

具 体 数 学

计算机科学基础

[美] R.L.格雷厄姆
D.E.克努特 著
O.帕塔希尼克

庄心谷 译

西安电子科技大学出版社

1992

(陕)新登字 010 号

CONCRETE MATHEMATICS
A Foundation for Computer Science
Ronald L. Graham
Donald E. Knuth
Oren Patashnik
ADDISON-WESLEY (1989)

内 容 简 介

本书介绍支持高级计算机程序设计和算法分析的数学。它的主要目的是给计算机科学的学生和专业人员提供数学技巧方面的一个坚实的和合适的基础。介绍的内容对其他领域的学生和专业人员也很有用。

具体数学既是抽象数学的一个伙伴，又是连续数学和离散数学的一个混合体。作者“更具体地”解释，“具体数学是数学公式的控制操作，利用了解问题的一批技巧”。这些技巧使程序设计者、理论工作者和其他科学工作者能计算看来可怕的和，解复杂的递归关系，以及发现数据中的巧妙的型式。

具体数学是一项重要的新工作，通过作者卓越的工作形成了一本实用教科书。同时，它也是 Knuth 的经典著作《计算机程序设计技巧》一书中“数学预备知识”的扩展，深化了课题涉及的范围，且表述更加从容。书中包含 500 多个习题（分为六类），除了研究的问题外，给出了所有习题的完整解答。

主要课题包括，和、递归、整函数、初等数论、二项系数、母函数、离散概率、渐近方法。

具 体 数 学

计算机科学基础

[美] R L 格雷厄姆

D E 克努特 著

O. 帕塔希尼克

庄心谷 译

责任编辑 梁家新

西安电子科技大学出版社出版发行

西安电子科技大学印刷厂印刷

新华书店经销

开本 787×1092 1/16 印张 34 8/16 字数 822 千字

1992 年 1 月第 1 版 1992 年 1 月第 1 次印刷 印数 1-1 000

ISBN 7-5606-0174-X / TP · 0056 定价：29.00 元

中文版序言

本书介绍我在过去 30 年间研究计算机科学时经常用到的一些数学技巧。我为该书这样快地译成中文而感到高兴和荣幸。我希望中国的计算机科学的学生也像我一样喜爱钻研这些引人入胜的数学模型。

D.E.Knuth(高德纳)^①

1991 年 1 月

^①Knuth教授希望本书使用 Frances Yao 教授为他取的中文姓名“高德纳”(Gao Dunah)。

序 言

从 1970 年起, Stanford 大学每年开设一次名为具体数学的课程, 本书就是作者在讲授这门课程的基础上写成的。每年大约有 50 名学生选修这门课, 大部分为研究生, 也有三年级和四年级的学生, 而这些学生毕业后又在别的学校开设类似的课程。由此看来, 本教材提供给广大读者(包括大学二年级学生)的时机已经成熟了。

具体数学诞生于昏暗的、变动剧烈的十年。在那动荡的年代里, 大学校园是论战的温床, 经常对保持长时期有用性表示怀疑。对大学课程本身也表示异议, 而数学也未能逃脱详尽的考查。John Hammersley 就写了一篇发人深思的文章“由于学院和大学中的‘近世数学’以及类似的模糊的、费脑筋的无用的内容削弱了数学技巧”^[145], 还有一些担忧的数学家^[272]甚至问道:“数学能被拯救吗?”本书的作者之一, Knuth 写了一套名为《计算机程序设计技巧》的书, 他在写第一卷时发现, 在当时那些数学工具书里得不到他所需要的完整的组成部分。他需要的数学是完全不同于他在大学中主修的数学, 而是一种能非常精确领会以及有充分根据判断计算机程序的数学。所以他引入了一门新的课程, 在这门课程中教他希望过去有人教给他的那些内容。

由于通常称为“抽象数学”的抽象观念的新浪潮把具体的经典结果迅速地扫出近代数学课程, 所以原先打算把“具体数学”课程作为对“抽象数学”的一种矫正。抽象数学是一门极好的学科, 它没有任何错误: 它是优美的, 具有普遍意义的和有用的。但是抽象数学的追随者误入歧途, 因为抽象数学的余下部分是次要的内容, 且不再有值得注意的价值。一般化的目标非常流行, 而一代数学家变得不乐于研究特殊项目的妙处, 不喜欢接受解定量问题的挑战, 或者意识不到技巧的价值。抽象数学变成了限于狭隘范围内的科目, 且与现实失去联系。为了恢复合理的平衡, 数学教育需要一种具体的措施。

当 Knuth 第一次在史坦福大学讲授具体数学时, 他说了下面一段话用来解释有点陌生的课程名称。他尝试教一门难对付的而不是轻松的数学课程。和他的一些同事们的希望相反, 他不准备讲集合论, 也不准备讲 Stone 嵌入定理, 甚至也不讲 Stone-Čech 的紧化。(几个建筑工程系的学生起立并悄悄地离开了教室)。

虽然具体数学开始是相对于其他潮流的一种反应, 但是它存在的主要理由是明确的, 而不是否定的。作为一门课程, 它继续处于通用地位, 有着“充实”的内容, 且证明在许多新的应用中, 这些内容是有价值的。在此期间, Z.A.Melzak 出版了名为《具体数学指南》^[214]的两卷书, 这从另一方面独立证实了这种名称是合适的。

具体数学的内容初看起来好像是各不相同的各种方法, 但是实践使它成为训练有素的

一套工具。的确，这些技巧有潜在的一致性，且对许多人有很强的吸引力。当另一位作者Graham在1979年第一次讲授此课程时，学生很感兴趣，他们决定一年后再继续重开一个班。

但是具体数学究竟是什么呢？它是连续数学和离散数学的混合物。更具体地说，利用解问题的一组技巧，控制操作数学公式。一旦读者学了本书中的内容，为了计算看来可怕的和，解复杂的递归关系，以及在数据中发现微妙的模型，你仅需一个冷静的头脑，一大张纸，以及写得相当不错的字。在代数技巧方面，你会那样流畅，你常会感到，求准确的结果比解决在以极限意义下成立的近似解答更容易。

本书中讨论的主要课题包含和，递归，初等数论，二项系数，母函数，离散概率和渐近方法。重点是放在操作技巧方面，而不是放在存在性定理或组合推理上；目标是使每个读者熟悉离散运算（像最大整函数和有限求和），就像学微积分的学生熟悉连续运算（像绝对值函数和无穷积分）一样。

注意，这些课题完全不同于大学生课程“离散数学”通常所教的内容，所以需要一个新的学科名称，而“具体数学”已成为合适的名称。

原先Stanford所用的具体数学的教材为《计算机程序设计技巧》^[173]中“数学预备知识”那一节，但是这110页的内容过于简略，所以另一个作者Palashnik起草了篇幅较大的补充草稿，本书就是在这些草稿基础上形成的。它是数学预备知识材料的扩充，且较从容地引入数学预备知识的内容，删去了一些较高深的部分；另一方面，这里增加了几个原先没有的课题，使内容更完整。

作者们乐意共同来完成这本书，因为课程的内容已经明确，其内容应具有的特点也已展现在我们面前；本书似乎写出了它自己的特色。而且，在几处我们所采用的有点不同于一般的方法看来已很好地配合在一起，经过这些年的实践我们不禁感到，本书有点象是我们所喜爱的研究数学方法的宣言书。所以我们认为本书的叙述体现了数学的美妙和惊人之处。我们希望读者至少也将分享我们在写此书时所感受到的喜悦。

由于本书诞生在大学环境中，所以我们尝试采用口语体以同时活跃课堂教学。有些人认为数学是一项严肃的工作，它一定总是枯燥无味的；但是我们却认为数学是有趣味的，我们不羞于承认这一实际情况。为什么要在工作和游戏之间画一条严格的界线呢？具体数学富有吸引人的模型，虽然计算操作并不一定是容易的，但解答却能令人惊讶地吸引人。本书中明显反映了教学工作的甘苦，因为他们已是我们生活中不可缺少的组成部分。

本书包含500多个习题，分为六类：

- 准备部分是每个读者初次阅读本书时应尽量完成的习题。
- 基本部分是推导论据的习题，学习论据最好的方法是，通过自己的努力推导，而不是阅读别人的推导。
- 课外习题是为加深对本章内容的理解而设置的问题。
- 考查性问题一般同时涉及两章或两章以上的概念；一般把它们作为课外考查(由于时间关系不作为课堂考查)。
- 额外问题是超出了学习具体数学的一般学生(指以本书作为该课程教材时)所能处理的问题的范围，这些问题以有趣的方式扩展了本教材的内容。
- 研究性问题在人力所及范围可能或不可能求得解答，但是这里提供的问题似乎值得

一试(不受时间限制)。

在附录 A 中给出了所有习题的解答,并常常附有有关结论的补充信息。(当然,研究性问题的“解答”是不完全的,但是即使这样,给出部分结果或提示可能是有帮助的。)鼓励读者查看解答,特别是准备部分问题的解答,但是在认真尝试解问题之前不要去查看解答。

在附录 C 中,我们曾想适当地注明每个习题的出处,因为一个启发性问题的设计往往包含有大量创造性和/或运气。数学工作者已遗憾地形成了一种习惯,就是引用习题不必表示任何感谢;我们认为相反的习惯,譬如一些关于国际象棋的杂志的做法(书中按惯例详细注明原象棋问题的名称,日期和地点)是十分好的,然而,我们不能确定许多已变成民间传说的问题的出处。如果读者知道我们引用的习题出处遗漏了或不准确,我们将十分高兴获悉详情,以便我们能在本书的后续版本中修改这些错漏。

作者非常感谢 Andrei Broder, Ernst Mayr, Andrew Yao 和 Frances Yao, 这些年来他们在 Stanford 大学讲授具体数学过程中为本书作出了很大贡献。此外我们十分感谢助教们,他们每年在班上创造性地收录资料并帮助设计考题,他们的名字列在附录 C 中。本书实质上是 16 年可贵的讲稿的一个概要,没有他们的一流的工作,就不可能有这本书。

还有许多人为完成本书提供了帮助。例如,我们要称赞 Brown, Columbia, CUNY, Princeton, Rice 和 Stanford 等大学的学生们,他们发现并修正了初稿中的错误。我们与 Addison-Wesley 出版社的交往特别有效和有帮助;我们特别要感谢我们的出版者 Peter Gordon, 生产主管人 Bette Aaronson, 制图者 Roy Brown, 以及校订者 Lyn Dupré。国家科学基金会和海军研究局给予了非常宝贵的支持。在我们准备索引的时候,Cheryl Graham 提供了很大帮助。我们尤其要感谢我们的妻子(Fan, Jill 和 Amy),感谢她们的耐心,支持,鼓励和意见。

我们试图完成一本完美的书,但是我们是不完美的作者。所以我们要求大家帮助来改正任何错误。对于第一个发现任何错误者,不管是数学的错误,还是历史的错误或印刷错误,将奖赏 \$2.56 以示感谢。

Ronald L. Graham

Donald E. Knuth

Oren Patashnik

1988 年 5 月

符号注释

本书所采用的符号有一些尚未成为标准符号，这里列出了对一些读者(指学过其他相似内容教材的读者)可能不熟悉的符号。

符号	名称
$\ln x$	自然对数: $\log_e x$
$\lg x$	以 2 为底的对数: $\log_2 x$
$\log x$	常用对数: $\log_{10} x$
$\lfloor x \rfloor$	下整: $\max\{n n \leq x, \text{ 整数 } n\}$
$\lceil x \rceil$	上整: $\min\{n n \geq x, \text{ 整数 } n\}$
$x \bmod y$	剩余: $x - y \lfloor x / y \rfloor$
$\{x\}$	分数部分: $x \bmod 1$
$\sum f(x) \delta x$	不定求和
$\sum_a^b f(x) \delta x$	定求和
$x^{\overline{n}}$	下降阶乘幂: $x! / (x - n)!$
$x^{\underline{n}}$	上升阶乘幂: $\Gamma(x + n) / \Gamma(x)$
$n!$	次阶乘: $n! / 0! - n! / 1! + \cdots + (-1)^n n! / n!$
$\mathcal{R} z$	实部: x , 如果 $z = x + iy$
$\mathcal{I} z$	虚部: y , 如果 $z = x + iy$
H_n	调和数: $1/1 + \cdots + 1/n$
$H_n^{(x)}$	广义调和数: $1/1^x + \cdots + 1/n^x$
$f^{(m)}(z)$	在 z 处 f 的 m 阶导数
$\left[\begin{smallmatrix} n \\ m \end{smallmatrix} \right]$	Stirling 轮换数("第一类")
$\left\{ \begin{smallmatrix} n \\ m \end{smallmatrix} \right\}$	Stirling 子集数("第二类")

• 2 •

符号	名称
$\langle \frac{n}{m} \rangle$	欧拉数
$\langle\langle \frac{n}{m} \rangle\rangle$	第二阶欧拉数
$(a_m \cdots a_0)_b$	$\sum_{k=0}^m a_k b^k$ 的根值记数法
$K(a_1, \cdots, a_n)$	延拓多项式
$F\left(\begin{matrix} a, b \\ c \end{matrix} \middle z\right)$	超几何函数
$\#A$	基数; 集合 A 中的元素个数
$[z^n]f(z)$	$f(z)$ 中 z^n 的系数
$[x.. \beta]$	闭区间: 集合 $\{x \alpha \leq x \leq \beta\}$
$[m=n]$	1 如果 $m=n$, 否则为 0*
$[m \setminus n]$	1 如果 m 除尽 n , 否则为 0*
$[m \setminus \setminus n]$	1 如果 m 整除 n , 否则为 0*
$[m \perp n]$	1 如果 m 和 n 互素, 否则为 0*

* 一般, 如果 S 是能真或假的任何命题, 加括号的记法 $[S]$ 表示如果 S 是真, 为 1, 否则为 0.

书中我们用单引号('...')表示书写的, 双引号("...")内的一个短语表示说的。因此, 字母串 'String' 有时称为一个 "string".

形式 ' a/bc ' 的一个表达式和 ' $a/(bc)$ ' 的意思相同。此外, $\log x / \log y = (\log x) / (\log y)$, $2n! = 2(n!)$.

目 录

第一章 递归问题		4.9 φ 和 μ	115
1.1 汉诺塔	1	习题	124
1.2 平面中的直线	4	第五章 二项系数	
1.3 Josephus 问题	7	5.1 基本等式	132
习题	14	5.2 基本的实用	148
第二章 和		5.3 处理的诀窍	160
2.1 表示法	18	5.4 母函数	170
2.2 和以及递归	21	5.5 超几何函数	176
2.3 和的操作	25	5.6 超几何变换	187
2.4 多重和	29	5.7 部分超几何和	193
2.5 一般的方法	36	习题	199
2.6 有限和无限演算	41	第六章 特殊数	
2.7 无限和	50	6.1 Stirling 数	211
习题	55	6.2 欧拉数	220
第三章 整函数		6.3 调和数	225
3.1 下整函数和上整函数	60	6.4 调和的求和	231
3.2 下整 / 上整的应用	62	6.5 伯努利数	234
3.3 下整 / 上整递归	69	6.6 Fibonacci 数	242
3.4 'MOD': 二元运算	72	6.7 延拓	251
3.5 下整 / 上整的和	75	习题	258
习题	83	第七章 母函数	
第四章 数论		7.1 多米诺(骨牌)理论和兑换	267
4.1 可除性	90	7.2 基本操作	276
4.2 素数	92	7.3 解递归	281
4.3 素数例子	94	7.4 特殊的母函数	293
4.4 阶乘因子	97	7.5 卷积	295
4.5 互素性	100	7.6 指数型母函数	305
4.6 'MOD': 同余关系	107	7.7 Dirichlet 母函数	310
4.7 独立剩余	109	习题	312
4.8 附加的应用	112		

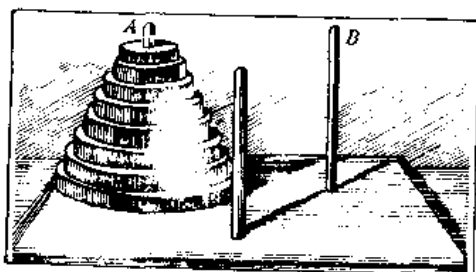
第八章 离散概率		9.2 O 表示法	371
8.1 定义	321	9.3 O 操作	377
8.2 平均值和方差	326	9.4 两个渐近的特殊技巧	389
8.3 概率母函数	332	9.5 Euler 的求和公式	395
8.4 掷硬币	337	9.6 最后的求和	400
8.5 散列法	345	习题	413
习题	358	附录 A 习题解答	420
第九章 渐近		附录 B 参考文献	512
9.1 级别	369	附录 C 注明习题的出处	535

第一章 递归问题

本章研究三个样本问题，这三个样本问题给出了递归问题的感性知识。它们有两个共同的特点：它们都是数学家们一直反复地研究的问题；它们的解都用了递归的概念，按递归概念，每个问题的解都依赖于相同问题的若干较小场合的解。

1.1 汉诺塔

首先让我们看一个称为汉诺塔的巧妙的测验智力的小问题，这是法国数学家 Edouard Lucas 在 1883 年提出的。题目给我们 8 个圆盘组成的一个塔，最初圆盘按尺寸递减的次序堆放在三根杆中的一根杆上(总共有三根杆)：目的是把整个塔移到另一根杆上，一次仅移 1 个圆盘，且 1 个较大的盘子不得移到比它小的盘子的上面。



Lucas^[208]按照一个虚构的传说，把他的玩具安装成印度教的非常大的塔，想象有 64 个纯金的圆盘安放在三根金刚石针上。开始时，他说，上帝把这些金盘放在第一根针上，且命令一群教士按照上述的规则把它们移到第三根针上去。据传说，教士们日夜做他们的工作。当他们完成时，塔将碎裂且世界将临末日。

一下很难看出问题有一个解，但稍加思索后我们确信它有一个解。现在出现了问题：我们该怎么干才好呢？也就是说，为了完成任务，多少次移动是必要和充分的？

处理这样一个问题的最好方法是把它一般化一点。印度教的塔有 64 个圆盘，汉诺塔有 8 个；让我们考虑如果有 n 个圆盘将发生什么情况。

这种推广的一个好处是，我们能把问题的 n 大大地降低。事实上，在本书中我们将反复见到首先看小的情形是有利的。我们一下就能看出如何移动仅包含 1 或 2 个盘的塔，并且用少量试验就能说明如何移 3 个盘的塔。

解问题的下一步是引入适当的表示法：取名并解出它。让我们说 T_n 是在 Lucas 的规则下 n 个盘从一根杆移到另一根杆的最小移动次数，则 T_1 显然是 1， $T_2 = 3$ 。为了方便，通过考虑最小情形，我们还能取得另一个数据：显然 $T_0 = 0$ ，因为转移 $n=0$ 个盘的塔不需要任何移动！精明的数学家不为思索小的情形而感到害臊，因为当极端的情形完全了解时，一般型式较容易理解（即使当它们是平凡的情形时）。

但是现在让我们改变自己的观点，尝试去思索大的情形。我们如何能转移一个大的塔？3 个盘的试验表明获胜的思想是把顶上 2 个盘转移到中间杆，然后移第 3 个盘，再把另外 2 个盘放到它的上面。这就提供给我们一个关于 n 个盘一般转移的思路：首先把 $n-1$ 个最小的盘转移到一个不同的杆（要求 T_{n-1} 次移动），然后移动最大的盘（要求一次移动），最后再把 $n-1$ 个最小盘转移回最大的盘上（要求另外的 T_{n-1} 次移动）。因此，我们至多用 $2T_{n-1} + 1$ 次移动能转移 n 个盘 ($n > 0$)：

$$T_n \leq 2T_{n-1} + 1, (n > 0).$$

此公式用了 ' \leq ' 而不是 ' $=$ '，因为我们的构造仅证明 $2T_{n-1} + 1$ 次移动是充分的，并未证明 $2T_{n-1} + 1$ 次移动是必要的。一个聪明的人可能想到一种捷径来证明它是必要的。

但是是否有一个较好的方法呢？实际上没有，在某个时刻我们一定要移动最大的盘。当我们移动最大盘时， $n-1$ 个最小的盘一定在一根杆上，且至少用 T_{n-1} 次移动把它们放到那里。若我们不太注意，则移动最大盘的次数可能大于一次。但是最后一次移动最大盘之后，我们一定要把 $n-1$ 个最小盘（一定还在一根杆上）转移回最大盘上，这还要求 T_{n-1} 次移动。因此

$$T_n \geq 2T_{n-1} + 1, (n > 0).$$

这样两个不等式和 $n=0$ 的平凡解一起，产生

$$\begin{aligned} T_0 &= 0 \\ T_n &= 2T_{n-1} + 1, (n > 0). \end{aligned} \tag{1.1}$$

（注意这些公式是和值 $T_1 = 1$ 和 $T_2 = 3$ 一致的。我们对小的情形的试验不仅有助于发现一个一般的公式，而且还提供了一种便于检查我们是否犯了可笑错误的方法。在后面几章中，当我们进入较复杂的运用时，这种检查将特别有价值。）

像式(1.1)那样一组等式被称为一个递归（递推关系或递归关系）。它给定一个边界值以及根据较早值表达一般值的一个方程。我们有时把这个一般的方程单独称作为一个递归，但从学术上来说它需要一个边界值才是完整的。

递归使我们能计算任何 n 的 T_n ，但是当 n 比较大的时候没有人真正想用一个递归来计算，因为运算太长了。递归仅给出间接的“局部”信息。递归的一个解使我们非常高兴，也就是说，我们希望有一个良好的，简洁的 T_n 的“闭形式”，即使对于大的 n ，使我们能迅速地计算它。依据一个闭形式，我们能明白 T_n 真正是什么。

那么我们如何解一个递归呢？一种方式是猜测正确的解，然后证明我们的猜测是正确的，并且对于猜测解来说我们最希望做的事是看一下 n 值小的情况。因而我们接连计算

$T_3 = 2 \cdot 3 + 1 = 7$; $T_4 = 2 \cdot 7 + 1 = 15$; $T_5 = 2 \cdot 15 + 1 = 31$; $T_6 = 2 \cdot 31 + 1 = 63$. 它确实看来像

$$T_n = 2^n - 1, (n \geq 0). \quad (1.2)$$

至少对于 $n \leq 6$, 此公式成立。

数学归纳法是证明某个命题关于整数 n (对所有 $n \geq n_0$) 成立的一种一般的方法。首先, 当 n 为最小值 n_0 时我们证明命题, 这称为基础。然后假设对于包含在 n_0 和 $n-1$ 之间的所有值, 已经证明命题成立, 对于 $n > n_0$ 证明命题, 这种方法称为归纳。这样一个证明仅用有限量工作便产生无限多个结果。

数学归纳法完美地为递归作了准备。例如, 在我们讨论的情形, 容易从式(1.1)得出式(1.2): 由于 $T_0 = 2^0 - 1 = 0$, 所以基础是平凡的。若假设当 $n-1$ 取代 n 时式(1.2)成立, 且对于 $n > 0$ 建立归纳:

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

因此对于 n , 式(1.2)同样成立。好! 我们关于 T_n 的探索成功地完成了。

当然教士的任务还未终止, 他们仍在尽职地移动圆盘, 休息一会儿再移, 因为对于 $n = 64$, 有 $2^{64} - 1$ 次移动(差不多 18×1000^6)。即使以每微秒移动一次这个办不到的速率, 他们也将需要多于 5 000 世纪的时间来转移印度教的塔。Lucas 原先的问题较实际一点, 它要求 $2^8 - 1 = 255$ 次移动, 敏捷的手大约需花费 4min。

在各种应用提出的许多问题中, 汉诺塔递归有典型性。在找像 T_n 那样某个有趣量的一个闭形式的表达式中, 我们经历三个阶段:

1. 看看小的情形。这使我们深入了解问题, 这在第二和第三阶段中将对我们有帮助。

2. 求出和证明关心的量的一个数学表达式。对于汉诺塔, 就是递归式(1.1), 它使我们能就给定的量来计算任何 n 的 T_n 。

3. 求出和证明我们的数学表达式的一个闭形式。对于汉诺塔, 就是递归解(1.2)。第三阶段是我们在本书中始终钻研的阶段。事实上, 我们常常完全略过阶段一和二, 因为我们有一个作为出发点的数学表达式。但是尽管这样, 我们仍将进入其解将使我们经历三个阶段的子问题。

汉诺塔的分析导致正确的解答, 但是它要求一个“归纳跳跃”, 这里依赖于解答的侥幸的猜中。本书的主要目的之一是阐明解递归式的方法, 而并不要求读者具有超人的洞察力。例如, 我们将看到递归式(1.1)能通过加 1 来简化:

$$T_0 + 1 = 1;$$

$$T_n + 1 = 2T_{n-1} + 2, (n > 0).$$

如果令 $U_n = T_{n-1} + 1$, 我们得到

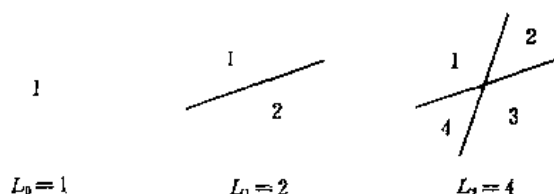
$$\begin{aligned} U_n &= 1 \\ U_n &= 2U_{n-1}, \quad (n > 0). \end{aligned} \quad (1.3)$$

发现此递归式的解就是 $U_n = 2^n$ ，这并不需要天才。因此， $T_n = 2^n - 1$ ，甚至于计算机都能发现这个解。

1.2 平面中的直线

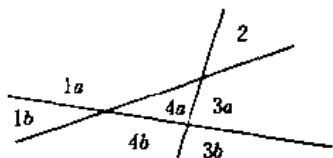
我们的第二个样本问题还具有一种几何特点：一个人用刀在意大利馅饼上做 n 次直线切割，能把馅饼切成多少份？或者，较正式地说：由平面中 n 条直线确定的最大区域数 L_n 是多少？瑞士数学家 Jacob Steiner^[278] 在 1826 年首先解决了这个问题。

我们还是先看小的情形，记住是从最小的情形开始。没有直线的平面具有 1 个区域；有 1 条直线的平面具有 2 个区域；有 2 条直线的平面有 4 个区域：



(每条直线在两个方向作无限延伸。)

自然，我们认为 $L_n = 2^n$ ，加入 1 条新的直线仅仅使区域数加倍。不幸，这是错误的。若第 n 条直线把每个老区域切成 2 个区域，则能使区域数加倍；由于每个老区域是凸的，当然它能把一个老区域至多切成两部分。(1 条直线能把 1 个凸区域至多切成 2 个新区域，新区域也将是凸的。) 但是当我们加入第 3 条直线，即在下图中的一条粗线时，我们立即发现，不管如何放前 2 条直线，它至多能切 3 个老区域：



因此 $L_3 = 4 + 3 = 7$ 是我们能做到的最好结果。

作一些思考之后，我们获得适当的推广。第 n 条直线 ($n > 0$) 增加的区域数为 k 当且仅当它切了 k 个老区域，而它切 k 个老区域当且仅当它在 $k-1$ 个不同的位置碰到前面的直线，2 条直线至多相交于一点。所以新的直线至多与 $n-1$ 条老的直线相交于 $n-1$ 个不同点，且一定有 $k \leq n$ 。这样我们就建立了上界

$$L_n \leq L_{n-1} + n, \quad (n > 0).$$

此外, 用归纳法易证我们能达到此公式中的等式. 我们只要使第 n 条直线不和其他任何直线平行(因此它和其他任何直线相交), 且使它不通过任何存在的交点(因此它和其他任何直线相交于不同的位置). 所以递归式为

$$\begin{aligned} L_0 &= 1; \\ L_n &= L_{n-1} + n, \quad (n > 0). \end{aligned} \quad (1.4)$$

在这里正确地检验了已知值 L_1 , L_2 和 L_3 , 所以我们将接受此式.

现在我们需要一个闭形式的解. 我们可以再做一次猜测游戏, 但是看来我们不熟悉 1, 2, 4, 7, 11, 16, ..., 所以让我们改变一下方针. 我们常常能把一个递归式“展开”或“摊开”到结束来了解它, 如下:


$$\begin{aligned} L_n &= L_{n-1} + n \\ &= L_{n-2} + (n-1) + n \\ &= L_{n-3} + (n-2) + (n-1) + n \\ &\vdots \\ &= L_0 + 1 + 2 + \cdots + (n-2) + (n-1) + n \\ &= 1 + S_n \end{aligned}$$

其中 $S_n = 1+2+3+\cdots+(n-1)+n$.

换句话说, L_n 是前 n 个正整数的和 S_n 加 1.

量 S_n 经常出现, 所以把它的小值作一个表是有价值的. 这样, 当下一次再见到它们时就可以认出这些数:

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

这些值也称为三角形数, 因为 S_n 是 n 行三角形阵中球状针孔的个数. 例如, 通常 4 行阵  有 $S_4 = 10$ 个针孔.

为了计算 S_n , 我们可用 Gauss 在 1786 年提出的一种技巧, 当时他为 9 岁^[73] (还可见 Euler[92, 部分 1, § 415]):

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

我们仅仅把 S_n 和它的“颠倒”相加, 以致右边 n 列的每一列之和均为 $n+1$. 化简可得

$$S_n = \frac{n(n+1)}{2}, \quad (n > 0). \quad (1.5)$$

我们得到了解:

$$L_n = \frac{n(n+1)}{2} + 1, (n \geq 0). \quad (1.6)$$

作为有经验的人, 我们可满足于这个推导, 且认为这就是证明, 即使在做展开以及思考时我们有点儿波动也无妨. 学数学的学生应能对付较严格的标准, 所以最好的办法是用归纳法建立一个严格的证明. 关键的归纳步是

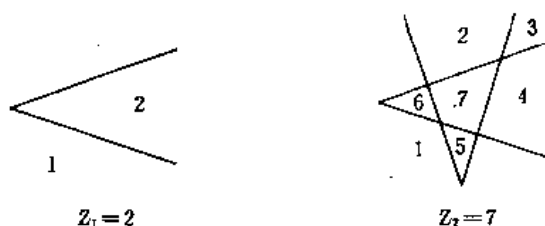
$$L_n = L_{n-1} + n = \left(\frac{1}{2}(n-1)n + 1 \right) + n = \frac{1}{2}n(n+1) + 1.$$

现在, 对闭形式(1.6)可确信无疑了.

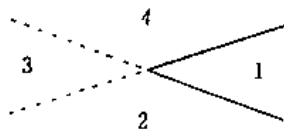
顺便提一句, 我们谈到“闭形式”时并没有明确说出指的是什么, 通常它是相当清楚的. 像式(1.1)和(1.4)的递归不是闭形式, 它们是根据量本身来表达一个量的; 而像式(1.2)和(1.6)是闭形式. 像 $1+2+\cdots+n$ 的和不是闭形式, 它们用“ \cdots ”来哄骗; 而像 $n(n+1)/2$ 的表达式是闭形式. 我们能给出一个如下的粗略定义: 如果我们能至多用独立于 n 的固定次“众所周知”的标准运算来计算它, 关于一个量 $f(n)$ 的表达式是闭形式. 例如, $2^n - 1$ 和 $n(n+1)/2$ 是闭形式, 因为它们以明确的方式仅涉及到加, 减, 乘, 除和指数.

简单的闭形式的总数是有限的, 而且不具有简单闭形式的递归式, 当这样的递归式反复出现而是重要的递归式时, 我们把新的运算加入我们的组成部分, 这就能大大地扩展以“简单”闭形式求解问题的范围. 例如, 已经证实前 n 个整数的乘积 $n!$ 很重要, 现在我们把考虑为一个基本运算. 所以公式“ $n!$ ”是闭形式, 虽然它的等式“ $1 \cdot 2 \cdot \cdots \cdot n$ ”不是闭形式.

现在再简要地考虑一下平面中直线的变形问题: 假设我们用弯曲的线来代替直线, 每个弯曲线含有一个“锯齿形的转角”, 平面中 n 个这样的弯曲线确定的区域的最大个数 Z_n 是多少? 我们可能认为 Z_n 是 L_n 两倍那样大, 或者可能是三倍那样大. 让我们来看:



根据这些小的情形, 且稍加思索, 我们看出: 1 条弯曲线像“2 条”直线过它们的交点不作延伸而合并了一些区域的 2 条直线.



就 2 条直线来说, 区域 2, 3 和 4 是不同的, 当 1 条弯曲线时, 变成 1 个区域。不管怎样, 如果我们作适当安排: 锯齿形转角点一定位于其他线的交点的“远处”, 这就是我们失去的所有区域; 也就是说, 每条线我们仅失去 2 个区域。因此,

$$\begin{aligned} Z_n &= L_{2n} - 2n = 2n(2n+1)/2 + 1 - 2n \\ &= 2n^2 - n + 1, \quad (n \geq 0) \end{aligned} \quad (1.7)$$

与闭形式(1.6)和(1.7)比较, 我们发现对于大的 n ,

$$L_n \sim \frac{1}{2}n^2,$$

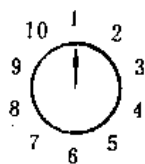
$$Z_n \sim 2n^2;$$

所以在弯曲线情况下所能获得的区域数大约是直线情况下的 4 倍。(在后面几章中, 我们将讨论如何分析当 n 很大时整函数的近似状态。)

1.3 Josephus 问题

我们最后介绍的例子是一个古老问题的变形, 它是以 1 世纪著名历史学家 Flavius Josephus 命名的。据传说, Josephus 若没有数学才能, 他就不会在活着的时候出名, 在犹太人和古罗马人战争期间, 他是陷入罗马人陷阱的 41 个犹太反抗者之一。反抗者宁死不做俘虏, 他们决定围成一个圆圈, 且环绕圆圈来进行, 杀死所有第三个剩下的人直到没有一个人留下。但是 Josephus 和一个不告发的同谋者感到自杀是愚蠢的行为, 所以他快速计算出在此恶性循环中他和他的朋友该站的地方。

在我们的变形中, 我们从围着一个圆的编号为 1 到 n 的 n 个人开始, 接着我们排除所有第 2 个剩下的人直到仅仅 1 个幸存者。例如, 这里是 $n=10$ 的开始形状:



排除的次序是 2, 4, 6, 8, 10, 3, 7, 1, 9, 所以 5 是幸存者。问题: 确定幸存者的号码 $J(n)$ 。

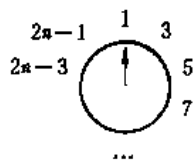
我们刚才看到 $J(10)=5$ 。我们可能猜测当 n 是偶数时 $J(n)=n/2$; 而且情形 $n=2$ 支持了猜测: $J(2)=1$ 。但是其他几种小的情形阻止了我们, 对于 $n=4$ 和 $n=6$, 猜测失败。

n	1	2	3	4	5	6
$J(n)$	1	1	3	1	3	5

返回到抽取位置, 让我们尝试作一种较好的猜测。看来 $J(n)$ 总是奇数。而事实上, 关于这

一点有一个充分的理由：围绕圈的第一次来回消去了所有偶数。此外，若 n 本身是偶数，除了人数减半以及改变了它们的编号，我们得出一种与开始时相似的场合。

所以让我们假设原先有 $2n$ 个人。在绕第一圈之后，留下



而下一个死的人将是 3。这就像 n 个人开始的那样，只是每个人的编号加倍，然后减 1。也就是说，

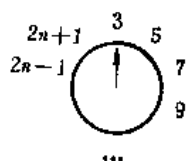
$$J(2n) = 2J(n) - 1, \quad (n \geq 1).$$

现在我们能很快地到达大的 n 。例如，我们知道 $J(10) = 5$ ，所以

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

相似地有 $J(40) = 17$ ，而且我们推出 $J(5 \cdot 2^m) = 2^{m+1} + 1$ 。

但是奇数情形呢？具有 $2n+1$ 个人，结果是在编号 $2n$ 被除去之后就除去编号为 1 的人，那我们留下



我们差不多还是得到原来具有 n 个人的情况，但是此时它们的编号加倍，且加 1，于是

$$J(2n+1) = 2J(n) + 1, \quad (n \geq 1).$$

这些方程和 $J(1) = 1$ 一起给出了就一切情形定义 J 的一个递归式：

$$\begin{aligned} J(1) &= 1; \\ J(2n) &= 2J(n) - 1, \quad (n \geq 1); \\ J(2n+1) &= 2J(n) + 1, \quad (n \geq 1). \end{aligned} \tag{1.8}$$

这个递归式不是从 $J(n-1)$ 取得 $J(n)$ ，它是非常“有效”的，因为每次应用此式时，把 n 减少二分之一或更多。比如说，我们仅用式(1.8)19 次能计算出 $J(1\,000\,000)$ 。但是我们仍然要找一闭形式，它将更快和更有益，毕竟这是生死攸关的事情。

我们的递归式可能十分快地建立一个小值的表，也许我们将能发现一种型式或猜测解答。

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

看来我们能以 2 的幂分组(表中由竖线表明的)，在一组的开始处 $J(n)$ 总是 1，而且在

一组内它增加 2。所以若把 n 记为形式 $n = 2^m + l$, 其中 2^m 是不超过 n 的 2 的最大幂, 而 l 是留下的量, 我们的递归式的解看来是

$$J(2^m + l) = 2l + 1, \quad (m \geq 0, 0 \leq l < 2^m). \quad (1.9)$$

(注意, 若 $2^m \leq n < 2^{m+1}$, 则剩余部分 $l = n - 2^m$ 满足 $0 \leq l < 2^{m+1} - 2^m = 2^m$.)

现在我们一定要证明式(1.9), 如同过去我们所用的归纳法那样, 但是这里是对 m 归纳。当 $m = 0$ 时, 我们一定有 $l = 0$; 因此式(1.9)的基础化为 $J(1) = 1$, 这是真的。归纳步有两部分, 依赖于 l 是偶数还是奇数。若 $m > 0$ 和 $2^m + l = 2n$, 则 l 是偶数并由式(1.8)和归纳假设可得

$J(2^m + l) = 2J(2^{m-1} + l/2) - 1 = 2(2l/2 + 1) - 1 = 2l + 1$, 这恰好是我们所要的。相似的证明引进 $2^m + l = 2n + 1$ 的奇数情形。我们还可能注意到式(1.8)意味着关系

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

不管怎么说, 归纳法完成, 建立了式(1.9)。

为了说明解式(1.9), 让我们计算 $J(100)$ 。此时我们得到 $100 = 2^6 + 36$, 所以 $J(100) = 2 \cdot 36 + 1 = 73$ 。

我们已完成了(解问题的)费力的工作, 现在来寻找不太费力的工作: 推广问题的每一个解以使它应用于一个较广的问题类。一旦我们学会了一种技巧, 仔细地查看它是有启发的, 并看和它一起我们能走多远。因此, 本节的余下部分, 我们将研究解(1.9), 且探讨递归式(1.8)的一些推广。这些探讨将揭示构成所有这样的问题的结构。

我们发现解中 2 的幂起了重要的作用, 所以自然要看二进制表示 n 和 $J(n)$ 。假设 n 的二进制展开为

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

也就是说,

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

其中每个 b_i 是 0 或 1, 且其中的第一位 b_m 是 1。想到 $n = 2^m + l$, 我们相继得到

$$n = (1b_{m-1}b_{m-2}\cdots b_1b_0)_2,$$

$$l = (0b_{m-1}b_{m-2}\cdots b_1b_0)_2,$$

$$2l = (b_{m-1}b_{m-2}\cdots b_1b_00)_2,$$

$$2l + 1 = (b_{m-1}b_{m-2}\cdots b_1b_01)_2,$$

$$J(n) = (b_{m-1}b_{m-2}\cdots b_1b_0b_m)_2.$$

(最后一步是由 $J(n) = 2l + 1$ 和 $b_m = 1$ 而得出的。)我们证明了

$$J((b_m b_{m-1} \cdots b_1 b_0)_2) = (b_{m-1} \cdots b_1 b_0 b_m)_2; \quad (1.10)$$

也就是说, 以计算机程序设计的用语, 循环左移一位我们从 n 取得 $J(n)$ 不可思议。例如,

若 $n = 100 = (1100100)_2$, 则 $J(n) = J((1100100)_2) = (1001001)_2$, 它是 $64+8+1=73$. 若我们全沿用二进制表示计算, 也许会立即发现此型式。

若从 n 开始, 且重复作 J 函数 $m+1$ 次, 则我们作 $m+1$ 次 1 位的循环移动; 由于 n 是一个 $(m+1)$ 位的数, 所以可期望最后再得出 n . 但是事情并非真是这样。例如, 若 $n=13$, 则我们得到 $J((1101)_2) = (1011)_2$, 但是 $J((1011)_2) = (111)_2$, 且过程中止, 当 0 变成第 1 位时, 它消失了。事实上, 由定义 $J(n)$ 总是 $\leq n$, 因为 $J(n)$ 是幸存者的编号; 因此若 $J(n) < n$, 连续重复, 我们不可能倒转到 n .

重复应用 J 产生一系列下降值, 最终到达一个“指定点”, 在那里 $J(n)=n$. 循环移的性质使我们容易看出指定点是什么: 重复函数足够多次总将产生一个全为 1 的型式, 它的值是 $2^{v(n)}-1$, 其中 $v(n)$ 是 n 的二进制表示中 1 的位数。因此, 由于 $v(13)=3$, 我们得到

$$\begin{array}{l} \text{2个或多于2个 } J \\ J(\overbrace{J(\cdots J(13)\cdots)}) = 2^3 - 1 = 7; \end{array}$$

类似

$$\begin{array}{l} \text{8个或多于8个 } J \\ J(\overbrace{J(\cdots J((101101101101011)_2)\cdots)}) = 2^{10} - 1 = 1023. \end{array}$$

难以理解, 但是为真。

让我们短暂地回到第一次猜测, 当 n 是偶数时 $J(n)=n/2$. 一般这显然是不真的, 但是当它为真时, 我们现在能确切地决定它:

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

若此数 $l = (1/3)(2^m-2)$ 是一个整数, 则 $n = 2^m + l$ 将是一个解, 因为 l 将小于 2^m . 不难验证, 当 m 是奇数时 2^m-2 是 3 的倍数, 但当 m 是偶数时则不是。(在第四章我们将研究这样的事情。) 所以方程 $J(n)=n/2$ 有无限多解, 开头部分如下:

m	l	$n = 2^m + l$	$J(n) = 2l+1 = n/2$	n (二进制)
1	0	2	1	10
3	2	10	5	1010
5	10	42	21	101010
7	42	170	85	10101010

注意最右一系列中的型式。这些数是二进制数, 对于这些数向左循环移 1 位产生通常右移 1 位(减半)的相同结果。

好，我们相当明白了 J 函数，下一步是推广它。若我们的问题产生出一个有点像式(1.8)的递归式，但具有不同的常数，将发生什么样的结果？对于解的猜测我们可能不很幸运，因为解可能实在古怪。让我们引入常数 α , β 和 γ 来研究这一点，且尝试找出较一般递归式

$$\begin{aligned} f(1) &= \alpha; \\ f(2n) &= 2f(n) + \beta, \quad (n \geq 1); \\ f(2n+1) &= 2f(n) + \gamma, \quad (n \geq 1) \end{aligned} \quad (1.11)$$

的一个闭形式。(我们原先的递归式有 $\alpha=1$, $\beta=-1$ 和 $\gamma=1$ 。)从 $f(1)=\alpha$ 开始，且用我们的方法逐步建立，我们能构造下列对于小的 n 值的一般表：

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)

看来 α 的系数是 n 的最大的 2 的幂。此外，在 2 的幂之间， β 的系数下降 1，向下到 0，而 γ 的系数从 0 向上上升 1。所以，如果我们通过分离出它对 α , β 和 γ 的依赖而把它表达为

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

看来

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

像往常一样，这里 $n = 2^m + l$ ，且 $0 \leq l < 2^m$ ($n \geq 1$)。

用归纳法证明式(1.13)和(1.14)不是很难，但是计算是凌乱和没有指导性的。幸运的是，有一种较好的方法，通过选取特殊值，然后合并它们来处理。让我们通过考虑特殊情形 $\alpha=1$, $\beta=\gamma=0$ ，且假设 $f(n)$ 等于 $A(n)$ 时来说明这一点。递归式(1.11)变成

$$\begin{aligned} A(1) &= 1; \\ A(2n) &= 2A(n), \quad (n \geq 1); \\ A(2n+1) &= 2A(n), \quad (n \geq 1). \end{aligned}$$

果然(通过对 m 归纳) $A(2^m + l) = 2^m$ 是真的。

接着, 让我们反过来用递归式(1.11)和解(1.13), 从简单函数 $f(n)$ 开始, 然后寻找是否有确定它的任何常数 (α, β, γ) 。比如说, 插入常数函数 $f(n) = 1$,

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

因此值 $(\alpha, \beta, \gamma) = (1, -1, -1)$ 满足这些方程, 将产生 $A(n) - B(n) - C(n) = f(n) = 1$ 。类似, 我们可插入 $f(n) = n$:

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

当 $\alpha = 1$, $\beta = 0$ 和 $\gamma = 1$ 时, 这些方程对一切 n 成立, 所以不需要用归纳法证明这些参数将产生 $f(n) = n$ 。我们已知在此时 $f(n) = n$ 将是解, 因为对于 n 的每一个值, 递归式(1.11)唯一地确定 $f(n)$ 。

现在我们的工作实质上已完成了! 我们已表明式(1.13)的函数 $A(n)$, $B(n)$ 和 $C(n)$ 满足方程

$$\begin{aligned} A(n) &= 2^m, \text{ 其中 } n = 2^m + l \text{ 且 } 0 \leq l < 2^m; \\ A(n) - B(n) - C(n) &= 1; \\ A(n) + C(n) &= n. \end{aligned}$$

概括地说, 它们解了式(1.11)。由于我们能取 $C(n) = n - A(n) = l$ 和 $B(n) = A(n) - 1 - C(n) = 2^m - 1 - l$ 解出这些方程, 所以可立即得出式(1.14)中的猜测。

此探讨说明了一种解递归的意外有用的包含各种组成部分的方法。首先, 我们找出对于已知解设置的一般参数, 这给出了我们能解的特殊情形的各种组成部分。然后把特殊情形结合起来得到一般情形。我们需要与独立参数一样多的独立的特殊解 (对于 α , β 和 γ , 此时有 3 个)。习题 16 和 20 提供了包含各种组成部分方法的进一步的例子。

我们知道原来的 J 递归有一个奇妙的解, 二进制表示为:

$$J((b_m b_{m-1} \cdots b_1 b_0)_2) = (b_{m-1} \cdots b_1 b_0 b_m)_2, \text{ 其中 } b_m = 1.$$

是否广义 Josephus 递归有这样奇妙的可能呢?

肯定的, 为什么不能呢? 若设 $\beta_0 = \beta$ 和 $\beta_1 = \gamma$, 我们能把广义递归(1.11)改写成

$$\begin{aligned} f(1) &= \alpha; \\ f(2n + j) &= 2f(n) + \beta_j, \quad j = 0, 1 \text{ 且 } n \geq 1. \end{aligned} \tag{1.15}$$

然后递归展开, 二进制方式为:

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

假设我们现在解除二进制表示, 允许任意数字位而不只是 0 和 1. 上述推导告诉我们

$$f((b_m b_{m-1} \cdots b_1 b_0)_2) = (\alpha \beta_{b_{m-1}} \beta_{b_{m-2}} \cdots \beta_{b_1} \beta_{b_0})_2. \quad (1.16)$$

好. 若我们以另一种方式改写式(1.12), 则在前面已看出这个型式:

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	$4\alpha + 2\gamma + \gamma$

例如, 当 $n = 100 = (1100100)_2$ 时, 如同前面那样, 我们原来的 Josephus 值 $\alpha = 1$, $\beta = -1$ 和 $\gamma = 1$ 产生

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

因为在 n 的表示中每组二进制数字位 $(10 \cdots 00)_2$ 被转换为

$$(1 \ -1 \cdots -1 \ -1)_2 = (00 \cdots 01)_2,$$

所以得出循环移性质。

因此, 表示法的改变给我们提供了一般递归式(1.15)的简洁的解(1.16). 若如实地不加约束, 则我们现在能更一般化. 递归式

$$\begin{aligned}
f(j) &= \alpha_j, \quad (1 \leq j < d); \\
f(dn + j) &= cf(n) + \beta_j, \quad (0 \leq j < d \text{ 且 } n \geq 1),
\end{aligned} \quad (1.17)$$

除从 d 进制的数开始, 且产生的值以 c 进制表示外, 与前面的一个相同. 也就是说, 它具

有基数改变的解

$$f((b_m b_{m-1} \cdots b_1 b_0)_d) = (\alpha_{b_m} \beta_{b_{m-1}} \beta_{b_{m-2}} \cdots \beta_{b_1} \beta_{b_0})_c.$$

例如, 假设走运, 给定递归式

$$f(1) = 34,$$

$$f(2) = 5,$$

$$f(3n) = 10f(n) + 76, (n \geq 1),$$

$$f(3n+1) = 10f(n) - 2, (n \geq 1),$$

$$f(3n+2) = 10f(n) + 8, (n \geq 1),$$

且假设我们要计算 $f(19)$ 。这里我们有 $d=3$ 和 $c=10$ 。现在 $19 = (201)_3$, 且基数改变的解告诉我们来完成 1 个数字位 1 个数字位地从基数 3 变到基数 10 的替换。所以第 1 位 2 变成 5, 0 和 1 变成 76 和 -2, 给出

$$f(19) = f((201)_3) = (576 - 2)_{10} = 1\,258,$$

这是我们的解答。

因此, Joscphus 和犹太人—古罗马人战争引出了一些有趣的一般递归式。

习 题

准备部分

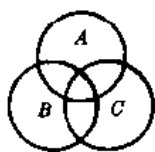
1. 所有的马是相同颜色的, 通过对一个给定集合中的马数进行归纳, 我们能证明这一点。这里是证明的方法: “若仅有一匹马, 则该马的颜色是相同的颜色, 所以基础是平凡的。对于归纳步, 假设有 n 匹编号 1 到 n 的马。根据归纳假设, 马 1 到 $n-1$ 是相同颜色, 且类似, 马 2 到 n 是相同颜色。但是中间 2 到 $n-1$ 的马, 当它们在不同的组中时, 不能改变颜色, 这些是马而不是变色龙, 所以根据传递性, 马 1 与马 n 同样一定是相同颜色。因此, 所有 n 匹马是相同颜色, 证毕。”用此推理, 如果有错误的话, 它的错误是什么?

2. 把 n 个圆盘组成的塔从左杆 A 转移到右杆 B , 如果 A 和 B 之间直接转移是不允许的, 找出移动的最短序列。(每次移动一定是移到中间杆或从中间杆移出。照例, 一个较大的盘一定不能出现在一个较小的盘上。)

3. 在前面习题的限制之下的转移过程中, 证明我们将实际遇到每一种适当堆放在三个杆上的 n 个盘的排列。

4. 在 Lucas 的原来规则下, 是否有三根杆上的任何起始和终止的形状, 它们相隔多于 $2^n - 1$ 次移动?

5. 3 个重叠圆的一个 “Venn 图” 常常用来说明结合于 3 个给定集的 8 个可能子集:



用4个重叠圆说明4个给定集，能否产生16个子集的可能性？

6. 平面中 n 条直线确定的有些区域是无限的，而其他的区域是有界的，有界区域的最大可能的个数是多少？

7. 设 $H(n) = J(n+1) - J(n)$ ，方程(1.8)告诉我们 $H(2n) = 2$ ，且 $H(2n+1) = J(2n+2) - J(2n+1) = (2J(n+1) - 1) - (2J(n) + 1) = 2H(n) - 2$ ，（对于所有 $n \geq 1$ ）。所以似乎可能对 n 归纳来证明对于所有 n ， $H(n) = 2$ 。这里的错误是什么？

课外习题

8. 解递归

$$Q_0 = \alpha; \quad Q_1 = \beta;$$

$$Q_n = (1 + Q_{n-1}) / Q_{n-2}, \quad (n > 1),$$

假设对于所有 $n \geq 0$ ， $Q_n \neq 0$ 。提示： $Q_4 = (1 + \alpha) / \beta$ 。

9. 有时用向后归纳是可能的，采用从 n 到 $n-1$ 来证明而不是从 $n-1$ 到 n 例如，考虑命题

$$P(n): x_1 \cdots x_n \leq \left(\frac{x_1 + \cdots + x_n}{n} \right)^n, \text{ 若 } x_1, \dots, x_n \geq 0.$$

由于 $(x_1 + x_2)^2 - 4x_1x_2 = (x_1 - x_2)^2 \geq 0$ ，所以当 $n=2$ 时，这是真的。

(a) 通过置 $x_n = (x_1 + \cdots + x_{n-1}) / (n-1)$ ，证明每逢 $n > 1$ ， $P(n)$ 蕴涵 $P(n-1)$ 。

(b) 证明 $P(n)$ 和 $P(2)$ 蕴涵 $P(2n)$ 。

(c) 阐明为什么对所有 n 蕴涵着 $P(n)$ 的真实性。

10. 设 Q_n 是把 n 个盘组成的塔从 A 转移到 B 所需的最小移动次数，如果所有移动一定是顺时针方向的，也就是说，从 A 到 B ，或从 B 到另一根杆，或从另一根杆到 A 。设 R_n 也是在此限制下从 B 离去再回到 A 所需的最小移动次数。证明

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

(你不需解这些递归式，在第七章中我们将看到如何做。)

11. 一个加倍的汉诺塔包含 n 种不同大小的 $2n$ 个盘，每种大小有2个。照例，要求我们一次仅移动1个盘，不把1个较大的盘放在1个较小的盘上。

(a) 若相同大小的盘彼此不加区别, 把一个加倍塔从一根杆转移到另一根需用多少次移动?

(b) 若要求我们再产生最后排列中所有相等大小盘的原来从上到下次序, 将会怎样?

[提示: 这是困难的, 实际它是一个“额外的问题”。]

12. 让我们更进一步推广习题 11(a), 假设有 m 种不同大小的盘, 且恰好有 n_k 个盘的大小为 k . 当考虑相等大小的盘是不能区别的盘时, 确定转移一个塔所需的最小移动次数 $A(n_1, \dots, n_m)$.

13. 由 n 个锯齿形线可确定的区域的最大个数是多少?



(每个锯齿形线包含由一直线段连接的两条平行的无限半直线).

14. 一块厚的干乳酪作 5 次直线切开, 你能获得多少份干乳酪? (当你作所有切割时, 干乳酪一定停留在它的原来的位置, 且每次切开对应于 3 维中的一个平面.) 找出由 n 个不同的平面所确定的 3 维区域的最大个数 P_n 的递归关系.

15. Josephus 有一个朋友, 由于他进入紧接最后位置而被救. 当每一个第 2 人被处死时, 倒数第 2 个幸存者的编号 $I(n)$ 是什么?

16. 用包含各种组合部分的方法来解一般 4 个参数的递归式

$$g(1) = \alpha;$$

$$g(2n + j) = 3g(n) + \gamma n + \beta_j, \quad (j = 0, 1 \text{ 且 } n \geq 1).$$

提示: 尝试函数 $g(n) = n$.

考查性问题

17. 若当有四根杆而不是三根时, 把 n 个盘组成的塔从一根杆转移到另一根杆所需移动的最小次数为 W_n , 证明

$$W_{n(n+1)/2} \leq 2W_{n(n-1)/2} + T_n, \quad (n > 0).$$

(这里 $T_n = 2^n - 1$ 是通常三根杆的次数.) 用此来找出一个闭形式 $f(n)$ 使得对于所有 $n \geq 0$, $W_{n(n+1)/2} \leq f(n)$.

18. 证明下列 n 条弯曲线的集合确定 Z_n 个区域, 其中 Z_n 由式 (1.7) 确定: 对于 $1 \leq j \leq n$, 第 j 条弯曲线的锯齿形转角在 $(n^{2j}, 0)$, 且上升通过点 $(n^{2j} - n^j, 1)$ 和 $(n^{2j} - n^j - n^{-n}, 1)$.

19. 当每个锯齿形转角是 30° 时, 是否可能用 n 条弯曲线获得 Z_n 个区域?

20. 用包含各种组成部分的方法来解一般 5 个参数的递归式

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

$$h(2n+j) = 4h(n) + \gamma_j n + \beta_j, \quad (j=0, 1 \text{ 且 } n \geq 1).$$

提示: 尝试函数 $h(n) = n$ 和 $h(n) = n^2$ 。

21. 假设在一个圆上有 $2n$ 个人, 前 n 个是“好人”, 而后 n 个是“坏人”。证明总有一个整数 m (依赖于 n), 使得若围绕圆我们着手处死所有第 m 个人, 所有坏人首先处死。(例如, 当 $n=3$ 时, 我们能取 $m=5$; 当 $n=4$ 时, 我们能取 $m=30$ 。)

额外的问题

22. 证明可能构造 n 个给定集合的所有 2^n 个可能子集的一个 Venn 图, 用 n 个凸多边形, 彼此叠合且在一个公共中心周围转动。

23. 假设 Josephus 发现他自己在一个给定位置 j , 但是他有机会指定消除的参数 q 使得处死所有第 q 个人, 他是否总能救出他自己?

研究性问题

24. 找出所有形为

$$X_n = \frac{a_0 + a_1 X_{n-1} + \cdots + a_k X_{n-k}}{b_1 X_{n-1} + \cdots + b_k X_{n-k}}$$

的递归关系, 它的解是周期的。

25. 通过证明习题 17 的关系中等式成立来解四根杆的汉诺塔问题的无限多情形。

26. 推广习题 23, 让我们说 $\{1, 2, \dots, n\}$ 的 Josephus 子集是 k 个数的集合, 使得对某个 q , 其他 $n-k$ 个数的人将首先被处死。(这些是 Josephus 要救的“好人”的 k 个位置。) 当 $n=9$ 时出现 2^9 个可能子集的 3 个子集是非 Josephus 的, 即 $\{1, 2, 5, 8, 9\}$, $\{2, 3, 4, 5, 8\}$ 和 $\{2, 5, 6, 7, 8\}$ 。当 $n=12$ 时有 13 个非 Josephus 集合, $n \leq 12$ 的任何其他值没有。是否对于大的 n 非 Josephus 子集是罕见的?

第二章 和

数学中到处有和，所以我们需要基本工具来处理它们。本章阐述表示法和方便地处理和的一般技巧。

2.1 表示法

在第一章中我们遇到前 n 个整数的和，我们写成 $1+2+3+\cdots+(n-1)+n$ ，在此公式中的 ‘ \cdots ’ 告诉我们完成由周围的项建立的型式。当然我们必须注意像 $1+7+\cdots+41.7$ 的和，没有一种显著的前后关系它是无意义的。另一方面，包含像 3 和 $(n-1)$ 的项过多了一点，若我们简单地记 $1+2+\cdots+n$ ，型式将推测清楚。有时我们甚至可能大胆到这样的程度而就写 $1+\cdots+n$ 。

我们将研究一般形式的和

$$a_1 + a_2 + \cdots + a_n, \quad (2.1)$$

其中每个 a_k 是以某种方式确定的一个数。这种表示法具有我们能“看出”整个和的优点，若我们有足够好的想象力，则好像全部把它写出来了。

和的每一个元素 a_k 称为一项。这些项常按照一个容易看出型式的公式含蓄地说明，即使是这样，我们有时也一定要把它们写成展开的形式，以致意思清楚。例如，若

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

意味着表示 n 项的和，而不是 2^{n-1} ，我们应把它更明确地记成

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

3 个小点表示法有许多用处，可是它能有多种解释，而且有点冗长。另一种表示法是可用的，就是值得注意的定界形式

$$\sum_{k=1}^n a_k, \quad (2.2)$$

称它为 Σ 表示法，因为它用了希腊字母 Σ (希腊语第十八个字母的大写)。此表示法告诉我

们在和中精确包含的那些项 a_k ，它的指标 k 是位于下界1和上界 n 之间所含的一个整数。口头上说，我们“就 k 从1到 n 求和”。在1820年Joseph Fourier引入了这个定界的 \sum 表示法，它立即使数学界大为倾倒。

顺便提一句， \sum 后面的数(这里为 a_k)称为被加数。

指标变量 k 称为式(2.2)中 \sum 记号的界限，因为 a_k 中的 k 与 \sum 表示法外边 k 的外表不相关。任何其他字母能替换这里的 k 而不改变式(2.2)的意思，常常用字母 i (也许因为它代表“指标”)来替换，但是由于知道 i 留作表示 $\sqrt{-1}$ ，所以我们一般将在 k 上求和。

结果是一个推广的 \sum 表示法比定界限的形式更有用：我们简单地在 \sum 之下写出一个或几个条件，来指定将作求和上面的指标集合。例如，式(2.1)和(2.2)中的和还能记成

$$\sum_{1 \leq k \leq n} a_k. \quad (2.3)$$

在此特殊的例子中，新形式与式(2.2)之间没有很大的不同，但是一般形式允许我们在不限于相继整数的指标集合上取和。例如，我们可表达100以下所有奇正整数平方的和如下：

$$\sum_{\substack{1 \leq k < 100 \\ k \text{ 奇数}}} k^2.$$

此和的定界限的等价式

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

是较不方便，且不太清楚的。类似，1和 N 之间的所有素数的倒数的和是

$$\sum_{\substack{p \leq N \\ p \text{ 素数}}} \frac{1}{p};$$

定界限形式要求我们记成

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

其中 p_k 表示第 k 个素数，且 $\pi(N)$ 是素数($\leq N$)的个数。(顺便提一句，此和给出接近 N 的一个随机整数的不同素因子的近似平均个数，因为关于这些整数的 $1/p$ 可被 p 整除。对于大的 N ，它的值近似地为 $\ln \ln N + 0.261\,972\,128$ ， $\ln x$ 代表 x 的自然对数，而 $\ln \ln x$ 代表 $\ln(\ln x)$ 。)

一般 \sum 表示法的最大优点是我们能比定界限的形式较容易地处理它。例如，假设我们要把指标变量 k 改变到 $k+1$ 。用一般形式，我们有

$$\sum_{1 \leq k \leq n} a_k = \sum_{1 \leq k+1 \leq n} a_{k+1},$$

易见这里发生的情况，且几乎不加思索我们能做此代换。但是用定界限形式，我们得到

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

较难看清发生的情况，且较可能引起错误。

另一方面，定界限形式也不是完全无用。它是好的和令人满意的，且我们能很快写出它，因为与式(2.3)的 8 个符号比较，式(2.2)有 7 个符号。所以当我们陈述一个问题或描述一个结果时，常用具有上界和下界的 \sum ，但是当我们处理一个和而它的指标需要变换时，我们宁愿用 \sum 下面的关系来表达。

在本书中 \sum 符号出现了 1 000 多次，所以可以肯定，我们确切知道它的意思。我们形式地把

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

记为所有项 a_k 的和的缩写，使得 k 是一个满足给定性质 $P(k)$ 的整数。(一个“性质 $P(k)$ ”是关于 k 的任何陈述，它能为真或为假。)我们将暂时假设仅有限多个整数 k 满足 $P(k)$ 具有 $a_k \neq 0$ ，否则无限多个非零和一起加入，事情就有点儿复杂。另一个极端，若 $P(k)$ 对于所有整数 k 为假，则我们得到一个“空的”和，一个空和的值被确定为零。

当一个和出现在书的一段中而不是一个展开的方程时，用式(2.4)的一种微小修改的形式：我们写成 ‘ $\sum_{P(k)} a_k$ ’，附加的性质像 \sum 的一个下标，以致公式将不伸出太多。类似，当我们要限于表示成一行时，‘ $\sum_{k=1}^n a_k$ ’ 是式(2.2)的一种合适选择。

常常吸引人们写成

$$\sum_{k=2}^{n-1} k(k-1)(n-k) \text{ 而不是 } \sum_{k=0}^n k(k-1)(n-k)$$

因为在此和中 $k=0, 1$ 和 n 的项为零。由于某种原因，看来合计 $n-2$ 项较有效而不是 $n+1$ 项。但是应抵制这样的引诱，计算的功效不同于理解的效能！我们将发现尽可能保持和的一个指标上的上下界简单是有利的，因为当界限简单时，处理和非常容易。实际上，形式 $\sum_{k=2}^{n-1}$ 甚至极易意义不清，因为当 $n=0$ 或 $n=1$ 时它的意思并不清楚(见习题 1)。零值项并不引起危害，而它们常避免一大堆麻烦。

到现在为止，我们已经讨论的表示法是十分标准的，但是现在我们即将产生一种基本背离传统的表示法。Kenneth Iverson 在他的程序设计语言 APL 中引入奇妙的概念^[16], p.11]，我们将看到它成功地简化了本书中我们要做的许多事情。此概念简单地把一个真或假的陈述放入方括号，若陈述为真，就表明结果是 1，若陈述是假，就表明结果是 0。例如，

$$[p \text{ 素数}] = \begin{cases} 1, & \text{若 } p \text{ 是一个素数;} \\ 0, & \text{若 } p \text{ 不是一个素数.} \end{cases}$$

Iverson 的约定允许我们用和的指标上不加约束的形式来表达和，因为我们能把式(2.4)改写成形式

$$\sum_k a_k [P(k)] . \quad (2.5)$$

若 $P(k)$ 为假，则项 $a_k [P(k)]$ 为零，所以我们能在被加项中间安全地包含它，这就使得容易处理和的指标，因为我们不必为界限条件骚扰。

需提到一个专门性事项：有时并不对所有的整数 k 定义 a_k ，我们可通过假设当 $P(k)$ 为假时 $[P(k)]$ 为“非常强的零”来克服这个困难。它是那么多的零，使得甚至当 a_k 为未定义时 $a_k [P(k)]$ 等于零。例如，如果我们用 Iverson 的约定把倒素数 ($\leq N$) 的和写成

$$\sum_p [p \text{ 素数}] [p \leq N] / p ,$$

当 $p=0$ 时用零来除就不成问题了，因为约定告诉我们 $[0 \text{ 素数}] [0 \leq N] / 0 = 0$ 。

让我们总结一下到目前为止关于和的讨论，有两种好的方法来表达项的一个和：一种方法用 ‘...’，另一种方法用 ‘ \sum ’。3 个小点的形式常呈现出有用的处理，特别组合邻近项，因为如果让整个和呈现在我们的眼前，则我们可能发现一种简化的型式。但是太琐碎也会过多而不合适。 \sum 表示法是紧凑的，给人深刻印象的记法，且常常启发那些以 3 个小点形式不明显的操作。当我们用 \sum 表示法处理时，零项一般无害；事实上，零常常使得 \sum 操作更容易。

2.2 和以及递归

好，我们现在明白了如何用所选的表示法来表达和。但是一个人如何实际去求得一个和的值呢？一种方法是注意到在和以及递归式之间的一种密切的关系。和

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

等价于递归

$$\begin{aligned} S_0 &= a_0 ; \\ S_n &= S_{n-1} + a_n ; (n > 0) . \end{aligned} \quad (2.6)$$

所以通过用我们在第一章中学到的方法来解递归式为闭形式，我们能把和计算为闭形式。

例如，若 a_n 等于一个常数加一个 n 的倍数，和的递归式(2.6)得到下列一般的形式：

$$\begin{aligned} R_0 &= \alpha ; \\ R_n &= R_{n-1} + \beta + \gamma n, (n > 0). \end{aligned} \quad (2.7)$$

如同第一章那样处理，我们求得 $R_1 = \alpha + \beta + \gamma$ ， $R_2 = \alpha + 2\beta + 3\gamma$ ，等等。一般解能写成

形式

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

其中 $A(n)$, $B(n)$ 和 $C(n)$ 是依赖于一般参数 α , β 和 γ 的系数。

包含各种组成部分的方法告诉我们来尝试对于 R_n 插入 n 的简单函数, 希望求出参数 α , β 和 γ , 此时解特别简单。置 $R_n = 1$ 意味着 $\alpha = 1$, $\beta = 0$, $\gamma = 0$, 因此

$$A(n) = 1.$$

置 $R_n = n$ 意味着 $\alpha = 0$, $\beta = 1$, $\gamma = 0$, 因此

$$B(n) = n.$$

置 $R_n = n^2$ 意味着 $\alpha = 0$, $\beta = -1$, $\gamma = 2$, 因此

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

我们得到 $C(n) = (n^2 + n)/2$ 是极容易的。

所以如果我们希望计算

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

则和的递归式(2.6)简化为式(2.7), 具有 $\alpha = \beta = a$, $\gamma = b$, 且解答为 $aA(n) + aB(n) + bC(n) = a(n+1) + b(n+1)n/2$.

相反, 许多递归能化为和; 所以本章后面学到的计算和的特殊方法将有助于解递归, 否则这些递归可能是困难的。汉诺塔递归是恰当的例子:

$$T_0 = 0,$$

$$T_n = 2T_{n-1} + 1, \quad (n > 0).$$

若我们用 2^n 除两边, 则能把它放进特殊形式(2.6):

$$\frac{T_n}{2^n} = 0,$$

$$\frac{T_n}{2^n} = \frac{T_{n-1}}{2^{n-1}} + \frac{1}{2^n}, \quad (n > 0).$$

现在可置 $S_n = T_n / 2^n$, 我们得到

$$S_0 = 0,$$

$$S_n = S_{n-1} + 2^{-n}, \quad (n > 0).$$

由此得到

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

(注意, 我们省去了此和的 $k=0$ 的项。) 在本章后面将导出几何级数 $2^{-1} + 2^{-2} + \dots + 2^{-n} = (1/2)^1 + (1/2)^2 + \dots + (1/2)^n$ 的和, 结果是 $1 - (1/2)^n$. 因此 $T_n = 2^n S_n = 2^n - 1$.

注意在此推导中, 我们用 2^n 除递归式而把 T_n 转换到 S_n . 此手段实际上是把形式

$$a_n T_n = b_n T_{n-1} + c_n \quad (2.9)$$

的任何递归化为和的一般技巧的特殊情形. 目的是用求和因子 s_n 乘两边:

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

巧妙地选择此因子 s_n 可使

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

于是如果我们记 $S_n = s_n a_n T_n$, 则得一个和的递归,

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

因此

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

原来递归(2.9)的解为

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

例如, 当 $n=1$, 我们得到 $T_1 = (s_1 b_1 T_0 + s_1 c_1) / s_1 a_1 = (b_1 T_0 + c_1) / a_1$.

但是我们如何能足够巧妙地找出真正的 s_n ? 没有问题: 展开关系 $s_n = s_{n-1} a_{n-1} / b_n$ 告诉我们, 因子

$$s_n = \frac{a_{n-1} a_{n-2} \cdots a_1}{b_n b_{n-1} \cdots b_2}, \quad (2.11)$$

或者此值的任何合适的常数倍数将是一个合适的求和因子. 例如, 汉诺塔递归有 $a_n = 1$ 和 $b_n = 2$. 我们刚才导出的一般方法表明: 如果要把递归式化为一个和, 则乘 $s_n = 2^{-n}$ 是有效的做法, 且并不需要卓越的灵机来发现这个乘数.

我们始终要小心不除零。每当所有 a 和所有 b 非零时，才能用求和因子方法。

让我们把这些想法应用到“快速分类”研究中提出的一个递归上去。快速分类是计算机内部分类数据的重要方法之一，当应用到随机次序的 n 个项时，快速分类所用的比较的平均次数满足递归

$$C_0 = 0;$$

$$C_n = n + 1 + \frac{2}{n} \sum_{k=0}^{n-1} C_k, \quad (n > 0). \quad (2.12)$$

比起前面见到的递归来说这看来十分害怕，它包含了所有前面值的和，且被 n 除。尝试小的情形给我们一些数据 ($C_1 = 2$, $C_2 = 5$, $C_3 = 26/3$)，但是并没减轻我们的害怕。

不过我们能有条不紊地简化式(2.12)的复杂性，首先把除法去掉，然后去掉 \sum 符号。这种方法是在两边乘 n ，得到关系式

$$nC_n = n^2 + n + 2 \sum_{k=0}^{n-1} C_k, \quad (n > 0);$$

因此，如果用 $(n-1)$ 替代 n ，则

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

我们现在可用第一个方程来减第二个方程， \sum 符号不再出现。

$$nC_n - (n-1)C_{n-1} = 2n + 2C_{n-1}, \quad (n > 1)$$

结果当 $n=1$ 时此关系也成立，因为 $C_1 = 2$ 。所以原来的 C_n 的递归式简化为十分简单的一个递归式：

$$C_0 = 0;$$

$$nC_n = (n+1)C_{n-1} + 2n, \quad (n > 0)$$

进行下去。现在我们能够应用一个求和因子，因为此递归式具有式(2.9)的形式，其中 $a_n = n$, $b_n = n+1$ 和 $c_n = 2n$ 。前页描述的一般方法告诉我们通过用

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

的某个倍数来乘此递归式。依照式(2.10)，解就是

$$C_n = 2(n+1) \sum_{k=1}^n \frac{1}{k+1}.$$

剩下的和十分相似于应用中常常出现的一个量。事实上，由于它时常出现，所以我们

给它取一个特殊名称和特殊记法:

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

字母 H 代表“调和”， H_n 是一个调和数，这样称呼是因为小提琴弦产生的第 k 个泛音是由一条 $1/k$ 倍长的弦产生的基音。

我们把 C_n 变成闭形式来结束快速分类递归式(2.12)的研究；若我们能用 H_n 表达 C_n ，则这将是可能的， C_n 的公式中的和是

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

把 k 改成 $k-1$ ，且修改边界条件，我们就能不太麻烦地把此式和 H_n 联系起来：

$$\begin{aligned} \sum_{1 \leq k \leq n} \frac{1}{k+1} &= \sum_{1 \leq k-1 \leq n} \frac{1}{k} = \sum_{2 \leq k \leq n+1} \frac{1}{k} \\ &= \left(\sum_{1 \leq k \leq n} \frac{1}{k} \right) - \frac{1}{1} + \frac{1}{n+1} = H_n - \frac{n}{n+1} . \end{aligned}$$

不错！我们已求出所需的和，完成了式(2.12)的解。当应用于 n 个随机次序的数据项时，快速分类所用的平均比较次数为

$$C_n = 2(n+1)H_n - 2n . \quad (2.14)$$

像往常一样，我们检验小的情形是正确的： $C_0 = 0$ ， $C_1 = 2$ ， $C_2 = 5$ 。

2.3 和的操作

和的操作方面成功的关键是把一个 \sum 改变成另一个较简单或更接近某个目标的能力。在学习了几个基本变换规则以及熟悉了它们的用法后容易做到这一点。

设 K 是任何有限的整数集合，可用三个简单的规则来变换 K 的元素上的和：

$$\sum_{k \in K} c a_k = c \sum_{k \in K} a_k ; \text{ (分配律)} \quad (2.15)$$

$$\sum_{k \in K} (a_k + b_k) = \sum_{k \in K} a_k + \sum_{k \in K} b_k ; \text{ (结合律)} \quad (2.16)$$

$$\sum_{k \in K} a_k = \sum_{p(k) \in K} a_{p(k)} . \text{ (交换律)} \quad (2.17)$$

分配律允许我们把常数移入和移出 \sum ；结合律允许我们把一个 \sum 分成两部分，或者把两个 \sum 合成一个；交换律表明我们能以任何想要的形式重排各项，此处 $p(k)$ 是所有整数的集合

的任何排列。例如，若 $K = \{-1, 0, +1\}$ ，且如果 $p(k) = -k$ ，这样三个定律分别告诉我们：

$$ca_{-1} + ca_0 + ca_1 = c(a_{-1} + a_0 + a_1) ; \text{ (分配律)}$$

$$(a_{-1} + b_{-1}) + (a_0 + b_0) + (a_1 + b_1)$$

$$= (a_{-1} + a_0 + a_1) + (b_{-1} + b_0 + b_1) ; \text{ (结合律)}$$

$$a_{-1} + a_0 + a_1 = a_1 + a_0 + a_{-1} . \text{ (交换律)}$$

第一章中 Gauss 的巧妙方法可视为这样三个基本定律的一种应用。假设我们要计算一个算术级数的一般形式，

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

根据交换律，我能用 $n-k$ 替代 k ，得到

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

利用结合律可把这两个方程相加：

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

而我们现在能应用分配律，且计算出一个平凡的和：

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

两边各除 2，我们证明了

$$\sum_{k=0}^n (a + bk) = \left(a + \frac{1}{2}bn \right) (n+1) . \quad (2.18)$$

右端可记成第一项和最后项的平均值，即 $\frac{1}{2}(a + (a + bn))$ ，乘上项数，即 $n+1$ 。

重要的是牢记一般交换律(2.17)中的函数 $p(k)$ 假设为所有整数的一种排列。换句话说，对于每一个整数 n ，恰好就有一个整数 k 使得 $p(k) = n$ 。否则交换律可不成立，习题 3 完全说明了这一点。像 $p(k) = k+c$ 或 $p(k) = c-k$ 的变换总是排列，其中 c 是一个整常数，所以它们总是成立的。

另一方面，我们能稍放宽排列的限制：我们仅需要求当 n 是指标集合 K 的一个元素时，恰好有一个整数 k 具有 $p(k) = n$ 。若 $n \notin K$ (也就是说，若 n 不在 K 中)， $p(k) = n$ 出现多少次都无关紧要，因为这样的 k 不参与此和。因此，例如，我们能表明

$$\sum_{\substack{k \in K \\ k \text{ 偶}}} a_k = \sum_{\substack{n \in K \\ n \text{ 偶}}} a_n = \sum_{\substack{2k \in K \\ 2k \text{ 偶}}} a_{2k} = \sum_{2k \in K} a_{2k} . \quad (2.19)$$

因为当 $n \in K$ 且 n 为偶数时, 恰好有一个 k 使得 $2k = n$.

Iverson 约定允许我们从公式中间的逻辑语句得到值 0 或 1, 它与分配律、结合律和交换律一起可用来推出另外的和的性质. 例如, 这是一个结合不同指标集合的重要规则: 若 K 和 K' 是任何整数集合, 则

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

这是根据一般公式

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

和

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

得出的. 我们典型地把规则(2.20)用到结合两个几乎不相交的指标集合, 如同

$$\sum_{k=1}^m a_k + \sum_{k=m}^n a_k = a_m + \sum_{k=1}^n a_k, \quad (1 \leq m \leq n);$$

或者从一个和分出一项, 如同

$$\sum_{0 \leq k \leq n} a_k = a_0 + \sum_{1 \leq k \leq n} a_k, \quad (n \geq 0). \quad (2.23)$$

此分出一项的操作是摄动法的基础, 它常常允许我们计算一个和为闭形式. 想法是从一个未知和且称它为 S_n 开始,

$$S_n = \sum_{0 \leq k \leq n} a_k,$$

(取名并解出它)然后我们以两种方式改写 S_{n+1} , 通过既分出它的最后一项又分出它的第一项:

$$\begin{aligned} S_n + a_{n+1} &= \sum_{0 \leq k \leq n+1} a_k = a_0 + \sum_{1 \leq k \leq n+1} a_k \\ &= a_0 + \sum_{1 \leq k+1 \leq n+1} a_{k+1} \end{aligned}$$

$$= a_0 + \sum_{0 \leq k \leq n} a_{k+1} . \quad (2.24)$$

现在我们在此最后和上处理，且尝试由 S_n 表达它。若成功，则可获得一个方程，它的解是我们所要找的和。

例如，让我们用此方法来求出一般几何级数的和，

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

式(2.24)中的一般摄动方法告诉我们

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

且根据分配律，右边的和是 $x \sum_{0 \leq k \leq n} ax^k = xS_n$ 。所以 $S_n + ax^{n+1} = a + xS_n$ ，且我们能解出 S_n 而得到

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

(当 $x = 1$ 时，和当然简单地为 $(n+1)a$ 。)右边可记为和中的第一项减不在和中(最后项后面的所有项)的第一项，除以 1 减 x (项的比)。

这似乎太容易了。让我们在一个稍困难的和上试用摄动技巧，

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

此时，我们有 $S_0 = 0$, $S_1 = 2$, $S_2 = 10$, $S_3 = 34$, $S_4 = 98$ ，一般公式是什么呢？按照式(2.24)我们有

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

所以我们要用 S_n 来表达右边的和。好，我们能借助结合律把它分为两个和，

$$\sum_{0 \leq k \leq n} k2^{k+1} + \sum_{0 \leq k \leq n} 2^{k+1} ,$$

剩下和的第一个是 $2S_n$ 。另一个和是几何级数，根据式(2.25)它等于 $(2 - 2^{n+2}) / (1 - 2) = 2^{n+2} - 2$ 。所以我们得到 $S_n + (n+1)2^{n+1} = 2S_n + 2^{n+2} - 2$ ，而代数计算得到

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

现在我们明白为什么 $S_3 = 34$: 它是 $32+2$, 而不是 $2 \cdot 17$.

用 x 替换 2 , 相似推导给出方程 $S_n + (n+1)x^{n+1} = xS_n + (x - x^{n+2})/(1-x)$, 因此我们能推出

$$\sum_{k=0}^n kx^k = \frac{x - (n+1)x^{n+1} + nx^{n+2}}{(1-x)^2}, \quad (x \neq 1). \quad (2.26)$$

注意到我们能以完全不同的方法推导出这个闭形式是有趣的, 此方法运用了微分的基本技巧. 若我们从方程

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

开始, 且两边对 x 求导, 我们得到

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

因为和的求导是它的各项求导的和. 在后续各章中, 我们将看到微积分和离散数学之间许多其他的联系.

2.4 多重和

一个和的项可能由两个或多个指标来指定, 而不仅仅由一个指标指定. 例如, 这里是 9 项的一个二重和, 由两个指标 j 和 k 产生:

$$\begin{aligned} \sum_{1 \leq j, k \leq 3} a_j b_k &= a_1 b_1 + a_1 b_2 + a_1 b_3 + a_2 b_1 + a_2 b_2 + a_2 b_3 \\ &\quad + a_3 b_1 + a_3 b_2 + a_3 b_3. \end{aligned}$$

对于这样的和, 我们用相同的表示法和方法, 就像我们对一个指标的和所做的那样. 因此, 若 $P(j, k)$ 是 j 和 k 的一个性质, 则所有项 $a_{j, k}$ 的和使得 $P(j, k)$ 是真能以两种方式写出, 一种用 Iverson 的约定, 且在所有整数对 j 和 k 上求和,

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

虽然求和的指标多于一个, 仅需一个 \sum 符号. \sum 表示在所用的所有指标的组合上的一个和.

当讨论到几个和的总和，我们也有必要用两个 \sum 。例如，

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

是关于

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

的缩写，它是 $\sum_k a_{j,k} [P(j, k)]$ 的所有整数 j 上的和，后者是在所有 $P(j, k)$ 为真的项 $a_{j,k}$ 的所有整数 k 上的和。此时，我们说二重和“先在 k 上求和”，依赖于多个指标的一个和能先在它的指标的任何一个上求和。

关于此事我们有一个称为交换求和次序的基本定律，它推广了前面见到的结合律(2.16):

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

此定律的中间项是两个指标上的一个和。左边， $\sum_j \sum_k$ 代表先在 k 上求和，然后在 j 上求和。右边， $\sum_k \sum_j$ 代表先在 j 上求和，然后在 k 上求和。实际上，当我们要以闭形式求二重和的值，通常较易首先在一个指标上求它的和而不是在另一个指标上；我们开始选取的是较方便的那一个。

没有理由对于和的和惊慌失措，但是对于初学者来说它们能呈现混淆，所以让我们举一些另外的例子。我们开始给的9项和提供了一种处理二重和的好的说明，因为那个和实际上可被简化，而简化过程是我们对 $\sum \sum$ 所做的典型的过程：

$$\begin{aligned} \sum_{1 \leq j, k \leq 3} a_j b_k &= \sum_{j, k} a_j b_k [1 \leq j, k \leq 3] = \sum_{j, k} a_j b_k [1 \leq j \leq 3] [1 \leq k \leq 3] \\ &= \sum_j \sum_k a_j b_k [1 \leq j \leq 3] [1 \leq k \leq 3] \\ &= \sum_j a_j [1 \leq j \leq 3] \sum_k b_k [1 \leq k \leq 3] = \sum_j a_j [1 \leq j \leq 3] \left(\sum_k b_k [1 \leq k \leq 3] \right) \\ &= \left(\sum_j a_j [1 \leq j \leq 3] \right) \left(\sum_k b_k [1 \leq k \leq 3] \right) \\ &= \left(\sum_{j=1}^3 a_j \right) \left(\sum_{k=1}^3 b_k \right) . \end{aligned}$$

这里的第一行表示没有特定次序的9项的和。第二行把它们分为三组， $(a_1 b_1 + a_1 b_2 + a_1 b_3) + (a_2 b_1 + a_2 b_2 + a_2 b_3) + (a_3 b_1 + a_3 b_2 + a_3 b_3)$ 。第三行把分配律用到因子上去，括出 a_j ，因为 a_j 和 $[1 \leq j \leq 3]$ 不依赖于 k ；这就给出 $a_1(b_1 + b_2 + b_3) + a_2(b_1 + b_2 + b_3) + a_3(b_1 + b_2 + b_3)$ 。第四行与第三行相同，但是添加了多余的一对

括号以致第五行看来不致难以理解。第五行提出对于每个 j 值出现的公因子 $(b_1 + b_2 + b_3)$: $(a_1 + a_2 + a_3)(b_1 + b_2 + b_3)$ 。最后行仅是另一种方式写出前面的一行。此推导方法能用来证明一般的分配律,

$$\sum_{\substack{j \in J \\ k \in K}} a_j b_k = \left(\sum_{j \in J} a_j \right) \left(\sum_{k \in K} b_k \right), \quad (2.28)$$

对于所有指标集 J 和 K 成立。

交换求和次序的基本定律式(2.27)有许多变形, 当我们要限制指标的范围而不是在所有整数 j 和 k 上求和时它就出现。这些变形形成两种形式, 第一种形式:

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

这就是以另一种方法写出的式(2.27), 因为 Iverson 的 $[j \in J, k \in K]$ 因子分解为 $[j \in J][k \in K]$, 每逢 j 和 k 的范围彼此独立时应用此定律。

关于交换的另一个公式有一点复杂, 当内层和的范围依赖于外层和的指标变量时应用它:

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

这里一定要以这样一种方式把集合 J , $K(j)$, K' 和 $J'(k)$ 联系起来, 即

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

像这样的因子分解原则上总是可能的, 因为我们可设 $J = K'$ 是所有整数的集合, 且 $K(j) = J'(k)$ 是支配二重和的基本性质 $P(j; k)$ 。但是有重要的特殊情形, 在此情形中集合 J , $K(j)$, K' 和 $J'(k)$ 有简单的形式, 在应用中常出现这些情形。例如, 此处是一种特别有用的因子分解:

$$[1 \leq j \leq n][j \leq k \leq n] = [1 \leq j \leq k \leq n] = [1 \leq k \leq n][1 \leq j \leq k]. \quad (2.31)$$

这个 Iverson 方程允许我们写成

$$\sum_{j=1}^n \sum_{k=j}^n a_{j, k} = \sum_{1 \leq j \leq k \leq n} a_{j, k} = \sum_{k=1}^n \sum_{j=1}^k a_{j, k}. \quad (2.32)$$

这两个和的和之一通常比另一个容易计算, 我们可用式(2.32)把难的转换成容易的。

让我们把这些想法应用到一个有用的例子上去。考虑 n^2 个 $a_j a_k$ 的数组

$$\begin{bmatrix} a_1 a_1 & a_1 a_2 & a_1 a_3 & \cdots & a_1 a_n \\ a_2 a_1 & a_2 a_2 & a_2 a_3 & \cdots & a_2 a_n \\ a_3 a_1 & a_3 a_2 & a_3 a_3 & \cdots & a_3 a_n \\ \vdots & \vdots & \vdots & & \vdots \\ a_n a_1 & a_n a_2 & a_n a_3 & \cdots & a_n a_n \end{bmatrix}.$$

我们的目标将是求出此数组的主对角线或主对角线上的所有元素之和

$$S_q = \sum_{1 \leq j < k \leq n} a_j a_k$$

的一个简单公式。因为 $a_j a_k = a_k a_j$ ，所以数组对于它的主对角线是对称的， S_q 将近似于所有元素和的一半(除了计及主对角线的无用的因子)。

这些考虑导出下列的处理。我们有

$$S_q = \sum_{1 \leq j < k \leq n} a_j a_k = \sum_{1 \leq k < j \leq n} a_k a_j = \sum_{1 \leq k < j \leq n} a_j a_k = S_b,$$

因为我们能给 (j, k) 再命名为 (k, j) 。此外，由于

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

我们得到

$$2S_q = S_q + S_b = \sum_{1 \leq j, k \leq n} a_j a_k + \sum_{1 \leq j=k \leq n} a_j a_k,$$

根据一般分配律 (2.28)，第一个和为 $\left(\sum_{j=1}^n a_j\right)\left(\sum_{k=1}^n a_k\right) = \left(\sum_{k=1}^n a_k\right)^2$ ，第二个和为 $\sum_{k=1}^n a_k^2$ 。所以我们有

$$S_q = \sum_{1 \leq j < k \leq n} a_j a_k = \frac{1}{2} \left(\left(\sum_{k=1}^n a_k \right)^2 + \sum_{k=1}^n a_k^2 \right), \quad (2.33)$$

这是由较简单的单重和表达上三角形和的一个表达式。

由此成功鼓励我们看另一个二重和：

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

当交换 j 和 k 时，也有对称性：

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

所以利用等式

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

我们可把 S 本身相加而推出

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

这里的第二个和是零, 第一个和如何? 它展开成四个分离的和, 每一个和是第一种形式:

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

在最后一步中, 按照一般的分配律(2.28)简化了两个和, 若第一个和的处理看来难以理解, 这里再详细地简化一下:

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

要是我们乘留下的变指标集的大小(这里为 n), 则能简单地消去和中不出现的一个指标变量(这里为 j).

回到我们刚才讨论的地方, 现在能两边除以 2 而得到一个有趣的公式:

$$\left(\sum_{k=1}^n a_k \right) \left(\sum_{k=1}^n b_k \right) = n \sum_{k=1}^n a_k b_k - \sum_{1 \leq j < k \leq n} (a_k - a_j)(b_k - b_j), \quad (2.34)$$

此等式产生了作为一种特殊情形的 Chebyshev 的和不等式:

$$\begin{aligned} \left(\sum_{k=1}^n a_k \right) \left(\sum_{k=1}^n b_k \right) &\leq n \sum_{k=1}^n a_k b_k, \quad a_1 \leq \cdots \leq a_n, \quad b_1 \leq \cdots \leq b_n; \\ \left(\sum_{k=1}^n a_k \right) \left(\sum_{k=1}^n b_k \right) &\geq n \sum_{k=1}^n a_k b_k, \quad a_1 \leq \cdots \leq a_n, \quad b_1 \geq \cdots \geq b_n. \end{aligned}$$

(一般, 若 $a_1 \leq \cdots \leq a_n$ 以及若 p 是 $\{1, \cdots, n\}$ 的一个排列, 则可能证明当 $b_{p(1)}$

$\leq \dots \leq b_{p(n)}$ 时 $\sum_{k=1}^n a_k b_{p(k)}$ 的最大值出现, 当 $b_{p(1)} \geq \dots \geq b_{p(n)}$ 时最小值出现。))

多重求和与单个和中改变求和的指标的一般操作有一种有趣的联系。如果 $p(k)$ 是整数的任何排列, 由交换律我们知道

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

但是当我们用 $f(j)$ 替换 k , 其中 f 是一个任意函数

$$f: J \rightarrow K$$

取一个整数 $j \in J$ 进入一个整数 $f(j) \in K$ 时, 发生什么情况? 指标替换的一般公式是

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

其中 $\#f^{-}(k)$ 表明集合

$$f^{-}(k) = \{j | f(j) = k\}$$

中的元素个数, 也就是使得 $f(j)$ 等于 k 的 $j \in J$ 的值的个数。

通过交换求和的次序容易证明式(2.35), 由于 $\sum_{j \in J} [f(j) = k] = \#f^{-}(k)$,

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

在 f 是 J 和 K 之间的一个一一对应的特殊情形中, 对于所有 k 我们有 $\#f^{-}(k) = 1$, 且一般公式(2.35)化成

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

这就是我们以前提到的交换律(2.17), 有点不易识别。

到目前为止, 我们的多重和的例子全用了像 a_k 或 b_k 的一般项。但是本书被认为是具体的, 所以让我们看一看实际数的一个多重和:

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

$$\text{例如, } S_1 = 0; \quad S_2 = 1; \quad S_3 = \frac{1}{2-1} + \frac{1}{3-1} + \frac{1}{3-2} = \frac{5}{2} .$$

计算一个双重和的标准方法是先对 j 或先对 k 求和, 所以让我们讨论两种选取。

$$S_n = \sum_{1 \leq k \leq n} \sum_{1 \leq j < k} \frac{1}{k-j} \quad \text{先对 } j \text{ 求和}$$

$$\begin{aligned}
&= \sum_{1 \leq k \leq n} \sum_{1 \leq k-j < k} \frac{1}{j} && \text{用 } k-j \text{ 替换 } j \\
&= \sum_{1 \leq k \leq n} \sum_{0 < j \leq k-1} \frac{1}{j} && \text{简化 } j \text{ 的界限} \\
&= \sum_{1 \leq k \leq n} H_{k-1} && \text{依据 } H_{k-1} \text{ 的定义式(2.13)} \\
&= \sum_{1 \leq k+1 \leq n} H_k && \text{用 } k+1 \text{ 替换 } k \\
&= \sum_{0 \leq k < n} H_k. && \text{简化 } k \text{ 的界限}
\end{aligned}$$

我们不知道如何得到调和数的和的闭形式。

若尝试先对 k 求和, 则得到

$$\begin{aligned}
S_n &= \sum_{1 \leq j \leq n} \sum_{j < k \leq n} \frac{1}{k-j} && \text{先对 } k \text{ 求和} \\
&= \sum_{1 \leq j \leq n} \sum_{j < k+j \leq n} \frac{1}{k} && \text{用 } k+j \text{ 替换 } k \\
&= \sum_{1 \leq j \leq n} \sum_{0 < k \leq n-j} \frac{1}{k} && \text{简化 } k \text{ 的界限} \\
&= \sum_{1 \leq j \leq n} H_{n-j} && \text{由 } H_{n-j} \text{ 的定义式(2.13)} \\
&= \sum_{1 \leq n-j \leq n} H_j && \text{用 } n-j \text{ 替换 } j \\
&= \sum_{0 \leq j < n} H_j && \text{简化 } j \text{ 的界限}
\end{aligned}$$

我们回到了相同的绝境。

但是有另一种方法来处理, 若在决定把 S_n 化为和的和之前用 $k+j$ 替换 k :

$$\begin{aligned}
S_n &= \sum_{1 \leq j < k \leq n} \frac{1}{k-j} && \text{再复制给定的和} \\
&= \sum_{1 \leq j < k+j \leq n} \frac{1}{k} && \text{用 } k+j \text{ 替换 } k \\
&= \sum_{1 \leq k \leq n} \sum_{1 \leq j \leq n-k} \frac{1}{k} && \text{先对 } j \text{ 求和} \\
&= \sum_{1 \leq k \leq n} \frac{n-k}{k} && \text{对 } j \text{ 求和是平凡的}
\end{aligned}$$

$$\begin{aligned}
 &= \sum_{1 \leq k \leq n} \frac{n}{k} - \sum_{1 \leq k \leq n} 1 && \text{依据结合律} \\
 &= n \left(\sum_{1 \leq k \leq n} \frac{1}{k} \right) - n && \text{啊呀!} \\
 &= nH_n - n. && \text{由 } H_n \text{ 的定义式(2.13)}
 \end{aligned}$$

我们找到了 S_n ，把它和冒失的开端结合起来，作为一个额外的结果我们得到一个进一步的等式：

$$\sum_{0 \leq k \leq n} H_k = nH_n - n. \quad (2.36)$$

我们能理解这里以两种方法所用的技巧，一个代数的方法和一个几何的方法。(1)用代数方法，若我们有一个双重和，它的项涉及 $k+f(j)$ ，其中 f 是一个任意的函数，此例表明试用 $k-f(j)$ 替换 k 且对 j 求和是好的主意。(2)用几何方法，我们可看如下的这个特殊和 S_n ，此时 $n=4$ ，

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

我们首先尝试，先对 j (列)或对 k (行)求和，给出 $H_1 + H_2 + H_3 = H_3 + H_2 + H_1$ ，获胜的主意实质上是按对角线求和，得到 $3/1+2/2+1/3$ 。

2.5 一般的方法

现在我们通过从不同的几个角度观察一个例子来巩固已学的内容。在以下几页，我们将尝试找前 n 个平方和的一个闭形式，称它为 \square_n ：

$$\square_n = \sum_{0 \leq k \leq n} k^2, \quad n \geq 0. \quad (2.37)$$

我们将看到至少有 7 种不同的方法来解这个问题，且在处理中我们将学到一般处理和的有用的策略。

照例我们先看一些小的情形。

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

没有 \square_n 的闭形式立即清楚,但是当我们在寻找一个闭形式时,能用这些值来检验。

方法 0: 在手册中查找

像前 n 个平方和的问题很可能以前已经解出,所以我们很可能在手边的一本参考书中找到解。果然在CRC出版社的标准数学表^[24]的第72页有解答:

$$\square_n = \frac{n(n+1)(2n+1)}{6}, \quad n \geq 0. \quad (2.38)$$

仅为了查明我们没有看错它,检验此公式正确地给出口 $_5 = 5 \cdot 6 \cdot 11 / 6 = 55$ 。顺便提一句,CRC表的第72页还有关于立方和, ..., 10次方和的资料。

数学公式的权威性的参考书是由Abramowitz和Stegun^[2]编辑的数学函数手册。在此书的第813页至814页列出了关于 $n \leq 100$ 的 \square_n 的值;第804和809页展示了等价于式(2.38)的公式,加之立方和, ..., 15次方和的相似公式,交错正负号或不交错正负号。

但是关于序列的问题的解答的最好来源是由Sloane^[270]编的一本称为整序列手册的令人惊异的小书,书中列出了几千个由数值表明的序列。如果你提出一个你所怀疑的已经研究过的递归,则你仅需计算足够多的项从其他著名的递归中识别出你的递归,然后你将可能找到Sloane手册中有关文献的一个指示数。例如,1, 5, 14, 30, ..., 原来是Sloane的序列数1574,且称它为“正方角锥数”(因为在一个具有 n^2 个球的正方底的角锥有 \square_n 个球)。Sloane提出三本参考书,一本就是我们已经提到的Abramowitz和Stegun的手册。

还有另一种方法来查世界上积累的数学知识是用一个计算机程序(譬如MACSYMA),它提供了符号操作的工具。这样的程序尤其对于需要与大的公式打交道的人来说是必不可少的。

熟悉标准的资料来源是有益的,因为它们很有用。但是方法0实际上与本书的精神不一致,因为我们要了解如何解出解答。查表的方法限于用在这样一些问题,其他人已判定这些问题是有价值研究的问题或在那里无新的问题。

方法 1: 猜测解答,用归纳法证明

也许有人私下告诉我们一个问题的解答,或者用某个不太严格的方法得出一个闭的形式,于是我们仅需证明它是正确的。

例如,我们可能注意到 \square_n 的值有相当小的素因子,所以我们可能提出公式(2.38),就像对于所有小的 n 值成立的那样。我们也可能猜测等价的公式

$$\square_n = \frac{n \left(n + \frac{1}{2} \right) (n+1)}{3}, \quad n \geq 0, \quad (2.39)$$

因为它较容易记忆, 所以此式较好. 在这里, 明显偏重于支持式(2.39), 但是无疑, 我们一定要证明我们的猜测. 为此目的想出了数学归纳法.

在 Judge Wapner 的极多的成果中, 他认为以下论证正确, “好吧, 先生, 我们知道 $\square_0 = 0 = 0(0+1/2)(0+1)/3$, 所以基础是不费力的. 对于归纳法来说, 假设 $n > 0$, 且设当 $n-1$ 替换 n 时公式(2.39)成立. 由于

$$\square_n = \square_{n-1} + n^2,$$

我们得到

$$\begin{aligned} 3\square_n &= (n-1)\left(n-\frac{1}{2}\right)(n) + 3n^2 \\ &= \left(n^3 - \frac{3}{2}n^2 + \frac{1}{2}n\right) + 3n^2 \\ &= \left(n^3 + \frac{3}{2}n^2 + \frac{1}{2}n\right) \\ &= n\left(n + \frac{1}{2}\right)(n+1). \end{aligned}$$

所以无疑对于所有 $n \geq 0$, 公式(2.39)确实成立.”

归纳法有它的地位, 且它比尝试查找解答更有说服力, 但是我们实际还未去寻找解答. 到现在为止, 本章中已求的所有其他和都没有用归纳法解决, 像 \square_n 那样我们同样能够从头开始决定一个和. 灵机的闪现将是不必要的, 甚至在目前较小创造性的时候, 我们能够求和.

方法 2: 摄动和

同样, 让我们回到对几何级数式(2.25)工作得很好的摄动法. 为了取得 \square_n 的一个方程, 我们取出 \square_{n+1} 的第一项和最后一项:

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

\square_n 彼此消去了. 尽管我们作了最大努力, 偶然摄动法产生像 $\square_n = \square_n$ 的结果, 所以我们的工作白费了.

另一方面, 这个推导的结果并不是一点没有用, 它展示了一种方法, 把前 n 个整数相加为闭形式,

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

纵然我们希望发现前面整数平方的和。如果我们从整数立方的和开始，称它为 $\sum_{k=0}^n k^3$ ，能否取得整数平方的一个表达式？让我们试试看。

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

确实被消去，且我们有足够的资料来确定 $\sum_{k=0}^n k^2$ ，而不依赖于归纳法：

$$\begin{aligned} 3 \sum_{k=0}^n k^2 &= (n+1)^3 - 3(n+1)n/2 - (n+1) \\ &= (n+1) \left(n^2 + 2n + 1 - \frac{3}{2}n - 1 \right) = (n+1) \left(n + \frac{1}{2} \right) n. \end{aligned}$$

方法 3: 建立所有组合部分

对于涉及 n^2 的被加数，稍加推广递归式(2.7)也就够了。

$$R_0 = \alpha;$$

$$R_n = R_{n-1} + \beta + \gamma n + \delta n^2, \quad n > 0, \quad (2.40)$$

的解将是一般形式的解

$$R_n = A(n)\alpha + B(n)\beta + C(n)\gamma + D(n)\delta; \quad (2.41)$$

且我们已经确定了 $A(n)$, $B(n)$ 和 $C(n)$ ，因为当 $\delta=0$ 时，式(2.41)和(2.7)相同。若现在插入 $R_n = n^3$ ，我们发现当 $\alpha=0$, $\beta=1$, $\gamma=-3$, $\delta=3$ 时 n^3 是解。因此

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

这样就确定了 $D(n)$ 。

我们对和 $\sum_{k=0}^n k^2$ 感兴趣，它等于 $\sum_{k=0}^{n-1} k^2 + n^2$ 。于是，如果在式(2.41)中置 $\alpha=\beta=\gamma=0$ 和 $\delta=1$ ，则取得 $\sum_{k=0}^n k^2 = R_n$ ，所以 $\sum_{k=0}^n k^2 = D(n)$ 。我们不需要做从 $B(n)$ 和 $C(n)$ 计算 $D(n)$ 的代数运算，因为我们已经知道了解答是什么；但是找到

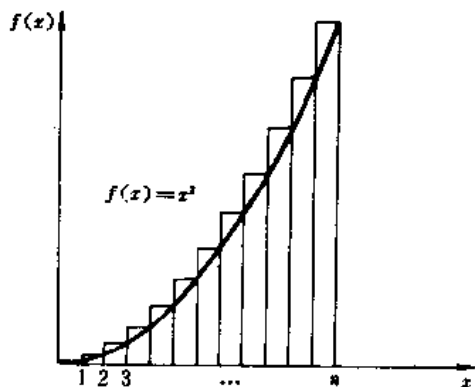
$$3D(n) = n^3 + 3C(n) - B(n) = n^3 + 3 \frac{(n+1)n}{2} - n = n \left(n + \frac{1}{2} \right) (n+1),$$

将使怀疑者消除疑虑。

方法 4: 用积分替换和

提出用微积分代替离散数学的人们倾向于熟悉 \int , 所以他们自然尝试把 \sum 改成 \int , 本书的目标之一是使 \sum 用得相当不错, 我们想象(至少对于确切结果来说) \int 比 \sum 困难. 因为和与积分是基于十分相似的思想, 所以探讨 \sum 和 \int 之间的关系仍然是一个好的主意.

在微积分中, 积分可视为曲线下的面积, 且我们能把接触曲线的长的小矩形加起来近似此面积. 如果给出长的小矩形的集合, 我们还能进行另一种方式. 由于 \square_n 是大小为 $1 \times 1, 1 \times 4, \dots, 1 \times n^2$ 的矩形面积之和, 所以它近似等于 0 和 n 之间曲线 $f(x) = x^2$ 之下的面积.



此曲线下的面积为 $\int_0^n x^2 dx = n^3 / 3$, 所以我们知道 \square_n 近似为 $(1/3)n^3$.

用这个事实的一种方法是研究近似的误差 $E_n = \square_n - (1/3)n^3$. 由于 \square_n 满足递归式 $\square_n = \square_{n-1} + n^2$, 我们发现 E_n 满足较简单的递归

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

寻找积分近似的另一种方法是通过相加楔形误差项来找 E_n 的公式. 我们得到

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

不管哪一种方法, 我们能找出 E_n , 然后求得 \square_n .

方法 5: 伸展和收缩

还有另一种方法来找 \square_n 的闭形式, 用看来较复杂的双重和来替换原来的和, 如果我们适当地来回处理双重和, 则实际上能简化它:

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

首先, 从单个和变成双重和好像是倒退的一步, 但是它实际上是前进了, 因为它产生的和较容易处理。我们不能期望通过不断的简化, 简化, 再简化来解每一个问题: 你不能仅向上爬达到最高山峰的顶点!

方法 6: 用有限演算

方法 7: 用母函数

随着在下一节以及后续各章中进一步学到一些技巧, 我们还要找出另外的 $\square_n = \sum_{k=0}^n k^2$ 的振奋人心的计算。

2.6 有限和无限演算

我们已经学了直接处理和的各种方法。现在是通过以更高的高度来看求和问题, 来学习更广泛的观点的时候了。数学家建立了“有限演算”, 相似于更传统的无限演算, 通过它可能以好的有规则的方式来处理求和。

无限演算是基于由

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

定义的微分算子 D 的性质上。有限演算是基于由

$$\Delta f(x) = f(x+1) - f(x) \tag{2.42}$$

定义的差分算子 Δ 的性质上。这是微商的有限类似物, 在此微商中我们限制了 h 为正整

数值。于是，当 $h \rightarrow 0$ 时， $h=1$ 是我们能达到“极限”的最接近的值，且当 $h=1$ 时， $\Delta f(x)$ 是 $(f(x+h)-f(x))/h$ 的值。

符号 D 和 Δ 称为算子，因为它们在函数上运算得到新的函数；它们是产生函数的函数的函数。若 f 是实数到实数的适当的平滑函数，则 Df 也是从实数到实数的一个函数。若 f 是任何实到实的函数，则 Δf 也是。在点 x 处的 Df 和 Δf 的函数值由上述定义给定。

在微积分的初期，我们学了 D 在幂函数 $f(x)=x^m$ 上的运算。此时 $Df(x)=mx^{m-1}$ 。通常我们写出它时略去 f

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

如果 Δ 算子将产生一个同样优美的结果，则将是很好的；不幸，它不行。例如，我们得到

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

但是有一类“ m 次幂”，在 Δ 之下能很好地变换，这就是为什么对有限演算感兴趣。由规则

$$x^{\overline{m}} = \overset{m \text{ 个因子}}{x(x-1)\cdots(x-m+1)}, \text{ 整数 } m \geq 0, \quad (2.43)$$

定义这样新奇的 m 次幂。注意 m 之下的短直线，这意味着 m 个因子应该逐渐下降。还有一个对应的定义，定义中因子上升，

$$x^{\underline{m}} = \overset{m \text{ 个因子}}{x(x+1)\cdots(x+m-1)}, \text{ 整数 } m \geq 0. \quad (2.44)$$

当 $m=0$ 时，我们得到 $x^{\overline{0}} = x^{\underline{0}} = 1$ ，因为一个没有因子的乘积惯常取为 1（就像一个没有项的和惯常为 0）。

如果我们必需朗读它，则量 $x^{\overline{m}}$ 称为“ x 按照 m 下降”；相似， $x^{\underline{m}}$ 是“ x 按照 m 上升”，这些函数也称为下降阶乘幂和上升阶乘幂，因为它们和阶乘函数 $n! = n(n-1)\cdots(1)$ 紧密相关。事实上， $n! = n^{\underline{n}} = 1^{\overline{n}}$ 。

在数学文献中出现的其他几种阶乘幂的表示法，著名的有 $x^{\overline{m}}$ 或 $x^{\underline{m}}$ 的“Pochhammer 的符号” $(x)_m$ ；对于 $x^{\overline{m}}$ 还看到像 $x^{(m)}$ 或 $x_{(m)}$ 的表示法。但是在 m 之下与 m 之上划线惯常是受欢迎的，因为它易写易记，并且没有多余的括号。

下降幂 $x^{\overline{m}}$ 关于 Δ 特别好。我们得到

$$\begin{aligned} \Delta(x^{\overline{m}}) &= (x+1)^{\overline{m}} - x^{\overline{m}} \\ &= (x+1)x\cdots(x-m+2) - x\cdots(x-m+2)(x-m+1) \end{aligned}$$

$$= mx(x-1)\cdots(x-m+2),$$

因此, 有限演算有一个便于使用的定律与 $D(x^m) = mx^{m-1}$ 相匹配:

$$\Delta(x^m) = mx^{m-1} \quad (2.45)$$

这是基本的阶乘事实.

无限演算的算子 D 有一个逆, 逆微商 (或积分) 算子 \int . 联系 D 到 \int 的演算的基本定理:

$$g(x) = Df(x) \quad \text{当且仅当} \quad \int g(x)dx = f(x) + C.$$

这里 $\int g(x)dx$, $g(x)$ 的不定积分, 是微商为 $g(x)$ 的函数类. 类似, Δ 有一个逆, 逆差分 (或求和) 算子 \sum , 且有另一个基本定理:

$$g(x) = \Delta f(x) \quad \text{当且仅当} \quad \sum g(x)\delta x = f(x) + C. \quad (2.46)$$

这里 $\sum g(x)\delta x$, $g(x)$ 的不定和, 是差分为 $g(x)$ 的函数类. (注意, 像 d 与 D 有关那样, 小写 δ 与大写 Δ 有关.) 不定积分的“ C ”是一个任意常数; 不定和的“ C ”是任何使得 $p(x+1) = p(x)$ 的函数 $p(x)$. 例如, C 可以是周期函数 $a + b\sin 2\pi x$; 当我们取差分时, 这样的函数被去掉, 就像取微商时常数被去掉一样. 在 x 的整数值处, 函数 C 是常数.

现在我们基本上准备好了. 无限演算也有定积分: 若 $g(x) = Df(x)$, 则

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

所以有限演算——模拟它的极好的同等级位的内容——有确定的和: 若 $g(x) = \Delta f(x)$, 则

$$\sum_a^b g(x)\delta x = f(x)\Big|_a^b = f(b) - f(a). \quad (2.47)$$

此公式给出了记法 $\sum_a^b g(x)\delta x$ 的一种含意, 就像前面的公式确定 $\int_a^b g(x)dx$ 那样.

但是 $\sum_a^b g(x)\delta x$ 的实际的直观意义是什么呢? 我们用类似的方法来定义它, 并不是必要的. 我们要相似性成立, 以致我们能容易记住有限演算的规则; 如果我们不了解它的意义, 则记法将是无用的. 让我们尝试先看一些特殊情形来推出它的意义, 假设 $g(x) = \Delta f(x) = f(x+1) - f(x)$. 若 $b=a$, 则我们得到

$$\sum_a^a g(x)\delta x = f(a) - f(a) = 0.$$

其次, 若 $b=a+1$, 则结果是

$$\sum_a^{a+1} g(x)\delta x = f(a+1) - f(a) = g(a)$$

更一般, 若 b 增加 1, 则我们得到

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

这些观察以及数学归纳法, 允许我们确切地推出, 当 a 和 b 是整数, 且 $b \geq a$ 时, $\sum_a^b g(x)\delta x$ 的一般意义:

$$\sum_a^b g(x)\delta x = \sum_{k=a}^{b-1} g(k) = \sum_{a \leq k < b} g(k) , \text{ 整数 } b \geq a . \quad (2.48)$$

换句话说, 确定的和与通常具有界限的和相同, 而不带上限的值。

让我们尝试以稍微不同的方式概括这一点。假设给定一个未知和, 想象它以闭形式来表示, 且想象我们能把它写成形式 $\sum_{a \leq k < b} g(k) = \sum_a^b g(x)\delta x$ 。如果我们仅能找到一个不确定的和或者使得 $g(x) = f(x+1) - f(x)$ 的逆差分函数 f , 有限演算理论告诉我们能把解答表达成 $f(b) - f(a)$ 。了解此原理的一种方法是完全写出 $\sum_{a \leq k < b} g(k)$, 用三个点的表示法:

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

除了 $f(b) - f(a)$ 外, 右边所有项消去, 所以 $f(b) - f(a)$ 是和的值。(形式 $\sum_{a \leq k < b} (f(k+1) - f(k))$ 的和相似于折叠望远镜而常称为可伸缩的, 因为折叠望远镜的厚度单独由最外管的外径和最内管的内径所确定。)

但是仅当 $b \geq a$ 时规则(2.48)应用; 如果 $b < a$ 发生什么情况呢? 式(2.47)表明我们一定有

$$\begin{aligned}\sum_a^b g(x)\delta x &= f(b) - f(a) \\ &= -(f(a) - f(b)) = -\sum_b^a g(x)\delta x .\end{aligned}$$

这相似于定积分的对应方程。相似论证证明 $\sum_a^b + \sum_b^c = \sum_a^c$, 它是等式 $\int_a^b + \int_b^c = \int_a^c$ 的和的相似结果, 完全的形式为对于所有整数 a , b 和 c ,

$$\sum_a^b g(x)\delta x + \sum_b^c g(x)\delta x = \sum_a^c g(x)\delta x . \quad (2.49)$$

此处, 我们中间几个人也许开始对我们相信的这些平行的和类似的结果感到奇怪。举

出以下这一点使我们足以相信，确定的求和给我们一个简单的方法来计算下降幂的和：基本定律(2.45)，(2.47)和(2.48)意味着一般定律

$$\sum_{0 \leq k < n} k^{\overline{m}} = \left. \frac{k^{\overline{m+1}}}{m+1} \right|_0^n = \frac{n^{\overline{m+1}}}{m+1}, \text{ 整数 } m, n \geq 0. \quad (2.50)$$

这个公式容易记住，因为它十分像熟悉的 $\int_0^n x^m dx = n^{m+1} / (m+1)$ 。

特别，当 $m=1$ 时，我们得到 $k^{\overline{1}} = k$ ，所以有限演算的原理给我们一种容易的方法来记住事实

$$\sum_{0 \leq k < n} k = \frac{n^{\overline{2}}}{2} = n(n-1)/2.$$

确定和方法还给我们一个暗示，范围 $0 \leq k < n$ 上的和常常比 $1 \leq k \leq n$ 上的和简单；前者就是 $f(n) - f(0)$ ，而后者一定被计算为 $f(n+1) - f(1)$ 。

平常的幂也能以此新方法求和，若我们首先以下降幂来表达它们。例如，

$$k^2 = k^{\overline{2}} + k^{\overline{1}},$$

因此，

$$\sum_{0 \leq k < n} k^2 = \frac{n^{\overline{3}}}{3} + \frac{n^{\overline{2}}}{2} = \frac{1}{3} n(n-1) \left(n-2 + \frac{3}{2} \right) = \frac{1}{3} n \left(n - \frac{1}{2} \right) (n-1).$$

用 $n+1$ 替代 n ，还可给出另一种方法来计算我们前面见过的 $\square_n = \sum_{0 \leq k \leq n} k^2$ 的值为闭形式。

那是相当容易的。事实上，它比前节中许多把此公式解决到底的其他方法的任何方法容易。所以让我们尝试增加一个等级，从平方到立方：简单的计算表明

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

(总是可能用 Stirling 数在通常幂和阶乘幂之间变换，在第六章中将研究 Stirling 数。)于是，

$$\sum_{a \leq k < b} k^3 = \left. \frac{k^{\overline{4}}}{4} + k^{\overline{3}} + \frac{k^{\overline{2}}}{2} \right|_a^b.$$

所以对于和来说下降幂是十分合宜的。但是它们是否有任何其他改善的特性呢？在求和之前是否我们一定要把通常的幂变换到下降幂，另一方面在我们能做什么事之前变换回来？不是的，常常直接处理阶乘幂是可能的，因为它们有另外的性质。例如，就像

$(x+y)^2 = x^2 + 2xy + y^2$ 那样, 它结果是 $(x+y)^2 = x^2 + 2x^{\frac{1}{2}}y^{\frac{1}{2}} + y^2$, 且 $(x+y)^m$ 和 $(x+y)^{\frac{m}{2}}$ 之间相同的类似结果成立。(在习题 5.37 中证明此“阶乘二项定理”。)

到现在为止, 我们仅考虑了非负指数的下降幂。为了推广通常幂的相似结果到负指数, 我们需要一个 $m < 0$ 的 $x^{\frac{m}{2}}$ 的适当定义。看一下序列

$$x^{\frac{3}{2}} = x(x-1)(x-2),$$

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

$$x^{\frac{1}{2}} = x,$$

$$x^{\frac{0}{2}} = 1,$$

我们注意到, 用 $x-2$ 去除, 从 $x^{\frac{3}{2}}$ 得出 $x^{\frac{2}{2}}$, 除以 $x-1$ 从 $x^{\frac{2}{2}}$ 得出 $x^{\frac{1}{2}}$, 然后除以 x 从 $x^{\frac{1}{2}}$ 得出 $x^{\frac{0}{2}}$ 。看来我们应接着除以 $x+1$ 从 $x^{\frac{0}{2}}$ 得出 $x^{\frac{-1}{2}}$ 是合情合理的(如果不是绝对必要), 从而作出 $x^{\frac{-1}{2}} = 1/(x+1)$ 。继续进行下去, 前面几个负指数下降幂是

$$x^{\frac{-1}{2}} = \frac{1}{x+1},$$

$$x^{\frac{-2}{2}} = \frac{1}{(x+1)(x+2)},$$

$$x^{\frac{-3}{2}} = \frac{1}{(x+1)(x+2)(x+3)},$$

且我们的负下降幂的一般定义为

$$x^{\frac{-m}{2}} = \frac{1}{(x+1)(x+2)\cdots(x+m)}, \quad m > 0. \quad (2.51)$$

(对于实数或甚至复数 m 来定义下降幂也是可能的, 但是将推迟到第五章。)

用此定义, 下降幂有另外很好的性质。也许最重要的是相似于通常幂的定律

$$x^{\frac{m}{2} + \frac{n}{2}} = x^{\frac{m}{2}} x^{\frac{n}{2}}$$

的一般的指数定律。下降幂形式是

$$x^{\frac{m+n}{2}} = x^{\frac{m}{2}} (x-m)^{\frac{n}{2}}, \quad \text{整数 } m \text{ 和 } n. \quad (2.52)$$

例如, $x^{\frac{2+3}{2}} = x^{\frac{2}{2}} (x-2)^{\frac{3}{2}}$, 且对于负的 n 我们得到

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

若我们取选定义 $x^{\frac{-1}{2}}$ 为 $1/x$ 来替代 $1/(x+1)$, 则指数律 (2.52) 在情形 $m=-1$ 和 $n=1$ 时将不成立。事实上, 我们能用式 (2.52) 来确切地说明在负指数情形置 $m=-n$ 应该如何定

义下降幂。当推广存在的一种表示法适用于较多的情形时，以一般定律继续成立的这样一种方式来提出定义总是最好的。

现在让我们肯定，对于新定义的下降幂，关键性的差分性质成立，当 $m < 0$ 时， $\Delta x^{\overline{m}} = mx^{\overline{m-1}}$ 成立吗？若 $m = -2$ ，例如，差分是

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

差分性质成立！对于所有 $m < 0$ 可用相似的论证。

所以像正下降幂一样，对于负下降幂求和性质(2.50)成立，只要不出现被零除：

$$\sum_a^b x^{\overline{m}} \delta x = \left. \frac{x^{\overline{m+1}}}{m+1} \right|_a^b, \quad m \neq -1.$$

但是当 $m = -1$ 时发生什么情况？当 $m = -1$ 时对于积分我们用了

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

我们希望有一个 $\ln x$ 的有限的相似函数；换句话说，我们寻找一个函数 $f(x)$ 使得

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

不难看出

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

是这样的一个函数，当 x 是一个整数时，此等式就是式(2.13)的调和数 H_x 。于是 H_x 是连续的 $\ln x$ 的离散类似函数。（在第六章中我们将对非整数 x 定义 H_x ，但是对于目前的目的来说整数值已经足够好了。在第九章中我们还将看到，对于大的 x ， $H_x - \ln x$ 的值近似为 $0.577 + 1/(2x)$ 。因此 H_x 和 $\ln x$ 不仅相似，而且它们值之差小于 1。）

我们现在能给出下降幂的求和的完全叙述：

$$\sum_a^b x^{\overline{m}} \delta x = \begin{cases} \left. \frac{x^{\overline{m+1}}}{m+1} \right|_a^b, & \text{若 } m \neq -1, \\ H_x \Big|_a^b, & \text{若 } m = -1. \end{cases} \quad (2.53)$$

此公式表明为什么如同快速分类的分析那样，调和数会在离散问题的解中提出，就像在连续问题的解中自然提出的所谓自然对数。

现在我们已找到 $\ln x$ 的一个相似函数，让我们看一看是否对于 e^x 有一个对应于等式 $D e^x = e^x$ ，哪一个函数具有性质 $\Delta f(x) = f(x)$ ？显见：

$$f(x+1) - f(x) = f(x) \Leftrightarrow f(x+1) = 2f(x).$$

所以我们涉及一个简单的递归，且可取 $f(x) = 2^x$ 作为离散指数函数。

对于任意 c ， c^x 的差分也是十分简单的，即

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

因此，如果 $c \neq 1$ ， c^x 的逆差分是 $c^x / (c-1)$ 。此事实和基本定律(2.47)和(2.48)一起给出一个令人满意的方法来了解几何级数和的一般公式：

$$\sum_{a \leq x \leq b} c^x = \sum_a^b c^x \delta x = \left. \frac{c^x}{c-1} \right|_a^b = \frac{c^b - c^a}{c-1}, \quad c \neq 1.$$

当遇到一个可能有用的作为闭形式的函数，我们能计算它的差分 $\Delta f = g$ ；于是我们得到一个函数 g ，它的不确定和 $\sum g(x) \delta x$ 是已知的。表 2.1 是对求和有用的差分以及逆差分的表的开头部分。

尽管连续和离散数学之间有相似之处，但有些连续概念并没有离散的相似概念。例如，无限演算的链规则是函数的函数的微商的方便规则；但是没有对应的有限演算的链规则，因为没有 $\Delta f(g(x))$ 的合适形式。除了像 $c \pm x$ 替换 x 那些有把握的情形外，离散的变量改变是困难的。

然而 $\Delta(f(x)g(x))$ 有一种相当合宜的形式，它用分部求和的规则提供给我们，这就是无限演算称为分部积分的有限相似结果。让我们回想起无限演算的公式

$$D(uv) = uDv + vDu$$

导致分部积分规则，积分之后再移项得

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

在有限演算中我们能做相似的事情。

表 2.1 差分是什么？

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

我们从应用差分算子到两个函数 $u(x)$ 和 $v(x)$ 的乘积上开始:

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

用

$$Ef(x) = f(x+1)$$

定义的移位算子 E , 此公式可改写成一种合宜的形式. 用它替代 $v(x+1)$ 产生乘积的差分的一种紧凑的规则:

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

(E 有点麻烦, 但是它使方程正确.) 在此方程的两端取不定和, 再移项产生分部求和的通用的规则:

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

正如无限演算的情况一样, 在所有三项上放上界限, 使得不定和确定.

当左边的和比右边的和难算时, 此规则是有帮助的. 让我们来看一个例子. 函数 $\int xe^x dx$ 典型地由分部积分来计算; 它的离散的相似形式为 $\sum x2^x \delta x$, 本章前面部分在形式 $\sum_{k=0}^n k2^k$ 中我们曾遇到它. 为了用分部求此和, 我们设 $u(x)=x$ 和 $\Delta v(x)=2^x$; 因此 $\Delta u(x)=1$, $v(x)=2^x$, 以及 $Ev(x)=2^{x+1}$. 代入式(2.56)得到

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

像以前所做的那样, 我们能用它来计算和, 附加上界限:

$$\begin{aligned}\sum_{k=0}^n k2^k &= \sum_0^{n+1} x2^x \delta x \\ &= x2^x - 2^{x+1} \Big|_0^{n+1} \\ &= ((n+1)2^{n+1} - 2^{n+2}) - (0 \cdot 2^0 - 2^1) = (n-1)2^{n+1} + 2.\end{aligned}$$

此法比用摄动法较易找到和, 因为我们不必思索.

在本章前面部分, 我们偶然找到 $\sum_{0 \leq k \leq n} H_k$ 的一个公式, 且认为我们自己幸运. 但是如果知道分部求和, 就能系统地找出我们的公式(2.36). 让我们通过处理看来甚至较难的和 $\sum_{0 \leq k \leq n} kH_k$ 来展示这个断言. 如果通过 $\int x \ln x dx$ 的相似形式的引导, 解并不困难: 我们取 $u(x) = H_x$ 和 $\Delta v(x) = x = x^1$, 因此 $\Delta u(x) = x^{-1}$, $v(x) = x^2/2$, $Ev(x) = (x+1)^2$

/2, 我们得到

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

(从第一行到第二行中, 我们利用了 $m=-1$, $n=2$ 的指数定律(2.52)来合并两个下降幂 $(x+1)^{\frac{2}{2}} x^{\frac{-1}{2}}$.) 现在我们能附加上界限, 得到

$$\sum_{0 \leq k \leq n} k H_k = \sum_0^n x H_x \delta x = \frac{n^{\frac{2}{2}}}{2} \left(H_n - \frac{1}{2} \right). \quad (2.57)$$

2.7 无限和

当我们在本章开始处定义 \sum 记法时, 实际上我们说“稍后讨论无限和问题。现在, 我们能设我们碰到的所有和仅有有限多个非零项”来巧妙处理无限和的问题。但是计算时刻最终来到; 我们一定要面对和为无限的事实。实际情况是无限和既有好消息也有坏消息。

首先, 坏消息: 当涉及无限和时, 对于操作 \sum , 我们所用的方法并不总是成立。但是接着的好消息: 有一个大的容易了解的无限类, 对于此类的无限和, 我们执行的所有运算是完全合法的。在我们更仔细地看下面意义的和之后将明白这些好消息和坏消息的原因。

每个人都知道有限和是: 一项一项地把一串项加起来, 直到把它们加完。但是一个无限和需要更小心地定义, 免得我们陷入不合理的情况。

例如, 看来定义无限和

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

等于 2 是自然的, 因为如果把它加倍, 我们得到

$$2S = 2 + 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 2 + S.$$

另一方面, 此相同理由建议我们应定义

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

为 -1, 如果我们把它加倍, 得到

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

发生了希奇的事情，把正数加起来怎么会得到一个负数？看来把 T 留下作为不确定，或者我们也许该说 $T = \infty$ ，因为 T 中逐项相加比任何指定的有限数更大。（注意 ∞ 是方程 $2T = T - 1$ 的另一个“解”；也“解了”方程 $2S = 2 + S$ 。）

让我们尝试提出一般和 $\sum_{k \in K} a_k$ 的值的好的定义，其中 K 可以是无限的。开始，让我们设所有项 a_k 是非负的。然后不难找到一个合适的定义：如果对于所有有限子集 $F \subset K$ ，有一个有界常数 A 使得

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

则我们定义 $\sum_{k \in K} a_k$ 是最小的这样的 A 。（根据众所周知的实数性质，即所有这样的 A 的集合总包含一个最小元素得出。）但是如果没有有限常数 A ，我们说 $\sum_{k \in K} a_k = \infty$ ；这意味着如果 A 是任何实数，则有有限多个项 a_k 的集合，它的和超过 A 。

前段中已小心地提出定义，使它不依赖于指标集 K 中可能存在的任何次序。所以我们即将做的论证不仅应用于整数集上的和，还将应用于许多指标 k_1, k_2, \dots 的多重和上。

在特殊情形中， K 是非负整数集合，我们的非负项 a_k 的定义意味着

$$\sum_{k \geq 0} a_k = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k.$$

由此说明为什么任何非降实数序列都有一个极限(可能为 ∞)。若极限是 A ，且如果 F 是任何非负整数的有限集，它的元素全 $\leq n$ ，则我们有 $\sum_{k \in F} a_k \leq \sum_{k=0}^n a_k \leq A$ ；因此， $A = \infty$ 或 A 是一个有界的常数。若 A' 是小于固定极限 A 的任何数，则总有一个 n 使得 $\sum_{k=0}^n a_k > A'$ ；因此有限集合 $F = \{0, 1, \dots, n\}$ 证明了 A' 不是一个有界常数。

按照刚才给出的定义，现在我们能容易地计算出某个无限和的值。例如，若 $a_k = x^k$ ，则得到

$$\sum_{k \geq 0} x^k = \lim_{n \rightarrow \infty} \frac{1 - x^{n+1}}{1 - x} = \begin{cases} 1/(1-x), & \text{若 } 0 \leq x < 1; \\ \infty, & \text{若 } x \geq 1. \end{cases}$$

特别，刚才我们考虑的无限和 S 和 T 有各自的值 2 和 ∞ ，恰与我们猜想的相同。另一个有趣的例子是

$$\sum_{k \geq 0} \frac{1}{(k+1)(k+2)} = \sum_{k \geq 0} k^{-2} = \lim_{n \rightarrow \infty} \sum_{k=0}^n k^{-2} = \lim_{n \rightarrow \infty} \left. \frac{k^{-1}}{-1} \right|_0^n = 1.$$

现在让我们考虑既有非负项也可以有负项的和。例如，

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

的值应该是什么？若我们两项两项地括在一起，得到

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

这样由此产生的和为零；但是如果在第一项之后两项两项地开始括在一起，则得到

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

和是 1。

我们也可以尝试在公式 $\sum_{k \geq 0} x^k = 1/(1-x)$ 中置 $x = -1$ ，由于当 $0 \leq x < 1$ 时我们证明了此公式成立；虽然它是整数的和，但我们不得不作出无限和为 $1/2$ 的断定。

另一个有趣的例子是双重表示的无限和 $\sum_k a_k$ ，其中对于 $k \geq 0$ ， $a_k = 1/(k+1)$ ， $k < 0$ ， $a_k = 1/(k-1)$ 。我们可把它写成

$$\cdots + \left(-\frac{1}{4}\right) + \left(-\frac{1}{3}\right) + \left(-\frac{1}{2}\right) + 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots, \quad (2.58)$$

若我们这样来计算此和，从“中心”元素处开始，且向外处理，

$$\cdots + \left(-\frac{1}{4} + \left(-\frac{1}{3} + \left(-\frac{1}{2} + (1) + \frac{1}{2}\right) + \frac{1}{3}\right) + \frac{1}{4}\right) + \cdots,$$

我们得到值 1；若我们把所有括号向左移一步

$$\cdots + \left(-\frac{1}{5} + \left(-\frac{1}{4} + \left(-\frac{1}{3} + \left(-\frac{1}{2} + 1\right) + \frac{1}{2}\right) + \frac{1}{3}\right) + \frac{1}{4}\right) + \cdots,$$

我们得到相同的值 1，因为最里面的 n 个括号内部的所有数的和是

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

类似的论证表明：如果括号向左或向右移任何指定的数量，值为 1，这就促使我们相信并确认是 1。另一方面，如果我们以下列方式归并项

$$\cdots + \left(-\frac{1}{4} + \left(-\frac{1}{3} + \left(-\frac{1}{2} + 1 + \frac{1}{2}\right) + \frac{1}{3} + \frac{1}{4}\right) + \frac{1}{5} + \frac{1}{6}\right) + \cdots,$$

从里到外第 n 对括号包含数

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

在第九章中我们将证明 $\lim_{n \rightarrow \infty} (H_{2n} - H_{n+1}) = \ln 2$ ；因此这种归并法使人联想到双重表示的无限和实际上应等于 $1 + \ln 2$ 。

关于和有极古怪的事情, 当以不同方式把它的项加起来时给出不同的值. 分析方面的高级教科书有许多定义, 依据这些定义能赋予这样的病态和以有意义的值; 但是如果我们采用这些定义, 则不能自由地用 \sum 记法运算, 如同我们所做的那样. 就本书的目的来说, 我们不需要难以处理的“条件收敛”的改进; 所以我们将保留无限和的一个定义, 使本章中我们所做的所有运算保持成立.

事实上, 我们的无限和的定义是十分简单的. 设 K 是任何集合, a_k 是对于 $k \in K$ 定义的实值项. (这里“ k ”可以实际表示几个指标 k_1, k_2, \dots , 所以 K 可能是多维的.) 任何实数 x 可记为它的正部和负部之差,

$$x = x^+ - x^-, \text{ 其中 } x^+ = x \cdot [x > 0], \quad x^- = -x \cdot [x < 0].$$

(或者 $x^+ = 0$ 或 $x^- = 0$.) 因为 a_k^+ 和 a_k^- 是非负的, 所以我们已说明了怎样定义无限和 $\sum_{k \in K} a_k^+$ 和 $\sum_{k \in K} a_k^-$ 的值. 我们的一般定义是

$$\sum_{k \in K} a_k = \sum_{k \in K} a_k^+ - \sum_{k \in K} a_k^-, \quad (2.59)$$

除非右边的和都等于 ∞ . 在后面的情形, 我们把 $\sum_{k \in K} a_k$ 按没有定义看待.

设 $A^+ = \sum_{k \in K} a_k^+$, $A^- = \sum_{k \in K} a_k^-$. 若 A^+ 和 A^- 都有限, 则称 $\sum_{k \in K} a_k$ 绝对收敛于值 $A = A^+ - A^-$. 若 $A^+ = \infty$, 但 A^- 有限, 则称和 $\sum_{k \in K} a_k$ 发散到 $+\infty$. 类似, 若 $A^- = \infty$, 但 A^+ 有限, 则称 $\sum_{k \in K} a_k$ 发散到 $-\infty$. 若 $A^+ = A^- = \infty$, 则所有解决问题的办法被中断.

我们从处理非负项的定义开始, 然后把它推广到实值项. 若项 a_k 是复数, 则我们能以显然的方法再一次推广定义: 定义 $\sum_{k \in K} a_k$ 为 $\sum_{k \in K} \mathcal{R} a_k + i \sum_{k \in K} \mathcal{I} a_k$, 其中 $\mathcal{R} a_k$ 和 $\mathcal{I} a_k$ 是 a_k 的实部和虚部, 假设这些和都被定义, 否则 $\sum_{k \in K} a_k$ 没有定义. (见习题 18.)

如同前面部分所述, 坏消息是有些无限和没有定义, 因为我们所做的操作在所有这样的情形中会产生矛盾. (见习题 34.) 好消息是每当我们讨论刚才定义的绝对收敛的和时, 本章所有的操作完全成立.

通过表明每一个变换规则保持所有绝对收敛和的值不变, 我们能验证好消息. 更明确地说, 这意味着我们一定要证明分配律, 结合律和交换律, 加上首先在一个指标变量上相加的规则; 从这样 4 个和上的基本运算能推出我们已做的其他事情.

分配律(2.15)能更精确地提出如下: 若 $\sum_{k \in K} a_k$ 绝对收敛于 A , 且 c 是任何复数, 则 $\sum_{k \in K} ca_k$ 绝对收敛于 cA . 我们把和分为上面所说的实部和虚部, 正和负部, 并且证明 $c > 0$, 每一项 a_k 非负的特殊情形能证明这一点. 此特殊情形中的证明见效是因为对于所有有限集合 F , $\sum_{k \in F} ca_k = c \sum_{k \in F} a_k$; 就 F 的大小进行归纳得到后面的事实.

结合律(2.16)可说明如下: 若 $\sum_{k \in K} a_k$ 和 $\sum_{k \in K} b_k$ 绝对收敛到 A 和 B , 则 $\sum_{k \in K} (a_k + b_k)$ 绝对收敛到 $A+B$. 这原来是我们即将证明的更一般定理的一个特殊情形.

交换律(2.17)实际上并不需要证明, 因为我们在讨论中已表明, 作为交换求和次序的

一般规则的特殊情形根据式(2.35)怎样推出它。

需要证明的主要结果是多重和的基本原理：在两个或多个指标上的绝对收敛和总能先对任何这样的指标求和。形式地说，我们将证明，如果 J 和 $\{K_j | j \in J\}$ 的元素是任何指标集合使得

$$\sum_{\substack{j \in J \\ k \in K_j}} a_{j,k} \quad \text{绝对收敛到 } A,$$

则对于每个 $j \in J$ 存在复数 A_j ，使得

$$\sum_{k \in K_j} a_{j,k} \quad \text{绝对收敛到 } A_j, \text{ 且}$$

$$\sum_{j \in J} A_j \quad \text{绝对收敛到 } A.$$

当所有项是非负时证明此结果就够了，因为我们能像前面那样把每个和分成实部和虚部，正部和负部来证明一般情形。所以让我们设对所有对 $(j, k) \in M$ ， $a_{j,k} \geq 0$ ，其中 M 是主要的指标集合 $\{(j, k) | j \in J, k \in K_j\}$ 。

我们给定 $\sum_{(j,k) \in M} a_{j,k}$ 是有限的，即对所有有限子集 $F \subseteq M$ ，

$$\sum_{(j,k) \in F} a_{j,k} \leq A$$

且 A 是最小的这样的上界。若 j 是 J 的任何元素，则每个形式 $\sum_{k \in F_j} a_{j,k}$ 的和的上界为 A ，其中 F_j 是 K_j 的有限子集。因此这些有限和有一个最小上界 $A_j \geq 0$ ，且根据定义 $\sum_{k \in K_j} a_{j,k} = A_j$ 。

我们还需证明对于所有有限子集 $G \subseteq J$ ， A 是 $\sum_{j \in G} A_j$ 的最小上界。假设 G 是具有 $\sum_{j \in G} A_j = A' > A$ 的 J 的有限子集，我们能找到有限子集 $F_j \subseteq K_j$ 使得对于每个具有 $A_j > 0$ 的 $j \in G$ ， $\sum_{k \in F_j} a_{j,k} > (A/A')A_j$ 。至少有一个这样的 j 。但是 $\sum_{j \in G, k \in F_j} a_{j,k} > (A/A') \sum_{j \in G} A_j = A$ ，这和对于所有有限集 $F \subseteq M$ ， $\sum_{(j,k) \in F} a_{j,k} \leq A$ 相矛盾。因此对于所有有限集 $G \subseteq J$ ， $\sum_{j \in G} A_j \leq A$ 。

最后，设 A' 是小于 A 的任何实数。若能找到一个有限集 $G \subseteq J$ 使得 $\sum_{j \in G} A_j > A'$ ，则将完成我们的证明。我们知道有一个有限集合 $F \subseteq M$ 使得 $\sum_{(j,k) \in F} a_{j,k} > A'$ ，设 G 是这个 F 中的 j 的集合，且设 $F_j = \{k | (j, k) \in F\}$ 。则 $\sum_{j \in G} A_j \geq \sum_{j \in G} \sum_{k \in F_j} a_{j,k} = \sum_{(j,k) \in F} a_{j,k} > A'$ ，证完。

现在我们是合法了！我们证明了就无限和所做的每件事是有理的，只要对项的绝对值的所有有限和有一个有限的界限。由于当我们以两种不同方式计算双重表示的无限和 (2.58) 时，给出两个不同的解答，它的正的项 $1 + 1/2 + 1/3 + \dots$ 一定发散到 ∞ ；否则不

管怎样归并项我们将取得相同的解答。

习 题

准备部分

1. 记法

$$\sum_{k=4}^0 q_k$$

的意义是什么?

2. 简化表达式 $x \cdot ([x > 0] - [x < 0])$ 。

3. 通过全部写出和

$$\sum_{0 \leq k \leq 5} a_k \quad \text{和} \quad \sum_{0 \leq k^2 \leq 5} a_{k^2}$$

来表出你对 \sum 记法的理解。(注意, 第二个和有点复杂。)

4. 把三重和

$$\sum_{1 \leq i < j < k \leq 4} a_{ijk}$$

表为(用三个 \sum 的)三重求和。

(a) 首先对 k 求和, 然后对 j 求和, 接着对 i 求和。

(b) 首先对 i 求和, 然后对 j 求和, 接着对 k 求和。

不用 \sum 记号再把你的三重和全部写出, 用括号表明首先加在一起的项。

5. 下列推导的错误是什么?

$$\left(\sum_{j=1}^n a_j \right) \left(\sum_{k=1}^n \frac{1}{a_k} \right) = \sum_{j=1}^n \sum_{k=1}^n \frac{a_j}{a_k} = \sum_{k=1}^n \sum_{j=1}^n \frac{a_k}{a_k} = \sum_{k=1}^n n = n^2.$$

6. 作为 j 和 n 的一个函数, $\sum_k [1 \leq j \leq k \leq n]$ 的值是什么?

7. 设 $\nabla f(x) = f(x) - f(x-1)$, $\nabla(x^{\overline{m}})$ 是什么?

8. 当 m 是一个给定的整数时, $0^{\overline{m}}$ 的值是什么?

9. 相似于式(2.52), 上升阶乘幂的指数律是什么? 用此来定义 $x^{\overline{-n}}$ 。

10. 课文推出乘积差分的下列公式:

$$\Delta(uv) = u\Delta v + Ev\Delta u.$$

这个公式怎么能是正确的, 左边关于 u 和 v 是对称的, 但是右边关于 u 和 v 是不对称的?

基本部分

11. 分部求和的一般规则(2.56)等价于

$$\sum_{0 \leq k < n} (a_{k+1} - a_k) b_k = a_n b_n - a_0 b_0 - \sum_{0 \leq k < n} a_{k+1} (b_{k+1} - b_k), \quad n \geq 0.$$

用分配律, 结合律和交换律直接证明此公式.

12. 证明函数 $p(k) = k + (-1)^k c$ 是所有整数集合的一个置换, c 是一个整数.

13. 用包含各种组成部分的方法来找 $\sum_{k=0}^n (-1)^k k^2$ 的一个闭形式.

14. 通过把 $\sum_{k=1}^n k 2^k$ 改写成多重和 $\sum_{1 \leq j \leq k \leq n} 2^k$ 来计算它.

15. 用课本的方法 5 来计算 $\sum_{k=1}^n k^3$ 如下: 首先写出 $\sum_{k=1}^n k^3 + \sum_{k=1}^n k^3$
 $= 2 \sum_{1 \leq j \leq k \leq n} jk$, 然后应用式(2.33).

16. 证明 $x^m / (x-n)^m = x^n / (x-m)^n$, 除非分母之一为零.

17. 对于所有整数 m , 证明下列公式能用来作为上升阶乘幂和下降阶乘幂之间的转换:

$$x^{\overline{m}} = (-1)^m (-x)^{\overline{m}} = (x+m-1)^{\overline{m}} = \frac{1}{(x-1)^{\overline{-m}}};$$

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

(习题 9 的解答定义 $x^{\overline{-m}}$.)

18. 设 $\mathcal{R}z$ 和 $\mathcal{I}z$ 是复数 z 的实部和虚部. 绝对值 $|z|$ 是 $\sqrt{(\mathcal{R}z)^2 + (\mathcal{I}z)^2}$. 复数项 a_k 的和 $\sum_{k \in K} a_k$ 称为绝对收敛, 当实值和 $\sum_{k \in K} \mathcal{R}a_k$ 和 $\sum_{k \in K} \mathcal{I}a_k$ 都绝对收敛时, 证明 $\sum_{k \in K} a_k$ 绝对收敛当且仅当有一个有界常数 B 使得对于所有有限子集 $F \subseteq K$, $\sum_{k \in F} |a_k| \leq B$.

课外作业的习题

19. 用一个求和因子来解递归

$$T_0 = 5;$$

$$2T_n = nT_{n-1} + 3 \cdot n!, \quad n > 0.$$

20. 试用摄动法计算 $\sum_{k=0}^n k H_k$, 并推出 $\sum_{k=0}^n H_k$ 的值.

21. 假设 $n \geq 0$, 用摄动法计算 $S_n = \sum_{k=0}^n (-1)^{n-k}$, $T_n = \sum_{k=0}^n (-1)^{n-k} k$ 和 $U_n = \sum_{k=0}^n (-1)^{n-k} k^2$.

22. 证明 Lagrange 的等式(不用归纳法):

$$\sum_{1 \leq j < k \leq n} (a_j b_k - a_k b_j)^2 = \left(\sum_{k=1}^n a_k^2 \right) \left(\sum_{k=1}^n b_k^2 \right) - \left(\sum_{k=1}^n a_k b_k \right)^2.$$

顺便提一句, 此式蕴含 Cauchy 不等式,

$$\left(\sum_{k=1}^n a_k b_k \right)^2 \leq \left(\sum_{k=1}^n a_k^2 \right) \left(\sum_{k=1}^n b_k^2 \right).$$

23. 用两种方式来计算和 $\sum_{k=1}^n (2k+1)/(k(k+1))$:

(a) 用“部分分数” $1/k - 1/(k+1)$ 代替 $1/(k(k+1))$.

(b) 分部求和.

24. $\sum_{0 \leq k \leq n} H_k / (k+1)(k+2)$ 是什么? 提示: 推广式(2.57)的推导.

25. 记法 $\prod_{k \in K} a_k$ 意味着对于所有 $k \in K$, 数 a_k 的乘积. 为了简单起见, 假设仅仅是有限多个 k , $a_k \neq 1$, 因为不需定义无限乘积. 相似于 \sum 成立的分配律, 结合律和交换律, \prod 满足什么定律?

26. 通过操作 \prod 记法, 用单个乘积 $\prod_{k=1}^n a_k$ 来表达双重乘积 $\prod_{1 \leq j < k \leq n} a_j a_k$. (此习题给出上三角形等式(2.33)的一个乘积相似形.)

27. 计算 $\Delta(c^{\frac{x}{k}})$, 且用它来推出 $\sum_{k=1}^n (-2)^{\frac{k}{k}} / k$ 的值.

28. 下列推导的什么地方出错?

$$\begin{aligned} 1 &= \sum_{k \geq 1} \frac{1}{k(k+1)} = \sum_{k \geq 1} \left(\frac{k}{k+1} - \frac{k-1}{k} \right) \\ &= \sum_{k \geq 1} \sum_{j \geq 1} \left(\frac{k}{j} [j=k+1] - \frac{j}{k} [j=k-1] \right) \\ &= \sum_{j \geq 1} \sum_{k \geq 1} \left(\frac{k}{j} [j=k+1] - \frac{j}{k} [j=k-1] \right) \\ &= \sum_{j \geq 1} \sum_{k \geq 1} \left(\frac{k}{j} [k=j-1] - \frac{j}{k} [k=j+1] \right) \\ &= \sum_{j \geq 1} \left(\frac{j-1}{j} - \frac{j}{j+1} \right) = \sum_{j \geq 1} \frac{-1}{j(j+1)} = -1. \end{aligned}$$

考查性问题

29. 计算和 $\sum_{k=1}^n (-1)^k k / (4k^2 - 1)$.

30. 一种纸牌游戏者很久才知道 $15 = 7+8 = 4+5+6 = 1+2+3+4+5$. 求出用相继的正整数的和来代表 1050 的方法数. (用它本身平凡的代表 '1050' 作为一种方法计入; 因此有

四种方法用相继的正整数的和代表 15, 而不是三种方法。顺便提一句, 在此问题中纸牌游戏规则的知识是无用的。)

31. Riemann 的 zeta 函数 $\zeta(k)$ 定义为无限和

$$1 + \frac{1}{2^k} + \frac{1}{3^k} + \cdots = \sum_{j=1}^{\infty} \frac{1}{j^k}.$$

证明 $\sum_{k \geq 2} (\zeta(k) - 1) = 1$. $\sum_{k \geq 1} (\zeta(2k) - 1)$ 的值是什么?

32. 设 $a \dot{-} b = \max(0, a - b)$. 证明对于所有实 $x \geq 0$,

$$\sum_{k \geq 0} \min(k, x \dot{-} k) = \sum_{k \geq 0} (x \dot{-} (2k + 1)),$$

且求和的闭形式。

额外问题

33. 令 $\bigwedge_{k \in K} a_k$ 表示数 a_k 的最小者(或者它们的最大下界, 如果 K 是无限), 假设每个 a_k 或者为实数, 或者为 $\pm \infty$. 相似于 Σ 和 Π 成立的那些定律, 对于 \bigwedge 记号有哪些定律成立? (见习题 25.)

34. 证明如果按照式(2.59), 和 $\sum_{k \in K} a_k$ 是没有定义的, 则它以下列意义是极古怪的: 若 A^- 和 A^+ 是任何给定的实数, 可能找出 K 的有限子集的序列 $F_1 \subset F_2 \subset F_3 \subset \cdots$ 使得当 n 是奇数时

$$\sum_{k \in F_n} a_k \leq A^-; \text{ 当 } n \text{ 是偶数时, } \sum_{k \in F_n} a_k \geq A^+.$$

35. 证明 Goldbach 定理

$$1 = \frac{1}{3} + \frac{1}{7} + \frac{1}{8} + \frac{1}{15} + \frac{1}{24} + \frac{1}{26} + \frac{1}{31} + \frac{1}{35} + \cdots = \sum_{k \in P} \frac{1}{k-1}.$$

其中 P 是如下递归定义的“完全幂”的集合:

$$P = \{m^n \mid m \geq 2, n \geq 2, m \notin P\}.$$

36. Solomon Golomb“自描述序列” $\langle f(1), f(2), f(3), \cdots \rangle$ 是具有一列性质的正整数的非降序列, 即对于每个 k , 序列恰好包含 $f(k)$ 次 k 的出现。稍加思索显示序列一定开始如下:

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

设 $g(n)$ 是使得 $f(m) = n$ 的最大整数 m . 证明

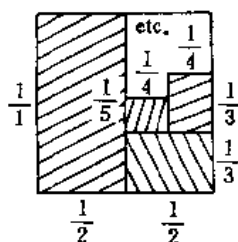
$$(a) \ g(n) = \sum_{k=1}^n f(k).$$

$$(b) \ g(g(n)) = \sum_{k=1}^n kf(k).$$

$$(c) \ g(g(g(n))) = \frac{1}{2}ng(n)(g(n)+1) - \frac{1}{2}\sum_{k=1}^{n-1}g(k)(g(k)+1).$$

研究性问题

37. 是否能将所有长为 $1/k$ 宽为 $1/(k+1)$ 的矩形 ($k \geq 1$) 一起装填在长为 1 宽为 1 的正方形内部? (记住它们的面积相加达到 1.)



第三章 整函数

整数构成了离散数学的主要成分，且我们常常需要把分数或任意实数变换到整数。本章的目的是熟悉这种变换且能应用自如，同时学一些值得注意的性质。

3.1 下整函数和上整函数

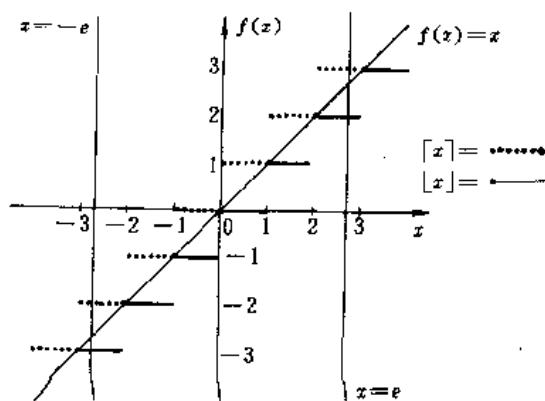
我们先从说明下整(最大整数)和上整(最小整数)函数开始，对于所有实数 x ，它们被定义如下：

$\lfloor x \rfloor$ = 小于或等于 x 的最大整数；

$\lceil x \rceil$ = 大于或等于 x 的最小整数。 (3.1)

早在 60 年代 Kenneth E. Iverson 引入了这个记法并取名“下整”和“上整”[161, P.12页]。他发现排字工人能通过修剪‘[’和‘]’的顶部和底部来处理符号。他的记法用得非常普遍，现在在一篇专业论文中能用下整和上整括号，而不必说明它们的意义。直到最近，人们常常用‘ $\lfloor x \rfloor$ ’记 $\leq x$ 的最大整数，而对于最小整函数没有一个好的对应词。有些作者还试用‘ $\lceil x \rceil$ ’来表示，缺乏成功的希望。

除了记法方面的变化之外，在函数方面也有变化。例如，有些袖珍计算器有一个 INT 函数，当 x 为正时定义为 $\lfloor x \rfloor$ ，当 x 为负时定义为 $\lceil x \rceil$ 。这些计算器的设计者也许要他们的 INT 函数满足等式 $\text{INT}(-x) = -\text{INT}(x)$ 。但是我们将坚持用我们的下整和上整函数，因为它们甚至有比此更好的性质。



熟悉下整函数和上整函数的一种好方法是了解它们的图形，这些图形形成了直线 $f(x)=x$ 上下像阶梯的形状：

例如，根据图我们看出

$$\lfloor e \rfloor = 2, \lfloor -e \rfloor = -3,$$

$$\lceil e \rceil = 3, \lceil -e \rceil = -2,$$

因为 $e = 2.718\,28\dots$ 。

从这个说明开始，我们能看出几个下整和上整的事实。首先，由于下整函数位于对角线 $f(x)=x$ 上或 $f(x)=x$ 之下，所以我们有 $\lfloor x \rfloor \leq x$ ；同样 $\lceil x \rceil \geq x$ 。（当然根据定义这是完全明显的。）在整数点处两个函数恰好相等：

$$\lfloor x \rfloor = x \Leftrightarrow x \text{ 是一个整数} \Leftrightarrow \lceil x \rceil = x.$$

（我们用记号“ \Leftrightarrow ”意味着“当且仅当”。）此外，当它们不同的时候，上整恰好比下整大 1：

$$\lceil x \rceil - \lfloor x \rfloor = [x \text{ 不是一个整数}]. \quad (3.2)$$

如果我们将对角线向下平移一个单位，则它完全位于下整函数之下，所以 $x-1 < \lfloor x \rfloor$ ；同样有 $x+1 > \lceil x \rceil$ 。把这些结果结合起来得到

$$x-1 < \lfloor x \rfloor \leq x \leq \lceil x \rceil < x+1. \quad (3.3)$$

最后，函数关于两个轴彼此是反射的：

$$\lfloor -x \rfloor = -\lceil x \rceil, \lceil -x \rceil = -\lfloor x \rfloor. \quad (3.4)$$

因此每一个容易由另一个表达出来。这一点有助于说明上整函数曾经没有它自己的记号的原因。但是正如我们对于上升幂和下降幂所采用的特殊记号那样，我们看出上整函数常常有足够的根据给出它们特殊的符号。很久以来数学家们有正弦和余弦，正切和余切，正割和余割，最大和最小；现在我们还有下整和上整函数。

为了实际证明下整函数和上整函数的性质，而不是仅仅从图形上看出这些事实，以下四个规则是特别有用的：

$$\begin{aligned} \lfloor x \rfloor = n &\Leftrightarrow n \leq x < n+1, \quad (a) \\ \lfloor x \rfloor = n &\Leftrightarrow x-1 < n \leq x, \quad (b) \\ \lceil x \rceil = n &\Leftrightarrow n-1 < x \leq n, \quad (c) \\ \lceil x \rceil = n &\Leftrightarrow x \leq n < x+1. \quad (d) \end{aligned} \quad (3.5)$$

（在所有四种情形中我们假设 n 是整数以及 x 是实数。）规则(a)和(c)是定义式(3.1)的直接结果；规则(b)和(d)也是定义的直接结果，仅把 n 移到不等式的中部。

把整数项移进或移出一个下整函数(或上整函数)是可能的：

$$\lfloor x+n \rfloor = \lfloor x \rfloor + n, \text{ 整数 } n. \quad (3.6)$$

（因为规则(3.5(a))表明此结论等价于不等式 $\lfloor x \rfloor + n \leq x+n < \lfloor x \rfloor + n+1$ 。）但是像移出一个常数

因子那样的相似运算一般不能做。例如，当 $n=2$ 和 $x=1/2$ 时，我们有 $\lfloor nx \rfloor \neq n \lfloor x \rfloor$ 。这意味着下整和上整括号是比较僵硬的。如果我们能除去它们，或者当描述它们时我们能证明一切，我们往往感到高兴。

在许多场合下整括号和上整括号是多余的，以致我们能任意插入或删除它们。例如，一个实数和一个整数间的任何不等式等价于整数间的一个下整不等式或一个上整不等式：

$$\begin{aligned} x < n &\Leftrightarrow \lfloor x \rfloor < n, & (a) \\ n < x &\Leftrightarrow n < \lceil x \rceil, & (b) \\ x \leq n &\Leftrightarrow \lceil x \rceil \leq n, & (c) \\ n \leq x &\Leftrightarrow n \leq \lfloor x \rfloor, & (d) \end{aligned} \tag{3.7}$$

这些规则是容易证明的。例如，若 $x < n$ ，则肯定 $\lfloor x \rfloor < n$ ，因为 $\lfloor x \rfloor \leq x$ 。相反，若 $\lfloor x \rfloor < n$ ，则一定有 $x < n$ ，因为 $x \leq \lfloor x \rfloor + 1$ 和 $\lfloor x \rfloor + 1 \leq n$ 。

如果在式(3.7)中的四个规则如同证明它们那样容易记住将是合适的。没有下整或下整的每个不等式对应于相同的具有下整或具有上整的不等式，但是在决定哪一种合适之前我们需要重新考虑。

x 和 $\lfloor x \rfloor$ 间的差称为 x 的分数部分，在应用中它的出现常常值得用另一个记号来标记：

$$\{x\} = x - \lfloor x \rfloor. \tag{3.8}$$

我们有时候称 $\lfloor x \rfloor$ 为 x 的整数部分，因为 $x = \lfloor x \rfloor + \{x\}$ 。若一个实数 x 能记为形式 $x = n + \theta$ ，其中 n 是整数，且 $0 \leq \theta \leq 1$ ，由式(3.5(a))我们能得出 $n = \lfloor x \rfloor$ 和 $\theta = \{x\}$ 。

如果 n 是任意实数，则等式(3.6)不成立。但是我们能推出，一般对 $\lfloor x+y \rfloor$ 仅有两种可能性：若写出 $x = \lfloor x \rfloor + \{x\}$ 和 $y = \lfloor y \rfloor + \{y\}$ ，则我们有 $\lfloor x+y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor + \lfloor \{x\} + \{y\} \rfloor$ 且由于 $0 \leq \{x\} + \{y\} < 2$ ，我们有时得到 $\lfloor x+y \rfloor$ 为 $\lfloor x \rfloor + \lfloor y \rfloor$ ，否则它为 $\lfloor x \rfloor + \lfloor y \rfloor + 1$ 。

3.2 下整 / 上整的应用

我们现在已明白了处理下整和上整的基本工具。让我们开始把它们用在一个容易的问题上： $\lceil \lg 35 \rceil$ 等于多少？(我们用 ‘lg’ 表示底为 2 的对数。) 由于 $2^5 < 35 \leq 2^6$ ，取对数得到 $5 < \lg 35 \leq 6$ ，所以由式(3.5(c))可知 $\lceil \lg 35 \rceil = 6$ 。

注意，当用二进制记法时，数 35 是 6 位长： $35 = (100111)_2$ 。是否 $\lceil \lg n \rceil$ 总是 n 记为二进制的长度？不完全是这样。我们也需 6 位来记 $32 = (100000)_2$ 。所以 $\lceil \lg n \rceil$ 是问题的错误解答。(当 n 是 2 的幂时它失效，但是那是无限多次失效。) 通过实现取 m 位来写出每个数 n 使得 $2^{m-1} \leq n < 2^m$ ，我们能找到一个正确解答；于是式(3.5(a))告诉我们 $m-1 = \lfloor \lg n \rfloor$ ，所以 $m = \lfloor \lg n \rfloor + 1$ 。也就是说，对于所有 $n > 0$ ，我们需要 $\lfloor \lg n \rfloor + 1$ 位来二进制表达 n 。换句话说，相似推导产生解答 $\lceil \lg(n+1) \rceil$ ；如果我们乐于说取 0 位来二进制写出 $n=0$ ，此公式对于 $n=0$ 也成立。

让我们接着看几个下整或上整表达式。 $\lfloor \lfloor x \rfloor \rfloor$ 是什么？由于 $\lfloor x \rfloor$ 是一个整数，所以

$\lfloor \lfloor x \rfloor \rfloor$ 就是 $\lfloor x \rfloor$ ，这是不费力的。这就是把最内层的 $\lfloor x \rfloor$ 围上许多次下整或上整的其他表达式。

下面是一个较难对付的问题：证明或推翻结论

$$\lfloor \sqrt{\lfloor x \rfloor} \rfloor = \lfloor \sqrt{x} \rfloor, \text{ 实数 } x \geq 0. \quad (3.9)$$

当 x 是整数时等式明显成立，因为 $x = \lfloor x \rfloor$ 。而且在特殊情形 $\pi = 3.141\,59\dots$ ， $e = 2.718\,28\dots$ 和 $\varphi = (1 + \sqrt{5})/2 = 1.618\,03\dots$ ，有等式，因为我们得到 $1=1$ 。我们没有找到一个反例而联想到一般等式成立，所以尝试来证明它。

顺便提到，当我们面临“证明或推翻”时，通常首先尝试用反例来推翻较好，有两方面原因：推翻可能较容易（我们仅需一个反例）；且挑选激发了我们的创造性的思维。即使给定的结论是真的，我们寻找反例期间明白一个反例为不可能的原因，常常就引出一个证明。此外，对证明的事实怀疑是有益的。

如果我们尝试借助于微积分来证明 $\lfloor \sqrt{\lfloor x \rfloor} \rfloor = \lfloor \sqrt{x} \rfloor$ 可从分解 x 为它的整数部分和分数部分 $\lfloor x \rfloor + \{x\} = n + \theta$ 开始，然后用二项定理展开平方根： $(n + \theta)^{1/2} = n^{1/2} + n^{-1/2}\theta/2 - n^{-3/2}\theta^2/8 + \dots$ 。但是这种方式相当难做。

使用我们已建立的工具是很容易的。此处是一种可能的对策：以某种方式除去 $\lfloor \sqrt{\lfloor x \rfloor} \rfloor$ 的外面的下整括号和平方根，然后移去内部的下整括号，再添加回外部下整括号而取得 $\lfloor \sqrt{x} \rfloor$ 。我们设 $m = \lfloor \sqrt{\lfloor x \rfloor} \rfloor$ ，且用式(3.5(a))，给出 $m \leq \sqrt{\lfloor x \rfloor} < m + 1$ 。如此移去外部下整括号并不丢失任何信息。由于所有三个表达式是非负的，所以乘方得到 $m^2 \leq \lfloor x \rfloor < (m + 1)^2$ ，这就除去了平方根。接着我们移去下整括号，对于左边不等式用式(3.7(d))，对于右边不等式用式(3.7(a))： $m^2 \leq x < (m + 1)^2$ 。现在返回我们的步骤是简单的事情了，取平方根得 $m \leq \sqrt{x} < m + 1$ ，且用式(3.5(a))得 $m = \lfloor \sqrt{x} \rfloor$ ；因此 $\lfloor \sqrt{\lfloor x \rfloor} \rfloor = m = \lfloor \sqrt{x} \rfloor$ ，结论是真的。类似，我们能证明

$$\lceil \sqrt{\lceil x \rceil} \rceil = \lceil \sqrt{x} \rceil, \text{ 实 } x \geq 0.$$

刚才找到的证明并不十分依赖平方根的性质。更进一步的考虑表明我们能推广此想法且证明更多的结果：设 $f(x)$ 是任何连续，单调上升函数，且具有性质

$$f(x) = \text{整数} \Rightarrow x = \text{整数}.$$

(符号“ \Rightarrow ”意思是“蕴涵”。)于是每当 $f(x)$ ， $f(\lfloor x \rfloor)$ 和 $f(\lceil x \rceil)$ 被定义，我们有

$$\lfloor f(x) \rfloor = \lfloor f(\lfloor x \rfloor) \rfloor \text{ 和 } \lceil f(x) \rceil = \lceil f(\lceil x \rceil) \rceil. \quad (3.10)$$

让我们就上整来证明这个一般的性质，因为我们前面对下整的讨论以及下整的证明几乎是相同的。若 $x = \lceil x \rceil$ ，则没有什么可证的。否则 $x < \lceil x \rceil$ ，且 $f(x) < f(\lceil x \rceil)$ ，因为 f 是上升的。因此 $\lceil f(x) \rceil \leq \lceil f(\lceil x \rceil) \rceil$ ，因为 $\lceil \cdot \rceil$ 是非降的。若 $\lceil f(x) \rceil < \lceil f(\lceil x \rceil) \rceil$ ，则一定有一个数 y 使得 $x \leq y < \lceil x \rceil$ ，且 $f(y) = \lceil f(x) \rceil$ ，因为 f 是连续的。这个 y 是整数，因为 f 的特殊性质。但是在 x 和 $\lceil x \rceil$ 之间的确不能有一个整数。这个矛盾意味着我们一定有 $\lceil f(x) \rceil = \lceil f(\lceil x \rceil) \rceil$ 。

此定理的一个重要的特殊情形值得加以注意，若 m 和 n 是整数且分母 n 是正的，则有

$$\left\lfloor \frac{x+m}{n} \right\rfloor = \left\lfloor \frac{\lfloor x \rfloor + m}{n} \right\rfloor \quad \text{和} \quad \left\lceil \frac{x+m}{n} \right\rceil = \left\lceil \frac{\lceil x \rceil + m}{n} \right\rceil. \quad (3.11)$$

例如, 设 $m=0$, 我们有 $\lfloor \lfloor x/10 \rfloor / 10 \rfloor / 10 \rfloor = \lfloor x/1000 \rfloor$. 除 10 并去掉数字三次和除 1000 并去掉余数相同.

现在让我们尝试证明或推翻另一个命题:

$$\lfloor \sqrt{\lfloor x \rfloor} \rfloor = \lfloor \sqrt{x} \rfloor \quad \text{实 } x \geq 0.$$

当 $x=\pi$ 和 $x=e$ 时此等式成立, 但是当 $x=\varphi$ 时它不成立; 所以我们知道此等式一般不真.

在进一步讨论之前, 让我们离开片刻来讨论书中询问到的数学方面的不同“层次”的问题.

层次 1 给定一个明显的元素和一个明显的性质 $P(x)$, 证明 $P(x)$ 为真. 例如, “证明 $\lfloor \pi \rfloor = 3$ ”. 这里问题涉及到找某个表明事实的证明.

层次 2 给定一个明显的集合 X 和一个明显的性质 $P(x)$, 证明对于所有 $x \in X$, $P(x)$ 为真. 例如, “证明对于所有实 x , $\lfloor x \rfloor \leq x$ ”. 问题又涉及找一个证明, 但是此时证明一定要是一般的. 我们正在研究代数, 而不仅仅是算术.

层次 3 给定一个明显的集合 X 和一个明显的性质 $P(x)$, 证明或推翻对于所有 $x \in X$, $P(x)$ 为真, 例如, “证明或推翻对于所有实 $x \geq 0$, $\lfloor \sqrt{\lfloor x \rfloor} \rfloor = \lfloor \sqrt{x} \rfloor$ ”. 这里有一个不确定的附加层次, 但总归结果可以达到. 这是接近于一个数学家常面临的实际场合: 进入书本的断言往往会真, 但是新事物必须用有偏见的眼光来考虑. 若命题为假, 我们的任务是找一个反例. 若命题为真, 我们一定要找像层次 2 那样的一个证明.

层次 4 给定一个明显的集合 X 和一个明显的性质 $P(x)$, 找 $P(x)$ 为真的一个必要和充分条件 $Q(x)$. 例如, “找 $\lfloor x \rfloor \geq \lceil x \rceil$ 的一个必要和充分条件”. 问题是找 Q 使得 $P(x) \Leftrightarrow Q(x)$. 当然, 总有一个平凡的解答, 我们能取 $Q(x) = P(x)$. 但是暗指的要求是找一个尽可能简单的条件, 要求创造性地发现一个成立的简单条件. (例如, 假使这样的话, “ $\lfloor x \rfloor \geq \lceil x \rceil \Leftrightarrow x$ 是一个整数”.) 找 $Q(x)$ 所需发现的特佳元素使这类问题相当困难, 但是在“现实世界”中, 这是数学家一定要做的较典型的问题. 最后, 当然给出的一个证明一定要 $P(x)$ 为真当且仅当 $Q(x)$ 为真.

层次 5 给定一个明显的集合 X , 找出它的元素的一个有趣的性质 $P(x)$. 现在我们处于纯研究的可怕的范围, 在这里学生可能认为这完全是不规则的领域. 这是实际的数学, 教科书的作者很少敢询问层次 5 的问题.

结束离题的话, 而让我们把最后的问题从层次 3 变换到层次 4: $\lfloor \sqrt{\lfloor x \rfloor} \rfloor = \lfloor \sqrt{x} \rfloor$ 的必要和充分条件是什么? 我们已看到当 $x=3.142$ 时等式成立, 但是当 $x=1.618$ 时等式不成立; 进一步试验表明当 x 在 9 和 10 之间时它也不成立. 我们看到每当 $m^2 < x < m^2 + 1$ 时不成立的情形出现, 因为这就在左边给出 m 而在右边给出 $m+1$. 在所有其他 \sqrt{x} 被定义的情形中, 即当 $x=0$ 或 $m^2 + 1 \leq x \leq (m+1)^2$ 时, 我们取得等式. 所以下列陈述是相等的必要和充分条件: x 是一个整数, 或者 $\sqrt{\lfloor x \rfloor}$ 不是一个整数.

对于下一个问题让我们考虑由 C.A.R.Hoare 和 Lyle Ramshaw 提出的一种便于使用

的新记法: $[\alpha, \beta]$ 表示实数 x 使得 $\alpha \leq x \leq \beta$ 的集合。此集合称为闭区间, 因为它包含端点 α 和 β 。不包含端点而由所有 x 使得 $\alpha < x < \beta$ 组成的区间, 记为 (α, β) , 称为开区间。仅包含一个端点的区间 $[\alpha, \beta)$ 和 $(\alpha, \beta]$ 相似定义且称为半开区间。

在这样的区间中包含多少个整数? 半开区间较容易, 所以我们从半开区间开始。事实上半开区间几乎总是比开或闭区间好。例如, 它们是加性的, 我们能合并半开区间 $[\alpha, \beta)$ 和 $[\beta, \gamma)$ 形成半开区间 $[\alpha, \gamma)$ 。对于开区间这将行不通, 因为点 β 排斥在外, 而对于闭区间也将发生问题, 因为 β 将包含二次。

回到我们的问题上来。若 α 和 β 是整数, 解答是容易的: 假设 $\alpha \leq \beta$, 则 $[\alpha, \beta]$ 包含 $\beta - \alpha + 1$ 个整数 $\alpha, \alpha+1, \dots, \beta$ 。相似, $(\alpha, \beta]$ 包含 $\beta - \alpha$ 个整数。但是我们的问题较难, 因为 α 和 β 是任意实数。然而我们能把它转换成较容易的问题, 因为当 n 是一个整数时可按照式(3.7)

$$\begin{aligned}\alpha \leq n < \beta &\Leftrightarrow \lceil \alpha \rceil \leq n < \lceil \beta \rceil, \\ \alpha < n \leq \beta &\Leftrightarrow \lfloor \alpha \rfloor < n \leq \lfloor \beta \rfloor.\end{aligned}$$

右边的区间有整数端点且包含和左边区间相同个数的整数, 左边区间有实数端点。所以区间 $[\alpha, \beta]$ 恰好包含 $\lceil \beta \rceil - \lceil \alpha \rceil + 1$ 个整数, 且 $(\alpha, \beta]$ 包含 $\lceil \beta \rceil - \lceil \alpha \rceil$ 个整数。这就是我们实际要引入下整括号和上整括号的情形, 而不是除去它们。

顺便提一句, 有一种记忆法来记住用下整和上整的情形: 包含左端点而不包含右端点的半开区间(例如 $0 \leq \theta < 1$)稍微比包含右端点而不包含左端点的半开区间更常见, 且下整稍微比上整更常见。所以由 Murphy 定律, 符合一般性的规则是和我们所期望的(对 $[\alpha, \beta)$ 上整以及对 $(\alpha, \beta]$ 下整)相反。

相似的分析表明闭区间 $[\alpha, \beta]$ 包含 $\lfloor \beta \rfloor - \lfloor \alpha \rfloor + 1$ 个整数, 开区间 (α, β) 包含 $\lfloor \beta \rfloor - \lfloor \alpha \rfloor - 1$ 个整数, 虽然我们对后者加附加的限制 $\alpha \neq \beta$, 以致公式不会老是要表明空区间 (α, α) 总共包含 -1 个整数而使我们为难。为了概括起见, 我们已推出下列事实:

区间	包含的整数个数	限制
$[\alpha, \beta]$	$\lfloor \beta \rfloor - \lfloor \alpha \rfloor + 1$	$\alpha \leq \beta$,
$[\alpha, \beta)$	$\lceil \beta \rceil - \lceil \alpha \rceil$	$\alpha \leq \beta$,
(3.12)		
$(\alpha, \beta]$	$\lfloor \beta \rfloor - \lfloor \alpha \rfloor$	$\alpha < \beta$,
(α, β)	$\lceil \beta \rceil - \lceil \alpha \rceil - 1$	$\alpha < \beta$.

现在这里是一个我们接受的问题。具体数学俱乐部的娱乐场(仅对购买本书的人开放)有一个轮盘赌的轮, 具有编号 1 到 1000 的 1000 个位置。如果一次旋转发生的数 n 可被它的立方根的下整除尽, 也就是说, 如果

$$\lfloor \sqrt[3]{n} \rfloor \mid n,$$

则成为赢者, 且娱乐场付我们 5 元; 否则成为输者, 且我们一定要付 1 元。(记号 $a \mid b$ 读作“ a 除尽 b ”, 意味着 b 是 a 的确切的倍数; 第四章将仔细研究此关系。)如果我们玩此游戏, 能否指望赢钱?

通过首先算赢者个数 W 和输者个数 $L = 1000 - W$, 能计算平均赢的钱数. 也就是说每次游戏我们将赢(或输)的钱数. 若在 1000 次游戏期间每个数发生一次, 则我们赢 $5W$ 元和输 L 元, 所以平均赢的钱数将是

$$\frac{5W - L}{1000} = \frac{5W - (1000 - W)}{1000} = \frac{6W - 1000}{1000}$$

如果有 167 个或多于 167 个赢者, 则我们有好处; 否则好处属于娱乐场.

在 1 到 1000 中间如何能算出赢者的个数? 认出一种型式是不困难的. 从 1 到 $2^3 - 1 = 7$ 的数全是赢者, 因为对于每个这样的数 $\lfloor \sqrt[3]{n} \rfloor = 1$. 数 $2^3 = 8$ 到 $3^3 - 1 = 26$ 中间仅仅偶数是赢者, $3^3 = 27$ 到 $4^3 - 1 = 63$ 中间仅仅那些被 3 整除的数是赢者, 等等.

如果我们用第二章的求和技巧, 采用关于合理命题赋值 0 或 1 的 Iverson 的约定, 能系统分析整个安排:

$$\begin{aligned} W &= \sum_{n=1}^{1000} [n \text{ 是一个赢者}] \\ &= \sum_{1 \leq n \leq 1000} [\lfloor \sqrt[3]{n} \rfloor \setminus n] = \sum_{k,n} [k = \lfloor \sqrt[3]{n} \rfloor] [k \setminus n] [1 \leq n \leq 1000] \\ &= \sum_{k,m,n} [k^3 \leq n < (k+1)^3] [n = km] [1 \leq n \leq 1000] \\ &= 1 + \sum_{k,m} [k^3 \leq km < (k+1)^3] [1 \leq k < 10] \\ &= 1 + \sum_{k,m} [m \in [k^2, (k+1)^3/k]] [1 \leq k < 10] \\ &= 1 + \sum_{1 \leq k < 10} (\lceil k^2 + 3k + 3 + 1/k \rceil - \lceil k^2 \rceil) \\ &= 1 + \sum_{1 \leq k < 10} (3k + 4) = 1 + \frac{7+31}{2} \cdot 9 = 172. \end{aligned}$$

这个推导值得仔细研究. 注意, 第 6 行用到了一个半开区间中整数个数的公式(3.12). “困难”操作仅是在第 3 行和第 4 行之间作出决定来处理作为一种特殊情形的 $n = 1000$. (当 $k = 10$ 时不等式 $k^3 \leq n < (k+1)^3$ 不容易和 $1 \leq n \leq 1000$ 合并.) 一般, 有界条件往往会 Σ 操作的最关键部分.

最后一行表明 $W = 172$; 因此每次游戏的平均赢得的钱的公式化为 $(6 \cdot 172 - 1000) / 1000$ 元, 它为 3.2 分. 我们能期望在做 100 次每次 1 元的打赌之后近于有 3.20 元. (当然, 娱乐场可能使有些数比其他数更合适.)

我们刚才解的娱乐场问题是更常见问题的一种修整形式, “有多少个 n , 其中 $1 \leq n \leq 1000$, 满足关系 $\lfloor \sqrt[3]{n} \rfloor \setminus n$?” 从数学上来说这两个问题是相同的. 但是有时候修整一个问题是一种好主意. 我们开始用更多的用语(像“赢者”和“输者”), 它有助于了解发生的情况.

让我们使其一般化. 假设改变 1000 为 1 000 000, 或改变为更大的数 N , (我们假设娱乐场有顾客, 且能有一个较大的轮.) 现在在那里有多少个赢者?

用相同论证, 但是我们需要更仔细地讨论 k 的最大值, 为了方便起见我们称它为 K :
 $K = \lfloor \sqrt[3]{N} \rfloor$.

(以前 K 是 10.) 对于一般的 N , 赢者总数共计

$$\begin{aligned} W &= \sum_{1 \leq k < K} (3k + 4) + \sum_{m} [K^3 \leq Km \leq N] \\ &= \frac{1}{2}(7 + 3K + 1)(K - 1) + \sum_{m} [m \in [K^2 \dots N/K]]. \\ &= \frac{3}{2}K^2 + \frac{5}{2}K - 4 + \sum_{m} [m \in [K^2 \dots N/K]]. \end{aligned}$$

我们知道剩下的和是 $\lfloor N/K \rfloor - \lfloor K^2 \rfloor + 1 = \lfloor N/K \rfloor - K^2 + 1$; 因此公式

$$W = \lfloor N/K \rfloor + \frac{1}{2}K^2 + \frac{5}{2}K - 3, \quad K = \lfloor \sqrt[3]{N} \rfloor \quad (3.13)$$

给出了大小为 N 的轮的一般解答.

此公式的前两项近似为 $N^{2/3} + (1/2)N^{2/3} = (3/2)N^{2/3}$, 当 N 大时, 其他项与前两项比较是十分小的. 在第九章中我们将学到如何来推得像

$$W = \frac{3}{2}N^{2/3} + O(N^{1/3})$$

的表达式, 其中 $O(N^{1/3})$ 表示一个不大于常数乘 $N^{1/3}$ 的量. 不管是什么常数, 我们知道它是独立于 N 的; 所以对于大的 N 来说, 在 W 中 O 项与 $(3/2)N^{2/3}$ 相比是十分小的. 例如, 下表表明 $(3/2)N^{2/3}$ 和 W 有多么接近:

N	$\frac{3}{2}N^{2/3}$	W	误差(%)
1 000	150.0	172	12.791
10 000	696.2	746	6.670
100 000	3 231.7	3 343	3.331
1 000 000	15 000.0	15 247	1.620
10 000 000	69 623.8	70 158	0.761
100 000 000	323 165.2	324 322	0.357
1 000 000 000	1 500 000.0	1 502 496	0.166

这是一个十分好的近似.

近似公式是有用的, 因为它们比具有下整和上整的公式简单. 而特别对于实际中往往会出现较小的 N 值, 确切的真值常常也是重要的. 例如, 娱乐场所有者可能虚假地假定当 $N = 1\,000$ 时仅有 $(3/2)N^{2/3} = 150$ 个赢者(此时对于娱乐场将有 10 分钱的好处).

本节中最后的应用考虑所谓的谱. 我们定义实数 α 的谱是无限的整数多重集,
 $\text{Spec}(\alpha) = \{\lfloor \alpha \rfloor, \lfloor 2\alpha \rfloor, \lfloor 3\alpha \rfloor, \dots\}$.

(一个多重集是能有重复元素的一个集合。)例如, $1/2$ 的谱开始有 $\{0, 1, 1, 2, 2, 3, 3, \dots\}$ 。

易证没有两个谱是相等的, $\alpha \neq \beta$ 意味着 $\text{Spec}(\alpha) \neq \text{Spec}(\beta)$ 。不失一般性, 假设 $\alpha < \beta$, 有正整数 m 使得 $m(\beta - \alpha) \geq 1$ 。(事实上, 有 $m \geq \lceil 1/(\beta - \alpha) \rceil$ 就够了, 不需炫耀下整和上整的知识。)因此 $m\beta - m\alpha \geq 1$, 且 $\lfloor m\beta \rfloor > \lfloor m\alpha \rfloor$ 。于是当 $\text{Spec}(\alpha)$ 至少具有 m 时, $\text{Spec}(\beta)$ 少于 m 个 ($\leq \lfloor m\alpha \rfloor$) 元素。

谱具有许多极好的性质。例如, 考虑两个多重集

$$\begin{aligned}\text{Spec}(\sqrt{2}) &= \{1, 2, 4, 5, 7, 8, 9, 11, 12, 14, 15, 16, 18, 19, 21, 22, 24, \dots\}, \\ \text{Spec}(2 + \sqrt{2}) &= \{3, 6, 10, 13, 17, 20, 23, 27, 30, 34, 37, 40, 44, 47, 51, \dots\}.\end{aligned}$$

用袖珍计算器容易计算 $\text{Spec}(\sqrt{2})$, 依据式(3.6), $\text{Spec}(2 + \sqrt{2})$ 的第 n 个元素就是 $2n$, 它大于 $\text{Spec}(\sqrt{2})$ 的第 n 个元素。再仔细看一看显示, 这两个谱以一种十分意外的方式联系起来: 从外表上显出, 在一个中缺掉的任何数存在于另一个中, 但是没有数同时在两个中! 确实: 正整数是 $\text{Spec}(\sqrt{2})$ 和 $\text{Spec}(2 + \sqrt{2})$ 的不相交的并。我们称这些谱形成正整数的一个划分。

为了证明这个结论, 我们将算有多少个 $\text{Spec}(\sqrt{2})$ 的元素是 $\leq n$, 有多少个 $\text{Spec}(2 + \sqrt{2})$ 的元素 $\leq n$ 。如果对于每个 n , 总和是 n , 则这样两个谱确实划分了整数。

设 α 是正的。 $\text{Spec}(\alpha)$ 中元素的个数 ($\leq n$) 为

$$\begin{aligned}N(\alpha, n) &= \sum_{k \geq 0} \lfloor k\alpha \rfloor \leq n \\ &= \sum_{k \geq 0} \lfloor k\alpha \rfloor < n + 1 \\ &= \sum_{k \geq 0} [k\alpha < n + 1] \\ &= \sum_k [0 < k < (n + 1)/\alpha] \\ &= \lfloor (n + 1)/\alpha \rfloor - 1.\end{aligned}\tag{3.14}$$

此推导有两处特别有趣的地方。第一, 它用了定律

$$m \leq n \iff m < n + 1, \text{ 整数 } m \text{ 和 } n \tag{3.15}$$

把 ' \leq ' 改变成 ' $<$ ', 以致依据式(3.7)能移去下整括号。而且更微妙的是它在范围 $k > 0$ 上求和而不是 $k \geq 1$, 因为对于某个 n 和 α , $(n+1)/\alpha$ 可以小于 1。若尝试应用式(3.12)来决定 $[1..(n+1)/\alpha]$ 中的整数个数, 而不是 $(0..(n+1)/\alpha)$ 中的整数个数, 我们将得到正确的答案; 但是我们的推导将有错误, 因为将遇不到可应用的条件。

我们有了 $N(\alpha, n)$ 的公式, 现在可通过用式(3.14)测试是否对于所有整数 $n > 0$, $N(\sqrt{2}, n) + N(2 + \sqrt{2}, n) = n$ 来验证是否 $\text{Spec}(\sqrt{2})$ 和 $\text{Spec}(2 + \sqrt{2})$ 划分正整数:

$$\left\lceil \frac{n+1}{\sqrt{2}} \right\rceil - 1 + \left\lceil \frac{n+1}{2+\sqrt{2}} \right\rceil - 1 = n$$

$$\Leftrightarrow \left\lfloor \frac{n+1}{\sqrt{2}} \right\rfloor + \left\lfloor \frac{n+1}{2+\sqrt{2}} \right\rfloor = n, \quad \text{依据式(3.2);}$$

$$\Leftrightarrow \frac{n+1}{\sqrt{2}} - \left\{ \frac{n+1}{\sqrt{2}} \right\} + \frac{n+1}{2+\sqrt{2}} - \left\{ \frac{n+1}{2+\sqrt{2}} \right\} = n, \quad \text{依据式(3.8).}$$

现在由于有简洁的等式

$$\frac{1}{\sqrt{2}} + \frac{1}{2+\sqrt{2}} = 1$$

而一切易做; 我们的条件也化为测试是否对于所有 $n > 0$,

$$\left\{ \frac{n+1}{\sqrt{2}} \right\} + \left\{ \frac{n+1}{2+\sqrt{2}} \right\} = 1.$$

因为这是两个非整数的数合计为整数 $n+1$ 的分数部分, 我们成功了. 它是一个划分.

3.3 下整 / 上整递归

我们进而讲下整和上整的一个有趣的新的方面, 就是研究递归关系. 下面首先研究递归

$$K_0 = 1;$$

$$K_{n+1} = 1 + \min(2K_{\lfloor n/2 \rfloor}, 3K_{\lfloor n/3 \rfloor}), \quad n \geq 0. \quad (3.16)$$

因而, 例如, K_1 是 $1 + \min(2K_0, 3K_0) = 3$; 序列起始为 1, 3, 3, 4, 7, 7, 7, 9, 9, 10, 13, ... 本书的作者之一 Knuth 称这些数为 Knuth 数.

习题 25 要求证明或推翻对于所有 $n \geq 0$, $K_n \geq n$. 刚才列出的前面少数几个 K 满足不等式, 所以有希望这是一般成立的. 让我们尝试用归纳法证明, 基础 $n=0$ 直接得自定义的递归. 对于归纳步, 假设对于由小到大的某个指定的非负 n 的所有值不等式成立, 我们尝试证明 $K_{n+1} \geq n+1$. 由递归可知 $K_{n+1} = 1 + \min(2K_{\lfloor n/2 \rfloor}, 3K_{\lfloor n/3 \rfloor})$. 归纳假设告诉我们 $2K_{\lfloor n/2 \rfloor} \geq 2\lfloor n/2 \rfloor$, $3K_{\lfloor n/3 \rfloor} \geq 3\lfloor n/3 \rfloor$. 可是, $2\lfloor n/2 \rfloor$ 能和 $n-1$ 一样小, $3\lfloor n/3 \rfloor$ 能和 $n-2$ 一样小. 根据我们的归纳假设最多能推得 $K_{n+1} \geq 1 + (n-2)$, 这远达不到 $K_{n+1} \geq n+1$.

我们现在有理由关心 $K_n \geq n$ 的真值性, 所以让我们尝试推翻它. 若能找到一个 n 使得 $2K_{\lfloor n/2 \rfloor} < n$ 或者 $3K_{\lfloor n/3 \rfloor} < n$ 或换句话说使得

$$K_{\lfloor n/2 \rfloor} < n/2 \quad \text{或} \quad K_{\lfloor n/3 \rfloor} < n/3,$$

我们将有 $K_{n+1} < n+1$. 这可能吗? 在这里我们不给出解答, 而留作习题 25.

涉及下整和/或上整的递归关系经常在计算机科学中出现,因为基于重要的“公治”算法常把大小为 n 的一个问题化成大小为 n 的几分之一(整数)的相同问题的解。例如,分类 n 个记录的一种方法(若 $n > 1$)是把它们分成两个近似相等的部分,一部分大小为 $\lceil n/2 \rceil$,另一部分大小为 $\lfloor n/2 \rfloor$ 。(顺便提一句,注意

$$n = \lceil n/2 \rceil + \lfloor n/2 \rfloor, \quad (3.17)$$

此公式迟早要用到。)每部分分开分类之后(用相同方法,递归应用),至多再作 $n-1$ 次比较能把记录合并为最终的次序。所以执行的比较的总次数至多为 $f(n)$,其中

$$\begin{aligned} f(1) &= 0; \\ f(n) &= f(\lceil n/2 \rceil) + f(\lfloor n/2 \rfloor) + n - 1, \quad n > 1. \end{aligned} \quad (3.18)$$

在习题 34 中出现此递归的解。

第一章中的 Josephus 问题具有一个类似的递归,它能变成形式

$$\begin{aligned} J(1) &= 1; \\ J(n) &= 2J(\lfloor n/2 \rfloor) - (-1)^n, \quad n > 1. \end{aligned}$$

我们已经取得比第一章多的使用工具,所以让我们考虑更正式的 Josephus 问题,问题中消除所有第 3 个人,而不是所有第 2 个人。若我们应用第一章中使用的方法到这个更困难的问题,我们最终得递归。

$$J_3(n) = \left[\frac{3}{2} J_3\left(\left\lfloor \frac{3}{2} n \right\rfloor\right) + a_n \right] \bmod n + 1,$$

其中‘mod’是我们将要作简短研究的一个函数,且按照 $n \bmod 3 = 0, 1$ 或 2 而有 $a_n = -2, +1$,或 $-1/2$ 。但是这个递归作进一步讨论太可怕了。

有另一种探讨 Josephus 问题的方式,给出了一种十分好的构造。每当一个人越过,我们能赋予一个新数。于是,1 和 2 变成 $n+1$ 和 $n+2$,则 3 被执行;4 和 5 变成 $n+3$ 和 $n+4$,则 6 被执行;..., $3k+1$ 和 $3k+2$ 变成 $n+2k+1$ 和 $n+2k+2$,则 $3k+3$ 被执行;...,则 $3n$ 被执行(或留下幸存)。例如,当 $n=10$,数是

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						

数 $3k$ 最后以消除第 k 个人而告终。所以如果我们能计算出人数 $3n$ 的原先的数,则我们能断定谁是幸存者。

若 $N > n$, 人数 N 一定要有一个以前的数,且我们能找到它。如下:我们有

$N = n + 2k + 1$ 或 $N = n + 2k + 2$, 因此 $k = \lfloor (N - n - 1) / 2 \rfloor$, 以前的数分别是 $3k + 1$ 或 $3k + 2$. 也就是说, 它是 $3k + (N - n - 2k) = k + N - n$. 因此我们能计算幸存者数 $J_3(n)$ 如下:

```

 $N := 3n;$ 
while  $N > n$  do  $N := \left\lfloor \frac{N - n - 1}{2} \right\rfloor + N - n;$ 
 $J_3(n) := N.$ 

```

这不是 $J_3(n)$ 的闭形式; 甚至连递归都不是. 但是至少它告诉我们如果 n 是大的数, 如何相当快地计算解答.

幸运的是, 如果我们用变量 $D = 3n + 1 - N$ 来替换 N , 有一种方法来简化此算法. (此表示法的改变对应于赋予 $3n$ 下降到 1 的数, 代替从 1 上升到 $3n$, 就像一种递减计数的分类.) 于是结构复杂的 N 的赋值变成

$$\begin{aligned} D &:= 3n + 1 - \left(\left\lfloor \frac{(3n + 1 - D) - n - 1}{2} \right\rfloor + (3n + 1 - D) - n \right) \\ &= n + D - \left\lfloor \frac{2n - D}{2} \right\rfloor = D - \left\lfloor \frac{-D}{2} \right\rfloor = D + \left\lceil \frac{D}{2} \right\rceil = \left\lceil \frac{3}{2} D \right\rceil, \end{aligned}$$

我们可把算法改写如下:

```

 $D := 1;$ 
while  $D \leq 2n$  do  $D := \left\lceil \frac{3}{2} D \right\rceil;$ 
 $J_3(n) := 3n + 1 - D.$ 

```

这看来十分好, 因为 n 以十分简单的方法进入计算. 事实上, 以同样理由我们可表明当消除所有第 q 个人时, 幸存者 $J_q(n)$ 能计算如下:

```

 $D := 1;$ 
while  $D \leq (q - 1)n$  do  $D := \left\lceil \frac{q}{q - 1} D \right\rceil;$ 
 $J_q(n) := qn + 1 - D.$ 

```

(3.19)

$q = 2$ 的情形我们十分了解, 当 $n = 2^m + l$ 时, 这使 D 增长到 2^{m+1} ; 因此 $J_2(n) = 2(2^m + l) + 1 - 2^{m+1} = 2l + 1$.

式(3.19)中的方法计算出能由下列递归定义的一个整数序列:

$$\begin{aligned} D_0^{(q)} &= 1; \\ D_n^{(q)} &= \left\lceil \frac{q}{q-1} D_{n-1}^{(q)} \right\rceil \quad n > 0. \end{aligned} \quad (3.20)$$

除了当 $q = 2$ 之外, 看来这些数并不以简单的方式和任何熟悉的函数有联系, 因此它们也许没有一种好的闭形式. 但是如果如同“已知”那样, 我们愿意接受序列 $D_n^{(q)}$, 则容易描述

广义 Josephus 问题的解: 幸存者 $J_q(n)$ 为 $qn + 1 - D_k^{(q)}$, 其中 k 是尽可能小的数使得 $D_k^{(q)} > (q-1)n$.

3.4 ‘MOD’: 二元运算

当 m 和 n 是正整数时, m 除 n 的商是 $\lfloor n/m \rfloor$. 对于此除法的剩余也有一个便于使用的简单表示法, 且我们称它为 ‘ $n \bmod m$ ’. 基本公式

$$n = m \underbrace{\lfloor n/m \rfloor}_{\text{商}} + \underbrace{n \bmod m}_{\text{剩余}}$$

告诉我们, 我们能把 $n \bmod m$ 表达为 $n - m\lfloor n/m \rfloor$. 我们能把此推广到负整数, 实际上推广到任意实数:

$$x \bmod y = x - y\lfloor x/y \rfloor, \quad y \neq 0. \quad (3.21)$$

这就把 ‘mod’ 定义为一个二元运算, 就像加法和减法是二元运算那样. 数学已经非形式地这样用 mod 很久了, 取得各种量 mod 10, mod 2π 等等, 但是最近 20 年才形式地理解它. 老的概念, 新的表示法.

当 x 和 y 是正实数时, 我们容易领会 $x \bmod y$ 的直观意义, 如果我们想象圆周 y 的一个圆, 它的点被赋予区间 $[0, y)$ 中的实数. 如果我们围绕圆移动距离 x , 从 0 开始, 在 $x \bmod y$ 处结束. (当我们行进时遇到 0 的次数为 $\lfloor x/y \rfloor$.)

当 x 或 y 是负时, 为了恰切地理解它的意义, 我们需要仔细看一看定义. 此处是一些整值的例子:

$$\begin{aligned} 5 \bmod 3 &= 5 - 3\lfloor 5/3 \rfloor = 2; \\ 5 \bmod -3 &= 5 - (-3)\lfloor 5/(-3) \rfloor = -1; \\ -5 \bmod 3 &= -5 - 3\lfloor -5/3 \rfloor = 1; \\ -5 \bmod -3 &= -5 - (-3)\lfloor -5/(-3) \rfloor = -2. \end{aligned}$$

‘mod’ 后面的数称为模数; 至今无人来称呼 ‘mod’ 前面的数. 在应用中, 模数通常是正的, 但是当模数为负时, 定义完全有意义. 不管哪一种情形, $x \bmod y$ 的值在 0 和模数之间:

$$\begin{aligned} 0 &\leq x \bmod y < y, \quad y > 0; \\ 0 &\geq x \bmod y > y, \quad y < 0. \end{aligned}$$

关于 $y=0$ 如何呢? 此时定义式(3.21)无定义, 为了避免被零除, 同时也为了完整起见, 我们可定义

$$x \bmod 0 = x. \quad (3.22)$$

此约定保持了 $x \bmod y$ 总与 y 的倍数相差为 x 的性质。(看来通过定义 $x \bmod 0 = \lim_{y \rightarrow 0} x \bmod y = 0$ 可以较自然地使函数在 0 处连续。但是在第四章中我们将看到这是无用的。连续性不是 mod 运算的重要方面。)

我们已经看到 mod 的一种不易识别的特殊情形，当我们用 x 的整数和分数部分记它时， $x = \lfloor x \rfloor + \{x\}$ 。分数部分也能记为 $x \bmod 1$ ，因为我们有

$$x = \lfloor x \rfloor + x \bmod 1.$$

注意，在此公式中不需要括号，我们把 mod 理解为比加法或减法结合更紧的运算。

下整函数用来定义 mod，而同时没有用上整函数。我们也许能用上整来定义一个类似 mod 的记法。像

$$x \text{ mumble } y = y[x / y] - x;$$

在我们的圆的相似形中，此表示行进者需继续行进的距离，在行进距离 x 之后，返回到起始点 0。当然我们需要一个比‘mumble’更好的名字。如果出现足够的应用，也许将给它取一个适当的名字。

分配律是 mod 的最重要的代数性质，我们有

$$c(x \bmod y) = (cx) \bmod (cy), \quad (3.23)$$

对于所有实数 c ， x 和 y 。(就像 mod 比乘法结合得松，还可以从右边移去括号。)根据定义(3.21)易证分配律，因为若 $cy \neq 0$ ，则

$$c(x \bmod y) = c(x - y\lfloor x / y \rfloor) = cx - cy\lfloor cx / cy \rfloor = cx \bmod cy,$$

且零模数情形平凡地成立。我们的四个例子(用 ± 5 和 ± 3)两次说明了这个定律，用 $c = -1$ 。像式(3.23)的等式使人放心，因为它使我们相信‘mod’已适当地定义。

在本节余下部分中，我们将考虑一种应用，‘mod’虽然不起主要作用，但它有用。在各种场合常出现问题：我们要把 n 个事物划分为尽可能相等的 m 组。

例如，假设我们有 n 个短行的版本，我们希望把它排成 m 列。由于整齐的原因，我们要把列按长度的下降次序排列(实际上为非升次序)；且长度将接近相同，任何两列之差不大于版本的一行的数量。若把版本的 37 行分成 5 列，我们宁愿要右端的排列：

8	8	8	8	5	8	8	7	7	7
行 1	行 9	行 17	行 25	行 33	行 1	行 9	行 17	行 24	行 31
行 2	行 10	行 18	行 26	行 34	行 2	行 10	行 18	行 25	行 32
行 3	行 11	行 19	行 27	行 35	行 3	行 11	行 19	行 26	行 33
行 4	行 12	行 20	行 28	行 36	行 4	行 12	行 20	行 27	行 34
行 5	行 13	行 21	行 29	行 37	行 5	行 13	行 21	行 28	行 35
行 6	行 14	行 22	行 30		行 6	行 14	行 22	行 29	行 36
行 7	行 15	行 23	行 31		行 7	行 15	行 23	行 30	行 37
行 8	行 16	行 24	行 32		行 8	行 16			

此外我们要按列分配版本的行, 首先决定多少行进入第一列, 然后移到第二列, 第三列, 等等, 因为这是人们阅读的方式。逐行分配将在每列中给出正确的行数, 但是次序是不对的。(我们将取右端那样的排列, 但是列 1 将包含行 1, 6, 11, ..., 36 而不是希望的行 1, 2, 3, ..., 8。)

不能用逐行分配对策, 但是它告诉我们放入每列中的多少行。若 n 不是 m 的倍数, 逐行分配过程清楚地表明长列的每一列将包含 $\lceil n/m \rceil$ 行, 而短列的每一列将包含 $\lfloor n/m \rfloor$ 行。恰好有 $n \bmod m$ 个长列(正如产生的那样, 恰好有 $n \bmod m$ 个短列)。

让我们推广术语, 谈论‘事物’和‘组’来代替‘行’和‘列’。我们刚才已决定, 第一组将包含 $\lceil n/m \rceil$ 个事物; 所以下列的序列分配方案应利用: 当 $m > 0$ 时, 为了分配 n 个事物成 m 组, 把 $\lceil n/m \rceil$ 个事物放入一组, 然后递归地用相同过程把余下的 $n' = n - \lceil n/m \rceil$ 个事物放入 $m' = m - 1$ 个另外的组。

例如, 若 $n = 314$, $m = 6$, 分配进行如下:

余下事物	余下组	「事物/组」
314	6	53
261	5	53
208	4	52
156	3	52
104	2	52
52	1	52

这是行得通的。我们取接近相同大小的组, 纵然除数保持改变。

为何它见效? 一般我们能假设 $n = qm + r$, 其中 $q = \lfloor n/m \rfloor$, $r = n \bmod m$ 。若 $r = 0$, 处理简单:

我们把 $\lceil n/m \rceil = q$ 个事物放入第一组, 且用 $n' = n - q$ 替换 n , 把留下的 $n' = qm'$ 个事物放入余下的 $m' = m - 1$ 个组。如果 $r > 0$, 我们把 $\lceil n/m \rceil = q + 1$ 个事物放入第一组, 且把 $n' = n - q - 1$ 替换 n , 留下的 $n' = qm' + r - 1$ 个事物供后继的组。新剩余为 $r' = r - 1$, 但是 q 保留相同。由此得出将有 r 个具有 $q + 1$ 个事物的组, 后面是 $m - r$ 个具有 q 个事物的组。

在第 k 组中有多少个事物呢? 我们希望有一个公式, 当 $k \leq n \bmod m$ 时它给出 $\lceil n/m \rceil$, 其他情况给出 $\lfloor n/m \rfloor$ 。不难验证

$$\left\lceil \frac{n - k + 1}{m} \right\rceil$$

具有希望的性质, 因为正如在前一段中那样, 如果我们记 $n = qm + r$, 这就化为 $q + \lceil (r - k + 1)/m \rceil$, 这里 $q = \lfloor n/m \rfloor$ 。如果 $1 \leq k \leq m$ 和 $0 \leq r < m$, 我们有 $\lceil (r - k + 1)/m \rceil = \lfloor (k - 1)/m \rfloor$ 。所以我们能定出一个等式, 它表示出 n 划分成以非升次序尽可能相等的 m 个部分:

$$n = - \left\lfloor \frac{n}{m} \right\rfloor + \left\lfloor \frac{n-1}{m} \right\rfloor + \cdots + \left\lfloor \frac{n-m+1}{m} \right\rfloor. \quad (3.24)$$

此等式对所有正整数 m , 以及所有整数 n (是正的, 负的, 或零)成立。在式(3.17)中我们已

遇到 $m=2$ 的情形, 在那里我们以一种稍微不同的形式写出了它, $n = \lceil n/2 \rceil + \lfloor n/2 \rfloor$.

