代数学讲义

基础结构篇：集合、群、线性代数、环、模、张量与域

摘要　从集合公理出发，构造映射、商集与自然数，证明归纳和递归原理。继而讨论整数、群、向量空间、环、理想、模和张量积。由主理想整环上的模结构定理推导线性变换的有理标准形与 Jordan 标准形。继而构造域扩张与分裂域，证明有限 Galois 对应，并计算有限域和具体扩张。正文展开对象的构造、运算、结构定理及其证明。

中文修订稿　第九版（审阅修订稿）

目录

记号与约定 **iv**

1 逻辑语言与证明 **1**
1.1 从数学断言到逻辑形式 . . . 1
1.2 变量、量词与对象的依赖 . . . 2
1.3 证明怎样服从断言的结构 . . . 3

2 集合公理与基本构造 **4**
2.1 用元素描述集合 . . . 4
2.2 外延、空集与配对 . . . 5
2.3 分离：在已有集合内部筛选 . . . 6
2.4 集合族的并与交：先分清两层对象 . . . 8
2.5 幂集：把所有子集收成一个集合 . . . 9
2.6 有序对与笛卡尔积 . . . 10
2.7 替换与带指标的集合族 . . . 11
2.8 正则性、无穷与选择 . . . 13
2.9 这些构造怎样用于下一章 . . . 14

3 关系、映射与商集 **15**
3.1 二元关系与函数图像 . . . 15
3.2 映射、复合与可逆性 . . . 15
3.3 像、原像与集合运算 . . . 16
3.4 等价关系：表示不同与对象相同 . . . 17
3.5 划分与商集：把一类作为一个元素 . . . 18
3.6 良定义性：计算能否只依赖对象 . . . 19
3.7 每个映射都带来一种分类 . . . 20
3.8 选择函数与满射的右逆 . . . 21
3.9 幂集比原集合有更多元素 . . . 22
3.10 偏序与幂集中的包含关系 . . . 22

4 自然数、归纳与递归 **24**
4.1 最小归纳集 . . . 24
4.2 归纳法来自什么 . . . 24
4.3 成员关系成为自然数的次序 . . . 25
4.4 递归定义为什么存在 . . . 26
4.5 加法的构造与运算律 . . . 27
4.6 乘法、分配律与幂 . . . 29

5 整数与整除 **31**
5.1 从自然数到整数：让差也成为数 . . . 31
5.2 整除如何转化为等式 . . . 32
5.3 带余除法：存在性与唯一性 . . . 33
5.4 欧几里得算法与 Bézout 等式 . . . 33
5.5 素数与唯一分解 . . . 34
5.6 同余：只保留除法的余数 . . . 35

6 群论 **37**
6.1 二元运算与群公理 . . . 37
6.2 整数次幂与具体群 . . . 38
6.3 子群、交与生成 . . . 40
6.4 同态、核与纤维 . . . 41

6.5 循环群的分类与子群 . . . 42
6.6 陪集、等价关系与 Lagrange 定理 . . . 42
6.7 正规子群与商群的乘法 . . . 44
6.8 商群的因子化与第一同构定理 . . . 45
6.9 子群乘积、第二与第三同构定理 . . . 46
6.10 群作用：用群元素规定置换 . . . 48
6.11 共轭、中心、中心化子与正规化子 . . . 50
6.12 自同构、内自同构与交换化 . . . 50
6.13 类方程与素数幂阶群 . . . 52
6.14 Cauchy 定理：对元素组作循环置换 . . . 53
6.15 Sylow 子群的存在性 . . . 53
6.16 Sylow 子群的包含、共轭与数目 . . . 54
6.17 直积与交换群的直和 . . . 56
6.18 自由群的字与约简 . . . 56
6.19 正规闭包与生成元关系 . . . 58
6.20 自由积与融合自由积的商构造 . . . 59

7 线性代数 61
7.1 从坐标运算到向量空间 . . . 61
7.2 子空间与最小封闭集合 . . . 62
7.3 线性无关来自表示的唯一性 . . . 63
7.4 基、交换引理与维数 . . . 63
7.5 线性映射怎样由一组基确定 . . . 66
7.6 商空间、纤维与线性方程的解 . . . 69
7.7 直和：分解是否唯一 . . . 70
7.8 矩阵：把映射写成坐标运算 . . . 71
7.9 换基：输入与输出坐标的转换 . . . 74
7.10 消元、主元与线性方程的计算 . . . 75
7.11 对偶空间：用标量测量向量 . . . 77
7.12 行列式：交替多线性与可逆性 . . . 79
7.13 特征向量：映射保留哪些方向 . . . 82
7.14 双线性型、内积与正交分解 . . . 84
7.15 四个基本子空间与解方程的整体结构 . . . 88
7.16 仿射子空间与凸组合 . . . 89
7.17 子空间的维数公式 . . . 90
7.18 综合算例 . . . 90

8 环、理想与多项式 93
8.1 为什么需要环 . . . 93
8.2 子环、理想与由生成元产生的结构 . . . 94
8.3 商环与同态 . . . 95
8.4 中国剩余定理 . . . 97
8.5 多项式环与代入 . . . 98
8.6 整环中的不可约分解 . . . 100
8.7 极大理想与素理想 . . . 102

