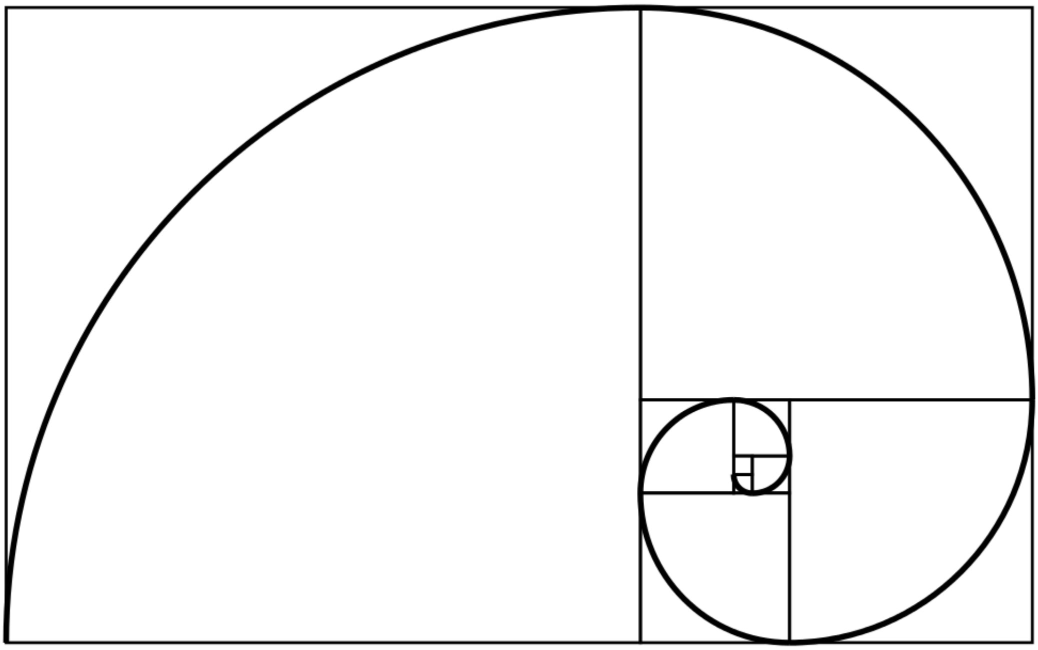
数的旅程



刘新宇 [1]

April 22, 2025

[1] **刘新宇**

Version: $\phi = \dfrac{\sqrt{5}-1}{2} = 0.618$

https://github.com/liuxinyu95

目录

第 1 章　计数

> 一二三四五, 金木水火土。
> 天地分上下, 日月照今古。
>
> 部编小学一年级
> 语文课本第一课

数充满了我们的生活。例如这则新闻报道:“2024 年巴黎奥运会(第 33 届奥运会)已于当地时间 8 月 1 日闭幕。巴黎是继伦敦后的世界第 2 个至少 3 次举办夏奥会的城市。这是首届男女比例完全平衡的奥运会, 男女运动员各为 5250 名。本届奥运会共设有 32 个大项, 329 个小项, 共有 206 个国家和地区参赛, 新增了滑板、冲浪、竞技攀岩和霹雳舞四个大项。中国代表团最终在巴黎奥运会上夺得 40 金 27 银 24 铜的优异成绩。”

这短短的 180 字中有 14 个数字。数是谁发明的? 历史书上没有答案。数出现在所有历史文字记录中, 数也许诞生在史前, 伴随着语言和文字。要找到答案, 我们有两条线索: 1) 追寻古老的历史物证, 石刻、壁画、器物上关于数的印记; 2) 追溯数在语言演变中的痕迹。比如英文中的 eleven (11) 来自古英语 endleofan, 意思是(数到 10 还)剩余 1; twelve (12) 来自 twelf, 意思是剩余 2。

1.1　罗塞塔石碑

走进大英博物馆第 4 展厅, 有一件展品吸引着观众。这是一块残缺的石碑, 长 112 厘米, 宽 76 厘米, 上面刻满了文字, 如图 1.1 所示。人们能一眼辨认出位于底部的内容是希腊字母(见图 1.2a, 表 B.1), 但上面的部分犹如天书。仔细观察会发现余下的文字大致分成两种: 一种是弯弯曲曲的符号, 如图 1.2b 所示, 位于石碑中部; 另一种是奇妙的图案, 如图 1.2c 所示, 位于石碑上部。

这一展品叫做罗塞塔石碑。1799 年, 拿破仑率领法军远征埃及。远征军有 3 万 8 千人, 各类舰只 350 艘, 另有由著名数学家蒙日带队的学者科学家 167 人[1]。这位未来的法兰西皇帝从一名炮兵军官脱颖而出。他敏锐地看出数学不仅可以算出大炮的弹道, 还关乎法兰西的国运。拿破仑的大军巧妙地避开了英国海军的封锁, 7 月 2 日攻占了埃及的亚历山大城, 随即向开罗进军。7 月 15 日, 一位士兵在尼罗河三角洲前线

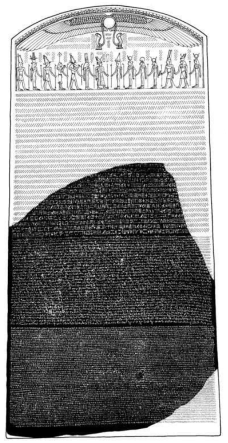
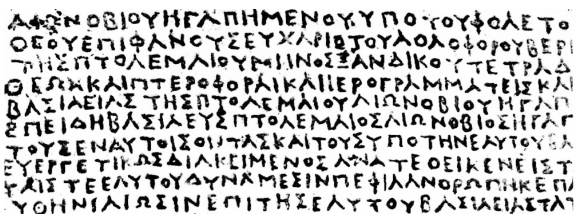
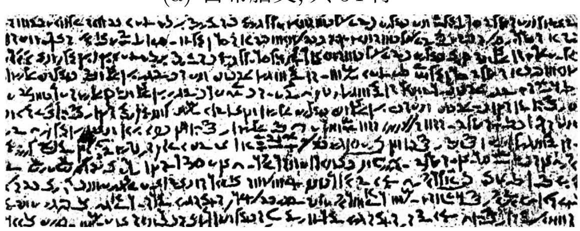


图 1.1: 古埃及罗赛塔石碑。石碑不完整, 从中间断裂。



(a) 古希腊文, 共 54 行



(b) 古埃及世俗文字, 是当时埃及平民使用的文字。共 32 行

(c) 古埃及象形文字, 又称圣书体, 代表献给神明的文字。共 14 行

图 1.2: 罗赛塔石碑上的三种文字

的小镇罗塞塔(Rosetta)挖掘防御工事。他偶然发现了一块断碑砌在一段古老的墙中。尽管并不完整,这仍是一个重大的发现。法军指挥官决定把它送到拿破仑设立于开罗的埃及研究所,并于8月运抵开罗。根据发现地点,它被命名为罗塞塔石碑。1801年,英军击败了拿破仑,罗塞塔石碑也落入了英军之手。1802年2月,它被运抵英国朴次茅斯港,并最终收藏于大英博物馆。

这块石碑有何特殊之处呢?它上面用三种不同的文字记录了同一内容:公元前196年,13岁的国王托勒密5世加冕一周年。他从父亲托勒密4世袭得正统王位,并做了许多善行,如捐助神庙、减免税收等等。埃及当时处于托勒密王朝的统治之下,统治者是希腊人[1]。国王宣称自己是法老。因此铭文使用了古埃及象形文字、古埃及世俗文字、希腊官方文字三种文字镌刻,并颁布全埃及各个神庙勒石立碑。而罗塞塔石碑就是其中之一,并且是至今唯一发现的一块[2]。罗塞塔石碑于是成为了破解古埃及文字的钥匙。英国物理学家托马斯·杨[3]和法国学者商博良通过研究此碑,成功破译了古埃及象形文字[2]。

1.2 古埃及和古巴比伦的计数系统

通过考古发现和对古文字的破译,我们发现古埃及大约在公元前3400年就出现了表示数字的符号。是众多古文明中最早的(美索不达米亚大约在公元前3000年出现了数字符号,中国大约在公元前1600年出现了数字符号[5])。最早的数字符号多是"|"、"||"、"|||"或"—"、"="、"≡"。这表明我们的老祖先是从"数数"开始认识数字的。很可能是采集狩猎到更多的东西。随着生产生活的进步,人们不断认识更大的数,逐渐从一数到了十——显然我们的祖先板着手指头数数。但接下来遇到了困难。人只有两只手十个手指。手指不够用了怎么办?这时有三种解决方法:(1)把脚也用上可以数到二十,但接下来会遇到同样的问题。(2)自由分组。在西伯利亚的尤卡吉尔语中有这样的例子:三和一、两个三、两个四、十差一等等。(3)固定分组。数到某一固定数目,例如十,分成一组,把这个组看成一个单位并起个名字。比如古埃及用 ∩ 表示十。这样 ∩∩ 表示二十(古罗马用X表示十,I表示一。23表示为XXIII)。

为了建造宏伟的金字塔,古埃及人还创造了更大的单位,如图1.3所示。他们用永恒之神(Heh)代表一百万,用蝌蚪表示十万,用弯曲的手指表示一万,莲花表示一千、弯曲的绳子表示一百……计数时按照单位分组,超出后就组成更大的单位。最后把每个单位的数目和大小乘起来。图1.4给出了两个例子。

尽管方法(3)能够处理很大的数,但方法(2)对付小数字也很方便。古巴比伦人生活在两河流域的美索不达米亚平原(位于底格里斯河、幼发拉底河之间的平原,现今的伊拉克境内)。他们在湿泥板上刻划文字,然后在阳光下晒干或用火烘干变硬。这些文

[1] 马其顿国王亚历山大大帝征服了埃及。他死后,埃及总督托勒密一世于公元前305年自称国王建立王朝,统治埃及275年。

[2] 1913年于丹德拉神庙附近出土了一块罗马时期(公元前30年~295年)的三种文字石碑。自上而下为四行不完整的圣书体象形文字、七行世俗体文字、七行希腊文字。类似罗赛塔碑的情况,顶部象形文字残损严重[3]。

[3] 托马斯·杨(1773~1829),就是高中物理课本中杨氏双缝实验——发现光的干涉现象的物理学家。

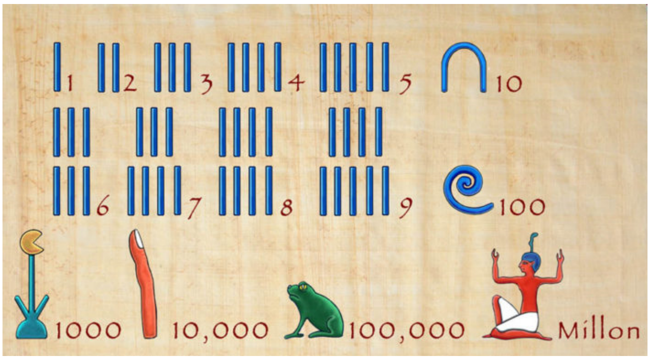


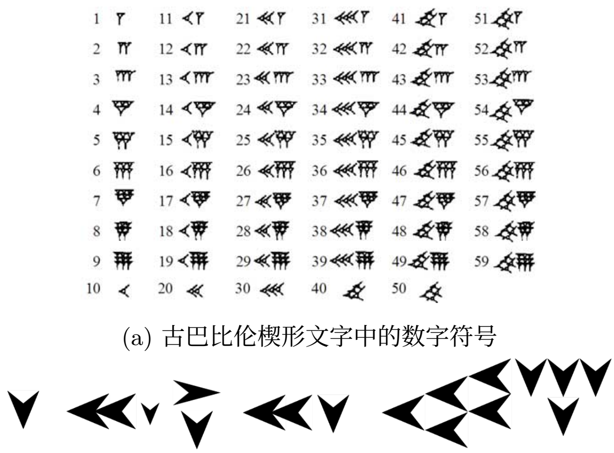
图 1.3: 古埃及象形文字中的数字符号。

(a) 212427 表示为 $2(100000)+1(10000)+2(1000)+4(100)+2(10)+7(1)$

(b) 古埃及埃德富(Edfu)神庙中的象形数字, 表示 1333330

图 1.4: 古埃及象形文字数字的例子

字呈楔形。古巴比伦人使用 60 进制(分组单位是 60)。这种进制也出现在我们的日常生活中。60 秒是 1 分, 60 分是 1 小时。所以 1 小时 12 分 30 秒, 或写成 1:12:30, 包括 $1(60 \times 60) + 12(60) + 30 = 4350$ 秒。60 足够大, 所以古巴比伦人在 60 以内使用自由分组法, 超过 60 用 60^2、60^3、60^4……分组。图 1.5a 是楔形文字 60 以内的符号。可以明显看出其中的规律: 一是一个纵向的楔形, 从一到九每数一就增加一个楔形。十变成了一个横向的楔形, 从十一到十九是一个横向的楔形(10)加上相应纵向的楔形。二十是两个横向的楔形, 然后重复这个规律。60 以内的数字是横向的楔形个数乘以 10 加上纵向的楔形个数。



(a) 古巴比伦楔形文字中的数字符号

(b) 古巴比伦楔形数字 $1(60^3) + 19(60^2) + 21(60) + 54$

图 1.5: 古巴比伦数字

但 60 以上的数字展现出不一样的规律, 如图 1.5b 所示。285714 的 60 进制表示是 1:19:21:54, 但 19 写成了 $20 - 1$。其中较小的纵向楔形表示减法, 减数的上方有一个向右的楔形。可以看出, 这并非是我们现代意义上严格的 60 进制, 而是一种混合进制—— 60 以内用十进制或减法表示。和古埃及相比, 古巴比伦没有为 60、60^2、60^3……创建单独的符号, 而是依赖数字所在的位置决定单位的大小。最左边的数字表示有多少个 1, 向右的第二个位置上的数字表示有多少个 60, 第三个位置上的数字表示有多少个 60^2, 以此类推。这种方法叫做“位值制系统”。但由于缺乏 0, 古巴比伦的数字有歧义。直到约公元前 300 年后才偶尔出现了 0, 但只用于数字的中间而非结尾, 无法区分出 11 和 1100. 古文字学者至今仍面临这个问题, 只能通过上下文推测数值。

1.3 古罗马和中国计数系统

随着古罗马文明的崛起并建立起横跨欧亚非的帝国, 罗马计数系统产生了巨大的影响, 直到今天我们仍然可以在钟表盘上、建筑上、著作的章节编号上看到罗马数字(见图 1.6)。罗马数字的优点是符号简单, 只包含 I、V、X、C、M 几个符号, 对于较小的数字很直观。例如 III 表示 3, VII 表示 $5 + 2 = 7$, 2025 可表示为 MMXXV (2 个

1000, 2 个 10 加 5)。但是罗马数字的缺点是用法混乱。早期的罗马数字只包含加法,后来出现了“左减右加”的用法, 如 IV 表示 5 - 1 = 4, 而 VI 表示 5 + 1 = 6。这样 IXX = (10 - 1) + 10 = 19, 但 19 = 10 + (10 - 1) = XIX 就有歧义了: XIX 还可以理解为 (XI)X = 11 + 10 = 21。同样 XVIII = 16, 但 16 = 8 + 8 = IIXIIX 和 16 = (10 - 4) + 10 = IIIIXX 都有歧义。其中 IIIIXX 出现在 1388 年巴黎的一份协定中表示 88[4]。

1	5	10	100	1000
I	V	X	C	M

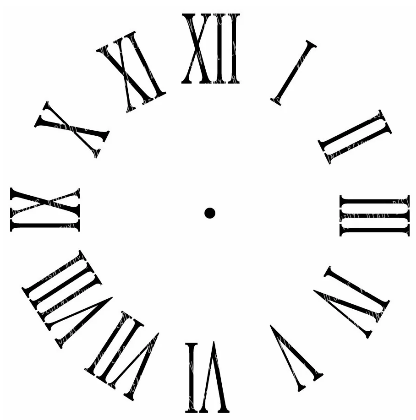


图 1.6: 钟表盘上的罗马数字

在古代中国产生了非常接近现代计数系统的十进制“乘法分组”计数法。汉字的前 10 个数字符号是:

一 二 三 四 五 六 七 八 九 十

接下来中文没有出现英文中 eleven(剩一)和 twelve(剩二)的问题, 而是直接进展到十一、十二……十九、二十……九十九。数字较小时, 也使用廿(20)和卅(30)这样的单位, 通常用在日历中。然后增加了百、千、万的分组单位。例如 6147 写为六千一百四十七, 表示 6 个千、1 个百、4 个十、7 个一, 即 6 千 1 百 4 十 7 个, 换成罗马单位相当于 6M1C4X7I。每个数字乘以单位的大小, 然后加到一起就是数值。这看起来和我们日常生活中的数字几乎一样了, 可是几乎毕竟是几乎, 究竟还差了一些。比如六千一百七, 究竟是 6170 还是 6107? 汉字中不是有“零”么? 可是在公元十世纪前的古代中国, 零从来没有用在数学上。零是一个形声字, 雨表形, 令表声。本意是下雨。例如《诗经·豳风·东山》: 我来自东, 零雨其濛。引申意是落下, 例如《诗经·郑风·野有蔓草》: 野有蔓草, 零露漙(tuán)兮。《楚辞·离骚》: 惟草木之零落兮, 恐美人之迟暮。在可能成书于东汉时期的《九章算术》中, 没有出现任何含有汉字零的数字。即使在宋元以后引入了零, 用法也不一致。例如 103 可以写成一百零三, 也可以写成一百有三。《水浒传》里

的好汉数目常读作一百单八将。北京土话中用“二百五”形容一个人愚钝,指的是 250。在没有零的情况下,要想避免歧义就必须明确每个数字所代表的单位。这就需要给不同大小的分组单位命名或创造符号。随着数字的增大就需要更多的单位。但数的增加是无限的,而符号的数目是有限的,早晚有用光的时候。下表列出了中文和英文中越来越大的单位名称。

百	100	秭 (zǐ)	10^{24}	恒河沙	10^{52}
千	1000	穰 (ráng)	10^{28}	阿僧祇 (zhī)	10^{56}
万	10000	沟	10^{32}	那由他	10^{60}
亿	10^8	涧	10^{36}	不可思议	10^{64}
兆	10^{12}	正	10^{40}	无量大数	10^{68}
京	10^{16}	载	10^{44}		
垓 (gāi)	10^{20}	极	10^{48}		

可以看到,汉语中这些大单位词汇,有许多来自佛教[6]。包括恒河沙,它表示 1 后面跟着 52 个 0。英语中的大单位如下表。从一开始,每增加一千倍就有一个对应的单位。万进位和千进位的不同,也是文化上的一种差异。

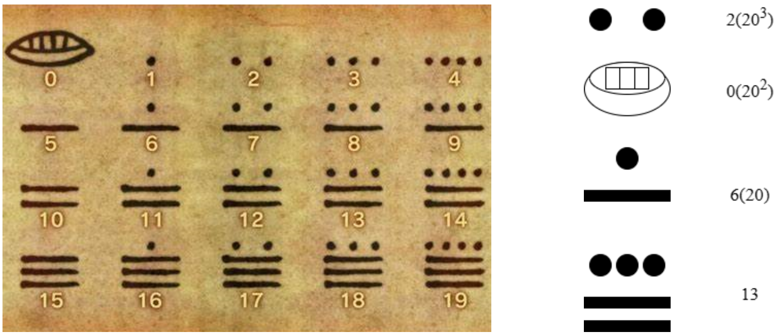
thousand	10^3	quattuordecillion	10^{45}	octovigintillion	10^{87}
million	10^6	quindecillion	10^{48}	novemvigintillion	10^{90}
billion	10^9	sexdecillion	10^{51}	trigintillion	10^{93}
trillion	10^{12}	septdecillion	10^{54}	untrigintillion	10^{96}
quadrillion	10^{15}	octodecillion	10^{57}	duotrigintillion	10^{99}
quintillion	10^{18}	novemdecillion	10^{60}	googol	10^{100}
sexillion	10^{21}	vigintillion	10^{63}		
septillion	10^{24}	unvigintillion	10^{66}		
octillion	10^{27}	duovigintillion	10^{69}		
noniliion	10^{30}	trevigintillion	10^{72}		
decillion	10^{33}	quattuorvigintillion	10^{75}		
undecillion	10^{36}	quinvigintillion	10^{78}		
duodecillion	10^{39}	sexvigintillion	10^{81}		
tredecillion	10^{42}	seprvigintillion	10^{84}		

表中最后一个大单位古格尔(googol)是在 1920 年由 9 岁的米尔顿 · 西洛塔(Milton Sirotta)想出的名字。这个数字是 1 后面跟着 100 个零。著名的互联网公司谷歌的名字就来自它。

1.4 位值制计数系统

在大航海时代,西班牙探险者在中美洲的尤卡坦半岛(位于墨西哥湾和加勒比海之间)发现玛雅文明创造出了完美的计数系统。图 1.7 给出了玛雅数字。类似古巴比

伦的 10-60 混合进制，玛雅人使用 5-20 混合进制。20 以内采用 5 进制，每数 1 画一个点，每数到 5 划一横线。20 以上的数竖着写，并规范使用 0 做占位符，如图 1.7 所示。玛雅人的计数系统和玛雅历法相互影响。以地球围绕太阳旋转一周作为一个太阳年，每月 20 天。玛雅人很快发现 $20 \times 18 = 360$，接近一年。于是规定 1 年 18 个月，在加上 5 个禁忌日，一年恰好是 365 天。每 4 年再加上 1 天。这和我们现今太阳年的天数几乎一样。



(a) 20 以内采用 5 进制　　(b) $2(20^3) + 0(20^2) + 6(20) + 13 = 16133$

图 1.7: 玛雅数字

无论是玛雅的 20 进制还是古巴比伦的 60 进制，都不需要额外命名单位，而是通过数字所在位置赋予大小的意义。这种方法叫做位值制计数系统（positional numeral system），归纳起来有 3 个要求：

1) 确定一个进制 [4] b。如我们日常使用的 $b = 10$、古巴比伦使用的 $b = 60$、玛雅使用的 $b = 20$。

2) 给 1、2……$b - 1$ 这些数字命名。例如汉语的一、二、三……九；英语的 one, two, three, ..., nine；玛雅的点、点点……四点三划。

3) 代表 0 的符号。

这样任何一个数 n 等于：

$$n = a_m b^m + a_{m-1} b^{m-1} + \cdots + a_a b + a_0 \tag{1.1}$$

其中 a_i 是被命名的数字，包括 0、1、2……$b - 1$。n 写为 $a_m a_{m-1} \cdots a_0$。例如 $2024 = 2 \times 10^3 + 0 \times 10^2 + 2 \times 10 + 4$，其中 $b = 10, m = 3, a_3 = 2, a_2 = 0, a_1 = 2, a_0 = 4$，恰好也写为 2024。注意 0 在这里起到了关键的“占位”作用——总共有 2 个千、2 个十、4 个一，但是没有百。0 占据在百的位置上，这样其它数字才能“各就各位”，避免出现 224 这样的问题。我们再看一正一反两个例子：

[4] 英文 base 的首字母

例 1.4.1. 猫咪王国中, 每只猫爪只有 *4* 个可动的指(见图 *1.8*), 左右共 *8* 个可动的指。如果采用 *8* 进制, 并规定 *0* 的符号为“喵”, *1* 到 *7* 为: (*1*) 苗、(*2*) 秒、(*3*) 妙、(*4*) 咪、(*5*) 谜、(*6*) 米、(*7*) 蜜。王国一年的天数 $365 = 5 \times 64 + 5 \times 8 + 5$, 写成“谜谜谜”。$2024 = 5 \times 8^3 + 3 \times 8^2 + 7 \times 8 + 0$, 写成“谜妙蜜喵”。

图 1.8: 猫爪

例 1.4.2. 反例: 我国古代用天干地支纪年。天干有 *10* 个符号: 甲、乙、丙、丁、戊、己、庚、辛、壬、癸; 地支有 *12* 个符号: 子、丑、寅、卯、辰、巳、午、未、申、酉、戌、亥。组合起来可以记录 *60* 以内的数字, 从甲子开始, 接下来天干地支各向前进一到乙丑, 然后是丙寅……到癸亥结束(见图 *1.9*)。比如 *2025* 年是农历乙巳年。乙的上一个天干符号是甲, 巳的上一个地支符号是辰, 所以前一年 *2024* 年是甲辰年, 后一年 *2026* 年是丙午年。尽管我们不说“丙天午地”, 看似用位置表示(第一个符号是天干、第二个符号是地支), 但干支纪年不是位值制计数法。例如丙午并非一个甲子中的第 $3 \times 10 + 7 = 37$ 年, 而是第 *43* 年 [5]。

罗马数字的一个问题就是歧义 —— 一个数字有多种写法或一个写法可以代表不同的数字。位值制能解决这个问题么? 明显 $8 = 08 = 008$。我们在电梯、数字钟, 汽车里程表上都见过类似的情况。可以在一个数字前放任意多个零而不改变值。如果对最高位的数字加以限制, 规定式 (1.1) 中的 $a_m \neq 0$, 那么一个数的位值制表示是唯一的(见练习 1.4)。

在历史的长河中, 玛雅人和古巴比伦人的计数系统都湮没了。20 和 60 以内的数字并非单一符号, 混合进制难于计算。现代位值制系统是印度文明发展出并经由阿拉伯人传入西方的, 被称作“印度 —— 阿拉伯系统”(Hindu-Arabic system)。在公元前三世纪, 印度的碑铭中出现了 1、4、6 这样的符号; 在公元一、二世纪的岩洞中, 形如 2、3、4、5、6、7、9 的符号也出现了。可以看出 2 从二、3 从三中演化的迹象, 在公元七世纪, 印度数学家婆罗摩笈多定义了零[8]。但直到公元九世纪以前, 符号 0 仍未出现(见图 1.10)。印度人首先用一个点或小圆圈表示零, 命名为 sunya, 梵文的意思是“空位”。传入阿拉伯后译为 sifr, 意思是“保持完整”。1140 年, 它伴随着阿拉伯学者花拉子密的著作被音译为拉丁文, 最后变为今天英语中的 zero。

