$. (Кстати, Майкл Фредман (Michael Fredman) [134] показал, что $\lim_{n\to\infty} K_n/n = 1$, т.е. между элементами S нет слишком больших промежутков.)

3.44 Пусть $d_n^{(q)} = D_{n-1}^{(q)}\,\mathrm{mumble}(q-1)$, так что $D_n^{(q)} = (qD_{n-1}^{(q)} + d_n^{(q)})/(q-1)$ и $a_n^{(q)} = \lceil D_{n-1}^{(q)}/(q-1)\rceil$. Тогда $D_{k-1}^{(q)} \leqslant (q-1)n \iff a_k^{(q)} \leqslant n$, откуда вытекает искомый результат. (Это решение найдено Эйлером (Euler) [116], который устанавливал a и d последовательно, не догадываясь, что можно обойтись одной последовательностью $D_n^{(q)}$.)

3.45 Пусть $\alpha > 1$ удовлетворяет условию $\alpha + 1/\alpha = 2m$. Тогда находим, что $2Y_n = \alpha^{2^n} + \alpha^{-2^n}$, и отсюда вытекает, что $Y_n = \lceil \alpha^{2^n}/2\rceil$.

Слишком просто...

3.46 Данное указание вытекает из (3.9), поскольку $2n(n+1) = \lfloor 2(n+\frac{1}{2})^2\rfloor$. Пусть $n+\theta = (\sqrt{2}^{\,l} + \sqrt{2}^{\,l-1})m$ и $n'+\theta' = (\sqrt{2}^{\,l+1} + \sqrt{2}^{\,l})m$, где $0 \leqslant \theta, \theta' < 1$. Тогда $\theta' = 2\theta \bmod 1 = 2\theta - d$, где d равно 0 или 1. Мы хотим доказать, что $n' = \lfloor \sqrt{2}(n+\frac{1}{2})\rfloor$; это равенство выполняется тогда и только тогда, когда

$$0 \leqslant \theta'(2-\sqrt{2}) + \sqrt{2}(1-d) < 2.$$

Что касается решения рекуррентного соотношения, то обратите внимание, что $\mathrm{Spec}(1+1/\sqrt{2})$ и $\mathrm{Spec}(1+\sqrt{2})$ образуют разбиение положительных целых чисел; значит, любое положительное целое число a может быть единственным образом записано в виде $a = \lfloor(\sqrt{2}^{\,l} + \sqrt{2}^{\,l-1})m\rfloor$, где l и m — целые числа, причем m нечетное, а $l \geqslant 0$. Отсюда вытекает, что $L_n = \lfloor(\sqrt{2}^{\,l+n} + \sqrt{2}^{\,l+n-1})m\rfloor$.

3.47 (а) $c = -\frac{1}{2}$. (б) c — целое. (в) $c = 0$. (г) c — произвольное число. В ответе к упр. 1.2.4–40 в [207] получены более общие результаты.

3.48 Пусть $x^{:0} = 1$ и $x^{:(k+1)} = x\lfloor x^{:k}\rfloor$; а также пусть $a_k = \{x^{:k}\}$ и $b_k = \lfloor x^{:k}\rfloor$. Тогда указанное тождество принимает вид $x^3 = 3x^{:3} + 3a_1a_2 + a_1^3 - 3b_1b_2 + b_1^3$. Поскольку $a_k + b_k = x^{:k} = xb_{k-1}$ для $k \geqslant 0$, имеем $(1-xz)(1+b_1z+b_2z^2+\cdots) = 1 - a_1z - a_2z^2 - \cdots$.

Таким образом,

$$\frac{1}{1-xz} = \frac{1+b_1z+b_2z^2+\cdots}{1-a_1z-a_2z^2-\cdots}.$$

Возьмем логарифм от обеих частей, чтобы отделить a от b. Затем продифференцируем по z и получим

$$\frac{x}{1-xz} = \frac{a_1+2a_2z+3a_3z^2+\cdots}{1-a_1z-a_2z^2-\cdots} + \frac{b_1+2b_2z+3b_3z^2+\cdots}{1+b_1z+b_2z^2+\cdots}.$$

Коэффициент при z^{n-1} в левой части равен x^n; в правой части для этого коэффициента имеется формула, совпадающая с нашим тождеством при $n=3$.

Можно вывести аналогичные тождества и для более общего произведения $x_0x_1\ldots x_{n-1}$ [170].

Более интересная (и еще не решенная) задача: выяснить, когда заданное мультимножество определяет неупорядоченную пару $\{\alpha,\beta\}$, если и α, и β меньше 1.

3.49 (Решение Генриха Роллетчека (Heinrich Rolletschek).) Можно заменить (α,β) на $(\{\beta\},\alpha+\lfloor\beta\rfloor)$ без изменения $\lfloor n\alpha\rfloor+\lfloor n\beta\rfloor$. Следовательно, условие $\alpha=\{\beta\}$ является необходимым. Оно же и достаточно: пусть $m=\lfloor\beta\rfloor$ — наименьший элемент данного мультимножества и пусть S — мультимножество, полученное из заданного вычитанием mn из n-го в порядке возрастания элемента при всех n. Если $\alpha=\{\beta\}$, то последовательные элементы S отличаются либо на 0, либо на 2; следовательно, мультимножество $\frac{1}{2}S=\mathrm{Spec}(\alpha)$ определяет α.

3.50 Согласно неопубликованным заметкам Уильяма О. Вича (William A. Veech), достаточно, чтобы $\alpha\beta$, β и 1 были линейно независимы над множеством рациональных чисел.

3.51 Как замечает Г. С. Вильф (H. S. Wilf), функциональное уравнение $f(x^2-1)=f(x)^2$ определяло бы $f(x)$ для всех $x\geqslant\phi$, если бы функция $f(x)$ была задана на произвольном интервале $(\phi\,..\,\phi+\epsilon)$.

3.52 Существует бесконечное количество способов разбиения всех положительных целых чисел на три или более обобщенных спектров при *иррациональных* α_k, например

$$\mathrm{Spec}(2\alpha;0)\cup\mathrm{Spec}(4\alpha;-\alpha)\cup\mathrm{Spec}(4\alpha;-3\alpha)\cup\mathrm{Spec}(\beta;0).$$

Но в точном смысле все эти разбиения возникают в результате “расширения” базового $\mathrm{Spec}(\alpha)\cup\mathrm{Spec}(\beta)$; см. [158]. Все известные разбиения с рациональными α_k, например

$$\mathrm{Spec}(7;-3)\cup\mathrm{Spec}(\tfrac{7}{2};-1)\cup\mathrm{Spec}(\tfrac{7}{4};0),$$

используют параметры, как в сформулированном предположении, которое принадлежит А. С. Френкелю (A. S. Fraenkel) [128].

3.53 Частичные результаты рассматриваются в [95, с. 30–31]. Жадный алгоритм, вероятно, не завершается.

4.1 1, 2, 4, 6, 16, 12.

4.2 Заметим, что $m_p+n_p=\min(m_p,n_p)+\max(m_p,n_p)$. Рекуррентное соотношение $\text{НОК}(m,n) = \left(\frac{n}{n \bmod m}\right)\text{НОК}(n \bmod m, m)$ корректно, но на практике для вычисления НОК непригодно; лучший известный способ вычисления $\text{НОК}(m,n)$ состоит в вычислении $\text{НОД}(m,n)$ и делении mn на полученное значение.

4.3 Это справедливо для целого значения x, но $\pi(x)$ определена для действительных чисел x. Правильная формула —

$$\pi(x)-\pi(x-1) \;=\; \bigl[\lfloor x\rfloor — \text{простое число}\bigr]\,;$$

в ее корректности очень легко убедиться.

4.4 Между дробями $\frac{1}{0}$ и $\frac{0}{-1}$ оказалось бы отраженное слева направо дерево Штерна–Броко с отрицательными знаменателями у всех дробей и т.д. Так что получившееся в результате дерево содержало бы *все* дроби m/n с $m \perp n$. Условие $m'n - mn' = 1$ в процессе построения остается справедливым. (Такая конструкция называется *венком Штерна–Броко*, поскольку для удобства последнюю дробь $\frac{0}{1}$ можно считать идентичной первой дроби $\frac{0}{1}$, соединяя тем самым наверху деревья в замкнутый круг. Венки Штерна–Броко находят интересные применения в машинной графике, поскольку с их помощью можно представить любое рациональное направление на плоскости.)

4.5 $L^k = \left(\begin{smallmatrix}1 & k\\ 0 & 1\end{smallmatrix}\right)$ и $R^k = \left(\begin{smallmatrix}1 & 0\\ k & 1\end{smallmatrix}\right)$; это справедливо даже при $k<0$. (Общая формула для любого произведения матриц L и R будет найдена в главе 6.)

4.6 $a=b$. (В главе 3 определено $x \bmod 0 = x$, в первую очередь для того, чтобы это было именно так.)

4.7 Для этого необходимо, чтобы $m \bmod 10 = 0$, $m \bmod 9 = k$ и $m \bmod 8 = 1$. Но m не может быть одновременно и четным, и нечетным.

4.8 Надо, чтобы $10x+6y \equiv 10x+y \pmod{15}$; следовательно, $5y \equiv 0 \pmod{15}$; откуда $y \equiv 0 \pmod 3$. Должно быть $y=0$ или 3 и $x=0$ или 1.

4.9 $3^{2k+1} \bmod 4 = 3$, так что $(3^{2k+1}-1)/2$ нечетное. Указанное число делится на $(3^7-1)/2$ и $(3^{11}-1)/2$ (и на другие числа).

4.10 $999(1-\frac{1}{3})(1-\frac{1}{37}) = 648$.

4.11 $\sigma(0) = 1$, $\sigma(1) = -1$, $\sigma(n) = 0$ при $n > 1$. (Обобщенные функции Мёбиуса, определенные на произвольных частично упорядоченных множествах, обладают интересными и важными свойствами, впервые установленными Вейснером (Weisner) [366] и развитыми многими другими, в особенности Джан-Карло Рота (Gian-Carlo Rota) [313].)

4.12 $\sum_{d\backslash m}\sum_{k\backslash d}\mu(d/k)\,g(k) = \sum_{k\backslash m}\sum_{d\backslash(m/k)}\mu(d)\,g(k) = \sum_{k\backslash m} g(k)[m/k=1] = g(m)$ согласно (4.7) и (4.9).

4.13 (а) $n_p \leqslant 1$ при всех p; (б) $\mu(n) \neq 0$.

4.14 Истинно при $k > 0$. Используйте (4.12), (4.14) и (4.15).

4.15 Нет. Например, $e_n \bmod 5 = [2\,\text{или}\,3]$; $e_n \bmod 11 = [2, 3, 7$ или $10]$.

4.16 $1/e_1+1/e_2+\cdots+1/e_n = 1-1/\bigl(e_n(e_n-1)\bigr) = 1-1/(e_{n+1}-1)$.

4.17 Имеем $f_n \bmod f_m{=}2$; так что НОД(f_n, f_m)=НОД$(2, f_m)$=1. (Кстати, соотношение $f_n = f_0 f_1 \ldots f_{n-1} + 2$ очень похоже на рекуррентное соотношение, определяющее числа Евклида e_n.)

4.18 Если $n = qm$ и q нечетное, то $2^n + 1 = (2^m + 1)(2^{n-m} - 2^{n-2m} + \cdots - 2^m + 1)$.

4.19 Первая сумма равна $\pi(n)$, поскольку ее общий член — $[k+1$ — простое число$]$. Внутренняя сумма во втором равенстве представляет собой $\sum_{1\leqslant k<m}[k\backslash m]$, так что она больше 1 тогда и только тогда, когда m — составное число; и мы снова получаем $\pi(n)$. Наконец, $\lceil\{m/n\}\rceil = [n\nmid m]$, так что третья сумма представляет собой приложение теоремы Вильсона. Но, конечно же, вычисление $\pi(n)$ по любой из приведенных формул — безумие.

4.20 Пусть $p_1 = 2$ и пусть p_n — наименьшее простое число, большее, чем $2^{p_{n-1}}$. Тогда $2^{p_{n-1}} < p_n < 2^{p_{n-1}+1}$, откуда следует, что можно взять $b = \lim_{n\to\infty} \lg^{(n)} p_n$, где $\lg^{(n)}$ представляет собой функцию lg, итерированную n раз. Указанное числовое значение получается из $p_2 = 5$ и $p_3 = 37$. Оказывается, что $p_4 = 2^{37} + 9$, что дает более точное значение

$$b \approx 1.2516475977905$$

(но не дает ключ к значению p_5).

4.21 В силу постулата Бертрана $P_n < 10^n$. Пусть

$$K = \sum_{k\geqslant 1} 10^{-k^2} P_k = .200300005\ldots$$

Тогда $10^{n^2}K \equiv P_n +$ дробь $(\mathrm{mod}\ 10^{2n-1})$.

4.22 $(b^{mn}-1)/(b-1) = \big((b^m-1)/(b-1)\big)(b^{mn-m}+\cdots+1)$. (Числа вида $(10^p-1)/9$ являются простыми при $p < 49081$ только при значениях $p = 2$, 19, 23, 317, 1031.) Числа такого вида носят название "повторни́цы" ("repunits").

4.23 $\rho(2k+1) = 0$ и $\rho(2k) = \rho(k)+1$ при $k \geqslant 1$. Можно показать по индукции, что $\rho(n) = \rho(n-2^m)$, если $n > 2^m$ и $m > \rho(n)$. На k-м шаге перекладывается $\rho(k)$-й диск башни, если перенумеровать диски как $0, 1, \ldots, n-1$. Если k представляет собой степень 2, то понятно, что это так. А если $2^m < k < 2^{m+1}$, то $\rho(k) < m$; k-е и $k-2^m$-е перемещения соответствуют друг другу в последовательности перекладываний, перемещающей $m+1$ дисков за $T_m + 1 + T_m$ шагов.

4.24 Цифра, которая вносит вклад dp^m в n, вносит в $\epsilon_p(n!)$ вклад $dp^{m-1}+\cdots+d = d(p^m-1)/(p-1)$, следовательно, $\epsilon_p(n!) = \big(n-\nu_p(n)\big)/(p-1)$.

4.25 При всех p справедливо $m\backslash\backslash n \iff m_p = 0$ или $m_p = n_p$. Отсюда следует, что (а) истинно. Однако (б) ложно — например, для нашего любимого примера $m = 12$ и $n = 18$. (Это распространенное заблуждение.)

4.26 Да, поскольку последовательность $\mathcal{G}_N$ определяет поддерево дерева Штерна–Броко.

4.27 Дополните более короткую строку символами M (поскольку M лежит в алфавите между L и R), пока обе строки не станут одинаковой длины, а затем сравнивайте строки в алфавитном порядке. Например, верхние уровни дерева упорядочены следующим образом: LL < LM < LR < MM < RL < RM < RR. (Другое решение заключается в добавлении к строкам бесконечной строки RL^∞ и сравнении до нахождения L < R.)

4.28 Нам надо воспользоваться только первой частью этого представления:

$$\begin{array}{cccccccccccccccc} \mathrm{R} & \mathrm{R} & \mathrm{R} & \mathrm{L} & \mathrm{L} & \mathrm{L} & \mathrm{L} & \mathrm{L} & \mathrm{L} & \mathrm{L} & \mathrm{R} & \mathrm{R} & \mathrm{R} & \mathrm{R} & \mathrm{R} & \mathrm{R} \\ \frac{1}{1}, & \frac{2}{1}, & \frac{3}{1}, & \frac{4}{1}, & \frac{7}{2}, & \frac{10}{3}, & \frac{13}{4}, & \frac{16}{5}, & \frac{19}{6}, & \frac{22}{7}, & \frac{25}{8}, & \frac{47}{15}, & \frac{69}{22}, & \frac{91}{29}, & \frac{113}{36}, & \frac{135}{43}, \ldots \end{array}$$

Дробь $\frac{4}{1}$ имеется здесь потому, что она оценивает значение π сверху лучше, чем $\frac{1}{0}$, а не потому, что она ближе, чем $\frac{3}{1}$. Аналогично $\frac{25}{8}$ является лучшей нижней границей, чем $\frac{3}{1}$. Здесь наличествуют все простейшие верхние и нижние границы, но очередное действительно хорошее приближение не встретится до тех пор, пока строка символов R не перейдет обратно в строку символов L.

4.29 $1/\alpha$. Чтобы получить $1-x$ из x в двоичной записи, мы заменяем все 0 на 1, и наоборот; чтобы получить $1/\alpha$ из α в записи Штерна–Броко, мы заменяем все L на R, и наоборот. (Следует также рассмотреть случаи с конечным количеством символов, но они также должны работать, поскольку данное соответствие сохраняет упорядочение.)

4.30 Поскольку m целых чисел $x \in [A\,..\,A+m)$ различны по модулю m, то их остатки $(x \bmod m_1, \ldots, x \bmod m_r)$ пробегают все $m_1 \ldots m_r = m$ возможных значений, одно из которых должно быть равно $(a_1 \bmod m_1, \ldots, a_r \bmod m_r)$ в силу принципа "голубиных гнезд".

4.31 Число, записанное в системе счисления с основанием b, делится на d тогда и только тогда, когда сумма его цифр делится на d, при условии, что $b \equiv 1 \pmod d$. Это следует из того, что $(a_m \ldots a_0)_b = a_m b^m + \cdots + a_0 b^0 \equiv a_m + \cdots + a_0$.

4.32 Система $\varphi(m)$ чисел $\{\, kn \bmod m \mid k \perp m \text{ и } 0 \leqslant k < m \,\}$ представляет собой систему чисел $\{\, k \mid k \perp m \text{ и } 0 \leqslant k < m \,\}$ в некотором порядке. Перемножьте их и поделите на $\prod_{0 \leqslant k < m,\, k \perp m} k$.

4.33 Очевидно, что $h(1) = 1$. Если $m \perp n$, то

$$h(mn) = \sum_{d \backslash mn} f(d)\, g(mn/d) = \sum_{c \backslash m,\, d \backslash n} f(cd)\, g\bigl((m/c)(n/d)\bigr) =$$
$$= \sum_{c \backslash m} \sum_{d \backslash n} f(c)\, g(m/c)\, f(d)\, g(n/d);$$

это равно $h(m)\,h(n)$, так как в каждом члене этой суммы $c \perp d$.

4.34 $g(m) = \sum_{d \backslash m} f(d) = \sum_{d \backslash m} f(m/d) = \sum_{d \geqslant 1} f(m/d)$, если функция $f(x)$ равна нулю при нецелом значении x.

4.35 Начальные условия представляют собой

$$I(0,n) = 0; \qquad I(m,0) = 1.$$

При $m, n > 0$ имеются два правила, первое из которых тривиально при $m > n$, а второе — при $m < n$:

$$I(m,n) = I(m, n \bmod m) - \lfloor n/m \rfloor I(n \bmod m, m);$$
$$I(m,n) = I(m \bmod n, n).$$

4.36 Разложение любой из заданных величин на необратимые множители должно содержать $m^2 - 10n^2 = \pm 2$ или ± 3, но это невозможно по модулю 10.

4.37 Пусть $a_n = 2^{-n}\ln(e_n - \frac{1}{2})$ и $b_n = 2^{-n}\ln(e_n + \frac{1}{2})$. Тогда

$$e_n = \lfloor E^{2^n} + \tfrac{1}{2} \rfloor \iff a_n \leqslant \ln E < b_n .$$

Но $a_{n-1} < a_n < b_n < b_{n-1}$, так что можно взять $E = \lim_{n\to\infty} e^{a_n}$. На самом деле оказывается, что

$$E^2 = \frac{3}{2}\prod_{n\geqslant 1}\left(1 + \frac{1}{(2e_n - 1)^2}\right)^{1/2^n}$$

представляет собой произведение, которое быстро сходится к значению $(1.26408473530530111\ldots)^2$. Но эти наблюдения не позволяют выяснить, чему равно число e_n, если только мы не сможем найти другое выражение для E, не зависящее от чисел Евклида.

4.38 Пусть $r = n \bmod m$. Тогда

$$a^n - b^n = (a^m - b^m)(a^{n-m}b^0 + a^{n-2m}b^m + \cdots + a^r b^{n-m-r}) + \\ + b^{m\lfloor n/m \rfloor}(a^r - b^r).$$

4.39 Если $a_1 \ldots a_t$ и $b_1 \ldots b_u$ — полные квадраты, то таковым же является и число

$$a_1 \ldots a_t b_1 \ldots b_u / c_1^2 \ldots c_\nu^2 ,$$

где $\{a_1, \ldots, a_t\} \cap \{b_1, \ldots, b_u\} = \{c_1, \ldots, c_\nu\}$. (Фактически можно показать, что последовательность $\langle S(1), S(2), S(3), \ldots \rangle$ содержит каждое непростое положительное целое число ровно один раз.)

4.40 Пусть $f(n) = \prod_{1\leqslant k\leqslant n,\, p\nmid k} k = n!/p^{\lfloor n/p \rfloor}\lfloor n/p \rfloor!$ и $g(n) = n!/p^{\epsilon_p(n!)}$. Тогда

$$g(n) = f(n)f(\lfloor n/p \rfloor)f(\lfloor n/p^2 \rfloor)\ldots = f(n)g(\lfloor n/p \rfloor) .$$

Кроме того, $f(n) \equiv a_0!(p-1)!^{\lfloor n/p \rfloor} \equiv a_0!(-1)^{\lfloor n/p \rfloor} \pmod p$, и $\epsilon_p(n!) = \lfloor n/p \rfloor + \epsilon_p(\lfloor n/p \rfloor!)$. Эти рекуррентные соотношения позволяют легко доказать требуемое по индукции. (Возможны и иные решения.)

4.41 (а) Если $n^2 \equiv -1 \pmod p$, то $(n^2)^{(p-1)/2} \equiv -1$; но теорема Ферма гласит, что это — $+1$.
(б) Пусть $n = ((p-1)/2)!$; тогда $n \equiv (-1)^{(p-1)/2}\prod_{1\leqslant k<p/2}(p - k) = (p-1)!/n$, следовательно, $n^2 \equiv (p-1)!$.

4.42 Для начала заметим, что $k \perp l \iff k \perp l + ak$ для любого целого a, так как $\text{НОД}(k, l) = \text{НОД}(k, l + ak)$ в соответствии с алгоритмом Евклида. Итак,

$$m \perp n \text{ и } n' \perp n \iff mn' \perp n$$
$$\iff mn' + nm' \perp n.$$

Аналогично

$$m' \perp n' \text{ и } n \perp n' \iff mn' + nm' \perp n'.$$

Следовательно,

$$m \perp n \text{ и } m' \perp n' \text{ и } n \perp n' \iff mn' + nm' \perp nn'.$$

4.43 Надо выполнить умножение на $L^{-1}R$, затем на $R^{-1}L^{-1}RL$, затем на $L^{-1}R$, затем на $R^{-2}L^{-1}RL^2$ и т.д.; n-й множитель равен $R^{-\rho(n)}L^{-1}RL^{\rho(n)}$, поскольку мы должны сократить $\rho(n)$ R. А $R^{-m}L^{-1}RL^m = \binom{0\ \ -1}{1\ 2m+1}$.

John .316

Такой лозунг можно было видеть на чемпионате 1993 года, когда в игру вступил Джон Крук.

4.44 Можно найти простейшее рациональное число, лежащее в интервале

$$[0.3155\,..\,0.3165) = \left[\tfrac{631}{2000}\,..\,\tfrac{633}{2000}\right),$$

просматривая представления Штерна–Броко дробей $\frac{631}{2000}$ и $\frac{633}{2000}$ и останавливаясь непосредственно перед тем, как в первой из них встречается символ L, а во второй — символ R:

$(m_1, n_1, m_2, n_2) := (631, 2000, 633, 2000)$;
пока $m_1 > n_1$ **или** $m_2 < n_2$ **выполнять**
  если $m_2 < n_2$ **то** (вывод(L); $(n_1, n_2) := (n_1, n_2) - (m_1, m_2)$)
  иначе (вывод(R); $(m_1, m_2) := (m_1, m_2) - (n_1, n_2)$).

Искомый результат — $LLLRRRRR = \frac{6}{19} \approx .3158$. Кстати, коэффициент .334 требует по меньшей мере 287 ударов.

Зато коэффициент .333 — всего трех. Но все это мелочь по сравнению с 667 ударами, необходимыми для достижения коэффициента .999 (или .001).
— Переводчик

(Приведенный алгоритм дает неверный результат, когда простейшей дробью оказывается правая граница интервала (рассмотрите, например, значения 0.187 и 0.188 — для них приведенный алгоритм даст одно и то же значение LLLLLRR (3/16), в то время как для 0.187 правильным ответом является LLLLLRRLRRR (14/75)). Для получения корректного ответа следует добавить к строке представления правой границы бесконечную строку LRRR… (так как правая граница *не* может быть решением). Провести коррекцию алгоритма несложно — достаточно в двух местах заменить сравнение $m_2 < n_2$ на $m_2 \leqslant n_2$. — *Переводчик*.)

4.45 $x^2 \equiv x \pmod{10^n} \iff x(x-1) \equiv 0 \pmod{2^n}$ и $x(x-1) \equiv 0 \pmod{5^n} \iff x \bmod 2^n = [0 \text{ или } 1]$ и $x \bmod 5^n = [0 \text{ или } 1]$. (Последний шаг оправдан, поскольку $x(x-1) \bmod 5 = 0$ означает, что либо x, либо $x-1$ кратно 5; в этом случае другой множитель взаимно прост с 5^n и может быть выделен из данного сравнения.)

Так что имеется не более четырех решений, ($x = 0$ и $x = 1$) нельзя рассматривать как "n-значное число", кроме разве что случая $n = 1$. Два другие решения имеют вид x и $10^n + 1 - x$, и как минимум одно из этих чисел $\geqslant 10^{n-1}$. При $n = 4$ вторым решением является *трех*значное число $10001 - 9376 = 625$. Мы ожидаем, что два n-значных решения будут примерно у 80% всех n, но это предположение еще не доказано.

(Подобные самовоспроизводящиеся числа называются автоморфными.)

4.46 (а) Если $j'j - k'k = \text{НОД}(j,k)$, то $n^{k'k} n^{\text{НОД}(j,k)} = n^{j'j} \equiv 1$ и $n^{k'k} \equiv 1$. (б) Пусть $n = pq$, где p — наименьший простой делитель числа n. Если $2^n \equiv 1 \pmod n$, то $2^n \equiv 1 \pmod p$. Кроме того, $2^{p-1} \equiv 1 \pmod p$; следовательно, $2^{\text{НОД}(p-1,n)} \equiv 1 \pmod p$. Но $\text{НОД}(p-1,n) = 1$ по определению p.

4.47 Если $n^{m-1} \equiv 1 \pmod m$, то должно быть $n \perp m$. Если $n^k \equiv n^j$ при некоторых $1 \leqslant j < k < m$, то $n^{k-j} \equiv 1$, так как мы можем выполнить деление на n^j. Таким образом, если не все числа $n^1 \bmod m, \ldots, n^{m-1} \bmod m$ различны, существует $k < m-1$, для которого $n^k \equiv 1$. Наименьшее такое k делит $m-1$ согласно упр. 46, а. Но тогда $kq = (m-1)/p$ при некотором простом p и некотором положительном целом числе q; но это невозможно, так как $n^{kq} \not\equiv 1$. Поэтому все числа $n^1 \bmod m, \ldots, n^{m-1} \bmod m$ различны и взаимно просты с m. Таким образом, числа $1, \ldots, m-1$ взаимно просты с m, и m должно быть простым.

4.48 Попарно объединяя числа с обратными к ним, можно свести произведение $\pmod m$ к $\prod_{1 \leqslant n < m,\, n^2 \bmod m = 1} n$. Теперь можно воспользоваться нашим умением решать уравнения $n^2 \bmod m = 1$. С помощью модулярной арифметики выясняется, что искомый результат равен $m-1$, если $m = 4$, p^k или $2p^k$ $(p > 2)$; в противном случае он равен $+1$.

4.49 (а) Либо $m < n$ ($\Phi(N) - 1$ случаев), либо $m = n$ (один случай), либо $m > n$ (вновь $\Phi(N) - 1$ случаев). Следовательно, $R(N) = 2\Phi(N) - 1$. (б) Из (4.62) получаем

$$2\Phi(N) - 1 = -1 + \sum_{d \geqslant 1} \mu(d) \lfloor N/d \rfloor \lfloor 1 + N/d \rfloor ;$$

следовательно, требуемый результат справедлив тогда и только тогда, когда

$$\sum_{d\geqslant 1} \mu(d)\lfloor N/d\rfloor = 1 \qquad \text{при } N \geqslant 1.$$

А это — частный случай (4.61), если положить $f(x) = [x \geqslant 1]$.

4.50 (а) Если f — некоторая функция, то

$$\begin{aligned}\sum_{0\leqslant k<m} f(k) &= \sum_{d\backslash m}\ \sum_{0\leqslant k<m} f(k)\bigl[d = \text{НОД}(k,m)\bigr] = \\ &= \sum_{d\backslash m}\ \sum_{0\leqslant k<m} f(k)\bigl[k/d \perp m/d\bigr] = \\ &= \sum_{d\backslash m}\ \sum_{0\leqslant k<m/d} f(kd)\bigl[k \perp m/d\bigr] = \\ &= \sum_{d\backslash m}\ \sum_{0\leqslant k<d} f(km/d)\bigl[k \perp d\bigr].\end{aligned}$$

С частным случаем этого вывода мы знакомы по выводу формулы (4.63). Аналогичный вывод выполняется и при замене $\sum$ на $\prod$. Таким образом, имеем

$$z^m - 1 = \prod_{0\leqslant k<m} (z-\omega^k) = \prod_{d\backslash m}\ \prod_{\substack{0\leqslant k<d\\ k\perp d}} (z-\omega^{km/d}) = \prod_{d\backslash m} \Psi_d(z),$$

потому что $\omega^{m/d} = e^{2\pi i/d}$.

Часть (б) следует из части (а) по аналогии с принципом (4.56) для произведений вместо сумм. Кстати, эта формула показывает, что $\Psi_m(z)$ имеет целые коэффициенты, поскольку $\Psi_m(z)$ получается путем умножения и деления полиномов со старшим коэффициентом 1.

4.51 $(x_1+\cdots+x_n)^p = \sum_{k_1+\cdots+k_n=p} p!/(k_1!\ldots k_n!)x_1^{k_1}\ldots x_n^{k_n}$, и коэффициенты делятся на p, за исключением ситуации, когда некоторое $k_j = p$. Следовательно, $(x_1+\cdots+x_n)^p \equiv x_1^p+\cdots+x_n^p \pmod p$. Теперь можно положить все x равными 1 и получить $n^p \equiv n$.

4.52 Если $p > n$, то и доказывать нечего. В противном случае $x \perp p$, так что $x^{k(p-1)} \equiv 1 \pmod p$; это означает, что по меньшей мере $\lfloor (n-1)/(p-1)\rfloor$ данных чисел кратны p. А $(n-1)/(p-1) \geqslant n/p$, поскольку $n \geqslant p$.

4.53 Вначале покажите, что если $m \geqslant 6$ и m — не простое число, то $(m-2)! \equiv 0 \pmod{m}$. (Если $m = p^2$, то произведение для $(m-2)!$ включает p и $2p$; в противном случае оно включает d и m/d, где $d < m/d$.) Затем рассмотрите следующие случаи.

"Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk".
— Кронекер (Kronecker) [365]

Случай 0, $n < 5$. Данное условие выполняется только при $n = 1$.

Случай 1, $n \geqslant 5$ и n — простое. Тогда $(n-1)!/(n+1)$ — целое число и оно не может быть кратно n.

Случай 2, $n \geqslant 5$, n — составное и $n+1$ — составное. Тогда n и $n+1$ делят $(n-1)!$, и $n \perp n+1$; следовательно, $n(n+1)\backslash(n-1)!$.

Случай 3, $n \geqslant 5$, n — составное, а $n+1$ — простое. Тогда $(n-1)! \equiv 1 \pmod{n+1}$ по теореме Вильсона, и

$$\lceil (n-1)!/(n+1) \rceil \;=\; \bigl((n-1)!+n\bigr)/(n+1)\,;$$

а это число делится на n.

Значит, ответ таков: либо $n = 1$, либо $n \neq 4$ — составное число.

4.54 $\epsilon_2(1000!) > 500$ и $\epsilon_5(1000!) = 249$, следовательно, $1000! = a \cdot 10^{249}$ при некотором целом четном a. Поскольку $1000 = (13000)_5$, упр. 40 гласит, что $a \cdot 2^{249} = 1000!/5^{249} \equiv -1 \pmod 5$. Кроме того, $2^{249} \equiv 2$, следовательно, $a \equiv 2$, а значит, $a \bmod 10 = 2$ или 7, так что окончательный ответ — $2 \cdot 10^{249}$.

4.55 Один из способов доказательства — доказательство по индукции того факта, что $P_{2n}/P_n^4(n+1)$ — целое число. Этот более сильный результат помогает в дальнейшем выполнении индукции. Другой способ основан на доказательстве того, что каждое простое число делит числитель по меньшей мере столько же раз, сколько оно делит и знаменатель. Этот способ сводится к доказательству неравенства

$$\sum_{k=1}^{2n} \lfloor k/m \rfloor \;\geqslant\; 4\sum_{k=1}^{n} \lfloor k/m \rfloor\,, \qquad \text{целое } m \geqslant 2,$$

которое следует из неравенства

$$\left\lfloor \frac{2n-1}{m} \right\rfloor + \left\lfloor \frac{2n}{m} \right\rfloor \;\geqslant\; 4\left\lfloor \frac{n}{m} \right\rfloor + [n \bmod m = m-1] - [n \bmod m = 0]\,.$$

Последнее неравенство истинно при $0 \leqslant n < m$, и обе его части увеличиваются на 4 при увеличении n на m.

4.56 Пусть

$$f(m) \;=\; \sum_{k=1}^{2n-1} \min(k, 2n-k)[m\backslash k],$$

$$g(m) = \sum_{k=1}^{n-1} (2n{-}2k{-}1)\bigl[m\backslash(2k+1)\bigr].$$

Количество раз, которое p делит числитель указанного в условии отношения, равно $f(p) + f(p^2) + f(p^3) + \cdots$, а число раз, которое p делит его знаменатель, равно $g(p) + g(p^2) + g(p^3) + \cdots$. Но $f(m) = g(m)$ при нечетном m согласно упр. 2.32. Таким образом, указанное отношение сводится к $2^{n(n-1)}$ согласно упр. 3.22.

4.57 Указание предполагает стандартное изменение порядка суммирования, поскольку

$$\sum_{1\leqslant m\leqslant n} [d\backslash m] = \sum_{0<k\leqslant n/d} [m=dk] = \lfloor n/d \rfloor .$$

Обозначая сумму в указании через $\Sigma(n)$, получим

$$\Sigma(m+n) - \Sigma(m) - \Sigma(n) = \sum_{d\in S(m,n)} \varphi(d) .$$

С другой стороны, из (4.54) известно, что $\Sigma(n) = \frac{1}{2}n(n+1)$. Следовательно, $\Sigma(m+n) - \Sigma(m) - \Sigma(n) = mn$.

4.58 Функция $f(m)$ мультипликативна, а когда $m = p^k$, она равна $1+p+\cdots+p^k$. Это значение представляет собой степень 2 тогда и только тогда, когда p — простое число Мерсенна и $k = 1$. Действительно, k должно быть нечетным, а в этом случае сумма равна

$$(1+p)(1+p^2+p^4+\cdots+p^{k-1})$$

и $(k-1)/2$ должно быть нечетным и т.д. Необходимое и достаточное условие состоит в том, что m является произведением различных простых чисел Мерсенна.

4.59 Доказательство указания. Если $n = 1$, то $x_1 = \alpha = 2$, так что нет никаких проблем. Если $n > 1$, можно считать, что $x_1 \leqslant \cdots \leqslant x_n$. Случай 1: $x_1^{-1}+\cdots+x_{n-1}^{-1}+(x_n-1)^{-1} \geqslant 1$ и $x_n > x_{n-1}$. Тогда можно найти $\beta \geqslant x_n-1 \geqslant x_{n-1}$, такое, что $x_1^{-1}+\cdots+x_{n-1}^{-1}+\beta^{-1} = 1$; следовательно, по индукции $x_n \leqslant \beta+1 \leqslant e_n$ и $x_1\ldots x_n \leqslant x_1\ldots x_{n-1}(\beta+1) \leqslant e_1\ldots e_n$. Существует такое положительное целое число m, что $\alpha = x_1\ldots x_n/m$; следовательно, $\alpha \leqslant e_1\ldots e_n = e_{n+1}-1$, и мы имеем $x_1\ldots x_n(\alpha+1) \leqslant e_1\ldots e_n e_{n+1}$. Случай 2: $x_1^{-1}+\cdots+x_{n-1}^{-1}+(x_n-1)^{-1} \geqslant 1$ и $x_n = x_{n-1}$. Пусть $a = x_n$ и $a^{-1}+(a-1)^{-1} = (a-2)^{-1}+\zeta^{-1}$. Тогда можно показать, что $a \geqslant 4$ и $(a-2)(\zeta+1) \geqslant a^2$. Так что существует некоторое $\beta \geqslant \zeta$, такое, что $x_1^{-1}+\cdots+x_{n-2}^{-1}+(a-2)^{-1}+\beta^{-1} = 1$; отсюда по индукции следует, что $x_1 \ldots x_n \leqslant x_1 \ldots x_{n-2}(a-2)(\zeta+1) \leqslant x_1 \ldots x_{n-2}(a-2)(\beta+1) \leqslant e_1 \ldots e_n$, и можно завершить доказательство, как и ранее. Случай 3: $x_1^{-1}+\cdots+x_{n-1}^{-1}+(x_n-1)^{-1} < 1$. Пусть $a = x_n$ и пусть $a^{-1}+\alpha^{-1} = (a-1)^{-1}+\beta^{-1}$. Можно показать, что $(a-1)(\beta+1) > a(\alpha+1)$, так как это соотношение эквивалентно соотношению

$$a\alpha^2 - a^2\alpha + a\alpha - a^2 + \alpha + a > 0,$$

которое является следствием неравенства $a\alpha(\alpha-a)+(1+a)\alpha \geqslant (1+a)\alpha > a^2 - a$. Следовательно, можно заменить x_n и α на $a-1$ и β, повторяя это преобразование до тех пор, пока не будут применимы случаи 1 или 2.

Другим следствием указания является то, что $1/x_1+\cdots+1/x_n < 1$ влечет за собой $1/x_1+\cdots+1/x_n \leqslant 1/e_1+\cdots+1/e_n$; см. упр. 16.

4.60 Суть дела в том, что $\theta < \frac{2}{3}$. Тогда можно выбрать p_1 достаточно большим (чтобы удовлетворить приведенным ниже условиям), а p_n — наименьшим простым числом, бóльшим p_{n-1}^3. При таком определении p_n положим $a_n = 3^{-n}\ln p_n$ и $b_n = 3^{-n}\ln(p_n+1)$. Если мы сможем показать, что $a_{n-1} \leqslant a_n < b_n \leqslant b_{n-1}$, то сможем взять $P = \lim_{n\to\infty} e^{a_n}$, как в упр. 37. Но это предположение эквивалентно тому, что $p_{n-1}^3 \leqslant p_n < (p_{n-1}+1)^3$. Если в этих пределах не найдется простого числа p_n, то должно быть простое число $p < p_{n-1}^3$, такое, что $p + p^\theta > (p_{n-1}+1)^3$. Но отсюда вытекает, что $p^\theta > 3p^{2/3}$, а это невозможно, когда p достаточно велико.

"Человек создал целые числа; все остальное — Дьёдонне (Dieudonné)".
— Р. К. Ги (R. K. Guy)
("От Бога", если переводить фамилию Dieudonné буквально. — Переводчик)

Можно почти наверняка взять $p_1 = 2$, поскольку все имеющиеся данные указывают на то, что известные оценки для промежутков между простыми числами много слабее по сравнению с истинными промежутками (см. упр. 69). Тогда $p_2 = 11$, $p_3 = 1361$, $p_4 = 2521008887$ и $1.306377883863 < P < 1.306377883869$.

4.61 Обозначим через $\hat{m}$ и $\hat{n}$ правые части; заметим, что $\hat{m}n'-m'\hat{n} = 1$, следовательно, $\hat{m} \perp \hat{n}$. Кроме того, $\hat{m}/\hat{n} > m'/n'$ и $N = \big((n+N)/n'\big)n'-n \geqslant \hat{n} > \big((n+N)/n'-1\big)n'-n = N-n' \geqslant 0$. Так что мы имеем $\hat{m}/\hat{n} \geqslant m''/n''$. Если же равенство не выполняется, то получаем, что $n'' = (\hat{m}n'-m'\hat{n})n'' = n'(\hat{m}n''-m''\hat{n})+\hat{n}(m''n'-m'n'') \geqslant n'+\hat{n} > N$, т.е. противоречие.

Кстати, из этого упражнения следует, что $(m+m'')/(n+n'') = m'/n'$, хотя первая из этих дробей не всегда сократима.

4.62 Это число $2^{-1}+2^{-2}+2^{-3}-2^{-6}-2^{-7}+2^{-12}+2^{-13}-2^{-20}-2^{-21}+2^{-30}+2^{-31}-2^{-42}-2^{-43}+\cdots$ можно записать как

$$\frac{1}{2}+3\sum_{k\geqslant 0}\big(2^{-4k^2-6k-3}-2^{-4k^2-10k-7}\big).$$

Кстати, эту сумму можно выразить и в аналитическом виде с использованием “тета-функции” $\theta(z,\lambda)=\sum_k e^{-\pi\lambda k^2+2izk}$:

$$e \quad\leftrightarrow\quad \tfrac12+\tfrac38\theta(\tfrac4\pi\ln 2,\, 3i\ln 2)-\tfrac{3}{128}\theta(\tfrac4\pi\ln 2,\, 5i\ln 2)\,.$$

4.63 Любое число $n>2$ либо имеет простой делитель $d>2$, либо делится на $d=4$. В любом случае разрешимость уравнения с показателем n влечет разрешимость уравнения $(a^{n/d})^d+(b^{n/d})^d=(c^{n/d})^d$ с показателем d. Поскольку при $d=4$ решения не существует, d должно быть простым числом.

Я нашел поистине чудесное доказательство большой теоремы Ферма, но поля слишком малы, чтобы привести его здесь.

Указание следует из биномиальной теоремы, поскольку $(a^p+(x-a)^p)/x\equiv pa^{p-1} \pmod x$ при нечетном p. В наименьшем контрпримере, если (4.46) неверно, то должно быть $a\perp x$. Если x не делится на p, то x взаимно простое с c^p/x; это означает, что всякий раз, когда q — простое и обладает свойствами $q^e\backslash\backslash x$ и $q^f\backslash\backslash c$, мы будем иметь $e=fp$. Следовательно, $x=m^p$ для некоторого m. С другой стороны, если x делится на p, то c^p/x делится на p, но не на p^2, и c^p не имеет других общих с x множителей.

4.64 Одинаковые дроби в последовательности $\mathcal{P}_N$ появляются в “порядке органных труб”:

$$\frac{2m}{2n},\ \frac{4m}{4n},\ \ldots,\ \frac{rm}{rn},\ \ldots,\ \frac{3m}{3n},\ \frac{m}{n}\,.$$

Предположим, что описание $\mathcal{P}_N$ корректное; мы хотим доказать, что корректное и описание $\mathcal{P}_{N+1}$. Это означает, что если kN нечетное, то нужно показать, что

$$\frac{k-1}{N+1}=\mathcal{P}_{N,kN}\,;$$

а если kN четное, то нужно показать, что

$$\mathcal{P}_{N,kN-1}\quad \mathcal{P}_{N,kN}\quad \frac{k-1}{N+1}\quad \mathcal{P}_{N,kN}\quad \mathcal{P}_{N,kN+1}\,.$$

В обоих случаях полезно знать количество дробей, строго меньших $(k-1)/(N+1)$ в $\mathcal{P}_N$; это число равно

$$\sum_{n=1}^{N}\sum_{m}\Bigl[0\leqslant\frac{m}{n}<\frac{k-1}{N+1}\Bigr]=\sum_{n=1}^{N}\left\lceil\frac{(k-1)n}{N+1}\right\rceil=\sum_{n=0}^{N}\left\lfloor\frac{(k-1)n+N}{N+1}\right\rfloor=$$
$$=\frac{(k-2)N}{2}+\frac{d-1}{2}+d\left\lfloor\frac{N}{d}\right\rfloor$$

согласно (3.32), где $d=\text{НОД}(k-1,N+1)$. Это сводится к $\frac12(kN-d+1)$, поскольку $N \bmod d=d-1$.

Кроме того, количество дробей в $\mathcal{P}_N$, равных $(k-1)/(N+1)$, которые должны предшествовать ей в $\mathcal{P}_{N+1}$, равно $\frac{1}{2}(d-1-[d \text{ четное}])$ в соответствии с природой порядка органных труб.

Если kN нечетное, то d четное и $(k-1)/(N+1)$ предшествуют $\frac{1}{2}(kN-1)$ элементов $\mathcal{P}_N$; это именно то число, которое нужно. Если kN четное, то d нечетное и $(k-1)/(N+1)$ предшествуют $\frac{1}{2}(kN)$ элементов $\mathcal{P}_N$. Если $d=1$, ни один из них не равен $(k-1)/(N+1)$ и $\mathcal{P}_{N,kN}$ представляет собой символ '<'; в противном случае $(k-1)/(N+1)$ попадает между двумя равными элементами и $\mathcal{P}_{N,kN}$ представляет собой символ '='. (Ч. С. Пирс (C. S. Peirce) [288] независимо открыл дерево Штерна–Броко примерно в то же время, когда он открыл $\mathcal{P}_N$.)

4.65 Аналогичный вопрос для (аналогичных) чисел Ферма f_n — одна из знаменитых нерешенных задач. Наш вопрос может оказаться более легким... или более трудным.

"Ни один квадрат, меньший 25×10^{14}, не делит никакое число Евклида".
— Илан Варди (Ilan Vardi)

4.66 Известно, что ни один квадрат, меньший 36×10^{18}, не делит никакое число Мерсенна или число Ферма. Но гипотеза Шинцеля (Schinzel) о том, что существует бесконечно много чисел Мерсенна, свободных от квадратов, пока что не доказана. Неизвестно даже, существует ли бесконечно много таких p, для которых $p\backslash\backslash(a\pm b)$, где все простые множители чисел a и b не больше 31.

4.67 М. Сегеди (M. Szegedy) доказал эту гипотезу для всех больших n; см. [348], [95, с. 78–79] и [55].

4.68 Это предположение гораздо слабее по сравнению с результатом из следующего упражнения.

4.69 Крамер (Cramér) [66] на основе вероятностных соображений показал, что данная гипотеза вполне правдоподобна, и это подтверждается вычислительными экспериментами: Брент (Brent) [37] показал, что $P_{n+1}-P_n \leqslant 602$ для $P_{n+1} < 2.686\times10^{12}$. Наилучшее, что известно к 1994 году — это оценка из упр. 60 [255]. Упр. 68 имеет ответ "да", если $P_{n+1}-P_n < 2P_n^{1/2}$ для всех достаточно больших n. Согласно Ги (Guy) [169, задача A8] П. Эрдеш (P. Erdős) предложил \$10 000 за доказательство того, что существует бесконечно много n, таких, что

$$P_{n+1}-P_n > \frac{c\ln n \ln\ln n \ln\ln\ln\ln n}{(\ln\ln\ln n)^2}$$

для всех $c>0$.

4.70 Согласно упр. 24 это выполняется тогда и только тогда, когда $\nu_2(n)=\nu_3(n)$. Методы из [96] могут помочь расколоть этот орешек.

4.71 Если $k = 3$, то наименьшим решением является $n = 4700063497 = 19 \cdot 47 \cdot 5263229$; другое решение для данного случая неизвестно.

4.72 Известно, что это справедливо для бесконечного количества значений a, включая -1 (очевидно) и 0 (не столь очевидно). Лемер (Lehmer) [244] выдвинул знаменитое предположение, что $\varphi(n)\backslash(n-1)$ тогда и только тогда, когда n — простое число.

4.73 Известно, что это эквивалентно гипотезе Римана (согласно которой комплексная дзета-функция $\zeta(z)$ отлична от нуля, когда действительная часть z больше $1/2$).

4.74 Экспериментальные данные дают основания предполагать, что существует около $p(1-1/e)$ различных величин — точно так же, как если бы факториалы были случайным образом распределены по модулю p.

Чему равно 11^4 в системе счисления с основанием 11?

5.1 $(11)_r^4 = (14641)_r$ в любой системе счисления с основанием $r \geqslant 7$ согласно биномиальной теореме.

5.2 Отношение $\binom{n}{k+1}/\binom{n}{k} = (n-k)/(k+1)$ оказывается $\leqslant 1$ при $k \geqslant \lfloor n/2 \rfloor$ и $\geqslant 1$ при $k < \lceil n/2 \rceil$, так что максимум достигается при $k = \lfloor n/2 \rfloor$ и $k = \lceil n/2 \rceil$.

5.3 Выразите его через факториалы. Оба произведения представляют собой $f(n)/f(n-k)f(k)$, где $f(n) = (n+1)!\,n!\,(n-1)!$.

5.4 $\binom{-1}{k} = (-1)^k\binom{k+1-1}{k} = (-1)^k\binom{k}{k} = (-1)^k[k \geqslant 0]$.

5.5 Если $0 < k < p$, то в числителе дроби $\binom{p}{k}$ имеется p, которое не сокращается ни с чем в знаменателе. Поскольку $\binom{p}{k} = \binom{p-1}{k} + \binom{p-1}{k-1}$, должно быть $\binom{p-1}{k} \equiv (-1)^k \pmod p$ при $0 \leqslant k < p$.

5.6 Решающим шагом (после избавления от второго k) должен быть следующий:

$$\frac{1}{n+1}\sum_k \binom{n+k}{k}\binom{n+1}{k+1}(-1)^k =$$
$$= \frac{1}{n+1}\sum_{k\geqslant 0} \binom{n+k}{n}\binom{n+1}{k+1}(-1)^k =$$
$$= \frac{1}{n+1}\sum_k \binom{n+k}{n}\binom{n+1}{k+1}(-1)^k -$$
$$- \frac{1}{n+1}\binom{n-1}{n}\binom{n+1}{0}(-1)^{-1}.$$

В исходном выводе этот дополнительный член, представляющий собой $[n=0]$, был забыт.

5.7 Да, так как $r^{\underline{-k}} = (-1)^k/(-r-1)^{\underline{k}}$. Кроме того,

$$r^{\overline{k}}(r+\tfrac{1}{2})^{\overline{k}} = (2r)^{\overline{2k}}/2^{2k}.$$

5.8 Функция $f(k) = (k/n-1)^n$ представляет собой полином степени n, старший коэффициент которого равен n^{-n}. Согласно (5.40) эта сумма равна $n!/n^n$. При больших значениях n в соответствии с приближением Стирлинга это значение приближенно равно $\sqrt{2\pi n}/e^n$. (Что весьма отличается от значения $(1-1/e)$, которое получается, если воспользоваться приближением $(1-k/n)^n \sim e^{-k}$, справедливым при фиксированном k и $n \to \infty$.)

5.9 $\mathcal{E}_t(z)^t = \sum_{k\geqslant 0} t(tk+t)^{k-1}z^k/k! = \sum_{k\geqslant 0}(k+1)^{k-1}(tz)^k/k! = \mathcal{E}_1(tz)$ в соответствии с (5.60).

5.10 $\sum_{k\geqslant 0} 2z^k/(k+2) = F(2,1;3;z)$, поскольку $t_{k+1}/t_k = (k+2)z/(k+3)$.

5.11 Первая функция является бесселевой, а вторая — гауссовой:

Но не бесовской.

$$z^{-1}\sin z = \sum_{k\geqslant 0}(-1)^k z^{2k}/(2k+1)! = F(1;1,\tfrac{3}{2};-z^2/4);$$
$$z^{-1}\arcsin z = \sum_{k\geqslant 0} z^{2k}(\tfrac{1}{2})^{\overline{k}}/(2k+1)k! = F(\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{2};z^2).$$

5.12 (а) Является, если $n \neq 0$, поскольку отношение членов равно n. (б) Является, если n — целое число; отношение членов равно $(k+1)^n/k^n$. Обратите внимание, что для получения этого члена из (5.115) мы полагаем $m = n+1$, $a_1 = \cdots = a_m = 1$, $b_1 = \cdots = b_n = 0$, $z = 1$ и умножаем на 0^n. (в) Является. Отношение членов равно $(k+1)(k+3)/(k+2)$. (г) Не является. Отношение членов равно $1 + 1/(k+1)H_k$, и $H_k \sim \ln k$ не является рациональной функцией. (д) Является. Обратное к любому гипергеометрическому члену представляет собой гипергеометрический член. Тот факт, что $t(k) = \infty$ при $k < 0$ и $k > n$ не исключает $t(k)$ из гипергеометрических членов. (е) Конечно, да. (ж) Не всегда — например, при $t(k) = 2^k$ и $T(k) = 1$ не является. (з) Является. Отношение членов $t(n-1-k)/t(n-1-(k+1))$ представляет собой рациональную функцию (обратное к отношению членов для t при замене k на $n-1-k$) при произвольном n. (и) Является. Отношение членов может быть записано как

$$\frac{a\,t(k+1)/t(k) + b\,t(k+2)/t(k) + c\,t(k+3)/t(k)}{a + b\,t(k+1)/t(k) + c\,t(k+2)/t(k)},$$

и $t(k+m)/t(k) = \big(t(k+m)/t(k+m-1)\big)\ldots\big(t(k+1)/t(k)\big)$ представляет собой рациональную функцию от k. (й) Не является. Если

Любое значение гипергеометрического члена $t(k)$ может быть записано как $0^{e(k)}v(k)$, где $e(k)$ — целое и $v(k) \neq 0$. Предположим, что отношение членов $t(k+1)/t(k)$ равно $p(k)/q(k)$ и что p и q полностью факторизованы в комплексных числах. Тогда для каждого k $e(k+1)$ равно $e(k)$ плюс количество нулевых множителей в разложении $p(k)$ минус количество нулевых множителей в разложении $q(k)$. $v(k+1)$ равно $v(k)$, умноженному на произведение ненулевых множителей $p(k)$ и деленному на произведение ненулевых множителей $q(k)$.

две рациональные функции, $p_1(k)/q_1(k)$ и $p_2(k)/q_2(k)$, совпадают в бесконечном количестве значений k, то они совпадают при всех k, поскольку $p_1(k)q_2(k) = q_1(k)p_2(k)$ — полиномиальное тождество. Поэтому отношение членов $\lceil (k+1)/2\rceil/\lceil k/2\rceil$ должно быть равно 1, если это рациональная функция. (к) Нет. Отношение членов должно быть равным $(k+1)/k$, поскольку это так для всех $k > 0$; но тогда $t(-1)$ может быть нулевым, только если $t(0)$ кратно 0^2, тогда как $t(1)$ может быть равно 1, только если $t(0) = 0^1$.

5.13 $R_n = n!^{n+1}/P_n^2 = Q_n/P_n = Q_n^2/n!^{n+1}$.

5.14 Первый множитель в (5.25) представляет собой $\binom{l-k}{l-k-m}$ при $k \leqslant l$, т.е. $(-1)^{l-k-m}\binom{-m-1}{l-k-m}$. Сумма при $k \leqslant l$ представляет собой сумму по всем k, поскольку $m \geqslant 0$. (Условие $n \geqslant 0$ в действительности не требуется, хотя переменная k должна принимать отрицательные значения при $n < 0$.)

Чтобы перейти от (5.25) к (5.26), сначала замените s на $-1-n-q$.

5.15 Если n нечетное, сумма равна нулю, поскольку k можно заменить на $n-k$. Если $n = 2m$, в соответствии с (5.29) при $a = b = c = m$ сумма равна $(-1)^m(3m)!/m!^3$.

5.16 Это просто $(2a)!\,(2b)!\,(2c)!/(a+b)!\,(b+c)!\,(c+a)!$, умноженное на (5.29), если записать слагаемые через факториалы.