9 交换群与模论 104
9.1 从向量空间到模 . . . 104
9.2 子模、商模与良定义性 . . . 105
9.3 生成、自由模与直和 . . . 107
9.4 挠与无挠 . . . 108
9.5 PID 上的子模 . . . 109
9.6 整数矩阵与 Smith 标准形 . . . 109
9.7 有限生成交换群结构定理 . . . 111
9.8 一般主理想整环上的结构定理 . . . 112
9.9 初等因子与不变因子的唯一性 . . . 116

10 多项式模与线性变换的标准形 **119**
10.1 给向量规定多项式的作用 . . . 119
10.2 最小多项式与循环子空间 . . . 120
10.3 伴随矩阵与有理标准形 . . . 122
10.4 从矩阵实际计算不变因子 . . . 124
10.5 互素因子给出的投影与主分解 . . . 127
10.6 从素幂循环模得到 Jordan 链 . . . 128
10.7 怎样从核的维数读出块大小 . . . 129
10.8 实际换基、迭代与标量域 . . . 131

11 多重线性代数 **133**
11.1 从双线性运算到张量积 . . . 133
11.2 存在性、唯一性与生成元 . . . 134
11.3 映射的张量积与括号的处理 . . . 135
11.4 张量的坐标、矩阵与对偶 . . . 136
11.5 直和、商模与循环模的计算 . . . 137
11.6 扩张标量与平坦性 . . . 138
11.7 张量代数：把任意长度的乘积放在一起 . . . 139
11.8 交替映射与外幂的构造 . . . 140
11.9 外积与外代数 . . . 141
11.10 外幂的基、诱导映射与行列式 . . . 142
11.11 对称代数与多项式 . . . 143
11.12 分次交换与张量乘法的符号 . . . 144

12 域扩张与 Galois 理论 **146**
12.1 域中的标量与特征 . . . 146
12.2 扩张次数与生成子域 . . . 147
12.3 代数元与最小多项式 . . . 149
12.4 嵌入由根的像决定 . . . 150
12.5 分裂域的构造、延拓与唯一性 . . . 151
12.6 形式导数与可分性 . . . 152
12.7 有限自同构群的固定域 . . . 154
12.8 正规性、可分性与 Galois 扩张 . . . 155
12.9 Galois 对应与正规子群 . . . 156
12.10 有限域的构造与全部子域 . . . 158
12.11 双二次扩张的基、固定域与图 . . . 159
12.12 一个非交换的 Galois 群 . . . 161
12.13 本原元：把有限生成改为单生成 . . . 161

记号与约定

$\mathbb{N}=\{0,1,2,\dots\}$，$\mathbb{Z},\mathbb{Q},\mathbb{R},\mathbb{C}$ 分别表示整数、有理数、实数和复数。空集记为 $\varnothing$；$\subseteq$ 允许相等。id_A 表示集合 A 的恒等映射。符号 $\stackrel{\mathrm{def}}{=}$ 表示定义，$\cong$ 表示相应结构的同构。集合论中的无界量词遍历集合；在其余章节中，变量按已声明的对象范围使用。

群一般使用乘法记号，单位元为 e；交换群、向量空间和模使用加法记号，零元为 0。有限和的空和等于 0，有限积的空积等于 1。本讲义的环均结合且有单位元，环同态保持单位元；整环和域额外要求 $1\neq 0$。除专门说明外，模论和多重线性代数中的 R 是交换有单位环，R-模为满足 $1m=m$ 的模。向量空间的标量域记为 $\mathbb{F}$。

集合论采用带等号的经典一阶逻辑及 ZFC 公理体系，公理与基本构造在第 2 章展开。线性代数中的基与维数论证以有限维空间为范围；涉及任意集合族的选择时，在相应位置说明。一个图称为交换图，指具有相同起点和终点的已标明复合映射相等；虚线箭头表示由相应命题构造出的映射。文献说明与正文引用标出概念和定理的标准来源；本文展开的证明采用本书的记号，直接引用而未展开证明的结果另行注明。

1 逻辑语言与证明

1.1 从数学断言到逻辑形式

考察两个断言：“若一个整数是 6 的倍数，它就是偶数”；“若两个集合相等，它们的元素完全相同”。它们谈论的对象不同，却都有“若某个条件成立，则另一个条件成立”的形式。要检验具体的数学结论，必须了解其中的整数或集合；要研究这种推理形式，则可以暂时把两个条件分别记为 P 和 Q。

逻辑研究的一个基本问题是：结论是否由给定的前提保证？将断言的形式单独写出，能够使这个问题接受精确检验。命题逻辑研究整个断言之间的组合；一阶逻辑进一步分析断言中的变量、关系和量词。数学逻辑还研究形式证明、解释及其相互关系。这里先建立证明所需的语言。

定义 1.1 (命题与真值)**.** 在所讨论的解释下，具有确定真假的陈述称为命题。真和假是它的两个真值，分别记为 1 和 0。例如“$2+3=5$”是真命题，“$2+3=6$”是假命题。“求一个偶数”是要求，不是命题。

“x 是偶数”尚未指定 x，不能直接判断真假；指定 $x=4$ 或 $x=5$ 后，得到不同的命题。保留变量的这种陈述将在量词一节讨论。

定义 1.2 (命题公式)**.** 以 $p_0, p_1, \ldots$ 为命题变元。每个变元都是公式；若 P, Q 是公式，则

$$\neg P, \quad (P \wedge Q), \quad (P \vee Q), \quad (P \Rightarrow Q), \quad (P \Leftrightarrow Q)$$