---

[5] 除了查看图 1.9 外, 也可以这样计算: 天干每 10 年循环, 地支每 12 年循环。丙是天干中的第 3 个符号, 午是地支中的第 7 个符号, 所以丙午年除以 10 余 3 且除以 12 余 7。60 中除以 10 余 3 的数字有 3、13……43、53, 其中只有 43 除以 12 余 7。

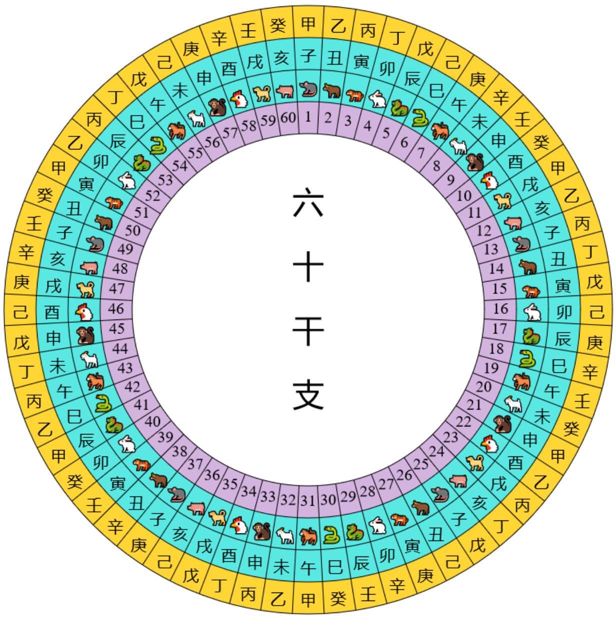


图 1.9: 干支纪年

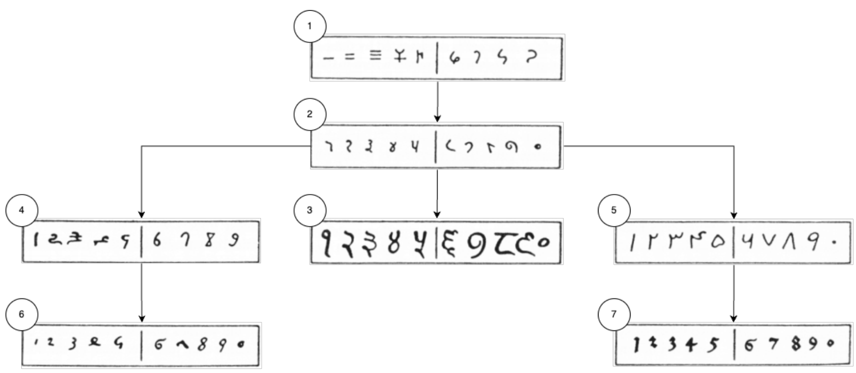


图 1.10: 印度——阿拉伯数字的演进:1、婆罗门数字, 公元一世纪;2、印度瓜廖尔数字包括 0, 公元九世纪;3、4、5、梵文天成体、东、西阿拉伯数字, 公元十一世纪, 其中西阿拉伯数字至今仍在土耳其使用;6、7、十五和十六世纪。

图 1.11: 前苏联 1983 年发行的花拉子密纪念邮票

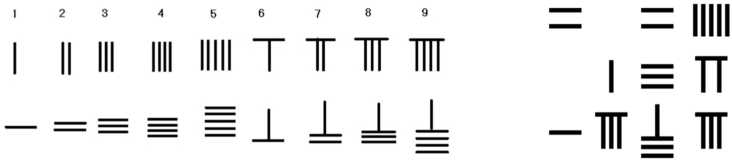
公元七世纪初, 阿拉伯文明崛起。罗马帝国分裂后, 古希腊、拜占庭、印度的数学成果逐渐在八世纪汇集到阿拉伯帝国。阿拔斯王朝于公元 830 年在巴格达设立了著名的“智慧宫”, 聚集了大批学者研究整理来自世界各地的天文地理资料和学术成果。公元 813 年, 著名学者阿尔·花拉子密(al-Khwārizmī, 约 780～约 850, 见图 1.11)来到巴格达, 并成为智慧宫的主要学者之一。关于花拉子密的生平, 我们知之甚少。甚至他的名字可能只是个“绰号”。阿拉伯文 al 是“来自”的意思, Khwārizmī 是地名“花剌子模”, 位于今天中亚的乌兹别克斯坦境内。金庸先生的小说《射雕英雄传》中有黄蓉帮着成吉思汗用计攻下花剌子模的故事, 说的就是这个地方。花拉子密在数学方面完成了两部传世之作:《代数学》(成书于 820 年)和《印度数字算术》(成书于 825 年)。此外他还完成了地理和天文方面的著作。《代数学》的拉丁文译名为 Hisab al-jabr w'al-muqabala, 意思是还原与对消的科学。其中 al-jabr 后来演变为英文 algebra。1859 年清代学者李善兰与传教士伟烈亚力创造性地把它译为中文“代数”[7]。花拉子密在书中系统地给出了一元一次、二次方程的代数解法。但是他所在的时代还没有代数符号系统, 因此完全使用文字来描述问题和解题步骤, 并附加上几何解释来说明正确性。《印度数字算术》在十二世纪被翻译为拉丁文(Liber Algorismi de numero Indorum)并传入欧洲, 促进了印度——阿拉伯十进制计数系统的传播。花拉子密名字的拉丁音译 Algorismi 后来成了“算法”一词 Algorithm[9]。

十进制位值制系统带来了巨大优势, 便于计算并可以处理任意大的数字。我们可以想象用罗马数字计算 MMXXV - CXXXVII 有多么困难, 而小学生也可以用竖式快速算出结果:

$$\begin{array}{r} 2\ 0\ 2\ 5 \\ -\quad 1\ 3\ 7 \\ \hline 1\ 8\ 8\ 8 \end{array}$$

这仅仅是减法, 更不用说乘除法了。尽管古汉语中的计数系统并非位值制(见上一节), 古代中国却发展出了一套严格意义上的位值制计算方法——筹算。筹算使用

若干小棍进行计算。这些小棍用竹子制成叫做算筹[6]。它的发明是一个漫长的过程，但春秋战国时期已经普遍使用了。成语"运筹帷幄"、"一筹莫展"、"技高一筹"都反映了算筹的使用情况。算筹有横竖两种摆法，如图 1.12a 所示。如果个位竖着摆，则十位横着摆。这样交错摆放避免混淆。如果某一位是零，则空出不摆放[7]。算筹进行竖式计算非常方便，尤其是处理进位和借位。例如进位时只要把一个小棍摆到下一位就可以了。如图 1.12b 所示。



(a) 算筹数字的横竖摆法 (b) 算筹计算 2025 - 137

图 1.12: 算筹

花拉子密的著作一开始只有少数学者了解。1202 年斐波那契通过他的著作《算盘书》(Liber Abaci) 进一步把印度 —— 阿拉伯系统和计算方法介绍到欧洲。无独有偶，"斐波那契"其实也是个绰号。这位中世纪的数学家叫列奥纳多，来自比萨，所以称作比萨的列奥纳多。斐波那契来自拉丁文 filius Bonacci，意思是波那契之子。斐波那契的父亲当时是商人，在北非以及地中海一带经商。斐波那契逐渐从埃及人、希腊人、西西里人、阿拉伯人那里学到了各种数字系统。他逐渐认识到了印度 —— 阿拉伯数字的先进性。《算盘书》不仅包括位值制的记法，还讲解了如何计算利润率、利息、汇率、度量衡换算等实际问题。一经面试大受欢迎，被广泛传抄。神圣罗马帝国皇帝腓特烈二世听说了斐波那契，在 1220 年邀请他到比萨。斐波那契接受了宫廷数学家的挑战，成功解决了所有题目，包括一道一元三次方程。他采用的数值解法给出了 9 位小数精度[10]。今天，以他命名的"斐波那契数列"，即 1、1、2、3、5、8……已经家喻户晓。我们将在第 5 章详细介绍这个有趣的数列。

图 1.13: 比萨的列奥纳多(斐波那契)1175-1250

[6]形声字"筹"的形旁是竹。也有用木头、兽骨、象牙、金属等材料的。

[7]在计算过程中空位可能不明显造成混淆。后来人们在书写记录算筹运算结果时，就在应该是零的数位上写"□"或"○"。

x	$1-x$
0	1
1	0

x	非 x
假	真
真	假

x	y	$x \times y$
0	0	0
0	1	0
1	0	0
1	1	1

x	y	x 与 y
假	假	假
假	真	假
真	假	假
真	真	真

图 1.14: 二进制 $1-x$ 与逻辑非的对应、乘法和逻辑与的对应

越来越多的学者和在各国往来贸易的商人把印度——阿拉伯计数系统传播到了全球。它成了全人类的通用语言。十七世纪时，德国数学家莱布尼茨提出了二进制，它导致了计数系统在二十世纪产生了革命性的应用。电子计算机内部使用二进制对数字和信息编码。尽管二进制只有 0、1 两种数字，表示很长，例如 2025 写成二进制是 1111 1101 001，但 0、1 两种状态可以通过电路的通、断或电平（电压）的高、低方便地实现。还有一个好处：0、1 可以对应逻辑的真、假；事物的有、无；答案的对、错，从而方便用机器统一处理数和逻辑，如图 1.14。

我们的祖先从狩猎、采集、计数开始，经过了四千余年才发展出了现代十进制位值制计数系统。数的诞生实在是一个漫长曲折的过程。它没有一个确定的发明者，而是在各个文明中独立产生于生产劳动，互相影响，逐渐发展，是人类群体性的进步。

练习 1

1.1. 2025 年深圳市南山区小学 4 年级期末数学考试中出现了这样一道题：“计算 114×21，同学们用了不同的方法。在思路上这些方法有什么相同的地方？”

①

$$114 \times 20 = 2280$$
$$114 \times 1 = 114$$
$$2280 + 114 = 2394$$

②

$$\begin{aligned} & 114 \times 21 \\ =& 114 \times 7 \times 3 \\ =& 798 \times 3 \\ =& 2394 \end{aligned}$$

③

×	100	10	4
20	2000	200	80
1	100	10	4

$$\begin{array}{r} 2\,2\,8\,0 \\ +\quad 1\,1\,4 \\ \hline 2\,3\,9\,4 \end{array}$$

④

$$\begin{array}{rl} 1\,1\,4 & \\ \times \quad 2\,1 & \\ \hline 1\,1\,4 & \cdots\ 114 \times 1 = 114 \\ 2\,2\,8\,\cdot & \cdots\ 114 \times 20 = 2280 \\ \hline 2\,3\,9\,4 & \end{array}$$

这四个方法中，哪些利用了十进制位值制计数系统的优势？

1.2. 可以利用 $n = a_m b^m + a_{m-1} b^{m-1} + \cdots + a_a b + a_0$ 进行进制转换：

1) 用 n 除以 b 得到余数 a_0、商 q_0。

2) 把 n 替换为 q_0, 再次用 n 除以 b 得到余数 a_1、商 q_1。

3) 把 n 替换为 q_1, 重复上述步骤直到商 $q_m = 0$。

则 n 的 b 进制表示为 $a_m \cdots a_1 a_0$。请将十进制数 123 转换为(a)玛雅文;(b)古巴比伦文;(c)计算机二进制表示。

1.3. ★[8] 如果你会编程, 请用自己熟悉的语言实现二进制到十进制的相互转换。

1.4. ★ 证明一个数的位值制表示是唯一的。提示: 考虑如果一个数有两种表示会怎样?

答案: 61 页

[8]标星号的题目较难。

第 2 章　零

> 因此"有"与"无"的真理,就是两者的统一。
>
> 黑格尔《小逻辑》

尽管数有超过 5000 年的历史, 零却只有 1200 年的历史, 是个"新生事物"。零诞生于印度,经由阿拉伯人引入欧洲。但它仅仅具有"占位符"的特殊身份,和 1、2、3……比起来是个"二等公民",甚至面临被"开除数籍"的风险。

2.1　质疑与否定

零一出生就代表"无"、"没有"。印度人给它起名叫 sunya, 意为"空位"。所以 2025 中的 0 表示没有百。"无 [1]"在西方的文化传统、宗教哲学上是负面否定的。它往往和黑暗、虚空、死亡等意象关联。而 1 代表有一个, 2 代表有两个……人们不说 0, 而说没有, 表示对有、生存、实在的否定。莎士比亚《王子复仇记》中的名句"To be or not to be, that is the question."译为"生存还是死亡?"人们说 3 只羊跑了 1 只羊, 还有两只羊;如果再跑了两只羊, 人们说没有羊, 而不说有 0 只羊。

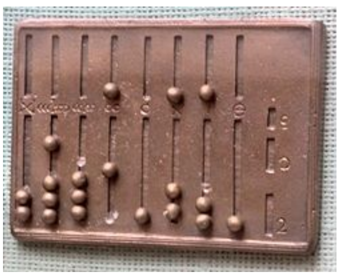


图 2.1: 现代复原的罗马算盘, 在盘上开槽, 槽中摆放石子, 槽上标有罗马数字单位 I, V, X, C, M。

印度-阿拉伯计数系统传入欧洲后, 尽管计算方便, 人们还是把计算结果转换成罗

[1] 对应的英文是 void, 表示虚空。

马数字,而避免使用零[2]。教会宣布 0 是邪恶的符号,禁止在公开场合使用。僧侣学者们用算盘(见图 2.1, 不是中国明代发明的珠算)计算。欧洲的算盘源自古巴比伦的泥板。就是在一块板上用小石子按照罗马数字演算,算好后抄下结果。但商人和银行家们看到了印度-阿拉伯数字的巨大好处, 于是产生了十六世纪的"人机大战", 如图 2.2 所示。结果可想而知。自然是代表印度 —— 阿拉伯计数的"机"胜了。这种"人机大战"在历史上一再上演: 斯蒂芬森的火箭号列车[3]和马赛跑; 深蓝挑战国际象棋大师卡斯帕罗夫; Alpha-Go 挑战九段棋手李世石、柯洁; ChatGPT 挑战大学生入学考试……每当有新事物出现,就有精彩的人机大战。

图 2.2: 版画"算术", 描绘了一个人用印度 —— 阿拉伯数字计算, 另一个人用罗马算盘计算的比赛。出自格雷戈尔·赖施 1503 年的《哲学之珠》。

"机"的胜利让 0 具有了"二等公民"的身份 —— 可以被使用了。可是为何 0 不能像 1、2、3……那样成为"一等公民"呢?

2.2 真正的数

要想成为一等公民, 0 和 1、2、3……比差了什么呢? 是值, 或者说大小。1 这个符号代表的大小是 1, 比如 1 厘米的长度、1 克的质量、1 只羊……同样, 2 的值可以代表 2 厘米的长度、2 克质量、2 只羊……可是 0 呢? 它没有值。"0 是个占位符, 不是一个数, 因为它没有值[11]。"0 极特殊, 它表现得像一个破坏者, 而非 1、2、3……那样的"良民"。我们看看 0 是如何"破坏规矩"的:

1) 阿基米德公理。一个数[4]加上自己越来越大。阿基米德甚至指出这个规律是一个公理, 今天叫做阿基米德公理: 任意非零的 $m < n$, 则反复 $m + m + \cdots$ 将会超过

[2]古代中国在进行算筹计算时(见第 1.4 节)用空位代表 0。但计算完成后就把结果用乘法分组系统表示出来, 如一百、两千一十五、一百有二。汉字"零"直到宋、元后才出现。

[3]英国工程师乔治·斯提芬森和其子罗伯特·斯蒂芬森于 1830 年设计制造的蒸汽机车。

[4]本节的数都是自然数。

n。

$$\begin{array}{lll} 1+1=2 & 1+1+1=3 & \dots \\ 2+2=4 & 2+2+2=6 & \dots \\ 3+3=6 & 3+3+3=9 & \dots \\ \dots & \dots & \end{array}$$

但 $0+0=0$ 没有变大, $0+0+0+\cdots$ 不会超过任何数, 包括 1。

2) 积与面积。各个文明在丈量土地时都独立发展出了面积的概念, 并将面积和乘法联系了起来。一块长方形的土地面积就是长与宽的积。$1\times 5=5$ 代表宽 1 长 5 的矩形面积; $2\times 5=10$ 代表宽 2 长 5 的矩形面积; $3\times 5=15$ 代表宽 3 长 5 的矩形面积是……但是 0×5 呢? 矩形不见了, "坍缩"成了一个点[5], 如图 2.3 所示, 矩形消失了。

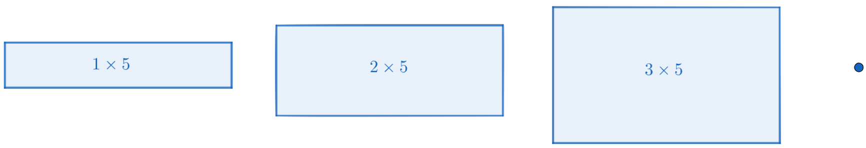


图 2.3: 0×5 把矩形"压缩"成了一个没有大小的点。

乘法还表示缩放。把一把尺子放大 2 倍, 不仅尺子长度乘以 2, 尺子上的刻度间隔也乘以 2; 同样把尺子放大 3 倍, 尺子的长度乘以 3, 刻度间隔也乘以 3, 如图 2.4 所示。相反, 把图中最大的尺子缩小 3 倍, 即乘以 $\frac{1}{3}$, 则得到图中最小的尺子。但是如果缩放 0 倍(乘以 0)整条尺子连同上面的刻度都"坍缩"成一点。尺子消失了。

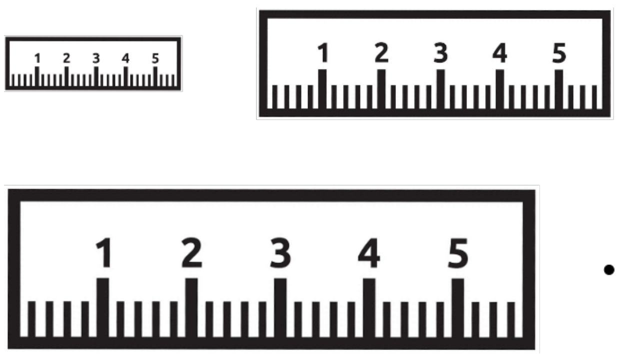


图 2.4: 缩放 0 倍把尺子连同刻度都"压缩"成了一个没有大小的点。

尺子代表了线段, 乘以 0 破坏了矩形和线段。在古希腊发展几何学的时候还没有零

[5]按欧几里得的定义, 线段是没有宽度的, 为什么矩形不是退化成了一条长 5 的线段? 如果我们考虑长度的单位, 例如厘米, 则面积的单位是平方厘米。$0\times 5=0$ 平方厘米, 而线段只有长度, 单位是 5 厘米。0 平方厘米 $\neq$ 5 厘米。

的概念。古希腊的计数系统中也没有零。欧几里得定义“**点**是没有部分的”，“**线**只有长度而没有宽度”[6]。0 对应几何学中的一个点么？

3) 乘除法的可逆性。除法被定义为乘法的逆运算。既然 $2\times 3=6$ 则 $6\div 2=3$；既然 $3\times 4=12$ 则 $12\div 3=4$……据此 $2\times 0=0$ 则 $0\div 2=0$，并且 $a\times 0=0$ 故 $0\div a=0$。但是 0 可以作除数么？根据逆运算，除法 $a\div b$ 相当于求 c 使得 $bc=a$，则 $a\div 0$ 相当于求一个数 x 使得 $0x=a$。但 $0x=0$，若 $a\neq 0$，则不存在任何数 x 满足这个要求，因为：$0x=0\neq a$；若 $a=0$，虽然 $0\times 0=0$，看似 $0\div 0=0$；但 $1\times 0=0$，这样 $0\div 0=1$ 看似也可以；同样 $2\times 0=0$，故似乎 $0\div 0=2$……既然任何 $a\times 0=0$ 则 $0\div 0=a$，即 $0\div 0$ 等于任何值。但这意味着 $(0\div 0)=0=1=2=3=\cdots$，代表“无”的 0，等于代表“有”的 1、2、3……无即是有，空即是色。逻辑学家罗素指出从一个假命题可以推出任何命题。如果 $0=1$，我们可以推出 $1+1=3$：

$$\begin{aligned} 0=1 &\Rightarrow 1=2 && \text{两边}+1 \\ &\Rightarrow 1+1=3 && \text{把}1=1\text{加到两边} \end{aligned}$$

伊恩·斯图尔特给出了这样一个荒唐的例子[12]：

一只猫有一只尾巴；

没有猫有八只尾巴。

把这两句话加起来：一只猫有九只尾巴。

鉴于 0 的这些奇特的“破坏”性质，人们把 0 归为 1、2、3……之外的异类，如图 2.5 所示，0 被放到了 1 到 9 的后面。

图 2.5: 0 在键盘上的位置

[6] 欧几里得《几何原本》定义 1、2。

2.2.1 序数

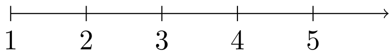
除了表示大小外,数还表示顺序。1、2、3……中,1 在 2 的前面,3 在 2 的右边。随着人们数数自然而然地感知到了顺序。迈克尔·阿廷在记下了他和女儿的对话[13]:

"一、二、三、五、四……"

"不对,爸爸,应该是一、二、三、四、五"

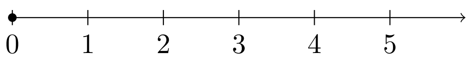
"是的,但为什么我不能说一、二、三、五、四呢?"