5.17 Поскольку $\binom{2n-1/2}{n} = \binom{4n}{2n}/2^{2n}$ и $\binom{2n-1/2}{2n} = \binom{4n}{2n}/2^{4n}$, получаем $\binom{2n-1/2}{n} = 2^{2n}\binom{2n-1/2}{2n}$.

5.18 $\binom{3r}{3k}\binom{3k}{k,k,k}/3^{3k}$.

5.19 $\mathcal{B}_{1-t}(-z)^{-1} = \sum_{k\geqslant 0}\binom{k-tk-1}{k}\bigl(-1/(k-tk-1)\bigr)(-z)^k$ согласно (5.60), а это есть $\sum_{k\geqslant 0}\binom{tk}{k}\bigl(1/(tk-k+1)\bigr)z^k = \mathcal{B}_t(z)$.

5.20 Этот ряд равен $F(-a_1,\ldots,-a_m;-b_1,\ldots,-b_n;(-1)^{m+n}z)$; см. упр. 2.17.

5.21 $\lim_{n\to\infty}(n+m)^{\underline{m}}/n^m = 1$.

5.22 Умножение и деление отдельных значений (5.83) дает

$$\begin{aligned}\frac{(-1/2)!}{x!\,(x-1/2)!} &= \lim_{n\to\infty}\binom{n+x}{n}\binom{n+x-1/2}{n}n^{-2x}\Bigg/\binom{n-1/2}{n}\\ &= \lim_{n\to\infty}\binom{2n+2x}{2n}n^{-2x}\end{aligned}$$

согласно (5.34) и (5.36). Кроме того,

$$1/(2x)! = \lim_{n\to\infty}\binom{2n+2x}{2n}(2n)^{-2x}.$$

Следовательно... и т.д. Между прочим, гамма-функциональным эквивалентом этой формулы является

$$\Gamma(x)\,\Gamma(x+\tfrac{1}{2}) \;=\; \Gamma(2x)\,\Gamma(\tfrac{1}{2})/2^{2x-1}\,.$$

5.23 $(-1)^n n_{\mathrm{i}}$, см. (5.50).

5.24 Согласно (5.35) и (5.93) эта сумма равна

$$\binom{n}{m} F\left(\begin{matrix} m-n, -m \\ 1/2 \end{matrix}\,\middle|\, 1\right) = \binom{2n}{2m}.$$

5.25 Это эквивалентно легко доказываемому соотношению

$$(a-b)\frac{a^{\overline{k}}}{(b+1)^{\overline{k}}} \;=\; a\,\frac{(a+1)^{\overline{k}}}{(b+1)^{\overline{k}}} - b\,\frac{a^{\overline{k}}}{b^{\overline{k}}}$$

или, в операторной форме, $a-b = (\vartheta+a)-(\vartheta+b)$.

Аналогично

$$(a_1-a_2)\,F\left(\begin{matrix} a_1, a_2, a_3, \ldots, a_m \\ b_1, \ldots, b_n \end{matrix}\,\middle|\, z\right) \;=$$

$$=\; a_1\,F\left(\begin{matrix} a_1+1, a_2, a_3, \ldots, a_m \\ b_1, \ldots, b_n \end{matrix}\,\middle|\, z\right) -$$

$$-\,a_2\,F\left(\begin{matrix} a_1, a_2+1, a_3, \ldots, a_m \\ b_1, \ldots, b_n \end{matrix}\,\middle|\, z\right),$$

поскольку $a_1-a_2 = (a_1+k)-(a_2+k)$. Если a_1-b_1 — целое неотрицательное число d, это второе соотношение позволяет выразить $F(a_1,\ldots,a_m; b_1,\ldots,b_n; z)$ в виде линейной комбинации функций $F(a_2+j, a_3, \ldots, a_m; b_2, \ldots, b_n; z)$ при $0 \leqslant j \leqslant d$, тем самым удаляя один верхний параметр и один нижний параметр. Так, например, получается аналитический вид для $F(a,b;a-1;z)$, $F(a,b;a-2;z)$ и т.д.

Гаусс [143, §7] вывел аналогичные соотношения между функцией $F(a,b;c;z)$ и любыми двумя "смежными" гипергеометрическими функциями, в которых один из параметров изменен на ± 1. Рейнвилль (Rainville) [301] обобщил их на случай нескольких параметров.

5.26 Если отношение членов в исходном гипергеометрическом ряду равно $t_{k+1}/t_k = r(k)$, то отношение членов в новом ряду есть $t_{k+2}/t_{k+1} = r(k+1)$. Следовательно,

$$F\left(\begin{matrix} a_1, \ldots, a_m \\ b_1, \ldots, b_n \end{matrix}\,\middle|\, z\right) \;=\; 1 + \frac{a_1 \ldots a_m\, z}{b_1 \ldots b_n}\, F\left(\begin{matrix} a_1+1, \ldots, a_m+1, 1 \\ b_1+1, \ldots, b_n+1, 2 \end{matrix}\,\middle|\, z\right).$$

5.27 Это — сумма четных членов ряда $F(2a_1, \ldots, 2a_m; 2b_1, \ldots, 2b_m; z)$. Имеем $(2a)^{\overline{2k+2}}/(2a)^{\overline{2k}} = 4(k+a)(k+a+\frac{1}{2})$ и т.д.

Приравнивая коэффициенты при z^n, получаем формулу Пфаффа–Заальшютца (5.97).

5.28 $F\binom{a,b}{c}\big|z\big) = (1-z)^{-a} F\binom{a,c-b}{c}\big|\frac{-z}{1-z}\big) = (1-z)^{-a} F\binom{c-b,a}{c}\big|\frac{-z}{1-z}\big) = (1-z)^{c-a-b} F\binom{c-a,c-b}{c}\big|z\big)$. (Эйлер доказал это тождество, показав, что обе его части удовлетворяют одному и тому же дифференциальному уравнению. Закон отражения часто приписывается Эйлеру, но он, похоже, не встречается нигде в опубликованных им работах.)

5.29 В соответствии с правилом свертки Вандермонда коэффициенты при z^n равны. (Оригинальное доказательство Куммера (Kummer) было иным: он рассматривал $\lim_{m\to\infty} F(m, b-a; b; z/m)$ в законе отражения (5.101).)

5.30 Продифференцируйте еще раз, получив $z(1-z)F''(z) + (2-3z)F'(z) - F(z) = 0$. Поэтому согласно (5.108) $F(z) = F(1,1;2;z)$.

5.31 Условие $f(k) = T(k+1) - T(k)$ означает, что $f(k+1)/f(k) = \big(T(k+2)/T(k+1) - 1\big)/\big(1 - T(k)/T(k+1)\big)$ является рациональной функцией от k.

5.32 При суммировании полинома относительно k метод Госпера сводится к "методу неопределенных коэффициентов". Имеем $q(k) = r(k) = 1$ и попытаемся решить $p(k) = s(k+1) - s(k)$. В методе предполагается, что $s(k)$ является полиномом, степень которого $d = \deg(p) + 1$.

5.33 Решением соотношения $k = (k-1)s(k+1) - (k+1)s(k)$ является $s(k) = -k + \frac{1}{2}$; следовательно, искомый ответ — $(1-2k)/2k(k-1) + C$.

5.34 Предельное соотношение выполняется, поскольку все члены при $k > c$ стремятся к нулю, а при вычислении предела остальных членов $\epsilon - c$ сокращается с $-c$. Таким образом, вторая частичная сумма есть $\lim_{\epsilon\to 0} F(-m, -n; \epsilon - m; 1) = \lim_{\epsilon\to 0}(\epsilon + n - m)^{\overline{m}}/(\epsilon - m)^{\overline{m}} = (-1)^m\binom{n-1}{m}$.

5.35 (а) $2^{-n}3^n[n \geqslant 0]$. (б) $(1-\frac{1}{2})^{-k-1}[k \geqslant 0] = 2^{k+1}[k \geqslant 0]$.

5.36 Сумма цифр числа $m+n$ является суммой цифр числа m плюс сумма цифр числа n минус $p-1$, умноженное на количество переносов, так как каждый перенос уменьшает сумму цифр на $p-1$. (См. в [226] распространение этого результата на обобщенные биномиальные коэффициенты.)

5.37 При делении первого тождества на $n!$ получаем свертку Вандермонда $\binom{x+y}{n} = \sum_k \binom{x}{k}\binom{y}{n-k}$. Второе тождество вытекает, например, из формулы $x^{\overline{k}} = (-1)^k(-x)^{\underline{k}}$, если обратить как x, так и y.

5.38 Выберите c как можно бо́льшим так, чтобы $\binom{c}{3} \leqslant n$. Тогда $0 \leqslant n - \binom{c}{3} < \binom{c+1}{3} - \binom{c}{3} = \binom{c}{2}$; замените n на $n - \binom{c}{3}$ и продолжайте в том же духе. И наоборот, любое такое представление получается тем же способом. (Можно проделать то же самое с представлением

$$n = \binom{a_1}{1} + \binom{a_2}{2} + \cdots + \binom{a_m}{m}, \quad 0 \leqslant a_1 < a_2 < \cdots < a_m$$

при любом фиксированном m.)

5.39 $x^m y^n = \sum_{k=1}^{m} \binom{m+n-1-k}{n-1} a^n b^{m-k} x^k + \sum_{k=1}^{n} \binom{m+n-1-k}{m-1} \times a^{n-k} b^m y^k$ при любых $m > 0$ и $n > 0$ по индукции по $m+n$.

5.40 $(-1)^{m+1} \sum_{k=1}^{n} \sum_{j=1}^{m} \binom{r}{j} \binom{m-rk-s-1}{m-j} = (-1)^m \times \sum_{k=1}^{n} \Bigl(\binom{m-rk-s-1}{m} - \binom{m-r(k-1)-s-1}{m} \Bigr) = (-1)^m \Bigl(\binom{m-rn-s-1}{m} - \binom{m-s-1}{m} \Bigr) = \binom{rn+s}{m} - \binom{s}{m}$.

5.41 $\sum_{k \geqslant 0} n!/(n-k)!\,(n+k+1)! = \bigl(n!/(2n+1)!\bigr) \sum_{k>n} \binom{2n+1}{k}$, что равно $2^{2n} n!/(2n+1)!$.

5.42 Будем рассматривать n как неопределенную действительную переменную. Метод Госпера при $q(k) = k+1$ и $r(k) = k-1-n$ дает решение $s(k) = 1/(n+2)$; следовательно, искомая неопределенная сумма равна $(-1)^{x-1} \frac{n+1}{n+2} / \binom{n+1}{x}$. И

$$\sum_{k=0}^{n} (-1)^k \Big/ \binom{n}{k} = (-1)^{x-1} \frac{n+1}{n+2} \Big/ \binom{n+1}{x} \bigg|_0^{n+1} = $$
$$= 2\,\frac{n+1}{n+2}\,[n \text{ четное}]\,.$$

Кстати, из этого упражнения следует формула

$$\frac{1}{n\binom{n-1}{k}} = \frac{1}{(n+1)\binom{n}{k+1}} + \frac{1}{(n+1)\binom{n}{k}},$$

которая "дуальна" по отношению к основному рекуррентному соотношению (5.8).

5.43 После выполнения подсказанного первого шага можно применить правило (5.21) и просуммировать по k. Затем правило (5.21) применяется еще раз, и завершает дело свертка Вандермонда. (Комбинаторное доказательство этого тождества было дано Эндрюсом (Andrews) [10]. Способ быстрого перехода от этого тождества к доказательству (5.29) поясняется в [207, упр. 1.2.6–62].)

5.44 Сокращение факториалов показывает, что

$$\binom{m}{j}\binom{n}{k}\binom{m+n}{m} = \binom{m+n-j-k}{m-j}\binom{j+k}{j}\binom{m+n}{j+k},$$

так что вторая сумма — это дробь $1/\binom{m+n}{m}$, умноженная на первую сумму. А первая сумма представляет собой просто частный случай (5.32) для $l=0$, $n=b$, $r=a$, $s=m+n-b$, так что она равна $\binom{a+b}{a}\binom{m+n-a-b}{n-a}$.

5.45 Согласно (5.9) $\sum_{k\leqslant n}\binom{k-1/2}{k}=\binom{n+1/2}{n}$. Если это выражение недостаточно "аналитичное", можно применить (5.35) и получить $(2n+1)\binom{2n}{n}4^{-n}$.

5.46 По формуле (5.69) эта свертка противоположна коэффициенту при z^{2n} в разложении $\mathcal{B}_{-1}(z)\mathcal{B}_{-1}(-z)$. Далее, $(2\mathcal{B}_{-1}(z)-1)(2\mathcal{B}_{-1}(-z)-1)=\sqrt{1-16z^2}$; следовательно, $\mathcal{B}_{-1}(z)\mathcal{B}_{-1}(-z)=\frac14\sqrt{1-16z^2}+\frac12\mathcal{B}_{-1}(z)+\frac12\mathcal{B}_{-1}(-z)-\frac14$. По биномиальной теореме

$$(1-16z^2)^{1/2} = \sum_n\binom{1/2}{n}(-16)^n z^{2n} = -\sum_n\binom{2n}{n}\frac{4^n z^{2n}}{2n-1},$$

так что окончательный ответ — $\binom{2n}{n}4^{n-1}/(2n-1)+\binom{4n-1}{2n}/(4n-1)$.

5.47 В соответствии с (5.61), эта сумма представляет собой коэффициент при z^n в разложении $(\mathcal{B}_r(z)^s/Q_r(z))(\mathcal{B}_r(z)^{-s}/Q_r(z))=Q_r(z)^{-2}$, где $Q_r(z)=1-r+r\mathcal{B}_r(z)^{-1}$.

Заключенное в рамку утверждение на обороте этой страницы истинно

5.48 $F(2n+2,1;n+2;\frac12)=2^{2n+1}/\binom{2n+1}{n+1}$ — частный случай соотношения (5.111).

5.49 Тождество Заальшютца (5.97) дает

$$\binom{x+n}{n}\frac{y}{y+n}\,F\!\left(\left.{-x,\,-n,\,-n-y \atop -x-n,\,1-n-y}\right|1\right) = \frac{(y-x)^{\overline{n}}}{(y+1)^{\overline{n}}}.$$

5.50 Левая часть представляет собой

$$\sum_{k\geqslant0}\frac{a^{\overline{k}}\,b^{\overline{k}}}{c^{\overline{k}}}\frac{(-z)^k}{k!}\sum_{m\geqslant0}\binom{k+a+m-1}{m}z^m =$$

$$= \sum_{n\geqslant0}z^n\sum_{k\geqslant0}\frac{a^{\overline{k}}\,b^{\overline{k}}}{c^{\overline{k}}\,k!}(-1)^k\binom{n+a-1}{n-k},$$

а в соответствии со сверткой Вандермонда (5.92) коэффициент при z^n равен

$$\binom{n+a-1}{n}F\!\left(\left.{a,b,-n \atop c,a}\right|1\right) = \frac{a^{\overline{n}}}{n!}\frac{(c-b)^{\overline{n}}}{c^{\overline{n}}}.$$

5.51 (а) Отражение дает $F(a,-n;2a;2) = (-1)^n F(a,-n;2a;2)$. (Кстати, из этой формулы вытекает замечательное соотношение $\Delta^{2m+1} f(0) = 0$, когда $f(n) = 2^n x^{\underline{n}}/(2x)^{\underline{n}}$.)

(б) Почленный предел равен $\sum_{0\leqslant k\leqslant m} \binom{m}{k} \frac{2m+1}{2m+1-k}(-2)^k$ плюс дополнительный член при $k = 2m-1$. Этот дополнительный член равен

$$\frac{(-m)\dots(-1)\,(1)\dots(m)\,(-2m+1)\dots(-1)\,2^{2m+1}}{(-2m)\dots(-1)\,(2m-1)!} =$$
$$= (-1)^{m+1}\frac{m!\,m!\,2^{2m+1}}{(2m)!} = \frac{-2}{\binom{-1/2}{m}};$$

следовательно, согласно (5.104) искомый предел равен $-1/\binom{-1/2}{m}$, противоположному значению того, что мы имели.

5.52 При $k > N$ члены обоих рядов равны нулю. Это тождество соответствует замене k на $N-k$. Обратите внимание, что

$$a^{\overline{N}} = a^{\overline{N-k}}(a+N-k)^{\overline{k}} =$$
$$= a^{\overline{N-k}}(a+N-1)^{\underline{k}} = a^{\overline{N-k}}(1-a-N)^{\overline{k}}(-1)^k.$$

5.53 Если $b = -\frac{1}{2}$, то независимо от a левая часть (5.110) равна $1-2z$, а правая часть равна $(1-4z+4z^2)^{1/2}$. Правая часть представляет собой формальный степенной ряд

$$1 + \binom{1/2}{1}4z(z-1) + \binom{1/2}{2}16z^2(z-1)^2 + \cdots,$$

который может быть расписан по степеням z и переупорядочен так, чтобы получить $1-2z+0z^2+0z^3+\cdots$; однако такое переупорядочение включает в качестве промежуточного шага расходящийся при $z=1$ степенной ряд, так что оно незаконно.

Заключенное в рамку утверждение на обороте этой страницы ложно

5.54 Если $m+n$ нечетное, скажем, $2N-1$, то нужно показать, что

$$\lim_{\epsilon\to 0} F\left({N-m-\frac{1}{2},\ -N+\epsilon \atop -m+\epsilon}\middle| 1\right) = 0.$$

Равенство (5.92) применимо, поскольку $-m+\epsilon > -m-\frac{1}{2}+\epsilon$, а сомножитель $\Gamma(c-b) = \Gamma(N-m)$ в знаменателе бесконечен, поскольку $N \leqslant m$; прочие сомножители конечны. В противном случае $m+n$ четное; подставляя $n = m-2N$, согласно (5.93) получаем

$$\lim_{\epsilon\to 0} F\left({-N,\ N-m-\frac{1}{2}+\epsilon \atop -m+\epsilon}\middle| 1\right) = \frac{(N-1/2)^{\underline{N}}}{m^{\underline{N}}}.$$

Остается показать, что

$$\binom{m}{m-2N}\frac{(N-1/2)!}{(-1/2)!}\frac{(m-N)!}{m!} = \binom{m-N}{m-2N}2^{-2N},$$

а это случай $x = N$ из упр. 22.

5.55 Пусть $Q(k) = (k+A_1)\ldots(k+A_M)Z$ и $R(k) = (k+B_1)\ldots(k+B_N)$. Тогда $t(k+1)/t(k) = P(k)Q(k-1)/P(k-1)R(k)$, где $P(k) = Q(k) - R(k)$ — ненулевой полином.

5.56 Решением соотношения $-(k+1)(k+2) = s(k+1)+s(k)$ служит $s(k) = -\frac{1}{2}k^2 - k - \frac{1}{4}$; следовательно, $\sum \binom{-3}{k}\,\delta k = \frac{1}{8}(-1)^{k-1} \times (2k^2+4k+1) + C$. К тому же

$$\begin{aligned}(-1)^{k-1}\left\lfloor\frac{k+1}{2}\right\rfloor\left\lfloor\frac{k+2}{2}\right\rfloor &= \\ &= \frac{(-1)^{k-1}}{4}\left(k+1-\frac{1+(-1)^k}{2}\right)\left(k+2-\frac{1-(-1)^k}{2}\right) = \\ &= \frac{(-1)^{k-1}}{8}(2k^2+4k+1)+\frac{1}{8}.\end{aligned}$$

5.57 Имеем $t(k+1)/t(k) = (k-n)(k+1+\theta)(-z)/(k+1)(k+\theta)$. Поэтому можно взять $p(k) = k+\theta$, $q(k) = (k-n)(-z)$, $r(k) = k$. Секретная функция $s(k)$ должна быть константой α_0, и мы получаем

$$k+\theta = (-z(k-n)-k)\,\alpha_0\,;$$

следовательно, $\alpha_0 = -1/(1+z)$ и $\theta = -nz/(1+z)$. Сумма равна

$$\sum \binom{n}{k} z^k\left(k-\frac{nz}{1+z}\right)\delta k = -\frac{n}{1+z}\binom{n-1}{k-1}z^k + C.$$

(Частный случай $z = 1$ упоминался в (5.18).)

5.58 Если $m > 0$, то можно заменить $\binom{k}{m}$ на $\frac{k}{m}\binom{k-1}{m-1}$ и вывести формулу $T_{m,n} = \frac{n}{m}T_{m-1,n-1} - \frac{1}{m}\binom{n-1}{m}$. Таким образом, подойдет суммирующий множитель $\binom{n}{m}^{-1}$:

$$\frac{T_{m,n}}{\binom{n}{m}} = \frac{T_{m-1,n-1}}{\binom{n-1}{m-1}} - \frac{1}{m} + \frac{1}{n}.$$

Это можно развернуть и получить

$$\frac{T_{m,n}}{\binom{n}{m}} = T_{0,n-m} - H_m + H_n - H_{n-m}.$$

Наконец, $T_{0,n-m} = H_{n-m}$, так что $T_{m,n} = \binom{n}{m}(H_n - H_m)$. (Этот результат можно также получить, воспользовавшись производящими функциями; см. пример 2 в разделе 7.5.)

5.59 $\sum_{j\geqslant 0,\, k\geqslant 1} \binom{n}{j}\bigl[j = \lfloor \log_m k\rfloor\bigr] = \sum_{j\geqslant 0,\, k\geqslant 1} \binom{n}{j}[m^j \leqslant k < m^{j+1}]$, что равно $\sum_{j\geqslant 0} \binom{n}{j}(m^{j+1} - m^j) = (m-1)\sum_{j\geqslant 0}\binom{n}{j}m^j = (m-1)\times(m+1)^n$.

5.60 В случае $m = n$ приближение

$$\binom{m+n}{n} \approx \sqrt{\frac{1}{2\pi}\Bigl(\frac{1}{m}+\frac{1}{n}\Bigr)}\Bigl(1+\frac{m}{n}\Bigr)^n\Bigl(1+\frac{n}{m}\Bigr)^m$$

сводится к $\binom{2n}{n} \approx 4^n/\sqrt{\pi n}$.

5.61 Пусть $\lfloor n/p\rfloor = q$ и $n \bmod p = r$. Полиномиальное тождество $(x+1)^p \equiv x^p + 1 \pmod p$ означает, что

$$(x+1)^{pq+r} \equiv (x+1)^r(x^p+1)^q \pmod p.$$

Коэффициент при x^m слева равен $\binom{n}{m}$. Справа же этот коэффициент равен $\sum_k \binom{r}{m-pk}\binom{q}{k}$, что равно просто $\binom{r}{m \bmod p}\binom{q}{\lfloor m/p\rfloor}$, поскольку $0 \leqslant r < p$.

5.62 $\binom{np}{mp} = \sum_{k_1+\cdots+k_n=mp}\binom{p}{k_1}\ldots\binom{p}{k_n} \equiv \binom{n}{m} \pmod{p^2}$, поскольку все члены данной суммы кратны p^2, за исключением $\binom{n}{m}$ членов, из которых ровно m из $k_1,\ldots,k_n$ равны p. (Стенли (Stanley) [335, упр. 1.6(d)] показал, что данное сравнение в действительности выполняется по модулю p^3 при $p > 3$.)

5.63 Это $S_n = \sum_{k=0}^{n}(-4)^k\binom{n+k}{n-k} = \sum_{k=0}^{n}(-4)^{n-k}\binom{2n-k}{k}$. Знаменатель в (5.74) обращается в нуль при $z = -1/4$, так что просто осуществить подстановку в эту формулу нельзя. Рекуррентное соотношение $S_n = -2S_{n-1} - S_{n-2}$ приводит к решению $S_n = (-1)^n(2n+1)$.

5.64 Это $\sum_{k\geqslant 0}\bigl(\binom{n}{2k}+\binom{n}{2k+1}\bigr)/(k+1) = \sum_{k\geqslant 0}\binom{n+1}{2k+1}/(k+1)$, что равно

$$\frac{2}{n+2}\sum_{k\geqslant 0}\binom{n+2}{2k+2} = \frac{2^{n+2}-2}{n+2}.$$

5.65 Умножьте обе части на n^{n-1} и замените k на $n-1-k$, что даст

$$\sum_k \binom{n-1}{k} n^k (n-k)! = (n-1)!\sum_{k=0}^{n-1}\bigl(n^{k+1}/k! - n^k/(k-1)!\bigr) = $$
$$= (n-1)!\,n^n/(n-1)!.$$

(Частичные суммы на самом деле могут быть установлены с помощью алгоритма Госпера.) С другой стороны, $\binom{n}{k}kn^{n-1-k}k!$ можно интерпретировать как число отображений множества $\{1,\ldots,n\}$ в само себя при условии, что $f(1)$, ..., $f(k)$ различны, а $f(k+1) \in \{f(1),\ldots,f(k)\}$; суммирование по k должно дать n^n.

5.66 Это задача о прогулке по аллеям парка, когда имеется только один "очевидный" способ продолжения на каждом шаге. Сначала заменим $k-j$ на l, затем — $\lfloor\sqrt{l}\,\rfloor$ на k и получим

$$\sum_{j,k\geqslant 0}\binom{-1}{j-k}\binom{j}{m}\frac{2k+1}{2^j}\,.$$

Бесконечный ряд сходится, потому что все его члены при фиксированном j не превосходят некоторый полином от j, деленный на 2^j. Суммируя по k, получим

$$\sum_{j\geqslant 0}\binom{j}{m}\frac{j+1}{2^j}\,.$$

Внеся в скобки $j+1$ и применив (5.57), получаем ответ: $4(m+1)$.

5.67 Согласно (5.26) это $3\binom{2n+2}{n+5}$, поскольку

$$\binom{\binom{k}{2}}{2} = 3\binom{k+1}{4}\,.$$

5.68 Используя тот факт, что

$$\sum_{k\leqslant n/2}\binom{n}{k} = 2^{n-1}+\frac{1}{2}\binom{n}{n/2}[n \text{ четное}]\,,$$

получаем $n\left(2^{n-1}-\binom{n-1}{\lfloor n/2\rfloor}\right)$.

5.69 Поскольку $\binom{k+1}{2}+\binom{l-1}{2}\leqslant\binom{k}{2}+\binom{l}{2} \iff k<l$, минимум достигается при как можно более одинаковых k. Следовательно, согласно формуле разбиения на как можно более равные части из главы 3 минимум равен

$$(n \bmod m)\binom{\lceil n/m\rceil}{2}+(m-(n \bmod m))\binom{\lfloor n/m\rfloor}{2} = \\ = m\binom{\lfloor n/m\rfloor}{2}+(n \bmod m)\left\lfloor\frac{n}{m}\right\rfloor\,.$$

Аналогичный результат справедлив при любом нижнем индексе вместо 2.

Заключенное в рамку утверждение на обороте этой страницы не является утверждением

5.70 Это $F(-n, \frac{1}{2}; 1; 2)$; но это также и $(-2)^{-n}\binom{2n}{n}F(-n, -n; \frac{1}{2} - n; \frac{1}{2})$, если заменить k на $n-k$. Однако в соответствии с тождеством Гаусса (5.111) $F(-n, -n; \frac{1}{2} - n; \frac{1}{2}) = F(-\frac{n}{2}, -\frac{n}{2}; \frac{1}{2} - n; 1)$. (С другой стороны, $F(-n, -n; \frac{1}{2}-n; \frac{1}{2}) = 2^{-n}F(-n, \frac{1}{2}; \frac{1}{2}-n; -1)$ по закону отражения (5.101), а формула Куммера (5.94) сводит это к (5.55).) Окончательный ответ: 0 — при нечетном n, $2^{-n}\binom{n}{n/2}$ — при четном n. (См. [164, §1.2] по поводу другого вывода. Эта сумма возникает при изучении одного простого алгоритма [195].)

5.71 (а) Заметим, что

$$S(z) = \sum_{k\geqslant 0} a_k \frac{z^{m+k}}{(1-z)^{m+2k+1}} = \frac{z^m}{(1-z)^{m+1}} A\big(z/(1-z)^2\big).$$

(б) $A(z) = \sum_{k\geqslant 0}\binom{2k}{k}(-z)^k/(k+1) = \big(\sqrt{1+4z}-1\big)/2z$, так что мы имеем $A\big(z/(1-z)^2\big) = 1-z$. Таким образом, $S_n = [z^n]\,\big(z/(1-z)\big)^m = \binom{n-1}{n-m}$.

5.72 Указанную величину можно представить как $m(m-n)\ldots(m-(k-1)n)n^{k-\nu(k)}/k!$. Любой простой делитель p числа n делит числитель минимум $k-\nu(k)$ раз и знаменатель — максимум $k-\nu(k)$ раз, так как именно столько раз 2 делит $k!$. Простое число p, которое не делит n, должно делить произведение $m(m-n)\ldots(m-(k-1)n)$ как минимум столько же раз, сколько оно делит $k!$, поскольку j сомножителей $m(m-n)\ldots(m-(p^r-1)n)$ кратны p^{r-j} при любом $r \geqslant j \geqslant 1$ и любом m.

5.73 Подстановка $X_n = n!$ дает $\alpha = \beta = 1$; подстановка $X_n = n{¡}$ дает $\alpha = 1$, $\beta = 0$. Таким образом, общее решение таково: $X_n = \alpha n{¡} + \beta(n! - n{¡})$.

5.74 $\binom{n+1}{k} - \binom{n-1}{k-1}$ при $1 \leqslant k \leqslant n$.

5.75 Рекуррентное соотношение $S_k(n+1) = S_k(n) + S_{(k-1) \bmod 3}(n)$ дает возможность индуктивной проверки того, что два из чисел S одинаковы и что $S_{(-n) \bmod 3}(n)$ отличается от них на $(-1)^n$. Эти три значения разбивают свою сумму $S_0(n)+S_1(n)+S_2(n) = 2^n$ на как можно более близкие по значениям части, так что здесь должно быть $2^n \bmod 3$ вхождений величины $\lceil 2^n/3\rceil$ и $3-(2^n \bmod 3)$ вхождений $\lfloor 2^n/3\rfloor$.

Заключенное в рамку утверждение на обороте этой страницы не заключено в рамку

5.76 $Q_{n,k} = (n+1)\binom{n}{k} - \binom{n}{k+1}$.

5.77 Все члены суммы равны нулю, за исключением случая $k_1 \leqslant \cdots \leqslant k_m$, когда данное произведение представляет собой мультиномиальный коэффициент

$$\binom{k_m}{k_1,\, k_2-k_1,\, \ldots,\, k_m-k_{m-1}}.$$

Таким образом, сумма по $k_1, \ldots, k_{m-1}$ равна m^{k_m}, а последняя сумма по k_m дает $(m^{n+1}-1)/(m-1)$.

5.78 Расширьте область суммирования до $k = 2m^2 + m - 1$; тогда новые члены представляют собой $\binom{1}{4} + \binom{2}{6} + \cdots + \binom{m-1}{2m} = 0$. Поскольку $m \perp (2m+1)$, пары $\bigl(k \bmod m, k \bmod (2m+1)\bigr)$ различны. Кроме того, числа $(2j+1) \bmod (2m+1)$, когда j изменяется от 0 до $2m$, — это числа $0, 1, \ldots, 2m$ в некотором порядке. Следовательно, данная сумма равна

$$\sum_{\substack{0\leqslant k<m\\ 0\leqslant j<2m+1}} \binom{k}{j} = \sum_{0\leqslant k<m} 2^k = 2^m - 1\,.$$

5.79 (a) Сумма равна 2^{2n-1}, так что искомый НОД должен быть степенью 2. Если $n = 2^k q$, где q нечетное, $\binom{2n}{1}$ делится на 2^{k+1} и не делится на 2^{k+2}. Каждое из чисел $\binom{2n}{2j+1}$ делится на число 2^{k+1} (см. упр. 36), так что это число должно быть их НОД. (б) Если $p^r \leqslant n+1 < p^{r+1}$, то наибольшее количество переносов при сложении k с $n-k$ в системе счисления с основанием p получается при $k = p^r - 1$. В этом случае количество переносов равно $r - \epsilon_p(n+1)$ и $r = \epsilon_p\bigl(L(n+1)\bigr)$.

5.80 Сначала докажите по индукции, что $k! \geqslant (k/e)^k$.

5.81 Обозначим через $f_{l,m,n}(x)$ левую часть. Достаточно показать, что $f_{l,m,n}(1) > 0$ и $f'_{l,m,n}(x) < 0$ при $0 \leqslant x \leqslant 1$. Величина $f_{l,m,n}(1)$ равна $(-1)^{n-m-1}\binom{l+m+\theta}{l+n}$ согласно (5.23) и эта величина положительна, потому что данный биномиальный коэффициент имеет ровно $n-m-1$ отрицательных сомножителей. По той же причине это неравенство справедливо и при $l = 0$. Если $l > 0$, то $f'_{l,m,n}(x) = -l\, f_{l-1,m,n+1}(x)$, а по индукции эта величина отрицательна.

5.82 Пусть $\epsilon_p(a)$ — показатель степени, в которой простое число p делит a и пусть $m = n - k$. Тогда доказываемое тождество сводится к

$$\begin{aligned}\min\bigl(&\epsilon_p(m)-\epsilon_p(m+k), \epsilon_p(m+k+1)-\epsilon_p(k+1), \epsilon_p(k)-\epsilon_p(m+1)\bigr) = \\ &= \min\bigl(\epsilon_p(k)-\epsilon_p(m+k), \epsilon_p(m)-\epsilon_p(k+1), \epsilon_p(m+k+1)-\epsilon_p(m+1)\bigr)\,.\end{aligned}$$

Для краткости запишем это как $\min(x_1, y_1, z_1) = \min(x_2, y_2, z_2)$. Обратите внимание, что $x_1 + y_1 + z_1 = x_2 + y_2 + z_2$. Общее соотношение

$$\epsilon_p(a) < \epsilon_p(b) \quad\Longrightarrow\quad \epsilon_p(a) = \epsilon_p\bigl(|a \pm b|\bigr)$$

позволяет заключить, что $x_1 \neq x_2 \Longrightarrow \min(x_1, x_2) = 0$; то же самое справедливо и для (y_1, y_2) и (z_1, z_2). Теперь завершить доказательство не составляет труда.

5.83 (Решение П. Пауле (P. Paule).) Пусть r — неотрицательное целое число. Данная сумма представляет собой коэффициент при $x^l y^m$ в

$$\begin{aligned}\sum_{j,k}(-1)^{j+k}\frac{(1+x)^{j+k}}{x^k}\binom{r}{j}\binom{n}{k}(1+y)^{s+n-j-k}y^j &= \\ = \left(1-\frac{(1+x)y}{1+y}\right)^r\left(1-\frac{1+x}{(1+y)x}\right)^n(1+y)^{s+n} &= \\ = (-1)^n(1-xy)^{n+r}(1+y)^{s-r}/x^n\,,&\end{aligned}$$

так что понятно, что она равна $(-1)^l\binom{n+r}{n+l}\binom{s-r}{m-n-l}$. (См. также упр. 106.)

5.84 Следуя указанию, получаем

$$z\mathcal{B}_t(z)^{r-1}\mathcal{B}_t'(z) = \sum_{k\geqslant 0}\binom{tk+r}{k}\frac{kz^k}{tk+r}$$

и аналогичную формулу для $\mathcal{E}_t(z)$. Таким образом, формулы $\bigl(zt\mathcal{B}_t^{-1}(z)\mathcal{B}_t'(z)+1\bigr)\mathcal{B}_t(z)^r$ и $\bigl(zt\mathcal{E}_t^{-1}(z)\mathcal{E}_t'(z)+1\bigr)\mathcal{E}_t(z)^r$ дают соответствующие правые части формул (5.61). Поэтому необходимо доказать, что

$$\begin{aligned}zt\mathcal{B}_t^{-1}(z)\mathcal{B}_t'(z)+1 &= \frac{1}{1-t+t\mathcal{B}_t(z)^{-1}}\,,\\ zt\mathcal{E}_t^{-1}(z)\mathcal{E}_t'(z)+1 &= \frac{1}{1-zt\mathcal{E}(z)^t}\,,\end{aligned}$$

а это вытекает из (5.59).

5.85 Если $f(x) = a_n x^n + \cdots + a_1 x + a_0$ является некоторым полиномом степени $\leqslant n$, то по индукции можно доказать, что

$$\sum_{0\leqslant\epsilon_1,\ldots,\epsilon_n\leqslant 1}(-1)^{\epsilon_1+\cdots+\epsilon_n}f(\epsilon_1x_1+\cdots+\epsilon_nx_n) = (-1)^n n!\,a_n x_1\ldots x_n\,.$$

Предложенное в упражнении тождество является частным случаем данного при $a_n = 1/n!$ и $x_k = k^3$.

5.86 (а) Сначала представьте его с $n(n-1)$ индексными переменными l_{ij} при $i \neq j$. Подстановка $k_{ij} = l_{ij} - l_{ji}$ при $1 \leqslant i < j < n$ и использование ограничений $\sum_{i\neq j}(l_{ij} - l_{ji}) = 0$ при всех $i < n$ позволяет вычислить суммы по l_{jn} для $1 \leqslant j < n$, а затем по l_{ji} при $1 \leqslant i < j < n$ по правилу свертки Вандермонда. (б) $f(z) - 1$

представляет собой полином степени $< n$, имеющий n корней, так что он должен быть равен нулю. (в) Рассмотрите постоянные члены в

$$\prod_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}}\left(1-\frac{z_i}{z_j}\right)^{a_i} = \sum_{k=1}^{n}\prod_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}}\left(1-\frac{z_i}{z_j}\right)^{a_i-[i=k]}.$$

5.87 Согласно (5.61) первый член представляет собой $\sum_k\binom{n-k}{k}z^{mk}$. Слагаемые во втором члене — это

$$\frac{1}{m}\sum_{k\geqslant 0}\binom{(n+1)/m+(1+1/m)k}{k}(\zeta z)^{k+n+1} = \\ = \frac{1}{m}\sum_{k>n}\binom{(1+1/m)k-n-1}{k-n-1}(\zeta z)^{k}.$$

Поскольку $\sum_{0\leqslant j<m}(\zeta^{2j+1})^k = m(-1)^l[k=ml]$, эти члены дают в сумме

Заключенное в рамку утверждение на обороте этой страницы ссылается само на себя

$$\sum_{k>n/m}\binom{(1+1/m)mk-n-1}{mk-n-1}(-z^m)^k = \\ = \sum_{k>n/m}\binom{(m+1)k-n-1}{k}(-z^m)^k = \sum_{k>n/m}\binom{n-mk}{k}z^{mk}.$$

Кстати, функции $\mathcal{B}_m(z^m)$ и $\zeta^{2j+1}z\,\mathcal{B}_{1+1/m}(\zeta^{2j+1}z)^{1/m}$ представляют собой $m+1$ комплексных корней уравнения $w^{m+1}-w^m=z^m$.

5.88 Воспользуемся теми фактами, что $\int_0^\infty(e^{-t}-e^{-nt})\,dt/t = \ln n$ и $(1-e^{-t})/t \leqslant 1$. (Согласно (5.83) $\binom{x}{k} = O(k^{-x-1})$ при $k\to\infty$; так что эта оценка означает, что ряд Стирлинга $\sum_k s_k\binom{x}{k}$ сходится при $x > -1$. Эрмит (Hermite) [186] показал, что его сумма равна $\ln\Gamma(1+x)$.)

5.89 Сложение этого соотношения с (5.19) дает в обеих частях $y^{-r}(x+y)^{m+r}$ в силу биномиальной теоремы. Дифференцирование дает

$$\sum_{k>m}\binom{m+r}{k}\binom{m-k}{n}x^k y^{m-k-n} = \\ = \sum_{k>m}\binom{-r}{k}\binom{m-k}{n}(-x)^k(x+y)^{m-k-n},$$

и можно заменить k на $k+m+1$ и применить (5.15), чтобы получить

$$\sum_{k\geqslant 0}\binom{m+r}{m+1+k}\binom{-n-1}{k}(-x)^{m+1+k}y^{-1-k-n} =$$
$$= \sum_{k\geqslant 0}\binom{-r}{m+1+k}\binom{-n-1}{k}x^{m+1+k}(x+y)^{-1-k-n}.$$

В гипергеометрическом виде это сводится к

$$F\left({1-r,\ n+1 \atop m+2}\,\middle|\,\frac{-x}{y}\right) = \left(1+\frac{x}{y}\right)^{-n-1} F\left({m+1+r,\ n+1 \atop m+2}\,\middle|\,\frac{x}{x+y}\right),$$

что является частным случаем закона отражения (5.101) при $(a,b,c,z) = (n+1, m+1+r, m+2, -x/y)$. (Таким образом, преобразование (5.105) связано с законом отражения и формулой из упр. 52.)

5.90 Если r — неотрицательное целое число, сумма конечна и вывод в тексте корректен постольку, поскольку ни в одном из членов этой суммы при $0 \leqslant k \leqslant r$ нет нуля в знаменателе. В противном случае данная сумма бесконечна, и согласно (5.83) ее k-й член $\binom{k-r-1}{k}/\binom{k-s-1}{k}$ приблизительно равен $k^{s-r}(-s-1)!/(-r-1)!$. Поэтому для того, чтобы данный бесконечный ряд сходился, необходимо условие $r > s+1$. (Если r и s комплексные, то этим условием будет $\Re r > \Re s + 1$, поскольку $|k^z| = k^{\Re z}$.) Согласно (5.92) искомая сумма равна

Заключенное в рамку утверждение на обороте этой страницы не ссылается само на себя

$$F\left({-r,\ 1 \atop -s}\,\middle|\,1\right) = \frac{\Gamma(r-s-1)\Gamma(-s)}{\Gamma(r-s)\Gamma(-s-1)} = \frac{s+1}{s+1-r}.$$

Это та же формула, что и найденная для целых r и s.

5.91 (В этой задаче стóит прибегнуть к помощи компьютера.) Кстати, ввиду закона отражения Пфаффа при $c = (a+1)/2$ это тождество сводится к тождеству, которое эквивалентно тождеству Гаусса (5.110). Действительно, если $w = -z/(1-z)$, то $4w(1-w) = -4z/(1-z)^2$, и

$$F\left({\frac{1}{2}a,\ \frac{1}{2}a+\frac{1}{2}-b \atop 1+a-b}\,\middle|\,4w(1-w)\right) = F\left({a,\ a+1-2b \atop 1+a-b}\,\middle|\,\frac{-z}{1-z}\right) =$$
$$= (1-z)^a F\left({a,\ b \atop 1+a-b}\,\middle|\,z\right).$$

5.92 Эти тождества можно доказать так, как их доказал сам Клаузен (Clausen) более 150 лет назад, показав, что обе части удовлетворяют одному и тому же дифференциальному уравнению. Один из способов записать получающиеся равенства между коэффициентами при z^n — в терминах биномиальных коэффициентов:

$$\sum_k \frac{\binom{r}{k}\binom{s}{k}\binom{r}{n-k}\binom{s}{n-k}}{\binom{r+s-1/2}{k}\binom{r+s-1/2}{n-k}} = \frac{\binom{2r}{n}\binom{r+s}{n}\binom{2s}{n}}{\binom{2r+2s}{n}\binom{r+s-1/2}{n}};$$

$$\sum_k \frac{\binom{-1/4+r}{k}\binom{-1/4+s}{k}\binom{-1/4-r}{n-k}\binom{-1/4-s}{n-k}}{\binom{-1+r+s}{k}\binom{-1-r-s}{n-k}} =$$

$$= \frac{\binom{-1/2}{n}\binom{-1/2+r-s}{n}\binom{-1/2-r+s}{n}}{\binom{-1+r+s}{n}\binom{-1-r-s}{n}}.$$

Другой способ — выразить их в гипергеометрическом виде:

$$F\left(\begin{matrix}a, b, \frac{1}{2}-a-b-n, -n\\ \frac{1}{2}+a+b, 1-a-n, 1-b-n\end{matrix}\,\middle|\,1\right) = \frac{(2a)^{\overline{n}}\,(a+b)^{\overline{n}}\,(2b)^{\overline{n}}}{(2a+2b)^{\overline{n}}\,a^{\overline{n}}\,b^{\overline{n}}};$$

$$F\left(\begin{matrix}\frac{1}{4}+a, \frac{1}{4}+b, a+b-n, -n\\ 1+a+b, \frac{3}{4}+a-n, \frac{3}{4}+b-n\end{matrix}\,\middle|\,1\right) =$$

$$= \frac{(1/2)^{\overline{n}}\,(1/2+a-b)^{\overline{n}}\,(1/2-a+b)^{\overline{n}}}{(1+a+b)^{\overline{n}}\,(1/4-a)^{\overline{n}}\,(1/4-b)^{\overline{n}}}.$$

5.93 Этот вид — $\alpha^{-1}\prod_{j=1}^{k-1}\bigl(f(j)+\alpha\bigr)/f(j)$.

5.94 Алгоритм Госпера находит ответ $-\binom{a-1}{k-1}\binom{-a-1}{n-k}a/n + C$. Следовательно, если $m \geqslant 0$ — целое число, меньшее n, получаем

$$\sum \binom{a}{k}\binom{m-a}{n-k}\delta k = \sum_j \binom{m}{j}\frac{-a}{n-j}\binom{a-1}{k-1}\binom{-a-1}{n-j-k} + C.$$

5.95 Старшие коэффициенты p и r должны быть единицами, p не должен иметь общих множителей с q или r. Эти дополнительные условия легко выполнимы при помощи перестановки множителей из одного полинома в другой.

Предположим теперь, что $p(k+1)q(k)/p(k)r(k+1) = P(k+1)Q(k)/P(k)R(k+1)$, причем тройки полиномов (p, q, r) и (P, Q, R)

удовлетворяют новым критериям. Пусть $p_0(k) = p(k)/g(k)$ и $P_0(k) = P(k)/g(k)$, где $g(k) = \text{НОД}\bigl(p(k), P(k)\bigr)$ представляет собой произведение всех общих множителей p и P. Тогда

$$p_0(k+1)q(k)P_0(k)R(k+1) = p_0(k)r(k+1)P_0(k+1)Q(k).$$

Предположим, что $p_0(k) \neq 1$. Тогда существует такое комплексное число α, что $p_0(\alpha) = 0$; отсюда вытекает, что $q(\alpha) \neq 0$, $r(\alpha) \neq 0$ и $P_0(\alpha) \neq 0$. Следовательно, должно быть $p_0(\alpha+1)R(\alpha+1) = 0$ и $p_0(\alpha-1)Q(\alpha-1) = 0$. Пусть N — положительное целое число, такое, что $p_0(\alpha+N) \neq 0$ и $p_0(\alpha-N) \neq 0$. Повторяя те же рассуждения N раз, найдем, что $R(\alpha+1)\ldots R(\alpha+N) = 0 = Q(\alpha-1)\ldots Q(\alpha-N)$, а это противоречит (5.118). Таким образом, $p_0(k) = 1$. Аналогично $P_0(k) = 1$, так что $p(k) = P(k)$. Далее, из $q(\alpha) = 0$ в силу (5.118) вытекает $r(\alpha+1) \neq 0$, следовательно, $q(k)\backslash Q(k)$. Аналогично $Q(k)\backslash q(k)$, так что $q(k) = Q(k)$, поскольку у них один и тот же старший коэффициент. Это влечет за собой $r(k) = R(k)$.

5.96 Если $r(k)$ — ненулевая рациональная функция и $T(k)$ — гипергеометрический член, то $r(k)T(k)$ — тоже гипергеометрический член, который называется *подобным* $T(k)$. (Мы допускаем обращение $r(k)$ в ∞, а $T(k)$ — 0 или наоборот для конечного множества значений k.) В частности, $T(k+1)$ всегда подобен $T(k)$. Если $T_1(k)$ и $T_2(k)$ — два подобных гипергеометрических члена, то $T_1(k)+T_2(k)$ также является гипергеометрическим членом. Если $T_1(k)$, ..., $T_m(k)$ взаимно неподобны и $m > 1$, то $T_1(k)+\cdots+T_m(k)$ не может быть нулем при всех k, кроме конечного множества. Действительно, если бы это было возможно, мы взяли бы контрпример с минимальным m и положили бы $r_j(k) = T_j(k+1)/T_j(k)$. Поскольку $T_1(k)+\cdots+T_m(k) = 0$, имеем $r_m(k)T_1(k)+\cdots+r_m(k)T_m(k) = 0$ и $r_1(k)T_1(k)+\cdots+r_m(k)T_m(k) = T_1(k+1)+\cdots+T_m(k+1) = 0$; следовательно, $\bigl(r_m(k)-r_1(k)\bigr)T_1(k)+\cdots+\bigl(r_m(k)-r_{m-1}(k)\bigr)T_{m-1}(k) = 0$. Мы не можем получить $r_m(k)-r_j(k) = 0$ при всех $j < m$, поскольку T_j и T_m неподобны. Но m выбрано минимальным, так что выбранный набор не может быть контрпримером; отсюда следует, что $m = 2$. Но тогда $T_1(k)$ и $T_2(k)$ должны быть подобными, поскольку они оба равны нулю при всех k, кроме конечного множества.

Пусть теперь $t(k)$ — любой гипергеометрический член, такой, что $t(k+1)/t(k) = r(k)$, и предположим, что $t(k) = \bigl(T_1(k+1)+\cdots+T_m(k+1)\bigr)-\bigl(T_1(k)+\cdots+T_m(k)\bigr)$, где m минимально. Тогда T_1, ..., T_m должны быть взаимно неподобны. Пусть $r_j(k)$ —

рациональная функция, такая, что

$$r(k)\big(T_j(k+1)-T_j(k)\big)-\big(T_j(k+2)-T_j(k+1)\big) \;=\; r_j(k)T_j(k)\,.$$

Предположим, что $m>1$. Поскольку $0=r(k)t(k)-t(k+1)=r_1(k)T_1(k)+\cdots+r_m(k)T_m(k)$, должно быть $r_j(k)=0$ при всех j, за исключением не более одного значения. Если $r_j(k)=0$, функция $\bar{t}(k)=T_j(k+1)-T_j(k)$ удовлетворяет соотношению $\bar{t}(k+1)/\bar{t}(k)=t(k+1)/t(k)$. Поэтому алгоритм Госпера обнаружит решение.

5.97 Предположим сначала, что z не равно $-d-1/d$ ни при каком целом $d>0$. Тогда в алгоритме Госпера мы имеем $p(k)=1$, $q(k)=(k+1)^2$, $r(k)=k^2+kz+1$. Поскольку $\deg(Q)<\deg(R)$ и $\deg(p)-\deg(R)+1=-1$, единственно возможным выбором является $z=d+2$, где d— неотрицательное целое число. Попытка взять $s(k)=\alpha_d k^d+\cdots+\alpha_0$ терпит неудачу при $d=0$, но оказывается успешной при $d>0$. (Система линейных уравнений, получаемая приравниванием коэффициентов при k^d, k^{d-1}, …, k^1 в (5.122), выражает α_{d-1}, …, α_0 как положительные кратные α_d, и тогда оставшееся уравнение $1=\alpha_d+\cdots+\alpha_1$ определяет α_d.) Например, для $z=3$ неопределенная сумма равна $(k+2)k!^2/\prod_{j=1}^{k-1}(j^2+3j+1)+C$.

Если, напротив, $z=-d-1/d$, то указанные члены $t(k)$ для $k\geqslant d$ бесконечны. В этой ситуации есть два разумных способа действия. Можно убрать нули из знаменателя, переопределив

$$t(k) \;=\; \frac{k!^2}{\prod_{j=d+1}^{k}\big(j^2-j(d+1/d)+1\big)} \;=\; \frac{(d-1/d)!\,k!^2}{(k-1/d)!\,(k-d)!}\,,$$

и тем самым сделав $t(k)=0$ для $0\leqslant k<d$ и положительным для $k\geqslant d$. В этом случае алгоритм Госпера дает $p(k)=k^{\underline{d}}$, $q(k)=k+1$, $r(k)=k-1/d$, и мы можем решить (5.122) относительно $s(k)$, потому что коэффициент при k^j в правой части равен $(j+1+1/d)\alpha^j$ плюс кратные $\{\alpha_{j+1},\dots,\alpha_d\}$. Например, когда $d=2$, неопределенная сумма равна $(3/2)!\,k!\,(\frac{2}{7}k^2-\frac{26}{35}k+\frac{32}{105})/(k-3/2)!+C$.

В качестве альтернативного пути можно попытаться суммировать исходные члены, но только в диапазоне $0\leqslant k<d$. Тогда можно заменить $p(k)=k^{\underline{d}}$ на

Заметьте, любая конечная последовательность тривиально суммируема, потому что можно найти полином, который совпадает с $t(k)$ при $0\leqslant k<d$.

$$p'(k) \;=\; \sum_{j=1}^{d}(-1)^{d-j}j\begin{bmatrix}d\\j\end{bmatrix}k^{j-1}\,.$$

Эта замена допустима, поскольку (5.117) по-прежнему выполнено для $0\leqslant k<d-1$; мы имеем $p'(k)=\lim_{\epsilon\to0}\big((k+\epsilon)^{\underline{d}}-k^{\underline{d}}\big)/\epsilon=$

$\lim_{\epsilon\to 0}(k+\epsilon)^{\underline{d}}/\epsilon$, так что этот трюк по сути сокращает нули в числителе и знаменателе (5.117), как в правиле Лопиталя. Метод Госпера после этого вычисляет неопределенную сумму.

5.98 $nS_{n+1} = 2nS_n$. (Будьте осторожны: это рекуррентное соотношение не дает информации о S_1/S_0.)

5.99 Пусть $p(n,k) = (n{+}1{+}k)\beta_0(n){+}(n{+}1{+}a{+}b{+}c{+}k)\beta_1(n) = \hat{p}(n,k)$, $\bar{t}(n,k) = t(n,k)/(n+1+k)$, $q(n,k) = (n+1+a+b+c+k)(a-k)(b-k)$, $r(n,k) = (n+1+k)(c+k)k$. Тогда решением (5.129) будет $\beta_0(n) = (n+1+a+b+c)(n+1+a+b)$, $\beta_1(n) = -(n+1+a)(n+1+b)$, $\alpha_0(n) = s(n,k) = -1$. Мы находим (5.134), заметив, что это равенство истинно при $n = -a$, и используя индукцию по n.

5.100 Алгоритм Госпера–Зайльбергера легко находит, что

$$\frac{n+2}{\binom{n}{k}} - \frac{2n+2}{\binom{n+1}{k}} = \frac{n-k}{\binom{n}{k+1}} - \frac{n+1-k}{\binom{n}{k}}, \qquad 0 \leqslant k < n.$$

Суммирование от $k=0$ до $n{-}1$ дает $(n{+}2)(S_n{-}1){-}(2n{+}2)(S_{n+1}{-}1-\frac{1}{n+1}) = -n$. Следовательно, $(2n{+}2)S_{n+1} = (n{+}2)S_n + 2n{+}2$. Теперь применение суммирующего множителя приводит к выражению $S_n = (n+1)2^{-n}\sum_{k=0}^{n} 2^k/(k+1)$.

5.101 (а) Если зафиксировать m, то алгоритм Госпера–Зайльбергера обнаруживает, что $(n{+}2)S_{m,n+2}(z) = (z{-}1)(n{+}1)S_{m,n}(z) + (2n+3-z(n-m+1))S_{m,n+1}(z)$. Можно также применить этот метод к слагаемому

$$\beta_0(m,n)t(m,n,k) + \beta_1(m,n)t(m{+}1,n,k) + \beta_2(m,n)t(m,n{+}1,k),$$

что дает более простое рекуррентное соотношение

$$(m+1)S_{m+1,n}(z) - (n+1)S_{m,n+1}(z) = (1-z)(m-n)S_{m,n}(z).$$

(б) Здесь нам придется поработать немного больше — у нас пять уравнений с шестью неизвестными. Алгоритм находит

$$(n+1)(z-1)^2\binom{n}{k}^2 z^k - (2n+3)(z+1)\binom{n+1}{k}^2 z^k + {}$$
$$+ (n+2)\binom{n+2}{k}^2 z^k = T(n,k+1) - T(n,k),$$

$$T(n,k) = \binom{n+1}{k-1}^2 \frac{s(n,k)}{n+1} z^k,$$
$$s(n,k) = (z{-}1)k^2 - 2((n{+}2)z{-}2n{-}3)k + (n{+}2)((n{+}2)z{-}4n{-}5).$$

Таким образом, $(n+1)(z-1)^2 S_n(z) - (2n+3)(z+1)S_{n+1}(z) + (n+2)S_{n+2}(z) = 0$. Кстати, это рекуррентное соотношение выполняется и для отрицательных n, и мы имеем $S_{-n-1}(z) = S_n(z)/(1-z)^{2n+1}$.

