

# 数学基础选讲

—少年班先修课程

程艺 编著

中国科学技术大学  
二〇二〇年十月

## 前　　言

从上世纪五十年代开始, 美国在高中开设了具有大学水平的“先修课程 ( Advanced Placement , 简称 AP 课程)”. 不仅为解决高中教育和大学教育的衔接问题起到较好的作用, 还使得高中学生可提前接触并完成一些大学的学分课程. 国内一些大学和学术机构也注意到类似的问题, 尝试建立符合中国国情的 AP 课程体系.

AP 课程并不是简单地在高中开设几门大学入门课程, 而是需要有一个系统的设计和考虑. 一个显而易见的问题是 AP 课程能否被大学认可和接受, 并可替代同类的大学课程和学分. 以《单变量微积分》为例, 该课程在大学一般需要100-120 学时, 约 5 或 6 个学分. 如果开设可替代的 AP 课程, 过多的学时势必会加重高中生的学业负担, 而一般介绍性质的 AP 课程, 又难以达到大学对该课程的要求.

开设 AP 课程的另一个初衷是希望为学有余力的高中生尽早接受大学课程的思维方式和学习方法. 笔者在大学从教中也深感高中生进入大学后, 确实需要尽快提高抽象概念的理解能力, 适应分析问题的思维方式. 鉴于此, 不妨让学生对一些大学数学课程中的基本概念和思想, 对分析和处理问题的思维方式等方面“先修”一步, 为大学课程学习打基础、做铺垫, 而不必拘泥于针对某门大学课程, 选择“先修”内容.

编写本《数学基础选讲》正是基于上述考虑. 《选讲》选择了七个专题. 每个专题既不针对某个高中数学选修模块, 也不针对某门大学数学课程. 但从整体上看, 更加聚焦与大学数学课程的衔接, 为分析、几何和代数等基础课程做一些基础性的铺垫. 七个专题并不追求内容的完整性和深度, 而是试图以更加简洁的方式, 讲清楚一些基本数学概念的形成和背景, 帮助学生理解如何从问题入手, 产生新的数学思想和理论的过程, 使学生在抽象思维的养成和分析问题能力的锻炼上有“渐入佳境”的感觉, 因此仍然定位为大学数学的“先修课程”.

第 1 讲主要从四个方面, 让学生理解一个看似自然, 又难以说清楚的概念: “无限”. 从自然数归纳公理看自然数的无限性以及熟知的数学归纳法在处理“无限”问题中的重要性; 从有限集合之间元素的对应, 引出无限集合“基数”的建立; 从两直线上点的“射影”, 在几何上如何理解“无限远点”. 从有限求和到无限求和, 借用中学所学习的“任意”和“存在”两个逻辑量词, 准确刻画无限求和的含义. 这里, 实际上已经涉及到数项级数(数列)极限的定义, 为微积分中的极限做了基本介绍.

第 2 讲是关于整数, 把学生们习以为常的一些运算和数之间的大小关系做了归纳提炼, 希望学生逐步建立“公理”的概念. 再从初等数论的角度介绍了素数、带余除法、辗转相除、同余和同余类以及多项式等, 为学习抽象代数提供一个基本模型.

第 3 讲从有理数、可公度和不可公度讲起, 从序到有界, 再到确界原理, 最后给出

---

实数的定义. 尽可能避免涉及极限的内容. 同时解释了十进制无穷小数, 最后选择了 Dedekind 分割作为实数构造的一种方法.

第 4 讲是关于复数, 从复数的起源, 直到给出 Euler 公式, 介绍了代数学基本定理和单位根的有关问题, 通过一些例子简单介绍了复变数函数的一些基本特点.

第 5 讲以空间解析几何作为基础, 从直观上的坐标系, 到向量的代数运算、再到如何把几何问题代数化. 最后从三维空间向量的代数、内积, 如何抽象为一般向量空间的定义和内积, 完成从具体到抽象的过程. 通过例子介绍了线性方程组. 为学生进一步学习线性代数做了铺垫.

第 6 讲是关于用直尺和圆规几何作图问题. 让学生感受到如何将一个纯几何的问题代数化, 最终能够用学生理解的方式, 解决著名的三个几何作图不可解的问题. 虽然内容中出现的如“扩域”这样一些看似高难度的概念, 但从具体问题入手, 自然产生扩域的概念, 也易于让学生接受.

第 7 讲是关于群的介绍. 通过回顾二次和三次代数方程的代数求解方法, 逐步引进根的置换, 再抽象出“对称群”的概念, 并以对称群为模型, 介绍了有限群的一些基本知识. 考虑到这个专题所涉内容的广泛性和深刻性, 这里只是从问题出发, 引入了群的概念, 使学生掌握群的基本性质, 并为进一步选修抽象代数、或 Galois 理论、或有关群论的课程, 提供了先期的接触.

需要说明的是, 以上内容尚未在任何普通高中或大学新生中开设过. 只是在为中国科学技术大学少年班和创新班一年级学生的教学中, 时常穿插部分内容, 或作为自学材料提供给学生, 帮助他们更好地学习和理解数学, 起到了较好作用.

这里对中国科大少年班做一个简单介绍. 中国科大少年班的办学历史与我国改革开放、恢复高考几乎同步. 简要地说, 其主要特点体现在如下两个方面:

一是在人才选拔上, 从早先的中学或社会推荐、中国科大组织考试(笔试和面试), 到后期以高考为初试, 再经中国科大自行组织的笔试和面试作为复试, 最终择优录取, 实际上开创了我国高校“自主招生录取”的先河.

二是在培养模式上, 始终坚持“强化数理基础, 自主选择专业”的基本原则, 学生入学后不确定专业, 后期根据个人志向和特点, 自主选择专业. 体现了个性化培养的理念.

给少年班(以及后来扩展的创新班)讲授数学基础课程, 看似没有专业对课程的要求, 实际上不管他们今后选择什么样的专业, 都要保证为他们打下“宽、厚、实”的数学基础. 这样促使教师不得不为符合他们特点而改革教学方法, 调整课程内容结构. 这些改革已经在系统地推进和实践, 这里不展开介绍. 编写本《选讲》的目的也是针对少年班和创新班的特点, 从打牢数学基础、有利于大学数学课程学习角度出发, 尝试设计一门数学“先修课程”或选修课程.

顾名思义，“先修课程”是指在进入大学之前进行先修。现阶段高中的部分选修内容尚未纳入高考，因此，指望在高中生中普遍讲授这样的专题不大现实。但是，面向那些对数学特别感兴趣的高中生，可根据他们的精力和兴趣，选择若干专题进行讲授。也可为刚入学的大学新生开设选修课程。从提高素养、拓宽视野、夯实基础、培养抽象思维和分析问题的能力角度看，应该是有益的。

本《选讲》在编写过程中大量参考借鉴了已经出版的一些书籍和网上的一些资料，一些思路和内容直接来自于这些参考书籍和资料。在此向有关书籍和资料的作者深表感谢。尽管如此，鉴于编写的水平有限，选材未必合适，错误也难以避免，望读者批评指正。

### 参考书目：

- [1] R. Courant, H. Robbins 著, I Stewart 修订, 左平、张饴慈译:《什么是数学》, 复旦大学出版社, 2018年第4版.
- [2] 谷超豪著:《谈谈数学中的无限》, 上海教育出版社出版, 1988年8月第1版.
- [3] 冯克勤、余红兵编著:《整数与多项式》, 高等教育出版社, 1999年10月第1版.
- [4] W. Rudin 著, 赵慈庚、蒋铎译:《数学分析原理》, 机械工业出版社 2005年1月第1版.
- [5] 程艺、陈卿、李平、许斌编著:《数学分析讲义》(第三册), 高等教育出版社, 2020年8月第1版.
- [6] K. F. Riley, M. P. Hobson, S. J. Bence: Mathematical Methods for Physics and Engineering, Cambridge University Press, 2006, Third Editon
- [7] 陈发来、陈效群、李思敏、王新茂编著:《线性代数与解析几何》, 高等教育出版社, 2011年7月第1版.
- [8] 李尚志主编:《三等分角与数域扩充》, 湖南教育出版社, 2004年6月第1版.
- [9] 李世雄著:《代数方程与置换群》, 上海教育出版社, 1981年2月第1版.
- [10] 欧阳毅、申伊境内编著:《代数学I: 代数学基础》, 高等教育出版社, 2016年6月第1版.

编者

2020年10月于中国科学技术大学

# 目 录

前言 ······	I
<b>第 1 讲 无限 ······</b>	<b>1</b>
§1.1 自然数的无限性和数学归纳法 ······	1
§1.2 无限集合 ······	4
§1.3 无限求和 ······	11
§1.4 无限远点 ······	19
第 1 讲习题 ······	31
<b>第 2 讲 整数 ······</b>	<b>32</b>
§2.1 正整数与整数 ······	32
§2.2 数的整除性 ······	34
§2.3 Euclid 辗转相除法 ······	36
§2.4 素数和整数的素因子分解 ······	38
§2.5 同余和同余类 ······	43
§2.6 同余方程(组) ······	50
§2.7 多项式 ······	54
第 2 讲习题 ······	59
<b>第 3 讲 实数 ······</b>	<b>61</b>
§3.1 有理数 ······	61
§3.2 可公度与不可公度 ······	63
§3.3 实数 ······	65
§3.4 十进制小数 ······	70
§3.5 Dedekind 分割* ······	72
第 3 讲习题 ······	80
<b>第 4 讲 复数 ······</b>	<b>81</b>
§4.1 复数的起源和运算 ······	81
§4.2 复数的几何含义和 Euler 公式 ······	83
§4.3 代数基本定理 ······	87
§4.4 单位根 ······	88
§4.5 复变数的函数 ······	92
第 4 讲习题 ······	96

---

<b>第 5 讲 解析几何与向量空间</b>	<b>98</b>
§5.1 向量和及其代数运算	99
§5.2 向量的坐标表示和坐标系	104
§5.3 平面、直线与曲面	107
§5.4 其它常用坐标系	113
§5.5 向量空间	115
§5.6 线性方程组的解	123
第 5 讲习题	128
<b>第 6 讲 几何作图</b>	<b>129</b>
§6.1 尺规在作图中的功能	129
§6.2 作图的代数表示	130
§6.3 数域的扩充	136
§6.4 几何作图与代数方程的根	138
§6.5 几何作图中三个不可解的问题	140
§6.6 等分圆周的几何作图问题*	143
第 6 讲习题	150
<b>第 7 讲 群与对称性</b>	<b>151</b>
§7.1 从代数方程求根谈起	151
§7.2 对称多项式和代数方程的 Viete 公式	154
§7.3 代数方程根的置换和求解	156
§7.4 置换及其表示	160
§7.5 有限群及其性质	164
第 7 讲习题	177

# 第 1 讲 无限

无限(或无穷)在数学中起着重要作用. 自然数序列的无限性是一个最简单的例子, 直线或平面上所有点的集合等等, 都是包含无穷多个对象(或称为元素)的集合. 在数学中, 我们将面临大量的如何处理具有无穷多个元素的集合, 或无穷多个对象所共有的性质. 甚至还将面临如何从有限过渡到无限等一系列问题.

## §1.1 自然数的无限性和数学归纳法

自然数就是熟知的

$$\mathbb{N} = \{1, 2, 3, \dots\},$$

这里, 暂且不考虑“0”, 因此也称上述数为“正整数”. 自然数的公理化定义是由 Peano(皮亚诺 1858 -1932)作出的. 对此不作进一步介绍. 自然数的基本性质首先是它所满足的算术, 也就是自然数之间的加法和乘法, 以及加法和乘法所服从的交换律、结合律和分配律. 其次是自然数的有序性, 即任何两个自然数之间一定存在大小关系. 以上两点将在第 2 讲中介绍. 第三是自然数的无限性, 也就是说, 自然数是没有止境的. 因为任何自然数  $n$  后, 还可以写出下一个自然数  $n + 1$ . 这种性质总结为下列归纳公理:

**归纳公理:** 设  $S \subseteq \mathbb{N}$ , 如果  $S$  满足

- (a)  $1 \in S$ ,
- (b) 若  $n \in S$ , 则  $n + 1 \in S$ ,

那么  $S = \mathbb{N}$ .

由归纳公理, 不难得到下列自然数的最小数原理.

**定理 1.1 (最小数原理)** 设  $T$  是  $\mathbb{N}$  的非空子集. 则  $T$  中必有最小的整数.

**证明** 设

$$S = \{s \mid s \in \mathbb{N}, s \leq t, \text{ 对任意的 } t \in T \text{ 成立}\}.$$

显然,  $1 \in S$ , 因此  $S$  非空. 又因为  $T$  非空, 对  $t \in T$ , 有  $t + 1 > t$ , 所以  $t + 1 \notin S$ , 也就是  $S \neq \mathbb{N}$ .

据此推出: 存在  $s_0 \in S$ , 而  $s_0 + 1 \notin S$ . 否则, 若这样的  $s_0$  不存在, 就意味着对任意的  $s \in S$ , 都有  $s + 1 \in S$ , 根据归纳原理推出  $S = \mathbb{N}$ , 这与  $S \neq \mathbb{N}$  的结论相矛盾.

若  $s_0 \notin T$ , 则对于任意的  $t \in T$ , 都有  $s_0 < t$ , 也就是  $s_0 + 1 \leq t$ , 推出  $s_0 + 1 \in S$ , 这与  $s_0$  的选取矛盾. 所以  $s_0 \in T$ , 由  $S$  的定义推出  $s_0$  是  $T$  的最小数.

需要说明的是, 这里谈论的“最小数原理”仅是“自然数集的最小数原理”. 例如,

数集

$$\left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, \frac{1}{n}, \dots\right\}$$

中是没有最小数的.

归纳公理,或者说从自然数  $n$  到  $n+1$  这样一步一步地产生自然数的无限序列是常用的数学归纳法的基础.

所谓数学归纳法是数学推理的一个最基本的方式之一. 用来证明一个具有无限序列情形都是正确的数学定理.

**例 1.1.1** 任何有  $n+2$  条边的凸多边形的内角之和等于  $180^\circ$  的  $n$  倍.

这是一个对每一个自然数  $n$  都成立的定理. 如果采用一个一个地验证的方式去证明, 不管验证到多少, 仍然不能说明定理为真. 因此必须采用一种严格的数学推理方法来证明.

显然, 当  $n=1$  时, 凸多边形就是三角形, 我们可以用独立于该定理的其他结果知道, 三角形内角之和为  $180^\circ$ .

对于  $n=2$  的四边形, 可以画一条对角线把四边形分成两个三角形, 因此利用三角形内角之和的结论可知四边形内角之和为  $2 \cdot 180^\circ$ .

接着  $n=3$  五边形的情形, 可以把它分解成三角形和四边形, 再利用已经证明的关于三角形和四边形内角之和的结论得到五边形内角之和为  $180^\circ + 2 \cdot 180^\circ = 3 \cdot 180^\circ$ .

以此类推, 可以逐次证明  $n=4, n=5$ , 等情形. 而且每一步都以同样的方式由前面的结论推出.

上述证明的基本思想是: 为了证明定理对所有  $n$  成立, 我们基于以下两点:

一是对  $n=1$  或前几个情形通过验证可知是正确的, 在例1.1.1中, 借助熟知的三角形内角之和为  $180^\circ$ , 可以直接验证前几种情形的正确性.

二是存在一种一般的方法, 能够表明: 如果定理对  $n$  成立, 那么对下一个  $n+1$  也成立. 在在例1.1.1中, 这个一般的方法就是把  $n+2$  条边的凸多边形分成三角形和  $n+1$  条边的凸多边形.

把上述分析抽象为如下定理.

**定理 1.2** 设  $A_n$  是关于正整数的一个命题, 如果

(1) 当  $n=1$  时,  $A_1$  成立;

(2) 对任意的正整数  $n > 1$ , 由  $A_n$  成立可推出  $A_{n+1}$  成立. 或者说, 由前  $n$  个  $A_1, A_2, \dots, A_n$  成立可推出  $A_{n+1}$  成立.

那么,  $A_n$  对所有正整数  $n$  成立.

**证明** 采取反证法, 在定理的条件下, 假设有一个命题  $A_r$  不成立, 记

$$S = \{n \mid A_n \text{ 不成立}\} \subset \mathbb{N},$$

根据假设,  $r \in S$ , 因此  $S$  是一个非空集合, 根据最小数原理, 存在最小数  $m \in S$ . 由定理条件(1)可知  $1 \notin S$ , 所以  $m > 1$ . 又因为  $m$  是  $S$  中最小数, 所以  $m - 1 \notin S$ , 因此  $A_{m-1}$  成立, 由条件(2)立即推出  $A_m$  也成立. 但是  $m \in S$ , 即  $A_m$  不成立, 矛盾. 矛盾说明我们的假设是错误的, 所以结论是在定理条件下所有  $A_n$  都成立.  $\square$

**例 1.1.2** 对任意的正整数  $n$ , 求证  $f(n) = n^4 + 2n^3 + 2n^2 + n$  能被 6 整除.

**证明** 当  $n = 1$  时,  $f(1) = 6$ , 显然能被 6 整除.

假设  $f(n)$  能够被 6 整除, 那么

$$\begin{aligned} f(n+1) &= (n+1)^4 + 2(n+1)^3 + 2(n+1)^2 + (n+1) \\ &= (n^4 + 4n^3 + 6n^2 + 4n + 1) \\ &\quad + 2(n^3 + 3n^2 + 3n + 1) + 2(n^2 + 2n + 1) + (n+1) \\ &= (n^4 + 2n^3 + 2n^2 + n) + (4n^3 + 12n^2 + 14n + 6) \\ &= f(n) + (4n^3 + 12n^2 + 14n + 6) \end{aligned}$$

根据归纳假设, 上式右边第一项能够被 6 整除, 第二项中的  $12n^2 + 6$  也能够被 6 整除. 因此只要证明剩余的  $4n^3 + 14n$  能够被 6 整除即可, 也就是要证明  $2n^3 + 7n$  能够被 3 整除.

为此, 需要再次用归纳法去证明: 对任意的正整数  $n$ ,  $2n^3 + 7n$  能够被 3 整除.

这件事对  $n = 1$  显然成立. 假设对一般  $n$ ,  $2n^3 + 7n$  能够被 3 整除, 那么对  $n+1$ , 有

$$2(n+1)^3 + 7(n+1) = (2n^3 + 7n) + 3(2n^2 + 2n + 3).$$

根据归纳假设, 右边第一项能被 3 整除, 而第二项是 3 的倍数, 当然也能被 3 整除. 这样就证明了对任何正整数  $n$ ,  $2n^3 + 7n$  能够被 3 整除. 因此也就证明了对任何正整数  $n$ ,  $f(n) = n^4 + 2n^3 + 2n^2 + n$  能被 6 整除.

这里再次强调使用数学归纳法必须确保定理1.2中的条件(1)和(2)真正被满足.

**例 1.1.3** 记  $\max\{a, b\}$  为  $a$  和  $b$  中较大的一个数. 考虑下列命题

对给定的正整数  $n$ , 如果  $a, b$  是使得  $\max\{a, b\} = n$  成立的任意的正整数, 那么  $a = b$ .

按照归纳法的步骤, 不难验证  $A_1$  显然成立, 这是因为对任意两个满足  $\max\{a, b\} = 1$  的正整数, 一定有  $a = b = 1$

假设命题对  $n$  成立, 那么对于  $n+1$ , 设  $a, b$  是任意两个使得  $\max\{a, b\} = n+1$  的正整数, 令

$$a' = a - 1, \quad b' = b - 1$$

则  $\max\{a', b'\} = n$ , 由归纳假设, 命题对  $n$  成立, 所以  $a' = b'$ , 即  $a = b$ , 推出命题对  $n+1$  也成立. 这样根据数学归纳法, 就有对任意的  $n$ , 命题都成立.

现在任取两个正整数  $a = 5, b = 2$ , 则  $n = \max\{5, 2\} = 5$ , 命题对  $n = 5$  成立, 意味着  $5 = 2$ . 这样的结论显然是荒唐的. 但是问题出在哪里呢?

不难看出在上述推导中取  $a' = a - 1, b' = b - 1$ , 要保证  $a', b'$  是正整数, 那么就已经限定了  $a > 1, b > 1$ , 因此并不是针对所有正整数. 也就是命题的条件并没有真正被满足, 所以得出一个荒诞的结论. 由此可见在进行数学推导时, 必须严格缜密.

## §1.2 无限集合

有限个元素所组成的集合称为有限集合. 对有限集合来说, 它的元素的个数是一个最基本的量, 称为集合的基数. 因此, 有限集合的基数是可以通过计算集合中元素的个数得到的, 不管该集合的元素是什么.

例如集合  $\{a, b, c, d\}$  的基数是 4, 集合  $\{2, 4, 6, 8\}$  的基数也是 4. 空集  $\phi$  的基数为 0.

两个有限集合的基数大小也可以通过下列方法比较: 例如一群小朋友与一篮苹果. 如果每个小朋友仅拿一个苹果, 到最后, 若苹果分光了, 还有小朋友没有拿到, 说明小朋友集合的基数大于苹果集合的基数. 若每个小朋友都拿到一个, 篮子里还有剩余的苹果, 说明小朋友集合的基数小于苹果集合的基数. 如果正好每人分得一个苹果, 没有剩余, 说明小朋友与苹果集合的基数相等.

将上述通过 1-1 对应来考察有限集合基数是否相等的做法, 推广到其它集合中去.

**定义 1.3** 设  $A, B$  为两个集合,  $f$  是两个集合之间的映射

$$f : A \longrightarrow B$$

(1) 若对  $A$  中任意两个元素  $a, a'$ , 只要  $a \neq a'$ , 就有  $f(a) \neq f(a')$ , 则称映射为单射.

(2) 若对  $B$  中任何一个元素  $b$ , 至少存在  $A$  中的元素  $a$ , 使得  $f(a) = b$ , 则称映射为满射.

(3) 若映射既是单射又是满射, 那么称映射为 1-1 映射(或称 1-1 对应).

利用 1-1 对应, 可以比较两个集合之间的基数, 特别是当选择自然数集的子集合, 或

自然数集自身作为一个标准, 那么通过与它们的1-1 对应, 可以定义有限集合与无限集合.

**定义 1.4** 设  $A, B$  为两个集合,  $\mathbb{N}$  为自然数集合.

(1) 若存在  $A \rightarrow B$  的1-1 对应, 则称  $A$  和  $B$  有相同的基数, 或者称  $A$  和  $B$  是等势的.

(2) 若存在  $A \rightarrow B$  的满射, 但不存在  $A \rightarrow B$  的单射 (因此也就不存在  $A$  和  $B$  之间的1-1 对应), 则称集合  $A$  比集合  $B$  具有更大的基数.

(3) 若存在自然数  $n$ , 使得  $A$  与自然数的子集合  $\{1, 2, \dots, n\}$  之间有1-1 对应, 则称  $A$  为有限集合; 若这样的  $n$  不存在, 则称  $A$  为无限集合.

(4) 称自然数集合  $\mathbb{N}$  的基数为可数的, 因此任何与  $\mathbb{N}$  的1-1 对应的集合  $A$ , 其基数也是可数的. 具有可数的基数的集合称为可数集合 (因此自然数集合  $\mathbb{N}$  是可数集合). 所有可数集合都有相同的基数, 或者说都是等势的.

对于有限集合来说, 基数相等和通常集合个数相等的概念时一致的. 因为当数一数一个有限集合  $A$  中元素的个数时, 就是把  $A$  中的元素与自然数的子集合  $\{1, 2, \dots, n\}$  1-1 对应

$$\begin{array}{cccccc} 1 & 2 & 3 & \cdots & n \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\ a_1 & a_2 & a_3 & \cdots & a_n \end{array}$$

从另一个角度看, 上述1-1 对应就是给有限集合的元素进行一种不重复的排列 (或者说编一个号码), 当然, 排列方式不唯一. 显然, 一个有限集合, 去掉一个元素, 或增加一个元素, 都不可能与原集合1-1 对应, 也就是基数不可能再相等.

同样, 对于一个可数集合  $A$ , 也可以通过与自然数集合的1-1 对应进行不重复的排列 (编号):

$$\begin{array}{ccccccc} 1 & 2 & 3 & \cdots & n & \cdots \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow & \cdots \\ a_1 & a_2 & a_3 & \cdots & a_n & \cdots \end{array}$$

因此我们通常记可数集合为

$$A = \{a_1, a_2, \dots, a_n, \dots\}.$$

但是对于可数集合, 其中无限个元素组成的子集合也具有相同的基数. 例如自然数集合  $\mathbb{N}$  中, 所有偶数组成的子集合也具有和  $\mathbb{N}$  相同的基数, 即偶数组成的集合与自然数集合

是1-1 对应的, 因此也是可数的集合:

$$\begin{array}{ccccccc} 1 & 2 & 3 & \cdots & n & \cdots \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow & \cdots \\ 2 & 4 & 6 & \cdots & 2n & \cdots \end{array}$$

可见, 无限集合具有一些独特的性质. 我们把可数集合的部分性质罗列如下.

**性质 1.5** 若存在从自然数集合  $\mathbb{N}$  到无限集合  $U$  的满射

$$f : \mathbb{N} \longrightarrow U,$$

则  $U$  为可数集合. 特别地,  $\mathbb{N}$  的任何无限真子集是可数集.

**证明** 证明的关键是通过满射, 构造  $U$  中所有元素的一个不重复排列, 也就是建立与  $\mathbb{N}$  的1-1 对应. 记  $j_1 = 1$ , 并令

$$a_1 = f(j_1) = f(1),$$

因为  $f$  为满射且  $U$  是无限集, 集合

$$E_1 = \{n \in \mathbb{N} \mid f(n) \neq a_1\}$$

不是空集. 根据自然数的最小数原理,  $E_1$  存在最小数, 记为  $j_2$ , 令

$$a_2 = f(j_2).$$

此时我们有  $j_2 > j_1$ ,  $a_2 \neq a_1$ . 由于  $f$  仅仅是满射, 所以数  $1, 2, \dots, j_2$  在  $f$  下的像为  $a_1$  或者  $a_2$ :

$$f : \{1, 2, \dots, j_2\} \longrightarrow \{a_1, a_2\}.$$

再令

$$E_2 = \{n \in \mathbb{N} \mid f(n) \neq a_1, f(n) \neq a_2\},$$

并取  $E_2$  的最小数  $j_3$ , 显然  $j_3 > j_2$ . 记

$$a_3 = f(j_3),$$

因此  $a_1, a_2, a_3$  两两不等, 且

$$f : \{1, 2, \dots, j_2, \dots, j_3\} \longrightarrow \{a_1, a_2, a_3\}.$$

假设已经取到  $j_k > j_{k-1} > \dots > j_1$ , 使得  $a_l = f(j_l)$ ,  $l = 1, 2, \dots, k$  两两互不相等, 并且  $1, 2, 3, \dots, j_k$  在  $f$  下的像为  $\{a_1, \dots, a_k\}$ :

$$f : \{1, 2, \dots, j_k\} \longrightarrow \{a_1, a_2, \dots, a_k\}.$$

那么记

$$E_{k+1} = \{n \in \mathbb{N} \mid f(n) \neq a_l, l = 1, 2, \dots, k\},$$

取  $E_{k+1}$  的最小数  $j_{k+1}$  并记

$$a_{k+1} = f(j_{k+1}),$$

则有  $j_{k+1} > j_k > \dots > j_1, a_{k+1}, a_k, \dots, a_1$  互不相等, 且

$$f : \{1, 2, 3, \dots, j_k, \dots, j_{k+1}\} \longrightarrow \{a_1, \dots, a_k, a_{k+1}\}.$$

由于  $f$  是满射, 这样得到一个严格递增的自然数列  $j_1 < j_2 < j_3 < \dots$  和一个  $U$  中元素的一个排列

$$U = f(\mathbb{N}) = \{a_1, a_2, a_3, \dots\}.$$

因此  $U$  是可数的.

设  $V$  为  $\mathbb{N}$  的无限真子集. 取定一个  $V$  中的元素  $m$ , 定义映射  $f : \mathbb{N} \rightarrow V$  如下:

$$\text{对任意的 } n \in \mathbb{N}, f : n \longmapsto f(n) = \begin{cases} n, & \text{当 } n \in V \\ m, & \text{当 } n \notin V. \end{cases}$$

显然  $f : \mathbb{N} \rightarrow V$  是满射. 由前面的结论可知  $V$  为可数集. □

根据性质 1.5, 在自然数集  $\mathbb{N}$  中去掉有限个数的集合 (一定是无限集合), 或即使去掉无限个数, 只要剩余的数仍然是无限集合, 那么基数是不变的. 例如, 集合  $\{2, 3, \dots, n, \dots\}$  以及偶数集合或奇数集合都是可数集合, 它们与自然数集合是 1-1 对应的, 有相同的基数.

**性质 1.6** 有限集合与可数集合的并集是可数集合. 有限个可数集合的并集仍然是可数集合. 可数个可数集合的并集还是可数集.

**证明** 这里我们只讨论第三种情形. 设  $A_1, A_2, \dots$  为可数个可数集. 记

$$A_k = \{a_{k1}, a_{k2}, \dots, a_{kn}, \dots\}, \quad k = 1, 2, 3, \dots.$$

那么并集  $\bigcup_{k=1}^{\infty} A_k$  里的所有元素可以表示为如下无穷矩阵,

$$\begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{array}$$

把纸面沿顺时针方向旋转  $45^\circ$ , 我们可以把矩阵看成一个大三角形, 沿着箭头我们得到元素的一个排列

$$\begin{aligned}
& a_{11} \rightarrow \\
& \rightarrow a_{21} \rightarrow a_{12} \rightarrow \\
& \rightarrow a_{31} \rightarrow a_{22} \rightarrow a_{13} \rightarrow \\
& \rightarrow \cdots \rightarrow \cdots \rightarrow \cdots \rightarrow \cdots \rightarrow
\end{aligned}$$

因此给出了  $\mathbb{N}$  到并集的一个满射, 但不一定是单射, 因为所有元素  $a_{ij}$  中可能含有重复的元素. 根据性质1.5 可知结论成立.  $\square$

性质1.6 说明, 在一个可数集合中增加有限个元素, 或增加可数个元素, 结果仍然是可数的, 也就是基数不变.

需要说明的是性质1.6 的证明中使用的排列方法并不是唯一的.

### 推论 1.7 整数集合

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots, \pm n, \dots\}$$

是可数集, 因此与自然数集  $\mathbb{N}$  有相同的基数.

证明 记

$$\mathbb{Z}_+ = \mathbb{N} = \{1, 2, 3, \dots, n, \dots\},$$

$$\mathbb{Z}_- = \{-1, -2, -3, \dots, -n, \dots\},$$

那么  $\mathbb{Z}_-$  与  $\mathbb{N}$  是1-1 对应的, 因此是可数集. 所以,

$$\mathbb{Z} = \mathbb{Z}_- \cup \{0\} \cup \mathbb{Z}_+$$

是一个可数集.  $\square$

集合  $A, B$  的直积  $A \times B$  (或者笛卡尔积) 定义为

$$A \times B = \{(x, y) \mid x \in A, y \in B\}.$$

类似地可以定义有限个集合  $A_1, A_2, \dots, A_n$  的直积的定义为

$$\prod_{k=1}^n A_k = A_1 \times A_2 \times \cdots \times A_n = \{(x_1, x_2, \dots, x_n) \mid x_k \in A_k, k = 1, 2, \dots, n\}.$$

类似于性质1.6 的证明, 我们可得到如下命题

**性质 1.8** 有限个可数集合的直积也是可数集.

证明 这里我们只考虑两个可数集合的直积. 设

$$A = \{a_1, a_2, \dots, a_n, \dots\}, B = \{b_1, b_2, \dots, b_n, \dots\},$$

那么

$$A \times B = \{(a_i, b_j) \mid a_i \in A, b_j \in B\}$$

记  $a_{ij} = (a_i, b_j)$ , 类似性质1.6的做法, 可以给出从  $\mathbb{N}$  到  $A \times B$  的一个满射, 因此  $A \times B$  是可数的.  $\square$

**推论 1.9** 有理数集合是可数集, 也就是有理数集与自然数集的基数相等.

**证明** 设有理数集为

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \right\}.$$

那么

$$f_1 : \mathbb{Z} \times \mathbb{Z}_+ \longrightarrow \mathbb{Q}, f_1 : (p, q) \longmapsto \frac{p}{q}$$

是  $\mathbb{Z} \times \mathbb{Z}_+$  到  $\mathbb{Q}$  的满射. 因为  $\mathbb{Z} \times \mathbb{Z}_+$  与自然数集合  $\mathbb{N}$  1-1 对应:

$$f_2 : \mathbb{N} \longrightarrow \mathbb{Z} \times \mathbb{Z}_+,$$

所以复合映射

$$f = f_1 \circ f_2 : \mathbb{N} \longrightarrow \mathbb{Q}$$

是满射, 根据性质1.5, 有理数集  $\mathbb{Q}$  是可数集合.  $\square$

通常把有理数集排序为

$$\mathbb{Q} = \{r_1, r_2, \dots, r_n, \dots\}.$$

以上是与自然数集有相同基数的可数集合的主要性质. 一个自然的问题是: 是否所有的无限集合都是可数集合? 或者问是否存在基数小于或大于可数集合基数的无限集合?

考虑一个无限集合  $A$ , 若  $A_0 \subset A$  是  $A$  的子集, 定义

$$A \setminus A_0 = \{a \mid a \in A, a \notin A_0\},$$

即从  $A$  中去掉  $A_0$  的元素, 称为  $A_0$  在  $A$  中的余集.

从  $A$  中取一个元素  $a_1$ , 又从  $A \setminus \{a_1\}$  中取一个元素  $a_2$ , 显然  $a_2 \neq a_1$ . 若已经取到互不相等的  $a_1, a_2, \dots, a_n$ , 则由于  $A$  是无限集合,  $A \setminus \{a_1, a_2, \dots, a_n\}$  非空, 所以从中取  $a_{n+1}$ , 显然  $a_{n+1} \neq a_i, i = 1, 2, \dots, n$ . 这样的操作可以一直继续下去. 因此得到  $A$  的一个可数的子集  $A_0 = \{a_1, a_2, \dots, a_n, \dots\}$ .

定义映射  $f : A \longrightarrow \mathbb{N}$  如下:

$$f(a) = \begin{cases} n, & \text{当 } a = a_n \in A_0 \\ 1, & \text{当 } a \in A \setminus A_0 \end{cases}$$

它是从  $A$  到  $\mathbb{N}$  的一个满射. 根据定义1.4, 如果  $f$  又是单射, 那么  $A$  的基数与  $\mathbb{N}$  相同, 如果  $f$  不是单射, 那么  $A$  有比  $\mathbb{N}$  更大的基数. 所以

**定理 1.10** 不存在基数比自然数集(可数集合)基数更小的无限集合.

但是, 却存在比可数集合的基数更大的无限集合.

**定义 1.11** 无限集合称为**不可数**, 是指不存在它与 $\mathbb{N}$ 之间的1-1对应. 换而言之, 它有比 $\mathbb{N}$ 更大的基数.

为了构造一个不可数集合, 需要引入一个新的概念:

**定义 1.12** 对于一个非空集合 $A$ , 记 $2^A$ 是它的所有子集构成的集合

$$2^A = \{X \mid X \subset A\}.$$

并称为 $A$ 的**幂集**.

例如(以下 $\phi$ 表示空集), 当 $A = \{a, b\}$ 时

$$2^A = \{\phi, \{a\}, \{b\}, \{a, b\}\}.$$

当 $A = \{x, y, z\}$ 时

$$2^A = \{\phi, \{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \{x, y, z\}\}.$$

不难看出, 当 $A$ 中元素的个数(基数)为2时, $2^A$ 中元素的个数(基数)为 $2^2 = 4$ , 当 $A$ 中元素的个数(基数)为3时, $2^A$ 中元素的个数(基数)为 $2^3 = 8$ .

一般情况下, 对一个具有 $n$ 个元素的有限集合 $A$ , 它的各种组合给出幂集 $2^A$ 中的元素, 其个数为 $2^n$ . 翻译成“基数”, 即是

$$2^A \text{ 的基数} = 2^A \text{ 的基数}.$$

下面, 我们重点考虑可数的自然数集的幂集

$$2^{\mathbb{N}} = \{X \mid X \subseteq \mathbb{N}\}.$$

也就是 $\mathbb{N}$ 所有子集组成的集合.

**定理 1.13** (Cantor 康托1845~1918)  $2^{\mathbb{N}}$ 是不可数集合.

**证明** 采用反证法. 假设 $2^{\mathbb{N}}$ 是可数集, 则存在 $\mathbb{N}$ 与 $2^{\mathbb{N}}$ 的1-1对应. 这样 $2^{\mathbb{N}}$ 中的元素通过1-1对应可以排列如下

$$2^{\mathbb{N}} = \{U_1, U_2, \dots\}.$$

这里每个 $U_1, U_2, \dots$ 都是 $\mathbb{N}$ 的子集. 令

$$V = \{k \in \mathbb{N} \mid k \notin U_k\}.$$

如果 $V$ 是 $2^{\mathbb{N}} = \{U_1, U_2, \dots\}$ 中某一个:  $V = U_n$ , 那么当 $n \in U_n$ 时, 有 $n \in V = U_n$ , 这与 $V$ 的定义矛盾, 当 $n \notin U_n$ 时, 有 $n \notin V = U_n$ 也与 $V$ 的定义矛盾. 所以 $V$ 不可能是 $2^{\mathbb{N}} = \{U_1, U_2, \dots\}$ 中某一个, 也就是 $V \notin 2^{\mathbb{N}}$ .

但是  $V$  是  $\mathbb{N}$  的子集合, 根据  $\mathbb{N}$  的幂集的定义, 有  $V \in 2^{\mathbb{N}}$ . 所以矛盾. 这样我们就证明了  $2^{\mathbb{N}}$  不可能是可数集.

注记 Cantor 的定理说明  $2^{\mathbb{N}}$  的基数比  $\mathbb{N}$  的基数大, 这引出另一个自然的问题, 是否存在一个集合, 它的基数比  $\mathbb{N}$  的基数大, 但比  $2^{\mathbb{N}}$  的基数小? 著名的 Cantor 连续统假设说不存在这样的集合. Kurt Gödel (哥德尔 1906 ~ 1978) 与 Paul Cohen (寇恩 1934 ~ 2007) 的工作说明以上问题不可能有答案. 他们分别构造了两个集合论的“模型”, 在这两个模型里普通的集合论公理都成立, 但是连续统假设在一个里成立, 在另一个里不成立. 这些内容已经超出本书的范围, 此处不再讨论.

### §1.3 无限求和

众所周知, 有限个数相加, 结果还是一个数. 但是, 当考虑由无限个数相加(简称“无限求和”), 问题变得复杂了. 例如, 考虑如下三种无限求和的问题

- (1)  $1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n} + \cdots,$
- (2)  $1 + 2 + 2^2 + \cdots + 2^n + \cdots,$
- (3)  $1 - 1 + 1 - 1 + \cdots + (-1)^{n-1} + \cdots.$

对于 (1), 假如相加的结果是一个数, 那么不妨将该数记为

$$S = 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n} + \cdots,$$

那么按照通常的算术运算, 有

$$2S = 2 + S,$$

因此得到  $S = 2$ .

同样, 对于 (2), 假如相加的结果仍然是一个数, 记为

$$S = 1 + 2 + 2^2 + \cdots + 2^n + \cdots,$$

按照计算应该有

$$2S = S + 1,$$

得到的结果是  $S = -1$ , 即无限个正数相加的结果是  $-1$ .

而对于 (3), 如果结果是一个数

$$S = 1 - 1 + 1 - 1 + \cdots + (-1)^{n-1} + \cdots.$$

根据加法结合律, 就有

$$S = \begin{cases} (1 - 1) + (1 - 1) + \cdots + (1 - 1) + \cdots = 0; \\ 1 - (1 - 1) - (1 - 1) - \cdots - (1 - 1) - \cdots = 1 \end{cases}$$

却出现了两种结果.

因此, 我们面临这样一些问题: 如何判断无限个数相加, 其结果仍然是一个数? 如果是, 加法的运算规律是否仍然能够保持?

仍然以上述三种情形为例, 如果先把前  $n$  项加起来, 对于(1)有

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

因此, 随着求和的项数越来越多, 也就是随着  $n$  越来越大, 结果也越来越接近一个数 2.

而对于(2), 它的前  $n$  项的和为

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

随着相加的项越来越多, 结果也随着  $n$  的增大而变得越来越大, 加到最后不可能等于某一个数.

对于(3), 我们发现它的前  $n$  项的和满足

$$1 - 1 + 1 - 1 + \cdots + (-1)^{n-1} = \begin{cases} 0 & \text{当 } n \text{ 是偶数} \\ 1 & \text{当 } n \text{ 是基数} \end{cases}$$

因此每加一项, 结果会从 0 变到 1, 或从 1 变到 0, 总是在 0 和 1 之间摇摆不定.

可见, 无限个数相加的总和, 有时可以用一个数来表示, 有时又不能用一个数来表示. 现在的问题是, 如何描述和判断无限个数相加的总和是一个数?

在展开讨论之前, 还是进一步观察上述第一个例子. 记前  $n$  项的和为

$$S_n = 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n},$$

我们看到,  $S_n$  与 2 的误差满足

$$|S_n - 2| = \frac{1}{2^n}.$$

随着求和的项越来越多, 也就是  $n$  越来越大,  $S_n$  与 2 的误差越来越小. 例如要想  $S_n$  与 2 的误差小于  $\frac{1}{10}$ , 只要求和的项超过 4 项就足够了, 即

$$|S_n - 2| = \frac{1}{2^n} < \frac{1}{10} \quad \text{当 } n > 4.$$

要想  $S_n$  与 2 的误差小于  $\frac{1}{100}$ , 只要求和的项超过7项就足够了, 即

$$|S_n - 2| = \frac{1}{2^n} < \frac{1}{100} \text{ 当 } n > 7.$$

总之, 无论你要求误差如何小, 比如小于任何一个正数  $\varepsilon$ , 只要求和的项数  $n$  是一个满足  $n > \left\lceil \frac{\ln \varepsilon}{\ln 2} \right\rceil$  的正整数. 即

$$|S_n - 2| = \frac{1}{2^n} < \varepsilon, \text{ 当 } n > \left\lceil \frac{\ln \varepsilon}{\ln 2} \right\rceil.$$

因此, 我们有理由相信, 无限多个数  $1, \frac{1}{2}, \frac{1}{2^2}, \dots, \frac{1}{2^n}, \dots$  相加的总和应该是2.

对一般情形, 为了从数学上更加准确地表述, 我需要借助中学所学的两个逻辑量词: 全称量词 (简称“任意”, 常用符号  $\forall$  表示) 与 存在量词 (简称“存在”, 常用符号  $\exists$  表示). 这里对它们的使用规则再作回顾. 首先观察以下例子:

### 例 1.3.1 下列等式

$$2 + 1 = 1 + 2, 2 + 2 = 2 + 2, 2 + 3 = 3 + 2, 2 + 4 = 4 + 2, \dots$$

其中的省略号, 隐含了一个“无限”事实, 如果利用全称量词, 它写成

对任意正整数  $x$ , 有  $2 + x = x + 2$ .

**例 1.3.2** 无限多个数  $a_1, a_2, \dots, a_n, \dots$  (也称为“数列”, 或简记为  $\{a_n\}$ ) 有界. 这个命题蕴含了存在一个数  $M$ , 使得  $|a_n| \leq M$  对任意的  $n$  都成立. 用全称量词可以表述为

存在一个实数  $M$ , 使得对任意的正整数  $n$ , 有

$$|a_n| \leq M.$$

一个含有全称量词的命题称为全称命题, 通常表述为

对任意  $x \in U$ ,  $A(x)$  成立,

这里  $U$  是给定的集合(范围),  $A$  是一个含变量  $x$  的命题. 它的意思是说“对所有  $U$  中的  $x$ ,  $A(x)$  成立”. 因此“任意”是在一定范围内的“任意”.

类似地, 存在命题:

存在  $x \in U$  使得  $A(x)$  成立

就是说“至少有一个  $U$  中的元素  $x$  使得  $A(x)$  成立”. 因此“存在”也是在一定范围内的“存在”.

当一个命题中出现两个以上的量词, 有些情况是比较简单的. 例如: 对任意整数  $x$ , 对任意整数  $y$ ,  $x + y = y + x$ . 显然这里两个全称量词“任意-任意”的顺序无关大局, 因

此可以把这个命题简写为: 对任意整数  $x$  和  $y$ ,  $x + y = y + x$ . 同样, 两个以上的存在量词相邻出现, 它们的顺序也不要紧. 例如: “存在整数  $x$ , 存在整数  $y$ , 使得  $x + y = 2$ ,  $x + 2y = 3$ ”. 可以说“存在整数  $x$  和  $y$ , 使得  $x + y = 2$ ,  $x + 2y = 3$ ”. 因此, 相同类型的量词可以交换次序或者合并.

但是对于不同类型的量词来说, 这条规则不成立.

**例 1.3.3** 对任意的整数  $a$ , 存在一个整数  $b$ , 满足  $b = a + 1$ .

这是一个“任意-存在”命题, 如果改变任意和存在的顺序, 则原命题就会变成“存在一个整数  $b$ , 对任意的整数  $a$ , 满足  $b = a + 1$ ”. 这显然是一个错误的命题.

否定一个全称命题只需要找到一个反例, 即全称命题“对任意  $x \in U$ ,  $A(x)$ ”的否定, 等价于存在命题“存在  $x \in U$  使得非  $A(x)$ ”.

否定一个存在命题则需要说明所有的情形都不成立, 即存在命题“存在  $x \in U$  使得  $A(x)$ ”的否定为全称命题“对任意  $x \in U$ , 非  $A(x)$ ”.

对一个含有不同类型量词的命题来说, 它的否命题可以通过改变量词的顺序(或者说是改变量词的类型)得到.

**例 1.3.4** 例1.3.2 的否命题: 数列  $\{a_n\}$  无界.

表示方式为对任意的正数  $M$ , 存在一个正整数  $n$  使得  $|a_n| > M$ .

可以看出, 数列有界的“存在-任意”以及最后陈述“ $|a_n| \leq M$ ”的命题的否命题变成了“任意-存在”并且用“ $|a_n| > M$ ”否定最后陈述的命题.

以上关于全称量词和存在量词以及它们的使用规则的简单回顾, 对用来表述无限求和已经足够了.

下面, 我们给出无限个数求和的一般定义.

设有无限多个数  $a_1, a_2, \dots, a_n, \dots$ , 把它们形式上相加称为无穷级数, 或简称为级数, 记为

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \dots + a_n + \dots$$

前  $n$  项相加称为级数的部分和, 记为

$$S_n = \sum_{k=1}^n a_k = a_1 + a_2 + \dots + a_n.$$

**定义 1.14** 对于级数  $\sum_{n=1}^{\infty} a_n$ , 如果有一个数  $a$ , 使得下列命题成立

对任意的正数  $\varepsilon$ , 存在一个正整数  $N$ , 使得当  $n > N$  时(即对任意大于  $N$  的正整

数  $n$  ), 下列不等式成立

$$|S_n - a| = \left| \sum_{k=1}^n a_k - a \right| < \varepsilon.$$

那么称级数的总和为  $a$ , 记

$$a = \sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots + a_n + \cdots,$$

或者说级数收敛于  $a$ . 否则称级数是发散的.

根据上述定义, 先前所举的例子  $1 + 2 + 2^2 + \cdots + 2^n + \cdots$  根本就不是一个数, 当然也就不能按照数来进行运算.

**例 1.3.5** 讨论级数  $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)} + \cdots$  的敛散性.

解 因为级数的前  $n$  项部分和为

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

因此, 对任意的正数  $\varepsilon$ , 只要取  $N$  是满足  $N > \frac{1}{\varepsilon} - 1$  的正整数, 那么当  $n > N$  时, 就有

$$|S_n - 1| = \frac{1}{n+1} < \frac{1}{N+1} < \varepsilon.$$

也就是级数收敛于 1.

**例 1.3.6** 设  $|q| < 1$ , 讨论几何级数  $\sum_{n=1}^{\infty} q^n = 1 + q + q^2 + \cdots$  的敛散性.

解 因为级数的前  $n$  项部分和为

$$S_n = 1 + q + q^2 + \cdots + q^{n-1} = \frac{1 - q^n}{1 - q},$$

所以对任意的正数  $\varepsilon$ , 只要取  $N$  是满足  $N > \left| \frac{\ln \varepsilon (1-q)}{\ln |q|} \right|$  的一个正整数, 那么当  $n > N$  时, 就有

$$\left| S_n - \frac{1}{1-q} \right| = \frac{|q|^n}{1-q} < \frac{|q|^N}{1-q} < \varepsilon,$$

所以对于公比  $|q| < 1$ , 几何级数收敛

$$1 + q + q^2 + \cdots + q^n + \cdots = \frac{1}{1-q}, \quad |q| < 1.$$

从级数收敛的定义中不难看出, 级数的部分和  $S_n$  实际上形成了一个数列  $\{S_1, S_2, S_3, \dots, S_n, \dots\}$ , 因此上述关于无穷级数的收敛性定义, 也可以独立地用来

定义数列的收敛性, 也就是数列  $\{S_n\}$  的极限, 记为

$$\lim_{n \rightarrow +\infty} S_n = a.$$

这里不再重复给出数列收敛(有有限极限)的定义, 而是通过下列例子给予说明.

**例 1.3.7** 设  $|q| < 1$ , 讨论数列  $\{q^n\} = \{1, q, q^2, \dots, q^n, \dots\}$  的敛散性.

**解** 对任意小的正数  $\varepsilon$ , 不妨设  $0 < \varepsilon < 1$ , 只要取  $N$  是满足  $N > \frac{\ln \varepsilon}{\ln |q|}$  的正整数, 那么当  $n > N$  时, 就有

$$|q^n| = |q|^n < |q|^N < \varepsilon,$$

因此

$$\lim_{n \rightarrow +\infty} q^n = 0.$$

这里, 我们限制  $\varepsilon$  的取值范围并不失一般性, 因为当  $\varepsilon \geq 1$  时, 显然有  $|q|^n < 1 \leq \varepsilon$  对任何  $n$  都成立.

**定理 1.15** 如果级数  $\sum_{n=1}^{\infty} a_n$  收敛, 那么数列  $\{a_n\}$  收敛于 0, 或者说

$$\lim_{n \rightarrow +\infty} a_n = 0.$$

**证明** 不妨设级数收敛于  $a$ . 对任意的正数  $\varepsilon$ ,  $\frac{\varepsilon}{2}$  也是任意的正数, 因此, 存在正整数  $N$ , 使得当  $n > N$  (当然  $n+1 > N$ ) 时, 有

$$|S_n - a| < \frac{\varepsilon}{2}, \quad |S_{n+1} - a| < \frac{\varepsilon}{2},$$

因此有

$$|a_{n+1}| = |S_{n+1} - S_n| \leq |S_{n+1} - a| + |S_n - a| < \varepsilon$$

对  $n > N$  成立, 即  $\lim_{n \rightarrow +\infty} a_n = 0$ . □

读者不难发现, 要用定义说明级数收敛, 必须事先知道收敛的值. 自然要问, 是否存在一些方法, 直接判断级数是否收敛, 而无需事先知道收敛的值. 这个问题在《数学分析》中将会有明确和丰富的答案. 这里作如下分析.

假设级数  $\sum_{n=1}^{\infty} a_n$  收敛, 那么存在一个数  $a$  使得级数满足定义中的命题, 对任意的正数  $\varepsilon$ , 当然  $\frac{\varepsilon}{2}$  也是任意的正数, 因此存在一个正整数  $N$ , 使得当  $n > N$  时, 有

$$|S_n - a| < \frac{\varepsilon}{2}.$$

另取任意的正整数  $p > 0$ ,  $n+p > N$ , 因此也有

$$|S_{n+p} - a| < \frac{\varepsilon}{2}.$$

这样我们就推得

$$|S_{n+p} - S_n| \leq |S_{n+p} - a| + |S_n - a| < \varepsilon,$$

上式就是

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| = |S_{n+p} - S_n| < \varepsilon.$$

其中  $|a_{n+1} + a_{n+2} + \cdots + a_{n+p}|$  是级数求和中从第  $n+1$  ( $n > N$ ) 项开始的任何有限段的和式的绝对值.

反之, 对于任意正数  $\varepsilon$ , 如果存在一个正整数  $N$ , 使得对超过  $N$  的任意一段和满足

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| < \varepsilon,$$

那么级数是收敛的. 这样我们就有了下列定理

**定理 1.16** (Cauchy 柯西 1789-1857 收敛准则) 级数  $\sum_{n=1}^{\infty} a_n$  收敛的充分必要条件是: 于任意正数  $\varepsilon$ , 存在一个正整数  $N$ , 使得不等式

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| = |S_{n+p} - S_n| < \varepsilon,$$

对任意满足  $n > N$  的正整数  $n$  和任意的正整数  $p$  成立.

上述分析实际上已经给出定理的必要性的证明, 而充分性的证明需要用到实数理论中相关的等价命题, 在此不做进一步讨论.

该定理的基本性体现在给出一个与定义等价的命题, 重要性体现在只要看对于充分靠后的、任意有限长度的一段和式的和充分小, 就能断定级数是收敛的, 而无需事先知道级数的收敛的值.

**推论 1.17** 设有两个级数  $\sum_{n=1}^{\infty} a_n$  和  $\sum_{n=1}^{\infty} b_n$ , 如果  $|a_n| \leq b_n$ ,  $n = 1, 2, \dots$ , 那么级数  $\sum_{n=1}^{\infty} b_n$  收敛就一定推出级数  $\sum_{n=1}^{\infty} a_n$  收敛.

**证明** 因为  $\sum_{n=1}^{\infty} b_n$  收敛, 根据Cauchy收敛准则, 对任意的正数  $\varepsilon$ , 一定存在  $N > 0$ , 使得

$$|b_{n+1} + \cdots + b_{n+p}| < \varepsilon$$

对任意的  $n > N$  和任意的正整数  $p$  成立. 那么

$$|a_{n+1} + \cdots + a_{n+p}| \leq |b_{n+1} + \cdots + b_{n+p}| < \varepsilon$$

对任意的  $n > N$  和任意的正整数  $p$  也成立. 所以  $\sum_{n=1}^{\infty} a_n$  收敛.  $\square$

**例 1.3.8** 判断下列级数的收敛性.

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots .$$

解 我们无法猜测该级数具体的收敛值, 因此也就无法验证它是否满足收敛的定义, 但是有了Cauchy 收敛准则, 对于任意的正数  $\varepsilon$ , 取  $N$  是满足  $N > \frac{\varepsilon}{2}$  的正整数, 对于满足  $n > N$  的正整数  $n$  和任意的正整数  $p$ , 有

$$\begin{aligned} & \left| \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots + \frac{1}{(n+p)!} \right| \\ & < \left| \frac{1}{n(n+1)} + \frac{1}{(n+1)(n+2)} + \cdots + \frac{1}{(n+p-1)(n+p)} \right| \\ & < \left| \frac{1}{n} - \frac{1}{n+1} + \frac{1}{n+1} - \frac{1}{n+2} + \cdots + \frac{1}{n+p-1} - \frac{1}{n+p} \right| \\ & = \left| \frac{1}{n} - \frac{1}{n+p} \right| < \frac{2}{n} < \frac{2}{N} < \varepsilon, \end{aligned}$$

因此, 我们不需要事先知道级数收敛的值, 根据Cauchy 收敛准则就能判断出该级数一定收敛于一个数. 记这个数为  $e$ :

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots.$$

对于数  $e$ , 不难发现  $e > 2$ , 同时有

$$\frac{1}{n!} < \frac{1}{2^{n-1}}, \quad n \geq 2,$$

因此

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \cdots < 1 + 1 + \frac{1}{2} + \cdots \frac{1}{2^{n-1}} + \cdots = 3,$$

也就是说, 数  $e$  满足  $2 < e < 3$ . 不仅如此, 我们还可以证明

**定理 1.18** 数  $e$  是一个无理数.

**证明** 因为  $2 < e < 3$ , 所以  $e$  不能是整数, 假设  $e = \frac{p}{q}$  是有理数, 其中  $p, q$  是正整数, 且  $q$  必然满足  $q \geq 2$ . 这样我们就有

$$\begin{aligned} q!e = (q-1)!p &= \left[ q! + q! + \frac{q!}{2!} + \cdots + \frac{q!}{(q-1)!} + 1 \right] \\ &\quad + \frac{1}{q+1} + \frac{1}{(q+2)(q+1)} + \frac{1}{(q+3)(q+2)(q+1)} + \cdots. \end{aligned}$$

注意到上式左边和右边方括号内都是整数, 而右边剩余的部分满足

$$\begin{aligned} 0 &< \frac{1}{q+1} + \frac{1}{(q+2)(q+1)} + \frac{1}{(q+3)(q+2)(q+1)} + \cdots \\ &< \frac{1}{3} \left( 1 + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \cdots \right) = \frac{1}{2}. \end{aligned}$$

不是一个整数, 因此矛盾. 矛盾说明  $e$  只能是无理数.  $\square$

我们知道, 有限个有理数相加, 其结果还是有理数. 但是数  $e$  的例子却表明, 无限个有理数相加, 其结果可能不再是有理数. 这种现象我们实际上已经是接触过的. 例如关

于十进制小数

$$a = a_0.a_1a_2a_3 \dots,$$

也可以表示为一个级数

$$a = a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3} + \dots$$

其中,  $a_0$  是整数,  $a_1, a_2, a_3, \dots$  是取值于  $\{0, 1, 2, \dots, 9\}$  的整数.

独立地看, 根据 Cauchy 收敛准则, 任何这样的级数一定是收敛的, 因为

$$\left| \frac{a_{n+1}}{10^{n+1}} + \dots + \frac{a_{n+p}}{10^{n+p}} \right| \leq \frac{9}{10^{n+1}} \left( 1 + \frac{1}{10} + \dots + \frac{1}{10^{n+p-1}} \right) < \frac{1}{10^n}$$

当小数部分是有限的, 即从某项开始  $a_k = 0$ ,  $k > k_0$ , 显然  $a$  是有理数.

当小数部分出现无限循环时, 例如当  $a_{3k+1} = a_1, a_{3k+2} = a_2, a_{3k+3} = a_3, k = 1, 2, \dots$  时, 该数表示为

$$a = a_0 + 0.\dot{a}_1\dot{a}_2\dot{a}_3$$

其中小数部分满足

$$10^3(0.\dot{a}_1\dot{a}_2\dot{a}_3) = a_1a_2a_3 + 0.\dot{a}_1\dot{a}_2\dot{a}_3,$$

因此

$$0.\dot{a}_1\dot{a}_2\dot{a}_3 = \frac{a_1a_2a_3}{999}$$

是一个有理数, 推出  $a$  也是有理数.

当小数部分出现无限不循环时, 我们知道它一定收敛于一个数, 这个数称为无理数.

**注记** 在数学中, 实数理论和极限理论是基础性的内容. 这里, 虽然用到了有关极限的知识, 仅仅是用来讨论无限求和的问题, 关于极限的内容, 在大学任何一门《数学分析》或《微积分》课程中还会详细介绍.

关于实数, 尤其是无理数, 可通过十进制小数构造, 也可借助 Cauchy 收敛准则的思想, 通过定义一种“有理数的 Cauchy 数列”来构造, 还可以通过所谓的 Dedekind (戴德金 1831-1916) 分割来构造. 我们将在第 3 讲中介绍 Dedekind 的方法.

## §1.4 无限远点

本节将从几何的角度讨论“无限”的问题

### 1° 射影变换下的无限远点

**两直线上的射影** 设想平面上有两条相交直线  $l$  和  $l'$ , 交点为  $O$ . 取两直线外一点  $Q$ , 那么连接  $Q$  与  $l$  上任意一点  $P$  的直线  $\overline{QP}$ , 交直线  $l'$  于  $P'$  点. 形象地说, 设想有

从  $Q$  发出的灯光, 把  $l$  上的点  $P$  投射到  $l'$  上, 因此  $P'$  就是  $l$  上的点  $P$  在  $l'$  上的影子, 当  $P$  在  $l$  上移动时, 影子  $P'$  在  $l'$  上也随之移动. 这种对应称为透视 或者称为射影.