"事情不是这样的。"



把数字按照顺序排成一列叫做数轴。数的这种含义叫"序数"(ordinal)。在 0 引入西方前,制定历法的人不知道 0 的概念。于是规定:耶稣基督诞生后的年份为公元 1 年 [7], 以后顺序为公元 2 年、公元 3 年……而之前的一年叫做公元前 1 年, 再前一年叫公元前 2 年……注意这里没有零。这带来了一个有趣的问题, 例如公元前 202 年 5 月,刘邦称帝建立西汉。到了公元 9 年正月王莽篡位西汉灭亡。那么西汉王朝有多长? 如果用 $202 + 9 = 211$ 年计算则多了 1 年。为了看出问题, 我们只要考虑公元前 1 年 12 月 25 日出生的婴儿到了公元 1 年 11 月 25 日时几岁? 当然不满 1 周岁, 只有 11 个月大。那到了公元 1 年 12 月 31 日呢? 有 1 岁了, 但绝不可能是 $1 + 1 = 2$ 岁。如果一个公元 1 年生的婴儿到了公元 2 年生日时几岁呢? $2 - 1 = 1$ 岁。一般公元 x 年到公元 y 年过了 $y - x$ 年。那么 1999 年 12 月 31 日晚上 11:59 分 50 秒时, 全世界都在倒数进入新千年。可严格说来公元 2000 年距公元 1 年过了 $2000 - 1 = 1999$ 年。全世界提前庆祝了 1 年。为什么会有这个错误? 因为跳过了公元 0 年(没有公元 0 年)。如果你没想明白, 请考虑这个例子:午夜 11 点 59 分 59 秒后的 1 秒, 标志着新的一天的开始。此时是 0 时 0 分 0 秒, 新一天的第一秒是 0 时 0 分 1 秒。新的一天始于 0 时而不是 1 时。请仔细看看钟表盘。

我们看看温度计, 如图 2.6 所示。一个大气压下冰水混合物的温度被定义为 0°C;水沸腾的温度时 100°C。之间均匀分成 100 份。1 格代表 1°C。这样温度计上的 3 对应的大小是 3°C、4 对应 4°C……x 对应 x°C。零下 5°C 到 10°C 间差了 15°C, 避免了公元纪年出现的问题。如果把温度计横过来, 我们就得到了一个正常的数轴。0 处于 1、2、3……的左侧。在这里 0 的含义不再是"无", 而是"起始"、"原点"。



一个数 a 在 b 的左边, 表明 a 在 b 之前, 即 $a < b$, 它们距离 $b - a$。反之, a 在 b 的右边, 表明 a 在 b 之后, 即 $a > b$, 它们距离 $a - b$[8]。从运动的观点看 $a + d$ 表示向前(右)移动 d, $a - d$ 表示向后(左)移动 d。如果 $d = 0$, 则 $a + 0 = a - 0 = a$, 表示固定 a

[7] 耶稣生于 12 月 25 日, 即西方的圣诞节。要到接下来的 1 月 1 日才算公元 1 年。

[8] 从节约一致的观点出发, 我们可只用 $a < b$ 定义"前", 用它的否定定义另一种情况。

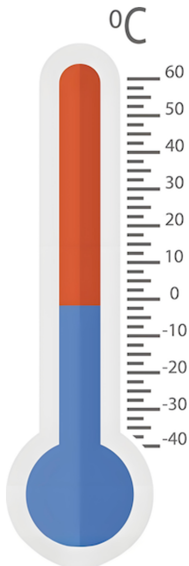


图 2.6: 温度计相当于一个垂直的数轴

不动。在序数的意义下, 0 被赋予了合法的地位, 不再是"二等公民", 而与 1、2、3……一样是一个序数。

在计算机系统中, 有一种叫做数组(array)的数据结构。它使用一段连续的空间存储一组值, 比如:

```
int a[10];
```

声明一个数组 a, 它连续存储 10 个整数。在计算机内部, 一个整数占 16 个二进制位, 即用 16 个 bit 表示一个整数。对于正整数, 最大可以表示到 $2^{16}-1=65535$, 即 16 个二进制位都是 1。例如 314 写成二进制 100111010(=256 + 32 + 16 + 8 + 2), 即第 9、6、5、4、2 位是 1, 其它 11 位是 0(见第 1.4 节)。每 8 个二进制位称为一个字节(Byte), 所以一个 16 位整数占用 2 个字节, 记为: `sizeof(int) = 2B = 16`。一个容纳 10 个整数的数组占用 $2\times 10=20$ 字节的存储空间。

计算机执行到 `int a[10]` 时, 就会在系统存储器 [9] 中找到一个位置, 称为地址, 让 `a` 指向它。从这之后预留出连续 20 个字节来, 供程序使用。接下来程序用这个地址索引数组的各个元素。比如从 `a` 开始两个字节保存首个元素, `a` 后面第 3、4 字节保存次一个元素, 如图 2.7 所示把 1、2……9、10 存入了数组。

注意首个元素的地址从 `a` 开始, 它与地址 `a` 的距离为 0, 记作 `a[0]`, 次一个元素的地址从 `a + sizeof(int) = a + 16` 开始, 记作 `a[1]`, 下一个地址从 `a + 2 * sizeof(int)` 开始 [10], 记作 `a[2]`……计算机从 `a + i * sizeof(int)` 找到 `a[i]`, 所以数组中的 10 个元素从 0 开始索引, 依次是 `a[0], a[1], ..., a[9]`。很多人初学编程不知为何数组从 0 开始到 $n-1$ 结束。其原因即在于此: 数组的下标 i 表示了元素 `a[i]` 到数组起始位置的距

[9] 有两种情况: 栈和堆。这个例子使用栈。

[10] 计算机系统用"*"表示乘法

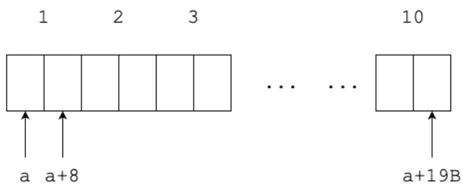


图 2.7: 数组

离。距离从 0 开始, 0 才是真正的起点。以下例子程序将 1 到 10 装入数组:

```
int a[10];
for (int i = 0; i < 10; ++i) { a[i] = i + 1; }
```

有一个程序员自嘲的笑话。某程序员经过痛苦的熬夜加班, 终得休息, 决心全家吃顿饺子庆贺。他问孩子:“妈妈包了多少个饺子?”小孩子耳濡目染, 指着盘中的饺子数道:“0、1、2、3……”妈妈无奈地感叹:“天生加班的程序员命。”

类似的困扰也出现在小学数学中的“植树问题”中。一条 100 米长的路, 每隔 10 米种一棵树, 共需要几棵树? 围绕 400 米长的操场跑道, 每 20 米种一棵树, 需要多少棵数? 对于前者, 我们只要把路想成数轴, 意识到 0 是起点, 从 0 数起: 第 0 棵、第 1 棵、第 2 棵……第 10 棵。第 i 棵到起点的距离就是 $10i$ 米。$100 \div 10 = 10$ 故第 10 棵树距起点 100 米。而 0、1、2……10 共 11 个数, 故共种 11 棵数。对于后者, 也从 0 开始数。400 米是一圈, 所以 $400 \div 20 = 20$, 第 20 棵树距离起点恰好是一圈 400 米, 它和第 0 棵树位置重合, 不需要种了, 故 0、1、2……19、~~20~~, 共 20 棵树。

生活中我们一般从 1 开始数数 1、2、3……, 除了程序员和他可爱的孩子, 似乎没有人从 0 开始数数。读者朋友们, 你觉得什么情况下应该从 1 开始数, 什么情况下应该从 0 开始数呢?(见练习 2.1)

2.2.2 基数

从“无”到“起始”的变化是认识的转变, 世界上各个文明对“无”有着不同的理解。基督教的创世纪神话描述世界开始之前是混沌与无, 天地、光明、万物是世界的开始。佛教认为我们感知到的物质世界只是幻象,“无”才是世界的本质。古代中国在万物的变化中寻找规律, 认为阴阳是推动变化的力量, 从混沌的太极中演生大千世界。这些不同的认识和文化传统, 导致了对于零的不同态度。随着东西文明的交融, 人们对于无的认识也发生了变化。例如在西方哲学中, 黑格尔系统地应用辩证法考虑绝对的有与绝对的无, 并以纯粹的有无作为逻辑学的开始。今天的宇宙大爆炸理论认为我们的宇宙在诞生自 138 亿年前的一次大爆炸, 根据相对论的观点, 在此之前没有时间, 没有空间。

既然代表“无”的 0 在序数上意味着起始, 那么 0 在基数上是否可以“无中生有”呢? 与序数对应的概念是基数 (cardinal)。在前面的例子中, 3 的值可以代表 3 厘米的长度、3 克的质量、3 只羊……这些都是具体的大小。那么抽象的基数 3, 应该能

代表任何包含 3 个事物的“复多”。用集合论的语言就是任何有 3 个元素的集合，如 $\{\bigstar, \square, \bigcirc\}$ 或 $\{a, b, c\}$。这正是逻辑学家弗雷格对抽象的数的定义。代表“无”的 0 是不包含任何元素的集合（即空集）的基数。布尔巴基学派在 1939 年用符号 $\varnothing$ 代表空集。1 是包含一个元素的集合的基数，一个元素的集合可以是 $\{\bigstar\}$、$\{a\}$，甚至是 $\{\varnothing\}$。这是一个集合的集合，它包含的唯一元素是空集 $\varnothing$。$\{\varnothing\}$ 的基数是 1。这样就用“无”构造出了“有”、从 0 到 1。2 是包含了两个元素的集合的基数。我们可以从 1 到 2 构造出 $\{\varnothing, \{\varnothing\}\}$。它包含两个不同的元素（都是集合）：空集 $\varnothing$ 和 $\{\varnothing\}$。3 是包含三个元素的集合的基数，可以从 2 到 3 构造出：$\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$。以此类推。

基数	集合
0	$\varnothing$
1	$\{\varnothing\}$
2	$\{\varnothing, \{\varnothing\}\}$
3	$\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$
$\cdots$	$\cdots$
n	$\{\varnothing, \{\varnothing\}, \cdots\}$

这样就“无中生有”，以至万物；从 0 到 1，从 1 到多。

2.3 负数

0 是打开魔盒的钥匙，它引发着人们的好奇心。一旦接受了 0，自然就有这些问题：(1) 1 的前面是 0，0 的前面是谁？(2) 1 比 2、3……小，0 比 1 还小，还有比 0 更小的数么？(3) 做减法 $a - b$ 时，如果 b 比 a 还大会怎样？

到 1500 年左右，零已被人接受作为一个数，但负数的历史比零更短。16 世纪和 17 世纪的大多数数学家并不承认它们是数[20]-208。最早使用负数的是古代中国人。在汉代的《九章算术》中，已经有完整的“正负术”，如卷八中：“同名相除，异名相益……其异名相除，同名相益……”翻译成白话文的意思是：“（做减法时）正负相同值相减，正负相反值相加……（做加法时）正负不同值相减，正负相同值相加。”刘徽在注释中说：“今两算得失相反，要令正负以名之。正算赤，负算黑，否则以邪正为异。”翻译成白话文的意思是：“如果在计算中两个量所代表的意思相反，一个代表得到，一个代表失去，就需要用“正”、“负”来命名它们。用红色算筹表示正数，用黑色算筹表示负数（见第 1.4 节算筹部分）。如果手头没有不同颜色的算筹，就把一个正着放、一个斜着放加以区分。”[15]

可见中国古人在用算筹计算时，有两种不同颜色的算筹：红色代表正数，黑色代表负数。并形成了系统的计算规则。对正负数的实际意义也有解释，用正数表示卖出（对应收入），用负数表示买入（对应支出）。这既源于战国时代后日益繁荣的商业贸易，又源于传统思想文化中的阴阳概念。如图 2.9 所示，红色的阳对应正，黑色的阴对应负。直到今天，我们仍然在医疗检测时说结果是阴性、阳性 [11]，在电化学中说电极是阴极、阳极。

[11] 阴性的英文为 negative（负），阳性的英文为 postive（正）。

图 2.8: 阴阳

古希腊文明将数的概念建立在几何之上。长度、面积、体积对应着几何图形的线段、形状、占据的空间。它们自然都是正的, 没有形成负数的概念。伴随着十进制位值制系统和零的演进, 印度数学在 7 世纪出现了负数。数学家婆罗摩笈多(公元 589～670 年)在他的著作中明确地区分正负数。他用财富和负债赋予正负数实际意义, 使用特殊的符号标记负数, 并给出了一系列的规则, 如[16]:

1) 债务减零是债务。
2) 财富减零是财富。
3) 零减零为零。
4) 零减负债是财富。
5) 零减财富是负债。
6) 零乘以负债或财富是零。
7) 零乘以零为零。
8) 财富与财富的积或商是财富。
9) 负债与负债的积或商是财富。
10) 财富与负债的积或商是负债。
11) 负债与财富的积或商是负债。

图 2.9: 婆罗摩笈多(安德里亚斯•施特里克绘)

古希腊亚历山大时代的数学家丢番图(公元 200～284 年)开始独立于几何之外研究算术问题。他使用符号代表未知量, 考虑了一次方程和二次方程的解。在他的著作《丢番图算术》中有一道题目: 求方程 $4 = 4x + 20$ 的解。面对引出的负数解, 丢番图认为它是荒谬的。公元 9 世纪, 阿拉伯数学家从印度的数学和天文学著作中接触到了负数, 但他们同时又受到古希腊数学的影响, 从而对负数产生了矛盾的态度。数学家花拉子密一方面接受负数的概念及其运算规则, 并将其应用到解方程的移项与对消中, 但另一方面, 他使用几何论证方法的正确性, 因此不得不排斥负数。花拉子密不能像我们今天一样用 $ax + b = 0$ 概括所有一次方程, 也不能能用 $ax^2 + bx + c = 0$ 概括所有二

次方程。为了避免负系数, 他把方程划分为 6 类, 如 $ax = b$、$b - ax = 0$、$ax^2 + bx = c$、$bx + c = ax^2$、$ax^2 + c = bx$ 等。直到 15 世纪, 欧洲学者才开始接触、研究负数。16 世纪, 意大利数学家卡尔达诺[12](1501～1576 年) 在他的著作《大数》[13]中给出了一般三次、四次方程的公式解法。他同样遵循古希腊传统, 使用几何证明正确性。把一次项解释为长度、二次项解释为面积、三次项解释为体积。不接受负数作为方程的系数。卡尔达诺不得不给出了 60 种不同类型的方程。

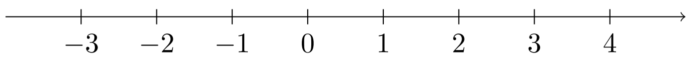
2.3.1 序数

17 世纪, 英国数学家沃利斯(1616～1703 年)引入了数轴, 并通过序数理解负数的意义。从序数上说, 负数是 0 之前的部分, 我们看下面两个例子。

例 2.3.1. 公元纪年。以耶稣基督诞生的一年为公元 *1* 年, 此后为公元 *2* 年、*3* 年……那么在耶稣诞生前呢? 依然有时间, 有历史事件发生。这些历史事件发生在公元前。耶稣诞生前的年份依次是公元前 *1* 年、公元前 *2* 年……公元前的年份蕴含着负数的意义(注意, 这里缺少公元 *0* 年)。

例 2.3.2. 温度。进入冬天, 气温逐渐下降, 终于 *3* 度、*2* 度、*1* 度、*0* 度, 降到了冰点。但 *0* 度不是终结, 尽管水凝固成冰, 但随着温度继续下降, 温度计上的红色煤油液面跨过了 *0* 度, 降到了零下 *1* 度、零下 *2* 度……零下的温度比 *0* 度低, 它们是负温度。

把温度计横过来, 并向两边无限延伸, 我们就得到了数轴:



负数作为序数在 0 和任何正数的左侧。0 分开了正数、负数, 位于中间。我们在第 2.2.1 介绍的规则依然成立: 若 $a < b$, a 在 b 的左边, 如果 $b = 0$, a 为负数, 如果 $a = 0$ 则 b 为正数。在序数意义下, 小于“<”和大于“>”关系还可以用差的正负来表示。$a - b < 0$ 则 $a < b$, a 在 b 的左边[14]。$a - b > 0$ 则 $a > b$, a 在 b 的右边。若 $d > 0$ 则 $a + d$ 向前(右)移动 d, $a - d$ 向后(左)移动 d, 若 $a < d$, 则移动到负数的区域。如果 $d < 0$ 则移动的方向相反: $a + d$ 向后(左), $a - d$ 向前(右)。

2.3.2 基数

尽管沃利斯用数轴给出了负数的意义, 但他却奇怪地认为负数比无穷大, 但不小于零。在《无穷大的算术》[15]中, 沃利斯写到: 由于比例 $a/0$ 在 a 为正数时是无穷大[16], 故当分母变为负数时, 例如当 a/b 中的 b 是负数时, 这个比必定大于无穷大[20]。这都

[12]也有人译作卡尔丹

[13]拉丁文名为 Ars Magna, 出版于 1545 年。

[14]$a < b$, 两边减 b: $a - b < b - b = 0$

[15]拉丁文名为 Arithmetica Infinitorum, 1655 年出版

[16]沃利斯是微积分的早期先驱之一, 当时尚未形成严密的极限概念。若 a、b 都是正数, 当观察到随着 b 减小到接近 0 时, a/b 变得很大, 于是随意地认为 $a/0$ 是无穷大。

说明人们对于负数作为基数的意义感到困惑。总体来说,16 世纪和 17 世纪,并没有很多数学家对于使用负数心安理得或者承认它是数,更谈不上承认它们可以作为方程的真实的根[20]。笛卡尔的坐标系只有第一象限,帕斯卡则认为从 0 减去 4 纯粹是胡说。天文学家第谷·布拉赫使用了符号"-"来标识负数,但他认为负数只能私下使用。

负数引起的困惑是可以理解的。以温度为例,零下三十几度的南极洲显然比零下二十几度的东北地区更寒冷。从寒冷的程度看,如果基数表现大小,似乎南极洲温度的基数更大,但从数轴上的序数看 $-30 < -20$。从年代上看,孔子(公元前 551～公元前 479 年)生活的时代比孟子(公元前 372～公元前 289 年)生活的时代更久远。如果基数表现大小,似乎孔子时代年份的基数更大,但从纪年的序数看 $-479 < -372$。这些问题逐渐使我们认识到绝对值的概念。在数轴上,一个数 a 的绝对值是 a 到 0 的距离(长度),例如正数 5 的绝对值是 $5 - 0 = 5$,即 5 本身;负数 -5 的绝对值是 $0 - (-5) = 5$。在数轴上-5 和 5 关于 0 对称,彼此称为对方的相反数。a 的相反数是 $-a$。而 0 到自己的距离是 0 - 0 = 0。这样归纳下来,数轴上一个数 a 的绝对值记作 $|a|$,有 3 种情况:

$$|a| = \begin{cases} a & a > 0 \\ 0 & a = 0 \\ -a & a < 0 \end{cases}$$

用语言描述就是,正数和零的绝对值是它本身,负数的绝对值是它的相反数。二十世纪初,范·德·瓦尔登在《代数学》[17] 一书中给出了一个"构造"负数的方法[17]-8。用一对数偶 (a, b) 来表示整数如下:

- 用 $(a + b, b)$ 表示自然数 a。
- 用 (b, b) 表示 0。
- 用 $(b, a + b)$ 表示负数 $-a$。

例如 (3, 1) 表示 2, (5, 5) 表示 0, (100, 101) 表示-1。我们可以把 (a, b) 想象成天平两边的质量,天平平衡时 (b, b) 两边质量相等,左右的的质量差为 0;天平左侧质量更大的时候,左右质量差为正数;天平右侧质量更大的时候,左右质量差为负数。每个数都有许多表示,例如 $(3, 1) = (9, 7) = (360, 358) = \cdots$ 都表示 2;$(5, 5) = (100, 100) = (128, 128) = \cdots$ 都表示 0;$(100, 101) = (1, 2) = (8, 9) = \cdots$ 都表示 -1。但每个符号 (a, b) 定义一个,且只定义一个整数,即:

- 用 $a > b$ 表示自然数 $a - b$。
- 用 $a = b$ 表示 0。
- 用 $a < b$ 表示负数 $-(b - a)$。

[17] 原名《近世代数》。数学家范·德·瓦尔登当时从荷兰来到德国哥廷根求学,他根据埃米尔·阿廷和艾米·诺特的演讲和讨论班内容整理出了这本抽象代数讲义,影响了一代数学家。

接下来为数偶定义加法和乘法:

$$(a,b)+(c,d)=(a+c,b+d) \tag{2.1}$$

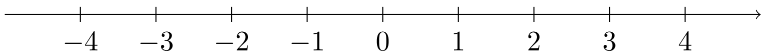
$$(a,b)\cdot(c,d)=(ac+bd,ad+bc) \tag{2.2}$$

例如加法 $2+(-1)=(5,3)+(2,3)=(7,6)=1$, 乘法 $-2\times 3=(1,3)\cdot(5,2)=(1\times 5+3\times 2,1\times 2+3\times 5)=(11,17)=-6$。练习 2.3 要求读者验证加法、乘法的交换律、结合律、交换律(合称五大定律)对数偶都适用。

这个方法完全不使用零和负数, 只用 1、2、3……形式化地定义出了正数、零、负数, 绕开了历史上引起争议的问题。不禁令人想起克罗内克的话:"上帝创造了自然数, 其余都是人的工作。"

2.4　数轴

引入零和负数后, 数轴上就标记出了向两侧无限延伸的整数。



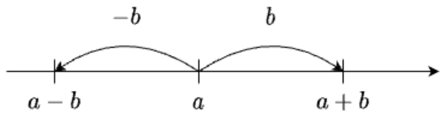
我们介绍一下标准符号的由来。0、1、2、3……叫做自然数, 记作 $\mathbb{N}$。是英文 nature numbers 的首字母。在没有认识到 0 以前, 自然数不包括 0, 只有 1、2、3……但随着 0 的引入, 我们认识到 0 才是真正"自然"的起点、原点(见第 2.6 节皮亚诺公理)。今天大多数定义(包括我国的部编义务教育数学教材)都把 0 算作自然数。通常用小写字母 n 表示某个自然数。引入负数后, 负整数、零、正整数, 即……−3、−2、−1、0、1、2、3……合称整数, 记作 $\mathbb{Z}$。它来自德语"数"Zahlen 的首字母。正整数有时记作 $\mathbb{Z}^+$, 负整数记作 $\mathbb{Z}^-$。整数的英文是 integral number 或 integer, 我们有时也用其首字母 i 表示某个整数。在枚举一些列的事物时, 通常用 i 作下标, 例如:$a_1,a_2,...,a_i,...,a_n$, 其中 a_i 表示第 i 个。但这里的 i 不是 integer, 而是索引的英文 index 的首字母。a_n 通常指最后一个, 即第 n 个。数学家们非常重视符号的使用。好的符号形象、生动、易于理解, 表达能力强, 甚至具有意想不到的力量(如莱布尼茨的微积分符号)。韦达、笛卡尔开创了代数符号的传统, 他们用 a,b,c 表示已知量, 用 x,y,z 表示未知量。这些习惯影响至今。

利用数轴, 我们可以给出四则运算($+,-,\times,\div$)的直观解释(第 3 章会给出几何解释)。在 2.2.1 节, 我们说加减 $a\pm b$ 将数轴上的数 a 移动 b 个单位, 共有五种移动情况:

<table>
<tr><th></th><th>$b>0$</th><th>$b<0$</th><th>$b=0$</th></tr>
<tr><td>$a+b$</td><td>向右</td><td>向左</td><td rowspan="2">不动</td></tr>
<tr><td>$a-b$</td><td>向左</td><td>向右</td></tr>
</table>

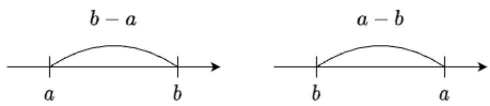
我们看到 $b<0$ 时 $a+b$ 相当于减法, 而 $a-b$ 相当于 a 加上 $-b$。b 和 $-b$ 互为相反数, 这样减法无非是加法的一种情况。我们可以用加法概括减法:

	$b>0$	$b<0$	$b=0$
$a+b$	向右	向左	不动

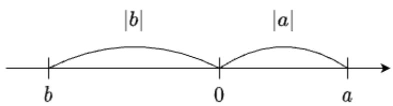


任给数轴上两个数 a, b, 它们的距离是:

	$a<b$	$a>b$	$a=b$
距离	$b-a$	$a-b$	0

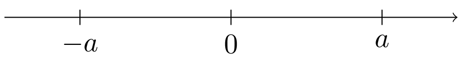


这恰是绝对值的概念, 距离等于 $|a-b|$。特殊情况下, 若 $b=0$, 一个数 a 到原点 0 的距离是 $|a|$。



如果我们把“差”理解为“差距”, 可正可负。在数轴上给定一个数 a, 它与另一个数 b 的差(距)就是 $a-b$。若 $a-b>0$ 则 $a>b$, a 在 b 的右侧; 若 $a-b<0$ 则 $a<b$, a 在 b 的左侧; 若 $a-b=0$ 则 $a=b$, a 和 b 重合。特别地, a 到原点 0 的差(距)$a-0=a$, 就是它本身。

在数轴上, 一个数 a 的相反数就是它关于原点 0 的对称点 $-a$; 0 的相反数是它自己。



0 能“无中生有”: $0\begin{cases} n \\ -n \end{cases}$, “分裂”成任何数 n 和它的相反数 $-n$。反之, n 和 $-n$ 遇到一起“湮灭”为 0, 如同一对正负电荷。以上我们固定数轴, 移动数轴上的数(点)。运动是相对的, 我们也可以移动变换整条数轴:

2.4.1 平移

把数轴上的所有数 $+1$ 相当于把整条数轴怎样移动呢? 向右? 请看图 2.10:

奇怪, 竟然向左移动了 1 格。$\{\cdots-2,-1,0,1,2\cdots\}$ 各自加 1 后变成了 $\{\cdots-1,0,1,2,3\cdots\}$, 的确向左移动了 1 个单位。那么 -1 一定是反向, 也就是向右移动了, 如图 2.11 所示。

一般地, $+a$ 将数轴移动 $-a$, $-a$ 将数轴移动 a。

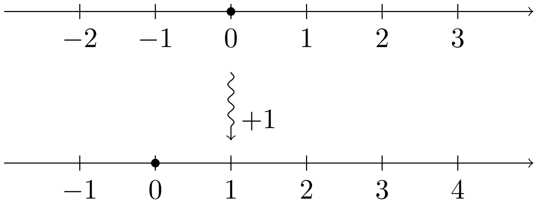


图 2.10: 将数轴上每个数 $+1$

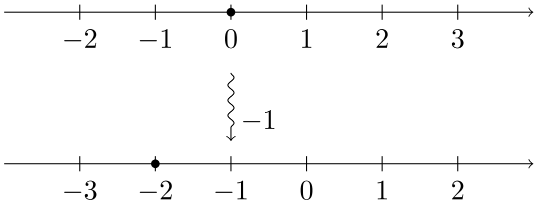


图 2.11: 将数轴上每个数 -1

	$a > 0$	$a < 0$	$a = 0$
$\mathbb{Z} + a$	向左	向右	不动
$\mathbb{Z} - a$	向右	向左	

这是为什么呢? 记原数轴为 A, 移动后的数轴为 A'。A 上的每个数 $x \rightsquigarrow x + a$, 变成了 $x' = x + a$, A 上的 $-a \rightsquigarrow -a + a = 0$, 变成了 A' 上的原点。所以 A' 上的原点对应着原数轴上的 $-a$。若 $a > 0$, 自然新原点(对应 $-a$)在老原点(对应 0)的左边, 如图 2.12 所示。

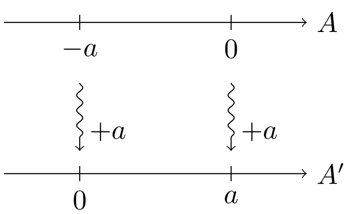


图 2.12: 数轴平移

2.4.2 相反数

把数轴上的所有数取相反数(简称取反, 英文 negate)相当于把数轴反向(绕原点旋转 180°)。

取两次反自然会复原, 即旋转 180° 再旋转 180° 等效于旋转 360°, 等效于不动。

$\rightarrow$	反向	$\leftarrow$	反向	$\rightarrow$
a	相反数	$-a$	相反数	$-(-a) = a$

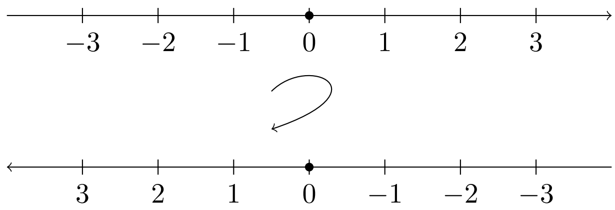


图 2.13: 数轴反转

取相反数 $a \rightsquigarrow -a$ 等效于乘以 -1, 即 $-a = a(-1) = (-1)a$, 两次取反复原就解释了“负负为正”: $-(-a) = (-1)(-a) = a$。

2.4.3　缩放

第 2 节说乘除相当于缩放。$\times 2$ 把数轴放大 2 倍, $\times 3$ 把数轴放大 3 倍, $\div 2$ 把数轴缩小到 $\frac{1}{2}$, $\div 3$ 缩小到 $\frac{1}{3}$。除法是乘法的逆运算, 所以 $\div 2$ 也可以认为是 $\times \frac{1}{2}$, 即缩放到 $\frac{1}{2}$ 倍, $\div 3$ 即缩放到 $\frac{1}{3}$。这样乘除法就统一成乘法一种情况:

	$a > 1$	$0 < a < 1$	$a = 1$
$\times a$	放大	缩小	不变

引入负数后, 如上节所说, $\times(-1)$ 相当于取相反数, 数轴反向。$\times(-2)$ 相当于放大 2 倍后反转数轴, 即 $(-1) \times 2$, 当然也等效于先反转数轴再放大 2 倍, 即 $2 \times (-1) = (-1) \times 2$。同样 $\times(-3)$ 相当于放大 3 倍后反向, 或反向后放大 3 倍。这样乘除法可以归纳如下:

	$a > 1$	$a = 1$	$0 < a < 1$	$-1 < a < 0$	$a = -1$	$a < -1$
$\times a$	放大	不变	缩小	缩小 + 反向	反向	放大 + 反向

有一种特殊情况: $\times 0$ 数轴“坍缩”成一点。上述“动作”可以组合: 先平移 a 再平移 b, 当然这等效于先平移 b 再平移 a, 即 $a + b = b + a$, 这说明了加法的交换律。同样, 先缩放 a 倍, 再缩放 b 倍和先缩放 b 倍, 再缩放 a 倍等效, 即 $ab = ba$, 这说明了乘法交换率。我们把结合率的数轴意义作为练习 2.6。

更复杂的组合, 如先平移再缩放或反向, 如 $x \rightsquigarrow ax + b$, 即先缩放 a 倍, 再平移 $-b$。练习 2.7 要求读者考虑如何先平移再缩放。

2.5　单位元

0 和负数成了 1、2、3……一样的一等公民。它们有序数(再数轴上合法、确定的位置), 有基数(确定的值)。0 和其它数比起来仍是如此特殊。它是起点, 它是原点, 它是产生万物的“无”, 还有谁像它这样独特么? 1 站了出来要和 0 比一比。

$$\begin{array}{cc}
0 & 1 \\
1+0=0+1=1 & 1\times 0=0\times 1=0 \\
2+0=0+2=2 & 1\times 2=2\times 1=2 \\
3+0=0+3=3 & 1\times 3=3\times 1=3 \\
\cdots & \cdots \\
a+0=0+a=a & 1\times a=a\times 1=a
\end{array}$$

0 加上任何数都不变, 而 1 乘以任何数(包括 0)也都不变。我们看看 0 和 1 对数轴的作用:

1) 加法把数轴左右平移, $+a$ 后数轴移动 $-a$ 个单位($a > 0$ 向左, $a < 0$ 向右), 0 的作用是固定数轴不动。

2) 乘法把数轴放大缩小, $\times a$ 则缩放 a 倍($|a| > 1$ 放大, $0 < |a| < 1$ 缩小), 1 的作用是固定数轴不变。

对于 0, 任何一个数 a 都能找到它的"加法双胞胎"兄弟, 相反数 $-a$, 使得 $a + (-a) = 0$。对于 1, 任何一个数 a 也能找到它的"乘法双胞胎"兄弟, 倒数 $\frac{1}{a}$, 使得 $a \times \frac{1}{a} = 1$, 不过 0 没有"乘法双胞胎"兄弟。读者朋友们, 你们觉得 0 和 1 谁更特殊呢?