如果我们要的部分是以非降次序, 小的组在大的组之前进行, 我们能以相同方式处理, 但是第一组中具有 $\lfloor n/m \rfloor$ 个事物. 于是我们将导出对应的等式

$$n = \left\lfloor \frac{n}{m} \right\rfloor + \left\lfloor \frac{n+1}{m} \right\rfloor + \cdots + \left\lfloor \frac{n+m-1}{m} \right\rfloor. \quad (3.25)$$

用式(3.4)或习题 12 的等式, 式(3.25)和(3.24)之间的转换是可能的.

现在如果我们在式(3.25)中用 $\lfloor mx \rfloor$ 替换 n , 且应用规则(3.11)移去下整内部的下整记号, 我们获得一个对于所有实数 x 成立的等式:

$$\lfloor mx \rfloor = \lfloor x \rfloor + \left\lfloor x + \frac{1}{m} \right\rfloor + \cdots + \left\lfloor x + \frac{m-1}{m} \right\rfloor. \quad (3.26)$$

这是令人惊异的, 因为下整函数是接近一个实值的整数, 而左边的一个近似等于右边它们的一串和. 如果我们设 $\lfloor x \rfloor$ 平均来说约略为 $x-1/2$, 左边约略为 $mx-1/2$, 而右边约略为 $(x-1/2) + (x-1/2+1/m) + \cdots + (x-1/2+(m-1)/m) = mx-1/2$, 出现的所有这些粗略近似的和是正确的!

3.5 下整/上整的和

方程(3.26)展示了至少一种涉及 $\lfloor \cdot \rfloor$ 的和取得闭形式是可能的. 有其他的吗? 是的, 在这样的情形中通常用的技巧是通过引入一个新的变量除去下整或上整.

例如, 让我们来看一看是否可能把和

$$\sum_{0 \leq k < n} \lfloor \sqrt{k} \rfloor$$

化成闭形式. 一种想法是引入变量 $m = \lfloor \sqrt{k} \rfloor$, 我们能如同轮盘赌问题中所做的那样“机械地”处理:

$$\begin{aligned} \sum_{0 \leq k < n} \lfloor \sqrt{k} \rfloor &= \sum_{k, m \geq 0} m [k < n] [m = \lfloor \sqrt{k} \rfloor] \\ &= \sum_{k, m \geq 0} m [k < n] [m \leq \sqrt{k} < m+1] \\ &= \sum_{k, m \geq 0} m [k < n] [m^2 \leq k < (m+1)^2] \\ &= \sum_{k, m \geq 0} m [m^2 \leq k < (m+1)^2 \leq n] \\ &\quad + \sum_{k, m \geq 0} m [m^2 \leq k < n < (m+1)^2]. \end{aligned}$$

但有界条件有些难处理。让我们首先设 $n = a^2$ 是完全平方。于是第二个和为零，而第一个和能以通常方式求值：

$$\begin{aligned}
 \sum_{k, m \geq 0} m[m^2 \leq k < (m+1)^2 \leq a^2] \\
 &= \sum_{m \geq 0} m((m+1)^2 - m^2)[m+1 \leq a] \\
 &= \sum_{m \geq 0} m(2m+1)[m < a] \\
 &= \sum_{m \geq 0} (2m^2 + 3m)[m < a] \\
 &= \sum_0^a (2m^2 + 3m) \delta m \\
 &= \frac{2}{3} a(a-1)(a-2) + \frac{3}{2} a(a-1) = \frac{1}{6} (4a+1)a(a-1).
 \end{aligned}$$

在一般情形中，我们能设 $a = \lfloor \sqrt{n} \rfloor$ ；我们仅需加 $a^2 \leq k < n$ 的项，这些项全等于 a ，所以它们共计 $(n - a^2)a$ 。这就给出希望的闭形式，

$$\sum_{0 \leq k < n} \lfloor \sqrt{k} \rfloor = na - \frac{1}{3}a^3 - \frac{1}{2}a^2 - \frac{1}{6}a, \quad a = \lfloor \sqrt{n} \rfloor. \quad (3.27)$$

另一种处理这样的和的方法是用 $\sum_j [1 \leq j \leq x]$ 替换形式 $\lfloor x \rfloor$ 的表达式，每当 $x \geq 0$ 时这是合法的。这里是如何处理 $\lfloor \sqrt{k} \rfloor$ 的和的方法，如果为了方便我们设 $n = a^2$ ，

$$\begin{aligned}
 \sum_{0 \leq k < n} \lfloor \sqrt{k} \rfloor &= \sum_{j, k} [1 \leq j \leq \sqrt{k}] [0 \leq k < a^2] \\
 &= \sum_{1 \leq j < a} \sum_k [j^2 \leq k < a^2] \\
 &= \sum_{1 \leq j < a} (a^2 - j^2) = a^3 - \frac{1}{3}a \left(a + \frac{1}{2} \right) (a+1).
 \end{aligned}$$

现在这里就是改变变量，引出一个变换的和的另一个例子。大约在 1909 年的相同时期，三个数学家 Bohl^[28]，Sierpiński^[265]和 Weyl^[300]独立地发现了一个值得注意的定理：如果 α 是无理数，则当 $n \rightarrow \infty$ 时，分数部分 $\{n\alpha\}$ 非常均匀地分布在 0 和 1 之间。叙述这一点的一种方法是，对于所有无理数 α 和所有几乎处处连续的函数 f ，

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{0 \leq k < n} f(\{k\alpha\}) = \int_0^1 f(x) dx. \quad (3.28)$$

例如，置 $f(x) = x$ 能找到 $\{n\alpha\}$ 的平均值，我们获得 $1/2$ 。（这恰好就是我们所期望的，但是知道不管无理数 α 为何确实可证明成立是有益的。）

让我们通过简单函数

$$f_v(x) = [0 \leq x < v]$$

的线性组合的“阶梯”函数上下逼近 $f(x)$ 来证明 Bohl, Sierpiński 和 Weyl 的定理。在这里我们的目的不是证明定理，那是微积分书的任务；但是让我们尝试通过看出在特殊情形 $f(x) = f_v(x)$ 时它处理得多么好，而领会到它成立的基本原因。换句话说，让我们尝试看出，当 n 很大且 α 是无理数时，和

$$\sum_{0 \leq k < n} [\{k\alpha\} < v]$$

到达“理想”值 nv 多近。

为此，我们定义偏差 $D(\alpha, n)$ 为所有 $0 \leq v \leq 1$ 上和

$$s(\alpha, n, v) = \sum_{0 \leq k < n} ([\{k\alpha\} < v] - v) \quad (3.29)$$

的最大绝对值。

我们的目标是通过表明 $|s(\alpha, n, v)|$ 总是适当地小，来证明与 n 比较 $D(\alpha, n)$ “不太大”。

首先我们能把 $s(\alpha, n, v)$ 改写成较简单的形式，然后引入一个新指标变量 j ：

$$\begin{aligned} \sum_{0 \leq k < n} ([\{k\alpha\} < v] - v) &= \sum_{0 \leq k < n} (k\alpha - [k\alpha] - v) \\ &= -nv + \sum_{0 \leq k < n} \sum_j [k\alpha - v < j \leq k\alpha] \\ &= -nv + \sum_{0 \leq j < [n\alpha]} \sum_{k < n} [j\alpha^{-1} \leq k < (j+v)\alpha^{-1}]. \end{aligned}$$

如果幸运，我们能在 k 上处理和，但是应引入某些新变换，以使公式将不致这样凌乱。不失一般性，我们能设 $0 < \alpha < 1$ ，让我们记

$$\begin{aligned} a &= [\alpha^{-1}], \quad \alpha^{-1} = a + \alpha'; \\ b &= [v\alpha^{-1}], \quad v\alpha^{-1} = b + v'. \end{aligned}$$

于是 $\alpha' = \{\alpha^{-1}\}$ 是 α^{-1} 的分数部分，而 v' 是 $v\alpha^{-1}$ 的 mumble 分数部分。

有界条件再次成为我们遭遇困难仅有的根源。现在让我们不考虑限制 ' $k < n$ '，且在它的条件下在 k 上计算和：

$$\begin{aligned} \sum_k \left[k \in [j\alpha^{-1}, (j+v)\alpha^{-1}] \right] &= [(j+v)(a + \alpha')] - [j(a + \alpha')] \\ &= b + [j\alpha' + v'] - [j\alpha']. \end{aligned}$$

这是相当简单的。我们由此式得到：

$$s(\alpha, n, v) = -nv + \lfloor n\alpha \rfloor b + \sum_{0 \leq j < \lfloor n\alpha \rfloor} (\lceil j\alpha' - v' \rceil - \lceil j\alpha' \rceil) - S, \quad (3.30)$$

其中 S 是未排除的 $k \geq n$ 情形的校正值。量 $j\alpha'$ 将不是一个整数，因为 α (因此 α') 是无理数，且对于至多 j 的一个值 $j\alpha' - v'$ 将是一个整数。所以能把上整项改变成下整项：

$$s(\alpha, n, v) = -nv + \lceil n\alpha \rceil b - \sum_{0 \leq j < \lceil n\alpha \rceil} (\lfloor j\alpha' \rfloor - \lfloor j\alpha' - v' \rfloor) - S + [0 \text{ 或 } 1].$$

有趣的是，我们取得了一个看起来颇像 $s(\alpha, n, v)$ 的和，但具有不同的参数： α' 代替 α ， $\lceil n\alpha \rceil$ 代替 n ， v' 代替 v ，而不是一个闭形式。所以我们将得到一个 $s(\alpha, n, v)$ 的递归式，它将 (有希望) 引出偏差 $D(\alpha, n)$ 的一个递归式。这意味着，我们要

$$s(\alpha', \lceil n\alpha \rceil, v') = \sum_{0 \leq j < \lceil n\alpha \rceil} (\lfloor j\alpha' \rfloor - \lfloor j\alpha' - v' \rfloor) - v'$$

参加进来，

$$s(\alpha, n, v) = -nv + \lceil n\alpha \rceil b - \lceil n\alpha \rceil v' - s(\alpha', \lceil n\alpha \rceil, v') - S + [0 \text{ 或 } 1].$$

再用 $b - v' = v\alpha^{-1}$ ，我们看到，如果用 $n\alpha(b - v') = nv$ 替换 $\lceil n\alpha \rceil(b - v')$ ，一切将极好地简化。

$$s(\alpha, n, v) = -s(\alpha', \lceil n\alpha \rceil, v') - S + \epsilon + [0 \text{ 或 } 1].$$

这里 ϵ 是至多 $v\alpha^{-1}$ 的一个正误差。习题 18 证明 S 同样在 0 和 α^{-1} 之间。我们也能从和中除去 $j = \lceil n\alpha \rceil - 1 = \lfloor n\alpha \rfloor$ 的项，因为它提供 v' 或 $v' - 1$ 。因此，如果我们在所有 v 上取最大绝对值，将得到

$$D(\alpha, n) \leq D(\alpha', \lfloor \alpha n \rfloor) + \alpha^{-1} + 2. \quad (3.31)$$

在后续几章中我们将学到一些方法，当 n 充分大时，这些方法使我们能从此递归式推出 $D(\alpha, n)$ 总比 n 小得多。因此定理(3.28)不仅是真实的，它还能被增强：‘收敛到极限十分快’。

这完全是和，下整和上整计算的一个习题。不习惯于“证明误差是小”的读者可发现，很难相信当面对这样看来古怪的和时，谁都有勇气去继续做下去。但是实际上，第二次查看表明，贯穿整个计算有一个简单的目的，明确的线索。主要思想是能把 n 项的某个和 $s(\alpha, n, v)$ 化成至多 αn 项的相似和。除了从接近界限的项中留下小的剩余外，别的都消去。

现在让我们深吸一口气，且另给一个和，此和不平凡，但 (与我们刚才给的和相比) 有很大优点，它产生闭形式以致我们易检验答案。现在我们的目标是通过找

$$\sum_{0 \leq k < m} \left\lfloor \frac{nk + x}{m} \right\rfloor, \text{ 整数 } m > 0, \text{ 整数 } n,$$

的一个表达式来推广式(3.26)中的和。找此和的闭形式比我们至今所做的更难对付（也许除了我们刚见的偏差问题）。但是它是有启发的，所以在本章余下部分我们将继续讨论。

通常，特别对于难对付的问题，我们从讨论小的情形开始。特殊情形 $n=1$ 是式(3.26)，用 x/m 替换 x ，

$$\left\lfloor \frac{x}{m} \right\rfloor + \left\lfloor \frac{1+x}{m} \right\rfloor + \cdots + \left\lfloor \frac{m-1+x}{m} \right\rfloor = \lfloor x \rfloor.$$

且如同第一章中那样，我们发现通过向下推广到情形 $n=0$ 取得更多数据是有用的：

$$\left\lfloor \frac{x}{m} \right\rfloor + \left\lfloor \frac{x}{m} \right\rfloor + \cdots + \left\lfloor \frac{x}{m} \right\rfloor = m \left\lfloor \frac{x}{m} \right\rfloor.$$

我们的问题有二个参数， m 和 n ，让我们看一些小的 m 的情形。当 $m=1$ 时，和中刚好有一项，且它的值为 $\lfloor x \rfloor$ 。当 $m=2$ 时，和为 $\lfloor x/2 \rfloor + \lfloor (x+n)/2 \rfloor$ 。我们能通过把 n 从下整函数中移出来消除 x 和 n 之间的相互影响，但是为了做这一点一定要分开考虑偶数和奇数 n 。若 n 是偶数，则 $n/2$ 是整数，所以我们能从下整移去它，

$$\left\lfloor \frac{n}{2} \right\rfloor + \left(\left\lfloor \frac{x}{2} \right\rfloor + \frac{n}{2} \right) = 2 \left\lfloor \frac{x}{2} \right\rfloor + \frac{n}{2}.$$

若 n 是奇数，则 $(n-1)/2$ 是整数，所以我们得到

$$\left\lfloor \frac{x}{2} \right\rfloor + \left(\left\lfloor \frac{x+1}{2} \right\rfloor + \frac{n-1}{2} \right) = \lfloor x \rfloor + \frac{n-1}{2}.$$

最后一步是根据具有 $m=2$ 的式(3.26)得出的。

对于偶数和奇数 n 的这些公式有点像 $n=0$ 和 1 的公式，但是还未形成明确的型式，所以我们最好还是继续探讨一些小的情形。对于 $m=3$ ，和是

$$\left\lfloor \frac{x}{3} \right\rfloor + \left\lfloor \frac{x+n}{3} \right\rfloor + \left\lfloor \frac{x+2n}{3} \right\rfloor,$$

而且我们考虑 n 的三个情形：它是 3 的倍数，或者它是比一个倍数多 1，或者它是比一个倍数多 2。也就是说， $n \bmod 3 = 0, 1$ ，或 2。若 $n \bmod 3 = 0$ ，则 $n/3$ 和 $2n/3$ 是整数，所以和是

$$\left\lfloor \frac{x}{3} \right\rfloor + \left(\left\lfloor \frac{x}{3} \right\rfloor + \frac{n}{3} \right) + \left(\left\lfloor \frac{x}{3} \right\rfloor + \frac{2n}{3} \right) = 3 \left\lfloor \frac{x}{3} \right\rfloor + n.$$

若 $n \bmod 3 = 1$ ，则 $(n-1)/3$ 和 $(2n-2)/3$ 是整数，所以我们有

$$\left\lfloor \frac{x}{3} \right\rfloor + \left(\left\lfloor \frac{x+1}{3} \right\rfloor + \frac{n-1}{3} \right) + \left(\left\lfloor \frac{x+2}{3} \right\rfloor + \frac{2n-2}{3} \right) = \lfloor x \rfloor + n - 1.$$

这里的最后一步也得自式(3.26)，此时 $m=3$ 。最后如果 $n \bmod 3 = 2$ ，则

$$\left\lfloor \frac{x}{3} \right\rfloor + \left(\left\lfloor \frac{x+2}{3} \right\rfloor + \frac{n-2}{3} \right) + \left(\left\lfloor \frac{x+1}{3} \right\rfloor + \frac{2n-1}{3} \right) = \lfloor x \rfloor + n - 1.$$

我们已完成情形 $m=3$, 但仍不能识别出型式, 所以我们处理 $m=4$:

$$\left\lfloor \frac{x}{4} \right\rfloor + \left\lfloor \frac{x+n}{4} \right\rfloor + \left\lfloor \frac{x+2n}{4} \right\rfloor + \left\lfloor \frac{x+3n}{4} \right\rfloor.$$

现在可以考虑基于 $n \bmod m$ 的情形. 若 $n \bmod 4 = 0$, 则

$$\left\lfloor \frac{x}{4} \right\rfloor + \left(\left\lfloor \frac{x}{4} \right\rfloor + \frac{n}{4} \right) + \left(\left\lfloor \frac{x}{4} \right\rfloor + \frac{2n}{4} \right) + \left(\left\lfloor \frac{x}{4} \right\rfloor + \frac{3n}{4} \right) = 4 \left\lfloor \frac{x}{4} \right\rfloor + \frac{3n}{2}.$$

若 $n \bmod 4 = 1$, 则

$$\begin{aligned} & \left\lfloor \frac{x}{4} \right\rfloor + \left(\left\lfloor \frac{x+1}{4} \right\rfloor + \frac{n-1}{4} \right) + \left(\left\lfloor \frac{x+2}{4} \right\rfloor + \frac{2n-2}{4} \right) + \left(\left\lfloor \frac{x+3}{4} \right\rfloor + \frac{3n-3}{4} \right) \\ &= \lfloor x \rfloor + \frac{3n}{2} - \frac{3}{2}. \end{aligned}$$

情形 $n \bmod 4 = 3$ 的结果给出相同解答. 最后, 从 $n \bmod 4 = 2$ 的情形我们获得的结果有点不同, 而这就产生了一般情况的一个重要线索:

$$\begin{aligned} & \left\lfloor \frac{x}{4} \right\rfloor + \left(\left\lfloor \frac{x+2}{4} \right\rfloor + \frac{n-2}{4} \right) + \left(\left\lfloor \frac{x}{4} \right\rfloor + \frac{2n}{4} \right) + \left(\left\lfloor \frac{x+2}{4} \right\rfloor + \frac{3n-2}{4} \right) \\ &= 2 \left(\left\lfloor \frac{x}{4} \right\rfloor + \left\lfloor \frac{x+2}{4} \right\rfloor \right) + \frac{3n}{2} - 1 = 2 \left\lfloor \frac{x}{2} \right\rfloor + \frac{3n}{2} - 1. \end{aligned}$$

此最后一步简化了类似于 $\lfloor y/2 \rfloor + \lfloor (y+1)/2 \rfloor$ 的形式, 它也是式(3.26)的一种特殊情形.

对于小 m 的值的和, 总结如下:

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	$4 \left\lfloor \frac{x}{4} \right\rfloor + \frac{3n}{2}$	$\lfloor x \rfloor + \frac{3n}{2} - \frac{3}{2}$	$2 \left\lfloor \frac{x}{2} \right\rfloor + \frac{3n}{2} - 1$	$\lfloor x \rfloor + \frac{3n}{2} - \frac{3}{2}$

看来好像我们取了类似于

$$a \left\lfloor \frac{x}{a} \right\rfloor + bn + c$$

的形式, 其中 a , b 和 c 以某种方式依赖于 m 和 n . 甚至缺乏辨别的人也能看出 b 也许是 $(m-1)/2$. a 的一个表达式较难看出; 但是情形 $n \bmod 4 = 2$ 给出一个暗示, a 也许是

$\gcd(m, n)$, 即 m 和 n 的最大公因子。这是有意义的, 因为当把分数 n/m 化为最小的项时, $\gcd(m, n)$ 是从 m 和 n 移去的因子, 而我们的和含有分数 n/m 。(在第四章中我们将仔细地看 \gcd 运算。) c 的值看来较难理解, 但是也许对于 a 和 b , 它将退出我们的证明。

在计算小 m 的和中, 我们实际再把和的每一项写成

$$\left\lfloor \frac{x + kn}{m} \right\rfloor = \left\lfloor \frac{x + kn \bmod m}{m} \right\rfloor + \frac{kn}{m} - \frac{kn \bmod m}{m},$$

因为 $(kn - kn \bmod m)/m$ 是一个整数, 它可从下整括号内提出。因此原来的和能展开成下列的表:

$$\begin{array}{rcl} \left\lfloor \frac{x}{m} \right\rfloor & + & \frac{0}{m} - \frac{0 \bmod m}{m} \\ + \left\lfloor \frac{x + n \bmod m}{m} \right\rfloor & + & \frac{n}{m} - \frac{n \bmod m}{m} \\ + \left\lfloor \frac{x + 2n \bmod m}{m} \right\rfloor & + & \frac{2n}{m} - \frac{2n \bmod m}{m} \\ & \vdots & \vdots \\ + \left\lfloor \frac{x + (m-1)n \bmod m}{m} \right\rfloor & + & \frac{(m-1)n}{m} - \frac{(m-1)n \bmod m}{m} \end{array}$$

当我们就 m 的小值试验时, 这样三列分别产生 $a \lfloor x/a \rfloor$, bn 和 c 。

特别, 我们能看出 b 是如何产生的。第二列是一个算术级数, 我们知道它的和是第一项和最后一项的平均值, 乘上项数:

$$\frac{1}{2} \left(0 + \frac{(m-1)n}{m} \right) \cdot m = \frac{(m-1)n}{2},$$

所以证实了 $b = (m-1)/2$ 的推断。

第一列和第三列看来较难对付, 为了决定 a 和 c , 我们一定要较仔细地查看数序列

$$0 \bmod m, n \bmod m, 2n \bmod m, \dots, (m-1)n \bmod m.$$

例如, 假设 $m=12$ 和 $n=5$ 。如果把序列想象为钟的时刻, 数为 0 点钟(我们把 12 点钟看作 0 点钟), 然后为 5 点钟, 10 点钟, 3 点钟(=15 点钟), 8 点钟, 等等。结果我们达到每一个小时恰好一次。

现在假设 $m=12$ 和 $n=8$ 。数为 0 点钟, 8 点钟, 4 点钟(=16 点钟), 但 0, 8 和 4 重复。由于 8 和 12 都是 4 的倍数, 而且因为数是从 0 开始(也是 4 的倍数), 所以决不突破此型式, 它们一定全是 4 的倍数。

在这样两种情形中, 我们有 $\gcd(12, 5)=1$ 和 $\gcd(12, 8)=4$ 。下一章我们将证明一般的规则, 它指出如果 $d=\gcd(m, n)$, 则我们以某种次序获得数 $0, d, 2d, \dots, m-d$, 后面是相同序列的 $d-1$ 个复制品。例如, $m=12$ 和 $n=8$ 时, 型式 0, 8, 4 出现 4 次。

和的第一列现在完全有意义。它包含以某个次序项 $\lfloor x/m \rfloor, \lfloor (x+d)/m \rfloor, \dots, \lfloor (x+m-d)/m \rfloor$ 的 d 个复制品, 所以它的和是

$$\begin{aligned}
 & d \left(\left\lfloor \frac{x}{m} \right\rfloor + \left\lfloor \frac{x+d}{m} \right\rfloor + \cdots + \left\lfloor \frac{x+m-d}{m} \right\rfloor \right) \\
 &= d \left(\left\lfloor \frac{x/d}{m/d} \right\rfloor + \left\lfloor \frac{x/d+1}{m/d} \right\rfloor + \cdots + \left\lfloor \frac{x/d+m/d-1}{m/d} \right\rfloor \right) \\
 &= d \left\lfloor \frac{x}{d} \right\rfloor.
 \end{aligned}$$

最后一步也是式(3.26)的另一应用，证实了 a 的推断：

$$a = d = \gcd(m, n).$$

正如我们推断的那样，现在也能计算 c ，因为第三列变得容易探索。它包含算术级数 $0/m, d/m, 2d/m, \dots, (m-d)/m$ 的 d 个复制品，所以它的和是

$$d \left(\frac{1}{2} \left(0 + \frac{m-d}{m} \right) \cdot \frac{m}{d} \right) = \frac{m-d}{2};$$

第三列实际被减，不是被加，所以我们得到

$$c = \frac{d}{2} \frac{m}{d}.$$

难以理解的问题的结尾，完成了探索。希望的闭形式为

$$\sum_{0 \leq k < m} \left\lfloor \frac{nk+x}{m} \right\rfloor = d \left\lfloor \frac{x}{d} \right\rfloor + \frac{m-1}{2} n + \frac{d-m}{2},$$

其中 $d = \gcd(m, n)$ 。作为一种检查，我们能查明，这证实了我们以前知道的 $n=0$ 和 $n=1$ 的特殊情形：当 $n=0$ ，取得 $d = \gcd(m, 0) = m$ ，公式的最后二项为零，所以公式适当地给出 $m \lfloor x/m \rfloor$ 。对于 $n=1$ ，取得 $d = \gcd(m, 1) = 1$ ，最后两项恰好消去，和恰好为 $\lfloor x \rfloor$ 。

对闭形式稍作处理，实际能使它变成按 m 和 n 对称：

$$\begin{aligned}
 \sum_{0 \leq k < m} \left\lfloor \frac{nk+x}{m} \right\rfloor &= d \left\lfloor \frac{x}{d} \right\rfloor + \frac{m-1}{2} n + \frac{d-m}{2} \\
 &= d \left\lfloor \frac{x}{d} \right\rfloor + \frac{(m-1)(n-1)}{2} + \frac{m-1}{2} + \frac{d-m}{2} \\
 &= d \left\lfloor \frac{x}{d} \right\rfloor + \frac{(m-1)(n-1)}{2} + \frac{d-1}{2}. \tag{3.32}
 \end{aligned}$$

这是令人惊讶的，因为没有理由觉得这样的一个和该是对称的。我们证明了一个“互反律”，

$$\sum_{0 \leq k < m} \left\lfloor \frac{nk+x}{m} \right\rfloor = \sum_{0 \leq k < n} \left\lfloor \frac{mk+x}{n} \right\rfloor \quad \text{整数 } m, n > 0.$$

例如，若 $m=41$ 和 $n=127$ ，左边的和有 41 项，右边的和有 127 项；但是对于所有实数

x , 它们的结果仍然相等.

习 题

准备部分

1. 当我们在第一章中分析 Josephus 问题时, 我们把任意正整数 n 表为形式 $n = 2^m + l$, 其中 $0 \leq l < 2^m$. 用下整和 / 或上整括号给出如同 n 的函数那样的 l 和 m 的显式.
2. 对于给定的一个实数 x , 最近整数的公式是什么? 如果有联系, 当 x 恰好是两个整数之间一半的距离时, (a) 给出由上聚拢到 $\lceil x \rceil$ 的表达式; (b) 给出由下聚拢到 $\lfloor x \rfloor$ 的表达式.
3. 当 m 和 n 是正整数, α 是大于 n 的无理数时, 计算 $\lfloor \lfloor m\alpha \rfloor n / \alpha \rfloor$.
4. 课文中描述了层次 1 到层次 5 的问题. 层次 0 的问题是什么? (顺便说一句, 这不是一个层次 0 的问题.)
5. 找当 n 是正整数时, $\lfloor nx \rfloor = n \lfloor x \rfloor$ 的充分必要条件. (你的条件该含有 $\{x\}$.)
6. 当 $f(x)$ 为连续的, 单调下降函数, 它仅当 x 是整数时取整数值, 关于 $\lfloor f(x) \rfloor$ 能表明什么有趣的结果?
7. 解递归

$$X_n = n, \quad 0 \leq n < m;$$

$$X_n = X_{n-m} + 1, \quad n \geq m.$$

8. 证明 Dirichlet 匣子原理: 若把 n 个物体放入 m 个匣子, 则某个匣子一定包含 $\geq \lceil n/m \rceil$ 个物体, 且某个匣子一定包含 $\leq \lfloor n/m \rfloor$ 个物体.
9. 在公元前 1800 年, 埃及数学家把 0 和 1 之间的无理数表为一个分数 $1/x_1 + \dots + 1/x_k$ 的和, 其中 x 是不相同的正整数. 例如, 记 $1/3 + 1/15$ 代替 $2/5$. 证明总可能以一种有规则的方式做到以下这一点: 若 $0 < m/n < 1$, 则

$$\frac{m}{n} = \frac{1}{q} + \left\{ \frac{m}{n} - \frac{1}{q} \right\} \text{ 的表达式}, \quad q = \left\lceil \frac{n}{m} \right\rceil.$$

(这是公元 1202 年 Leonardo Fibonacci 提出的 Fibonacci 算法.)

基本部分

10. 证明表达式

$$\left\lceil \frac{2x+1}{2} \right\rceil - \left\lceil \frac{2x+1}{4} \right\rceil + \left\lfloor \frac{2x+1}{4} \right\rfloor$$

总为 $\lfloor x \rfloor$ 或 $\lceil x \rceil$. 在什么情况产生每种情形?

11. 给出课文中提到的当 $\alpha < \beta$ 时, 开区间 (α, β) 恰好包含 $\lceil \beta \rceil - \lfloor \alpha \rfloor - 1$ 个整数的详细证明. 为了使证明正确, 为什么一定要把 $\alpha = \beta$ 情形排除在外?

12. 证明对于所有整数 n 和所有正整数 m ,

$$\left\lceil \frac{n}{m} \right\rceil = \left\lfloor \frac{n+m-1}{m} \right\rfloor.$$

此等式给出了把上整转化为下整的另一种方法, 且反之亦对, 而不是用反射律(3.4).)

13. 设 α 和 β 是正实数. 证明 $\text{Spec}(\alpha)$ 和 $\text{Spec}(\beta)$ 划分正整数当且仅当 α 和 β 是无理数且 $1/\alpha + 1/\beta = 1$.

14. 证明或推翻

$$(x \bmod ny) \bmod y = x \bmod y, \text{ 整数 } n.$$

15. 是否有类似于式(3.26)的等式, 它用上整而不用下整.

16. 证明 $n \bmod 2 = (1 - (-1)^n)/2$. 找出或证明 $n \bmod 3$ 的形式 $a + b\omega^n + c\omega^{2n}$ 的相似表达式, 其中 ω 是复数 $(-1 + i\sqrt{3})/2$. 提示: $\omega^3 = 1$ 和 $1 + \omega + \omega^2 = 0$.

17. 通过用 $\sum_{j=1}^{\lfloor x+k/m \rfloor} [1 \leq j \leq x+k/m]$ 代替 $\lfloor x+k/m \rfloor$, 且先对 k 求和, 计算 $x \geq 0$ 的和 $\sum_{0 \leq k < m} \lfloor x+k/m \rfloor$. 你的解答是否与式(3.26)相同?

18. 证明式(3.30)中边界值误差项 S 最多为 $\alpha^{-1}v$. 提示: 指出不包含 j 的小值.

课外作业的习题

19. 找出实数 $b > 1$ 使得对于所有实数 $x \geq 1$,

$$\lfloor \log_b x \rfloor = \lfloor \log_b \lfloor x \rfloor \rfloor$$

的一个充分必要条件.

20. 求当 $x > 0$ 时, 闭区间 $[x, \beta]$ 中 x 的所有倍数之和.

21. 对于 $0 \leq m \leq M$, 十进表示中有多少个数 2^m 第一位数字为 1?

22. 计算和 $S_n = \sum_{k \geq 1} \lfloor n/2^k + 1/2 \rfloor$ 和 $T_n = \sum_{k \geq 1} 2^k \lfloor n/2^k + 1/2 \rfloor^2$.

23. 证明序列

$$1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, \dots$$

的第 n 个元素是 $\lfloor \sqrt{2n} + 1/2 \rfloor$. (序列恰好包含 m 个 m .)

24. 习题 13 建立了两个多重集 $\text{Spec}(\alpha)$ 和 $\text{Spec}(\alpha/(\alpha-1))$ 之间的有趣的关系, 当 α 是任何无理数 > 1 , $1/\alpha + (\alpha-1)/\alpha = 1$. 找出(或证明)当 α 是任何正实数时, 两个多重集 $\text{Spec}(\alpha)$ 和 $\text{Spec}(\alpha/(\alpha-1))$ 之间的一个有趣的关系.

25. 证明或推翻式(3.16)定义的 Knuth 数满足对于所有非负的 n , $K_n \geq n$.

26. 证明附属的 Josephus 数(3.20)满足

$$\left(\frac{q}{q-1}\right)^n \leq D_n^{(q)} \leq q \left(\frac{q}{q-1}\right)^n, \quad n \geq 0.$$

27. 证明式(3.20)定义的无限多个数 $D_n^{(q)}$ 是偶数, 且无限多个是奇数.

28. 解递归

$$a_0 = 1;$$

$$a_n = a_{n-1} + \lfloor \sqrt{a_{n-1}} \rfloor, \quad n > 0.$$

29. 证明除式(3.31)之外, 我们有

$$D(x, n) \geq D(x', \lfloor xn \rfloor) - \alpha^{-1} - 2.$$

30. 证明递归

$$X_0 = m,$$

$$X_n = X_{n-1}^2 - 2, \quad n > 0$$

有解 $X_n = \lfloor \alpha^{2^n} \rfloor$, 如果 m 是大于 2 的整数, 其中 $\alpha + \alpha^{-1} = m$, 且 $\alpha > 1$. 例如, 若 $m = 3$, 解为

$$X_n = \lfloor \varphi^{2^{n+1}} \rfloor, \quad \varphi = \frac{1 + \sqrt{5}}{2}, \quad \alpha = \varphi^2.$$

31. 证明或推翻:

$$\lfloor x \rfloor + \lfloor y \rfloor + \lfloor x + y \rfloor \leq \lfloor 2x \rfloor + \lfloor 2y \rfloor.$$

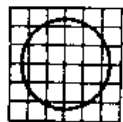
32. 设 $\|x\| = \min(x - \lfloor x \rfloor, \lceil x \rceil - x)$ 表示从 x 到最近整数的距离.

$$\sum_k 2^k \|x / 2^k\|^2 \text{ 的值是什么?}$$

(注意这个和能双向无限. 例如, 当 $x = 1/3$ 时, $k \rightarrow -\infty$ 时项为非零, $k \rightarrow +\infty$ 时项也为非零.)

考查性问题

33. 直径为 $2n-1$ 个单位长, 对称地画在 $2n \times 2n$ 的棋盘上的一个圆, 这里说明 $n=3$ 的情形:



(a) 棋盘的多少个单元包含弓形?

(b) 求一个函数 $f(n, k)$, 使得恰好棋盘的 $\sum_{k=1}^{n-1} f(n, k)$ 个单元完全在圆内.

$$34. \text{ 设 } f(n) = \sum_{k=1}^n \lceil \lg k \rceil.$$

(a) 当 $n \geq 1$ 时, 求 $f(n)$ 的一个闭形式.

(b) 证明对于所有 $n \geq 1$,

$$f(n) = n - 1 + f(\lceil n/2 \rceil) + f(\lfloor n/2 \rfloor).$$

35. 简化公式 $\lfloor (n+1)^2 n! / e \rfloor \bmod n$.

36. 假设 n 是非负整数, 求和

$$\sum_{1 \leq k < 2^n} \frac{1}{2 \lfloor \lg k \rfloor 4 \lfloor \lg \lg k \rfloor}$$

的一个闭形式.

37. 证明对于所有正整数 m 和 n , 等式

$$\sum_{0 \leq k < m} \left(\left\lfloor \frac{m+k}{n} \right\rfloor - \left\lfloor \frac{k}{n} \right\rfloor \right) = \left\lfloor \frac{m^2}{n} \right\rfloor - \left\lfloor \frac{\min(m \bmod n, (-m) \bmod n)^2}{n} \right\rfloor$$

成立.

38. 设 x_1, \dots, x_n 是实数, 并使得对于所有正整数 m 等式

$$\sum_{k=1}^n \lfloor mx_k \rfloor = \left\lfloor m \sum_{1 \leq k \leq n} x_k \right\rfloor$$

成立. 证明关于 x_1, \dots, x_n 的某个有趣的结果.

39. 证明对于每一个实数 $x \geq 1$ 和每一个整数 $b > 1$, 双重和 $\sum_{0 \leq k \leq \log_b x} \sum_{0 \leq j < b^k} \lfloor (x + jb^k) / b^{k+1} \rfloor = (b-1)(\lfloor \log_b x \rfloor + 1) + \lfloor x \rfloor - 1$.

40. 下面的图形中表明, 螺线函数 $\sigma(n)$ 把非负整数 n 映射到有序的整数对 $(x(n), y(n))$. 例如, 它把 $n=9$ 映射到有序对 $(1, 2)$.

(a) 证明如果 $m = \lfloor \sqrt{n} \rfloor$,

$$x(n) = (-1)^m \left((n - m(m+1)) \cdot \lfloor 2\sqrt{n} \rfloor \text{ 是偶数} \rfloor + \left\lfloor \frac{1}{2}m \right\rfloor \right)$$

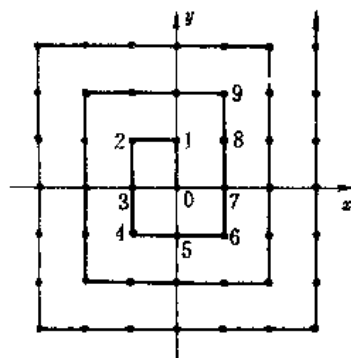
且求 $y(n)$ 的相似公式. 提示: 螺线按照 $\lfloor 2\sqrt{n} \rfloor = 4k-2, 4k-1, 4k, 4k+1$, 归成 W_k, S_k, E_k, N_k 四部分.

(b) 相反地证明由形式

$$n = (2k)^2 \pm (2k + x(n) + y(n)), \quad k = \max(|x(n)|, |y(n)|)$$

的公式, 我们能从 $\sigma(n)$ 来确定 n .

给定符号为 + 和符号为 - 的一个规则.



提示: $\lfloor \sqrt{2n(n+1)} \rfloor = \left\lfloor \sqrt{2\left(n + \frac{1}{2}\right)} \right\rfloor$.

47. 如果函数 $f(x)$ 对于每一个正整数 m 满足

$$f(mx) = f(x) + f\left(x + \frac{1}{m}\right) + \cdots + f\left(x + \frac{m-1}{m}\right),$$

则称它为仿作的函数。找出使下列函数为仿作的关于实数 c 的充分必要条件:

(a) $f(x) = x + c$.

(b) $f(x) = [x + c \text{ 是一个整数}]$.

(c) $f(x) = \max(\lfloor x \rfloor, c)$.

(d) $f(x) = x + c\lfloor x \rfloor - \frac{1}{2}[x \text{ 不是整数}]$.

48. 找出实数 $0 \leq \alpha < 1$ 和 $\beta \geq 0$ 的一个充分必要条件, 使得能根据值的无限多重集

$$\{\lfloor nx \rfloor + \lfloor n\beta \rfloor | n > 0\}$$

来确定 α 和 β .

研究性问题

49. 找出非负实数 α 和 β 的一个充分必要条件, 使得能根据值的无限多重集

$$\{\lfloor nx \rfloor \beta | n > 0\}$$

来确定 α 和 β .

50. 设 x 是实数 $\geq \varphi = (1/\sqrt{2})(1 + \sqrt{5})$. 递归

$$Z_0(x) = x,$$

$$Z_n(x) = Z_{n-1}(x)^2 - 1, \quad n > 0,$$

的解可记为 $Z_n(x) = \lceil f(x)^{2^n} \rceil$, 若 x 为整数, 其中

$$f(x) = \lim_{n \rightarrow \infty} Z_n(x)^{1/2^n},$$

因为 $Z_n(x) - 1 < f(x)^{2^n} < Z_n(x)$. 这个函数 $f(x)$ 有什么有趣的性质?

51. 给定非负实数 α 和 β , 设

$$\text{Spec}(\alpha; \beta) = \{\lfloor \alpha + \beta \rfloor, \lfloor 2\alpha + \beta \rfloor, \lfloor 3\alpha + \beta \rfloor, \cdots\}$$

是推广 $\text{Spec}(\alpha) = \text{Spec}(\alpha; 0)$ 的一个多重集. 证明或推翻: 若 $m \geq 3$, 则多重集 $\text{Spec}(\alpha_1; \beta_1), \text{Spec}(\alpha_2; \beta_2), \cdots, \text{Spec}(\alpha_m; \beta_m)$ 划分正整数, 且若参数 $\alpha_1 < \alpha_2 < \cdots < \alpha_m$ 是有理的, 则

$$\alpha_k = \frac{2^m - 1}{2^{k-1}}, \quad 1 \leq k \leq m.$$

52. Fibonacci 算法(习题 9)是下列意义的“贪婪”算法, 即它在每一步处选取可能最小的 q . 知道一个较麻烦的算法, 把 n 为奇数的每一个分数 m/n 表为奇数分母的不同单位分数的和 $1/q_1 + \cdots + 1/q_k$. 这样一种表示的贪婪算法总是终止的吗?

第四章 数 论

离散数学是本书的重点，而整数对离散数学是极为重要的。数论是讨论整数性质的数学重要分支，所以我们要研究它。

在前面一章，通过介绍二元运算‘mod’和‘gcd’，我们初步接触了数论的内容。现在让我们真正深入到这个课题中去。

4.1 可除性

如果 $m > 0$ ，且 n/m 是一个整数，我们说 m 除尽 n (或 m 可除尽 n)。此性质是数论的基础，用一个特别记法来记它是方便的，所以我们记

$$m \backslash n \Leftrightarrow m > 0 \text{ 且对于某个整数 } k, n = mk. \quad (4.1)$$

(在当前的数学文献中，记法 ‘ $m|n$ ’ 实际比 ‘ $m \backslash n$ ’ 更普遍。但是竖线使用太多，如绝对值，集合定义符，条件概率等，——而使用倒的切割。而且，‘ $m \backslash n$ ’ 给出了 m 是一个暗指的比的分母的印象。所以我们大胆地让可除性符号向左倾斜。)

如果 m 除不尽 n ，我们记 ‘ $m \nmid n$ ’。

有一个相似的关系“ n 是 m 的一个倍数”，除了 m 不必为正外，它几乎意味着相同的意思。此时我们简明地表示对某个整数 k ， $n = mk$ 的意思。于是，例如，仅有 0 的一个倍数 (即 0)，但是没有数被 0 除尽，每个整数是 -1 的一个倍数，但是 (严格地说) 没有整数可被 -1 除尽。当 m 和 n 是任何实数时，这些定义可以应用，例如， π 可除尽 2π 。但是当 m 和 n 是整数时，我们几乎总将用到它们。这毕竟是数论。

两个整数 m 和 n 的最大公因子是除尽它们两者的最大整数：

$$\gcd(m, n) = \max \{k | k \backslash m \text{ 和 } k \backslash n\}. \quad (4.2)$$

例如， $\gcd(12, 18) = 6$ 。这是一个熟悉的概念，因为这是四年级学生学习把分数 m/n 化到最小项取出的公因子： $12/18 = (12/6)(18/6) = 2/3$ 。注意，如果 $n > 0$ ，我们得到 $\gcd(0, n) = n$ ，因为任何正数除尽 0，因为 n 是它自己的最大因子。 $\gcd(0, 0)$ 的值是没有定义的。

另一个熟悉的概念是最小公倍数，

$$\text{lcm}(m, n) = \min\{k | k > 0, m \setminus k \text{ 和 } n \setminus k\}, \quad (4.3)$$

如果 $m \leq 0$ 或 $n \leq 0$, 则无定义。学算术的学生知道这就像最小公分母, 当具有分母 m 和 n 的分式相加时用到它。例如, $\text{lcm}(12, 18) = 36$, 四年级的学生知道 $(7/12) + (1/18) = (21/36) + (2/36) = (23/36)$ 。lcm 有点相似于 gcd, 但是因为 gcd 有较好性质, 所以我们给予更多的时间来讨论它。

gcd 的一个最好性质是容易用 2300 年前的欧几里德算法来计算它。对于给定值 $0 \leq m < n$, 为了计算 $\text{gcd}(m, n)$, 欧几里德算法用了递归。

$$\begin{aligned} \text{gcd}(0, n) &= n; \\ \text{gcd}(m, n) &= \text{gcd}(n \bmod m, m), \text{ 对于 } m > 0. \end{aligned} \quad (4.4)$$

于是, 例如, $\text{gcd}(12, 18) = \text{gcd}(6, 12) = \text{gcd}(0, 6) = 6$ 。所述的递归是成立的, 因为 m 和 n 的任何公因子一定也是 m 和 $n \bmod m$ 的一个公因子, 它是 $n - \lfloor n/m \rfloor m$ 。看来根本没有任何 $\text{lcm}(m, n)$ 的递归像这个递归那样简单(见习题 2)。

欧几里德算法还给出如下结果: 我们能把它推广以致它将计算出整数 m' 和 n' 满足

$$m'm + n'n = \text{gcd}(m, n). \quad (4.5)$$

这是由于: 若 $m = 0$, 我们简单地取 $m' = 0$ 和 $n' = 1$ 。否则设 $r = n \bmod m$, 且以 r 和 m 代替 m 和 n 递归地应用方法, 计算 \bar{r} 和 \bar{m} 使得

$$\bar{r}r + \bar{m}m = \text{gcd}(r, m).$$

由于 $r = n - \lfloor n/m \rfloor m$ 和 $\text{gcd}(r, m) = \text{gcd}(m, n)$, 这个方程告诉我们

$$\bar{r}(n - \lfloor n/m \rfloor m) + \bar{m}m = \text{gcd}(m, n).$$

可把右边改写成它依赖于 m 和 n :

$$(\bar{m} - \lfloor n/m \rfloor \bar{r})m + \bar{r}n = \text{gcd}(m, n),$$

因此, $m' = \bar{m} - \lfloor n/m \rfloor \bar{r}$ 和 $n' = \bar{r}$ 是我们在式(4.5)中需要的整数。例如, 在我们的适用的情况 $m = 12, n = 18$ 中, 此法给出 $6 = 0 \cdot 0 + 1 \cdot 6 = 1 \cdot 6 + 0 \cdot 12 = (-1) \cdot 12 + 1 \cdot 18$ 。

但是为什么式(4.5)是这样一个精巧的结果? 主要理由是数 m' 和 n' 实际证明了欧几里德算法产生任何特殊情形中的准确解答。让我们假设在冗长的计算之后计算机告诉我们 $\text{gcd}(m, n) = d$ 和 $m'm + n'n = d$, 但是我们怀疑, 且想像实际有一个较大公因子, 机器不知为什么忽略了它。然而这是不会的, 因为任何 m 和 n 的公因子必须除尽 $m'm + n'n$, 所以它必须除尽 d , 它必须是 $\leq d$ 。此外, 我们能容易地检验出 d 的确除尽 m 和 n 两者。(输出它们自身证明正确性的算法被称为自证明算法。)

在本章余下的部分我们将经常用到式(4.5)。它的重要结论之一是下列小的定理:

$$k \setminus m \text{ 和 } k \setminus n \Leftrightarrow k \setminus \text{gcd}(m, n). \quad (4.6)$$

(证明: 如果 k 除尽 m 和 n 两者, 它除尽 $m'm + n'n$, 所以它除尽 $\text{gcd}(m, n)$ 。反之, 如果 k 除尽 $\text{gcd}(m, n)$, 它除尽 m 的一个因子和 n 的一个因子, 所以它除尽 m 和 n 两者。)我们总知道 m 和 n 的任何公因子一定小于或等于它们的最大公因子, 这就是最大公因子的

定义。但是我们事实上知道任何公因子是它们的最大公因子的一个因子。

有时我们需要处理在 n 的所有因子上的和，此时用方便的规则

$$\sum_{m|n} a_m = \sum_{m|n} a_{n/m}, \quad \text{整数 } n > 0, \quad (4.7)$$

常常是有用的，由于就 m 来算时 n/m 通过 n 的所有因子，所以它成立。例如，当 $n=12$ 时，它就是 $a_1 + a_2 + a_3 + a_4 + a_6 + a_{12} = a_{12} + a_6 + a_4 + a_3 + a_2 + a_1$ 。

还有一个稍微更一般的等式，

$$\sum_{m|n} a_m = \sum_k \sum_{m>0} a_m [n = mk], \quad (4.8)$$

这是定义式(4.1)的直接结果，如果 n 是正的，式(4.8)的右边为 $\sum_{k|n} a_{n/k}$ ，因此式(4.8)意味着式(4.7)。当 n 是负的时，方程式(4.8)也行得通。(在这种情形，当 k 是 n 的一个因子的负值时，右边的非零项出现。)

并且，由定律

$$\sum_{m|n} \sum_{k|m} a_{k, m} = \sum_{k|n} \sum_{l|(n/k)} a_{k, kl}, \quad (4.9)$$

能“交换”因子上的双重和。例如，当 $n=12$ 时，此定律取下列形式：

$$\begin{aligned} & a_{1, 1} + (a_{1, 2} + a_{2, 2}) + (a_{1, 3} + a_{3, 3}) + (a_{1, 4} + a_{2, 4} + a_{4, 4}) + (a_{1, 6} + a_{2, 6} + a_{3, 6} \\ & \quad + a_{6, 6}) + (a_{1, 12} + a_{2, 12} + a_{3, 12} + a_{4, 12} + a_{6, 12} + a_{12, 12}) \\ &= (a_{1, 1} + a_{1, 2} + a_{1, 3} + a_{1, 4} + a_{1, 6} + a_{1, 12}) + (a_{2, 2} + a_{2, 4} + a_{2, 6} + a_{2, 12}) \\ & \quad + (a_{3, 3} + a_{3, 6} + a_{3, 12}) + (a_{4, 4} + a_{4, 12}) + (a_{6, 6} + a_{6, 12}) + a_{12, 12}. \end{aligned}$$

我们能用 Iverson 的操作证明式(4.9)，左边是

$$\sum_{j, l} \sum_{k, m>0} a_{n, m} [n = jm] [m = kl] = \sum_j \sum_{k, l>0} a_{k, kl} [n = jkl],$$

右边是

$$\sum_{j, m} \sum_{k, l>0} a_{k, kl} [n = jk] [n/k = ml] = \sum_m \sum_{k, l>0} a_{k, kl} [n = mlk],$$

除了重新命名指标外，它们是相同的。此例表明，当我们研究数论时，第二章中所学的技巧是挺合用的。

4.2 素数

如果一个正整数只有两个因子，即 1 和 p ，则称 p 为素数。在本章的余下部分，当不明显说出时，总以字母 p 表示一个素数。我们约定 1 不是素数，所以开始列出的素数序列如下：

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, ...

有些数看来像素数，但它不是，如 $91(=7 \cdot 13)$ 和 $161(=7 \cdot 23)$ 。这些数和其他的数有三个或三个以上的因子，被称为合成数，每个大于 1 的整数或者是素数或者是合成数，但是不会既是素数又是合成数。

素数很重要，因为它们是所有正整数的基本的组成部分，任何正整数 n 能记为素数的一个乘积，

$$n = p_1 \cdots p_m = \prod_{k=1}^m p_k, \quad p_1 \leq \cdots \leq p_m. \quad (4.10)$$

例如， $12=2 \cdot 2 \cdot 3$ ； $11011=7 \cdot 11 \cdot 11 \cdot 13$ ； $11111=41 \cdot 271$ 。(用 Π 表示乘积，相似于用 Σ 表示和，如同习题 2.25 表明的那样，如果 $m=0$ ，我们考虑它为一个空乘积，由定义它的值为 1，这就是 $n=1$ 由式(4.10)所表示的方式。)这样一种因子分解总是可能的，因为如果 $n>1$ 不是素数，它有一个因子 n_1 使得 $1 < n_1 < n$ ；因此我们能记 $n = n_1 \cdot n_2$ ，且(由归纳法)我们知道 n_1 和 n_2 能记为素数的乘积。

并且，式(4.10)中的展开是唯一的：仅有一种方式把 n 记为以非降次序的素数的乘积。此命题称为算术的基本定理，它看来很明显，我们可能奇怪为什么还需证明它。怎么能有二个不同的素数集合具有相同的乘积？不能有，但是理由不能“由素数的定义”简单地说明。例如，如果我们考虑当 m 和 n 是整数时形式 $m + n\sqrt{10}$ 的所有实数的集合，任何两个这样的数的乘积仍具有相同形式，且如果不能以非平凡方法因子分解它，我们可称这样一个数为“素数”。数 6 有两种代表 $2 \cdot 3 = (4+\sqrt{10})(4-\sqrt{10})$ ，然而习题 36 证明在此系统中 2, 3, $4+\sqrt{10}$, $4-\sqrt{10}$ 全为“素数”。

所以要严格证明式(4.10)是唯一的。当 $n=1$ 时的确仅有一个可能性，因为此时乘积一定是空的；所以让我们假设 $n>1$ ，且所有较小数因子唯一，假设有两种因子分解

$$n = p_1 \cdots p_m = q_1 \cdots q_k, \quad p_1 \leq \cdots \leq p_m \text{ 和 } q_1 \leq \cdots \leq q_k,$$

其中 p 和 q 全为素数。我们将证明 $p_1 = q_1$ 。如果不相等，我们能设 $p_1 < q_1$ ，使 p_1 小于所有 q ，由于 p_1 和 q_1 是素数，它们的 gcd 一定是 1；因此欧几里德的自证明算法给出整数 a 和 b 使得 $ap_1 + bq_1 = 1$ ，所以

$$ap_1q_2 \cdots q_k + bq_1q_2 \cdots q_k = q_2 \cdots q_k.$$

由于 $q_1q_2 \cdots q_k = n$ ，现在 p_1 除尽左边的两项，因此 p_1 除尽右边的 $q_2 \cdots q_k$ 。因此 $q_2 \cdots q_k / p_1$ 是一个整数，且 $q_2 \cdots q_k$ 有一个素数因子分解，在此因子分解中 p_1 出现。但是 $q_2 \cdots q_k < n$ ，所以它有一个唯一的因子分解(由归纳法)。此矛盾表明， p_1 终于一定等于 q_1 ，所以 p_1 能除尽两个 n 的因子分解，得到 $p_2 \cdots p_m = q_2 \cdots q_k < n$ 。另一个因子也相等(由归纳法)，由此完成了唯一性的证明。

有时用另一种方式叙述基本定理较有用：每一个正整数能唯一地记为形式

$$n = \prod_p p^{n_p}, \quad \text{其中每个 } n_p \geq 0. \quad (4.11)$$

右边是无限多个素数上的一个乘积；但是对于任何特殊的 n ，除少数指数外全为零，所以对应的因子为 1。所以它实际是一个有限乘积，如同许多“无限”和实际是有限那样，它们

的项几乎全部为零。

公式(4.11)唯一地代表 n , 所以我们能想象序列 $\langle n_2, n_3, n_5, \dots \rangle$ 是正整数的数系。例如, 12 的素数指数代表是 $\langle 2, 1, 0, 0, \dots \rangle$, 18 的素数指数代表是 $\langle 1, 2, 0, 0, \dots \rangle$ 。为了乘两个数, 我们简单地加它们的代表。换句话说,

$$k = mn \Leftrightarrow k_p = m_p + n_p \quad \text{对所有 } p. \quad (4.12)$$

这意味着

$$m \mid n \Leftrightarrow m_p \leq n_p \quad \text{对所有 } p. \quad (4.13)$$

立即得到

$$k = \gcd(m, n) \Leftrightarrow k_p = \min(m_p, n_p) \quad \text{对所有 } p. \quad (4.14)$$

$$k = \text{lcm}(m, n) \Leftrightarrow k_p = \max(m_p, n_p) \quad \text{对所有 } p. \quad (4.15)$$

例如, 由于 $12 = 2^2 \cdot 3^1$ 和 $18 = 2^1 \cdot 3^2$, 通过取公共指数的 \min 和 \max , 我们能得到它们的 \gcd 和 lcm :

$$\gcd(12, 18) = 2^{\min(2, 1)} \cdot 3^{\min(1, 2)} = 2^1 \cdot 3^1 = 6,$$

$$\text{lcm}(12, 18) = 2^{\max(2, 1)} \cdot 3^{\max(1, 2)} = 2^2 \cdot 3^2 = 36.$$

如果素数 p 除尽乘积 mn , 由唯一因子分解定理, 则它除尽 m 或 n , 也许两者。但是合成数没有这个性质。例如, 非素数 4 除尽 $60 = 6 \cdot 10$, 但是它既不除尽 6 也不除尽 10, 原因是简单的: 在因子分解中 $60 = 6 \cdot 10 = (2 \cdot 3)(2 \cdot 5)$, $4 = 2 \cdot 2$ 的两个素数因子被分成两部分, 因此 4 都不除尽这两部分。但是一个素数是不可分的, 所以它一定除尽原来因子之一。

4.3 素数例子

有多少个素数? 很多, 事实上有无限多, 很久以前欧几里德证明了这一点, 如下。假设仅有有限多个素数, 譬如说 k 个, $2, 3, 5, \dots, P_k$ 。于是我们该考虑数

$$M = 2 \cdot 3 \cdot 5 \cdot \dots \cdot P_k + 1,$$

k 个素数中没有一个能除尽 M , 因为每个都除尽 $M-1$ 。因此一定有其他某一个素数除尽 M , 也许 M 本身是素数。这就与我们的假设 $2, 3, \dots, P_k$ 为仅有的素数相矛盾, 所以确实一定有无限多个素数。

欧几里德证明提供了用递归

$$e_n = e_1 e_2 \cdots e_{n-1} + 1, \quad \text{当 } n \geq 1, \quad (4.16)$$

定义欧几里德数, 开始的序列

$$e_1 = 1 + 1 = 2;$$

$$e_2 = 2 + 1 = 3;$$

$$e_3 = 2 \cdot 3 + 1 = 7;$$

$$e_4 = 2 \cdot 3 \cdot 7 + 1 = 43;$$

这些全为素数。但是接着的情形, e_5 是 $1807=13 \cdot 139$, 结果 $e_6=3\,263\,443$ 是素数, 然而

$$e_7 = 547 \cdot 607 \cdot 1\,033 \cdot 31\,051;$$

$$e_8 = 29\,881 \cdot 67\,003 \cdot 9\,119\,521 \cdot 6\,212\,157\,481.$$

知道 e_9, \dots, e_{17} 是合成数, 余下的 e_n 可能同样为合成数。然而, 欧几里德数全部彼此互素, 也就是说,

$$\gcd(e_m, e_n) = 1, \text{ 当 } m \neq n \text{ 时.}$$

欧几里德算法 (还有别的什么?) 以短的三个步骤告诉了我们这一点, 因为当 $n > m$ 时 $e_n \bmod e_m = 1$:

$$\gcd(e_m, e_n) = \gcd(1, e_m) = \gcd(0, 1) = 1.$$

所以, 如果我们让 q_j 是 e_j 的最小因子 (对所有 $j \geq 1$), 则素数 q_1, q_2, q_3, \dots 全是不同的。这是一个无限多素数的序列。

让我们暂停考虑第一章的观点的欧几里德数。我们是否能以闭形式表达 e_n ? 移去三个点能简化递归(4.16): 如果 $n > 1$ 我们有

$$e_n = e_1 \cdots e_{n-2} e_{n-1} + 1 = (e_{n-1} - 1) e_{n-1} + 1 = e_{n-1}^2 - e_{n-1} + 1.$$

因此 e_n 大约有 e_{n-1} 的两倍那样多的十进制数字位。习题 37 证明了有一个常数 $E \approx 1.264$ 使得

$$e_n = \left\lfloor E^{2^n} + \frac{1}{2} \right\rfloor. \quad (4.17)$$

习题 60 提供一个相似的公式, 它只给出素数

$$p_n = \lfloor P^{3^n} \rfloor, \text{ 对于某个常数 } P. \quad (4.18)$$

但是像方程(4.17)和(4.18), 因为以一类不公开的方式从数 e_n 和 p_n 来计算常数 E 和 P , 所以实际不能考虑为闭形式, 还不知道它们和其他值得注意的数学常数相联系的独立关系。

确实没有人知道任何给出任意大素数的有用公式。然而 1984 年在 Chevron Geosciences 的计算机科学工作者发现了可能性极小的石油。他们用了由 David Slowinski 建立的一个程序, 当测试一个新的 Cray X-MP 巨型计算机时, 发现了那时所知的最大素数,

$$2^{216\,091} - 1.$$

在个人计算机上用几 ms 计算这个数是容易的, 因为近代计算机以二进制记法工作, 此数

简单地 $(11\cdots 1)_2$ 。所有它的 216 091 位是‘1’。但是证明这个数是素数是很困难的。事实上，几乎任何关于它的计算需许多时间，因为它是这样大。例如，在个人计算机上甚至一个成熟的算法仅把 $2^{216\,091}-1$ 转换到基数 10 也要花几分钟。当打印时，它的 65 050 个十进制数字位要花第一类邮件的 65 分邮票。

顺便说到，当有 216 091 个圆盘时解 Hanoi 塔问题必要的移动次数为 $2^{216\,091}-1$ 。形式 2^p-1

的数(其中 p 是素数)称为 Mersenne 数，以纪念 17 世纪研究它的一些性质的先辈 Marin Mersenne。至今知道的 Mersenne 素数 $p=2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1\,279, 2\,203, 2\,281, 3\,217, 4\,253, 4\,423, 9\,689, 9\,941, 11\,213, 19\,937, 21\,701, 23\,209, 44\,497, 86\,243, 110\,503, 132\,049$ 和 216 091。

如果 n 是合成数，则数 2^n-1 不可能为素数，因为 $2^{km}-1$ 有 2^m-1 作为一个因子：

$$2^{km}-1=(2^m-1)(2^{m(k-1)}+2^{m(k-2)}+\cdots+1).$$

但是当 p 是素数时， 2^p-1 不总是素数； $2^{11}-1=2\,047=23\cdot 89$ 是最小的这样的非素数。(Mersenne 明白这一点。)

大数的因子分解和素性测试是目前的热门课题。在[174]的 4.5.4 节中概括了到 1981 年的结果，且许多新结果连续被发现。该书的 391 页到 394 页说明了测试 Mersenne 数素性的特殊方法。

就最近两百年的大部分时间而言，虽然仅知 31 个 Mersenne 素数，但知道的最大素数是一个 Mersenne 素数。许多人尝试找大的，但是不易解决。所以那确实有趣，且吉尼斯世界记录大全中的一点可代替试验形式 2^k+1 的数(对于像 3 或 5 那样小的 k 值)。测试这些数的素性几乎能像测试 Mersenne 数的素性一样快；[174]的习题 4.5.4-27 给出详细的讨论。

我们没有完全回答关于有多少个素数的原问题。有无限多个，但是有些无限集比其他一些集合稠密。例如，在正整数中间有无限多个偶数和无限多个完全平方。然而在几种重要意义下，偶数比完全平方多。一种这样的意义是看第 n 个值的大小。第 n 个偶整数是 $2n$ 而第 n 个完全平方是 n^2 ；由于对大的 n ， $2n$ 比 n^2 小得多，第 n 个偶整数比第 n 个完全平方出现早得多，所以我们能说，偶整数比完全平方多，用相似意思来查看不超过 x 的值的个数。有 $\lfloor x/2 \rfloor$ 个这样的偶整数和 $\lfloor \sqrt{x} \rfloor$ 个完全平方；由于对于大的 x ， $x/2$ 比 \sqrt{x} 大得多，我们可再次说，有更多的偶整数。

在这样两种意义下，我们能对素数说些什么？结果第 n 个素数 P_n 大约为 n 乘 n 的自然对数：

$$P_n \sim n \ln n.$$

(符号‘ \sim ’可读为“渐近于”；它意味着当 n 趋于无穷时，比 $P_n/n \ln n$ 的极限是 1.)

类似，对于不超过 x 的素数个数 $\pi(x)$ ，我们有知名的素数定理：

$$\pi(x) \sim \frac{x}{\ln x}.$$

我们能容易证明它们的一个结论包含着另一个结论, 但证明这样两个结论超出了本书范围。在第九章中我们将讨论函数趋于无穷的速度, 且我们将看到(P_n 的近似)函数 $n \ln n$ 渐近地位于 $2n$ 和 n^2 之间。因此素数少于偶整数, 但是素数多于完全平方。

仅当 n 或 $x \rightarrow \infty$ 时取极限成立的这些公式能用更精确的估计来代替。例如, Rosser 和 Schoenfeld^[253]建立了便于使用的界限

$$\ln x - \frac{3}{2} < \frac{x}{\pi(x)} < \ln x - \frac{1}{2}, \quad \text{对 } x \geq 67; \quad (4.19)$$

$$n \left(\ln n + \ln \ln n - \frac{3}{2} \right) < P_n < n \left(\ln n + \ln \ln n - \frac{1}{2} \right), \quad \text{对 } n \geq 20. \quad (4.20)$$

如果我们看一个“随机”整数 n , 则它是素数的机会大约在 $\ln n$ 个中有一个。例如, 如果我们查看接近 10^{16} 的数, 则在找到一个素数之前, 我们必须检验它们的 $16 \ln 10 \approx 36.8$ 个。(结果恰好在 $10^{16} - 370$ 和 $10^{16} - 1$ 之间有 10 个素数。)然而素数的分布有许多不规则性。例如, 包括在 $P_1 P_2 \cdots P_n + 2$ 和 $P_1 P_2 \cdots P_n + P_{n+1} - 1$ 之间的所有数是合成数。已经知道许多“孪生”素数 p 和 $p + 2$ (如 5 和 7, 11 和 13, 17 和 19, 29 和 31, ..., 9 999 999 999 999 641 和 9 999 999 999 999 643, ...), 然而没有人知道是否有无穷多对孪生素数。(见 Hardy 和 Wright^[150], § 1.4 和 § 2.8.)

计算 $\leq x$ 的所有 $\pi(x)$ 个素数的一种简单的方法是形成所谓 Eratosthenes 的筛: 首先写下所有 2 到 x 的整数。接着把 2 圈起来, 记它为素数, 且划掉所有其他 2 的倍数。然后再圈最小的未圈的未划掉的数, 且划掉它的其他的倍数。当圈或划掉所有数时, 圈的数是素数。例如当 $x = 10$ 时, 我们写下 2 到 10, 圈 2, 然后划掉它的倍数 4, 6, 8 和 10。接着 3 是最小的未圈, 未划掉的数, 所以我们圈它, 且划掉 6 和 9。现在 5 是最小的, 所以我们圈它, 且划掉 10。最后我们圈 7。被圈的数是 2, 3, 5 和 7; 所以这就是不超过 10 的 $\pi(10) = 4$ 个素数。

4.4 阶乘因子

现在让我们查看一些很有趣的合成数, 阶乘的因子分解:

$$n! = 1 \cdot 2 \cdot \cdots \cdot n = \prod_{k=1}^n k, \quad \text{整数 } n \geq 0. \quad (4.21)$$

按照一个空乘积的约定, 定义 $0!$ 是 1。于是 $n! = (n-1)!n$ (对每一个正整数 n)。这是 n 个不同元素的排列数。也就是说, 在一行上排列 n 个东西的方式数: 对于第一个东西有 n 种选取; 对于第一个东西的每种选取, 第二个有 $n-1$ 种选取; 对于这 $n(n-1)$ 种选取的每个, 第三个有 $n-2$ 种选取, 等等, 一起给出 $n(n-1)(n-2)\cdots(1)$ 种排列, 这些是阶乘函数的前几个值。

n	0	1	2	3	4	5	6	7	8	9	10
$n!$	1	1	2	6	24	120	720	5 040	40 320	362 880	3 628 800

知道几个阶乘的事实是有用的, 像前 6 个左右的值, 以及 $10!$ 大约为 $3\frac{1}{2}$ 百万加零头的事实; 另一个有趣的事实是当 $n \geq 25$ 时, $n!$ 中的数字位数超过 n .

用有点像第一章的 Gauss 诀窍, 我们能证明 $n!$ 充分大:

$$n!^2 = (1 \cdot 2 \cdot \dots \cdot n)(n \cdot \dots \cdot 2 \cdot 1) = \prod_{k=1}^n k(n+1-k).$$

我们有 $n \leq k(n+1-k) \leq (1/4)(n+1)^2$, 因为二次多项式 $k(n+1-k) = (1/4)(n+1)^2 - (k - 1/2(n+1))^2$ 在 $k=1$ 处有它的最小值, 在 $k=1/2(n+1)$ 处有它的最大值. 所以

$$\prod_{k=1}^n n \leq n!^2 \leq \prod_{k=1}^n \frac{(n+1)^2}{4};$$

也就是说,

$$n^{n/2} \leq n! \leq \frac{(n+1)^n}{2^n}. \quad (4.22)$$

这个关系告诉我们, 阶乘函数以指数律增长!!

对于大的 n , 我们能用 Stirling 公式来更精确地近似 $n!$, 在第九章中我们将推出 Stirling 公式:

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \quad (4.23)$$

且一个更精确的近似告诉我们渐近相对误差: Stirling 公式渐近 $n!$ 差不多差 $1/(12n)$ 的一个因子. 甚至对适当小的 n , 此十分精确的估计也是很好的. 例如, 当 $n=10$ 时, Stirling 近似式(4.23)给出接近 3 598 696 的一个值, 且这是大约 $0.83\% \approx 1/120$, 很小. 对于渐近来说是优良的.

让我们回到素数上来. 对于任何给定的系数 p , 我们希望确定除尽 $n!$ 的 p 的最大幂; 也就是说, 我们要 $n!$ 的唯一的因子分解中 p 的指数. 我们用 $e_p(n!)$ 记这个数, 且从小的情形 $p=2$ 和 $n=10$ 开始研究. 由于 $10!$ 是 10 个数的乘积, 通过那 10 个数提供的 2 的幂的求和能求出 $e_2(10!)$; 这个计算对应于下列数组的列的求和:

	1	2	3	4	5	6	7	8	9	10	2 的 幂
被 2 除尽		×		×		×		×		×	$5 = \lfloor 10/2 \rfloor$
被 4 除尽				×				×			$2 = \lfloor 10/4 \rfloor$
被 8 除尽								×			$1 = \lfloor 10/8 \rfloor$
2 的 幂	0	1	0	2	0	1	0	3	0	1	8

(形成的列和有时称为直尺函数 $\rho(k)$, 因为它们相似于 “|||||”, 线的长度记 1 英寸的一部分.) 这样的 10 个和的总和是 8, 因此 2^8 除尽 $10!$, 但是 2^9 除不尽 $10!$.

还有另一种方法: 我们能计算行的贡献的总数. 第一行记提供 2 的幂的数(因此可被

2 除尽), 它们有 $\lfloor 10/2 \rfloor = 5$ 个。第二行记提供另外的一个 2 的幂的数。它们有 $\lfloor 10/4 \rfloor = 2$ 个。第三行记提供又一个 2 的幂的数, 它们有 $\lfloor 10/8 \rfloor = 1$ 个。这说出了所有成分, 所以我们有 $e_2(10!) = 5 + 2 + 1 = 8$ 。

对于一般的 n , 此法给出

$$e_2(n!) = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n}{4} \right\rfloor + \left\lfloor \frac{n}{8} \right\rfloor + \cdots = \sum_{k \geq 1} \left\lfloor \frac{n}{2^k} \right\rfloor.$$

此和实际为有限的, 因为当 $2^k > n$ 时被加数为零。所有它仅有 $\lfloor \lg n \rfloor$ 个非零项, 且它是完全容易计算的。例如, 当 $n = 100$ 时, 我们有

$$e_2(100!) = 50 + 25 + 12 + 6 + 3 + 1 = 97.$$

每一项就是前一项一半取下整。对所有 n , 这是真的, 因为如同式(3.11)的特殊情形, 我们有 $\lfloor n/2^{k+1} \rfloor = \lfloor \lfloor n/2^k \rfloor / 2 \rfloor$ 。当我们以二进制记数时, 容易明白这里的结果:

$$100 = (1\ 100\ 100)_2 = 100$$

$$\lfloor 100/2 \rfloor = (110\ 010)_2 = 50$$

$$\lfloor 100/4 \rfloor = (11\ 001)_2 = 25$$

$$\lfloor 100/8 \rfloor = (1\ 100)_2 = 12$$

$$\lfloor 100/16 \rfloor = (110)_2 = 6$$

$$\lfloor 100/32 \rfloor = (11)_2 = 3$$

$$\lfloor 100/64 \rfloor = (1)_2 = 1$$

我们仅略去一项中最小有意义的一位就得到下一项。

二进制表示还显示如何推出另一个公式

$$e_2(n!) = n - v_2(n), \quad (4.24)$$

其中 $v_2(n)$ 是 n 的二进制表示中 1 的个数。这个简化行得通, 因为每个 1 提供 2^m 给 n 的值, 提供 $2^{m-1} + 2^{m-2} + \cdots + 2^0 = 2^m - 1$ 给 $e_2(n!)$ 的值。

把结果推广到一个任意素数 p , 我们有

$$e_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots = \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor \quad (4.25)$$

理由如前。

$e_p(n!)$ 有多大呢? 通过从被加数移去下整符号, 然后求无限几何级数的和, 我们得到一个容易的(但是为好的)上界:

$$e_p(n!) < \frac{n}{p} + \frac{n}{p^2} + \frac{n}{p^3} + \cdots$$

$$\begin{aligned}
 &= \frac{n}{p} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots \right) \\
 &= \frac{n}{p} \left(\frac{p}{p-1} \right) = \frac{n}{p-1}.
 \end{aligned}$$

对于 $p=2$ 和 $n=100$, 此不等式说出 $97 < 100$, 因此上界 100 不仅正确, 它也靠近真值 97. 事实上, 真值 $n - v_2(n)$ 一般是 $\sim n$, 因为 $v_2(n) \leq \lceil \lg n \rceil$ 是渐近地比 n 小得多.

当 $p=2$ 和 3 时, 我们的公式给出 $\varepsilon_2(n!) \sim n$ 和 $\varepsilon_3(n!) \sim n/2$, 所以就这一次 $\varepsilon_3(n!)$ 恰好为 $\varepsilon_2(n!)$ 的一半那样大看来是合理的. 例如, 当 $n=6$ 和 $n=7$ 时就出现这种情况, 因为 $6! = 2^4 \cdot 3^2 \cdot 5 = 7!/7$. 但是还没有人证明这种巧合经常出现.

$\varepsilon_p(n!)$ 的界限又转过来给出 $p^{e_p(n!)}$ 的一个界限, $p^{e_p(n!)}$ 是 p 对 $n!$ 的贡献:

$$p^{e_p(n!)} < p^{n/(p-1)},$$

且通过注意到 $p \leq 2^{p-1}$, 我们能(冒大大地放松上界的风险)简化此公式, 因此 $p^{n/(p-1)} \leq (2^{p-1})^{n/(p-1)} = 2^n$. 换句话说, 任何素数对 $n!$ 所做的贡献小于 2^n .

我们能利用此结果来取得有无限多素数的另一个证明. 如果仅有 k 个素数 $2, 3, \dots, P_k$, 则对所有 $n > 1$, 我们有 $n! < (2^n)^k = 2^{nk}$, 因为每个素数至多能提供 $2^n - 1$ 的一个因子. 但是选足够大的 n , 譬如说 $n = 2^{2k}$, 容易与 $n! < 2^{nk}$ 相矛盾. 于是

$$n! < 2^{nk} = 2^{2^{2k}k} = n^{n/2}$$

与式(4.22)中推出的不等式 $n! \geq n^{n/2}$ 相矛盾. 仍有无限多个素数.

我们甚至能加强此论证来取得不超过 n 的素数个数 $\pi(n)$ 的一个粗糙界限. 每一个这样的素数提供一个小于 2^n 的因子给 $n!$, 所以(如前)

$$n! < 2^{n\pi(n)}.$$

如果我们用 Stirling 近似式(4.23)来替换这里的 $n!$, 它是一个下界, 取对数, 我们得到

$$n\pi(n) > n \lg \left(\frac{n}{e} \right) + \frac{1}{2} \lg(2\pi n),$$

因此

$$\pi(n) > \lg \left(\frac{n}{e} \right).$$

这个下界比起实际值 $\pi(n) \sim n / \ln n$ 来说相当弱, 因为当 n 大时, $\lg n$ 比 $n / \lg n$ 小得多, 但是我们不必十分努力来取得它, 它仅为一个界限而已.

4.5 互素性

当 $\gcd(m, n) = 1$ 时, 整数 m 和 n 没有公共的素数因子, 我们称它们是互素的.

在实用中此概念很重要，我们该有一个特殊记号来记它，但是数论工作者还未提供一个很好的记法。所以我们呼吁：世界上的数学家们！不要让我们等太久！通过定义一个新记号，我们能使许多公式更清楚！让我们约定记‘ $m \perp n$ ’，且如果 m 和 n 是互素的，则称“ m 素于 n ”，换句话说，让我们表明

$$m \perp n \Leftrightarrow m, n \text{ 是整数且 } \gcd(m, n) = 1. \quad (4.26)$$

分数 m/n 在最小项中当且仅当 $m \perp n$ ，因为消去分子和分母的最大公因子，把分数减小到最小项，一般我们猜想

$$m / \gcd(m, n) \perp n / \gcd(m, n), \quad (4.27)$$

确实这是真的。它是依据习题 14 中证明的更一般的定律 $\gcd(km, kn) = k \gcd(m, n)$ 而得到的结果。

由 \gcd 规则(4.14)，当我们处理数的素数指数表示时， \perp 关系有一个简单的表述：

$$m \perp n \Leftrightarrow \min(m_p, n_p) = 0 \quad \text{对所有 } p. \quad (4.28)$$

此外，由于 m_p 和 n_p 是非负的，我们能把它改写成

$$m \perp n \Leftrightarrow m_p n_p = 0 \quad \text{对所有 } p. \quad (4.29)$$

现在我们能证明一个重要定律，由此能分离以及合并两个左边相同的 \perp 关系：

$$k \perp m \text{ 和 } k \perp n \Leftrightarrow k \perp mn. \quad (4.30)$$

考虑到式(4.29)，此定律的另一种说法是，当 m_p 和 n_p 是非负时， $k_p m_p = 0$ 和 $k_p n_p = 0$ 当且仅当 $k_p (m_p + n_p) = 0$ 。

有一种构造所有具有 $m \perp n$ 的非负分数 m/n 集合的极好方法，称为“Stern-Brocot 树”，由德国数学家 Moriz Stern^[279] 以及法国造表者 Achille Brocot^[35] 独立发现。方法的思想是从两个分数 $0/1$ ， $1/0$ 开始，然后如同希望的那样，重复下列操作多次：

在两个邻接的分数 m/n 和 m'/n' 之间插入 $(m+m')/(n+n')$ ，新的分数 $(m+m')/(n+n')$ 称为 m/n 和 m'/n' 的中间数。例如，第一步给出 $0/1$ 和 $1/0$ 之间的 1 个新项，

$$\frac{0}{1}, \frac{1}{1}, \frac{1}{0};$$

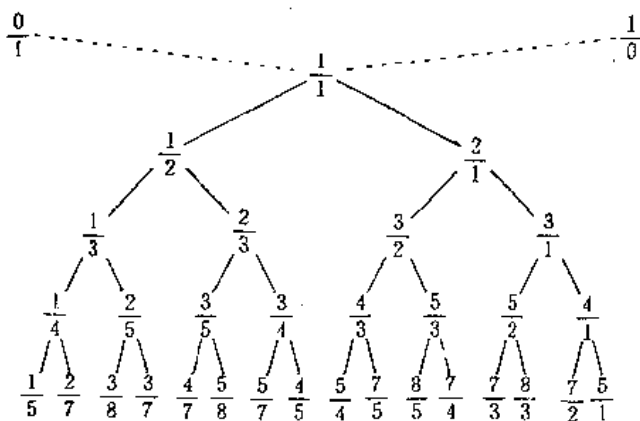
接着给出 2 个：

$$\frac{0}{1}, \frac{1}{2}, \frac{1}{1}, \frac{2}{1}, \frac{1}{0};$$

接着给出 4 个，

$$\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}, \frac{3}{2}, \frac{2}{1}, \frac{3}{1}, \frac{1}{0};$$

然后我们将给出 8 个，16 个，等等。整个数组能视为无限二叉树结构，它的顶部级如下：



每个分数是 $(m+m')/(n+n')$, 其中 m/n 是左边的上面的最接近的祖先, m'/n' 是右边的上面的最接近的祖先。(一个“祖先”是接着的分支向上可达到的一个分数。)在此树中能見到許多型式。

这个构造为什么行得通? 例如, 当每个中间分数 $(m+m')/(n+n')$ 出现在此树中时, 它为什么是在最小项中? (如果 m, m', n 和 n' 都是奇数, 我们将取得偶数/偶数; 不知怎么地, 构造保证几乎相互不出现具有奇分子和分母的分数。)且所有可能的分数 m/n 为什么恰好出现一次? 为什么一个特别的分数不能出现两次, 或者完全不出现?

所有这些问题令人惊异地有简单的回答, 基于下列基本事实: 如果在任何构造阶段 m/n 和 m'/n' 是相继的分数, 我们有

$$m'n - mn' = 1. \quad (4.31)$$

此关系开始是真的 ($1 \cdot 1 - 0 \cdot 0 = 1$); 当我们插入一个新中间值 $(m+m')/(n+n')$ 时, 需要检验的新情形是

$$\begin{aligned} (m+m')n - m(n+n') &= 1, \\ m'(n+n') - (m+m')n' &= 1. \end{aligned}$$

这样两个方程等价于替换的原来条件式(4.31), 所以在构造的所有阶段式(4.31)是不变式。

此外, 如果 $m/n < m'/n'$, 且如果所有值是非负的, 则易验证

$$\frac{m}{n} < \frac{m+m'}{n+n'} < \frac{m'}{n'}.$$

一个中间的分数不是它祖先间的一半的长度, 但是它位于中间的某处。所以构造维持次序, 且我们不可能在两个不同的地方取得相同的分数。

仍留下一个问题, 任何 a/b 的正分数 a/b 可能被略去吗? 回答是不能, 因为我们限于构造 a/b 的最接近的邻域, 在此区域中, 情况容易分析: 我们开始有

$$\frac{m}{n} = \frac{0}{1} < \left(\frac{a}{b}\right) < \frac{1}{0} = \frac{m'}{n'},$$

其中对 b/a 加括号表明它还未实际表示出, 于是如果在某阶段有

$$\frac{m}{n} < \left(\frac{a}{b}\right) < \frac{m'}{n'},$$

则构造形成 $(m+m')/(n+n')$, 且有三种情形。或者 $(m+m')/(n+n') = a/b$, 我们获得了; 或者 $(m+m')/(n+n') < a/b$, 我们能置 $m \leftarrow m+m'$, $n \leftarrow n+n'$; 或者 $(m+m')/(n+n') > a/b$, 我们能置 $m' \leftarrow m+m'$, $n' \leftarrow n+n'$ 。此过程不能无限地进行下去, 因为条件

$$\frac{a}{b} - \frac{m}{n} > 0 \quad \text{和} \quad \frac{m'}{n'} - \frac{a}{b} > 0,$$

意味着

$$an - bm \geq 1 \quad \text{和} \quad bm' - an' \geq 1,$$

因此

$$(m' + n')(an - bm) + (m + n)(bm' - an') \geq m' + n' + m + n,$$

依据式(4.31), 这与 $a+b \geq m'+n'+m+n$ 相同。在每步处 m 增加或者 n 增加或者 m' 增加或者 n' 增加, 所以至多 $a+b$ 步后我们一定获得。

阶 N 的 Farey 序列, 记为 \mathcal{F}_N , 是所有简约的 0 和 1 之间的分数, 它的分母小于或等于 N , 以上升次序排列。例如, 如果 $N=6$, 我们有

$$\mathcal{F}_6 = \frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1}.$$

我们一般能从 $\mathcal{F}_1 = 0/1, 1/1$ 开始, 然后插入中间数(每当可能这样做, 不取太大的一个分母)来得到 \mathcal{F}_N 。以此法不会漏掉任何分数, 因为我们知道 Stern-Brocot 构造不会漏掉任何分数, 且因为从分母 $> N$ 的一个分数不会形成一个分母 $\leq N$ 的中间数。(换句话说, \mathcal{F}_N 定义通过修剪不要的分支得到的 Stern-Brocot 树的子树。)由此得到每当 m/n 和 m'/n' 是 Farey 序列的相继元素时, $m'n - mn' = 1$ 。

这个构造的方法显示能以一种简单的方式从 \mathcal{F}_{N-1} 获得 \mathcal{F}_N : 我们在 \mathcal{F}_{N-1} 的相继分数 $m/n, m'/n'$ 之间简单地插入分数 $(m+m')/N$, 相继分数的分母共计为 N 。例如, 按照所述的规则插入 $1/7, 2/7, \dots, 6/7$, 从 \mathcal{F}_6 的元素得到 \mathcal{F}_7 是容易的,

$$\mathcal{F}_7 = \frac{0}{1}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{1}{1}.$$

当 N 是素数时, $N-1$ 个新分数将出现; 否则将有少于 $N-1$ 个数, 因为这个过程仅产生与 N 互素的分子。

很久以前, 在式(4.5)中我们以不同的说法证明了每当 $m \perp n$ 和 $0 < m \leq n$ 时, 能找到整数 a 和 b , 使得

$$ma - nb = 1. \quad (4.32)$$

(我们实际说 $m'm+n'n = \gcd(m, n)$, 但是对于 $\gcd(m, n)$ 我们能写 1, 对于 m' 写 a , 对于 $-n'$ 写 b 。)Farey 序列给出式(4.23)的另一个证明, 因为我们能让 b/a 是在 \mathcal{F}_n 中 m/n 前

面的分数, 因此式(4.5)也就是式(4.31)。例如, $3a-7b=1$ 的一个解是 $a=5$, $b=2$, 因在 \mathcal{F}_7 中 $2/5$ 在 $3/7$ 前面。这个构造意味着, 如果 $0 < m \leq n$, 对于 $0 \leq b < a < n$, 我们总能找到式(4.32)的解。相似, 如果 $0 \leq n < m$ 和 $m \perp n$, 让 a/b 是 \mathcal{F}_m 中跟在 n/m 后面的分数, 对于 $0 < a \leq b \leq m$, 我们能解式(4.32)。

在 Farey 序列中, 三个相继项的序列有一个习题 61 中证明的令人惊奇的性质。但是我们最好不再讨论 Farey 序列, 因为弄清楚整个 Stern-Brocot 树更加有趣。

我们事实上能把 Stern-Brocot 树视为代表有理数的一个数系, 因为每个正的, 简约分数恰好出现一次。当我们从树的根转到一个特殊分数时, 让我们用字母 L 和 R 表示下到左分支或右分支, 于是 L 和 R 的一个串唯一确定树中的一个位置。例如, $LRRL$ 意味着我们从 $1/1$ 下到左边的 $1/2$, 然后下到右边的 $2/3$, 然后下到右边的 $3/4$, 然后下到左边的 $5/7$ 。我们能把 $LRRL$ 考虑为 $5/7$ 的一种表示法。每一个正分数以此方式表示为唯一的一个 L 和 R 的串。

实际有一个小问题: 分数 $1/1$ 对应于空串, 对于它我们需要一个记法, 让我们称它为 I , 因为这看来有点像 1 , 而它表示“单位元素”。

这个表示法自然引出两个问题: (1) 给定 $m \perp n$ 的正整数 m 和 n , 对应于 m/n 的 L 和 R 的串是什么? (2) 给定 L 和 R 的一个串, 对应于它的分数是什么? 问题(2)看来较容易, 所以让我们首先处理它, 我们定义

$f(S)$ = 对应于 S 的分数,

其中 S 是 L 和 R 的串。例如, $f(LRRL) = 5/7$ 。

按照构造, 如果 m/n 和 m'/n' 是树的较高级中位于 S 前和位于 S 后最接近的分数, 则 $f(S) = (m+m')/(n+n')$ 。开始 $m/n = 0/1$ 和 $m'/n' = 1/0$; 然后如同我们分别在树中移到右边或左边那样, 相继用中间数 $(m+m')/(n+n')$ 替换 m/n 或 m'/n' 。

我们如何能抓住数学公式中的这种容易处理的情况? 一点儿试验启示, 最好的方法是运用 2×2 的一个矩阵

$$M(S) = \begin{pmatrix} n & n' \\ m & m' \end{pmatrix}$$

它有 4 个量, 涉及到围住 S 的祖先的分数 m/n 和 m'/n' 。在分数的位置方面, 我们能把 m 放在上面, n 放在下面; 但是这个颠倒的排列制订得相当好, 因为当过程开始时我们有 $M(I) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, 通常称 $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 为单位矩阵 I 。

向左边的一步用 $n+n'$ 替换 n' , 用 $m+m'$ 替换 m' , 因此

$$M(SL) = \begin{pmatrix} n & n+n' \\ m & m+m' \end{pmatrix} = \begin{pmatrix} n & n' \\ m & m' \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = M(S) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

(这是 2×2 矩阵相乘的一般规则

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} aw+by & ax+bz \\ cw+dy & cx+dz \end{pmatrix})$$

的特殊情形。)相似有

$$M(SR) = \begin{pmatrix} n+n' & n' \\ m+m' & m' \end{pmatrix} = M(S) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

所以如果定义 L 和 R 为 2×2 矩阵

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad (4.33)$$

通过对 S 的长度进行归纳, 我们得到简单的公式 $M(S) = S$. 这不是很好吗? (字母 L 和 R 起双重角色的作用, 在串表示中为矩阵, 也为字母。)例如,

$$\begin{aligned} M(LRRL) &= LRRL = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}, \end{aligned}$$

围住 $LRRL = 5/7$ 的祖先分数是 $2/3$ 和 $3/4$, 这个构造给出了问题(2)的解答:

$$f(S) = f\left(\begin{pmatrix} n & n' \\ m & m' \end{pmatrix}\right) = \frac{m+m'}{n+n'}. \quad (4.34)$$

问题(1)的情况如何呢? 这是容易的, 现在我们知道树结点和 2×2 矩阵之间的基本联系. 给定 $m \perp n$ 的一对正整数 m 和 n , 我们能通过“对半查找”找出 Stern-Brocot 树中 m/n 的位置如下:

```
S := I;
While m/n ≠ f(s) do
if m/n < f(s) then (输出(L); S := SL)
else (输出(R); S := SR).
```

此输出为希望的 L 和 R 的串.

通过改变 m 和 n , 而不是运用状态 S , 还有另一种方式来处理同样的工作. 如果 S 是任何 2×2 矩阵, 我们有

$$f(RS) = f(S) + 1,$$

因为 RS 和 S 相象, 但是上面的行加到下面的行上去. 让我们慢慢地查看它:

$$S = \begin{pmatrix} n & n' \\ m & m' \end{pmatrix}, \quad RS = \begin{pmatrix} n & n' \\ m+n & m'+n' \end{pmatrix},$$

因此, $f(S) = (m+m')/(n+n')$ 和 $f(RS) = ((m+n) + (m'+n'))/(n+n')$. 如果我们在 $m > n$ 的分数 m/n 上进行对半查找算法, 第一个输出将为 R ; 因此如果从 $(m-n)/n$ 开始而不是 m/n , 则算法的随后情况将得到的恰好比 $f(S)$ 大 1. 相似的性质对 L 成立,

$$\frac{m}{n} = f(RS) \Leftrightarrow \frac{m-n}{n} = f(S), \text{ 当 } m > n;$$

$$\frac{m}{n} = f(LS) \Leftrightarrow \frac{m}{n-m} = f(S), \text{ 当 } m < n.$$

```

While    $m \neq n$    do
    if     $m < n$    then (输出( $L$ );  $n := n - m$ )
    else  (输出( $R$ );  $m := m - n$ ).

```

$m=5$	5	3	1	1
$n=7$	2	2	2	1
输出	L	R	R	L

RRRLRLLLLRLRRRRRLRLLLLLLLRLR...

顺便指出, e 的代表有 Stern-Brocot 系中的一种规则的类型:

$$e = RL^0RLR^2LRL^4RLR^6LRL^8RLR^{10}LRL^{12}RL\cdots,$$

依据此表示, 我们能推出分数

$$\begin{array}{cccccccccccccccccccccccc} R & R & L & R & R & L & R & L & L & L & L & R & L & R & R & R & R & R & R \\ \frac{1}{1} & \frac{2}{1} & \frac{3}{1} & \frac{5}{2} & \frac{8}{3} & \frac{11}{4} & \frac{19}{7} & \frac{30}{11} & \frac{49}{18} & \frac{68}{25} & \frac{87}{32} & \frac{106}{39} & \frac{193}{71} & \frac{299}{110} & \frac{492}{181} & \frac{685}{252} & \frac{878}{323} & \frac{1071}{394} & \frac{1264}{465} & \dots \end{array}$$

是 e 的最简单的有理上近似和下近似。如果 m/n 不在此列表中出现, 则此列表中的某个分数(它的分子 $\leq m$, 它的分母 $\leq n$)位于 m/n 和 e 之间。例如, $27/10$ 不像 $19/7=2.714\cdots$ 为简单的一个近似, $19/7$ 在列表中出现且较接近 e 。我们能明白这一点, Stern-Brocot 树不仅包括所有有理数(它按次序包含它们), 且因为上面出现和具有小分子和分母的所有分数都小于原始的分数的。因此, $27/10=RRLRLL$ 小于 $19/7=RRLRRL$, 它小于 $e=RRLRRLR\cdots$ 。以此法能找到极好的近似。例如, $1264/465\approx 2.718280$ 和 e 具有相同的 6 个十进位; 从 e 的 Stern-Brocot 表示的前 19 个

字母, 我们得到此分数, 精确度大约到 e 的二进制表示的 19 位。

简单修改无矩阵对半查找过程, 我们能找到无理数 α 的无限表示:

```
if  $\alpha < 1$  then (输出( $L$ );  $\alpha := \alpha / (1 - \alpha)$ )
      else (输出( $R$ );  $\alpha := \alpha - 1$ ).
```

(这些步被重复无限多次, 或者直到我们感到厌倦了。)如果 α 是有理数, 以此法得到的无限表示和前面的相同, 但是在 α 的(有限)表示的右边添加 RL^∞ , 例如, 若 $\alpha = 1$, 我们取得 $RLLL\cdots$, 对应于极限趋于 1 的分数无限序列 $1/1, 2/1, 3/2, 4/3, 5/4, \cdots$. 如果我们把 L 想象为 0, R 想象为 1, 这种情况恰好相似于通常的二进制表示: 就像 $[0, 1)$ 中每个实数 x 具有不以全为 1 终止的一个无限二进制表示 $(.b_1 b_2 b_3 \cdots)_2$, $[0, \infty)$ 中每个实数 α 具有不以全为 R 终止的一个无限 Stern-Brocot 表示 $B_1 B_2 B_3 \cdots$. 如果让 $0 \longleftrightarrow L$ 和 $1 \longleftrightarrow R$, 在 $[0, 1)$ 和 $[0, \infty)$ 之间我们有一种保持次序的一一对应。

在欧几里德算法和有理数的 Stern-Brocot 表示之间有一个明白表示的关系。给定 $\alpha = m/n$, 我们取得 $\lfloor m/n \rfloor$ 个 R , 接着 $\lfloor n/(m \bmod n) \rfloor$ 个 L , 接着 $\lfloor (m \bmod n)/(n \bmod (m \bmod n)) \rfloor$ 个 R , 等等。这些数 $m \bmod n, n \bmod (m \bmod n), \cdots$ 恰好是欧几里德算法中检验的值。(确实没有无限多个 R 。)在第六章中我们将进一步探讨这个关系。

4.6 ‘MOD’: 同余关系

模数算术是数论提供了一种主要工具。当我们在第三章中用到二元运算 ‘mod’ 时, 已见到它, 通常为一个表达式中的一个运算。在本章中, 我们将也结合整个方程用 ‘mod’, 对于 ‘mod’ 用一个不同的记法是方便的:

$$a \equiv b \pmod{m} \Leftrightarrow a \bmod m = b \bmod m. \quad (4.35)$$

例如, $9 \equiv -16 \pmod{5}$, 因为 $9 \bmod 5 = 4 = (-16) \bmod 5$, 公式 ‘ $a \equiv b \pmod{m}$ ’ 能读作 “对模数 m , a 和 b 是同余”。当 a, b 和 m 是任意实数时, 定义有意义, 但是我们几乎仅对整数用它。

由于 $x \bmod m$ 和 x 相差一个 m 的倍数, 我们能以另一种方式理解同余:

$$a \equiv b \pmod{m} \Leftrightarrow a - b \text{ 是 } m \text{ 的一个倍数}. \quad (4.36)$$

如果 $a \bmod m = b \bmod m$, 则式 (3.21) 中的 ‘mod’ 的定义告诉我们, 对于某些整数 k 和 l , $a - b = a \bmod m + km - (b \bmod m + lm) = (k - l)m$. 反之, 如果 $a - b = km$, 那么如果 $m = 0$, 则 $a = b$; 否则

$$\begin{aligned} a \bmod m &= a - \lfloor a/m \rfloor m = b + km - \lfloor (b + km)/m \rfloor m \\ &= b - \lfloor b/m \rfloor m = b \bmod m. \end{aligned}$$

式 (4.36) 中的一的表征常比式 (4.35) 容易应用。例如, 我们有 $8 \equiv 23 \pmod{5}$, 因为 $8 - 23 = -15$ 是 5 的一个倍数; 我们不必计算 $8 \bmod 5$ 和 $23 \bmod 5$ 。

同余符号‘ \equiv ’看来像‘ $=$ ’一样方便，因为同余几乎像方程一样。例如，同余是一个等价关系；也就是说，它满足自反定律‘ $a \equiv a$ ’，对称律‘ $a \equiv b \Rightarrow b \equiv a$ ’和传递律‘ $a \equiv b \equiv c \Rightarrow a \equiv c$ ’。所有这些性质容易证明，因为任何关系‘ \equiv ’满足‘ $a \equiv b \Leftrightarrow f(a) = f(b)$ ’对于某个函数 f 是等价关系。（此时， $f(x) = x \bmod m$ 。）并且，我们能加减同余元素而不失去同余：

$$a \equiv b \text{ 和 } c \equiv d \Rightarrow a + c \equiv b + d \pmod{m};$$

$$a \equiv b \text{ 和 } c \equiv d \Rightarrow a - c \equiv b - d \pmod{m}.$$

如果 $a-b$ 和 $c-d$ 都是 m 的倍数，则 $(a+c)-(b+d) = (a-b)+(c-d)$ 和 $(a-c)-(b-d) = (a-b)-(c-d)$ 也是 m 的倍数。顺便提到，对于每次‘ \equiv ’的出现，完全不必写‘ \pmod{m} ’；如果模数是常量，我们仅需命名它一次来建立前后关系。这是一个同余记法的极大方便。

还有涉及整数提供的乘法计算：

$$a \equiv b \text{ 和 } c \equiv d \Rightarrow ac \equiv bd \pmod{m}, \text{ 整数 } b, c.$$

证明： $ac-bd = (a-b)c + b(c-d)$ 。现在重复应用这个乘法性质来取得幂：

$$a \equiv b \Rightarrow a^n \equiv b^n \pmod{m}, \text{ 整数 } a, b;$$

$$\text{整数 } n \geq 0.$$

例如，由于 $2 \equiv -1 \pmod{3}$ ，我们有 $2^n \equiv (-1)^n \pmod{3}$ ；这意味着 $2^n - 1$ 是 3 的倍数当且仅当 n 是偶数。

因此，我们通常对方程做的大多数代数运算也能对同余做。这里说的是大多数而不是全部。除法运算明显不存在。如果 $ad \equiv bd \pmod{m}$ ，我们不总能推出 $a \equiv b$ 。例如， $3 \cdot 2 \equiv 5 \cdot 2 \pmod{4}$ ，但是 $3 \not\equiv 5$ 。

然而在 d 和 m 互素的一般情形，我们能挽救同余的相消性质：

$$ad \equiv bd \Leftrightarrow a \equiv b \pmod{m}, \text{ 整数 } a, b, d, m \text{ 且 } d \perp m. \quad (4.37)$$

例如，如果模数 m 不是 5 的倍数，则从 $15 \equiv 35 \pmod{m}$ 推出 $3 \equiv 7 \pmod{m}$ 是合法的。

为了证明此性质，我们再用推广的 gcd 定律(4.5)，求 d' 和 m' 使得 $d'd + m'm = 1$ 。于是如果 $ad \equiv bd$ ，我们能在两边乘同余的 d' 得到 $ad'd \equiv bd'd$ 。由于 $d'd \equiv 1$ ，从而有 $ad'd \equiv a$ 和 $bd'd \equiv b$ ，因此 $a \equiv b$ 。此证明表明当考虑 \pmod{m} 同余时，数 d' 所起的作用几乎像 $1/d$ ；所以我们称它为“对模数 m 的 d 的倒数”。

应用除法到同余的另一种方式是像除其他数一样除模数：

$$ad \equiv bd \pmod{md} \Leftrightarrow a \equiv b \pmod{m}, \text{ 对 } d \neq 0. \quad (4.38)$$

此定律对所有实数 a, b, d 和 m 成立，因为它仅依赖于分配律 $(a \bmod m)d = ad \bmod md$ ：我们有 $a \bmod m \equiv b \bmod m \Leftrightarrow (a \bmod m)d \equiv (b \bmod m)d \Leftrightarrow ad \bmod md \equiv bd \bmod md$ 。因此，例如，从 $3 \cdot 2 \equiv 5 \cdot 2 \pmod{4}$ 我们推出 $3 \equiv 5 \pmod{2}$ 。

我们能把式(4.37)和(4.38)合在一起而取得尽可能减小模数的一般定律：

$$ad \equiv bd \pmod{m}$$

$$\Leftrightarrow a \equiv b \left(\text{mod} \frac{m}{\gcd(d, m)} \right), \text{ 整数 } a, b, d, m. \quad (4.39)$$

由于我们能用 d' 乘 $ad \equiv bd$, 其中 $d'd + m'm = \gcd(d, m)$; 这就给出同余 $a \cdot \gcd(d, m) \equiv b \cdot \gcd(d, m) \pmod{m}$, 它被 $\gcd(d, m)$ 除尽。

让我们再查看一下改变模数的思想。如果知道 $a \equiv b \pmod{100}$, 则我们也一定有 $a \equiv b \pmod{10}$, 或模 100 的任何因子。说 $a-b$ 是 100 的倍数比说 $a-b$ 是 10 的倍数要强。一般,

$$a \equiv b \pmod{md} \Rightarrow a \equiv b \pmod{m}, \quad \text{整数 } d. \quad (4.40)$$

因为 md 的任何倍数是 m 的倍数。

相反, 如果知道关于两个小模数, $a \equiv b$, 我们能否推出关于一个较大的模数 $a \equiv b$? 能推出; 规则是

$$\begin{aligned} a \equiv b \pmod{m} \quad \text{和} \quad a \equiv b \pmod{n} \\ \Leftrightarrow a \equiv b \pmod{\text{lcm}(m, n)}, \quad \text{整数 } m, n > 0. \end{aligned} \quad (4.41)$$

例如, 如果知道对模数 12 和 18, $a \equiv b$, 我们能可靠地推出 $a \equiv b \pmod{36}$ 。理由是, 如果 $a-b$ 是 m 和 n 的一个公倍数, 它是 $\text{lcm}(m, n)$ 的一个倍数。根据唯一的因子分解原理得到此结果。

此定律的特殊情形 $m \perp n$ 是极重要的, 因为当 m 和 n 互素时, $\text{lcm}(m, n) = mn$ 。所以我们将明确地叙述它:

$$\begin{aligned} a \equiv b \pmod{mn} \\ \Leftrightarrow a \equiv b \pmod{m} \quad \text{和} \quad a \equiv b \pmod{n}, \quad \text{若 } m \perp n. \end{aligned} \quad (4.42)$$

例如, $a \equiv b \pmod{100}$ 当且仅当 $a \equiv b \pmod{25}$ 和 $a \equiv b \pmod{4}$ 。以另一种方法表达这一点, 如果知道 $x \pmod{25}$ 和 $x \pmod{4}$, 则我们就有足够的论据来确定 $x \pmod{100}$ 。这是中国的剩余定理(见习题 30)的特殊情形, 这样称它是因为大约公元 350 年中国数学家孙子发现了它。

式(4.42)中的模数 m 和 n 能进一步被分解为互素的因子直到分离出每一个不同的素数。所以如果 m 的素数因子分解为 $\prod_p p^{m_p}$, 则

$$a \equiv b \pmod{m} \Leftrightarrow a \equiv b \pmod{p^{m_p}} \text{ 对所有 } p.$$

模数为素数的幂的同余建立了模数为整数的所有同余。

4.7 独立剩余

同余的一种重要应用是剩余数系, 在此数系中把整数 x 表示为剩余的一个序列, 剩

余是关于彼此互素的模数的

$$\text{Res}(x) = (x \bmod m_1, \dots, x \bmod m_r), \text{ 若 } m_j \perp m_k, \text{ 对 } 1 \leq j < k \leq r.$$

知道 $x \bmod m_1, \dots, x \bmod m_r$ 并没有告诉我们关于 x 的一切事情。但是它的确允许我们确定 $x \bmod m$, 其中 m 是乘积 $m_1 \cdots m_r$ 。在实际应用中我们常常知道 x 在一个确定的范围内; 于是如果知道 $x \bmod m$, 且若 m 足够大, 我们将知道 x 的一切事情。

例如, 让我们查看仅有二个模数 3 和 5 的剩余数系的小的情形:

$x \bmod 15$	$x \bmod 3$	$x \bmod 5$	$x \bmod 15$	$x \bmod 3$	$x \bmod 5$
0	0	0	8	2	3
1	1	1	9	0	4
2	2	2	10	1	0
3	0	3	11	2	1
4	1	4	12	0	2
5	2	0	13	1	3
6	0	1	14	2	4
7	1	2			

每个有序对 $(x \bmod 3, x \bmod 5)$ 是不同的, 因为 $x \bmod 3 = y \bmod 3$ 和 $x \bmod 5 = y \bmod 5$ 当且仅当 $x \bmod 15 = y \bmod 15$ 。

由于同余的规则, 我们能在两个组成部分上独立完成加、减和乘。例如, 如果要以模数 15 用 $13 = (1, 3)$ 乘 $7 = (1, 2)$, 我们计算 $1 \cdot 1 \bmod 3 = 1$ 和 $2 \cdot 3 \bmod 5 = 1$ 。解答是 $(1, 1) = 1$, 因此 $7 \cdot 13 \bmod 15$ 一定等于 1。果然它算出了。

在计算机应用中, 这个独立的原理是有用的, 因为不同的组成部分能分开处理(例如, 用不同的计算机。)如果每个模数 m_k 是一个不同的素数 p_k , 选取稍微小于 2^{31} , 则基本算术运算处理范围 $[-2^{31}, 2^{31})$ 中的整数的计算机能容易地计算模数 p_k 的和, 差和乘积。 r 个这样的素数的集合使它可能来加, 减和乘“多倍精度数”直到几乎 $31r$ 个位, 而剩余系使它可能做这一点比以其他方式加, 减和乘这样大的数要快。

在适当的情况中, 我们甚至能做除法。例如, 假设我们要计算整数的一个大行列式的准确值。结果将是一个整数 D , 根据它的项的大小能给出 $|D|$ 的界限。但是计算行列式的快速方法只知道要求除法, 而这将引出分数(且损失精确度, 如果我们采用二进制近似。)补救的办法是对不同的大素数 p_k , 计算 $D \bmod p_k = D_k$ 。如果除数不出现为 p_k 的倍数, 我们确实能除尽模数。这是很靠不住的。但是如果这的确出现, 我们能选另一个素数。最后, 对于充分多的素数, 知道 D_k , 我们将有足够信息来确定 D 。

但是我们还未说明, 从剩余的给定序列 $(x \bmod m_1, \dots, x \bmod m_r)$ 如何取回 $x \bmod m$ 。我们已表明原则上能完成此转换, 但是计算可能是很可怕的, 在应用中可能取消此主意。幸好有一种相当简单的方法来做这种工作, 且我们能就表中表明的情況来说明它。中心思想是就两种情形 $(1, 0)$ 和 $(0, 1)$ 来解问题; 如果 $(1, 0) = a$ 和 $(0, 1) = b$, 则

$(x, y) = (ax + by) \pmod{15}$, 这是由于同余能相乘和相加。

此时通过查表 $a = 10$ 和 $b = 6$; 但是当模数非常大时, 我们如何能求得 a 和 b ? 换句话说, 如果 $m \perp n$, 找数 a 和 b 使得方程

$$a \pmod{m} = 1, a \pmod{n} = 0, b \pmod{m} = 0, b \pmod{n} = 1$$

全成立的好方法是什么? 再一次用式(4.5)来挽救: 用欧几里德算法, 我们能找到 m' 和 n' 使得

$$m'm + n'n = 1.$$

所以我们能取 $a = n'n$ 和 $b = m'm$, 把它们都化为希望的 \pmod{mn} .

当模数大时, 为了减少计算需要进一步的诀窍; 详细讨论超出了本书的范围, 但是在 [174, p.274] 中能找到它们。从剩余转换到对应的原来数是可实现的, 但它是十分慢的, 仅当在转换回去之前剩余数系中完成所有的运算序列时, 我们节省了总时间。

让我们试解一个小的问题来加强这些同余的思想: 如果当 $x \equiv x'$ 时我们考虑两个解 x 和 x' 是相同的, 则同余式

$$x^2 \equiv 1 \pmod{m} \quad (4.43)$$

有多少个解?

依据前面说明的一般原理, 我们该首先考虑 m 是一个素数幂 p^k ($k > 0$) 的情形。则同余式 $x^2 - 1$ 能写成

$$(x-1)(x+1) \equiv 0 \pmod{p^k},$$

所以 p 一定除尽 $x-1$, 或 $x+1$, 或者两者。但是除非 $p=2$, p 不能除尽 $x-1$ 和 $x+1$ 两者, 我们将把此情形留在后面讨论。如果 $p > 2$, 则 $p^k \mid (x-1)(x+1) \Leftrightarrow p^k \mid (x-1)$ 或 $p^k \mid (x+1)$, 所以恰好有两个解 $x \equiv +1$ 和 $x \equiv -1$ 。

情形 $p=2$ 有点不同。如果 $2^k \mid (x-1)(x+1)$ 则 $x-1$ 或者 $x+1$ 被 2 除尽, 但不被 4 除尽。所以另一个一定被 2^{k-1} 除尽。这意味着当 $k \geq 3$ 时我们有 4 个解, 即 $x \equiv +1$ 和 $x \equiv -1 \pm 1$ 。(例如, 当 $p^k=8$ 时 4 个解是 $x \equiv 1, 3, 5, 7 \pmod{8}$; 知道任何具有形式 $8n+1$ 的奇整数的平方是有用的。)

现在 $x^2 \equiv 1 \pmod{m}$ 当且仅当对于在 m 的完全因子分解中 $m_p > 0$ 的所有素数 p , $x^2 \equiv 1 \pmod{p^{m_p}}$ 。每个素数和其他素数是独立的, 且除当 $p=2$ 之外, 对于 $x \pmod{p^{m_p}}$ 恰好有两种可能性。所以如果 m 恰好有 r 个不同的素数因子, $x^2 \equiv 1$ 的解的总数是 2^r , 除了当 m 是偶数时有一校正。

确切数一般是

$$2^{r + \lceil 8\sqrt{m} \rceil + \lceil 4\sqrt{m} \rceil - \lceil 2\sqrt{m} \rceil} \quad (4.44)$$

例如, 有 4 个“模数 12 的单位的平方根”, 即 1, 5, 7 和 11。当 $m=15$ 时, 这就是剩余 $\pmod{3}$ 和 $\pmod{5}$ 是 ± 1 的 4 个, 即剩余数系中的 (1, 1), (1, 4), (2, 1) 和 (2, 4), 这些

解是通常(十进制)数系中的 1, 4, 11 和 14.

4.8 附加的应用

第三章还留下一些未完成的工作: 我们希望证明 m 个数

$$0 \bmod m, n \bmod m, 2n \bmod m, \dots, (m-1)n \bmod m \quad (4.45)$$

精确地由以某种次序的 m/d 个数的 d 个拷贝所组成, 其中 $d = \gcd(m, n)$.

$$0, d, 2d, \dots, m-d$$

例如, $m=12$ 和 $n=8$, 我们有 $d=4$, 而数为 0, 8, 4, 0, 8, 4, 0, 8, 4, 0, 8, 4.

证明的第一部分表明我们取得前 m/d 个数的 d 个拷贝, 是平凡的. 依据式(4.38), 我们有

$$jn \equiv kn \pmod{m} \Leftrightarrow j\left(\frac{n}{d}\right) \equiv k\left(\frac{n}{d}\right) \pmod{\frac{m}{d}},$$

因此当 $0 \leq k < m/d$ 时我们取得出现值的 d 个拷贝.

现在我们一定要证明这 m/d 个数是以某种次序的 $\{0, d, 2d, \dots, m-d\}$. 让我们记 $m = m'd$ 和 $n = n'd$. 依据分配律(3.23), $kn \bmod m = d(kn' \bmod m')$, 所以当 $0 \leq k \leq m'$ 出现的值是 d 倍数

$$0 \bmod m', n' \bmod m', 2n' \bmod m', \dots, (m'-1)n' \bmod m'.$$

但是我们由式(4.27)知道 $m' \perp n'$, 我们已分出它们的最大公因子. 所以我们仅需考虑 $d=1$ 的情形, 即 m 和 n 互素的情形.

让我们假设 $m \perp n$. 此时用“鸽巢原理”容易看出式(4.45)的数就是以某种次序的 $\{0, 1, \dots, m-1\}$. 鸽巢原理说, 如果把 n 只鸽子放入 m 个鸽巢, 有一个空巢当且仅当有一个鸽巢有多于一只鸽子. (与在习题 3.8 中证明的 Dirichlet 匣子原理是相同的.) 我们知道式(4.45)的数是不同的, 因为当 $m \perp n$ 时

$$jn \equiv kn \pmod{m} \Leftrightarrow j \equiv k \pmod{m},$$

这是式(4.37). 所以 m 个不同的数一定填满所有鸽巢 $0, 1, \dots, m-1$. 所以完成了第三章未结束的工作.

完成了证明, 但是我们还能用一种直接的方法来证明, 而不是依赖于间接的鸽巢论证. 如果 $m \perp n$ 以及如果给定一个值 $j \in [0..m)$, 我们通过对 k 解同余式

$$kn \equiv j \pmod{m}$$

能明确地计算出 $k \in [0..m)$ 使得 $kn \bmod m = j$. 我们用 n' 简单地乘两边, 其中

$m'm+n'n=1$, 取得

$$k = jn' \pmod{m},$$

因此 $k = jn' \pmod{m}$.

我们能利用刚证的结果来建立 1640 年 Pierre de Fermat 发现的一个重要结果。Fermat 是一个大数学家, 他对微积分的发现和数学的许多其他分支都有贡献。在他留下的笔记中包含许多未加证明的定理, 这些定理除一个之外相继被证实。这剩下的一个定理称为“Fermat 的最后定理”, 该定理陈述为, 当 $n > 2$, 对于所有正整数 a, b, c 和 n ,

$$a^n + b^n \neq c^n. \quad (4.46)$$

(当然方程 $a+b=c$ 和 $a^2 + b^2 = c^2$ 有许多解。)Tanner 和 Wagstaff^[285]对 $n \leq 150\,000$, 证实了这个猜测。

1640 年的 Fermat 定理是一个可证明的定理。现在称它为 Fermat 小定理(或简称 Fermat 定理), 陈述为

$$n^{p-1} \equiv 1 \pmod{p}, \text{ 如果 } n \perp p. \quad (4.47)$$

证明: 我们通常假设 p 表示一个素数。我们知道 $p-1$ 个数 $n \pmod{p}, 2n \pmod{p}, \dots, (p-1)n \pmod{p}$ 是以某种次序排列的数 $1, 2, \dots, p-1$, 所以如果把它们乘在一起, 我们取得

$$\begin{aligned} n \cdot (2n) \cdot \dots \cdot ((p-1)n) \\ &= (n \pmod{p}) \cdot (2n \pmod{p}) \cdot \dots \cdot ((p-1)n \pmod{p}) \\ &\equiv (p-1)! \pmod{p}, \end{aligned}$$

其中同余是模数 p 。这意味着

$$(p-1)!n^{p-1} \equiv (p-1)! \pmod{p},$$

并能消去 $(p-1)!$, 由于它不被 p 所除尽。证毕。

Fermat 定理的另一种形式有时更方便:

$$n^p \equiv n \pmod{p}, \quad \text{整数 } n. \quad (4.48)$$

此同余式对所有整数 n 成立。证明是容易的: 如果 $n \perp p$, 我们只要将式(4.47)乘以 n ; 如果不是 $n \perp p$, 则 $p \mid n$, 所以 $n^p \equiv 0 \equiv n$.

同一年他发现了式(4.47), Fermat 给 Mersenne 写了一封信, 说他猜想对于所有 $n \geq 0$, 数

$$f_n = 2^{2^n} + 1$$

将是素数。他知道前 5 种情形给出素数:

$2^1 + 1 = 3$; $2^2 + 1 = 5$; $2^4 + 1 = 17$; $2^8 + 1 = 257$; $2^{16} + 1 = 65\,537$; 但是他不明白如何证明下一个情形, $2^{32} + 1 = 4\,294\,967\,297$ 将是素数。

注意到用 Fermat 发现的定理可证明 $2^{32} + 1$ 不是素数是有趣的：在式(4.47)中我们能置 $n=3$ ，推得

$$3^{2^{32}} \equiv 1 \pmod{2^{32} + 1}, \text{ 如果 } 2^{32} + 1 \text{ 是素数.}$$

用手来试验这个关系是可能的，从 3 开始，且平方 32 次，保存模数 $2^{32} + 1$ 的剩余。首先我们有 $3^2 = 9$ ，接着 $3^{2^2} = 81$ ，然后 $3^{2^3} = 6561$ ，等等直到我们达到

$$3^{2^{32}} \equiv 3\,029\,026\,160 \pmod{2^{32} + 1}.$$

结果不是 1，所以 $2^{32} + 1$ 不是素数。这个推翻的方法没有给我们有关因子可能是什么的线索，但是它证明了因子存在。（它们是 641 和 6700417。）

如果对模数 $2^{32} + 1$ ， $3^{2^{32}}$ 结果为 1，计算并不证明 $2^{32} + 1$ 是素数；这仅不推翻它。但是习题 47 讨论了 Fermat 定理的逆定理，用它我们能证明大量基本数是素数，而不必作大量费力的算术。

我们从同余式的两边消去 $(p-1)!$ 来证明 Fermat 定理。结果对模数 p ， $(p-1)!$ 总和 -1 同余；这是称为 Wilson 定理的经典结果的一部分：

$$(n-1)! \equiv -1 \pmod{n} \Leftrightarrow n \text{ 是素数, 若 } n > 1. \quad (4.49)$$

此定理的一半是平凡的：如果 $n > 1$ 不是素数，则它有一个素因子 p ，呈现为 $(n-1)!$ 的一个因子，所以 $(n-1)!$ 不能和 -1 同余。（如果对模数 n ， $(n-1)!$ 和 -1 同余，则对模数 p ，它也将和 -1 同余，但并不是。）

Wilson 定理的另一半陈述 $(p-1)! \equiv -1 \pmod{p}$ 。通过对模数 p 配合数和它们的逆，我们能证明这一半。如果 $n \leq p$ ，我们知道存在 n' 使得

$$n'n \equiv 1 \pmod{p};$$

这里 n' 是 n 的逆， n 也是 n' 的逆。 n 的任何 2 个逆一定彼此同余，因为 $nn' \equiv nn''$ 意味着 $n' \equiv n''$ 。

现在假设我们对 1 和 $p-1$ 之间的每个数和它的逆配成对。由于一个数和它的逆的乘积和 1 同余，所以所有逆对中的所有数的乘积也和 1 同余，由此看来 $(p-1)!$ 和 1 同余。让我们来检验，譬如说对 $p=5$ 。我们取得 $4! = 24$ ；但是对模数 5，这是和 4 同余，而不是和 1 同余。错在哪里？让我们仔细查看逆：

$$1' = 1, \quad 2' = 3, \quad 3' = 2, \quad 4' = 4.$$

2 和 3 配对，但 1 和 4 不配对——它们的逆就是它们自身。

为了恢复我们的分析，一定要确定哪些数是它们自身的逆。如果 x 是它的自身的逆，则 $x^2 \equiv 1 \pmod{p}$ ；且我们已经证明，当 $p > 2$ 时，这个同余式恰好有两个根。（如果 $p=2$ ，这是明显的， $(p-1)! \equiv -1$ ，所以我们不必为这种情形操心。）根是 1 和 $p-1$ ，（在 1 和 $p-1$ 之间的）其他数配成对，因此正如希望的那样

$$(p-1)! \equiv 1 \cdot (p-1) \equiv -1.$$

不幸, 我们不能有效地计算阶乘, 所以 Wilson 定理不能用来实际测试素性, 它仅为一个定理.

4.9 φ 和 μ

整数 $\{0, 1, \dots, m-1\}$ 中有多少个数与 m 互素? 这是一个称为 m 的“装载”的重要的量 $\varphi(m)$ (是由爱取新名字的英国数学家 J.J. Sylvester^[283]命名的). 我们有 $\varphi(1)=1$, $\varphi(p)=p-1$, 以及对于所有合成数 m , $\varphi(m) < m-1$.

φ 函数被称为欧拉装载函数, 因为欧拉首先研究它. 例如, 欧拉发现 Fermat 定理 (4.47) 能以下列方式推广到非素数模数:

$$n^{\varphi(m)} \equiv 1 \pmod{m}, \text{ 如果 } n \perp m. \quad (4.50)$$

(习题 32 要求欧拉定理的一种证明.)

如果 m 是一个素数幂 p^k , 则易算 $\varphi(m)$, 因为 $n \perp p^k \Leftrightarrow p \nmid n$. 在 $\{0, 1, \dots, p^k-1\}$ 中的 p 的倍数是 $\{0, p, 2p, \dots, p^k-p\}$, 因此有它们的 p^{k-1} 个, $\varphi(p^k)$ 计数留下的:

$$\varphi(p^k) = p^k - p^{k-1}.$$

注意, 此公式正确地给出了当 $k=1$ 时 $\varphi(p)=p-1$.

如果 $m > 1$ 不是一个素数幂, 我们能记 $m = m_1 m_2$, 其中 $m_1 \perp m_2$. 于是数 $0 \leq n < m$ 能在剩余数系中表示为 $(n \bmod m_1, n \bmod m_2)$. 我们用式 (4.30) 和 (4.4) 得到

$$n \perp m \Leftrightarrow n \bmod m_1 \perp m_1 \text{ 和 } n \bmod m_2 \perp m_2.$$

因此, 如果我们考虑互素性是一种性能, 则 $n \bmod m$ 是“适合的”当且仅当 $n \bmod m_1$ 和 $n \bmod m_2$ 都是“适合的”. 现在能递归地计算模数 m 的适合值的总个数: 它是 $\varphi(m_1) \cdot \varphi(m_2)$, 因为有 $\varphi(m_1)$ 种适合的方式在剩余表示中选第一部分 $n \bmod m_1$, 有 $\varphi(m_2)$ 种适合的方式在剩余表示中选第二部分 $n \bmod m_2$.

例如, $\varphi(12) = \varphi(4)\varphi(3) = 2 \cdot 2 = 4$, 因为 n 对 12 是素数当且仅当 $n \bmod 4 = (1 \text{ 或 } 3)$ 和 $n \bmod 3 = (1 \text{ 或 } 2)$. 在剩余数系中, 对 12 为素数的 4 个值为 $(1, 1), (1, 2), (3, 1), (3, 2)$; 在通常十进制表示中, 它们是 1, 5, 7, 11. 欧拉定理陈述为: 每当 $n \perp 12$ 时 $n^4 \equiv 1 \pmod{12}$.

如果 $f(1)=1$ 且每当 $m_1 \perp m_2$ 时

$$f(m_1 m_2) = f(m_1) f(m_2) \quad (4.51)$$

则称正整数的函数 $f(m)$ 为积性函数. 我们刚才已证 $\varphi(m)$ 是积性函数. 在本章前面我们也已看到另一个积性函数的实例: $x^2 \equiv 1 \pmod{m}$ 的不同余解的个数是积性的. 还有另一个例子是对于任何幂 α , $f(m) = m^\alpha$.

一个积性函数完全由在素数幂处它的值完全确定，因为我们能把任何正整数 m 分解为它的素数幂因子，而因子是彼此互素的，一般公式

$$f(m) = \prod_p f(p^{m_p}), \text{ 如果 } m = \prod_p p^{m_p} \quad (4.52)$$

成立当且仅当 f 是积性的。

此公式特别给出对于一般 m 欧拉装载函数的值：

$$\varphi(m) = \prod_{p|m} (p^{m_p} - p^{m_p-1}) = m \prod_{p|m} \left(1 - \frac{1}{p}\right). \quad (4.53)$$

例如， $\varphi(12) = (4-2)(3-1) = 12(1-1/2)(1-1/3)$ 。

现在让我们来看把 φ 函数应用到有理数 mod 1 的研究。我们说如果 $0 \leq m < n$ ，分数 m/n 是基本的分数。所以 $\varphi(n)$ 是具有分母 n 的简约基本分数的个数；且 Farey 序列 \mathcal{F}_n 包含所有分母为 n 或小于 n 的简约基本分数，还有非基本分数 $1/1$ 。

在简化到最小项之前，具有分母 12 的所有基本分数的集合是

$$\frac{0}{12}, \frac{1}{12}, \frac{2}{12}, \frac{3}{12}, \frac{4}{12}, \frac{5}{12}, \frac{6}{12}, \frac{7}{12}, \frac{8}{12}, \frac{9}{12}, \frac{10}{12}, \frac{11}{12}.$$

简化产生

$$\frac{0}{1}, \frac{1}{12}, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{5}{12}, \frac{1}{2}, \frac{7}{12}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, \frac{11}{12}.$$

且以它们的分母把这些分数归类：

$$\frac{0}{1}, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{6}, \frac{5}{6}, \frac{1}{12}, \frac{5}{12}, \frac{7}{12}, \frac{11}{12}.$$

这样处理我们能得到什么呢？12 的每一个因子 d 连同它的分子的所有 $\varphi(d)$ 出现成一个分母。仅出现的分母是 12 的因子。因此

$$\varphi(1) + \varphi(2) + \varphi(3) + \varphi(4) + \varphi(6) + \varphi(12) = 12.$$

如果我们从未简约的分数 $0/m, 1/m, \dots, (m-1)/m$ (任何 m) 开始，相似的结果显然发生，因此

$$\sum_{d|m} \varphi(d) = m. \quad (4.54)$$

在本章接近开始处，我们说到数论中的问题常要求在一个数的因子上求和，公式 (4.54) 是这样的和，所以证明了我们的论断是正确的。(我们将看其他例子。)

现在这里是一个好奇的事实：如果 f 是任何函数使得和

$$g(m) = \sum_{d|m} f(d)$$

是积性的函数，则 f 自身是积性的函数。(这个结果和式 (4.45) 以及 $g(m) = m$ 显然是积性的事实一起，给出 $\varphi(m)$ 是积性的另一个理由)。我们对 m 归纳能证明这个好奇的事实：基

础步是容易的, 因为 $f(1)=g(1)=1$. 设 $m>1$, 且假设每当 $m_1 \mid m_2$ 以及 $m_1 m_2 < m$ 时, $f(m_1 m_2)=f(m_1)f(m_2)$. 如果 $m=m_1 m_2$ 和 $m_1 \perp m_2$, 我们有

$$g(m_1 m_2) = \sum_{d \mid m_1 m_2} f(d) = \sum_{d_1 \mid m_1} \sum_{d_2 \mid m_2} f(d_1 d_2)$$

和 $d_1 \perp d_2$, 因为 m_1 的所有因子和 m_2 的所有因子互素. 依据归纳假设, 除了可能当 $d_1 = m_1$ 和 $d_2 = m_2$ 外, $f(d_1 d_2)=f(d_1)f(d_2)$; 因此我们得到

$$\begin{aligned} & \left(\sum_{d_1 \mid m_1} f(d_1) \right) \left(\sum_{d_2 \mid m_2} f(d_2) \right) = f(m_1)f(m_2) + f(m_1 m_2) \\ & = g(m_1)g(m_2) - f(m_1)f(m_2) + f(m_1 m_2). \end{aligned}$$

但是这等于 $g(m_1 m_2) = g(m_1)g(m_2)$, 所以 $f(m_1 m_2)=f(m_1)f(m_2)$.

反之, 如果 $f(m)$ 是积性的, 则对应的因子上求和的函数 $g(m)=\sum_{d \mid m} f(d)$ 总是积性的. 事实上, 习题 33 证明还是真的. 因此好奇的事情和它的逆都是真实的.

对于所有 $m \geq 1$, 由方程

$$\sum_{d \mid m} \mu(d) = [m=1] \quad (4.55)$$

定义 Möbius 函数 $\mu(m)$. (这样命名为了纪念 19 世纪数学家 August Möbius.) 此方程实际是一个递归, 因为左边是由 $\mu(m)$ 和 $\mu(d)(d < m)$ 的一些值组成的和. 例如, 如果我们相继代入 $m=1, 2, \dots, 12$, 我们能计算前 12 个值:

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

Möbius 用递归公式(4.55)提出, 因为他注意到它对应下列重要的“反演原理”:

$$g(m) = \sum_{d \mid m} f(d) \Leftrightarrow f(m) = \sum_{d \mid m} \mu(d) g\left(\frac{m}{d}\right). \quad (4.56)$$

按照这个原理, μ 函数给出一种新方式来理解任何函数 $f(m)$, 对于它我们知道 $\sum_{d \mid m} f(d)$.

公式(4.56)的证明用了本章开始处我们所描述的两个诀窍式(4.7)和(4.9): 如果 $g(m) = \sum_{d \mid m} f(d)$, 则

$$\begin{aligned} \sum_{d \mid m} \mu(d) g\left(\frac{m}{d}\right) &= \sum_{d \mid m} \mu\left(\frac{m}{d}\right) g(d) = \sum_{d \mid m} \mu\left(\frac{m}{d}\right) \sum_{k \mid d} f(k) \\ &= \sum_{k \mid m} \sum_{d \mid (m/k)} \mu\left(\frac{m}{kd}\right) f(k) = \sum_{k \mid m} \sum_{d \mid (m/k)} \mu(d) f(k) \\ &= \sum_{k \mid m} [m/k=1] f(k) = f(m). \end{aligned}$$

相似的方法可证明式(4.56)的另一半(见习题 12).

关系式(4.56)给出 Möbius 函数的一个有用的性质, 且我们列出前 12 个值; 但是当 m

大时 $\mu(m)$ 的值是什么? 我们如何能解递归(4.55)? 函数 $g(m) = [m=1]$ 显然是积性的, 除了当 $m=1$ 之外终究是零。所以像刚才证明的那样, 式(4.55)定义的 Möbius 函数一定是积性的。所以如果计算 $\mu(p^k)$, 我们能计算出 $\mu(m)$ 是什么。

当 $m = p^k$ 时, 式(4.55)说明对所有 $k \geq 1$,

$$\mu(1) + \mu(p) + \mu(p^2) + \cdots + \mu(p^k) = 0,$$

因为 p^k 的因子是 $1, \dots, p^k$ 。由此得到

$$\mu(p) = -1; \quad \mu(p^k) = 0 \quad \text{对 } k > 1.$$

所以依据式(4.52), 我们得到一般公式

$$\mu(m) = \prod_{p|m} \mu(p^{m_p}) = \begin{cases} (-1)^r, & \text{如果 } m = p_1 p_2 \cdots p_r; \\ 0, & \text{如果某个 } p^2 \text{ 可除尽 } m. \end{cases} \quad (4.57)$$

这就是 μ 。

如果我们把式(4.54)视为函数 $\varphi(m)$ 的一个递归, 则能用 Möbius 规则(4.56)解递归。产生的解是

$$\varphi(m) = \sum_{d|m} \mu(d) \frac{m}{d}. \quad (4.58)$$

例如,

$$\begin{aligned} \varphi(12) &= \mu(1) \cdot 12 + \mu(2) \cdot 6 + \mu(3) \cdot 4 + \mu(4) \cdot 3 + \mu(6) \cdot 2 + \mu(12) \cdot 1 \\ &= 12 - 6 - 4 + 0 + 2 + 0 = 4 \end{aligned}$$

如果 r 个不同的素数, 譬如说 $\{p_1, \dots, p_r\}$ 可除尽 m , 和式(4.58)仅有 2^r 个非零项, 因为 μ 函数常为零。因此我们能看出式(4.58)和(4.53)相符合, 它标成

$$\varphi(m) = m \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right);$$

如果我们乘出 r 个因子 $(1 - 1/p_i)$, 则精确地取得式(4.58)的 2^r 个非零项。与此相比, Möbius 函数的优点是它可用于许多场合。

例如, 让我们试计算出 Farey 序列 \mathcal{F}_n 中有多少个分数。这是 $[0, 1]$ 中分母不超过 n 的简约分数的个数, 所以它比 $\Phi(n)$ 大 1, 其中我们定义

$$\Phi(x) = \sum_{1 \leq k \leq x} \varphi(k). \quad (4.59)$$

(我们一定要把 1 加到 $\Phi(n)$, 因为最后分数 $1/1$ 。)求式(4.59)中的和看来困难, 但是注意对于所有实数 $x \geq 0$,

$$\sum_{d \geq 1} \Phi\left(\frac{x}{d}\right) = \frac{1}{2} \lfloor x \rfloor + x \quad (4.60)$$

我们能间接确定 $\Phi(x)$ 。这个等式为什么成立? 看来它有点害怕, 然而实际并不超出我们

的知识范围。计简约的和未简约的分数，有 $1/2 \lfloor x \rfloor \lfloor 1+x \rfloor$ 个基本的分数 m/n ($0 \leq m < n \leq x$)，这给出右边。具有 $\gcd(m, n) = d$ 的这样的分数的个数为 $\Phi(x/d)$ ，因为用 $m'd$ 替换 m ， $n'd$ 替换 n 之后，这样的分数是 m'/n' ($0 \leq m' < n' \leq x/d$)。所以左边以不同方式计算了相同的分数，等式一定是真的。

让我们更仔细地查看此场合，以致方程(4.59)和(4.60)变得更清楚。 $\Phi(x)$ 的定义意味着 $\Phi(x) = \Phi(\lfloor x \rfloor)$ ，但是对于任意实值定义 $\Phi(x)$ 是方便的，不仅仅对整数定义 $\Phi(x)$ 。在整数数值处我们有表

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

当 $x = 12$ 时，我们能检验式(4.60)：

$$\begin{aligned} \Phi(12) + \Phi(6) + \Phi(4) + \Phi(3) + \Phi(2) + \Phi(2) + 6 \cdot \Phi(1) \\ = 46 + 12 + 6 + 4 + 2 + 2 + 6 = 78 = \frac{1}{2} \cdot 12 \cdot 13. \end{aligned}$$

令人惊异。

等式(4.60)能视为 $\Phi(x)$ 的一个明显的递归；例如，我们刚才看到，我们用它从 $\Phi(m)$ ($m < 12$) 的一些值来计算 $\Phi(12)$ 。我们还能用另一个 Möbius 函数的极好性质来解这样的递归：

$$g(x) = \sum_{d \geq 1} f\left(\frac{x}{d}\right) \Leftrightarrow f(x) = \sum_{d \geq 1} \mu(d) g\left(\frac{x}{d}\right). \quad (4.61)$$

对于所有函数 f 使得 $\sum_{k, d \geq 1} |f(x/kd)| < \infty$ ，这个反演律成立，我们能证明它如下。假设 $g(x) = \sum_{d \geq 1} f(x/d)$ ，则

$$\begin{aligned} \sum_{d \geq 1} \mu(d) g\left(\frac{x}{d}\right) &= \sum_{d \geq 1} \mu(d) \sum_{k \geq 1} f\left(\frac{x}{kd}\right) \\ &= \sum_{m \geq 1} f\left(\frac{x}{m}\right) \sum_{d, k \geq 1} \mu(d) [m = kd] \\ &= \sum_{m \geq 1} f\left(\frac{x}{m}\right) \sum_{d|m} \mu(d) = \sum_{m \geq 1} f\left(\frac{x}{m}\right) [m = 1] = f(x). \end{aligned}$$

另一方向的证明本质上是相同的。

现在我们能解 $\Phi(x)$ 的递归(4.60)：

$$\Phi(x) = \frac{1}{2} \sum_{d \geq 1} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \left\lfloor 1 + \frac{x}{d} \right\rfloor. \quad (4.62)$$

这总是一个有限和。例如，

$$\begin{aligned}\Phi(12) &= \frac{1}{2}(12 \cdot 13 - 6 \cdot 7 - 4 \cdot 5 + 0 - 2 \cdot 3 + 2 \cdot 3 \\ &\quad - 1 \cdot 2 + 0 + 0 + 1 \cdot 2 - 1 \cdot 2 + 0) \\ &= 78 - 21 - 10 - 3 + 3 - 1 + 1 - 1 = 46.\end{aligned}$$

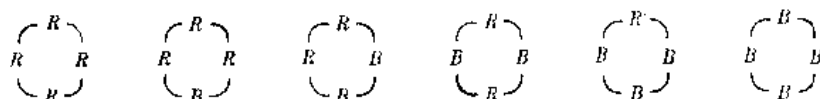
在第九章中我们将看到，如何用式(4.62)来取得 $\Phi(x)$ 的一个好的近似；事实上，我们将证明

$$\Phi(x) = \frac{3}{\pi^2} x^2 + O(x \log x).$$

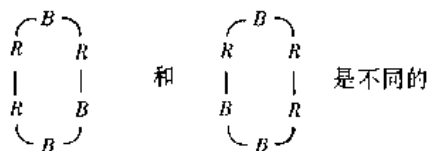
所以函数 $\Phi(x)$ “平滑地”增长，它最终平衡了 $\varphi(k)$ 的不稳定状态。

与上一章建立的传统相一致，让我们用一个问题来结束本章，这个问题说明了许多我们刚见的内容并也提前说明了下一章。假设我们有 n 种不同颜色的珠子，我们的目标是计算有多少种不同方式把它们串成长度为 m 的圆形项链。称可能的项链数为 $N(m, n)$ ，我们能尝试“命名和解决”这个问题。

例如，具有两种颜色的珠子 R 和 B ，我们能以 $N(4, 2) = 6$ 种不同的方式做出长为 4 的项链：



所有其他方式等价于这几种中的一种，因为一个项链的旋转并不改变它。然而考虑反射是不同的，在 $m = 6$ 的情形中，例如，



在 1892 年 P.A. MacMahon 首先解决了计算这些配置的问题^[212]。

没有 $N(m, n)$ 的明显递归式，但是能够通过 m 种方式把每个项链切断为线性串且考虑产生的生成物。例如，当 $m = 4$ 和 $n = 2$ ，我们取得

<i>RRRR</i>	<i>RRRR</i>	<i>RRRR</i>	<i>RRRR</i>
<i>RRBR</i>	<i>RRRB</i>	<i>BRRR</i>	<i>RBRR</i>
<i>RBBR</i>	<i>RKBB</i>	<i>BRRB</i>	<i>BBRR</i>
<i>RBRB</i>	<i>BRBR</i>	<i>KBRB</i>	<i>BRBR</i>
<i>RBBB</i>	<i>BRBB</i>	<i>BBRB</i>	<i>BBBB</i>
<i>BBBB</i>	<i>BBBB</i>	<i>BBBB</i>	<i>BBBB</i>

在 $mN(m, n)$ 个串的这种排列中, n^m 种可能型式的每一个至少出现一次, 某些型式出现多于一次。型式 $a_0 \cdots a_{m-1}$ 出现多少次? 这是容易的: 这就是循环转移 $a_k \cdots a_{m-1} a_0 \cdots a_{k-1}$ 产生和原来的 $a_0 \cdots a_{m-1}$ 相同型式的个数。例如, $BRBR$ 出现 2 次, 因为 4 种方式切断 $BRBR$ 形成的项链产生 4 种循环转移 ($BRBR, RBRB, BRBR, RBRB$); 4 个中的 2 个与 $BRBR$ 一致。这个论证表明

$$\begin{aligned} mN(m, n) &= \sum_{a_0, \dots, a_{m-1} \in S_n} \sum_{0 \leq k < m} [a_0 \cdots a_{m-1} = a_k \cdots a_{m-1} a_0 \cdots a_{k-1}] \\ &= \sum_{0 \leq k < m} \sum_{a_0, \dots, a_{m-1} \in S_n} [a_0 \cdots a_{m-1} = a_k \cdots a_{m-1} a_0 \cdots a_{k-1}] \end{aligned}$$

这里 S_n 是 n 种不同颜色的集合。

让我们看当给定 k 时, 有多少种型式满足 $a_0 \cdots a_{m-1} = a_k \cdots a_{m-1} a_0 \cdots a_{k-1}$ 。例如, 如果 $m=12$ 和 $k=8$, 则要计算

$$a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 a_{10} a_{11} = a_8 a_9 a_{10} a_{11} a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7$$

的解的个数。这意味着 $a_0 = a_8 = a_4$, $a_1 = a_9 = a_5$, $a_2 = a_{10} = a_6$ 和 $a_3 = a_{11} = a_7$ 。所以能以 n^4 种方式选取 a_0, a_1, a_2 和 a_3 的值, 且剩下的 a 依赖于它们。这看来熟悉吗? 一般,

$$a_j = a_{(j+k) \bmod m}, \quad \text{对 } 0 \leq j < m$$

的解使 a_j 和 $a_{(j+kl) \bmod m}$ (对 $l=1, 2, \dots$) 相等; 而我们知道模数 m 的 k 的倍数是 $\{0, d, 2d, \dots, m-d\}$, 其中 $d = \gcd(k, m)$ 。所以一般的解是独立选取 a_0, \dots, a_{d-1} , 然后置 $a_j = a_{j-d}$ ($d \leq j < m$), 共有 n^d 个解。

我们刚好证明了

$$mN(m, n) = \sum_{0 \leq k < m} n^{\gcd(k, m)}.$$

由于它仅包含项 n^d , 其中 $d \mid m$, 所以能简化此和。代入 $d = \gcd(k, m)$ 产生

$$\begin{aligned} N(m, n) &= \frac{1}{m} \sum_{d \mid m} n^d \sum_{0 \leq k < m} [d = \gcd(k, m)] \\ &= \frac{1}{m} \sum_{d \mid m} n^d \sum_{0 \leq k < m} [k/d \perp m/d] \\ &= \frac{1}{m} \sum_{d \mid m} n^d \sum_{0 \leq k < m/d} [k \perp m/d]. \end{aligned}$$

(允许用 k 替换 k/d , 因为 k 一定是 d 的倍数。)最后, 我们依据定义得到 $\sum_{0 \leq k < m/d} [k \perp m/d] = \varphi(m/d)$, 所以我们得到 MacMahon 公式:

$$N(m, n) = \frac{1}{m} \sum_{d|m} n^d \varphi\left(\frac{m}{d}\right) = \frac{1}{m} \sum_{d|m} \varphi(d) n^{m/d}. \quad (4.63)$$

例如, 当 $m=4$ 和 $n=2$ 时, 项链数为 $(1/4)(1 \cdot 2^4 + 1 \cdot 2^2 + 2 \cdot 2^1) = 6$, 正如我们猜想的那样。

由 MacMahon 和定义的值 $N(m, n)$ 是一个整数并不是明显的! 让我们试直接证明

$$\sum_{d|m} \varphi(d) n^{m/d} \equiv 0 \pmod{m}, \quad (4.64)$$

而不用与项链有关的提示。在 m 是素数的特殊情形中, 这同余式化为 $n^p + (p-1)n \equiv 0 \pmod{p}$; 也就是说, 它化为 $n^p \equiv n$ 。在式(4.48)中我们已看到, 此同余式是 Fermat 定理的另一种形式。所以当 $m=p$ 时式(4.64)成立, 我们可把它视为 Fermat 定理的一种推广, 推广到模数不是素数的情形。(欧拉的推广式(4.50)是不同的。)

对于所有素数模数已证明了式(4.64), 所以让我们查看剩下的最小情形 $m=4$ 。我们一定要证明

$$n^4 + n^2 + 2n \equiv 0 \pmod{4}.$$

如果我们分开考虑偶数和奇数情形, 证明是容易的。如果 n 是偶数, 左边的所有三项是对模数 4 的 0 同余, 所以它们的和也是这样。如果 n 是奇数, n^4 和 n^2 的每一个对 1 同余, $2n$ 是对 2 同余; 因此左边对 $1+1+2$ 同余, 对模数 4 的 0 同余, 我们完成了证明。

接着, 让我们大胆一点来试 $m=12$ 。 m 的这个值应是有趣的, 因为它有许多因子, 包含一个素数的平方, 然而它相当小。(还有好的机会, 我们将能把对 12 的证明推广到对一般 m 的证明。)我们一定要证明的同余式是

$$n^{12} + n^6 + 2n^4 + 2n^3 + 2n^2 + 4n \equiv 0 \pmod{12}.$$

现在该怎么办? 依据式(4.42), 此同余式成立当且仅当它对模数 3 和模数 4 也成立。所以让我们证明它对模数 3 成立。对素数, 同余式(4.64)成立, 所以我们有 $n^3 + 2n \equiv 0 \pmod{3}$ 。仔细研究显示, 我们能这个事实来归并较大和的项:

$$\begin{aligned} n^{12} + n^6 + 2n^4 + 2n^3 + 2n^2 + 4n \\ &= (n^{12} + 2n^4) + (n^6 + 2n^2) + 2(n^3 + 2n) \\ &= 0 + 0 + 2 \cdot 0 \equiv 0 \pmod{3}. \end{aligned}$$

所以对模数 3 它行得通。

我们完成了一半。为了证明模数 4 同余, 采用相同的诀窍。我们已证明了 $n^4 + n^2 + 2n \equiv 0 \pmod{4}$, 所以用此型式来归类:

$$\begin{aligned} n^{12} + n^6 + 2n^4 + 2n^3 + 2n^2 + 4n \\ &= (n^{12} + n^6 + 2n^4) + 2(n^3 + n^2 + 2n) \\ &= 0 + 2 \cdot 0 \equiv 0 \pmod{4}. \end{aligned}$$

对 $m=12$, 证毕.

到现在为止, 我们对于素数 m , 对于 $m=4$ 和 $m=12$, 证明了同余式. 现在让我们对素数幂来试证它. 为了具体起见, 可假设对于某个素数 p , $m=p^3$. 则式(4.64)的左边是

$$\begin{aligned} n^{p^3} + \varphi(p)n^{p^2} + \varphi(p^2)n^p + \varphi(p^3)n \\ = n^{p^3} + (p-1)n^{p^2} + (p^2-p)n^p + (p^3-p^2)n \\ = (n^{p^3} - n^{p^2}) + p(n^{p^2} - n^p) + p^2(n^p - n) + p^3n. \end{aligned}$$

如果我们能证明 p^3 可除尽 $n^{p^3} - n^{p^2}$, p^2 可除尽 $n^{p^2} - n^p$, p 可除尽 $n^p - n$, 则能证明这是对于模数 p^3 的 0 同余, 因为整个式将被 p^3 除尽. 依据 Fermat 定理的另一种形式, 我们有 $n^p \equiv n \pmod{p}$, 所以 p 除尽 $n^p - n$; 因此有一个整数 q 使得

$$n^p = n + pq.$$

现在我们把两边增加 p 幂, 按照二项定理展开右边(第五章中将见到), 再归并, 给出

$$\begin{aligned} n^{p^2} &= (n + pq)^p = n^p + (pq)^1 n^{p-1} \binom{p}{1} + (pq)^2 n^{p-2} \binom{p}{2} + \cdots \\ &= n^p + p^2 Q \end{aligned}$$

(对于某个其他整数 Q). 这里我们能得出 p^2 的一个因子, 因为在第二项中 $\binom{p}{1} = p$, 且因为跟着的所有项中出现 $(pq)^2$ 的一个因子, 所以我们发现 p^2 除尽 $n^{p^2} - n^p$.

我们把两边再增加 p 幂, 展开, 再归并, 取得

$$\begin{aligned} n^{p^3} &= (n^p + p^2 Q)^p = n^{p^2} + (p^2 Q)^1 n^{p(p-1)} \binom{p}{1} + (p^2 Q)^2 n^{p(p-2)} \binom{p}{2} + \cdots \\ &= n^{p^2} + p^3 Q \end{aligned}$$

(对另一个整数 Q). 所以 p^3 除尽 $n^{p^3} - n^{p^2}$. 这就结束了 $m=p^3$ 的证明, 因为我们已证出 p^3 除尽式(4.64)的左边.

并且我们用归纳法能证明

$$n^{p^k} = n^{p^{k-1}} + p^k Q$$

(对某个最后的整数 Q); 因此

$$n^{p^k} \equiv n^{p^{k-1}} \pmod{p^k}, \quad \text{对 } k > 0. \quad (4.65)$$

因此 p^k 可除尽式(4.64)的左边

$$(n^{p^k} - n^{p^{k-1}}) + p(n^{p^{k-1}} - n^{p^{k-2}}) + \cdots + p^{k-1}(n^p - n) + p^k n,$$

所以它对于模数 p_k 为 0 同余.

现在我们对于素数幂已证明了式(4.64), 所以仅需对 $m=m_1 m_2$, 其中 $m_1 \perp m_2$, 假

设对 m_1 和 m_2 , 同余式是真的, 来证明它。在研究 $m=12$ 情形时, 因子分解成 $m=3$ 和 $m=4$ 的实例, 这使我们想到这种方式行得通。

我们知道 φ 函数是积性的, 所以我们能写成

$$\begin{aligned}\sum_{d|m} \varphi(d) n^{m/d} &= \sum_{d_1|m_1, d_2|m_2} \varphi(d_1 d_2) n^{m_1 m_2 / d_1 d_2} \\ &= \sum_{d_1|m_1} \varphi(d_1) \left(\sum_{d_2|m_2} \varphi(d_2) (n^{m_1/d_1})^{m_2/d_2} \right).\end{aligned}$$

但是内部的和是对模数 m_2 的 0 同余, 因为我们假设对于 m_2 , 式(4.64)成立; 所以整个和是对模数 m_2 的 0 同余。对称的论证, 我们发现整个和同样是对模数 m_1 的 0 同余。因此依据式(4.42), 它是对模数 m 的 0 同余, 证毕。

习 题

准备部分

1. 对于 $1 \leq k \leq 6$, 恰好有 k 个因子的最小正整数是什么?
2. 证明 $\gcd(m, n) \cdot \text{lcm}(m, n) = m \cdot n$, 当 $n \bmod m \neq 0$ 时, 用此等式依据 $\text{lcm}(n \bmod m, m)$ 表达 $\text{lcm}(m, n)$ 。提示: 用公式(4.12), (4.14)和(4.15)。
3. 让 $\pi(x)$ 是不超过 x 的素数的个数。证明或推翻

$$\pi(x) - \pi(x-1) = [x \text{ 是素数}].$$

4. 如果 Stern-Brocot 构造从 5 个分数 $(0/1, 1/0, 0/-1, -1/0, 0/1)$ 开始, 而不是从 $(0/1, 1/0)$ 开始, 则出现什么?
5. 当 L 和 R 是式(4.33)的 2×2 矩阵时, 求 L^k 和 R^k 的简单公式。
6. ' $a \equiv b \pmod{0}$ ' 意味着什么?
7. 把编号为 1 到 10 的 10 个人排列成 Josephus 问题中的一个圆, 且每第 m 个人将被处死。(m 的值可能比 10 大很多。)证明对于任何 k , 前 3 个人离去不能是 10, k 和 $k+1$ (以此次序)。
8. 书中考虑的剩余数系 $(x \bmod 3, x \bmod 5)$ 有一个好奇的性质, 即 13 对应于 $(1, 3)$ 看来几乎相同。说明不计算所有 15 对剩余, 如何找出这样一种一致的所有实例。换句话说, 找同余式

$$10x + y \equiv x \pmod{3}, \quad 10x + y \equiv y \pmod{5}$$

的所有解。提示: 用 $10u + 6v \equiv u \pmod{3}$ 和 $10u + 6v \equiv v \pmod{5}$ 。

9. 证明 $(3^{77} - 1)/2$ 是奇数和合成数。提示: $3^{77} \bmod 4$ 是什么?
10. 计算 $\varphi(999)$ 。

11. 求具有性质

$$g(n) = \sum_{0 \leq k \leq n} f(k) \Leftrightarrow f(n) = \sum_{0 \leq k \leq n} \sigma(k)g(n-k)$$

的一个函数 $\sigma(n)$ 。(这相似于 Möbius 函数, 见式(4.56).)

12. 简化公式 $\sum_{d|m} \sum_{k|d} \mu(k)g(d/k)$.

13. 如果对于任何 $m > 1$, m^2 不可除尽一个正整数 n , 则称 n 为无平方的。找出 n 是无平方的一个充分和必要的条件。

(a) 利用 n 的素数指数表示式(4.11);

(b) 利用 $\mu(n)$ 。

基本部分

14. 证明或推翻

(a) $\gcd(km, kn) = k \gcd(m, n)$;

(b) $\text{lcm}(km, kn) = k \text{lcm}(m, n)$ 。

15. 每个素数呈现为某个欧几里德数 e_n 的一个因子吗?

16. 前 n 个欧几里德数的倒数之和是什么?

17. 设 f_n 是“Fermat 数” $2^{2^n} + 1$ 。证明 $f_m \perp f_n$, 如果 $m < n$ 。

18. 证明如果 $2^n + 1$ 是素数, 则 n 是 2 的一个幂。

19. 对于每一个正整数 n , 有一个素数 p 使得 $n < p \leq 2n$ 。(这本质上是“Bertrand 的假定”, 在 1845 年 Joseph Bertrand 对 $n < 3\,000\,000$ 证实了这一点, 在 1850 年 Chebyshev 对所有 n 证明了这一点。)用 Bertrand 的假定来证明有一个常数 $b \approx 1.25$ 使得数

$$\lfloor 2^b \rfloor, \lfloor 2^{2^b} \rfloor, \lfloor 2^{2^{2^b}} \rfloor, \dots$$

全为素数。

20. 设 p_n 是第 n 个素数。求出一个常数 K 使得

$$\lfloor (10^{n^2} K) \bmod 10^n \rfloor = p_n.$$

21. 证明下列等式(n 是正整数):

$$\begin{aligned} \sum_{1 \leq k \leq n} \left\lfloor \frac{\varphi(k+1)}{k} \right\rfloor &= \sum_{1 \leq m \leq n} \left[\left(\sum_{1 \leq k \leq m} \left\lfloor \frac{(m/k)}{\lfloor m/k \rfloor} \right\rfloor \right)^{-1} \right] \\ &= n - 1 - \sum_{k=1}^n \left\lfloor \left\{ \frac{(k-1)! + 1}{k} \right\} \right\rfloor. \end{aligned}$$

提示: 这是一个技巧问题, 解答是相当容易的。

22. 数 $1111\,111\,111\,111\,111\,111$ 是素数。证明任何基数 b , $(11\cdots 1)_b$ 能为素数仅当 1 的个数是素数。

23. 确定 $\rho(k)$ 的一个递归, $\rho(k)$ 是讨论 $\varepsilon_2(n!)$ 中的直尺函数. 表明 $\rho(k)$ 和当用 $2^k - 1$ 次移动转换 n 个圆盘的 Hanoi 塔时在第 k 步 ($1 \leq k \leq 2^k - 1$) 移动的圆盘之间有一个联系.

24. 利用 $v_p(n)$ 表达 $\varepsilon_p(n!)$, $v_p(n)$ 是 n 的基数 p 表示中数字位的和, 从而推广了式 (4.24).

25. 如果 $m \setminus n$ 和 $m \perp n/m$, 则称 m 恰好除尽 n , 记为 $m \setminus \setminus n$. 例如, 在阶乘因子讨论中 $p^{e_p(n)} \setminus \setminus n!$. 证明或推翻下列结论:

(a) $k \setminus \setminus n$ 和 $m \setminus \setminus n \Leftrightarrow km \setminus \setminus n$, 如果 $k \perp m$.

(b) 对于所有 $m, n > 0$, $\gcd(m, n) \setminus \setminus m$ 或者 $\gcd(m, n) \setminus \setminus n$.

26. 考虑所有非负简约分数 m/n , 使得 $mn \leq N$ 的序列 g_N , 例如,

$$\mathcal{G}_{10} = \frac{0}{1}, \frac{1}{10}, \frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}, \frac{3}{2}, \frac{2}{1}, \frac{5}{2}, \frac{3}{1}, \frac{4}{1}, \\ \frac{5}{1}, \frac{6}{1}, \frac{7}{1}, \frac{8}{1}, \frac{9}{1}, \frac{10}{1}.$$

每当在 \mathcal{G}_N 中 m/n 是紧靠 m'/n' 前面的分数时, $m'n - mn' = 1$ 是真的吗?

27. 基于 Stern-Brocot 数系中有理数表示为 L 和 R , 给出比较有理数的一个简单规则.

28. π 的 Stern-Brocot 表示是

$$\pi = R^3 L^7 R^{15} L R^{292} L R L R^2 L R^3 L R^{14} L^2 R \dots,$$

用它来求 π 的所有最简有理近似, 其分母小于 50. $22/7$ 是它们中的一个吗?

29. 书中描述了 $[0, 1)$ 中的二进制实数 $x = (b_1 b_2 b_3 \dots)_2$ 和 $[0, \infty)$ 中的 Stern-Brocot 实数 $\alpha = B_1 B_2 B_3 \dots$ 之间的一种对应. 如果 x 对应于 α , 且 $x \neq 0$, 什么数对应于 $1-x$?

30. 证明下列命题(中国剩余定理): 设 m_1, \dots, m_r 是具有 $m_i \perp m_k$ ($1 \leq j < k \leq r$) 的整数: 设 $m = m_1 \dots m_r$, 且设 a_1, \dots, a_r, A 是整数. 则恰好有一个整数 a 使得

$$a \equiv a_k \pmod{m_k} \quad (1 \leq k \leq r) \quad \text{和} \quad A \leq a < A + m.$$

31. 3 可除尽十进制表示的一个数当且仅当 3 可除尽它的数字位的和. 证明这个出名的规则, 且推广它.

32. 推广式 (4.47) 的证明来证欧拉定理式 (4.50).

33. 证明如果 $f(m)$ 和 $g(m)$ 是积性函数, 则 $h(m) = \sum_{d|m} f(d) g(m/d)$ 也是积性的.

34. 证明式 (4.56) 是式 (4.61) 的特殊情形.

课外习题

35. 当 m 和 n 是 $m \neq n$ 的非负整数时, 设 $I(m, n)$ 是一个函数满足关系

$$I(m, n)m + I(n, m)n = \gcd(m, n).$$

因此在式 (4.5) 中 $I(m, n) = m'$ 和 $I(n, m) = n'$; $I(m, n)$ 的值是 m 关于 n 的逆. 找出确定 $I(m, n)$ 的一个递归.

36. 考虑集合 $Z(\sqrt{10}) = \{m + n\sqrt{10} \mid \text{整数 } m, n\}$. 如果 $m^2 - 10n^2 = \pm 1$, 则称数

$m+n\sqrt{10}$ 是一个单位, 因为它有一个逆(也就是说, 由于 $(m+n\sqrt{10}) \cdot \pm(m-n\sqrt{10}) = 1$). 例如, $3+\sqrt{10}$ 是一个单位, 所以 $19-6\sqrt{10}$ 也是. 相消单位对能被插入任何因子分解, 所以我们忽略它们. $Z(\sqrt{10})$ 的非单位数如果不能被写成两个非单位的乘积, 则称为素数. 证明 2, 3 和 $4 \pm \sqrt{10}$ 是 $Z(\sqrt{10})$ 的素数. 提示: 如果 $2 = (k+l\sqrt{10}) \times (m+n\sqrt{10})$ 则 $4 = (k^2-10l^2)(m^2-10n^2)$. 此外, 任何整数 mod 10 的平方为 0, 1, 4, 5, 6, 或 9.

37. 证明式 (4.17). 提示: 证明 $e_n - (1/2) = (e_{n-1} - (1/2))^2 + 1/4$, 且考虑 $2^{-n} \cdot \log(e_n - (1/2))$.

38. 证明如果 $a \perp b$ 且 $a < b$, 则

$$\gcd(a^m - b^m, a^n - b^n) = a^{\gcd(m, n)} - b^{\gcd(m, n)}, \quad 0 \leq m < n.$$

(所有变量是整数.) 提示: 用欧几里德算法.

39. 设 $S(m)$ 是最小正整数 n , 对于它存在一个整数的上升序列

$$m = a_1 < a_2 < \cdots < a_t = n$$

使得 $a_1 a_2 \cdots a_t$ 是一个完全平方. (如果 m 是一个完全平方, 我们可设 $t=1$ 和 $n=m$.) 例如, $S(2)=6$, 因为最好的这样的序列是 $2 \cdot 3 \cdot 6$. 我们得到

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

证明每当 $0 < m < m'$ 时 $S(m) \neq S(m')$.

40. 如果 n 的基数 p 表示是 $(a_m \cdots a_1 a_0)_p$, 证明

$$\frac{n!}{p^{\varepsilon_p(n)}} \equiv (-1)^{\varepsilon_p(n)} a_m! \cdots a_1! a_0! \pmod{p}.$$

(左边移去所有 p 因子简单地为 $n!$. 当 $n=p$ 时, 此式化成 Wilson 定理.)

41. (a) 证明如果 $p \bmod 4 = 3$, 无整数 n 使得 p 除尽 n^2+1 . 提示: 用 Fermat 定理.

(b) 但是如果 $p \bmod 4 = 1$, 证明有这样一个整数.

提示: 把 $(p-1)!$ 写成 $\left(\prod_{k=1}^{(p-1)/2} k(p-k) \right)$, 且想到 Wilson 定理.

42. 在最小项中考虑两个分数 m/n 和 m'/n' . 证明当和 $m/n + m'/n'$ 被化为最小项时, 分母将为 nn' 当且仅当 $n \perp n'$. (换句话说, $(mn' + m'n)/nn'$ 已在最小项中当且仅当 n 和 n' 无公因子.)

43. 在 Stern-Brocot 树的 k 级处有 2^k 个结点, 对应于矩阵 $L^k, L^{k-1}R, \dots, R^k$, 证明从 L^k 开始, 然后相继乘

$$\begin{pmatrix} 0 & -1 \\ 1 & 2\rho(n)+1 \end{pmatrix} \quad 1 \leq n < 2^k, \text{ 其中 } \rho(n) \text{ 是直尺函数}$$

能得到这个序列。

44. 证明一个棒球手的平均成功率是.316, 则他一定至少击球 19 次。(如果他击球 n 次中击中 m 次, 则 $.3155 \leq m/n < .3165$ 。)

45. 数 9 376 有独特的自复制性质

$$9\,376^2 = 87\,909\,376.$$

有多少个 4 个数字位的数 x 满足方程 $x^2 \bmod 10\,000 = x$? 有多少个 n 个数字位的数 x 满足方程 $x^2 \bmod 10^n = x$?

46. (a) 证明如果 $n^i \equiv 1$ 和 $n^k \equiv 1 \pmod{m}$, 则 $n^{\gcd(i, k)} \equiv 1$ 。

(b) 证明如果 $n > 1$, $2^n \not\equiv 1 \pmod{n}$ 。提示: 考虑 n 的最小素因子。

47. 证明如果 $n^{m-1} \equiv 1 \pmod{m}$, 且如果对所有素数 $n^{(m-1)/p} \not\equiv 1 \pmod{m}$ 使得 $p \mid (m-1)$, 则 m 是素数。提示: 证明如果此条件成立, 数 $n^k \bmod m$ 是不同的 ($1 \leq k < m$)。

48. 通过确定表达式 $(\prod_{1 \leq n < m, n \perp m} n) \bmod m$ ($m > 1$) 的值来推广 Wilson 定理式 (4.49)。

49. 设 $R(N)$ 是整数对 (m, n) 的个数使得 $0 \leq m < N$, $0 \leq n < N$, 且 $m \perp n$ 。

(a). 利用 Φ 函数表达 $R(N)$ 。

$$(b) \text{ 证明 } R(N) = \sum_{d \geq N} \left\lfloor \frac{N}{d} \right\rfloor^2 \mu(d).$$

50. 设 m 是一个正整数, 且设

$$\omega = e^{2\pi i/m} = \cos\left(\frac{2\pi}{m}\right) + i \sin\left(\frac{2\pi}{m}\right).$$

我们称 ω 是单位的第 m 个根, 由于 $\omega^m = e^{2\pi i} = 1$ 。事实上, m 个复数 $\omega^0, \omega^1, \dots, \omega^{m-1}$ 的每一个是单位的第 m 个根, 因为 $(\omega^k)^m = e^{2\pi ki} = 1$, 所以 $z - \omega^k$ 是多项式 $z^m - 1$ ($0 \leq k < m$) 的一个因子。由于这些因子是不同的, 则复数上的 $z^m - 1$ 的完全因子分解一定是

$$z^m - 1 = \prod_{0 \leq k < m} (z - \omega^k).$$

(a) 设 $\Psi_m(z) = \prod_{0 \leq k < m, k \perp m} (z - \omega^k)$ 。(次数 $\phi(m)$ 的这个多项式称为 m 阶割圆多项式。) 证明

$$z^m - 1 = \prod_{d \mid m} \Psi_d(z).$$

$$(b) \text{ 证明 } \Psi_m(z) = \prod_{d \mid m} (z^d - 1)^{\mu(m/d)}.$$

考查性问题

51. 用多项式定理展开 $(1+1+\cdots+1)^p$ 来证明 Fermat 定理式(4.48).

52. 设 n 和 x 是正整数, 使得 x 没有因子 $\leq n$ (除 1 外), 且设 p 是素数, 证明数 $\{x-1, x^2-1, \cdots, x^{n-1}-1\}$ 中至少 $\lfloor n/p \rfloor$ 个是 p 的倍数.

53. 求出所有正整数 n 使得 $n \mid \lfloor (n-1)! / (n+1) \rfloor$.

54. 用手工计算来确定 $1000! \bmod 10^{250}$ 的值.

55. 设 P_n 是前 n 个阶乘的乘积 $\prod_{k=1}^n k!$. 证明 P_{2n} / P_n^4 是一个整数 (对于所有正整数 n).

56. 证明

$$\left(\prod_{k=1}^{2n-1} k^{\min(k, 2n-k)} \right) / \left(\prod_{k=1}^{n-1} (2k+1)^{2n-2k-1} \right)$$

是一个 2 的幂.

57. 设 $S(m, n)$ 是所有整数 k 使得

$$m \bmod k + n \bmod k \geq k$$

的集合. 例如, $S(7, 9) = \{2, 4, 5, 8, 10, 11, 12, 13, 14, 15, 16\}$. 证明

$$\sum_{k \in S(m, n)} \varphi(k) = mn.$$

提示: 首先证明 $\sum_{1 \leq m \leq n} \sum_{d|m} \varphi(d) = \sum_{d>1} \varphi(d) \lfloor n/d \rfloor$. 然后考虑 $\lfloor (m+n)/d \rfloor - \lfloor m/d \rfloor - \lfloor n/d \rfloor$.

58. 设 $f(m) = \sum_{d|m} d$. 找出 $f(m)$ 是一个 2 的幂的充分必要条件.

额外问题

59. 证明如果 x_1, \cdots, x_n 是具有 $1/x_1 + \cdots + 1/x_n = 1$ 的正整数, 则 $\max(x_1, \cdots, x_n) < e_n$. 提示: 用归纳法证明下列较强的结果: “如果 $1/x_1 + \cdots + 1/x_n + 1/\alpha = 1$, 其中 x_1, \cdots, x_n 是正整数, 且 α 是一个有理数 $\geq \max(x_1, \cdots, x_n)$, 则 $\alpha+1 \leq e_{n+1}$ 和 $x_1 \cdots x_n (\alpha+1) \leq e_1 \cdots e_n e_{n+1}$ ”. (证明是不平凡的.)

60. 证明存在一个常数 P 使得式(4.18)仅给出素数. 你可用下列(很不平凡的)事实: 对于某个常数 c 和所有充分大的 p , p 和 $p+cp^\theta$ 之间有一个素数, 其中 $\theta = 1051/1920$.

61. 证明如果 m/n , m'/n' 和 m''/n'' 是 \mathcal{F}_N 的相继元素, 则

$$m'' = \left\lfloor \frac{n+N}{n'} \right\rfloor m' - m,$$

$$n'' = \left\lfloor \frac{n+N}{n'} \right\rfloor n' - n.$$

(这个递归允许我们从 $0/1$ 和 $1/N$ 开始, 依次计算出 \mathcal{P}_N 的元素.)

62. 在二进制 \longleftrightarrow Stern-Brocot 对应中, 对应于 e 的二进制数是什么? (把解答表示为一个无限和, 不必求出它的闭形式.)

63. 证明如果 Fermat 最后定理式(4.46)是不真实的, 则使它不真实的最小 n 是素数. (你可假设当 $n=4$ 时结果成立.) 此外, 如果 $a^p + b^p = c^p$ 和 $a \perp b$, 证明存在一个整数 m 使得

$$a + b = \begin{cases} m^p, & \text{如果 } p \nmid c, \\ p^{p-1} m^p, & \text{如果 } p \mid c. \end{cases}$$

因此 c 一定是很大的. 提示: 设 $x = a + b$, 且注意到, $\gcd(x, (a^p + (x-a)^p)/x) = \gcd(x, pa^{p-1})$.

64. N 阶 Peirce 序列 \mathcal{P}_N 是用 ' $<$ ' 或 ' $=$ ' 符号分开的分数的无限串, 它包含所有非负分数 m/n ($m \geq 0$ 和 $n \leq N$, 包含未简约的分数). 从

$$\mathcal{P}_1 = \frac{0}{1} < \frac{1}{1} < \frac{2}{1} < \frac{3}{1} < \frac{4}{1} < \frac{5}{1} < \frac{6}{1} < \frac{7}{1} < \frac{8}{1} < \frac{9}{1} < \frac{10}{1} < \dots$$

开始递归地定义它. 对于 $N \geq 1$, 我们就在 \mathcal{P}_N 的第 kN 个符号前面插入两个符号(所有 $k > 0$)来形成 \mathcal{P}_{N+1} . 两个插入符号是

$$\frac{k-1}{N+1} \quad , \quad \text{如果 } kN \text{ 是奇数};$$

$$\mathcal{P}_{N, kN+1} = \frac{k-1}{N+1}, \quad \text{如果 } kN \text{ 是偶数}.$$

这里的 $\mathcal{P}_{N, j}$ 表示 \mathcal{P}_N 的第 j 个符号, 当 j 是偶数时, 它将是 ' $<$ ' 或者 ' $=$ '; 当 j 是奇数时, 它将是一个分数. 例如,

$$\mathcal{P}_2 = \frac{0}{2} = \frac{0}{1} < \frac{1}{2} < \frac{2}{2} = \frac{1}{1} < \frac{3}{2} < \frac{4}{2} = \frac{2}{1} < \frac{5}{2} < \frac{6}{2} = \frac{3}{1} < \frac{7}{2} < \frac{8}{2} = \frac{4}{1} < \frac{9}{2} < \frac{10}{2} = \frac{5}{1} < \dots;$$

$$\mathcal{P}_3 = \frac{0}{2} = \frac{0}{3} = \frac{0}{1} < \frac{1}{3} < \frac{1}{2} < \frac{2}{3} < \frac{2}{2} = \frac{3}{3} = \frac{1}{1} < \frac{4}{3} < \frac{3}{2} < \frac{5}{3} < \frac{4}{2} = \frac{6}{3} = \frac{2}{1} < \frac{7}{3} < \frac{5}{2} < \dots;$$

$$\mathcal{P}_4 = \frac{0}{2} = \frac{0}{4} = \frac{0}{3} = \frac{0}{1} < \frac{1}{4} < \frac{1}{3} < \frac{2}{4} = \frac{1}{2} < \frac{3}{4} < \frac{2}{3} = \frac{4}{4} = \frac{3}{3} = \frac{1}{1} < \frac{5}{4} < \frac{4}{3} < \frac{6}{4} = \dots;$$

$$\mathcal{P}_5 = \frac{0}{2} = \frac{0}{4} = \frac{0}{5} = \frac{0}{3} = \frac{0}{1} < \frac{1}{5} < \frac{1}{4} < \frac{1}{3} < \frac{2}{5} < \frac{2}{4} = \frac{1}{2} < \frac{3}{5} < \frac{2}{3} < \frac{4}{5} < \frac{3}{4} < \frac{5}{5} = \frac{4}{4} = \dots;$$

$$\mathcal{P}_6 = \frac{0}{2} = \frac{0}{4} = \frac{0}{6} = \frac{0}{5} = \frac{0}{3} = \frac{0}{1} < \frac{1}{6} < \frac{1}{5} < \frac{1}{4} < \frac{2}{6} = \frac{1}{3} < \frac{3}{6} < \frac{2}{5} = \frac{3}{5} = \frac{1}{2} < \frac{4}{6} < \frac{3}{4} < \frac{5}{6} = \dots.$$

(相等元素以稍微独特的次序出现.) 证明上述规则定义的 ' $<$ ' 和 ' $=$ ' 符号正确描述了 Peirce 序列中邻近分数之间的关系.

研究性问题

65. 是否所有欧几里德数 e_n 是无平方的?

66. 是否所有 Mersenne 数 $2^p - 1$ 是无平方的?

67. 证明或推翻对于所有整数序列 $0 < a_1 < \cdots < a_n$,

$$\max_{1 \leq j < k \leq n} a_k / \gcd(a_j, a_k) \geq n.$$

68. 是否有一个常数 Q 使得对所有 $n \geq 0$, $\lfloor Q^{2^n} \rfloor$ 是素数?

69. 设 P_n 表示第 n 个素数. 证明或推翻 $P_{n+1} - P_n = O(\log P_n)^2$.

70. 对于无限多 n , $\varepsilon_3(n!) = \varepsilon_2(n!) / 2$ 吗?

71. 证明或推翻: 如果 $k \neq 1$, 存在 $n > 1$ 使得 $2^n \equiv k \pmod{n}$. 是否有无限多个这样的 n ?

72. 证明或推翻: 对于所有整数 a , 存在无限多个 n 使得 $\varphi(n) \nmid (n+a)$.

73. 如果 Faery 序列

$$\mathcal{F}_n = \langle \mathcal{F}_n(0), \mathcal{F}_n(1), \dots, \mathcal{F}_n(\Phi(n)) \rangle$$

的 $\Phi(n)+1$ 项完全均匀分布, 我们将期望 $\mathcal{F}_n(k) \approx k / \Phi(n)$. 所以和 $D(n)$

$$= \sum_{k=0}^{\Phi(n)} |\mathcal{F}_n(k) - k / \Phi(n)| \text{ 测量“}\mathcal{F}_n\text{关于均匀性的偏差”}.$$

对于所有 $\varepsilon > 0$, $D(n) = O(n^{1/2+\varepsilon})$

是真的吗?

74. 当 $p \rightarrow \infty$ 时, 集合 $\{0! \bmod p, 1! \bmod p, \dots, (p-1)! \bmod p\}$ 中大约有多少个不同的值?

第五章 二项系数

让我们松一口气。前面几章已经看到了一些繁重的工作状况，涉及下整，上整，mod， φ 和 μ 函数的和。现在我们将研究二项系数，它在应用中相当重要，并且比所有其他的那些量容易运算。

5.1 基本等式

由于我们在本节后面将看到的二项定理，而称符号 $\binom{n}{k}$ 为二项系数，读作“ n 选取 k ”。此叫法来自它的组合解释，它是从 n 个元素的集合中选取 k 个元素的子集的方式数。例如，从集合 $\{1, 2, 3, 4\}$ 我们可有 6 种方式选取 2 个元素，

$\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}$;

所以 $\binom{4}{2} = 6$ 。

为了以较熟悉的项来表达数 $\binom{n}{k}$ ，最容易的是首先确定从 n 个元素的集合中选取 k 个元素序列的个数，而不是子集的个数；对于序列，考虑元素的次序。我们用第四章中所用的相同论证表明 $n!$ 是 n 个物体的排列数。序列的第一个元素有 n 种选择；对于每一个，第二个元素有 $n-1$ 种选择；等等，直到第 k 个元素有 $n-k+1$ 种选择。总之它给出 $n(n-1)\cdots(n-k+1) = n^{\overline{k}}$ 种选择。且由于 k 个元素的每个子集恰好有 $k!$ 种不同的次序，此序列数把每个子集恰好计算了 $k!$ 次。为了获得解答，我们简单地用 $k!$ 来除：

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots(1)}.$$

例如，

$$\binom{4}{2} = \frac{4 \cdot 3}{2 \cdot 1} = 6,$$

这和前面的枚举一致。

我们称 n 为上指数, k 为下指数。用组合说明指数限制为非负整数; 因为集合不会有负数个元素或分数个元素。但是除了它的组合说明之外二项系数还有许多用处, 所以我们将移去某些限制。最有用的是允许一个实 (甚至复) 数出现在上指数处, 允许一个任意整数出现在下指数处。所以形式的定义取下列形式:

$$\binom{r}{k} = \begin{cases} \frac{r(r-1)\cdots(r-k+1)}{k(k-1)\cdots(1)} = \frac{r^{\underline{k}}}{k!}, & \text{整数 } k \geq 0; \\ 0, & \text{整数 } k < 0. \end{cases} \quad (5.1)$$

此定义有几个值得注意的特点。首先, 称上指数为 r , 而不称 n , 字母 r 强调当任何实数出现在此位置处二项系数有意义。例如, 我们得到 $\binom{-1}{3} = (-1)(-2)(-3)/(3 \cdot 2 \cdot 1) = -1$ 。这里没有组合解释, 但是 $r = -1$ 原来是一种重要的特殊情形。像 $r = -1/2$ 的非整数指数也是有用的。

其次, 我们能把 $\binom{r}{k}$ 看作为 r 的 k 次多项式。我们将看到这种观点常常是有用的。

第三, 未就非整数下指数定义二项系数。我们可以给出一种合理的定义, 但实际应用是不常发生的, 所以把这种推广留在本章后面再讲。

最后的注意点: 我们在定义的右边列出了限制 ‘整数 $k \geq 0$ ’ 和 ‘整数 $k < 0$ ’。这样的限制将在研究的所有等式中列出, 以致可应用的范围将是清楚的。一般较少的限制更好, 因为一个无限制的等式最有用; 运用的任何限制还是等式的重要部分, 当我们运算二项系数时, 暂时忽略难记的限制且以后再检验不违背限制是较容易的, 但是必需做检验。

例如, 几乎每次遇到 $\binom{n}{n}$, 它等于 1, 所以我们会误认为它总是 1。但是仔细看一看定义式 (5.1), 告诉我们仅当 $n \geq 0$ 时 $\binom{n}{n}$ 是 1 (设 n 是整数); 当 $n < 0$ 时, 我们得到 $\binom{n}{n} = 0$ 。像这样的陷阱能 (且将) 引起使用期间冒险。

在接触到等式之前, 我们将用到平淡的二项系数, 让我们看一些小的值。表 5.1 中的数形成 Pascal 三角形的开始部分, 为了纪念 Blaise Pascal 而取此名 (1623~1662), 因为他写出了关于它们的一篇有影响的专题论文^[227]。表中的空的项目实际为 0, 因为式 (5.1) 的分子为零; 例如, $\binom{1}{2} = (1 \cdot 0)/(2 \cdot 1) = 0$ 。这些项目简单地留作空白有助于突出表的剩下部分。

值得记住前三列的公式,

$$\binom{r}{0} = 1, \quad \binom{r}{1} = r, \quad \binom{r}{2} = \frac{r(r-1)}{2}; \quad (5.2)$$

这些公式对于任意实数成立。(回想起 $\binom{n+1}{2} = \frac{1}{2}n(n+1)$ 是第一章中我们关于三角形数引出的公式, 三角形数明显地呈现在表 5.1 的 $\binom{n}{2}$ 列中。) 记住前五行或者这样的 Pascal

表 5.1 Pascal 三角形

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

三角形也是一个好主意，可使在某个问题中出现型式 1, 4, 6, 4, 1 时，我们将获得一个线索，即二项系数也许就潜藏在附近。实际上，Pascal 三角形中的数满足无限多个等式，所以仔细研究一下能找到一些出人意外的关系，这是不太奇怪的。例如，有一个好奇的“六边形性质”，通过表 5.1 的下右部分中围绕 84 的 6 个数 56, 28, 36, 120, 210, 126 来说明。由此六边形中交错的数相乘的二种方式给出相同的乘积： $56 \cdot 36 \cdot 210 = 28 \cdot 120 \cdot 126 = 423\,360$ 。如果从 Pascal 三角形的任何其他部分取出这样一个六边形，则相同的结果成立。

现在来讨论等式。本节的目的学习几个简单的规则，用这些规则我们能解大多数涉及二项系数的实际问题。

一般情形根据阶乘能改动定义式(5.1)，上指数 r 是整数 n ，它大于或等于下指数 k ：

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad \text{整数 } n \geq k \geq 0. \quad (5.3)$$

为了取得此公式，我们只要用 $(n-k)!$ 乘式(5.1)的分子和分母即可。把二项系数展开成这样的阶乘形式通常是有用的（例如，当证明六边形性质时），而且我们常常要把阶乘改变成二项系数。

阶乘表示暗示 Pascal 三角形中的一种对称性：每行从左到右读和从右到左读相同。把 k 改成 $n-k$ 得到反映对称的等式，称为对称等式：

$$\binom{n}{k} = \binom{n}{n-k}, \quad \text{整数 } n \geq 0, \quad \text{整数 } k. \quad (5.4)$$

此公式有组合意义，因为从 n 中选出指定的 k 个事物就相当于指定 $n-k$ 个不选取的事

物。

在等式(5.4)中限制 n 和 k 是整数是明显的, 因为每个下指数一定是一个整数。但是为什么 n 不能是负的? 例如, 假设 $n = -1$, 是否

$$\binom{-1}{k} \stackrel{?}{=} \binom{-1}{-1-k}$$

是成立的方程? 不是。例如, 当 $k=0$ 时, 我们在左边得到 1, 而在右边得到 0。事实上, 对于任何整数 $k \geq 0$, 左边是

$$\binom{-1}{k} = \frac{(-1)(-2)\cdots(-k)}{k!} = (-1)^k,$$

它或者是 1 或者是 -1; 但是右边是 0, 因为下指数是负的。且对于负的 k , 左边为 0, 但右边是

$$\binom{-1}{-1-k} = (-1)^{-1-k},$$

它或者是 1 或者是 -1, 所以方程 $\binom{-1}{k} = \binom{-1}{-1-k}$ 总不成立!

对于所有其他负整数 n , 对称等式也不成立, 但是不幸的是, 最容易忘记这个限制, 因为上指数中的表示有时仅对它的变量的不分明 (但合法) 的值为负的。许多运算二项系数的人至少陷入此陷阱三次。

但是对称等式有一个重要的补救的特性: 对于 k 的所有值, 甚至当 $k < 0$ 或 $k > n$ 它成立。(因为在这些情形两边为零。) 否则 $0 \leq k \leq n$, 根据式(5.3)立即得到对称性:

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

下一个重要的等式让我们作移进或移出二项系数:

$$\binom{r}{k} = \frac{r}{k} \binom{r-1}{k-1}, \text{ 整数 } k \neq 0. \quad (5.5)$$

这里限制 k 不让用 0 除。我们称式(5.5)为吸收等式, 因为当变量在外面是妨害时, 我们用此等式把变量吸收进一个二项系数。因为 $r^{\frac{k}{k}} = r(r-1)^{\frac{k-1}{k-1}}$ 和 $k! = k(k-1)!$ (当 $k > 0$ 时), 根据定义式(5.1)得到方程, 当 $k < 0$ 时两边为零。

若我们在式(5.5)的两边乘 k , 则可得到甚至当 $k=0$ 时吸收等式成立:

$$k \binom{r}{k} = r \binom{r-1}{k-1}, \text{ 整数 } k. \quad (5.6)$$

此等式还有一个下指数保持原样的相伴等式:

$$(r-k) \binom{r}{k} = r \binom{r-1}{k}, \text{ 整数 } k. \quad (5.7)$$

在两个对称性的应用中间加进式(5.6)的应用我们能引出式(5.7):

$$\begin{aligned}(r-k)\binom{r}{k} &= (r-k)\binom{r}{r-k} && \text{(由对称性)} \\ &= r\binom{r-1}{r-k-1} && \text{(由式(5.6))} \\ &= r\binom{r-1}{k}. && \text{(由对称性)}\end{aligned}$$

等一等。我们要求等式对所有实数 r 成立, 然而我们刚给的推导仅当 r 是正整数时成立。(如果我们泰然地用对称性质, 则上指数 $r-1$ 一定是非负整数。) 我们被骗了吗? 没有。的确推导仅对正整数 r 成立; 但是我们断定等式对于所有 r 的值成立, 因为式(5.7)的两边都是 r 的 $k+1$ 次多项式。一个非零的 d 次或低于 d 次的多项式至多能有 d 个不同的零点, 所以两个如此的多项式之差也是 d 次或低于 d , 它不能在多于 d 点处为零, 除非它恒等于零。换句话说, 如果两个 d 次或低于 d 次的多项式在多于 d 点处相同, 则它们一定完全相同。我们已证明了 $(r-k)\binom{r}{k} = r\binom{r-1}{k}$ (每当 r 是一个正整数), 所以这样两个多项式在无穷多点处相同, 且它们一定是恒等。

前段中的证明技巧对于从整数到实数推广许多等式是有用的, 我们称此技巧为多项式论证。我们将一次又一次地见到它。某些方程, 像对称等式(5.4), 不是多项式之间的等式, 所以我們不总能用此方法。但是许多等式有必需的形式。

例如, 这里是另一个多项式等式, 也许是最重要的二项等式, 称为加法公式:

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}, \text{ 整数 } k. \quad (5.8)$$

当 r 是正整数时, 加法公式告诉我们, Pascal 三角形中的每一个数是前一行中两个数的和, 一个数直接在它上面, 另一个是刚才左边的一个。当 r 为负的, 实数, 或复数也可应用; 仅限制 k 是整数, 以致定义了二项系数。

证明加法公式的一种方法是设 r 是正整数且用组合解释。 $\binom{r}{k}$ 就是从 r 个元素集合中选取可能的 k 个元素子集的个数。若我们有 r 个鸡蛋的集合, 恰好包含一个坏鸡蛋, 有 $\binom{r}{k}$ 种方式来选 k 个鸡蛋, 恰好 $\binom{r-1}{k}$ 种选取包含好鸡蛋, $\binom{r-1}{k-1}$ 种方式包含坏鸡蛋, 因为这样的选取为 $r-1$ 个好鸡蛋的 $k-1$ 个。把这样两个数加在一起给出式(5.8)。此推导假设 r 是正整数, 且 $k \geq 0$ 。但是当 $k < 0$ 时等式两边都是零, 多项式论证对所有剩下的情形建立了式(5.8)。

我们还能把两个吸收等式(5.7)和(5.6)加在一起引出式(5.8):

$$(r-k)\binom{r}{k} + k\binom{r}{k} = r\binom{r-1}{k} + r\binom{r-1}{k-1},$$

左边为 $r \binom{r}{k}$ ，且两边用 r 除。此推导除了 $r=0$ 外都成立，且容易检验剩下的情形。

有些人不注意发现这样的巧妙的证明，或者有些人用另一种方法，宁愿用定义的直接运算来引出式(5.8)。若 $k > 0$,

$$\begin{aligned} \binom{r-1}{k} + \binom{r-1}{k-1} &= \frac{(r-1)^{\underline{k}}}{k!} + \frac{(r-1)^{\underline{k-1}}}{(k-1)!} \\ &= \frac{(r-1)^{\underline{k-1}}(r-k)}{k!} + \frac{(r-1)^{\underline{k-1}}k}{k!} \\ &= \frac{(r-1)^{\underline{k-1}}r}{k!} = \frac{r^{\underline{k}}}{k!} = \binom{r}{k}. \end{aligned}$$

还有 $k \leq 0$ 的情形容易处理。

我们刚才已看到三种相当不同的加法公式的证明，这并不意外；二项系数有许多有用的性质，有几个性质不得不引导到现有的一个等式的证明。

加法公式本质上是 Pascal 三角形的数的一个递归，所以我们将看到用归纳法证明其他等式特别有用。我们还能通过展开递归立即得到一个新的等式。例如，

$$\begin{aligned} \binom{5}{3} &= \binom{4}{3} + \binom{4}{2} \\ &= \binom{4}{3} + \binom{3}{2} + \binom{3}{1} \\ &= \binom{4}{3} + \binom{3}{2} + \binom{2}{1} + \binom{2}{0} \\ &= \binom{4}{3} + \binom{3}{2} + \binom{2}{1} + \binom{1}{0} + \binom{1}{-1}. \end{aligned}$$

由于 $\binom{1}{-1} = 0$ ，此项不出现，我们停止展开。此方法产生一般公式

$$\begin{aligned} \sum_{k \leq n} \binom{r+k}{k} &= \binom{r}{0} + \binom{r+1}{1} + \cdots + \binom{r+n}{n} \\ &= \binom{r+n+1}{n}, \text{ 整数 } n. \end{aligned} \tag{5.9}$$

注意在求和的指数上我们不需要下限 $k \geq 0$ ，因为 $k < 0$ 的项为零。

此公式把一个二项系数表达为其他的二项系数的和，这些二项系数的上指数和下指数保持相同距离间隔。重复展开最小下指数的二项系数我们找到它：首先 $\binom{5}{3}$ ，然后 $\binom{4}{2}$ ，然后 $\binom{3}{1}$ ，然后 $\binom{2}{0}$ 。如果我们用另一种方式展开，即重复展开最大下指数的二项系数，将出现什么情形？我们取得

$$\begin{aligned}
 \binom{5}{3} &= \binom{4}{3} + \binom{4}{2} \\
 &= \binom{3}{3} + \binom{3}{2} + \binom{4}{2} \\
 &= \binom{2}{3} + \binom{2}{2} + \binom{3}{2} + \binom{4}{2} \\
 &= \binom{1}{3} + \binom{1}{2} + \binom{2}{2} + \binom{3}{2} + \binom{4}{2} \\
 &= \binom{0}{3} + \binom{0}{2} + \binom{1}{2} + \binom{2}{2} + \binom{3}{2} + \binom{4}{2}.
 \end{aligned}$$

现在 $\binom{0}{3}$ 为零(同样 $\binom{0}{2}$ 和 $\binom{1}{2}$ 为零, 但是这些项使等式较整齐), 我们能发现一般型式:

$$\begin{aligned}
 \sum_{0 \leq k \leq n} \binom{k}{m} &= \binom{0}{m} + \binom{1}{m} + \cdots + \binom{n}{m} \\
 &= \binom{n+1}{m+1}, \text{ 整数 } m, n \geq 0.
 \end{aligned} \tag{5.10}$$

我们把这个等式称为在上指数上求和, 它把一个二项系数表达为其他的二项系数的和, 这些二项系数的下指数是常量。这种情形中的和需下限 $k \geq 0$, 因为 $k < 0$ 的项不为零。一般 m 和 n 也不能是负的。

等式(5.10)有一种有趣的组合解释。如果我们要在编号0到 n 的 $n+1$ 张票的集合中选取 $m+1$ 张票, 当选取的最大票是数 k 时, 有 $\binom{k}{m}$ 种方式来实现。

我们能通过归纳法用加法公式来证明式(5.9)和(5.10), 但是我们也能互相证明它们。例如, 让我们由式(5.10)来证明式(5.9), 我们的证明将说明一些普通的二项系数的操作。我们的一般打算是处理式(5.9)的左边 $\sum \binom{r+k}{k}$, 以致它看上去像式(5.10)的左边 $\sum \binom{k}{m}$ 。于是我们调用由一个二项系数替换和的等式, 最后把此系数转换成式(5.9)的右边。

为了方便, 我们设 r 和 n 是非负整数, 由多项式论证, 根据此特殊情形得到式(5.9)的一般情形。让我们用 m 代替 r , 以致这个变量看来很像非负整数。现在能系统地实行方法如下:

$$\begin{aligned}
 \sum_{k \leq n} \binom{m+k}{k} &= \sum_{-m \leq k \leq n} \binom{m+k}{k} \\
 &= \sum_{-m \leq k \leq n} \binom{m+k}{m} \\
 &= \sum_{0 \leq k \leq m+n} \binom{k}{m}
 \end{aligned}$$

$$= \binom{m+n+1}{m+1} = \binom{m+n+1}{n}.$$

让我们详细地来看这里的推导。关键的一步是第二行，在那里我们应用对称法则式(5.4)把 $\binom{m+k}{k}$ 换成 $\binom{m+k}{m}$ 。由于仅当 $m+k \geq 0$ 时允许这样做，所以第一步除去 $k < -m$ 的项来限制 k 的范围。（这是合法的，因为那些项是零。）现在我们几乎准备应用式(5.10)，第三行建立了这一点，用 $k-m$ 替换 k ，且整理求和的范围。这一步像第一步一样，仅仅环绕 \sum 记号操作。现在 k 出现在上指数，且和的界限是适当的形式，所以第四行应用式(5.10)。再用一次对称性就完成了证明。

在第一章和第二章中得出的一些和实际上是式(5.10)的特殊情形，或者是此等式的不易识别的形式。例如，情形 $m=1$ 给出直到 n 的非负整数的和：

$$\binom{0}{1} + \binom{1}{1} + \cdots + \binom{n}{1} = 0 + 1 + \cdots + n = \frac{(n+1)n}{2} = \binom{n+1}{2}.$$

如果我们用 $m!$ 除此公式的两边，则一般情形等价于第二章的规则

$$\sum_{0 \leq k \leq n} k^m = \frac{(n+1)^{\overline{m+1}}}{m+1}, \text{ 整数 } m, n \geq 0.$$

事实上，加法公式(5.8)告诉我们，如果我们用 $x+1$ 和 m 分别替换 r 和 k ，

$$\Delta \left(\binom{x}{m} \right) = \binom{x+1}{m} - \binom{x}{m} = \binom{x}{m-1}.$$

因此第二章的方法给出方便的不定的求和公式

$$\sum \binom{x}{m} \delta x = \binom{x}{m+1} + C. \quad (5.11)$$

二项系数的名称是从涉及二项表达式 $x+y$ 的幂的二项定理获得的。让我们看此定理的最小的情形：

$$(x+y)^0 = 1x^0y^0$$

$$(x+y)^1 = 1x^1y^0 + 1x^0y^1$$

$$(x+y)^2 = 1x^2y^0 + 2x^1y^1 + 1x^0y^2$$

$$(x+y)^3 = 1x^3y^0 + 3x^2y^1 + 3x^1y^2 + 1x^0y^3$$

$$(x+y)^4 = 1x^4y^0 + 4x^3y^1 + 6x^2y^2 + 4x^1y^3 + 1x^0y^4$$

不难看出为什么这些系数和 Pascal 三角形中的数相同：当我们展开乘积

$$(x+y)^n = \underbrace{(x+y)(x+y)\cdots(x+y)}_{n \text{ 个因子}},$$

其每一项本身是 n 个因子的乘积, 每个因子或者是一个 x 或者是一个 y . k 个 x 因子和 $n-k$ 个 y 因子的这样的项的个数是合并这样的项后 $x^k y^{n-k}$ 的系数. 而这恰好是 n 个二项中选取 k 个提供的 x 的方式数; 也就是说, 它是 $\binom{n}{k}$.

有些教科书中把量 0^0 作为没有定义, 因为函数 x^0 和 0^x 当 x 下降到 0 时具有不同的极限值. 但是这是一种误解, 我们必须定义

$$x^0 = 1, \text{ 对于所有 } x,$$

如果二项定理当 $x=0, y=0$, 与 $/$ 或 $x=-y$ 时成立. 任意限定的定理是太重要了! 和预料相反, 函数 0^x 完全不重要.

但是确切的二项定理是什么? 它是下列等式:

$$(x+y)^r = \sum_k \binom{r}{k} x^k y^{r-k}, \quad \begin{array}{l} \text{整数 } r \geq 1, \\ \text{或 } |x/y| < 1. \end{array} \quad (5.12)$$

和是在所有整数 k 上求的, 但是当 r 是负整数时, 它确实是有限和, 因为除了 $0 \leq k \leq r$ 的那些项外所有的项为零. 另一方面, 当 r 是负, 或甚至当 r 是任意实数或复数时, 定理也成立. 在这样的情形中, 和确实为无限的, 且我们一定有 $|x/y| < 1$ 来保证和的绝对收敛.