Сумму $S_n(z)$ можно рассматривать как модифицированный полином Лежандра $P_n(z) = \sum_k \binom{n}{k}^2 (z-1)^{n-k}(z+1)^k/2^n$, поскольку можно записать $S_n(z) = (1-z)^n P_n\left(\frac{1+z}{1-z}\right)$. Аналогично сумма $S_{m,n}(z) = (1-z)^n P_n^{(0,m-n)}\left(\frac{1+z}{1-z}\right)$ представляет собой модификацию полинома Якоби.

А случай $z = 0$?

5.102 Эта сумма равна $F(a - \frac{1}{3}n, -n; b - \frac{4}{3}n; -z)$, так что нам не надо рассматривать случай $z = -1$. Пусть $n = 3m$. Мы ищем решение (5.129) для

$$\begin{aligned} p(m,k) &= (3m+3-k)^{\underline{3}}(m+1-k)\beta_0 + (4m+4-b-k)^{\underline{4}}\beta_1\,, \\ q(m,k) &= (3m+3-k)(m+1-a-k)z\,, \\ r(m,k) &= k(4m+1-b-k)\,, \\ s(m,k) &= \alpha_2 k^2 + \alpha_1 k + \alpha_0\,. \end{aligned}$$

Получающиеся в результате пять однородных уравнений имеют ненулевое решение $(\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1)$ тогда и только тогда, когда матрица коэффициентов имеет нулевой определитель. Этот определитель, полином от m, обращается в нуль только в восьми случаях. Один из этих случаев, конечно, (5.113); но теперь мы можем вычислить сумму для всех неотрицательных целых чисел n, а не только для $n \not\equiv 2 \pmod 3$:

$$\sum_k \binom{n}{k}\binom{\frac{1}{3}n - \frac{1}{6}}{k} 8^k \Big/ \binom{\frac{4}{3}n - \frac{2}{3}}{k} = [1, 1, -\tfrac{1}{2}]\binom{2n}{n}\Big/\binom{\frac{4}{3}n - \frac{2}{3}}{n}\,.$$

Здесь запись $[c_0, c_1, c_2]$ обозначает одно значение $c_{n \bmod 3}$. Другой случай, $(a, b, z) = (\frac{1}{2}, 0, 8)$, дает тождество

$$\sum_k \binom{n}{k}\binom{\frac{1}{3}n - \frac{1}{2}}{k} 8^k \Big/ \binom{\frac{4}{3}n}{k} = [1, 0, 0]\, 16^{n/3}\binom{\frac{2}{3}n}{\frac{1}{3}n}\Big/\binom{\frac{4}{3}n}{n}\,.$$

(Удивительным образом эта сумма равна нулю, если n не кратно 3; в последнем случае получается тождество

$$\sum_k \binom{3m}{k}\binom{2m}{2k}\binom{2k}{k} 2^k \Big/ \binom{4m}{k}\binom{m}{k} = 16^m \frac{(3m)!\,(2m)!}{(4m)!\,m!}\,,$$

которое может даже оказаться полезным.) Остальные шесть случаев порождают еще более таинственные суммы

$$\sum_k \binom{n}{k}\binom{\frac{1}{3}n-a}{k} z^k \Big/ \binom{\frac{4}{3}n-b}{k} =$$

$$= [c_0, c_1, c_2] \frac{\binom{\frac{1}{3}n-a}{\lfloor n/3\rfloor}\binom{\frac{1}{3}n-a'}{\lfloor n/3\rfloor} x^{\lfloor n/3\rfloor}}{\binom{\frac{4}{3}n-b}{n}\binom{\frac{1}{3}n-b}{\lfloor n/3\rfloor}\binom{\frac{1}{3}n-b'}{\lfloor n/3\rfloor}},$$

где значения $(a, b, z, c_0, c_1, c_2, a', b', x)$ равны соответственно

$(\frac{7}{12}, \frac{1}{3}, 8, 1, -1, 0, \frac{1}{4}, 0, 64)$; $(\frac{1}{4}, 0, 8, 1, 2, 0, \frac{7}{12}, \frac{1}{3}, 64)$;
$(\frac{5}{12}, \frac{2}{3}, 8, 1, 0, -3, \frac{3}{4}, 0, 64)$; $(\frac{1}{12}, \frac{1}{3}, 8, 1, 3, 0, \frac{3}{4}, 0, 64)$;
$(\frac{1}{2}, 0, -4, 1, 2, 0, \frac{1}{6}, \frac{1}{3}, -16)$; $(\frac{1}{6}, \frac{2}{3}, -4, 1, 0, -3, \frac{5}{6}, 0, -16)$.

5.103 Будем считать, что все a_i' и b_i' ненулевые, поскольку в противном случае соответствующие множители не окажут влияния на степени относительно k. Пусть $\hat{t}(n,k) = \hat{p}(n,k)\bar{t}(n,k)$, где

$$\bar{t}(n,k) = \frac{\prod_{i=1}^{p}\bigl(a_in + a_i'k + a_il[a_i<0] + a_i''\bigr)!}{\prod_{i=1}^{q}\bigl(b_in + b_i'k + b_il[b_i>0] + b_i''\bigr)!} z^k .$$

В таком случае $\deg(\hat{p}) = \deg(f) + \max\bigl(\sum_{i=1}^{q} b_i[b_i>0] - \sum_{i=1}^{p} a_i[a_i<0], \sum_{i=1}^{p} a_i[a_i>0] - \sum_{i=1}^{q} b_i[b_i<0]\bigr) \geqslant \deg(f) + \frac{1}{2}l\bigl(|a_1| + \cdots + |a_p| + |b_1| + \cdots + |b_q|\bigr)$, за исключениями случаев, когда в старшем коэффициенте происходит сокращение. Также $\deg(q) = \sum_{i=1}^{p} a_i'[a_i'>0] - \sum_{i=1}^{q} b_i'[b_i'<0]$, $\deg(r) = \sum_{i=1}^{q} b_i'[b_i'>0] - \sum_{i=1}^{p} a_i'[a_i'<0]$, вновь кроме исключительных случаев.

(Этими оценками можно воспользоваться для того, чтобы непосредственно показать, что с увеличением l степень $\hat{p}$ в конце концов становится достаточно большой, чтобы обеспечить существование полинома $s(n,k)$, а количество неизвестных α_j и β_j в конце концов становится больше числа однородных уравнений в решаемой системе. Так что мы получаем еще одно доказательство сходимости алгоритма Госпера–Зайльбергера, если, как в тексте, прибегнем к аргументу о существовании решения, в котором не все $\beta_0(n)$, …, $\beta_l(n)$ нулевые.)

5.104 Пусть $t(n,k) = (-1)^k(r-s-k)!\,(r-2k)!/\bigl((r-s-2k)!\,(r-n-k+1)!\,(n-k)!\,k!\bigr)$. Тогда $\beta_0(n)t(n,k) + \beta_1(n)t(n+1,k)$ не суммируемо в гипергеометрических членах, потому что $\deg(\hat{p}) = 1$, $\deg(q-r) = 3$, $\deg(q+r) = 4$, $\lambda = -8$, $\lambda' = -4$; но $\beta_0(n)t(n,k) +$

$\beta_1(n)t(n+1,k) + \beta_2(n)t(n+2,k)$ уже суммируемо, в основном потому, что $\lambda' = 0$ при $q(n,k) = -(r-s-2k)(r-s-2k-1)(n+2-k)(r-n-k+1)$ и $r(k) = (r-s-k+1)(r-2k+2)(r-2k+1)k$. Решением является

$$\begin{aligned}\beta_0(n) &= (s-n)(r-n+1)(r-2n+1)\,,\\ \beta_1(n) &= (rs-s^2-2rn+2n^2-2r+2n)(r-2n-1)\,,\\ \beta_2(n) &= (s-r+n+1)(n+2)(r-2n-3)\,,\\ \alpha_0(n) &= r-2n-1\,,\end{aligned}$$

и мы можем заключить, что $\beta_0(n)S_n + \beta_1(n)S_{n+1} + \beta_2(n)S_{n+2} = 0$, где S_n обозначает указанную сумму. Этого достаточно, чтобы доказать тождество по индукции после проверки случаев $n = 0$ и $n = 1$.

Но S_n удовлетворяет также и более простому рекуррентному соотношению $\bar\beta_0(n)S_n + \bar\beta_1(n)S_{n+1} = 0$, где $\bar\beta_0(n) = (s-n)(r-2n+1)$ и $\bar\beta_1(n) = -(n+1)(r-2n-1)$. Почему же наш метод не обнаружил это соотношение? Ну, никто не утверждал, что такое рекуррентное соотношение с необходимостью влечет за собой неопределенную суммируемость членов $\bar\beta_0(n)t(n,k) + \bar\beta_1(n)t(n+1,k)$. Удивительно то, что метод Госпера–Зайльбергера и во многих других случаях не находит простейшего рекуррентного соотношения.

Заметим, что найденное нами рекуррентное соотношение второго порядка может быть факторизовано:

$$\begin{aligned}&\beta_0(n) + \beta_1(n)N + \beta_2(n)N^2 =\\ &\qquad = \bigl((r-n+1)N + (r-s-n-1)\bigr)\ \bigl(\bar\beta_0(n) + \bar\beta_1(n)N\bigr)\,,\end{aligned}$$

где N — оператор сдвига из (5.145).

5.105 Положите $a = 1$ и сравните коэффициенты при z^{3n} в обеих частях тождества Хенричи (Henrici), “дружественного монстра”,

$$\begin{aligned}&f(a,z)\,f(a,\omega z)\,f(a,\omega^2 z) =\\ &\quad = F\left({\tfrac12 a-\tfrac14,\ \tfrac12 a+\tfrac14 \atop \tfrac13 a,\ \tfrac13 a+\tfrac13,\ \tfrac13 a+\tfrac23,\ \tfrac23 a-\tfrac13,\ \tfrac23 a,\ \tfrac23 a+\tfrac13,\ a}\,\middle|\,\left(\frac{4z}{9}\right)^3\right),\end{aligned}$$

где $f(a,z) = F(1;a,1;z)$. Это тождество может быть доказано, если показать, что обе части удовлетворяют одному и тому же дифференциальному уравнению.

Питер Пауле (Peter Paule) нашел еще один интересный способ вычислить эту сумму:

$$\sum_{k,l}\binom{N}{k,l,N-k-l}^2\omega^{k+2l} = \sum_{k,l}\binom{N}{k-l,l,N-k}^2\omega^{k+l} =$$

$$= \sum_{k,l}\binom{N}{k}^2\binom{k}{l}^2\omega^{k+l} =$$

$$= \sum_{k}\binom{N}{k}^2\omega^k\,[z^k]\,\bigl((1+z)(\omega+z)\bigr)^k =$$

$$= [z^0]\sum_{k}\binom{N}{k}^2\left(\frac{\omega(1+z)(\omega+z)}{z}\right)^k =$$

$$= [z^0]\sum_{k,j}\binom{N}{k}^2\binom{k}{j}\left(\frac{\omega(1+z)(\omega+z)}{z}-1\right)^j =$$

$$= [z^0]\sum_{k,j}\binom{N}{k}\binom{N-j}{N-k}\binom{N}{j}\left(\frac{(\omega z-1)^2}{\omega z}\right)^j =$$

$$= \sum_{j}\binom{2N-j}{N}\binom{N}{j}[z^j]\,(z-1)^{2j} =$$

$$= \sum_{j}\binom{2N-j}{N}\binom{N}{j}\binom{2j}{j}(-1)^j,$$

использовав биномиальную теорему, свертку Вандермонда и тот факт, что $[z^0]g(az) = [z^0]g(z)$. Теперь можно положить $N = 3n$, применить алгоритм Госпера–Зайльбергера к данной сумме S_n и получить волшебным образом рекуррентное соотношение первого порядка $(n+1)^2S_{n+1} = 4(4n+1)(4n+3)S_n$; окончательный результат следует отсюда по индукции.

Если заменить $3n$ на $3n+1$ или $3n+2$, указанная сумма становится равной нулю. В самом деле, $\sum_{k+l+m=N} t(k,l,m)\omega^{l-m}$ всегда равно нулю, когда $N \bmod 3 \neq 0$ и $t(k,l,m) = t(l,m,k)$.

5.106 (Решение Шалоша Б. Экхада (Shalosh B. Ekhad).) Положим

$$T(r,j,k) = \frac{((1+n+s)(1+r)-(1+n+r)j+(s-r)k)(j-l)j}{(l-m+n-r+s)(n+r+1)(j-r-1)(j+k)}t(r,j,k);$$

$$U(r,j,k) = \frac{(s+n+1)(k+l)k}{(l-m+n-r+s)(n+r+1)(j+k)}t(r,j,k).$$

Проверка сформулированного тождества оказывается чисто рутинной работой, а (5.32) получается путем суммирования по j и k. (Мы суммируем $T(r,j+1,k)-T(r,j,k)$ сначала по j, а затем — по k; другие члены $U(r,j,k+1)-U(r,j,k)$ суммируем сначала по k, а затем — по j.)

Кроме того, нам надо проверить (5.32) при $r = 0$. В этом случае после триномиальной перестройки оно приводится к виду $\sum_k (-1)^k \binom{n}{n+l}\binom{n+l}{k+l}\binom{s+n-k}{m} = (-1)^l \binom{n}{n+l}\binom{s}{m-n-l}$. Мы считаем, что l, m и n — целые числа и $n \geqslant 0$. Обе части очевидным образом нулевые, если не выполнено условие $n + l \geqslant 0$. В противном случае можно заменить k на $n - k$ и воспользоваться (5.24).

Заметим, что $1/nk$ — подходящий член, поскольку он равен $(n-1)!(k-1)!/n!\,k!$. $1/(n^2 - k^2)$ — также подходящий член. А вот $1/(n^2 + k^2)$ — нет.

5.107 Если бы это был подходящий член, то должен был бы найтись линейный разностный оператор, аннулирующий его. Другими словами, мы бы имели тождество с конечными суммами

$$\sum_{i=0}^{I}\sum_{j=0}^{J} \alpha_{i,j}(n)/\bigl((n+i)(k+j)+1\bigr) = 0,$$

где α — некоторые полиномы от n, не все нулевые. Выберем целые числа i, j и n так, чтобы было $n > 1$ и $\alpha_{i,j}(n) \neq 0$. Тогда при $k = -1/(n+i) - j$ член суммы с индексами (i, j) становится бесконечным, тогда как все остальные члены конечны.

5.108 Заменим в двойной сумме k на $m-k$, воспользуемся (5.28) для суммирования по k, и получим

$$A_{m,n} = \sum_j \binom{m}{j}^2 \binom{m+n-j}{m}^2,$$

после чего триномиальная перестройка (5.21) даст одну из требуемых формул.

Оказывается, достаточно трудно найти прямое доказательство того, что две симметричные суммы для $A_{m,n}$ равны. Можно, однако, доказать равенство косвенным путем, применив алгоритм Госпера–Зайльбергера, чтобы показать, что обе суммы удовлетворяют рекуррентному соотношению

$$(n+1)^3 A_{m,n} - f(m,n) A_{m,n+1} + (n+2)^3 A_{m,n+2} = 0,$$

где $f(m,n) = (2n+3)(n^2+3n+2m^2+2m+3)$. Полагая $t_1(n,k) = \binom{m}{k}\binom{n}{k}\binom{m+k}{k}\binom{n+k}{k}$ и $t_2(n,k) = \binom{m+n-k}{k}^2\binom{m+n-2k}{m-k}^2$, находим

$$\begin{aligned}(n+1)^2 t_j(n,k) - f(m,n)t_j(n+1,k) + (n+2)^2 t_j(n+2,k)\\ = T_j(n,k+1) - T_j(n,k),\end{aligned}$$

где $T_1(n,k) = -2(2n+3)k^4 t_1(n,k)/(n+1-k)(n+2-k)$ и $T_2(n,k) = -\bigl((n+2)(4mn+n+3m^2+8m+2) - 2(3mn+n+m^2+6m+2)k + (2m+1)k^2\bigr)k^2(m+n+1-k)^2 t_2(n,k)/(n+2-k)^2$. Это доказывает рекуррентное соотношение, так что остается только проверить равенство для $n = 0$ и $n = 1$. (Можно было бы также использовать более простое рекуррентное соотношение

$$m^3 A_{m,n-1} - n^3 A_{m-1,n} = (m-n)(m^2+n^2-mn)A_{m-1,n-1},$$

к которому можно прийти при помощи методов из упр. 101.)

Тот факт, что первое выражение для $A_{m,n}$ равно третьему, влечет за собой замечательное тождество между производящими функциями $\sum_{m,n} A_{m,n} w^m z^n$:

$$\sum_k \frac{w^k S_k(z)^2}{(1-z)^{2k+1}} = \sum_k \binom{2k}{k}^2 \frac{w^k}{(1-w)^{2k+1}} \frac{z^k}{(1-z)^{2k+1}},$$

где $S_k(z) = \sum_j \binom{k}{j}^2 z^j$. На самом деле оказывается, что

$$\sum_k \frac{w^k S_k(x) S_k(y)}{(1-x)^k (1-y)^k} = \sum_k \binom{2k}{k} \frac{w^k}{(1-w)^{2k+1}} \frac{\sum_j \binom{k}{j}^2 x^j y^{k-j}}{(1-x)^k(1-y)^k};$$

это частный случай тождества, открытого Бейли (Bailey) [19].

5.109 Пусть $X_n = \sum_k \binom{n}{k}^{a_0} \binom{n+k}{k}^{a_1} \ldots \binom{n+lk}{k}^{a_l} x^k$ для любых положительных целых чисел $a_0, a_1, \ldots, a_l$ и для любого целого x. Тогда если $0 \leqslant m < p$, то мы имеем

$$X_{m+pn} = \sum_{j=0}^{p-1} \sum_k \binom{m+pn}{j+pk}^{a_0} \ldots \binom{m+pn+l(j+pk)}{j+pk}^{a_l} x^{j+pk},$$

$$X_m X_n = \sum_{j=0}^{p-1} \sum_k \binom{m}{j}^{a_0} \binom{n}{k}^{a_0} \ldots \binom{m+lj}{j}^{a_l} \binom{n+lk}{k}^{a_l} x^{j+k}.$$

Соответствующие члены сравнимы $(\mathrm{mod}\ p)$, потому что из упр. 36 следует, что они кратны p при $lj+m \geqslant p$, из упр. 61 вытекает, что биномиальные коэффициенты сравнимы при $lj+m < p$, а из (4.48) — что $x^p \equiv x$.

5.110 Сравнение, несомненно, выполняется, если $2n+1$ — простое число. Стивен Скиена (Steven Skiena) нашел также пример $n = 2953$, где $2n+1 = 3 \cdot 11 \cdot 179$.

Илан Варди (Ilan Vardi) заметил, что это условие выполняется для $2n+1 = p^2$, *где* p *— простое, тогда и только тогда, когда* $2^{p-1} \bmod p^2 = 1$. *Это дает еще два примера:* $n = (1093^2-1)/2$; $n = (3511^2-1)/2$.

5.111 Частичные результаты представлены в [96]. Вычислительные эксперименты были проведены В. А. Высоцким.

5.112 Если n — не степень 2, то из упр. 36 мы знаем, что $\binom{2n}{n}$ кратно 4. Для противного случая указанное свойство было проверено для всех значений $n \leqslant 2^{22000}$ Э. Гранвилем (A. Granville) и О. Рамаре (O. Ramaré), которые также усилили теорему Саркёзи (Sárközy) [317], показав, что $\binom{2n}{n}$ делится на квадрат простого числа при всех $n > 2^{22000}$. Это подтверждает давно выдвинутую гипотезу о том, что $\binom{2n}{n}$ никогда не свободно от квадратов при $n > 4$.

Аналогичные предположения для кубов таковы: $\binom{2n}{n}$ делится на куб простого числа для всех $n > 1056$ и либо на 2^3, либо на 3^3 для всех $n > 2^{29} + 2^{23}$. Эти гипотезы были проверены для всех $n < 2^{10000}$. Поль Эрдёш (Paul Erdős) предположил, что на самом деле $\max_p \epsilon_p\bigl(\binom{2n}{n}\bigr) \to \infty$ при $n \to \infty$; это должно быть справедливо, даже если мы ограничим p значениями 2 и 3.

5.113 Теорема о производящих функциях из упр. 7.20 может помочь в ответе на этот вопрос.

5.114 Штрел (Strehl) [344] показал, что $c_n^{(2)} = \sum_k \binom{n}{k}^3 = \sum_k \binom{n}{k}^2\binom{2k}{n}$ представляют собой так называемые числа Фрейнела [132] и что $c_n^{(3)} = \sum_k \binom{n}{k}^2\binom{2k}{k}^2\binom{2k}{n-k}$. В другом направлении Г. С. Вильф (H. S. Wilf) показал, что $c_n^{(m)}$ — целые для всех m, когда $n \leqslant 9$.

6.1 2314, 2431, 3241, 1342, 3124, 4132, 4213, 1423, 2143, 3412, 4321.

6.2 Ровно $\left\{ n \atop k \right\} m^{\underline{k}}$, потому что область определения всякой такой функции разбивается на k непустых подмножеств, а для каждого разбиения имеется $m^{\underline{k}}$ способов приписать функции значения. (Суммирование по k дает комбинаторное доказательство формулы (6.10).)

6.3 В этом случае имеем $d_{k+1} \leqslant$ (центр тяжести) $- \epsilon = 1 - \epsilon + (d_1 + \cdots + d_k)/k$. Это рекуррентное соотношение подобно (6.55), но с $1 - \epsilon$ вместо 1; следовательно, оптимальным решением является $d_{k+1} = (1-\epsilon)H_k$. Эта величина не ограничена при $\epsilon < 1$.

6.4 $H_{2n+1} - \frac{1}{2}H_n$. (Аналогично $\sum_{k=1}^{2n}(-1)^{k-1}/k = H_{2n} - H_n$.)

6.5 Величина $U_n(x, y)$ равна

$$x\sum_{k\geqslant 1}\binom{n}{k}(-1)^{k-1}k^{-1}(x+ky)^{n-1} + y\sum_{k\geqslant 1}\binom{n}{k}(-1)^{k-1}(x+ky)^{n-1},$$

и первая сумма есть

$$U_{n-1}(x, y) + \sum_{k\geqslant 1}\binom{n-1}{k-1}(-1)^{k-1}k^{-1}(x+ky)^{n-1}.$$

Лишнее k^{-1} может быть внесено в биномиальный коэффициент, и мы получим, что

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

Тем самым доказывается соотношение (6.75). Пусть $R_n(x,y) = x^{-n}U_n(x,y)$; тогда $R_0(x,y)=0$ и $R_n(x,y) = R_{n-1}(x,y)+1/n+y/x$, следовательно, $R_n(x,y) = H_n + ny/x$. (Между прочим, первоначальная сумма $U_n = U_n(n,-1)$ не приводит к рекуррентному соотношению, подобному этому, так что проще вычислить индуктивно более общую сумму, которая лишена зависимости x от n, нежели ее частный случай. Вот еще один поучительный пример, когда сильное индуктивное предположение существенно меняет исход дела.)

6.6 Каждая пара крольчат ьь, рожденных в конце одного месяца, становится парой половозрелых кроликов **аа** в конце следующего месяца; каждая пара **аа** становится парой **аа** и ьь. Таким образом, каждая пара ьь ведет себя наподобие трутня в родословном дереве, а каждая пара **аа** — наподобие матки, за тем исключением, что родословное дерево пчел обращено в прошлое, а родословное дерево кроликов — в будущее. Через n месяцев имеется в наличии F_{n+1} пар кроликов, из которых F_n — половозрелые, а F_{n-1} — крольчата. (Именно в этом контексте Фибоначчи ввел в употребление свои числа.)

Рекуррентное соотношение Фибоначчи — аддитивное, а кролики почему-то множатся...

6.7 (а) Положите $k = 1-n$ и примените (6.107). (б) Положите $m = 1$ и $k = n-1$ и примените (6.128).

6.8 $55+8+2$ превращается в $89+13+3 = 105$; точное значение — 104.60736.

Это "точное значение" представляет собой длину 65 международных миль, но международная миля равна .999998 американской мили. В 3937 американских милях содержится ровно 6336 километров; метод Фибоначчи преобразует 3937 в 6370.

6.9 21. (При возведении единиц измерения в квадрат мы переходим от F_n к F_{n+2}. Точное значение — приблизительно 20.72.)

6.10 Все неполные частные $a_0, a_1, a_2, \ldots$ равны 1, потому что $\phi = 1+1/\phi$. (Так что представление Штерна–Броко данной величины имеет вид RLRLRLRLRL... .)

6.11 $(-1)^{\overline{n}} = [n=0]-[n=1]$; см. (6.11).

6.12 Это следствие (6.31) и дуальной формулы из табл. 325.

6.13 Согласно упр. 12 эти две формулы эквивалентны. Можно воспользоваться индукцией, а можно заметить, что z^nD^n применительно к $f(z) = z^x$ дает $x^{\underline{n}}z^x$, в то время как ϑ^n применительно к той же самой функции дает $x^n z^x$; поэтому последовательность $\langle \vartheta^0, \vartheta^1, \vartheta^2, \ldots \rangle$ должна быть связана с $\langle z^0 D^0, z^1 D^1, z^2 D^2, \ldots \rangle$ так же, как последовательность $\langle x^0, x^1, x^2, \ldots \rangle$ связана с $\langle x^{\underline{0}}, x^{\underline{1}}, x^{\underline{2}}, \ldots \rangle$.

6.14 Имеем

$$x\binom{x+k}{n} = (k+1)\binom{x+k}{n+1} + (n-k)\binom{x+k+1}{n+1},$$

потому что $(n+1)x = (k+1)(x+k-n) + (n-k)(x+k+1)$. (Достаточно проверить последнее тождество при $k=0$, $k=-1$ и $k=n$.)

6.15 Поскольку $\Delta\bigl(\binom{x+k}{n}\bigr) = \binom{x+k}{n-1}$, общая формула такова:

$$\sum_k \left\langle {n \atop k} \right\rangle \binom{x+k}{n-m} = \Delta^m(x^n) = \sum_j \binom{m}{j}(-1)^{m-j}(x+j)^n.$$

Положите $x=0$ и обратитесь к (6.20).

6.16 $A_{n,k} = \sum_{j\geqslant 0} a_j \left\{ {n-j \atop k} \right\}$; данная сумма всегда конечна.

6.17 (а) $\left|{n \atop k}\right| = \left[{n+1 \atop n+1-k}\right]$. (б) $\left|{n \atop k}\right| = n^{\underline{n-k}} = n!\,[n\geqslant k]/k!$. (в) $\left|{n \atop k}\right| = k!\left\{{n \atop k}\right\}$.

6.18 Это эквивалентно правилу (6.3) или рекуррентному соотношению (6.8).

6.19 Воспользуйтесь табл. 334.

6.20 $\sum_{1\leqslant j\leqslant k\leqslant n} 1/j^2 = \sum_{1\leqslant j\leqslant n}(n+1-j)/j^2 = (n+1)H_n^{(2)} - H_n$.

6.21 Указанное число — это сумма дробей с нечетными знаменателями, так что оно имеет вид a/b с нечетным b. (Между прочим, из постулата Бертрана вытекает, что b_n также делится по меньшей мере на одно простое число, когда $n > 2$.)

6.22 $\bigl|z/k(k+z)\bigr| \leqslant 2|z|/k^2$ при $k > 2|z|$, так что эта сумма определена, когда ее знаменатели не равны нулю. Если $z = n$, то $\sum_{k=1}^{m}\bigl(1/k - 1/(k+n)\bigr) = H_m - H_{m+n} + H_n$, что стремится к H_n при $m \to \infty$. (Величина $H_{z-1} - \gamma$ часто называется пси-функцией $\psi(z)$.)

6.23 $z/(e^z+1) = z/(e^z-1) - 2z/(e^{2z}-1) = \sum_{n\geqslant 0}(1-2^n)B_n z^n/n!$.

6.24 Если n — нечетное число, то $T_n(x)$ является полиномом от x^2, следовательно, когда мы берем производную и вычисляем $T_{n+1}(x)$ по формуле (6.95), его коэффициенты умножаются на четные числа. (На самом деле можно доказать большее: согласно упр. 54 в знаменателе числа Бернулли B_{2n} всегда содержится

2 в первой степени, следовательно, $2^{2n-k}\backslash\backslash T_{2n+1} \iff 2^k\backslash\backslash(n+1)$. Нечетные положительные целые числа $(n+1)T_{2n+1}/2^{2n}$ называются числами Дженоччи $\langle 1,1,3,17,155,2073,\ldots\rangle$ в честь Дженоччи (Genocchi) [145].)

Само собой, вездесущий Эйлер знал о числах Дженоччи задолго до того, как тот появился на свет: см. [110], том 2, глава 7, §181.

6.25 $100n - nH_n < 100(n-1) - (n-1)H_{n-1} \iff H_{n-1} > 99$. (Наименьшее такое n приблизительно равно $e^{99-\gamma}$, в то время как на весь путь червяк затрачивает время $N \approx e^{100-\gamma}$, т.е. примерно в e раз больше. Таким образом, червяк становится ближе к цели своего путешествия на протяжении последних 63% пути.)

6.26 Пусть $u(k) = H_{k-1}$ и $\Delta v(k) = 1/k$, так что $u(k) = v(k)$. Тогда $S_n - H_n^{(2)} = \sum_{k=1}^n H_{k-1}/k = H_{k-1}^2\big|_1^{n+1} - S_n = H_n^2 - S_n$.

6.27 Заметьте, что если $m>n$, то НОД$(F_m, F_n)=$НОД(F_{m-n}, F_n) согласно (6.108). Это позволяет доказать требуемое по индукции.

6.28 (a) $Q_n = \alpha(L_n - F_n)/2 + \beta F_n$. (Решение можно также записать в виде $Q_n = \alpha F_{n-1} + \beta F_n$.) (б) $L_n = \phi^n + \hat{\phi}^n$.

6.29 При $k = 0$ это правило (6.133). При $k = 1$ это, в сущности, соотношение

$$\begin{aligned} K(x_1,\ldots,x_n)x_m &= K(x_1,\ldots,x_m)\,K(x_m,\ldots,x_n) - \\ &\quad - K(x_1,\ldots,x_{m-2})\,K(x_{m+2},\ldots,x_n)\,; \end{aligned}$$

на языке азбуки Морзе второе произведение справа отбрасывает те случаи, где первое произведение содержит накладывающиеся тире. При $k > 1$ достаточно индукции по k с использованием (6.127) и (6.132). (Это тождество справедливо также в тех случаях, когда один или более индексов при K становятся равными -1, если условиться, что $K_{-1} = 0$. Когда умножение некоммутативно, тождество Эйлера остается в силе для $k = n-1$, если записать его в виде

$$\begin{aligned} &K_{m+n}(x_1,\ldots,x_{m+n})\,K_{n-1}(x_{m+n-1},\ldots,x_{m+1}) = \\ &\quad = K_{m+n-1}(x_1,\ldots,x_{m+n-1})\,K_n(x_{m+n},\ldots,x_{m+1}) - \\ &\quad\quad - (-1)^n K_{m-1}(x_1,\ldots,x_{m-1})\,. \end{aligned}$$

Например, в случае $m = 0$, $n = 3$ мы получаем несколько неожиданные некоммутативные разложения

$$(abc + a + c)(1 + ba) = (ab + 1)(cba + a + c).$$

6.30 Производная $K(x_1,\ldots,x_n)$ по x_m представляет собой

$$K(x_1,\ldots,x_{m-1})\,K(x_{m+1},\ldots,x_n)\,,$$

а вторая производная равна нулю; следовательно, ответ таков:

$$K(x_1,\ldots,x_n)+K(x_1,\ldots,x_{m-1})\,K(x_{m+1},\ldots,x_n)y\,.$$

$\left|{n \atop k}\right| = \left|{-k \atop -n}\right|$.

6.31 Поскольку $x^{\overline{n}} = (x+n-1)^{\underline{n}} = \sum_k \binom{n}{k} x^{\underline{k}}(n-1)^{\underline{n-k}}$, имеем $\left|{n \atop k}\right| = \binom{n}{k}(n-1)^{\underline{n-k}}$. Эти коэффициенты, как оказывается, удовлетворяют рекуррентному соотношению

$$\left|{n \atop k}\right| = (n-1+k)\left|{n-1 \atop k}\right| + \left|{n-1 \atop k-1}\right|, \qquad \text{целые } n,k>0.$$

6.32 $\sum_{k\leqslant m} k\left\{{n+k \atop k}\right\} = \left\{{m+n+1 \atop m}\right\}$ и $\sum_{0\leqslant k\leqslant n}\left\{{k \atop m}\right\}(m+1)^{n-k} = \left\{{n+1 \atop m+1}\right\}$, они оба приводятся в табл. 326.

6.33 Если $n>0$, то согласно (6.71) $\left[{n \atop 3}\right] = \frac12(n-1)!\,(H_{n-1}^2 - H_{n-1}^{(2)})$, а согласно (6.20) $\left\{{n \atop 3}\right\} = \frac16(3^n - 3\cdot 2^n + 3)$.

6.34 Имеем $\left\langle{-1 \atop k}\right\rangle = 1/(k+1)$, $\left\langle{-2 \atop k}\right\rangle = H_{k+1}^{(2)}$, и вообще при любом целом n величина $\left\langle{n \atop k}\right\rangle$ задается соотношением (6.38).

6.35 Выберите в качестве n наименьшее целое $>1/\epsilon$, такое, что $\lfloor H_n\rfloor > \lfloor H_{n-1}\rfloor$.

6.36 В этом случае $d_{k+1} = \bigl(100+(1+d_1)+\cdots+(1+d_k)\bigr)/(100+k)$, и решением является $d_{k+1} = H_{k+100} - H_{101} + 1$ при $k\geqslant 1$. Это значение превышает 2 при $k \geqslant 176$.

6.37 Суммирование (по частям) дает $H_{mn} - \bigl(\frac{m}{m} + \frac{m}{2m} + \cdots + \frac{m}{mn}\bigr) = H_{mn} - H_n$. Поэтому бесконечная сумма равна $\ln m$. (Отсюда следует, что

$$\sum_{k\geqslant 1}\frac{\nu_m(k)}{k(k+1)} = \frac{m}{m-1}\ln m,$$

потому что $\nu_m(k) = (m-1)\sum_{j\geqslant 1}(k \bmod m^j)/m^j$.)

6.38 $(-1)^k\bigl(\binom{r-1}{k}r^{-1} - \binom{r-1}{k-1}H_k\bigr) + C$. (По частям, с использованием (5.16).)

6.39 Запишите ее в виде $\sum_{1\leqslant j\leqslant n} j^{-1}\sum_{j\leqslant k\leqslant n} H_k$ и, исходя из (6.67), просуммируйте вначале по k, получив

$$(n+1)H_n^2 - (2n+1)H_n + 2n\,.$$

6.40 Если $6n-1$ — простое число, числитель суммы

$$\sum_{k=1}^{4n-1}\frac{(-1)^{k-1}}{k} = H_{4n-1} - H_{2n-1}$$

делится на $6n-1$, потому что эта сумма представляет собой

$$\sum_{k=2n}^{4n-1} \frac{1}{k} = \sum_{k=2n}^{3n-1} \left(\frac{1}{k} + \frac{1}{6n-1-k}\right) = \sum_{k=2n}^{3n-1} \frac{6n-1}{k(6n-1-k)}.$$

Аналогично, если $6n+1$ является простым числом, то числитель суммы $\sum_{k=1}^{4n}(-1)^{k-1}/k = H_{4n} - H_{2n}$ кратен $6n+1$. Для 1987 следует суммировать до $k = 1324$, а для 2011 — до 1340.

6.41 $S_{n+1} = \sum_k \binom{\lfloor (n+1+k)/2 \rfloor}{k} = \sum_k \binom{\lfloor (n+k)/2 \rfloor}{k-1}$, следовательно, $S_{n+1} + S_n = \sum_k \binom{\lfloor (n+k)/2+1 \rfloor}{k} = S_{n+2}$. Ответ: F_{n+2}.

6.42 F_n.

6.43 Положив $z = \frac{1}{10}$ в $\sum_{n\geqslant 0} F_n z^n = z/(1-z-z^2)$, получим $\frac{10}{89}$. Данная сумма представляет собой периодическую десятичную дробь с длиной периода 44:

0.11235 95505 61797 75280 89887 64044 94382 02247 19101 12359 55+ .

6.44 Замените при необходимости (m, k) либо на $(-m, -k)$, либо на $(k, -m)$, либо на $(-k, m)$ так, чтобы было $m \geqslant k \geqslant 0$. Если $m = k$, то все ясно. Если же $m > k$, то можно заменить (m, k) на $(m-k, m)$ и воспользоваться индукцией.

6.45 $X_n = A(n)\alpha + B(n)\beta + C(n)\gamma + D(n)\delta$, где $B(n) = F_n$, $A(n) = F_{n-1}$, $A(n) + B(n) - D(n) = 1$ и $B(n) - C(n) + 3D(n) = n$.

6.46 $\phi/2$ и $\phi^{-1}/2$. Пусть $u = \cos 72°$ и $v = \cos 36°$; тогда $u = 2v^2 - 1$ и $v = 1 - 2\sin^2 18° = 1 - 2u^2$. Следовательно, $u + v = 2(u+v)(v-u)$, и $4v^2 - 2v - 1 = 0$. Продолжая, можно найти пять комплексных корней пятой степени из единицы:

$$1, \quad \frac{\phi^{-1} \pm i\sqrt{2+\phi}}{2}, \quad \frac{-\phi \pm i\sqrt{3-\phi}}{2}.$$

6.47 $2^n\sqrt{5}F_n = (1+\sqrt{5})^n - (1-\sqrt{5})^n$, и четные степени $\sqrt{5}$ сокращаются. Пусть теперь p — нечетное простое число. Тогда $\binom{p}{2k+1} \equiv 0$ за исключением случая, когда $k = (p-1)/2$, и $\binom{p+1}{2k+1} \equiv 0$ за исключением случая, когда $k = 0$ или $k = (p-1)/2$; следовательно, $F_p \equiv 5^{(p-1)/2}$ и $2F_{p+1} \equiv 1 + 5^{(p-1)/2} \pmod p$. Можно показать, что $5^{(p-1)/2} \equiv 1$, когда p имеет вид $10k \pm 1$, и $5^{(p-1)/2} \equiv -1$, когда p имеет вид $10k \pm 3$.

6.48 Пусть $K_{i,j} = K_{j-i+1}(x_i, \ldots, x_j)$. Многократное применение (6.133) сводит обе части к $(K_{1,m-2}(x_{m-1} + x_{m+1}) + K_{1,m-3}) \times K_{m+2,n} + K_{1,m-2}K_{m+3,n}$.

6.49 Положите $z = \frac{1}{2}$ в (6.146); неполные частные — это 0, 2^{F_0}, 2^{F_1}, 2^{F_2}, ... (Как отмечал Кнут [206], это число является трансцендентным.)

6.50 (a) $f(n)$ четное $\iff 3\backslash n$. (б) Если двоичное представление n есть $(1^{a_1}0^{a_2}\ldots 1^{a_{m-1}}0^{a_m})_2$, где m четное, то $f(n) = K(a_1, a_2, \ldots, a_{m-1})$.

6.51 (a) Комбинаторное доказательство. Представление множества $\{1, 2, \ldots, p\}$ в виде k подмножеств или циклов делится на “орбиты” из 1 или p представлений каждая, если к каждому элементу добавить 1 по модулю p, например

$$\{1,2,4\} \cup \{3,5\} \to \{2,3,5\} \cup \{4,1\} \to \{3,4,1\} \cup \{5,2\} \to$$
$$\to \{4,5,2\} \cup \{1,3\} \to \{5,1,3\} \cup \{2,4\} \to \{1,2,4\} \cup \{3,5\}.$$

Орбита размером 1 получается только тогда, когда это преобразование переводит представление само в себя; но тогда $k = 1$ или $k = p$. С другой стороны, можно дать алгебраическое доказательство: имеем $x^{\overline{p}} \equiv x^{\underline{p}} + x^{\underline{1}}$ и $x^{\underline{p}} \equiv x^p - x \pmod p$, поскольку согласно теореме Ферма $x^p - x$ делится на $(x-0)(x-1)\ldots\bigl(x-(p-1)\bigr)$.

(б) Этот результат следует из п. (a) и теоремы Вильсона; можно также воспользоваться тем, что $x^{\overline{p-1}} \equiv x^{\overline{p}}/(x-1) \equiv (x^p - x)/(x-1) = x^{p-1} + x^{p-2} + \cdots + x$.

(в) Имеем $\left\{ {p+1 \atop k} \right\} \equiv \left[{p+1 \atop k} \right] \equiv 0$ при $3 \leqslant k \leqslant p$, затем — $\left\{ {p+2 \atop k} \right\} \equiv \left[{p+2 \atop k} \right] \equiv 0$ при $4 \leqslant k \leqslant p$ и т.д. (Аналогично $\left[{2p-1 \atop p} \right] \equiv -\left\{ {2p-1 \atop p} \right\} \equiv 1$.)

(г) $p! = p^{\overline{p}} = \sum_k (-1)^{p-k} p^k \left[{p \atop k} \right] = p^p \left[{p \atop p} \right] - p^{p-1} \left[{p \atop p-1} \right] + \cdots + p^3 \left[{p \atop 3} \right] - p^2 \left[{p \atop 2} \right] + p \left[{p \atop 1} \right]$. Но $p \left[{p \atop 1} \right] = p!$, так что

$$\begin{bmatrix} p \\ 2 \end{bmatrix} = p \begin{bmatrix} p \\ 3 \end{bmatrix} - p^2 \begin{bmatrix} p \\ 4 \end{bmatrix} + \cdots + p^{p-2} \begin{bmatrix} p \\ p \end{bmatrix}$$

кратно p^2. (Это называется теоремой Вольштенхольма (Wolstenholme).)

6.52 (a) Заметьте, что $H_n = H_n^* + H_{\lfloor n/p \rfloor}/p$, где $H_n^* = \sum_{k=1}^{n} [k \perp p]/k$. (б) Действуя по mod 5, получаем $H_r = \langle 0, 1, 4, 1, 0 \rangle$ при $0 \leqslant r \leqslant 4$. Таким образом, первым решением служит $n = 4$. Из п. (a) известно, что $5\backslash a_n \implies 5\backslash a_{\lfloor n/5 \rfloor}$; поэтому следующим возможным интервалом служит $n = 20 + r$, $0 \leqslant r \leqslant 4$, где мы имеем $H_n = H_n^* + \frac{1}{5}H_4 = H_{20}^* + \frac{1}{5}H_4 + H_r - \sum_{k=1}^{r} 20/k(20+k)$. Числитель дроби H_{20}^*, как и числитель дроби H_4, делится на 25. Следовательно, в этом интервале есть только два решения: $n = 20$ и $n = 24$. Следующим возможным интервалом является $n = 100 + r$; в этом случае $H_n = H_n^* + \frac{1}{5}H_{20}$, что равно

$\frac{1}{5}H_{20} + H_r$ плюс некая дробь, числитель которой кратен 5. Если $\frac{1}{5}H_{20} \equiv m \pmod 5$, где m — целое число, то гармоническое число H_{100+r} будет иметь числитель, который делится на 5 тогда и только тогда, когда $m + H_r \equiv 0 \pmod 5$; следовательно, m должно быть $\equiv 0$, 1 или 4. Действуя по модулю 5, найдем, что $\frac{1}{5}H_{20} = \frac{1}{5}H^*_{20} + \frac{1}{25}H_4 \equiv \frac{1}{25}H_4 = \frac{1}{12} \equiv 3$; следовательно, при $100 \leqslant n \leqslant 104$ решений нет. Аналогично нет решений и при $120 \leqslant n \leqslant 124$; итак, мы нашли все три решения.

(Согласно упр. 6.51, г мы всегда получаем, что $p^2\backslash a_{p-1}$, $p\backslash a_{p^2-p}$ и $p\backslash a_{p^2-1}$, если p — некоторое простое число $\geqslant 5$. Только что приведенное рассуждение показывает, что этим исчерпываются все решения для $p\backslash a_n$ тогда и только тогда, когда не существует решения для $p^{-2}H_{p-1} + H_r \equiv 0 \pmod p$ при $0 \leqslant r < p$. Последнее условие справедливо не только при $p = 5$, но и при $p = 13$, 17, 23, 41 и 67, а может, и для бесконечного количества простых чисел. Числитель H_n делится на 3, только когда $n = 2$, 7 и 22; он делится на 7, только когда $n = 6$, 42, 48, 295, 299, 337, 341, 2096, 2390, 14675, 16731, 16735 и 102728. См. ответ к упр. 92.)

Вниманию программистов: вот интересное условие для проверки на стольких простых числах, сколько вы сможете.

6.53 Суммирование по частям дает

$$\frac{n+1}{(n+2)^2}\left(\frac{(-1)^m}{\binom{n+1}{m+1}}\big((n+2)H_{m+1} - 1\big) - 1\right).$$

6.54 (а) Если $m \geqslant p$, то $S_m(p) \equiv S_{m-(p-1)}(p) \pmod p$, поскольку $k^{p-1} \equiv 1$ при $1 \leqslant k < p$. Кроме того, $S_{p-1}(p) \equiv p - 1 \equiv -1$. Если $0 < m < p - 1$, то можно записать, что

$$S_m(p) = \sum_{k=0}^{p-1}\sum_{j=0}^{m}\left\{{m \atop j}\right\}k^{\underline{j}} = \sum_{j=0}^{m}\left\{{m \atop j}\right\}\frac{p^{\underline{j+1}}}{j+1} \equiv 0.$$

(б) Обстоятельство, содержащееся в указании к упражнению, означает, что знаменатель числа I_{2n} не делится ни на какое простое p; следовательно, I_{2n} должно быть целым числом. Чтобы доказать утверждение из указания, можно считать, что $n>1$. Тогда

$$B_{2n} + \frac{[(p-1)\backslash(2n)]}{p} + \sum_{k=0}^{2n-2}\binom{2n+1}{k}B_k\frac{p^{2n-k}}{2n+1}$$

представляет собой целое число согласно (6.78), (6.84) и п. (а). Так что надо убедиться, что ни один из знаменателей $\binom{2n+1}{k} \times B_k p^{2n-k}/(2n+1) = \binom{2n}{k}B_k p^{2n-k}/(2n-k+1)$ не делится на p. Знаменатель числа $\binom{2n}{k}B_k p$ не делится на p, поскольку в знаменателе B_k не содержится p^2 (по индукции), а знаменатель $p^{2n-k-1}/(2n-k+1)$ не делится на p, поскольку $2n-k+1 < p^{2n-k}$

(Числители чисел Бернулли играли важную роль в ранних исследованиях большой теоремы Ферма; см. Рибенбойм (Ribenboim) [308].)

при $k \leqslant 2n-2$. QED (Числа I_{2n} табулированы в [224]. Эрмит (Hermite) вычислил I_{2n} вплоть до I_{18} в 1875 году [184]. Оказывается, что $I_2 = I_4 = I_6 = I_8 = I_{10} = I_{12} = 1$; следовательно, на самом деле для чисел Бернулли, приведенных в тексте главы, включая $\frac{-691}{2730}$(!), все же *имеется* "простая" закономерность. Но при $2n > 12$ числа I_{2n}, похоже, не обладают сколь-нибудь примечательными свойствами. Например, $B_{24} = -86579 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} - \frac{1}{7} - \frac{1}{13}$, и число 86579 — простое.)

(в) Числа $2-1$ и $3-1$ всегда делят $2n$. Если n — простое число, то делителями $2n$ являются только 1, 2, n и $2n$, так что при простом $n > 2$ знаменатель числа B_{2n} будет равен 6, если только число $2n+1$ также не является простым. В последнем случае можно испытывать числа $4n+3$, $8n+7$, $\ldots$ до тех пор, пока в конце концов не попадется непростое число (поскольку n делит $2^{n-1}n + 2^{n-1} - 1$). (Для этого доказательства не требуется более трудная, но справедливая теорема о том, что существует бесконечно много простых чисел вида $6k+1$.) Число 6 также может быть знаменателем B_{2n} и для непростых значений n, таких как 49.

6.55 Указанная сумма равна $\frac{m+1}{x+m+1}\binom{x+n}{n}\binom{n}{m+1}$ согласно правилу свертки Вандермонда. Чтобы получить (6.70), продифференцируйте и положите $x = 0$.

6.56 Сперва замените k^{n+1} на $\bigl((k-m)+m\bigr)^{n+1}$ и разложите по степеням $k-m$; результат упростится, как при выводе (6.72). Если $m > n$ или $m < 0$, то ответ представляет собой $(-1)^n n! - m^n/\binom{n-m}{n}$. В противном случае необходимо взять предел (5.41) при $x \to -m$ за вычетом члена при $k = m$; в результате получится $(-1)^n n! + (-1)^{m+1}\binom{n}{m} m^n (n + 1 + mH_{n-m} - mH_m)$.

6.57 Сначала докажите по индукции, что n-я строка содержит самое большее три различные величины $A_n \geqslant B_n \geqslant C_n$; если n четное, то они располагаются в циклическом порядке $[C_n, B_n, A_n, B_n, C_n]$, а если n нечетное, то в циклическом порядке $[C_n, B_n, A_n, A_n, B_n]$. Кроме того,

$$\begin{aligned} A_{2n+1} &= A_{2n} + B_{2n}\,; & A_{2n} &= 2A_{2n-1}\,;\\ B_{2n+1} &= B_{2n} + C_{2n}\,; & B_{2n} &= A_{2n-1} + B_{2n-1}\,;\\ C_{2n+1} &= 2C_{2n}\,; & C_{2n} &= B_{2n-1} + C_{2n-1}\,. \end{aligned}$$

Отсюда следует, что $Q_n = A_n - C_n = F_{n+1}$. (Относительно обернутых биномиальных коэффициентов порядка 3 см. упр. 5.75.)

6.58 (а) $\sum_{n\geqslant 0} F_n^2 z^n = z(1-z)/(1+z)(1-3z+z^2) = \frac{1}{5}\bigl((2-3z)/(1-3z+z^2) - 2/(1+z)\bigr)$. (Возведите в квадрат формулу Бине

(6.123) и просуммируйте по n, затем сгруппируйте члены так, чтобы ϕ и $\hat\phi$ исчезли.) (б) Аналогично

$$\sum_{n\geqslant 0} F_n^3 z^n = \frac{z(1-2z-z^2)}{(1-4z-z^2)(1+z-z^2)} = \frac{1}{5}\left(\frac{2z}{1-4z-z^2}+\frac{3z}{1+z-z^2}\right).$$

Отсюда следует, что $F_{n+1}^3 - 4F_n^3 - F_{n-1}^3 = 3(-1)^n F_n$. (Соответствующее рекуррентное соотношение для m-х степеней включает фибоначчиевы коэффициенты из упр. 86; оно было открыто Джарденом (Jarden) и Моцкиным (Motzkin) [194].)

6.59 Пусть m — фиксированное число. Можно доказать по индукции по n, что в самом деле можно найти такое x с дополнительным условием $x \not\equiv 2 \pmod 4$. Если x является таким решением, то можно перейти к решению по модулю 3^{n+1}, потому что

$$F_{8\,3^{n-1}} \equiv 3^n, \qquad F_{8\cdot 3^{n-1}-1} \equiv 3^n+1 \pmod{3^{n+1}};$$

подойдет либо x, либо $x+8\cdot 3^{n-1}$, либо $x+16\cdot 3^{n-1}$.

6.60 Единственно возможные случаи — F_1+1, F_2+1, F_3+1, F_4-1 и F_6-1. В противном случае в разложениях возникают числа Люка из упр. 28:

$$\begin{aligned} F_{2m}+(-1)^m &= L_{m+1}F_{m-1}; & F_{2m+1}+(-1)^m &= L_m F_{m+1};\\ F_{2m}-(-1)^m &= L_{m-1}F_{m+1}; & F_{2m+1}-(-1)^m &= L_{m+1}F_m. \end{aligned}$$

(В общем случае $F_{m+n}-(-1)^n F_{m-n} = L_m F_n$.)

6.61 $1/F_{2m} = F_{m-1}/F_m - F_{2m-1}/F_{2m}$, когда m четное и положительное. Вторая сумма равна $5/4 - F_{3\cdot 2^n-1}/F_{3\cdot 2^n}$ при $n \geqslant 1$.

6.62 (а) $A_n = \sqrt5\,A_{n-1} - A_{n-2}$ и $B_n = \sqrt5\,B_{n-1} - B_{n-2}$. Кстати, кроме того, $\sqrt5\,A_n + B_n = 2A_{n+1}$ и $\sqrt5\,B_n - A_n = 2B_{n-1}$. (б) Таблица начальных значений показывает, что

$$A_n = \begin{cases} L_n, & n \text{ четное};\\ \sqrt5\,F_n, & n \text{ нечетное}; \end{cases} \qquad B_n = \begin{cases} \sqrt5\,F_n, & n \text{ четное};\\ L_n, & n \text{ нечетное}. \end{cases}$$

(в) $B_n/A_{n+1} - B_{n-1}/A_n = 1/(F_{2n+1}+1)$, потому что $B_nA_n - B_{n-1}A_{n+1} = \sqrt5$ и $A_nA_{n+1} = \sqrt5\,(F_{2n+1}+1)$. Обратите внимание, что $B_n/A_{n+1} = (F_n/F_{n+1})[n \text{ четное}] + (L_n/L_{n+1})[n \text{ нечетное}]$. (г) Аналогично $\sum_{k=1}^{n} 1/(F_{2k+1}-1) = (A_0/B_1 - A_1/B_2) + \cdots + (A_{n-1}/B_n - A_n/B_{n+1}) = 2 - A_n/B_{n+1}$. Эту величину можно записать и как $(5F_n/L_{n+1})[n \text{ четное}] + (L_n/F_{n+1})[n \text{ нечетное}]$.

6.63 (а) $\left[{n \atop k}\right]$. Всего имеется $\left[{n-1 \atop k-1}\right]$ перестановок с $\pi_n = n$ и $(n-1)\left[{n-1 \atop k}\right]$ перестановок с $\pi_n < n$. (б) $\left\langle{n \atop k}\right\rangle$. Каждая перестановка $\rho_1 \ldots \rho_{n-1}$ множества $\{1, \ldots, n-1\}$ приводит к n перестановкам $\pi_1\pi_2 \ldots \pi_n = \rho_1 \ldots \rho_{j-1}\, n\, \rho_{j+1} \ldots \rho_{n-1}\rho_j$. Если $\rho_1 \ldots \rho_{n-1}$ содержит k превышений, то имеется $k+1$ значений j, которые дают k превышений в перестановке $\pi_1\pi_2 \ldots \pi_n$; остальные $n-1-k$ значений дают $k+1$. Следовательно, общее количество способов получения k превышений в перестановке $\pi_1\pi_2 \ldots \pi_n$ равно $(k+1)\left\langle{n-1 \atop k}\right\rangle + \bigl((n-1)-(k-1)\bigr)\left\langle{n-1 \atop k-1}\right\rangle = \left\langle{n \atop k}\right\rangle$.

6.64 В соответствии с доказанным в упр. 5.72 знаменатель дроби $\binom{1/2}{2n}$ равен $2^{4n-\nu_2(n)}$. Согласно (6.44) дробь $\left[{1/2 \atop 1/2-n}\right]$ имеет тот же знаменатель, поскольку $\left\langle\!\left\langle{n \atop 0}\right\rangle\!\right\rangle = 1$ и $\left\langle\!\left\langle{n \atop k}\right\rangle\!\right\rangle$ четное при $k > 0$.

6.65 Это равносильно утверждению, что $\left\langle{n \atop k}\right\rangle/n!$ представляет собой вероятность того, что $\lfloor x_1 + \cdots + x_n \rfloor = k$, если $x_1, \ldots, x_n$ — независимые случайные величины, равномерно распределенные между 0 и 1. Пусть $y_j = (x_1 + \cdots + x_j) \bmod 1$. Тогда $y_1, \ldots, y_n$ независимы и равномерно распределены, а $\lfloor x_1 + \cdots + x_n \rfloor$ — это количество спадов в последовательности y. Перестановка y также случайна, и вероятность k спадов такая же, как и вероятность k подъемов.