像 0 和 1 这样"鹤立鸡群"的例子比比皆是。英语中有单词、短语、句子、段落……我们把一列英语字符叫做"字符串", 如 "apple"、"hello"、"to be or not to be"、" "(空格)。我们可以把字符串连接在一起, 如: "hello" $+\!\!+$ " " = "hello ", "hello " $+\!\!+$ "apple" = "hello apple"。不包含任何字符的串叫空串 "", 把空串连接到任何字符上都不变:

$$\text{``''} +\!\!+\, S = S +\!\!+ \text{``''} = S$$

如 "" $+\!\!+$ "hello" = "hello" $+\!\!+$ "" = "hello"。再比如集合的并。把两个集合中包含的元素不重复地放在一起组成新的集合, 如 $\{a, b, c\} \cup \{1, 2, 3\} = \{a, b, c, 1, 2, 3\}$, $\{a, b, c\} \cup \{c, d, e\} = \{a, b, c, d, e\}$。空集与任何集合的并都不变:

$$\varnothing \cup S = S \cup \varnothing = S$$

如 $\varnothing \cup \{a, b, c\} = \{a, b, c\} \cup \varnothing = \{a, b, c\}$。数学上我们把 0 叫做加法的单位元, 1 叫做乘法的单位元, 空串叫做字符串连接的单位元, 空集叫集合并的单位元。一般地在集合(可以有限也可以无限) M 中, 如果存在一个元素 e, 在某种二元运算 $\odot$ 下使得:

$$e \odot a = a \odot e = a$$

对 M 中的所有 a 都成立, 我们把 e 叫做运算 $\odot$ 在 M 中的单位元。例如 0 是加法在数的集合中的单位元, 1 是乘法在除去 0 之外的数的集合中的单位元, 空串是连接操作在字符串集合中的单位元, 空集是并在所有集合的集合中的单位元。单位元和集合中的其它元素一样是一等公民, 但它是特殊的、唯一的(见练习 2.8)。

对集合中的元素 a, 如果存在某个元素 b 使得 $a \odot b = e$, 则称 b 是 a 的逆。例如 a 在加法运算下的逆就是相反数 $-a$, 因为 $a + (-a) = 0$。非 0 数 a 在乘法运算下的逆是 $\frac{1}{a}$, 因为 $a \times \frac{1}{a} = 1$。并非所有元素都有逆, 例如除了空串外, 所有字符串都没有连接操作的逆; 除了空集外, 所有集合都没有并操作的逆。有了逆, 我们就可以把加减乘除四则运算简化为“二则运算”, $a - b$ 只不过是 a 加上 b 的逆(相反数), $a \div b$ 只不过是 a 乘以 b 的逆(倒数)。

2.6 皮亚诺公理

1889 年, 意大利数学家皮亚诺为自然数建立起严格的公理化系统。它使用零作为起点, 在加上“数数”这个动作, 定义出了自然数。这组公理一共有五条, 合称皮亚诺公理。

公理 1) 0 是自然数。

公理 2) (数数动作)每个自然数 n 都有它的下一个自然数 n', 称为它的后继。

似乎仅仅有这两条公理, 已经能够定义出无穷无尽的自然数了, 从 0 开始, 下一个是 1, 下一个是 2……接下来是某个 n, 下一个是 $n + 1$……以至无穷。但是好挑剌的数学家给出了一个反例: 考虑只有两个元素 $\{0, 1\}$ 组成的数字系统, 定义 1 的后继为 0, 0 的后继为 1。这样也满足上面的两条公理, 却不是我们想像中的自然数。为此我们还需要第三条皮亚诺公理来排除这种情况。

公理 3) 0 不是任何自然数的后继。

仅仅有这三条公理就够了么? 我们还可以给出一个反例: 考虑有限元素 $\{0, 1, 2\}$ 组成的数字系统, 定义 0 的后继是 1, 1 的后继是 2, 2 的后继还是 2。这样也能满足上述三条公理。为此我们还需要第四条皮亚诺公理。

公理 4) 不同的自然数有不同的后继。如果两个自然数 m 和 n 的后继相同, 则它们相等。即 $m = n$ 当且仅当 $m' = n'$;

但是, 仅仅用这四条公理仍然不够, 因为存在这样的反例: 考虑集合{0, 0.5, 1, 1.5, 2, 2.5, ...}, 定义 0 的后继是 1、1 的后继是 2……, 0.5 的后继是 1.5、1.5 的后继是 2.5……但 0.5 不是任何数的后继。为了排除这种“不可达”的数, 还需要最后一条皮亚诺公理。

公理 5) 如果自然数的某个子集 S 包含 0, 并且其中每个元素都有后继元素。那么 S 就是全体自然数 $\mathbb{N}$。

为什么公理 5 可以排除掉上述反例呢? 考虑 $\{0, 0.5, 1, 1.5, 2, 2.5, ...\}$ 的一个子集 $\{0, 1, 2, ...\}$。它包含 0, 并且每个元素都有后继元素, 但是它不等于原集合。因为 1.5、2.5……都不在这个子集中。所以它不满足第五条公理。公理五还有另外一个响亮的名字 —— 归纳公理, 它可以这样被等价地描述:

公理 5) 任意关于自然数的命题,如果证明了它对自然数 0 是对的,又假定它对自然数 n 为真时,可以证明它对 n' 也真,那么命题对所有自然数都真。(这条公理保证了数学归纳法的正确性)

例如,我们想证明能爬上任意层的大厦。用数学归纳法证明需要两个步骤: 1) 证明我们能进入大楼一层大厅。2) 证明在任一层楼我们有能力"更上一层楼"。以上就是完整的五条皮亚诺公理,也称为皮亚诺算术系统。我们接下来展示皮亚诺公理系统,包括数学归纳法的威力。利用皮亚诺公理,我们可以定义出自然数的加法和乘法。加法定义包含两部分:任何自然数和 0 相加不变; a 和 b 的后继相加,等于 $a+b$ 的后继,即:

$$\begin{aligned} a+0 &= a \\ a+b' &= (a+b)' \end{aligned} \tag{2.3}$$

我们来验证一下 2+3。2 是 0 的二重后继 $0''$, 3 是 0 的三重后继 $0'''$, 根据加法定义 2+3 为:

$$0''+0''' = (0''+0'')' = (0''+0')'' = (0''+0)''' = (0'')''' = 0'''''$$

结果的确是 0 的五重后继,也就是 5。加法定义可以直观地解释为从数轴上的 a 移动到 b' 相当于先移动到 b 再前进 1。这种在定义中包含了自身的定义称为递归的。它把一个问题转换为规模较小的问题。在本例中,原问题 $a+b'$ 被递归地转换为问题 $a+b$, 规模降低了 1。此时如果 $b \neq 0$ 就继续减小规模; 如果 $b=0$ 则 $a+0=a$, 递归结束。0 在这里被称为递归出口。我们接下来定义自然数的乘法:

$$\begin{aligned} a \cdot 0 &= 0 \\ a \cdot b' &= a \cdot b + a \end{aligned} \tag{2.4}$$

乘法定义也包括两部分:任何数乘以 0 为 0; a 和 b 的后继相乘等于 a 乘以 b 再加上 a。其中第二部分可以解释为: $a \cdot (b+1) = a \cdot b + a \cdot 1$。我们可以验证一下 2×3:

$$\begin{aligned} 0'' \times 0''' &= 0'' \times 0'' + 0'' = 0'' \times 0' + 0'' + 0'' = 0'' \times 0 + 0'' + 0'' + 0'' \\ &= 0 + 0'' + 0'' + 0'' = 0'' + 0'' + 0'' = (0'' + 0')' + 0'' = (0'' + 0)'' + 0'' \\ &= (0'')'' + 0'' = 0'''' + 0'' \\ &= (0'''' + 0')' = (0'''' + 0)'' = (0'''')'' = 0'''''' \end{aligned}$$

结果的确是 0 的 6 重后继。乘法定义也是递归的。它将问题 $a \cdot b'$ 转化为两个子问题:(1) 递归地求 $a \cdot b$(规模降低了 1);(2) 在此基础上 $+a$, 直到 b 的规模降低到 0, 递归到达出口结束。递归是一种常见的方法,将复杂问题分解,各个击破。这种策略叫做"分而治之",特别适合于利用机器解决问题,在现代计算机系统中得到大量应用。

与通常的观念不同,加法和乘法的交换律、结合律既不是公理也不是公设。它们都是定理,可以用皮亚诺公理和定义证明。以加法结合律为例: $(a+b)+c = a+(b+c)$。我们利用数学归纳法,先证明当 $c=0$ 时它是对的。根据加法定义的第一条规则:

$$(a+b)+0 = a+b = a+(b+0)$$

然后是递推步骤, 假设 $(a+b)+c = a+(b+c)$ 成立, 我们要推出 $(a+b)+c' = a+(b+c')$。

$$\begin{aligned}
(a+b)+c' &= (a+b+c)' && \text{加法定义的规则二, 反向} \\
&= (a+(b+c))' && \text{递推假设} \\
&= a+(b+c)' && \text{加法定义的规则二} \\
&= a+(b+c') && \text{加法定义的规则二, 反向} \quad \square
\end{aligned}$$

这样就证明了加法的结合律。但加法交换律的证明却并不简单, 参见附录 C 的证明。

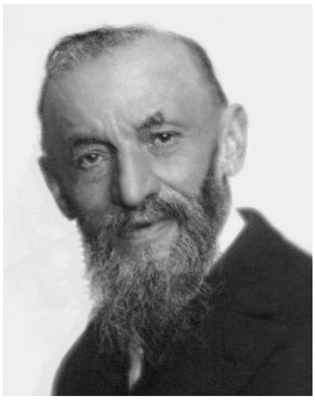


图 2.14: 朱塞佩·皮亚诺(Giuseppe Peano)1858 - 1932。

朱塞佩·皮亚诺 1858 年 8 月 27 日生于意大利库内奥(Cuneo)附近的斯宾尼塔(Spinetta)村。这是位于意大利北部都灵附近的农村。他出生的时候, 正值意大利统一。1876 年皮亚诺考入都灵大学学习, 1880 年毕业后就留校任教。皮亚诺最开始讲授微积分课程, 1887 年他和克罗西奥(Carola Crosio)结婚。1886 年起, 皮亚诺还同时担任都灵军事学院的教授。从 19 世纪 80 年代起, 皮亚诺开始研究数理逻辑, 并致力于数学基础的构建工作。他撰写了《数学公式汇编》这本巨著, 力图把所有的数学成果都用形式化的方法汇集起来。这本书可以说为数学的严密化奠定了基础。1900 年在第二届国际数学家大会上, 罗素遇到了皮亚诺, 他在自传(1951)中说[18]:

“这次大会是我的精神生活的一个转折点, 因为在那里我遇到了皮亚诺。在此之前, 我已经听说过他的名字, 也知道他的一些工作。我突然明白了, 他的符号提供了我多年来一直试图寻找的分析的工具, 而且从他那里我获得了一直以来想要从事的工作的一种新的有效的技术。”

皮亚诺强烈地影响了罗素和怀特海合著的《数学原理》一书, 对早期的计算理论起了重要的作用。皮亚诺最初用法语发表研究著作, 但他对自然语言固有的歧义感到不满。为了解决这个问题, 他于 1900 年左右发明了一套没有歧义的统一

语言,称为“无屈折拉丁语”。这门语言后来被称为“国际语”(拉丁文 Interlingua 和世界语(Esperanto)是两种不同的人造语言)。皮亚诺努力推广他的新语言,但是事实并不如他所愿,几乎没有人愿意读他用国际语重写的《数学公式汇编》。相反倒是他早期的法语著作使数学家的观点发生了深刻的变化,尤其对法国的布尔巴基学派的纲领,产生了很大影响。1932 年 4 月 20 日,皮亚诺因心脏病逝世于都灵。

练习 2

2.1. 什么情况下应从 1 开始数数,什么情况下应从 0 开始数数?

2.2. ★ 定义数偶的减法和除法

2.3. 验证数偶的加法乘法满足交换律、结合律、分配律

2.4. 用乘法对加法的分配律证明 $(-1) \times (-1) = 1$

2.5. 证明一个数的相反数是唯一的。

2.6. 说明加法、乘法结合律的数轴意义。

2.7. 考虑 $x \rightsquigarrow 2x + 1$,如果先平移再缩放,应先向哪个方向平移多少?再缩放几倍?一般地,$x \rightsquigarrow ax + b$ 应如何变换?

2.8. 证明单位元是唯一的。

2.9. ★ 定义 0 的后继为 1,证明对于任何自然数都有 $a \cdot 1 = a$,提示:利用附录 C 的结论。

2.10. ★ 证明乘法分配律。

2.11. ★ 证明乘法结合律和交换律。

2.12. ★ 如何利用皮亚诺公理验证 $3 + 147 = 150$?

答案: 62 页

第 3 章　分数

此曲只应天上有，人间能得几回闻。

杜甫《赠花卿》

2015 年 9 月 14 日，美国 LIGO[1] 天文台探测到一个神秘信号 GW150914。它来自 13 亿光年之外的一次惊天动地的奇观：一对双黑洞天体彼此靠近、吸引，旋转着合并到一起（如图 3.1）。它们巨大引力激起的“涟漪”穿越 13 亿年时空 [2] 到达了地球。物理学家们花了几个月的时间进行数据分析，排除了可能的干扰因素，最终于 2016 年 2 月 11 日正式宣布：这是人类首次探测到了引力波，证实了爱因斯坦在百年前通过广义相对论做出的预言。LIGO 的物理学家们把引力波转换成声音信号，我们得以首次听到来自宇宙深处的“天籁之音”。

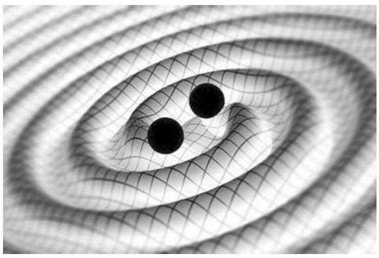


图 3.1: 双黑洞合并过程中产生引力波示意。

2500 年前，正是追寻天籁之音的过程，使得古希腊的先贤毕达哥拉斯发现了音乐背后的数学秘密。相传有一天毕达哥拉斯经过铁匠铺，听到从里面传出了悦耳的声音。这些声音是铁匠们用不同重量的铁锤一起敲打铁砧时产生的。巴洛克时期的音乐家亨德尔有一部作品叫做《快乐的铁匠》（作品编号：HWV430）。毕达哥拉斯注意到有些声音是和谐的，而另一些不和谐。他进一步发现当铁锤的重量比恰好是 12、9、8、6 时，敲打的声音是和谐的。毕达哥拉斯敏锐地发现了音乐背后的数字规律：

- 比例 12 : 6（即 2 : 1）对应纯 8 度音；
- 比例 9 : 6（即 3 : 2）对应纯 5 度音；

[1] 激光干涉引力波

[2] 根据爱因斯坦的广义相对论，引力波以光速传播。

- 比例 $12:9$(即 $4:3$)对应纯 4 度音;
- 比例 $9:8$ 对应纯 2 度音。

这个故事非常流行, 如 3.2 所示。这幅四联木刻连环画展示了毕达哥拉斯先是聆听到铁匠们挥舞铁锤敲打, 然后他开始进行定量研究, 包括敲打不同重量的钟, 敲打盛有不同水的杯子, 弹拨坠有不同重量的琴弦, 吹奏不同长度的笛子。但这个故事禁不起推敲, 第二幅图和第一幅图是相互矛盾的。我们可以自己动手验证一下。依照第二幅图中的样子, 用几个同样的杯子盛上不同的水, 然后用不同大小的勺子敲击。我们会发现音调的高低是由杯子而不是勺子决定的。同理, 铁匠铺传出的音调高低是由铁砧而不是锤子决定的。但第一幅图中只有一个铁砧。这组连环画还有别的问题。铁锤、钟、杯子、琴弦上的砝码、笛子上的数字 4、6、8、9、12、16 是印度——阿拉伯数字(见第 1.4 节), 要到十三、十四世纪才传入欧洲。古希腊的毕达哥拉斯不可能用这样的数字。尽管如此, 这组连环画仍然反映了毕达哥拉斯求知好学, 善于思考, 动手实践进行定量研究的治学传统。

图 3.2: 毕达哥拉斯与音乐, 出自 1492 年(或 1480 年)弗兰奇诺·加富里奥《乐理》中的木刻插页。

古希腊的乐器叫做里尔琴(Lyre, 也译作里拉琴、莱雅琴、诗琴), 是一种七弦琴(见图 3.3a), 在很多场合已经成为代表音乐的符号。我们推测毕达哥拉斯把琴弦的一半, 也就是 $\frac{1}{2}$ 张紧弹奏, 获得了 8 度音; 把琴弦的 $\frac{2}{3}$ 张紧弹奏, 获得了 5 度音; 把琴弦的 $\frac{3}{4}$ 张紧弹奏, 获得了 4 度音; 这些音调之间的关系如图 3.4 所示。

(a) 绘有太阳神阿波罗手持里尔琴的盘子，约公元前 480～470 年，藏于雅典德尔菲博物馆。

(b) 里尔琴符号

毕达哥拉斯通过数学，具体说是分数与比例奠定了西方音乐的理论基础。他认为整个宇宙是一把巨大的里尔琴，古希腊人所知道的七个天体（日月和五大行星水星、金星、火星、木星、土星）是琴上的七根弦，在不同的音调上震动。而数与比例代表着宇宙和谐的天籁之音。

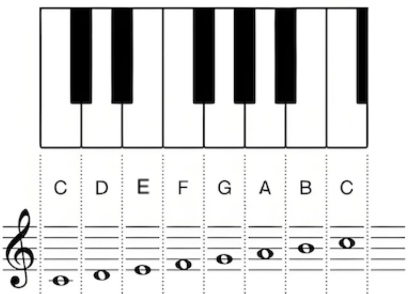
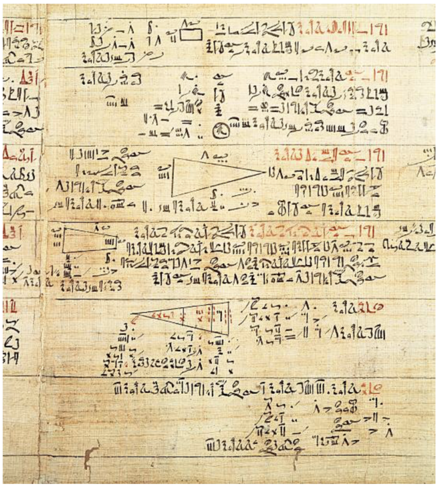


图 3.4: 两个 C 间隔 8 度，琴弦比 2 : 1；C 与 G 间隔 5 度，琴弦比 3 : 2；C 与 F 间隔 4 度，琴弦比 4 : 3。

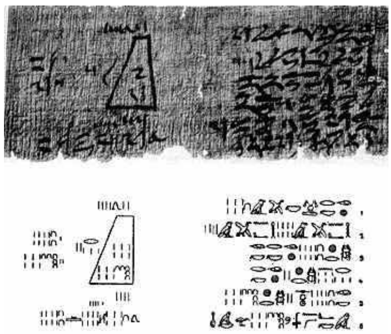
分数的历史比零和负数还要长。古代中国在春秋时代（公元前 770 年～前 476 年）的《左传》中，规定了诸侯的都城大小：最大不可超过周文王国都的三分之一，中等的不可超过五分之一，小的不可超过九分之一。但它们只是表达某种部分的量的词语，而不是真正意义上的分数，因为它们没有直接参与加减乘除运算。古代中国要到汉代才发展出完整的分数计算规则，见 3.3 节。最早的分数出现在古埃及。

3.1　埃及分数

我们是通过两份重要的古代文件了解到埃及分数的。它们是莱茵德纸草书（见图 3.5a）和莫斯科纸草书（见图 3.5b）。所谓纸草书 [3] 是古埃及广泛使用的书写载体。它用当时盛产于尼罗河三角洲的纸莎草的茎作为材料，经过切片、浸泡、压制、干燥制成。在公元前 3000 年左右甚至出口到希腊。莎草纸在埃及的干燥气候下可以很好地保存，经历千年而不腐坏。但在潮湿的环境下，很容易霉变损毁。这两份保留了古埃及数学成果的纸草书异常珍贵。它们一份由英国埃及学者莱茵德（Rhind）于 1858 年购得，现藏于大英博物馆。内容是公元前 1650 年前后的教科书，作者是书记官阿梅斯 [4]，包含有 85 道数学问题。莫斯科纸草书又名戈列尼雪夫纸草书，由俄罗斯贵族戈列尼雪夫于 1893 年在埃及购得，现藏于莫斯科普希金精细艺术博物馆。作者是约公元前 1890 年的一位佚名作者，包含有 25 道数学问题。



(a) 莱茵德纸草书局部



(b) 莫斯科纸草书局部

这两份文物反映了古埃及人已使用分数参与运算解决问题。但古埃及的分数有一个既奇怪又合理的特点——所有的分子都是 1。因此被称为“单位分数”。在古埃及象形文字（圣书体，见 1.1 节）中，在数字 a 上方写（画）一个“眼睛”来表示 $\frac{1}{a}$，如所图 3.6 示，表示 $\frac{1}{5}$ 和 $\frac{1}{15}$。单位分数的意义是直观易懂的。$\frac{1}{3}$ 表示把某物均分 3 份后每份的量，$\frac{1}{5}$ 是均分 5 份后每份的量，$\frac{1}{a}$ 是均分 a 份后每份的量。

但是 $\frac{3}{4}$、$\frac{2}{7}$ 这样的值分子不为 1，古埃及人就把它们表示为单位分数之和。例如 $\frac{3}{4} = \frac{1}{2} + \frac{1}{4}$，写成象形文字如图 3.7 左侧所示。中间的符号象征着一个人行走的双腿。如果行走的方向同书写方向一致表示相加，行走的方向同书写的方向相反表示相减。图 3.7 右侧表示 $\frac{2}{5}(= \frac{1}{3} + \frac{1}{15})$。将一个值表示为单位分数的和时，古埃及人要求每个单位分数都不同，不允许重复。因此 $\frac{2}{7}$ 不能写成 $\frac{1}{7} + \frac{1}{7}$ 而要写成 $\frac{1}{4} + \frac{1}{28}$。某些情况下，仅仅分解为两个单位分数是不够的。例如莱茵德纸草书上将 $\frac{2}{29}$ 分解为