两个二项定理的特殊情形是值得特别注意, 纵然它们很简单. 若 $x=y=1$ 和 $r=n$ 是非负, 则我们得到

$$2^n = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n}, \text{ 整数 } n \geq 0.$$

此方程告诉我们, Pascal 三角形的行 n 总和为 2^n . 当 x 为 -1 而不是 $+1$ 时, 我们得到

$$0^n = \binom{n}{0} - \binom{n}{1} + \cdots + (-1)^n \binom{n}{n}, \text{ 整数 } n \geq 0.$$

例如, $1-4+6-4+1=0$. 除了最高行(当 $n=0$ 和 $0^0=1$)外, 若我们对行 n 的元素轮流给正负号, 则其和为零.

当 r 不是非负整数时, 我们最常用特殊情形 $y=1$ 的二项定理. 让我们详细地说明这种特殊情形, 用 z 代替 x 来强调这里可含有任意复数:

$$(1+z)^r = \sum_k \binom{r}{k} z^k, \quad |z| < 1. \quad (5.13)$$

如果我们在此式中置 $z=x/y$ 且两边乘 y^r , 则得到式(5.12)中的一般公式.

仅当 r 是非负整数时, 我们用组合解释证明了二项定理. 用多项式论证我们不能从非负整数情形导出一般情形, 因为在一般情形和是无限的. 但是当 r 是任意时, 我们能利用泰勒级数和复变数理论:

$$\begin{aligned}
 f(z) &= \frac{f(0)}{0!} z^0 + \frac{f'(0)}{1!} z^1 + \frac{f''(0)}{2!} z^2 + \cdots \\
 &= \sum_{k \geq 0} \frac{f^{(k)}(0)}{k!} z^k.
 \end{aligned}$$

函数 $f(z) = (1+z)^r$ 的导数容易计算, 事实上, $f^{(k)}(z) = r^{\underline{k}} (1+z)^{r-k}$. 置 $z=0$ 得出式 (5.13).

我们还需证明当 $|z| < 1$ 时无穷和收敛. 因为由下面的方程 (5.83), $\binom{r}{k} = O(k^{-1-r})$, 所以它收敛.

现在让我们更仔细地看当 n 是负整数时 $\binom{n}{k}$ 的值. 处理这些值的一种方法是用加法法则 (5.8) 把项填入表 5.1 中数的上面, 从而得到表 5.2, 例如, 我们一定有 $\binom{-1}{0} = 1$, 因为 $\binom{0}{0} = \binom{-1}{0} + \binom{-1}{-1}$ 和 $\binom{-1}{-1} = 0$; 于是我们一定有 $\binom{-1}{1} = -1$, 因为 $\binom{0}{1} = \binom{-1}{1} + \binom{-1}{0}$; 等等.

表 5.2 Pascal 三角形, 向上扩展

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

所有这样的数是熟悉的数. 实际上, 表 5.2 的行和列如同表 5.1 中列的出现(但是没有负号). 所以对于负的 n 的 $\binom{n}{k}$ 值和正的 n 的值之间一定有一个联系. 一般规则为

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}, \text{ 整数 } k; \quad (5.14)$$

它是易证的, 因为

$$\begin{aligned}
 r^{\underline{k}} &= r(r-1)\cdots(r-k+1) \\
 &= (-1)^k (-r)(1-r)\cdots(k-1-r) = (-1)^k (k-r-1)^{\underline{k}},
 \end{aligned}$$

当 $k \geq 0$, 而当 $k < 0$ 时两边为零.

等式(5.14)特别有价值,因为它成立没有任何限制。(当然,下指数一定是整数以致二项系数是有定义的。)式(5.14)中的变换称为上指数求反,或“上求反”。

但是如何能记住这个重要的公式呢?我们看到的其他等式,对称,吸收,加法等十分简单,但是这一个等式看来相当凌乱。仍然有一种不太坏的记忆法:为了求反上指数,我们从记下 $(-1)^k$ 开始,其中 k 是下指数(下指数没有改变),然后立即在下指数和上指数位置再写 k 两次。然后我们把新上指数减去原来的上指数,且再减1而完成工作(总为减,不是加,因为这是一个求反过程)。

为了实践操作,让我们相继求反上指数两次,得到

$$\begin{aligned}\binom{r}{k} &= (-1)^k \binom{k-r-1}{k} \\ &= (-1)^{2k} \binom{k-(k-r-1)-1}{k} = \binom{r}{k},\end{aligned}$$

所以我们正确地回到开始之处。这也许不是等式创造者的打算,但是我们很清楚,并没有入歧途。

当然式(5.14)的某些应用比此更有用。例如,我们能用上求反来移动上和下指数位置的量。等式有一种对称的公式表示,

$$(-1)^m \binom{-n-1}{m} = (-1)^n \binom{-m-1}{n}, \text{ 整数 } m, n \geq 0, \quad (5.15)$$

因为两边都等于 $\binom{m+n}{n}$,它成立。

上求反也能用来引出下列有趣的和:

$$\begin{aligned}\sum_{k \leq m} \binom{r}{k} (-1)^k &= \binom{r}{0} - \binom{r}{1} + \cdots + (-1)^m \binom{r}{m} \\ &= (-1)^m \binom{r-1}{m}, \text{ 整数 } m,\end{aligned} \quad (5.16)$$

想法是求反上指数,则应用式(5.9),且再求反:

$$\begin{aligned}\sum_{k \leq m} \binom{r}{k} (-1)^k &= \sum_{k \leq m} \binom{k-r-1}{k} \\ &= \binom{-r+m}{m} \\ &= (-1)^m \binom{r-1}{m}.\end{aligned}$$

此公式给出 Pascal 三角形的第 r 行的部分和,规定行的各项给出交替的正负号。例如,若 $r=5$ 和 $m=2$,公式给出 $1-5+10=6=(-1)^2 \binom{4}{2}$ 。

注意如果 $m \geq r$,则式(5.16)给出整个行的交错和,而当 r 是正整数时此和为零。当我

们用二项定理展开 $(1-1)^r$ 时我们在前面证明了这一点,明白此表达式的部分和也能计算成闭形式是有趣的.

较简单的部分和

$$\sum_{k \leq m} \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m} \quad (5.17)$$

情况如何,如果我们能计算具有交错正负号的对应和,我们是否确实该能算出这一个呢?但是不能,Pascal 三角形的一行的部分和没有闭形式.我们能算出列,即式(5.10),但不能算出行.然而,难以理解的是,有一种方法能用来部分地求行元素的和,如果用行元素到中心的距离来乘它们:

$$\sum_{k \leq m} \binom{r}{k} \left(\frac{r}{2} - k \right) = \frac{m+1}{2} \binom{r}{m+1}, \text{ 整数 } m. \quad (5.18)$$

(用对 m 的归纳法容易验证此公式.)在被加数中有因子 $(r/2-k)$ 和没有因子 $(r/2-k)$ 的这些部分和之间的关系相似于积分

$$\int_{-\infty}^a x e^{-x^2} dx = -\frac{1}{2} e^{-a^2} \text{ 和 } \int_{-\infty}^a e^{-x^2} dx$$

之间的关系.外观上具有因子 x 的左边较复杂的积分有闭形式,而右边无因子的看来较简单的积分没有闭形式.外表能看错.

本章末我们研究一种方法,这种方法以一种直接一般的调整可能决定是否涉及二项系数的给定级数的部分和有一个闭形式.此法有发现等式(5.16)和(5.18)的能力,且它也将告诉我们式(5.17)是死胡同.

二项级数的部分和导致另一类型的一个好奇的关系:

$$\sum_{k \leq m} \binom{m+r}{k} x^k y^{m-k} = \sum_{k \leq m} \binom{-r}{k} (-x)^k (x+y)^{m-k}, \text{ 整数 } m. \quad (5.19)$$

用归纳法不难证明这个等式:当 $m < 0$ 两边为零, $m = 0$ 时为 1. 如果让 S_m 代表左边的和,我们能应用加法公式(5.8),且容易证明当 $m > 0$ 时

$$S_m = \sum_{k \leq m} \binom{m-1+r}{k} x^k y^{m-k} + \sum_{k \leq m} \binom{m-1+r}{k-1} x^k y^{m-k};$$

和

$$\sum_{k \leq m} \binom{m-1+r}{k} x^k y^{m-k} = y S_{m-1} + \binom{m-1+r}{m-1} x^m,$$

$$\sum_{k \leq m} \binom{m-1+r}{k-1} x^k y^{m-k} = x S_{m-1}.$$

因此,

$$S_m = (x+y)S_{m-1} + \binom{-r}{m}(-x)^m,$$

而且式(5.19)的右边也满足此递归。由归纳法, 两边一定相等, 证毕。

但是有一个较简洁的证明。当 r 是范围 $0 \geq r \geq -m$ 中的整数时, 二项定理告诉我们式(5.19)的两边为 $(x+y)^{m+r}y^{-r}$ 。而且因为两边是次数 m 或小于 m 的 r 的多项式, 一般在 $m+1$ 个不同值处相同足以(但仅勉强够!)证明相等。

看来有一个等式把一个和等于另一个和可能是愚蠢的, 两个和没有一个是闭形式, 但是有时产生的一边比另一边容易计算。例如, 如果我们置 $x=-1$ 和 $y=1$, 取得

$$\sum_{k \leq m} \binom{m+r}{k} (-1)^k = \binom{-r}{m}, \text{ 整数 } m \geq 0,$$

为等式(5.16)的另一种形式。如果我们置 $x=y=1$ 和 $r=m+1$, 取得

$$\sum_{k \leq m} \binom{2m+1}{k} = \sum_{k \leq m} \binom{m+k}{k} 2^{m-k}.$$

左边和刚好为具有上指数 $2m+1$ 的二项系数的一半, 而这些相同于另一半中它们的副本部分, 因为 Pascal 三角形左右对称。因此左边刚好是 $(1/2)2^{2m+1} = 2^{2m}$ 。这就产生一个完全没有想到的公式,

$$\sum_{k \leq m} \binom{m+k}{k} 2^{-k} = 2^m, \text{ 整数 } m \geq 0. \quad (5.20)$$

让我们检验它, 当 $m=2$ 时: $\binom{2}{0} + \frac{1}{2}\binom{3}{1} + \frac{1}{4}\binom{4}{2} = 1 + \frac{3}{2} + \frac{6}{4} = 4$, 使人震惊。

到现在为止, 我们所见到的是二项系数或者每项仅有一个二项系数的和。但是我们面临许多复杂的问题涉及到两个或多个二项系数的乘积, 所以我们在本节的余下部分考虑如何处理这样的情形。

这里是一个方便的规则, 常常有助于简化两个二项系数的乘积:

$$\binom{r}{m} \binom{m}{k} = \binom{r}{k} \binom{r-k}{m-k}, \text{ 整数 } m, k. \quad (5.21)$$

我们已见到特殊情形 $k=1$, 它是吸收等式(5.6)。虽然式(5.21)的两边是二项系数的乘积, 由于公式的余下部分的影响, 一边常常较易求和。例如, 左边用 m 二次, 右边仅用它一次。所以当对 m 求和时, 我们通常用 $\binom{r}{k} \binom{r-k}{m-k}$ 来替换 $\binom{r}{m} \binom{m}{k}$ 。

首先因为在 $\binom{r}{m}$ 和 $\binom{m}{k}$ 的阶乘表示中消去 $m!$, 所以方程(5.21)成立。若所有变量是整数且 $r \geq m \geq k \geq 0$, 我们得到

$$\binom{r}{m} \binom{m}{k} = \frac{r!}{m!(r-m)!} \frac{m!}{k!(m-k)!}$$

$$\begin{aligned}
 &= \frac{r!}{k!(m-k)!(r-m)!} \\
 &= \frac{r!}{k!(r-k)!} \frac{(r-k)!}{(m-k)!(r-m)!} = \binom{r}{k} \binom{r-k}{m-k},
 \end{aligned}$$

这是容易的。此外，如果 $m < k$ 或 $k < 0$ ，式(5.21)的两边是零，所以对所有整数 m 和 k 等式成立。最后，多项式论证把它推广到对所有实数 r 成立。

在适当的重新命名变量之后，二项系数 $\binom{r}{k} = r! / (r-k)!k!$ 能写成形式 $(a+b)! / a!b!$ 。类似，上面推导的中间的 $r! / k!(m-k)!(r-m)!$ 能写成形式 $(a+b+c)! / a!b!c!$ 。这是一个“三项系数”，在“三项定理”中产生它：

$$\begin{aligned}
 (x+y+z)^n &= \sum_{\substack{0 \leq a, b, c \leq n \\ a+b+c=n}} \frac{(a+b+c)!}{a!b!c!} x^a y^b z^c \\
 &= \sum_{\substack{0 \leq a, b, c \leq n \\ a+b+c=n}} \binom{a+b+c}{b+c} \binom{b+c}{c} x^a y^b z^c.
 \end{aligned}$$

所以 $\binom{r}{m} \binom{m}{k}$ 确实是一个乔装着的三项系数。在应用中偶然提出三项系数，为了强调对称形式，我们能方便地记它们为

$$\binom{a+b+c}{a, b, c} = \frac{(a+b+c)!}{a!b!c!}.$$

二项和三项系数推广到多项系数，它总可表达为二项系数的乘积：

$$\begin{aligned}
 \binom{a_1 + a_2 + \cdots + a_m}{a_1, a_2, \dots, a_m} &= \frac{(a_1 + a_2 + \cdots + a_m)!}{a_1! a_2! \cdots a_m!} \\
 &= \binom{a_1 + a_2 + \cdots + a_m}{a_2 + \cdots + a_m} \cdots \binom{a_{m-1} + a_m}{a_m}.
 \end{aligned}$$

所以，当我们偶然碰见这种情况时，就应用我们的标准的技巧。

现在我们来列表 5.3，它列出了最重要的标准技巧中间的一些等式。这些等式是当跟涉及两个二项系数乘积的和奋斗时有帮助的等式。每个这样的等式是在 k 上求和，在每个二项系数中出现一次 k ；还有 4 个几乎独立的参数，称为 m, n, r ，等，在每个指数位置一个。产生不同的情形依赖于是否 k 出现在上或下指数，且依赖于是否它出现一个正号或负号。有时有一个附加因子 $(-1)^k$ ，为了使项求和成闭形式而需要它。

要完全记住表 5.3 是太复杂；仅希望把它作为参考。但是此表中的第一个等式是最可记忆的，且应记住它。它阐明了两个二项系数的乘积的和（在所有整数 k 上求和），二项系数中上指数是常数以及下指数对于所有 k 有一个常数和，是相加下和上指数得到的二项系数。此等式称为 **Vandermonde** 的卷积，因为他在 1700 年稍晚的时候写了一篇关于它的重要论文^[293]；然而，早在 1303 年中国的 Chu Shih-Chieh 就知道了它。表 5.3 中的所

有其他等式通过小心地用像上指数求反或应用对称法则等等, 可从 Vandermonde 卷积得到。所以 Vandermonde 卷积是所有等式中最基本的等式。

表 5.3 二项系数的乘积的和

$$\sum_k \binom{r}{m+k} \binom{s}{n-k} = \binom{r+s}{m+n}, \quad \text{整数 } m, n. \quad (5.22)$$

$$\sum_k \binom{l}{m+k} \binom{s}{n+k} = \binom{l+s}{l-m+n}, \quad \begin{array}{l} \text{整数 } l \geq 0, \\ \text{整数 } m, n. \end{array} \quad (5.23)$$

$$\sum_k \binom{l}{m+k} \binom{s+k}{n} (-1)^k = (-1)^{l+m} \binom{s-m}{n-l}, \quad \begin{array}{l} \text{整数 } l \geq 0, \\ \text{整数 } m, n. \end{array} \quad (5.24)$$

$$\sum_{k \leq l} \binom{l-k}{m} \binom{s}{k-n} (-1)^k = (-1)^{l+m} \binom{s-m-1}{l-m-n}, \quad \begin{array}{l} \text{整数} \\ l, m, n \geq 0 \end{array} \quad (5.25)$$

$$\sum_{0 \leq k \leq l} \binom{l-k}{m} \binom{q+k}{n} = \binom{l+q+1}{m+n+1}, \quad \begin{array}{l} \text{整数 } l, m \geq 0, \\ \text{整数 } n \geq q \geq 0. \end{array} \quad (5.26)$$

我们通过给出一种合适的组合解释来证明 Vandermonde 卷积。如果用 $k-m$ 替换 k , $n-m$ 替换 n , 我们能设 $m=0$, 因此证明的等式为

$$\sum_k \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}, \quad \text{整数 } n. \quad (5.27)$$

设 r 和 s 是非负整数, 于是用多项式论证得到一般情形。右边 $\binom{r+s}{n}$ 是从 r 个男人和 s 个女人中间选取 n 个人的方式数。左边, 和的每一项是选取 k 个男人和 $n-k$ 个女人的方式数, 对所有 k 求和计算了每种可能性恰好一次。

我们常常从左到右用这些等式, 因为那是简化的方向。但是我们偶尔也以另一方向进行, 暂时使一表达式较复杂。当这样处理时, 我们通常产生一个双重和, 对此我们能交换求和的次序, 然后简化。

进一步讨论之前让我们看表 5.3 中还有两个等式的证明。证明式(5.23)是容易的, 我们仅需用 $\binom{l}{l-m-k}$ 来替换第一个二项系数, 然后应用 Vandermonde 的式(5.22)。

下一个式(5.24)的证明稍有一点困难。我们可通过一系列变换把它变为 Vandermonde 卷积, 但是我们能采用老的可靠的数学归纳法的技巧容易地证明它。当我们不能立即有明显的结论时, 常常试用归纳法, 且这里对 l 归纳恰很好。

对于基础 $l=0$, 除了当 $k=-m$ 时所有项为零, 所以方程的两边为 $(-1)^m \binom{s-m}{n}$ 。现在假设等式对于所有小于某个指定的 l 的所有值成立, 其中 $l>0$ 。我们能用法公式, 用 $\binom{l-1}{m+k} + \binom{l-1}{m+k-1}$ 来替换 $\binom{l}{m+k}$; 原来的和分成两个和, 用归纳假设能计算每一个和:

$$\begin{aligned} & \sum_k \binom{l-1}{m+k} \binom{s+k}{n} (-1)^k + \sum_k \binom{l-1}{m+k-1} \binom{s+k}{n} (-1)^k \\ &= (-1)^{l-1+m} \binom{s-m}{n-l+1} + (-1)^{l+m} \binom{s-m+1}{n-l+1}. \end{aligned}$$

如果我们再一次应用加法公式, 由此则简化了式(5.24)的右边.

关于这个推导有两点值得注意. 首先, 我们再次看到不在某个范围上而面对所有整数 k 求和的极大方便, 因为不需关心界限条件. 其次, 用数学归纳法加法公式很好处理, 因为它是二项系数的一个递归式. 上指数为 l 的二项系数表示为两个上指数为 $l-1$ 的二项系数, 这恰好是应用归纳假设所需要的.

表 5.3 就讲这么多. 3 个或多于 3 个二项系数乘积的和的结果是什么? 如果求和的指数伸展到所有的系数, 则求闭形式的机会不妙: 仅知道少数几个这种类型和的闭形式, 因此我们需要的和可以不和给定的样本相配. 一个在习题 43 中证明的希罕的等式是

$$\sum_k \binom{m-r+s}{k} \binom{n+r-s}{n-k} \binom{r+k}{m+n} = \binom{r}{m} \binom{s}{n}, \quad \text{整数 } m, n \geq 0. \quad (5.28)$$

这里是另一个较对称的例子:

$$\sum_k \binom{a+b}{a+k} \binom{b+c}{b+k} \binom{c+a}{c+k} (-1)^k = \frac{(a+b+c)!}{a!b!c!}, \quad \text{整数 } a, b, c \geq 0. \quad (5.29)$$

这一个有 2 个系数的相似部分,

$$\begin{aligned} & \sum_k \binom{a+b}{a+k} \binom{b+a}{b+k} (-1)^k = \frac{(a+b)!}{a!b!}, \\ & \text{整数 } a, b \geq 0, \end{aligned} \quad (5.30)$$

它在表 5.3 中不出现. 相似的 4 个系数和没有闭形式, 但是一个类似的和有闭形式:

$$\begin{aligned} & \sum_k (-1)^k \binom{a+b}{a+k} \binom{b+c}{b+k} \binom{c+d}{c+k} \binom{d+a}{d+k} / \binom{2a+2b+2c+2d}{a+b+c+d+k} \\ &= \frac{(a+b+c+d)!(a+b+c)!(a+b+d)!(a+c+d)!(b+c+d)!}{(2a+2b+2c+2d)!(a+c)!(b+d)!a!b!c!d!} \\ & \text{整数 } a, b, c, d \geq 0. \end{aligned}$$

这是在 20 世纪初期由 John Dougall^[69]发现的.

是否 Dougall 等式是所知道的最使人不快的二项系数的和? 不是! 到目前为止它是

$$\begin{aligned} & \sum_{k_{ij}} (-1)^{\sum_{1 \leq i < j \leq n} k_{ij}} \left(\prod_{1 \leq i < j \leq n} \binom{a_i + a_j}{a_j + k_{ij}} \right) \left(\prod_{1 \leq i < n} \binom{a_i + a_n}{a_n + \sum_{i < j} k_{ij} - \sum_{i > j} k_{ij}} \right) \\ &= \binom{a_1 + \cdots + a_n}{a_1, a_2, \cdots, a_n}, \quad \text{整数 } a_1, a_2, \cdots, a_n \geq 0. \end{aligned} \quad (5.31)$$

这里的和是对 $\binom{n-1}{2}$ 个指数变量 k_{ij} 求的, $1 < i < j \leq n$. 方程(5.29)是特殊情形, $n=3$; 情形 $n=4$ 可写出如下, 如果对于 (a_1, a_2, a_3, a_4) 我们用 (a, b, c, d) , 且对于 (k_{12}, k_{13}, k_{23}) 用 (i, j, k) :

$$\sum_{i,j,k} (-1)^{i+j+k} \binom{a+b}{b+i} \binom{a+c}{c+j} \binom{b+c}{c+k} \binom{a+d}{d-i-j} \binom{b+d}{d+i-k} \binom{c+d}{d+j+k} \\ = \frac{(a+b+c+d)!}{a!b!c!d!}, \text{ 整数 } a, b, c, d \geq 0.$$

在 $n(n-1)$ 个分数的乘积

$$\prod_{\substack{1 \leq i, j \leq n \\ i \neq j}} \left(1 - \frac{z_i}{z_j} \right)^{a_i}$$

全部展成 z 的正和负幂之后, 式(5.31)的左边就是 $z_1^0 z_2^0 \cdots z_n^0$ 的系数. 在1962年Freeman Dyson猜测了右边, 且此后不久几个人证明了它. 习题86给出了式(5.31)的一个“简单的”证明.

另一个涉及很多二项系数的值得注意的等式是

$$\sum_{j,k} (-1)^{j+k} \binom{j+k}{j} \binom{r}{j} \binom{n}{k} \binom{m+n-j-k}{m-j} \\ = \binom{n+r}{n} \binom{m-r}{m-n}, \quad \text{整数 } m, n \geq 0. \quad (5.32)$$

在习题83中证明的这一等式还有在实际应用中呈现的机会. 但是我们正远离了“基本等式”的主题, 所以停下来更好, 且观察我们所学到的内容.

我们已看到二项系数满足几乎使人迷惑的各种等式. 幸运的是, 有些等式容易记住, 且我们能记住的等式以少数几步推出大多数其他的等式. 表5.4收集了10个最有用的公式放在一起; 这些等式是知道的最适合的等式.

5.2 基本的实用

在前节中我们通过处理和以及组合其他等式而导出一些等式. 找这些推导是不太难对付的, 我们知道尝试证明的是, 所以能提出一个一般的计划, 且不太麻烦地填入细节. 然而, 通常不是面对一个要证明的等式, 我们面对的是要简化的一个和, 且不知道简化的形式什么 (即使简化的形式存在). 在这一节和下一节中处理许多这样的和, 我们将掌握二项系数的工具.

下面让我们尝试处理涉及一个二项系数的少数几个和.

问题 1: 一个比值的和

我们希望得到

$$\sum_{k=0}^m \binom{m}{k} / \binom{n}{k}, \text{ 整数 } n \geq m \geq 0,$$

的闭形式。第一眼看到此和有点莫名其妙, 因为我们没有看到过关于二项系数的商的任何等式。(此外, 和涉及二个二项系数, 它看来与这个问题前面的结论相抵触。) 然而, 正如能用阶乘表示来把二项系数的乘积再表达为另一个乘积, 这就是如何取得等式(5.21), 我们能照样地处理一个商。事实上, 设 $r=n$ 以及用 $\binom{n}{k} \binom{n}{m}$ 除方程(5.21)的两边, 能避免不简洁的阶乘表示, 同时产生

$$\binom{m}{k} / \binom{n}{k} = \binom{n-k}{m-k} / \binom{n}{m}.$$

所以我们用右边的商来替换出现在和中的左边的商, 和变成

$$\sum_{k=0}^m \binom{n-k}{m-k} / \binom{n}{m}.$$

我们仍得到一个商, 但是分母中的二项系数不涉及求和的指数 k , 所以能把它移出和。我们将在后面恢复它。

我们还能通过对所有 $k \geq 0$ 求和来简化界限条件, 由于对于 $k > m$ 的项为零, 留下的和就不那么吓人:

$$\sum_{k \geq 0} \binom{n-k}{m-k}.$$

它相似于式(5.9)中的一个等式, 因为具有相同正负号的指数 k 出现二次。但是这里它是 $-k$, 而在式(5.9)中不是。所以下一步是明显的, 仅有一个合适的事可做:

$$\sum_{k \geq 0} \binom{n-k}{m-k} = \sum_{m-k \geq 0} \binom{n-(m-k)}{m-(m-k)} = \sum_{k \leq m} \binom{n-m+k}{k}.$$

现在我们能应用平行求和等式, 即(5.9):

$$\sum_{k \leq m} \binom{n-m+k}{k} = \binom{(n-m)+m+1}{m} = \binom{n+1}{m}.$$

最后, 我们把前面从和中移出的分母中的 $\binom{n}{m}$ 恢复原状, 然后应用式(5.7)取得希望的闭形式:

$$\binom{n+1}{m} / \binom{n}{m} = \frac{n+1}{n+1-m}.$$

表 5.4 居首位的 10 个二项系数等式

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$,	整数 $n \geq k \geq 0$.	阶乘展开
$\binom{n}{k} = \binom{n}{n-k}$,	整数 $n \geq 0$, 整数 k .	对称
$\binom{r}{k} = \frac{r}{k} \binom{r-1}{k-1}$,	整数 $k \neq 0$.	吸收 / 抽出
$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}$,	整数 k .	加法 / 归纳
$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}$,	整数 k .	上求反
$\binom{r}{m} \binom{m}{k} = \binom{r}{k} \binom{r-k}{m-k}$,	整数 m, k .	三项修正
$\sum_k \binom{r}{k} x^k y^{r-k} = (x+y)^r$,	整数 $r \geq 0$ 或 $ x/y < 1$.	二项定理
$\sum_{k \leq n} \binom{r+k}{k} = \binom{r+n+1}{n}$.	整数 n .	类似求和
$\sum_{0 \leq k \leq n} \binom{k}{m} = \binom{n+1}{m+1}$,	整数 $m, n \geq 0$.	上求和
$\sum_k \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}$,	整数 n .	Vandermonde 结合式

这个推导实际对 n 的任何实值成立, 只要不出现被零除, 也就是说, 只要 n 不为整数 $0, 1, \dots, m-1$.

此外, 还有更重要的推导是检验解答。这并不太复杂, 但是我们总将检验一下。在小的情形 $m=2$ 和 $n=4$, 我们有

$$\binom{2}{0} \binom{4}{0} + \binom{2}{1} \binom{4}{1} + \binom{2}{2} \binom{4}{2} = 1 + \frac{1}{2} + \frac{1}{6} = \frac{5}{3}$$

这完全和闭形式 $(4+1)/(4+1-2)$ 一致。

问题 2: 来自分类文献中的问题

下一个出现的和由来已久(70 年代初期), 是在人们对二项系数运用自如之前。一篇介绍一个改进合并技巧的论文^[165]以下列陈述结束: “能证明节省转移的期望数…由表达式

$$T = \sum_{r=0}^{\infty} r \frac{m-r-1}{m} \frac{C_{m-n-1}}{C_n}$$

给出, 这里的 m 和 n 如同上面所定义的, $\frac{C_n}{m}$ 是每次 m 个元素取 n 个的组合数的符号。…作者感谢受托处理者把节省的期望转移的一个较复杂的方程化成这里所给出的形

式。”

我们将看到这确实不是作者的问题的最后解答，甚至它不是中间的解答。

首先我们应把和变换成我们能处理的某种形式，任何人总不愿用可怕的记法 $\sum_{k=0}^{m-n-1} C_{n-m-1}^{m-k-1}$ ，热心的受托处理者也受不了。用我们的语言来说，我们应写成

$$T = \sum_{k=0}^n k \binom{m-k-1}{m-n-1} / \binom{m}{n}, \text{ 整数 } m > n \geq 0.$$

分母中的二项系数不涉及求和的指数，所以我们能把它移出，且处理新和

$$S = \sum_{k=0}^n k \binom{m-k-1}{m-n-1}.$$

还有什么？求和的指数出现在二项系数的上指数而不是下指数。所以如果另一个 k 不在那里，我们能处理和，且应用上指数的求和式(5.10)。虽然我们不能有外加的 k ，但如果能以某种方式把 k 并入二项系数，用一个吸收等式，则能对上指数求和。不幸的是，这里不能用那些等式处理。但是如果用 $m-k$ 替换 k ，我们能应用吸收等式(5.6)：

$$(m-k) \binom{m-k-1}{m-n-1} = (m-n) \binom{m-k}{m-n}.$$

所以这里是关键：我们把 k 再写成 $m-(m-k)$ ，把和 S 分成两个和：

$$\begin{aligned} \sum_{k=0}^n k \binom{m-k-1}{m-n-1} &= \sum_{k=0}^n (m-(m-k)) \binom{m-k-1}{m-n-1} \\ &= \sum_{k=0}^n m \binom{m-k-1}{m-n-1} - \sum_{k=0}^n (m-k) \binom{m-k-1}{m-n-1} \\ &= m \sum_{k=0}^n \binom{m-k-1}{m-n-1} - \sum_{k=0}^n (m-n) \binom{m-k}{m-n} \\ &= mA - (m-n)B, \end{aligned}$$

其中

$$A = \sum_{k=0}^n \binom{m-k-1}{m-n-1}, \quad B = \sum_{k=0}^n \binom{m-k}{m-n}.$$

留下的和 A 以及 B 与以前上指数变而下指数固定的形式没有不同，让我们首先处理 B ，因为它看来较简单。稍作处理就足以使被加数与式(5.10)的左边相配：

$$\begin{aligned} \sum_{0 \leq k \leq n} \binom{m-k}{m-n} &= \sum_{0 \leq m-k \leq n} \binom{m-(m-k)}{m-n} \\ &= \sum_{m-n \leq k \leq m} \binom{k}{m-n} \end{aligned}$$

$$= \sum_{0 \leq k \leq m} \binom{k}{m-n}.$$

在最后一步中, 包括和中 $0 \leq k < m-n$ 的项, 它们全为零, 因为上指数小于下指数. 现在我们在上指数求和, 由式(5.10)得到

$$B = \sum_{0 \leq k \leq m} \binom{k}{m-n} = \binom{m+1}{m-n+1}.$$

另一个和 A 是相同的, 但是用 $m-1$ 替换 m . 因此我们有一个给定和 S 的闭形式, 可更进一步简化它:

$$\begin{aligned} S &= mA - (m-n)B = m \binom{m}{m-n} - (m-n) \binom{m+1}{m-n+1} \\ &= \left(m - (m-n) \frac{m+1}{m-n+1} \right) \binom{m}{m-n} \\ &= \left(\frac{n}{m-n+1} \right) \binom{m}{m-n}. \end{aligned}$$

这就给出了原先和的一个闭形式:

$$\begin{aligned} T &= S / \binom{m}{n} = \frac{n}{m-n+1} \binom{m}{m-n} / \binom{m}{n} \\ &= \frac{n}{m-n+1}. \end{aligned}$$

甚至处理者不能简化这个形式.

我们再用一个小的情形来检验解答. 当 $m=4$ 和 $n=2$ 时, 我们得到

$$T = 0 \cdot \binom{3}{1} / \binom{4}{2} + 1 \cdot \binom{2}{1} / \binom{4}{2} + 2 \cdot \binom{1}{1} / \binom{4}{2} = 0 + \frac{2}{6} + \frac{2}{6} = \frac{2}{3},$$

它与公式 $2/(4-2+1)$ 相同.

问题 3: 测验中的一个问题

让我们再处理一个涉及单个二项系数的和. 这一个与上一个不相同, 它起源于学术界, 是课外测验的一个问题. 当

$$Q_n = \sum_{k \leq 2^n} \binom{2^n - k}{k} (-1)^k, \text{ 整数 } n \geq 0$$

时, 我们要 $Q_{1\,000\,000}$ 的值. 这个问题比其他问题难, 我们不能应用到目前为止已见到的任何一个等式, 所以我们不能就把它加起来. 求和的指数 k 出现在上和下指数中, 正是正负号相反. 上指数求反也无帮助; 移去因子 $(-1)^k$, 但是它把 $2k$ 引入上指数.

当无从下手时, 我们知道最好看一看小的情形. 如果我们不能发现一个型式, 且用归

纳法证明它, 至少我们将有检验结果的一些数据。这里是一些非零项以及前 4 个 n 值的它们的和。

n			Q_n
0	$\binom{1}{0}$	$= 1$	$= 1$
1	$\binom{2}{0} - \binom{1}{1}$	$= 1 - 1$	$= 0$
2	$\binom{4}{0} - \binom{3}{1} + \binom{2}{2}$	$= 1 - 3 + 1$	$= -1$
3	$\binom{8}{0} - \binom{7}{1} + \binom{6}{2} - \binom{5}{3} + \binom{4}{4}$	$= 1 - 7 + 15 - 10 + 1$	$= 0$

我们最好不试下一个情形 $n=4$, 这使算术误差太大。(用手计算像 $\binom{12}{4}$ 和 $\binom{11}{5}$ 的项, 单独把它们和其他项结合起来, 仅当成功希望很小时才值得花时间。)

所以型式以 1, 0, -1, 0 开始, 即使我们知道下一项或两项, 闭形式也不明显。但是如果能找到并证明 Q_n 的一个递归式, 我们也许能猜测和证明它的闭形式。为了找一个递归式, 我们需要把 Q_n 和 Q_{n-1} 联系起来(或和 $Q_{\text{较小值}}$ 联系起来); 为了做到这一点, 我们需要把像 $\binom{128-13}{13}$ (当 $n=7$ 和 $k=13$ 时产生它) 的项和像 $\binom{64-13}{13}$ 的一些项联系起来。这看来没有希望, 因为我们不知道相距 64 行的 Pascal 三角形中项目之间的任何简洁的关系。归纳证明的主要工具加法公式仅联系相距 1 行的项目。

但是这一点导致一个关键的观察: 不需讨论相距 2^{n-1} 行的项目。变量 n 不单独出现, 在此意义上它总是 2^n , 所以 2^n 是不相干的东西! 若我们用 m 替换 2^n , 则只需找较一般的(但是较易的)和

$$R_m = \sum_{k \leq m} \binom{m-k}{k} (-1)^k, \text{ 整数 } m \geq 0$$

的闭形式, 于是我们也将得到 $Q_n = R_{2^n}$ 的闭形式。而有加法公式将给出序列 R_m 的递归式的好机会。

对于小的 m 的 R_m 的值可来自表 5.1, 如果我们交错加和减出现在西南到东北对角线中的值。结果是:

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

看来将有一大堆被消去。

让我们现在来看 R_m 的公式, 看是否它确定一个递归式。我们的策略是应用加法公式 (5.8), 且找在产生的表达式中有形式 R_k 的和, 有点像第二章的摄动法中所做的那样:

$$\begin{aligned}
R_m &= \sum_{k \leq m} \binom{m-k}{k} (-1)^k \\
&= \sum_{k \leq m} \binom{m-1-k}{k} (-1)^k + \sum_{k \leq m} \binom{m-1-k}{k-1} (-1)^k \\
&= \sum_{k \leq m} \binom{m-1-k}{k} (-1)^k + \sum_{k+1 \leq m} \binom{m-2-k}{k} (-1)^{k+1} \\
&= \sum_{k \leq m-1} \binom{m-1-k}{k} (-1)^k + \binom{-1}{m} (-1)^m \\
&\quad - \sum_{k \leq m-2} \binom{m-2-k}{k} (-1)^k - \binom{-1}{m-1} (-1)^{m-1} \\
&= R_{m-1} + (-1)^{2m} - R_{m-2} - (-1)^{2(m-1)} = R_{m-1} - R_{m-2}.
\end{aligned}$$

(在最后第二步我们用了公式 $\binom{-1}{m} = (-1)^m$, 我们知道当 $m \geq 0$ 时它为真。) 对于 $m \geq 2$, 此推导成立。

根据此递归很快能产生 R_m 的值, 且我们立即看出序列是周期的。实际上,

$$R_m = \begin{cases} 1 \\ 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{cases} \quad \text{如果 } m \bmod 6 = \begin{cases} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{cases}$$

用归纳法证明是用检验。或者, 如果一定要给出一个更学术的证明, 我们能把递归展开一步得到

$$R_m = (R_{m-2} - R_{m-3}) - R_{m-2} = -R_{m-3},$$

每当 $m \geq 3$ 。因此每当 $m \geq 6$ 时, $R_m = R_{m-6}$ 。

最后, 由于 $Q_n = R_{2^n}$, 我们能通过确定 $2^n \bmod 6$ 和用 R_m 的闭形式来确定 Q_n 。当 $n=0$ 时, 我们有 $2^0 \bmod 6 = 1$; 继续乘 $2 \pmod{6}$ 之后, 所以型式 2, 4 重复。因此

$$Q_n = R_{2^n} = \begin{cases} R_1 = 1, & \text{若 } n = 0; \\ R_2 = 0, & \text{若 } n \text{ 是奇数;} \\ R_4 = -1, & \text{若 } n > 0 \text{ 是偶数.} \end{cases}$$

这个 Q_n 的闭形式和问题开始时我们计算的前 4 个值相同。我们断定 $Q_{1\,000\,000} = R_4 = -1$ 。

问题 4: 涉及两个二项系数的一个和。

我们的下一个任务是找

$$\sum_{k=0}^n k \binom{m-k-1}{m-n-1}, \text{ 整数 } m > n \geq 0$$

的闭形式。等一等，此问题的标题中约定的第二个二项系数在哪里？且为什么我们该尝试简化我们已简化的一个和？（这是问题 2 中的和 S 。）

如果我们视被加数为两个二项系数的乘积，则此为较容易简化的一个和，然后利用表 5.3 中找到的一个一般等式。当我们把 k 再写为 $\binom{k}{1}$ 时，第二个二项系数出现：

$$\sum_{k=0}^n k \binom{m-k-1}{m-n-1} = \sum_{0 \leq k \leq n} \binom{k}{1} \binom{m-k-1}{m-n-1}.$$

等式(5.26)是应用的一个等式，因为它的求和指数出现在两个上指数中且有相反正负号。

但是我们的和还不是完全恰当的形式。如果要有一个与式(5.26)相配的形式，求和的上限该是 $m-1$ 。没有问题， $n < k \leq m-1$ 的项为零。所以我们能插入 $(l, m, n, q) \leftarrow (m-1, m-n-1, 1, 0)$ ，解答为

$$S = \binom{m}{m-n+1}.$$

这比我们前面取得的公式更清楚。用式(5.7)能把它转换到前面的公式：

$$\binom{m}{m-n+1} = \frac{n}{m-n+1} \binom{m}{m-n}.$$

类似，通过把特殊值插入到已见过的其他一般等式中去，我们能取得有趣的结果。例如，在式(5.26)中假设置 $m=n=1$ 和 $q=0$ ，则等式读成

$$\sum_{0 \leq k \leq l} (l-k)k = \binom{l+1}{3}.$$

左边为 $l((l+1)l/2) - (1^2 + 2^2 + \cdots + l^2)$ ，这就给出一种解第二章中费力弄明白的平方和问题的新方法。

这个描述的意思是：非常一般的和的特殊情形有时在一般形式中处理最合适。当学一般形式时，弄清楚它们的简单的特殊情形是聪明的。

问题 5：有三个因子的和

这里是另一个不太坏的和。我们希望简化

$$\sum_k \binom{n}{k} \binom{s}{k} k, \text{ 整数 } n \geq 0.$$

求和的指数 k 出现在两个下指数中，且有相同正负号，所以表 5.3 中的等式(5.23)看来接近于我们需要的。作一点处理，我们该能用它。

式(5.23)与我们所得到的和之间的最大不同是在我们的和中额外的 k ，但是能用一个吸收等式把 k 吸入一个二项系数：

$$\sum_k \binom{n}{k} \binom{s}{k} k = \sum_k \binom{n}{k} \binom{s-1}{k-1} s = s \sum_k \binom{n}{k} \binom{s-1}{k-1}.$$

我们不关心当 k 不出现时 s 出现，因为它是常量。现在我们准备应用等式且取得闭形式，

$$s \sum_k \binom{n}{k} \binom{s-1}{k-1} = s \binom{n+s-1}{n-1}.$$

若我们第一步中选 k 吸入 $\binom{n}{k}$ ，而不是 $\binom{s}{k}$ ，将不允许直接应用式(5.23)，因为 $n-1$ 可以是负的；至少在一个上指数中等式要求一个非负值。

问题 6：具有威胁特征的和

下一个和是更复杂的。我们寻找

$$\sum_{k \geq 0} \binom{n+k}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1}, \text{ 整数 } n \geq 0,$$

的闭形式。一种和的困难的有用测度是求和指数出现的次数。根据此测度我们的麻烦很大， k 出现了 6 次。此外，前面问题中处理的关键步，把二项系数外的某个量吸入一个二项系数不能在这里起作用。若我们吸收 $k+1$ ，则在它的位置中我们刚好取得 k 的另一个出现。且不仅如此：指数 k 在一个二项系数内二次以系数 2 束缚它。乘法常数通常比加法常数较难移动。

虽然此时我们幸运。 $2k$ 是恰当的，为了应用等式(5.21)我们需要它们，所以我们取得

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

两个 2 不出现，且这样使 k 出现一次，以致第一个障碍排除了，但还有 5 个。

分母中的 $k+1$ 是留下的最讨厌的特征，现在我们用等式(5.6)把它吸收入 $\binom{n}{k}$ ：

$$\begin{aligned} \sum_{k \geq 0} \binom{n+k}{k} \binom{n}{k} \frac{(-1)^k}{k+1} &= \sum_k \binom{n+k}{k} \binom{n+1}{k+1} \frac{(-1)^k}{n+1} \\ &= \frac{1}{n+1} \sum_k \binom{n+k}{k} \binom{n+1}{k+1} (-1)^k. \end{aligned}$$

(记住 $n \geq 0$) 两个障碍排除了，还有 4 个。

为了消去另一个 k 我们有两个有希望的选择。我们能利用 $\binom{n+k}{k}$ 的对称性，或者求反上指数 $n+k$ ，从而既消去因子 $(-1)^k$ 又消去 k 。让我们探索两种可能性，用对称选择开始：

$$\frac{1}{n+1} \sum_k \binom{n+k}{k} \binom{n+1}{k+1} (-1)^k = \frac{1}{n+1} \sum_k \binom{n+k}{n} \binom{n+1}{k+1} (-1)^k.$$

第三个障碍排除了, 还有 3 个, 且我们能够通过代入式(5.24)得出很大好处: 用 $(n+1, 1, n, n)$ 替换 (l, m, n, s) , 我们取得

$$\frac{1}{n+1} \sum_k \binom{n+k}{n} \binom{n+1}{k+1} (-1)^k = \frac{1}{n+1} (-1)^n \binom{n+1}{-1} = 0.$$

零, 是吗? 这样到底行得通吗? 当 $n=2$ 时让我们检验它: $\binom{2}{0}\binom{0}{0}\frac{1}{1} - \binom{3}{2}\binom{2}{1}\frac{1}{2} + \binom{4}{4}\binom{4}{2}\frac{1}{3} = 1 - \frac{6}{2} + \frac{6}{3} = 0$. 检查了.

只是为了好玩, 让我们探索另一个选择, 求反 $\binom{n+k}{k}$ 的上指数:

$$\frac{1}{n+1} \sum_k \binom{n+k}{k} \binom{n+1}{k+1} (-1)^k = \frac{1}{n+1} \sum_k \binom{-n-1}{k} \binom{n+1}{k+1}.$$

现在应用式(5.23), 以 $(l, m, n, s) \leftarrow (n+1, 1, 0, -n-1)$, 和

$$\frac{1}{n+1} \sum_k \binom{-n-1}{k} \binom{n+1}{k+1} = \frac{1}{n+1} \binom{0}{n}.$$

等一等. 当 $n > 0$ 时这是零, 但是当 $n=0$ 时它是 1. 另一条解的路径告诉我们, 所有情形和是零! 给出的情况怎样? 当 $n=0$ 时, 和实际得出的是 1, 而正确的解答是 $‘n=0’$. 我们在前面推导中一定造成一个错误.

为了看出第一次产生的差异, 让我们再重现一次 $n=0$ 时的推导. 对, 我们陷入了前面提到的老的陷阱: 当上指数能负时我们尝试应用对称性! 当 k 的范围为所有整数时我们用 $\binom{n+k}{n}$ 替换 $\binom{n+k}{k}$ 是没有道理的, 因为当 $k < -n$ 时把零转换到了一个非零值. (这很遗憾.)

和中另一个因子 $\binom{n+1}{k+1}$, 除了当 $n=0$ 和 $k=-1$ 外, 当 $k < -n$ 时原来是零. 因此当我们检验 $n=2$ 的情形时错误不显露. 习题 6 阐明了我们应完成的事情.

问题 7: 新的障碍

这一个也难对付, 我们要

$$\sum_{k \geq 0} \binom{n+k}{m+2k} \binom{2k}{k} \frac{(-1)^k}{k+1}, \text{ 整数 } m, n > 0$$

的闭形式, 若 m 为 0, 则根据刚结束的问题我们得到和, 但是它不是, 留下了一个实际的困境, 在问题 6 中所用的在这里行不通. (特别不能用关键的第一步.)

然而, 如果我们能以某种方式除去 m , 就能用刚才导出的结果. 所以我们的策略

是：对于某个非负整数 l ，用像 $\binom{l+k}{2k}$ 的一些项的和替换 $\binom{n+k}{m+2k}$ ；于是被加数看来像问题 6 中的被加数，且能交换求和的次序。

$\binom{n+k}{m+2k}$ 的代用品是什么？本章中较早导出的等式的过细的检查仅发现一个适当的选择物，即表 5.3 中的方程(5.26)。用它的一种方法是用 $(n+k-1, 2k, m-1, 0, j)$ 分别替换参数 (l, m, n, q, k) ：

$$\begin{aligned} \sum_{k \geq 0} \binom{n+k}{m+2k} \binom{2k}{k} \frac{(-1)^k}{k+1} \\ &= \sum_{k \geq 0} \sum_{0 \leq j \leq n+k-1} \binom{n+k-1-j}{2k} \binom{j}{m-1} \binom{2k}{k} \frac{(-1)^k}{k+1} \\ &= \sum_{j \geq 0} \binom{j}{m-1} \sum_{\substack{k \geq j-n+1 \\ k \geq 0}} \binom{n+k-1-j}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1}. \end{aligned}$$

在最后一步中我们改变求和的次序，按照第二章的规则处置 \sum 下面的条件。

我们不能用问题 6 的结果完全替换内部和，因为它有额外条件 $k \geq j-n+1$ 。但是此额外条件是多余的，除非 $j-n+1 > 0$ ；也就是说，除非 $j \geq n$ 。且当 $j \geq n$ 时，内部和的第一个二项系数为零，因为它的上指数在 0 和 $k-1$ 之间，于是严格地小于下指数 $2k$ 。所以我们可以把附加限制 $j < n$ 放在外部和，不影响包含的非零项。这使限制 $k \geq j-n+1$ 多余，而我们能问题 6 的结果。现在双重和终于明白：

$$\begin{aligned} \sum_{j \geq 0} \binom{j}{m-1} \sum_{\substack{k \geq j-n+1 \\ k \geq 0}} \binom{n+k-1-j}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1} \\ &= \sum_{0 \leq j < n} \binom{j}{m-1} \sum_{k \geq 0} \binom{n+k-1-j}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1} \\ &= \sum_{0 \leq j < n} \binom{j}{m-1} [n-1-j=0] = \binom{n-1}{m-1}. \end{aligned}$$

除当 $j=n-1$ 时外，内部和消失，由此我们取得一个简单的闭形式作为我们的解答。

问题 8：一个不同的障碍

让我们以另一种考虑和

$$S_m = \sum_{k \geq 0} \binom{n+k}{2k} \binom{2k}{k} \frac{(-1)^k}{k+1+m}, \quad \text{整数 } m, n \geq 0.$$

的方法来扩充问题 6 的范围。另外，当 $m=0$ 时我们得到以前处理过的和，但是现在 m 出现在不同的位置。这个问题比问题 7 稍难，但是(幸运的是)较有利于找解。如同问题 6，我们能从

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

开始。现在(如同问题 7)我们尝试把依赖 m 的部分展开成我们知道如何处理的一些项。当 m 为零时, 我们把 $k+1$ 吸入 $\binom{n}{k}$; 若 $m > 0$, 如果把 $1/(k+1+m)$ 展开为可吸收的项, 我们能做同样的事情。而我们的运气仍然保持: 在问题 1 中证明一个适当的等式

$$\sum_{j=0}^n \binom{m}{j} \binom{r}{j}^{-1} = \frac{r+1}{r+1-m}, \quad \text{整数 } m \geq 0, \quad r \notin \{0, 1, \dots, m-1\}. \quad (5.33)$$

用 $-k-2$ 替换 r 给出希望的展开,

$$S_m = \sum_{k \geq 0} \binom{n+k}{k} \binom{n}{k} \frac{(-1)^k}{k+1} \sum_{j \geq 0} \binom{m}{j} \binom{-k-2}{j}^{-1}$$

现在如同计划的那样, 能把 $(k+1)^{-1}$ 吸入 $\binom{n}{k}$ 。事实上, 它也能吸入 $\binom{-k-2}{j}^{-1}$ 。双重吸收的建议在幕后甚至更多相消成为可能。是的, 把我们的新的被加数中的东西展成阶乘且回到二项系数, 得到我们能对 k 求和的公式:

$$\begin{aligned} S_m &= \frac{m!n!}{(m+n+1)!} \sum_{i \geq 0} (-1)^i \binom{m+n+1}{n+1+j} \sum_k \binom{n+1+j}{k+j+1} \binom{-n-1}{k} \\ &= \frac{m!n!}{(m+n+1)!} \sum_{j \geq 0} (-1)^j \binom{m+n+1}{n+1+j} \binom{j}{n}. \end{aligned}$$

由式(5.24), 在所有整数 j 上的和为零。因此 $-S_m$ 是 $j < 0$ 的和。

为了计算 $j < 0$ 的 $-S_m$, 让我们用 $-k-1$ 替换 j 且对 $k \geq 0$ 求和:

$$\begin{aligned} S_m &= \frac{m!n!}{(m+n+1)!} \sum_{k \geq 0} (-1)^k \binom{m+n+1}{n-k} \binom{-k-1}{n} \\ &= \frac{m!n!}{(m+n+1)!} \sum_{k \leq n} (-1)^{n-k} \binom{m+n+1}{k} \binom{k-n-1}{n} \\ &= \frac{m!n!}{(m+n+1)!} \sum_{k \leq n} (-1)^k \binom{m+n+1}{k} \binom{2n-k}{n} \\ &= \frac{m!n!}{(m+n+1)!} \sum_{k \leq 2n} (-1)^k \binom{m+n+1}{k} \binom{2n-k}{n}. \end{aligned}$$

最后应用式(5.25), 我们得到解答:

$$S_m = (-1)^n \frac{m!n!}{(m+n+1)!} \binom{m}{n} = (-1)^n m^{\overline{n}} m^{-n-1}.$$

我们最好还是检验它。当 $n=2$ 时, 求得

$$S_m = \frac{1}{m+1} - \frac{6}{m+2} + \frac{6}{m+3} = \frac{m(m-1)}{(m+1)(m+2)(m+3)}.$$

我们的推导要求 m 是整数, 但是结果对所有实数 m 成立, 因为 $(m+1)^{n+1} S_m$ 是次数 $\leq n$ 的 m 的多项式。

5.3 处理的诀窍

让我们接着看三个技巧, 这些技巧是我们已学的方法的有意义的引伸。

诀窍 1: 平分

我们的许多等式涉及一个任意实数 r 。当 r 具有特殊形式“整数减二分之一”, 二项系数 $\binom{r}{k}$ 能记为看来完全不同的二项系数的乘积, 这导致一族能意外灵活操作的新等式。

为了看清这样的一种方法如何工作, 从倍量公式

$$r^k \left(r - \frac{1}{2}\right)^k = (2r)^{2k} / 2^{2k}, \quad \text{整数 } k \geq 0, \quad (5.34)$$

开始。如果我们展开下降幂, 且交插左边的因子, 此等式是明显的:

$$r \left(r - \frac{1}{2}\right) (r-1) \left(r - \frac{3}{2}\right) \cdots (r-k+1) \left(r - k + \frac{1}{2}\right) = \frac{(2r)(2r-1)\cdots(2r-2k+1)}{2 \cdot 2 \cdot \cdots \cdot 2}.$$

现在用 $k!^2$ 除两边, 我们取得

$$\binom{r}{k} \binom{r-1/2}{k} = \binom{2r}{2k} \binom{2k}{k} / 2^{2k}, \quad \text{整数 } k. \quad (5.35)$$

如果我们置 $k=r=n$, 其中 n 是整数, 这就产生

$$\binom{n-1/2}{n} = \binom{2n}{n} / 2^{2n}, \quad \text{整数 } n. \quad (5.36)$$

求反上指数还得到另一个有用公式,

$$\binom{-1/2}{n} = \left(\frac{-1}{4}\right)^n \binom{2n}{n}, \quad \text{整数 } n. \quad (5.37)$$

例如, 当 $n=4$ 时我们得到

$$\begin{aligned} \binom{-1/2}{4} &= \frac{(-1/2)(-3/2)(-5/2)(-7/2)}{4!} \\ &= \left(\frac{-1}{2}\right)^4 \frac{1 \cdot 3 \cdot 5 \cdot 7}{1 \cdot 2 \cdot 3 \cdot 4} \end{aligned}$$

$$= \left(\frac{-1}{4} \right)^4 \frac{1 \cdot 3 \cdot 5 \cdot 7 \cdot 2 \cdot 4 \cdot 6 \cdot 8}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 1 \cdot 2 \cdot 3 \cdot 4} = \left(\frac{-1}{4} \right)^4 \binom{8}{4}.$$

注意, 如何把奇数个数的乘积改变为一个阶乘.

等式(5.35)有一个有趣的推论. 设 $r = \frac{1}{2}n$, 且在所有整数 k 上求和. 用式(5.23), 结果是

$$\begin{aligned} \sum_k \binom{n}{2k} \binom{2k}{k} 2^{-2k} &= \sum_k \binom{n/2}{k} \binom{(n-1)/2}{k} \\ &= \binom{n-1/2}{\lfloor n/2 \rfloor}, \quad \text{整数 } n \geq 0, \end{aligned} \quad (5.38)$$

因为 $n/2$ 或者 $(n-1)/2$ 是 $\lfloor n/2 \rfloor$, 一个非负整数!

我们也能用 Vandermonde 卷积(5.27)推出

$$\sum_k \binom{-1/2}{k} \binom{-1/2}{n-k} = \binom{-1}{n} = (-1)^n, \quad \text{整数 } n \geq 0.$$

根据式(5.37)代入值得出

$$\begin{aligned} \binom{-1/2}{k} \binom{-1/2}{n-k} &= \left(\frac{-1}{4} \right)^k \binom{2k}{k} \left(\frac{-1}{4} \right)^{n-k} \binom{2(n-k)}{n-k} \\ &= \frac{(-1)^n}{4^n} \binom{2k}{k} \binom{2n-2k}{n-k}, \end{aligned}$$

这就是和达到 $(-1)^n$. 因此我们有一个关于 Pascal 三角形的“中间”元素的值得注意的性质:

$$\sum_k \binom{2k}{k} \binom{2n-2k}{n-k} = 4^n, \quad \text{整数 } n \geq 0. \quad (5.39)$$

例如, $\binom{0}{0} \binom{6}{3} + \binom{2}{1} \binom{4}{2} + \binom{4}{2} \binom{2}{1} + \binom{6}{3} \binom{0}{0} = 1 \cdot 20 + 2 \cdot 6 + 6 \cdot 2 + 20 \cdot 1 = 64 = 4^3$.

第一个诀窍的这些说明表明尝试把形式 $\binom{2k}{k}$ 的二项系数改为形式 $\binom{n-1/2}{k}$ 的二项系数是聪明的, 其中 n 是某个适当的整数(通常为 0, 1, 或 k), 产生的结果可能简单多了.

诀窍 2: 高阶差分

我们在前面看到, 可能计算数列 $\binom{n}{k} (-1)^k$ 的部分和, 而不是数列 $\binom{n}{k}$ 的部分和,

• 结果有交错正负号的二项系数 $\binom{n}{k} (-1)^k$ 的许多重要应用. 关于这一点的一个原因是这

样的系数密切地和 2.6 节中定义的差分算子 Δ 结合在一起。

函数在点 x 处的差分 Δf 是

$$\Delta f(x) = f(x+1) - f(x);$$

若我们再运用 Δ ，则得到第二次差分

$$\begin{aligned}\Delta^2 f(x) &= \Delta f(x+1) - \Delta f(x) = (f(x+2) - f(x+1)) - (f(x+1) - f(x)) \\ &= f(x+2) - 2f(x+1) + f(x),\end{aligned}$$

它相似于第二次导数。类似，我们得到

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

$$\Delta^4 f(x) = f(x+4) - 4f(x+3) + 6f(x+2) - 4f(x+1) + f(x),$$

等等，二项系数进入这些具有交错正负号的公式。

一般，第 n 次差分是

$$\Delta^n f(x) = \sum_k \binom{n}{k} (-1)^{n-k} f(x+k), \text{ 整数 } n \geq 0. \quad (5.40)$$

此公式用归纳法容易证明，但是也有一种好方法直接用算子的初等理论来证明它。记得 2.6 节由规则

$$Ef(x) = f(x+1)$$

定义的移位算子 E ，因此算子 Δ 是 $E-1$ ，其中 1 是由规则 $1f(x) = f(x)$ 定义的恒等算子，由二项定理，

$$\Delta^n = (E-1)^n = \sum_k \binom{n}{k} E^k (-1)^{n-k}.$$

这是一个元素为算子的方程，它等价于式(5.40)，由于 E^k 是把 $f(x)$ 处理为 $f(x+k)$ 的算子。

当我们考虑负下降幂时，产生一种有趣和重要的情形。设 $f(x) = (x-1)^{-1} = 1/x$ 。然后，由规则 (2.45)，我们得到 $\Delta f(x) = (-1)(x-1)^{-2}$ ， $\Delta^2 f(x) = (-1) \cdot (-2)(x-1)^{-3}$ ，一般

$$\Delta^n ((x-1)^{-1}) = (-1)^n (x-1)^{-n-1} = (-1)^n \frac{n!}{x(x+1)\cdots(x+n)}.$$

现在方程(5.40)告诉我们

$$\begin{aligned}\sum_k \binom{n}{k} \frac{(-1)^k}{x+k} &= \frac{n!}{x(x+1)\cdots(x+n)} \\ &= x^{-1} \binom{x+n}{n}^{-1}, \quad x \notin \{0, -1, \dots, -n\}.\end{aligned}\quad (5.41)$$

例如,

$$\begin{aligned}\frac{1}{x} - \frac{4}{x+1} + \frac{6}{x+2} - \frac{4}{x+3} + \frac{1}{x+4} \\ = \frac{4!}{x(x+1)(x+2)(x+3)(x+4)} = 1/x \binom{x+4}{4}.\end{aligned}$$

式(5.41)中的和是 $n!/(x(x+1)\cdots(x+n))$ 的部分分式展开.

也能根据正下降幂获得有意义的结果. 若 $f(x)$ 是次数为 d 的多项式, 差分 $\Delta f(x)$ 是次数 $d-1$ 的多项式; 所以 $\Delta^d f(x)$ 是常量, 且如果 $n > d$, $\Delta^n f(x) = 0$. 这个极重要的事实简化了许多公式.

仔细一看, 可得到更进一步的信息: 设

$$f(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_1 x^1 + a_0 x^0$$

是次数 d 的任何多项式. 在第六章中我们将看到能把通常幂表达为下降幂的和 (例如, $x^2 = x^{\underline{2}} + x^{\underline{1}}$), 因此有系数 $b_d, b_{d-1}, \dots, b_1, b_0$ 使得

$$f(x) = b_d x^{\underline{d}} + b_{d-1} x^{\underline{d-1}} + \cdots + b_1 x^{\underline{1}} + b_0 x^{\underline{0}}.$$

(结果 $b_d = a_d$ 和 $b_0 = a_0$, 但是介于中间的系数以较复杂的方式联系起来.) 设 $c_k = k!b_k$, $0 \leq k \leq d$, 则

$$f(x) = c_d \binom{x}{d} + c_{d-1} \binom{x}{d-1} + \cdots + c_1 \binom{x}{1} + c_0 \binom{x}{0},$$

于是, 任何多项式能表示为二项系数的倍数的和. 这样一个展开称为 $f(x)$ 的牛顿级数, 因为牛顿多方面地用了它.

我们在本章前面观察到加法公式意味着

$$\Delta \left(\binom{x}{k} \right) = \binom{x}{k-1}.$$

所以, 由归纳法, 牛顿级数的第 n 次差分是十分简单的:

$$\Delta^n f(x) = c_d \binom{x}{d-n} + c_{d-1} \binom{x}{d-1-n} + \cdots + c_1 \binom{x}{1-n} + c_0 \binom{x}{-n}.$$

若我们现在置 $x=0$, 除 $k=n=0$ 的项外, 右边的所有项 $c^k \binom{x}{k-n}$ 为零; 因此

$$\Delta^n f(0) = \begin{cases} c_n, & \text{若 } n \leq d, \\ 0, & \text{若 } n > d. \end{cases}$$

$f(x)$ 的牛顿级数是

$$f(x) = \Delta^d f(0) \binom{x}{d} + \Delta^{d-1} f(0) \binom{x}{d-1} + \cdots + \Delta f(0) \binom{x}{1} + f(0) \binom{x}{0}.$$

例如, 假设 $f(x) = x^3$. 容易计算

$$f(0) = 0, \quad f(1) = 1, \quad f(2) = 8, \quad f(3) = 27;$$

$$\Delta f(0) = 1, \quad \Delta f(1) = 7, \quad \Delta f(2) = 19;$$

$$\Delta^2 f(0) = 6, \quad \Delta^2 f(1) = 12;$$

$$\Delta^3 f(0) = 6.$$

所以牛顿级数为 $x^3 = 6 \binom{x}{3} + 6 \binom{x}{2} + 1 \binom{x}{1} + 0 \binom{x}{0}$.

公式 $\Delta^n f(0) = c_n$ 也能以如下方式陈述, 用 $x=0$ 的式(5.40):

$$\sum_k \binom{n}{k} (-1)^k \left(c_0 \binom{k}{0} + c_1 \binom{k}{1} + c_2 \binom{k}{2} + \cdots \right) = (-1)^n c_n, \quad \text{整数 } n \geq 0.$$

这里 $\langle c_0, c_1, c_2, \cdots \rangle$ 是任意系数序列, 对于所有 k , 无限和 $c_0 \binom{k}{0} + c_1 \binom{k}{1} + c_2 \binom{k}{2} + \cdots$ 实际是有限的, 所以收敛不是一个问题。特别, 我们能证明重要的等式

$$\sum_k \binom{n}{k} (-1)^k (a_0 + a_1 k + \cdots + a_n k^n) = (-1)^n n! a_n, \quad \text{整数 } n \geq 0, \quad (5.42)$$

因为多项式 $a_0 + a_1 k + \cdots + a_n k^n$ 总能记为牛顿级数 $c_0 \binom{k}{0} + c_1 \binom{k}{1} + \cdots + c_n \binom{k}{n}$, $c_n = n! a_n$.

初看, 出现的许多和用第 n 次差分的思想来平凡地求和实际是没有希望的。例如, 让我们考虑等式

$$\sum_k \binom{n}{k} \binom{r-sk}{n} (-1)^k = s^n, \quad \text{整数 } n \geq 0. \quad (5.43)$$

这看来印象十分深刻, 因为它完全不同于以前见到的结果。但是一旦我们注意到被加数中说明问题的因子 $\binom{n}{k} (-1)^k$, 它实际是易了解的, 因为函数

$$f(k) = \binom{r-sk}{n} = \frac{1}{n!} (-1)^n s^n k^n + \cdots - (-1)^n s^n \binom{k}{n} + \cdots$$

是次数 n 的 k 的多项式, 第一项系数为 $(-1)^n s^n / n!$. 所以式(5.43)就是式(5.42)的一种应用.

我们已讨论了 $f(x)$ 是多项式的假设下的牛顿级数. 但是我们还看到无限牛顿级数

$$f(x) = c_0 \binom{x}{0} + c_1 \binom{x}{1} + c_2 \binom{x}{2} + \cdots$$

也有意义, 因为当 x 是非负整数时这样的和总是有限的, 就像多项式情形那样, 在无限情形公式 $\Delta^n f(0) = c_n$ 的推导成立; 所以我们有一般等式

$$f(x) = f(0) \binom{x}{0} + \Delta f(0) \binom{x}{1} + \Delta^2 f(0) \binom{x}{2} + \Delta^3 f(0) \binom{x}{3} + \cdots, \text{ 整数 } x \geq 0. \quad (5.44)$$

对于非负整数 x 定义的任何函数此公式成立. 此外, 如果对于 x 的其他值右边收敛, 则它定义了一个以一种自然方式“插入” $f(x)$ 的函数. (有无限多种方式来插入函数值, 所以我们不能断定对于使无限级数收敛的所有 x 式(5.44)成立. 例如, 如果我们设 $f(x) = \sin(\pi x)$, 我们在所有整数点处得到 $f(x) = 0$. 所以式(5.44)的右边恒等于零, 但是在所有非整数 x 处左边为非零.)

牛顿级数是无限演算的泰勒级数的有限演算的解答. 就像泰勒级数能记为

$$g(a+x) = \frac{g(a)}{0!} x^0 + \frac{g'(a)}{1!} x^1 + \frac{g''(a)}{2!} x^2 + \frac{g'''(a)}{3!} x^3 + \cdots,$$

$f(x) = g(a+x)$ 的牛顿级数能记为

$$g(a+x) = \frac{g(a)}{0!} x^0 + \frac{\Delta g(a)}{1!} x^1 + \frac{\Delta^2 g(a)}{2!} x^2 + \frac{\Delta^3 g(a)}{3!} x^3 + \cdots, \quad (5.45)$$

(这与式(5.44)相同, 因为当 $f(x) = g(a+x)$ 时对于所有 $n \geq 0$, $\Delta^n f(0) = \Delta^n g(a)$.) 当 g 是多项式, 或当 $x=0$ 时, 泰勒和牛顿级数都是有限的; 此外, 当 x 是正整数时, 牛顿级数是有限的. 否则对于 x 的特殊值, 和可能或不可能收敛. 如果当 x 不是非负整数时牛顿级数收敛, 它实际可以收敛到一个不同于 $g(a+x)$ 的值, 因为牛顿级数式(5.45)仅依赖于分隔开的函数值 $g(a)$, $g(a+1)$, $g(a+2)$, \cdots .

由二项定理提供了收敛牛顿级数的一个例子. 设 $g(x) = (1+z)^x$, 其中 z 是固定的复数使得 $|z| < 1$. 于是 $\Delta g(x) = (1+z)^{x+1} - (1+z)^x = z(1+z)^x$, 因此 $\Delta^n g(x) = z^n (1+z)^x$. 此时无限牛顿级数

$$g(a+x) = \sum_n \Delta^n g(a) \binom{x}{n} = (1+z)^a \sum_n \binom{x}{n} z^n$$

收敛到“正确的”值 $(1+z)^{a+x}$ (对于所有 x).

James Stirling 尝试用牛顿级数来推广阶乘函数到非整数值。首先他找到系数 S_n 使得

$$x! = \sum_n S_n \binom{x}{n} = S_0 \binom{x}{0} + S_1 \binom{x}{1} + S_2 \binom{x}{2} + \cdots \quad (5.46)$$

是对于 $x=0, x=1, x=2$ 等等的一个等式。但是他发现除 x 是非负整数外, 产生的级数不收敛, 所以他再试, 此时记

$$\ln x! = \sum_n s_n \binom{x}{n} = s_0 \binom{x}{0} + s_1 \binom{x}{1} + s_2 \binom{x}{2} + \cdots \quad (5.47)$$

现在 $\Delta(\ln x!) = \ln(x+1)! - \ln x! = \ln(x+1)$, 因此由式(5.40)

$$\begin{aligned} s_n &= \Delta^n (\ln x!)|_{x=0} \\ &= \Delta^{n-1} (\ln(x+1))|_{x=0} \\ &= \sum_k \binom{n-1}{k} (-1)^{n-1-k} \ln(k+1). \end{aligned}$$

所以系数为 $s_0 = s_1 = 0, s_2 = \ln 2, s_3 = \ln 3 - 2\ln 2 = \ln(3/4), s_4 = \ln 4 - 3\ln 3 + 3\ln 2 = \ln(32/27)$ 等等. 在此法中 Stirling 获得收敛的一个级数(虽然他没有证明它); 事实上, 他的级数对所有 $x > -1$ 收敛. 从而他能满意地求 $(1/2)!$ 的值. 习题 88 说明其余的问题.

诀窍 3: 反演

我们刚才对牛顿级数的推导的规则(5.45)的特殊情形可以下列方式再记成:

$$g(n) = \sum_k \binom{n}{k} (-1)^k f(k) \Leftrightarrow f(n) = \sum_k \binom{n}{k} (-1)^k g(k). \quad (5.48)$$

这个 f 和 g 之间的双重关系称为反演公式; 它颇像第四章中遇到的 Mobius 公式(4.56)和(4.61). 反演公式告诉我们如何来解“隐递归”, 在那里一个未知序列被嵌入一个和.

例如, $g(n)$ 可以为已知函数, $f(n)$ 可以为未知函数, 我们可找一种方式来证明 $g(n) = \sum_k \binom{n}{k} (-1)^k f(k)$. 于是式(5.48)让我们表达 $f(n)$ 为已知值的和.

我们能用本章开始处的基本方法直接证明式(5.48). 若 $g(n) = \sum_k \binom{n}{k} (-1)^k f(k)$, (对于所有 $n \geq 0$, 则

$$\begin{aligned} \sum_k \binom{n}{k} (-1)^k g(k) &= \sum_k \binom{n}{k} (-1)^k \sum_j \binom{k}{j} (-1)^j f(j) \\ &= \sum_j f(j) \sum_k \binom{n}{k} (-1)^{k+j} \binom{k}{j} \\ &= \sum_j f(j) \sum_k \binom{n}{j} (-1)^{k+j} \binom{n-j}{k-j} \end{aligned}$$

$$\begin{aligned}
 &= \sum_j f(j) \binom{n}{j} \sum_k (-1)^k \binom{n-j}{k} \\
 &= \sum_j f(j) \binom{n}{j} [n-j=0] = f(n).
 \end{aligned}$$

当然从另一方向证明是相同的, 因为 f 和 g 之间的关系是对称的.

让我们通过把它应用到“足球胜利问题”来说明式(5.48): 得胜足球队 n 个球迷的一组把他们的帽子抛向空中, 帽子随机地返回, 一顶帽子回到 n 个球迷中的一个球迷. 恰好 k 个球迷取得他们自己的返回的帽子有多少种方式 $h(n, k)$?

例如, 若 $n=4$, 帽子和球迷取名 A, B, C, D , 帽子下落的 $4! = 24$ 可能方式产生下列合适所有人的个数:

$ABCD$	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

所以 $h(4, 4)=1$, $h(4, 3)=0$, $h(4, 2)=6$, $h(4, 1)=8$, $h(4, 0)=9$.

注意到 $h(n, k)$ 是选取 k 个幸运的帽子所有者的方式数, 即 $\binom{n}{k}$, 乘上排列余下的 $n-k$ 顶帽子以致没有一顶帽子归正确的所有者的方式数, 即 $h(n-k, 0)$, 我们能确定 $h(n, k)$. 一个排列称为错排, 如果它改变了每一项目的位置, 有时把 n 个元素的错排的个数记为符号 ' ni ', 读为“ n 次阶乘”. 所以 $h(n-k, 0) = (n-k)i$, 我们得到一般公式

$$h(n, k) = \binom{n}{k} (n-k)i.$$

(次阶乘表示法不是标准的, 它显然不是一个好主意, 但是让我们试用它片刻来看是否能渐渐喜欢它. 如果 ' ni ' 不用, 我们总能采用 ' D_n ' 或某种记法.)

如果得到 ni 的闭形式, 则将解出我们的问题, 所以让我们看能找出什么东西. 有一个容易的方式来取得一个递归, 因为对于所有 k , $h(n, k)$ 的和是 n 顶帽子的排列的总数:

$$\begin{aligned}
 n! &= \sum_k h(n, k) = \sum_k \binom{n}{k} (n-k)i \\
 &= \sum_k \binom{n}{k} ki, \text{ 整数 } n \geq 0.
 \end{aligned} \tag{5.49}$$

(在最后一步中把 k 改成 $n-k$, 把 $\binom{n}{n-k}$ 改成 $\binom{n}{k}$.) 因此不明显的递归能计算所有我们想算的 $h(n, k)$:

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

例如, 这里是如何计算 $n=4$ 的行: 两个最右项是显然的, 仅有一种方式所有帽子正确地下落, 且没有 3 个球迷取得他们自己帽子的方式。(第 4 个球迷取得谁的帽子?) 当 $k=2$ 和 $k=1$ 时, 我们能用 $h(n, k)$ 的方程, 得到 $h(4, 2) = \binom{4}{2} h(2, 0) = 6 \cdot 1 = 6$ 和 $h(4, 1) = \binom{4}{1} h(3, 0) = 4 \cdot 2 = 8$. 对于 $h(4, 0)$, 不能用此方程; 更确切地说, 我们能, 但是它给出 $h(4, 0) = \binom{4}{0} h(4, 0)$, 它是真的但是无用. 取另一种方法, 我们能用关系 $h(4, 0) + 8 + 6 + 0 + 1 = 4!$ 来推出 $h(4, 0) = 9$, 这是 $4!$ 的值. 对于 $k < n$, $n!$ 相似依赖于 $k!$ 的值.

我们如何能解像式 (5.49) 的递归? 容易, 它具有式 (5.48) 的形式, $g(n) = n!$ 和 $f(k) = (-1)^k k!$. 因此它的解是

$$n! = (-1)^n \sum_k \binom{n}{k} (-1)^k k!.$$

这实际不是一个解; 它是一个和, 如果可能, 它应变成闭形式. 但是它比递归好, 和能被简化, 因为在 $\binom{n}{k}$ 中隐藏的 $k!$ 消去 $k!$, 所以让我们尝试这一点: 我们取得

$$n! = \sum_{0 \leq k \leq n} \frac{n!}{(n-k)!} (-1)^{n+k} = n! \sum_{0 \leq k \leq n} \frac{(-1)^k}{k!}. \quad (5.50)$$

余下的和迅速收敛到数 $\sum_{k \geq 0} (-1)^k / k! = e^{-1}$. 事实上, 排除在和外的一些项为

$$\begin{aligned} n! \sum_{k > n} \frac{(-1)^k}{k!} &= \frac{(-1)^{n+1}}{n+1} \sum_{k \geq 0} (-1)^k \frac{(n+1)!}{(k+n+1)!} \\ &= \frac{(-1)^{n+1}}{n+1} \left(1 - \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} - \dots \right), \end{aligned}$$

括在括号内的量位于 1 和 $1 - \frac{1}{n+2} = \frac{n+1}{n+2}$ 之间. 所以 $n!$ 和 $n!/e$ 之间的差绝对值粗略地

为 $1/n$; 更精确地说, 它位于 $1/(n+1)$ 和 $1/(n+2)$ 之间. 但是 $n!$ 是整数, 如果 $n > 0$, 它一定是 $n!/e$ 四舍五入所取得的最接近的整数. 所以得到我们所找的闭形式:

$$n! = \left[\frac{n!}{e} + \frac{1}{2} \right] + [n = 0]. \quad (5.51)$$

这就是无球迷取得返回的正确的帽子的方式数. 当 n 大时, 知道这发生的概率是很有意义的. 若我们设 $n!$ 种排列的每一种都是等可能的, 因为帽子抛得极高, 此概率为

$$\frac{n!}{n!} = \frac{n!/e + O(1)}{n!} \sim \frac{1}{e} = 0.367\ldots$$

所以当 n 取大时, 所有帽子误放的概率差不多为 37%.

顺便提一句, 次阶乘的递归(5.49)恰好与式(5.46)相同, 当 Stirling 尝试推广阶乘函数时, 这是他考虑的第一个递归. 因此 $S_k = k!$. 这些系数这样大, 对于非整数 x 无穷级数式(5.46)发散是并不奇怪的.

在结束此问题之前, 让我们简要地看两个有趣的型式, 即抓住小 $h(n, k)$ 的表中显示的型式. 首先, 看来所有 0 对角线下面的数 1, 3, 6, 10, 15, ... 是三角形数. 这个观察容易证明, 因为那些表的项目是 $h(n, n-2)$, 我们有

$$h(n, n-2) = \binom{n}{n-2} 2! = \binom{n}{2}.$$

还看出在前两列中的数之差为 ± 1 . 是否总是对的? 是,

$$\begin{aligned} h(n, 0) - h(n, 1) &= n! - n(n-1)! \\ &= \left(n! \sum_{0 \leq k \leq n} \frac{(-1)^k}{k!} \right) - \left(n(n-1)! \sum_{0 \leq k \leq n-1} \frac{(-1)^k}{k!} \right) \\ &= n! \frac{(-1)^n}{n!} = (-1)^n. \end{aligned}$$

换句话说, $n! = n(n-1)! + (-1)^n$. 这是比以前得到的更简单的错排数的递归.

现在让我们转换另外一些东西. 如果应用反演到式(5.41)中引出公式

$$\sum_k \binom{n}{k} \frac{(-1)^k}{x+k} = \frac{1}{x} \binom{x+n}{n}^{-1},$$

我们找到

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

这是有趣的, 但实际不是新的. 如果我们在 $\binom{x+k}{k}$ 中求反上指数, 只不过是再显示等式(5.33).

5.4 母函数

我们现在达到了本书中极重要的概念，母函数的概念。我们希望以某种方式处理无限序列 $\langle a_0, a_1, a_2, \dots \rangle$ 能方便地表为辅助变量 z 的幂级数，

$$A(z) = a_0 + a_1 z + a_2 z^2 + \dots = \sum_{k \geq 0} a_k z^k. \quad (5.52)$$

用字母 z 作为辅助变量的名称是合适的，因为我们常想象 z 为复数。复变数理论在它的公式中约定用 ' z '，幂级数(解析函数或全纯函数)是这种理论的主要内容。

在后续几章中，我们将看到许多母函数。实际上，第七章完全专心讨论它们。现在的目的是简要介绍基本概念，并展示母函数和二项系数研究的关系。

因为母函数是表示整个无限序列的简单的量，所以它是有用的。我们常能以这样的方式来解问题，首先设置一个或多个母函数，然后环绕这些函数极简单的处理直到我们十分了解它们，最后再来看系数。有一点儿运气，我们将充分了解函数来得到我们需要知道的关于它们的系数的内容。

若 $A(z)$ 是任何幂级数 $\sum_{k \geq 0} a_k z^k$ ，我们发现记

$$[z^n] A(z) = a_n \quad (5.53)$$

是方便的；换句话说，用 $[z^n] A(z)$ 表示 $A(z)$ 中的 z^n 的系数。

设 $A(z)$ 是像式 (5.52) 中的 $\langle a_0, a_1, a_2, \dots \rangle$ 的母函数，且设 $B(z)$ 是另一个序列 $\langle b_0, b_1, b_2, \dots \rangle$ 的母函数，则乘积 $A(z)B(z)$ 是幂级数

$$\begin{aligned} (a_0 + a_1 z + a_2 z^2 + \dots)(b_0 + b_1 z + b_2 z^2 + \dots) \\ = a_0 b_0 + (a_0 b_1 + a_1 b_0)z + (a_0 b_2 + a_1 b_1 + a_2 b_0)z^2 + \dots \end{aligned}$$

此乘积中 z^n 的系数是

$$a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0 = \sum_{k=0}^n a_k b_{n-k}.$$

所以如果我们希望计算具有一般形式

$$c_n = \sum_{k=0}^n a_k b_{n-k} \quad (5.54)$$

的任何和，且如果知道母函数 $A(z)$ 和 $B(z)$ ，我们得到

$$c_n = [z^n] A(z)B(z).$$

由式(5.54)定义的序列 $\langle c_n \rangle$ 称为序列 $\langle a_n \rangle$ 和 $\langle b_n \rangle$ 的卷积, 通过形成所有下标合计为指定量的乘积的和来“卷”两个序列。前面一段的要点是序列的卷积对应于它们的母函数的乘积。

母函数提供了有力的方法来发现和/或证明等式。例如, 二项定理告诉我们 $(1+z)^r$ 是序列 $\langle \binom{r}{0}, \binom{r}{1}, \binom{r}{2}, \dots \rangle$ 的母函数:

$$(1+z)^r = \sum_{k \geq 0} \binom{r}{k} z^k.$$

类似,

$$(1+z)^s = \sum_{k \geq 0} \binom{s}{k} z^k.$$

如果把这些乘在一起, 我们取得另一个母函数:

$$(1+z)^r (1+z)^s = (1+z)^{r+s}.$$

现在出现了巧妙的结果: 使此方程两边的 z^n 的系数相等给出

$$\sum_{k=0}^n \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}.$$

这是我们已发现的 Vandermonde 卷积, 式(5.27)!

这是合宜的和容易的, 让我们尝试另一个。此时用 $(1-z)^r$, 它是序列 $\langle (-1)^n \binom{r}{n} \rangle = \langle \binom{r}{0}, -\binom{r}{1}, \binom{r}{2}, \dots \rangle$ 的母函数。乘 $(1+z)^r$ 给出另一个母函数, 我们知道它的系数:

$$(1-z)^r (1+z)^r = (1-z^2)^r.$$

现在使 z^n 的系数相等给出方程

$$\sum_{k=0}^n \binom{r}{k} \binom{r}{n-k} (-1)^k = (-1)^{n/2} \binom{r}{n/2} [n \text{ 偶}]. \quad (5.55)$$

我们应对一个或两个小的情形检验此式。当 $n=3$ 时, 结果为

$$\binom{r}{0} \binom{r}{3} - \binom{r}{1} \binom{r}{2} + \binom{r}{2} \binom{r}{1} - \binom{r}{3} \binom{r}{0} = 0.$$

每个正项被对应的负项消去。每当 n 为奇数时相同情况发生, 此时和不十分有趣, 但是当 n 为偶数时, 比如说 $n=2$, 则取得一个不同于 Vandermonde 卷积的非平凡和:

$$\binom{r}{0} \binom{r}{2} - \binom{r}{1} \binom{r}{1} + \binom{r}{2} \binom{r}{0} = 2 \binom{r}{2} - r^2 = -r.$$

所以当 $n=2$ 时式(5.55)顺利通过检验, 原来式(5.30)是新等式(5.55)的特殊情形。

二项系数也显示出了一些其他母函数, 最值得注意的下列重要等式, 它的下指数保持固定, 上指数变动:

$$\frac{1}{(1-z)^{n+1}} = \sum_{k \geq 0} \binom{n+k}{n} z^k, \quad \text{整数 } n \geq 0, \quad (5.56)$$

$$\frac{z^n}{(1-z)^{n+1}} = \sum_{k \geq 0} \binom{k}{n} z^k, \quad \text{整数 } n \geq 0, \quad (5.57)$$

这里第二个等式就是第一个等式乘 z^n , 也就是说, “向右移” n 个位置。第一个等式就是有点隐蔽的二项定理的特殊情形: 如果用式(5.13)展开 $(1-z)^{-n-1}$, z^k 的系数为 $\binom{-n-1}{k} \cdot (-1)^k$, 求反上指数我们能再记为 $\binom{k+n}{k}$ 或 $\binom{n+k}{n}$ 。明确地指明这些特殊情形是值得的, 因为它们在应用中经常出现。

当 $n=0$ 时我们取得特殊情形的一个特殊情况, 几何级数

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \cdots = \sum_{k \geq 0} z^k.$$

这是序列 $\langle 1, 1, 1, \cdots \rangle$ 的母函数, 它特别有用, 因为任何其他序列与这个序列的卷积是和的序列: 当对于所有 k , $b_k = 1$ 时, 式(5.54)化为

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

所以如果 $A(z)$ 是被加数 $\langle a_0, a_1, a_2, \cdots \rangle$ 的母函数, 则 $A(z)/(1-z)$ 是和 $\langle c_0, c_1, c_2, \cdots \rangle$ 的母函数。

用帽子和足球迷间联系中的反演来解错排问题, 也能以有趣的方式用母函数再解它。基本的递归

$$n! = \sum_k \binom{n}{k} (n-k)!$$

能使它成为一个卷积的形式, 如果把 $\binom{n}{k}$ 展开为阶乘, 再两边除 $n!$,

$$1 = \sum_{k=0}^n \frac{1}{k!} \frac{(n-k)!}{(n-k)!}.$$

序列 $\langle 1/0!, 1/1!, 1/2!, \cdots \rangle$ 的母函数为 e^z , 因此如果设

$$D(z) = \sum_{k \geq 0} \frac{k!}{k!} z^k,$$

卷积 / 递归告诉我们

$$\frac{1}{1-z} = e^z D(z).$$

解 $D(z)$ 给出

$$D(z) = \frac{1}{1-z} e^{-z} = \frac{1}{1-z} \left(\frac{1}{0!} z^0 - \frac{1}{1!} z^1 + \frac{1}{2!} z^2 - \dots \right).$$

使 z^n 的系数相等, 现在告诉我们

$$\frac{n!}{n!} = \sum_{k=0}^n \frac{(-1)^k}{k!},$$

这是前面用反演导出的公式.

到目前为止, 母函数的钻研给出了一些结果的巧妙的证明, 这些结果我们已经知道是由较麻烦的方法推出来的. 但是除式(5.55)外还未用母函数获得任何新的结果. 现在我们准备一些新的和较意外的东西. 有两族幂级数产生一类特别有意义的二项系数等式: 让我们定义广义二项级数 $\mathcal{B}_t(z)$ 和广义指数级数 $\mathcal{E}_t(z)$ 如下:

$$\mathcal{B}_t(z) = \sum_{k \geq 0} (tk)^{\overline{k-1}} \frac{z^k}{k!}; \quad \mathcal{E}_t(z) = \sum_{k \geq 0} (tk+1)^{\overline{k-1}} \frac{z^k}{k!}. \quad (5.58)$$

可以证明这些函数满足等式

$$\mathcal{B}_t(z)^{1-t} - \mathcal{B}_t(z)^{-t} = z; \quad \mathcal{E}_t(z)^{-t} \ln \mathcal{E}_t(z) = z. \quad (5.59)$$

由特殊情形 $t=0$, 得到

$$\mathcal{B}_0(z) = 1+z; \quad \mathcal{E}_0(z) = e^z.$$

这就说明为什么具有参数 t 的级数称为“广义”二项级数和“广义”指数级数.

对于所有实数 r , 下列两对等式成立:

$$\mathcal{B}_t(z)' = \sum_{k \geq 0} \binom{tk+r}{k} \frac{r}{tk+r} z^k, \quad \mathcal{E}_t(z)' = \sum_{k \geq 0} r \frac{(tk+r)^{\overline{k-1}}}{k!} z^k; \quad (5.60)$$

$$\frac{\mathcal{B}_t(z)'}{1-t+t\mathcal{B}_t(z)^{-1}} = \sum_{k \geq 0} \binom{tk+r}{k} z^k, \quad \frac{\mathcal{E}_t(z)'}{1-zt\mathcal{E}_t(z)'} = \sum_{k \geq 0} \frac{(tk+r)^{\overline{k}}}{k!} z^k. \quad (5.61)$$

(当 $tk+r=0$ 时, 我们必须仔细一点来说明 z^k 的系数, 每个系数是 r 的多项式. 例如, $\mathcal{E}_t(z)'$ 的常数项为 $r(0+r)^{-1}$, 且甚至当 $r=0$ 时这是等于 1.)

由于方程(5.60)和(5.61)对所有 r 成立, 当我们把对应于不同幂 r 和 s 的级数相乘, 取得十分一般的等式. 例如,

$$\begin{aligned}\mathcal{B}_t(z)^r \frac{\mathcal{B}_t(z)^s}{1-t-t\mathcal{B}_t(z)^{-1}} &= \sum_{k \geq 0} \binom{tk+r}{k} \frac{r}{tk+r} z^k \sum_{j \geq 0} \binom{tj+s}{j} z^j \\ &= \sum_{n \geq 0} z^n \sum_{k \geq 0} \binom{tk+r}{k} \frac{r}{tk+r} \binom{t(n-k)+s}{n-k}.\end{aligned}$$

此幂级数一定等于

$$\frac{\mathcal{B}_t(z)^{r+s}}{1-t-t\mathcal{B}_t(z)^{-1}} = \sum_{n \geq 0} \binom{tn+r+s}{n} z^n;$$

因此我们能使 z^n 的系数相等, 取得等式

$$\sum_k \binom{tk+r}{k} \binom{t(n-k)+s}{n-k} \frac{r}{tk+r} = \binom{tn+r+s}{n}, \text{ 整数 } n,$$

对于所有实数 r, s 和 t 成立。当 $t=0$ 时, 此等式化为 Vandermonde 卷积。(如果在此公式中意外地发生 $tk+r$ 等于零, 分母因子 $tk+r$ 应考虑与二项系数的分子中的 $tk+r$ 消去。等式两边是 r, s 和 t 的多项式。)当我们用 $\mathcal{B}_t(z)^s$ 乘 $\mathcal{B}_t(z)^r$ 时相似等式成立, 等等; 表 5.5 示出一些结果。

表 5.5 一般卷积等式, 对于整数 $n \geq 0$ 成立

$$\sum_k \binom{tk+r}{k} \binom{tn-tk+s}{n-k} \frac{r}{tk+r} = \binom{tn+r+s}{n}. \quad (5.62)$$

$$\begin{aligned}\sum_k \binom{tk+r}{k} \binom{tn-tk+s}{n-k} \frac{r}{tk+r} \cdot \frac{s}{tn-tk+s} \\ = \binom{tn+r+s}{n} \frac{r+s}{tn+r+s}.\end{aligned} \quad (5.63)$$

$$\sum_k \binom{n}{k} (tk+r)^k (tn-tk+s)^{n-k} \frac{r}{tk+r} = (tn+r+s)^n. \quad (5.64)$$

$$\begin{aligned}\sum_k \binom{n}{k} (tk+r)^k (tn-tk+s)^{n-k} \frac{r}{tk+r} \cdot \frac{s}{tn-tk+s} \\ = (tn+r+s)^n \frac{r+s}{tn+r+s}.\end{aligned} \quad (5.65)$$

我们学到查看一般结果的特殊情形一般是一个好主意。例如, 若置 $t=1$, 发生什么? 广义二项级数 $\mathcal{B}_1(z)$ 是十分简单的, 它就是

$$\mathcal{B}_1(z) = \sum_{k \geq 0} z^k = \frac{1}{1-z},$$

所以 $\mathcal{B}_1(z)$ 未给出 Vandermonde 卷积中不知道的东西。但是 $\mathcal{E}_1(z)$ 是重要的函数,

$$\mathcal{E}(z) = \sum_{k \geq 0} (k+1)^{k-1} \frac{z^k}{k!} = 1 + z + \frac{3}{2}z^2 + \frac{8}{3}z^3 + \frac{125}{24}z^4 + \cdots \quad (5.66)$$

这是在前面没有见过的, 它满足基本等式

$$\mathcal{E}(z) = e^{z\mathcal{E}(z)}. \quad (5.67)$$

在许多应用中出现此函数, Eisenstein 首先研究了它^[75].

广义二项级数的特殊情形 $t=2$ 和 $t=-1$ 是特别有趣的情形, 因为它们的系数反复出现在具有递归结构的问题中。所以为了将来参考而明确地展示这些级数是有用的。

$$\begin{aligned} \mathcal{B}_2(z) &= \sum_k \binom{2k}{k} \frac{z^k}{1+k} \\ &= \sum_k \binom{2k+1}{k} \frac{z^k}{1+2k} = \frac{1 - \sqrt{1-4z}}{2z} \end{aligned} \quad (5.68)$$

$$\begin{aligned} \mathcal{B}_{-1}(z) &= \sum_k \binom{1-k}{k} \frac{z^k}{1-k} \\ &= \sum_k \binom{2k-1}{k} \frac{(-z)^k}{1-2k} = \frac{1 + \sqrt{1+4z}}{2}. \end{aligned} \quad (5.69)$$

$$\mathcal{B}_2(z)' = \sum_k \binom{2k+r}{k} \frac{r}{2k+r} z^k. \quad (5.70)$$

$$\mathcal{B}_{-1}(z)' = \sum_k \binom{r-k}{k} \frac{r}{r-k} z^k. \quad (5.71)$$

$$\frac{\mathcal{B}_2(z)'}{\sqrt{1-4z}} = \sum_k \binom{2k+r}{k} z^k. \quad (5.72)$$

$$\frac{\mathcal{B}_{-1}(z)'^{+1}}{\sqrt{1+4z}} = \sum_k \binom{r-k}{k} z^k. \quad (5.73)$$

$\mathcal{B}_2(z)$ 的系数 $\binom{2n}{n} \frac{1}{n+1}$ 称为 Catalan 数 C_n , 因为在 1830 年期间 Eugène Catalan 写了一篇关于它们的有影响的论文^[46]。开始的序列如下:

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

$\mathcal{B}_{-1}(z)$ 的系数本质上是相同的, 但是在开始处有一个额外的 1, 其他数正负交错: $\langle 1, 1, -1, 2, -5, 14, \dots \rangle$ 。于是 $\mathcal{B}_{-1}(z) = 1 + z \mathcal{B}_2(-z)$, 我们还得到

$$\mathcal{B}_{-1}(z) = \mathcal{B}_2(-z)^{-1}.$$

让我们导出式(5.72)和(5.73)的一个重要的结论, 即函数 $\mathcal{B}_{-1}(z)$ 和 $\mathcal{B}_2(-z)$ 之间进一步联系的一个关系, 来结束本节:

$$\frac{\mathcal{B}_{-1}(z)^{n+1} - (-z)^{n+1} \mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} = \sum_{k \leq n} \binom{n-k}{k} z^k.$$

此式成立是因为在 $(-z)^{n+1} \mathcal{B}_2(-z)^{n+1} / \sqrt{1+4z}$ 中的 z^k 的系数 (当 $k > n$ 时) 为

$$\begin{aligned} [z^k] \frac{(-z)^{n+1} \mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} &= (-1)^{n+1} [z^{k-n-1}] \frac{\mathcal{B}_2(-z)^{n+1}}{\sqrt{1+4z}} \\ &= (-1)^{n+1} (-1)^{k-n-1} [z^{k-n-1}] \frac{\mathcal{B}_2(z)^{n+1}}{\sqrt{1-4z}} \\ &= (-1)^k \binom{2(k-n-1)+n+1}{k-n-1} \\ &= (-1)^k \binom{2k-n-1}{k-n-1} = (-1)^k \binom{2k-n-1}{k} \\ &= \binom{n-k}{k} = [z^k] \frac{\mathcal{B}_{-1}(z)^{n+1}}{\sqrt{1+4z}}. \end{aligned}$$

一些项彼此消去。现在我们能用式(5.68)和(5.69)来获得闭形式

$$\sum_{k \leq n} \binom{n-k}{k} z^k = \frac{1}{\sqrt{1+4z}} \left(\left(\frac{1+\sqrt{1+4z}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{1+4z}}{2} \right)^{n+1} \right),$$

整数 $n \geq 0$. (5.74)

(特殊情形 $z=-1$ 在 5.2 节的问题 3 中提出。由于数 $\frac{1}{2}(1 \pm \sqrt{3})$ 是单位的第 6 个根, 在那个问题中我们看到和 $\sum_{k \leq n} \binom{n-k}{k} (-1)^k$ 有周期的特性。) 我们相似地能把式(5.70)和(5.71)结合起来消去大系数, 取得

$$\sum_{k \leq n} \binom{n-k}{k} \frac{n}{n-k} z^k = \left(\frac{1+\sqrt{1+4z}}{2} \right)^n + \left(\frac{1-\sqrt{1+4z}}{2} \right)^n, \text{ 整数 } n > 0. \quad (5.75)$$

5.5 超几何函数

我们应用到二项系数的方法, 当它们处理时十分有效, 但是我们一定承认常出现的方

法是特定的, 比起技巧来说更像诀窍。当我们处理一个问题时, 常有许多方向来进行, 并可能忙得团团转。二项系数像是反复无常的, 它们的外表容易改变。所以自然问到是否没有某个统一的原理将有规则地同时处理各式各样的二项系数求和。幸运的是, 回答为有。统一原理基于称为超几何级数的无限和的理论。

Euler, Gauss 和 Riemann 在很久以前就着手超几何级数的研究; 事实上, 这样的级数仍然是值得重视的研究的主题。但是超几何的表示法有点令人生畏, 为了用它需花一点儿时间。

一般超几何级数是具有 $m+n$ 个参数的 z 的幂级数, 根据上升阶乘幂它被定义如下:

$$F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) = \sum_{k \geq 0} \frac{a_1^{\bar{k}} \cdots a_m^{\bar{k}}}{b_1^{\bar{k}} \cdots b_n^{\bar{k}}} \frac{z^k}{k!} \quad (5.76)$$

为了避免用零除, 没有一个 b 可以为零或一个负整数。除了这一点, a 和 b 可以是我们希望的任何数。记法 ' $F(a_1, \dots, a_m; b_1, \dots, b_n; z)$ ' 也用来作为两行形式(5.76)的替换记法, 由于有时一行形式印刷上处理较好。 a 称为上参数, 它们出现在 F 的项的分子中。 b 称为下参数, 它们出现在分母。最后的量 z 称为自变量。

一般参考书常用 ' ${}_mF_n$ ' 替代 ' F ' 作为有 m 个上参数和 n 个下参数的超几何的名称。但是如果不得反复写出它们, 额外的下标使公式零乱且消耗时间。我们能数出有多少个参数, 所以通常不需额外附加的不必要的累赘。

许多重要函数作为一般超几何的特殊情形出现: 实际上, 这就是为什么超几何是如此有力的原因。例如, 当 $m=n=0$ 时出现最简单的情形: 没有参数, 我们取得熟悉的级数

$$F(|z) = \sum_{k \geq 0} \frac{z^k}{k!} = e^z.$$

当 m 或 n 为零时, 记法实际看来有点不确定, 为了避免这一点我们可在上下加一个额外的 '1',

$$F\left(\begin{matrix} 1 \\ 1 \end{matrix} \middle| z\right) = e^z.$$

一般, 如果消去分子和分母中同时出现的一个参数, 或者插入两个相同的参数, 则不改变函数。

下一个最简单情形有 $m=1$, $a_1=1$ 和 $n=0$; 我们改变参数为 $m=2$, $a_1=a_2=1$, $n=1$ 和 $b_1=1$, 以至 $n>0$ 。此级数也是原来熟悉的结果, 因为 $1^{\bar{k}}=k!$:

$$F\left(\begin{matrix} 1, 1 \\ 1 \end{matrix} \middle| z\right) = \sum_{k \geq 0} z^k = \frac{1}{1-z}.$$

它是熟悉的几何级数: $F(a_1, \dots, a_m; b_1, \dots, b_n; z)$ 称为超几何, 因为它包含几何级数 $F(1, 1; 1; z)$ 作为一个十分特别的情形。

事实上, 运用式(5.56)一般情形 $m=1$ 和 $n=0$ 容易相加为闭形式

$$F\left(\begin{matrix} a, & 1 \\ & 1 \end{matrix} \middle| z\right) = \sum_{k \geq 0} a^{\bar{k}} \frac{z^k}{k!} = \sum_k \binom{a+k-1}{k} z^k = \frac{1}{(1-z)^a}. \quad (5.77)$$

如果用 $-a$ 替换 a , $-z$ 替换 z , 我们取得二项定理

$$F\left(\begin{matrix} -a, & 1 \\ & 1 \end{matrix} \middle| -z\right) = (1+z)^a.$$

一个负整数作为上参数引出无限级数变成有限, 因为每当 $k > a \geq 0$ 且 a 是整数时 $(-a)^{\bar{k}} = 0$.

一般情形 $m=0$, $n=1$ 是另一个著名的级数, 但是它不是离散数学文献中熟知的:

$$F\left(\begin{matrix} & 1 \\ b, & 1 \end{matrix} \middle| z\right) = \sum_{k \geq 0} \frac{(b-1)!}{(b-1+k)!} \frac{z^k}{k!} = I_{b-1}(2\sqrt{z}) \frac{(b-1)!}{z^{(b-1)/2}}. \quad (5.78)$$

这个函数 I_{b-1} 称为阶 $b-1$ 的“修正 Bessel 函数”。特殊情形 $b=1$ 给出 $F\left(\begin{matrix} & 1 \\ 1, & 1 \end{matrix} \middle| z\right) = I_0(2\sqrt{z})$, 它是有趣的级数 $\sum_{k \geq 0} z^k / k!^2$.

特殊情形 $m=n=1$ 称为“合流超几何级数”, 常用字母 M 来表示:

$$F\left(\begin{matrix} a \\ b \end{matrix} \middle| z\right) = \sum_{k \geq 0} \frac{a^{\bar{k}}}{b^{\bar{k}}} \frac{z^k}{k!} = M(a, b, z). \quad (5.79)$$

Ernst Kummer 引入这个工程中有重要应用的函数。

到现在为止有些人感到奇怪, 为什么不讨论无穷级数(5.76)的收敛性。如果用 z 仅作为一个形式符号, 回答是我们可忽略收敛性。不难验证, 如果系数 α_k 在域中, 形式 $\sum_{k \geq n} \alpha_k z^k$ 的外形的无限和形成一个域。我们能对这样的形式和加, 减, 乘, 除, 微分和作函数复合, 而不必担心收敛性; 我们推出的任何等式仍将形式地真。例如, 对于任何非零 z , 超几何 $F\left(\begin{matrix} 1, & 1 \\ & 1 \end{matrix} \middle| z\right) = \sum_{k \geq 0} k! z^k$ 不收敛, 在第七章中还将看到我们仍能用它来解问题。另一方面, 每当用一个特殊数值替换 z 时, 我们的确必须肯定无穷和是完全定义的。

在复杂情况中接着研究的内容实际上是大家非常满意的超几何级数。事实上, 直到 1870 年左右, 当每样东西被推广到任意的 m 和 n 时, 它就成为超几何级数。下面这个有两个上参数和一个下参数的级数

$$F\left(\begin{matrix} a, & b \\ & c \end{matrix} \middle| z\right) = \sum_{k \geq 0} \frac{a^{\bar{k}} b^{\bar{k}}}{c^{\bar{k}} k!} \quad (5.80)$$

常称为高斯的超几何级数。虽然 Euler^[95]和 Pfaff^[232]已发现了一些关于它的值得注意的东西, 但是它的许多精巧的性质首先是由高斯在他的 1812 年的博士论文中证明的^[116]。它的

重要的特殊情形之一是

$$\begin{aligned}\ln(1+z) &= zF\left(\begin{matrix} 1, 1 \\ 2 \end{matrix} \middle| -z\right) = z \sum_{k \geq 0} \frac{k!k!}{(k+1)!} \frac{(-z)^k}{k!} \\ &= z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots\end{aligned}$$

注意, $z^{-1}\ln(1+z)$ 是超几何函数, 但是 $\ln(1+z)$ 不能是超几何的, 因为当 $z=0$ 时超几何级数总有值 1.

到现在为止, 除了提供取名的理由之外, 对我们来说实际上超几何还未完成任何事情. 但是我们已看到几种很不同的函数都能视为超几何, 这将是以下感兴趣的主要点. 我们将看到一大类和能以“标准的”方式记为超几何级数, 因此我们将有关于二项系数结论的一个好的资料体系.

什么样的级数是超几何的? 如果查看相继项的比, 容易回答这个问题:

$$F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) = \sum_{k \geq 0} t_k, \quad t_k = \frac{a_1^k \cdots a_m^k z^k}{b_1^k \cdots b_n^k k!}.$$

第一项为 $t_0 = 1$, 其他项有

$$\begin{aligned}\frac{t_{k+1}}{t_k} &= \frac{a_1^{k+1} \cdots a_m^{k+1}}{a_1^k \cdots a_m^k} \frac{b_1^k \cdots b_n^k}{b_1^{k+1} \cdots b_n^{k+1}} \frac{k!}{(k+1)!} \frac{z^{k+1}}{z^k} \\ &= \frac{(k+a_1) \cdots (k+a_m)z}{(k+b_1) \cdots (k+b_n)(k+1)}\end{aligned}\tag{5.81}$$

给出的比. 这是 k 的有理函数, 也就是说, k 的多项式的商. 任何 k 的有理函数能在复数上分解因子且把它变成这种形式. 分子中的 a 是分子中多项式的根的负数, 分母中的 b 是分母中多项式的根的负数. 如果分母还未包含特殊因子 $(k+1)$, 可在分子和分母中都包含 $(k+1)$. 剩下一个常数因子, 称之为 z . 所以超几何级数正好是这样的级数, 它的第一项是 1, 它的项比 t_{k+1}/t_k 是 k 的有理函数.

例如, 假设给定一个具有项比

$$\frac{t_{k+1}}{t_k} = \frac{k^2 + 7k + 10}{4k^2 + 1}$$

(k 的有理函数) 的无穷级数. 分子多项式分解成两个因子 $(k+2)(k+5)$, 分母为 $4(k+i/2)(k-i/2)$. 因为分母缺要求的因子 $(k+1)$, 我们把项比记为

$$\frac{t_{k+1}}{t_k} = \frac{(k+2)(k+5)(k+1)(1/4)}{(k+i/2)(k-i/2)(k+1)},$$

我们能读出结果：给定级数为

$$\sum_{k \geq 0} t_k = t_0 F\left(\begin{matrix} 2, 5, 1 \\ i/2, -i/2 \end{matrix} \middle| 1/4\right).$$

于是，我们有一个一般的方法来寻找给定量 S 的超几何表示，当这样的一种表示是可能的时候：首先把 S 记为无穷级数，它的第一项是非零的。我们选取一个记法以致级数为 $t_0 \neq 0$ 的 $\sum_{k \geq 0} t_k$ ，然后计算 t_{k+1}/t_k 。如果项比不是 k 的有理函数，则运气不好。否则我们把它表示为形式(5.81)；这就给出了参数 $a_1, \dots, a_m, b_1, \dots, b_n$ 和一个自变量 z ，使得 $S = t_0 F(a_1, \dots, a_m; b_1, \dots, b_n; z)$ 。

如果希望强调项比的重要性，高斯超几何级数可记为递归分解因子形式

$$F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| z\right) = 1 + \frac{ab}{1c} z \left(1 + \frac{a+1}{2} \frac{b+1}{c+1} z \left(1 + \frac{a+2}{3} \frac{b+2}{c+2} z (1 + \dots)\right)\right).$$

现在让我们再列出本章前面引出的二项系数等式，把它们表示为超几何。例如，让我们解决类似求和定律，

$$\sum_{k \leq n} \binom{r+k}{k} = \binom{r+n+1}{n}, \text{ 整数 } n,$$

看来像超几何表示法，我们需把和记为从 $k=0$ 开始的无穷级数，所以用 $n-k$ 替换 k ：

$$\sum_{k \geq 0} \binom{r+n-k}{n-k} = \sum_{k \geq 0} \frac{(r+n-k)!}{r!(n-k)!} = \sum_{k \geq 0} t_k.$$

这个级数形式上为无穷，但实际上有限，因为在分母中的 $(n-k)!$ ，当 $k > n$ 时使 $t_k = 0$ 。（在后面将看到对于所有 x ，定义了 $1/x!$ ，且当 x 是负整数时 $1/x! = 0$ 。但是现在让我们在获得更多超几何经验之前冒失地忽略这样的技巧。）项比为

$$\begin{aligned} \frac{t_{k+1}}{t_k} &= \frac{(r+n-k-1)!r!(n-k)!}{r!(n-k-1)!(r+n-k)!} = \frac{n-k}{r+n-k} \\ &= \frac{(k+1)(k-n)(1)}{(k-n-r)(k+1)}, \end{aligned}$$

而且 $t_0 = \binom{r+n}{n}$ ，因此类似求和定律等价于超几何等式

$$\binom{r+n}{n} F\left(\begin{matrix} 1, -n \\ -n-r \end{matrix} \middle| 1\right) = \binom{r+n+1}{n}.$$

通过用 $\binom{r+n}{n}$ 除给出稍简单的形式，

$$F\left(\begin{matrix} 1, -n \\ -n-r \end{matrix} \middle| 1\right) = \frac{r+n+1}{r+1}, \text{ 如果 } \binom{r+n}{n} \neq 0. \quad (5.82)$$

让我们做另外一个。等式(5.16)

$$\sum_{k \leq m} \binom{r}{k} (-1)^k = (-1)^m \binom{r-1}{m}, \text{ 整数 } m,$$

的项比在 $m-k$ 替换 k 之后为 $(k-m)/(r-m+k+1) = (k+1)(k-m)(1)/(k-m+r+1)(k+1)$; 因此式(5.16)给出一个闭形式

$$F\left(\begin{matrix} 1, -m \\ -m+r+1 \end{matrix} \middle| 1\right).$$

这本质上与式(5.82)左边的超几何函数相同, 但是用 m 代替 n , 用 $r+1$ 代替 $-r$. 所以能从式(5.9)的超几何形式(5.82)引出等式(5.16). (我们找到它并不奇怪, 容易用式(5.9)来证明式(5.16).)

在继续进行下去之前, 我们该考虑退化情形, 因为当一个下参数为零或一个负整数时, 超几何是没有定义的。当 r 和 n 是正整数时通常应用类似求和等式; 但另一方面 $-n-r$ 是负整数, 超几何式(5.76)无定义。于是我们如何能考虑式(5.82)为合法呢? 回答是当 $\varepsilon \rightarrow 0$ 时可取 $F\left(\begin{matrix} 1, -n \\ -n-r+\varepsilon \end{matrix} \middle| 1\right)$ 的极限。

在本章后面我们将较仔细地查看这样的事情, 但是现在我们就能认识到某些分母会完全失败。然而, 有趣的是, 我们尝试超几何表达最早的和原来是退化的。

在推导式(5.82)中另一个可能的恼火点是把 $\binom{r+n-k}{n-k}$ 表达为 $(r+n-k)!/r!(n-k)!$. 当 r 是负整数时此表达式失效, 因为如果定律

$$0! = 0 \cdot (-1) \cdot (-2) \cdot \cdots \cdot (-m+1) \cdot (-m)!$$

成立, 则 $(-m)$ 必须为 ∞ . 另一方面, 我们需通过考虑当 $\varepsilon \rightarrow 0$ 时 $r+\varepsilon$ 的极限来接近整数结果。

但是仅当 r 是整数时定义了阶乘表示 $\binom{r}{k} = r!/k!(r-k)!$. 如果要用超几何有效地处理, 我们需要对所有复数定义的阶乘函数。幸好有这样一个函数, 且能以许多方式定义它。这里是一种最有用的 $z!$ 的定义, 实际为 $1/z!$ 的定义:

$$\frac{1}{z!} = \lim_{n \rightarrow \infty} \binom{n+z}{n} n^{-z}. \quad (5.83)$$

(见习题 21. Euler^[81]22 岁时发现了它。)式(5.83)能证对于所有复数 z 极限存在, 且仅当 z 是负整数时它是零。另一个有意义的定义是

$$z! = \int_0^{\infty} t^z e^{-t} dt, \text{ 若 } \Re_z > -1. \quad (5.84)$$

这个积分仅当 z 的实部大于 -1 时存在, 但是我们能公式

$$z! = z(z-1)! \quad (5.85)$$

推广定义到所有复数 z (除负整数外), 还有来自式(5.47)中的 $\ln z!$ 的 Stirling 的插值的另一个定义。所有这些处理导致相同的阶乘函数。

有一个称 γ 函数的十分熟悉的函数, 它和通常阶乘的联系有点像升幂和降幂的联系。标准参考书常同时用阶乘和 γ 函数, 如果必须用下列公式, 则它们之间的转换是方便的:

$$\Gamma(z+1) = z!, \quad (5.86)$$

$$(-z)!\Gamma(z) = \frac{\pi}{\sin \pi z}. \quad (5.87)$$

当 z 和 w 是任意复数时, 能用这些广义阶乘来定义广义阶乘幂:

$$z^{\overline{w}} = \frac{z!}{(z-w)!}, \quad (5.88)$$

$$z^{\underline{w}} = \frac{\Gamma(z+w)}{\Gamma(z)}. \quad (5.89)$$

仅有的附带条件是当这些公式给出 ∞/∞ 时我们一定要用适当的极限值。(公式不会有 $0/0$, 因为阶乘和 γ 函数值不会为零。)当 z 和 w 是任何不管什么样的复数时, 二项系数能记为

$$\binom{z}{w} = \lim_{\zeta \rightarrow z} \lim_{\omega \rightarrow w} \frac{\zeta!}{\omega!(\zeta-\omega)!}. \quad (5.90)$$

准备了广义阶乘工具, 把前面引出的等式化成它们的超几何实体, 能返回到我们的目标。正如我们所期望的那样, 二项定理(5.13)原来恰好是式(5.77)。所以接着尝试最有趣的等式是 Vandermonde 卷积(5.27):

$$\sum_k \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}, \quad \text{整数 } n.$$

这里第 k 项是

$$t_k = \frac{r!}{(r-k)!k!} \frac{s!}{(s-n+k)!(n-k)!},$$

且在上述表达式中不再怕用广义阶乘。每当 t_k 包含像 $(\alpha+k)!$ 的因子时, k 之前有一加号, 依据式(5.85)我们在项比 t_{k+1}/t_k 中取得 $(\alpha+k+1)!/(\alpha+k)! = k+\alpha+1$; 这提供参数 ' $\alpha+1$ ' 到对应的超几何, 如果 $(\alpha+k)!$ 是在 t_k 的分子中, 如同一个上参数, 否则如同一个下参数。类似, 像 $(\alpha-k)!$ 的因子引出 $(\alpha-k-1)!/(\alpha-k)! = (-1)/(k-\alpha)$; 这提供 ' $-\alpha$ ' 到相对参数集合(颠倒上和下的角色), 且取消超几何自变量。像 $n!$ 的因子, 它和 k 独立, 加入 t_0 但是从项比中消失。用这样的诀窍, 我们进一步计算就能预言式(5.27)的项比为

$$\frac{t_{k+1}}{t_k} = \frac{k-r}{k+1} \frac{k-n}{k+s-n+1}$$

乘 $(-1)^2 = 1$, Vandermonde 的卷积变成

$$\binom{s}{n} F\left(\begin{matrix} -r, -n \\ s-n+1 \end{matrix} \middle| 1\right) = \binom{r+s}{n}. \quad (5.91)$$

当 $z=1$ 和 b 是负整数时, 一般能用此方程来确定 $F(a, b; c; z)$.

让我们把式(5.91)改写成一种形式以致当需计算新的和时, 表查找容易. 结果是

$$F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| 1\right) = \frac{\Gamma(c-a-b)\Gamma(c)}{\Gamma(c-a)\Gamma(c-b)}; \quad \begin{matrix} \text{整数 } b \leq 0 \\ \text{或 } \Re c > \Re a + \Re b \end{matrix} \quad (5.92)$$

Vandermonde 卷积(5.27)仅适合于上参数之一, 比如说 b , 是非正整数的情形; 但是高斯证明当 a, b, c 是复数, 它们的实部满足 $\Re c > \Re a + \Re b$ 时, 式(5.92)也成立. 在其他情形, 无穷级数 $F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| 1\right)$ 不收敛. 当 $b = -n$ 时, 用阶乘代替 Γ 函数可更方便地写出等式:

$$F\left(\begin{matrix} a, -n \\ c \end{matrix} \middle| 1\right) = \frac{(c-a)^{\bar{n}}}{c^{\bar{n}}} = \frac{(a-c)^{\bar{n}}}{(-c)^{\bar{n}}}, \quad \text{整数 } n \geq 0. \quad (5.93)$$

原来表 5.3 中所有的 5 个等式是 Vandermonde 卷积的特殊情形; 当对退化情况适当注意时, 公式(5.93)包括它们的一切.

注意, 式(5.82)就是式(5.93)的特殊情形 $a=1$. 所以实际上不需记住式(5.82), 不需等式(5.9)导至式(5.82), 虽然表 5.4 说它是值得注意的. 公式处理的计算机程序, 面临计算 $\sum_k \binom{r+k}{k}$ 的问题, 能把和转换成超几何且插入 Vandermonde 卷积的一般等式.

5.2 节中的问题 1 需要

$$\sum_{k \geq 0} \binom{m}{k} \bigg/ \binom{n}{k}$$

的值. 对于超几何来说这个问题是一个自然的问题, 在一点儿实践之后任何超几何学者能立即看出参数, 像 $F(1, -m; -n; 1)$. 这个问题还是 Vandermonde 上的另一个特殊的副本!

问题 2 和问题 4 中的和也产生 $F(2, 1, -n; 2-m; 1)$, (首先需用 $k+1$ 代替 k .) 而问题 6 中的“威胁性的”和原来就是 $F(n+1, -n; 2; 1)$. 除了 Vandermonde 的有力的卷积的不易识别的形式外, 没有另外的求和了吗?

是的, 问题 3 有点不同. 它处理了式(5.74)中考虑的一般和 $\sum_k \binom{n-k}{k} z^k$ 的特殊情形, 且这导至

$$F\left(\begin{matrix} 1 + 2\lceil n/2 \rceil, -n \\ 1/2 \end{matrix} \middle| -z/4\right)$$

的一个闭形式表达式。

当我们查看 $(1-z)^c(1+z)^c$ 的系数时, 还证明了式(5.55)中的很新的结果:

$$F\left(\begin{matrix} 1-c-2n, -2n \\ c \end{matrix} \middle| -1\right) = (-1)^n \frac{(2n)!}{n!} \frac{(c-1)!}{(c+n-1)!}, \text{ 整数 } n \geq 0.$$

当它被推广到复数时, 称它为 **Kummer 公式**:

$$F\left(\begin{matrix} a, b \\ 1+b-a \end{matrix} \middle| -1\right) = \frac{(b/2)!}{b!} (b-a)^{b/2}. \quad (5.94)$$

(1836 年 Ernst Kummer^[187]证明了它。)

比较这两个公式是有趣的, 用 $1-2n-a$ 代替 c , 我们发现结果是一致的当且仅当 n 是正整数时,

$$(-1)^n \frac{(2n)!}{n!} = \lim_{b \rightarrow -2n} \frac{(b/2)!}{b!} = \lim_{x \rightarrow -n} \frac{x!}{(2x)!} \quad (5.95)$$

例如, 假设 $n=3$; 则我们该有 $-6!/3! = \lim_{x \rightarrow -3} x!/(2x)!$ 。我们知道 $(-3)!$ 和 $(-6)!$ 都是无限的,

但是我们可忽视此困难且设想 $(-3)! = (-3)(-4)(-5)(-6)!$, 以致 $(-6)!$ 的两次出现将消去。然而, 这样的试验一定被抵制, 因为它们导致错误解答! 当 $x \rightarrow -3$ 时 $x!/(2x)!$ 的极限不是 $(-3)(-4)(-5)$, 但是按照式(5.95), 的确 $-6!/3! = (-4)(-5)(-6)$ 。

计算式(5.95)中极限的正确方法是用方程(5.87), 它把负自变量阶乘联系到正自变量 γ 函数。如果用 $-n+\varepsilon$ 代替 x 且让 $\varepsilon \rightarrow 0$, 式(5.87)的两次应用给出

$$\frac{(-n-\varepsilon)!}{(-2n-2\varepsilon)!} \frac{\Gamma(n+\varepsilon)}{\Gamma(2n+2\varepsilon)} = \frac{\sin(2n+2\varepsilon)\pi}{\sin(n+\varepsilon)\pi}.$$

现在 $\sin(x+y) = \sin x \cos y + \cos x \sin y$, 所以用第九章的方法这个正弦的比是

$$\frac{\cos 2n\pi \sin 2\varepsilon\pi}{\cos n\pi \sin \varepsilon\pi} = (-1)^n (2 + O(\varepsilon)).$$

由式(5.86), 我们得到

$$\lim_{\varepsilon \rightarrow 0} \frac{(-n-\varepsilon)!}{(-2n-2\varepsilon)!} = 2(-1)^n \frac{\Gamma(2n)}{\Gamma(n)} = 2(-1)^n \frac{(2n-1)!}{(n-1)!} = (-1)^n \frac{(2n)!}{n!},$$

正如希望的那样。

通过重新陈述本章中到目前为止我们见到的其他等式, 让我们完成概括的研究, 以超几何外表来表达它们, 式(5.29)中三重二项和能写成

$$F\left(\begin{matrix} 1-a-2n, 1-b-2n, -2n \\ a, b \end{matrix} \middle| 1\right)$$

$$= (-1)^n \frac{(2n)!}{n!} \frac{(a+b+2n-2)^{\bar{n}}}{a^{\bar{n}} b^{\bar{n}}}, \text{ 整数 } n \geq 0.$$

当此式被推广到复数时, 称它为 **Dixon 公式**:

$$F\left(\begin{matrix} a, b, c \\ 1+c-a, 1+c-b \end{matrix} \middle| 1\right) = \frac{(c/2)! (c-a)^{c/2} (c-b)^{c/2}}{c! (c-a-b)^{c/2}}, \quad (5.96)$$

$$\Re a + \Re b < 1 + \Re c/2.$$

我们遇到的最一般公式之一是三重二项和(5.28), 它产生 **Saalschütz 等式**:

$$F\left(\begin{matrix} a, b, -n \\ c, a+b-c-n+1 \end{matrix} \middle| 1\right) = \frac{(c-a)^{\bar{n}} (c-b)^{\bar{n}}}{c^{\bar{n}} (c-a-b)^{\bar{n}}},$$

$$= \frac{(a-c)^{\bar{n}} (b-c)^{\bar{n}}}{(-c)^{\bar{n}} (a+b-c)^{\bar{n}}}, \quad \text{整数 } n \geq 0. \quad (5.97)$$

此公式给出一般超几何级数在 $z=1$ 处的值, 具有 3 个上参数和 2 个下参数, 提供的一个上参数是非正整数且 $b_1 + b_2 = a_1 + a_2 + a_3 + 1$. (如果下参数之和超过上参数之和为 2 而不是 1, 则习题 25 的公式能用来表示 $F(a_1, a_2, a_3; b_1, b_2; 1)$, 依据满足 Saalschütz 等式的两个超几何。)

5.2 节的问题 8 中来之不易的等式化成

$$\frac{1}{1+x} F\left(\begin{matrix} x+1, n+1, -n \\ 1, x+2 \end{matrix} \middle| 1\right) = (-1)^n x^{\bar{n}} x^{\frac{-n-1}{2}}.$$

这就是 Saalschütz 等式(5.97)的特殊情形 $c=1$, 所以直接研究超几何能省许多工作!

问题 7 怎么样? 这个特别威胁性的和给出公式

$$F\left(\begin{matrix} n+1, m-n, 1, \frac{1}{2} \\ \frac{1}{2}m+1, \frac{1}{2}m+\frac{1}{2}, 2 \end{matrix} \middle| 1\right) = \frac{m}{n},$$

这是我们看到的第一个具有 3 个下参数的情形, 所以它看来是新的。但是实际上它不是, 由

$$F\left(\begin{matrix} n, m-n-1, -\frac{1}{2} \\ \frac{1}{2}m, \frac{1}{2}m-\frac{1}{2} \end{matrix} \middle| 1\right) = 1$$

能替换左边, 用习题 26, 且又到达 Saalschütz 等式。

这是另一个减少下参数的经历, 但是它也是另一个理解超几何方法能力的原因。

表 5.5 中的卷积等式没有超几何的等价式。因为仅当 t 是整数时它们的项比是 k 的有理函数。甚至当 $t=1$ 时方程(5.64)和(5.65)都不是超几何的。但是当 t 有小整数值时，我们注意到式(5.62)所告诉我们的：

$$F\left(\begin{matrix} \frac{1}{2}r, \frac{1}{2}r + \frac{1}{2}, -n, -n-s \\ r+1, -n-\frac{1}{2}s, -n-\frac{1}{2}s+\frac{1}{2} \end{matrix} \middle| 1\right) = \binom{r+s+2n}{n} / \binom{s+2n}{n};$$

$$F\left(\begin{matrix} \frac{1}{3}r, \frac{1}{3}r + \frac{1}{3}, \frac{1}{3}r + \frac{2}{3}, -n, -n-\frac{1}{2}s, -n-\frac{1}{2}s-\frac{1}{2} \\ \frac{1}{2}r + \frac{1}{2}, \frac{1}{2}r + 1, -n-\frac{1}{3}s, -n-\frac{1}{3}s+\frac{1}{3}, -n-\frac{1}{3}s+\frac{2}{3} \end{matrix} \middle| 1\right)$$

$$= \binom{r+s+3n}{n} / \binom{s+3n}{n}.$$

当用 $(1, 2n+1-m, -1-n)$ 替换量 (r, s, n) 时，这些公式的第一个又给出了问题 7 的结果。

最后，“不期望的”和(5.20)给出一个不期望的超几何等式，它是相当有启发的。让我们以慢动作来查看它。首先转到一个无穷和

$$\sum_{k \leq m} \binom{m+k}{k} 2^{-k} = 2^m \Leftrightarrow \sum_{k \geq 0} \binom{2m-k}{m-k} 2^k = 2^{2m}.$$

$(2m-k)!2^k / m!(m-k)!$ 的项比是 $2(k-m)/(k-2m)$ ，所以我们得到 $z=2$ 的超几何等式：

$$\binom{2m}{m} F\left(\begin{matrix} 1, -m \\ -2m \end{matrix} \middle| 2\right) = 2^{2m}, \text{ 整数 } m \geq 0. \quad (5.98)$$

而查看下参数‘ $-2m$ ’，负整数是禁止的，所以此等式无定义！

如同前面所指望的那样，该仔细查看这样的极限的情形，因为退化超几何常常可通过从附近非退化点接近它们来计算。当做这一点时一定要小心，因为以不同方式取极限，则能得到不同的结果。例如，这里是两个极限，当一个上参数增加 ε 时，结果两个极限是完全不同的：

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} F\left(\begin{matrix} -1+\varepsilon, -3 \\ -2+\varepsilon \end{matrix} \middle| 1\right) &= \lim_{\varepsilon \rightarrow 0} \left(1 + \frac{(-1+\varepsilon)(-3)}{(-2+\varepsilon)1!} + \frac{(-1+\varepsilon)(\varepsilon)(-3)(-2)}{(-2+\varepsilon)(-1+\varepsilon)2!}\right. \\ &\quad \left.+ \frac{(-1+\varepsilon)(\varepsilon)(1+\varepsilon)(-3)(-2)(-1)}{(-2+\varepsilon)(-1+\varepsilon)(\varepsilon)3!}\right) \\ &= 1 - \frac{3}{2} + 0 + \frac{1}{2} = 0; \end{aligned}$$

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} F\left(\begin{matrix} -1, -3 \\ -2+\varepsilon \end{matrix} \middle| 1\right) &= \lim_{\varepsilon \rightarrow 0} \left(1 + \frac{(-1)(-3)}{(-2+\varepsilon)1!} + 0 + 0\right) \\ &= 1 - \frac{3}{2} + 0 + 0 = -\frac{1}{2}. \end{aligned}$$

类似, 我们定义 $\binom{-1}{-1} = 0 = \lim_{\varepsilon \rightarrow 0} \binom{-1+\varepsilon}{-1+\varepsilon}$, 这与 $\lim_{\varepsilon \rightarrow 0} \binom{-1+\varepsilon}{-1+\varepsilon} = 1$ 不相同. 适当的方法是把式(5.98)看作一个极限, 用上参数 $-m$ 使对于 $k > m$ 级数 $\sum_{k=0}^{\infty} \binom{2m-k}{m-k} 2^k$ 的所有项为零来实现; 这意味着我们要作出下列更精确的陈述:

$$\binom{2m}{m} \lim_{\varepsilon \rightarrow 0} F\left(\begin{matrix} 1, -m \\ -2m + \varepsilon \end{matrix} \middle| 2\right) = 2^{2m}, \quad \text{整数 } m \geq 0. \quad (5.99)$$

此极限的每一项是恰当定义的, 因为直到 $k > 2m$ 分母因子 $(-2m)^k$ 不变成零. 所以此极限恰好给出我们开始的和(5.20).

5.6 超几何变换

现在该明白, 已知的超几何闭形式的数据库是处理二项系数和的有用工具. 我们简单地把任何给定和转换成它的标准超几何形式, 然后在表中查找它. 若表中有, 则取得解答. 若没有, 如果和可表达为闭形式, 我们能把它加进数据库. 我们也可在表中包含这些项目, 并说明“此和没有简单的一般闭形式.”例如, 和 $\sum_{k=0}^n \binom{n}{k}$ 对应于超几何函数

$$\binom{n}{m} F\left(\begin{matrix} 1, -m \\ n - m + 1 \end{matrix} \middle| -1\right), \quad \text{整数 } n \geq m \geq 0, \quad (5.100)$$

仅当 m 接近 0, $(1/2)n$ 或 n 时, 这个和有简单的闭形式.

但是还有更多的内容, 因为超几何函数服从一些它们自身的等式. 这意味着每一个超几何闭形式导致另外的闭形式以及引出数据库中的另外的项目. 例如, 习题 25 和 26 中的等式告诉我们, 如何把一个超几何变换到两个其他的相似的超几何, 但参数不同, 这些又能依次变换.

1793 年 J.F. Pfaff 发现一个意外的反射律,

$$\frac{1}{(1-z)^a} F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| \frac{-z}{1-z}\right) = F\left(\begin{matrix} a, c-b \\ c \end{matrix} \middle| z\right), \quad (5.101)$$

它是另一种类型的变换. 这里幂级数的形式等式, 如果当展开左边时用无穷级数 $(-z)^k \left(1 + \binom{k+a}{1} z + \binom{k+a+1}{2} z^2 + \dots\right)$ 代替量 $(-z)^k / (1-z)^{k+a}$ (见习题 50). 当 $z \neq 1$ 时我们能从此定律从已知等式引出新的公式.

例如, Kummer 公式(5.94)能和反射律(5.101)结合起来, 如果选取适当参数, 二者都

可应用:

$$2^{-a} F\left(\begin{matrix} a, 1-a \\ 1+b-a \end{matrix} \middle| \frac{1}{2}\right) = F\left(\begin{matrix} a, b \\ 1+b-a \end{matrix} \middle| -1\right) \\ = \frac{(b/2)!}{b!} (b-a)^{b/2}. \quad (5.102)$$

我们现在能置 $a = -n$, 且总有一天我们可能需要从此方程回到二项系数的新等式:

$$\sum_{k \geq 0} \frac{(-n)^k (1+n)^{\bar{k}}}{(1+b+n)^k} \frac{2^{-k}}{k!} = \sum_k \binom{n}{k} \left(\frac{-1}{2}\right)^k \binom{n+k}{k} / \binom{n+b+k}{k} \\ = 2^{-n} \frac{(b/2)!(b+n)!}{b!(b/2+n)!}, \text{ 整数 } n \geq 0. \quad (5.103)$$

例如, 当 $n=3$ 时此等式表明

$$1 - 3 \frac{4}{2(4+b)} + 3 \frac{4 \cdot 5}{4(4+b)(5+b)} - \frac{4 \cdot 5 \cdot 6}{8(4+b)(5+b)(6+b)} \\ = \frac{(b+3)(b+2)(b+1)}{(b+6)(b+4)(b+2)}.$$

几乎是难以相信, 但对于所有 b 为真。(除了当分母中的一个因子成为零外。)

这是有趣的事, 让我们再试验。我们将可能发现一个公式, 它将使我们的朋友确实感到惊讶。如果应用 Pfaff 的反射律到不熟悉的形式(5.99), 那里 $z=2$, 反射律告诉我们什么? 此时置 $a=-m$, $b=1$ 和 $c=-2m+\varepsilon$, 得

$$(-1)^m \lim_{\varepsilon \rightarrow 0} F\left(\begin{matrix} -m, 1 \\ -2m+\varepsilon \end{matrix} \middle| 2\right) = \lim_{\varepsilon \rightarrow 0} F\left(\begin{matrix} -m, -2m-1+\varepsilon \\ -2m+\varepsilon \end{matrix} \middle| 2\right) \\ = \lim_{\varepsilon \rightarrow 0} \sum_{k \geq 0} \frac{(-m)^{\bar{k}} (-2m-1+\varepsilon)^{\bar{k}}}{(-2m+\varepsilon)^{\bar{k}}} \frac{2^k}{k!} \\ = \sum_{k \leq m} \binom{m}{k} \frac{(2m+1)^{\bar{k}}}{(2m)^{\bar{k}}} (-2)^k,$$

因为没有极限项接近零。这导致另一个非凡的公式,

$$\sum_{k \leq m} \binom{m}{k} \frac{2m+1}{2m+1-k} (-2)^k = (-1)^m 2^{2m} / \binom{2m}{m} \\ = 1 / \binom{-1/2}{m}, \text{ 整数 } m \geq 0. \quad (5.104)$$

例如, 当 $m=3$ 时, 和是

$$1 - 7 + \frac{84}{5} - 14 = -\frac{16}{5},$$

且 $\left(-\frac{1}{3}\right)^2$ 确实等于 $-\frac{5}{16}$.

当我们查看二项系数等式以及把它们转换到超几何形式时, 检查式(5.19), 因为它是两个和而不是闭形式之间的一个关系。但是现在我们能把式(5.19)视为超几何级数间的一个等式。如果对 y 微分它 n 次, 然后用 $m-n-k$ 代替 k 。则取得

$$\begin{aligned} \sum_{k \geq 0} \binom{m+r}{m-n-k} \binom{n+k}{n} x^{m-n-k} y^k \\ = \sum_{k \geq 0} \binom{-r}{m-n-k} \binom{n+k}{n} (-x)^{m-n-k} (x+y)^k. \end{aligned}$$

这产生了下列超几何变换:

$$F\left(\begin{matrix} a, -n \\ c \end{matrix} \middle| z\right) = \frac{(a-c)^n}{(-c)^n} F\left(\begin{matrix} a, -n \\ 1-n+a-c \end{matrix} \middle| 1-z\right), \quad \begin{matrix} \text{整数} \\ n \geq 0. \end{matrix} \quad (5.105)$$

注意, 当 $z=1$ 时, 这化为 Vandermonde 卷积, 式(5.93)。

微分看来是有用的, 如果这个例子是任何表示法, 在第二章中当求和 $x + 2x^2 + \dots + nx^n$ 时, 我们也发现它是有帮助的。让我们来看当对 z 微分一般的超几何级数时发生什么:

$$\begin{aligned} \frac{d}{dz} F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) &= \sum_{k \geq 1} \frac{a_1^k \dots a_m^k z^{k-1}}{b_1^k \dots b_n^k (k-1)!} \\ &= \sum_{k+1 \geq 1} \frac{a_1^{k+1} \dots a_m^{k+1} z^k}{b_1^{k+1} \dots b_n^{k+1} k!} \\ &= \sum_{k \geq 0} \frac{a_1(a_1+1)^k \dots a_m(a_m+1)^k z^k}{b_1(b_1+1)^k \dots b_n(b_n+1)^k k!} \\ &= \frac{a_1 \dots a_m}{b_1 \dots b_n} F\left(\begin{matrix} a_1+1, \dots, a_m+1 \\ b_1+1, \dots, b_n+1 \end{matrix} \middle| z\right). \end{aligned} \quad (5.106)$$

参数移出和由小变大。

也可能用微分仅捏住一个参数而剩下的参数保持固定。关于这一点我们用算子

$$\theta = z \frac{d}{dz},$$

它作用在一个函数上, 微分它, 然后乘 z 。这个算子给出

$$\theta F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) = z \sum_{k \geq 1} \frac{a_1^k \dots a_m^k z^{k-1}}{b_1^k \dots b_n^k (k-1)!} = \sum_{k \geq 0} \frac{k a_1^k \dots a_m^k z^k}{b_1^k \dots b_n^k k!},$$

它本身不是太有用。但是如采用一个 F 的上参数, 比如说 a_1 , 乘 F , 且加 θF , 我们取得

$$\begin{aligned} (\theta + a_1)F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) &= \sum_{k \geq 0} \frac{(k + a_1) a_1^{\bar{k}} \dots a_m^{\bar{k}} z^k}{b_1^{\bar{k}} \dots b_n^{\bar{k}} k!} = \sum_{k \geq 0} \frac{a_1 (a_1 + 1)^{\bar{k}} a_2^{\bar{k}} \dots a_m^{\bar{k}} z^k}{b_1^{\bar{k}} \dots b_n^{\bar{k}} k!} \\ &= a_1 F\left(\begin{matrix} a_1 + 1, a_2, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right). \end{aligned}$$

仅一个参数替换。

类似的诀窍用于下参数, 但是这种情形中减少代替增加:

$$\begin{aligned} (\theta + b_1 - 1)F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) &= \sum_{k \geq 0} \frac{(k + b_1 - 1) a_1^{\bar{k}} \dots a_m^{\bar{k}} z^k}{b_1^{\bar{k}} \dots b_n^{\bar{k}} k!} \\ &= \sum_{k \geq 0} \frac{(b_1 - 1) a_1^{\bar{k}} \dots a_m^{\bar{k}} z^k}{(b_1 - 1)^{\bar{k}} b_2^{\bar{k}} \dots b_n^{\bar{k}} k!} \\ &= (b_1 - 1)F\left(\begin{matrix} a_1, \dots, a_m \\ b_1 - 1, b_2, \dots, b_n \end{matrix} \middle| z\right). \end{aligned}$$

我们现在能结合所有这些运算, 且以两种不同方式表达相同量来作出一个数学的“双关语”, 即, 我们有

$$\begin{aligned} (\theta + a_1) \dots (\theta + a_m) F &= a_1 \dots a_m F\left(\begin{matrix} a_1 + 1, \dots, a_m + 1 \\ b_1, \dots, b_n \end{matrix} \middle| z\right), \text{ 和} \\ (\theta + b_1 - 1) \dots (\theta + b_n - 1) F &= (b_1 - 1) \dots (b_n - 1) F\left(\begin{matrix} a_1, \dots, a_m \\ b_1 - 1, \dots, b_n - 1 \end{matrix} \middle| z\right), \end{aligned}$$

其中 $F = F(a_1, \dots, a_m; b_1, \dots, b_n; z)$, 式(5.106)告诉我们, 顶行是底行的导数。所以一般的超几何函数 F 满足微分方程

$$D(\theta + b_1 - 1) \dots (\theta + b_n - 1) F = (\theta + a_1) \dots (\theta + a_m) F, \quad (5.107)$$

其中 D 是算子 $\frac{d}{dz}$ 。

就一个例子来看这一点, 让我们找标准的超几何级数 $F(z) = F(a, b; c; z)$, 按照式(5.107), 我们有

$$D(\theta + c - 1)F = (\theta + a)(\theta + b)F.$$

以通常记法这是什么意思? $(\theta + c - 1)F$ 为 $zF'(z) + (c - 1)F(z)$, 它的导数产生左边,

$$F'(z) + zF''(z) + (c - 1)F'(z).$$

在右边我们有

$$\begin{aligned}
 (\theta + a)(zF'(z) + bF(z)) &= z \frac{d}{dz} (zF'(z) + bF(z)) + a(zF'(z) + bF(z)) \\
 &= zF'(z) + z^2 F''(z) + bzF'(z) + azF'(z) + abF(z).
 \end{aligned}$$

使两边相等告诉我们

$$z(1-z)F''(z) + (c - z(a+b+1))F'(z) + abF(z) = 0. \quad (5.108)$$

此方程等价于因子分解形式(5.107).

相反, 我们能从微分方程返回到幂级数. 让我们假设 $F(z) = \sum_{k \geq 0} t_k z^k$ 是满足式(5.107)的幂级数.

简单计算表明一定得到

$$\frac{t_{k+1}}{t_k} = \frac{(k+a_1) \cdots (k+a_m)}{(k+b_1) \cdots (k+b_n)(k+1)};$$

因此 $F(z)$ 一定是 ${}_0F(a_1, \dots, a_m; b_1, \dots, b_n; z)$. 我们证明了超几何级数(5.76)是满足微分方程(5.107)的唯一的幂级数且有常数项 1.

如果超几何解出了所有世界上的微分方程, 那将是好的, 但是它们并不是这样. 式(5.107)的右边总展成形式 $\alpha_k z^k F^{(k)}(z)$ 的项的和, 其中 $F^{(k)}(z)$ 是 k 阶导数 $D^k F(z)$; 左边总展成形式 $\beta_k z^{k-1} F^{(k)}(z)$ ($k > 0$) 的项的和. 所以微分方程(5.107)总取特殊形式

$$z^{n-1}(\beta_n - z\alpha_n)F^{(n)}(z) + \cdots + (\beta_1 - z\alpha_1)F'(z) - \alpha_0 F(z) = 0.$$

情形 $n=2$ 的方程(5.108)说明了这一点. 相反, 在习题 6.13 我们将证明这种形式的任何微分方程能根据 θ 算子因子分解, 给出像式(5.107)的一个方程. 所以这些是解为具有有理项比的幂级数的微分方程.

式(5.107)的两边乘 z 免除了 D 算子且给出一个有益的全为 θ 的形式,

$$\theta(\theta + b_1 - 1) \cdots (\theta + b_n - 1)F = z(\theta + a_1) \cdots (\theta + a_m)F. \quad (5.109)$$

在项比式(5.81)中左边的第一个因子 $\theta = (\theta + 1 - 1)$ 对应于 $(k+1)$, 在一般超几何中它对应于第 k 项的分母中的 $k!$. 其他因子 $(\theta + b_j - 1)$ 对应于分母因子 $(k + b_j)$, 它对应于式(5.76)中的 b_j^k . 右边, z 对应于 z^k , 而 $(\theta + a_j)$ 对应于 a_j^k .

这个微分理论的一个用处是找并且证明新的变换. 例如, 我们能容易验证两个超几何

$$F\left(\begin{matrix} 2a, 2b \\ a+b+\frac{1}{2} \end{matrix} \middle| z\right) \text{ 和 } F\left(\begin{matrix} a, b \\ a+b-\frac{1}{2} \end{matrix} \middle| 4z(1-z)\right)$$

满足微分方程

$$z(1-z)F''(z) + (a+b+\frac{1}{2})(1-2z)F'(z) - 4abF(z) = 0;$$

因此高斯等式^[116, 方程102]

$$F\left(\begin{matrix} 2a, 2b \\ a+b+\frac{1}{2} \end{matrix} \middle| z\right) = F\left(\begin{matrix} a, b \\ a+b+\frac{1}{2} \end{matrix} \middle| 4z(1-z)\right) \quad (5.110)$$

一定是真的, 特别, 每当两个无穷和收敛时,

$$F\left(\begin{matrix} 2a, 2b \\ a+b+\frac{1}{2} \end{matrix} \middle| \frac{1}{2}\right) = F\left(\begin{matrix} a, b \\ a+b+\frac{1}{2} \end{matrix} \middle| 1\right). \quad (5.111)$$

每一个超几何的新等式有二项系数的推断, 且这一个也不例外. 让我们考虑和

$$\sum_{k \leq m} \binom{m-k}{n} \binom{m+n+1}{k} \left(\frac{-1}{2}\right)^k, \text{ 整数 } m \geq n \geq 0.$$

对 $0 \leq k \leq m-n$, 项是非零的, 如同前面那样, 有点巧妙地取极限, 我们能把此和表达为超几何

$$\lim_{\epsilon \rightarrow 0} \binom{m}{n} F\left(\begin{matrix} n-m, -n-m-1+\alpha\epsilon \\ -m+\epsilon \end{matrix} \middle| \frac{1}{2}\right).$$

α 的值不影响极限, 因为在初期非正上参数 $n-m$ 从和中删去. 我们能置 $\alpha=2$, 以致式(5.111)应用. 现在能计算极限, 因为右边是式(5.92)的特殊情形. 如同习题 54 中表明的那样, 结果能表为简单的形式

$$\sum_{k \leq m} \binom{m-k}{n} \binom{m+n+1}{k} \left(\frac{-1}{2}\right)^k \\ \left(\binom{(m+n)/2}{n}\right) 2^{n-m} [m+n \text{ 是偶数}], \text{ 整数 } m \geq n \geq 0. \quad (5.112)$$

例如, 当 $m=5$ 和 $n=2$ 时我们取得 $\binom{5}{2}\binom{8}{0} - \binom{4}{2}\binom{8}{1}/2 + \binom{3}{2}\binom{8}{2}/4 - \binom{2}{2}\binom{8}{3}/8 = 10 - 24 + 21 - 7 = 0$; 当 $m=4$ 和 $n=2$ 时, 两边给出 $3/4$.

我们也能找到, 当 $z=-1$ 时式(5.110)给出二项和的情形, 但是这些确实是离奇的. 若置 $a=(1/6)-(n/3)$ 和 $b=-n$, 我们取得奇怪的公式

$$F\left(\begin{matrix} \frac{1}{3}, \frac{2}{3}n, -2n \\ \frac{2}{3} - \frac{4}{3}n \end{matrix} \middle| -1\right) = F\left(\begin{matrix} \frac{1}{6}, -\frac{1}{3}n, -n \\ \frac{2}{3} - \frac{4}{3}n \end{matrix} \middle| -8\right).$$

当 $n \not\equiv 2 \pmod{3}$ 时这些超几何是非退化多项式, 且巧妙地选取参数以致左边能用式(5.94)

计算。所以我们引导到一个确实伤脑筋的结果。

$$\begin{aligned} \sum_k \binom{n}{k} \left(\frac{1}{3}n - \frac{1}{6} \right) 8^k / \binom{\frac{4}{3}n - \frac{2}{3}}{k} \\ = \binom{2n}{n} / \binom{\frac{4}{3}n - \frac{2}{3}}{n}, \text{ 整数 } n \geq 0, n \not\equiv 2 \pmod{3}. \end{aligned} \quad (5.113)$$

这是在二项系数中我们已见到的最令人吃惊的等式。等式的小的情形不易用手来检查。(结果, 当 $n=3$ 时两边给出 $81/7$ 。)但是等式完全无用, 当然, 在实际问题中它确实不出现。

以上是关于超几何的人为例子。我们已看到, 超几何级数提供了一种高级的方法来了解二项系数和中所发生的事情。在 Wilfred N. Bailey^[15]写的传统的书中以及 Lucy Joan Slater^[269]的续集中可找到大量另外的资料。

5.7 部分超几何和

本章中计算的大多数和是在所有指标 $k \geq 0$ 的范围上求的, 但是有时也能找到一个闭形式, 它在一般范围 $0 \leq k < m$ 上成立。例如, 根据式(5.16)可知

$$\sum_{k < m} \binom{n}{k} (-1)^k = (-1)^{m-1} \binom{n-1}{m-1}, \text{ 整数 } m. \quad (5.114)$$

第二章中的理论给出一种好方法来了解像这样的公式: 若 $f(k) = \Delta g(k) = g(k+1) - g(k)$, 则我们答应记 $\sum f(k) \delta k = g(k) + C$ 和

$$\sum_a^b f(k) \delta k = g(k) \Big|_a^b = g(b) - g(a).$$

此外, 当 a 和 b 是整数且 $a \leq b$ 时, 我们得到

$$\sum_a^b f(k) \delta k = \sum_{a \leq k < b} f(k) = g(b) - g(a).$$

所以等式(5.114)对应于不定求和公式

$$\sum \binom{n}{k} (-1)^k \delta k = (-1)^{k-1} \binom{n-1}{k-1} + C;$$

并且对应于差分公式

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

从函数 $g(k)$ 开始并计算 $\Delta_k(k) = f(k)$ 是容易的, 它的和将是 $g(k) + C$. 但是从 $f(k)$ 开始并计算出它的不定和 $\sum f(k)\delta k = g(k) + C$ 是十分难的, 这个函数 g 可以没有简单形式. 例如, 显而易见 $\sum \binom{n}{k} \delta k$ 没有简单形式, 否则就能计算像 $\sum_{k \leq n/3} \binom{n}{k}$ 的和, 关于此我们没有线索.

1977 年, R.W.Gosper^[124]发现了一个极好的方法来判定是否一个给定的函数关于称为超几何项的一般函数类是可求无定限的和的. 让我们对于超几何级数 $F(a_1, \dots, a_m; b_1, \dots, b_n; z)$ 的第 k 项记为

$$F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right)_k = \frac{a_1^{\bar{k}} \dots a_m^{\bar{k}}}{b_1^{\bar{k}} \dots b_n^{\bar{k}}} \frac{z^k}{k!}. \quad (5.115)$$

我们将 $F(a_1, \dots, a_m; b_1, \dots, b_n; z)_k$ 视为 k 的一个函数, 不是 z 的函数. Gosper 的判定过程允许我们来判定是否存在参数 $c, A_1, \dots, A_M, B_1, \dots, B_N$ 和 Z , 使得

$$\sum F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right)_k \delta k = c F\left(\begin{matrix} A_1, \dots, A_M \\ B_1, \dots, B_N \end{matrix} \middle| Z\right)_k + C, \quad (5.116)$$

给出 $a_1, \dots, a_m, b_1, \dots, b_n$ 和 z . 我们将说给定的函数 $F(a_1, \dots, a_m; b_1, \dots, b_n; z)$ 是以超几何项可求和的, 如果这样的常数 $c, A_1, \dots, A_M, B_1, \dots, B_N, Z$ 存在.

让我们分别以 $t(k)$ 和 $T(k)$ 作为 $F(a_1, \dots, a_m; b_1, \dots, b_n; z)_k$ 和 $F(A_1, \dots, A_M; B_1, \dots, B_N; Z)_k$ 的缩写. Gosper 的判定过程中的第一步是表达项比

$$\frac{t(k+1)}{t(k)} = \frac{(k+a_1) \dots (k+a_m)z}{(k+b_1) \dots (k+b_n)(k+1)}$$

成特殊形式

$$\frac{t(k+1)}{t(k)} = \frac{p(k+1)}{p(k)} \frac{q(k)}{r(k+1)}, \quad (5.117)$$

其中 p, q 和 r 是服从下列条件的多项式:

$$(k+\alpha) \nmid q(k) \text{ 和 } (k+\beta) \nmid r(k) \Rightarrow \alpha - \beta \text{ 不是一个正整数.} \quad (5.118)$$

此条件容易达到: 我们从暂时置 $p(k) = 1, q(k) = (k+a_1) \dots (k+a_m)z$ 和 $r(k) = (k+b_1-1) \dots (k+b_n-1)k$ 开始, 然后检查是否违背式(5.118). 如果 q 和 r 有因子 $(k+\alpha)$ 和 $(k+\beta)$, 其中 $\alpha - \beta = N > 0$, 则把它们分离出 q 和 r 且用

$$p(k)(k+\alpha-1)^{N-1} = p(k)(k+\alpha-1)(k+\alpha-2) \dots (k+\beta+1)$$

代替 $p(k)$, 新的 p, q 和 r 仍满足式(5.117), 可重复此过程直到式(5.118)成立.

我们的目的是找一个超几何项 $T(k)$, 使得对某个常数 c ,

$$t(k) = cT(k+1) - cT(k). \quad (5.119)$$

让我们记

$$cT(k) = \frac{r(k)s(k)t(k)}{p(k)}, \quad (5.120)$$

其中 $s(k)$ 是以某种方式一定被发现的隐蔽的函数。将式(5.120)代入式(5.117)和(5.119), 给出 $s(k)$ 一定满足的方程:

$$p(k) = q(k)s(k+1) - r(k)s(k). \quad (5.121)$$

如果能找到满足此递归的 $s(k)$, 则我们找到 $\sum t(k)\delta k$.

假设 $T(k+1)/T(k)$ 是 k 的有理函数, 所以由式(5.120)和(5.119), $r(k)s(k)/p(k) = T(k)/(T(k+1) - T(k))$ 是 k 的有理函数, 且 $s(k)$ 一定是多项式的商:

$$s(k) = f(k)/g(k). \quad (5.122)$$

但是事实上我们能证明 $s(k)$ 是一个多项式。由于如果 $g(k) \neq 1$, 且如果 $f(k)$ 和 $g(k)$ 没有公因子, 设 N 是最大整数使得对于某个复数 β , $(k+\beta)$ 和 $(k+\beta+N-1)$ 都出现为 $g(k)$ 的因子。 N 的值是正的, 因为 $N=1$ 总满足此条件。方程(5.121)可改写为

$$p(k)g(k+1)g(k) = q(k)f(k+1)g(k) - r(k)g(k+1)f(k),$$

且若置 $k = -\beta$ 和 $k = -\beta - N$, 我们取得

$$r(-\beta)g(1-\beta)f(-\beta) = 0 = q(-\beta-N)f(1-\beta-N)g(-\beta-N).$$

现在 $f(-\beta) \neq 0$, $f(1-\beta-N) \neq 0$, 因为 f 和 g 没有公共根。同样 $g(1-\beta) \neq 0$, $g(-\beta-N) \neq 0$, 否则 $g(k)$ 将包含因子 $(k+\beta-1)$ 或 $(k+\beta+N)$, 违反 N 的最大性。所以

$$r(-\beta) = q(-\beta-N) = 0.$$

但这与条件(5.118)矛盾。因此 $s(k)$ 一定是一个多项式。

剩下的任务是当 $p(k)$, $q(k)$ 和 $r(k)$ 是给定多项式时, 判定是否存在一个满足式(5.121)的多项式 $s(k)$ 。对于任何特定次数 d 的多项式判定这一点是容易的, 因为对于未知系数 $(\alpha_d, \dots, \alpha_0)$ 我们能写

$$s(k) = \alpha_d k^d + \alpha_{d-1} k^{d-1} + \dots + \alpha_0, \quad \alpha_d \neq 0$$

且把此表达式代入确定的方程。多项式 $s(k)$ 将满足递归当且仅当 α 满足某个线性方程, 因为每个 k 的幂一定有式(5.121)两边上的相同系数。

但是如何能确定 s 的次数呢? 实际至多有两种可能性。我们可把式(5.121)再写成形式

$$2p(k) = Q(k)(s(k+1) + s(k)) + R(k)(s(k+1) - s(k)), \quad (5.123)$$

其中 $Q(k) = q(k) + r(k)$ 和 $R(k) = q(k) + r(k)$ 。如果 $s(k)$ 有次数 d ，则和 $s(k+1) + s(k) = 2\alpha_d k^d + \dots$ 也有次数 d ，而差分 $s(k+1) - s(k) = \Delta s(k) = d\alpha_d k^{d-1} + \dots$ 有次数 $d-1$ 。（假设零多项式有次数 -1 。）让我们记多项式 p 的次数为 $\deg(p)$ 。若 $\deg(Q) \geq \deg(R)$ ，则式 (5.123) 右边的次数为 $\deg(Q) + d$ ，所以我们一定有 $d = \deg(p) - \deg(Q)$ 。另一方面，若 $\deg(Q) < \deg(R) = d'$ ，我们能写出 $Q(k) = \beta k^{d'-1} + \dots$ 和 $R(k) = \gamma k^{d'} + \dots$ ，其中 $\gamma \neq 0$ ；式 (5.123) 的右边有形式

$$(2\beta\alpha_d + \gamma d\alpha_d)k^{d+d'-1} + \dots$$

因此，两种可能性：或者 $2\beta + \gamma d \neq 0$ ， $d = \deg(p) - \deg(R) + 1$ ；或者 $2\beta + \gamma d = 0$ ， $d > \deg(p) - \deg(R) + 1$ 。第二种情形仅当 $-2\beta/\gamma$ 是大于 $\deg(p) - \deg(R) + 1$ 的整数 d 时需要检查。

于是我们有足够的论据来判定是否一个适当的多项式 $s(k)$ 存在。若存在，则能把它代入式 (5.120)，得到 T 。若不存在，我们证明了 $\sum t(k)\delta k$ 不是一个超几何项。

现在是举一个例子的时候了。让我们试验部分和 (5.114)，对于任何指定的 n ，Gosper 方法能推出

$$\sum \binom{n}{k} (-1)^k \delta k$$

的值。忽略不包含 k 的因子，我们要

$$t(k) = F\left(\begin{matrix} 1, -n \\ 1 \end{matrix} \middle| 1\right)_k$$

的和。第一步把项比改写成要求的形式 (5.117)，我们得到

$$\frac{t(k+1)}{t(k)} = \frac{(k-n)}{(k+1)} = \frac{p(k+1)q(k)}{p(k)r(k+1)},$$

所以简单地取 $p(k) = 1$ ， $q(k) = k-n$ 和 $r(k) = k$ 。这样选取的 p 、 q 和 r 满足式 (5.118)，除非 n 是负整数，我们假设不是这样。按照式 (5.123)，我们应考虑多项式 $Q(k) = -n$ 和 $R(k) = 2k-n$ 。由于 R 比 Q 的次数大，我们需查看两种情形。 $d = \deg(p) - \deg(R) + 1$ ，它是 0；或者 $d = -2\beta/\gamma$ ，其中 $\beta = -n$ ， $\gamma = 2$ ，因此 $d = n$ 。第一种情形较好，所以让我们先试它。

方程 (5.121) 是

$$1 = (k-n)\alpha_0 - kx_0$$

所以选 $\alpha_0 = -1/n$ 。这满足要求的条件且给出

$$cT(k) = \frac{r(k)s(k)t(k)}{p(k)} = k \cdot \frac{(-1)}{n} \cdot \binom{n}{k} (-1)^k = \binom{n-1}{k-1} (-1)^{k-1},$$

这就是我们希望确定的解答.

如果应用相同方法来找不定和 $\sum \binom{n}{k} \delta k$, 没有 $(-1)^k$. 除了 $q(k)$ 将为 $n-k$ 之外, 其他几乎相同; 因此 $Q(k) = n-2k$ 将有比 $R(k) = n$ 大的次数, 且我们将断定 d 有不可能的值 $\deg(p) - \deg(Q) = -1$. 所以函数 $\binom{n}{k}$ 不是以超几何项可求和的.

然而, 一旦除去了不可能的情形, 剩下的什么情形是不大可能发生的, 一定是真的 (按照 S.Holmes^[70]). 当定义 p , q 和 r 时, 我们决定忽略 n 可能为负整数的可能性. 如果是, 怎么办呢? 让我们置 $n = -N$, 其中 N 是正的. 于是 $\sum \binom{n}{k} \delta k$ 的项比是

$$\frac{r(k+1)}{r(k)} = \frac{-(k+N)}{(k+1)} = \frac{p(k+1)}{p(k)} \frac{q(k)}{r(k+1)},$$

且由 $p(k) = (k+1)^{N-1}$, $q(k) = -1$, $r(k) = 1$ 来表达它. Gosper 方法现在告诉我们来寻找次数为 $d = N-1$ 的多项式 $s(k)$, 毕竟可能有希望. 例如, 当 $N=2$ 时, 我们要来解

$$k+1 = -((k+1)\alpha_1 + \alpha_0) - (k\alpha_1 + \alpha_0).$$

使 k 的系数与 1 相等, 得到

$$1 = -\alpha_1 - \alpha_1, \quad 1 = -\alpha_1 - \alpha_0 - \alpha_0,$$

因此, $s(k) = -\frac{1}{2}k - \frac{1}{4}$ 是一个解, 且

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

这能为欲求的和吗? 是的, 它被认为无误:

$$(-1)^k \frac{2k+3}{4} - (-1)^{k-1} \frac{2k+1}{4} = (-1)^k (k+1) = \binom{-2}{k}.$$

我们可把求和公式写成另一种形式,

$$\begin{aligned} \sum_{k < m} \binom{-2}{k} &= (-1)^{k-1} \frac{2k+1}{4} \Big|_0^m \\ &= \frac{(-1)^{m-1}}{2} \left(m + \frac{1 - (-1)^m}{2} \right) \\ &= (-1)^{m-1} \left\lceil \frac{m}{2} \right\rceil. \end{aligned}$$

这种表示隐瞒了 $\binom{-2}{k}$ 按超几何项可求和的事实, 因为 $\lceil m/2 \rceil$ 不是一个超几何项.

对于本章前面提到的超几何和的数据库,可求和的超几何的目录作了有用的增建部分. 让我们尝试搜集我们知道的超几何项的和的一个表. 几何级数 $\sum z^k \delta k$ 是一种十分特殊的情形, 它能记为 $\sum z^k \delta k = (z-1)^{-1} z^k + C$ 或

$$\sum F\left(\begin{matrix} 1, 1 \\ 1 \end{matrix} \middle| z\right)_k \delta k = \frac{1}{z-1} F\left(\begin{matrix} 1, 1 \\ 1 \end{matrix} \middle| z\right)_k + C. \quad (5.124)$$

在第二章中我们还计算了 $\sum k z^k \delta k$. 当 $k=0$ 时, 这个被加项为零, 所以考虑和 $\sum (k+1) \cdot z^k \delta k$ 代替而取得一个合适的超几何项. 适当的公式以超几何记法就是

$$\sum F\left(\begin{matrix} 2, 1 \\ 1 \end{matrix} \middle| z\right)_k \delta k = \frac{-1}{(1-z)^2} F\left(\begin{matrix} 1+1/(1-z), 1 \\ 1/(1-z) \end{matrix} \middle| z\right)_k. \quad (5.125)$$

还有公式 $\sum \binom{k}{n} \delta k = \binom{k}{n+1}$; 方程(5.10); 为了避免被零除, 把它写成 $\sum \binom{k+n+1}{n} \delta k = \binom{k+n+1}{n+1}$, 且取得

$$\sum F\left(\begin{matrix} n+2, 1 \\ 2 \end{matrix} \middle| 1\right)_k \delta k = \frac{1}{n+1} F\left(\begin{matrix} n+2, 1 \\ 1 \end{matrix} \middle| 1\right)_k, \quad n \neq -1. \quad (5.126)$$

当我们超几何表达它时, 等式(5.9)原来和此式等价.

一般, 如果有形式

$$\sum F\left(\begin{matrix} a_1, \dots, a_m, 1 \\ b_1, \dots, b_n \end{matrix} \middle| z\right)_k \delta k = c F\left(\begin{matrix} A_1, \dots, A_M, 1 \\ B_1, \dots, B_N \end{matrix} \middle| z\right)_k \quad (5.127)$$

的求和公式, 则我们对于任何整数 l 有

$$\sum F\left(\begin{matrix} a_1, \dots, a_m, 1 \\ b_1, \dots, b_n \end{matrix} \middle| z\right)_{k+l} \delta k = c F\left(\begin{matrix} A_1, \dots, A_M, 1 \\ B_1, \dots, B_N \end{matrix} \middle| z\right)_{k+l}.$$

把指数移位 l 有一般公式:

$$F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right)_{k+l} = \frac{a_1^l \dots a_m^l z^l}{b_1^l \dots b_n^l l!} F\left(\begin{matrix} a_1+l, \dots, a_m+l, 1 \\ b_1+l, \dots, b_n+l, l+1 \end{matrix} \middle| z\right)_k.$$

因此任何给定的等式(5.127)有无限多个移位形式:

$$\begin{aligned} & \sum F\left(\begin{matrix} a_1+l, \dots, a_m+l, 1 \\ b_1+l, \dots, b_n+l \end{matrix} \middle| z\right)_k \delta k \\ &= c \frac{b_1^l \dots b_n^l}{a_1^l \dots a_m^l} \frac{A_1^l \dots A_M^l}{B_1^l \dots B_N^l} F\left(\begin{matrix} A_1+l, \dots, A_M+l, 1 \\ B_1+l, \dots, B_N+l \end{matrix} \middle| z\right)_k. \end{aligned} \quad (5.128)$$

在这里的 a, A, b 和 B 中间通常有令人满意的数量的删去。例如, 如果应用这个移位公式到式(5.126), 我们取得一般的等式

$$\sum F\left(\begin{matrix} n+l+2, 1 \\ l+2 \end{matrix} \middle| 1\right)_k \delta k = \frac{l+1}{n+1} F\left(\begin{matrix} n+l+2, 1 \\ l+1 \end{matrix} \middle| 1\right)_k, \quad (5.129)$$

对于所有 $n \neq 1$ 成立。式(5.125)的移位形式是

$$\sum F\left(\begin{matrix} l+2, 1 \\ l+1 \end{matrix} \middle| z\right)_k \delta k = \frac{-1}{1-z} \frac{l+1/(1-z)}{l+1} F\left(\begin{matrix} l+1+1/(1-z), 1 \\ l+1/(1-z) \end{matrix} \middle| z\right)_k. \quad (5.130)$$

有一点耐心, 我们能计算出不少可能有用的不定求和等式:

$$\sum F\left(\begin{matrix} a, 2+(1-a)z/(1-z), 1 \\ 1+(1-a)z/(1-z), 2 \end{matrix} \middle| z\right)_k \delta k = \frac{1}{az-1} F\left(\begin{matrix} a, 1 \\ 1 \end{matrix} \middle| z\right)_k; \quad (5.131)$$

$$\sum F\left(\begin{matrix} a, b, 1+(c-ab)/(c-a-b+1), 1 \\ c+1, (c-ab)/(c-a-b+1), 2 \end{matrix} \middle| 1\right)_k \delta k = \frac{c}{ab-c} F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| 1\right)_k; \quad (5.132)$$

$$\sum F\left(\begin{matrix} a, b, 1 \\ c+1, a+b-c+1 \end{matrix} \middle| 1\right)_k \delta k = \frac{(c)(c-b-a)}{(c-a)(c-b)} F\left(\begin{matrix} a, b, 1 \\ c, a+b-c \end{matrix} \middle| 1\right)_k. \quad (5.133)$$

习 题

准备部分

1. 11^4 是多少? 对于懂得二项系数的人, 为什么此数容易计算?
2. 当 n 是给定正整数时, k 的哪一个值使 $\binom{n}{k}$ 为最大? 证明你的回答。
3. 证明六边形性质, $\binom{n-1}{k-1} \binom{n}{k+1} \binom{n+1}{k} = \binom{n-1}{k} \binom{n+1}{k+1} \binom{n}{k-1}$.
4. 通过求反(实际未求反)它的上指数计算 $\binom{-1}{k}$ 。
5. 设 p 是素数, 证明 $\binom{p}{k} \bmod p = 0 (0 < k < p)$ 。其中含有的关于二项系数 $\binom{p-1}{k}$ 的意思是什么?
6. 通过正确地应用对称性, 解决 5.2 节问题 6 中的课文的推导。
7. 当 $k < 0$ 时, 式(5.34)也真吗?
8. 计算 $\sum_k \binom{n}{k} (-1)^k (1-k/n)^n$ 。当 n 十分大时, 此和的近似值是多少? 提示: 对于某个函数 f , 此和是 $\Delta^n f(0)$
9. 证明式(5.58)的广义指数服从定律
 $\mathcal{G}_t(z) = \mathcal{G}(tz)^{1/t}$, 若 $t \neq 0$.

其中 $\mathcal{E}(z)$ 是 $\mathcal{E}_1(z)$ 的缩写.

10. 证明 $-2(\ln(1-z) + z)/z^2$ 是超几何函数.

11. 按照超几何级数表达两个函数

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots$$

$$\arcsin z = z + \frac{1 \cdot z^3}{2 \cdot 3} + \frac{1 \cdot 3 \cdot z^5}{2 \cdot 4 \cdot 5} + \frac{1 \cdot 3 \cdot 5 \cdot z^7}{2 \cdot 4 \cdot 6 \cdot 7} + \cdots$$

12. 下列哪一个 k 的函数是一个“超几何项”(按照式(5.115)的意义)? 说明为什么是或为什么不是.

(a) n^k . (b) k^n . (c) $(k! + (k+1)!)/2$.

(d) H_k , 也就是, $\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{k}$.

(e) $t(k)T(n-k)/T(n)$, 当 t 和 T 是超几何项.

(f) $(t(k) + T(k))/2$, 当 t 和 T 是超几何项.

(g) $(at(k) + bt(k+1) + ct(k+2))/(a + bt(1) + ct(2))$, 当 t 是一个超几何项.

基本部分

13. 求习题 4.55 的超阶乘函数 $P_n = \prod_{k=1}^n k!$, 超阶乘函数 $Q_n = \prod_{k=1}^n k^k$ 以及乘积 $R_n = \prod_{k=0}^n \binom{n}{k}$ 之间的关系.

14. 在 Vandermonde 卷积(5.22)中求反上指数来证明等式(5.25). 然后表明另一个求反产生式(5.26).

15. $\sum_k \binom{n}{k}^3 (-1)^k$ 是什么? 提示: 见式(5.29).

16. 当 a, b, c 为非负整数时, 计算和

$$\sum_k \binom{2a}{a+k} \binom{2b}{b+k} \binom{2c}{c+k} (-1)^k.$$

17. 求 $\binom{2n-1/2}{n}$ 和 $\binom{2n-1/2}{2n}$ 之间的一个简单关系.

18. 对于乘积

$$\binom{r}{k} \binom{r-1/3}{k} \binom{r-2/3}{k}.$$

求一个类似于式(5.35)的另一种形式.

19. 证明式(5.58)的广义二项服从定律

$$\mathcal{B}_i(z) = \mathcal{B}_{1-i}(-z)^{-1}.$$

20. 在式(5.76)中用降幂代替升幂, 由公式

$$G\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) = \sum_{k \geq 0} \frac{a_1^k \dots a_m^k}{b_1^k \dots b_n^k} \frac{z^k}{k!}$$

定义一个“广义非正式的几何级数”, 说明 G 如何与 F 有关系。

21. 通过表明当 $z=m$ 是正整数时, 式(5.83)中的极限为 $1/((m-1)\cdots(1))$, 来指出欧拉的阶乘定义和通常的定义一致。

22. 用式(5.83)来证明阶乘倍量公式:

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

23. $F(-n, 1; ; 1)$ 的值是多少?

24. 用超几何级数来求 $\sum_k \binom{n}{m+k} \binom{m+k}{2k} 4^k$.

25. 证明

$$\begin{aligned} (a_1 - b_1)F\left(\begin{matrix} a_1, a_2, \dots, a_m \\ b_1 + 1, b_2, \dots, b_n \end{matrix} \middle| z\right) \\ = a_1 F\left(\begin{matrix} a_1 + 1, a_2, \dots, a_m \\ b_1 + 1, b_2, \dots, b_n \end{matrix} \middle| z\right) - b_1 F\left(\begin{matrix} a_1, a_2, \dots, a_m \\ b_1, b_2, \dots, b_n \end{matrix} \middle| z\right). \end{aligned}$$

在超几何 $F(a_1, a_2, a_3, \dots, a_m; b_1, \dots, b_n; z)$, $F(a_1 + 1, a_2, \dots, a_m; b_1, \dots, b_n; z)$ 和 $F(a_1, a_2 + 1, a_3, \dots, a_m; b_1, \dots, b_n; z)$ 之间找一个相似的关系。

26. 如同超几何级数的倍数那样在公式

$$F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) = 1 + G(z)$$

中表达函数 $G(z)$,

27. 证明

$$\begin{aligned} F\left(\begin{matrix} a_1, a_1 + \frac{1}{2}, \dots, a_m, a_m + \frac{1}{2} \\ b_1, b_1 + \frac{1}{2}, \dots, b_n, b_n + \frac{1}{2} \end{matrix} \middle| (2^{m-n-1}z)^2\right) \\ = \frac{1}{2} \left(F\left(\begin{matrix} 2a_1, \dots, 2a_m \\ 2b_1, \dots, 2b_n \end{matrix} \middle| z\right) + F\left(\begin{matrix} 2a_1, \dots, 2a_m \\ 2b_1, \dots, 2b_n \end{matrix} \middle| -z\right) \right). \end{aligned}$$

28. 应用 Pfaff 反射定律(5.101)两次来证明欧拉等式

$$F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| z\right) = (1-z)^{c-a-b} F\left(\begin{matrix} c-a, c-b \\ c \end{matrix} \middle| z\right).$$

29. 表明合流超几何满足

$$e^z F\left(\begin{matrix} a \\ b \end{matrix} \middle| -z\right) = F\left(\begin{matrix} b-a \\ b \end{matrix} \middle| z\right).$$

30. 什么样的超几何级数 F 满足 $zF'(z) + F(z) = 1/(1-z)$?

31. 表明如果 $f(k)$ 是按超几何项可求和的任何函数, 则 f 是超几何项的倍数. 换句话说, 如果 $\sum f(k)\delta k = cF(A_1, \dots, A_M; B_1, \dots, B_N; Z)_k + C$, 则存在常数 $a_1, \dots, a_m, b_1, \dots, b_n$ 和 z 使得 $f(k)$ 是一个常数乘 $F(a_1, \dots, a_m; b_1, \dots, b_n; z)_k$.

32. 用 Gosper 方法求 $\sum k^2 \delta k$.

33. 用 Gosper 方法求 $\sum \delta k / (k^2 - 1)$.

34. 指出一个部分超几何和总能表为通常超几何的极限: 当 c 是非负整数时

$$\sum_{k \leq n} F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right)_k = \lim_{\varepsilon \rightarrow 0} F\left(\begin{matrix} -c, a_1, \dots, a_m \\ \varepsilon - c, b_1, \dots, b_n \end{matrix} \middle| z\right).$$

用此思想来计算 $\sum_{k \leq n} \binom{n}{k} (-1)^k$.

课外习题

35. 没有前后关系, 记法 $\sum_{k \leq n} \binom{n}{k} 2^{k-n}$ 有两种解释.

计算它:

(a) 作为 k 的和. (b) 作为 n 的和.

36. 设 p^k 是除尽 $\binom{m+n}{m}$ 的素数 p 的最大幂, 其中 m 和 n 是非负整数. 证明 k 是当在基数 p 数系中 m 被加到 n 时出现的进位次数. 提示: 这里用习题 4.24.

37. 证明阶乘幂成立的类似的二项定理. 也就是说, 证明等式: 对于所有非负整数 n

$$(x+y)^{\overline{n}} = \sum_k \binom{n}{k} x^{\overline{k}} y^{\overline{n-k}},$$

$$(x+y)^{\bar{n}} = \sum_k \binom{n}{k} x^{\bar{k}} y^{\bar{n-k}}.$$

38. 证明所有非负整数 n 能唯一表为形式 $n = \binom{a}{1} + \binom{b}{2} + \binom{c}{3}$, 其中 a, b 和 c 是整数, 且 $0 \leq a < b < c$. (称这为二项数系.)

39. 证明对于所有 $n > 0$, 如果 $xy = ax + by$, 则 $x^{\overline{n}} y^{\overline{m}} = \sum_{k=1}^n \binom{2n-1-k}{n-1} (a^{\overline{k}} b^{\overline{n-k}} x^{\overline{k}} + a^{\overline{n-k}} b^{\overline{n}} y^{\overline{k}})$. 找较一般的乘积 $x^{\overline{m}} y^{\overline{n}}$ 的一个相似公式.

40. 求

$$\sum_{j=1}^m (-1)^{j+1} \binom{r}{j} \sum_{k=1}^n \binom{-j+rk-s}{m-j}, \text{ 整数 } m, n \geq 0$$

的一个闭形式。

41. 当 n 是非负整数时, 计算 $\sum_k \binom{n}{k} k! / (n+1+k)!$ 。

42. 求不定和 $\sum \left((-1)^x / \binom{n}{x} \right) \delta x$, 且用它来计算和 $\sum_{k=0}^n (-1)^k / \binom{n}{k}$ 为闭形式。

43. 证明三个二项系数等式 (5.28)。提示: 首先用 $\sum_j \binom{r}{m+n-j} \binom{k}{j}$ 替换 $\binom{r+k}{m+n}$ 。

44. 用等式 (5.32) 来找双重和

$$\sum_{j,k} (-1)^{j+k} \binom{j+k}{j} \binom{a}{j} \binom{b}{k} \binom{m+n-j-k}{m-j} \text{ 和}$$

$$\sum_{j,k \geq 0} (-1)^{j+k} \binom{a}{j} \binom{m}{j} \binom{b}{k} \binom{n}{k} / \binom{m+n}{j+k}$$

的闭形式, 给定整数 $m \geq a \geq 0$ 和 $n \geq b \geq 0$ 。

45. 求 $\sum_{k \leq n} \binom{2k}{k} 4^{-k}$ 的闭形式。

46. 当 n 是正整数时, 计算下列和为闭形式:

$$\sum_k \binom{2k-1}{k} \binom{4n-2k-1}{2n-k} \frac{(-1)^{k-1}}{(2k-1)(4n-2k-1)}.$$

提示: 母函数再次获得成功。

47. 和 $\sum_k \binom{rk+s}{k} \binom{rn-rk-s}{n-k}$ 是 r 和 s 的一个多项式。指出它不依赖于 s 。

48. 等式 $\sum_{k \leq n} \binom{n+k}{n} 2^{-k} = 2^n$ 与 $\sum_{k \geq 0} \binom{n+k}{n} z^k = 1 / (1-z)^{n+1}$ 结合起来能产生 $\sum_{k \geq n} \binom{n+k}{n} 2^{-k} = 2^n$ 。后面等式的超几何形式是什么?

49. 用超几何方法来计算

$$\sum_k (-1)^k \binom{x}{k} \binom{x+n-k}{n-k} \frac{y}{y+n-k}.$$

50. 用比较方程两边的 z^n 的系数来证明 Pfaff 反射定律 (5.101)。

51. 式 (5.104) 的推导表明

$$\lim_{\varepsilon \rightarrow 0} F(-m, -2m-1+\varepsilon; -2m+\varepsilon; 2) = 1 / \binom{-1/2}{m}.$$

在此习题中我们将看出, 稍微不同的极限过程导致明显不同的退化超几何级数 $F(-m, -2m-1; -2m; 2)$ 的解答.

(a) 证明 $\lim_{\varepsilon \rightarrow 0} F(-m+\varepsilon, -2m-1; -2m+2\varepsilon; 2) = 0$, 通过用 Pfaff 反射定律证明对于所有整数 $m \geq 0$, 等式 $F(a, -2m-1; 2a; 2) = 0$.

(b) $\lim_{\varepsilon \rightarrow 0} F(-m+\varepsilon, -2m-1; -2m+\varepsilon; 2)$ 是多少?

52. 证明如果 N 是非负整数,

$$\begin{aligned} b_1^{\bar{N}} \cdots b_n^{\bar{N}} F\left(\begin{matrix} a_1, \dots, a_m, -N \\ b_1, \dots, b_n \end{matrix} \middle| z\right) \\ = a_1^{\bar{N}} \cdots a_m^{\bar{N}} (-z)^N F\left(\begin{matrix} 1-b_1-N, \dots, 1-b_n-N, -N \\ 1-a_1-N, \dots, 1-a_m-N \end{matrix} \middle| \frac{(-1)^{m+n}}{z}\right). \end{aligned}$$

53. 如果在高斯等式(5.110)中置 $b = -1/2$ 和 $z = 1$, 则左边化为 -1 而右边为 $+1$. 为什么这并不证明 $-1 = +1$?

54. 说明怎样获得式(5.112)的右边.

55. 如果对于所有 $k \geq 0$ 超几何项 $t(k) = F(a_1, \dots, a_m; b_1, \dots, b_n; z)_k$ 以及 $T(k) = F(A_1, \dots, A_M; B_1, \dots, B_N; Z)_k$ 满足 $t(k) = c(T(k+1) - T(k))$, 证明 $z = Z$ 和 $m - n = M - N$.

56. 用 Gosper 的方法求 $\sum \binom{-3}{k} \delta k$ 的一般公式, 指出 $(-1)^{k-1} \lfloor (k+1)/2 \rfloor \lceil (k+2)/2 \rceil$ 也是一个解.

57. 用 Gosper 的方法来找一个常数 θ 使得

$$\sum \binom{n}{k} z^k (k + \theta) \delta k$$

是以超几何项可求和的.

58. 若 m 和 n 是整数 ($0 \leq m \leq n$), 设

$$T_{m,n} = \sum_{0 \leq k \leq n} \binom{k}{m} \frac{1}{n-k}.$$

求 $T_{m,n}$ 和 $T_{m-1,n-1}$ 之间的一个关系, 然后应用一个求和因子来解你的递归.

考查性问题

59. 当 m 和 n 是正整数时, 求

$$\sum_{k \geq 1} \binom{n}{\lfloor \log_m k \rfloor}$$

的一个闭形式。

60. 当 m 和 n 都大时用 Stirling 近似式(4.23)来估计 $\binom{m+n}{n}$, 当 $m=n$ 时你的公式化为什么?

61. 证明当 p 是素数时, 对于所有非负整数 m 和 n , 我们得到

$$\binom{n}{m} \equiv \binom{\lfloor n/p \rfloor}{\lfloor m/p \rfloor} \binom{n \bmod p}{m \bmod p} \pmod{p}.$$

62. 假设 p 是素数, m 和 n 是正整数, 确定 $\binom{np}{mp} \bmod p^2$ 的值. 提示: 你可用下列 Vandermonde 卷积的推广式:

$$\sum_{k_1+k_2+\dots+k_m=n} \binom{r_1}{k_1} \binom{r_2}{k_2} \dots \binom{r_m}{k_m} = \binom{r_1+r_2+\dots+r_m}{n}.$$

63. 给定一个整数 $n \geq 0$, 求

$$\sum_{k=0}^n (-4)^k \binom{n+k}{2k}$$

的一个闭形式。

64. 给定整数 $n \geq 0$, 计算 $\sum_{k=0}^n \binom{n}{k} \left\lceil \frac{k+1}{2} \right\rceil$.

65. 证明

$$\sum_k \binom{n-1}{k} n^{-k} (k+1)! = n.$$

66. 作为 m 的函数, 计算“Harry 双重和”

$$\sum_{0 \leq j \leq k} \binom{-1}{j - \lfloor \sqrt{k-j} \rfloor} \binom{j}{m} \frac{1}{2^j}, \text{ 整数 } m \geq 0.$$

, (在 j 和 k 上求和。)

67. 求

$$\sum_{k=0}^n \binom{k}{2} \binom{2n-k}{n}, \text{ 整数 } n \geq 0$$

的一个闭形式。

68. 求

$$\sum_k \binom{n}{k} \min(k, n-k), \text{ 整数 } n \geq 0$$

的一个闭形式。

69. 作为 m 和 n 的一个函数, 求

$$\min_{\substack{k_1, \dots, k_m \geq 0 \\ k_1 + \dots + k_m = n}} \sum_{j=1}^m \binom{k_j}{2}$$

的一个闭形式。

70. 求

$$\sum_k \binom{n}{k} \binom{2k}{k} \left(\frac{-1}{2}\right)^k, \text{ 整数 } n \geq 0$$

的一个闭形式。

71. 设

$$S_n = \sum_{k \geq 0} \binom{n+k}{m+2k} a_k,$$

其中 m 和 n 是非负整数, 且设 $A(z) = \sum_{k \geq 0} a_k z^k$ 是序列 $\langle a_0, a_1, a_2, \dots \rangle$ 的母函数。

(a) 以 $A(z)$ 表达母函数 $S(z) = \sum_{n \geq 0} S_n z^n$ 。

(b) 用此技巧来解 5.2 节中的问题 7。

72. 证明, 如果 m, n 和 k 是整数, 且 $n > 0$, 则 $\binom{m/n}{k} n^{2k - v(k)}$ 是一个整数。其中 $v(k)$ 是 k 的二进制表示中的 1 的个数。

73. 用包含各种组成部分的方法来解决递归

$$X_0 = \alpha; \quad X_1 = \beta;$$

$$X_n = (n-1)(X_{n-1} + X_{n-2}), \quad n > 1.$$

提示: $n!$ 和 $n!$ 都满足此递归。

74. 下图为 Pascal 三角形的一种异常的形式, 虽然内部数仍满足加法公式, 但其两侧包含数 1, 2, 3, 4, ... 而不是全为 1。

$$\begin{array}{ccccccc} & & & & 1 & & & & \\ & & & & & & & & \\ & & 2 & & & 2 & & & \\ & & & & & & & & \\ & 3 & & 4 & & 3 & & & \\ & & & & & & & & \\ & 4 & & 7 & & 7 & & 4 & \\ & & & & & & & & \\ & 5 & & 11 & & 14 & & 11 & & 5 \\ & & & & & & & & & \\ & \cdot & & \cdot & & \cdot & & \cdot & & \cdot \end{array}$$

如果对于 $1 \leq k \leq n$, $\left(\binom{n}{k}\right)$ 表示行 n 中第 k 个数, 我们有 $\left(\binom{n}{1}\right) = \left(\binom{n}{n}\right) = n$ 和 $\left(\binom{n}{k}\right) = \left(\binom{n-1}{k}\right) + \left(\binom{n-1}{k-1}\right)$, ($1 < k < n$). 表达量 $\left(\binom{n}{k}\right)$ 为闭形式.

75. 求函数

$$S_0(n) = \sum_k \binom{n}{3k},$$

$$S_1(n) = \sum_k \binom{n}{3k+1},$$

$$S_2(n) = \sum_k \binom{n}{3k+2}$$

之间的一个关系以及量 $\lfloor 2^n/3 \rfloor$ 和 $\lceil 2^n/3 \rceil$.

76. 对于 $n, k \geq 0$ 解下列递归:

$$Q_{n,0} = 1; Q_{0,k} = [k=0];$$

$$Q_{n,k} = Q_{n-1,k} + Q_{n-1,k-1} + \binom{n}{k}, \quad n, k > 0.$$

77. $\sum_{0 \leq k_1, \dots, k_m \leq n} \prod_{1 \leq i \leq m} \binom{k_{i+1}}{k_i}$ (如果 $m > 1$) 的值是多少?

78. 假设 m 是正整数, 求

$$\sum_{k=0}^{2m^2} \binom{k \bmod m}{(2k+1) \bmod (2m+1)}$$

的一个闭形式.

79. (a) $\binom{2n}{1}, \binom{2n}{3}, \dots, \binom{2n}{2n-1}$ 的最大公约数是什么? 提示: 考虑这 n 个数的和.

(b) 证明 $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$ 的最小公倍数等于 $L(n+1)/(n+1)$, 其中 $L(n) = \text{lcm}(1, 2, \dots, n)$.

80. 对于所有整数 $k, n \geq 0$, 证明 $\binom{n}{k} \leq (en/k)^k$.

81. 若 $0 < \theta < 1$ 且 $0 \leq x \leq 1$, 如果 l, m, n 是非负整数 ($m < n$), 证明不等式

$$(-1)^{n-m-1} \sum_k \binom{l}{k} \binom{m+\theta}{n+k} x^k > 0.$$

提示: 考虑取 x 的导数.

额外问题

82. 证明 Pascal 三角形有一个比课文中引用的更意外的六边形性质: 若 $0 < k < n$, 则

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

例如, $\gcd(56, 36, 210) = \gcd(28, 120, 126) = 2$.

83. 证明令人惊异的等式(5.32), 首先证明每当右边为零时它为真.

84. 根据第一对卷积公式(5.60)证明第二对卷积公式(5.61). 提示: 对 z 微分.

85. 证明

$$\sum_{m=1}^n (-1)^m \sum_{1 \leq k_1 \leq k_2 \leq \dots \leq k_m \leq n} \binom{k_1^3 + k_2^3 + \dots + k_m^3 + 2^n}{n} = (-1)^n n!^3 - \binom{2^n}{n}.$$

(左边是 $2^n - 1$ 项的和.) 提示: 更加是真的了.

86. 当 $n(n-1)$ 个因子

$$\prod_{\substack{1 \leq i, j \leq n \\ i \neq j}} \left(1 - \frac{z_i}{z_j}\right)^{a_i}$$

全部展成复变数 z_1, \dots, z_n 的正幂和负幂时, 设 a_1, \dots, a_n 为非负整数, $C(a_1, \dots, a_n)$ 是常数项 $z_1^0 \dots z_n^0$ 的系数.

(a) 证明 $C(a_1, \dots, a_n)$ 等于式(5.31)的左边.

(b) 证明如果 z_1, \dots, z_n 是不同的复数, 则多项式

$$f(z) = \sum_{k=1}^n \prod_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{z - z_j}{z_k - z_j}$$

同样等于 1.

(c) 用 $f(0)$ 乘原来的 $n(n-1)$ 个因子的乘积, 且推出 $C(a_1, \dots, a_n)$ 等于

$$C(a_1 - 1, a_2, \dots, a_n) + C(a_1, a_2 - 1, \dots, a_n) + \dots + C(a_1, a_2, \dots, a_n - 1).$$

(此递归定义多项系数, 所以 $C(a_1, \dots, a_n)$ 一定等于式(5.31)的右边.)

87. 设 m 是正整数, 且设 $\zeta = e^{2\pi i/m}$. 证明

$$\begin{aligned} & \sum_{k \leq n/m} \binom{n-mk}{k} z^{mk} \\ &= \frac{\mathcal{B}_{-m}(z^m)^{n+1}}{(1+m)\mathcal{B}_{-m}(z^m) - m} - \sum_{0 \leq i < m} \frac{\left(\zeta^{2i+1} z \mathcal{B}_{1+1/m}(\zeta^{2i+1} z)^{1/m}\right)^{n+1}}{(m+1)\mathcal{B}_{1+1/m}(\zeta^{2i+1} z)^{-1} - 1}. \end{aligned}$$

(在特殊情形 $m=1$ 时, 此式化为式(5.74).)

88. 证明对于所有 $k > 1$, 式(5.47)中的系数 s_k 等于

$$(-1)^k \int_0^{\infty} e^{-t} (1 - e^{-t})^{k-1} \frac{dt}{t},$$

因此 $|s_k| < 1/(k-1)$.

89. 证明如果 $|x| < |y|$ 和 $|x| < |x+y|$, 式(5.19)有一个无限配对式

$$\sum_{k \geq m} \binom{m+r}{k} x^k y^{m-k} = \sum_{k \geq m} \binom{-r}{k} (-x)^k (x+y)^{m-k}, \text{ 整数 } m.$$

对 y 微分此等式 n 次, 且按超几何表达它, 你取得什么关系?

90. 当 r 和 s 是整数时 ($s \geq r \geq 0$), 5.2 节中的问题 1 考虑 $\sum_{k \geq 0} \binom{r}{k} / \binom{s}{k}$. 如果 r 和 s 不是整数, 此和的值是什么?

91. 证明 Clausen 乘积等式

$$F\left(\begin{matrix} a, b \\ a+b+\frac{1}{2} \end{matrix} \middle| z\right)^2 = F\left(\begin{matrix} 2a, a+b, 2b \\ 2a+2b, a+b+\frac{1}{2} \end{matrix} \middle| z\right);$$

$$F\left(\begin{matrix} \frac{1}{4}+a, \frac{1}{4}+b \\ 1+a+b \end{matrix} \middle| z\right) F\left(\begin{matrix} \frac{1}{4}-a, \frac{1}{4}-b \\ 1-a-b \end{matrix} \middle| z\right) = F\left(\begin{matrix} \frac{1}{2}, \frac{1}{2}+a-b, \frac{1}{2}-a-b \\ 1+a+b, 1-a-b \end{matrix} \middle| z\right).$$

当这些公式两边的 z^n 的系数相等时, 产生什么等式?

92. 通过指明两边满足相同的微分方程来证明 Whipple 等式:

$$F\left(\begin{matrix} \frac{1}{2}a, \frac{1}{2}a+\frac{1}{2}, 1-a-b-c \\ 1+a-b, 1+a-c \end{matrix} \middle| \frac{-4z}{(1-z)^2}\right) = (1-z)^a F\left(\begin{matrix} a, b, c \\ 1+a-b, 1+a-c \end{matrix} \middle| z\right).$$

93. 给定任何函数 f 和任何常数 α , 证明不定和

$$\sum \left(\prod_{j=1}^{k-1} (f(j) + \alpha) / \prod_{j=1}^k f(j) \right) \delta k$$

有一个(相当)简单的形式.

94. 证明如果 $\omega = e^{2\pi i/3}$, 我们有

$$\sum_{k+l+m=3n} \binom{3n}{k, l, m} \omega^{l+2m} = \binom{4n}{n, n, 2n}, \text{ 整数 } n \geq 0.$$

研究性问题

95. 设 $q(n)$ 是中间的二项系数 $\binom{2n}{n}$ 的最小奇素数因子. 按照习题 36, 除不尽 $\binom{2n}{n}$

的奇素数 p 是这样一些数, 对于这些数在 n 的基数 p 的表示中的所有数字位是 $(p-1)/2$ 或更小. 计算机试验表明, 除了 $q(3\,160)=13$ 之外, 对于所有 $n < 10^{10\,000}$, $q(n) \leq 11$.

(a) 对于所有 $n > 3\,160$, $q(n) \leq 11$ 吗?

(b) 对于无限多 n , $q(n) = 11$ 吗?

96. 对于所有 $n > 4$, 一个素数的平方可除尽 $\binom{2n}{n}$ 吗?

97. 对于 n 的什么值产生 $\binom{2n}{n} \equiv (-1)^n \pmod{(2n+1)}$?

第六章 特殊数

在数学中常出现一些数序列，我们可立刻辨认并给它们以特殊名称。例如，学过算术的人知道平方数序列 $\langle 1, 4, 9, 16, \dots \rangle$ 。在第一章中我们遇到三角形数 $\langle 1, 3, 6, 10, \dots \rangle$ ；在第四章中我们研究素数 $\langle 2, 3, 5, 7, \dots \rangle$ ；在第五章中我们简要地查看了 Catalan 数 $\langle 1, 2, 5, 14, \dots \rangle$ 。

在本章中将接触到几个其他的重要序列。首先我们将研究 Stirling 数 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 和 $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ 以及 Eulerian 数 $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle$ ，这些数形成的系数的三角形型式相似于 Pascal 三角形中的二项系数 $\binom{n}{k}$ 。然后我们将充分查看调和数 H_n 和 Bernoulli 数 B_n ，这些数不同于研究过的其他序列，因为它们是分数，而不是整数。最后，我们将查看吸引人的 Fibonacci 数 F_n 和它们的某些重要的推广。

6.1 Stirling 数

我们从与二项系数有某些紧密联系的 Stirling 数开始，以 Jame Stirling(1692~1770) 的名字取名。这些数有两种形式，通常称为“第一类和第二类 Stirling 数”。虽然它们有悠久的历史 and 许多应用，但仍然没有标准的记法。我们将用 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 记第二类 Stirling 数，用 $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ 记第一类 Stirling 数，因为这些符号比许多其他人们所用的记法更为用户所支持。

当 n 和 k 小时，表 6.1 和 6.2 表明了 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 和 $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ 的值。涉及数“1, 7, 6, 1”的问题可能与 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 有关，涉及“6, 11, 6, 1”的问题可能与 $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ 有关，就像我们设想涉及“1, 4, 6, 4, 1”的问题可能与 $\binom{n}{k}$ 有关；这些是当 $n=4$ 时出现的标志序列。

第二类 Stirling 数比其他类更常出现，所以让我们首先考虑它。符号 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 代表把 n 个事物的集合划分为 k 个非空子集的方式数。例如，把 4 个元素的集合分为两部分有 7 种方式：

$$\{1, 2, 3\} \cup \{4\}, \{1, 2, 4\} \cup \{3\}, \{1, 3, 4\} \cup \{2\}, \{2, 3, 4\} \cup \{1\}, \\ \{1, 2\} \cup \{3, 4\}, \{1, 3\} \cup \{2, 4\}, \{1, 4\} \cup \{2, 3\}; \quad (6.1)$$

于是 $\left\{ \begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right\} = 7$ 。注意，如同数 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 那样又用大括号表示集合，这种记法关系帮助我们记住 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 的意义，把它读作“ n 子集 k ”。

表 6.1 子集的 Stirling 三角形

n	$\left\{ \begin{smallmatrix} n \\ 0 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 1 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 2 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 3 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 4 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 5 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 6 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 7 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 8 \end{smallmatrix} \right\}$	$\left\{ \begin{smallmatrix} n \\ 9 \end{smallmatrix} \right\}$
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

让我们查看小的 k 。恰有一种方式把 n 个元素放入单个非空集合，因此 $\left\{ \begin{smallmatrix} n \\ 1 \end{smallmatrix} \right\} = 1$ ，对于所有 $n > 0$ 。另一方面 $\left\{ \begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right\} = 0$ ，因为 0 个元素的集合是空的。

情形 $k=0$ 有点微妙。若我们约定恰有一种方式把一个空集划分为零个非空部分，则情况变好，因此 $\left\{ \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right\} = 1$ 。但是一个非空集合至少需要一个部分，所以 $n > 0$ ， $\left\{ \begin{smallmatrix} n \\ 0 \end{smallmatrix} \right\} = 0$ 。

当 $k=2$ 时出现什么？确实 $\left\{ \begin{smallmatrix} 0 \\ 2 \end{smallmatrix} \right\} = 0$ 。如果 $n > 0$ 个元素的一个集合分为两个非空部分，这些部分之一包含最后一个元素和前 $n-1$ 个元素的某个子集。有 2^{n-1} 种方式选取后面的子集，因为前 $n-1$ 个元素的每一个或者在它之内或者在它之外；但是不一定要把所有这样的元素放入它的内部，因为我们要以两个非空部分结束。所以我们减 1：

$$\left\{ \begin{smallmatrix} n \\ 2 \end{smallmatrix} \right\} = 2^{n-1} - 1, \text{ 整数 } n > 0. \quad (6.2)$$

(这与我们在上面枚举的 $\left\{ \begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right\} = 7 = 2^3 - 1$ 方式一致。)

修改这个论证引出一个递归，用它能对所有 k 计算 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ ：给定 $n > 0$ 个元素的一个集合，把它划分为 k 个非空部分，我们或者把最后元素单独放入一类(以 $\left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\}$ 种方式)，

或者把它和前 $n-1$ 个元素的某个非空子集放在一起。后一种情形有 $k\left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\}$ 种可能性, 因为 $\left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\}$ 种方式的每一个方式把前 $n-1$ 个元素分配入 k 个非空部分, 给出第 n 个元素能加入的 k 个子集。因此

$$\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = k \left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\} + \left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\}, \text{ 整数 } n > 0. \quad (6.3)$$

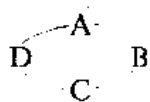
表 6.2 轮换的 Stirling 三角形

n	$\left[\begin{smallmatrix} n \\ 0 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 1 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 2 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 3 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 4 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 5 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 6 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 7 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 8 \end{smallmatrix} \right]$	$\left[\begin{smallmatrix} n \\ 9 \end{smallmatrix} \right]$
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

这是产生表 6.1 的定律。没有 k 的因子, 它将化为产生 Pascal 三角形的加法公式(5.8)。

现在考虑第一类 Stirling 数。这些数有点不同, $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ 计算 n 个元素安排成 k 个轮换 (而不是子集) 的方式数。我们把 $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ 称为 ' n 轮换 k '。

轮换是循环排列, 像第四章中考虑的项链。轮换



能更紧凑地记为 $[A, B, C, D]$, 据此条件

$$[A, B, C, D] = [B, C, D, A] = [C, D, A, B] = [D, A, B, C];$$

一个轮换“环绕整整一圈”, 因为它的终端连接它的开始。另一方面, 轮换 $[A, B, C, D]$ 不同于 $[A, B, D, C]$ 或 $[D, C, B, A]$ 。

有 11 种不同方式从 4 个元素来形成两个轮换:

$$\begin{aligned} & [1, 2, 3][4], [1, 2, 4][3], [1, 3, 4][2], [2, 3, 4][1], \\ & [1, 3, 2][4], [1, 4, 2][3], [1, 4, 3][2], [2, 4, 3][1], \end{aligned}$$

$$[1, 2][3, 4], [1, 3][2, 4], [1, 4][2, 3]; \quad (6.4)$$

因此 $\left[\begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right] = 11$.