也是公式，分别读作“非 P”“P 且 Q”“P 或 Q”“若 P 则 Q”“P 当且仅当 Q”。所有公式均由有限次应用这些规则得到。括号指定运算的范围。

变元允许我们讨论共同形式。例如将“n 是 6 的倍数”代入 P，将“n 是偶数”代入 Q，就得到前面的整除断言。给各命题变元指定真值称为一个赋值，记为 v。复合公式的真值按下表确定。

$v(P)$	$v(Q)$	$v(\neg P)$	$v(P \wedge Q)$	$v(P \vee Q)$	$v(P \Rightarrow Q)$	$v(P \Leftrightarrow Q)$
1	1	0	1	1	1	1
1	0	0	0	1	0	0
0	1	1	0	1	1	0
0	0	1	0	0	1	1

这里的“或”允许两者同时成立。“若 P 则 Q”只排除一种情形：P 真而 Q 假。以整数断言为例，反驳它需要找出一个是 6 的倍数却不是偶数的整数。一个不满足前提的整数，无论是否为偶数，都不能成为它的反例。这解释了蕴含在前提为假时取真值 1 的约定。

定义 1.3 (逻辑等价与重言式)**.** 若公式 P, Q 在每个赋值下都有相同真值，称它们逻辑等价。若一个公式在每个赋值下均为真，称为重言式。

命题 1.4 (蕴含、逆否与否定)**.** 有如下逻辑等价：

$$\begin{aligned} P \Rightarrow Q \quad &\text{与} \quad \neg P \vee Q, \\ P \Rightarrow Q \quad &\text{与} \quad \neg Q \Rightarrow \neg P, \\ \neg(P \wedge Q) \quad &\text{与} \quad \neg P \vee \neg Q, \\ \neg(P \vee Q) \quad &\text{与} \quad \neg P \wedge \neg Q. \end{aligned}$$

证明. 前两行中的各个公式，都恰在 P 真且 Q 假时为假。第三行两边为真，都表示至少有一个命题为假；第四行两边为真，都表示两个命题均为假。因而每行两边在所有赋值下真值一致。 □

定义 1.5 (充分条件与必要条件)**.** 若 $P \Rightarrow Q$ 成立，则 P 为 Q 的充分条件，Q 为 P 的必要条件。双向蕴含同时成立时，称 P, Q 互为充要条件。

“充分” 表示满足 P 已经足以保证 Q；“必要” 表示要有 P 就必须有 Q。例如，是 6 的倍数足以保证是偶数；成为偶数是成为 6 的倍数必须满足的条件。但偶数 4 不是 6 的倍数，因此逆命题一般不能由原命题得到。逆否命题与原命题的等价，来自上面的真值分析。

1.2 变量、量词与对象的依赖

命题联结词处理已有命题的组合。陈述 “每个整数都满足某个性质” 时，我们还要说明性质对哪些对象成立。整数有无限多个，不能靠列出有限个实例表达这项要求。

定义 1.6 (谓词与量词)**.** 设 X 为给定的集合，称为论域。含变量的陈述 $P(x)$，若每当给定 $x \in X$ 就成为命题，称为 X 上的谓词。记

$$\forall x \in X\ P(x), \qquad \exists x \in X\ P(x),$$

分别表示每个 $x \in X$ 都满足 $P(x)$，以及至少有一个 $x \in X$ 满足 $P(x)$。符号 $\forall, \exists$ 分别为全称量词和存在量词。

例 1.7 (同一公式在不同论域中)**.** “存在 x 使 $x^2 = 2$” 在实数论域中为真，在整数论域中为假。因为任何整数 x 或满足 $|x| \le 1$，从而 $x^2 \le 1$；或满足 $|x| \ge 2$，从而 $x^2 \ge 4$。论域参与了断言的含义。

量词还规定一个选择可以依赖哪些对象。比较

$$\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}\ (y > x), \qquad \exists y \in \mathbb{R}\ \forall x \in \mathbb{R}\ (y > x).$$

第一个断言允许先给 x，再取 $y = x + 1$。第二个要求先固定同一个 y，使它大于所有实数；但此后可以取 $x = y$，要求便不成立。逐个对象都能作选择，与能作一个对所有对象通用的选择，是不同的数学条件。

变量出现的位置也有区别。在 $\forall x \in X\ P(x, y)$ 中，x 在这个量词的作用范围内被约束，y 仍为自由变量；其真值还可能随 y 改变。没有自由变量的公式称为语句。

命题 1.8 (量词的否定)**.** 在经典逻辑中，

$$\begin{aligned} \neg\big(\forall x \in X\ P(x)\big) &\iff \exists x \in X\ \neg P(x), \\ \neg\big(\exists x \in X\ P(x)\big) &\iff \forall x \in X\ \neg P(x). \end{aligned}$$

证明. 否定“每个对象都满足”，意味着至少有一个对象不满足；否定“至少有一个对象满足”，意味着每个对象都不满足。这分别给出两行的等价。 □

因此，否定多重量词断言时，应当逐层处理。例如

$$\neg\big(\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}\ (y > x)\big) \quad\Longleftrightarrow\quad \exists x \in \mathbb{R}\ \forall y \in \mathbb{R}\ (y \le x).$$

右边要求一个实数 x 挡住所有候选 y；只对某一个 x 找到某一个不大于它的 y，不足以完成否定。

若 $X = \varnothing$，则 $\forall x \in X\ P(x)$ 为真，而 $\exists x \in X\ P(x)$ 为假。前者没有可使其失败的对象，后者没有可供选取的对象。这些约定使集合包含、交与原像的公式在空集情形仍保持一致。