6.66 $2^{n+1}(2^{n+1}-1)B_{n+1}/(n+1)$, если $n > 0$. (См. (7.56) и (6.92); требуемые числа являются, в сущности, коэффициентами разложения $1 - \operatorname{th} z$.)

6.67 Эта сумма — $\sum_k \left(\left\{{n \atop k+1}\right\}(k+1)! + \left\{{n \atop k}\right\}k!\right)\binom{n-k}{n-m}(-1)^{m-k} = \sum_k \left\{{n \atop k}\right\}k!(-1)^{m-k}\left(\binom{n-k}{n-m} - \binom{n+1-k}{n-m}\right) = \sum_k \left\{{n \atop k}\right\}k!(-1)^{m+1-k} \times \binom{n-k}{n-m-1} = \left\langle{n \atop n-m-1}\right\rangle$ в соответствии с (6.3) и (6.40). Теперь примените (6.34). (Это тождество допускает и комбинаторное истолкование [59].)

6.68 Имеет место общая формула

$$\left\langle\!\!\left\langle{n \atop m}\right\rangle\!\!\right\rangle = \sum_{k=0}^{m} \binom{2n+1}{k} \left\{{n+m+1-k \atop m+1-k}\right\} (-1)^k \qquad \text{при } n > m \geqslant 0,$$

аналогичная формуле (6.38). При $m = 2$ она представляет собой

$$\begin{aligned}\left\langle\!\!\left\langle{n \atop 2}\right\rangle\!\!\right\rangle &= \left\{{n+3 \atop 3}\right\} - (2n+1)\left\{{n+2 \atop 2}\right\} + \binom{2n+1}{2}\left\{{n+1 \atop 1}\right\} \\ &= \tfrac{1}{2}3^{n+2} - (2n+3)2^{n+1} + \tfrac{1}{2}(4n^2+6n+3).\end{aligned}$$

6.69 $\frac{1}{3}n(n+\frac{1}{2})(n+1)(2H_{2n}-H_n) - \frac{1}{36}n(10n^2+9n-1)$. (Было бы неплохо автоматизировать вывод подобных формул.)

6.70 $1/k-1/(k+z) = z/k^2-z^2/k^3+\cdots$; этот ряд сходится при $|z| < 1$.

6.71 Заметьте, что $\prod_{k=1}^{n}(1+z/k)e^{-z/k} = \binom{n+z}{n}n^{-z}e^{(\ln n-H_n)z}$. Если $f(z) = \frac{d}{dz}(z!)$, то $f(z)/z!+\gamma = H_z$.

6.72 Для $\operatorname{tg} z$ можно использовать тождество $\operatorname{tg} z=\operatorname{ctg} z-2\operatorname{ctg} 2z$ (которое эквивалентно тождеству из упр. 23). Кроме того, функция $z/\sin z = z\operatorname{ctg} z+z\operatorname{tg}\frac{1}{2}z$ раскладывается в степенной ряд $\sum_{n\geqslant 0}(-1)^{n-1}(4^n-2)B_{2n}z^{2n}/(2n)!$; а

$$\begin{aligned}\ln\frac{\operatorname{tg} z}{z} &= \ln\frac{\sin z}{z}-\ln\cos z = \\ &= \sum_{n\geqslant 1}(-1)^n\frac{4^nB_{2n}z^{2n}}{(2n)(2n)!} - \sum_{n\geqslant 1}(-1)^n\frac{4^n(4^n-1)B_{2n}z^{2n}}{(2n)(2n)!} = \\ &= \sum_{n\geqslant 1}(-1)^{n-1}\,\frac{4^n(4^n-2)B_{2n}z^{2n}}{(2n)(2n)!},\end{aligned}$$

потому что $\frac{d}{dz}\ln\sin z = \operatorname{ctg} z$ и $\frac{d}{dz}\ln\cos z = -\operatorname{tg} z$.

6.73 $\operatorname{ctg}(z+\pi) = \operatorname{ctg} z$ и $\operatorname{ctg}(z+\frac{1}{2}\pi) = -\operatorname{tg} z$; следовательно, тождество эквивалентно тождеству

$$\operatorname{ctg} z = \frac{1}{2^n}\sum_{k=0}^{2^n-1}\operatorname{ctg}\frac{z+k\pi}{2^n},$$

которое получается по индукции из случая $n = 1$. Указанный предел справедлив, поскольку $z\operatorname{ctg} z \to 1$ при $z \to 0$. Можно показать, что допустим почленный переход к пределу, так что справедлива формула (6.88). (Между прочим, имеет место и общая формула

$$\operatorname{ctg} z = \frac{1}{n}\sum_{k=0}^{n-1}\operatorname{ctg}\frac{z+k\pi}{n}.$$

Она может быть установлена либо из формулы (6.88), либо из формулы

$$\frac{1}{e^{nz}-1} = \frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{e^{z+2k\pi i/n}-1},$$

эквивалентной разложению $1/(z^n-1)$ в элементарные дроби.)

6.74 Поскольку $\operatorname{tg} 2z+\sec 2z = (\sin z+\cos z)/(\cos z-\sin z)$, подстановка $x = 1$ в (6.94) дает $T_n(1) = 2^nE_n$, где $1/\cos z=\sum_{n\geqslant 0}E_{2n}z^{2n}/(2n)!$. (Коэффициенты E_n называются в комбинаторике *Эйлеровыми числами* (не путайте их с числами Эйлера $\left\langle {n \atop k} \right\rangle$). Имеем $\langle E_0, E_1, E_2, \ldots\rangle = \langle 1, 1, 1, 2, 5, 16, 61, 272, 1385, 7936, 50521, \ldots\rangle$.

В числовом анализе Эйлеровы числа определяются иначе: здесь они равны $(-1)^{n/2}E_n[n]$, где E_n — определенные выше Эйлеровы числа.)

6.75 Пусть $G(w,z) = \sin z/\cos(w+z)$ и $H(w,z) = \cos z/\cos(w+z)$ и пусть $G(w,z)+H(w,z) = \sum_{m,n} E_{m,n} w^m z^n/m!\,n!$. Тогда из уравнений $G(w,0) = 0$ и $(\frac{\partial}{\partial z}-\frac{\partial}{\partial w})G(w,z) = H(w,z)$ вытекает, что $E_{m,0} = 0$, когда m нечетное, $E_{m,n+1} = E_{m+1,n}+E_{m,n}$, когда $m+n$ четное; а из уравнений $H(0,z) = 1$ и $(\frac{\partial}{\partial w}-\frac{\partial}{\partial z})H(w,z) = G(w,z)$ вытекает, что $E_{0,n} = [n=0]$ при четном n, $E_{m+1,n} = E_{m,n+1}+E_{m,n}$ при нечетном $m+n$. Следовательно, n-я строка треугольника содержит числа $E_{n,0}$, $E_{n-1,1}$, ..., $E_{0,n}$. Слева число $E_{n,0}$ представляет собой секансное число $E_n[n]$; справа имеем число $E_{0,n} = T_n+[n=0]$.

6.76 Пусть A_n обозначает искомую сумму. Заглянув вперед, в (7.49), мы увидим, что $\sum_n A_n z^n/n! = \sum_{n,k}(-1)^k\left\{{n \atop k}\right\}2^{n-k}k! \times z^n/n! = \sum_k(-1)^k 2^{-k}(e^{2z}-1)^k = 2/(e^{2z}+1) = 1-\operatorname{th} z$. Следовательно, согласно упр. 23 или 72

$$A_n = (2^{n+1}-4^{n+1})B_{n+1}/(n+1) = (-1)^{(n+1)/2}T_n+[n=0].$$

6.77 Формула доказывается индукцией по m с использованием рекуррентного соотношения из упр. 18. Можно также доказать ее, отталкиваясь от (6.46), с учетом того факта, что

$$\frac{(-1)^{m-1}(m-1)!}{(e^z-1)^m} = (D+1)^{\overline{m-1}}\frac{1}{e^z-1} = \\ = \sum_{k=0}^{m-1}\left[{m \atop m-k}\right]\frac{d^{m-k-1}}{dz^{m-k-1}}\frac{1}{e^z-1}, \quad \text{целое } m > 0.$$

Последнее уравнение оказывается эквивалентным следующему:

$$\frac{d^m}{dz^m}\frac{1}{e^z-1} = (-1)^m\sum_k\left\{{m+1 \atop k}\right\}\frac{(k-1)!}{(e^z-1)^k}, \quad \text{целое } m \geqslant 0.$$

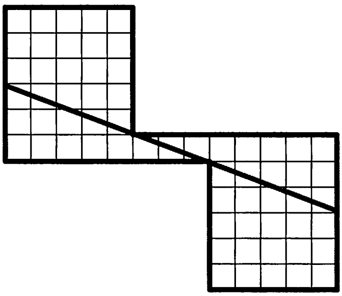
6.78 Если $p(x)$ — некоторый полином степени $\leqslant n$, то

$$p(x) = \sum_k p(-k)\binom{-x}{k}\binom{x+n}{n-k},$$

потому что это равенство справедливо при $x = 0, -1, \ldots, -n$. Указанное в упражнении тождество представляет собой частный случай, когда $p(x) = x\sigma_n(x)$ и $x = 1$. Кстати, мы получаем более простое выражение чисел Бернулли через числа Стирлинга, положив $k = 1$ в (6.99):

$$\sum_{k\geqslant 0} \left\{ {m \atop k} \right\} (-1)^k \frac{k!}{k+1} = B_m .$$

6.79 Сэм Лойд (Sam Loyd) [256, с. 288 и 378] привел конструкцию



и утверждал, что еще в 1858 году он изобрел (но не опубликовал) конструкцию "64 = 65". (Аналогичные парадоксы восходят по крайней мере к XVIII веку, но Лойд нашел более удачные способы их представления.)

6.80 Можно ожидать $A_m/A_{m-1} \approx \phi$, так что давайте испытаем $A_{m-1} = 618034+r$ и $A_{m-2} = 381966-r$. Тогда $A_{m-3} = 236068+2r$ и т.д., и мы найдем, что $A_{m-18} = 144-2584r$, $A_{m-19} = 154+4181r$. Следовательно, $r = 0$, $x = 154$, $y = 144$, $m = 20$.

6.81 Если $P(F_{n+1}, F_n) = 0$ при бесконечном количестве *четных* значений n, то $P(x, y)$ делится на $U(x, y)-1$, где $U(x, y) = x^2-xy-y^2$. Действительно, если t — совокупная степень полинома P, то можно записать

$$P(x, y) = \sum_{k=0}^{t} q_k x^k y^{t-k} + \sum_{j+k<t} r_{j,k} x^j y^k = Q(x, y) + R(x, y) .$$

Тогда

$$\frac{P(F_{n+1}, F_n)}{F_n^t} = \sum_{k=0}^{t} q_k \left(\frac{F_{n+1}}{F_n}\right)^k + O(1/F_n),$$

и, переходя к пределу $n \to \infty$, получаем $\sum_{k=0}^{t} q_k \phi^k = 0$. Следовательно, $Q(x, y)$ кратно $U(x, y)$, скажем, $A(x, y)U(x, y)$. Но $U(F_{n+1}, F_n) = (-1)^n$ и n четное, так что $P_0(x, y) = P(x, y) - (U(x, y)-1)A(x, y)$ представляет собой другой полином, такой, что $P_0(F_{n+1}, F_n) = 0$. Совокупная степень P_0 меньше t, так что по индукции по t полином P_0 кратен $U-1$.

Аналогично полином $P(x,y)$ делится на $U(x,y)+1$, если условие $P(F_{n+1},F_n)=0$ выполняется для бесконечного количества *нечетных* значений n. Объединение этих двух фактов дает требуемое необходимое и достаточное условие: $P(x,y)$ делится на $U(x,y)^2-1$.

6.82 Сначала складываем цифры без переноса и получаем цифры 0, 1 и 2. Затем используем два правила переноса,

$$\begin{aligned} 0\,(d{+}1)\,(e{+}1) &\rightarrow 1\,d\,e\,, \\ 0\,(d{+}2)\,0\,e &\rightarrow 1\,d\,0\,(e{+}1)\,, \end{aligned}$$

всегда выполняя крайний слева допустимый перенос. Этот процесс завершается, поскольку двоичное значение $(b_m \dots b_2)_2$, получаемое при чтении $(b_m \dots b_2)_F$, увеличивается при каждом переносе. Но перенос мог бы распространиться вправо от "точки Фибоначчи", например $(1)_F+(1)_F$ становится $(10.01)_F$. Такое распространение вправо захватывает не более двух позиций, и эти две цифровые позиции при необходимости вновь могут быть обнулены при помощи алгоритма "добавления 1" из текста главы.

Между прочим, имеется соответствующая операция "умножения" неотрицательных целых чисел. Если $m = F_{j_1}+\cdots+F_{j_q}$ и $n = F_{k_1}+\cdots+F_{k_r}$ в фибоначчиевой системе счисления, то по аналогии с умножением двоичных чисел положим $m\circ n = \sum_{b=1}^{q}\sum_{c=1}^{r} F_{j_b+k_c}$. (Это определение означает, что при больших m и n значение $m\circ n \approx \sqrt{5}\,mn$, хотя $1\circ n \approx \phi^2 n$.) Фибоначчиево сложение приводит к доказательству ассоциативного закона $l\circ(m\circ n) = (l\circ m)\circ n$.

Упражнение: $m\circ n = mn + \lfloor (m+1)/\phi \rfloor n + m\lfloor (n+1)/\phi \rfloor$.

6.83 Да; например, можно взять

$$\begin{aligned} A_0 &= 331635635998274737472200656430763\,; \\ A_1 &= 1510028911088401971189590305498785\,. \end{aligned}$$

Получившаяся последовательность обладает тем свойством, что A_n делится на p_k (не будучи ему равно) при $n \bmod m_k = r_k$, где числа (p_k, m_k, r_k) соответственно принимают 18 следующих значений:

$(3,4,1)$	$(2,3,2)$	$(5,5,1)$
$(7,8,3)$	$(17,9,4)$	$(11,10,2)$
$(47,16,7)$	$(19,18,10)$	$(61,15,3)$
$(2207,32,15)$	$(53,27,16)$	$(31,30,24)$
$(1087,64,31)$	$(109,27,7)$	$(41,20,10)$
$(4481,64,63)$	$(5779,54,52)$	$(2521,60,60)$

Каждому целому n соответствует одна из этих троек; например, все нечетные n содержатся в одной из шести троек в первом столбце, а тройки в среднем столбце включают все четные n, которые не делятся на 6. Оставшаяся часть доказательства основана на том факте, что $A_{m+n} = A_m F_{n-1} + A_{m+1} F_n$, а также на сравнениях

$$\begin{aligned} A_0 &\equiv F_{m_k - r_k} \bmod p_k\,, \\ A_1 &\equiv F_{m_k - r_k + 1} \bmod p_k \end{aligned}$$

для каждой из троек (p_k, m_k, r_k). (Возможен “усовершенствованный” вариант решения, в котором A_0 и A_1 представляют собой числа, состоящие “всего лишь” из 17 цифр каждое [218].)

6.84 Последовательности из упр. 62 удовлетворяют следующим соотношениям: $A_{-m} = A_m$, $B_{-m} = -B_m$ и

$$\begin{aligned} A_m A_n &= A_{m+n} + A_{m-n}\,; \\ A_m B_n &= B_{m+n} - B_{m-n}\,; \\ B_m B_n &= A_{m+n} - A_{m-n}\,. \end{aligned}$$

Пусть $f_k = B_{mk}/A_{mk+l}$ и $g_k = A_{mk}/B_{mk+l}$, где $l = \frac{1}{2}(n-m)$. Тогда $f_{k+1} - f_k = A_l B_m / (A_{2mk+n} + A_m)$ и $g_k - g_{k+1} = A_l B_m / (A_{2mk+n} - A_m)$; следовательно,

$$\begin{aligned} S^+_{m,n} &= \frac{\sqrt{5}}{A_l B_m} \lim_{k\to\infty} (f_k - f_0) &&= \frac{\sqrt{5}}{\phi^l A_l L_m}\,; \\ S^-_{m,n} &= \frac{\sqrt{5}}{A_l B_m} \lim_{k\to\infty} (g_0 - g_k) &&= \frac{\sqrt{5}}{A_l L_m}\left(\frac{2}{B_l} - \frac{1}{\phi^l}\right) = \\ & &&= \frac{2}{F_l L_l L_m} - S^+_{m,n}\,. \end{aligned}$$

6.85 Это свойство выполняется тогда и только тогда, когда N имеет один из семи видов: 5^k, $2\cdot 5^k$, $4\cdot 5^k$, $3^j\cdot 5^k$, $6\cdot 5^k$, $7\cdot 5^k$, $14\cdot 5^k$.

6.86 Для любого положительного целого числа m пусть $r(m)$ будет наименьшим индексом j, таким, что C_j делится на m; если такого j не существует, положим $r(m) = \infty$. В таком случае имеет место следующая цепочка эквивалентностей: C_n делится на m тогда и только тогда, когда НОД$(C_n, C_{r(m)})$ делится на m тогда и только тогда, когда $C_{\text{НОД}(n, r(m))}$ делится на m тогда и только тогда, когда НОД$(n, r(m)) = r(m)$ тогда и только тогда, когда n делится на $r(m)$.

(И обратно, как легко видеть, сформулированное НОД-условие вытекает из условия, что последовательность C_1, C_2, …

имеет (возможно, бесконечную) функцию $r(m)$, такую, что C_n делится на m тогда и только тогда, когда n делится на $r(m)$.)

Положим теперь $\Pi(n) = C_1 C_2 \ldots C_n$, так что

$$\binom{m+n}{m}_{\mathcal{C}} = \frac{\Pi(m+n)}{\Pi(m)\,\Pi(n)}\,.$$

Если p — простое число, то кратность, с которой p делит $\Pi(n)$, равна $f_p(n) = \sum_{k\geqslant 1} \lfloor n/r(p^k)\rfloor$, поскольку $\lfloor n/p^k\rfloor$ есть количество элементов $\{C_1, \ldots, C_n\}$, делящихся на p^k. Поэтому $f_p(m+n) \geqslant f_p(m)+f_p(n)$ для всех p и $\binom{m+n}{m}_{\mathcal{C}}$ — целое число.

6.87 Данное произведение матриц равно

$$\begin{pmatrix} K_{n-2}(x_2, \ldots, x_{n-1}) & K_{n-1}(x_2, \ldots, x_{n-1}, x_n) \\ K_{n-1}(x_1, x_2, \ldots, x_{n-1}) & K_n(x_1, x_2, \ldots, x_{n-1}, x_n) \end{pmatrix}.$$

Это же относится и к произведениям L и R, таким как в (6.137), потому что

$$R^a \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} L^a.$$

Определитель же равен $K_n(x_1, \ldots, x_n)$; более общий трехдиагональный определитель

$$\det \begin{pmatrix} x_1 & 1 & 0 & \ldots & 0 \\ y_2 & x_2 & 1 & & 0 \\ 0 & y_3 & x_3 & 1 & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & 0 & \ldots & y_n & x_n \end{pmatrix}$$

удовлетворяет рекуррентному соотношению $D_n = x_n D_{n-1} - y_n \times D_{n-2}$.

6.88 Пусть $\alpha^{-1} = a_0 + 1/\bigl(a_1 + 1/(a_2 + \cdots)\bigr)$ — представление α^{-1} в виде непрерывной дроби. Тогда

$$\frac{a_0}{z} + \cfrac{1}{A_0(z) + \cfrac{1}{A_1(z) + \cfrac{1}{A_2(z) + \cfrac{1}{\ddots}}}} = \frac{1-z}{z} \sum_{n\geqslant 1} z^{\lfloor n\alpha\rfloor},$$

где

$$A_m(z) = \frac{z^{-q_{m+1}} - z^{-q_{m-1}}}{z^{-q_m} - 1}, \qquad q_m = K_m(a_1, \ldots, a_m).$$

В доказательстве, аналогичном доказательству (6.146) в тексте главы, используется обобщение теоремы Цеккендорфа (Френкель (Fraenkel) [129, §4]). Если $z = 1/b$, где целое $b \geqslant 2$, то это дает представление в виде непрерывной дроби трансцендентного числа $(b-1)\sum_{n\geqslant 1} b^{-\lfloor n\alpha \rfloor}$, как в упр. 49.

6.89 Пусть $p = K(0, a_1, a_2, \ldots, a_m)$, так что p/n представляет собой m-ю подходящую дробь к данной непрерывной дроби. Тогда $\alpha = p/n + (-1)^m/nq$, где $q = K(a_1, \ldots, a_m, \beta)$ и $\beta > 1$. Поэтому точки $\{k\alpha\}$ при $0 \leqslant k < n$ могут быть записаны в виде

$$\frac{0}{n},\ \frac{1}{n}+\frac{(-1)^m\pi_1}{nq},\ \ldots,\ \frac{n-1}{n}+\frac{(-1)^m\pi_{n-1}}{nq},$$

где $\pi_1 \ldots \pi_{n-1}$ — перестановка множества $\{1, \ldots, n-1\}$. Пусть $f(\nu)$ — количество таких точек $< \nu$; тогда как $f(\nu)$, так и νn увеличиваются на 1 при возрастании ν от k/n до $(k+1)/n$, за исключением случая, когда $k = 0$ или $k = n-1$, так что они никогда не различаются более чем на 2.

6.90 Согласно (6.139) и (6.136) нам нужно максимизировать $K(a_1, \ldots, a_m)$ по всем последовательностям положительных целых чисел, сумма которых $\leqslant n+1$. Данный максимум достигается тогда, когда все a равны 1, поскольку, если $j \geqslant 1$ и $a \geqslant 1$, то

$$\begin{aligned}
&K_{j+k+1}(1, \ldots, 1, a+1, b_1, \ldots, b_k) = \\
&\quad = K_{j+k+1}(1, \ldots, 1, a, b_1, \ldots, b_k) + K_j(1, \ldots, 1)\, K_k(b_1, \ldots, b_k) \leqslant \\
&\quad \leqslant K_{j+k+1}(1, \ldots, 1, a, b_1, \ldots, b_k) + K_{j+k}(1, \ldots, 1, a, b_1, \ldots, b_k) = \\
&\quad = K_{j+k+2}(1, \ldots, 1, a, b_1, \ldots, b_k).
\end{aligned}$$

(Моцкин (Motzkin) и Штраус (Straus) [278] показывают, как решать более общие задачи максимизации, связанные с континуантами.)

6.91 Один из кандидатов для случая $n \bmod 1 = \frac{1}{2}$ представлен в [213, §6], хотя, быть может, лучше всего умножить обсуждаемые там целые числа на некоторую константу, включающую в себя $\sqrt{\pi}$. Элегантное решение этой исследовательской проблемы было найдено Филиппом Флажоле (Philippe Flajolet) и Гельмутом Продингером (Helmut Prodinger) в *SIAM Journal on Discrete Mathematics* **12** (1999), 155–159.

6.92 (а) Дэвид Бойд (David Boyd) показал, что существует только конечное количество решений при любом $p < 500$, за исключением, возможно, $p = 83, 127, 397$. (б) Поведение b_n достаточно странное: с одной стороны, $b_n = \text{НОК}(1, \ldots, n)$ при $968 \leqslant n \leqslant 1066$; с другой стороны, $b_{600} = \text{НОК}(1, \ldots, 600)/(3^3\cdot$

$5^2 43$). Эндрю Одлыжко (Andrew Odlyzko) заметил, что p делит НОК$(1,\ldots,n)/b_n$ тогда и только тогда, когда $kp^m \leqslant n < (k+1)p^m$ при некотором $m \geqslant 1$ и некотором $k < p$, таком, что p делит числитель числа H_k. Поэтому бесконечно много таких n существует в том случае, когда, к примеру, может быть показано, что почти все простые числа имеют только одно такое значение k (а именно — $k = p-1$).

6.93 (Брент (Brent) [38] обнаружил у числа e^γ удивительно большое неполное частное 1568705, но, по-видимому, это просто случайное стечение обстоятельств. Например, Госпер (Gosper) обнаружил еще бо́льшие неполные частные у числа π: 453 294-е равно 12996958, а 11 504 931-е равно 878783625.)

6.94 Рассмотрите производящую функцию $\sum_{m,n\geqslant 0} \left|{m+n \atop m}\right| w^m z^n$, которая равна $\sum_k \bigl(wF(\alpha'+\beta'+\gamma', \alpha'+\beta', \alpha')+zF(\alpha+\gamma, \alpha+\beta, \alpha)\bigr)^k(1)$, где $F(a,b,c)$ — дифференциальный оператор $a+b\vartheta_w+c\vartheta_z$.

Кауэрс нашел решение несмотря на то, что числа Стирлинга не “голономны” в смысле [383].

6.95 Элегантное решение этой исследовательской проблемы найдено Мануэлем Кауэрсом (Manual Kauers), *Journal of Symbolic Computation* **42** (2007), 948–970.

7.1 Подставим в производящую функцию z^4 вместо □ и z вместо ▭, что даст нам $1/(1-z^4-z^2)$. Это похоже на производящую функцию для T, в которой z заменено на z^2. Следовательно, ответом будет 0 для нечетных m и $F_{m/2+1}$ — для четных.

7.2 $G(z) = 1/(1-2z)+1/(1-3z)$; $\widehat{G}(z) = e^{2z}+e^{3z}$.

7.3 Положив в производящей функции $z = 1/10$, получим $\frac{10}{9}\ln\frac{10}{9}$.

7.4 Поделим $P(z)$ на $Q(z)$ и найдем частное $T(z)$ и остаток $P_0(z)$, степень которого меньше степени Q. Коэффициенты $T(z)$ следует добавить к коэффициентам $[z^n]\,P_0(z)/Q(z)$ для малых значений n. (Это — полином $T(z)$ из (7.28).)

7.5 Это свертка $(1+z^2)^r$ и $(1+z)^r$, так что

$$S(z) = (1+z+z^2+z^3)^r .$$

Кстати, для коэффициентов этой производящей функции не известно никакой простой формулы; следовательно, указанная сумма, скорее всего, не может быть выражена в достаточно простом аналитическом виде. (Производящие функции пригодны и для получения отрицательных результатов.)

7.6 Запишем решение соотношений $g_0 = \alpha$, $g_1 = \beta$, $g_n = g_{n-1}+2g_{n-2}+(-1)^n\gamma$ в виде $g_n = A(n)\alpha+B(n)\beta+C(n)\gamma$. Функция 2^n подходит в качестве решения при $\alpha = 1$, $\beta = 2$, $\gamma = 0$;

функция $(-1)^n$ подходит в качестве решения при $\alpha=1$, $\beta=-1$, $\gamma=0$; функция $(-1)^n n$ подходит в качестве решения при $\alpha=0$, $\beta=-1$, $\gamma=3$. Следовательно, $A(n){+}2B(n)=2^n$, $A(n){-}B(n)=(-1)^n$ и $-B(n){+}3C(n)=(-1)^n n$.

7.7 $G(z)=\bigl(z/(1{-}z)^2\bigr)G(z){+}1$, следовательно,

$$G(z)=\frac{1-2z+z^2}{1-3z+z^2}=1+\frac{z}{1-3z+z^2};$$

поэтому имеем $g_n=F_{2n}+[n=0]$.

Готов спорить, что у сомнительного "фана порядка 0" все же есть одно остовное дерево.

7.8 Продифференцировав $(1{-}z)^{-x-1}$ по x дважды, получим

$$\binom{x+n}{n}\bigl((H_{x+n}-H_x)^2-(H^{(2)}_{x+n}-H^{(2)}_x)\bigr).$$

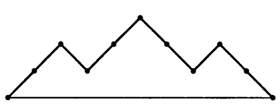
Теперь положим $x=m$.

7.9 $(n{+}1)(H_n^2-H_n^{(2)})-2n(H_n-1)$.

7.10 Из тождества $H_{k-1/2}-H_{-1/2}=\frac{2}{2k-1}+\cdots+\frac{2}{1}=2H_{2k}-H_k$ вытекает, что $\sum_k\binom{2k}{k}\binom{2n-2k}{n-k}(2H_{2k}-H_k)=4^nH_n$.

7.11 (а) $C(z)=A(z)B(z^2)/(1{-}z)$. (б) $zB'(z)=A(2z)e^z$, следовательно, $A(z)=\frac{z}{2}e^{-z/2}B'(\frac{z}{2})$. (в) $A(z)=B(z)/(1{-}z)^{r+1}$, следовательно, $B(z)=(1{-}z)^{r+1}A(z)$ и поэтому $f_k(r)=\binom{r+1}{k}(-1)^k$.

7.12 C_n. Числа в верхней строке соответствуют позициям '$+1$' в последовательности из $+1$ и -1, определяющей "горную гряду"; числа в нижней строке соответствуют позициям '-1'. Например, указанный в условии массив соответствует



.

7.13 Продолжим последовательность периодическим образом (положив $x_{m+k}=x_k$) и определим $s_n=x_1+\cdots+x_n$. Имеем $s_m=l$, $s_{2m}=2l$ и т.д. Должен существовать максимальный индекс k_j, такой, что $s_{k_j}=j$, $s_{k_j+m}=l+j$ и т.д. Эти индексы k_1, …, k_l (modulo m) определяют требуемый циклический сдвиг.

Например, в последовательности $\langle -2,1,-1,0,1,1,-1,1,1,1\rangle$ с $m=10$ и $l=2$ имеем $k_1=17$, $k_2=24$.

7.14 Уравнение $\widehat{G}(z)=-2z\widehat{G}(z)+\widehat{G}(z)^2+z$ (отнеситесь к последнему слагаемому с должным вниманием!) с помощью формулы для корней квадратного уравнения приводит к

$$\widehat{G}(z)=\frac{1+2z-\sqrt{1+4z^2}}{2}.$$

Следовательно, $g_{2n+1} = 0$ и $g_{2n} = (-1)^n(2n)!\,C_{n-1}$ для всех $n > 0$.

7.15 Существует $\binom{n}{k}\varpi_{n-k}$ разбиений, в которых подмножество, включающее $n+1$, содержит еще k других объектов. Следовательно, $\widehat{P}'(z) = e^z\widehat{P}(z)$. Решением этого дифференциального уравнения будет $\widehat{P}(z) = e^{e^z+c}$, причем $c = -1$, поскольку $\widehat{P}(0) = 1$. (Этот результат можно также получить суммированием (7.49) по m, поскольку $\varpi_n = \sum_m {n \brace m}$.)

7.16 Один из способов состоит в том, чтобы прологарифмировать соотношение

$$B(z) = 1\big/\big((1-z)^{a_1}(1-z^2)^{a_2}(1-z^3)^{a_3}(1-z^4)^{a_4}\ldots\big),$$

а затем использовать формулу для $\ln\frac{1}{1-z}$ и поменять порядок суммирования.

7.17 Это следует из того, что $\int_0^\infty t^n e^{-t}\,dt = n!$. Имеется также обратная формула:

$$\widehat{G}(z) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} G(ze^{-i\theta})\,e^{e^{i\theta}}\,d\theta\,.$$

7.18 (а) $\zeta(z-\frac{1}{2})$; (б) $-\zeta'(z)$; (в) $\zeta(z)/\zeta(2z)$. Любое положительное целое число однозначно представимо в виде m^2q, где q свободно от квадратов.

7.19 При $n > 0$ коэффициент $[z^n]\exp\big(x\ln F(z)\big)$ является полиномом степени n от x, который кратен x. Первая формула свертки получается из приравнивания коэффициентов при z^n в равенстве $F(z)^xF(z)^y = F(z)^{x+y}$. Вторая формула получается из приравнивания коэффициентов при z^{n-1} в $F'(z)F(z)^{x-1}F(z)^y = F'(z)F(z)^{x+y-1}$, потому что

$$F'(z)F(z)^{x-1} = x^{-1}\frac{\partial}{\partial z}\big(F(z)^x\big) = x^{-1}\sum_{n\geqslant 0} nf_n(x)z^{n-1}\,.$$

(Последующие свертки получаются путем взятия частной производной $\partial/\partial x$, как в (7.43).)

Как показано в [221], можно утверждать даже большее, а именно

$$\sum_{k=0}^{n}\frac{xf_k(x+tk)}{x+tk}\,\frac{yf_{n-k}(y+t(n-k))}{y+t(n-k)} = \frac{(x+y)f_n(x+y+tn)}{x+y+tn}$$

для произвольных x, y и t. В действительности последовательность полиномов $xf_n(x+tn)/(x+tn)$ дает коэффициенты функции $\mathcal{F}_t(z)^x$, где

$$\mathcal{F}_t(z) = F\big(z\mathcal{F}_t(z)^t\big)\,.$$

(В (5.59) и (6.48) мы встречались с частными случаями этого тождества.)

7.20 Пусть $G(z) = \sum_{n\geqslant 0} g_n z^n$. Тогда

$$z^l G^{(k)}(z) = \sum_{n\geqslant 0} n^{\underline{k}} g_n z^{n-k+l} = \sum_{n\geqslant 0} (n+k-l)^{\underline{k}} g_{n+k-l} z^n$$

для всех $k, l \geqslant 0$, если положить $g_n = 0$ при $n < 0$. Следовательно, если $P_0(z), \ldots, P_m(z)$ — полиномы с максимальной степенью d, не все равные нулю, то существуют полиномы $p_0(n), \ldots, p_{m+d}(n)$, такие, что

$$P_0(z)G(z)+\cdots+P_m(z)G^{(m)}(z) = \sum_{n\geqslant 0} \sum_{j=0}^{m+d} p_j(n) g_{n+j-d} z^n .$$

Поэтому из условия дифференцируемой конечности функции $G(z)$ следует, что

$$\sum_{j=0}^{m+d} p_j(n+d) g_{n+j} = 0 \qquad \text{при всех } n \geqslant 0.$$

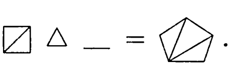
Обратное утверждение доказывается аналогично. (Одно из следствий заключается в том, что производящая функция $G(z)$ дифференцируемо конечна тогда и только тогда, когда соответствующая экспоненциальная производящая функция $\widehat{G}(z)$ дифференцируемо конечная.)

7.21 Это задача размена с номиналами монет 10 и 20, так что $G(z) = 1/(1-z^{10})(1-z^{20}) = \check{G}(z^{10})$, где $\check{G}(z) = 1/(1-z)(1-z^2)$. (а) Разложением $\check{G}(z)$ на элементарные дроби является $\frac{1}{2}(1-z)^{-2}+\frac{1}{4}(1-z)^{-1}+\frac{1}{4}(1+z)^{-1}$, так что $[z^n]\,\check{G}(z) = \frac{1}{4}\bigl(2n+3+(-1)^n\bigr)$. Подстановка $n = 50$ дает 26 способов выплаты требуемой суммы. (б) $\check{G}(z) = (1+z)/(1-z^2)^2 = (1+z)(1+2z^2+3z^4+\cdots)$, так что $[z^n]\,\check{G}(z) = \lfloor n/2 \rfloor + 1$. (Сравните со значением $N_n = \lfloor n/5 \rfloor + 1$ в задаче о размене монет в тексте. Задача грабителя эквивалентна задаче размена при помощи копеек и двушек.)

Этот медленный способ решения поможет кассиру потянуть время до приезда полиции.

В США формально имеется монета достоинством два цента, но с 1873 года она не чеканится.

7.22 Каждый многоугольник имеет "основание" (отрезок внизу). Если A и B — триангулированные многоугольники, положим, что $A \triangle B$ является результатом приклеивания основания A к верхней левой стороне $\triangle$, а основания B — к верхней правой стороне. Так, например,



(Исходные многоугольники, возможно, придется слегка искривить для подгонки к нужной форме.) Таким способом можно получить любую триангуляцию, поскольку основание входит ровно в один треугольник, при этом слева и справа от него лежат некоторые триангулированные треугольники A и B.

Замена каждого треугольника на z дает степенной ряд, в котором коэффициентом при z^n служит число триангуляций с n треугольниками, а именно количество способов разбиений $(n+2)$-угольника на треугольники. Поскольку $P = 1+zP^2$, это производящая функция для чисел Каталана $C_0+C_1z+C_2z^2+\cdots$; количество триангуляций n-угольника равно $C_{n-2} = \binom{2n-4}{n-2}/(n-1)$.

7.23 Обозначим через a_n искомое число, а через b_n — количество разбиений колонны с выемкой размером $2\times1\times1$ на вершине. Рассматривая возможные варианты расположения видимых сверху кирпичей, получим

$$\begin{aligned} a_n &= 2a_{n-1}+4b_{n-1}+a_{n-2}+[n=0]\,; \\ b_n &= a_{n-1}+b_{n-1}\,. \end{aligned}$$

Следовательно, производящие функции удовлетворяют уравнениям $A = 2zA+4zB+z^2A+1$, $B = zA+zB$, откуда находим

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

Эта формула имеет отношение к задаче о покрытии костями домино прямоугольника размером $3\times n$; имеем $a_n = \frac{1}{3}\big(U_{2n}+V_{2n+1}+(-1)^n\big) = \frac{1}{6}(2+\sqrt{3}\,)^{n+1}+\frac{1}{6}(2-\sqrt{3}\,)^{n+1}+\frac{1}{3}(-1)^n$, что есть просто $(2+\sqrt{3}\,)^{n+1}/6$, округленное до ближайшего целого.

"Интересно, что a_{2n} равно U_{2n}^2, квадрату количества способов покрытия прямоугольника размером $3\times 2n$ костями домино, а $a_{2n+1} = 2V_{2n+1}^2$."
— И. Капланси (I. Kaplansky)

7.24 $n\sum_{k_1+\cdots+k_m=n} k_1\cdot\ldots\cdot k_m/m = F_{2n+1}+F_{2n-1}-2$. (Рассмотрите коэффициент $[z^{n-1}]\,\frac{d}{dz}\ln\big(1/(1-G(z))\big)$, где $G(z) = z/(1-z)^2$.)

7.25 Производящая функция — $P(z)/(1-z^m)$, где $P(z)=z+2z^2+\cdots+(m-1)z^{m-1} = \big((m-1)z^{m+1}-mz^m+z\big)/(1-z)^2$. В знаменателе находится полином $Q(z) = 1-z^m = (1-\omega^0z)(1-\omega^1z)\ldots(1-\omega^{m-1}z)$. По теореме о разложении рациональных функций для случая различных корней получаем

$$n \bmod m = \frac{m-1}{2}+\sum_{k=1}^{m-1}\frac{\omega^{-kn}}{\omega^k-1}\,.$$

7.26 $(1-z-z^2)\mathfrak{F}(z)=F(z)$ приводит к $\mathfrak{F}_n = \big(2(n+1)F_n+nF_{n+1}\big)/5$, как в уравнении (7.61).

7.27 Каждая ориентированная циклическая схема начинается с  или , или $2\times k$-цикла (для некоторого $k \geqslant 2$), ориентированного одним из двух способов. Следовательно,

$$Q_n = Q_{n-1}+Q_{n-2}+2Q_{n-2}+2Q_{n-3}+\cdots+2Q_0$$

для $n \geqslant 2$; $Q_0 = Q_1 = 1$. Таким образом, производящая функция представляет собой

$$\begin{aligned} Q(z) &= zQ(z)+z^2Q(z)+2z^2Q(z)/(1-z)+1 = \\ &= 1/\bigl(1-z-z^2-2z^2/(1-z)\bigr) = \\ &= \frac{(1-z)}{(1-2z-2z^2+z^3)} = \\ &= \frac{\phi^2/5}{1-\phi^2 z}+\frac{\phi^{-2}/5}{1-\phi^{-2}z}+\frac{2/5}{1+z}, \end{aligned}$$

и $Q_n = \bigl(\phi^{2n+2}+\phi^{-2n-2}+2(-1)^n\bigr)/5 = \bigl((\phi^{n+1}-\hat\phi^{n+1})/\sqrt5\,\bigr)^2 = F_{n+1}^2$.

7.28 В общем случае, если $A(z) = (1+z+\cdots+z^{m-1})B(z)$, то мы имеем $A_r+A_{r+m}+A_{r+2m}+\cdots = B(1)$ для $0 \leqslant r < m$. В нашем случае $m = 10$ и $B(z) = (1+z+\cdots+z^9)(1+z^2+z^4+z^6+z^8)(1+z^5)$.

7.29 $F(z)+F(z)^2+F(z)^3+\cdots = z/(1-z-z^2-z) = \bigl(1/(1-(1+\sqrt2\,)z)-1/(1-(1-\sqrt2\,)z)\bigr)/\sqrt8$, так что искомый ответ — $\bigl((1+\sqrt2\,)^n-(1-\sqrt2\,)^n\bigr)/\sqrt8$.

7.30 $\sum_{k=1}^{n}\binom{2n-1-k}{n-1}\bigl(a^nb^{n-k}/(1-\alpha z)^k+a^{n-k}b^n/(1-\beta z)^k\bigr)$, согласно упр. 5.39.

7.31 Производящая функция Дирихле будет равна $\zeta(z)^2/\zeta(z-1)$; следовательно, находим, что $g(n)$ представляет собой произведение $(k+1-kp)$ по всем степеням простых чисел p^k, делящим n нацело.

7.32 Можно считать, что все $b_k \geqslant 0$. Множество арифметических прогрессий будет точным покрытием тогда и только тогда, когда

$$\frac{1}{1-z} = \frac{z^{b_1}}{1-z^{a_1}}+\cdots+\frac{z^{b_m}}{1-z^{a_m}}.$$

Вычтем $z^{b_m}/(1-z^{a_m})$ из обеих частей и положим $z = e^{2\pi i/a_m}$. Левая часть обратится в бесконечность, а правая останется конечной, если только a_{m-1} не будет равно a_m.

7.33 $(-1)^{n-m+1}[n>m]/(n-m)$.

7.34 Можно также записать

$$G_n(z) = \sum_{k_1+(m+1)k_{m+1}=n} \binom{k_1+k_{m+1}}{k_{m+1}} (z^m)^{k_{m+1}}.$$

В общем случае, если

$$G_n = \sum_{k_1+2k_2+\cdots+rk_r=n} \binom{k_1+k_2+\cdots+k_r}{k_1, k_2, \ldots, k_r} z_1^{k_1} z_2^{k_2} \ldots z_r^{k_r},$$

то $G_n = z_1G_{n-1}+z_2G_{n-2}+\cdots+z_rG_{n-r}+[n=0]$, и производящей функцией будет $1/(1-z_1w-z_2w^2-\cdots-z_rw^r)$. В рассматриваемом частном случае ответом является $1/(1-w-z^mw^{m+1})$. (См. частный случай $m=1$ в (5.74).)

7.35 (а) $\frac{1}{n}\sum_{0<k<n}\bigl(1/k+1/(n-k)\bigr) = \frac{2}{n}H_{n-1}$. (б) $[z^n]\left(\ln\frac{1}{1-z}\right)^2 = \frac{2!}{n!}\left[{n \atop 2}\right] = \frac{2}{n}H_{n-1}$ из (7.50) и (6.58). Еще один способ в п. (б) состоит в использовании правила $[z^n]\,F(z) = \frac{1}{n}[z^{n-1}]\,F'(z)$ с $F(z) = \left(\ln\frac{1}{1-z}\right)^2$.

7.36 $\frac{1-z^m}{1-z}A(z^m)$.

7.37 (а) В таблице выполняется соотношение $a_{2n}=a_{2n+1}=b_n$:

n	0	1	2	3	4	5	6	7	8	9	10
a_n	1	1	2	2	4	4	6	6	10	10	14
b_n	1	2	4	6	10	14	20	26	36	46	60

(б) $A(z) = 1/\bigl((1-z)(1-z^2)(1-z^4)(1-z^8)\ldots\bigr)$. (в) $B(z) = A(z)/(1-z)$, а мы хотим доказать, что $A(z) = (1+z)B(z^2)$. Это следует из $A(z) = A(z^2)/(1-z)$.

7.38 $(1-wz)M(w,z) = \sum_{m,n\geqslant 1}\bigl(\min(m,n)-\min(m-1,n-1)\bigr)\times w^mz^n = \sum_{m,n\geqslant 1} w^mz^n = wz/(1-w)(1-z)$. В общем случае

$$M(z_1,\ldots,z_m) = \frac{z_1\ldots z_m}{(1-z_1)\ldots(1-z_m)(1-z_1\ldots z_m)}.$$

7.39 Ответами на вопросы из указания будут соответственно

$$\sum_{1\leqslant k_1<k_2<\cdots<k_m\leqslant n} a_{k_1}a_{k_2}\ldots a_{k_m} \quad \text{и} \quad \sum_{1\leqslant k_1\leqslant k_2\leqslant\cdots\leqslant k_m\leqslant n} a_{k_1}a_{k_2}\ldots a_{k_m}.$$

Следовательно: (а) нам нужен коэффициент при z^m в произведении $(1+z)(1+2z)\ldots(1+nz)$. Этот полином является зеркальным к $(z+1)^{\overline{n}}$, поэтому он равен $\left[{n+1 \atop n+1}\right]+\left[{n+1 \atop n}\right]z+\cdots+\left[{n+1 \atop 1}\right]z^n$ и ответом будет $\left[{n+1 \atop n+1-m}\right]$. (б) Коэффициент при z^m в $1/\bigl((1-z)(1-2z)\ldots(1-nz)\bigr)$ согласно (7.47) равен $\left\{{m+n \atop n}\right\}$.

7.40 Экспоненциальная производящая функция для $\langle nF_{n-1}-F_n\rangle$ будет $(z-1)\widehat{F}(z)$, где $\widehat{F}(z)=\sum_{n\geqslant 0}F_n z^n/n! = (e^{\phi z}-e^{\hat\phi z})/\sqrt{5}$. Экспоненциальная производящая функция для $\langle n\text{¡}\rangle$ равна $e^{-z}/(1-z)$. Их произведение равно

$$5^{-1/2}\bigl(e^{(\hat\phi-1)z}-e^{(\phi-1)z}\bigr) = 5^{-1/2}(e^{-\phi z}-e^{-\hat\phi z}).$$

Имеем $\widehat{F}(z)e^{-z}=-\widehat{F}(-z)$. Таким образом, ответом будет $(-1)^n F_n$.

7.41 Количество возрастающе-убывающих перестановок с максимальным элементом n в позиции $2k$ равно $\binom{n-1}{2k-1}A_{2k-1}A_{n-2k}$. Аналогично количество возрастающе-убывающих перестановок с наименьшим элементом 1 в позиции $2k+1$ равно $\binom{n-1}{2k}A_{2k}A_{n-2k-1}$, потому что убывающе-возрастающих перестановок столько же, сколько и возрастающе-убывающих. Суммирование по всем возможностям дает

$$2A_n = \sum_k \binom{n-1}{k} A_k\, A_{n-1-k} + 2[n=0] + [n=1].$$

Следовательно, экспоненциальная производящая функция $\widehat{A}$ удовлетворяет уравнению $2\widehat{A}'(z)=\widehat{A}(z)^2+1$, и $\widehat{A}(0)=1$; указанная функция является решением этого дифференциального уравнения. (Следовательно, A_n представляет собой число Эйлера E_n из упр. 6.74, а именно — число секанса при четном n и число тангенса при нечетном n.)

7.42 Пусть a_n — количество цепочек марсианских ДНК, которые не заканчиваются на c или e; и пусть b_n — количество цепочек, заканчивающихся этими символами. Тогда

$$\begin{aligned}
a_n &= 3a_{n-1}+2b_{n-1}+[n=0], & b_n &= 2a_{n-1}+b_{n-1};\\
A(z) &= 3zA(z)+2zB(z)+1, & B(z) &= 2zA(z)+zB(z);\\
A(z) &= \frac{1-z}{1-4z-z^2}, & B(z) &= \frac{2z}{1-4z-z^2};
\end{aligned}$$

и общее количество цепочек равно $[z^n](1+z)/(1-4z-z^2)=F_{3n+2}$.

7.43 Из (5.45) имеем $g_n=\Delta^n\dot{G}(0)$. Можно записать n-ю разность произведения как

$$\Delta^n A(z)B(z) = \sum_k \binom{n}{k}\bigl(\Delta^k E^{n-k}A(z)\bigr)\bigl(\Delta^{n-k}B(z)\bigr),$$

и $E^{n-k}=(1+\Delta)^{n-k}=\sum_j\binom{n-k}{j}\Delta^j$. Таким образом, находим

$$h_n = \sum_{j,k}\binom{n}{k}\binom{n-k}{j} f_{j+k}\, g_{n-k}.$$

Это сумма по всем триномиальным коэффициентам; ее можно привести к более симметричному виду

$$h_n = \sum_{j+k+l=n} \binom{n}{j,k,l} f_{j+k}\, g_{k+l}\,.$$

7.44 Каждое разбиение на k непустых подмножеств может быть упорядочено $k!$ способами, так что $b_k = k!$. Таким образом, можно записать $\widehat{Q}(z) = \sum_{n,k\geqslant 0} {n \brace k} k!\, z^n/n! = \sum_{k\geqslant 0}(e^z-1)^k = 1/(2-e^z)$. Это — сумма геометрической прогрессии $\sum_{k\geqslant 0} e^{kz}/2^{k+1}$, следовательно, $a_k = 1/2^{k+1}$. Наконец, $c_k = 2^k$. Рассмотрите все перестановки для случая различных x, замените каждое отношение '>' между индексами на '<', а если индексы связаны отношением '<', то для значений допускается любое из отношений '<' или '='. (Например, перестановка $x_1x_3x_2$ порождает два варианта сравнения: $x_1 < x_3 < x_2$ и $x_1 = x_3 < x_2$, потому что $1 < 3 > 2$.)

7.45 Эта сумма равна $\sum_{n\geqslant 1} r(n)/n^2$, где $r(n)$ — количество способов записать n в виде произведения двух взаимно простых сомножителей. Если n делится на t различных простых чисел, то $r(n) = 2^t$. Следовательно, $r(n)/n^2$ — мультипликативная функция и интересующая нас сумма равна

$$\prod_p \left(1+\frac{2}{p^2}+\frac{2}{p^4}\cdots\right) = \prod_p \left(1+\frac{2}{p^2-1}\right) =$$
$$= \prod_p \left(\frac{p^2+1}{p^2-1}\right) = \zeta(2)^2/\zeta(4) = \frac{5}{2}\,.$$

7.46 Пусть $S_n = \sum_{0\leqslant k\leqslant n/2}\binom{n-2k}{k}\alpha^k$. Тогда $S_n = S_{n-1}+\alpha S_{n-3}+[n=0]$, и производящая функция равна $1/(1-z-\alpha z^3)$. При $\alpha = -\frac{4}{27}$, как рекомендуется в указании, эта функция имеет хорошее разложение на множители $1/(1+\frac{1}{3}z)(1-\frac{2}{3}z)^2$. Общая теорема о разложении дает теперь $S_n = (\frac{2}{3}n+c)(\frac{2}{3})^n+\frac{1}{9}(-\frac{1}{3})^n$, и константа c оказывается равной $\frac{8}{9}$.

7.47 Представление Штерна–Броко для $\sqrt{3}$ есть $R(LR^2)^\infty$, потому что

$$\sqrt{3}+1 = 2+\cfrac{1}{1+\cfrac{1}{\sqrt{3}+1}}\,.$$

Сами дроби равны $\frac{1}{1}, \frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \frac{7}{4}, \frac{12}{7}, \frac{19}{11}, \frac{26}{15}, \ldots$; начиная с некоторого момента они задаются поочередно следующими выражениями:

$$\frac{V_{2n-1}+V_{2n+1}}{U_{2n}},\ \frac{U_{2n}+V_{2n+1}}{V_{2n+1}},\ \frac{U_{2n+2}+V_{2n-1}}{U_{2n}+V_{2n+1}},\ \frac{V_{2n+1}+V_{2n+3}}{U_{2n+2}},\ \ldots .$$

7.48 Имеем $g_0 = 0$, и, если $g_1 = m$, производящая функция удовлетворяет уравнению

$$aG(z)+bz^{-1}G(z)+cz^{-2}\bigl(G(z)-mz\bigr)+\frac{d}{1-z} = 0.$$

Следовательно, $G(z) = P(z)/(az^2+bz+c)(1-z)$ для некоторого полинома $P(z)$. Пусть ρ_1 и ρ_2 — корни cz^2+bz+a, причем $|\rho_1| \geqslant |\rho_2|$. Если $b^2-4ac \leqslant 0$, то $|\rho_1|^2 = \rho_1\rho_2 = a/c$ — рациональное число, что противоречит тому факту, что $\sqrt[n]{g_n}$ стремится к $1+\sqrt{2}$. Следовательно, $\rho_1 = (-b+\sqrt{b^2-4ca})/2c = 1+\sqrt{2}$; а отсюда следует, что $a = -c$, $b = -2c$, $\rho_2 = 1-\sqrt{2}$. Производящая функция теперь принимает вид

$$\begin{aligned} G(z) &= \frac{z\bigl(m-(r+m)z\bigr)}{(1-2z-z^2)(1-z)} = \\ &= \frac{-r+(2m+r)z}{2(1-2z-z^2)}+\frac{r}{2(1-z)} = mz+(2m-r)z^2+\cdots, \end{aligned}$$

где $r = d/c$. Поскольку g_2 — целое, r также целое. Имеем также

$$g_n = \alpha(1+\sqrt{2}\,)^n+\hat{\alpha}(1-\sqrt{2}\,)^n+\tfrac{1}{2}r = \bigl\lfloor \alpha(1+\sqrt{2}\,)^n \bigr\rfloor,$$

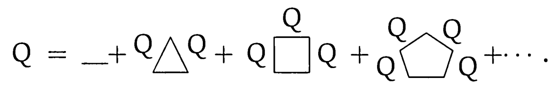
а это может быть, только если $r = -1$, потому что $(1-\sqrt{2}\,)^n$ стремится к нулю, чередуясь в знаке. Следовательно, $(a,b,c,d) = \pm(1,2,-1,1)$. Теперь находим $\alpha = \frac{1}{4}(1+\sqrt{2}\,m)$, которое находится между 0 и 1, только когда $0 \leqslant m \leqslant 2$. Каждое из этих значений в действительности дает некоторое решение; соответствующие последовательности $\langle g_n\rangle$ представляют собой $\langle 0,0,1,3,8,\ldots\rangle$, $\langle 0,1,3,8,20,\ldots\rangle$ и $\langle 0,2,5,13,32,\ldots\rangle$.

7.49 (а) Общим знаменателем выражения $\bigl(1/\bigl(1-(1+\sqrt{2})z\bigr)+1/\bigl(1-(1-\sqrt{2})z\bigr)\bigr)$ является $1-2z-z^2$; следовательно, $a_n = 2a_{n-1}+a_{n-2}$ для $n \geqslant 2$. (б) Это верно ввиду того, что a_n четное и $-1 < 1-\sqrt{2} < 0$. (в) Положим

$$b_n = \left(\frac{p+\sqrt{q}}{2}\right)^n+\left(\frac{p-\sqrt{q}}{2}\right)^n.$$

Нас устроило бы, если бы b_n было нечетным для всех $n > 0$ и $-1 < (p-\sqrt{q})/2 < 0$. Действуя, как в п. (а), мы находим $b_0 = 2$, $b_1 = p$ и $b_n = pb_{n-1}+\frac{1}{4}(q-p^2)b_{n-2}$ для $n \geqslant 2$. Одно из возможных решений получается при $p = 3$ и $q = 17$.

7.50 Расширив идею умножения многоугольников из упр. 22, получим



Заменим каждый n-угольник на z^{n-2}. Такая замена правильно сочетается с умножением, поскольку операция склейки преобразует m- и n-угольники в $(m+n-2)$-угольник. Таким образом, производящая функция равна

$$Q = 1+zQ^2+z^2Q^3+z^3Q^4+\cdots = 1+\frac{zQ^2}{1-zQ},$$

и формула для вычисления корней квадратного уравнения дает $Q = \bigl(1+z-\sqrt{1-6z+z^2}\,\bigr)/4z$. Коэффициент при z^{n-2} в этом степенном ряду равняется числу способов провести непересекающиеся диагонали в выпуклом n-угольнике. Эти коэффициенты, по-видимому, не имеют простого выражения в аналитическом виде через другие величины, которые мы обсуждали в этой книге. Однако их асимптотическое поведение известно [207, упр. 2.2.1–12].

Дайте мне полиномы Лежандра, и я найду аналитический вид.