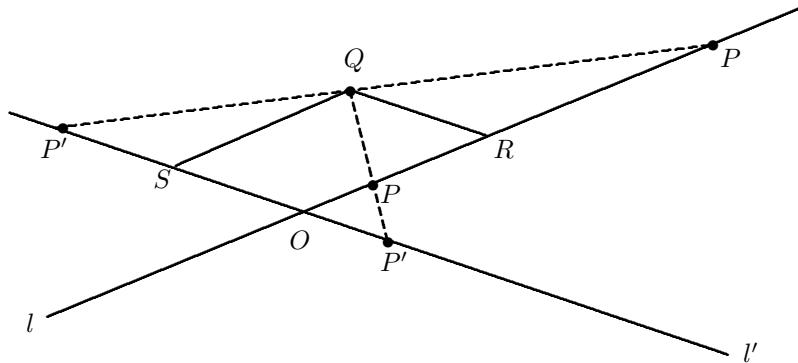


图 1.1

注意到在  $l$  上有一个点比较特殊. 设  $R$  是  $l$  上的点, 使得直线  $\overline{QR}$  与  $l'$  平行. 当  $P$  沿着  $l$  从左边越来越接近  $R$  时, 它的影子  $P'$  在  $l'$  上向右移动越来越远. 当  $P$  到达  $R$  时, 它的影子消失在  $l'$  右边的尽头. 当  $P$  越过  $R$  处在  $R$  的右边时, 它的影子(也就是直线  $\overline{QP}$  与  $l'$  的交点)却出现在  $l'$  的左边, 而且  $P$  从右边越来越接近  $R$  时, 影子  $P'$  就在  $l'$  的左边越来越远之处, 直至消失在左边的尽头.

因此, 可以设想两条平行线  $\overline{QR}$  与  $l'$  在一个假想的点: 无限远点(或称为理想点)相交, 那么  $R$  的影子就是那个无限远点, 而且  $l'$  上右去的无限远点和左去的无限远点应该是同一个点, 因为它们是同一个点  $R$  的影子, 也就是通过射影对应同一个点  $R$ .

不但如此,任何与  $l'$  平行的直线,它们的无限远点也是  $R$  的影子.因此,任何两条平行线的无限远点是同一点,或者说平行线在无限远点相交.

同样, 在  $l'$  上也有一个特殊点. 设  $S$  是  $l'$  上一点, 使得直线  $\overline{QS}$  与  $l$  平行. 当  $P$  沿  $l$  向左越走越远时,  $P'$  就从  $S$  的右边越来越接近  $S$ , 当  $P$  沿  $l$  向右越走越远时,  $P'$  就从  $S$  的左边越来越接近  $S$ . 类似地, 在  $l$  上引进无限远点, 使它的影子对应  $l'$  上的点  $S$ .

这样, 只有引进直线上的无限远点之后, 从  $Q$  点作直线  $l$  和  $l'$  的射影对应, 才能使  $l$  上的每一点都能与  $l'$  上的点对应. 特别  $l$  上的点  $R$  对应  $l'$  上的无限远点, 而  $l$  上的无限远点对应  $l'$  上的点  $S$ .

我们规定，在每一条直线上除了普通点之外，再加上一个“理想点”，也就是“无限远点”。这个点属于和给定直线平行的所有直线。相互不平行的直线有不同的无限远点，再把无限远点的全体看做一条无限远的直线。这样构成的平面称为射影平面。

**性质 1.19** 在射影平面上, 任何两条直线都有唯一的交点.

- (1) 如果两条直线都不是无限远直线, 又不平行, 那么交点就是有限处的一点.
- (2) 如果两条直线都不是无限远直线, 但相互平行, 那么它们的交点就是所共有的无限远点.
- (3) 如果其中一条直线是无限远直线, 另一条直线是非无限远直线, 那么它们的交点就是那条非无限远直线的无限远点.

**两平面之间的射影** 在平面上引进无限远点以及无限远直线后, 再来看一看一个平面到另一个平面的透视问题. 设有平面  $\Pi$  和  $\Pi'$  以及两平面外一点  $Q$ . 对于平面  $\Pi$  上任何一点  $P$ , 直线  $\overline{QP}$  交  $\Pi'$  于  $P'$ , 也就是说  $P'$  是  $P$  在  $\Pi'$  上的影子. 当  $P$  在  $\Pi$  上一条直线上移动时, 它的影子  $P'$  也在  $\Pi'$  的一条直线上移动. 因此投射将  $\Pi$  上的点和直线对应到  $\Pi'$  上的点和直线. 但是有下面的例外, 当  $\overline{QP}$  平行于  $\Pi'$  时,  $P$  在  $\Pi'$  上没有普通的点与之对应.  $\Pi$  上这些特殊点构成一条平行于  $\Pi'$  的直线  $l$ . 这条直线在  $\Pi'$  上没有普通的直线与之对应. 如果我们规定这样的点  $P$  对应  $\overline{QP}$  上无限远点, 那么直线  $l$  就对应  $\Pi'$  上无限远直线.

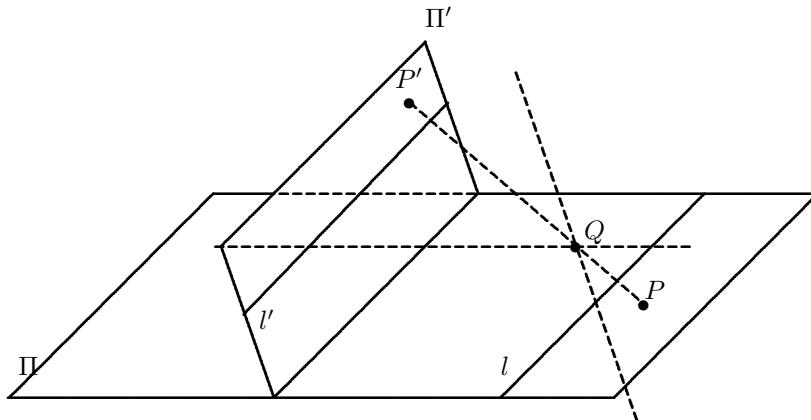


图 1.2

同样,  $\Pi'$  上也有一条直线  $l'$ , 该直线平行于  $\Pi$ , 且直线上任何一点  $P'$  是  $\overline{QP'}$  上无限远点的影子, 因此整个直线是  $\Pi$  无限远直线的对应.

**射影变换** 通过在直线(或平面)上引进无限远点(或无限直线)所形成的直线(或平面)之间就确立了一个1-1 对应, 而且没有例外. 或者说是建立了直线(或平面)上点之间的射影变换. 由此产生一门学科称为**射影几何**, 其主要内容就是要研究在有限次射影变换下, 哪些几何性质是不变的.

需要注意的是, 本节讨论的射影是以一点为中心的**中心射影**, 除了中心射影, 还可建立所谓的**平行射影**, 如同从太阳上照射到地球平面上的阳光一样. 当然, 引进了无限远以后, 也可以把平行射影看成是中心在无限远点的中心射影.

以下所说的“点”或“直线”，都包括无限远点或无限远直线. 特别，用“普通点”或“普通直线”表示那些不是无限远点或无限远直线的点或直线.

**性质 1.20** 在射影变换下保持不变的一些简单性质如下：

(1) 直线上任何一点都可以通过某个射影，对应到无限远点；平面上任何一条直线都可以通过某个射影对应到无限远直线，因此该直线上所有点都对应到无限远点.

(2) 如果平面上三个点，或者更多的点共线，则通过射影后对应的点仍然共线.

(3) 平面上三个或更多的直线交于一点（包括无限远点），那么通过射影后对应的直线仍然交于一点.

根据上述性质，三角形在射影变换下仍然是三角形. 但是诸如直线段的长度，两直线的夹角等等，在射影变换下一般是会改变的，正如人的影子的长度不等于人的身高一样. 等腰三角形或等边三角形可以射影成一个各边不等的三角形.

本节的目的不是全面介绍射影几何，或者说在射影变换下不变的几何性质. 因此我们仅通过一些例子来解释引进无限远点后的一些基本情况.

**Desargues (笛沙格 1591-1661) 定理** Desargues 定理是射影几何中最早发现的结果之一，该定理的具体描述如下：

**定理 1.21** 设  $\triangle ABC$  和  $\triangle A'B'C'$  是平面上两个三角形. 如果连接它们对应顶点的直线  $\overline{AA'}$ ,  $\overline{BB'}$ ,  $\overline{CC'}$  交于一点，那么对应边的延长线的三个交点一定共线（即在一条直线上）.

**证明** 因为引进了无限远点和无限远直线，对应边的延长线  $\overline{BC}$  和  $\overline{B'C'}$ ,  $\overline{AB}$  和  $\overline{A'B'}$ ,  $\overline{AC}$  和  $\overline{A'C'}$  分别都有交点（见性质1.19）. 设交点分别为  $P, Q, R$ . 我们的目的是证明这三点在一条直线上.

连接  $Q, R$ ，并做一射影变换，使  $Q, R$  所在直线投影到无限远直线（也就是将  $Q, R$  投影到无限远点）. 投影后的三角形还是三角形，只是投影后的三角形的两条对应边  $AB$  将平行于  $A'B'$ ,  $AC$  将平行于  $A'C'$ . 因此我们只要证明投影后的第三条对应边  $BC$  和  $B'C'$  延长线的交点也在无限远点，那么投影后三个交点都在无限远直线上. 根据性质1.20，就可知道原来的三个交点一定也在一条直线上.

根据上述分析，现在我们只需证明当三角形  $\triangle ABC$  和  $\triangle A'B'C'$  的两条对应边  $AB$  平行于  $A'B'$ ,  $AC$  平行于  $A'C'$  时，如果对应顶点交于一点  $O$ ，则  $BC$  和  $B'C'$  相互平行，因此它们的延长线的交点在无限远点.

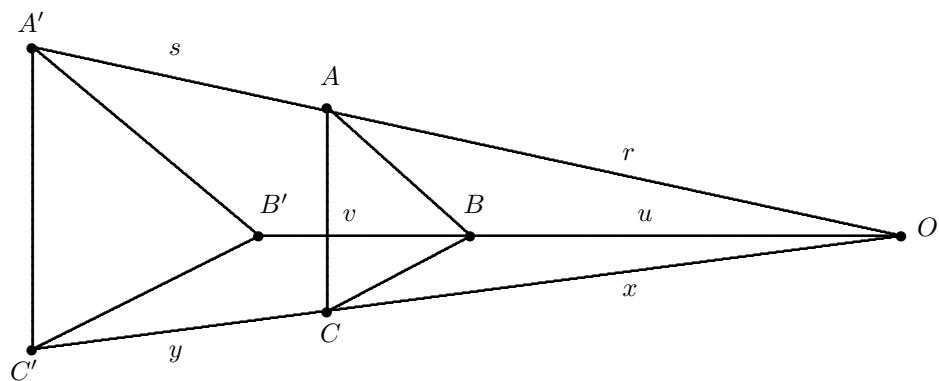


图 1.3

如图所示, 我们有

$$\begin{aligned} AB \parallel A'B' &\text{ 推出 } \frac{u}{v} = \frac{r}{s}, \\ AC \parallel A'C' &\text{ 推出 } \frac{x}{y} = \frac{r}{s}, \end{aligned}$$

因此

$$\frac{u}{v} = \frac{x}{y},$$

即  $BC \parallel B'C'$ . 这样我们就完成了定理的证明.  $\square$

需要说明的是, 这里我们只考虑了平面上的 Desargues 定理, 有关更加详细的内容可参见任何一本关于射影几何的教材.

**齐次坐标** 大家知道, 在平面上引进坐标系后, 平面上的点  $P$  就与一个数组<sup>1-1</sup> 对应, 在直角坐标系下, 这个数组记为  $(x, y)$ , 称为  $P$  的坐标. 有了坐标, 平面上一条直线  $l$  是由这样一些点的轨迹, 这些点的坐标  $(x, y)$  满足一个线性方程

$$l: ax + by + c = 0,$$

其中,  $a, b, c$  是常数, 它确定了直线的位置和方向. 例如

当  $c = 0$ , 表示直线过原点  $O(0, 0)$ ;

当  $a = 0, b = 1, c = 0$ , 对应的直线方程是  $y = 0$ , 即是平面上的  $x$  轴;

当  $a = 1, b = -1, c = 0$ , 确定的直线为  $x = y$ , 它表示过  $O$  点并平分  $x$  轴和  $y$  轴的正向之间的夹角.

同样, 一个坐标满足的二次方程就确定了一个二次曲线, 例如圆、椭圆、抛物线和双

曲线上的点分别满足下列二次方程

$$x^2 + y^2 = r^2, (r > 0)$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, (a > 0, b > 0)$$

$$y = ax^2 + b,$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, (a > 0, b > 0)$$

总之, 引进坐标系后, 把一些诸如点和直线等几何概念加以“数字化”.

反过来, 对一组有序的数组  $(x, y)$ , 我们又可以将它“几何化”, 称它为一个“点”,  $(x, y)$  所满足的方程称为一条“曲线”, 特别当该方程是线性方程时, 就称为“直线”.

平面如此, 空间也是如此, 两个数的数组如此, 三个数, 乃至多个数的数组也是如此, 详情见第 5 讲.

现在面临的问题是, 在添加了无限远点和无限直线的直线或平面上, 无限远点的坐标是什么? 无限远直线的方程是什么? 如果我们仍然把解析几何方法应用到射影几何中去, 就需要建立一个既能包括普通点又能包括无限远点的坐标系.

以下以平面为例.

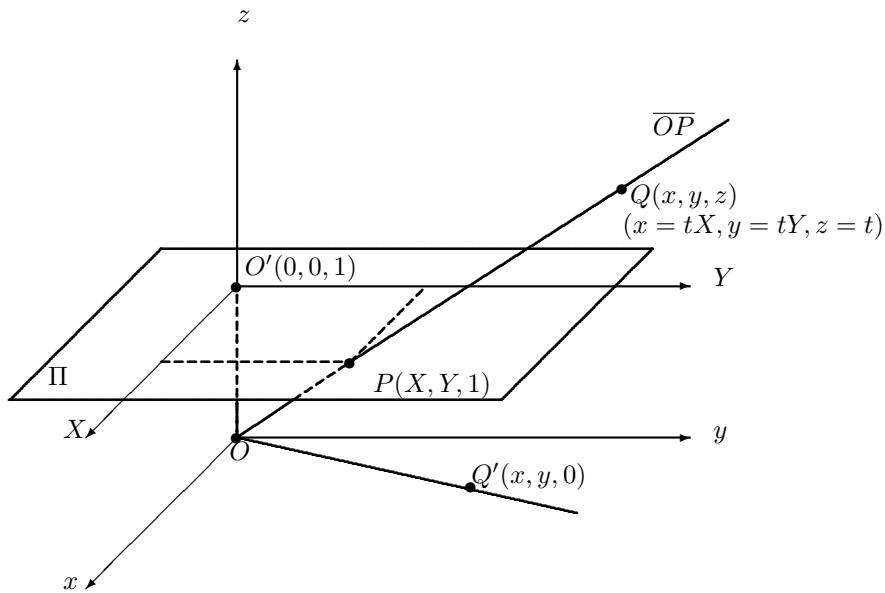


图 1.4

设平面  $\Pi$  上的直角坐标系为  $O'XY$ , 把  $\Pi$  放到一个具有三维直角坐标系的空间  $Oxyz$  中, 使得  $\Pi$  的原点位于  $Oxyz$  中的  $(0, 0, 1)$  点, 并且使得  $\Pi$  平行于  $Oxyz$  中的  $Oxy$  平面. 这样  $\Pi$  上任何一点  $P(X, Y)$  在  $Oxyz$  空间就有了一个三维的坐标  $P(X, Y, 1)$ .

取  $O$  为射影中心点, 那么  $\Pi$  上每一点  $P$  与过该点和  $O$  的直线  $\overline{OP}$  对应. 该直线上任意一个异于  $O$  的点  $Q$  都是  $P$  点的影子, 而  $\Pi$  上的无限远点则与过  $O$  点并与  $\Pi$  平行的直线  $l$  对应.

首先考虑  $P$  是  $\Pi$  上的普通点, 那么  $\overline{OP}$  上任何异于  $O$  的点  $Q$  的坐标  $Q(x, y, z)$  称为  $P$  的齐次坐标, 特别,  $P$  本身的坐标  $(X, Y, 1)$  也是  $P$  的一个齐次坐标. 其它齐次坐标(即直线  $\overline{OP}$  上其它点的坐标)  $(x, y, z)$  满足

$$x = tX, \quad y = tY, \quad z = t, \quad t \neq 0.$$

因此  $P$  的齐次坐标表示为  $(tX, tY, t)$ .

对平面  $\Pi$  上一个普通点  $P$  引进的齐次坐标, 需要用三个数的数组而不是我们熟悉的两个数的数组, 并且不是唯一的, 它带有一个任意非零因子  $t$ .

但是对  $\Pi$  上的无限远点而言, 它对应的是过  $O$  与  $\Pi$  平行的直线(也就是空间中  $Oxy$  平面上的直线), 这条直线上任意异于  $O$  的一点的坐标为  $Q'(x, y, 0)$ , 所以定义  $(x, y, 0)$  为  $\Pi$  上无限远点的齐次坐标.

这样就把包含  $\Pi$  和  $\Pi'$  的无限远点在内的射影平面上任何一点的齐次坐标用三个数的数组  $(x, y, z)$  表示, 其中当  $z \neq 0$  时, 表示普通点  $(X, Y, 1) = (\frac{x}{z}, \frac{y}{z}, 1)$  的齐次坐标,  $z = 0$  时表示无限远点的齐次坐标.

现在考虑平面  $\Pi$  上直线  $l$  的方程

$$l : aX + bY + c = 0.$$

因此  $l$  上的点的齐次坐标满足的方程(简称直线的齐次坐标方程)为

$$l : ax + by + cz = 0,$$

其中  $a, b, c$  为不全为零的常数. 特别,  $\Pi$  上无限远直线  $l_\infty$  的齐次坐标方程如下

$$l_\infty : z = 0.$$

不难看出, 写出直线的齐次坐标方程时, 无限远直线的方程不过是一个特例而已.

从三维空间  $Oxyz$  上看,  $\Pi$  上直线  $l$  的齐次坐标方程就是  $Oxyz$  空间中过直线  $l$  和直线外一点  $O$  (原点) 的平面方程, 这个平面是从  $O$  点发出的光投射到  $l$  上的影子在  $Oxyz$  空间中形成的平面, 而无限远直线的影子所形成的平面正是空间中的  $Oxy$  平面. 平面上每一点的坐标, 都是直线  $l$  上某一点的齐次坐标.

称三元数组  $(a, b, c)$  为直线(包括无限远直线)  $l$  的齐次坐标. 对任意的  $t \neq 0$ ,  $(ta, tb, tc)$  也是同一直线的齐次坐标.

不难发现, 刻画平面上直线(含无限远直线)方程中, 点和直线是完全对称的, 即在

$$ax + by + cz = 0$$

中, 三个齐次坐标  $(x, y, z)$  表示点, 而三个齐次坐标  $(a, b, c)$  表示直线.

## 2° 反演变换下的无限远点

**反演变换** 给定常数  $k$  (这里始终设  $k > 0$ ) 和平面上一点  $O$ , 对于平面上任何异于  $O$  的点  $P$ , 作射线  $OP$ , 在射线上取一点  $P'$  使得  $OP$  的长度和  $OP'$  的长度满足

$$|OP| \cdot |OP'| = k^2,$$

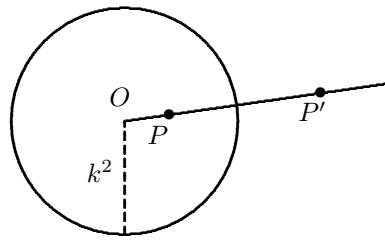


图 1.5

这样就定义了平面上除  $O$  点之外任意一点  $P$  到  $P'$  的一个变换, 称为以  $O$  为反演中心、以  $k$  为反演幂的反演变换. 或简称为反演.  $k$  称为反演的  $P'$  称为  $P$  关于圆  $C$  的反演点.

**性质 1.22** 对给定的反演中心和反演幂, 有

- (1) 如果  $P'$  是  $P$  的反演点, 那么  $P$  是  $P'$  的反演点.
- (2) 设  $C$  是以反演中心  $O$  为圆心,  $k$  为半径的圆, 并称为反演基圆. 那么当  $P$  是圆内的点时,  $P'$  是圆外的一点, 反之如果  $P$  在圆外, 那么  $P'$  就在圆内.

当  $P$  在圆上时, 反演是不变的.

- (3) 当  $P$  沿着固定的射线越来越接近  $O$  时, 反演点沿着同一射线向相反的方向越走越远. 只要  $P$  和  $O$  充分接近,  $P'$  就可以到距离  $O$  任意远的地方.

**无限远点和扩充平面** 按照定义, 在平面上的反演变换只有一点是例外, 即反演中心  $O$  没有对应的反演点, 也没有任何一点以  $O$  作为反演点.

为此, 类似于在射影变换时的做法, 我们引进一个“理想点”, 称为**无限远点**, 记为  $\infty$ , 使它与  $O$  互为反演点. 但是, 与射影变换情形的区别在于, 因为  $O$  只有一个点, 因此在不同的射线上得到的无限远点只能认为是一个点. 这样, 我们在考虑反演变换时, 在平面上添加无限远点, 而且只能添加一个无限远点. 通过添加一个无限远点  $\infty$ , 使得平面上任何一点 (包括  $O$  和  $\infty$ ) 都有反演点与之对应.

如果把平面看成是复平面  $\mathbb{C}$ , 并设反演中心为  $z_0$ , 那么反演变换就转化为复数

$z \rightarrow z'$  的变换

$$z \mapsto z' = z_0 + \frac{k^2}{\bar{z} - \bar{z}_0} = z_0 + \frac{k^2}{|\bar{z} - \bar{z}_0|^2}(z - z_0),$$

这里对一个复数  $z = x + iy$ ,  $\bar{z} = x - iy$  表示  $z$  的复共轭. 添加无限远点后, 我们记

$$\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\},$$

并称为扩充复平面.

对于反演变换的如下结论.

**定理 1.23** 设  $O$  是反演中心,  $k > 0$  是反演幂.  $C$  是以  $O$  为圆心, 以  $k$  为半径的基圆, 则在反演变换下

- (1) 过  $O$  的直线变成过  $O$  的直线;
- (2) 过  $O$  的圆变成一条不过  $O$  的直线;
- (3) 不过  $O$  的直线变成过  $O$  点的圆; 过  $O$  点的直线仍然变成过  $O$  点的直线.
- (4) 不过  $O$  的圆变成不过  $O$  点的圆.

当引进无限远点以后, 任何直线可以看成是广义的圆. 因此上述结论概括起来就是: **反演变换把圆变成圆.**

**证明** (1) 是显然的, 因为有反演的定义, 过  $O$  的任何一点的反演点仍然在这条直线上.

关于(2)的证明, 我们采取初等的办法, 设  $K$  是过  $O$  的任意一个圆, 其半径为  $r$ , 圆心在  $A$  点,  $OAB$  为圆  $K$  的一条直径. 取  $P$  为  $K$  上任意一个动点, 并设  $\angle POA = \theta$ , 那么在直角三角形  $\triangle OPB$  中,

$$|OP| \sec \theta = |OB|.$$

设  $P'$  是  $P$  的反演点, 因此  $P'$  在射线  $OP$  上, 并根据定义有

$$|OP| \cdot |OP'| = k^2,$$

所以

$$|OP'| = \frac{k^2}{|OP|} = \frac{k^2}{|OB|} \sec \theta.$$

再设  $B$  的反演点为  $B'$ , 因此

$$|OB| \cdot |OB'| = k^2, |OB'| = \frac{k^2}{|OB|},$$

所以

$$|OB'| = |OP'| \cos \theta,$$

这样就推出  $\angle OB'P'$  为直角. 也就是圆  $K$  上任何一点  $P$  关于圆  $C$  的反演点都在垂直  $OB'$ , 垂足在  $B'$  的直线上. 于是(2)得证.

(3)的结论是性质1.22和(2)的直接推论.

为了证明(4), 我们取复平面的原点为反演中心  $O$ ,  $k > 0$  为反演幂, 因此反演变换为

$$z \mapsto z' = \frac{k^2}{\bar{z}}.$$

设  $K$  是一个不过  $O$  的, 以  $z_1$  为圆心, 以  $a$  为半径的圆, 因此圆的方程为

$$|z - z_1|^2 = (z - z_1)(\bar{z} - \bar{z}_1) = a^2.$$

因为  $K$  不过  $O$  点, 所以

$$|z_1|^2 \neq a^2,$$

把  $z = \frac{k^2}{\bar{z}'}$  代入得

$$\left(\frac{k^2}{\bar{z}'} - z_1\right) \left(\frac{k^2}{z'} - \bar{z}_1\right) = a^2,$$

整理得

$$k^4 + |z_1|^2|z'|^2 - k^2(z_1\bar{z}' + \bar{z}_1z') = a^2|z'|^2,$$

或

$$|z'|^2 - \frac{k^2}{|z_1|^2 - a^2}(z_1\bar{z}' + \bar{z}_1z') + \frac{k^4}{|z_1|^2 - a^2} = 0$$

令

$$z_2 = \frac{k^2 z_1}{|z_1|^2 - a^2}, \quad b = \frac{k^2 a}{||z_1|^2 - a^2|},$$

那么  $z'$  满足的方程是一个以  $z_2$  为圆心, 以  $b$  为半径的圆:

$$K' : (z' - z_2)(\bar{z}' - \bar{z}_2) = b^2$$

所以反演把圆  $K$  变换到圆  $K'$ . □

**球极投影** 为了进一步理解扩充复平面, 引进如下球极投影: 在三维坐标空间  $Ouvw$  中, 作一个以  $(0, 0, r)$  为圆心,  $r$  为半径的球

$$u^2 + v^2 + (w - r)^2 = r^2,$$

球上的点  $N(0, 0, 2r)$  称为“北极点”,  $S(0, 0, 0)$  称为“南极点”. 把平面  $Ouv$  看成是复平面, 因此复平面的原点与球的南极点重合.

用直线段将球面的北极点  $N$  与复平面上任意一点  $z$  相连, 此线段交球面于一点  $P(z)$ , 这样就建立起球面上的点与复平面上点(也就是一个复数)的1-1 对应  $z \mapsto P(z)$

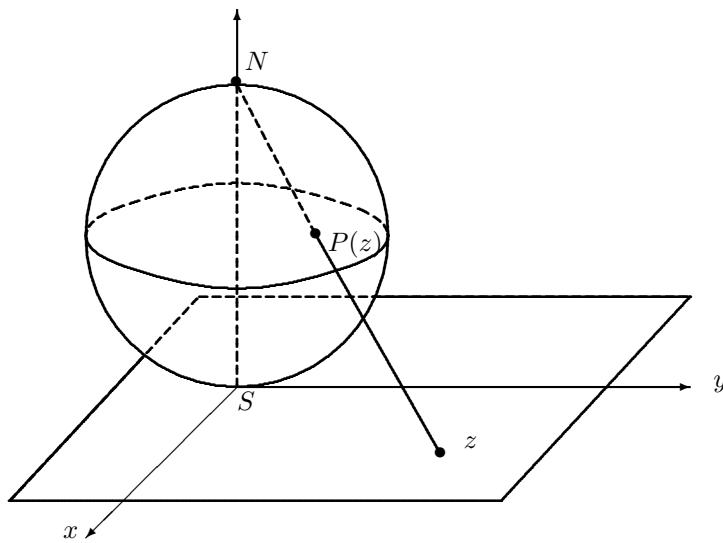


图 1.6

只有北极点例外. 称这种对应为球极投影. 当我们在复平面上引进无限远点  $\infty$  后, 规定  $\infty$  对应球面上的北极点, 因此, 添加了无限远点的扩张复平面  $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$  通过球极投影就与一个二维球面是1-1 对应的. 称这样的球面为 复球面.

为了进一步讨论球极投影的解析表达式, 不妨设球的半径为  $\frac{1}{2}$ , 因此球的圆心的空间坐标为  $(0, 0, \frac{1}{2})$ , 北极点的坐标为  $(0, 0, 1)$ .

设球面上对应  $z = x + iy$  的点  $P(z)$  的空间坐标为  $(u, v, w)$ , 因此

$$u^2 + v^2 + \left(w - \frac{1}{2}\right)^2 = \frac{1}{4}, \quad \text{或} \quad u^2 + v^2 = w(1 - w).$$

因为  $z$  对应的空间坐标为  $(x, y, 0)$ , 根据球极投影,  $(0, 0, 1)$ ,  $(u, v, w)$  和  $(x, y, 0)$  在一条直线上,

$$\frac{u - 0}{x - 0} = \frac{v - 0}{y - 0} = \frac{w - 1}{0 - 1},$$

由此得

$$x = \frac{u}{1 - w}, \quad y = \frac{v}{1 - w}, \quad \text{即} \quad z = \frac{u + iv}{1 - w}.$$

从上式以及  $u^2 + v^2 = w(1 - w)$  不难得到

$$x^2 + y^2 = |z|^2 = \frac{u^2 + v^2}{(1 - w)^2} = \frac{w}{1 - w},$$

解出  $w$  并得到

$$\begin{aligned} u &= \frac{x}{x^2 + y^2 + 1} = \frac{z + \bar{z}}{2(|z|^2 + 1)}, \\ v &= \frac{y}{x^2 + y^2 + 1} = \frac{z - \bar{z}}{2i(|z|^2 + 1)}, \\ w &= \frac{x^2 + y^2}{x^2 + y^2 + 1} = \frac{|z|^2}{|z|^2 + 1}. \end{aligned}$$

这样我们就得到了  $z = x + iy$  和对应点  $P(z) = (u, v, w)$  之间的变换关系.

**球极投影的性质** 有了上述球极投影中复数  $z$  与球面上点的对应关系, 不难看出

**性质 1.24** 设球的圆心的空间坐标为  $\left(0, 0, \frac{1}{2}\right)$ , 半径为  $\frac{1}{2}$ . 那么在球极投影下

(1) 球上的纬线投影到复平面上以  $O$  为圆心的圆, 特别是球上的赤道投影到复平面上以  $O$  为圆心的单位圆.

(2) 球上的经线投影到复平面上过  $O$  的直线.

所谓球面上的纬线即是球面上满足

$$w = c \quad (0 \leq c < 1 \text{ 是常数}), \quad u^2 + v^2 = c(1 - c)$$

的点构成的圆, 因此对应的  $z$  满足

$$|z|^2 = \frac{c}{1 - c},$$

它表示复平面上以  $O$  为圆心的圆. 特别, 球上的赤道 (即  $c = \frac{1}{2}$  的纬线) 就投影到复平面上单位圆  $|z|^2 = 1$ .

所谓球上的经线即是球面上满足

$$\frac{v}{u} = c \quad (\text{常数})$$

的点  $(u, v, w)$ , 对应复平面上是过  $O$  所有满足幅角为常数  $\frac{y}{x} = \tan \theta = c$  的复数, 因此是从  $O$  出发的射线.

如果考虑复平面上以  $O$  为反演中心,  $1$  为反演幂的反演变换, 那么  $z$  和反演点  $z'$  在一条射线上, 并分别在单位圆的两侧. 对应到球面上,  $P(z)$  和  $P(z')$  都在经线上, 并且分别在南北两个半球. 而  $O$  和  $O$  的反演点  $\infty$  对应球面上的南极点和北极点. 这样扩充复平面  $C_\infty$  上的反演变换, 投影到复球面上, 即是南北半球上处在同一经线上的点之间的变换, 其中  $C_\infty$  上原点  $O$  和无限远点  $\infty$  之间的反演变换, 在球面上无非是南极点和北极点之间的变换而已.

## 第 1 讲习题

1. 证明: 对任意的正整数  $n$ , 有

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \left[ \frac{n(n+1)}{2} \right]^2.$$

2. 证明: 对任意的正整数  $n$  和  $p > -1$ , 有

$$(1+p)^n \geqslant 1 + np.$$

3. (平均不等式) 设  $a_1, a_2, \dots, a_n$  是  $n$  个正实数, 则有

$$\frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}} \leqslant \sqrt[n]{a_1 a_2 \cdots a_n} \leqslant \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

4. 设  $f_1 : E_1 \rightarrow E_2$ ,  $f_1 : E_2 \rightarrow E_3$  分别是满射, 证明  $f_2 \circ f_1 : E_1 \rightarrow E_3$  也是满射.

5. 用定义证明:

$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)(2n+1)} = \frac{1}{2}.$$

6. 设  $A, B, C, D$  是直线  $l$  上四个有序点, 定义它们的交比为

$$x = \frac{CA}{CB} / \frac{DA}{DB}.$$

证明: 在中心射影下, 交比是不变的. 即 (如图1.7) 直线  $l$  上四个有序点  $A, B, C, D$  的交比等于直线  $l'$  上与之投影对应的四个有序点  $A', B', C', D'$  的交比相等.

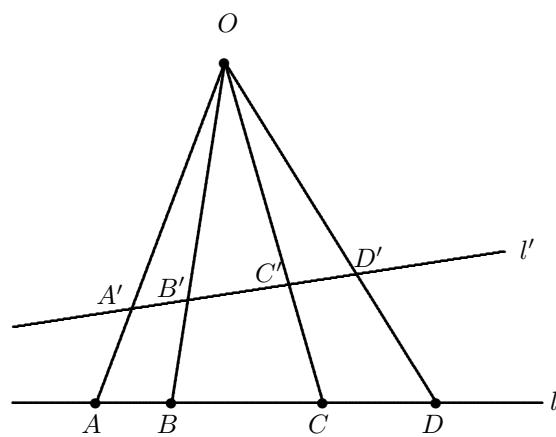


图 1.7

## 第 2 讲 整数

数是数学的基础, 为了更好地理解数系, 首先返回到最简单的基础. 这个基础就是自然数. 自然数在数系中, 乃至在整个数学中扮演的基础性角色, 使得 Kronecker (克罗内克 1823-1891) 发出这样的感慨: “上帝创造了自然数, 其余的都是人的工作”.

本专题介绍整数的有关基本知识, 更详细的内容将会在《初等数论》或《代数学基础》中讲授.

### §2.1 正整数与整数

这里采用与第1讲一致的记号, 记正整数(不包括 0 的自然数)集合如下,

$$\mathbb{N} = \{1, 2, 3, \dots\},$$

在这个集合中有最基本的算术, 算术的基础在于正整数的加法和乘法服从某些规律. 我们往往感兴趣的正是这些具有普遍性的规律, 现将加法和乘法运算等规律罗列如下.

**加法运算:** 对任意的  $a, b \in \mathbb{N}$ , 有唯一确定的正整数, 称为  $a$  与  $b$  的和, 记为  $a + b$ . 加法运算满足下列规则

- (1) 交换律  $a + b = b + a$ .
- (2) 结合律  $(a + b) + c = a + (b + c)$ .

**乘法运算:** 对任意的  $a, b \in \mathbb{N}$ , 有唯一确定的正整数, 称为  $a$  与  $b$  的积, 记为  $a \cdot b$  或  $ab$ . 乘法运算满足下列规则

- (3) 交换律  $ab = ba$ .
- (4) 结合律  $(ab)c = a(bc)$ .
- (5) 分配律  $a(b + c) = ab + ac$ .
- (6) 单位元  $1 \cdot a = a \cdot 1 = a$ .

**顺序关系:** 对任意的  $a, b \in \mathbb{N}$ , 如果存在  $c \in \mathbb{N}$  使得  $a = b + c$ , 则记为  $a > b$  或  $b < a$ . 可以证明, 对任意两个数  $a$  和  $b$ , 下列三个关系有且仅有一个成立.

- (7)  $a < b$ ,  $a = b$ ,  $a > b$ .

顺序关系满足

- (8) 由  $a < b, b < c$  推出  $a < c$ .
- (9) 由  $a < b$  推出  $a + c < b + c$  对任何  $c \in \mathbb{N}$  成立.
- (10) 由  $a < b$  推出  $ac < bc$  对任何  $c \in \mathbb{N}$  成立.

正整数最重要, 最本质的性质是**归纳公理**, 在第 1 讲已经做了介绍, 并由归纳公理出发, 讨论了最小数原理和数学归纳法. 需要补充的是, 除了最小数原理外, 还可以得到所谓**最大数原理**. 为了保持完整性, 现将它们罗列如下:

**归纳公理:** 设  $S \subseteq \mathbb{N}$ , 如果  $S$  满足 (a)  $1 \in S$ , (b) 若  $n \in S$ , 则  $n + 1 \in S$ , 那么  $S = \mathbb{N}$ .

**定理 2.1 (最小数原理)** 设  $S$  是  $\mathbb{N}$  的非空子集. 则  $S$  中必有最小的正整数.

**定理 2.2 (最大数原理)** 设  $S$  是  $\mathbb{N}$  的非空子集. 称  $S$  有上界, 是指存在正整数  $m$ , 使得对任意的  $s \in S$ , 有  $s \leq m$ . 若  $S$  有上界, 则  $S$  中必有最大正整数, 也就是存在  $s_0 \in S$ , 使得对任意的  $s \in S$ , 有  $s \leq s_0$ .

现在将正整数扩展为整数, 一是引进数 "0", 二是引进负数:  $-a, a \in \mathbb{N}$ , 因此整数集合为

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}.$$

整数  $\mathbb{Z}$  中可定义加法, 它适用加法运算 (1) 和 (2), 并且 0 满足

$$0 + a = a + 0 = a, \quad a \in \mathbb{Z}.$$

而且对每个  $a \in \mathbb{Z}$  有唯一的  $x \in \mathbb{Z}$ , 使得

$$a + x = 0.$$

将  $x$  记作  $-a$ , 于是对任意的  $a, b \in \mathbb{Z}$ , 方程

$$a + x = b$$

在  $\mathbb{Z}$  中有唯一解  $x = b - a$ . 这样  $\mathbb{Z}$  中可以作加法的逆运算, 称为减法. 在代数学中, 我们将可作加、减法运算并适合上述运算规则的集合称为**加法群**, 因此  $\mathbb{Z}$  是一个加法群.

整数  $\mathbb{Z}$  中的乘法适合运算规律 (3), (4), (5), (6). 对既能作加法、减运算, 又能作乘法运算的集合, 称为**(交换)环**. 注意, 对乘法, 有

$$0 \cdot a = 0 \quad (a \in \mathbb{Z})$$

$$ab = 0 \quad \text{当且仅当 } a = 0 \text{ 或 } b = 0.$$

具有这种性质的环称为无零因子环. 因此,  $\mathbb{Z}$  是无零因子环, 从而乘法的消去率成立, 即对任何  $b \neq 0$ ,

$$ab = cb \text{ 推出 } a = c.$$

整数  $\mathbb{Z}$  中也有顺序的概念, 它适合规律(7),(8),(9), 只是在规律(10)应增加条件  $c > 0$ , 即修改成

(10') 由  $a < b$  推得  $ac < bc$ , 其中  $c > 0$ .

其他类似正整数集  $\mathbb{N}$  的性质还有

**例 2.1.1** 对整数集  $\mathbb{Z}$  中的非空子集  $E$ , 若  $E$  有下(或上)界, 即若存在  $m \in \mathbb{Z}$ , 使得对任意的  $n \in E$ , 有  $n \geq m$  (或  $n \leq m$ ), 则  $E$  中必有最小(或大)整数.

**证明** 设  $E$  的下界为  $l \in \mathbb{Z}$ , 则对任意的  $n \in E$ ,  $n - l + 1 > 0$ , 并且  $E' = \{n - l + 1 \mid n \in E\} \subset \mathbb{N}$  是正整数的非空子集, 因此  $E'$  有最小数  $m > 0$ , 这样  $n = m + l - 1$  就是  $E$  的最小数.  $\square$

整数  $\mathbb{Z}$  中还可以定义绝对值:

$$|a| = \begin{cases} a, & a \in \mathbb{N} \\ 0, & a = 0 \\ -a, & -a \in \mathbb{N} \end{cases}$$

因此  $|a|$  是非负整数. 绝对值由如下性质:

$$(11) |ab| = |a||b|.$$

$$(12) (\text{三角不等式}) |a + b| \leq |a| + |b|.$$

以上就是正整数集合  $\mathbb{N}$  与整数集合  $\mathbb{Z}$  的一些基本常识.

## §2.2 数的整除性

在上述讨论中, 整数集合  $\mathbb{Z}$  除了加法和乘法外, 还可以做减法, 即加法的逆运算. 但是一般不能作除法, 也就是说对两个整数  $a, b$ ,  $b \neq 0$ ,  $\frac{a}{b}$  不一定是整数, 或者说不一定存在整数  $c$ , 使得  $a = bc$ , 由此引出整数的第一个基本概念: 数的整除性.

**定义 2.3** 设  $a, b$  是两个整数,  $b \neq 0$ . 如果存在一个整数  $c$  使得  $a = bc$ , 则称  $b$  整除  $a$ , 用  $b|a$  表示, 并称  $b$  是  $a$  的一个约数(或因子), 而  $a$  为  $b$  的倍数. 如果不存在上述整数  $c$ , 则称  $b$  不整除  $a$ , 用  $b \nmid a$  表示.

整除的上述定义, 立即导出下列基本性质.

**定理 2.4** 设  $a, b, c$  是整数.

(1) 若  $b|a$ , 且  $c \neq 0$ , 则  $bc|ac$ , 反之亦然, 特别地  $b|a$  等价于  $(\pm b)|(\pm a)$ .

(2) 若  $b|c$ , 且  $c|a$ , 则  $b|a$ .

(3) 若  $b|a$ , 且  $b|c$ , 则  $b|(xa + yc)$ , 其中  $x, y \in \mathbb{Z}$ .

(4) 若  $b|a$ , 且  $a \neq 0$ , 则  $|b| \leq |a|$ ; 于是若  $b|a$ , 且  $a|b$ , 则  $|a| = |b|$ .

从该定理出发, 我们有下列两个定理

**定理 2.5 (带余除法)** 设  $a, b$  为整数  $b \neq 0$ , 则存在唯一的一对整数  $q, r$ , 使得

$$a = bq + r, \quad 0 \leq r < |b|,$$

整数  $q$  称为  $a$  被  $b$  除的商,  $r$  称为  $a$  被  $b$  除的余数. 显然  $b|a$  当且仅当  $r = 0$ , 此时  $q = \frac{a}{b}$  是一个整数.

**证明** 设  $b \nmid a$ , 令

$$E = \{a - bk \mid k \in \mathbb{Z}\}$$

易知  $E \cap \mathbb{N}$  非空 (例如取  $k = -nb$ , 所以  $a - bk = a + nb^2$ , 当  $n$  是足够大的正整数时, 能保证  $a - bk > 0$ , 也就是  $E$  中有正整数), 根据最小数原理,  $E \cap \mathbb{N}$  有最小数, 也就是  $E$  有最小正整数, 设为

$$r = a - bq > 0, \quad q \in \mathbb{Z}.$$

显然  $r \neq |b|$ , 否则  $a = b(q \pm 1)$ , 这与  $b \nmid a$  矛盾.

若  $r > |b|$ , 记  $r' = r - |b| > 0$ , 那么  $r' = r - (\pm b) = a - b(q \pm 1) \in E$ , 而且  $r > r' > 0$ , 这与  $r$  是  $E$  中的最小正整数相矛盾. 所以  $r < |b|$ .

要证明唯一性, 可假设如果另有一对整数  $q_1, r_1$  满足定理要求

$$a = bq_1 + r_1, \quad 0 \leq r_1 < |b|,$$

那么

$$b(q - q_1) + (r - r_1) = 0$$

由此推出  $b|(r - r_1)$ . 但是  $0 \leq r - r_1 < |b|$ , 所以只有  $r - r_1 = 0$ , 从而  $q - q_1 = 0$ .  $\square$

作为带余除法最简单的例子是  $b = 2$  的情形, 也就是整数被 2 除的余数只有两种可能的值, 0 或 1. 因此余数为 0 (即能被 2 整除) 的整数为偶数, 余数为 1 的整数为奇数, 偶数和奇数的一般表达式为

$$2k, \quad 2k + 1, \quad k \in \mathbb{Z}.$$

类似地考虑被 3 除的整数, 其余数只能有 0, 1, 2 三种可能. 因此被 3 除的整数分为三类, 分别是

$$3k; \quad 3k + 1; \quad 3k + 2, \quad k \in \mathbb{Z}.$$

数的整除性引出另一个重要概念, 即

**定义 2.6** 设  $a, b$  是不全为零的整数. 如果整数  $d$  满足:

(1)  $d$  是  $a$  和  $b$  的公共约数(简称公约数或公因子), 即  $d|a, d|b$ .

(2)  $d$  是  $a$  和  $b$  的公约数中最大的, 即如果另有一个公约数  $d'$ , 则  $d' \leq d$ .

则称  $d$  是  $a, b$  的最大公约数(或称最大公因子).

注意, 如果  $d$  是  $a$  和  $b$  的公约数, 则  $-d$  也是公约数, 因此最大公约数一定是正整数. 记两个数  $a$  和  $b$  的最大公约数为

$$d = (a, b).$$

例如  $(4, 6) = 2, (6, 9) = 3, (10, 9) = 1$ .

特别, 如果两个整数  $a$  和  $b$  的最大公约数  $(a, b) = 1$ , 则称  $a$  和  $b$  互素.

这里我们只考虑两个数的最大公约数问题, 对包含两个以上数组的最大公约数记为

$$d = (a_1, a_2, \dots, a_n)$$

它表示  $d$  是  $a_1, a_2, \dots, a_n$  最大的公共约数.

特别当  $(a_1, a_2, \dots, a_n) = 1$  时, 我们称数组  $a_1, a_2, \dots, a_n$  互素.

如果数组  $a_1, a_2, \dots, a_n$  任何两个数是互素的:  $(a_i, a_j) = 1, i \neq j$ , 那么称数组  $a_1, a_2, \dots, a_n$  两两互素.

很明显, 两两互素的数组一定是互素的. 但一组互素的数未必是两两互素的, 例如  $(2, 3, 4) = 1$ , 但其中的 2 和 4 不是互素的.

### §2.3 Euclid 辗转相除法

现在的问题是如何求两个数的最大公约数. 为了解决这个问题, 首先根据带余除法, 给出下面的引理.

**引理 2.7** 对任意两个整数  $a$  和  $b$ , 由带余除法可知

$$a = bq + r.$$

那么

$$(a, b) = (b, r)$$

也就是说  $a$  和  $b$  的最大公约数是  $a$  被  $b$  除的余数  $r$  与  $b$  的最大公约数.

这是因为如果一个数  $u$  是  $a$  和  $b$  的公约数, 则存在整数  $s, t$  使得

$$a = su, \quad b = tu,$$

所以  $u$  也能整除  $r = a - bq = su - tuq = (s - tq)u$ , 因此  $a$  和  $b$  的任何公约数也是  $b$  和  $r$  的公约数, 反过来也是对的: 如果一个整数能够整除  $b$  和  $r$ , 当然也能整除  $a = bq + r$  和  $b$ . 所以  $a, b$  和  $b, r$  的公约数是一样的, 当然就有  $(a, b) = (b, r)$ . 特别当  $r = 0$  时 (即  $a$  能被  $b$  整除), 显然有  $(a, b) = (b, 0) = b$ .

下面, 我们给出著名的Euclid (欧几里得约公元前330 - 公元前275) 辗转相除法.

对任意两个整数  $a$  和  $b$ , 不妨设  $b > 0$ , 反复利用带余除法,

$$\begin{aligned} a &= bq_0 + r_0, & 0 < r_0 < b, \\ b &= r_0q_1 + r_1, & 0 < r_1 < r_0, \\ r_0 &= r_1q_2 + r_2, & 0 < r_2 < r_1, \\ r_1 &= r_2q_3 + r_3, & 0 < r_3 < r_2, \\ &\dots & \dots \\ r_{n-2} &= r_{n-1}q_n + r_n, & 0 < r_n < r_{n-1}, \\ r_{n-1} &= r_nq_{n+1} + 0. \end{aligned}$$

只要余数  $r_1, r_2, \dots$  不出现 0, 就一直做下去. 因为

$$b > r_1 > r_2 > r_3 > \dots > 0$$

而且余数都是正整数, 所以辗转相除的过程不会无限制的继续下去, 最多  $n \leq b$  步就结束了, 也就有  $r_{n-1} = r_nq_{n+1} + 0$ .

利用引理2.7, 我们得到

$$(a, b) = (b, r_0) = (r_0, r_1) = \dots = (r_{n-1}, r_n) = (r_n, 0) = r_n$$

也就是说通过上述“辗转相除”, 直到最后一个不为零的余数就是  $a$  和  $b$  的最大公约数.

上述过程还直接导致一个有意思的结果.

从辗转相除的第一个方程得到  $r_0 = a - bq_0$ , 因此  $r_0$  能表示为  $a$  和  $b$  的整系数线性组合

$$r_0 = k_0a + l_0b,$$

这里  $k_0 = 1, l_0 = -q_0$  都是整数. 代入下一个方程有

$$r_1 = b - r_0q_1 = b - (k_0a + l_0b)q_1 = k_1a + l_1b$$

其中  $k_1, l_1$  也是整数. 重复上述过程, 最后有

$$r_n = (a, b) = ka + lb.$$

其中  $k, l$  为整数. 这样不但求出任意两个整数  $a$  和  $b$  的最大公约数  $(a, b)$ , 还得到方程  $ka + lb = (a, b)$  一对整数解  $k, l$ , 也就是得到下列结论:

**定理 2.8 (Bezout)** 对任意两个整数  $a$  和  $b$ , 存在一对整数  $k, l$  使得  $a$  和  $b$  的最大公约数表示为  $a$  和  $b$  的整系数线性组合

$$ka + lb = (a, b)$$

这个公式称为 *Bezout* 公式. 特别当  $a$  和  $b$  互素时(即  $(a, b) = 1$ ), 下列方程有整数解

$$ka + lb = 1.$$

顺便指出, *Bezout* 公式可以独立于辗转相除法证明, 以便向多个数的数组进行推广. 这里为了简化, 我们把对包含两个数的数组的 *Bezout* 公式看作是辗转相除法的推论.

利用 *Bezout* 公式, 可以得到下列结果.

**推论 2.9** 设  $a$  和  $b$  是任意两个整数, 则

(1)  $a$  和  $b$  的每个公约数都是它们最大公约数的约数.

(2)  $(ma, mb) = m(a, b)$ , 这里  $m$  是任意正整数.

**证明** 设  $u$  是  $a$  和  $b$  的公约数, 即  $u|a, u|b$ , 因此  $u$  也能整除  $ka + lb$ , 根据 *Bezout* 公式知,  $u$  能够整除  $(a, b)$ .

为了证明(2), 设  $d = (a, b), e = (ma, mb)$ . 因为  $md|ma, md|mb$ , 所以由(1)可知  $md|(ma, mb)$ , 即  $md|e$ .

另一方面, 由 *Bezout* 公式, 存在整数  $k, l$ , 使得  $d = ka + lb$ . 由此推得

$$md = mka + mlb = k(ma) + l(mb)$$

所以  $e|md$ , 综上得  $e = md$ . □

## §2.4 素数和整数的素因子分解

**定义 2.10** 对于任何大于 1 的正整数  $p$ , 如果  $p$  没有真因子, 即  $p$  的正约数只有 1 和  $p$  自身, 则称  $p$  为素数(或质数), 否则称为合数.

于是, 整个正整数被分为三类: 数 1 作为单独一类, 两外还有素数类和合数类.

素数有很多, 例如

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \dots$$

关于素数最经典的一个结果是

**定理 2.11** 素数有无穷多个.

该定理的证明早在公元前 3 世纪, 就由 Euclid 给出, 证明的方法至今仍是数学推理的一个典范. 也就是今天被称为的“反证法”.

**定理的证明** 假设定理不成立, 也就是说素数的个数只有有限个, 因此用  $p_1, p_2, \dots, p_n$  来表示这全部有限个素数. 考虑整数

$$N = p_1 p_2 \cdots p_n + 1$$

由于  $N > 1$  且  $N \neq p_j, j = 1, \dots, n$ , 因此  $N$  是合数, 故有素因子(素约数)  $p$ , 由于已经假设所有的素数都在这里, 所以  $p$  必然等于  $p_1, \dots, p_n$  中某一个, 因而  $p$  能整除  $N - p_1 p_2 \cdots p_n = 1$ , 也就是  $p$  能整除 1. 这显然是一个荒谬的结论. 导致出现荒谬结论的原因是假设. 因此假设是错误的, 因而它的反面必然是正确的.  $\square$

素数的重要性在于这样一个事实, 即任何大于 1 的正整数都能表示为素数的乘积. 这是因为对于大于 1 的整数  $p$ , 如果它是合数就一定有非平凡的约数  $p = p_1 p_2$ , 如果  $p_1$  和  $p_2$  两者至少有一个是合数, 则可继续分解. 这样的分解只有有限步, 直至分解到不能分解为止.

当一个整数被分解成素数的乘积后, 根据整数乘法的交换性, 这些素因子乘积的次序无关紧要. 那么在不计次序时, 一个整数的素因子分解是不是唯一的? 答案是肯定的, 即下列被称为算术基本定理:

**定理 2.12** (算术基本定理) 每个大于 1 的正整数均可分解成有限个素数的乘积,如果不计素因子在乘积中的次序, 那么分解是唯一的.

我们仍然采取 Euclid 的证明方法, 为此做一些准备.

**引理 2.13** 设  $p$  是素数,  $a$  和  $b$  是两个整数. 如果  $p$  能整除  $ab$ , 那么  $p$  至少能够整除  $a$  和  $b$  中某一个.

推而广之, 设  $p$  是素数,  $a_1, a_2, \dots, a_n$  是  $n$  个整数. 如果  $p$  能整除  $a_1 a_2 \cdots a_n$ , 那么  $p$  至少能够整除  $a_1, a_2, \dots, a_n$  中某一个.

**引理的证明** 因为素数  $p$  只有 1 和自身  $p$  是它的约数. 如果  $p$  不能整除  $a$ , 那么两者一定互素  $(a, p) = 1$ , 根据 Bezout 公式, 存在整数  $k, l$  使得

$$1 = ka + lp.$$

两边乘以  $b$ , 得到

$$b = kab + lpb.$$

因为  $p$  整除  $ab$ , 所以有  $ab = pr$ , 这样

$$b = kpr + lpb = p(kr + lb)$$

也就是  $p$  能整除  $b$ . 所以我们证明了如果  $p$  不能整除  $a$ , 那么一定能够整除  $b$ , 也就是  $p$  至少能够整除  $a$  和  $b$  中某一个.

关于第二个断言的证明, 采用归纳的方法.

当  $n = 2$  时, 即两个数的情形下, 证明已经给出了.

假设  $p$  能整除  $a_1 \cdots a_{n-1}$  推出  $p$  能整除  $a_1, \dots, a_{n-1}$  中至少某一个.

那么当  $p$  能整除  $a_1 \cdots a_{n-1} a_n = (a_1 \cdots a_{n-1}) a_n$  时. 根据  $n = 2$  的情形可知  $p$  能整除  $a_1 \cdots a_{n-1}$  和  $a_n$  中某一个. 如果  $p$  能整除  $a_n$ , 则结论得证, 如果  $p$  能整除  $a_1 \cdots a_{n-1}$ , 根据归纳假设,  $p$  能整除  $a_1, \dots, a_{n-1}$  中的某一个. 无论哪种情形,  $p$  一定能整除  $a_1, \dots, a_{n-1}, a_n$  中某一个.  $\square$

**算术基本定理的证明** 有了上述引理, 我们假设正整数  $a$  有两种素因子分解

$$a = p_1 \cdots p_r = q_1 \cdots q_s,$$

这里  $p_1, \dots, p_r$  和  $q_1, \dots, q_s$  都是素数. 显然  $p_1$  能整除上式左边, 当然也能整除上式右边, 即  $p_1$  能整除  $q_1 \cdots q_s$ . 根据引理  $p_1$  一定能整除  $q_1, \dots, q_s$  至少某一个, 不妨设  $p_1$  能整除  $q_k$ , 由于  $q_k$  是素数, 所以  $q_k = p_1$ , 两边消去这个共同素因子, 在剩余的部分继续上述过程, 最后得到  $p_1, \dots, p_r$  一定是  $q_1, \dots, q_s$  中的一部分因此  $r \leq s$ .

上述做法完全可对称地对  $q_1, \dots, q_s$  实施, 因此得到  $s \leq r$  且  $q_1, \dots, q_s$  是  $p_1, \dots, p_r$  的一部分. 综上分析只有两者完全重合.  $\square$

算术基本定理虽然针对正整数, 但不难推广到所有整数, 如果再把素因子中相等的因子合并, 最后有

**定理 2.14 (标准分解)** 设  $n$  是任意整数, 则他可唯一分解为

$$n = \varepsilon p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$$

这里  $\varepsilon = \pm 1$ ,  $p_1, \dots, p_k$  是两两不同的素数,  $\alpha_1, \dots, \alpha_k$  是正整数. 也称  $\alpha_j$  是分解中素因子  $p_j$  的重数,  $j = 1, \dots, k$ .

通过以上讨论, 我们发现素数是构成正整数的最基本单元, 如同构成物质的基本粒子, 或像人类的“基因”. 人们自然要问, 如何在众多的正整数中筛选出所有的素数? 或者通过一个什么样的表达式产生素数? 在整个正整数中有没有素数分布的“图谱”?

首先看看如何在正整数中筛选出素数. 一种古老和简单的方法, 是对小于一个给定的数  $n$ , 按照  $1, 2, 3, \dots, n$  排列, 然后逐一划掉所有 2 的倍数的那些数(不含 2 本身), 再逐一划掉 3 的倍数(不含 3 本身), 如此下去留下的就是  $n$  以内的素数. 这个过程称为“Eratosthemens (爱拉陀塞姆约公元前274-194) 筛法”.

另一种考虑是研究产生素数的数学公式. Fermat (费马 1601-1665) 曾给出下列公式

$$F(n) = 2^{2^n} + 1$$

(通常称由上式给出的数为“Fermat 数”). 可以验证对  $n = 1, 2, 3, 4$ , Fermat 数是素数, 但是对  $n = 5$ , Fermat 数是合数

$$2^{2^5} + 1 = 641 \cdot 6700417$$

还有其它一些能够产生素数的简洁公式, 例如

$$f(n) = n^2 - n + 41,$$

当  $n = 1, 2, \dots, 40$ ,  $f(n)$  都是素数, 但是  $f(41) = 41^2$  不再是素数. 公式

$$f(n) = n^2 - 79n + 1601$$

从  $n = 1$  直到  $n = 79$ , 都是素数, 但是当  $n = 80$  时就不再是素数了.

虽然人们没有寻找到求出素数的规律性公式, 但发现在  $n$  以内素数的“密度”(或者说“平均分布”)却有一定的规律.

设  $A_n$  表示  $n$  以内的素数的个数, 则  $A_n/n$  近似于  $1/\ln n$ , 而且随着  $n$  的增大, 这样的近似越来越精确. 细节就不再讨论了.

Euler (欧拉 1707-1783) 研究了另外一个问题, 即任意正整数  $n$  以内的, 与  $n$  互素的正整数的个数问题. 为此 Euler 定义了如下函数:

**定义 2.15** 对任意正整数  $n$ , 用  $\varphi(n)$  表示  $n$  以内(即从 1 到  $n$ )且与  $n$  互素的正整数个数. 称为 Euler 函数.

例如  $\varphi(1) = 1$ , 而  $\varphi(12) = 4$ , 因为在不超过 12 的正整数中, 只有 1, 5, 7, 11 与 12 互素. 为了进一步研究 Euler 函数, 首先有

**引理 2.16** 对任意一个素数  $p$  以及  $p^\alpha$  ( $\alpha$  是正整数), 我们有

$$\varphi(p) = p - 1,$$

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha \left(1 - \frac{1}{p}\right).$$

**证明**  $\varphi(p) = p - 1$  是显然的. 现在设正整数  $\alpha > 1$ , 那么在 1 到  $p^\alpha$  中那些是  $p$  倍数的数为

$$p, 2p, 3p, \dots, p^{\alpha-1}p,$$

共有  $p^{\alpha-1}$  个, 在 1 到  $p^\alpha$  中除掉这些  $p$  的倍数的数, 剩余的都与  $p^\alpha$  互素, 互素的数的个数共有  $p^\alpha - p^{\alpha-1}$  个, 这样我们就证明了第二个等式.  $\square$

下列定理表明, 对一般的  $n$ , Euler 函数的函数值也是可以严格计算的.

**定理 2.17** (Euler 定理) 对任意的正整数  $n$ , 设  $n$  的素因子分解为

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k},$$

其中,  $p_1, p_2, \dots, p_k$  是互不相等的  $n$  的素因子,  $\alpha_j, j = 1, \dots, k$  是素因子  $p_j, j = 1, \dots, k$  的重数. 那么 Euler 函数在  $n$  的取值为

$$\varphi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right).$$

**证明** 证明分以下步骤.

(1) 在 1 到  $n$  中是  $p_1$  的倍数的数共有  $p_1^{\alpha_1-1} p_2^{\alpha_2} \cdots p_k^{\alpha_k} = \frac{n}{p_1}$  个, 这些数分别是

$$p_1, 2p_1, 3p_1, \dots, \frac{n}{p_1} p_1,$$

因此, 排除这些  $p_1$  的倍数的数, 就得到在 1 到  $n$  中与  $p_1$  互素的整数的个数

$$n - \frac{n}{p_1} = n \left(1 - \frac{1}{p_1}\right).$$

(2) 同理, 在 1 到  $n$  中是  $p_2$  的倍数的数共有  $\frac{n}{p_2}$  个:

$$p_2, 2p_2, 3p_2, \dots, \frac{n}{p_2} p_2,$$

但其中有些数也是  $p_1$  的倍数. 这些  $p_1$  的倍数出现在

$$1, 2, 3, \dots, \frac{n}{p_2}$$

之中 (因为  $p_1$  与  $p_2$  互素). 除去这些重复的数, 剩余的数就是那些  $p_2$  的倍数但是与  $p_1$  互素的数, 也就是  $1, 2, 3, \dots, \frac{n}{p_2}$  中与  $p_1$  互素的数. 根据 (1), 它们的个数为  $\frac{n}{p_2} \left(1 - \frac{1}{p_1}\right)$ .