1.3 证明怎样服从断言的结构

证明全称命题时，取一个任意对象，只使用已给条件推出结论。任意性保证同一论证适用于论域中的每个对象。证明存在命题时，则必须给出一个满足要求的对象，或用已经证明的存在性结果保证它存在。

例 1.9 (从定义得到直接证明)**.** 整数 n 为偶数，指存在整数 k 使 $n = 2k$。证明“两个偶数之和仍为偶数”时，设 a, b 为任意两个偶数。按定义分别存在整数 r, s 使 $a = 2r, b = 2s$，于是

$$a + b = 2r + 2s = 2(r + s).$$

由于 $r + s$ 为整数，这个等式提供了偶数定义要求的整数，因此 $a + b$ 为偶数。这里两个整数 r, s 分别由 a, b 的偶性得到，没有理由要求它们相同。

对于蕴含 $P \Rightarrow Q$，直接证明在假设 P 下建立 Q。逆否证明在假设 $\neg Q$ 下建立 $\neg P$。反证法则在假设原结论不成立后推出矛盾，从而由经典逻辑得到原结论。

例 1.10 (逆否证明)**.** 证明：整数 n 若满足 n^2 为偶数，则 n 为偶数。其逆否命题是“n 非偶数，则 n^2 非偶数”。整数按奇偶分为两类，故 n 非偶数时可写成 $n = 2k + 1$，其中 $k \in \mathbb{Z}$。于是

$$n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,$$

为奇数。逆否命题成立，原命题随之成立。

注 1.11 (语法与语义)**.** 设 Γ 为一组假设，P 为公式。$\Gamma \vdash P$ 表示在指定规则下存在从这些假设得到 P 的有限形式推导；$\Gamma \models P$ 表示任何使所有假设成立的解释也使 P 成立。前者讨论推导，后者讨论解释。日常数学证明通常用自然语言写出论证；将证明完全形式化，需要进一步指定语言、公理和推导规则。

文献：命题逻辑和一阶语言的系统处理见 [1]；基本证明写法可参照 [3] 第 2、4–7 章。

2 集合公理与基本构造

从一个具体问题开始。设 a, b, c 是三个互不相同的对象，

$$A = \{a, b\}, \qquad C = \{b, c\}.$$

如果只保留 A 中也属于 C 的元素，结果就是 $\{b\}$。这件事通常叫作“取交集”，写成 $A \cap C = \{b\}$。在这个例子中，可以列出全部元素；换成不能逐项列出的集合时，我们需要用条件说明该保留哪些元素。

本章围绕这种集合构造展开。我们要弄清楚：怎样准确描述一个集合的元素，以及哪些公理保证描述出来的集合存在。先处理空集、把两个对象收在一起、从已有集合里筛选元素等构造，再处理有序对和按指标收集集合。下一章的关系与映射需要这些对象；自然数的构造还需要无穷公理。

2.1 用元素描述集合

花括号里列出的是哪些对象？

写 $A = \{a, b\}$，表示 A 的元素恰为 a, b。所以 $a \in A$ 和 $b \in A$ 成立，$c \notin A$ 成立。符号 $\in$ 读作“属于”，$\notin$ 读作“不属于”。它们都在检查一个完整对象是不是右边集合的元素。

集合本身也可以作为元素。比较

$$D = \{a, b\}, \qquad E = \big\{\{a\}, \{b\}\big\}.$$

D 收入的对象是 a, b；E 收入的对象是 $\{a\}, \{b\}$。外层花括号列出的是整个集合的元素，不能把内层花括号直接去掉。例如 $\{a\} \in E$，因为 $\{a\}$ 就是外层列出的一个元素；要判断 a 是否属于 E，则要另行比较 a 是否等于 $\{a\}$ 或 $\{b\}$。

本书的集合基础采用 ZFC。在这个体系中，上面的对象最终都用集合表示，小写字母和大写字母不表示不同种类的对象。开头的例子是在假定 a, b, c 已给定时解释记号；配对公理将保证例子中配出的集合存在。

包含关系检查的是左边的每个元素

$A \subseteq B$ 读作“A 包含于 B”，其含义是：

$$\text{若} x \in A, \text{就有} x \in B \qquad (\text{对每个} x).$$

例如 $\{a\} \subseteq \{a, b\}$，因为左边唯一的元素 a 属于右边。任意集合都满足 $A \subseteq A$；要求两集合不相等时，另外写 $A \subsetneq B$，称为真包含。

因此，$a \in A$ 与 $\{a\} \subseteq A$ 等价：两句话都要求 a 是 A 的元素。但是 $\{a\} \in A$ 要求的是整个集合 $\{a\}$ 被收入 A，这是另一项条件。下节会用空集给出一个能逐项核验的例子。

这些记号的基本读法可对照 Hammack[3, 第 1 章，§1.1、§1.3]。该书用熟悉的数集解释记号；本章进一步说明相应构造所依据的公理。

2.2 外延、空集与配对

外延公理：用元素判别集合相等

外延公理说：两个集合只要元素完全相同，就相等。对任意给定的集合 A, B，其公式为

$$\big[\forall x\,(x \in A \iff x \in B)\big] \quad \Rightarrow \quad A = B.$$

中括号里的条件同时要求 “属于 A 的都属于 B” 和 “属于 B 的都属于 A”。反方向，即 $A = B$ 时成员条件相同，来自等号的代入规则。

命题 2.1 (双向包含判别)**.** *$A = B$ 当且仅当 $A \subseteq B$ 且 $B \subseteq A$。*

证明. 若 $A = B$，任意 $x \in A$ 都属于 B，任意 $x \in B$ 都属于 A，所以两个包含成立。反之，假设两个包含都成立。任取 x，$A \subseteq B$ 保证 $x \in A$ 推出 $x \in B$；$B \subseteq A$ 保证反向蕴含。因此 $x \in A$ 当且仅当 $x \in B$，由外延公理得 $A = B$。 □