[3] 英文 Papyrus，是纸的英文 paper 的字源

[4] Ahmes，也译作阿默斯。

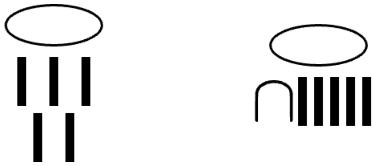


图 3.6: 埃及分数 $\frac{1}{5}$ 和 $\frac{1}{15}$

$\frac{2}{29}=\frac{1}{24}+\frac{1}{58}+\frac{1}{174}+\frac{1}{232}$。并且分解方式是不唯一的，例如 $\frac{2}{29}=\frac{1}{15}+\frac{1}{435}=\frac{1}{16}+\frac{1}{232}+\frac{1}{464}$。这样的分解没有明显的规律，普通人难以掌握。为此，古埃及人制作了大量的表格以供查找计算。

图 3.7: 左: $\frac{1}{2}+\frac{1}{4}$ 右: $\frac{1}{3}+\frac{1}{15}$

迄今为止，没有发现直接的文献，揭示为何古埃及人将分数系统设计为不同的单位分数之和。我们只能加以猜测。有一则传说故事说：老父亲在临终前希望把自己的遗产分给三个儿子。给大儿子 $\frac{1}{2}$，二儿子 $\frac{1}{3}$，小儿子 $\frac{1}{9}$。但全部遗产只是 17 匹马。三个儿子不知道如何分配，也不愿杀死马进行分割。他们于是去请教村里的长者。老人将自己的一匹马借给了他们。于是大儿子分得 $18\times\frac{1}{2}=9$ 匹马，二儿子分得 $18\times\frac{1}{3}=6$ 匹马，小儿子分得 $18\times\frac{1}{9}=2$ 匹马，剩余 $18-9-6-2=1$ 匹马，恰好就是智慧的老人借给他们的那一匹。于是物归原主，老人牵走了他自己的马。这个传说故事用埃及分数解释就是

$$\frac{17}{18}=\frac{1}{2}+\frac{1}{3}+\frac{1}{9}$$

尽管脍炙人口，这个传说不可能是古埃及人发展分数系统的初衷。我们在各个民族，包括阿拉伯、印度、犹太、中国的民间故事中都发现了类似的故事。分配的遗产有马、骆驼、大象等动物，数量也有不同。例如三个儿子按照 $\frac{1}{2}$、$\frac{1}{4}$、$\frac{1}{6}$ 分配 11 个动物，体现为：

$$\frac{11}{12}=\frac{1}{2}+\frac{1}{4}+\frac{1}{6}$$

练习 3.5 要求给出所有能这样进行分配的数量和比例。我们最早在十八世纪伊朗哲学家纳拉奇（Mulla Muhammad Mahdi Naraqi）的著作中看到此故事的文字记载。其次，尽管在莱茵德纸草书中有关于遗产分配的问题，但主要的分数应用是平均分配。例如第 63 题[20]-15：把 700 块面包分给四个人，第一人得 $\frac{2}{3}$，第二人得 $\frac{1}{2}$，第三人得 $\frac{1}{3}$，第四人得 $\frac{1}{4}$。阿梅斯的解法在我们今天看来相当于解方程 [5]：

[5] $\frac{2}{3}$ 是古埃及极少数特殊的量，不需要写成单位分数之和。

$$\frac{2}{3}x+\frac{1}{2}x+\frac{1}{3}x+\frac{1}{4}x=700$$

他把 $\frac{2}{3}$、$\frac{1}{2}$ 、$\frac{1}{3}$ 、$\frac{1}{4}$ 加起来得到 $1\frac{3}{4}(=1+\frac{1}{2}+\frac{1}{4})$,然后用 1 除以 $1\frac{3}{4}$ 得 $\frac{4}{7}=(\frac{1}{2}+\frac{1}{14})$。最后 $700\times\frac{4}{7}=400$,从而得到 $x=400$。这相当于我们今天小学高年级的一元一次方程。注意到第一个人和第三个人分得的面包不是整数块。

另一种观点认为,这种分数系统可以用更经济的方式分配物品。考虑将 5 块面包分给 8 个人。一种方法是将每个面包平均切分成 8 份,然后每人拿 5 块,如图 3.8 所示。

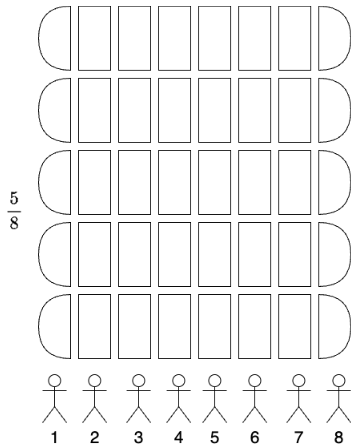


图 3.8: 共切成了 40 小块,每人分得 5 块。

古埃及人可能发现了更好的切割方式,如图 3.9 所示。即每个人拿走半块面包和 $\frac{1}{8}$ 面包。

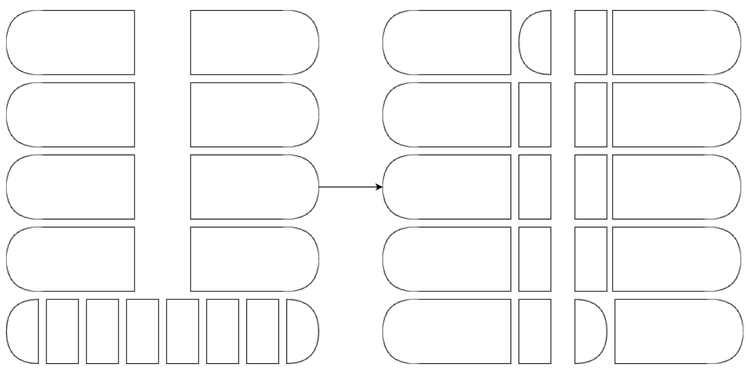


图 3.9: 共切成了 16 块,每人分得 2 块(一大 $\frac{1}{2}$ 和一小 $\frac{1}{8}$)。

读者朋友们,你觉得古埃及人为何如此设计他们的分数系统呢?以后来的眼光来看埃及分数,一方面它的计算很复杂,逐渐被历史的长河淘汰了,另一方面,数学家思考这样的问题:1)是否每个分数都可以分解成埃及分数?2)如果可以分解,怎样的分解方式最好?其中第一个问题在 1202 年由斐波那契在《算盘书》中解决了。而第二个问题催生了至今仍未解决的数学猜想。在此之前,我们先看两个相对简单的问题。

命题 3.1.1. 每个埃及分数都可以分解为两个不同的埃及分数之和，即 $\frac{1}{n} = \frac{1}{a} + \frac{1}{b}$。

证明.

$$\begin{aligned}
\frac{1}{n} &= \frac{n+1}{n(n+1)} && \text{上下同} \times (n+1) \\
&= \frac{\cancel{n}}{\cancel{n}(n+1)} + \frac{1}{n(n+1)} \\
&= \frac{1}{n+1} + \frac{1}{n(n+1)}
\end{aligned}$$

□

我们可以立即用这个结论推出：

命题 3.1.2. 每个 $\frac{2}{n}$ 都能分解为埃及分数。

证明.

$$\begin{aligned}
\frac{2}{n} &= \frac{1}{n} + \frac{1}{n} \\
&= \frac{1}{n} + \frac{1}{n+1} + \frac{1}{n(n+1)} && \text{上面证明的结论}
\end{aligned}$$

□

当 n 是奇数时，我们能得到更好的结果：$\frac{2}{n}$ 一定能分解为两个埃及分数：

证明.

$$\begin{aligned}
\frac{1}{n} &= \frac{1}{n+1} + \frac{1}{n(n+1)} && \text{由命题 3.1.1} \\
\frac{2}{n} &= \frac{2}{n+1} + \frac{2}{n(n+1)} && \text{左右同} \times 2 \\
&= \frac{1}{\frac{1}{2}(n+1)} + \frac{1}{n\frac{1}{2}(n+1)} && n\text{是奇数，上下同} \div 2
\end{aligned}$$

□

进一步，练习 3.1 要求证明当 n 是奇素数时，这种分解是唯一的。莱茵德纸草书中有一张表记录了从 $\frac{2}{5}$ 到 $\frac{2}{101}$ 的所有分解，印证了这个结论。回到斐波那契证明的定理，考虑任何即约分数 [6]。如果它是假分数，可以先转化成带分数，然后只考虑分解真分数部分。所以只需要证明：

定理 3.1.3 (斐波那契). 任何即约真分数 $\frac{b}{a}$ 可分解为埃及分数。

证明. 斐波那契利用带余数的除法：$a = bq + r$，其中 q 是商、r 是余数，且 $0 < r < b$。最接近并小于 $\frac{b}{a}$ 的分数是 $\frac{1}{q+1}$。斐波那契考虑

$$\frac{b}{a} = \frac{1}{q+1} + x$$

然后再把 x 分解为埃及分数。通过这一分而治之的思想，原来的问题就转化为分解 x 的问题。

$$\begin{aligned}
x &= \frac{b}{a} - \frac{1}{q+1} = \frac{bq+b-a}{a(q+1)} && \text{通分} \\
&= \frac{b-(a-bq)}{a(q+1)} = \frac{b-r}{a(q+1)} && \text{余数}r = a - bq
\end{aligned}$$

[6] 分子、分母不能进一步约分的分数。顾名思义，“即约”指完成了约分。

注意到 x 的分子 $b' = b - r$。由于余数 $0 < r < b$, 所以 $b' < b$, 它比原分数 $\frac{b}{a}$ 的分子 b 至少减小了 1。如果 $b' = 1$, 则分解完成, 否则我们继续分解 x。这样就可以得到一系列不断缩小的分子序列 $b > b' > b'' > \cdots$ 因为 b 是一个确定的正整数, 它不会无限减小, 所以必定在某次达到 1 从而结束分解。因此每个即约真分数都可以分解为埃及分数。 □

我们用一个例子 $\frac{5}{11}$ 来理解斐波那契的证明。$11 = 2 \times 5 + 1$, 商 $q = 2$、余数 $r = 1$。第一步分解:

$$\begin{aligned}
\frac{5}{11} &= \frac{1}{q+1} + x = \frac{1}{3} + x \\
x &= \frac{5}{11} - \frac{1}{3} = \frac{4}{33} \\
\frac{5}{11} &= \frac{1}{3} + \frac{4}{33}
\end{aligned}$$

接下来分解 $\frac{4}{33}$。再次用带余数除法 $33 = 4 \times 8 + 1$, 商 $q = 8$、余数 $r = 1$。

$$\begin{aligned}
\frac{4}{33} &= \frac{1}{q+1} + x = \frac{1}{9} + x \\
x &= \frac{4}{33} - \frac{1}{9} = \frac{12 - 11}{99} = \frac{1}{9}
\end{aligned}$$

分解结束, 得到:

$$\frac{5}{11} = \frac{1}{3} + \frac{1}{9} + \frac{1}{99}$$

斐波那契使用的策略是每次用最接近原分数 $\frac{b}{a}$ 的埃及分数进行分解。这种策略叫做贪心策略。但是贪心策略并不一定给出最优分解。例如

$$\frac{5}{121} = \frac{1}{25} + \frac{1}{757} + \frac{1}{763309} + \frac{1}{873960180913} + \frac{1}{1527612795642093418846225}$$

但还有另一个分解:

$$\frac{5}{121} = \frac{1}{33} + \frac{1}{121} + \frac{1}{363}$$

所谓最优的含义是: 用最少的埃及分数分解, 如果两个分解的埃及分数个数相同, 则分母越小越好。斐波那契的方法每次至少将分子减 1, 所以 $\frac{b}{a}$ 最多一定被分解为 b 个埃及分数, 但 $\frac{5}{121}$ 的例子说明可能存在着更好的分解方法。匈牙利数学家保罗·埃尔德什和德国数学家斯特劳斯在 1950 年猜想对于大于 1 的自然数 n,

$$\frac{4}{n} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}$$

总成立, 其中 $a > b > c$。也就是说 $\frac{4}{n}$ 总能分解为 3 个两两不同的埃及分数。这个著名的数论问题称为埃尔德什 —— 斯特劳斯猜想 [7], 迄今(2015)仍未解决。

[7] Erdős - Straus

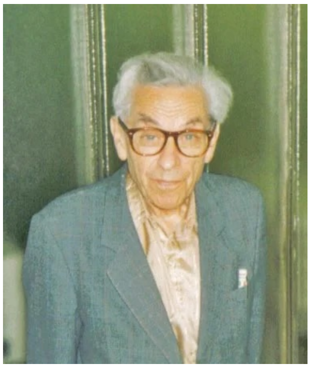


图 3.10: 保罗·埃尔德什(1913～1996)

保罗·埃尔德什(Paul Erdős), 匈牙利籍犹太人, 著名的"自由职业"数学家。一生发表论文 1500 多篇(包括与其他作者合作), 是迄今发表论文数最多的数学家(第二位为欧拉)。埃尔德什在数论、组合数学等方面作出了很大贡献[23]。他四处游历, 探访当地的数学家, 与他们一起工作, 曾和 507[a] 人合写论文。这使得他事实上成为了一个巨大的数学合作网络的中心人物。因此有人定义了"埃尔德什数", 简称埃数。埃尔德什自己的埃数为 0, 与他直接合作写论文的人的埃数为 1, 与埃数为 1 的人合写论文的人埃数为 2, 依此类推。例如爱因斯坦的埃数为 2, 菲尔兹奖获奖者的埃数中值最低时为 3。目前已知最大的埃数为 15(不包括非数学家, 他们的埃数为无穷大)。

1913 年, 埃尔德什出生在匈牙利首都布达佩斯, 父母都是犹太人。埃尔德什出生前两个姐姐都死于可怕的猩红热, 父母生怕他夭折, 因此对他格外呵护。可是生逢乱世, 他还不到 1 岁第一次世界大战就爆发了。父亲在沙俄与奥匈帝国的战争中被俘, 被关在西伯利亚的战俘营 6 年直至战争结束才生还。令人惊讶的是, 埃尔德什的父亲在被俘期间自学了英语。由于没有人教授发音, 父亲读音很奇怪。埃尔德什从他父亲那里也学到了这种奇怪的口音, 并持续终生。一战后的匈牙利政局混乱, 罢工不断。作为中学数学教师的母亲为了不影响孩子们的教育坚持上课, 结果政局变乱时, 她因不参与罢工而丢掉了工作。1920 年匈牙利已经开始迫害犹太人, 13 年后希特勒上台, 欧洲进一步走向了战争。

尽管环境如此恶劣, 埃尔德什在匈牙利全国考试中脱颖而出, 考入了布达佩斯的帕兹马尼·彼得大学, 并于 1934 年获得了博士学位。可是犹太人的处境恶劣, 他选择了去英国曼彻斯特大学做博士后研究。在剑桥大学, 他认识了数学家哈代和乌拉姆, 后者成了他的终身朋友。埃尔德什想探望在匈牙利的父母, 但希特勒在 1938 年占领了奥地利和捷克, 他被迫在中途逃回, 并辗转到了美国普林斯屯高等研究院。埃尔德什花了 1 年时间与他人合作创立了概率数论, 但研究院不知为何只答应延长 6 个月任期。乌拉姆在这个困难的时候伸出援手, 邀请埃尔德什去麦迪逊的威斯康星大学。这开启了埃尔德什四处游历, 随遇而安研究数学的"流浪生活"。

埃尔德什坚信上帝手中有一本“天书”, 里面包含着最优美、最简洁的数学证明。他毕生都在追寻那些“数学天书中的证明”。尽管有些定理已经被证明了, 埃尔德什会继续寻找更完美, 有时甚至是初等的证明。比如 1845 年约瑟夫·伯特兰猜想任何整数 n 和 $2n$ 之间至少存在着一个素数。1850 年切比雪夫证明了这个猜想, 因此今天它叫做“波特兰 —— 切比雪夫定理”。埃尔德什在 19 岁时重新用初等方法证明了这个猜想。他对自己的证明很满意, 相信这就是“天书”中的证明。“这简直来自数学天书!”成了埃尔德什对他人工作的最高评价。1998 年, 埃尔德什诞生 85 周年之际, 数学家们收集了最优美的 45 个证明汇集成了一本《数学天书中的证明》。

第二次世界大战中, 埃尔德什和匈牙利的家人完全失去了联系。战争结束后他才获知, 父亲已经于 1942 年因心脏病突发去世。家里有 4 位亲属被害, 有位表兄弟是奥斯维辛集中营的幸存者。最幸运的是, 母亲还活着。从 1943 年起, 埃尔德什“云游”在普渡大学、斯坦福大学、圣母大学、约翰斯·霍普金斯大学之间。他没有全职教职, 但反而可以自由选择和任何人合作研究任何喜欢的问题。这种生活即工作, 工作即生活的人生持续了半个世纪。居无定所、没有妻子、没有孩子、没有固定工作的羁绊。足迹遍布 22 个国家, 尽管有时不得不在某些国境线前返回: 冷战期间, 他的祖国匈牙利怀疑他是间谍而拒绝其入境。美国在麦卡锡主义盛行时也禁止他入境。尽管没有披露原因, 有人猜测他与 1949 年回国的中国数学家华罗庚在通信中讨论数学问题, 移民局的官员们担心信中那些宛如天书的数学符号可能是密码……[25] 不速之客埃尔德什经常突然出现在某位数学家的门口: “我的大脑敞开了!”然后住在别人家里一段时间, 一起研究有趣的数学问题。

埃尔德什是匈牙利科学院、美国科学院、英国皇家学会会员。他在 1951 年获得科尔奖, 在 1984 年获得沃尔夫奖。面对 5 万美元奖金, 他只花掉了 720 美元, 而把剩下的全部捐出设立奖学金以纪念他的父母。1996 年, 83 岁的埃尔德什在波兰华沙突发心脏病去世。在几个小时前, 他刚刚在华沙的学术会议上解决了一个棘手的几何问题。

[a]一说为 511 人。在埃尔德什死后, 仍有 70 多篇以他为作者的论文发表[24]-p588。

3.2 古巴比伦的分数

以今天的眼光来看, 埃及分数不便于计算, 像是在数学上走了一段弯路。它的形式意义大于它的实用意义。此后的古希腊、古罗马并未发展出独立的分数系统, 而是用一些特定的词汇表示部分的量。古罗马表示部分的词汇来源于其重量单位系统。单位重量叫做 as, 它的 $\frac{1}{12}$ 叫做 uncia, 是英制重量盎司 ounce 的词源。表 3.1 列出了常用的罗马分数名词:

有许多词至今仍影响着我们的语言。表示 $\frac{1}{2}$ 的 semi, 常用作“半”的词根, 例如半圆 semicircle、分号 semicolon、半导体 semiconductor 等。表示 $\frac{1}{4} = \frac{3}{12}$ 的 quadrans 在英语中是 quarter, 仍表示一刻钟、一季度。

值	词汇	意义	值	词汇	意义
$\frac{1}{12}$	uncia	十二分之一	$\frac{9}{12}$	dodrans	去掉四分之一
$\frac{2}{12}$	sextans	六分之一	$\frac{10}{12}$	dextans	去掉六分之一
$\frac{3}{12}$	quadrans	四分之一	$\frac{11}{12}$	deunx	去掉一个 uncia
$\frac{4}{12}$	triens	三分之一	$\frac{1}{24}$	semuncia	uncia 的一半
$\frac{5}{12}$	quincunx	五个 uncia	$\frac{1}{48}$	sicilicus	
$\frac{6}{12}$	semis	一半	$\frac{1}{72}$	scriptulum	
$\frac{7}{12}$	septunx	七个 uncia	$\frac{1}{144}$	scripulum	$\frac{1}{12} \times \frac{1}{12}$
$\frac{8}{12}$	bes	三分之二	$\frac{1}{288}$	scrupulum	$\frac{1}{24} \times \frac{1}{12}$

表 3.1: 罗马分数名称

古巴比伦发展出了分数系统。我们从耶鲁大学收藏的一块泥板中找到了“物证”，如图 3.11 所示。此图描绘了一个正方形和对角线。正方形的一边刻有楔形文字 30(右上角的 3 个楔形)，在对角线下刻有文字 42, 25, 35，贯穿着这同一对角线还刻有文字 1, 24, 51, 10。我们在初中数学课上学到过正方形的对角线长度是边长的 $\sqrt{2}$ 倍。简单计算 $30 \times \sqrt{2} \approx 30 \times 1.4142 = 42.426$，其中 42 恰好是对角线下所刻文字的第一个。古巴比伦人使用 60 进制，我们怀疑接下来的 25, 35 是小数部分。这很容易验证：

图 3.11: 耶鲁大学古巴比伦泥板，编号 Y7289

$$\frac{25}{60} + \frac{35}{60^2} \approx 0.4264$$

这说明古巴比伦人正确地计算出了对角线长度。42, 25, 35 代表 $42.426 \approx 30 \times \sqrt{2}$。那么 1, 24, 51, 10 会不会也是个小数呢？我们试一试：

$$1 + \frac{24}{60} + \frac{51}{60^2} + \frac{10}{60^3} \approx 1.414213$$

竟然是 $\sqrt{2}$ 的小数表示，并且精度达到了百万分位(小数点后 6 位)！我们由此确知古巴比伦人发展出了 60 进制小数系统。但由于没有小数点，他们的数字是有歧义

的。例如 12, 15 既可能是 $12 \times 60 + 15 = 720$ 也可能是 $12 + \frac{15}{60} = 12.25$。再加上 0 没有被统一使用，我们只能通过上下文猜测一个数的真正含义。

由于使用 60 进制，古巴比伦人可以比较精细、准确地表示 $\frac{1}{60} \sim \frac{59}{60}$ 之间的部分量。当使用 2 到 3 位 60 进制小数时，已足够生产、生活、乃至天文观测所需的精度。此外，60 的真因数[8] 足够丰富（包括 2, 3, 4, 5, 6, 10, 12, 15, 20, 30），除不尽产生循环小数的机会较少（10 进制只有 2 和 5 两个真因数，只有少数情况能除尽而不循环，参见定理 3.4.6）。可能由于这些因素，古巴比伦人只使用 60 进制小数而没有发展出一般的分数系统。

3.3 古代中国的分数

在汉代的《九章算术》中，已经包含了分数四则运算规则和各种应用题。这些内容集中出现在第一卷“方田”中（见图 3.12）。由于篇幅不大，我们不妨一起赏析一下中国古人的智慧。首先是分数的记法。《九章算术》中已经使用了“a 分之 b”这样的描述，和现代汉语完全一致，如：“十五分之一”。古人用步数表示长度，例如一块田长九步、宽五步。当用分数表示长度时，说“a 分步之 b”，如：“七分步之四”、“五分步之三”。而在现代汉语中，我们说七分之四步、五分之三步。又如表示金额：“八钱三分钱之一”，可见已经有了带分数的概念。对应现代汉语中的八又三分之一钱。《九章算术》中的分数直接来源于除法，如：“七人卖四马，一人卖七分马之四。”这就是 $\frac{4}{7}$ 的意义。

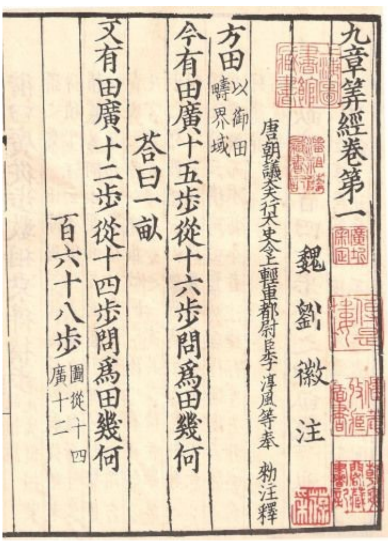

九章筭經卷第一 魏 劉徽 注
唐朝議大夫行太史令上輕車都尉臣李淳風等奉 勑注釋
方田 以御田疇界域
今有田廣十五步從十六步問爲田幾何
荅曰一畝
又有田廣十二步從十四步問爲田幾何
荅曰 百六十八步 圖從十四廣十二

图 3.12:《九章算术》书影

有了分数记法，接下来《九章算术》要处理一个分数的值是否“唯一”的问题。这就引出了“约分”的概念。作者用两道例题说明：

1) 今有十八分之十二，问约之得几何？ 答曰：三分之二。（即：$\frac{12}{18} = \frac{2}{3}$）

[8] 除 1 和 n 以外的因数称 n 的真因数。

2) 又有九十一分之四十九,问约之得几何? 答曰:十三分之七。(即:$\frac{49}{91}=\frac{7}{13}$)

具体怎样约分呢?《九章算术》给出了"辗转相除"法,即著名的欧几里得算法求最大公约数(也叫做最大公因数):"可半者半之;不可半者,副置分母、子之数,以少减多,更相减损,求其等也。以等数约之。"解释成现代汉语:首先是特殊情况。若分子分母都是偶数,先不断 $\div 2$,化为奇数。然后是一般情况。记分数为 $\frac{b}{a}$,记 a 与 b 的最大公约数为 (a,b),则:

$$(a,b)=\begin{cases} a>b: & (a-b,b) \\ a=b: & a \\ a<b: & (a,b-a) \end{cases} \tag{3.1}$$

就是"以少减多,更相减损、求其等也"的意思。这种方法在欧几里得的《几何原本》中有系统的描述和正确性证明,今天被称做"欧几里得算法",在中国常被称做辗转相除法。对比一般的方法:把 a、b 分解为因子的积,然后找出最多的公因子相乘,欧几里得算法的效率极高 [9],便于机械化计算,是当今数论和密码学的基础算法 [10]。我们将在下一章中详细介绍它的原理。这里举一个例子:求 18 和 12 的最大公约数。

$$(18,12)=(18-12,12)=(6,12)=(6,12-6)=(6,6)=6$$

只用了 3 步就求出了结果。最后分子分母同除以最大公约数得:$\frac{12}{18}=\frac{12/6}{18/6}=\frac{2}{3}$。

《九章算术》接下来通过例题引入分数加法:

1) 今有三分之一,五分之二,问合之得几何? 答曰:十五分之十一。(即:$\frac{1}{3}+\frac{2}{5}=\frac{11}{15}$)

2) 又有三分之二,七分之四,九分之五,问合之得几何? 答曰:得一、六十三分之五十。(即:$\frac{2}{3}+\frac{4}{7}+\frac{5}{9}=1\frac{50}{63}$)

作者解释分数加法原理:"母互乘子,并以为实。母相乘为法。"即:

$$\frac{b}{a}+\frac{d}{c}=\frac{bc+ad}{ac} \tag{3.2}$$

新的分子叫"实",分母叫"法"。接下来:"实如法而一。不满法者,以法命之。"即 $\frac{n}{n}=1$,抽取整数部分化为带分数。最后说:"其母同者,直相从之。"即:

$$\frac{b}{a}+\frac{c}{a}=\frac{b+c}{a} \tag{3.3}$$

分母相同时,直接将分子相加。可见《九章算术》的通分并未使用最小公倍数,而是直接将分母相乘。

《九章算术》接下来介绍分数减法,基本上类似于加法,我们省略跳过。之后介绍分数大小的比较。《九章算术》的作者采用和减法一样的处理,通过母子互乘比大小。

[9] 利用除法代替式 (3.1) 中的减法可以达到"对数复杂度"。100 位的整数,大约递归 8 次左右即可算出。

[10] 在抽象代数,如环论、域论中也有重要应用

即将 $\frac{b}{a}$ 与 $\frac{d}{c}$ 的大小比较转化为 bc 与 ad 的大小比较。接下来在介绍多个分数的平均数之后讲解分数除以整数, 例如:“今有七人, 分八钱三分钱之一。问人得几何? 答曰: 人得一钱二十一分钱之四。”即 $8\frac{1}{3} \div 7 = 1\frac{4}{21}$。作者解释说:“以人数为法, 钱数为实, 实如法而一。有分者通之。”即用金额作分子, 人数作分母。这里并没有考虑除数是分数的情况。最后是分数乘法:

1) 今有田广七分步之四, 从五分步之三, 问为田几何? 答曰: 三十五分步之十二。(即: $\frac{4}{7} \times \frac{3}{5} = \frac{12}{35}$)
2) 又有田广九分步之七, 从十一分步之九, 问为田几何? 答曰: 十一分步之七。(即: $\frac{7}{9} \times \frac{9}{11} = \frac{7}{11}$)

作者给出的计算规则是:“母相乘为法, 子相乘为实, 实如法而一。”即分子分母相乘再化带分数。有趣的是《九章算术》对带分数乘法进行了单独处理:

1) 今有田广三步三分步之一, 从五步五分步之二, 问为田几何? 答曰: 十八步。(即: $3\frac{1}{3} \times 5\frac{2}{5} = 18$)
2) 又有田广七步四分步之三, 从十五步九分步之五, 问为田几何? 答曰: 一百二十步九分步之五。(即: $7\frac{3}{4} \times 15\frac{5}{9} = 120\frac{5}{9}$)

作者解释说:“分母各乘其全, 分子从之, 相乘为实。分母相乘为法。实如法而一。”即化成假分数, 相乘再化带分数。例如:

$$3\frac{1}{3} \times 5\frac{2}{5} = \frac{10}{3} \times \frac{27}{5} = \frac{10 \times 27}{3 \times 5} = 18$$

《九章算术》一开篇就系统地引入了分数及其四则运算。可见中国古人对分数的重视。后继的各种图形面积、方程计算都建立在这个基础上。

3.4 印度分数和小数

我们今天使用的分数记法源自古印度。印度人将表示分子的数写在上方, 将表示分母的数写在下方, 但并未使用分数线, 如:

$$\begin{matrix} 3 \\ 4 \end{matrix}$$

表示四分之三。后来阿拉伯人在约公元 1200 年引入了分数线来分开分子与分母[21], 如:

$$\frac{3}{4}$$

斐波那契最早将现代分数记法介绍到欧洲, 但在写带分数时, 他照搬了阿拉伯人从右向左的书写习惯, 将分数写在整数的左边[22]。分数线带来了不小的印刷困难。在

1718 年, 英国著名的川宁茶叶公司创世人托马斯·川宁[11] 在记账时用斜线代替水平分数线, 如: 1/4 磅茶叶。这样就便于印刷和用打字机书写分数。今天很多字处理和编辑软件可以方便输入分数。不少网页支持了 LaTeX 排版, 可以用 `\frac{b}{a}`来输入 $\frac{b}{a}$。

随着印度——阿拉伯计数系统的确立, 阿拉伯学者开始引入了十进制小数。数学家阿尔·卡西(Al-Kashi, 公元 920 年～980 年)在他的著作《圆周的研究》[12] 中将圆周率 π 的值写为: “sah-hah 3 14159”, 其中 sah-hah 的意思是整数部分(今天土耳其语中的 sahih), 修饰了 3, 接下来是小数部分。可见此时尚未引入小数点。中国魏晋时的数学家刘徽(见图 3.13b)在注解《九章算术》时, 不满于古代用 3 作为圆周率, 于是: “微数无名知以为分子, 以十为分母”, 决定采用十进分数来求得更精确的值。当他用正 96 边形逼近圆时求得: “三百一十四寸六百二十五分寸之一百六十九”, 这个值的表示已经非常接近小数的形式了[13]。

(a) 纳皮尔肖像, 1616 年油画, 爱丁堡大学收藏

(b) 蒋兆和创作的刘徽像, 部编数学课本

1530 年, 鲁道夫(Christoff Rudolff, 1499?～1545?年)使用一条数线分隔整数和小数部分。后来马吉尼(Magini, 1555 年～1617 年)将竖线改进成了小数点, 但并未得到广泛使用。直到 20 年后英国学者约翰·纳皮尔(John Napier, 1550 年～1617 年, 见图 3.13a)发明对数时使用了小数点, 现代小数形式才被广泛接受。我们常用的百分数符号%是佚名的某位意大利人在 1425 年引入的。

从另一角度看, 小数是分母为 10、100、1000......的分数的和。百分数看似是分母为 100 的分数, 实则只是小数乘以 100 的结果。这是因为百分数的分子不一定是整数。例如某网站说它的可靠性非常高, 达 99.999%(俗称 5 个 9), 这相当于小数 0.99999。在日常生活中, 人们还会用千分数, 写作 999.99‰。在金融领域还会用基点(BPS: base points 的缩写)这个量。1bps = 0.01% = 0.0001。比如我们看见新闻说美联储(美国联邦储备委员会, 是美国的中央银行的核心管理机构)宣布降息 25 个基点。实际上就

[11] Thomas Twining

[12] *al-Risali al-mohitije* 译作英文 *Treatise on the circumference*

[13] 关于这个值, 刘徽接着写道: “则出圆之表矣。故还就一百九十二觚之全幂三百一十四寸以为圆幂之定率而弃其余分。”他觉得大了, 故得到此时的圆周率 3.14。但他不甘心, 又继续用 192 边形计算到“三百一十四寸二十五分寸之四”。这个带分数是 $314\frac{4}{25} = 314.16$ 寸。可见刘徽是混合使用小数和分数的。

是降低利息 0.25% = 0.0025。例如利率从 2.5%降低 25 个基点就变成 2.5% - 0.25% = 2.25%。

3.4.1 分数与小数的关系

现在我们回过头来说为何小数是分母为 10、100、1000……的分数和？并提出一个问题：是否每个分数都可以唯一地表示成小数？举例来说 $0.125 = \frac{1}{10} + \frac{2}{100} + \frac{5}{1000}$，恰好是分母为 10 的 $\frac{1}{10}$，分母为 100 的 $\frac{2}{100}$，分母为 1000 的 $\frac{5}{1000}$ 的和。一般来说，小数 $0.a_1a_2\ldots a_n$ 可表示为：

$$a_1\frac{1}{10} + a_2\frac{1}{100} + \cdots + a_n\frac{1}{10^n} \tag{3.4}$$

我们在数学课上知道，小数分为有限小数（如 0.125）和无限小数（如 $0.\dot{3} = 0.333\cdots$，$0.\dot{1}\dot{5} = 0.1515\cdots$，$\pi = 3.14159\cdots$）。根据式 (3.4)，无限小数可写成无限个分数的和。那么这无限个正数的和是个有限值还是无限大？为了解答这个问题，我们先引入一个看似反直觉的命题：

引理 3.4.1. 无限循环小数 $0.999\cdots$ 的值是 *1*。

乍看上去左边的 $0.999\cdots$ 和右边的 1 明显是两个不同的数，它们怎么会相等呢？

证明. 令 $x = 0.999\cdots$，扩大 10 倍：$10x = 9.999\cdots$，相减：

$$\begin{aligned} 10x - x &= 9.999\cdots - 0.999\cdots \\ 9x &= 9 && \text{右侧的小数部分减光了} \\ x &= 1 \end{aligned}$$

□

注意，小数部分只有有限个 9 时等号不成立，哪怕有 $n = 10000$ 个 9。因为 $9.99\cdots9$（个位的 9 和小数部分的 9999 个 9）减 $0.99\cdots9$（小数部分的 1 万个 9）等于 $8.99\cdots91$。因此 $x = \frac{8.99\cdots91}{9} \neq 1$。即 1 万位有限小数 $0.99\cdots9 \neq 1$。这样无限个正数的和 $\frac{9}{10} + \frac{9}{100} + \cdots = 1$，是个确定的有限值 [14]。

推论 3.4.2. 把任何无限小数 $0.a_1a_2\cdots$ 写成无穷多个分数的和：$a_1\dfrac{1}{10} + a_2\dfrac{1}{100} + \cdots$ 这个和是有限的。

证明. 由于每个小数位 $0 \leq a_i \leq 9$，所以

$$0 \leq 0.a_1a_2\cdots = a_1\frac{1}{10} + a_2\frac{1}{10^2} + \cdots \leq \frac{9}{10} + \frac{9}{10^2} + \cdots = 0.999\cdots = 1$$

□

这个结论：$0.a_1a_2\cdots \leq 1$ 看似说了一句废话，但无限多个（正）值加起来是有限的是一个极为惊人的数学与逻辑断言。它是解决一个有着两千多年历史的悖论的钥匙。

[14] 我们也可以用高中数学的极限概念进行证明。注意到 n 位有限小数：$0.99\cdots9 = 1 - 0.00\cdots01 = 1 - \frac{1}{10^{n+1}}$，所以 $0.999\cdots = \lim\limits_{n\to\infty}(1 - \frac{1}{10^{n+1}}) = 1 - \lim\limits_{n\to\infty}\frac{1}{10^{n+1}} = 1 - 0 = 1$。

古希腊的哲学家，埃利亚的芝诺（Zeno of Elea，约公元前 490～425 年）提出了著名的四个悖论。第一个悖论最为人们所津津乐道，名叫阿基里斯与乌龟悖论。阿基里斯是荷马史诗《伊里亚特》中的英雄，以英勇著称。这个悖论说：如果让爬得很慢的乌龟在阿基里斯前面一段路程出发，那么阿基里斯将永远追不上乌龟。这是因为，阿基里斯为了赶上乌龟，必须先到达乌龟的出发点 A，但当阿基里斯到达 A 点时，乌龟已经在这段时间前进到了 B 点。但当阿基里斯到达 B 点时，乌龟又已经到了前面的 C 点……以此类推，两者间的距离虽然越来越近，但阿基里斯总是落在乌龟的后面而追不上乌龟。如图 3.14 所示。

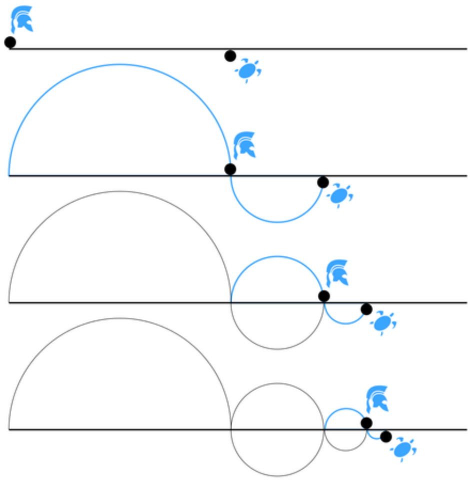


图 3.14: 阿基里斯与乌龟悖论

但这与我们生活中的常识是不相符的。我们在小学数学课上学习过"追及问题"，通常利用相对运动加以解决。令乌龟的速度为 v_1，阿基里斯的速度为 v_2，则阿基里斯相对于乌龟的运动速度为 $v_2 - v_1$。若乌龟在阿基里斯前方距离 s 处，则阿基里斯在 $t = \frac{s}{v_2 - v_1}$ 时与乌龟相遇并接下来超过乌龟。但系统严密的相对运动概念要等到现代科学之父伽利略在两千多年之后才提出。芝诺的推理是如此让人信服，以至于千百年来吸引了无数学者的研究。刘易斯·卡罗尔（Lewis Carrol）、侯世达（Douglas Hofstadter）甚至拿乌龟和阿基里斯作为文学作品中的主人公。

阿基里斯与乌龟悖论的解决通常需要极限的概念。但我们手里的武器——作为分数和的小数已经足够强大到解决它。为了让推理更直观，假设阿基里斯的速度是乌龟的 10 倍，即 $v_2 : v_1 = 10$，乌龟在阿基里斯前 $s = 100$ 米 [15] 处。按照芝诺的推理，阿基里斯需要先跑完 $s = 100$ 米，此时乌龟向前爬了 $\frac{1}{10}s = 10$ 米；接下来阿基里斯要继续跑这 10 米，但乌龟又向前爬了 $\frac{1}{100}s = 1$ 米……所以阿基里斯要跑的距离是：

$$s(1 + \frac{1}{10} + \frac{1}{100} + \cdots) = 100 \times (1 + 0.1 + 0.01 + \cdots) = 100 \times 1.11\cdots = 111.11\cdots$$

假设阿基里斯是个"百米飞人"，只用 9 秒就可以跑完 100 米。但接下来他需要再

[15] 古希腊类似的长度单位叫做 Pous，翻译为"尺"，也叫做"脚长"，是一种基于人体尺度的长度单位，大约合 0.308 米。

用 0.9 秒跑完 10 米,然后再用 0.09 秒跑完 1 米……总共需要的时间是:

$$
\begin{aligned}
9 + 0.9 + 0.09 + \cdots = 9.99\cdots &= 10 \times 0.99\cdots \\
&= 10 \times 1 && \text{据引理 3.4.1} \\
&= 10
\end{aligned}
$$

这样即使按照芝诺的推理,阿基里斯在 10 秒时就可以追上并超越乌龟。

3.4.2 无限循环小数

为了把即约分数 $\frac{b}{a}$ 转化为小数,我们用 b 除以 a。有的能除尽,如 $\frac{1}{2} = 0.5$、$\frac{1}{8} = 0.125$、$\frac{1}{25} = 0.4$,有的除不尽产生循环,如 $\frac{1}{3} = 0.\dot{3} = 0.333\cdots$、$\frac{1}{7} = 0.\dot{1}\dot{4}\dot{2}\dot{8}\dot{5}\dot{7}$。什么时候能除尽?什么时候除不尽?这好像是一个运气问题。实际上我们的运气非常"差",通常是除不尽的。下面的定理揭示了我们要有多幸运才能除尽。

定理 3.4.3. 当且仅当分数 $\dfrac{b}{a} = \dfrac{b}{2^\alpha 5^\beta}$ 时能化成有限小数,其中 α、β 是非负整数。

证明. 令 $n = \max(\alpha, \beta)$,即 α、β 中的较大者。故 $n - \alpha \geq 0, n - \beta \geq 0$。

$$
10^n \frac{b}{a} = \frac{2^n 5^n b}{2^\alpha 5^\beta} = 2^{n-\alpha} 5^{n-\beta} b
$$

一定是整数。所以 $\frac{b}{a}$ 的小数部分有 n 位,即 $0.a_1 a_2 \cdots a_n$ 是有限小数。

反过来,n 位有限小数 $0.a_1 a_2 \cdots a_n$ 展开成分数和为:

$$
\begin{aligned}
\frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n} &= \frac{a_1 10^{n-1} + a_2 10^{n-2} + \cdots + a_n}{10^n} && \text{通分} \\
&= \frac{B}{10^n} && \text{记分子为}B \\
&= \frac{B}{2^n 5^n} = \frac{b}{a} && \text{约分}
\end{aligned}
$$

因为分母 $10^n = 2^n 5^n$ 只有因子 2 和 5 的幂,所以约分后 a 也只有 2 和 5 的幂。 □

有了这个结论,当我们看到一个分数的分母含有 2 和 5 以外的其它因子时,比如 3、7、11……则它一定不能化成有限小数。这个定理的证明还可以扩展到 b 进制小数。如果分母含有任何不能整除 b 的因子,则除不尽,不能转换成 b 进制有限小数。所以 $\frac{5}{12} = 0.41666\cdots$ 尽管不能化成 10 进制有限小数,但可以化成古巴比伦 60 进制小数 0, 25。而 $\frac{5}{14}$ 却不能化成 60 进制有限小数。特别地,形如 $\frac{b}{2^n}$ 以外的任何即约分数都不能化成二进制有限小数,这就造成了电子计算机内部的小数计算的误差。

除此之外的分数就只能化成无限循环小数了(下一章我们将看到分数不可能产生无限不循环小数)。但我们发现有的有一位循环节,如 $\frac{1}{3} = 0.\dot{3}$,有的有两位,如 $\frac{5}{11} = 0.\dot{4}\dot{5}$,有的更长,如 $\frac{1}{7} = 0.\dot{1}\dot{4}\dot{2}\dot{8}\dot{5}\dot{7}$,有的一开始不循环,后面才开始循环,如 $\frac{5}{12} = 0.41\dot{6}$,这里面有什么规律么?

命题 3.4.4. $\dfrac{a}{9}$ 的小数形式是 $0.\dot{a} = 0.aaa\cdots$

证明.

$$
\begin{aligned}
\text{令}x &= 0.\dot{a} = 0.aaa\cdots \\
10x &= a.aaa\cdots \\
10x - x &= a.aaa\cdots - 0.aaa = a \\
9x &= a \\
x &= \frac{a}{9}
\end{aligned}
$$

□

所以 $\frac{1}{3} = \frac{3}{9} = 0.\dot{3} = 0.333\cdots$ 进一步:

命题 3.4.5. 无限循环小数 $0.\dot{a_1}\dot{a_2}\cdots\dot{a_n}$ 的分数形式是 $\dfrac{a_1a_2\cdots a_n}{99\cdots 9} = \dfrac{a_1a_2\cdots a_n}{10^{n+1}-1}$。

我们把此命题的证明留作练习 3.7。可以用 $\frac{1}{7}$ 验证一下:

$$
\begin{array}{r}
1\ 4\ 2\ 8\ 5\ 7 \\
\times \qquad\quad 7 \\
\hline
9\ 9\ 9\ 9\ 9\ 9
\end{array}
$$

故:

$$
0.\dot{1}\dot{4}\dot{2}\dot{8}\dot{5}\dot{7} = \frac{142857}{999999} = \frac{142857}{142857\times 7} = \frac{1}{7}
$$

这个命题使得我们可以轻松看出 $\frac{2}{11}$、$\frac{4}{33}$ 等分数的循环小数表示，例如 $\frac{2}{11} = \frac{2\times 9}{11\times 9} = \frac{18}{99} = 0.\dot{1}\dot{8}$。通过这个命题还可以发现循环节长度的规律。如果 $\frac{a_1a_2\cdots a_n}{99\cdots 9}$ 约分后是 $\frac{b}{a}$, 那么 $ak = 999\cdots 9$。也就是说 $ak+1 = 100\cdots 0$。因此如果 $100\cdots 0$ (n 个 0)除以 a 的余数是 1, 那么 $\frac{b}{a}$ 的循环节的长度最长为 n。循环节的长度仅和分母有关 [16]。我们还需要最后一个定理来回答什么时候开始循环的问题。

定理 3.4.6. 即约真分数 $\dfrac{b}{2^\alpha 5^\beta q}$, 其中 q 不含任何 2、5 为因子。令 $n = \max(\alpha, \beta)$, 则其小数形式从 n 位后开始循环, 循环节的长度是 m, 并且 10^{m+1} 除以 q 余 1。

证明. 我们组合之前的证明过程。首先证明前 n 位不循环:

$$
10^n\frac{b}{2^\alpha 5^\beta q} = \frac{2^n5^nb}{2^\alpha 5^\beta q} = \frac{2^{n-\alpha}5^{n-\beta}b}{q} = X + \frac{b'}{q}
$$

这相当于扩大 10^n 后化带分数, 其中 X 是整数部分, $0 \le X < 10^n$。接下来由于 10^{m+1} 除以 q 余 1, 所以它减 1 后能被 q 整除:

$$
\begin{aligned}
(10^{m+1}-1)\frac{b'}{q} &= kb' = a_1a_2\cdots a_m && \text{是整数} \\
\frac{b'}{q} &= \frac{a_1a_2\cdots a_m}{99\cdots 9} && m\text{个}9
\end{aligned}
$$

由命题 3.4.5 知它是无限循环小数。 □

[16]数论中的费马小定理断言，素数 p 除 10^{p-1} 的余数等于 1。例如 1000000 除以 7 余 1，$\frac{b}{7}$ 的循环节长 6。尽管 $10^{11-1} = 10000000000$ 除以 11 余 1, 但 $\frac{2}{11} = 0.\dot{1}\dot{8}$ 的循环节长 2。我们需要更多的数论知识, 例如欧拉定理, 才能唯一确定循环节的长度。这超出了本书的范围。参见《哈代数论》IX

例如 $\frac{5}{12} = 0.41\dot{6}$, 其中 $12 = 2^2 \times 5^0 \times 3$, 指数 2 和 0 的最大值是 2, 所以 2 位后开始循环; $q = 3$, 循环节长 1。

芝诺（约公元前 490 年 —— 公元前 425 年）, 古希腊哲学家, 生于意大利半岛南部的埃利亚。所以我们常称其为埃利亚的芝诺。关于他的生平, 缺少可靠的文字记载。据传, 他早年是一个自学成才的乡村孩子, 一生经历坎坷, 最终遭到一位暴君的陷害而被拘捕、拷打、直至被处死[26]。芝诺是著名的哲学学派 —— 埃利亚学派的代表人物之一, 这一学派的领袖是芝诺的老师巴门尼德。巴门尼德认为整个世界是个不变的整体, 即"不变的一", 运动、变化与多样性都只是幻象。芝诺以其悖论闻名。他一生曾巧妙地构想出 40 多个悖论, 在流传下来的悖论中以关于运动的四个"无限微妙、无限深邃"的悖论最为著名。他提出这些悖论很可能是为他老师的哲学观点辩护。

除了阿基里斯与乌龟悖论外, 另外三个悖论分别叫做"二分悖论"、"飞矢不动悖论"、"运动场悖论"。二分悖论的主人公是希腊神话中善于疾走的女猎手阿塔兰塔。如果阿塔兰塔想从 A 到 B, 那么她必须先走到 1/2 的位置。同样在此之前, 她必须要到达 1/4 的位置。而为了到达这一位置, 她必须先到达 1/8 的位置……以此类推。由于这样的中点有无限多个, 阿塔兰塔永远也也无法到达目的。芝诺利用这个悖论来说明了运动根本无法发生。练习 3.8 要求读者扩展引理 3.4.1 来解决二分和的问题。

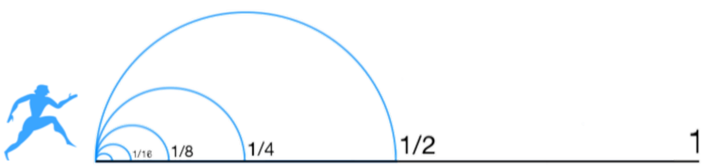


图 3.15: 二分悖论

飞矢不动悖论从另一个角度描述无穷导致运动无法发生。芝诺指出, 任何物体停在相同的位置都不叫运动, 可是飞行的箭矢在任一时刻不也是停在一个地方么? 这样说来, 自然飞失也是不动的。如果说前两个悖论是由于分割空间导致的, 则这个悖论是由对时间的分割导致的。

图 3.16: 飞矢不动悖论

第四个悖论叫作“运动场悖论”。这一悖论主要针对时间原子论的观点，即认为存在最小的不可分割的时间单位。如图 3.17 所示，运动场中有 3 列人。最初他们都首尾对齐。在最小的时间单元内，A 列不动，B 向右移一个单位，而 Γ 向左移动一个单位。容易得知，相对 B 而言，Γ 其实移动了两个单位。这就意味着，应该存在这一让 Γ 相对于 B 移动一个单位的时间。而这一时间应该是最小单位时间的一半。但如果存在不可分割的“时间原子”，那么这两个时间就是相同的，即最小时间和它的一半相等。

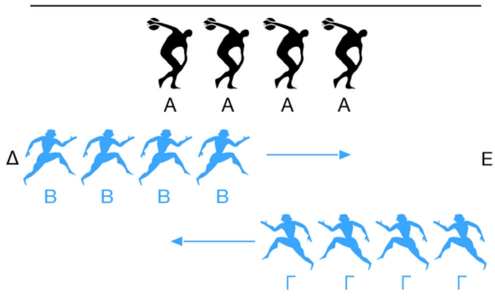


图 3.17: 运动场悖论

芝诺悖论并不复杂，稍加琢磨就能理解。但是导出的结果却出人意料。根据生活中的常识，运动和时间是如此真实，阿基里斯不可能赶不上乌龟。可是驳倒这些悖论却并不容易，从亚里士多德到罗素，从阿基米德到赫尔曼 · 外尔，都对芝诺悖论提出了各种不同的解法[27]。芝诺悖论在当时曾给古希腊人造成深深的困惑。而芝诺悖论所涉及的对时间、空间、无限、连续、运动的看法，也都在极长的历史岁月中困扰着后来的哲学家和数学家。