把 1 到  $n$  中与  $p_1$  互素的数中, 再排除那些是  $p_2$  的倍数但与  $p_1$  互素的数, 就得到与  $p_1$  和  $p_2$  都互素的数, 它们的个数为

$$n - \frac{n}{p_1} - \frac{n}{p_2} \left(1 - \frac{1}{p_1}\right) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right).$$

(3) 1到  $n$  中, 是  $p_3$  的倍数的数共有  $\frac{n}{p_3}$  个:

$$p_3, 2p_3, 3p_3, \dots, \frac{n}{p_3}p_3,$$

其中有些也是  $p_1$  或  $p_2$  的倍数, 并在(1)和(2)中都已经排除过. 同样因为  $p_1, p_2$  与  $p_3$  互素, 所以这些数出现在

$$1, 2, 3, \dots, \frac{n}{p_3}$$

中, 因此除去这些重复排除的数, 剩余的数正是1 到  $\frac{n}{p_3}$  中与  $p_1$  和  $p_2$  互素的那些数, 它们的个数由(2)可知是  $\frac{n}{p_3} \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right)$ . 因此在1 到  $n$  中与  $p_1, p_2, p_3$  都互素的数的个数为

$$n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) - \frac{n}{p_3} \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \left(1 - \frac{1}{p_3}\right).$$

按同样的程序递推下去, 就得到定理中的结果.  $\square$

由 Euler 定理, 我们还能得到 Euler 函数所满足的一个性质, 这个性质也称为 Euler 函数的积性

**推论 2.18** 对任意两个正整数  $m, n$ , 若  $(m, n) = 1$ , 则

$$\varphi(mn) = \varphi(m)\varphi(n).$$

证明是显然的, 设  $m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ ,  $n = q_1^{\beta_1} q_2^{\beta_2} \cdots q_l^{\beta_l}$ . 因为  $(m, n) = 1$ , 所以两个正整数的素数因子不重复, 利用 Euler 定理就得  $\varphi(mn) = \varphi(m)\varphi(n)$ .

## §2.5 同余和同余类

对一个固定的正整数  $m$ , 根据带余除法, 任何一个整数被  $m$  除的余数只可能是  $0, 1, \dots, m-1$  中的某一个. 因此, 对所有的整数, 根据它们被  $m$  除后的余数, 可以将它们进行分类, 这就是下列关于同余类的定义.

**定义 2.19** 如果两个整数  $a$  和  $b$  同被  $m$  除的余数相等, 也就是  $a - b$  能被  $m$  整除:  $m|(a - b)$ . 则称它们被  $m$  除是“同余”的. 标准的说法是  $a$  和  $b$  模  $m$  同余, 或者称  $a$  模  $m$  同余于  $b$ ,  $m$  称为模. 记为

$$a \equiv b \pmod{m}$$

并称为(模  $m$  的)同余式. 如果  $m \nmid (a - b)$ , 则称  $a$  和  $b$  模  $m$  不同余. 两个数  $a$  和  $b$  同余, 当且仅当

$$m|(a - b), \text{ 或 } a = b + rm.$$

这里  $r$  是某个整数.

两个数之间的同余关系与通常两个数相等有许多类似性质, 我们把它们罗列如下

### 性质 2.20

(1) (自反性)  $a \equiv a \pmod{m}$ ;

(2) (对称性)  $a \equiv b \pmod{m}$ , 则  $b \equiv a \pmod{m}$ ;

(3) (传递性)  $a \equiv b \pmod{m}$ ,  $b \equiv c \pmod{m}$ , 则  $a \equiv c \pmod{m}$ .

如果  $a \equiv a' \pmod{m}$ ,  $b \equiv b' \pmod{m}$ , 那么

(4)  $a \pm b \equiv a' \pm b' \pmod{m}$ ;

(5)  $ab \equiv a'b' \pmod{m}$ .

(6) 设  $m$  为素数, 则

$a \equiv 0 \pmod{m}$  或  $b \equiv 0 \pmod{m}$ , 当且仅当  $ab \equiv 0 \pmod{m}$ .

证明 对(1),(2),(3)直接根据定义可证,

对于(4)和(5), 只要注意到根据同余的等价条件有  $a = a' + rm$ ,  $b = b' + sm$ , 所以

$$a \pm b = a' + b' + (r \pm s)m$$

$$ab = a'b' + (srm \pm a's \pm b'r)m$$

即可完成证明.

对于(6): 若  $m|a$  或  $m|b$ , 则  $m|ab$ , 反之, 若  $m|ab$ , 则  $m$  至少能整除  $a, b$  中某一个(引理 2.13).  $\square$

下面是若干关于模的简单实用的定理.

### 定理 2.21

(1) 若  $ac \equiv bc \pmod{m}$ , 则

$$ac \equiv bc \pmod{\frac{m}{(c, m)}}.$$

特别当  $(c, m) = 1$  时, 有  $a \equiv b \pmod{m}$ .

(2) 若  $a \equiv b \pmod{m}$ ,  $d|m$ , 则  $a \equiv b \pmod{d}$ .

证明 记

$$m_1 = (c, m), \quad m_2 = \frac{m}{(c, m)},$$

因此  $m = m_1 m_2$ ,  $c = km_1$ , 由已知条件知

$$m|(a - b)c, \text{ 或 } m_1 m_2|(a - b)km_1,$$

所以  $m_2|(a - b)k_1$ , 因此

$$m_2|(a - b)c.$$

特别当  $m_1 = (c, m) = 1$  时,  $m = m_2$  不能整除  $c$ , 只能整除  $a - b$ .

若  $d|m$ , 记  $m = dm_1$ . 如果  $m|(a - b)$ , 即存在整数  $k$ , 使得  $a - b = mk = dm_1k$ , 推得  $d|(a - b)$ .  $\square$

**定义 2.22** 设集合  $E$  中的元素之间存在一种关系记为  $a \sim b$ ,  $a, b \in E$ . 如果这种关系满足

自反性:  $a \sim a$ ,  $a \in E$ ;

对称性: 若  $a \sim b$ , 则  $b \sim a$ ;

传递性: 若  $a \sim b$ ,  $b \sim c$ , 则  $a \sim c$ .

那么称这种关系为等价关系. 两个元素具有等价关系  $a \sim b$  也称为  $a$  和  $b$  彼此等价.

因此性质2.20 中的(1)(2)和(3)表明整数之间模  $m$  的同余的关系是一种等价关系. 平面上三角形之间的相似性是一种等价关系, 两个数相等的关系当然也是等价关系. 性质2.20 中的(4)和(5)还说明, 对于相同的模  $m$ , 同余关系如同相等关系一样还保持“加法”、“减法”和“乘法”. 对于相等关系, 有下列结果

$$ab = 0, \text{ 当且仅当 } a = 0 \text{ 或 } b = 0.$$

但是对于同余关系, 必须对模  $m$  加以限制, 也就是性质2.20 中的(6). 如果模不是素数, 上面的结论是不成立的. 例如

$$8 \equiv 2 \pmod{6}, 9 \equiv 3 \pmod{6}, \text{ 但是, } 72 \equiv 0 \pmod{6}.$$

一般来说, 等价关系  $\sim$  将集合  $E$  中元素按彼此等价归为一类, 称为关系  $\sim$  的等价类.

同样, 同余关系也可将整数按模  $m$  是否同余分为若干个两两不相交的类(“家族”), 使得同一个类中的任何两个整数(类中的成员)模  $m$  同余, 不同类中的两个整数(成员)模  $m$  不同余. 每一个这样的类称为模  $m$  的同余类. 具体来讲对一个整数  $a$ , 我们将  $a$  所属的同余类记为  $[a] \pmod{m}$  或简记为  $[a]$ , 这个类包含所有模  $m$  与  $a$  同余的整数

$$[a] = \{x \mid x \in \mathbb{Z}, x \equiv a \pmod{m}\}$$

$a$  是这个类的代表元.

如果  $a \equiv b \pmod{m}$ , 说明  $a$  和  $b$  属于同一类, 所以  $[a] = [b]$ , 否则  $[a]$  和  $[b]$  作为  $a$  和  $b$  为代表的两个类是不相交的.

根据带余除法, 任何整数必与数组  $0, 1, 2, \dots, m - 1$  中某一个模  $m$  同余, 而数组  $0, 1, 2, \dots, m - 1$  内任何两个数彼此模  $m$  不同余, 因此模  $m$  恰好有  $m$  个同余类, 它们分别以余数  $0, 1, 2, \dots, m - 1$  为表元. 记这些同余类形成的集合为

$$\mathbb{Z}_m = \{[0], [1], [2], \dots, [m - 1]\}.$$

这是一个以“类”为元素的集合, 集合的元素的个数为  $m$  个. 例如, 取  $m = 2$ , 则模 2 的同余类只有两个, 一个是模 2 与 0 同余的所有整数集  $[0]$ , 即所有偶数构成的类. 另一个是模 2 与 1 同余的所有整数集  $[1]$ , 即所有奇数构成的类, 集合

$$\mathbb{Z}_2 = \{[0], [1]\}$$

共有两个元素, 一个是偶数类, 另一个是奇数类.

有趣的是, 根据同余的性质2.20, 我们可以定义集合  $\mathbb{Z}_m$  中元素的加法、减法和乘法运算.

**同余类的加减法** 设  $[a], [b] \in \mathbb{Z}_m$ , 则将模  $m$  的同余类  $[a \pm b]$  定义为  $[a]$  与  $[b]$  的和(差)

$$[a] \pm [b] = [a \pm b].$$

首先要验证这样定义的合理性, 即不管选取什么样的表元, 两个类之间的加减是唯一确定的: 假如在  $[a]$  和  $[b]$  中分别另取两个表元  $a' \in [a], b' \in [b]$ , 即  $a \equiv a' \pmod{m}, b \equiv b' \pmod{m}$ , 从而由同余性质2.20 中的(4)得

$$a \pm b \equiv a' \pm b' \pmod{m}$$

也就是  $a' \pm b'$  与  $a \pm b$  属于同一类, 所以  $[a' \pm b'] = [a \pm b]$ .

**同余类的乘法** 同理, 可以定义  $[a]$  和  $[b]$  的乘积如下:

$$[a][b] = [ab],$$

根据性质2.20 中的(5), 这样的定义的类之间的乘积与表元的选取无关.

**同余类中的零元和单位元** 分别称同余类  $[0]$  和  $[1]$  为  $\mathbb{Z}_m$  中的零元和单位元, 它们分别满足:

$$[a] + [0] = [0] + [a] = [a], [1][a] = [a][1] = [a].$$

容易验证上述定义的同余类之间的加法和乘法满足交换律、结合律以及乘法对加法的分配律.

因此, 虽然  $\mathbb{Z}_m$  只有有限个元素, 但是  $\mathbb{Z}_m$  与整数集合  $\mathbb{Z}$  有同样的运算规律, 所以也是加法群和交换环, 称为模  $m$  的同余类环.

下面以  $\mathbb{Z}_5 = \{[0], [1], [2], [3], [4]\}$  为例, 分别给出元素之间的加法和乘法.

加法	[0]	[1]	[2]	[3]	[4]
[0]	[0]	[1]	[2]	[3]	[4]
[1]	[1]	[2]	[3]	[4]	[0]
[2]	[2]	[3]	[4]	[0]	[1]
[3]	[3]	[4]	[0]	[1]	[2]
[4]	[4]	[0]	[1]	[2]	[3]

乘法	[0]	[1]	[2]	[3]	[4]
[0]	[0]	[0]	0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]
[2]	[0]	[2]	[4]	[1]	[3]
[3]	[0]	[3]	[1]	[4]	[2]
[4]	[0]	[4]	[3]	[2]	[1]

几何上看, 整数  $\mathbb{Z}$  对应实数轴上所有整数点, 也就是以长度为 1 的等分点 (图2.1).

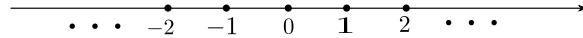


图 2.1

而  $\mathbb{Z}_m$  对应的是圆上  $m$  个等分点, 例如钟表上的 12 个刻度, 就是  $\mathbb{Z}_{12}$  中元素对应的点. 每过 12个小时, 指针指向同一个位置. 例如 3 小时, 15 小时, 27 小时等等, 都指向 [3] 的位置, 因为  $3, 15, 27, \dots$ , 模 12 同余, 属于模12 的同余类 [3] (图2.2).

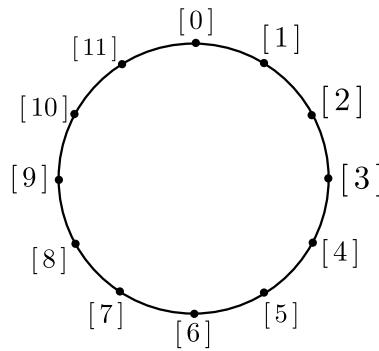


图 2.2

关于  $\mathbb{Z}_m$  中元素在乘法运算下是否可逆, 需要进一步讨论.

回顾一下在有理数  $\mathbb{Q}$  中, 一个非零有理数  $\alpha$  的所谓逆元, 是指方程  $\alpha x = 1$  在  $\mathbb{Q}$  中有唯一解. 这个解  $x$  就称为有理数  $\alpha$  的逆, 记为  $x = \alpha^{-1} \in \mathbb{Q}$ .

但是在  $\mathbb{Z}_m$  中, 并非对所有的非零元素  $[a]$ , 方程  $[a][x] = [ax] = [1]$  在  $\mathbb{Z}_m$  中都有解, 也就是  $[a]$  在  $\mathbb{Z}_m$  中不一定都有逆. 例如对  $\mathbb{Z}_6$  中的  $[2]$ , 就不存在  $[x]$ , 使得  $[2][x] = [2x] = 1$  有解, 因为不会存在整数  $x$  使得  $2x \equiv 1 \pmod{6}$ .

另外, 对  $\mathbb{Z}_6$  中两个非零元素  $[2]$  和  $[3]$ , 有  $[2][3] = [0]$ , 这也说明  $\mathbb{Z}_6$  中的  $[2]$  和  $[3]$  不可能有逆.

然而, 如果考虑  $\mathbb{Z}_m$  的子集合

$$\mathbb{Z}_m^* = \{[a] \mid [a] \in \mathbb{Z}_m, (a, m) = 1\},$$

则  $\mathbb{Z}_m^*$  中的元素却有逆, 也就是说, 只有当  $(a, m) = 1$  时,  $[a]$  可逆且  $[a]^{-1} \in \mathbb{Z}_m^*$ .

**定理 2.23** 集合  $\mathbb{Z}_m^*$  中元素满足

$$(1) [1] \in \mathbb{Z}_m^*.$$

$$(2) \text{若 } [a], [b] \in \mathbb{Z}_m^*, \text{ 则 } [a][b] \in \mathbb{Z}_m^*.$$

$$(3) \text{若 } [a] \in \mathbb{Z}_m^*, \text{ 则方程 } [a][x] = [1] \text{ 在 } \mathbb{Z}_m^* \text{ 中有唯一解, 记为 } [x] = [a]^{-1}.$$

任何满足上述三条的集合称为乘法群, 因此  $\mathbb{Z}_m^*$  是(可交换的)乘法群, 且元素的个数为  $1, 2, \dots, m$  中与  $m$  互素的整数个数, 即  $\varphi(m)$  (定义 2.15 中的 Euler 函数).

**证明** (1) 是显然的, 因为  $(1, m) = 1$ .

对于 (2), 只要用到若  $(a, m) = 1, (b, m) = 1$  就有  $(ab, m) = 1$  这个简单事实即可.

对于 (3) 的证明如下:

对任意的  $[a] \in \mathbb{Z}_m^*$ , 因为  $(a, m) = 1$ , 所以

$$0a - 1, a - 1, 2a - 1, \dots, (m - 1)a - 1$$

这  $m$  个数两两互不同余, 否则若  $j > i$ , 使得  $ia - 1$  与  $ja - 1$  同余, 就导致  $m|(j - i)a$ , 但  $(a, m) = 1$ , 而  $0 < j - i < m$ , 这显然不可能.

因此, 这  $m$  个数所在的同余类共  $m$  个, 与  $\mathbb{Z}_m$  一致:

$$\{[0a - 1], [a - 1], [2a - 1], \dots, [(m - 1)a - 1]\} = \mathbb{Z}_m,$$

这样就意味着, 必存在唯一的  $k$ , 使得  $[ka - 1] = [0]$ , 即存在唯一的  $k$ , 使得  $ka \equiv 1 \pmod{m}$ , 所以  $[x] = [k]$  是方程  $[a][x] = [1]$  的唯一解.

因  $m|ka - 1$ ,  $(a, m) = 1$ , 所以  $(k, m) = 1$ , 这就推得解  $[x] = [k] \in \mathbb{Z}_m^*$ . □

**例 2.5.1** 当  $m = 5$ , 有

$$\mathbb{Z}_5^* = \{[1], [2], [3], [4]\},$$

元素的逆分别为

$$[1]^{-1} = [1], [2]^{-1} = [3], [3]^{-1} = [2], [4]^{-1} = [4].$$

当  $m = 12$ , 有

$$\mathbb{Z}_{12}^* = \{[1], [5], [7], [11]\},$$

元素的逆分别为

$$[1]^{-1} = [1], [5]^{-1} = [5], [7]^{-1} = [7], [11]^{-1} = [11].$$

当  $m = 14$ , 有

$$\mathbb{Z}_{14}^* = \{[1], [3], [5], [9], [11], [13]\},$$

元素的逆分别为

$$[1]^{-1} = [1], [3]^{-1} = [5], [5]^{-1} = [3], [9]^{-1} = [11], [11]^{-1} = [9], [13]^{-1} = [13].$$

因此要使得  $\mathbb{Z}_m$  中除  $[0]$  外每个元素都可逆, 只有当  $m = p$  是素数. 因为此时除 0 以外,  $1, 2, \dots, p - 1$  都与  $p$  互素, 也就有  $\mathbb{Z}_p = \{[0]\} \cup \mathbb{Z}_p^*$ .

**定理 2.24** 设  $p$  是任意素数, 则  $\mathbb{Z}_p$  中元素满足

(1) 对任意的  $[a], [b] \in \mathbb{Z}_p$ , 有

$$[a] \pm [b] = [a \pm b] \in \mathbb{Z}_p, [a][b] \in \mathbb{Z}_p;$$

(2)  $\mathbb{Z}_p$  中有 0 元  $[0]$  和单位元  $[1]$ , 即对任意的  $[a] \in \mathbb{Z}_p$ , 有

$$[a] + [0] = [a], [a][1] = [a];$$

(3) 对  $\mathbb{Z}_p$  中任何非零的  $[a]$ , 存在逆元  $[x] = [a]^{-1} \in \mathbb{Z}_p$ , 使得

$$[a][a]^{-1} = [a]^{-1}[a] = [1].$$

任何满足上述三条的集合称为域, 因此对素数  $p$ ,  $\mathbb{Z}_p$  是一个域. 因为  $\mathbb{Z}_p$  由  $p$  个同余类构成, 因此称为有限域.

总结一下, 通过对同余类的讨论, 我们接触到这样几个数学概念:

满足加法、乘法以及交换律、结合律、分配律并有 0 元和单位元的集合称为(交换)环.

满足乘法和除法(乘法的逆运算)以及交换律、结合律并有单位元的集合称为(交换的)乘法群.

既是交换环又是乘法群(除法除 0 元素外)的集合称为域.

对我们熟悉的集合(数集)来说, 整数集合  $\mathbb{Z}$  是环, 但不是域, 有理数集合  $\mathbb{Q}$  和实数集合  $\mathbb{R}$  是域.

无论是  $\mathbb{Z}$ , 还是  $\mathbb{Q}$  或  $\mathbb{R}$ , 其中的元素个数都是无限的. 这里, 通过整数的同余类, 给出了元素个数只有有限个  $p$  ( $p$  是素数) 的有限域的一个经典的例子.

最后, 我们给出一个重要的定理及其证明.

**定理 2.25** (Fermat 定理) 设  $a$  是一个整数, 如果  $p$  是任意一个不能整除  $a$  的素数 (即  $p \nmid a$ ), 那么

$$a^{p-1} \equiv 1 \pmod{p} \text{ 或 } a^p \equiv a \pmod{p}.$$

**证明** 考虑  $a$  的倍数:

$$m_1 = a, m_2 = 2a, \dots, m_{p-1} = (p-1)a.$$

这些数中任何两个都不会模  $p$  同余, 这是因为对任意的  $m_r, m_s$  ( $1 \leq r < s \leq p-1$ ),  $m_s - m_r = (s-r)a$ , 无论是  $s-r$  还是  $a$  都不能被  $p$  整除. 同样, 任何  $m_r = ra$  都不会模  $p$  同余 0, 因为它不能被  $p$  整除.

以上分析说明  $m_1, m_2, \dots, m_{p-1}$  这  $p-1$  个数被  $p$  除的余数两两不等且不等于 0, 所以所有余数只能是  $1, 2, \dots, p-1$ . 虽然有可能不会按次序对应, 但根据同余性质 2.20 中的 (5), 有

$$m_1 m_2 \cdots m_{p-1} \equiv 1 \cdot 2 \cdots (p-1) \pmod{p},$$

注意到

$$m_1 m_2 \cdots m_{p-1} = 1 \cdot 2 \cdots (p-1) a^{p-1} = (p-1)! a^{p-1},$$

由上述两式得

$$(p-1)! (a^{p-1} - 1) \equiv 0 \pmod{p},$$

但是  $p$  不能整除  $(p-1)!$ , 这样由性质 2.20 中的 (6) 得,  $p$  只能整除  $a^{p-1} - 1$ , 也就是

$$a^{p-1} - 1 \equiv 0 \pmod{p}.$$

□

## §2.6 同余方程(组)

给定的正整数  $m$ , 以及一个  $n$  次整系数多项式

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

下列方程

$$f(x) \equiv 0 \pmod{m}$$

称为模  $m$  的  $n$  次同余方程。显然, 如果  $x = c$  是方程的一个解, 那么同余类  $[c]$  中的任意数(即形如  $c + km$  的数, 其中  $k$  是任意整数)都是方程的解。我们将属于同一同余类中的解视为相同的, 只有那些模  $m$  不同余的解才视为不同的解。同余方程组就是对给定的  $r$  个正整数  $m_1, \dots, m_r$  以及  $r$  个整系数多项式  $f_1(x), \dots, f_r(x)$ , 一组模  $m_i$  的同余方程

$$f_i(x) \equiv 0 \pmod{m_i}, \quad i = 1, 2, \dots, r.$$

的联立方程组。以下我们主要讨论一次同余方程(组)的求解问题。

**定理 2.26** 给定  $m$  和  $a, b$ 。考虑一次同余方程

$$ax \equiv b \pmod{m}.$$

(1) 若  $(a, m) = 1$ , 则方程存在唯一解。

(2) 若  $(a, m) = d$ , 则方程的解存在的充分必要条件是  $d|b$ 。当此条件成立时, 共有  $d$  个解。

**证明** (1) 由 Bezout 定理 2.8 知, 存在整数  $k, l$  使得

$$ka + lm = 1,$$

两边同时乘以  $b$  得  $kba + lbm = b$ , 推得  $m|(kba - b)$ , 所以  $x = kb$  是方程的解。

如果方程另有一解  $x'$ , 则  $m|a(x - x')$ , 而  $(a, m) = 1$ , 推得  $m|(x - x')$ , 即  $x \equiv x' \pmod{m}$ 。因此解是唯一的。

(2) 设方程有解  $x$ , 所以  $m|(ax - b)$ , 从而  $d|(ax - b)$ 。但  $d|a$ , 所以  $d|b$ 。反之如果  $d|b$ , 同(1) 由 Bezout 定理 2.8 知, 存在整数  $k, l$  使得

$$ka + lm = (a, m) = d,$$

推得

$$kba + lbm = db, \text{ 或 } k\frac{b}{d}a + l\frac{b}{d}m = b,$$

因此  $k\frac{b}{d}a \equiv b \pmod{m}$ , 即  $x = k\frac{b}{d}$  是解。

设  $x_0$  是方程的一个解, 即  $m|(ax_0 - b)$ , 那么

$$x_k = x_0 + k\frac{m}{d}, \quad k = 0, 1, \dots, d-1$$

也是方程的解。这是因为  $ax_k = ax_0 + k\frac{am}{d}$ , 其中  $\frac{a}{d}$  是整数, 而

$$ax_k - b = ax_0 - b + k\frac{am}{d},$$

所以  $m|(ax_k - b)$ 。另一方面,

$$x_k - x_{k'} = (k - k')\frac{m}{d},$$

所以  $m \nmid (x_k - x_{k'})$ . 这样我们就证明了方程有  $d$  个解.  $\square$

早在公元400年前后, 我国有一本数学著作:《孙子算经》, 其中记载了这样一个问题:

**定理 2.27 (孙子定理)** 今有物不知其数, 三三数之余二, 五五数之余三, 七七数之余二, 问该物几何?

该问题实际上是要解三个一次同余方程联立的方程组的问题. 设该物共有  $N$  个, 因此  $N$  满足

$$N \equiv 2 \pmod{3}, N \equiv 3 \pmod{5}, N \equiv 2 \pmod{7}.$$

我们把上述具体问题稍加推广, 就是著名的**中国剩余定理**:

**定理 2.28 (中国剩余定理)** 设  $m_1, m_2, m_3$  是两两互素的正整数, 则对任意的整数  $b_1, b_2, b_3$ , 下列关于  $x$  的同余方程组必有解.

$$x \equiv b_1 \pmod{m_1}, x \equiv b_2 \pmod{m_2}, x \equiv b_3 \pmod{m_3}. \quad (1)$$

方程 (1) 如果有解  $x$ , 那么任何  $x + nm_1m_2m_3$  也是解, 所以一旦有解, 解不唯一.

**证明** 把同余方程组 (1) 分解为三组同余方程组

$$x_1 \equiv 1 \pmod{m_1}, x_1 \equiv 0 \pmod{m_2}, x_1 \equiv 0 \pmod{m_3}; \quad (2)$$

$$x_2 \equiv 0 \pmod{m_1}, x_2 \equiv 1 \pmod{m_2}, x_2 \equiv 0 \pmod{m_3}; \quad (3)$$

$$x_3 \equiv 0 \pmod{m_1}, x_3 \equiv 0 \pmod{m_2}, x_3 \equiv 1 \pmod{m_3}; \quad (4)$$

如果这三组同余方程组 (2)、(3)、(4) 分别有解  $x_1, x_2, x_3$ , 那么 (1) 的解就为

$$x = b_1x_1 + b_2x_2 + b_3x_3.$$

关于方程组 (2), 由条件可知  $(m_2m_3, m_1) = 1$ , 根据 Bezout 公式, 存在整数  $k_1, l_1$ , 使得整系数线性组合满足

$$k_1(m_2m_3) + l_1m_1 = 1$$

也就是  $k_1(m_2m_3) \equiv 1 \pmod{m_1}$ , 取  $x_1 = k_1m_2m_3$ , 则  $x$  满足 (2) 中第一个方程, 而 (2) 第二个、第三个方程自然满足. 所以  $x_1 = k_1m_2m_3$  是 (2) 的一个解.

同理, 我们可分别得到 (3) 和 (4) 的解.

$$x_2 = k_2(m_1m_3), x_3 = k_3(m_1m_2),$$

其中如同  $k_1$  一样,  $k_2$  和  $k_3$  是满足 Bezout 公式中的整系数

$$k_2(m_1m_3) + l_2m_2 = 1, k_3(m_1m_3) + l_3m_3 = 1.$$

$\square$

从上面的证明可以看出, 解同余方程组(1), 最终转化为利用 Euclid 辗转相除法求出相关的 Bezout 公式(见定理2.8), 也就是求出  $k_1, k_2, k_3$ .

现在回到孙子定理, 具体来说

$$m_1 = 3, m_2 = 5, m_3 = 7,$$

$$b_1 = 2, b_2 = 3, b_3 = 2.$$

因此欲求  $x_1$ , 只需求整数  $k_1, l_1$  使得  $35k_1 + 3l_1 = 1$ , 根据辗转相除法  $k_1 = 35, l_1 = -23$ . 所以  $x_1 = 70$ .

同理欲求  $x_2$ , 只需求整数  $k_2, l_2$  使得  $21k_2 + 5l_2 = 1$ , 解得  $k_2 = 1, l_2 = -20$ , 所以  $x_2 = 21$ .

欲求  $x_3$ , 只需求整数  $k_3, l_3$  使得  $15k_3 + 7l_3 = 1$ , 解得  $k_3 = 1, l_3 = -2$ , 所以  $x_3 = 15$ .

最终孙子定理的解为

$$x = 2 \cdot 70 + 3 \cdot 21 + 2 \cdot 15 = 233.$$

显然这个问题的解并不是唯一的, 因此减去  $3 \cdot 5 \cdot 7 = 105$  的倍数仍是解. 这样我们最终得到最小正整数解为  $x = 23$ .

有趣的是约在 1593 年, 我国古代数学家程大位(1533-1606)就给出了非常精确的答案: 他在《算法统宗》一书中把这一问题的解法总结成了如下歌诀:

**三人同行七十稀, 五树梅花廿一枝, 七子团圆半个月, 除百零五便得知。**

歌诀中前三句话出现的“三人”、“五树”和“七子”指的是三个模  $3, 5, 7$ . 而“七十稀”、“廿一枝”和“半个月”指的是求出的三个解  $x_1 = 70, x_2 = 21, x_3 = 15$ . 最后一句话中“百零五”指的是三个模的乘积  $3 \cdot 5 \cdot 7 = 105$ . 并说明了如何得到最小正整数解. 可见我国古代科学家对辗转相除法已经是十分熟悉的.

**注记:**

一, 类似孙子定理的问题在我国古代还有不少, 比较有名的如“韩信点兵”和“百鸡问题”等等, 最终都是求解同余方程组的问题.

二, 中国剩余定理可以推广到包含  $n$  个同余方程的方程组. 设  $m_1, m_2, \dots, m_r$  是两两互素的正整数, 则对任意的整数  $b_1, b_2, \dots, b_r$ , 下列同余方程组

$$x \equiv b_1 \pmod{m_1}, x \equiv b_2 \pmod{m_2}, \dots, x \equiv b_r \pmod{m_r}.$$

必有解.

## §2.7 多项式

在本专题的最后, 简要介绍有关多项式的基本内容, 这是因为多项式与整数有很多相似之处, 同时在后续各讲中, 将会不断用到多项式的这些基本内容.

所谓多项式是指

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \quad (a_n \neq 0),$$

这里,  $n$  是一个正整数, 称为多项式的次数, 记为  $\deg f = n$ . 次数为 0 的多项式称为平凡多项式, 次数大于 0 的多项式称为非平凡多项式. 对于首项(最高次数项)系数为 1 的多项式, 称为首一多项式.

多项式的系数  $a_n, a_{n-1}, \dots, a_0$  可以是整数, 也可以属于有理数域  $\mathbb{Q}$ 、实数域  $\mathbb{R}$ 、复数域  $\mathbb{C}$  或其他数域.

当然, 如果多项式系数是有理数, 那么利用通分, 可以使多项式等价于整系数的多项式, 例如

$$\frac{1}{3}x^2 + \frac{2}{5}x + 1 = 0,$$

通分后就化为

$$5x^2 + 6x + 15 = 0.$$

因此比较多的情况下, 是讨论整系数的多项式而不是有理系数的多项式.

称系数属于整数  $\mathbb{Z}$ , 或某个域  $\mathbb{F}$  的多项式为“ $\mathbb{Z}$  上多项式”, 或“域  $\mathbb{F}$  上多项式”, 记整数  $\mathbb{Z}$ , 或域  $\mathbb{F}$  上多项式的全体为  $\mathbb{Z}[x]$ , 或  $\mathbb{F}[x]$ .

有时把多项式按照幂次从低到高排列

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

这与次数从高到低排列没有本质区别.

### 1° 多项式的加法与乘法

同一个域上两个多项式

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \quad (a_n \neq 0),$$

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_0 \quad (b_m \neq 0)$$

之间可以进行加法和乘法运算, 其结果还是这个域上的多项式, 并且有

$$\deg(f + g) \leq \deg f + \deg g, \quad \deg(fg) = \deg f + \deg g.$$

对于加法和乘法来说, 即使多项式的系数属于整数集合  $\mathbb{Z}$ , 经过加法和乘法后, 所得到的多项式的系数还是属于  $\mathbb{Z}$ .

## 2° 多项式的整除和带余除法

设  $\mathbb{F}$  是给定的域,  $f(x), g(x) \in \mathbb{F}[x]$ ,  $g(x) \neq 0$ , 如果存在  $h(x) \in \mathbb{F}[x]$ , 使得

$$f(x) = g(x)h(x),$$

那么称  $g(x)$  在  $\mathbb{F}[x]$  中整除  $f(x)$ , 记为  $g(x)|f(x)$ , 并称  $g(x)$  是  $f(x)$  的一个因式,  $f(x)$  是  $g(x)$  的一个倍式. 对于两个多项式  $f(x), g(x)$ , 如果  $h(x)|f(x), h(x)|g(x)$ , 那么称  $h(x)$  是  $f(x), g(x)$  的公因式.

对于不能整除的情形, 我们有下列多项式的带余除法.

**定理 2.29** 设  $\mathbb{F}$  是给定的域,  $f(x), g(x) \in \mathbb{F}[x], g(x) \neq 0$ , 那么在  $\mathbb{F}[x]$  存在  $\mathbb{F}$  上唯一的一组多项式  $q(x), r(x)$ , 使得

$$f(x) = q(x)g(x) + r(x),$$

其中  $r(x)$  或者为 0, 或者  $\deg r < \deg g$ .

**证明** 若果  $\deg f < \deg g$ , 那么取  $q(x) = 0, r(x) = f(x)$ , 如果  $g(x) = c \neq 0$  (常数), 那么取  $q(x) = \frac{1}{c}f(x), r(x) = 0$ , 所以我们设

$$n = \deg f \geq m = \deg g,$$

并记  $f(x)$  和  $g(x)$  的首项系数分别为  $a_n, b_m$ . 令  $q_1(x) = \frac{a_n}{b_m}x^{n-m}$ , 那么

$$f_1(x) = f(x) - q_1(x)g(x)$$

的次数满足  $\deg f_1 < \deg f$ . 如果  $\deg f_1 > \deg g$ , 那么对  $f_1(x)$  重复上述过程, 这样我们得到一列多项式

$$f_0(x) = f(x), f_1(x), f_2(x), \dots, \text{ 和 } q_1(x), q_2(x), \dots,$$

使得

$$f_{i+1}(x) = f_i(x) - q_{i+1}(x)g(x), i = 0, 1, 2, \dots.$$

而且

$$\deg f > \deg f_1 > \deg f_2 > \dots \geq \deg g,$$

这样的过程有限步后必然会停止, 即存在  $l$ , 使得  $\deg f_l < \deg g$ . 于是  $q(x) = q_1(x) + \dots + q_l(x)$  及  $r(x) = f_l(x)$  便给出定理中的表示.

假如有两组表示

$$f(x) = q_1(x)g(x) + r_1(x), f(x) = q_2(x)g(x) + r_2(x),$$

那么推出  $(q_1(x) - q_2(x))g(x) = r_1(x) - r_2(x)$ , 比较两边次数就可得到  $q_1(x) = q_2(x), r_1(x) = r_2(x)$ , 因此唯一性得证.  $\square$

由于整数不可以做除法, 所以上述定理对系数是整数(整系数)的多项式  $f(x), g(x) \in \mathbb{Z}[x], g(x) \neq 0$  不适用, 但是如果限制  $g(x)$  的首项系数为 1, 那么上述证明过程和结论都适用于整系数多项式.

通过带余除法, 可以证明下列定理

**定理 2.30** 设  $J$  是  $\mathbb{F}[x]$  的一个非空子集合, 如果  $J$  满足下列条件:

- (1) 对任意的  $f(x), g(x) \in J$ , 有  $f(x) - g(x) \in J$ ;
- (2) 对任意的  $f(x) \in J$  和任意的  $h(x) \in \mathbb{F}[x]$ , 有  $f(x)h(x) \in J$ .

那么或者  $J = \{0\}$ , 或者存在唯一的首项系数为 1 的多项式  $d(x) \in \mathbb{F}[x]$ , 使得  $J = \langle d(x) \rangle$ . 这里

$$\langle d(x) \rangle = \{h(x)d(x) \mid h(x) \in \mathbb{F}[x]\}.$$

这样的  $d(x)$  也属于  $J$ . 也就是说, 满足上述两个条件的集合  $J$  中的多项式要么是 0, 要么是其中一个元素  $d(x)$  的倍式.

**证明** 如果  $J \neq \{0\}$ , 则  $J$  中存在非零多项式, 设  $d(x)$  是  $J$  中次数最低的一个(不唯一), 但由条件中的(2)可设  $d(x)$  的首项系数为 1.

这样, 对任意的  $f(x) \in J$ , 根据带余除法存在  $q(x), r(x) \in \mathbb{F}[x]$ , 使得

$$f(x) = q(x)d(x) + r(x),$$

其中  $r(x) = 0$  或者  $\deg r < \deg d$ . 若  $r(x) \neq 0$ , 则根据  $J$  满足条件(2), 有  $q(x)d(x) \in J$ . 再由(1)可知

$$r(x) = f(x) - q(x)d(x) \in J,$$

但是  $\deg r < \deg d$ , 这与  $d(x)$  的选取矛盾. 所以  $r(x) = 0$ , 即对任意的  $f(x) \in J$ , 存在  $q(x) \in \mathbb{F}[x]$ , 使得  $f(x) = q(x)d(x) \in J$ .

如果存另一个首项系数为 1 的多项式  $d_1(x)$  使得  $J = \{h(x)d_1(x) \mid h(x) \in \mathbb{F}[x]\}$ , 那么  $d_1(x)|d(x)$ ,  $d(x)|d_1(x)$ , 从而必有  $d_1(x) = d(x)$ , 所以唯一性得证.  $\square$

**定理 2.31** 设  $f(x), g(x)$  是  $\mathbb{F}[x]$  中两个不全为零的多项式, 则集合

$$J = \{\alpha(x)f(x) + \beta(x)g(x) \mid \alpha(x), \beta(x) \in \mathbb{F}[x]\}$$

满足定理 2.30 中条件, 因此存在  $d(x) \in J$ , 使得  $J = \langle d(x) \rangle$ .

显然,  $f(x) = f(x) + 0g(x) \in \langle d(x) \rangle$ , 因此  $f(x)$  是  $d(x)$  的倍式, 推得  $d(x)|f(x)$ , 同理可得  $d(x)|g(x)$ , 所以  $d(x)$  是  $f(x)$  和  $g(x)$  的公因式. 假设  $f(x)$  和  $g(x)$  有另外一个公因式  $d_1(x)$ , 那么由于  $d(x) \in J$ , 所以存在  $\alpha(x), \beta(x) \in \mathbb{F}[x]$ , 使得  $d(x) = \alpha(x)f(x) + \beta(x)g(x)$ , 因此  $d_1(x)|d(x)$ , 从而  $\deg d_1(x) \leq \deg d(x)$ , 也就是说  $d(x)$  是  $f(x)$

和  $g(x)$  的次数最高且首项系数为 1 的公因式. 这样的公因式称为最大公因式, 记为

$$d(x) = (f(x), g(x)).$$

当  $(f(x), g(x)) = 1$  时, 称  $f(x)$  和  $g(x)$  互素.

**推论 2.32** (多项式的Bezout 公式) 设  $f(x), g(x) \in \mathbb{F}[x]$  且不全为零, 则存在  $\alpha(x), \beta(x) \in \mathbb{F}[x]$ , 使得

$$\alpha(x)f(x) + \beta(x)g(x) = (f(x), g(x)),$$

特别当  $f(x), g(x)$  互素时, 存在  $\alpha(x), \beta(x) \in \mathbb{F}[x]$ , 使得

$$\alpha(x)f(x) + \beta(x)g(x) = 1.$$

上述结果和概念, 不难推广到  $\mathbb{F}[x]$  中有限个不全为零的多项式情形. 对于如何具体求出两个多项式  $f(x)$  和  $g(x)$  的最大公因式以及 Bezout 公式中的  $\alpha(x)$  和  $\beta(x)$ , 我们可以仿照整数求最大公约数的做法, 通过建立多项式的 Euclid 辗转相除法来解决, 这里就不在讨论了.

### 3° 多项式的可约性

一个系数属于整数  $\mathbb{Z}$ , 或数域  $\mathbb{F}$  的多项式  $p(x)$  如果能够分解成两个系数同样属于数域  $\mathbb{F}$  非平凡的多项式  $p_1(x), p_2(x)$  的乘积:

$$p(x) = p_1(x)p_2(x),$$

那么就称  $p(x)$  是整数  $\mathbb{Z}$ , 或数域  $\mathbb{F}$  上可约多项式, 否则称为整数  $\mathbb{Z}$ , 或数域  $\mathbb{F}$  上不可约多项式.

**例 2.7.1** 注意到, 可约或不可约与多项式系数所在的范围密切相关. 这里给出最简单的几种情况

$x^2 - 4$  在  $\mathbb{Z}$  上是可约的:  $x^2 - 4 = (x - 2)(x + 2)$ .

$9x^2 - 4$  在  $\mathbb{Z}$  上不可约, 但是在  $\mathbb{Q}$  上可约:  $9x^2 - 4 = \left(x - \frac{2}{3}\right)\left(x + \frac{2}{3}\right)$ .

$x^2 - 2$  在  $\mathbb{Q}$  上不可约, 但在  $\mathbb{R}$  上可约:  $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ .

$x^2 + 1$  在  $\mathbb{R}$  上不可约, 但在  $\mathbb{C}$  上可约:  $x^2 + 1 = (x - i)(x + i)$

虽然上述例中的多项式都是整系数多项式. 但随着系数所在范围的扩大

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C},$$

可约性也在发生变化.

### 4° 因式分解

如同整数可以分解为素数的乘积一样, 我们希望任何一个数域  $\mathbb{F}$  上的多项式可以分解为  $\mathbb{F}$  上不可约多项式的乘积. 为此有

**定理 2.33** 设  $\mathbb{F}$  是一个数域, 则

(1)  $\mathbb{F}[x]$  中有无穷多个首项系数为 1 的不可约多项式.

(2)  $\mathbb{F}[x]$  中任何多项式  $f(x)$  可唯一分解为不可约多项式的乘积:

$$f(x) = ap_1^{m_1}(x) \cdots p_r^{m_r}(x),$$

这里  $p_1(x), \dots, p_r(x)$  是  $\mathbb{F}[x]$  中首项系数为 1 的不可约多项式,  $m_1, \dots, m_r$  表示重数.

关于结论(1), 对包含无穷多个数的数域  $\mathbb{F}$  情形, 如  $\mathbb{Q}, \mathbb{R}, \mathbb{C}$  等, 是非常简单地, 因为对任意的  $a \in \mathbb{F}$ , 一次式  $x - a \in \mathbb{F}[x]$  都是不可约的, 因此有无穷多个. 其它情形在此就不讨论了.

关于结论(2), 可仿照整数的素因子分解唯一性的证明, 从多项式的 Bezout 公式, 导出多项式最大公因式的基本性质即可证得.

## 第 2 讲习题

1. 设  $n > 1$  为整数, 如果对于任何整数  $m$ , 或者  $n|m$ , 或者  $(n, m) = 1$ , 那么  $n$  必定是素数.
2. 设  $n$  为正整数, 且  $n \geq 2$ . 如果对任何不超过  $\sqrt{n}$  的素数都不能整除  $n$ , 那么  $n$  一定是素数.
3. 设正整数  $n > 2$ , 证明: 在  $n$  和  $n!$  之间一定存在素数.
4. 考虑正整数的素因子标准分解

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$$

这里  $p_1, \dots, p_k$  是两两不同的素数,  $\alpha_1, \dots, \alpha_k$  是对应素因子的重数.

证明:  $n$  的所有因子 (即能整除  $n$  的数, 包括  $n$  和 1) 是这样的数

$$m = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$$

其中  $0 \leq \beta_1 \leq \alpha_1$ ,  $0 \leq \beta_2 \leq \alpha_2, \dots, 0 \leq \beta_k \leq \alpha_k$ . 而且  $n$  所有不同因子 (包括  $n$  和 1) 的个数为

$$(\alpha_1 + 1)(\alpha_2 + 1) \cdots (\alpha_k + 1).$$

例如

$$144 = 2^4 3^2,$$

的所有因子共  $5 \cdot 3$  个, 具体为

$$1, 2, 4, 8, 16, 3, 6, 12, 24, 48, 9, 18, 36, 72, 144$$

5. 如果把自然数看成是一个等差的 (无穷) 数列, 那么定理2.11表明, 该数列中有无穷多个素数.

试证明等差数列

$$4n + 3, \quad n = 1, 2, 3, \dots,$$

中也有无穷多个素数.

提示: 任何大于 2 的素数一定是奇数, 因而一定是  $4n + 1$  或  $4n + 3$  这种形式. 注意到两个  $4n + 1$  形式的数相乘还是  $4n + 1$  形式的数. 现在假设只有有限个  $4n + 3$  形式的素数  $p_1, p_2, \dots, p_r$ , 考虑

$$N = 4(p_1 p_2 \cdots p_r) - 1 = 4(p_1 p_2 \cdots p_r - 1) + 3.$$

类似定理2.11的证明, 以及利用两个  $4n + 1$  形式的数相乘还是  $4n + 1$  形式的数的结论推出假设是错误的.

6. 对互素的两个数 15, 7, 求出整数  $k, l$  使得  $15k + 7l = 1$ .

7. 求解同余方程组

$$x \equiv 1 \pmod{4}, \quad x \equiv 2 \pmod{5}, \quad x \equiv 4 \pmod{7}.$$

8. 求解同余方程组

$$5x \equiv 7 \pmod{12}, \quad 7x \equiv 1 \pmod{10}.$$

提示: 对于同余方程

$$ax \equiv b \pmod{m},$$

当  $(a, m) = 1$  时, 存在整数  $l, k$ , 使得  $al + mk = 1$ . 因此

$$(ax - b)l + mkx = x - bl,$$

这样, 原方程可化为等价的形式

$$x \equiv bl \pmod{m}.$$

另一方面, 如果  $m = m_1m_2$ , 那么  $(a, m) = 1$  等价于  $(a, m_1) = 1, (a, m_2) = 1$ , 且  $m|(ax - b)$  等价于  $m_1|(ax - b), m_2|(ax - b)$ , 因此原方程等价于同余方程组

$$ax \equiv b \pmod{m_1}, \quad ax \equiv b \pmod{m_2}.$$

结合上述分析, 本题的同余方程组中

$$5x \equiv 7 \pmod{12} \text{ 等价于 } x \equiv 2 \pmod{3}, \quad x \equiv 3 \pmod{4},$$

$$7x \equiv 1 \pmod{10} \text{ 等价于 } x \equiv 1 \pmod{2}, \quad x \equiv 3 \pmod{5}.$$

注意到  $x \equiv 3 \pmod{4}$  和  $x \equiv 1 \pmod{2}$  等价于  $x \equiv 3 \pmod{4}$ . 所以本题同余方程组等价于下列同余方程组

$$x \equiv 2 \pmod{3}, \quad x \equiv 3 \pmod{4}, \quad x \equiv 3 \pmod{5}.$$

9. 求有理系数多项式  $\alpha(x), \beta(x)$ , 使得

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

## 第 3 讲 实数

虽然对实数的认识,甚至包括无理数,可追述到古希腊时代,但真正把实数系构建起来还是十九世纪的事情.随着微积分学的严格化,极限理论逐步建立,也使得实数理论日臻完善.主要以 Weierstrass (1815 ~ 1897), Cantor (1845 ~ 1918) 和 Dedekind (1831 ~ 1916) 的工作最具代表性.在本专题中,我们将重点介绍 Dedekind 的方法,其它方法构造出的实数系与 Dedekind 方法构造的实数系本质上是等价的.

### §3.1 有理数

自然数是从计算一个集合(如一群羊,一篮苹果)的元素个数的过程中抽象出来的.但在实际生活中,我们不仅要计数还需要进行如长度、面积、时间、重量等度量.为了把度量的问题变为计数的问题,首先要选择一个度量单位,也就是度量标准,如公尺、亩、小时、斤等等,然后看被度量的量包含多少个单位.比如一段距离正好是15个公尺,那么这个距离就是15公尺.

但是,往往出现这样的情况,计算一段距离的单位个数未必恰到好处,可能会出现15个公尺还多一点,16个公尺又不够的情况.那么只有把原有的单位分成  $n$  等分,引进新的更小的单位继续度量,例如1公尺分为100厘米,1亩分为10分,1小时分为60分钟,还有我国古代把1斤分为16两(这也就有了“半斤八两,彼此彼此”的说法)等等.继续丈量的结果就是1公尺39厘米,1小时18分钟,1亩半(5分)或古制的1斤12两等等.

如果把原单位确定为1,那么再等分就是  $\frac{1}{n}$ ,如果一个量正好等于小单位  $\frac{1}{n}$  的  $m$  倍,那么它的度量用  $\frac{m}{n}$  表示.显然  $n$  个  $\frac{1}{n}$  就是1.

这些量之间也可以做加法,仍以长度为例,把长度单位分成3分,一段距离的长度正好等于2个  $\frac{1}{3}$ ,即  $\frac{2}{3}$ .另外把长度单位分成4分,另一个距离正好为3个  $\frac{1}{4}$ ,为了把两个距离相加,我们把长度单位分成12分(3和4的倍数,相当于把  $\frac{1}{3}$  分成4分,把  $\frac{1}{4}$  分成3分),那么第一段距离就是9个  $\frac{1}{12}$ ,而第二段距离就是8个  $\frac{1}{12}$ ,两者相加就是17个  $\frac{1}{12}$ ,因为12个  $\frac{1}{12}$  等于一个单位,所以总距离为1个单位又5个  $\frac{1}{12}$ .用数学式子表示就是

$$\frac{3}{4} + \frac{2}{3} = \frac{9+8}{12} = \frac{17}{12} = 1 + \frac{5}{12}.$$

把这些量抽象出来(也就是抛开公尺、亩、小时、磅等具体度量,如同在计数时抛开是羊还是苹果一样),就产生了两个正整数之比,也就是分数  $\frac{m}{n}$ ,其中  $m, n$  是正整数( $m$  可以是0,但  $n \neq 0$ ,因为把一个单位分成0等分是没有意义的).

我们把上述从实际度量中抽象出来的正整数之比扩大到所有整数之比也就产生了由分数表示的**有理数**, 记为

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0 \right\}.$$

这里  $\mathbb{Z}$  表示所有整数集合(见第1讲)显然, 它包含了所有整数(即  $n=1$  的分数).

**有理数的算术** 定义有理数之间的加法和减法:

$$\frac{m}{n} \pm \frac{m'}{n'} = \frac{mn' \pm nm'}{nn'},$$

以及乘法:

$$\frac{m}{n} \cdot \frac{m'}{n'} = \frac{mm'}{nn'},$$

不难验证这些运算如同整数一样满足交换律、结合律以及乘法对加法的分配律. 与整数不同的是, 有理数之间可以做除法, 或者说任何一个非零的有理数  $\frac{m}{n} \in \mathbb{Q}, m \neq 0$  一定有逆元

$$\left(\frac{m}{n}\right)^{-1} = \frac{n}{m} \in \mathbb{Q}, \text{ 满足 } \frac{m}{n} \cdot \frac{n}{m} = 1,$$

这样两个有理数的除法为

$$\frac{m}{n} \div \frac{m'}{n'} = \frac{m}{n} \cdot \left(\frac{m'}{n'}\right)^{-1} = \frac{mn'}{nm'} \in \mathbb{Q}.$$

**有理数的序** 有理数之间的序(或大小)定义如下

$$\frac{m}{n} - \frac{m'}{n'} = \frac{mn' - nm'}{nn'} \begin{cases} > 0, & \text{当 } mn' - nm' \text{ 与 } nn' \text{ 同号,} \\ = 0, & \text{当 } mn' - nm' = 0, \\ < 0, & \text{当 } mn' - nm' \text{ 与 } nn' \text{ 异号.} \end{cases}$$

这样, 对任意两个有理数  $a, b \in \mathbb{Q}$ , 三种情况:  $a < b, a = b, a > b$  有且仅有一种成立, 而且若  $a < b, b < c$ , 则  $a < c$ .

有理数的序与有理数的加法和乘法是相容的, 也就是若  $a < b$ , 那么  $a + c < b + c$  对任何  $c \in \mathbb{Q}$  成立, 若  $a > 0, b > 0$ , 则  $ab > 0$ .

**有理数的几何解释** 在直线  $L$  上任取一点为原点或点  $0$ , 再将任意一点取作  $1$ (一般我们习惯上取在  $0$  的右边), 以两点之间的距离为度量的尺度或单位(这样就有了数轴), 如果从原点开始, 依单位长度往正向逐次丈量, 就得到自然数在数轴上对应的点, 同理往原点左侧的丈量得到负整数对应的点.

为描述有理数对应的点, 就要把单位进行等份, 几何上可以这么做: 设数轴上  $1$  对应的点为  $A$ , 过原点作一条异于数轴的直线  $L'$ , 对任意的自然数  $n > 1$ , 在  $L'$  上取点  $A', B'$  满足  $OA' = n, OB' = 1$ , 过点  $B'$  作线段  $\overline{AA'}$  的平行线, 交数轴于  $B$  点, 则  $OB = \frac{1}{n}$ . 利

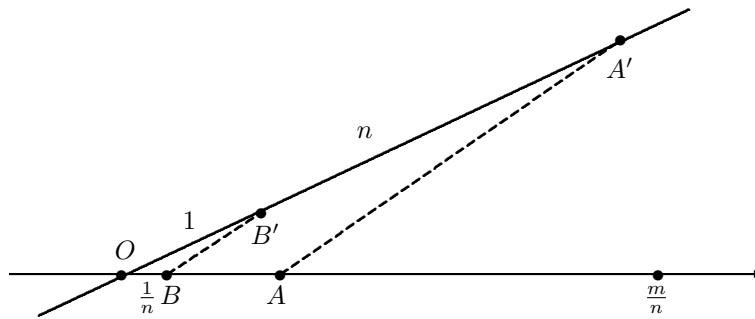


图 3.1

用  $OB$  作为新的尺度逐次丈量  $m$  次, 就得到  $\frac{m}{n}$ . 因此可以在数轴上表出所有有理数. 我们把有理数对应数轴上的点称为有理点, 并不再区分有理数还是有理点.

**有理数的稠密性** 有理数是“处处稠密”的, 即任何两个有理数  $a$  和  $b$  之间必然存在无穷多个有理数.

不妨设  $a < b$ , 显然有理数  $c = \frac{a+b}{2}$  满足  $a < c < b$ , 对  $a, c$  和  $c, b$  不断重复这个过程就会发现  $a$  和  $b$  之间有无穷多个有理数.

总之, 引进有理数以后, 在实际问题中使得度量长度、面积、时间、重量等等变得更加精细.

从数学上看, 有理数不但保持了整数所有算术性质, 对加法和乘法满足的交换律、结合律和分配律, 同时还具有乘法的逆运算 (或者说任何非零的有理数都有逆元), 满足这些规律的集合称为域, 再加上有理数所具有的序的性质, 所以称有理数集合  $\mathbb{Q}$  是一个有序域.

远在古希腊时代, 人们就已经掌握了有理数的这些性质. 以Pythagoras (毕达哥拉斯约公元前580~约前500) 为代表的学派认为, 数 (就是整数和整数之比, 也就是有理数) 不但用来计数、丈量, 甚至所有的音阶都可以用数来表示. 他们甚至把数上升到哲学层面, 提出“数是万物之源”, “数是万物的本质”, 并赋予了诸如  $1, 2, 3, \dots$  等数以“灵性”.

## §3.2 可公度与不可公度

设有两个线段  $a$  和  $b$ , 如果  $b$  的长度是  $a$  长度的正整数  $m$  倍, 那么可以用  $a$  作为  $b$  的度量, 即  $b = ma$ . 另一种情况是  $b$  不等于  $a$  的整数倍, 那么如果存在正整数  $n$ , 使得  $a$  分为  $n$  等分后,  $b$  恰好是每个等分的长度为  $\frac{a}{n}$  的  $m$  倍:

$$b = \frac{m}{n}a.$$

这样  $a$  和  $a$  就有一个公共度量线段  $\frac{a}{n}$ , 它的  $n$  倍等于  $a$ , 而它的  $m$  倍等于  $b$ , 称  $a$  和  $b$

是可公度的,或可通约的,否则称为不可公度的或不可通约的.

显然,可公度性具有传递性:若  $a$  和  $b$  可公度,  $b$  和  $c$  可公度,则  $a$  和  $c$  可公度.

如果选  $a$  为单位线段 ( $a = 1$ ), 则与之可公度的线段对应数轴上的有理点.

随后,人们就发现了与单位线段不可公度的线段,这就是边长为一个单位的正方形的对角线. 设单位正方形对角线的长度为  $x$ , 则根据勾股定理

$$x^2 = 1^2 + 1^2 = 2.$$

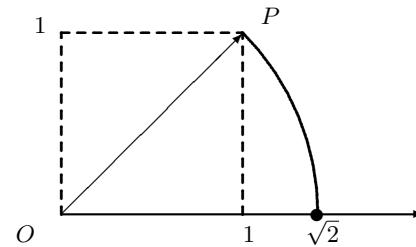


图 3.2

**定理 3.1** 单位正方形对角线线段与单位线段不可公度,也就没有有理数满足  $x^2 = 2$ .

**证明** 采用反证法. 假如存在两个正整数  $m, n$ , 使得

$$x = \frac{m}{n}, \quad (m, n) = 1.$$

因此

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

由此推出  $m^2 = 2n^2$ , 即  $m^2$  是偶数, 所以  $m$  自身也是偶数  $m = 2k$ , 代入得  $4k^2 = 2n^2$ , 又得  $n^2$  是偶数, 继而  $n$  是偶数. 这与  $(m, n) = 1$  相矛盾.  $\square$

所以单位正方形对角线长度与 1 之比不再是两个整数之比,即不再是有理数.这样使得本来看似完美无瑕、代表万物、具有灵性的有理数家族中突然闯进了一个“魔鬼”. 我们把那些与 1 不可公度(当然与所有有理数不可公度)的数称为无理数.

顺便说一句,“有理数”的名称起源于古希腊,英文的词根是“ratio”,即是“比例”的意思,与古希腊人的定义是相同的,后来出现了“rational number”这个词,中文翻译为“有理数”,把不是有理数(不能表示为整数之比)的数“irrational number”干脆翻译成了“无理数”,其实本意与“有理”还是“没理”毫无关系.

定理3.1还表明,虽然有理数是稠密的,但有理数之间是不“连续”的,也就是有理数之间存在“空隙”.例如,把有理数  $\mathbb{Q}$  分成两组

$$X = \{x \mid x \in \mathbb{Q}, x < 0; \text{ 或 } x^2 < 2, x > 0\},$$

$$Y = \{y \mid y \in \mathbb{Q}, y^2 > 2, y > 0\},$$

则,在两组有理数之间,不存在有理数.进一步我们发现

**定理 3.2** 对于上述两组有理数,  $X$  在  $\mathbb{Q}$  中无最大数,  $Y$  在  $\mathbb{Q}$  中无最小数.也就是说  $X$  中没有最大的有理数,  $Y$  中没有最小的有理数.

**证明** 设  $a \in \mathbb{Q}$ , 且  $a > 0$ , 令

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

于是

$$a'^2 - 2 = \frac{2(a^2 - 2)}{a + 2}.$$

若  $a \in X$ , 则  $a^2 - 2 < 0$ , 因此  $a' > a$ , 且  $a'^2 - 2 < 0$ , 即  $a' \in X$ . 即对任何  $X$  中的有理数  $a > 0$ , 在  $X$  中一定还存在比  $a$  大的有理数  $a' \in X$ .

若  $a \in Y$ , 则  $a^2 - 2 > 0$ , 因此  $a' < a$ , 且  $a'^2 - 2 > 0$ , 即  $a' \in Y$ , 所以任何  $Y$  中的有理数  $a$ , 必存在比  $a$  小的有理数  $a' \in Y$ .  $\square$

有理数的不连续性或者说有理数之间存在空隙这一缺憾, 正是我们需要弥补的, 而填满这些空隙的就是无理数. 只有这样才能使微积分建立在坚实的数的基础上. 同时将看到, 有理数空隙的“数量”远远多于有理数.