空集公理：存在一个没有元素的集合

这里把空集的存在单独列出：存在集合 E，对每个 x 都有 $x \notin E$。如果 E' 也没有元素，那么 E, E' 的成员条件完全相同，外延公理给出 $E = E'$。所以没有元素的集合存在且唯一，记为 $\varnothing$，称为空集。

命题 2.2. *对任意集合 A，有 $\varnothing \subseteq A$；且 $A \subseteq \varnothing$ 当且仅当 $A = \varnothing$。*

证明. 若 $\varnothing \subseteq A$ 不成立，就应当有一个元素属于 $\varnothing$ 却不属于 A。但 $\varnothing$ 没有元素，因此不存在这样的反例，故包含成立。若 $A \subseteq \varnothing$，与刚证明的 $\varnothing \subseteq A$ 合用，双向包含判别给出 $A = \varnothing$。若 $A = \varnothing$，则 A 的每个元素都属于 $\varnothing$，所以 $A \subseteq \varnothing$。 □

配对公理：把给定对象收作元素

对任意给定集合 a, b，存在集合 D，其元素恰为 a 与 b：

$$x \in D \quad \iff \quad x = a \text{ 或} x = b.$$

外延公理使 D 唯一，因此可将它记为 $\{a, b\}$。取 $b = a$，得到 $\{a\}$，称为单元素集。花括号里的顺序和重复书写都不改变成员条件，所以

$$\{a, b\} = \{b, a\}, \qquad \{a, a\} = \{a\}.$$

这里 $\varnothing$ 与 $\{\varnothing\}$ 必须区分：前者没有元素，后者恰有一个元素，而这个元素本身是空集。

例 2.3 (属于与包含要分别检查)**.** 令

$$u = \varnothing, \qquad v = \{u\}, \qquad A = \{v\}.$$

$u \neq v$，因为 u 没有元素，而 v 有元素 u。现在逐项检查：

- $u \notin A$，因为 A 的唯一元素是 v，而 $u \neq v$；但 $u \subseteq A$，因为空集包含于每个集合。
- $v \in A$，这是 $A = \{v\}$ 的定义；但 $v \not\subseteq A$，因为 $u \in v$ 而 $u \notin A$。

因此，同一个左边对象可以“属于右边却不包含于右边”，也可以“包含于右边却不属于右边”。判别的关键是逐项使用两个定义。

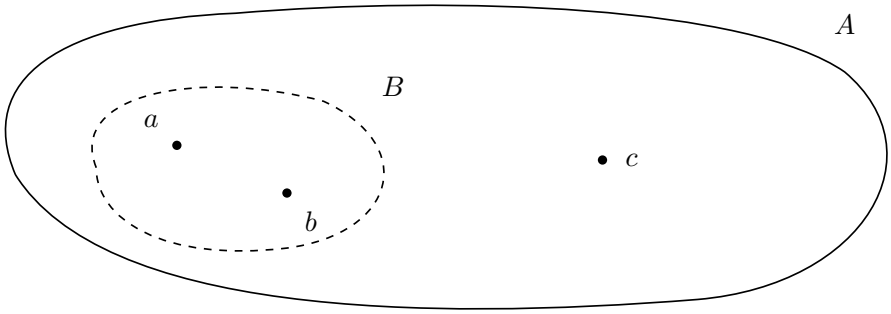


图 2.1: 属于与包含

图 2.1 表示另一种常见情形：$A = \{a, b, c\}$，$B = \{a, b\}$，其中 a, b, c 互不相同。点表示元素，包围曲线表示把这些元素收在一起的集合。于是 $a \in B$、$B \subseteq A$，而 $c \in A$ 且 $c \notin B$。图中没有把整个 B 另画成 A 的一个元素。当一个元素本身是集合时，把它画成点只表示当前观察层次。

2.3 分离：在已有集合内部筛选

回到章首的问题：从 $A = \{a, b\}$ 中，只保留也属于 $C = \{b, c\}$ 的元素。对每个候选元素，我们检查两件事：它是否属于 A，以及它是否属于 C。a 通过第一项检查，未通过第二项；b 两项都通过。所以结果只有 b。

先固定范围，再给出筛选条件

分离公理模式把这个过程推广到任意集合：给定集合 A 和一个关于 x 的合法公式，存在集合 B，其元素恰为 A 中使该公式成立的元素。把这个公式暂记为 $\varphi(x)$，成员条件就是

$$x \in B \quad \Longleftrightarrow \quad x \in A \text{ 且} \varphi(x).$$

φ 是希腊字母 phi，$\varphi(x)$ 在此只是“关于 x 的条件”的记号。例如，取 $\varphi(x)$ 为“$x \in C$”，就回到开头的取交问题。它不是一个待求的集合，也没有在这里定义一个函数。