3.5 数系的扩展

在某种意义上，分数是人类在历史上第一次正式扩展我们的数系。分数看起来和 0、1、2、3……大相径庭。它有两个部分：分子和分母；它抽象了多种不同的东西：小数、百分数、除法、比例 $a : b$，它们本质上都是分数。尤其是后两种，它们展示出 0、1、2、3……不具备的一个特性：$\frac{1}{2} = \frac{2}{4} = \frac{50}{100} = 3 \div 6 = 7 : 14 = \cdots$ 这说明一个值的分数表示是不唯一的。这反映了不同的除法式子可以得出相等的值（或比例）。如果要把分数加入“数”的大家庭，我们就必须解决一系列问题：

问题 1. 哪些分数（比例）是等价的？

问题 2. 怎样比较分数间的大小？怎样比较分数与其它数的大小？分数在数轴上的位置是什么？

问题 3. 分数间的四则运算是怎样的？分数与其它数的四则运算是怎样的？

问题 4. 五大运算定律对分数适用么？加法、乘法的单位元对分数的意义一样么？

先看问题 1, 若 $\frac{b}{a}=\frac{d}{c}$, 两边乘以 ac 有 $bc=ad$, 这个结果被形象地称为"交叉相乘"(见图 3.18)。在小学数学课中, 我们知道比例相等时 $b:a=d:c$ 有外项积等于内项积, 即 $bc=ad$, 这和分数等价是一致的。在数学上, 一个关系 $\sim$ 要成为等价关系需要满足三个条件:

条件 1. 自反性, 即 $a\sim a$。

条件 2. 对称性, 若 $a\sim b$ 则 $b\sim a$。

条件 3. 传递性, 若 $a\sim b, b\sim c$ 则 $a\sim c$。

我们不难验证 $bc=ad$ 是一个等价关系。首先验证自反性, 比较 $\frac{b}{a}$ 和 $\frac{b}{a}$ 自己, 显然 $ab=ab$, 满足自反性。其次验证对称性, 如果 $\frac{b}{a}$ 和 $\frac{d}{c}$ 满足 $bc=ad$, 显然也有 $ad=bc$, 这样 $\frac{d}{c}$ 和 $\frac{b}{a}$ 之间也成立, 满足对称性。最后验证传递性, 若 $\frac{b}{a}$ 和 $\frac{d}{c}$ 满足 $bc=ad$, 并且 $\frac{d}{c}$ 和 $\frac{f}{e}$ 满足 $de=cf$。等量相乘有 $bcde=adcf$。因为 $c\neq 0$ 是分母, 可以消去得 $bde=adf$。若 $d\neq 0$ 也可以消去, 进一步有 $be=af$; 否则 $d=0$, 则 $bc=ad=0=de=cf$, 因此 $b=f=0$, 同样有 $be=af$。满足传递性。

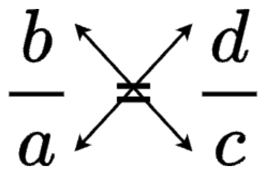


图 3.18: 交叉相乘

有了这个等价关系, 就可以判断任何两个分数(或比例)是否等价。我们可以在众多的分数表示中选取 $\frac{b}{a}$, 且 a,b 没有除 1 以外的公因子(也称为互素, 记作 $(a,b)=1$)的作为代表, 称为"即约分数"。唯一的例外是 0, 我们选取 0 作为 $\frac{0}{a}$ 的代表。这样的代表只有一个, 因为若 $\frac{b}{a}=\frac{d}{c}$, 且各自的分子分母互素, 由等价关系 $bc=ad$, 所以 a 整除 bc。又因为 a 和 b 互素, 所以 a 必然整除 c; 反过来由等价关系 $bc=ad$ 还可以推出 c 整除 ad。又因为 c 和 d 互素, 所以 c 必然整除 a。这样 a 和 c 互相能够整除对方, 所以它们必然相等, 即 $a=c$。同理可证 $b=d$, 这就证明了即约分数的代表是唯一的。

也许有读者问, 为什么要引入这么复杂的等价关系？为什么不直接约分看结果是否一样？古人一开始的确是自然地采用约分来判断相等的, 如同上节《九章算术》中介绍的那样。并且在小学数学课上我们也是这样做的。但随着我们用分数这个武器挑战更大的数和更复杂的问题, 就发现新的情况了。乘法远远比除法更容易。面对 $\frac{8723}{22814}$

和 $\frac{143}{374}$ 这样的分数时, 恐怕很少有人能一眼看出如何约分, 即使借助计算器或使用欧几里得算法。但做乘法就快多了, 用计算器很快就算出 $8723 \times 374 = 3262402$ 并且 $22814 \times 143 = 3262402$, 所以它们相等。中学数学把分数拓展到了分式, 我们发现这个等价关系依然好用。例如 $\dfrac{x^5-1}{x^3-1}$ 和 $\dfrac{1+x+x^2+x^3+x^4}{1+x+x^2}$, 同学们对它们进行因式分解再约分通常会遇到困难 [17]。但多项式乘法却容易得多:

$$\begin{aligned}
(x^3-1)(1+x+x^2+x^3+x^4) &= \cancel{x^3}+\cancel{x^4}+x^5+x^6+x^7-1-x-x^2-\cancel{x^3}-\cancel{x^4}\\
&= x^7+x^6+x^5-x^2-x-1\\
(x^5-1)(1+x+x^2) &= x^5+x^6+x^7-1-x-x^2\\
&= x^7+x^6+x^5-x^2-x-1
\end{aligned}$$

所以这两个分式是相等的。这就是数学家们为何要抽象出等价关系的概念, 并精挑细选地用“交叉相乘”来判断分数、比例、分式相等。

对于问题 2, 我们可以利用《九章算术》中的办法, 通过 $\frac{b}{a}-\frac{d}{c}$ 与 0 的关系来判断大小。这个过程相当于通分后比较新的分子间大小。我们可以把任何整数 n 转化为 $\frac{n}{1}$, 这样所有的大小关系就转化为分数之间的比较。这也意味着任何一个分数都对应到数轴上的一点。对于任何多个数, 包括整数、分数、小数都可以对应到数轴上的各自位置。然后从左到右整理出它们的序关系。

对于问题 3,《九章算术》给出了分数间的四则运算规则。概括为加减先通分, 再加减分子, 最后约分; 乘法把分子分母各自相乘、约分。一个数 $n \neq 0$ 的倒数是 $\frac{1}{n}$, 因为 $n\frac{1}{n}=\frac{n}{n}=1$。如果 n 是分数 $\frac{b}{a}$, 它的倒数是 $\frac{a}{b}$, 因为 $\frac{b}{a}\frac{a}{b}=1$。分数本质上相当于除法, 我们可以通过倒数在分数 —— 除法 —— 乘法间转换:

$$\frac{b}{a} = b \div a = b \times 1 \div a = b \times \frac{1}{a}$$

以及:

$$\frac{b}{a} \div \frac{d}{c} = \frac{b}{a} \times 1 \div \frac{d}{c} = \frac{b}{a} \times \frac{c}{d}$$

或:

$$\frac{b}{a} \div \frac{d}{c} = \frac{\frac{b}{a}}{\frac{d}{c}} = \frac{bc}{ad} \qquad \text{上下同} \times ac$$

同样, 任何整数 n 可以转化为 $\frac{n}{1}$, 从而将整数与分数的四则运算转化为分数间的四则运算。

对于问题 4, 我们发现将分数引入数系后, 任何数都可表示为分数: 整数 n 包括 0, 表示为 $\frac{n}{1}$; 分数、有限小数、无限循环小数表示为 $\frac{b}{a}, a \neq 0$。这个数系称做**有理数**, 记作

[17] 其中一个方法是: 看出 $1^n-1=0$, 所以 $x-1$ 是 x^n-1 的一个因子。接下来用多项式长除法得出另一个因子。

$\mathbb{Q}$。我们将在第 4 章解释这个名字的含义。在整数 $\mathbb{Z}$ 中, 每个数 x 都有相反数(加法的逆元)$-x$, 但只有 ± 1 有倒数(乘法的逆元)。在有理数 $\mathbb{Q}$ 中, 每个不等于 0 的数 x 都有倒数(乘法的逆元) $\frac{1}{x}$。自然数 $\mathbb{N}$ 对加法减运算不是封闭的, 例如 $1-3=-2 \notin \mathbb{N}$。但加减运算在整数 $\mathbb{Z}$ 内是封闭的, 任何两个整数 m, n, 总有 $m \pm n \in \mathbb{Z}$。整数 $\mathbb{Z}$ 对四则运算不是封闭的, 因为 $1 \div 2 = \frac{1}{2} = 0.5 \notin \mathbb{Z}$。但四则运算在有理数 $\mathbb{Q}$ 内是封闭的, 任何两个有理数 p, q, 总有 $p \pm q, pq, \frac{p}{q} (q \neq 0)$ 都在 $\mathbb{Q}$ 内。

可以验证五大运算定律对于分数都成立, 我们这里验证加法的结合率、乘法对加法的分配律, 其余三个留作练习 3.9。

证明. 分数加法的结合律:

$$
\begin{aligned}
(\frac{b}{a}+\frac{d}{c})+\frac{f}{e} &= \frac{bc+ad}{ac}+\frac{f}{e} = \frac{bce+ade+acf}{ace} \\
&= \frac{bce+(ade+acf)}{ace} && \text{对分子和用结合律} \\
&= \frac{b\cancel{ce}}{a\cancel{ce}}+\frac{ade+acf}{ace} \\
&= \frac{b}{a}+(\frac{\cancel{a}d\cancel{e}}{\cancel{a}c\cancel{e}}+\frac{\cancel{ac}f}{\cancel{ac}e}) \\
&= \frac{b}{a}+(\frac{d}{c}+\frac{f}{e})
\end{aligned}
$$

□

证明. 我们只证明左侧的分数分配律, 右侧的证明可以利用分数乘法交换律。

$$
\begin{aligned}
\frac{b}{a}(\frac{d}{c}+\frac{f}{e}) &= \frac{b}{a}\frac{de+cf}{ce} \\
&= \frac{b(de+cf)}{ace} = \frac{bde+bcf}{ace} && \text{分子用分配律} \\
&= \frac{bd\cancel{e}}{ac\cancel{e}}+\frac{b\cancel{c}f}{a\cancel{c}e} \\
&= \frac{b}{a}\frac{d}{c}+\frac{b}{a}\frac{f}{e}
\end{aligned}
$$

□

让我们回到毕达哥拉斯追寻天籁之音的故事。发现了分数导致和谐的乐音后, 如何调整里尔琴上的七根琴弦呢? 如果加上第八根琴弦, 那么它应该是第一根琴弦的一半, 这样才能跨越八度音程。毕达哥拉斯发现纯五度音程 $3:2$ 叠加一个纯四度音程 $4:3$ 恰好是一个八度音程, 因为: $\frac{3}{2} \times \frac{4}{3} = 2$。他用纯五度音程的比例 $3:2$ 作为倍数, 不断提高音调, 得到一组等比数列:

$$
(\frac{3}{2})^{-1}, (\frac{3}{2})^{0}, (\frac{3}{2})^{1}, (\frac{3}{2})^{2}, (\frac{3}{2})^{3}, (\frac{3}{2})^{4}, (\frac{3}{2})^{5} = \frac{2}{3}, 1, \frac{3}{2}, \frac{9}{4}, \frac{27}{8}, \frac{81}{16}, \frac{243}{32}
$$

其中后面 4 个数大于 2, 超过了八度音程。因此毕达哥拉斯把它们每个数都除以 2 或 4, 找到其对应的前一个八度音程的数, 然后重新从小到大排列得到:

$$
1, \frac{9}{8}, \frac{81}{64}, \frac{4}{3}, \frac{3}{2}, \frac{27}{16}, \frac{243}{128}, \frac{2}{1}
$$

它表示的是音阶中每个音与最低音的比例关系。为了知道每个音的相对关系，于是让相邻的两个数相除，又得到：

$$\frac{9}{8}, \frac{9}{8}, \frac{256}{243}, \frac{9}{8}, \frac{9}{8}, \frac{9}{8}, \frac{256}{243}$$

较大的 9:8 叫全音，较小的 256:243 叫半音。这个列表就是“全、全、半、全、全、全、半”，在音乐上叫做“大调”。通常表现欢快的，阳刚的主题。而音乐上的“小调”，对应的列表移动一下是“全、半、全、全、半、全、全”。通常表现含蓄的、柔美的主题。可是有没有可能使得每个比例相同，得到完美的“平均律”呢？这超出了毕达哥拉斯所能想到的所有设计。为了达到这样的天籁之音，我们需要在一个八度音程 $2:1$ 中设置 5 个全音，2 个半音，共 $5 \times 2 + 2 = 12$ 个比例，使得：

$$1 = \alpha^0, \alpha^1, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6, \alpha^7, \alpha^8, \alpha^9, \alpha^{10}, \alpha^{11}, \alpha^{12} = 2$$

显然 $\alpha = \sqrt[12]{2}$，但它不是一个分数！让我们沿着数的旅程继续追寻天籁之音吧。

练习 3

3.1. ★ 若 p 为大于 2 的素数，证明 $\frac{2}{p}$ 可唯一分解为两个埃及分数 $\frac{2}{p} = \frac{1}{a} + \frac{1}{b} (a \neq b)$。提示：这个式子等价于 $(2a - p)(2b - p) = p^2$。

3.2. 证明埃尔德什——斯特劳斯猜想对所有偶数 n 成立。

3.3. 实际上我们只要证明埃尔德什——斯特劳斯猜想对于所有素数成立即可。这是为什么？

3.4. 证明埃尔德什——斯特劳斯猜想对所有 $4k + 3$ 型的素数成立。提示：利用练习 3.1 的结论。

3.5. ★ 找出三个儿子按照 $\frac{1}{a}$、$\frac{1}{b}$、$\frac{1}{c}$ 分配 n 个动物的所有可能数值和比例。要求大儿子分得的财产多于二儿子，二儿子分得的财产多于小儿子。恰好借来一只动物后可以分好并余下一只动物。如果你会编程，可以用计算机枚举出所有解。但此题要求手工找出所有可能的 n, a, b, c。提示：本题本质是求方程 $\frac{n}{n+1} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}, a < b < c$ 的所有正整数解。三儿子至少分得一只动物，那么 n 的最小值是多少？你能据此推出 a 不可能超过 3，因而只能是 2 么？

3.6. ★ 证明任何数值 x 的小数表示是唯一的，其中 $0 < x < 1$。

3.7. 证明无限循环小数 $0.\dot{a_1}\dot{a_2}\cdots\dot{a_n}$ 的分数形式是 $\frac{a_1a_2\cdots a_n}{99\cdots9} = \frac{a_1a_2\cdots a_n}{10^{n+1}-1}$。

3.8. 利用分数和引理 3.4.1 解释：如果阿塔兰塔先跑完全程的 $\frac{1}{2}$，接下来再跑完剩余的一半，即 $\frac{1}{4}$，然后不断跑完剩余路程的一半，则她最终可以跑完全程。提示：考虑二进制分数和。

3.9. 验证分数的加法交换律、乘法交换律，乘法结合率。

答案：65 页

附录 A　参考答案

答案 1

1.1. ①、③、④

1.2. 将十进制数 123 转换为(a)玛雅文;(b)古巴比伦文;(c)计算机二进制表示。

玛雅文:

$$
\begin{aligned}
123 \div 20 &= 6 \cdots 3 \\
6 \div 20 &= 0 \cdots 6 \\
120 &= 6(20) + 3
\end{aligned}
$$

$\underline{\circ}$

$\circ \circ \circ$

从上到下 6、3

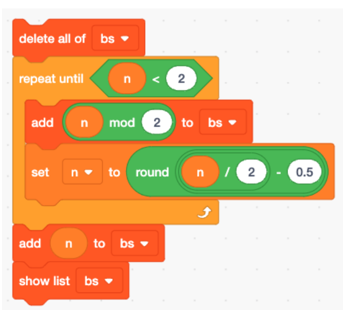
古巴比伦文:YY YYY

$$
\begin{aligned}
123 \div 60 &= 2 \cdots 3 \\
2 \div 20 &= 0 \cdots 2 \\
120 &= 2(60) + 3
\end{aligned}
$$

计算机二进制表示:1111011

$$
\begin{array}{ccc}
123 \div 2 = 61 \cdots 1 & 61 \div 2 = 30 \cdots 1 & 30 \div 2 = 15 \cdots 0 \\
15 \div 2 = 7 \cdots 1 & 7 \div 2 = 3 \cdots 1 & 3 \div 2 = 1 \cdots 1 \\
1 \div 2 = 0 \cdots 1 & &
\end{array}
$$

1.3. 编程实现二进制到十进制的相互转换。



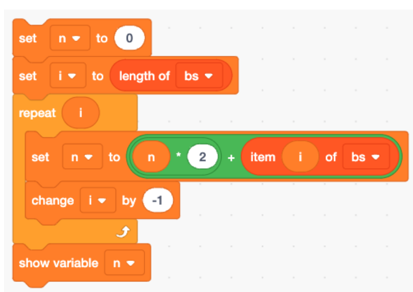


图 1.15: Scratch 例子程序:十进制 n 与二进制 bs 相互转换,左:$10 \to 2$;右:$2 \to 10$

说明：round 是四舍五入。对 $n/2$ - 0.5 四舍五入相当于对 $n/2$ 取整，求得 n 除以 2 的除数。bs 中的二进制数字低位在前，高位在后。与 Scratch 的循环不同，下面的 Haskell 例子给出了递归实现。其中的二进制数字高位在前、低位在后。

```
bin x = if x < 2 then [x] else (bin (x `div` 2)) ++ [x `mod` 2]
dec = foldl (λn x → n*2 + x) 0
```

1.4. 证明一个数的位值制表示是唯一的。

证明. 假设一个数还有另一个表示。如果它们的位数不同，我们在较短的前面添加 0，使它们一样长。令补 0 后的表示分别为 $a_m a_{m-1} \cdots a_1 a_0$ 和 $c_m c_{m-1} \cdots c_1 c_0$。按照式 (1.1) 计算的值相等，即：

$$a_m b^m + a_{m-1} b^{m-1} + \cdots + a_1 b + a_0 = c_m b^m + c_{m-1} b^{m-1} + \cdots + c_1 b + c_0$$

将 a_0 和 c_0 移到右边，剩下的移到左边：

$$(a_m - c_m) b^m + (a_{m-1} - c_{m-1}) b^{m-1} + \cdots + (a_1 - c_1) b = c_0 - a_0$$

左边可以被 b 整除，所以 $c_0 - a_0$ 也可以被 b 整除。由于 c_0、a_0 都只能是 0 到 $b-1$ 的整数，所以它们的差只能是 $1-b, \ldots, -1, 0, 1, \ldots b-1$ 中的一个。但其中只有 0 能被 b 整除。所以 $c_0 - a_0 = 0$，即 $c_0 = a_0$。
接下来 $(a_m - c_m) b^m + (a_{m-1} - c_{m-1}) b^{m-1} + \cdots + (a_1 - c_1) b = 0$。由于 $b \neq 0$，两边除以 b，然后将 a_1 和 c_1 移动到左边：

$$(a_m - c_m) b^m + (a_{m-1} - c_{m-1}) b^{m-1} + \cdots + (a_2 - c_2) b = c_1 - a_1$$

同样可以推出 $c_1 = a_1$。重复这个步骤 m 次，最后得到 $a_m = c_m$。这就证明了两种表示是一样的，即位值制表示是唯一的。 □

答案 2

2.1. 什么情况下应从 1 开始数数，什么情况下应从 0 开始数数？
考虑计数的情形是序数还是基数，是否包含到起点（原点）距离的含义，是否可能有负数的意义等因素。

2.2. 定义数偶的减法和除法

(a) 减法相当于加上相反数，而数偶 (a, b) 的相反数是 (b, a)。定义减法：$(a, b) - (c, d) = (a, b) + (d, c) = (a + d, b + c)$。

(b) 为了定义除法，注意到两个性质：(1) (ka, kb) 相当于把 (a, b) 扩大 k 倍，$(\frac{1}{k}a, \frac{1}{k}b)$ 缩小到 $\frac{1}{k}$。(2) $(-a, -b) = (b, a)$。利用它们得到：

$$(a, b)/(c, d) = \begin{cases} c > d: & (\dfrac{a}{c-d}, \dfrac{b}{c-d}) \\ c < d: & (\dfrac{b}{d-c}, \dfrac{a}{d-c}) \end{cases}$$

其中 $c \neq d$。

2.3. 验证数偶的加法乘法满足交换律、结合律、分配律

(a) 交换率:

$$\begin{aligned}(a,b)+(c,d) &= (a+c,b+d) && \text{数偶加法定义}\\ &= (c+a,d+b) && \text{自然数的加法交换律}\\ &= (c,d)+(a,b) && \text{反向用数偶加法定义}\end{aligned}$$

$$\begin{aligned}(a,b)\cdot(c,d) &= (ac+bd,ad+bc) && \text{数偶乘法定义}\\ &= (ca+db,da+cb) && \text{自然数乘法交换律}\\ &= (c,d)\cdot(a,b) && \text{反向用数偶乘法定义}\end{aligned}$$

(b) 结合律:略

(c) 分配律:我们只给出左侧分配律, 略去右侧。

$$\begin{aligned}&(a,b)\cdot((c,d)+(e,f))\\ =&(a,b)\cdot(c+e,d+f)\\ =&(a(c+e)+b(d+f),a(d+f)+b(c+e))\\ =&(ac+bd+ae+bf,ad+bc+af+be)\\ =&(ac+bd,ad+bc)+(ae+bf,af+be) && \text{反向用数偶加法定义}\\ =&(a,b)\cdot(c,d)+(a,b)\cdot(e,f) && \text{反向用数偶乘法定义}\end{aligned}$$

2.4. 用乘法对加法的分配律证明 $(-1)\cdot(-1)=1$

证明.

$$\begin{aligned}(-1)\times(-1) &= (-1)\times(-1)+0 && \text{0 加上任何数不变}\\ &= (-1)\times(-1)+(-1+1) && -1\text{和 1 互为相反数}\\ &= (-1)\times(-1)+(-1)\times 1+1 && \text{1 乘以任何数不变}\\ &= (-1)\times(-1+1)+1 && \text{反向用乘法分配律}\\ &= (-1)\times 0+1=0+1=1 && \square\end{aligned}$$

2.5. 证明一个数的相反数是唯一的。

证明. 假设 b 是 a 的一个相反数, 我们要证明 $b=-a$:

$$\begin{aligned}0 &= b+a && \text{相反数}\\ 0-a &= b+a-a && \text{两边}-a\\ -a &= b+(a-a)=b+0=b && \square\end{aligned}$$

2.6. 说明加法、乘法结合律的数轴意义。

加法:数轴先平移 $a+b$ 再平移 c 等效于先平移 a 再平移 $b+c$。乘法:数轴先缩放 ab 再缩放 c 等效于先缩放 a 再缩放 bc。

2.7. 考虑 $x \rightsquigarrow 2x+1$, 如果先平移再缩放, 应先向哪个方向平移多少? 再缩放几倍? 一般地, $x \rightsquigarrow ax+b$ 应如何变换?

注意到 $2x+1 = 2(x+\frac{1}{2})$, 所以先向左平移 $\frac{1}{2}$ 再放大 2 倍。一般地, 由于 $ax+b=a(x+\frac{b}{a})$, 故先平移 $-\frac{b}{a}$(若 $\frac{b}{a}>0$ 则向左, 否则向右)再放大 a 倍。

2.8. 证明单位元是唯一的。

证明. 假设存在另一单位元 e', 使得对任意元素 a 满足 $e' \odot a = a \odot e' = a$。我们要证明 $e=e'$:

$$
\begin{aligned}
e &= e \odot e' && e'\text{是单位元} \\
&= e' && e\text{是单位元}
\end{aligned}
$$

□

2.9. 定义 0 的后继为 1, 证明对于任何自然数都有 $a \cdot 1 = a$

证明. 利用附录 C 证明的结论:$0+a=a$:

$$
\begin{aligned}
a' \cdot 1 &= a' \cdot 0' && \text{定义 0 的后继为 1} \\
&= a' \cdot 0 + a' && \text{乘法定义规则二} \\
&= 0 + a' && \text{乘法定义规则一} \\
&= a'
\end{aligned}
$$

□

2.10. 证明乘法分配律。

证明. 只证明左侧的分配律 $c(a+b)=ca+cb$。对 b 使用数学归纳法, $b=0$ 时:

$$
\begin{aligned}
c(a+0) &= ca && \text{加法规则一} \\
&= ca + 0 && \text{反向用加法规则一} \\
&= ca + c0 && \text{反向用乘法规则一}
\end{aligned}
$$

递推假设 $c(a+b)=ca+cb$, 接下来证明 $c(a+b')=ca+cb'$

$$
\begin{aligned}
c(a+b') &= c(a+b)' && \text{加法规则二} \\
&= c(a+b)+c && \text{乘法规则二} \\
&= ca+cb+c && \text{递推假设} \\
&= ca+(cb+c) && \text{加法结合律} \\
&= ca+cb' && \text{反向用乘法规则二}
\end{aligned}
$$

□

2.11. 证明乘法结合律和交换律。

我们只证明乘法结合律 $(ab)c = a(bc)$, 乘法交换律的证明则给出一个提纲。

证明. 对 c 使用数学归纳法, $c = 0$ 时:

$$
\begin{aligned}
(ab)0 &= 0 && \text{乘法规则一} \\
&= a0 && \text{反向用乘法规则一} \\
&= a(b0) && \text{反向用乘法规则一}
\end{aligned}
$$

递推假设 $(ab)c = a(bc)$, 接下来要证明 $(ab)c' = a(bc')$

$$
\begin{aligned}
(ab)c' &= (ab)c + ab && \text{乘法规则二} \\
&= a(bc) + ab && \text{递推假设} \\
&= a(bc + b) && \text{上题证明的分配律} \\
&= a(bc') && \text{反向用乘法规则二}
\end{aligned}
$$

□

证明乘法交换律可以分为三步, 都使用数学归纳法。首先证明 $1a = a$, 然后再证明右侧的分配律 $(a + b)c = ac + bc$, 最后再证明交换律。

2.12. 如何利用皮亚诺公理验证 $3 + 147 = 150$?