### §3.3 实数

要想把有理数系扩展到连续的, 没有空隙的数系, 自然的想法是要把有理数系的优点继承下来, 通过补充一些要求, 并有具体的方法, 达到扩充的目的, 正如从整数扩展到有理数一样. 为此我们先把有理数的性质, 抽象成下列一系列定义和概念.

**定义 3.3 (有序集)** 设集合  $S$ ,  $S$  上的序是一种关系, 记为  $<$ , 它满足

(1) 对任意的  $x, y \in S$   $x < y$ ,  $x = y$ ,  $y < x$  有且仅有一种成立.

(2) 对任意的  $x, y, z \in S$ , 如果  $x < y$ ,  $y < z$ , 那么  $x < z$ .

定义了“序”的集合称为有序集.

有时也用  $y > x$  代替  $x < y$ , 用  $x \leq y$  表示  $x < y$  或  $x = y$ , 也就是对  $y < x$  的否定.

显然有理数集  $\mathbb{Q}$  是有序集.

对任意一个有序集, 有

**定义 3.4 (有界)** 设  $S$  是有序集,  $E \subset S$ , 若存在  $\beta \in S$ , 使得对每个  $x \in E$ , 有  $x \leq \beta$ , 则称  $E$  有上界,  $\beta$  为  $E$  的一个上界. 同理定义下界. 同时有上下界的集合  $E$  称为有界集合.

**定义 3.5 (确界)** 设  $S$  是有序集,  $E \subset S$  有界, 若  $\alpha \in S$ , 满足

(a)  $\alpha$  是  $E$  的上界;

(b) 对任何  $\gamma < \alpha$ , 那么  $\gamma$  就不是  $E$  的上界.

则称  $\alpha$  是  $E$  的上确界, 也就是  $E$  的最小上界. 记为  $\alpha = \sup E$ . 由 (b) 不难得得到如果上确界存在, 则一定是唯一的.

同理定义  $E$  的下确界  $\alpha$ , 记为  $\alpha = \inf E$

**例 3.3.1**  $S = \mathbb{Q}$ ,  $E = \left\{1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots\right\}$ , 则  $\sup E = 1 \in S$ ,  $\inf E = 0 \in S$

但是  $1 \in E$ , 因此 1 是  $E$  中的最大数. 而  $0 \notin E$ , 所以  $E$  中没有最小数.

**例 3.3.2** 集合

$$X = \{x \mid x \in \mathbb{Q}, x < 0; \text{ 或 } x^2 < 2, x > 0\}, \quad Y = \{y \mid y \in \mathbb{Q}, y^2 > 2, y > 0\}$$

都是  $\mathbb{Q}$  的子集, 显然  $Y$  也是  $X$  在  $\mathbb{Q}$  中所有上界的集合,  $X$  是  $Y$  在  $\mathbb{Q}$  中所有下界的集合. 因为  $Y$  中没有最小数, 所以  $X$  在  $\mathbb{Q}$  中没有最小的上界, 即在  $\mathbb{Q}$  中不存在上确界. 同理  $Y$  在  $\mathbb{Q}$  中没有下确界.

**定义 3.6 (确界原理)** 设  $S$  是有序集, 如果  $S$  中的任何非空、有上界的子集  $E$ , 在  $S$  中必有上确界, 即  $\sup E$  存在并且  $\sup E \in S$ , 那么称  $S$  满足上确界原理.

如果  $S$  中的任何非空、有下界的子集  $E$ , 在  $S$  中必有下确界, 即  $\inf E$  存在并且  $\inf E \in S$ , 那么称  $S$  满足下确界原理.

同时满足上、下确界原理的有序集  $S$ , 称为满足确界原理.

上例说明,  $\mathbb{Q}$  是不满足确界原理的. 下面的定理告诉我们满足上确界原理和满足下确界原理本质上是等价的.

**定理 3.7** 设  $S$  是有序集, 则  $S$  满足上确界原理当且仅当  $S$  满足下确界原理.

**证明** 不妨设  $S$  满足上确界原理. 那么对  $S$  的任意有下界的子集  $E$ , 要证  $E$  在  $S$  中有下确界. 记  $E'$  所有下界的集合为

$$E' = \{x \mid x \in S, x \text{ 是 } E \text{ 的下界}\}$$

显然,  $E$  中的元素都是  $E'$  的上界, 根据条件存在  $\alpha = \sup E' \in S$ .

下面要证明  $\alpha$  也是  $E$  在  $S$  中的下确界.

首先证明  $\alpha$  是  $E$  的下界. 若不然, 存在  $\gamma \in E$ , 使得  $\gamma < \alpha$ , 因此  $\gamma$  不是  $E'$  的上界, 这与  $E$  中的元素都是  $E'$  的上界矛盾.

其次证明  $\alpha$  是  $E$  的最大下界. 对  $E$  的任意的一个下界  $\beta$ , 根据  $E'$  的定义, 有  $\beta \in E'$ , 所以  $\beta \leq \alpha$ .

**定义 3.8 (连续(完备)性公理)** 设  $S$  是有序集, 如果对  $S$  中任意两个非空子集

$X$  和  $Y$ , 只要满足

$$x \leqslant y, \quad x \in X, \quad y \in Y,$$

就一定存在  $c \in S$ , 使得

$$x \leqslant c \leqslant y.$$

那么称  $S$  满足连续(或完备)性公理.

我们已经看到  $\mathbb{Q}$  的两个子集

$$X = \{x \mid x \in \mathbb{Q}, x < 0; \text{ 或 } x^2 < 2, x > 0\}, \quad Y = \{y \mid y \in \mathbb{Q}, y^2 > 2, y > 0\}$$

之间不存在有理数, 也就不满足完备性公理.

**定理 3.9** 有序集  $S$  满足连续性公理, 当且仅当  $S$  满足确界原理.

**证明** 若连续性公理成立, 对任何非空有上界集合  $X \subset S$ , 令

$$Y = \{y \mid y \in S \text{ 是 } X \text{ 的上界}\},$$

因此对任意的  $x \in X, y \in Y$ , 有  $x \leqslant y$ . 根据连续性公理, 一定存在  $c \in S$ , 使得对任意的  $x \in X, y \in Y$ , 有  $x \leqslant c \leqslant y$ . 第一个不等式说明  $c$  是  $X$  的上界, 第二个不等式说明  $c$  是最小上界, 即  $c = \sup X$ .

反之, 若  $S$  满足确界原理, 那么任取两个非空子集  $X$  和  $Y$ , 它们满足对于任何  $x \in X, y \in Y$ , 有  $x \leqslant y$ . 则  $Y$  中的任何元素都是  $X$  的上界. 因此  $X$  有上确界, 记为  $c = \sup X$ . 因此  $c \leqslant y$  对任何  $y \in Y$  成立, 同时  $x \leqslant c$ , 对任意的  $x \in X$  成立.

**定义 3.10 (域)** 设  $F$  是一个集合, 它具有加法和乘法运算, 若这些运算满足下列公理, 则称  $F$  为**域**

**加法公理:** 对任意的  $x, y \in F$ , 可定义  $x + y \in F$ . 加法运算满足

- 1、有零元 0: 且  $x + 0 = 0 + x = x$ .
- 2、有负元: 对每个  $x \in F$ , 有  $-x \in F$ , 且  $x + (-x) = -x + x = 0$ .
- 3、交换律:  $x + y = y + x$ .
- 4、结合律:  $x + (y + z) = (x + y) + z$ .

**乘法公理:** 对任意的  $x, y \in F$ , 可定义  $x \cdot y \in F$  且满足

- 1、有单位元 1:  $1 \cdot x = x \cdot 1 = x$ .
- 2、有逆元: 对任意的  $x \in F, x \neq 0$ , 有  $x^{-1} \in F$ , 使得  $x \cdot x^{-1} = x^{-1} \cdot x = 1$ .
- 3、交换律:  $x \cdot y = y \cdot z$ .
- 4、结合律:  $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ .
- 5、分配律:  $(x + y) \cdot z = x \cdot z + y \cdot z$ .

**定义 3.11 (有序域)** 设  $F$  即是有序集, 又是域. 若还满足如下条件, 则称为有序域.

1、当  $x, y, z \in F$ , 且  $y < z$  时, 有  $x + y < x + z$ ;

2、当  $x, y \in F$ , 且  $x > 0, y > 0$ , 则  $xy > 0$ .

至此, 我们抽象地定义了集合的序、域、确界原理以及与之等价的连续(完备)性公理, 这三方面的概念统称为**实数公理**.

虽然有理数  $\mathbb{Q}$  有序, 而且是域, 但它不满足确界原理或与之等价的连续性公理. 我们的目的, 就是要构造一个满足上述公理的, 并且还包含  $\mathbb{Q}$  作为子集合的集合.

下面的定理是本专题的核心, 它说明满足实数公理的集合是存在的.

**定理 3.12** 存在满足实数公理的集合, 称为实数域, 记为  $\mathbb{R}$ , 其中的元素称为实数. 而且包含有理数域  $\mathbb{Q} \subset \mathbb{R}$ .

证明的过程实际上是一个构造实数域的过程, 我们将在最后一节给出.

实数域  $\mathbb{R}$  还有两个特点:

一是几何上看, 实数与数轴上的点 1-1 对应. 因此“实数”和“点”不再区分, 或称数是点的坐标. 整数、有理数、无理数对应的点分别称为“整数点”、“有理点”和“无理点”.

二是  $\mathbb{R}$  的基数大于  $\mathbb{Q}$  的基数, 因此是不可数的, 这里就不再证明了.

**定理 3.13** 实数域  $\mathbb{R}$  满足

(1) **Archimedes 性:** 若  $x, y \in \mathbb{R}$  且  $x > 0$ , 则一定存在最小整数  $n$ , 使得

$$(n - 1)x \leqslant y < nx.$$

若  $y > 0$ , 则  $n > 0$ . 若  $x = 1$ , 则对任意实数  $y$ , 一定存在整数  $n$ , 使得

$$n - 1 \leqslant y < n.$$

(2) **有理数在实数中的稠密性:** 若  $x, y \in \mathbb{R}$  且  $x < y$ , 则一定存在  $c \in \mathbb{Q}$ , 使得

$$x < c < y.$$

**证明** 设  $E = \{nx \mid n \in \mathbb{Z}\}$ , 若对任意的  $n \in \mathbb{Z}$ , 有  $nx \leqslant y$ , 即  $y$  是  $E$  在  $\mathbb{R}$  中的上界, 因此有上确界  $\alpha = \sup E \in \mathbb{R}$ .

又因为  $x > 0$ , 所以  $\alpha - x$  不是  $E$  的上界, 也就是存在整数  $m$ , 使得  $mx \in E$ ,  $\alpha - x < mx$ , 这样就有  $\alpha < (m + 1)x \in E$ , 这与  $\alpha$  是上确界矛盾. 因此存在  $n_1 \in \mathbb{Z}$ , 使得  $n_1x > y$

将上述结果应用到  $x > 0$  和  $-y \in \mathbb{R}$  上, 则存在  $n_2 \in \mathbb{Z}$ , 使得  $n_2x > -y$ . 两者结

合起来, 就是存在  $n_1, n_2 \in \mathbb{Z}$ , 使得

$$-n_2x < y < n_1x.$$

记

$$S = \{k \mid k \in \mathbb{N}, \text{ 使得 } kx > y\},$$

则  $S$  包含  $n_1$  因此非空, 同时对  $k \in S$ , 有

$$k > \frac{y}{x} > -n_2,$$

即  $-n'$  是  $S$  的一个下界, 根据第2讲例2.1.1,  $S$  中有最小数  $n$ , 使得

$$(n-1) \leq y < nx.$$

关于(2) 的证明如下, 由于  $x < y$ , 得  $y - x > 0$ , 对  $y - x$  和 1, 根据 (1), 存在整数  $n$ , 使得

$$n(y - x) > 1, \text{ 或 } ny > nx + 1.$$

因  $y - x > 0$ , 所以  $n > 0$  是正整数. 再对 1 和  $nx$  利用 (1), 存在整数  $m_1$  使得

$$m - 1 \leq nx < m.$$

综合上述不等式, 有

$$nx < m \leq nx + 1 < ny.$$

因  $n > 0$ , 从而

$$x < \frac{m}{n} < y.$$

□

实数的绝对值定义为, 对任意的  $a \in \mathbb{R}$ ,

$$|a| = \begin{cases} a & \text{当 } a \geq 0; \\ -a, & \text{当 } a < 0. \end{cases}$$

绝对值满足

1°  $|a| \geq 0$ , 等号成立当且仅当  $a = 0$ , 即正定性;

2°  $|a - b| = |b - a|$ , 即对称性;

3°  $|a + b| \leq |a| + |b|$ , 即三角不等式.

因此两个数差的绝对值给出对应数轴上点的距离.

### §3.4 十进制小数

下面, 主要讨论熟知的实数的十进制小数(以下简称小数)与本专题所定义的实数之间的关系. 根据上节的定义, 假设满足确界原理或等价的连续性公理实数域  $\mathbb{R}$  存在.

设  $x > 0$  是实数, 根据Archimedes 性, 存在非负整数  $a_0 \geq 0$ , 使得

$$a_0 \leq x < a_0 + 1.$$

对  $0 \leq 10(x - a_0) < 10$  在利用Archimedes 性, 存在非负整数  $a_1$ , 使得

$$a_1 \leq 10(x - a_0) < a_1 + 1 \quad \text{或} \quad a_0 + \frac{a_1}{10} \leq x < a_0 + \frac{a_1}{10} + \frac{1}{10}.$$

显然  $a_1 + 1 \leq 10$  或  $0 \leq a_1 \leq 9$ , 若不然, 有  $a_1 \geq 10$ , 推得  $10(x - a_0) \geq a_1 \geq 10$ , 进而得  $x \geq a_0 + 1$ , 这与  $a_0$  的选取矛盾.

若存在 0 和 9 之间的非负整数  $a_1, a_2, \dots, a_n$  使得

$$a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n} \leq x < a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n} + \frac{1}{10^n}$$

对  $n$  成立, 则对

$$0 \leq 10^{n+1} \left( x - a_0 - \frac{a_1}{10} - \frac{a_2}{10^2} - \cdots - \frac{a_n}{10^n} \right) < 10$$

利用Archimedes 性, 存在非负整数  $0 \leq a_{n+1} \leq 9$ , 使得上式对  $n + 1$  也成立.

令  $E$  是按上述步骤得到的数

$$a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n}, \quad n = 0, 1, 2, \dots,$$

所组成的集合. 因此,  $x$  是  $E$  在  $\mathbb{R}$  中的一个上界, 因此  $E$  在  $\mathbb{R}$  中有上确界. 下面要说明  $x = \sup E$ .

若不然, 令  $x' = \sup E < x$ , 那么

$$0 < x - x' < \frac{1}{10^n}, \quad n = 1, 2, \dots,$$

这是不可能的. 因此  $x = \sup E$ , 它可表示为下列小数形式

$$x = a_0.a_1a_2\cdots a_n\cdots = a_0 + \frac{a_1}{10} + \cdots + \frac{a_n}{10^n} + \cdots$$

其中  $0 \leq a_1, a_2, \dots, a_n, \dots \leq 9$ .

几何上看比较直观, 即选定  $a_0 \leq x < a_0 + 1$  后, 将数轴上以  $a_0$  和  $a_0 + 1$  为端点的区间  $[a_0, a_0 + 1]$  进行 10 等分, 选择一个等分区间

$$\left[ a_0 + \frac{a_1}{10}, a_0 + \frac{a_1 + 1}{10} \right], \quad 0 \leq a_1 \leq 9,$$

使得

$$a_0 + \frac{a_1}{10} \leqslant x < a_0 + \frac{a_1}{10} + \frac{1}{10},$$

如此下去就得到  $x$  的十进制小数表达式. 当然这种几何描述并不严谨, 这里仅做一个直观上的展示.

在十进制小数中  $a_1, a_2, \dots, a_n, \dots$  都是介于 0 和 9 之间的非负整数. 这就产生了三种可能.

### 1° 有限小数

若  $a_1, a_2, \dots, a_n, \dots$  中只有有限项非零, 不妨设  $a_j = 0, j > m$ , 那么

$$x = a_0.a_1a_2 \cdots a_m = a_0 + \frac{a_1}{10} + \cdots + \frac{a_m}{10^m}.$$

称为有限小数. 通过通分, 有限的小数可以表示成分数形式  $x = \frac{p}{q}$ , 其中  $q = 10^m$ . 如果  $p$  和  $q$  有公因子, 可以简化成不可约分数.

但是并不是所有有理数都可以表示成有限小数. 例如  $\frac{5}{11}$  是不能表示为有限小数的. 这是因为假如

$$\frac{5}{11} = \frac{b}{10^n}$$

则  $5 \cdot 10^n = 11 \cdot b$ , 推出 11 能整除  $10^n$ . 这显然是不可能的.

### 2° 无限循环小数

若  $a_1, a_2, \dots, a_n, \dots$  中出现无限循环情况, 即从某项  $a_n$  开始, 存在一个正整数  $k$ , 使得  $a_{n+1}, a_{n+2}, \dots, a_{n+k}$  重复出现, 对下标使用同余的表述, 就是当  $n+l$  与  $n+j$  模  $k$  相等时, 对应的非负整数相等:

$$a_{n+l} = a_{n+j} \text{ 当 } l \equiv j \pmod{k}, \quad j = 1, 2, \dots, k$$

这样的小数称为无限循环小数, 记为

$$x = a_0.a_1a_2 \cdots a_n \dot{a}_{n+1} \cdots \dot{a}_{n+k}$$

那么

$$10^n(x - a_0.a_1a_2 \cdots a_n) = 0.\dot{a}_{n+1} \cdots \dot{a}_{n+k}$$

所以

$$\begin{aligned} 10^{n+k}(x - a_0.a_1a_2 \cdots a_n) &= a_{n+1} \cdots a_{n+k} + 0.\dot{a}_{n+1} \cdots \dot{a}_{n+k} \\ &= a_{n+1} \cdots a_{n+k} + 10^n(x - a_0.a_1a_2 \cdots a_n) \end{aligned}$$

解得

$$x = a_0.a_1a_2 \cdots a_n + \frac{a_{n+1} \cdots a_{n+k}}{10^{n+k} - 10^n}$$

所以无限循环小数也是有理数.

反之,任意有理数,如果不是有限小数,则一定是无限循环小数.例如,  $\frac{5}{11} = 0.\dot{4}\dot{5}$ .

一般情况下,设  $\frac{q}{p}$  是有理数,其中  $0 < q < p$ ,  $(q, p) = 1$ , 则存在正整数  $r$  使得

$$10^{r-1}q < p, \quad 10^r q > p$$

为了简化,不妨设  $r = 1$  也就是

$$10q > p$$

因此存在整数  $a_1, 0 \leq a_1 \leq 9$  使得

$$10q = a_1 p + q_1, \quad 0 \leq q_1 < p$$

所以

$$\frac{q}{p} = \frac{a_1}{10} + \frac{1}{10} \frac{q_1}{p}$$

对  $\frac{q_1}{p}$  重复上述过程

$$\frac{q_1}{p} = \frac{a_2}{10} + \frac{1}{10} \frac{q_2}{p}$$

所以

$$\frac{q}{p} = \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{1}{10^2} \frac{q_2}{p}$$

如此下去  $0 \leq a_n \leq 9, 0 \leq q_n < p$ , 如果有一个  $q_n = 0$ , 上述过程终止,  $\frac{q}{p}$  就是有限小数,  
如果  $q_n \neq 0$ , 则在 0 到  $p$  的有限个整数中, 必有相等, 因此出现无限循环小数.

### 3° 无限不循环小数

除了前两种情况之外,  $a_1, a_2, \dots, a_n, \dots$  中既不是有限个非零, 也不出现循环, 因此称为无限不循环小数, 它正是我们构造的无理数.

## §3.5 Dedekind分割\*

Dedekind 用有理数域  $\mathbb{Q}$  的分割来定义实数, 并证明这样定义的实数满足实数公理, 并包含  $\mathbb{Q}$  因此就是定理3.12 一种构造性证明. 具体思想如下:

**定义 3.14** 把有理数域  $\mathbb{Q}$  分割为两个非空且不相交的子集  $A$  和  $A'$ ,

$$\mathbb{Q} = A \cup A', \quad A \cap A' = \emptyset,$$

使得对任意的  $a \in A$  和  $a' \in A'$ , 有  $a < a'$ . 集合  $A$  称为分割的下组, 集合  $A'$  称为分割的上组, 并将这种分割记成  $A|A'$ . 这里  $\emptyset$  表示空集.

在上述定义下,  $A$  和  $A'$  在  $\mathbb{Q}$  中互为余集:  $A = \mathbb{Q} \setminus A'$ ,  $A' = \mathbb{Q} \setminus A$ . 即对任意的集合  $A \subset \mathbb{Q}$ ,  $A$  在  $\mathbb{Q}$  中的余集为

$$\mathbb{Q} \setminus A = \{a \mid a \in \mathbb{Q}, \text{ 但 } a \notin A\}.$$

从数轴上看, 一个分割实际上是把数轴上所有有理点分成左边一部分(即是下组)和右边一部分(即是上组). 因此就得出分割的三种可能:

(1) 在下组  $A$  内无最大数, 而在上组  $A'$  内有最小数  $r$ ,

$$A = \{a \mid a \in \mathbb{Q}, a < r\}, \quad A' = \{a' \mid a' \in \mathbb{Q}, a' \geq r\};$$

(2) 在下组  $A$  内有最大数, 而在上组  $A'$  内无最小数  $r$ ,

$$A = \{a \mid a \in \mathbb{Q}, a \leq r\}, \quad A' = \{a' \mid a' \in \mathbb{Q}, a' > r\};$$

(3) 在下组  $A$  内无最大数, 而在上组  $A'$  内也无最小数, 例如

$$A = \{a \mid a \in \mathbb{Q}, a < 0; \text{ 或 } a > 0 \text{ 但 } a^2 < 2\},$$

$$A' = \{a' \mid a' \in \mathbb{Q}, a' > 0 \text{ 且 } a'^2 > 2\}.$$

因为分割是对有理数域  $\mathbb{Q}$  的分割, 所以以下我们说的“数”当然是指有理数.

上述三种分割, 前两种是有理数产生的, 称这样的分割为有理分割. 为了确定起见, 我们约定: 凡是说到有理数  $r$  确定的有理分割时, 常把这数放在上组内, 这样我们只考虑第一种和第三种分割. 不管是第一种还是第三种分割, 分割的下组内都无最大数.

那么会不会有第四种可能的分割呢? 即分割  $A|A'$  中, 下组有最大数, 上组有最小数. 事实上, 这样的分割是不存在的. 我们可以采取反证法来说明:

假如存在  $\mathbb{Q}$  的一个分割  $A|A'$ , 使得  $A$  有最大数  $a^* \in A$ , 而  $A'$  有最小数  $b^* \in A'$ . 根据分割的定义, 有  $a^* < b^*$ . 但是, 有理数  $c = \frac{a^* + b^*}{2}$  满足  $a^* < c < b^*$ , 即  $c$  大于下组的最大数, 小于上组的最小数, 所以  $c$  既不属于下组  $A$ , 又不属于上组  $A'$ , 这显然与  $A$  和  $A'$  是  $\mathbb{Q}$  的一个分割矛盾.

Dedekind 的思想是一个分割对应一个“数”, 当分割是有理分割时, 对应的数就是有理数, 例如 0, 1 和一般的有理数  $r$  分别为

$$0 = A_0|A'_0, \quad A_0 = \{a \mid a < 0\}, \quad A'_0 = \{a' \mid a' \geq 0\};$$

$$1 = A_1|A'_1, \quad A_1 = \{a \mid a < 1\}, \quad A'_1 = \{a' \mid a' \geq 1\};$$

$$r = A_r|A'_r, \quad A_r = \{a \mid a < r\}, \quad A'_r = \{a' \mid a' \geq r\}.$$

当分割不是有理分割, 即是(3)中形式的分割时, 就对应一个新的“数”, 称为“无理数”. 因此把所有分割构成的集合定义成实数集合.

**定理 3.15** 下列集合满足实数公理, 因此给出了一个实数模型.

$$\mathbb{R} = \{A|A' : \mathbb{Q} \text{ 的所有分割}\}$$

下面, 我们就要逐一验证上述集合是有序域, 同时满足确界原理, 因此它就是定理3.12 中所要求的实数域. 这里我们仅仅验证以下四点, 一是如何在两个分割之间定义“大小”, 二是两个分割怎么做加法, 三是两个分割如何相乘, 四是这些由分割形成的集合  $\mathbb{R}$  满足确界原理.

### 1° 分割的序

设  $\alpha = A_\alpha|A'_\alpha$ ,  $\beta = A_\beta|A'_\beta$  是两个分割, 则根据分割的定义, 要么  $A_\alpha \subset A_\beta$ ,  $A'_\beta \subset A'_\alpha$  要么  $A_\beta \subset A_\alpha$ ,  $A'_\alpha \subset A'_\beta$ . 因此

当  $A_\alpha \not\subset A_\beta$ ,  $A'_\beta \not\subset A'_\alpha$  时, 就定义  $\alpha < \beta$ .

当  $A_\beta \not\subset A_\alpha$ ,  $A'_\alpha \not\subset A'_\beta$  时, 就定义  $\alpha > \beta$ .

当  $A_\alpha = A_\beta$ ,  $A'_\alpha = A'_\beta$  时, 就定义  $\alpha = \beta$ .

这样我们就定义了分割之间的一种序(也就是“大小”). 可以验证上述定义满足序公理的有关性质.

特别, 对于一个分割  $\alpha = A_\alpha|A'_\alpha$ , 如果  $A_\alpha$  中存在正的有理数, 则  $A_0 \subset A_\alpha$ , 所以  $\alpha > 0$ , 反之  $\alpha \leq 0$ .

### 2° 分割的加法

**性质 3.16** 设  $\alpha = A_\alpha|A'_\alpha$ ,  $\beta = A_\beta|A'_\beta$  是两个分割, 令

$$A_\gamma = \{a + b \mid a \in A_\alpha, b \in A_\beta\}, A'_\gamma = \mathbb{Q} \setminus A_\gamma,$$

则它们给出  $\mathbb{Q}$  的一个分割  $\gamma = A_\gamma|A'_\gamma$ , 并定义这个分割为

$$\gamma = \alpha + \beta.$$

这里再次强调  $\mathbb{Q} \setminus A_\gamma$  表示  $\mathbb{Q}$  中不属于  $A_\gamma$  的那些有理数:

$$\mathbb{Q} \setminus A_\gamma = \{a \mid a \in \mathbb{Q}, \text{ 但 } a \notin A_\gamma\}.$$

**证明** 显然  $A_\gamma$ ,  $A'_\gamma$  都是有理数的子集合, 要证明  $A_\gamma|A'_\gamma$  是一个分割, 也就是要证明对任意的  $c = a + b \in A_\gamma$ ,  $c' \in A'_\gamma$ , 有  $c < c'$ .

这是因为对任意的  $a \in A_\alpha$ ,  $c' - a \notin A_\beta$ , 因此  $c' - a \in A'_\beta$ . 否则存在  $d \in A_\beta$  使得  $c' - a = d$ , 推出  $c' = a + d \in A_\gamma$ , 这与  $A'_\gamma$  的定义相矛盾.

因为  $c' - a \in A'_\beta$ , 所以对任意的  $b \in A_\beta$ , 有  $c' - a > b$ , 即  $c' > a + b = c$ .  $\square$

这样我们就定义了两个分割之间的加法, 可以验证这样定义的加法满足交换律和结合律.

以下两种情况还需要进一步明确, 为此我们有下列两个性质.

**性质 3.17** 对任意的  $\alpha = A_\alpha | A'_\alpha$ , 有

$$\alpha + 0 = \alpha,$$

这里  $0 = A_0 | A'_0$ ,  $A_0 = \{a \mid a < 0\}$ ,  $A'_0 = \{a' \mid a' \geq 0\}$ .

**证明** 设

$$A_\gamma = \{a + b \mid a \in A_\alpha, b \in A_0\},$$

由此给出分割  $\gamma = A_\gamma | A'_\gamma$ . 一方面由于  $b < 0$ , 因此, 对任意的  $a \in A_\alpha$ ,  $a' \in A'_\alpha$ , 推得  $a + b < a < a'$ , 所以  $a + b \in A_\alpha$ . 这样我们就证明了  $A_\gamma \subset A_\alpha$ .

另一方面, 由于  $A_\alpha$  没有最大数, 所以对任意的  $a \in A_\alpha$ , 一定存在一个有理数  $\tilde{a}$  满足  $\tilde{a} \in A_\alpha$ ,  $\tilde{a} > a$ . 记  $b = a - \tilde{a} < 0$ , 所以  $b \in A_0$ , 就推得  $a = \tilde{a} + b \in A_\gamma$ . 这样我们就证明了  $A_\gamma \supset A_\alpha$ .

综合两方面结论, 有  $A_\gamma = A_\alpha$ , 也就是  $\alpha + 0 = \alpha$ .  $\square$

**性质 3.18** 对于任何一个分割  $\alpha = A_\alpha | A'_\alpha$ , 定义

$$A_\beta = \{b \mid \text{存在 } a' \in A'_\alpha, \text{ 使得 } b < -a'\}, A'_\beta = \mathbb{Q} \setminus A_\beta,$$

则  $\beta = A_\beta | A'_\beta$  是一个分割, 且  $\beta = -\alpha$ , 即  $\beta$  是  $\alpha$  的负元.

**证明** 任取  $b \in A_\beta$ , 所以存在  $a' \in A'_\alpha$ , 使得  $b < -a'$ .

任取  $b' \in A'_\beta$ , 即  $b' \notin A_\beta$ , 所以对于任意的  $a' \in A'_\alpha$ , 有  $b' \geq -a' > b$ . 这样我们首先验证了  $A_\beta | A'_\beta$  是一个分割.

其次, 我们要证明  $\beta = -\alpha$ , 为此, 只要证明  $\alpha + \beta = 0$ , 也就是要证明

$$A_\gamma = \{a + b \mid a \in A_\alpha, b \in A_\beta\} = A_0.$$

为此, 取任意的  $a + b \in A_\gamma$ , 由  $A_\gamma$  的定义知,  $-b \in A'_\alpha$ , 所以  $-b > a$ , 也就是  $a + b < 0$ , 推得  $a + b \in A_0$ , 即  $A_\gamma \subset A_0$ .

反过来, 对任意的  $d \in A_0$ , 有  $d < 0$ , 所以  $-d > 0$ , 取  $a \in A_\alpha$ , 那么一定存在一个非负整数  $n$ , 使得  $a + (-nd) \in A_\alpha$ , 但是  $a + (-n-1)d \in A'_\alpha$ , 也就是  $(n+1)d - a \in A_\beta$ . 这样就可以把  $d$  表示成  $A_\alpha$  与  $A_\beta$  中两个数的和:

$$d = (a - nd) + (n+1)d - a \in A_\gamma.$$

所以我们有证明了  $A_0 \subset A_\gamma$ . 因此也就完成了证明.  $\square$

3° 分割的乘法

设  $\alpha = A_\alpha | A'_\alpha$ ,  $\beta = A_\beta | A'_\beta$  是两个分割, 为了定义两者之间的乘法, 我们分如下四种情况

$$(1) \alpha \geq 0, \beta \geq 0;$$

$$(2) \alpha < 0, \beta < 0;$$

$$(3) \alpha \geq 0, \beta < 0;$$

$$(4) \alpha < 0, \beta \geq 0.$$

对于(1), 我们有

**性质 3.19** 设  $\alpha \geq 0, \beta \geq 0$ , 令

$$A_\gamma = \{c \mid c < 0, \text{ 或者存在 } a \in A_\alpha, b \in A_\beta, a > 0, b > 0 \text{ 使得 } c = ab\},$$

$$A'_\gamma = \mathbb{Q} \setminus A_\gamma,$$

那么  $\gamma = A_\gamma | A'_\gamma$  是  $\mathbb{Q}$  的一个分割, 记为  $\gamma = \alpha \cdot \beta$ .

**证明** 若  $\beta = 0$ , 即  $A_\beta = A_0$ . 根据定义,  $A_\gamma = \{c \mid c < 0\} = A_0$ . 因此有  $A'_\gamma = A'_0$ , 显然  $A_\gamma | A'_\gamma$  是一个分割, 在这种情况下  $0 = \alpha \cdot 0$ .

若  $\beta > 0$ , 任取  $c' \in A'_\gamma$ , 因为  $\alpha \geq 0$ , 所以  $c' \geq 0$ .

任取  $c \in A_\gamma$ , 如果  $c < 0$ , 显然有  $c < c'$ ; 如果  $c > 0$ , 意味着存在  $a > 0, b > 0$ , 使得  $c = ab$ . 那么  $\frac{c'}{b} \notin A_\alpha$ , 推得  $\frac{c'}{b} \in A'_\alpha$ , 也就是  $\frac{c'}{b} > a$ , 最后得  $c' > ab = c$ . 所以  $A_\gamma | A'_\gamma$  分割.  $\square$

为了定义其他情形下分割之间的乘法, 我们首先要定义分割的“绝对值”.

**定义 3.20** 分割  $\alpha = A_\alpha | A'_\alpha$  的绝对值  $|\alpha|$  定义为

$$|\alpha| = \begin{cases} \alpha, & \text{若 } \alpha \geq 0; \\ -\alpha, & \text{若 } \alpha < 0. \end{cases}$$

或者说  $|\alpha|$  对应的分割  $|\alpha| = A_{|\alpha|} | A'_{|\alpha|}$  满足

$$A_{|\alpha|} = \begin{cases} A_\alpha, & \text{若 } \alpha \geq 0; \\ A_{-\alpha}, & \text{若 } \alpha < 0. \end{cases}$$

这里  $A_{-\alpha}$  的定义由性质 3.18 给出.

显然  $|\alpha| \geq 0$ , 并且  $|\alpha| = 0$  当且仅当  $\alpha = 0$ .

**定义 3.21** 设  $\alpha, \beta$  是任意两个分割, 它们的积  $\alpha\beta$  定义为:

$$\alpha\beta = \begin{cases} -|\alpha||\beta|, & \text{若 } \alpha > 0, \beta \leq 0, \\ -|\alpha||\beta|, & \text{若 } \alpha \leq 0, \beta > 0, \\ |\alpha||\beta|, & \text{若 } \alpha < 0, \beta < 0. \end{cases}$$

因为  $|\alpha| \geq 0, |\beta| \geq 0$ , 所以  $|\alpha|$  和  $|\beta|$  的乘法已经由性质3.19给出, 因此上述定义是合理的.

最后, 我们讨论一个非零分割的逆.

**性质 3.22** 如果分割  $\alpha > 0, \alpha = A_\alpha | A'_\alpha$ , 定义

$$A_\beta = \left\{ b \mid b \leq 0 \text{ 或者存在 } a' \in A'_\alpha, \text{ 使得 } b < \frac{1}{a'} \right\}, \quad A'_\beta = \mathbb{Q} \setminus A_\beta,$$

则  $\beta = A_\beta | A'_\beta$  是一个分割, 且  $\alpha\beta = 1$ .  $\beta$  称为  $\alpha$  的逆, 记为  $\beta = \alpha^{-1}$ .

如果  $\alpha < 0$ , 那么它的逆定义为  $-(-\alpha)^{-1}$ .

**证明** 任取  $b \in A_\beta, b > 0$ , 根据定义, 存在  $a' \in A'_\alpha$ , 使得  $b < \frac{1}{a'}$ .

任取  $b' \in A'_\beta$ , 则  $b' \notin A_\beta$ , 所以  $b' > 0$ , 且对任意的  $a' \in A'_\alpha$ , 有  $b' \geq \frac{1}{a'}$ , 所以  $b < b'$ , 即  $\beta = A_\beta | A'_\beta$  是一个分割.

根据乘法, 记  $\gamma = \alpha \cdot \beta, \gamma = A_\gamma | A'_\gamma$ , 其中  $A_\gamma$  和  $A'_\gamma$  的定义由性质3.19给出. 现在要证明  $\gamma = 1$ , 即是要证明  $A_\gamma = A_1$ .

任取  $c \in A_\gamma$ , 若  $c < 0$ , 显然有  $c \in A_1$ . 若  $c > 0$ , 则存在  $a \in A_\alpha, b \in A_\beta, a > 0, b > 0$ , 使得  $c = ab$ . 由于  $b > 0$ , 所以存在  $a' \in A'_\alpha$  使得  $b < \frac{1}{a'}$ , 由此得  $c = ab < \frac{a}{a'} < 1$ . 也就是我们首先证明了  $A_\gamma \subset A_1$ .

反之, 任取  $c \in A_1$ , 不妨设  $c > 0$ , 由于  $c < 1$ , 我们有  $\frac{1}{c} > 1$ . 因此对任意的  $a \in A_\alpha$ , 存在一个非负整数  $n$ , 使得

$$a_1 = \frac{a}{c^n} \in A_\alpha, \text{ 但是 } a'_1 = \frac{a}{c^{n+1}} \in A'_\alpha,$$

(1) 如果  $a'_1 = \frac{a}{c^{n+1}}$  不是  $A'_\alpha$  的最小数, 那么存在  $a' \in A'_\alpha$ , 使得  $a'_1 > a'$ . 记  $b_1 = \frac{1}{a'_1}$ , 则

$$b_1 = \frac{1}{a'_1} < \frac{1}{a'},$$

所以  $b_1 \in A_\beta$ , 这样我们就得到

$$c = \frac{a}{c^n} \frac{c^{n+1}}{a} = \frac{a}{c^n} \frac{1}{a'_1} = a_1 b_1 \in A_\gamma,$$

也就是  $A_1 \subset A_\gamma$ .

(2) 如果  $a'_1 = \frac{a}{c^{n+1}}$  是  $A'_\alpha$  的最小数, 记

$$\tilde{a} = \frac{a_1 + a'_1}{2} = \frac{1}{c^n} \frac{a}{2} \left( 1 + \frac{1}{c} \right),$$

那么  $\tilde{a} < a'_1$ , 所以  $\tilde{a} \in A_\alpha$ , 并且存在非负整数  $m$ , 使得

$$a_2 = \frac{1}{c^m} \tilde{a} \in A_\alpha, \text{ 但是 } a'_2 = \frac{1}{c^{m+1}} \tilde{a} \in A'_\alpha,$$

不难计算

$$a'_2 = \frac{1}{c^{m+1}} \tilde{a} = \frac{1}{c^{m+n+1}} \frac{1}{2} \left(1 + \frac{1}{c}\right) a > \frac{1}{c^{n+1}} a = a'_1.$$

所以类似(1), 存在  $a' \in A'_\alpha$ , 使得  $a'_2 > a'$ . 记  $b_2 = \frac{1}{a'_2}$ , 则  $b_2 < \frac{1}{a'}$ , 推得  $b_2 \in A_\beta$ , 最后我们有  $c = a_2 b_2 \in A_\gamma$ .

不管是情形(1)还是情形(2), 都有  $c \in A_\gamma$ , 即  $A_1 \subset A_\gamma$ . 这样我们就证明了  $A_\gamma = A_1$ , 也就是  $\alpha \cdot \beta = \gamma = 1$ .  $\square$

至此, 我们在分割集合  $\mathbb{R}$  上定义了序、加法和乘法运算, 定义了加法的 0 元、乘法的单位元和非零元的逆元. 关于序的传递性、加法和乘法的交换律、结合律以及乘法对加法的分配律比较简单, 读者可以自行完成验证.

#### 4° 完备性

最后, 我们要验证

**性质 3.23** 定理 3.15 中给出的集合

$$\mathbb{R} = \{A|A' : \mathbb{Q} \text{ 的所有分割}\}$$

满足确界原理.

**证明** 设

$$X = \{\alpha = A_\alpha|A'_\alpha\} \subset \mathbb{R}$$

是  $\mathbb{R}$  中有上界的子集合,  $\beta = B_\beta|B'_\beta \in \mathbb{R}$  是  $X$  的任意一个上界, 即对于任意的  $\alpha = A_\alpha|A'_\alpha \in X$ , 有  $\alpha \leq \beta$ , 也就是  $A_\alpha \subset B_\beta$ . 那么首先验证

$$A_{\alpha_0} = \bigcup_{\alpha \in X} A_\alpha, \quad A'_{\alpha_0} = \mathbb{Q} \setminus A_0$$

是  $\mathbb{Q}$  的一个分割. 这是因为对任意的  $a \in A_{\alpha_0}$ , 推得  $a \in A_\alpha$ , 对某个  $\alpha \in X$  成立; 对任意的  $b \in A'_{\alpha_0}$ , 推得  $b \notin A_\alpha$ ,  $\alpha \in X$ , 所以  $a < b$ .

其次验证  $\alpha_0 = A_{\alpha_0}|A'_{\alpha_0}$  是  $X$  的上确界. 这是因为对任意的  $\alpha = A_\alpha|A'_\alpha \in X$ ,

$$A_\alpha \subset A_{\alpha_0}, \quad A'_{\alpha_0} \supset A'_\alpha,$$

所以  $\alpha_0 \geq \alpha$ . 而对任意的  $\alpha = A_\alpha|A'_\alpha \in X$ ,  $A_\alpha \subset B_\beta$ , 所以  $A_{\alpha_0} \subset B_\beta$ , 即  $\alpha_0 \leq \beta$ . 所以  $\alpha_0$  是最小上界.  $\square$

这样, 我们就证明了定理 3.15.

Dedekind 的思想是伟大的, 他实际上已经涉及到一个哲学问题: “数是什么? 数应当是什么?”在他的思想里, 任何满足实数公理的集合, 不管这个集合的元素是什么, 或怎么构造的, 它们都是一个实数模型!

### 第3讲习题

1. 设  $n$  不是完全平方数的正整数, 证明  $\sqrt{n}$  是无理数.

2. (1) 设  $r, s$  是有理数, 若  $r + s\sqrt{2} = 0$ , 则  $r = s = 0$ .

(2) 验证下列集合

$$\mathbb{F} = \{r + s\sqrt{2} \mid r, s \in \mathbb{Q}\}$$

是一个数域.

3. 证明:  $\frac{1}{3}$  不能表示为有限的十进制小数.

4. 证明:  $\sqrt{2} + \sqrt{3} + \sqrt{5}$  不是有理数.

5. 设  $E$  是区间  $[0, 1]$  中所有有理数的集合. 试证可通过挖去以  $E$  中所有有理数为中心的开区间, 并使得这些开区间的总长度不超过  $\frac{1}{2}$ .

提示: 因为  $E$  是可数的, 因此  $E$  中有理数可排列  $E = \{r_1, r_2, \dots\}$ , 以每个  $r_n$  为中心挖去一个开区间, 通过控制开区间的长度, 并利用无穷求和方式求出开区间总长度的上界.

6. 设  $\xi$  是一个无理数, 证明下列集合

$$S = \{m + n\xi \mid m, n \in \mathbb{Z}\}$$

在  $\mathbb{R}$  中稠密.

提示: 集合  $S$  具有如下性质:  $l(m+n\xi) \in S$ ,  $(m+n\xi) \pm (m'+n'\xi) \in S$ , 这里  $l \in \mathbb{Z}$ . 要证明  $S$  在  $\mathbb{R}$  中稠密, 只要证明对任意的实数  $a, b$ , 必有  $S$  中的元素介于两者之间. 不妨设  $0 < a < b$ . 对于任意正整数  $i$ , 记  $n_i = -[i\xi]$ , 这里  $[x]$  表示一个实数  $x$  的整数部分, 例如  $[\pi] = 3, [2.14] = 2$ . 那么  $x_i = n_i + i\xi \in S$ , 而且  $x_i$  实际上表示了  $i\xi$  的小数部分, 因此  $0 < x_i < 1$ . 取正整数  $k$ , 使得  $\frac{1}{k} < b - a$ , 那么  $S$  中  $k+1$  个数  $x_1, x_2, \dots, x_{k+1}$ , 至少有一对  $x_i, x_j$  满足  $0 < x_j - x_i < \frac{1}{k}$ . 再利用 Archimedes 性 (见定理 3.13), 存在满足  $n(x_j - x_i) > a$  的最小正整数  $n$ , 然后说明  $n(x_j - x_i) < b$ , 最后利用  $S$  的性质及可完成证明.

## 第 4 讲 复数

复数虽然是中学必学内容, 但为了保持完整性, 我们还是简要回顾复数的起源、复数的基本运算以及复数的几何含义.

### §4.1 复数的起源和运算

复数的历史可以追溯到16世纪. 问题的起源现在看就是一个代数方程的求根问题. 例如, 对于整数为系数的一次代数方程

$$ax + b = 0, \quad a, b \in \mathbb{Z}, \quad a \neq 0,$$

如果  $a$  不能整除  $b$ , 那么该方程在整数范围内没有解, 必须把整数扩充到有理数才有有理数解  $x = -\frac{a}{b}$ . 对于二次方程

$$x^2 - 2 = 0,$$

在有理数域内不存在解, 因此有理数也不够用了, 促使人们不得不构造一个更广的实数域, 使得在实数范围内方程有解  $x = \sqrt{2}$ .

然而, 即使是实数范围内, 也无法解决所有二次代数方程的求解问题. 例如方程

$$x^2 + 1 = 0$$

就没有实数解. 历史上还有一个十分著名的问题, 即

**例 4.1.1** 如何把 10 分成两部分, 使其积等于 40.

显然这个问题就是求二次代数方程的解

$$x(10 - x) = 40,$$

结果发现, 这个方程也没有实数解.

如果我们继续尝试能否像从整数扩充到有理数, 从有理数扩充到实数一样, 使得在原来范围内不可解的方程, 在扩充的数域中是可解的. 问题是如何扩充实数域使得上述方程可解.

对于上述两个方程, 如果按照通常求解方法, 则方程  $x^2 + 1 = 0$  的解为  $x_{\pm} = \pm\sqrt{-1}$ , 而方程  $x(10 - x) = 40$  的解是  $x_{\pm} = 5 \pm \sqrt{-15}$ . 这就碰到了一个新的问题, 负数如何开平方根?

要使在实数域中没有意义的负数开方变得有意义, 我们引进一个新的“数”

$$i = \sqrt{-1},$$

它服从基本的运算规则：

$$i^2 = -1.$$

引进的这个“数” $i$ 就是一个符号，它无法像其它数一样用来计数，因此称为虚数单位。如果 $i$ 能够与其它实数一样进行加法和乘法运算，这样上述方程的解就分别由新的符号 $\pm i$ 和 $5 \pm i\sqrt{5}$ 表示。类似地，对任意的两个实数 $x, y$ ，就产生了一个新的符号

$$z = x + iy$$

称为复数， $x$ 称为复数 $z$ 的实部， $y$ 称为复数 $z$ 的虚部，并分别记为

$$x = \operatorname{Re}(z), y = \operatorname{Im}(z).$$

两个复数相等，当且仅当两者的实部和虚部分别相等。用记号 $\mathbb{C}$ 表示全部复数的集合：

$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}.$$

对于实部为零的复数我们称为纯虚数，而虚部为零的复数就是实数。

下面简单回顾关于复数的运算。

对任意两个复数

$$z_1 = x_1 + iy_1, z_2 = x_2 + iy_2,$$

它们之间的运算定义如下

$$\begin{aligned} z_1 \pm z_2 &= (x_1 \pm x_2) + i(y_1 \pm y_2) \\ z_1 z_2 &= (x_1 + iy_1)(x_2 + iy_2) = x_1 x_2 + i(x_1 y_2 + x_2 y_1) + i^2 y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1) \end{aligned}$$

因此两个复数相加减还是复数，两个复数相乘还是复数。不难验证这样的运算规则满足加法和乘法的交换律、结合律和分配律。特别

$$(x + iy)(x - iy) = x^2 + y^2$$

是一个实数，我们称复数 $\bar{z} = x - iy$ 为复数 $z = x + iy$ 的共轭复数，记

$$|z|^2 = z\bar{z} = x^2 + y^2.$$

并称 $|z|$ 为 $z$ 的模。同时可以定义零元 $0 = 0 + i0$ ，显然 $z$ 非零当且仅当 $|z| \neq 0$ 。

对任何非零的 $z = x + iy$ ，它的逆定义为

$$z^{-1} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}$$

满足  $zz^{-1} = z^{-1}z = 1$ . 因此两个复数  $z_1 = x_1 + iy_1$ ,  $z_2 = x_2 + iy_2$  的除法为

$$\begin{aligned}\frac{z_1}{z_2} &= \frac{x_1 + iy_1}{x_2 + iy_2} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} \\ &= \left( \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} \right) + i \left( \frac{y_1 x_2 - x_1 y_2}{x_2^2 + y_2^2} \right)\end{aligned}$$

其中  $|z_2| \neq 0$ .

因此复数的集合  $\mathbb{C}$  在加法和乘法运算下是封闭的, 并且具有零元和单位元  $1 = 1 + i0$  以及任何非零元都有逆元. 称  $\mathbb{C}$  为复数域. 因此复数域是实数域  $\mathbb{R}$  的扩充.

## §4.2 复数的几何含义和 Euler 公式

一个复数  $z = x + iy$  就一一对应一组数  $(x, y)$ , 把这组数看成是直角坐标系  $Oxy$  下平面上的点  $P$  的坐标, 这样一个复数就对应平面上的一个点, 就像我们把实数与数轴上的点一一对应一样. 通常也用复数  $z = x + iy$  表示  $Oxy$  平面上的点  $P(x, y)$ . 在这样的对应之下,  $0 = 0 + i0$  对应坐标平面上的原点  $O$ , 实数 (即那些虚部为零的复数) 对应的正是  $x$  轴上的点, 而那些纯虚数 (即实部为零的复数) 对应的正是  $y$  轴上的点. 因此  $z = x + iy$  就成为复数的直角坐标表示. 我们称坐标平面为 **复平面**,  $x$  轴称为**实轴**, 而  $y$  轴称为**虚轴**.

这样就建立了复数  $z = x + iy$  与平面上点  $P(x, y)$ , 进而与平面上向量  $\overrightarrow{OP}$  之间的对应.

复数除了可以用直角坐标表示外, 还可以用极坐标表示. 为此, 取  $x$  轴的正向半轴作为极轴, 坐标原点作为极点, 于是如果点  $z = x + iy$  的极径记作  $r$ , 极角记作  $\theta$ , 那么就有

$$z = x + iy = r(\cos \theta + i \sin \theta).$$

显然, 极径就是复数  $z$  的模, 是被唯一确定的:

$$r = |z| = \sqrt{x^2 + y^2}.$$

极角  $\theta$  称为复数  $z$  的幅角, 用符号  $\operatorname{Arg} z$  表示. 但是当  $z \neq 0$  时,  $z$  的幅角可以相差  $2\pi$  的任何一个整倍数:

$$\theta = \operatorname{Arg} z = \begin{cases} \arctan \frac{y}{x} + 2k\pi & (\text{第一、第四象限}), \\ \arctan \frac{y}{x} + (2k+1)\pi & (\text{第二、第三象限}). \end{cases}$$

这里,  $k$  为任何整数. 函数  $\arctan$  的值域为  $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ . 我们将用  $\arg$  来表示  $\operatorname{Arg}$  的值中的一个值, 比如  $\arg z$  来表示  $\operatorname{Arg} z$  中对应  $k = 0$  的值, 并称  $\arg z$  是  $\operatorname{Arg} z$  的主值.

它的模  $|z|$  就是对应的点到坐标原点  $O$  的距离, 而  $\frac{y}{x}$  就是点到原点的直线的斜率. 记  $|z| = r$  (向径), 横坐标 (即  $x$  轴) 的正向逆时针到复数对应的点和原点之间的直线的夹角为  $\theta$  (幅角), 则复数在极坐标下的表示为

有了复数  $z = x + iy$  与平面上点  $P(x, y)$  以及和向量  $\overrightarrow{OP}$  的对应, 就可以把复数的代数运算赋予几何上的解释.

首先观察复数  $z = x + iy$  的模正是对应点  $P(x, y)$  到原点的距离, 或者说是向量  $\overrightarrow{OP}$  的长度

$$|z| = |\overrightarrow{OP}|.$$

$z$  的共轭复数  $\bar{z} = x - iy$  对应的点是  $P(x, y)$  关于  $x$  轴对称的点  $\bar{P}(x, -y)$ .

两个复数  $z = x + iy$ ,  $z' = x' + iy'$  相减,  $z \pm z'$  对应的点的坐标为  $(x \pm x', y \pm y')$ , 也是向量  $\overrightarrow{OP} \pm \overrightarrow{OP'}$ , 或者说  $z + z'$  表示以  $O$ ,  $z$ ,  $z'$  为三个顶点的平行四边形的第四个顶点, 其模  $|z + z'|$  正是平行四边形从  $O$  到  $z + z'$  的对角线长度, 因此有

$$|z + z'| \leq |z| + |z'|$$

而  $z - z'$  的模

$$|z - z'| = \sqrt{(x - x')^2 + (y - y')^2}$$

表示  $z$  与  $z'$  之间的距离.

用极坐标表示复数, 更容易看出两个复数乘积的几何含义, 设

$$z = r(\cos \theta + i \sin \theta), \quad z' = r'(\cos \theta' + i \sin \theta').$$

那么

$$\begin{aligned} zz' &= rr'[(\cos \theta \cos \theta' - \sin \theta \sin \theta') + i(\cos \theta \sin \theta' + \sin \theta \cos \theta')] \\ &= rr'[\cos(\theta + \theta') + i \sin(\theta + \theta')]. \end{aligned}$$

因此两个复数乘积  $zz'$  的模是两个复数  $z$  和  $z'$  模的乘积  $|zz'| = |z||z'|$ ,  $zz'$  的幅角正是  $z$  和  $z'$  幅角的和. 换句话说复数  $z'$  乘以复数  $z$  就是把  $z$  逆时针旋转  $\theta'$ , 长度乘以  $r'$  得到的复数.

$$\begin{aligned} \frac{z}{z'} &= z \frac{\overline{z'}}{|z'|^2} = \frac{r}{r'}[(\cos \theta \cos \theta' + \sin \theta \sin \theta') + i(-\cos \theta \sin \theta' + \sin \theta \cos \theta')] \\ &= \frac{r}{r'}[\cos(\theta - \theta') + i \sin(\theta - \theta')]. \end{aligned}$$

一般情况下, 对任意三个互不相等的复数  $z_1, z_2, z_3$ , 有

$$z_3 - z_1 = (z_2 - z_1)r(\cos \theta + i \sin \theta),$$

这里  $\theta$  是以  $z_1, z_2, z_3$  为顶点的三角形中  $z_1$  的角,

$$r = \frac{|z_3 - z_1|}{|z_2 - z_1|}.$$

如果考虑复数  $z$  反复相乘

$$z^2 = r^2(\cos 2\theta + i \sin 2\theta),$$

以及

$$z^n = r^n(\cos n\theta + i \sin n\theta),$$

这里  $n$  是任意正整数. 将  $z = r(\cos \theta + i \sin \theta)$  代入上式左边, 并消去  $r^n$ , 我们最终得到著名的 De Moivre (棣莫弗 1667-1754) 公式:

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta.$$

De Moivre 公式一个深刻的含义是一个模长为 1 的复数, 它的任何  $n$  倍的复数的幅角是原复数幅角的  $n$  倍. 在此基础上, Euler 给出了下列著名的公式:

$$e^{i\theta} = \cos \theta + i \sin \theta$$

特别取  $\theta = \pi$ , Euler 公式就给出了

$$e^{i\pi} + 1 = 0$$

这个公式把人们最常见的几个数 0, 1,  $\pi$ ,  $e$  以及虚数单位  $i$  联系在一起.

这里我们无法给出 Euler 公式严格证明, 只能从形式上给出解释.

在数学分析或微积分中, 我们会给出  $e^x$  以及  $\sin x, \cos x$  的 Taylor (泰勒 1685-1731) 展开式. 所谓 Taylor 展开式无非是将上述三个函数展开成一个无穷次数的“多项式”(称为幂级数)

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \\ \cos x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{2n!} + \dots \\ \sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots \end{aligned}$$

虽然上述展开式是对实数  $x$  而言, 但如果形式上把实数换成纯虚数  $x = i\theta$ , 代入  $e^x$  的表达式, 并借助  $i$  的运算规则

$$i^2 = -1, i^3 = -i, i^4 = 1, \dots$$

不难发现

$$\begin{aligned} e^{i\theta} &= \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} = \left( 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots + (-1)^m \frac{\theta^{2m}}{(2m)!} + \cdots \right) \\ &\quad + i \left( \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots + (-1)^m \frac{\theta^{2m+1}}{(2m+1)!} + \cdots \right) \\ &= \cos \theta + i \sin \theta. \end{aligned}$$

这里, 我们只给出了 Euler 公式一个形式上的“证明”, 并没有用到有关复变函数的任何知识. 利用 Euler 公式, 对于任何复数  $x + iy$ , 设  $r$  是它的模长,  $\theta$  是幅角, 则复数可以表示成

$$z = x + iy = r(\cos \theta + i \sin \theta) = r e^{i\theta}$$

因此, 也可用  $e^{i\theta}$  分别表示  $\cos \theta$  和  $\sin \theta$ :

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

这里我们给出 Euler 公式一个简单应用.

**例 4.2.1** 设  $z = e^{i\theta} = \cos \theta + i \sin \theta$ , 因此

$$z^k = e^{ik\theta} = \cos k\theta + i \sin k\theta, \quad k = 1, 2, \dots,$$

对  $k$  求和得

$$\begin{aligned} \sum_{k=1}^n e^{ik\theta} &= \frac{e^{i\theta} - e^{i(n+1)\theta}}{1 - e^{i\theta}} = \frac{e^{i\frac{\theta}{2}} - e^{i(n+\frac{1}{2})\theta}}{e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}} \\ &= \frac{[\cos \frac{\theta}{2} - \cos (n + \frac{1}{2})\theta] + i [\sin \frac{\theta}{2} - \sin (n + \frac{1}{2})\theta]}{-2i \sin \frac{\theta}{2}} \end{aligned}$$

因此, 取实部, 得

$$\sum_{k=1}^n \cos k\theta = \frac{\sin (n + \frac{1}{2})\theta - \sin \frac{\theta}{2}}{2 \sin \frac{\theta}{2}}$$

即

$$\frac{1}{2} + \sum_{k=1}^n \cos k\theta = \frac{\sin (n + \frac{1}{2})\theta}{2 \sin \frac{\theta}{2}}, \quad \theta \neq 2l\pi.$$

当然上述恒等式也可以直接证明. 利用三角函数的积化和差, 我们有

$$\begin{aligned} 2 \sin \frac{\theta}{2} \left( \frac{1}{2} + \sum_{k=1}^n \cos k\theta \right) &= \sin \frac{\theta}{2} + \sum_{k=1}^n \left[ \sin \left( k + \frac{1}{2} \right) \theta - \sin \left( k - \frac{1}{2} \right) \theta \right] \\ &= \sin \left( n + \frac{1}{2} \right) \theta. \end{aligned}$$

### §4.3 代数基本定理

现在重新回到关于代数方程求解的问题. 引进复数以后, 是不是仅仅解决如方程  $x^2 + 1 = 0$  或  $x(10 - x) = 40$  的求解问题? 对其它二次, 乃至更高次数的代数方程在复数范围内是否也可以求解?

首先考虑一般的实系数 ( $a, b, c$  均为实数, 且  $a \neq 0$ ) 二次代数方程

$$ax^2 + bx + c = 0,$$

根据求根公式, 方程的两个解 (有时也称为方程的两个根) 为

$$x_{\pm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

当  $\Delta = b^2 - 4ac > 0$  时, 方程有两个实数根.

当  $\Delta = b^2 - 4ac = 0$  时, 方程有实的重根  $x = -\frac{b}{2a}$ .

当  $\Delta = b^2 - 4ac < 0$  时, 方程没有实数根, 也就是在实数范围内方程没有解. 但在复数范围内方程有两个互为共轭的复数根

$$x_{\pm} = \frac{-b}{2a} \pm i \frac{\sqrt{|b^2 - 4ac|}}{2a}.$$

因此, 引进复数后, 对任何实系数二次方程在复数域范围内都是可解的.

我们自然希望像求解二次代数方程那样, 利用系数之间的加、减、乘、除等四则运算和开平方根、开立方根、开四方根等得到更高次数的代数方程的根. 然而, 这种“用根式”的求解方法到了五次及以上次数的代数方程就行不通了. 事实上, Ruffini (鲁菲尼 1756- 1822) 和 Abel (阿贝尔 1802 ~ 1829) 证明了不可能用根式的方法解一般的  $n$  代数方程. 这一结果进一步促成了新的数学的重大发展. 具体内容这里不做过多介绍. 需要指出的是, 五次及以上的代数方程的求根问题, 虽然不能像二次代数方程那样用开方的方法求根, 但并不表示高次代数方程的根不存在.