外延公理保证这样的 B 唯一，于是引入写法

$$B = \{x \in A : \varphi(x)\}.$$

这句话读作“由 A 中所有满足条件 φ 的元素 x 组成的集合”。冒号左边给出候选元素的范围，右边给出保留条件。

什么是参数？ 在条件 “$x \in C$” 中，x 是接受检查的变量，C 是预先给定的集合。这个给定的 C 叫作条件的参数。换一个 C，保留的元素就可能改变。一般公式可以使用有限多个给定参数，教材常写成 $\varphi(x, p_1, \dots, p_r)$。这串写法只是在列出公式允许依赖哪些对象。

要得到 B，条件必须先确定。例如，“保留不属于 B 的元素” 不能用作构造这个尚未确定的 B 的筛选条件。形式上要求结果变量 B 不在该公式中自由出现。这里的 “自由出现” 指没有被公式内部的量词约束、仍需外部指定取值的出现。

为什么叫公理模式？ 用条件 “$x \in C$” 得到一条公理，换成 “$x \notin C$” 又得到一条。每个合法的一阶公式都对应一条这样的公理，“模式” 给出统一的写法。公式必须由成员关系、等式及第一章的逻辑规则构成，允许使用已经定义、能够展开的缩写。这个要求不等于允许任意一句未经说明的自然语言条件。公理模式的标准表述可参见 Weiss[2, 第 3 章，pp. 25–27]。

交集与差集是分离的直接应用

给定 A, B，分别选择条件 “$x \in B$” 和 “$x \notin B$”，得到

$$A \cap B = \{x \in A : x \in B\},$$
$$A \setminus B = \{x \in A : x \notin B\}.$$

$\cap$ 读作 “交”，$\setminus$ 读作 “差”。一个元素属于 $A \cap B$，要求它同时属于两个集合；属于 $A \setminus B$，要求它属于 A 且不属于 B。后者没有使用包含所有集合的 “全集”。

为什么筛选必须有一个已有范围？

如果没有范围限制，尝试把所有满足 “$x \notin x$” 的集合收成 R，就应有

$$x \in R \quad \Longleftrightarrow \quad x \notin x.$$

将 x 换成 R，得到 $R \in R$ 当且仅当 $R \notin R$。若前者成立，便推出它不成立；若前者不成立，又推出它成立，矛盾。这是 Russell 论证，参见 Weiss[2, p. 21]。

限制在已有的 A 内，就能合法构造一个集合；它的结论如下。

定理 2.4 (不存在以所有集合为元素的集合)**.** *对每个集合 A，都存在集合 R_A 不属于 A。*

证明. 固定一个集合 A，在 A 内使用分离，得到

$$R_A = \{x \in A : x \notin x\}.$$

下标 A 只提醒这个集合是从给定的 A 中筛选出来的。它的完整成员条件是

$$x \in R_A \quad \Longleftrightarrow \quad x \in A \text{ 且} x \notin x.$$

假设 $R_A \in A$。在上式中令 $x = R_A$，得到

$$R_A \in R_A \quad \Longleftrightarrow \quad R_A \in A \text{ 且} R_A \notin R_A.$$

在假设 $R_A \in A$ 下，右边等价于 $R_A \notin R_A$。若 $R_A \in R_A$ 成立就推出它不成立；若它不成立就推出它成立，矛盾。所以 $R_A \notin A$。由于 A 任意，没有一个集合能够把所有集合都收作自己的元素。 □

本节确实得到了集合 R_A，结论只是它不属于 A。尤其要同时保留 $R_A \subseteq A$ 与 $R_A \notin A$：前者由筛选范围得到，后者由证明得到，二者并不矛盾。这一证明只用分离及逻辑推理，不需要先用后面的正则公理证明 $x \notin x$。

补充术语：类

教材有时用 $\{x : \varphi(x)\}$ 泛指满足条件的全部集合，称为由公式描述的*类*。这类描述本身不保证存在一个相应的集合。例如 $\{x : x = x\}$ 描述所有集合，刚才的定理说明它不是集合，因此称为*真类*。在 ZFC 中，这里只是用类记号简写公式，没有增加一种新的基本对象。需要把描述的结果当成集合使用时，仍须证明它的存在性。相关记号见 Weiss[2, pp. 20–22]。

2.4 集合族的并与交：先分清两层对象

集合与集合族

当一个集合的元素本身都是我们正在讨论的集合时，常称它为*集合族*。为方便阅读，本章把集合族写成 $\mathcal{F}$，其中的一员写成 T。$\mathcal{F}$ 是花体字母 F，$\mathcal{F}$ 与 T 只是角色不同，仍都属于集合。

例如给定 a, b，令

$$\mathcal{F} = \big\{\{a\}, \{a, b\}\big\}.$$

阅读时分两层：外层花括号列出 $\mathcal{F}$ 的成员，内层花括号分别描述这些成员含有哪些元素。所以 $\{a\} \in \mathcal{F}$，而 $a \in \{a\}$。这两句话检查的是不同层次；成员关系一般不具有传递性，不能从它们直接断言 $a \in \mathcal{F}$。

并集公理：把成员集合中的元素合在一起

对每个集合族 $\mathcal{F}$，存在集合 U，满足

$$x \in U \quad \Longleftrightarrow \quad \text{存在} T \in \mathcal{F}, \text{使} x \in T.$$