单个元素的一个轮换(也就是说, 仅有一个元素的一个轮换)实际上与单个元素的一个集合(仅有一个元素的一个集合)相同。类似, 2个元素的一个轮换像2个元素的一个集合, 因为我们有 $[A, B] = [B, A]$ 就像 $\{A, B\} = \{B, A\}$ 。但是有两个不同的3个元素的轮换 $[A, B, C]$ 和 $[A, C, B]$ 。例如, 注意到, 通过从3个元素集合的每一个集合形成二个轮换, 再从式(6.1)中7个集合对能获得式(6.4)中的11个轮换对。

一般, 每当 $n > 0$ 时, 从任何 n 个元素集合能形成 $n! / n = (n-1)!$ 个轮换(有 $n!$ 个排列, 且每个轮换对应于它们的 n , 因为它的任何一个元素能列在第一个。)所以我们有

$$\left[\begin{smallmatrix} n \\ 1 \end{smallmatrix} \right] = (n-1)!, \text{ 整数 } n > 0. \quad (6.5)$$

这比 Stirling 子集数的值 $\left\{ \begin{smallmatrix} n \\ 1 \end{smallmatrix} \right\} = 1$ 大得多。事实上, 易见轮换数至少与子集数一样大,

$$\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] \geq \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}, \text{ 整数 } n, k \geq 0, \quad (6.6)$$

因为每一个划分为非空子集至少引出轮换的一个排列。

当所有的轮换一定是单个元素或两个元素时, 式(6.6)中的等号成立, 因为在这种情形轮换等价于子集。当 $k=n$ 和 $k=n-1$ 时出现这种情形, 因此

$$\left[\begin{smallmatrix} n \\ n \end{smallmatrix} \right] = \left\{ \begin{smallmatrix} n \\ n \end{smallmatrix} \right\}, \quad \left[\begin{smallmatrix} n \\ n-1 \end{smallmatrix} \right] = \left\{ \begin{smallmatrix} n \\ n-1 \end{smallmatrix} \right\}.$$

事实上, 易见

$$\left[\begin{smallmatrix} n \\ n \end{smallmatrix} \right] = \left\{ \begin{smallmatrix} n \\ n \end{smallmatrix} \right\} = 1, \quad \left[\begin{smallmatrix} n \\ n-1 \end{smallmatrix} \right] = \left\{ \begin{smallmatrix} n \\ n-1 \end{smallmatrix} \right\} = \binom{n}{2}. \quad (6.7)$$

(排列 n 个元素为 $n-1$ 个轮换或子集的方式数是相同轮换或子集中选取2个元素的方式数。)在表6.1和表6.2中都明显地提出了三角形数 $\binom{n}{2} = 1, 3, 6, 10, \dots$ 。

修改我们对 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 所用的论证, 导出 $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ 的一个递归。在 k 个轮换中, n 个元素的每一种排列或者把最后元素作为单独一个轮换(以 $\left[\begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right]$ 种方式)或者把这元素插入前 $n-1$ 个元素的 $\left[\begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right]$ 种轮换排列的一个中。在后一种情形, 有 $n-1$ 种不同方式做插入。(这需作某些思考, 但是不难验证, 为了形成一个 $(j+1)$ 个元素的轮换, 有 j 种方式把一个新元素放入一个 j 个元素的轮换。当 $j=3$ 时, 例如, 当插入新元素 D 时, 轮换 $[A, B, C]$ 引出

$[A, B, C, D], [A, B, D, C]$ 或 $[A, D, B, C]$, 且没有其他可能性。计算所

有 j 上的总数给出总共 $n-1$ 种方式把第 n 个元素插入 $n-1$ 个元素的一个轮换分解。所以希望的递归是

$$\left[\begin{matrix} n \\ k \end{matrix} \right] = (n-1) \left[\begin{matrix} n-1 \\ k \end{matrix} \right] + \left[\begin{matrix} n-1 \\ k-1 \end{matrix} \right], \text{ 整数 } n > 0. \quad (6.8)$$

这是产生表 6.2 的相似的加法公式。

比较式(6.8)和(6.3)表明, 在 Stirling 轮换数的情形中右边第一项用它的上指数 $(n-1)$ 乘, 而在 Stirling 子集数的情形用它的下指数 k 乘。所以当我们用数学归纳法证明时, 能在像 $n \left[\begin{matrix} n \\ k \end{matrix} \right]$ 和 $k \left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ 的项中完成“吸收”。

每一个排列对应于轮换的一个集合。例如, 考虑 123456789 变为 384729156 的排列。我们能方便地以两行表示它

$$\begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 4 & 7 & 2 & 9 & 1 & 5 & 6, \end{array}$$

表明 1 到 3, 2 到 8, 等等。由 1 到 3, 3 到 4, 4 到 7, 7 回到 1 轮换结构; 这是轮换 $[1, 3, 4, 7]$ 。此排列中的另一个轮换为 $[2, 8, 5]$, 还有一个轮换为 $[6, 9]$ 。所以排列 384729156 等价于轮换配置

$$[1, 3, 4, 7][2, 8, 5][6, 9].$$

如果有 $[1, 2, \dots, n]$ 的任何排列 $\pi_1 \pi_2 \dots \pi_n$, 则每一个元素在一个唯一的轮换中, 如果从 $m_0 = m$ 开始, 且查看 $m_1 = \pi_{m_0}$, $m_2 = \pi_{m_1}$, 等等, 最终一定回到 $m_k = m_0$ 。(数迟早一定要重复, 且重新出现的第一个数一定是 m_0 , 因为我们知道它为其他数 m_1, m_2, \dots, m_{k-1} 的唯一的前面的数。)所以每一个排列确定一个轮换配置。相反, 我们把构造颠倒过来, 每一个轮换配置显然确定一个排列, 此一一对应表明轮换配置和排列本质上是相同的。

所以 $\left[\begin{matrix} n \\ k \end{matrix} \right]$ 是恰好包含 k 个轮换的 n 个元素的排列个数。如果对所有 k 上的 $\left[\begin{matrix} n \\ k \end{matrix} \right]$ 求和, 我们一定取得总排列数:

$$\sum_{k=0}^n \left[\begin{matrix} n \\ k \end{matrix} \right] = n!, \text{ 整数 } n \geq 0. \quad (6.9)$$

例如, $6+11+6+1=24=4!$ 。

Stirling 数是有用的, 因为在许多问题中出现递归关系(6.3)和(6.8)。例如, 如果要用下降幂 $x^{\underline{n}}$ 表示通常的幂 x^n , 我们求得前几种情形是:

$$\begin{aligned} x^0 &= x^{\underline{0}}; \\ x^1 &= x^{\underline{1}}; \\ x^2 &= x^{\underline{2}} + x^{\underline{1}}; \\ x^3 &= x^{\underline{3}} + 3x^{\underline{2}} + x^{\underline{1}}; \end{aligned}$$

$$x^4 = x^4 + 6x^3 + 7x^2 + x^1.$$

这些系数怀疑像表 6.1 中的数，左右之间反射；所以我们能颇相信一般公式为

$$x^n = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^k, \text{ 整数 } n \geq 0. \quad (6.10)$$

且确实是如此，归纳法的简单证明确定这个论证：我们有 $x \cdot x^k = x^{k+1} + kx^k$ ，因为 $x^{k+1} = x^k(x-k)$ ；因此 $x \cdot x^{n-1}$ 是

$$\begin{aligned} x \sum_k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} x^k &= \sum_k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} x^{k+1} + \sum_k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} kx^k \\ &= \sum_k \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\} x^k + \sum_k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} kx^k \\ &= \sum_k \left(k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} + \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\} \right) x^k = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^k. \end{aligned}$$

换句话说，Stirling 子集数是产生通常幂的阶乘幂的系数。

我们还可求助于其他方式，因为 Stirling 轮换数是产生阶乘幂的通常幂的系数：

$$\begin{aligned} x^{\bar{0}} &= x^0; \\ x^{\bar{1}} &= x^1; \\ x^{\bar{2}} &= x^2 + x^1; \\ x^{\bar{3}} &= x^3 + 3x^2 + 2x^1; \\ x^{\bar{4}} &= x^4 + 6x^3 + 11x^2 + 6x^1. \end{aligned}$$

我们有 $(x+n-1) \cdot x^k = x^{k+1} + (n-1)x^k$ ，所以像刚才给出的证明那样表明

$$(x+n-1)x^{n-1} = (x+n-1) \sum_k \left[\begin{matrix} n-1 \\ k \end{matrix} \right] x^k = \sum_k \left[\begin{matrix} n \\ k \end{matrix} \right] x^k.$$

这导致一般公式

$$x^n = \sum_k \left[\begin{matrix} n \\ k \end{matrix} \right] x^k, \text{ 整数 } n \geq 0 \quad (6.11)$$

的归纳法的证明。（置 $x=1$ 再给出式(6.9).）

但是等一等。此方程涉及上升阶乘幂 $x^{\bar{n}}$ ，而式(6.10)涉及下降阶乘幂 $x^{\underline{n}}$ 。如果要依据通常幂表达 $x^{\bar{n}}$ ，或者要依据上升幂表达 $x^{\underline{n}}$ ，怎么办呢？容易；我们只要增添一些负号且取得

$$x^n = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (-1)^{n-k} x^{\bar{k}}, \text{ 整数 } n \geq 0, \quad (6.12)$$

$$x^{\bar{n}} = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (-1)^{n-k} x^k, \text{ 整数 } n \geq 0, \quad (6.13)$$

这样就行得通。因为，如公式

$$x^4 = x(x-1)(x-2)(x-3) = x^4 - 6x^3 + 11x^2 - 6x,$$

就像公式

$$x^{\bar{4}} = x(x+1)(x+2)(x+3) = x^4 + 6x^3 + 11x^2 + 6x,$$

但具有交错的正负号，如果 x 取为负，习题 2.17 的一般等式

$$x^{\bar{n}} = (-1)^n (-x)^{\bar{n}}$$

把式(6.10)转变成式(6.12)，把式(6.11)转变成式(6.13)。

表 6.3 基本的 Stirling 数等式，整数 $n \geq 0$

递归:

$$\begin{aligned} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} &= k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} + \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\} \\ \left[\begin{matrix} n \\ k \end{matrix} \right] &= (n-1) \left[\begin{matrix} n-1 \\ k \end{matrix} \right] + \left[\begin{matrix} n-1 \\ k-1 \end{matrix} \right] \end{aligned}$$

特殊值:

$$\begin{aligned} \left\{ \begin{matrix} n \\ 0 \end{matrix} \right\} &= \left[\begin{matrix} n \\ 0 \end{matrix} \right] = [n=0] \\ \left\{ \begin{matrix} n \\ 1 \end{matrix} \right\} &= [n>0]; \quad \left[\begin{matrix} n \\ 1 \end{matrix} \right] = (n-1)! [n>0] \\ \left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} &= (2^{n-1} - 1) [n>0]; \quad \left[\begin{matrix} n \\ 2 \end{matrix} \right] = (n-1)! H_{n-1} [n>0] \\ \left\{ \begin{matrix} n \\ n-1 \end{matrix} \right\} &= \left[\begin{matrix} n \\ n-1 \end{matrix} \right] = \binom{n}{2} \\ \left\{ \begin{matrix} n \\ n \end{matrix} \right\} &= \left[\begin{matrix} n \\ n \end{matrix} \right] = \binom{n}{n} = 1 \\ \left\{ \begin{matrix} n \\ k \end{matrix} \right\} &= \left[\begin{matrix} n \\ k \end{matrix} \right] = \binom{n}{k} = 0, \text{ 若 } k > n \end{aligned}$$

幂之间的转换:

$$x^n = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^{\bar{k}} = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (-1)^{n-k} x^{\bar{k}}$$

$$x^n = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} (-1)^{n-k} x^k$$

$$x^n = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} x^k$$

反演公式:

$$\sum_k \begin{bmatrix} n \\ k \end{bmatrix} \begin{Bmatrix} k \\ m \end{Bmatrix} (-1)^{n-k} = [m=n]$$

$$\sum_k \begin{Bmatrix} n \\ k \end{Bmatrix} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^{n-k} = [m=n]$$

表 6.4 另外的 Stirling 数等式, 整数 $l, m, n \geq 0$

$$\begin{Bmatrix} n+1 \\ m+1 \end{Bmatrix} = \sum_k \begin{pmatrix} n \\ k \end{pmatrix} \begin{Bmatrix} k \\ m \end{Bmatrix} \quad (6.15)$$

$$\begin{bmatrix} n+1 \\ m+1 \end{bmatrix} = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} \begin{pmatrix} k \\ m \end{pmatrix} \quad (6.16)$$

$$\begin{Bmatrix} n \\ m \end{Bmatrix} = \sum_k \begin{pmatrix} n \\ k \end{pmatrix} \begin{Bmatrix} k+1 \\ m+1 \end{Bmatrix} (-1)^{n-k} \quad (6.17)$$

$$\begin{bmatrix} n \\ m \end{bmatrix} = \sum_k \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \begin{pmatrix} k \\ m \end{pmatrix} (-1)^{m-k} \quad (6.18)$$

$$m! \begin{Bmatrix} n \\ m \end{Bmatrix} = \sum_k \begin{pmatrix} m \\ k \end{pmatrix} k^n (-1)^{m-k} \quad (6.19)$$

$$\begin{Bmatrix} n+1 \\ m+1 \end{Bmatrix} = \sum_{k=0}^n \begin{Bmatrix} k \\ m \end{Bmatrix} (m+1)^{n-k} \quad (6.20)$$

$$\begin{bmatrix} n+1 \\ m+1 \end{bmatrix} = \sum_{k=0}^n \begin{bmatrix} k \\ m \end{bmatrix} n^{n-k} = n! \sum_{k=0}^n \begin{bmatrix} k \\ m \end{bmatrix} / k! \quad (6.21)$$

$$\begin{Bmatrix} m+n+1 \\ m \end{Bmatrix} = \sum_{k=0}^m k \begin{Bmatrix} n+k \\ k \end{Bmatrix} \quad (6.22)$$

$$\begin{bmatrix} m+n+1 \\ m \end{bmatrix} = \sum_{k=0}^m (n+k) \begin{bmatrix} n+k \\ k \end{bmatrix} \quad (6.23)$$

$$\begin{pmatrix} n \\ m \end{pmatrix} = \sum_k \begin{Bmatrix} n+1 \\ k+1 \end{Bmatrix} \begin{bmatrix} k \\ m \end{bmatrix} (-1)^{m-k} \quad (6.24)$$

$$(n-m)! \begin{pmatrix} n \\ m \end{pmatrix} [n \geq m] = \sum_k \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \begin{Bmatrix} k \\ m \end{Bmatrix} (-1)^{m-k} \quad (6.25)$$

$$\begin{Bmatrix} n \\ n-m \end{Bmatrix} = \sum_k \begin{pmatrix} m-n \\ m+k \end{pmatrix} \begin{pmatrix} m+n \\ n+k \end{pmatrix} \begin{bmatrix} m+k \\ k \end{bmatrix} \quad (6.26)$$

$$\begin{bmatrix} n \\ n-m \end{bmatrix} = \sum_k \begin{pmatrix} m-n \\ m+k \end{pmatrix} \begin{pmatrix} m+n \\ n+k \end{pmatrix} \begin{Bmatrix} m+k \\ k \end{Bmatrix} \quad (6.27)$$

$$\left\{ \begin{matrix} n \\ l+m \end{matrix} \right\} \binom{l+m}{l} = \sum_k \left\{ \begin{matrix} k \\ l \end{matrix} \right\} \left\{ \begin{matrix} n-k \\ m \end{matrix} \right\} \binom{n}{k} \quad (6.28)$$

$$\left[\begin{matrix} n \\ l+m \end{matrix} \right] \binom{l+m}{l} = \sum_k \left[\begin{matrix} k \\ l \end{matrix} \right] \left[\begin{matrix} n-k \\ m \end{matrix} \right] \binom{n}{k} \quad (6.29)$$

我们能记住, 什么时候把 $(-1)^{n-k}$ 因子插入像式(6.12)那样的一个公式, 因为当 x 大时, 有幂的一个自然次序:

$$x^n > x^n > x^{\frac{n}{2}}, \text{ 对所有 } x > n > 1. \quad (6.30)$$

Stirling 数 $\left[\begin{matrix} n \\ k \end{matrix} \right]$ 和 $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ 是非负的, 所以当依据“大”幂来展开一个“小”幂时, 必需用负号,

我们能把式(6.11)代入式(6.12), 且得到一个双重和:

$$x^n = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (-1)^{n-k} x^k = \sum_{k,m} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \left[\begin{matrix} k \\ m \end{matrix} \right] (-1)^{n-k} x^m.$$

对所有 x , 此式成立, 所以右边的 $x^0, x^1, \dots, x^{n-1}, x^{n+1}, x^{n+2}, \dots$ 的系数一定全为零, 且一定有等式

$$\sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \left[\begin{matrix} k \\ m \end{matrix} \right] (-1)^{n-k} = [m=n], \text{ 整数 } m, n \geq 0. \quad (6.31)$$

像二项系数那样, Stirling 数满足许多意外的等式。但是这些等式不像第五章中的等式那样通用, 所以它们不常用。我们仅列出最简单的一些等式, 以备将来需要解 Stirling 难题时引用。表 6.3 和表 6.4 包含了最常用的一些公式, 我们已推出的最重要的等式也放在表中。

当我们在第五章中研究二项系数时, 发现, 等式 $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ 成立无任何限制, 以此方式对负的 n 定义 $\binom{n}{k}$ 是有利的。用此等式推广 $\binom{n}{k}$, 不会有组合的意义, 我们(在表 5.2 中)发现, 当向上扩展它时, Pascal 三角形实质上以替换的形式再造自身。让我们对 Stirling 三角形试验相同的事情: 如果我们确定对于所有整数 n 和 k , 基本递归

$$\begin{aligned} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} &= k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} + \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\} \\ \left[\begin{matrix} n \\ k \end{matrix} \right] &= (n-1) \left[\begin{matrix} n-1 \\ k \end{matrix} \right] + \left[\begin{matrix} n-1 \\ k-1 \end{matrix} \right] \end{aligned}$$

成立, 出现什么情况呢? 如果作合理的附加规定

$$\left\{ \begin{matrix} 0 \\ k \end{matrix} \right\} = \left[\begin{matrix} 0 \\ k \end{matrix} \right] = [k=0] \text{ 和 } \left\{ \begin{matrix} n \\ 0 \end{matrix} \right\} = \left[\begin{matrix} n \\ 0 \end{matrix} \right] = [n=0], \quad (6.32)$$

则解变成唯一。

表 6.5 一前一后的 Stirling 三角形

n	$\left\{ \begin{matrix} n \\ -5 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ -4 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ -3 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ -2 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ -1 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ 0 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ 1 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ 3 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ 4 \end{matrix} \right\}$	$\left\{ \begin{matrix} n \\ 5 \end{matrix} \right\}$
-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

事实上，一个意外好的型式出现：轮换的 Stirling 三角形显示出上面的子集 Stirling 三角形，且反过来也对！通过极简单的定律：

$$\left[\begin{matrix} n \\ k \end{matrix} \right] = \left\{ \begin{matrix} -k \\ -n \end{matrix} \right\}, \text{ 整数 } k, n, \quad (6.33)$$

把两类 Stirling 数联系起来。我们得到“对偶性”，有点像 \min 和 \max , $\lfloor x \rfloor$ 和 $\lceil x \rceil$, $x^{\underline{n}}$ 和 $x^{\bar{n}}$, \gcd 和 lcm 之间的关系。容易验证在此对应之下，递归 $\left[\begin{matrix} n \\ k \end{matrix} \right] = (n-1) \left[\begin{matrix} n-1 \\ k \end{matrix} \right] + \left[\begin{matrix} n-1 \\ k-1 \end{matrix} \right]$ 和 $\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} + \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\}$ 两者结果相同。

6.2 欧拉数

现在再提出另一个有价值的三角形，这归功于 Euler^[88,p.485]，我们用 $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle$ 记它的元素。这种情形的角括号使人联想到“小于”和“大于”符号； $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle$ 是有 n 次升高的 $\{1, 2, \dots, n\}$ 的排列 $\pi_1 \pi_2 \dots \pi_n$ 的个数，即有 k 处 $\pi_j < \pi_{j+1}$ 。（注意：这种记法比 Stirling 数的记法 $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$, $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 更加不标准。但是我们将看到它有好的意义。）

例如， $\{1, 2, 3, 4\}$ 的 11 个排列有两个升高：

1324, 1423, 2314, 2413, 3412;

1243, 1342, 2341, 2134, 3124, 4123.

(第一行列出具有 $\pi_1 < \pi_2 > \pi_3 < \pi_4$ 的排列, 第二行列出具有 $\pi_1 < \pi_2 < \pi_3 > \pi_4$ 和 $\pi_1 > \pi_2 < \pi_3 < \pi_4$ 的排列。)因此 $\langle \frac{4}{2} \rangle = 11$.

表 6.6 欧拉三角形

n	$\langle \frac{n}{0} \rangle$	$\langle \frac{n}{1} \rangle$	$\langle \frac{n}{2} \rangle$	$\langle \frac{n}{3} \rangle$	$\langle \frac{n}{4} \rangle$	$\langle \frac{n}{5} \rangle$	$\langle \frac{n}{6} \rangle$	$\langle \frac{n}{7} \rangle$	$\langle \frac{n}{8} \rangle$	$\langle \frac{n}{9} \rangle$
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

表 6.6 列出最小的欧拉数; 注意到, 此次标志序列为 1, 11, 11, 1. 当 $n > 0$ 时至多能有 $n-1$ 次升高, 所以在三角形的对角线上我们有 $\langle \frac{n}{n} \rangle = [n=0]$.

像 Pascal 三角形那样, 欧拉三角形在左右之间是对称的. 但是此时对称律稍有不同:

$$\langle \frac{n}{k} \rangle = \langle \frac{n}{n-1-k} \rangle, \text{ 整数 } n > 0, \quad (6.34)$$

排列 $\pi_1 \pi_2 \cdots \pi_n$ 有 $n-1-k$ 次升高当且仅当它的“反射” $\pi_n \cdots \pi_2 \pi_1$ 有 k 次升高.

让我们试求 $\langle \frac{n}{k} \rangle$ 的一个递归. 如果以所有可能的方式插入新元素 n , 则 $\{1, \dots, n-1\}$ 的每个排列 $\rho = \rho_1 \cdots \rho_{n-1}$ 引出 $\{1, 2, \dots, n\}$ 的 n 个排列. 假设把 n 放在位置 j , 则获得排列 $\pi = \rho_1 \cdots \rho_{j-1} n \rho_j \cdots \rho_{n-1}$. 如果 $j=1$ 或 $\rho_{j-1} < \rho_j$, 则 π 中的升高数和 ρ 中的升高数相同; 如果 $\rho_{j-1} > \rho_j$ 或 $j=n$, 则它比 ρ 中的升高数多 1 个. 所以来自有 k 个升高的排列 ρ 总共有 $(k+1)\langle \frac{n-1}{k} \rangle$ 种方式 π 有 k 个升高, 加上来自具有 $k-1$ 次升高的排列 ρ 总共有 $((n-2)-(k-1)+1)\langle \frac{n-1}{k-1} \rangle$ 种方式 π 有 k 次升高. 希望的递归是

$$\langle \frac{n}{k} \rangle = (k+1)\langle \frac{n-1}{k} \rangle + (n-k)\langle \frac{n-1}{k-1} \rangle, \text{ 整数 } n > 0. \quad (6.35)$$

我们再一次通过置

$$\langle \begin{smallmatrix} 0 \\ k \end{smallmatrix} \rangle = [k = 0], \text{ 整数 } k, \quad (6.36)$$

来起始递归, 当 $k < 0$ 时我们将假设 $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle = 0$.

欧拉数是有用的, 因为它们主要提供了通常幂和相继的二项系数之间的一个独特的联系:

$$x^n = \sum_k \langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle \binom{x+k}{n}, \text{ 整数 } n \geq 0. \quad (6.37)$$

(这是“Worpitzky 等式”^[308].) 例如, 我们有

$$\begin{aligned} x^2 &= \binom{x}{2} + \binom{x+1}{2}, \\ x^3 &= \binom{x}{3} + 4\binom{x+1}{3} + \binom{x+2}{3}, \\ x^4 &= \binom{x}{4} + 11\binom{x+1}{4} + 11\binom{x+2}{4} + \binom{x+3}{4}, \end{aligned}$$

等等. 用归纳法易证式(6.37)(习题 14).

顺便提一句, 式(6.37)还给出另一种方式来获得前 n 个平方的和: 我们有 $k^2 = \langle \begin{smallmatrix} 2 \\ 0 \end{smallmatrix} \rangle$

$$\langle \begin{smallmatrix} k \\ 2 \end{smallmatrix} \rangle + \langle \begin{smallmatrix} 2 \\ 1 \end{smallmatrix} \rangle \langle \begin{smallmatrix} k+1 \\ 2 \end{smallmatrix} \rangle = \binom{k}{2} + \binom{k+1}{2}, \text{ 因此}$$

$$\begin{aligned} 1^2 + 2^2 + \cdots + n^2 &= \left(\binom{1}{2} + \binom{2}{2} + \cdots + \binom{n}{2} \right) + \left(\binom{2}{2} + \binom{3}{2} + \cdots + \binom{n+1}{2} \right) \\ &= \binom{n+1}{3} + \binom{n+2}{3} = \frac{1}{6}(n+1)n((n-1) + (n+2)). \end{aligned}$$

欧拉递归(6.35)比 Stirling 递归(6.3)和(6.8)更复杂一些, 所以我们不期望数 $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle$ 满足许多简单的等式. 但仍然有几个简单的等式:

$$\langle \begin{smallmatrix} n \\ m \end{smallmatrix} \rangle = \sum_{k=0}^m \binom{n+1}{k} (m+1-k)^n (-1)^k; \quad (6.38)$$

$$m! \left\{ \begin{smallmatrix} n \\ m \end{smallmatrix} \right\} = \sum_k \langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle \binom{k}{n-m}; \quad (6.39)$$

$$\langle \begin{smallmatrix} n \\ m \end{smallmatrix} \rangle = \sum_k \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} \binom{n-k}{m} (-1)^{n-k-m} k!. \quad (6.40)$$

如果用 z^{n-m} 乘式(6.39)且对 m 求和, 我们取得 $\sum_m \left\{ \begin{smallmatrix} n \\ m \end{smallmatrix} \right\} m! z^{n-m} = \sum_k \langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle (z+1)^k$. 用 $z=1$

代替 z 且使 z^k 的系数相等给出式(6.40). 因此, 最后两个等式本质上是等价的, 第一个等式(6.38)当 m 小时给出特殊的值:

$$\langle \frac{n}{0} \rangle = 1; \langle \frac{n}{1} \rangle = 2^n - n - 1; \langle \frac{n}{2} \rangle = 3^n - (n+1)2^n + \binom{n+1}{2}.$$

表 6.7 二阶欧拉三角形

n	$\langle \frac{n}{0} \rangle$	$\langle \frac{n}{1} \rangle$	$\langle \frac{n}{2} \rangle$	$\langle \frac{n}{3} \rangle$	$\langle \frac{n}{4} \rangle$	$\langle \frac{n}{5} \rangle$	$\langle \frac{n}{6} \rangle$	$\langle \frac{n}{7} \rangle$	$\langle \frac{n}{8} \rangle$
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

这里我们不需详细研究欧拉数; 通常简单地知道它们存在就够了, 当需要产生时可求助于一组基本等式. 然而, 在我们结束此课题之前, 还应注意到表 6.7 中表明的另一个三角形型式的系数. 我们称这些系数为“二阶欧拉数” $\langle \frac{n}{k} \rangle$, 因为它们满足相似于式(6.35)的一个递归, 但是在一处用 $2n-1$ 代替 n :

$$\langle \frac{n}{k} \rangle = (k+1) \langle \frac{n-1}{k} \rangle + (2n-1-k) \langle \frac{n-1}{k-1} \rangle. \quad (6.41)$$

这些数有一种好奇的组合解释, 由 Gessel 和 Stanley^[118]首先提到: 如果我们形成具有特殊性质的多重集 $\{1, 1, 2, 2, \dots, n, n\}$ 的排列, 性质是对于 $1 \leq m \leq n$ 两次出现 m 之间的所有数大于 m , 则 $\langle \frac{n}{k} \rangle$ 是这种具有 k 次升高的排列数. 例如, 有 8 个适合的单个升高的 $\{1, 1, 2, 2, 3, 3\}$ 的排列:

113322, 133221, 221331, 221133, 223311, 233211, 331122, 331221.

因此 $\langle \frac{3}{1} \rangle = 8$. 多重集 $\{1, 1, 2, 2, \dots, n, n\}$ 总共有

$$\sum_k \langle \frac{n}{k} \rangle = (2n-1)(2n-3)\cdots(1) = \frac{(2n)!}{2^n} \quad (6.42)$$

个适合的排列, 因为 n 的两次出现一定是相邻的, 且 $n-1$ 的一个排列中间有 $2n-1$ 个位置插入它们。例如, 当 $n=3$, 排列 1221 有 5 个插入处, 产生 331221, 133221, 123321, 122331 和 122133。推广对于通常欧拉数的论证能证明递归式(6.41)。

二阶欧拉数是重要的, 主要由于它们和 Stirling 数的联系^[19]。对 n 归纳, 我们得到

$$\left\{ \begin{matrix} x \\ x-n \end{matrix} \right\} = \sum_k \langle \langle n \rangle \rangle_k \binom{x+n-1-k}{2n}, \text{ 整数 } n \geq 0; \quad (6.43)$$

$$\left[\begin{matrix} x \\ x-n \end{matrix} \right] = \sum_k \langle \langle n \rangle \rangle_k \binom{x+k}{2n}, \text{ 整数 } n \geq 0. \quad (6.44)$$

例如,

$$\begin{aligned} \left\{ \begin{matrix} x \\ x-1 \end{matrix} \right\} &= \binom{x}{2}, & \left[\begin{matrix} x \\ x-1 \end{matrix} \right] &= \binom{x}{2}; \\ \left\{ \begin{matrix} x \\ x-2 \end{matrix} \right\} &= \binom{x+1}{4} + 2\binom{x}{4}, & \left[\begin{matrix} x \\ x-2 \end{matrix} \right] &= \binom{x}{4} + 2\binom{x+1}{4}; \\ \left\{ \begin{matrix} x \\ x-3 \end{matrix} \right\} &= \binom{x+2}{6} + 8\binom{x+1}{6} + 6\binom{x}{6}, & \left[\begin{matrix} x \\ x-3 \end{matrix} \right] &= \binom{x}{6} + 8\binom{x+1}{6} + 6\binom{x+2}{6}. \end{aligned}$$

(在式(6.7)中已遇到情形 $n=1$ 。)每当 x 是整数和 n 是非负整数时这些等式成立。由于右边是 x 的多项式, 我们能用式(6.43)和(6.44), 对于任意实(或复)的 x 值定义 Stirling 数 $\left\{ \begin{matrix} x \\ x-n \end{matrix} \right\}$ 和 $\left[\begin{matrix} x \\ x-n \end{matrix} \right]$ 。

如果 $n > 0$, 当 $x=0, x=1, \dots$ 和 $x=n$ 时这些多项式 $\left\{ \begin{matrix} x \\ x-n \end{matrix} \right\}$ 和 $\left[\begin{matrix} x \\ x-n \end{matrix} \right]$ 是零; 所以它们可用 $(x-0), (x-1), \dots$ 和 $(x-n)$ 除尽。查看除这些已知因子外左边成为什么是有意义的。我们用规则

$$\sigma_n(x) = \frac{\left[\begin{matrix} x \\ x-n \end{matrix} \right]}{(x(x-1)\cdots(x-n))} \quad (6.45)$$

来定义 Stirling 多项式 $\sigma_n(x)$ 。($\sigma_n(x)$ 的次数为 $n-1$ 。)前几个情形为:

$$\sigma_0(x) = \frac{1}{x};$$

$$\sigma_1(x) = \frac{1}{2};$$

$$\sigma_2(x) = \frac{(3x-1)}{24};$$

$$\sigma_3(x) = \frac{(x^2-x)}{48};$$

$$\sigma_4(x) = \frac{(15x^3 - 30x^2 + 5x + 2)}{5760}.$$

通过二阶欧拉数能计算它们, 例如,

$$\sigma_3(x) = \frac{((x-4)(x-5) + 8(x+1)(x-4) + 6(x+2)(x+1))}{6!}.$$

表 6.8 Stirling 卷积公式

$$rs \sum_{k=0}^n \sigma_k(r) \sigma_{n-k}(s) = (r+s) \sigma_n(r+s) \quad (6.46)$$

$$s \sum_{k=0}^n k \sigma_k(r) \sigma_{n-k}(s) = n \sigma_n(r+s) \quad (6.47)$$

$$rs \sum_{k=0}^n \sigma_k(r+k) \sigma_{n-k}(s+n-k) = (r+s) \sigma_n(r+s+n) \quad (6.48)$$

$$s \sum_{k=0}^n k \sigma_k(r+k) \sigma_{n-k}(s+n-k) = n \sigma_n(r+s+n) \quad (6.49)$$

$$\left\{ \begin{matrix} n \\ m \end{matrix} \right\} = (-1)^{n-m+1} \frac{n!}{(m-1)!} \sigma_{n-m}(-m) \quad (6.50)$$

$$\left[\begin{matrix} n \\ m \end{matrix} \right] = \frac{n!}{(m-1)!} \sigma_{n-m}(n) \quad (6.51)$$

结果, 这些多项式满足两个十分巧妙的等式:

$$\left(\frac{ze^z}{e^z - 1} \right)^x = x \sum_{n \geq 0} \sigma_n(x) z^n; \quad (6.52)$$

$$\left(\frac{1}{z} \ln \frac{1}{1-z} \right)^x = x \sum_{n \geq 0} \sigma_n(x+n) z^n. \quad (6.53)$$

所以如同在表 5.5 中对二项系数所做的那样, 我们能获得 Stirling 数的一般卷积公式, 结果显示在表 6.8 中. 当 Stirling 数的一个和不适合表 6.3 或 6.4 的等式时, 表 6.8 就可能是合适的公式. (一个例子出现在本章后面, 如下面的方程(6.100). 习题 7.19 讨论了基于像式(6.52)和(6.53)等的卷积的一般原理.)

6.3 调和数

现在是仔细查看调和数的时候了, 我们首先在第二章后面遇到过它:

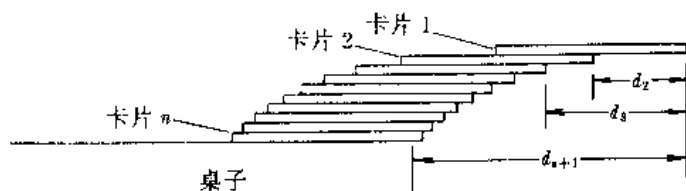
$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = \sum_{k=1}^n \frac{1}{k}, \text{ 整数 } n \geq 0. \quad (6.54)$$

在算法分析中常常出现这些数，因而计算机科学工作者需要它们的一种特殊记法。我们用 H_n ，‘H’表示“调和”，因为波长 $1/n$ 的一个音称为波长为 1 的一个音的第 n 个调和。前几个值为：

n	0	1	2	3	4	5	6	7	8	9	10
H_n	0	1	$\frac{3}{2}$	$\frac{11}{6}$	$\frac{25}{12}$	$\frac{137}{60}$	$\frac{49}{20}$	$\frac{363}{140}$	$\frac{761}{280}$	$\frac{7129}{2520}$	$\frac{7381}{2520}$

习题 21 指明当 $n > 1$ 时 H_n 不是一个整数。

这里是一种卡片诀窍，基于 R·T·Sharp^[264] 的思想，它说明了在简单场合中如何自然地产生调和数。给定 n 张卡片和一个桌子，我们希望通过把卡片堆放在桌边产生最大可能的伸出，服从重力定律。为了使确定的问题更多一点，要求卡片边和桌边平行，否则转动卡片以致它们的角更突出一点能使伸出增加。为了使解答简单，假设每张卡片是 2 个单位长。



关于 1 张卡片，当它的重力刚好在桌边上我们取为最大伸出。重心是在卡片中部，所以可伸出卡长的一半，或 1 个单位。

关于 2 张卡片，不难确信我们取得的最大伸出为上面的一张卡片的重心刚好在第 2 张卡片的边的上面，且 2 张卡片结合起来的重心就在桌边上面。2 张卡片的联合重心将在它们公共部分的中间，所以我们能完成另外的伸出的半个单位。

这种型式提供一个一般的方法，我们安置这些卡片以致上面 k 张卡片的重心就在第 $k+1$ 张卡片的上面(它支撑上面的 k 张卡片)。桌子起到第 $n+1$ 张卡片的作用。为了代数地表达这个条件，可设 d_k 为最上面的一张卡片的端边到从顶端开始第 k 张卡片的对应边的距离，于是 $d_1 = 0$ ，且我们要作出前 k 张卡片的重心 d_{k+1} ：

$$d_{k+1} = \frac{(d_1 + 1) + (d_2 + 1) + \cdots + (d_k + 1)}{k}, \quad 1 \leq k \leq n. \quad (6.55)$$

(重量分别为 w_1, \dots, w_k ，重心分别在位置 p_1, \dots, p_k 的 k 个元素的重心在位置 $(w_1 p_1 + \cdots + w_k p_k) / (w_1 + \cdots + w_k)$)。我们可把这个递归再写为二个等价形式

$$k d_{k+1} = k + d_1 + \cdots + d_{k-1} + d_k, \quad k \geq 0;$$

$$(k-1) d_k = k-1 + d_1 + \cdots + d_{k-1}, \quad k \geq 1.$$

两个方程相减得

$$kd_{k+1} - (k-1)d_k = 1 + d_k, \quad k \geq 1.$$

因此 $d_{k+1} = d_k + 1/k$. 第 2 张卡片将位移第 3 张二分之一一个单位, 过第 4 张三分之一一个单位, 等等. 根据归纳法一般公式

$$d_{k+1} = H_k, \quad (6.56)$$

如果置 $k=n$, 当如描述的那样堆放 n 张卡片时, 我们得到 $d_{n+1} = H_n$ 作为总伸出.

能否通过退缩, 不把每张卡片推进到极端位置而给后面前进储备“重力位能”, 以使们达到更大的伸出呢? 不, 任何完全平衡卡片的安置有

$$d_{k+1} \leq \frac{(1+d_1) + (1+d_2) + \cdots + (1+d_k)}{k}, \quad 1 \leq k \leq n.$$

此外 $d_1 = 0$, 由归纳法得到 $d_{k+1} \leq H_n$.

注意关于完全过桌边的顶上的 1 张卡片并不占用太多卡片. 我们需要多于 2 个单位的 1 张卡片长的一个伸出. 超过 2 的第 1 个调和数是 $H_4 = 25/12$, 所以仅需 4 张卡片.

对于 52 张卡片, 有 H_{52} 单位的伸出, 结果它是 $H_{52}/2 \approx 2.27$ 个卡片长. (我们将学到一个公式, 说明如何对大的 n 来计算 H_n 的一个近似值, 而不必把整串分数加起来.)

一个称为“在橡胶带上爬行”的有趣问题以另一种形式表明调和数. 一次慢而坚持不懈地爬行 W , 从 1m 长的橡胶带的一端开始, 朝另一端以每分钟 1 cm 爬行. 在每分钟末, 一个同样坚持不懈的带的持有者 K , 把带延伸 1 m, 他的唯一的目的是阻挠 W . 因此在爬行 1 min 后, W 离开始处为 1 cm 且离结束处为 99 cm; 然后把它延伸 1 m. 在延伸操作过程中 W 维持他的相对位置, 离开始处为 1% 且离结束处为 99%; 所以 W 现在离开始点为 2 cm 且离目标为 198 cm. 在 W 爬行另外 1 min 后, 行进的记录是 3 cm 且还有 197 cm 要走; 但是 K 延伸, 距离变成 4.5 和 295.5, 等等. 爬行能到达结束处吗? 他保持移动, 但是目标看来远离更快. (我们设想 K 和 W 有无限耐久性, 带的无限弹性, 无限微小的爬行.)

让我们写下一些公式. 当 K 延伸橡胶带时, W 爬行的带的一部分停留在同样的地方. 于是第一分钟他爬行带的 $1/100$, 第二分钟为 $1/200$, 第三分钟为 $1/300$, 等等. n min 后, 他爬行的带的一部分是

$$\frac{1}{100} \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \right) = \frac{H_n}{100}. \quad (6.57)$$

所以如果 H_n 在任何时候超过 100, 则他到达结束处.

我们将立即看到对于大的 n 如何估计 H_n . 现在让我们考虑在相同情况中“特大爬行”如何执行, 从而简单地检验我们的分析. 不同于 W , 特大爬行每分钟能爬行 50 cm, 所以按照刚才给的论证, 它在 n min 后将爬行带长的 $H_n/2$. 如果我们的推理正确, 则由于 $H_4 > 2$, 在 n 到达 4 之前特大爬行将结束. 是的, 简单的计算表明过 3 min 后特大爬行仅留下 $33(1/3)$ cm 要行进. 它正好用了 3 min 40 s 结束.

调和数也出现在 Stirling 三角形中. 让我们试求 $\left[\begin{smallmatrix} n \\ 2 \end{smallmatrix} \right]$ 的一个闭形式, 即恰有二个轮换的 n 个元素的排列数的一种闭形式. 递归(6.8)告诉我们

$$\left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] = n \left[\begin{smallmatrix} n \\ 2 \end{smallmatrix} \right] + \left[\begin{smallmatrix} n \\ 1 \end{smallmatrix} \right] = n \left[\begin{smallmatrix} n \\ 2 \end{smallmatrix} \right] + (n-1)!, \text{ 如果 } n > 0;$$

且此递归是第二章的求和因子技巧的一个自然的选择者

$$\frac{1}{n!} \left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] = \frac{1}{(n-1)!} \left[\begin{smallmatrix} n \\ 2 \end{smallmatrix} \right] + \frac{1}{n}.$$

展开此递归, $\frac{1}{n!} \left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] = H_n$; 因此

$$\left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] = n! H_n. \quad (6.58)$$

在第二章中我们证明调和级数 $\sum_k 1/k$ 发散, 这意味着当 $n \rightarrow \infty$ 时 H_n 取任意大. 但是证明是间接的; 我们找某一个无限和(2.58), 当它被重新排列时给出不同的解答, 因此 $\sum_k 1/k$ 不能是有界的. $H_n \rightarrow \infty$ 看来不直观, 因为它意指充分大的堆放卡片将伸出桌子 1 英里以上, 以及蠕行 W 将最终到达他的带的终端. 所以让我们仔细查看当 n 大时 H_n 的大小.

观察 $H_n \rightarrow \infty$ 的最简单的方法也许是按照 2 的幂来分组它的项. 我们把 1 项放入组 1, 2 项放入组 2, 4 项放入组 3, 8 项放入组 4, 等等:

$$\underbrace{\frac{1}{1}}_{\text{组1}} + \underbrace{\frac{1}{2}}_{\text{组2}} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\text{组3}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\text{组4}} + \frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12} + \frac{1}{13} + \frac{1}{14} + \frac{1}{15} + \cdots$$

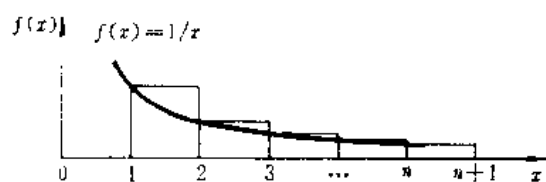
组 2 中的 2 项是在 $1/4$ 和 $1/2$ 之间, 所以此组的和是在 $2 \cdot 1/4 = 1/2$ 和 $2 \cdot 1/2 = 1$ 之间. 组 3 中的所有 4 项是在 $1/8$ 和 $1/4$ 之间, 所以它们的和也在 $1/2$ 和 1 之间. 事实上, 组 k 中 2^{k-1} 项的每一项在 2^{-k} 和 2^{1-k} 之间; 因此每个别组的和是在 $1/2$ 和 1 之间.

此分组过程告诉我们, 如果 n 是在组 k 中, 一定有 $H_n > k/2$ 和 $H_n \leq k$ (通过对 k 归纳). 于是 $H_n \rightarrow \infty$, 且事实上

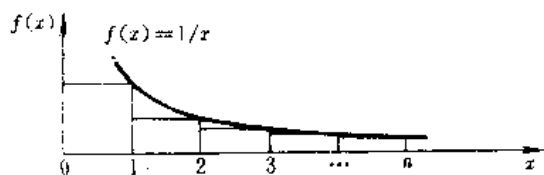
$$\frac{\lfloor \lg n \rfloor + 1}{2} < H_n \leq \lfloor \lg n \rfloor + 1. \quad (6.59)$$

我们现在知道 H_n 在 2 的一个因子内部. 虽然调和数逼近无穷, 但它们仅对数地逼近它——也就是说, 十分慢.

稍作一点工作和一番计算能求得较好的界限. 在第二章中学习了 H_n 是连续函数 $\ln n$ 的离散的相似函数. 自然对数定义为一条曲线下的面积, 所以提出了一种几何对照:



1 和 n 之间曲线下的面积 $\int_1^n dx/x = \ln n$ 小于 n 个矩形的面积 $\sum_{k=1}^n 1/k = H_n$, 因此 $\ln n < H_n$; 这比式(6.59)所有的结果更强。安置稍有不同矩形, 我们取得一个相似的上界:



此时 n 个矩形的面积 H_n 小于第一个矩形的面积加上曲线下的面积。我们证明了

$$\ln n < H_n < \ln n + 1, \quad n > 1. \quad (6.60)$$

我们现在知道 H_n 的值具有至多为 1 的误差。

当求倒数平方的和而不是简单地求倒数的和时, 产生“二阶”调和数 $H_n^{(2)}$:

$$H_n^{(2)} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2} = \sum_{k=1}^n \frac{1}{k^2}.$$

相似, 通过求 $(-r)$ 次幂的和定义 r 阶调和数:

$$H_n^{(r)} = \sum_{k=1}^n \frac{1}{k^r}. \quad (6.61)$$

如果 $r > 1$, 当 $n \rightarrow \infty$ 时这些数趋向一个极限; 在习题 2.31 中我们注意到此极限通常称为黎曼 ζ 函数:

$$\zeta(r) = H_\infty^{(r)} = \sum_{k=1}^{\infty} \frac{1}{k^r}. \quad (6.62)$$

欧拉发现了一种简洁的方式, 用广义的调和来逼近通常的 $H_n^{(1)}$. 让我们考虑无穷级数

$$\ln\left(\frac{k}{k-1}\right) = \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{3k^3} + \frac{1}{4k^4} + \cdots, \quad (6.63)$$

当 $k > 1$ 时它收敛。左边为 $\ln k - \ln(k-1)$; 所以如果对 $2 \leq k \leq n$ 求两边的和, 左边和缩短且我们取得

$$\begin{aligned}\ln n - \ln 1 &= \sum_{k=2}^n \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{3k^3} + \frac{1}{4k^4} + \cdots \right) \\ &= (H_n - 1) + \frac{1}{2} (H_n^{(2)} - 1) + \frac{1}{3} (H_n^{(3)} - 1) + \frac{1}{4} (H_n^{(4)} - 1) + \cdots.\end{aligned}$$

重新整理, 我们得到 H_n 和 $\ln n$ 之间的差的一个表达式:

$$H_n - \ln n = 1 - \frac{1}{2} (H_n^{(2)} - 1) - \frac{1}{3} (H_n^{(3)} - 1) - \frac{1}{4} (H_n^{(4)} - 1) - \cdots.$$

当 $n \rightarrow \infty$ 时, 右边趋向极限值

$$1 - \frac{1}{2} (\zeta(2) - 1) - \frac{1}{3} (\zeta(3) - 1) - \frac{1}{4} (\zeta(4) - 1) - \cdots,$$

它现在被认为是欧拉常数, 通常用希腊字母 γ 来记. 事实上, $\zeta(r) - 1$ 接近 $1/2^r$, 所以这个无穷级数收敛得相当快且我们能计算十进制值

$$\gamma = 0.577\,215\,664\,9\cdots. \quad (6.64)$$

欧拉的论证建立极限关系

$$\lim_{n \rightarrow \infty} (H_n - \ln n) = \gamma; \quad (6.65)$$

于是位于大约式(6.60)中的两端之间范围的 58%. 我们逐步引导到它的值.

如同将在第九章中见到的那样, 进一步改进是可能的. 例如, 我们将证明

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{\varepsilon_n}{120n^4}, \quad 0 < \varepsilon_n < 1. \quad (6.66)$$

此公式让我们推得第一百万个调和数是

$$H_{1\,000\,000} \approx 14.392\,726\,722\,865\,723\,631\,381\,127\,5,$$

不必把一百万个分数相加. 除了别的以外, 这意味着堆起一百万张卡片能伸出桌边超过 7 张卡片长.

关于在橡胶带上蠕行, 式(6.66)告诉我们什么? 由于 H_n 是无界的, 当 H_n 首次超过 100 时, 蠕行将确实达到终点. 对于 H_n 我们的近似表明当 n 接近

$$e^{100-\gamma} \approx e^{99.423}$$

时这将发生. 事实上, 习题 9.49 证明 n 的临界值或者是 $\lfloor e^{100-\gamma} \rfloor$ 或者是 $\lceil e^{100-\gamma} \rceil$. 我们能想象 W 成功, 当他最终越过终点线, K 十分懊恼, 长期的爬行开始之后需经历大约 $287 \times (1\,000)^{11}$ 世纪. (橡胶带将延伸多于 10^{27} 光年长, 它的微小颗粒将相当远离.)

6.4 调和的求和

现在让我们从回顾在第二章中学过的几个概念开始, 查看涉及调和数的一些和. 在式(2.36)和(2.57)中我们证明

$$\sum_{0 \leq k < n} H_k = nH_n - n; \quad (6.67)$$

$$\sum_{0 \leq k < n} kH_k = \frac{n(n-1)}{2} H_n - \frac{n(n-1)}{4}. \quad (6.68)$$

让我们大胆取一个更一般的和, 它包含这两者作为特殊情形. 当 m 是非负整数时,

$$\sum_{0 \leq k < n} \binom{k}{m} H_k$$

的值是什么?

在第二章中, 最适合于处理式(6.67)和(6.68)的方法称为分部求和法. 我们把被加数写成形式 $u(k)\Delta v(k)$, 且应用一般等式

$$\sum_a^b u(x)\Delta v(x)\delta x = u(x)v(x)\Big|_a^b - \sum_a^b v(x+1)\Delta u(x)\delta x. \quad (6.69)$$

记住了吗? 现在我们面对的和 $\sum_{0 \leq k < n} \binom{k}{m} H_k$ 就此法来说是自然的, 因为我们能让

$$\begin{aligned} u(k) &= H_k, & \Delta u(k) &= H_{k+1} - H_k = \frac{1}{k+1}; \\ v(k) &= \binom{k}{m+1}, & \Delta v(k) &= \binom{k+1}{m+1} - \binom{k}{m+1} = \binom{k}{m}. \end{aligned}$$

(换句话说, 调和数有一个简单的 Δ 且二项系数有一个简单的 Δ^{-1} , 所以我们正可用分部求和的方法。)代入式(6.69)产生

$$\begin{aligned} \sum_{0 \leq k < n} \binom{k}{m} H_k &= \sum_0^n \binom{x}{m} H_x \delta x = \binom{x}{m+1} H_x \Big|_0^n - \sum_0^n \binom{x+1}{m+1} \frac{\delta x}{x+1} \\ &= \binom{n}{m+1} H_n - \sum_{0 \leq k < n} \binom{k+1}{m+1} \frac{1}{k+1}. \end{aligned}$$

剩下的和是容易的, 因为我们能用备用的方程(5.5)来并入 $(k+1)^{-1}$:

$$\sum_{0 \leq k < n} \binom{k+1}{m+1} \frac{1}{k+1} = \sum_{0 \leq k < n} \binom{k}{m} \frac{1}{m+1} = \binom{n}{m+1} \frac{1}{m+1}.$$

因此得到我们寻找的解答:

$$\sum_{0 \leq k \leq n} \binom{k}{m} H_k = \binom{n}{m+1} \left(H_n - \frac{1}{m+1} \right). \quad (6.70)$$

(当 $m=0$ 和 $m=1$ 时, 用式(6.67)和(6.68)检验很好。)

下一个和的例子用除代替乘: 让我们试计算

$$S_n = \sum_{k=1}^n \frac{H_k}{k}.$$

如果用它的定义展开, 得双重和,

$$S_n = \sum_{1 \leq j \leq k \leq n} \frac{1}{j \cdot k}.$$

现在第二章的另一个方法来作为我们的助手, 方程(2.33)告诉我们

$$S_n = \frac{1}{2} \left(\left(\sum_{k=1}^n \frac{1}{k} \right)^2 + \sum_{k=1}^n \frac{1}{k^2} \right) = \frac{1}{2} (H_n^2 + H_n^{(2)}). \quad (6.71)$$

结果, 如果试用分部求和, 也能以另一种方式获得此解答(见习题 26)。

现在让我们尝试处理较困难的问题[291], 它并不服从分部求和法:

$$U_n = \sum_{k \geq 1} \binom{n}{k} \frac{(-1)^{k-1}}{k} (n-k)^n, \text{ 整数 } n \geq 1.$$

(此和也不明显提到调和数, 但是谁知道什么时候它们可出现?)

我们将以两种方式解此问题, 一种通过费事地求出解答, 另一种通过巧妙的和/或侥幸的方法。首先, 费事的方法。我们用二项定理展开 $(n-k)^n$, 以致分母中麻烦的 k 将和分子结合起来:

$$\begin{aligned} U_n &= \sum_{k \geq 1} \binom{n}{k} \frac{(-1)^{k-1}}{k} \sum_j \binom{n}{j} (-k)^j n^{n-j} \\ &= \sum_j \binom{n}{j} (-1)^{j-1} n^{n-j} \sum_{k \geq 1} \binom{n}{k} (-1)^k k^{j-1}. \end{aligned}$$

看来这不完全是凌乱的, 因为内和中的 k^{j-1} 是 k 的多项式, 且等式(5.40)告诉我们, 我们简单地取此多项式的第 n 次差分。首先一定要清理几件事情。一件事情是, 如果 $j=0$, k^{j-1} 不是一个多项式, 我们将需要分裂项且分开处理它。另一件事情是, 第 n 次差分的公式中省去了项 $k=0$, 当 $j=1$ 时此项为非零, 所以我们需较好地重建它(且再减去它)。结果是

$$U_n = \sum_{j \geq 1} \binom{n}{j} (-1)^{j-1} n^{n-j} \sum_{k \geq 0} \binom{n}{k} (-1)^k k^{j-1}$$

$$\begin{aligned}
&= \sum_{j=1}^n \binom{n}{j} (-1)^{j-1} n^{n-j} \binom{n}{0} 0^{j-1} \\
&= \binom{n}{0} n^n \sum_{k=1}^n \binom{n}{k} (-1)^k k^{-1}.
\end{aligned}$$

现在最高行(仅剩的双重和)为零: 它为次数小于 n 的多项式的第 n 次差分倍数之和, 且这样的第 n 次差分为零。除 $j=1$ 外第二行为零, 此时它等于 $-n^n$ 。所以第三项是仅剩下的困难; 我们把原来的问题化成一个十分简单的和:

$$U_n = n^n (T_n - 1), \text{ 其中 } T_n = \sum_{k=1}^n \binom{n}{k} \frac{(-1)^{k-1}}{k}. \quad (6.72)$$

例如, $U_3 = \binom{3}{1} \frac{8}{1} - \binom{3}{2} \frac{1}{2} = \frac{45}{2}$; $T_3 = \binom{3}{1} \frac{1}{1} - \binom{3}{2} \frac{1}{2} + \binom{3}{3} \frac{1}{3} = \frac{11}{6}$; 因此 $U_3 = 27(T_3 - 1)$, 如同要求的那样。

我们如何能计算 T_n 呢? 一种方式是用 $\binom{n-1}{k} + \binom{n-1}{k-1}$ 来代替 $\binom{n}{k}$, 获得依据 T_{n-1} 的 T_n 的一个简单递归。但是有一个更启发的方法: 在式(5.41)中有一个相似的公式, 即

$$\sum_k \binom{n}{k} \frac{(-1)^k}{x+k} = \frac{n!}{x(x+1)\cdots(x+n)}.$$

如果去掉 $k=0$ 的项且置 $x=0$, 我们得到负 T_n 。所以让我们处理它:

$$\begin{aligned}
T_n &= \left(\frac{1}{x} - \frac{n!}{x(x+1)\cdots(x+n)} \right) \Big|_{x=0} \\
&= \left(\frac{(x+1)\cdots(x+n) - n!}{x(x+1)\cdots(x+n)} \right) \Big|_{x=0} \\
&= \left(\frac{x^n \left[\begin{smallmatrix} n+1 \\ n+1 \end{smallmatrix} \right] + \cdots + x \left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] + \left[\begin{smallmatrix} n+1 \\ 1 \end{smallmatrix} \right] - n!}{x(x+1)\cdots(x+n)} \right) \Big|_{x=0} = \frac{1}{n!} \left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right].
\end{aligned}$$

(我们用 $(x+1)\cdots(x+n) = x^{\overline{n+1}}/x$ 的展开式 (6.11); 我们能除尽从分子中出来的 x , 因为 $\left[\begin{smallmatrix} n+1 \\ 1 \end{smallmatrix} \right] = n!$ 。)但是从式 (6.58) 中知道 $\left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] = n!H_n$, 因此 $T_n = H_n$, 我们得到解答:

$$U_n = n^n (H_n - 1). \quad (6.73)$$

这是一种方式。另一种方式将试计算很一般的和

$$U_n(x, y) = \sum_{k \geq 1} \binom{n}{k} \frac{(-1)^{k-1}}{k} (x + ky)^n, \text{ 整数 } n \geq 0, \quad (6.74)$$

原来的 U_n 的值将作为特殊情形 $U_n(n, -1)$ 的值。(但愿我们尝试更多的一般性, 因为前面的推导“放过了”给定问题的大部细节(第 n 次差分把它们除去了), 故这些细节一定是不相干的。)

作一点小的变化, 我们能再用前面的推导且发现 $U_n(x, y)$ 的值, 或者能用 $(x + ky)^{n-1}(x + ky)$ 代替 $(x + ky)^n$, 然后用 $\binom{n-1}{k} + \binom{n-1}{k-1}$ 代替 $\binom{n}{k}$, 导出递归

$$U_n(x, y) = xU_{n-1}(x, y) + \frac{x^n}{n} + yx^{n-1}; \quad (6.75)$$

这能用一个求和因子容易地解出(习题 5)。

但是用另一种诀窍是最容易的, 它使用了第二章中的有利条件: 微分。 $U_n(x, y)$ 对 y 的求导产生一个 k , 它和分母中的 k 相消, 且产生的和是平凡的:

$$\begin{aligned} \frac{\partial}{\partial y} U_n(x, y) &= \sum_{k \geq 1} \binom{n}{k} (-1)^{k-1} n(x + ky)^{n-1} \\ &= \binom{n}{0} nx^{n-1} - \sum_{k \geq 0} \binom{n}{k} (-1)^k n(x + ky)^{n-1} = nx^{n-1}. \end{aligned}$$

(再一次, 次数 $< n$ 的一个多项式的第 n 次差分成为零。)

我们证明了 $U_n(x, y)$ 对 y 的求导为 nx^{n-1} , 与 y 无关。一般, 如果 $f'(y) = c$ 则 $f(y) = f(0) + cy$; 所以我们一定得到 $U_n(x, y) = U_n(x, 0) + nx^{n-1}y$ 。

余下的工作是确定 $U_n(x, 0)$ 。 $U_n(x, 0)$ 就是 x^n 乘上式(6.72)中我们已研究过的和 $T_n = H_n$, 所以式(6.74)中的一般和有闭形式

$$U_n(x, y) = x^n H_n + nx^{n-1}y. \quad (6.76)$$

特别, 原问题的解为 $U_n(n, -1) = n^n(H_n - 1)$ 。

6.5 伯努利数

我们讨论的下一个重要的数序列取名为伯努利数, 以纪念 Jakob Bernoulli(1654 ~ 1705), 他在作出 m 次幂的和的公式时发现了好奇的关系^[22]。让我们记

$$S_m(n) = 0^m + 1^m + \cdots + (n-1)^m = \sum_{k=0}^{n-1} k^m = \sum_0^n x^m \delta x. \quad (6.77)$$

(于是当 $m > 0$ 时, 在广义调和数的表示法中我们有 $S_m(n) = H_{n-1}^{(-m)}$.) 伯努利查看了下列公式序列且发现一种型式:

$$S_0(n) = n$$

$$S_1(n) = \frac{1}{2}n^2 - \frac{1}{2}n$$

$$S_2(n) = \frac{1}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{6}n$$

$$S_3(n) = \frac{1}{4}n^4 - \frac{1}{2}n^3 + \frac{1}{4}n^2$$

$$S_4(n) = \frac{1}{5}n^5 - \frac{1}{2}n^4 + \frac{1}{3}n^3 - \frac{1}{30}n$$

$$S_5(n) = \frac{1}{6}n^6 - \frac{1}{2}n^5 + \frac{5}{12}n^4 - \frac{1}{12}n^2$$

$$S_6(n) = \frac{1}{7}n^7 - \frac{1}{2}n^6 + \frac{1}{2}n^5 - \frac{1}{6}n^3 + \frac{1}{42}n$$

$$S_7(n) = \frac{1}{8}n^8 - \frac{1}{2}n^7 + \frac{7}{12}n^6 - \frac{7}{24}n^4 + \frac{1}{12}n^2$$

$$S_8(n) = \frac{1}{9}n^9 - \frac{1}{2}n^8 + \frac{2}{3}n^7 - \frac{7}{15}n^5 + \frac{2}{9}n^3 - \frac{1}{30}n$$

$$S_9(n) = \frac{1}{10}n^{10} - \frac{1}{2}n^9 + \frac{3}{4}n^8 - \frac{7}{10}n^6 + \frac{1}{2}n^4 - \frac{3}{20}n^2$$

$$S_{10}(n) = \frac{1}{11}n^{11} - \frac{1}{2}n^{10} + \frac{5}{6}n^9 - n^7 + n^5 - \frac{1}{2}n^3 + \frac{5}{66}n$$

你还能明白它吗? $S_m(n)$ 中 n^{m+1} 的系数总是 $1/(m+1)$, n^m 的系数总是 $-1/2$, n^{m-1} 的系数总是 $m/12$, n^{m-2} 的系数总是零, n^{m-3} 的系数总是 $-m(m-1)(m-2)/720$, n^{m-4} 的系数总是零. 看来好像型式将延伸, n^{m-k} 的系数总是某个常数乘 $m^{\frac{k}{2}}$.

这是伯努利的发现, 在近代的记法中把系数写成形式

$$\begin{aligned} S_m(n) &= \frac{1}{m+1} \left(B_0 n^{m+1} + \binom{m+1}{1} B_1 n^m + \cdots + \binom{m+1}{m} B_m n \right) \\ &= \frac{1}{m+1} \sum_{k=0}^m \binom{m+1}{k} B_k n^{m+1-k}. \end{aligned} \quad (6.78)$$

由一个隐递归关系定义伯努利数,

$$\sum_{j=0}^m \binom{m+1}{j} B_j = [m=0], \text{ 对于所有 } m \geq 0. \quad (6.79)$$

例如, $\binom{2}{0} B_0 + \binom{2}{1} B_1 = 0$. 前面几个值为

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	$-\frac{691}{2730}$

(通过显示奇怪的分数 $-691/2730$ 消除了关于一种 B_n 的简单闭形式的所有猜测。)

通过对 m 归纳, 用摄动法(在第二章中求 $S_2(n) = \square_n$ 的方式之一)能证明伯努利公式(6.78):

$$\begin{aligned} S_{m+1}(n) + n^{m+1} &= \sum_{k=0}^{n-1} (k+1)^{m+1} \\ &= \sum_{k=0}^{n-1} \sum_{j=0}^{m+1} \binom{m+1}{j} k^j = \sum_{j=0}^{m+1} \binom{m+1}{j} S_j(n). \end{aligned} \quad (6.80)$$

设 $\hat{S}_m(n)$ 为式(6.78)的右边; 我们希望指出 $S_m(n) = \hat{S}_m(n)$, 假设 $S_j(n) = \hat{S}_j(n)$ ($0 \leq j < m$). 如同在第二章中对 $m=2$ 所做的那样, 我们开始从式(6.80)的两边减去 $S_{m+1}(n)$, 然后用式(6.78)展开每个 $S_j(n)$, 再重新组合以使右边 n 的幂的系数放在一起且简化为:

$$\begin{aligned} n^{m+1} &= \sum_{j=0}^m \binom{m+1}{j} S_j(n) = \sum_{j=0}^m \binom{m+1}{j} \hat{S}_j(n) + \binom{m+1}{m} \Delta \\ &= \sum_{j=0}^m \binom{m+1}{j} \frac{1}{j+1} \sum_{k=0}^j \binom{j+1}{k} B_k n^{j+1-k} + (m+1)\Delta \\ &= \sum_{0 \leq k \leq j \leq m} \binom{m+1}{j} \binom{j+1}{k} \frac{B_k}{j+1} n^{j+1-k} + (m+1)\Delta \\ &= \sum_{0 \leq k \leq j \leq m} \binom{m+1}{j} \binom{j+1}{j-k} \frac{B_{j-k}}{j+1} n^{k+1} + (m+1)\Delta \\ &= \sum_{0 \leq k \leq j \leq m} \binom{m+1}{j} \binom{j+1}{k+1} \frac{B_{j-k}}{j+1} n^{k+1} + (m+1)\Delta \\ &= \sum_{0 \leq k \leq m} \frac{n^{k+1}}{k+1} \sum_{k \leq j \leq m} \binom{m+1}{j} \binom{j}{k} B_{j-k} + (m+1)\Delta \\ &= \sum_{0 \leq k \leq m} \frac{n^{k+1}}{k+1} \binom{m+1}{k} \sum_{k \leq j \leq m} \binom{m+1-k}{j-k} B_{j-k} + (m+1)\Delta \\ &= \sum_{0 \leq k \leq m} \frac{n^{k+1}}{k+1} \binom{m+1}{k} \sum_{0 \leq j \leq m-k} \binom{m+1-k}{j} B_j + (m+1)\Delta \\ &= \sum_{0 \leq k \leq m} \frac{n^{k+1}}{k+1} \binom{m+1}{k} [m-k=0] + (m+1)\Delta \\ &= \frac{n^{m+1}}{m+1} \binom{m+1}{m} + (m+1)\Delta \end{aligned}$$

$$= n^{m+1} + (m+1)\Delta, \text{ 其中 } \Delta = S_m(n) - \hat{S}_m(n).$$

(这个推导是在第五章中所学的标准操作的一种好的复习。)因此 $\Delta = 0$, $S_m(n) = \hat{S}_m(n)$, 证毕。

在第七章中, 我们将用母函数得到式(6.78)的一种十分简单的证明。关键的思想将是指出伯努利数是幂级数

$$\frac{z}{e^z - 1} = \sum_{n \geq 0} B_n \frac{z^n}{n!} \quad (6.81)$$

的系数。现在让我们简单地设方程(6.81)成立, 这样我们能推出它的一些惊奇的结果。如果把 $(1/2)z$ 加到式(6.81)两边, 从而从右边消去项 $B_1 z/1! = -(1/2)z$, 我们取得

$$\frac{z}{e^z - 1} + \frac{z}{2} = \frac{z}{2} \frac{e^z + 1}{e^z - 1} = \frac{z}{2} \frac{e^{z/2} + e^{-z/2}}{e^{z/2} - e^{-z/2}} = \frac{z}{2} \coth \frac{z}{2}. \quad (6.82)$$

这里的 \coth 是“双曲余切”函数, 在微积分书中另外被认为 $\cosh z / \sinh z$; 我们有

$$\sinh z = \frac{e^z - e^{-z}}{2}, \quad \cosh z = \frac{e^z + e^{-z}}{2}. \quad (6.83)$$

把 z 改成 $-z$ 给出 $(-z/2)\coth(-z/2) = (z/2)\coth(z/2)$; 因此每个奇数编号的 $(z/2)\coth(z/2)$ 的系数一定为零, 且我们有

$$B_3 = B_5 = B_7 = B_9 = B_{11} = B_{13} = \cdots = 0. \quad (6.84)$$

此外式(6.82)导致 \coth 的系数的一个闭形式:

$$z \coth z = \frac{2z}{e^{2z} - 1} + \frac{2z}{2} = \sum_{n \geq 0} B_{2n} \frac{(2z)^{2n}}{(2n)!} = \sum_{n \geq 0} 4^n B_{2n} \frac{z^{2n}}{(2n)!}. \quad (6.85)$$

但是并不十分需要双曲函数, 人们对三角学的“实”函数较感兴趣。我们能由规则

$$\sin z = -i \sinh iz, \quad \cos z = \cosh iz, \quad (6.86)$$

依据双曲函数表达通常的三角函数。对应的幂级数为

$$\begin{aligned} \sin z &= \frac{z^1}{1!} - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots, & \sinh z &= \frac{z^1}{1!} + \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots; \\ \cos z &= \frac{z^0}{0!} - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots, & \cosh z &= \frac{z^0}{0!} + \frac{z^2}{2!} + \frac{z^4}{4!} + \cdots. \end{aligned}$$

因此 $\cot z = \cos z / \sin z = i \cosh iz / \sinh iz = i \coth iz$, 且有

$$z \cot z = \sum_{n \geq 0} B_{2n} \frac{(2iz)^{2n}}{(2n)!} = \sum_{n \geq 0} (-4)^n B_{2n} \frac{z^{2n}}{(2n)!} \quad (6.87)$$

欧拉找到了另一个值得注意的 $z \cot z$ 的公式(习题 73):

$$z \cot z = 1 - 2 \sum_{k \geq 1} \frac{z^2}{k^2 \pi^2 - z^2} \quad (6.88)$$

我们能把欧拉公式展成 z^2 的幂, 获得

$$\begin{aligned} z \cot z &= 1 - 2 \sum_{k \geq 1} \left(\frac{z^2}{k^2 \pi^2} + \frac{z^4}{k^4 \pi^4} + \frac{z^6}{k^6 \pi^6} + \cdots \right) \\ &= 1 - 2 \left(\frac{z^2}{\pi^2} H_{\infty}^{(2)} + \frac{z^4}{\pi^4} H_{\infty}^{(4)} + \frac{z^6}{\pi^6} H_{\infty}^{(6)} + \cdots \right) \end{aligned}$$

在以前的式(6.87)中, 使 z^{2n} 的系数相等, 给出一个几乎非凡的无限多无限和的闭形式:

$$\zeta(2n) = H_{\infty}^{(2n)} = (-1)^{n-1} \frac{2^{2n-1} \pi^{2n} B_{2n}}{(2n)!}, \text{ 整数 } n > 0 \quad (6.89)$$

例如,

$$\zeta(2) = H_{\infty}^{(2)} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots = \pi^2 B_2 = \frac{\pi^2}{6}; \quad (6.90)$$

$$\zeta(4) = H_{\infty}^{(4)} = 1 + \frac{1}{16} + \frac{1}{81} + \cdots = -\frac{\pi^4 B_4}{3} = \frac{\pi^4}{90}; \quad (6.91)$$

公式(6.89)不仅是 $H_{\infty}^{(2n)}$ 的一个闭形式, 它也告诉我们 B_{2n} 的近似大小, 由于当 n 大时 $H_{\infty}^{(2n)}$ 十分接近 1. 它还告诉我们对所有 $n > 0$, $(-1)^{n-1} B_{2n} > 0$, 于是非零伯努利数交错正负号.

此外伯努利数还出现在正切函数的系数中,

$$\tan z = \frac{\sin z}{\cos z} = \sum_{n \geq 0} (-1)^{n-1} 4^n (4^n - 1) B_{2n} \frac{z^{2n-1}}{(2n)!}, \quad (6.92)$$

以及其他三角函数中(习题 70). 公式(6.92)导致另一个伯努利数的重要事实, 即,

$$T_{2n-1} = (-1)^{n-1} \frac{4^n (4^n - 1)}{2n} B_{2n} \text{ 是正整数.} \quad (6.93)$$

例如, 我们有

n	1	3	5	7	9	11	13
T_n	1	2	16	272	7 936	353 792	22 368 256

(称 T 为正切数。)

根据 B · F · Logan 的思想, 证明式(6.93)的一种方法是考虑幂级数

$$\begin{aligned}\frac{\sin z + x \cos z}{\cos z - x \sin z} &= x + (1 + x^2)z + (2x^3 + 2x)\frac{z^2}{2} + (6x^4 + 8x^2 + 2)\frac{z^3}{6} + \cdots \\ &= \sum_{n \geq 0} T_n(x) \frac{z^n}{n!},\end{aligned}\quad (6.94)$$

其中 $T_n(x)$ 是 x 的多项式; 置 $x=0$ 给出 $T_n(0) = T_n$, 第 n 个正切数。如果把式(6.94)对 x 微分, 我们得到

$$\frac{1}{(\cos z - x \sin z)^2} = \sum_{n \geq 0} T'_n(x) \frac{z^n}{n!};$$

但是如果对 z 微分, 我们得到

$$\frac{1+x^2}{(\cos z - x \sin z)^2} = \sum_{n \geq 1} T_n(x) \frac{z^{n-1}}{(n-1)!} = \sum_{n \geq 0} T_{n+1}(x) \frac{z^n}{n!}.$$

(尝试它——相消是十分巧妙的。)所以我们有

$$T_{n+1}(x) = (1+x^2)T'_n(x), \quad T_0(x) = x, \quad (6.95)$$

一个简单的递归, 从它推得 $T_n(x)$ 的系数为非负整数。此外, 容易证明 $T_n(x)$ 次数为 $n+1$, 且它的系数交错地为零和正的。所以如同式(6.93)中所要求的那样, $T_{2n+1}(0) = T_{2n+1}$ 是正整数。

递归(6.95)给出一种简单方法来计算伯努利数, 经由正切数, 仅用整数的简单运算; 对比之下, 定义的递归(6.79)涉及困难的具有分数的算术运算。

如果我们要计算从 a 到 $b-1$ 的 n 次幂的和, 而不是从 0 到 $n-1$, 第二章的理论告诉我们

$$\sum_{k=a}^{b-1} k^m = \sum_a^b x^m \delta x = S_m(b) - S_m(a). \quad (6.96)$$

当考虑 k 的负值时此等式有有趣的结果:

$$\sum_{k=-n+1}^{-1} k^m = (-1)^m \sum_{k=0}^{n-1} k^m, \quad \text{当 } m > 0,$$

因此

$$S_m(0) - S_m(-n+1) = (-1)^m (S_m(n) - S_m(0)).$$

但是 $S_m(0) = 0$, 所以我们有等式

$$S_m(1-n) = (-1)^{m+1} S_m(n), \quad m > 0. \quad (6.97)$$

所以 $S_m(1) = 0$. 如果以因子形式写多项式 $S_m(n)$, 它总将有因子 n 和 $(n-1)$, 因为它有根 0 和 1. 一般, $S_m(n)$ 是次数 $m+1$ 的多项式, 首项为 $(1/m+1)n^{m+1}$. 而且, 我们能在式 (6.97) 中置 $n = 1/2$ 取得 $S_m(1/2) = (-1)^{m+1} S_m(1/2)$; 如果 m 是偶数, 这使 $S_m(1/2) = 0$, 所以 $(n - (1/2))$ 将是一个附加因子. 这些观察说明为什么在第二章中我们找到简单的因子分解

$$S_2(n) = \frac{1}{3} n \left(n - \frac{1}{2} \right) (n-1),$$

我们能用这样的理由来导出 $S_2(n)$ 的值, 而不计算它! 此外, 式 (6.97) 意指具有剩下因子的多项式, $\hat{S}_m(n) = S_m(n) / (n - (1/2))$, 总满足

$$\hat{S}_m(1-n) = \hat{S}_m(n), \quad n \text{ 偶数}, \quad m > 0.$$

由此推知 $S_m(n)$ 总能记为因子形式

$$S_m(n) = \begin{cases} \frac{1}{m+1} \prod_{k=1}^{[m/2]} \left(n - \frac{1}{2} - \alpha_k \right) \left(n - \frac{1}{2} + \alpha_k \right), & m \text{ 奇数}; \\ \left(n - \frac{1}{2} \right) \frac{1}{m+1} \prod_{k=1}^{m/2} \left(n - \frac{1}{2} - \alpha_k \right) \left(n - \frac{1}{2} + \alpha_k \right), & m \text{ 偶数}. \end{cases} \quad (6.98)$$

这里 $\alpha_1 = 1/2$, $\alpha_2, \dots, \alpha_{[m/2]}$ 是适当的复数, 它的值依赖于 m . 例如,

$$S_3(n) = \frac{n^2(n-1)^2}{4};$$

$$S_4(n) = \frac{n \left(n - \frac{1}{2} \right) (n-1) \left(n - \frac{1}{2} + \sqrt{\frac{7}{12}} \right) \left(n - \frac{1}{2} - \sqrt{\frac{7}{12}} \right)}{5};$$

$$S_5(n) = \frac{n^2(n-1)^2 \left(n - \frac{1}{2} + \sqrt{\frac{3}{4}} \right) \left(n - \frac{1}{2} - \sqrt{\frac{3}{4}} \right)}{6};$$

$$S_6(n) = n \left(n - \frac{1}{2} \right) (n-1) \left(n - \frac{1}{2} + \alpha \right) \left(n - \frac{1}{2} - \alpha \right) \left(n - \frac{1}{2} + \bar{\alpha} \right) \left(n - \frac{1}{2} - \bar{\alpha} \right),$$

$$\text{其中 } \alpha = 2^{-5/2} 3^{-1/2} 31^{1/4} (\sqrt{\sqrt{31} + \sqrt{27}} + i\sqrt{\sqrt{31} - \sqrt{27}}).$$

如果 m 是奇数且大于 1, 我们有 $B_m = 0$; 因此 n^2 可除尽 $S_m(n)$ (以及 $(n-1)^2$ 可除尽 $S_m(n)$). 否则 $S_m(n)$ 的根看来不服从简单的定律.

让我们查看它们如何与 Stirling 数相联系来结束伯努利数的研究。计算 $S_m(n)$ 的一种方法是把原来的幂改变成下降幂, 因为下降幂有容易的和。在处理容易的和之后我们能转回到原来的幂:

$$\begin{aligned} S_m(n) &= \sum_{k=0}^{n-1} k^m = \sum_{k=0}^{n-1} \sum_{j \geq 0} \left\{ \begin{matrix} m \\ j \end{matrix} \right\} k^{\underline{j}} = \sum_{j \geq 0} \left\{ \begin{matrix} m \\ j \end{matrix} \right\} \sum_{k=0}^{n-1} k^{\underline{j}} \\ &= \sum_{j \geq 0} \left\{ \begin{matrix} m \\ j \end{matrix} \right\} \frac{n^{\underline{j+1}}}{j+1} \\ &= \sum_{j \geq 0} \left\{ \begin{matrix} m \\ j \end{matrix} \right\} \frac{1}{j+1} \sum_{k \geq 0} (-1)^{j+1-k} \left[\begin{matrix} j+1 \\ k \end{matrix} \right] n^k. \end{aligned}$$

所以, 使和(6.78)中那些系数相等, 我们一定有等式

$$\sum_{j \geq 0} \left\{ \begin{matrix} m \\ j \end{matrix} \right\} \left[\begin{matrix} j+1 \\ k \end{matrix} \right] \frac{(-1)^{j+1-k}}{j+1} = \frac{1}{m+1} \binom{m+1}{k} B_{m+1-k}. \quad (6.99)$$

直接证明此关系将是好的, 从而以一种新方法来说伯努利数。但是表 6.3 或 6.4 中的等式不给出任何归纳法证明的明显处理, 即式(6.99)中左边的和是一个常数乘 $m^{\frac{k-1}{2}}$ 。若 $k=m+1$, 左边和就是 $\left\{ \begin{matrix} m \\ m \end{matrix} \right\} \left[\begin{matrix} m+1 \\ m+1 \end{matrix} \right] / (m+1) = 1 / (m+1)$, 所以这是容易的情形。如果 $k=m$, 左边共计为 $\left\{ \begin{matrix} m \\ m-1 \end{matrix} \right\} \left[\begin{matrix} m \\ m \end{matrix} \right] m^{-1} - \left\{ \begin{matrix} m \\ m \end{matrix} \right\} \left[\begin{matrix} m+1 \\ m \end{matrix} \right] (m+1)^{-1} = (1/2)(m-1) - (1/2)m = -1/2$, 所以这也是相当容易的情形。但是如果 $k < m$, 左边和看来令人不快。如果伯努利采取这条路线, 也许不能发现他的数。

我们能做的一件事情是用 $\left\{ \begin{matrix} m+1 \\ j+1 \end{matrix} \right\} - (j+1) \left\{ \begin{matrix} m \\ j+1 \end{matrix} \right\}$ 代替 $\left\{ \begin{matrix} m \\ j \end{matrix} \right\}$ 。(j+1)恰好消去了难处理的分子, 左边变成

$$\sum_{j \geq 0} \left\{ \begin{matrix} m+1 \\ j+1 \end{matrix} \right\} \left[\begin{matrix} j+1 \\ k \end{matrix} \right] \frac{(-1)^{j+1-k}}{j+1} - \sum_{j \geq 0} \left\{ \begin{matrix} m \\ j+1 \end{matrix} \right\} \left[\begin{matrix} j+1 \\ k \end{matrix} \right] (-1)^{j+1-k}.$$

当 $k < m$ 时, 由式(6.31), 第二个和为零。这就留下第一个和, 它要求记法上的改变; 让我们再命名所有变量使求和的指数为 k , 且使其他参数为 m 和 n 。于是等式(6.99)等价于

$$\sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \left[\begin{matrix} k \\ m \end{matrix} \right] \frac{(-1)^{k-m}}{k} = \frac{1}{n} \binom{n}{m} B_{n-m} + [m = n-1]. \quad (6.100)$$

虽然表 6.4 仍未提供任何明显的下一步, 看来我们具有较合意的事情。

现在利用表 6.8 中的卷积公式。我们能用式(6.51)和(6.50)并依据 Stirling 多项式改写被加数:

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} \left[\begin{matrix} k \\ m \end{matrix} \right] = (-1)^{n-k+1} \frac{n!}{(k-1)!} \sigma_{n-k}(-k) \cdot \frac{k!}{(m-1)!} \sigma_{k-m}(k);$$

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} \left[\begin{matrix} k \\ m \end{matrix} \right] \frac{(-1)^{k-m}}{k} = (-1)^{n+1-m} \frac{n!}{(m-1)!} \sigma_{n-k}(-k) \sigma_{k-m}(k).$$

看来事情是好的；式(6.48)中的卷积产生

$$\begin{aligned} \sum_{k=0}^n \sigma_{n-k}(-k) \sigma_{k-m}(k) &= \sum_{k=0}^{n-m} \sigma_{n-m-k}(-n + (n-m-k)) \sigma_k(m+k) \\ &= \frac{m-n}{(m)(-n)} \sigma_{n-m}(m-n + (n-m)). \end{aligned}$$

现在验证了公式(6.100)，我们发现伯努利数与 Stirling 多项式中的常数项相联系：

$$(-1)^{m-1} m \sigma_m(0) = \frac{B_m}{m!} + [m=1]. \quad (6.101)$$

6.6 Fibonacci 数

现在我们来到也许是最合意的特殊数序列，Fibonacci 序列 $\langle F_n \rangle$ ：

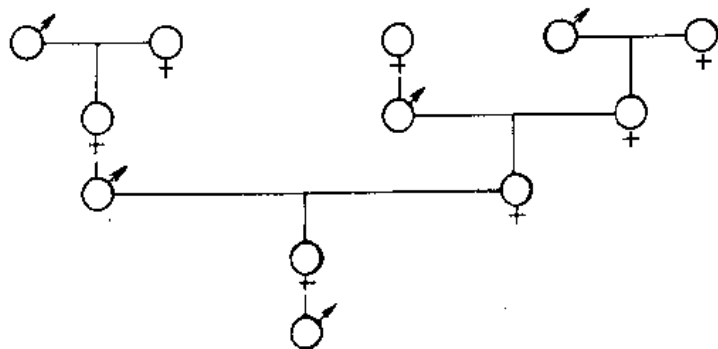
n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
F_n	0	1	1	2	3	5	8	13	21	34	55	89	144	233	377

不像调和数和伯努利数，Fibonacci 数是合宜的简单整数。由递归

$$\begin{aligned} F_0 &= 0; \\ F_1 &= 1; \\ F_n &= F_{n-1} + F_{n-2}, \quad n > 1, \end{aligned} \quad (6.102)$$

定义它们。这个简明的规则，每个数依赖于前面两个数的最简单的可能的递归解决了在各种场合出现的 Fibonacci 数。

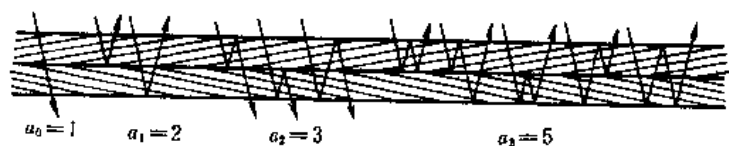
“蜂树”提供了如何自然地产生 Fibonacci 数的一个好例子。让我们考虑一个公蜂的家谱。每个公蜂(也称为雄蜂)由一个雌蜂(也称为一个女王)无性地生殖；然而，每个雌蜂有双亲，一个公的和一个雌的。这里是树的前几级：



雄蜂有一个祖父和一个祖母，它有一个曾祖父和两个曾祖母，他有两个曾曾祖父和三个曾曾祖母。一般，由归纳容易看出它恰好有 F_{n+1} 个曾...曾祖父和 F_{n+2} 个曾...曾祖母。

自然界常常出现 Fibonacci 数，也许是相似于蜂树定律的原因。例如，典型的向日葵有一个包含螺旋形紧靠着的大量小花的大头，通常一个方向有 34 圈且另一方向有 55 圈，较小的头将有 21 和 34，或 13 和 21 圈；在英国有一次曾展出一个具有 89 和 144 个螺旋形的巨大的向日葵。在某些种类的松果中发现相似的型式。

这里是一个不同性质的例^[219]：假设我们一前一后地放两片玻璃，光线在改变方向 n 次后或通过或反射，有多少种方式 a_n ？前几种情形是：



当 n 是偶数时，有偶数次弹起且射线通过；当 n 是奇数时，射线被反射且在它进入的相同侧再冒出。 a_n 看来是 Fibonacci 数，稍凝视一下图就告诉我们为什么是 Fibonacci 数：对 $n \geq 2$ ， n 次弹起或者它们第一次弹起止于相对面且以 a_{n-1} 种方式继续，或者它们由弹起止于中间面开始，然后以 a_{n-2} 种方式再弹跳回来结束。因此我们得到 Fibonacci 递归 $a_n = a_{n-1} + a_{n-2}$ 。起始条件不同，但相差不太大，因为我们有 $a_0 = 1 = F_2$ 和 $a_1 = 2 = F_3$ ；所以仅简单地移动两个位置，且 $a_n = F_{n+2}$ 。

Leonardo Fibonacci 在 1202 年引入了这些数，数学家们逐渐发现了它们的许多有趣的结果。Edouard Lucas，第一章中讨论的汉诺塔游戏的制作者，在 19 世纪后半期广泛地研究了它们(事实上，是 Lucas 普及了名称“Fibonacci 数”)。他的令人惊异的结果之一是用 Fibonacci 数的性质来证明 39 位的 Mersenne 数 $2^{127} - 1$ 是素数。

1680 年由法国天文学家 Jean-Dominique Cassini^[45]提出的 Fibonacci 数的最老定理之一是等式

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n, \quad n > 0. \quad (6.103)$$

当 $n=6$ 时，例如，Cassini 等式正确地断言 $13 \cdot 5 - 8^2 = 1$ 。

对于小的 k 值涉及形式 $F_{n \pm k}$ 的 Fibonacci 数的多项式公式能变换成仅涉及 F_n 和 F_{n+1} 的公式，因为当 $m < n$ 时我们能用规则

$$F_m = F_{m+2} - F_{m+1} \quad (6.104)$$

依据较高的 Fibonacci 数来表达 F_m ，当 $m > n+1$ 时，我们能

$$F_m = F_{m-2} + F_{m-1} \quad (6.105)$$

以较低的 Fibonacci 数来代替 F_m 。于是，例如，在式(6.103)中我们能用 $F_{n+1} - F_n$ 来代替

F_{n-1} 取得形式为

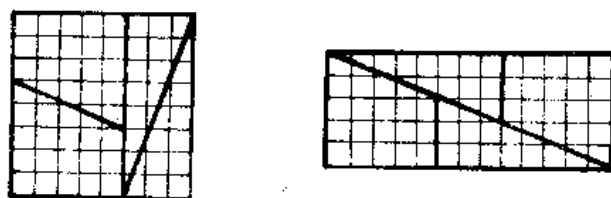
$$F_{n+1}^2 - F_{n+1}F_n - F_n^2 = (-1)^n \quad (6.106)$$

的 Cassini 等式。而且，当用 $n+1$ 代替 n 时，Cassini 等式成为

$$F_{n+2}F_n - F_{n+1}^2 = (-1)^{n+1},$$

这与 $(F_{n+1} + F_n)F_n - F_{n+1}^2 = (-1)^{n+1}$ 相同，与式(6.106)相同。因此 Cassini(n)是真当且仅当 Cassini($n+1$)是真，通过归纳法可证对所有 n 方程(6.103)成立。

Cassini 等式是一个几何谬论的基础，它是 Lewis Carroll 喜欢的游戏之一^{[54][258][298]}。想法是取一个棋盘，且把它分成 4 片，如同这里表明的那样，把这 4 片再装配成一个矩形：



变戏法似的：原来 $8 \times 8 = 64$ 个方块，的面积再排列产生 $5 \times 13 = 65$ 个方块！相似的构造把任何 $F_n \times F_n$ 个方块分成 4 片，图例分别为 13, 8, 5 和 3 之处用 F_{n+1} , F_n , F_{n-1} 和 F_{n-2} 作为尺寸。其结果为一个 $F_{n-1} \times F_{n+1}$ 矩形；依据式(6.103)，增加一个方块或减少一个方块，依赖于 n 是偶数还是奇数。

严格地说，除非 $m \geq 2$ 不能应用分解式(6.105)，因为对于负的 n 没有定义 F_n 。如果消去这个界限条件且用式(6.104)和(6.105)来定义负指数的 Fibonacci 数，则一大堆操作变得较容易。例如， F_{-1} 是 $F_1 - F_0 = 1$ ，于是 F_{-2} 是 $F_0 - F_{-1} = -1$ 。以此方式我们引出值

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

且(由归纳法)它很快清楚地变成

$$F_{-n} = (-1)^{n-1} F_n, \text{ 整数 } n. \quad (6.107)$$

当我们以此方式推广 Fibonacci 序列时，不仅对 $n > 0$ ，而对所有整数 n ，Cassini 等式(6.103)是真的。

用式(6.105)和(6.104)把 F_{n+k} 化成 F_n 和 F_{n+1} 的一种结合的过程导致公式序列

$$\begin{aligned} F_{n+2} &= F_{n+1} + F_n & F_{n-1} &= F_{n+1} - F_n \\ F_{n+3} &= 2F_{n+1} + F_n & F_{n-2} &= -F_{n+1} + 2F_n \\ F_{n+4} &= 3F_{n+1} + 2F_n & F_{n-3} &= 2F_{n+1} - 3F_n \end{aligned}$$

$$F_{n+5} = 5F_{n+1} + 3F_n \quad F_{n-4} = -3F_{n+1} + 5F_n$$

在这些公式中间另一种型式变得明显:

$$F_{n+k} = F_k F_{n+1} + F_{k-1} F_n. \quad (6.108)$$

用归纳法容易证明, 此等式对所有整数 k 和 n (正, 负或零) 成立.

如果我们在式(6.108)中置 $k=n$, 求得

$$F_{2n} = F_n F_{n+1} + F_{n-1} F_n, \quad (6.109)$$

因此 F_{2n} 是 F_n 的一个倍数. 相似,

$$F_{3n} = F_{2n} F_{n+1} + F_{2n-1} F_n,$$

并可推得 F_{3n} 也是 F_n 的一个倍数. 由归纳法证明对于所有整数 k 和 n ,

$$F_{kn} \text{ 为 } F_n \text{ 的一个倍数.} \quad (6.110)$$

例如, 这说明了为什么 F_{15} (它等于 610) 为 F_3 和 F_5 (等于 2 和 5) 两者的一个倍数. 事实上, 还有更多等式是真的, 习题 27 证明

$$\gcd(F_m, F_n) = F_{\gcd(m, n)}. \quad (6.111)$$

例如, $\gcd(F_{12}, F_{18}) = \gcd(144, 2584) = 8 = F_6$.

现在我们能证明式(6.110)的逆: 如果 $n > 2$ 且如果 F_m 是 F_n 的一个倍数, 则 m 是 n 的一个倍数. 如果 $F_n \nmid F_m$, 则 $F_n \nmid \gcd(F_m, F_n) = F_{\gcd(m, n)} \leq F_n$. 仅当 $F_{\gcd(m, n)} = F_n$ 时这是可能的, 我们的假设 $n > 2$ 使它必须有 $\gcd(m, n) = n$, 因此 $n \mid m$.

Yuri Matijasevich 在他的著名的证明^[213]中用可除性概念的一种推广, 即没有算法决定是否一个给定的整系数的多变量多项式方程有一个整数解. Matijasevich 的引理说明如果 $n > 2$, Fibonacci 数 F_m 是 F_n^2 的一个倍数当且仅当 m 是 nF_n 的一个倍数.

让我们通过查看 $k=1, 2, 3, \dots$ 的序列 $\langle F_{kn} \bmod F_n^2 \rangle$ 来证明这一点, 并注意什么时候 $F_{kn} \bmod F_n^2 = 0$. (我们知道 m 一定有形式 kn 如果 $F_m \bmod F_n = 0$) 首先我们有 $F_n \bmod F_n^2 = F_n$, 这不是零. 接着由式(6.108)我们得到

$$F_{2n} = F_n F_{n+1} + F_{n-1} F_n \equiv 2F_n F_{n+1} \pmod{F_n^2},$$

因为 $F_{n+1} \equiv F_{n-1} \pmod{F_n}$. 相似,

$$F_{2n+1} = F_{n+1}^2 + F_n^2 \equiv F_{n+1}^2 \pmod{F_n^2}.$$

这个同余允许我们计算

$$\begin{aligned}
F_{3n} &= F_{2n+1}F_n + F_{2n}F_{n+1} \\
&\equiv F_{n+1}^2F_n + 2(F_nF_{n+1})F_{n+1} = 3F_{n+1}^2F_n \pmod{F_n^2}; \\
F_{3n+1} &= F_{2n+1}F_{n+1} + F_{2n}F_n \\
&\equiv F_{n+1}^3 + (2F_nF_{n+1})F_n \equiv F_{n+1}^3 \pmod{F_n^2}.
\end{aligned}$$

一般, 我们对 k 归纳求得

$$F_{kn} \equiv kF_nF_{n+1}^{k-1} \text{ 和 } F_{kn+1} \equiv F_{n+1}^k \pmod{F_n^2}.$$

现在 F_{n+1} 和 F_n 互素, 所以

$$\begin{aligned}
F_{kn} \equiv 0 \pmod{F_n^2} &\Leftrightarrow kF_n \equiv 0 \pmod{F_n^2} \\
&\Leftrightarrow k \equiv 0 \pmod{F_n}.
\end{aligned}$$

我们证明了 Matijasevich 引理。

Fibonacci 数的最重要性质之一是以特殊方式把它们用来表示整数。让我们记下

$$j \gg k \Leftrightarrow j \geq k + 2. \quad (6.112)$$

于是每一个正整数有唯一的一种表示形式

$$n = F_{k_1} + F_{k_2} + \cdots + F_{k_r}, \quad k_1 \gg k_2 \gg \cdots \gg k_r \gg 0. \quad (6.113)$$

(这是“Zeckendorf 定理”^{[201][312]}。)例如, 一百万的表示为

$$\begin{aligned}
1\,000\,000 &= 832\,040 + 121\,393 + 46\,368 + 144 + 55 \\
&= F_{30} + F_{26} + F_{24} + F_{12} + F_{10}.
\end{aligned}$$

我们用一种“贪婪”法总能找到这样一种表示, 选取 F_{k_1} 为最大 Fibonacci 数 $\leq n$, 然后选取 F_{k_2} 为最大 Fibonacci 数 $\leq n - F_{k_1}$, 等等。(更确切地说, 假设 $F_k \leq n < F_{k+1}$, 则我们有 $0 \leq n - F_k < F_{k+1} - F_k = F_{k-1}$ 。如果 n 是 Fibonacci 数, 则 $r=1$ 和 $k_1=k$ 时式(6.113)成立。否则对 n 归纳, $n - F_k$ 有 Fibonacci 表示 $F_{k_2} + \cdots + F_{k_r}$; 如果置 $k_1=k$, 式(6.113)成立, 因为不等式 $F_{k_2} \leq n - F_k < F_{k-1}$ 意味着 $k \gg k_2$ 。)相反, 形式(6.113)的任何表示意味着

$$F_{k_1} \leq n < F_{k_1+1},$$

因为当 $k \gg k_2 \gg \cdots \gg k_r \gg 0$ 时 $F_{k_2} + \cdots + F_{k_r}$ 的最大可能值为

$$F_{k-2} + F_{k-4} + \cdots + F_{k \bmod 2 + 2} = F_{k-1} - 1, \quad \text{若 } k \geq 2. \quad (6.114)$$

(对 k 归纳容易证明此公式, 当 k 为 2 或 3 时左边为零。)所以 k_1 是前面所说的贪婪选取值, 且表示一定是唯一的。

任何唯一的表示系统是一个数系, 所以 Zeckendorf 定理导致 Fibonacci 数系。我们能

把任何非负整数 n 表为 0 和 1 的序列, 记为

$$n = (b_m b_{m-1} \cdots b_2)_F \Leftrightarrow n = \sum_{k=2}^m b_k F_k. \quad (6.115)$$

此数系有些像二进(基数 2)表示, 除了没有两个相邻 1 之外。例如, 这里是 1 到 20 的数, 以 Fibonacci 数系表达:

$$\begin{array}{llll} 1 = (000001)_F & 6 = (001001)_F & 11 = (010100)_F & 16 = (100100)_F \\ 2 = (000010)_F & 7 = (001010)_F & 12 = (010101)_F & 17 = (100101)_F \\ 3 = (000100)_F & 8 = (010000)_F & 13 = (100000)_F & 18 = (101000)_F \\ 4 = (000101)_F & 9 = (010001)_F & 14 = (100001)_F & 19 = (101001)_F \\ 5 = (001000)_F & 10 = (010010)_F & 15 = (100010)_F & 20 = (101010)_F \end{array}$$

就刚才指出的一百万的 Fibonacci 表示能和它的二进制表示 $2^{19} + 2^{18} + 2^{17} + 2^{16} + 2^{14} + 2^9 + 2^6$ 对照:

$$\begin{aligned} (1000000)_{10} &= (100010100000000000101000000)_F \\ &= (11110100001001000000)_F \end{aligned}$$

因为 Fibonacci 表示不允许有相邻 1, 所以 Fibonacci 表示需要多一些的位; 但是两种表示是相似的。

在 Fibonacci 数系中加 1, 有两种情形: 如果“单位数字位”是 0, 则我们把它改成 1; 由于单位数字位涉及 F_2 , 所以添加 $F_2 = 1$ 。另外, 两个最小有意义的数字位将是 01, 把它们改成 10(从而加了 $F_3 - F_2 = 1$)。最后, 为了改变数字位型式‘011’为‘100’直到一行中没有两个 1, 我们一定要作必需那么多的“进位”(进位规则等价于用 F_{m+2} 代替 $F_{m+1} + F_m$)。例如, 从 $5 = (1000)_F$ 到 $6 = (1001)_F$ 或从 $6 = (1001)_F$ 到 $7 = (1010)_F$ 不要求进位; 但是从 $7 = (1010)_F$ 到 $8 = (10000)_F$ 一定要进位二次。

到现在为止, 我们已讨论了许多 Fibonacci 数的性质, 但是还没有提出它们的闭公式。我们也没有找 Stirling 数, 欧拉数, 或伯努利数的闭形式; 但是我们能发现调和数的闭形式 $H_n = \left[\begin{smallmatrix} n+1 \\ 2 \end{smallmatrix} \right] / n!$ 。是否在 F_n 和我们知道的其他量之间有一个关系? 能“解”定义 F_n 的递归吗?

回答是肯定的。事实上, 用第五章中简单提到的母函数的概念有一种简单方法解递归。让我们考虑无穷级数

$$F(z) = F_0 + F_1 z + F_2 z^2 + \cdots = \sum_{n \geq 0} F_n z^n. \quad (6.116)$$

如果能找到 $F(z)$ 的一个简单公式, 我们就有机会找到它的系数 F_n 的简单公式。

在第七章中将详细研究母函数, 但是有这个例子到那时是有帮助的。如果我们查看当用 z 和 z^2 乘它时, 幂级数 $F(z)$ 的好的性质有:

$$\begin{aligned}
 F(z) &= F_0 + F_1 z + F_2 z^2 + F_3 z^3 + F_4 z^4 + F_5 z^5 + \cdots, \\
 zF(z) &= F_0 z + F_1 z^2 + F_2 z^3 + F_3 z^4 + F_4 z^5 + \cdots, \\
 z^2 F(z) &= F_0 z^2 + F_1 z^3 + F_2 z^4 + F_3 z^5 + \cdots.
 \end{aligned}$$

如果现在从第一个方程减去后两个方程, 由于 Fibonacci 递归, 涉及 z^2 , z^3 和 z 的高次幂的项将不出现。因为 $F_0 = 0$, 在第一个位置常数项 F_0 实际上不出现。所以减后仅留下 $(F_1 - F_0)z$, 这就是 z 。换句话说,

$$F(z) - zF(z) - z^2 F(z) = z,$$

解 $F(z)$ 给出紧凑的公式

$$F(z) = \frac{z}{1 - z - z^2}. \quad (6.117)$$

我们现在已把 Fibonacci 序列的所有信息归结为一个简单的(虽然没有认出的)表达式 $z/(1 - z - z^2)$ 。这(信不信由你)是前进了一步, 因为能把分母因子分解, 然后用部分分式得到一个能容易展成幂级数的公式。此幂级数中的系数将是 Fibonacci 数的一个闭形式。

如果倒着处理它, 则刚叙述的处理方法也许能理解得较好。如果我们有一个较简单的母函数, 比如说 $1/(1 - \alpha z)$, 其中 α 是常数, 则我们知道所有 z 幂的系数, 因为

$$\frac{1}{1 - \alpha z} = 1 + \alpha z + \alpha^2 z^2 + \alpha^3 z^3 + \cdots.$$

相似, 如果我们有一个形式 $A/(1 - \alpha z) + B/(1 - \beta z)$ 的母函数, 容易确定系数, 因为

$$\begin{aligned}
 \frac{A}{1 - \alpha z} + \frac{B}{1 - \beta z} &= A \sum_{n \geq 0} (\alpha z)^n + B \sum_{n \geq 0} (\beta z)^n \\
 &= \sum_{n \geq 0} (A\alpha^n + B\beta^n) z^n.
 \end{aligned} \quad (6.118)$$

所以我们仅需找常数 A , B , α 和 β , 使得

$$\frac{A}{1 - \alpha z} + \frac{B}{1 - \beta z} = \frac{z}{1 - z - z^2},$$

而且我们将找 $F(z)$ 中 z^n 的系数 F_n 的闭形式 $A\alpha^n + B\beta^n$ 。左边能改写成

$$\frac{A}{1 - \alpha z} + \frac{B}{1 - \beta z} = \frac{A - A\beta z + B - B\alpha z}{(1 - \alpha z)(1 - \beta z)},$$

所以寻找的 4 个常数是两个多项式方程的解:

$$(1 - \alpha z)(1 - \beta z) = 1 - z - z^2; \quad (6.119)$$

$$(A + B) - (A\beta + B\alpha)z = z. \quad (6.120)$$

我们要把 $F(z)$ 的分母因子分解为形式 $(1-\alpha z)(1-\beta z)$, 这样我们将能把 $F(z)$ 表达为两个分式之和, 因子 $(1-\alpha z)$ 和 $(1-\beta z)$ 彼此容易分开.

注意式(6.119)中的分母因子已写成形式 $(1-\alpha z)(1-\beta z)$, 而不是较通常的形式 $c(z-\rho_1) \cdot (z-\rho_2)$, 其中 ρ_1 和 ρ_2 是根. 因为 $(1-\alpha z)(1-\beta z)$ 导致较好的幂级数展开.

我们可用几种方法来找 α 和 β , 其中之一用了巧妙的方法: 让我们引入一个新变量 w 且试求因子分解

$$w^2 - wz - z^2 = (w - \alpha z)(w - \beta z).$$

然后简单地置 $w = 1$, 我们将得到 $1 - z - z^2$ 的因子. 用二次公式能找 $w^2 - wz - z^2 = 0$ 的根, 它们是

$$\frac{z \pm \sqrt{z^2 + 4z^2}}{2} = \frac{1 \pm \sqrt{5}}{2} z.$$

所以

$$w^2 - wz - z^2 = \left(w - \frac{1 + \sqrt{5}}{2} z \right) \left(w - \frac{1 - \sqrt{5}}{2} z \right),$$

且我们得到所要找的常数 α 和 β .

如同在许多技术方面那样, 在数学的许多方面数 $(1 + \sqrt{5})/2 \approx 1.61803$ 是重要的, 在古代它被认为是许多设计类型的最中意的比率. 所以它有一个特殊名称, 黄金分割. 我们用希腊字母 φ 记它, 以纪念 Phidias, 他在雕刻术中使用了它. 另一个根 $(1 - \sqrt{5})/2 = -1/\varphi \approx -.61803$ 共享 φ 的许多性质, 所以它有特殊名称 $\hat{\varphi}$. 这些数是方程 $w^2 - w - 1 = 0$ 的根, 所以我们有

$$\varphi^2 = \varphi + 1; \quad \hat{\varphi}^2 = \hat{\varphi} + 1. \quad (6.121)$$

(后面再论 φ 和 $\hat{\varphi}$.)

我们已找到式(6.119)中需要的常数 $\alpha = \varphi$ 和 $\beta = \hat{\varphi}$, 现在仅需找式(6.120)中的 A 和 B . 在这方程中置 $z = 0$ 得到 $B = -A$, 所以式(6.120)归结为

$$-\hat{\varphi}A + \varphi A = 1.$$

解是 $A = 1/(\varphi - \hat{\varphi}) = 1/\sqrt{5}$, 式(6.117)的部分分式展开成

$$F(z) = \frac{1}{\sqrt{5}} \left(\frac{1}{1 - \varphi z} - \frac{1}{1 - \hat{\varphi} z} \right). \quad (6.122)$$

我们已取得了所要的恰当的 $F(z)$. 把分式展成像式(6.118)中的幂级数, 给出 z^n 的系数的闭形式:

$$F_n = \frac{1}{\sqrt{5}} (\varphi^n - \hat{\varphi}^n). \quad (6.123)$$

(此公式首先由 Leonhard Euler^[91]于 1765 年发表, 但是人们忘了它, 直到 1843 年 Jacques

Binet^[23]重新发现。)

在我们停止奇特的推导之前, 应检验它的准确性。对 $n=0$ 公式正确地给出 $F_0=0$; 对于 $n=1$, 它给出 $F_1=(\varphi-\hat{\varphi})/\sqrt{5}$, 它确实为 1, 对于较高幂, 方程(6.121)表明由式(6.123)定义的数满足 Fibonacci 递归, 所以由归纳法它们一定是 Fibonacci 数。(用二项定理还能展开 φ^n 和 $\hat{\varphi}^n$, 且找出 $\sqrt{5}$ 的各种幂, 但是这相当凌乱。一个闭形式不一定提供一种快速计算的方法, 但是可更确切地告诉我们在数学方面 F_n 如何与其他量联系起来。)

稍用一点洞察力, 我们将简单地猜出公式(6.123)且用归纳法证明它。但是母函数法是发现它的一种有力方法, 在第七章中我们将看到同一种方法引出困难得多的递归的解。此外, 我们不用操心在推导式(6.123)中无穷和是否收敛; 结果, 幂级数的系数上的大多数运算能严格证明有理, 不管和实际收敛还是不收敛^[15]。还有, 觉得无穷和推理靠不住而怀疑的读者, 可用一个可靠的归纳证明验证用无穷级数发现的方程(6.123)而自慰。

式(6.123)的有趣结果之一是当 n 大时整数 F_n 极接近于无理数 $\varphi^n/\sqrt{5}$ 。(因为 $\hat{\varphi}$ 的绝对值小于 1, $\hat{\varphi}^n$ 以指数律地变小且它几乎可忽略。)例如, $F_{10}=55$ 且 $F_{11}=89$ 是十分接近

$$\frac{\varphi^{10}}{\sqrt{5}} \approx 55.003\ 64 \quad \text{和} \quad \frac{\varphi^{11}}{\sqrt{5}} \approx 88.997\ 75.$$

我们能利用此观察引出另一个闭形式,

$$F_n = \left\lfloor \frac{\varphi^n}{\sqrt{5}} + \frac{1}{2} \right\rfloor = \frac{\varphi^n}{\sqrt{5}} \quad \text{四舍五入到最近整数,} \quad (6.124)$$

因为对所有 $n \geq 0$, $|\hat{\varphi}^n/\sqrt{5}| < 1/2$, 当 n 是偶数时, F_n 稍小于 $\varphi^n/\sqrt{5}$; 否则它稍大。

Cassini 等式(6.103)可改写为

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

当 n 大时, $1/F_{n-1}F_n$ 非常小, 所以 F_{n+1}/F_n 一定非常接近 F_n/F_{n-1} ; 且式(6.124)告诉我们这个比率近似于 φ 。事实上, 我们有

$$F_{n+1} = \varphi F_n + \hat{\varphi}^n. \quad (6.125)$$

(当 $n=0$ 或 $n=1$ 时, 通过检验此等式为真; 当 $n>1$ 时, 用归纳法证此等式为真; 通过代入式(6.123)也能直接证明它。)比率 F_{n+1}/F_n 起伏地接近 φ 。

由于巧合, φ 也十分接近 1 英里中的公里数(确切数为 1.609 344, 由于 1 英寸恰好为 2.54 cm。)这给出了一种公里和英里之间转换的方便的方法, 因为 F_{n+1} 公里的距离(十分接近)为 F_n 英里的距离。

假设我们要把一个非 Fibonacci 数从公里转换到英里; 30 km 是多少, 美国型式? 容

易：我们就用 Fibonacci 数系，用前面说明的贪婪方式把 30 转换到它的 Fibonacci 表示 $21+8+1$ 。现在我们能将每个数移下一个等级，取得 $13+5+1$ 。（前一个‘1’为 F_2 ，因为在式 (6.113) 中 $k_r \gg 0$ ；新的‘1’为 F_1 。） φ 近乎除尽移下的量，因此 19 英里是我们的估计。（这相当接近，正确解答是 18.64 英里左右。）相似，从英里到公里我们能移上一级；30 英里近似于 $34+13+2=49$ km。（这不是非常接近，正确数为 48.28 左右。）

结果，此“移下”规则给出，对所有 $n \leq 100$ 每 n km 正确的四舍五入的英里数，除了情形 $n=4, 12, 62, 75, 91$ 和 96 外，其差小于 $2/3$ 英里。而“移上”规则，或者给出正确的四舍五入的 n 英里的公里数，或者对于所有 $n \leq 126$ ，1 km 太多了。（真正麻烦的情形仅为 $n=4$ ，此时 $n=3+1$ 的特殊的四舍五入误差都成为同一方向，而不是彼此相消。）

6.7 延拓

Fibonacci 数和第四章研究的 Stern-Brocot 树有重要的联系，且它们具有欧拉广泛研究的多项式序列的重要的归纳结果。这些多项式称为延拓多项式，因为它们是连分式

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \frac{1}{a_5 + \frac{1}{a_6 + \frac{1}{a_7}}}}}}} \quad (6.126)$$

的研究的关键。

延拓多项式 $K_n(x_1, x_2, \dots, x_n)$ 有 n 个参数，它由下列递归定义：

$$\begin{aligned} K_0(\) &= 1 ; \\ K_1(x_1) &= x_1 ; \\ K_n(x_1, \dots, x_n) &= K_{n-1}(x_1, \dots, x_{n-1})x_n + K_{n-2}(x_1, \dots, x_{n-2}) . \end{aligned} \quad (6.127)$$

例如， $K_1(x_1)$ 后面的三个是：

$$\begin{aligned} K_2(x_1, x_2) &= x_1 x_2 + 1 ; \\ K_3(x_1, x_2, x_3) &= x_1 x_2 x_3 + x_1 + x_3 ; \\ K_4(x_1, x_2, x_3, x_4) &= x_1 x_2 x_3 x_4 + x_1 x_2 + x_1 x_4 + x_3 x_4 + 1 . \end{aligned}$$

由此容易归纳地看到项数是一个 Fibonacci 数：

$$K_n(1, 1, \dots, 1) = F_{n+1} . \quad (6.128)$$

当由前后关系暗示参数的个数时，我们可简单地写 ‘ K ’ 而不是 ‘ K_n ’，就像我们

一个长度为 n 的莫尔斯电码序列，总共有 k 个划有 $n-2k$ 个点和 $n-k$ 个符号。这些点和划能以 $\binom{n-k}{k}$ 种方式排列，所以如果我们用 z 代替每点用 1 代替每划，则取得

我们还知道一个延拓中的总项数是一个 Fibonacci 数。因此有等式

(式(6.129)的一个闭形式, 推广了出现在式(5.84)中 Fibonacci 数的 Euler-Binet 公式(6.123).)

$$K(x_n, \dots, x_2, x_1) = K(x_1, x_2, \dots, x_n). \quad (6.131)$$
$$K_n(x_1, \dots, x_n) = x_1 K_{n-1}(x_2, \dots, x_n) + K_{n-2}(x_1, \dots, x_n). \quad (6.132)$$
$$\begin{aligned} & K_{m+n}(x_1, \dots, x_m, x_{m+1}, \dots, x_{m+n}) \\ &= K_m(x_1, \dots, x_m)K_n(x_{m+1}, \dots, x_{m+n}) \\ &+ K_{m-1}(x_1, \dots, x_{m-1})K_{n+1}(x_{m+2}, \dots, x_{m+n}) . \end{aligned} \quad (6.133)$$

欧拉^[90]发现延拓还服从一个值得注意的定律, 它推广了 Cassini 等式:

$$\begin{aligned}
& K_{m+n}(x_1, \dots, x_{m+n})K_k(x_{m+1}, \dots, x_{m+k}) \\
&= K_{m+k}(x_1, \dots, x_{m+k})K_m(x_{m+1}, \dots, x_{m+n}) \\
&+ (-1)^k K_{m-1}(x_1, \dots, x_{m-1})K_{n-k-1}(x_{m+k+2}, \dots, x_{m+n}). \quad (6.134)
\end{aligned}$$

每当 K 上的下标全为非负时, 此定律成立(在习题 29 中证明)。例如, 当 $k=2$, $m=1$ 和 $n=3$ 时, 我们有

$$K(x_1, x_2, x_3, x_4)K(x_2, x_3) = K(x_1, x_2, x_3)K(x_2, x_3, x_4) + 1.$$

延拓多项式和欧几里德算法有密切联系。例如, 假设在 4 步中结束 $\gcd(m, n)$ 的计算:

$$\begin{aligned}
\gcd(m, n) &= \gcd(n_0, n_1) & n_0 &= m, \quad n_1 = n; \\
&= \gcd(n_1, n_2) & n_2 &= n_0 \bmod n_1 = n_0 - q_1 n_1; \\
&= \gcd(n_2, n_3) & n_3 &= n_1 \bmod n_2 = n_1 - q_2 n_2; \\
&= \gcd(n_3, n_4) & n_4 &= n_2 \bmod n_3 = n_2 - q_3 n_3; \\
&= \gcd(n_4, 0) = n_4 & 0 &= n_3 \bmod n_4 = n_3 - q_4 n_4.
\end{aligned}$$

于是我们有

$$\begin{aligned}
n_4 &= n_4 & &= K(\quad)n_4; \\
n_3 &= q_4 n_4 & &= K(q_4)n_4; \\
n_2 &= q_3 n_3 + n_4 & &= K(q_3, q_4)n_4; \\
n_1 &= q_2 n_2 + n_3 & &= K(q_2, q_3, q_4)n_4; \\
n_0 &= q_1 n_1 + n_2 & &= K(q_1, q_2, q_3, q_4)n_4.
\end{aligned}$$

一般, 如果在计算商序列 q_1, \dots, q_k 之后欧几里德算法在 k 步中找到最大公因子 d , 则开始数为 $K(q_1, q_2, \dots, q_k)d$ 和 $K(q_2, \dots, q_k)d$ 。(18 世纪初期 Thomas Fantet de Lagny^[190]注意到此事实, 他看来是明确研究延拓的第一个人。Lagny 指出, 当 q 取它们的最小值时作为延拓出现的连贯 Fibonacci 数是引起欧几里德算法取一个给定步数的最小输入。)

延拓还和连分式有密切联系, 因此把它们取名为延拓。例如, 我们有

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3}}} = \frac{K(a_0, a_1, a_2, a_3)}{K(a_1, a_2, a_3)}. \quad (6.135)$$

对于任何深度的连分式相同型式成立。用归纳法易证; 例如, 我们有

$$\frac{K\left(a_0, a_1, a_2, a_3 + \frac{1}{a_4}\right)}{K\left(a_1, a_2, a_3 + \frac{1}{a_4}\right)} = \frac{K(a_0, a_1, a_2, a_3, a_4)}{K(a_1, a_2, a_3, a_4)},$$

由于等式

$$\begin{aligned} K_n(x_1, \dots, x_{n-1}, x_n + y) \\ = K(x_1, \dots, x_{n-1}, x_n) + K_{n-1}(x_1, \dots, x_{n-1})y. \end{aligned} \quad (6.136)$$

(在习题 30 中证明并推广此等式。)

此外, 延拓与第四章中讨论的 Stern-Brocot 树有紧密联系。此树中的每个结点可表示为一个 L 和 R 的序列, 比如说

$$R^{a_0} L^{a_1} R^{a_2} L^{a_3} \dots R^{a_{n-2}} L^{a_{n-1}}, \quad (6.137)$$

其中 $a_0 \geq 0, a_1 \geq 1, a_2 \geq 1, a_3 \geq 1, \dots, a_{n-2} \geq 1, a_{n-1} \geq 0$, 且 n 为偶数, 用式 (4.33) 的 2×2 矩阵 L 和 R , 由归纳法不难证明式 (6.137) 的等价矩阵为

$$\begin{pmatrix} K_{n-2}(a_1, \dots, a_{n-2}) & K_{n-1}(a_1, \dots, a_{n-2}, a_{n-1}) \\ K_{n-1}(a_0, a_1, \dots, a_{n-2}) & K_n(a_0, a_1, \dots, a_{n-2}, a_{n-1}) \end{pmatrix} \quad (6.138)$$

(证明是习题 80 的一部分。)例如,

$$R^a L^b R^c L^d = \begin{pmatrix} bc + 1 & bcd + b + d \\ abc + a + c & abcd + ab + ad + cd + 1 \end{pmatrix}.$$

所以我们最后能用式 (4.34) 写出 Stern-Brocot 树的 L 和 R 表示为式 (6.137) 的分式的一个闭形式:

$$f(R^{a_0} \dots L^{a_{n-1}}) = \frac{K_{n+1}(a_0, a_1, \dots, a_{n-1}, 1)}{K_n(a_1, \dots, a_{n-1}, 1)}. \quad (6.139)$$

(这是“Halphen 定理”^[143]。)例如, 找 $LRRL$ 的分式, 我们有 $a_0 = 0, a_1 = 1, a_2 = 2, a_3 = 1$ 和 $n = 4$; 方程 (6.139) 给出

$$\frac{K(0, 1, 2, 1, 1)}{K(1, 2, 1, 1)} = \frac{K(2, 1, 1)}{K(1, 2, 1, 1)} = \frac{K(2, 2)}{K(3, 2)} = \frac{5}{7}.$$

(我们用规则 $K_n(x_1, \dots, x_{n-1}, x_n + 1) = K_{n+1}(x_1, \dots, x_{n-1}, x_n, 1)$ 在参数列表中并入第一位的 1 和尾部的 1, 在式 (6.136) 中置 $y = 1$ 获得此规则。)

比较式 (6.135) 和 (6.139) 表明, 对应于 Stern-Brocot 树中的一般结点式 (6.137) 的分式有连分式表示

$$f(R^{a_0} \dots L^{a_{n-1}}) = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_{n-1} + \frac{1}{1}}}}} \quad (6.140)$$

因此立刻就能在连分式和 Stern-Brocot 树中对应的结点之间转换。例如,

$$f(LRRL) = 0 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{1}}}}$$

在第四章我们看到无理数确定 Stern-Brocot 树中的无限路径, 它们能表为 L 和 R 的一个无限串。如果 α 的无限串为 $R^{a_0} L^{a_1} R^{a_2} L^{a_3} \dots$, 则有一个对应的无限连分式

$$\alpha = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \frac{1}{a_5 + \ddots}}}}} \quad (6.141)$$

此无限连分式也能直接获得: 设 $\alpha_0 = \alpha$, 且对于 $k \geq 0$ 设

$$a_k = \lfloor \alpha_k \rfloor, \quad \alpha_k = a_k + \frac{1}{\alpha_{k+1}} \quad (6.142)$$

a 称为 α 的“部分商”。如果 α 是有理数, 比如说 m/n , 此过程遍及欧几里德算法求得的商, 然后(和 $\alpha_{k+1} = \infty$ 一起)停止。

欧拉常数 γ 是有理数还是无理数? 没有人知道。我们能通过在 Stern-Brocot 树中找 γ 取得关于这个有名的未解决问题的部分信息; 如果它是有理数我们将发现它, 而如果它是无理数我们将找到它的所有最接近的有理近似。 γ 的连分式从下列部分商开始:

k	0	1	2	3	4	5	6	7	8
a_k	0	1	1	2	1	2	1	4	3

所以它的 Stern-Brocot 表示开始 $LRLLRLLRLLLRRL\dots$, 没有明显的型式。依据 Richard Brent 的计算^[33]指出, 如果 γ 是有理数, 它的分母一定比 10 000 个十进数字位更长。所以没有人相信 γ 是有理数, 但是至今没有人能证明它不是有理数。

让我们证明一个把一大堆概念结合在一起的值得注意的等式来结束本章。我们引入第

三章中的谱的概念; α 的谱是数 $\lfloor n\alpha \rfloor$ 的多重集合, 其中 α 是给定常数. 所以无穷级数

$$\sum_{n \geq 1} z^{\lfloor n\alpha \rfloor} = z + z^3 + z^4 + z^6 + z^8 + z^9 + \cdots$$

能称为 φ 的谱的母函数, 其中 $\varphi = (1 + \sqrt{5})/2$ 为黄金分割. 我们将证明 J. L. Davison 在 1976 年发现的等式^[61], 它是联系到这个 Fibonacci 序列的母函数的一个无穷连分式:

$$\frac{z^{F_1}}{1 + \frac{z^{F_2}}{1 + \frac{z^{F_3}}{1 + \frac{z^{F_4}}{\ddots}}}}} = (1 - z) \sum_{n \geq 1} z^{\lfloor n\varphi \rfloor}. \quad (6.143)$$

式(6.143)的两边是有趣的, 让我们首先查看数 $\lfloor n\varphi \rfloor$. 如果 n 的 Fibonacci 表示式(6.113)为 $F_{k_1} + \cdots + F_{k_r}$, 我们期望 $n\varphi$ 逼近 $F_{k_1+1} + \cdots + F_{k_r+1}$, (如同当从英里转换到公里那样)我们把数的 Fibonacci 表示左移. 事实上, 我们从式(6.125)知道

$$n\varphi = F_{k_1+1} + \cdots + F_{k_r+1} - (\hat{\varphi}^{k_1} + \cdots + \hat{\varphi}^{k_r}).$$

现在 $\hat{\varphi} = -1/\varphi$ 和 $k_1 \gg \cdots \gg k_r \gg 0$, 所以有

$$\begin{aligned} |\hat{\varphi}^{k_1} + \cdots + \hat{\varphi}^{k_r}| &< \varphi^{-k_1} + \varphi^{-k_1-2} + \varphi^{-k_1-4} + \cdots \\ &= \frac{\varphi^{-k_1}}{1 - \varphi^{-2}} = \varphi^{1-k_1} \leq \varphi^{-1} < 1; \end{aligned}$$

同样论证, $\hat{\varphi}^{k_1} + \cdots + \hat{\varphi}^{k_r}$ 和 $(-1)^{k_r}$ 有相同正负号. 因此

$$\lfloor n\varphi \rfloor = F_{k_1+1} + \cdots + F_{k_r+1} - [k_r (n \text{ 是偶数})]. \quad (6.144)$$

让我们称一个数 n 为 **Fibonacci 奇数**, 如果它的最小有意义的 Fibonacci 位是 1; 这与称 $K_r(n) = 2$ 相同. 否则 n 是 **Fibonacci 偶数**. 例如, 最小的 Fibonacci 奇数是 1, 4, 6, 9, 12, 14, 17 和 19. 若 $K_r(n)$ 是偶数, 则由式(6.114) $n-1$ 是 Fibonacci 偶数; 相似, 如果 $K_r(n)$ 是奇数, 则 $n-1$ 是 Fibonacci 奇数. 所以

$$K_r(n) \text{ 是偶数} \Leftrightarrow n-1 \text{ 是 Fibonacci 偶数.}$$

此外, 如果 $K_r(n)$ 是偶数, 式(6.144)暗指 $K_r(\lfloor n\varphi \rfloor) = 2$; 如果 $K_r(n)$ 是奇数, 式(6.144)表明 $K_r(\lfloor n\varphi \rfloor) = K_r(n) + 1$. 所以 $K_r(\lfloor n\varphi \rfloor)$ 总是偶数, 我们证明了

$$\lfloor n\varphi \rfloor - 1 \text{ 总是 Fibonacci 偶数.}$$

反之, 如果 m 是任何 Fibonacci 偶数, 我们能把这个计算颠倒过来且求一个 n 使得 $m+1 = \lfloor n\varphi \rfloor$. (如同前面说明的那样, 首先在 F 记法中加 1. 如果没有进位出现, n 是 $(m+2)$ 右移; 否则 n 是 $(m+1)$ 右移). 所以式(6.143)的右边和能写成

$$\sum_{n \geq 1} z^{\lfloor n\phi \rfloor} = z \sum_{m \geq 0} z^m [m \text{ 是 Fibonacci 偶数}]. \quad (6.145)$$

左边的分式怎么样？让我们改写式(6.143)以致连分式看来像式(6.141)，所有分子为 1：

$$\frac{1}{z^{-F_0} + \frac{1}{z^{-F_1} + \frac{1}{z^{-F_2} + \frac{1}{\ddots}}}}} = \frac{1-z}{z} \sum_{n \geq 1} z^{\lfloor n\phi \rfloor}. \quad (6.146)$$

(这个变换有点复杂！具有 z^{F_n} 作为分子的原来分式的分子和分母应被 z^{F_n-1} 除尽。)如果在 $1/z^{-F_n}$ 处中止这个新连分式，如同式(6.135)中的那样，它的值将是延拓的比，

$$\frac{K_{n+2}(0, z^{-F_0}, z^{-F_1}, \dots, z^{-F_n})}{K_{n+1}(z^{-F_0}, z^{-F_1}, \dots, z^{-F_n})} = \frac{K_n(z^{-F_1}, \dots, z^{-F_n})}{K_{n+1}(z^{-F_0}, z^{-F_1}, \dots, z^{-F_n})}.$$

让我们首先查看分母，希望它将易处理。置 $Q_n = K_{n+1}(z^{-F_0}, \dots, z^{-F_n})$ ，我们求得 $Q_0 = 1$ ， $Q_1 = 1 + z^{-1}$ ， $Q_2 = 1 + z^{-1} + z^{-2}$ ， $Q_3 = 1 + z^{-1} + z^{-2} + z^{-3} + z^{-4}$ ，且一般都配合得极好，得到一个几何级数

$$Q_n = 1 + z^{-1} + z^{-2} + \dots + z^{-(F_{n+2}-1)}.$$

对应的分子是 $P_n = K_n(z^{-F_1}, \dots, z^{-F_n})$ ，这产生像 Q_n 的形式但具有较少的项。例如，与 $Q_5 = 1 + z^{-1} + \dots + z^{-12}$ 比较，我们得到

$$P_5 = z^{-1} + z^{-2} + z^{-4} + z^{-5} + z^{-7} + z^{-9} + z^{-10} + z^{-12}.$$

较仔细查看呈现出所出示的项的型式：我们得到

$$P_5 = \frac{1 + z^2 + z^3 + z^5 + z^7 + z^8 + z^{10} + z^{11}}{z^{12}} = z^{-12} \sum_{m=0}^{12} z^m [m \text{ 是 Fibonacci 偶数}];$$

一般我们用归纳法能证明

$$P_n = z^{1-F_{n+2}} \sum_{m=0}^{F_{n+2}-1} z^m [m \text{ 是 Fibonacci 偶数}].$$

所以

$$\frac{P_n}{Q_n} = \frac{\sum_{m=0}^{F_{n+2}-1} z^m [m \text{ 是 Fibonacci 偶数}]}{\sum_{m=0}^{F_{n+2}-1} z^m}.$$

由于式(6.145)，现在当 $n \rightarrow \infty$ 取极限得到式(6.146)。

习题

准备部分

1. 恰好有两个轮换的 $\{1, 2, 3, 4\}$ 的 $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 11$ 个排列是什么? (轮换形式出现在式(6.4)中, 希望像 2314 的非轮换形式来代替.)
2. 从 n 个元素的一个集合到 m 个元素的一个集合有 m^n 个函数. 它们中间有多少个恰好在 k 个不同的函数值的范围内变化?
3. 在现实世界中堆卡片者都知道, 允许放松一点是明智的, 以使当出现一点风时卡片将不致倒塌. 假设要求顶上 k 张卡片的重心从第 $k+1$ 张卡片的边起至少为 ε 个单位, (例如, 因此第一张卡片对第二张卡片至多伸出 $1-\varepsilon$ 个单位.) 如果我们有足够的卡片, 是否仍能达到任何大的伸出?
4. 依据调和数表达 $1/1 + 1/3 + \cdots + 1/(2n+1)$.
5. 说明如何根据式(6.74)中的 $U_n(x, y)$ 的定义得到递归(6.75), 且解此递归.
6. 一个探险者在一个岛上留下一对幼兔. 如果幼兔在一个月后变成熟, 且每对成熟兔每月生产一对幼兔, n 个月后出现多少对兔? (两个月后有两对, 一对是新生的.) 在这个问题和书中的“蜂树”之间找出一个联系.
7. 证明 Cassini 等式(6.103)是式(6.108)以及(6.134)的特殊情形.
8. 用 Fibonacci 数系把 65 英里 / n 转换成 km / h 的近似数.
9. 在 8 平方英里中大约有多少平方公里?
10. φ 的连分式表示是什么?

基本部分

11. 当 n 是非负整数时, 具有交错正负号的 Stirling 轮换数三角形的行和 $\sum_k (-1)^k \cdot \begin{bmatrix} n \\ k \end{bmatrix}$ 是多少?

12. 证明 Stirling 数有相似于式(5.48)的反演定律形式:

$$g(n) = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (-1)^k f(k) \Leftrightarrow f(n) = \sum_k \left[\begin{matrix} n \\ k \end{matrix} \right] (-1)^k g(k).$$

13. 在第二章和第五章中提到的微分算子 $D = \frac{d}{dz}$ 和 $\theta = zD$. 我们得到

$$\theta^2 = z^2 D^2 + zD,$$

因为 $\theta^2 f(z) = \theta z f'(z) = z \frac{d}{dz} z f'(z) = z^2 f''(z) + z f'(z)$, 它是 $(z^2 D^2 + zD)f(z)$. 由此能

相似指出 $\theta^3 = z^3 D^3 + 3z^2 D^2 + zD$. 证明一般公式, 对于所有 $n \geq 0$.

$$\theta^n = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} z^k D^k,$$

$$z^n D^n = \sum_k \left[\begin{matrix} n \\ k \end{matrix} \right] (-1)^{n-k} \theta^k.$$

(如同式(5.109)中的那样, 可用这些公式在形式 $\sum_k \alpha_k z^k f^{(k)}(z)$ 和 $\sum_k \beta_k \theta^k f(z)$ 的微分表达式之间作转换.)

14. 证明欧拉数的幂等式(6.37).

15. 通过取式(6.37)的第 m 次差分来证明欧拉等式(6.39).

16. 当 k 和 n 的取值范围是所有整数的集合, 双重递归:

$$A_{n,0} = a_n [n \geq 0]; \quad A_{0,k} = 0, \text{ 如果 } k > 0;$$

$$A_{n,k} = k A_{n-1,k} + A_{n-1,k-1}, \text{ 整数 } k, n,$$

的一般解是什么?

17. 解下列递归, 假设当 $n < 0$ 或 $k < 0$ 时 $\left| \begin{matrix} n \\ k \end{matrix} \right|$ 为零:

$$(a) \left| \begin{matrix} n \\ k \end{matrix} \right| = \left| \begin{matrix} n-1 \\ k \end{matrix} \right| + n \left| \begin{matrix} n-1 \\ k-1 \end{matrix} \right| + [n=k=0], \quad n, k \geq 0.$$

$$(b) \left| \begin{matrix} n \\ k \end{matrix} \right| = (n-k) \left| \begin{matrix} n-1 \\ k \end{matrix} \right| + \left| \begin{matrix} n-1 \\ k-1 \end{matrix} \right| + [n=k=0], \quad n, k \geq 0.$$

$$(c) \left| \begin{matrix} n \\ k \end{matrix} \right| = k \left| \begin{matrix} n-1 \\ k \end{matrix} \right| + k \left| \begin{matrix} n-1 \\ k-1 \end{matrix} \right| + [n=k=0], \quad n, k \geq 0.$$

18. 证明 Stirling 多项式满足

$$(x+1)\sigma_n(x+1) = (x-n)\sigma_n(x) + x\sigma_{n-1}(x).$$

19. 证明广义 Stirling 数满足

$$\frac{\sum_{k=0}^n \left\{ \begin{matrix} x+k \\ x \end{matrix} \right\} \left[\begin{matrix} x \\ x-n+k \end{matrix} \right] (-1)^k}{\binom{x+k}{n+1}} = 0, \text{ 整数 } n > 0.$$

$$\frac{\sum_{k=0}^n \left[\begin{matrix} x+k \\ x \end{matrix} \right] \left\{ \begin{matrix} x \\ x-n+k \end{matrix} \right\} (-1)^k}{\binom{x+k}{n+1}} = 0, \text{ 整数 } n > 0.$$

20. 求 $\sum_{k=1}^n H_k^{(2)}$ 的一个闭形式.

21. 证明如果 $H_n = a_n / b_n$, 其中 a_n 和 b_n 是整数, 则分母 b_n 是 $2^{\lfloor \lg n \rfloor}$ 的倍数. 提示:

考虑数 $2^{\lfloor \log n \rfloor - 1} H_n - (1/2)$.

22. 证明除了当 z 是负整数外, 对于所有复数 z 无穷和

$$\sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+z} \right)$$

收敛; 且指出当 z 是非负整数时它等于 H_z . (所以当 z 是复数时, 能用此公式来定义调和数 H_z .)

23. 当展成 z 的幂时, 方程(6.81)给出 $z/(e^z - 1)$ 的系数, $z/(e^z + 1)$ 的系数是什么? 提示: 考虑等式 $(e^z + 1)(e^z - 1) = e^{2z} - 1$.

24. 证明正切数 T_{2n+1} 是 2^n 的倍数. 提示: 证明 $T_{2n}(x)$ 和 $T_{2n+1}(x)$ 的所有系数是 2^n 的倍数.

25. 方程(6.57)证明爬行在某个时刻 N 处最终将达到橡胶带的终端. 所以当比起 $n-1$ 分钟后来讲, 他在 n 分钟后更接近终端时, 第一个时刻 n 一定来到. 指出 $n < (1/2)N$.

26. 用分部求和法来计算 $S_n = \sum_{k=1}^n H_k / k$. 提示: 还考虑有关的和 $\sum_{k=1}^n H_{k-1} / k$.

27. 证明 Fibonacci 数的 gcd 定律(6.111).

28. Lucas 数 L_n 定义为 $F_{n+1} + F_{n-1}$. 于是, 按照式(6.109), 我们得到 $F_{2n} = F_n L_n$. 这里是前几个值的表:

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

(a) 用包含各种组成部分方法证明一般递归

$$Q_0 = \alpha, Q_1 = \beta, Q_n = Q_{n-1} + Q_{n-2}, n > 1$$

的解 Q_n 能依据 F_n 和 L_n 表达.

(b) 依据 φ 和 $\bar{\varphi}$ 求 L_n 的闭形式.

29. 证明延拓的欧拉等式, 方程(6.134).

30. 推广式(6.136)来求增量延拓 $K(x_1, \dots, x_{m-1}, x_m + y, x_{m+1}, \dots, x_n)$ 的一个表达式, 当 $1 \leq m \leq n$.

课外习题

31. 求下降幂表示上升幂中的系数 $\left| \begin{smallmatrix} n \\ k \end{smallmatrix} \right|$ 的闭形式:

$$x^{\bar{n}} = \sum_k \left| \begin{smallmatrix} n \\ k \end{smallmatrix} \right| x^{\underline{k}}, \text{ 整数 } n \geq 0.$$

(例如, $x^{\bar{4}} = x^{\underline{4}} + 12x^{\underline{3}} + 36x^{\underline{2}} + 24x^{\underline{1}}$, 因此 $\left| \begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right| = 36$.)

32. 在第五章中, 我们通过以两种方式展开递归 $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ 获得公式

$$\sum_{k \leq n} \binom{n+k}{k} = \binom{n+m+1}{m} \text{ 和 } \sum_{0 \leq k \leq n} \binom{k}{n} = \binom{m+1}{n+1}.$$

当展开相似递归 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = k \left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\} + \left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\}$ 时, 出现什么等式?

33. 表 6.3 给出 $\left[\begin{smallmatrix} n \\ 2 \end{smallmatrix} \right]$ 和 $\left\{ \begin{smallmatrix} n \\ 2 \end{smallmatrix} \right\}$ 的值, 接着的 $\left[\begin{smallmatrix} n \\ 3 \end{smallmatrix} \right]$ 和 $\left\{ \begin{smallmatrix} n \\ 3 \end{smallmatrix} \right\}$ 的闭形式(不涉及 Stirling 数)是什么?

34. 如果假设对所有整数 k 和 n 基本递归关系(6.35)成立, 且对于所有 $k < 0$, $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle = 0$, $\langle \begin{smallmatrix} -1 \\ k \end{smallmatrix} \rangle$ 和 $\langle \begin{smallmatrix} -2 \\ k \end{smallmatrix} \rangle$ 是什么?

35. 证明对于每一个 $\varepsilon > 0$ 存在一个整数 $n > 1$ (依赖 ε) 使得 $H_n \bmod 1 < \varepsilon$.

36. 是否可能以这样一种方式堆放 n 块砖, 最顶上的一块砖不在最底下一块砖的任何点的上面, 与 100 块砖一样重的一个人还能平衡在顶上一块砖的中间而这个砖堆不倒塌?

37. 假设 m 和 n 是正整数, 依据调和数来表达出 $\sum_{k=1}^n (k \bmod m) / k(k+1)$. 当 $n \rightarrow \infty$ 时极限值是什么?

38. 求不定和 $\sum \binom{r}{k} (-1)^k H_k \delta k$.

39. 依据 n 和 H_n 表达 $\sum_{k=1}^n H_k^2$.

40. 证明 1979 整除 $\sum_{k=1}^{1319} (-1)^{k-1} / k$ 的分子, 且给出 1987 的相似结果. 提示: 用 Gauss 方法来得到分式的一个和, 它的分子是 1979. 见习题 4.

41. 当 n 是整数(可能负)时, 以闭形式计算和 $\sum_k \binom{\lfloor (n+k)/2 \rfloor}{k}$.

42. 如果 S 是整数集, $S+1$ 是“移位”集 $\{x+1 | x \in S\}$. 具有性质 $S \cup (S+1) = \{1, 2, \dots, n+1\}$ 的 $\{1, 2, \dots, n\}$ 的子集有多少个?

43. 证明无穷和

$$\begin{aligned} & \cdot 1 \\ & + \cdot 01 \\ & + \cdot 002 \\ & + \cdot 0003 \\ & + \cdot 00005 \\ & + \cdot 000008 \\ & + \cdot 0000013 \\ & \vdots \end{aligned}$$

收敛到一个有理数.

44. 证明 Cassini 等式(6.106)的逆: 如果 k 和 m 是整数使得 $|m^2 - km - k^2| = 1$, 则有一个整数 n 使得 $k = \pm F_n$ 和 $m = \pm F_{n+1}$.

45. 用包含各种组成部分的方法来解一般的递归

$$X_0 = \alpha, X_1 = \beta, X_n = X_{n-1} + X_{n-2} + \gamma n + \delta.$$

46. $\cos 36^\circ$ 和 $\cos 72^\circ$ 是多少?

47. 证明

$$2^{n-1} F_n = \sum_k \binom{n}{2k+1} 5^k,$$

且用此等式来导出 $F_n \bmod p$ 和 $F_{n+1} \bmod p$ 的值, 当 p 是素数时.

48. 证明通过把邻近的参数折叠在一起能从延拓多项式中移去零值参数:

$$\begin{aligned} K_n(x_1, \dots, x_{m-1}, 0, x_{m+1}, \dots, x_n) \\ = K_{n-2}(x_1, \dots, x_{m-2}, x_{m-1} + x_{m+1}, x_{m+2}, \dots, x_n), \quad 1 < m < n. \end{aligned}$$

49. 求数 $\sum_{n \geq 1} 2^{-\lfloor \log n \rfloor}$ 的分式表示.

50. 由递归

$$f(1) = 1,$$

$$f(2n) = f(n),$$

$$f(2n+1) = f(n) + f(n+1),$$

对所有正整数 n 定义 $f(n)$.

(a) 对哪一些 n , $f(n)$ 为偶?

(b) 表明依据延拓能表达 $f(n)$.

考查性问题

51. 设 p 是素数.

(a) 证明 $\left\{ \begin{smallmatrix} p \\ k \end{smallmatrix} \right\} \equiv \left[\begin{smallmatrix} p \\ k \end{smallmatrix} \right] \equiv 0 \pmod{p}$, $1 < k < p$.

(b) 证明 $\left[\begin{smallmatrix} p-1 \\ k \end{smallmatrix} \right] \equiv 1 \pmod{p}$, $1 \leq k \leq p$.

(c) 证明 $\left\{ \begin{smallmatrix} 2p-2 \\ p \end{smallmatrix} \right\} \equiv \left[\begin{smallmatrix} 2p-2 \\ p \end{smallmatrix} \right] \equiv 0 \pmod{p}$.

(d) 证明如果 $p > 3$ 我们有 $\left[\begin{smallmatrix} p \\ 2 \end{smallmatrix} \right] \equiv 0 \pmod{p^2}$. 提示: 考虑 $p^{\frac{p}{2}}$.

52. 将 H_n 如同 a_n / b_n 记成最小项.

(a) 证明 $p \setminus b_n \Leftrightarrow p \times a_{\lfloor n/p \rfloor}$, 如果 p 是素数.

(b) 求出所有 $n > 0$ 使得 5 可除尽 a_n .

53. 当 $0 \leq m \leq n$ 时, 求 $\sum_{k=0}^m \binom{n}{k} (-1)^k H_k$ 的闭形式. 提示: 习题 5.42 得到 H_k 因子的和.

54. 设 $n > 0$, 此习题的目的是指出 B_{2n} 的分母是所有素数 p 的乘积使得 $(p-1) \nmid (2n)$.

(a) 当 p 是素数且 $m > 0$ 时, 指出 $S_m(p) + [(p-1) \setminus m]$ 是 p 的倍数.

(b) 用 (a) 的结果证明

$$B_{2n} + \sum_{p \text{ 素数}} \frac{[(p-1) \setminus (2n)]}{p} = I_{2n} \text{ 是一个整数.}$$

提示: 证明下列结果就够了, 如果 p 是任何素数, 分式 $B_{2n} + [(p-1) \setminus (2n)]/p$ 的分母不为 p 所除尽.

(c) 证明 B_{2n} 的分母总是 6 的一个奇倍数, 且对无限多个 n 它等于 6.

55. 证明式 (6.70) 作为一个更一般等式的一个推论, 通过求和

$$\sum_{0 \leq k \leq m} \binom{k}{m} \binom{x+k}{k}$$

以及对 x 求微分.

56. 计算 $\sum_{k \neq m} \binom{n}{k} (-1)^k k^{n+1} / (k-m)$ 为整数 m 和 n 的一个函数的闭形式 (除了 $k=m$ 值外, 在所有整数 k 上求和.)

57. 由

$$\binom{\binom{n}{k}}{k} = \binom{\binom{n-1}{k}}{k} + \binom{\binom{n-1}{(k-1) \bmod 5}}{k}, \quad n > 0,$$

以及 $\binom{\binom{0}{k}}{k} = [k=0]$ 定义“序模 5 的环绕二项系数”. 设 Q_n 是行 n 中这些数的最大和最小之间的差:

$$Q_n = \max_{0 \leq k < 5} \binom{\binom{n}{k}}{k} - \min_{0 \leq k < 5} \binom{\binom{n}{k}}{k}.$$

求和证明 Q_n 和 Fibonacci 数之间的一个关系.

58. 求 $\sum_{n \geq 0} F_n^2 z^n$ 和 $\sum_{n \geq 0} F_n^3 z^n$ 的闭形式. 关于数量 $F_{n+1}^3 - 4F_n^3 - F_{n-1}^3$ 你推断出什么?

59. 证明如果 m 和 n 是正整数, 则存在一个整数 x 使得 $F_x \equiv m \pmod{3^n}$.

60. 求出所有正整数 n 使得 $F_n + 1$ 或者 $F_n - 1$ 是一个素数.

61. 证明等式

$$\sum_{k=0}^n \frac{1}{F_{2^k}} = 3 - \frac{F_{2^n-1}}{F_{2^n}}, \quad \text{整数 } n \geq 1.$$

$\sum_{k=0}^n 1/F_{3 \cdot 2^k}$ 是什么?

62. 设 $A_n = \varphi^n + \varphi^{-n}$ 和 $B_n = \varphi^n - \varphi^{-n}$.

(a) 求常数 α 和 β , 使得对所有 $n \geq 0$, $A_n = \alpha A_{n-1} + \beta A_{n-2}$ 和 $B_n = \alpha B_{n-1} + \beta B_{n-2}$.

(b) 依据 F_n 和 L_n 表达 A_n 和 B_n (见习题 28).

(c) 证明 $\sum_{k=1}^n 1/(F_{2k+1} + 1) = B_n / A_{n+1}$.

(d) 求 $\sum_{k=1}^n 1/(F_{2k+1} - 1)$ 的闭形式.

额外问题

63. 有多少个 $\{1, 2, \dots, n\}$ 的排列 $\pi_1 \pi_2 \dots \pi_n$ 恰好有 k 个指标 j 使得

(a) 对所有 $i < j$, $\pi_i < \pi_j$? (这样的 j 称为“左对右的极大值.”)

(b) $\pi_j > j$? (这样的 j 称为“超过数.”)

64. $\left[\frac{1/2}{1/2 - n} \right]$ 的分母是什么, 当这个分式化为最小项时?

65. 证明等式

$$\int_0^1 \cdots \int_0^1 f(Lx_1 + \cdots + x_n) dx_1 \cdots dx_n = \sum_k \langle n \rangle_k \frac{f(k)}{n!}.$$

66. 证明 $\langle \frac{n}{1} \rangle = 2 \langle \frac{n}{1} \rangle$, 且求 $\langle \frac{n}{2} \rangle$ 的闭形式.

67. 求 $\sum_{k=1}^n k^2 H_{n+k}$ 的闭形式.

68. 指明习题 22 的广义调和数有幂级数展开

$$H_z = \sum_{n \geq 1} (-1)^n H_{\infty}^{(n)} z^{n-1}.$$

69. 通过考虑此无限乘积前 n 个因子的极限 (当 $n \rightarrow \infty$) 证明方程 (5.83) 的广义阶乘可记为

$$\prod_{k \geq 1} \left(1 + \frac{z}{k} \right) e^{-z/k} = \frac{e^{z^2}}{z!}.$$

指出 $(d/dz)(z!)$ 和习题 22 的一般调和数有关.

70. 证明正切函数有幂级数 (6.92), 且求 $z/\sin z$ 和 $\ln((\tan z)/z)$ 的对应的级数.

71. 求数 $T_n(1)$ 和 $1/\cos z$ 的系数之间的一个关系.

72. 具有交错正负号的欧拉三角形的行和 $\sum_k (-1)^k \langle \frac{n}{k} \rangle$ 是什么?

73. 证明, 对所有整数 $n \geq 1$,

$$z \cot z = \frac{z}{2^n} \cot \frac{z}{2^n} - \frac{z}{2^n} \tan \frac{z}{2^n} + \sum_{k=1}^{\infty} \frac{z}{2^n} \left(\cot \frac{z+k\pi}{2^n} + \cot \frac{z-k\pi}{2^n} \right),$$

且证明对固定的 k 当 $n \rightarrow \infty$ 时第 k 个被加数的极限是 $2z^2 / (z^2 - k^2 \pi^2)$.

74. 证明联系 Stirling 数, 伯努利数和 Cantalan 数的下列关系:

$$\sum_{k=0}^n \left\{ \begin{matrix} n+k \\ k \end{matrix} \right\} \binom{2n}{n+k} \frac{(-1)^k}{k+1} = B_n \binom{2n}{n} \frac{1}{n+1}.$$

75. 指明 $64=65$ 的谬论的 4 个棋盘块还能再装配来证明 $64=63$.

76. 由递归

$$A_1 = x, A_2 = y, A_n = A_{n-1} + A_{n-2}$$

定义的一个序列对某个 m 有 $A_m = 1\,000\,000$. 什么正整数 x 和 y 使 m 尽可能大?

77. 书中描述了一种方法把涉及 F_{n+k} 的一个公式改变成仅涉及 F_n 和 F_{n+1} 的一个公式. 所以我们自然想知道当在形式上它们不相同时, 是否两个这样的“简化”公式能相等. 设 $P(x, y)$ 是具有整系数的 x 和 y 的多项式, 找一个对所有 $n \geq 0$, $P(F_{n+1}, F_n) = 0$ 的必要和充分条件.

78. 说明如何加正整数, 在 Fibonacci 数系中完全行得通.

79. 如果 A_0 和 A_1 互素, 一个满足 Fibonacci 递归 $A_n = A_{n-1} + A_{n-2}$ 的序列 $\langle A_n \rangle$ 是否可能不包含素数?

80. 表明在矩阵乘积

$$\begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_n \end{pmatrix}$$

中以及在行列式

$$\det \begin{vmatrix} x_1 & 1 & 0 & 0 & \cdots & 0 \\ -1 & x_2 & 1 & 0 & & 0 \\ 0 & -1 & x_3 & 1 & & \vdots \\ \vdots & & -1 & & \ddots & \\ & & & & & 1 \\ 0 & 0 & \cdots & -1 & x_n \end{vmatrix}$$

中出现延拓多项式.

81. 当 α 是任何正无理数时, 推广式(6.146), 找一个与母函数 $\sum_{n \geq 1} z^{\lfloor n\alpha \rfloor}$ 有关系的连分式.

82. 设 m 和 n 是奇正整数, 求

$$S_{m,n}^+ = \sum_{k \geq 0} \frac{1}{F_{2mk+n} + F_m}, \quad S_{m,n}^- = \sum_{k \geq 0} \frac{1}{F_{2mk+n} - F_m}.$$

的闭形式。提示：习题 62 中的和是 $S_{1,3}^+ - S_{1,2n+3}^+$ 和 $S_{1,3}^- - S_{1,2n+3}^-$ 。

83. 设 α 是 $(0, 1)$ 中的无理数，且设 a_1, a_2, a_3, \dots 是它的连分式表示中的部分商，证明当 $n = K(a_1, \dots, a_n)$ 时 $|D(\alpha, n)| < 2$ ，其中 D 是第三章中定义的偏差。

84. 设 Q_n 是 Stern-Brocot 树的 n 级上的最大分母，(按照第四章中的图 $\langle Q_0, Q_1, Q_2, Q_3, Q_4, \dots \rangle = \langle 1, 2, 3, 5, 8, \dots \rangle$ 。) 证明 $Q_n = F_{n+2}$ 。

85. 表出所有 N 使得 Fibonacci 剩余

$$\{F_0 \bmod N, F_1 \bmod N, F_2 \bmod N, \dots\}$$

形成完全系 $\{0, 1, \dots, N-1\}$ 。(见习题 59。)

研究性问题

86. 把 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ 的定义推广到 n 和 k 的任意实值的最好方式是什么?

87. 如同习题 52 那样，将 H_n 如同 a_n / b_n 记成最小项。

(a) 是否有无限多个 n 具有 $11 \nmid a_n$?

(b) 是否有无限多个 n 具有 $b_n = \text{lcm}(1, 2, \dots, n)$? (两个这样的值是 $n=250$ 和 $n=1000$ 。)

88. 证明 γ 和 e^γ 是无理数。

89. 阐述两个参数递归

$$\begin{aligned} \left| \begin{smallmatrix} n \\ k \end{smallmatrix} \right| &= (\alpha n + \beta k + \gamma) \left| \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right| \\ &\quad + (\alpha' n + \beta' k + \gamma') \left| \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right| + [n=k=0], \quad n, k \geq 0, \end{aligned}$$

的解的一般理论，假设当 $n < 0$ 或 $k < 0$ 时 $\left| \begin{smallmatrix} n \\ k \end{smallmatrix} \right| = 0$ 。(二项系数，Stirling 数，欧拉数以及习题 17 和 31 的序列是特殊情形。) 利用表达的一般解，什么特殊值 $(\alpha, \beta, \gamma, \alpha', \beta', \gamma')$ 产生“基本解”?

第七章 母函数

就大家所知，处理数序列的最有力的方法是来操作“生成”这些序列的无穷级数，我们已学了一大堆序列，也看到少数几个母函数；现在我们准备深入探索母函数，将能看到它们是多么有用。

7.1 多米诺(骨牌)理论和兑换


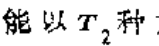
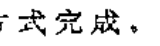
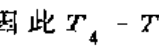
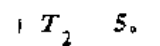
母函数是很重要的，但对大多数人来说尚无经验，当我们开始更深入查看它们时理应用一种轻松的探讨方式，所以让我们在本章开始用一些娱乐和游戏来尝试阐述母函数的直观知识。我们将研究概念的两种应用，一个涉及多米诺骨牌，另一个涉及硬币。

用 2×1 的多米诺骨牌完全覆盖一个 $2 \times n$ 矩形有多少种方式 T_n ？假设多米诺牌是相同的(或者因为它们面朝下的，或者因为某人已表示它们是不可区别的，比如说把它们全涂成红色)。因此仅仅是它们的方向，垂直和水平的确定问题，我们想象正用形状为多米诺的骨牌在处理。例如，一个 2×3 矩形有三种骨牌盖法，即 $\begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \end{smallmatrix}$ ， $\begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \end{smallmatrix}$ 和 $\begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \end{smallmatrix}$ ，所以 $T_3 = 3$ 。

为了找一般的 T_n 的闭形式，我们通常先查看小的情形。当 $n=1$ 时显然仅有一种盖法， $\begin{smallmatrix} \square \\ \square \end{smallmatrix}$ ；当 $n=2$ 时有两种盖法， $\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}$ 和 $\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}$ 。

关于 $n=0$ 的情形如何；一个 2×0 的矩形有多少种盖法？此问题的含意不是立即清楚的，但是看来相似于前面的情况：有零个元素的一个排列(即空排列)，所以 $0! = 1$ 。有一种方式从 n 个物体中选取零个物体(即不选取物体)，所以 $\binom{n}{0} = 1$ 。有一种方法把空集划分为零个非空子集，但是没有这样的方法来划分一个非空集，所以 $\left\{ \begin{smallmatrix} n \\ 0 \end{smallmatrix} \right\} = [n=0]$ 。依据这样的理由我们能推出用多米诺牌来盖一个 2×0 矩形仅有一种方式，即没有用多米诺牌，所以 $T_0 = 1$ 。(这损害了当 $n=1, 2$ 和 3 成立的简单型式 $T_n = n$ ，但是无论如何这种型式很可能注定要失败，因为按照情况的推理 T_0 应该为 1。)每当我们解一个枚举问题时，适当理解空的情形是有用的。

让我们再看一种小的情形， $n=4$ 。有两种可能的盖法来盖矩形的左面的边——在那里我们或者放一块垂直的多米诺牌或者两块水平的多米诺牌。如果选一块垂直的牌，不完全的解是 $\begin{smallmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \end{smallmatrix}$ ，且余下的 2×3 矩形能以 T_3 种方式覆盖。如果我们选取两个水平的牌，不

完全解  能以 T_2 种方式完成, 因此 $T_4 = T_3 + T_2 = 5$. (5 种盖法为 , ,  和 )

现在我们知道 T_n 的前 5 个值:

n	0	1	2	3	4
T_n	1	1	2	3	5

这些数猜想像 Fibonacci 数, 不难看出为什么是这样: 我们在建立 $T_4 = T_3 + T_2$ 时所用的理由容易推广到 $T_n = T_{n-1} + T_{n-2}$ ($n \geq 2$). 除了初始值 $T_0 = 1$ 和 $T_1 = 1$ 有点不同外, 这里我们有与 Fibonacci 数相同的递归. 但是这些初始值是相继的 Fibonacci 数 F_1 和 F_2 , 所以 T 就是向上移 1 位的 Fibonacci 数:

$$T_n = F_{n+1}, \quad n \geq 0.$$






(我们考虑这个为 T_n 的闭形式, 因为 Fibonacci 数是考虑为“已知的”很重要的数, 而且 F_n 本身有一个关于代数运算的闭形式(6.123).) 注意这个方程坚定了置 $T_0 = 1$ 的常识.

但是就母函数来说必须做什么呢? 我们即将接触到这一点——有另一种方式解决 T_n 是什么. 这个新方式基于一种醒目的主意. 让我们考虑所有可能的 $2 \times n$ 种盖法(所有 $n \geq 0$)的“和”, 且称它为 T :

$$T = 1 + \square + \square\square + \square\square\square + \square\square\square\square + \square\square\square\square\square + \dots \quad (7.1)$$

(左边第一项‘1’意味着一个 2×0 矩形的空盖法.) 这个和 T 代表一大堆信息. 它是有用的, 因为它使我们能够作为一个整体来证明 T 的结果, 而不是迫使我们对其个别项来证明它们(用归纳法).

这个和的项意味着盖法, 它是组合元素. 当无限多个盖法加在一起时, 我们不要过分注意它们的合法性; 本来样样事情都可以做得很严谨, 但是现在我们的目标是超出通常的代数公式来扩展我们的能动性.

我们可把型式加在一起, 还能通过毗连来乘它们. 例如, 可把盖法  和  相乘而取得新的盖法 , 但是注意乘法不可交换, 也就是说, 乘法的次序在考虑之列:  不同于 .

由此乘法概念不难看出, 空盖法起特殊的作用, 它是乘法单元. 例如, $1 \times \square = \square \times 1 = \square$.

现在我们能用多米诺算术来处理无限和 T :

$$\begin{aligned} T &= 1 + \square + \square\square + \square\square\square + \square\square\square\square + \dots \\ &= 1 + \square(1 + \square + \square\square + \square\square\square + \dots) + \square\square(1 + \square + \square\square + \dots) \\ &= 1 + \square T + \square\square T. \end{aligned} \quad (7.2)$$

右边中每一个成立的盖法恰好出现一次, 所以我们所做的是适当的, 即使忽略了第二章中关于“绝对收敛”的告诫. 此方程的最低一行告诉我们, T 中的一切或者是空盖法, 或者是一个垂直骨牌后面是 T 中的另外一些东西, 或者是两块水平骨牌后面是 T 中的另外一些

东西。

所以让我们试解 T 的方程。用 $|T$ 代替左边的 T 且在方程两边减去右边的最后两项，我们得到

$$(1 - \square - \square)T = 1. \quad (7.3)$$

为了一致性检验起见，这里是一种展开形式：

$$\begin{array}{r} 1 + \square + \square\square + \square\square\square + \square\square\square\square + \square\square\square\square\square + \square\square\square\square\square\square + \dots \\ - \square - \square\square - \square\square\square - \square\square\square\square - \square\square\square\square\square - \square\square\square\square\square\square - \dots \\ - \square\square - \square\square\square - \square\square\square\square - \square\square\square\square\square - \square\square\square\square\square\square - \square\square\square\square\square\square\square - \dots \\ \hline 1 \end{array}$$

顶上的每一项，除了第一项，均被第二行或第三行中的一项消去，所以我们的方程是正确的。

到目前为止，相当容易作出我们处理的方程的组合意义。然而，为了取得 T 的紧凑表达式，我们勾消了组合分配、跳出代数的约定用 $| - \square - \square$ 除方程(7.3)的两边取得

$$T = \frac{1}{1 - \square - \square}. \quad (7.4)$$

(乘法不是可变换的，所以在左右部分之间不加区别的情况下，我们差点受骗。在我们的应用中不碍事，因为 $|$ 和任何东西都可交换。但是让我们不过分讲究，除非我们的原始的主意导致谬论。)

下一步是用规则

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \dots$$

把这个分数展开成幂级数。空盖法 $|$ 是组合算术的乘法单元，它起到通常乘法单元 1 的作用； $\square + \square$ 扮演 z 的角色。所以我们得到展开式

$$\begin{aligned} \frac{1}{1 - \square - \square} &= 1 + (\square + \square) + (\square + \square)^2 + (\square + \square)^3 + \dots \\ &= 1 + (\square + \square) + (\square\square + \square\square + \square\square) \\ &\quad + (\square\square\square + \square\square\square + \square\square\square + \square\square\square + \square\square\square + \square\square\square) + \dots \end{aligned}$$

这是 T ，但是骨牌盖法以前面不同的次序排列。在此和中每一个盖法恰好出现一次。例如， $\square\square\square\square$ 出现在展开式 $(\square + \square)^4$ 中。

通过简缩此无限和，略去不感兴趣的细节可取得有用的信息。例如，我们能想象型式变成不胶合的，且个别的多米诺牌彼此交换，于是像 $\square\square\square\square$ 的一项变成 $\square^4 \square^6$ ，因为它包含 4 个垂直的和 6 个水平的骨牌。归并同类项给出级数

$$T = 1 + \square + \square^2 + \square^2 + \square^3 + 2\square\square^2 + \square^4 + 3\square^2\square^2 + \square^4 + \dots$$

这里的 $2\square\square^2$ 代表两项老的展开式 $\square\square\square$ 和 $\square\square\square$ ，它有一个垂直的和两个水平的多米诺

牌, 类似, $3\sqcup^2\sqcup^2$ 代表三项 $\begin{array}{|c|c|c|}\hline \hline \hline\end{array}$, $\begin{array}{|c|c|}\hline \hline \hline\end{array}$ 和 $\begin{array}{|c|}\hline \hline \hline\end{array}$ 。实质上我们把 \sqcup 和 \sqcup^2 看作为通常(交换的)变量。

用二项定理能找到 T 的交换形式中系数的一种闭形式:

$$\begin{aligned}\frac{1}{1-(\sqcup+\sqcup^2)} &= 1 + (\sqcup + \sqcup^2) + (\sqcup + \sqcup^2)^2 + (\sqcup + \sqcup^2)^3 + \cdots \\ &= \sum_{k \geq 0} (\sqcup + \sqcup^2)^k \\ &= \sum_{k \geq 0} \binom{k}{j} \sqcup^j \sqcup^{2k-2j} \\ &= \sum_{j, m \geq 0} \binom{j+m}{j} \sqcup^j \sqcup^{2m}\end{aligned}\quad (7.5)$$

(最后一步用 m 替换 $k-j$, 这是合法的, 因为当 $0 \leq k < j$ 时我们有 $\binom{k}{j} = 0$ 。)我们推出 $\binom{j+m}{j}$ 是用 j 个垂直多米诺牌和 $2m$ 个水平多米诺牌来盖一个 $2 \times (j+2m)$ 矩形的方式数。例如, 我们最近考虑的涉及 4 个垂直和 6 个水平多米诺牌的 2×10 盖法 $\begin{array}{|c|c|c|c|}\hline \hline \hline\end{array}$, 共有 $\binom{4+3}{4} = 35$ 种这样的盖法, 所以 T 的交换形式中的一项是 $35\sqcup^4\sqcup^6$ 。

我们甚至能通过忽略多米诺牌的方位来消除一部分。假设我们不关心水平/垂直的分类, 仅要知道 $2 \times n$ 盖法的总数。(事实上, 这就是我们着手发现的数 T_n 。)对于 \sqcup 和 \sqcup^2 , 通过简单地代替一个量 z , 我们能归并必要的信息。同样我们也能用 1 替换 \sqcup , 取得

$$T = \frac{1}{1-z-z^2} \quad (7.6)$$

除了分子中 z 的一个失去的因子, 这是 Fibonacci 数的母函数(6.117), 所以我们推出 T 中 z^n 的系数是 F_{n+1} 。

关于 T 所推演的紧凑的表示 $1/(1-\sqcup-\sqcup^2)$, $1/(1-\sqcup-\sqcup^2)$ 和 $1/(1-z-z^2)$ 称为母函数, 因为它们产生有趣的系数。

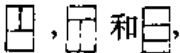
顺便提一句, 我们的推导意味着恰好有 m 对水平多米诺牌的 $2 \times n$ 多米诺盖法数为 $\binom{n-m}{m}$ 。(因为有 $j=n-2m$ 个垂直多米诺牌, 所以这是必然的结果, 按照我们的公式有

$$\binom{j+m}{j} = \binom{j+m}{m} = \binom{n-m}{m}$$

种方式来覆盖。在第六章中我们看到 $\binom{n-m}{m}$ 是长度 n 的莫尔斯码序列包含 m 个长划的数目; 事实上, 易见 $2 \times n$ 多米诺牌盖法直接对应于莫尔斯码序列。(盖法 $\begin{array}{|c|c|c|c|}\hline \hline \hline\end{array}$ 对应于 ‘·—·—·—’。)因此, 多米诺牌盖法与我们在第六章中研究的延拓多项式紧密相联。世界太小了。





我们已经以两种方式来解决 T_n 问题。第一种猜测解答且用归纳法证明它的方法是较容易的；第二种方法，用多米诺牌型式的无限和，且提取有趣的系数，是行家的方法。但是用第二种方法是否仅因为用多米诺牌作游戏感觉有趣，好像它们是代数变量？不是，引入第二种方法的实际原因是由于无限和的方法是非常有力的方法。第二种方法应用于许多问题，因为它不要求我们作不可思议的猜测。

让我们推广到高一级的问题，对于此问题我们不能作猜测。用多米诺牌盖一个 $3 \times n$ 矩形有多少种方式 U_n ？


这个问题的前几种情形告诉我们一点：空盖给出 $U_0 = 1$ 。当 $n=1$ 时没有有效的盖法，因为一个 2×1 多米诺牌填不满一个 3×1 矩形，而且没有 2 个的空间。下一个情形， $n=2$ ，用手容易作出；有三种盖法，，所以 $U_2 = 3$ 。（考虑一下我们已经知道的这一点，因为前面的问题告诉我们 $T_3 = 3$ ；盖一个 3×2 矩形的方法数与盖一个 2×3 的数目相同。）当 $n=3$ 时，如同 $n=1$ ，没有盖法。通过快速彻底的查找或通过从较高级来考虑问题，我们能确信这一点：一个 3×3 矩形的面积是奇数，所以不可能用面积为偶的多米诺牌来盖它。（同样的论证显然应用于任何奇 n 。）最后，当 $n=4$ 时，看来约有 12 个盖法，不花费许多时间很难保证表是完全的而肯定确切的数目。

所以让我们尝试上次所用的无限和方法：

$$U = 1 + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \dots \quad (7.7)$$

每一个非空盖法从  或  或  开始，但不幸的是，这三种可能的前两种不是简单的提因子且再次把 U 留给我们。然而， U 中所有以  开始的项可记为 $\begin{array}{|c|} \hline \text{ } \\ \hline \end{array} V$ ，其中

$$V = \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \dots$$

是所有多米诺牌盖法缺左下角的一个残缺的 $3 \times n$ 矩形。类似，以  开始的 U 的项可记为 $\begin{array}{|c|} \hline \text{ } \\ \hline \end{array} \Lambda$ ，其中

$$\Lambda = \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} + \dots$$

包含所有缺左上角的矩形盖法，级数 Λ 是 V 的镜像。这样的分解允许我们记

$$U = 1 + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} V + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} \Lambda + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U$$

且能同样分解 V 和 Λ ，因为这样的盖法仅以两种方式开始：

$$V = \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} V;$$

$$\Lambda = \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} \Lambda.$$

现在我们有三个未知量 (U 、 V 和 Λ) 的三个方程。我们首先解 V 和 Λ 用 U 来表示，然后把结果代入 U 的方程：

$$V = (1 - \begin{array}{|c|} \hline \text{ } \\ \hline \end{array})^{-1} \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U, \quad \Lambda = (1 - \begin{array}{|c|} \hline \text{ } \\ \hline \end{array})^{-1} \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U;$$

$$U = 1 + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} (1 - \begin{array}{|c|} \hline \text{ } \\ \hline \end{array})^{-1} \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} (1 - \begin{array}{|c|} \hline \text{ } \\ \hline \end{array})^{-1} \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U + \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} U.$$

最后的方程能解出 U ，给出紧凑公式。

$$U = \frac{1}{1 - \frac{1}{2}(1 - \frac{1}{2})^{-1} + \frac{1}{2}(1 - \frac{1}{2})^{-1} - \frac{1}{2}}. \quad (7.8)$$

此表达式确定了无限和 U ，就像式(7.4)确定了 T 。

下一步是进行交换。当我们拆开所有多米诺牌且仅用 \square 和 \sqcap 的幂而极好地简化：

$$\begin{aligned} U &= \frac{1}{1 - \square^2 \sqcap (1 - \sqcap^3)^{-1} - \square^2 \sqcap (1 - \sqcap^3)^{-1} - \sqcap^3} \\ &= \frac{1 - \sqcap^3}{(1 - \sqcap^3)^2 - 2 \square^2 \sqcap} \\ &= \frac{(1 - \sqcap^3)^{-1}}{1 - 2 \square^2 \sqcap (1 - \sqcap^3)^{-2}} \\ &= \frac{1}{1 - \sqcap^3} + \frac{2 \square^2 \sqcap}{(1 - \sqcap^3)^3} + \frac{4 \square^4 \sqcap^2}{(1 - \sqcap^3)^5} + \frac{8 \square^6 \sqcap^3}{(1 - \sqcap^3)^7} + \dots \\ &= \sum_{k \geq 0} \frac{2^k \square^{2k} \sqcap^k}{(1 - \sqcap^3)^{2k+1}} \\ &= \sum_{k, m \geq 0} \binom{m+2k}{m} 2^k \square^{2k} \sqcap^{k+3m}, \end{aligned}$$

(此推导值得仔细研究，最后一步用了公式 $(1-w)^{-2k-1} = \sum_m \binom{m+2k}{m} w^m$ ，等式(5.56)。)让我们再来看看最底下一行说明什么。首先，它说明每一个 $3 \times n$ 盖法用了一种偶数个垂直多米诺牌。如果有 $2k$ 个垂直的，则一定至少有 k 个水平的，且对某个 $m \geq 0$ 水平的总数一定是 $k+3m$ 。最后，用 $2k$ 个垂直和 $k+3m$ 个水平的可能盖法数恰好是 $\binom{m+2k}{m} 2^k$ 。

现在我们能分析当开始考虑 $3 \times n$ 问题时所留下不确定的 3×4 盖法。当 $n=4$ 时总面积为 12，所以总共需要 6 个多米诺牌。对于某个 k 和 m ，有 $2k$ 个垂直和 $k+3m$ 个水平的牌，因此 $2k+k+3m=6$ 。换句话说， $k+m=2$ 。如果我们不用垂直的，对于 $k=0$ 和 $m=2$ ，可能的数目是 $\binom{2+0}{2} 2^0 = 1$ 。(这就是计算出的盖法 $\begin{smallmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{smallmatrix}$ 。)如果用 2 个垂直的牌，对于 $k=1$ 和 $m=1$ ，有 $\binom{1+2}{1} 2^1 = 6$ 个这样的盖法。如果用 4 个垂直的牌，对于 $k=2$ 和 $m=0$ ，有 $\binom{0+4}{0} 2^2 = 4$ 个这样的盖法，得到总数 $U_4 = 11$ 。一般，如果 n 是偶数，则这个推理表明 $k+m = \frac{1}{2}n$ ，因此 $\binom{m+2k}{m} = \binom{n/2+k}{n/2-k}$ ，且 $3 \times n$ 盖法的总数是

$$U_n = \sum_k \binom{n/2+k}{n/2-k} 2^k = \sum_m \binom{n-m}{m} 2^{n/2-m} \quad (7.9)$$

如同前面那样, 我们也能对于 $\textcircled{1}$ 和 $\textcircled{4}$ 用 z 代替, 取得在特殊种类的多米诺牌之间不加区别的一个母函数。结果是

$$U = \frac{1}{1 - z^3(1 - z^3)^{-1} - z^3(1 - z^3)^{-1} - z^3} = \frac{1 - z^3}{1 - 4z^3 + z^6}. \quad (7.10)$$

如果把这个商展成幂级数, 我们得到

$$U = 1 + U_2 z^3 + U_4 z^6 + U_6 z^9 + U_8 z^{12} + \cdots,$$

数 U_n 的一个母函数。(在此公式中足标和指数之间有一种难以理解的配合不当, 但是这是容易说明的。例如, z^9 的系数是 U_6 , 它计算了 3×6 矩形的盖法。这就是我们所要的, 因为每一个这样的盖法包含 9 个多米诺牌。)

我们能继续分析式(7.10)且取得系数的一个闭形式, 但是在取得较多经验后再在本章后面分析较好。所以让我们暂时不谈多米诺牌, 进行下一个典型的问题, “兑换”。

有多少种方式来付 50 分? 我们假设一定要用 1 分 $\textcircled{1}$, 5 分 $\textcircled{5}$, 1 角 $\textcircled{10}$, 2 角 5 分 $\textcircled{25}$ 和 0.5 元 $\textcircled{50}$ 。George Polya^[239]以一种启发的方法用母函数解这个问题来推广它。

让我们建立无限和来表示所有可能给出兑换的方式, 就像用无限和表示所有可能的多米诺牌型式来处理多米诺牌问题。开始用较少品种的硬币是最简单的, 所以首先假设我们仅有 1 分硬币。在兑换中, 交某个 1 分硬币的个数的所有方式之和可记为

$$\begin{aligned} P &= 1 + \textcircled{1} + \textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{1}\textcircled{1}\textcircled{1} + \cdots \\ &= 1 + \textcircled{1} + \textcircled{1}^2 + \textcircled{1}^3 + \textcircled{1}^4 + \cdots \end{aligned}$$

第一项表示没有交 1 分硬币的方式, 第二项表示 1 个 1 分, 然后 2 个 1 分, 3 个 1 分, 等等。现在如果允许用 1 分和 5 分, 则所有可能方式之和是

$$\begin{aligned} N &= P + \textcircled{5}P + \textcircled{5}\textcircled{5}P + \textcircled{5}\textcircled{5}\textcircled{5}P + \textcircled{5}\textcircled{5}\textcircled{5}\textcircled{5}P + \cdots \\ &= (1 + \textcircled{5} + \textcircled{5}^2 + \textcircled{5}^3 + \textcircled{5}^4 + \cdots)P, \end{aligned}$$

由于每种支付有一个选自第一个因子的确定的 5 分的个数和一个选自 P 的确定的 1 分的个数。(注意, N 不是和 $1 + \textcircled{1} + \textcircled{5} + (\textcircled{1} + \textcircled{5})^2 + (\textcircled{1} + \textcircled{5})^3 + \cdots$, 因为这样一个和包含许多类型的支付多于一次。例如, 项 $(\textcircled{1} + \textcircled{5})^2 = \textcircled{1}\textcircled{1} + \textcircled{1}\textcircled{5} + \textcircled{5}\textcircled{1} + \textcircled{5}\textcircled{5}$ 看作 $\textcircled{1}\textcircled{5}$ 和 $\textcircled{5}\textcircled{1}$ 好像是不同的, 但是我们要列出的每个硬币集合仅仅一次不考虑次序。)

类似, 如果同样允许用 1 角, 我们得到无限和

$$D = (10 + \textcircled{10} + \textcircled{10}^2 + \textcircled{10}^3 + \textcircled{10}^4 + \cdots)N,$$

当全部展开它时, 包含像 $\textcircled{10}^3 \textcircled{5}^3 \textcircled{1}^5 = \textcircled{10}\textcircled{10}\textcircled{10}\textcircled{5}\textcircled{5}\textcircled{5}\textcircled{1}\textcircled{1}\textcircled{1}\textcircled{1}\textcircled{1}$ 的项。每个这样的项是作兑换的一个不同方式。加入 2 角 5 分, 然后加入 0.5 元, 可能给出

$$Q = (25 + \textcircled{25} + \textcircled{25}^2 + \textcircled{25}^3 + \textcircled{25}^4 + \cdots)D;$$

$$C = (50 + \textcircled{50} + \textcircled{50}^2 + \textcircled{50}^3 + \textcircled{50}^4 + \cdots)Q.$$

我们的问题是找 C 中恰好价值为 50 分的项数。

令人满意地解此问题的简单诀窍是：用 z 替换①，用 z^5 替换⑤，用 z^{10} 替换⑩，用 z^{25} 替换②⑤和用 z^{50} 替换⑤⑩。于是每项用 z^n 来替换，其中 n 是原来项的钱值。例如，项⑤⑩⑤⑤①变成 $z^{50+10+5+5+1} = z^{71}$ 。支付 13 分的 4 种方式，即⑩①³，⑤①³，⑤²①³和①¹³，每个化成 z^{13} ；因此作 z 的替代之后 z^{13} 的系数将是 4。

设 P_n ， N_n ， D_n ， Q_n 和 C_n 是付 n 分的方式数，且分别允许至多用硬币值 1，5，10，25 和 50 分，分析告诉我们，这些方式数是各自的幂级数

$$\begin{aligned} P &= 1 + z + z^2 + z^3 + z^4 + \cdots, \\ N &= (1 + z^5 + z^{10} + z^{15} + z^{20} + \cdots)P, \\ D &= (1 + z^{10} + z^{20} + z^{30} + z^{40} + \cdots)N, \\ Q &= (1 + z^{25} + z^{50} + z^{75} + z^{100} + \cdots)D, \\ C &= (1 + z^{50} + z^{100} + z^{150} + z^{200} + \cdots)Q, \end{aligned}$$

中 z^n 的系数。显然，对所有 $n \geq 0$ $P_n = 1$ 。稍加思索证明我们可得到 $N_n = \lfloor n/5 \rfloor + 1$ ；为了使 n 分出自 1 分和 5 分，我们一定要选取 0 或 1 或 \cdots 或 $\lfloor n/5 \rfloor$ 个 5 分，之后仅有一种方式提供需要的 1 分的个数。因此 P_n 和 N_n 是简单的，但是 D_n ， Q_n 和 C_n 的值变得较复杂。

讨论这些公式的一种方式了解 $1 + z^m + z^{2m} + \cdots$ 就是 $1/(1 - z^m)$ 。因此我们能记

$$\begin{aligned} P &= 1/(1 - z), \\ N &= P/(1 - z^5), \\ D &= N/(1 - z^{10}), \\ Q &= D/(1 - z^{25}), \\ C &= Q/(1 - z^{50}). \end{aligned}$$

两边乘以分母，我们得到

$$\begin{aligned} (1 - z)P &= 1, \\ (1 - z^5)N &= P, \\ (1 - z^{10})D &= N, \\ (1 - z^{25})Q &= D, \\ (1 - z^{50})C &= Q. \end{aligned}$$

现在我们能在这组方程中使 z^n 的系数相等，取得递归关系，从而很快计算出希望的系数：

$$\begin{aligned} P_n &= P_{n-1} + [n=0], \\ N_n &= N_{n-5} + P_n, \\ D_n &= D_{n-10} + N_n, \end{aligned}$$

$$Q_n = Q_{n-25} + D_n,$$

$$C_n = C_{n-50} + Q_n.$$

例如, 在 $D = (1 - z^{25})Q$ 中 z^n 的系数等于 $Q_n - Q_{n-25}$; 所以像要求的那样, 一定有 $Q_n - Q_{n-25} = D_n$.

我们能展开这些递归, 且找到, 例如, $Q_n = D_n + D_{n-25} + D_{n-50} + D_{n-75} + \dots$, 当足标取负时终止. 但是非重复形式是方便的, 如同在 Pascal 三角形中那样, 仅用一次加法来计算每个系数.

让我们用递归来找 C_{50} . 首先, $C_{50} = C_0 + Q_{50}$, 我们要知道 Q_{50} . 然后 $Q_{50} = Q_{25} + D_{50}$, $Q_{25} = Q_0 + D_{25}$, 我们还要知道 D_{50} 和 D_{25} . 这些 D_n 依次依赖于 D_{40} , D_{30} , D_{20} , D_{15} , D_{10} , D_5 , 以及 N_{50} , N_{45} , \dots , N_5 . 所以一次简单计算足以确定所有必要的系数:

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

表中最后的值给出解答 C_{50} : 恰好有 50 种方式交 50 分的小费.

关于 C_n 的一个闭形式怎么样? 方程乘在一起给出紧凑表达式

$$C = \frac{1}{1-z} \frac{1}{1-z^5} \frac{1}{1-z^{10}} \frac{1}{1-z^{25}} \frac{1}{1-z^{50}}, \quad (7.11)$$

但是如何由此取得 z^n 的系数不明显. 幸好有一种方法, 我们在本章后面将回到这个问题.

如果考虑给定兑换的问题. 当我们生活在一个地方, 能铸造每一个正整数单位的硬币 (①, ②, ③, \dots), 而不是以前仅允许的 5 种. 对应的母函数是一个无限的分式乘积,

$$\frac{1}{(1-z)(1-z^2)(1-z^3)\cdots},$$

且当完全乘出这些因子时 z^n 的系数称为 n 的划分数 $p(n)$. n 的一个划分是把 n 表为正整数的一个和, 不考虑次序. 例如, 有 7 种不同的 5 的划分, 即,

$$5 = 4+1 = 3+2 = 3+1+1 = 2+2+1 = 2+1+1+1 = 1+1+1+1+1,$$

因此 $p(5)=7$. (还有 $p(2)=2$, $p(3)=3$, $p(4)=5$ 和 $p(6)=11$, 开始看来好像 $p(n)$ 总是素数. 但是 $p(7)=15$, 型式变了.) 没有 $p(n)$ 的闭形式, 但是划分理论是数学的一个吸引人的分支, 在此分支中已有许多值得注意的发现. 例如, Ramanujan 通过作母函数的精巧的变换, 证明了 $p(5n+4) \equiv 0 \pmod{5}$, $p(7n+5) \equiv 0 \pmod{7}$ 和 $p(11n+6) \equiv 0 \pmod{11}$ (见 Andrews[11, 第 10 章]).

7.2 基本操作

现在让我们仔细考虑一些使幂级数更有力的技巧。

首先是几个术语和表示法。我们的一般母函数有形式

$$G(z) = g_0 + g_1 z + g_2 z^2 + \cdots = \sum_{n \geq 0} g_n z^n, \quad (7.12)$$

且称 $G(z)$ 是序列 $\langle g_0, g_1, g_2, \cdots \rangle$ (也称它为 $\langle g_n \rangle$) 的母函数。 $G(z)$ 中 z^n 的系数 g_n 有时记为 $[z^n]G(z)$ 。

式(7.12)中的和是对所有 $n \geq 0$ 求的, 但是常常发现和推广到对所有整数求更方便。简单地考虑 $g_{-1} = g_{-2} = \cdots = 0$ 是能做到的, 在这样的情形我们仍可谈序列 $\langle g_0, g_1, g_2, \cdots \rangle$, 好像负的 n , g_n 不存在。

当处理母函数时, 两种“闭形式”出现。依据 z 表达, 可有 $G(z)$ 的闭形式; 或者依据 n 表达, 可有 g_n 的一个闭形式。例如, Fibonacci 数的母函数有闭形式 $z/(1-z-z^2)$, Fibonacci 数有闭形式 $(\varphi^n - \hat{\varphi}^n)/\sqrt{5}$ 。这里将说明哪一类闭形式是有意义的。

现在谈一点关于观点的话: 母函数 $G(z)$ 呈现为两种不同的实体, 依赖于我们如何看待它。有时它是一个复变数 z 的函数, 满足微积分书中证明的所有标准的性质。有时它简单地是一个形式的幂级数, z 作为一个位置的占有者, 在前节中, 例如, 我们用了第二种解释; 我们看到几个例子, 在例子中, 对于组合元素的“和”中元素的某个特性用 z 替代。于是 z^n 的系数是这种特性出现 n 次的组合元素数。

当把 $G(z)$ 视为一个复变数的函数时, 它的收敛变成一个问题。在第二章中说过, 无穷级数 $\sum_{n \geq 0} g_n z^n$ (绝对) 收敛当且仅当对任何 N 有一个有界常数 A 使得有限和 $\sum_{0 \leq n \leq N} |g_n z^n|$ 不超过 A , 所以易见如果对于某个值 $z = z_0$, $\sum_{n \geq 0} g_n z^n$ 收敛, 则对于所有 $|z| < |z_0|$ 的 z 它亦收敛。而且一定有 $\lim_{n \rightarrow \infty} |g_n z_0^n| = 0$; 因此, 按照第九章的记法, 如果在 z_0 处收敛, $g_n = O(|1/z_0|^n)$ 。反之, 如果 $g_n = O(M^n)$, 则对所有 $|z| < 1/M$, 级数 $\sum_{n \geq 0} g_n z^n$ 收敛, 这些是幂级数收敛的基本事实。

但是对于我们的目的收敛通常是不相干的东西, 除非尝试研究系数的渐近状态。像在正规幂级数一个运算是合法的那样, 几乎我们在母函数上执行的每一个运算都能严格证明, 甚至当级数不收敛时这样的运算也是合法的。(例如, 在 Bell[19], Niven[225] 和 Henrici[15], 第 1 章]中找到有关的理论。)

此外, 如果不再考虑所有的告诫且不加严格证明地推出公式, 则我们一般能取得推导的结果且用归纳法证明它们。例如, 仅当 $|z| < 1/\varphi \approx 0.618$ 时 Fibonacci 数的母函数收敛, 但是当证明公式 $F_n = (\varphi^n - \hat{\varphi}^n)/\sqrt{5}$ 时我们不需知道这一点。如果不相信形式幂级

数的理论, 后面的公式一旦发现, 能直接验证。所以在本章中将忽略收敛的问题, 因为它有碍于问题的讨论。

关于观点就讲那么多。接着来研究给母函数以新形式的主要工具, 加, 移位, 改变变量, 微分, 积分和乘, 除非另外声明, 以下内容中假设 $F(z)$ 和 $G(z)$ 是序列 $\langle f_n \rangle$ 和 $\langle g_n \rangle$ 的母函数。我们还假设对负的 n , f_n 和 g_n 是零, 因为这省去了一些关于求和界限的争论。

当我们把常数乘 F 和 G 再加在一起时, 结果十分明显:

$$\alpha F(z) + \beta G(z) = \alpha \sum_n f_n z^n + \beta \sum_n g_n z^n = \sum_n (\alpha f_n + \beta g_n) z^n. \quad (7.13)$$

这给出了序列 $\langle \alpha f_n + \beta g_n \rangle$ 的母函数。

母函数移位不十分难, 为了把 $G(z)$ 向右移 m 个位置, 也就是说, 形成前面有 m 个 0 的序列 $\langle 0, \dots, 0, g_0, g_1, \dots \rangle = \langle g_{n-m} \rangle$ 的母函数, 我们简单地用 z^m 乘:

$$z^m G(z) = \sum_n g_n z^{n+m} = \sum_n g_{n-m} z^n, \text{ 整数 } m \geq 0. \quad (7.14)$$

这是我们所用的运算(2次), 与加法一起推演方程 $(1 - z - z^2)F(z) = z$ 来找出第六章中 Fibonacci 数的一个闭形式。

为了把 $G(z)$ 向左移 m 个位置, 也就是说形成丢弃前 m 个元素的序列 $\langle g_m, g_{m+1}, g_{m+2}, \dots \rangle = \langle g_{n+m} \rangle$ 的母函数, 减掉前 m 项, 然后用 z^m 除:

$$\frac{G(z) - g_0 - g_1 z - \dots - g_{m-1} z^{m-1}}{z^m} = \sum_{n \geq m} g_n z^{n-m} = \sum_{n \geq 0} g_{n+m} z^n \quad (7.15)$$

(除非 $g_0 = \dots = g_{m-1} = 0$, 我们不能把最后这个和延伸在所有 n 上。)

用一个常数倍数替代 z 是另一种诀窍:

$$G(cz) = \sum_n g_n (cz)^n = \sum_n c^n g_n z^n \quad (7.16)$$

这产生了序列 $\langle c^n g_n \rangle$ 的母函数。 $c = -1$ 的特殊情形特别有用。

我们常常要把 n 的一个因子写入系数, 微分可使我们做到这一点:

$$G'(z) = g_1 + 2g_2 z + 3g_3 z^2 + \dots = \sum_n (n+1) g_{n+1} z^n \quad (7.17)$$

把它右移一位, 可给出一种有时较有用的形式,

$$zG'(z) = \sum_n n g_n z^n \quad (7.18)$$

这是序列 $\langle ng_n \rangle$ 的母函数。重复微分将允许我们用任何希望的 n 的多项式乘 g_n 。

逆运算积分，让我们用 n 除项：

$$\int_0^x G(t) dt = g_0 z + \frac{1}{2} g_1 z^2 + \frac{1}{3} g_2 z^3 + \cdots = \sum_{n \geq 1} \frac{1}{n} g_{n-1} z^n. \quad (7.19)$$

(注意常数项为零。) 如果我们要 $\langle g_n/n \rangle$ 的母函数，而不是 $\langle g_{n-1}/n \rangle$ 的母函数，应先向左移一个位置，在积分中用 $(G(t) - g_0)/t$ 替换 $G(t)$ 。

最后，我们说明如何把母函数乘在一起：

$$\begin{aligned} F(z)G(z) &= (f_0 + f_1 z + f_2 z^2 + \cdots)(g_0 + g_1 z + g_2 z^2 + \cdots) \\ &= (f_0 g_0) + (f_0 g_1 + f_1 g_0)z + (f_0 g_2 + f_1 g_1 + f_2 g_0)z^2 + \cdots \\ &= \sum_n \left(\sum_k f_k g_{n-k} \right) z^n. \end{aligned} \quad (7.20)$$

如同在第 7 章中见到的那样，这给出 $\langle f_n \rangle$ 和 $\langle g_n \rangle$ 的卷积 $\langle h_n \rangle$ 序列的母函数。和 $h_n = \sum_k f_k g_{n-k}$ 也能记为 $h_n = \sum_{k=0}^n f_k g_{n-k}$ ，因为当 $k < 0$ 时 $f_k = 0$ ，当 $k > n$ 时 $g_{n-k} = 0$ 。乘/卷积比其他运算复杂一点，但是它十分有用，在下面的 7.5 节我们将看到极为有用的例子。

乘法有几个特殊情形作为运算是值得研究的，我们已看到一个这样的运算：当 $F(z) = z^m$ 时，我们取得移位运算(7.14)。此时和 h_n 变成单项 g_{n-m} ，因为所有 f_k 为 0，除了 $f_m = 1$ 。

当 $F(z)$ 是熟悉的函数 $1/(1-z) = 1 + z + z^2 + \cdots$ 时，另一个有用的特殊情形产生：所有 $f_k (k \geq 0)$ 为 1 且我们得到重要公式：

$$\frac{1}{1-z} G(z) = \sum_n \left(\sum_{k \geq 0} g_{n-k} \right) z^n = \sum_n \left(\sum_{k \leq n} g_k \right) z^n. \quad (7.21)$$

用 $1/(1-z)$ 乘母函数可给出原来序列的累积的和的母函数。

表 7.1 总结了至今我们讨论的运算。为了用所有这些有效的操作来帮助得到库存的相当多的母函数的组成部分，表 7.2 列出了最简的序列与母函数，我们可用这些作为出发点来解相当多的问题。

表 7.2 中的每一个母函数都是十分重要的，应该记住，其中许多母函数是其他母函数的特殊情形，用表 7.1 的基本运算能依据其他母函数快速推出其中许多母函数，所以记住它们并不十分难。

例如，让我们考虑序列 $\langle 1, 2, 3, 4, \cdots \rangle$ ，它的母函数 $1/(1-z)^2$ 常常是有用的。

此母函数在表 7.2 的中间出现, 它也是 $\langle 1, \binom{m+1}{m}, \binom{m+2}{m}, \binom{m+3}{m}, \dots \rangle$ 的特殊情形 $m=1$, 它出现在更下面. 它还是紧密相关的序列 $\langle 1, c, \binom{c+1}{2}, \binom{c+2}{3}, \dots \rangle$ 的特殊情形 $c=2$. 通过如同式(7.21)中取累积和我们能从 $\langle 1, 1, 1, 1, \dots \rangle$ 的母函数导出它, 也就是说用 $(1-z)$ 除 $1/(1-z)$. 用式(7.17)通过微分我们能从 $\langle 1, 1, 1, 1, \dots \rangle$ 导出它.

表 7.1 母函数操作

$$\begin{aligned}
 \alpha F(z) + \beta G(z) &= \sum_n (\alpha f_n + \beta g_n) z^n \\
 z^m G(z) &= \sum_n g_{n-m} z^n, \text{ 整数 } m \geq 0 \\
 \frac{G(z) - g_0 - g_1 z - \dots - g_{m-1} z^{m-1}}{z^m} &= \sum_{n \geq 0} g_{n+m} z^n, \text{ 整数 } m \geq 0 \\
 G(cz) &= \sum_n c^n g_n z^n \\
 G'(z) &= \sum_n (n+1) g_{n+1} z^n \\
 zG'(z) &= \sum_n n g_n z^n \\
 \int_0^1 G(t) dt &= \sum_{n \geq 1} \frac{1}{n} g_{n-1} z^n \\
 F(z)G(z) &= \sum_n \left(\sum_k f_k g_{n-k} \right) z^n \\
 \frac{1}{1-z} G(z) &= \sum_n \left(\sum_{k \leq n} g_k \right) z^n
 \end{aligned}$$

序列 $\langle 1, 0, 1, 0, \dots \rangle$ 是另一个以许多方法获得它的母函数的序列. 通过在等式 $\sum_n z^n = 1/(1-z)$ 中的 z 用 z^2 替代我们显然能导出公式 $\sum_n z^{2n} = 1/(1-z^2)$; 我们还能对序列 $\langle 1, -1, 1, -1, \dots \rangle$ 应用累积和, 它的母函数为 $1/(1+z)$, 取得 $1/(1+z)(1-z) = 1/(1-z^2)$. 还有一个第三种方法, 它是基于任何给定序列抽取偶编号的项 $\langle g_0, 0, g_2, 0, g_4, 0, \dots \rangle$ 的一般方法: 如果我们把 $G(-z)$ 加到 $G(z)$, 可得:

$$G(z) + G(-z) = \sum_n g_n (1 + (-1)^n) z^n = 2 \sum_n g_n [n \text{ 偶}] z^n,$$

所以

$$\frac{G(z) + G(-z)}{2} = \sum_n g_{2n} z^{2n} \quad (7.22)$$

表 7.2 简单序列和它们的母函数

序列	母函数	闭形式
$\langle 1, 0, 0, 0, 0, 0, \dots \rangle$	$\sum_{n \geq 0} [n=0] z^n$	1
$\langle 0, \dots, 0, 1, 0, 0, \dots \rangle$	$\sum_{n \geq 0} [n=m] z^n$	z^m
$\langle 1, 1, 1, 1, 1, 1, \dots \rangle$	$\sum_{n \geq 0} z^n$	$\frac{1}{1-z}$
$\langle 1, -1, 1, -1, 1, -1, \dots \rangle$	$\sum_{n \geq 0} (-1)^n z^n$	$\frac{1}{1+z}$
$\langle 1, 0, 1, 0, 1, 0, \dots \rangle$	$\sum_{n \geq 0} [2 \nmid n] z^n$	$\frac{1}{1-z^2}$
$\langle 1, 0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots \rangle$	$\sum_{n \geq 0} [m \nmid n] z^n$	$\frac{1}{1-z^m}$
$\langle 1, 2, 3, 4, 5, 6, \dots \rangle$	$\sum_{n \geq 0} (n+1) z^n$	$\frac{1}{(1-z)^2}$
$\langle 1, 2, 4, 8, 16, 32, \dots \rangle$	$\sum_{n \geq 0} 2^n z^n$	$\frac{1}{1-2z}$
$\langle 1, 4, 6, 4, 1, 0, 0, \dots \rangle$	$\sum_{n \geq 0} \binom{4}{n} z^n$	$(1+z)^4$
$\langle 1, c, \binom{c}{2}, \binom{c}{3}, \dots \rangle$	$\sum_{n \geq 0} \binom{c}{n} z^n$	$(1+z)^c$
$\langle 1, c, \binom{c+1}{2}, \binom{c+2}{3}, \dots \rangle$	$\sum_{n \geq 0} \binom{c+n-1}{n} z^n$	$\frac{1}{(1-z)^c}$
$\langle 1, c, c^2, c^3, \dots \rangle$	$\sum_{n \geq 0} c^n z^n$	$\frac{1}{1-cz}$
$\langle 1, \binom{m+1}{m}, \binom{m+2}{m}, \binom{m+3}{m}, \dots \rangle$	$\sum_{n \geq 0} \binom{m+n}{m} z^n$	$\frac{1}{(1-z)^{m+1}}$
$\langle 0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots \rangle$	$\sum_{n \geq 1} \frac{1}{n} z^n$	$\ln \frac{1}{1-z}$
$\langle 0, 1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \dots \rangle$	$\sum_{n \geq 1} \frac{(-1)^{n+1}}{n} z^n$	$\ln(1+z)$
$\langle 1, 1, \frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \frac{1}{120}, \dots \rangle$	$\sum_{n \geq 0} \frac{1}{n!} z^n$	e^z

以相似方式能抽出奇编号项,

$$\frac{G(z) - G(-z)}{2} = \sum_n g_{2n+1} z^{2n+1}. \quad (7.23)$$

在特殊情形 $g_n = 1$ 和 $G(z) = 1/(1-z)$, $\langle 1, 0, 1, 0, \dots \rangle$ 的母函数是 $(1/2)(G(z) + G(-z)) = (1/2)(1/(1-z) + 1/(1+z)) = 1/(1-z^2)$.