事实上, 我们有下列著名的代数基本定理:

**定理 4.1** 下列实系数或复系数的  $n$  次代数方程

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0.$$

在复数范围内 (当然包含了实数) 至少存在一个根. 这里  $a_{n-1}, \dots, a_0$  为实数或复数. 不失一般性, 我们总假设首项系数为 1.

由代数基本定理, 可以得到得到

**推论 4.2** 任何一个多项式

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

可以分解为  $n$  个因式的乘积

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

其中  $\alpha_1, \alpha_2, \dots, \alpha_n$  是复数. 显然它们是  $n$  次代数方程  $f(x) = 0$  的根. 或者说任何一个  $n$  次代数方程一定有  $n$  个根(重根的重数计作根的个数).

关于代数基本定理的证明, 已经超出了本专题的范围, 在此不再介绍. 这里着重强调这样一件事情, 虽然  $n(n \geq 5)$  次代数方程的根无法像二次代数方程那样通过系数之间的四则运算和开根号求出, 但代数基本定理告诉我们,  $n$  次代数方程在复数范围内存在  $n$  个根. 并不需要随着  $n$  的不同去创造其它什么新的数.

由上述推论, 还可以证明下列结果(读者可作为习题自证).

**推论 4.3** 如果

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

是实系数多项式(即  $a_{n-1}, \dots, a_0$  均为实数), 那么

1、若  $z_0 = x_0 + iy_0$  是方程  $f(x) = 0$  的复数根, 则它的共轭  $\bar{z}_0 = x_0 - iy_0$  也是方程的根, 也就是实系数代数方程的复根成对( $z_0, \bar{z}_0$ )出现.

2、实系数多项式一定能够分解为下列形式

$$f(x) = (x - \alpha_1)^{r_1} \cdots (x - \alpha_k)^{r_k} (x^2 + \beta_1x + \gamma_1)^{s_1} \cdots (x^2 + \beta_lx + \gamma_l)^{s_l},$$

这里  $\alpha_i, \beta_j, \gamma_j$  都是实数, 满足

$$\beta_j^2 - 4\gamma_j < 0, \quad j = 1, 2, \dots, l.$$

$r_j, j = 1, \dots, k$  是方程实根的重数,  $s_j, j = 1, \dots, l$  是成对出现的复根的重数, 它们满足

$$r_1 + \cdots + r_k + 2s_1 + \cdots + 2s_l = n.$$

## §4.4 单位根

**定义 4.4** 一个复数  $\xi$  称为(复)单位根, 如果存在某个正整数  $n$ , 使  $\xi$  满足方程

$$z^n - 1 = 0.$$

对固定的  $n$ , 上述方程的  $n$  个根记为  $\xi_1, \xi_2, \dots, \xi_n$ , 称为  $n$  次单位根.

例如当  $n = 2$  时, 方程  $z^2 - 1 = 0$  有两个 2 次单位根:  $\xi_1 = -1, \xi_2 = 1$ .

当  $n = 3$  时, 方程  $z^3 - 1 = 0$  有三个 3 次单位根:

$$\begin{aligned}\xi_1 &= \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} = -\frac{1}{2} + i \frac{\sqrt{3}}{2}, \\ \xi_2 &= \cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3} = -\frac{1}{2} - i \frac{\sqrt{3}}{2}, \\ \xi_3 &= 1.\end{aligned}$$

当  $n = 4$  时, 方程  $z^4 - 1 = 0$  有四个 4 次单位根:

$$\begin{aligned}\xi_1 &= \cos \frac{2\pi}{4} + i \sin \frac{2\pi}{4} = i \\ \xi_2 &= \cos \frac{4\pi}{4} + i \sin \frac{4\pi}{4} = -1, \\ \xi_3 &= \cos \frac{6\pi}{4} + i \sin \frac{6\pi}{4} = -i, \\ \xi_4 &= 1.\end{aligned}$$

一般情况下, 根据 De Moivre 公式, 方程  $z^n - 1 = 0$  有  $n$  个互不相等的单位根:

$$\xi_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \quad k = 1, 2, \dots, n.$$

在复平面上, 这些单位根对应的点恰好把单位元  $|z| = 1$  分成  $n$  等分, 其中一个分点为  $\xi_n = 1$ . 我们把互不相等的  $n$  次单位根的集合记为

$$A_n = \{\xi_1, \xi_2, \dots, \xi_n\}.$$

注意到在 4 次单位根中,  $\xi_2 = -1$  实际上也是 2 次单位根:  $\xi_2^2 = 1$ , 但  $\xi_1, \xi_3$  不是. 自然要问, 在  $A_n$  中, 哪些  $n$  次的单位根同时也是更低次数的单位根, 哪些不是?

**定义 4.5** 设  $\xi$  是单位根, 若  $d$  是使得  $\xi^d = 1$  成立的最小正整数, 则称  $d$  为  $\xi$  的阶. 若  $\xi \in A_n$  且  $\xi$  的阶恰好是  $d = n$ , 则称  $\xi$  是本原的  $n$  次单位根.

根据这个定义, 在 4 次单位根的集合  $A_4$  中,  $\xi_2$  的阶是 2,  $\xi_4$  的阶是 1, 而  $\xi_1, \xi_3$  的阶是 4, 因此是本原的. 在  $A_3$  中, 除了  $\xi_3$  外,  $\xi_1, \xi_2$  都是本原的.

下面的任务是在单位根中区分哪些是本原的, 并讨论本原单位根的性质.

**定理 4.6** 设单位根  $\xi$  的阶是  $d$ .

(1)  $\xi$  是  $n$  次单位根当且仅当  $d|n$ .

(2) 对于整数  $k$ ,  $\xi^k$  的阶仍为  $d$  当且仅当  $k$  与  $d$  互素:  $(k, d) = 1$ .

**证明** (1) 的证明: 若  $\xi^n = 1$ , 显然  $n \geq d$ , 令  $n = qd + r$ ,  $0 \leq r < d$ . 由  $\xi^d = 1$  得

$$1 = \xi^n = \xi^{qd+r} = (\xi^d)^q \xi^r = \xi^r,$$

又因为  $d$  是使得  $\xi^d = 1$  成立的最小正整数, 所以  $r = 0$ , 即  $d|m$ . 反之显然.

(2) 的证明: 设  $\xi^k$  的阶为  $d$ . 若  $(k, d) = c > 1$ , 则存在  $k' < k, d' < d$  使得  $k = k'c, d = d'c$ . 所以

$$(\xi^k)^{d'} = \xi^{k'd'c} = (\xi^d)^{k'} = 1,$$

但  $d' < d$ , 这与  $\xi^k$  的阶是  $d$  矛盾, 因此  $(k, d) = 1$ .

设  $\xi^k$  的阶是  $d'$ , 因为  $(\xi^k)^d = (\xi^d)^k = 1$ , 所以  $\xi^k$  的阶  $d'$  满足  $d' \leq d$ .

若  $d' < d$ , 根据定理4.6, 从  $\xi^{kd'} = 1$  推出  $d|kd'$ , 而  $(k, d) = 1$ , 所以  $d|d'$ , 这与  $d' < d$  矛盾, 因此  $d' = d$ .  $\square$

**定理 4.7** 若  $\xi$  是一个本原的  $n$  次单位根, 则  $A_n$  中所有  $n$  次单位根可由  $\xi$  的幂次生成:

$$A_n = \{\xi^0, \xi, \xi^2, \dots, \xi^{n-1}\}$$

反之, 若  $A_n$  中的单位根都可由一个  $n$  次单位根  $\xi$  按上述方式生成, 则  $\xi$  一定是本原的.

**证明** 显然,  $n$  个数  $\xi^0, \xi, \xi^2, \dots, \xi^{n-1}$  都是  $n$  次单位根, 因此只要证明它们互不相同. 假如有

$$\xi^i = \xi^j \quad (0 \leq i < j \leq n-1),$$

则  $\xi^{j-i} = 1$ , 但  $0 < j-i < n$ , 这与  $\xi$  是  $n$  次本原单位根矛盾.

设  $A_n$  由  $n$  次单位根  $\xi$  按上述方式生成  $A_n = \{\xi^0, \xi, \xi^2, \dots, \xi^{n-1}\}$ , 记  $\xi$  的阶为  $d$ . 若  $d < n$ , 那么  $\xi$  的任何幂次  $\xi^i$  都满足  $(\xi^i)^d = 1$ , 也就是它们都是  $d$  次单位根. 即  $A_n \subset A_d$ , 这是不可能的, 因为  $A_d$  中只有  $d$  个元素.  $\square$

对方程  $z^n - 1 = 0$  来说, 它的  $n$  次单位根中

$$\xi = \xi_1 = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$$

是本原的, 因为  $(1, n) = 1$ . 因此  $n$  次单位根的集合可以表示为

$$A_n = \{\xi^0, \xi, \xi^2, \dots, \xi^{n-1}\}.$$

在这个集合中, 根据定理4.7, 只要  $(k, n) = 1$ , 则  $\xi^k$  就是本原的, 因此,  $A_n$  中本原的单位根的个数正是  $1, 2, \dots, n$  中与  $n$  互素的数的个数, 这个数就是 Euler 函数  $\varphi(n)$  (见第二讲定义2.15!) 这样, 本原的  $n$  次单位根的集合为

$$B_n = \{\xi^k \mid 1 \leq k \leq n, (k, n) = 1\}$$

每个本原的  $n$  次单位根  $\xi^k \in B_n$ , 都可通过幂次生成所有的  $n$  次单位根

$$A_n = \{1, \xi^k, \xi^{2k}, \dots, \xi^{(n-1)k}\}, \quad \xi^k \in B_n.$$

方程  $z^n - 1 = 0$  的  $n$  个单位根的集合  $A_n$  具有下列性质:

(1) 在复数的乘法运算下是封闭的, 即对  $\xi_i, \xi_j \in A_n$ , 有

$$\xi_i \xi_j = \xi_k \in A_n, \quad k : (i + j) \equiv k \pmod{n},$$

(2) 有单位元  $\xi_n = 1 \in A_n$ , 并对任何  $\xi_i \in A_n$  有逆元

$$\xi_i^{-1} = \xi_{n-i} \in A_n.$$

因此形成一个乘法群. 而  $A_n$  的所有元素可用一个本原的  $n$  次根生成, 例如取  $\xi = \xi_1$ ,

$$A_n = \{\xi^0, \xi, \xi^2, \dots, \xi^{n-1}\}.$$

因此这样的群为循环群.

因为  $A_n$  中的数对应复平面中单位圆上的  $n$  个等分点, 其中  $\xi_n = 1$  在实轴上. 因此, 从几何上看, 即是从实轴出发, 沿单位圆每次旋转  $\frac{2\pi}{n}$ , 则可跑遍所有的等分点. 若对本原的  $\xi^k$ , 因为  $(k, n) = 1$ , 所以每次旋转  $\frac{2k\pi}{n}$ , 也能跑遍所有的等分点.

继续考虑方程  $z^n - 1 = 0$ , 利用

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

不难看出除 1 以外的  $n$  次单位根  $\xi, \xi^2, \dots, \xi^{n-1}$  满足方程

$$x^{n-1} + x^{n-2} + \dots + 1 = 0,$$

我们称这个方程为分圆方程. 例如 1 的三次复根

$$\begin{aligned} \xi &= \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} = \frac{1}{2}(-1 + i\sqrt{3}), \\ \xi^2 &= \cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3} = \frac{1}{2}(-1 - i\sqrt{3}). \end{aligned}$$

是分原方程

$$x^2 + x + 1 = 0$$

的根. 同样地, 1 的五次根除 1 以外都满足分圆方程

$$x^4 + x^3 + x^2 + x + 1 = 0.$$

因此要想用直尺和圆规作出一个正五边形, 我们必须求解这个方程. 为此作一些简单的代数变换, 先在方程两边除以  $x^2$  得

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

再作变换

$$z = x + \frac{1}{x},$$

方程就化为一个 2 次方程

$$z^2 + z - 1 = 0.$$

求出它的两个根为

$$z_1 = \frac{-1 + \sqrt{5}}{2}, \quad z_2 = \frac{-1 - \sqrt{5}}{2}.$$

再分别求解

$$x + \frac{1}{x} = \frac{-1 \pm \sqrt{5}}{2}$$

就得到 1 的五次复根.

## §4.5 复变数的函数

在第 2 节中, 我们实际上已经把一个常见的实变数的函数  $f(x) = e^x$  换成了复变数的函数  $f(z) = e^z$ , 只不过限制  $z$  为纯虚数  $z = i\theta$ . 那么是不是可以把定义在实数范围内的初等函数的定义域都换成复数的某个范围. 下面, 我们通过一些具体的例子可以看到, 事情并不是那么简单, 即使可以定义类似的复变数函数, 函数的性质也发生了变化.

一般来说, 所谓复变数函数是指从复数的某个范围内  $D \subset \mathbb{C}$  到复数的一个映射

$$f : D \subset \mathbb{C} \longrightarrow \mathbb{C},$$

或写成

$$f : z \in D \longmapsto w \in \mathbb{C}$$

记为  $w = f(z)$ . 如果对于每一个  $z \in D$ , 有唯一的复数  $w$  与之对应, 则称映射(或函数)是单值的. 如果有多个  $w$  与之对应, 则称映射时多值的.

如果对任意两个  $z_1, z_2 \in D$ , 对应两个不同的  $w_1, w_2$ , 则称映射在  $D$  上是一一的, 或单叶的.

以下, 我们仅讨论几个具体的例子.

### 1° 幂函数

首先讨论  $w = z^n$ , 这里  $n$  是正整数.

这是  $n$  个相同的复数  $z$  相乘, 其结果还是一个复数, 因此是单值的. 如果令

$$z = r(\cos \theta + i \sin \theta), \quad w = \rho(\cos \varphi + i \sin \varphi),$$

那么

$$\rho = r^n, \quad \varphi = n\theta.$$

从上式不难看出, 两个模相等的复数  $z_1, z_2$  映射到同一个点 (即  $z_1^n = z_2^n$ ), 当且仅当它们的幅角相差  $\frac{2\pi}{n}$  的整数倍. 因此要使映射  $w = z^n$  是单叶的, 其充分必要条件是定义域  $D$  中任意两个点  $z_1, z_2$ , 都不满足下列等式

$$|z_1| = |z_2|, \arg z_1 - \arg z_2 = \frac{2k\pi}{n}, k \neq 0 \text{ 是整数.}$$

例如, 扇形区域

$$D = \{z \mid 0 < \arg z < \frac{2\pi}{n}\}$$

就是一个能保证映射  $w = z^n$  的单叶性. 这个扇形区域被映射  $w = z^n$  一一映射到

$$\tilde{D} = \{w \mid 0 < \arg w < 2\pi\}$$

即映到去掉正半轴的复平面.

其次讨论

$$w = \sqrt[n]{z} = z^{\frac{1}{n}}.$$

为了说明问题, 我们只考虑  $w = \sqrt{z} = z^{\frac{1}{2}}$  的情况.

设  $z = r(\cos \theta + i \sin \theta)$ , 则

$$z \longrightarrow z^{\frac{1}{2}} = \begin{cases} r^{\frac{1}{2}} \left( \cos \frac{\theta}{2} + i \sin \frac{\theta}{2} \right) = w \\ r^{\frac{1}{2}} \left( \cos \frac{\theta+2\pi}{2} + i \sin \frac{\theta+2\pi}{2} \right) = -w. \end{cases}$$

因此,  $w = z^{\frac{1}{2}}$  是一个多值函数.

进一步分析发现, 当  $z$  绕一个不包含原点的封闭曲线旋转一圈并回到原处, 那么  $z$  的幅角不会发生改变, 如果  $z$  绕包含原点的封闭曲线旋转一圈并回到原处, 我们发现幅角  $\theta \rightarrow \theta + 2\pi$ . 此时  $w = z^{\frac{1}{2}}$  的函数值会产生如下变化

$$\begin{aligned} w &= r^{\frac{1}{2}} \left( \cos \frac{\theta}{2} + i \sin \frac{\theta}{2} \right) \\ &\longrightarrow r^{\frac{1}{2}} \left( \cos \frac{\theta+2\pi}{2} + i \sin \frac{\theta+2\pi}{2} \right) \\ &\longrightarrow -r^{\frac{1}{2}} \left( \cos \frac{\theta}{2} + i \sin \frac{\theta}{2} \right) = -w. \end{aligned}$$

为了避免多值性, 我们把函数  $w = z^{\frac{1}{2}}$  定义域所在的复平面  $\mathbb{C}$  从原点 0 到无穷  $\infty$  剪开一个口子, 例如, 沿  $x$  轴的正半轴剪开, 这样使得在剪开后的定义域内无法形成任何包含原点的封闭曲线, 也就避免了出现多值性. 或者说如果限制  $z$  的幅角取值在  $0 \leq \theta < 2\pi$  范围内, 就可以保证函数的单值性.

最后讨论

$$w = z^{-1} = \frac{1}{z}.$$

显然该映射的定义域为  $z \neq 0$ . 因为  $|w| = \frac{1}{|z|}$ , 所以映射把  $z$  所在的复平面上单位圆盘内除 0 以外的点映射到  $w$  所在的平面的圆盘外的点. 如果在闭的复平面上考虑, 那么映射把 0 映射到  $\infty$ .

### 2° 指数函数

这里, 我们仍然不加证明地引进指数函数

$$w = e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y),$$

那么, 它满足:

$$e^{z_1} e^{z_2} = e^{z_1+z_2},$$

并且是以  $2\pi i$  为周期的周期函数

$$e^{z+2\pi i} = e^z.$$

和通常一样, 对数函数定义为指数函数的反函数. 我们把满足方程

$$e^w = z (z \neq 0)$$

的函数  $w = f(z)$  称为对数函数. 令  $w = u + iv$ ,  $z = re^{i\theta}$ , 那么

$$e^{u+iv} = re^{i\theta},$$

所以  $u = \ln r$ ,  $v = \theta$ , 因此

$$w = \ln |z| + i \operatorname{Arg} z.$$

由于  $\operatorname{Arg} z$  为多值函数, 所以对数函数  $w = f(z)$  是多值函数, 并且每两个值相差  $2\pi i$  的整数倍. 如果规定  $\operatorname{Arg} z$  取主值  $\operatorname{arg} z$ , 那么对数函数就是一个单值函数, 记为

$$\ln z = \ln |z| + i \operatorname{arg} z.$$

### 3° 三角函数

借助指数函数, 我们定义

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}$$

为复数  $z$  的余弦函数和正弦函数. 其他三角函数可按实变数三角函数同样的方式定义. 由指数函数的性质, 读者不难验证这样定义的复变数的三角函数满足周期性、奇偶性以及积化和差与和差化积的公式.

### 4° 双曲函数

类似复变数三角函数的定义, 我们还可以定义下列复变数的双曲函数:

$$\cosh z = \frac{e^z + e^{-z}}{2}, \quad \sinh z = \frac{e^z - e^{-z}}{2}$$

它们也具有实变数的双曲函数所满足的性质.

### 5° Rokovsky (茹科夫斯基 1847-1921) 函数

$$w = f(z) = \frac{1}{2} \left( z + \frac{1}{z} \right) \quad (z \neq 0).$$

首先讨论该函数的单叶性. 若有两个  $z_1 \neq z_2$ , 使得  $f(z_1) = f(z_2)$ , 那么

$$(z_1 - z_2) \left( 1 - \frac{1}{z_1 z_2} \right) = 0,$$

因此推出

$$z_1 z_2 = 1.$$

所以在任何区域  $D$  内, 映射  $w = f(z)$  是单叶映射的充分必要条件是  $D$  中没有任何两个点  $z_1, z_2$  满足关系  $z_1 z_2 = 1$ . 例如使得映射保持单叶性的区域有单位圆内部  $|z| < 1$ , 或它的外部  $|z| > 1$ .

令  $z = r(\cos \theta + i \sin \theta)$ ,  $w = u + iv$ , 则映射可以表示为

$$u = \frac{1}{2} \left( r + \frac{1}{r} \right) \cos \theta, \quad v = \frac{1}{2} \left( r - \frac{1}{r} \right) \sin \theta$$

的形式.

如果取定  $r = r_0 < 1$ , 那么  $|z| = r_0$  给出单位圆内部一个同心圆. 令

$$a = \left( r_0 + \frac{1}{r_0} \right), \quad b = \left( \frac{1}{r_0} - r_0 \right)$$

则

$$u = a \cos \theta, \quad v = -\sin \theta, \quad \text{或} \quad \frac{u^2}{a^2} + \frac{v^2}{b^2} = 1$$

是  $w$  所在复平面上一个椭圆. 若  $z$  所在复平面上的点沿圆  $|z| = r_0$  从  $\theta = 0$  起步逆时针转动, 那么  $w$  所在复平面上对应的点从  $w = a$  出发顺时针转动.

当  $r_0 \rightarrow 1$  时,  $w$  所在平面上的椭圆压缩成  $u$  轴上一段线段  $[-1, 1]$ . 当  $r_0 \rightarrow 0$  时, 椭圆趋向无穷, 所以茹科夫斯基函数把 0 映到  $\infty$ .

如果取定  $\theta = \theta_0$ , 那么  $z = r(\cos \theta_0 + i \sin \theta_0)$ ,  $0 < r < 1$  是  $z$  所在平面上一段射线段, 在茹科夫斯基函数映射下对应  $w$  所在复平面上的双曲线

$$\frac{u^2}{\cos^2 \theta_0} - \frac{v^2}{\sin^2 \theta_0} = 1.$$

## 第4讲习题

1. 计算  $\sqrt{5 + i12}$ .

提示: 令  $\sqrt{5 + i12} = x + iy$ , 平方后比较实部和虚部.

2. 根据复数运算规则, 证明: 对任意的两个复数  $z_1, z_2$ , 有

$$|z_1| - |z_2| \leq |z_1 - z_2|$$

并给出等号相等的充要条件.

3. 设四个任意复数  $z_1, z_2, z_3, z_4$ , 证明: 这四个数同在一个圆上或一条直线上当且仅当  $\frac{z_4 - z_1}{z_4 - z_2}$  和  $\frac{z_3 - z_1}{z_3 - z_2}$  有相同的幅角, 或者说

$$\frac{z_4 - z_1}{z_4 - z_2} / \frac{z_3 - z_1}{z_3 - z_2}$$

是实数.

4. 利用复数的极坐标表示, 证明余弦公式

$$c^2 = a^2 + b^2 - 2ab \cos \theta.$$

其中  $a, b, c$  分别是三角形的三条边的长度,  $\theta$  是边  $c$  的对应角.

5. 列出所有的 12 次本原单位根.

6. 设  $z$  是任意的复数, 试证: 对任意的正整数  $n$ ,  $z^n + \frac{1}{z^n}$  可以表示为  $w = z + \frac{1}{z}$  的  $n$  次多项式.

7. 证明茹科夫斯基函数

$$f(z) = \frac{1}{2} \left( z + \frac{1}{z} \right),$$

当  $|z| = 1$  时,  $f(z) \in [-1, 1]$ .

8. 求下列函数的最大值

$$f(x) = \sqrt{x^4 - 3x^2 - 6x + 13} - \sqrt{x^4 - x^2 + 1}$$

提示: 首先将两个根号内部表示成平方和的形式

$$f(x) = \sqrt{(x-3)^2 + (x^2-2)^2} - \sqrt{x^2 + (x^2-1)^2}$$

并令

$$z_1 = (x-3) + i(x^2-2), \quad z_2 = x + i(x^2-1),$$

这样

$$f(x) = |z_1| - |z_2| \leq |z_1 - z_2| = \sqrt{10}.$$

再利用第 2 题的结果.

9. 设  $D$  是锐角三角形  $\triangle ABC$  内部一点,  $\angle ADB = \angle ACB + 90^\circ$ , 并且

$$AC \cdot BD = AD \cdot BC.$$

求

$$\frac{AB \cdot CD}{AC \cdot BD}$$

的值.

提示: 如图所示

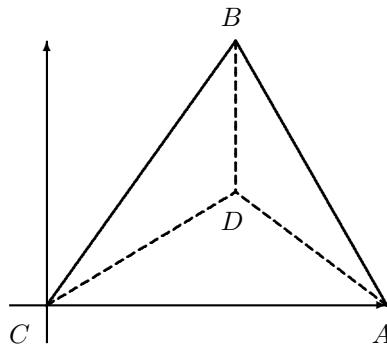


图 4.1

分别记点  $A, B, D$  对应的复数为  $z_A, z_B, z_D$ , 那么题目的条件为

$$|z_A||z_B - z_D| = |z_B||z_A - z_D|,$$

要证

$$\frac{|z_A - z_B||z_D|}{|z_A||z_B - z_D|}.$$

利用

$$(z_B - z_D) = (z_A - z_D)r(\cos \theta + i \sin \theta),$$

其中

$$r = \frac{|z_B - z_D|}{|z_A - z_D|}, \quad \theta = \angle ACB + 90^\circ.$$

并注意到

$$\cos \theta + i \sin \theta = i(\cos \angle ACB + i \sin \angle ACB) = \frac{1}{|z_B|}z_B,$$

即可征得结果.

## 第 5 讲 解析几何与向量空间

解析几何是几何与代数之间的桥梁, 通过引进坐标, 使得一些几何问题可以用代数方法来研究. 首先对平面解析几何做一个简要回顾.

在平面上取定一个点  $O$  (称为原点)、交于该点的相互垂直的有向数轴, 就构成了平面坐标系. 平面上任何一点  $P$  都唯一地对应一个数组  $(x, y)$ , 称为  $P$  点的坐标. 也可以说, 坐标系的建立, 使得平面上的点完全“数字化”了.

平面上任何两点  $P_1(x_1, y_1)$ ,  $P_2(x_2, y_2)$  之间的距离也可由点的坐标之间的代数运算给出:

$$\rho(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

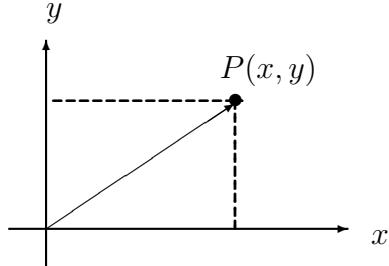


图 5.1

在坐标系中, 可以把一些几何图形用代数方程表示. 例如, 椭圆上任何一点到两定点距离等于常数. 不妨取坐标系使得两定点的坐标分别为  $F_1(p, 0)$  和  $F_2(-p, 0)$ . 那么, 椭圆上任何一点  $P(x, y)$  到  $F_1(p, 0)$  和  $F_2(-p, 0)$  的距离为

$$r_1 = |PF_1| = \sqrt{(x - p)^2 + y^2}, \quad r_2 = |PF_2| = \sqrt{(x + p)^2 + y^2},$$

如果  $r_1 + r_2 = 2a$  为常数 (显然  $a > p$ , 否则该问题无解), 那么有

$$r_2^2 - r_1^2 = \begin{cases} (x + p)^2 - (x - p)^2 = 4px, \\ (r_2 + r_1)(r_2 - r_1) = 2a(r_2 - r_1) \end{cases}$$

因此得到  $r_2 - r_1 = \frac{2px}{a}$ , 与  $r_2 + r_1 = 2a$  联立得

$$r_1 = a - \frac{px}{a}, \quad r_2 = a + \frac{px}{a}.$$

把  $r_1$  代入  $r_1^2 = (x - p)^2 + y^2$ :

$$\left(a - \frac{px}{a}\right)^2 = (x - p)^2 + y^2.$$

记  $b = \sqrt{a^2 - p^2}$ , 整理后即可得椭圆上任何一点  $P(x, y)$  的坐标满足的代数方程

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

其中  $a > b > 0$  分别称为椭圆的长半轴和短半轴,  $F_1(p, 0)$  和  $F_2(-p, 0)$  称为椭圆的焦点.

## §5.1 向量和及其代数运算

为了更好地用代数表示并解决几何问题, 需要借助“向量”的概念. 所谓向量来源于物理学, 一些物理量不仅有大小, 还有方向. 例如位移、力、速度、加速度等等, 抛开它们的物理意义, 只保留大小和方向两个要素, 就抽象为数学中向量概念. 因此, 向量就是既有大小, 又有方向的量.

向量可以用平面或空间中有向线段来表示. 设  $\overline{AB}$  为连接  $A, B$  两点的线段, 它的长度等于  $a$  的大小, 从  $A$  指向  $B$  的方向与  $a$  的方向相同, 我们就可以用  $\overrightarrow{AB}$  表示向量  $a$ , 记为  $a = \overrightarrow{AB}$ .

两个向量, 只要大小相等、方向相同, 两者就是相等的, 因此相等的向量可以通过不改变大小和方向的平移使之完全重合. 通过平移, 可以把任何向量  $a$  等同于起点在原点的一个向量  $a = \overrightarrow{OA}$ .

有三种向量比较特别, 大小为 0 的零向量; 大小等于 1 的向量称为单位向量; 与向量  $a$  大小相等, 但方向相反的向量, 称为  $a$  的负向量(或反向量), 记为  $-a$ .

### 1° 向量的加法和数乘

设  $a, b$  是两个向量, 用同一起点  $O$  的有向线段表示它们:  $a = \overrightarrow{OA}, b = \overrightarrow{OB}$ , 则以  $\overrightarrow{OA}, \overrightarrow{OB}$  为邻边的平行四边形的对角线向量  $c = \overrightarrow{OC}$  就称为这两个向量的和, 记作

$$\overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{OB} \text{ 或简写成 } c = a + b.$$

这种求和的方法称为平行四边形法则(图 5.4). 也可以用三角形法则(图 5.5), 即以  $a$  的终点  $A$  为起点, 做一有向线段使其等于  $b$ :  $\overrightarrow{AB} = b$ , 则以  $O$  为起点,  $B$  为终点的向量  $\overrightarrow{OB}$  就是  $a$  与  $b$  的和

$$a + b = \overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OB}.$$

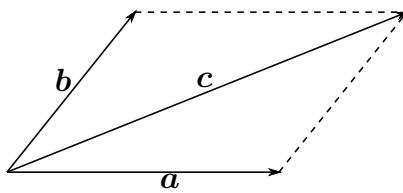


图 5.2

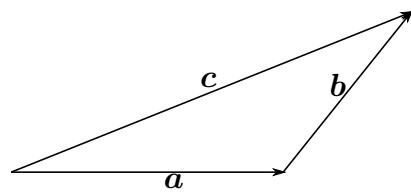


图 5.3

设  $a = \overrightarrow{OA}$ ,  $b = \overrightarrow{OB}$ , 则  $-b = \overrightarrow{BO}$ , 因此

$$a - b = a + (-b) = \overrightarrow{OA} + (-\overrightarrow{OB}) = \overrightarrow{OA} + \overrightarrow{BO} = \overrightarrow{BA}$$

**性质 5.1** 向量加法满足如下性质:

1.  $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ , (交换律)
2.  $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ , (结合律)
3.  $\mathbf{a} + \mathbf{0} = \mathbf{a}$ ,  $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$ .

**证明** 从向量求和的平行四边形法则容易看出交换律成立. 而用三角形法则较易证明结合律(图 5.6). 设  $\mathbf{a} = \overrightarrow{OA}$ ,  $\mathbf{b} = \overrightarrow{AB}$ ,  $\mathbf{c} = \overrightarrow{BC}$ , 则  $\mathbf{a} + \mathbf{b} = \overrightarrow{OB}$ ,  $\mathbf{b} + \mathbf{c} = \overrightarrow{AC}$ , 所以

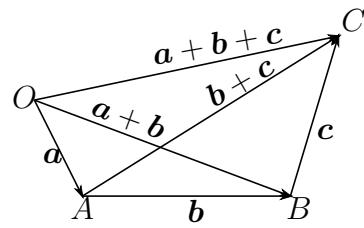


图 5.4

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \overrightarrow{OB} + \overrightarrow{BC} = \overrightarrow{OC}, \quad \mathbf{a} + (\mathbf{b} + \mathbf{c}) = \overrightarrow{OA} + \overrightarrow{AC} = \overrightarrow{OC}.$$

向量的另一个基本运算是数与向量的乘法. 设  $\lambda$  是一个实数,  $\lambda$  与  $\mathbf{a}$  相乘是一个向量, 记为  $\lambda\mathbf{a}$ , 其长度等于  $|\lambda||\mathbf{a}|$ , 当  $\lambda > 0$  时  $\lambda\mathbf{a}$  的方向与  $\mathbf{a}$  相同, 当  $\lambda < 0$  时  $\lambda\mathbf{a}$  的方向与  $\mathbf{a}$  相反. 称  $\lambda\mathbf{a}$  为数  $\lambda$  与向量  $\mathbf{a}$  的数乘. 不难验证, 数乘满足下列性质

**性质 5.2** 设  $\lambda, \mu$  是实数,  $\mathbf{a}, \mathbf{b}$  是向量, 则

- 1°  $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$ ,
- 2°  $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$ ,
- 3°  $(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a})$ .
- 4°  $0\mathbf{a} = \mathbf{0}$ ,  $1\mathbf{a} = \mathbf{a}$ ,  $(-1)\mathbf{a} = -\mathbf{a}$ .

**证明** 这里只证明 1°. 不妨设  $\lambda > 0$ , 则以  $\mathbf{a}, \mathbf{b}$  为边的平行四边形与以  $\lambda\mathbf{a}, \lambda\mathbf{b}$  为边的平行四边形对应边平行, 因此长度成比例, 即可得 1°.

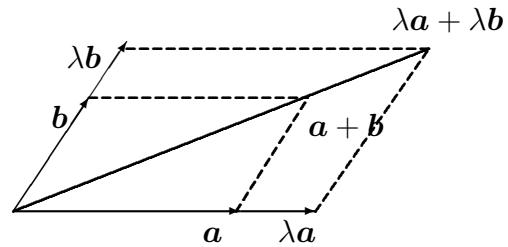


图 5.5

利用数乘, 可以把一个非零向量  $\mathbf{a}$  分解为它的大小与方向的乘积

记与  $\mathbf{a}$  同向的单位向量  $\mathbf{e}_a = \frac{\mathbf{a}}{|\mathbf{a}|}$ , 它可以用来表示  $\mathbf{a}$  的方向. 因此

$$\mathbf{a} = |\mathbf{a}|\mathbf{e}_a$$

## 2° 向量的共线和共面

**定义 5.3** 一组向量, 如果通过平移使它们同处一条直线上, 那么称它们是共线的. 共线的两个非零向量  $\mathbf{a}$  和  $\mathbf{b}$  也称为是相互平行的, 用  $\mathbf{a} \parallel \mathbf{b}$  表示.

一组向量, 如果通过平移使它们同处一个平面上, 那么称它们是共面的. 不难看出, 两个向量通过平移, 使得起点一致, 因此是一定共面的.

根据上述定义, 可以把共线和共面转化为等价的代数表示.

**定理 5.4** 两个向量  $\mathbf{a}, \mathbf{b}$  共线, 当且仅当  $\mathbf{a}, \mathbf{b}$  线性相关, 即存在不全为零的实数  $\mu, \nu$  使得

$$\mu\mathbf{a} + \nu\mathbf{b} = 0.$$

或存在不为零的实数  $\lambda$ , 使得

$$\mathbf{a} = \lambda\mathbf{b}.$$

三个向量  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  共面, 当且仅当  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  线性相关, 即存在不全为零的三个数  $\lambda, \mu, \nu$  使得

$$\lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{c} = 0,$$

或其中一个可以表示成另外两个的线性组合,

$$\mathbf{a} = \mu\mathbf{b} + \nu\mathbf{c}, \mu, \nu \text{ 不全为零.}$$

### 3° 向量的内积和夹角

设  $\mathbf{a} = \overrightarrow{OA}, \mathbf{b} = \overrightarrow{OB}$  是两个向量,  $\theta(\mathbf{a}, \mathbf{b})$  为它们的夹角, 取值范围规定为 0 到  $\pi$ .

**定义 5.5** 两个向量  $\mathbf{a}, \mathbf{b}$  的内积  $\mathbf{a} \cdot \mathbf{b}$  是一个实数, 定义为

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos \theta(\mathbf{a}, \mathbf{b}).$$

若其中一个向量是零向量, 则内积规定为 0. 内积也称作向量的数量积或点乘.

从定义直接看出  $\mathbf{a} \cdot \mathbf{b} > 0$ , 当且仅当  $\theta(\mathbf{a}, \mathbf{b}) < \frac{\pi}{2}$ .  $\mathbf{a} \cdot \mathbf{b} = 0$ , 当且仅当  $\theta(\mathbf{a}, \mathbf{b}) = \frac{\pi}{2}$ , 此时两个向量称为相互正交(或垂直), 记为  $\mathbf{a} \perp \mathbf{b}$ .

从几何上看(图 5.8),  $|\mathbf{a}| \cos \theta$  是向量  $\mathbf{a}$  在向量  $\mathbf{b}$  所在的直线上(简称在  $\mathbf{b}$  上)的投影向量的有向长度, 当夹角  $\theta$  是锐角时, 投影向量的方向与  $\mathbf{b}$  一致; 当夹角  $\theta$  是钝角时, 投影向量的方向与  $\mathbf{b}$  相反.

记  $\mathbf{a}$  在  $\mathbf{b}$  上的投影向量为  $\mathbf{a}_b$ , 则

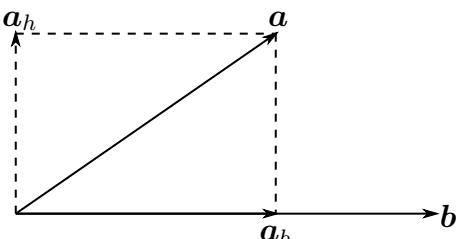
$$\mathbf{a}_b = |\mathbf{a}| \cos \theta \mathbf{e}_b = (\mathbf{a} \cdot \mathbf{e}_b) \mathbf{e}_b,$$

这里  $\mathbf{e}_b = \frac{1}{|\mathbf{b}|} \mathbf{b}$  是  $\mathbf{b}$  的单位向量.

**性质 5.6** 内积满足下列代数性质.

1°  $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ . (交换律)

图 5.6



2°  $(\lambda\mathbf{a}) \cdot \mathbf{b} = \lambda\mathbf{a} \cdot \mathbf{b}$ , 这里  $\lambda$  是任意实数. (结合律)

3°  $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$ . (分配律)

4°  $(\mathbf{a}, \mathbf{a}) = |\mathbf{a}|^2 \geq 0$ , 等号成立当且仅当  $\mathbf{a} = 0$  (正定性)

**证明**  $1^\circ$  的证明是显然的. 在  $2^\circ$  的证明中, 若  $\lambda > 0$ , 则  $|\lambda \mathbf{a}| = \lambda |\mathbf{a}|$ , 且  $\theta(\lambda \mathbf{a}, \mathbf{b}) = \theta(\mathbf{a}, \mathbf{b})$ , 所以等式成立. 若  $\lambda < 0$ , 则  $|\lambda \mathbf{a}| = -\lambda |\mathbf{a}|$ , 且  $\theta(\lambda \mathbf{a}, \mathbf{b}) = \pi - \theta(\mathbf{a}, \mathbf{b})$ , 所以  $\cos \theta(\lambda \mathbf{a}, \mathbf{b}) = \theta(\mathbf{a}, \mathbf{b}) = -\cos \theta(\mathbf{a}, \mathbf{b})$ , 等式也成立.

关于  $3^\circ$  的证明如下: 不妨设  $\mathbf{c}$  是单位向量  $|\mathbf{c}| = 1$ . 因此  $\mathbf{a} + \mathbf{b}$  以及  $\mathbf{a}, \mathbf{b}$  在  $\mathbf{c}$  上的投影向量满足

$$(\mathbf{a} + \mathbf{b})_c = \mathbf{a}_c + \mathbf{b}_c,$$

所以

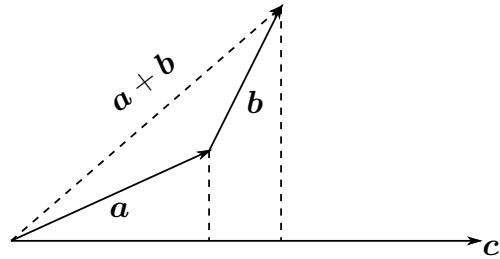


图 5.7

$$\begin{aligned} (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} &= (\mathbf{a} + \mathbf{b})_c \cdot \mathbf{c} = (\mathbf{a}_c + \mathbf{b}_c) \cdot \mathbf{c} \\ &= ((\mathbf{a} \cdot \mathbf{c})\mathbf{c} + (\mathbf{b} \cdot \mathbf{c})\mathbf{c}) \cdot \mathbf{c} = (\mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c})(\mathbf{c} \cdot \mathbf{c}) \\ &= \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}. \end{aligned}$$

这样就完成了该性质的证明.  $\square$

#### 4° 向量的外积和有向面积

向量的另一个重要运算称为向量的外积. 为此我们先引进“右手系”的概念.

设  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  是三个不共面向量构成的有序向量组, 并有同一个起点. 前两个向量  $\mathbf{a}, \mathbf{b}$  就决定了一个平面, 而  $\mathbf{c}$  的方向指向平面的某一侧. 当右手四指顺着平面, 按照从  $\mathbf{a}$  到  $\mathbf{b}$  转动时, 如果拇指与  $\mathbf{c}$  都指向平面的同一侧, 那么称  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  为右手系, 否则称为左手系. 容易看出, 如果  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  是右手系, 那么  $\{\mathbf{b}, \mathbf{c}, \mathbf{a}\}$  和  $\{\mathbf{c}, \mathbf{a}, \mathbf{b}\}$  仍是右手系. 但是  $\{\mathbf{b}, \mathbf{a}, \mathbf{c}\}$  和  $\{\mathbf{a}, \mathbf{b}, -\mathbf{c}\}$  都是左手系.

**定义 5.7** 两个向量  $\mathbf{a}$  和  $\mathbf{b}$  的外积是一个向量, 记为  $\mathbf{a} \times \mathbf{b}$ , 它的大小规定为以  $\mathbf{a}, \mathbf{b}$  为邻边的平行四边形的面积

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin \theta(\mathbf{a}, \mathbf{b}),$$

它的方向规定为:  $\mathbf{a} \times \mathbf{b}$  与  $\mathbf{a}$  和  $\mathbf{b}$  垂直, 且向量组  $\{\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}\}$  构成右手系 (图 5.10). 若  $\mathbf{a}$  和  $\mathbf{b}$  中有一个是零向量, 则两者的外积规定为零向量. 外积也称为向量积或叉乘.

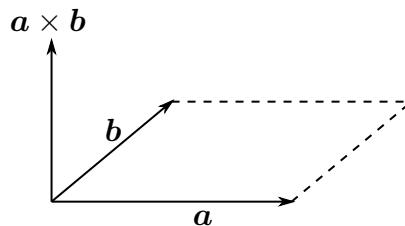


图 5.8

显然, 如果两个向量平行, 那么它们的外积等于零向量.

**性质 5.8** 向量的外积运算满足如下性质:

$1^\circ$   $\mathbf{a} \times \mathbf{b} = 0$ , 当且仅当  $\mathbf{a}$  和  $\mathbf{b}$  平行.

2°  $(\lambda \mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda \mathbf{b}) = \lambda(\mathbf{a} \times \mathbf{b})$ , 这里  $\lambda$  是一个实数.

3°  $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ , (反称性)

4°  $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ . (分配律)

1°, 2°, 3° 可以利用定义直接验证, 关于 4° 的证明, 将在引进坐标后给出, 那里证明将变得更加简单.

正如定义中所提到的, 向量  $\mathbf{a} \times \mathbf{b}$  的大小以  $\mathbf{a}, \mathbf{b}$  为邻边的平行四边形的面积, 但是由性质 5.8 可知  $\mathbf{b} \times \mathbf{a}$  的大小也是这个四边形的面积, 只是方向与  $\mathbf{a} \times \mathbf{b}$  相反. 因此我们实际上定义了面积的有向性, 即如果规定  $\mathbf{a} \times \mathbf{b}$  为四边形面积的正向, 那么  $\mathbf{b} \times \mathbf{a}$  就表示四边形面积的负项. 称  $\mathbf{a} \times \mathbf{b}$  为以  $\mathbf{a}, \mathbf{b}$  为邻边的平行四边形的有向面积.

### 5° 向量的混合积和有向体积

**定义 5.9** 向量  $\mathbf{a} \times \mathbf{b}$  与向量  $\mathbf{c}$  的内积  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$  为三个向量  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  的混合积.

**性质 5.10** 设  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  为三个非零向量, 则

1°  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  的混合积的绝对值是以  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  为棱的平行六面体的体积.

2°  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 0$  当且仅当  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  共面.

3°  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b}$ .

**证明** 以  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  为棱的平行六面体 (图 5.10) 是指: 将向量表示为同一起点的有向线段时, 以三个有向线段为棱的平行六面体. 它的底面积  $S$  等于  $|\mathbf{a} \times \mathbf{b}|$ , 高  $h$  等于向量  $\mathbf{c}$  在底面单位法向  $\frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|}$  上投影向量的长度, 即

$$h = \frac{|\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})|}{|\mathbf{a} \times \mathbf{b}|},$$

所以体积

$$V = S \cdot h = |\mathbf{a} \times \mathbf{b} \cdot \mathbf{c}|.$$

当体积为零时, 当且仅当三个向量共面.

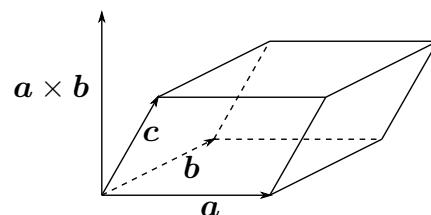


图 5.9

关于 3° 的证明, 只要注意到  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$  的正负取决于有序向量组  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  是否是右手系, 如果是右手系, 向量  $\mathbf{a} \times \mathbf{b}$  与  $\mathbf{c}$  的夹角为锐角, 因此  $\mathbf{a} \times \mathbf{b} \cdot \mathbf{c} > 0$ , 当  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  是左手系时,  $\mathbf{a} \times \mathbf{b}$  与  $\mathbf{c}$  的夹角为钝角, 所以  $\mathbf{a} \times \mathbf{b} \cdot \mathbf{c} < 0$ . 3° 中  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}, (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a}$  和  $(\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b}$  的大小都是同一个平行六面体的体积, 而且对应的有序向量组  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}, \{\mathbf{b}, \mathbf{c}, \mathbf{a}\}$  和  $\{\mathbf{c}, \mathbf{a}, \mathbf{b}\}$  要么同是右手系, 要么同是左手系, 所以 3° 成立.

根据上述性质, 我们把混合积  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$  定义为以  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  为棱的平行六面体的有向

体积, 当三个向量为右手系时, 体积为正, 当为左手系时, 体积为负.

## §5.2 向量的坐标表示和坐标系

以上, 我们从几何上定义了向量的加法和数乘, 并引进了共线共面的概念. 在此基础上将建立坐标系.

### 1° 仿射坐标系

**定理 5.11** 设向量组  $\{e_1, e_2, e_3\}$  是空间中三个不共面的向量, 则任一向量  $x$  都存在唯一的一组数  $(x_1, x_2, x_3)$ , 使得  $x$  表示为  $\{e_1, e_2, e_3\}$  的线性组合

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3.$$

**证明** 设  $x = \overrightarrow{OP}$ ,  $e_1 = \overrightarrow{OA}$ ,  $e_2 = \overrightarrow{OB}$ ,  $e_3 = \overrightarrow{OC}$ . 过  $P$  点做平行直线  $OC$  的平行线, 交  $AOB$  所在平面于  $Q$  点. 再过  $Q$  作平行于  $OB$  的平行线, 交  $OC$  于  $R$ . 则

$$x = \overrightarrow{OP} = \overrightarrow{OR} + \overrightarrow{RQ} + \overrightarrow{QP}.$$

由于

$$\overrightarrow{OR} \parallel \overrightarrow{OA}, \overrightarrow{RQ} \parallel \overrightarrow{OB}, \overrightarrow{QP} \parallel \overrightarrow{OC},$$

因此分别共线, 即存在实数  $x_1, x_2, x_3$  使得

$$\overrightarrow{OR} = x_1 \overrightarrow{OA} = x_1 e_1, \overrightarrow{RQ} = x_2 \overrightarrow{OB} = x_2 e_2, \overrightarrow{QP} = x_3 \overrightarrow{OC} = x_3 e_3,$$

即  $x = x_1 e_1 + x_2 e_2 + x_3 e_3$ . 假如存在另外一组坐标  $x'_1, x'_2, x'_3$ , 使得  $x = x'_1 e_1 + x'_2 e_2 + x'_3 e_3$ , 那么

$$(x_1 - x'_1) e_1 + (x_2 - x'_2) e_2 + (x_3 - x'_3) e_3 = 0,$$

由于  $e_1, e_2, e_3$  不共面, 所以由定理 5.4 知  $x_1 = x'_1, x_2 = x'_2, x_3 = x'_3$ , 即坐标是唯一的.  $\square$

**定义 5.12** 空间中任意三个不共面的向量  $e_1, e_2, e_3$  称为空间的一组基, 对于向量  $x$ , 称

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3$$

中的系数  $(x_1, x_2, x_3)$  为向量  $x$  在基  $e_1, e_2, e_3$  下的仿射坐标或简称坐标.

若选定原点  $O$ , 则  $O$  与一组基  $e_1, e_2, e_3$  (通常选定满足右手系的一组基) 合在一起记为  $[O; e_1, e_2, e_3]$ , 并称为空间的仿射坐标系. 在坐标系中, 三个向量  $e_1, e_2, e_3$  也分别称为坐标向量.

给定仿射坐标系  $[O; e_1, e_2, e_3]$ , 结合空间中点  $P$  与向量的对应  $x = \overrightarrow{OP}$ , 就建立了空间中点、向量和坐标(数组)之间的一一对应关系:

$$\text{空间中的点 } P \longleftrightarrow \text{向量 } \overrightarrow{OP} \longleftrightarrow \text{坐标 } (x_1, x_2, x_3).$$

因此也称  $(x_1, x_2, x_3)$  是点  $P$  的坐标.

一切关于向量从几何上定义的加法、数乘、内积、外积以及混合积都可以转化为坐标之间的运算. 例如, 两个向量的加法就是对应坐标的加法, 数乘就是对应坐标的数乘:

$$\lambda \mathbf{x} + \mu \mathbf{y} \longleftrightarrow (\lambda x_1 + \mu x'_1, \lambda x_2 + \mu x'_2, \lambda x_3 + \mu x'_3).$$

其中  $\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3, \mathbf{y} = x'_1 \mathbf{e}_1 + x'_2 \mathbf{e}_2 + x'_3 \mathbf{e}_3$ .

### 2° 直角坐标系

为了进一步简化计算, 取空间中两两相互垂直、并形成右手系的三个单位向量作为坐标向量, 记为  $\mathbf{i}, \mathbf{j}, \mathbf{k}$ , 并称这一特殊的仿射坐标系  $[O; \mathbf{i}, \mathbf{j}, \mathbf{k}]$  为直角坐标系. 相应的坐标轴分别称为  $x$  轴、 $y$  轴和  $z$  轴,

在直角坐标系  $[O; \mathbf{i}, \mathbf{j}, \mathbf{k}]$  中, 任何向量表示为

$$\mathbf{x} = \overrightarrow{OP} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

有时也直接用向量的坐标表示该向量

$$\mathbf{x} = \overrightarrow{OP} = (x, y, z),$$

这样, 三个坐标向量分别表示为

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1).$$

有时, 也把直角坐标系记为  $Oxyz$ .

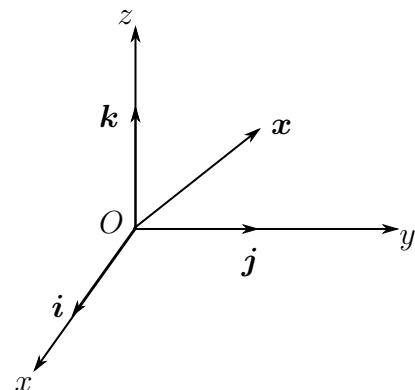


图 5.10

关于仿射坐标系的所有概念与结论都适用于直角坐标系, 而直角坐标系的特殊性使得利用向量的坐标进行有关计算变得更加容易和简洁.

### 3° 直角坐标系下向量的运算

不难验证, 三个坐标向量  $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$  满足如下关系:

$$(1) \mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0, \quad \mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1.$$

$$(2) \mathbf{i} \times \mathbf{j} = \mathbf{k}; \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}; \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}, \quad \mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0.$$

这是下列一系列计算的基础.

设  $\mathbf{a} = a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}, \mathbf{b} = b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k}, \mathbf{c} = c_1 \mathbf{i} + c_2 \mathbf{j} + c_3 \mathbf{k}$ , 则

(1) 向量的加法和数乘即是坐标对应的加法和数乘(这里  $\mu, \nu \in \mathbb{R}$ ):

$$\mu \mathbf{a} + \nu \mathbf{b} = (\mu a_1 + \nu b_1) \mathbf{i} + (\mu a_2 + \nu b_2) \mathbf{j} + (\mu a_3 + \nu b_3) \mathbf{k}.$$

(2)  $\mathbf{a}$  和  $\mathbf{b}$  的内积:

$$\mathbf{a} \cdot \mathbf{b} = (a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}) \cdot (b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k}) = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

因此两向量的夹角、向量的模长(大小)以及两点之间的距离可分别表示为

$$\begin{aligned}\theta(\mathbf{a}, \mathbf{b}) &= \arccos \left( \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} \sqrt{b_1^2 + b_2^2 + b_3^2}} \right), \\ |\mathbf{a}| &= \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{a_1^2 + a_2^2 + a_3^2}, \\ d(P_1, P_2) &= |\overrightarrow{P_1 P_2}| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}.\end{aligned}$$

这里两点的坐标分别是  $P_1(a_1, a_2, a_3)$  和  $P_2(b_1, b_2, b_3)$ .

(3)  $\mathbf{a}$  和  $\mathbf{b}$  的外积:

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= (a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}) \times (b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k}) \\ &= (a_2 b_3 - a_3 b_2) \mathbf{i} + (a_3 b_1 - a_1 b_3) \mathbf{j} + (a_1 b_2 - a_2 b_1) \mathbf{k} \\ &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} + \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}.\end{aligned}$$

这里  $\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$  为三个二阶行列式, 它们分别是有向面积  $\mathbf{a} \times \mathbf{b}$  在三个坐标平面  $Oyz, Ozx, Oxy$  上的有向投影.

(4)  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  的混合积:

$$\mathbf{a} \times \mathbf{b} \cdot \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

关于向量运算的其他性质的证明, 在坐标表示下也会变得更加简单.

**例 5.2.1** 性质5.8中外积分配律的证明.

设  $\mathbf{a} = a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}$ ,  $\mathbf{b} = b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k}$ ,  $\mathbf{c} = c_1 \mathbf{i} + c_2 \mathbf{j} + c_3 \mathbf{k}$ , 则

$$\begin{aligned}\mathbf{a} \times (\mathbf{b} + \mathbf{c}) &= (a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k}) \times [(b_1 + c_1) \mathbf{i} + (b_2 + c_2) \mathbf{j} + (b_3 + c_3) \mathbf{k}] \\ &= [a_2(b_3 + c_3) - a_3(b_2 + c_2)] \mathbf{i} + [a_3(b_1 + c_1) - a_1(b_3 + c_3)] \mathbf{j} \\ &\quad + [a_1(b_2 + c_2) - a_2(b_1 + c_1)] \mathbf{k} \\ &= [(a_2 b_3 - a_3 b_2) + (a_2 c_3 - a_3 c_2)] \mathbf{i} + [(a_3 b_1 - a_1 b_3) + (a_3 c_1 - a_1 c_3)] \mathbf{j} \\ &\quad + [(a_1 b_2 - a_2 b_1) + (a_1 c_2 - a_2 c_1)] \mathbf{k} \\ &= \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}.\end{aligned}$$

**例 5.2.2** 证明 Cauchy 不等式

$$(a_1 b_1 + a_2 b_2 + a_3 b_3)^2 \leq (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2),$$

等号成立当且仅当存在实数  $\lambda$  满足  $a_i = \lambda b_i$ ,  $i = 1, 2, 3$ .

**证明** 设向量  $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ ,  $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ , 则 Cauchy 不等式等价于

$$(\mathbf{a} \cdot \mathbf{b})^2 = (|\mathbf{a}| |\mathbf{b}| \cos \theta)^2 \leq |\mathbf{a}|^2 |\mathbf{b}|^2$$

而且, 等号成立当且仅当  $\cos \theta = \pm 1$ , 即  $\mathbf{a}$  与  $\mathbf{b}$  共线.

#### 4° 行列式与有向面(体)积

对于任何一个实的二阶行列式

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix},$$

把两行分别看成平面上两点的坐标  $P_1(a_1, a_2)$ ,  $P_2(b_1, b_2)$ , 那么该行列式几何上就表示了以  $O, P_1, P_2$  为三个顶点或以  $\mathbf{a} = \overrightarrow{OP_1}$ ,  $\mathbf{b} = \overrightarrow{OP_2}$  为棱的平行四边形的面积. 因此,  $\mathbf{a}, \mathbf{b}$  不共线, 当且仅当对应的行列式不为零. 行列式取正或负, 体现了面积的有向性.

对任意一个三阶行列式

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

只要把每一行作为空间中一点的坐标

$$P_1(a_1, a_2, a_3), P_2(b_1, b_2, b_3), P_3(c_1, c_2, c_3),$$

那么三阶行列式几何上就表示以  $O, P_1, P_2, P_3$  为顶点或以

$$\mathbf{a} = \overrightarrow{OP_1}, \mathbf{b} = \overrightarrow{OP_2}, \mathbf{c} = \overrightarrow{OP_3}$$

为棱的平行六面体的体积. 因此三个向量  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  不共面, 当且仅当对应的行列式不为零. 行列式取正或负, 体现了体积的有向性.

不难验证, 关于行列式的运算规则与向量外积和混合积是一致的. 特别是两行交换对应外积的交换, 或混合积的左右手系的互换. 而在三阶行列式中, 行之间的轮换, 对应的是混合积性质5.10 中的 3°. 其他对应也均可验证.

### §5.3 平面、直线与曲面

利用向量和坐标系, 空间中的一些曲面和曲线可以表示为代数方程. 我们首先讨论最简单的平面和直线的代数方程.

#### 1° 平面

几何上看, 过一定点并垂直一个确定方向就可确定一个平面.

在坐标系  $Oxyz$  中, 设平面  $\Pi$  过定点  $P_0(x_0, y_0, z_0)$ , 并与给定向量  $\mathbf{n} = ai + bj + ck$  垂直,  $\mathbf{n}$  称为平面  $\Pi$  的 法向量.

对  $\Pi$  上任意一点  $P(x, y, z)$ , 则  $\overrightarrow{P_0P} \perp \mathbf{n}$ , 也就是

$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0 \text{ 或者 } \mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0,$$

其中  $\mathbf{r} = \overrightarrow{OP}$ ,  $\mathbf{r}_0 = \overrightarrow{OP_0}$  分别是  $P$  和  $P_0$  的位置向量. 这就是平面  $\Pi$  向量形式的方程.

注意到

$$\mathbf{r} - \mathbf{r}_0 = \overrightarrow{P_0P} = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k},$$

因此平面上任意一点  $P(x, y, z)$  的坐标满足下列三元一次方程

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

或

$$ax + by + cz + d = 0,$$

其中

$$d = -(ax_0 + by_0 + cz_0)$$

是一个已知数. 称上述方程为平面的一般方程. 反之, 对任意一个三元一次方程  $ax + by + cz + d = 0$ , 方程的解都在以系数构成的  $\mathbf{n} = ai + bj + ck$  为法向量的平面上. 这样我们就把平面的几何描述, 转化为一个代数方程.

平面方程的一些特殊情形是值得关注的.

- (1)  $d = 0$ , 方程退化为  $ax + by + cz = 0$ , 因此  $(0, 0, 0)$  满足方程, 即平面过原点.
- (2)  $c = 0$ , 此时法向量  $\mathbf{n} = (a, b, 0)$  垂直于  $z$  轴, 所以方程  $ax + by + d = 0$  表示平行于  $z$  轴的平面. 对于  $a = 0$  或  $b = 0$  的情形类似.
- (3)  $a = b = 0$ , 此时法向量  $\mathbf{n} = (0, 0, c)$  平行于  $z$  轴, 所以方程  $cz + d = 0$  或者  $z = -\frac{d}{c}$  表示过点  $(0, 0, -\frac{d}{c})$  且平行于  $Oxy$  平面的平面. 其他情形类似.

**例 5.3.1** 设给定空间不在一条直线上的三点  $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2), P_3(x_3, y_3, z_3)$ , 以  $P_1$  为定点,  $\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}$  为法向量, 就可确定过三点的平面

$$\overrightarrow{P_1P} \cdot (\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}) = \begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0,$$

特别, 当三点分别位于三个坐标轴上  
 $P_1(\alpha, 0, 0), P_2(0, \beta, 0), P_3(0, 0, \gamma)$ , 则方程为

$$\begin{vmatrix} x - \alpha & y & z \\ -\alpha & \beta & 0 \\ -\alpha & 0 & \gamma \end{vmatrix} = 0,$$

或简化为

$$\frac{x}{\alpha} + \frac{y}{\beta} + \frac{z}{\gamma} = 1.$$

该平面分别于三个坐标轴在给定三点相截并称  $\alpha, \beta, \gamma$  为平面与数轴的截距.