外延公理使 U 唯一，记为 $\bigcup \mathcal{F}$，读作“集合族 $\mathcal{F}$ 的并集”。$\bigcup$ 是“大并”符号。上例中

$$\bigcup \big\{\{a\}, \{a, b\}\big\} = \{a, b\}.$$

这里的操作是先取出族中的成员集合，再收集它们的元素。

命题 2.5 (两个集合的并)**.** *对给定 A, B，存在唯一集合 $A \cup B$，满足*

$$x \in A \cup B \quad \Longleftrightarrow \quad x \in A \text{ 或} x \in B.$$

证明. 用配对公理得到集合族 $\{A, B\}$，再用并集公理取出 $\bigcup\{A, B\}$。若 x 属于这个并集，就存在 $T \in \{A, B\}$ 使 $x \in T$。这样的 T 是 A 或 B，所以 $x \in A$ 或 $x \in B$。反之，若 $x \in A$，取 $T = A$；若 $x \in B$，取 $T = B$。两种情形都使 $x \in \bigcup\{A, B\}$。于是它具有所需的成员条件，外延公理保证唯一性。 □

$\{A, B\}$ 是以 A, B 为元素的集合，$A \cup B$ 则收集 A, B 各自的元素，两种写法要区分。如果族 $\mathcal{F}$ 为空，就找不到 $T \in \mathcal{F}$，所以 $\bigcup \varnothing = \varnothing$。

非空集合族的交

“大交” 符号 $\bigcap\mathcal{F}$ 表示同时属于族中每个成员集合的元素所成的集合。其存在性可从分离得到，不需要单列一项交集公理。

命题 2.6 (非空集合族的交)**.** *若 $\mathcal{F}\neq\varnothing$，则存在唯一集合 $\bigcap\mathcal{F}$，满足*

$$x\in\bigcap\mathcal{F}\quad\Longleftrightarrow\quad \text{对每个}T\in\mathcal{F},\ \text{都有}x\in T.$$

证明. 因为 $\mathcal{F}$ 非空，取它的一个成员 T_0。在已经存在的集合 T_0 内用分离，得到

$$C=\{x\in T_0:\text{对每个}T\in\mathcal{F}\ \text{都有}x\in T\}.$$

若 $x\in C$，它按定义属于每个 $T\in\mathcal{F}$。反之，若 x 属于每个 $T\in\mathcal{F}$，它也属于其中的 T_0，因而满足 C 的全部条件，得到 $x\in C$。所以 C 具有所需成员条件，外延公理保证唯一性；选用另一个 T_0 也只能得到同一个集合。 □

例如

$$\bigcap\big\{\{a\},\{a,b\}\big\}=\{a\}.$$

若 $\mathcal{F}=\varnothing$，“属于族中每个成员” 对任何 x 都空真。没有共同范围时，按这个条件收集所有 x 就会要求以所有集合为元素的集合，而它不存在。

在给定集合 X 内讨论子集族时，可以另行规定交集的范围：

$$\{x\in X:\text{对每个}T\in\mathcal{F}\ \text{都有}x\in T\}.$$

当 $\mathcal{F}$ 为空，这个条件不排除任何 $x\in X$，所以结果是 X。因此，“空族的交等于 X” 必须连同预先固定的范围 X 一起理解。

命题 2.7 (分配律)**.** *对任意集合 A,B,C，有*

$$A\cap(B\cup C)=(A\cap B)\cup(A\cap C).$$

证明. 取 $x\in A\cap(B\cup C)$。它属于 A，并且属于 B 或 C。若属于 B，则属于 $A\cap B$；若属于 C，则属于 $A\cap C$。所以它属于右边，得到从左向右的包含。反过来，取属于右边的 x。它属于 $A\cap B$ 或 $A\cap C$。前一种情形给出 $x\in A$ 且 $x\in B$，后一种给出 $x\in A$ 且 $x\in C$。两种情形都得到 $x\in A$ 且 $x\in B\cup C$，所以 x 属于左边。由双向包含判别得到等式。 □

2.5 幂集：把所有子集收成一个集合

分离能从 A 中取出某一个子集。下一章讨论集合上的关系时，还要把符合要求的子集本身作为选择对象。因此需要另一个集合，能够把 A 的所有子集收在一起。

幂集公理说：对每个给定集合 A，存在一个集合，以 A 的全部子集为元素。这个集合称为 A 的幂集，记为 $\mathcal{P}(A)$；$\mathcal{P}$ 是花体字母 P。它的成员条件为

$$S\in\mathcal{P}(A)\quad\Longleftrightarrow\quad S\subseteq A.$$

左边把 S 看成幂集的一个元素，右边检查 S 的元素是否都属于 A。

例 2.8 (二元素集合的幂集)**.** 设 $a \neq b$，$A = \{a, b\}$，则

$$\mathcal{P}(A) = \{\varnothing, \{a\}, \{b\}, \{a, b\}\}.$$

若 $S \subseteq A$，S 中不能有 a, b 以外的元素。按“含不含 a”“含不含 b”分成四种情形：都不含得到 $\varnothing$，只含 a 得到 $\{a\}$，只含 b 得到 $\{b\}$，都含得到 $\{a, b\}$。这四个集合都确实是 A 的子集，因此既没有漏项，也没有额外项。当 $A = \varnothing$ 时，它唯一的子集是 $\varnothing$，所以 $\mathcal{P}(\varnothing) = \{\varnothing\}$。

分离得到的是 A 的某一个子集，幂集则把 A 的所有子集收作元素。尤其是 $A \subseteq A$，所以 $A \in \mathcal{P}(A)$；这句话不等于 $A \in A$。