让我们就 Fibonacci 数的母函数尝试抽取诀窍。我们知道 $\sum_n F_n z^n = z/(1-z-z^2)$, 因此

$$\begin{aligned} \sum_n F_{2n} z^{2n} &= \frac{1}{2} \left(\frac{z}{1-z-z^2} + \frac{-z}{1+z-z^2} \right) \\ &= \frac{1}{2} \left(\frac{z + z^2 - z^3 - z + z^2 + z^3}{(1-z^2)^2 - z^2} \right) = \frac{z^2}{1-3z^2+z^4}. \end{aligned}$$

这产生序列 $\langle F_0, 0, F_2, 0, F_4, \dots \rangle$, 因此交错 F 的序列 $\langle F_0, F_2, F_4, F_6, \dots \rangle = \langle 0, 1, 3, 8, \dots \rangle$ 有简单的母函数:

$$\sum_n F_{2n} z^n = \frac{z}{1-3z+z^2} \quad (7.24)$$

7.3 解递归

现在让我们把注意力转到母函数的一个最重要的应用: 递归关系的解。

给定满足一个给定递归的序列 $\langle g_n \rangle$, 我们寻找依据 n 的 g_n 的闭形式。这个问题的解通过 4 步母函数的处理即可求得。这些步骤几乎是机械的, 足以在计算机上完成。

(1) 写出依据序列的其他元素表达 g_n 的单个方程。这个方程应对所有整数 n 成立, 假设 $g_{-1} = g_{-2} = \dots = 0$ 。

(2) 用 z^n 乘方程两边, 且在所有 n 上求和。左边给出和 $\sum_n g_n z^n$, 它是母函数 $G(z)$ 。右边应进行操作, 以致它变成涉及 $G(z)$ 的某个其他表达式。

(3) 解产生的方程, 取得 $G(z)$ 的闭形式。

(4) 把 $G(z)$ 展成幂级数且读出 z^n 的系数, 这是 g_n 的闭形式。

这个方法行得通, 因为以这种方式单个函数 $G(z)$ 表示整个序列 $\langle g_n \rangle$, 许多操作是可能的。

例 1: 再论 Fibonacci 数。

例如, 让我们重温第六章中的 Fibonacci 数的推导。在那一章中我们学到了一种新方法; 现在我们更有条理、给定的递归是

$$g_0 = 0; g_1 = 1;$$

$$g_n = g_{n-1} + g_{n-2}, n \geq 2.$$

由上面 4 步我们将找到 g_n 的闭形式。

第 1 步说明写出 g_n 的递归像“单个方程”。我们能说

$$g_n = \begin{cases} 0 & \text{若 } n \leq 0; \\ 1 & \text{若 } n = 1; \\ g_{n-1} + g_{n-2} & \text{若 } n > 1. \end{cases}$$

但是这使人失望。第 1 步实际想要一个公式，它不涉及一种情形一种情形的构造。单个方程

$$g_n = g_{n-1} + g_{n-2}$$

对 $n \geq 2$ 成立，且当 $n \leq 0$ 时它也成立(因为我们有 $g_0 = 0$ 和 $g_{\text{负的}} = 0$)。但是当 $n = 1$ 时我们在左边取得 1，在右边取得 0。幸运的是，问题容易确定，因为我们把 $[n = 1]$ 加到右边；当 $n = 1$ 时加 1，当 $n \neq 1$ 时它不改变。所以我们得到

$$g_n = g_{n-1} + g_{n-2} + [n = 1],$$

这是第 1 步所要的方程。

现在第 2 步要求把 $\langle g_n \rangle$ 的方程变换成 $G(z) = \sum g_n z^n$ 的方程。此工作不难：

$$\begin{aligned} G(z) &= \sum g_n z^n = \sum g_{n-1} z^n + \sum g_{n-2} z^n + \sum [n = 1] z^n \\ &= \sum g_n z^{n+1} + \sum g_n z^{n+2} + z \\ &= zG(z) + z^2 G(z) + z \end{aligned}$$

此时第 3 步也是简单的；我们有

$$G(z) = \frac{z}{1 - z - z^2},$$

当然它来得并不意外。

第 4 步是无可争辩的。在第六章中我们由一闪灵机来实现它；现在让我们进行得慢一些，以致当我们遇到较困难的问题时也能可靠地到达第 4 步。当 $z / (1 - z - z^2)$ 展成幂级数时 z^n 的系数

$$[z^n] \frac{z}{1 - z - z^2} \text{ 是什么? 更一般, 如果给定任何有理函数}$$

$$R(z) = \frac{P(z)}{Q(z)},$$

其中 P 和 Q 是多项式，系数 $[z^n]R(z)$ 是什么?

有一类系数特别好的有理函数, 即

$$\frac{a}{(1-\rho z)^{m+1}} = \sum_{n \geq 0} \binom{m+n}{m} a \rho^n z^n. \quad (7.25)$$

(情形 $\rho=1$ 呈现在表 7.2 中, 且我们能对 z 替代 ρz 取得这里说明的一般公式。)像式(7.25)那样函数的有限和,

$$S(z) = \frac{a_1}{(1-\rho_1 z)^{m_1+1}} + \frac{a_2}{(1-\rho_2 z)^{m_2+1}} + \cdots + \frac{a_l}{(1-\rho_l z)^{m_l+1}}, \quad (7.26)$$

也有好的系数,

$$[z^n]S(z) = a_1 \binom{m_1+n}{m_1} \rho_1^n + a_2 \binom{m_2+n}{m_2} \rho_2^n + \cdots + a_l \binom{m_l+n}{m_l} \rho_l^n. \quad (7.27)$$

我们将表明使得 $R(0) \neq \infty$ 的每个有理函数 $R(z)$ 能表达为形式

$$R(z) = S(z) + T(z), \quad (7.28)$$

其中 $S(z)$ 有形式(7.26), $T(z)$ 是多项式. 所以有系数 $[z^n]R(z)$ 的一个闭形式. 找 $S(z)$ 和 $T(z)$ 等价于找 $R(z)$ 的“部分分式展开”.

注意当 z 有值 $1/\rho_1, \dots, 1/\rho_l$ 时 $S(z) = \infty$, 所以需找数 ρ_k . 如果我们将成功地把 $R(z)$ 表达为希望的形式 $S(z) + T(z)$, 则一定是数 α_k 的倒数, 在那里 $Q(\alpha_k) = 0$. (回想起 $R(z) = P(z)/Q(z)$, 其中 P 和 Q 是多项式, 仅当 $Q(z) = 0$ 时有 $R(z) = \infty$.)

假设 $Q(z)$ 有形式

$$Q(z) = q_0 + q_1 z + \cdots + q_m z^m, \text{ 其中 } q_0 \neq 0 \text{ 和 } q_m \neq 0.$$

“反射”多项式

$$Q^R(z) = q_0 z^m + q_1 z^{m-1} + \cdots + q_m$$

和 $Q(z)$ 有一个重要关系:

$$\begin{aligned} Q^R(z) &= q_0(z - \rho_1) \cdots (z - \rho_m) \\ \Leftrightarrow Q(z) &= q_0(1 - \rho_1 z) \cdots (1 - \rho_m z). \end{aligned}$$

于是, Q^R 的根是 Q 的根的倒数, 反之亦对. 所以通过反射多项式 $Q^R(z)$ 的因子分解能找到数 ρ_k .

例如, 在 Fibonacci 情形中我们有

$$Q(z) = 1 - z - z^2; \quad Q^R(z) = z^2 - z - 1.$$

在二次公式 $(-b \pm \sqrt{b^2 - 4ac})/2a$ 中置 $(a, b, c) = (1, -1, -1)$ 能找到 Q^R 的根; 我们发现它们是

$$\varphi = \frac{1 + \sqrt{5}}{2} \text{ 和 } \hat{\varphi} = \frac{1 - \sqrt{5}}{2}.$$

所以 $Q^n(z) = (z - \varphi)(z - \hat{\varphi})$ 和 $Q(z) = (1 - \varphi z)(1 - \hat{\varphi} z)$.

一旦找到 ρ , 我们能进行找部分分式展开. 如果所有根不同, 情况是简单的, 让我们先考虑特殊情形. 我们可同样说明和形式证明一般的结果:

不同根的有理展开定理.

若 $R(z) = P(z)/Q(z)$, 其中 $Q(z) = q_0(1 - \rho_1 z) \cdots (1 - \rho_l z)$, 且数 (ρ_1, \dots, ρ_l) 是不同的, 若 $P(z)$ 是次数低于 l 的多项式, 则

$$[z^n]R(z) = a_1 \rho_1^n + \cdots + a_l \rho_l^n, \text{ 其中 } a_k = \frac{-\rho_k P(1/\rho_k)}{Q'(1/\rho_k)}. \quad (7.29)$$

证明: 设 a_1, \dots, a_l 是固定的常数. 如果 $R(z) = P(z)/Q(z)$ 等于

$$S(z) = \frac{a_1}{1 - \rho_1 z} + \cdots + \frac{a_l}{1 - \rho_l z},$$

公式 (7.29) 成立. 且表明当 $z \rightarrow 1/\rho_k$ 时函数 $T(z) = R(z) - S(z)$ 不是无限, 我们能证 $R(z) = S(z)$. 关于这一点将表明有理函数 $T(z)$ 不为无限, 因此 $T(z)$ 一定是多项式. 我们也能证明 $z \rightarrow \infty$ 时 $T(z) \rightarrow 0$, 因此 $T(z)$ 一定是零.

设 $\alpha_k = 1/\rho_k$. 为了证明 $\lim_{z \rightarrow \alpha_k} T(z) \neq \infty$, 证明 $\lim_{z \rightarrow \alpha_k} (z - \alpha_k)T(z) = 0$ 就够了, 因为 $T(z)$ 是 z 的有理函数, 因此我们要证

$$\lim_{z \rightarrow \alpha_k} (z - \alpha_k)R(z) = \lim_{z \rightarrow \alpha_k} (z - \alpha_k)S(z).$$

右边的极限等于 $\lim_{z \rightarrow \alpha_k} a_k(z - \alpha_k)/(1 - \rho_k z) = -a_k/\rho_k$, 因为 $(1 - \rho_k z) = -\rho_k(z - \alpha_k)$

和 $(z - \alpha_k)/(1 - \rho_j z) \rightarrow 0 (j \neq k)$. 左边的极限依据 L'Hospital 规则为

$$\lim_{z \rightarrow \alpha_k} (z - \alpha_k) \frac{P(z)}{Q(z)} = P(\alpha_k) \lim_{z \rightarrow \alpha_k} \frac{z - \alpha_k}{Q(z)} = \frac{P(\alpha_k)}{Q'(\alpha_k)}.$$

于是证明了定理.

回到 Fibonacci 例子上来, 我们有 $P(z) = z$ 和 $Q(z) = 1 - z - z^2 = (1 - \varphi z)(1 - \hat{\varphi} z)$; 因此 $Q'(z) = -1 - 2z$,

$$\frac{-\rho P(1/\rho)}{Q'(1/\rho)} = \frac{-1}{-1 - 2/\rho} = \frac{\rho}{\rho + 2}.$$

按照式 (7.29), $[z^n]R(z)$ 中 φ^n 的系数为 $\varphi/(\varphi + 2) = 1/\sqrt{5}$; $\hat{\varphi}$ 的系数为 $\hat{\varphi}/(\hat{\varphi} + 2) = -1/\sqrt{5}$. 所以如同式 (6.123) 中那样, 定理告诉我们 $F_n = (\varphi^n - \hat{\varphi}^n)/\sqrt{5}$.

当 $Q(z)$ 有重根时, 计算变得更困难, 但是我们能加强定理和证明且证明下列更一般的结论:

有理母函数的一般展开定理

若 $R(z) = P(z)/Q(z)$, 其中 $Q(z) = q_0(1 - \rho_1 z)^{d_1} \cdots (1 - \rho_l z)^{d_l}$ 且数 (ρ_1, \dots, ρ_l) 是不同的, 如果 $P(z)$ 是次数小于 $d_1 + \dots + d_l$ 的多项式, 则

$$[z^n]R(z) = f_1(n)\rho_1^n + \dots + f_l(n)\rho_l^n, \text{ 对所有 } n \geq 0, \quad (7.30)$$

其中每个 $f_k(n)$ 是首项系数为

$$a_k = \frac{(-\rho_k)^{d_k} P(1/\rho_k) d_k}{Q^{(d_k)}(1/\rho_k)} = \frac{P(1/\rho_k)}{(d_k - 1)! q_0 \prod_{j \neq k} (1 - \rho_j / \rho_k)^{d_j}} \quad (7.31)$$

次数 $d_k - 1$ 的多项式。这能通过 $\max(d_1, \dots, d_l)$ 上归纳来证明, 用这样一个事实, 即

$$R(z) = \frac{a_1(d_1 - 1)!}{(1 - \rho_1 z)^{d_1}} + \dots + \frac{a_l(d_l - 1)!}{(1 - \rho_l z)^{d_l}}$$

是有理函数, 它的分母多项式不可被 $(1 - \rho_k z)^{d_k}$ (任何 k) 除尽。

例 2: 一个近乎随机的递归。

现在已看到一些一般的方法, 我们准备处理新的问题。让我们尝试找递归

$$\begin{aligned} g_0 &= g_1 = 1; \\ g_n &= g_{n-1} + 2g_{n-2} + (-1)^n, \quad n \geq 2 \end{aligned} \quad (7.32)$$

的闭形式, 先作出小的情形的一个表总是一个好主意, 且让我们做的递归也容易:

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

无闭形式是明显的, 且这个序列甚至在 Sloane 手册^[270]中没有列出; 所以如果要发现解我们需通过 4 个步骤过程。

第 1 步是容易的, 因为当 $n < 2$ 时仅需把捏造的因子插入确定的情况:

$$g_n = g_{n-1} + 2g_{n-2} + (-1)^n [n \geq 0] + [n = 1]$$

对所有整数 n 成立。现在能实现第 2 步:

$$\begin{aligned} G(z) &= \sum_n g_n z^n = \sum_n g_{n-1} z^n + 2 \sum_n g_{n-2} z^n + \sum_{n \geq 0} (-1)^n z^n + \sum_{n=1} z^n \\ &= zG(z) + 2z^2 G(z) + \frac{1}{1+z} + z \end{aligned}$$

(顺便提一句, 我们还能用 $\binom{-1}{n}$ 而不是 $(-1)^n [n \geq 0]$, 从而由二项定理取得 $\sum_n \binom{-1}{n} z^n = (1+z)^{-1}$.) 第3步是初等代数, 它产生

$$G(z) = \frac{1+z(1+z)}{(1+z)(1-z-2z^2)} = \frac{1+z+z^2}{(1-2z)(1+z)^2}.$$

留下第4步.

在分母中的二次因子有点麻烦, 因为我们知道重根比不同根更复杂; 但是下面就是我们的解答. 我们有两个根 $\rho_1 = 2$ 和 $\rho_2 = -1$, 一般展开定理(7.30)告诉我们对某个常数 c

$$g_n = a_1 2^n + (a_2 n + c)(-1)^n,$$

其中

$$a_1 = \frac{1+1/2+1/4}{(1+1/2)^2} = \frac{7}{9}, \quad a_2 = \frac{1-1+1}{1-2/(-1)} = \frac{1}{3}.$$

(当分母有好因子时式(7.31)中 a_k 的第2个公式比第1个公式容易使用. 我们在 $R(z)$ 中简单地代 $z = 1/\rho_k$, 除了在因子中这给出零, 且用 $(d_k - 1)!$ 除; 这给出 $n^{d_k-1} \rho_k^n$ 的系数.) 代入 $n=0$, 剩下常数 c 的值可得为 $2/9$. 因此解答为

$$g_n = \frac{7}{9} 2^n + \left(\frac{1}{3} n + \frac{2}{9} \right) (-1)^n. \quad (7.33)$$

检验情形 $n=1$ 和 2 没有坏处, 恰好肯定我们没有做错. 甚至也应试 $n=3$, 因为这个公式看来古怪. 但是它是正确的, 没有错.

能否用猜测来发现式(7.33)呢? 也许列出更多一些值, 我们可发现当 n 大时 $g_{n+1} \approx 2g_n$. 由于侥幸我们甚至可能抽出常数 $7/9$. 但是可以确信得到的母函数作为一个工具是较简单和更可靠的.

例3: 相互递归序列

有时我们有两个或多个相互依赖的递归. 于是我们能形成它们两者的母函数, 且通过4个步骤方法的简单推广来解这两个母函数.

例如, 让我们回到本章前面探讨的 $3 \times n$ 多米诺骨牌覆盖的问题. 如果仅知道总方式数 U_n 用多米诺骨牌覆盖一个 $3 \times n$ 矩形, 不把此数分解为垂直多米诺骨牌和水平多米诺骨牌, 我们不需进行像前面那样详细的讨论. 我们能仅设置递归

$$\begin{aligned} U_0 &= 1, \quad U_1 = 0; \quad V_0 = 0, \quad V_1 = 1; \\ U_n &= 2V_{n-1} + U_{n-2}, \quad V_n = U_{n-1} + V_{n-2}, \quad n \geq 2. \end{aligned}$$

这 V_n 是用 $(3n-1)/2$ 块多米诺骨牌覆盖缺一角的 $3 \times n$ 矩形的方式数。如同前面那样, 如果考虑矩形左边处可能的多米诺配置, 这些递归是容易发现的, 这里是小 n 的 U_n 和 V_n 的值:

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

让我们在 4 步中找闭形式。首先(第 1 步), 对于所有 n , 我们有

$$U_n = 2V_{n-1} + U_{n-2} + [n=0], \quad V_n = U_{n-1} + V_{n-2}.$$

因此(第 2 步),

$$U(z) = 2zV(z) + z^2U(z) + 1, \quad V(z) = zU(z) + z^2V(z).$$

现在(第 3 步)我们一定要解两个未知数的两个方程; 但是这是容易的, 因为第二个方程产生 $V(z) = zU(z)/(1-z^2)$; 我们找到

$$U(z) = \frac{1-z^2}{1-4z^2+z^4}, \quad \text{且 } V(z) = \frac{z}{1-4z^2+z^4}. \quad (7.35)$$

(在式(7.10)中我们有 $U(z)$ 的这个公式, 但是用 z^3 替代 z^2 。在那个推导中, n 是多米诺骨牌数, 现在它是矩形的宽。)

分母 $1-4z^2+z^4$ 是 z^2 的函数, 这就是使 $U_{2n+1}=0$ 和 $V_{2n}=0$ 的事物。当我们分解分母时能通过保留 z^2 而利用 z^2 的这个好的性质: 我们不需自始至终把 $1-4z^2+z^4$ 引到 4 个因子 $(1-\rho_k z)$ 的乘积, 因为形式 $(1-\rho_k z^2)$ 的 2 个因子足以告诉我们系数。换句话说, 如果考虑母函数

$$W(z) = \frac{1}{1-4z+z^2} = W_0 + W_1z + W_2z^2 + \cdots, \quad (7.36)$$

我们将得到 $V(z) = zW(z^2)$ 和 $U(z) = (1-z^2)W(z^2)$; 因此 $V_{2n+1} = W_n$, $U_{2n} = W_n - W_{n-1}$ 。通过处理较简单的函数 $W(z)$ 我们节省了时间和精力。

$1-4z+z^2$ 的因子是 $(z-2-\sqrt{3})$ 和 $(z-2+\sqrt{3})$, 且它们也能写成 $(1-(2+\sqrt{3})z)$ 和 $(1-(2-\sqrt{3})z)$, 因为这个多项式就是它的反射。于是, 我们得到

$$V_{2n+1} = W_n = \frac{3+2\sqrt{3}}{6} (2+\sqrt{3})^n + \frac{3-2\sqrt{3}}{6} (2-\sqrt{3})^n;$$

$$U_{2n} = W_n - W_{n-1} = \frac{3+\sqrt{3}}{6} (2+\sqrt{3})^n + \frac{3-\sqrt{3}}{6} (2-\sqrt{3})^n$$

$$= \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} + \frac{(2-\sqrt{3})^n}{3+\sqrt{3}}, \quad (7.37)$$

这就是 $3 \times n$ 多米诺覆盖数的希望闭形式。

顺便提一句，认识到第二项总在 0 和 1 之间能简化 U_{2n} 的公式。数 U_{2n} 是整数，所以我们有

$$U_{2n} = \left\lceil \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} \right\rceil, \quad n \geq 0. \quad (7.38)$$

事实上，另一项 $(2-\sqrt{3})^n / (3+\sqrt{3})$ 当 n 大时极端小，因为 $2-\sqrt{3} \approx 0.268$ 。如果在数值计算中试用公式(7.38)，这需要重视。例如，当要求计算 $(2+\sqrt{3})^{10} / (3-\sqrt{3})$ 时，一个中等名牌的计算器提供 413 403.000 5。这正确到 9 位有意义的数字；但是准确值稍小于 413 403，不是稍大。所以取 413 403.000 5 的上整数将是一个错误；正确解答 $U_{20} = 413 403$ 通过四舍五入到最近整数得到。取上整数是冒险的。

例 4: 兑换的一个闭形式。

当我们离开作兑换的问题时，仅计算了付 50 分的方式数。让我们来计算兑换 1 元，或 1 百万元的方式数。

前面导出的母函数为

$$C(z) = \frac{1}{1-z} \frac{1}{1-z^5} \frac{1}{1-z^{10}} \frac{1}{1-z^{25}} \frac{1}{1-z^{50}},$$

这是一个分母次数为 91 的 z 的有理函数，所以我们能分解分母为 91 个因子，且提供在兑换中给出 n 分的方式数 C_n 的 91 项的“闭形式”。但是仔细考虑这是极糟的。在此特殊情形，能不能比提供的一般方法处理得更好？

当我们注意到分母几乎是 z^5 的函数，立即浮现一线希望。如果用 $(1+z+z^2+z^3+z^4)/(1-z^5)$ 替换 $1/(1-z)$ ，注意到 $1-4z^2+z^4$ 是 z^2 的函数，我们刚用的诀窍来简化计算能应用到 $C(z)$ ：

$$\begin{aligned} C(z) &= \frac{1+z+z^2+z^3+z^4}{1-z^5} \frac{1}{1-z^5} \frac{1}{1-z^{10}} \frac{1}{1-z^{25}} \frac{1}{1-z^{50}} \\ &= (1+z+z^2+z^3+z^4) \check{C}(z^5), \\ \check{C}(z) &= \frac{1}{1-z} \frac{1}{1-z} \frac{1}{1-z^2} \frac{1}{1-z^5} \frac{1}{1-z^{10}}. \end{aligned}$$

简练的函数 $\check{C}(z)$ 有次数仅为 19 的一个分母，所以它比原来的易处理。顺便提一句， $C(z)$ 的新表达式显示 $C_{5n} = C_{5n+1} = C_{5n+2} = C_{5n+3} + C_{5n+4}$ ；回顾此方程的集合是显然的：留下 53 分小费的方式数和留下 50 分小费的方式数相同，因为对模 5 来说 1 分的个数是预先定的。

但是基于分母的根 $\check{C}(z)$ 仍没有真正简单的闭形式。计算 $\check{C}(z)$ 的系数的最简单方式也

许是认识到每个分母因子是 $1 - z^{10}$ 的一个因子。因此我们能写出

$$\check{C}(z) = \frac{A(z)}{(1 - z^{10})^5}, \text{ 其中 } A(z) = A_0 + A_1 z + \cdots + A_{31} z^{31}. \quad (7.39)$$

想知道的 $A(z)$ 的值是

$$\begin{aligned} & (1 + z + \cdots + z^9)^2 (1 + z^2 + \cdots + z^8)(1 + z^5) \\ &= 1 + 2z + 4z^2 + 6z^3 + 9z^4 + 13z^5 + 18z^6 + 24z^7 \\ & \quad + 31z^8 + 39z^9 + 45z^{10} + 52z^{11} + 57z^{12} + 63z^{13} + 67z^{14} + 69z^{15} \\ & \quad + 69z^{16} + 67z^{17} + 63z^{18} + 57z^{19} + 52z^{20} + 45z^{21} + 39z^{22} + 31z^{23} \\ & \quad + 24z^{24} + 18z^{25} + 13z^{26} + 9z^{27} + 6z^{28} + 4z^{29} + 2z^{30} + z^{31} \end{aligned}$$

最后, 由于 $1/(1 - z^{10})^5 = \sum_{k \geq 0} \binom{k+4}{4} z^{10k}$, 我们能确定当 $n = 10q + r$ 和 $0 \leq r < 10$ 时

$\check{C}_n = [z^n] \check{C}(z)$ 的系数如下:

$$\begin{aligned} \check{C}_{10q+r} &= \sum_{i,k} A_i \binom{k+4}{4} [10q+r=10k+i] \\ &= A_r \binom{q+4}{4} + A_{r+10} \binom{q+3}{4} + A_{r+20} \binom{q+2}{4} + A_{r+30} \binom{q+1}{4}. \end{aligned} \quad (7.40)$$

这给出 10 种情形, 每个 r 值对应一种情形; 但是和涉及复数的幂的选择对象比较, 它是一个十分好的闭形式。

例如, 我们能用来表达式导出 $C_{50q} = \check{C}_{10q}$ 的值。于是由 $r=0$, 我们得到

$$C_{50q} = \binom{q+4}{4} + 45 \binom{q+3}{4} + 52 \binom{q+2}{4} + 2 \binom{q+1}{4}.$$

兑换 50 分的方式数为 $\binom{5}{4} + 45 \binom{4}{4} = 50$; 兑换 1 元的方式数为 $\binom{6}{4} + 45 \binom{5}{4} + 52 \binom{4}{4} = 292$; 兑换 1 000 000 元的方式数为

$$\begin{aligned} & \binom{2\,000\,004}{4} + 45 \binom{2\,000\,003}{4} + 52 \binom{2\,000\,002}{4} + 2 \binom{2\,000\,001}{4} \\ &= 66\,666\,793\,333\,412\,666\,685\,000\,001. \end{aligned}$$

例 5: 一个发散级数。

现在让我们尝试取得由

$$g_0 = 1,$$

$$g_n = n g_{n-1}, \quad n > 1$$

定义的数 g_n 的一个闭形式。在此开始之后我们马上认识到 g_n 就是 $n!$; 事实上, 第二章中描述的求和因子的方法立即提出了这个解答。但是让我们试用母函数来解递归, 就看出发

生了什么情况。(一个有力的技巧应能处理像这样容易的递归,也就像我们不能十分容易猜测解答的其它递归。)

方程

$$g_n = ng_{n-1} + [n=0]$$

对所有 n 成立,且导出

$$G(z) = \sum_n g_n z^n = \sum_n ng_{n-1} z^n + \sum_{n=0} z^n$$

为了完成第2步,我们要依据 $G(z)$ 表达 $\sum_n ng_{n-1} z^n$, 表 7.1 中的基本操作提供了以某种方

式涉及的导数 $G'(z) = \sum_n ng_n z^{n-1}$. 所以我们筹划接近这样的和的类型:

$$G(z) = 1 + \sum_n (n+1)g_n z^{n+1} = 1 + \sum_n ng_n z^{n+1} + \sum_n g_n z^{n+1} = 1 + z^2 G'(z) + zG(z).$$

让我们检验此方程,对于小 n 用 g_n 的值. 由于

$$G = 1 + z + 2z^2 + 6z^3 + 24z^4 + \dots,$$

$$G' = 1 + 4z + 18z^2 + 96z^3 + \dots,$$

我们得到

$$z^2 G' = z^2 + 4z^3 + 18z^4 + 96z^5 + \dots,$$

$$zG = z + z^2 + 2z^3 + 6z^4 + 24z^5 + \dots,$$

$$1 = 1.$$

这三行加起来为 G , 所以至今我们的解是好极了. 顺便提一句, 我们常发现写 ' G ' 来替代 ' $G(z)$ ' 是方便的; 当不改变 z 时, 额外的 ' (z) ' 使公式显得杂乱.

接着是第3步, 这与前面我们所做的不同, 因为要解一个微分方程. 但是这是一个能用 5.6 节的超几何级数技巧处理的微分方程, 这些技巧不太拙劣. (不熟悉超几何的读者不必担忧, 这将是易懂的。)

首先我们一定要除去常数 '1', 所以对两边取导数:

$$\begin{aligned} G' &= (z^2 G' + zG + 1)' = (2zG' + z^2 G'') + (G + zG') \\ &= z^2 G'' + 3zG' + G \end{aligned}$$

第五章中的理论告诉我们用 θ 算子改写它, 且由习题 6.13 可知

$$\theta G = zG', \quad \theta^2 G = z^2 G'' + zG'.$$

所以微分方程的希望形式是

$$\theta G = z\theta^2 G + 2z\theta G + zG - z(\theta + 1)^2 G.$$

按照式(5.119), 具有 $g_0 = 1$ 的解为超几何级数 $F(1, 1; z)$.

第3步我们没有料到; 但是现在知道了函数 G 是什么, 第4步是容易的, 超几何定义(5.76)给出幂级数展开:

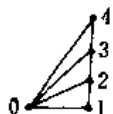
$$G(z) = F\left(\begin{matrix} 1, 1 \\ 1 \end{matrix} \middle| z\right) = \sum_{n \geq 0} \frac{1^{\bar{n}} 1^{\bar{n}} z^n}{n!} = \sum_{n \geq 0} n! z^n.$$

我们确定了一直知道的闭形式 $g_n = n!$.

注意, 即使 $G(z)$ 对所有非零 z 发散但这种技巧仍给出正确的解答. 序列 $n!$ 增长得极快, 当 $n \rightarrow \infty$ 项 $|n! z^n|$ 趋于 ∞ , 除非 $z = 0$. 这表明形式幂级数可用代数方法操作而不必担忧收敛性.

例6: 自始至终返回的一个递归.

让我们应用母函数到图论中的一个问题来结束这一节. n 级的扇是具有 $2n-1$ 条边的顶点 $\{0, 1, \dots, n\}$ 上的一个图, 定义如下: 顶点 0 用一条边连接每一个其他的 n 个顶点, 且顶点 k 用一条边连接顶点 $k+1$, $1 \leq k < n$. 例如, 这里是 4 级的扇, 它有 5 个顶点和 7 条边.



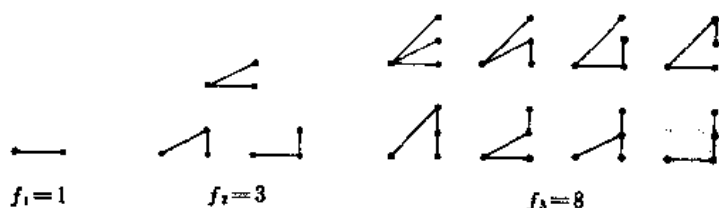
有趣的问题: 在这样一个图中有多少个支撑树? 一个支撑树是包含所有顶点的一个子图, 且包含足够的边来构造子图仍然连通但不形成一个循环. 结果 $n+1$ 个顶点上的一个图的每一个支撑树恰好有 n 条边. 少于 n 条边子图将不连通, 多于 n 条边子图将有一个循环; 图论的书中证明了这一点.

在 n 级的扇中从 $2n-1$ 条边中间选取 n 条边有 $\binom{2n-1}{n}$ 种方式, 但是这些选取不总产生一个支撑树. 例如子图

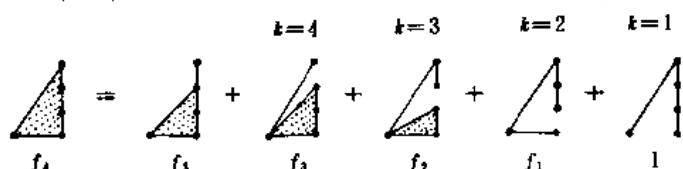


有 4 条边, 但不是支撑树; 它有一个从 0 到 4 到 3 到 0 的循环, 且在 $\{1, 2\}$ 和其他顶点之间没有连接. 我们要计算 $\binom{2n-1}{n}$ 种选取的多少种实际产生支撑树.

让我们看一些小的情形. 枚举 $n=1, 2$ 和 3 的支撑树是很容易的: (如果我们总在左边画出顶点 0, 则不需表明顶点的标号.) $n=0$ 的情形是什么? 首先看来置 $f_0 = 1$ 是合理的; 但是我们将取 $f_0 = 0$, 因为 0 级的扇的存在性(它应有 $2n-1 = -1$ 条边)是可疑的.



我们的 4 步过程告诉我们来找对所有 n 成立的 f_n 的一个递归。通过观察最上面的顶点(顶点 n)如何连接到支撑树的剩余部分, 我们能取得一个递归。如果它不和顶点 0 连接, 则它一定和顶点 $n-1$ 连接, 因为它一定和图的剩余部分连接。此时, $\{0$ 直到 $n-1$ 的顶点上)余下扇的 f_n 个支撑树的任何一个将完成整个图的一个支撑树。否则顶点 n 和 0 连接, 且有某个数 $k \leq n$ 使得直接连接顶点 $n, n-1, \dots, k$, 但是 k 和 $k-1$ 之间的边不在子树中。于是不能有 0 和 $\{n-1, \dots, k\}$ 之间的边, 或者将有一个循环。如果 $k=1$, 支撑树完全被确定。如果 $k>1$, f_{k-1} 种方式的任何一种来产生 $\{0, 1, \dots, k-1\}$ 上的一个支撑树将产生整个图上的一个支撑树。例如, 这里就是 $n=4$ 时所作的分析:



对于 $n \geq 1$ 成立的一般方程是

$$f_n = f_{n-1} + f_{n-1} + f_{n-2} + f_{n-3} + \dots + f_1 + 1.$$

(在末端的 '1' 看来好像差不多为 f_0 , 我们应选取 $f_0 = 1$; 但是要牢记住我们的选取。)稍许改变足以作出对所有整数 n 成立的方程:

$$f_n = f_{n-1} + \sum_{k < n} f_k + [n > 0]. \quad (7.41)$$

这是从 f_{n-1} 直到所有前面值的“自始至终返回”的一个递归, 所以它不同于本章至今所见的其他递归。我们用了一种特殊方法来除去第二章中当解快速分类递归(2.12)时的一个相似的右边和; 即, 我们从另一个 $(f_{n+1} - f_n)$ 中除去递归的一个实例。这个诀窍现在将除去 Σ , 如同它当时做的那样; 但是将看到母函数允许我们直接用这样的和处理。(这是母函数所做的一件好事, 因为不久将看到十分复杂的递归。)

第 1 步已结束, 第 2 步需做一件新的事情:

$$\begin{aligned}
 F(z) &= \sum_n f_n z^n = \sum_n f_{n-1} z^n + \sum_{k < n} f_k z^n + \sum_n [n > 0] z^n \\
 &= zF(z) + \sum_k f_k z^k \sum_n [n > k] z^{n-k} + \frac{z}{1-z}
 \end{aligned}$$

$$\begin{aligned}
 &= zF(z) + F(z) \sum_{m \geq 0} z^m + \frac{z}{1-z} \\
 &= zF(z) + F(z) \frac{z}{1-z} + \frac{z}{1-z}.
 \end{aligned}$$

这里的关键诀窍是把 z^n 改变成 $z^k z^{n-k}$ ；这使它可能依据 $F(z)$ 表达双重和的值，正如第 2 步中所要求的。

现在第 3 步是简单的代数，我们找到

$$F(z) = \frac{z}{1 - 3z + z^2}.$$

有点记性的人将识别这就像偶数编号 Fibonacci 数的母函数(7.24)。所以我们不需再做第 4 步；我们找到扇的支撑问题的有点意外的解答：

$$f_n = F_{2n}, \quad n \geq 0. \quad (7.42)$$

7.4 特殊的母函数

如果我们知道许多不同幂级数的系数，则 4 步过程的第 4 步变得十分容易。就其本身来说表 7.2 中的展开式是十分有用的，而许多其他类型的闭形式是可能的。所以我们应该用对应于第六章中考虑的“特殊数”所列出的幂级数来补充表。

表 7.3 特殊数的母函数

$$\frac{1}{(1-z)^{m+1}} \ln \frac{1}{1-z} = \sum_n (H_{m+n} - H_n) \binom{m+n}{n} z^n \quad (7.43)$$

$$\frac{z}{e^z - 1} = \sum_{n \geq 0} B_n \frac{z^n}{n!} \quad (7.44)$$

$$\frac{F_m z}{1 - (F_{m-1} + F_{m+1})z + (-1)^m z^2} = \sum_n F_{mn} z^n \quad (7.45)$$

$$\sum_k \left\{ \begin{matrix} m \\ k \end{matrix} \right\} \frac{k! z^k}{(1-z)^{k+1}} = \sum_n n^m z^n \quad (7.46)$$

$$(z^{-1})^{-m} = \frac{z^m}{(1-z)(1-2z)\cdots(1-mz)} = \sum_n \left\{ \begin{matrix} n \\ m \end{matrix} \right\} z^n \quad (7.47)$$

$$z^m = z(z+1)\cdots(z+m-1) = \sum_n \left[\begin{matrix} m \\ n \end{matrix} \right] z^n \quad (7.48)$$

$$(e^z - 1)^m = m! \sum_{n \geq 0} \left\{ \begin{matrix} n \\ m \end{matrix} \right\} \frac{z^n}{n!} \quad (7.49)$$

$$\left(\ln \frac{1}{1-z} \right)^m = m! \sum_{n \geq 0} \left[\begin{matrix} n \\ m \end{matrix} \right] \frac{z^n}{n!} \quad (7.50)$$

$$\left(\frac{z}{\ln(1+z)} \right)^m = \sum_{n \geq 0} \frac{z^n}{n!} \left\{ \begin{matrix} m \\ m-n \end{matrix} \right\} / \binom{m-1}{n} \quad (7.51)$$

$$\left(\frac{z}{1-e^{-z}} \right)^m = \sum_{n \geq 0} \frac{z^n}{n!} \left[\begin{matrix} m \\ m-n \end{matrix} \right] / \binom{m-1}{n} \quad (7.52)$$

$$e^{z+ws} = \sum_{n \geq 0} \binom{n}{m} w^m \frac{z^n}{n!} \quad (7.53)$$

$$e^{w(e^z-1)} = \sum_{n \geq 0} \left\{ \begin{matrix} n \\ m \end{matrix} \right\} w^m \frac{z^n}{n!} \quad (7.54)$$

$$\frac{1}{(1-z)^w} = \sum_{n \geq 0} \left[\begin{matrix} n \\ m \end{matrix} \right] w^m \frac{z^n}{n!} \quad (7.55)$$

$$\frac{1-w}{e^{(w-1)z} - w} = \sum_{n \geq 0} \langle \begin{matrix} n \\ m \end{matrix} \rangle w^m \frac{z^n}{n!} \quad (7.56)$$

表 7.3 是我们需要的数据库。表中的等式不难证明，所以不需详细研究它们；当我们遇到一个新问题时，此表主要为了参考，但是值得提一提第一个公式(7.43)的一个有趣的证明：我们从等式

$$\frac{1}{(1-z)^{x+1}} = \sum_n \binom{x+n}{n} z^n$$

开始，且把它对 x 微分。在左边， $(1-z)^{-x-1}$ 是等于 $e^{(x+1)\ln(1/(1-z))}$ ，所以 d/dx 提供一个因子 $\ln(1/(1-z))$ 。在右边， $\binom{x+n}{n}$ 的分子是 $(x+n) \cdots (x+1)$ ，而 d/dx 把此分为 n 项，它的和等价于用

$$\frac{1}{x+n} + \cdots + \frac{1}{x+1} = H_{x+n} - H_x$$

乘 $\binom{x+n}{n}$ 。用 m 替换 x 给出式(7.43)。注意即使当 x 不是整数时， $H_{x+n} - H_x$ 也是有意义的。

顺便提一句，这个微分一个复杂乘积的方法，留下为一个乘积，通常比表达导数为一个和更好。例如

$$\frac{d}{dx} ((x+n)^n \cdots (x+1)^1) = (x+n)^n \cdots (x+1)^1 \left(\frac{n}{x+n} + \cdots + \frac{1}{x+1} \right)$$

的右边像一个和那样，是比较凌乱的。

表 7.3 中的一般等式包含许多重要的特殊情形。例如，当 $m=0$ 时式(7.43)简化成 H_n 的母函数：

$$\frac{1}{1-z} \ln \frac{1}{1-z} = \sum_n H_n z^n. \quad (7.57)$$

用其他方式也能推出此方程；例如，我们能取 $\ln(1/(1-z))$ 的幂级数且用 $1-z$ 除它而取得累积和。

等式(7.51)和(7.52)各自涉及到比 $\left\{ \begin{smallmatrix} m \\ m-n \end{smallmatrix} \right\} / \binom{m-1}{n}$ 和 $\left[\begin{smallmatrix} m \\ m-n \end{smallmatrix} \right] / \binom{m-1}{n}$ ，当 $n \geq m$ 时它有不定形式 $0/0$ 。然而，用式(6.45)的 Stirling 多项式有一种方式给出它们一种适当的含意，因为

$$\left\{ \begin{smallmatrix} m \\ m-n \end{smallmatrix} \right\} / \binom{m-1}{n} = (-1)^{n+1} n! m \sigma_n(n-m); \quad (7.58)$$

$$\left[\begin{smallmatrix} m \\ m-n \end{smallmatrix} \right] / \binom{m-1}{n} = n! m \sigma_n(m). \quad (7.59)$$

于是，例如，式(7.51)的情形 $m=1$ 不应认为是幂级数 $\sum_{n \geq 0} (z^n/n!) \left\{ \begin{smallmatrix} 1 \\ 1-n \end{smallmatrix} \right\} / \binom{0}{n}$ ，而认为是

$$\frac{z}{\ln(1+z)} = - \sum_{n \geq 0} (-z)^n \sigma_n(n-1) = 1 + \frac{1}{2}z - \frac{1}{12}z^2 + \cdots.$$

等式(7.53)，(7.55)，(7.54)和(7.56)是“双重母函数”或“超母函数”，因为它们有形式 $G(w, z) = \sum_{m,n} g_{m,n} w^m z^n$ ， w^m 的系数是变量 z 的一个母函数， z^n 的系数是变量 w 的一个母函数。

7.5 卷 积

两个给定序列 $\langle f_0, f_1, \dots \rangle = \langle f_n \rangle$ 和 $\langle g_0, g_1, \dots \rangle = \langle g_n \rangle$ 的卷积是序列 $\langle f_0 g_0, f_0 g_1 + f_1 g_0, \dots \rangle = \langle \sum_k f_k g_{n-k} \rangle$ 。在 5.4 节和 7.2 节中我们看到序列的卷积对应于它们的母函数的乘积。这个事实使它容易计算许多和，否则这些和将难处理。

例 1: Fibonacci 卷积

例如，让我们试计算 $\sum_{k=0}^n F_k F_{n-k}$ 为闭形式。这是 $\langle F_n \rangle$ 和它自己的卷积，所以和一定是 $F(z)^2$ 中 z^n 的系数，其中 $F(z)$ 是 $\langle F_n \rangle$ 的母函数。我们只需计算出这个系数的值。

母函数 $F(z)$ 是 $z/(1-z-z^2)$ ，多项式的一个商；所以有理函数的一般展开定理告诉我们能从部分分式表示得到解答。我们能一般展开定理(7.30)且用心地做出，或者能用事实

$$\begin{aligned}
 F(z)^2 &= \left(\frac{1}{\sqrt{5}} \left(\frac{1}{1-\varphi z} - \frac{1}{1-\hat{\varphi}z} \right) \right)^2 \\
 &= \frac{1}{5} \left(\frac{1}{(1-\varphi z)^2} - \frac{2}{(1-\varphi z)(1-\hat{\varphi}z)} + \frac{1}{(1-\hat{\varphi}z)^2} \right) \\
 &= \frac{1}{5} \sum_{n \geq 0} (n+1)\varphi^n z^n - \frac{2}{5} \sum_{n \geq 0} F_{n+1} z^n + \frac{1}{5} \sum_{n \geq 0} (n+1)\hat{\varphi}^n z^n.
 \end{aligned}$$

让我们尝试依据 Fibonacci 数的闭形式，而不是依据 φ 和 $\hat{\varphi}$ 表达解答。想起 $\varphi + \hat{\varphi} = 1$ ，我们得到

$$\begin{aligned}
 \varphi^n + \hat{\varphi}^n &= [z^n] \left(\frac{1}{1-\varphi z} + \frac{1}{1-\hat{\varphi}z} \right) \\
 &= [z^n] \frac{2 - (\varphi + \hat{\varphi})z}{(1-\varphi z)(1-\hat{\varphi}z)} = [z^n] \frac{2-z}{1-z-z^2} = 2F_{n+1} - F_n.
 \end{aligned}$$

因此

$$F(z)^2 = \frac{1}{5} \sum_{n \geq 0} (n+1)(2F_{n+1} - F_n) z^n - \frac{2}{5} \sum_{n \geq 0} F_{n+1} z^n,$$

且我们得到寻找的解答：

$$\sum_{k=0}^n F_k F_{n-k} = \frac{2nF_{n+1} - (n+1)F_n}{5}. \quad (7.60)$$

例如，当 $n=3$ 时，此公式在左边给出 $F_0 F_3 + F_1 F_2 + F_2 F_1 + F_3 F_0 = 0+1+1+0=2$ ，右边给出 $(6F_4 - 4F_3)/5 = (18-8)/5 = 2$ 。

例 2: 调和卷积

某种称为“样本类”的计算机方法的效率依赖于和

$$T_{m,n} = \sum_{0 \leq k \leq n} \binom{k}{m} \frac{1}{n-k}, \quad \text{整数 } m, n \geq 0$$

的值。习题 5.58 通过有点复杂的双重归纳使用求和因子获得这个和的值。十分容易了解 $T_{m,n}$ 就是 $\langle \binom{0}{m}, \binom{1}{m}, \binom{2}{m}, \dots \rangle$ 和 $\langle 0, \frac{1}{1}, \frac{1}{2}, \dots \rangle$ 的卷积的第 n 项。在表 7.2 中两个序列有简单的母函数：

$$\sum_{n \geq 0} \binom{n}{m} z^n = \frac{z^m}{(1-z)^{m+1}}; \quad \sum_{n \geq 0} \frac{z^n}{n} = \ln \frac{1}{1-z}.$$

所以，由式(7.43)，

$$\begin{aligned}
 T_{m,n} &= [z^n] \frac{z^m}{(1-z)^{m+1}} \ln \frac{1}{1-z} = [z^{n-m}] \frac{1}{(1-z)^{m+1}} \ln \frac{1}{1-z} \\
 &= (H_n - H_m) \binom{n}{n-m}.
 \end{aligned}$$

事实上, 有许多和归结为这个卷积的相同种类, 因为我们得到对于所有 r 和 s

$$\frac{1}{(1-z)^{r+1}} \ln \frac{1}{1-z} \cdot \frac{1}{(1-z)^{s+1}} = \frac{1}{(1-z)^{r+s+2}} \ln \frac{1}{1-z}.$$

使 z^n 的系数相等给出一般的等式

$$\sum_k \binom{r+k}{k} \binom{s+n-k}{n-k} (H_{r+k} - H_r) = \binom{r+s+n+1}{n} (H_{r+s+n+1} - H_{r+s+1}). \quad (7.61)$$

这似乎哪有这么好的事. 但是至少当 $n=2$ 时它检验了

$$\begin{aligned} & \binom{r+1}{1} \binom{s+1}{1} \frac{1}{r+1} + \binom{r+2}{2} \binom{s+0}{0} \left(\frac{1}{r+2} + \frac{1}{r+1} \right) \\ &= \binom{r+s+3}{2} \left(\frac{1}{r+s+3} + \frac{1}{r+s+2} \right). \end{aligned}$$

像 $s=0$ 的特殊情形是和一般情形一样值得注意.

还有另外一些, 我们能利用卷积等式

$$\sum_k \binom{r+k}{k} \binom{s+n-k}{n-k} = \binom{r+s+n+1}{n}$$

把 H_r 换到另一侧, 因为 H_r 和 k 独立:

$$\begin{aligned} & \sum_k \binom{r+k}{k} \binom{s+n-k}{n-k} H_{r+k} \\ &= \binom{r+s+n+1}{n} (H_{r+s+n+1} - H_{r+s+1} + H_r). \end{aligned} \quad (7.62)$$

还有: 如果 r 和 s 是非负整数 l 和 m , 我们能用 $\binom{l+k}{l}$ 替换 $\binom{r+k}{k}$, 用 $\binom{m+n-k}{m}$ 替换 $\binom{s+n-k}{n-k}$; 然后能把 k 改变为 $k-l$, n 改变为 $n-m-l$, 取得

$$\sum_{k=0}^n \binom{k}{l} \binom{n-k}{m} H_k = \binom{n+1}{l+m+1} (H_{n+1} - H_{l+m+1} + H_l), \text{ 整数 } l, m, n \geq 0. \quad (7.63)$$

即使这个等式的特殊情形 $l=m=0$, 在第二章中处理也是困难的! (见式(2.36))我们经过了长途跋涉.

例 3: 卷积的卷积

如果我们形成 $\langle f_n \rangle$ 和 $\langle g_n \rangle$ 的卷积, 然后这个卷积再与第三个序列 $\langle h_n \rangle$ 作卷积, 将取得一个序列, 它的第 n 项是

$$\sum_{j+k+l=n} f_j g_k h_l.$$

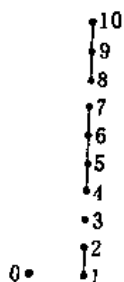
这个三重卷积的母函数当然是三重乘积 $F(z)G(z)H(z)$. 用类似方式, 序列 $\langle g_n \rangle$ 和它自己

的 m 重卷积的第 n 项等于

$$\sum_{k_1+k_2+\cdots+k_m=n} g_{k_1} g_{k_2} \cdots g_{k_m}$$

且它的母函数为 $G(z)^m$.

我们可把这些看到的结果应用到前面所研究的扇支撑问题上去(7.3 节中的例 6). 结果有另外一个方法来计算 n 级扇的支撑树个数 f_n , 基于顶点 $\{1, 2, \dots, n\}$ 之间三条边的构造: 顶点 k 和顶点 $k+1$ 之间的边关于子树可被选取或者不被选取, 并选这些边的每种方式向上连接邻近顶点的一些部分. 例如, 当 $n=10$ 时可连接顶点 $\{1, 2\}$, $\{3\}$, $\{4, 5, 6, 7\}$ 和 $\{8, 9, 10\}$:



通过加入附加的边到顶点 0 我们能作多少个支撑树? 我们需连接 0 到四部分的每一部分; 有 2 种方式连接 0 和 $\{1, 2\}$, 1 种方式连接 0 和 $\{3\}$, 4 种方式连接 0 和 $\{4, 5, 6, 7\}$, 3 种方式连接 0 和 $\{8, 9, 10\}$, 或者总共 $2 \cdot 1 \cdot 4 \cdot 3 = 24$ 种方式. 为了形成部分而在所有可能方式上求和给出下列支撑树总数的表达式:

$$f_n = \sum_{m \geq 0} \sum_{\substack{k_1+k_2+\cdots+k_m=n \\ k_1, k_2, \dots, k_m \geq 0}} k_1 k_2 \cdots k_m. \quad (7.64)$$

例如, $f_4 = 4 + 3 \cdot 1 + 2 \cdot 2 + 1 \cdot 3 + 2 \cdot 1 \cdot 1 + 1 \cdot 2 \cdot 1 + 1 \cdot 1 \cdot 2 + 1 \cdot 1 \cdot 1 \cdot 1 = 21$.

对于 $m=1, 2, 3, \dots$ 这是序列 $\langle 0, 1, 2, 3, \dots \rangle$ 的 m 重卷积的和; 因此 $\langle f_n \rangle$ 的母函数是

$$F(z) = G(z) + G(z)^2 + G(z)^3 + \cdots = \frac{G(z)}{1 - G(z)},$$

其中 $G(z)$ 是 $\langle 0, 1, 2, 3, \dots \rangle$ 的母函数, 即 $z/(1-z)^2$. 因而像前面那样我们得到

$$F(z) = \frac{z}{(1-z)^2 - z} = \frac{z}{1-3z+z^2}.$$

这种研究 $\langle f_n \rangle$ 的方法比以前我们得到的复杂递归较对称和吸引人.

例 4: 一个卷积递归.

下一个例子特别重要, 事实上, 它是母函数用在递归的解方面的“经典例子”.

假设我们有 $n+1$ 个变量 x_0, x_1, \dots, x_n . 通过 n 次相乘计算它的乘积. 把括号插入乘积 $x_0 \cdot x_1 \cdot \dots \cdot x_n$ 以致完全指定相乘的次序, 问有多少种方式? 例如, 当 $n=2$ 有 2 种方式, $x_0 \cdot (x_1 \cdot x_2)$ 和 $(x_0 \cdot x_1) \cdot x_2$. 且当 $n=3$ 时有 5 种方式,

$$x_0 \cdot (x_1 \cdot (x_2 \cdot x_3)), x_0 \cdot ((x_1 \cdot x_2) \cdot x_3), (x_0 \cdot x_1) \cdot (x_2 \cdot x_3), \\ (x_0 \cdot (x_1 \cdot x_2)) \cdot x_3, ((x_0 \cdot x_1) \cdot x_2) \cdot x_3.$$

因此 $C_2 = 2$, $C_3 = 5$, 我们还有 $C_1 = 1$ 和 $C_0 = 1$.

让我们用 7.3 节的 4 步过程, C 的递归是什么? 关键的观察看到当 $n > 0$ 时恰有一个 '·' 运算在所有括号外面, 这是最后的相乘把一切联系在一起. 如果这个 '·' 出现在 x_k 和 x_{k+1} 之间, 则有 C_k 种方式完全把括号插入 $x_0 \cdot \dots \cdot x_k$, 有 C_{n-k-1} 种方式完全把括号插入 $x_{k+1} \cdot \dots \cdot x_n$; 因此

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \dots + C_{n-1} C_0, \text{ 如果 } n > 0.$$

现在我们认出了这个表达式像一个卷积, 并且知道如何来修补公式以致对所有整数 n 成立:

$$C_n = \sum_k C_k C_{n-1-k} + [n=0]. \quad (7.65)$$

第 1 步现在完成了. 第 2 步让我们用 z^n 乘且相加:

$$C(z) = \sum_n C_n z^n = \sum_{k,n} C_k C_{n-1-k} z^n + \sum_{n=0} z^n = \sum_k C_k z^k \sum_n C_{n-1-k} z^{n-k} + 1 \\ = C(z) \cdot z C(z) + 1$$

嗨, 你瞧, 关于母函数, 卷积变成了一个乘积. 事情十分意外.

第 3 步也是容易的, 我们用二次公式解 $C(z)$:

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

但是我们应该选+号还是-号呢? 两种选取产生满足 $C(z) = zC(z)^2 + 1$ 的一个函数, 但是仅仅一种选取对问题是合适的. 依据独断的考虑可能选取+号是最好的; 但是我们立即发现这种选取给出 $C(0) = \infty$, 与事实矛盾. (假设正确的函数 $C(z)$ 有 $C(0) = C_0 = 1$.) 所以我们推得

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

最后, 第 4 步, $[z^n]C(z)$ 是什么? 二项定理告诉我们

$$\sqrt{1-4z} = \sum_{k=0} \binom{1/2}{k} (-4z)^k = 1 + \sum_{k=1} \frac{1}{2k} \binom{-1/2}{k-1} (-4z)^k;$$

图例. 图式(5.17).

$$\begin{aligned}\frac{1 - \sqrt{1 - 4z}}{2z} &= \sum_{k \geq 1} \frac{1}{k} \binom{-1/2}{k-1} (-4z)^{k-1} \\ &= \sum_{n \geq 0} \binom{-1/2}{n} \frac{(-4z)^n}{n+1} = \sum_{n \geq 0} \binom{2n}{n} \frac{z^n}{n+1}.\end{aligned}$$

插入括号的方式数 C_n 是 $\binom{2n}{n} \frac{1}{n+1}$.

当引入 Catalan 数的序列 $\langle 1, 1, 2, 5, 14, \dots \rangle = \langle C_n \rangle$ 时, 我们在第五章中提前用了这个结果. 在几十个问题中产生这个序列, 而这些问题起先看来是彼此无关的^[41], 因为许多场合具有对应于卷积递归(7.65)的一个递归结构.

例如, 让我们考虑下列问题: 有多少个+1和-1的序列 $\langle a_1, a_2, \dots, a_{2n} \rangle$ 有性质

$$a_1 + a_2 + \dots + a_{2n} = 0,$$

并具有它们的所有部分和

$$a_1, a_1 + a_2, \dots, a_1 + a_2 + \dots + a_{2n}$$

非负? 一定有 n 个+1和 n 个-1出现. 通过作为 n 的一个函数绘制部分和 $s_n = \sum_{k=1}^n a_k$ 的序列, 我们能以图形表示这个问题: $n=3$ 的 5 个解是:



这些是用形式“ \cdot ”和“ \cdot ”的线段绘制的宽 $2n$ 的“山脉”. 结果恰有 C_n 种方式做这一点, 以下列方式序列能和括号问题联系起来: 围绕整个公式放一对外加的括号, 以致有 n 对括号对应于 n 个相乘. 现在用+1替换第一个“ \cdot ”, 用-1替换每一个“ \cdot ”, 并把其他一切除去. 例如, 由此规则公式 $x_0 \cdot ((x_1 \cdot x_2) \cdot (x_3 \cdot x_4))$ 对应于 $\langle +1, +1, -1, +1, +1, -1, -1, -1 \rangle$. 把括号插入 $x_0 \cdot x_1 \cdot x_2 \cdot x_3$ 的 5 种方式对应于上面表明的 $n=3$ 的 5 个山脉.

此外, 把序列计数问题稍微改造一下可导致避开用母函数的意外简单的组合解: 当所有部分和

$$a_0, a_0 + a_1, a_0 + a_1 + a_2, \dots, a_0 + a_1 + \dots + a_{2n}$$

要求为正时, 有多少个+1和-1的序列 $\langle a_0, a_1, a_2, \dots, a_{2n} \rangle$ 有性质

$a_0 + a_1 + a_2 + \dots + a_{2n} = 1$? 显然这些就是前面问题的序列, 具有附加元素 $a_0 = +1$ 放在前面. 但是使用 George Rancy^[243]在 1959 年发现的一个值得注意的事实, 通过简单的计数论证能枚举新问题中的序列: 如果 $\langle x_1, x_2, \dots, x_m \rangle$ 是任何整数序列, 它的和是+1, 恰好循环移位

$$\langle x_1, x_2, \dots, x_m \rangle, \langle x_2, \dots, x_m, x_1 \rangle, \dots, \langle x_m, x_1, \dots, x_{m-1} \rangle$$

之一，它的所有部分和是正的。例如，考虑序列 $\langle 3, -5, 2, -2, 3, 0 \rangle$ ，它的循环移位是

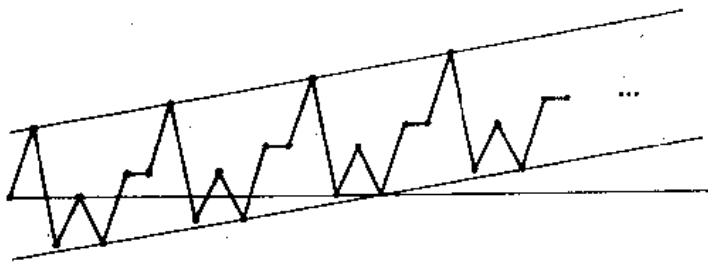
$$\begin{array}{ll} \langle 3, -5, 2, -2, 3, 0 \rangle & \langle -2, 3, 0, 3, -5, 2 \rangle \\ \langle -5, 2, -2, 3, 0, 3 \rangle & \langle 3, 0, 3, -5, 2, -2 \rangle \checkmark \\ \langle 2, -2, 3, 0, 3, -5 \rangle & \langle 0, 3, -5, 2, -2, 3 \rangle \end{array}$$

且仅查出一个有完全正的部分和。

用简单的几何论证能证明 Rancy 引理。让我们周期地推广序列取得一个无穷序列

$$\langle x_1, x_2, \dots, x_m, x_1, x_2, \dots, x_m, x_1, x_2, \dots \rangle,$$

因此对于所有 $k \geq 0$ 设 $x_{m+k} = x_k$ 。如果我们现在作为 n 的函数画出部分和 $s_n = x_1 + \dots + x_n$ ，则 s_n 的图有一个“平均斜率” $1/m$ ，因为 $s_{m+n} = s_n + 1$ 。例如，对应于我们开始时的例子序列 $\langle 3, -5, 2, -2, 3, 0, 3, -5, 2, \dots \rangle$ 的图如下：



如同表明的那样，整个图包含在斜率为 $1/m$ 的两条直线之间，在说明中 $m=6$ 。一般这些界限直线在 m 个点的每个循环中仅接触图一次，因为斜率 $1/m$ 的直线每 m 个单位长仅一次到达整数坐标的点。唯一的较低的交点是循环中仅有的位置，此处所有部分和将为正，因为曲线上每一个其他点对于它的右边在 m 个单位长之内有一个交点。

用 Rancy 引理容易枚举 $+1$ 和 -1 的序列 $\langle a_0, \dots, a_{2n} \rangle$ ，它的部分和全部为正且总和为 $+1$ 。出现 n 个 -1 和 $n+1$ 个 $+1$ 的序列有 $\binom{2n+1}{n}$ 个，而 Rancy 引理告诉我们恰好 $1/(2n+1)$ 个这样的序列所有部分和为正。(在一个 $N \times (2n+1)$ 的阵列中，列出所有 $N = \binom{2n+1}{n}$ 个这样的序列以及所有 $2n+1$ 个它们的循环移位。每一行恰好包含一个解。在每一列中每个解恰好出现一次。所在阵列中有 $N/(2n+1)$ 个不同的解，每个出现 $(2n+1)$ 次。)具有正部分和的序列总数为

$$\binom{2n+1}{n} \frac{1}{2n+1} = \binom{2n}{n} \frac{1}{n+1} = C_n.$$

例 5: n 重卷积的一个递归。

我们能推广刚才考虑的问题，通过研究 $+1$ 和 $(1-m)$ 的序列 $\langle a_0, \dots, a_{mn} \rangle$ ，它的部

分和全为正且总和为+1. 这样的序列称为 m -Raney 序列. 如果 k 个 $(1-m)$ 出现, $mn+1-k$ 个 $+1$ 出现, 我们得到

$$k(1-m) + (mn+1-k) = 1,$$

因此 $k=n$. 出现 n 个 $(1-m)$ 和 $mn+1-n$ 个 $+1$ 的序列有 $\binom{mn+1}{n}$ 个, 且 Raney 引理告诉我们, 所有部分和正的这种序列的个数恰好为

$$\binom{mn+1}{n} \frac{1}{mn+1} = \binom{mn}{n} \frac{1}{(m-1)n+1}. \quad (7.66)$$

所以这是 m -Raney 序列数. 让我们称这是 Fuss-Catalan 数 $C_n^{(m)}$, 因为是 N.I. Fuss^[109] 在 1791 年首先研究了序列 $\langle C_n^{(m)} \rangle$ (比 Catalan 早许多年). 通常的 Catalan 数是 $C_n = C_n^{(2)}$.

现在我们知道解答 (7.66), 让我们对付“危难”且解决一个导出它的问题. 在 $m=2$ 的情形, 问题是: “满足递归 $C_n = \sum_k C_k C_{n-1-k} + [n=0]$ 的数 C_n 是什么? 我们将尝试找一般情形的一个类似问题 (一个相似的递归).

长度 1 的平凡序列 $\langle +1 \rangle$ 显然是 m -Raney 序列. 如果把数 $(1-m)$ 放在任何 m 序列 (为 m -Raney) 的右边, 我们取得一个 m -Raney 序列; 当部分和增加到 $+2$, 然后 $+3, \dots, +m$ 和 $+1$ 时, 它们保持为正、相反, 如果 $n > 0$, 我们能表明以下列方式产生所有 m -Raney 序列 $\langle a_0, \dots, a_{mn} \rangle$: 最后项 a_{mn} 一定是 $(1-m)$. 对于 $1 \leq j \leq mn$, 部分和 $s_j = a_0 + \dots + a_{j-1}$ 是正, 且 $s_{mn} = m$ 因为 $s_{mn} + a_{mn} = 1$. 设 k_1 是 $\leq mn$ 的最大指标使得 $s_{k_1} = 1$, k_2 是使得 $s_{k_2} = 2$ 的最大指标, 等等. 因此对于 $k_j < k \leq mn$ 和 $1 \leq j \leq m$, $s_{k_j} = j$ 和 $s_k > j$. 由此得到 $k_m = mn$, 并能容易地验证每个子序列 $\langle a_0, \dots, a_{k_1-1} \rangle, \langle a_{k_1}, \dots, a_{k_2-1} \rangle, \dots, \langle a_{k_{m-1}}, \dots, a_{k_m-1} \rangle$ 是一个 m -Raney 序列, 对于某些非负整数 n_1, n_2, \dots, n_m , 一定有 $k_1 = mn_1 + 1, k_2 - k_1 = mn_2 + 1, \dots, k_m - k_{m-1} = mn_m + 1$.

所以 $\binom{mn+1}{n} 1/(mn+1)$ 是下列两个有趣问题的解答: “对于所有整数 n , 递归

$$C_n^{(m)} = \left(\sum_{n_1+n_2+\dots+n_m=n-1} C_{n_1}^{(m)} C_{n_2}^{(m)} \dots C_{n_m}^{(m)} \right) + [n=0] \quad (7.67)$$

所定义的数 $C_n^{(m)}$ 是什么?” “如果 $G(Z)$ 是满足

$$G(Z) = zG(z)^m + 1 \quad (7.68)$$

的一个幂级数, $[z^n]G(z)$ 是什么?

注意这些不是容易的问题. 通常在 Catalan 情形 ($m=2$), 我们用二次公式和二项定理解式 (7.68) 求出 $G(z)$ 和它的系数; 但是当 $m=3$ 时没有标准技巧给出如何解三次方程 $G = zG^3 + 1$ 的任何思路. 所以在询问它之前把它转成较易回答这个问题的形式.

然而现在我们知道足以询问较难的问题且导出它们的解答。下面一个问题怎么样:“如果 l 是正整数并且 $G(z)$ 是式(7.68)定义的幂级数, $[z^n]G(z)^l$ 是什么?”用我们刚才给出的论证表明 $[z^n]G(z)^l$ 是长度 $mn+l$ 的序列数且有下列三个性质:

- 每个元素是+1 或者是 $(1-m)$ 。
- 部分和全为正的。
- 总和是 l 。

为了通过把具有 m -Raney 性质的 l 个序列放在一起而取得所有这样的序列。做这一点的方式数为

$$\sum_{n_1+n_2+\cdots+n_l=n} C_{n_1}^{(m)} C_{n_2}^{(m)} \cdots C_{n_l}^{(m)} = [z^n]G(z)^l.$$

Raney 证明了他的引理的推广, 由此我们知道如何计算这些序列: 如果 $\langle x_1, x_2, \dots, x_m \rangle$ 是任何整数序列, 对于所有 j , $x_j \leq 1$, 且具有 $x_1 + x_2 + \cdots + x_m = l > 0$, 则恰好 l 个循环移位

$$\langle x_1, x_2, \dots, x_m \rangle, \langle x_2, \dots, x_m, x_1 \rangle, \dots, \langle x_m, x_1, \dots, x_{m-1} \rangle$$

有全为正的部分和。

例如, 我们能在序列 $\langle -2, 1, -1, 0, 1, 1, -1, 1, 1, 1 \rangle$ 上检验这个命题。循环移位是

$$\begin{array}{ll} \langle -2, 1, -1, 0, 1, 1, -1, 1, 1, 1 \rangle & \langle 1, -1, 1, 1, 1, -2, 1, -1, 0, 1 \rangle \\ \langle 1, -1, 0, 1, 1, -1, 1, 1, 1, -2 \rangle & \langle -1, 1, 1, 1, -2, 1, -1, 0, 1, 1 \rangle \\ \langle -1, 0, 1, 1, -1, 1, 1, 1, -2, 1 \rangle & \langle 1, 1, 1, -2, 1, -1, 0, 1, 1, -1 \rangle \checkmark \\ \langle 0, 1, 1, -1, 1, 1, 1, -2, 1, -1 \rangle & \langle 1, 1, -2, 1, -1, 0, 1, 1, -1, 1 \rangle \\ \langle 1, 1, -1, 1, 1, 1, -2, 1, -1, 0 \rangle \checkmark & \langle 1, -2, 1, -1, 0, 1, 1, -1, 1, 1 \rangle \end{array}$$

仅有两个标记‘ \checkmark ’的例子所有部分和为正。在习题 13 中证明这个推广的引理。

长度 $mn+l$ 且总和为 l 的 $+1$ 和 $(1-m)$ 的序列一定恰好有 n 个 $(1-m)$ 出现。推广的引理告诉我们这样的 $\binom{mn+l}{n}$ 个序列的 $l/(mn+l)$ 个其所有部分和为正; 因此我们的困难问题有一个意外的简单解答: 对所有整数 $l > 0$,

$$[z^n]G(z)^l = \binom{mn+l}{n} \frac{l}{mn+l} \quad (7.69)$$

没有忘记第五章的读者可能很有经验: “这公式看来熟悉, 前面我们见过它吗?” 确实见过, 方程(5.60)表明

$$[z^n]\mathcal{B}_r(z)^l = \binom{ln+r}{n} \frac{r}{ln+r}.$$

所以式(7.68)中的母函数 $G(z)$ 实际一定是广义二项级数 $\mathcal{B}_m(z)$ 。果然, 方程(5.59)表明

$$\mathcal{B}_m(z)^{1-m} - \mathcal{B}_m(z)^{-m} = z.$$

它就是

$$\mathcal{B}_m(z) - 1 = z \mathcal{B}_m(z)^m.$$

让我们变到第五章的记法, 现在论述广义二项式。第五章不加证明地叙述了一串等式。现在我们通过证明每当 t 和 r 是正整数时由

$$\mathcal{B}_t(z) = \sum_n \binom{tn+1}{n} \frac{z^n}{tn+1}$$

定义的幂级数 $\mathcal{B}_t(z)$ 有值得注意的性质

$$\mathcal{B}_t(z)' = \sum_n \binom{tn+1}{n} \frac{rz^n}{tn+r}$$

封闭了缺口部分。

我们能否推广这些结果到任意的 t 和 r 值? 是的, 因为系数 $\binom{tn+r}{r} \frac{r}{tn+r}$ 是 t 和 r 的多项式。由

$$\mathcal{B}_t(z)' = e^{r \ln \mathcal{B}_t(z)} = \sum_{n \geq 0} \frac{(r \ln \mathcal{B}_t(z))^n}{n!} = \sum_{n \geq 0} \frac{r^n}{n!} \left(- \sum_{m \geq 1} \frac{(1 - \mathcal{B}_t(z))^m}{m} \right)^n$$

定义的一般第 r 次幂的系数是 t 和 r 的多项式, 对无限多 t 和 r 的值这些多项式等于 $\binom{tn+r}{n} \frac{r}{tn+r}$ 。所以两个多项式的序列一定是恒等的。

第五章还提到广义指数级数

$$\mathcal{E}_t(z) = \sum_{n \geq 0} \frac{(tn+1)^{n-1}}{n!} z^n,$$

在式(5.60)中表明了它有一个相同值得注意的性质:

$$[z^n] \mathcal{E}_t(z)' = \frac{r(tn+r)^{n-1}}{n!} \quad (7.70)$$

我们能证明这一点作为 $\mathcal{B}_t(z)$ 的公式的极限情形, 因为证明

$$\mathcal{E}_t(z)' = \lim_{x \rightarrow \infty} \mathcal{B}_{xt}(z/x)^{xr}$$

并不困难。

7.6 指数型母函数

有时序列 $\langle g_n \rangle$ 所具有的母函数的性质十分复杂, 而有关序列 $\langle g_n / n! \rangle$ 所具有的母函数的性质十分简单. 在这种情形下, 我们自然宁愿选择 $\langle g_n / n! \rangle$ 来研究, 然后在最后乘 $n!$. 这种方法经常用, 我们给它一个特殊名称: 称幂级数

$$\hat{G}(z) = \sum_{n \geq 0} g_n \frac{z^n}{n!} \quad (7.71)$$

为序列 $\langle g_0, g_1, g_2, \dots \rangle$ 的指数型母函数. 因为指数函数 e^z 是 $\langle 1, 1, 1, \dots \rangle$ 的指数型母函数而产生这个名称.

表 7.3 中许多母函数实际是指数型母函数. 例如, 方程 (7.50) 表明 $(\ln \frac{1}{1-z})^m / m!$ 是序列 $\langle \begin{bmatrix} 0 \\ m \end{bmatrix}, \begin{bmatrix} 1 \\ m \end{bmatrix}, \begin{bmatrix} 2 \\ m \end{bmatrix}, \dots \rangle$ 的指数型母函数. 这个序列的通常的母函数是十分复杂的 (而且还是发散的).

指数型母函数有它们自身的基本操作特点, 相似于 7.2 节中学过的运算. 例如, 如果用 z 乘 $\langle g_n \rangle$ 的指数型母函数, 我们得到

$$\sum_{n \geq 0} g_n \frac{z^{n+1}}{n!} = \sum_{n \geq 1} g_{n-1} \frac{z^n}{(n-1)!} = \sum_{n \geq 0} n g_{n-1} \frac{z^n}{n!},$$

这是 $\langle 0, g_0, 2g_1, \dots \rangle = \langle n g_{n-1} \rangle$ 的指数型母函数.

$\langle g_0, g_1, g_2, \dots \rangle$ 的指数型母函数对 z 求导数给出

$$\sum_{n \geq 0} n g_n \frac{z^{n-1}}{n!} = \sum_{n \geq 1} g_n \frac{z^{n-1}}{(n-1)!} = \sum_{n \geq 0} g_{n+1} \frac{z^n}{n!}, \quad (7.72)$$

这是 $\langle g_1, g_2, \dots \rangle$ 的指数型母函数. 指数型母函数的微分对应于通常母函数的左移运算 $(G(z) - g_0) / z$. (在我们研究超几何级数时, 用了这个左移性质, 式 (5.106).) 一个指数型母函数的积分给出

$$\int_0^z \sum_{n \geq 0} g_n \frac{t^n}{n!} dt = \sum_{n \geq 0} g_n \frac{z^{n+1}}{(n+1)!} = \sum_{n \geq 1} g_{n-1} \frac{z^n}{n!}, \quad (7.73)$$

这是一个右移, $\langle 0, g_0, g_1, \dots \rangle$ 的指数型母函数.

如同通常母函数的运算那样, 指数型母函数的极有趣的运算是乘积. 如果 $\hat{F}(z)$ 和 $\hat{G}(z)$ 是 $\langle f_n \rangle$ 和 $\langle g_n \rangle$ 的指数型母函数, 则 $\hat{F}(z)\hat{G}(z) = \hat{H}(z)$ 是称为 $\langle f_n \rangle$ 和 $\langle g_n \rangle$ 的二项卷积 $\langle h_n \rangle$ 的指数型母函数:

$$h_n = \sum_k \binom{n}{k} f_k g_{n-k}. \quad (7.74)$$

这里出现二项系数, 因为 $\binom{n}{k} = n! / k!(n-k)!$, 因此

$$\frac{h_n}{n!} = \sum_{k=0}^n \frac{f_k}{k!} \frac{g_{n-k}}{(n-k)!};$$

换句话说, $\langle h_n / n! \rangle$ 是 $\langle f_n / n! \rangle$ 和 $\langle g_n / n! \rangle$ 的通常的卷积。

在应用中常常出现二项卷积。例如, 在式(6.79)中我们用隐含的递归式

$$\sum_{j=0}^m \binom{m+1}{j} B_j = [m=0], \text{ 所有 } m \geq 0$$

定义 Bernoulli 数; 这能改写为二项卷积, 如果我们用 $m+1$ 替换 n , 并且在两边加一项 B_n :

$$\sum_k \binom{n}{k} B_k = B_n + [n=1], \text{ 所有 } n \geq 0. \quad (7.75)$$

现在我们能通过引入 Bernoulli 数的指数型母函数 $\hat{B}(z) = \sum_{n \geq 0} B_n z^n / n!$ 把这个递归和幂级数联系起来(如同第六章中断定的那样)。式(7.75)的左边是 $\langle B_n \rangle$ 和常数序列 $\langle 1, 1, 1, \dots \rangle$ 的二项卷积, 因此左边的指数型母函数为 $\hat{B}(z)e^z$ 。右边的指数型母函数是 $\sum_{n \geq 0} (B_n + [n=1])z^n / n! = \hat{B}(z) + z$ 。所以我们一定有 $\hat{B}(z) = z / (e^z - 1)$; 我们证明了方程(6.81), 如同方程(7.44)那样, 它也出现在表 7.3 中。

现在让我们再考虑本书中常常出现的一个和

$$S_m(n) = 0^m + 1^m + 2^m + \dots + (n-1)^m = \sum_{0 \leq k < n} k^m.$$

这次我们将尝试分析母函数的问题, 希望它将意外地变得较简单。我们将考虑 n 是指定的, m 是变量; 因此我们的目标是推出幂级数

$$S(z) = S_0(n) + S_1(n)z + S_2(n)z^2 + \dots = \sum_{m \geq 0} S_m(n)z^m$$

的系数。我们知道 $\langle 1, k, k^2, \dots \rangle$ 的母函数是

$$\frac{1}{1-kz} = \sum_{m \geq 0} k^m z^m,$$

因此通过交换求和的次序

$$S(z) = \sum_{m \geq 0} \sum_{0 \leq k < n} k^m z^m = \sum_{0 \leq k < n} \frac{1}{1-kz}.$$

我们能将此和变成闭形式,

$$\begin{aligned} S(z) &= \frac{1}{z} \left(\frac{1}{z^{-1}-0} + \frac{1}{z^{-1}-1} + \cdots + \frac{1}{z^{-1}-n+1} \right) \\ &= \frac{1}{z} (H_{z^{-1}} - H_{z^{-1}-n}); \end{aligned} \quad (7.76)$$

但是我们不知道把这样一个闭形式展成 z 的幂.

指数型母函数来援救. 序列 $\langle S_0(n), S_1(n), S_2(n), \cdots \rangle$ 的指数型母函数是

$$\hat{S}(z, n) = S_0(n) + S_1(n) \frac{z}{1!} + S_2(n) \frac{z^2}{2!} + \cdots = \sum_{m \geq 0} S_m(n) \frac{z^m}{m!}.$$

为了取得这些系数 $S_m(n)$, 我们能用 $\langle 1, k, k^2, \cdots \rangle$ 的指数型母函数, 即

$$e^{kz} = \sum_{m \geq 0} k^m \frac{z^m}{m!},$$

我们得到

$$\hat{S}(z, n) = \sum_{m \geq 0} \sum_{0 \leq k \leq n} k^m \frac{z^m}{m!} = \sum_{0 \leq k \leq n} e^{kz}.$$

后面的和是等比级数, 所以有闭形式

$$\hat{S}(z, n) = \frac{e^{nz} - 1}{e^z - 1}. \quad (7.77)$$

我们只需计算出这个比较简单的函数的系数, 并且我们将知道 $S_m(n)$, 因为 $S_m(n) = m! [z^m] \hat{S}(z, n)$.

这里牵涉到 Bernoulli 数. 刚才我们看到 Bernoulli 数的指数型母函数是

$$\hat{B}(z) = \sum_{k \geq 0} B_k \frac{z^k}{k!} = \frac{z}{e^z - 1};$$

因此我们能写下

$$\begin{aligned} \hat{S}(z) &= \hat{B}(z) \frac{e^{nz} - 1}{z} \\ &= (B_0 \frac{z^0}{0!} + B_1 \frac{z^1}{1!} + B_2 \frac{z^2}{2!} + \cdots) (n \frac{z^0}{1!} + n^2 \frac{z^1}{2!} + n^3 \frac{z^2}{3!} + \cdots). \end{aligned}$$

和 $S_m(n)$ 是此乘积中 $m!$ 乘 z^m 的系数。例如,

$$\begin{aligned} S_0(n) &= 0! \left(B_0 \frac{n}{1!0!} \right) &&= n; \\ S_1(n) &= 1! \left(B_0 \frac{n^2}{2!0!} + B_1 \frac{n}{1!1!} \right) &&= \frac{1}{2}n^2 - \frac{1}{2}n; \\ S_2(n) &= 2! \left(B_0 \frac{n^3}{3!0!} + B_1 \frac{n^2}{2!1!} + B_2 \frac{n}{1!2!} \right) &&= \frac{1}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{6}n. \end{aligned}$$

所以我们又一次推出公式 $\square_n = S_2(n) = \frac{1}{3}n(n - \frac{1}{2})(n - 1)$, 并且这是最简单的推导: 几行之后我们就可找到 $S_m(n)$ 的一般情况(对所有 m).

一般公式可写成

$$S_{m-1}(n) = \frac{1}{m} (B_m(n) - B_m(0)), \quad (7.78)$$

其中 $B_m(x)$ 是由

$$B_m(x) = \sum_k \binom{m}{k} B_k x^{m-k}$$

定义的 Bernoulli 多项式。

这里是理由: Bernoulli 多项式是序列 $\langle B_0, B_1, B_2, \dots \rangle$ 和 $\langle 1, x, x^2, \dots \rangle$ 的二项卷积, 因此 $\langle B_0(x), B_1(x), B_2(x), \dots \rangle$ 的指数型母函数是它们的指数型母函数的乘积,

$$\hat{B}(z, x) = \sum_{m \geq 0} B_m(x) \frac{z^m}{m!} = \frac{z}{e^z - 1} \sum_{m \geq 0} x^m \frac{z^m}{m!} = \frac{ze^{xz}}{e^z - 1}. \quad (7.80)$$

因为依据式(7.77), $\langle 0, S_0(n), 2S_1(n), \dots \rangle$ 的指数型母函数是

$$z \frac{e^{nz} - 1}{e^z - 1} = \hat{B}(z, n) - \hat{B}(z, 0)$$

而得到方程(7.78)。

现在让我们转到另外一个指数型母函数的问题: n 个顶点 $\{1, 2, \dots, n\}$ 的完全图中可能有多少棵生成树? 让我们称此数为 t_n 。完全图有 $\frac{1}{2}n(n-1)$ 条边, 每条边连接每对不同的顶点; 所以我们实质上是寻找在 n 个给定顶点间画 $n-1$ 条线来联结它们的总方式数。

我们有 $t_1 = t_2 = 1$, 还有 $t_3 = 3$, 因为 3 个顶点的完全图是 2 阶扇形, 我们知道 $f_2 = 3$ 。当 $n=4$ 时有 16 棵生成树:

$$\begin{array}{cccc} \text{K} \times \Sigma & \Pi \nabla \square & N \times N & \\ \text{U} \sqcup & \times \times & \square \Sigma & \searrow \end{array}$$

(7.81)

因此 $t_4 = 16$.

扇形的相似问题的经验提供了处理这个问题的最好方法, 就是选出 1 个顶点, 且查看当忽略接触特殊顶点的所有边时生成树连接在一起的部分. 如果非特殊顶点形成了大小为 k_1, k_2, \dots, k_m 的 m 个部分, 则有 $k_1 k_2 \dots k_m$ 种方式把它们连接到特殊顶点. 例如, 在 $n=4$ 的情形中, 我们把下左顶点考虑为特殊顶点. (7.81) 的上面一排表明 $3t_3$ 种情形, 在那里其他 3 个顶点有 t_3 种方式连接它们, 且有 3 种方式连接下左顶点. 下面一排表明 $2 \cdot 1 \times t_2 t_1 \times \binom{3}{2}$ 个解, 在那里其他 3 个顶点用 $\binom{3}{2}$ 种方式被分成大小为 2 和 1 的部分;

还有情形 $\begin{array}{c} \bullet \\ \bullet \end{array}$, 在那里其他 3 个顶点完全不连接.

这个推理方式导致递归

$$t_n = \sum_{m \geq 0} \frac{1}{m!} \sum_{k_1 + k_2 + \dots + k_m = n-1} \binom{n-1}{k_1, k_2, \dots, k_m} k_1 k_2 \dots k_m t_{k_1} t_{k_2} \dots t_{k_m} \quad (\text{对所有 } n > 1).$$

这里是理由: 有 $\binom{n-1}{k_1, k_2, \dots, k_m}$ 种方式把 $n-1$ 个元素分配成大小分别为 k_1, k_2, \dots, k_m 的 m 个部分的一个序列; 有 $t_{k_1} t_{k_2} \dots t_{k_m}$ 种方式用生成树连接这些单独的部分; 有 $k_1 k_2 \dots k_m$ 种方式把顶点 n 连接到这些部分; 且我们除以 $m!$, 因为我们不考虑部分的次序. 例如, 当 $n=4$ 时递归表明

$$t_4 = 3t_3 + \frac{1}{2} \left(\binom{3}{1,2} 2t_1 t_2 + \binom{3}{2,1} 2t_2 t_1 \right) + \frac{1}{6} \left(\binom{3}{1,1,1} t_1^3 \right) = 3t_3 + 6t_2 t_1 + t_1^3.$$

开始时 t_n 的递归看来难对付, 甚至或许使人吃惊; 但是它确实不错, 仅仅复杂一些. 我们能定义

$$u_n = n! t_n$$

并有相当大的简化:

$$\frac{u_n}{n!} = \sum_{m \geq 0} \frac{1}{m!} \sum_{k_1 + k_2 + \dots + k_m = n-1} \frac{u_{k_1}}{k_1!} \frac{u_{k_2}}{k_2!} \dots \frac{u_{k_m}}{k_m!}, \quad n > 1. \quad (7.82)$$

内和是指数型母函数 $\hat{U}(z)$ 中 z^{n-1} 的系数, 上升到 m 次幂; 如果把对应于情形 $m=0$ 的项 $\hat{U}(z)^0$ 包括在内, 当 $n=1$ 时我们也得到正确的公式, 所以对于所有 $n > 0$,

$$\frac{u_n}{n!} = [z^{n-1}] \sum_{m \geq 0} \frac{1}{m!} \hat{U}(z)^m = [z^{n-1}] e^{\hat{U}(z)} = [z^n] z e^{\hat{U}(z)}$$

且我们有方程

$$\hat{U}(z) = z e^{\hat{U}(z)}. \quad (7.83)$$

方程(7.83)差不多像

$$\mathcal{G}(z) = e^{z \mathcal{G}(z)},$$

它定义了式(5.59)和(7.70)中的广义指数级数 $\mathcal{G}(z) = \mathcal{G}_1(z)$; 事实上, 我们有

$$\hat{U}(z) = z \mathcal{G}(z).$$

所以我们能获得问题的解答:

$$t_n = \frac{u_n}{n} = \frac{n!}{n} [z^n] \hat{U}(z) = (n-1)! [z^{n-1}] \mathcal{G}(z) = n^{n-2}. \quad (7.84)$$

对于所有 $n > 0$, $\{1, 2, \dots, n\}$ 上的完全图恰好有 n^{n-2} 棵生成树。

7.7 Dirichlet 母函数

有许多其他的可能方式从一个级数生成一个序列, 至少原则上能用任何“核”函数 $K_n(z)$ 系使得

$$\sum_n g_n K_n(z) = 0 \Rightarrow g_n = 0, \text{ 对于所有 } n.$$

通常母函数用 $K_n(z) = z^n$, 指数型母函数用 $K_n(z) = z^n / n!$; 我们也能试用下降阶乘幂 $z^{\underline{n}}$, 或二项系数 $z^n / n! = \binom{z}{n}$ 。

母函数和指数型母函数的最重要的替换核函数是用 $1/n^s$; 打算用于序列 $\langle g_1, g_2, \dots \rangle$, 序列从 $n=1$ 开始而不是 $n=0$:

$$\tilde{G}(z) = \sum_{n \geq 1} \frac{g_n}{n^s}. \quad (7.85)$$

称它为 **Dirichlet 母函数**, 因为德国数学家 Gustav Lejeune Dirichlet (1805–1859) 对它做了许多研究。

例如, 常数序列 $\langle 1, 1, 1, \dots \rangle$ 的 Dirichlet 母函数是

$$\sum_{n>1} \frac{1}{n^z} = \zeta(z). \quad (7.86)$$

这是 Riemann zeta 函数, 我们也称它为广义调和数 $H_{\infty}^{(z)}$ (当 $z > 1$ 时).

Dirichlet 母函数的乘积对应于一种特殊类型的卷积:

$$\tilde{F}(z)\tilde{G}(z) = \sum_{l,m>1} \frac{f_l}{l^z} \frac{g_m}{m^z} = \sum_{n>1} \frac{1}{n^z} \sum_{l,m>1} f_l g_m [l \cdot m = n].$$

因此 $\tilde{F}(z)\tilde{G}(z) = \tilde{H}(z)$ 是序列

$$h_n = \sum_{d|n} f_d g_{n/d} \quad (7.87)$$

的 Dirichlet 母函数.

例如, 我们从式(4.55)知道 $\sum_{d|n} \mu(d) = [n=1]$, 这是 Mobius 序列 $\langle \mu(1), \mu(2), \mu(3), \dots \rangle$ 和 $\langle 1, 1, 1, \dots \rangle$ 的 Dirichlet 卷积, 因此

$$\tilde{M}(z)\zeta(z) = \sum_{n>1} \frac{[n=1]}{n^z} = 1. \quad (7.88)$$

换句话说, $\langle \mu(1), \mu(2), \mu(3), \dots \rangle$ 的 Dirichlet 母函数是 $\zeta(z)^{-1}$.

当序列 $\langle g_1, g_2, \dots \rangle$ 是积性函数时, 即当

$$g_{mn} = g_m g_n \quad (m \perp n)$$

时, Dirichlet 母函数特别有价值.

在这样的情形中, 当 n 是一个素数的幂时, 由 g_n 的值确定所有 n 的 g_n 的值, 并能把 Dirichlet 母函数分解成素数上的一个乘积:

$$\tilde{G}(z) = \prod_{p \text{ 素数}} \left(1 + \frac{g_p}{p^z} + \frac{g_{p^2}}{p^{2z}} + \frac{g_{p^3}}{p^{3z}} + \dots \right). \quad (7.89)$$

例如, 若对于所有 n 置 $g_n = 1$, 我们获得一个 Riemann zeta 函数的乘积表示:

$$\zeta(z) = \prod_{p \text{ 素数}} \left(\frac{1}{1 - p^{-z}} \right). \quad (7.90)$$

Möbius 函数有 $\mu(p) = 1$, 对于 $k > 1$ $\mu(p^k) = 0$, 因此它的 Dirichlet 母函数是

$$\tilde{M}(z) = \prod_{p \text{ 素数}} (1 - p^{-z}); \quad (7.91)$$

当然, 这与式(7.88)和(7.90)一致. Euler φ 函数有 $\varphi(p^k) = p^k - p^{k-1}$, 因为它的 Dirichlet

母函数有因子分解的形式

$$\tilde{\Phi}(z) = \prod_{p \text{ 素数}} \left(1 + \frac{p-1}{p^z - p} \right) = \prod_{p \text{ 素数}} \frac{1-p^{-z}}{1-p^{1-z}}. \quad (7.92)$$

我们推出 $\tilde{\Phi}(z) = \zeta(z-1)/\zeta(z)$ 。

习 题

准备部分

1. 一个古怪的收藏者，他收藏 $2 \times n$ 多米诺骨牌，对于每个垂直的多米诺骨牌付 \$4，对于每个水平的多米诺骨牌付 \$1。依据这个准则有多少个 $2 \times n$ 多米诺骨牌恰好价值为 \$ m ？例如，当 $m=6$ 时，有三个解： $\begin{smallmatrix} \text{图1} \\ \text{图2} \\ \text{图3} \end{smallmatrix}$ 和 \dots 。

2. 以闭形式给出序列 $\langle 2, 5, 13, 35, \dots \rangle = \langle 2^n + 3^n \rangle$ 的母函数和指数型母函数。

3. $\sum_{n \geq 0} H_n / 10^n$ 是什么？

4. 有理函数 $P(z)/Q(z)$ 的一般展开定理不完全一般，因为它限制 P 的次数小于 Q 的次数。如果 P 具有比这个更大的次数将出现什么情况？

5. 找一个母函数 $S(z)$ 使得

$$[z^n]S(z) = \sum_k \binom{r}{k} \binom{r}{n-2k}.$$

基本部分

6. 不用母函数，证明依据所有组成部分的方法能解递归(7.32)。

7. 解递归

$$g_0 = 1;$$

$$g_n = g_{n-1} + 2g_{n-2} + \dots + ng_0, \quad n > 0.$$

8. $[z^n](\ln(1-z))^2 / (1-z)^{m+1}$ 是什么？

9. 用前面习题的结果来计算 $\sum_{k=0}^n H_k H_{n-k}$ 。

10. 在等式(7.61)中置 $r=s=-1/2$ ，且利用像式(5.36)的方法移去所有出现的 $1/2$ ，你推出令人惊异的等式是什么？

11. 这个问题(问题的三部分是独立的)给出了母函数操作中的常规做法。我们假设

$$A(z) = \sum_n a_n z^n, \quad B(z) = \sum_n b_n z^n, \quad C(z) = \sum_n c_n z^n, \quad \text{且对于负的 } n \text{ 系数为零.}$$

(a) 如果 $c_n = \sum_{j+2k \leq n} a_j b_k$ ，依据 A 和 B 表达 C 。

(b) 如果 $nb_n = \sum_{k=0}^n 2^k a_k / (n-k)!$ ，依据 B 表达 A 。

(c) 如果 r 是实数, 且若 $a_n = \sum_{k=0}^n \binom{r+k}{k} b_{n-k}$, 依据 B 表达 A ; 然后用你的公式来找系数 $f_k(r)$, 使得 $b_n = \sum_{k=0}^n f_k(r) a_{n-k}$.

12. 把数 $\{1, 2, \dots, 2n\}$ 排成一个 $2 \times n$ 数组以致行和列从左到右和从上到下是上升次序, 问有多少种方式? 例如, 当 $n=5$ 时, 一个解为

$$\begin{pmatrix} 1 & 2 & 4 & 5 & 8 \\ 3 & 6 & 7 & 9 & 10 \end{pmatrix}$$

13. 证明式(7.69)前面陈述的 Raney 推广的引理.

14. 用指数型母函数解递归

$$g_0 = 0, \quad g_1 = 1,$$

$$g_n = -2ng_{n-1} + \sum_k \binom{n}{k} g_k g_{n-k}, \quad n > 1.$$

15. 把 n 个元素划分成子集的方式数是 Bell 数 b_n . 例如, $b_3 = 5$, 因为我们能以下列方式划分 $\{1, 2, 3\}$:

$$\{1, 2, 3\}; \{1, 2\} \cup \{3\}; \{1, 3\} \cup \{2\}; \{1\} \cup \{2, 3\}; \{1\} \cup \{2\} \cup \{3\}.$$

证明 $b_{n+1} = \sum_k \binom{n}{k} b_k b_{n-k}$, 且用此公式来找指数型母函数 $\sum_n b_n z^n / n!$ 的一个闭形式.

16. 由卷积公式

$$b_n = \sum_{k_1 + 2k_2 + \dots + nk_n = n} \binom{a_1 + k_1 - 1}{k_1} \binom{a_2 + k_2 - 1}{k_2} \dots \binom{a_n + k_n - 1}{k_n}$$

把两个序列 $\langle a_n \rangle$ 和 $\langle b_n \rangle$ 联系起来, 且 $a_0 = 0$ 和 $b_0 = 1$. 证明对应的母函数满足 $\ln B(z) = A(z) + \frac{1}{2} A(z^2) + \frac{1}{3} A(z^3) + \dots$.

17. 证明一个序列的指数型母函数通过公式

$$\int_0^\infty \hat{G}(zt) e^{-t} dt = G(z) \quad (\text{如果积分存在})$$

和通常母函数 $G(z)$ 联系起来.

18. 求下列序列的 Dirichlet 母函数:

$$(a) \quad g_n = \sqrt{n};$$

$$(b) \quad g_n = \ln n;$$

$$(c) \quad g_n = [n \text{ 是没有平方的数}].$$

依据 zeta 函数表达你的解答. (习题 4.13 中定义了没有平方的数.)

19. 每个级数 $F(z) = \sum_{n \geq 0} f_n z^n$ (其中 $f_0 = 1$) 定义一个由规则

$$F(z)^* = \sum_{n \geq 0} f_n(x) z^n, \text{ 其中 } f_n(1) = f_n \text{ 和 } f_n(0) = [n=0]$$

确定的多项式 $f_n(x)$ 的序列. $f_n(x)$ 一般有次数 n . 证明这样的多项式满足卷积公式

$$\sum_{k=0}^n f_k(x) f_{n-k}(y) = f_n(x+y);$$

$$(x+y) \sum_{k=0}^n k f_k(x) f_{n-k}(y) = x n f_n(x+y).$$

(表 5.5 和 6.8 中等式是此诀窍的特殊情形.)

20. 如果存在有限多个不全为零的多项式 $P_0(z), \dots, P_m(z)$, 使得

$$P_0(z)G(z) + P_1(z)G'(z) + \dots + P_m(z)G^{(m)}(z) = 0$$

则称幂级数 $G(z)$ 是可微分有限的幂级数. 如果存在有限多个不全为零的多项式 $p_0(z), \dots, p_m(z)$, 使得对于所有整数 $n \geq 0$

$$p_0(n)g_n + p_1(n)g_{n+1} + \dots + p_m(n)g_{n+m} = 0$$

则称数序列 $\langle g_0, g_1, g_2, \dots \rangle$ 是多项式递归的. 证明一个母函数是可微分有限的, 当且仅当它的系数序列是多项式递归的.

课外习题

21. 一个强盗抢劫银行, 且需要 10 元和 20 元票面的 \$ 500. 他还要求知道出纳员能给他的钱的方式数. 求母函数 $G(z)$, 对于这个母函数, 此数为 $[z^{500}]G(z)$, 并且求一个更紧凑的母函数 $\check{G}(z)$, 对于它此数为 $[z^{50}] \check{G}(z)$. 通过 (a) 用部分分式, (b) 用像式 (7.39) 的方法确定要求的方式数.

22. 设 P 是“三角剖分”多边形的所有方式的和:

$$P = \text{---} + \triangle + \square + \square + \text{---} + \text{---} + \text{---} + \text{---} + \text{---} + \dots$$

(第一项表示一个仅有两个顶点的退化多边形, 每一个其他的项表明一个多边形被分割为三角形. 例如, 5 边形能以 5 种方式三角剖分.) 在三角剖分的多边形 A 和 B 上定义一个“乘法”运算 $A \triangle B$ 以致方程

$$P = \text{---} + P \triangle P$$

成立. 然后用 ‘ z ’ 替换每个三角形; 关于把一个 n 边形分解成三角形的方式数, 这告诉你什么结果?

23. 用一个 $2 \times 2 \times n$ 的柱状物造 $2 \times 1 \times 1$ 的砖状物有多少种方式?

24. 当 $n \geq 3$ 时, 在一个 n -轮形图中(在一个圈中有 n 个“外层”顶点, 每个外层顶点和第 $n+1$ 个“中心”顶点相连接)有多少棵生成树?

25. 设 $m \geq 2$ 是一个整数, 作为 x 和 m 的一个函数, 序列 $\langle n \bmod m \rangle$ 的母函数的闭形式是什么? 使用这个母函数, 按照复数 $\omega = e^{2\pi i/m}$ 来表达 ' $n \bmod m$ '. (例如, 当 $m=2$ 时, 有 $\omega = -1$ 和 $n \bmod 2 = \frac{1}{2} - \frac{1}{2}(-1)^n$.)

26. 由递归

$$\mathcal{F}_0 = 0, \mathcal{F}_1 = 1,$$

$$\mathcal{F}_n = \mathcal{F}_{n-1} + \mathcal{F}_{n-2} + F_n, \quad n > 1$$

定义第二种 Fibonacci 数 $\langle \mathcal{F}_n \rangle$, 依据通常 Fibonacci 数 F_n 和 F_{n+1} 表达 \mathcal{F}_n .

27. $2 \times n$ 多米诺骨牌还能看成在 $2 \times n$ 的点组中画 n 条不相交线的一种方式:

1 2 2 1 2 1 1

如果我们将两个这样的模式叠加起来, 取得圈的集合, 因为两条线邻接每一个点, 例如, 如果上面的线和线

1 1 2 2 2 2

结合起来, 结果为

1 1 1 1 2 2

结合

1 1 2 2 2 1 1 和 1 2 2 1 2 2

也得到相同的圈的集合. 如果我们将方向分配给垂直线, 在第一个模式中用箭头上/下/上/下/...交错地进行, 在第二个模式中用箭头下/上/下/上/...交错地进行, 我们得到一种唯一的方式来再构造原来的模式. 例如,

1 2 2 1 2 1 1 + 1 1 2 2 2 2 = 1 1 1 1 2 2

所以这种定向圈模式数一定是 $T_n^2 = F_{n+1}^2$, 我们能够通过代数来证明这一点. 设 Q_n 是定向 $2 \times n$ 圈模式数. 求 Q_n 的递归, 用母函数解它, 且用代数方法推出 $Q_n = F_{n+1}^2$.

28. 式(7.39)中 $A(z)$ 的系数满足 $A_r + A_{r+10} + A_{r+20} + A_{r+30} = 100 (0 \leq r < 10)$, 寻找得出这一点的“简单的”解释.

29. Fibonacci 乘积的和

$$\sum_{n>0} \sum_{\substack{k_1+k_2+\dots+k_m=n \\ k_1, k_2, \dots, k_m > 0}} F_{k_1} F_{k_2} \dots F_{k_m}$$

是什么?

30. 如果母函数 $G(z) = 1/(1-\alpha z)(1-\beta z)$ 有部分分式分解 $a/(1-\alpha z) + b/(1-\beta z)$, $G(z)^n$ 的部分分式分解是什么?

31. 正整数 n 的什么函数 $g(n)$ 满足递归

$$\sum_{d|n} g(d)\varphi(n/d) = 1$$

其中 φ 是 Euler φ 函数?

32. 算术级数是整数的一个无限集合

$$\{an+b\} = \{b, a+b, 2a+b, 3a+b, \dots\}$$

算术级数集合 $\{a_1n+b_1\}, \dots, \{a_mn+b_m\}$ 称为一个正好覆盖, 如果每一个非负整数在级数中出现一次且仅出现一次. 例如, 三个级数 $\{2n\}, \{4n+1\}, \{4n+3\}$ 构成一个正好覆盖. 证明如果 $\{a_1n+b_1\}, \dots, \{a_mn+b_m\}$ 是一个正好覆盖, 使得 $2 \leq a_1 \leq \dots \leq a_m$, 则 $a_{m-1} = a_m$. 提示: 用母函数.

考查性问题

33. $[w^m z^n](\ln(1+z))/(1-wz)$ 是什么?

34. 求母函数 $\sum_{n \geq 0} G_n(z)w^n$ 的一个闭形式, 如果

$$G_n(z) = \sum_{k \leq n/m} \binom{n-mk}{k} z^{mk}.$$

(这里 m 是一个指定的正整数.)

35. 用两种方式计算和 $\sum_{0 < k < n} 1/k(n-k)$:

(a) 以部分分式展开被加数.

(b) 把和看作为卷积, 再用母函数.

36. 设 $A(z)$ 是 $\langle a_0, a_1, a_2, a_3, \dots \rangle$ 的母函数. 依据 A, z 和 m 来表达

$$\sum_n a_{\lfloor n/m \rfloor} z^n.$$

37. 设 a_n 是把正整数 n 写成 2 的幂的和的方式数, 不管次序. 例如, $a_4 = 4$, 因为 $4 = 2+2 = 2+1+1 = 1+1+1+1$. 按照惯例, 我们设 $a_0 = 1$, 设 $b_n = \sum_{k=0}^n a_k$ 是开始的 a 的累积和.

(a) 由小变大作 a 和 b 的表, 直到 $n=10$. 在你的表中能看到什么令人惊异的关系?

(还未证明它.)

(b) 把母函数 $A(z)$ 表达为一个无限乘积.

(c) 用(b)所得的表达式证明(a)的结果.

38. 求双重母函数

$$M(w, z) = \sum_{m, n \geq 0} \min(m, n) w^m z^n$$

的闭形式. 对于指定的 $m \geq 2$, 推广你的解答来获得

$$M(z_1, \dots, z_m) = \sum_{n_1, \dots, n_m \geq 0} \min(n_1, \dots, n_m) z_1^{n_1} \cdots z_m^{n_m}$$

的闭形式.

39. 给定正整数 m 和 n , 求

$$\sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} k_1 k_2 \cdots k_m \text{ 和 } \sum_{1 \leq k_1 \leq k_2 \leq \dots \leq k_m \leq n} k_1 k_2 \cdots k_m$$

的闭形式. (例如, 当 $m=2$ 和 $n=3$ 时, 和是 $1 \cdot 2 + 1 \cdot 3 + 2 \cdot 3$ 和 $1 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 + 2 \cdot 2 + 2 \cdot 3 + 3 \cdot 3$.) 提示: 在母函数 $(1+a_1 z) \cdots (1+a_n z)$ 和 $1 / (1-a_1 z) \cdots (1-a_n z)$ 中 z^m 的系数是什么?

40. 把 $\sum_k \binom{n}{k} (kF_{k-1} - F_k)(n-k)!$ 表为闭形式.

41. n 阶的上升-下降排列是整数 $\{1, 2, \dots, n\}$ 的一种排列 $a_1 a_2 \cdots a_n$, 它交错上升和下降:

$$a_1 < a_2 > a_3 < a_4 > \cdots$$

例如, 35142 是 5 阶上升-下降排列. 如果 A_n 表示 n 阶上升-下降排列的个数, 证明 $\langle A_n \rangle$ 的指数型母函数是 $(1+\sin z) / \cos z$.

42. 空间探测器发现火星上的有机物有 5 个符号组成的脱氧核糖核酸, 记为 (a, b, c, d, e) , 而不是地球上的脱氧核糖核酸中的 4 种成分. 在火星的脱氧核糖核酸的串中不连续出现 4 对 cd, ce, ed 和 ee , 但是没有被禁止的对的任何串是可能存在的. (因此 $bbcd a$ 是被禁止的, 但是 $bbdca$ 是可以的.) 可能存在多少种长度 n 的火星脱氧核糖核酸串? (当 $n=2$ 时, 解答是 21, 因为串的左端和右端是可辨别的.)

43. 序列 $\langle g_n \rangle$ 的 Newton 母函数定义为

$$\hat{G}(z) = \sum_n g_n \binom{z}{n}.$$

求确定序列 $\langle f_n \rangle$, $\langle g_n \rangle$ 和 $\langle h_n \rangle$ 之间关系的一个卷积公式, 序列的 Newton 母函数由方程 $\hat{F}(z)\hat{G}(z) = \hat{H}(z)$ 联系起来. 试尽可能简单和对称地作出你的公式.

44. 当 n 个数 $\{x_1, \dots, x_n\}$ 彼此比较时, 设 q_n 是可能的结果数. 例如, $q_3 = 13$, 因为可能的结果是:

$$\begin{aligned} x_1 < x_2 < x_3; & \quad x_1 < x_2 = x_3; & \quad x_1 < x_3 < x_2; & \quad x_1 = x_2 < x_3; \\ x_1 = x_2 = x_3; & \quad x_1 = x_3 < x_2; & \quad x_2 < x_1 < x_3; \\ x_2 < x_1 = x_3; & \quad x_2 < x_3 < x_1; & \quad x_2 = x_3 < x_1; \\ x_3 < x_1 < x_2; & \quad x_3 < x_1 = x_2; & \quad x_3 < x_2 < x_1; \end{aligned}$$

求指数型母函数 $\hat{Q}(z) = \sum_n g_n z^n / n!$ 的一个闭形式, 并求序列 $\langle a_n \rangle$, $\langle b_n \rangle$, $\langle c_n \rangle$ 使得

$$q_n = \sum_{k \geq 0} k^n a_k = \sum_k \left\{ \begin{matrix} n \\ k \end{matrix} \right\} b_k = \sum_k \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle c_k, \text{ 所有 } n \geq 0.$$

$$45. \text{ 计算 } \sum_{m, n \geq 0} [m, n] / m^2 n^2.$$

46. 把

$$\sum_{0 \leq k \leq n/2} \binom{n-2k}{k} \left(\frac{-4}{27} \right)^k$$

化成闭形式. 提示: $z^3 - z^2 + \frac{4}{27} = (z + \frac{1}{3})(z - \frac{2}{3})^2$.

47. 证明式(7.34)中给定的 $3 \times n$ 多米诺骨牌的数 U_n 和 V_n 与 Stern-Brocot 树中收敛到 $\sqrt{3}$ 的部分紧密相关.

48. 某一个序列 $\langle g_n \rangle$ 对于整数 (a, b, c, d) 且 $\gcd(a, b, c, d) = 1$ 满足递归

$$ag_n + bg_{n+1} + cg_{n+2} + d = 0, \text{ 整数 } n \geq 0.$$

对于某个在 0 和 1 之间的实数 α , 它还有闭形式

$$g_n = \lfloor \alpha(1 + \sqrt{2})^n \rfloor, \text{ 整数 } n \geq 0.$$

求 a, b, c, d 和 α .

49. 这是关于幂和奇偶性的问题.

(a) 考虑由公式

$$a_n = (1 + \sqrt{2})^n + (1 - \sqrt{2})^n$$

定义的序列 $\langle a_0, a_1, a_2, \dots \rangle = \langle 2, 2, 6, \dots \rangle$. 求这个序列满足的简单的递归关系.

(b) 证明对于所有整数 $n > 0$, $\lceil (1 + \sqrt{2})^n \rceil \equiv n \pmod{2}$.

(c) 求形式为 $(p + \sqrt{q})/2$ 的一个数 α , 其中 p 和 q 是正整数, 使得对于所有整数

$n > 0, \lfloor \alpha^n \rfloor \equiv n \pmod{2}.$

额外问题

50. 接习题 22, 考虑把多边形分解为多边形的方式之和:

$$Q = \text{---} + \triangle + \square + \begin{array}{|c|} \hline \diagup \\ \hline \end{array} + \begin{array}{|c|} \hline \diagdown \\ \hline \end{array} \\ + \text{五边形} + \text{带一条对角线的五边形} + \text{带两条对角线的五边形} + \dots$$

求 Q 的符号方程, 并用它来求凸 n 边形内的画不相交对角线的方式数的母函数。(作为 z 的函数给出母函数的闭形式, 你不需作系数的闭形式。)

51. 证明乘积

$$2^{mn/2} \prod_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}} \left(\left(\cos^2 \frac{j\pi}{m+1} \right) \boxplus^2 + \left(\cos^2 \frac{k\pi}{n+1} \right) \boxminus^2 \right)$$

是 $m \times n$ 矩形的多米诺骨牌的母函数。(有 mn 个因子, 我们能想象把它写成矩形的 mn 个单元。如果 mn 是奇数, 则中间的因子为零。 $\boxplus^j \boxminus^k$ 的系数是作 j 个垂直的多米诺骨牌和 k 个水平的多米诺骨牌的方式数。) 提示: 这是一个困难的问题, 确实超出了本书的范围。你可就情形 $m=3, n=4$ 简单地验证公式。

52. 证明由递归

$$p_n(y) = \left(y - \frac{1}{4}\right)^n - \sum_{k=0}^{n-1} \binom{2n}{2k} \left(\frac{-1}{4}\right)^{n-k} p_k(y), \text{ 整数 } n \geq 0$$

定义的多项式具有形式 $p_n(y) = \sum_{m=0}^n \left| \frac{n}{m} \right| y^m$, 其中 $\left| \frac{n}{m} \right|$ 是一个正整数 ($1 \leq m \leq n$)。提示: 这个习题是非常有启发的, 但是不很容易。

53. 以一种明显的方式把 5 边形数序列 $\langle 1, 5, 12, 22, \dots \rangle$ 推广到三角数和平方数:



设第 n 个三角数是 $T_n = n(n+1)/2$, 设第 n 个 5 边形数是 $P_n = n(3n-1)/2$, 并设 U_n 是式 (7.38) 中定义的 $3 \times n$ 多米诺骨牌数。证明三角数 $T_{(U_{2n+2}-1)/2}$ 也是一个 5 边形数。提示:

$$3U_{2n}^2 = (V_{2n-1} + V_{2n+1})^2 + 2.$$

54. 考虑下列古怪的构造:

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				1 024 ...

(从一行包含所有正整数开始，然后删除一切的第 m 列，这里 $m=5$ ；然后用部分和替换剩下的项；然后删除一切的第 $(m-1)$ 列；然后再用部分和替换，等等。)用母函数证明最后的结果是第 m 次幂的序列。例如，当 $m=5$ 时，如同表明的那样我们得到 $\langle 1^5, 2^5, 3^5, 4^5, \dots \rangle$ 。

55. 证明如果幂级数 $F(z)$ 和 $G(z)$ 是可微分有限的(如同习题 20 中定义的那样)，则 $F(z)+G(z)$ 和 $F(z)G(z)$ 也是可微分有限的。

研究性问题

56. 证明在某一类“简单闭形式”中，作为 n 的函数，在 $(1+z+z^2)^n$ 中的 z^n 的系数没有“简单闭形式”。

57. 证明或推翻：如果所有 $G(z)$ 的系数是 0 或者 1，且若所有 $G(z)^2$ 的系数小于某个常数 M ，则无限多个 $G(z)^2$ 的系数为零。

第八章 离散概率

许多试图理解现实世界的问题带上或然率的成分。如果我们假设事件服从适当的公理，则允许用概率论来计算复杂事件的可能性，在科学的所有分支中此理论有重要的应用，且它和前面几章中已研究的技巧有强的联系。

如果用求和而不是积分能计算所有事件的概率，则称概率为“离散的”概率。我们对求和掌握得相当好，所以随时可以用我们的知识来做一些有趣的概率和平均值的计算，这是并不出乎意料之外的。

8.1 定义

概率论从概率空间的概念开始，它是一个给定的问题中能发生的所有事物的集合 Ω ，是把概率 $\Pr(\omega)$ 赋予每个基本事件 $\omega \in \Omega$ 的一个规则。概率 $\Pr(\omega)$ 一定是一个非负实数，且在每一个离散概率空间中条件

$$\sum_{\omega \in \Omega} \Pr(\omega) = 1 \quad (8.1)$$

一定成立。因此，每个值 $\Pr(\omega)$ 一定在区间 $[0..1]$ 中。我们把 \Pr 称为一个概率分布，因为它在事件 ω 中间分配总概率 1。

这里是一个例子：如果滚动一对骰子，则基本事件集合 Ω 是 $D^2 = \{\square\square, \square\blacksquare, \dots, \blacksquare\blacksquare\}$ 其中

$$D = \{\square, \blacksquare, \blacklozenge, \blacktriangle, \blacktriangledown, \boxtimes\}$$

是一个给定骰子能降落的所有 6 种方式的集合。两个滚动，譬如 $\square\blacksquare$ 和 $\blacksquare\square$ 认为是不同的，因此这个概率空间总共有 $6^2 = 36$ 个元素。

我们通常假设骰子是“完好的”，即一个指定骰子的 6 种可能结果的每一种具有概率 $1/6$ ，且 Ω 中 36 种可能滚动的每一种具有概率 $1/36$ 。但是我们也能考虑“灌过铅的”骰子，此时有一种不同的概率分布。例如，设

$$\Pr_1(\square) = \Pr_1(\blacksquare) = \frac{1}{4};$$

$$\Pr_1(\square) = \Pr_1(\blacksquare) = \Pr_1(\boxtimes) = \Pr_1(\boxdot) = \frac{1}{8}.$$

于是 $\sum_{d \in D} \Pr_1(d) = 1$, 所以 \Pr_1 是集合 D 上的一个概率分布, 且能通过规则

$$\Pr_{11}(dd') = \Pr_1(d)\Pr_1(d') \quad (8.2)$$

把概率赋予 $\Omega = D^2$ 的元素。例如, $\Pr_{11}(\boxtimes\blacksquare) = 1/4 \cdot 1/8 = 1/32$ 。这是一个正确的分布, 因为

$$\begin{aligned} \sum_{\omega \in \Omega} \Pr_{11}(\omega) &= \sum_{dd' \in D^2} \Pr_{11}(dd') = \sum_{d, d' \in D} \Pr_1(d)\Pr_1(d') \\ &= \sum_{d \in D} \Pr_1(d) \sum_{d' \in D} \Pr_1(d') = 1 \cdot 1 = 1. \end{aligned}$$

我们也能考虑一个完好骰子和一个灌铅骰子的情形,

$$\Pr_{01}(dd') = \Pr_0(d)\Pr_1(d'), \quad \text{其中 } \Pr_0(d) = \frac{1}{6}, \quad (8.3)$$

此时 $\Pr_{01}(\boxtimes\blacksquare) = 1/6 \cdot 1/8 = 1/48$ 。在“现实世界”中实际不能期望骰子每一面同等经常出现, 因为骰子不完全对称; 但是 $1/6$ 通常相当接近实际情况。

一个事件是 Ω 的一个子集。在骰子游戏中, 例如, 集合

$$\{\square\square, \square\blacksquare, \square\boxtimes, \square\boxdot, \blacksquare\square, \blacksquare\blacksquare, \blacksquare\boxtimes, \blacksquare\boxdot, \boxtimes\square, \boxtimes\blacksquare, \boxtimes\boxtimes, \boxtimes\boxdot, \boxdot\square, \boxdot\blacksquare, \boxdot\boxtimes, \boxdot\boxdot\}$$

是“掷成双”的事件。 Ω 的单独元素称为基本事件, 因为不能分解它们为较小的子集; 我们能把 ω 想象为一个元素事件 $\{\omega\}$ 。

由公式

$$\Pr(\omega \in A) = \sum_{\omega \in A} \Pr(\omega) \quad (8.4)$$

定义一个事件 A 的概率; 且一般如果 $R(\omega)$ 是关于 ω 的任何语句, 我们把使得 $R(\omega)$ 为真的所有 $\Pr(\omega)$ 的和记为 ' $\Pr(R(\omega))$ '. 因此, 例如, 完好的骰子成双的概率为 $1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36 = 1/6$; 但是当两个骰子灌铅, 且概率分布为 \Pr_1 时, 成双的概率为 $1/16 + 1/64 + 1/64 + 1/64 + 1/64 + 1/16 = 3/16 > 1/6$ 。把骰子灌铅使“掷成双”事件更可能发生。

(我们已经以一种比第二章定义的更一般的意义用了 \sum 符号: 式(8.1)和(8.4)中的和出现在任意集合的所有元素 ω 上, 而不仅仅在整数上。然而, 这新发展实际不使人惊恐; 每当用非整数时, 我们商定在 \sum 之下用特殊的记法, 所以与我们通常的惯例没有混淆。第二章中的其它定义仍成立; 特别, 在那一章中无限和的定义给出当集合 Ω 是无限时我们的和的相应解释。每个概率是非负的, 且所有概率的和是有界的, 所以对于所有子集 $A \subseteq \Omega$ 明确定义了式(8.4)中事件 A 的概率。)

一个随机变量是定义在概率空间的基本事件 ω 上的一个函数。例如，如果 $\Omega = D^2$ ，我们能定义 $S(\omega)$ 是骰子滚动 ω 上的点数和，以致 $S(\text{四四}) = 6+3=9$ 。点数总共 7 的概率是事件 $S(\omega)=7$ 的概率，即

$$\Pr(\text{四四}) + \Pr(\text{三四}) + \Pr(\text{二四}) + \Pr(\text{五四}) + \Pr(\text{二五}) + \Pr(\text{三五}).$$

对于完好骰子 ($\Pr = \Pr_{00}$)，它出现的概率为 $1/6$ ；对于灌铅骰子 ($\Pr = \Pr_{11}$)，它出现的概率为 $1/16 + 1/64 + 1/64 + 1/64 + 1/64 + 1/16 = 3/16$ ，这和我们成双所看到的相同。

当我们谈到随机变量时，通常省略 ' (ω) '，因为当处理任何指定问题时，通常仅涉及一个概率空间。因此对于滚出 7 的事件，我们简单地称 ' $S=7$ '，对于事件 $\{\square\square, \square\square, \square\square\}$ ，' $S=4$ '。

通过随机变量的值的概率分布能表征随机变量。因此，例如， S 在 7 个可能值 $\{2, 3, \dots, 12\}$ 上取值，对于此集中每个 s 我们能列表列出 $S=s$ 的概率：

s	2	3	4	5	6	7	8	9	10	11	12
$\Pr_{00}(S=s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
$\Pr_{11}(S=s)$	$\frac{4}{64}$	$\frac{4}{64}$	$\frac{5}{64}$	$\frac{6}{64}$	$\frac{7}{64}$	$\frac{12}{64}$	$\frac{7}{64}$	$\frac{6}{64}$	$\frac{5}{64}$	$\frac{4}{64}$	$\frac{4}{64}$

如果处理仅涉及随机变量 S 的一个问题，且无其它骰子的性质，我们能单独从这些概率计算解答，不考虑集合 $\Omega = D^2$ 的详细情况。事实上，我们能定义概率空间为较小的集合 $\Omega = \{2, 3, \dots, 12\}$ ，具有想要的概率分布 $\Pr(s)$ 。于是 ' $S=4$ ' 是一个基本事件。因此我们常常忽略基础的概率空间 Ω ，而直接处理随机变量和它们的分布。

如果在相同概率空间 Ω 上定义两个随机变量 X 和 Y ，若对于 X 范围中的每个 x 和 Y 范围中的每个 y 我们知道“联合分布”

$$\Pr(X=x \text{ 和 } Y=y),$$

则不必知道 Ω 的一切就能表征它们的情况。如果对于所有 x 和 y ,

$$\Pr = (X=x \text{ 和 } Y=y) = \Pr(X=x) \cdot \Pr(Y=y) \quad (8.5)$$

我们称 X 和 Y 是独立的随机变量。直观地说，这意味着 X 的值不影响 Y 的值。

例如，如果 Ω 是骰子滚动集合 D^2 ，我们能设 S_1 是第一个骰子上的点数， S_2 是第二个骰子上的点数。则关于前面讨论的概率分布 \Pr_{00} ， \Pr_{11} 和 \Pr_{01} 的每一个，随机变量 S_1 和 S_2 是独立的，因为我们定义每个基本事件 dd' 的骰子概率为 $S_1=d$ 的概率乘 $S_2=d'$ 的概率的一个乘积。我们能不同地定义概率，以致说，

$$\frac{\Pr(\text{四四})}{\Pr(\text{四四})} \neq \frac{\Pr(\text{四四})}{\Pr(\text{四四})};$$

但是不做这一点，因为没有假设不同的骰子彼此影响。用我们的定义，这两个比为

$$\Pr(S_2 = 5) / \Pr(S_2 = 6).$$

我们已定义 S 是两个点值之和, $S_1 + S_2$. 让我们考虑另一个随机变量 P , 即乘积 $S_1 S_2$. S 和 P 独立吗? 非形式地说, 不独立; 如果告诉 $S=2$, 我们知道 P 一定是 1. 形式地说, 还是不独立, 因为独立条件(8.5)显然不成立(至少在完好骰子情形): 对于所有合理的 s 和 p 值, 我们有 $0 < \Pr_{00}[S=s] \cdot \Pr_{00}[P=p] \leq 1/6 \cdot 1/9$; 这不等于 $\Pr_{00}[S=s \text{ 和 } P=p]$, 它是 $1/36$ 的倍数.

如果要知道一个给定随机变量的典型情况, 我们常问到它的“平均”值. 但是“平均”的概念意义不明确; 当给定一个数序列时, 人们一般说到三种不同的平均类型:

- 平均值(所有值的和除以值的个数);
- 中位数(数值上来说, 它是中值);
- 众数(最常出现的值).

例如, $(3, 1, 4, 1, 5)$ 的平均值为 $(3+1+4+1+5)/5=2.8$; 中位数是 3; 众数是 1.

但是概率论处理的是随机变量, 而不是数序列, 所以我们还要定义随机变量的“平均”的概念. 假设我们反复做一个试验, 以这样一种方式做独立试验, X 的每个值出现的频率近似地和它的概率成正比. (例如, 我们可滚一对骰子许多次, 观察 S 与 / 或 P 的值.) 我们希望定义一个随机变量的平均值以致这样的试验通常将产生一个数序列, 它的平均值, 中位数, 或众数和我们定义的 X 的平均值, 中位数, 或众数近似相等.

这里能完成它: 在概率空间 Ω 上的实随机变量 X 的平均值定义为

$$\sum_{x \in X(\Omega)} x \cdot \Pr(X = x), \quad (8.6)$$

如果这个可能的无限和存在. ($X(\Omega)$ 表示 X 能取的所有值的集合.) X 的中位数定义为所有 x 的集合使得

$$\Pr(X \leq x) \geq \frac{1}{2} \quad \text{和} \quad \Pr(X \geq x) \geq \frac{1}{2}. \quad (8.7)$$

X 的众数定义为所有 x 的集合使得

$$\Pr(X = x) \geq \Pr(X = x'), \quad \text{所有 } x' \in X(\Omega). \quad (8.8)$$

在我们的掷骰子例中, S 的平均值在分布 \Pr_{00} 中是 $2 \cdot 1/36 + 3 \cdot 2/36 + \dots + 12 \cdot 1/36 = 7$. 在分布 \Pr_{11} 它也是 7. 在两种分布中中位数和众数两者同样是 $\{7\}$. 所以在所有三个定义下 S 有同一个平均. 另一方面, 在分布 \Pr_{00} 中 P 有平均值 $49/4 = 12.25$; 它的中位数是 $\{10\}$, 它的众数是 $\{6, 12\}$. 如果我们把分布 \Pr_{11} 给灌铅骰子, P 的平均值不改变, 但是中位数为 $\{8\}$, 而众数变成单独一个数 $\{6\}$.

对于一个随机变量的平均值, 概率论有一个特殊名称和记法: 称它为期望值, 且记成

$$EX = \sum_{\omega \in \Omega} X(\omega) \Pr(\omega). \quad (8.9)$$

在我们的掷骰子例中, 此和有 36 项(对 Ω 的每个元素有一项), 而式(8.6)是仅有 7 项的

和具有相同值，因为它们都等于

$$\sum_{\substack{\omega \in \Omega \\ x \in X(\Omega)}} x \Pr(\omega) [x = X(\omega)].$$

在应用方面随机变量的平均值比其他类型的平均更有意义，所以从现在开始我们基本上忘记了中位数和众数。在本章剩下部分中我们几乎交替用术语“期望值”，“平均值”和“平均”。

如果 X 和 Y 是相同概率空间上定义的两个随机变量，则 $X+Y$ 也是这个空间上的随机变量。用公式(8.9)，它们的和的平均是它们的平均的和：

$$E(X+Y) = \sum_{\omega \in \Omega} (X(\omega) + Y(\omega)) \Pr(\omega) = EX + EY. \quad (8.10)$$

如果 α 是任何常数，则我们相似地有简单规则

$$E(\alpha X) = \alpha EX. \quad (8.11)$$

但是随机变量乘积的对应规则一般较复杂；定义期望值为基本事件上的一个和，乘积的和常常没有简单的形式。虽然有此困难，但是在随机变量独立的特殊情形中乘积的平均值有一个非常好的公式：

$$E(XY) = (EX)(EY), \quad \text{如果 } X \text{ 和 } Y \text{ 独立.} \quad (8.12)$$

我们能够通过乘积的分布律证明这一点，

$$\begin{aligned} E(XY) &= \sum_{\omega \in \Omega} X(\omega)Y(\omega) \cdot \Pr(\omega) \\ &= \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} xy \cdot \Pr(X=x \text{ 和 } Y=y) \\ &= \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} xy \cdot \Pr(X=x) \Pr(Y=y) \\ &= \sum_{x \in X(\Omega)} x \Pr(X=x) \cdot \sum_{y \in Y(\Omega)} y \Pr(Y=y) = (EX)(EY). \end{aligned}$$

例如，当 S_1 和 S_2 是一对随机骰子的第一个和第二个骰子上的点数时，我们知道 $S = S_1 + S_2$ 和 $P = S_1 S_2$ 。我们有 $ES_1 = ES_2 = 7/2$ ，因此 $ES = 7$ ；而且 S_1 和 S_2 是独立的，所以如同前面断定的那样， $EP = (7/2) \cdot (7/2) = 49/4$ 。我们还有 $E(S+P) = ES + EP = 7 + 49/4$ 。但是 S 和 P 不独立，所以我们不能断定 $E(SP) = 7 \cdot 49/4 = 343/4$ 。事实上，在分布 \Pr_{00} 中 SP 的期望值等于 $637/6$ ，在分布 \Pr_{11} 中(恰好)为 112 。

8.2 平均值和方差

在知道了随机变量的期望值之后，下一个很重要的性质是它的方差，定义为偏离期望值的平方的平均值：

$$VX = E((X - EX)^2). \quad (8.13)$$

如果用 μ 表示 EX ，则 VX 是 $(X - \mu)^2$ 的期望值。它测量了 X 的分布的“散布”。

作为方差计算的简单例子，让我们假设刚才已接受了一个提供：某人给我们以某种抽彩给奖法的 2 张礼券。抽彩给奖的组织者卖出每周抽签的 100 张彩票。以均匀随机过程选取这些彩票之一，也就是说等可能地选取每张彩票，幸运的彩票持有者获得 1 亿元。其他 99 张彩票持有者什么东西也没有获得。

我们能以两种方式使用我们的礼券：或者买 2 张相同抽彩给奖法的彩票，或者在两种抽彩给奖法的每一种中买 1 张彩票。哪一种是较好的策略？让我们试分析这一点，设 X_1 和 X_2 是随机变量，它们代表在第 1 张和第 2 张彩票上我们获得的总数。 X_1 的期望值(以百万为单位)是

$$EX_1 = \frac{99}{100} \cdot 0 + \frac{1}{100} \cdot 100 = 1.$$

对于 X_2 ，同样成立。期望值有可加性质，所以我们平均总共获得的将是

$$E(X_1 + X_2) = EX_1 + EX_2 = 200 \text{ 万元},$$

不管采用哪一种策略。

看来两种策略仍然是不同的。让我们查看期望值之外的内容，且研究 $X_1 + X_2$ 的精确概率分布：

	获得(百万)		
	0	100	200
相同抽签	.980 0	.020 0	
不同抽签	.980 1	.019 8	.000 1

如果买相同抽彩给奖的 2 张彩票，有 98% 的机会不中奖，2% 的机会获得 1 亿元。如果买不同抽彩给奖的 2 张彩票，有 98.01% 的机会不中奖，所以这比前面稍多一些可能性；而我们有 0.01% 机会获得 2 亿元，也比前面稍多一些可能性；而我们获得 1 亿元的机会现在是 1.98%。所以在第二种场合中的 $X_1 + X_2$ 的分布稍微更散开；中间的值，1 亿元有稍小的可能性，但是末端值有稍大的可能性。这种随机变量的散布概念就是想要获得的方差。我们依据随机变量对它的平均值的平方偏差来测量散布。情形 1 中，方差是

$$.98(0M - 2M)^2 + .02(100M - 2M)^2 = 196M^2;$$

在情形 2 中, 方差是

$$.9801(0M - 2M)^2 + .0198(100M - 2M)^2 + .0001(200M - 2M)^2 = 198M^2.$$

正如我们期望的那样, 后一个方差稍大, 因为情形 2 的分布稍微更散开。

当我们处理方差时, 都取平方, 所以使数相当大。(因子 M^2 是百万的乘方, 有点像大赌注赌徒下的赌注。)为了把数变换回更有意义的原来的级别, 我们常取方差的平方根。产生的数称为标准差, 通常把它记为希腊字母 σ :

$$\sigma = \sqrt{VX}. \quad (8.14)$$

在两种抽彩给奖的策略中, 随机变量 $X_1 + X_2$ 的标准差是 $\sqrt{196M^2} = 14.00M$ 和 $\sqrt{198M^2} \approx 14.071247M$ 。从某种意义上来说, 第二种选择的策略的风险费大约 71 247 元。

方差如何帮助我们选取一种策略呢? 这是不清楚的。具有较大方差的策略有小的风险费; 但是对于我们的钱, 通过承担较大风险或者通过使它安全而使我们的获得最多呢? 假设我们有机会买 100 张彩票, 而不是 2 张。则我们在单独一种抽彩给奖法中保证胜利(且方差为零); 或者我们能在 100 种不同抽彩给奖法上赌博, 具有 $.99^{100} \approx .366$ 机会不中奖, 但是获得 100 亿元也具有非零概率。在这些选择的策略中间作决定超出了本书的范围, 我们这里仅阐明如何作计算。

事实上, 有较简单的方法来计算方差, 而不用定义(8.13)。(我们觉得事情的背后一定有数学方面的道理, 因为在抽彩给奖例子中的方差不可思议地产生 M^2 的整倍数。)我们有

$$\begin{aligned} E((X - EX)^2) &= E(X^2 - 2X(EX) + (EX)^2) \\ &= E(X^2) - 2(EX)(EX) + (EX)^2, \end{aligned}$$

由于 (EX) 是常数, 因此

$$VX = E(X^2) - E(X)^2. \quad (8.15)$$

“方差是平方的平均值减平均值的平方”。

例如, 在抽彩给奖问题中, $(X_1 + X_2)^2$ 的平均值产生 $.98(0M)^2 + .02(100M)^2 = 200M^2$, 或产生 $.9801(0M)^2 + .0198(100M)^2 + .0001(200M)^2 = 202M^2$ 。减去 $4M^2$ (平均值的平方)得到结果, 我们得到费力的方法。

还有一个更容易的公式, 如果当 X 和 Y 独立时我们要计算 $V(X+Y)$, 则有

$$\begin{aligned} E((X + Y)^2) &= E(X^2 + 2XY + Y^2) \\ &= E(X^2) + 2(EX)(EY) + E(Y^2), \end{aligned}$$

因为我们知道在独立情形中 $E(XY) = (EX)(EY)$ 。所以

$$V(X + Y) = E((X + Y)^2) - (EX + EY)^2$$

$$\begin{aligned}
& -E(X^2) + 2(EX)(EY) + E(Y^2) - (EX)^2 - 2(EX)(EY) - (EY)^2 \\
& = E(X^2) - (EX)^2 + E(Y^2) - (EY)^2 \\
& = VX + VY.
\end{aligned} \tag{8.16}$$

“独立随机变量的和的方差是它们的方差的和”。例如，单独 1 张抽彩给奖彩票我们能获得总数的方差是

$$E(X_1^2) - (EX_1)^2 = .99(0M)^2 + .01(100M)^2 - (1M)^2 = 99M^2.$$

所以两种分开(独立)的抽彩给奖中，2 张抽彩给奖彩票的总获得的方差是 $2 \times 99M^2 = 198M^2$ 。 n 张独立抽彩给奖彩票的对应方差是 $n \times 99M^2$ 。

此相同公式得出骰子滚动和 S 的方差，因为 $S = S_1 + S_2$ 是两个独立随机变量的和。当骰子是完好的时候，我们有

$$VS_1 = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2}\right)^2 = \frac{35}{12},$$

因此 $VS = 35/12 + 35/12 = 35/6$ 。灌铅骰子有

$$VS_1 = \frac{1}{8}(2 \cdot 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 2 \cdot 6^2) - \left(\frac{7}{2}\right)^2 = \frac{45}{12};$$

当两个骰子被灌铅时， $VS = 45/6 = 7.5$ 。注意，灌铅骰子给出 S 一个较大方差，虽然 S 实际比完好骰子更经常取它的平均值 7。如果我们的目标是抛出许多幸运的 7，则方差不是成功的最好的指示量。

我们已学了如何计算方差，但是没有实际看到一个充分的理由，说明计算方差是自然的事情。每个人计算了它，但是为什么呢？主要理由是 Chebyshev 不等式^{[24][50]}，它说明了方差有一个重要的性质：

$$\Pr((X - EX)^2 \geq \alpha) \leq VX / \alpha, \text{ 对所有 } \alpha > 0. \tag{8.17}$$

(这与第二章中我们遇到的 Chebyshev 求和不等式是不同的。)粗略地说，式(8.17)告诉我们，如果随机变量 X 的方差 VX 小，早它将很少远离它的平均值 EX 。证明是非常简单的。我们有

$$\begin{aligned}
VX &= \sum_{\omega \in \Omega} (X(\omega) - EX)^2 \Pr(\omega) \\
&\geq \sum_{\substack{\omega \in \Omega \\ (X(\omega) - EX)^2 \geq \alpha}} (X(\omega) - EX)^2 \Pr(\omega) \\
&\geq \sum_{\substack{\omega \in \Omega \\ (X(\omega) - EX)^2 \geq \alpha}} \alpha \Pr(\omega) = \alpha \cdot \Pr((X - EX)^2 \geq \alpha),
\end{aligned}$$

除以 α , 就结束了证明。

如果用 μ 记平均值, 并用 σ 记标准差, 而且如果在式 (8.17) 中用 $c^2 V X$ 替换 α , 则条件 $(X - EX)^2 \geq c^2 V X$ 与 $(X - \mu)^2 \geq (c\sigma)^2$ 相同; 因此式 (8.17) 表为

$$\Pr(|X - \mu| \geq c\sigma) \leq 1/c^2. \quad (8.18)$$

因此, X 位于它的平均值的 c 个标准差之外的概率至多为 $1/c^2$ 。一个随机变量将至少有 75% 的机会位于 μ 的 2σ 之内; 它将至少有 99% 的机会位于 $\mu - 10\sigma$ 和 $\mu + 10\sigma$ 之间。这些是 Chebyshev 不等式的 $\alpha = 4VX$ 和 $\alpha = 100VX$ 的情形。

如果滚动一对完好骰子 n 次, 对于大的 n , n 次滚动的总值几乎总接近 $7n$ 。这里是理由: n 次独立的滚动的方差是 $35n/6$ 。方差 $35n/6$ 意味着标准差仅为 $\sqrt{35n/6}$ 。所以 Chebyshev 不等式告诉我们, 当滚动 n 个完好的骰子时, 至少所有试验的 99% 中最后和将位于

$$7n - 10\sqrt{\frac{35}{6}n} \text{ 和 } 7n + 10\sqrt{\frac{35}{6}n}$$

之间。例如, 百万次滚动的总值在 6.976 百万和 7.024 百万之间的可能性比 99 比 1 更好。

一般, 设 X 是概率空间 Ω 上的任何随机变量, 并具有有限平均值 μ 和有限标准差 σ 。然后我们能考虑概率空间 Ω^n , 它的基本事件是 n 元组 $(\omega_1, \omega_2, \dots, \omega_n)$, 其中每个 $\omega_k \in \Omega$, 且它的概率是

$$\Pr(\omega_1, \omega_2, \dots, \omega_n) = \Pr(\omega_1)\Pr(\omega_2)\cdots\Pr(\omega_n).$$

如果我们现在用公式

$$X_k(\omega_1, \omega_2, \dots, \omega_n) = X(\omega_k)$$

定义随机变量 X_k , 量

$$X_1 + X_2 + \cdots + X_n$$

是 n 个独立随机变量的和, 它对应于在 Ω 上取 X 的 n 个独立“样本”, 且把它们加在一起。 $X_1 + X_2 + \cdots + X_n$ 的平均值是 $n\mu$, 标准差是 $\sqrt{n}\sigma$; 因此 n 个样本的平均

$$\frac{1}{n}(X_1 + X_2 + \cdots + X_n)$$

至少 99% 的机会位于 $\mu - 10\sigma/\sqrt{n}$ 和 $\mu + 10\sigma/\sqrt{n}$ 之间。换句话说, 如果选足够大的 n , 则 n 个独立样本的平均几乎总非常接近期望值 EX 。(在概率论书中证明了称为强大数定律的更强的定理, 但是对于我们的目的来说, 刚得到的 Chebyshev 不等式的简单结果就够了。)

有时我们不知道概率空间的特征, 而我们要通过重复抽样随机变量 X 的值来估计它的平均值。(例如, 我们可能要知道旧金山 1 月份一天中午的平均温度, 或者可能希望知

诸保险代理人的平均估计寿命。)如果我们已得到独立的实验观察 X_1, X_2, \dots, X_n , 我们能猜测其平均值近似为

$$\hat{E}X = \frac{X_1 + X_2 + \dots + X_n}{n}. \quad (8.19)$$

且我们还能用公式

$$\hat{V}X = \frac{X_1^2 + X_2^2 + \dots + X_n^2}{n-1} - \frac{(X_1 + X_2 + \dots + X_n)^2}{n(n-1)} \quad (8.20)$$

作出方差的一个估计。此公式中的 $(n-1)$ 看来像印刷错误; 如同式(8.19)中的那样, 看来它们该是 n , 因为式(8.15)中的期望值定义了真方差 VX 。然而用 $n-1$ 代替 n 我们取得较好的估计, 因为定义式(8.20)意味着

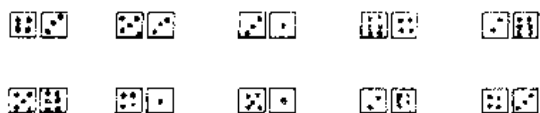
$$E(\hat{V}X) = VX. \quad (8.21)$$

这里是理由:

$$\begin{aligned} E(\hat{V}X) &= \frac{1}{n-1} E\left(\sum_{k=1}^n X_k^2 - \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^n X_j X_k\right) \\ &= \frac{1}{n-1} \left(\sum_{k=1}^n E(X_k^2) - \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^n E(X_j X_k)\right) \\ &= \frac{1}{n-1} \left(\sum_{k=1}^n E(X^2) - \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^n (E(X)^2 [j \neq k] + E(X^2) [j = k])\right) \\ &= \frac{1}{n-1} \left(nE(X^2) - \frac{1}{n} (nE(X^2) + n(n-1)E(X)^2)\right) \\ &= E(X^2) - E(X)^2 = VX. \end{aligned}$$

(当它用 $E(X)^2 [j \neq k] + E(X^2) [j = k]$ 替换 $E(X_j X_k)$ 时, 这个推导用了观察的独立性。)

实际上, 通常随机变量 X 的试验结果通过计算样本平均值 $\hat{\mu} = \hat{E}X$ 和样本标准差 $\hat{\sigma} = \sqrt{\hat{V}X}$ 得到, 且用形式 $\hat{\mu} \pm \hat{\sigma} / \sqrt{n}$ 表示解答。例如, 这里是假设两个完好骰子的10次滚动:



点数和 S 的样本平均是

$$\hat{\mu} = (7 + 11 + 8 + 5 + 4 + 6 + 10 + 8 + 8 + 7) / 10 = 7.4;$$

样本方差是

$$(7^2 + 11^2 + 8^2 + 5^2 + 4^2 + 6^2 + 10^2 + 8^2 + 8^2 + 7^2 - 10\hat{\mu}^2)/9 \approx 2.1^2.$$

在这些试验基础上, 我们估计这些骰子的平均点数和是 $7.4 \pm 2.1/\sqrt{10} = 7.4 \pm 0.7$.

为了表明如何推理地而不是经验地计算平均值和方差, 让我们再举一个平均值和方差的例子。在第五章中我们考虑了一个“足球胜利问题”, 其中把 n 顶帽子抛向空中, 而结果是帽子的一种随机排列。在方程(5.51)中我们证明了没有人取回正确的帽子的概率为 $n!/n! \approx 1/e$ 。我们也推出 k 个人恰好接到他们自己的帽子的概率公式

$$P(n, k) = \frac{1}{n!} \binom{n}{k} (n-k)! = \frac{1}{k!} \frac{(n-k)!}{(n-k)!}. \quad (8.22)$$

再形式地说明刚才学到的这些结果, 我们能考虑 $\{1, 2, \dots, n\}$ 的所有 $n!$ 个排列 π 的概率空间 Π_n , 其中 $\Pr(\pi) = 1/n!$ (所有 $\pi \in \Pi_n$), 随机变量

$$F_n(\pi) = \pi \text{ 的“不变点”的个数, } \pi \in \Pi_n$$

计量足球胜利问题中正确落下帽子的数目。方程(8.22)给出 $\Pr(F_n = k)$, 但是让我们自称不知道任何这样的公式, 仅要研究 F_n 的平均值和它的标准差。

事实上, 避开第 V 章的复杂内容, 非常容易计算平均值。我们简单地看到

$$F_n(\pi) = F_{n,1}(\pi) + F_{n,2}(\pi) + \dots + F_{n,n}(\pi),$$

$$F_{n,k}(\pi) = [\pi \text{ 的位置 } k \text{ 是一个不变点}], \pi \in \Pi_n.$$

因此

$$EF_n = EF_{n,1} + EF_{n,2} + \dots + EF_{n,n}.$$

且 $F_{n,k}$ 的期望值简单地为 $F_{n,k} = 1$ 的概率, 它是 $1/n$, 因为 $n!$ 个排列 $\pi = \pi_1 \pi_2 \dots \pi_n \in \Pi_n$ 的 $(n-1)!$ 个恰好有 $\pi_k = k$ 。所以

$$EF_n = n/n = 1, \quad n > 0. \quad (8.23)$$

平均一个帽子将处于它的正确位置。“一个随机排列平均有一个不变点。”

现在标准差是什么呢? 这个问题较困难, 因为 $F_{n,k}$ 不彼此独立。但是我们通过分析它们中间的彼此依赖能计算方差:

$$\begin{aligned} E(F_n^2) &= E\left(\left(\sum_{k=1}^n F_{n,k}\right)^2\right) = E\left(\sum_{j=1}^n \sum_{k=1}^n F_{n,j} F_{n,k}\right) \\ &= \sum_{j=1}^n \sum_{k=1}^n E(F_{n,j} F_{n,k}) = \sum_{1 \leq k \leq n} E(F_{n,k}^2) + 2 \sum_{1 \leq j < k \leq n} E(F_{n,j} F_{n,k}). \end{aligned}$$

(我们用第二章中推出式(2.33)的相似方法。)现在 $F_{n,k}^2 = F_{n,k}$, 因为 $F_{n,k}$ 是 0 或 1, 因此如同前面那样 $E(F_{n,k}^2) = EF_{n,k} = 1/n$, 且如果 $j < k$, 我们得到 $E(F_{n,j} F_{n,k}) = \Pr(\pi \text{ 有 } j \text{ 和 } k \text{ 两个为不变点}) = (n-2)!/n! = 1/n(n-1)$ 。所以

$$E(F_n^4) = \frac{n}{n} + \binom{n}{2} \frac{2}{n(n-1)} = 2, \quad n \geq 2. \quad (8.24)$$

(作为一个检验, 当 $n=3$ 时, 我们有 $(2/6) \cdot 0^2 + (3/6) \cdot 1^2 + (0/6) \cdot 2^2 + (1/6) \cdot 3^2 = 2$.) 方差是 $E(F_n^2) - (EF_n)^2 = 1$, 所以标准差(像平均值)是 1. “ $n \geq 2$ 个元素的一个随机排列有 1 ± 1 个不变点.”

8.3 概率母函数

如果 X 是仅取非负整值的随机变量, 则我们用七章的技巧能很好地获得概率分布. X 的概率母函数是

$$G_X(z) = \sum_{k \geq 0} \Pr(X = k) z^k. \quad (8.25)$$

这个 z 的幂级数包含随机变量 X 的所有信息. 我们也能以两种另外的方式表达它:

$$G_X(z) = \sum_{\omega \in \Omega} \Pr(\omega) z^{X(\omega)} = E(z^X). \quad (8.26)$$

$G_X(z)$ 的系数是非负的, 且它们总和为 1; 后面的条件可记为

$$G_X(1) = 1. \quad (8.27)$$

相反, 任何非负系数的幂级数 $G(z)$ (且 $G(1)=1$) 是某个随机变量的概率母函数.

关于概率母函数的令人满意的事情是它通常简化了平均值和方差的计算. 例如, 容易表达平均值:

$$EX = \sum_{k \geq 0} k \cdot \Pr(X = k) = \sum_{k \geq 0} \Pr(X = k) \cdot k z^{k-1} \Big|_{z=1} = G'_X(1). \quad (8.28)$$

我们简单地对 z 求导概率母函数, 且置 $z=1$.

方差稍微复杂一点:

$$\begin{aligned} E(X^2) &= \sum_{k \geq 0} k^2 \cdot \Pr(X = k) \\ &= \sum_{k \geq 0} \Pr(X = k) \cdot (k(k-1)z^{k-2} + kz^{k-1}) \Big|_{z=1} = G''_X(1) + G'_X(1). \end{aligned}$$

所以

$$VX = G''_X(1) + G'_X(1) - (G'_X(1))^2. \quad (8.29)$$

如果能计算两个导数值 $G'_x(1)$ 和 $G''_x(1)$, 则方程(8.28)和(8.29)说明, 我们能计算平均值和方差, 我们不需知道概率的闭形式; 我们甚至不需知道 $G_x(z)$ 的闭形式。

当 G 是任何函数时, 记

$$\text{平均值}(G) = G'(1), \quad (8.30)$$

$$\text{方差}(G) = G''(1) + G'(1) - G'(1)^2 \quad (8.31)$$

是方便的, 因为我们经常要计算这些导数的组合。

在许多重要的情形, 关于概率母函数的第二件令人满意的事情是它们是 z 的比较简单函数。例如, 让我们看 n 阶的均匀分布, 此时随机变量在 $\{0, 1, \dots, n-1\}$ 的每个值上以概率 $1/n$ 取值。此时概率母函数是

$$U_n(z) = \frac{1}{n} (1 + z + \dots + z^{n-1}) = \frac{1}{n} \frac{1 - z^n}{1 - z}, \quad n \geq 1. \quad (8.32)$$

我们有 $U_n(z)$ 的一个闭形式, 因为这是一个几何级数。

但是此闭形式证明是有点令人为难的: 当我们代入 $z=1$ (对于概率母函数最关键性的 z 值), 我们得到不定比 $0/0$, 即使 $U_n(x)$ 是对 z 的任何值完全明确定义的一个多项式。根据非闭形式 $(1 + z + \dots + z^{n-1})/n$, 值 $U_n(1) = 1$ 是明显的, 然而看来如果我们要从闭形式来确定 $U_n(1)$, 我们一定要借助 L'Hospital 法则来求 $\lim_{z \rightarrow 1} U_n(z)$ 。用 L'Hospital 法则确定 $U'_n(1)$ 将是费劲的, 因为在分母中有一个因子 $(z-1)^2$; $U''_n(1)$ 也是费劲的。

幸运地有一个摆脱困境的很好方法。如果 $G(z) = \sum_{n \geq 0} g_n z^n$ 是任何幂级数, 它至少对 z 的一个值 ($|z| > 1$) 收敛, 幂级数 $G'(z) = \sum_{n \geq 0} n g_n z^{n-1}$ 也有这个性质, 且 $G''(z)$, $G'''(z)$ 等等也都有这个性质。所以由 Taylor 定理, 我们可记

$$G(1+t) = G(1) + \frac{G'(1)}{1!} t + \frac{G''(1)}{2!} t^2 + \frac{G'''(1)}{3!} t^3 + \dots; \quad (8.33)$$

当把 $G(1+t)$ 按 t 的幂展开时, 所有 $G(z)$ 在 $z=1$ 处的导数将出现为系数。

例如, 以此法容易求得均匀概率母函数的导数:

$$\begin{aligned} U_n(1+t) &= \frac{1}{n} \frac{(1+t)^n - 1}{t} \\ &= \frac{1}{n} \binom{n}{1} + \frac{1}{n} \binom{n}{2} t + \frac{1}{n} \binom{n}{3} t^2 + \dots + \frac{1}{n} \binom{n}{n} t^{n-1}. \end{aligned}$$

把此式与式 (8.33) 比较得到

$$U_n(1) = 1; \quad U'_n(1) = \frac{n-1}{2}; \quad U''_n(1) = \frac{(n-1)(n-2)}{3}; \quad (8.34)$$

一般 $U_n^{(m)} = (n-1)^m / (m+1)$, 虽然我们仅需 $m=1$ 和 $m=2$ 的情形来计算平均值和方

差、均匀分布的平均值是

$$U'_n(1) = \frac{n-1}{2}, \quad (8.35)$$

方差是

$$U''_n(1) + U'_n(1) - U'_n(1)^2 = 4\frac{(n-1)(n-2)}{12} + 6\frac{(n-1)}{12} - 3\frac{(n-1)^2}{12} = \frac{n^2-1}{12}. \quad (8.36)$$

关于概率母函数的第三件令人满意的事情是概率母函数的乘积对应于独立随机变量的和。我们在第五章和第七章中学到母函数的乘积对应于序列的卷积，但是在应用中更重要的是知道概率的卷积对应于独立随机变量的和。事实上，如果 X 和 Y 是取整数值的随机变量，则 $X+Y=n$ 的概率是

$$\Pr(X+Y=n) = \sum_k \Pr(X=k \text{ 和 } Y=n-k).$$

若 X 和 Y 是独立的，我们现在有一个卷积

$$\Pr(X+Y=n) = \sum_k \Pr(X=k)\Pr(Y=n-k),$$

所以

$$G_{X+Y}(z) = G_X(z)G_Y(z), \text{ 若 } X \text{ 和 } Y \text{ 是独立的.} \quad (8.37)$$

在本章前面，我们看到，当 X 和 Y 独立时， $V(X+Y) = VX + VY$ 。设 $F(z)$ 和 $G(z)$ 是 X 和 Y 的概率母函数，且设 $H(z)$ 是 $X+Y$ 的概率母函数。则

$$H(z) = F(z)G(z),$$

且平均值和方差的公式(8.28)到(8.31)告诉我们，一定有

$$\text{平均值}(H) = \text{平均值}(F) + \text{平均值}(G); \quad (8.38)$$

$$\text{方差}(H) = \text{方差}(F) + \text{方差}(G). \quad (8.39)$$

这些导数(平均值(H)= $H'(1)$ 和方差(H)= $H''(1)+H'(1)-H'(1)^2$)的性质的公式对于任意函数乘积 $H(z)=F(z)G(z)$ 是不成立的；我们有

$$H'(z) = F'(z)G(z) + F(z)G'(z),$$

$$H''(z) = F''(z)G(z) + 2F'(z)G'(z) + F(z)G''(z).$$

但是如果置 $z=1$ ，我们能看到只要

$$F(1) = G(1) = 1 \quad (8.40)$$

且导数存在，式(8.38)和(8.39)将成立。为了这些公式成立，“概率”不必在 $[0..1]$ 中。每当 $F(1)$ 和 $G(1)$ 是非零时，为了使这个条件成立，我们能通过除以 $F(1)$ 和 $G(1)$ 来正规化函数 $F(z)$ 和 $G(z)$ 。

平均值和方差不是全部描述。它们仅是由丹麦天文学家 Thorvald Nicolai

Thiele^[288]在 1903 年引入的称为累积量统计的一个无穷级数的两个累积量。一个随机变量的前两个累积量 κ_1 和 κ_2 就是我们需要的平均值和方差，高阶累积量表达一个分布的更细微的性质。当 $G(z)$ 是一个随机变量的概率母函数，一般的公式

$$\ln G(e^t) = \frac{\kappa_1}{1!}t + \frac{\kappa_2}{2!}t^2 + \frac{\kappa_3}{3!}t^3 + \frac{\kappa_4}{4!}t^4 + \cdots \quad (8.41)$$

定义了所有阶的累积量。

让我们更仔细地研究累积量。如果 $G(z)$ 是 X 的概率母函数，我们有

$$\begin{aligned} G(e^t) &= \sum_{k \geq 0} \Pr(X=k) e^{kt} = \sum_{k, m \geq 0} \Pr(X=k) \frac{k^m t^m}{m!} \\ &= 1 + \frac{\mu_1}{1!}t + \frac{\mu_2}{2!}t^2 + \frac{\mu_3}{3!}t^3 + \cdots, \end{aligned} \quad (8.43)$$

其中

$$\mu_m = \sum_{k \geq 0} k^m \Pr(X=k) = E(X^m).$$

这个量 μ_m 被称为 X 的“第 m 个矩”。我们可在式(8.41)两边取指数，得到 $G(e^t)$ 的另一个公式：

$$\begin{aligned} G(e^t) &= 1 + \frac{\left(\kappa_1 t + \frac{1}{2}\kappa_2 t^2 + \cdots\right)}{1!} + \frac{\left(\kappa_1 t + \frac{1}{2}\kappa_2 t^2 + \cdots\right)^2}{2!} + \cdots \\ &= 1 + \kappa_1 t + \frac{1}{2}(\kappa_2 + \kappa_1^2)t^2 + \cdots. \end{aligned}$$

使 t 的幂的系数相等导致一系列公式

$$\kappa_1 = \mu_1, \quad (8.44)$$

$$\kappa_2 = \mu_2 - \mu_1^2, \quad (8.45)$$

$$\kappa_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3, \quad (8.46)$$

$$\kappa_4 = \mu_4 - 4\mu_1\mu_3 + 12\mu_1^2\mu_2 - 3\mu_2^2 - 6\mu_1^4, \quad (8.47)$$

$$\kappa_5 = \mu_5 - 5\mu_1\mu_4 + 20\mu_1^2\mu_3 - 10\mu_2\mu_3 + 30\mu_1\mu_2^2 - 60\mu_1^3\mu_2 + 24\mu_1^5, \quad (8.48)$$

⋮

依据矩确定累积量。注意，正如断言的那样， κ_2 确实是方差， $E(X)^2 - (EX)^2$ 。

方程(8.41)清楚地使得两个概率母函数的乘积 $F(z)G(z)$ 所定义的累积量将是 $F(z)$ 和 $G(z)$ 的对应累积量的和，因为乘积的对数是和。所以就像平均值和方差那样，独立随机变量的和的所有累积量满足加法性质。这个性质使累积量比矩更重要。

如果我们采取稍微不同的步骤，写

$$G(1+t) = 1 + \frac{\alpha_1}{1!}t + \frac{\alpha_2}{2!}t^2 + \frac{\alpha_3}{3!}t^3 + \dots,$$

方程(8.33)告诉我们, α 是“阶乘矩”

$$\begin{aligned}\alpha_m &= G^{(m)}(1) \\ &= \sum_{k \geq 0} \Pr(X=k) k^m z^{k-m} \Big|_{z=1} \\ &= \sum_{k \geq 0} k^m \Pr(X=k) \\ &= E(X^m).\end{aligned}\tag{8.49}$$

由此得到

$$\begin{aligned}G(e^t) &= 1 + \frac{\alpha_1}{1!}(e^t - 1) + \frac{\alpha_2}{2!}(e^t - 1)^2 + \dots \\ &= 1 + \frac{\alpha_1}{1!} \left(t + \frac{1}{2}t^2 + \dots \right) + \frac{\alpha_2}{2!}(t^2 + t^3 + \dots) + \dots \\ &= 1 + \alpha_1 t + \frac{1}{2}(\alpha_2 + \alpha_1^2)t^2 + \dots,\end{aligned}$$

且我们能依据导数 $G^{(m)}(1)$ 表达累积量:

$$\kappa_1 = \alpha_1,\tag{8.50}$$

$$\kappa_2 = \alpha_2 + \alpha_1^2 - \alpha_1^2,\tag{8.51}$$

$$\kappa_3 = \alpha_3 + 3\alpha_2\alpha_1 + \alpha_1^3 - 3\alpha_2\alpha_1 - 3\alpha_1^2 + 2\alpha_1^3,\tag{8.52}$$

⋮

此公式的序列产生了把式(8.38)和(8.39)推广到所有累积量的“加法”等式。

让我们回到现实中来, 且应用这些思想到简单例子。一个随机变量的最简单情形是一个“随机常数”, 其中 X 以概率 1 取某个指定值 x 。此时 $G_x(z) = z^x$, $\ln G_x(e^t) = xt$, 因此平均值是 x , 所有其他累积量是零。由此得到任何概率母函数乘以 z^x 的运算, 平均值增加 x , 但是剩下的方差和所有其他累积量不改变。

如何把概率母函数应用到骰子? 一个完好骰子上的点数分布有概率母函数

$$G(z) = \frac{z + z^2 + z^3 + z^4 + z^5 + z^6}{6} = z U_6(z),$$

其中 U_6 是 6 阶的均匀分布的概率母函数。因子 ‘ z ’ 把平均值增加 1, 所以平均值是 3.5, 而不是式(8.35)所给的 $(n-1)/2 = 2.5$; 但是外加的 ‘ z ’ 不影响方差(8.36), 它等于 $35/12$ 。

两个独立的骰子的总点数的概率母函数是一个骰子上的点数的概率母函数的平方,

$$\begin{aligned} G_S(z) &= \frac{z^2 + 2z^3 + 3z^4 + 4z^5 + 5z^6 + 6z^7 + 5z^8 + 4z^9 + 3z^{10} + 2z^{11} + z^{12}}{36} \\ &= z^2 U_6(z)^2. \end{aligned}$$

如果我们滚动一对完好的骰子 n 次, 从头至尾取得总共 k 点的概率相似地为

$$[z^k]G_S(z)^n = [z^k]z^{2n}U_6(z)^{2n} = [z^{k-2n}]U_6(z)^{2n}.$$

在前面考虑的足球胜利的脱帽问题中, 按另一种方式认为是枚举随机排列的不变点问题, 根据式(5.59)我们知道概率母函数是

$$F_n(z) = \sum_{0 \leq k \leq n} \frac{(n-k)!}{(n-k)!} \frac{z^k}{k!}, \quad n \geq 0. \quad (8.53)$$

所以

$$F'_n(z) = \sum_{1 \leq k \leq n} \frac{(n-k)!}{(n-k)!} \frac{z^{k-1}}{(k-1)!} = \sum_{0 \leq k \leq n-1} \frac{(n-1-k)!}{(n-1-k)!} \frac{z^k}{k!} = F_{n-1}(z).$$

不必知道系数的细节, 我们根据此递归 $F'_n(z) = F_{n-1}(z)$ 推得 $F_n^{(m)}(z) = F_{n-m}(z)$; 因此

$$F_n^{(m)}(1) = F_{n-m}(1) = [n \geq m]. \quad (8.54)$$

此公式容易计算平均值和方差; 如同前面那样, 我们(更快)求得当 $n \geq 2$ 时它们都为 1。

事实上, 现在能证明每当 $n \geq m$ 时此随机变量的第 m 个累积量 κ_m 等于 1。对于仅依赖于 $F'_n(1)$, $F''_n(1)$, \dots , $F_n^{(m)}(1)$ 的第 m 个累积量, 这些量全为 1; 因此我们得到, 如同当用极限概率母函数

$$F_\infty(z) = e^{z-1} \quad (8.55)$$

替换 $F_n(z)$ 所做的第 m 个累积量的相同解答(对于所有阶的导数有 $F_\infty^{(m)}(1) = 1$)。 F_∞ 的累积量全等于 1, 因为

$$\ln F_\infty(e^t) = \ln e^{e^t-1} = e^t - 1 = \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots$$

8.4 掷硬币

现在让我们转到处理仅有两个结局的情况。如果掷一个硬币, 它出现正面的概率为 p , 出现反面的概率为 q , 其中

$$p + q = 1.$$

(假设硬币不在它的边沿上停止, 或者落入一个洞内, 等等。)整个这一节中, 数 p 和 q 总是相加为 1. 如果硬币是完好的, 则有 $p = q = 1/2$; 否则称硬币是偏倚的.

在抛一个硬币一次后正面个数的概率母函数是

$$H(z) = q + pz. \quad (8.56)$$

如果我们抛硬币 n 次, 并总假设不同的硬币抛掷是独立的, 按照二项定理由

$$H(z)^n = (q + pz)^n = \sum_{k \geq 0} \binom{n}{k} p^k q^{n-k} z^k \quad (8.57)$$

产生正面数. 因此我们在 n 次抛掷中恰好得到 k 个正面的机会是 $\binom{n}{k} p^k q^{n-k}$. 这个概率序列称为二项分布.

假设我们反复抛一个硬币, 直到正面第一次出现. 恰好要求 k 次抛掷的概率是什么? $k=1$ 的概率为 p (因为这是第一次掷时正面的概率); $k=2$ 的概率为 qp (因为这是首先为反面, 然后正面的概率); 且对于一般 k 概率为 $q^{k-1}p$. 所以母函数是

$$pz + qpz^2 + q^2pz^3 + \cdots = \frac{pz}{1 - qz}. \quad (8.58)$$

重复此过程直到获得 n 个正面, 给出概率母函数

$$\left(\frac{pz}{1 - qz} \right)^n = p^n z^n \sum_k \binom{n+k-1}{k} (qz)^k = \sum_k \binom{k-1}{k-n} p^n q^{k-n} z^k. \quad (8.59)$$

顺便提一句, 这是 z^n 乘负二项分布的母函数

$$\left(\frac{p}{1 - qz} \right)^n = \sum_k \binom{n+k-1}{k} p^n q^k z^k. \quad (8.60)$$

在掷一个硬币直到出现 n 个正面的例(8.59)中的概率空间不同于本章前面见到的概率空间, 因为它包含无限多个元素. 每个元素是正面与 / 或反面的一个有限序列, 正好一起包含 n 个正面, 且以正面终止; 这样一个序列的概率是 $p^n q^{k-n}$, 其中 $k-n$ 是反面数. 因此, 例如, 若 $n=3$ 且如果用 H 记正面, 用 T 记反面, 序列 $THTTTTHH$ 是概率空间的一个元素, 且它的概率是 $qpqqqp = p^3 q^4$.

设 X 是具有二项分布(8.57)的随机变量, 且设 Y 是具有负二项分布(8.60)的随机变量. 这些分布依赖于 n 和 p . X 的平均值是 $nH'(1) = np$, 由于它的概率母函数是 $H(z)^n$; 方差是

$$n(H''(1) + H'(1) - H'(1)^2) = n(0 + p - p^2) = npq. \quad (8.61)$$

因此标准差是 \sqrt{npq} : 如果我们抛一个骰子 n 次, 我们期望取得大约 $np \pm \sqrt{npq}$ 次正面. 用类似的方式能求得 Y 的平均值和方差: 如果我们设

$$G(z) = \frac{p}{1 - qz},$$

则有

$$G'(z) = \frac{pq}{(1 - qz)^2},$$

$$G''(z) = \frac{2pq^2}{(1 - qz)^3};$$

因此 $G'(1) = pq/p^2 = q/p$ 和 $G''(1) = 2pq^2/p^3 = 2q^2/p^2$. 由此得到 Y 的平均值为 nq/p 和方差为 nq/p^2 .

导出 Y 的平均值和方差的较简单的方法是用倒数母函数

$$F(z) = \frac{1 - qz}{p} = \frac{1}{p} - \frac{q}{p}z \quad (8.62)$$

且记

$$G(z)^n = F(z)^{-n}. \quad (8.63)$$

这个多项式 $F(z)$ 不是一个概率母函数, 因为它有一个负系数, 但是它满足关键条件 $F(1) = 1$. 因此 $F(z)$ 形式上是对应于一个硬币取得正面的“概率”等于 $-q/p$ 的一个二项式; $G(z)$ 形式等价于掷这样一个硬币 -1 次. 所以具有参数 (n, p) 的负二项分布可视为具有参数 $(n', p') = (-n, -q/p)$ 的通常二项分布. 形式地做下去, 平均值一定是 $n'p' = (-n)(-q/p) = nq/p$, 方差一定是 $n'p'q' = (-n)(-q/p)(1 + q/p) = nq/p^2$. 这个涉及负概率的形式推导是成立的, 因为通常二项式的推导是基于形式幂级数之间的等式上面, 此时未用假设 $0 \leq p \leq 1$.

让我们移到另一个例上来: 我们必需掷一个硬币多少次直到我们在一行中取得正面两次? 现在概率空间由所有以 HH 终止但在最后位置以前没有相继的 H 的序列:

$$\Omega = \{HH, THH, TTHH, HTHH, TTTHH, THTHH, HTTHH, \dots\}.$$

用 p 替换 H, q 替换 T 获得任何给定序列的概率; 例如, 序列 THTHH 将以概率

$$\Pr(\text{THTHH}) = qpqp p = p^3 q^2$$

出现.

如同第七章开始处所做的那样, 现在我们能利用母函数, 设 S 是 Ω 的所有元素的无限和

$$S = HH + THH + TTHH + HTHH + TTTHH + THTHH + HTTHH + \dots$$

如果我们用 p 替换每个 H, 用 q 替换每个 T, 我们得到直至两个相继正面出现的抛掷次数的概率母函数.

S 和方程(7.1)中的多米诺骨牌的和

$$T = | + \square + \square\square + \square\square\square + \square\square\square\square + \square\square\square\square\square + \square\square\square\square\square\square + \dots$$

之间有一个不寻常的关系。事实上,如果我们用 T 替换每个 \square , 用 HT 替换每个 $\square\square$, 然后在最后添加一个 HH , 则我们从 T 得到 S 。这个对应容易证明, 因为对于某个 $n \geq 0$, Ω 的每个元素有形式 $(T+HT)^n HH$, 而 T 的每一项有形式 $(\square+\square\square)^n$ 。所以根据式(7.4)我们有

$$S = (1 - T - HT)^{-1} HH$$

且问题的概率母函数为

$$G(z) = (1 - qz + (pz)(qz))^{-1} (pz)^2 = \frac{p^2 z^2}{1 - qz - pqz^2}. \quad (8.64)$$

关于负二项分布的经验给我们提供了一个线索, 即通过记

$$G(z) = \frac{z^2}{F(z)} \quad \text{其中} \quad F(z) = \frac{1 - qz - pqz^2}{p^2}$$

以及计算这个伪概率母函数 $F(z)$ 的“平均值”和“方差”, 我们能很容易地计算式(8.64)的平均值和方差。(再一次引进 $F(1)=1$ 的一个函数。)我们得到

$$F'(1) = (-q - 2pq) / p^2 = 2 - p^{-1} - p^{-2};$$

$$F''(1) = -2pq / p^2 = 2 - 2p^{-1}.$$

所以, 由于 $z^2 = F(z)G(z)$, 平均值 $\langle z^2 \rangle = 2$, 方差 $\langle z^2 \rangle = 0$, 分布 $G(z)$ 的平均值和方差是

$$\text{平均值}(G) = 2 - \text{平均值}(F) = p^{-2} + p^{-1}; \quad (8.65)$$

$$\text{方差}(G) = -\text{方差}(F) = p^{-4} + 2p^{-3} - 2p^{-2} - p^{-1}. \quad (8.66)$$

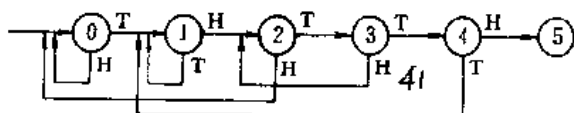
当 $p = \frac{1}{2}$ 时, 平均值和方差分别是 6 和 22。(习题 4 讨论了通过减法计算平均值和方差。)

现在让我们试做一个较复杂的试验: 我们抛掷硬币直到首次获得型式 $THTTH$ 。获得的状态之和是

$$S = THTTH + HTHTTH + TTHTTH \\ + HHTHTTH + HTTHTTH + THTHTTH + TTTHTTH + \dots,$$

这一个和比前一个和更难描述。如果回到第七章中解多米诺问题的方法上去, 我们能得到 S 的一个公式, 把它考虑为由下列“自动机”确定的“有限状态语言”:

概率空间中的基本事件是 H 和 T 的序列, 它引导状态 0 到状态 5。例如, 假设我们正好



看到 THT, 则我们处于状态 3。抛掷反面, 现在把我们带到状态 4; 抛掷正面状态 3 将把我们带到状态 2(不是自始至终回到状态 0, 由于我们正好见到 TH, 可能跟随的是 TTH)。

在这种表达中, 可设 S_k 是引导到状态 k 的 H 和 T 的所有序列的和; 由此得到

$$S_0 = 1 + S_0H + S_2H,$$

$$S_1 = S_0T + S_1T + S_4T,$$

$$S_2 = S_1H + S_3H,$$

$$S_3 = S_2T,$$

$$S_4 = S_3T,$$

$$S_5 = S_4H.$$

现在问题中的和 S 是 S_5 , 通过解 6 个未知量的 6 个方程可得到它。用 pz 替换 H, qz 替换 T 给出母函数, 其中 S_k 中 z^n 的系数是 n 次抛掷后在状态 k 中的概率。

以相同方式, 状态之间的转变的任何图形(其中从状态 j 转变到状态 k 以给定概率 p_{jk} 出现)导致一组线性联立方程, 它的解是出现 n 次转变后状态概率的母函数。这类系统称为马尔科夫过程, 它们的状态理论和线性方程的理论有内在联系。

但是能以一种没有一般有限状态方法复杂的很简单的方法解硬币抛掷问题。代替 6 个未知量的 6 个方程, 我们可仅用 2 个未知量的 2 个方程来表示 S 。诀窍是考虑所有不包含任何给定型式 THTHH 的结局的抛掷序列的辅助和 $N = S_0 + S_1 + S_2 + S_3 + S_4$:

$$N = 1 + H + T + HH + \dots + THTHT + THTTT + \dots,$$

我们得到

$$1 + N(H + T) = N + S, \quad (8.67)$$

因为左边每一项或者以 THTTH 终止(且属于 S)或者不以 THTTH 终止(且属于 N); 相反地说, 右边的每一项或者是空的或者属于 NH 或者属于 NT 。我们还有重要的附加方程

$$NTHTTH = S + STTH, \quad (8.68)$$

因为左边的每一项在第一个 H 或第二个 H 后面完成 S 的一项, 且右边的每一项属于左边,

容易得到这 2 个联立方程的解: 根据式(8.67)我们有 $N = (1 - S)(1 - H - T)^{-1}$, 因此

$$(1 - S)(1 - T + H)^{-1}THTTH = S(1 + TTH).$$

如同前面那样, 如果用 pz 替换 H, qz 替换 T, 我们得到抛掷数的概率母函数 $G(z)$ 。由于 $p+q=1$, 出现一点简化, 我们发现

$$\frac{(1-G(z))p^2q^3z^5}{1-z} = G(z)(1+pq^2z^3);$$

因此解是

$$G(z) = \frac{p^2q^3z^5}{p^2q^3z^5 + (1+pq^2z^3)(1-z)}. \quad (8.69)$$

注意, 如果 $pq \neq 0$, $G(1)=1$; 我们终于以概率 1 遇到型式 THHTH, 除非硬币被装填以致它总出现正面, 或总发现反面。

为了取得分布(8.69)的平均值和方差, 如同我们在前一个问题所做的那样, 我们转换 $G(z)$, 记 $G(z)=z^5/F(z)$, 其中 F 是多项式:

$$F(z) = \frac{p^2q^3z^5 + (1+pq^2z^3)(1-z)}{p^2q^3}. \quad (8.70)$$

相应的导数是

$$F'(1) = 5 - (1+pq^2)/p^2q^3,$$

$$F''(1) = 20 - 6pq^2/p^2q^3;$$

如果 X 是抛掷数, 我们得到

$$EX = \text{平均值}(G) = 5 - \text{平均值}(F) = p^{-2}q^{-3} + p^{-1}q^{-1}; \quad (8.71)$$

$$\begin{aligned} VX = \text{方差}(G) &= -\text{方差}(F) \\ &= -25 + p^{-2}q^{-3} + 7p^{-1}q^{-1} + \text{平均值}(F)^2 \\ &= (EX)^2 - 9p^{-2}q^{-3} - 3p^{-1}q^{-1}. \end{aligned} \quad (8.72)$$

当 $p=1/2$ 时, 平均值和方差是 36 和 996。

让我们来一般化: 刚才解的问题足够“随机”地表明如何分析等待首次出现正面和反面的任意型式 A 的情形。我们再设 S 是 H 和 T 的所有获胜序列的和, 设 N 是还未遇到型式 A 的所有序列的和。方程(8.67)保持相同; 方程(8.68)将变成

$$\begin{aligned} NA = S(1 + A^{(1)}[A^{(m-1)} = A_{(m-1)}] + A^{(2)}[A^{(m-2)} = A_{(m-2)}] \\ + \cdots + A^{(m-1)}[A^{(1)} = A_{(1)}]), \end{aligned} \quad (8.73)$$

其中 m 是 A 的长度, $A^{(k)}$ 和 $A_{(k)}$ 分别表示 A 的最后 k 个符号和前 k 个符号。例如, 如果 A 是刚才研究的型式 THHTH, 则有

$$A^{(1)} = H, \quad A^{(2)} = TH, \quad A^{(3)} = TTH, \quad A^{(4)} = HTTH;$$

$$A_{(1)} = T, \quad A_{(2)} = TH, \quad A_{(3)} = THT, \quad A_{(4)} = THTT.$$

由于只有完全的匹配 $A^{(2)} = A_{(2)}$, 方程(8.73)才化为(8.68)。

设 \tilde{A} 是型式 A 中 p^{-1} 替换 H, q^{-1} 替换 T 的结果。于是不难推广式(8.71)和(8.72)的推导, 得到(习题 20)一般的平均值和方差

$$EX = \sum_{k=1}^m \tilde{A}_{(k)} [A^{(k)} = A_{(k)}]; \quad (8.74)$$

$$VX = (EX)^2 - \sum_{k=1}^m (2k-1) \tilde{A}_k [A^{(k)} = A_{(k)}]. \quad (8.75)$$

在特殊情形 $p=1/2$ 中, 我们能以特别简单的方式解释这些公式。给定 m 个正面和反面的型式 A , 设

$$A : A = \sum_{k=1}^m 2^{k-1} [A^{(k)} = A_{(k)}]. \quad (8.76)$$

我们能容易地找到此数的二进制表示, 在这样的位置下置‘1’, 当串重叠于移到此位置开始的一个拷贝时使得串完全匹配于它本身:

$$A = \text{HTHTHHHTHTH}$$

$$A : A = (1000010101)_2 = 512 + 16 + 4 + 1 = 533$$

$$\text{HTHTHHHTHTH} \quad \checkmark$$

$$\text{HTHTHHHTHTH}$$

$$\text{HTHTHHHTHTH}$$

$$\text{HTHTHHHTHTH}$$

$$\text{HTHTHHHTHTH}$$

$$\text{HTHTHHHTHTH} \quad \checkmark$$

$$\text{HTHTHHHTHTH}$$

$$\text{HTHTHHHTHTH} \quad \checkmark$$

$$\text{HTHTHHHTHTH}$$

$$\text{HTHTHHHTHTH} \quad \checkmark$$

方程(8.74)告诉我们, 如果用一个完好的硬币, 直到型式 A 出现的抛掷期望数恰好是 $2(A : A)$, 因为当 $p=q=1/2$ 时 $\tilde{A}_k = 2^k$ 。由苏联数学家 A·D·Solov'ev 在 1966 年^[271]首先发现的这个结果初看起来似自相矛盾: 宁可无自重叠的型式出现而不愿重叠型式产生! 遇到 HHHHT 或 THHHH 几乎是遇到 HHHHH 的两倍。

现在让我们考虑一个由 Walter Penney^[231]在 1969 年创造的有趣的游戏。Alice 和 Bill 抛掷一个硬币直到 HHT 或 HTT 出现; 如果型式 HHT 先出现, 则 Alice 获胜, 如果型式 HTT 先出现, 则 Bill 获胜。此游戏现称它为“Penney 赌注”, 如果用一个完好的硬币, 看来确实是公平的, 因为孤立地看两种型式 HHT 和 HTT 时它们有相同的特征: 等待直到 HHT 第一次出现的概率母函数是

$$G(z) = \frac{z^2}{z^3 - 8(z-1)},$$

对 HTT 也有相同的概率母函数。所以如果他们玩单人纸牌戏, Alice 和 Bill 都没有有利条件。

但是当同时考虑两者时, 型式之间有一种有趣的相互影响。设 S_A 是 Alice 获胜构造的和, S_B 是 Bill 获胜构造的和:

$$S_A = \text{HHT} + \text{HHHT} + \text{THHT} + \text{HHHHT} + \text{HTHHT} + \text{THHHT} + \cdots;$$

$$S_B = \text{HTT} + \text{THTT} + \text{HTHTT} + \text{TTHTT} + \text{THTHTT} + \text{TTTHTT} + \cdots.$$

再(从处理仅涉及一种型式的诀窍取得暗示)让我们用 N 记至今没有胜者的所有序列的和:

$$N = 1 + \text{H} + \text{T} + \text{HH} + \text{HT} + \text{TH} + \text{TT} + \text{HHH} + \text{HTH} + \text{THH} + \cdots. \quad (8.77)$$

于是我们容易验证下列方程组:

$$1 + N(\text{H} + \text{T}) = N + S_A + S_B;$$

$$N\text{HHT} = S_A; \quad (8.78)$$

$$N\text{HTT} = S_A T + S_B.$$

如果现在置 $\text{H} = \text{T} = 1/2$, S_A 的产生的值变成 Alice 获胜的概率, S_B 变成 Bill 获胜的概率。三个方程化为

$$1 + N = N + S_A + S_B, \quad \frac{1}{8}N = S_A, \quad \frac{1}{8}N = \frac{1}{2}S_A + S_B;$$

且我们求得 $S_A = 2/3$, $S_B = 1/3$ 。Alice 获胜大约是 Bill 获胜的两倍!

在此游戏的推广中, Alice 和 Bill 选择正面和反面的型式 A 和 B , 且他们抛掷硬币直到 A 或者 B 出现。两种型式不必有相同长度, 但是我们假设在 B 内 A 不出现, 在 A 内 B 也不出现。(否则游戏将退化。例如, 若 $A = \text{HT}$ 和 $B = \text{THTH}$, 不高明的 Bill 永不获胜; 若 $A = \text{HTH}$ 和 $B = \text{TH}$, 两个游戏者可能同时宣称胜利。)相似于式(8.73)和(8.78), 我们能写出三个方程:

$$1 + N(\text{H} + \text{T}) = N + S_A + S_B;$$

$$NA = S_A \sum_{k=1}^l A^{(l-k)} [A^{(k)} = A_{(k)}] + S_B \sum_{k=1}^{\min(l, m)} A^{(l-k)} [B^{(k)} = A_{(k)}];$$

$$NB = S_A \sum_{k=1}^{\min(l, m)} B^{(m-k)} [A^{(k)} = B_{(k)}] + S_B \sum_{k=1}^m B^{(m-k)} [B^{(k)} = B_{(k)}]. \quad (8.79)$$

这里 l 是 A 的长度, m 是 B 的长度。例如, 如果我们有 $A = \text{HTTHTHTH}$ 和 $B = \text{THTHTTH}$, 两个型式依赖方程是

$$NHTTHTHTH = S_A THTHTH + S_A + S_B THTHTH + S_B THTH;$$

$$NTHTHTTH = S_A THTTH + S_A TTH + S_B THTTH + S_B.$$

如果假设用一个完好的硬币，我们通过置 $H=T=1/2$ 获得胜利的概率；这就把两个关键方程化成

$$\begin{aligned} N &= S_A \sum_{k=1}^l 2^k [A^{(k)} = A_{(k)}] + S_B \sum_{k=1}^{\min(l, m)} 2^k [B^{(k)} = A_{(k)}]; \\ N &= S_A \sum_{k=1}^{\min(l, m)} 2^k [A^{(k)} = B_{(k)}] + S_B \sum_{k=1}^m 2^k [B^{(k)} = B_{(k)}]. \end{aligned} \quad (8.80)$$

如果把式(8.76)的 $A:A$ 运算推广到两个独立串 A 和 B 的函数:

$$A:B = \sum_{k=1}^{\min(l, m)} 2^{k-1} [A^{(k)} = B_{(k)}], \quad (8.81)$$

现在方程(8.80)简单地变成

$$S_A(A:A) + S_B(B:A) = S_A(A:B) + S_B(B:B);$$

Alice 喜爱发生的可能性为

$$\frac{S_A}{S_B} = \frac{B:B - B:A}{A:A - A:B}. \quad (8.82)$$

(John Horton Conway 发现此优美的公式^[11].)

例如，若像上面那样 $A=HTTHTHTH$ 和 $B=THTHTTH$ ，我们有 $A:A = (10000001)_2 = 129$ ， $A:B = (0001010)_2 = 10$ ， $B:A = (0001001)_2 = 9$ ， $B:B = (1000010)_2 = 66$ ；所以比 S_A/S_B 是 $(66-9)/(129-10) = 57/119$ 。平均来说，每 176 次 Alice 将仅胜 57 次。

在 Penney 游戏中能发生奇怪的事情。例如，以 $3/2$ 可能性型式 HHTH 获胜超过型式 HTHH，以 $7/5$ 可能性型式 HTHH 获胜超过 THHH。所以 HHTH 应比 THHH 好许多。然而 THHH 实际以 $7/5$ 可能性获胜超过 HHTH！型式之间的关系不是传递的。事实上，习题 57 证明了如果 Alice 选择长度 $l \geq 3$ 的任何型式 $\tau_1 \tau_2 \cdots \tau_l$ ，若 Bill 选择型式 $\bar{\tau}_2 \tau_1 \tau_2 \cdots \tau_{l-1}$ ，其中 $\bar{\tau}_2$ 是 τ_2 (正面/反面) 的反面，它总能保证还有较好的获胜机会。

8.5 散列法

让我们通过把概率论应用到计算机程序设计来结束本章。关于计算机内部存贮和检索信息的几个重要算法是基于称为“散列法”的技巧。一般的问题是维持一组记录，每个记录包含一个“键”值 K ，以及维持关于此键的某个数据 $D(K)$ ；当给定 K 时，我们要能快速找

到 $D(K)$ 。例如, 每个键可以是一个学生的名字, 有关的数据可以是学生的家庭作业的分。

在实践中, 计算机没有足够容量对每个可能键拨出一个存贮单元; 可能有几十亿个键, 但是在任何一种应用中只有相当少的键实际出示。问题的一个解是维持两个表 $KEY[j]$ 和 $DATA[j]$ ($1 \leq j \leq N$), 其中 N 为能提供的记录的总数; 另一个变量 n 说出实际出示多少个记录。于是我们能以一种显而易见的方式序列地通过表来查找给定的键 K :

- S1 置 $j := 1$ 。(查找完所有位置 $< j$ 。)
- S2 若 $j > n$, 结束。(查找不成功。)
- S3 若 $KEY[j] = K$, 结束。(查找成功。)
- S4 j 增加 1, 返回到 S2。(我们再试。)

在一次成功查找之后, 在 $DATA[j]$ 中出现希望的数据项 $D(k)$ 。在一次不成功查找之后, 假设表还未填满容量, 通过置

$$n := j, \text{ KEY}[n] := K, \text{ DATA}[n] := D(K)$$

我们能将 K 和 $D(K)$ 插入表。

这个方法行得通, 但是它极慢; 每当作一次不成功的查找, 且 n 相当大时, 我们需重复 S2 $n+1$ 次。

发现的散列法使速度加快。在一种流行的形式中, 基本思想是用 m 个分离的表来代替一个巨大的表。一个“散列函数”把每个可能键 K 转换为 1 和 m 之间的一个列表数 $h(K)$ 。一个附加表 $FIRST[i]$ 指向表 i 中的第一个记录 ($1 \leq i \leq m$); 另一个附加表 $NEXT[j]$ 指向它的列表中的跟随记录 j 后面的记录 ($1 \leq j \leq N$)。我们假设

$FIRST[i] = -1$, 若列表 i 是空的;

$NEXT[j] = 0$, 若记录 j 是它的列表中的最后一个。如同前面那样, 有一个变量 n 告知全部存贮多少个记录。

例如, 假设键是名字, 且假设有基于名字的第一个字母的 $m = 4$ 个列表:

$$h(\text{名字}) = \begin{cases} 1, & \text{A} - \text{F}; \\ 2, & \text{G} - \text{L}; \\ 3, & \text{M} - \text{R}; \\ 4, & \text{S} - \text{Z}. \end{cases}$$

我们从 4 个空的列表开始, 且具有 $n = 0$ 。譬如说, 如果第一个记录有 Nora 作为它的键, 我们有 $h(\text{Nora}) = 3$, 所以 Nora 变成表 3 中第一项的键。如果接着两个名字是 Glenn 和 Jim, 它们都进入表 2。现在内存中的表似为如下所示:

$FIRST[1] = -1, FIRST[2] = 2, FIRST[3] = 1, FIRST[4] = -1.$
 $KEY[1] = \text{Nora}, NEXT[1] = 0;$
 $KEY[2] = \text{Glenn}, NEXT[2] = 3;$
 $KEY[3] = \text{Jim}, NEXT[3] = 0; n = 3.$

($DATA[1], DATA[2]$ 和 $DATA[3]$ 的值是保密的, 将不显示。)插入 18 个记录之后, 列表可能包含名字

表 1	表 2	表 3	表 4
Dianne	Glenn	Nora	Scott
Ari	Jim	Mike	Tina
Brian	Jennifer	Michael	
Fran	Joan	Ray	
Doug	Jerry	Paula	
	Jean		

且这些名字和 NEXT 项目混合在一起在 KEY 数组中出现, 使表保持有效地分离。如果现在要找 John, 我们必须扫描完表 2 中的 6 个名字(碰巧是最长的表); 但是远非查看所有 18 名字一样差。

下列是依照此方案查找键 K 的算法的精确说明:

H1 置 $i := h(K)$ 和 $j := \text{FIRST}[i]$.

H2 若 $j \leq 0$, 终止。(查找不成功。)

H3 若 $\text{KEY}[j] = K$, 终止。(查找成功。)

H4 置 $i := j$, 则置 $j := \text{NEXT}[i]$ 且返回到 H2。(我们将再试。)

例如, 在给定的例子中查找 Jennifer, H1 将置 $i := 2$ 和 $j := 2$; H3 将发现 $\text{Glenn} \neq \text{Jennifer}$; H4 将置 $j := 3$; H3 将发现 $\text{Jim} \neq \text{Jennifer}$ 。

在前面的算法中, 一次成功查找之后, 在 $\text{DATA}[j]$ 中出现希望的数据 $D(K)$ 。在一次不成功查找之后, 通过作下列操作:

$n := n + 1$;

if $j < 0$ then $\text{FIRST}[i] := n$ else $\text{NEXT}[i] := n$;

$\text{KEY}[n] := K$; $\text{DATA}[n] := D(K)$; $\text{NEXT}[n] := 0$. (8.83)

我们能把 K 和 $D(K)$ 送进表。现在的表将是最近的表。

我们希望取得大致相等长度的表, 因为这将使大约 m 次查找任务较快完成。 m 的值通常比 4 大很多, 所以 $1/m$ 的一个因子将是有意改进。

我们预先不知道将显示什么键, 但是一般可能选取散列函数以致能把 $h(K)$ 考虑为在 1 和 m 之间均匀分布, 独立于显示的其他键的散列值的随机变量。这种情形下, 计算散列函数就像滚动有 m 面的一个骰子。有一种可能性是所有记录将落入同一个表, 就像有一种可能性是一个骰子总是面 1 朝上; 但是概率论告诉我们, 表几乎总是相当一致地均衡。

散列法的分析: 引言

“算法分析”是计算机科学的一个分支, 它引出计算机方法有效性的量的信息。“一个算法的概率分析”是算法运行时间的研究, 考虑为依赖于输入数据的呈现特性的随机变量。散列法是概率分析的特别好的候选方法, 因为平均来说它是极有效的方法, 虽然它的最坏情形太糟。(当所有键有相同散列值时最坏情形出现。)事实上, 用散列法的计算机程序设计者最好还是一个概率论的信仰者。

当用上面的算法进行查找时, 执行 H3 的次数记为 P 。(H3 的每次执行称为表中的一个“探索”。)如果我们知道 P , 依据查找是成功还是不成功, 可知执行每步多少次:

步	不成功查找	成功查找
H1	1 次	1 次
H2	$P+1$ 次	P 次
H3	P 次	P 次
H4	P 次	$P-1$ 次

因此决定查找过程的运行时间的主要量是探索次数 P 。

我们能取得算法的一种好的易记的描述，想象我们保存有一本以一种特殊方式组织的地址本，每页仅有一个项目的地位。在本子的封面上记下 m 个表的每个表中第一个项目的页数；每个名字 K 确定它所属的表 $h(K)$ 。本子内的每一页指的是它的表中的后继页。在这样的本子中找一个地址所需的探索次数是我们一定要查阅的页数。

如果已插入 n 项，在表中它们的位置仅依赖于各自的散列值 $\langle h_1, h_2, \dots, h_n \rangle$ 。考虑的 m^n 个可能序列 $\langle h_1, h_2, \dots, h_n \rangle$ 的每一个序列是等可能的，且 P 是依赖于这样一个序列的随机变量。

情形 1: 不显示的键

让我们首先考虑在一次不成功查找中的 P 的情况，假设预先已把 n 个记录插入散列表，此时，相应的概率空间包含 m^{n+1} 个基本事件

$$\omega = (h_1, h_2, \dots, h_n; h_{n+1})$$

其中 h_j 是插入的第 j 个键的散列值， h_{n+1} 是查找不成功的键的散列值。我们假设已适当选取散列函数 h 以致对每一个这样的 ω ， $\Pr(\omega) = 1/m^{n+1}$ 。

例如，如果 $m=n=2$ ，有 8 种等可能的可能结果：

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

若 $h_1 = h_2 = h_3$ ，在断定不显示新键 K 之前我们作两次不成功的探索；若 $h_1 = h_2 \neq h_3$ ，我们不作；等等。这个所有可能结果的表表明 P 有概率母函数 $(2/8 + 4/8 \cdot z + 2/8 \cdot z^2) = (1/2 + 1/2 \cdot z)^2$ (当 $m=n=2$ 时) 给定的概率分布。

对于表中每一项，一次不成功查找作一次探索计入 h_{n+1} ，所以我们有一般公式

$$P = [h_1 = h_{n+1}] + [h_2 = h_{n+1}] + \dots + [h_n = h_{n+1}]. \quad (8.84)$$

$h_j = h_{n+1}$ 的概率为 $1/m$ ($1 \leq j \leq n$)。所以由此得到

$$EP = E[h_1 = h_{n+1}] + E[h_2 = h_{n+1}] + \cdots + E[h_n = h_{n+1}] = \frac{n}{m}.$$

我们可能更慢地算出: 设 X_j 是随机变量

$$X_j = X_j(\omega) = [h_j = h_{n+1}].$$

于是 $P = X_1 + \cdots + X_n$ 和 $EX_j = 1/m$ (所有 $j \leq n$); 因此

$$EP = EX_1 + \cdots + EX_n = n/m.$$

正如我们希望的那样, 无散列的平均探索次数是 $1/m$ 次。而且随机变量 X_j 是独立的, 且它们的每一个有相同的概率母函数

$$X_j(z) = \frac{m-1+z}{m};$$

所以一次不成功查找中探索总次数的概率母函数是

$$F(z) = X_1(z) \cdots X_n(z) = \left(\frac{m-1+z}{m} \right)^n. \quad (8.85)$$

这是 $p = 1/m$, $q = (m-1)/m$ 的二项分布; 换句话说, 在一次不成功查找中探索次数的情况就像当我们掷一个无偏硬币时的正面数, 在每次抛掷中正面的概率为 $1/m$ 。方程 (8.61) 告诉我们, P 的方差为

$$npq = \frac{n(m-1)}{m^2}.$$

当 m 大时, P 的方差近似为 n/m , 所以标准差近似为 $\sqrt{n/m}$ 。

情形 2: 显示的键

现在让我们查看成功查找。此时依赖于应用的适合的概率空间有点复杂: 我们设 Ω 是所有基本事件

$$\omega = (h_1, \cdots, h_n; k) \quad (8.86)$$

的集合, 其中 h_j 是像前面的第 j 个键的散列值, k 是寻找的键的指标 (键的散列值是 h_k)。因此我们有 $1 \leq h_j \leq m$ ($1 \leq j \leq n$), 以及 $1 \leq k \leq n$; 总共有 $m^n \cdot n$ 个基本事件 ω 。

设 s_j 是查找插入表中的第 j 个键的概率。如果 ω 是事件 (8.86), 则

$$\Pr(\omega) = s_k / m^n. \quad (8.87)$$

(在某些应用中最常查找首先插入的或最后插入的项, 所以我们将不假设每个 $s_j = 1/n$ 。)

注意, $\sum_{\omega \in \Omega} \Pr(\omega) = \sum_{k=1}^n s_k = 1$, 因此式 (8.87) 定义了一个合法的概率分布。

如果键 K 是插入它的表中的第 p 个键, 则在一次成功查找中探索次数 P 是 p 。所以

$$P = [h_1 = h_k] + [h_2 = h_k] + \cdots + [h_k = h_k];$$

或, 如果设 X_j 是随机变量 $[h_j = h_k]$, 我们有

$$P = X_1 + X_2 + \cdots + X_k. \quad (8.88)$$

例如, 假设我们有 $m=10$ 和 $n=16$, 且散列值有下列“随机”型式:

$$(h_1, \dots, h_{16}) = 3141592653589793;$$

$$(P_1, \dots, P_{16}) = 1112111122312133.$$

找第 j 个键所需的探索次数 P_j 表明在 h_j 下面.

方程 (8.88) 把 P 表为随机变量的和, 但是我们不能简单地把 EP 计算为 $EX_1 + \cdots + EX_k$, 因为量 k 本身是一个随机变量. P 的概率母函数是什么? 为了回答这个问题, 我们暂时离开主题来谈条件概率.

如果 A 和 B 是概率空间中的事件, 我们称给定 B 的条件下, A 的条件概率为

$$\Pr(\omega \in A | \omega \in B) = \frac{\Pr(\omega \in A \cap B)}{\Pr(\omega \in B)}. \quad (8.89)$$

例如, 若 X 和 Y 是随机变量, 给定 $Y=y$ 条件下事件 $X=x$ 的条件概率为

$$\Pr(X=x | Y=y) = \frac{\Pr(X=x \text{ 和 } Y=y)}{\Pr(Y=y)}. \quad (8.90)$$

对于 Y 中的任何指定的 y , X 中所有 x 上的这些条件概率的和为 $\Pr(Y=y) / \Pr(Y=y) = 1$; 所以式 (8.90) 定义了一个概率分布, 且我们能定义一个新的随机变量 ' $X|y$ ' 使得 $\Pr(X|y=x) = \Pr(X=x | Y=y)$.

若 X 和 Y 是独立的, 则随机变量本质上将与 X 相同, 而不管 y 的值, 因为根据式 (8.5), $\Pr(X=x | Y=y)$ 等于 $\Pr(X=x)$; 这就是独立的意思. 但是如果 X 和 Y 是依赖的, 当 $y \neq y'$ 时, 不管怎样随机变量 $X|y$ 和 $X|y'$ 不必彼此相象.