有了平面方程, 我们可以讨论与平面有关的一系列问题.

**例 5.3.2** 讨论两个平面之间的夹角. 设两个平面方程为

$$a_1x + b_1y + c_1z + d_1 = 0, \quad \mathbf{n}_1 = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k};$$

$$a_2x + b_2y + c_2z + d_2 = 0, \quad \mathbf{n}_2 = a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}.$$

它们的夹角  $\theta$  定义为两个平面法向量  $\mathbf{n}_1$  与  $\mathbf{n}_2$  的夹角, 并称为两平面的二面角. 因此

$$\cos \phi = \frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{|\mathbf{n}_1||\mathbf{n}_2|} = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2}\sqrt{a_2^2 + b_2^2 + c_2^2}},$$

当  $\phi = 0$  时, 两平面平行; 即  $\mathbf{n}_1 \parallel \mathbf{n}_2$ , 当  $\phi = \pi$  时, 两平面垂直, 也就是  $\mathbf{n}_1 \perp \mathbf{n}_2$ , 即  $\mathbf{n}_1 \cdot \mathbf{n}_2 = a_1a_2 + b_1b_2 + c_1c_2 = 0$ .

空间中一个平面实际上把空间分成三个部分: 法向量所指的“上”半部分, 上半部分的反面“下”半部分, 以及平面本身. 判断空间中一点  $P_0$  落在哪部分, 只要看平面上任何一个点  $P$  到  $P_0$  的向量  $\overrightarrow{PP_0}$  与平面法向量的夹角是锐角、钝角或是直角, 也就是看  $\mathbf{n} \cdot \overrightarrow{PP_0}$  是正的、负的或是零. 因为  $P$  满足方程, 所以

$$\mathbf{n} \cdot \overrightarrow{PP_0} = a(x_0 - x) + b(y_0 - y) + c(z_0 - z) = ax_0 + by_0 + cz_0 + d,$$

根据上式就有

$P_0$  在平面的上半部分, 当且仅当  $ax_0 + by_0 + cz_0 + d > 0$ ,

$P_0$  在平面的下半部分, 当且仅当  $ax_0 + by_0 + cz_0 + d < 0$ ,

$P_0$  在平面上, 当且仅当  $ax_0 + by_0 + cz_0 + d = 0$ .

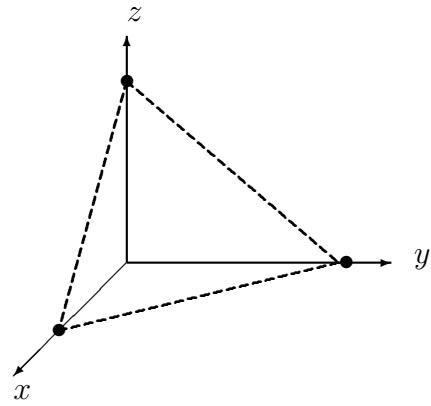


图 5.11

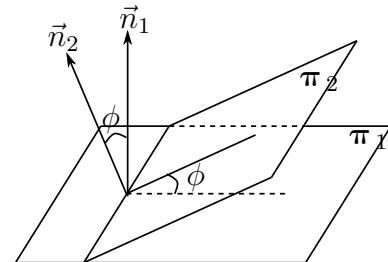


图 5.12

## 2° 直线

空间中过一定点  $P_0(x_0, y_0, z_0)$  并沿着给定的方向  $\mathbf{v} = li + mj + nk$ , 就可确定一条直线, 对该直线上任意一点  $P(x, y, z)$ ,  $\overrightarrow{P_0P} = \mathbf{r} - \mathbf{r}_0$  与  $\mathbf{v}$  共线, 因此有

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}, \quad \text{或} \quad \begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases}$$

也就是直线上任意一点  $P$  的坐标  $(x, y, z)$  表示为参数  $t$  的线性方程, 称为直线的点向式方程, 向量  $\mathbf{v}$  称为直线的方向向量. 如果消去参数  $t$  就得到

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}.$$

它实际上是含有两个三元一次方程的方程组.

如果方向向量的坐标之一为零, 例如  $l = 0$ , 此时直线的方向与  $x$  轴垂直, 点向式方程应理解为下述的方程组

$$x = x_0, \quad \frac{y - y_0}{m} = \frac{z - z_0}{n},$$

如果有两个为零, 例如  $l = 0, m = 0$ , 则直线与  $z$  轴平行, 方程为

$$x = x_0, \quad y = y_0.$$

设直线  $L$  为两个不平行平面的交线, 因此点  $P(x, y, z)$  在交线  $L$  上当且仅当它在两个平面上, 也就是它的坐标满足三元一次方程组

$$\begin{cases} a_1x + b_1y + c_1z + d_1 = 0 \\ a_2x + b_2y + c_2z + d_2 = 0 \end{cases}$$

这个方程组称为直线的一般方程. 因为直线  $L$  垂直于两个平面的法向量  $\mathbf{n}_1$  和  $\mathbf{n}_2$ , 所以它的方向向量  $\mathbf{v} = \mathbf{n}_1 \times \mathbf{n}_2$ . 再从上述方程组求出一个特解  $P_0(x_0, y_0, z_0)$ , 则直线的一般方程也可以写成点向式方程.

**例 5.3.3** 求过两个给定点  $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2)$  的直线方程.

只要取方向向量为  $\mathbf{v} = \overrightarrow{P_2P_1} = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k}$ , 就有

$$\begin{aligned} \mathbf{r} &= \mathbf{r}_1 + t\mathbf{v} = \mathbf{r}_1 + t(\mathbf{r}_2 - \mathbf{r}_1) \\ \frac{x - x_1}{x_2 - x_1} &= \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}. \end{aligned}$$

**例 5.3.4** 两直线的关系. 给定两条直线

$$L_1 : \mathbf{r} = \mathbf{r}_1 + t\mathbf{v}_1, \quad \text{其中 } \mathbf{r}_1 = (x_1, y_1, z_1), \mathbf{v}_1 = (l_1, m_1, n_1),$$

$$L_2 : \mathbf{r} = \mathbf{r}_2 + t\mathbf{v}_2, \quad \text{其中 } \mathbf{r}_2 = (x_2, y_2, z_2), \mathbf{v}_2 = (l_2, m_2, n_2).$$

它们的位置关系可分“共面”和“异面”两种情形.

直线  $L_1$  与  $L_2$  共面等价于  $\mathbf{v}_1, \mathbf{v}_2$  与  $\mathbf{r}_1 - \mathbf{r}_2$  共面:

$$(\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{v}_1 \times \mathbf{v}_2 = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \end{vmatrix} = 0.$$

而直线  $L_1$  与  $L_2$  异面的充分必要条件是  $(\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{v}_1 \times \mathbf{v}_2 \neq 0$

在  $L_1$  与  $L_2$  共面情况下, 两直线方向向量  $\mathbf{v}_1$  与  $\mathbf{v}_2$  的夹角  $\theta$  就是两直线的夹角.

$$\cos \theta = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1| |\mathbf{v}_2|} = \frac{l_1 l_2 + m_1 m_2 + n_1 n_2}{\sqrt{l_1^2 + m_1^2 + n_1^2} \sqrt{l_2^2 + m_2^2 + n_2^2}}$$

当  $\cos \theta = \pm 1$  时,  $\mathbf{v}_1 \parallel \mathbf{v}_2$ , 因此两直线平行当且仅当

$$\frac{l_1}{l_2} = \frac{m_1}{m_2} = \frac{n_1}{n_2};$$

当  $|\cos \theta| < 1$  时两直线相交, 特别, 两直线互相垂直当且仅当

$$l_1 l_2 + m_1 m_2 + n_1 n_2 = 0.$$

### 3° 二次曲面

所谓二次曲面是指空间中点  $P(x, y, z)$  的坐标  $(x, y, z)$  满足二次代数方程所形成的曲面. 例如

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2$$

表示以  $P_0(x_0, y_0, z_0)$  为球心,  $R$  为半径的球面.

下面给出一些典型的二次曲面.

(1) 椭球面 设  $a > 0, b > 0, c > 0$ ,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

如果与  $Oxy$  平面相交, 则交线是一个椭圆, 一般情况下若与平行  $Oxy$  的平面  $z = h$  ( $|h| < c$ ) 相交, 交线也是一个椭圆, 只是椭圆的长半轴和短半轴分别为

$$a' = a \sqrt{1 - \frac{h^2}{c^2}}, \quad b' = b \sqrt{1 - \frac{h^2}{c^2}}.$$

(2) 单叶和双叶双曲面 设  $a > 0, b > 0, c > 0$ ,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = \pm 1.$$

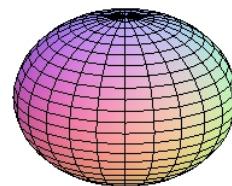


图 5.13

上式右边取正号为单叶双曲面, 去负号为双叶双曲面.

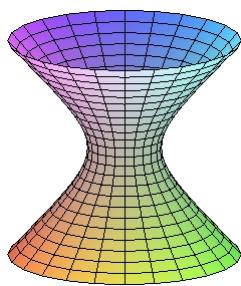


图 5.14

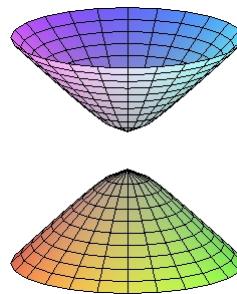


图 5.15

(3) 椭圆抛物面和双曲抛物面 设  $a > 0, b > 0$ ,

$$z = \frac{x^2}{a^2} \pm \frac{y^2}{b^2},$$

其中取正号对应椭圆抛物面, 取负号对应双曲抛物面. 双曲抛物面也称为马鞍面.

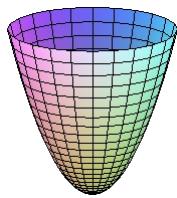


图 5.16

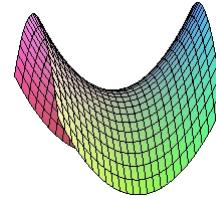


图 5.17

读者会提出这样的问题: 二次代数方程的一般形式为

$$a_1x^2 + a_2y^2 + a_3z^2 + b_1xy + b_2yz + b_3xz + c_1x + c_2y + c_3z + d = 0,$$

除了上述例子外, 其他二次代数方程表示什么样的图形?

注意到一个几何事实, 即空间中的图形在平移和旋转之下形状是不变的. 因此, 尽管三个变量  $(x, y, z)$  的二次方程种类繁多, 但除一些退化情形外, 可以通过坐标系的平移和旋转变换到 9 类非平凡的标准方程, 因此二次代数方程所表示的图形也仅仅有(非平凡的)9 种, 这里我们仅仅列出了其中的 5 种.

坐标系的平移和旋转本质上正是平面解析几何中所熟知 配方法, 通过配方法可把二元二次方程变换成标准的椭圆、抛物和双曲线. 这里仅给出一个坐标系的平移和旋转的例子, 详细情况就不展开讨论了.

### 例 5.3.5 二次方程

$$z = xy.$$

通过绕 $z$  轴旋转  $\frac{\pi}{4}$ , 或令

$$x' = \frac{\sqrt{2}}{2}(x+y), \quad y' = \frac{\sqrt{2}}{2}(x-y), \quad z' = z,$$

则方程变换为  $(x', y', z')$  满足的方程

$$z' = \frac{x'^2}{2} - \frac{y'^2}{2}.$$

因此  $z = xy$  表示的也是马鞍面, 只是图5.17 中的马鞍面绕  $z$  轴旋转了  $\frac{\pi}{4}$  而已.

## §5.4 其它常用坐标系

直角坐标系是通过确定原点和空间三个两两正交的单位向量构成的. 这样坐标系称为**线性坐标系**. 它们的特征之一是, 一个坐标分量等于常数对应的都是平面. 例如  $z = c$  表示空间中平行于坐标平面  $Oxy$  的平面. 下面介绍常用的极坐标系、柱坐标系和球坐标系, 但这些坐标系已经不再是线性坐标系.

### 1° 平面的极坐标系

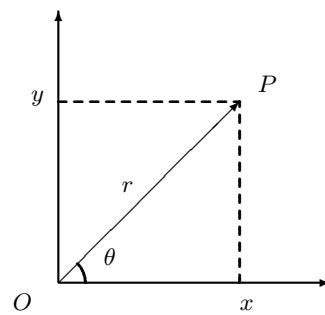
为了保持完整性, 首先介绍平面极坐标系. 在平面上取定一点  $O$  (称为极点), 从极点引一条射线  $Ox$  (称为极轴), 再选定一个长度单位和角度的正向 (通常取极轴正向的逆时针方向), 这样就构成了平面上的**极坐标系**. 对于平面上任意一点  $P$ , 用  $r$  表示  $P$  到极点  $O$  的距离 (线段  $\overline{OP}$  的长度或向量  $\overrightarrow{OP}$  的大小),  $\theta$  表示从极轴到向量  $\overrightarrow{OP}$  的正向夹角 (幅角), 则数组  $(r, \theta)$  可以用来确定点  $P$  在空间的位置, 并称为  $P$  点的极坐标. 这里,  $r$  的取值范围为  $[0, +\infty)$ ,  $\theta$  的取值范围为  $[0, 2\pi]$ .

如果在平面直角坐标系  $Oxy$  中, 取原点为极坐标系的极点,  $x$  轴为极轴, 那么平面上任意一点  $P$  的直角坐标和极坐标之间的关系 (变换) 如下:

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases} \quad \text{或者} \quad \begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = \arctan \frac{y}{x} \end{cases}.$$

$P$  的位置向量可以表示为

$$\mathbf{r} = xi + yj = r \cos \theta i + r \sin \theta j.$$



不难发现,  $r = \text{常数}$  是平面上以原点为圆心的同心圆,

$\theta = \text{常数}$  是从原点出发的射线.

图 5.18

### 2° 柱面坐标系

设空间取定直角坐标系  $Oxyz$  对任意一点  $P(x, y, z)$ , 位置向量  $\overrightarrow{OP}$  在  $Oxy$  平面上的投影向量

$$\overrightarrow{OP'} = xi + yj,$$

用极坐标表示

$$\overrightarrow{OP'} = r \cos \theta i + r \sin \theta j,$$

因此, 位置向量可表示为

$$\overrightarrow{OP} = \overrightarrow{OP'} + zk = r \cos \theta i + r \sin \theta j + zk.$$

或者

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = z,$$

其中

$$0 \leq r < +\infty, \quad 0 \leq \theta < 2\pi, \quad -\infty < z < +\infty.$$

这样就给出了空间的柱面坐标系. 数组  $(r, \theta, z)$  称为点  $P$  的柱面坐标. 在柱面坐标系中,  $r = c$  (正的常数) 表示以  $c$  为半径的圆柱面  $x^2 + y^2 = c^2$ ,  $\theta = \theta_0$  (常数) 是以  $z$  轴为边的半平面.

### 3° 球面坐标系

设位置向量  $\overrightarrow{OP}$  与  $z$  轴的方向角为  $\theta$ ,  $\overrightarrow{OP}$  在  $Oxy$  平面上的投影向量为  $\overrightarrow{OP'}$ , 则

$$z = |OP| \cos \theta, \quad \mathbf{r} = \overrightarrow{OP} = \overrightarrow{OP'} + |OP| \cos \theta \mathbf{k}.$$

将点  $P'$  用  $Oxy$  平面上的极坐标表示出, 设  $\varphi$  是  $\overrightarrow{OP'}$  在  $Oxy$  平面上的幅角, 则

$$\overrightarrow{OP'} = |OP'| \cos \varphi i + |OP'| \sin \varphi j.$$

令  $r = |OP| = |\mathbf{r}|$ , 则  $|OP'| = |OP| \sin \theta$ , 因此

$$\mathbf{r} = \overrightarrow{OP} = r \sin \theta \cos \varphi i + r \sin \theta \sin \varphi j + r \cos \theta \mathbf{k},$$

或

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta.$$

其中

$$0 \leq r < +\infty, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \varphi < 2\pi.$$

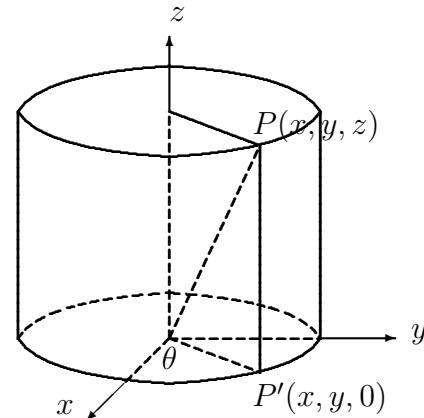


图 5.19

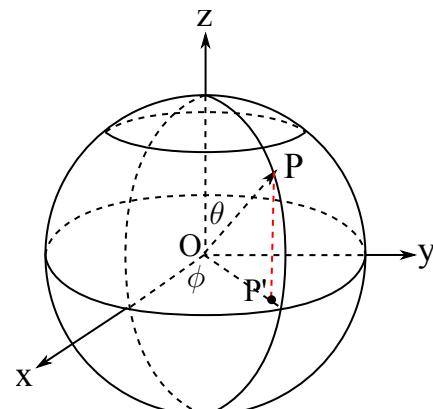


图 5.20

数组  $(r, \theta, \varphi)$  称为点  $P$  的球面座标, 所形成的坐标系称为 **球面座标系**.

## §5.5 向量空间

读者或已注意到, 在二维平面和三维空间中, 无论是“有大小, 有方向”的向量, 还是向量的共线或共面, 从几何上看是十分直观的. 但是, 如果在更高维数的抽象空间, 自然要问, 如何定义向量和向量的空间? 如何定义基和坐标? 如何定义向量的长度和向量之间的夹角? 回答这一系列问题, 需要从具体的三维空间的模型出发, 给出抽象的定义. 这里, 大致介绍有关思路:

### 1° $n$ 维向量空间

首先, 把三维空间中向量的加法和数乘所满足的性质5.1和性质5.2, 以及“0 向量”、“负向量”等概念抽象出来, 作为一般向量空间的定义.

**定义 5.13** 设  $V$  是一个集合,  $V$  中的元素仍然称为“向量”,  $\mathbb{F}$  是一个给定的数域, 如果对  $V$  中的“向量”可定义加法和数乘:

- (1) 加法:  $\mathbf{a} + \mathbf{b} \in V$ ,  $\mathbf{a}, \mathbf{b} \in V$ ;
- (2) 数乘:  $\lambda \mathbf{a} \in V$ ,  $\mathbf{a} \in V$ ,  $\lambda \in \mathbb{F}$ ;
- (3)  $V$  中存在一个元素称为“0 向量”, 满足  $\mathbf{a} + 0 = 0 + \mathbf{a} = \mathbf{a}$ ,  $\mathbf{a} \in V$ ;
- (4) 对任何  $\mathbf{a} \in V$ , 存在  $\mathbf{b} \in V$ , 使得  $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} = 0$ , 记  $\mathbf{b} = -\mathbf{a}$  并称为  $\mathbf{a}$  的负向量, 因此  $\mathbf{b} - \mathbf{a} = \mathbf{b} + (-\mathbf{a})$ ;

加法和数乘满足下列运算规律

- (A1)  $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ ,  $\mathbf{a}, \mathbf{b} \in V$  (交换律);
- (A2)  $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ ,  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$  (结合律);
- (A3)  $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$ ,  $\mathbf{a}, \mathbf{b} \in V$ ,  $\lambda \in \mathbb{F}$ ;
- (A4)  $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$ ,  $\mathbf{a} \in V$ ,  $\lambda, \mu \in \mathbb{F}$ ;
- (A5)  $(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a})$ ,  $\mathbf{a} \in V$ ,  $\lambda, \mu \in \mathbb{F}$ ;

那么称  $V$  是数域  $\mathbb{F}$  上一个**向量空间**或**线性空间**.

今后为了方便, 将始终取定  $\mathbb{F}$  为实数域  $\mathbb{R}$ .

注意到, 在三维空间中, 选取三个不共面的向量作为基, 进而得到仿射坐标系. 为此, 根据定理5.4, 把共线共面等几何概念, 用等价的“线性相关”的代数概念代替, 并将线性相关的概念推广到向量空间  $V$  上任何一组向量, 给出下列线性相关一般的定义.

**定义 5.14** 设  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k$  为  $V$  中一组向量, 若存在不全为零的实数  $\lambda_1, \lambda_2, \dots, \lambda_k$ , 使得

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \dots + \lambda_k\mathbf{a}_k = 0,$$

则称这组向量**线性相关**. 否则称为**线性无关**, 因此, 对线性无关的向量组, 若

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \dots + \lambda_k\mathbf{a}_k = 0,$$

必有

$$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0.$$

显然, 对一组线性相关的向量, 增加若干个向量后仍然线性相关, 对一组线性无关的向量, 减少若干向量后仍然线性无关. 那么, 自然会问是否存在一组线性无关的向量, 使得增加任何一个新的向量, 就不再线性无关了.

**定义 5.15** 设  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  是  $V$  中有限个向量构成的一组线性无关向量, 如果增加任意一个向量  $\mathbf{x}$  后, 向量组  $\{\mathbf{x}, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  不再是线性无关的 (因此线性相关), 则称  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  为**极大线性无关组**.

**定理 5.16** 设  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  是  $V$  的极大线性无关组, 则  $V$  中任意一个向量  $\mathbf{x}$  都可以唯一地表示为  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  的线性组合

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n = \sum_{i=1}^n x_i\mathbf{e}_i.$$

**证明** 因为  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  是  $V$  中极大线性无关组, 因此添加  $\mathbf{x}$  后形成的向量组线性相关, 即存在不全为零的数  $\lambda_0, \lambda_1, \dots, \lambda_n$  使得

$$\lambda_0\mathbf{x} + \lambda_1\mathbf{e}_1 + \dots + \lambda_n\mathbf{e}_n = 0,$$

显然  $\lambda_0 \neq 0$ , 否则  $\lambda_1, \dots, \lambda_n$  不全为零, 并且  $\lambda_1\mathbf{e}_1 + \dots + \lambda_n\mathbf{e}_n = 0$ , 这与  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  线性无关矛盾. 等式两边除以  $\lambda_0$  就得到定理的结论. 若

$$\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i = \sum_{i=1}^n x'_i\mathbf{e}_i,$$

则

$$\sum_{i=1}^n (x_i - x'_i)\mathbf{e}_i = 0,$$

因  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  线性无关, 所以  $x_i = x'_i, i = 1, \dots, n$ , 唯一性得证.

**定义 5.17** 若  $V$  中存在由有限个向量  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  组成的极大线性无关组, 则称其为  $V$  的**基**, 其中的向量  $\mathbf{e}_i, i = 1, \dots, n$  称为**基向量**;  $n$  称为  $V$  的**维数**, 记为  $\dim V = n$ ;  $V$  称为  $n$  维**向量空间**.  $V$  中的元素称为  $n$  维**向量**. 定理 5.16 中对任何向量  $\mathbf{x}$  的线性组合中的数组  $(x_1, x_2, \dots, x_n)$  称为向量  $\mathbf{x}$  在基  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  下的**坐标**.

在定义中, 必须说明: 如果另有一组极大线性无关组  $\{e'_1, \dots, e'_m\}$ , 要证明  $n = m$  才能使维数的定义具有合理性. 为此要用到线性方程组的有关理论. 这里我们仅从具体例子加以说明.

若分别有两个向量和三个向量组成的极大线性无关组  $\{e_1, e_2\}$  和  $\{e'_1, e'_2, e'_3\}$ , 根据定理5.16, 向量  $e'_1, e'_2, e'_3$  可表示为  $e_1, e_2$  的线性组合:

$$e'_1 = a_1 e_1 + a_2 e_2, \quad e'_2 = b_1 e_1 + b_2 e_2, \quad e'_3 = c_1 e_1 + c_2 e_2.$$

若有  $\lambda_1 e'_1 + \lambda_2 e'_2 + \lambda_3 e'_3 = 0$ , 则

$$(a_1 \lambda_1 + b_1 \lambda_2 + c_1 \lambda_3) e_1 + (a_2 \lambda_1 + b_2 \lambda_2 + c_2 \lambda_3) e_2 = 0,$$

因为  $\{e_1, e_2\}$  线性无关, 所以有

$$\begin{cases} a_1 \lambda_1 + b_1 \lambda_2 + c_1 \lambda_3 = 0, \\ a_2 \lambda_1 + b_2 \lambda_2 + c_2 \lambda_3 = 0. \end{cases}$$

这是两个方程、三个变量的三元一次方程组, 因此必有非零解  $\lambda_1, \lambda_2, \lambda_3$ , 使得  $\lambda_1 e'_1 + \lambda_2 e'_2 + \lambda_3 e'_3 = 0$ , 这与  $\{e'_1, e'_2, e'_3\}$  线性无关矛盾. 因此

**定理 5.18** 若向量空间  $V$  存在有限个向量组成的极大线性无关组, 那么  $V$  中任何极大线性无关组中向量的个数是相等的.

至此, 我们完成了从具体的三维空间到一般向量空间的抽象.

**例 5.5.1** 设实数域  $\mathbb{R}$  上由下列有序数组组成的集合

$$E_n = \{\mathbf{a} = (a_1, a_2, \dots, a_n) \mid a_1, a_2, \dots, a_n \in \mathbb{R}\},$$

其中  $(0, 0, \dots, 0)$  为 0 向量. 对  $E_n$  中任意两个元素 (称为向量)

$$\mathbf{a} = (a_1, a_2, \dots, a_n) \in E_n, \quad \mathbf{b} = (b_1, b_2, \dots, b_n) \in E_n$$

和实数  $\lambda \in \mathbb{R}$  定义加法和数乘如下:

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$$

$$\lambda \mathbf{a} = (\lambda a_1, \lambda a_2, \dots, \lambda a_n)$$

不难验证  $E_n$  满足向量空间的定义, 并且以

$$\mathbf{e}_1 = (1, 0, \dots, 0),$$

$$\mathbf{e}_2 = (0, 1, \dots, 0),$$

.....

$$\mathbf{e}_n = (0, 0, \dots, 1)$$

为一组基向量, 即

$$\mathbf{a} = (a_1, a_2, \dots, a_n) = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \dots + a_n\mathbf{e}_n.$$

因此  $E_n$  是  $n$  维向量空间, 称为  $n$  维数组向量空间或简称为数组空间.

**例 5.5.2** 设  $F_n[x]$  表示实数域上次数不超过  $n$  的多项式的全体

$$F_n[x] = \{f(x) = a_0 + a_1x + \dots + a_nx^n \mid a_0, a_1, \dots, a_n \in \mathbb{F}\},$$

那么在多项式的加法和数乘定义下,  $F_n[x]$  是实数域上的向量空间, 并以

$$1, x, x^2, \dots, x^n \in F_n[x]$$

为基向量, 因此是  $n+1$  维的向量空间.

## 2° 向量空间的同构

**定义 5.19** 设  $V_1, V_2$  是数域  $\mathbb{R}$  上两个向量空间, 如果存在 1-1 映射

$$\sigma : V_1 \longrightarrow V_2$$

满足

- (1)  $\sigma(\mathbf{a} + \mathbf{b}) = \sigma(\mathbf{a}) + \sigma(\mathbf{b}), \mathbf{a}, \mathbf{b} \in V_1;$
- (2)  $\sigma(\lambda\mathbf{a}) = \lambda\sigma(\mathbf{a}), \lambda \in \mathbb{R}, \mathbf{a} \in V_1,$

那么称向量空间  $V_1$  与  $V_2$  同构,  $\sigma$  称为同构映射. 当  $V_1 = V_2$  时,  $\sigma$  称为自同构.

根据定义, 显然有

$$\begin{aligned} \sigma(0) &= 0, \\ \sigma(-\mathbf{a}) &= -\sigma(\mathbf{a}), \\ \sigma(\lambda_1\mathbf{a}_1 + \dots + \lambda_m\mathbf{a}_m) &= \lambda_1\sigma(\mathbf{a}_1) + \dots + \lambda_m\sigma(\mathbf{a}_m), \end{aligned}$$

这里 0 分别表示  $V_1$  和  $V_2$  中的零向量,  $\lambda_1, \dots, \lambda_m$  是任意实数,  $\mathbf{a}_1, \dots, \mathbf{a}_m \in V_1$ .

**定理 5.20** 设  $V_1, V_2, V_3$  是同一数域上三个向量空间, 则

- (1) 若  $\sigma : V_1 \longrightarrow V_2$  的同构映射, 那么  $\sigma^{-1} : V_2 \longrightarrow V_1$  的同构映射.
- (2) 若  $V_1$  与  $V_2$  同构,  $V_2$  与  $V_3$  同构, 则  $V_1$  与  $V_3$  同构.
- (3) 若  $V_1$  与  $V_2$  同构, 则同构映射把  $V_1$  中的任意线性无关组  $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  映射到  $V_2$  中的线性无关组  $\{\sigma(\mathbf{a}_1), \dots, \sigma(\mathbf{a}_m)\}$ .
- (4)  $V_1$  与  $V_2$  同构, 当且仅当  $\dim V_1 = \dim V_2$  (即维数相等).

因此, 两个向量空间的同构表明空间代数结构是一样的, 并不在意两个空间中的向量的具体含义.

**证明** (1) 设  $\sigma : V_1 \rightarrow V_2$  是同构映射, 因此是 1-1 映射, 所以映射  $\sigma^{-1} : V_2 \rightarrow V_1$  存在. 对任意  $\mathbf{a}', \mathbf{b}' \in V_2$ , 记

$$\mathbf{a} = \sigma^{-1}(\mathbf{a}'), \quad \mathbf{b} = \sigma^{-1}(\mathbf{b}'),$$

那么  $\mathbf{a}' = \sigma(\mathbf{a}), \mathbf{b}' = \sigma(\mathbf{b})$ , 推得  $\mathbf{a}' + \mathbf{b}' = \sigma(\mathbf{a} + \mathbf{b})$ , 所以

$$\sigma^{-1}(\mathbf{a}' + \mathbf{b}') = \mathbf{a} + \mathbf{b} = \sigma^{-1}(\mathbf{a}') + \sigma^{-1}(\mathbf{b}').$$

对任意实数  $\lambda$ ,

$$\sigma^{-1}(\lambda \mathbf{a}') = \sigma^{-1}(\lambda \sigma(\mathbf{a})) = \sigma^{-1}(\sigma(\lambda \mathbf{a})) = \lambda \mathbf{a} = \lambda \sigma^{-1}(\mathbf{a}')$$

(2) 设  $\sigma_1 : V_1 \rightarrow V_2, \sigma_2 : V_2 \rightarrow V_3$ , 则  $\sigma_2 \circ \sigma_1 : V_1 \rightarrow V_3$  是 1-1 映射, 且对任意  $\mathbf{a}, \mathbf{b} \in V_1$  有

$$\begin{aligned}\sigma_2 \circ \sigma_1(\mathbf{a} + \mathbf{b}) &= \sigma_2(\sigma_1(\mathbf{a}) + \sigma_1(\mathbf{b})) = \sigma_2(\sigma_1(\mathbf{a})) + \sigma_2(\sigma_1(\mathbf{b})) \\ &= \sigma_2 \circ \sigma_1(\mathbf{a}) + \sigma_2 \circ \sigma_1(\mathbf{b}) \\ \sigma_2 \circ \sigma_1(\lambda \mathbf{a}) &= \sigma_2(\sigma_1(\lambda \mathbf{a})) = \sigma_2(\lambda \sigma_1(\mathbf{a})) = \lambda \sigma_2(\sigma_1(\mathbf{a})) \\ &= \lambda \sigma_2 \circ \sigma_1(\mathbf{a})\end{aligned}$$

(3) 设  $\sigma : V_1 \rightarrow V_2$  是同构映射,  $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  是  $V_1$  的一组线性无关向量. 那么对任意实数  $\lambda_1, \dots, \lambda_m$ , 如果

$$\lambda_1 \sigma(\mathbf{a}_1) + \dots + \lambda_m \sigma(\mathbf{a}_1) = 0,$$

就有

$$\sigma(\lambda_1 \mathbf{a}_1 + \dots + \lambda_m \mathbf{a}_1) = 0,$$

因此

$$\lambda_1 \mathbf{a}_1 + \dots + \lambda_m \mathbf{a}_1 = 0,$$

所以得  $\lambda_1 = \dots = \lambda_m = 0$ , 即  $V_2$  中向量组  $\{\sigma(\mathbf{a}_1), \dots, \sigma(\mathbf{a}_m)\}$  线性无关.

(4) 设  $\sigma : V_1 \rightarrow V_2$  是同构映射,  $\dim V_1 = n, \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  是  $V_1$  的一组基, 根据(3),  $V_2$  中向量组  $\{\sigma(\mathbf{e}_1), \dots, \sigma(\mathbf{e}_n)\}$  是线性无关组, 因此  $\dim V_2 \geq \dim V_1$ , 再结合 (1),  $\sigma^{-1} : V_2 \rightarrow V_1$  是同构映射, 可证  $\dim V_1 \geq \dim V_2$ , 因此  $\dim V_1 = \dim V_2$ .

反之若  $\dim V_1 = \dim V_2$ , 分别取  $V_1$  的一组基  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  和  $V_2$  的一组基  $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ , 按下列方式定义一个映射, 先令

$$\sigma(\mathbf{e}_1) = \mathbf{e}'_1, \dots, \sigma(\mathbf{e}_n) = \mathbf{e}'_n$$

对任意的  $\mathbf{a} = \lambda_1 \mathbf{e}_1 + \dots + \lambda_n \mathbf{e}_n$ , 定义

$$\sigma(\mathbf{a}) = \lambda_1 \sigma(\mathbf{e}_1) + \dots + \lambda_n \sigma(\mathbf{e}_n) = \lambda_1 \mathbf{e}'_1 + \dots + \lambda_n \mathbf{e}'_n,$$

可以验证  $\sigma$  是同构映射.  $\square$

**例 5.5.3** 实数域上任何  $n$  维向量空间  $V$  与例 5.5.1 中的数组向量空间  $E$  是同构的. 当选定  $V$  的一组基  $\{e_1, \dots, e_n\}$  后, 同构映射就是

$$\sigma : \mathbf{a} \longrightarrow (a_1, \dots, a_n),$$

其中  $(a_1, \dots, a_n)$  是  $\mathbf{a} = a_1 e_1 + \dots + a_n e_n$  在基下的坐标. 反之任何一组实数作为坐标唯一对应  $V$  中一个向量.

这个例子说明, 从同构的角度看, 实数域上  $n$  维向量空间本质上与  $n$  维数组构成的空间无异.

**例 5.5.4** 实数域上次数不超过  $n$  的多项式全体构成的向量空间与  $n+1$  维的数组向量空间  $E$  是同构的.

### 3° 向量空间的子空间

**定义 5.21** 设  $V$  是实数域上向量空间,  $W$  是  $V$  的非空子集, 如果  $W$  对  $V$  中的加法和数乘运算保持封闭, 综合而言即是

$$\lambda \mathbf{a} + \mu \mathbf{b} \in W \quad \mathbf{a}, \mathbf{b} \in W, \lambda, \mu \in \mathbb{R},$$

那么  $W$  是  $V$  的子空间.

显然, 若取  $V$  的任意子集  $S$  (不一定是子空间), 则

$$W = \{\lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \dots + \lambda_l \mathbf{a}_l \mid \lambda_i \in \mathbb{R}, \mathbf{a}_i \in S, i = 1, \dots, l, l \in \mathbb{N}\}$$

是  $V$  的子空间, 称这样的子空间为由  $V$  的子集  $S$  生成的子空间, 记为  $W = \langle S \rangle$ . 特别由  $V$  中若干个线性无关的向量  $e_1, \dots, e_l$  生成的子空间为  $W = \langle e_1, \dots, e_l \rangle$ ,  $W$  的维数是  $l$ .

**例 5.5.5** 在  $n$  次多项式构成的  $n+1$  维向量空间  $F_n[x]$  中, 设

$$W = \{p(x) \mid p(-x) = p(x), p(x) \in F_n[x]\},$$

不难验证  $W$  是  $F_n[x]$  的子空间, 该子空间也是线性无关向量

$$1, x^2, x^4, \dots, x^{2m}$$

生成的子空间, 这里  $m = \left[ \frac{n}{2} \right]$ ,  $[a]$  表示不超过  $a$  的最大整数. 因此  $W$  的维数是  $m+1$ .

**4° 向量空间的内积** 在三维空间的具体模型中, 内积是通过几何的方式定义的(见定义 5.5). 但从代数上看, 内积实际上是一个满足性质 5.6 的两个向量对应一个实数的函数关系, 因此将具体内积的代数性质抽象出来, 我们给出下列定义:

**定义 5.22** 设  $V$  是实数域上向量空间, 如果  $V$  中任意两个向量  $\mathbf{a}, \mathbf{b} \in V$ , 对应一个实数  $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}$ , 满足:

(1) **对称性:** 对任意两个向量  $\mathbf{a}, \mathbf{b} \in V$ , 有

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a});$$

(2) **线性性:** 对任意两个实数  $\mu, \nu$  和任意三个向量  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$ , 有

$$(\mu\mathbf{a} + \nu\mathbf{b}, \mathbf{c}) = \mu(\mathbf{a}, \mathbf{c}) + \nu(\mathbf{b}, \mathbf{c});$$

(3) **正定性:** 对任意一个向量  $\mathbf{a} \in V$ , 有

$$(\mathbf{a}, \mathbf{a}) \geq 0, \text{ 等号成立当且仅当 } \mathbf{a} = 0.$$

那么称  $(\mathbf{a}, \mathbf{b})$  为向量  $\mathbf{a}$  和  $\mathbf{b}$  的内积. 实数域上定义了内积的向量空间  $V$  称为欧氏空间.

**定理 5.23** (Cauchy-Schwarz 不等式) 设  $V$  是欧氏空间,  $(\cdot, \cdot)$  是  $V$  的一个内积, 则对任意两个向量  $\mathbf{a}, \mathbf{b} \in V$ , 有

$$|(\mathbf{a}, \mathbf{b})| \leq \sqrt{(\mathbf{a}, \mathbf{a})(\mathbf{b}, \mathbf{b})}.$$

**证明** 对任意实数  $\lambda$  和  $\mathbf{a}, \mathbf{b} \in V$ , 有

$$0 \leq (\lambda\mathbf{a} + \mathbf{b}, \lambda\mathbf{a} + \mathbf{b}) = (\mathbf{a}, \mathbf{a})\lambda^2 + 2(\mathbf{a}, \mathbf{b})\lambda + (\mathbf{b}, \mathbf{b}),$$

因此根据  $\lambda$  的二次多项式判别式即可证得定理. □

根据 Cauchy-Schwarz 不等式, 我们就可以定义任意向量  $\mathbf{a}$  的大小(称为模)

$$|\mathbf{a}| = \sqrt{(\mathbf{a}, \mathbf{a})},$$

以及通过

$$\cos \theta = \frac{(\mathbf{a}, \mathbf{b})}{|\mathbf{a}||\mathbf{b}|}$$

定义任意两个非零向量  $\mathbf{a}, \mathbf{b} \in V$  的夹角  $\theta$ . 因此若两个非零向量的内积为零  $(\mathbf{a}, \mathbf{b}) = 0$ , 则称  $\mathbf{a}$  和  $\mathbf{b}$  相互正交或相互垂直.

从 Cauchy-Schwarz 不等式, 还可以得到下列不等式

**推论 5.24** (三角不等式)

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|, \quad \mathbf{a}, \mathbf{b} \in V.$$

**证明**

$$\begin{aligned} |\mathbf{a} + \mathbf{b}|^2 &= (\mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2(\mathbf{a}, \mathbf{b}) \\ &\leq |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|\mathbf{a}||\mathbf{b}| = (|\mathbf{a}| + |\mathbf{b}|)^2, \end{aligned}$$

两边开方就得到三角不等式.

**推论 5.25** (平行四边形等式)

$$|\mathbf{a} + \mathbf{b}|^2 + |\mathbf{a} - \mathbf{b}|^2 = 2(|\mathbf{a}|^2 + |\mathbf{b}|^2), \quad \mathbf{a}, \mathbf{b} \in V.$$

**证明** 直接由下列两式相加就可得到

$$|\mathbf{a} + \mathbf{b}|^2 = (\mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2(\mathbf{a}, \mathbf{b})$$

$$|\mathbf{a} - \mathbf{b}|^2 = (\mathbf{a} - \mathbf{b}, \mathbf{a} - \mathbf{b}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2(\mathbf{a}, \mathbf{b})$$

现在有一个疑问, 对于  $n$  维向量空间, 是否可以定义一个内积? 下面的定理给出构造性的答复.

**定理 5.26** 设  $V$  是  $n$  维向量空间,  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  是  $V$  的一组基, 对任意两个向量  $\mathbf{a}, \mathbf{b} \in V$ :

$$\mathbf{a} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \cdots + a_n \mathbf{e}_n, \quad \mathbf{b} = b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + \cdots + b_n \mathbf{e}_n,$$

定义

$$(\mathbf{a}, \mathbf{b}) = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

不难验证它是  $V$  上的一个内积.

根据这个内积, 基向量是两两正交的:

$$(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

(1) 向量的模长:

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

(2) 两向量的夹角:

$$\cos \theta = \frac{a_1 b_1 + a_2 b_2 + \cdots + a_n b_n}{\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} \sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}}.$$

而且对任何实数  $a_1, \dots, a_n$  和  $b_1, \dots, b_n$ , 有

(3) Cauchy-Schwarz 不等式:

$$(a_1 b_1 + \cdots + a_n b_n)^2 \leq (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2).$$

(4) 三角不等式

$$\sqrt{(a_1 + b_1)^2 + \cdots + (a_n + b_n)^2} \leq \sqrt{a_1^2 + \cdots + a_n^2} + \sqrt{b_1^2 + \cdots + b_n^2}.$$

(5) 平行四边形等式:

$$\begin{aligned}(a_1 + b_1)^2 + \cdots + (a_n + b_n)^2 + (a_1 - b_1)^2 + \cdots + (a_n - b_n)^2 \\ = 2((a_1^2 + \cdots + a_n^2) + (b_1^2 + \cdots + b_n^2)).\end{aligned}$$

当然, 向量空间上的内积不是唯一的, 《线性代数》将会有详细介绍, 不再赘述.

## §5.6 线性方程组的解

所谓线性方程组是指包含  $n$  个未知量  $x_1, x_2, \dots, x_n$  并由  $m$  个一次方程组成的方程组

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots, \cdots, \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{array} \right. \quad (1)$$

其中系数  $a_{ij}, i = 1, \dots, m, j = 1, \dots, n$  以及  $b_i, i = 1, \dots, m$  是给定的实数. 当所有的  $b_i = 0, i = 1, \dots, m$  时, 称对应的方程组为齐次线性方程组, 否则称为非齐次线性方程组. 把方程组的系数按下列方式组成矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

称为  $m \times n$  矩阵, 特别, 当  $m = n$  时,  $n \times n$  矩阵  $A$  称为  $n$  阶方阵. 方程 (1) 也可简写为

$$Ax = \mathbf{b}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

这里把向量写成列的形式称为列向量, 而通常写成行的形式称为行向量.

我们不打算给出齐次和非齐次线性方程组解的结构的一般理论, 而是通过一些例子做出说明.

1° 系数矩阵为方阵的非齐次线性方程组的解

例 5.6.1 考虑下列线性方程组

$$\begin{cases} x + y + z = b_1, \\ x + 2y + 3z = b_2, \\ 2x - z = b_3. \end{cases} \quad (2)$$

这里  $b_1, b_2, b_3$  是任意给定实数. 用矩阵表示, 方程 (2) 为

$$A\mathbf{x} = \mathbf{b}, \quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 0 & -1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

其中方程组的系数矩阵是 3 阶方阵.

从几何上看, 在  $Oxyz$  坐标系中, 方程组中三个方程

$$x + y + z = b_1, \quad x + 2y + 3z = b_2, \quad 2x - z = b_3$$

分别表示三张平面, 因此方程组的解即是三张平面的交点的坐标  $(x, y, z)$ . 而三张平面相交于一点当且仅当三张平面的法向量

$$\mathbf{v}_1 = (1, 1, 1), \quad \mathbf{v}_2 = (1, 2, 3), \quad \mathbf{v}_3 = (2, 0, -1)$$

不共面. 注意到  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  正是矩阵  $A$  的三行构成的行向量.

从代数上看, 把矩阵  $A$  的三列分别记为三个列向量

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix},$$

那么方程组可以表示为

$$x\mathbf{a}_1 + y\mathbf{a}_2 + z\mathbf{a}_3 = \mathbf{b},$$

也就是  $\mathbf{b}$  能够表示为  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  的线性组合, 组合系数  $x, y, z$  即是方程组的解. 因此方程组对任何  $\mathbf{b}$  有唯一解, 当且仅当  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  线性无关.

不管是从几何上看, 还是从代数上看, 不管是矩阵  $A$  的行构成的行向量不共面, 还是列向量线性无关, 它们的共同特点是矩阵的行列式不为零 (参见§5.2 的最后关于行列式部分)!

$$\det A = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 0 & -1 \end{vmatrix} = 1 \neq 0.$$

因此系数矩阵是方阵的非齐次线性方程组有唯一解, 当且仅当系数矩阵的行列式不等于零. 当然, 对于  $n$  阶方阵  $A$ , 需要事先定义  $A$  的行列式, 这里就不再展开了.

2° 系数矩阵为  $m \times n$  矩阵的线性方程组的解

首先考虑齐次线性方程组, 即方程组 (1) 中  $b_1 = b_2 = \cdots = b_m = 0$ . 记齐次线性方程组所有解构成的集合为

$$V = \{(x_1, x_2, \dots, x_n)\},$$

不难验证  $V$  满足向量空间的定义. 例如  $(0, 0, \dots, 0)$  显然是方程组的一组解; 若  $(x_1, x_2, \dots, x_n), (x'_1, x'_2, \dots, x'_n)$  分别是方程组的解, 那么  $(\lambda x_1, \lambda x_2, \dots, \lambda x_n)$  以及  $(x_1 + x'_1, x_2 + x'_2, \dots, x_n + x'_n)$  也是解, 这里  $\lambda \in \mathbb{R}$ . 其它验证读者可自行完成.

显然  $V$  是例5.5.1 中数组空间  $E_n$  的子空间, 因此  $\dim V \leq n$ . 设  $e_1, \dots, e_m$  是  $V$  的一组基向量, 它们也是齐次线性方程组的一组解, 称为**基本解组**, 那么线性方程组的通解就是这组基本解组的线性组合

$$\mathbf{x} = \lambda_1 e_1 + \cdots + \lambda_m e_m.$$

解空间  $V$  的维数与方程组的系数  $a_{ij}$  密切相关, 具体情况将通过下面第一个例子加以说明.

对于非齐次线性方程组 (1), 它的任意两个解的差满足对应的齐次线性方程组, 即

$$\text{若 } A\mathbf{x} = \mathbf{b}, A\mathbf{x}' = \mathbf{b}, \text{ 则 } A(\mathbf{x} - \mathbf{x}') = 0.$$

因此, 若  $\mathbf{x}_0 = (x_1^0, x_2^0, \dots, x_n^0)$  是非齐次的一个特解,  $e_1, \dots, e_m$  是对应齐次线性方程的基本解组, 那么非齐次方程 (1) 的通解为

$$\mathbf{x} = \mathbf{x}_0 + \lambda_1 e_1 + \cdots + \lambda_m e_m, \quad \lambda_1, \dots, \lambda_m \in \mathbb{R}.$$

当然, 并非所有非齐次线性方程组的解都存在, 这一点将通过下面的第二个例子给予解释.

## 例 5.6.2 设齐次线性方程组

$$\left\{ \begin{array}{l} x_1 + x_2 + x_3 + x_4 = 0, \\ x_1 + 2x_2 + 3x_3 - 2x_4 = 0, \\ x_1 - x_3 + 4x_4 = 0. \end{array} \right. \quad (3)$$

利用消元法可得方程组的通解为

$$\mathbf{x} = (\lambda - 4\mu, -2\lambda + 3\mu, \lambda, \mu) = \lambda e_1 + \mu e_2,$$

其中

$$e_1 = (1, -2, 1, 0), \quad e_2 = (-4, 3, 0, 1).$$

是两个线性无关的解, 构成方程的基本解组. 因此解空间  $V$  是 4 维数组空间  $E_4$  的 2 维子空间.

现在考察方程组的系数矩阵

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & -2 \\ 1 & 0 & -1 & 4 \end{pmatrix}$$

其中的列向量为

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \quad \mathbf{a}_4 = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix},$$

此时方程组可表示为

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 + x_4\mathbf{a}_4 = 0,$$

不难看出向量组  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$  中,  $\mathbf{a}_1, \mathbf{a}_2$  是极大线性无关组, 且

$$\mathbf{a}_3 = 2\mathbf{a}_2 - \mathbf{a}_1, \quad \mathbf{a}_4 = 4\mathbf{a}_1 - 3\mathbf{a}_2.$$

因此方程简化为

$$(x_1 - x_3 + 4x_4)\mathbf{a}_1 + (x_2 + 2x_3 - 3x_4)\mathbf{a}_2 = 0,$$

因为  $\mathbf{a}_1, \mathbf{a}_2$  线性无关, 所以有

$$x_1 - x_3 + 4x_4 = 0, \quad x_2 + 2x_3 - 3x_4 = 0.$$

由此看出, 未知变量  $x_3, x_4$  可以取任意实数, 不妨设  $x_3 = \lambda, x_4 = \mu$ . 那么  $x_1, x_2$  分别为  $x_1 = \lambda - 4\mu, x_2 = -2\lambda + 3\mu$ . 这样就得到方程组的通解. 解空间的维数等于  $E_4$  的维数减去系数矩阵中列向量极大线性无关组的个数.

对于一般情况, 记

$$\text{rank}(A) = A \text{ 的列向量中极大线性无关组的个数},$$

那么齐次线性方程组的解空间的维数满足

$$\dim V = n - \text{rank}(A).$$

有关证明可参见任何一本线性代数教材.

**例 5.6.3** 考虑方程组 (3) 的非齐次情形

$$\left\{ \begin{array}{l} x_1 + x_2 + x_3 + x_4 = b_1, \\ x_1 + 2x_2 + 3x_3 - 2x_4 = b_2, \\ x_1 - x_3 + 4x_4 = b_3. \end{array} \right. \quad (4)$$

方程 (4) 有解  $(x_1^0, x_2^0, x_3^0, x_4^0)$ , 当且仅当

$$x_1^0 \mathbf{a}_1 + x_2^0 \mathbf{a}_2 + x_3^0 \mathbf{a}_3 + x_4^0 \mathbf{a}_4 = \mathbf{b}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

这里  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$  同例 5.6.2,  $\mathbf{a}_1, \mathbf{a}_2$  线性无关, 因此

$$(x_1^0 - x_3^0 + 4x_4^0)\mathbf{a}_1 + (x_2^0 + 2x_3^0 - 3x_4^0)\mathbf{a}_2 = \mathbf{b},$$

也就意味着 (4) 有解当且仅当  $\mathbf{b}$  能表示为  $\mathbf{a}_1, \mathbf{a}_2$  的线性组合. 或者说向量组  $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$  与向量组  $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{b}\}$  中极大线性无关组的个数是一样的. 把两者分别作为矩阵  $A$  和  $\tilde{A}$  的列向量

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & -2 \\ 1 & 0 & -1 & 4 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1 & 1 & 1 & 1 & b_1 \\ 1 & 2 & 3 & -2 & b_2 \\ 1 & 0 & -1 & 4 & b_3 \end{pmatrix}$$

那么 (4) 有解的充分必要条件是两个矩阵的秩相等:

$$\text{rank}(A) = \text{rank}(\tilde{A}).$$

矩阵  $\tilde{A}$  称为矩阵  $A$  的增广矩阵. 对一般的线性方程组 (1), 它的增广矩阵就是在系数矩阵  $A$  中再添加一列  $\mathbf{b}$ .

## 第 5 讲习题

1. 已知空间中三个向量满足  $\mathbf{a} + \mathbf{b} + \mathbf{c} = 0$ , 试证:

$$\mathbf{a} \times \mathbf{b} = \mathbf{b} \times \mathbf{c} = \mathbf{c} \times \mathbf{a}.$$

反之, 若上式成立, 且  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  不共线, 证明  $\mathbf{a} + \mathbf{b} + \mathbf{c} = 0$ .

2. 证明下列两条直线是异面直线 (即不平行也不相交的两条直线)

$$\begin{cases} x + y - z - 1 = 0, \\ 2x + y - z - 2 = 0 \end{cases} \quad \text{和} \quad \begin{cases} x + 2y - z - 2 = 0, \\ x + 2y + 2z + 4 = 0. \end{cases}$$

并求两条直线的距离 (即两直线的公垂线段之长度).

3. 考虑两平面  $\Pi_1$  和  $\Pi_2$  之间的投影映射 (见第 1 讲第 4 节), 试证明  $\Pi_1$  上的圆可投影为  $\Pi_2$  上的椭圆、抛物线和双曲线.

提示: 不妨取  $\Pi_1$  为  $Oxz$  平面,  $\Pi_2$  为  $Oxy$  平面. 取  $Oxz$  平面上的圆

$$x^2 + (z - c)^2 = 1,$$

以及  $Q(0, -1, \lambda)$  作为投影的中心点. 给出圆上任意一点  $P(x, 0, z)$  在  $Oxy$  平面上的投影  $P'(x', y', 0)$ , 求出  $P'(x', y', 0)$  满足的方程, 并讨论  $\lambda$  的取值.

4. 设  $\mathbf{e}_1 = \mathbf{i} + \mathbf{j}$ ,  $\mathbf{e}_2 = \mathbf{j} + \mathbf{k}$  是  $\mathbb{R}^3$  中两个向量. 试证明

$$V = \{\alpha \mathbf{e}_1 + \beta \mathbf{e}_2 \mid \alpha, \beta \in \mathbb{R}\}$$

是 3 维向量空间中的 2 维子空间.

5. 求下列齐次线性方程组的通解

$$x + z = 0,$$

$$y + z = 0,$$

$$x + y + 2z = 0$$

并确定  $V$  的维数, 给出一组基向量.

## 第 6 讲 几何作图

所谓几何作图是指从一些已知图形出发, 仅限于用没有刻度的直尺和圆规 (以下均简称尺规) 作出新的图, 因此又称尺规作图.

### §6.1 尺规在作图中的功能

先从一个例子谈起.

**例 6.1.1** 已知直线段  $AB$ , 作出它的垂直平分线 (当然也就得到它的平分点).

具体做法是: 分别以  $A$  和  $B$  作为圆心, 以超过  $AB$  长度一半的任意长度  $r$  为半径, 用圆规作两个相交的圆, 再用直尺连接圆的两个交点, 就得到  $AB$  的垂直平分线和平分点.

类似的问题统称为作图问题. 为了方便, 今后我们把能够用尺规作出的作图问题称为是可解的, 不能作出的称为是不可解的.

但是, 并不是每个作图问题都是可解的, 例如三等分角问题, 倍立方体问题, 求作正七边形问题以及化圆成方问题等等, 是无法用圆规和直尺来完成的 (将在§6.5 讨论). 这样就产生了一个问题, 究竟哪些作图问题是可解的, 哪些是不解的. 用什么方法类判断一个作图问题可解还是不可解.

本专题主要从几何作图问题, 如何发现它背后的数学, 但是并不讨论对于一个可解的几何作图问题具体作图过程. 为此具体分析一下直尺和圆规在作图中的基本功能.

**尺规作图基本功能:**

- (1) 过两个已知点作一条直线;
- (2) 以已知点为圆心, 以已知长度为半径作一个圆;
- (3) 作两条已知直线的交点;
- (4) 作一个圆与已知圆或已知直线的交点.

功能 (1) 和 (2) 分别是直尺和圆规仅有的功能, 根据给定的条件 (点和定长) 用直尺或圆规直接完成. 后两种功能需要根据给定的条件, 用直尺和圆规找出交点. 几何作图正是按上述功能, 一个步骤接着一个步骤的过程, 也就是已知前面一步, 如何用直尺和

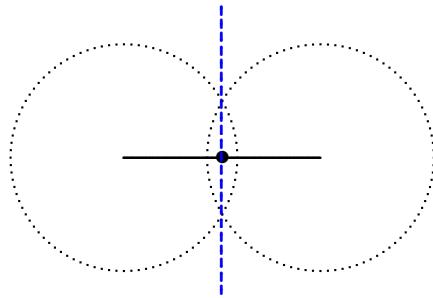


图 6.1

圆规作出下一步的图形.

根据解析几何的观点, 在平面上建立直角坐标系之后, 平面上的点可由坐标表示, 知道了点就可以作出它的坐标, 知道了坐标就可以作出点. 而点的坐标是一对实数, 因此从给定的点, 用直尺和圆规作出新的点, 就是从已知数用直尺和圆规作出新的数. 因此, 关键是把几何作图问题“翻译”成代数语言. 形象地说, 就是把几何作图问题转化为“数字化”的问题.

**例 6.1.2** 已知两条直线  $OA$  和  $OB$  相交于  $O$ , 作出它们夹角的平分线.

具体做法是: 不妨设  $OA$  和  $OB$  长度都等于 1 (若不然, 可以通过以  $O$  为圆心, 用圆规以已知的定长 1 为半径, 在两直线上截出两点, 也就是圆与直线的交点, 代替  $A$  和  $B$ ). 分别以  $A$  和  $B$  为圆心, 用圆规作两个半径相等的圆交于  $P$ . 最后用直尺连接  $O$  和  $P$ , 就得到两直线的夹角的平分线.

若在坐标系下考虑, 以  $O$  为原点, 以  $OA$  作为  $x$  轴的正方向,  $OA$  的长度作为单位长度, 建立坐标系.

设已知角  $\angle AOB = \theta$ , 则三个已知点的坐标分别为

$$O(0, 0), A(1, 0), B(\cos \theta, \sin \theta),$$

直尺和圆规能够做出的  $P$  点的坐标为

$$P\left(r \cos \frac{\theta}{2}, r \sin \frac{\theta}{2}\right),$$

就相当于根据已知点  $O, A, B$  的坐标作出  $P$  点的坐标, 或者说根据给定的数  $1, \cos \theta$  用直尺和圆规作出新的数  $\cos \frac{\theta}{2}$ .

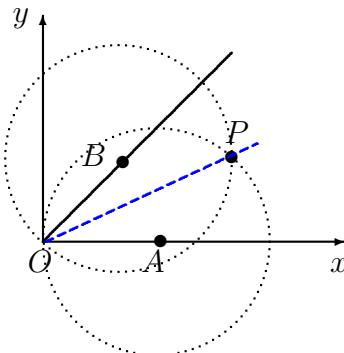


图 6.2

## §6.2 作图的代数表示

于是, 几何作图问题就转化为从已知数出发, 用直尺和圆规能够作出哪些数的问题.

### 1° 用尺规作出已知数的四则运算

即任给两个数  $a$  和  $b$ , 用尺规可以作出  $a \pm b, ra, ab$ , 和  $\frac{a}{b}$ . 这里  $r$  是任意有理数.

(a) 给定两个正实数  $a$  和  $b$  分别代表两个线段的长度, 用直尺画一条数轴, 用圆规依次向数轴正向标出距离  $OA = a$ ,  $AB = b$ , 则线段  $OB$  的长度就是  $a + b$ , 若沿  $OA$  方向接连标出距离  $a$ , 则可作长度为  $na$  的直线;

若沿相反方向标出  $AB = b$ , 则  $OB = a - b$ . 所以两个数  $a$  和  $b$  之间的加减法可由几何作图来实现. 若  $a, b$  中有负数, 仍然以其绝对值为线段长度, 只是在数轴上丈量时, 取数轴正向的反方向即可.

(b) 设  $OA = a$ , 类似图3.1, 过  $O$  作另一条直线  $OB$ , 使得  $OB = q$  也就是数轴上单位长度的  $q$  倍. 在  $OB$  上取一点  $D$  使得  $OD = 1$ . 连接  $AB$ , 并过  $D$  点作  $AB$  的平行线交  $OA$  于  $C$ , 因此三角形  $\triangle OAB$  与  $\triangle OCD$  相似, 所以  $\frac{OC}{OD} = \frac{OA}{OB}$ , 这样得  $OC = \frac{1}{q}a$  (图6.3).

再将  $OC$  扩大  $p$  倍, 就得到长度为  $\frac{p}{q}a$  的线段.

(c) 同样, 用直尺和圆规还可以实现两个数的乘法和除法:

在任意角的两边分别标出  $OA = a$ ,  $OC = b$ , 在  $OA$  上作  $OB = 1$ , 连接  $BC$  并过  $A$  点作  $BC$  的平行线交  $OC$  (或延长线) 于  $D$ , 则  $OD = ab$ .