我们先看看经典的 $2 + 2 = 4$ 是怎么证明的:

$$
\begin{aligned}
2 + 2 &= 0'' + 0'' && \text{2 是 0 的两次后继} \\
&= (0'' + 0')' && \text{加法定义规则二} \\
&= ((0'' + 0)')' && \text{加法定义规则二} \\
&= ((0'')')' && \text{加法定义规则一} \\
&= 0'''' = 4 && \text{0 的 4 次后继}
\end{aligned}
$$

显然用这个方法证明 $3 + 147 = 150$ 太冗长了, 我们可以用先前证明的加法交换律证明 $147 + 3 = 150$ 会容易一些。另一个方法是通过数学归纳法证明 $3 + a = a'''$。

答案 3

3.1. 若 p 为大于 2 的素数, 证明 $\frac{2}{p}$ 可唯一分解为两个埃及分数 $\frac{2}{p} = \frac{1}{a} + \frac{1}{b} (a \neq b)$。

证明. 上式等价于 $(2a-p)(2b-p)=p^2$, 我们可验证如下:

$$
\begin{aligned}
\frac{2}{p} = \frac{1}{a} + \frac{1}{b} &= \frac{a+b}{ab} && \text{通分} \\
2ab &= pa + pb && p, a, b\text{是分母, 不为}0 \\
4ab - 2pa - 2pb &= 0 && \text{移项} \times 2 \\
4ab - 2pa - 2pb + p^2 &= p^2 && \text{两边} + p^2 \\
(2a-p)(2b-p) &= p^2 && \text{因式分解}
\end{aligned}
$$

因为 p 是素数, p^2 的因子只有 1、p、p^2。由 $a \neq b$ 可以排除掉 p, 所以 $2a-p=1$, $2b-p=p^2$, 即:

$$a = \frac{p+1}{2}, b = \frac{p(p+1)}{2}$$

因为 p 是奇素数, 所以 a 是整数。因此:

$$\frac{2}{p} = \frac{1}{\frac{1}{2}(p+1)} + \frac{1}{\frac{1}{2}p(p+1)}$$ □

3.2. 令 $n=2m$, 则 $\frac{4}{2m} = \frac{2}{m}$, 这就是定理 3.1.2。

3.3. 注意到:

$$\frac{4}{mp} = \frac{1}{ma} + \frac{1}{mb} + \frac{1}{mc}$$

3.4. 证明. 利用练习 3.1 的结论

$$
\begin{aligned}
\frac{2}{p} &= \frac{1}{\frac{1}{2}(p+1)} + \frac{1}{\frac{1}{2}p(p+1)} \\
\frac{4}{p} &= \frac{2}{\frac{1}{2}(p+1)} + \frac{2}{\frac{1}{2}p(p+1)} && \text{左右} \times 2 \\
\frac{4}{4k+3} &= \frac{2}{\frac{1}{2}(4k+4)} + \frac{2}{\frac{1}{2}(4k+3)(4k+4)} && \text{带入}p = 4k+3 \\
&= \frac{1}{k+1} + \frac{1}{(4k+3)(k+1)} \\
&= \frac{1}{k+2} + \frac{1}{(k+1)(k+2)} + \frac{1}{(4k+3)(k+1)} && \text{由定理 3.1.1}
\end{aligned}
$$ □

3.5. 我们先给出一个纯数学解法, 然后再给出用计算机穷举的解法。读者朋友们可以加以对比。本题本质上是求方程

$$\frac{n}{n+1} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}, a < b < c$$

的所有正整数解。即借来一只后, 把 $n+1$ 只动物分成三份 $\frac{n+1}{a}$、$\frac{n+1}{b}$、$\frac{n+1}{c}$。它们的和恰好是 n, 因而可以把剩余的那一只归还。为了简洁, 我们令 $n' = n+1$, 表示借来一只动物后的数量, 把上述方程改写为:

$$\frac{n'-1}{n'} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \tag{3.5}$$

最后只要把 n' 减一就可以了。借来一只动物后能恰好分配，意味着 a, b, c 都能整除 n'。我们先证明一个重要的结论：大儿子只能分得 $\frac{1}{2}$ 的财产，即 a 只能等于 2。

证明. 若 $a \geq 3$，由于 $a < b < c$，所以：

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \leq \frac{1}{3} + \frac{1}{4} + \frac{1}{5} = \frac{47}{60}$$

小儿子至少分得 1 只；二儿子分得的比小儿子多，所以至少分得 2 只；同理，大儿子至少分得 3 只动物，再加上借来的一只。所以 $n' > 3 + 2 + 1 = 6$。于是：

$$\frac{n'-1}{n'} = 1 - \frac{1}{n'} > 1 - \frac{1}{6} = \frac{5}{6} = \frac{50}{60} > \frac{47}{60} \geq \frac{1}{a} + \frac{1}{b} + \frac{1}{c}$$

也就是说 $a \geq 3$ 时，方程式 (3.5) 不可能成立。所以 a 只能是 2。 □

问题于是就化简为，求方程

$$\frac{n'-1}{n'} = \frac{1}{2} + \frac{1}{b} + \frac{1}{c}$$

的正整数解。我们接下来证明 b 只能是 3 和 4，从而进一步化简问题：

证明. 如果 $b \geq 5$，由于 $b < c$，故：

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \leq \frac{1}{2} + \frac{1}{5} + \frac{1}{6} = \frac{13}{15} \approx 0.87$$

由于 a, b, c 都能整除 n'，而 $a = 2$，所以 n' 必然是偶数。我们上面推出 $n' > 6$，比 6 大的第一个偶数是 8。因此：

$$\frac{n'-1}{n'} = 1 - \frac{1}{n'} \geq 1 - \frac{1}{8} = \frac{7}{8} = 0.875 > 0.87 \geq \frac{1}{a} + \frac{1}{b} + \frac{1}{c}$$

也就是说 $b \geq 5$ 时，方程式 (3.5) 不可能成立。所以 b 只能是 3 或 4。 □

情况 1) $a = 2, b = 3$

$$\begin{aligned}
\frac{n'-1}{n'} &= \frac{1}{2} + \frac{1}{3} + \frac{1}{c} \\
1 - \frac{1}{n'} &= \frac{5}{6} + \frac{1}{c} \\
\frac{1}{6} &= \frac{1}{n'} + \frac{1}{c}
\end{aligned}$$

由于 $a = 2, b = 3$ 都整除 n'，所以 6 整除 n'。不妨设 $n' = 6k$，其中 k 为大于 1 的整数(因为我们前面确定了 n' 最小是 8)。代入上式，整理得：

$$k = \frac{c}{c-6} \geq 2$$

解不等式 $\frac{c}{c-6} \geq 2$ 得 $c \leq 12$。同时分母 $c-6$ 是正数，所以 $c > 6$。这样 c 的值只可能是 $7, 8, 9, \cancel{10}, \cancel{11}, 12$。其中 $10, 11$ 不可能，因为此时 $k = \frac{c}{c-6}$ 不是整数。我们得到 4 个解：

$$
\begin{array}{ll}
a=2, b=3, c=7, n'=42 & 42\frac{1}{2} + 42\frac{1}{3} + 42\frac{1}{7} = 21 + 14 + 6 = 41 \\
a=2, b=3, c=8, n'=24 & 24\frac{1}{2} + 24\frac{1}{3} + 24\frac{1}{8} = 12 + 8 + 3 = 23 \\
a=2, b=3, c=9, n'=18 & 18\frac{1}{2} + 18\frac{1}{3} + 18\frac{1}{9} = 9 + 6 + 2 = 17 \\
a=2, b=3, c=12, n'=12 & 12\frac{1}{2} + 12\frac{1}{3} + 12\frac{1}{12} = 6 + 4 + 1 = 11
\end{array}
$$

情况 2) $a = 2, b = 4$

$$
\begin{aligned}
\frac{n'-1}{n'} &= \frac{1}{2} + \frac{1}{4} + \frac{1}{c} \\
1 - \frac{1}{n'} &= \frac{3}{4} + \frac{1}{c} \\
\frac{1}{4} &= \frac{1}{n'} + \frac{1}{c}
\end{aligned}
$$

由于 $a = 2, b = 4$ 都整除 n'，所以 4 整除 n'。不妨设 $n' = 4k$，其中 $k \geq 2$ 为整数(同理 n' 最小是 8)。代入上式，整理得：

$$
k = \frac{c}{c-4} \geq 2
$$

解不等式 $\frac{c}{c-2} \geq 2$ 得 $c \leq 8$。同时分母 $c-4$ 是正数，所以 $c > 4$。这样 c 的值只可能是 $5, 6, \cancel{7}, 8$。其中 7 不可能，因为此时 $k = \frac{c}{c-4}$ 不是整数。我们得到 3 个解：

$$
\begin{array}{ll}
a=2, b=4, c=5, n'=20 & 20\frac{1}{2} + 20\frac{1}{4} + 20\frac{1}{5} = 10 + 5 + 4 = 19 \\
a=2, b=4, c=6, n'=12 & 12\frac{1}{2} + 12\frac{1}{4} + 12\frac{1}{6} = 6 + 3 + 2 = 11 \\
a=2, b=4, c=8, n'=8 & 8\frac{1}{2} + 8\frac{1}{4} + 8\frac{1}{8} = 4 + 2 + 1 = 7
\end{array}
$$

注意到两种情况都有分 11+1 只动物的可能，但分配比例不同。总共有 7 个解。

用计算机求解时，可以让机器枚举所有可能的 a, b, c。判断其倒数和是否满足 $\frac{n+1}{n}$ 的形式。显然，a 从 2 开始，依次尝试 2、3、4……并且由于 $a < b < c$，所以我们从 $a + 1$ 开始枚举 b，从 $b + 1$ 开始枚举 c。接下来我们需要估计 a, b, c 的一个上限，使得枚举的过程可以结束。因为 $a < b < c$，所以 c 的上限一定也是 a, b 的上限。此外，三儿子最少分到 1 只动物，所以 $c \leq n'$。从式 (3.5) 可以推出：

$$
\begin{aligned}
1 - \frac{1}{n'} &= \frac{1}{a} + \frac{1}{b} + \frac{1}{c} \\
1 - \frac{1}{n'} - \frac{1}{c} &= \frac{1}{a} + \frac{1}{b} \le \frac{1}{2} + \frac{1}{3} = \frac{5}{6} \\
\frac{1}{6} &\le \frac{1}{n'} + \frac{1}{c} \le \frac{1}{c} + \frac{1}{c} && \text{三儿子至少分得一只, 故} \frac{1}{n'} \le \frac{1}{c} \\
\frac{1}{6} &\le \frac{2}{c} \Rightarrow c \le 12
\end{aligned}
$$

因此枚举的上限是 12。其次从 $1 - \frac{1}{n'} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}$ 还可以解出 $n' = \frac{abc}{abc - ab - bc - ac}$，如果其分母不为 0，且 a, b, c 都整除 n'，则找到了一组解。图 3.19 给出了例子程序。

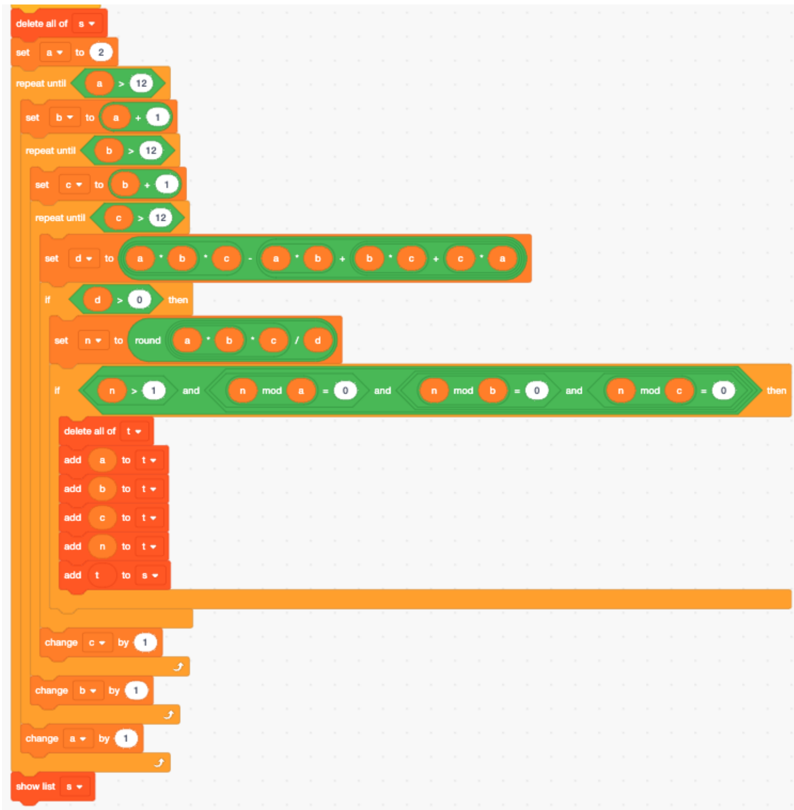


图 3.19: 三子继承问题的 Scratch 程序

下面是对应的 Haskell 例子程序。

```
[(a, b, c, n) | a ← [2..10], b ← [a+1..11], c ← [b+1..12],
     let d = a∗b∗c - a∗b - b∗c - a∗c, d > 0,
     let n = a∗b∗c `div` d, n > 1,
         n `mod` a == 0, n `mod` b == 0, n `mod` c == 0]
```

3.6. 证明. 用反证法，令：

$$
\frac{a_1}{10} + \frac{a_2}{100} + \cdots = \frac{b_1}{10} + \frac{b_2}{100} + \cdots \tag{3.6}
$$

设 a_n 和 b_n 是第一对不相等的。$|a_n - b_n| \geq 1$。于是:

$$\begin{aligned}\frac{a_1}{10}+\frac{a_2}{100}+\cdots-(\frac{b_1}{10}+\frac{b_2}{100}+\cdots) &\geq \frac{1}{10^n}-(\frac{|a_{n+1}-b_{n+1}|}{10^{n+1}}+\frac{|a_{n+2}-b_{n+2}|}{10^{n+2}}+\cdots) \\ &\geq \frac{1}{10^n}-(\frac{9}{10^{n+1}}+\frac{9}{10^{n+2}}+\cdots) \\ &= \frac{1}{10^n}-\frac{1}{10^n}(0.99\cdots)=0\end{aligned}$$

这与式 (3.6) 矛盾。 □

3.7. 证明.

$$\begin{aligned}\text{设}x &= 0.\dot{a_1}\dot{a_2}\cdots\dot{a_n} \\ 10^n x &= a_1a_2\cdots a_n.\dot{a_1}\dot{a_2}\cdots\dot{a_n} \\ 10^n x - x &= a_1a_2\cdots a_n \\ x &= \frac{a_1a_2\cdots a_n}{10^n-1} = \frac{a_1a_2\cdots a_n}{99\cdots9}\end{aligned}$$

□

3.8. 十进制无限循环小数 $0.99\cdots = 1$, 对应的二进制无限循环小数 $0.11\cdots = 1$。证明与十进制类似:

证明.

$$x = 0.11\cdots \qquad 2x = 1.11\cdots$$

$$x = 2x - x = 1$$

□

阿塔兰塔先跑到 $\frac{1}{2}$ 处, 即二进制 0.1 处; 再跑到 $\frac{1}{4}$ 处, 即二进制 0.01 处……她跑过的路程和等于二进制小数和:

$$0.1 + 0.01 + 0.001 + \cdots = 0.11\cdots = 1$$

3.9. 加法交换律:

$$\begin{aligned}\frac{b}{a}+\frac{d}{c} &= \frac{bc+ad}{ac} && \text{通分} \\ &= \frac{ad+bc}{ac} && \text{分子用整数加法交换律} \\ &= \frac{ad}{ac}+\frac{bc}{ac}=\frac{d}{c}+\frac{b}{a}\end{aligned}$$

乘法交换律:

$$\frac{b}{a}\times\frac{d}{c}=\frac{bd}{ac}=\frac{db}{ca}=\frac{d}{c}\times\frac{b}{a}$$

乘法结合律:

$$\frac{b}{a}\times\frac{d}{c}\times\frac{f}{e}=\frac{bdf}{ace}=\frac{b(df)}{a(ce)}=\frac{b}{a}\frac{df}{cd}=\frac{b}{a}\times(\frac{d}{c}\times\frac{f}{e})$$

附录 B　希腊字母表

大写	小写	读音	大写	小写	读音
A	α	Alpha	N	ν	Nu
B	β	Beta	Ξ	ξ	Xi
Γ	γ	Gamma	O	o	Omicron
Δ	δ	Delta	Π	π	Pi
E	ϵ	Epsilon	P	ρ	Rho
Z	ζ	Zeta	Σ	σ	Sigma
H	η	Eta	T	τ	Tau
Θ	θ	Theta	Υ	υ	Upsilon
I	ι	Iota	Φ	ϕ	Phi
K	κ	Kappa	X	χ	Chi
Λ	λ	Lambda	Ψ	ψ	Psi
M	μ	Mu	Ω	ω	Omega

表 B.1: 希腊字母表

附录 C 加法交换律的证明

为了证明加法交换律 $a+b=b+a$, 我们首先证明三个结论。一、对于任何自然数都有:

$$0+a=a \tag{C.1}$$

即加法左侧的零可以消去。利用数学归纳法, 当 $a=0$ 时, 根据加法定义的第一条规则有:

$$0+0=0$$

其次是递推情况, 设 $0+a=a$, 我们要推出 $0+a'=a'$。

$$\begin{aligned} 0+a' &= (0+a)' && \text{加法定义的规则二} \\ &= a' && \text{递推假设} \end{aligned}$$

接下来定义 0 的后继为 1, 并证明第二个结论:

$$a'=a+1 \tag{C.2}$$

也就是说, 任何自然的后继, 等于这个自然数加一。这是因为:

$$\begin{aligned} a' &= (a+0)' && \text{加法定义的规则一} \\ &= a+0' && \text{加法定义的规则二} \\ &= a+1 && \text{定义 0 的后继是 1} \end{aligned}$$

第三个要证明的结论是交换律的一个特例:

$$a+1=1+a \tag{C.3}$$

用数学归纳法, $a=0$ 时:

$$\begin{aligned} 0+1 &= 1 && \text{加法左侧零可消去} \\ &= 1+0 && \text{加法定义的第一条规则} \end{aligned}$$

然后是递推情况, 设 $a+1=1+a$ 成立, 我们要推出 $a'+1=1+a'$。

$$
\begin{aligned}
a'+1 &= a'+0' && \text{1 是 0 的后继} \\
&= (a'+0)' && \text{加法定义的第一条规则} \\
&= ((a+1)+0)' && \text{结论二:式 (C.2)} \\
&= (a+1)' && \text{加法定义的规则一} \\
&= (1+a)' && \text{递推假设} \\
&= 1+a' && \text{加法定义的规则二}
\end{aligned}
$$

有了这三个结论，就可以着手证明加法交换律了。再次使用数学归纳法，首先证明 $b=0$ 时交换律成立。根据加法定义的规则一，有 $a+0=a$；同时根据刚才证明的结论一，又有 $0+a=a$。这就证明了 $a+0=0+a$。然后证明递推情况。假设 $a+b=b+a$ 成立，我们要推出 $a+b'=b'+a$。

$$
\begin{aligned}
a+b' &= (a+b)' && \text{根据加法定义的第二条规则} \\
&= (b+a)' && \text{递推假设} \\
&= b+a' && \text{加法定义的第二条规则} \\
&= b+a+1 && \text{结论二:式 (C.2)} \\
&= b+1+a && \text{结论三:式 (C.3)} \\
&= (b+1)+a && \text{第 2.6 证明的加法结合律} \\
&= b'+a && \text{结论三:式 (C.3)}
\end{aligned}
$$

这样就使用皮亚诺公理，完整地证明了加法的交换律[19]-p147-148。

参考文献

[1] 马克 · 哈里森. “埃及和叙利亚的拿破仑战争”. 世界史百科, 2023 年 4 月 27 日, https://www.worldhistory.org/Napoleon's_Campaign_in_Egypt_and_Syria/. 2025 年 2 月 26 日访问。

[2] 大英博物馆. “Everything you ever wanted to know about the Rosetta Stone”. 大英博物馆博客, 2017 年 7 月 14 日, https://www.britishmuseum.org/blog/everything-you-ever-wanted-know-about-rosetta-stone. 2025 年 2 月 14 日访问.

[3] 上海博物馆. “金字塔之巅:古埃及文明”. 上海书画出版社, 2024. p114.

[4] LeVeque, William Judson, Smith, David Eugene. “数和计数系统”. 大英百科全书, 2025 年 1 月 7 日, https://www.britannica.com/science/numeral. 2025 年 2 月 14 日访问.

[5] [美] 卡尔文 ·C· 克劳森 著 袁向东、袁钧 译. “数学旅行家:漫游数王国” [M]. 上海教育出版社. 2001.

[6] [日] 野口哲典 著, 刘慧 韩丽红 译. 数学原来可以这样学 [M]. 湖南人民出版社. 2014. p31.

[7] 韩雪涛. 好的数学 —— “下金蛋” 的数学问题 [M]. 长沙: 湖南科学技术出版社, 2009. p23.

[8] J J O'Connor and E F Robertson. “Brahmagupta”. MacTutor, https://mathshistory.st-andrews.ac.uk/Biographies/Brahmagupta/. Accessed 10 March 2025.

[9] 大英百科全书编辑. “阿尔·花拉子密”. 大英百科全书, 2025 年 2 月 13 日, https://www.britannica.com/biography/al-Khwarizmi. 2025 年 2 月 18 日访问.

[10] Gies, Frances Carney. “斐波那契”. 大英百科全书, 2024 年 12 月 2 日, https://www.britannica.com/biography/Fibonacci. 2025 年 2 月 18 日访问.

[11] Charles Seife. “Zero: the biography of a number” Penguin Books, 2000. ISBN: 9-781-1011-9960-2

[12] [英] 伊恩·斯图尔特 著, 何生 译. “不可思议的数” 人民邮电出版社. 2019. ISBN: 978-7-115-51051-8

[13] Michael Artin. 代数(英文版, 第二版) [M]. 北京: 机械工业出版社, 2011.

[14] 莫里斯 · 克莱因. 古今数学思想, 第一册 [M]. 张理京, 等译. 上海: 上海科学技术出版社, 2014.

[15] 郭书春 译注. 九章算术 [M]. 上海古籍出版社. 2009. ISBN: 9787532554331

[16] 利奥 · 罗杰斯. “负数的历史”. NRICH, 剑桥大学. https://nrich.maths.org/articles/history-negative-numbers. 2011. 2025 年 3 月 10 日访问.

[17] [荷] 范·德·瓦尔登 著 丁石孙等 译. 代数学 I [M]. 北京:科学出版社. 2009, ISBN: 978-7-03-024562-5

[18] M · 克莱因. 数学: 确定性的丧失 [M]. 李宏魁, 译. 湖南: 湖南科学技术出版社, 2007.

[19] [美] 亚历山大 A· 斯捷潘诺夫, 丹尼尔 E· 罗斯著, 爱飞翔译. “数学与泛型编程: 高效编程的奥秘”. 机械工业出版社. 2017, ISBN: 9787111576587

[20] 莫里斯 · 克莱因. 古今数学思想, 第一册 [M]. 张理京, 等译. 上海: 上海科学技术出版社, 2014.

[21] Liz Pumfrey. “History of fractions”. NRICH, 剑桥大学. https://nrich.maths.org/articles/history-fractions. 2011. 2025 年 4 月 11 日访问.

[22] Jeff Miller. “Earliest Uses of Symbols for Fractions”. MacTutor, School of Mathematics and Statistics, University of St Andrews, Scotland. https://mathshistory.st-andrews.ac.uk/Miller/mathsym/fractions/. Accessed 11 April 2025.

[23] 霍夫曼·保罗. “保罗·埃尔德什”. 大英百科全书, 2025 年 3 月 22 日, https://www.britannica.com/biography/Paul-Erdos. 2025 年 4 月 16 日访问.

[24] 约翰·史迪威. “Mathematics and its history. Third edition”. Springer 2010. ISBN: 9781441960535

[25] J J O’Connor and E F Robertson. “Paul Erdős”. MacTutor, Jan, 2000, https://mathshistory.st-andrews.ac.uk/Biographies/Erdos/. Accessed 16 April 2025.

[26] 韩雪涛. 数学悖论与三次数学危机 [M]. 北京:人民邮电出版社, 2016.

[27] 大英百科全书编辑. “埃利亚的芝诺”. 大英百科全书, 2024 年 4 月 17 日, https://www.britannica.com/biography/Zeno-of-Elea. 2025 年 4 月 22 日访问.

索引