若 X 仅取非负整数值, 我们能把它概率母函数分解为关于任何其他的随机变量 Y 的条件概率母函数的和:

$$G_X(z) = \sum_{y \in Y(\Omega)} \Pr(Y=y) G_{X|y}(z). \quad (8.91)$$

因为对于所有 $x \in X(\Omega)$, 左边 z^x 的系数是 $\Pr(X=x)$, 右边 z^x 的系数是

$$\sum_{y \in Y(\Omega)} \Pr(Y=y) \Pr(X=x | Y=y) = \sum_{y \in Y(\Omega)} \Pr(X=x \text{ 和 } Y=y) = \Pr(X=x),$$

所以此式成立. 例如, 若 X 是两个完好的骰子的点数的乘积, Y 是点数的和, $X|6$ 的概率母函数是

$$G_{x_{16}}(z) = \frac{2}{5}z^5 + \frac{2}{5}z^8 + \frac{1}{5}z^9$$

因为 $Y=6$ 的条件概率包含 5 个等可能事件 $\{\square\square\square, \square\square\square, \square\square\square, \square\square\square\square\square, \square\square\square\square\square\}$ 。此时方程 (8.91) 化成

$$\begin{aligned} G_X(z) = & \frac{1}{36}G_{x_{12}}(z) + \frac{2}{36}G_{x_{13}}(z) + \frac{3}{36}G_{x_{14}}(z) + \frac{4}{36}G_{x_{15}}(z) \\ & + \frac{5}{36}G_{x_{16}}(z) + \frac{6}{36}G_{x_{17}}(z) + \frac{5}{36}G_{x_{18}}(z) + \frac{4}{36}G_{x_{19}}(z) \\ & + \frac{3}{36}G_{x_{110}}(z) + \frac{2}{36}G_{x_{111}}(z) + \frac{1}{36}G_{x_{112}}(z), \end{aligned}$$

显而易见这是以前了解的公式。(结束离题的话。)

在散列法的情形中, 如果设 $X=P$, $Y=K$, 式(8.91)告诉我们在一次成功查找中如何写出探索的概率母函数。对于 1 和 n 之间的任何指定的 k , 随机变量 P/k 定义为独立随机变量的和 $X_1 + \dots + X_k$, 这就是式(8.88)。所以它有概率母函数

$$G_{P/k}(z) = \left(\frac{m-1+z}{m} \right)^{k-1} z.$$

所以 P 的概率母函数是

$$G_P(z) = \sum_{k=1}^n s_k G_{P/k}(z) = \sum_{k=1}^n s_k \left(\frac{m-1+z}{m} \right)^{k-1} z = z S \left(\frac{m-1+z}{m} \right), \quad (8.92)$$

其中

$$S(z) = s_1 + s_2 z + s_3 z^2 + \dots + s_n z^{n-1} \quad (8.93)$$

是查找概率 s_k 的概率母函数(为了方便除以 z)。

我们有 P 的概率母函数; 现在我们能通过导数求得平均值和方差, 如同前面所做的那样, 首先移去 z 因子稍微容易, 因此用找 $P-1$ 的平均值和方差来替代:

$$F(z) = G_P(z) / z = S \left(\frac{m-1+z}{m} \right);$$

$$F'(z) = \frac{1}{m} S' \left(\frac{m-1+z}{m} \right);$$

$$F''(z) = \frac{1}{m^2} S'' \left(\frac{m-1+z}{m} \right).$$

所以

$$EP = 1 + \text{平均值}(F) = 1 + F'(1) = 1 + m^{-1} \text{平均值}(S); \quad (8.94)$$

$$\begin{aligned} VP = \text{方差}(F) &= F''(1) + F'(1) - F'(1)^2 \\ &= m^{-2} S''(1) + m^{-1} S'(1) - m^{-2} S'(1)^2 \end{aligned}$$

$$= m^{-2} \text{方差}(S) + (m^{-1} - m^{-2}) \text{平均值}(S). \quad (8.95)$$

这些是依据采用的查找分布 S 来表达探索次数 P 的平均值和方差的一般公式。

例如, 假设我们有 $s_k = 1/n$ ($1 \leq k \leq n$), 这意味着我们正在做完全“随机的”成功查找, 表中所有键为等可能的, 于是 $S(z)$ 是式(8.32)中的均匀概率分布 $U_n(z)$, 且我们有平均值 $\langle S \rangle = (n-1)/2$, 方差 $\langle S^2 \rangle = (n^2 - 1)/12$ 。因此

$$EP = \frac{n-1}{2m} + 1; \quad (8.96)$$

$$VP = \frac{n^2 - 1}{12m^2} + \frac{(m-1)(n-1)}{2m^2} = \frac{(n-1)(6m+n-5)}{12m^2} \quad (8.97)$$

我们再一次获得希望的加速因子 $1/m$ 。如果 $m = n/\ln n$ 且 $n \rightarrow \infty$, 此时每次成功查找的平均探索次数大约为 $(1/2)\ln n$, 标准差渐近为 $(\ln n)/\sqrt{12}$ 。

另一方面, 我们可假设 $s_k = (kH_n)^{-1}$ ($1 \leq k \leq n$); 称这个分布为“Zipf 定律”, 于是平均值 $\langle G \rangle = n/H_n$, 方差 $\langle G^2 \rangle = 1/2n(n+1)/H_n - n^2/H_n^2$ 。当 $n \rightarrow \infty$ 时 $m = n/\ln n$ 的平均探索次数近似为 2, 标准差渐近于 $\sqrt{\ln n}/\sqrt{2}$ 。

在两种情形中, 分析让担心最坏情形的人放心: Chebyshev 不等式告诉我们, 除了极罕见的情形外, 表将是良好的和短的。

情形 2, 继续讨论: 方差的变形

刚才我们通过考虑 P 是具有 $m^n \cdot n$ 个元素 $(h_1, \dots, h_n; k)$ 的概率空间上的一个随机变量, 计算了一次成功查找中探索次数的方差。但是我们还能采用另一种观点: 散列值的每个型式 (h_1, \dots, h_n) 定义一个随机变量 $P|(h_1, \dots, h_n)$, 代表我们在 n 个给定键的一个指定散列表的一次成功查找中所做的探索。 $P|(h_1, \dots, h_n)$ 的平均值

$$A(h_1, \dots, h_n) = \sum_{p=1}^n p \cdot \Pr(P|(h_1, \dots, h_n) = p) \quad (8.98)$$

可称为代表一次成功查找的运行时间。此量 $A(h_1, \dots, h_n)$ 是一个随机变量, 它依赖于 (h_1, \dots, h_n) , 而不依赖于最后的成分 k ; 我们可把它写成形式

$$A(h_1, \dots, h_n) = \sum_{k=1}^n s_k P(h_1, \dots, h_n; k),$$

由于 $P|(h_1, \dots, h_n) = p$ 具有概率

$$\begin{aligned} \frac{\sum_{k=1}^n \Pr(P(h_1, \dots, h_n; k) = p)}{\sum_{k=1}^n \Pr(h_1, \dots, h_n; k)} &= \frac{\sum_{k=1}^n m^{-n} s_k [P(h_1, \dots, h_n; k) = p]}{\sum_{k=1}^n m^{-n} s_k} \\ &= \sum_{k=1}^n s_k [P(h_1, \dots, h_n; k) = p]. \end{aligned}$$

$A(h_1, \dots, h_n)$ 的平均值得自在所有 m^n 个可能结果上求和且除以 m^n ，它将与前面我们得到的式(8.94)中的平均值相同。但是 $A(h_1, \dots, h_n)$ 的方差有点不一样了，这是 m^n 个平均的方差，而不是 $m^n \cdot n$ 个探索计数的方差。例如，若 $m=1$ (以致仅有一个表)，“平均”值 $A(h_1, \dots, h_n) = A(1, \dots, 1)$ 实际是常数，所以它的方差 VA 是零；但是一次成功查找中的探索次数不是常数，所以方差是非零。

我们可通过当 $s_k = 1/n (1 \leq k \leq n)$ 时，在最简单情形中对一般的 m 和 n 进行计算，来说明方差之间的这种不同。换句话说，我们暂时假设有一个查找键的均匀分布。任何给定的散列值的序列 (h_1, \dots, h_n) 定义 m 个表，对于某些数 n_j 分别包含 (n_1, n_2, \dots, n_m) 项，其中

$$n_1 + n_2 + \dots + n_m = n.$$

表中 n 个键的每一个是等可能的，一次成功查找将得到探索的平均运行时间

$$\begin{aligned} A(h_1, \dots, h_n) &= \frac{(1 + \dots + n_1) + (1 + \dots + n_2) + \dots + (1 + \dots + n_m)}{n} \\ &= \frac{n_1(n_1 + 1) + n_2(n_2 + 1) + \dots + n_m(n_m + 1)}{2n} \\ &= \frac{n_1^2 + n_2^2 + \dots + n_m^2 + n}{2n}. \end{aligned}$$

我们的目的是在包含所有 m^n 个序列 (h_1, \dots, h_n) 的概率空间上计算这个量 $A(h_1, \dots, h_n)$ 的方差。

计算将是简单的，如果我们计算一个稍微不同的量

$$B(h_1, \dots, h_n) = \binom{n_1}{2} + \binom{n_2}{2} + \dots + \binom{n_m}{2}$$

的方差，就得到结果。我们有

$$A(h_1, \dots, h_n) = 1 + B(h_1, \dots, h_n)/n,$$

因此 A 的平均值和方差满足

$$EA = 1 + \frac{EB}{n}; \quad VA = \frac{VB}{n^2}. \quad (8.99)$$

表的大小为 n_1, n_2, \dots, n_m 的概率是多项式系数

$$\binom{n}{n_1, n_2, \dots, n_m} = \frac{n!}{n_1! n_2! \dots n_m!}$$

除以 m^n ；因此 $B(h_1, \dots, h_n)$ 的概率母函数是

$$B_n(z) = \sum_{\substack{n_1, n_2, \dots, n_m \geq 0 \\ n_1 + n_2 + \dots + n_m = n}} \binom{n}{n_1, n_2, \dots, n_m} z^{\binom{n_1}{2} + \binom{n_2}{2} + \dots + \binom{n_m}{2}} m^{-n}.$$

对于无经验者这个和看来有点害怕，但是第七章中已教我们把它认为是一个 m 次卷叠的卷积。事实上，如果考虑指数超母函数

$$G(w, z) = \sum_{n \geq 0} B_n(z) \frac{m^n w^n}{n!},$$

容易验证， $G(w, z)$ 仅为 m 次幂：

$$G(w, z) = \left(\sum_{k \geq 0} z^{\binom{k}{2}} \frac{w^k}{k!} \right)^m.$$

作为一个检验，试置 $z=1$ ；我们得到 $G(w, 1) = (e^w)^m$ ，所以 $m^n w^n / n!$ 的系数是 $B_n(1) = 1$ 。

如果知道 $B'_n(1)$ 和 $B''_n(1)$ ，我们将能计算方差 (B_n) ，所以我们取 $G(w, z)$ 对 z 的偏导数：

$$\begin{aligned} \frac{\partial}{\partial z} G(w, z) &= \sum_{n \geq 0} B'_n(z) \frac{m^n w^n}{n!} \\ &= m \left(\sum_{k \geq 0} z^{\binom{k}{2}} \frac{w^k}{k!} \right)^{m-1} \sum_{k \geq 0} \binom{k}{2} z^{\binom{k}{2}-1} \frac{w^k}{k!}; \\ \frac{\partial^2}{\partial z^2} G(w, z) &= \sum_{n \geq 0} B''_n(z) \frac{m^n w^n}{n!} \\ &= m(m-1) \left(\sum_{k \geq 0} z^{\binom{k}{2}} \frac{w^k}{k!} \right)^{m-2} \left(\sum_{k \geq 0} \binom{k}{2} z^{\binom{k}{2}-1} \frac{w^k}{k!} \right)^2 \\ &\quad + m \left(\sum_{k \geq 0} z^{\binom{k}{2}} \frac{w^k}{k!} \right)^{m-1} \sum_{k \geq 0} \binom{k}{2} \left(\binom{k}{2} - 1 \right) z^{\binom{k}{2}-2} \frac{w^k}{k!}. \end{aligned}$$

复杂的结果；但是当我们置 $z=1$ 时有很大的简化。例如，我们有

$$\begin{aligned} \sum_{n \geq 0} B'_n(1) \frac{m^n w^n}{n!} &= m e^{(m-1)w} \sum_{k \geq 2} \frac{w^k}{2(k-2)!} \\ &= m e^{(m-1)w} \sum_{k \geq 0} \frac{w^{k+2}}{2k!} \\ &= \frac{m w^2 e^{(m-1)w}}{2} e^w = \sum_{n \geq 0} \frac{(m w)^{n+2}}{2 m n!} = \sum_{n \geq 0} \frac{n(n-1) m^n w^n}{2 m n!}, \end{aligned}$$

并由此得到

$$B'_n(1) = \binom{n}{2} \frac{1}{m}. \quad (8.100)$$

现在式(8.99)中 EA 的表达式给出 $EA = 1 + (n-1)/2m$, 与式(8.96)相同.

$B'_n(1)$ 的公式涉及类似的和

$$\begin{aligned} \sum_{k \geq 0} \binom{k}{2} \left(\binom{k}{2} - 1 \right) \frac{w^k}{k!} &= \frac{1}{4} \sum_{k \geq 0} \frac{(k+1)k(k-1)(k-2)w^k}{k!} \\ &= \frac{1}{4} \sum_{k \geq 3} \frac{(k+1)w^k}{(k-3)!} = \frac{1}{4} \sum_{k \geq 0} \frac{(k+4)w^{k+3}}{k!} = \left(\frac{1}{4} w^4 + w^3 \right) e^w; \end{aligned}$$

因此我们求得

$$\begin{aligned} \sum_{n \geq 0} B'_n(1) \frac{m^n w^n}{n!} &= m(m-1) e^{w(m-2)} \left(\frac{1}{2} w^2 e^w \right)^2 + m e^{w(m-1)} \left(\frac{1}{4} w^4 + w^3 \right) e^w \\ &= m e^{wm} \left(\frac{1}{4} m w^4 + w^3 \right); \\ B'_n(1) &= \binom{n}{2} \left(\binom{n}{2} - 1 \right) \frac{1}{m}. \end{aligned} \quad (8.101)$$

现在可把所有结果放在一起, 且计算希望的方差 VA . 出现大量相消, 且结果意外的简单:

$$\begin{aligned} VA = \frac{VB}{n^2} &= \frac{B''_n(1) + B'_n(1) - B'_n(1)^2}{n^2} \\ &= \frac{n(n-1)}{m^2 n^2} \left(\frac{(n+1)(n-2)}{4} + \frac{m}{2} - \frac{n(n-1)}{4} \right) \\ &= \frac{(m-1)(n-1)}{2m^2 n}. \end{aligned} \quad (8.102)$$

当出现这种“巧合”时, 我们觉得有数学的原因; 可能有另一种方式来处理问题, 说明为什么解答有这样一个简单的形式. 事实上, 有另一种方法(习题 60), 它表明当 s_k 是寻找的第 k 个插入元素的概率时, 平均成功查找的方差有一般形式

$$VA = \frac{m-1}{m^2} \sum_{k=1}^n s_k^2 (k-1). \quad (8.103)$$

方程(8.102)是特殊情形 $s_k = 1/n (1 \leq k \leq n)$.

除了平均的方差外, 我们还可考虑方差的平均. 换句话说, 定义散列表的每个序列 (h_1, \dots, h_n) 也定义了成功查找的概率分布, 且此概率分布的方差告知在不同的成功查找中如何散布探索次数. 例如, 让我们回到把 $n=16$ 个插入 $m=10$ 个表的情形:

$$(h_1, \dots, h_{16}) = 3141592653589793$$

$$(P_1, \dots, P_{16}) = 1112111122312133.$$

在产生的散列表中一次成功查找有概率母函数

$$\begin{aligned} G(3, 1, 4, 1, \dots, 3) &= \sum_{k=1}^{16} s_k z^{P(3, 1, 4, 1, \dots, 3; k)} \\ &= s_1 z + s_2 z + s_3 z + s_4 z^2 + \dots + s_{16} z^3. \end{aligned}$$

刚才我们已考虑此表的一次成功查找中的平均探索次数, 即 $A(3, 1, 4, 1, \dots, 3)$ = 平均值($G(3, 1, 4, 1, \dots, 3)$). 我们还能考虑方差,

$$\begin{aligned} &s_1 \cdot 1^2 + s_2 \cdot 1^2 + s_3 \cdot 1^2 + s_4 \cdot 2^2 + \dots + s_{16} \cdot 3^2 \\ &\quad - (s_1 \cdot 1 + s_2 \cdot 1 + s_3 \cdot 1 + s_4 \cdot 2 + \dots + s_{16} \cdot 3)^2. \end{aligned}$$

这个方差是依赖于 (h_1, \dots, h_n) 的随机变量, 所以考虑它的平均值是自然的.

换句话说, 为了懂得成功查找的情况, 我们可能希望知道的方差有三种自然的类型: 探索次数的总和方差, 在所有 (h_1, \dots, h_n) 和 k 上取得; 探索的平均次数的方差, 其中平均是在所有 k 上取得, 方差是在所有 (h_1, \dots, h_n) 上取得; 探索次数的方差的平均, 其中方差是在所有 k 上取得, 平均是在所有 (h_1, \dots, h_n) 上取得. 以符号表示, 总方差为

$$\begin{aligned} VP &= \sum_{1 \leq h_1, \dots, h_n \leq m} \sum_{k=1}^n \frac{s_k}{m^n} P(h_1, \dots, h_n; k)^2 \\ &\quad - \left(\sum_{1 \leq h_1, \dots, h_n \leq m} \sum_{k=1}^n \frac{s_k}{m^n} P(h_1, \dots, h_n; k) \right)^2; \end{aligned}$$

平均的方差为

$$\begin{aligned} VA &= \sum_{1 \leq h_1, \dots, h_n \leq m} \frac{1}{m^n} \left(\sum_{k=1}^n s_k P(h_1, \dots, h_n; k) \right)^2 \\ &\quad - \left(\sum_{1 \leq h_1, \dots, h_n \leq m} \frac{1}{m^n} \sum_{k=1}^n s_k P(h_1, \dots, h_n; k) \right)^2; \end{aligned}$$

方差的平均为

$$\begin{aligned} AV &= \sum_{1 \leq h_1, \dots, h_n \leq m} \frac{1}{m^n} \left(\sum_{k=1}^n s_k P(h_1, \dots, h_n; k) \right)^2 \\ &\quad - \left(\sum_{k=1}^n s_k P(h_1, \dots, h_n; k)^2 \right). \end{aligned}$$

结果, 以简单的方式联系这样三个量:

$$VP = VA + AV. \quad (8.104)$$

事实上, 如果 X 和 Y 是任何概率空间中的随机变量, X 取实值, 条件概率分布总满足等式

$$VX = V(E(X|Y)) + E(V(X|Y)). \quad (8.105)$$

(习题 22 中证明此等式。)方程(8.104)是 X 为一次成功查找中的探索次数, Y 是散列值的序列 (h_1, \dots, h_n) 的特殊情形。

一般方程(8.105)需要仔细地理解, 因为记法往往会隐蔽不同的随机变量和概率空间中计算的期望和方差。对于 Y 中的每个 y , 在式(8.90)中我们定义了随机变量 $X|y$, 且这个随机变量有一个依赖于 y 的期望值 $E(X|y)$ 。现在 $E(X|Y)$ 表示随机变量, 当 y 在所有 Y 的可能值上变化时, 它的值是 $E(X|y)$, 且 $V(E(X|Y))$ 是这个随机变量关于 Y 的概率分布的方差。当 y 变化时, $E(V(X|Y))$ 相似地为随机变量 $V(X|y)$ 的平均。式(8.105)的左边是 X 的无条件方差 VX 。由于方差是非负的, 我们总有

$$VX \geq V(E(X|Y)) \text{ 和 } VX \geq E(V(X|Y)). \quad (8.106)$$

情形 1, 再回到不成功查找

让我们对散列法的细致研究再作一次算法分析的典型计算来结束。这次让我们仔细查看与一次不成功查找联系的总运行时间, 假设计算机把以前未知的键插入它的内存。

式(8.83)中的插入过程有两种情形, 依赖于 j 是负还是零。我们有 $j < 0$ 当且仅当 $P=0$, 由于一个负值来自一个空表的 FIRST 项。因此, 如果表以前是空的, 有 $P=0$, 我们一定要置 $\text{FIRST}[h_{n+1}] := n+1$ 。(新的记录将插入位置 $n+1$ 。)否则有 $P > 0$, 且我们一定要把一个 LINK 项置入 $n+1$ 。这样两种情形可取不同的时间量; 所以一次不成功查找的总运行时间有形式

$$T = \alpha + \beta P + \delta[P=0], \quad (8.107)$$

其中 α , β 和 δ 是常数, 它们依赖于所用的计算机以及以机器内部语言编码散列法的方式, 知道 T 的平均值和方差是好的, 因为在实践中这种信息比 P 的平均值和方差更合适。

到现在为止, 我们所用的概率母函数仅与取非负整数值的随机变量有关。但是当 X 是任何实值随机变量时, 结果, 我们实质上能以处理

$$G_X(z) = \sum_{\omega \in \Omega} \text{Pr}(\omega) z^{X(\omega)}$$

以相同方式处理, 因为 X 的主要特性仅依赖于接近 $z=1$ 的 G_X 的情况, 在那里 z 的幂是正且明确的。例如, 一次不成功查找的运行时间式(8.107)是一个随机变量, 它定义于等可能的散列值 $(h_1, \dots, h_n; h_{n+1}) (1 \leq h_i \leq m)$ 的概率空间上; 我们考虑级数

$$G_T(z) = \frac{1}{m^{n+1}} \sum_{h_1=1}^m \dots \sum_{h_n=1}^m \sum_{h_{n+1}=1}^m z^{\alpha + \beta P(h_1, \dots, h_n; h_{n+1}) + \delta[P(h_1, \dots, h_n; h_{n+1})=0]}$$

是概率母函数, 甚至当 α , β 和 δ 不是整数时。(事实上, 参数 α , β , δ 是有时间量纲的物理量; 它们甚至不是纯粹的数! 然而我们能在 z 的幂中用它们。)通过计算 $G_T'(1)$ 和

$G_T'(1)$ 且以惯常方式结合这些值, 我们仍能计算 T 的平均值和方差.

P 而不是 T 的母函数是

$$P(z) = \left(\frac{m-1+z}{m} \right)^n = \sum_{p \geq 0} \Pr(P=p) z^p.$$

所以我们有

$$\begin{aligned} G_T(z) &= \sum_{p \geq 0} \Pr(P=p) z^{\alpha + \beta p + \delta(p-0)} \\ &= z^\alpha \left((z^\delta - 1) \Pr(P=0) + \sum_{p \geq 0} \Pr(P=p) z^{\beta p} \right) \\ &= z^\alpha \left((z^\delta - 1) \left(\frac{m-1}{m} \right)^n + \left(\frac{m-1+z}{m} \right)^n \right). \end{aligned}$$

现在确定平均值(G_T)和方差(G_T)是例行的工作:

$$\text{平均值}(G_T) = G_T'(1) = \alpha + \beta \frac{n}{m} + \delta \left(\frac{m-1}{m} \right)^n; \quad (8.108)$$

$$\begin{aligned} G_T''(1) &= \alpha(\alpha-1) + 2\alpha\beta \frac{n}{m} + \beta(\beta-1) \frac{n}{m} + \beta^2 \frac{n(n-1)}{m^2} \\ &\quad + 2\alpha\delta \left(\frac{m-1}{m} \right)^n + \delta(\delta-1) \left(\frac{m-1}{m} \right)^n; \end{aligned}$$

$$\begin{aligned} \text{方差}(G_T) &= G_T''(1) + G_T'(1) - G_T'(1)^2 = \beta^2 \frac{n(m-1)}{m^2} - 2\beta\delta \left(\frac{m-1}{m} \right)^n \frac{n}{m} \\ &\quad + \delta^2 \left(\left(\frac{m-1}{m} \right)^n - \left(\frac{m-1}{m} \right)^{2n} \right). \end{aligned} \quad (8.109)$$

在第九章中我们将学习当 m 和 n 大时如何来估计像这样的量. 例如, 若 $m=n$ 和 $n \rightarrow \infty$, 第九章的技巧将证明 T 的平均值和方差分别是 $\alpha + \beta + \delta e^{-1} + O(n^{-1})$ 和 $\beta^2 - 2\beta\delta e^{-1} + \delta^2(e^{-1} - e^{-2}) + O(n^{-1})$. 若 $m = n / \ln n$ 和 $n \rightarrow \infty$, 对应的结果是

$$\text{平均值}(G_T) = \beta \ln n + \alpha + \delta / n + O((\log n)^2 / n^2);$$

$$\text{方差}(G_T) = \beta^2 \ln n - ((\beta \ln n)^2 + 2\beta\delta \ln n - \delta^2) / n + O((\log n)^3 / n^2).$$

习 题

准备部分

1. 当一个骰子是完好的, 另一个是灌铅的, 式(8.3)的概率分布 \Pr_{01} 中成双的概率是

什么? 滚出 $S=7$ 的概率是多少?

2. 随机洗一副纸牌, 最上和最下的两张牌全为 A 的概率是多少? (所有 $52!$ 个排列有概率 $1/52!$.)

3. 在 1979 年问 Stanford 的具体数学学生, 抛掷硬币直到它们相继取得两次正面, 报告要求的抛掷次数。回答是

3, 2, 3, 5, 10, 2, 6, 6, 9, 2.

在 1987 年问 Princeton 的具体数学学生, 做相同事情, 有下列结果:

10, 2, 10, 7, 5, 2, 10, 6, 10, 2.

基于 (a) Stanford 样本, (b) Princeton 样本, 估计平均值和方差。

4. 设 $H(z) = F(z)/G(z)$, 其中 $F(1) = G(1) = 1$. 证明如果在 $z=1$ 处需要的导数存在, 相似于式 (8.38) 和 (8.39),

平均值(H) = 平均值(F) - 平均值(G),

方差(H) = 方差(F) - 方差(G).

5. 假设 Alice 和 Bill 用一个偏的硬币玩游戏 (8.78), 出现正面的概率为 p . 是否有一个 p 值, 对于此值游戏变成公平?

6. 当 X 和 Y 是独立随机变量时, 条件方差律 (8.105) 化成什么形式?

基本部分

7. 证明如果两个灌铅的骰子具有相同的概率分布, 则成双的概率总是至少 $1/6$.

8. 设 A 和 B 是事件使得 $A \cap B = \Omega$. 证明

$$\Pr(\omega \in A \cap B) = \Pr(\omega \in A)\Pr(\omega \in B) - \Pr(\omega \notin A)\Pr(\omega \notin B).$$

9. 证明或推翻: 如果 X 和 Y 是独立随机变量, 则当 F 和 G 是任何函数时, $F(X)$ 和 $G(Y)$ 也是独立随机变量。

10. 按照定义 (8.7), 能为随机变量 X 的中位数的元素的最大个数是多少?

11. 构造具有有限平均值和无限方差的一个随机变量。

12. (a) 如果 $P(z)$ 是随机变量 X 的概率母函数, 证明

$$\Pr(X \leq r) \leq x^{-r} P(x) \quad \text{对 } 0 < x \leq 1;$$

$$\Pr(X \geq r) \leq x^{-r} P(x) \quad \text{对 } x \geq 1.$$

(这些重要的关系称为末尾不等式。)

(b) 在特别情形 $P(z) = (1+z)^n / 2^n$ 下, 当 $0 < \alpha < 1/2$ 时, 用第一个末尾不等式来证明

$$\sum_{k \leq \alpha n} \binom{n}{k} \leq 1 / \alpha^{\alpha n} (1-\alpha)^{(1-\alpha)n}.$$

13. 如果 X_1, \dots, X_{2n} 是具有相同分布的独立随机变量, 且若 α 是任何实数, 证明

$$\Pr\left(\left|\frac{X_1 + \dots + X_{2n}}{2n} - \alpha\right| \leq \left|\frac{X_1 + \dots + X_n}{n} - \alpha\right|\right) \geq \frac{1}{2}.$$

14. 设 $F(z)$ 和 $G(z)$ 是概率母函数, 且设

$$H(z) = pF(z) + qG(z)$$

其中 $p+q=1$ 。(这称为 F 和 G 的混合; 它对应于抛掷一个硬币, 且依赖于硬币出现正面还是反面来选取概率分布 F 或 G 。)根据 p, q 以及 F 或 G 的平均值和方差来求 H 的平均值和方差。

15. 如果 $F(z)$ 和 $G(z)$ 是概率母函数, 我们能通过“复合”

$$H(z) = F(G(z))$$

来定义另一个概率母函数 $H(z)$ 。根据平均值(F), 方差(F), 平均值(G)和方差(G)来表达平均值(H)和方差(H)。(方程(8.92)是特殊情形。)

16. 当 $F_n(z)$ 是式(8.53)中定义的足球确定的母函数时, 求超母函数 $\sum_{n \geq 0} F_n(z) w^n$ 的闭形式。

17. 设 $X_{n,p}$ 和 $Y_{n,p}$ 分别有二项分布和负二项分布(具有参数 (n, p))。(在式(8.57)和(8.60)中定义这些分布。)证明 $\Pr(Y_{n,p} \leq m) = \Pr(X_{m+1,1-p} \geq n)$, 这含有什么二项系数的等式?

18. 如果对于所有 $k \geq 0$, $\Pr(X = k) = e^{-\mu} \mu^k / k!$, 则称随机变量 X 为具有平均值 μ 的 Poisson 分布。

(a) 这样一个随机变量的概率母函数是什么?

(b) 它的平均值, 方差和其他累积量是什么?

19. 继续前面的习题, 设 X_1 是具有平均值 μ_1 的随机 Poisson 变量, X_2 是具有平均值 μ_2 的随机 Poisson 变量, 与 X_1 独立。

(a) $X_1 + X_2 = n$ 的概率是多少?

(b) $2X_1 + 3X_2$ 的平均值, 方差和其他累积量是多少?

20. 证明等待给定正面和反面的型式所需次数的平均值和方差的一般公式(8.74)和(8.75)。

21. 如果在式(8.77)中都置 H 和 T 等于 $1/2$, Δ 的值代表什么?

22. 证明条件期望和条件方差定律(8.105)。

课外习题

23. 设 Pr_{00} 是两个完好骰子的概率分布, Pr_{11} 是式(8.2)中给出的两个灌铅骰子的概率分布。求所有事件 A 使得 $\text{Pr}_{00}(A) = \text{Pr}_{11}(A)$ 。这些事件中的哪些仅依赖于随机变量 S ? ($\Omega = D_2$ 的概率空间有 2^{36} 个事件, 仅仅这些事件的 2^{11} 个事件只依赖于 S 。)

24. 玩者 J 滚动 $2n+1$ 个完好的骰子, 且把出现四的那些骰子移去, 然后玩者 K 说出 1 和 6 之间的一个数, 滚动剩下的骰子, 且把呈现说出数的那些骰子移去。移去总的骰子

最多($n+1$ 或更多)的玩家是胜者。

(a) J 移去的总骰子数的平均值和方差是什么?

提示: 骰子是独立的。

(b) 当 $n=2$ 时, J 胜的概率是多少?

25. 考虑一种赌博游戏, 在游戏中你下给定数量 A 的赌注, 并滚动一个完好的骰子。如果出现 k 点, 你用 $2(k-1)/5$ 乘你的赌注。(特别每当你滚出 4 时你把赌注加倍, 但是如果你滚出 1 则你输尽。)任何时候你都能停止和收回现时的赌注。 n 次滚动之后你的赌注的平均值和方差是多少? (忽略任何四舍五入到货币的整数量的结果。)

26. 求在 n 个元素的一个随机置换中 l 个轮换的个数的平均值和方差。(在式(8.23), (8.24)和(8.53)中讨论的足球胜利问题是 $l=1$ 的特殊情形。)

27. 设 X_1, X_2, \dots, X_n 是随机变量 X 的独立样本。方程(8.19)和(8.20)说明在这些观察的基础上如何来估计 X 的平均值和方差, 对于第三个累积量 κ_3 给出一个估计。(在 κ_3 的期望值为 κ_3 的意义下, 你的公式应是一个“无偏”估计。)

28. (a) 假设 Alice 胜, 硬币抛掷游戏(8.78)的平均长度是多少?

(b) 假设 Bill 胜, 硬币抛掷游戏(8.78)的平均长度是多少?

29. Alice, Bill 和计算机抛掷一个完好的硬币, 直到第一次出现各自的型式 $A=HHTH$, $B=HTHH$, 或 $C=THHH$ 之一。(如果仅涉及这些型式的两个, 依据式(8.82)我们知道 A 可能胜 B , B 可能胜 C , C 可能胜 A ; 但是所有三种型式同时在游戏中。)问每个玩者胜的机会是多少?

30. 书中考虑了散列表中成功查找的三类方差, 实际上还有两种: 我们能考虑 $P(h_1, \dots, h_n; k)$ 的(在 h_1, \dots, h_n 上的)方差的平均(对 k 平均); 同时我们能考虑平均(在 h, \dots, h_n 上的)方差(在 k 上)。计算这些量。

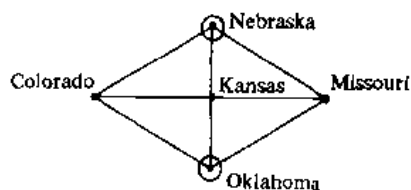
31. 在 5 边形 $ABCDE$ 的顶点 A 处放一个苹果, 离苹果两个顶点远的 C 处有一个蛙虫。蛙虫每天以相同概率跳到两个相邻顶点之一。因此一天之后蛙虫以概率 $1/2$ 处于顶点 B , 以概率 $1/2$ 处于顶点 D , 两天后, 蛙虫可能再回到 C , 因为它不记忆前面的位置。当它达到顶点 A 时停下来吃苹果。

(a) 直到吃苹果, 天数的平均值和方差是什么?

(b) 设 p 是天数为 100 或多于 100 的概率, Chebyshev 不等式关于 p 表明了什么?

(c) (习题 12 的)末尾不等式关于 p 告诉了我们什么?

32. Alice 和 Bill 所在部队驻扎在 5 个州 Kansas, Nebraska, Missouri, Oklahoma 或 Colorado 中的一个州。开始 Alice 在 Nebraska, Bill 在 Oklahoma。每个月每个人被调派到一个相邻的州, 每个相邻州是等可能的。(下列是相邻的图: 开始的州加圈。)例如, 第一个月之后 Alice 被调派到 Colorado, Kansas 或 Missouri, 调派到每个州的概率为 $1/3$ 。求使 Alice 和 Bill 彼此相遇的月数的平均值和方差。(你可望利用计算机的帮助。)



33. 式(8.88)中的随机变量 X_1 和 X_2 是独立的吗?

34. Gina 是一个高尔夫球员, 她每一击为(在高尔夫球的标准打数上获得一击)“超级一击”的概率为 $p = .05$, 为通常一击的概率为 $q = .91$, 为(关于标准打数花费她的一击)“次一击”的概率为 $r = .04$. (无高尔夫球员: 每一圈上她分别以概率 p , q 或 r 朝她的目标前进 2, 1 或 0 步. 在 m 个穴的标准打数上, 她的得分是最小的 n 使得在取得 n 圈后她前进 m 步或更多步. 低分比高分好.)

(a) 证明当 Gina 对击标准打数的对手比赛时, Gina 多半胜 4 穴标准打数. (换句话说, 她的得分小于 4 的概率比她的得分大于 4 的概率大.)

(b) 证明她在 4 穴标准打数上的平均得分大于 4. (所以在总的得分上, 她对于一个“沉着”的对手倾向于输, 虽然在按穴的对手比赛中她倾向于胜.)

考查性问题

35. 灌铅的一个骰子有概率分布

$$\Pr(\square) = p_1; \Pr(\blacksquare) = p_2; \dots; \Pr(\blacksquare) = p_6.$$

设 S_n 是滚动此骰子 n 次后点数的和. 求“灌铅分布”上一个必要和充分条件使得对于所有 n , 两个随机变量 $S_n \bmod 2$ 和 $S_n \bmod 3$ 彼此独立.

36. 某个骰子的 6 个面包含点的型式



而不是通常的 \square 到 \blacksquare .

(a) 证明有一种方式把点数赋予另一个骰子的 6 个面, 以致当掷这样两个骰子时, 点数的和具有像两个通常骰子上点数和的相同概率分布. (假设所有 36 面的配对是等可能的.)

(b) 推广, 求点数赋予 n 个骰子的 $6n$ 个面的所有方式以致点数和的分布和 n 个通常骰子上点数的分布相同. (每个面接受一个正整数点数.)

37. 设 p_n 是在一行中看到正面两次之前恰好需要抛一个完好的硬币 n 次的概率, 且设 $q_n = \sum_{k \geq n} p_k$. 求 p_n 和 q_n 依据 Fibonacci 数的闭形式.

38. 滚动一个完好的骰子直到出现所有 6 个面所需的次数的概率母函数是什么? 推广到 m 个面的完好骰子: 给出看到 m 个面的 l 个面所需滚动次数的平均值和方差的闭形式. 此数恰好为 n 的概率是多少?

39. Dirichlet 概率母函数有形式

$$P(z) = \sum_{n \geq 1} \frac{p_n}{z^n}.$$

因此 $P(0)=1$. 若 X 是具有 $\Pr(X=n)=p_n$ 的随机变量, 依据 $P(z)$ 和它的导数表达 $E(X)$, $V(X)$ 和 $E(\ln X)$.

40. 二项分布 (8.57) 的第 m 个累积量 κ_m 有形式 $nf_m(p)$, 其中 f_m 是 m 次多项式. (例如, $f_1(p)=p$, $f_2(p)=p-p^2$, 因为平均值和方差是 np 和 npq .)

(a) 求 $f_m(p)$ 中 p^k 的系数的闭形式.

(b) 证明 $f_m\left(\frac{1}{2}\right) = (2^m - 1)B_m / m + [m=1]$, 其中 B_m 是第 m 个 Bernoulli 数.

41. 设随机变量 X_n 是一个完好骰子直到正面总共出现 n 次的抛掷次数. 证明 $E(X_{n+1}^{-1}) = (-1)^n (\ln 2 + H_{\lfloor n/2 \rfloor} - H_n)$. 用第九章的方法以绝对误差 $O(n^{-3})$ 来估计此值.

42. 某人有找工作的问题. 如果在任何给定的早晨他未受雇用, 则有常数概率 p_n (独立于过去的历史) 他在该天晚上之前被雇用; 但是如果当这天开始时他取得一个工作, 则有常数概率 p_f 他将在黄昏被解雇. 求他均匀地有一个工作的晚上的平均个数, 假设开始他被雇用且此过程持续 n 天. (例如, 若 $n=1$, 解答为 $1-p_f$.)

43. 求概率母函数 $G_n(z) = \sum_{k \geq 0} p_{k,n} z^k$, 其中 $p_{k,n}$ 是 n 个元素的随机置换恰好有 k 个轮换的概率. 轮换数的平均值和标准差是什么?

44. 体育系举行 2^n 个网球队员的校内“分离比赛”如下: 在第一轮中, 每对等可能地把球员随机分成一对一对, 且举行 2^{n-1} 次比赛. 胜者进入第二轮, 其中相同处理产生 2^{n-2} 个胜者, 等等. 第 k 轮有仍未击败的 2^{n-k+1} 个球员中间随机选取的 2^{n-k} 次比赛. 第 n 轮产生冠军. 在比赛组织者不知情的情况下, 在球员间实际有一个次序, 以致 x_1 是最好的, x_2 是第二名, \dots , x_{2^n} 最差. 当 x_j 对 x_k 且 $j < k$ 时, 胜者为 x_j 的概率为 p , 胜者为 x_k 的概率为 $1-p$, 独立于其他比赛. 我们假设相同概率 p 应用到所有 j 和 k .

(a) x_1 比赛获胜的概率是多少?

(b) 第 n 轮(最后的比赛)是在居首位的两个球员 x_1 和 x_2 之间进行的概率是多少?

(c) 最好的 2^k 个球员是第 k 轮到最后轮中的竞争者的概率是多少? (前面的两个问题是 $k=0$ 和 $k=1$ 的情形.)

(d) 设 $N(n)$ 是本质上不同的比赛结果的个数; 如果在相同的球员间举行比赛且有相同的胜者, 则两次比赛本质上是相同的. 证明 $N(n) = 2^{n!}$.

(e) x_2 比赛获胜的概率是多少?

(f) 证明如果 $1/2 < p < 1$, 则对于 $1 \leq j < 2^n$, x_j 胜的概率严格大于 x_{j+1} 胜的概率.

45. 西班牙的真葡萄酒是按照称为“Solera”的多阶段方式制造的. 为了简单起见, 我们假设制酒者仅有 3 个桶, 称为 A , B 和 C . 每年桶 C 中的三分之一的酒装瓶且用 B 中的酒来替换; 然后以 A 的三分之一的酒来补充 B ; 最后用新酒补充 A . 设 $A(z)$, $B(z)$, $C(z)$ 是概率母函数, 其中 z^n 的系数是刚做转换之后的对应桶中 n 年老酒的部分.

(a) 假设自古以来操作继续下去,以致我们有一种稳定状态,此时 $A(z)$, $B(z)$ 和 $C(z)$ 在每年初是相同的,求这些母函数的闭形式。

(b) 在相同的假设下,求每桶中酒的龄期的平均值和标准差。当葡萄酒被装瓶时,它的平均龄期是多少?它的多少恰好是 25 年老?

(c) 现在考虑有限时间:假设在 0 年开始处所有 3 个桶包含新酒,在 n 年开始处装瓶的葡萄酒的平均龄期是多少?

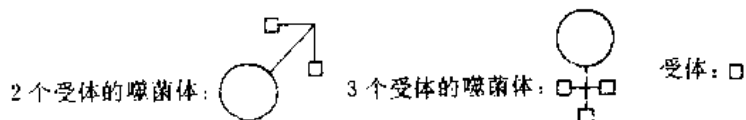
46. Stefan Banach 常带两匣火柴,开始每匣包含 n 根,每当他需要火柴时以概率 $1/2$ 随机选取一匣,独立于前面的选取。在取出 1 根火柴后,把匣放回口袋(即使匣变成空——所有著名数学家常这样做)。当他所选的匣是空时,他将扔掉它且去取另一匣。

(a) 一旦他发现另一匣也是空的,这种情况出现的概率是多少?(对 $n=1$,它一半时间出现,对 $n=2$,它 $3/8$ 时间出现。)为了回答这一部分,求母函数 $P(w, z) = \sum_{m,n} p_{m,n} w^m z^n$ 的闭形式,其中 $p_{m,n}$ 是开始一匣中有 m 根火柴且另一匣有 n 根,当首先选取一个空匣时两匣都空的概率。然后求 $p_{n,n}$ 的闭形式。

(b) 推广部分(a)的解答。求当首先扔掉一个空匣时另一匣中恰有 k 根火柴的概率的闭形式。

(c) 求另一匣中火柴的平均数的闭形式。

47. 一些医生和一些物理学家合作,最近发现一对以特有方式繁殖的微生物。雄性微生物称为 2 个受体的噬菌体,在它的表面上有 2 个受体;雌性微生物称为 3 个受体的噬菌体,有 3 个受体:

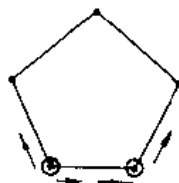


当用 1 个 ψ 粒子扩散来培养 2 个受体的噬菌体和 3 个受体的噬菌体时,恰好噬菌体之一上的 1 个受体吸收粒子;每个受体是等可能的。如果它是 2 个受体的噬菌体的受体,则 2 个受体的噬菌体变成 3 个受体的噬菌体;如果它是 3 个受体的噬菌体的受体,则 3 个受体的噬菌体分裂成两个 2 个受体的噬菌体。因此如果一个试验以具有 2 个受体的噬菌体开始,第 1 个 ψ 粒子把它改变成 3 个受体的噬菌体,第 2 个粒子把 3 个受体的噬菌体分裂为两个 2 个受体的噬菌体,第 3 个粒子把 2 个受体的噬菌体之一改变成一个 3 个受体的噬菌体。第 4 个粒子到达 2 个受体的噬菌体或 3 个受体的噬菌体;则或者有两个 3 个受体的噬菌体(概率 $2/5$)或者三个 2 个受体的噬菌体(概率 $3/5$)。如果从一个 2 个受体的噬菌体开始,且以扩散单个 ψ 粒子来培养 n 次,求出现的 2 个受体的噬菌体的平均个数的闭形式。

48. 5 个人位于 5 边形的顶点,彼此抛掷飞盘。他们有两个飞盘,开始位于图示的相邻顶点。在每个时间间隔内,每个飞盘以相等概率(沿着 5 边形的一条边)被抛掷到左面或右面。此过程继续下去直到一个人同时为两个飞盘的目标,则游戏停止。(所有抛掷独立于过去的历史。)

(a) 求抛掷对数的平均值和方差。

(b) 依据 Fibonacci 数, 求游戏持续多于 100 步的概率的闭形式.



49. 雪地散步者 Luke 在他的山岳小屋渡寒假. 前走廊有 m 双长统靴, 后走廊有 n 双. 每次他去散步, 抛掷一个(完好的)硬币来决定从前走廊离开还是从后走廊离开, 且他穿上这个走廊的一双长统靴出发. 他返回每个走廊(独立于他的出发地点)且在返回的走廊处留下长统靴有 50/50 可能性. 因此在一次散步之后在前走廊将有 $m+[-1, 0 \text{ 或 } +1]$ 双, 在后走廊将有 $n+[-1, 0 \text{ 或 } +1]$ 双. 如果所有长统靴堆积在一个走廊且他决定从另一个走廊离开, 他无长统靴离开且冻伤, 终止假期. 假设他把散步继续下去直到终止, 设 $P_N(m, n)$ 是完成恰好 N 次无冻伤的往返(开始前走廊 m 双后走廊 n 双)的概率, 如果 m 和 n 都是正的, 则

$$P_N(m, n) = \frac{1}{4} P_{N-1}(m-1, n+1) + \frac{1}{2} P_{N-1}(m, n) + \frac{1}{4} P_{N-1}(m+1, n-1).$$

这是因为这个第一次往返或者是前/后, 前/前, 后/后或者后/前, 每个具有概率 $1/4$, 且剩下 $N-1$ 次往返.

(a) 通过求当 $m=0$ 或 $n=0$ 成立的公式来完成 $P_N(m, n)$ 的递归. 用递归来得到概率母函数中间成立的方程

$$g_{m,n}(z) = \sum_{N \geq 0} P_N(m, n) z^N.$$

(b) 微分你的方程, 且置 $z=1$. 从而得到量 $g'_{m,n}(1)$ 间的关系. 解这些方程, 从而决定在冻伤之前往返的平均次数.

(c) 证明如果我们代入 $z = 1/\cos^2 \theta$, $g_{m,n}$ 有一种闭形式:

$$g_{m,n}\left(\frac{1}{\cos^2 \theta}\right) = \frac{\sin(2m+1)\theta + \sin(2n+1)\theta}{\sin(2m+2n+2)\theta} \cos \theta.$$

50. 考虑函数

$$H(z) = 1 + \frac{1-z}{2z} (z-3 + \sqrt{(1-z)(9-z)}) .$$

此问题的目的是证明 $H(z) = \sum_{k \geq 0} h_k z^k$ 是概率母函数, 且获得关于它的一些基本事实.

(a) 设 $(1-z)^{3/2} (9-z)^{1/2} = \sum_{k \geq 0} c_k z^k$. 证明 $c_0 = 3$, $c_1 = -14/3$, $c_2 = 37/27$,

对于所有 $l \geq 0$, $c_{3+l} = 3 \sum_k \binom{l}{k} \binom{1/2}{3+k} \left(\frac{8}{9}\right)^{k+1}$. 提示: 用等式

$$(9-z)^{1/2} = 3(1-z)^{1/2} \left(1 + \frac{8}{9}z/(1-z)\right)^{1/2}$$

且把最后的因子展成 $z/(1-z)$ 的幂。

(b) 用(a)和习题 5.81 证明 $H(z)$ 的系数全为正。

(c) 证明令人惊异的等式

$$\sqrt{\frac{9-H(z)}{1-H(z)}} = \sqrt{\frac{9-z}{1-z}} + 2.$$

(d) H 的平均值和方差是什么?

51. (假想中的南美洲的)黄金国中的抽彩给奖法用前面问题中定义的支付分布 H 。每张彩票花费 1 个已废的西班牙金币, 且以概率 h_k 支付 k 个金币。你的每张彩票胜的可能性完全独立于其他彩票的可能性; 换句话说, 一张彩票的胜或负不影响在相同彩票抽奖法中购买的任何其他彩票的胜的概率。

(a) 假设你从 1 个金币开始, 玩此游戏。如果你赢 k 个金币, 在第 2 次游戏中你买 k 张彩票; 然后你在第 2 次游戏中取得全部赢得的钱且把它们用到第 3 次游戏; 等等。如果你的彩票没有一张是赢者, 你破了产且你必须停止赌博。证明在这样玩 n 轮之后你现时的钱的概率母函数是

$$1 - \frac{4}{\sqrt{(9-x)/(1-z)} + 2n - 1} + \frac{4}{\sqrt{(9-x)/(1-z)} - 2n + 1}.$$

(b) 设 g_n 是第 n 次游戏首次把你所有的钱输了的概率, 且设 $G(z) = g_1 z + g_2 z^2 + \dots$. 证明 $G(1) = 1$. (这意味着以概率 1 你迟早必然要输, 虽然你可能在此期间玩得很开心。) G 的平均值和方差是什么?

(c) 如果你持续地玩直到破产, 你买的彩票的平均总数是多少?

(d) 如果你以 2 个金币开始, 而不是 1 个金币, 直到你输光的游戏的平均数是多少?

额外问题

52. 证明当概率空间有限时, 书中随机变量中位数和众数的定义在某种含意的意义下对应于序列的中位数和从数的定义。

53. 证明或推翻: 如果 X , Y 和 Z 是具有性质(所有 3 个对 (X, Y) , (X, Z) 和 (Y, Z) 是独立的)的随机变量, 则 $X+Y$ 独立于 Z 。

54. 方程(8.20)证明了 $\hat{V}X$ 的平均值是 VX . $\hat{V}X$ 的方差是什么?

55. 通常玩的纸牌包含 52 张牌, 4 种花色的每 1 张具有集合 $\{A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K\}$ 的面值, 设 X 和 Y 表示最顶上和最底下牌的各自的面值, 且考虑下列洗牌的算法:

S1 随机置换纸牌以致以概率 $1/52!$ 出现每种排列。

S2 如果 $X \neq Y$, 抛掷 1 个偏的硬币, 出现正面的概率为 p , 如果出现正面返回到 S1. 否则停止.

假设每个硬币抛掷和每个置换独立于所有其他的随机策略. 在此过程停止之后使 X 和 Y 为独立随机变量的 p 值是多少?

56. 把习题 48 的飞盘问题从 5 边形推广到 m 边形. 当飞盘开始位于相邻的顶点时, 一般无碰撞抛掷次数的平均值和方差是多少? 证明, 如果 m 是奇数, 抛掷次数的概率母函数能写成硬币抛掷分布的乘积:

$$G_m(z) = \prod_{k=1}^{(m-1)/2} \frac{p_k z}{1 - q_k z},$$

$$\text{其中 } p_k = \sin^2 \frac{(2k-1)\pi}{2m}, \quad q_k = \cos^2 \frac{(2k-1)\pi}{2m}$$

提示: 试代入 $z = 1 / \cos^2 \theta$.

57. 证明如果 $l \geq 3$, 当抛掷一个完好硬币时, Penney 赌注型式 $\tau_1, \tau_2, \dots, \tau_{l-1}, \tau_l$ 总比型式 $\bar{\tau}_1, \bar{\tau}_2, \dots, \bar{\tau}_{l-1}$ 差.

58. 是否有正面和反面的型式 A 和 B 使得 A 比 B 长, 然而当抛掷 1 个完好硬币时, 多于一半时间 A 出现在 B 的前面?

59. 设 k 和 n 是指定正整数 ($k < n$).

(a) 求找插入具有 m 个表的散列表的第 k 项和第 n 项所需探索数的联合分布的概率母函数.

$$G(w, z) = \frac{1}{m^n} \sum_{h_1=1}^m \cdots \sum_{h_n=1}^m w^{P(h_1, \dots, h_n; k)} z^{P(h_1, \dots, h_n; n)}$$

的闭形式.

(d) 虽然随机变量 $P(h_1, \dots, h_n; k)$ 和 $P(h_1, \dots, h_n; n)$ 是不独立的, 证明它们多少有点独立:

$$\begin{aligned} E(P(h_1, \dots, h_n; k)P(h_1, \dots, h_n; n)) \\ = (EP(h_1, \dots, h_n; k))(EP(h_1, \dots, h_n; n)). \end{aligned}$$

60. 用前一习题的结果来证明(8.103).

61. 按习题 47, 求 n 次扩散后 2 个受体的噬菌体数的方差.

研究性问题

62. 正态分布是除了平均值和方差外所有累积量为零的非离散概率分布. 是否有一种容易的方法告知是否累积量的给定序列 $\langle \kappa_1, \kappa_2, \kappa_3, \dots \rangle$ 来自离散分布? (在一个离散分布中所有概率一定是“极微的”。)

63. 在 Penney 赌注游戏中是否有 $l \geq 3$ 正面和反面的任何序列 $A = \tau_1 \tau_2 \cdots \tau_{l-1} \tau_l$ 使得序列 $H\tau_1 \tau_2 \cdots \tau_{l-1}$ 和 $\bar{H}\tau_1 \tau_2 \cdots \tau_{l-1}$ 相对于 A 都同样地执行良好?

第九章 渐 近

当我们能找到准确解答时，固然是好极了，因为对于完整的理解有了一些令人满意的结果。但有时近似解的情况也是良好的。如果我们遇到一个和或者一个递归，它的解没有闭形式(就我们能告知的)，我们仍希望知道关于解答的事情；我们不必坚持要知道全部或者什么也不知道。即使的确有一个闭形式，我们的认识可能是不完全的，因为我们可能不知道如何把它和其他闭形式比较。

例如，和

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

(明显地)没有闭形式。但是恰好知道当 $n \rightarrow \infty$ 时

$$S_n \sim 2 \binom{3n}{n}.$$

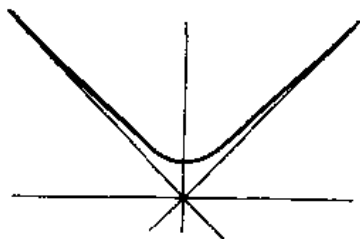
我们说和是“渐近到” $2 \binom{3n}{n}$ 。有较详细的信息更好，像

$$S_n = \binom{3n}{n} \left(2 - \frac{4}{n} + O\left(\frac{1}{n^2}\right) \right), \quad (9.1)$$

它给出一个“阶 $1/n^2$ 的相对误差”。但是就是这个公式也不足以告诉我们，与其他量比较 S_n 为多大、 S_n 或 Fibonacci 数 F_{4n} ，哪一个较大？回答：我们有 $S_2 = 22 > F_8 = 21$ (当 $n=2$ 时)；但是最后 F_{4n} 较大，因为 $F_{4n} \sim \varphi^{4n} / \sqrt{5}$ 和 $\varphi^4 \approx 6.8541$ ，而

$$S_n = \sqrt{\frac{3}{\pi n}} (6.75)^n \left(1 - \frac{151}{72n} + O\left(\frac{1}{n^2}\right) \right). \quad (9.2)$$

本章的目标是学习如何不费太大功夫来熟悉和推出像以上的结果。渐近(asymptotic)这个词起源于希腊语，意思是“不落在一起”。当古希腊的数学家研究圆锥曲线时，他们考虑了像 $y = \sqrt{1+x^2}$ 的图的抛物线。



它有直线 $y=x$ 和 $y=-x$ 作为“渐近线”。当 $x \rightarrow \infty$ 时, 曲线向渐近线靠近, 但是不完全达到这些渐近线。现在我们在较宽的意义下用“渐近”来表示当某个参数靠近一个极限值时任何近似值越来越接近真值。对于我们来说, 渐近意味着“几乎落在一起。”

有些渐近公式推导很困难, 超出了本书的范围。我们将满足于这个主题的一个引言, 我们希望获得一个适当的基础, 在此基础上能建立起进一步的技巧。我们特别对于懂得 ‘ \sim ’ 和 ‘ O ’ 以及相似符号的定义感兴趣, 且我们将研究处理渐近量的一些基本方法。

9.1 级别

通常实际中出现的 n 的函数有不同的“渐近增长比”, 一个比另一个靠近无穷大更快。我们通过公式

$$f(n) \prec g(n) \Leftrightarrow \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0 \quad (9.3)$$

来表示这一点。这个关系是传递的: 如果 $f(n) \prec g(n)$ 和 $g(n) \prec h(n)$, 则 $f(n) \prec h(n)$ 。如果 $f(n) \prec g(n)$, 我们也可以写为 $g(n) \succ f(n)$ 。这种记法是 Paul du Bois-Reymond^[29] 在 1871 年引入的。

例如, $n \prec n^2$; 我们形式地说 n 比 n^2 增长得更慢。事实上,

$$n^\alpha \prec n^\beta \Leftrightarrow \alpha < \beta, \quad (9.4)$$

其中 α 和 β 是任意实数。

当然, 有许多与 n 的幂无关的函数。我们能用 \prec 关系把函数的位置列成渐近等级的次序, 像下列包含的项目:

$$1 \prec \log \log n \prec \log n \prec n^\varepsilon \prec n^c \prec n^{\log n} \prec c^n \prec n^n \prec c^{c^n}.$$

(这里 ε 和 c 是任意常数, 且 $0 < \varepsilon < 1 < c$.)

除了 1 之外, 当 n 趋向于无穷大时, 这里列出的所有函数都趋向于无穷大。因此, 当在这个级别中试安置一个新函数时, 我们不去决定是否它变成无穷, 而要决定它变得多快。

当我们作渐近分析时, 它有助于培养一种扩张的状态: 当想象一个变量靠近无穷大时, 应认为大。例如, 级别说出 $\log n \prec n^{0.0001}$; 如果我们限制像 $n = 10^{100}$ 那样的数, 这

似乎是错误的。此时, $\log n = 100$, 而 $n^{0.0001}$ 仅为 $10^{0.01} \approx 1.0233$ 。但是如果增长到 $n = 10^{100}$, 则 $\log n = 10^{100}$ 比 $n^{0.0001} = 10^{10^96}$ 小多了。

即使 ε 极小 (譬如说, 小于 $1/10^{100}$), 如果 n 足够大, $\log n$ 的值将比 n^ε 的值小许多。如果置 $n = 10^{10^{2k}}$, 其中 k 是使得 $\varepsilon \geq 10^{-k}$, 我们有 $\log n = 10^{2k}$, 而 $n^\varepsilon \geq 10^{10^k}$ 。所以当 $n \rightarrow \infty$ 时, 比 $(\log n)/n^\varepsilon$ 趋于零。

上面表明的级别是涉及趋向于无穷大的函数。然而, 我们常常对趋向于零的函数感兴趣, 所以那样的函数的相似级别是有用的。我们通过取倒数得到一个, 因为当 $f(n)$ 和 $g(n)$ 不为零时, 我们得到

$$f(n) < g(n) \Leftrightarrow \frac{1}{g(n)} < \frac{1}{f(n)}. \quad (9.5)$$

因此, 例如, 下列函数(除 1 之外)都趋向于零:

$$\frac{1}{c^{\frac{1}{\varepsilon n}}} < \frac{1}{n^n} < \frac{1}{c^n} < \frac{1}{n^{\log n}} < \frac{1}{n^\varepsilon} < \frac{1}{n^{\frac{1}{\varepsilon}}} < \frac{1}{\log n} < \frac{1}{\log \log n} < 1.$$

让我们查看另外几个函数, 看它们该安置在何处。我们知道, 小于或等于 n 的素数个数 $\pi(n)$ 近似于 $n/\ln n$ 。由于 $1/n^\varepsilon < 1/\ln n < 1$, 用 n 相乘得到

$$n^{1-\varepsilon} < \pi(n) < n.$$

事实上, 例如通过注意

$$n^{\alpha_1} (\log n)^{\alpha_2} (\log \log n)^{\alpha_3} < n^{\beta_1} (\log n)^{\beta_2} (\log \log n)^{\beta_3} \Leftrightarrow (\alpha_1, \alpha_2, \alpha_3) < (\beta_1, \beta_2, \beta_3) \quad (9.6)$$

我们能推广 (9.4)。这里 $(\alpha_1, \alpha_2, \alpha_3) < (\beta_1, \beta_2, \beta_3)$ 意味着字典次序; 换句话说, $\alpha_1 < \beta_1$, 或者 $\alpha_1 = \beta_1$ 和 $\alpha_2 < \beta_2$, 或者 $\alpha_1 = \beta_1$ 和 $\alpha_2 = \beta_2$ 和 $\alpha_3 < \beta_3$ 。

关于函数 $e^{\sqrt{\log n}}$ 怎么样, 在级别中它处于何处? 我们能通过用

$$e^{f(n)} < e^{g(n)} \Leftrightarrow \lim_{n \rightarrow \infty} (f(n) - g(n)) = -\infty \quad (9.7)$$

来回答问题, 依据定义 (9.3) 取对数的两步中产生此结果。因而

$$1 < f(n) < g(n) \Rightarrow e^{f(n)} < e^{g(n)}.$$

且由于 $1 < \log \log n < \sqrt{\log n} < c \log n$, 我们有 $\log n < e^{\sqrt{\log n}} < n^\varepsilon$ 。

当两个函数 $f(n)$ 和 $g(n)$ 有相同增长率时, 我们记 ' $f(n) \asymp g(n)$ '。正式定义是: 对于某个 C 和所有充分大的 n ,

$$f(n) \asymp g(n) \Leftrightarrow |f(n)| \leq C|g(n)| \quad \text{和} \quad |g(n)| \leq C|f(n)|. \quad (9.8)$$

例如, 如果 $f(n)$ 是常数, $g(n) = \cos n + \arctan n$, 上式成立。每当 $f(n)$ 和 $g(n)$ 是相同次数的

多项式时, 我们将简单地证明上式成立. 还有一个较强的关系, 由规则

$$f(n) \sim g(n) \Leftrightarrow \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1 \quad (9.9)$$

来定义. 此时我们称“ $f(n)$ 渐近于 $g(n)$ ”.

G.H.Hardy^[148]引入一类称为对数-指数函数的有趣和重要的概念, 这类函数递归地定义为满足下列性质的最小函数族 \mathcal{L} :

- 对于所有实数 α , 常数函数 $f(n) = \alpha$ 在 \mathcal{L} 中.
- 恒等函数 $f(n) = n$ 是在 \mathcal{L} 中.
- 如果 $f(n)$ 和 $g(n)$ 在 \mathcal{L} 中, 则 $f(n) - g(n)$ 也在 \mathcal{L} 中.
- 如果 $f(n)$ 在 \mathcal{L} 中, 则 $e^{f(n)}$ 也在 \mathcal{L} 中.
- 如果 $f(n)$ 在 \mathcal{L} 中, 且“最终为正”, 则 $\ln f(n)$ 在 \mathcal{L} 中.

如果有一个整数 n_0 使得每当 $n \geq n_0$ 时 $f(n) > 0$, 称函数为“最终为正”.

例如, 我们能用这些规则来表明每当 $f(n)$ 和 $g(n)$ 在 \mathcal{L} 中时, $f(n) + g(n)$ 在 \mathcal{L} 中, 因为 $f(n) + g(n) = f(n) - (0 - g(n))$. 如果 $f(n)$ 和 $g(n)$ 是 \mathcal{L} 的最终为正的成员, 它们的乘积 $f(n)g(n) = e^{\ln f(n) + \ln g(n)}$ 和商 $f(n)/g(n) = e^{\ln f(n) - \ln g(n)}$ 在 \mathcal{L} 中; 所以像 $\sqrt{f(n)} = e^{\frac{1}{2}\ln f(n)}$ 的函数在 \mathcal{L} 中, 等等. Hardy 证明了每一个对数-指数函数是最终为正, 最终为负, 或者恒为零. 所以除了不能用恒为零的函数除之外, 任何两个 \mathcal{L} 函数的乘积和商在 \mathcal{L} 中.

关于对数-指数函数的 Hardy 的主要定理是形成一个渐近组别: 如果 $f(n)$ 和 $g(n)$ 是 \mathcal{L} 中的任何函数, 则 $f(n) < g(n)$, 或者 $f(n) \sim g(n)$, 或者 $f(n) \asymp g(n)$. 在最后的情形, 事实上有一个常数 α 使得

$$f(n) \sim \alpha g(n).$$

Hardy 定理的证明超出了本书的范围; 但是知道定理存在是好的, 因为几乎我们任何时候需要讨论的每一个函数都在 \mathcal{L} 中. 实际上, 我们一般能不太困难地为一个给定的函数提供一个给定的级别.

9.2 O 表示法

渐近分析的一种极好的表示惯例是由 Paul Bachmann 1894 引入的, 后来由 Edmund Landau 和其他人普及. 在像

$$H_n = \ln n + \gamma + O(1/n) \quad (9.10)$$

的公式中已经看到了它. 公式告诉我们第 n 个调和数等于 n 的自然对数的欧位常数, 加上一个“ n 分之1的大 O ”的量. 最后的这个量不精确地确定; 但是不管它是什么, 表示法断定它的绝对值不大于一个常数乘 $1/n$.

O 表示法好在它去掉了不重要的细节, 让我们去钻研特征: 如果 $1/n$ 的常数倍不重

要, 量 $O(1/n)$ 是可忽略的小。

此外, 我们开始在公式的中间正确地用 O 。如果要依据 9.1 节中的表示法来表达式 (9.10), 一定要把 ' $\ln n + \gamma$ ' 转换到左边, 且确定像

$$H_n - \ln n - \gamma < \frac{\log \log n}{n}$$

的一个较弱结果, 或像

$$H_n - \ln n - \gamma \asymp \frac{1}{n}$$

的一个较强结果。大 O 表示法允许我们可以不转换而确定相称细节的一个恰当的量。

如果考虑一些附加的例子, 则能使非确切的指定量的思想更清楚。有时用表示法 ' ± 1 ' 来代替 -1 或 $+1$, 我们不知道(或者我们也许不关心)它是哪一个, 然而我们能在公式中处理它。

N.G.de Bruijn 在他的渐近方法一书中考虑了一个大 L 来帮助我们理解大 O 。如果用 $L(5)$ 记绝对值小于 5 的一个数(但是不说数是多少), 则我们能执行一些计算而不知道完全的真值。例如, 我们能推导公式, 譬如 $1+L(5)=L(6)$; $L(2)+L(3)=L(5)$; $L(2)L(3)=L(6)$; $e^{L(5)}=L(e^5)$; 等等。但是我们不能推得 $L(5)-L(3)=L(2)$, 因为左边可能是 $4-0$ 。事实上, 我们最多能说 $L(5)-L(3)=L(8)$ 。

Bachmann 的 O 表示法相似于 L 表示法, 但是它比较不精确: $O(\alpha)$ 代表绝对值至多为一个常数乘 $|\alpha|$ 的一个数。我们不说数是多少, 且我们也不说常数是多少。当然如果在描述中没有变量, “常数”的概念是没有意义的, 所以当至少一个值为变化的量(譬如说 n) 时, 我们在上下文中仅用 O 表示法。于是在本书中公式

$$f(n) = O(g(n)) \quad \text{对所有 } n \quad (9.11)$$

意味着有一个常数 C 使得

$$|f(n)| \leq C|g(n)| \quad \text{对所有 } n; \quad (9.12)$$

且当 $O(g(n))$ 位于一个公式中间时, 它代表满足式 (9.12) 的一个函数 $f(n)$ 。 $f(n)$ 的值是未知的, 但是我们知道它们不太大。de Bruijn 的 ' $L(n)$ ' 相似地代表一个不确定的函数 $f(n)$, 它的值满足 $|f(n)| < |n|$ 。 L 和 O 之间的主要不同是涉及一个不确定的常数 C 的 O 表示法; 每次 O 的出现可能涉及一个不同的 C , 但是每个 C 与 n 无关。

例如, 我们知道前 n 个平方的和是

$$\square_n = \frac{1}{3}n(n+\frac{1}{2})(n+1) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n.$$

我们能记

$$\square_n = O(n^3)$$

因为对于所有整数 n , $\left|\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n\right| \leq \frac{1}{3}|n|^3 + \frac{1}{2}|n|^2 + \frac{1}{6}|n| \leq \frac{1}{3}|n^3| + \frac{1}{2}|n^3| + \frac{1}{6}|n^3| = |n^3|$. 我们相似地有更明确的公式

$$\square_n = \frac{1}{3}n^3 + O(n^2);$$

我们也能草率地把信息丢掉, 说

$$\square_n = O(n^{10}).$$

在 O 的定义中不要求给出尽可能好的界限.

但是等一等. 如果变量 n 不是整数, 将会怎样? 如果有一个像 $S(x) = (1/3)x^3 + (1/2)x^2 + (1/6)x$ 那样的公式, 其中 x 是实数, 将会怎样? 于是我们不能说 $S(x) = O(x^3)$, 因为当 $x \rightarrow 0$ 时, 比 $S(x)/x^3 = (1/3) + (1/2)x^{-1} + (1/6)x^{-2}$ 变成无界. 并且我们不能说 $S(x) = O(x)$, 因为当 $x \rightarrow \infty$ 时, 比 $S(x)/x = (1/3)x^2 + (1/2)x + (1/6)$ 变成无界. 所以显然不能对 $S(x)$ 用 O 表示法.

关于这种困境的解答是用 O 的那些变量一般服从附带的条件. 例如, 如果规定 $|x| \geq 1$, 或者 $|x| \geq \varepsilon$, 其中 ε 是任何正常数, 或者 x 是整数, 则我们能写 $S(x) = O(x^3)$. 如果我们规定 $|x| \leq 1$, 或者 $|x| \leq c$, 其中 c 是正常数, 则我们能写 $S(x) = O(x)$. 由涉及的变量的约束, 由它的环境产生 O 表示法.

这些约束常常由一个极限关系来指定. 例如, 我们可说

$$f(n) = O(g(n)) \quad \text{当 } n \rightarrow \infty. \quad (9.13)$$

这意味着当 n “接近” ∞ 时, O 表示法被认为成立; 我们不关心 n 不十分大时发生什么情况, 并且我们甚至不恰切地指出“接近”的意义; 此时 O 的每次出现不明显地要求存在两个常数 C 和 n_0 , 使得

$$|f(n)| \leq C|g(n)| \quad \text{每当 } n \geq n_0. \quad (9.14)$$

对于每个 O , C 和 n_0 的值可能不同, 但是它们不依赖于 n . 表示法

$$f(x) = O(g(x)) \quad \text{当 } x \rightarrow 0$$

相似地意味着存在两个常数 C 和 ε 使得

$$|f(x)| \leq C|g(x)| \quad \text{每当 } |x| \leq \varepsilon. \quad (9.15)$$

极限值不必是 ∞ 或 0 ; 我们能写

$$\ln z = z - 1 + O((z - 1)^2) \quad \text{当 } z \rightarrow 1,$$

因为当 $|z - 1| \leq \frac{1}{2}$ 时, 可证明 $|\ln z - z + 1| \leq |z - 1|^2$.

在上面逐步建立了 O 的定义, 从看来十分明显的内容发展到看来相当复杂的内容; 现在我们得到 O 代表一个未定函数以及依赖于环境的一个或者二个未指定的常数。对于任何合理的表示法, 这可能看来相当复杂了, 但它仍不是全部描述! 在基本的情况中埋伏着另外一个巧妙的考虑。即, 我们需要认识到记

$$\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n = O(n^3)$$

是好的, 但是我们从另一方面写出这个等式。否则会导出不合理的结论, 像从等式 $n = O(n^2)$ 和 $n^2 = O(n^2)$ 导出 $n = n^2$ 。当处理 O 表示法以及任何其他涉及不确切的指定量的公式时, 我们论述单向等式。一个方程的右边不比左边给出更多信息, 且它可能给出更少信息; 右边是左边的“未加工部分”。

根据严格的正式观点, 表示法 $O(g(n))$ 不代表单个函数 $f(n)$, 而代表所有函数 $f(n)$ 使得对于某个常数 C , $|f(n)| \leq C|g(n)|$ 的集合。一个不涉及 O 表示法的通常函数 $g(n)$ 代表包含单个函数 $f(n) = g(n)$ 的集合。如果 S 和 T 是 n 的函数的集合, 表示法 $S + T$ 代表形式 $f(n) + g(n)$ 的所有函数的集合, 其中 $f(n) \in S$ 和 $g(n) \in T$; 其他表示法, 像 $S - T$, ST , S/T , \sqrt{S} , e^S , $\ln S$ 也相似地定义。严格地说, 这样的函数集合之间的一个“方程”是集合包含; ‘=’ 记号实际意味着 ‘ \subseteq ’。这些正式的定义把所有 O 操作置于坚实逻辑的基础上。

例如, “方程”

$$\frac{1}{3}n^3 + O(n^2) = O(n^3)$$

意味着 $S_1 \subseteq S_2$, 其中 S_1 是形式 $(1/3)n^3 + f_1(n)$ 使得存在一个常数 C_1 , $|f_1(n)| \leq C_1|n^2|$ 的所有函数的集合, S_2 是所有函数 $f_2(n)$ 使得存在常数 C_2 , $|f_2(n)| \leq C_2|n^3|$ 的集合。我们可通过取左边的一个任意元素且证明它属于右边来形式地证明这个“方程”: 给定 $(1/3)n^3 + f_1(n)$ 使得 $|f_1(n)| \leq C_1|n^2|$, 我们一定要证明有一个常数 C_2 使得 $|(1/3)n^3 + f_1(n)| \leq C_2|n^3|$ 。由于对于所有整数 n , $n^2 \leq |n^3|$, 常数 $C_2 = 1/3 + C_1$ 达到预期的目的。

如果 ‘=’ 实际意味着 ‘ \subseteq ’。为什么我们不用 ‘ \subseteq ’ 而妄用等号? 有 4 个原因。

第一, 传统上的原因。数论学家从用 O 表示法的等号开始且习惯保留下来。至今所建立的内容, 我们不希望数学上的一致性有改变。

第二, 传统上的原因。用计算机的人很习惯于看到妄用的等号——多年来 FORTRAN 和 BASIC 的程序员写像 ‘ $N = N + 1$ ’ 那样的赋值语句。多一次妄用也不算多。

第三, 传统上的原因。我们常把 ‘=’ 读成字 ‘是’。例如, 我们把公式 $H_n = O(\log n)$ 说成 “ H 下标 n 是 $\log n$ 的大 O ”。而在英语中, 这个 ‘是’ 是单向的。我们说一个鸟是一个动物, 但是不能说一个动物是一个鸟; “动物” 是 “鸟” 的一种未加工形式。

第四, 对于我们的目的, 这样用是自然的。如果把使用 O 表示法限于它处于公式的整个右边的场合, 像调和近似 $H_n = O(\log n)$, 或像描述分类算法的运行时间

$T(n) = O(n \log n)$, 它不在乎我们用 '=' 或另一个记法。但是在一个表达式的中间用 O 表示法时, 像我们通常在渐近计算中所做的那样, 如果想象等号为一个等式, 且若想象 $O(1/n)$ 为十分小的量, 那么我们的直觉是很满意的。

所以我们将继续用 '=', 且继续将 $O(g(n))$ 认为是一个不完全确切的函数, 如果需要, 我们总能求助于集合论的定义来了解此函数。

但是当我们对定义挑剔时, 应再提醒一个专门事项: 如果在讨论中有几个变量, O 表示法形式地代表两个或两个以上变量的函数的集合, 而不仅是一个变量。每个函数的定义域是通常“自由”变化的每一个变量。

这个概念有点难捉摸, 因为一个变量可能仅在表达式的部分中被定义, 当由一个 \sum 或者相似的东西控制它。例如, 让我们仔细查看方程

$$\sum_{k=0}^n (k^2 + O(k)) = \frac{1}{3}n^3 + O(n^2), \text{ 整数 } n \geq 0. \quad (9.16)$$

左边的表达式 $k^2 + O(k)$ 代表形式 $k^2 + f(k, n)$ 使得对于 $0 \leq k \leq n$, $|f(k, n)| < Ck$ 的所有两个变量函数的集合存在一个常数 C 。对于 $0 \leq k \leq n$, 此函数集合的和是形式

$$\sum_{k=0}^n (k^2 + f(k, n)) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n + f(0, n) + f(1, n) + \cdots + f(n, n)$$

的所有函数 $g(n)$ 的集合, 其中 f 具有所说的性质。由于我们有

$$\begin{aligned} & \left| \frac{1}{2}n^2 + \frac{1}{6}n + f(0, n) + f(1, n) + \cdots + f(n, n) \right| \\ & \leq \frac{1}{2}n^2 + \frac{1}{6}n^2 + C \cdot 0 + C \cdot 1 + \cdots + C \cdot n \\ & < n^2 + C(n^2 + n) / 2 < (C+1)n^2, \end{aligned}$$

所有这样的函数 $g(n)$ 属于式(9.16)的右边; 所以式(9.16)是成立的。

有时人们通过假设 O 表示法给出一个增长的确切阶, 他们用它似乎它指出了一个下界而且还指出一个上界。例如, 可以说一个分类 n 个数的算法效率差, “因为它的运行时间为 $O(n^2)$ ”。但是一个 $O(n^2)$ 的运行时间不意味着运行时间也不是 $O(n)$ 。对于下界, 另有一个大 Ω 的表示法:

$$f(n) = \Omega(g(n)) \Leftrightarrow |f(n)| \geq C|g(n)| \quad \text{对某个 } C > 0. \quad (9.17)$$

我们有 $f(n) = \Omega(g(n))$ 当且仅当 $g(n) = O(f(n))$ 。如果 n 足够大, 运行时间为 $\Omega(n^2)$ 的分类算法比起运行时间为 $\Omega(n \log n)$ 的分类算法效率差。

最后有大 Θ , 它指定了一个确切的生长阶:

$$f(n) = \Theta(g(n)) \Leftrightarrow \begin{cases} f(n) = O(g(n)) \\ \text{和 } f(n) = \Omega(g(n)) \end{cases}. \quad (9.18)$$

我们有 $f(n) = \Theta(g(n))$ 当且仅当 $f(n) \asymp g(n)$, 前面我们看到的方程(9.8)的表示法。

Edmund Landau^[194]创造了一个“小 o ”表示法,

$$f(n) = o(g(n)) \Leftrightarrow |f(n)| \leq \varepsilon |g(n)| \quad \begin{cases} \text{对于所有 } n \geq n_0(\varepsilon) \text{ 和} \\ \text{对于所有常数 } \varepsilon > 0. \end{cases} \quad (9.19)$$

这本质上是式(9.3)的关系 $f(n) \sim g(n)$ 。我们还有

$$f(n) \sim g(n) \Leftrightarrow f(n) = g(n) + o(g(n)). \quad (9.20)$$

许多作者在渐近公式中用 ‘ o ’, 但是一个更明确的 ‘ O ’ 表达式几乎总是可取的。例如, 称为“冒泡分类”的计算机方法的平均运行时间依赖于和 $P(n) = \sum_{k=0}^n k^{n-k} k! / n!$ 的渐近值。初等渐近方法足以证明 $P(n) \sim \sqrt{\pi n / 2}$, 这意味着当 $n \rightarrow \infty$ 时, 比 $P(n) / \sqrt{\pi n / 2}$ 趋于 1。然而, 通过考虑差 $P(n) - \sqrt{\pi n / 2}$, 不是比, 最好地了解 $P(n)$ 的真实情况:

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

在中间列中的数值资料不十分使人信服, 因为它确实是离引人注目地证明 $P(n) / \sqrt{\pi n / 2}$ 迅速趋于 1 很远。但是右边的列表表明 $P(n)$ 确实十分接近 $\sqrt{\pi n / 2}$ 。因此如果能导出形式

$$P(n) = \sqrt{\pi n / 2} + O(1)$$

的公式, 或者更加成形的估计, 像

$$P(n) = \sqrt{\pi n / 2} - \frac{2}{3} + O(1 / \sqrt{n}),$$

则我们能说明 $P(n)$ 的情况。渐近分析的较强的方法需要证明 O 的结果, 但是另外的研究要求学习这些较强的方法, 我们通过进一步理解 O -界限来详细地补充这些方法。

并且, 许多分类算法对某些常数 A 和 B 具有形式

$$T(n) = A n \lg n + B n + O(\lg n),$$

的运行时间。在 $T(n) \sim A n \lg n$ 处停止分析没有告知整个情况, 而结果是一个差的对策来选取一个仅依据它的 A 值的分类算法。在差的 ‘ B ’ 的花费处, 常具有一个好的 ‘ A ’ 的算法。由于 $n \lg n$ 比 n 仅稍微增长快一点, 渐近较快的算法 (具有稍微小的 A 值) 可能仅对 n 的值

较快，而在实践中实际从不出现。因此，渐近方法让我们越过第一项，且如果我们要作出方法的正确选取，估计 B 是必要的。

在进行研究 O 之前，让我们再讨论一个数学记法的细小的方面。本章中用三种不同的对数表示法： \lg ， \ln ， \log 。在联系到计算机方法中我们常用 ' \lg '，因为此时常与 2 为底的对数有关；在纯粹数学计算中我们常用 ' \ln '，因为自然对数的公式是好而简单的。但是关于 ' \log ' 怎么样？这不就是学生在中学所学的“通常”底为 10 的对数，在数学和计算机科学中“通常”的对数很不平常吗？是的，通过用 ' \log ' 代表自然对数或 2 为底的对数，许多数学家混淆了结果。这里没有一般的一致意见。但是当在 O 表示法内出现一个对数时，我们通常能松了一口气，因为 O 忽略了相乘的常数。当 $n \rightarrow \infty$ 时，在 $O(\lg n)$ ， $O(\ln n)$ 和 $O(\log n)$ 间没有不同；在 $O(\lg \lg n)$ ， $O(\ln \ln n)$ 和 $O(\log \log n)$ 间相似地也没有不同。我们可选取我们想要的随便哪一个；因为 ' \log ' 更能读出，看来更支持用 ' \log '。所以我们在上下文中一般用 ' \log '，它使得叙述更清楚而并不引入多义性。

9.3 O 操作

像任何数学的形式化那样， O 表示法有使我们能从定义的细节中摆脱出来的操作规则。一旦用定义证明规则是正确的，从此我们能在较高水平上处理，而不必考虑实际验证在另一个中包含一个函数的集合。我们甚至不必计算每个 O 所包含的常数 C ，只要按照保证这样的常数存在的规则。

例如，我们最后能证明

$$n^m = O(n^{m'}), \text{ 当 } m \leq m'; \quad (9.21)$$

$$O(f(n)) + O(g(n)) = O(|f(n)| + |g(n)|). \quad (9.22)$$

于是我们立即能说 $(1/3)n^3 + (1/2)n^2 + (1/6)n = O(n^3) + O(n^3) + O(n^3) = O(n^3)$ ，而不用前节中的麻烦的计算。

下面是一些从定义易得的规则：

$$f(n) = O(f(n)); \quad (9.23)$$

$$c \cdot O(f(n)) = O(f(n)), \text{ 如果 } c \text{ 是常数}; \quad (9.24)$$

$$O(O(f(n))) = O(f(n)); \quad (9.25)$$

$$O(f(n))O(g(n)) = O(f(n)g(n)); \quad (9.26)$$

$$O(f(n)g(n)) = f(n)O(g(n)). \quad (9.27)$$

习题 9 证明了式(9.22)，其他的证明是相似的。不管变量 n 的附带的条件如何，我们总能用右边的形式来替换左边的形式。

方程(9.27)和(9.23)使我们导出等式 $O(f(n)^2) = O(f(n))^2$ 。由于我们能记 $O(\log n)^2$ 而不是 $O((\log n)^2)$ ，有时这有助于避免括号。这两者都优于 ' $O(\log^2 n)$ '， $O(\log^2 n)$ 是不明确的，因为有些作者用它来指 ' $O(\log \log n)$ '。

我们是否还能写

$$O(\log n)^{-1} \text{ 代替 } O((\log n)^{-1})?$$

不能! 这是一种乱用的表示法, 因为函数集合 $1/O(\log n)$ 既不是一个子集, 也不是 $O(1/\log n)$ 的母集。我们能用 $\Omega(\log n)^{-1}$ 合法取代 $O((\log n)^{-1})$, 但是这将是难处理的。所以我们将限于把“ O 外面的指数”使用到固定的正整数指数。

幂级数给出了一些最有用的运算。如果和

$$S(z) = \sum_{n \geq 0} a_n z^n$$

对于某个复数 $z = z_0$ 绝对收敛, 则

$$S(z) = O(1), \text{ 对于所有 } |z| \leq |z_0|.$$

这是显然的, 因为

$$|S(z)| \leq \sum_{n \geq 0} |a_n| |n|^n \leq \sum_{n \geq 0} |a_n| |z_0|^n = C < \infty.$$

特别当 $z \rightarrow 0$ 时, $S(z) = O(1)$, 当 $n \rightarrow \infty$ 时, $S(1/n) = O(1)$, 仅保证至少 z 的一个非零值 $S(z)$ 收敛。我们能用此原理在任何合适之处截断幂级数。且估计 O 的剩余部分。例如, 不仅 $S(z) = O(1)$, 而且

$$S(z) = a_0 + O(z),$$

$$S(z) = a_0 + a_1 z + O(z^2),$$

等等, 因为

$$S(z) = \sum_{0 \leq k < m} a_k z^k + z^m \sum_{n \geq m} a_n z^{n-m}$$

而后面一个和(跟 $S(z)$ 一样)对 $z = z_0$ 绝对收敛, 且为 $O(1)$ 。表 9.1 列出了一些最有用的渐近公式, 其中一半公式是按照这个规则简单地截断幂级数。

以相似的方式能截断 Dirichlet 级数 $\sum_{k \geq 1} a_k / k^z$: 如果当 $z = z_0$ 时 Dirichlet 级数绝对收敛, 我们能在任何项处截断它, 且取得近似式

$$\sum_{1 \leq k < m} a_k / k^z + O(m^{-z}), \text{ 对 } \Re z \geq \Re z_0 \text{ 成立}.$$

表 9.1 中 Bernoulli 数 B_n 的渐近公式说明了这个原理。

另一方面, 表 9.1 中的 H_n , $n!$ 和 $\pi(n)$ 的渐近公式不是截断收敛级数; 如果我们无限地推广, 它们将对所有 n 值发散。在 $\pi(n)$ 的情形中, 这是特别容易看清楚, 因为在 7.3 节例 5 中我们已看到幂级数 $\sum_{k \geq 0} k! / (\ln n)^k$ 处处发散。然而把这些发散级数截断

产生有用的近似。

表 9.1 渐近近似, 当 $n \rightarrow \infty$ 和 $z \rightarrow 0$ 时成立

$$H_n = \ln n + \gamma + \frac{1}{2n} + \frac{1}{12n^2} + \frac{1}{120n^4} + O\left(\frac{1}{n^6}\right). \quad (9.28)$$

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} + O\left(\frac{1}{n^4}\right)\right). \quad (9.29)$$

$$B_n = 2[n\text{偶}](-1)^{n/2-1} \frac{n!}{(2\pi)^2} (1 + 2^{-n} + 3^{-n} + O(4^{-n})). \quad (9.30)$$

$$\pi(n) = \frac{n}{\ln n} + \frac{n}{(\ln n)^2} + \frac{2!}{(\ln n)^3} + \frac{3!}{(\ln n)^4} + O\left(\frac{n}{(\ln n)^5}\right). \quad (9.31)$$

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + O(z^5). \quad (9.32)$$

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + O(z^5). \quad (9.33)$$

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + z^4 + O(z^5). \quad (9.34)$$

$$(1+z)^\alpha = 1 + \alpha z + \binom{\alpha}{2} z^2 + \binom{\alpha}{3} z^3 + \binom{\alpha}{4} z^4 + O(z^5). \quad (9.35)$$

如果渐近近似有形式 $f(n) + O(g(n))$, 其中 $f(n)$ 不涉及 O , 则称它有绝对误差 $O(g(n))$ 。如果渐近近似有形式 $f(n)(1 + O(g(n)))$, 其中 $f(n)$ 不涉及 O , 则它有相对误差 $O(g(n))$ 。例如, 表 9.1 中 H_n 的近似有绝对误差 $O(n^{-6})$; $n!$ 的近似有相对误差 $O(n^{-4})$ 。(式(9.29)的右边实际不具有要求的形式 $f(n) \times (1 + O(n^{-4}))$), 但是如果我们要的话, 能把它再写成

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3}\right) (1 + O(n^{-4}));$$

相似的计算是习题 12 的内容。)此近似的绝对误差是 $O(n^{n-3.5}e^{-n})$ 。如果忽略 O 项, 绝对误差与十进小数点后面正确的十进数字位数有关; 相对误差对应于正确的“有效数字”位数。

我们能利用截断幂级数来证明一般定律

$$\ln(1 + O(f(n))) = O(f(n)), \text{ 如果 } f(n) < 1; \quad (9.36)$$

$$e^{O(f(n))} = 1 + O(f(n)), \text{ 如果 } f(n) = O(1). \quad (9.37)$$

(这里我们假设 $n \rightarrow \infty$; 当 $n \rightarrow 0$ 时 $\ln(1 + O(f(x)))$ 和 $e^{O(f(x))}$ 的相似公式成立。)例如, 设 $\ln(1+g(x))$ 是属于式(9.36)的左边的任何函数。则有常数 C , n_0 和 c 使得

$$|g(n)| \leq C|f(n)| \leq c < 1, \text{ 对于所有 } n \geq n_0.$$

由此得到对于所有 $n \geq n_0$, 无穷和

$$\ln(1 + g(n)) = g(n) \cdot (1 - \frac{1}{2}g(n) + \frac{1}{3}g(n)^2 - \cdots)$$

收敛, 且括号内的级数以常数 $1 + (1/2)c + (1/3)c^2 + \cdots$ 为界限。这就证明了式(9.36), 并且式(9.37)的证明是相似的, 把方程(9.36)和(9.37)结合起来给出有用的公式

$$(1 + O(f(n)))^{O(g(n))} = 1 + O(f(n)g(n)), \begin{cases} \text{如果 } f(n) < 1 \text{ 和} \\ f(n)g(n) = O(1). \end{cases} \quad (9.38)$$

问题 1: 回到轮盘赌

现在让我们在几个渐近问题处碰碰运气。在第三章中我们在一个确定的游戏导出获胜位置的个数的方程(3.13):

$$W = \lfloor N/K \rfloor + \frac{1}{2}K^2 + \frac{5}{2}K - 3, \quad K = \lfloor \sqrt[3]{N} \rfloor.$$

且我们有可能在第九章中导出 W 的渐近形式。当 $N \rightarrow \infty$ 时让我们来估计 W 。

这里的主要思想是用 $N^{1/3} + O(1)$ 替换 K 来去掉下整数括号。于是我们能进一步进行下去且写成

$$K = N^{1/3}(1 + O(N^{-1/3}));$$

称这为“分出大部分”。(我们将大量使用这个诀窍。)现在我们由式(9.38)和(9.26)得到

$$K^2 = N^{2/3}(1 + O(N^{-1/3}))^2 = N^{2/3}(1 + O(N^{-1/3})) = N^{2/3} + O(N^{1/3}).$$

相似地有

$$\begin{aligned} \lfloor N/K \rfloor &= N^{1-1/3}(1 + O(N^{-1/3}))^{-1} + O(1) = N^{2/3}(1 + O(N^{-1/3})) + O(1) \\ &= N^{2/3} + O(N^{1/3}). \end{aligned}$$

由此得到获胜位置的个数为

$$\begin{aligned} W &= N^{2/3} + O(N^{1/3}) + \frac{1}{2}(N^{2/3} + O(N^{1/3})) + O(N^{1/3}) + O(1) \\ &= \frac{3}{2}N^{2/3} + O(N^{1/3}). \end{aligned} \quad (9.39)$$

注意 O 项如何吸收另一项直到仅剩一项; 这是有代表性的, 且它说明了为什么在一个公式中间 O 表示法是有用的。

问题 2: Stirling 公式的扰动

$n!$ 的 Stirling 近似无疑是最有名的渐近公式, 在本章稍后将证明它, 现在让我们更好地了解它的性质。我们能把一种近似形式记成

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{a}{n} + \frac{b}{n^2} + O(n^{-3})\right), \text{ 当 } n \rightarrow \infty, \quad (9.40)$$

(对于确定的常数 a 和 b .) 由于对所有大的 n 此式成立, 当用 $n-1$ 替换 n 时, 它一定也渐近地成立:

$$(n-1)! = \sqrt{2\pi(n-1)} \left(\frac{n-1}{e}\right)^{n-1} \times \left(1 + \frac{a}{n-1} + \frac{b}{(n-1)^2} + O((n-1)^{-3})\right). \quad (9.41)$$

当然, 我们知道 $(n-1)! = n! / n$, 因此这个公式的右边一定简化为式(9.40)的右边除以 n .

所以让我们简化式(9.41). 如果分出大的部分, 第一个因子变得易处理:

$$\begin{aligned} \sqrt{2\pi(n-1)} &= \sqrt{2\pi n} (1 - n^{-1})^{1/2} \\ &= \sqrt{2\pi n} \left(1 - \frac{1}{2n} - \frac{1}{8n^2} + O(n^{-3})\right). \end{aligned}$$

这里用了方程(9.35).

相似得到

$$\begin{aligned} \frac{a}{n-1} &= \frac{a}{n} (1 - n^{-1})^{-1} = \frac{a}{n} + \frac{a}{n^2} + O(n^{-3}); \\ \frac{b}{(n-1)^2} &= \frac{b}{n^2} (1 - n^{-1})^{-2} = \frac{b}{n^2} + O(n^{-3}); \\ O((n-1)^{-3}) &= O(n^{-3} (1 - n^{-1})^{-3}) = O(n^{-3}). \end{aligned}$$

式(9.41)中稍微微妙的事情仅涉及到因子 $(n-1)^{n-1}$, 它等于

$$n^{n-1} (1 - n^{-1})^{n-1} = n^{n-1} (1 - n^{-1})^n (1 + n^{-1} + n^{-2} + O(n^{-3})).$$

(我们每个因子展开直到取得 $O(n^{-3})$ 的相对误差, 因为乘积的相对误差是各个因子的相对误差的和. 所有 $O(n^{-3})$ 的项将合并.)

为了展开 $(1 - n^{-1})^n$, 我们先计算 $\ln(1 - n^{-1})$, 然后形成指数 $e^{n \ln(1 - n^{-1})}$:

$$\begin{aligned} (1 - n^{-1})^n &= \exp(n \ln(1 - n^{-1})) \\ &= \exp\left(n\left(-n^{-1} - \frac{1}{2}n^{-2} - \frac{1}{3}n^{-3} + O(n^{-4})\right)\right) \\ &= \exp\left(-1 - \frac{1}{2}n^{-1} - \frac{1}{3}n^{-2} + O(n^{-3})\right) \\ &= \exp(-1) \cdot \exp\left(-\frac{1}{2}n^{-1}\right) \cdot \exp\left(-\frac{1}{3}n^{-2}\right) \cdot \exp(O(n^{-3})) \\ &= \exp(-1) \cdot \left(1 - \frac{1}{2}n^{-1} + \frac{1}{8}n^{-2} + O(n^{-3})\right) \end{aligned}$$

$$\begin{aligned} & \cdot (1 - \frac{1}{3}n^{-2} + O(n^{-4})) \cdot (1 + O(n^{-3})) \\ & = e^{-1} (1 - \frac{1}{2}n^{-1} - \frac{5}{24}n^{-2} + O(n^{-3})). \end{aligned}$$

这里我们用表示法 $\exp z$ ，而不是 e^z ，因为它可使我们在公式的主行上处理复杂的指数，而不是在上标位置。我们一定要展开具有绝对误差 $O(n^{-4})$ 的 $\ln(1-n^{-1})$ ，为了最后得出 $O(n^{-3})$ 的相对误差，因为对数被 n 相乘。

式(9.41)的右边现在被化为 $\sqrt{2\pi n}$ 乘 n^{n-1}/e^n 乘几个因子的乘积：

$$\begin{aligned} & (1 - \frac{1}{2}n^{-1} - \frac{1}{8}n^{-2} + O(n^{-3})) \cdot (1 + n^{-1} + n^{-2} + O(n^{-3})) \\ & \cdot (1 - \frac{1}{2}n^{-1} - \frac{5}{24}n^{-2} + O(n^{-3})) \cdot (1 + an^{-1} + (a+b)n^{-2} + O(n^{-3})). \end{aligned}$$

把这些因子乘出且吸收所有渐近项为一个 $O(n^{-3})$ 产生

$$1 + an^{-1} + (a+b - \frac{1}{12})n^{-2} + O(n^{-3}).$$

我们希望取得 $1 + an^{-1} + bn^{-2} + O(n^{-3})$ ，因为这就是我们所需要的和(9.40)右边相配的部分。有什么东西弄错了吗？没有，全没有问题；表 9.1 告知 $a=1/12$ ，因此 $a+b-1/12=b$ 。

这个扰动论证并没有证明 Stirling 近似的正确性，但是它证明了结果：公式(9.40)不能成立除非 $a=1/12$ 。如果在式(9.40)中我们用 $cn^{-3} + O(n^{-4})$ 替换 $O(n^{-3})$ ，且作相对误差 $O(n^{-4})$ 的计算，则能导出 $b=1/288$ 。（这不是确定 a 和 b 值的最简方式，但是它行得通。

问题 3：第 n 个素数

方程(9.31)是 $\pi(n)$ 的渐近公式，素数的个数不超过 n 。如果用第 n 个素数 $p = P_n$ 来替换 n ，我们有 $\pi(p) = n$ ；因此当 $n \rightarrow \infty$ 时

$$n = \frac{p}{\ln p} + O\left(\frac{p}{(\log p)^2}\right). \quad (9.42)$$

让我们就 p 来“解”这个方程，我们将知道第 n 个素数的近似大小。

第一步是简化 O 项。如果用 $p/\ln p$ 除两边，我们发现 $n \ln p / p \rightarrow 1$ ；因此 $p/\ln p = O(n)$ 和

$$O\left(\frac{p}{(\log p)^2}\right) = O\left(\frac{n}{\log p}\right) = O\left(\frac{n}{\log n}\right).$$

(我们有 $(\log p)^{-1} \leq (\log n)^{-1}$ ，因为 $p \geq n$ 。)

第二步是除 O 项外式(9.42)两边移项。这是合法的，因为一般规则

$$a_n = b_n + O(f(n)) \Leftrightarrow b_n = a_n + O(f(n)). \quad (9.43)$$

(如果我们在两边乘 -1 , 且把 $a_n + b_n$ 加到两边, 则这些方程的每一个方程能从另一个方程得到。)因此

$$\frac{p}{\ln p} = n + O\left(\frac{n}{\log n}\right) = n(1 + O(1/\log n)),$$

且我们得到

$$p = n \ln p (1 + O(1/\log n)). \quad (9.44)$$

这是 $p = P_n$ 的“近似递归”. 我们的目标是把它改成“近似闭形式”, 我们能渐近地展现递归来做这一点. 所以让我试展现式(9.44).

两边取对数得出

$$\ln p = \ln n + \ln \ln p + O(1/\log n). \quad (9.45)$$

这个值能替换式(9.44)中的 $\ln p$, 但是 we 希望在作替换之前除去所有 p . 在整个过程中的某处, 最后的 p 一定要不出现; 对于递归来说我们不能以标准的方法除去它, 因为对于小的 p 式(9.44)不指定初始条件.

做这项工作的一种方式是从证明较弱的结果 $p = O(n^2)$ 开始. 如果我们平方式(9.44)且除以 pn^2 就得到

$$\frac{p}{n^2} = \frac{(\ln p)^2}{p} (1 + O(1/\log n)),$$

因为当 $n \rightarrow \infty$ 时右边趋于零. 我们知道了 $p = O(n^2)$, 所以 $\log p = O(\log n)$ 和 $\log \log p = O(\log \log n)$. 现在我们根据式(9.45)推得

$$\ln p = \ln n + O(\log \log n);$$

事实上, 用处理中的这个新估计, 我们能推得 $\ln \ln p = \ln \ln n + O(\log \log n / \log n)$, 现在式(9.45)产生

$$\ln p = \ln n + \ln \ln n + O(\log \log n / \log n).$$

我们能把它代入式(9.44)的右边, 得到

$$p = n \ln n + n \ln \ln n + O(n).$$

这是第 n 个素数的近似大小.

我们能用 $\pi(n)$ 的较好近似代替式(9.42)来改进这个估计. 式(9.31)的下一项告知

$$n = \frac{p}{\ln p} + \frac{p}{(\ln p)^2} + O\left(\frac{p}{(\log p)^3}\right); \quad (9.46)$$

如同前面处理的那样, 我们得到递归

$$p = n \ln p (1 + (\ln p)^{-1})^{-1} (1 + O(1/\log n)^2), \quad (9.47)$$

它有相对误差 $O(1/\log n)^2$, 而不是 $O(1/\log n)$ 。取对数且保持适当的精确性(但是不太高)就产生

$$\begin{aligned} \ln p &= \ln n + \ln \ln p + O(1/\log n) \\ &= \ln n (1 + \frac{\ln \ln p}{\ln n} + O(1/\log n)^2); \\ \ln \ln p &= \ln \ln n + \frac{\ln \ln n}{\ln n} + O\left(\frac{\log \log n}{\log n}\right)^2. \end{aligned}$$

最后把这些结果代入式(9.47), 我们的解答弄清楚了它的途径:

$$P_n = n \ln n + n \ln \ln n - n + n \frac{\ln \ln n}{\ln n} + O\left(\frac{n}{\log n}\right). \quad (9.48)$$

例如, 当 $n = 10^6$ 时, 这估计达到 $15\,631\,363.8 + O(n/\log n)$; 第 10^6 个素数实际是 $15\,485\,863$ 。习题 21 证明了如果用更精确的 $\pi(n)$ 的近似开始来代替式(9.46), 则更精确的 P_n 的近似产生。

问题 4: 过去的期终考试的一个和

当 1970~1971 学期间首次在史坦福大学教具体数学时, 学生问到具有绝对误差 $O(n^{-7})$ 的和

$$S_n = \frac{1}{n^2+1} + \frac{1}{n^2+2} + \cdots + \frac{1}{n^2+n} \quad (9.49)$$

的渐近值。让我们设想一下, 这个问题交给我们时, 我们的本能反应首先是什么?

我们并不惊慌失措, 第一个反应是想到大。如果置 $n = 10^{100}$, 且查看和, 我们来看由 n 项组成的和, 每项稍比 $1/n^2$ 小; 因此和稍微小于 $1/n$ 。一般, 通过观察情况且取得解答的粗略估计, 我们通常能取得渐近问题的一个不错的开始。

让我们尝试通过抽出每项的最大部分来改进粗略的估计。我们有

$$\frac{1}{n^2+k} = \frac{1}{n^2(1+k/n^2)} = \frac{1}{n^2} \left(1 - \frac{k}{n^2} + \frac{k^2}{n^4} - \frac{k^3}{n^6} + O\left(\frac{k^4}{n^8}\right) \right),$$

所以自然来尝试求所有这些近似的和:

$$\frac{1}{n^2+1} = \frac{1}{n^2} - \frac{1}{n^4} + \frac{1^2}{n^6} - \frac{1^3}{n^8} + O\left(\frac{1^4}{n^{10}}\right)$$

$$\begin{aligned}
 \frac{1}{n^2+2} &= \frac{1}{n^2} - \frac{2}{n^4} + \frac{2^2}{n^6} - \frac{2^3}{n^8} + O\left(\frac{2^4}{n^{10}}\right) \\
 &\vdots \\
 \frac{1}{n^2+n} &= \frac{1}{n^2} - \frac{n}{n^4} + \frac{n^2}{n^6} - \frac{n^3}{n^8} + O\left(\frac{n^4}{n^{10}}\right) \\
 S_n &= \frac{n}{n^2} - \frac{n(n+1)}{2n^4} + \dots
 \end{aligned}$$

依据前两列的和, 看来似乎取得 $S_n = n^{-1} - (1/2)n^{-2} + O(n^{-3})$; 但是计算不令人满意。

如果我们坚持用此法, 最终将达到目标; 但是有两个原因使得我们将不去操心求其他列的和: 首先, 当 $n/2 \leq k \leq n$ 时, 最后列将给我们的项为 $O(n^{-6})$, 所以我们将有一个 $O(n^{-5})$ 的误差; 这误差太大, 我们将还需包含展开式中另一列。出题者是否存心让别人失望? 我们觉得一定有一个较好的方法。其次, 确实有一种很好的方法, 实况就在我们眼前。

即我们知道 S_n 的闭形式: 就是 $H_{n^2+n} - H_{n^2}$ 。我们知道调和数的好的近似, 所以我们就用它两次:

$$\begin{aligned}
 H_{n^2+n} &= \ln(n^2+n) + \gamma + \frac{1}{2(n^2+n)} - \frac{1}{12(n^2+n)^2} + O\left(\frac{1}{n^8}\right); \\
 H_{n^2} &= \ln n^2 + \gamma + \frac{1}{2n^2} - \frac{1}{12n^4} + O\left(\frac{1}{n^8}\right).
 \end{aligned}$$

现在能抽出大的项且简化, 如同研究 Stirling 的近似所做的那样。我们得到

$$\begin{aligned}
 \ln(n^2+n) &= \ln n^2 + \ln\left(1 + \frac{1}{n}\right) = \ln n^2 + \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \dots; \\
 \frac{1}{n^2+n} &= \frac{1}{n^2} - \frac{1}{n^3} + \frac{1}{n^4} - \dots; \\
 \frac{1}{(n^2+n)^2} &= \frac{1}{n^4} - \frac{2}{n^5} + \frac{3}{n^6} - \dots.
 \end{aligned}$$

所以有许多有帮助的相消, 我们求得

$$\begin{aligned}
 S_n &= n^{-1} - \frac{1}{2}n^{-2} + \frac{1}{3}n^{-3} - \frac{1}{4}n^{-4} + \frac{1}{5}n^{-5} - \frac{1}{6}n^{-6} \\
 &\quad - \frac{1}{2}n^{-3} + \frac{1}{2}n^{-4} - \frac{1}{2}n^{-5} + \frac{1}{2}n^{-6} \\
 &\quad + \frac{1}{6}n^{-5} - \frac{1}{4}n^{-6}
 \end{aligned}$$

附加的项为 $O(n^{-7})$ 。作一点算术运算得:

$$S_n = n^{-1} - \frac{1}{2}n^{-2} - \frac{1}{6}n^{-3} + \frac{1}{4}n^{-4} - \frac{2}{15}n^{-5} + \frac{1}{12}n^{-6} + O(n^{-7}). \quad (9.50)$$

如同前几章中导出确切结果时我们所做的那样, 如果能用数值来检验此解答, 则它将是好的。渐近公式是较难来验证; 一个任意大的常数可能藏在 O 项内, 所以任何数值测试是没有确定结论的。但是在实践中, 没有理由使我们相信陷于困境, 所以我们能假设未知的 O 常数是适当地小。用计算器可求得 $S_4 = 1/17 + 1/18 + 1/19 + 1/20 = 0.2170107$; 当 $n=4$ 时我们的渐近估计为

$$\frac{1}{4} \left(1 + \frac{1}{4} \left(-\frac{1}{2} + \frac{1}{4} \left(-\frac{1}{6} + \frac{1}{4} \left(\frac{1}{4} + \frac{1}{4} \left(-\frac{2}{15} + \frac{1}{4} \cdot \frac{1}{12} \right) \right) \right) \right) \right) = 0.2170125.$$

如果在项 n^{-6} 中产生譬如说误差 $1/12$, 在第 5 个十进位将显出一个 $(1/12)(1/4096)$ 的差; 所以我们的渐近解答很可能是正确的。

问题 5: 一个无限和

现在我们转到由 Solmon Golomb^[122]提出的一个渐近问题:

$$S_n = \sum_{k \geq 1} \frac{1}{k N_n(k)^2} \quad (9.51)$$

的近似值是什么, 其中 $N_n(k)$ 是把 k 记为底 n 表示法所要求的数字的个数?

首先让我们再试作一个粗略估计。数字的个数 $N_n(k)$ 近似为 $\log_n k = \log k / \log n$, 所以此和的项粗略地为 $(\log n)^2 / k (\log k)^2$ 。对 k 求和得到 $\approx (\log n)^2 \cdot \sum_{k \geq 2} 1 / k (\log k)^2$, 而因它与积分

$$\int_2^{\infty} \frac{dx}{x (\ln x)^2} = -\frac{1}{\ln x} \Big|_2^{\infty} = \frac{1}{\ln 2}$$

对照, 此和收敛于一个常数, 所以我们期望 S_n 大约为 $C(\log n)^2$ (对某个常数 C)。

像这样的曲折的分析对确定方向来说是有用的, 但是需要较好的估计来解问题。一种想法是恰切表达 $N_n(k)$:

$$N_n(k) = \lfloor \log_n k \rfloor + 1. \quad (9.52)$$

因此, 例如, 当 $n^2 \leq k < n^3$ 时, k 有 3 个基数 n 的数字, 当 $\lfloor \log_n k \rfloor = 2$ 时这就正好出现。

由此得到 $N_n(k) > \log_n k$, 因此 $S_n = \sum_{k \geq 1} 1 / k N_n(k)^2 < 1 + (\log n)^2 \sum_{k \geq 2} 1 / k (\log k)^2$ 。

如同问题 1 那样处理, 我们能试记 $N_n(k) = \log_n k + O(1)$, 把它代入 S_n 的公式。这里所表示的项 $O(1)$ 总在 0 和 1 之间, 且平均大约为 $1/2$, 所以它看来情况相当好。但是, 关于 S_n , 它仍不是足够好的近似; 当 k 小时, 它给出零有效数字 (也就是说, 高相

对误差), 而这些是提供给和的大部分项, 我们需要一种不同的想法。

在我们采用渐近估计之前, 关键(如同问题 4)是用我们操作技巧把和变成易处理的形式。我们可引入求和的新变量 $m = N_n(k)$:

$$\begin{aligned} S_n &= \sum_{k, m \geq 1} \frac{[m = N_n(k)]}{km^2} \\ &= \sum_{k, m \geq 1} \frac{[n^{m-1} \leq k < n^m]}{km^2} \\ &= \sum_{m \geq 1} \frac{1}{m^2} (H_{n^{m-1}} - H_{n^{m-1}-1}). \end{aligned}$$

这可能看来比开始的和差, 但是实际上前进了一步, 因为对于调和数我们有非常好的近似。

但我们还是踌躇, 且尝试再简化一些, 不必匆忙做渐近。分部求和让我们对需要近似的 $H_{n^{k-1}}$ 的每个值来归并项:

$$S_n = \sum_{k \geq 1} H_{n^{k-1}} \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right).$$

例如, $1/2^2$ 乘 H_{n^2-1} , 然后用 $-1/3^2$ 乘。(我们用了 $H_{n^0-1} = H_0 = 0$ 的事实。)

现在我们准备展开调和数, 估计 $(n-1)!$ 的经验告诉我们估计 H_{n^k} 比估计 $H_{n^{k-1}}$ 容易, 因为 $(n^k - 1)$ 将是凌乱的; 所以我们记

$$\begin{aligned} H_{n^k-1} &= H_{n^k} - \frac{1}{n^k} = \ln n^k + \gamma + \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right) - \frac{1}{n^k} \\ &= k \ln n + \gamma - \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right). \end{aligned}$$

现在我们的和化为

$$\begin{aligned} S_n &= \sum_{k \geq 1} \left(k \ln n + \gamma - \frac{1}{2n^k} + O\left(\frac{1}{n^{2k}}\right) \right) \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) \\ &= (\ln n) \sum_1 + \gamma \sum_2 - \frac{1}{2} \sum_3(n) + O(\sum_3(n^2)). \end{aligned} \quad (9.53)$$

容易的四部分留下: \sum_1 , \sum_2 , $\sum_3(n)$ 和 $\sum_3(n^2)$ 。

让我们首先处理 \sum_3 , 因为 $\sum_3(n^2)$ 是 O 项; 接着将看到所取得的误差类型。(如果完全精确的其他计算总将并入一个 O , 则进行这些计算是无意义的。)这个和简单地为幂级数,

$$\sum_3(x) = \sum_{k \geq 1} \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) x^{-k}.$$

当 $x \geq 1$ 时级数收敛, 所以能在任何要求的地方截断它。如果我们在 $k=1$ 的项处截断 $\sum_3(n^2)$, 得到 $\sum_3(n^2) = O(n^{-2})$, 因此式(9.53)有绝对误差 $O(n^2)$ 。(为了减小这个绝对误差, 我们能用 H_n 的较好的近似, 但是现在 $O(n^{-2})$ 够好了。)如果在 $k=2$ 的项处截断 $\sum_3(n)$, 我们取得

$$\sum_3(n) = \frac{3}{4}n^{-1} + O(n^{-2});$$

这是我们所需要的整个准确度。

现在我们可同样处理 \sum_2 , 因为它是很容易的:

$$\sum_2 = \sum_{k \geq 1} \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right).$$

这是缩短的级数 $(1-1/4)+(1/4-1/9)+(1/9-1/16)+\dots=1$ 。

最后, \sum_1 给出 S_n 的首项, 式(9.53)中 $\ln n$ 的系数:

$$\sum_1 = \sum_{k \geq 1} k \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right).$$

这是 $(1-1/4)+(2/4-2/9)+(3/9-3/16)+\dots=1/1+1/4+1/9+\dots=H_\infty^{(2)}=\pi^2/6$ 。(如果前面不用分部求和, 我们直接看到 $S_n \sim \sum_{k \geq 1} (\ln n)/k^2$, 因为 $H_{n^{k-1}} - H_{n^{k-1}-1} \sim \ln n$; 所以虽然分部求和使其他一些工作做起来容易, 但对计算首项没有帮助。)

现在已计算了式(9.53)中的每个 \sum , 所以我们可以把这些结果合在一起, 得到 Golomb 问题的解答:

$$S_n = \frac{\pi^2}{6} \ln n + \gamma - \frac{3}{8n} + O\left(\frac{1}{n^2}\right). \quad (9.54)$$

注意, 此式比原先的曲折估计 $C(\log n)^2$ 增长得更慢。有时一个离散的和不服从连续的直观。

问题 6: 大 Φ

在第四章的结束处, 我们看到 Farey 级数 \mathcal{F}_n 中的因子数为 $1+\Phi(n)$, 其中

$$\Phi(n) = \varphi(1) + \varphi(2) + \dots + \varphi(n);$$

而在式(4.62)中我们表明

$$\Phi(n) = \frac{1}{2} \sum_{k \geq 1} \mu(k) \lfloor n/k \rfloor \lfloor 1+n/k \rfloor. \quad (9.55)$$

现在当 n 大时让我们试估计 $\Phi(n)$ 。(原先引导 Bachmann 创造 O -表示法的就是像这样的和。)

想起大, 也许 $\Phi(n)$ 将和 n^2 成正比。如果就用 $\lfloor n/k \rfloor$ 代替最后因子 $\lfloor 1+n/k \rfloor$, 我们将有 $|\Phi(n)| \leq (1/2) \sum_{k \geq 1} \lfloor n/k \rfloor^2 \leq (1/2) \sum_{k \geq 1} (n/k)^2 = (\pi^2/12)n^2$, 因为 Möbius 函数 $\mu(k)$ 或者为 -1 , 或者为 0 , 或者为 $+1$ 。在最后因子中, 另外的 ‘ $1+$ ’ 加了 $\sum_{k \geq 1} \mu(k) \lfloor n/k \rfloor$; 但是对于 $k > N$, 这是零, 所以其绝对值不能大于 $nH_n = O(n \log n)$ 。

这个初步分析表明, 我们将发现把它记为

$$\begin{aligned} \Phi(n) &= \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\left(\frac{n}{k} \right) + O(1) \right)^2 = \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\left(\frac{n}{k} \right)^2 + O\left(\frac{n}{k}\right) \right) \\ &= \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\frac{n}{k} \right)^2 + \sum_{k=1}^n O\left(\frac{n}{k}\right) \\ &= \frac{1}{2} \sum_{k=1}^n \mu(k) \left(\frac{n}{k} \right)^2 + O(n \log n) \end{aligned}$$

是有助的。此表示移去了下整记号; 剩下的问题是计算具有准确度 $O(n \log n)$ 的无下整记号的和 $(1/2) \sum_{k=1}^n \mu(k) n^2 / k^2$; 换句话说, 我们要计算具有准确度 $O(n^{-1} \log n)$ 的 $\sum_{k=1}^n \mu(k) 1/k^2$ 。但是这是容易的; 我们能简单地从头到尾上升到 $k = \infty$, 因为新近添加的项是

$$\begin{aligned} \sum_{k \geq n} \frac{\mu(k)}{k^2} &= O\left(\sum_{k \geq n} \frac{1}{k^2}\right) = O\left(\sum_{k \geq n} \frac{1}{k(k-1)}\right) \\ &= O\left(\sum_{k \geq n} \left(\frac{1}{k-1} - \frac{1}{k}\right)\right) = O\left(\frac{1}{n}\right). \end{aligned}$$

在式 (7.88) 中我们证明了 $\sum_{k \geq 1} \mu(k)/k^z = 1/\zeta(z)$, 因此 $\sum_{k \geq 1} \mu(k)/k^2 = 1/(\sum_{k \geq 1} 1/k^2) = 6/\pi^2$, 我们得到解答:

$$\Phi(n) = \frac{3}{\pi^2} n^2 + O(n \log n). \quad (9.56)$$

9.4 两个渐近的特殊技巧

现在我们有一些 O 操作的工具, 让我们从稍微高的角度来考虑我们所做的事情。当

需要处理难对付的问题时，在我们的渐近武器库中将有一些重要的武器。

特殊技巧 1: 自益

当我们在 9.3 节的问题 3 中估计第 n 个素数 P_n 时，我们解一个形式

$$P_n = n \ln P_n (1 + O(1/\log n))$$

的渐近递归。通过首先用递归来表明较弱结果 $O(n^2)$ ，我们证明了 $P_n = n \ln n + O(n)$ 。这是一般的称为自益的方法的特殊情形，此法中我们渐近地解一个递归，由粗糙的估计开始，再把它代入递归；用此法常能导出越来越好的估计，“通过自益来处理”。

这里是另一个恰好说明自益的问题：当 $n \rightarrow \infty$ 时，母函数

$$G(z) = \exp\left(\sum_{k \geq 1} \frac{z^k}{k^2}\right) \quad (9.57)$$

中的系数 $g_n = [z^n]G(z)$ 的渐近值是什么？如果这个方程对 z 求导数，我们求得

$$G'(z) = \sum_{n=0}^{\infty} n g_n z^{n-1} = \left(\sum_{k \geq 1} \frac{z^{k-1}}{k}\right) G(z);$$

使两边的 z^{n-1} 的系数相等得到递归

$$n g_n = \sum_{0 \leq k < n} \frac{g_k}{n-k}. \quad (9.58)$$

我们的问题等价于找式(9.58)解的一个渐近公式，具有初始条件 $g_0 = 1$ 。前几个值

n	0	1	2	3	4	5	6
g_n	1	1	$\frac{3}{4}$	$\frac{19}{36}$	$\frac{107}{288}$	$\frac{641}{2400}$	$\frac{51103}{259200}$

并没有显示出一种型式，且在 Sloane 的手册中^[270]并不出现整数序列 $\langle n!^2 g_n \rangle$ ；所以 g_n 的一个闭形式看来是绝对做不到的，也许最多有希望推导渐近信息。

我们处理这个问题，首先看出对于所有 $n \geq 0$ ， $0 < g_n \leq 1$ ；用归纳法这是容易证明的。所以我们有一个起始：

$$g_n = O(1).$$

事实上，对于自益操作来说，这个方程被用来“采取措施促其发展”：把它代入式(9.58)的右边产生

$$n g_n = \sum_{0 \leq k < n} \frac{O(1)}{n-k} = H_n O(1) = O(\log n);$$

因此我们得到

$$g_n = O\left(\frac{\log n}{n}\right), \text{ 对 } n > 1.$$

我们还能再自益:

$$\begin{aligned} ng_n &= \frac{1}{n} + \sum_{0 < k < n} \frac{O((1 + \log k)/k)}{n - k} \\ &= \frac{1}{n} + \sum_{0 < k < n} \frac{O(\log n)}{k(n - k)} \\ &= \frac{1}{n} + \sum_{0 < k < n} \left(\frac{1}{k} + \frac{1}{n - k}\right) \frac{O(\log n)}{n} \\ &= \frac{1}{n} + \frac{2}{n} H_{n-1} O(\log n) = \frac{1}{n} O(\log n)^2, \end{aligned}$$

得到

$$g_n = O\left(\frac{\log n}{n}\right)^2. \quad (9.59)$$

想永远这样进行下去吗? 也许对于所有 m 我们将有 $g_n = O(n^{-1} \log n)^m$.

实际上不是, 且我们仅达到缩小返回的目的. 下一个自益的尝试涉及到和

$$\begin{aligned} \sum_{0 < k < n} \frac{1}{k^2(n - k)} &= \sum_{0 < k < n} \left(\frac{1}{nk^2} + \frac{1}{n^2k} + \frac{1}{n^2(n - k)} \right) \\ &= \frac{1}{n} H_{n-1}^{(2)} + \frac{2}{n^2} H_{n-1}, \end{aligned}$$

它是 $\Omega(n^{-1})$; 所以我们不能取得落在 $\Omega(n^{-2})$ 之下的 g_n 的一个估计.

事实上, 现在应用老的技巧取出最大部分足以了解 g_n :

$$\begin{aligned} ng_n &= \sum_{0 \leq k < n} \frac{g_k}{n} + \sum_{0 \leq k < n} g_k \left(\frac{1}{n - k} - \frac{1}{n} \right) \\ &= \frac{1}{n} \sum_{k \geq 0} g_k - \frac{1}{n} \sum_{k \geq n} g_k + \frac{1}{n} \sum_{0 \leq k < n} \frac{k g_k}{n - k}. \end{aligned} \quad (9.60)$$

这里的第一个和是 $G(1) = \exp(1/1 + 1/4 + 1/9 + \dots) = e^{\pi^2/6}$, 因为对所有 $|z| \leq 1$, $G(z)$ 收敛. 第二个和是第一个和的末尾部分, 用式(9.59)可取得一个上界:

$$\sum_{k \geq n} g_k = O\left(\sum_{k \geq n} \frac{(\log k)^2}{k^2}\right) = O\left(\frac{(\log n)^2}{n}\right).$$

例如因为

$$\sum_{k>n} \frac{(\log k)^2}{k^2} < \sum_{m>1} \sum_{n < k \leq n+m-1} \frac{(\log n^{m+1})^2}{k(k-1)} < \sum_{m>1} \frac{(m+1)^2 (\log n)^2}{n^m},$$

这最后的估计随之而来。

(习题 54 讨论了估计这种末尾部分的更一般的方法。)

通过已熟悉的论证, 式(9.60)中的第三个和是

$$O\left(\sum_{0 \leq k < n} \frac{(\log n)^2}{k(n-k)}\right) = O\left(\frac{(\log n)^3}{n}\right).$$

所以式(9.60)证明了

$$g_n = \frac{e^{n^2/6}}{n^2} + O(\log n / n^3) \quad (9.61)$$

最后, 我们能把此公式送回到递归, 再次用自益: 结果是

$$g_n = \frac{e^{n(2/6)}}{n^2} + O(\log n / n^3). \quad (9.62)$$

(习题 23 窥视剩下的 O 项的内情。)

特殊技巧 2: 交换末尾部分

我们推出式(9.62), 有点像推导 $\Phi(n)$ 的渐近值式(9.56)的相同方法: 在两种情形中, 我们由一个有限和开始, 但是通过研究一个无限和而取得一个渐近值。我们不能简单地把 O 引入被加数而得到无限和; 当 k 小时且另外当 k 大时, 我们必须小心地用一个方法。

这些推导是重要的 3 步渐近求和方法的特殊情形, 现在我们将以非常一般的形式来讨论。每当要估计 $\sum_k a_k(n)$ 的值时, 我们能试用下列方法:

1. 首先把和分成两个不相交的范围, D_n 和 T_n 。在 D_n 上的求和应是“主要的”部分, 就此意义当 n 大时, 它包含足够多的项来确定和的有效数字位。在另一个范围 T_n 上的求和仅为“末尾”端, 它对全体来说影响很小。

2. 找一个渐近估计

$$a_k(n) = b_k(n) + O(c_k(n))$$

当 $k \in D_n$ 时它成立。当 $k \in T_n$ 时 O 界不必成立。

3. 现在证明三个和的每一个和是小的:

$$\sum_a(n) = \sum_{k \in T_n} a_k(n); \quad \sum_b(n) = \sum_{k \in T_n} b_k(n); \quad \sum_c(n) = \sum_{k \in D_n} |c_k(n)|. \quad (9.63)$$

如果能成功地完成所有的 3 步, 我们得到一个好的估计:

$$\sum_{k \in D_n \cup T_n} a_k(n) = \sum_{k \in D_n \cup T_n} b_k(n) + O(\sum_a(n)) + O(\sum_b(n)) + O(\sum_c(n)).$$

为什么是这样。我们能“切去”给定和的末尾，取得范围 D_n 中的一个好估计，对于一个好估计 D_n 是必要的：

$$\sum_{k \in D_n} a_k(n) = \sum_{k \in D_n} (b_k(n) + O(c_k(n))) = \sum_{k \in D_n} b_k(n) + O(\sum_c(n)).$$

我们能用另一个末尾部分来替换此末尾部分，即使对于老的来说新的末尾部分可能是一个很糟的近似，因为末尾部分确实无关紧要：

$$\begin{aligned} \sum_{k \in T_n} a_k(n) &= \sum_{k \in T_n} (b_k(n) - b_k(n) + a_k(n)) \\ &= \sum_{k \in T_n} b_k(n) + O(\sum_b(n)) + O(\sum_a(n)). \end{aligned}$$

当我们计算式(9.60)中的和时，例如得到

$$a_k(n) = [0 \leq k < n] g_k / (n - k),$$

$$b_k(n) = g_k / n,$$

$$c_k(n) = k g_k / n(n - k);$$

求和的范围是

$$D_n = \{0, 1, \dots, n-1\}, \quad T_n = \{n, n+1, \dots\};$$

我们发现

$$\sum_a(n) = 0, \quad \sum_b(n) = O((\log n)^2 / n^2), \quad \sum_c(n) = O((\log n)^3 / n^2).$$

这就引出式(9.61)。

当估计式(9.55)中的 $\Phi(n)$ 时，类似地得到

$$a_k(n) = \mu(k) \lfloor n/k \rfloor \lfloor 1 + n/k \rfloor, \quad b_k(n) = \mu(k) n^2 / k^2, \quad c_k(n) = n/k;$$

$$D_n = \{1, 2, \dots, n\}, \quad T_n = \{n+1, n+2, \dots\}.$$

注意到 $\sum_a(n) = 0$, $\sum_b(n) = O(n)$, $\sum_c(n) = O(n \log n)$, 我们导出式(9.56)。

这里是另一个可行末尾部分转换的例子。(不像前面的例子那样，此例完全一般地说明特殊技巧，具有 $\sum_a(n) \neq 0$ 。) 我们找

$$L_n = \sum_{k \geq 0} \frac{\ln(n+2^k)}{k!}$$

的渐近值。当 k 小时，出现对此和的大的作用，因为 $k!$ 在分母中。在此范围中我们有

$$\ln(n + 2^k) = \ln n + \frac{2^k}{n} - \frac{2^{2k}}{2n^2} + O\left(\frac{2^{3k}}{n^3}\right). \quad (9.64)$$

我们能证明，对于 $0 \leq k < \lfloor \lg n \rfloor$ ，此估计成立，因为由 O 截断的开始的项以收敛级数

$$\sum_{m \geq 3} \frac{2^{km}}{m! n^m} \leq \frac{2^{3k}}{n^3} \sum_{m \geq 3} \frac{2^{k(m-3)}}{n^{m-3}} \leq \frac{2^{3k}}{n^3} (1 + \frac{1}{2} + \frac{1}{4} + \cdots) = \frac{2^{3k}}{n^3} \cdot 2$$

(在此范围中， $2^k/n \leq 2^{\lfloor \lg n \rfloor - 1}/n \leq 1/2$ 。)为界限。

所以我们能应用刚才所说的 3 步法，具有

$$a_k(n) = \ln(n + 2^k) / k!,$$

$$b_k(n) = (\ln n + 2^k/n - 4^k/2n^2) / k!,$$

$$c_k(n) = 8^k / n^3 k!;$$

$$D_n = \{0, 1, \dots, \lfloor \lg n \rfloor - 1\},$$

$$T_n = \{\lfloor \lg n \rfloor, \lfloor \lg n \rfloor + 1, \dots\}.$$

我们只需找出式(9.63)中三个 \sum 的好的界限，且我们将知道 $\sum_{k \geq 0} a_k(n) \approx \sum_{k \geq 0} b_k(n)$ 。

用于和 $\sum_c(n) = \sum_{k \in D_n} 8^k / n^3 k!$ 的主要部分的误差显然以 $\sum_{k \geq 0} 8^k / n^3 k! = e^8 / n^3$ 为界限，所以能以 $O(n^{-3})$ 来替换它。新的末尾部分误差是

$$\begin{aligned} \left| \sum_b(n) \right| &= \left| \sum_{k \geq \lfloor \lg n \rfloor} b_k(n) \right| < \sum_{k \geq \lfloor \lg n \rfloor} \frac{\ln n + 2^k + 4^k}{k!} \\ &< \frac{\ln n + 2^{\lfloor \lg n \rfloor} + 4^{\lfloor \lg n \rfloor}}{\lfloor \lg n \rfloor!} \sum_{k \geq 0} \frac{4^k}{k!} = O\left(\frac{n^2}{\lfloor \lg n \rfloor!}\right). \end{aligned}$$

由于 $\lfloor \lg n \rfloor!$ 比任何 n 的幂增长更快，所以 $\sum_c(n) = O(n^{-3})$ 覆盖了这个很小的误差。来自最先末尾部分的误差，

$$\sum_a(n) = \sum_{k \geq \lfloor \lg n \rfloor} a_k(n) < \sum_{k \geq \lfloor \lg n \rfloor} \frac{k + \ln n}{k!},$$

还是较小。

最后，容易把 $\sum_{k \geq 0} b_k(n)$ 相加成一个闭形式，且我们得到希望的渐近公式：

$$\sum_{k \geq 0} \frac{\ln(n + 2^k)}{k!} = e \ln n + \frac{e^2}{n} - \frac{e^4}{2n^2} + O\left(\frac{1}{n^3}\right). \quad (9.65)$$

事实上，我们所用的方法显然使得它对于任何指定的 $m > 0$,

$$\sum_{k=0}^{\infty} \frac{\ln(n+2^k)}{k!} = e \ln n + \sum_{k=1}^{m-1} (-1)^{k+1} \frac{e^{2^k}}{k n^k} + O\left(\frac{1}{n^m}\right). \quad (9.66)$$

(如果我们让 $m \rightarrow \infty$, 这是对于所有指定 n 发散的级数的一个截断。)

在我们的解中仅有一个缺陷: 我们真要注意。我们在假设 $k < \lceil \lg n \rceil$ 基础上推出式 (9.64), 但是习题 53 证明了对于所有 k 值所说的估计实际是成立的。如果我们知道较强的一般结果, 将不必用两种末尾部分的技巧, 我们能直接达到最后的公式! 但是后面将遇到交换末尾部分达到最适当处理的问题。

9.5 Euler 的求和公式

现在讨论下一个特殊技巧, 它是本书将要讨论的最后一个重要的技巧, 1732 年 Leonhard Euler^[82] 首先提出的近似和的一般方法。(此思想也联系到稍后独立发现它的 Edinburgh 的数学教授 Colin Maclaurin 的名字^[211, p.305])

这里是公式

$$\sum_{a \leq k < b} f(k) = \int_a^b f(x) dx + \sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(x) \Big|_a^b + R_m, \quad (9.67)$$

$$\text{其中 } R_m = (-1)^{m+1} \int_a^b \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx, \quad \begin{array}{l} \text{整数 } a \leq b, \\ \text{整数 } m \geq 1. \end{array} \quad (9.68)$$

左边是我们可能要计算的一个典型的和; 右边是涉及积分和导数的和的另一个表达式。如果 $f(x)$ 是充分“平滑”的函数, 它将有 m 个导数 $f'(x), \dots, f^{(m)}(x)$, 此公式结果是一个等式。在剩余 R_m 小的意义下, 右边常是左边和的极好的近似。例如, 我们将看到 $n!$ 的 Stirling 近似是 Euler 求和公式的结果, 所以是调和数 H_n 的渐近近似。

式 (9.67) 中的数 B_k 是第六章中遇到的 Bernoulli 数; 式 (9.68) 中的函数 $B_m(\{x\})$ 是第七章中遇到的 Bernoulli 多项式。如同第三章中那样, 记号 $\{x\}$ 表示分数部分 $x - \lfloor x \rfloor$ 。Euler 求和公式可以说是把有关的一切合在一起。

让我们回想起小的 Bernoulli 数的值, 因为由它们列出接近的 Euler 一般公式总是方便的:

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_8 = -\frac{1}{30};$$

$$B_3 = B_5 = B_7 = B_9 = B_{11} = \dots = 0.$$

Jakob Bernoulli 在研究整数幂的和时发现了这些数, 而 Euler 公式说明了为什么: 如果置 $f(x) = x^{m-1}$, 我们有 $f^{(m)}(x) = 0$; 因此 $R_m = 0$, 式 (9.67) 化为

$$\begin{aligned}\sum_{a \leq k < b} k^{m-1} &= \left. \frac{x^m}{m} \right|_a^b + \sum_{k=1}^m \frac{B_k}{k!} (m-1)^{k-1} \left. x^{m-k} \right|_a^b \\ &= \frac{1}{m} \sum_{k=0}^m \binom{m}{k} B_k \cdot (b^{m-k} - a^{m-k}).\end{aligned}$$

例如, 当 $m=3$ 时, 我们得到求和的适用的例子:

$$\sum_{0 \leq k < n} k^2 = \frac{1}{3} \left(\binom{3}{0} B_0 n^3 + \binom{3}{1} B_1 n^2 + \binom{3}{2} B_2 n \right) = \frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6}.$$

(这是最后一次在本书中推出这个有名的公式。)

在证明 Euler 公式之前, 让我们查看一个高水平的理由(Lagrange^[192]提出的), 为什么这样一个公式该存在。第二章定义了差分算子 Δ , 且说明了 \sum 是 Δ 的逆算子, 就像 \int 是微分算子 D 的逆算子, 我们能用 Taylor 公式, 依据 D 表达 Δ 如下:

$$f(x+\varepsilon) = f(x) + \frac{f'(x)}{1!} \varepsilon + \frac{f''(x)}{2!} \varepsilon^2 + \dots$$

置 $\varepsilon=1$ 得出

$$\begin{aligned}\Delta f(x) &= f(x+1) - f(x) \\ &= f'(x)/1! + f''(x)/2! + f'''(x)/3! + \dots \\ &= (D/1! + D^2/2! + D^3/3! + \dots)f(x) = (e^D - 1)f(x).\end{aligned}\quad (9.69)$$

这里 e^D 表示微分运算 $1 + D/1! + D^2/2! + D^3/3! + \dots$, 由于 $\Delta = e^D - 1$, 逆算子 $\sum = 1/\Delta$ 应是 $1/(e^D - 1)$; 我们从表 7.3 知道 $z/(e^z - 1) = \sum_{k \geq 0} B_k z^k/k!$ 是关于 Bernoulli 数的幂级数。因此

$$\sum = \frac{B_0}{D} + \frac{B_1}{1!} + \frac{B_2}{2!} D + \frac{B_3}{3!} D^2 + \dots = \int + \sum_{k \geq 1} \frac{B_k}{k!} D^{k-1}.\quad (9.70)$$

把此算子方程应用到 $f(x)$, 加上下限产生

$$\sum_a^b f(x) \delta x = \int_a^b f(x) dx + \sum_{k \geq 1} \frac{B_k}{k!} f^{(k-1)}(x) \Big|_a^b \quad (9.71)$$

这就是无剩余项的 Euler 求和公式(9.67)。(事实上, Euler 没有考虑剩余, 直到 1823 年 S.D.Poisson 发表的一篇近似求和的重要论文^[236]之前。剩余项是重要的, 因为无限和 $\sum_{k \geq 1} (B_k/k!) f^{(k-1)}(x) \Big|_a^b$ 常常发散。式(9.71)的推导纯粹是形式的, 不考虑收敛。)

现在让我们包含剩余来证明式(9.67)。证明情形 $a=0$ 和 $b=1$, 即

$$f(0) = \int_0^1 f(x) dx + \sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(x) \Big|_0^1 - (-1)^m \int_0^1 \frac{B_m(x)}{m!} f^{(m)}(x) dx$$

就够了, 因为对于任何整数 l 我们能用 $f(x+l)$ 替换 $f(x)$, 取得

$$f(l) = \int_l^{l+1} f(x)dx + \sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(x) \Big|_l^{l+1} - (-1)^m \int_l^{l+1} \frac{B_m(\{x\})}{m!} f^{(m)}(x)dx.$$

一般公式(9.67)就是范围 $a \leq l < b$ 上这个等式的和, 因为中间的项恰好缩短了.

当 $a=0$, $b=1$ 时, 用归纳法对 m 归纳证明, 从 $m=1$ 开始:

$$f(0) = \int_0^1 f(x)dx - \frac{1}{2}(f(1) - f(0)) + \int_0^1 (x - \frac{1}{2})f'(x)dx.$$

(一般由方程

$$B_m(x) = \binom{m}{0} B_0 x^m + \binom{m}{1} B_1 x^{m-1} + \cdots + \binom{m}{m} B_m x^0 \quad (9.72)$$

定义 Bernoulli 多项式 $B_m(x)$, 因此特别有 $B_1(x) = x - 1/2$. 换句话说, 我们要证明

$$\frac{f(0) + f(1)}{2} = \int_0^1 f(x)dx + \int_0^1 (x - \frac{1}{2})f'(x)dx.$$

但是这就是分部积分公式

$$u(x)v(x) \Big|_0^1 = \int_0^1 u(x)dv(x) + \int_0^1 v(x)du(x) \quad (9.73)$$

中 $u(x) = f(x)$ 和 $v(x) = x - 1/2$ 的特别情况. 因此情形 $m=1$ 是容易的.

当 $m > 1$, 经过 $m-1$ 到 m 且完成归纳, 我们需要表明 $R_{m-1} = (B_{m-1}/(m-1)!)f^{(m-1)}(x) \Big|_0^1 + R_m$, 即

$$(-1)^m \int_0^1 \frac{B_{m-1}(x)}{(m-1)!} f^{(m-1)}(x)dx = \frac{B_m}{m!} f^{(m-1)}(x) \Big|_0^1 - (-1)^m \int_0^1 \frac{B_m(x)}{m!} f^{(m)}(x)dx.$$

此式化为方程

$$(-1)^m B_m f^{(m-1)}(x) \Big|_0^1 = m \int_0^1 B_{m-1}(x) f^{(m-1)}(x)dx + \int_0^1 B_m(x) f^{(m)}(x)dx.$$

再一次把式(9.73)应用于这两个积分, $u(x) = f^{(m-1)}(x)$ 和 $v(x) = B_m(x)$, 因为 Bernoulli 多项式(9.72)的导数是

$$\begin{aligned} \frac{d}{dx} \sum_k \binom{m}{k} B_k x^{m-k} &= \sum_k \binom{m}{k} (m-k) B_k x^{m-k-1} \\ &= m \sum_k \binom{m-1}{k} B_k x^{m-1-k} = m B_{m-1}(x). \end{aligned} \quad (9.74)$$

(吸收等式(5.7)在这里是有用的。)所以要求的公式将成立当且仅当

$$(-1)^m B_m f^{(m-1)}(x) \Big|_0^1 = B_m(x) f^{(m-1)}(x) \Big|_0^1.$$

换句话说, 我们需要有

$$(-1)^m B_m = B_m(1) = B_m(0), \quad m > 1. \quad (9.75)$$

这有点麻烦, 因为 $B_m(0)$ 显然等于 B_m , 不等于 $(-1)^m B_m$, 但是实际上不成问题, 因为 $m > 1$; 我们知道当 m 是奇数时, B_m 是零。(但还是侥幸地脱险了。)

为了完成 Euler 求和公式的证明, 我们需要表明 $B_m(1) = B_m(0)$, 这相当于说

$$\sum_k \binom{m}{k} B_k = B_m, \quad m > 1.$$

但这就是 Bernoulli 数的定义式(6.79), 所以我们完成了。

等式 $B'_m(x) = m B_{m-1}(x)$ 意味着

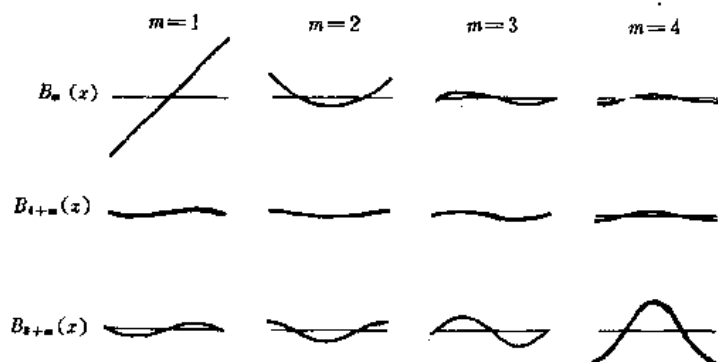
$$\int_0^1 B_m(x) dx = \frac{B_{m+1}(1) - B_{m+1}(0)}{m+1},$$

我们现在知道当 $m \geq 1$ 时, 此积分为零。因此 Euler 公式中的剩余项

$$R_m = \frac{(-1)^{m+1}}{m!} \int_a^b B_m(\{x\}) f^{(m)}(x) dx,$$

用函数 $B_m(\{x\})$ 乘 $f^{(m)}(x)$, 它的平均值是零。这意味着 R_m 有相当的机会是小的。

由于 $B_m(x)$ 支配 R_m , 让我们更详细地查看 $B_m(x)$ $0 \leq x \leq 1$ 。这里是前 12 个 m 值的 $B_m(x)$ 的图:



虽然 $B_3(x)$ 到 $B_9(x)$ 很小, 但是 Bernoulli 多项式和数最终取得很大。幸亏 R_m 有一个补偿的因子 $1/m!$, 它有助于把结果稳定下来。

当 $m \geq 3$ 时, $B_m(x)$ 的图开始看来很像正弦波; 习题 58 证明 $B_m(x)$ 事实上能用

$\cos(2\pi x - (1/2)\pi m)$ 的负倍数来很好地近似, 具有相对误差 $1/2^m$.

一般对于 $0 < x < 1/2$, $B_{4k+1}(x)$ 是负的, 对于 $1/2 < x < 1$, 是正的. 所以对于 $0 < x < 1/2$, 它的积分 $B_{4k+2}(x)/(4k+2)$ 下降, 对于 $1/2 < x < 1$, 上升. 并且, 我们有

$$B_{4k+1}(1-x) = -B_{4k+1}(x), \quad 0 \leq x \leq 1,$$

由此得到

$$B_{4k+2}(1-x) = B_{4k+2}(x), \quad 0 \leq x \leq 1.$$

常数项 B_{4k+2} 引起的积分 $\int_0^1 B_{4k+2}(x)dx$ 为零, 因此 $B_{4k+2} > 0$. $B_{4k+2}(x)$ 的积分为 $B_{4k+3}(x)/(4k+3)$, 所以当 $0 < x < 1/2$ 时它一定为正, 当 $1/2 < x < 1$ 时为负; 此外, $B_{4k+3}(1-x) = -B_{4k+3}(x)$, 所以 $B_{4k+3}(x)$ 有 $B_{4k+1}(x)$ 所述的性质, 但取消了. 所以 $B_{4k+4}(x)$ 有 $B_{4k+2}(x)$ 所说的性质, 但取消了. 所以 $B_{4k+5}(x)$ 有 $B_{4k+1}(x)$ 所说的性质; 我们完成一个循环, 对所有 k 归纳地建立了所说的性质.

依据此分析, $B_{2m}(x)$ 的最大值一定出现在 $x=0$ 处或 $x=1/2$ 处. 习题 17 证明

$$B_{2m}\left(\frac{1}{2}\right) = (2^{1-2m} - 1)B_{2m}; \quad (9.76)$$

因此我们有

$$|B_{2m}(\{x\})| \leq |B_{2m}|. \quad (9.77)$$

这能用来建立 Euler 求和公式中剩余的一个有用的上界, 因为我们从式(6.89)知道

$$\frac{|B_{2m}|}{(2m)!} = \frac{2}{(2\pi)^{2m}} \sum_{k \geq 1} \frac{1}{k^{2m}} = O((2\pi)^{-2m}), \quad \text{当 } m > 0.$$

所以能把 Euler 公式(9.67)再写成如下:

$$\begin{aligned} \sum_{a \leq k \leq b} f(k) &= \int_a^b f(x)dx - \frac{1}{2}f(x)\Big|_a^b + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} f^{(2k-1)}(x)\Big|_a^b \\ &\quad + O((2\pi)^{-2m}) \int_a^b |f^{(2m)}(x)|dx. \end{aligned} \quad (9.78)$$

例如, 若 $f(x) = e^x$, 所有导数是相同的, 此公式告诉我们 $\sum_{a \leq k \leq b} e^k = (e^b - e^a)(1 - 1/2 + B_2/2! + B_4/4! + \cdots + B_{2m}/(2m)! + O((2\pi)^{-2m}))$. 当然, 我们知道此和实际是几何级数, 等于 $(e^b - e^a)/(e - 1) = (e^b - e^a) \sum_{k \geq 0} B_k/k!$.

如果对于 $a \leq x \leq b$, $f^{(2m)}(x) \geq 0$, 则积分 $\int_a^b |f^{(2m)}(x)|dx$ 就是 $f^{(2m-1)}(x)\Big|_a^b$, 所以我们有

$$|R_{2m}| \leq \left| \frac{B_{2m}}{(2m)!} f^{(2m-1)}(x) \right|_a^b;$$

换句话说, 最后项的数值是剩余的界限(就是剩余前的一项), 此时, 我们能给出一个更好的估计, 如果我们知道

$$f^{(2m+2)}(x) \geq 0 \text{ 和 } f^{(2m+4)}(x) \geq 0, \quad a \leq x \leq b. \quad (9.79)$$

由此出现, 这意味着关系

$$R_{2m} = \theta \frac{B_{2m+2}}{(2m+2)!} f^{(2m+1)}(x) \Big|_a^b, \quad \text{对某个 } 0 < \theta_m < 1; \quad (9.80)$$

换句话说, 剩余将位于 0 和式(9.78)中第一个抛弃项(如果我们增加 m , 此项将跟在最后项后面)之间。

这里是证明: 对于所有 m , Euler 求和公式成立, 且当 $m > 0$ 时 $B_{2m+1} = 0$; 因此 $R_{2m} = R_{2m+1}$, 第一个抛弃项一定是

$$R_{2m} - R_{2m+2}.$$

所以我们要表明 R_{2m} 位于 0 和 $R_{2m} - R_{2m+2}$ 之间, 且这是真的当且仅当 R_{2m} 和 R_{2m+2} 有相反的正负号。我们断定

$$f^{(2m+2)}(x) \geq 0 \quad \text{对 } a \leq x \leq b \text{ 意味着 } (-1)^m R_{2m} \geq 0. \quad (9.81)$$

此式与式(9.79)一起将证明 R_{2m} 和 R_{2m+2} 有相反的正负号, 所以完成了式(9.80)的证明。

如果回想起 R_{2m+1} 的定义以及对 $B_{2m+1}(\{x\})$ 的图证明的结果, 那末证明式(9.81)是不困难的。即, 我们有

$$R_{2m} = R_{2m+1} = \int_a^b \frac{B_{2m+1}(\{x\})}{(2m+1)!} f^{(2m+1)}(x) dx,$$

且 $f^{(2m+1)}(x)$ 是上升的, 因为它的导数 $f^{(2m+2)}(x)$ 是正的。(更精确地说, $f^{(2m+1)}(x)$ 是非降的, 因为它的导数是非负的。) $B_{2m+1}(\{x\})$ 的图看来好像是 $(-1)^{m+1}$ 乘正弦波, 所以当用一个上升函数来乘它时, 在几何上显然每个正弦波的另一半比第一半影响大。正如所希望的那样, 这使得 $(-1)^m R_{2m+1} \geq 0$ 。习题 16 形式地证明了此结论。

9.6 最后的求和

正如我们准备结束本书那样, 现在来总结一下。我们将把 Euler 求和公式用到一些有趣和重要的例子。

求和 1: 这个求和是很容易的

但是首先我们将考虑一个有趣的不重要的例子, 即已经知道如何处理的一个和。让我们来看, 如果把 Euler 求和公式用到缩短的和

$$S_n = \sum_{1 \leq k \leq n} \frac{1}{k(k+1)} = \sum_{1 \leq k \leq n} \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{n+1},$$

Euler 求和公式告诉我们什么。这并不妨碍用一系列渐近方式来处理 Euler 公式的第一个重大的应用。

我们同样可把函数 $f(x) = 1/x(x+1)$ 写成部分分式开始,

$$f(x) = \frac{1}{x} - \frac{1}{x+1},$$

因为这使它较容易来积分和微分。事实上, 我们有 $f'(x) = -1/x^2 + 1/(x+1)^2$ 和 $f''(x) = 2/x^3 - 2/(x+1)^3$; 一般有

$$f^{(k)}(x) = (-1)^k k! \left(\frac{1}{x^{k+1}} - \frac{1}{(x+1)^{k+1}} \right), \quad k \geq 0.$$

并且

$$\int_0^n f(x) dx = \ln x - \ln(x+1) \Big|_1^n = \ln \frac{2n}{n+1}.$$

把此代入求和公式(9.67), 得到

$$S_n = \ln \frac{2n}{n+1} - \sum_{k=1}^{\infty} (-1)^k \frac{B_k}{k} \left(\frac{1}{n^k} - \frac{1}{(n+1)^k} - 1 + \frac{1}{2^k} \right) + R_m(n),$$

其中 $R_m(x) = - \int_1^n B_m(\{x\}) (1/x^{m+1} - 1/(x+1)^{m+1}) dx$.

例如, 当 $m=4$ 时右边为

$$\begin{aligned} \ln \frac{2n}{n+1} - \frac{1}{2} \left(\frac{1}{n} - \frac{1}{n+1} - \frac{1}{2} \right) - \frac{1}{12} \left(\frac{1}{n^2} - \frac{1}{(n+1)^2} - \frac{3}{4} \right) \\ + \frac{1}{120} \left(\frac{1}{n^4} - \frac{1}{(n+1)^4} - \frac{15}{16} \right) + R_4(n). \end{aligned}$$

这有点糟了, 确实看来它不像实际答案 $1 - n^{-1}$. 但是让我们无论以什么方式继续下去, 看看得到什么。我们知道如何把右边的项展开成 n 的负幂, 譬如说展到 $O(n^{-5})$:

$$\begin{aligned} \ln \frac{n}{n+1} &= -n^{-1} + \frac{1}{2}n^{-2} - \frac{1}{3}n^{-3} + \frac{1}{4}n^{-4} + O(n^{-5}); \\ \frac{1}{n+1} &= n^{-1} - n^{-2} + n^{-3} - n^{-4} + O(n^{-5}); \\ \frac{1}{(n+1)^2} &= n^{-2} - 2n^{-3} + 3n^{-4} + O(n^{-5}); \end{aligned}$$

$$\frac{1}{(n+1)^4} \sim n^{-4} + O(n^{-5}).$$

所以近似式的右边的项总计为

$$\begin{aligned} & \ln 2 + \frac{1}{4} + \frac{1}{16} - \frac{1}{128} + (-1 - \frac{1}{2} + \frac{1}{2})n^{-1} + (\frac{1}{2} - \frac{1}{2} - \frac{1}{12} + \frac{1}{12})n^{-2} \\ & + (-\frac{1}{3} + \frac{1}{2} - \frac{2}{12})n^{-3} + (\frac{1}{4} - \frac{1}{2} + \frac{3}{12} + \frac{1}{120} - \frac{1}{120})n^{-4} + R_4(n) \\ & = \ln 2 + \frac{39}{128} - n^{-1} + R_4(n) + O(n^{-5}). \end{aligned}$$

n^{-2} , n^{-3} 和 n^{-4} 的系数刚好消去, 正如它们该消去那样。

如果一切都好, 将能证明 $R_4(n)$ 渐近地小, 可能是 $O(n^{-5})$, 我们将得到和的一个近似。但是我们不可能证明这一点, 因为偶然知道精确的常数项为 1, 不是 $\ln 2 + (39/128)$ (它近似为 0.997 8)。所以 $R_4(n)$ 实际等于 $(89/128) - \ln 2 + O(n^{-4})$, 但是 Euler 求和公式不告诉我们这一点。

换句话说, 我们失败了。

如果我们让 m 取得越来越大, 作为尝试确定结果的一种方式, 我们注意到近似式中的常数项形成一种型式:

$$\ln 2 - \frac{1}{2}B_1 + \frac{1}{2} \cdot \frac{3}{4}B_2 - \frac{1}{3} \cdot \frac{7}{8}B_3 + \frac{1}{4} \cdot \frac{15}{16}B_4 - \frac{1}{5} \cdot \frac{31}{32}B_5 + \cdots.$$

我们也许能证明当项数变成无限时, 此级数趋于 1? 但是不是这样, Bernoulli 数取十分大。例如, $B_{22} = (854\,513/138) > 6\,192$, 所以 $|R_{22}(n)|$ 将比 $|R_4(n)|$ 大很多。我们完全失败了。

然而, 有一个方法显现出来, 这个没有被注意的方法在 Euler 公式的其他应用中是重要的。关键是注意当 $n \rightarrow \infty$ 时 $R_4(n)$ 趋于一个确定的极限

$$\lim_{n \rightarrow \infty} R_4(n) = - \int_1^{\infty} B_4(\{x\}) \left(\frac{1}{x^5} - \frac{1}{(x+1)^5} \right) dx = R_4(\infty).$$

每当 $x \rightarrow \infty$, $f^{(m)}(x) = O(x^{-2})$ 时, 积分 $\int_1^{\infty} B_4(\{x\}) f^{(m)}(x) dx$ 将存在, 此时 $f^{(4)}(x)$ 确实合法。并且, 我们有

$$\begin{aligned} R_4(n) &= R_4(\infty) + \int_n^{\infty} B_4(\{x\}) \left(\frac{1}{x^5} - \frac{1}{(x+1)^5} \right) dx \\ &= R_4(\infty) + O\left(\int_n^{\infty} x^{-6} dx\right) = R_4(\infty) + O(n^{-5}). \end{aligned}$$

于是对于某个常数 C , 我们用 Euler 求和公式证明

$$\begin{aligned}\sum_{1 \leq k \leq n} \frac{1}{k(k+1)} &= \ln 2 + \frac{39}{128} - n^{-1} + R_4(\infty) + O(n^{-1}) \\ &= C - n^{-1} + O(n^{-5}).\end{aligned}$$

我们不知道常数是什么，一定要用其他方法建立它，但是 Euler 求和公式能让我们推出常数存在。

假设我们选取一个很大的 m 值。于是相同理由将告诉我们

$$R_m(n) = R_m(\infty) + O(n^{-m-1}),$$

对于一些常数 c_2, c_3, \dots ，我们有公式

$$\sum_{1 \leq k \leq n} \frac{1}{k(k+1)} = C - n^{-1} + c_2 n^{-2} + c_3 n^{-3} + \dots + c_m n^{-m} + O(n^{-m-1}).$$

此时，我们知道这些 c 恰好为零，但是让我们证明它，恰好增强对 Euler 公式的置信度。

项 $\ln(n/(n+1))$ 提供 $(-1)^m/m$ 给 c_m ；项 $(-1)^{m+1}(B_m/m)n^{-m}$ 提供 $(-1)^{m+1}B_m/m$ ；

项 $(-1)^k(B_k/k)(n+1)^{-k}$ 提供 $(-1)^m \binom{m-1}{k-1} B_k/k$ 。所以

$$\begin{aligned}(-1)^m c_m &= \frac{1}{m} - \frac{B_m}{m} + \sum_{k=1}^m \binom{m-1}{k-1} \frac{B_k}{k} \\ &= \frac{1}{m} - \frac{B_m}{m} + \frac{1}{m} \sum_{k=1}^m \binom{m}{k} B_k = \frac{1}{m} (1 - B_m + B_m(1) - 1).\end{aligned}$$

当 $m > 1$ 时，它确实为零。我们证明了

$$\sum_{1 \leq k \leq n} \frac{1}{k(k+1)} = C - n^{-1} + O(n^{-m-1}), \text{ 对所有 } m \geq 1. \quad (9.82)$$

但这不足以证明此和恰好等于 $C - n^{-1}$ ，实际值大概可能是 $C - n^{-1} + 2^{-n}$ 。纵然我们没有明确地计算任何剩余，但是 Euler 求和公式对于任意大的 m ，给出 $O(n^{-m-1})$ 。

再讨论求和 1：概括和一般化

在离开一系列讨论之前，让我们再考察从较高的观点来看我们做了什么。我们从和

$$S_n = \sum_{1 \leq k \leq n} f(k)$$

开始，且用 Euler 求和公式写出

$$S_n = F(n) - F(1) + \sum_{k=1}^n (T_k(n) - T_k(1)) + R_m(n), \quad (9.83)$$

其中 $F(x)$ 是 $\int f(x)dx$, $T_k(x)$ 是涉及 B_k 和 $f^{(k-1)}(x)$ 的一个确定的项。我们还注意到, 有一个常数 c 使得

$$f^{(m)}(x) = O(x^{c-m}) \quad \text{当 } x \rightarrow \infty, \text{ 对所有大的 } m.$$

(即, $f(k)$ 是 $1/k(k+1)$; $F(x)$ 是 $\ln(x/(x+1))$; $T_k(x)$ 是 $(-1)^{k+1} \times (B_k/k)(x^{-k} - (x+1)^{-k})$; 且 c 是 -2 。) 对于所有充分大的 m 值, 这意味着剩余有一个小的尾部,

$$\begin{aligned} R'_m(n) &= R'_m(\infty) - R'_m(n) \\ &= (-1)^{m+1} \int_n^\infty \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx = O(n^{c+1-m}). \end{aligned} \quad (9.84)$$

所以我们推得, 存在一个常数 C 使得

$$S_n = F(n) + C + \sum_{k=1}^n T_k(n) - R'_m(n). \quad (9.85)$$

(注意到, C 很好地并入了 $T_k(1)$ 项, 它是一个讨厌的项。)

每当 $R'_m(\infty)$ 存在时, 简单地断言 C 的存在, 省去了将来问题中不必要的工作。

现在让我们假设对于 $1 \leq x < n$, $f^{(2m+2)}(x) \geq 0$ 和 $f^{(2m+4)}(x) \geq 0$ 。我们已证明, 这意味着剩余的一个简单的界限式(9.80),

$$R'_{2m}(n) = \theta_{m,n} (T_{2m+2}(n) - T_{2m+2}(1)),$$

其中 $\theta_{m,n}$ 位于 0 和 1 之间的某处。但是我们实际并不涉及 $R'_{2m}(n)$ 和 $T_{2m+2}(1)$ 的界限; 当引入常数 C 时, 我们毕竟除去了 $T_k(1)$ 。我们实际要的是像

$$-R'_{2m}(n) = \varphi_{m,n} T_{2m+2}(n)$$

的一个界限, 其中 $0 < \varphi_{m,n} < 1$; 这就使我们从式(9.85)推得

$$S_n = F(n) + C + T_1(n) + \sum_{k=1}^n T_{2k}(n) + \varphi_{m,n} T_{2m+2}(n), \quad (9.86)$$

因此剩余确实将在零和第一个抛弃项之间。

把前面的论证稍加修改将使结果更完善。让我们假设

$$f^{(2m+2)}(x) \geq 0 \text{ 和 } f^{(2m+4)}(x) \geq 0, \text{ 当 } x \rightarrow \infty. \quad (9.87)$$

就剩余项而论, 式(9.85)的右边就像 $a=n$ 和 $b=\infty$ 的 Euler 求和公式的右边的负值, 通过对 m 归纳, 产生接连的剩余。所以能用前面的论证。

求和 2: 协调调和数

现在我们已经从一个平凡的(但是可靠的)例子学到许多东西, 能容易地处理一个不平凡的问题。让我们用 Euler 求和公式推出 H_n 的近似式, 我们已经承认了一些时候。

此时, $f(x) = 1/x$. 由求和 1, 我们已经知道 f 的积分和导数, 并且当 $x \rightarrow \infty$ 时, $f^{(m)}(x) = O(x^{-m-1})$. 所以能立刻代入公式(9.85): 对某个常数 C ,

$$\sum_{1 \leq k \leq n} \frac{1}{k} = \ln n + C + B_1 n^{-1} - \sum_{k=1}^m \frac{B_{2k}}{2kn^{2k}} - R'_{2m}(n).$$

左边的和是 H_{n-1} , 不是 H_n ; 但是处理 H_{n-1} , 然后加 $1/n$ 比右边用 $(n+1)$ 更方便.

于是 $B_1 n^{-1}$ 将变成 $(B_1 + 1)n^{-1} = 1/(2n)$. 让我们称常数为 γ 而不是 C , 因为 Euler 常

数 γ 实际定义为 $\lim_{n \rightarrow \infty} (H_n - \ln n)$.

用我们刚才阐述的理论能很好地估计剩余项, 因为对所有 $x > 0$, $f^{(2m)}(x) = (2m)!/x^{2m+1} \geq 0$. 所以式(9.86)告诉我们

$$H_n = \ln n + \gamma + \frac{1}{2n} - \sum_{k=1}^m \frac{B_{2k}}{2kn^{2k}} + \theta_{m,n} \frac{B_{2m+2}}{(2m+2)n^{2m+2}}, \quad (9.88)$$

其中 $\theta_{m,n}$ 是 0 和 1 之间的某个分数. 这是一般公式, 它的前几项列在表 9.1 中. 例如, 当 $m=2$ 时, 我们取得

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \frac{\theta_{2,n}}{252n^6}. \quad (9.89)$$

顺便说一下, 当 $n=2$ 时, 此方程给出 γ 的一个好的近似:

$$\gamma = H_2 - \ln 2 - \frac{1}{4} + \frac{1}{48} - \frac{1}{1920} + \varepsilon = 0.577\,165\cdots + \varepsilon,$$

其中 ε 在 0 和 $1/16\,128$ 之间. 如果取 $n=10^4$ 和 $m=250$, 我们得到 γ 的值精确到 1 271 个小数位, 从而开头部分为^[171]:

$$\gamma = 0.577\,215\,664\,901\,532\,860\,606\,512\,090\,082\,402\,43\cdots. \quad (9.90)$$

但是 Euler 常数也出现在其他公式中, 甚至使得更有效地计算它^[281].

求和 3: Stirling 近似式

如果 $f(x) = \ln x$, 则有 $f'(x) = 1/x$, 我们几乎能用倒数求和所做的相同计算来计算对数的和. Euler 求和公式产生

$$\begin{aligned} \sum_{1 \leq k \leq n} \ln k &= n \ln n - n + \sigma - \frac{\ln}{2} + \sum_{k=1}^m \frac{B_{2k}}{2k(2k-1)n^{2k-1}} \\ &\quad + \varphi_{m,n} \frac{B_{2m+2}}{(2m+2)(2m+1)n^{2m+1}} \end{aligned}$$

其中 σ 是一个确定的常数, “Stirling 常数”, $0 < \varphi_{m,n} < 1$. (此时, $f^{(2m)}(x)$ 是负的, 不是正的; 但是我们仍能说, 通过第一个抛弃项产生剩余, 因为我们能从 $f(x) = -\ln x$ 开始, 而

不是 $f(x) = \ln x$.) 两边加 $\ln n$ 得到

$$\ln n! = n \ln n - n + \frac{\ln n}{2} + \sigma + \frac{1}{12n} - \frac{1}{360n^3} + \frac{\varphi_{2,n}}{1260n^5} \quad (9.91)$$

(当 $m=2$ 时), 且通过两边取 'exp' 能得到表 9.1 中的近似. (e^σ 的值证明是 $\sqrt{2\pi}$, 但是我们不准备引出这个公式. 事实上, 在 de Moivre^[64] 证明了常数存在之后几年, Stirling 才发现 σ 的闭形式.)

如果指定 m 且 $n \rightarrow \infty$, 在绝对误差的意义上, 一般公式给出 $\ln n!$ 的一个越来越好的近似, 因此, 在相对误差的意义上, 它给出 $n!$ 的一个越来越好的近似. 但是如果指定 n 且 m 上升, 则误差界限 $|B_{2m+2}| / (2m+2)(2m+1)n^{2m+1}$ 下降到一个确定点, 然后开始上升. 所以近似到达了这样一点, 在这点处不能以一类不定的原则限制逼近 $n!$ 的数值.

在第五章的方程(5.83)中, 我们用 Euler 提出的定义

$$\frac{1}{\alpha!} = \lim_{n \rightarrow \infty} \binom{n+\alpha}{n} n^{-\alpha}$$

把阶乘推广到任意实数 α . 假设 α 是一个大数; 则

$$\ln \alpha! = \lim_{n \rightarrow \infty} (\alpha \ln n + \ln n! - \sum_{k=1}^n \ln(\alpha + k)),$$

且能以 $f(x) = \ln(x + \alpha)$ 用 Euler 求和公式来估计此和:

$$\sum_{k=1}^n \ln(k + \alpha) = F_m(\alpha, n) - F_m(\alpha, 0) + R_{2m}(\alpha, n),$$

$$F_m(\alpha, x) = (x + \alpha) \ln(x + \alpha) - x + \frac{\ln(x + \alpha)}{2} + \sum_{k=1}^m \frac{B_{2k}}{2k(2k-1)(x + \alpha)^{2k-1}},$$

$$R_{2m}(\alpha, n) = \int_0^n \frac{B_{2m}(\{x\})}{2m} \frac{dx}{(x + \alpha)^{2m}}.$$

(这里我们用了式(9.67), $a=0$ 和 $b=n$, 然后在两边加 $\ln(n+\alpha) - \ln \alpha$.) 如果从 $\ln n!$ 的 Stirling 近似式中减去这个 $\sum_{k=1}^n \ln(k + \alpha)$ 近似式, 然后加上 $\alpha \ln n$ 且取 $n \rightarrow \infty$ 的极限, 我们得到

$$\ln \alpha! = \alpha \ln \alpha - \alpha + \frac{\ln \alpha}{2} + \sigma + \sum_{k=1}^m \frac{B_{2k}}{(2k)(2k-1)\alpha^{2k-1}} - \int_0^\infty \frac{B_{2m}(\{x\})}{2m} \frac{dx}{(x + \alpha)^{2m}}.$$

因为 $\alpha \ln n + n \ln n - n + (1/2) \ln n - (n + \alpha) \ln(n + \alpha) + n - (1/2) \ln(n + \alpha) \rightarrow -\alpha$, 而不在这里的其他项趋于零. 因此 Stirling 近似对于推广的阶乘(对于 Gamma 函数 $\Gamma(\alpha+1) = \alpha!$)恰好相当于通常阶乘一样.

求和 4: 钟形被加数

让我们现在转到一个完全不同风格的和:

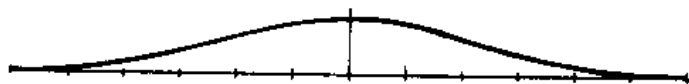
$$\Theta_n = \sum_k e^{-k^2/n} \quad (9.92)$$

$$= \cdots + e^{-9/n} + e^{-4/n} + e^{-1/n} + 1 + e^{-1/n} + e^{-4/n} + e^{-9/n} + \cdots.$$

这是一个双重无限和, 当 $k=0$ 它的项达到最大值 $e^0 = 1$. 我们称它为 Θ_n , 因为这是一个涉及到量 $e^{-1/n}$ 自乘 $p(k)$ 次幂的幂级数, 其中 $p(k)$ 是次数 2 的多项式; 这种幂级数通常称为“theta 函数”, 如果 $n = 10^{100}$, 则有

$$e^{k^2/n} = \begin{cases} e^{-.01} \approx 0.990\,05, & \text{当 } k = 10^{49}; \\ e^{-1} \approx 0.367\,88, & \text{当 } k = 10^{50}; \\ e^{-100} < 10^{-43}, & \text{当 } k = 10^{51}. \end{cases}$$

当它下降且停留在十分靠近零的地方, 被加数停留在十分靠近 1 的地方直到 k 到达 \sqrt{n} 附近. 我们猜测 Θ_n 将和 \sqrt{n} 成比例. 这里是 $e^{-k^2/n}$ 的图, 当 $n = 10$ 时:



n 的较大值恰好由 \sqrt{n} 的因子水平拉直图形.

在 Euler 求和公式中令 $f(x) = e^{-x^2/n}$ 且取 $a = -\infty$, $b = +\infty$, 我们能估计 Θ_n . (如果无穷看来太可怕, 设 $a = -A$ 和 $b = +B$, 然后取 $A, B \rightarrow \infty$ 的极限.)

如果用 $u\sqrt{n}$ 替换 x , 则 $f(x)$ 的积分为

$$\int_{-\infty}^{+\infty} e^{-x^2/n} dx = \sqrt{n} \int_{-\infty}^{+\infty} e^{-u^2} du = \sqrt{n} C.$$

$\int_{-\infty}^{+\infty} e^{-u^2} du$ 的值是众所周知的, 但是现在称它为 C . 当我们完成代入 Euler 求和公式之后再恢复它.

接着我们需要知道的是导数序列 $f'(x)$, $f''(x)$, \cdots , 为此而置

$$f(x) = g(x/\sqrt{n}), \quad g(x) = e^{-x^2}$$

是方便的. 于是微积分的导数规则告诉我们

$$\frac{df(x)}{dx} = \frac{dg(y)}{dy} \frac{dy}{dx}, \quad y = \frac{x}{\sqrt{n}};$$

这相当于说

$$f'(x) = \frac{1}{\sqrt{n}} g'(x/\sqrt{n}).$$

由归纳法我们得到

$$f^{(k)}(x) = n^{-k/2} g^{(k)}(x/\sqrt{n}).$$

例如, 我们有 $g'(x) = -2xe^{-x^2}$ 和 $g''(x) = (4x^2 - 2)e^{-x^2}$; 因此

$$f'(x) = \frac{1}{\sqrt{n}} \left(-2\frac{x}{\sqrt{n}} \right) e^{-x^2/n}, \quad f''(x) = \frac{1}{n} \left(4\left(\frac{x}{\sqrt{n}}\right)^2 - 2 \right) e^{-x^2/n}.$$

如果我们处理较简单的函数 $g(x)$, 易见发生的结果。

我们不必完全计算 $g(x)$ 的导数, 因为我们仅关心当 $x = \pm\infty$ 时的极限值。为此, 只需注意到 $g(x)$ 的每一个导数是 e^{-x^2} 乘一个 x 的多项式:

$$g^{(k)}(x) = P_k(x)e^{-x^2}, \quad \text{其中 } P_k \text{ 是次数为 } k \text{ 的多项式.}$$

由归纳法获得此结果。

当 $x \rightarrow \pm\infty$ 时, 负指数 e^{-x^2} 趋于零比 $P_k(x)$ 趋于无穷更快, 所以我们对于所有 $k \geq 0$ 有

$$f^{(k)}(+\infty) = f^{(k)}(-\infty) = 0.$$

所以所有项

$$\sum_{k=1}^{\infty} \frac{B_k}{k!} f^{(k-1)}(x) \Big|_{-\infty}^{+\infty}$$

变成零, 留下出自 $\int f(x)dx$ 的项和剩余:

$$\begin{aligned} \Theta_n &= C\sqrt{n} + (-1)^{m+1} \int_{-\infty}^{+\infty} \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx \\ &= C\sqrt{n} + \frac{(-1)^{m+1}}{n^{m/2}} \int_{-\infty}^{+\infty} \frac{B_m(\{x\})}{m!} g^{(m)}\left(\frac{x}{\sqrt{n}}\right) dx \\ &= C\sqrt{n} + \frac{(-1)^{m+1}}{n^{(m-1)/2}} \int_{-\infty}^{+\infty} \frac{B_m(\{u\sqrt{n}\})}{m!} P_m(u) e^{-u^2} du \quad (x = u\sqrt{n}) \\ &= C\sqrt{n} + O(n^{(1-m)/2}). \end{aligned}$$

由于 $|B_m(\{u\sqrt{n}\})|$ 是有界的, 且每当 P 是多项式时积分 $\int_{-\infty}^{+\infty} |P(u)|e^{-u^2} du$ 存在, 所以这里有 O 估计。(此 O 包含的常数依赖于 m .)

对于任意大的 M , 我们证明了 $\Theta_n = C\sqrt{n} + O(n^{-M})$; Θ_n 和 $C\sqrt{n}$ 之间的差“指数律地小”, 所以让我们来决定 Θ_n 的值中扮演这样一个大的角色的常数 C .

决定 C 的一种方法是查表中熟悉的积分; 但是我们宁愿要知道如何能推出值, 以致

遇到表中未列出的积分时, 我们能处理它们. 如果我们巧妙地处理重积分, 初等微积分足以计算 C ,

$$C^2 = \int_{-\infty}^{+\infty} e^{-x^2} dx \int_{-\infty}^{+\infty} e^{-y^2} dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy.$$

转换成极坐标得到

$$\begin{aligned} C^2 &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \\ &= \frac{1}{2} \int_0^{2\pi} d\theta \int_0^{\infty} e^{-u} du \quad (u=r^2) \\ &= \frac{1}{2} \int_0^{2\pi} d\theta = \pi. \end{aligned}$$

所以 $C = \sqrt{\pi}$. $x^2 + y^2 = r^2$ 是周长为 $2\pi r$ 的圆的方程这一事实说明了 π 进入的理由.

计算 C 的另一种方法是用 \sqrt{t} 替换 x , 用 $\frac{1}{2}t^{-1/2}dt$ 替换 dx :

$$C = \int_{-\infty}^{+\infty} e^{-x^2} dx = 2 \int_0^{\infty} e^{-x^2} dx = \int_0^{\infty} t^{-1/2} e^{-t} dt.$$

依据式(5.84), $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$, 所以这个积分等于 $\Gamma(1/2)$. 我们证明了 $\Gamma(1/2) = \sqrt{\pi}$.

于是我们的最后公式为

$$\Theta_n = \sum_k e^{-k^2/n} = \sqrt{\pi n} + O(n^{-M}), \text{ 对所有指定的 } M. \quad (9.93)$$

O 中的常数依赖于 M , 这就是我们称 M 是“指定的”理由.

例如, 当 $n=2$ 时, 无限和 Θ_2 等于 2.506 628 288; 虽然 n 很小, 这已是 $\sqrt{2\pi} = 2.506 628 275$ 的一个非常好的近似. Θ_{100} 的值有 427 个小数位与 $10\sqrt{\pi}$ 一致! 习题 59 用改进的方法推出 Θ_n 的一个快速收敛的级数, 它得出

$$\Theta_n / \sqrt{\pi n} = 1 + 2e^{-n\pi^2} + O(e^{-4n\pi^2}). \quad (9.94)$$

求和 5: 决定性的事实的讨论

现在我们将处理最后一个和, 它的结果将告诉我们 Stirling 常数 σ 的值. 最后这个和也说明了本章(以及全书)的许多其他技巧, 所以这是一种恰当的方式来结束我们对具体数学的探讨.

最后的任务看来几乎不合理地容易: 我们尝试用 Euler 求和公式来找

$$A_n = \sum_k \binom{2n}{k}$$

的渐近值.

这是我们已经了解的解答的另一种情形(恰当吗?), 但是对于老问题试用新方法总是有趣的, 以致我们对照实际情况, 可能发现一些新东西.

所以我们想象大且了解到 A_n 的主要部分来自接近 $k=n$ 的中间的项. 选取一种记法使得一个和的最大部分出现在 $k=0$ 附近, 这总是一个好主意, 因为这样我们就能用交换尾部的诀窍来除去具有大的 $|k|$ 的项. 所以我们用 $n+k$ 替换 n :

$$A_n = \sum_k \binom{2n}{n+k} = \sum_k \frac{(2n)!}{(n+k)! (n-k)!}.$$

由于当 n 大 k 小时, 我们知道近似的 $(n \pm k)!$, 所以看来情况相当好.

现在我们要结合交换尾部诀窍进行 3 个步骤的过程. 即, 我们要写

$$\frac{(2n)!}{(n+k)! (n-k)!} = a_k(n) = b_k(n) + O(c_k(n)), \text{ 对 } k \in D_n,$$

以致能得估计

$$A_n = \sum_k b_k(n) + O\left(\sum_{k \notin D_n} a_k(n)\right) + O\left(\sum_{k \in D_n} b_k(n)\right) + \sum_{k \in D_n} O(c_k(n)).$$

所以让我们尝试在 $|k|$ 小的范围内估计 $\binom{2n}{n+k}$, 我们可用表 9.1 中出现的 Stirling 近似, 但是用式(9.91)中的对数等价式容易处理:

$$\begin{aligned} \ln a_k(n) &= \ln(2n)! - \ln(n+k)! - \ln(n-k)! \\ &= 2n \ln 2n - 2n + \frac{1}{2} \ln 2n + \sigma + O(n^{-1}) \\ &\quad - (n+k) \ln(n+k) + n+k - \frac{1}{2} \ln(n+k) - \sigma + O((n+k)^{-1}) \\ &\quad - (n-k) \ln(n-k) + n-k - \frac{1}{2} \ln(n-k) - \sigma + O((n-k)^{-1}). \end{aligned} \quad (9.95)$$

我们要把它转换成一个好的, 简单的 O 估计.

交换尾部方法允许我们处理仅当 k 在“主要的”集合 D_n 中时成立的估计. 但是应如何定义 D_n 呢? 我们必须使 D_n 足够小, 我们能作出一个好的估计; 例如, 最好不让 k 接近 n , 或者在式(9.95)中的项 $O((n-k)^{-1})$ 将放大. 然而 D_n 一定要足够大, 与全部和相比尾部项 ($k \notin D_n$ 的项) 可忽略地小. 对于找一个合适的集合 D_n , 尝试法通常是必要的, 在此问题中我们将要做的计算表明定义:

$$k \in D_n \Leftrightarrow |k| \leq n^{1/2+\varepsilon} \quad (9.96)$$

是合理的。在我们开始知道范围之后，这里的 ε 是我们后面可选取的一个小的正常数。（ O 估计将依赖于 ε 的值。）方程(9.95)现在化为

$$\begin{aligned} \ln a_k(n) = & (2n + \frac{1}{2})\ln 2 - \sigma - \frac{1}{2}\ln n + O(n^{-1}) \\ & - (n + k + \frac{1}{2})\ln(1 + k/n) - (n - k + \frac{1}{2})\ln(1 - k/n). \end{aligned} \quad (9.97)$$

（我们分出大的对数部分，记

$$\ln(n \pm k) = \ln n + \ln(1 \pm k/n),$$

这样消去了许多 $\ln n$ 项。）

现在我们需要渐近地展开项 $\ln(1 \pm k/n)$ ，直到当 $n \rightarrow \infty$ 时有一个趋于零的误差项。我们用 $(n \pm k + 1/2)$ 乘 $\ln(1 \pm k/n)$ ，所以应展开对数，直到达到 $o(n^{-1})$ ，用假设 $|k| \leq n^{1/2+\varepsilon}$ ：

$$\ln(1 \pm \frac{k}{n}) = \pm \frac{k}{n} - \frac{k^2}{2n^2} + O(n^{-3/2+3\varepsilon}).$$

乘以 $n \pm k + 1/2$ 产生

$$\pm k - \frac{k^2}{2n} + \frac{k^2}{n} + O(n^{-1/2+3\varepsilon}),$$

加的其他项被吸收入 $O(n^{-1/2+3\varepsilon})$ 。所以式(9.97)变成

$$\ln a_k(n) = (2n + \frac{1}{2})\ln 2 - \sigma - \frac{1}{2}\ln n - k^2/n + O(n^{-1/2+3\varepsilon}).$$

取指数，我们得到

$$a_k(n) = \frac{2^{2n+1/2}}{e^\sigma \sqrt{n}} e^{-k^2/n} (1 + O(n^{-1/2+3\varepsilon})), \quad (9.98)$$

加上

$$b_k(n) = \frac{2^{2n+1/2}}{e^\sigma \sqrt{n}} e^{-k^2/n}, \quad c_k(n) = 2^{2n} n^{-1+3\varepsilon} e^{-k^2/n}$$

就是我们的近似。注意到， k 以一种很简单的方式参与 $b_k(n)$ 和 $c_k(n)$ 。我们很幸运，因为我们将 k 上求和。

交换尾部诀窍告诉我们，如果完成好估计工作，则 $\sum_k a_k(n)$ 将近似 $\sum_k b_k(n)$ ，所以

让我们计算

$$\begin{aligned}\sum_k b_k(n) &= \frac{e^{2n+1/2}}{e^\sigma \sqrt{n}} \sum_k e^{k^2/n} \\ &= \frac{2^{2n+1/2}}{e^\sigma \sqrt{n}} \Theta_n = \frac{2^{2n} \sqrt{2\pi}}{e^\sigma} (1 + O(n^{-M})).\end{aligned}$$

(另一个侥幸: 我们将用到前面例子的和 Θ_n .) 这是令人鼓舞的, 因为我们知道原来和实际是

$$A_n = \sum_k \binom{2n}{k} = (1+1)^{2n} = 2^{2n}.$$

所以看来好像我们将得到 $e^\sigma = \sqrt{2\pi}$.

但是有蹊跷: 仍需证明我们的估计足够好. 所以让我们首先考虑由 $c_k(n)$ 提供的误差:

$$\sum_{|k| \leq n^{1/2+\varepsilon}} c_k(n) = \sum_{|k| \leq n^{1/2+\varepsilon}} 2^{2n} n^{-1+3\varepsilon} e^{-k^2/n} \leq 2^{2n} n^{-1+3\varepsilon} \Theta_n = O(2^{2n} n^{-\frac{1}{2}+3\varepsilon}).$$

如果 $3\varepsilon < (1/2)$, 此式渐近地小于前面的和.

接着我们一定要检查尾部. 我们得到

$$\begin{aligned}\sum_{k > n^{1/2+\varepsilon}} e^{-k^2/n} &< \exp(-\lfloor n^{1/2+\varepsilon} \rfloor^2 / n) (1 + e^{-1/n} + e^{-2/n} + \cdots) \\ &= O(e^{-n^{2\varepsilon}}) \cdot O(n),\end{aligned}$$

对于所有 M , 它为 $O(n^{-M})$, 所以 $\sum_{k \notin D_n} b_k(n)$ 渐近地忽略不计. (我们刚好选在 $n^{1/2+\varepsilon}$ 处切断, 以致在 D_n 的外部将指数律地小. 像 $n^{1/2} \log n$ 的其他选取也将足够好, 产生的估计稍微准确一点, 但是公式将出现更复杂. 我们不需作最强的可能估计, 因为我们的主要目的是建立常数 σ 的值.) 另一个尾部

$$\sum_{k > n^{1/2+\varepsilon}} \binom{2n}{n+k}$$

的界限为 $2n$ 乘它的最大项, 此项在切断点 $k \approx n^{1/2+\varepsilon}$ 处出现. 知道此项近似 $b_k(n)$, 与 A_n 相比它是指数律地小; 一个指数律小的乘数消除了 $2n$ 的因子.

于是我们成功地用交换尾部诀窍证明了估计

$$2^{2n} = \sum_k \binom{2n}{k} = \frac{\sqrt{2\pi}}{e^\sigma} 2^{2n} + O(2^{2n} n^{-\frac{1}{2}+3\varepsilon}), \text{ 如果 } 0 < \varepsilon < \frac{1}{6}. \quad (9.99)$$

我们可选 $\varepsilon = \frac{1}{8}$, 推出

$$\sigma = \frac{1}{2} \ln 2\pi$$

证毕。

习 题

准备部分

1. 证明或推翻: 如果 $f_1(n) < g_1(n)$ 和 $f_2(n) < g_2(n)$, 则我们得到 $f_1(n) + f_2(n) < g_1(n) + g_2(n)$.
2. 哪个函数增长较快:
 - (a) $n^{(\ln n)}$ 或 $(\ln n)^n$?
 - (b) $n^{(\ln \ln \ln n)}$ 或 $(\ln n)!$?
 - (c) $(n!)!$ 或 $((n-1)!)!(n-1)!^n$?
 - (d) $F_{\lceil H_n \rceil}^2$ 或 H_{F_n} ?
3. 下列论证的错误是什么? “由于 $n = O(n)$ 和 $2n = O(n)$, 等等, 我们得到 $\sum_{k=1}^n kn = \sum_{k=1}^n O(n) = O(n^2)$ ”.
4. 给出一个记号 O 在左边而不在右边的成立的方程的例子 (不用乘零的诀窍, 那是太容易了。) 提示: 考虑取极限。
5. 证明或推翻: 如果 $f(n)$ 和 $g(n)$ 对所有 n 为正, 则 $O(f(n) + g(n)) = f(n) + O(g(n))$. (与式 (9.27) 比较.)
6. $(\ln n + \gamma + O(1/n))$ 乘以 $(n + O(\sqrt{n}))$, 且用记号 O 来表达你的解答。
7. 估计 $\sum_{k=0}^{\infty} e^{-k/n}$, 具有绝对误差 $O(n^{-1})$.

基本部分

8. 给出一个函数 $f(n)$ 和 $g(n)$ 的例子, 使得三种关系 $f(n) < g(n)$, $f(n) > g(n)$, $f(n) \asymp g(n)$ 都不成立, 虽然 $f(n)$ 和 $g(n)$ 都单调上升到 ∞ .
9. 依据 O 的函数集的定义, 通过表明左边是右边的子集来严格证明式 (9.22).
10. 证明或推翻: 对所有实数 $x > 0$, $\cos O(1 + f(n))^2 = O(f(n))^x$.

11. 证明或推翻: $O(x+y)^2 = O(x^2) + O(y^2)$.

12. 证明当 $n \rightarrow \infty$ 时,

$$1 + \frac{2}{n} + O(n^{-2}) = (1 + \frac{2}{n})(1 + O(n^{-2})).$$

13. 计算 $(n + 2 + O(n^{-1}))^n$, 具有相对误差 $O(n^{-1})$.

14. 证明 $(n + \alpha)^{n+\beta} = n^{n+\beta} e^{\alpha} (1 + \alpha(\beta - (1/2)\alpha)n^{-1} + O(n^{-2}))$.

15. 给出“中间的”三项系数 $\binom{3n}{n, n, n}$ 的一个渐近公式, 准确到相对误差 $O(n^{-3})$.

16. 证明如果 $0 < x < 1/2$, $B(1-x) = -B(x) \geq 0$, 我们得到

$$\int_a^b B(\{x\})f(x)dx \geq 0$$

若我们还设对 $a \leq x \leq b$, $f'(x) \geq 0$.

17. 对所有 $m \geq 0$, 用母函数证明 $B_m(1/2) = (2^{1-m} - 1)B_m$.

18. 当 $x > 0$ 时, 求 $\sum_k \binom{2n}{k}^x$, 具有相对误差 $O(n^{-1/4})$.

课外习题

19. 当 $n = 10$, $z = \alpha = 0.1$ 和 $O(f(n)) = O(f(z)) = 0$ 时, 用计算机来比较表 9.1 中近似的左边和右边.

20. 证明或推翻下列估计, 当 $n \rightarrow \infty$ 时:

$$(a) O\left(\left(\frac{n^2}{\log \log n}\right)^{1/2}\right) = O(\lfloor \sqrt{n} \rfloor^2).$$

$$(b) e^{(1+O(1/n))^2} = e + O(1/n).$$

$$(c) n! = O(((1-1/n)^n n)^n).$$

21. 方程 (9.48) 给出第 n 个素数, 具有相对误差 $O(\log n)^{-2}$. 由式 (9.31) 的另一项开始, 在式 (9.46) 中把相对误差改进到 $O(\log n)^{-3}$.

22. 把式 (9.54) 改进到 $O(n^{-3})$.

23. 进一步推进近似式 (9.62), 取得绝对误差 $O(n^{-3})$.

提示: 设 $g_n = c/(n+1)(n+2) + h_n$, h_n 所满足的递归是什么?

24. 假设 $a_n = O(f(n))$ 和 $b_n = O(f(n))$. 证明或推翻卷积 $\sum_{k=0}^n a_k b_{n-k}$ 在下列情形也是 $O(f(n))$:

$$(a) f(n) = n^{-\alpha}, \alpha > 1.$$

$$(b) f(n) = \alpha^{-n}, \alpha > 1.$$

25. 就本章所阐述的内容来证明式 (9.1) 和 (9.2).

26. 方程 (9.91) 表明了如何来计算 $\ln 10!$, 具有绝对误差 $< 1/126\,000\,000$. 所以如果取指数, 将得到 $10!$, 具有小于 $e^{1/126\,000\,000} - 1 < 10^{-8}$ 的相对误差. (事实上, 近似给出 $3\,628\,799.971\,4$.) 如果现在四舍五入到最近整数, 知道 $10!$ 是一个整数, 我们得到一个精确的结果. 如果计算 Stirling 近似的足够多项, 是否总可能以相似的方式计算 $n!$? 当 n 是一个指定的 (大的) 整数时, 估计 m 的值给出 $\ln n!$ 的最好近似. 与 $n!$ 一起比较这个近似中的绝对误差.

27. 用 Euler 求和公式找 $H_n^{(-\alpha)} = \sum_{k=1}^n k^{-\alpha}$ 的渐近值, 其中 α 是任何指定的实数. (你的解答可能涉及到你不知道闭形式的一个常数.)

28. 习题 5.13 定义了超阶乘函数 $Q_n = 1^1 2^2 \cdots n^n$. 求 Q_n 的渐近值, 具有相对误差 $O(n^{-1})$. (你的解答可能涉及到一个你不知道闭形式的常数.)

29. 如同前一习题那样估计函数 $1^{1/1} 2^{1/2} \cdots n^{1/n}$.

30. 当 l 是指定非负的整数时, 求 $\sum_{k \geq 0} k^l e^{-k^2/n}$ 的渐近值, 具有绝对误差 $O(n^{-1})$.

31. 当 $c > 1$ 和 m 是正整数时, 计算 $\sum_{k \geq 0} 1/(c^k + c^m)$, 具有绝对误差 $O(c^{-3m})$.

考查性问题

32. 计算 $e^{H_n + H_n^{(2)}}$, 具有绝对误差 $O(n^{-1})$.

33. 计算 $\sum_{k \geq 0} \binom{n}{k} / n^k$, 具有绝对误差 $O(n^{-3})$.

34. 确定值 A 到 F , 使得 $(1 + 1/n)^{nH_n}$ 是

$$An + B(\ln n)^2 + C \ln n + D + \frac{E(\ln n)^2}{n} + \frac{F \ln n}{n} + O(n^{-1}).$$

35. 计算 $\sum_{k=1}^n 1/kH_k$, 具有绝对误差 $O(1)$.

36. 计算 $S_n = \sum_{k=1}^n 1/(n^2 + k^2)$, 具有绝对误差 $O(n^{-5})$.

37. 计算 $\sum_{k=1}^n (n \bmod k)$, 具有绝对误差 $O(n \log n)$.

38. 计算 $\sum_{k \geq 0} k^k \binom{n}{k}$, 具有相对误差 $O(n^{-1})$.

39. 计算 $\sum_{0 \leq k < n} \ln(n-k)(\ln n)^k / k!$, 具有绝对误差 $O(n^{-1})$.

提示: 证明 $k \geq 10 \ln n$ 的项可忽略.

40. 设 m 是 (指定的) 正整数. 计算 $\sum_{k=1}^n (-1)^k H_k^m$, 具有绝对误差 $O(1)$.

41. 计算“Fibonacci阶乘” $\prod_{k=1}^n F_k$, 具有相对误差 $O(n^{-1})$ 或更好. (你的解答可能涉及到你不知道闭形式的一个常数.)

42. 设 α 是范围 $0 < \alpha < 1/2$ 中的一个常数. 在前几章中我们看到没有和 $\sum_{k \leq \alpha n} \binom{n}{k}$ 的一般闭形式. 证明有一个渐近公式

$$\sum_{k \leq \alpha n} \binom{n}{k} = 2^{nH(\alpha) - \frac{1}{2} \lg n + O(1)},$$

其中 $H(\alpha) = \alpha \lg \frac{1}{\alpha} + (1 - \alpha) \lg \left(\frac{1}{1 - \alpha} \right)$. 提示: 对于 $0 < k \leq \alpha n$, 证明 $\binom{n}{k-1} < \frac{\alpha}{1 - \alpha} \binom{n}{k}$.

43. 证明兑换 n 分的方式数(第七章中考虑的那样) C_n 渐近地为 $cn^4 + O(n^3)$ (对某个常数 c). 这个常数是什么?

44. 证明当 $x \rightarrow \infty$ 时,

$$x^{\frac{1}{2}} = x^{\frac{1}{2}} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} - x^{-1/2} \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} + x^{-3/2} \begin{bmatrix} 1/2 \\ -3/2 \end{bmatrix} + O(x^{-5/2}).$$

(记住式(5.98)中的定义 $x^{\frac{1}{2}} = x! / (x-1/2)!$, 以及表 6.8 中的广义 Stirling 数的定义.)

45. 设 α 是 0 和 1 间的无理数. 第三章讨论了量 $D(\alpha, n)$, 它测量了分数部分 $\{k\alpha\}$ ($0 \leq k < n$) 偏离均匀分布的最大偏差. 递归

$$D(\alpha, n) \leq D(\{\alpha^{-1}\}, \lfloor xn \rfloor) + \alpha^{-1} + 2$$

在式(3.31)中已证明了; 我们还有一个明显的界限

$$0 \leq D(\alpha, n) \leq n.$$

证明 $\lim_{n \rightarrow \infty} D(\alpha, n)/n = 0$. 提示: 第六章讨论了连分式.

46. 证明习题 7.15 的 Bell 数 $b_n = e^{-1} \sum_{k \geq 0} k^n / k!$ 渐近地等于

$$m(n)^n e^{m(n) - n - 1/2} / \sqrt{\ln n},$$

其中 $m(n) \ln m(n) = n - 1/2$, 且估计此近似的相对误差.

47. 设 m 是一个 ≥ 2 的整数. 分析两个和

$$\sum_{k=1}^n \lfloor \log_m k \rfloor \quad \text{和} \quad \sum_{k=1}^n \lceil \log_m n \rceil,$$

哪一个渐近地接近 $\log_m n!$?

48. 考虑以十进记法的调和数 H_k ($1 \leq k \leq n$) 的表, 第 k 项 \hat{H}_k 准确地四舍五入到 d_k 个有意义的数字位, 其中 d_k 恰好是大到足以把此值和 H_{k-1} 和 H_{k+1} 区别. 例如, 这里是取自表中的 5 项, 其中 H_k 超越了 10:

k	H_k	\hat{H}_k	d_k
12 364	9.999 800 41-	9.999 8	5
12 365	9.999 881 28+	9.999 9	5
12 366	9.999 962 15-	9.999 96	6
12 367	10.000 043 01-	10.000 0	6
12 368	10.000 123 86+	10.000 1	6

估计表中全部数字位的个数, $\sum_{k=1}^n d_k$ 具有绝对误差 $O(n)$.

49. 在第六章中我们考虑了 $n(s)$ 后达到橡皮带的终端的蠕行的事, 其中 $H_{n-1} < 100 \leq H_n$. 证明如果 n 是正整数使得 $H_{n-1} \leq \alpha \leq H_n$, 则

$$\lfloor e^{\alpha-1} \rfloor \leq n \leq \lceil e^{\alpha-1} \rceil.$$

50. 给硅谷的冒险的资本家提供了一种机会, 对他们的投资给以指数的支付: 对于一项 n ($n \geq 2$) 百万元的投資, 企业财团允诺在一年之后付清 N 百万元, 其中 $N = 10^n$. 当然有一些风险; 实际分配是企业财团以概率 $1/(k^2 H_n^{(2)})$ 支付 k 百万元 (对于范围 $1 \leq k \leq N$ 中的每个整数 k). (所有支付是以一百万元支付, 也就是说, 以 \$1 000 000 的确切的倍数支付, 由一个真正的随机过程确定支付。)注意到, 一个投资者总至少取回一百万元。

(a) 如果投資 n 百万元, 一年后渐近期望返回的是什么? (换句话说, 支付的均值是什么?) 你的解答应正确到绝对误差 $O(10^{-n})$ 元之内。

(b) 如果你投資 n 百万元, 你获利的渐近概率是什么? (换句话说, 你取回的比投入的多的机会是多少?) 你的解答应正确到绝对误差 $O(n^{-3})$ 元之内。

额外问题

51. 证明或推翻: 当 $n \rightarrow \infty$ 时, $\int_n^\infty O(x^{-2})dx = O(x^{-1})$.

52. 证明存在一个幂级数 $A(z) = \sum_{k \geq 0} a_k z^k$, 对所有复数 z 收敛, 使得

$$A(n) \geq \left\{ n^{n^{n^{\cdot^{\cdot^{\cdot^n}}}}} \right\}_n.$$

53. 证明如果 $f(x)$ 是这样的函数, 它的导数对于所有 $x \geq 0$ 满足

$$f'(x) \leq 0, -f''(x) \leq 0, f'''(x) \leq 0, \dots, (-1)^m f^{(m+1)}(x) \leq 0.$$

则我们得到

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \dots + \frac{f^{(m-1)}(0)}{(m-1)!}x^{m-1} + O(x^m), \text{ 对 } x \geq 0.$$

特别, 对于所有 $k, n > 0$, 情形 $f(x) = -\ln(1+x)$ 证明了式(9.64).

54. 设 $f(x)$ 是正的, 可微分的函数使得当 $x \rightarrow \infty$ 时 $xf'(x) < f(x)$. 证明

$$\sum_{k \geq n} \frac{f(k)}{1+k} = O\left(\frac{f(n)}{n^\alpha}\right), \text{ 如果 } \alpha > 0.$$

提示: 考虑量 $f(k-1/2)/(k-1/2)^\alpha - f(k+1/2)/(k+1/2)^\alpha$.

55. 把式(9.99)改进到相对误差 $O(n^{-3/2+5\epsilon})$.

56. 在许多算法分析中出现量 $Q(n) = 1 + ((n-1)/n) + ((n-1)/n)((n-2)/n) + \dots = \sum_{k \geq 1} n^k / n^k$. 求它的渐近值, 具有绝对误差 $O(1)$.

57. 在式(9.54)中推出 Golomb 和 $\sum_{k \geq 1} 1/k \lfloor 1 + \log_n k \rfloor^2$ 的渐近公式. 求无下整括号的相似和 $\sum_{k \geq 1} 1/k(1 + \log_n k)^2$ 的渐近公式. 提示: 我们有 $\int_0^\infty u e^{-uk} k^{-u} du = 1/(1 + i \ln k)^2$.

58. 用留数计算, 在方形围路 $z = x + iy$ 上积分

$$\frac{1}{2\pi i} \oint \frac{2\pi i e^{2\pi i z}}{e^{2\pi i z} - 1} z^m dz$$

其中, $\max(|x|, |y|) = M + 1/2$, 设整数 M 趋于 ∞ . 来证明

$$B_m(\{x\}) = -2 \frac{m!}{(2\pi)^m} \sum_{k \geq 1} \frac{\cos(2\pi kx - \frac{1}{2}\pi m)}{k^m}, \quad (m \geq 2).$$

59. 设 $\Theta_n(t) = \sum_k e^{-(k+t)^2/n}$, t 的一个周期函数, 证明把 $\Theta_n(t)$ 展开成傅里叶级数为

$$\begin{aligned} \Theta_n(t) = & \sqrt{\pi n} (1 + 2e^{-\pi^2 n} (\cos 2\pi t) + 2e^{-4\pi^2 n} (\cos 4\pi t) \\ & + 2e^{-9\pi^2 n} (\cos 6\pi t) + \dots). \end{aligned}$$

(对于方程(9.93)中的和 $\Theta_n = \Theta_n(0)$, 此公式给出一个快速收敛的级数.)

60. 说明为什么渐近展式

$$\binom{2n}{n} = \frac{4^n}{\sqrt{\pi n}} \left(1 - \frac{1}{8n} + \frac{1}{128n^2} + \frac{5}{1024n^3} - \frac{21}{32768n^4} + O(n^{-5}) \right)$$

中的系数都有 2 的幂的分母。

61. 习题 45 证明了对于所有无理数 α , 偏差 $D(\alpha, n)$ 是 $o(n)$. 展示一个无理数 α 使得对任何 $\varepsilon > 0$, $D(\alpha, n)$ 不是 $O(n^{1-\varepsilon})$.