Кстати, если заменить каждый n-угольник в Q на wz^{n-2}, можно получить

$$Q = \frac{1+z-\sqrt{1-(4w+2)z+z^2}}{2(1+w)z},$$

формулу, в которой коэффициент при $w^m z^{n-2}$ представляет собой количество способов разбить n-угольник непересекающимися диагоналями на m многоугольников.

7.51 Ключевой первый шаг состоит в наблюдении, что квадрат интересующего нас количества способов есть количество циклических схем определенного вида, обобщающих упр. 27. Эти схемы могут быть подсчитаны посредством вычисления определителя некоторой матрицы, для которой несложно найти собственные значения. В случае $m = 3$ и $n = 4$ оказывается полезным тот факт, что $\cos 36° = \phi/2$ (упр. 6.46).

7.52 Несколько начальных случаев таковы: $p_0(y)=1$, $p_1(y)=y$, $p_2(y) = y^2+y$, $p_3(y) = y^3+3y^2+3y$. Пусть $p_n(y) = q_{2n}(x)$, где $y = x(1-x)$; мы ищем производящую функцию, которая подходящим образом определит $q_{2n+1}(x)$. Одна такая функция — это $\sum_n q_n(x)z^n/n! = 2e^{ixz}/(e^{iz}+1)$, откуда следует, что $q_n(x) = i^n E_n(x)$, где $E_n(x)$ называется полиномом Эйлера. Имеем $\sum(-1)^x x^n\,\delta x = \frac{1}{2}(-1)^{x+1}E_n(x)$, поэтому полиномы Эйлера аналогичны полиномам Бернулли и раскладываются на множители, аналогичные (6.98). Из упр. 6.23 получаем $nE_{n-1}(x) =$

$\sum_{k=0}^{n}\binom{n}{k}B_k x^{n-k}(2-2^{k+1})$; этот полином согласно упр. 6.54 имеет целые коэффициенты. Следовательно, коэффициенты полинома $q_{2n}(x)$, со знаменателями в виде степеней двойки, также должны быть целыми. Значит, $p_n(y)$ имеет целые коэффициенты. Наконец, соотношение $(4y-1)p_n''(y)+2p_n'(y) = 2n(2n-1)p_{n-1}(y)$ показывает, что

$$2m(2m-1)\left|{n \atop m}\right| = m(m+1)\left|{n \atop m+1}\right|+2n(2n-1)\left|{n-1 \atop m-1}\right|,$$

а отсюда вытекает, что числа $\left|{n \atop m}\right|$ — положительны. (Подобное доказательство позволяет установить, что аналогичная величина $(-1)^n(2n+2)E_{2n+1}(x)/(2x-1)$, если ее выразить в виде полинома n-й степени от y, имеет положительные целые коэффициенты.) Можно показать, что $\left|{n \atop 1}\right|$ — число Дженоччи $(-1)^{n-1}(2^{2n+1}-2)\times B_{2n}$ (см. упр. 6.24) и что $\left|{n \atop n-1}\right| = \binom{n}{2}$, $\left|{n \atop n-2}\right| = 2\binom{n+1}{4}+3\binom{n}{4}$, и т.д.

7.53 Это число $P_{(1+V_{4n+1}+V_{4n+3})/6}$. Так, например, $T_{20} = P_{12} = 210$; $T_{285} = P_{165} = 40755$.

7.54 Пусть E_k означает операцию над степенными рядами, состоящую в обнулении всех коэффициентов при z^n, за исключением $n \bmod m = k$. Тогда описанная конструкция эквивалентна операции

$$E_0\, S\, E_0\, S\, (E_0+E_1)\, S\, \ldots\, S\, (E_0+E_1+\cdots+E_{m-1}),$$

примененной к $1/(1-z)$, где S означает "умножить на $1/(1-z)$". Это выражение содержит $m!$ членов

$$E_0\, S\, E_{k_1}\, S\, E_{k_2}\, S\, \ldots\, S\, E_{k_m}\,.$$

Здесь $0 \leqslant k_j < j$, и каждый такой член дает $z^{rm}/(1-z^m)^{m+1}$, где r — число мест, где $k_j < k_{j+1}$. Имеется ровно $\left\langle{m \atop r}\right\rangle$ слагаемых с данным значением r, так что коэффициентом при z^{mn} в соответствии с (6.37) будет $\sum_{r=0}^{m-1}\left\langle{m \atop r}\right\rangle\binom{n+m-r}{m} = (n+1)^m$. (Тот факт, что операция E_k может быть выражена через комплексные корни единицы, по всей видимости, нисколько не помогает в решении поставленной задачи.)

7.55 Пусть $P_0(z)F(z) + \cdots + P_m(z)F^{(m)}(z) = Q_0(z)G(z) + \cdots + Q_n(z)G^{(n)}(z) = 0$, где $P_m(z)$ и $Q_n(z)$ — ненулевые. (a) Пусть $H(z) = F(z)+G(z)$. Тогда найдутся рациональные функции $R_{k,l}(z)$ для $0 \leqslant l < m+n$, такие, что $H^{(k)}(z) = R_{k,0}(z)F^{(0)}(z) + \cdots + R_{k,m-1}(z)F^{(m-1)}(z) + R_{k,m}(z)G^{(0)}(z) + \cdots + R_{k,m+n-1}(z)G^{(n-1)}(z)$. Векторы $\bigl(R_{k,0}(z), \ldots, R_{k,m+n-1}(z)\bigr)$ в количестве $m+n+1$ линейно зависимы в $(m+n)$-мерном векторном пространстве рациональных функций; следовательно, существуют рациональные функции $S_l(z)$, не все нулевые, такие, что $S_0(z)H^{(0)}(z) + \cdots + S_{m+n}(z)H^{(m+n)}(z) = 0$. (б) Аналогично положим $H(z)=F(z)G(z)$. Имеются рациональные функции $R_{k,l}(z)$ для $0 \leqslant l < mn$, такие, что $H^{(k)}(z) = \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} R_{k,ni+j}(z)F^{(i)}(z)G^{(j)}(z)$, следовательно, $S_0(z)H^{(0)}(z) + \cdots + S_{mn}(z)H^{(mn)}(z) = 0$ для некоторых рациональных $S_l(z)$, не все из которых нулевые. (Аналогичным образом можно доказать, что если $\langle f_n\rangle$ и $\langle g_n\rangle$ полиномиально рекурсивны, то таковы и последовательности $\langle f_n + g_n\rangle$ и $\langle f_n g_n\rangle$. Кстати, для частных подобный результат отсутствует; например, $\cos z$ является дифференцируемо конечной функцией, тогда как $1/\cos z$ таковой не является.)

7.56 Эйлер [113] показал, что это число есть также $[z^n]\,1/\sqrt{1-2z-3z^2}$, и дал формулу $t_n = \sum_{k\geqslant 0} n^{\underline{2k}}/k!^2 = \sum_k \binom{n}{k}\binom{n-k}{k}$. При исследовании этих чисел он также открыл “достопамятное фиаско индукции”: хотя $3t_n - t_{n+1}$ и равно $F_{n-1}(F_{n-1}+1)$ при $0 \leqslant n \leqslant 8$, этот эмпирический закон таинственным образом нарушается при n, равном 9, или большем! Джордж Эндрюс (George Andrews) [12] разъяснил тайну, показав, что сумма $\sum_k [z^{n+10k}] \times (1+z+z^2)^n$ может быть выражена в аналитическом виде через числа Фибоначчи.

Г. С. Вильф (H. S. Wilf) заметил, что $[z^n]\,(a+bz+cz^2)^n = [z^n]\,1/f(z)$, где $f(z) = \sqrt{1-2bz+(b^2-4ac)z^2}$ (см. [373, с. 159]); отсюда следует, что коэффициенты удовлетворяют соотношению

$$(n+1)A_{n+1} - (2n+1)bA_n + n(b^2-4ac)A_{n-1} \;=\; 0\,.$$

Применив алгоритм Петковшека (Petkovšek) [291], можно доказать, что это рекуррентное соотношение имеет аналитическое решение в виде конечной суммы гипергеометрических членов тогда и только тогда, когда $abc(b^2-4ac)=0$. В частности, поэтому средний триномиальный коэффициент не имеет такого выражения в аналитическом виде. Следующий шаг должен, вероятно, заключаться в распространении этого результата на более широкий класс аналитических выражений (включая, например, гармонические числа и/или числа Стирлинга).

Дайте мне полиномы Лежандра, и я найду аналитический вид.

7.57 (Поль Эрдёш (Paul Erdős) неоднократно предлагал за решение этой задачи 500 долларов.)

8.1 $\frac{1}{24}+\frac{1}{48}+\frac{1}{48}+\frac{1}{48}+\frac{1}{48}+\frac{1}{24}=\frac{1}{6}$. (На самом деле мы *всегда* получим дубль с вероятностью $\frac{1}{6}$, если хотя бы одна игральная кость правильная.) Любые две грани с суммой 7 имеют одинаковые вероятности в распределении $\Pr_1$, так что событие $S=7$ имеет ту же вероятность, что и выпадение дубля.

8.2 Имеется 12 способов задать верхнюю и нижнюю карты и 50! способов разместить остальные карты, так что искомая вероятность равна $12\cdot 50!/52! = 12/(51\cdot 52) = \frac{1}{1713} = \frac{1}{221}$.

8.3 $\frac{1}{10}(3+2+\cdots+9+2) = 4.8$; $\frac{1}{9}(3^2+2^2+\cdots+9^2+2^2-10(4.8)^2) = \frac{388}{45}$, что приближенно равно 8.6. Истинные математическое ожидание и дисперсия для правильной монеты равны 6 и 22, так что Станфордским студентам везет на орлы. Соответствующие числа для Принстона равны 6.4 и $\frac{562}{45} \approx 12.5$. (Это распределение имеет $\kappa_4 = 2974$, что весьма много. Следовательно, стандартное отклонение этой оценки дисперсии для $n = 10$ также довольно велико, $\sqrt{2974/10 + 2(22)^2/9} \approx 20.1$ в соответствии с упр. 54. Студентов нельзя обвинить в жульничестве.)

8.4 Это следует из (8.38) и (8.39), потому что $F(z) = G(z)H(z)$. (Подобные формулы имеют место для всех кумулянтов, несмотря на то что $F(z)$ и $G(z)$ могут иметь отрицательные коэффициенты.)

8.5 Замените 0 на p и P на $q = 1 - p$. Если $S_A = S_B = \frac{1}{2}$, то $p^2qN = \frac{1}{2}$ и $pq^2N = \frac{1}{2}q + \frac{1}{2}$; решением будет $p = 1/\phi^2$, $q = 1/\phi$.

8.6 В этом случае $X|y$ имеет для всех y то же распределение, что и X, следовательно, $E(X|Y) = EX$ есть константа, и $V\big(E(X|Y)\big) = 0$. Также $V(X|Y)$ — константа и совпадает со своим ожидаемым значением.

8.7 Имеем $1 = (p_1 + p_2 + \cdots + p_6)^2 \leqslant 6(p_1^2 + p_2^2 + \cdots + p_6^2)$ по неравенству Чебышева для сумм из главы 2.

8.8 Пусть $p = \Pr(\omega \in A \cap B)$, $q = \Pr(\omega \notin A)$ и $r = \Pr(\omega \notin B)$. Тогда $p + q + r = 1$ и доказываемое тождество имеет вид $p = (p+r)(p+q) - qr$.

8.9 Это верно (при очевидном условии, что F и G определены на областях изменения случайных величин X и Y), потому что

$$\begin{aligned}\Pr\big(F(X)=f \text{ и } G(Y)=g\big) &= \sum_{\substack{x\in F^{-1}(f)\\ y\in G^{-1}(g)}} \Pr(X=x \text{ и } Y=y) = \\ &= \sum_{\substack{x\in F^{-1}(f)\\ y\in G^{-1}(g)}} \Pr(X=x)\cdot\Pr(Y=y) = \\ &= \Pr\big(F(X)=f\big)\cdot\Pr\big(G(Y)=g\big).\end{aligned}$$

8.10 Два. Пусть $x_1 < x_2$ — медианы; тогда $1 \leqslant \Pr(X \leqslant x_1) + \Pr(X \geqslant x_2) \leqslant 1$, следовательно, имеет место равенство. (Некоторые дискретные распределения вовсе не имеют медианы. Пусть, например, Ω представляет собой множество всех дробей вида $\pm 1/n$ с вероятностями $\Pr(+1/n) = \Pr(-1/n) = \frac{\pi^2}{12}n^{-2}$.)

8.11 Например, пусть $K = k$ с вероятностью $4/(k+1)(k+2)(k+3)$ для всех целых $k \geqslant 0$. Тогда $EK = 1$, но $E(K^2) = \infty$. (Аналогично можно построить случайную величину со всеми конечными кумулянтами до κ_m, но с $\kappa_{m+1} = \infty$.)

8.12 (а) Пусть $p_k = \Pr(X{=}k)$. Если $0 < x \leqslant 1$, имеем $\Pr(X \leqslant r) = \sum_{k\leqslant r} p_k \leqslant \sum_{k\leqslant r} x^{k-r}p_k \leqslant \sum_k x^{k-r}p_k = x^{-r}P(x)$. Второе неравенство доказывается аналогично. (б) Положим $x = \alpha/(1-\alpha)$ с целью минимизации правой части. (Более точная оценка данной суммы имеется в упр. 9.42.)

8.13 (Решение Б. Питтеля (B. Pittel).) Положим $Y = (X_1 + \cdots + X_n)/n$ и $Z = (X_{n+1} + \cdots + X_{2n})/n$. Тогда

$$\begin{aligned} \Pr\left(\left|\frac{Y+Z}{2} - \alpha\right| \leqslant |Y-\alpha|\right) &\geqslant \\ \geqslant \Pr\left(\left|\frac{Y-\alpha}{2}\right| + \left|\frac{Z-\alpha}{2}\right| \leqslant |Y-\alpha|\right) &= \\ = \Pr\big(|Z-\alpha| \leqslant |Y-\alpha|\big) &\geqslant \tfrac{1}{2}. \end{aligned}$$

Последнее неравенство фактически можно заменить на '$>$' для любого дискретного распределения вероятностей, поскольку $\Pr(Y{=}Z) > 0$.

8.14 $\mathrm{Mean}(H) = p\,\mathrm{Mean}(F) + q\,\mathrm{Mean}(G)$; $\mathrm{Var}(H) = p\,\mathrm{Var}(F) + q\,\mathrm{Var}(G) + pq\big(\mathrm{Mean}(F) - \mathrm{Mean}(G)\big)^2$. (Смесь на самом деле есть частный случай условного распределения: пусть Y отвечает бросанию монеты, $X|0$ порождается функцией $F(z)$, а $X|P$ — функцией $G(z)$. Тогда $VX = EV(X|Y) + VE(X|Y)$, где $EV(X|Y) = pV(X|0) + qV(X|P)$, а $VE(X|Y)$ есть дисперсия функции $pz^{\mathrm{Mean}(F)} + qz^{\mathrm{Mean}(G)}$.)

8.15 Согласно правилу дифференцирования сложной функции $H'(z) = G'(z)F'\big(G(z)\big)$; $H''(z) = G''(z)F'\big(G(z)\big) + G'(z)^2F''\big(G(z)\big)$. Следовательно,

$$\begin{aligned} \mathrm{Mean}(H) &= \mathrm{Mean}(F)\,\mathrm{Mean}(G)\,; \\ \mathrm{Var}(H) &= \mathrm{Var}(F)\,\mathrm{Mean}(G)^2 + \mathrm{Mean}(F)\,\mathrm{Var}(G)\,. \end{aligned}$$

(Случайную величину, отвечающую распределению H, можно описать следующим образом: определим неотрицательное целое число n в соответствии с распределением F; затем сложим n независимых случайных величин, имеющих распределение G. Тождество для дисперсии из этого упражнения является частным случаем (8.106), где X имеет распределение H, а Y — распределение F.)

8.16 $e^{w(z-1)}/(1-w)$.

8.17 $\Pr(Y_{n,p} \leqslant m) = \Pr(Y_{n,p} + n \leqslant m+n)$ = вероятность того, что для выпадания n орлов требуется $\leqslant m+n$ бросаний = вероятность того, что $m+n$ бросаний дадут $\geqslant n$ орлов $= \Pr(X_{m+n,p} \geqslant n)$. Следовательно,

$$\sum_{k\leqslant m} \binom{n+k-1}{k} p^n q^k = \sum_{k\geqslant n} \binom{m+n}{k} p^k q^{m+n-k} = \\ = \sum_{k\leqslant m} \binom{m+n}{k} p^{m+n-k} q^k;$$

а это есть тождество (5.19) с $n = r$, $x = q$, $y = p$.

8.18 (а) $G_X(z) = e^{\mu(z-1)}$. (б) Для всех $m \geqslant 1$ m-й кумулянт равен μ. (Случай $\mu = 1$ в (8.55) назван F_∞.)

8.19 (а) $G_{X_1+X_2}(z) = G_{X_1}(z)G_{X_2}(z) = e^{(\mu_1+\mu_2)(z-1)}$. Следовательно, вероятность равна $e^{-\mu_1-\mu_2}(\mu_1+\mu_2)^n/n!$; сумма независимых пуассоновских случайных величин есть пуассоновская случайная величина. (б) В общем случае, если K_mX обозначает m-й кумулянт случайной величины X, то для $a, b \geqslant 0$ имеем $K_m(aX_1 + bX_2) = a^m(K_mX_1) + b^m(K_mX_2)$. Следовательно, ответом будет $2^m\mu_1 + 3^m\mu_2$.

8.20 В общем случае производящая функция случайной величины имеет вид $G(z) = z^m/F(z)$, где

$$F(z) = z^m + (1-z)\sum_{k=1}^{m} \widetilde{A}_{(k)}[A^{(k)} = A_{(k)}]z^{m-k},$$

$$F'(1) = m - \sum_{k=1}^{m} \widetilde{A}_{(k)}[A^{(k)} = A_{(k)}],$$

$$F''(1) = m(m-1) - 2\sum_{k=1}^{m}(m-k)\widetilde{A}_{(k)}[A^{(k)} = A_{(k)}].$$

8.21 Это значение есть $\sum_{n\geqslant 0} q_n$, где q_n — вероятность того, что игра Алисы и Билла после n бросаний не завершена. Пусть p_n означает вероятность окончания игры на n-м бросании; тогда $p_n + q_n = q_{n-1}$. Следовательно, среднее время игры равняется $\sum_{n\geqslant 1} np_n = (q_0 - q_1) + 2(q_1 - q_2) + 3(q_2 - q_3) + \cdots = q_0 + q_1 + q_2 + \cdots = N$, поскольку $\lim_{n\to\infty} nq_n = 0$.

Другой способ получить тот же ответ состоит в том, чтобы заменить 0 и P на $\frac{1}{2}z$. Тогда, продифференцировав первое уравнение в (8.78), получим $N(1)+N'(1) = N'(1)+S'_A(1)+S'_B(1)$.

Кстати, $N = \frac{16}{3}$.

8.22 По определению имеем $V(X|Y) = E(X^2|Y) - \big(E(X|Y)\big)^2$ и $V\big(E(X|Y)\big) = E\big((E(X|Y))^2\big) - \big(E\big(E(X|Y)\big)\big)^2$; следовательно, $E\big(V(X|Y)\big) + V\big(E(X|Y)\big) = E\big(E(X^2|Y)\big) - \big(E\big(E(X|Y)\big)\big)^2$. Однако $E\big(E(X|Y)\big) = \sum_y \Pr(Y{=}y)E(X|y) = \sum_{x,y}\Pr(Y{=}y)\Pr\big((X|y){=}x\big)x = EX$ и $E\big(E(X^2|Y)\big) = E(X^2)$, так что результат представляет собой просто VX.

8.23 Пусть $\Omega_0 = \{$⚀, ⚅$\}^2$ и $\Omega_1 = \{$⚁, ⚂, ⚃, ⚄$\}^2$; и пусть Ω_2 — множество из остальных 16 элементов Ω. Тогда $\Pr_{11}(\omega) - \Pr_{00}(\omega) = \frac{+20}{576}, \frac{-7}{576}, \frac{+2}{576}$, если $\omega \in \Omega_0, \Omega_1, \Omega_2$. Следовательно, при выборе событий A нужно взять k_j элементов из Ω_j, где (k_0, k_1, k_2) — один из следующих наборов: $(0,0,0)$, $(0,2,7)$, $(0,4,14)$, $(1,4,4)$, $(1,6,11)$, $(2,6,1)$, $(2,8,8)$, $(2,10,15)$, $(3,10,5)$, $(3,12,12)$, $(4,12,2)$, $(4,14,9)$, $(4,16,16)$. Например, имеется $\binom{4}{2}\binom{16}{6}\binom{16}{1}$ событий типа $(2,6,1)$. Общее количество таких событий составляет $[z^0](1+z^{20})^4(1+z^{-7})^{16}(1+z^2)^{16}$, что равно 1304872090. Если ограничиться событиями, зависящими только от S, можно получить 40 решений $S \in A$, где $A = \emptyset$, $\{{}^{\,2}_{12}, {}^{\,4}_{10}, {}^{6}_{8}\}$, $\{{}^{\,2}_{12}, 5, 9\}$, $\{2, 12, {}^{\,4}_{10}, {}^{6}_{8}, 5, 9\}$, $\{2,4,6,8,10,12\}$, $\{{}^{\,3}_{11}, 7, {}^{5}_{9}, 4, 10\}$, а также дополнения к этим множествам. (Здесь запись '${}^{\,2}_{12}$' означает 2 или 12, но не оба вместе.)

8.24 (a) Любая игральная кость в конечном счете окажется во владении игрока J с вероятностью $p = \frac{1}{6} + \left(\frac{5}{6}\right)^2 p$; следовательно, $p = \frac{6}{11}$. Пусть $q = \frac{5}{11}$. Тогда производящая функция случайной величины числа игральных костей у J будет $(q+pz)^{2n+1}$ с математическим ожиданием $(2n+1)p$ и дисперсией $(2n+1)pq$, как следует из (8.61). (б) $\binom{5}{3}p^3q^2 + \binom{5}{4}p^4q + \binom{5}{5}p^5 = \frac{94176}{161051} \approx .585$.

8.25 Производящая функция случайной величины суммы ставки после n бросков равна $G_n(z)$, где

$$
\begin{aligned}
G_0(z) &= z^A;\\
G_n(z) &= \textstyle\sum_{k=1}^{6} G_{n-1}(z^{2(k-1)/5})/6 \qquad \text{для } n > 0.
\end{aligned}
$$

Эту задачу, пожалуй, проще решить без применения производящих функций, чем с ними.

(Нецелые показатели степеней не вызывают трудностей.) Отсюда следует, что $\mathrm{Mean}(G_n) = \mathrm{Mean}(G_{n-1})$ и $\mathrm{Var}(G_n) + \mathrm{Mean}(G_n)^2 = \frac{22}{15}\big(\mathrm{Var}(G_{n-1}) + \mathrm{Mean}(G_{n-1})^2\big)$. Таким образом, математическое ожидание всегда равно A, в то время как дисперсия возрастает до значения $\big(\left(\frac{22}{15}\right)^n - 1\big)A^2$.

8.26 Производящая функция случайной величины $F_{l,n}(z)$ удовлетворяет уравнению $F'_{l,n}(z) = F_{l,n-l}(z)/l$; следовательно, $\mathrm{Mean}(F_{l,n}) = F'_{l,n}(1) = [n \geqslant l]/l$ и $F''_{l,n}(1) = [n \geqslant 2l]/l^2$; дисперсия вычисляется очень легко. (Фактически мы имеем

$$F_{l,n}(z) = \sum_{0 \leqslant k \leqslant n/l} \frac{1}{k!}\left(\frac{z-1}{l}\right)^k,$$

что стремится при $n \to \infty$ к распределению Пуассона с математическим ожиданием $1/l$.)

8.27 Величина $(n^2\Sigma_3 - 3n\Sigma_2\Sigma_1 + 2\Sigma_1^3)/n(n-1)(n-2)$, где $\Sigma_k = X_1^k + \cdots + X_n^k$, имеет желаемое значение. Это вытекает из тождеств

$$\begin{aligned} E\Sigma_3 &= n\mu_3\,; \\ E(\Sigma_2\Sigma_1) &= n\mu_3 + n(n-1)\mu_2\mu_1\,; \\ E(\Sigma_1^3) &= n\mu_3 + 3n(n-1)\mu_2\mu_1 + n(n-1)(n-2)\mu_1^3\,. \end{aligned}$$

Кстати, третий кумулянт равен $\kappa_3 = E\bigl((X - EX)^3\bigr)$, но уже для четвертого кумулянта столь простое соотношение не выполняется: $\kappa_4 = E\bigl((X-EX)^4\bigr) - 3(VX)^2$.

8.28 (В упражнении неявно подразумевается, что $p = q = \frac{1}{2}$, но далее для полноты приводится ответ в общем случае.) Замена O на pz и P на qz дает $S_A(z) = p^2qz^3/(1-pz)(1-qz)(1-pqz^2)$ и $S_B(z) = pq^2z^3/(1-qz)(1-pqz^2)$. Для условной вероятности выигрыша Алисы после n-го броска, при условии, что она выигрывает, получаем следующую производящую функцию случайной величины:

$$\frac{S_A(z)}{S_A(1)} = z^3 \cdot \frac{q}{1-pz} \cdot \frac{p}{1-qz} \cdot \frac{1-pq}{1-pqz^2}\,.$$

Это произведение псевдопроизводящих функций случайных величин, математическое ожидание которого равно $3 + p/q + q/p + 2pq/(1-pq)$. Формулы для Билла отличаются отсутствием множителя $q/(1-pz)$, так что математическое ожидание для Билла равно $3 + q/p + 2pq/(1-pq)$. При $p = q = \frac{1}{2}$ ответом в случае (а) будет $\frac{17}{3}$; в случае (б) — $\frac{14}{3}$. Билл выигрывает в два раза реже, но если ему случается выиграть, то это происходит быстрее. Общее среднее число бросаний составляет $\frac{2}{3} \cdot \frac{17}{3} + \frac{1}{3} \cdot \frac{14}{3} = \frac{16}{3}$, что согласуется с упр. 21. При игре в одиночку для любой последовательности время ожидания равно 8.

8.29 Подстановка $\mathrm{O} = \mathrm{P} = \frac{1}{2}$ в

$$\begin{aligned} 1 + N(\mathrm{O} + \mathrm{P}) &= N + S_A + S_B + S_C \\ N\,\mathrm{OOPO} &= S_A(\mathrm{OPO} + 1) + S_B(\mathrm{OPO} + \mathrm{PO}) + S_C(\mathrm{OPO} + \mathrm{PO}) \end{aligned}$$

$$\begin{aligned} N\,0P00 &= S_A(P00+0)+S_B(P00+1)+S_C(P00)\\ N\,P000 &= S_A(00)+S_B(0)+S_C \end{aligned}$$

дает вероятности выигрыша. В общем случае будет иметь место $S_A+S_B+S_C=1$ и

$$\begin{aligned} S_A(A{:}A)+S_B(B{:}A)+S_C(C{:}A) &= S_A(A{:}B)+S_B(B{:}B)+S_C(C{:}B) =\\ &= S_A(A{:}B)+S_B(B{:}C)+S_C(C{:}C)\,. \end{aligned}$$

В частности, уравнения $9S_A+3S_B+3S_C = 5S_A+9S_B+S_C = 2S_A+4S_B+8S_C$ дают $S_A=\frac{16}{52}$, $S_B=\frac{17}{52}$, $S_C=\frac{19}{52}$.

8.30 Дисперсия $P(h_1,\ldots,h_n;k)|k$ есть дисперсия сдвинутого биномиального распределения $\bigl((m-1+z)/m\bigr)^{k-1}z$, которая составляет $(k-1)(\frac{1}{m})(1-\frac{1}{m})$ в силу (8.61). Следовательно, среднее дисперсий будет равно $\mathrm{Mean}(S)(m-1)/m^2$. Дисперсия средних — это дисперсия величины $(k-1)/m$, а именно $\mathrm{Var}(S)/m^2$. В соответствии с (8.106), сумма этих двух чисел должна быть равна VP, и это действительно так. Фактически мы в слегка замаскированном виде повторили вывод (8.96). (См. упр. 15.)

8.31 (a) Действуя методом грубой силы, мы получаем систему из пяти уравнений с пятью неизвестными:

$$\begin{aligned} A &= \tfrac12 zB+\tfrac12 zE\,; \quad B = \tfrac12 zC\,; \quad C = 1+\tfrac12 zB+\tfrac12 zD\,;\\ D &= \tfrac12 zC+\tfrac12 zE\,; \quad E = \tfrac12 zD\,. \end{aligned}$$

Однако позиции C и D равноудалены от цели, так же как и B и E, так что их можно объединить друг с другом. Если принять $X=B+E$ и $Y=C+D$, то у нас останется три уравнения:

$$A = \tfrac12 zX\,;\quad X = \tfrac12 zY\,;\quad Y = 1+\tfrac12 zX+\tfrac12 zY\,.$$

Следовательно, $A = z^2/(4-2z-z^2)$; отсюда $\mathrm{Mean}(A)=6$ и $\mathrm{Var}(A)=22$. (Слышите звон? Эта задача фактически эквивалентна подбрасыванию правильной монеты до появления двух орлов подряд: орел означает "сделать шаг к яблоку", а решка — "сделать шаг назад".) (б) Неравенство Чебышева говорит нам, что $\Pr(S\geqslant 100) = \Pr\bigl((S-6)^2\geqslant 94^2\bigr) \leqslant 22/94^2 \approx .0025$. (в) Второе неравенство хвоста говорит, что $\Pr(S\geqslant 100)\leqslant 1/x^{98}(4-2x-x^2)$ для всех $x\geqslant 1$; подставив $x=(\sqrt{49001}-99)/100$, мы получим верхнюю границу 0.00000005. (В действительности эта вероятность равна приблизительно 0.0000000009 согласно упр. 37.)

8.32 Из соображений симметрии ситуацию в каждом месяце можно свести к одной из четырех возможностей:

"Элли... знала, что только Великий Гудвин вернет ее в Канзас".

D, два штата расположены по диагонали;
A, штаты расположены по соседству, и среди них нет Канзаса;
K, два штата — это Канзас и какой-то иной;
S, два штата совпадают.

Рассматривая переходы в полученной цепи Маркова, получаем четыре уравнения,

$$\begin{aligned} D &= 1 + z(\tfrac{2}{9}D + \tfrac{2}{12}K), \\ A &= z(\tfrac{4}{9}A + \tfrac{4}{12}K), \\ K &= z(\tfrac{4}{9}D + \tfrac{4}{9}A + \tfrac{4}{12}K), \\ S &= z(\tfrac{3}{9}D + \tfrac{1}{9}A + \tfrac{2}{12}K), \end{aligned}$$

сумма которых равна $D+K+A+S = 1+z(D+A+K)$. Решением является

$$S = \frac{81z - 45z^2 - 4z^3}{243 - 243z + 24z^2 + 8z^3},$$

но простейший способ найти математическое ожидание и дисперсию, по-видимому, в том, чтобы подставить $z = 1+w$ и выполнить разложение по степеням w, игнорируя кратные w^2:

$$\begin{aligned} D &= \tfrac{27}{16} + \tfrac{1593}{512}w + \cdots; \\ A &= \tfrac{9}{8} + \tfrac{2115}{256}w + \cdots; \\ K &= \tfrac{15}{8} + \tfrac{2661}{256}w + \cdots. \end{aligned}$$

Теперь $S'(1) = \frac{27}{16} + \frac{9}{8} + \frac{15}{8} = \frac{75}{16}$ и $\frac{1}{2}S''(1) = \frac{1593}{512} + \frac{2115}{256} + \frac{2661}{256} = \frac{11145}{512}$. Математическое ожидание равно $\frac{75}{16}$, а дисперсия — $\frac{105}{4}$. (Есть ли более простой путь?)

8.33 Первый ответ: конечно, да, ведь хеш-значения $h_1, \ldots, h_n$ независимы! По здравом размышлении второй ответ — конечно, нет, несмотря на независимость хеш-значений $h_1, \ldots, h_n$. Имеем $\Pr(X_j = 0) = \sum_{k=1}^{n} s_k\bigl([j \ne k](m-1)/m\bigr) = (1-s_j)(m-1)/m$, но $\Pr(X_1 = X_2 = 0) = \sum_{k=1}^{n} s_k[k>2](m-1)^2/m^2 = (1-s_1-s_2)(m-1)^2/m^2 \ne \Pr(X_1 = 0)\Pr(X_2 = 0)$.

8.34 Пусть $[z^n]\,S_m(z)$ — вероятность того, что Джина продвинется менее чем на m шагов за n кругов. Тогда $S_m(1)$ — ее средний результат в m-ударной игре; $[z^m]\,S_m(z)$ — вероятность ее проигрыша в игре со стабильным партнером; а $1-[z^{m-1}]\,S_m(z)$ — вероятность выигрыша. Имеем рекуррентное соотношение

$$\begin{aligned} S_0(z) &= 0\,; \\ S_m(z) &= \bigl(1+pzS_{m-2}(z)+qzS_{m-1}(z)\bigr)/(1-rz) \qquad \text{для } m>0. \end{aligned}$$

Для решения п. (а) достаточно вычислить коэффициенты для $m, n \leqslant 4$; при этом удобно заменить z на $100w$, чтобы в вычислениях участвовали только целые числа. Получаем следующую таблицу коэффициентов:

S_0	0	0	0	0	0
S_1	1	4	16	64	256
S_2	1	95	744	4432	23552
S_3	1	100	9065	104044	819808
S_4	1	100	9975	868535	12964304

Следовательно, Джина выигрывает с вероятностью $1-.868535 = .131465$; проигрывает Джина с вероятностью .12964304. (б) Чтобы найти математическое ожидание количества ударов, вычисляем

$$S_1(1) = \tfrac{25}{24}\,; \quad S_2(1) = \tfrac{4675}{2304}\,; \quad S_3(1) = \tfrac{667825}{221184}\,; \quad S_4(1) = \tfrac{85134475}{21233664}\,.$$

(Оказывается, $S_5(1) \approx 4.9995$; она выигрывает по обоим критериям — лункам и ударам в игре с пятиударной лункой, но проигрывает обоими способами в игре, рассчитанной на три удара.)

8.35 Как следует из китайской теоремы об остатках, требуемое условие будет выполняться для всех n тогда и только тогда, когда оно выполняется для $n=1$. Одно необходимое и достаточное условие — тождество полиномов

$$\begin{aligned} &\bigl(p_2+p_4+p_6+(p_1+p_3+p_5)w\bigr)\bigl(p_3+p_6+(p_1+p_4)z+(p_2+p_5)z^2\bigr) \\ &\qquad = \bigl(p_1wz+p_2z^2+p_3w+p_4z+p_5wz^2+p_6\bigr)\,, \end{aligned}$$

но это только незначительная переформулировка исходной задачи. Более простым критерием является

$$(p_2+p_4+p_6)(p_3+p_6) = p_6\,, \quad (p_1+p_3+p_5)(p_2+p_5) = p_5\,,$$

что требует проверки всего двух коэффициентов в предыдущем произведении. Общее решение имеет три степени свободы: пусть $a_0+a_1 = b_0+b_1+b_2 = 1$; тогда можно взять $p_1 = a_1b_1$, $p_2 = a_0b_2$, $p_3 = a_1b_0$, $p_4 = a_0b_1$, $p_5 = a_1b_2$, $p_6 = a_0b_0$.

8.36 (a) ⚀ ⚁ ⚁ ⚂ ⚂ ⚃. (б) Если на гранях k-й игральной кости нанесено $s_1, \ldots, s_6$ очков, то положим $p_k(z) = z^{s_1} + \cdots + z^{s_6}$. Мы хотим найти такие полиномы p_k, для которых $p_1(z) \ldots p_n(z) = (z + z^2 + z^3 + z^4 + z^5 + z^6)^n$. Разложение этого полинома в произведение неприводимых множителей с рациональными коэффициентами представляет собой $z^n(z+1)^n \times (z^2+z+1)^n(z^2-z+1)^n$; следовательно, $p_k(z)$ должен иметь вид $z^{a_k}(z+1)^{b_k}(z^2+z+1)^{c_k}(z^2-z+1)^{d_k}$. При этом должно быть $a_k \geqslant 1$, поскольку $p_k(0) = 0$; и, более того, $a_k = 1$, поскольку $a_1 + \cdots + a_n = n$. Далее, условие $p_k(1) = 6$ означает, что $b_k = c_k = 1$. Теперь легко понять, что $0 \leqslant d_k \leqslant 2$, поскольку $d_k > 2$ дает отрицательные коэффициенты. Когда $d = 0$ и $d = 2$, мы получаем две игральные кости из п. (a); следовательно, все решения исчерпываются k парами игральных костей из п. (a) плюс $n - 2k$ обычных игральных костей для некоторого $k \leqslant \frac{1}{2}n$.

8.37 Количество последовательностей длиной n, описывающих возможные результаты бросания монеты, равно F_{n-1} для $n > 0$ в силу связи бросаний монеты и покрытий костями домино. Следовательно, вероятность того, что потребуется ровно n бросаний, составляет $F_{n-1}/2^n$, если монета правильная. Кроме того, $q_n = F_{n+1}/2^{n-1}$, поскольку $\sum_{k\geqslant n} F_k z^k = (F_n z^n + F_{n-1} z^{n+1})/(1-z-z^2)$. (Разумеется, возможно и систематическое решение поставленной задачи с применением производящих функций.)

8.38 После появления k граней задача выбрасывания еще одной эквивалентна бросанию монеты с вероятностью успеха $p_k = (m-k)/m$. Следовательно, соответствующая производящая функция случайной величины равна $\prod_{k=0}^{l-1} p_k z/(1-q_k z) = \prod_{k=0}^{l-1}(m-k)z/(m-kz)$. Математическое ожидание равно $\sum_{k=0}^{l-1} p_k^{-1} = m \times (H_m - H_{m-l})$; дисперсия равна $m^2\bigl(H_m^{(2)} - H_{m-l}^{(2)}\bigr) - m(H_m - H_{m-l})$; и уравнение (7.47) дает выражение в аналитическом виде для требуемой вероятности, а именно $m^{-n} m! \left\{ {n-1 \atop l-1} \right\}/(m-l)!$. (Обсуждаемая в этом упражнении задача обычно называется задачей "сбора купонов".)

8.39 $E(X) = P(-1)$; $V(X) = P(-2) - P(-1)^2$; $E(\ln X) = -P'(0)$.

8.40 (a) Имеем $\kappa_m = n\bigl(0!\left\{ {m \atop 1} \right\}p - 1!\left\{ {m \atop 2} \right\}p^2 + 2!\left\{ {m \atop 3} \right\}p^3 - \cdots\bigr)$ согласно (7.49). Кстати, третий кумулянт равен $npq(q-p)$, а четвертый — $npq(1-6pq)$. Тождество $q+pe^t = (p+qe^{-t})e^t$ показывает, что $f_m(p) = (-1)^m f_m(q) + [m=1]$; следовательно, можно записать $f_m(p) = g_m(pq)(q-p)^{[m\,]}$, где g_m — полином степени $\lfloor m/2 \rfloor$, если только $m > 1$. (б) Положим $p = \frac{1}{2}$ и $F(t) = \ln(\frac{1}{2} + \frac{1}{2}e^t)$. Тогда $\sum_{m\geqslant 1}\kappa_m t^{m-1}/(m-1)! = F'(t) = 1-1/(e^t+1)$, и можно воспользоваться упр. 6.23.

8.41 Если $G(z)$ представляет собой производящую функцию случайной величины X, которая принимает только целые положительные значения, то $\int_0^1 G(z)\,dz/z = \sum_{k\geqslant 1}\Pr(X=k)/k = E(X^{-1})$. Если X— распределение количества бросаний до получения $n+1$ орлов, то из (8.59) имеем $G(z) = \bigl(pz/(1-qz)\bigr)^{n+1}$, и выписанный интеграл можно преобразовать:

$$\int_0^1 \left(\frac{pz}{1-qz}\right)^{n+1}\frac{dz}{z} = \int_0^1 \frac{w^n\,dw}{1+(q/p)w},$$

если выполнить замену $w = pz/(1-qz)$. Для $p = q$ подынтегральное выражение можно записать как $(-1)^n\bigl((1+w)^{-1}-1+w-w^2+\cdots+(-1)^n w^{n-1}\bigr)$, так что интеграл равен $(-1)^n\bigl(\ln 2-1+\frac12-\frac13+\cdots+(-1)^n/n\bigr)$. Из (9.28) имеем $H_{2n}-H_n = \ln 2-\frac14 n^{-1}+\frac1{16}n^{-2}+O(n^{-4})$; отсюда следует, что $E(X_{n+1}^{-1}) = \frac12 n^{-1}-\frac14 n^{-2}+O(n^{-4})$.

8.42 Пусть $F_n(z)$ и $G_n(z)$— производящие функции случайных величин для количества рабочих вечеров, если человек первоначально был соответственно безработным или имел работу. Пусть $q_h = 1-p_h$ и $q_f = 1-p_f$. Тогда $F_0(z) = G_0(z) = 1$ и

$$\begin{aligned} F_n(z) &= p_h z G_{n-1}(z) + q_h F_{n-1}(z)\,; \\ G_n(z) &= p_f F_{n-1}(z) + q_f z G_{n-1}(z)\,. \end{aligned}$$

Решение дается суперпроизводящей функцией случайной величины

$$G(w,z) = \sum_{n\geqslant 0} G_n(z)w^n = A(w)/\bigl(1-zB(w)\bigr)\,,$$

где $B(w) = w\bigl(q_f-(q_f-p_h)w\bigr)/(1-q_h w)$ и $A(w) = \bigl(1-B(w)\bigr)/(1-w)$. Теперь получаем $\sum_{n\geqslant 0} G_n'(1)w^n = \alpha w/(1-w)^2+\beta/(1-w)-\beta/\bigl(1-(q_f-p_h)w\bigr)$, где

$$\alpha = \frac{p_h}{p_h+p_f}\,, \qquad \beta = \frac{p_f(q_f-p_h)}{(p_h+p_f)^2}\,;$$

следовательно, $G_n'(1) = \alpha n + \beta\bigl(1-(q_f-p_h)^n\bigr)$. (Аналогично $G_n''(1) = \alpha^2 n^2 + O(n)$, так что дисперсия равна $O(n)$.)

8.43 В силу (6.11) $G_n(z) = \sum_{k\geqslant 0}\left[{n \atop k}\right] z^k/n! = z^{\overline{n}}/n!$. Это— произведение биномиальных производящих функций случайных величин, $\prod_{k=1}^n\bigl((k-1+z)/k\bigr)$, где k-я функция имеет математическое ожидание $1/k$ и дисперсию $(k-1)/k^2$; следовательно, $\mathrm{Mean}(G_n) = H_n$ и $\mathrm{Var}(G_n) = H_n - H_n^{(2)}$.

8.44 (а) Чемпион должен пройти без поражений n кругов, так что искомая вероятность равна p^n. (б, в) Игроки $x_1, \ldots, x_{2^k}$ должны быть случайным образом "рассеяны" по разным подтурнирам и должны выиграть все $2^k(n-k)$ своих встреч. Игроков можно расположить по 2^n листьям дерева турнира $2^n!$ способами; чтобы обеспечить нужное рассеивание, имеется $2^k!(2^{n-k})^{2^k}$ способов размещения 2^k лучших игроков и $(2^n-2^k)!$ способов размещения остальных. Следовательно, искомая вероятность равна $(2p)^{2^k(n-k)}/\binom{2^n}{2^k}$. Для $k=1$ это выражение упрощается до $(2p^2)^{n-1}/(2^n-1)$. (г) Каждый результат турнира соответствует некоторой перестановке игроков: обозначим через y_1 чемпиона, через y_2 — второго финалиста; через y_3 и y_4 — игроков, проигравших y_1 и y_2 в полуфинале; через $(y_5, \ldots, y_8)$ — игроков, проигравших в четвертьфинале игрокам $(y_1, \ldots, y_4)$; и т.д. (Действуя иначе, можно показать, что на первом круге может быть $2^n!/2^{n-1}!$ существенно различных результатов; на втором — $2^{n-1}!/2^{n-2}!$ и т.д.) (д) Пусть S_k есть множество всех 2^{k-1} потенциальных противников x_2 на k-м круге. Условная вероятность того, что x_2 выиграет турнир при условии, что x_1 принадлежит S_k, равна

$$\begin{aligned}\Pr(x_1 \text{ играет с } x_2)\cdot p^{n-1}(1-p) + \Pr(x_1 \text{ не играет с } x_2)\cdot p^n &= \\ = p^{k-1}p^{n-1}(1-p) + (1-p^{k-1})p^n\,.&\end{aligned}$$

Вероятность того, что $x_1 \in S_k$, равна $2^{k-1}/(2^n-1)$; суммирование по k дает ответ:

$$\sum_{k=1}^{n} \frac{2^{k-1}}{2^n-1}\bigl(p^{k-1}p^{n-1}(1-p)+(1-p^{k-1})p^n\bigr) = p^n - \frac{(2p)^n-1}{2^n-1}\,p^{n-1}\,.$$

(е) Каждый из $2^n!$ результатов турнира имеет определенную вероятность; вероятность выигрыша x_j есть сумма этих вероятностей по всем $(2^n-1)!$ результатам, в которых x_j — чемпион. Переставим во всех этих турнирах x_j и x_{j+1}; такое изменение не влияет на вероятность, если x_j и x_{j+1} никогда не встречались один с другим, однако в случае их встречи вероятность победы x_j умножается на $(1-p)/p < 1$.

8.45 (а) $A(z) = 1/(3-2z)$; $B(z) = zA(z)^2$; $C(z) = z^2A(z)^3$. Производящая функция случайной величины для бутилируемого вина равна $z^3A(z)^3$, что есть произведение z^3 на отрицательное биномиальное распределение с параметрами $n=3$, $p=\frac{1}{3}$. (б) $\mathrm{Mean}(A) = 2$, $\mathrm{Var}(A) = 6$; $\mathrm{Mean}(B) = 5$, $\mathrm{Var}(B) = 2\,\mathrm{Var}(A) = 12$; $\mathrm{Mean}(C) = 8$, $\mathrm{Var}(C) = 18$. В среднем херес имеет возраст 9 лет. Доля 25-летнего вина составляет $\binom{-3}{22}(-2)^{22}3^{-25} =$

$\binom{24}{22}2^{22}3^{-25} = 23\cdot(\frac{2}{3})^{24} \approx .00137$. (в) Пусть коэффициентом при w^n будет производящая функция случайной величины на начало года n. Тогда

$$\begin{aligned} A &= \bigl(1+\tfrac{1}{3}w/(1-w)\bigr)/(1-\tfrac{2}{3}zw)\,;\\ B &= \bigl(1+\tfrac{1}{3}zwA\bigr)/(1-\tfrac{2}{3}zw)\,;\\ C &= \bigl(1+\tfrac{1}{3}zwB\bigr)/(1-\tfrac{2}{3}zw)\,. \end{aligned}$$

Продифференцируем по z и положим $z=1$; это даст

$$C' = \frac{8}{1-w} - \frac{1/2}{(1-\frac{2}{3}w)^3} - \frac{3/2}{(1-\frac{2}{3}w)^2} - \frac{6}{1-\frac{2}{3}w}\,.$$

Средний возраст бутилируемого хереса через n лет после начала процесса на 1 больше коэффициента при w^{n-1}, а именно он равен $9-(\frac{2}{3})^n(3n^2+21n+72)/8$. (Он превышает 8 уже при $n=11$.)

8.46 (а) $P(w,z) = 1+\frac{1}{2}\bigl(wP(w,z)+zP(w,z)\bigr) = \bigl(1-\frac{1}{2}(w+z)\bigr)^{-1}$, следовательно, $p_{m,n} = 2^{-m-n}\binom{m+n}{n}$. (б) $P_k(w,z) = \frac{1}{2}(w^k + z^k)P(w,z)$; следовательно,

$$p_{k,m,n} = 2^{k-1-m-n}\left(\binom{m+n-k}{m} + \binom{m+n-k}{n}\right).$$

(в) $\sum_k kp_{k,n,n} = \sum_{k=0}^{n} k2^{k-2n}\binom{2n-k}{n} = \sum_{k=0}^{n}(n-k)2^{-n-k}\binom{n+k}{n}$; эту сумму можно вычислить с помощью (5.20):

$$\begin{aligned} &\sum_{k=0}^{n} 2^{-n-k}\left((2n+1)\binom{n+k}{n} - (n+1)\binom{n+1+k}{n+1}\right) =\\ &\qquad = (2n+1)-(n+1)2^{-n}\left(2^{n+1} - 2^{-n-1}\binom{2n+2}{n+1}\right) =\\ &\qquad = \frac{2n+1}{2^{2n}}\binom{2n}{n} - 1\,. \end{aligned}$$

(Методы главы 9 показывают, что это $2\sqrt{n/\pi}-1+O(n^{-1/2})$.)

8.47 После поглощения n частиц имеется $n+2$ равноправных рецепторов. Пусть случайная величина X_n обозначает количество имеющихся дифагов; тогда $X_{n+1} = X_n + Y_n$, где $Y_n = -1$, если $(n+1)$-я ψ-частица попадает в рецептор дифага (условная вероятность — $2X_n/(n+2)$); в противном случае $Y_n = +2$. Следовательно,

$$EX_{n+1} = EX_n + EY_n = EX_n - 2EX_n/(n{+}2) + 2\bigl(1-2EX_n/(n{+}2)\bigr)\,.$$

Рекуррентное соотношение $(n+2)EX_{n+1} = (n-4)EX_n+2n+4$ может быть решено, если умножить обе его части на суммирующий множитель $(n+1)^{\underline{5}}$; можно также угадать ответ $EX_n = (2n+4)/7$ для всех $n > 4$ и доказать его по индукции. (Оказывается, после пятого шага всегда образуются два дифага и один трифаг независимо от состояния после четвертого шага.)

8.48 (a) Расстояние между дисками (измеряемое так, чтобы всегда получалось четное число) составляет 0, 2 или 4 единицы и изначально равно 4. Соответствующие производящие функции случайных величин A, B, C (где, скажем, $[z^n]\,C$ равняется вероятности расстояния 4 после n бросков) удовлетворяют соотношениям

$$A \;=\; \tfrac{1}{4}zB\,, \quad B \;=\; \tfrac{1}{2}zB+\tfrac{1}{4}zC\,, \quad C \;=\; 1+\tfrac{1}{4}zB+\tfrac{3}{4}zC\,.$$

Отсюда следует, что $A = z^2/(16-20z+5z^2) = z^2/F(z)$, и мы получаем $\mathrm{Mean}(A) = 2-\mathrm{Mean}(F) = 12$, $\mathrm{Var}(A) = -\mathrm{Var}(F) = 100$. (Более сложное, но и более интересное решение основано на разложении A следующим образом:

$$A \;=\; \frac{p_1z}{1-q_1z}\cdot\frac{p_2z}{1-q_2z} \;=\; \frac{p_2}{p_2-p_1}\frac{p_1z}{1-q_1z}+\frac{p_1}{p_1-p_2}\frac{p_2z}{1-q_2z}\,,$$

где $p_1 = \phi^2/4 = (3+\sqrt{5}\,)/8$, $p_2 = \hat{\phi}^2/4 = (3-\sqrt{5}\,)/8$ и $p_1+q_1 = p_2+q_2 = 1$. Таким образом, игра эквивалентна бросанию двух монет с вероятностями выпадания орла p_1 и p_2; монеты бросаются по одной, пока обе они не выпадут орлами вверх. Суммарное количество требуемых бросаний будет иметь то же распределение, что и число бросков дисков. Математическое ожидание и дисперсия времени ожидания для этих двух монет будут соответственно $6\mp2\sqrt{5}$ и $50\mp22\sqrt{5}$, следовательно, для общего математического ожидания и дисперсии получаем 12 и 100, как и ранее.)

(б) Разложение производящей функции на правильные дроби дает возможность просуммировать вероятности. (Заметьте, что $\sqrt{5}/(4\phi)+\phi^2/4 = 1$, так что ответ можно выразить через степени ϕ.) Игра будет длиться более n шагов с вероятностью $5^{(n-1)/2}4^{-n}(\phi^{n+2}-\phi^{-n-2})$; для четного n это равно $5^{n/2}4^{-n}F_{n+2}$. Таким образом, окончательный ответ — $5^{50}\times 4^{-100}F_{102}\approx .00006$.

8.49 (a) $P_N(0,n) = \frac{1}{2}[N=0]+\frac{1}{4}P_{N-1}(0,n)+\frac{1}{4}P_{N-1}(1,n-1)$ при $n>0$; для $P_N(m,0)$ имеет место аналогичная формула; $P_N(0,0) =$

$[N=0]$. Следовательно,

$$\begin{aligned} g_{m,n} &= \tfrac14 z g_{m-1,n+1} + \tfrac12 z g_{m,n} + \tfrac14 z g_{m+1,n-1}\,; \\ g_{0,n} &= \tfrac12 + \tfrac14 z g_{0,n} + \tfrac14 g_{1,n-1}\,; \quad \text{и т.д.} \end{aligned}$$

(б) $g'_{m,n} = 1 + \frac14 g'_{m-1,n+1} + \frac12 g'_{m,n} + \frac14 g'_{m+1,n-1}$; $g'_{0,n} = \frac12 + \frac14 g'_{0,n} + \frac14 g'_{1,n-1}$; и т.д. Индукцией по m получаем $g'_{m,n} = (2m+1)g'_{0,m+n} - 2m^2$ для всех $m, n \geqslant 0$. А поскольку $g'_{m,0} = g'_{0,m}$, должно быть $g'_{m,n} = m+n+2mn$. (в) Данная формула удовлетворяет рекуррентному соотношению при $mn > 0$, потому что

$$\sin(2m+1)\theta = \frac{1}{\cos^2\theta}\biggl(\frac{\sin(2m-1)\theta}{4} + \frac{\sin(2m+1)\theta}{2} + \frac{\sin(2m+3)\theta}{4}\biggr)\,;$$

это следует из тождества $\sin(x-y)+\sin(x+y) = 2\sin x\cos y$. Таким образом, остается только проверить граничные условия.

8.50 (а) Используя указание, получаем

$$3(1-z)^2 \sum_k \binom{1/2}{k}\Bigl(\frac89 z\Bigr)^k (1-z)^{2-k} = \\ = 3(1-z)^2 \sum_k \binom{1/2}{k}\Bigl(\frac89\Bigr)^k \sum_j \binom{k+j-3}{j} z^{j+k}\,;$$

далее следует посмотреть на коэффициент при z^{3+l}. (б) $H(z) = \frac23 + \frac{5}{27}z + \frac12\sum_{l\geqslant 0} c_{3+l}z^{2+l}$. (в) Пусть $r = \sqrt{(1-z)(9-z)}$. Можно показать, что $(z-3+r)(z-3-r) = 4z$, а следовательно, $\bigl(r/(1-z)+2\bigr)^2 = (13-5z+4r)/(1-z) = \bigl(9-H(z)\bigr)/\bigl(1-H(z)\bigr)$. (г) Вычисление первой производной при $z=1$ показывает, что $\mathrm{Mean}(H) = 1$. Вторая производная при $z=1$ расходится, так что дисперсия бесконечна.

8.51 (а) Пусть $H_n(z)$ — производящая функция случайной величины ваших денег после n кругов игры; $H_0(z) = z$. Для n кругов имеется распределение

$$H_{n+1}(z) = H_n\bigl(H(z)\bigr)\,,$$

так что требуемый результат получается индукцией (с использованием удивительного тождества из предыдущей задачи). (б) $g_n = H_n(0) - H_{n-1}(0) = 4/n(n+1)(n+2) = 4(n-1)^{\underline{-3}}$. Математическое ожидание равно 2, а дисперсия бесконечна. (в) Среднее количество билетов, которые будут куплены на n-м круге, равно

$\mathrm{Mean}(H_n) = 1$ согласно упр. 15. Поэтому общее количество билетов бесконечно. (Таким образом, в конце концов вы почти наверняка проиграете, и проигрыша можно ждать уже после второй игры, но в то же время приобретете, как можно ожидать, бесконечно много билетов.) (г) В этом случае производящая функция случайной величины после n игр равна $H_n(z)^2$, и метод из п. (б) дает математическое ожидание $16-\frac{4}{3}\pi^2 \approx 2.8$. (Здесь появляется сумма $\sum_{k\geqslant 1} 1/k^2 = \pi^2/6$.)