若对任意角的两边分别标出  $OA = a$ ,  $OB = b$ , 且在  $OB$  上标出  $OD = 1$ , 过  $D$  作平行于  $AB$  的直线交  $OA$  (或延长线) 于  $C$ , 则  $OC = \frac{a}{b}$ .

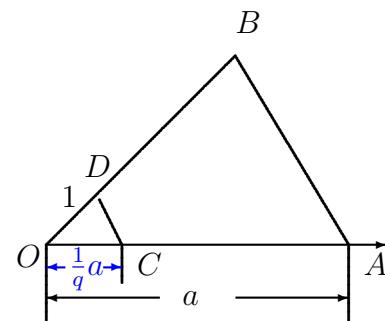


图 6.3

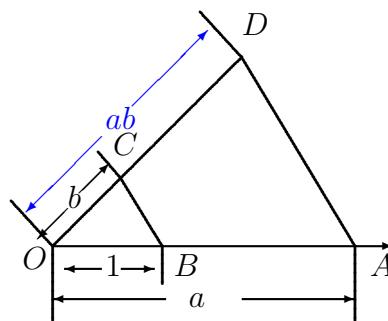


图 6.4

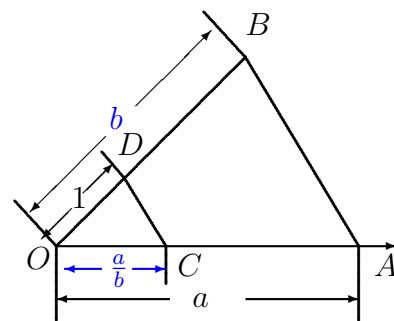


图 6.5

上面从已知数通过尺规, 逐步作出  $a \pm b, ra, ab, \frac{a}{b}$  的过程说明如下结论:

**定理 6.1** 已知一个包含 1 的数集  $S = \{1, a, b, c, \dots\}$ , 那么用尺规可作出通过  $S$  中的数之间的有限次加、减(加法逆运算)、乘、除(乘法逆运算)等运算所得到的数.

**定义 6.2** 设  $S = \{1, a, b, c, \dots\}$  是若干个数的集合, 集合中的数经过任意有限次加减乘除得到所有可能的数的集合记为  $\mathbb{F}(S)$ , 不难验证  $\mathbb{F}(S)$  满足第 3 讲中关于数域的定义3.10, 因此如同有理数集和实数集一样,  $\mathbb{F}(S)$  是一个数域, 称为由  $S$  生成的数域.

按照上述定义, 定理6.1 表明, 若已知  $S$  中的数, 通过尺规可以作出数域  $\mathbb{F}(S)$  中的任何数.

例如取  $S = \{1\}$ , 不难看出  $S$  生成的数域是有理数域  $\mathbb{F}(S) = \mathbb{Q}$ , 因此有

**推论 6.3** 从 1 出发, 通过尺规, 可以作出有理数域  $\mathbb{Q}$  中所有的有理数.

今后, 我们说“从数集  $S$  或数域  $\mathbb{F}$  出发”, 是指已经假设  $S$  或  $\mathbb{F}$  中的数是已知的, 或者已经通过直尺和圆规作出的数.

显然, 如果从任何有理数构成的集合出发, 用尺规按上述方式 (a), (b), (c), 我们得到的还是有理数, 自然要问, 能否用尺规通过其它方法作出无理数? 答案是肯定的, 而且是有决定意义的.

### 2° 用尺规作出已知数的平方根

设  $d > 0$  是已知数, 在直线上标出  $OA = d$  和  $AB = 1$ , 首先根据例 6.1.1, 作直线段  $OB$  的平分点, 并以平分点为圆心,  $OB$  为直径作一个圆. 再过  $A$  作  $OB$  的垂线交圆于  $C$ , 不难看出直角三角形  $\Delta OAC$  和  $\Delta BAC$  相似, 因此有

$$\frac{AC}{AB} = \frac{OA}{AC},$$

得

$$AC = \sqrt{d}.$$

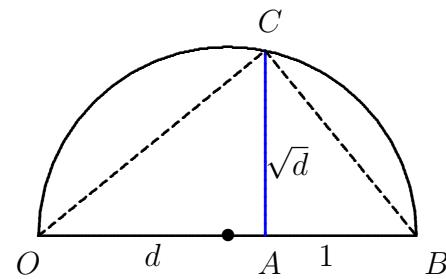


图 6.6

把  $\mathbb{F}$  和  $\sqrt{d}$  作为已知数, 进而可以用直尺和圆规作出形如

$$\alpha + \beta\sqrt{d}, \alpha, \beta \in \mathbb{F}$$

的数. 显然, 只要  $\sqrt{d} \notin \mathbb{F}$ , 上述形式的数的范围比数域  $\mathbb{F}$  更广, 并包含  $\mathbb{F}$  中的所有数. 这里需要说明的是, 如果  $\sqrt{d} \in \mathbb{F}$ , 那么通过开平方运算, 并没有扩大  $\mathbb{F}$  的范围, 只有当  $\sqrt{d} \notin \mathbb{F}$  时, 我们才真正在  $\mathbb{F}$  中增加了那些开平方运算得到的数. 例如只要取  $\sqrt{d}$  不是有理数, 那么就得到了比有理数域  $\mathbb{Q}$  范围更广的数, 它不但包含了所有有理数, 还增加了有理数的平方根.

下面这个例子, 说明如何通过作出形如  $\alpha + \beta\sqrt{d}$  的数, 解决几何作图问题.

**例 6.2.1** 单位圆内内接正十边形的作图问题是可解的.

这里, 我们只讨论上述作图问题用尺规可解还是不可解, 并不讨论怎么作出内接正十边形的具体作图过程.

假设圆心为  $O$  的单位圆内有一内接正十边形, 记某条边为  $AB$  长度为  $x$ .  $\triangle AOB$  是等腰三角形, 圆心角  $\angle AOB = 36^\circ$ , 其它两角  $\angle OAB = \angle OBA = 72^\circ$ . 作角  $\angle OAB$  的角平分线, 交  $OB$  于  $C$ , 因此  $AC$  将  $\triangle AOB$  分为两个等腰三角形  $\triangle ABC$  和  $\triangle ACO$ . 因此有

$$AB = AC = OC = x, BC = 1 - x.$$

又因为  $\triangle AOB$  与  $\triangle ABC$  相似, 这样就有

$$\frac{1}{x} = \frac{x}{1-x},$$

即  $x$  是二次方程

$$x^2 + x - 1 = 0$$

的解, 它的解为

$$x = \frac{\sqrt{5}-1}{2},$$

另一个解为负的, 可以不考虑. 这是一个有理数域添加了平方根  $\sqrt{5}$  的数, 因此是可以用尺规作出的, 这样用圆规从单位圆上一点出发, 以  $x$  为半径依次作与单位圆的交点, 再用直尺连接这些交点就得到用尺规作出的内接正十边形.

顺便指出,  $\frac{\sqrt{5}-1}{2} \approx 0.618$  正是所谓的“黄金分割”, 也就是如果矩形的宽和长取成这样的比值从审美观点看是最好的. 早在公元前 6 世纪 Pythagoras (毕达哥拉斯, 约公元前 580 - 约前 500 (490)) 学派就研究过正五边形和正十边形的作图问题, 由此推断应该触及并掌握了黄金分割.

利用简单的三角函数余弦定理, 还可以得到

$$\cos 36^\circ = \frac{\sqrt{5}+1}{4}, \cos 72^\circ = \frac{\sqrt{5}-1}{4}.$$

因此这些数也是通过尺规可以作出的.

归纳  $1^\circ$  和  $2^\circ$ , 我们得出如下结论:

仅用尺的几何作图, 可以实现已知数的加减乘除以及开平方等代数运算.

现在要问, 假如仅从  $\mathbb{F}$  出发, 除了  $\alpha + \beta\sqrt{d}$  形式的数以外, 用尺规还能不能作出其它什么数? 答案是否定的.

**定理 6.4** 若仅从一个已知数域  $\mathbb{F}$  出发, 用直尺和圆规只能作出形如  $\alpha + \beta\sqrt{d}$  的数, 其中  $\alpha, \beta \in \mathbb{F}$ .

**证明** 为了证明上述结论, 从代数的角度, 对尺规作图的基本功能分析如下:

(1) 两个已知点的直线的长度 设  $(a_1, b_1), (a_2, b_2)$  为两个已知点, 坐标  $a_1, b_1, a_2, b_2 \in \mathbb{F}$ , 那么连接  $(a_1, b_1), (a_2, b_2)$  两点线段的长度为

$$\rho = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2},$$

它是  $\mathbb{F}$  中的数经过开平方所得到的数, 因此是形如  $\alpha + \beta\sqrt{d}$ ,  $\alpha, \beta \in \mathbb{F}$  的数.

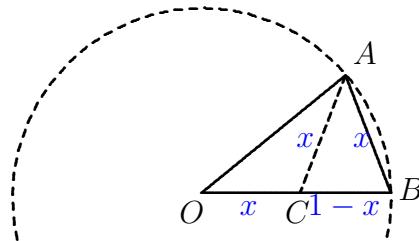


图 6.7

同时根据解析几何, 过已知两点的直线方程为

$$(b_1 - b_2)x + (a_1 - a_2)y + (a_1b_2 - a_2b_1) = 0,$$

由于  $\mathbb{F}$  是数域, 所以上述方程的系数仍然属于  $\mathbb{F}$ .

### (2) 两条已知直线的交点.

设两条不平行直线的方程为

$$ax + by + c = 0,$$

$$a'x + b'y + c' = 0,$$

其中系数  $a, b, c, a', b', c' \in \mathbb{F}$ , 那么它们的交点坐标为

$$x_0 = \frac{cb' - bc'}{ab' - ba'}, \quad y_0 = \frac{ac' - a'c}{ab' - ba'}.$$

所以交点坐标是  $\mathbb{F}$  中的数经过加减乘除所得到的, 因此仍然是  $\mathbb{F}$  中的数, 也就是用直尺作出两条相交直线的交点坐标, 仍然是  $\mathbb{F}$  中的数.

### (3) 圆以及圆和直线的交点.

设以  $(\xi, \eta)$  为圆心, 以  $r$  为半径,  $\xi, \eta, r \in \mathbb{F}$ , 那么圆的方程为

$$(x - \xi)^2 + (y - \eta)^2 = r^2$$

或表示成下列 2 次代数方程

$$x^2 + y^2 - 2\xi x - 2\eta y + \gamma = 0.$$

其中  $\gamma = \xi^2 + \eta^2 - r^2 \in \mathbb{F}$ , 因此上述二次方程的系数仍然是  $\mathbb{F}$  中的数.

圆与直线的交点坐标, 就是系数都在  $\mathbb{F}$  中的圆和直线方程的联立方程的解

$$x^2 + y^2 - 2\xi x - 2\eta y + \gamma = 0,$$

$$ax + by + c = 0,$$

从联立方程中消去变量  $y$ , 得到一个关于  $x$  的二次代数方程

$$Ax^2 + Bx + C = 0,$$

其中系数

$$A = a^2 + b^2, \quad B = 2(ac - b^2\xi + ab\eta), \quad C = c^2 + 2bc\eta + b^2\gamma,$$

都是  $\mathbb{F}$  中的数经过加减乘除得到的, 因此仍是  $\mathbb{F}$  中的数. 这样圆与直线相交就意味着这个2次代数方程有实数解

$$x = \frac{-B \pm \sqrt{\Delta}}{2A},$$

其中  $\Delta = B^2 - 4AC \in \mathbb{F}$ . 若  $\sqrt{\Delta} \in \mathbb{F}$ , 则  $x \in \mathbb{F}$ ; 若  $\sqrt{\Delta} \notin \mathbb{F}$ , 则  $x$  是形如  $\alpha + \beta\sqrt{\Delta}$ ,  $\alpha, \beta \in \mathbb{F}$  的数.

对于  $y$  也有类似结果  $y = p' + q'\sqrt{\Delta}$ ,  $p', q' \in \mathbb{F}$ . 这样我们得到的结论是用尺规作出的圆与直线的交点坐标都是形如  $\alpha + \beta\sqrt{d}$ ,  $\alpha, \beta \in \mathbb{F}$  的数.

#### (4) 两个圆相交的交点.

设系数都是  $\mathbb{F}$  中数的两个圆的代数方程为

$$x^2 + y^2 - 2\xi x - 2\eta y + \gamma = 0,$$

$$x^2 + y^2 - 2\xi' x - 2\eta' y + \gamma' = 0.$$

消去平方项得一次代数方程

$$2(\xi - \xi')x + 2(\eta - \eta')y - (\gamma - \gamma') = 0,$$

方程的系数都是  $\mathbb{F}$  中的数. 如同 (3)一样, 通过上述方程与第一个圆的方程联立得到一个二次代数方程, 通过求解该方程, 得到用尺规作出的两个圆的交点, 其坐标也是形如  $\alpha + \beta\sqrt{d}$ ,  $\alpha, \beta \in \mathbb{F}$  的数.

例如关于角平分线的问题 (例6.1.2). 不妨设所讨论的夹角为锐角  $0 < \theta < \frac{\pi}{2}$ , 因此以  $A$  和  $B$  为圆心, 以 1 为半径的两个圆有交点, 设交点坐标为  $P(x, y)$ , 那么  $x, y$  满足方程

$$(x - 1)^2 + y^2 = 1,$$

$$(x - \cos \theta)^2 + (y - \sin \theta)^2 = 1$$

或

$$x^2 - 2x + y^2 = 0,$$

$$x^2 - 2x \cos \theta + y^2 - 2y \sin \theta = 0,$$

消去平方项得

$$x(1 - \cos \theta) - y \sin \theta = 0,$$

与第一个圆的方程联立, 可解的交点坐标为

$$x = 1 + \cos \theta, \quad y = \sin \theta = \sqrt{1 - \cos^2 \theta}.$$

它们都是从已知数  $1, \cos \theta$  出发, 通过加减乘除和开方运算得到的数.

总之, 从已知数  $\mathbb{F}$  出发用尺规作出的线段长度, 交点坐标只是形如  $\alpha + \beta\sqrt{d}$ ,  $\alpha, \beta \in \mathbb{F}$  的数, 不会产生其它形式的数.  $\square$

**注记** 在几何作图中, 有时会有“以任意一点为圆心”、“以任意长度为半径”、“作任意一条直线”等作图要求. 作图中出现这样或那样“任意”的要求, 说明作图的效果与这些任意值无关. 因此我们可以选择那些任意的点或任意的长度为有理点或有理数, 那些任意的直线为一条由有理系数的 1 次代数方程表示的直线.

### §6.3 数域的扩充

我们把上节关于用直尺和圆规作图的几何问题转化为代数过程稍作总结.

首先, 建立坐标系, 在坐标系中, 从已知的条件, 用尺规作一个交点的问题转化为从已知数作出新的数的问题.

这样, 当给定一些数 (这些数的集合记为  $S$ ) 后, 我们能用直尺和圆规作出一个范围更广的数集  $\mathbb{F}$ . 这个新的数集从代数上看对加法、乘法以及它们的逆运算是封闭的, 而且包含 0元和单位元 1. 根据定义,  $\mathbb{F}$  是数域.

再接着用尺规作出  $\mathbb{F}$  中数的平方根, 这样再次把数域  $\mathbb{F}$  扩大到包含形如  $\alpha + \beta\sqrt{d}$  的数集.

最后我们证明了若把  $\mathbb{F}$  的数当作已知数, 那么用尺规能够作出的数也只能是形如  $\alpha + \beta\sqrt{d}$  的数.

取  $d \in \mathbb{F}$ , 但  $\sqrt{d} \notin \mathbb{F}$ , 并将形如  $\alpha + \beta\sqrt{d}$  的数的全体记为

$$\mathbb{F}(\sqrt{d}) = \{\alpha + \beta\sqrt{d} \mid \alpha, \beta \in \mathbb{F}\}.$$

**定理 6.5** 从几何作图角度看  $\mathbb{F}(\sqrt{d})$  中的数均可以用直尺和圆规作出, 从代数上看,  $\mathbb{F}(\sqrt{d})$  对加法、乘法以及它们的逆运算是封闭的, 而且包含 0元和单位 1. 因而  $\mathbb{F}(\sqrt{d})$  是一个包含  $\mathbb{F}$  的数域.

**证明** 仅需如下简单验证. 任取  $\mathbb{F}(\sqrt{d})$  中两个数  $\alpha + \beta\sqrt{d}, \alpha' + \beta'\sqrt{d}$ , 那么

$$(\alpha + \beta\sqrt{d}) \pm (\alpha' + \beta'\sqrt{d}) = (\alpha \pm \alpha') + (\beta \pm \beta')\sqrt{d} \in \mathbb{F}(\sqrt{d}),$$

$$(\alpha + \beta\sqrt{d}) \cdot (\alpha' + \beta'\sqrt{d}) = \alpha\alpha' + d\beta\beta' + (\alpha\beta' + \beta\alpha')\sqrt{d} \in \mathbb{F}(\sqrt{d}),$$

$$\frac{\alpha + \beta\sqrt{d}}{\alpha' + \beta'\sqrt{d}} = \frac{\alpha\alpha' - d\beta\beta'}{\alpha'^2 - d\beta'^2} + \frac{\alpha\beta' - \beta\alpha'}{\alpha'^2 - d\beta'^2}\sqrt{d} \in \mathbb{F}(\sqrt{d}).$$

这里, 由于  $\sqrt{d} \notin \mathbb{F}$ , 所以第三式中的分母  $\alpha'^2 - d\beta'^2 \neq 0$ , 否则就有  $\sqrt{d} = \pm \frac{\alpha'}{\beta'} \in \mathbb{F}$ , 这与  $\sqrt{d} \notin \mathbb{F}$  相矛盾, 因此第三式中除法是有意义的.  $\square$

**定义 6.6** 称数域  $\mathbb{F}(\sqrt{d})$  为数域  $\mathbb{F}$  的扩域,  $\mathbb{F}$  为  $\mathbb{F}(\sqrt{d})$  的子域 .

现在, 从任何一个已知数集  $S$  出发, 用尺规作出数域  $\mathbb{F}_0 = \mathbb{F}(S)$ , 取这样的  $d_1 \in \mathbb{F}_0$ , 用尺规作出它的平方根, 并使得  $\sqrt{d_1} \notin \mathbb{F}_0$ , 这样我们就得到  $\mathbb{F}_0$  的扩域, 记为

$$\mathbb{F}_1 = \{\alpha + \beta\sqrt{d_1} \mid \alpha, \beta \in \mathbb{F}_0\},$$

然后从  $\mathbb{F}_1$  出发, 重复上述过程: 取  $d_2 \in \mathbb{F}_1$ , 但  $\sqrt{d_2} \notin \mathbb{F}_1$ , 就得到  $\mathbb{F}_1$  的扩域:

$$\mathbb{F}_2 = \{\alpha + \beta\sqrt{d_2} \mid \alpha, \beta \in \mathbb{F}_1\},$$

如此下去, 我们就得到  $\mathbb{F}_{i-1}$  的扩域

$$\mathbb{F}_i = \{\alpha + \beta\sqrt{d_i} \mid \alpha, \beta \in \mathbb{F}_{i-1}\},$$

其中  $d_i \in \mathbb{F}_{i-1}$  但  $\sqrt{d_i} \notin \mathbb{F}_{i-1}$ ,  $i = 1, \dots, n, \dots$ .

这一系列扩域满足下列包含关系

$$\mathbb{F}_0 \subset \mathbb{F}_1 \subset \dots \subset \mathbb{F}_n \subset \dots.$$

正如定理6.5所指出的那样, 从已知数集  $S$  出发, 用尺规作出数域  $\mathbb{F}_0$ , 再以数域  $\mathbb{F}_0$  为出发点, 用尺规作出它的扩域  $\mathbb{F}_1$ . 这样每一次从新的数域出发, 一步一步地进行域的扩张, 那么扩域  $\mathbb{F}_i$ ,  $i = 1, 2, \dots$  中的数都可以用尺规一步一地作出来.

**例 6.3.1** 设  $\mathbb{F}_0 = \mathbb{Q}$  (有理数域),

(1) 取  $2 \in \mathbb{F}_0$ , 显然  $\sqrt{2} \notin \mathbb{F}_0$ , 则  $\mathbb{F}_0$  的扩域为

$$\mathbb{F}_1 = \{\alpha + \beta\sqrt{2} \mid \alpha, \beta \in \mathbb{F}_0\},$$

再取  $\sqrt{2} \in \mathbb{F}_1$ , 显然  $\sqrt[4]{2} \notin \mathbb{F}_1$ , 若不然, 有

$$\sqrt[4]{2} = \alpha + \beta\sqrt{2} \in \mathbb{F}_1$$

对某个  $\alpha, \beta \in \mathbb{F}_0$  成立, 两边平方得

$$\sqrt{2} = \alpha^2 + 2\beta^2 + 2\alpha\beta\sqrt{2},$$

推出  $\alpha^2 + 2\beta^2 = 0$ ,  $2\alpha\beta = 1$ , 这是不可能的.

这样扩域  $\mathbb{F}_2$  中包含了形如

$$p + q\sqrt[4]{2}$$

的数, 其中

$$p = \alpha + \beta\sqrt{2} \in \mathbb{F}_1, \quad q = \alpha' + \beta'\sqrt{2} \in \mathbb{F}_1, \quad \alpha, \beta, \alpha', \beta' \in \mathbb{F}_0.$$

或写成

$$p + q\sqrt[4]{2} = \alpha + \alpha'\sqrt[4]{2} + \beta\sqrt[4]{2^2} + \beta'\sqrt[4]{2^3},$$

其中  $\alpha, \beta, \alpha', \beta' \in \mathbb{F}_0$  (有理数).

(2) 取  $3 = 3 + 0\sqrt{2} \in \mathbb{F}_1$ , 那么由  $\sqrt{3}$  产生的  $\mathbb{F}_1$  的扩域  $\mathbb{F}_2$  包含了形如

$$p + q\sqrt{3}$$

的数, 其中

$$p = \alpha + \beta\sqrt{2} \in \mathbb{F}_1, \quad q = \alpha' + \beta'\sqrt{2} \in \mathbb{F}_1, \quad \alpha, \beta, \alpha', \beta' \in \mathbb{F}_0,$$

或写成

$$p + q\sqrt{3} = \alpha + \beta\sqrt{2} + \alpha'\sqrt{3} + \beta'\sqrt{6},$$

其中  $\alpha, \beta, \alpha', \beta' \in \mathbb{F}_0$ . 例如, 形如  $\alpha\sqrt{2} + \beta\sqrt{3}$  也是  $\mathbb{F}_2$  中的数, 其中  $\alpha, \beta \in \mathbb{F}_0$ .

(3) 取  $1 + \sqrt{2} \in \mathbb{F}_1$ , 那么由  $\sqrt{1 + \sqrt{2}}$  产生的  $\mathbb{F}_1$  的扩域  $\mathbb{F}_2$  包含了形如

$$p + q\sqrt{1 + \sqrt{2}} = \alpha + \beta\sqrt{2} + \alpha'\sqrt{1 + \sqrt{2}} + \beta'\sqrt{2 + 2\sqrt{2}},$$

的数, 其中  $\alpha, \beta, \alpha', \beta' \in \mathbb{F}_0$ .

## §6.4 几何作图与代数方程的根

用尺规作出的图形仅是直线或圆, 它们的代数方程是以系数属于已知数域  $\mathbb{F}$  中的 1 次或 2 次代数方程, 通过尺规作图得到的交点坐标也不会越过数域  $\mathbb{F}$  和其中数开平方根的范围.

例如用尺规作出已知圆与直线的交点问题, 实际上归结于求一个二次代数方程

$$Ax^2 + Bx + C = 0$$

的实数解问题, 其中系数  $A, B, C$  都是数域已知的数域  $\mathbb{F}$ . 因此交点坐标是方程的系数通过加减乘除和开平方根得到的, 因此是可以用尺规作出这个解的. 例6.2.1也是把用尺规作出单位元内接正十边形的问题, 转化为二次代数方程

$$x^2 + x - 1 = 0$$

的解的问题.

### 定义 6.7 多项式

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

称为数域  $\mathbb{F}$  上多项式, 是指它的系数  $a_n, \dots, a_0 \in \mathbb{F}$ .

若方程  $f(x) = 0$  的一个根 (或解)  $x_0 \in \mathbb{F}_n$ , 其中  $\mathbb{F}_n$  是  $\mathbb{F}$  经过有限次逐步添加平方根得到的扩域, 则称  $x_0$  是可以用尺规作出的.

因此我们转而讨论一个新的问题, 哪些代数方程的实根可以由尺规作出.

下面只讨论一些较特殊的情形. 关于一般的  $n$  次代数方程的求根问题显然超出了本专题讨论的范围, 读者可参考其他书籍.

### 1° 三次方程的求解问题

设

$$f(x) = x^3 + ax^2 + bx + c.$$

是数域  $\mathbb{F}$  上多项式 ( $a, b, c \in \mathbb{F}$ ). 从第 2 讲中的代数基本定理可知, 3 次代数方程  $f(x) = 0$  有三个根  $x_1, x_2, x_3$ , 且根与系数的关系为 (Viète 定理)

$$a = -(x_1 + x_2 + x_3),$$

$$b = x_1x_2 + x_2x_3 + x_1x_3,$$

$$c = -x_1x_2x_3.$$

**引理 6.8** 如果数域  $\mathbb{F}$  上 3 次代数方程  $f(x) = 0$  有形如  $\alpha + \beta\sqrt{d}$  的根, 这里  $\alpha, \beta \in \mathbb{F}$ ,  $\beta \neq 0$ ,  $d \in \mathbb{F}$  但  $\sqrt{d} \notin \mathbb{F}$ , 那么

(1)  $\alpha - \beta\sqrt{d}$  也是方程  $f(x) = 0$  的根.

(2) 方程  $f(x) = 0$  一定有属于  $\mathbb{F}$  的根.

**证明** 设  $x_1 = \alpha + \beta\sqrt{d}$  是满足引理条件的方程  $f(x) = 0$  的根, 令  $x_2 = \alpha - \beta\sqrt{d}$ . 那么 2 次多项式

$$g(x) = (x - x_1)(x - x_2) = x^2 - 2\alpha x + \alpha^2 + d\beta^2,$$

也是数域  $\mathbb{F}$  上的多项式, 用  $g(x)$  除  $f(x)$  得 (参见第 2 讲)

$$f(x) = g(x)q(x) + r(x),$$

这里商式  $q(x)$  和余式  $r(x)$  的系数是  $f(x)$  和  $g(x)$  的系数通过加减乘除等算术运算得到的, 因此  $q(x)$  和  $r(x)$  是  $\mathbb{F}$  上的多项式.

因为除式  $g(x)$  是 2 次多项式, 所以余式  $r(x)$  的次数不超过 1, 不妨设

$$r(x) = \lambda x + \rho \quad (\lambda, \rho \in \mathbb{F}).$$

将  $x = x_1$  代入  $f(x) = g(x)q(x) + r(x)$ , 并注意到  $f(x_1) = g(x_1) = 0$ , 推出  $r(x_1) = 0$ . 如果  $\lambda \neq 0$ , 那么

$$x_1 = -\frac{\rho}{\lambda} \in \mathbb{F},$$

这与  $x_1 = \alpha + \beta\sqrt{d}, \beta \neq 0$  矛盾. 所以  $\lambda = 0$ , 进而  $\rho = 0$ . 这样我们就有

$$f(x) = g(x)q(x),$$

即  $x_2 = \alpha - \beta\sqrt{d}$  也是方程  $f(x) = 0$  的根.

再利用根与系数的关系,  $f(x) = 0$  的第三个根  $x_3$  满足

$$a = -(x_1 + x_2 + x_3) = -(2\alpha + x_3),$$

即  $x_3 = -a + 2\alpha \in \mathbb{F}$ . 这样就完成了引理的证明.  $\square$

**定理 6.9** 设

$$f(x) = x^3 + ax^2 + bx + c$$

是数域  $\mathbb{F}$  上 3 次多项式. 如果方程  $f(x) = 0$  在数域  $\mathbb{F}$  内没有根, 那么方程  $f(x) = 0$  的任何一个根都不可能是用尺规可以作出的数.

**证明** (反证法) 假设方程  $f(x) = 0$  存在一个根, 记为  $x_n$  是能够用尺规作出的数, 也就是数域  $\mathbb{F}$  中的数通过有限次加减乘除以及有限次开平方运算得到的数. 设整个过程用到了  $n$  次开平方根的运算, 并记每次被开平方的数分别为  $d_1, d_2, \dots, d_n$ . 这样我们就得到一串扩域  $\mathbb{F}_0, \mathbb{F}_1, \dots, \mathbb{F}_n$ .

根据假设, 方程的根  $x_n$  在  $\mathbb{F}_n$  内, 因此

$$x_n = \alpha_n + \beta_n \sqrt{d_n} \in \mathbb{F}_n,$$

这里  $\alpha_n, \beta_n \in \mathbb{F}_{n-1}$ ,  $d_n \in \mathbb{F}_{n-1}$  但  $\sqrt{d_n} \notin \mathbb{F}_{n-1}$ . 根据引理 6.8 的第二个结论, 方程  $f(x) = 0$  一定有属于数域  $\mathbb{F}_{n-1}$  的根, 记为

$$x_{n-1} = \alpha_{n-1} + \beta_{n-1} \sqrt{d_{n-1}} \in \mathbb{F}_{n-1},$$

这里  $\alpha_{n-1}, \beta_{n-1} \in \mathbb{F}_{n-2}$ ,  $d_{n-1} \in \mathbb{F}_{n-2}$  但  $\sqrt{d_{n-1}} \notin \mathbb{F}_{n-2}$ . 同理可得方程  $f(x) = 0$  在  $\mathbb{F}_{n-2}$  中有根, 以此类推, 最终得到方程在  $\mathbb{F}_0 = \mathbb{F}$  中有根, 因此与定理的条件相矛盾.  $\square$

## §6.5 几何作图中三个不可解的问题

有了上面各节的准备, 我们可以把几何作图问题转化成一个代数问题, 也就是转化为有理数域  $\mathbb{Q}$  上代数方程是否在  $\mathbb{Q}$  内存在根的问题. 简而言之, 就是在有理数域  $\mathbb{Q}$  上几何作图是否可解, 转化为代数方程在  $\mathbb{Q}$  中是否可解.

现在, 我们可以讨论历史上关于几何作图的三个不可解的著名问题了.

### 1° 倍立方问题

所谓倍立方问题是给定边长为 1 的立方体, 能否用尺规作出一个体积是它二倍的正立方体. 设新的正方体的边长为  $x > 0$ , 问题归结为

已知 1, 用尺规是否能作出数  $x$  使得  $x^3 = 2$ .

从已知数 1 出发, 可以用尺规作出有理数域  $\mathbb{F}_0 = \mathbb{Q}$ . 由于方程

$$x^3 - 2 = 0$$

唯一的实根  $x = \sqrt[3]{2}$  不是有理数, 所以方程在  $\mathbb{Q}$  中没有根, 根据定理6.9, 几何作图的倍立方问题不可解!

为了更好地解释定理6.9, 这里不妨以倍立方问题作为具体例子, 重复定理6.9 的证明过程.

假设这个实根  $x$  可以用尺规作出, 那么它一定需要经过有限次开平方运算才能得到.

设  $x$  可经过  $n$  次开平方运算得到, 并记第  $n$  次开平方的数为  $d$ , 因此

$$x = \alpha + \beta\sqrt{d} \in \mathbb{F}_n, \quad \alpha, \beta, d \in \mathbb{F}_{n-1}, \text{ 但 } \sqrt{d} \notin \mathbb{F}_{n-1}.$$

根据引理6.8中第一个结论,

$$x' = \alpha - \beta\sqrt{d}$$

也是方程  $x^3 - 2 = 0$  的实根.

但是方程只有一个实根  $\sqrt[3]{2}$ , 所以  $x = x' = \sqrt[3]{2}$  推出  $\beta = 0$ , 且

$$x = x' = \sqrt[3]{2} = \alpha \in \mathbb{F}_{n-1}.$$

也就是这个实根  $x$  是  $\mathbb{F}_{n-1}$  中的数, 记为

$$x = \alpha' + \beta'\sqrt{d'}, \quad \alpha', \beta', d' \in \mathbb{F}_{n-2}, \text{ 但 } \sqrt{d'} \notin \mathbb{F}_{n-2}.$$

重复上述推导, 并一步一步继续下去, 知道最后得到一个荒谬的结论:  $x = \sqrt[3]{2} \in \mathbb{F}_0$  是有理数. 因此假设是错误的, 因此就证明了方程  $x^3 - 2 = 0$  的根不可能用尺规作出, 也就是倍立方问题是不可解的!

## 2° 三等分任意角问题

现在证明只用尺规三等分任意角一般来说是不可能的. 当然, 对一些特殊情况, 如像  $90^\circ, 180^\circ$  那样的角是可以用尺规三等分的. 我们要说明用尺规对每一个角的三等分的方法是不存在的.

设给定一个角  $\theta$  的余弦  $\cos \theta$ , 这时, 我们要从已知数集  $S = \{1, \cos \theta\}$  出发, 用尺规作出数  $\cos \left(\frac{\theta}{3}\right)$ . 利用三角函数的有关公式, 我们有

$$\cos \theta = 4 \cos^3 \left(\frac{\theta}{3}\right) - 3 \cos \left(\frac{\theta}{3}\right)$$

因此从已知数  $\cos \theta$  出发, 用尺规作出  $x = \cos\left(\frac{\theta}{3}\right)$  相当于用尺规作出 3 次代数方程

$$4x^3 - 3x - \cos \theta = 0$$

的解. 这里, 我们取一种特殊情况来说明三等分角问题不可解.

设  $\theta = 60^\circ$ , 则  $\cos \theta = \frac{1}{2}$ . 因此从数集  $S = \left\{1, \frac{1}{2}\right\}$  出发, 用尺规可作出有理数域  $\mathbb{Q}$ . 而方程简化为

$$8x^3 - 6x - 1 = 0.$$

或经过简单变换  $v = 2x$ , 方程变换为

$$v^3 - 3v - 1 = 0$$

根据定理6.9, 只需说明上述方程在  $\mathbb{Q}$  中没有根, 因此就推出关于三等分角的几何作图问题不可解.

假设该方程有有理根  $v = \frac{p}{q}$ , 其中  $(p, q) = 1$ . 代入方程有

$$p^3 - 3q^2p - q^3 = 0$$

由此推出

$$q^3 = p(p^2 - 3q^2), \quad \text{或} \quad p^3 = q^2(3p + q)$$

从  $q^3 = p(p^2 - 3q^2)$  看出,  $q^3$  能被  $p$  整除, 因此  $p, q$  有公因子, 除非  $p = \pm 1$ .

从  $p^3 = q^2(3p + q)$  看出  $p^3$  能被  $q$  整除, 因此  $p, q$  有公因子, 除非  $q = \pm 1$ .

所以方程的有理解只可能是  $v = \pm 1$ , 这显然是不成立的. 所以方程没有有理根, 当然也就不可能有用尺规作出的解.

### 3° 正七边形的作图问题

考虑单位圆的内接正七边形问题, 类似正十边形情形, 设每边对应的圆心角为  $\theta = \frac{360^\circ}{7}$ , 那么问题归结为

已知 1, 用尺规是否可以作出  $\cos \theta$ .

借助第 4 讲中的复数, 正七边形的顶点正是方程

$$z^7 - 1 = 0,$$

的单位根. 由于方程有一个显然的根  $z = 1$ , 而其余的根都满足分圆方程

$$\frac{z^7 - 1}{z - 1} = z^6 + z^5 + z^4 + z^3 + z^2 + z + 1 = 0.$$

在上述方程中两边除以  $z^3$  得

$$z^3 + \frac{1}{z^3} + z^2 + \frac{1}{z^2} + z + \frac{1}{z} + 1 = 0,$$

或写成

$$\left(z + \frac{1}{z}\right)^3 + \left(z + \frac{1}{z}\right)^2 - 2\left(z + \frac{1}{z}\right) - 1 = 0$$

设

$$z = \cos \theta + i \sin \theta, \quad \theta = \frac{360^\circ}{7}$$

令

$$x = z + \frac{1}{z} = 2 \cos \theta,$$

则  $x$  满足整系数 3 次代数方程

$$x^3 + x^2 - 2x - 1 = 0.$$

现在从 1 出发, 得到有理数域  $\mathbb{Q}$ . 能否用尺规作出正七边形, 就是能否用尺规作出数  $x$ , 也就是上述方程是否存在有理根.

我们用反证法证明上述方程确实不存在有理根, 也就推出正七边形的作图问题是不可解的.

假设方程有有理根  $x = \frac{p}{q}$ ,  $(p, q) = 1$ , 那么

$$p^3 + p^2q - 2pq^2 - q^3 = 0,$$

上式分别导出

$$\begin{aligned} p^3 &= q(q^2 + 2pq - p^2), \\ q^3 &= p(p^2 + pq - 2q^2), \end{aligned}$$

因此  $p$  和  $q$  有公因子, 这与  $(p, q) = 1$  矛盾.

**注记** 在几何作图不可解的例子中还有一个著名的例子是化圆成方问题, 即能否用尺规作出与一个单位圆面积相等的正方形. 也就是问能否用尺规作出一个满足方程

$$x^2 = \pi$$

的数.

事实上这个问题已经被证明了是不可解的. 但证明的过程需要用到  $\pi$  是超越数这个事实. 所谓超越数是那些不可能成为任何一个有理系数的代数方程的根的数. 这些内容超出了本专题的范围, 在此就不展开讨论了.

## §6.6 等分圆周的几何作图问题\*

除了上节讨论的几何作图中的三个著名问题外, 还有一个饶有兴趣的历史悠久问

题, 就是给定任意正整数  $n$ , 用尺规等分任意圆周问题, 依序连接等分点, 因此上述问题也是作出圆周内接正  $n$  边形问题.

我们已经看到, 内接正十边形问题是可解的, 但是内接正七边形是不可解的. 自然要问, 对于什么样的  $n$ , 上述问题是可解的.

等分圆周, 就是要  $n$  等分圆周角  $360^\circ$ . 与三等分角问题相比较, 很容易看出, 前者只是等分固定的圆周角, 但等分的分数  $n$  可以不同; 而后者虽是三等分, 但是角度却是任意的.

可类比的是,  $n$  等分圆周问题, 就是从已知  $\cos 360^\circ = 1$  的数出发, 用尺规作出  $\cos \frac{360^\circ}{n}$  的数的问题. 或者说, 当  $n$  满足什么条件时, 用尺规可以作出  $\cos \frac{360^\circ}{n}$ .

### 1° 几个例子

**例 6.6.1** 当  $n = 2^k$  时,  $n$  等分圆周的几何作图问题是可解的.

**证明** 显然 2 等分圆周非常简单, 只要过圆心, 用直尺作以直线就可以了, 紧接着按例6.1.1 作该直线的垂直平分线, 则可将圆周得 4 等分. 根据例6.1.2, 用尺规可平分任意角, 因此可 8 等分圆周. 这样一直做下去, 就可作圆周的  $2^k$  等分.

**例 6.6.2** 若  $m$  等分圆周问题可解, 则  $m2^k$  等分圆周问题也可解. 若  $m$  等分圆周问题不可解, 则对  $m$  的任何倍数  $mk$ , 对应的等分问题也不可解.

**证明** 从前一个例子可以看出, 若尺规能够  $m$  等分圆周, 再对等分角不断进行平分, 那么就会得到圆周的  $m2^k$  等分.

设对于  $m$ , 等分问题不可解, 如果存在  $k$ , 使得  $mk$  等分问题可解, 那么将相邻的  $k$  等分合并, 就得到  $m$  等分的圆周, 这与假设是矛盾的, 因此例子中的第二个结论成立.

**例 6.6.3** 对  $m = 3, 5$ , 尺规是可以等分圆周的, 因此也可  $3 \cdot 2^k$  和  $5 \cdot 2^k$  等分圆周, 例如可 6 等分圆周等. 但对 7 和 9, 尺规是无法等分的, 因此对 7 和 9 的倍数  $n = 7m$ , 或  $n = 9m$  的等分问题也是不可解的.

**证明** 三等分圆周对应的圆周角为  $\cos \frac{360^\circ}{3} = \cos 120^\circ = -\frac{1}{2}$ , 因此 3 等分圆周的尺规作图问题是可解的.

注意, 这里三等分的是一个特殊值的圆心角  $\theta = 260^\circ$ , 因此与三等分任意角的结论并不矛盾.

对于 5 等分问题, 只要在例6.2.1 中十个等分点中, 间隔选择五个等分点即可, 因此 5 等分问题用尺规是可以作出的.

7 等分问题如同内接正 7 边形的尺规作图问题, 因此不可解.

对于 9 等分问题, 可仿照 7 等分问题的证明, 从已知数 1 出发, 用尺规作出有理数

域  $\mathbb{Q}$ , 接下来的问题就是对 9 等分角  $\theta = \frac{360^\circ}{9}$ ,  $\cos \theta$  可否能用尺规作出的问题. 设

$$z = \cos \theta + i \sin \theta, \quad \theta = \frac{360^\circ}{9}.$$

那么  $z$  满足

$$z^9 - 1 = (z^3 - 1)(z^6 + z^3 + 1) = 0,$$

因  $z^3 - 1 \neq 0$ , 所以

$$z^6 + z^3 + 1 = 0,$$

同除  $z^3$  得

$$z^3 + 1 + z^{-3} = 0.$$

令  $x = z + z^{-1} = 2 \cos \theta$ , 则  $z^3 + z^{-3} = x^3 - 3x$ , 因此

$$x^3 - 3x + 1 = 0.$$

如果该方程有有理根  $x = \frac{p}{q}$ ,  $(p, q) = 1$ , 则导致

$$p^3 - pq^2 + q^3 = 0,$$

或

$$p^3 = q^2(p - q), \quad q^3 = p(p^2 - q^2)$$

无论哪种情况都与  $(p, q) = 1$  矛盾, 因此没有有理解, 根据定理6.9, 尺规 9 等分圆周问题不可解.

## 2° 17等分圆周(正17边形的作图)问题

下面给出一个著名的等分问题, 它是由Gauss在他19岁那年完成的.

**例 6.6.4** 17 等分圆周的几何作图问题(也是圆内接正17边形的作图问题)是可解的.

**证明** 考虑方程

$$z^{17} - 1 = 0$$

则方程的 17 个根对应单位圆上 17 个等分点. 其中,  $z_0 = 1$  对应的等分点的坐标为  $P_0(1, 0)$ , 记

$$\theta = \frac{360^\circ}{17},$$

则其他 16 个等分点用复数表示分别为

$$z_1 = \cos \theta + i \sin \theta = e^{i\theta},$$

$$z_k = z_1^k = \cos k\theta + i \sin k\theta = e^{ik\theta}, \quad k = 1, \dots, 16.$$

这里用到了复数的 Euler 表示(第 4 讲§4.2). 除  $z_0 = 1$  外, 其它 16 个解共轭成对出现.

$$\bar{z}_k = z_{17-k}, \quad k = 0, 1, \dots, 17.$$

或者说, 对  $\theta = \frac{360^\circ}{17}$ , 有

$$\cos m\theta = \cos k\theta, \quad (m + k = 17)$$

根据根与系数的关系, 有

$$\sum_{k=0}^{17} z_k = 0, \quad \text{或} \quad \sum_{k=1}^8 (z_k + \bar{z}_k) = 2 \sum_{k=1}^8 \cos k\theta = -1.$$

现在从已知的 1 出发, 用尺规作出有理数域  $\mathbb{Q}$ . 如果能证明  $\cos \theta$  是能够通过尺规作出的数, 那么

$$\sin \theta = \sqrt{1 - \cos^2 \theta}$$

是通过加减乘除运算和已知数开平方运算得到的数, 因此也是可以通过尺规作出的. 这样就得到对应  $z$  的等分点的坐标

$$P_1(\cos \theta, \sin \theta),$$

因此线段  $P_0P_1$  的长度也是可以通过尺规作出的数. 从  $P_1$  开始以该长度为半径, 依次作圆交单位圆的交点就是其它的等分点. 这样就完成了用尺规解决 17 等分圆周的问题.

要证明从已知的有理数域  $\mathbb{Q}$  出发, 能够用尺规作出数  $\cos \theta$ . 过程虽然繁琐, 但基本上是初等的. 具体思路如下: 把  $\cos \theta, \cos 2\theta, \dots, \cos 8\theta$  分为如下两组

$$a_1 = 2(\cos \theta + \cos 2\theta + \cos 4\theta + \cos 8\theta),$$

$$a_2 = 2(\cos 3\theta + \cos 5\theta + \cos 6\theta + \cos 7\theta),$$

由根与系数关系直接得到

$$a_1 + a_2 = -1.$$

利用三角函数的和差化积公式, 和

$$\cos m\theta = \cos k\theta, \quad (m + k = 17).$$

可以证明

$$\begin{aligned} a_1 a_2 &= 4(\cos \theta + \cos 2\theta + \cos 4\theta + \cos 8\theta)(\cos 3\theta + \cos 5\theta + \cos 6\theta + \cos 7\theta) \\ &= -4. \end{aligned}$$

综合上述结果, 实数  $a_1, a_2$  满足

$$a_1 + a_2 = -1, \quad a_1 a_2 = -4$$

因此是数域  $\mathbb{Q}$  上二次方程的解

$$x^2 + x - 4 = 0$$

该方程的判别式为  $\Delta = 17 > 0$ , 因此两个实解可通过有理数的加减乘除和开平方根得到, 即  $a_1, a_2$  可以用尺规作出.

再将  $a_1$  中求和项分为以下两组

$$b_1 = 2(\cos \theta + \cos 4\theta),$$

$$b_2 = 2(\cos 2\theta + \cos 8\theta),$$

因此, 首先有

$$b_1 + b_2 = a_1,$$

同时用类似证明  $a_1 a_2 = -4$  的方法得

$$b_1 b_2 = 4(\cos \theta + \cos 4\theta)(\cos 2\theta + \cos 8\theta) = -1,$$

因此  $b_1, b_2$  满足数域  $\mathbb{Q}(\sqrt{17})$  上的二次方程

$$x^2 - a_1 x - 1 = 0$$

其判别式  $\Delta = a_1^2 + 1 > 0$ , 因此两个实根可以通过  $\mathbb{Q}(\sqrt{17})$  中的数经过加减乘除和开平方根得到, 所以是可以用尺规作出的数.

同理把  $a_2$  中的求和项分为以下两组

$$c_1 = 2(\cos 3\theta + \cos 5\theta),$$

$$c_2 = 2(\cos 6\theta + \cos 7\theta),$$

得

$$c_1 + c_2 = a_2, \quad c_1 c_2 = -1.$$

所以  $c_1, c_2$  是  $\mathbb{Q}(\sqrt{17})$  上方程

$$x^2 - a_2 x - 1 = 0, \quad \Delta = a_2^2 + 1 > 0$$

的两个解, 它们也是可以用尺规作出的数.

注意到  $c_1$  的定义, 通过和差化积有

$$c_1 = 2(\cos 3\theta + \cos 5\theta) = 4 \cos \theta \cos 4\theta,$$

将该方程与

$$b_1 = 2(\cos \theta + \cos 4\theta)$$

联立, 不难发现数  $2 \cos \theta$  和  $2 \cos 4\theta$  是下列二次方程

$$x^2 - b_1 x + c_1 = 0$$

的解, 该方程的系数都是可用尺规作出的数, 且判别式  $\Delta = b_1^2 + c_1^2 > 0$ , 因此  $2 \cos \theta$  和  $2 \cos 4\theta$  是在用尺规作出的已知数的基础上, 再次通过加减乘除和开平方根运算得到, 因此也是可以用尺规作出的. 这样, 我们最终证明了用尺规可以作出  $\cos \theta$ ,  $\theta = \frac{260^\circ}{17}$ .  $\square$

**注记** 这里我们只关心用尺规作出  $\cos \theta$  的问题, 如要具体算出每一步得到的一对根, 只需要利用求根公式即可, 但需要比较一对根之间的大小. 详细情况这里就不再讨论了.

### 3° 一般性结论

**引理 6.10** 设  $mn$  是互素的两个正整数  $(m, n) = 1$ , 若  $m$  等分圆周和  $n$  等分圆周问题可解, 则  $mn$  等分圆周问题也可解.

**证明** 因为  $(m, n) = 1$ , 根据 Bezout 定理2.8 (第 2 讲§2.4), 存在两个整数  $l, k$ , 使得

$$lm + kn = 1.$$

因此

$$\frac{1}{mn} = \frac{l}{n} + \frac{k}{m}.$$

记

$$\theta = \frac{360^\circ}{mn}, \quad \theta_1 = \frac{360^\circ}{m}, \quad \theta_2 = \frac{360^\circ}{n}$$

这样用复数的 Euler 表示, 有

$$e^\theta = e^{k\theta_1} e^{l\theta_2}$$

两边取实部得

$$\cos \theta = \operatorname{Re} ((\cos \theta_1 + i \sin \theta_1)^k (\cos \theta_2 + i \sin \theta_2)^l).$$

根据引理条件,  $\cos \theta_1$   $\cos \theta_2$  用尺规是可以作出的, 因此,

$$\sin \theta_1 = \sqrt{1 - \cos^2 \theta_1}, \quad \sin \theta_2 = \sqrt{1 - \cos^2 \theta_2}$$

也是可用尺规作出的. 进而得出  $\cos \theta$  是通过  $\cos \theta_1$   $\cos \theta_2$  的加减乘除和开平方根运算得到的, 因此也是可以用尺规作出的数.  $\square$

例如 15 等分圆周问题是可解的, 因为  $(3, 5) = 1$ , 且  $\frac{1}{15} = \frac{2}{3} + \frac{1}{5}$ .

现在总结如下:

根据引理6.10, 如果两个素数  $p_1, p_2$  对应的等分问题是可解的, 那么  $p_1 p_2$  对应的等分问题也是可解的. 但是一些例子告诉我们即使是素数, 一些素数对应的等分问题是可解的, 如 2, 3, 5, 17 等等, 一些素数对应的等分问题是不可解的, 如 7.

从 17 等分圆周问题的可解性看出, 素数 17 有其特殊性. 也就是

$$17 = 2^2 + 1$$

正是 Fermat 数(第 2 讲§2.4). 虽然 Fermat 数并不都是素数, 如果我们取 Fermat 素数  $2^{2^n} + 1$ , 则对应的等分问题是可解的. 结合引理和例6.6.2, 我们有下列结论.

**定理 6.11**  $n$  等分圆周(即圆内接正  $n$  边形)可用尺规作图的充分必要条件是

$$n = 2^k p_1 p_2 \cdots p_m,$$

这里  $p_1, p_2, \dots, p_m$  是不同的 Fermat 素数.

详细证明就不再讨论了.

## 第 6 讲习题

1. 设  $p = 1 + \sqrt{2}$ ,  $q = 2 - \sqrt{2}$ , 把  $\frac{p}{q}$ ,  $p + p^2$  表示成  $a + b\sqrt{2}$  的形式.

2. 设  $\mathbb{F}_0 = \mathbb{Q}$ , 证明

(1) 对  $d_1 \in \mathbb{F}_0$ , 但  $\sqrt{d_1} \notin \mathbb{F}_0$ , 那么

$$\mathbb{F}_1 = \{\alpha + \beta\sqrt{d_1} \mid \alpha, \beta \in \mathbb{F}_0\},$$

中任何数都是某个有理系数 2 次代数方程的根.

(2) 对  $d_2 \in \mathbb{F}_1$ , 但  $\sqrt{d_2} \notin \mathbb{F}_1$ , 那么

$$\mathbb{F}_2 = \{\alpha + \beta\sqrt{d_2} \mid \alpha, \beta \in \mathbb{F}_1\},$$

中任何数也是某个有理系数的 4 次代数方程的根.

注记: 一个数  $x$ , 如果是某个整系数代数方程

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0 \quad (a_n, \dots, a_0 \in \mathbb{Z})$$

的根, 则称  $x$  是代数数.

显然, 任何有理数  $x = \frac{p}{q}$  是整系数一次代数方程

$$qx - p = 0$$

的根. 因此有理数是代数数. 也可以把代数数看成是有理数的一种推广.

由于任何有理系数的代数方程, 通过通分都等价于整系数的代数方程, 因此有理系数代数方程的根也一定是整系数代数方程的根, 也就是说任何有理系数代数方程的根一定是代数数.

本题目中, 实际上要证明对于有理数域  $\mathbb{F}_0 = \mathbb{Q}$  的扩域  $\mathbb{F}_1 = \mathbb{F}_0(\sqrt{d_1})$  以及  $\mathbb{F}_2 = \mathbb{F}_1(\sqrt{d_2})$ , 其中的数都是代数数.

进一步用数学归纳法可证明: 扩域  $\mathbb{F}_n$  中的数任何数  $x$ , 都是系数属于  $\mathbb{F}_{n-k}$  的一个  $2^k$  次代数方程的根,  $0 < k \leq n$ . 当  $k = n$  时, 就可证明  $\mathbb{F}_n$  中的数任何数  $x$  都是代数数. 用尺规作图的语言说, 就是从有理数域出发, 经过尺规作出的数都是代数数.

不是代数数的数称为超越数. 事实上, 超越数是存在的, 典型的例子是  $\pi$ . 有关讨论已超出本专题的范围, 不再赘述.

## 第 7 讲 群与对称性

一个  $n$  次的一元代数方程(以下简称方程)通常表示为

$$a_0x^n + a_1x^{n-1} + \cdots + a_n = 0, \quad (a_0 \neq 0)$$

其中  $n$  是正整数, 称为方程的次数,  $a_0, a_1, \dots, a_n$  称为方程的系数, 一般取值为复数, 或者是实数. 以下我们均假设  $a_0, a_1, \dots, a_n$  是实数. 方程的一个解有时也称为方程的根. 在第4讲§4.3中, 已经提到著名的代数基本定理: 任何  $n$  次代数方程一定有  $n$  个根. 但是, 代数基本定理只是指出方程根的存在性, 并没有给出根的具体求法.

我们知道, 一个二次代数方程

$$ax^2 + bx + c = 0, \quad (a \neq 0)$$

的两个根可以通过对系数  $a, b, c$  进行有限次加、减、乘、除和开平方根运算给出. 事实上, 古代巴比伦人就已经用上述方法求解二次代数方程了, 但是对于三次、四次乃至更高次数的代数方程, 能否通过系数的加、减、乘、除和开平方根、开立方根或开四次方根、甚至开更高次方的根求出方程的解. 直到16世纪中叶才得到解决. 时间跨度上千年. 继而人们自然希望也能像二次、三次、四次方程那样找到一个公式, 使五次及以上方程的根能用方程系数的代数式表示出来. 但遗憾的是, 人们又花费了两百多年的时间却毫无进展.

正是这种“遗憾”, 使人们开始意识到对五次方程, 这样的求解公式也许根本不存在. 因此不得不思考背后的原因. 十九世纪早期, 两位年轻的数学家 Galois 和 Abel 考虑了代数方程根的置换, 由此产生了“群”的概念, 不但揭示了五次及五次以上方程用代数求解方法不可解的深层次原因, 也使得群的理论在物理、化学、力学等领域发挥了重要作用.

本专题仅限于介绍如何由求解代数方程引进群的概念, 以及关于群的一些基本内容.

### §7.1 从代数方程求根谈起

首先给出一个一般性的定义

**定义 7.1** 对  $n$  次代数方程

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n = 0, \quad (a_0 \neq 0),$$

这里, 方程的系数  $a_0, a_1, \dots, a_n$  为复数或实数.

如果一个数  $a$  是通过方程系数  $a_0, a_1, \dots, a_n$  的加、减、乘、除以及各种次数的开方得到, 那么称这个数由  $a_0, a_1, \dots, a_n$  代数表示 (简称“表示”).

如果通过某种方法使得方程的根由方程的系数  $a_0, a_1, \dots, a_n$  代数表示, 则称为代数求解方法.

### 1° 二次代数方程的求解

不失一般性, 取二次方程的首项系数为 1, 因此有

$$x^2 + px + q = 0.$$

通过配方, 该方程可以化为

$$\left(x + \frac{p}{2}\right)^2 + q - \frac{p^2}{4} = 0,$$

因此方程的两个根可以由系数  $p, q$  代数表示

$$x_1 = \frac{-p + \sqrt{p^2 - 4q}}{2}, \quad x_2 = \frac{-p - \sqrt{p^2 - 4q}}{2}.$$

记

$$\Delta = p^2 - 4q,$$

根据  $\Delta$  的取值, 可以判断方程的两个根是两个实根、重根还是一对互为共轭的复根.

### 2° 三次代数方程的求解 关于三次代数方程

$$x^3 + ax^2 + bx + c = 0,$$

情况虽然较为复杂. 但仍然可以用代数求解方法求解. 经过代换  $y = x + \frac{a}{3}$ , 方程化为

$$y^3 + \frac{3b - a^2}{3}y + \frac{2a^3 - 9ab + 27c}{27} = 0,$$

因此, 一般三次的代数方程求解问题可归结为对下列方程的求解问题

$$x^3 + px + q = 0.$$

这里  $p, q$  由原方程的系数  $a, b, c$  代数表示. 再作如下变换

$$x = z - \frac{p}{3z},$$

就有

$$z^3 - \frac{p^3}{27z^3} + q = 0,$$

或转化成一个关于  $z^3$  的二次方程

$$z^6 + qz^3 - \frac{p^3}{27} = 0$$

根据二次方程的求根公式, 有

$$z^3 = \frac{-q \pm \sqrt{\Delta}}{2}.$$

这里

$$\Delta = q^2 + \frac{4}{27}p^3.$$

如果  $u$  是满足上式中右边取正号的方程

$$u^3 = \frac{-q + \sqrt{\Delta}}{2},$$

那么  $u\omega, u\omega^2$  也满足上述方程, 即  $u, u\omega, u\omega^2$  是  $\frac{-q+\sqrt{\Delta}}{2}$  的三个立方根. 这里

$$\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \omega^2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$$

是三次的单位根

$$\omega^3 = 1.$$

为了利用  $x = u - \frac{p}{3u}$  求  $x$ , 先要计算  $v = -\frac{p}{3u}$ , 不难得到  $v$  满足

$$v^3 = \left(-\frac{p}{3u}\right)^3 = \frac{-q - \sqrt{\Delta}}{2},$$

那么,  $v\omega, v\omega^2$  也满足同样的方程, 即  $v, v\omega, v\omega^2$  是  $\frac{-q-\sqrt{\Delta}}{2}$  的三个立方根. 要使得  $u, u\omega, u\omega^2$  和  $v, v\omega, v\omega^2$  配对组合成原三次方程的根,  $u, v$  必须满足

$$uv = -\frac{p}{3},$$

因此  $(u\omega)(v\omega^2) = -\frac{p}{3}, (u\omega^2)(v\omega) = -\frac{p}{3}$ , 这样原方程的三个根为

$$x_1 = u + v,$$

$$x_2 = u\omega + v\omega^2 = -\frac{1}{2}(u+v) + i\frac{\sqrt{3}}{2}(u-v),$$

$$x_3 = u\omega^2 + v\omega = -\frac{1}{2}(u+v) - i\frac{\sqrt{3}}{2}(u-v),$$

其中

$$u = \sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}}, \quad v = \sqrt[3]{\frac{-q - \sqrt{\Delta}}{2}}$$

是使得  $uv = -\frac{p}{3}$  成立的  $\frac{-q+\sqrt{\Delta}}{2}$  和  $\frac{-q-\sqrt{\Delta}}{2}$  的立方根. 综合上述过程, 不难发现方程的三个根都可以由方程的系数代数表示, 因此三次代数方程可以用代数求解方法求解.

类似二次方程情形, 由  $\Delta = q^2 + \frac{4}{27}p^3$  可对方程的根做如下判断.

当  $\Delta > 0$  时, 取  $u, v$  是实的立方根, 且  $u \neq v$ , 因此  $x_1$  是实根, 而  $x_2, x_3$  是互为共轭的一对复根.

当  $\Delta = 0$  时, 取  $u, v$  是  $-\frac{q}{2}$  实的立方根, 因此  $u = v, x_1 = 2u, x_2 = x_3 = u$  为三个实根, 其中一对实根是重根.

当  $\Delta < 0$  时, 此时  $p < 0$ , 并且

$$u = \sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}} = \sqrt[3]{\frac{-q + i\sqrt{-\Delta}}{2}}$$

不难验证

$$u\bar{u} = |u|^2 = \sqrt[3]{\left(\frac{-q + i\sqrt{-\Delta}}{2}\right)\left(\frac{-q + i\sqrt{-\Delta}}{2}\right)} = -\frac{p}{3} = uv,$$

也就是  $v = \bar{u}$ . 令  $u = \alpha + i\beta$ , 那么  $v = \alpha - i\beta$ , 所以三个根

$$x_1 = 2\alpha, x_2 = -\alpha - \sqrt{3}\beta, x_3 = -\alpha + \sqrt{3}\beta$$

为互不相等的实根.