62. 给定 n , 设 $\left\{ \binom{n}{m(n)} \right\} = \max_k \left\{ \binom{n}{k} \right\}$ 是 Stirling 子集三角形的行 n 中的最大项. 证明对于所有充分大的 n , 我们有 $m(n) = \lfloor \bar{m}(n) \rfloor$ 或 $m(n) = \lceil \bar{m}(n) \rceil$, 其中 $\bar{m}(n)(\bar{m}(n) + 2)\ln(\bar{m}(n) + 2) = n(\bar{m}(n) + 1)$.

提示: 这是困难的.

63. 证明习题 2.36 的 S.W.Golomb 的描述的序列满足

$$f(n) = \varphi^{2-\varphi} n^{\varphi-1} + O(n^{\varphi-1} / \log n).$$

64. 仅用“欧拉”(18 世纪)数学找等式

$$\sum_{n \geq 1} \frac{\cos 2n\pi x}{n^2} = \pi^2 \left(x^2 - x + \frac{1}{6} \right) \quad \text{对 } 0 \leq x \leq 1$$

的一种证明.

研究性问题

65. 找 Stirling 近似的一种“组合的”证明。(注意, n^n 是 $\{1, 2, \dots, n\}$ 映射入它自身的映射数, $n!$ 是 $\{1, 2, \dots, n\}$ 映到它自身的映射数.)

66. 考虑一个 $n \times n$ 点组 ($n \geq 3$), 每个点有 4 个邻近点。(在边界处我们对 n 模“环绕”。) 设 χ_n 是以这样一种方式把红色、白色和蓝色赋予这些点的方式数, 无相邻点具有相同颜色。(因此 $\chi_3 = 12$.) 证明

$$\chi_n \sim \left(\frac{4}{3} \right)^{3n^2/2} e^{-\pi/6}.$$

67. 设 Q_n 是最小整数 m 使得 $H_m > n$. 求最小整数 n 使得 $Q_n \neq \lfloor e^{n-1} + 1/2 \rfloor$, 或证明没有这样的 n 存在.

附录 A 习题解答

这里对每一习题都作了解答(至少是简略解答),有些解答还超出了所问的问题,读者若通过自己认真地求解,然后再参看本附录将会收到很好的效果。

关于研究性问题的任何解法(或部分解法)或解非研究性问题的任何更简便的(或更恰当的)方法,请告知作者。

1.1 除了当 $n=2$ 外,证明是好的,如果两匹马的所有集合具有相同颜色的马,则对于任何马数,命题是真的。

1.2 如果 X_n 是移动次数,我们有 $X_0=0$ 和 $X_n=X_{n-1}+1+X_{n-1}+1+X_{n-1}$ ($n>0$)。由此得出(例如通过两边加 1) $X_n=3^n-1$ 。(($1/2$) X_n 次移动后,结果整个塔将在中间杆上,中间处!)

1.3 有 3^n 种可能的排列,因为每个圆盘能在任何一个杆上,我们一定要碰上它们的全部,因为最短的解用了 3^n-1 次移动。(这个构造等价于一个“三元 Gray 码”,它从 $(0\cdots 0)_3$ 到 $(2\cdots 2)_3$ 贯穿所有数,一次仅改变一个数字位。)

1.4 没有。如果最大的圆盘不必移动,(由归纳法证明) $2^{n-1}-1$ 次移动就够了,否则需 $(2^{n-1}-1)+1+(2^{n-1}-1)$ 次(由归纳法证明)。

1.5 不能;不同的圆至多交于两点,所以第 4 个圆能增加区域数至多 14 个。然而用卵形做这次工作是可能的:



Venn^[294]断言用椭圆没有办法来做 5 个集合的情形,但是 Grünbaum^[137]发现了用椭圆的一种 5 个集合的构造。

1.6 如果第 n 条直线和前面的直线相交于 $k>0$ 个不同的点,我们得到 $k-1$ 个新的有界域(假设前面的直线相互不平行)和 2 个新的无界域。因此有界域的最大个数是 $(n-2)+(n-3)+\cdots=S_{n-2}=(n-1)(n-2)/2=L_n-2n$ 。

1.7 基础是未被证明的,事实上, $H(1)\neq 2$ 。

1.8 $Q_2=(1+\beta)/\alpha$; $Q_3=(1+\alpha+\beta)/\alpha\beta$; $Q_4=(1+\alpha)/\beta$; $Q_5=\alpha$; $Q_6=\beta$ 。所以序列是周期的!

1.9 (a) 从不等式

$$x_1 \cdots x_{n-1} \left(\frac{x_1 + \cdots + x_{n-1}}{n-1} \right) \leq \left(\frac{x_1 + \cdots + x_{n-1}}{n-1} \right)^n$$

我们取得 $P(n-1)$ 。

(b) 依据 $P(n)$, $x_1 \cdots x_n x_{n+1} \cdots x_{2n} \leq (((x_1 + \cdots + x_n)/n)((x_{n+1} + \cdots + x_{2n})/n))^n$; 依据 $P(2)$, 乘积内部 $\leq ((x_1 + \cdots + x_{2n})/2n)^2$ 。

(c) 例如, 从 $P(2)$ 推得 $P(4)$, 再推得 $P(3)$, 再推得 $P(6)$, 再推得 $P(5)$ 。

1.10 首先证明当 $n > 0$ 时 $R_n = R_{n-1} + 1 + Q_{n-1} + 1 + R_{n-1}$ 。顺便提到, 第七章的方法将告诉我们 $Q_n = ((1+\sqrt{3})^{n+1} - (1-\sqrt{3})^{n+1}) / (2\sqrt{3}) - 1$ 。

1.11 (a) 我们不能比移加倍的 $(n-1)$ 塔, 然后移两个最大圆盘(且转换两个最大圆盘的次序), 然后再移加倍的 $(n-1)$ 塔做得更好; 因此 $A_n = 2A_{n-1} + 2$ 和 $A_n = 2T_n = 2^{n+1} - 2$ 。这个解交换了两个最大圆盘, 但是把其他 $2n-2$ 个圆盘返回到它们原来的次序。

(b) 设 B_n 是最小移动次数。于是 $B_1 = 3$, 且当 $n > 1$ 时能证没有比 $B_n = A_{n-1} + 2 + A_{n-1} + 2 + B_{n-1}$ 更好的对策。因此对所有 $n > 0$, $B_n = 2^{n+2} - 5$ 。这不寻常的就是 $2A_n - 1$, 且我们还有 $B_n = A_{n-1} + 1 + A_{n-1} + 1 + A_{n-1} + 1 + A_{n-1}$ 。

1.12 如果所有 $m_k > 0$, 则 $A(m_1, \cdots, m_n) = 2A(m_1, \cdots, m_{n-1}) + m_n$, 这是“广义的 Josephus”型的一个方程, 具有解 $(m_1, \cdots, m_n)_2 = 2^{n-1}m_1 + \cdots + 2m_{n-1} + m_n$ 。

顺便提到, 习题 11(b) 的对应的推广呈现满足递归

$$B(m_1, \cdots, m_n) = \begin{cases} A(m_1, \cdots, m_n), & \text{如果 } m_n = 1; \\ 2m_n - 1, & \text{如果 } n = 1; \\ 2A(m_1, \cdots, m_{n-1}) + 2m_n \\ \quad + B(m_1, \cdots, m_{n-1}), & \text{如果 } n > 1 \text{ 和 } m_n > 1. \end{cases}$$

1.13 给定确定 L_n 个区域的 n 条直线, 我们能用充分长线段的极狭的锯齿形线条替换它们, 每对锯齿形线条之间有 9 处相交。这表明 $ZZ_n = ZZ_{n-1} + 9n - 8$ (对所有 $n > 0$); 因而 $ZZ_n = 9S_n - 8n + 1 = (9/2)n^2 - (7/2)n + 1$ 。

1.14 每次新切割确定的新 3 维区域数是由新平面和前面平面相交而在新平面中确定的 2 维区域数。因此 $P_n = P_{n-1} + L_{n-1}$, 结果 $P_5 = 26$ 。(一块立方的乳酪中 6 次切割能切成 27 块小立方体, 或者最多到 $P_6 = 42$ 份古怪的形状。)

顺便提到, 此递归的解符合于一种很好的型式, 如果我们用二项系数(见第五章)来表达, 则有:

$$X_n = \binom{n}{0} + \binom{n}{1};$$

$$L_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2}$$

$$P_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3}.$$

这里 X_n 是由一直线上的 n 个点所确定的 1 维区域的最大数.

1.15 当 $n > 1$ 时, 函数 I 满足 J 的相同递归, 但是 $I(1)$ 未定义. 由于 $I(2) = 2$ 和 $I(3) = 1$, 没有 $I(1) = \alpha$ 的值将允许我们用一般的方法; “末端游戏”的展开依赖于 n 的二进制表示中的前两位.

如果 $n = 2^m + 2^{m-1} + k$, 其中 $0 \leq k < 2^{m+1} + 2^m - (2^m + 2^{m-1}) = 2^m + 2^{m-1}$, 解为 $I(n) = 2k + 1$ (所有 $n > 2$). 采用表示 $n = 2^m + l$, 另一种方法可表达这一点为

$$I(n) = \begin{cases} J(n) + 2^{m-1}, & \text{如果 } 0 \leq l < 2^{m-1}; \\ J(n) - 2^m, & \text{如果 } 2^{m-1} \leq l < 2^m. \end{cases}$$

1.16 设 $g(n) = a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 + d(n)\gamma$. 当 $n = (1b_{m-1} \cdots b_1 b_0)_2$ 时, 我们从式 (1.18) 知道 $a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 = (\alpha\beta_{b_{m-1}} \beta_{b_{m-2}} \cdots \beta_{b_1} \beta_{b_0})_3$; 这就确定了 $a(n)$, $b(n)$ 和 $c(n)$. 在递归中置 $g(n) = n$ 意味着 $a(n) + c(n) - d(n) = n$, 因此我们知道了一切. [置 $g(n) = 1$ 给出附加的等式 $a(n) - 2b(n) - 2c(n) = 1$, 利用较简单的函数 $a(n)$ 和 $a(n) + c(n)$ 能用它来确定 $b(n)$.]

1.17 我们一般有 $W_m \leq 2W_{m-k} + T_k$ ($0 \leq k \leq m$). (此关系对应于转移顶端 $n-k$, 然后仅用三个杆来移底部 k , 然后用顶端的 $n-k$ 来结束.) 当 $m = n(n+1)/2$ 时, 所说的关系是基于唯一的 k 值, 它使这个一般不等式的右边减至最小. (然而, 我们不能推得等式成立, 转移塔的许多其他策略是可能的.) 如果置 $Y_n = (W_{n(n+1)/2} - 1)/2^n$, 我们找到 $Y_n \leq Y_{n-1} + 1$, 因此 $W_{n(n+1)/2} \leq 2^n(n-1) + 1$.

1.18 证明从 $(n^{2k}, 0)$ 起的两条线和从 $(n^{2j}, 0)$ 起的两条线相交就够了, 且所有这些交点是不同的.

从 $(x_j, 0)$ 起通过 $(x_j - a_j, 1)$ 的一条线和从 $(x_k, 0)$ 起通过 $(x_k - a_k, 1)$ 的一条线相交于点 $(x_j - ta_j, t)$, 其中 $t = (x_k - x_j)/(a_k - a_j)$. 设 $x_j = n^{2j}$ 和 $a_j = n^j + (0 \text{ 或 } n^{-n})$. 则比 $t = (n^{2k} - n^{2j}) / (n^k - n^j + (-n^{-n} \text{ 或 } 0 \text{ 或 } n^{-n}))$ 严格地位于 $n^j + n^k - 1$ 和 $n^j + n^k + 1$ 之间, 因此交点的 y 坐标唯一地确定 j 和 k . 具有相同 j 和 k 的 4 个交点也是不同的.

1.19 当 $n > 11$ 时不能. 从一条弯线的顶开始, 它的部分线在角 θ 和 $\theta + 30^\circ$ 处变, 另一条弯线的部分线在角 φ 和 $\varphi + 30^\circ$ 处变, 仅当 $|\theta - \varphi| > 30^\circ$ 时这两条弯线能相交 4 次. 我们不能选取多于 11 个角使彼此相距这样远. (可能选 11 吗?)

1.20 设 $h(n) = a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 + d(n)\gamma_0 + e(n)\gamma_1$. 当 $n = (1b_{m-1} \cdots b_1 b_0)_2$ 时, 我们从式 (1.18) 知道 $a(n)\alpha + b(n)\beta_0 + c(n)\beta_1 = (\alpha\beta_{b_{m-1}} \beta_{b_{m-2}} \cdots \beta_{b_1} \beta_{b_0})_4$, 这就确定了 $a(n)$, $b(n)$ 和 $c(n)$. 在递归中置 $h(n) = n$ 意味着 $a(n) + c(n) - 2d(n) - 2e(n) = n$; 置 $h(n) = n^2$ 意味着 $a(n) + c(n) + 4e(n) = n^2$. 因此 $d(n) = (3a(n) + 3c(n) - n^2 - 2n)/4$; $e(n) = (n^2$

$-a(n) - c(n))/4$.

1.21 我们可设 m 是 $2n, 2n-1, \dots, n+1$ 的最小(或任何)公倍数。(不严格的论证提出 m 的一个“随机”值将以概率

$$\frac{n}{2n} \frac{n-1}{2n-1} \cdots \frac{1}{n+1} = 1/\binom{2n}{n} \sim \frac{\sqrt{\pi n}}{4^n}$$

接替, 所以我们可期望找到小于 4^n 的这样一个 m .)

1.22 取一个具有 2^n 个边的正多边形, 用长度为 2^n 的“de Bruijn 循环”的元素来标记边。(这是 0 和 1 的循环序列, 所有相邻元素的 n 元组是不同的, 见[173, 习题 2.3.4.2-23]和[174, 习题 3.2.2-17]。在标记 1 的每个边上附加一个很细小的凸部分。 n 个集合是转动 k 个边的长度($k=0, 1, \dots, n-1$)所产生的多边形的拷贝。

1.23 能。(我们需要第四章的初等数论原理。)设 $L(n)=\text{lcm}(1, 2, \dots, n)$, 我们能假设 $n>2$; 因此由 Bertrand 假定, 在 $n/2$ 和 n 之间有一个素数 p 。我们也能假设 $j>n/2$ 。由于 $q'=L(n)+1-q$ 剩下 $j'=n+1-j$ 当且仅当 q 剩下 j 。选取 q 以致 $q \equiv 1 \pmod{L(n)/p}$ 和 $q \equiv j+1-n \pmod{p}$ 。现在以次序 $1, 2, \dots, n-p, j+1, j+2, \dots, n, n-p+1, \dots, j-1$ 移动人。

1.24 仅知道的例子是: $X_n = a/X_{n-1}$, 它有周期 2; 习题 8 中周期 5 的 R.C.Lyness 递归; 具有周期 8 的 H.Todd 递归 $X_n = (1+X_{n-1}+X_{n-2})/X_{n-3}$; 通过形式 $Y_n = \alpha X_{m_n}$ 的替换能从这些递归导出递归。当 $k=2$ 时, 通过 Bill Gosper 的详尽查找发现无周期 4 的非平凡解。Lyness^[210]和 Kurshan 以及 Gopinath^[189]建立了部分理论。当开始值为实数时, 周期为 9 的另一类型的有趣例子是 Morton Brown^[38]发现的递归 $X_n = |X_{n-1} - X_{n-2}|$ 。任何希望周期 ≥ 5 的非线性递归是基于延拓的^[55]。

1.25 如果 $T^{(k)}(n)$ 表示具有 k 个辅助杆转移 n 个圆盘所需的最小移动次数 (因此 $T^{(1)}(n)=T_n$ 和 $T^{(2)}(n)=W_n$), 我们有 $T^{(k)}\left(\binom{n+1}{k}\right) \leq 2T^{(k)}\left(\binom{n}{k}\right) + T^{(k-1)}\left(\binom{n}{k-1}\right)$ 。还不知道 (n, k) 使此不等式不能是等式的例子。当 k 比 n 小时, 公式 $2^{n+1-k}\binom{n-1}{k-1}$ 给出 $T^{(k)}\left(\binom{n}{k}\right)$ 的一个合适的(但不是最优的)上界。

1.26 对于所有 m 和 n , 用 $O(n \log n)$ 步能计算执行次序的排列^[175, 习题 5.1.1-2 和 5.1.1-5]。每当 $n \equiv 0 \pmod{3}$ 以及 $n \geq 9$ 时, Bjorn Poonen^[241]证明了存在恰好有 4 个“坏人”的非 Josephus 集合; 事实上, 对于某个 $\varepsilon > 0$, 这样的集合数至少为 $\varepsilon \binom{n}{4}$ 。通过进一步计算他还发现具有非 Josephus 集合的另一个 ($n < 24$) 仅为 $n=20$, 关于 $k=14$, 它有 236 个这样的集合, 关于 $k=13$, 有两个(一个是 $\{1, 2, 3, 4, 5, 6, 7, 8, 11, 14, 15, 16, 17\}$; 另一个是关于 21 的它的反射)。对于 $n=15$ 和 $k=9$, 有一个唯一的非 Josephus 集合, 即 $\{3, 4, 5, 6, 8, 10, 11, 12, 13\}$ 。

2.1 没有同意这样记法; 三种回答是正当的: (1) 我们能说 $\sum_{k=m}^n q_k$ 总等价于

$\sum_{n \leq k \leq m} q_k$; 则所说的和为零。(2) 一个人可能说, 在 k 的下降值上求和给出和为 $q_4 + q_3 + q_2$

$+q_1+q_0$ 。但是这与一般接受的约定当 $n=0$ 时 $\sum_{k=1}^n q_k = 0$ 相冲突。(3) 我们能说 $\sum_{k=m}^n q_k = \sum_{k \leq n} q_k - \sum_{k < m} q_k$; 则所说的和为 $-q_1 - q_2 - q_3$ 。这个约定看来不熟悉, 但是它服从有用的定律 $\sum_{k=a}^b + \sum_{k=b+1}^c = \sum_{k=a}^c$ (对所有 a, b, c)。

最好仅当 $n-m \geq -1$ 时用表示法 $\sum_{k=m}^n$, 则两种约定(1)和(3)一致。

2.2 这是 $|x|$, 顺便提到, 常称量 $([x > 0] - [x < 0])$ 为 $\text{Sign}(x)$ 或 $\text{Signum}(x)$; 当 $x > 0$ 时它为 $+1$, 当 $x = 0$ 时为 0 , 当 $x < 0$ 时为 -1 。

2.3 第一个和当然是 $a_0 + a_1 + a_2 + a_3 + a_4 + a_5$; 第二个和是 $a_4 + a_1 + a_0 + a_1 + a_4$, 因为和是在值 $k \in \{-2, -1, 0, +1, +2\}$ 上的。交换律这里不成立, 因为函数 $p(k) = k^2$ 不是一种排列, 有些 n 值(譬如说 $n=3$)没有 k 使得 $p(k)=n$; 其他(譬如说, $n=4$)有两个这种的 k 。

$$2.4 \quad (a) \sum_{i=1}^4 \sum_{j=i+1}^4 \sum_{k=j+1}^4 a_{ijk} = \sum_{i=1}^2 \sum_{j=i+1}^3 \sum_{k=j+1}^4 a_{ijk} = ((a_{123} + a_{124}) + a_{134}) + a_{234}.$$

$$(b) \sum_{k=1}^4 \sum_{j=1}^{k-1} \sum_{i=1}^{j-1} a_{ijk} = \sum_{k=3}^4 \sum_{j=2}^{k-1} \sum_{i=1}^{j-1} a_{ijk} = a_{123} + (a_{124} + (a_{134} + a_{234})).$$

2.5 相同的指标 ' k ' 被用作两个不同的指标变量, 虽然 k 限制在内和中, 这是在数学(和计算机程序设计)中的出名的错误。如果 $a_j = a_k$ (对所有 j 和 k , $1 \leq j, k \leq n$), 结果是正确的。

2.6 它是 $[1 \leq j \leq n](n-j+1)$ 。在这里第一个因子是必要的, 因为当 $j < 1$ 或 $j > n$ 时我们应取零。

2.7 mx^{m-1} 。基于 ∇ 而不是 Δ 的有限演算的形式将突出升阶乘幂。

2.8 如果 $m \geq 1$, 为 0 ; 如果 $m \leq 0$, 为 $1/|m|!$ 。

2.9 对于整数 m 和 n , $x^{\overline{m+n}} = x^{\overline{m}}(x+n)^{\overline{n}}$ 。置 $m = -n$, 告诉我们 $x^{\overline{-n}} = 1/(x-n)^{\overline{n}} = 1/(x-1)^{\overline{n}}$ 。

2.10 另一个可能的右边为 $Eu\Delta v + v\Delta u$ 。

2.11 把左边分成两个和, 且在第二个和中把 k 改成 $k+1$ 。

2.12 如果 $p(k)=n$, 则 $n+c = k + ((-1)^k + 1)c$ 和 $((-1)^k + 1)$ 是偶数; 因此, $(-1)^{n+c} = (-1)^k$ 和 $k = n - (-1)^{n+c}c$ 。反之, 这个 k 的值产生 $p(k)=n$ 。

2.13 设 $R_0 = \alpha$, 且对于 $n > 0$, $R_n = R_{n-1} + (-1)^n(\beta + n\gamma + n^2\delta)$, 则 $R(n) = A(n)\alpha + B(n)\beta + C(n)\gamma + D(n)\delta$ 。置 $R_n = 1$ 产生 $A(n) = 1$ 。置 $R_n = (-1)^n$ 产生 $A(n) + 2B(n) = (-1)^n$ 。置 $R_n = (-1)^n n$ 产生 $-B(n) + 2C(n) = (-1)^n n$ 。置 $R_n = (-1)^n n^2$ 产生 $B(n) - 2C(n) + 2D(n) = (-1)^n n^2$ 。所以 $2D(n) = (-1)^n(n^2 + n)$, 所说的和是 $D(n)$ 。

2.14 因为当 $1 \leq k \leq n$ 时我们有 $k = \sum_{1 \leq j \leq k} 1$, 提出的改写是合法的。首先对 k 求和; 多重和化为

$$\sum_{1 \leq j \leq n} (2^{n+1} - 2^j) = n2^{n+1} - (2^{n+1} - 2).$$

2.15 第一步用 $2\sum_{1 \leq j \leq k} j$ 替换 $k(k+1)$. 第二步给出 $\binom{n}{2} + \square_n = \left(\sum_{k=1}^n k\right)^2 + \square_n$.

2.16 依据式(2.52), $x^{\frac{m}{n}}(x-m)^{\frac{n}{n}} = x^{\frac{m+n}{n}} = x^{\frac{n}{n}}(x-n)^{\frac{m}{n}}$.

2.17 对于前两个用归纳法, 对于第三个用式(2.52). 根据第一行得到第二行.

2.18 用下列事实, $(\mathcal{R}z)^+ \leq |z|$, $(\mathcal{R}z)^- \leq |z|$, $(\mathcal{I}z)^+ \leq |z|$, $(\mathcal{I}z)^- \leq |z|$, 以及 $|z| \leq (\mathcal{R}z)^+ + (\mathcal{R}z)^- + (\mathcal{I}z)^+ + (\mathcal{I}z)^-$.

2.19 两边乘 $2^{n-1}/n!$ 且设 $S_n = 2^n T_n / n! = S_{n-1} + 3 \cdot 2^{n-1} = 3(2^n - 1) + S_0$. 解是 $T_n = 3 \cdot n! + n! / 2^{n-1}$. (在第四章中我们将看出仅当 n 是 0 或 2 的幂时 T_n 是整数.)

2.20 摄动法给出

$$S_n + (n+1)H_{n+1} = S_n + \left(\sum_{0 \leq k \leq n} H_k\right) + n+1.$$

2.21 抽出 S_{n+1} 的最后项给出 $S_{n+1} = 1 - S_n$; 抽出第一项给出

$$\begin{aligned} S_{n+1} &= (-1)^{n+1} + \sum_{1 \leq k \leq n+1} (-1)^{n+1-k} = (-1)^{n+1} + \sum_{0 \leq k \leq n} (-1)^{n-k} \\ &= (-1)^{n+1} + S_n. \end{aligned}$$

因此 $2S_n = 1 + (-1)^n$, 我们得到 $S_n = [n \text{ 是偶}]$, 我们相似地发现

$$T_{n+1} = n+1 - T_n = \sum_{k=0}^n (-1)^{n-k} (k+1) = T_n + S_n,$$

因此 $2T_n = n+1 - S_n$. 我们得到 $T_n = (1/2)(n + [n \text{ 是奇}])$. 最后, 相同的方式产生

$$\begin{aligned} U_{n+1} &= (n+1)^2 - U_n = U_n + 2T_n + S_n \\ &= U_n + n + [n \text{ 是奇}] + [n \text{ 是偶}] \\ &= U_n + n + 1. \end{aligned}$$

因此 U_n 是三角形数 $(1/2)(n+1)n$.

2.22 两倍和给出 $1 \leq j, k \leq n$ 上的一个和, 它分成 3 个容易处理的和.

2.23 (a) 此方式给出计算 $2n + H_n - 2n + (H_n + (1/(n+1) - 1))$ 的 4 个和. (用 $1/k + 1/(k+1)$ 替换被加数比较容易.)

(b) 设 $u(x) = 2x + 1$ 和 $\Delta v(x) = 1/x(x+1) = (x-1)^{-2}$, 则 $\Delta u(x) = 2$ 和 $v(x) = -(x-1)^{-1} = -1/x$. 解答是 $2H_n + n/(n+1)$.

2.24 分部求和, $\sum x^m H_x \delta x = x^{\frac{m+1}{2}} H_x / (m+1) - x^{\frac{m+1}{2}} / (m+1)^2 + C$, 因此

$\sum_{0 \leq k < n} k^m H_k = n^{\frac{m+1}{2}} (H_n - 1 / (m+1)) / (m+1) + O(n^{\frac{m+1}{2}} / (m+1)^2)$. 此时 $m = -2$, 所以和达到 $1 - (H_n + 1) / (n+1)$.

2.25 这里是一些基本的相似结果:

$$\begin{aligned} \sum_{k \in K} c a_k &= c \sum_{k \in K} a_k & \longleftrightarrow & \prod_{k \in K} a_k^c = (\prod_{k \in K} a_k)^c \\ \sum_{k \in K} (a_k + b_k) &= \sum_{k \in K} a_k + \sum_{k \in K} b_k & \longleftrightarrow & \prod_{k \in K} (a_k b_k) = (\prod_{k \in K} a_k) (\prod_{k \in K} b_k) \\ \sum_{k \in K} a_k &= \sum_{p(k) \in K} a_{p(k)} & \longleftrightarrow & \prod_{k \in K} a_k = \prod_{p(k) \in K} a_{p(k)} \\ \sum_{\substack{j \in J \\ k \in K}} a_{j,k} &= \sum_{j \in J} \sum_{k \in K} a_{j,k} & \longleftrightarrow & \prod_{\substack{j \in J \\ k \in K}} a_{j,k} = \prod_{j \in J} \prod_{k \in K} a_{j,k} \\ \sum_{k \in K} a_k &= \sum_k a_k [k \in K] & \longleftrightarrow & \prod_{k \in K} a_k = \prod_k a_k^{[k \in K]} \\ \sum_{k \in K} 1 &= \#K & \longleftrightarrow & \prod_{k \in K} c = c^{\#K} \end{aligned}$$

2.26 $P^2 = (\prod_{1 \leq i, k \leq n} a_i a_k) (\prod_{1 \leq j, k \leq n} a_j a_k)$. 第一个因子是 $(\prod_{k=1}^n a_k^n)^2$, 第二个因子是 $\prod_{k=1}^n a_k^2$. 因此 $P = (\prod_{k=1}^n a_k)^{n+1}$.

2.27 $\Lambda(c^{\frac{x}{2}}) = c^{\frac{x}{2}} (c - x - 1) = c^{\frac{x+2}{2}} / (c - x)$. 置 $c = -2$, 且 x 下降 2 产生 $\Lambda(-(-2)^{\frac{x-2}{2}}) = (-2)^{\frac{x}{2}} / x$, 因此所说的和是 $(-2)^{\frac{-1}{2}} - (-2)^{\frac{n-1}{2}} = (-1)^n n! - 1$.

2.28 第二行和第三行之间和的交换不是合理的, 此和的项不绝对收敛. 除了 $\sum_{k \geq 1} [k = j-1] k / j$ 的结果也许该写成 $[j-1 \geq 1] (j-1) / j$, 其他一切是完全正确的, 且明确地简化.

2.29 用部分分式取得

$$\frac{k}{4k^2 - 1} = \frac{1}{4} \left(\frac{1}{2k+1} + \frac{1}{2k-1} \right).$$

$(-1)^k$ 因子现在使每项的两半和它们的相邻的项相消. 因此解答为 $-1/4 + (-1)^n / (8n+4)$.

2.30 $\sum_a^b x dx = (1/2)(b^2 - a^2) = (1/2)(b-a)(b+a-1)$, 所以我们有

$$(b-a)(b+a-1) = 2100 = 2^2 \cdot 3 \cdot 5^2 \cdot 7.$$

对于写 $2100 = x \cdot y$ (其中 x 是偶, y 是奇) 的每种方法有一个解, 我们设 $a = (1/2)|x-y| + 1/2$ 和 $b = (1/2)(x+y) + 1/2$. 所以解的个数是 $3 \cdot 5^2 \cdot 7$ 的因子的个数, 即 12, 一般有 $\prod_{p \geq 2} (n_p + 1)$ 种方式表示 $\prod_p p^{n_p}$, 其中乘积是在素数范围上.

2.31 $\sum_{j,k \geq 2} j^{-k} = \sum_{j \geq 2} 1/j^2 (1 - 1/j) = \sum_{j \geq 2} 1/j(j-1)$. 第二个和相似地为 $3/4$.

2.32 如果 $2n \leq x < 2n+1$, 和是 $0 + \dots + n + (x-n-1) + \dots + (x-2n) = n(x-n) = (x-1) + (x-3) + \dots + (x-2n+1)$. 如果 $2n-1 \leq x < 2n$, 它们相似地都等于 $n(x-n)$. (提前看第三章, 公式 $\lfloor (1/2)(x+1) \rfloor \lfloor (x - \lfloor (1/2)(x+1) \rfloor) \rfloor$ 包括两种情形.)

2.33 如果 K 是空的, $\bigwedge_{k \in K} a_k = \infty$. 基本的相似结果是:

$$\sum_{k \in K} c a_k = c \sum_{k \in K} a_k \quad \longleftrightarrow \quad \bigwedge_{k \in K} (c + a_k) = c + \bigwedge_{k \in K} a_k$$

$$\sum_{k \in K} (a_k + b_k) = \sum_{k \in K} a_k + \sum_{k \in K} b_k \quad \longleftrightarrow \quad \bigwedge_{k \in K} \min(a_k, b_k) = \min(\bigwedge_{k \in K} a_k, \bigwedge_{k \in K} b_k)$$

$$\sum_{k \in K} a_k = \sum_{p(k) \in K} a_{p(k)} \quad \longleftrightarrow \quad \bigwedge_{k \in K} a_k = \bigwedge_{p(k) \in K} a_{p(k)}$$

$$\sum_{\substack{j \in J \\ k \in K}} a_{j,k} = \sum_{j \in J} \sum_{k \in K} a_{j,k} \quad \longleftrightarrow \quad \bigwedge_{\substack{j \in J \\ k \in K}} a_{j,k} = \bigwedge_{j \in J} \bigwedge_{k \in K} a_{j,k}$$

$$\sum_{k \in K} a_k = \sum_k a_k [k \in K] \quad \longleftrightarrow \quad \bigwedge_{k \in K} a_k = \bigwedge_k a_k \cdot \infty^{(k \notin K)}$$

2.34 设 $K^+ = \{k | a_k \geq 0\}$ 和 $K^- = \{k | a_k < 0\}$, 于是如果 n 为奇, 我们选取 F_n 是 $F_{n-1} \cup E_n$, 其中 $E_n \subseteq K^-$ 是充分大, $\sum_{k \in (F_{n-1} \cap K^+)} a_k - \sum_{k \in E_n} (-a_k) < A^-$.

2.35 能如下证明 Goldbach 的和等于

$$\sum_{m,n \geq 2} m^{-n} = \sum_{m \geq 2} \frac{1}{m(m-1)} = 1.$$

不求几何级数的和, 它等于 $\sum_{k \in P, l \geq 1} k^{-l}$; 所以如果我们能找到有序对 $(m, n) (m, n \geq 2)$ 和有序对 $(k, l) (k \in P, l \geq 1)$ 当对应时 $m^n = k^l$ 之间的一一对应, 则将完成证明. 如果 $m \notin P$, 我们设 $(m, n) \longleftrightarrow (m^n, 1)$; 但是如果 $m = a^b \in P$, 我们设 $(m, n) \longleftrightarrow (a^n, b)$.

2.36 (a) 依据定义, $g(n) - g(n-1) = f(n)$.

(b) 依据 (a), $g(g(n)) - g(g(n-1)) = \sum_k f(k) [g(n-1) < k \leq g(n)] = n(g(n) - g(n-1)) = n f(n)$.

(c) 再依据 (a),

$g(g(g(n))) = g(g(g(n-1)))$ 是

$$\begin{aligned} & \sum_k f(k)[g(g(n-1)) < k \leq g(g(n))] \\ &= \sum_{j,k} j[j = f(k)][g(g(n-1)) < k \leq g(g(n))] \\ &= \sum_{k,k} j[j = f(k)][g(n-1) < j \leq g(n)] \\ &= \sum_j j(g(j) - g(j-1))[g(n-1) < j \leq g(n)] \\ &= \sum_j j f(j)[g(n-1) < j \leq g(n)] = n \sum_j j[g(n-1) < j \leq g(n)]. \end{aligned}$$

Colin Mallows 看出序列还能用递归

$$f(1) = 1, f(n+1) = 1 + f(n+1 - f(f(n))), \text{ (对 } n \geq 0 \text{)}$$

来定义。

2.37 作者 RLG 想它们可能不能装填；作者 DEK 想它们可能装填；作者 OP 未提出他的想法。

$$3.1 \quad m = \lfloor \lg n \rfloor; \quad l = n - 2^m = n - 2^{\lfloor \lg n \rfloor}.$$

$$3.2 \quad (a) \lfloor x+0.5 \rfloor; \quad (b) \lceil x-0.5 \rceil.$$

$$3.3 \quad \text{由于 } 0 < \{mx\} < 1, \text{ 这是 } \lfloor mn - \{m\alpha\}n/\alpha \rfloor.$$

$$3.4 \quad \text{不要求证明, 仅为侥幸的猜中.}$$

$$3.5 \quad \text{我们有 } \lfloor nx \rfloor = n \lfloor x \rfloor \Leftrightarrow n \lfloor x \rfloor \leq \lfloor nx \rfloor < n \lfloor x \rfloor + 1 \Leftrightarrow n \lfloor x \rfloor \leq nx < n \lfloor x \rfloor + 1 \Leftrightarrow nx - n \{x\} \leq nx < nx - n \{x\} + 1, \text{ 依据式 (3.5(a)), (3.7(a)), (3.7(d)) 和 (3.8), 当 } n \text{ 是正整数时, 这等价于 } n \{x\} < 1. \text{ (此时对于所有 } x, n \lfloor x \rfloor \leq \lfloor nx \rfloor \text{).)}$$

$$3.6 \quad \lfloor f(x) \rfloor = \lfloor f(\lceil x \rceil) \rfloor.$$

$$3.7 \quad \lfloor n/m \rfloor + n \bmod m.$$

$$3.8 \quad \text{如果所有匣子包含 } < \lceil n/m \rceil \text{ 个物体, 则 } n \leq (\lceil n/m \rceil - 1)m, \text{ 所以 } n/m + 1 \leq \lceil n/m \rceil \text{ 与式 (3.5) 相矛盾. 其他证明是相似的.}$$

$$3.9 \quad \text{我们有 } m/n - 1/q = (n \text{ mumble } m)/qn, \text{ 过程一定终止, 因为 } 0 \leq n \text{ mumble } m < m. \text{ 表达式的分母严格上升, 因此是不同的, 因为 } qn/(n \text{ mumble } m) > q.$$

$$3.10 \quad \text{如果 } \{x\} \neq 1/2, \lceil x + (1/2) \rceil - \lfloor (2x+1)/4 \rfloor \text{ 不是一个整数是 } x \text{ 的最接近的整数; 否则它是最接近的偶整数 (见习题 2). 因此公式给出一种“无偏”方式来四舍五入.}$$

$$3.11 \quad \text{如果 } n \text{ 是整数, } x < n < \beta \Leftrightarrow \lfloor \alpha \rfloor < n < \lceil \beta \rceil. \text{ 当 } a \text{ 和 } b \text{ 是整数时满足 } a < n < b \text{ 的整数的个数是 } (b-a-1)(b > a). \text{ 如果 } \alpha = \beta = \text{整数, 我们将取得错误解答.}$$

$$3.12 \quad \text{依据式 (3.6), 从两边减 } \lfloor n/m \rfloor, \text{ 取得 } \lceil (n \bmod m)/m \rceil = \lfloor (n \bmod m + m - 1)/n \rfloor. \text{ 由于 } 0 \leq n \bmod m < n, \text{ 两边现在都等于 } \lfloor n \bmod m > 0 \rfloor.$$

一个较短而不太直接的简单证明看到, 式(3.24)中的第一项一定等于式(3.25)中的最后

一项.

3.13 如果它们形成一个划分, 书中的 $N(\alpha, n)$ 的公式意味着 $1/\alpha + 1/\beta = 1$, 因为方程 $N(\alpha, N) + N(\beta, N) = n$ 中的 n 的系数一定相同, 如果对大的 n 方程成立. 因此 α 和 β 都为有理数或都为无理数. 如果两者为无理数, 正如书中表明的那样, 我们的确取得一个划分. 如果两者能写成具有分子 m , 值 $m-1$ 不出现在两个范围中. (然而, 当 $1/\alpha + 1/\beta = 1$ 时, Golomb^[121]看到集合 $\{\lfloor n\alpha \rfloor | n \geq 1\}$ 和 $\{\lceil n\beta \rceil - 1 | n \geq 1\}$ 总形成一个划分.)

3.14 如果 $ny = 0$, 是显然的, 否则依据式(3.21)和(3.6)为真的.

3.15 在式(3.24)中把 $\lceil mx \rceil$ 代入 n : $\lceil mx \rceil = \lceil x \rceil + \lceil x - (1/m) \rceil + \cdots + \lceil x - (m-1)/m \rceil$.

3.16 当 $0 \leq n < 3$ 时, 通过检验能证实公式 $n \bmod 3 = 1 + (1/3)((\omega-1)\omega^n - (\omega+2)\omega^{2n})$.

当 m 是任何正整数时, $n \bmod m$ 的一般公式出现在习题 7.25 中.

$$\begin{aligned} 3.17 \quad \sum_{j,k} [0 \leq k < m] [1 \leq j \leq x + k/m] &= \sum_{j,k} [0 \leq k < m] [1 \leq j \leq \lceil x \rceil] [k \geq m(j-x)] \\ &= \sum_{1 \leq j \leq \lceil x \rceil} \sum_k [0 \leq k < m] - \sum_{j=\lceil x \rceil} \sum_k [0 \leq k < m(j-x)] = m\lceil x \rceil - \lceil m(\lceil x \rceil - x) \rceil = -\lceil -mx \rceil \\ &= \lfloor mx \rfloor. \end{aligned}$$

3.18 我们有

$$S = \sum_{0 \leq j < \lfloor n\alpha \rfloor} \sum_{\lceil j\alpha^{-1} \rceil \leq k < (j+v)\alpha^{-1}} 1.$$

如果 $j \leq n\alpha - 1 \leq n\alpha - v$, 没有贡献, 因为 $(j+v)\alpha^{-1} \leq n$. 因此 $j = \lfloor n\alpha \rfloor$ 是这样的仅有的情形, 此时值等于 $\lceil (\lfloor n\alpha \rfloor + v)\alpha^{-1} \rceil - n \leq \lceil v\alpha^{-1} \rceil$.

3.19 当且仅当 b 是整数. (如果 b 是整数, $\log_b x$ 是连续上升函数, 它仅在整数点取整数值. 如果 b 不是整数, 则当 $x=b$ 时条件失效.)

$$3.20 \quad \sum_k kx[\alpha \leq kx \leq \beta] = x \sum_k k[\lceil \alpha/x \rceil \leq k \leq \lfloor \beta/x \rfloor], \text{ 这共计 } (1/2)x(\lfloor \beta/x \rfloor + \lceil \alpha/x \rceil - \lceil \alpha/x \rceil \lceil \alpha/x - 1 \rceil).$$

3.21 如果 $10^n \leq 2^M < 10^{n+1}$, 恰好有 $n+1$ 个这样的 2 的幂, 因为对于每个 n 恰好有一个 n 个数字位的 2 的幂. 所以解答是 $1 + \lfloor M \log 2 \rfloor$.

注意: 当 $l > 1$ 时具有首数字位 l 的 2 的幂的个数是较困难的, 它是 $\sum_{0 \leq n \leq M} (\lfloor n \log 2 - \log l \rfloor - \lfloor n \log 2 - \log(l+1) \rfloor)$.

3.22 除了第 k 个外, 对于 n 和 $n-1$ 的所有项是相同的, 其中 $n = 2^{k-1}q$, q 是奇数; 我们有 $S_n = S_{n-1} + 1$, 和 $T_n = T_{n-1} + 2^k q$. 因此 $S_n = n$ 和 $T_n = n(n+1)$.

$$3.23 \quad X_n = m \Leftrightarrow (1/2)m(m-1) < n \leq (1/2)m(m+1) \Leftrightarrow m^2 - m + (1/4) < 2n < m^2 - m + (1/4) \Leftrightarrow m - (1/2) < \sqrt{2n} < m + (1/2).$$

3.24 设 $\beta = \alpha/(\alpha+1)$. 于是在 $\text{Spec}(\beta)$ 中出现非负整数 m 的次数恰好比 $\text{Spec}(\alpha)$ 中出现 m 的次数多 1. 为什么? 因为 $N(\beta, n) = N(\alpha, n) + n + 1$.

3.25 继续课文中的推导. 如果能找到 m 的一个值使得 $K_m \leq m$, 则当 $n=2m+1$ 时, 在 $n+1$ 处我们能违背所述的不等式. (当 $n=3m+1$ 和 $n=3m+2$ 时也是如此.) 但是存在要求 $2K_{\lfloor n'/2 \rfloor} \leq n'$ 或 $3K_{\lfloor n'/3 \rfloor} \leq n'$ 即,

$$K_{\lfloor n'/2 \rfloor} \leq \lfloor n'/2 \rfloor \text{ 或 } K_{\lfloor n'/3 \rfloor} \leq \lfloor n'/3 \rfloor$$

的这样的一个 $m=n'+1$. 这样一步一步下去, 意味着 $K_0 \leq 0$, 但是 $K_0 = 1$.

我们实际要证明的是对于所有 $n > 0$, K_n 严格大于 n . 事实上, 用归纳法容易证明这一点, 虽然它是一个比我们不能证明的结果的较强结论!

(由这个习题可认识到一个重要的经验, 与其说一个下整数函数性质的习题不如说一个归纳法特性的习题.)

3.26 归纳法, 用较强假设

$$D_n^{(q)} \leq (q-1) \left(\left(\frac{q}{q-1} \right)^{n+1} - 1 \right), \text{ 对 } n \geq 0.$$

3.27 如果 $D_n^{(3)} = 2^m b - a$, 其中 b 是奇数, a 是 0 或 1, 则 $D_{n+b}^{(3)} = 3^m b - a$.

3.28 关键观察是 $a_n = m^2$ 意味着 $a_{n+2k+1} = (m+k)^2 + m - k$ 和 $a_{n+2k+2} = (m+k)^2 + 2m$ ($0 \leq k \leq m$), 因此 $a_{n+2m+1} = (2m)^2$. 可把解写成 Carl Witty 发现的一种优美的形式:

$$a_{n-1} = 2^l + \left\lfloor \left(\frac{n-l}{2} \right)^2 \right\rfloor, \text{ 当 } 2^l + l \leq n < 2^{l+1} + l + 1 \text{ 时}.$$

3.29 $D(\alpha', \lfloor n\alpha \rfloor)$ 至多是

$$s(\alpha', \lfloor n\alpha \rfloor, v') = -s(\alpha, n, v) + S - \varepsilon - [0 \text{ 或 } 1] - v' + [0 \text{ 或 } 1]$$

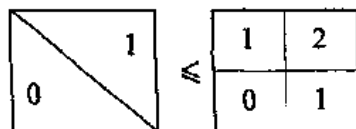
的右边的最大值.

3.30 依据归纳法, $X_n = \alpha^{2^n} + \alpha^{-2^n}$, X_n 是一个整数.

3.31 这里是一个“优美的”, “印象深刻的”证明, 没有给出如何发现它的思路:

$$\begin{aligned} \lfloor x \rfloor + \lfloor y \rfloor + \lfloor x+y \rfloor &= \lfloor x + \lfloor y \rfloor \rfloor + \lfloor x+y \rfloor \\ &\leq \left\lfloor x + \frac{1}{2} \lfloor 2y \rfloor \right\rfloor + \left\lfloor x + \frac{1}{2} \lfloor 2y \rfloor + \frac{1}{2} \right\rfloor \\ &= \lfloor 2x + \lfloor 2y \rfloor \rfloor = \lfloor 2x \rfloor + \lfloor 2y \rfloor. \end{aligned}$$

但是还有一种依据观察的简单的图示证明, 我们仅需考虑情形 $0 \leq x, y < 1$. 于是在平面中函数看来像:



一个稍强的结果是可能的, 即

$$\lceil x \rceil + \lfloor y \rfloor + \lfloor x + y \rfloor \leq \lceil 2x \rceil + \lfloor 2y \rfloor;$$

但是仅当 $\{x\} = 1/2$ 时, 这是较强的。如果在此等式中用 $(-x, x+y)$ 替换 (x, y) , 且应用反射律(3.4), 我们得到

$$\lfloor y \rfloor + \lfloor x + y \rfloor + \lfloor 2x \rfloor \leq \lfloor x \rfloor + \lfloor 2x + 2y \rfloor.$$

3.32 设 $f(x)$ 是问题中的和, 由于 $f(x) = f(-x)$, 我们可假设 $x \geq 0$. 当 $k \rightarrow -\infty$ 时, 2^k 为项的界限, 当 $k \rightarrow +\infty$ 时, $x^2 / 2^k$ 为项的界限, 所以对所有实数 x , 和存在.

我们有 $f(2x) = 2 \sum_k 2^{k-1} \|x / 2^{k-1}\|^2 = 2f(x)$. 设 $f(x) = l(x) + r(x)$, 其中 $l(x)$ 是 $k \leq 0$ 的和, $r(x)$ 是 $k > 0$ 的和, 则 $l(x+1) = l(x)$, 且对于所有 x , $l(x) \leq 1/2$. 当 $0 \leq x < 1$ 时, 我们有 $r(x) = x^2 / 2 + x^2 / 4 + \dots = x^2$ 和 $r(x+1) = (x-1)^2 / 2 + (x+1)^2 / 4 + (x+1)^2 / 8 + \dots = x^2 + 1$. 因此当 $0 \leq x < 1$ 时, $f(x+1) = f(x) + 1$.

现在我们用归纳法证明当 $0 \leq x < 1$ 时, 对于所有整数 $n \geq 0$, $f(x+n) = f(x) + n$. 特别, $f(n) = n$. 所以一般, $f(x) = 2^{-m} f(2^m x) = 2^{-m} \lfloor 2^m x \rfloor + 2^{-m} f(\{2^m x\})$. 但是 $f(\{2^m x\}) = l(\{2^m x\}) + r(\{2^m x\}) \leq (1/2) + 1$, 所以对所有的 m , $|f(x) - x| \leq |2^{-m} \lfloor 2^m x \rfloor - x| + 2^{-m} \cdot (3/2) \leq 2^{-m} \cdot (5/2)$.

必然发生的结论是对所有实数 x , $f(x) = |x|$.

3.33 设 $r = n - (1/2)$ 是圆的半径.

(a) 在棋盘的单元中间有 $2n-1$ 条水平线和 $2n-1$ 条垂直线, 且圆和每条这样的线相交两次. 由于 r^2 不是整数, Pythagorean 定理告诉我们, 圆不通过任何单元的角. 因此圆通过有交点的单元一样多的单元, 即 $8n-4 = 8r$. (相同的公式给出棋盘边处的单元数.)

$$(b) f(n, k) = 4 \lfloor \sqrt{r^2 - k^2} \rfloor.$$

根据(a)和(b)得到

$$\frac{1}{4} \pi r^2 - 2r \leq \sum_{0 < k < r} \lfloor \sqrt{r^2 - k^2} \rfloor \leq \frac{1}{4} \pi r^2, \quad r = n - \frac{1}{2}.$$

获得此和的更精确估计的工作是数论中的一个著名问题, Gauss 和其他许多人进行了研究, 见 Dickson^[65, 第2卷, 第6章].

3.34 (a) 设 $m = \lceil \lg n \rceil$. 我们能加上 $2^m - n$ 项来简化边界处的计算:

$$\begin{aligned} f(n) + (2^m - n)m &= \sum_{k=1}^{2^m} \lceil \lg k \rceil = \sum_{j,k} j \lfloor j \rfloor = \lceil \lg k \rceil [1 \leq k \leq 2^m] \\ &= \sum_{j,k} j [2^{j-1} < k \leq 2^j] [1 \leq j \leq m] \\ &= \sum_{j=1}^m j 2^{j-1} = 2^m (m-1) + 1. \end{aligned}$$

结果 $f(n) = nm - 2^m + 1$.

(b) 我们有 $\lceil n/2 \rceil = \lfloor (n+1)/2 \rfloor$, 由此得到一般递归 $g(n) = a(n) + g(\lceil n/2 \rceil) + g(\lfloor n/2 \rfloor)$ 的解一定满足 $\Delta g(n) = \Delta a(n) + \Delta g(\lfloor n/2 \rfloor)$. 特别, 当 $a(n) = n-1$ 时, n 的二进制表示中的位数, 即 $\lceil \lg(n+1) \rceil$ 满足 $\Delta f(n) = 1 + \Delta f(\lfloor n/2 \rfloor)$. 现在从 Δ 转换到 Σ .

一个更直接的解可依据等式 $\lceil \lg 2j \rceil = \lceil \lg j \rceil + 1$ 和 $\lceil \lg(2j-1) \rceil = \lceil \lg j \rceil + \lfloor j > 1 \rfloor$, (对于 $j \geq 1$) 而得到.

$$3.35 \quad (n+1)^2 n! \cdot e = A_n + (n+1)^2 + (n+1) + B_n.$$

其中 $A_n = \frac{(n+1)^2 n!}{0!} + \frac{(n+1)^2 n!}{1!} + \cdots + \frac{(n+1)^2 n!}{(n-1)!}$ 是 n 的倍数,

$$\begin{aligned} B_n &= \frac{(n+1)^2 n!}{(n+2)!} + \frac{(n+1)^2 n!}{(n+3)!} + \cdots \\ &= \frac{(n+1)}{(n+2)} \left(1 + \frac{1}{n+3} + \frac{1}{(n+3)(n+4)} + \cdots \right) \\ &< \frac{n+1}{n+2} \left(1 + \frac{1}{n+3} + \frac{1}{(n+3)(n+3)} + \cdots \right) \\ &= \frac{(n+1)(n+3)}{(n+2)^2} < 1. \end{aligned}$$

因此解答是 $2 \bmod n$.

3.36 和是

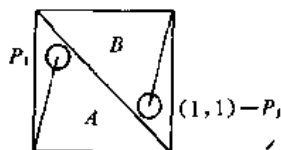
$$\begin{aligned} &\sum_{k, l, m} 2^{-l} 4^{-m} [m = \lfloor \lg l \rfloor] [l = \lfloor \lg k \rfloor] [1 < k < 2^{2^m}] \\ &= \sum_{k, l, m} 2^{-l} 4^{-m} [2^m \leq l < 2^{m+1}] [2^l \leq k < 2^{l+1}] [0 \leq m < n] \\ &= \sum_{l, m} 4^{-m} [2^m \leq l < 2^{m+1}] [0 \leq m < n] \\ &= \sum_m 2^{-m} [0 \leq m < n] = 2(1 - 2^{-n}). \end{aligned}$$

3.37 首先考虑情形 $m < n$, 依据是否 $m < (1/2)n$, 把它分成子情形; 然后表明当 m 依据 n 增加时, 两边以相同方式改变.

3.38 至多一个 x_k 可是非整数. 丢弃所有整数 x_k , 且假设 n 是留下的, 当 $\{x\} \neq 0$ 时, 当 $m \rightarrow \infty$ 时, $\{mx\}$ 的平均位于 $1/4$ 和 $1/2$ 之间; 因此当 $n > 1$ 时, $\{mx_1\} + \cdots + \{mx_n\} - \{mx_1 + \cdots + mx_n\}$ 不能有平均值零.

但是刚才给出的论证依赖于均匀分布的一个困难的定理. 一个初等的证明是可能的, 对 $n=2$, 概略地叙述如下: 设 P_m 是点 $(\{mx\}, \{my\})$. 把单位正方形 $0 \leq x, y < 1$ 分成依照 $x+y < 1$ 或 $x+y \geq 1$ 的三角形区域 A 和 B . 我们要证明如果 $\{x\}$ 和 $\{y\}$ 是非零, 则对某个 m , $P_m \in B$. 如果 $P_1 \in B$, 则我们完成了. 否则有一个圆心在 P_1 , 半径 $\varepsilon > 0$ 的圆盘 D 使得 $D \subseteq A$. 根据 Dirichlet 匣子原理, 如果 N 足够大, 序列 P_1, \dots, P_N 一定包含具有

$|P_k - P_j| < \varepsilon$ 的两个点 ($k > j$).



由此得出 P_{k-j-1} 在 $(1, 1) - P_1$ 的 ε 之内, 因此 $P_{k-j-1} \in B$.

3.39 用 $b-j$ 替换 j 且把项 $j=0$ 加入和, 以致对于 j 的和能用习题 15. 结果, 当对 k 求和时, 插进

$$\lceil x/b^k \rceil - \lceil x/b^{k+1} \rceil + b - 1.$$

3.40 设 $\lfloor 2\sqrt{n} \rfloor = 4k+r$, 其中 $-2 \leq r < 2$, 且设 $m = \lfloor \sqrt{n} \rfloor$. 于是用归纳法能证下列关系:

部分	r	m	x	y	当且仅当
W_k	-2	$2k-1$	$m(m+1) - n - k$	k	$(2k-1)(2k-1) \leq n \leq (2k-1)(2k)$
S_k	-1	$2k-1$	$-k$	$m(m+1) - n + k$	$(2k-1)(2k) < n < (2k)(2k)$
E_k	0	$2k$	$n - m(m+1) + k$	$-k$	$(2k)(2k) \leq n \leq (2k)(2k+1)$
N_k	1	$2k$	k	$n - m(m+1) - k$	$(2k)(2k+1) < n < (2k+1)(2k+1)$

因此, 当 $k \geq 1$ 时, W_k 是长度为 $2k$ 的一部分, 在那里路径行进西部且 $y(n) = k$; S_k 是长度为 $2k-2$ 的一部分, 在那里路径行进南部且 $x(n) = -k$; 等等. (a) 所示希望的公式是

$$y(n) = (-1)^m ((n - m(m+1)) \cdot [\lfloor 2\sqrt{n} \rfloor \text{ 是奇}] - \left\lceil \frac{1}{2} m \right\rceil).$$

(b) 在所有部分上, $k = \max(|x(n)|, |y(n)|)$. 在部分 W_k 和 S_k 上, 我们有 $x < y$ 和 $n+x+y = m(m+1) = (2k)^2 - 2k$. 在 E_k 和 N_k 上, 我们有 $x \geq y$ 和 $n-x-y = m(m+1) = (2k)^2 + 2k$. 因此正负号是 $(-1)^{(x(n) < y(n))}$.

3.41 由于 $1/\varphi + 1/\varphi^2 = 1$, 所述序列的确划分正整数. 由于条件 $g(n) = f(f(n)) + 1$ 唯一确定 f 和 g , 我们仅需证明对所有 $n > 0$, $\lfloor \lfloor n\varphi \rfloor \varphi \rfloor + 1 = \lfloor n\varphi^2 \rfloor$. 根据习题 3 得到这一点 ($\alpha = \varphi$, $n = 1$).

3.42 不存在, 一个论证像书中两个范围情形的分析以及在习题 13 中证明一个三划分出现当且仅当 $1/\alpha + 1/\beta + 1/\gamma = 1$ 和

$$\left\{ \frac{n+1}{\alpha} \right\} + \left\{ \frac{n+1}{\beta} \right\} + \left\{ \frac{n+1}{\gamma} \right\} = 1, \text{ (对所有 } n > 0 \text{)}.$$

但是依据均匀分布的定理, 如果 α 是无理数, 则 $\{(n+1)/\alpha\}$ 的平均值是 $1/2$. 参数不能全为有理数, 且若 $\gamma = m/n$, 平均是 $3/2 - 1/(2n)$. 因此 γ 一定是整数, 但是这也行不通. (有一种不可能性的证明, 仅用简单的原理, 不用均匀分布的定理, 见 [125].)

3.43 展开 K_n 的递归一步给出 4 个数 $1+a+a \cdot b \cdot K_{\lfloor (n-1-a)/(a \cdot b) \rfloor}$, 其中 a 和 b 每个为 2 或 3. (这种简化涉及式(3.11)的一种应用, 它与等式 $x+\min(y, z)=\min(x+y, x+z)$ 一起移去下整之内的下整. 我们一定要略去具有负下标的项, 即具有 $n-1-a < 0$ 的项.)

现在沿着这样的路线继续下去, 导致下列解释: K_n 是形式

$$1 + a_1 + a_1 a_2 + a_1 a_2 a_3 + \cdots + a_1 a_2 a_3 \cdots a_m$$

的所有数的多重集 S 中的最小数 $> n$, 其中 $m \geq 0$, 每个 a_k 是 2 或 3. 因此,

$$S = \{1, 3, 4, 7, 9, 10, 13, 15, 19, 21, 22, 27, 28, 31, 31, \dots\};$$

数 31 在 S 中出现“两次”, 因为它有两种表示 $1+2+4+8+16=1+3+9+18$.

(Michael Fredman^[108]证明了 $\lim_{n \rightarrow \infty} K_n / n = 1$, 即 S 没有巨大的间隙.)

3.44 设 $d_n^{(q)} = D_{n-1}^{(q)} \bmod (q-1)$, 以致 $D_n^{(q)} = (qD_{n-1}^{(q)} + d_n^{(q)}) / (q-1)$ 和 $a_n^{(q)} = \lceil D_{n-1}^{(q)} / (q-1) \rceil$. 现在 $D_{k-1}^{(q)} \leq (q-1)n \Leftrightarrow a_k^{(q)} \leq n$, 且得到结果. (这是 Euler^[94]发现的解, 他相继确定了 a 和 d 而不知道序列 $D_n^{(q)}$ 就够了.)

3.45 设 $\alpha > 1$ 满足 $\alpha+1/\alpha=2m$. 则我们找到 $2Y_n = \alpha^{2^n} + \alpha^{-2^n}$, 且由此得到 $Y_n = \lceil \alpha^{2^n} / 2 \rceil$.

3.46 由于 $2n(n+1) = \lfloor 2(n+(1/2))^2 \rfloor$, 依据式(3.9)得到提示, 设 $n+\theta = (\sqrt{2}^l + \sqrt{2}^{l-1})m$ 和 $n'+\theta' = (\sqrt{2}^{l+1} + \sqrt{2}^l)m$, 其中 $0 \leq \theta, \theta' < 1$. 于是 $\theta' = 2\theta \bmod 1 = 2\theta - d$, 其中 d 是 0 或 1. 我们要证明 $n' = \lfloor \sqrt{2}(n+(1/2)) \rfloor$, 此等式成立当且仅当

$$0 \leq \theta'(2 - \sqrt{2}) + \sqrt{2}(1 - d) < 2.$$

为了解递归, 注意 $\text{Spec}(1+1/\sqrt{2})$ 和 $\text{Spec}(1+\sqrt{2})$ 划分正整数; 因此任何正整数 a 能唯一地写成形式 $a = \lfloor (\sqrt{2}^l + \sqrt{2}^{l-1})m \rfloor$, 其中 l 和 m 是整数, m 为奇且 $l \geq 0$. 由此得到 $L_n = \lfloor (\sqrt{2}^{l+n} + \sqrt{2}^{l+n-1})m \rfloor$.

3.47 (a) $C = -(1/2)$. (b) c 是整数. (c) $c=0$. (d) c 是任意的. 更一般的结果见 [173] 中习题 1.2.4-40 的解答.

3.48 (Heinrich Rolletschek 的解.) 我们可用 $(\{\beta\}, \alpha + \lfloor \beta \rfloor)$ 替换 (α, β) , 不改变 $\lfloor n\alpha \rfloor + \lfloor n\beta \rfloor$. 因此条件 $\alpha = \{\beta\}$ 是必要的, 也是充分的: 设 $m = \lfloor \beta \rfloor$ 是给定多重集的最小元素, 且对所有 n , 设 S 是由给定的一个元素获得的多重集, 这个元素为第 n 个最小元素减去 nm . 如果 $\alpha = \{\beta\}$, S 的接连的元素相差为 0 或 2, 因此多重集 $(1/2)S = \text{Spec}(\alpha)$ 确定 α .

3.49 依据 William A. Veech 的未出版的笔记, 在有理数上 $\alpha\beta$, β 和 1 线性独立就够了.

3.50 H. S. Wilf 看到如果我们在任何区间 $(\varphi, \varphi+\varepsilon)$ 上知道 $f(x)$, 则对于所有 $x \geq \varphi$, 函数方程 $f(x^2 - 1) = f(x)^2$ 确定 $f(x)$.

3.51 有无限多种方式把正整数划分为 3 个或 3 个以上的具有无理数 α_k 的推广的范

图; 例如,

$$\text{Spec}(2\alpha; 0) \cup \text{Spec}(4\alpha; -\alpha) \cup \text{Spec}(4\alpha; -3\alpha) \cup \text{Spec}(\beta; 0)$$

是行得通的, 但是有一种明确的意义, 通过“扩张”一个基本的 $\text{Spec}(\alpha) \cup \text{Spec}(\beta)$ 而产生所有这样的划分^[128]. 仅知道有理数的例子, 譬如,

$$\text{Spec}(7; -3) \cup \text{Spec}\left(\frac{7}{2}; -1\right) \cup \text{Spec}\left(\frac{7}{4}; 0\right).$$

是依赖于所述猜测的那些参数, 这是 A. S. Fraenkel^[103]提出的.

3.52 在[77, pp.30~31]讨论了部分结果.

4.1 1, 2, 4, 6, 16, 12.

4.2 注意 $m_p + n_p = \min(m_p, n_p) + \max(m_p, n_p)$. 递归 $\text{lcm}(m, n) = (n / (n \bmod m)) \text{lcm}(n \bmod m, m)$ 是成立的, 但是对于计算 lcm , 它实际是不可取的; 知道计算 $\text{lcm}(m, n)$ 的最好方法是先计算 $\text{gcd}(m, n)$, 然后用 gcd 除 mn .

4.3 如果 x 是整数, 这是成立的, 但是对所有实数 x , $\pi(x)$ 是被定义的. 正确的公式 $\pi(x) - \pi(x-1) = [Lx]$ 是素数]

容易证实.

4.4 在 $1/0$ 和 $0/(-1)$ 之间我们有左右反射的 Stern-Brocot 树, 具有否定的所有分母, 等等, 所以结果是所有 $m \perp n$ 的分数 m/n . 整个构造期间条件 $m'n - mn' = 1$ 仍成立. (称这为 Stern-Brocot 圈, 因为我们能合适地把最后的 $0/1$ 视为与第一个 $0/1$ 相同, 从而在顶处把树连接为一个圈. Stern-Brocot 圈对计算机图形学有有趣的应用, 因为它表示出平面中所有合理的方向.)

4.5 $L^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ 和 $R^k = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$, 甚至当 $k < 0$ 时也成立. (在第六章中我们将发现一个任何 L 和 R 乘积的一般公式.)

4.6 $a = b$. (第三章定义 $x \bmod 0 = x$, 以致这将是真的.)

4.7 我们需要 $m \bmod 10 = 0$, $m \bmod 9 = k$ 和 $m \bmod 8 = 1$. 但是 m 不能既是偶数又是奇数.

4.8 我们要 $10x + 6y \equiv 10x + y \pmod{15}$, 因此 $5y \equiv 0 \pmod{15}$, 因此 $y \equiv 0 \pmod{3}$. 我们一定有 $y = 0$ 或 3 , 以及 $x = 0$ 或 1 .

4.9 $3^{2k+1} \bmod 4 = 3$, 所以 $(3^{2k+1} - 1)/2$ 是奇数. 所述的数可被 $(3^7 - 1)/2$ 和 $(3^{11} - 1)/2$ (以及其他数) 除尽.

4.10 $999(1 - (1/3))(1 - (1/37)) = 648$.

4.11 $\sigma(0) = 1$; $\sigma(1) = -1$; $\sigma(n) = 0$ ($n > 1$). (在具有有趣的和重要的性质的任意偏序结构上定义广义 Mobius 函数, 首先由 Weisner^[299] 研究, 且由其他许多人发展, 特别是 Gian-Carlo Rota^[254].)

$$4.12 \quad \sum_{d \wedge m} \sum_{k \wedge d} \mu(d/k) g(k) = \sum_{k \wedge m} \sum_{d \wedge (m/k)} \mu(d) g(k) = \sum_{k \wedge m} g(k) \cdot [m/k = 1] = g(m),$$

依据式(4.7)和(4.9)。

4.13 (a) 对所有 p , $n_p \leq 1$; (b) $\mu(n) \neq 0$ 。

4.14 当 $k > 0$ 时真。用式(4.12), (4.14)和(4.15)。

4.15 不。例如, $e_n \bmod 5 = [2 \text{ 或 } 3]$; $e_n \bmod 11 = [2, 3, 7 \text{ 或 } 10]$ 。

4.16 $1/e_1 + 1/e_2 + \cdots + 1/e_n = 1 - 1/(e_n(e_n - 1)) = 1 - 1/(e_{n+1} - 1)$ 。

4.17 我们有 $f_n \bmod f_m = 2$; 因此 $\gcd(f_n, f_m) = \gcd(2, f_m) = 1$ 。(顺便提到, 关系 $f_n = f_0 f_1 \cdots f_{n-1} + 2$ 十分相似于定义欧几里德数 e_n 的递归。)

4.18 如果 $n = qm$ 和 q 是奇数, 则 $2^n + 1 = (2^m + 1)(2^{n-m} - 2^{n-2m} + \cdots - 2^m + 1)$ 。

4.19 设 $p_1 = 2$, 且设 p_n 是大于 $2^{p_{n-1}}$ 的最小素数。于是 $2^{p_{n-1}} < p_n < 2^{p_{n-1}+1}$, 由此得到 $b = \lim_{n \rightarrow \infty} \lg^{(n)} p_n$, 其中 $\lg^{(n)}$ 是函数 \lg 重复 n 次。所述的数值来自 $p_2 = 5$, $p_3 = 37$ 。结果 $p_4 = 2^{37} + 9$, 且这给出更精确的值

$$b \approx 1.251\,647\,597\,790\,5$$

(但是关于 p_5 没有线索)。

4.20 依据 Bertrand 公设, $P_n < 10^n$ 。设

$$K = \sum_{k \geq 1} 10^{-k^2} P_k = .200\,300\,005\ldots$$

于是 $10^{n^2} K \equiv P_n + \text{分数} \pmod{10^{2n-1}}$ 。

4.21 第一个和是 $\pi(n)$, 因为被加数是 $(k+1)$ 是素数。第二个和中的内和是 $\sum_{1 \leq k < m} [k \setminus m]$ 。

所以它大于 1 当且仅当 m 是合成数; 我们又取 $\pi(n)$ 。最后, $\lceil \{m/n\} \rceil = (n \times m)$, 所以第三个和是 Wilson 定理的一个应用。当然, 这些公式的任何一个来计算 $\pi(n)$ 是完全愚蠢的。

4.22 $(b^m - 1)/(b - 1) = ((b^m - 1)/(b - 1))(b^{m-m} + \cdots + 1)$, [当 $p = 2, 19, 23, 317, 1\,031$ 时, 仅仅形式 $(10^p - 1)/9$ ($p < 2\,000$) 的素数出现。]

4.23 $\rho(2k+1) = 0$; $\rho(2k) = \rho(k) + 1$ ($k \geq 1$)。如果 $n > 2^m$ 和 $m > \rho(n)$, 我们能归纳法证明 $\rho(n) = \rho(n - 2^m)$ 。如果我们把圆盘编号为 $0, 1, \dots, n-1$, 则第 k 次 Hanoi 移动是圆盘 $\rho(k)$ 。如果 k 是 2 的幂, 这是明显的。如果 $2^m < k < 2^{m+1}$, 我们有 $\rho(k) < m$; 移动 k 和 $k - 2^m$ 对应于用 $T_m + 1 + T_m$ 步转换 $m+1$ 个圆盘的序列。

4.24 提供 dp^n 到 n 的数字位贡献 $dp^{n-1} + \cdots + d = d(p^m - 1)/(p-1)$ 到 $\varepsilon_p(n!)$, 因此 $\varepsilon_p(n!) = (n - v_p(n))/(p-1)$ 。

4.25 $m \setminus n \Leftrightarrow m_p = 0$ 或 $m_p = n_p$ (对所有 p)。由此得到在我们感兴趣的例子 $m = 12, n = 18$ 中 (a) 是真的, 但是 (b) 失效。(这是一种普通的谬误。)

4.26 是真的, 因为 \mathcal{S}_N 定义 Stern-Brocot 树的一个子树。

4.27 用 M 来延伸较短的串(因为 M 依字母顺序位于 L 和 R 之间), 直到两个串长度相等, 然后用字典次序。例如, 树的最顶部的级是 $LL < LM < LR < MM < RL < RM < RR$ 。(另一个解是把无限串 RL^∞ 添加到两个输入, 且继续比较直到发现 $L < R$ 。)

4.28 我们仅需用表示的第一部分:

	R	R	R	L	L	L	L	L	L	L	R	R	R	R	R	R
$\frac{1}{1}$	$\frac{2}{1}$	$\frac{3}{1}$	$\frac{4}{1}$	$\frac{7}{2}$	$\frac{10}{3}$	$\frac{13}{4}$	$\frac{16}{5}$	$\frac{19}{6}$	$\frac{22}{7}$	$\frac{25}{8}$	$\frac{47}{15}$	$\frac{69}{22}$	$\frac{91}{29}$	$\frac{113}{36}$	$\frac{135}{43}$...

分数 $4/1$ 出现, 因为比起 $1/0$ 来它是一个较好的上界, 而不是因为它比 $3/1$ 更接近。类似, 比起 $3/1$ 来 $25/8$ 是一个较好的下界。最简单的上界和最简单的下界全出现了, 但是直到 R 的串前恰好转换回 L 时下一个确实好的近似才出现。

4.29 $1/\alpha$ 。为了在二进表示中从 x 取得 $1-x$, 我们交换 0 和 1; 在 Stern-Brocot 表示中从 α 取得 $1/\alpha$, 我们交换 L 和 R 。(一定还要考虑有限情形, 但是它们一定行得通, 因为对应是维持次序的。)

4.30 m 个整数 $x \in [A..A+m)$ 是不同的 $\bmod m$; 因此它们的剩余 $(x \bmod m_1, \dots, x \bmod m_r)$ 穿过所有 $m_1 \dots m_r = m$ 个可能值, 依据鸽巢原理, 其中之一一定是 $(a_1 \bmod m_1, \dots, a_r \bmod m_r)$ 。

4.31 基数 b 表示中的一个数可被 d 除尽, 当且仅当它的数字位的和可被 d 除尽, 每当 $b \equiv 1 \pmod{d}$ 时, 因为 $(a_m \dots a_0)_b = a_m b^m + \dots + a_0 b^0 \equiv a_m + \dots + a_0$, 得到这一点。

4.32 $\varphi(m)$ 个数 $\{kn \bmod m \mid k \perp m \text{ 和 } 0 \leq k < m\}$ 是以某个次序的数 $\{k \mid k \perp m \text{ 和 } 0 \leq k < m\}$, 把它们乘在一起且用 $\prod_{0 \leq k < m, k \perp m} k$ 除。

4.33 显然 $h(1)=1$ 。如果 $m \perp n$, 则 $h(mn) = \sum_{d \mid mn} f(d)g(mn/d) = \sum_{c \mid m, d \mid n} f(cd)g((m/c)(n/d))$
 $= \sum_{c \mid m} \sum_{d \mid n} f(c)g(m/c)f(d)g(n/d)$; 这是 $h(m)h(n)$, 因为对和中每一项, $c \perp d$ 。

4.34 如果当 x 不是整数时 $f(x)$ 是零, $g(m) = \sum_{d \mid m} f(d) = \sum_{d \mid m} f(m/d) = \sum_{d > 1} f(m/d)$ 。

4.35 基础情形为

$$I(0, n) = 0; \quad I(m, 0) = 1.$$

当 $m, n > 0$ 时, 有两个规则, 如果 $m > n$, 第一个规则是平凡的, 如果 $m < n$, 第二个规则是平凡的:

$$I(m, n) = I(m, n \bmod m) - \lfloor n/m \rfloor I(n \bmod m, m);$$

$$I(m, n) = I(m \bmod n, n).$$

4.36 把任何给定量因子分解为非单位, 一定有 $m^2 - 10n^2 = \pm 2$ 或 ± 3 , 但是这是不可能 $\bmod 10$ 。

4.37 设 $a_n = 2^{-n} \ln\left(e_n - \frac{1}{2}\right)$ 和 $b_n = 2^{-n} \ln\left(e_n + \frac{1}{2}\right)$ 。则

$$e_n = \left\lfloor E^{2^n} + \frac{1}{2} \right\rfloor \Leftrightarrow a_n \leq \ln E < b_n,$$

且 $a_{n-1} < a_n < b_n < b_{n-1}$ ，所以我们能取 $E = \lim_{n \rightarrow \infty} e^{a_n}$ ，事实上，得出

$$E^2 = \frac{3}{2} \prod_{n \geq 1} \left(1 + \frac{1}{(2e_n - 1)^2} \right)^{1/2^n},$$

乘积快速收敛到 $(1.264\ 084\ 735\ 305\ 301\ 11)^2$ 。但是这些观察并不告诉我们 e_n 是什么，除非能找出不依赖于欧几里德数的 E 的另一个表达式。

$$4.38 \quad a^n - b^n = (a^m - b^m)(a^{n-m}b^0 + a^{n-2m}b^m + \cdots + a^{n \bmod m}b^{n-m-n \bmod m}) \\ + b^{m \lfloor n/m \rfloor} (a^{n \bmod m} - b^{n \bmod m}).$$

4.39 如果 $a_1 \cdots a_t$ 和 $b_1 \cdots b_u$ 是完全平方，则

$$a_1 \cdots a_t b_1 \cdots b_u / c_1^2 \cdots c_v^2$$

也是，其中 $\{a_1, \dots, a_t\} \cap \{b_1, \dots, b_u\} = \{c_1, \dots, c_v\}$ 。(事实上，能证明序列 $\langle S(1), S(2), S(3), \dots \rangle$ 包含每一个非系数正整数恰好一次。)

4.40 设 $f(n) = \prod_{1 \leq k \leq n, p \nmid k} k = n! / p^{\lfloor n/p \rfloor} \lfloor n/p \rfloor!$ 和 $g(n) = n! / p^{e_p(n)}$ 。于是

$$g(n) = f(n)f\left(\left\lfloor \frac{n}{p} \right\rfloor\right)f\left(\left\lfloor \frac{n}{p^2} \right\rfloor\right) \cdots = f(n)g\left(\left\lfloor \frac{n}{p} \right\rfloor\right),$$

还有 $f(n) \equiv a_0!(p-1)!^{\lfloor n/p \rfloor} \equiv a_0!(-1)^{\lfloor n/p \rfloor} \pmod{p}$ 和 $e_p(n!) = \lfloor n/p \rfloor + e_p(\lfloor n/p \rfloor!)$ 。这些递归使它容易通过归纳法证明结果。(其他几个解是可能的。)

4.41 (a) 如果 $n^2 \equiv -1 \pmod{p}$ ，则 $(n^2)^{(p-1)/2} \equiv -1$ ，但是 Fermat 说，它是 +1。

(b) 设 $n = ((p-1)/2)!$ ，我们有 $n \equiv (-1)^{(p-1)/2} \prod_{1 \leq k < p/2} (p-k) = (p-1)! / n$ ，

因此 $n^2 \equiv (p-1)!$ 。

4.42 首先我们看到 $k \perp l \Leftrightarrow k \perp l+ak$ (对任何整数 a)，因为依据欧几里德算法 $\gcd(k, l) = \gcd(k, l+ak)$ 。现在

$$m \perp n \text{ 和 } n' \perp n \Leftrightarrow mn' \perp n \\ \Leftrightarrow mn' + nm' \perp n.$$

相似，

$$m' \perp n' \text{ 和 } n \perp n' \Leftrightarrow mn' + nm' \perp n'.$$

因此

$$m \perp n \text{ 和 } m' \perp n' \text{ 和 } n \perp n' \Leftrightarrow mn' + nm' \perp nn'.$$

4.43 我们要用 $L^{-1}R$ 乘, 然后用 $R^{-1}L^{-1}RL$ 乘, 然后 $L^{-1}R$, 然后 $R^{-2}L^{-1}RL^2$, 等等; 第 n 个乘数是 $R^{-\rho(n)}L^{-1}RL^{\rho(n)}$, 因为我们一定要消去 $\rho(n)$ 个 R . 且 $R^{-m}L^{-1}RL^m = \begin{pmatrix} 0 & -1 \\ 1 & 2m+1 \end{pmatrix}$.

4.44 通过查看 $631/2000$ 和 $633/2000$ 的 Stern-Brocot 表示, 且恰好在前者有 L 后者有 R 之前停止, 我们能找到位于

$$[.3155, .3165) = \left[\frac{631}{2000}, \frac{633}{2000} \right)$$

中的最简单的有理数:

$$(m_1, n_1, m_2, n_2) := (631, 2000, 633, 2000);$$

while $m_1 > n_1$ or $m_2 < n_2$ do

if $m_2 < n_2$ then (输出(L); $(n_1, n_2) := (n_1, n_2) - (m_1, m_2)$)

else (输出(R); $(m_1, m_2) := (m_1, m_2) - (n_1, n_2)$).

输出是 $LLLLRRRRR = 6/19 \approx .3158$. 顺便提到, .344 的一个平均意味着至少击球 287 次.

4.45 $x^2 \equiv x \pmod{10^n} \Leftrightarrow x(x-1) \equiv 0 \pmod{2^n}$ 和 $x(x-1) \equiv 0 \pmod{5^n} \Leftrightarrow x \pmod{2^n} = [0 \text{ 或 } 1]$, $x \pmod{5^n} = [0 \text{ 或 } 1]$. (最后步被证明是正确的, 因为 $x(x-1) \pmod{5} = 0$ 意味着 x 或者 $x-1$ 是 5 的倍数, 此时其他因子与 5^n 互素, 且能把它与同余隔开.)

所以至多有 4 个解, 除非 $n=1$, ($x=0$ 与 $x=1$) 两个解没有资格称“ n 个数字位的数”, 其他两个解有形式 x 和 10^n+1-x , 至少这些数的一个数是 $\geq 10^{n-1}$. 当 $n=4$ 时, 另一解 $10001-9376=625$ 不是一个 4 个数字位的数. 我们期望对于 90% 的 n 取得 n 个数字位的解, 但是这个猜测还未被证明.

(这样的自复制的数称为“自同构的”。)

4.46 (a) 如果 $j'j-k'k = \gcd(j, k)$, 我们有 $n^{k/k} n^{\gcd(j, k)} = n^{j'j} \equiv 1$ 和 $n^{k/k} \equiv 1$.

(b) 设 $n=pq$, 其中 p 是 n 的最小素数因子. 如果 $2^n \equiv 1 \pmod{n}$, 则 $2^n \equiv 1 \pmod{p}$, 还有 $2^{p-1} \equiv 1 \pmod{p}$, 因此 $2^{\gcd(p-1, n)} \equiv 1 \pmod{p}$. 但是根据 p 的定义, $\gcd(p-1, n)=1$.

4.47 如果 $n^{m-1} \equiv 1 \pmod{m}$, 我们一定有 $n \perp m$. 如果 $n^k \equiv n^j$ (对某个 $1 \leq j < k < m$), 则 $n^{k-j} \equiv 1$, 因为我们能被 n^j 除尽. 所以如果数 $n^1 \pmod{m}, \dots, n^{m-1} \pmod{m}$ 是不同的, 则有一个 $k < m-1$ (具有 $n^k \equiv 1$). 依据习题 46(a), 最小的这样的 k 除尽 $m-1$. 于是 $kq = (m-1)/p$ (对某个素数 p 和某个正整数 q); 这是不可能的, 因为 $n^{kq} \not\equiv 1$. 所以数 $n^1 \pmod{m}, \dots, n^{m-1} \pmod{m}$ 是不同的, 且与 m 互素. 所以数 $1, \dots, m-1$ 与 m 互素, 且 m 一定是素数.

4.48 通过用数的逆来把数配对, 我们能把乘积(mod m)化为 $\prod_{1 \leq a < m, a^2 \bmod m = 1} a$. 现在把解的知识用到 $a^2 \bmod m = 1$. 如果 $m = 4$, p^k 或 $2p^k$ ($p > 2$), 根据剩余算术我们求得的结果是 $m-1$, 否则它是 $+1$.

4.49 (a) 或 $m < n(\Phi(N-1)$ 种情形), 或 $m = n$ (一种情形), 或 $m > n(\Phi(N-1)$ 种情形), 因此 $R(N) = 2\Phi(N-1) + 1$.

(b) 根据式(4.62)我们取得

$$2\Phi(N-1) + 1 = 1 + \sum_{d \geq 1} \mu(d) \left\lfloor \frac{N}{d} \right\rfloor \left\lfloor \frac{N}{d} - 1 \right\rfloor;$$

因此所述结果成立当且仅当

$$\sum_{d \geq 1} \mu(d) \left\lfloor \frac{N}{d} \right\rfloor = 1, \quad (N \geq 1).$$

如果我们置 $f(x) = [x \geq 1]$, 这是式(4.61)的特殊情形.

4.50 (a) 如果 f 是任何函数,

$$\begin{aligned} \sum_{0 \leq k < m} f(k) &= \sum_{d \wedge m} \sum_{0 \leq k < m} f(k) [d = \gcd(k, m)] \\ &= \sum_{d \wedge m} \sum_{0 \leq k < m} f(k) [k/d \perp m/d] \\ &= \sum_{d \wedge m} \sum_{0 \leq k < m/d} f(kd) [k \perp m/d] \\ &= \sum_{d \wedge m} \sum_{0 \leq k < d} f(km/d) [k \perp d], \end{aligned}$$

在式(4.63)的推导中我们看到这样的一种特殊情形. \prod 代替 \sum , 相似推导成立, 因此我们有

$$z^m - 1 = \prod_{0 \leq k < m} (z - \omega^k) = \prod_{d \wedge m} \prod_{\substack{0 \leq k < d \\ k \perp d}} (z - \omega^{km/d}) = \prod_{d \wedge m} \Psi_d(z)$$

因为 $\omega^{m/d} = e^{2\pi i/d}$.

据乘积代替和的式(4.56)的相似结果, 从(a)得到(b), 顺便提到, 此公式证明 $\Psi_m(z)$ 有整数系数, 因此通过乘和除首项系数为 1 的多项式得到 $\Psi_m(z)$.

$$4.51 \quad (x_1 + \cdots + x_n)^p = \sum_{k_1 + \cdots + k_n = p} p! / (k_1! \cdots k_n!) x_1^{k_1} \cdots x_n^{k_n}, \text{ 且除非某个 } k_j = p,$$

系数被 p 除尽. 因此 $(x_1 + \cdots + x_n)^p \equiv x_1^p + \cdots + x_n^p \pmod{p}$. 现在我们能把所有 x 置为 1, 得到 $n^p \equiv n$.

4.52 如果 $p > n$, 没有什么证明的. 否则 $x \perp p$, 所以 $x^{k(p-1)} \equiv 1 \pmod{p}$; 这意味着给定数中至少 $\lfloor (n-1)/(p-1) \rfloor$ 个是 p 的倍数, 且由于 $n \geq p$, 所以 $(n-1)/(p-1) \geq n/p$.

4.53 首先证明如果 $m \geq 6$ 和 m 不是素数, 则 $(m-2)! \equiv 0 \pmod{m}$. (如果 $m = p^2$, $(m-2)!$ 的乘积包含 p 和 $2p$; 否则它包含 d 和 m/d , 其中 $d < m/d$.) 接着考虑下列情形:

情形 0, $n < 5$. 仅对 $n=1$, 条件成立.

情形 1, $n \geq 5$ 和 n 是素数. 则 $(n-1)!/(n+1)$ 是整数且它不能为 n 的倍数.

情形 2, $n \geq 5$, n 是合成数, 且 $n+1$ 是合成数. 则 n 和 $n+1$ 除尽 $(n-1)!$, 且 $n \perp n+1$, 因此 $n(n+1) \nmid (n-1)!$.

情形 3, $n \geq 5$, n 是合成数, 且 $n+1$ 是素数. 则 $(n-1)! \equiv 1 \pmod{n+1}$ (依据 Wilson 定理), 且

$$\left\lfloor \frac{(n-1)!}{(n+1)} \right\rfloor = \frac{((n-1)! + n)}{(n+1)};$$

n 可除尽它.

所以解答是: 或者 $n=1$, 或者 $n \neq 4$ 是合成数.

4.54 $\varepsilon_2(1000!) > 500$ 和 $\varepsilon_5(1000!) = 249$, 因此对某个偶整数 a , $1000! = a \cdot 10^{249}$. 由于 $1000 = (1300)_5$, 习题 40 告诉我们 $a \cdot 2^{249} = 1000! / 5^{249} \equiv -1 \pmod{5}$, 还有 $2^{249} \equiv 2$, 因此 $a \equiv 2$, 因此 $a \bmod 10 = 2$ 或 7 ; 解答是 $2 \cdot 10^{249}$.

4.55 一种方法是用归纳法证明 $P_{2n}/P_n^4(n+1)$ 是一个整数, 这个较强结果帮助做完归纳法. 另一种方法是基于证明至少每当每个素数 p 除尽分母时它除尽分子. 这就化成证明不等式

$$\sum_{k=1}^{2n} \left\lfloor \frac{k}{m} \right\rfloor \geq 4 \sum_{k=1}^n \left\lfloor \frac{k}{m} \right\rfloor,$$

它来自

$$\left\lfloor \frac{(2n-1)}{m} \right\rfloor + \left\lfloor \frac{2n}{m} \right\rfloor \geq \left\lfloor \frac{n}{m} \right\rfloor.$$

当 $0 \leq n < m$ 时, 后者为真, 当 n 增加 m 时两边增加 4.

4.56 设 $f(m) = \sum_{k=1}^{2n-1} \min(k, 2n-k)[m \setminus k]$, $g(m) = \sum_{k=1}^{n-1} (2n-2k-1) \times [m \setminus (2k+1)]$. p 除尽所述乘积的分子的次数是 $f(p) + f(p^2) + f(p^3) + \dots$, p 除尽分母的次数是 $g(p) + g(p^2) + g(p^3) + \dots$. 但是根据习题 2.32, 每当 m 为奇时 $f(m) = g(m)$. 根据习题 3.22, 所述乘积化为 $2^{n(n-1)}$.

4.57 由于

$$\sum_{1 \leq m \leq n} [d \setminus m] = \sum_{0 < k \leq n/d} [m = dk] = \lfloor n/d \rfloor,$$

提示提供了一种标准的交换求和, 调用提示的和 $\sum(n)$, 我们有

$$\sum(m+n) - \sum(m) - \sum(n) = \sum_{d \in S(m, n)} \varphi(d)$$

另一方面, 我们从式(4.54)知道 $\sum(n) = \frac{1}{2}n(n+1)$. 因此 $\sum(m+n) - \sum(m) - \sum(n) = mn$.

4.58 函数 $f(m)$ 是积性的, 且当 $m = p^k$ 时它等于 $1+p+\cdots+p^k$. 这是2的幂当且仅当 p 是一个 Mersenne 素数且 $k=1$. k 一定是奇数, 且此时和为

$$(1+p)(1+p^2+p^4+\cdots+p^{k-1})$$

且 $(k-1)/2$ 一定是奇数, 等等. 充分必要条件是 m 为不同的 Mersenne 素数的乘积.

4.59 提示的证明: 如果 $n=1$, 我们有 $x_1 = \alpha = 2$, 所以没有问题. 如果 $n>1$, 我们可设 $x_1 \leq \cdots < x_n$. 情形 1: $x_1^{-1} + \cdots + x_{n-1}^{-1} + (x_n - 1)^{-1} \geq 1$ 和 $x_n > x_{n-1}$. 于是我们能找到 $\beta \geq x_n - 1 \geq x_{n-1}$ 使得 $x_1^{-1} + \cdots + x_{n-1}^{-1} + \beta^{-1} = 1$; 因此 $x_n \leq \beta + 1 \leq e_n$ 和 $x_1 \cdots x_n \leq x_1 \cdots x_{n-1}(\beta + 1) \leq e_1 \cdots e_n$. 用归纳法, 有一个正整数 m 使得 $\alpha = x_1 \cdots x_n / m$, 因此 $\alpha \leq e_1 \cdots e_n = e_{n+1} - 1$, 并且我们有 $x_1 \cdots x_n(\alpha + 1) \leq e_1 \cdots e_n e_{n+1}$. 情形 2: $x_1^{-1} + \cdots + x_{n-1}^{-1} + (x_n - 1)^{-1} \geq 1$ 和 $x_n = x_{n-1}$. 设 $a = x_n$ 和 $a^{-1} + (a-1)^{-1} = (a-2)^{-1} + \zeta^{-1}$. 于是我们能证明 $a \geq 4$ 和 $(a-2)(\zeta + 1) \geq a^2$. 所以有一个 $\beta \geq \zeta$ 使得 $x_1^{-1} + \cdots + x_{n-2}^{-1} + (a-2)^{-1} + \beta^{-1} = 1$; 由此用归纳法得到 $x_1 \cdots x_n \leq x_1 \cdots x_{n-2}(a-2)(\zeta + 1) \leq x_1 \cdots x_{n-2}(a-2)(\beta + 1) \leq e_1 \cdots e_n$, 且如同前面那样我们能结束. 情形 3: $x_1^{-1} + \cdots + x_{n-1}^{-1} + (x_n - 1)^{-1} < 1$. 设 $a = x_n$, 且设 $a^{-1} + \alpha^{-1} = (a-1)^{-1} + \beta^{-1}$. 由此能证 $(a-1)(\beta + 1) > a(\alpha + 1)$, 因为此等式等价于

$$a\alpha^2 + a^2\alpha + a\alpha - a^2 + \alpha + a > 0,$$

它是 $a\alpha(\alpha - a) + (1 + a)\alpha \geq (1 + a)\alpha > a^2 - a$ 的一个推论. 因此我们能用 $\alpha - 1$ 和 β 替换 x_n 和 α , 重复此变换直到情形 1 或 2 应用.

提示的另一个推论是 $1/x_1 + \cdots + 1/x_n < 1$ 意味着 $1/x_1 + \cdots + 1/x_n \leq 1/e_1 + \cdots + 1/e_n$, 见习题 16.

4.60 主要点是 $\theta < 2/3$. 于是我们能取 p_1 充分大(下面遇到这些条件)以及 p_n 是比 p_{n-1}^3 大的最小素数. 用此定义设 $a_n = 3^{-n} \ln p_n$ 和 $b_n = 3^{-n} \ln(p_n + 1)$. 如果我们能证明 $a_{n-1} \leq a_n < b_n \leq b_{n-1}$, 如同习题 37 中那样, 我们能取 $P = \lim_{n \rightarrow \infty} e^{a_n}$. 但是此假设是等价于 $p_{n-1}^3 \leq p_n < (p_{n-1} + 1)^3$. 如果在此范围中没有素数 p_n , 一定有一个素数 $p < p_{n-1}^3$ 使得 $p + cp^{\theta} > (p_{n-1} + 1)^3$. 但是这意味着 $cp^{\theta} > 3p^{2/3}$, 当 p 充分大时, 这是不可能的.

我们几乎一定取 $p_1 = 2$, 因为所有可用的根据表示, 素数间间隙的已知界限比实际情

况弱得多(见习题 69)。于是 $p_2 = 11$, $p_3 = 1361$, $p_4 = 2521008887$, $1.306377883863 < P < 1.306377883869$ 。

4.61 设 \hat{m} 和 \hat{n} 是右边; 看到 $\hat{m}n' - m'\hat{n} = 1$, 因此 $\hat{m} \perp \hat{n}$ 。还有 $\hat{m}/\hat{n} > m'/n'$ 和 $N = ((n+N)/n')n' - n \geq \hat{n} > ((n+N)/(n'-1))n' - n = N - n' \geq 0$ 。所以我们有 $\hat{m}/\hat{n} \geq m''/n''$ 。如果等式不成立, 我们有 $n'' = (\hat{m}n' - m'\hat{n})n'' = n'(\hat{m}n'' - m''\hat{n}) + \hat{n}(m''n' - m'n'') \geq n' + \hat{n} > N$, 一个矛盾。

顺便提到, 此习题意味着 $(m+m'')/(n+n'') = m'/n'$, 虽然前面的分数总不能化简。

4.62 $2^{-1} + 2^{-2} + 2^{-3} - 2^{-6} - 2^{-7} + 2^{-12} + 2^{-13} - 2^{-20} - 2^{-21} + 2^{-30} + 2^{-31} - 2^{-42} - 2^{-43} + \dots$ 能写成

$$\frac{1}{2} + 3 \sum_{k \geq 0} (2^{-4k^2 - 6k - 3} - 2^{-4k^2 - 10k - 7}).$$

顺便提到, 此和能用“ θ 函数” $\theta(z, \lambda) = \sum_k e^{-\pi \lambda k^2 + 2izk}$ 表为闭形式, 我们有

$$e \leftrightarrow \frac{1}{2} + \frac{3}{8} \theta\left(\frac{4}{\pi} \ln 2, 3i \ln 2\right) - \frac{3}{128} \theta\left(\frac{4}{\pi} \ln 2, 5i \ln 2\right).$$

4.63 任何 $n > 2$ 或者有一个系数因子 d , 或者可被 $d=4$ 除尽。不论发生哪一种情形, 具有幂 n 的一个解意味着具有幂 d 的一个解 $(a^{n/d})^d + (b^{n/d})^d = (c^{n/d})^d$ 。由于 $d=4$ 没有解, 所以 d 一定是素数。

提示来自二项定理, 因为当 p 是奇数时, $a^p + (x-a)^p - pa^{p-1}$ 是 x 的倍数。假设 $a \perp x$ 。如果 p 不可除尽 x , 则 x 和 c^p/x 互素; 因此对某个 m , $x = m^p$ 。如果 p 可除尽 x , 则 p 可除尽 c^p/x , 但是 p^2 不可除尽 c^p/x , 且 c^p 和 x 一样, 没有其他因子。

(事实上, a, b, c 的值一定还高于这个结果指出的! Inkeri^[160]证明了

$$\min(a, b) > \left(\frac{2p^3 + p}{\ln(3p)} \right)^p.$$

他的证明可见[249, pp.228–229], 该书包含 Fermat 最后定理进展的广泛的评述。)

4.64 \mathcal{P}_N 中的相等分数以“管风琴次序”出现

$$\frac{2m}{2n}, \frac{4m}{4n}, \dots, \frac{rm}{rn}, \dots, \frac{3m}{3n}, \frac{m}{n}.$$

假设 \mathcal{P}_N 是正确的, 我们要证明 \mathcal{P}_{N+1} 是正确的。这意味着如果 kN 是奇数, 我们要证明

$$\frac{k-1}{N+1} = \mathcal{P}_{N, kN};$$

如果 kN 是偶数, 我们要证明

$$\mathcal{P}_{N, kN-1} \mathcal{P}_{N, kN} \frac{k-1}{N+1} \mathcal{P}_{N, kN} \mathcal{P}_{N, kN+1}.$$

在两种情形中, 知道 \mathcal{P}_N 中分数的个数严格小于 $(k-1)/(N+1)$ 将是有益的; 依据式(3.32), 这是

$$\begin{aligned}\sum_{n=1}^N \sum_m \left[0 \leq \frac{m}{n} < \frac{k-1}{N+1} \right] &= \sum_{n=1}^N \left\lfloor \frac{(k-1)n}{N+1} \right\rfloor = \sum_{n=0}^N \left\lfloor \frac{(k-1)n + N}{N+1} \right\rfloor \\ &= \frac{(k-2)N}{2} + \frac{d-1}{2} + d \left\lfloor \frac{N}{d} \right\rfloor \\ &= \frac{1}{2}(kN - d + 1), \quad d = \gcd(k-1, N+1).\end{aligned}$$

此外, 依据管风琴次序的性质, 在 \mathcal{P}_N 中分数个数等于 $(k-1)/(N+1)$, 在 \mathcal{P}_{N+1} 中位于它之前的应是 $\frac{1}{2}(d-1 - [d \text{ 偶数}])$.

如果 kN 是奇数, 则 d 是偶数且 \mathcal{P}_N 的 $\frac{1}{2}(kN-1)$ 个元素前有 $(k-1)/(N+1)$ 个; 这就是使事情行得通的正确个数. 如果 kN 是偶数, 则 d 是奇数且 \mathcal{P}_N 的 $\frac{1}{2}(kN)$ 个元素前有 $(k-1)/(N+1)$ 个. 如果 $d=1$, 没有一个等于 $(k-1)/(N+1)$, 且 $\mathcal{P}_{N, kN}$ 是' $<$ '; 否则 $(k-1)/(N+1)$ 落在两个相同元素之间且 $\mathcal{P}_{N, kN}$ 是' $=$ '. (差不多在 C.S. Peirce 发现 \mathcal{P}_N 的同时, 他独立发现了 Stern-Brocot 树^[230].)

4.65 (类似)Fermat 数 f_n 的相似问题是一个著名的未解决的问题, 这一问题可能较容易或者较难.

4.66 知道没有小于 36×10^{18} 的平方除尽一个 Mersenne 数或 Fermat 数, 但是仍没有 Schinzel 猜测的证明, 即存在无限多个无平方的 Mersenne 数. 还不知道是否有无限多个 p 使得 $p \nmid (a \pm b)$, 其中所有 a 和 b 的素数因子都 ≤ 31 .

4.67 M. Szegedy 证明了这个猜测(对所有大的 n), 见[77, 78-79 页], [49]和[284].

4.68 与下面的习题相比, 这是很弱的猜测.

4.69 Cramer^[56]证明在概率统计基础上这个猜测似乎是可能的, 且计算经验证实了这一点: Brent^[32]证明了对于 $P_{n+1} < 2.686 \times 10^{12}$, $P_{n+1} - P_n \leq 602$, 但是习题 60 中很弱的界限是目前最好的证明^[221]. 对于所有充分大的 n , 如果 $P_{n+1} - P_n < 2P_n^{1/2}$, 习题 68 有一个“是”的回答, 依据 Guy^[139, 问题A8], 对于所有 $c > 0$ 有无限多个 n 使得

$$P_{n+1} - P_n > \frac{c \ln n \ln \ln n \ln \ln \ln n}{(\ln \ln n)^2} \quad (\text{对所有 } c > 0)$$

的证明, Paul Erdos 出价 \$10 000.

4.70 依据习题 24, 此成立当且仅当 $v_2(n) = v_3(n)$, [78]的方法可能有助于揭开这个猜测.

4.71 当 $k=3$ 时, 最小解是 $n = 4\,700\,063\,497 = 19 \cdot 47 \cdot 5\,263\,229$, 在此情形中还不知道其他解.

4.72 知道对于无限多 a 值, (当然)包含 -1 和(不太明显地)包含 0 , 是真的, Lehmer^[199]

有一个著名的猜测, $\varphi(n) \setminus (n-1)$ 当且仅当 n 是素数.

4.73 知道这是等价于 Riemann 假设(实部在 0 和 1 之间的复 ζ 函数的所有零点具有等于 $1/2$ 的实部).

4.74 经验证据提出差不多有 $p(1-1/e)$ 个不同的值, 就好像对模 p 随机分配阶乘.

5.1 由二项定理, 在任何基数 $r \geq 7$ 的数系中, $(11)_r^4 = (14641)_r$.

5.2 当 $k \geq \lfloor n/2 \rfloor$ 时, 比值 $\binom{n}{k+1} / \binom{n}{k} = (n-k)/(k+1)$ 是 ≤ 1 , 当 $k < \lceil n/2 \rceil$ 时, ≥ 1 , 所以当 $k = \lfloor n/2 \rfloor$ 和 $k = \lceil n/2 \rceil$ 时, 最大值出现.

5.3 展开成阶乘. 两个乘积都等于 $f(n)/f(n-k)f(k)$, 其中 $f(n) = (n+1)!n!(n-1)!$.

$$5.4 \quad \binom{-1}{k} = (-1)^k \binom{k+1-1}{k} = (-1)^k \binom{k}{k} = (-1)^k [k \geq 0].$$

5.5 如果 $0 < k < p$, $\binom{p}{k}$ 的分子中有一个 p , 在分母中没有消去它. 由于 $\binom{p}{k} = \binom{p-1}{k} + \binom{p-1}{k-1}$, 对于 $0 \leq k < p$, 我们一定有 $\binom{p-1}{k} \equiv (-1)^k \pmod{p}$.

5.6 (在第二次降下之后)关键步应是

$$\begin{aligned} & \frac{1}{n+1} \sum_k \binom{n+k}{k} \binom{n+1}{k+1} (-1)^k \\ &= \frac{1}{n+1} \sum_{k \geq 0} \binom{n+k}{n} \binom{n+1}{k+1} (-1)^k \\ &= \frac{1}{n+1} \sum_k \binom{n+k}{n} \binom{n+1}{k+1} (-1)^k \\ & \quad - \frac{1}{n+1} \binom{n-1}{n} \binom{n+1}{0} (-1)^{-1}. \end{aligned}$$

原来的推导忘记包含这个附加的项, 它是 $[n=0]$.

5.7 真的, 因为 $r^{-k} = (-1)^k / (-r-1)^k$, 我们还有

$$r^k \left(r + \frac{1}{2}\right)^k = (2r)^{2k} / 2^{2k}.$$

5.8 $f(k) = (k/n-1)^n$ 是次数为 n 的多项式, 它的首项系数是 n^{-n} , 依据式(5.40), 和为 $n! / n^n$. 当 n 大时, Stirling 近似式表明这近似于 $\sqrt{2\pi n} / e^n$. (这是完全不同于 $(1-1/e)$, 如果我们用近似式 $(1-k/n)^n \sim e^{-k}$ (当 $n \rightarrow \infty$ 时对指定的 k 成立), 这就是我们所得到的.)

$$\begin{aligned} 5.9 \quad \text{依据式 (5.70), } \mathcal{E}_i(z)^i &= \sum_{k \geq 0} i(ik+i)^{k-1} z^k / k! = \sum_{k \geq 0} (k+1)^{k-1} (iz)^k / k! \\ &= \mathcal{E}_1(iz). \end{aligned}$$

5.10 由于 $t_{k+1}/t_k = (k+2)z/(k+3)$, $\sum_{k \geq 0} 2z^k/(k+2) = F(2, 1; 3; z)$.

5.11 第一个是 Bessel 的, 第二个是 Gauss 的:

$$z^{-1} \sin z = \sum_{k \geq 0} (-1)^k z^{2k} / (2k+1)! = F\left(1; 1, \frac{3}{2}; -z^2/4\right);$$

$$z^{-1} \arcsin z = \sum_{k \geq 0} z^{2k} \left(\frac{1}{2}\right)^{\bar{k}} / (2k+1)k! = F\left(\frac{1}{2}, \frac{1}{2}; \frac{3}{2}; z^2\right).$$

5.12 (a) 是的, 项比是 n , (b) 不是, 当 $k=0$ 时, 值该是 1; 但是如果 n 是整数, $(k+1)^n$ 行得通. (c) 是的, 项比是 $(k+1)(k+3)/(k+2)$. (d) 不是, 项比是 $1+1/(k+1)H_k$; 且 $H_k \sim \ln k$ 不是一个有理函数. (e) 是的, 项比是

$$\frac{t(k+1)}{t(k)} / \frac{T(n-k)}{T(n-k-1)}.$$

(f) 不总是; 譬如, 当 $t(k) = 2^k$, $T(k) = 1$ 时, 不是, (g) 是的, 项比能写成

$$\frac{at(k+1)/t(k) + bt(k+2)/t(k) + ct(k+3)/t(k)}{a + bt(k+1)/t(k) + ct(k+2)/t(k)},$$

且 $t(k+m)/t(k) = (t(k+m)/t(k+m-1)) \cdots (t(k+1)/t(k))$ 是 k 的一个有理函数.

$$5.13 \quad R_n = n!^{n+1} / P_n^2 = Q_n / P_n = Q_n^2 / n!^{n+1}.$$

5.14 当 $k \leq l$ 时, 式(5.25)中的第一个因子是 $\binom{l-k}{l-k-m}$, 且这是 $(-1)^{l-k-m} \times \binom{-m-1}{l-k-m}$. 由于 $m \geq 0$, $k \leq l$ 的和是所有 k 上的总和. (条件 $n \geq 0$ 实际不是必须的, 虽然如果 $n < 0$, k 一定取负值.)

为了从式(5.25)得到式(5.26), 首先用 $-1-n-q$ 替换 s .

5.15 如果 n 是奇数, 和是零, 因为我们能用 $n-k$ 替换 k . 如果 $n=2m$, 依据式(5.29) ($a=b=c=m$), 和是 $(-1)^m (3m)! / m!^3$.

5.16 如果我们用阶乘来写出被加数, 这就是

$$(2a)! (2b)! (2c)! / (a+b)! (b+c)! (c+a)! \text{ 乘式(5.29).}$$

$$5.17 \quad \binom{2n-1/2}{n} = \binom{4n}{2n} / 2^{2n}; \quad \binom{2n-1/2}{2n} = \binom{4n}{2n} / 2^{4n}; \quad \text{所以} \quad \binom{2n-1/2}{n} = 2^{2n} \binom{2n-1/2}{2n}.$$

$$5.18 \quad \binom{3r}{3k} = \binom{3k}{k, k, k} / 3^{3k}.$$

5.19 依据式(5.60), $\mathcal{B}_{1-1}(-z)^{-1} = \sum_{k \geq 0} \binom{k-1k-1}{k} (-1/(k-1k-1)) (-z)^k$, 且这

是 $\sum_{k \geq 0} \binom{tk}{k} (1/(tk-k+1)) z^k = \mathcal{B}_t(z)$.

5.20 它等于 $F(-a_1, \dots, -a_m; -b_1, \dots, -b_n; (-1)^{m+n} z)$, 见习题 2.17.

5.21 $\lim_{n \rightarrow \infty} (n+m)^{\frac{m}{n}} / n^m = 1$.

5.22 式(5.83)的乘和除的情况给出(依据式(5.44)和(5.46))

$$\begin{aligned} \frac{(-1/2)!}{x!(x-1/2)!} &= \lim_{n \rightarrow \infty} \binom{n+x}{n} \binom{n+x-1/2}{n} n^{-2x} / \binom{n-1/2}{n} \\ &= \lim_{n \rightarrow \infty} \binom{2n+2x}{2n} n^{-2x} \end{aligned}$$

还有

$$\frac{1}{(2x)!} = \lim_{n \rightarrow \infty} \binom{2n+2x}{2n} (2n)^{-2x}$$

等等, 顺便提到, Γ 函数等价式为

$$\Gamma(x)\Gamma\left(x+\frac{1}{2}\right) = \Gamma(2x)\Gamma\left(\frac{1}{2}\right) / 2^{2x-1}.$$

5.23 $(-1)^n n!$, 见式(5.50).

5.24 依据式(5.35)和(5.93), 此和为 $\binom{n}{m} F\left(\begin{matrix} m-n, -m \\ 1/2 \end{matrix} \middle| 1\right) = \binom{2n}{2m}$.

5.25 这是等价于易证的等式

$$(a-b) \frac{a^{\bar{k}}}{(b+1)^{\bar{k}}} = a \frac{(a+1)^{\bar{k}}}{(b+1)^{\bar{k}}} - b \frac{a^{\bar{k}}}{b^{\bar{k}}}$$

像算子公式 $a-b = (0+a) - (0+b)$.

我们相似有

$$\begin{aligned} (a_1 - a_2) F\left(\begin{matrix} a_1, a_2, a_3, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) \\ = a_1 F\left(\begin{matrix} a_1 + 1, a_2, a_3, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) - a_2 F\left(\begin{matrix} a_1, a_2 + 1, a_3, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right), \end{aligned}$$

因为 $a_1 - a_2 = (a_1 + k) - (a_2 + k)$. 如果 $a_1 - b_1$ 是非负整数 d , 这个第二个等式允许我们把 $F(a_1, \dots, a_m; b_1, \dots, b_n; z)$ 表为 $F(a_2+j, a_3, \dots, a_m; b_2, \dots, b_n; z)$ ($0 \leq j \leq d$) 的线性组合, 因此消去一个上参数和一个下参数. 因此, 例如, 我们取得 $F(a, b; a-1; z)$, $F(a, b; a-2; z)$ 等的闭形式.

Gauss^[116, § 7]导出 $F(a, b; c; z)$ 和任何两个“接近的”超几何之间的相似关系, 超几何中一个参数改变 ± 1 . Rainville^[242]把它推广到多个参数的情形.

5.26 如果原来的超几何级数中项比是 $t_{k+1}/t_k = r(k)$, 则新的超几何级数中项比是 $t_{k+2}/t_{k+1} = r(k+1)$. 因此

$$F\left(\begin{matrix} a_1, \dots, a_m \\ b_1, \dots, b_n \end{matrix} \middle| z\right) = 1 + \frac{a_1 \cdots a_m z}{b_1 \cdots b_n} F\left(\begin{matrix} a_1 + 1, \dots, a_m + 1, 1 \\ b_1 + 1, \dots, b_n + 1, 2 \end{matrix} \middle| z\right)$$

5.27 这是 $F(2a_1, \dots, 2a_m; 2b_1, \dots, 2b_n; z)$ 的偶数项的和. 我们有 $(2a)^{\overline{2k+2}} / (2a)^{\overline{2k}} = 4(k+a)\left(k+a+\frac{1}{2}\right)$, 等等.

5.28 我们有 $F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| z\right) = (1-z)^{-a} F\left(\begin{matrix} a, c-b \\ c \end{matrix} \middle| \frac{-z}{1-z}\right) = (1-z)^{-a} F\left(\begin{matrix} c-b, a \\ c \end{matrix} \middle| \frac{-z}{1-z}\right) = (1-z)^{c-a-b} F\left(\begin{matrix} c-a, c-b \\ c \end{matrix} \middle| z\right)$. (Euler 通过证明两边满足相同的微分方程, 证明了等式. 反射律常归功于 Euler, 但是在他发表的文章中似乎未出现它.)

5.29 依据 Vandermonde 卷积, z^n 的系数是相等的. (Kummer 原来的证明是不同的, 他考虑反射律(5.101)中的 $\lim_{m \rightarrow \infty} F(m, b-a; b; z/m)$.)

5.30 再微分, 取得 $z(1-z)F''(z) + (2-3z)F'(z) - F(z) = 0$. 所以依据式(5.108), $F(z) = F(1, 1; 2; z)$.

5.31 条件 $f(k) = cT(k+1) - cT(k)$ 意味着 $f(k+1)/f(k) = (T(k+2)/T(k+1) - 1)/(1 - T(k)/T(k+1))$ 是 k 的有理函数.

5.32 当以 k 求一个多项式的和时, Gosper 方法化成“未定系数法”. 我们有 $q(k) = r(k) = 1$, 且试解 $p(k) = s(k+1) - s(k)$. 方法提出设 $s(k)$ 是一个次数为 $d = \deg(p) + 1$ 的多项式.

5.33 $k = (k-1)s(k+1) - (k+1)s(k)$ 的解是 $s(k) = -k + \frac{1}{2}$, 因此解答是 $(1-2k)/2k(k-1) + C$.

5.34 极限关系成立, 因为 $k > c$ 的所有项成为零, 且在其他项的极限中用 $-c$ 消去 $\varepsilon - c$. 所以第二个部分和是 $\lim_{\varepsilon \rightarrow 0} F(-m, -n; \varepsilon - m; 1) = \lim_{\varepsilon \rightarrow 0} (\varepsilon + n - m)^{\overline{m}} / (\varepsilon - m)^{\overline{m}} = (-1)^m \binom{n-1}{m}$.

5.35 (a) $2^{-n} 3^n [n \geq 0]$. (b) $\left(1 - \frac{1}{2}\right)^{-k-1} [k \geq 0] = 2^{k+1} [k \geq 0]$.

5.36 $m+n$ 的数字位的和是 m 的数字位的和加上 n 的数字位的和, 减去 $p-1$ 乘进位的次数, 因为每次进位数字位和减少 $p-1$.

5.37 用 $n!$ 除第一个等式产生 $\binom{x+y}{n} = \sum_k \binom{x}{k} \binom{y}{n-k}$, Vandermonde 卷积.

例如, 如果我们把 x 和 y 都取负, 根据公式 $x^k = (-1)^k (-x)^k$ 得到第二个等式。

5.38 选尽可能大的 c 使得 $\binom{c}{3} \leq n$, 则 $0 \leq n - \binom{c}{3} < \binom{c+1}{3} - \binom{c}{3} = \binom{c}{2}$; 用 $n - \binom{c}{3}$ 替换 n , 且继续用相同方法。相反, 此法中获得任何这样的表示。(对于任何指定的 m , 我们能

$$n = \binom{a_1}{1} + \binom{a_2}{2} + \cdots + \binom{a_m}{m}, \quad 0 \leq a_1 < a_2 < \cdots < a_m$$

处理相同的事情。)

5.39 通过对 $m+n$ 的归纳, 对于所有 $mn > 0$, $x^m y^n = \sum_{k=1}^m \binom{m+n-1-k}{n-1} a^n b^{m-k} x^k + \sum_{k=1}^n \binom{m+n-1-k}{m-1} a^{n-k} b^m y^k$.

$$\begin{aligned} 5.40 \quad & (-1)^{m+1} \sum_{k=1}^n \sum_{j=1}^m \binom{r}{j} \binom{m-rk-s-1}{m-j} = (-1)^{m+1} \sum_{k=1}^n \left(\binom{m-r(k-1)-s-1}{m} \right. \\ & \left. - \binom{m-rk-s-1}{m} \right) = (-1)^{m+1} \left(\binom{m-s-1}{m} - \binom{m-rn-s-1}{m} \right) = \binom{rn+s}{m} \\ & - \binom{s}{m}. \end{aligned}$$

5.41 $\sum_{k \geq 0} n! / (n-k)! (n+k+1)! = (n! / (2n+1)!) \sum_{k \geq n} \binom{2n+1}{k}$, 它是 $2^{2n} n! / (2n+1)!$.

5.42 我们把 n 作为一个不确定的实变量处理, 具有 $q(k) = k+1$ 和 $r(k) = k-1-n$ 的 Gosper 方法有解 $s(k) = 1/(n+2)$; 因此希望的不确定的和是 $(-1)^{x-1} \frac{n+1}{n+2} / \binom{n+1}{k}$ 且

$$\sum_{k=0}^n (-1)^k / \binom{n}{k} = (-1)^{x-1} \frac{n+1}{n+2} / \left(\binom{n+1}{x} \right)_0^{n+1} = 2 \frac{n+1}{n+2} [n \text{ 偶}].$$

顺便提到, 这个习题意味着公式

$$\frac{1}{n \binom{n-1}{k}} = \frac{1}{(n+1) \binom{n}{k+1}} + \frac{1}{(n+1) \binom{n}{k}},$$

是基本递归(5.8)的一个“对偶”。

5.43 在示意的第一步之后, 我们能应用式(5.21)并在 k 上求和, 于是再应用式(5.21)和 Vandermonde 卷积结束工作。(Andrews^[10]给出了这个等式的组合证明。在[173, 习题 1.2.6–62]中说明有一种快速方法给出从此等式到式(5.29)的证明。)

5.44 消阶乘证明

$$\binom{m}{j} \binom{n}{k} \binom{m+n}{m} = \binom{m+n-j-k}{m-j} \binom{j+k}{j} \binom{m+n}{j+k},$$

所以第二个和是 $1/\binom{m+n}{m}$ 乘第一个。每当 $n \geq b$, 甚至 $m < a$, 我们能证明第一个和是 $\binom{a+b}{a} \binom{m+n-a-b}{m-b}$: 设 a 和 b 是指定的, 且称第一个和为 $S(m, n)$. 等式(5.32)包含情形 $n=b$, 且由于 $\binom{m+n-j-k}{m-j} = \binom{m+n-1-j-k}{m-j} + \binom{m-1+n-j-k}{m-1-j}$, 我们有 $S(m, n) = S(m, n-1) + S(m-1, n) + (-1)^{m+n} \binom{m+n}{m} \binom{a}{m} \binom{b}{n}$. 由于当 $n > b$ 时, $\binom{b}{n} = 0$, 通过对 $m+n$ 归纳得到结果, 情形 $m=0$ 是平凡的. 依据对称性, 每当 $m \geq a$, 甚至 $n < b$ 时, 公式 $\binom{a+b}{a} \binom{m+n-a-b}{n-a}$ 成立.

5.45 依据式(5.9), $\sum_{k \leq n} \binom{k-1/2}{k} = \binom{n+1/2}{n}$. 如果此形式不够“闭”, 我们能使用式(5.35)且取得 $(2n+1) \binom{2n}{n} 4^{-n}$.

5.46 依据式(5.79), 此卷积是 $\mathcal{B}_{-1}(z) \mathcal{B}_{-1}(-z)$ 中 z^{2n} 的系数的负值. 现在 $(2\mathcal{B}_{-1}(z)-1)(2\mathcal{B}_{-1}(-z)-1) = \sqrt{1-16z^2}$, 因此 $\mathcal{B}_{-1}(z) \mathcal{B}_{-1}(-z) = \frac{1}{4} \sqrt{1-16z^2} + \frac{1}{2} \mathcal{B}_{-1}(z) + \frac{1}{2} \mathcal{B}_{-1}(-z) - \frac{1}{4}$. 用二项式定理:

$$(1-16z^2)^{1/2} = \sum_n \binom{1/2}{n} (-16)^n z^{2n} = - \sum_n \binom{2n}{n} \frac{4^n z^{2n}}{2n-1}$$

所以解答是 $\binom{2n}{n} 4^{n-1} / (2n-1) + \binom{4n-1}{2n} / (4n-1)$.

5.47 用式(5.61), 它是 $(\mathcal{B}_r(z)^s / Q_r(z)) (\mathcal{B}_r(z)^{-s} / Q_r(z)) = 1 / Q_r(z)^2$ 中 z^n 的系数, 其中 $Q_r(z) = 1 - r + r \mathcal{B}_r(z)^{-1}$.

5.48 $F\left(2n+2, 1; n+2; \frac{1}{2}\right) = 2^{2n+1} / \binom{2n+1}{n+1}$, 式(5.111)的一种特殊情形.

5.49 Saalschutz 的等式(5.97)产生

$$\binom{x+n}{n} \frac{y}{y+n} F\left(\begin{matrix} -x, -n, -n-y \\ -x-n, 1-n-y \end{matrix} \middle| 1\right) = \frac{(y-x)^n}{(y+1)^n}.$$

5.50 左边为

$$\sum_{k \geq 0} \frac{a^{\bar{k}} b^{\bar{k}}}{c^{\bar{k}}} \frac{(-z)^k}{k!} \sum_{m \geq 0} \binom{k+a+m-1}{m} z^m = \sum_{n \geq 0} z^n \sum_{k \geq 0} \frac{a^{\bar{k}} b^{\bar{k}}}{c^{\bar{k}}} \frac{(-1)^k}{k!} \binom{n+a-1}{n-k}$$

且 z^n 的系数为

$$\binom{n+a-1}{n} F\left(\begin{matrix} a, b, -n \\ c, a \end{matrix} \middle| 1\right) \frac{a^{\bar{n}}}{n!} = \frac{(c-b)^{\bar{n}}}{c^{\bar{n}}}$$

(用 Vandermonde 卷积(5.92)).

5.51 (a) 反射给出 $F(a, -n; 2a; 2) = (-1)^n F(a, -n; 2a; 2)$. (顺便提到, 此公式意味着值得注意的等式 $\Delta^{2m+1} f(0) = 0$, 当 $f(n) = 2^n x^n / (2x)^n$ 时.)

(b) 逐项取极限是 $\sum_{0 \leq k \leq m} \binom{m}{k} \frac{2m+1}{2m+1-k} (-2)^k$ 加一个附加项(对 $k=2m-1$); 附加项是

$$\frac{(-m) \cdots (-1)(1) \cdots (m)(-2m+1) \cdots (-1) 2^{2m+1}}{(-2m) \cdots (-1)(2m-1)!} = (-1)^{m+1} \frac{m!}{(2m)!} \frac{m! 2^{2m+1}}{(2m)!}$$

$$= \frac{-2}{\binom{-1/2}{m}};$$

因此, 用式(5.104), 此极限是 $-1 / \binom{-1/2}{m}$, 我们所有的负值.

5.52 对于 $k > N$, 两个级数的项是零. 此等式对应于用 $N-k$ 替换 k . 注意

$$\begin{aligned} a^{\bar{N}} &= a^{\bar{N}-k} (a+N-k)^{\bar{k}} \\ &= a^{\bar{N}-k} (a+N-1)^{\bar{k}} = a^{\bar{N}-k} (1-a-N)^{\bar{k}} (-1)^k. \end{aligned}$$

5.53 当 $b = -\frac{1}{2}$ 时, 式(5.110)的左边是 $1-2z$, 右边是 $(1-4z+4z^2)^{1/2}$, 与 a 独立. 右边是正式的幂级数

$$1 + \binom{1/2}{1} 4z(z-1) + \binom{1/2}{2} 16z^2(z-1)^2 + \cdots,$$

能展开它且重新排列得到 $1-2z+0z^2+0z^3+\cdots$; 但是重排当 $z=1$ 时, 在它的中间步骤中涉及发散级数, 所以它是不合法的.

5.54 如果 $m+n$ 是奇数, 譬如说 $2N-1$, 我们要证明

$$\lim_{\varepsilon \rightarrow 0} F\left(\begin{matrix} N-m-\frac{1}{2}, -N+\varepsilon \\ -m+\varepsilon \end{matrix} \middle| 1\right) = 0.$$

由于 $-m+\varepsilon > -m-\frac{1}{2}+\varepsilon$, 应用方程(5.92), 且由于 $N \leq m$, 分母因子 $\Gamma(c-b)$

$= \Gamma(N-m)$ 是有限的, 其他因子是有限的. 否则 $m+n$ 是偶数; 置 $n = m - 2N$, 用式(5.93), 我们得到

$$\lim_{\varepsilon \rightarrow 0} F \left(\begin{matrix} -N, N-m-\frac{1}{2}+\varepsilon \\ -m+\varepsilon \end{matrix} \middle| 1 \right) = \frac{(n-1/2)^N}{m}.$$

剩下的工作是证明

$$\binom{m}{m-2N} \frac{(N-1/2)!}{(-1/2)!} \frac{(m-N)!}{m!} = \binom{m-N}{m-2N} 2^{-2N}$$

这是习题 22 的情形 $x = N$.

5.55 设 $Q(k) = (k+A_1) \cdots (k+A_M)Z$ 和 $R(k) = (k+B_1) \cdots (k+B_N)$. 于是 $t(k+1)/t(k) = P(k)Q(k-1)/P(k-1)R(k)$, 其中 $P(k) = Q(k) - R(k)$ 是一个非零多项式.

5.56 $-(k+1)(k+2) = s(k+1) + s(k)$ 的解是 $s(k) = -(1/2)k^2 - k - (1/4)$, 因

此 $\sum \binom{-3}{k} \delta k = \frac{1}{8} (-1)^{k-1} (2k^2 + 4k + 1) + C$, 而且

$$\begin{aligned} & (-1)^{k-1} \left\lfloor \frac{k+1}{2} \right\rfloor \left\lfloor \frac{k+2}{2} \right\rfloor \\ &= \frac{(-1)^{k-1}}{4} \left(k+1 - \frac{1+(-1)^k}{2} \right) \left(k+2 - \frac{1-(-1)^k}{2} \right) \\ &= \frac{(-1)^{k-1}}{8} (2k^2 + 4k + 1) + \frac{1}{8}. \end{aligned}$$

5.57 我们有 $t(k+1)/t(k) = (k-n)(k+1+\theta)(-z)/(k+1)(k+\theta)$. 所以我们设 $p(k) = k+\theta$, $q(k) = (k-n)(-z)$, $r(k) = k$. 隐蔽的函数 $s(k)$ 一定是一个常数 α_0 , 且我们有

$$k+\theta = (-z(k-n)-k)\alpha_0;$$

因此 $\alpha_0 = -1/(1+z)$, $\theta = -nz/(1+z)$. 和是

$$\sum \binom{n}{k} z^k \left(k - \frac{nz}{1+z} \right) \delta k = -\frac{n}{1+z} \binom{n-1}{k-1} z^k + C.$$

(特殊情形 $z=1$ 是式(5.18)中提到的, 一般情形等价于式(5.131).)

5.58 如果 $m > 0$, 我们能用 $\frac{k}{m} \binom{k-1}{m-1}$ 替换 $\binom{k}{m}$, 且推出公式 $T_{m,n} = \frac{n}{m} T_{m-1,n-1}$

$\frac{1}{m} \binom{n-1}{m}$. 所以求和因子 $\binom{n}{m}^{-1}$ 是合适的:

$$\frac{T_{m,n}}{\binom{n}{m}} = \frac{T_{m-1,n-1}}{\binom{n-1}{m-1}} - \frac{1}{m} + \frac{1}{n}.$$

我们能展开这个来取得

$$\frac{T_{m,n}}{\binom{n}{m}} = T_{0,n-m} - H_m + H_n - H_{n-m}.$$

最后 $T_{0,n-m} = H_{n-m}$, 所以 $T_{m,n} = \binom{n}{m}(H_n - H_m)$. (也可能用母函数来推出这个结果, 见 7.5 节中的例 2.)

$$5.59 \quad \sum_{j \geq 0, k \geq 1} \binom{n}{j} [j = \lfloor \log_m k \rfloor] = \sum_{j \geq 0, k \geq 1} \binom{n}{j} [m^j \leq k < m^{j+1}], \text{ 它是 } \sum_{j \geq 0} \binom{n}{j} (m^{j+1} - m^j) = (m-1) \sum_{j \geq 0} \binom{n}{j} m^j = (m-1)(m+1)^n.$$

$$5.60 \quad \binom{2n}{n} \approx 4^n / \sqrt{\pi n} \text{ 是}$$

$$\binom{m+n}{n} \approx \sqrt{\frac{1}{2\pi} \left(\frac{1}{m} + \frac{1}{n} \right)} \left(1 + \frac{m}{n} \right)^n \left(1 + \frac{n}{m} \right)^m$$

的情形 $m=n$.

5.61 设 $\lfloor n/p \rfloor = q$ 和 $n \bmod p = r$. 多项式等式 $(x+1)^n \equiv x^r + 1 \pmod{p}$ 意味着

$$(x+1)^{pq+r} \equiv (x+1)^r (x^p+1)^q \pmod{p}.$$

左边的 x^m 的系数是 $\binom{n}{m}$. 右边它是 $\sum_k \binom{r}{m-pk} \binom{q}{k}$, 也就是 $\binom{r}{m \bmod p}$ $\binom{q}{\lfloor m/p \rfloor}$, 因为 $0 \leq r < p$.

$$5.62 \quad \binom{np}{mp} = \sum_{k_1+\dots+k_p=np} \binom{p}{k_1} \dots \binom{p}{k_p} \equiv \binom{n}{m} \pmod{p^2}, \text{ 因为和的所有项是 } p^2$$

的倍数, 除了 $\binom{n}{m}$ 个项外, 它们中恰好 m 个 k 是等于 p . (Stanley^[275, 习题 1.6(d)] 证明当 $p > 3$ 时, 实际上对模 p^3 同余成立.)

$$5.63 \quad \text{这是 } S_n = \sum_{k=0}^n (-4)^k \binom{n+k}{n-k} = \sum_{k=0}^n (-4)^{n-k} \binom{2n-k}{k}. \text{ 式 (5.74) 的分母为零}$$

(当 $z = -1/4$ 时), 所以我们不能简单地代入公式. 递归 $S_n = -2S_{n-1} - S_{n-2}$ 引出解 $S_n = (-1)^n (2n+1)$.

$$5.64 \quad \sum_{k \geq 0} \left(\binom{n}{2k} + \binom{n}{2k+1} \right) / (k+1) = \sum_{k \geq 0} \binom{n+1}{2k+1} / (k+1), \text{ 它是}$$

$$\frac{2}{n+2} \sum_{k \geq 0} \binom{n+2}{2k+2} = \frac{2^{n+2} - 2}{n+2}.$$

5.65 两边用 n^{n-1} 相乘, 且用 $n-1-k$ 替换 k , 取得

$$\begin{aligned} \sum_k \binom{n-1}{k} n^k (n-k)! &= (n-1)! \sum_{k=0}^{n-1} (n^{k+1} / k! - n^k / (k-1)!) \\ &= (n-1)! n^n / (n-1)! \end{aligned}$$

(事实上, Gosper 算法发现了部分和。)用另一种方法, 把 $\binom{n}{k} k n^{n-1-k} k!$ 解释为 $\{1, \dots, n\}$ 到它自身的映射的个数, 同时 $f(1), \dots, f(k)$ 不同, 但是 $f(k+1) \in \{f(1), \dots, f(k)\}$; 在 k 上求和一定给出 n^n .

5.66 这是一个“走遍公园路径”问题, 其中每一步处仅有一种“明显的”方法来处理, 首先用 l 替换 $k-j$, 然后用 k 替换 $\lfloor \sqrt{l} \rfloor$, 取得

$$\sum_{k \geq 0} \binom{-1}{j-k} \binom{j}{m} \frac{2k+1}{2^j}.$$

因为 j 的一个多项式用 2^j 除超出指定 j 的项, 所以无穷级数收敛. 现在在 k 上求和, 取得

$$\sum_{j \geq 0} \binom{j}{m} \frac{j+1}{2^j}.$$

并 $j+1$ 且应用式(5.67), 取得解答 $4(m+1)$.

5.67 因为

$$\binom{\binom{k}{2}}{2} = 3 \binom{k+1}{4}$$

由式(5.26)得 $3 \binom{2n+2}{n+5}$.

5.68 用结论

$$\sum_{k \leq n/2} \binom{n}{k} = 2^{n-1} + \frac{1}{2} \binom{n}{n/2} [n \text{ 是偶数}],$$

我们取得 $n \left(2^{n-1} - \binom{n-1}{\lfloor n/2 \rfloor} \right)$.

5.69 由于 $\binom{k+1}{2} + \binom{l-1}{2} \leq \binom{k}{2} + \binom{l}{2} \Leftrightarrow k < l$, 当 k 尽可能相等时, 最小

出现。因此, 用第三章的均分公式, 最小是

$$\begin{aligned} & (n \bmod m) \binom{\lceil n/m \rceil}{2} + (n - (n \bmod m)) \binom{\lfloor n/m \rfloor}{2} \\ &= n \binom{\lfloor n/m \rfloor}{2} + (n \bmod m) \left\lfloor \frac{n}{m} \right\rfloor. \end{aligned}$$

对于用任何下部指标替换 2, 相似结果成立。

5.70 这是 $F\left(-n, \frac{1}{2}; 1, 2\right)$, 但是如果我们用 $n-k$ 替换 k , 它也是 $(-2)^{-n} \binom{2n}{n} \cdot F\left[-n, -n; \frac{1}{2} - n; \frac{1}{2}\right]$. 现在由 Gauss 等式 (5.111), $F\left(-n, -n; \frac{1}{2} - n; \frac{1}{2}\right) = F\left(-\frac{n}{2}, -\frac{n}{2}; \frac{1}{2} - n; 1\right)$. (用另一种方法, 用反射律 (5.101), 以及 Kummer 公式 (5.94) 把此联系到式 (5.55), $F\left(-n, -n; \frac{1}{2} - n; \frac{1}{2}\right) = 2^{-n} F\left(-n, \frac{1}{2}; \frac{1}{2} - n; -1\right)$.) 当 n 是奇数时, 解答为 0, 当 n 是偶数时, 解答为 $2^{-n} \binom{n}{n/2}$. (另一个推导见 [134, § 1.2], 在简单的查找算法的研究中出现此和 [164].)

$$5.71 \quad (a) \quad S(z) = \sum_{k \geq 0} a_k z^{m+k} / (1-z)^{m+2k+1} = z^m (1-z)^{-m-1} A(z/(1-z)^2).$$

(b) 这里 $A(z) = \sum_{k \geq 0} \binom{2k}{k} (-z)^k / (k+1) = (\sqrt{1+4z} - 1) / 2z$, 所以我们有 $A(z/(1-z)^2) = 1-z$. 因此 $S_n = [z^n](z/(1-z))^m = \binom{n-1}{n-m}$.

5.72 所说的量是 $m(m-n)\cdots(m-(k-1)n)n^{k-v(k)}/k!$. 任何 n 的系数因子 p 至少除尽分子 $k-v(k)$ 次, 且至多除尽分母 $k-v(k)$ 次, 因为这是 2 除尽 $k!$ 的次数. 一个素数 p 不除尽 n , 至少每当它除尽 $k!$ 时, 它一定除尽乘积 $m(m-n)\cdots(m-(k-1)n)$, 因为对所有 $r \geq 1$ 和所有 m , $m(m-n)\cdots(m-(p^r-1)n)$ 是 p^r 的一个倍数.

5.73 代入 $X_n = n!$ 产生 $\alpha = \beta = 1$; 代 $X_n = n!$ 产生 $\alpha = 1, \beta = 0$. 所以一般解是 $X_n = \alpha n! + \beta(n! - n!)$.

$$5.74 \quad \text{对 } 1 \leq k \leq n, \quad \binom{n+1}{k} - \binom{n-1}{k-1}.$$

5.75 递归 $S_k(n+1) = S_k(n) + S_{(k-1) \bmod 3}(n)$ 使它可能归纳地证实两个 S 是相等的, 以及 $S_{(-n) \bmod 3}(n)$ 与它们相差 $(-1)^n$. 这样 3 个值把它们的和 $S_0(n) + S_1(n) + S_2(n) = 2^n$ 分得尽可能相等, 所以一定有 $2^n \bmod 3$ 次出现 $\lfloor 2^n/3 \rfloor$, $3 - (2^n \bmod 3)$ 次出现 $\lceil 2^n/3 \rceil$.

$$5.76 \quad Q_{n,k} = (n+1) \binom{n}{k} - \binom{n}{k+1}.$$

5.77 当乘积是多项系数

$$\binom{k_m}{k_1, k_2 - k_1, \dots, k_m - k_{m-1}}$$

时, 除 $k_1 \leq \dots \leq k_m$ 外, 项为零, 所以 k_1, \dots, k_{m-1} 上的和是 m^{k_m} , 在 k_m 上的最后和产生 $(m^{n+1} - 1) / (m - 1)$.

5.78 把和延伸到 $k = 2m^2 + m - 1$, 新项是 $\binom{1}{4} + \binom{2}{6} + \dots + \binom{m-1}{2m} = 0$. 由于 $m \perp (2m+1)$, 对 $(k \bmod m; k \bmod (2m+1))$ 是不同的. 而且, 当 j 从 0 到 $2m$ 变化时, 数 $(2j+1) \bmod (2m+1)$ 是以某种次序排列 $0, 1, \dots, 2m$ 的数. 因此和是

$$\sum_{\substack{0 \leq k < n \\ 0 \leq j < 2m+1}} \binom{k}{j} = \sum_{0 \leq k < n} 2^k = 2^n - 1.$$

5.79 (a) 和是 2^{2n-1} , 所以 gcd 一定是 2 的一个幂. 如果 $n = 2^k q$, 其中 q 是奇数, 则 $\binom{2n}{1}$ 可被 2^{k+1} 除尽, 不被 2^{k+2} 除尽. 每个 $\binom{2n}{2j+1}$ 可被 2^{k+1} 除尽 (见习题 36), 所以这一定是 gcd.

(b) 如果当 $k = p' - 1$ 时, $p' \leq n+1 < p'^{+1}$, 我们通过加 k 到 $n-k$ 取得最多的基数 p 进位. 此时进位数是 $r - \varepsilon_p(n+1)$, 以及 $r = \varepsilon_p(L(n+1))$.

5.80 首先用归纳法证明 $k! \geq (k/e)^k$.

5.81 设 $f_{l, m, n}(x)$ 是左边. 证明我们有 $f_{l, m, n}(1) > 0$ 以及对于 $0 \leq x \leq 1$, $f'_{l, m, n}(x) < 0$ 就够了. 用式 (5.23), $f_{l, m, n}(1)$ 的值是 $(-1)^{n-m-1} \binom{l+m+\theta}{l+n}$, 且这是正的, 因为二项系数恰好有 $n-m-1$ 个负因子. 同样理由, 当 $l=0$ 时, 不等式是真的. 如果 $l>0$, 我们有 $f'_{l, m, n}(x) = -lf_{l-1, m, n+1}(x)$, 依据归纳法它是负的.

5.82 设 $\varepsilon_p(a)$ 是指数, 素数 p 除尽 a , 且设 $m = n - k$. 证明等式化为

$$\begin{aligned} & \min(\varepsilon_p(m) - \varepsilon_p(m+k), \varepsilon_p(m+k+1) - \varepsilon_p(k+1), \varepsilon_p(k) - \varepsilon_p(m+1)) \\ &= \min(\varepsilon_p(k) - \varepsilon_p(m+k), \varepsilon_p(m) - \varepsilon_p(k+1), \varepsilon_p(m+k+1) - \varepsilon_p(m+1)). \end{aligned}$$

为了简洁起见, 让我们把此记为 $\min(x_1, y_1, z_1) = \min(x_2, y_2, z_2)$. 注意到, $x_1 + y_1 + z_1 = x_2 + y_2 + z_2$. 一般关系

$$\varepsilon_p(a) < \varepsilon_p(b) \Rightarrow \varepsilon_p(a) = \varepsilon_p(|a \pm b|)$$

让我们推得 $x_1 \neq x_2 \Rightarrow \min(x_1, x_2) = 0$; 对 (y_1, y_2) 和 (z_1, z_2) , 也同样成立, 现在完成证明是件简单的事情了.

5.83 如果 $m < n$, 对于每个指定的 j , 量 $\binom{j+k}{j} \binom{m+n-j-k}{m-j}$ 是次数小于 n 的

k 的多项式; 因此 k 上的和是零. 如果 $m \geq n$ 以及如果 r 是范围 $n < r \leq m$ 中的一个整数, 则对于每个指定的 k , 量 $\binom{j+k}{k} \binom{m+n-j-k}{n-k}$ 是次数小于 r 的 j 的多项式; 因此 j 上的和是零. 如果 $m \geq n$ 以及如果对于 $0 \leq d < n$, $r = -d-1$ 是一个整数, 我们有 $\binom{r}{j}$

$= (-1)^j \binom{j+d}{j} = (-1)^j \sum_l \binom{j}{l} \binom{d}{l}$; 因此给定的和能写成

$$\begin{aligned} & \sum_{k,l} (-1)^k \binom{j+k}{j} \binom{d}{l} \binom{j}{l} \binom{n}{k} \binom{m+n-j-k}{m-j} \\ &= \sum_{k,l} (-1)^k \binom{n}{k} \binom{d}{l} \binom{l+k}{l} \binom{j+k}{j-l} \binom{m+n-j-k}{m-j} \\ &= \sum_{k,l} (-1)^{k+m+l} \binom{n}{k} \binom{d}{l} \binom{l+k}{l} \binom{-l-k-1}{j-l} \binom{-n+k-1}{m-j} \\ &= \sum_{k,l} (-1)^{k+m+l} \binom{n}{k} \binom{d}{l} \binom{l+k}{l} \binom{-l-n-2}{m-l}. \end{aligned}$$

由于 $\binom{l+k}{l}$ 是次数 $d < n$ 的 k 的多项式, 这是零.

如果 $m \geq n$, 对于 r 的 m 个不同值, 我们已证明了等式. 一般我们仅需考虑另外一种情形来证明它. 设 $r=0$, 则 $j=0$ 以及用式(5.25), 和是

$$\sum_k (-1)^k \binom{n}{k} \binom{m+n-k}{m} = \binom{m}{n}.$$

(是否有一个实际较短的证明?)

5.84 沿着提示, 我们取得

$$z \mathcal{B}_l(z)^{r-1} \mathcal{B}_l'(z) = \sum_{k \geq 0} \binom{rk+r}{k} \frac{kz^k}{rk+r}.$$

以及 $\mathcal{E}_l(z)$ 的相似公式. 于是公式 $(zt \mathcal{B}_l^{-1}(z) \mathcal{B}_l'(z) + 1) \mathcal{B}_l(z)'$ 和 $(zt \mathcal{E}_l^{-1}(z) \mathcal{E}_l'(z) + 1) \cdot \mathcal{E}_l(z)'$ 给出式(5.61)的各自的右边, 所以我们一定要证明

$$(zt \mathcal{B}_l^{-1}(z) \mathcal{B}_l'(z) + 1) \mathcal{B}_l(z)' = \frac{1}{1 - t + t \mathcal{B}_l(z)^{-1}},$$

$$(zt \mathcal{E}_l^{-1}(z) \mathcal{E}_l'(z) + 1) \mathcal{E}_l(z)' = \frac{1}{1 - zt \mathcal{E}_l(z)'},$$

根据式(5.59)得到这些结果.

5.85 如果 $f(x) = a_n x^n + \cdots + a_1 x + a_0$ 是次数 $\leq n$ 的任何多项式, 我们能归纳证明

$$\sum_{0 \leq \varepsilon_1, \dots, \varepsilon_n \leq 1} (-1)^{\varepsilon_1 + \dots + \varepsilon_n} f(\varepsilon_1 x_1 + \dots + \varepsilon_n x_n) = (-1)^n n! a_n x_1 \cdots x_n.$$

所述等式是特殊情形，其中 $a_n = 1/n!$ 和 $x_k = k^3$ 。

5.86 (a) 首先对所有 $i \neq j$ ，用 $n(n-1)$ 个指标变量 l_{ij} 展开。对 $1 \leq i < j < n$ ，置 $k_{ij} = l_{ij} - l_{ji}$ 且对于所有 $i < n$ 用约束 $\sum_{i+j=n} (l_{ii} - l_{ji}) = 0$ 让我们依据 Vandermonde 卷积对 $1 \leq j < n$ 实现在 l_{jn} 上求和，并且然后对 $1 \leq i < j < n$ 在 l_{ji} 上求和。

(b) $f(z) - 1$ 是次数 $< n$ 的多项式，有 n 个根，所以它一定是零。

(c) 考虑

$$\prod_{\substack{1 \leq i, j < n \\ i \neq j}} \left(1 - \frac{z_i}{z_j}\right)^{a_i} = \sum_{k=1}^n \prod_{\substack{1 \leq i, j < n \\ i \neq j}} \left(1 - \frac{z_i}{z_j}\right)^{a_i - (i=k)}$$

中的常数项。

5.87 用式(5.61)，第一项是 $\sum_k \binom{n-k}{k} z^{mk}$ ，第二项中的被加数是

$$\begin{aligned} & \frac{1}{m} \sum_{k \geq 0} \binom{(n+1)/m + (1+1/m)k}{k} (\zeta z)^{k+n+1} \\ &= \frac{1}{m} \sum_{k \geq n} \binom{(1+1/m)k - n - 1}{k - n - 1} (\zeta z)^k. \end{aligned}$$

由于 $\sum_{0 \leq k < m} (\zeta^{2j+1})^k = m(-1)^j [k = m/2]$ ，这些项共计

$$\begin{aligned} & \sum_{k \geq n/m} \binom{(1+1/m)mk - n - 1}{mk - n - 1} (-z^m)^k \\ &= \sum_{k \geq n/m} \binom{(n+1)k - n - 1}{k} (-z^m)^k = \sum_{k \geq n/m} \binom{n-mk}{k} z^{mk}. \end{aligned}$$

顺便提到，函数 $\mathcal{B}_m(z^m)$ 和 $\zeta^{2j+1} z \mathcal{B}_{1+1/m}(\zeta^{2j+1} z)^{1/m}$ 是方程 $w^{m+1} - w^m = z^m$ 的 $m+1$ 个复根。

5.88 利用 $\int_0^\infty (e^{-t} - e^{-nt}) dt / t = \ln n$ 和 $(1 + e^{-t}) / t \leq 1$ 两个事实。(依据式(5.83)，

当 $k \rightarrow \infty$ 时我们有 $\binom{x}{k} = O(k^{-x-1})$ ；所以这个界限意味着当 $x > -1$ 时，Stirling 级数

$\sum_k s_k \binom{x}{k}$ 收敛。Hermite^[155] 证明和是 $\ln \Gamma(1+x)$ 。)

5.89 依据二项定理，把这个加到式(5.19)，在两边产生 $y^{-x}(x+y)^{m+x}$ ，微分产生

$$\begin{aligned} & \sum_{k>m} \binom{m+r}{k} \binom{m-k}{n} x^k y^{m-k-n} \\ &= \sum_{k>m} \binom{-r}{k} \binom{m-k}{n} (-x)^k (x+y)^{m-k-n}, \end{aligned}$$

且我们用 $k+m+1$ 替换 k , 应用式(5.15)得到

$$\begin{aligned} & \sum_{k>0} \binom{m+r}{m+1+k} \binom{-n-1}{k} (-x)^{m+1+k} y^{-1-k-n} \\ &= \sum_{k>0} \binom{-r}{m+1+k} \binom{-n-1}{k} x^{m+1+k} (x+y)^{-1-k-n}. \end{aligned}$$

在超几何形式中, 此式化为

$$F\left(\begin{matrix} 1-r, n+1 \\ m+2 \end{matrix} \middle| \frac{-x}{y}\right) = \left(1+\frac{x}{y}\right)^{-n-1} F\left(\begin{matrix} m+1+r, n+1 \\ m+2 \end{matrix} \middle| \frac{x}{x+y}\right),$$

它是反射律(5.101)的特殊情形(a, b, c, z) = ($n+1, m+1+r, m+2, -x/y$). (因此式(5.105)与反射有关, 与习题 52 中的公式有关.)

5.90 如果 r 是一个非负整数, 和是有限的, 文中的推导成立, 只要分母中对于 $0 \leq k \leq r$ 没有和的项为零. 否则和是无限的, 且依据式(5.83), 第 k 项 $\binom{k-r-1}{k} / \binom{k-s-1}{k}$ 近似为 $k^{s-r} (-s-1)! / (-r-1)!$. 所以如果无限级数将收敛, 我们需要 $r > s+1$. (如果 r 和 s 是复数, 条件为 $\Re r > \Re s+1$, 因为 $|k^z| = k^{\Re z}$.) 依据式(5.92), 和为

$$F\left(\begin{matrix} -r, 1 \\ -s \end{matrix} \middle| 1\right) = \frac{\Gamma(r-s-1)\Gamma(-s)}{\Gamma(r-s)\Gamma(-s-1)} = \frac{s+1}{s+1-r};$$

当 r 和 s 是整数时, 这是我们找到的相同公式.

5.91 (对于这个最好用像 MACSYMA 的一个程序.) 顺便提到, 当 $c = (a+1)/2$ 时, 由于 Pfaff 的反射律, 化成等价于式(5.110)的一个等式. 如果 $w = -z/(1-z)$, 我们有 $4w(1-w) = -4z/(1-z)^2$, 以及

$$\begin{aligned} F\left(\begin{matrix} \frac{1}{2}a, \frac{1}{2}a+\frac{1}{2}-b \\ 1+a-b \end{matrix} \middle| 4w(1-w)\right) &= F\left(\begin{matrix} a, a+1-2b \\ 1+a-b \end{matrix} \middle| \frac{-z}{1-z}\right) \\ &= (1-z)^a F\left(\begin{matrix} a, b \\ 1+a-b \end{matrix} \middle| z\right). \end{aligned}$$

5.92 像 Clausen 在 150 年之前证明它们那样, 通过证明两边满足相同的微分方程, 能证明这些等式. 一种方法是依据二项系数写出 z^n 的系数之间得到的方程:

$$\begin{aligned} \sum_k \frac{\binom{r}{k} \binom{s}{k} \binom{r}{n-k} \binom{s}{n-k}}{\binom{r+s-1/2}{k} \binom{r+s-1/2}{n-k}} &= \frac{\binom{2r}{n} \binom{r+s}{n} \binom{2s}{n}}{\binom{2r+2s}{n} \binom{r+s-1/2}{n}}; \\ \sum_k \frac{\binom{-1/4+r}{k} \binom{-1/4+s}{k} \binom{-1/4-r}{n-k} \binom{-1/4-s}{n-k}}{\binom{-1+r+s}{k} \binom{-1-r-s}{n-k}} \\ &= \frac{\binom{-1/2}{n} \binom{-1/2+r-s}{n} \binom{-1/2-r+s}{n}}{\binom{-1+r+s}{n} \binom{-1-r-s}{n}}. \end{aligned}$$

另一种方法是依据超几何:

$$\begin{aligned} F\left(\begin{matrix} a, b, \frac{1}{2}-a-b-n, -n \\ \frac{1}{2}+a+b, 1-a-n, 1-b-n \end{matrix} \middle| 1\right) &= \frac{(2a)^{\bar{n}} (a+b)^{\bar{n}} (2b)^{\bar{n}}}{(2a+2b)^{\bar{n}} a^{\bar{n}} b^{\bar{n}}}; \\ F\left(\begin{matrix} \frac{1}{4}+a, \frac{1}{4}+b, a+b-n, -n \\ 1+a+b, \frac{3}{4}+a-n, \frac{3}{4}+b-n \end{matrix} \middle| 1\right) \\ &= \frac{(1/2)^{\bar{n}} (1/2+a-b)^{\bar{n}} (1/2-a+b)^{\bar{n}}}{(1+a+b)^{\bar{n}} (1/4-a)^{\bar{n}} (1/4-b)^{\bar{n}}}. \end{aligned}$$

5.93 $\alpha^{-1} \prod_{j=1}^k (f(j) + \alpha) / f(j)$. (当 f 是次数 2 的多项式时的特殊情形等价于等式 (5.13).)

5.94 这是 Henrici 的“支持巨大”等式的一个结论,

$$\begin{aligned} f(a, z) f(a, \omega z) f(a, \omega^2 z) \\ = F\left(\begin{matrix} \frac{1}{2}a - \frac{1}{4}, \frac{1}{2}a + \frac{1}{4} \\ \frac{1}{3}a, \frac{1}{3}a + \frac{1}{3}, \frac{1}{3}a + \frac{2}{3}, \frac{2}{3}a - \frac{1}{3}, \frac{2}{3}a, \frac{2}{3}a + \frac{1}{3}, a \end{matrix} \middle| \left(\frac{4z}{9}\right)^3\right), \end{aligned}$$

其中 $f(a, z) = F(; a; z)$. 通过证明两边满足相同微分方程能证明这个等式. 如果我们用 $3n+1$ 或 $3n+2$ 替换 $3n$, 给出的和是零.

5.95 部分结果见[78]. V. A. Vyssotsky 完成了计算机试验.

5.96 按照 Sarkozy^[256], 所有大的 n 具有所说的性质. 事实上, Paul Erdos 猜测当 $n \rightarrow \infty$ 时 $\max_p \varepsilon_p \left(\binom{2n}{n} \right)$ 趋于无限.

5.97 如果 $2n+1$ 是素数, 同余一定成立. 当 $2n+1 = 3 \cdot 11 \cdot 179$ 时, Steven Skiena 还得到例子 $n = 2953$.

6.1 2314, 2431, 3241, 1342, 3124, 4132, 4213, 1423, 2143, 3412, 4321.

6.2 $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} m^k$, 因为每一个这样的函数把它的域划分成 k 个非空子集, 且有 m^k 种方式把函数值赋给每个划分. (k 上求和给出式(6.10)的一个组合证明.)

6.3 现在 $d_{k+1} \leq (\text{重心}) - \varepsilon = 1 - \varepsilon + (d_1 + \cdots + d_k)/k$. 这个递归像式(6.55), 但是用 $1 - \varepsilon$ 代替 1; 因此最优解是 $d_{k+1} = (1 - \varepsilon)H_k$. 只要 $\varepsilon < 1$, 这是无界的.

6.4 $H_{2n+1} - (1/2)H_n$. (类似 $\sum_{k=1}^{2n} (-1)^{k-1}/k = H_{2n} - H_n$.)

6.5 $U_n(x, y)$ 等于

$$x \sum_{k \geq 1} \binom{n}{k} (-1)^{k-1} k^{-1} (x+ky)^{n-1} + y \sum_{k \geq 1} \binom{n}{k} (-1)^{k-1} (x+ky)^{n-1},$$

且第一个和是 $U_{n-1}(x, y) + \sum_{k \geq 1} \binom{n-1}{k-1} (-1)^{k-1} k^{-1} (x+ky)^{n-1}$. 余下的 k^{-1} 可被吸收,

且我们有 $\sum_{k \geq 1} \binom{n}{k} (-1)^{k-1} (x+ky)^{n-1} = x^{n-1} + \sum_{k \geq 0} \binom{n}{k} (-1)^{k-1} (x+ky)^{n-1} = \dots^{n-1}$.

这证明了式(6.75). 设 $R(x, y) = x^{-n} U_n(x, y)$, 于是 $R_0(x, y) = 0$ 和 $R_n(x, y) = R_{n-1}(x, y) + 1/n + y/x$, 因此 $R_n(x, y) = H_n + ny/x$. (顺便提到, 原来和 $U_n = U_n(n, -1)$ 导不出像这样的一个递归, 所以更一般的和(它对 n 的依赖关系移去 x) 比它的特殊情形更容易来归纳地解, 这是另一个启发性的例子, 其中一个强归纳假设形成成功和失败之间的区别.)

6.6 一个月的月末出现的一对幼兔 bb , 在下一个月的月末变成一对成熟的兔 aa , 且每对 aa 变成一对 aa 和一对 bb . 因此每对 bb 相当于蜂树中的一个雄蜂, 每对 aa 相当于一个蜂王, 除了当兔子向前时恰好蜂树向后. n 个月之后有 F_{n+1} 对兔子, F_n 是成熟的兔, F_{n-1} 是幼兔. (这就是 Fibonacci 原先引入他的数的来龙去脉.)

6.7 (a) 置 $k = 1 - n$ 且应用式(6.107). (b) 置 $m = 1$ 和 $k = n - 1$, 且应用式(6.128).

6.8 $55 + 8 + 2$ 变成 $89 + 13 + 3 = 105$, 真值是 $104.607\ 361$.

6.9 21. (当单位被平方时, 我们从 F_n 到 F_{n+2} . 准确的解答接近 20.72.)

6.10 部分商 a_0, a_1, a_2, \dots 全等于 1, 因为 $\varphi = 1 + 1/\varphi$. (所以 Stern-Brocot 表示是 $RLRLRLRL\dots$.)

6.11 $(-1)^{\bar{n}} = [n=0] - [n=1]$, 见式(6.11).

6.12 这是式(6.31)的一个推论, 且它的对偶在表 6.3 中.

6.13 依据习题 12, 两个公式是等价的, 我们可用归纳法. 或者我们能看到, $z^n D^n$ 应用于 $f(z) = z^x$ 给出 $x^n z^x$, 而 θ^n 应用于相同函数给出 $x^n z^x$; 所以像 $\langle x^0, x^1, x^2, \dots \rangle$ 与 $\langle x^0, x^1, x^2, \dots \rangle$ 有关那样, 序列 $\langle \theta^0, \theta^1, \theta^2, \dots \rangle$ 一定和 $\langle z^0 D^0, z^1 D^1, z^2 D^2, \dots \rangle$ 有关.

6.14 我们有

$$x \binom{x+k}{n} = (k+1) \binom{x+k}{n+1} + (n-k) \binom{x+k+1}{n+1},$$

因为 $(n+1)x = (k+1)(x+k-n) + (n-k)(x+k+1)$ 。(当 $k=0$, $k=-1$ 和 $k=n$ 时, 验证后面的等式就够了。)

6.15 由于 $\Delta \left(\binom{x+k}{n} \right) = \binom{x+k}{n-1}$, 我们有一般公式

$$\sum_k \binom{n}{k} \binom{x+k}{n-m} = \Delta^m (x^n) = \sum_j \binom{m}{j} (-1)^{m-j} (x+j)^n.$$

置 $x=0$ 且求助于式(6.19),

$$6.16 \quad A_{n,k} = \sum_{j \geq 0} a_j \left\{ \binom{n-j}{k} \right\}, \text{ 此和总是有限的.}$$

$$6.17 \quad (a) \left| \binom{n}{k} \right| = \left[\binom{n+1}{n+1-k} \right]. \quad (b) \left| \binom{n}{k} \right| = n^{\overline{n-k}} = n! \lfloor n \geq k \rfloor / k!.$$

$$(c) \left| \binom{n}{k} \right| = k! \left\{ \binom{n}{k} \right\}.$$

6.18 这等价于式(6.3)或(6.8)。(特别, 由此得到当 $n > 1$ 时 $\sigma_n(1) = -n\sigma_n(0) = B_n / n!$ 。)

6.19 用表 6.8。

$$6.20 \quad \sum_{1 \leq j \leq k \leq n} 1/j^2 = \sum_{1 \leq j \leq n} (n+1-j)/j^2 = (n+1)H_n^{(2)} - H_n.$$

6.21 提示的数是具有奇分母的分数的一个和, 所以它有 a 和 b 为奇数的形式 a/b 。(顺便提到, Bertrand 的假定意味着当 $n > 2$ 时, 至少一个奇素数也可除尽 b_n 。)

6.22 当 $k > 2|z|$ 时, $|z/k(k+z)| \leq 2|z|/k^2$, 所以当分母不为零时, 明确定义了和, 如果 $z=n$ 我们有 $\sum_{k=1}^m (1/k - 1/(k+n)) = H_m - H_{m+n} + H_n$, 当 $m \rightarrow \infty$ 时它趋于 H_n 。(常称量 $H_{z-1} - \gamma$ 为 $\psi(z)$ 。)

$$6.23 \quad z/(e^z + 1) = z/(e^z - 1) - 2z/(e^{2z} - 1) = \sum_{n \geq 0} (1 - 2^n) B_n z^n / n!.$$

6.24 当 n 是奇数时, $T_n(x)$ 是 x^2 的一个多项式, 因此当我们形成微商和用式(6.95)计算 $T_{n+1}(x)$ 时, 偶数乘它的系数。(事实上我们还能证明: 依据习题 54, 在它的分母中对于第一个幂 Bernoulli 数 B_{2n} 总有 2, 因此 $2^{2n-k} \setminus T_{2n+1} \Leftrightarrow 2^k \setminus (n+1)$ 。奇正整数 $(n+1)T_{2n+1}/2^{2n}$ 称为 Genocchi 数 $\langle 1, 1, 3, 17, 155, 2073, \dots \rangle$, 为了纪念 Genocchi^[11]。)

6.25 $100n - nH_n < 100(n-1) - (n-1)H_{n-1} \Leftrightarrow H_{n-1} > 99$ 。(当他在 $N \approx e^{100-\gamma}$ 处结束, 最小这样的 n 近似于 $e^{99-\gamma}$, 差不多 e 倍那么长。所以在他的旅程的最后 63% 期间, 他变得较接近。)

$$6.26 \quad \text{设 } u(k) = H_{k-1} \text{ 和 } \Delta v(k) = 1/k, \text{ 以致 } u(k) = v(k). \text{ 于是我们有 } S_n = H_n^{(2)}$$

$$= \sum_{k=1}^n H_{k-1} / k = H_{k-1}^2 \Big|_1^{n+1} - S_n = H_n^2 - S_n.$$

6.27 看到依据式(6.108), 当 $m > n$ 时我们有 $\gcd(F_m, F_n) = \gcd(F_{m-n}, F_n)$. 这就产生一个归纳法的证明.

6.28 (a) $Q_n = \alpha(L_n - F_n)/2 + \beta F_n$. (解还能写成 $Q_n = \alpha F_{n-1} + \beta F_n$.)

(b) $L_n = \varphi^n + \hat{\varphi}^n$.

6.29 当 $k=0$ 时, 等式是式(6.133). 当 $k=1$ 时, 它实质上是

$$K(x_1, \dots, x_n)x_m = K(x_1, \dots, x_m)K(x_m, \dots, x_n) \\ - K(x_1, \dots, x_{m-2})K(x_{m+2}, \dots, x_n);$$

就 Morse 码项而论, 右边第二个乘积去掉第一个乘积具有交叉的长划的情形. 当 $k > 1$ 时, 用式(6.127)和(6.132)对 k 归纳就够了. (如果我们采用 $K_{-1} = 0$ 的约定, 当 K 的一个或多个下标变成 -1 时, 等式亦真. 当乘法不可交换时, 如果我们把它写成形式

$$K_{m+n}(x_1, \dots, x_{m+n})K_k(x_{m+k}, \dots, x_{m+1}) \\ = K_{m+k}(x_1, \dots, x_{m+k})K_n(x_{m+n}, \dots, x_{m+1}) \\ + (-1)^k K_{m-1}(x_1, \dots, x_{m-1})K_{m-k-1}(x_{m+1}, \dots, x_{m+k+2}).$$

例如, 对于情形 $k=2, m=0, n=3$, 我们得到有点意外的非交换因子分解

$$(abc + a + c)(1 + ba) = (ab + 1)(cba + a + c).$$

6.30 $K(x_1, \dots, x_n)$ 关于 x_m 的导数是

$$K(x_1, \dots, x_{m-1})K(x_{m+1}, \dots, x_n),$$

且第二个导数为零; 因此解答是

$$K(x_1, \dots, x_n) + K(x_1, \dots, x_{m-1})K(x_{m+1}, \dots, x_n)y.$$

6.31 由于 $x^{\bar{n}} = (x + n - 1)^{\bar{n}} = \sum_k \binom{n}{k} x^{\bar{k}} (n-1)^{\bar{n-k}}$, 我们有 $\left| \frac{n}{k} \right| = \binom{n}{k} (n-1)^{\bar{n-k}}$.

顺便提到, 这些系数满足递归

$$\left| \frac{n}{k} \right| = (n-1+k) \left| \frac{n-1}{k} \right| + \left| \frac{n-1}{k-1} \right|, \text{ 整数 } n, k > 0.$$

6.32 $\sum_{k \leq m} k \left\{ \frac{n+k}{k} \right\} = \left\{ \frac{m+n+1}{m} \right\}$ 和 $\sum_{0 \leq k \leq n} \left\{ \frac{k}{m} \right\} (m+1)^{n-k} = \left\{ \frac{n+1}{m+1} \right\}$, 两者在

表 6.4 中出现.

6.33 如果 $n > 0$, 依据式(6.71), 我们有 $\left[\frac{n}{3} \right] = (1/2)(n-1)! \left(H_{n-1}^2 - H_{n-1}^{(2)} \right)$; 依

据式(0.19), $\left\{ \frac{n}{3} \right\} = (1/6)(3^n - 3 \cdot 2^n + 3)$.

6.34 我们有 $\langle \frac{-1}{k} \rangle = 1/(k+1)$, $\langle \frac{-2}{k} \rangle = H_{k+1}^{(2)}$, 一般对于所有整数 n , 由式(6.38)给出 $\langle \frac{n}{k} \rangle$.

6.35 设 n 是使得 $\lfloor H_n \rfloor \geq \lfloor H_{n-1} \rfloor$ 的最小整数 $> 1/\varepsilon$.

6.36 现在 $d_{k+1} = (100 + (1 + d_1) + \cdots + (1 + d_k)) / (100 + k)$, 解是 $d_{k+1} = H_{k+100} - H_{101} + 1$ (对 $k \geq 1$). 当 $k \geq 176$ 时这超过 2.

6.37 和 (用分部) 是 $H_{mn} - \left(\frac{m}{m} + \frac{m}{2m} + \cdots + \frac{m}{mn} \right) = H_{mn} - H_n$, 所以无限和是 $\ln m$. (由此得到

$$\sum_{k \geq 1} \frac{v_m(k)}{k(k+1)} = \frac{m}{m-1} \ln m,$$

因为 $v_m(k) = (m-1) \sum_{i \geq 1} (k \bmod m^i) / m^i$.)

6.38 $(-1)^k \left(\binom{r-1}{k} r^{-1} - \binom{r-1}{k-1} H_k \right) + C$. (依据分部, 用式(5.16).)

6.39 把它写成 $\sum_{1 \leq i \leq n} i^{-1} \sum_{1 \leq k \leq n} H_k$. 且通过式(6.67)首先对 k 求和, 取得

$$(n+1)H_n^2 - (2n+1)H_n + 2n.$$

6.40 如果 $6n-1$ 是素数

$$\sum_{k=1}^{4n-1} \frac{(-1)^{k-1}}{k} = H_{4n-1} - H_{2n-1}$$

的分子可被 $6n-1$ 除尽, 因为和是

$$\sum_{k=2n}^{4n-1} \frac{1}{k} = \sum_{k=2n}^{3n-1} \left(\frac{1}{k} + \frac{1}{6n-1-k} \right) = \sum_{k=2n}^{3n-1} \frac{6n-1}{k(6n-1-k)}.$$

类似, 如果 $6n+1$ 是素数, $\sum_{k=1}^{4n} (-1)^{k-1} / k = H_{4n} - H_{2n}$ 的分子是 $6n+1$ 的一个倍数.

1987 年以来我们加到 $k=1324$.

6.41 $S_{n+1} = \sum_k \binom{\lfloor (n+1+k)/2 \rfloor}{k} = \sum_k \binom{\lfloor (n+k)/2 \rfloor}{k-1}$, 因此我们有 $S_{n+1} + S_n = \sum_k \binom{\lfloor (n+k)/2 \rfloor + 1}{k} = S_{n+2}$. 解答为 F_{n+2} .

6.42 F_n .

6.43 在 $\sum_{n \geq 0} F_n z^n = z / (1 - z - z^2)$ 中置 $z = 1/10$ 取得 $10/89$. 和是一个具有周期长 44 的循环十进制数:

0.112 359 550 561 797 752 808 988 764 044 943 820 224 719 101 123 595 5 + .

6.44 如果必要, 用 $(-m, -k)$ 或 $(k, -m)$ 或 $(-k, m)$ 替换 (m, k) , 以致 $m \geq k \geq 0$. 如果 $m = k$, 结果是清楚的. 如果 $m > k$, 我们能用 $(m-k, m)$ 替换 (m, k) 且用归纳法.

6.45 $X_n = A(n)\alpha + B(n)\beta + C(n)\gamma + D(n)\delta$, 其中 $B(n) = F_n$, $A(n) = F_{n-1}$, $A(n) + B(n) - D(n) = 1$, 和 $B(n) - C(n) + 3D(n) = n$.

6.46 $\varphi/2$ 和 $\varphi^{-1}/2$. 设 $u = \cos 72^\circ$ 和 $v = \cos 36^\circ$, 则 $u = 2v^2 - 1$ 和 $v = 1 - 2\sin^2 18^\circ = 1 - 2u^2$. 因此 $u + v = 2(u+v)(v-u)$, $4v^2 - 2v - 1 = 0$. 我们能作此研究找到 5 个单位的复根:

$$1, \frac{\varphi^{-1} \pm i\sqrt{2+\varphi}}{2}, \frac{-\varphi \pm i\sqrt{3-\varphi}}{2}.$$

6.47 $2^n \sqrt{5} F_n = (1 + \sqrt{5})^n - (1 - \sqrt{5})^n$, $\sqrt{5}$ 的偶次幂消去. 现在设 p 是奇素数. 则除了当 $k = (p-1)/2$ 时, $\binom{p}{2k+1} \equiv 0$, 除了当 $k = 0$ 或 $k = (p-1)/2$ 时, $\binom{p+1}{2k+1} \equiv 0$; 因此 $F_p \equiv 5^{(p-1)/2}$ 和 $2F_{p+1} \equiv 1 + 5^{(p-1)/2} \pmod{p}$. 能证当 p 有形式 $10k \pm 1$ 时, $5^{(p-1)/2} \equiv 1$; 当 p 有形式 $10k \pm 3$ 时, $5^{(p-1)/2} \equiv -1$.

6.48 这一定是真的, 因为式(6.138)是多项式等式且我们能置 $a_m = 0$.

6.49 在式(6.146)中置 $z = 1/2$; 部分商为 $0, 2^{F_0}, 2^{F_1}, 2^{F_2}, \dots$. (Knuth^[172]注意到此数是超越数.)

6.50 (a) $f(n)$ 是偶 $\Leftrightarrow 3 \nmid n$.

(b) 如果 n 的二进制表示为 $(1^{a_1} 0^{a_2} \dots 1^{a_m} 0^{a_n})_2$, 其中 m 是偶数, 我们有 $f(n) = K(a_1, a_2, \dots, a_{m-1})$.

6.51 (a) 如果我们把 1 加到每个元素(模 p), 组合证明: $\{1, 2, \dots, p\}$ 转化为 k 个子集或轮换的排列被分为各有 1 或 p 排列的“轨道”. 例如,

$$\begin{aligned} \{1, 2, 4\} \cup \{3, 5\} &\rightarrow \{2, 3, 5\} \cup \{4, 1\} \rightarrow \{3, 4, 1\} \cup \{5, 2\} \rightarrow \{4, 5, 2\} \cup \{1, 3\} \\ &\rightarrow \{5, 1, 3\} \cup \{2, 4\} \rightarrow \{1, 2, 4\} \cup \{3, 5\}. \end{aligned}$$

当此变换使得一个排列转成自身时, 我们仅取得大小 1 的一个轨道; 然而 $k=1$ 或 $k=p$.

另一方面, 有一个代数证明: 我们有 $x^p \equiv x^{\frac{p}{2}} + x^{\frac{1}{2}}$ 和 $\alpha^{\frac{p}{2}} \equiv x^{\frac{p}{2}} - x \pmod{p}$, 因为 Fermat 定理告诉我们, $(x-0)(x-1)\dots(x-(p-1))$ 可除尽 $x^p - x$.

(b) 依据(a)和 Wilson 定理得此结果, 或者用 $x^{p-1} \equiv x^{\frac{p}{2}} / (x-1) \equiv (x^{\frac{p}{2}} - x) / (x-1) = x^{p-1} + x^{p-2} + \dots + x$.

(c) 我们 有 $\left\{ \begin{smallmatrix} p+1 \\ k \end{smallmatrix} \right\} \equiv \left[\begin{smallmatrix} p+1 \\ k \end{smallmatrix} \right] \equiv 0 \ (3 \leq k \leq p)$, 于是对 $4 \leq k \leq p$, $\left\{ \begin{smallmatrix} p+2 \\ k \end{smallmatrix} \right\} \equiv \left[\begin{smallmatrix} p+2 \\ k \end{smallmatrix} \right] \equiv 0$, 等等。(类似, 我们有 $\left[\begin{smallmatrix} 2p-1 \\ p \end{smallmatrix} \right] \equiv - \left\{ \begin{smallmatrix} 2p-1 \\ p \end{smallmatrix} \right\} \equiv 1$.)

(d) $p! = p^p = \sum_k (-1)^{p-k} p^k \left[\begin{smallmatrix} p \\ k \end{smallmatrix} \right] = p^p \left[\begin{smallmatrix} p \\ p \end{smallmatrix} \right] - p^{p-1} \left[\begin{smallmatrix} p \\ p-1 \end{smallmatrix} \right] + \cdots + p^3 \left[\begin{smallmatrix} p \\ 3 \end{smallmatrix} \right] - p^2 \left[\begin{smallmatrix} p \\ 2 \end{smallmatrix} \right] + p \left[\begin{smallmatrix} p \\ 1 \end{smallmatrix} \right]$. 但是 $p \left[\begin{smallmatrix} p \\ 1 \end{smallmatrix} \right] = p!$, 所以

$$\left[\begin{smallmatrix} p \\ 2 \end{smallmatrix} \right] = p \left[\begin{smallmatrix} p \\ 3 \end{smallmatrix} \right] - p^2 \left[\begin{smallmatrix} p \\ 4 \end{smallmatrix} \right] + \cdots + p^{p-2} \left[\begin{smallmatrix} p \\ p \end{smallmatrix} \right]$$

是 p^2 的倍数。(称它为 Wolstenholme 定理.)

6.52 (a) 看出 $H_n = H_n^* + H_{[n/p]} / p$, 其中 $H_n^* = \sum_{k=1}^n (k \perp p) / k$.

(b) 作 mod 5, 我们对 $0 \leq r \leq 4$ 有 $H_r = \langle 0, 1, 4, 1, 0 \rangle$. 因此第一个解是 $n=4$. 依据 (a) 我们知道 $5 \setminus a_n \Rightarrow 5 \setminus a_{[n/5]}$; 所以当我们有 $H_n = H_n^* + (1/5)H_4 = H_{20}^* + (1/5)H_4 + H_r + \sum_{k=1}^r 20/k(20+k)$ 时, 下一个可能的范围是 $n=20+r$, $0 \leq r \leq 4$. 像 H_4 的分子那样, 25 可除尽 H_{20}^* 的分子. 因此此范围中的仅有解是 $n=20$ 和 $n=24$. 下一个可能范围是 $n=100+r$; 现在 $H_n = H_n^* + (1/5)H_{20}$, 它是 $(1/5)H_{20} + H_r$ 加上分子是 5 的倍数的一个分数. 如果 $(1/5)H_{20} \equiv m \pmod{5}$, 其中 m 是一个整数, 调和数 H_{100+r} 将有一个被 5 除尽的分子, 当且仅当 $m + H_r \equiv 0 \pmod{5}$; 因此 m 一定是 $\equiv 0, 1$ 或 4 . 作 mod 5 我们发现 $(1/5)H_{20} = (1/5)H_{20}^* + (1/25)H_4 - (1/25)H_4 = 1/12 \equiv 3$, 因此对 $100 \leq n \leq 104$ 无解. 类似对于 $120 \leq n \leq 124$ 无解; 我们已找到所有的 3 个解.

(依据习题 6.51(d), 如果 p 是任何素数 ≥ 5 , 我们总有 $p^2 \setminus a_{p-1}$, $p \setminus a_{p^2-p}$ 和 $p \setminus a_{p^2-1}$. 刚给的论证表明这些是对 $p \setminus a_n$ 仅有的解, 当且仅当对于 $0 \leq r < p$, $p^{-2}H_{p-1} + H_r \equiv 0 \pmod{p}$ 无解. 后面条件不仅对 $p=5$ 成立, 还对 $p=13, 17, 23, 41$ 和 67 也成立——也许对无限多个素数成立. 仅当 $n=2, 7$ 和 22 时, 3 除尽 H_n 的分子; 仅当 $n=6, 42, 48, 295, 299, 337, 341, 2096, 2390, 14675, 16731, 16735$ 和 102728 时, 7 除尽它.)

6.53 用分部求和产生

$$\frac{n+1}{(n+2)^2} \left(\frac{(-1)^m}{\binom{n+1}{m+1}} ((n+2)H_{m+1} - 1) - 1 \right).$$

6.54 (a) 如果 $m \geq p$, 我们有 $S_m(p) \equiv S_{m-(p-1)}(p) \pmod{p}$, 因为当 $1 \leq k < p$ 时 $k^{p-1} \equiv 1$.

同样 $S_{p-1}(p) \equiv p-1 \equiv -1$. 如果 $0 < m < p-1$, 我们能写下

$$S_m(p) = \sum_{j=0}^m \left[\begin{matrix} m \\ j \end{matrix} \right] (-1)^{m-j} \sum_{k=0}^{p-1} k^j = \sum_{j=0}^m \left[\begin{matrix} m \\ j \end{matrix} \right] (-1)^{m-j} \frac{p^{j+1}}{j+1} \equiv 0.$$

(b) 提示中的条件意味着 I_{2n} 的分母不被任何素数 p 除尽, 因此 I_{2n} 一定是一个整数. 为了证明提示, 我们可设 $n > 1$. 则依据式(6.78), (6.84)和(a),

$$B_{2n} + \frac{[(p-1) \setminus (2n)]}{p} + \sum_{k=0}^{2n-2} \binom{2n+1}{k} B_k \frac{p^{2n-k}}{2n+1}$$

是整数, 所以我们要验证, 分数 $\binom{2n+1}{k} B_k p^{2n-k} / (2n+1) = \binom{2n}{k} B_k p^{2n-k} / (2n-k+1)$ 没有被 p 除尽的分母. $\binom{2n}{k} B_k p$ 的分母不被 p 除尽, 因为 B_k 在它的分母中无 p^2 (用归纳法); $p^{2n-k-1} / (2n-k+1)$ 的分母不被 p 除尽, 因为当 $k \leq 2n-2$ 时 $2n-k+1 < p^{2n-k}$, 证毕. ([185]中用表格表出数 I_{2n} . 1875年^[153]Hermite 通过 I_{18} 计算它们, 结果为 $I_2 = I_4 = I_6 = I_8 = I_{10} = I_{12} = 1$; 因此课文中表示的 Bernoulli 数实际有一种“简单的”型式, 包括 $\frac{-691}{2730}$ (!). 但是当 $n > 6$ 时, 数 I_{2n} 看来没有任何重大的特性. 例如, $B_{24} = -86579 \frac{1}{2} \frac{1}{3} \frac{1}{5} \frac{1}{7} \frac{1}{13}$, 且 86579 是素数.

(c) 数 $2-1$ 和 $3-1$ 总除尽 $2n$, 如果 n 是素数, $2n$ 仅有的因子是 1, 2, n 和 $2n$, 所以除非 $2n+1$ 也是素数, 对于素数 $n > 2$ 的 B_{2n} 的分母将是 6. 在后一种情形, 我们能试 $4n+3, 8n+7, \dots$, 直到我们最终找到一个非素数(因为 n 除尽 $2^{n-1}n+2^{n-1}-1$). (这个证明不需要更困难的定理, 即有无限多个形式 $6k+1$ 的素数.) 当 n 有非素数值时, B_{2n} 的分母也能是 6, 像 49 那样.

6.55 用 Vandermonde 卷积, 所说的和为 $\frac{m+1}{x+m+1} \binom{x+n}{n} \binom{n}{m+1}$. 为了取得式(6.70), 微分且置 $x=0$.

6.56 像在式(6.72)的推导中出现的简化那样, 首先用 $((k-m)+m)n+1$ 替换 k^{n+1} , 且展开成 $k-m$ 的幂. 如果 $m > n$ 或 $m < 0$, 解答为 $(-1)^n n! - m^n / \binom{n-m}{n}$. 否则我们需要取, 当 $x \rightarrow -m$ 时式(5.51)的极限减 $k=m$ 的项, 解答归结为 $(-1)^n n! + (-1)^{m+1} \binom{n}{m} m^n (n+1+mH_{n-m}-mH_m)$.

6.57 首先用归纳法证明, 第 n 行至多包含 3 个不同值 $A_n \geq B_n \geq C_n$; 如果 n 是偶, 它们在循环次序 $[C_n, B_n, A_n, B_n, C_n]$ 中出现, 而如果 n 是奇, 它们在循环次序 $[C_n, B_n, A_n, A_n, B_n]$ 中出现. 此外

$$\begin{aligned} A_{2n+1} &= A_{2n} + B_{2n}; & A_{2n} &= 2A_{2n-1}; \\ B_{2n+1} &= B_{2n} + C_{2n}; & B_{2n} &= A_{2n-1} + B_{2n-1}; \end{aligned}$$

$$C_{2n+1} = 2C_{2n}; \quad C_{2n} = B_{2n-1} + C_{2n-1}.$$

由此得到 $Q_n = A_n - C_n = F_{n+1}$. (对于阶 3 的环绕二项系数见习题 5.75.)

$$6.58 \quad (a) \sum_{n \geq 0} F_n^2 z^n = z(1-z)/(1+z)(1-3z+z^2) = (1/5)((2-3z)/(1-3z+z^2) - 2/(1+z)).$$

$$(b) \sum_{n \geq 0} F_n^3 z^n = z(1-2z-z^2)/(1-4z-z^2)(1+z-z^2) = (1/5)(2z/(1-4z-z^2) + 3z/(1+z-z^2)).$$

(通过平方或立方 Binet 公式(6.123)且对 n 求和得到这些公式, 然后合并项, 以致不出现 ϕ 和 $\hat{\phi}$.) 由此得到 $F_{n+1}^3 - 4F_n^3 + F_{n-1}^3 = 3(-1)^n F_n$. (Jarden 和 Motzkin^[163]已发现第 m 次幂的对应递归.)

6.59 设 m 被指定. 我们能通过对 n 归纳来证明, 事实上, 可能找到具有附加条件 $x \not\equiv 2 \pmod{4}$ 的这样的一个 x . 如果 x 是如此的一个解, 我们能上升到一个模 3^{n+1} 的解, 因为

$$F_{8 \cdot 3^{n-1}} \equiv 3^n, \quad F_{8 \cdot 3^{n-1}-1} \equiv 3^n + 1 \pmod{3^{n+1}};$$

x 或 $x + 8 \cdot 3^{n-1}$ 或 $x + 16 \cdot 3^{n-1}$ 将做此工作.

6.60 $F_1 + 1, F_2 + 1, F_3 + 1, F_4 - 1$ 和 $F_6 - 1$ 是仅有的情形. 否则习题 28 的 Lucas 数出现在因子分解

$$F_{2m} + (-1)^m = L_{m+1} F_{m-1}; \quad F_{2m+1} + (-1)^m = L_m F_{m+1};$$

$$F_{2m} - (-1)^m = L_{m-1} F_{m+1}; \quad F_{2m+1} - (-1)^m = L_{m+1} F_m$$

中. (一般我们有 $F_{m+n} - (-1)^n F_{m-n} = L_n F_n$.)

6.61 当 m 是偶和正时, $1/F_{2m} = F_{m-1}/F_m - F_{2m-1}/F_{2m}$. 对于 $n \geq 1$, 第二个和是 $5/4 - F_{3 \cdot 2^{n-1}}/F_{3 \cdot 2^n}$.

6.62 (a) $A_n = \sqrt{5} A_{n-1} - A_{n-2}$ 和 $B_n = \sqrt{5} B_{n-1} - B_{n-2}$. 顺便提到, 我们还有 $\sqrt{5} A_n + B_n = 2A_{n+1}$ 和 $\sqrt{5} B_n - A_n = 2B_{n+1}$.

(b) 小值的表显示

$$A_n = \begin{cases} L_n, & n \text{ 偶}, \\ \sqrt{5} F_n, & n \text{ 奇}; \end{cases} \quad B_n = \begin{cases} \sqrt{5} F_n, & n \text{ 偶}, \\ L_n, & n \text{ 奇}. \end{cases}$$

(c) $B_n/A_{n+1} - B_{n-1}/A_n = 1/(F_{2n+1} + 1)$, 因为 $B_n A_n - B_{n-1} A_{n+1} = \sqrt{5}$, 以及 $A_n A_{n+1} = \sqrt{5}(F_{2n+1} + 1)$. 注意到, $B_n/A_{n+1} = (F_n/F_{n+1})[n \text{ 偶}] + (L_n/L_{n+1})[n \text{ 奇}]$.

(d) 类似, $\sum_{k=1}^n 1/(F_{2k+1} - 1) = (A_0/B_1 - A_1/B_2) + \cdots + (A_{n-1}/B_n - A_n/B_{n+1})$

$= 2 - A_n / B_{n+1}$. 此等式也能表为 $(5F_n / L_{n+1})[n \text{ 偶}] + (L_n / F_{n+1})[n \text{ 奇}]$.

6.63 (a) $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$. 具有 $\pi_n = n$ 的有 $\left[\begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right]$ 个, 具有 $\pi_n < n$ 的有 $(n-1)\left[\begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right]$ 个.

(b) $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle$. $\{1, \dots, n-1\}$ 的每个排列 $\rho_1 \dots \rho_{n-1}$ 导致 n 个排列 $\pi_1 \pi_2 \dots \pi_n = \rho_1 \dots \rho_{j-1} n \rho_{j+1} \dots \rho_{n-1} \rho_j$. 如果 $\rho_1 \dots \rho_{n-1}$ 有 k 个超过数, 有 j 的 $k+1$ 个值产生 $\pi_1 \pi_2 \dots \pi_n$ 中的 k 个超过数; 剩下的 $n-1-k$ 个值产生 $k+1$. 因此在 $\pi_1 \pi_2 \dots \pi_n$ 中取 k 个超过数的方法总数为 $(k+1) \langle \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \rangle + ((n-1)-(k+1)) \langle \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \rangle = \langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle$.

6.64 依据习题 5.72 中的证明, $\left(\begin{smallmatrix} 1/2 \\ 2n \end{smallmatrix} \right)$ 的分母是 $2^{4n-2n(n)}$. 依据式 (6.44), $\left[\begin{smallmatrix} 1/2 \\ 1/2-n \end{smallmatrix} \right]$ 的分母是相同的, 因为 $\langle \begin{smallmatrix} n \\ 0 \end{smallmatrix} \rangle = 1$ 以及对于 $k > 0$, $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle$ 是偶的.

6.65 当 x_1, \dots, x_n 是 0 和 1 之间均匀分布的独立的随机变量时, 这等价于说, $\langle \begin{smallmatrix} n \\ k \end{smallmatrix} \rangle / n!$ 是我们有 $\lfloor x_1 + \dots + x_n \rfloor = k$ 的概率. 设 $y_j = (x_1 + \dots + x_j) \bmod 1$. 于是 y_1, \dots, y_n 是独立和均匀分布, $\lfloor x_1 + \dots + x_n \rfloor$ 是 y 中的下降数. y 的排列是随机的, 且 k 个下降的概率相同于 k 个上升的概率.

6.66 相似于式 (6.38), 我们有一般公式

$$\langle \begin{smallmatrix} n \\ m \end{smallmatrix} \rangle = \sum_{k=0}^m \binom{2n+1}{k} \left\{ \begin{smallmatrix} n+m+1-k \\ m+1-k \end{smallmatrix} \right\} (-1)^k, \text{ 对 } n > m \geq 0.$$

当 $m=2$ 时这等于

$$\begin{aligned} \langle \begin{smallmatrix} n \\ 2 \end{smallmatrix} \rangle &= \left\{ \begin{smallmatrix} n+3 \\ 3 \end{smallmatrix} \right\} - (2n+1) \left\{ \begin{smallmatrix} n+2 \\ 2 \end{smallmatrix} \right\} + \binom{2n+1}{2} \left\{ \begin{smallmatrix} n+1 \\ 1 \end{smallmatrix} \right\} \\ &= \frac{1}{2} 3^{n+2} - (2n+3) 2^{n+1} + \frac{1}{2} (4n^2 + 6n + 3). \end{aligned}$$

6.67 $(1/3)n(n+(1/2))(n+1)(2H_{2n} - H_n) - (1/36)n(10n^2 + 9n - 1)$. (像这样使推导公式自动化是吸引人的.)

6.68 $1/k - 1/(k+z) = z/k^2 - z^2/k^3 + \dots$, 当 $|z| < 1$ 时全收敛.

6.69 注意 $\prod_{k=1}^n (1+z/k) e^{-z/k} = \binom{n+z}{n} n^{-z} e^{(n+1)z}$. 如果 $f(z) = \frac{d}{dz} (z!)$, 我们找到 $f(z)/z! + \gamma = H_z$.

6.70 对于 $\tan z$, 我们能用 $\tan z = \cot z - 2 \cot 2z$ (它等价于习题 23 的等式). 此外 $z/\sin z = z \cot z + z \tan(1/2)z$ 有幂级数 $\sum_{n \geq 0} (-1)^{n-1} (4^n - 2) B_{2n} z^{2n} / (2n)!$, 以及

$$\begin{aligned} \ln \frac{\tan z}{z} &= \ln \frac{\sin z}{z} - \ln \cos z \\ &= \sum_{n \geq 1} (-1)^n \frac{4^n B_{2n} z^{2n}}{(2n)(2n)!} - \sum_{n \geq 1} (-1)^n \frac{4^n (4^n - 1) B_{2n} z^{2n}}{(2n)(2n)!} \end{aligned}$$

$$= \sum_{n \geq 1} (-1)^n \frac{4^n (4^n - 2) B_{2n} z^{2n}}{(2n)(2n)!},$$

因为 $\frac{d}{dz} \ln \sin z = \cot z$ 和 $\frac{d}{dz} \ln \cos z = -\tan z$,

6.71 由于 $\tan 2z + \sec 2z = (\sin z + \cos z) / (\cos z - \sin z)$, 在式(6.94)中置 $x=1$ 给出 $T_n(1) = 2^n T_n$ (当 n 是奇数), 当 n 是偶数, $T_n(1) = 2^n E_n$, 其中 $1/\cos z = \sum_{n \geq 0} E_n z^{2n} / (2n)!$.

(E_n 称为 Euler 数, 不要误认为 $\langle \frac{n}{k} \rangle$.)

6.72 如果 $n > 0$, $2^{n+1} (2^{n+1} - 1) B_{n+1} / (n+1)$. (见式(7.56)和(6.92), 希望的数实际上是 $1 - \tanh z$ 的系数.)

6.73 $\cot(z+\pi) = \cot z$ 和 $\cot(z+(1/2)\pi) = -\tan z$. 因此等式等价于

$$\cot z = \frac{1}{2^n} \sum_{k=0}^{2^n-1} \cot \frac{z+k\pi}{2^n},$$

根据 $n=1$ 的情形用归纳法得到它. 由于当 $z \rightarrow 0$ 时 $z \cot z \rightarrow 1$, 故所述极限随之而来. 由此表明逐项取极限是正确的, 因此式(6.88)成立. (顺便提到, 一般公式

$$\cot z = \frac{1}{n} \sum_{k=0}^{n-1} \cot \frac{z+k\pi}{n}$$

也是真的. 依据式(6.88)或依据

$$1/(e^{nz} - 1) = \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{e^{z+2k\pi i/n} - 1}$$

可以证明, 它等价于 $1/(z^n - 1)$ 的部分分式展开.)

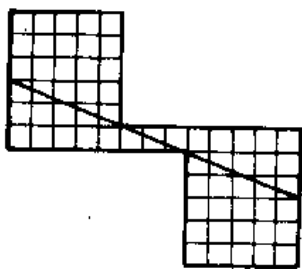
6.74 如果 $p(x)$ 是任何次数 $\leq n$ 的多项式, 我们有

$$p(x) = \sum_k p(-k) \binom{-x}{k} \binom{x+n}{n-k},$$

因为此方程对 $x=0, -1, \dots, -n$ 成立, 所述等式是 $p(x) = x \sigma_n(x)$ 和 $x=1$ 的特殊情形. 顺便提到, 依据在式(6.99)中置 $k=1$ 的 Stirling 数, 我们得到 Bernoulli 数的较简单的表达式:

$$\sum_{k \geq 0} \left\{ \begin{matrix} m \\ k \end{matrix} \right\} (-1)^k \frac{k!}{k+1} = B_m.$$

6.75 Sam Loyd^[204, p. 288, p. 378] 给出结构



且在 1858 年宣称已发现(但未出版)64=65 排列。(相似的谬论至少追溯到 18 世纪,但是 Loyd 找到了表示它们的较好方法。)

6.76 我们期望 $A_m / A_{m-1} \approx \varphi$, 所以我们试 $A_{m-1} = 618\,034 + r$ 和 $A_{m-2} = 381\,966 - r$. 然后 $A_{m-3} = 236\,068 + 2r$, 等等, 且我们发现 $A_{m-18} = 144 - 2\,584r$, $A_{m-19} = 154 + 4\,181r$. 因此 $r = 0$, $x = 154$, $y = 144$, $m = 20$.

6.77 如果对无限多个 n 的偶数值, $P(F_{n+1}, F_n) = 0$, 则 $U(x, y) - 1$ 可除尽 $P(x, y)$, 其中 $U(x, y) = x^2 - xy - y^2$. 考虑到如果 t 是 P 的总次数, 我们能写下

$$P(x, y) = \sum_{k=0}^t q_k x^k y^{t-k} + \sum_{i+k < t} r_{i,k} x^i y^k = Q(x, y) + R(x, y).$$

于是

$$\frac{P(F_{n+1}, F_n)}{F_n^t} = \sum_{k=0}^t q_k \left(\frac{F_{n+1}}{F_n} \right)^k + O\left(\frac{1}{F_n} \right)$$

且通过当 $n \rightarrow \infty$ 时取极限, 我们有 $\sum_{k=0}^t q_k \varphi^k = 0$. 因此 $Q(x, y)$ 是 $U(x, y)$ 的倍数, 比方说

$A(x, y)U(x, y)$. 但是 $U(F_{n+1}, F_n) = (-1)^n$ 且 n 为偶数, 所以 $P_0(x, y) = P(x, y) - (U(x, y) - 1)A(x, y)$ 是使得 $P_0(F_{n+1}, F_n) = 0$ 的另一个多项式. P_0 的总次数小于 t , 所以通过对 t 归纳, P_0 是 $U - 1$ 的倍数.

类似, 如果对无限多 n 的奇数值 $P(F_{n+1}, F_n) = 0$, $U(x, y) + 1$ 可除尽 $P(x, y)$. 合并这两个事实给出希望的必要和充分条件: $U(x, y)^2 - 1$ 可除尽 $P(x, y)$.

6.78 首先加不进位的数字位, 取得数字位 0, 1, 2. 然后用两个进位规则

$$0(d+1)(e+1) \rightarrow 1de,$$

$$0(d+2)0e \rightarrow 1d0(e+1),$$

总施加于最左的能用的进位. 此过程结束, 因为像完成一个进位时 $(b_m \cdots b_2)_2$ 增加那样, 通过读 $(b_m \cdots b_2)_F$ 得到二进制值. 但是一个进位可传送到“Fibonacci 点”的右边; 例如, $(1)_F + (1)_F$ 变成 $(10.01)_F$. 这种向右传送最多伸展两位; 如果必要, 用课文中的“加 1”算法把那两个数字位再调到零.

顺便提到, 在非负整数上有对应的“乘法”运算: 如果 $m = F_{i_1} + \dots + F_{i_s}$ 和 $n = F_{k_1} + \dots + F_{k_r}$ (在 Fibonacci 数系中), 用相似于二进制的乘法, 设 $m \circ n = \sum_{b=1}^s \sum_{c=1}^r F_{i_b + k_c}$. (此定义意味着当 m 和 n 大时, $m \circ n \approx \sqrt{5}mn$, 虽然 $l \circ n \approx \varphi^2 n$.) Fibonacci 加法导致结合律 $l \circ (m \circ n) = (l \circ m) \circ n$ 的一个证明.)

6.79 是的. 例如, 我们能取

$$A_0 = 331\ 635\ 635\ 998\ 274\ 737\ 472\ 200\ 656\ 430\ 763;$$

$$A_1 = 1\ 510\ 028\ 911\ 088\ 401\ 971\ 189\ 590\ 305\ 498\ 785.$$

当 $n \bmod m_k = r_k$ 时, 产生的序列有 A_n 被 p_k 除尽 (但不等于 p_k) 的性质, 其中数 (p_k, m_k, r_k) 有下列 18 个相应的值:

(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)

把这些三元组之一应用于每一个整数 n ; 例如, 第一列中的 6 个三元组覆盖每一个 n 的奇数值, 中间列覆盖所有不被 6 除尽的偶数 n . 证明的余下部分是基于 $A_{m+n} = A_m F_{n-1} + A_{m+1} F_n$ 以及对于每个三元组 (p_k, m_k, r_k) 的同余式

$$A_0 \equiv F_{m_k - r_k} \pmod{p_k},$$

$$A_1 \equiv F_{m_k - r_k + 1} \pmod{p_k}.$$

(一个改进的解, 其中 A_0 和 A_1 都是“仅有”17 个数字位的数, 也是可能的^[184].)

6.80 矩阵乘积是

$$\begin{pmatrix} K_{n-2}(x_2, \dots, x_{n-1}) & K_{n-1}(x_2, \dots, x_{n-1}, x_n) \\ K_{n-1}(x_1, x_2, \dots, x_{n-1}) & K_n(x_1, x_2, \dots, x_{n-1}, x_n) \end{pmatrix}.$$

这与式(6.137)中的 L 和 R 的乘积有关, 因为我们有

$$R^a \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} L^a.$$

行列式是 $K_n(x_1, \dots, x_n)$, 更一般的三对角线行列式

$$\det \begin{pmatrix} x_1 & 1 & 0 & \cdots & 0 \\ y_2 & x_2 & 1 & & 0 \\ 0 & y_3 & x_3 & 1 & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & 0 & \cdots & y_n & x_n \end{pmatrix}$$

满足递归 $D_n = x_n D_{n-1} - y_n D_{n-2}$.

6.81 设 $\alpha^{-1} = a_0 + 1/(a_1 + 1/(a_2 + \cdots))$ 是 α^{-1} 的连分式表示. 则我们有

$$\frac{a_0}{z} + \frac{1}{A_0(z) + \frac{1}{A_1(z) + \frac{1}{A_2(z) + \frac{1}{\ddots}}}} = \frac{1-z}{z} \sum_{n \geq 1} z^{\lfloor na \rfloor},$$

其中

$$A_m = \frac{z^{-q_{m+1}} - z^{-q_m-1}}{z^{-q_m} - 1}, \quad q_m = K_m(a_1, \cdots, a_m).$$

证明相似于课文的式(6.146)的证明, 用了 Zeckendorf 定理的推广(Fraenkel^[104, §4]). 如果 $z = 1/b$, 其中 b 是一个整数 ≥ 2 , 如同习题 49 那样, 这给出了超越数 $(b-1) \sum_{n \geq 1} b^{-\lfloor na \rfloor}$ 的连分数表示.

6.82 习题 62 的序列满足 $A_{-m} = A_m$, $B_{-m} = -B_m$, 且

$$A_m A_n = A_{m+n} + A_{m-n};$$

$$A_m B_n = B_{m+n} - B_{m-n};$$

$$B_m B_n = A_{m+n} - A_{m-n}.$$

设 $f_k = B_{mk} / A_{mk+l}$ 和 $g_k = A_{mk} / B_{mk+l}$, 其中 $l = (1/2)(n-m)$. 于是 $f_{k+1} - f_k = A_l B_m / (A_{2mk+n} + A_m)$ 和 $g_k - g_{k+1} = A_l B_m / (A_{2mk+n} - A_m)$. 因此我们有

$$\begin{aligned} S_{m,n}^+ &= \frac{\sqrt{5}}{A_l B_m} \lim_{k \rightarrow \infty} (f_k - f_0) = \frac{\sqrt{5}}{\varphi^l A_l L_m}; \\ S_{m,n}^- &= \frac{\sqrt{5}}{A_l B_m} \lim_{k \rightarrow \infty} (g_0 - g_k) = \frac{\sqrt{5}}{A_l L_m} \left(\frac{2}{B_l} - \frac{1}{\varphi^l} \right) \\ &= \frac{2}{F_l L_l L_m} - S_{m,n}^+. \end{aligned}$$

6.83 设 $p = K(0, a_1, a_2, \cdots, a_m)$, 以致 p/n 是第 m 个收敛到连分式. 于是

$\alpha = p/n + (-1)^m/nq$, 其中 $q = K(a_1, \dots, a_m, \beta)$ 和 $\beta > 1$. 所以对 $0 \leq k < n$, 点 $\{k\alpha\}$ 能写成

$$\frac{0}{n}, \frac{1}{n} + \frac{(-1)^m \pi_1}{nq}, \dots, \frac{n-1}{n} + \frac{(-1)^m \pi_{n-1}}{nq},$$

其中 π_1, \dots, π_{n-1} 是 $\{1, \dots, n-1\}$ 的一个排列, 设 $f(v)$ 是这样的点 $< v$ 的个数, 除 $k=0$ 或 $k=n-1$ 时外, 当 v 从 k/n 增加到 $(k+1)/n$ 时, $f(v)$ 和 vn 都增加 1, 所以它们不会相差 2 或相差更大.