8.52 Если ω и ω' такие события, что $\Pr(\omega) > \Pr(\omega')$, то в последовательности n независимых экспериментов ω будет встречаться чаще ω', поскольку количество появлений ω будет очень близко к $n\Pr(\omega)$. Следовательно, при $n \to \infty$ со стремящейся к 1 вероятностью медиана и мода значений X в последовательности независимых испытаний будут медианой и модой случайной величины X.

8.53 Это утверждение можно опровергнуть даже в том частном случае, когда каждая случайная величина принимает всего два значения — 0 и 1. Пусть $p_0 = \Pr(X=Y=Z=0)$, $p_1 = \Pr(X=Y=\overline{Z}=0)$, ..., $p_7 = \Pr(\overline{X}=\overline{Y}=\overline{Z}=0)$, где $\overline{X} = 1-X$. Тогда $p_0+p_1+\cdots+p_7 = 1$, и случайные величины попарно независимы тогда и только тогда, когда

$$\begin{aligned}(p_4+p_5+p_6+p_7)(p_2+p_3+p_6+p_7) &= p_6+p_7\,,\\ (p_4+p_5+p_6+p_7)(p_1+p_3+p_5+p_7) &= p_5+p_7\,,\\ (p_2+p_3+p_6+p_7)(p_1+p_3+p_5+p_7) &= p_3+p_7\,.\end{aligned}$$

Но $\Pr(X+Y=Z=0) \ne \Pr(X+Y=0)\Pr(Z=0) \iff p_0 \ne (p_0+p_1)(p_0+p_2+p_4+p_6)$. Одним из решений является

$$p_0 = p_3 = p_5 = p_6 = 1/4\,; \qquad p_1 = p_2 = p_4 = p_7 = 0\,.$$

Эквивалентное определение этих случайных величин, например, таково: бросить две правильные монеты и положить $X =$ (первая монета упала орлом вверх), $Y =$ (вторая монета упала орлом вверх), $Z =$ (монеты упали по-разному). Другой пример, в котором все вероятности ненулевые, может быть таким:

$$\begin{aligned}p_0 &= 4/64\,, \quad p_1 = p_2 = p_4 = 5/64\,,\\ p_3 &= p_5 = p_6 = 10/64\,, \quad p_7 = 15/64\,.\end{aligned}$$

По этой причине мы говорим, что n переменных X_1, ..., X_n независимы, если

$$\Pr(X_1 = x_1 \text{ и } \cdots \text{ и } X_n = x_n) = \Pr(X_1 = x_1)\ldots\Pr(X_n = x_n)\,;$$

для выполнения этого условия попарной независимости недостаточно.

8.54 (См. обозначения в упр. 27.) Имеем

$$\begin{aligned} E(\Sigma_2^2) &= n\mu_4+n(n-1)\mu_2^2\,; \\ E(\Sigma_2\Sigma_1^2) &= n\mu_4+2n(n-1)\mu_3\mu_1+n(n-1)\mu_2^2+n(n-1)(n-2)\mu_2\mu_1^2\,; \\ E(\Sigma_1^4) &= n\mu_4+4n(n-1)\mu_3\mu_1+3n(n-1)\mu_2^2+ \\ &\quad +6n(n-1)(n-2)\mu_2\mu_1^2+n(n-1)(n-2)(n-3)\mu_1^4\,; \end{aligned}$$

отсюда следует, что $V(\widehat{V}X) = \kappa_4/n + 2\kappa_2^2/(n-1)$.

8.55 Всего имеется $A = \frac{1}{17}\cdot 52!$ перестановок с $X = Y$ и $B = \frac{16}{17}\cdot 52!$ перестановок с $X \ne Y$. После описанной процедуры любая перестановка с $X = Y$ будет встречаться с вероятностью $\frac{1}{17}/\bigl((1-\frac{16}{17}p)A\bigr)$, потому что мы возвращаемся к шагу S1 с вероятностью $\frac{16}{17}p$. Аналогично любая перестановка с $X \ne Y$ будет встречаться с вероятностью $\frac{16}{17}(1-p)/\bigl((1-\frac{16}{17}p)B\bigr)$. Выбор $p = \frac{1}{4}$ делает $\Pr(X{=}x \text{ и } Y{=}y) = \frac{1}{169}$ для всех x и y. (Таким образом, можно дважды бросить правильную монету и вернуться к шагу S1, если оба раза выпадет орел.)

8.56 Если m четное, то диски всегда находятся на нечетном расстоянии друг от друга и игра продолжается до бесконечности. Если $m = 2l + 1$, то соответствующими производящими функциями являются

$$\begin{aligned} G_m &= \tfrac{1}{4}zA_1\,; \\ A_1 &= \tfrac{1}{2}zA_1 + \tfrac{1}{4}zA_2\,, \\ A_k &= \tfrac{1}{4}zA_{k-1} + \tfrac{1}{2}zA_k + \tfrac{1}{4}zA_{k+1} \qquad \text{для } 1 < k < l, \\ A_l &= \tfrac{1}{4}zA_{l-1} + \tfrac{3}{4}zA_l + 1\,. \end{aligned}$$

(Коэффициент $[z^n]\,A_k$ представляет собой вероятность того, что диски после n бросков окажутся на расстоянии $2k$.) Вспоминая подобные уравнения из упр. 49, положим $z = 1/\cos^2\theta$ и $A_1 = X\sin 2\theta$, где X требуется определить. По индукции получаем (не используя уравнение для A_l), что $A_k = X\sin 2k\theta$. Следовательно, нам надо выбрать X так, чтобы

$$\Bigl(1 - \frac{3}{4\cos^2\theta}\Bigr) X \sin 2l\theta \;=\; 1 + \frac{1}{4\cos^2\theta} X \sin(2l-2)\theta\,.$$

Оказывается, что $X = 2\cos^2\theta/\sin\theta\cos(2l+1)\theta$, следовательно,

$$G_m \;=\; \frac{\cos\theta}{\cos m\theta}\,.$$

Знаменатель обращается в нуль, когда θ является нечетным кратным $\pi/(2m)$; таким образом, $1 - q_k z$ — корень знаменателя для $1 \leqslant k \leqslant l$, и указанное в упражнении представление в виде произведения обязано иметь место. Чтобы найти математическое ожидание и дисперсию, можно записать

Опять тригонометрия! Может, дело в том, что броски выполняются вдоль сторон m-угольника?

$$\begin{aligned} G_m &= (1 - \tfrac{1}{2}\theta^2 + \tfrac{1}{24}\theta^4 - \cdots)/(1 - \tfrac{1}{2}m^2\theta^2 + \tfrac{1}{24}m^4\theta^4 - \cdots) = \\ &= 1 + \tfrac{1}{2}(m^2-1)\theta^2 + \tfrac{1}{24}(5m^4 - 6m^2 + 1)\theta^4 + \cdots = \\ &= 1 + \tfrac{1}{2}(m^2-1)(\operatorname{tg}\theta)^2 + \tfrac{1}{24}(5m^4 - 14m^2 + 9)(\operatorname{tg}\theta)^4 + \cdots = \\ &= 1 + G_m'(1)(\operatorname{tg}\theta)^2 + \tfrac{1}{2}G_m''(1)(\operatorname{tg}\theta)^4 + \cdots, \end{aligned}$$

потому что $\operatorname{tg}^2\theta = z - 1$ и $\operatorname{tg}\theta = \theta + \frac{1}{3}\theta^3 + \cdots$. Поэтому имеем $\operatorname{Mean}(G_m) = \frac{1}{2}(m^2-1)$ и $\operatorname{Var}(G_m) = \frac{1}{6}m^2(m^2-1)$. (Обратите внимание, что отсюда вытекают тождества

$$\frac{m^2-1}{2} = \sum_{k=1}^{(m-1)/2} \frac{1}{p_k} = \sum_{k=1}^{(m-1)/2} \Bigl(1\Big/\sin\frac{(2k-1)\pi}{2m}\Bigr)^2;$$

$$\frac{m^2(m^2-1)}{6} = \sum_{k=1}^{(m-1)/2} \Bigl(\operatorname{ctg}\frac{(2k-1)\pi}{2m}\Big/\sin\frac{(2k-1)\pi}{2m}\Bigr)^2.$$

Третий кумулянт этого распределения равен $\frac{1}{30}m^2(m^2-1)(4m^2-1)$; но на этом кумулянты с хорошим разложением заканчиваются. Математическое ожидание можно найти и существенно проще. Действительно, $G_m + A_1 + \cdots + A_l = z(A_1 + \cdots + A_l) + 1$, следовательно, для $z = 1$ мы имеем $G_m' = A_1 + \cdots + A_l$. Поскольку $G_m = 1$ при $z = 1$, простая индукция показывает, что $A_k = 4k$.)

8.57 Имеем $A{:}A \geqslant 2^{l-1}$, $B{:}B < 2^{l-1} + 2^{l-3}$ и $B{:}A \geqslant 2^{l-2}$, следовательно, $B{:}B - B{:}A \geqslant A{:}A - A{:}B$ возможно только при $A{:}B > 2^{l-3}$. Это означает, что $\overline{\tau}_2 = \tau_3$, $\tau_1 = \tau_4$, $\tau_2 = \tau_5$, …, $\tau_{l-3} = \tau_l$. Но тогда $A{:}A \approx 2^{l-1} + 2^{l-4} + \cdots$, $A{:}B \approx 2^{l-3} + 2^{l-6} + \cdots$, $B{:}A \approx 2^{l-2} + 2^{l-5} + \cdots$ и $B{:}B \approx 2^{l-1} + 2^{l-4} + \cdots$; следовательно, $B{:}B - B{:}A$ в конечном счете меньше, чем $A{:}A - A{:}B$. (Более сильные результаты были получены Гуибасом (Guibas) и Одлыжко (Odlyzko) [168], которые показали, что шансы Билла всегда максимизируются одной из последовательностей $0\tau_1 \ldots \tau_{l-1}$ и $P\tau_1 \ldots \tau_{l-1}$. На самом деле выигрышная стратегия Билла единственна — см. следующее упражнение.)

8.58 (Решение Я. Цирика (J. Csirik).) Если A есть 0^l или P^l, то одна из двух сравниваемых последовательностей совпадает с A

и, следовательно, не может быть использована. В противном случае положим $\hat{A} = \tau_1 \ldots \tau_{l-1}$, $O = 0\hat{A}$ и $P = P\hat{A}$. Нетрудно проверить, что $O{:}A = P{:}A = \hat{A}{:}\hat{A}$, $O{:}O + P{:}P = 2^{l-1} + 2(\hat{A}{:}\hat{A}) + 1$ и $A{:}O + A{:}P = 1 + 2(A{:}A) - 2^l$. Поэтому из уравнения

$$\frac{O{:}O - O{:}A}{A{:}A - A{:}O} = \frac{P{:}P - P{:}A}{A{:}A - A{:}P}$$

следует, что обе дроби равны

$$\frac{O{:}O - O{:}A + P{:}P - P{:}A}{A{:}A - A{:}O + A{:}A - A{:}P} = \frac{2^{l-1}+1}{2^l - 1}.$$

Далее можно записать дроби по-другому и показать, что

$$\frac{O{:}O - O{:}A}{P{:}P - P{:}A} = \frac{A{:}A - A{:}O}{A{:}A - A{:}P} = \frac{p}{q},$$

где $pq > 0$ и $p + q = \text{НОД}(2^{l-1}+1, 2^l - 1) = \text{НОД}(3, 2^l - 1)$; так что мы можем считать, что l четное и что $p = 1$, $q = 2$. Отсюда следует $A{:}A - A{:}O = (2^l - 1)/3$ и $A{:}A - A{:}P = (2^{l+1} - 2)/3$, значит, $A{:}O - A{:}P = (2^l - 1)/3 \geqslant 2^{l-2}$. Имеем $A{:}O \geqslant 2^{l-2}$ тогда и только тогда, когда $A = (PO)^{l/2}$. Но в этом случае $O{:}O - O{:}A = A{:}A - A{:}O$, так что $2^{l-1} + 1 = 2^l - 1$ и $l = 2$.

(Цирик [69] пошел дальше и показал, что для $l \geqslant 4$ у Алисы нет лучшего выбора, чем играть $OP^{l-3}O^2$. Но даже при этой стратегии Билл выигрывает с вероятностью около $\frac{2}{3}$.)

8.59 Согласно (8.82) нам надо, чтобы $B{:}B - B{:}A > A{:}A - A{:}B$. Одним из решений является $A = PPOO$, $B = OOO$.

8.60 (а) Возможны два случая: $h_k \neq h_n$ и $h_k = h_n$:

$$G(w,z) = \frac{m-1}{m}\Bigl(\frac{m-2+w+z}{m}\Bigr)^{k-1} w \Bigl(\frac{m-1+z}{m}\Bigr)^{n-k-1} z + {} + \frac{1}{m}\Bigl(\frac{m-1+wz}{m}\Bigr)^{k-1} wz \Bigl(\frac{m-1+z}{m}\Bigr)^{n-k-1} z.$$

(б) Можно обосновать это соотношение алгебраически, вычислив частные производные $G(w,z)$ по w и z и положив $w = z = 1$; возможно также комбинаторное обоснование. Какими бы ни были значения $h_1, \ldots, h_{n-1}$, ожидаемое значение $P(h_1, \ldots, h_{n-1}, h_n; n)$ (усреднение по h_n) постоянно, потому что хеш-последовательность $(h_1, \ldots, h_{n-1})$ определяет последовательность длин списков $(n_1, n_2, \ldots, n_m)$ и указанное среднее значение равно $\bigl((n_1 + 1) + (n_2 + 1) + \cdots + (n_m + 1)\bigr)/m = (n - 1 + m)/m$. Таким образом, случайная величина $EP(h_1, \ldots, h_n; n)$ не зависит от $(h_1, \ldots, h_{n-1})$, а следовательно, не зависит от $P(h_1, \ldots, h_n; k)$.

8.61 Если $1 \leqslant k < l \leqslant n$, то, как следует из предыдущего упражнения, коэффициент при $s_k s_l$ в дисперсии среднего равен нулю. Поэтому остается только рассмотреть коэффициент при s_k^2, который равен

$$\sum_{1\leqslant h_1,\ldots,h_n\leqslant m} \frac{P(h_1,\ldots,h_n;k)^2}{m^n} - \Biggl(\sum_{1\leqslant h_1,\ldots,h_n\leqslant m} \frac{P(h_1,\ldots,h_n;k)}{m^n}\Biggr)^2,$$

т.е. дисперсия функции $\bigl((m-1+z)/m\bigr)^{k-1}z$; это, в свою очередь, равно $(k-1)(m-1)/m^2$, как в упр. 30.

8.62 Производящая функция случайной величины $D_n(z)$ удовлетворяет рекуррентному соотношению

$$\begin{aligned} D_0(z) &= z; \\ D_n(z) &= z^2 D_{n-1}(z) + 2(1-z^3)D'_{n-1}(z)/(n+1) \qquad \text{для } n > 0. \end{aligned}$$

Отсюда можно вывести рекуррентное соотношение

$$D''_n(1) = (n-11)D''_{n-1}(1)/(n+1) + (8n-2)/7,$$

которое имеет решение $\frac{2}{637}(n+2)(26n+15)$ для всех $n \geqslant 11$ (независимо от начальных условий). Следовательно, дисперсия оказывается равной $\frac{108}{637}(n+2)$ для $n \geqslant 11$.

8.63 (Можно поставить другой вопрос: получается ли предполагаемая последовательность кумулянтов хоть из какого-нибудь распределения? Так, в последовательности кумулянтов κ_2 должно быть неотрицательно, $\kappa_4 + 3\kappa_2^2 = E\bigl((X-\mu)^4\bigr)$ должно быть не менее $\bigl(E\bigl((X-\mu)^2\bigr)\bigr)^2 = \kappa_2^2$, и т.д. Необходимые и достаточные условия для этого варианта задачи были найдены Гамбургером (Hamburger) [6, 175].)

9.1 Это верно, если все функции положительны. В противном случае можно, например, взять $f_1(n) = n^3 + n^2$, $f_2(n) = -n^3$, $g_1(n) = n^4 + n$, $g_2(n) = -n^4$.

9.2 (а) Поскольку $(\ln n)^2 \prec n \ln c \prec n \ln\ln n$, получаем $n^{\ln n} \prec c^n \prec (\ln n)^n$. (б) $n^{\ln\ln\ln n} \prec (\ln n)! \prec n^{\ln\ln n}$. (в) Прологарифмируйте, чтобы показать, что $(n!)!$ растет быстрее. (г) $F^2_{\lceil H_n\rceil} \asymp \phi^{2\ln n} = n^{2\ln\phi}$; $H_{F_n} \sim n\ln\phi$ растет быстрее, потому что $\phi^2 = \phi + 1 < e$.

9.3 Замена kn на $O(n)$ подразумевает различные C для различных k; нужно же, чтобы все O имели общую константу C. В действительности в данном контексте требуется, чтобы O обозначало множество функций двух переменных, k и n. Правильно будет записать $\sum_{k=1}^n kn = \sum_{k=1}^n O(n^2) = O(n^3)$.

9.4 Например, $\lim_{n\to\infty} O(1/n) = 0$. В левой части $O(1/n)$ есть множество всех функций $f(n)$, для которых найдутся константы C и n_0, такие, что $|f(n)| \leqslant C/n$ для всех $n \geqslant n_0$. Предел любой функции из этого множества равен нулю, поэтому левая часть представляет собой одноэлементное множество $\{0\}$. В правой части нет переменных; 0 представляет $\{0\}$, множество (из одного элемента) всех "функций без переменных, имеющих нулевое значение". (Понимаете ли вы логику этого высказывания? Если нет, вернитесь к этой задаче через год; вы, вероятно, сможете обращаться с O-обозначениями, даже если ваша интуиция не втискивается в рамки строгого формализма.)

9.5 Пусть $f(n) = n^2$ и $g(n) = 1$; тогда n принадлежит левому множеству, но не принадлежит правому, так что утверждение ложно.

9.6 $n \ln n + \gamma n + O(\sqrt{n} \ln n)$.

9.7 $(1-e^{-1/n})^{-1} = nB_0 - B_1 + B_2 n^{-1}/2! + \cdots = n + \frac{1}{2} + O(n^{-1})$.

9.8 Пусть, например, $f(n) = \lfloor n/2 \rfloor!^2 + n$, $g(n) = (\lceil n/2 \rceil - 1)!\,\lceil n/2 \rceil! + n$. Эти функции, как оказывается, удовлетворяют соотношениям $f(n) = O\bigl(ng(n)\bigr)$ и $g(n) = O\bigl(nf(n)\bigr)$; разумеется, возможны примеры и с более резким различием.

9.9 (Для полноты предположим существование граничных условий $n \to \infty$, так что каждое O подразумевает две константы.) Каждая функция из левой части имеет вид $a(n) + b(n)$, причем существуют константы m_0, B, n_0, C, такие, что $|a(n)| \leqslant B|f(n)|$ для $n \geqslant m_0$ и $|b(n)| \leqslant C|g(n)|$ для $n \geqslant n_0$. Следовательно, функция в левой части не превосходит $\max(B, C)\bigl(|f(n)| + |g(n)|\bigr)$ для $n \geqslant \max(m_0, n_0)$, так что она принадлежит правой части.

9.10 Если $g(x)$ принадлежит левой части, так что $g(x) = \cos y$ для некоторого y, причем $|y| \leqslant C|x|$ для некоторого C, то $0 \leqslant 1 - g(x) = 2\sin^2(y/2) \leqslant \frac{1}{2}y^2 \leqslant \frac{1}{2}C^2x^2$. Следовательно, множество из левой части содержится в правой, и формула верна.

9.11 Утверждение верно. Действительно, если, скажем, $|x| \leqslant |y|$, то $(x+y)^2 \leqslant 4y^2$. Таким образом, $(x+y)^2 = O(x^2) + O(y^2)$. Следовательно, $O(x+y)^2 = O\bigl((x+y)^2\bigr) = O\bigl(O(x^2) + O(y^2)\bigr) = O\bigl(O(x^2)\bigr) + O\bigl(O(y^2)\bigr) = O(x^2) + O(y^2)$.

9.12 $1+2/n+O(n^{-2}) = (1+2/n)\bigl(1+O(n^{-2})/(1+2/n)\bigr)$ согласно (9.26), кроме того, $1/(1+2/n) = O(1)$; теперь используйте (9.26).

9.13 $n^n\bigl(1 + 2n^{-1} + O(n^{-2})\bigr)^n = n^n \exp\bigl(n\bigl(2n^{-1} + O(n^{-2})\bigr)\bigr) = e^2 n^n + O(n^{n-1})$.

9.14 Это $n^{n+\beta}\exp\bigl((n+\beta)(\alpha/n-\frac{1}{2}\alpha^2/n^2+O(n^{-3}))\bigr)$.

9.15 $\ln\binom{3n}{n,n,n} = 3n\ln 3-\ln n+\frac{1}{2}\ln 3-\ln 2\pi+\bigl(\frac{1}{36}-\frac{1}{4}\bigr)n^{-1}+O(n^{-3})$, так что ответом будет

$$\frac{3^{3n+1/2}}{2\pi n}\bigl(1-\tfrac{2}{9}n^{-1}+\tfrac{2}{81}n^{-2}+O(n^{-3})\bigr)\,.$$

Интересно сравнить эту формулу с аналогичным результатом для среднего биномиального коэффициента, упр. 9.60.

9.16 Если l — любое целое число из диапазона $a\leqslant l<b$, то

$$\int_0^1 B(x)f(l+x)\,dx = \int_{1/2}^1 B(x)f(l+x)\,dx-\int_0^{1/2} B(1-x)f(l+x)\,dx = \int_{1/2}^1 B(x)\bigl(f(l+x)-f(l+1-x)\bigr)\,dx\,.$$

Поскольку $l+x\geqslant l+1-x$ при $x\geqslant\frac{1}{2}$, этот интеграл положителен, если $f(x)$ — неубывающая функция.

9.17 $\sum_{m\geqslant 0}B_m(\frac{1}{2})z^m/m! = ze^{z/2}/(e^z-1) = z/(e^{z/2}-1)-z/(e^z-1)$.

9.18 Вывод в тексте для $\alpha=1$ можно обобщить, получив

$$b_k(n)=\frac{2^{(2n+1/2)\alpha}}{(2\pi n)^{\alpha/2}}e^{-k^2\alpha/n}\,,\quad c_k(n)=2^{2n\alpha}n^{-(1+\alpha)/2+3\epsilon}e^{-k^2\alpha/n}\,;$$

ответом будет $2^{2n\alpha}(\pi n)^{(1-\alpha)/2}\alpha^{-1/2}\bigl(1+O(n^{-1/2+3\epsilon})\bigr)$.

9.19 $H_{10}=2.928968254\approx 2.928968256$; $10!=3628800\approx 3628712.4$; $B_{10}=0.075757576\approx 0.075757494$; $\pi(10)=4\approx 10.0017845$; $e^{0.1}=1.10517092\approx 1.10517083$; $\ln 1.1=0.0953102\approx 0.0953083$; $1.1111111\ldots\approx 1.1111$; $1.1^{0.1}=1.00957658\approx 1.00957643$. (Аппроксимация $\pi(n)$ дает больше верных цифр при бо́льших n, например $\pi(10^9)=50\,847\,534\approx 50\,840\,742$.)

9.20 (а) Верно; левая часть есть $o(n)$, тогда как правая часть эквивалентна $O(n)$. (б) Верно; левая часть есть $e\cdot e^{O(1/n)}$. (в) Неверно; левая часть примерно в $\sqrt{n}$ раз превосходит верхнюю границу для правой части.

9.21 Имеем $P_n=p=n\bigl(\ln p-1-1/\ln p+O(1/\log n)^2\bigr)$, при этом

$$\ln p = \ln n+\ln\ln p-1/\ln n+\ln\ln n/(\ln n)^2+O(1/\log n)^2\,;$$
$$\ln\ln p = \ln\ln n+\frac{\ln\ln n}{\ln n}-\frac{(\ln\ln n)^2}{2(\ln n)^2}+\frac{\ln\ln n}{(\ln n)^2}+O(1/\log n)^2\,.$$

Отсюда следует, что

$$P_n = n\Bigl(\ln n + \ln\ln n - 1 + \\ + \frac{\ln\ln n - 2}{\ln n} - \frac{(\ln\ln n)^2/2 - 3\ln\ln n}{(\ln n)^2} + O(1/\log n)^2\Bigr).$$

Что говорит, утопая, специалист по аналитической теории чисел?

log log log log ...

(Можно несколько улучшить аппроксимацию, заменив здесь $O(1/\log n)^2$ величиной $-5.5/(\ln n)^2 + O(\log\log n/\log n)^3$; в этом случае мы получаем $P_{1000000} \approx 15\,480\,992.8$.)

9.22 Заменим в разложении H_{n^k} член $O(n^{-2k})$ на $-\frac{1}{12}n^{-2k} + O(n^{-4k})$; в результате $O\bigl(\Sigma_3(n^2)\bigr)$ в (9.53) перейдет в $-\frac{1}{12}\Sigma_3(n^2) + O\bigl(\Sigma_3(n^4)\bigr)$. Имеем

$$\Sigma_3(n) = \tfrac{3}{4}n^{-1} + \tfrac{5}{36}n^{-2} + O(n^{-3}),$$

следовательно, член $O(n^{-2})$ в (9.54) может быть заменен на $-\frac{19}{144}n^{-2} + O(n^{-3})$.

9.23 $nh_n = \sum_{0\leqslant k<n} h_k/(n-k) + 2cH_n/(n+1)(n+2)$. Выберем $c = e^{\pi^2/6} = \sum_{k\geqslant 0} g_k$, так что $\sum_{k\geqslant 0} h_k = 0$ и $h_n = O(\log n)/n^3$. Развертывание $\sum_{0\leqslant k<n} h_k/(n-k)$, как в (9.60), дает теперь $nh_n = 2cH_n/(n+1)(n+2) + O(n^{-2})$, следовательно,

$$g_n = e^{\pi^2/6}\left(\frac{n + 2\ln n + O(1)}{n^3}\right).$$

9.24 (a) Если $\sum_{k\geqslant 0}|f(k)| < \infty$ и $f(n-k) = O\bigl(f(n)\bigr)$ для $0 \leqslant k \leqslant n/2$, то

$$\sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{n/2} O\bigl(f(k)\bigr)O\bigl(f(n)\bigr) + \sum_{k=n/2}^{n} O\bigl(f(n)\bigr)O\bigl(f(n-k)\bigr),$$

что составляет $2O\bigl(f(n)\sum_{k\geqslant 0}|f(k)|\bigr)$, так что в этом случае все доказано. (б) В случае $a_n = b_n = \alpha^{-n}$ свертка $(n+1)\alpha^{-n}$ не есть $O(\alpha^{-n})$.

9.25 $S_n/\binom{3n}{n} = \sum_{k=0}^{n} n^{\underline{k}}/(2n+1)^{\overline{k}}$. Диапазон суммирования можно сузить, скажем, до $0 \leqslant k \leqslant (\log n)^2$. В этом диапазоне $n^{\underline{k}} = n^k\bigl(1 - \binom{k}{2}/n + O(k^4/n^2)\bigr)$ и $(2n+1)^{\overline{k}} = (2n)^k\bigl(1 + \binom{k+1}{2}/2n + O(k^4/n^2)\bigr)$, так что слагаемое равно

$$\frac{1}{2^k}\left(1 - \frac{3k^2 - k}{4n} + O\Bigl(\frac{k^4}{n^2}\Bigr)\right).$$

Следовательно, суммирование по k дает $2 - 4/n + O(1/n^2)$. Теперь, чтобы доказать (9.2), можно применить аппроксимацию Стирлинга к $\binom{3n}{n} = (3n)!/(2n)!\,n!$.

9.26 Минимум достигается на члене $B_{2m}/(2m)(2m-1)n^{2m-1}$, когда $2m \approx 2\pi n + \frac{3}{2}$, и этот член приближенно равен $1/(\pi e^{2\pi n}\sqrt{n}\,)$. Следовательно, абсолютная погрешность $\ln n!$ слишком велика для точного определения значения $n!$ путем округления до целого, если n больше, чем примерно $e^{2\pi+1}$.

9.27 Можно считать, что $\alpha \neq -1$. Положим $f(x) = x^\alpha$; ответом будет

$$\sum_{k=1}^{n} k^\alpha = C_\alpha + \frac{n^{\alpha+1}}{\alpha+1} + \frac{n^\alpha}{2} + \sum_{k=1}^{m} \frac{B_{2k}}{2k}\binom{\alpha}{2k-1} n^{\alpha-2k+1} + O(n^{\alpha-2m-1})\,.$$

(Оказывается, константа C_α равна $\zeta(-\alpha)$, причем на самом деле эта формула служит *определением* $\zeta(-\alpha)$ при $\alpha > -1$.)

В частности, $\zeta(0) = -1/2$ *и* $\zeta(-n) = -B_{n+1}/(n+1)$ *для целого* $n > 0$.

9.28 В общем случае положим в формуле суммирования Эйлера $f(x) = x^\alpha \ln x$, где $\alpha \neq -1$. Действуя, как и в предыдущем упражнении, найдем

$$\sum_{k=1}^{n} k^\alpha \ln k = C'_\alpha + \frac{n^{\alpha+1}\ln n}{\alpha+1} - \frac{n^{\alpha+1}}{(\alpha+1)^2} + \frac{n^\alpha \ln n}{2} + \sum_{k=1}^{m} \frac{B_{2k}}{2k}\binom{\alpha}{2k-1} n^{\alpha-2k+1}(\ln n + H_\alpha - H_{\alpha-2k+1}) + O(n^{\alpha-2m-1}\log n)\,;$$

как можно показать [74, §3.7], константа C'_α равна $-\zeta'(-\alpha)$. (Можно исключить множитель $\log n$ из O-слагаемого, когда α — положительное целое число $\leqslant 2m$; в этом случае мы также заменяем k-й член суммы в правой части на $B_{2k}\alpha!\,(2k-2-\alpha)! \times (-1)^\alpha n^{\alpha-2k+1}/(2k)!$ при $\alpha < 2k-1$.) Для решения поставленной задачи положим $\alpha = 1$ и $m = 1$ и возьмем экспоненту от обеих частей, что даст нам

$$Q_n = A \cdot n^{n^2/2+n/2+1/12} e^{-n^2/4}\bigl(1 + O(n^{-2})\bigr)\,,$$

где $A = e^{1/12-\zeta'(-1)} \approx 1.2824271291$ — "константа Глейшера".

9.29 Положим $f(x) = x^{-1}\ln x$. Небольшая модификация вычислений из предыдущего упражнения дает

$$\sum_{k=1}^{n}\frac{\ln k}{k} = \frac{(\ln n)^2}{2} + \gamma_1 + \frac{\ln n}{2n} -$$
$$- \sum_{k=1}^{m}\frac{B_{2k}}{2k}n^{-2k}(\ln n - H_{2k-1}) + O(n^{-2m-1}\log n),$$

где $\gamma_1 \approx -0.07281584548367672486$ — "константа Стилтьеса" (см. ответ к упр. 9.57). Взятие экспоненты дает

$$e^{\gamma_1}\sqrt{n^{\ln n}}\left(1 + \frac{\ln n}{2n} + O\Bigl(\frac{\log n}{n}\Bigr)^2\right).$$

9.30 Положим $g(x) = x^l e^{-x^2}$ и $f(x) = g(x/\sqrt{n})$. Тогда $n^{-l/2} \times \sum_{k\geqslant 0} k^l e^{-k^2/n}$ есть

$$\int_0^\infty f(x)\,dx - \sum_{k=1}^{m}\frac{B_k}{k!}f^{(k-1)}(0) - (-1)^m\int_0^\infty \frac{B_m(\{x\})}{m!}f^{(m)}(x)\,dx =$$
$$= n^{1/2}\int_0^\infty g(x)\,dx - \sum_{k=1}^{m}\frac{B_k}{k!}n^{(k-1)/2}g^{(k-1)}(0) + O(n^{-m/2}).$$

Поскольку $g(x) = x^l - x^{2+l}/1! + x^{4+l}/2! - x^{6+l}/3! + \cdots$, производные $g^{(m)}(x)$ следуют простой схеме, и ответом является

$$\frac{1}{2}n^{(l+1)/2}\Gamma\Bigl(\frac{l+1}{2}\Bigr) - \frac{B_{l+1}}{(l+1)!\,0!} + \frac{B_{l+3}n^{-1}}{(l+3)!\,1!} - \frac{B_{l+5}n^{-2}}{(l+5)!\,2!} + O(n^{-3}).$$

9.31 Несколько неожиданное тождество $1/(c^{m-k} + c^m) + 1/(c^{m+k} + c^m) = 1/c^m$ показывает, что сумма членов для $0 \leqslant k \leqslant 2m$ равна $(m + \frac{1}{2})/c^m$. Остальная сумма равна

$$\sum_{k\geqslant 1}\frac{1}{c^{2m+k} + c^m} = \sum_{k\geqslant 1}\left(\frac{1}{c^{2m+k}} - \frac{1}{c^{3m+2k}} + \frac{1}{c^{4m+3k}} - \cdots\right) =$$
$$= \frac{1}{c^{2m+1} - c^{2m}} - \frac{1}{c^{3m+2} - c^{3m}} + \cdots,$$

и этот ряд можно оборвать в любом месте, получив погрешность, не превышающую первого отброшенного члена.

9.32 $H_n^{(2)} = \pi^2/6 - 1/n + O(n^{-2})$ по формуле суммирования Эйлера, поскольку мы знаем константу; H_n задается формулой (9.89). Так что ответ таков:

В этом ответе встречаются три мировые константы: e, π, γ.

$$ne^{\gamma+\pi^2/6}\bigl(1 - \tfrac{1}{2}n^{-1} + O(n^{-2})\bigr).$$

9.33 Имеем $n^{\underline{k}}/n^{\overline{k}} = 1-k(k-1)n^{-1}+\frac{1}{2}k^2(k-1)^2n^{-2}+O(k^6n^{-3})$; деление на $k!$ и суммирование по $k \geqslant 0$ дает $e - en^{-1} + \frac{7}{2}en^{-2} + O(n^{-3})$.

9.34 $A = e^\gamma$; $B = 0$; $C = -\frac{1}{2}e^\gamma$; $D = \frac{1}{2}e^\gamma(1-\gamma)$; $E = \frac{1}{8}e^\gamma$; $F = \frac{1}{12}e^\gamma(3\gamma+1)$.

9.35 Поскольку $1/k\bigl(\ln k+O(1)\bigr) = 1/k\ln k+O\bigl(1/k(\log k)^2\bigr)$, данная сумма равна $\sum_{k=2}^{n} 1/k\ln k + O(1)$. Оставшаяся сумма, по формуле суммирования Эйлера, равна $\ln\ln n + O(1)$.

9.36 Здесь замечательно работает формула суммирования Эйлера:

$$S_n = \sum_{0\leqslant k<n} \frac{1}{n^2+k^2} + \frac{1}{n^2+x^2}\bigg|_0^n =$$
$$= \int_0^n \frac{dx}{n^2+x^2} + \frac{1}{2}\frac{1}{n^2+x^2}\bigg|_0^n + \frac{B_2}{2!}\frac{-2x}{(n^2+x^2)^2}\bigg|_0^n + O(n^{-5}).$$

Следовательно, $S_n = \frac{1}{4}\pi n^{-1} - \frac{1}{4}n^{-2} - \frac{1}{24}n^{-3} + O(n^{-5})$.

9.37 Эта сумма равна

$$\sum_{k,q\geqslant 1} (n-qk)\bigl[n/(q+1) < k \leqslant n/q\bigr] =$$
$$= n^2 - \sum_{q\geqslant 1} q\left(\binom{\lfloor n/q\rfloor+1}{2} - \binom{\lfloor n/(q+1)\rfloor+1}{2}\right) =$$
$$= n^2 - \sum_{q\geqslant 1}\binom{\lfloor n/q\rfloor+1}{2}.$$

Оставшаяся сумма подобна (9.55), но без множителя $\mu(q)$. Использованный нами метод работает и здесь, но вместо $1/\zeta(2)$ мы получим $\zeta(2)$, так что ответом является $\bigl(1-\frac{\pi^2}{12}\bigr)n^2 + O(n\log n)$.

9.38 Заменим k на $n-k$ и положим $a_k(n) = (n-k)^{n-k}\binom{n}{k}$. Тогда $\ln a_k(n) = n\ln n - \ln k! - k + O(kn^{-1})$, и можно воспользоваться заменой хвоста с $b_k(n) = n^n e^{-k}/k!$, $c_k(n) = kb_k(n)/n$, $D_n = \{k \mid k \leqslant \ln n\}$ и получить $\sum_{k=0}^{n} a_k(n) = n^n e^{1/e}\bigl(1+O(n^{-1})\bigr)$.

9.39 Замена хвоста с $b_k(n) = (\ln n - k/n - \frac{1}{2}k^2/n^2)(\ln n)^k/k!$, $c_k(n) = n^{-3}(\ln n)^{k+3}/k!$, $D_n = \{k \mid 0 \leqslant k \leqslant 10\ln n\}$. При $k \approx 10\ln n$ имеем $k! \asymp \sqrt{k}\,(10/e)^k(\ln n)^k$, так что k-й член представляет собой $O(n^{-10\ln(10/e)}\log n)$. Окончательный ответ: $n\ln n - \ln n - \frac{1}{2}(\ln n)(1+\ln n)/n + O\bigl(n^{-2}(\log n)^3\bigr)$.

9.40 Объединяя члены попарно, получим $H_{2k}^m - (H_{2k} - \frac{1}{2k})^m = \frac{m}{2k}H_{2k}^{m-1}$ плюс члены, сумма которых по всем $k \geqslant 1$ равна $O(1)$.

Предположим, что n четное. Из формулы суммирования Эйлера

$$\sum_{k=1}^{n/2}\frac{H_{2k}^{m-1}}{k} = \sum_{k=1}^{n/2}\frac{(\ln 2e^{\gamma}k)^{m-1}+O(k^{-1}(\log k)^{m-2})}{k} = \\ = \frac{(\ln e^{\gamma}n)^{m}}{m} + O(1);$$

следовательно, сумма равна $\frac{1}{2}H_n^m + O(1)$. Ответ для общего случая имеет вид $\frac{1}{2}(-1)^n H_n^m + O(1)$.

9.41 Положим $\alpha = \hat{\phi}/\phi = -\phi^{-2}$. Имеем

$$\sum_{k=1}^{n}\ln F_k = \sum_{k=1}^{n}\bigl(\ln\phi^k - \ln\sqrt{5} + \ln(1-\alpha^k)\bigr) = \\ = \frac{n(n+1)}{2}\ln\phi - \frac{n}{2}\ln 5 + \sum_{k\geqslant 1}\ln(1-\alpha^k) - \sum_{k>n}\ln(1-\alpha^k).$$

Последняя сумма есть $\sum_{k>n} O(\alpha^k) = O(\alpha^n)$. Следовательно, ответ —

$$\phi^{n(n+1)/2}5^{-n/2}C + O(\phi^{n(n-3)/2}5^{-n/2}),$$

где $C = (1-\alpha)(1-\alpha^2)(1-\alpha^3)\ldots \approx 1.226742$.

9.42 Утверждение из указания следует из того, что $\binom{n}{k-1}/\binom{n}{k} = \frac{k}{n-k+1} \leqslant \frac{\alpha n}{n-\alpha n+1} < \frac{\alpha}{1-\alpha}$. Пусть $m = \lfloor \alpha n \rfloor = \alpha n - \epsilon$. Тогда

$$\binom{n}{m} < \sum_{k\leqslant m}\binom{n}{k} < \\ < \binom{n}{m}\left(1 + \frac{\alpha}{1-\alpha} + \left(\frac{\alpha}{1-\alpha}\right)^2 + \cdots\right) = \binom{n}{m}\frac{1-\alpha}{1-2\alpha}.$$

Следовательно, $\sum_{k\leqslant \alpha n}\binom{n}{k} = \binom{n}{m}O(1)$, и остается оценить $\binom{n}{m}$. Из приближения Стирлинга $\ln\binom{n}{m} = -\frac{1}{2}\ln n - (\alpha n - \epsilon)\ln(\alpha - \epsilon/n) - \bigl((1-\alpha)n+\epsilon\bigr)\ln(1-\alpha+\epsilon/n) + O(1) = -\frac{1}{2}\ln n - \alpha n\ln\alpha - (1-\alpha)n\ln(1-\alpha) + O(1)$.

9.43 Знаменатель содержит множители вида $z - \omega$, где ω — комплексный корень из единицы. Из них только множитель $z-1$ входит с кратностью 5. Поэтому из (7.31) следует, что только один корень имеет коэффициент $\Omega(n^4)$, и этот коэффициент есть $c = 5/(5!\cdot 1\cdot 5\cdot 10\cdot 25\cdot 50) = 1/1500000$.

9.44 Приближение Стирлинга гласит, что для $\ln\bigl(x^{-\alpha}x!/(x-\alpha)!\bigr)$ имеется асимптотический ряд

$$-\alpha-(x+\tfrac{1}{2}-\alpha)\ln(1-\alpha/x)-\frac{B_2}{2\cdot 1}\bigl(x^{-1}-(x-\alpha)^{-1}\bigr)- \\ -\frac{B_4}{4\cdot 3}\bigl(x^{-3}-(x-\alpha)^{-3}\bigr)-\cdots,$$

в котором каждый коэффициент при x^{-k} представляет собой полином от α. Следовательно, $x^{-\alpha}x!/(x-\alpha)!=c_0(\alpha)+c_1(\alpha)x^{-1}+\cdots+c_n(\alpha)x^{-n}+O(x^{-n-1})$ при $x\to\infty$, где $c_n(\alpha)$ является полиномом от α. Мы знаем, что $c_n(\alpha)=\left[{\alpha\atop \alpha-n}\right](-1)^n$, если α — целое, и $\left[{\alpha\atop \alpha-n}\right]$ есть полином от α степени $2n$; следовательно, $c_n(\alpha)=\left[{\alpha\atop \alpha-n}\right](-1)^n$ для всех действительных α. Другими словами, асимптотические формулы

Дальнейшее обсуждение см. в [220].

$$x^{\underline{\alpha}}=\sum_{k=0}^{n}\left[{\alpha\atop \alpha-k}\right](-1)^k x^{\alpha-k}+O(x^{\alpha-n-1}),$$

$$x^{\overline{\alpha}}=\sum_{k=0}^{n}\left[{\alpha\atop \alpha-k}\right]x^{\alpha-k}+O(x^{\alpha-n-1})$$

обобщают формулы (6.13) и (6.11), имеющие место в целочисленном случае.

9.45 Пусть неполными частными при разложении α в цепную дробь будут $\langle a_1,a_2,\ldots\rangle$. Пусть α_m означает цепную дробь $1/(a_m+\alpha_{m+1})$ для $m\geqslant 1$. Тогда для всех m выполняется соотношение $D(\alpha,n)=D(\alpha_1,n)<D\bigl(\alpha_2,\lfloor\alpha_1 n\rfloor\bigr)+a_1+3<D\bigl(\alpha_3,\lfloor\alpha_2\lfloor\alpha_1 n\rfloor\rfloor\bigr)+a_1+a_2+6<\cdots<D\bigl(\alpha_{m+1},\lfloor\alpha_m\lfloor\ldots\lfloor\alpha_1 n\rfloor\ldots\rfloor\rfloor\bigr)+a_1+\cdots+a_m+3m<\alpha_1\ldots\alpha_m n+a_1+\cdots+a_m+3m$. Поделим на n и устремим n к ∞; предел будет меньше, чем $\alpha_1\ldots\alpha_m$ для всех m. И, наконец, имеем

$$\alpha_1\ldots\alpha_m=\frac{1}{K(a_1,\ldots,a_{m-1},a_m+\alpha_m)}<\frac{1}{F_{m+1}}.$$

9.46 Для удобства будем вместо $m(n)$ писать просто m. Из аппроксимации Стирлинга получаем, что $k^n/k!$ достигает максимального значения при $k\approx m\approx n/\ln n$, поэтому заменим k на $m+k$ и найдем, что

$$\ln\frac{(m+k)^n}{(m+k)!}=n\ln m-m\ln m+m-\frac{\ln 2\pi m}{2}- \\ -\frac{(m+n)k^2}{2m^2}+O(k^3m^{-2}\log n)+O(m^{-1}).$$

На самом деле мы хотим заменить k на $\lfloor m\rfloor + k$; это добавит еще $O(km^{-1}\log n)$. Метод замены хвоста с $|k| \leqslant m^{1/2+\epsilon}$ позволяет выполнить суммирование по k и найти довольно хорошую асимптотическую оценку через величины Θ из (9.93):

$$\varpi_n = \frac{e^{m-1}m^{n-m}}{\sqrt{2\pi m}}\bigl(\Theta_{2m^2/(m+n)} + O(1)\bigr) = $$
$$= e^{m-n-1/2}m^n\sqrt{\frac{m}{m+n}}\biggl(1 + O\Bigl(\frac{\log n}{n^{1/2}}\Bigr)\biggr).$$

Отсюда получается требуемая формула, относительная погрешность которой равна $O(\log\log n/\log n)$.

9.47 Пусть $\log_m n = l+\theta$, где $0 \leqslant \theta < 1$. Нижняя сумма (с функцией “пол”) равна $l(n+1)+1-(m^{l+1}-1)/(m-1)$; верхняя сумма (с функцией “потолок”) равна $(l+1)n - (m^{l+1}-1)/(m-1)$; точная сумма есть $(l+\theta)n - n/\ln m + O(\log n)$. Если пренебречь членами порядка $o(n)$, то разность между верхней и точной суммами будет равна $\bigl(1-f(\theta)\bigr)n$, а разность между точной и нижней суммами будет равна $f(\theta)n$, где

$$f(\theta) = \frac{m^{1-\theta}}{m-1} + \theta - \frac{1}{\ln m}.$$

Максимальное значение этой функции есть $f(0) = f(1) = m/(m-1) - 1/\ln m$, а минимальное — $\ln\ln m/\ln m + 1 - \bigl(\ln(m-1)\bigr)/\ln m$. Верхнее значение ближе к точному, когда n близко к степени m, а нижнее точнее, когда θ лежит где-то между 0 и 1.

9.48 Пусть $d_k = a_k + b_k$, где a_k — количество цифр перед десятичной точкой. Тогда $a_k = 1 + \lfloor\log H_k\rfloor = \log\log k + O(1)$, где log обозначает $\log_{10}$. Для того чтобы оценить b_k, найдем число десятичных цифр, требуемое для различения некоторого числа y от близких к нему чисел $y-\epsilon$ и $y+\epsilon'$. Обозначим через $\delta = 10^{-b}$ длину диапазона чисел, округляемых до $\hat{y}$. Тогда имеем $|y-\hat{y}| \leqslant \frac{1}{2}\delta$, а также $y-\epsilon < \hat{y}-\frac{1}{2}\delta$ и $y+\epsilon' > \hat{y}+\frac{1}{2}\delta$. Следовательно, $\epsilon + \epsilon' > \delta$. А если $\delta < \min(\epsilon, \epsilon')$, то округление позволяет отличить $\hat{y}$ и от $y-\epsilon$, и от $y+\epsilon'$. Следовательно, $10^{-b_k} < 1/(k-1)+1/k$ и $10^{1-b_k} \geqslant 1/k$; получаем поэтому $b_k = \log k + O(1)$. Окончательно находим $\sum_{k=1}^{n} d_k = \sum_{k=1}^{n}(\log k + \log\log k + O(1))$, что по формуле суммирования Эйлера есть $n\log n + n\log\log n + O(n)$.

9.49 Имеем $H_n > \ln n + \gamma + \frac{1}{2}n^{-1} - \frac{1}{12}n^{-2} = f(n)$, где $f(x)$ — возрастающая при всех $x > 0$ функция; следовательно, если $n \geqslant e^{\alpha-\gamma}$, то $H_n \geqslant f(e^{\alpha-\gamma}) > \alpha$. Кроме того, $H_{n-1} < \ln n + \gamma - \frac{1}{2}n^{-1} =$

$g(n)$, где $g(x)$ — возрастающая при всех $x > 0$ функция; следовательно, если $n \leqslant e^{\alpha-\gamma}$, то $H_{n-1} \leqslant g(e^{\alpha-\gamma}) < \alpha$. Таким образом, $H_{n-1} \leqslant \alpha \leqslant H_n$ означает, что $e^{\alpha-\gamma}+1 > n > e^{\alpha+\gamma}-1$. (Более точные результаты были получены Боасом (Boas) и Ренчем (Wrench) [33].)

9.50 (a) Ожидаемый результат равен $\sum_{1\leqslant k\leqslant N} k/(k^2H_N^{(2)}) = H_N/H_N^{(2)}$, а нам надо асимптотическое значение с точностью $O(N^{-1})$:

$$\frac{\ln N+\gamma+O(N^{-1})}{\pi^2/6-N^{-1}+O(N^{-2})} = \frac{6\ln 10}{\pi^2}n+\frac{6\gamma}{\pi^2}+\frac{36\ln 10}{\pi^4}\frac{n}{10^n}+O(10^{-n}).$$

Коэффициент $(6\ln 10)/\pi^2 \approx 1.3998$ говорит, что мы ожидаем приблизительно 40% дохода.

(б) Вероятность получения прибыли равна $\sum_{n<k\leqslant N} 1/(k^2H_N^{(2)}) = 1-H_n^{(2)}/H_N^{(2)}$, и, поскольку $H_n^{(2)} = \frac{\pi^2}{6}-n^{-1}+\frac{1}{2}n^{-2}+O(n^{-3})$, эта вероятность равна

$$\frac{n^{-1}-\frac{1}{2}n^{-2}+O(n^{-3})}{\pi^2/6+O(N^{-1})} = \frac{6}{\pi^2}n^{-1}-\frac{3}{\pi^2}n^{-2}+O(n^{-3}),$$

т.е. она *убывает* с ростом n. (Ожидаемое значение в п. (а) велико потому, что оно включает такие огромные выплаты, которые повлияли бы на всю мировую экономику, если бы вдруг их действительно пришлось выплатить.)

9.51 Строго говоря, это неверно, поскольку функция, представленная $O(x^{-2})$, может быть неинтегрируемой. (Это может быть, например, функция '$[x\in S]/x^2$', где S — неизмеримое множество.) Но если предположить, что $f(x)$ — интегрируемая функция, такая, что $f(x) = O(x^{-2})$ при $x\to\infty$, то $\left|\int_n^\infty f(x)\,dx\right| \leqslant \int_n^\infty |f(x)|\,dx \leqslant \int_n^\infty Cx^{-2}\,dx = Cn^{-1}$.

Приличная и симпатичная.

9.52 На самом деле цепочку возведений в n-ю степень можно заменить любой функцией $f(n)$, сколь угодно быстро стремящейся к бесконечности. Определим последовательность $\langle m_0, m_1, m_2, \ldots\rangle$, положив $m_0 = 0$ и взяв в качестве m_k наименьшее целое число $> m_{k-1}$, такое, что

$$\left(\frac{k+1}{k}\right)^{m_k} \geqslant f(k+1)^2.$$

Пусть теперь $A(z) = \sum_{k\geqslant 1}(z/k)^{m_k}$. Этот степенной ряд сходится для всех z, потому что члены для $k > |z|$ ограничены геометрической прогрессией. Кроме того, $A(n+1) \geqslant ((n+1)/n)^{m_n} \geqslant f(n+1)^2$, следовательно, $\lim_{n\to\infty} f(n)/A(n) = 0$.

9.53 По индукции O-член представляет собой $(m-1)!^{-1} \times \int_0^x t^{m-1} f^{(m)}(x-t)\,dt$. Поскольку $f^{(m+1)}$ имеет знак, противоположный $f^{(m)}$, абсолютная величина этого интеграла ограничена значением $\bigl|f^{(m)}(0)\bigr| \int_0^x t^{m-1}\,dt$; так что погрешность ограничена абсолютной величиной первого отброшенного члена.

Звучит как "теорема о вредном".

9.54 Пусть $g(x) = f(x)/x^{\alpha}$. Тогда $g'(x) \sim -\alpha g(x)/x$ при $x \to \infty$. По теореме о среднем $g(x-\frac12) - g(x+\frac12) = -g'(y) \sim \alpha g(y)/y$ для некоторого y между $x-\frac12$ и $x+\frac12$. Далее, $g(y) = g(x)\bigl(1+O(1/x)\bigr)$, так что $g(x-\frac12) - g(x+\frac12) \sim \alpha g(x)/x = \alpha f(x)/x^{1+\alpha}$. Следовательно,

$$\sum_{k\geqslant n} \frac{f(k)}{k^{1+\alpha}} = O\Bigl(\sum_{k\geqslant n} \bigl(g(k-\tfrac12) - g(k+\tfrac12)\bigr)\Bigr) = O\bigl(g(n-\tfrac12)\bigr).$$

9.55 Оценка величины $(n+k+\frac12)\ln(1+k/n) + (n-k+\frac12)\ln(1-k/n)$ расширяется до $k^2/n + k^4/6n^3 + O(n^{-3/2+5\epsilon})$, так что нам, по-видимому, придется добавить в $b_k(n)$ множитель $e^{-k^4/6n^3}$ и взять $c_k(n) = 2^{2n} n^{-2+5\epsilon} e^{-k^2/n}$. Лучше, однако, не трогать $b_k(n)$ и положить

$$c_k(n) = 2^{2n} n^{-2+5\epsilon} e^{-k^2/n} + 2^{2n} n^{-5+5\epsilon} k^4 e^{-k^2/n},$$

заменив тем самым $e^{-k^4/6n^3}$ на $1+O(k^4/n^3)$. Сумма $\sum_k k^4 e^{-k^2/n}$ равна $O(n^{5/2})$, как показано в упр. 30.

9.56 Если $k \leqslant n^{1/2+\epsilon}$, то из приближения Стирлинга $\ln(n^{\underline{k}}/n^k) = -\frac12 k^2/n + \frac12 k/n - \frac16 k^3/n^2 + O(n^{-1+4\epsilon})$, следовательно,

$$n^{\underline{k}}/n^k = e^{-k^2/2n}\bigl(1 + k/2n - \tfrac23 k^3/(2n)^2 + O(n^{-1+4\epsilon})\bigr).$$

Просуммируем с учетом тождества из упр. 30, и не забывая опускать члены с $k=0$, и получим $-1 + \Theta_{2n} + \Theta_{2n}^{(1)} - \frac23 \Theta_{2n}^{(3)} + O(n^{-1/2+4\epsilon}) = \sqrt{\pi n/2} - \frac13 + O(n^{-1/2+4\epsilon})$.

9.57 Если воспользоваться указанием, то данная сумма превратится в $\int_0^\infty u e^{-u} \zeta(1+u/\ln n)\,du$. Одно из определений дзета-функции — ряд

$$\zeta(1+z) = z^{-1} + \sum_{m\geqslant 0} (-1)^m \gamma_m z^m/m!,$$

где $\gamma_0 = \gamma$, а γ_m — константа Стилтьеса [201, 341]

$$\lim_{n\to\infty}\Bigl(\sum_{k=1}^{n} \frac{(\ln k)^m}{k} - \frac{(\ln n)^{m+1}}{m+1}\Bigr).$$

Следовательно, наша сумма равна

$$\ln n + \gamma - 2\gamma_1 (\ln n)^{-1} + 3\gamma_2 (\ln n)^{-2} - \cdots .$$

9.58 Пусть $0 \leqslant \theta \leqslant 1$ и $f(z) = e^{2\pi i z\theta}/(e^{2\pi i z} - 1)$. Тогда имеем

$$\begin{aligned} |f(z)| &= \frac{e^{-2\pi y\theta}}{1 + e^{-2\pi y}} \leqslant 1, && \text{если } x \bmod 1 = \tfrac{1}{2}; \\ |f(z)| &\leqslant \frac{e^{-2\pi y\theta}}{|e^{-2\pi y} - 1|} \leqslant \frac{1}{1 - e^{-2\pi\epsilon}}, && \text{если } |y| \geqslant \epsilon. \end{aligned}$$

Следовательно, $|f(z)|$ ограничен на контуре и интеграл есть $O(M^{1-m})$. Вычет функции $2\pi i f(z)/z^m$ в точке $z = k \neq 0$ равен $e^{2\pi i k\theta}/k^m$; вычет в точке $z = 0$ есть коэффициент при z^{-1} в разложении

$$\frac{e^{2\pi i z\theta}}{z^{m+1}} \left(B_0 + B_1 \frac{2\pi i z}{1!} + \cdots \right) = \frac{1}{z^{m+1}} \left(B_0(\theta) + B_1(\theta) \frac{2\pi i z}{1!} + \cdots \right),$$

а именно — $(2\pi i)^m B_m(\theta)/m!$. Следовательно, сумма вычетов внутри контура равна

$$\frac{(2\pi i)^m}{m!} B_m(\theta) + 2 \sum_{k=1}^{M} e^{\pi i m/2} \frac{\cos(2\pi k\theta - \pi m/2)}{k^m}.$$

Эта сумма равна интегралу по контуру, т.е. $O(M^{1-m})$, и, следовательно, стремится к нулю при $M \to \infty$.