以上用代数求解方法, 分别给出了二次和三次代数方程详细的求解过程.

下面将精力集中到如何转变思路, 对代数方程的求解问题有一个更加深入的认识.

## §7.2 对称多项式和代数方程的 Viete 公式

设  $x_1, x_2, \dots, x_n$  是  $n$  个不定元, 称

$$F(x_1, \dots, x_n) = \sum_{i_1, \dots, i_n} a_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n}$$

为  $n$  元多项式. 这里求和都是有限求和  $i_k = 1, \dots, d_k$ ,  $d_k$  为正整数,  $k = 1, \dots, n$ .

### 1° 对称多项式

**定义 7.2** 对  $n$  元多项式  $F(x_1, \dots, x_n)$ , 若交换任意两个不定元均不改变  $F(x_1, \dots, x_n)$ , 则称其为**对称多项式**.

例如, 两个不定元  $x_1, x_2$  的多项式

$$F(x_1, x_2) = x_1^2 + x_2^2$$

满足  $F(x_1, x_2) = F(x_2, x_1)$ , 因此是对称多项式.

设  $x_1, \dots, x_n$  是不定元, 引进新的不定元  $x$ , 作  $x$  的多项式

$$\begin{aligned} f(x) &= (x - x_1)(x - x_2) \cdots (x - x_n) \\ &= x^n - \lambda_1 x^{n-1} + \lambda_2 x^{n-2} - \cdots + (-1)^n \lambda_n \end{aligned}$$

则多项式  $f(x)$  的系数

$$\begin{aligned} \lambda_1(x_1, \dots, x_n) &= x_1 + x_2 + \cdots + x_n, \\ \lambda_2(x_1, \dots, x_n) &= x_1 x_2 + x_1 x_3 + \cdots + x_2 x_3 + \cdots + \cdots + x_{n-1} x_n, \\ &\vdots \\ \lambda_n(x_1, \dots, x_n) &= x_1 x_2 \cdots x_n. \end{aligned}$$

显然是 ( $n$  元) 对称多项式, 因为任意交换  $x_i$  和  $x_j$ , 不改变  $f(x)$ , 当然也不会改变  $f(x)$  的系数. 称  $\lambda_1, \lambda_2, \dots, \lambda_n$  为  $x_1, x_2, \dots, x_n$  的初等对称多项式.

**定理 7.3** (对称多项式基本定理) 设  $F(x_1, x_2, \dots, x_n)$  是  $x_1, x_2, \dots, x_n$  的对称多项式, 则  $F(x_1, x_2, \dots, x_n)$  可表示为初等对称多项式的多项式, 即存在一个  $n$  元多项式  $G(y_1, y_2, \dots, y_n)$ , 使得

$$F(x_1, x_2, \dots, x_n) = G(\lambda_1, \lambda_2, \dots, \lambda_n)$$

**证明** 这里只证明  $n = 2$  的情形, 其它情形证明类似.

设  $F(x_1, x_2)$  是  $x_1, x_2$  的一个对称多项式. 对  $F(x_1, x_2)$  中每一项, 我们按  $x_1$  的次数从高到低排列. 设  $F(x_1, x_2)$  中关于  $x_1$  的最高次项为  $a x_1^n x_2^m$ , 其中  $a$  为该项的系数.

首先断定:  $n \geq m$ . 若不然, 根据对称性  $F(x_1, x_2) = F(x_2, x_1)$ ,  $F(x_1, x_2)$  中还包含  $a x_1^m x_2^n$ , 这与  $a x_1^n x_2^m$  是  $x_1$  的最高次矛盾.

另一方面, 注意到对于两个基本对称多项式的幂

$$a \lambda_1^{n-m} \lambda_2^m = (x_1 + x_2)^{n-m} (x_1 x_2)^m$$

也是  $x_1, x_2$  的对称多项式, 且  $x_1$  的最高次幂项为  $a x_1^n x_2^m$ , 所以在

$$F_1(x_1, x_2) = F(x_1, x_2) - a \lambda_1^{n-m} \lambda_2^m$$

中, 关于  $x_1$  的最高次项的次数一定小于  $n$ , 并且  $F_1(x_1, x_2)$  仍然是对称多项式. 重复上述过程就可把  $F(x_1, x_2)$  表示为  $\lambda_1, \lambda_2$  的多项式.  $\square$

**定理 7.4** (Viete 公式) 对实系数  $n$  次多项式

$$f(x) = x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

设  $x_1, x_2, \dots, x_n$  是它的  $n$  个根, 因此有因式分解

$$f(x) = (x - x_1)(x - x_2) \cdots (x - x_n),$$

则有下列 Viete 公式 (也称为根与系数的关系)

$$\lambda_1(x_1, \dots, x_n) = x_1 + x_2 + \cdots + x_n = -a_{n-1},$$

$$\lambda_2(x_1, \dots, x_n) = x_1x_2 + x_1x_3 + \cdots + x_2x_3 + \cdots + \cdots + x_{n-1}x_n = a_{n-2},$$

⋮

$$\lambda_n(x_1, \dots, x_n) = x_1x_2 \cdots x_n = (-1)^n a_0.$$

### §7.3 代数方程根的置换和求解

下面, 我们从一个新的角度重新审视代数方程的求解问题, 主要以二次和三次代数方程为主.

#### 1° 二次代数方程根的置换和求解

对于二次代数方程

$$x^2 + px + q = 0,$$

设  $x_1, x_2$  是它的两个根, 那么

$$x^2 + px + q = (x - x_1)(x - x_2).$$

因此根与系数关系的 Viete 公式为

$$\lambda_1(x_1, x_2) = x_1 + x_2 = -p,$$

$$\lambda_2(x_1, x_2) = x_1x_2 = q,$$

它们是  $x_1$  和  $x_2$  的对称多项式, 也就是在  $x_1$  和  $x_2$  的置换下具有不变性. 将这种置换记为

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

因此在置换作用下有

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} : \lambda_i(x_1, x_2) \Rightarrow \lambda_i(x_2, x_1) = \lambda_i(x_1, x_2), \quad i = 1, 2..$$

由定理7.3 和 Viete 公式可知

**定理 7.5** 设  $x_1, x_2$  是二次代数方程  $f(x) = x^2 + px + q = 0$  的两个根, 则任何关于  $x_1, x_2$  的对称多项式  $P(x_1, x_2)$  都可以表示为  $\lambda_1, \lambda_2$  的多项式. 因此  $P(x_1, x_2)$  可以由方程的系数  $p, q$  通过系数  $p, q$  的加减和乘法运算表示出来.

### 例 7.3.1

$$\begin{aligned}x_1^2 + x_2^2 &= \lambda_1^2 - 2\lambda_2 = p^2 - 2q; \\x_1^3 + x_2^3 &= \lambda_1^3 - 3\lambda_1\lambda_2 = -p^3 + 3pq; \\(x_1 - x_2)^2 &= \lambda_1^2 - 4\lambda_2 = p^2 - 4q.\end{aligned}$$

注意到对称多项式

$$(x_1 - x_2)^2 = (x_1 + \omega x_2)^2 = p^2 - 4q,$$

其中  $\omega = -1$  是二次的单位根, 即满足

$$\omega^2 = 1.$$

开平方根后与  $x_1 + x_2 = -p$  联立

$$x_1 + x_2 = -p, \quad x_1 - x_2 = \pm\sqrt{p^2 - 4q},$$

就可由方程的系数  $p, q$  (通过加减乘除和开平方根运算) 把根表示出来

$$x_1 = \frac{-p + \sqrt{p^2 - 4q}}{2}, \quad x_2 = \frac{-p - \sqrt{p^2 - 4q}}{2}.$$

这里的关键, 是我们通过对称多项式  $(x_1 - x_2)^2$  找到了  $x_1 - x_2 = x_1 + \omega x_2$ .

### 2° 三次代数方程根的置换和求解

如果说对于二次方程, 还难以体会利用根的置换求解方程的作用, 那么我们用同样的方法, 考察三次代数方程

$$x^3 + px^2 + qx + r = 0,$$

和 Viete 公式

$$\begin{aligned}\lambda_1(x_1, x_2, x_3) &= x_1 + x_2 + x_3 = -p, \\ \lambda_1(x_1, x_2, x_3) &= x_1x_2 + x_2x_3 + x_3x_1 = q, \\ \lambda_1(x_1, x_2, x_3) &= x_1x_2x_3 = -r.\end{aligned}$$

关于根的任何一个置换  $(x_1, x_2, x_3) \rightarrow (x_{i_1}, x_{i_2}, x_{i_3})$  就是对根的编号  $(1, 2, 3)$  进行

一次重新排列. 这样的排列共有六种, 记为:

$$\begin{aligned}\sigma_0 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \\ \sigma_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.\end{aligned}$$

与二次方程的情形完全一样, 任何一个三次方程根  $x_1, x_2, x_3$  的对称的多项式  $P(x_1, x_2, x_3)$ , 都可以表示为  $\lambda_1, \lambda_2, \lambda_3$  的多项式, 进而可以由方程的系数  $p, q, r$  通过加减乘法表示出来.

现在我们需要找到一个与二次情形中  $x_1 - x_2 = x_1 + \omega x_2, \omega^2 = 1$  类似的多项式. 因此我们借助三次单位根  $1, \omega, \omega^2$ , 它们满足

$$\omega^3 = 1.$$

设

$$\psi_1 = x_1 + \omega x_2 + \omega^2 x_3,$$

那么在六种置换的作用下,  $\psi_1$  分别变换为  $\psi_2, \psi_3, \psi_4, \psi_5, \psi_6$ , 具体形式如下

$$\begin{aligned}\sigma_0 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} : \psi_1 \longrightarrow x_1 + \omega x_2 + \omega^2 x_3 = \psi_1, \\ \sigma_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} : \psi_1 \longrightarrow x_1 + \omega x_3 + \omega^2 x_2 = \psi_2, \\ \sigma_2 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} : \psi_1 \longrightarrow x_3 + \omega x_1 + \omega^2 x_2 = \psi_3, \\ \sigma_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} : \psi_1 \longrightarrow x_3 + \omega x_2 + \omega^2 x_1 = \psi_4 \\ \sigma_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} : \psi_1 \longrightarrow x_2 + \omega x_1 + \omega^2 x_3 = \psi_5, \\ \sigma_5 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} : \psi_1 \longrightarrow x_2 + \omega x_3 + \omega^2 x_1 = \psi_6.\end{aligned}$$

注意到, 除置换  $\sigma_0$  外,  $\psi_1$  在其他置换下都发生了变化, 因此  $\psi_1$  不是  $x_1, x_2, x_3$  的对称多项式. 但是整体上看, 在所有置换下, 作为集合  $\{\psi_1, \psi_2, \psi_3, \psi_4, \psi_5, \psi_6\}$  是不变的, 只不过秩序重新排列了而已. 因此

#### 性质 7.6 关于 $t$ 的多项式

$$f(t) = (t - \psi_1)(t - \psi_2)(t - \psi_3)(t - \psi_4)(t - \psi_5)(t - \psi_6)$$

在  $x_1, x_2, x_3$  的任何置换下是不变的, 因此  $f(t)$  的系数是  $x_1, x_2, x_3$  的对称多项式.

进一步分析, 我们发现, 6 个置换

$$S_6 = \{\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5\}$$

中有一组, 记为

$$A_3 = \{\sigma_0, \sigma_2, \sigma_5\}$$

起到较为重要的作用, 其中的置换把  $\psi_1$  分别变换到

$$\psi_1 = \psi_1, \quad \psi_3 = \omega^2\psi_1, \quad \psi_6 = \omega\psi_1,$$

而剩下的  $\sigma_1, \sigma_3, \sigma_4$  把  $\psi_1$  分别变换为  $\psi_2, \psi_4, \psi_5$ , 并且满足

$$\psi_4 = \omega\psi_2, \quad \psi_5 = \omega^2\psi_2.$$

也就是  $\psi_1, \psi_2, \psi_3, \psi_4, \psi_5, \psi_6$  中只有  $\psi_1$  和  $\psi_2$  是独立的. 因此性质7.6 中的对称多项式可以表示为

$$\begin{aligned} f(t) &= (t - \psi_1)(t - \omega\psi_1)(t - \omega^2\psi_1)(t - \psi_2)(t - \omega\psi_2)(t - \omega^2\psi_2) \\ &= (t^3 - \psi_1^3)(t^3 - \psi_2^3) \\ &= t^6 - (\psi_1^3 + \psi_2^3)t^3 + \psi_1^3\psi_2^3. \end{aligned}$$

这里我们用到了  $\omega^3 = 1$  (因此也有  $\omega^2 + \omega + 1 = 0$ ).

### 性质 7.7 多项式 $f(t)$ 的系数

$$\psi_1^3 + \psi_2^3, \quad \text{和} \quad \psi_1^3\psi_2^3$$

是  $x_1, x_2, x_3$  的对称多项式, 因此能够用三次多项式的系数  $p, q, r$  表示

$$\begin{aligned} \psi_1^3 + \psi_2^3 &= -2p^3 + 9pq - 27r, \\ \psi_1^3\psi_2^3 &= (p^2 - 3q)^3. \end{aligned}$$

证明上式时, 只要把  $\psi_1 = x_1 + \omega x_2 + \omega^2 x_3$ ,  $\psi_2 = x_1 + \omega x_3 + \omega^2 x_2$  代入并利用三次多项式的 Viete 公式和单位根  $\omega^3 = 0$ ,  $\omega^2 + \omega + 1 = 0$  即可.

考虑方程  $f(t) = 0$ , 它的 6 个根  $\psi_1, \omega\psi_1, \omega^2\psi_1, \psi_2, \omega\psi_2, \omega^2\psi_2$  可以通过

$$\begin{aligned} t^3 &= \frac{(\psi_1^3 + \psi_2^3) \pm \sqrt{(\psi_1^3 + \psi_2^3)^2 - 4(\psi_1^3\psi_2^3)}}{2} \\ &= \frac{(-2p^3 + 9pq - 27r) \pm \sqrt{(-2p^3 + 9pq - 27r)^2 - 4(p^2 - 3q)^3}}{2} \end{aligned}$$

开立方根得到, 因此它们可以由  $p, q, r$  代数表示 (即通过系数  $p, q, r$  的加减乘除以及开平方根、开立方根得到).

另一方面, 由于只有  $\psi_1, \psi_2$  是独立的, 再从 Viete 公式中取  $x_1 + x_2 + x_3 = -p$ , 通过线性方程组

$$\begin{cases} x_1 + x_2 + x_3 = -p \\ x_1 + \omega x_2 + \omega^2 x_3 = \psi_1 \\ x_1 + \omega x_3 + \omega^2 x_2 = \psi_2 \end{cases}$$

解出  $x_1, x_2, x_3$  (由求解线性方程组的理论可知, 上述方程的系数行列式不为零, 所以有唯一一组解):

$$\begin{cases} x_1 = \frac{1}{3}(-p + \psi_1 + \psi_2), \\ x_2 = \frac{1}{3}(-p + \omega^2\psi_1 + \omega\psi_2), \\ x_3 = \frac{1}{3}(-p + \omega\psi_1 + \omega^2\psi_2), \end{cases}$$

也就是三次方程的三个根  $x_1, x_2, x_3$  可用方程系数  $p$  和  $\psi_1, \omega\psi_1, \omega^2\psi_1, \psi_2, \omega\psi_2, \omega^2\psi_2$  代数表示, 最终得到三次方程的根  $x_1, x_2, x_3$  可由其系数  $p, q, r$  代数表示.

至此, 我们从另一个角度解决了三次代数方程用代数求解问题.

在考虑四次代数方程求根问题时, 根的置换一共有 24 种, 置换的集合记为  $S_4$ , 我们发现  $S_4$  与  $S_3$  有类似的性质, 但是对一般的  $n$  次代数方程, 根的置换种类达到  $n!$  种, 这  $n!$  种置换的集合记为  $S_n$ , 远比  $S_3, S_4$  要复杂. 这就需要对置换所构成的集合的代数性质作进一步研究, 因此也就产生了数学中十分重要的一个方向-群的理论.

## §7.4 置换及其表示

我们首先从置换出发, 并通过置换满足的代数结构, 引进抽象的群的定义, 并以置换为模型, 讨论群的基本性质.

抽象地看, 所谓置换实际上是一个有  $n$  个元素的有限集合

$$X = \{x_1, x_2, \dots, x_n\}$$

自身到自身的 1-1 对应 (1-1 映射):

$$\sigma : X \longrightarrow X,$$

这里集合  $X$  中元素不一定是方程的根, 可以是任何元素, 例如  $X = \{1, 2, \dots, n\}$ . 因为是一对一的, 所以可以具体表示为

$$\sigma(x_1) = x_{\alpha_1}, \sigma(x_2) = x_{\alpha_2}, \dots, \sigma(x_n) = x_{\alpha_n},$$

这里  $\alpha_1, \alpha_2, \dots, \alpha_n$  是 1 到  $n$  之间两两不相等的整数, 因此它们是  $1, 2, \dots, n$  的一个排列.

称  $n$  个元素的置换为  $n$  阶置换, 显然  $n$  阶置换共有  $n!$  种.

因为  $\sigma$  是  $n$  个元素之间的 1-1 对应, 有时我们干脆将它记为

$$\sigma(1) = \alpha_1, \sigma(2) = \alpha_2, \dots, \sigma(n) = \alpha_n,$$

并按列表的方法, 将它表示为上下对应

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \end{pmatrix}$$

一般来说, 我们习惯将置换中的第一排按顺序记为  $1, 2, 3, \dots, n$ , 但是, 只要上下对应不变, 列的顺序并没有任何关系. 例如

$$\begin{pmatrix} 4 & 2 & 1 & 3 \\ 3 & 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

正如一个班级的同学与学号的 1-1 对应, 与同学排的座位没有关系.

记  $n$  个元素的置换的全体为

$$S_n = \left\{ \sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \end{pmatrix} \mid (\alpha_1, \dots, \alpha_n) \text{ 是 } (1, \dots, n) \text{ 的一个排列} \right\}$$

### 1° 置换的乘积

设有两个置换

$$\sigma_1 = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix},$$

它们的乘积定义为两个置换的复合, 即先进行  $\sigma_1$  置换, 再进行  $\sigma_2$  置换:

$$X \xrightarrow{\sigma_1} X \xrightarrow{\sigma_2} X$$

或者表示为

$$\sigma_2 \sigma_1 = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \cdots & n \\ \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix} = \sigma_3.$$

也就是, 在  $\sigma_1$  的第一行任何一个  $i$ , 对应的列找到  $\alpha_i$ , 再在  $\sigma_2$  的第一行中  $\alpha_i$  找到同一列对应的  $\beta_i$ , 这样我们就得到两个置换乘积将  $i$  置换到  $\beta_i$ .

注意到置换的乘积未必具有交换性, 例如

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix},$$

但是

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix},$$

两者是不相等的。虽然置换的乘法不满足交换律，但是满足结合律，即对于任意三个置换  $\sigma, \rho, \tau$ ，有

$$\tau(\rho\sigma) = (\tau\rho)\sigma.$$

### 2° 恒等置换

在所有置换中，下列置换称为恒等置换

$$I = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

它的特点是与任何置换  $s$  相乘，结果不变

$$I\sigma = \sigma I = \sigma.$$

### 3° 逆置换

任何一个置换

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix},$$

必存在逆置换

$$\sigma^{-1} = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ 1 & 2 & \cdots & n \end{pmatrix},$$

使得

$$\sigma^{-1}\sigma = \sigma\sigma^{-1} = I.$$

例如

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}.$$

### 4° 轮换与对换

设  $\alpha_1, \alpha_2, \dots, \alpha_d$  是  $\{1, 2, \dots, n\}$  中  $d$  个不相等的整数。若一个置换  $\sigma$  满足

$$\sigma(\alpha_1) = \alpha_2, \sigma(\alpha_2) = \alpha_3, \dots, \sigma(\alpha_{d-1}) = \alpha_d, \sigma(\alpha_d) = \alpha_1,$$

并且  $\sigma(\beta) = \beta$ ,  $\beta \neq \alpha_1, \alpha_2, \dots, \alpha_d$ . 则称置换为一个  $d$ - 轮换， $d$  称为轮换的长度。简记为

$$\sigma = (\alpha_1 \alpha_2 \cdots \alpha_d) = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{d-1} & \alpha_d & \beta_1 & \beta_2 & \cdots & \beta_r \\ \alpha_2 & \alpha_3 & \cdots & \alpha_d & \alpha_1 & \beta_1 & \beta_2 & \cdots & \beta_r \end{pmatrix}.$$

其中  $d + r = n$ . 因为置换表示与列的排序无关, 所以我们把通过置换发生变化的排在前面, 而那些在置换下不变的排在后面. 当然, 这样的记号不唯一, 例如  $(\alpha_2\alpha_3 \cdots \alpha_d\alpha_1), (\alpha_3\alpha_4 \cdots \alpha_d\alpha_1\alpha_2)$  等等都表示同一个轮换.

如果一个轮换  $\sigma_1$  中出现的数字在另一个轮换  $\sigma_2$  中不再出现, 那么就称这两个轮换是不相交的, 否则称为相交的. 例如 (134) 与 (589) 是不相交的, 但 (134) 与 (247) 是相交的.

两个不相交的轮换的乘积是可以交换的, 即

$$(\alpha_1\alpha_2 \cdots \alpha_d)(\beta_1\beta_2 \cdots \beta_r) = (\beta_1\beta_2 \cdots \beta_r)(\alpha_1\alpha_2 \cdots \alpha_d),$$

只要  $\alpha_k \neq \beta_l$ ,  $k = 1, 2, \dots, d$ ,  $l = 1, 2, \dots, r$ .

所谓对换是指最简单的轮换, 即

$$\sigma = (\alpha\beta) = \begin{pmatrix} \cdots & \alpha & \cdots & \beta & \cdots \\ \cdots & \beta & \cdots & \alpha & \cdots \end{pmatrix}.$$

它表示

$$\sigma(\alpha) = \beta, \quad \sigma(\beta) = \alpha, \quad \sigma(j) = j, \quad j \neq \alpha, \beta.$$

对于轮换和对换我们有下列结果

**定理 7.8** 除恒等置换外, 任何一个置换都可以表示为若干个不相交的轮换的乘积. 任何一个置换都可以表示为若干个对换(可以相交)的乘积.

**证明** 取  $\alpha_1 = 1$ , 若  $\sigma(\alpha_1) = \alpha_1$ , 即  $\alpha_1$  在置换下不发生变化, 则考虑  $\alpha_1 = 2$ , 搜索下去直到出现  $\sigma(\alpha_1) \neq \alpha_1$ . 记  $\alpha_2 = \sigma(\alpha_1)$ . 显然  $\sigma(\alpha_2) \neq \alpha_2$ , 否则与一一对应矛盾. 若  $\sigma(\alpha_2) = \alpha_1$ , 则  $(\alpha_1\alpha_2)$  是一个对换(最简单的轮换), 剩余的是不再出现  $\alpha_1$  和  $\alpha_2$  的置换, 继续上述过程. 若  $\sigma(\alpha_2) \neq \alpha_1$ , 则记  $\alpha_3 = \sigma(\alpha_2)$ . 如此下去直到出现  $\sigma(\alpha_{d_1}) = \alpha_1$ , 这样  $(\alpha_1\alpha_2 \cdots \alpha_{d_1})$  就是一个轮换. 剩余部分是不再包含  $\alpha_1, \alpha_2, \dots, \alpha_{d_1}$  的置换, 继续实施上述过程. 由于置换的个数是有限的, 所以经过有限步就可以把  $\sigma$  分解成不相交的轮换的乘积.

关于定理的第二个结论, 只要考慮任何轮换可以分解成对换的乘积即可, 事实上, 对于任何一个轮换, 有

$$(\alpha_1\alpha_2 \cdots \alpha_d) = (\alpha_1\alpha_2)(\alpha_2\alpha_3) \cdots (\alpha_{d-2}\alpha_{d-1})(\alpha_{d-1}\alpha_d),$$

例如

$$\begin{aligned}
 (1234) &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix} \\
 &= (12)(23)(34).
 \end{aligned}$$

这样我们就证明了定理.  $\square$

定理7.8 实际上给出了置换的另一种表示, 这种表示方法在某些场合更加方便.

## §7.5 有限群及其性质

在上节中, 我们定义了置换的乘法、单位置换以及置换的逆, 同时讨论了乘法、单位元以及逆元所满足的算术性质, 把这些性质抽象出来就得到了“群”的定义.

### 1° 有限群

**定义 7.9** 一个集合  $G$ , 如果满足如下条件, 那么称  $G$  是一个群:

(1)  $G$  有一种二元运算, 称为“乘法”, 即对任意的  $a, b \in G$ , 有  $ab \in G$ , 乘法运算满足结合律

$$(ab)c = a(bc), \quad a, b, c \in G.$$

(2)  $G$  中含有称为“单位元”的元素, 通常记为  $I$ , 它满足

$$Ia = aI = a.$$

(3) 对  $G$  中的任何元素  $a$ , 存在一个称为逆元的元素, 记为  $a^{-1}$ , 它使得

$$aa^{-1} = a^{-1}a = I.$$

如果群  $G$  的元素个数是有限的, 那么称为**有限群**. 元素的个数称为群的阶. 如果  $G$  中的乘法是可交换的, 即对任意的  $a, b \in G$ , 有  $ab = ba$ , 那么称  $G$  为**交换群**或**Abel群**.

由定义, 直接得到如下定理.

**定理 7.10 (消去律)** 若群  $G$  中的元素  $a, b, c$  满足  $ab = ac$ , 或  $ba = ca$ , 则  $b = c$ .

**证明** 设  $ab = ac$ , 两边左乘  $a^{-1}$  得  $a^{-1}(ab) = a^{-1}(ac)$ , 利用结合律得  $(a^{-1}a)b = (a^{-1}a)c$ , 也就是  $b = Ib = Ic = c$ . 对于右乘  $ba = ca$ , 证明类似  $\square$

其实, 在前面各专题中, 已经出现了群的例子, 这里连同一些新的例子一并展示.

**例 7.5.1** 模  $m$  的同余类  $\mathbb{Z}_m$  的子集合

$$\mathbb{Z}_m^* = \{[a] \mid [a] \in \mathbb{Z}_m, (a, m) = 1\},$$

是一个可交换的有限群 (第 2 讲§2.5). 特别, 当  $m = 8$  时,

$$\mathbb{Z}_8^* = \{[1], [3], [5], [7]\}$$

是一个群.

**例 7.5.2**  $n$  次的单位根组成的集合在复数乘法运算下是一个群 (第 4 讲§4.4). 特别当  $n = 4$  时, 四个单位根  $\{1, -1, i, -i\}$  在复数乘法下构成一个群.

**例 7.5.3** 考虑等边三角形的对称不变性. 设  $\{I, a, b\}$  分别是三角形以中心点为圆心顺时针旋转  $0^\circ, 120^\circ, 240^\circ$  的旋转,  $\{c, d, e\}$  分别是以三个角的平分线  $K, L, M$  的反射. 则等边三角形在上述旋转和反射下是不变的.

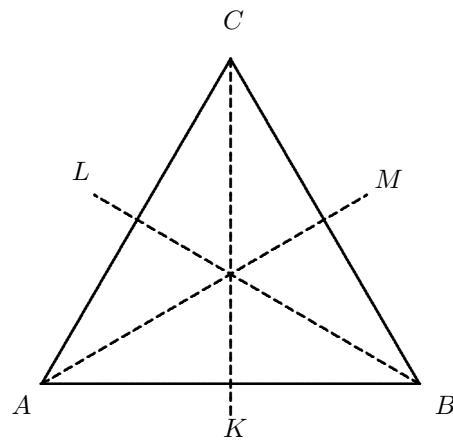


图 7.1

在六个元素的集合中

$$S = \{I, a, b, c, d, e\}$$

将连续作用定义为乘法, 那么  $S$  是一个群, 元素之间的乘积以及单位元和逆元等信息均可从下列乘法表中反映.

乘法	$I$	$a$	$b$	$c$	$d$	$e$
$I$	$I$	$a$	$b$	$c$	$d$	$e$
$a$	$a$	$b$	$I$	$e$	$c$	$d$
$b$	$b$	$I$	$a$	$d$	$e$	$c$
$c$	$c$	$d$	$e$	$I$	$a$	$b$
$d$	$d$	$e$	$c$	$b$	$I$	$a$
$e$	$e$	$c$	$d$	$a$	$b$	$I$

**例 7.5.4** 设下列六个函数

$$\begin{aligned} f_1(x) &= x, & f_2(x) &= \frac{1}{1-x}, & f_3(x) &= \frac{x-1}{x}, \\ f_4(x) &= \frac{1}{x}, & f_5(x) &= 1-x, & f_6(x) &= \frac{x}{x-1} \end{aligned}$$

若定义上述六个函数的乘法如下

$$(f_i \circ f_j)(x) = f_i(f_j(x)),$$

也就是函数的复合, 那么不难验证在该乘法的定义下, 六个函数之间满足群的乘法和性质. 例如,

$$\begin{aligned} (f_6 \circ f_3)(x) &= f_6(f_3(x)) = \frac{(x-1)/x}{(x-1)/x - 1} = 1-x = f_5(x), \\ (f_3 \circ f_6)(x) &= f_3(f_6(x)) = \frac{x/(x-1) - 1}{x/(x-1)} = \frac{1}{x} = f_4(x), \end{aligned}$$

各项之间的乘法由下面的乘法表给出.

乘法	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$
$f_1$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$
$f_2$	$f_2$	$f_3$	$f_1$	$f_6$	$f_4$	$f_5$
$f_3$	$f_3$	$f_1$	$f_2$	$f_5$	$f_6$	$f_4$
$f_4$	$f_4$	$f_5$	$f_6$	$f_1$	$f_2$	$f_3$
$f_5$	$f_5$	$f_6$	$f_4$	$f_3$	$f_1$	$f_2$
$f_6$	$f_6$	$f_4$	$f_5$	$f_2$	$f_3$	$f_1$

其中,  $f_1$  是单位元. 每个元素的逆元也在表中给出, 例如  $f_3$  的逆元为  $f_2$ :

$$f_3 \circ f_2 = f_2 \circ f_3 = f_1.$$

例7.5.2和例7.5.4虽然是完全不同的两个例子, 但是如果做一个对应

$$\{I, a, b, c, d, e\} \longrightarrow \{f_1, f_2, f_3, f_4, f_5, f_6\}$$

不难发现它们的乘法表是完全一样的, 称这种现象为群的同构.

**定义 7.11** 设有两个群  $G_1$  和  $G_2$  分别以  $I_1$  和  $I_2$  为单位元. 如果存在一个 1-1 映射  $\varphi: G_1 \longrightarrow G_2$  使得

$$\varphi(ab) = \varphi(a)\varphi(b) \quad \text{对任意 } a, b \in G_1 \text{ 成立}$$

那么称  $\varphi$  为群  $G_1$  和  $G_2$  的同构. 上式中  $ab$  是  $G_1$  中的乘法,  $\varphi(a)\varphi(b)$  是  $G_2$  中的乘法.

如果两个群  $G_1$  和  $G_2$  之间存在一个同构, 那么称这两个群是同构的, 记为

$$G_1 \cong G_2.$$

根据定义, 若存在  $G_1$  和  $G_2$  的同构  $\varphi$ , 则一定有

$$\varphi(I_1) = I_2, \quad (\varphi(a))^{-1} = \varphi(a^{-1}) \quad a \in G_1.$$

例7.5.2和例7.5.4中给出的两个群是同构的.

顾名思义, 两个同构的群意味着它们有相同的群结构, 与群无关的其它性质可以忽略不计. 例如不管是旋转反射还是函数的复合, 例7.5.2和例7.5.4中群的结构是一样的.

## 2° 子群

**定义 7.12** 设  $G$  是一个群,  $H$  是  $G$  的子集合. 如果  $H$  对  $G$  中的运算也构成一个群, 那么称  $H$  为  $G$  的子群.

显然, 只有单位元的集合(仍然用  $I$  表示)  $I = \{I\}$  以及  $G$  自身一定是  $G$  的子群, 我们称这两个子群为平凡子群, 其它子群称为非平凡子群. 一般情况下, 所讨论的子群都是指非平凡子群.

关于有限群和它的的子群之间有下列定理(也称Lagrange 定理).

**定理 7.13** (Lagrange) 设  $G$  是有限群,  $H$  是  $G$  的非平凡子群(当然也是有限群), 那么  $H$  的阶一定能够整除  $G$  的阶. 若  $G$  的阶为素数, 则  $G$  不存在非平凡子群.

**证明** 设  $G$  的阶为  $n$ ,  $H$  的阶为  $r$ , 因为  $H$  是非平凡子群, 所以  $n > r > 1$ . 记

$$H = \{g_1, g_2, \dots, g_r\} \subset S.$$

取  $G$  中一个元素  $a_1 \in G$  但是  $a_1 \notin H$  (所以  $a_1$  一定不是单位元, 因为单位元一定也在  $H$  中), 则集合

$$a_1H = \{a_1g_1, a_1g_2, \dots, a_1g_r\} \subset S$$

中的元素互不相等, 否则, 必有  $a_1g_i = a_1g_j$ , 根据消去率, 推出  $g_i = g_j$ , 这与  $H$  的元素个数是  $r$  相矛盾.

同时  $a_1H$  中的元素与  $H$  的元素相异, 否则, 必有某个  $a_1g_i$  与  $g_1, g_2, \dots, g_r$  中某一个相等  $a_1g_i = g_j$ , 这样就推出  $a_1 = g_jg_i^{-1} \in H$ , 这与  $a_1$  的选取矛盾.

这样  $G$  包含了  $H$  和  $a_1H$  中共  $2r$  个元素, 即  $n \geq 2r$ .

若  $n = 2r$ , 则  $n$  是  $r$  的倍数, 定理得证.

若  $n > 2r$ , 则一定可以再选取  $G$  中一个元素  $a_2$ , 使得  $a_2 \notin H, a_2 \notin a_1H$ , 令

$$a_2H = \{a_2g_1, a_2g_2, \dots, a_2g_r\} \subset S,$$

那么, 按照上述同样的道理可知  $a_2H$  中的元素互不相等.

如果存在某个  $a_2g_i$  使得  $a_2g_i = g_j$ , 则推出  $a_2 = g_jg_i^{-1} \in H$ .

如果存在某个  $a_2g_i$  使得  $a_2g_i = a_1g_j$ , 则推出  $a_2 = a_1g_jg_i^{-1}$ , 因为  $g_jg_i^{-1} \in H$ , 所以必有  $g_jg_i^{-1} = g_k$ , 推出  $a_2 = a_1g_k \in a_1H$ .

两种情况都与  $a_2$  的选取矛盾, 所以  $a_2H$  中的元素与  $H$  和  $a_1H$  中的元素相异.

因此  $G$  包含了这三组  $H$ ,  $a_1H$  和  $a_2H$  共  $3r$  个元素, 即  $n \geq 3r$ .

若  $n = 3r$ , 则定理得证, 若  $n > 3r$ , 则可以继续上述程序. 因为  $G$  的元素有限, 所以进行到第  $k$  步, 必然有  $n = kr$ , 即  $n$  是  $r$  的倍数.  $\square$

**例 7.5.5** 在例7.5.2中,  $H = \{I, a, b\} \subset S$  的元素满足下列乘法表

乘法	$I$	$a$	$b$
$I$	$I$	$a$	$b$
$a$	$a$	$b$	$I$
$b$	$b$	$I$	$a$

因此是  $S$  的子群.

### 3° 对称群与置换群

根据群的定义, 所有  $n$  阶置换构成的集合  $S_n$ , 满足群的定义, 因此  $S_n$  是一个  $n!$  阶的有限群, 称为**对称群**. 由于置换的乘积是不可交换的, 所以对称群是非 Abel 群.

对称群  $S_n$  中任何子群称为**置换群**.

虽然这里讨论置换群是从代数方程根的对称性开始的, 但是, 置换群有着广泛的应用, 并在不同的应用中以不同的形式出现, 其中几何图形的对称性就是一个重要的例子.

**例 7.5.6** 考虑对称群  $S_3$ , 它由六个置换组成

$$\begin{aligned}\sigma_0 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \\ \sigma_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.\end{aligned}$$

不难看出, 在例7.5.2 中, 若将等边三角形的三个角分别用  $A, B, C$  表示, 那么任何保持等边三角形不变的变换, 都对应于三个角  $(A, B, C)$  的一个置换:

$$I = \sigma_0, \quad a = \sigma_5, \quad b = \sigma_2, \quad c = \sigma_4, \quad d = \sigma_3, \quad e = \sigma_1.$$

因此例7.5.2中的  $S$  就是对称群  $S_3$ .

**例 7.5.7**  $S_3$  的子集合  $\{\sigma_0, \sigma_1\}$ ,  $\{\sigma_0, \sigma_3\}$ ,  $\{\sigma_0, \sigma_4\}$  和  $\{\sigma_0, \sigma_2, \sigma_5\}$  是它的子群, 也就是这四个子集合分别有单位元, 对置换的乘法封闭, 且包含任何元素的逆元. 因此是置换群. 从几何上看, 其中的子群  $\{\sigma_0, \sigma_1\}$ ,  $\{\sigma_0, \sigma_3\}$ ,  $\{\sigma_0, \sigma_4\}$  正是等边三角形沿对称轴的反射, 而子群  $\{\sigma_0, \sigma_2, \sigma_5\}$  是等边三角形的旋转.

继续考虑根的置换, 记

$$A_3 = \{\sigma_0, \sigma_2, \sigma_5\} = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \right\},$$

它的一个典型特征是下列多项式

$$\phi = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$$

仅在  $A_3$  中置换的作用下是不变的:

$$\sigma_i : \phi \longrightarrow \phi, \quad i = 0, 2, 5,$$

但是在  $S_3$  中其它置换的作用是变化的:

$$\sigma_i : \phi \longrightarrow -\phi, \quad i = 1, 3, 4.$$

一般地, 考虑多项式

$$\phi(x_1, x_2, \dots, x_n) = \prod_{i < j}^n (x_i - x_j)$$

即所有  $(x_i - x_j), i < j$  因子的乘积, 不难看出任何一个置换  $\sigma \in S_n$ , 必有

$$\sigma : \phi(x_1, x_2, \dots, x_n) \longrightarrow \pm \phi(x_1, x_2, \dots, x_n).$$

若  $\sigma : \phi \longrightarrow \phi$ , 则称  $\sigma$  为偶置换, 若  $\sigma : \phi \longrightarrow -\phi$ , 则称  $\sigma$  为奇置换.

**定理 7.14** 对于  $S_n$  中的置换, 有

- (1) 任何一个对换一定是奇置换.
- (2) 长度为奇数的轮换是偶置换, 长度为偶数的轮换为奇置换.
- (3) 任何一个偶置换可以分解为偶数个对换的乘积, 任何一个奇置换可以分解为奇数个对换的乘积.

(4)  $S_n$  中偶置换和奇置换的个数各为  $\frac{n!}{2}$ .

**证明** (1) 的证明是显然的. 根据定理 7.8, 直接可以得到 (2) 和 (3). 关于 (4) 的证明, 我们作偶置换和奇置换的如下对应: 设  $\sigma$  是一个偶置换, 那么  $(12)\sigma$  就是一个奇置换, 而且这样的对应是一对一的, 因此偶置换与奇置换的个数相等, 各等于  $S_n$  的一半.  $\square$

类似  $A_3$ , 我们有

**定理 7.15** 记偶置换的全体为  $A_n$ , 则  $A_n$  是  $S_n$  的子群, 称为交错群.

**证明** 设  $\sigma, \tau$  是两个偶置换, 那么

$$\phi(x_1, \dots, x_n) \xrightarrow{\sigma} \phi(x_1, \dots, x_n) \xrightarrow{\tau} \phi(x_1, \dots, x_n),$$

所以  $\tau\sigma$  还是偶置换. 偶置换的逆  $\sigma^{-1} : \phi \Rightarrow \phi$  仍是偶置换, 恒等置换当然是偶置换, 所以  $A_n$  是一个群.  $\square$

为了更好地掌握交错群, 我们利用置换可分解为对换的乘积这样的特点, 进一步分析交错群中元素的特征.

**定理 7.16** 交错群  $A_n$  中任意一个置换  $\sigma$  均可分解为长度为 3 的(可相交)轮换的乘积.

**证明** 根据定理 7.8, 对  $\sigma \in A_n$ , 一定可以分解为偶数个对换的乘积:

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_{2r},$$

若

(1)  $\sigma_1 = \sigma_2$ , 则  $\sigma_1 \sigma_2 = I$ .

(2)  $\sigma_1$  和  $\sigma_2$  有一个字符相同, 不妨设  $\sigma_1 = (\alpha\beta)$ ,  $\sigma_2 = (\beta\gamma)$ , 则

$$\begin{aligned} \sigma_1 \sigma_2 &= (\alpha\beta)(\beta\gamma) \\ &= \begin{pmatrix} \alpha & \beta & \gamma & \cdots \\ \beta & \alpha & \gamma & \cdots \end{pmatrix} \begin{pmatrix} \alpha & \beta & \gamma & \cdots \\ \alpha & \gamma & \beta & \cdots \end{pmatrix} \\ &= \begin{pmatrix} \alpha & \beta & \gamma & \cdots \\ \beta & \gamma & \alpha & \cdots \end{pmatrix} = (\alpha\beta\gamma). \end{aligned}$$

这里省略号表示上下不发生变化的对应.

(3)  $\sigma_1$  和  $\sigma_2$  无相同字符, 不妨设  $\sigma_1 = (\alpha\beta)$ ,  $\sigma_2 = (\gamma\delta)$ , 则

$$\sigma_1 \sigma_2 = (\alpha\beta)(\gamma\delta) = (\alpha\beta)(\beta\gamma)(\beta\gamma)(\gamma\delta) = (\alpha\beta\gamma)(\beta\gamma\delta)$$

这里我们插入的对换满足  $(\beta\gamma)(\beta\gamma) = I$ .

继续对余下的对换的乘积  $\sigma_3 \sigma_4 \cdots \sigma_{2r}$  进行类似讨论, 就得到定理的结果.  $\square$

**例 7.5.8** 根据上述讨论, 我们可以简洁地把交错群  $A_4$  中 12 个元素表示出来.

$$\begin{aligned} A_4 = &\{(1), (12)(34), (13)(24), (14)(23), \\ &(123), (124), (134), (234), (132), (142), (143), (243)\}. \end{aligned}$$

这里我们用 (1) 表示恒等置换.

#### 4° 循环群

一般地, 对阶为  $n$  的有限群  $G$ , 任取  $a \in G$ ,  $a \neq I$ , 那么

$$a, a^2, a^3, \cdots, a^{n+1} \in G,$$

因  $G$  中只有  $n$  个元素, 所以上面  $n+1$  个元素必有两个相等, 不妨设  $a^i = a^j$ ,  $i < j$ , 推得  $a^{j-i} = I$ , 也就是说, 对于有限群  $G$  中任何一个元素  $a \neq I$ , 一定存在正整数  $m$ , 使得  $a^m$  是单位元. 根据最小数原理(第 1 讲§1.1), 一定存在最小的正整数  $r$ , 使得  $a^r = I$ . 称  $r$  为  $a$  的周期. 一般来说, 不同的元素有不同的周期.

**定义 7.17** 设  $G$  是阶为  $n$  的有限群,  $a \in G$ ,  $a \neq I$ ,  $r$  是  $a$  的周期, 那么  $\{I, a, a^2, \dots, a^{r-1}\}$  构成  $G$  的一个  $r$  阶的子群, 并称为由  $a$  生成的循环子群. 如果存在一个元素  $a$ , 使得  $a$  的周期  $r = n$ , 即

$$G = \{I, a, a^2, \dots, a^{n-1}\},$$

那么称  $G$  为循环群.

作为定理7.13 的推论, 我们有

**推论 7.18** 若群  $G$  的阶是素数, 则  $G$  一定是循环群.

**证明** 任取  $a \in G$ ,  $a \neq I$ , 根据定理7.13, 由  $a$  生成的循环子群的阶  $r > 1$  一定整除  $n$ , 因为  $n$  是素数, 所以  $r = n$ , 也就是  $G = \{I, a, a^2, \dots, a^{n-1}\}$  是一个循环群.

**例 7.5.9** 单位元  $I$  的周期为 1. 一个对换  $(ij)$  的周期为了 2:  $(ij)^2 = I$ . 一个轮换  $(i_1 i_2 \cdots i_d)$  的周期为  $d$ :  $(i_1 i_2 \cdots i_d)^d = I$ . 置换

$$a = (12 \cdots n) = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 2 & 3 & 4 & \cdots & n & 1 \end{pmatrix}$$

的周期是  $n$ ,  $\{I, a, a^2, \dots, a^{n-1}\}$  是  $S_n$  的循环子群. 但是对  $n \geq 3$ ,  $S_n$  的阶数  $n!$  不可能是素数, 因此  $S_n$  ( $n \geq 3$ ) 不会是循环群.

#### 4° 正规子群与单群

**定义 7.19** 设  $H$  是  $G$  非平凡子群, 若对于任意的  $g \in G$ ,  $h \in H$ , 有  $g^{-1}hg \in H$ , 则称  $H$  为  $G$  的正规子群.

显然, 如果  $G$  是交换群, 那么  $g \in G$ ,  $h \in H$ , 有  $g^{-1}hg = h \in H$ , 所以交换群的任何子群都是正规子群.

**例 7.5.10** 考虑对称群  $S_3$  的子群  $A_3 = \{\sigma_0, \sigma_2, \sigma_5\}$ , 只要对于  $S_3$  中其它 3 个元素

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

逐一验证  $g^{-1}hg \in A_3$  对  $g = \sigma_1, \sigma_3, \sigma_4$  和  $h = \sigma_0, \sigma_2, \sigma_5$  成立, 那么就推出  $A_3$  是  $S_3$  的正规子群. 经验证有

$$\sigma_1^{-1}\sigma_2\sigma_1 = \sigma_5, \quad \sigma_1^{-1}\sigma_5\sigma_1 = \sigma_2, \quad \sigma_3^{-1}\sigma_2\sigma_3 = \sigma_5,$$

$$\sigma_3^{-1}\sigma_5\sigma_3 = \sigma_2, \quad \sigma_4^{-1}\sigma_2\sigma_4 = \sigma_5, \quad \sigma_4^{-1}\sigma_5\sigma_4 = \sigma_2,$$

例如,

$$\sigma_1^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

$$\begin{aligned} \sigma_1^{-1}\sigma_2\sigma_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \sigma_5, \end{aligned}$$

所以  $A_3$  是  $S_3$  的正规子群.

一般地, 有

**定理 7.20** 交错群  $A_n$  是对称群  $S_n$  的正规子群.

这是因为对任意的  $\tau \in S_n$ ,  $\sigma \in A_n$ ,  $\tau$  和  $\tau^{-1}$  要么同是偶置换, 要么同是奇置换, 所以当  $\sigma$  是偶置换时,  $\tau^{-1}\sigma\tau$  一定是偶置换.

**定义 7.21** 若  $H$  是  $G$  的正规子群, 且  $G$  中不再存在包含  $H$  的正规子群, 则称  $H$  是  $G$  的极大正规子群.

例 7.5.10 中, 因为  $A_3$  的阶是 3,  $S_3$  的阶是 6, 如果存在另一个正规子群  $F$ , 使得  $A_3 \subset F \subset S_3$ , 那么  $F$  的阶  $r$  一定满足  $3 < r < 6$ , 且 3 必须整除  $r$ ,  $r$  必须整除 6, 这是不可能的, 所以  $A_3$  是  $S_3$  的极大正规子群.

一般情况下, 由于  $A_n$  的阶是  $\frac{n!}{2}$ , 所以

$$\frac{S_n \text{ 的阶}}{A_n \text{ 的阶}} = \frac{n!}{n!/2} = 2,$$

如果存在另一个正规子群  $F$ , 使得  $A_n \subset F \subset S_n$ , 那么  $F$  的阶  $r$  一定满足

$$\frac{n!}{2} < r < n!,$$

并且  $r$  能整除  $n!$ , 但

$$\frac{n!}{r} < \frac{n!}{n!/2} = 2$$

推出  $r = n!$ , 或  $F = S_n$ . 因此不存在包含  $A_n$  的  $S_n$  的任何非平凡子群, 所以  $A_n$  是  $S_n$  的极大正规子群.

### 6° 可解群

对于一个有限群  $G$ , 如果  $G_1$  是  $G$  的极大正规子群,  $G_2$  是  $G_1$  的极大正规子群,  $G_3$  是  $G_2$  的极大正规子群,  $\dots$ , 这样得到

$$G \supset G_1 \supset G_2 \supset \dots \supset I,$$

其中最后一个  $I$  必定是一个只包含单位元的群. 这样的一列群, 称为  $G$  的合成群列.

这里需要特别指出, 若  $G_1$  是  $G$  的正规子群,  $G_2$  是  $G_1$  的正规子群, 不能推出  $G_2$  是  $G$  的正规子群!

**定义 7.22** 设  $G$  的合成列为

$$G \supset G_1 \supset G_2 \supset \cdots \supset I,$$

记合成群列中每个群的阶分别为  $n, n_1, n_2, \dots, 1$ , 根据定理 7.13,

$$\frac{n}{n_1}, \frac{n_1}{n_2}, \frac{n_2}{n_3}, \dots$$

是一列整数, 称为  $G$  的组合因数.

如果一个有限群  $G$  的组合因数都是素数, 那么称  $G$  为可解群.

**例 7.5.11**  $S_3$  和  $S_4$  是可解群.

这是因为  $A_3$  是  $S_3$  的极大正规子群,  $A_3$  中不再包含任何非平凡的极大正规子群, 因此  $S_3$  的合成列为

$$S_3 \supset A_3 \supset I,$$

组合因数

$$\frac{S_3 \text{的阶}}{A_3 \text{的阶}} = \frac{6}{3} = 2, \quad \frac{A_3 \text{的阶}}{I \text{的阶}} = 3,$$

都是素数, 所以  $S_3$  是可解的.

在例 7.5.8 中,  $A_4$  是  $S_4$  的极大正规子群, 可以验证,  $A_4$  中的子集合

$$K = \{(1), (12)(34), (13)(24), (14)(23)\}$$

是  $A_4$  的正规子群, 而

$$H = \{(1), (12)(34)\}$$

是  $K$  的正规子群. 注意以下两点, 一是  $H$  是  $K$  的正规子群, 但不是  $A_4$  的正规子群, 二是  $K$  中与  $H$  的同阶的正规子群不唯一. 这样我们有

$$S_4 \supset A_4 \supset K \supset H \supset I$$

它们的组合因数

$$\frac{S_4 \text{的阶}}{A_4 \text{的阶}} = 2, \quad \frac{A_4 \text{的阶}}{K \text{的阶}} = 3, \quad \frac{K \text{的阶}}{H \text{的阶}} = 2, \quad \frac{H \text{的阶}}{I \text{的阶}} = 2,$$

都是素数, 而且在  $S_4, A_4, K, H, I$  不可能再插入任何正规子群, 所以它们是  $S_4$  的合成群列,  $S_4$  是可解群.

**定义 7.23** 若群  $G$  中无非平凡的正规子群, 则称  $G$  为单群.

显然, 若群  $G$  的阶是素数, 那么由定理7.13,  $G$  没有非平凡子群, 当然也不会有非平凡正规子群. 例如  $A_3$  的阶是 3, 所以一定是单群.  $A_4$  不是单群. 但是, 并非有限单群的阶数一定是素数.

**定理 7.24** 当  $n \geq 5$  时,  $A_n$  是单群.

由此很快推出

**定理 7.25**

当  $n \leq 4$  时, 对称群  $S_n$  是可解的.

当  $n \geq 5$  时, 对称群  $S_n$  是不可解的.

定理中关于  $n \leq 4$  的情形已经在例7.5.11中给出. 当  $n \geq 5$  时, 因为  $A_n$  是  $S_n$  的极大正规子群, 由定理7.24 知,  $A_n$  又是单群, 即  $A_n$  中不存在非平凡的正规子群即可. 这样  $S_n$  的合成列只有

$$S_n \supset A_n \supset I,$$

( $I$  是只包含单位元的平凡群), 它们的阶分别为  $n!$ ,  $\frac{n!}{2}$ , 1, 因而组合因数为

$$\frac{S_n \text{的阶}}{A_n \text{的阶}} = 2, \quad \frac{A_n \text{的阶}}{I \text{的阶}} = \frac{n!}{2},$$

显然, 当  $n \geq 5$  时,  $\frac{n!}{2}$  不可能是素数, 所以得到  $S_n$  不可解.

定理7.24 和定理7.25 是最终解决五次及五次以上的代数方程不能够用代数方法求解的关键定理. 虽然在§7.3 中用置换求解三次代数方程时, 已经体会到  $S_3$  的可解性  $S_3 \supset A_3 \supset I$  在求解中扮演的特殊角色. 但是对于彻底解决代数方程求解问题还有很长的路要走. 本专题的目的也就仅限于从求解这样具体问题出发, 如何引进群的概念, 进一步理论可参考有关书籍. 最后附上定理7.24 的证明作为本专题的结束.

**定理7.24 的证明**

设  $n \geq 5$ , 为了证明  $A_n$  没有非平凡的正规子群, 我们将充分利用置换分解为轮换的乘积这个事实.

假设  $A_n$  有  $I$  以外的正规子群  $H$ , 我们将证明  $H$  只能等于  $A_n$  自身.

取  $h \in H$ ,  $h \neq I$ , 并设  $h$  是使得  $\{1, 2, \dots, n\}$  中数字发生变化最少的一个置换. 根据定理7.8,  $h$  能够分解为不相交的轮换的乘积. 我们将分成以下几个步骤进行.

(1) 在  $h$  分解为不相交的轮换乘积中, 各轮换的长度一定相等.

否则, 将  $h$  分解成不相交的轮换的乘积, 其中必有两个轮换的长度不相等:

$$h = (\alpha_1 \alpha_2 \cdots \alpha_k)(\beta_1 \beta_2 \cdots \beta_l)(\cdots) \cdots (\cdots),$$

这里  $k \neq l$ , 不妨设  $k < l$ , 省略的地方的表示不相交的其他轮换. 因为轮换不相交, 所以

在置换的乘法下是可以交换的, 因此有

$$h^k = (\alpha_1 \alpha_2 \cdots \alpha_k)^k h_1 = h_1,$$

这里我们用到了  $(\alpha_1 \alpha_2 \cdots \alpha_k)^k = I$ ,  $h_1$  表示  $h$  中不相交的其他轮换  $k$  次方的乘积. 因为  $H$  是一个子群, 所以  $h_1 = h^k \in H$ , 但是  $h_1$  使得  $\{1, 2, \dots, n\}$  中数字发生变化的个数比  $h$  使得  $\{1, 2, \dots, n\}$  中数字发生变化少了  $k$  个, 这与假设矛盾, 因此在  $h$  分解为不相交的轮换乘积中, 各轮换的长度一定相等.

(2) 在  $h$  分解为不相交的轮换乘积中, 各轮换的长度不会超过 3.

否则在  $h$  分解中, 有一个轮换的长度超过 3, 不妨设

$$h = (1234 \cdots) h_1,$$

这里  $h_1$  是与  $(1234 \cdots)$  不相交的轮换的乘积. 令偶置换  $g = (234) \in A_n$ , 那么

$$\begin{aligned} h_2 &= g^{-1}hg = g^{-1}(1234 \cdots)h_1g = g^{-1}(1234 \cdots)gh_1 \\ &= (243)(1234 \cdots)(234)h_1 = (1423 \cdots)h_1 \end{aligned}$$

这里  $h_2 = g^{-1}hg \in H$ , 所以  $hh_2^{-1} \in H$ , 但是

$$hh_2^{-1} = (1234 \cdots)h_1((1423 \cdots)h_1)^{-1} = (1234 \cdots)(3421) = (143),$$

这与  $h$  的定义矛盾.

(3)  $h$  只能是一个轮换, 而不可能是多个轮换的乘积.

由前面两步, 我们已经知道  $h$  的分解只能由两种可能, 一是分解为长度为 2 的轮换(对换)乘积, 二是分解为长度为 3 的轮换乘积. 我们将逐一排除上述两种可能性.

第一种可能性的排除: 不妨设  $h = (12)(34)$ , 也就是  $h$  使得  $\{1, 2, \dots, n\}$  中 4 个数字发生变化. 因  $n > 4$ , 取偶置换  $g = (125) \in A_n$ , 有

$$hg^{-1}hg = (12)(34)(152)(12)(34)(125) = (152),$$

它使得  $\{1, 2, \dots, n\}$  中 3 个数字发生变化. 但是  $hg^{-1}hg \in H$ , 因此与  $h$  的选取矛盾.

第二种可能性的排除: 不妨设  $h = (123)(456)$ , 也就是  $h$  使得  $\{1, 2, \dots, n\}$  中 6 个数字发生变化. 仍然取偶置换  $g = (125) \in A_n$ , 则

$$hg^{-1}hg = (24356) \in H,$$

它使得  $\{1, 2, \dots, n\}$  中 5 个数字发生变化, 因此与  $h$  的选取矛盾.

总结上面讨论, 我们证明了如果  $h$  是  $H$  中使得  $\{1, 2, \dots, n\}$  改变最少者, 那么  $h$  只能是

$$h = (\alpha_1 \alpha_2), \text{ 或 } h = (\alpha_1 \alpha_2 \alpha_3),$$

但是  $h \in H \subset A_n$  是一个偶置换, 所以  $h$  只能是

$$h = (\alpha_1 \alpha_2 \alpha_3).$$

对于任意一个长度为 3 的轮换  $(\beta_1 \beta_2 \beta_3)$ , 设偶置换

$$g = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{pmatrix}$$

因为  $H$  是  $A_n$  的规范子群, 所以  $g^{-1}hg \in H$ , 这样就推出

$$(\beta_1 \beta_2 \beta_3) = \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_2 & \alpha_3 & \alpha_1 \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{pmatrix} = g^{-1}hg$$

所以, 任何一个长度为 3 的轮换属于  $H$ . 再根据定理7.16,  $A_n$  中的置换可分解为长度为3的轮换的乘积, 因此也属于  $H$ . 最终我们证明了  $H = A_n$ .  $\square$

## 第 7 讲习题

1. 设  $G$  是一个有限群,  $H$  是  $G$  的非空子集. 若对任意的  $a, b \in H$ , 有  $ab \in H$ , 则  $H$  一定是  $G$  的子群.

提示: 要证明  $H$  是  $G$  的子群, 就是要根据已知条件, 证明单位元  $I \in H$ , 以及  $a^{-1} \in H$  对任意的  $a \in H$  成立. 取  $a \in H$ , 若  $a = I$ , 则  $I, a^{-1} \in H$ . 若  $a \neq I$ , 根据条件  $a \cdot a = a^2 \in H$ , 进而推出  $\{a, a^2, \dots, a^n, \dots\} \subset H$ . 再根据  $G$  是有限群的条件, 推出  $I, a^{-1} \in H$ .

2. 证明群  $G$  的两个正规子群的交仍是正规子群.  
 3. 设  $G$  是一个群,  $g \in G$ , 试证下列  $G \rightarrow G$  的映射是同构映射(称为  $G$  的自同构)

$$a \mapsto \varphi(a) = g^{-1}ag, \quad a \in G.$$

4. 类似例7.5.2, 试证保持正四边形不变的变换所构成的群是  $S_4$  的子群, 阶数为  $2 \cdot 4 = 8$ .

提示: 如图不难发现保持正四边形不变的变换就是四个角  $A, B, C, D$  之间的置换, 其中旋转变换可表示为  $A, B, C, D$  的轮换, 而关于对称轴的反射可表示为对换的乘积, 例如关于对称轴  $X$  的反射为  $(1, 4)(2, 3)$ .

注意, 表示正四边形对称性的群是  $S_4$  的子群, 而  $S_4$  表示的是空间正四面体的对称性.

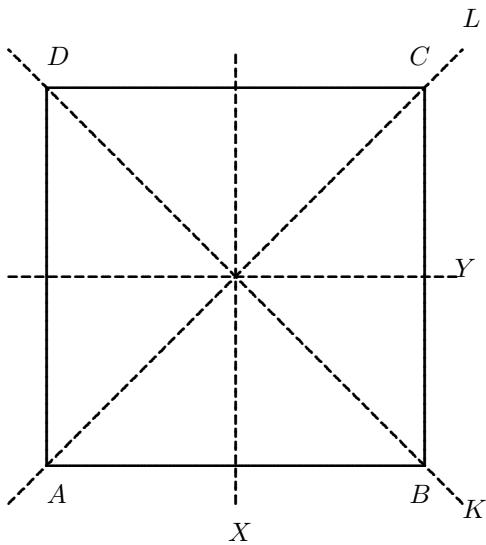


图 7.2

5. 求证不等边长的长方形的对称群有四个元素.