6.84 依据式 (6.139) 和 (6.136), 我们要在所有正整数和 $\leq n+1$ 的序列上使 $K(a_1, \dots, a_m)$ 最大. 当所有 a 是 1 时, 最大出现, 如果 $j \geq 1$ 和 $a \geq 1$, 我们有

$$\begin{aligned} & K_{j+k+1}(1, \dots, 1, a+1, b_1, \dots, b_k) \\ &= K_{j+k+1}(1, \dots, 1, a, b_1, \dots, b_k) + K_j(1, \dots, 1)K_k(b_1, \dots, b_k) \\ &\leq K_{j+k+1}(1, \dots, 1, a, b_1, \dots, b_k) + K_{j+k}(1, \dots, 1, a, b_1, \dots, b_k) \\ &= K_{j+k+2}(1, \dots, 1, a, b_1, \dots, b_k). \end{aligned}$$

(Motzkin 和 Straus^[220]在延拓上解更一般的最大值问题.)

6.85 性质成立当且仅当 N 有 7 种形式 $5^k, 2 \cdot 5^k, 4 \cdot 5^k, 3^j \cdot 5^k, 6 \cdot 5^k, 7 \cdot 5^k, 14 \cdot 5^k$ 之一.

6.86 在 [179, 第 6 节] 中出现情形 $n \bmod 1 = 1/2$ 的一个候选者, 虽然用涉及 $\sqrt{\pi}$ 的某个常数乘那里的整数可能是最好的.

6.87 (a) 如果仅有有限多个解, 猜测对所有素数同样成立是自然的.

(b) b_n 的情况是十分奇怪的: 对于 $968 \leq n \leq 1066$, 我们有 $b_n = \text{lcm}(1, \dots, n)$; 另一方面, $b_{600} = \text{lcm}(1, \dots, 600) / (3^3 \cdot 5^2 \cdot 43)$. Andrew Odlyzko 看出 p 除尽 $\text{lcm}(1, \dots, n) / b_n$ 当且仅当 $kp^m \leq n < (k+1)p^m$ (对于某个 $m \geq 1$ 和某个 $k < p$ 使得 p 除尽 H_k 的分子). 所以如果能表明, 例如, 几乎所有素数仅有一个这样的 k 值 (即 $k=p-1$), 则存在无限多个这样的 n .

6.88 (Brent^[33]发现在 e^7 中的意外大的部分商 1 568 705, 但是这看来是一种巧合. 例如, Gosper 发现 π 中的更大的部分商: 第 453 294 个是 12 996 958 以及第 11 504 913 个是 878 783 625.)

6.89 考虑母函数 $\sum_{m, n \geq 0} \binom{m+n}{m} w^m z^n$, 它有形式

$$\sum_n (wF(a, b, c) + zF(a', b', c'))^n,$$

其中 $F(a, b, c)$ 是微分算子 $a + b\theta_w + c\theta_z$.

7.1 在母函数中对 \square 代入 z^4 , 对 \sqsupset 代入 z , 取得 $1/(1-z^4-z^2)$. 这如同 T 的母函

7.13 周期地推广序列(设 $s_{m+k} = s_k$)且定义 $s_n = s_1 + \dots + s_m$. 我们有 $s_m = l$, $s_{2m} = 2l$, 等. 一定有一个最大指标 k_j 使得 $s_{k_j} = j$, $s_{k_j+m} = l+j$, 等等. 这些指标 k_1, \dots, k_l (模 m)指定问题中的循环移位.

例如, 具有 $m=10$ 和 $l=2$ 的序列 $\langle -2, 1, -1, 0, 1, 1, -1, 1, 1, 1 \rangle$ 中, 我们有 $k_1=17$, $k_2=24$.

7.14 通过二次公式, $\hat{G}(z) = -2z\hat{G}(z) + \hat{G}(z)^2 + z$ (关于最后项要小心!)导致

$$\hat{G}(z) = \frac{1 + 2z - \sqrt{1 + 4z^2}}{2}.$$

因此对所有 $n > 0$, $g_{2n+1} = 0$ 和 $g_{2n} = (-1)^n (2n)! C_{n-1}$.

7.15 包含 $n+1$ 的子集中有 k 个其他元素, 有 $\binom{n}{k} b_{n-k}$ 种划分, 因此 $\hat{B}'(z) = e^z \hat{B}(z)$. 这个微分方程的解是 $\hat{B}(z) = e^{e^z + c}$, 且由于 $\hat{B}(0)=1$, 所以 $c=-1$. (由于 $b_n = \sum_m \left\{ \begin{smallmatrix} n \\ m \end{smallmatrix} \right\}$, 通过式(7.49)对 m 求和, 我们也能取得这个结果.)

7.16 一种方法是取

$$B(z) = \frac{1}{((1-z)^{a_1} (1-z^2)^{a_2} (1-z^3)^{a_3} (1-z^4)^{a_4} \dots)}$$

的对数, 然后用 $\ln(1/(1-z))$ 的公式且交换求和的次序.

7.17 由于 $\int_0^\infty t^n e^{-t} dt = n!$, 这就随之而来. 还有以别的方向进行的公式:

$$\hat{G}(z) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} G(ze^{-i\theta}) e^{e^{i\theta}} d\theta.$$

7.18 (a) $\zeta(z - (1/2))$; (b) $-\zeta'(z)$; (c) $\zeta(z)/\zeta(2z)$. 每一个正整数唯一地可表示为 $m^2 q$, 其中 q 是无平方的.

7.19 如果 $n > 0$, 系数 $[z^n] \exp(x \ln F(z))$ 是 x 的次数 n 的多项式, 即 x 的倍数. 第一个卷积公式来自相等 $F(z)^x F(z)^y = F(z)^{x+y}$ 中 z^n 的系数. 第二个来自相等 $F'(z) F(z)^{x-1} F(z)^y = F'(z) F(z)^{x+y-1}$ 中 z^{n-1} 的系数, 因为我们有

$$F'(z) F(z)^{x-1} = x^{-1} \frac{\partial}{\partial z} (F(z)^x) = x^{-1} \sum_{n \geq 0} n f_n(x) z^{n-1}.$$

(如同式(7.43)中的那样, 作 $\partial/\partial x$ 得到另外的卷积.)

7.20 设 $G(z) = \sum_{n \geq 0} g_n z^n$. 如果对 $n < 0$ 我们考虑 $g_n = 0$, 则对于所有 $k, l \geq 0$,

$$z^l G^{(k)}(z) = \sum_{n \geq 0} n^{\underline{k}} g_n z^{n-k+l} = \sum_{n \geq 0} (n+k-l)^{\underline{k}} g_{n+k-l} z^n.$$

因此如果 $P_0(z), \dots, P_m(z)$ 是多项式(不全为零), 最大次数为 d , 则有多项式 $p_0(n), \dots, p_{m+d}(n)$ 使得

$$P_0(z)G(z) + \dots + P_m(z)G^{(m)}(z) = \sum_{n \geq 0} \sum_{j=0}^{m+d} p_j(n) g_{n+j-d} z^n.$$

所以一个可微有限的 $G(z)$ 意味着

$$\sum_{j=0}^{m+d} p_j(n+d) g_{n+j} = 0, \text{ 对所有 } n \geq 0.$$

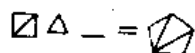
逆是相似的。(一个推论是, $G(z)$ 是可微有限当且仅当对应的指数型母函数 $\hat{G}(z)$ 是可微有限的。)

7.21 这是改变分母 10 和 20 的问题, 所以 $G(z) = 1/(1-z^{10})(1-z^{20}) = \check{G}(z^{10})$, 其中 $\check{G}(z) = 1/(1-z)(1-z^2)$ 。

(a) $\check{G}(z)$ 的部分分式分解是 $(1/2)(1-z)^{-2} + (1/4)(1-z)^{-1} + (1/4)(1+z)^{-1}$, 所以 $[z^n]\check{G}(z) = (1/4)(2n+3+(-1)^n)$ 。置 $n=50$ 产生 26 种方式作支付。

(b) $\check{G}(z) = (1+z)/(1-z^2)^2 = (1+z)(1+2z^2+3z^4+\dots)$, 所以 $[z^n]\check{G}(z) = \lfloor n/2 \rfloor + 1$ 。(把这与课文的硬币替换问题中的值 $N_n = \lfloor n/5 \rfloor + 1$ 比较, 银行盗贼的问题等价于分和 2 分替换的问题。)

7.22 每个多边形有一个“底”(下部处的线段)。如果 A 和 B 是三角剖分的多边形, 设 $A\Delta B$ 是把 A 的底粘贴到 Δ 的上左对角线, 且把 B 的底粘贴到上右对角线的结果。例如,



(多边形可能需要扭曲一点, 与 Δ 或使成形。)每个三角剖分以此方产生, 因为底线是一个唯一的三角形的一部分, 且有三角剖分的多边形 A 和 B 在它的左边和右边。

用 z 替换每个三角形得到一个幂级数, 在幂级数中 z^n 的系数是具有 n 个三角形的三角剖分的个数, 即把一个 $(n+2)$ 边形分解为三角形的方式数。由于 $P = 1 + zP^2$, 这是 Catalan 数 $C_0 + C_1z + C_2z^2 + \dots$ 的母函数; 三角剖分一个 n 边形的方式数是 $C_{n-2} = \binom{2n-4}{n-2} / (n-1)$ 。

7.23 设 a_n 是所述的数, b_n 是在顶部具有 $2 \times 1 \times 1$ 切口损失的方式数。通过考虑在顶部曲面上可能型式的外形, 我们有

$$a_n = 2a_{n-1} + 4b_{n-1} + a_{n-2} + [n=0];$$

$$b_n = a_{n-1} + b_{n-1}.$$

因此母函数满足 $A = 2zA + 4zB + z^2A + 1$, $B = zA + zB$, 且我们有

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



此公式和 $3 \times n$ 多米诺骨牌问题有关; 我们有 $a_n = (1/3)(U_{2n} + V_{2n+1} + (-1)^n) = (1/6)(2+\sqrt{3})^{n+1} + (1/6)(2-\sqrt{3})^{n+1} + (1/3)(-1)^n$, 它是 $(2+\sqrt{3})^{n+1}/6$ 化整到最近整数.

7.24 $n \sum_{k_1+\dots+k_m=n} k_1 \cdots k_m / m = F_{2n+1} + F_{2n-1} - 2$. (考虑系数 $[z^{n-1}](d/dz)\ln(1/(1-G(z)))$, 其中 $G(z) = z/(1-z)^2$.)

7.25 母函数是 $P(z)/(1-z^m)$, 其中 $P(z) = z + 2z^2 + \cdots + (m-1)z^{m-1} = ((m-1)z^{m+1} - mz^m + z)/(1-z)^2$. 分母是 $Q(z) = 1-z^m = (1-\omega^0 z)(1-\omega^1 z)\cdots(1-\omega^{m-1} z)$. 依据不同根的有理展开定理, 我们得到

$$n \bmod m = \frac{m-1}{2} + \sum_{k=1}^{m-1} \frac{\omega^{-kn}}{\omega^k - 1}.$$

7.26 如同方程 (7.60) 中的那样, $(1-z-z^2)\mathcal{F}(z) = F(z)$ 导致 $\mathcal{F}_n = (2(n+1)F_n + nF_{n+1})/5$.

7.27 以两种方式之一定向以  或  或一个 $2 \times k$ (某个 $k \geq 2$) 圈开始的每个定向圈型式. 因此对 $n \geq 2$,

$$Q_n = Q_{n-1} + Q_{n-2} + 2Q_{n-2} + 2Q_{n-3} + \cdots + 2Q_0,$$

$Q_0 = Q_1 = 1$. 所以母函数是

$$\begin{aligned} Q(z) &= zQ(z) + z^2Q(z) + 2z^2Q(z)/(1-z) + 1 \\ &= 1/(1-z-z^2-2z^2/(1-z)) \\ &= \frac{(1-z)}{(1-2z-2z^2+z^3)} \\ &= \frac{\varphi^2/5}{1-\varphi^2z} + \frac{\varphi^{-2}/5}{1-\varphi^{-2}z} + \frac{2/5}{1+z}, \end{aligned}$$

且 $Q_n = (\varphi^{2n+2} + \varphi^{-2n-2} + 2(-1)^n)/5 = ((\varphi^{n+1} - \hat{\varphi}^{n+1})/\sqrt{5})^2 = F_{n+1}^2$.

7.28 一般如果 $A(z) = (1+z+\cdots+z^{m-1})B(z)$, 对于 $0 \leq r < m$ 我们有 $A_r + A_{r+m} + A_{r+2m} + \cdots = B(1)$. 此时 $m=10$ 和 $B(z) = (1+z+\cdots+z^9)(1+z^2+z^4+z^6+z^8)(1+z^5)$.

7.29 $F(z) + F(z)^2 + F(z)^3 + \cdots = z/(1-z-z^2-z) = (1/(1-(1+\sqrt{2})z) - (1/(1-(1-\sqrt{2})z)))/\sqrt{8}$, 所以解答为 $((1+\sqrt{2})^n - (1-\sqrt{2})^n)/8$.

7.30 依据习题 5.39, $\sum_{k=1}^n \binom{2n-1-k}{n-1} (a^n b^{n-k} / (1-\alpha z)^k + a^{n-k} b^n / (1-\beta z)^k)$.

7.31 Dirichlet 母函数为 $\zeta(z)^2 / \zeta(z-1)$; 因此我们发现所有恰好除尽 n 的素数幂 p^k 上, $g(n)$ 是 $(k+1-kp)$ 的乘积.

7.32 我们可假设每个 $b_k \geq 0$. 算术级数的集合形成一个正好覆盖, 当且仅当

$$\frac{1}{1-z} = \frac{z^{b_1}}{1-z^{a_1}} + \cdots + \frac{z^{b_m}}{1-z^{a_m}}.$$

两边减 $z^{b_m} / (1-z^{a_m})$ 且置 $z = e^{2\pi i / a_m}$. 左边是无限的, 且除 $a_{m-1} = a_m$ 之外, 右边是有限的.

$$7.33 \quad (-1)^{n-m+1} [n > m] / (n-m),$$

7.34 我们还能写成 $G_n(z) = \sum_{k_1 + (m+1)k_{m+1} = n} \binom{k_1 + k_{m+1}}{k_{m+1}} (z^m)^{k_{m+1}}$. 一般, 如果

$$G_n = \sum_{k_1 + 2k_2 + \cdots + rk_r = n} \binom{k_1 + k_2 + \cdots + k_r}{k_1, k_2, \dots, k_r} z_1^{k_1} z_2^{k_2} \cdots z_r^{k_r},$$

我们有 $G_n = z_1 G_{n-1} + z_2 G_{n-2} + \cdots + z_r G_{n-r} + [n=0]$, 母函数为 $1 / (1 - z_1 w - z_2 w^2 - \cdots - z_r w^r)$. 在所述特殊情形中, 解答为 $1 / (1 - w - z^m w^{m+1})$. (情形 $m=1$, 见式 (5.74).)

$$7.35 \quad (a) \quad \frac{1}{n} \sum_{0 < k < n} (1/k + 1/(n-k)) = \frac{2}{n} H_{n-1}.$$

(b) $[z^n](\ln(1/(1-z)))^2 = (2!/n!) \left[\begin{smallmatrix} n \\ 2 \end{smallmatrix} \right] = (2/n) H_{n-1}$, 依据式 (7.50) 和 (6.58). 处理 (b) 的另一种方法是用规则 $[z^n]F(z) = (1/n)[z^{n-1}]F'(z)$, 其中 $F(z) = (\ln(1/(1-z)))^2$.

$$7.36 \quad \frac{1-z^m}{1-z} A(z^m).$$

7.37 (a) 在表

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

中惊人的等式 $a_{2n} = a_{2n+1} = b_n$ 成立.

$$(b) \quad A(z) = 1 / ((1-z)(1-z^2)(1-z^4)(1-z^8)\cdots).$$

(c) $B(z) = A(z) / (1-z)$, 且我们要证明 $A(z) = (1+z)B(z^2)$. 根据 $A(z) = A(z^2) / (1-z)$ 得到此式.

$$7.38 \quad (1-wz)M(w, z) = \sum_{n, n \geq 1} (\min(m, n) - \min(m-1, n-1)) w^m z^n = \sum_{n, n \geq 1} w^m z^n = wz / (1-w)(1-z). \text{ 一般, }$$

$$M(z_1, \dots, z_m) = \frac{z_1 \cdots z_m}{(1-z_1) \cdots (1-z_m)(1-z_1 \cdots z_m)}.$$

7.39 提示的解答分别是

$$\sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} a_{k_1} a_{k_2} \cdots a_{k_m} \text{ 和 } \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} a_{k_1} a_{k_2} \cdots a_{k_m}.$$

所以: (a) 在乘积 $(1+z)(1+2z)\cdots(1+nz)$ 中我们要 z^m 的系数。这是 $(z+1)^n$ 的反射, 所以它是 $\begin{bmatrix} n+1 \\ n+1 \end{bmatrix} + \begin{bmatrix} n+1 \\ n \end{bmatrix} z + \cdots + \begin{bmatrix} n+1 \\ 1 \end{bmatrix} z^n$, 且解答为 $\begin{bmatrix} n+1 \\ n+1-m \end{bmatrix}$. (b) 依据式 (7.47), 在 $1/(1-z)(1-2z)\cdots(1-nz)$ 中 z^m 的系数为 $\left\{ \begin{matrix} m+n \\ n \end{matrix} \right\}$.

7.40 $\langle nF_{n-1} - F_n \rangle$ 的指数型母函数为 $(z-1)\hat{F}(z)$, 其中 $\hat{F}(z) = \sum_{n \geq 0} F_n z^n / n! = (e^{\varphi z} - e^{\bar{\varphi} z}) / \sqrt{5}$. $\langle n! \rangle$ 的指数型母函数为 $e^{-z} / (1-z)$, 乘积是

$$5^{-1/2} (e^{(\varphi-1)z} - e^{(\bar{\varphi}-1)z}) = 5^{-1/2} (e^{-\varphi z} - e^{-\bar{\varphi} z}).$$

我们有 $\hat{F}(z)e^{-z} = -\hat{F}(-z)$. 所以解答是 $(-1)^n F_n$.

7.41 在位置 $2k$ 中具有最大元素 n 的上升-下降排列数为 $\binom{n-1}{2k-1} A_{2k-1} A_{n-2k}$. 类似, 在位置 $2k+1$ 中具有最小元素 1 的上升-下降排列数是 $\binom{n-1}{2k} A_{2k} A_{n-2k-1}$, 因为下降-上升排列和上升-下降排列是相同地为数众多. 在所有可能的值上求和得到

$$2A_n = \sum_k \binom{n-1}{k} A_k A_{n-1-k} + 2[n=0] + [n=1].$$

所以指数型母函数 \hat{A} 满足 $2\hat{A}'(z) = \hat{A}(z)^2 + 1$ 和 $\hat{A}(z) = 1$, 给定的函数解这个微分方程.

7.42 设 a_n 是不以 c 或 e 终止的火星脱氧核糖核酸的串数; 设 b_n 是以 c 或 e 终止的火星脱氧核糖核酸的串数. 则

$$\begin{aligned} a_n &= 3a_{n-1} + 2b_{n-1} + [n=0], \quad b_n = 2a_{n-1} + b_{n-1}; \\ A(z) &= 3zA(z) + 2zB(z) + 1, \quad B(z) = 2zA(z) + 2B(z); \\ A(z) &= \frac{1-z}{1-4z-z^2}, \quad B(z) = \frac{2z}{1-4z-z^2}; \end{aligned}$$

且总数为 $[z^n](1+z)/(1-4z-z^2) = F_{3n+2}$.

7.43 依据式 (5.45), $g_n = \Delta^n \hat{G}(0)$. 乘积的第 n 次差分能写成

$$\Delta^n A(z)B(z) = \sum_k \binom{n}{k} (\Delta^k E^{n-k} A(z)) (\Delta^{n-k} B(z)),$$

以及 $E^{n-k} = (1+\Delta)^{n-k} = \sum_j \binom{n-k}{j} \Delta^j$. 所以我们找到

$$h_n = \sum_{j,k} \binom{n}{k} \binom{n-k}{j} f_{j+k} g_{n-k}.$$

这是所有三项系数上的和, 它能变成更对称的形式

$$h_n = \sum_{j+k+l=n} \binom{n}{j, k, l} f_{j+k} g_{k+l}.$$

7.44 能以 $k!$ 种方式安排每个划分成的 k 个非空的子集, 所以 $b_k = k!$. 因此 $\hat{Q}(z) = \sum_{n,k \geq 0} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} k! z^n / n! = \sum_{k \geq 0} (e^z - 1)^k = 1 / (2 - e^z)$. 且这是几何级数 $\sum_{k \geq 0} e^{kz} / 2^{k+1}$, 因此 $a_k = 1 / 2^{k+1}$. 最后, $c_k = 2^k$; 当 x 不同时考虑所有排列, 在指标之间把每个 ' $>$ ' 改变为 ' $<$ ', 且在指标之间使每个 ' $<$ ' 可变成 ' $<$ ' 或 ' $=$ '. (例如, 因为 $1 < 3 > 2$, 排列 $x_1 x_3 x_2$ 产生 $x_1 < x_3 < x_2$ 和 $x_1 = x_3 < x_2$.)

7.45 此和是 $\sum_{n \geq 1} r(n) / n^2$, 其中 $r(n)$ 是 n 写成两个互素因子的乘积的方式数. 如果 t 个不同素数可除尽 n , $r(n) = 2^t$. 因此 $r(n) / n^2$ 是积性的且和为

$$\begin{aligned} \prod_p \left(1 + \frac{2}{p^2} + \frac{2}{p^4} + \cdots \right) &= \prod_p \left(1 + \frac{2}{p^2 - 1} \right) \\ &= \prod_p \left(\frac{p^2 + 1}{p^2 - 1} \right) = \frac{\zeta(2)^2}{\zeta(4)} = \frac{5}{2}. \end{aligned}$$

7.46 设 $S_n = \sum_{0 \leq k \leq n/2} \binom{n-2k}{k} \alpha^k$, 则 $S_n = S_{n-1} + \alpha S_{n-3} + [n=0]$, 且母函数为 $1 / (1 - z - \alpha z^3)$. 当 $\alpha = -(4/27)$ 时, 提示告诉我们, 这有一个好的因子分解 $1 / (1 + (1/3)z)(1 - (2/3)z)^2$. 现在, 一般展开定理产生 $S_n = ((2/3)n + c)(2/3)^n + (1/9)(-1/3)^n$, 剩下的常数 c 是 $8/9$.

7.47 $\sqrt{3}$ 的 Stern-Brocot 表示为 $R(LR^2)^\infty$, 因为

$$\sqrt{3} + 1 = 2 + \frac{1}{1 + \frac{1}{\sqrt{3} + 1}}.$$

分数是 $1/1, 2/1, 3/2, 5/3, 7/4, 12/7, 19/11, 26/15, \dots$; 它们实际有循环的型式

$$\frac{V_{2n-1} + V_{2n+1}}{U_{2n}}, \frac{U_{2n} + V_{2n+1}}{V_{2n+1}}, \frac{U_{2n+2} + V_{2n-1}}{U_{2n} + V_{2n+1}}, \frac{V_{2n+1} + V_{2n+3}}{U_{2n+2}}, \dots,$$

7.48 我们有 $x_0 = 0$, 且如果 $x_1 = m$, 母函数满足

$$aG(z) + bz^{-1}G(z) + cz^{-2}(G(z) - mz) + \frac{d}{1-z} = 0.$$

因此对某个多项式 $P(z)$, $G(z) = P(z) / (az^2 + bz + c)(1-z)$. 设 ρ_1 和 ρ_2 是 $cz^2 + bz + a$ 的根, $|\rho_1| \geq |\rho_2|$. 如果 $b^2 - 4ac \leq 0$, 则 $|\rho_1|^2 = \rho_1 \rho_2 = a/c$ 是有理的, 与 $\sqrt[n]{g_n}$ 趋于 $1 + \sqrt{2}$ 相矛盾. 因此 $\rho_1 = (-b + \sqrt{b^2 - 4ac}) / 2c = 1 + \sqrt{2}$, 且这意味着 $a = -c$, $b = -2c$, $\rho_2 = 1 - \sqrt{2}$. 现在母函数取形式

$$\begin{aligned} G(z) &= \frac{z(m - (r + m)z)}{(1 - 2z - z^2)(1 - z)} \\ &= \frac{-r + (m + 2r)z}{2(1 - 2z - z^2)} + \frac{r}{2(1 - z)} = mz + (2m - r)z^2 + \dots, \end{aligned}$$

其中 $r = d/c$. 由于 g_2 是一个整数, 所以 r 是一个整数. 我们还有

$$g_n = \alpha(1 + \sqrt{2})^n + \hat{\alpha}(1 - \sqrt{2})^n + \frac{1}{2}r = \lfloor \alpha(1 + \sqrt{2})^n \rfloor,$$

且仅当 $r = -1$ 时此式成立, 因为当它趋于零时 $(1 - \sqrt{2})^n$ 交替正负号. 因此 $(a, b, c, d) = \pm(1, 2, -1, 1)$. 现在我们找到 $\alpha = (1/4)(1 + \sqrt{2}m)$, 仅当 $0 \leq m \leq 2$ 时它在 0 和 1 之间. 每个这样的值实际给出一个解, 序列 $\langle g_n \rangle$ 是 $\langle 0, 0, 1, 3, 8, \dots \rangle$, $\langle 0, 1, 3, 8, 20, \dots \rangle$ 和 $\langle 0, 2, 5, 13, 32, \dots \rangle$.

7.49 (a) $(1 / (1 - (1 + \sqrt{2})z) + 1 / (1 - (1 - \sqrt{2})z))$ 的分母是 $1 - 2z - z^2$, 因此 $a_n = 2a_{n-1} + a_{n-2} (n \geq 2)$.

(b) 真的, 因为 a_n 是偶数, 且 $-1 < 1 - \sqrt{2} < 0$.

(c) 设

$$b_n = \left(\frac{p + \sqrt{q}}{2} \right)^n + \left(\frac{p - \sqrt{q}}{2} \right)^n.$$

对所有 $n > 0$, 我们希望 b_n 是奇, 且 $-1 < (p - \sqrt{q})/2 < 0$. 像(a)那样处理, 我们找到 $b_0 = 2$, $b_1 = p$, 且对 $n \geq 2$, $b_n = pb_{n-1} + (1/4)(q - p^2)b_{n-2}$. 一个满意的解有 $p = 3$ 和 $q = 17$.

7.50 推广习题 22 的乘积思想, 我们有

$$Q = \text{—} + Q \triangle Q + Q \square Q + \begin{array}{c} Q \quad \diagdown \quad Q \\ \quad \quad \quad Q \\ Q \quad \quad \quad Q \end{array} + \dots.$$

用 z^{n-2} 替换每个 n 多边形, 此替换在乘积之下呈现适当, 因为粘贴操作把一个 m 边形和一个 n 边形变成一个 $(m+n-2)$ 边形. 因此母函数为

$$Q = 1 + zQ^2 + z^2Q^3 + z^3Q^4 + \cdots = 1 + \frac{zQ^2}{1 - zQ}$$

且二次公式给出 $Q = (1 + z - \sqrt{1 - 6z + z^2}) / 2z$. 在此幂级数中 z^{n-2} 的系数是把不重叠的对角线放入一个凸 n 边形的方式数. 显然这些系数没有依据本书中已讨论过的其他量的闭形式, 但是它们的渐近情况是知道的^[173, 习题 2.2.1-12].

顺便提到, 用 wz^{n-2} 替换 Q 中的每个 n 边形, 我们取得

$$Q = \frac{1 + z - \sqrt{1 - (4w + 2)z + z^2}}{2(1 + w)z},$$

公式中的 $w^n z^{n-2}$ 的系数是用不相交对角线把一个 n 边形分为 m 个多边形的方式数.

7.51 推广习题 27. 关键的第一步是看出, 方式数的平方是一个确定类型的循环型式的个数. 通过计算本征值不确定的矩阵的行列式能计算这些. 当 $m=3$ 和 $n=4$ 时, $\cos 36^\circ = \varphi/2$ 是有用的(习题 6.46).

7.52 最初几种情形是 $p_0(y) = 1$, $p_1(y) = y$, $p_2(y) = y^2 + y$, $p_3(y) = y^3 + 3y^2 + 3y$. 设 $p_n(y) = q_n(x)$, 其中 $y = x(1-x)$, 我们寻找以一种方便的方式确定 $q_{2n+1}(x)$ 的一个母函数. 一个这样的函数是 $\sum q_n(x)z^n/n! = 2e^{xz}/(e^{2x}+1)$, 由它得到 $q_n(x) = i^n E_n(x)$, 其中 $E_n(x)$ 称为一个 Euler 多项式. 我们有 $\sum (-1)^x x^n \delta x = (1/2) \times (-1)^{x+1} E_n(x)$, 所以 Euler 多项式相似于 Bernoulli 多项式, 且它们也相似于式(6.98)中的那些因子. 依据习题 6.23, 我们有 $nE_{n-1}(x) = \sum_{k=0}^n \binom{n}{k} B_k x^{n-k} (2-2^{k+1})$; 依据习题 6.54, 这个多项式有整数系数. 因此系数有 2 的幂的分母的 $q_{2n}(x)$ 一定有整数系数. 因此 $p_n(y)$ 有整数系数. 最后, 关系式 $(4y-1)p_n''(y) + 2p_n'(y) = 2n(2n-1)p_{n-1}(y)$ 证明

$$2m(2m-1) \left| \begin{matrix} n \\ m \end{matrix} \right| = m(m+1) \left| \begin{matrix} n \\ m+1 \end{matrix} \right| + 2n(2n-1) \left| \begin{matrix} n-1 \\ m-1 \end{matrix} \right|,$$

由此得到 $\left| \begin{matrix} n \\ m \end{matrix} \right|$ 是正的. (一个相似的证明表明, 当表达成一个 y 的 n 次多项式时, 有关的量 $(-1)^n (2n+2)E_{2n+1}(x)/(2x-1)$ 有正的整数系数.) 由此表明 $\left| \begin{matrix} n \\ 1 \end{matrix} \right|$ 是 Genocchi 数 $(-1)^{n-1} (2^{2n+1} - 2)B_{2n}$ (见习题 6.24), 且 $\left| \begin{matrix} n \\ n-1 \end{matrix} \right| = \binom{n}{2}$, $\left| \begin{matrix} n \\ n-2 \end{matrix} \right| = 2\binom{n+1}{4} + 3\binom{n}{4}$, 等等.

7.53 它是 $P_{(1+V_{2n+1}+V_{2n+3})/6}$. 因此, 例如, $T_{20} = P_{12} = 210$; $T_{285} = P_{165} = 40\,755$.

7.54 除了那些 z^n (其中 $n \bmod m = k$) 之外, 设 E_k 是在幂级数上置所有系数为零的运算. 所述的结构等价于运算

$$E_0 S E_0 S (E_0 + E_1) S \cdots S (E_0 + E_1 + \cdots + E_{m-1})$$

应用于 $1/(1-z)$, 其中 S 意味着“用 $1/(1-z)$ 乘”. 有 $m!$ 个项

$$E_0 S E_{k_1} S E_{k_2} S \cdots S E_{k_m}$$

其中 $0 \leq k_j < j$, 且如果 r 是位数 ($k_j < k_{j+1}$), 每个这样的项计算到 $z^m / (1-z^m)^{m+1}$.

恰好 $\langle \begin{smallmatrix} m \\ r \end{smallmatrix} \rangle$ 项有一个 r 的给定值, 所以依据式 (6.37), z^m 的系数为

$$\sum_{r=0}^{m-1} \langle \begin{smallmatrix} m \\ r \end{smallmatrix} \rangle \binom{n+m-r}{m} = (n+1)^m. \quad (\text{运算 } E_k \text{ 能用单位的复数根来表达, 看来在此问题中没有用处.})$$

7.55 假设 $P_0(z)F(z) + \cdots + P_m(z)F^{(m)}(z) = Q_0(z)G(z) + \cdots + Q_n(z)G^{(n)}(z) = 0$, 其中 $P_m(z)$ 和 $Q_n(z)$ 为非零. (a) 设 $H(z) = F(z) + G(z)$, 则对于 $0 \leq l < m+n$, 有有理函数 $R_{k,l}(z)$ 使得 $H^{(k)}(z) = R_{k,0}(z)F^{(0)}(z) + \cdots + R_{k,m-1}(z)F^{(m-1)}(z) + R_{k,m}(z)G^{(0)}(z) + \cdots + R_{k,m+n-1}(z)G^{(n-1)}(z)$. 在 $(m+n)$ 维向量空间中, $m+n+1$ 个向量 $(R_{k,0}(z), \cdots, R_{k,m+n-1}(z))$ 是线性独立的, 它的分量是有理函数. 因此有不全为零的有理函数 $S_l(z)$, 使得 $S_0(z)H^{(0)}(z) + \cdots + S_{m+n}(z)H^{(m+n)}(z) = 0$. (b) 类似, 设 $H(z) = F(z)G(z)$, 对于 $0 \leq l < mn$, 有有理函数 $R_{k,l}(z)$, 具有 $H^{(k)}(z) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} R_{k,ni+j}(z)F^{(i)}(z)G^{(j)}(z)$, 因此对于某些不全为零的有理的 $S_l(z)$, $S_0(z)H^{(0)}(z) + \cdots + S_{mn}(z)H^{(mn)}(z) = 0$. (相似的证明表明如果 $\langle f_n \rangle$ 和 $\langle g_n \rangle$ 是多项式地递归的, 同样 $\langle f_n + g_n \rangle$ 和 $\langle f_n g_n \rangle$ 也是多项式地递归的. 顺便提到, 对于商没有类似的结果; 例如, $\cos z$ 是可微有限, 但是 $1/\cos z$ 不是.)

7.56 顺便提到, Euler 证明此数也是 $[z^n]1/\sqrt{1-2z-3z^2}$, 且他得到公式 $a_n = \sum_{k=0}^{2n} n^{2k} / k!^2$. 他还发现, 当研究这些数时, “归纳法失败”. 虽然对于 $0 \leq n < 9$, $3a_n - a_{n+1}$ 等于 $F_{n-1}(F_{n-1} + 1)$, 当 n 为 9 或比 9 大时, 这个经验律不可思议地失败!

7.57 (对于解出此问题者, Paul Erdos 现时提供 \$500.)

8.1 $(1/24) + (1/48) + (1/48) + (1/48) + (1/48) + (1/24) = 1/6$. (事实上, 当至少一个骰子是完好时, 我们总以概率 $1/6$ 取得成双.) 以分布 Pr_1 任何两个骰子的和为 7 具有相同概率, 所以如同成双那样, $S=7$ 具有相同概率.

8.2 指定最上和最下的两张牌有 12 种方式, 有 $50!$ 种方式排列其他的牌, 所以概率为 $12 \cdot 50! / 52! = 12 / (51 \cdot 52) = 1 / 17 \cdot 13 = 1 / 221$.

8.3 $(1/10)(3+2+\cdots+9+2) = 4.8$; $(1/9)(3^2 + 2^2 + \cdots + 9^2 + 2^2 - 10(4.8)^2) = 388/45 \approx 8.6$, 具有一个完好钱币的平均和方差为 6 和 22, 所以 Stanford 有一个不寻常的正面向

上的组, 对应的 Princeton 数字为 6.4 和 $562/45 \approx 12.5$. (此分布有 $\kappa_4 = 2974$, 它是相当大的, 因此当 $n=10$ 也相当大时, 按照习题 54, 此方差的标准差估计为 $\sqrt{2974/10 + 2(22)^2/9} \approx 20.1$, 人们不能抱怨学生作弊.)

8.4 依据式(8.38)和(8.39)得到这一点, 因为 $F(z) = G(z)H(z)$. (对所有累积量, 一个类似的公式成立, 纵然 $F(z)$ 和 $G(z)$ 可有负系数.)

8.5 用 p 替换 H , 用 $q=1-p$ 替换 T . 如果 $S_A = S_B = 1/2$, 我们有 $p^2 q N = 1/2$ 和 $p q^2 N = (1/2)q + (1/2)$; 解为 $p=1/\varphi^2$, $q=1/\varphi$.

8.6 此时对所有 y , $X|y$ 有 X 那样的相同分布, 因此 $E(X|Y) = EX$ 是常量和 $V(E(X|Y)) = 0$. 还有 $V(X|Y)$ 是常量且等于它的期望值.

8.7 用第二章的 Chebyshev 求和不等式, 我们有 $1 = (p_1 + p_2 + \cdots + p_6)^2 \leq 6(p_1^2 + p_2^2 + \cdots + p_6^2)$.

8.8 设 $p = \Pr(\omega \in A \cap B)$, $q = \Pr(\omega \notin A)$ 和 $r = \Pr(\omega \notin B)$. 则 $p+q+r=1$, 且被证明的等式是 $p = (p+r)(p+q) - qr$.

8.9 这是真的(服从 F 和 G 定义在 X 和 Y 的各自范围上的明显条件), 因为

$$\begin{aligned} \Pr(F(X)=f \text{ 和 } G(Y)=g) &= \sum_{\substack{x \in F^{-1}(f) \\ y \in G^{-1}(g)}} \Pr(X=x \text{ 和 } Y=y) \\ &= \sum_{\substack{x \in F^{-1}(f) \\ y \in G^{-1}(g)}} \Pr(X=x) \cdot \Pr(Y=y) \\ &= \Pr(F(X)=f) \cdot \Pr(G(Y)=g). \end{aligned}$$

8.10 2 个. 设 $x_1 < x_2$ 是中位数, 则 $1 \leq \Pr(X \leq x_1) + \Pr(X \geq x_2) \leq 1$, 因此等式成立. (有些离散分布没有中位元素. 例如, 设 Ω 是形式 $\pm 1/n$ 的所有分数的集合, 具有 $\Pr(+1/n) = \Pr(-1/n) = ((\pi^2/12) n^{-2})$.)

8.11 例如, 对于所有整数 $k \geq 0$, 设 $K=k$ 具有概率 $4/(k+1)(k+2)(k+3)$. 则 $EK=1$, 但是 $E(K^2) = \infty$. (我们能类似地通过 κ_m 构造有有限累积量的随机变量, 但是 $\kappa_{m+1} = \infty$.)

8.12 (a) 设 $p_k = \Pr(X=k)$. 如果 $0 < x \leq 1$, 我们有 $\Pr(X \leq r) = \sum_{k \leq r} p_k \leq \sum_{k \leq r} x^{k-r} p_k \leq \sum_k x^{k-r} p_k = x^{-r} P(x)$. 另一个不等式有类似的证明.

(b) 设 $x = \alpha/(1-\alpha)$ 使右边减至最小. (习题 9.42 中得到给定和的更精确的估计.)

8.13 (Boris Pittel 提出的解.) 让我们置 $Y = (X_1 + \cdots + X_n)/n$ 和 $Z = (X_{n+1} + \cdots + X_{2n})/n$. 则

$$\begin{aligned}
 & \Pr\left(\left|\frac{Y+Z}{2} - \alpha\right| \leq |Y - \alpha|\right) \\
 & \geq \Pr\left(\left|\frac{Y-\alpha}{2}\right| + \left|\frac{Z-\alpha}{2}\right| \leq |Y - \alpha|\right) \\
 & = \Pr(|Z - \alpha| \leq |Y - \alpha|) \geq \frac{1}{2}.
 \end{aligned}$$

关于任何离散概率分布, 最后的不等式为 ' $>$ ', 因为 $\Pr(Y=Z) > 0$.

8.14 均值 $(H) = p$ 均值 $(F) + q$ 均值 (G) ; 方差 $(H) = p$ 方差 $(F) + q$ 方差 $(G) + pq$ (均值 $(F) - \text{均值}(G))^2$. (一种混合实际是条件概率的特殊情形: 设 Y 是硬币, 设由 $F(z)$ 生成 $X|H$, 且设 $G(z)$ 生成 $X|T$, 则 $YX = EV(X|Y) + VE(X|Y)$, 其中 $EV(X|Y) = pV(X|H) + qV(X|T)$ 以及 $VE(X|Y)$ 是 $pz^{\text{均值}(F)} + qz^{\text{均值}(G)}$ 的方差.)

8.15 依据链规则, $H'(z) = G'(z)F'(G(z))$, $H''(z) = G''(z)F'(G(z)) + G'(z)^2 F''(G(z))$. 因此

均值 $(H) = \text{均值}(F)\text{均值}(G)$;

方差 $(H) = \text{方差}(F)\text{均值}(G)^2 + \text{均值}(F)\text{方差}(G)$.

(对应于概率分布 H 的随机变量可理解如下: 由分布 F 确定一个非负整数 n , 然后加 n 个具有分布 G 的独立随机变量的值. 当 X 有分布 H , Y 有分布 F 时, 此习题的方差等式是式(8.105)的特殊情形.)

$$8.16 \quad e^{w(x-1)} / (1-w).$$

8.17 $\Pr(Y_{n,p} \leq m) = \Pr(Y_{n,p} + n \leq m+n) =$ 我们需要 $\leq m+n$ 次抛掷得 n 次正面的概率 $= m+n$ 次抛掷产生 $\geq n$ 个正面的概率 $= \Pr(X_{m+n,p} \geq n)$. 因此

$$\begin{aligned}
 \sum_{k \leq n} \binom{n+k-1}{k} p^n q^k &= \sum_{k \geq n} \binom{m+n}{k} p^k q^{m+n-k} \\
 &= \sum_{k \leq n} \binom{m+n}{k} p^{m+n-k} q^k,
 \end{aligned}$$

这是式(5.19)具有 $n=r$, $x=q$, $y=p$ 的情形.

8.18 (a) $G_X(z) = e^{\mu(z-1)}$. (b) 对所有 $m \geq 1$, 第 m 个累积量为 μ . (在式(8.55)中, 情形 $\mu=1$ 称为 F_∞ .)

8.19 (a) $G_{X_1+X_2}(z) = G_{X_1}(z)G_{X_2}(z) = e^{(\mu_1+\mu_2)(z-1)}$. 因此概率为 $e^{\mu_1+\mu_2} (\mu_1+\mu_2)^n / n!$, 独立 Poisson 变量的和是 Poisson.

(b) 一般, 如果 $K_m X$ 表示随机变量 X 的第 m 次累积量, 当 $a, b \geq 0$ 时我们得到 $K_m(aX_1 + bX_2) = a^m(K_m X_1) + b^m(K_m X_2)$. 因此解答为 $2^m \mu_1 + 3^m \mu_2$.

8.20 一般概率母函数将为 $G(z) = z^m / F(z)$, 其中

$$F(z) = z^m + (1-z) \sum_{k=1}^m \tilde{A}_{(k)} \left[A^{(k)} = A_{(k)} \right] z^{m-k},$$

$$F'(1) = m - \sum_{k=1}^m \tilde{A}_{(k)} \left[A^{(k)} = A_{(k)} \right],$$

$$F''(1) = m(m-1) - 2 \sum_{k=1}^m (m-k) \tilde{A}_{(k)} \left[A^{(k)} = A_{(k)} \right].$$

8.21 这是 $\sum_{n \geq 0} q_n$, 其中 q_n 是 n 次抛掷后在 Alice 和 Bill 之间的游戏仍未完成的概率. 设 p_n 是第 n 次抛掷处游戏终止的概率, 则 $p_n + q_n = q_{n-1}$. 由于 $\lim_{n \rightarrow \infty} nq_n = 0$, 因此游戏的平均次数为 $\sum_{n \geq 1} np_n = (q_0 - q_1) + 2(q_1 - q_2) + 3(q_2 - q_3) + \cdots = q_0 + q_1 + q_2 + \cdots = N$.

建立此解答的另一种方法是用 $(1/2)z$ 替换 H 和 T. 于是式(8.78)中的第一个方程的导数告诉我们 $N(1) + N'(1) = N'(1) + S'_A(1) + S'_B(1)$.

依据方法, $N = 16/3$.

8.22 依据定义我们有 $V(X|Y) = E(X^2|Y) - (E(X|Y))^2$ 和 $V(E(X|Y)) = E((E(X|Y))^2) - (E(E(X|Y)))^2$; 因此 $E(V(X|Y)) + V(E(X|Y)) = E(E(X^2|Y)) - (E(E(X|Y)))^2$. 但是 $E(E(X|Y)) = EX$ 和 $E(E(X^2|Y)) = E(X^2)$, 所以结果就是 VX .

8.23 设 $\Omega_0 = \{\square, \blacksquare\}^2$ 和 $\Omega_1 = \{\square, \blacksquare, \boxtimes, \boxminus\}^2$, 且设 Ω_2 是 Ω 的其他 16 个元素. 于是根据 $\omega \in \Omega_0, \Omega_1, \Omega_2, \Pr_{11}(\omega) - \Pr_{00}(\omega) = +20/576, -7/576, +2/576$. 所以事件 A 一定是从 Ω_j 选取 k_j 个元素, 其中 (k_0, k_1, k_2) 是下列之一: $(0, 0, 0), (0, 2, 7), (0, 4, 14), (1, 4, 4), (1, 6, 11), (2, 6, 1), (2, 8, 8), (3, 8, 15), (3, 10, 5), (3, 12, 12), (4, 12, 2), (4, 14, 9), (4, 16, 16)$. 例如, 有形式 $(2, 6, 1)$ 的 $\binom{4}{2} \binom{16}{6} \binom{16}{1}$ 个事件. 这样的事件的总数为 $[z^0](1+z^{20})^4(1+z^{-7})^{16}(1+z^2)^{16}$, 结果是 1 304 927 002. 如果我们限制事件仅依赖于 S , 我们取得 40 个解 $S \in A$, 其中 $A = \emptyset, \{2, 4, 6\}, \{2, 12, 10, 8\}, \{2, 12, 10, 8, 5, 9\}, \{2, 4, 6, 8, 10, 12\}, \{3, 11, 7, 9, 4, 10\}$, 以及这些集合的补. (这里的记法 $\begin{smallmatrix} 2 \\ 12 \end{smallmatrix}$ 意味着 2 或 12, 但不是两者.)

8.24 (a) J 所占有的任何一个骰子一面朝上的概率为 $p = (1/6) + (5/6)p$, 因此 $p = 6/11$. 设 $q = 5/11$, 则由式(8.61), J 的总占有的概率母函数为 $(q + pz)^{2n+1}$, 具有均值 $(2n+1)p$ 和方差 $(2n+1)pq$.

$$(b) \binom{5}{3} p^3 q^2 + \binom{5}{4} p^4 q + \binom{5}{5} p^5 = 94\,176 / 161\,051 \approx .585.$$

8.25 n 次滚动后的现时赌注的概率母函数为 $G_n(z)$, 其中

$$G_0(z) = z^A;$$

$$G_n(z) = \sum_{k=1}^6 G_{n-1}(z^{2(k-1)/5}) / 6, \text{ 对于 } n > 0.$$

(非整数指数不引起麻烦.) 由此得到均值 $(G_n) = \text{均值}(G_{n-1})$, 以及方差 $(G_n) + \text{均值}(G_n)^2 = 22/15(\text{方差}(G_{n-1}) + \text{均值}(G_{n-1})^2)$. 所以均值总是 A , 但是方差变成 $((22/15)^n - 1)A^2$.

8.26 概率母函数 $F_{l,n}(z)$ 满足 $F'_{l,n}(z) = F_{l,n-1}(z)/l$, 因此均值 $(F_{l,n}) = F'_{l,n}(1) = [n \geq 1]/l$ 和 $F''_{l,n}(1) = [n \geq 2]/l^2$, 容易计算方差. (事实上, 我们有

$$F_{l,n}(z) = \sum_{0 \leq k \leq n/l} \frac{1}{k!} \left(\frac{z-1}{l} \right)^k,$$

当 $n \rightarrow \infty$ 时, 它趋于具有均值 $1/l$ 的 Poisson 分布.)

8.27 $(n^2 \sum_3 - 3n \sum_2 \sum_1 + 2 \sum_1^3) / n(n-1)(n-2)$ 有希望的均值, 其中 $\sum_k = X_1^k + \dots + X_n^k$. 根据等式

$$E \sum_3 = n \mu_3,$$

$$E(\sum_2 \sum_1) = n \mu_3 + n(n-1) \mu_2 \mu_1,$$

$$E(\sum_1^3) = n \mu_3 + 3n(n-1) \mu_2 \mu_1 + n(n-1)(n-2) \mu_1^3,$$

得到这个结果. 顺便提到, 第 3 个累积量为 $\kappa_3 = E((X - EX)^3)$. 但是第 4 个累积量没有这样一个简单的表达式, 我们有 $\kappa_4 = E((X - EX)^4) - 3(VX)^2$.

8.28 (习题含蓄地要求 $p = q = 1/2$, 但是为了完整性, 这里给的是一般解答.) 用 pz 替换 H , qz 替换 T , 得到 $S_A(z) = p^2 q z^3 / (1 - pz)(1 - qz)(1 - pqz^2)$ 和 $S_B(z) = pq^2 z^3 / (1 - qz)(1 - pqz^2)$. Alice 在第 n 次抛掷获胜的条件概率的概率母函数为

$$\frac{S_A(z)}{S_A(1)} = z^3 \cdot \frac{q}{1 - pz} \cdot \frac{p}{1 - qz} \cdot \frac{1 - pq}{1 - pqz^2}.$$

这是伪概率母函数的乘积, 它的均值为 $3 + p/q + q/p + 2pq/(1 - pq)$. Bill 的公式是相同的, 但是没有因子 $q/(1 - pz)$, 所以 Bill 的均值为 $3 + q/p + 2pq/(1 - pq)$. 当 $p = q = 1/2$ 时, 情形(a)中的解答为 $17/3$; 情形(b)为 $14/3$. Bill 往往仅胜一半, 但是当他获胜时, 他势必马上获胜. 抛掷的一般平均数为 $(2/3) \cdot (17/3) + (1/3) \cdot (14/3) = 16/3$, 与习题 21 一致. 每种型式的单人纸牌游戏有等待时间 8.

8.29 在

$$1 + N(H + T) = N + S_A + S_B + S_C$$

$$NHHTH = S_A(1 + HTH) + S_B(HTH + TH + 1) + S_C(HTH + TH)$$

$$NHTHH = S_A(THH + H) + S_B(THH + 1) + S_C(THH)$$

$$NTHHH = S_A(HH) + S_B(HH) + S_C$$

中置 $H = T = 1/2$, 取得获胜概率. 一般我们将有 $S_A + S_B + S_C = 1$ 和

$$\begin{aligned} S_A(A : A) + S_B(B : A) + S_C(C : A) &= S_A(A : B) + S_B(B : B) + S_C(C : B) \\ &= S_A(A : B) + S_B(B : C) + S_C(C : C). \end{aligned}$$

特别, 方程 $9S_A + 3S_B + 3S_C = 5S_A + 9S_B + S_C = 2S_A + 4S_B + 9S_C$ 意味着 $S_A = 16/52$, $S_B = 17/52$, $S_C = 19/52$.

8.30 $P(h_1, \dots, h_n; k)$ 的方差是改变的二项分布 $((m-1+z)/m)^{k-1}z$ 的方差, 依据式(8.61), 它是 $(k-1)(1/m)(1-(1/m))$. 因此方差的平均是均值 $(S)(m-1)/m^2$. 平均的方差是 $(k-1)/m$ 的方差, 即 $\text{Var}(S)/m^2$. 依据式(8.105), 这两个量的和应是 VP , 且它就是. 事实上, 我们以不易识别的方式再作式(8.95)的推导. (见习题 15.)

8.31 一种强力的解法设置 5 个未知数的 5 个方程: $A = (1/2)zB + (1/2)zE$; $B = (1/2)zC$; $C = 1 + (1/2)zB + (1/2)zD$; $D = (1/2)zC + (1/2)zE$; $E = (1/2)zD$. 但是位置 C 和 D 与目标是等距离的, B 和 E 也是, 所以我们可以把它们归在一起. 如果 $X = B + E$ 和 $Y = C + D$, 现在有 3 个方程:

$$A = \frac{1}{2}zX; X = \frac{1}{2}zY; Y = 1 + \frac{1}{2}zX + \frac{1}{2}zY.$$

因此 $A = z^2 / (4 - 2z - z^2)$, 我们有均值 $(A) = 6$ 和方差 $(A) = 22$. (想起这是什么吗? 事实上, 这个问题等价于抛掷一个完好硬币直到在一行中取得两次正面: 正面意味着“朝苹果前进”, 反面意味着“回去”.)(b) Chebyshev 不等式说, $\Pr(S \geq 100) = \Pr((S - 6)^2 \geq 94^2) \leq 22/94^2 \approx .0025$. (c) 第二个末端等式说, 对所有 $x \geq 1$, $\Pr(S \geq 100) \leq 1/x^{98} (4 - 2x - x^2)$, 当 $x = (\sqrt{49001} - 99)/100$, 我们得到上界 0.000 000 05. (按照习题 37, 实际概率近似为 0.000 000 000 9.)

8.32 依据对称性, 我们可把每个月的位置化为 4 种可能性之一:

D , 对角相对的州;

A , 邻接且不是 Kansas 的州;

K , Kansas 和另外一个州;

S , 相同的州.

考虑 Markovian 转换, 我们可得 4 个方程

$$D = 1 + z \left(\frac{2}{9}D + \frac{2}{12}K \right)$$

$$A = z \left(\frac{4}{9} A + \frac{4}{12} K \right)$$

$$K = z \left(\frac{4}{9} D + \frac{4}{9} A + \frac{4}{12} K \right)$$

$$S = z \left(\frac{3}{9} D + \frac{1}{9} A + \frac{2}{12} K \right)$$

它的和是 $D + K + A + S = 1 + z(D + A + K)$ 。解为

$$S = \frac{81z - 45z^2 - 4z^3}{243 - 243z + 24z^2 + 8z^3},$$

但是找均值和方差的最简方法可写 $z = 1+w$ ，且以 w 的幂展开，忽略 w^2 的倍数：

$$D = \frac{27}{16} + \frac{1593}{512} w + \cdots;$$

$$A = \frac{9}{8} + \frac{2115}{256} w + \cdots;$$

$$K = \frac{15}{8} + \frac{2661}{256} w + \cdots.$$

现在 $S'(1) = 27/16 + 9/8 + 15/8 = 75/16$ 和 $(1/2)S''(1) = 1593/512 + 2115/256 + 2661/256 = 11145/512$ ，均值为 $75/16$ 和方差为 $105/4$ 。(有较简的方法吗?)

8.33 第一个解答：显然是的，因为散列值 h_1, \dots, h_n 是独立的。第二个解答：一定不是，虽然散列值 h_1, \dots, h_n 是独立的。我们有 $\Pr(X_j = 0) = \sum_{k=1}^n s_k ([j \neq k] (m-1)/m) = (1-s_j)(m-1)/m$ ，但是 $\Pr(X_1 = X_2 = 0) = \sum_{k=1}^n s_k [k > 2] (m-1)^2/m^2 = (1-s_1-s_2)(m-1)^2/m^2 \neq \Pr(X_1 = 0)\Pr(X_2 = 0)$ 。

8.34 设 $[z^n]S_m(z)$ 是转 n 圈后 Gina 前进 $< m$ 步的概率。则 $S_m(1)$ 是 n 穴标准上她的平均得分； $[z^m]S_m(z)$ 是她相对于一个沉着对手输这样一个穴的概率； $1-[z^{m-1}]S_m(z)$ 是她胜它的概率。我们有递归

$$S_0(z) = 0;$$

$$S_m(z) = (1 + pzS_{m-2}(z) + qzS_{m-1}(z)) / (1 - rz), \quad m > 0.$$

为了解(a)，计算 $m, n \leq 4$ 的系数就够了；用 $100w$ 替换 z 是方便的，仅涉及整数计算。我们得到下列系数表：

S_0	0	0	0	0	0
S_1	1	4	16	64	256
S_2	1	95	744	4 432	23 552
S_3	1	100	9 065	104 044	819 808
S_4	1	100	9 975	868 535	12 964 304

所以 Gina 以概率 $1 - .868\,535 = .131\,465$ 胜, 以概率 $.129\,643\,04$ 输。(b) 为了找平均打数, 我们计算

$$S_1(1) = \frac{25}{24}; S_2(1) = \frac{4\,675}{2\,304}; S_3(1) = \frac{667\,825}{221\,184}; S_4(1) = \frac{85\,134\,475}{21\,233\,664}.$$

(顺便提到, $S_5(1) \approx 4.999\,5$; 在 5 穴标准上, 相对于穴和打数她胜, 但当标准为 3 时, 她都输。

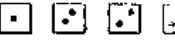
8.35 依据中国剩余定理, 对于所有 n 条件将为真, 当且仅当 $n=1$ 它为真。一个充分必要条件是多项式等式

$$\begin{aligned} & (p_2 + p_4 + p_6 + (p_1 + p_3 + p_5)w)(p_3 + p_6 + (p_1 + p_4)z + (p_2 + p_5)z^2) \\ & = (p_1 wz + p_2 z^2 + p_3 w + p_4 z + p_5 wz^2 + p_6), \end{aligned}$$

但是这就是有点儿重新陈述问题。一种较简单的特性是

$$(p_2 + p_4 + p_6)(p_3 + p_6) = p_6, (p_1 + p_3 + p_5)(p_2 + p_5) = p_5,$$

在原来乘积中它仅检验两个系数。一般解有 3 个自由度: 设 $a_0 + a_1 = b_0 + b_1 + b_2 = 1$, 且置 $p_1 = a_1 b_1$, $p_2 = a_0 b_2$, $p_3 = a_1 b_0$, $p_4 = a_0 b_1$, $p_5 = a_1 b_2$, $p_6 = a_0 b_0$ 。

8.36 (a) 。 (b) 如果第 k 个骰子有 s_1, \dots, s_6 点的面, 设 $p_k(z) = z^{s_1} + \dots + z^{s_6}$ 。我们要找具有 $p_1(z) \cdots p_n(z) = (z + z^2 + z^3 + z^4 + z^5 + z^6)^n$ 这样的多项式。具有有理系数的这个多项式的不可约因子为 $z^n(z+1)^n(z^2+z+1)^n(z^2-z+1)^n$, 因此 $p_k(z)$ 一定为形式 $z^{a_k}(z+1)^{b_k}(z^2+z+1)^{c_k}(z^2-z+1)^{d_k}$ 。我们一定有 $a_k > 1$, 因为 $p_k(0) = 0$; 且事实上 $a_k = 1$, 因为 $a_1 + \dots + a_n = n$ 。并且条件 $p_k(1) = 6$ 意味着 $b_k = c_k = 1$ 。现在容易明白 $0 < d_k < 2$, 因为 $d_k > 2$ 给出负系数。当 $d=0$ 和 $d=2$ 时, 在 (a) 中我们得到 2 个骰子; 所以如同 (a) 中那样对某个 $k < (1/2)n$ 仅有 k 对骰子的解, 加上 $n-2k$ 个通常的骰子。

8.37 由多米诺骨牌和硬币抛掷之间的关系, 对所有 $n > 0$, 长 n 的硬币上抛序列数为 F_{n-1} 。所以当硬币完好时, 确实需要 n 次上抛的概率为 $F_{n-1} / 2^n$ 。还有 $q_n = F_{n+1} / 2^{n-1}$, 因为 $\sum_{k \geq n} F_k z^n = (F_n z^n + F_{n+1} z^{n+1}) / (1 - z - z^2)$ 。(当然, 通过母函数的惯常的解也是可能的。)

8.38 当已看出 k 面时, 滚动一新骰子的工作等价于成功概率 $p_k = (m-k)/m$ 的抛掷硬币。因此概率母函数为 $\prod_{k=0}^{l-1} p_k z / (1 - q_k z) = \prod_{k=0}^{l-1} (m-k)z / (m-kz)$ 。均值是 $\sum_{k=0}^{l-1} p_k^{-1} = m(H_m - H_{m-l})$; 方差是 $m^2(H_m^{(2)} - H_{m-l}^{(2)}) - m(H_m - H_{m-l})$; 方程 (7.47) 提供了所要求的概率的闭形式, 即 $m^{-n} m! \left\{ \begin{matrix} n-1 \\ l-1 \end{matrix} \right\} / (m-l)!$ (通常称此习题中讨论的问题为“收集配给

票”。)

$$8.39 \quad E(x) = P(-1); \quad V(X) = P(-2) - P(-1)^2; \quad E(\ln X) = -P'(0).$$

8.40 (a) 依据式(7.49), 我们有 $\kappa_m = n(0! \left\{ \begin{matrix} m \\ 1 \end{matrix} \right\} p - 1! \left\{ \begin{matrix} m \\ 2 \end{matrix} \right\} p^2 + 2! \left\{ \begin{matrix} m \\ 3 \end{matrix} \right\} p^3 - \cdots)$. 顺便提到, 第3个累积量为 $npq(q-p)$, 第4个是 $npq(1-6pq)$. 等式 $q + pe^t = (p + qe^{-t})e^t$ 表明 $f_m(p) = (-1)^m f_m(q) + [m=1]$; 因此我们能写 $f_m(p) = g_m(pq) \times (q-p)^{[m/2]}$, 其中 g_m 是一个次数为 $\lfloor m/2 \rfloor$ 的多项式(每当 $m > 1$).

(b) 设 $p = 1/2$ 和 $F(t) = \ln((1/2) + (1/2)e^t)$. 则 $\sum_{m \geq 1} \kappa_m t^{m-1} / (m-1)! = F'(t) = 1 - 1/(e^t + 1)$, 且我们能用习题 6.23.

8.41 如果 $G(z)$ 是仅接受正整数值的随机变量 X 的概率母函数, 则 $\int_0^1 G(z) dz / z = \sum_{k \geq 1} \Pr(X=k) / k = E(X^{-1})$. 如果 X 是得到 $n+1$ 个正面的抛掷数的分布, 我们依据式(8.59)得到 $G(z) = (pz / (1 - qz))^{n+1}$, 如果我们代换 $w = pz / (1 - qz)$, 积分为

$$\int_0^1 \left(\frac{pz}{1 - qz} \right)^{n+1} \frac{dz}{z} = \int_0^1 \frac{w^n dw}{1 + (q/p)w}.$$

当 $p = q$ 时被积式可写为 $(-1)^n ((1+w)^{-1} - 1 + w - w^2 + \cdots + (-1)^n w^{n-1})$. 所以积分为 $(-1)^n (\ln 2 - 1 + 1/2 - 1/3 + \cdots + (-1)^n / n)$. 依据式(9.28)我们有 $H_{2n} - H_n = \ln 2 - (1/4)n^{-1} + (1/16)n^{-2} + O(n^{-4})$. 由此得到 $E(X_{n+1}^{-1}) = (1/2)n^{-1} - (1/4)n^{-2} + O(n^{-4})$.

8.42 如果某人开始分别未被雇用或被雇用, 设 $F_n(z)$ 和 $G_n(z)$ 分别是晚上雇用次数的概率母函数. 设 $q_h = 1 - p_h$ 和 $q_f = 1 - p_f$. 则 $F_0(z) = G_0(z) = 1$, 并且

$$F_n(z) = p_h z G_{n-1}(z) + q_h F_{n-1}(z);$$

$$G_n(z) = p_f F_{n-1}(z) + q_f z G_{n-1}(z).$$

由超母函数

$$G(w, z) = \sum_{n \geq 0} G_n(z) w^n = \frac{A(w)}{(1 - zB(w))}$$

给出解, 其中 $B(w) = w(q_f - (q_f - p_h)w) / (1 - q_h w)$ 和 $A(w) = (1 - B(w)) / (1 - w)$. 现在 $\sum_{n \geq 0} G_n'(1) w^n = \alpha w / (1 - w)^2 + \beta / (1 - w) - \beta / (1 - (q_f - p_h)w)$, 其中

$$\alpha = \frac{p_h}{p_h + p_f}, \quad \beta = \frac{p_f(q_f - p_h)}{(p_h + p_f)^2};$$

因此 $G'_n(1) = \alpha n + \beta(1 - (q_f - p_A)^n)$. (类似, $G''_n(1) = \alpha^2 n^2 + O(n)$, 所以方差为 $O(n)$.)

8.43 依据式(6.11), $G_n(z) = \sum_{k \geq 0} \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] z^k / n! = z^n / n!$. 这是二项概率母函数的乘积 $\prod_{k=1}^n ((k-1+z)/k)$, 其中第 k 个有均值 $1/k$ 和方差 $(k-1)/k^2$; 因此均值 $(G) = H_n$ 和方差 $(G_n) = H_n - H_n^{(2)}$.

8.44 (a) 冠军一定在 n 轮中不被击败, 所以解答为 p^n . (b, c) 在不同的局部比赛中队员 x_1, \dots, x_{2^k} 一定(按机会)被“安排”, 且他们一定胜所有 $2^k(n-k)$ 个他们的对手. 以 2^n 种方式能放比赛树的 2^n 个叶子. 为了安排它我们有 $2^k!(2^{n-k})^{2^k}$ 种方式来放最高的 2^k 个队员, $(2^n - 2^k)!$ 种方式放其他的. 因此概率为 $(2p)^{2^k(n-k)} / \binom{2^n}{2^k}$. 当 $k=1$ 时, 此简化为 $(2p^2)^{n-1} / (2^n - 1)$. (d) 每个比赛结果对应于队员的一个排列: 设 y_1 是冠军; 设 y_2 是另外的决赛选手; 设 y_3 和 y_4 在半决赛中输给 y_1 和 y_2 的队员; 设 (y_5, \dots, y_8) 是四分之一决赛中分别输给 (y_1, \dots, y_4) 的队员, 等等. (另一个证明表明第一轮本质上有 $2^n! / 2^{n-1}!$ 不同的结果; 第二轮有 $2^{n-1}! / 2^{n-2}!$, 等等.) (e) 设 S_k 是第 k 轮中 x_2 的 2^{k-1} 个可能对手的集合. 给定 x_1 属于 S_k , x_2 胜的条件概率为

$$\begin{aligned} \Pr(x_1 \text{ 赛 } x_2) \cdot p^{n-1}(1-p) + \Pr(x_1 \text{ 不赛 } x_2) \cdot p^n \\ = p^{k-1} p^{n-1}(1-p) + (1-p^{k-1}) p^n. \end{aligned}$$

$x_1 \in S_k$ 的机会是 $2^{k-1} / (2^n - 1)$, 对 k 求和得解答:

$$\sum_{k=1}^n \frac{2^{k-1}}{2^n - 1} (p^{k-1} p^{n-1}(1-p) + (1-p^{k-1}) p^n) = p^n - \frac{(2p)^n - 1}{2^n - 1} p^{n-1}.$$

(f) $2^n!$ 个比赛结果的每一个有一个确定的概率出现, x_j 胜的概率是 x_j 为冠军的所有 $(2^n - 1)!$ 个比赛结果上这些概率的和. 在所有那些结果中考虑用 x_{j+1} 替换 x_j ; 如果 x_j 和 x_{j+1} 不相遇, 这个替换不影响概率, 但是如果它们相遇, 则用 $(1-p)/p < 1$ 乘概率.

8.45 (a) $A(z) = 1/(3-2z)$; $B(z) = zA(z)^2$; $C(z) = z^2A(z)^3$. 当装瓶时, 葡萄酒的概率母函数是 $z^3A(z)^3$, 它是 z^3 乘上参数为 $n=3$, $p=1/3$ 的一个负二项分布.

(b) 平均值 $(A)=2$, 方差 $(A)=6$; 平均值 $(B)=5$, 方差 $(B)=2$ 方差 $(A)=12$; 平均值 $(C)=B$, 方差 $(C)=18$. 葡萄酒的平均龄期是 9 年老. 25 年老的部份是 $\binom{-3}{22}$

$$\times (-2)^{22} 3^{-25} = \binom{24}{22} 2^{22} 3^{-25} = 23 \cdot \left(\frac{2}{3}\right)^{24} \approx .00137.$$

(c) 设 w^n 的系数是年 n 开始处的概率母函数. 则

$$A = \left(1 + \frac{1}{3}w/(1-w)\right) / \left(1 - \frac{2}{3}zw\right);$$

$$B = \left(1 + \frac{1}{3}zwA\right) / \left(1 - \frac{2}{3}zw\right);$$

$$C = \left(1 + \frac{1}{3}zwB\right) / \left(1 - \frac{2}{3}zw\right).$$

对 z 求导且置 $z=1$, 这就有

$$C' = \frac{q}{1-w} - \frac{1/2}{\left(1 - \frac{2}{3}w\right)^3} - \frac{3/2}{\left(1 - \frac{2}{3}w\right)^2} - \frac{6}{1 - \frac{2}{3}w}.$$

在过程开始之后, 装瓶葡萄酒 n 年的平均龄期比 w^{n-1} 的系数大 1, 即 $9 - \left(\frac{2}{3}\right)^n (3n^2 + 21n + 72)/8$. (当 $n=11$ 时, 这已经超过 8.)

8.46 (a) $P(w, z) = 1 + (1/2)(wP(w, z) + zP(w, z)) = (1 - (1/2)(w+z))^{-1}$, 因此 $p_{mn} = 2^{-m-n} \binom{m+n}{n}$.

(b) $P_k(w, z) = (1/2)(w^k + z^k)P(w, z)$, 因此

$$p_{k, m, n} = 2^{k-1-m-n} \left(\binom{m+n-k}{m} + \binom{m+n-k}{n} \right).$$

(c) $\sum_k k p_{k, m, n} = \sum_{k=0}^n k 2^{k-2n} \binom{2n-k}{n} = \sum_{k=0}^n (n-k) 2^{-n-k} \binom{n+k}{n}$, 用式 (5.20)

可对此式求和:

$$\begin{aligned} & \sum_{k=0}^n 2^{-n-k} \left((2n+1) \binom{n+k}{n} - (n+1) \binom{n+1+k}{n+1} \right) \\ &= (2n+1) - (n+1) 2^{-n} \left(2^{n+1} - 2^{-n-1} \binom{2n+2}{n+1} \right) \\ &= \frac{2n+1}{2^{2n}} \binom{2n}{n} - 1. \end{aligned}$$

(第九章的方法证明这是 $2\sqrt{n/\pi} - 1 + O(n^{-1/2})$.)

8.47 在 n 次扩散后有 $n+2$ 个等可能受体. 设随机变量 X_n 表示目前 2 个受体的噬菌体的个数; 则 $X_{n+1} = X_n + Y_n$, 其中如果第 $n+1$ 个粒子击中一个 2 个受体的噬菌体(条件概率 $2X_n/(n+2)$), 则 $Y_n = -1$, 否则 $Y_n = +2$. 因此

$$EX_{n+1} = EX_n + EY_n = EX_n - 2EX_n/(n+2) + 2(1 - 2EX_n/(n+2)).$$

如果我们用求和因子 $(n+1)^{\frac{5}{2}}$ 乘两端, 能解递归式 $(n+2)EX_{n+1} = (n-4)EX_n + 2n+4$; 或者我们能猜出解答且用归纳法证明它: 对所有 $n \geq 4$, $EX_n = (2n+4)/7$. (顺便提到,

在 5 步后总有两个 2 受体的噬菌体和一个 3 受体的噬菌体。不管 4 步后的结构如何。)

8.48 (a) 飞盘之间的距离(测量使得它为一偶数)为 0, 2, 或 4 个单位, 起始为 4. 对应的母函数 A, B, C (其中, 譬如说, $[z^n]C$ 是 n 个抛掷后距离为 4 的概率)满足

$$A = \frac{1}{2}zB, \quad B = \frac{1}{2}zB + \frac{1}{4}zC, \quad C = 1 + \frac{1}{4}zB + \frac{3}{4}zC.$$

由此得到 $A = z^2 / (16 - 20z + 5z^2) = z^2 / F(z)$, 且我们有均值 $(A) = 2$ - 均值 $(F) = 12$, 方差 $(A) = -$ 方差 $(F) = 100$. (一个较困难, 但是很有趣的解因子 A 如下:

$$A = \frac{p_1 z}{1 - q_1 z} \cdot \frac{p_2 z}{1 - q_2 z} = \frac{p_2}{p_2 - p_1} \frac{p_1 z}{1 - q_1 z} + \frac{p_1}{p_1 - p_2} \frac{p_2 z}{1 - q_2 z},$$

其中 $p_1 = \varphi^2 / 4 = (3 + \sqrt{5}) / 8$, $p_2 = \hat{\varphi}^2 / 4 = (3 - \sqrt{5}) / 8$, 以及 $p_1 + q_1 = p_2 + q_2 = 1$. 因此, 游戏等价于有两个偏的硬币, 它的正面概率为 p_1 和 p_2 , 每次抛硬币直到它们出现正面, 且抛的总次数将与飞盘抛掷数有相同分布. 这样两个硬币等待时间的均值和方差分别为 $6 \pm 2\sqrt{5}$ 和 $50 \pm 22\sqrt{5}$, 因此如同前面那样, 总均值和方差为 12 和 100.)

(b) 展开母函数成部分分式使它可能来计算概率。(注意 $\sqrt{5} / (4\varphi) + \varphi^2 / 4 = 1$, 所以能以 φ 的幂来表解答。)游戏将持续多于 n 步的概率为 $5^{(n-1)/2} 4^{-n} (\varphi^{n+2} - \varphi^{-n-2})$, 当 n 为偶数时, 此为 $5^{n/2} 4^{-n} F_{n+2}$. 所以解答为 $5^{30} 4^{-100} F_{102} \approx .00006$.

8.49 (a) 如果 $n > 0$, $P_N(0, n) = (1/2)[N=0] + (1/4)P_{N-1}(0, n) + (1/4)P_{N-1}(1, n-1)$; $P_N(m, 0)$ 是相似的; $P_N(0, 0) = [N=0]$. 因此,

$$g_{m,n} = \frac{1}{4}zg_{m-1,n+1} + \frac{1}{2}zg_{m,n} + \frac{1}{4}zg_{m+1,n-1};$$

$$g_{0,n} = \frac{1}{2} + \frac{1}{4}zg_{0,n} + \frac{1}{4}g_{1,n-1}, \text{ 等等.}$$

(b) $g'_{m,n} = 1 + (1/4)g'_{m-1,n+1} + (1/2)g'_{m,n} + (1/4)g'_{m+1,n-1}$; $g'_{0,n} = (1/2) + (1/4)g'_{0,n} + (1/4)g'_{1,n-1}$, 等等. 对 m 用归纳法, 对所有 $m, n \geq 0$, 我们有 $g'_{m,n} = (2m+1)g'_{0,m+n} - 2m^2$. 且由于 $g'_{m,0} = g'_{0,m}$, 我们一定有 $g'_{m,n} = m+n+2mn$.

(c) 当 $mn > 0$ 时, 满足递归, 因为

$$\sin(2m+1)\theta = \frac{1}{\cos^2\theta} \left(\frac{\sin(2m-1)\theta}{4} + \frac{\sin(2m+1)\theta}{2} + \frac{\sin(2m+3)\theta}{4} \right).$$

这是等式 $\sin(x-y) + \sin(x+y) = 2\sin x \sin y$ 的一个推论. 所以剩下只要检验有界条件.

8.50 (a) 用提示. 我们取得

$$3(1-z)^2 \sum_k \binom{1/2}{k} \left(\frac{8}{9}z\right)^k (1-z)^{2-k}$$

$$= 3(1-z)^2 \sum_k \binom{1/2}{k} \left(\frac{8}{9}\right)^k \sum_j \binom{k+1-3}{j} z^{j+1},$$

现在查看 z^{3+l} 的系数。

$$(b) H(z) = 2/3 + (5/27)z + (1/2) \sum_{l \geq 0} c_{3+l} z^{2+l}.$$

(c) 设 $r = \sqrt{(1-z)(9-z)}$ ，可证明 $(z-3+r)(z-3-r) = 4z$ ，因此 $(r/(1-z) + 2)^2 = (13-5z+4r)/(1-z) = (9-H(z))/(1-H(z))$ 。

(d) 计算 $z=1$ 处第一个导数证明均值 $\langle H \rangle = 1$ ，在 $z=1$ 处的第二个导数发散，所以方差是无限的。

8.51 (a) 设 $H_n(z)$ 是玩 n 轮之后你所拥有的概率母函数，具有 $H_0(z) = z$ ， n 轮的分布是

$$H_{n+1}(z) = H_n(H(z)),$$

所以用归纳法结果是真的(用前面问题的令人惊异的等式)。

(b) $g_n = H_n(0) - H_{n-1}(0) = 4/n(n+1)(n+2) = 4(n-1)^{-3}$ ，均值是 2，方差是无限。

(c) 依据习题 15，你买的彩票的期望值(在第 n 轮)是均值 $\langle H_n \rangle = 1$ 。所以彩票的总期望数是无限的。(因此，最后你几乎肯定输，且在第 2 次游戏后你期望输，而你也期望买无限多张彩票。)

(d) 现在 n 次游戏后的概率母函数是 $H_n(z)^2$ ，且(b)的方法产生 $16 - (4/3)\pi^2 \approx 2.8$ 的一个均值。(这里显示和 $\sum_{k \geq 1} 1/k^2 = \pi^2/6$ 。)

8.52 如果 ω 和 ω' 是具有 $\Pr(\omega) > \Pr(\omega')$ 的事件，则 n 个独立试验序列中 ω 比 ω' 遇到多具有大的概率，因为 ω 出现将十分接近 $n\Pr(\omega)$ 次，所以，当 $n \rightarrow \infty$ 时，独立试验序列中 X 的值的中间数或众数将为随机变量 X 的中位数或众数的概率接近 1。

8.53 我们能推翻此命题，甚至在每个变量是 0 或 1 的特殊情形。设 $p_0 = \Pr(X=Y=Z=0)$ ， $p_1 = \Pr(X=Y=\bar{Z}=0)$ ， \dots ， $p_7 = \Pr(\bar{X}=\bar{Y}=\bar{Z}=0)$ ，其中 $\bar{X} = 1-X$ 。则 $p_0 + p_1 + \dots + p_7 = 1$ ，以及变量成对独立当且仅当我们有

$$(p_4 + p_5 + p_6 + p_7)(p_2 + p_3 + p_6 + p_7) = p_6 + p_7,$$

$$(p_4 + p_5 + p_6 + p_7)(p_1 + p_3 + p_5 + p_7) = p_5 + p_7,$$

$$(p_2 + p_3 + p_6 + p_7)(p_1 + p_3 + p_5 + p_7) = p_3 + p_7.$$

但是 $\Pr(X+Y=Z=0) \neq \Pr(X+Y=0)\Pr(Z=0) \Leftrightarrow p_0 \neq (p_0 + p_1)(p_0 + p_2 + p_4 + p_6)$ 。一个解是

$$p_0 = p_3 = p_5 = p_6 = 1/4; \quad p_1 = p_2 = p_4 = p_7 = 0.$$

这等价于抛两个无偏的硬币且设 $X =$ (第一个硬币是正面), $Y =$ (第二个硬币是正面), $Z =$ (硬币不同). 另一个例子, 所有概率非零, 是

$$p_0 = \frac{4}{64}, \quad p_1 = p_2 = p_4 = \frac{5}{64},$$

$$p_3 = p_5 = p_6 = \frac{10}{64}, \quad p_7 = \frac{15}{64}.$$

就此理由, 我们说 n 个变量 X_1, \dots, X_n 是独立的, 如果

$$\Pr(X_1 = x_1 \text{ 和 } \dots \text{ 和 } X_n = x_n) = \Pr(X_1 = x_1) \cdots \Pr(X_n = x_n),$$

成对独立不足以保证这一点.

8.54 (记号见习题 27.) 我们有

$$E\left(\sum_2^2\right) = n\mu_4 + n(n-1)\mu_2^2,$$

$$E\left(\sum_2 \sum_1^2\right) = n\mu_4 + 2n(n-1)\mu_3\mu_1 + n(n-1)\mu_2^2 + n(n-1)(n-2)\mu_2\mu_1^2,$$

$$E\left(\sum_1^4\right) = n\mu_4 + 4n(n-1)\mu_3\mu_1 + 3n(n-1)\mu_2^2 \\ + 6n(n-1)(n-2)\mu_2\mu_1^2 + n(n-1)(n-2)(n-3)\mu_1^4,$$

由此得到 $V(\hat{V}X) = \kappa_4/n + 2\kappa_2^2/(n-1)$.

8.55 关于 $X=Y$ 有 $A = (1/17) \cdot 52!$ 个置换, 关于 $X \neq Y$ 有 $B = (16/17) \cdot 52!$ 个置换. 所述的过程之后, 关于 $X=Y$ 每个置换以概率 $(1/17)/((1-(16/17)p)A)$ 出现, 因为我们以概率 $(16/17)p$ 返回到步 S1. 类似, 关于 $X \neq Y$ 的每个置换以概率 $(16/17)(1-p)/((1-(16/17)p)B)$ 出现. 选取 $p=1/4$ 使得对所有 x 和 y , $\Pr(X=x \text{ 和 } Y=y) = 1/169$. (所以我们能作两次抛 1 个无偏的硬币且返回 S1, 如果两者出现正面.)

8.56 如果 m 是偶数, 则飞盘总相距奇数距离而游戏永远持续下去. 如果 $m=2l+1$, 则相应的母函数是

$$G_m = \frac{1}{4}zA_1;$$

$$A_1 = \frac{1}{2}zA_1 + \frac{1}{4}zA_2,$$

$$A_k = \frac{1}{4}zA_{k-1} + \frac{1}{2}zA_k + \frac{1}{4}zA_{k+1}, \quad 1 < k < l,$$

$$A_l = \frac{1}{4}zA_{l-1} + \frac{3}{4}zA_l + 1.$$

(系数 $[z^n]A_k$ 是 n 次抛掷之后飞盘之间的距离为 $2k$ 的概率.) 取习题 49 中相似方程的提示, 我们置 $z = 1/\cos^2\theta$ 和 $A_1 = X\sin 2\theta$, 其中 X 是被确定的. 由此依据归纳法得到(不用

A_l 的方程) $A_k = X \sin 2k\theta$. 所以我们要选 X 使得

$$\left(1 - \frac{3}{4\cos^2\theta}\right) X \sin 2l\theta = 1 + \frac{1}{4\cos^2\theta} X \sin(2l-2)\theta.$$

结果 $X = 2\cos^2\theta / \sin\theta \cos(2l+1)\theta$, 因此

$$G_m = \frac{\cos\theta}{\cos m\theta}.$$

当 θ 是 $\pi/(2m)$ 的奇数倍时, 分母为零; 对于 $1 \leq k < l$, $1 - q_k z$ 是分母的根, 且所述的结果表达式一定成立. 为了找均值和方差, 我们能记

$$\begin{aligned} G_m &= \left(1 - \frac{1}{2}\theta^2 + \frac{1}{24}\theta^4 - \dots\right) / \left(1 - \frac{1}{2}m^2\theta^2 + \frac{1}{24}m^4\theta^4 - \dots\right) \\ &= 1 + \frac{1}{2}(m^2-1)\theta^2 + \frac{1}{24}(5m^4-6m^2+1)\theta^4 + \dots \\ &= 1 + \frac{1}{2}(m^2-1)(\tan\theta)^2 + \frac{1}{24}(5m^4-14m^2+9)(\tan\theta)^4 + \dots \\ &= 1 + G'_m(1)(\tan\theta)^2 + \frac{1}{2}G''_m(1)(\tan\theta)^4 + \dots, \end{aligned}$$

因为 $\tan^2\theta = z-1$ 和 $\tan\theta = \theta + (1/3)\theta^3 + \dots$, 所以我们有均值 $(G_m) = (1/2)(m^2-1)$ 和方差 $(G_m) = (1/6)m^2(m^2-1)$. (注意, 这意味着等式)

$$\begin{aligned} \frac{m^2-1}{2} &= \sum_{k=1}^{(m-1)/2} \frac{1}{p_k} = \sum_{k=1}^{(m-1)/2} \left(1 / \sin \frac{(2k-1)\pi}{2m}\right)^2; \\ \frac{m^2(m^2-1)}{6} &= \sum_{k=1}^{(m-1)/2} \left(\cot \frac{(2k-1)\pi}{2m} / \sin \frac{(2k-1)\pi}{2m}\right)^2. \end{aligned}$$

此分布的第三个累积量是 $(1/30)m^2(m^2-1)(4m^2-1)$, 但是在那里好的累积量的因子分解型式中止了. 有十分简单的方法引出均值: 我们有 $G_m + A_1 + \dots + A_l = z(A_1 + \dots + A_l) + 1$, 因此当 $z=1$ 时我们有 $G'_m = A_1 + \dots + A_l$. 由于当 $z=1$ 时 $G_m = 1$, 简单的归纳法证明 $A_k = 4k$.)

8.57 我们有 $A:A \geq 2^{l-1}$ 和 $B:B < 2^{l-1} + 2^{l-3}$ 以及 $B:A \geq 2^{l-2}$, 因此仅当 $A:B > 2^{l-3}$ 时 $B:B - B:A \geq A:A - A:B$ 是可能的. 这意味着 $\tau_2 = \tau_3$, $\tau_1 = \tau_4$, $\tau_2 = \tau_5$, \dots , $\tau_{l-3} = \tau_l$. 于是 $A:A \approx 2^{l-1} + 2^{l-4} + \dots$, $A:B \approx 2^{l-3} + 2^{l-6} + \dots$, $B:A \approx 2^{l-2} + 2^{l-5} + \dots$, $B:B \approx 2^{l-1} + 2^{l-4} + \dots$; 因此 $B:B - B:A$ 终究小于 $A:A - A:B$. (由 Guibas 和 Odlyzko^[138]已得到的漂亮的结果, 证明了 Bill 的可能性总增大到两种型式 $H\tau_1 \dots \tau_{l-1}$ 或 $T\tau_1 \dots \tau_{l-1}$ 之一.)

8.58 按照式(8.82), 我们要 $B:B - B:A > A:A - A:B$. 一个解是 $A = TTHH$, $B = HHH$.

8.59 (a) 依据 $h_k \neq h_n$ 或 $h_k = h_n$ 出现两种情形:

$$G(w, z) = \frac{m-1}{m} \left(\frac{m-2+w+z}{m} \right)^{k-1} w \left(\frac{m-1+z}{m} \right)^{n-k-1} z \\ + \frac{1}{m} \left(\frac{m-1+wz}{m} \right)^{k-1} wz \left(\frac{m-1+z}{m} \right)^{n-k-1} z.$$

(b) 我们能代数地证明, 对 $G(w, z)$ 取 w 和 z 的偏导数且置 $w=z=1$; 或者组合地证明: 无论 h_1, \dots, h_{n-1} 是什么值, $P(h_1, \dots, h_{n-1}, h_n; n)$ 的期望值是相同的 (对 h_n 取平均), 因为散列序列 (h_1, \dots, h_{n-1}) 确定列表大小的序列 (n_1, n_2, \dots, n_m) , 使得所述的期望值是 $((n_1+1) + (n_2+1) + \dots + (n_m+1))/m = (n-1+m)/m$. 所以随机变量 $EP(h_1, \dots, h_n; n)$ 是独立于 (h_1, \dots, h_{n-1}) , 因此独立于 $P(h_1, \dots, h_n; k)$.

8.60 如果 $1 \leq k < l \leq n$, 前面的习题证明在平均的方差中 $s_l s_k$ 的系数为零. 所以我们仅需考虑 s_k^2 的系数, 它是

$$\sum_{1 \leq k_1, \dots, k_n \leq m} \frac{P(h_1, \dots, h_n; k)^2}{m^n} - \left(\sum_{1 \leq k_1, \dots, k_n \leq m} \frac{P(h_1, \dots, h_n; k)}{m^n} \right)^2,$$

$((m-1+z)/m)^{k-1} z$ 的方差; 如同习题 30 中的那样, 这是 $(k-1)(m-1)/m^2$.

8.61 概率母函数 $D_n(z)$ 满足递归

$$D_0(z) = z;$$

$$D_n(z) = z^2 D_{n-1}(z) + 2(1-z^3) D'_{n-1}(z) / (n+1), \quad n > 0.$$

现在我们能导出递归

$$D''_n(1) = (n-11) D''_{n-1}(1) / (n+1) + (8n-2) / 7,$$

对所有 $n \geq 11$, 它有解 $(2/637)(n+2)(26n+15)$ (不管初始条件如何). 因此对 $n \geq 11$, 方差达到 $(12/49)(n+2)(212n+123)$.

8.62 (另一个问题问到, 是否一个给定的表明的累积量序列来自任何分布. 例如, κ_2 一定是非负的, 且 $\kappa_4 + 3\kappa_2^2 = E((X-\mu)^4)$ 至少为 $(E((X-\mu)^2))^2 = \kappa_2^2$, 等等. Hamburger^{[6][144]} 找到了另一问题的充分必要条件.)

8.63 (另外一个问题问到, 是否有一个简单规则告诉是 H 还是 T 更可取.) Conway 猜测无这样的联系存在, 此外 2^l 个顶点上的有向图中仅有一个圈, 从每个序列到它的“最好的打者”有一条弧.

9.1 如果函数全为正的, 为真. 否则我们可能有, 譬如说, $f_1(n) = n^3 + n^2$, $f_2(n) = -n^3$, $g_1(n) = n^4 + n$, $g_2(n) = -n^4$.

9.2 (a) 由于 $(\ln n)^2 < n \ln c$ $n \ln \ln n$, 我们有 $n^{\ln n} < c^n$ $(\ln n)^n$. (b) $n^{\ln \ln c n} < (\ln n)!$

$\sim n^{\ln \ln n}$. (c) 取对数证明 $(n!)!$ 次胜. (d) $P_{[H_n]}^2 \asymp \varphi^{2 \ln n} = n^{2 \ln \varphi}$. 因为 $\varphi^2 - \varphi + 1 < e$, $H_{P_n} \sim n \ln \varphi$ 次胜.

9.3 对于每个 k , $O(n)$ 替换 kn 要求一个不同的 C , 但是每个 O 意味着一个 C . 事实上, 文中的这个 O 要求它代表两个变量 k 和 n 的函数的一个集合. 正确地记为 $\sum_{k=1}^n kn = \sum_{k=1}^n O(n^2) = O(n^3)$.

9.4 例如, $\lim_{n \rightarrow \infty} O(1/n) = 0$, 左边的 $O(1/n)$ 是有常数 C 和 n_0 , 对所有 $n \geq n_0$ 有 $|f(n)| \leq C/n$ 的所有函数 $f(n)$ 的集合. 这集合中的所有函数的极限为 0, 所以左边为单个元素的集合 $\{0\}$. 右边, 没有变量; 0 代表 $\{0\}$, 即所有“无变量的函数, 它的值为零”的(单个元素)集合. (你能明白这里的内在的必然联系吗? 如果没有, 下一年再回到这个问题上来; 你很可能还要处理 O -记法, 即使你不能把你的直观变成严格的形式.)

9.5 设 $f(n) = n^2$ 和 $g(n) = 1$, 则 n 是在左边的集合中, 而不在右边的集合中, 所以命题不真.

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 例如, 设 $f(n) = \lfloor n/2 \rfloor! + n$, $g(n) = (\lceil n/2 \rceil - 1)! \lceil n/2 \rceil! + n$. 顺便提到, 这些函数满足 $f(n) = O(n g(n))$ 和 $g(n) = O(n f(n))$, 更极端的例子完全是可能的.

9.9 (为了完整起见, 我们假设有一个附带的条件 $n \rightarrow \infty$, 以致每个 O 意味着两个常数.) 左边的每个函数有形式 $a(n) + b(n)$, 这里存在常数 m_0, B, n_0, C 使得 $|a(n)| \leq B|f(n)|$ ($n \geq m_0$) 和 $|b(n)| \leq C|g(n)|$ ($n \geq n_0$). 所以对于 $n \geq \max(m_0, n_0)$, 左边函数至多为 $\max(B, C)(|f(n)| + |g(n)|)$, 所以它是右边的一个成员.

9.10 如果 $g(n)$ 属于左边, 以致对某个 y , $g(x) = \cos y$, 其中对某个 C , $|y| \leq C|x|$, 则 $0 \leq 1 - g(x) = 2 \sin^2(y/2) \leq (1/2)y^2 \leq (1/2)C^2 x^2$; 因此右边集合包含左边集合, 公式是真的.

9.11 命题是真的. 如果, 譬如说 $|x| \leq |y|$, 我们有 $(x+y)^2 \leq 4y^2$. 因此 $(x+y)^2 = O(x^2) + O(y^2)$. 因此 $O(x+y)^2 = O((x+y)^2) = O(O(x^2) + O(y^2)) = O(O(x^2)) + O(O(y^2)) = O(x^2) + O(y^2)$.

9.12 $1 + 2/n + O(n^{-2}) = (1 + 2/n)(1 + O(n^{-2})) / (1 + 2/n)$ (依据式 (9.26)), 且 $1/(1 + 2/n) = O(1)$. 现在用式 (9.26).

9.13 $n^n (1 + 2n^{-1} + O(n^{-2}))^n = n^n \exp(n(2n^{-1} + O(n^{-2}))) = e^2 n^n + O(n^{n-1})$.

9.14 它是 $n^{n+\beta} \exp((n+\beta)(\alpha/n - (1/2)\alpha^2/n^2 + O(n^{-3})))$.

9.15 $\ln \binom{3n}{n, n, n} = 3n \ln 3 - \ln n + (1/2) \ln 3 - \ln 2\pi + ((1/36) - (1/4))n^{-1} + O(n^{-3})$, 所以解答为

$$\frac{3^{2n+1/2}}{2\pi n} \left(1 - \frac{2}{9}n^{-1} + \frac{2}{81}n^{-2} + O(n^{-3}) \right).$$

9.16 如果 l 是范围 $a \leq l < b$ 中的任何整数, 我们有

$$\begin{aligned} \int_0^1 B(x)f(l+x)dx &= \int_{1/2}^1 B(x)f(l+x)dx - \int_0^{1/2} B(1-x)f(l+x)dx \\ &= \int_{1/2}^1 B(x)(f(l+x) - f(l+1-x))dx. \end{aligned}$$

由于当 $x \geq 1/2$ 时, $l+x \geq l+1-x$, 当 $f(x)$ 是非降时, 这个积分为正的.

$$9.17 \quad \sum_{m \geq 0} B_m \left(\frac{1}{2} \right) z^m / m! = ze^{z/2} / (e^z - 1) = z / (e^{z/2} - 1) - z / (e^z - 1).$$

9.18 推广情形 $\alpha = 1$ 的书的推导给出

$$b_k(n) = \frac{2^{(2n+1/2)\alpha}}{(2\pi n)^{\alpha/2}} e^{-k^2 \alpha / n}, \quad c_k(n) = 2^{2n\alpha} n^{-(1+\alpha)/2+3\alpha} e^{-k^2 \alpha / n};$$

解答是 $2^{2n\alpha} (\pi n)^{-(1+\alpha)/2} \alpha^{-1/2} (1 + O(n^{-1/2+3\alpha}))$.

9.19 $H_{10} = 2.928\,968\,254 \approx 2.928\,968\,256$; $10! = 3\,628\,800 \approx 3\,628\,712.4$; $B_{10} = 0.075\,757\,576 \approx 0.075\,757\,494$; $\pi(10) = 4 \approx 10.001\,784\,5$; $e^{0.1} = 1.105\,170\,92 \approx 1.105\,170\,83$; $\ln 1.1 = 0.095\,310\,2 \approx 0.095\,308\,3$; $1.111\,111\,1 \approx 1.111\,100\,0$; $1.1^{0.1} = 1.009\,576\,58 \approx 1.009\,576\,43$. (当 n 较大时, $\pi(n)$ 的近似给出更多有效数, 例如, $\pi(10^9) = 50\,847\,534 \approx 50\,840\,742$.)

9.20 (a) 是的, 左边为 $o(n)$ 而右边等价于 $O(n)$.

(b) 是的, 左边为 $e \cdot e^{O(1/n)}$.

(c) 不是, 左边近于右边界限的 \sqrt{n} 倍.

9.21 我们有 $P_n = m = n(\ln m - 1 - 1/\ln m + O(1/\log n)^2)$, 其中

$$\ln m = \ln n + \ln \ln m - 1/\ln n + \ln \ln n / (\ln n)^2 + O(1/\log n)^2;$$

$$\ln \ln m = \ln \ln n + \frac{\ln \ln n}{\ln n} - \frac{(\ln \ln n)^2}{2(\ln n)^2} + \frac{\ln \ln n}{(\ln n)^2} + O(1/\log n)^2.$$

由此得到

$$P_n = n \left(\ln n + \ln \ln n - 1 + \frac{\ln \ln n - 2}{\ln n} - \frac{\frac{1}{2}(\ln \ln n)^2 - 3 \ln \ln n}{(\ln n)^2} + O\left(\frac{1}{\log n}\right)^2 \right).$$

(稍好的近似用数量 $-5/(\ln n)^2 + O(\log \log n / \log n)^3$ 替换 $O(1/\log n)^2$, 于是我们估计 $P_{1\,000\,000} \approx 15\,483\,612.4$.)

9.22 在 H_n 的展开中用 $-(1/12)n^{-2k} + O(n^{-4k})$ 替换 $O(n^{-2k})$, 在式(9.53)中就用 $-(1/12)\sum_3(n^2) + O(\sum_3(n^4))$ 替换 $O(\sum_3(n^2))$. 我们有

$$\sum_3(n) = \frac{3}{4}n^{-1} + \frac{5}{36}n^{-2} + O(n^{-3}),$$

因此用 $-(19/144)n^{-2} + O(n^{-3})$ 能替换式(9.54)中的项 $O(n^{-2})$.

9.23 $nh_n = \sum_{0 \leq k < n} h_k / (n-k) + 2cH_n / (n+1)(n+2)$. 选取 $c = e^{\pi^2/6} = \sum_{k \geq 0} g_k$ 以

使 $\sum_{k \geq 0} h_k = 0$ 和 $h_n = O(\log n) / n^3$. 如同式(9.60)中那样的 $\sum_{0 \leq k < n} h_k / (n-k)$ 的展开现在产生 $nh_n = 2cH_n / (n+1)(n+2) + O(n^{-2})$, 因此

$$g_n = e^{\pi^2/6} \left(\frac{n + 2\ln n + O(1)}{n^3} \right).$$

9.24 (a) 如果 $\sum_{k \geq 0} |f(k)| < \infty$ 以及当 $0 \leq k \leq n/2$ 时, $f(n-k) = O(f(n))$, 我们有

$$\sum_{k=0}^n a_k b_{n-k} = \sum_{k=0}^{n/2} O(f(k))O(f(n)) + \sum_{k=n/2}^n O(f(n))O(f(n-k)),$$

它是 $2O(f(n) \sum_{k \geq 0} |f(k)|)$, 所以证明了此情形.

(b) 在此情形中如果 $a_n = b_n = x^{-n}$, 则卷积 $(n+1)x^{-n}$ 不是 $O(x^{-n})$.

9.25 $S_n / \binom{3n}{n} = \sum_{k=0}^n n^{\bar{k}} / (2n+1)^{\bar{k}}$. 譬如说, 我们可限制求和范围为 $0 \leq k \leq (\log n)^2$. 在此范围中 $n^{\bar{k}} = n^k \left(1 - \binom{k}{2} / n + O(k^4 / n^2) \right)$ 和 $(2n+1)^{\bar{k}} = (2n)^k (1 + \binom{k+1}{2} / 2n + O(k^4 / n^2))$, 所以被加数是

$$\frac{1}{2^k} \left(1 - \frac{3k^2 - k}{4n} + O\left(\frac{k^4}{n^2}\right) \right).$$

因此 k 上求和为 $2 - 4/n + O(1/n^2)$. 现在 Stirling 近似能应用到 $\binom{3n}{n} = (3n)! / (2n)!n!$, 证明了式(9.2).

9.26 最小出现在项 $B_{2m} / (2m)(2m-1)n^{2m-1}$ 处, 其中 $2m \approx 2\pi n + 3/2$, 而此项近似等于 $1 / (\pi e^{2\pi n} \sqrt{n})$, 所以 $\ln n!$ 的绝对误差太大不能用舍入的一个整数精确地确定 $n!$,

当 n 差不多大于 $e^{2\alpha+1}$ 时.

9.27 我们可设 $\alpha \neq -1$. 设 $f(x) = x^\alpha$, 解答为

$$\sum_{k=1}^n k^\alpha = C_\alpha + \frac{n^{\alpha+1}}{\alpha+1} + \frac{n^\alpha}{2} + \sum_{k=1}^n \frac{B_{2k}}{2k} \binom{\alpha}{2k-1} n^{\alpha-2k+1} + O(n^{\alpha-2m-1}).$$

(常数 C_α 结果是 $\zeta(-\alpha)$, 当 $\alpha > -1$ 时, 事实上由此公式定义.)

9.28 在 Euler 求和公式中取 $f(x) = x \ln x$ 得到

$$A \cdot n^{n^2/2+n/2+1/12} e^{-n^2/4} (1 + O(n^{-2})).$$

其中 $A \approx 1.282427$ 是“Glaisher 常数”.

9.29 设 $f(x) = x^{-1} \ln x$. 于是对所有大的 x , $f^{(2m)}(x) > 0$, 我们能记

$$\sum_{k=1}^n \frac{\ln k}{k} = \frac{(\ln n)^2}{2} + \ln S + \frac{\ln n}{2n} + \theta_n \frac{1 - \ln n}{12n^2}, \quad 0 < \theta_n < 1,$$

其中 $S \approx 0.929772$ 是常数. 取指数得到

$$S \sqrt{n^{\ln n}} \left(1 + \frac{\ln n}{2n} + O\left(\frac{\log n}{n}\right)^2 \right).$$

(一般, 如果 $f(x) = x^\alpha \ln x$, 如同习题 27 那样应用 Euler 求和公式, 产生的常数为 $-\zeta'(-\alpha)$ (如果 $\alpha \neq -1$). 于是, ζ 函数理论给出了前面习题中的 Glaisher 常数的闭形式. 在解答 9.57 的记法中, 我们得到 $\ln S = \gamma_1$.)

9.30 设 $g(x) = x^l e^{-x^2}$ 和 $f(x) = g(x/\sqrt{n})$. 则 $n^{-l/2} \sum_{k \geq 0} k^l e^{-k^2/n}$ 是

$$\begin{aligned} \int_0^\infty f(x) dx &= \sum_{k=1}^m \frac{B_k}{k!} f^{(k-1)}(0) - (-1)^m \int_0^\infty \frac{B_m(\{x\})}{m!} f^{(m)}(x) dx \\ &= n^{l/2} \int_0^\infty g(x) dx - \sum_{k=1}^m \frac{B_k}{k!} n^{(k-1)/2} g^{(k-1)}(0) + O(n^{-m/2}). \end{aligned}$$

由于 $g(x) = x^l - x^{2+l}/1! + x^{4+l}/2! - x^{6+l}/3! + \dots$, 导数 $g^{(m)}(x)$ 服从一个简单型式, 且解答是

$$\frac{1}{2} n^{(l+1)/2} \Gamma\left(\frac{l+1}{2}\right) - \frac{B_{l+1}}{(l+1)! 0!} + \frac{B_{l+3} n^{-1}}{(l+3)! 1!} - \frac{B_{l+5} n^{-2}}{(l+5)! 2!} + O(n^{-3}).$$

9.31 有点奇怪的等式 $1/(c^{m-k} + c^m) + 1/(c^{m+k} + c^m) = 1/c^m$ 使得 $0 \leq k \leq 2m$ 的项共计为 $(m+1/2)/c^m$. 剩下的项为

$$\sum_{k \geq 0} \frac{1}{c^{2m+k} + c^m} = \sum_{k \geq 1} \left(\frac{1}{c^{2m+k}} - \frac{1}{c^{3m+2k}} + \frac{1}{c^{4m+3k}} - \dots \right)$$

$$= \frac{1}{c^{2m+1} - c^{2m}} - \frac{1}{c^{3m+2} - c^{3m}} + \dots,$$

在任何希望处截此级数，具有的误差不超过第一个消去的项。

9.32 由于知道常数，所以我们应用 Euler 求和公式 $H_n^{(2)} = \pi^2/6 - 1/n + O(n^{-2})$ ，由式(9.89)给出 H_n ，所以解答为

$$ne^{1+x^2/6} \left(1 - \frac{1}{2}n^{-1} + O(n^{-2}) \right).$$

9.33 我们有 $n^k/n^k = 1 - k(k-1)n^{-1} + (1/2)k^2(k-1)^2n^{-2} + O(k^6n^{-3})$ ，用 $k!$ 除且对 $k \geq 0$ 上求和产生 $e - en^{-1} + (7/2)en^{-2} + O(n^{-3})$ 。

9.34 $A = e^7$; $B = 0$; $C = -(1/2)e^7$; $D = (1/2)e^7(1-\gamma)$; $E = (1/8)e^7$; $F = (1/12)e^7(3\gamma+1)$ 。

9.35 由于 $1/(k \ln k + O(1)) = 1/k \ln k + O(1/k(\log k)^2)$ ，给定和是 $\sum_{k=2}^n 1/k \ln k + O(1)$ 。由 Euler 求和公式，剩下和是 $\ln \ln n + O(1)$ 。

9.36 与 Euler 求和公式一起，很好地算出：

$$\begin{aligned} S_n &= \sum_{0 \leq k \leq n} \frac{1}{n^2 + k^2} + \frac{1}{n^2 + x^2} \Big|_0^n \\ &= \int_0^n \frac{dx}{n^2 + x^2} + \frac{1}{2} \frac{1}{n^2 + x^2} \Big|_0^n + \frac{B_2}{2!} \frac{-2x}{(n^2 + x^2)^2} \Big|_0^n + O(n^{-5}). \end{aligned}$$

因此 $S_n = (1/4)\pi n^{-1} - (1/4)n^{-2} - (1/24)n^{-3} + O(n^{-5})$ 。

9.37 这是

$$\begin{aligned} &\sum_{q \geq 1} (n - qk)[n/(q+1) < k \leq n/q] \\ &= n^2 - \sum_{q \geq 1} q \left(\binom{\lfloor n/q \rfloor + 1}{2} - \binom{\lfloor n/(q+1) \rfloor + 1}{2} \right) \\ &= n^2 - \sum_{q \geq 1} \left(\binom{\lfloor n/q \rfloor + 1}{2} \right). \end{aligned}$$

剩下的和像式(9.55)，但是没有因子 $\mu(q)$ 。在这里用与那里相同的方法，但是我们取 $\zeta(2)$ 代替 $1/\zeta(2)$ ，所以解答达到 $\left(1 - \frac{\pi^2}{12}\right)n^2 + O(n \log n)$ 。

9.38 用 $n-k$ 替换 k ，且设 $a_k(n) = (n-k)^{n-k} \binom{n}{k}$ 。于是 $\ln a_k(n) = n \ln n - \ln k! - k + O(kn^{-1})$ ，我们可用 $b_k(n) = n^n e^{-k}/k!$ ， $c_k(n) = kb_k(n)/n$ ， $D_n = \{k | k \leq \ln n\}$ 作尾

部交换, 取得 $\sum_{k=0}^n a_k(n) = n^n e^{1/e} (1 + O(n^{-1}))$.

9.39 用 $b_k(n) = (\ln n - k/n - (1/2)k^2/n^2)(\ln n)^k/k!$, $c_k(n) = n^{-3}(\ln n)^{k+3}/k!$, $D_n = \{k | 0 \leq k \leq 10 \ln n\}$ 尾部交换. 当 $k \approx 10 \ln n$ 时, 我们有 $k! \asymp \sqrt{k}(10/e)^k(\ln n)^k$, 所以第 k 项是 $O(n^{-10 \ln(10/e)} \log n)$. 解答是 $n \ln n - \ln n - (1/2)(\ln n)(1 + \ln n)/n + O(n^{-2}(\log n)^3)$.

9.40 两项两项合并, 我们发现 $H_{2k}^m - (H_{2k} - 1/2k)^m = (m/2k)H_{2k}^{m-1}$ 加上在所有 $k \geq 1$ 上的和的项为 $O(1)$. 假设 n 是偶数. Euler 求和公式意味着

$$\sum_{k=1}^{n/2} \frac{H_{2k}^{m-1}}{k} = \sum_{k=1}^{n/2} \frac{(\ln 2e^{\gamma}k)^{m-1} + O(1/k)}{k} + O(1) = \frac{(\ln e^{\gamma}n)^m}{m} + O(1),$$

因此和为 $(1/2)H_n^m + O(1)$. 一般解答为 $(1/2)(-1)^n H_n^m + O(1)$.

9.41 设 $\alpha = \hat{\varphi}/\varphi = -\varphi^{-2}$. 我们有

$$\begin{aligned} \sum_{k=1}^n \ln F_k &= \sum_{k=1}^n (\ln \varphi^k - \ln \sqrt{5} + \ln(1 - \alpha^k)) \\ &= \frac{n(n+1)}{2} \ln \varphi - \frac{n}{2} \ln 5 + \sum_{k=1}^n \ln(1 - \alpha^k) - \sum_{k>n} \ln(1 - \alpha^k). \end{aligned}$$

后面的和为 $\sum_{k>n} O(\alpha^k) = O(\alpha^n)$. 因此解答为

$$\varphi^{n(n+1)/2} 5^{-n/2} C + O(\varphi^{n(n-3)/2} 5^{-n/2}), \text{ 其中}$$

$$C = (1 - \alpha)(1 - \alpha^2)(1 - \alpha^3) \cdots \approx 1.226742.$$

9.42 由于 $\binom{n}{k-1} / \binom{n}{k} = \frac{k}{n-k-1} \leq \frac{\alpha n}{n - \alpha n + 1} < \frac{\alpha}{1 - \alpha}$, 提示随之而来, 设 $m = \lfloor \alpha n \rfloor = \alpha n - \varepsilon$. 则

$$\begin{aligned} \binom{n}{m} &< \sum_{k \leq m} \binom{n}{k} \\ &< \binom{n}{m} \left(1 + \frac{\alpha}{1 - \alpha} + \left(\frac{\alpha}{1 - \alpha} \right)^2 + \cdots \right) = \binom{n}{m} \frac{1 - \alpha}{1 - 2\alpha}. \end{aligned}$$

所以 $\sum_{k \leq m} \binom{n}{k} = \binom{n}{m} O(1)$, 剩下估计 $\binom{n}{m}$. 依据 Stirling 近似, 我们有 $\ln \binom{n}{m} = -(1/2) \ln n - (\alpha n - \varepsilon) \ln(\alpha - \varepsilon/n) - ((1 - \alpha)n + \varepsilon) \ln(1 - \alpha + \varepsilon/n) + O(1) = -(1/2) \ln n - \alpha n \ln \alpha - (1 - \alpha)n \ln(1 - \alpha) + O(1)$.

9.43 分母有形式 $z - \omega$ 的因子, 其中 ω 是单位的一个复数根. 仅仅因子 $z - 1$ 具有重数 5 出现. 所以依据式(7.31), 仅仅一个根有一个系数 $\Omega(n^4)$, 且系数为 $c = 5/(5! \cdot 1$

$$\cdot 5 \cdot 10 \cdot 25 \cdot 50) = 1 / 1\,500\,000.$$

9.44 Stirling 近似说到 $\ln(x^{-\alpha}x!/(x-\alpha)!)$ 有渐近级数

$$\begin{aligned} & -\alpha - \left(x + \frac{1}{2} - \alpha\right) \ln(1 - \alpha/x) - \frac{B_2}{2 \cdot 1} (x^{-1} - (x-\alpha)^{-1}) \\ & - \frac{B_4}{4 \cdot 3} (x^{-3} - (x-\alpha)^{-3}) - \dots, \end{aligned}$$

其中的 x^{-k} 的每个系数是 α 的一个多项式。因此当 $x \rightarrow \infty$ 时, $x^{-\alpha}x!/(x-\alpha)! = c_0(\alpha) + c_1(\alpha)x^{-1} + \dots + c_n(\alpha)x^{-n} + O(x^{-n-1})$, 其中 $c_n(\alpha)$ 是 α 的一个多项式。我们知道每当 α 是整数时, $c_n(\alpha) = \left[\begin{smallmatrix} \alpha \\ \alpha-n \end{smallmatrix} \right] (-1)^n$, 且 $\left[\begin{smallmatrix} \alpha \\ \alpha-n \end{smallmatrix} \right]$ 是次数 $2n$ 的 α 的一个多项式; 因此对所有实 α , $c_n(\alpha) = \left[\begin{smallmatrix} \alpha \\ \alpha-n \end{smallmatrix} \right] (-1)^n$ 。换句话说, 渐近公式

$$\begin{aligned} x^{\alpha} &= \sum_{k=0}^n \left[\begin{smallmatrix} \alpha \\ \alpha-k \end{smallmatrix} \right] (-1)^k x^{\alpha-k} + O(x^{\alpha-n-1}), \\ x^{\bar{\alpha}} &= \sum_{k=0}^n \left[\begin{smallmatrix} \alpha \\ \alpha-k \end{smallmatrix} \right] x^{\alpha-k} + O(x^{\alpha-n-1}) \end{aligned}$$

推广了方程(6.13)和(6.11), 其在所有整数情形中成立。

9.45 设 α 的部分商为 $\langle a_1, a_2, \dots \rangle$, 且设 α_m 是连分数 $1/(a_m + \alpha_{m+1})$, ($m \geq 1$)。于是 $D(\alpha, n) = D(\alpha_1, n) < D(\alpha_2, \lfloor \alpha_1 n \rfloor) + a_1 + 3 < D(\alpha_3, \lfloor \alpha_2 \lfloor \alpha_1 n \rfloor \rfloor) + a_1 + a_2 + 6 < \dots < D(\alpha_{m+1}, \lfloor \alpha_m \lfloor \dots \lfloor \alpha_1 n \rfloor \dots \rfloor \rfloor) + a_1 + \dots + a_m + 3m < \alpha_1 \dots \alpha_m n + a_1 + \dots + a_m + 3m$, (对所有 m)。用 n 除, 且设 $n \rightarrow \infty$; 对于所有 m , 极限小于 $\alpha_1 \dots \alpha_m$ 。我们最后有

$$\alpha_1 \dots \alpha_m = \frac{1}{K(a_1, \dots, a_{m-1}, a_m + \alpha_m)} < \frac{1}{F_{m+1}}.$$

9.46 为了方便起见, 我们就写 m , 而不是 $m(n)$ 。依据 Stirling 近似, 当 $k \approx m \approx n/\ln n$ 时, $k^n/k!$ 的最大值出现, 所以我们用 $m+k$ 替换 k , 发现

$$\begin{aligned} \ln \frac{(m+k)^n}{(m+k)!} &= n \ln m - m \ln m + m - \frac{\ln 2\pi m}{2} \\ &\quad - \frac{(m+n)k^2}{2m^2} + O(k^3 m^{-2} \log n) \end{aligned}$$

实际上我们要用 $\lfloor m \rfloor + k$ 替换 k , 这加了一个另外的 $O(km^{-1} \log n)$ 。现在 $|k| \leq m^{1/2+\epsilon}$ 时尾部交换方法允许我们在 k 上求和, 给出了相当准确的渐近估计

$$b_n = \frac{e^{m-1} m^{n-m}}{\sqrt{2\pi m}} (\Theta_{2m^2/(m+n)} + O(1))$$

$$= e^{m-n-1/2} m^n \sqrt{\frac{m}{m+n}} \left(1 + O\left(\frac{\log n}{n^{1/2}}\right) \right).$$

得到要求的公式, 具有相对误差 $O(\log \log n / \log n)$.

9.47 设 $\log_m n = l + \theta$, 其中 $0 \leq \theta < 1$. 下整和为 $l(n+1) + 1 - (m^{l+1} - 1) / (m - 1)$; 上整和为 $(l+1)n - (m^{l+1} - 1) / (m - 1)$; 恰切和为 $(l+\theta)n - n / \ln m + O(\log n)$. 忽略 $o(n)$ 的项, 上整和恰切之间的差为 $(1-f(\theta))n$, 恰切和下整之间的差为 $f(\theta)n$, 其中

$$f(\theta) = \frac{m^{1-\theta}}{m-1} + \theta - \frac{1}{\ln m}.$$

此函数有最大值 $f(0) = f(1) = m / (m - 1) - 1 / \ln m$, 且它的最小值是 $\ln \ln m / \ln m + 1 - (\ln(m-1)) / \ln m$. 当 n 接近 m 的一个幂时, 上整值较接近, 但是当 θ 在 0 和 1 之间某处时, 下整值较接近.

9.48 设 $d_k = a_k + b_k$, 其中 a_k 计数十进小数点左边的数字位. 则 $a_k = 1 + \lfloor \log H_k \rfloor = \log \log k + O(1)$, 其中 'log' 表示 \log_{10} . 为了估计 b_k , 为了区别 y 和附近的数 $y - \varepsilon$ 和 $y + \varepsilon'$, 我们查看十进位的个数: 设 $\delta = 10^{-b}$ 是四舍五入到 \hat{y} 的数的区间的长度. 我们有 $|y - \hat{y}| \leq (1/2)\delta$, 还有 $y - \varepsilon < \hat{y} - (1/2)\delta$ 和 $y + \varepsilon' > \hat{y} + (1/2)\delta$, 所以 $\varepsilon + \varepsilon' > \delta$. 且如果 $\delta < \min(\varepsilon, \varepsilon')$, 四舍五入区别了 \hat{y} 和 $y - \varepsilon$ 以及 $y + \varepsilon'$. 因此 $10^{-b_k} < 1 / (k-1) + 1/k$ 和 $10^{-b_k} \geq 1/k$, 我们有 $b_k = \log k + O(1)$. 所以最终 $\sum_{k=1}^n d_k = \sum_{k=1}^n (\log k + \log \log k + O(1))$, 依据 Euler 求和公式, 它是 $n \log n + n \log \log n + O(n)$.

9.49 我们有 $H_n > \ln n + \gamma + (1/2)n^{-1} - (1/12)n^{-2} = f(n)$, 对于所有 $x > 0$, 其中 $f(x)$ 是上升的. 因此如果 $n \geq e^{x-\gamma}$, 我们有 $H_n \geq f(e^{x-\gamma}) > \alpha$. 还有 $H_{n-1} < \ln n + \gamma - (1/2)n^{-1} = g(n)$, 对于所有 $x > 0$, 其中 $g(x)$ 是上升的. 因此如果 $n \leq e^{x-\gamma}$, 我们有 $H_{n-1} \leq g(e^{x-\gamma}) < \alpha$. 所以 $H_{n-1} \leq \alpha \leq H_n$ 意味着 $e^{x-\gamma} + 1 > n > e^{x+\gamma} - 1$. (Boas 和 Wrench^[27]已得到更精确的结果.)

9.50 (a) 期望的返回是 $\sum_{1 \leq k \leq N} k / (k^2 H_N^{(2)}) = H_N / H_N^{(2)}$, 我们需要的关于 $O(N^{-1})$ 的渐近值

$$\frac{\ln N + \gamma + O(N^{-1})}{\pi^2/6 - N^{-1} + O(N^{-2})} = \frac{6 \ln 10}{\pi^2} n + \frac{6\gamma}{\pi^2} + \frac{36 \ln 10}{\pi^4} \frac{n}{10^n} + O(10^{-n}).$$

系数 $(6 \ln 10) / \pi^2 \approx 1.3998$ 表明我们期望 40% 左右的得益.

(b) 得益的概率是 $\sum_{n \leq k \leq N} 1 / (k^2 H_N^{(2)}) = 1 - H_n^{(2)} / H_N^{(2)}$,

且由于 $H_n^{(2)} = \pi^2/6 - n^{-1} + (1/2)n^{-2} + O(n^{-3})$, 这是

$$\frac{n^{-1} - \frac{1}{2}n^{-2} + O(n^{-3})}{\pi^2/6 + O(N^{-1})} = \frac{6}{\pi^2}n^{-1} - \frac{3}{\pi^2}n^{-2} + O(n^{-3}).$$

实际关于 n 下降。(a)中的期望值是很高的,因为它包含的盈利如此巨大,如果在任何时候必需作出它们,这将影响整个世界经济。)

9.51 严格地说,这是不真的,因为由 $O(x^{-2})$ 表示的函数可能不可积。(它可能是 $[xS]/x^2$, 其中 S 是一个不可测集合。)但是如果约定 $f(x)$ 是可积函数,使得当 $x \rightarrow \infty$ 时 $f(x) = O(x^{-2})$, 则 $\left| \int_n^\infty f(x) dx \right| \leq \int_n^\infty |f(x)| dx \leq \int_n^\infty Cx^{-2} dx = Cn^{-1}$ 。

9.52 事实上,是用任何趋于无穷(不管怎样快)的函数替换一组 n , 通过置 $m_0 = 0$ 定义序列 $\langle m_0, m_1, m_2, \dots \rangle$, 且设 m_k 是 $> m_{k-1}$ 的最小整数,使得

$$\left(\frac{k+1}{k} \right)^{m_k} \geq f(k+1)^2.$$

现在设 $A(z) = \sum_{k \geq 1} (z/k)^{m_k}$, 对所有 z , 这个幂级数收敛, 因为一个几何级数界限了 $k > |z|$

的项。还有 $A(n+1) \geq ((n+1)/n)^{m_n} \geq f(n+1)^2$, 因此 $\lim_{n \rightarrow \infty} f(n)/A(n) = 0$ 。

9.53 用归纳法, O 项是 $(m-1)!^{-1} \int_0^x t^{m-1} f^{(m)}(x-t) dt$ 。由于 $f^{(m+1)}$ 有 $f^{(m)}$ 的相反正负号, 所以 $|f^{(m)}(0)| \int_0^x t^{m-1} dt$ 界限此积分的绝对值, 丢弃的第一项的绝对值界限误差。

9.54 设 $g(x) = f(x)/x^a$, 则 $g'(x) \sim -\alpha g(x)/x$ (当 $x \rightarrow \infty$), 依据均值定理, 对于 $x-1/2$ 和 $x+1/2$ 之间的某个 y , $g(x-1/2) - g(x+1/2) = -g'(y) \sim \alpha g(y)/y$ 。现在 $g(y) = g(x)(1 + O(1/x))$, 所以 $g(x-1/2) - g(x+1/2) \sim \alpha g(x)/x = \alpha f(x)/x^{1+a}$ 。所以

$$\sum_{k \geq n} \frac{f(k)}{k^{1+a}} = O\left(\sum_{k \geq n} \left(g\left(k - \frac{1}{2}\right) - g\left(k + \frac{1}{2}\right)\right) = O\left(g\left(n - \frac{1}{2}\right)\right)\right).$$

9.55 $(n+k+1/2)\ln(1+k/n) + (n-k+1/2)\ln(1-k/n)$ 的估计被展开到 $k^2/n + k^4/6n^3 + O(n^{-3/2+5\epsilon})$, 所以在 $b_k(n)$ 中显然要有一个额外因子 $e^{-k^4/6n^3}$, 且 $c_k(n) = 2^{2n} n^{-2+5\epsilon} e^{-k^2/n}$ 。但是结果是把 $b_k(n)$ 留下不动较好, 且设

$$c_k(n) = 2^{2n} n^{-2+5\epsilon} e^{-k^2/n} + 2^{2n} n^{-5+5\epsilon} k^4 e^{-k^2/n},$$

从而用 $1 + O(k^4/n^3)$ 替换 $e^{-k^4/6n^3}$ 。如同习题 30 表明的那样, 和 $\sum_k k^4 e^{-k^2/n}$ 是

$O(n^{3/2})$.

9.56 如果 $k \leq n^{1/2+\varepsilon}$, 依据 Stirling 近似我们有 $\ln(n^k/n^k) = -(1/2)k^2/n + (1/2)k/n - (1/6)k^3/n^2 + O(n^{-1+4\varepsilon})$, 因此

$$n^k/n^k = e^{-k^2/2n}(1 + k/2n - \frac{2}{3}k^3/(2n)^2 + O(n^{-1+4\varepsilon})).$$

用习题 30 中的等式求和, 且记住删去 $k=0$ 的项, 得到 $-1 + \Theta_{2n} + \Theta_{2n}^{(1)} - \frac{2}{3}\Theta_{2n}^{(3)} + O(n^{-1/2+4\varepsilon}) = \sqrt{\pi n/2} - \frac{1}{3} + O(n^{-1/2+4\varepsilon})$.

9.57 用提示, 给定的和变成 $\int_0^\infty u e^{-u} \zeta(1+u/\ln n) du$. 用级数

$$\zeta(1+z) = z^{-1} + \sum_{m \geq 0} (-1)^m \gamma_m z^m / m!$$

来定义 ζ 函数, 其中 $\gamma_0 = \gamma$, γ_m 是 Stieltjes 常数

$$\lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{(\ln k)^m}{k} - \frac{(\ln n)^{m+1}}{m+1} \right).$$

因此给定的和是

$$\ln n + \gamma - 2\gamma_1 (\ln n)^{-1} + 3\gamma_2 (\ln n)^{-2} - \dots.$$

9.58 设 $0 \leq \theta \leq 1$ 和 $f(z) = e^{2\pi i z \theta} / (e^{2\pi i z} - 1)$. 我们有

$$|f(z)| = \frac{e^{-2\pi y \theta}}{1 + e^{-2\pi y}} \leq 1, \text{ 当 } x \bmod 1 = \frac{1}{2};$$

$$|f(z)| \leq \frac{e^{-2\pi y \theta}}{|e^{-2\pi y} - 1|} \leq \frac{1}{1 - e^{-2\pi y}}, \text{ 当 } |y| \geq \varepsilon.$$

所以在围道上 $|f(z)|$ 是有界的, 且积分为 $O(M^{1-m})$. 在 $z = k \neq 0$ 处, $2\pi i f(z)/z^m$ 的残数是 $e^{2\pi i k \theta} / k^m$; 在 $z = 0$ 处残数是

$$\frac{e^{2\pi i \theta}}{z^{m+1}} \left(B_0 + B_1 \frac{2\pi i z}{1!} + \dots \right) = \frac{1}{z^{m+1}} \left(B_0(\theta) + B_1(\theta) \frac{2\pi i z}{1!} + \dots \right) \text{ 中 } z^{-1} \text{ 的系数,}$$

即 $(2\pi i)^m B_m(\theta) / m!$. 所以在围道内残数的和是

$$\frac{(2\pi i)^m}{m!} B_m(\theta) + 2 \sum_{k=1}^M e^{\pi i m / 2 \cos(2\pi k \theta - \pi m / 2)} k^m.$$

这等于围道积分 $O(M^{1-m})$, 所以当 $M \rightarrow \infty$ 时, 它趋于零.

9.59 如果 $F(x)$ 有充分好的性质, 我们有一般等式

$$\sum_k F(k+t) = \sum_n G(2\pi n) e^{2\pi i n t},$$

其中 $G(y) = \int_{-\infty}^{\infty} e^{-iyx} F(x) dx$. (这是“Poisson 求和公式”, 在 Henrici^[151], 定理 10.6c) 所写的书中可以找到.

9.60 依据习题 5.22, 所述公式等价于

$$n^{1/2} = n^{1/2} \left(1 - \frac{1}{8n} + \frac{1}{128n^2} + \frac{5}{1024n^3} - \frac{21}{32768n^4} + O(n^{-5}) \right).$$

因此从习题 6.64 和 9.44 得到结果.

9.61 思想是使 α^* “几乎”有理, 设 $a_k = 2^{2^k}$ 是 α 的第 k 个部分商, 且设 $n = (1/2)a_{m+1}q_m$, 其中 $q_m = K(a_1, \dots, a_m)$ 以及 m 是偶数. 于是 $0 < \{q_m \alpha\} < 1/2$, 而且如果我们取 $v = a_{m+1}/(4n)$, 得到一个偏差 $\geq (1/4)a_{m+1}$. 如果这小于 $n^{1-\epsilon}$, 我们将有

$$a_{m+1}^* = O(q_m^{1-\epsilon}),$$

但事实上 $a_{m+1} > q_m^{2^m}$.

9.62 两种类型的 Stirling 数的渐近可见 Canfield 的[43], 还可见 David 和 Barton 的[60, 第 16 章].

9.63 设 $c = \varphi^{2-\varphi}$. Fine^[120] 证明了估计 $cn^{\varphi-1} + o(n^{\varphi-1})$. Ilan Vardi 注意到误差项 $e(n) = f(n) - cn^{\varphi-1}$ 满足近似递归 $c^{\varphi} n^{2-\varphi} e(n) \approx -\sum_k e(k) [1 \leq k < cn^{\varphi-1}]$ 从而推出所述的较精确的估计. 如果 $u(x+1) = -u(x)$, 函数

$$\frac{n^{\varphi-1} u(\ln \ln n / \ln \varphi)}{\ln n}$$

渐近地满足这个递归. (对于某个这样的函数 u , Vardi 猜测

$$f(n) = n^{\varphi-1} \left(c + u \left(\frac{\ln \ln n}{\ln \varphi} \right) (\ln n)^{-1} + O((\log n)^{-2}) \right).$$

除了一种情形 $f(273) = 39 > c \cdot 273^{\varphi-1} \approx 38.4997$ 外, 对于小的 n 的计算表明, 对 $1 \leq n \leq 400$, $f(n)$ 等于 $cn^{\varphi-1}$ 的最近整数. 但是由于像习题 2.36 中那些结果, 小的误差可能被扩大了, 例如, $e(201\,636\,503) \approx 35.73$; $e(919\,986\,484\,788) \approx -1\,959.07$.

9.64 (通过对 m 归纳, 根据 $B_2(x)$ 的这个等式我们能容易推出习题 58 的等式。) 如果 $0 < x < 1$, 积分 $\int_x^{1/2} \sin N\pi t dt / \sin \pi t$ 能表为 N 个积分的和, 每个积分为 $O(N^{-2})$, 所以它为 $O(N^{-1})$; 此 O 包含的常数可依赖于 x . 积分等式 $\sum_{n=1}^N \cos 2n\pi t = \mathcal{R}(e^{2\pi i t}(e^{2N\pi i} - 1)/(e^{2\pi i} - 1)) = -(1/2) + (1/2)\sin(2N+1)\pi t / \sin \pi t$, 且设 $N \rightarrow \infty$, 现得到 $\sum_{n=1}^{\infty} (\sin 2n\pi x) / n = (\pi/2) - \pi x$, 这是 Euler 知道的一个关系^{[85][88, 部分 2, § 92]}. 再积分产生希望的公式. (这个解是由 E. M. E. Wermuth 提供的, Euler 的原先推导不符合近代标准的精确.)

9.65 当 f 是 $\{1, 2, \dots, n\}$ 到自身的一个随机映射时, 序列 $1, f(1), f(f(1)), \dots$ 中不同元素的期望个数是习题 56 的函数 $Q(n)$, 它的值是 $(1/2)\sqrt{2\pi n} + O(1)$; 这可设法估算因子 $\sqrt{2\pi n}$.

9.66 由此知道 $\ln \chi_n \sim (3/2)n^2 \ln(4/3)$, 常数 $e^{-x/6}$ 被检验到 8 位有意义的数字位.

9.67 例如, $e^{n^{-7}} = m + 1/2 + \varepsilon/n$ (对某个整数 m 和某个 $0 < \varepsilon < 1/8$), 这将不真; 但不知道反例.

附录 B 参考文献

HERE ARE THE WORKS cited in this book. Numbers in the margin specify the page numbers where citations occur.

References to published problems are generally made to the places where solutions can be found, instead of to the original problem statements, unless no solution has yet appeared in print.

- 1 N. H. Abel, letter to B. Holmboe (1823), in his *Œuvres Complètes*, first edition, 1839, volume 2, 264–265. Reprinted in the second edition, 1881, volume 2, 254–255.
- 2 Milton Abramowitz and Irene A. Stegun, editors, *Handbook of Mathematical Functions*. United States Government Printing Office, 1964. Reprinted by Dover, 1965.
- 3 William W. Adams and J. L. Davison, "A remarkable class of continued fractions," *Proceedings of the American Mathematical Society* 65 (1977), 194–198.
- 4 A. V. Aho and N. J. A. Sloane, "Some doubly exponential sequences," *Fibonacci Quarterly* 11 (1973), 429–437.
- 5 W. Ahrens, *Mathematische Unterhaltungen und Spiele*. Teubner, Leipzig, 1901. Second edition, in two volumes, 1910 and 1913.
- 6 Naum Il'ich Akhiezer, *Klassicheskaya Problema Momentov i Nekotoryye Voprosy Analiza, Sviyazannyye s Neyu*. Moscow, 1961. English translation, *The classical Moment Problem and Some Related Questions in Analysis*, Hafner, 1965.
- 7 R. E. Allardice and A. Y. Fraser, "La Tour d'Hanoi," *Proceedings of the Edinburgh Mathematical Society* 2 (1884), 50–53.
- 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.
- 10 George E. Andrews, "On sorting two ordered sets," *Discrete Mathematics* 11 (1975), 97-106.
- 11 George E. Andrews, *The Theory of Partitions*. Addison-Wesley, 1976.
- 12 George E. Andrews and K. Uchimura, "Identities in combinatorics IV: Differentiation and harmonic numbers," *Utilitas Mathematica* 28 (1985), 265-269.
- 13 M. D. Atkinson, "The cyclic towers of Hanoi," *Information Processing Letters* 13 (1981), 118-119.
- 14 Paul Bachmann, *Die analytische Zahlentheorie*. Teubner, Leipzig, 1894.
- 15 W. N. Bailey, *Generalized Hypergeometric Series*. Cambridge University Press, 1935; second edition, 1964.
- 16 W. W. Rouse Ball and H. S. M. Coxeter, *Mathematical Recreations and Essays*, twelfth edition. University of Toronto Press, 1974. (A revision of Ball's *Mathematical Recreations and Problems*, first published by Macmillan, 1892.)
- 17 P. Barlow, "Demonstration of a curious numerical proposition," *Journal of Natural Philosophy, Chemistry, and the Arts* 27 (1810), 193-205.
- 18 Samuel Beatty, "Problem 3177," *American Mathematical Monthly* 34 (1927), 159-160.
- 19 E. T. Bell, "Euler algebra," *Transactions of the American Mathematical Society* 25 (1923), 135-154.
- 20 E. T. Bell, "Exponential numbers," *American Mathematical Monthly* 41 (1934), 411-419.
- 21 Edward A. Bender, "Asymptotic methods in enumeration," *SIAM Review* 16 (1974), 485-515.
- 22 Jacobi Bernoulli, *Ars Conjectandi*, opus posthumum. Basel, 1713. Reprinted in *Die Werke von Jakob Bernoulli*, volume 3, 107-286.
- 23 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.
- 24 William H. Beyer, editor, *CRC Standard Mathematical Tables*, 25th edition. CRC Press, Boca Raton, Florida, 1978.

- 24' 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.
- 25 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.
- 26 Gunnar Blom, "Problem E 3043: Random walk until no shoes," *American Mathematical Monthly* 94 (1987), 78–79.
- 27 R. P. Boas, Jr. and J. W. Wrench, Jr., "Partial sums of the harmonic series," *American Mathematical Monthly* 78 (1971), 864–870.
- 28 P. Bohl, "Über ein in der Theorie der säkularen Störungen vorkommendes Problem," *Journal für die reine und angewandte Mathematik* 135 (1909), 189–283.
- 29 P. du Bois-Reymond, "Sur la grandeur relative des infinis des fonctions," *Annali di Matematica pura ed applicata*, series 2, 4 (1871), 338–353.
- 30 Émile Borel, *Leçons sur les séries à termes positifs*. Gauthier-Villars, 1902.
- 31 Jonathan M. Borwein and Peter B. Borwein, *Pi and the AGM*. Wiley, 1987.
- 32 Richard P. Brent, "The first occurrence of large gaps between successive primes," *Mathematics of Computation* 27 (1973), 959–963.
- 33 Richard P. Brent, "Computation of the regular continued fraction for Euler's constant," *Mathematics of Computation* 31 (1977), 771–777.
- 34 John Brillhart, "Some miscellaneous factorizations," *Mathematics of Computation* 17 (1963), 447–450.
- 35 Achille Brocot, "Calcul des rouages par approximation, nouvelle méthode," *Revue Chronométrique* 6 (1860), 186–194. (He also published a 97-page monograph with the same title in 1862.)
- 36 Maxey Brooke and C. R. Wall, "Problem B-14: A little surprise," *Fibonacci Quarterly* 1, 3 (1963), 80.
- 37 Brother U. Alfred [Brousseau], "A mathematician's progress," *Mathematics Teacher* 59 (1966), 722–727.
- 38 Morton Brown, "Problem 6439: A periodic sequence," *American Mathematical Monthly* 92 (1985), 218.

- 39 T. Brown, "Infinite multi-variable subpolynomial 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*. Cited in A. K. Austin, "Modern research in mathematics," *The Mathematical Gazette* 51 (1967), 149–150.
- 40 Thomas C. Brown, "Problem E2619: Squares in a recursive sequence," *American Mathematical Monthly* 85 (1978), 52–53.
- 41 William G. Brown, "Historical note on a recurrent combinatorial problem," *American Mathematical Monthly* 72 (1965), 973–977.
- 42 S. A. Burr, "On moduli for which the Fibonacci sequence contains a complete system of residues," *Fibonacci Quarterly* 9 (1971), 497–504.
- 43 E. Rodney Canfield, "On the location of the maximum Stirling number(s) of the second kind," *Studies in Applied Mathematics* 59 (1978), 83–93.
- 44 Lewis Carroll [pseudonym of C. L. Dodgson], *Through the Looking Glass and What Alice Found There*. Macmillan, 1871.
- 45 Jean-Dominique Cassini, "Une nouvelle progression de nombres," *Histoire de l'Académie Royale des Sciences*, Paris, volume 1, 201. (Cassini's work is summarized here as one of the mathematical results presented to the academy in 1680. This volume was published in 1733.)
- 46 E. Catalan, "Note sur une Équation aux différences finies," *Journal de Mathématiques pures et appliquées* 3 (1838), 508–516.
- 47 Augustin-Louis Cauchy, *Cours d'analyse de l'École Royale Polytechnique*. Imprimerie Royale, Paris, 1821. Reprinted in his *Œuvres Complètes*, series 2, volume 3.
- 48 Arnold Buffum Chace, *The Rhind Mathematical Papyrus*, volume 1. Mathematical Association of America, 1927. (Includes an excellent bibliography of Egyptian mathematics by R. C. Archibald.)
- 49 M. Chaimovich, G. Freiman, and J. Schönheim, "On exceptions to Szegedy's theorem," *Acta Arithmetica* 49 (1987), 107–112.
- 50 P. L. Tchebichef [Chebyshev], "Mémoire sur les nombres premiers," *Journal de Mathématiques pures et appliquées* 17 (1852), 366–390. Reprinted in his *Œuvres*, volume 1, 51–70.
- 50' P. L. Chebyshev, "O srednikh velichinakh," *Matematicheskii Sbornik* 2 (1867), 1–9. Reprinted in his *Polnoe Sobranie Sochinenii*, volume 2, 431–437. French translation, "Des valeurs moyennes," *Journal de Mathématiques pures et appliquées*, series 2, 12 (1867), 177–184; reprinted in his *Œuvres*, volume 1, 685–694.

- 51 Th. Clausen, "Ueber die Fälle, wenn die Reihe von der Form

$$y = 1 + \frac{\alpha}{1} \cdot \frac{\beta}{\gamma} x + \frac{\alpha \cdot \alpha + 1}{1 \cdot 2} \cdot \frac{\beta \cdot \beta + 1}{\gamma \cdot \gamma + 1} x^2 + \text{etc.}$$

ein Quadrat von der Form

$$z = 1 + \frac{\alpha'}{1} \cdot \frac{\beta'}{\gamma'} \cdot \frac{\delta'}{\epsilon'} x + \frac{\alpha' \cdot \alpha' + 1}{1 \cdot 2} \cdot \frac{\beta' \cdot \beta' + 1}{\gamma' \cdot \gamma' + 1} \cdot \frac{\delta' \cdot \delta' + 1}{\epsilon' \cdot \epsilon' + 1} x^2 + \text{etc. hat,}"$$

Journal für die reine und angewandte Mathematik 3 (1828), 89–91.

- 52 Th. Clausen, "Beitrag zur Theorie der Reihen," *Journal für die reine und angewandte Mathematik* 3 (1828), 92–95.
- 53 Th. Clausen, "Theorem," *Astronomische Nachrichten* 17 (1840), columns 351–352.
- 54 Stuart Dodgson Collingwood, *The Lewis Carroll Picture Book*. T. Fisher Unwin, 1899. Reprinted by Dover, 1961, with the new title *Diversions and Digressions of Lewis Carroll*.
- 55 J. H. Conway and R. L. Graham, "Problem E2567: A periodic recurrence," *American Mathematical Monthly* 84 (1977), 570–571.
- 56 Harald Cramér, "On the order of magnitude of the difference between consecutive prime numbers," *Acta Arithmetica* 2 (1937), 23–46.
- 57 A. L. Crelle, "Démonstration élémentaire du théorème de Wilson généralisé," *Journal für die reine und angewandte Mathematik* 20 (1840), 29–56.
- 58 D. W. Crowe, "The n-dimensional cube and the Tower of Hanoi," *American Mathematical Monthly* 63 (1956), 29–30.
- 59 D. R. Curtiss, "On Kellogg's Diophantine problem," *American Mathematical Monthly* 29 (1922), 380–387.
- 60 F. N. David and D. E. Barton, *Combinatorial Chance*. Hafner, 1962.
- 61 J. L. Davison, "A series and its associated continued fraction," *Proceedings of the American Mathematical Society* 63 (1977), 29–32.
- 62 N. G. de Bruijn, *Asymptotic Methods in Analysis*. North-Holland, 1958; third edition, 1970. Reprinted by Dover, 1981.
- 63 N. G. de Bruijn, "Problem 9," *Nieuw Archief voor Wiskunde*, series 3, 12 (1964), 68.
- 64 Abraham de Moivre, *Miscellanea analytica de seriebus et quadraturis*. London, 1730.

- 65 Leonard Eugene Dickson, *History of the Theory of Numbers*. Carnegie Institution of Washington, volume 1, 1919; volume 2, 1920; volume 3, 1923. Reprinted by Stechert, 1934; and by Chelsea, 1952, 1971.
- 66 Edsger W. Dijkstra, *Selected Writings on Computing: A Personal Perspective*. Springer-Verlag, 1982.
- 67 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. Reprinted in his *Werke*, volume 1, 635–638.
- 68 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.
- 69 John Dougall, "On Vandermonde's theorem, and some more general expansions," *Proceedings of the Edinburgh Mathematical Society* 25 (1907), 114–132.
- 70 A. Conan Doyle, "The sign of the four; or, The problem of the Sholtos," *Lippincott's Monthly Magazine* (Philadelphia) 45 (1890), 147–223.
- 71 A. Conan Doyle, "The adventure of the final problem," *The Strand Magazine* 6 (1893), 558–570.
- 72 Henry Ernest Dudeney, *The Canterbury Puzzles and Other Curious Problems*. E. P. Dutton, New York, 1908; 4th edition, Dover, 1958. (Dudeney had first considered the generalized Tower of Hanoi in *The Weekly Dispatch*, on 25 May 1902 and 15 March 1903.)
- 73 G. Waldo Dunnington, *Carl Friedrich Gauss: Titan of Science*. Exposition Press, New York, 1955.
- 74 A. W. F. Edwards, *Pascal's Arithmetical Triangle*. Oxford University Press, 1987.
- 75 G. Eisenstein, "Entwicklung von α^a ," *Journal für die reine und angewandte Mathematik* 28 (1844), 49–52. Reprinted in his *Mathematische Werke* 1, 122–125.
- 76 Erdős Pál, "Az $\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n} = \frac{a}{b}$ egyenlet egész számú megoldásairól," *Matematikai Lapok* 1 (1950), 192–209. English abstract on page 210.
- 77 P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*. Université de Genève, L'Enseignement Mathématique, 1980.

- 78 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.
- 79 Arulappah Eswarathasan and Eugene Levine, "p-integral harmonic sums," *Discrete Mathematics*, to appear.
- 80 Euclid, *ΣΤΟΙΧΕΙΑ*. Ancient manuscript first printed in Basel, 1533. Scholarly edition (Greek and Latin) by J. L. Heiberg in five volumes, Teubner, Leipzig, 1883–1888.
- 81 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.
- 82 Leonhard Euler, "Methodus generalis summandi progressionēs," *Commentarii academīæ scientiarum Petropolitanæ* 6 (1732), 68–97. Reprinted in his *Opera Omnia*, series 1, volume 14, 42–72.
- 83 Leonhard Euler, "De progressionibus harmonicis observationes," *Commentarii academīæ scientiarum Petropolitanæ* 7 (1734), 150–161. Reprinted in his *Opera Omnia*, series 1, volume 14, 87–100.
- 84 Leonhard Euler, "De fractionibus continuis, Dissertatio," *Commentarii academīæ scientiarum Petropolitanæ* 9 (1737), 98–137. Reprinted in his *Opera Omnia*, series 1, volume 14, 187–215.
- 85 Leonhard Euler, "Variæ observationes circa series infinitas," *Commentarii academīæ scientiarum Petropolitanæ* 9 (1737), 160–188. Reprinted in his *Opera Omnia*, series 1, volume 14, 216–244.
- 85' 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.
- 86 Leonhard Euler, *Introductio in Analysin Infinitorum*. Tomus primus, Lausanne, 1748. Reprinted in his *Opera Omnia*, series 1, volume 8. Translated into French, 1786; German, 1788.
- 87 Leonhard Euler, "De partitione numerorum," *Novi commentarii academīæ scientiarum Petropolitanæ* 3 (1750), 125–169. Reprinted in his *Commentationes arithmeticae collectæ*, volume 1, 73–101. Reprinted in his *Opera Omnia*, series 1, volume 2, 254–294.
- 88 Leonhard Euler, *Institutiones Calculi Differentialis cum eius usu in Analysisi Finitorum ac Doctrina Serierum*. Petrograd, Academiæ Imperialis Scientiarum, 1755. Reprinted in his *Opera Omnia*, series 1, volume 10. Translated into German, 1790.

- 89 Leonhard Euler, "Theoremata arithmetica nova methodo demonstrata," *Novi commentarii academiae scientiarum Petropolitanae* 8 (1760), 74–104. (Also presented in 1758 to the Berlin Academy.) Reprinted in his *Commentationes arithmeticae collectae*, volume 1, 274–286. Reprinted in his *Opera Omnia*, series 1, volume 2, 531–555.
- 90 Leonhard Euler, "Specimen algorithmi singularis," *Novi commentarii academiae scientiarum Petropolitanae* 9 (1762), 53–69. (Also presented in 1757 to the Berlin Academy.) Reprinted in his *Opera Omnia*, series 1, volume 15, 31–49.
- 91 Leonhard Euler, "Observationes analyticae," *Novi commentarii academiae scientiarum Petropolitanae* 11 (1765), 124–143. Reprinted in his *Opera Omnia*, series 1, volume 15, 50–69.
- 92 Leonhard Euler, *Vollständige Anleitung zur Algebra. Erster Theil. Von den verschiedenen Rechnungs-Arten, Verhältnissen und Proportionen*. St. Petersburg, 1770. Reprinted in his *Opera Omnia*, series 1, volume 1. Translated into Russian, 1768; Dutch, 1773; French, 1774; Latin, 1790; English, 1797.
- 93 Leonhard Euler, "Observationes circa bina biquadrata quorum summam in duo alia biquadrata resolvere liceat," *Novi commentarii academiae scientiarum Petropolitanae* 17 (1772), 64–69. Reprinted in his *Opera Omnia*, series 1, volume 3, 211–217.
- 94 Leonhard Euler, "Observationes circa novum et singulare progressionum genus," *Novi commentarii academiae scientiarum Petropolitanae* 20 (1775), 123–139. Reprinted in his *Opera Omnia*, series 1, volume 7, 246–261.
- 95 Leonhard Euler, "Specimen transformationis singularis serierum," *Nova acta academiae scientiarum Petropolitanae* 12 (1794), 58–70. Submitted for publication in 1778. Reprinted in his *Opera Omnia*, series 1, volume 16(2), 41–55.
- 96 William Feller, *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, 1950; second edition, 1957; third edition, 1968.
- 97 Pierre de Fermat, letter to Marin Mersenne (25 December 1640), in *Œuvres de Fermat*, volume 2, 212–217.
- 98 Leonardo Fibonacci [Pisano], *Liber Abaci*. First edition, 1202 (now lost); second edition 1228. Reprinted in *Scritti di Leonardo Pisano*, edited by Baldassarre Boncompagni, 1857, volume 1.
- 99 Michael E. Fisher, "Statistical mechanics of dimers on a plane lattice," *Physical Review* 124 (1961), 1664–1672.

- 100 R. A. Fisher, "Moments and product moments of sampling distributions," *Proceedings of the London Mathematical Society*, series 2, 30 (1929), 199–238.
- 101 Pierre Forcadet, *L'arithmétique*. Paris, 1557.
- 102 J. Fourier, "Refroidissement séculaire du globe terrestre," *Bulletin des Sciences par la Société philomathique de Paris*, series 3, 7 (1820), 58–70. Reprinted in *Œuvres de Fourier*, volume 2, 271–288.
- 103 Aviezri S. Fraenkel, "Complementing and exactly covering sequences," *Journal of Combinatorial Theory*, series A, 14 (1973), 8–20.
- 104 Aviezri S. Fraenkel, "How to beat your Wythoff games' opponent on three fronts," *American Mathematical Monthly* 89 (1982), 353–361.
- 105 J. S. Frame, B. M. Stewart, and Otto Dunkel, "Partial solution to problem 3918," *American Mathematical Monthly* 48 (1941), 216–219.
- 106 Piero della Francesca, *Libellus de quinque corporibus regularibus*. Vatican Library, manuscript Urbinas 632. Translated into Italian by Luca Pacioli, as part 3 of Pacioli's *Divine Proportione*, Venice, 1509.
- 107 W. D. Frazer and A. C. McKellar, "Samplesort: A sampling approach to minimal storage tree sorting," *Journal of the ACM* 27 (1970), 496–507.
- 108 Michael Lawrence Fredman, *Growth Properties of a Class of Recursively Defined Functions*. Ph.D. thesis, Stanford University, Computer Science Department, 1972.
- 109 Nikolaus I. Fuss, "Solutio quæstionis, quot modis polygonum n laterum in polygona m laterum, per diagonales resolvi quæat," *Nova acta academiciæ scientiarum Petropolitane* 9 (1791), 243–251.
- 110 Martin Gardner, "About phi, an irrational number that has some remarkable geometrical expressions," *Scientific American* 201, 2 (August 1959), 128–134. Reprinted with additions in his book *The 2nd Scientific American Book of Mathematical Puzzles & Diversions*, 1961, 89–103.
- 111 Martin Gardner, "On the paradoxical situations that arise from nontransitive relations," *Scientific American* 231, 4 (October 1974), 120–124. Reprinted with additions in his book *Time Travel and Other Mathematical Bewilderments*, 1988, 55–69.
- 112 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. Reprinted with additions in his book *Time Travel and Other Mathematical Bewilderments*, 1988, 111–124.
- 113 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.

- 114 J. Carfunkel, "Problem E 1816: An inequality related to Stirling's formula," *American Mathematical Monthly* 74 (1967), 202.
- 115 C. F. Gauss, *Disquisitiones Arithmeticae*. Leipzig, 1801. Reprinted in his *Werke*, volume 1.
- 116 Carolo Friderico Gauss, "Disquisitiones generales circa seriem infinitam

$$1 + \frac{\alpha\beta}{1 \cdot \gamma} x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1 \cdot 2 \cdot \gamma(\gamma+1)} x^2 + \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1 \cdot 2 \cdot 3 \cdot \gamma(\gamma+1)(\gamma+2)} x^3 + \text{etc.}$$

Pars prior," *Commentationes societatis regiae scientiarum Gottingensis recentiores* 2 (1813). (Thesis delivered to the Royal Society in Göttingen, 20 January 1812.) Reprinted in his *Werke*, volume 3, 123–163, together with an unpublished sequel on pages 207–229.

- 117 A. Genocchi, "Intorno all' espressione generale di numeri Bernoulliani," *Annali di Scienze Matematiche e Fisiche* 3 (1852), 395–405.
- 118 Ira Gessel and Richard P. Stanley, "Stirling polynomials," *Journal of Combinatorial Theory*, series A, 24 (1978), 24–33.
- 119 Jekuthiel Ginsburg, "Note on Stirling's numbers," *American Mathematical Monthly* 35 (1928), 77–80.
- 120 Solomon W. Golomb, "Problem 5407: A nondecreasing indicator function," *American Mathematical Monthly* 74 (1967), 740–743.
- 121 Solomon W. Golomb, "The 'Sales Tax' theorem," *Mathematics Magazine* 49 (1976), 187–189.
- 122 Solomon W. Golomb, "Problem E 2529: An application of $\psi(x)$," *American Mathematical Monthly* 83 (1976), 487–488.
- 123 I. J. Good, "Short proof of a conjecture by Dyson," *Journal of Mathematical Physics* 11 (1970), 1884.
- 124 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.
- 125 R. L. Graham, "On a theorem of Uspensky," *American Mathematical Monthly* 70 (1963), 407–409.
- 126 R. L. Graham, "A Fibonacci-like sequence of composite numbers," *Mathematics Magazine* 37 (1964), 322–324.
- 127 R. L. Graham, "Problem 5749," *American Mathematical Monthly* 77 (1970), 775.

- 128 Ronald L. Graham, "Covering the positive integers by disjoint sets of the form $\{n\alpha + \beta\} : n = 1, 2, \dots\}$," *Journal of Combinatorial Theory*, series A, 15 (1973), 354-358.
- 129 R. L. Graham, "Problem 1242: Bijection between integers and composites," *Mathematics Magazine* 60 (1987), 180.
- 130 R. L. Graham and D. E. Knuth, "Problem E2982: A double infinite sum for $|x|$," *American Mathematical Monthly* 96 (1989), 525-526.
- 131 Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, 1989. (The first printing had a different Iversonian notation.)
- 132 R. L. Graham and H. O. Pollak, "Note on a nonlinear recurrence related to $\sqrt{2}$," *Mathematics Magazine* 43 (1970), 143-145.
- 133 Guido Grandi, letter to Leibniz (July 1713), in *Leibnizens mathematische Schriften*, volume 4, 215-217.
- 134 Daniel H. Greene and Donald E. Knuth, *Mathematics for the Analysis of Algorithms*. Birkhäuser, Boston, 1981; third edition, 1990.
- 135 Samuel L. Greitzer, *International Mathematical Olympiads, 1959-1977*. Mathematical Association of America, 1978.
- 136 Oliver A. Gross, "Preferential arrangements," *American Mathematical Monthly* 69 (1962), 4-8.
- 137 Branko Grünbaum, "Venn diagrams and independent families of sets," *Mathematics Magazine* 48 (1975), 12-23.
- 138 L. J. Guibas and A. M. Odlyzko, "String overlaps, pattern matching, and nontransitive games," *Journal of Combinatorial Theory*, series A, 30 (1981), 183-208.
- 139 Richard K. Guy, *Unsolved Problems in Number Theory*. Springer-Verlag, 1981.
- 140 Marshall Hall, Jr., *The Theory of Groups*. Macmillan, 1959.
- 141 P. R. Halmos, "How to write mathematics," *L'Enseignement mathématique* 16 (1970), 123-152. Reprinted in *How to Write Mathematics*, American Mathematical Society, 1973, 19-48.
- 142 Paul R. Halmos, *I Want to Be a Mathematician: An Automathography*. Springer-Verlag, 1985. Reprinted by Mathematical Association of America, 1988.
- 143 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. Reprinted in his *Ouvres*, volume 2, 102-107.

- 144 Hans Hamburger, "Über eine Erweiterung des Stieltjesschen Momentenproblems," *Mathematische Annalen* 81 (1920), 235–319; 82 (1921), 120–164, 168–187.
- 145 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.
- 146 J. M. Hammersley, "An undergraduate exercise in manipulation," *The Mathematical Scientist* 14 (1989), 1–23.
- 147 Eldon R. Hansen, *A Table of Series and Products*. Prentice-Hall, 1975.
- 148 G. H. Hardy, *Orders of Infinity: The 'Infinitärcalcul' of Paul du Bois-Reymond*. Cambridge University Press, 1910; second edition, 1924.
- 149 G. H. Hardy, "A mathematical theorem about golf," *The Mathematical Gazette* 29 (1944), 226–227. Reprinted in his *Collected Papers*, volume 7, 488.
- 150 G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*. Clarendon Press, Oxford, 1938; fifth edition, 1979.
- 151 Peter Henrici, *Applied and Computational Complex Analysis*. Wiley, volume 1, 1974; volume 2, 1977; volume 3, 1986.
- 152 Peter Henrici, "De Branges' proof of the Bieberbach conjecture: A view from computational analysis," *Sitzungsberichte der Berliner Mathematischen Gesellschaft* (1987), 105–121.
- 153 Charles Hermite, letter to C. W. Borchardt (8 September 1875), in *Journal für die reine und angewandte Mathematik* 81 (1876), 93–95. Reprinted in his *Œuvres*, volume 3, 211–214.
- 154 Charles Hermite, *Cours de M. Hermite*. Faculté des Sciences de Paris, 1882. Third edition, 1887; fourth edition, 1891.
- 155 Charles Hermite, letter to S. Pincherle (10 May 1900), in *Annali di Matematica pura ed applicata*, series 3, 5 (1901), 57–60. Reprinted in his *Œuvres*, volume 4, 529–531.
- 156 I. N. Herstein and I. Kaplansky, *Matters Mathematical*. Harper & Row, 1974.
- 157 A. P. Hillman and V. E. Hoggatt, Jr., "A proof of Gould's Pascal hexagon conjecture," *Fibonacci Quarterly* 10 (1972), 565–568, 598.
- 158 C. A. R. Hoare, "Quicksort," *The Computer Journal* 5 (1962), 10–15.
- 159 L. C. Hsu, "Note on a combinatorial algebraic identity and its application," *Fibonacci Quarterly* 11 (1973), 480–484.

- 160 K. Inkeri, "Abschätzungen für eventuelle Lösungen der Gleichung im Fermatschen Problem," *Annales Universitatis Turkuensis*, series A, 16, 1 (1953), 3–9.
- 161 Kenneth E. Iverson, *A Programming Language*. Wiley, 1962.
- 162 C.G.J. Jacobi, *Fundamenta nova theoriæ functionum ellipticarum*. Königsberg, Bornträger, 1829. Reprinted in his *Gesammelte Werke*, volume 1, 49–239.
- 163 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 (Hebrew with English summary). English version reprinted in Dov Jarden, *Recurring Sequences*, Jerusalem, 1958, 42–45; second edition, Jerusalem, 1966, 30–33.
- 164 Arne Jonassen and Donald E. Knuth, "A trivial algorithm whose analysis isn't," *Journal of Computer and System Sciences* 16 (1978), 301–322.
- 165 Bush Jones, "Note on internal merging," *Software — Practice and Experience* 2 (1972), 241–243.
- 166 Flavius Josephus, *ΙΣΤΟΡΙΑ ΙΟΥΔΑΪΚΟΥ ΠΟΛΕΜΟΥ ΠΡΟΣ ΡΩΜΑΙΟΥΣ*. English translation, *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. (The "Josephus problem" may be based on an early manuscript now preserved only in the Slavonic version; see volume 2, page xi, and volume 3, page 654.)
- 167 R. Jungen, "Sur les séries de Taylor n'ayant que des singularités algébrique-logarithmiques sur leur cercle de convergence," *Commentarii Mathematici Helvetici* 3 (1931), 266–306.
- 168 I. Kaucký, "Problem E 2257: A harmonic identity," *American Mathematical Monthly* 78 (1971), 908.
- 169 Murray S. Klamkin, *International Mathematical Olympiads, 1978–1985, and Forty Supplementary Problems*. Mathematical Association of America, 1986.
- 170 Konrad Knopp, *Theorie und Anwendung der unendlichen Reihen*. Julius Springer, Berlin, 1922; second edition, 1924. Reprinted by Dover, 1945. Fourth edition, 1947; fifth edition, 1964. English translation, *Theory and Application of Infinite Series*, 1928; second edition, 1951.
- 171 Donald E. Knuth, "Euler's constant to 1271 places," *Mathematics of Computation* 16 (1962), 275–281.
- 172 Donald Knuth, "Transcendental numbers based on the Fibonacci sequence," *Fibonacci Quarterly* 2 (1964), 43–44, 52.

- 173 Donald E. Knuth, *The Art of Computer Programming*, volume 1: *Fundamental Algorithms*. Addison-Wesley, 1968; second edition, 1973.
- 174 Donald E. Knuth, *The Art of Computer Programming*, volume 2: *Seminumerical Algorithms*. Addison-Wesley, 1969; second edition, 1981.
- 175 Donald E. Knuth, *The Art of Computer Programming*, volume 3: *Sorting and Searching*. Addison-Wesley, 1973; second printing, 1975.
- 176 Donald E. Knuth, "Problem E2492: Some sum," *American Mathematical Monthly* 82 (1975), 855.
- 177 Donald E. Knuth, *Mariages stables et leurs relations avec d'autres problèmes combinatoires*. Les Presses de l'Université de Montréal, 1976. Revised and corrected edition, 1980.
- 178 Donald E. Knuth, *The TeXbook*. Addison-Wesley, 1984. Reprinted as volume A of *Computers & Typesetting*, 1986.
- 179 Donald E. Knuth, "An analysis of optimum caching," *Journal of Algorithms* 6 (1985), 181-199.
- 180 Donald E. Knuth, *Computers & Typesetting*, volume D: *METAFONT: The Program*. Addison-Wesley, 1986.
- 181 Donald E. Knuth, "Problem 1280," *Mathematics Magazine* 60 (1987), 329.
- 182 Donald E. Knuth, "Problem E3106: A new sum for n^2 ," *American Mathematical Monthly* 94 (1987), 795-797.
- 183 Donald E. Knuth, "Fibonacci multiplication," *Applied Mathematics Letters* 1 (1988), 57-60.
- 184 Donald E. Knuth, "A Fibonacci-like sequence of composite numbers," *Mathematics Magazine* 63 (1990), 21-25.
- 185 Donald E. Knuth and Thomas J. Buckholtz, "Computation of Tangent, Euler, and Bernoulli numbers," *Mathematics of Computation* 21 (1967), 663-688.
- 186 C. Kramp, *Éléments d'arithmétique universelle*. Cologne, 1808.
- 187 E. E. Kummer, "Ueber die hypergeometrische Reihe

$$1 + \frac{\alpha\beta}{1 \cdot \gamma} x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1 \cdot 2 \cdot \gamma(\gamma+1)} x^2 + \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1 \cdot 2 \cdot 3 \cdot \gamma(\gamma+1)(\gamma+2)} x^3 + \dots,$$

Journal für die reine und angewandte Mathematik 15 (1836), 39-83, 127-172. Reprinted in his *Collected Papers*, volume 2, 75-166.

- 188 E. E. Kummer, "Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen," *Journal für die reine und angewandte Mathematik*, 44 (1852), 93–146. Reprinted in his *Collected Papers*, volume 1, 485–538.
- 189 R. P. Kurshan and B. Gopinath, "Recursively generated periodic sequences," *Canadian Journal of Mathematics* 26 (1974), 1356–1371.
- 190 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*. Published as volume 11 of *Mémoires de l'Académie Royale des Sciences*, Paris, 1733.
- 191 J.-L. de la Grange [Lagrange], "Démonstration d'un théorème nouveau concernant les nombres premiers," *Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin* (1771), 125–137. Reprinted in his *Œuvres*, volume 3, 425–438.
- 192 J.-L. de la Grange [Lagrange], "Sur une nouvelle espèce de calcul relatif à la différentiation & à l'intégration des quantités variables," *Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin* (1772), 185–221. Reprinted in his *Œuvres*, volume 3, 441–476.
- 193 I. Lah, "Eine neue Art von Zahlen, ihre Eigenschaften und Anwendung in der mathematischen Statistik," *Mitteilungsblatt für Mathematische Statistik* 7 (1955), 203–212.
- 194 Edmund Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, two volumes. Teubner, Leipzig, 1909.
- 195 Edmund Landau, *Vorlesungen über Zahlentheorie*, three volumes. Hirzel, Leipzig, 1927.
- 195' P. S. de la Place [Laplace], "Mémoire sur les approximations des Formules qui sont fonctions de très-grands nombres," *Mémoires de l'Académie royale des Sciences de Paris* (1782), 1–88. Reprinted in his *Œuvres Complètes* 10, 207–291.
- 196 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.
- 197 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.
- 198 D. H. Lehmer, "On Stern's diatomic series," *American Mathematical Monthly* 36 (1929), 59–67.
- 199 D. H. Lehmer, "On Euler's totient function," *Bulletin of the American Mathematical Society*, series 2, 38 (1932), 745–751.

- 200 G. W. Leibniz, letter to Johann Bernoulli (May 1695), in *Leibnizens mathematische Schriften*, volume 3, 174-179.
- 201 C. G. Lekkerkerker, "Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci," *Simon Stevin* 29 (1952), 190-195.
- 201' Elliott H. Lieb, "Residual entropy of square ice," *Physical Review* 162 (1967), 162-172.
- 202 B. F. Logan, "The recovery of orthogonal polynomials from a sum of squares," *SIAM Journal on Mathematical Analysis* 21 (1990), 1031-1050.
- 202' B. F. Logan, "Polynomials related to the Stirling numbers," AT&T Bell Laboratories internal technical memorandum, August 10, 1987.
- 203 Calvin T. Long and Verner E. Hoggatt, Jr., "Sets of binomial coefficients with equal products," *Fibonacci Quarterly* 12 (1974), 71-79.
- 204 Sam Loyd, *Cyclopedia of Puzzles*. Franklin Bigelow Corporation, Morningside Press, New York, 1914.
- 205 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.
- 206 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 (1878), 49-54.
- 207 Edouard Lucas, *Théorie des Nombres*, volume 1. Gauthier-Villars, Paris, 1891.
- 208 Édouard Lucas, *Récréations mathématiques*, four volumes. Gauthier-Villars, Paris, 1891-1894. Reprinted by Albert Blanchard, Paris, 1960. (The Tower of Hanoi is discussed in volume 3, pages 55-59.)
- 209 R. C. Lyness, "Cycles," *The Mathematical Gazette* 26 (1942), 62.
- 210 R. C. Lyness, "Cycles," *The Mathematical Gazette* 29 (1945), 231-233.
- 211 Colin Maclaurin, *Collected Letters*, edited by Stella Mills. Shiva Publishing, Nantwich, Cheshire, 1982.
- 212 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.
- 213 Iu. V. Matijasevich, "Diofantovost' perechislimykh mnozhestv," *Doklady Akademii Nauk SSSR* 191 (1970), 279-282. English translation, with amendments by the author, "Enumerable sets are diophantine," *Soviet Mathematics* 11 (1970), 354-357.

- 214 Z. A. Melzak, *Companion to Concrete Mathematics*. Volume 1, *Mathematical Techniques and Various Applications*, Wiley, 1973; volume 2, *Mathematical Ideas, Modeling & Applications*, Wiley, 1976.
- 215 N. S. Mendelsohn, "Problem E 2227: Divisors of binomial coefficients," *American Mathematical Monthly* 78 (1971), 201.
- 216 W. H. Mills, "A prime representing function," *Bulletin of the American Mathematical Society*, series 2, 53 (1947), 604.
- 217 A. Moessner, "Eine Bemerkung über die Potenzen der natürlichen Zahlen," *Sitzungsberichte der Mathematisch - Naturwissenschaftliche Klasse der Bayerischen Akademie der Wissenschaften*, 1951, Heft 3, 29.
- 218 Peter L. Montgomery, "Problem E 2686: LCM of binomial coefficients," *American Mathematical Monthly* 86 (1979), 131.
- 219 Leo Moser, "Problem B-6: Some reflections," *Fibonacci Quarterly* 1, 4 (1963), 75-76.
- 220 T. S. Motzkin and E. G. Straus, "Some combinatorial extremum problems," *Proceedings of the American Mathematical Society* 7 (1956), 1014-1021.
- 221 C. J. Mozzochi, "On the difference between consecutive primes," *Journal of Number Theory* 24 (1986), 181-187.
- 222 B. R. Myers, "Problem 5795: The spanning trees of an n-wheel," *American Mathematical Monthly* 79 (1972), 914-915.
- 223 Isaac Newton, letter to John Collins (18 February 1670), in *The Correspondence of Isaac Newton*, volume 1, 27. Excerpted in *The Mathematical Papers of Isaac Newton*, volume 3, 563.
- 224 Ivan Niven, *Diophantine Approximations*. Interscience, 1963.
- 225 Ivan Niven, "Formal power series," *American Mathematical Monthly* 76 (1969), 871-889.
- 226 Blaise Pascal, "De numeris multiplicibus," presented to Académie Parisienne in 1654 and published with his *Traité du triangle arithmétique* [227]. Reprinted in *Œuvres de Blaise Pascal*, volume 3, 314-339.
- 227 Blaise Pascal, "Traité du triangle arithmétique," in his *Traité du Triangle Arithmétique, avec quelques autres petits traitez sur la mesme matiere*, Paris, 1665. Reprinted in *Œuvres de Blaise Pascal* (Hachette, 1904-1914), volume 3, 445-503; Latin editions from 1654 in volume 11, 366-390.
- 228 G. P. Patil, "On the evaluation of the negative binomial distribution with examples," *Technometrics* 2 (1960), 501-505.

- 229 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. (See also page 211.)
- 230 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. (See also “Ordinals,” an unpublished manuscript from circa 1905, in *Collected Papers of Charles Sanders Peirce*, volume 4, 268–280.)
- 231 Walter Penney, “Problem 95: Penney-Ante,” *Journal of Recreational Mathematics* 7 (1974), 321.
- 232 J. K. Percus, *Combinatorial Methods*. Springer-Verlag, 1971.
- 233 J. F. Pfaff, “Observationes analyticae ad L. Euleri institutiones calculi integralis, Vol. IV, Supplem. II & IV,” *Nova acta academicae scientiarum Petropolitanae* 11, Histoire section, 37–57. (This volume, printed in 1798, contains mostly proceedings from 1793, although Pfaff’s memoir was actually received in 1797.)
- 234 L. Pochhammer, “Ueber hypergeometrische Functionen n^{ter} Ordnung,” *Journal für die reine und angewandte Mathematik* 71 (1870), 316–352.
- 235 H. Poincaré, “Sur les fonctions à espaces lacunaires,” *American Journal of Mathematics* 14 (1892), 201–221.
- 236 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.
- 237 G. Pólya, “Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen,” *Acta Mathematica* 68 (1937), 145–254.
- 238 George Pólya, *Induction and Analogy in Mathematics*. Princeton University Press, 1954.
- 239 G. Pólya, “On picture-writing,” *American Mathematical Monthly* 63 (1956), 689–697.
- 240 G. Pólya and G. Szegő, *Aufgaben und Lehrsätze aus der Analysis*, two volumes. Julius Springer, Berlin, 1925; fourth edition, 1970 and 1971. English translation, *Problems and Theorems in Analysis*, 1972 and 1976.
- 241 Bjorn Poonen, “Josephus sets.” Unpublished manuscript, 1987.
- 242 R. Rado, “A note on the Bernoullian numbers,” *Journal of the London Mathematical Society* 9 (1934), 88–90.
- 242’ Earl D. Rainville, “The contiguous function relations for ${}_pF_q$ with applications to Bateman’s J_n^u and Rice’s $H_n(\zeta, p, v)$,” *Bulletin of the American Mathematical Society*, series 2, 51 (1945), 714–723.

- 243 George N. Raney, "Functional composition patterns and power series reversion," *Transactions of the American Mathematical Society* 94 (1960), 441-451.
- 244 D. Rameswar Rao, "Problem E2208: A divisibility problem," *American Mathematical Monthly* 78 (1971), 78-79.
- 245 John William Strutt, Third Baron Rayleigh, *The Theory of Sound*. First edition, 1877; second edition, 1894. (The cited material about irrational spectra is from section 92a of the second edition.)
- 246 Robert Recorde, *The Whetstone of Witte*. London, 1557.
- 247 Simeon Reich, "Problem 6056: Truncated exponential-type series," *American Mathematical Monthly* 84 (1977), 494-495.
- 248 Georges de Rham, "Un peu de mathématiques à propos d'une courbe plane," *Elemente der Mathematik* 2 (1947), 73-76, 89-97. Reprinted in his *Ouvres Mathématiques*, 678-689.
- 249 Paolo Ribenboim, *13 Lectures on Fermat's Last Theorem*. Springer-Verlag, 1979.
- 250 Bernhard Riemann, "Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe," Habilitationsschrift, Göttingen, 1854. Published in *Abhandlungen der mathematischen Classe der Königlischen Gesellschaft der Wissenschaften zu Göttingen* 13 (1868), 87-132. Reprinted in his *Gesammelte Mathematische Werke*, 227-264.
- 251 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.
- 252 Øystein Rødseth, "Problem E2273: Telescoping Vandermonde convolutions," *American Mathematical Monthly* 79 (1972), 88-89.
- 253 J. Barkley Rosser and Lowell Schoenfeld, "Approximate formulas for some functions of prime numbers," *Illinois Journal of Mathematics* 6 (1962), 64-94.
- 254 Gian-Carlo Rota, "On the foundations of combinatorial theory. I. Theory of Möbius functions," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 2 (1964), 340-368.
- 255 Ranjan Roy, "Binomial identities and hypergeometric series," *American Mathematical Monthly* 94 (1987), 36-46.
- 256 Louis Saalschütz, "Eine Summationsformel," *Zeitschrift für Mathematik und Physik* 35 (1890), 186-188.

- 256' A. Sárközy, "On divisors of binomial coefficients, I," *Journal of Number Theory* 20 (1985), 70-80.
- 257 W. W. Sawyer, *Prelude to Mathematics*. Baltimore, Penguin, 1955.
- 258 O. Schlömilch, "Ein geometrisches Paradoxon," *Zeitschrift für Mathematik und Physik* 13 (1868), 162.
- 259 Ernst Schröder, "Vier combinatorische Probleme," *Zeitschrift für Mathematik und Physik* 15 (1870), 361-376.
- 260 Heinrich Schröter, "Ableitung der Partialbruch- und Produkt-Entwicklungen für die trigonometrischen Funktionen," *Zeitschrift für Mathematik und Physik* 13 (1868), 254-259.
- 261 R. S. Scorer, P. M. Grundy, and C. A. B. Smith, "Some binary games," *The Mathematical Gazette* 28 (1944), 96-103.
- 262 J. Sedláček, "On the skeletons of a graph or digraph," in *Combinatorial Structures and their Applications*, Gordon and Breach, 1970, 387-391. (This volume contains proceedings of the Calgary International Conference of Combinatorial Structures and their Applications, 1969.)
- 263 J. O. Shallit, "Problem 6450: Two series," *American Mathematical Monthly* 92 (1985), 513-514.
- 264 R. T. Sharp, "Problem 52: Overhanging dominoes," *Pi Mu Epsilon Journal* 1, 10 (1954), 411-412.
- 265 W. Sierpiński, "Sur la valeur asymptotique d'une certaine somme," *Bulletin International Académie Polonaise des Sciences et des Lettres* (Cracovie), series A (1910), 9-11.
- 266 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.
- 267 Wacław Sierpiński, *A Selection of Problems in the Theory of Numbers*. Macmillan, 1964.
- 268 David L. Silverman, "Problematical Recreations 447: Numerical links," *Aviation Week & Space Technology* 89, 10 (1 September 1968), 71. Reprinted as Problem 147 in *Second Book of Mathematical Bafflers*, edited by Angela Fox Dunn, Dover, 1983.
- 269 Lucy Joan Slater, *Generalized Hypergeometric Series*. Cambridge University Press, 1966.
- 270 N. J. A. Sloane, *A Handbook of Integer Sequences*. Academic Press, 1973.

- 271 A. D. Solov'ev, "Oдно kombinatornoe tozhdestvo i ego primeneniye k zadache o pervom nastuplenii redkogo sobytiya," *Teoriya veroyatnostey i ee primeneniya* 11 (1966), 313–320. English translation, "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.
- 272 William G. Spohn, Jr., "Can mathematics be saved?" *Notices of the American Mathematical Society* 16 (1969), 890–894.
- 273 Richard P. Stanley, "Differentiably finite power series," *European Journal of Combinatorics* 1 (1980), 175–188.
- 274 Richard P. Stanley, "On dimer coverings of rectangles of fixed width," *Discrete Applied Mathematics* 12 (1985), 81–87.
- 275 Richard P. Stanley, *Enumerative Combinatorics*, volume 1. Wadsworth & Brooks/Cole, 1986.
- 276 K. G. C. von Staudt, "Beweis eines Lehrsatzes, die Bernoullischen Zahlen betreffend," *Journal für die reine und angewandte Mathematik* 21 (1840), 372–374.
- 277 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.
- 278 J. Steiner, "Einige Gesetze über die Theilung der Ebene und des Raumes," *Journal für die reine und angewandte Mathematik* 1 (1826), 349–364. Reprinted in his *Gesammelte Werke*, volume 1, 77–94.
- 279 M. A. Stern, "Ueber eine zahlentheoretische Funktion," *Journal für die reine und angewandte Mathematik* 55 (1858), 193–220.
- 280 L. Stickelberger, "Ueber eine Verallgemeinerung der Kreistheilung," *Mathematische Annalen* 37 (1890), 321–367.
- 281 James Stirling, *Methodus Differentialis*. London, 1730. English translation, *The Differential Method*, 1749.
- 282 Dora W. Sweeney, "On the computation of Euler's constant," *Mathematics of Computation* 17 (1963), 170–178.
- 283 J. J. Sylvester, "Problem 6919," *Mathematical Questions with their Solutions from the 'Educational Times'* 37 (1882), 42–43, 80.
- 284 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. Reprinted in his *Collected Mathematical Papers*, volume 4, 101–109.

- 284' M. Szegedy, "The solution of Graham's greatest common divisor problem," *Combinatorica* 6 (1986), 67-71.
- 285 Jonathan W. Tanner and Samuel S. Wagstaff, Jr., "New congruences for the Bernoulli numbers," *Mathematics of Computation* 48 (1987), 341-350.
- 286 S. Tanny, "A probabilistic interpretation of Eulerian numbers," *Duke Mathematical Journal* 40 (1973), 717-722.
- 287 L. Theisinger, "Bemerkung über die harmonische Reihe," *Monatshefte für Mathematik und Physik* 26 (1915), 132-134.
- 288 T.N. Thiele, *The Theory of Observations*. Charles & Edwin Layton, London, 1903. Reprinted in *The Annals of Mathematical Statistics* 2 (1931), 165-308.
- 289 E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*. Clarendon Press, Oxford, 1951; second edition, revised by D. R. Heath-Brown, 1986.
- 290 F.G. Tricomi and A. Erdélyi, "The asymptotic expansion of a ratio of gamma functions," *Pacific Journal of Mathematics* 1 (1951), 133-142.
- 291 Peter Ungar, "Problem E3052: A sum involving Stirling numbers," *American Mathematical Monthly* 94 (1987), 185-186.
- 292 J.V. Uspensky, "On a problem arising out of the theory of a certain game," *American Mathematical Monthly* 34 (1927), 516-521.
- 293 A. Vandermonde, "Mémoire sur des irrationnelles de différens ordres avec une application au cercle," *Histoire de l'Académie Royale des Sciences* (1772), part 1, 71-72; *Mémoires de Mathématique et de Physique, Tirés des Registres de l'Académie Royale des Sciences* (1772), 489-498.
- 294 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.
- 295 John Wallis, *A Treatise of Angular Sections*. Oxford, 1684.
- 296 Edward Waring, *Meditationes Algebraicae*. Cambridge, 1770; third edition, 1782.
- 296' William C. Waterhouse, "Problem E3117: Even odder than we thought," *American Mathematical Monthly* 94 (1987), 691-692.
- 297 Frederick V. Waugh and Margaret W. Maxfield, "Side-and-diagonal numbers," *Mathematics Magazine* 40 (1967), 74-83.
- 298 Warren Weaver, "Lewis Carroll and a geometrical paradox," *American Mathematical Monthly* 45 (1938), 234-236.

- 299 Louis Weisner, "Abstract theory of inversion of finite series," *Transactions of the American Mathematical Society* 38 (1935), 474-484.
- 300 Hermann Weyl, "Über die Gibbs'sche Erscheinung und verwandte Konvergenzphänomene," *Rendiconti del Circolo Matematico di Palermo* 30 (1910), 377-407.
- 301 F. J. W. Whipple, "Some transformations of generalized hypergeometric series," *Proceedings of the London Mathematical Society*, series 2, 26 (1927), 257-272.
- 302 Alfred North Whitehead, *An Introduction to Mathematics*. London and New York, 1911.
- 303 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. Reprinted as chapter 4 of *The Aims of Education and Other Essays*, New York, 1929.
- 304 Alfred North Whitehead, *Science and the Modern World*. New York, 1925. Chapter 2 reprinted in *The World of Mathematics*, edited by James R. Newman, 1956, volume 1, 402-416.
- 304' Herbert S. Wilf, *generatingfunctionology*. Academic Press, 1990.
- 305 H. C. Williams and H. Dubner, "The primality of R1031," *Mathematics of Computation* 47 (1986), 703-711.
- 306 J. Wolstenholme, "On certain properties of prime numbers," *Quarterly Journal of Pure and Applied Mathematics* 5 (1862), 35-39.
- 307 Derick Wood, "The Towers of Brahma and Hanoi revisited," *Journal of Recreational Mathematics* 14 (1981), 17-24.
- 308 J. Worpitzky, "Studien über die *Bernoullischen* und *Eulerschen* Zahlen," *Journal für die reine und angewandte Mathematik* 94 (1883), 203-232.
- 309 E. M. Wright, "A prime-representing function," *American Mathematical Monthly* 58 (1951), 616-618; errata in 59 (1952), 99.
- 310 Hermann Zapf, collected works, entitled *Hermann Zapf & His Design Philosophy*. Society of Typographic Arts, Chicago, 1987. (The AMS Euler typeface is mentioned on pages 97 and 136.)
- 311 Derek A. Zave, "A series expansion involving the harmonic numbers," *Information Processing Letters* 5 (1976), 75-77.
- 312 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.

附录 C 对习题出处的赞誉

本书中的习题取自许多方面，除很基本的习题外，作者试图查明以前已发表的所有问题的来龙去脉。

许多习题取自 Stanford 的具体数学班的考试题。助教和主讲教师常常为这些考试设计新的问题，所以在这里列出他们的名字：

年	主讲教师	助教
1970	Don Knuth	Vaughan Pratt
1971	Don Knuth	Leo Guibas
1973	Don Knuth	Henson Graves, Louis Jouaillie
1974	Don Knuth	Scot Drysdale, Tom Porter
1975	Don Knuth	Mark Brown, Luis Trabb Pardo
1976	Andy Yao	Mark Brown, Lyle Ramshaw
1977	Andy Yao	Yossi Shiloach
1978	frances Yao	Yossi Shiloach
1979	Ron Graham	Frank Liang, Chris Tong, Mark Haiman
1980	Andy Yao	Andrei Broder, Jim McGrath
1981	Ron Graham	Oren Patashnik
1982	Ernst Mayr	Joan Feigenbaum, Dave Helmbold
1983	Ernst Mayr	Anna Karlin
1984	Don Knuth	Oren Patashnik, Alex Schaffer
1985	Andrei Broder	Pang Chen, Stefan Sharkansky
1986	Don Knuth	Arif Merchant, Stefan Sharkansky

此外，David Klarner (1971)，Bob Sedgewick (1974)，Leo Guibas (1975) 和 Lyle Ramshaw (1979) 给班上分别作了 6 次以上的讲座。助教收集而主讲教师整理的这些讲稿是本书的基础。

-
- | | | | |
|------|----------------------------------------------------------------|------|--------------------------------------------------------------|
| 1.1 | Pólya [238, p. 120]. | 3.34 | 1970 midterm. |
| 1.2 | Scorer, Grundy, and Smith [261]. | 3.35 | 1975 midterm. |
| 1.5 | Venn [294]. | 3.36 | 1976 midterm. |
| 1.6 | Steiner [278]; Roberts [251]. | 3.37 | 1986 midterm; [181]. |
| 1.8 | Lyness [209]. | 3.38 | 1974 midterm. |
| 1.9 | Cauchy [47, note 2, theorem 17]. | 3.39 | 1971 midterm. |
| 1.10 | Atkinson [13]. | 3.40 | 1980 midterm. |
| 1.11 | Inspired by Wood [307]. | 3.41 | Klamkin [169, problem 1978/3]. |
| 1.14 | Steiner [278]; Pólya [238, chapter 3];
Brother Alfred [37]. | 3.42 | Uspensky [292]. |
| 1.17 | Dudeney [72, puzzle 1]. | 3.45 | Aho and Sloane [4]. |
| 1.21 | Ball [16] credits B. A. Swinden. | 3.46 | Graham and Pollak [132]. |
| 1.22 | Based on an idea of Peter Shor.* | 3.48 | R. L. Graham and D. R. Hofstadter.* |
| 1.23 | Bjorn Poonen.* | 3.51 | Fraenkel [103]. |
| 1.25 | Frame, Stewart, and Dunkel [105]. | 3.52 | S. K. Stein.* |
| 2.2 | Iverson [161, p. 11]. | 4.4 | [180, 4526]. |
| 2.3 | [173, exercise 1.2.3-2]. | 4.16 | Sylvester [283]. |
| 2.5 | [173, exercise 1.2.3-25]. | 4.19 | Bertrand [23, p. 129]; Chebyshev [50];
Wright [309]. |
| 2.22 | Cauchy [47, note 2, theorem 16]. | 4.21 | [178, pp. 148-149]. |
| 2.23 | 1982 final. | 4.22 | Brillhart [34]; Williams and Dub-
ner [305]. |
| 2.26 | [173, exercise 1.2.3-26]. | 4.23 | Crowe [58]. |
| 2.29 | 1979 midterm. | 4.24 | Legendre [196, second edition,
introduction]. |
| 2.30 | 1973 midterm. | 4.26 | [174, exercise 4.5.3-43]. |
| 2.34 | Riemann [250, section 3]. | 4.31 | Pascal [226]. |
| 2.35 | Euler [85] gave a fallacious "proof"
using divergent series | 4.36 | Hardy and Wright [150, 14.5]. |
| 2.36 | Golomb [120]; Ilan Vardi.* | 4.37 | Aho and Sloane [4]. |
| 2.37 | Leo Moser.* | 4.38 | Lucas [205]. |
| 3.6 | Ernst Mayr, 1982 homework. | 4.39 | [129]. |
| 3.8 | Dirichlet [67]. | 4.40 | Stickelberger [280]. |
| 3.9 | Chace [48]; Fibonacci [98, pp. 77-83]. | 4.41 | Legendre [196, 4135]; Hardy and
Wright [150, theorem 82]. |
| 3.12 | [173, exercise 1.2.4-48(a)]. | 4.42 | [174, exercise 4.5.1-6]. |
| 3.13 | Beatty [18]; Niven [224, theorem 3.7]. | 4.44 | [174, exercise 4.5.3-39]. |
| 3.19 | [173, exercise 1.2.4-34]. | 4.45 | [174, exercise 4.3.2-13]. |
| 3.21 | 1975 midterm. | 4.47 | Lehmer [197]. |
| 3.23 | [173, exercise 1.2.4-41]. | 4.48 | Gauss [115, 478]; Crete [57]. |
| 3.28 | Brown [40]. | 4.52 | 1974 midterm. |
| 3.30 | Aho and Sloane [4]. | 4.53 | 1973 midterm, inspired by Rao [244]. |
| 3.31 | Greitzer [135, problem 1972/3,
solution 2]. | 4.54 | 1974 midterm. |
| 3.32 | [130]. | 4.56 | Logan [202, eq. (6.15)]. |
| 3.33 | 1984 midterm. | | |

-
- | | | | |
|------|-----------------------------------------------------------------|------|---------------------------------------|
| 4.57 | A special case appears in [182] | 5.63 | 1974 midterm. |
| 4.58 | Sierpiński [266]. | 5.64 | 1980 midterm. |
| 4.59 | Curliiss [59], Erdős [76]. | 5.65 | 1983 midterm. |
| 4.60 | Mills [216] | 5.66 | 1984 midterm. |
| 4.61 | [173, exercise 1.3.2–19] | 5.67 | 1976 midterm. |
| 4.63 | Barlow [17], Abel [1]. | 5.68 | 1985 midterm. |
| 4.64 | Peirce [229]. | 5.69 | Lyle Ramshaw, guest lecture in 1986. |
| 4.66 | Rubenboim [249]; Sierpiński [267,
problem P ₁₉]. | 5.70 | Andrews [9, theorem 5.4]. |
| 4.67 | [127] | 5.71 | H. S. Wilf [304', exercise 4.16]. |
| 4.69 | Chamér [56]. | 5.72 | Hermite [154]. |
| 4.70 | P. Erdős.* | 5.74 | 1979 midterm. |
| 4.71 | [77, p. 96]. | 5.75 | 1971 midterm. |
| 4.72 | [77, p. 103]. | 5.76 | [173, exercise 1.2.6–59 (corrected)]. |
| 4.73 | Landau [195, volume 2, eq. 648]. | 5.77 | 1986 midterm. |
| 5.1 | Forcadel [101]. | 5.78 | [176]. |
| 5.3 | Long and Hoggatt [203]. | 5.79 | Mendelsohn [215]; Montgomery [218]. |
| 5.5 | 1983 in-class final. | 5.81 | 1986 final exam. |
| 5.13 | 1975 midterm. | 5.82 | Hillman and Hoggatt [157]. |
| 5.14 | [173, exercise 1.2.6–20]. | 5.85 | Hsu [159]. |
| 5.15 | Dixon [68]. | 5.86 | Good [123]. |
| 5.21 | Euler [81]. | 5.88 | Hermite [155]. |
| 5.25 | Gauss [116, 47]. | 5.91 | Whipple [301]. |
| 5.28 | Euler [95]. | 5.92 | Clausen [51], [52]. |
| 5.29 | Kummer [187, eq. 26.4]. | 5.93 | Gosper [124]. |
| 5.31 | Gosper [124]. | 5.94 | Henrici [152, p. 118]. |
| 5.34 | Bailey [15, §10.4]. | 5.95 | [77, p. 71]. |
| 5.36 | Kummer [188, p. 116]. | 5.96 | [77, p. 71]. |
| 5.37 | Vandermonde [293]. | 5.97 | R. William Gosper, Jr.* |
| 5.38 | [173, exercise 1.2.6–16]. | 6.6 | Fibonacci [98, p. 283]. |
| 5.40 | Rødseth [252]. | 6.15 | [175, exercise 5.1.3–2]. |
| 5.43 | Pfaff [233]; Saalschütz [256];
[173, exercise 1.2.6–31]. | 6.21 | Theisinger [287]. |
| 5.48 | Ranjan Roy.* | 6.25 | Gardner [112] credits Denys Wilquin. |
| 5.49 | Roy [255, eq. 3.13]. | 6.27 | Lucas [205]. |
| 5.53 | Gauss [116]; Richard Askey.* | 6.28 | Lucas [207, chapter 18]. |
| 5.58 | Frazer and McKellar [107]. | 6.31 | Lah [193]; R. W. Floyd.* |
| 5.59 | Stanford Computer Science Compre-
hensive Exam, Winter 1987. | 6.35 | 1977 midterm. |
| 5.60 | [173, exercise 1.2.6–41]. | 6.37 | Shallit [263]. |
| 5.61 | Lucas [206]. | 6.39 | [173, exercise 1.2.7–15]. |
| 5.62 | 1971 midterm. | 6.40 | Klamkin [169, problem 1979/1]. |
| | | 6.41 | 1973 midterm. |
| | | 6.43 | Brooke and Wall [36]. |
| | | 6.44 | Matinasevich [213]. |

-
- | | | | |
|------|-----------------------------------------------------------|------|-------------------------------------------------------|
| 6.46 | Francesca [106]; Wallis [295, chapter 4]. | 7.8 | Zave [311]. |
| 6.47 | Lucas [205]. | 7.9 | [173, exercise 1.2.7–22]. |
| 6.48 | [174, exercise 4.5.3–9(c)]. | 7.11 | 1971 final exam. |
| 6.49 | Davison [61]. | 7.12 | [175, pp. 63–64]. |
| 6.50 | 1985 midterm; Rham [248]; Dijkstra [66, pp. 230–232]. | 7.13 | Raney [243]. |
| 6.51 | Waring [296]; Lagrange [191]; Wolstenholme [306]. | 7.15 | Bell [20]. |
| 6.52 | Eswarathasan and Levine [79]. | 7.16 | Pólya [237, p. 149]; [173, exercise 2.3.4.4–1]. |
| 6.53 | Kaucký [168] treats a special case. | 7.20 | Jungen [167, p. 299] credits A. Hurwitz. |
| 6.54 | Staudt [276]; Clausen [53]; Rado [242]. | 7.22 | Pólya [239]. |
| 6.55 | Andrews and Uchimura [12]. | 7.23 | 1983 homework. |
| 6.56 | 1986 midterm. | 7.24 | Myers [222]; Sedláček [262]. |
| 6.57 | 1984 midterm, suggested by R. W. Floyd.* | 7.25 | [174, Carlitz's proof of lemma 3.3.3B]. |
| 6.58 | [173, exercise 1.2.8–30]; 1982 midterm. | 7.26 | [173, exercise 1.2.8–12]. |
| 6.59 | Burr [42]. | 7.32 | [77, pp. 25–26] credits L. Mirsky and M. Newman. |
| 6.61 | 1976 final exam. | 7.33 | 1971 final exam. |
| 6.62 | Borwein and Borwein [31, §3.7]. | 7.34 | Tomás Feder.* |
| 6.63 | [173, section 1.2.10]; Stanley [275, proposition 1.3.12]. | 7.36 | 1974 final exam. |
| 6.65 | Tanny [286]. | 7.37 | Euler [87, §50]; 1971 final exam. |
| 6.66 | Logan [202']. | 7.38 | 1973 final exam. |
| 6.67 | [175, exercise 6.1–13]. | 7.39 | [173, exercise 1.2.9–18]. |
| 6.70 | Euler [88, part 2, chapter 8]. | 7.41 | André [8]; [175, exercise 5.1.4–22]. |
| 6.72 | [175, exercise 5.1.3–3]. | 7.42 | 1974 final exam. |
| 6.73 | Euler [86, chapters 9 and 10]; Schröter [260]. | 7.44 | Gross [136]; [175, exercise 5.3.1–3]. |
| 6.74 | Logan [202']. | 7.45 | de Bruijn [63]. |
| 6.75 | Comic section, <i>Boston Herald</i> , August 21, 1904. | 7.47 | Waugh and Maxfield [297]. |
| 6.76 | Silverman and Dunn [268]. | 7.48 | 1984 final exam. |
| 6.78 | [183]. | 7.49 | Waterhouse [296']. |
| 6.79 | [126], modulo a numerical error. | 7.50 | Schröder [259]; [173, exercise 2.3.4.4–31]. |
| 6.80 | [174, exercises 4.5.3–2 and 3]. | 7.51 | Fisher [99]; Percus [232, pp. 89–123]; Stanley [274]. |
| 6.81 | Adams and Davison [3]. | 7.52 | Hammersley [146]. |
| 6.84 | Lehmer [198]. | 7.53 | Euler [92, part 2, section 2, chapter 6, §91]. |
| 6.85 | Burr [42]. | 7.54 | Moessner [217]. |
| 6.87 | Part (a) is from Eswarathasan and Levine [79]. | 7.55 | Stanley [273]. |
| 7.2 | [173, exercise 1.2.9–1]. | 7.56 | Euler [91]. |
| | | 7.57 | [77, p. 48] credits P. Erdős and P. Turán. |

- | | | | |
|------|---------------------------------------------------------------------------|------|----------------------------------------------|
| 8.13 | Thomas M. Cover * | 9.29 | de Bruijn [62, section 3.7]. |
| 8.15 | [173, exercise 1.2.10-17]. | 9.32 | 1976 final exam. |
| 8.17 | Patil [228]. | 9.34 | 1973 final exam. |
| 8.24 | John Knuth (age 4) and DEK; 1975 final | 9.35 | 1975 final exam. |
| 8.26 | [173, exercise 1.3.3-18] | 9.36 | 1980 class notes. |
| 8.27 | Fisher [166]. | 9.37 | [174, eq. 4.5.3-21]. |
| 8.29 | Guibas and Odlyzko [138]. | 9.38 | 1977 final exam. |
| 8.32 | 1977 final exam. | 9.39 | 1975 final exam, inspired by Reich [247]. |
| 8.34 | Hardy [149] has an incorrect analysis leading to the opposite conclusion. | 9.40 | 1977 final exam. |
| 8.35 | 1981 final exam. | 9.41 | 1980 final exam. |
| 8.36 | Gardner [113] credits George Sicherman. | 9.42 | 1979 final exam. |
| 8.38 | [174, exercise 3.3.2-10]. | 9.44 | Tricomi and Erdélyi [290]. |
| 8.39 | [177, exercise 4.3(a)]. | 9.46 | de Bruijn [62, §6.3]. |
| 8.41 | Feller [96, exercise IX.33]. | 9.47 | 1980 homework; [175, eq. 5.3.1-34]. |
| 8.43 | [173, sections 1.2.10 and 1.3.3]. | 9.48 | 1980 final exam. |
| 8.44 | 1984 final exam. | 9.49 | 1974 final exam. |
| 8.46 | Feller [96] credits Hugo Steinhaus. | 9.50 | 1984 final exam. |
| 8.47 | 1974 final, suggested by "fringe analysis" of 2-3 trees. | 9.51 | [134, §4.2.1]. |
| 8.48 | 1979 final exam. | 9.52 | Poincaré [235]; Borel [30, p. 27]. |
| 8.49 | Blom [26]; 1984 final exam. | 9.53 | Pólya and Szegő [240, part 1, problem 140]. |
| 8.50 | 1986 final exam. | 9.57 | Andrew M. Odlyzko * |
| 8.51 | 1986 final exam. | 9.58 | Henrici [151, exercise 4.9.8]. |
| 8.53 | Feller [96] credits S. N. Bernstein. | 9.60 | Ilan Vardi.* |
| 8.57 | Lyle Ramshaw.* | 9.62 | Canfield [43]. |
| 8.63 | Guibas and Odlyzko [138]. | 9.63 | Ilan Vardi.* |
| 9.1 | Hardy [148, 1.3(g)]. | 9.65 | M. P. Schützenberger.* |
| 9.2 | Part (c) is from Garfunkel [114]. | 9.66 | Lieb [201]; Stanley [275, exercise 4.37(c)]. |
| 9.3 | [173, exercise 1.2.1.1-6]. | 9.67 | Roas and Wrench [27]. |
| 9.6 | [173, exercise 1.2.11.1-3]. | | |
| 9.8 | Hardy [148, 1.2(iv)]. | | |
| 9.9 | Landau [194, vol. 1, p. 60]. | | |
| 9.14 | [173, exercise 1.2.11.3-6]. | | |
| 9.16 | Knopp [170, edition ≥ 2 , §64C]. | | |
| 9.18 | Bender [21, §3.1]. | | |
| 9.20 | 1971 final exam. | | |
| 9.24 | [134, §4.1.6]. | | |
| 9.27 | Titchmarsh [289]. | | |
| 9.28 | [173, exercise 1.2.11.2-7]. | | |

* Unpublished personal communication.