9.59 Если $F(x)$ — достаточно хорошо ведущая себя функция, то имеет место общее тождество

$$\sum_k F(k + t) = \sum_n G(2\pi n) e^{2\pi i n t},$$

где $G(y) = \int_{-\infty}^{+\infty} e^{-iyx} F(x)\, dx$. (Это "формула суммирования Пуассона", которую можно найти в обычных учебниках, таких как учебник Хенричи (Henrici) [182, теорема 10.6e].)

9.60 Согласно упр. 5.22 указанная формула эквивалентна

$$n^{\overline{1/2}} = n^{1/2} \left(1 - \frac{1}{8n} + \frac{1}{128n^2} + \frac{5}{1024n^3} - \frac{21}{32768n^4} + O(n^{-5}) \right).$$

Следовательно, нужный результат вытекает из упр. 6.64 и 9.44.

9.61 Главная идея — сделать α "почти" рациональным. Пусть $a_k = 2^{2^{2^k}}$ — k-е неполное частное в разложении α. Положим $n = \frac{1}{2}a_{m+1}q_m$, где $q_m = K(a_1, \ldots, a_m)$, причем m четное. Тогда $0 < \{q_m\alpha\} < 1/K(a_1, \ldots, a_{m+1}) < 1/(2n)$, и если взять $\nu = a_{m+1}/(4n)$, можно получить отклонение от равномерности $\geqslant \frac{1}{4}a_{m+1}$. Если бы оно было меньше, чем $n^{1-\epsilon}$, то мы бы имели $a_{m+1}^{\epsilon} = O(q_m^{1-\epsilon})$; но на самом деле $a_{m+1} > q_m^{2^m}$.

9.62 См. Кэнфилд (Canfield) [48]; в работе Дэйвид (David) и Бартона (Barton) [71, глава 16] можно найти асимптотику чисел Стирлинга обоих родов.

9.63 Положим $c = \phi^{2-\phi}$. Оценка $cn^{\phi-1} + o(n^{\phi-1})$ доказана Файном (Fine) [150]. Илан Варди (Ilan Vardi) заметил, что сформулированную более точную оценку можно вывести из приближенного рекуррентного соотношения $c^{\phi}n^{2-\phi}e(n) \approx -\sum_k e(k)[1 \leqslant k < cn^{\phi-1}]$, которому удовлетворяет остаточный член $e(n) = f(n) - cn^{\phi-1}$. Этому соотношению асимптотически удовлетворяет функция

$$\frac{n^{\phi-1}u(\ln\ln n/\ln\phi)}{\ln n},$$

Еще дальше в решении этой задачи продвинулся Жан-Люк Реми (Jean-Luc Rémy), Journal of Number Theory, vol. 66 (1997), 1–28.

если $u(x+1) = -u(x)$. (Варди высказал гипотезу, что

$$f(n) \;=\; n^{\phi-1}\left(c + u\Big(\frac{\ln\ln n}{\ln\phi}\Big)(\ln n)^{-1} + O\big((\log n)^{-2}\big)\right)$$

для некоторой такой функции u.) Вычисления для малых n показывают, что $f(n)$ равно ближайшему целому к $cn^{\phi-1}$ для $1 \leqslant n \leqslant 400$ с единственным исключением: $f(273) = 39 > c \cdot 273^{\phi-1} \approx 38.4997$. Однако эта малая погрешность постепенно возрастает из-за результатов, аналогичных приведенным в упр. 2.36, например $e(201636503) \approx 35.73$; $e(919986484788) \approx -1959.07$.

9.64 (Из этого тождества для $B_2(x)$ при помощи индукции по m можно легко вывести тождество из упр. 58.) Если $0 < x < 1$, то интеграл $\int_x^{1/2} \sin N\pi t\,dt/\sin\pi t$ можно записать как сумму N интегралов, каждый из которых есть $O(N^{-2})$, так что весь интеграл есть $O(N^{-1})$; константа в этом O может зависеть от x. Интегрирование тождества $\sum_{n=1}^{N}\cos 2n\pi t = \Re\big(e^{2\pi i t}(e^{2N\pi i t}-1)/(e^{2\pi i t}-1)\big) = -\frac{1}{2} + \frac{1}{2}\sin(2N+1)\pi t/\sin\pi t$ и устремление $N \to \infty$ дает $\sum_{n\geqslant 1}(\sin 2n\pi x)/n = \frac{\pi}{2} - \pi x$, соотношение, известное Эйлеру ([107, 110, часть 2, §92]). Очередное интегрирование дает желаемую формулу. (Данное решение предложено Э. М. Э. Вермутом (E. M. E. Wermuth) [367]; собственное решение Эйлера не удовлетворяет современным критериям строгости.)

9.65 Поскольку $a_0 + a_1 n^{-1} + a_2 n^{-2} + \cdots = 1 + (n-1)^{-1}(a_0 + a_1(n-1)^{-1} + a_2(n-1)^{-2} + \cdots)$, имеем рекуррентное соотношение $a_{m+1} = \sum_k \binom{m}{k} a_k$, которое совпадает с соотношением для чисел Белла. Следовательно, $a_m = \varpi_m$.

Несколько более длинное, но и более информативное доказательство основывается на том факте, что в силу (7.47) $1/(n-1)\ldots(n-m) = \sum_k \left\{ {k \atop m} \right\} / n^k$.

9.66 Ожидаемое количество различных элементов в последовательности 1, $f(1)$, $f(f(1))$, $\ldots$, когда f является случайным отображением множества $\{1, 2, \ldots, n\}$ в себя, есть функция $Q(n)$ из упр. 56, и ее значение равно $\frac{1}{2}\sqrt{2\pi n} + O(1)$; это может как-то помочь в установлении множителя $\sqrt{2\pi n}$.

"Сейчас уже твердо установлен тот парадоксальный факт, что самые крайние абстракции как раз и представляют орудия нашего познания конкретных фактов."
— А. Н. Уайтхед (A. N. Whitehead) [372]

9.67 Известно, что $\ln \chi_n \sim \frac{3}{2} n^2 \ln \frac{4}{3}$; константа $e^{-\pi/6}$ проверена эмпирически с точностью до восьми значащих цифр.

9.68 Равенство будет нарушено, если, например, $e^{n-\gamma} = m + \frac{1}{2} + \epsilon/m$ для некоторого целого m и некоторого $0 < \epsilon < \frac{1}{8}$; однако ни одного контрпримера не известно.

Б

Библиография

«Эта статья восполняет досадный пробел в литературе».
— *Math. Reviews*

ЗДЕСЬ ПРЕДСТАВЛЕНЫ РАБОТЫ, цитировавшиеся в книге. Числа на полях соответствуют номерам страниц, на которых встречаются соответствующие цитаты.

Ссылки на опубликованные задачи в основном указывают места, где можно найти решения, а не исходные постановки задач.

Где это возможно, фамилии авторов и названия работ приведены на языке оригинала. Библиография упорядочена в соответствии с английским написанием фамилий авторов.

1 N. H. Abel, letter to B. Holmboe (1823), в его *Œuvres Complètes*, first edition, 1839, volume 2, 264–265. Воспроизведено в second edition, 1881, volume 2, 254–255.

65 **2** Milton Abramowitz and Irene A. Stegun, editors, *Handbook of Mathematical Functions.* United States Government Printing Office, 1964. Воспроизведено Dover, 1965. (Имеется русский перевод: Абрамовиц М., Стиган И. (ред.) *Справочник по специальным функциям.* — М.: Наука, 1979.)

3 William W. Adams and J. L. Davison, "A remarkable class of continued fractions", *Proceedings of the American Mathematical Society* **65** (1977), 194–198. (См. также P. E. Böhmer, "Über die Transzendenz gewisser dyadischer Brüche", *Mathematische Annalen* **96** (1927), 367–377.)

4 A. V. Aho and N. J. A. Sloane, "Some doubly exponential sequences", *Fibonacci Quarterly* **11** (1973), 429–437.

26 **5** W. Ahrens, *Mathematische Unterhaltungen und Spiele.* Teubner, Leipzig, 1901. Second edition, in two volumes, 1910 and 1918. (Имеется русский перевод первого издания: Аренсъ В. *Математическія игры и развлеченія.* — Петроград: Физика, 1914.)

6 Ахиезер Н. И. *Классическая проблема моментов и некоторые вопросы анализа, связанные с нею.* — М.: Физматгиз, 1961. Английский перевод: *The Classical Moment Problem and Some Related Questions in Analysis*, Hafner, 1965. *704*

7 R. E. Allardice and A. Y. Fraser, "La Tour d'Hanoï", *Proceedings of the Edinburgh Mathematical Society* **2** (1884), 50–53. *19*

8 Désiré André, "Sur les permutations alternées", *Journal de Mathématiques pures et appliquées*, series 3, **7** (1881), 167–184.

9 George E. Andrews, "Applications of basic hypergeometric functions", *SIAM Review* **16** (1974), 441–484. *270*

10 George E. Andrews, "On sorting two ordered sets", *Discrete Mathematics* **11** (1975), 97–106. *634*

11 George E. Andrews, *The Theory of Partitions.* Addison-Wesley, 1976. (Имеется русский перевод: Эндрюс Г. *Теория разбиений.* — М.: Наука, 1982.) *402*

12 George E. Andrews, "Euler's 'exemplum memorabile inductionis fallacis' and q-trinomial coefficients", *Journal of the American Mathematical Society* **3** (1990), 653–669. *685*

13 George E. Andrews and K. Uchimura, "Identities in combinatorics IV: Differentiation and harmonic numbers", *Utilitas Mathematica* **28** (1985), 265–269.

14 Roger Apéry, "Interpolation de fractions continues et irrationalité de certaines constantes", in *Mathématiques*, Ministère des universités (France), Comité des travaux historiques et scientifiques, Section des sciences, *Bulletin de la Section des Sciences* **3** (1981), 37–53. *295*

15 M. D. Atkinson, "The cyclic towers of Hanoi", *Information Processing Letters* **13** (1981), 118–119.

16 M. D. Atkinson, "How to compute the series expansions of sec x and tg x", *American Mathematical Monthly* **93** (1986), 387–389. (Этот треугольник был впервые найден L. Seidel, "Ueber eine einfache Entstehungsweise der Bernoulli'schen Zahlen und einiger verwandten Reihen", *Sitzungsberichte der mathematisch-physikalischen Classe der königlich bayerischen Akademie der Wissenschaften zu München* **7** (1877), 157–187.)

17 Paul Bachmann, *Die analytische Zahlentheorie.* Teubner, Leipzig, 1894. *534*

18 W. N. Bailey, *Generalized Hypergeometric Series.* Cambridge University Press, 1935; second edition, 1964. *279*

654 **19** W. N. Bailey, "The generating function for Jacobi polynomials", *Journal of the London Mathematical Society* **13** (1938), 243–246.

26 **20** W. W. Rouse Ball and H. S. M. Coxeter, *Mathematical Recreations and Essays*, twelfth edition. University of Toronto Press, 1974. (Переработка книги W. W. Rouse Ball *Mathematical Recreations and Problems*, first published by Macmillan, 1892.) (Имеется русский перевод: Болл У., Коксетер Г. *Математические эссе и развлечения.* — М.: Мир, 1986.)

21 P. Barlow, "Demonstration of a curious numerical proposition", *Journal of Natural Philosophy, Chemistry, and the Arts* **27** (1810), 193–205.

22 Samuel Beatty, "Problem 3177", *American Mathematical Monthly* **34** (1927), 159–160.

403 **23** E. T. Bell, "Euler algebra", *Transactions of the American Mathematical Society* **25** (1923), 135–154.

24 E. T. Bell, "Exponential numbers", *American Mathematical Monthly* **41** (1934), 411–419.

25 Edward A. Bender, "Asymptotic methods in enumeration", *SIAM Review* **16** (1974), 485–515.

346 **26** Jacobi Bernoulli, *Ars Conjectandi*, opus posthumum. Basel, 1713. Воспроизведено в *Die Werke von Jakob Bernoulli*, volume 3, 107–286. (Имеется русский перевод четвертой части (числам Бернулли посвящена вторая часть): Бернулли Я. *Искусство предположения(й).* — СПб.: Имп. Академия наук, 1913; воспроизведено в сборнике: Бернулли Я. *О законе больших чисел.* — М.: Наука, 1986.)

27 J. Bertrand, "Mémoire sur le nombre de valeurs que peut prendre une fonction quand on y permute les lettres qu'elle renferme", *Journal de l'École Royale Polytechnique* **18**, cahier 30 (1845), 123–140.

65 **28** William H. Beyer, editor, *CRC Standard Mathematical Tables and Formulae*, 29th edition. CRC Press, Boca Raton, Florida, 1991.

471 **29** J. Bienaymé, "Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés", *Comptes Rendus hebdomadaires des séances de l'Académie des Sciences* (Paris) **37** (1853), 309–324.

30 J. Binet, “Mémoire sur un système de Formules analytiques, et leur application à des considérations géométriques”, *Journal de l'École Polytechnique* **9**, cahier 16 (1812), 280–354.

31 J. Binet, “Mémoire sur l'intégration des équations linéaires aux différences finies, d'un ordre quelconque, à coefficients variables”, *Comptes Rendus hebdomadaires des séances de l'Académie des Sciences* (Paris) **17** (1843), 559–567. *364*

32 Gunnar Blom, “Problem E 3043: Random walk until no shoes”, *American Mathematical Monthly* **94** (1987), 78–79.

33 R. P. Boas, Jr. and J. W. Wrench, Jr., “Partial sums of the harmonic series”, *American Mathematical Monthly* **78** (1971), 864–870. *714*

34 P. Bohl, “Über ein in der Theorie der säkularen Störungen vorkommendes Problem”, *Journal für die reine und angewandte Mathematik* **135** (1909), 189–283. (Имеется русский перевод: Боль П. Г. “Об одной проблеме в теории вековых возмущений”. *Собрание трудов.*— Рига: Зинатне, 1974, 338–453.) *119*

35 Émile Borel, *Leçons sur les séries à termes positifs.* Paris, 1902.

36 Jonathan M. Borwein and Peter B. Borwein, *Pi and the AGM.* Wiley, 1987.

37 Richard P. Brent, “The first occurrence of large gaps between successive primes”, *Mathematics of Computation* **27** (1973), 959–963. *628*

38 Richard P. Brent, “Computation of the regular continued fraction for Euler's constant”, *Mathematics of Computation* **31** (1977), 771–777. *372, 673*

39 John Brillhart, “Some miscellaneous factorizations”, *Mathematics of Computation* **17** (1963), 447–450.

40 Achille Brocot, “Calcul des rouages par approximation, nouvelle méthode”, *Revue Chronométrique* **3** (1861), 186–194. (He also published a 97-page monograph with the same title in 1862.) *155*

41 Maxey Brooke and C. R. Wall, “Problem B-14: A little surprise”, *Fibonacci Quarterly* **1**, 3 (1963), 80.

42 Brother U. Alfred (Brousseau), “A mathematician's progress”, *Mathematics Teacher* **59** (1966), 722–727.

43 Morton Brown, “Problem 6439: A periodic sequence”, *American Mathematical Monthly* **92** (1985), 218. *600*

(Такие статьи в этой книге не цитировались.)

44 T. Brown, “Infinite multi-variable subpolynormal Woffles which do not satisfy the lower regular Q-property (Piffles)”, in *A Collection of 250 Papers on Woffle Theory Dedicated to R. S. Green on His 23rd Birthday*. Цитировалось в A. K. Austin, “Modern research in mathematics”, *The Mathematical Gazette* **51** (1967), 149–150.

45 Thomas C. Brown, “Problem E 2619: Squares in a recursive sequence”, *American Mathematical Monthly* **85** (1978), 52–53.

433 46 William G. Brown, “Historical note on a recurrent combinatorial problem”, *American Mathematical Monthly* **72** (1965), 973–977.

47 S. A. Burr, “On moduli for which the Fibonacci sequence contains a complete system of residues”, *Fibonacci Quarterly* **9** (1971), 497–504.

717 48 E. Rodney Canfield, “On the location of the maximum Stirling number(s) of the second kind”, *Studies in Applied Mathematics* **59** (1978), 83–93.

49 L. Carlitz, “The generating function for $\max(n_1, n_2, \cdots, n_k)$”, *Portugaliae Mathematica* **21** (1962), 201–207.

52 50 Lewis Carroll (pseudonym of C. L. Dodgson), *Through the Looking Glass and What Alice Found There*. Macmillan, 1871. (Имеется русский перевод: Кэрролл Л. *Приключения Алисы в Стране чудес. Сквозь зеркало и что там увидела Алиса, или Алиса в Зазеркалье.* — М.: Наука, 1979.)

356 51 Jean-Dominique Cassini, “Une nouvelle progression de nombres”, *Histoire de l’Académie Royale des Sciences*, Paris, volume 1, 201. (В этой работе Кассини собраны математические результаты, представленные Королевской академии наук в 1680 году. Издание вышло в 1733 году.)

256 52 E. Catalan, “Note sur une Équation aux différences finies”, *Journal de Mathématiques pures et appliquées* **3** (1838), 508–516.

53 Augustin-Louis Cauchy, *Cours d’analyse de l’École Royale Polytechnique*. Imprimerie Royale, Paris, 1821. Воспроизведено в его *Œuvres complètes*, series 2, volume 3. (Имеется русский перевод: Коши О.-Л. *Краткое изложение уроков о дифференциальном и интегральном исчислении.* — СПб.: Имп. Академия наук, 1831.)

54 Arnold Buffum Chace, *The Rhind Mathematical Papyrus*, volume 1. Mathematical Association of America, 1927. (Включает великолепный список литературы египетских математиков, подготовленный Р. Арчибальдом (R. C. Archibald).)

55 M. Chaimovich, G. Freiman, and J. Schönheim, “On exceptions to Szegedy's theorem”, *Acta Arithmetica* **49** (1987), 107–112. *628*

56 P. L. Tchebichef (Chebyshev), “Mémoire sur les nombres premiers”, *Journal de Mathématiques pures et appliquées* **17** (1852), 366–390. Воспроизведено в его *Œuvres*, volume 1, 51–70. (Имеются русские переводы: Чебышев П. Л. “О простыхъ числахъ”, *Сочиненія П. Л. Чебышева*, томъ 1.— СПб.: Имп. Академия Наукъ, 1899, 51–70; Чебышев П. Л. *Полное собрание сочинений*, том 1.— М.-Л.: Изд. АН СССР, 1944, 191–207.)

57 Чебышев П. Л. “О среднихъ величинахъ”. *Математический сборникъ* **2** (1867), 1–9. Воспроизведено в: Чебышев П. Л. *Полное собрание сочинений*, том 2.— М.-Л.: Изд. АН СССР, 1944, 431–437. Французский перевод: “Des valeurs moyennes”, *Journal de Mathématiques pures et appliquées*, series 2, **12** (1867), 177–184; воспроизведено в его *Œuvres*, volume 1, 685–694. *471*

58 Чебышев П. Л. “О приближенныхъ выраженіяхъ одніхъ интеграловъ черезъ другие, взятые в тех же пределахъ”, *Сообщения и протоколы заседаний математическаго общества при Императорскомъ Харьковскомъ Университете* **4**, 2 (1882), 93–98. Воспроизведено в: Чебышев П. Л. *Полное собрание сочинений*, том 3.— М.-Л.: Изд. АН СССР, 1944, 191–207, 128–131. Французский перевод: “Sur les expressions approximatives des intégrales définies par les autres prises entre les mêmes limites”, в его *Œuvres*, volume 2, 716–719. *61*

59 F. R. K. Chung and R. L. Graham, “On the cover polynomial of a digraph”, *Journal of Combinatorial Theory*, series B, **65** (1995), 273–290. *665*

60 Th. Clausen, “Ueber die Fälle, wenn die Reihe von der Form

$$y = 1 + \frac{\alpha}{1}\cdot\frac{\beta}{\gamma}x + \frac{\alpha.\alpha+1}{1.2}\cdot\frac{\beta.\beta+1}{\gamma.\gamma+1}x^2 + \text{etc.}$$

ein Quadrat von der Form

$$z = 1 + \frac{\alpha'}{1}\cdot\frac{\beta'}{\gamma'}\cdot\frac{\delta'}{\epsilon'}x + {} + \frac{\alpha'.\alpha'+1}{1.2}\cdot\frac{\beta'.\beta'+1}{\gamma'.\gamma'+1}\cdot\frac{\delta'.\delta'+1}{\epsilon'.\epsilon'+1}x^2 + \text{etc.}\ \text{hat”},$$

Journal für die reine und angewandte Mathematik **3** (1828), 89–91.

61 Th. Clausen, "Beitrag zur Theorie der Reihen", *Journal für die reine und angewandte Mathematik* **3** (1828), 92–95.

62 Th. Clausen, "Theorem", *Astronomische Nachrichten* **17** (1840), columns 351–352.

357 **63** Stuart Dodgson Collingwood, *The Lewis Carroll Picture Book.* T. Fisher Unwin, 1899. Воспроизведено Dover в 1961 году под новым названием *Diversions and Digressions of Lewis Carroll.*

64 Louis Comtet, *Advanced Combinatorics.* Dordrecht, Reidel, 1974.

600 **65** J. H. Conway and R. L. Graham, "Problem E 2567: A periodic recurrence", *American Mathematical Monthly* **84** (1977), 570–571.

628 **66** Harald Cramér, "On the order of magnitude of the difference between consecutive prime numbers", *Acta Arithmetica* **2** (1937), 23–46.

67 A. L. Crelle, "Démonstration élémentaire du théorème de Wilson généralisé", *Journal für die reine und angewandte Mathematik* **20** (1840), 29–56.

68 D. W. Crowe, "The n-dimensional cube and the Tower of Hanoi", *American Mathematical Monthly* **63** (1956), 29–30.

703 **69** János A. Csirik, "Optimal strategy for the first player in the Penney ante game", *Combinatorics, Probability and Computing* **1** (1992), 311–321.

70 D. R. Curtiss, "On Kellogg's Diophantine problem", *American Mathematical Monthly* **29** (1922), 380–387.

717 **71** F. N. David and D. E. Barton, *Combinatorial Chance.* Hafner, 1962.

264 **72** Philip J. Davis, "Leonhard Euler's integral: A historical profile of the Gamma function", *American Mathematical Monthly* **66** (1959), 849–869.

373 **73** J. L. Davison, "A series and its associated continued fraction", *Proceedings of the American Mathematical Society* **63** (1977), 29–32.

535, 538, 708 **74** N. G. de Bruijn, *Asymptotic Methods in Analysis.* North-Holland, 1958; third edition, 1970. Воспроизведено Dover, 1981. (Имеется русский перевод первого издания: де Брейн Н.Г. *Асимптотические методы в анализе.* — М.: ИЛ, 1961.)

75 N. G. de Bruijn, "Problem 9", *Nieuw Archief voor Wiskunde*, series 3, **12** (1964), 68.

76 Abraham de Moivre, *Miscellanea analytica de seriebus et quadraturis.* London, 1730. *362, 577*

77 R. Dedekind, "Abriß einer Theorie der höheren Congruenzen in Bezug auf einen reellen Primzahl-Modulus", *Journal für die reine und angewandte Mathematik* **54** (1857), 1–26. Воспроизведено в его *Gesammelte mathematische Werke*, volume 1, 40–67. *178*

78 Leonard Eugene Dickson, *History of the Theory of Numbers.* Carnegie Institution of Washington, volume 1, 1919; volume 2, 1920; volume 3, 1923. Воспроизведено Stechert, 1934, and by Chelsea, 1952, 1971. *610*

79 Edsger W. Dijkstra, *Selected Writings on Computing: A Personal Perspective.* Springer-Verlag, 1982.

80 G. Lejeune Dirichlet, "Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einigen Anwendungen auf die Theorie der Zahlen", *Bericht über die Verhandlungen der Königlich-Preußischen Akademie der Wissenschaften zu Berlin* (1842), 93–95. Воспроизведено в его *Werke*, volume 1, 635–638.

81 A. C. Dixon, "On the sum of the cubes of the coefficients in a certain expansion by the binomial theorem", *Messenger of Mathematics* **20** (1891), 79–80.

82 John Dougall, "On Vandermonde's theorem, and some more general expansions", *Proceedings of the Edinburgh Mathematical Society* **25** (1907), 114–132. *220*

83 A. Conan Doyle, "The sign of the four; or, The problem of the Sholtos", *Lippincott's Monthly Magazine* (Philadelphia) **45** (1890), 147–223. (Имеется русский перевод: Конан Дойль А. "Знак четырех". *Собрание сочинений, том 1.*— М.: Правда, 1966, 151–264.) *285, 488*

84 A. Conan Doyle, "The adventure of the final problem", *The Strand Magazine* **6** (1893), 558–570. (Имеется русский перевод: Конан Дойль А. "Последнее дело Холмса". *Собрание сочинений, том 2.*— М.: Правда, 1966, 221–240.) *209*

85 P. du Bois-Reymond, "Sur la grandeur relative des infinis des fonctions", *Annali di Matematica pura ed applicata*, series 2, **4** (1871), 338–353. *531*

86 Harvey Dubner, "Generalized repunit primes", *Mathematics of Computation* **61** (1993), 927–930.

87 Henry Ernest Dudeney, *The Canterbury Puzzles and Other Curious Problems.* E. P. Dutton, New York, 1908; 4th edition, Dover, 1958. (Свое первое обобщение ханойской башни Дьюдени опубликовал в *The Weekly Dispatch*, on 15 November 1896, 25 May 1902 and 15 March 1903.) (Имеется русский перевод четвертого издания: Дьюдени Г. Э. *Кентерберийские головоломки.* — М.: Мир, 1979.)

24 **88** G. Waldo Dunnington, *Carl Friedrich Gauss: Titan of Science.* Exposition Press, New York, 1955.

297 **89** F. J. Dyson, "Some guesses in the theory of partitions", *Eureka* **8** (1944), 10–15.

201 **90** A. W. F. Edwards, *Pascal's Arithmetical Triangle.* Oxford University Press, 1987.

255 **91** G. Eisenstein, "Entwicklung von $\alpha^{\alpha^{\alpha^{\cdot^{\cdot}}}}$", *Journal für die reine und angewandte Mathematik* **28** (1844), 49–52. Воспроизведено в его *Mathematische Werke* **1**, 122–125.

172 **92** Noam D. Elkies, "On $A^4 + B^4 + C^4 = D^4$", *Mathematics of Computation* **51** (1988), 825–835.

93 Erdős Pál, "Az $\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n} = \frac{a}{b}$ egyenlet egész számú megoldásairól", *Matematikai Lapok* **1** (1950), 192–209. Аннотация на английском языке на стр. 210.

503 **94** Paul Erdős, "My Scottish Book 'problems'", in *The Scottish Book: Mathematics from the Scottish Café*, edited by R. Daniel Mauldin, 1981, 35–45.

616, 628 **95** P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory.* Université de Genève, L'Enseignement Mathématique, 1980.

628, 654 **96** P. Erdős, R. L. Graham, I. Z. Ruzsa, and E. G. Straus, "On the prime factors of $\binom{2n}{n}$", *Mathematics of Computation* **29** (1975), 83–92.

97 Arulappah Eswarathasan and Eugene Levine, "p-integral harmonic sums", *Discrete Mathematics* **91** (1991), 249–257.

144 **98** Euclid, *ΣΤΟΙΧΕΙΑ.* Старинный манускрипт, впервые напечатан в Базеле в 1533 году. Школьный учебник (на греческом и латыни) Гейберга (J. L. Heiberg) в пяти томах, Teubner, Leipzig, 1883–1888. (Имеется русский перевод: *Начала Евклида*, книги I–VI, VII-X, XI-XV. — М.-Л.: Гостехиздат, 1948–1950.)

99 Leonhard Euler, letter to Christian Goldbach (13 October 1729), in *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle*, edited by P. H. Fuss, St. Petersburg, 1843, volume 1, 3–7. *264*

100 L. Eulero, “De progressionibus transcendentibus seu quarum termini generales algebraice dari nequeunt”, *Commentarii academiæ scientiarum imperialis Petropolitanæ* **5** (1730), 36–57. Воспроизведено в его *Opera Omnia*, series 1, volume 14, 1–24. *264*

101 Leonh. Eulero, “Methodus generalis summandi progressiones”, *Commentarii academiæ scientiarum imperialis Petropolitanæ* **6** (1732), 68–97. Воспроизведено в его *Opera Omnia*, series 1, volume 14, 42–72. *564*

102 Leonh. Eulero, “Observationes de theoremate quodam Fermatiano, aliisque ad numeros primos spectantibus”, *Commentarii academiæ scientiarum imperialis Petropolitanæ* **6** (1732), 103–107. Воспроизведено в его *Opera Omnia*, series 1, volume 2, 1–5. Воспроизведено в его *Commentationes arithmeticæ collectæ*, volume 1, 1–3. *173*

103 Leonh. Eulero, “De progressionibus harmonicis observationes”, *Commentarii academiæ scientiarum imperialis Petropolitanæ* **7** (1734), 150–161. Воспроизведено в его *Opera Omnia*, series 1, volume 14, 87–100. *340*

104 Leonh. Eulero, “Methodus universalis series summandi ulterius promota”, *Commentarii academiæ scientiarum imperialis Petropolitanæ* **8** (1736), 147–158. Воспроизведено в его *Opera Omnia*, series 1, volume 14, 124–137. *328*

105 Leonh. Euler, “De fractionibus continuis, Dissertatio”, *Commentarii academiæ scientiarum imperialis Petropolitanæ* **9** (1737), 98–137. Воспроизведено в его *Opera Omnia*, series 1, volume 14, 187–215. *162*

106 Leonh. Euler, “Variæ observationes circa series infinitas”, *Commentarii academiæ scientiarum imperialis Petropolitanæ* **9** (1737), 160–188. Воспроизведено в его *Opera Omnia*, series 1, volume 14, 216–244.

107 Leonhard Euler, letter to Christian Goldbach (4 July 1744), in *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle*, edited by P. H. Fuss, St. Petersburg, 1843, volume 1, 278–293. *717*

108 Leonhardo Eulero, *Introductio in Analysin Infinitorum.* Tomus primus, Lausanne, 1748. Воспроизведено в его *Opera Omnia*, series 1, volume 8. Перевод на французский (1786), немецкий (1788), русский (1936), английский (1988). (Имеются русские переводы: Эйлер Л. *Введение в анализ бесконечных*, том 1.— М.-Л.: ОНТИ, 1936; 2-е изд.: М.: Физматгиз, 1961.)

109 L. Eulero, "De partitione numerorum", *Novi commentarii academiæ scientiarum imperialis Petropolitanæ* **3** (1750), 125–169. Воспроизведено в его *Commentationes arithmeticæ collectæ*, volume 1, 73–101. Воспроизведено в его *Opera Omnia*, series 1, volume 2, 254–294.

73, 328, 658, 717 **110** Leonhardo Eulero, *Institutiones Calculi Differentialis cum eius usu in Analysi Finitorum ac Doctrina Serierum.* St. Petersburg, Academiæ Imperialis Scientiarum Petropolitanæ, 1755. Воспроизведено в его *Opera Omnia*, series 1, volume 10. Перевод на немецкий (1790). (Имеется русский перевод: Эйлер Л. *Дифференциальное исчисление.*— М.-Л.: Гостехиздат, 1949.)

175 **111** L. Eulero, "Theoremata arithmetica nova methodo demonstrata", *Novi commentarii academiæ scientiarum imperialis Petropolitanæ* **8** (1760), 74–104. (Also presented in 1758 to the Berlin Academy.) Воспроизведено в его *Commentationes arithmeticæ collectæ*, volume 1, 274–286. Воспроизведено в его *Opera Omnia*, series 1, volume 2, 531–555.

367, 369 **112** L. Eulero, "Specimen algorithmi singularis", *Novi commentarii academiæ scientiarum imperialis Petropolitanæ* **9** (1762), 53–69. (В 1757 году представлена также в Берлинскую академию.) Воспроизведено в его *Opera Omnia*, series 1, volume 15, 31–49.

685 **113** L. Eulero, "Observationes analyticæ", *Novi commentarii academiæ scientiarum imperialis Petropolitanæ* **11** (1765), 124–143. Воспроизведено в его *Opera Omnia*, series 1, volume 15, 50–69.

114 Leonhard Euler, *Vollständige Anleitung zur Algebra. Erster Theil. Von den verschiedenen Rechnungs-Arten, Verhältnissen und Proportionen.* St. Petersburg, 1770. Воспроизведено в его *Opera Omnia*, series 1, volume 1. Перевод на русский (1768), голландский (1773), французский (1774), латынь (1790), английский (1797). (Имеются издания на русском языке: *Универсальная арифметика Леонарда Эйлера*, том 1.— СПб.: 1768, 2-е изд., 1787; *Основания алгебры Леонарда Эйлера*, том 1.— СПб.: 1912.)

172 **115** L. Eulero, "Observationes circa bina biquadrata quorum summam in duo alia biquadrata resolvere liceat", *Novi commentarii academiæ scientiarum imperialis Petropolitanæ* **17** (1772), 64–69. Воспроизведено в его *Commentationes arithmeticæ col-*

lectæ, volume 1, 473–476. Воспроизведено в его *Opera Omnia*, series 1, volume 3, 211–217.

116 L. Eulero, "Observationes circa novum et singulare progressionum genus", *Novi commentarii academiæ scientiarum imperialis Petropolitanæ* **20** (1775), 123–139. Воспроизведено в его *Opera Omnia*, series 1, volume 7, 246–261. *614*

117 L. Eulero, "De serie Lambertina, plurimisque eius insignibus proprietatibus", *Acta academiæ scientiarum imperialis Petropolitanæ* **3**,2 (1779), 29–51. Воспроизведено в его *Opera Omnia*, series 1, volume 6, 350–369. *255*

118 L. Eulero, "Specimen transformationis singularis serierum", *Nova acta academiæ scientiarum imperialis Petropolitanæ* **12** (1794), 58–70. Submitted for publication in 1778. Воспроизведено в его *Opera Omnia*, series 1, volume 16(2), 41–55. *260*

119 Johann Faulhaber, *Academia Algebræ*, Darinnen die miraculosische Inventiones zu den höchsten Cossen weiters *continuirt* und *profitiert* werden, ... biß auff die regulierte *Zensicubiccubic* Coß durch offnen Truck *publiciert* worden. Augsburg, 1631. *352*

120 William Feller, *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, 1950; second edition, 1957; third edition, 1968. (Имеется русский перевод третьего издания: Феллер У. *Введение в теорию вероятностей и ее приложения*, том 1.— М.: Мир, 1984.) *461*

121 Pierre de Fermat, letter to Marin Mersenne (25 December 1640), in *Œuvres de Fermat*, volume 2, 212–217. *173*

122 Leonardo filio Bonacii Pisano (Fibonacci), *Liber Abaci*. First edition, 1202 (now lost); second edition, 1228. Воспроизведено в *Scritti di Leonardo Pisano*, edited by Baldassarre Boncompagni, 1857, volume 1.

123 Bruno de Finetti, *Teoria delle Probabilità*. Turin, 1970. Английский перевод: *Theory of Probability*, Wiley, 1974–1975. *45*

124 Michael E. Fisher, "Statistical mechanics of dimers on a plane lattice", *Physical Review* **124** (1961), 1664–1672.

125 R. A. Fisher, "Moments and product moments of sampling distributions", *Proceedings of the London Mathematical Society*, series 2, **30** (1929), 199–238.

126 Pierre Forcadel, *L'arithmeticque*. Paris, 1557.

42 **127** J. Fourier, “Refroidissement séculaire du globe terrestre”, *Bulletin des Sciences par la Société philomathique de Paris*, series 3, **7** (1820), 58–70. Воспроизведено в *Œuvres de Fourier*, volume 2, 271–288.

616 **128** Aviezri S. Fraenkel, “Complementing and exactly covering sequences”, *Journal of Combinatorial Theory*, series A, **14** (1973), 8–20.

672 **129** Aviezri S. Fraenkel, “How to beat your Wythoff games' opponent on three fronts”, *American Mathematical Monthly* **89** (1982), 353–361.

130 J. S. Frame, B. M. Stewart, and Otto Dunkel, “Partial solution to problem 3918”, *American Mathematical Monthly* **48** (1941), 216–219.

131 Piero della Francesca, Франческа, Пьеро делла (Francesca, Piero della) Francesca, Piero della *Libellus de quinque corporibus regularibus.* Vatican Library, manuscript Urbinas 632. Перевод на итальянский Луки Пачоли (Luca Pacioli) в виде части 3 в Pacioli *Diuine Proportione*, Venice, 1509.

655 **132** J. Franel, Solutions to questions 42 and 170, in L'*Intermédiaire des Mathématiciens* **1** (1894), 45–47; **2** (1895), 33–35.

133 W. D. Frazer and A. C. McKellar, “Samplesort: A sampling approach to minimal storage tree sorting”, *Journal of the Association for Computing Machinery* **27** (1970), 496–507.

614 **134** Michael Lawrence Fredman, *Growth Properties of a Class of Recursively Defined Functions.* Ph.D. thesis, Stanford University, Computer Science Department, 1972.

436 **135** Nikolao Fuss, “Solutio quæstionis, quot modis polygonum n laterum in polygona m laterum, per diagonales resolvi quæat”, *Nova acta academiæ scientiarum imperialis Petropolitanæ* **9** (1791), 243–251.

364 **136** Martin Gardner, “About phi, an irrational number that has some remarkable geometrical expressions”, *Scientific American* **201**, 2 (August 1959), 128–134. Воспроизведено с дополнениями в его книге *The 2nd Scientific American Book of Mathematical Puzzles & Diversions*, 1961, 89–103. (Имеется русский перевод: Гарднер М. *Математические головоломки и развлечения.* — М.: Мир, 1971.)

494 **137** Martin Gardner, “On the paradoxical situations that arise from nontransitive relations”, *Scientific American* **231**, 4 (October

1974), 120–124. Воспроизведено с дополнениями в его книге *Time Travel and Other Mathematical Bewilderments*, 1988, 55–69. (Имеется русский перевод: Гарднер М. *Путешествие во времени.* — М.: Мир, 1990.)

138 Martin Gardner, "From rubber ropes to rolling cubes, a miscellany of refreshing problems", *Scientific American* **232**, 3 (March 1975), 112–114; **232**, 4 (April 1975), 130, 133. Воспроизведено с дополнениями в его книге *Time Travel and Other Mathematical Bewilderments*, 1988, 111–124. (Имеется русский перевод: Гарднер М. *Путешествие во времени.* — М.: Мир, 1990.)

139 Martin Gardner, "On checker jumping, the amazon game, weird dice, card tricks and other playful pastimes", *Scientific American* **238**, 2 (February 1978), 19, 22, 24, 25, 30, 32. Воспроизведено с дополнениями в его книге *Penrose Tiles to Trapdoor Ciphers*, 1989, 265–280. (Имеется русский перевод: Гарднер М. *От мозаик Пенроуза к надежным шифрам.* — М.: Мир, 1991.)

140 J. Garfunkel, "Problem E 1816: An inequality related to Stirling's formula", *American Mathematical Monthly* **74** (1967), 202.

141 George Gasper and Mizan Rahman, *Basic Hypergeometric Series.* Cambridge University Press, 1990. (Имеется русский перевод: Гаспер Дж., Рахман М. *Базисные гипергеометрические ряды.* — М.: Мир, 1993.) *279*

142 C. F. Gauss, *Disquisitiones Arithmeticæ.* Leipzig, 1801. Воспроизведено в его *Werke*, volume 1. *163*

143 Carolo Friderico Gauss, "Disquisitiones generales circa seriem infinitam *260, 266, 277, 632*

$$1+\frac{\alpha\beta}{1\,.\,\gamma}x+\frac{\alpha(\alpha+1)\beta(\beta+1)}{1\,.\,2\,.\,\gamma(\gamma+1)}xx+ \\ +\frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1\,.\,2\,.\,3\,.\,\gamma(\gamma+1)(\gamma+2)}x^3+\text{etc.}$$

Pars prior", *Commentationes societatis regiæ scientiarum Gottingensis recentiores* **2** (1813). (Опубликовано Королевским Геттингенским обществом 20 января 1812 года.) Воспроизведено в его *Werke*, volume 3, 123–163, вместе с неопубликованной частью на с. 207–229. (Имеется русский перевод: Гаусс К. Ф. *Труды по теории чисел. Арифметические исследования.* — М.: Изд. АН СССР, 1959, 15–583.)

144 C. F. Gauss, "Pentagramma mirificum", написано до 1836 года. Опубликовано в его *Werke*, volume 3, 480–490.

658 **145** Angelo Genocchi, "Intorno all'espressione generale de'numeri Bernulliani", *Annali di Scienze Matematiche e Fisiche* **3** (1852), 395–405.

146 Ira Gessel, "Some congruences for Apéry numbers", *Journal of Number Theory* **14** (1982), 362–368.

332 **147** Ira Gessel and Richard P. Stanley, "Stirling polynomials", *Journal of Combinatorial Theory*, series A, **24** (1978), 24–33.

332 **148** Jekuthiel Ginsburg, "Note on Stirling's numbers", *American Mathematical Monthly* **35** (1928), 77–80.

149 J. W. L. Glaisher, "On the product $1^1.2^2.3^3 \ldots n^n$", *Messenger of Mathematics* **7** (1877), 43–47.

717 **150** Solomon W. Golomb, "Problem 5407: A nondecreasing indicator function", *American Mathematical Monthly* **74** (1967), 740–743.

607 **151** Solomon W. Golomb, "The 'Sales Tax' theorem", *Mathematics Magazine* **49** (1976), 187–189.

553 **152** Solomon W. Golomb, "Problem E 2529: An application of $\psi(x)$", *American Mathematical Monthly* **83** (1976), 487–488.

153 I. J. Good, "Short proof of a conjecture by Dyson", *Journal of Mathematical Physics* **11** (1970), 1884.

280 **154** R. William Gosper, Jr., "Decision procedure for indefinite hypergeometric summation", *Proceedings of the National Academy of Sciences of the United States of America* **75** (1978), 40–42.

613 **155** R. L. Graham, "On a theorem of Uspensky", *American Mathematical Monthly* **70** (1963), 407–409.

156 R. L. Graham, "A Fibonacci-like sequence of composite numbers", *Mathematics Magazine* **37** (1964), 322–324.

157 R. L. Graham, "Problem 5749", *American Mathematical Monthly* **77** (1970), 775.

615 **158** Ronald L. Graham, "Covering the positive integers by disjoint sets of the form $\{\,[n\alpha+\beta] : n = 1, 2, \ldots\,\}$", *Journal of Combinatorial Theory*, series A, **15** (1973), 354–358.

159 R. L. Graham, "Problem 1242: Bijection between integers and composites", *Mathematics Magazine* **60** (1987), 180.

160 R. L. Graham and D. E. Knuth, "Problem E 2982: A double infinite sum for $|x|$", *American Mathematical Monthly* **96** (1989), 525–526.

161 Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics: A Foundation for Computer Science.* Addison-Wesley, 1989; second edition, 1994.

162 R. L. Graham and H. O. Pollak, "Note on a nonlinear recurrence related to $\sqrt{2}$", *Mathematics Magazine* **43** (1970), 143–145.

163 Guido Grandi, letter to Leibniz (July 1713), in *Leibnizens Mathematische Schriften*, volume 4, 215–217. *84*

164 Daniel H. Greene and Donald E. Knuth, *Mathematics for the Analysis of Algorithms.* Birkhäuser, Boston, 1981; third edition, 1990. (Имеется русский перевод второго издания: Грин Д.Х., Кнут Д.Э. *Математические методы анализа алгоритмов.* — М.: Мир, 1987.) *640*

165 Samuel L. Greitzer, *International Mathematical Olympiads, 1959–1977.* Mathematical Association of America, 1978.

166 Oliver A. Gross, "Preferential arrangements", *American Mathematical Monthly* **69** (1962), 4–8.

167 Branko Grünbaum, "Venn diagrams and independent families of sets", *Mathematics Magazine* **48** (1975), 12–23. *596*

168 L. J. Guibas and A. M. Odlyzko, "String overlaps, pattern matching, and nontransitive games", *Journal of Combinatorial Theory*, series A, **30** (1981), 183–208. *702*

169 Richard K. Guy, *Unsolved Problems in Number Theory.* Springer-Verlag, 1981. *628*

170 Inger Johanne Håland and Donald E. Knuth, "Polynomials involving the floor function", *Mathematica Scandinavica* **76** (1995), 194–200. Воспроизведено в Knuth D. E. *Selected Papers on Discrete Mathematics*, 257–264. *615*

171 Marshall Hall, Jr., *The Theory of Groups.* Macmillan, 1959. (Имеется русский перевод: Холл М. *Теория групп.* — М.: ИЛ, 1962.)

172 P. R. Halmos, "How to write Mathematics", *L'Enseignement mathématique* **16** (1970), 123–152. Воспроизведено в *How to Write Mathematics*, American Mathematical Society, 1973, 19–48. (Имеется русский перевод: Халмош П. Р. "Как писать математические тексты", *Успехи математических наук*, **26** (1971), 243–269.) *8*

173 Paul R. Halmos, *I Want to Be a Mathematician: An Automathography.* Springer-Verlag, 1985. Воспроизведено Mathematical Association of America, 1988. *7*

371 **174** G. H. Halphen, "Sur des suites de fractions analogues à la suite de Farey", *Bulletin de la Société mathématique de France* **5** (1876), 170–175. Воспроизведено в его *Œuvres*, volume 2, 102–107.

704 **175** Hans Hamburger, "Über eine Erweiterung des Stieltjesschen Momentenproblems", *Mathematische Annalen* **81** (1920), 235–319; **82** (1921), 120–164, 168–187.

7 **176** J. M. Hammersley, "On the enfeeblement of Mathematical skills by 'Modern Mathematics' and by similar soft intellectual trash in schools and universities", *Bulletin of the Institute of Mathematics and Its Applications* **4**, 4 (October 1968), 66–85.

177 J. M. Hammersley, "An undergraduate exercise in manipulation", *The Mathematical Scientist* **14** (1989), 1–23.

65 **178** Eldon R. Hansen, *A Table of Series and Products.* Prentice-Hall, 1975.

533 **179** G. H. Hardy, *Orders of Infinity: The 'Infinitärcalcül' of Paul du Bois-Reymond.* Cambridge University Press, 1910; second edition, 1924.

180 G. H. Hardy, "A Mathematical theorem about golf", *The Mathematical Gazette* **29** (1945), 226–227. Воспроизведено в его *Collected Papers*, volume 7, 488.

149 **181** G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers.* Clarendon Press, Oxford, 1938; fifth edition, 1979.

365, 403, 716 **182** Peter Henrici, *Applied and Computational Complex Analysis.* Wiley, volume 1, 1974; volume 2, 1977; volume 3, 1986.

183 Peter Henrici, "De Branges' proof of the Bieberbach conjecture: A view from computational analysis", *Sitzungsberichte der Berliner Mathematischen Gesellschaft* (1987), 105–121.

663 **184** Charles Hermite, letter to C. W. Borchardt (8 September 1875), in *Journal für die reine und angewandte Mathematik* **81** (1876), 93–95. Воспроизведено в его *Œuvres*, volume 3, 211–214.

185 Charles Hermite, *Cours de M. Hermite.* Faculté des Sciences de Paris, 1882. Third edition, 1887; fourth edition, 1891. (Имеется русский перевод четвертого издания: Эрмит Ш. *Курс анализа.*— М.-Л.: ОНТИ, 1936.)

643 **186** Charles Hermite, letter to S. Pincherle (10 May 1900), in *Annali di Matematica pura ed applicata*, series 3, **5** (1901), 57–60. Воспроизведено в его *Œuvres*, volume 4, 529–531.

187 I. N. Herstein and I. Kaplansky, *Matters Mathematical.* Harper & Row, 1974. *26*

188 A. P. Hillman and V. E. Hoggatt, Jr., "A proof of Gould's Pascal hexagon conjecture", *Fibonacci Quarterly* **10** (1972), 565–568, 598.

189 C. A. R. Hoare, "Quicksort", *The Computer Journal* **5** (1962), 10–15. *49*

190 L. C. Hsu, "Note on a combinatorial algebraic identity and its application", *Fibonacci Quarterly* **11** (1973), 480–484.

191 Kenneth E. Iverson, *A Programming Language.* Wiley, 1962. *45, 95*

192 C. G. J. Jacobi, *Fundamenta nova theoriæ functionum ellipticarum.* Königsberg, Bornträger, 1829. Воспроизведено в его *Gesammelte Werke*, volume 1, 49–239. *91*

193 Svante Janson, Donald E. Knuth, Tomasz Łuczak, and Boris Pittel, "The birth of the giant component", *Random Structures & Algorithms* **4** (1993), 233–358. Опубликовано с исправлениями в Donald E. Knuth *Selected Papers on Discrete Mathematics*, 643–792. *255*

194 Dov Jarden and Theodor Motzkin, "The product of sequences with a common linear recursion formula of order 2", *Riveon Lematematika* **3** (1949), 25–27, 38 (На иврите с английской аннотацией). Английская версия опубликована в Dov Jarden, *Recurring Sequences*, Jerusalem, 1958, 42–45; second edition, 1966, 30–33. *664*

195 Arne Jonassen and Donald E. Knuth, "A trivial algorithm whose analysis isn't", *Journal of Computer and System Sciences* **16** (1978), 301–322. Опубликовано с дополнениями в Donald E. Knuth *Selected Papers on Analysis of Algorithms*, 257–282. *640*

196 Bush Jones, "Note on internal merging", *Software — Practice and Experience* **2** (1972), 241–243. *225*

197 Flavius Josephus, *ΙΣΤΟΡΙΑ ΙΟΥΔΑΪΚΟΥ ΠΟΛΕΜΟΥ ΠΡΟΣ ΡΩΜΑΙΟΥΣ.* Английский перевод: *History of the Jewish War against the Romans*, by H. St. J. Thackeray, in the Loeb Classical Library edition of Josephus's works, volumes 2 and 3, Heinemann, London, 1927–1928. ("Задача Иосифа" взята, возможно, из более ранней рукописи, сохранившейся только в древнерусском варианте; см. том 2, с. xi, и том 3, с. 654.) (См. "История иудейской войны Иосифа Флавия в древнерусском переводе Н. А. Мещерского". — М.: Изд. АН СССР, 1958.) *26*

198 R. Jungen, “Sur les séries de Taylor n'ayant que des singularités algébrico-logarithmiques sur leur cercle de convergence”, *Commentarii Mathematici Helvetici* **3** (1931), 266–306.

317 **199** J. Karamata, “Théorèmes sur la sommabilité exponentielle et d'autres sommabilités rattachant”, *Mathematica* (Cluj) **9** (1935), 164–178.

200 I. Kaucký, “Problem E 2257: A harmonic identity”, *American Mathematical Monthly* **78** (1971), 908.

715 **201** J. B. Keiper, “Power series expansions of Riemann's ξ function”, *Mathematics of Computation* **58** (1992), 765–773.

356 **202** Johannes Kepler, letter to Joachim Tancke (12 May 1608), в его *Gesammelte Werke*, volume 16, 154–165.

203 Murray S. Klamkin, *International Mathematical Olympiads, 1978–1985, and Forty Supplementary Problems.* Mathematical Association of America, 1986.

255 **204** R. Arthur Knoebel, “Exponentials reiterated”, *American Mathematical Monthly* **88** (1981), 235–252.

205 Konrad Knopp, *Theorie und Anwendung der unendlichen Reihen.* Julius Springer, Berlin, 1922; second edition, 1924. Воспроизведено Dover, 1945. Fourth edition, 1947; fifth edition, 1964. Английский перевод: *Theory and Application of Infinite Series*, 1928; second edition, 1951.

661 **206** Donald Knuth, “Transcendental numbers based on the Fibonacci sequence”, *Fibonacci Quarterly* **2** (1964), 43–44, 52.

9, 599, 614, 634, 683 **207** Donald E. Knuth, *The Art of Computer Programming*, volume 1: *Fundamental Algorithms.* Addison-Wesley, 1968; third edition, 1997. (Имеется русский перевод: Кнут Д. Э. *Искусство программирования, том 1. Основные алгоритмы, 3-е изд.* — М.: Изд. дом “Вильямс”, 2000.)

147, 169, 599 **208** Donald E. Knuth, *The Art of Computer Programming*, volume 2: *Seminumerical Algorithms.* Addison-Wesley, 1969; third edition, 1997. (Имеется русский перевод: Кнут Д. Э. *Искусство программирования, том 2. Получисленные алгоритмы, 3-е изд.* — М.: Изд. дом “Вильямс”, 2000.)

328, 495, 600 **209** Donald E. Knuth, *The Art of Computer Programming*, volume 3: *Sorting and Searching.* Addison-Wesley, 1973; second edition, 1998. (Имеется русский перевод: Кнут Д. Э. *Искусство программирования, том 3. Сортировка и поиск, 2-е изд.* — М.: Изд. дом “Вильямс”, 2000.)

210 Donald E. Knuth, "Problem E 2492: Some sum", *American Mathematical Monthly* **82** (1975), 855.

211 Donald E. Knuth, *Mariages stables et leurs relations avec d'autres problèmes combinatoires.* Les Presses de l'Université de Montréal, 1976. Переработанное и исправленное издание, 1980. Английский перевод: *Stable Marriage and its Relation to Other Combinatorial Problems*, 1997.

212 Donald E. Knuth, *The TEXbook.* Addison-Wesley, 1984. Воспроизведено как volume A of *Computers & Typesetting*, 1986. (Имеется русский перевод: Кнут Д. Э. *Все про TEX.* — Протвино: АО RDTEX, 1993.)

213 Donald E. Knuth, "An analysis of optimum caching", *Journal of Algorithms* **6** (1985), 181–199. Воспроизведено с дополнениями в его *Selected Papers on Analysis of Algorithms*, 235–255. *672*

214 Donald E. Knuth, *Computers & Typesetting*, volume D: *METAFONT: The Program.* Addison-Wesley, 1986.

215 Donald E. Knuth, "Problem 1280: Floor function identity", *Mathematics Magazine* **61** (1988), 319–320.

216 Donald E. Knuth, "Problem E 3106: A new sum for n^2", *American Mathematical Monthly* **94** (1987), 795–797.

217 Donald E. Knuth, "Fibonacci multiplication", *Applied Mathematics Letters* **1** (1988), 57–60.

218 Donald E. Knuth, "A Fibonacci-like sequence of composite numbers", *Mathematics Magazine* **63** (1990), 21–25. *670*

219 Donald E. Knuth, "Problem E 3309: A binomial coefficient inequality", *American Mathematical Monthly* **97** (1990), 614.

220 Donald E. Knuth, "Two notes on notation", *American Mathematical Monthly* **99** (1992), 403–422. Воспроизведено с дополнениями в его *Selected Papers on Discrete Mathematics*, 15–44. *45, 210, 328, 712*

221 Donald E. Knuth, "Convolution polynomials", *The Mathematica Journal* **2**,4 (Fall 1992), 67–78. Воспроизведено с дополнениями в его *Selected Papers on Discrete Mathematics*, 225–256. *328, 675*

222 Donald E. Knuth, "Johann Faulhaber and sums of powers", *Mathematics of Computation* **61** (1993), 277–294. Воспроизведено с дополнениями в его *Selected Papers on Discrete Mathematics*, 61–84. *352*

223 Donald E. Knuth, "Bracket notation for the coefficient-of operator", in *A Classical Mind*, essays in honour of C. A. R. Hoare, *249* edited by A. W. Roscoe, Prentice-Hall, 1994, 247–258. Воспроизведено с дополнениями в его *Selected Papers on Discrete Mathematics*, 45–59.

663 **224** Donald E. Knuth and Thomas J. Buckholtz, "Computation of Tangent, Euler, and Bernoulli numbers", Числа Эйлера *Mathematics of Computation* **21** (1967), 663–688.

225 Donald E. Knuth and Ilan Vardi, "Problem 6581: The asymptotic expansion of the middle binomial coefficient", *American Mathematical Monthly* **97** (1990), 626–630.

633 **226** Donald E. Knuth and Herbert S. Wilf, "The power of a prime that divides a generalized binomial coefficient", *Journal für die reine und angewandte Mathematik* **396** (1989), 212–219. Воспроизведено в Knuth D. E. *Selected Papers on Discrete Mathematics*, 511–524.

12 **227** Donald E. Knuth and Hermann Zapf, "AMS Euler — A new typeface for Mathematics", *Scholarly Publishing* **20** (1989), 131–157. Воспроизведено в Knuth D. E. *Digital Typography*, 339–365.

149 **228** C. Kramp, *Élémens d'arithmétique universelle.* Cologne, 1808.

267 **229** E. E. Kummer, "Über die hypergeometrische Reihe

$$1 + \frac{\alpha\beta}{1.\gamma}x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1.2.\gamma(\gamma+1)}xx + \\ + \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1.2.3.\gamma(\gamma+1)(\gamma+2)}x^3 + \dots\text{"},$$

Journal für die reine und angewandte Mathematik **15** (1836), 39–83, 127–172. Воспроизведено в его *Collected Papers*, volume 2, 75–166.

230 E. E. Kummer, "Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen", *Journal für die reine und angewandte Mathematik* **44** (1852), 93–146. Воспроизведено в его *Collected Papers*, volume 1, 485–538.

600 **231** R. P. Kurshan and B. Gopinath, "Recursively generated periodic sequences", *Canadian Journal of Mathematics* **26** (1974), 1356–1371.

370 **232** Thomas Fantet de Lagny, *Analyse générale ou Méthodes nouvelles pour résoudre les problèmes de tous les genres et de tous les degrés à l'infini.* Опубликовано в volume 11 of *Mémoires de l'Académie Royale des Sciences*, Paris, 1733.

233 de la Grange (Lagrange), “Démonstration d’un théorème nouveau concernant les nombres premiers”, *Nouveaux Mémoires de l’Académie royale des Sciences et Belles-Lettres*, Berlin (1771), 125–137. Воспроизведено в его *Œuvres*, volume 3, 425–438.

234 de la Grange (Lagrange), “Sur une nouvelle espèce de calcul rélatif à la différentiation & à l’intégration des quantités variables”, *Nouveaux Mémoires de l’Académie royale des Sciences et Belles-Lettres*, Berlin (1772), 185–221. Воспроизведено в его *Œuvres*, volume 3, 441–476. *565*

235 I. Lah, “Eine neue Art von Zahlen, ihre Eigenschaften und Anwendung in der Mathematischen Statistik”, *Mitteilungsblatt für Mathematische Statistik* **7** (1955), 203–212. (Более общая формула опубликована в L. Toscano, *Commentationes* **3** (Vatican City: Accademia della Scienze, 1939), 721–757, Equations 17 and 117.)

236 I. H. Lambert, “Observationes variæ in Mathesin puram”, *Acta Helvetica* **3** (1758), 128–168. Воспроизведено в его *Opera Mathematica*, volume 1, 16–51. *253*

237 Lambert, “Observations analytiques”, *Nouveaux Mémoires de l’Académie royale des Sciences et Belles-Lettres*, Berlin (1770), 225–244. Воспроизведено в его *Opera Mathematica*, volume 2, 270–290. *253*

238 Edmund Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, two volumes. Teubner, Leipzig, 1909. *540*

239 Edmund Landau, *Vorlesungen über Zahlentheorie*, three volumes. Hirzel, Leipzig, 1927.

240 P. S. de la Place (Laplace), “Mémoire sur les approximations des Formules qui sont fonctions de très-grands nombres”, *Mémoires de l’Academie royale des Sciences de Paris* (1782), 1–88. Воспроизведено в его *Œuvres Complètes* **10**, 207–291. *560*

241 Adrien-Marie Legendre, *Essai sur la Théorie des Nombres*. Paris, 1798; second edition, 1808. Third edition (retitled *Théorie des Nombres*, in two volumes), 1830; fourth edition, Blanchard, 1955.

242 D. H. Lehmer, “Tests for primality by the converse of Fermat’s theorem”, *Bulletin of the American Mathematical Society*, series 2, **33** (1927), 327–340. Воспроизведено в его *Selected Papers*, volume 1, 69–82.

243 D. H. Lehmer, “On Stern’s diatomic series”, *American Mathematical Monthly* **36** (1929), 59–67.

629 **244** D. H. Lehmer, "On Euler's totient function", *Bulletin of the American Mathematical Society*, series 2, **38** (1932), 745–751. Воспроизведено в его *Selected Papers*, volume 1, 319–325.

217 **245** G. W. Leibniz, letter to Johann Bernoulli (May 1695), in *Leibnizens Mathematische Schriften*, volume 3, 174–179.

360 **246** C. G. Lekkerkerker, "Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci", *Simon Stevin* **29** (1952), 190–195.

215 **247** Tamás Lengyel, "A combinatorial identity and the world series", *SIAM Review* **35** (1993), 294–297.

248 Tamás Lengyel, "On some properties of the series $\sum_{k=0}^{\infty} k^n x^k$ and the Stirling numbers of the second kind", *Discrete Mathematics* **150** (1996), 281–292.

330 **249** Li Shan-Lan, *Duò Jī Bǐ Lèi* (Sums of Piles Obtained Inductively). В его *Zégǔxī Zhaī Suànxué* (Classically Inspired Meditations on Mathematics), Nanjing, 1867.

250 Elliott H. Lieb, "Residual entropy of square ice", *Physical Review* **162** (1967), 162–172.

178 **251** J. Liouville, "Sur l'expression $\varphi(n)$, qui marque combien la suite $1, 2, 3, \ldots, n$ contient de nombres premiers à n", *Journal de Mathématiques pures et appliquées*, series 2, **2** (1857), 110–112.

252 B. F. Logan, "The recovery of orthogonal polynomials from a sum of squares", *SIAM Journal on Mathematical Analysis* **21** (1990), 1031–1050.

253 B. F. Logan, "Polynomials related to the Stirling numbers", AT&T Bell Laboratories internal technical memorandum, August 10, 1987.

254 Calvin T. Long and Verner E. Hoggatt, Jr., "Sets of binomial coefficients with equal products", *Fibonacci Quarterly* **12** (1974), 71–79.

628 **255** Shituo Lou and Qi Yao, "A Chebychev's type of prime number theorem in a short interval-II", *Hardy-Ramanujan Journal* **15** (1992), 1–33.

668 **256** Sam Loyd, *Cyclopedia of Puzzles*. Franklin Bigelow Corporation, Morningside Press, New York, 1914.

257 E. Lucas, "Sur les rapports qui existent entre la théorie des nombres et le Calcul intégral", *Comptes Rendus hebdomadaires des séances de l'Académie des Sciences* (Paris) **82** (1876), 1303–1305.

258 Édouard Lucas, “Sur les congruences des nombres eulériens et des coefficients différentiels des fonctions trigonométriques, suivant un module premier”, *Bulletin de la Société mathématique de France* **6** (1877), 49–54.

259 Edouard Lucas, *Théorie des Nombres*, volume 1. Paris, 1891. *356*

260 Édouard Lucas, *Récréations mathématiques*, four volumes. Gauthier-Villars, Paris, 1891–1894. Воспроизведено Albert Blanchard, Paris, 1960. (Ханойская башня обсуждается в томе 3, с. 55–59.) *17*

261 R. C. Lyness, “Cycles”, *The Mathematical Gazette* **29** (1945), 231–233. *600*

262 R. C. Lyness, “Cycles”, *The Mathematical Gazette* **45** (1961), 207–209. *600*

263 Colin MacLaurin, *Collected Letters*, edited by Stella Mills. Shiva Publishing, Nantwich, Cheshire, 1982. *564*

264 P. A. MacMahon, “Application of a theory of permutations in circular procession to the theory of numbers”, *Proceedings of the London Mathematical Society* **23** (1892), 305–313. *182*

265 J.-C. Martzloff, *Histoire des Mathématiques Chinoises.* Paris, 1988. Английский перевод: *A History of Chinese Mathematics*, Springer-Verlag, 1997. *330*

266 Матиясевич Ю. В. “Диофантовость перечислимых множеств”, *Доклады АН СССР* **191** (1970), 279–282. Английский перевод с комментариями автора: “Enumerable sets are diophantine”, *Soviet Mathematics — Doklady* **11** (1970), 354–357. *359*

267 Z. A. Melzak, *Companion to Concrete Mathematics.* Volume 1, *Mathematical Techniques and Various Applications*, Wiley, 1973; volume 2, *Mathematical Ideas, Modeling & Applications*, Wiley, 1976. *8*

268 N. S. Mendelsohn, “Problem E 2227: Divisors of binomial coefficients”, *American Mathematical Monthly* **78** (1971), 201.

269 Marini Mersenni, *Cogitata Physico-Mathematica.* Paris, 1644. *146*

270 F. Mertens, “Ueber einige asymptotische Gesetze der Zahlentheorie”, *Journal für die reine und angewandte Mathematik* **77** (1874), 289–338. *181*

271 Mertens, “Ein Beitrag zur analytischen Zahlentheorie”, *Journal für die reine und angewandte Mathematik* **78** (1874), 46–62. *43*

272 W. H. Mills, "A prime representing function", *Bulletin of the American Mathematical Society*, series 2, **53** (1947), 604.

181 **273** A. F. Möbius, "Über eine besondere Art von Umkehrung der Reihen", *Journal für die reine und angewandte Mathematik* **9** (1832), 105–123. Воспроизведено в его *Gesammelte Werke*, volume 4, 589–612.

274 A. Moessner, "Eine Bemerkung über die Potenzen der natürlichen Zahlen", *Sitzungsberichte der Mathematisch Naturwissenschaftlichen Klasse der Bayerischen Akademie der Wissenschaften*, 1951, Heft 3, 29.

557 **275** Hugh L Montgomery, "Fluctuations in the mean of Euler's phi function", *Proceedings of the Indian Academy of Sciences*, Mathematical Sciences, **97** (1987), 239–245.

276 Peter L. Montgomery, "Problem E 2686: LCM of binomial coefficients", *American Mathematical Monthly* **86** (1979), 131.

355 **277** Leo Moser, "Problem B-6: Some reflections", *Fibonacci Quarterly* **1**, 4 (1963), 75–76.

672 **278** T. S. Motzkin and E. G. Straus, "Some combinatorial extremum problems", *Proceedings of the American Mathematical Society* **7** (1956), 1014–1021.

279 B. R. Myers, "Problem 5795: The spanning trees of an n-wheel", *American Mathematical Monthly* **79** (1972), 914–915.

339 **280** Isaac Newton, letter to John Collins (18 February 1670), в *The Correspondence of Isaac Newton*, volume 1, 27. Отрывки помещены в *The Mathematical Papers of Isaac Newton*, volume 3, 563.

281 Ivan Niven, *Diophantine Approximations.* Interscience, 1963.

403 **282** Ivan Niven, "Formal power series", *American Mathematical Monthly* **76** (1969), 871–889.

112 **283** Andrew M. Odlyzko and Herbert S. Wilf, "Functional iteration and the Josephus problem", *Glasgow Mathematical Journal* **33** (1991), 235–240.

284 Blaise Pascal, "De numeris multiplicibus", представлено в Парижской академии в 1654 году и опубликовано вместе с его *Traité du triangle arithmétique* [285]. Воспроизведено в *Œuvres de Blaise Pascal*, volume 3, 314–339.

201, 202, 743 **285** Blaise Pascal, "Traité du triangle arithmetique", в его *Traité du Triangle Arithmetique, avec quelques autres petits traitez sur la mesme matiere*, Paris, 1665. Воспроизведено в *Œuvres de*

Blaise Pascal (Hachette, 1904–1914), volume 3, 445–503; латинские издания от 1654 году в volume 11, 366–390.

286 G. P. Patil, "On the evaluation of the negative binomial distribution with examples", *Technometrics* **2** (1960), 501–505.

287 C. S. Peirce, letter to E. S. Holden (January 1901). In *The New Elements of Mathematics*, edited by Carolyn Eisele, Mouton, The Hague, 1976, volume 1, 247–253. (См. также с. 211.)

288 C. S. Peirce, letter to Henry B. Fine (17 July 1903). In *The New Elements of Mathematics*, edited by Carolyn Eisele, Mouton, The Hague, 1976, volume 3, 781–784. (См. также "Ordinals", an unpublished manuscript from circa 1905, в *Collected Papers of Charles Sanders Peirce*, volume 4, 268–280.) *628*

289 Walter Penney, "Problem 95: Penney-Ante", *Journal of Recreational Mathematics* **7** (1974), 321. *492*

290 J. K. Percus, *Combinatorial Methods.* Springer-Verlag, 1971.

291 Marko Petkovšek, "Hypergeometric solutions of linear recurrences with polynomial coefficients", *Journal of Symbolic Computation* **14** (1992), 243–264. *286, 685*

292 J. F. Pfaff, "Observationes analyticæ ad *L. Euleri* institutiones calculi integralis, Vol. IV, Supplem. II & IV", *Nova acta academiæ scientiarum imperialis Petropolitanæ* **11**, Histoire section, 37–57. (В этом томе, опубликованном в 1798 году, большинство работ относятся к 1793 году, хотя статья Пфаффа на самом деле написана в 1797 году.) *260, 269, 272*

293 L. Pochhammer, "Ueber hypergeometrische Functionen n$^{\text{ter}}$ Ordnung", *Journal für die reine und angewandte Mathematik* **71** (1870), 316–352. *72*

294 H. Poincaré, "Sur les fonctions à espaces lacunaires", *American Journal of Mathematics* **14** (1892), 201–221.

295 S. D. Poisson, "Mémoire sur le calcul numérique des intégrales définies", *Mémoires de l'Académie Royale des Sciences de l'Institut de France*, series 2, **6** (1823), 571–602. *566*

296 G. Pólya, "Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen", *Acta Mathematica* **68** (1937), 145–254. Английский перевод с комментариями Рональда Рида (Ronald C. Read): *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*, Springer-Verlag, 1987.

9, 35, 608 **297** George Pólya, *Induction and Analogy in Mathematics.* Princeton University Press, 1954. (Имеются два издания на русском языке: Пойа Д. *Математика и правдоподобные рассуждения.* — М.: ИЛ, 1957; М.: Наука, 1975.)

398 **298** G. Pólya, "On picture-writing", *American Mathematical Monthly* **63** (1956), 689–697.

299 G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, two volumes. Julius Springer, Berlin, 1925; fourth edition, 1970 and 1971. Английский перевод: *Problems and Theorems in Analysis*, 1972 and 1976. (Имеется несколько изданий на русском языке: Полиа Г., Сеге Г. *Задачи и теоремы из анализа*, в 2-х томах. — М.: Наука, 1937–1938; 1956; 1978.)

300 R. Rado, "A note on the Bernoullian numbers", *Journal of the London Mathematical Society* **9** (1934), 88–90.

632 **301** Earl D. Rainville, "The contiguous function relations for ${}_pF_q$ with applications to Bateman's $J_n^{u,v}$ and Rice's $H_n(\zeta, p, v)$", *Bulletin of the American Mathematical Society*, series 2, **51** (1945), 714–723.

434 **302** George N. Raney, "Functional composition patterns and power series reversion", *Transactions of the American Mathematical Society* **94** (1960), 441–451.

303 D. Rameswar Rao, "Problem E 2208: A divisibility problem", *American Mathematical Monthly* **78** (1971), 78–79.

107 **304** John William Strutt, Third Baron Rayleigh, *The Theory of Sound.* First edition, 1877; second edition, 1894. (Цитируемый материал относительно иррационального спектра взят из разд. 92a второго издания.) (Имеется русский перевод: Стретт Дж. В. (лорд Релей). *Теория звука*, в 2-х томах. — М.: Гостехиздат, 1955.)

537 **305** Robert Recorde, *The Whetstone of Witte.* London, 1557.

306 Simeon Reich, "Problem 6056: Truncated exponential-type series", *American Mathematical Monthly* **84** (1977), 494–495.

307 Georges de Rham, "Un peu de mathématiques à propos d'une courbe plane", *Elemente der Mathematik* **2** (1947), 73–76, 89–97. Воспроизведено в его *Œuvres Mathématiques*, 678–689.

662 **308** Paulo Ribenboim, *13 Lectures on Fermat's Last Theorem.* Springer-Verlag, 1979.

309 Bernhard Riemann, “Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe”, Habilitationsschrift, Göttingen, 1854. Опубликовано в *Abhandlungen der Mathematischen Classe der Königlichen Gesellschaft der Wissenschaften zu Göttingen* **13** (1868), 87–132. Воспроизведено в его *Gesammelte Mathematische Werke*, 227–264. (Имеются русские переводы: Бернштейн С. Н. (ред.) *Разложенiе функцiй в тригонометрическiе ряды.* — Харьковъ: Харьковское математическое общество, 1914, 25–90; Риман Б. *Сочинения.* — М.-Л.: ГИТТЛ, 1948, 225–261.)

310 Samuel Roberts, “On the figures formed by the intercepts of a system of straight lines in a plane, and on analogous relations in space of three dimensions”, *Proceedings of the London Mathematical Society* **19** (1889), 405–422.

311 Øystein Rødseth, “Problem E 2273: Telescoping Vandermonde convolutions”, *American Mathematical Monthly* **79** (1972), 88–89.

312 J. Barkley Rosser and Lowell Schoenfeld, “Approximate formulas for some functions of prime numbers”, *Illinois Journal of Mathematics* **6** (1962), 64–94. *148*

313 Gian-Carlo Rota, “On the foundations of combinatorial theory. I. Theory of Möbius functions”, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **2** (1964), 340–368. *617*

314 Ranjan Roy, “Binomial identities and hypergeometric series”, *American Mathematical Monthly* **94** (1987), 36–46.

315 Louis Saalschütz, “Eine Summationsformel”, *Zeitschrift für Mathematik und Physik* **35** (1890), 186–188. *269*

316 Салтыков А. И. “О функции Эйлера”. *Вестник Московского государственного университета*, серия 1, Математика, механика (1960), №6, 34–50. *557*

317 A. Sárközy, “On divisors of binomial coefficients, I”, *Journal of Number Theory* **20** (1985), 70–80. *655*

318 W. W. Sawyer, *Prelude to Mathematics.* Baltimore, Penguin, 1955. (Имеются два издания на русском языке: Сойер У. У. *Прелюдия к математике.* — М.: Просвещение, 1965; 1972.) *260*

319 O. Schlömilch, “Ein geometrisches Paradoxon”, *Zeitschrift für Mathematik und Physik* **13** (1868), 162. *357*

320 Ernst Schröder, “Vier combinatorische Probleme”, *Zeitschrift für Mathematik und Physik* **15** (1870), 361–376.

321 Heinrich Schröter, "Ableitung der Partialbruch- und Produkt-Entwickelungen für die trigonometrischen Funktionen", *Zeitschrift für Mathematik und Physik* **13** (1868), 254–259.

322 R. S. Scorer, P. M. Grundy, and C. A. B. Smith, "Some binary games", *The Mathematical Gazette* **28** (1944), 96–103.

323 J. Sedláček, "On the skeletons of a graph or digraph", in *Combinatorial Structures and their Applications*, Gordon and Breach, 1970, 387–391. (В этом томе представлены труды конференции Calgary International Conference on Combinatorial Structures and their Applications, 1969.)

324 J. O. Shallit, "Problem 6450: Two series", *American Mathematical Monthly* **92** (1985), 513–514.

334 **325** R. T. Sharp, "Problem 52: Overhanging dominoes", *Pi Mu Epsilon Journal* **1**, 10 (1954), 411–412.

119 **326** W. Sierpiński, "Sur la valeur asymptotique d'une certaine somme", *Bulletin International de l'Académie Polonaise des Sciences et des Lettres* (Cracovie), series A (1910), 9–11.

327 W. Sierpiński, "Sur les nombres dont la somme de diviseurs est une puissance du nombre 2", *Calcutta Mathematical Society Golden Jubilee Commemorative volume* (1958–1959), part 1, 7–9.

328 Wacław Sierpiński, *A Selection of Problems in the Theory of Numbers*. Macmillan, 1964.

329 David L. Silverman, "Problematical Recreations 447: Numerical links", *Aviation Week & Space Technology* **89**, 10 (1 September 1968), 71. Воспроизведено как Problem 147 в *Second Book of Mathematical Bafflers*, edited by Angela Fox Dunn, Dover, 1983.

65, 414, 558 **330** N. J. A. Sloane, *A Handbook of Integer Sequences*. Academic Press, 1973. Sequel, with Simon Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, 1995. `http://www.research.att.com/~njas/sequences`.

492 **331** Соловьев А. Д. "Одно комбинаторное тождество и его применение к задаче о первом наступлении редкого события", *Теория вероятностей и ее применения*, **11** (1966), 313–320. Английский перевод: "A combinatorial identity and its application to the problem concerning the first occurrence of a rare event", *Theory of Probability and Its Applications* **11** (1966), 276–282.

7 **332** William G. Spohn, Jr., "Can Mathematics be saved?" *Notices of the American Mathematical Society* **16** (1969), 890–894.

333 Richard P. Stanley, "Differentiably finite power series", *European Journal of Combinatorics* **1** (1980), 175–188.

334 Richard P. Stanley, "On dimer coverings of rectangles of fixed width", *Discrete Applied Mathematics* **12** (1985), 81–87.

335 Richard P. Stanley, *Enumerative Combinatorics*, volume 1. Wadsworth & Brooks/Cole, 1986. (Имеется русский перевод: Стенли Р. *Перечислительная комбинаторика.* — М.: Мир, 1990.) *638*

336 K. G. C. von Staudt, "Beweis eines Lehrsatzes, die Bernoullischen Zahlen betreffend", *Journal für die reine und angewandte Mathematik* **21** (1840), 372–374.

337 Guy L. Steele Jr., Donald R. Woods, Raphael A. Finkel, Mark R. Crispin, Richard M. Stallman, and Geoffrey S. Goodfellow, *The Hacker's Dictionary: A Guide to the World of Computer Wizards.* Harper & Row, 1983. *164*

338 J. Steiner, "Einige Gesetze über die Theilung der Ebene und des Raumes", *Journal für die reine und angewandte Mathematik* **1** (1826), 349–364. Воспроизведено в его *Gesammelte Werke*, volume 1, 77–94. *22*

339 M. A. Stern, "Ueber eine zahlentheoretische Funktion", *Journal für die reine und angewandte Mathematik* **55** (1858), 193–220. *155*

340 L. Stickelberger, "Ueber eine Verallgemeinerung der Kreistheilung", *Mathematische Annalen* **37** (1890), 321–367.

341 T. J. Stieltjes, letters to Hermite (June 1885), in *Correspondance d'Hermite et de Stieltjes*, volume 1, 146–159. *715*

342 T. J. Stieltjes, "Table des valeurs des sommes $S_k = \sum_1^\infty n^{-k}$", *Acta Mathematica* **10** (1887), 299–302. Воспроизведено в его *Œuvres Complètes*, volume 2, 100–103.

343 James Stirling, *Methodus Differentialis.* London, 1730. Английский перевод: *The Differential Method*, 1749. *244, 318, 362*

344 Volker Strehl, "Binomial identities — combinatorial and algorithmic aspects", *Discrete Mathematics* **136** (1994), 309–346. *655*

345 Dura W. Sweeney, "On the computation of Euler's constant", *Mathematics of Computation* **17** (1963), 170–178. *577*

346 J. J. Sylvester, "Problem 6919", *Mathematical Questions with their Solutions from the 'Educational Times'* **37** (1882), 42–43, 80.

175 **347** J. J. Sylvester, "On the number of fractions contained in any 'Farey series' of which the limiting number is given", *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, series 5, **15** (1883), 251–257. Воспроизведено в его *Collected Mathematical Papers*, volume 4, 101–109.

628 **348** M. Szegedy, "The solution of Graham's greatest common divisor problem", *Combinatorica* **6** (1986), 67–71.

349 S. Tanny, "A probabilistic interpretation of Eulerian numbers", *Duke Mathematical Journal* **40** (1973), 717–722.

350 L. Theisinger, "Bemerkung über die harmonische Reihe", *Monatshefte für Mathematik und Physik* **26** (1915), 132–134.

480, 481 **351** T. N. Thiele, *The Theory of Observations.* Charles & Edwin Layton, London, 1903. Воспроизведено в *The Annals of Mathematical Statistics* **2** (1931), 165–308.

352 E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function.* Clarendon Press, Oxford, 1951; second edition, revised by D. R. Heath-Brown, 1986. (Имеется русский перевод: Титчмарш Е. К. *Теория дзета-функции Римана.* — М.: ИЛ, 1953.)

353 F. G. Tricomi and A. Erdélyi, "The asymptotic expansion of a ratio of gamma functions", *Pacific Journal of Mathematics* **1** (1951), 133–142.

343 **354** Peter Ungar, "Problem E 3052: A sum involving Stirling numbers", *American Mathematical Monthly* **94** (1987), 185–186.

355 J. V. Uspensky, "On a problem arising out of the theory of a certain game", *American Mathematical Monthly* **34** (1927), 516–521.

296 **356** Alfred van der Poorten, "A proof that Euler missed ... Apéry's proof of the irrationality of $\zeta(3)$, an informal report", *The Mathematical Intelligencer* **1** (1979), 195–203.

218 **357** A. Vandermonde, "Mémoire sur des irrationnelles de différens ordres avec une application au cercle", *Mémoires de Mathématique et de Physique, tirés des registres de l'Académie Royale des Sciences* (1772), part 1, 489–498.

358 Ilan Vardi, "The error term in Golomb's sequence", *Journal of Number Theory* **40** (1992), 1–11.

596 **359** J. Venn, "On the diagrammatic and mechanical representation of propositions and reasonings", *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, series 5, **9** (1880), 1–18.

360 John Wallis, *A Treatise of Angular Sections.* Oxford, 1684.

361 Edward Waring, *Meditationes Algebraïcæ.* Cambridge, 1770; third edition, 1782.

362 William C. Waterhouse, "Problem E 3117: Even odder than we thought", *American Mathematical Monthly* **94** (1987), 691–692.

363 Frederick V. Waugh and Margaret W. Maxfield, "Side-and-diagonal numbers", *Mathematics Magazine* **40** (1967), 74–83.

364 Warren Weaver, "Lewis Carroll and a geometrical paradox", *American Mathematical Monthly* **45** (1938), 234–236. *357*

365 H. Weber, "Leopold Kronecker", *Jahresbericht der Deutschen Mathematiker-Vereinigung* **2** (1892), 5–31. Воспроизведено в *Mathematische Annalen* **43** (1893), 1–25. *624*

366 Louis Weisner, "Abstract theory of inversion of finite series", *Transactions of the American Mathematical Society* **38** (1935), 474–484. *617*

367 Edgar M. E. Wermuth, "Die erste Fourierreihe", *Mathematische Semesterberichte* **40** (1993), 133–145. *717*

368 Hermann Weyl, "Über die Gibbs'sche Erscheinung und verwandte Konvergenzphänomene", *Rendiconti del Circolo Matematico di Palermo* **30** (1910), 377–407. *119*

369 F. J. W. Whipple, "Some transformations of generalized hypergeometric series", *Proceedings of the London Mathematical Society*, series 2, **26** (1927), 257–272.

370 Alfred North Whitehead, *An Introduction to Mathematics.* London and New York, 1911. *602*

371 Alfred North Whitehead, "Technical education and its relation to science and literature", chapter 2 in *The Organization of Thought, Educational and Scientific*, London and New York, 1917. Воспроизведено как глава 4 в *The Aims of Education and Other Essays*, New York, 1929. (Имеется русский перевод: Уайтхэдъ А. Н. *Введение въ математику.* — Петроград: Физика, 1916.) *123*

372 Alfred North Whitehead, *Science and the Modern World.* New York, 1925. Глава 2 воспроизведена в *The World of Mathematics*, edited by James R. Newman, 1956, volume 1, 402–416. *718*

373 Herbert S. Wilf, *generatingfunctionology.* Academic Press, 1990; second edition, 1994. *685*

298, 300 **374** Herbert S. Wilf and Doron Zeilberger, "An algorithmic proof theory for hypergeometric (ordinary and 'q') multisum/integral identities", *Inventiones Mathematicae* **108** (1992), 575–633.

375 H. C. Williams and Harvey Dubner, "The primality of R1031", *Mathematics of Computation* **47** (1986), 703–711.

376 J. Wolstenholme, "On certain properties of prime numbers", *Quarterly Journal of Pure and Applied Mathematics* **5** (1862), 35–39.

377 Derick Wood, "The Towers of Brahma and Hanoi revisited", *Journal of Recreational Mathematics* **14** (1981), 17–24.

330 **378** J. Worpitzky, "Studien über die *Bernoulli*schen und *Euler*schen Zahlen", *Journal für die reine und angewandte Mathematik* **94** (1883), 203–232.

379 E. M. Wright, "A prime-representing function", *American Mathematical Monthly* **58** (1951), 616–618; errata in **59** (1952), 99.
collected works, entitled *Hermann Zapf & His Design*

380 *Derek A. Zave, "A series expansion involving the harmonic numbers", Information Processing Letters* **5** *(1976), 75–77.*

360 **381** *E. Zeckendorf, "Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas", Bulletin de la Société Royale des Sciences de Liège* **41** *(1972), 179–182.*

286 **382** *Doron Zeilberger, "Sister Celine's technique and its generalizations", Journal of Mathematical Analysis and Applications* **85** *(1982), 114–145. См. также Sister Mary Celine Fasenmyer, "A note on pure recurrence relations", American Mathematical Monthly* **56** *(1949), 14–17.*

673 **383** *Doron Zeilberger, "A holonomic systems approach to special functions identities", Journal of Computational and Applied Mathematics* **32** *(1990), 321–368.*

286 **384** *Doron Zeilberger, "The method of creative telescoping", Journal of Symbolic Computation* **11** *(1991), 195–204.*

B

Первоисточники упражнений

УПРАЖНЕНИЯ для этой книги собирались из множества источников. Авторы старались проследить происхождение всех задач, которые были опубликованы ранее, за исключением случаев, когда упражнение настолько элементарное, что его автор, вероятно, и не думал о том, чтобы что-то изобрести.

Многие упражнения взяты из стэнфордских экзаменационных задач по курсу конкретной математики. Зачастую новые задачи для этих экзаменов предлагались самими преподавателями и ассистентами. Вот, кстати, отличный повод перечислить их имена.

Год	*Преподаватель*	*Ассистент(ы)*
1970	Дон Кнут (Don Knuth)	Воган Пратт (Vaughan Pratt)
1971	Дон Кнут (Don Knuth)	Лео Гуибас (Leo Guibas)
1973	Дон Кнут (Don Knuth)	Хенсон Грейвс (Henson Graves), Луи Жуайек (Louis Jouaillec)
1974	Дон Кнут (Don Knuth)	Скот Драйсдел (Scot Drysdale), Том Портер (Tom Porter)
1975	Дон Кнут (Don Knuth)	Марк Браун (Mark Brown), Луи Трабб Пардо (Luis Trabb Pardo)
1976	Энди Яо (Andy Yao)	Марк Браун (Mark Brown), Лайл Рэмшоу (Lyle Ramshaw)
1977	Энди Яо (Andy Yao)	Йосси Шилоч (Yossi Shiloach)
1978	Френсис Яо (Frances Yao)	Йосси Шилоч (Yossi Shiloach)
1979	Рон Грэхэм (Ron Graham)	Френк Лианг (Frank Liang), Крис Тонг (Chris Tong), Марк Хайман (Mark Haiman)
1980	Энди Яо (Andy Yao)	Андрей Бродер (Andrei Broder), Джим Мак-Грат (Jim McGrath)

Практические занятия были очень ценными, можно сказать бесценными

Оставьте на следующий год того же преподавателя и тех же ассистентов.

Конспекты очень полезны.

Я никогда не "получал" числа Стирлинга.

1981	Рон Грэхэм (Ron Graham)	Орен Паташник (Oren Patashnik)
1982	Эрнст Мейр (Ernst Mayr)	Джоан Фейгенбаум (Joan Feigenbaum), Дэйв Хелмболд (Dave Helmbold)
1983	Эрнст Мейр (Ernst Mayr)	Анна Карлин (Anna Karlin)
1984	Дон Кнут (Don Knuth)	Орен Паташник (Oren Patashnik), Алекс Шаффер (Alex Schäffer)
1985	Андрей Бродер (Andrei Broder)	Панг Чен (Pang Chen), Стефан Шаркански (Stefan Sharkansky)
1986	Дон Кнут (Don Knuth)	Ариф Мерчант (Arif Merchant), Стефан Шаркански (Stefan Sharkansky)

Кроме того, свой вклад в данный курс внесли Дэвид Кларнер (David Klarner) (1971), Боб Седжевик (Bob Sedgewick) (1974), Лео Гуибас (Leo Guibas) (1975) и Лайл Рамшоу (Lyle Ramshaw) (1979), каждый из которых принял приглашение прочесть шесть или более лекций. Подробные записи лекций, которые каждый год составлялись ассистентами и редактировались преподавателями, послужили основой этой книги.

1.1 Пойа (Pólya) [297, с. 120].
1.2 Скорер (Scorer), Гранди (Grundy) и Смит (Smith) [322].
1.5 Венн (Venn) [359].
1.6 Штайнер (Steiner) [338]; Робертс (Roberts) [310].
1.8 Гаусс (Gauss) [144].
1.9 Коши (Cauchy) [53, пример 2, теорема 17].
1.10 Аткинсон (Atkinson) [15].
1.11 Предложена Вудом (Wood) [377].
1.14 Штайнер (Steiner) [338]; Пойа (Pólya) [297, глава 3]; Брат Альфред (Alfred) [42].
1.17 Дьюдени (Dudeney) [87, головоломка 1].
1.21 Болл (Ball) [20] приписывет задачу Свиндену (B. A. Swinden).
1.22 Основана на идее Питера Шора (Peter Shor).*
1.23 Бьёрн Поонен (Bjorn Poonen).*
1.25 Фрейм (Frame), Стюарт (Stewart) и Данкель (Dunkel) [130].
2.2 Айверсон (Iverson) [191, с. 11].
2.3 [207, упр. 1.2.3–2].
2.4 [207, упр. 1.2.3–25].
2.22 Бине (Binet) [30, §4].
2.23 Выпускной экзамен 1982 г.
2.26 [207, упр. 1.2.3–26].
2.29 Коллоквиум 1979 г.
2.30 Коллоквиум 1973 г.
2.31 Стилтьес (Stieltjes) [342].
2.34 Риман (Riemann) [309, §3].
2.35 Эйлер (Euler) [106] привел ошибочное "доказательство" с использованием расходящихся рядов.
2.36 Голомб (Golomb) [150]; Варди (Vardi) [358].
2.37 Лео Мозер (Leo Moser).*
3.6 Эрнст Мэйр (Ernst Mayr), домашняя работа, 1982.
3.8 Дирихле (Dirichlet) [80].
3.9 Чейс (Chace) [54]; Фибоначчи (Fibonacci) [122, с. 77–83].
3.12 [207, упр. 1.2.4–48(a)].
3.13 Битти (Beatty) [22]; Найвен (Niven) [281, теорема 3.7].
3.19 [207, упр. 1.2.4–34].
3.21 Коллоквиум 1975 г.
3.23 [207, упр. 1.2.4–41].
3.28 Браун (Brown) [45].
3.30 Ахо (Aho) и Слоан (Sloane) [4].
3.31 Грайцер (Greitzer) [165, задача 1972/3, решение 2].
3.32 [160].
3.33 Коллоквиум 1984 г.
3.34 Коллоквиум 1970 г.
3.35 Коллоквиум 1975 г.
3.36 Коллоквиум 1976 г.
3.37 Коллоквиум 1986 г. [215].
3.38 Коллоквиум 1974 г.
3.39 Коллоквиум 1971 г.
3.40 Коллоквиум 1980 г.
3.41 Кламкин (Klamkin) [203, задача 1978/3].
3.42 Успенский (Uspensky) [355].
3.45 Ахо (Aho) и Слоан (Sloane) [4].
3.46 Грэхем (Graham) и Поллак (Pollak) [162].
3.48 Халанд (Håland) и Кнут (Knuth) [170].
3.49 Р. Л. Грэхем (R. L. Graham) и Д. Р. Хофштадтер (D. R. Hofstadter).*
3.52 Френкель (Fraenkel) [128].
3.53 Ш. К. Штайн (S. K. Stein).*
4.4 [214, §526].
4.16 Сильвестр (Sylvester) [346].
4.19 [212, с. 148–149].
4.20 Бертран (Bertrand) [27, с. 129]; Чебышев [56]; Райт (Wright) [379].
4.22 Брилларт (Brillhart) [39]; Вильямс (Williams) и Дюбнер (Dubner) [375]; Дюбнер (Dubner) [86].

4.23 Кроу (Crowe) [68].
4.24 Лежандр (Legendre) [241, 2-е изд., введение].
4.26 [208, упр. 4.5.3–43].
4.31 Паскаль (Pascal) [284].
4.36 Харди (Hardy) и Райт (Wright) [181, §14.5].
4.37 Ахо (Aho) и Слоан (Sloane) [4].
4.38 Люка (Lucas) [257].
4.39 [159].
4.40 Стикельбергер (Stickelberger) [340].
4.41 Лежандр (Legendre) [241, §135]; Харди (Hardy) и Райт (Wright) [181, теорема 82].
4.42 [208, упр. 4.5.1–6].
4.44 [208, упр. 4.5.3–39].
4.45 [208, упр. 4.3.2–13].
4.47 Лемер (Lehmer) [242].
4.48 Гаусс (Gauss) [142, §78]; Крелль (Crelle) [67].
4.52 Коллоквиум 1974 г.
4.53 Коллоквиум 1973 г., идея Рао (Rao) [303].
4.54 Коллоквиум 1974 г.
4.56 Логан (Logan) [252, формула (6.15)].
4.57 Частный случай представлен в [216].
4.58 Серпиньски (Sierpiński) [327].
4.59 Кёртис (Curtiss) [70]; Эрдёш (Erdős) [93].
4.60 Миллз (Mills) [272].
4.61 [207, упр. 1.3.2–19].
4.63 Барлоу (Barlow) [21]; Абель (Abel) [1].
4.64 Пирс (Peirce) [287].
4.66 Рибенбойм (Ribenboim) [308]; Серпиньски (Sierpiński) [328, задача P^2_{10}].
4.67 [157].
4.69 Крамер (Cramér) [66].
4.70 Поль Эрдеш (Paul Erdős).*
4.71 [95, с. 96].
4.72 [95, с. 103].
4.73 Ландау (Landau) [239, том 2, формула 648].
5.1 Форкадель (Forcadel) [126].
5.3 Лонг (Long) и Хоггатт (Hoggatt) [254].
5.5 Курсовой выпускной экзамен 1983 г.
5.13 Коллоквиум 1975 г.
5.14 [207, упр. 1.2.6–20].
5.15 Диксон (Dixon) [81].
5.21 Эйлер (Euler) [99].
5.25 Гаусс [143, §7].
5.28 Эйлер [118].
5.29 Куммер (Kummer) [229, формула 26.4].
5.31 Госпер (Gosper) [154].
5.34 Бейли (Bailey) [18, §10.4].
5.36 Куммер (Kummer) [230, p. 116].
5.37 Вандермонд (Vandermonde) [357].
5.38 [207, упр. 1.2.6–56].
5.40 Рёдсет (Rødseth) [311].
5.43 Пфафф (Pfaff) [292]; [207, упр. 1.2.6–31].
5.48 Раньян Рой (Ranjan Roy).*
5.49 Рой (Roy) [314, формула 3.13].
5.53 Гаусс (Gauss) [143]; Ричард Аски (Richard Askey).*
5.58 Фрейзер (Frazer) и Мак-Келлар (McKellar) [133].
5.59 Станфордский общеобразовательный экзамен по информатике (зимний семестр 1987 г.)
5.60 [207, упр. 1.2.6–41].
5.61 Люка (Lucas) [258].
5.62 Коллоквиум 1971 г.
5.63 Коллоквиум 1974 г.
5.64 Коллоквиум 1980 г.
5.65 Коллоквиум 1983 г.
5.66 Коллоквиум 1984 г.
5.67 Коллоквиум 1976 г.
5.68 Коллоквиум 1985 г.

5.69 Лайл Рэмшоу (Lyle Ramshaw), лекция по приглашению в 1986 г.
5.70 Эндрюс (Andrews) [9, теорема 5.4].
5.71 Вильф (Wilf) [373, упр. 4.16].
5.72 Эрмит (Hermite) [185].
5.74 Коллоквиум 1979 г.
5.75 Коллоквиум 1971 г.
5.76 [207, упр. 1.2.6–59 (исправленное)].
5.77 Коллоквиум 1986 г.
5.78 [210].
5.79 Мендельсон (Mendelsohn) [268]; Монтгомери (Montgomery) [276].
5.81 Выпускной экзамен 1986 г. [219].
5.82 Хиллман (Hillman) и Хоггатт (Hoggatt) [188].
5.85 Су (Hsu) [190].
5.86 Гуд (Good) [153].
5.88 Эрмит (Hermite) [186].
5.91 Уиппл (Whipple) [369].
5.92 Клаузен (Clausen) [60, 61].
5.93 Госпер (Gosper) [154].
5.95 Петковшек (Petkovšek) [291, следствие 3.1].
5.96 Петковшек (Petkovšek) [291, следствие 5.1].
5.98 Ира Гессель (Ira Gessel).*
5.102 Г. С. Вильф (H. S. Wilf).*
5.104 Фолкер Штрел (Volker Strehl).*
5.105 Хенричи (Henrici) [183, стр. 118].
5.108 Апери (Apéry) [14].
5.109 Гессель (Gessel) [146].
5.110 Р. У. Госпер, мл. (R. William Gosper, Jr.)*
5.111 [95, с. 71].
5.112 [95, с. 71].
5.113 Вильф (Wilf) и Зайльбергер (Zeilberger) [374].
5.114 Штрел (Strehl) [344] указывает А. Шмидта (A. Schmidt).
6.6 Фибоначчи (Fibonacci) [122, стр. 283].
6.15 [209, упр. 5.1.3–2].
6.21 Тейзингер (Theisinger) [350].
6.25 Гарднер (Gardner) [138] приписывает авторство Денису Уилкину (Denys Wilquin).
6.27 Люка (Lucas) [257].
6.28 Люка (Lucas) [259, глава 18].
6.31 Лах (Lah) [235]; Р. У. Флойд (R. W. Floyd).*
6.35 Коллоквиум 1977 г.
6.37 Шаллит (Shallit) [324].
6.39 [207, упр. 1.2.7–15].
6.40 Кламкин (Klamkin) [203, задача 1979/1].
6.41 Коллоквиум 1973 г.
6.43 Брук (Brooke) и Уолл (Wall) [41].
6.44 Матиясевич [266].
6.46 Франческа Francesca, Piero della (Francesca) [131]; Уоллис (Wallis) [360, глава 4].
6.47 Люка (Lucas) [257].
6.48 [208, упр. 4.5.3–9(c)].
6.49 Дэйвисон (Davison) [73].
6.50 Коллоквиум 1985 г.; Рам (Rham) [307]; Дейкстра (Dijkstra) [79, с. 230–232].
6.51 Уоринг (Waring) [361]; Лагранж (Lagrange) [233]; Вольштенхольм (Wolstenholme) [376].
6.52 Эсваратасан и Левайн (Levine) [97].
6.53 Кауцки (Kaucký) [200] рассматривает частный случай.
6.54 Штаудт (Staudt) [336]; Клаузен (Clausen) [62]; Радо (Rado) [300].
6.55 Эндрюс (Andrews) и Учимура (Uchimura) [13].

6.56 Коллоквиум 1986 г.
6.57 Коллоквиум 1984 г., предложено Р. У. Флойдом (R. W. Floyd).*
6.58 [207, упр. 1.2.8–30]; коллоквиум 1982 г.
6.59 Барр (Burr) [47].
6.61 Выпускной экзамен 1976 г.
6.62 Д. Боруэн (Borwein, J.) и П. Боруэн (P. Borwein) [36, §3.7].
6.63 [207, раздел 1.2.10]; Стенли (Stanley) [335, предложение 1.3.12].
6.65 Танни (Tanny) [349].
6.66 [209, упр. 5.1.3–3].
6.67 Чанг (Chung) и Грэхем (Graham) [59].
6.68 Логан (Logan) [253].
6.69 [209, упр. 6.1–13].
6.72 Эйлер (Euler) [110, часть 2, глава 8].
6.73 Эйлер (Euler) [108, главы 9 и 10]; Шрётер (Schröter) [321].
6.75 Аткинсон (Atkinson) [16].
6.76 [209, ответ 5.1.3–3]; Ленжиель (Lengyel) [248].
6.78 Логан (Logan) [253].
6.79 *Boston Herald*, 21 августа 1904 г., раздел юмора
6.80 Сильверман (Silverman) и Данн (Dunn) [329].
6.82 [217].
6.83 [156], по модулю ошибки вычисления.
6.85 Барр (Burr) [47].
6.86 [226].
6.87 [208, упр. 4.5.3–2 и 3].
6.88 Адамс (Adams) и Дэйвисон (Davison) [3].
6.90 Лемер (Lehmer) [243].
6.92 Часть (a) из работы Эсваратасана (Eswarathasan) и Левайна (Levine) [97].
7.2 [207, упр. 1.2.9–1].
7.8 Зейв (Zave) [380].
7.9 [207, упр. 1.2.7–22].
7.11 Выпускной экзамен 1971 г.
7.12 [209, с. 63–64].
7.13 Рени (Raney) [302].
7.15 Белл (Bell) [24].
7.16 Пойа (Pólya) [296, с. 149; 207, упр. 2.3.4.4–1].
7.19 [221].
7.20 Юнген (Jungen) [198, с. 299] указывает в качестве автора А. Гурвица (A. Hurwitz).
7.22 Пойа (Pólya) [298].
7.23 Домашняя работа 1983 г.
7.24 Майерс (Myers) [279]; Седлачек (Sedláček) [323].
7.25 [208, доказательство Карлитца (Carlitz) леммы 3.3.3B].
7.26 [207, упр. 1.2.8–12].
7.32 [95, с. 25–26] указывает в качестве авторов Мирского (L. Mirsky) и Ньюмена (M. Newman).
7.33 Выпускной экзамен 1971 г.
7.34 Томас Федер (Tomás Feder).*
7.36 Выпускной экзамен 1974 г.
7.37 Эйлер (Euler) [109, §50]; выпускной экзамен 1971 г.
7.38 Карлитц (Carlitz) [49].
7.39 [207, упр. 1.2.9–18].
7.41 Андрэ (André) [8]; [209, упр. 5.1.4–22].
7.42 Выпускной экзамен 1974 г.
7.44 Гросс (Gross) [166]; [209, упр. 5.3.1–3].
7.45 де Брейн (de Bruijn) [75].
7.47 Во (Waugh) и Максфилд (Maxfield) [363].
7.48 Выпускной экзамен 1984 г.
7.49 Ватерхаус (Waterhouse) [362].
7.50 Шрёдер (Schröder) [320; 207, упр. 2.3.4.4–31].

7.51 Фишер (Fisher) [124]; Перкус (Percus) [290, с. 89–123]; Стенли (Stanley) [334].

7.52 Хаммерсли (Hammersley) [177].

7.53 Эйлер (Euler) [114, часть 2, раздел 2, глава 6, §91].

7.54 Месснер (Moessner) [274].

7.55 Стенли (Stanley) [333].

7.56 Эйлер (Euler) [113].

7.57 В [95, с. 48] имеется ссылка на П. Эрдёша (Erdős) и П. Турана (Turán).

8.13 Томас М. Ковер (Thomas M. Cover).*

8.15 [207, упр. 1.2.10–17].

8.17 Патил (Patil) [286].

8.24 Джон Кнут (John Knuth) (в возрасте 4 лет) и Д.Э.К.; выпускной экзамен 1975 г.

8.26 [207, упр. 1.3.3–18].

8.27 Фишер (Fisher) [125].

8.29 Гуибас (Guibas) и Одлыжко (Odlyzko) [168].

8.32 Выпускной экзамен 1977 г.

8.34 Харди (Hardy) [180] ошибся в ходе анализа этой задачи и пришел к противоположным выводам.

8.35 Выпускной экзамен 1981 г.

8.36 Гарднер (Gardner) [139] указывает в качестве автора Джорджа Сичермана (George Sicherman).

8.38 [208, упр. 3.3.2–10].

8.39 [211, упр. 4.3(a)].

8.41 Феллер (Feller) [120, упр. IX.33].

8.43 [207, разделы 1.2.10 и 1.3.3].

8.44 Выпускной экзамен 1984 г.

8.45 Выпускной экзамен 1985 г.

8.46 Феллер (Feller) [120] указывает в качестве автора Гуго Штейнгауза (Hugo Steinhaus).

8.47 Выпускной экзамен 1974 г.; задача возникла в ходе анализа 2-3-деревьев.

8.48 Выпускной экзамен 1979 г.

8.49 Блом (Blom) [32]; выпускной экзамен 1984 г.

8.50 Выпускной экзамен 1986 г.

8.51 Выпускной экзамен 1986 г.

8.53 Феллер (Feller) [120] указывает в качестве автора С. Н. Бернштейна.

8.57 Лайл Рэмшоу (Lyle Ramshaw).*

8.58 Гуибас (Guibas) и Одлыжко (Odlyzko) [168].

9.1 Харди (Hardy) [179, 1.3(g)].

9.2 Часть (в) взята у Гарфункеля (Garfunkel) [140].

9.3 [207, упр. 1.2.11.1–6].

9.6 [207, упр. 1.2.11.1–3].

9.8 Харди (Hardy) [179, 1.2(iv)].

9.9 Ландау (Landau) [238, том. 1, с. 60].

9.14 [207, упр. 1.2.11.3–6].

9.16 Кнопп (Knopp) [205, издание $\geqslant 2$, §64C].

9.18 Бендер (Bender) [25, §3.1].

9.20 Выпускной экзамен 1971 г.

9.24 [164, §4.1.6].

9.27 Титчмарш (Titchmarsh) [352].

9.28 Глейшер (Glaisher) [149].

9.29 де Брейн (de Bruijn) [74, §3.7].

9.32 Выпускной экзамен 1976 г.

9.34 Выпускной экзамен 1973 г.

9.35 Выпускной экзамен 1975 г.

9.36 Семинары 1980 г.

9.37 [208, ф. 4.5.3–21].

9.38 Выпускной экзамен 1977 г.

9.39 Выпускной экзамен 1975 г., идея Рейча (Reich) [306].

9.40 Выпускной экзамен 1977 г.

9.41 Выпускной экзамен 1980 г.

9.42 Выпускной экзамен 1979 г.

9.44 Трикоми (Tricomi) и Эрдейи (Erdélyi) [353].
9.46 де Брейн (de Bruijn) [74, §6.3].
9.47 Домашнее задание 1980 г. [209, формула 5.3.1–34].
9.48 Выпускной экзамен 1980 г.
9.49 Выпускной экзамен 1974 г.
9.50 Выпускной экзамен 1984 г.
9.51 [164, §4.2.1].
9.52 Пуанкаре (Poincaré) [294]; Борель (Borel) [35, с. 27].
9.53 Пойа (Pólya) и Сегё (Szegő) [299, ч. 1, задача 140].
9.57 Эндрю М. Одлыжко (Andrew M. Odlyzko).*
9.58 Хенричи (Henrici) [182, упр. 4.9.8].
9.60 [225].
9.62 Кэнфилд (Canfield) [48].
9.63 Варди (Vardi) [358].
9.65 Комте (Comtet) [64, глава 5, упр. 24].
9.66 М. П. Шюценбергер (M. P. Schützenberger).*
9.67 Либ (Lieb) [250]; Стенли (Stanley) [335, упр. 4.37(c)].
9.68 Боас (Boas) и Ренч (Wrench) [33].

* Неопубликованное личное сообщение.

Предметный указатель

Граффити также следовало бы снабдить указателями.

В СЛУЧАЕ, КОГДА УКАЗАТЕЛЬ отсылает к странице, содержащей соответствующее упражнение, из ответа к этому упражнению (в приложении А) можно извлечь дополнительную информацию; при этом номера страниц с ответами не указываются, если только указатель не отсылает к понятию, которое не фигурирует в формулировке соответствующего упражнения.

Список таблиц

Оригинал-макет американского издания подготовлен в Стенфордском университете при помощи издательской системы TeX для технических текстов, разработанной Д. Э. Кнутом. Для набора математических формул использовано новое семейство шрифтов AMS Euler (Version 2.1), разработанное Германом Цапфом по заказу Американского математического общества. Текст набран шрифтами Concrete Roman и Italic — новой версией семейства Computer Modern Д. Э. Кнута, гармонирующей с математическими шрифтами AMS Euler.

КОНКРЕТНАЯ МАТЕМАТИКА

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ

РОНАЛЬД Л. ГРЭХЕМ, *AT&T Bell Laboratories*
ДОНАЛЬД Э. КНУТ, *Stanford University*
ОРЕН ПАТАШНИК, *Center for Comm. Research*

Эта книга представляет собой введение в математику, служащую основой программирования и анализа алгоритмов. Главная цель ее знаменитых авторов — заложить теоретический математический фундамент и помочь овладеть практическими навыками, необходимыми для решения сложных задач, вычисления сумм устрашающего вида, обнаружения тонких закономерностей в данных и многого другого. Это книга не только для специалистов в области информатики — каковыми являются ее авторы, — но и для всех тех, кто всерьез использует математику независимо от области знаний, в которой они работают.

Название "конкретная математика" произошло от двух терминов: "КОНтинуальноая математика" и "дисКРЕТНАЯ математика". Его можно понимать и буквально: обучение общим методам ведется на многочисленных конкретных примерах и упражнениях разной степени сложности. Всего в книге представлено более 500 упражнений, разделенных на шесть категорий сложности. Ко всем упражнениям (кроме исследовательских проблем) приводятся полные ответы, что делает книгу особенно ценной для самостоятельного изучения.

Книгу можно рассматривать как расширенную версию "Математического введения" из *Искусства программирования* Д. Кнута, но с более подробным и обстоятельным изложением материала и более глубоким погружением в отдельные темы. В нее добавлен ряд новых тем, а развитие наиболее важных идей прослежено до исторических корней.

Основные темы книги

• Суммы • Рекуррентные соотношения • Целочисленные функции • Элементарная теория чисел • Биномиальные коэффициенты • Производящие функции • Дискретная вероятность • Асимптотические методы

Во второе издание книги вошли новые важные материалы о механическом суммировании. В ответ на широкое использование первого издания в качестве справочника авторы существенно доработали библиографию и предметный указатель. Впрочем, нетривиальные улучшения можно найти почти на каждой странице книги.

Читатели должны оценить неформальный стиль *Конкретной математики*, в частности многочисленные пометки на полях, в том числе шутки студентов. Авторы хотят, чтобы читатель не только изучил важные методы и получил нужные знания, но и применял их с удовольствием.

Об авторах

Рональд Л. Грэхем — научный руководитель в исследовательских лабораториях AT&T. Он является профессором математики Рутгерского университета и бывшим президентом Американского математического общества. Грэхем — автор еще шести книг по математике.

Дональд Э. Кнут — почетный профессор информатики Стэнфордского университета. Среди его трудов — знаменитый трехтомник *Искусство программирования* и пять томов, посвященных издательским системам TEX и METAFONT.

Орен Паташник — сотрудник Исследовательского центра средств связи в Ла-Холья и автор широко используемого библиографического пакета BibTEX.

ISBN 978-5-8459-1588-7

ВИЛЬЯМС
www.williamspublishing.com

ADDISON-WESLEY