2.6 有序对与笛卡尔积

下一章要把映射看成一组“输入与对应输出”的记录。一条记录需要保留两个位置：(a, b) 的第一项是输入，第二项是输出。如果只写 $\{a, b\}$，就无法分辨这两个位置。下面先构造有序对，再收集所有允许的有序对。

为什么要再定义一种括号？

$\{a, b\} = \{b, a\}$，所以无序配对无法记录“谁在前、谁在后”。表达坐标、关系和映射时，需要一个能区分两个位置的对象。圆括号 (a, b) 表示有序对；它的定义与花括号集合 $\{a, b\}$ 不同。

定义 2.9 (Kuratowski 有序对)**.** 对给定集合 a, b，定义

$$(a, b) = \big\{\{a\}, \{a, b\}\big\}.$$

先用配对公理得到 $\{a\}$ 与 $\{a, b\}$，再把这两个集合配对，得到右边的集合。内外两层花括号各有作用，不能删去其中一层。

下面的定理是 Weiss[2, p. 24，定理 1(5)] 的有序对相等判别，该书使用尖括号，我们使用圆括号。这一结论说明编码确实保存了两个位置。它是使用有序对的依据；读后面的代数时，可以使用相等判别，不必反复展开编码。

定理 2.10 (有序对的相等判别)**.** $(a, b) = (c, d)$ 当且仅当 $a = c$ 且 $b = d$。

证明. 假设两个有序对相等，即

$$\big\{\{a\}, \{a, b\}\big\} = \big\{\{c\}, \{c, d\}\big\}.$$

先证明 $a = c$。左边的每个成员集合都含元素 a，并且其中有成员 $\{a\}$，所以这些成员集合的公共元素恰为 a。对右边同样检查，公共元素恰为 c。两边是相同的集合族，公共元素集合相同，故 $\{a\} = \{c\}$，从而 $a = c$。这里取的是编码右边各成员集合的交，没有对两个坐标 a, b 本身取交。

再证明 $b = d$。左边各成员集合的并是 $\{a, b\}$，右边的并是 $\{c, d\}$，所以

$$\{a, b\} = \{c, d\} = \{a, d\}.$$

若 $b = a$，则右边的元素 d 必须属于左边的单元素集 $\{a\}$，所以 $d = a = b$。若 $b \neq a$，因为 $b \in \{a, d\}$，只能有 $b = d$。因此两种情形都得到第二坐标相等。

反之，若 $a = c$ 且 $b = d$，将这两个等式代入编码的定义，得到两个有序对相等。 □

当 $a=b$ 时，编码退化为 $(a,a)=\{\{a\}\}$，上面的判别仍成立。这种情况也说明集合不保留重复书写，但有序对的编码仍能恢复两个坐标。

定理 2.11 (笛卡尔积的存在)**.** *对任意集合 A,B，所有第一坐标在 A 中、第二坐标在 B 中的有序对构成一个集合，称为笛卡尔积，记为 $A\times B$。*

证明. 要使用分离，先找一个已经存在的集合，容纳全部候选有序对。令 $X=A\cup B$。幂集公理给出 $\mathcal{P}(X)$，再应用一次给出 $\mathcal{P}(\mathcal{P}(X))$。

若 $a\in A, b\in B$，则 $a,b\in X$，从而 $\{a\}\subseteq X$ 且 $\{a,b\}\subseteq X$。所以 $\{a\}$ 与 $\{a,b\}$ 都是 $\mathcal{P}(X)$ 的元素。由有序对的编码，它们组成的集合 (a,b) 是 $\mathcal{P}(X)$ 的一个子集，因此

$$(a,b)\in\mathcal{P}(\mathcal{P}(X)).$$

现在在这个共同范围内使用分离：取出满足如下条件的 p，“存在 $a\in A$ 与 $b\in B$，使 $p=(a,b)$”。所得集合为

$$A\times B=\{p\in\mathcal{P}(\mathcal{P}(X)) : \exists a\in A\ \exists b\in B\ (p=(a,b))\}.$$

它的成员恰为所需有序对，外延公理保证唯一性。 □

笛卡尔积的存在也见 Weiss[2, p. 26，定理 3]，那里使用替换给出另一种证明。本证明中的双重幂集提供了一个足够大的集合范围；它不表示 $A\times B$ 就等于整个双重幂集。实际使用时可写成更短的集合描述

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

若 A 或 B 为空，就不存在这样的有序对，所以 $A\times B=\varnothing$。

2.7 替换与带指标的集合族

此前的分离是在给定范围中保留元素。现在考虑另一件事：对每个 $x\in A$ 都产生一个由它唯一决定的集合，例如 $\{x\}$，这些结果能否共同组成集合？替换公理模式回答这个问题。它在后文收集等价类、集合族等对象时会用到。

替换：把唯一规定的结果收成集合

设 A 已给定，公式 $\varphi(x,y)$ 规定输入 x 与输出 y 的关系。假设对每个 $x\in A$，都存在唯一的 y 满足它，即

$$\forall x\in A\ \exists! y\ \varphi(x,y).$$

替换公理模式保证这些输出组成一个集合 B，其成员条件为

$$y\in B \quad\Longleftrightarrow\quad \text{存在}x\in A,\ \text{使}\varphi(x,y).$$

公式同样可以含有限多个已固定参数；结果集合 B 不能在该公式中自由出现。这里的“唯一”针对每个固定输入；两个不同输入可以有相同输出。

例如取条件 $y=\{x\}$，每个输入的输出由配对公理存在并由外延公理唯一确定。替换就保证 $\{\{x\} : x\in A\}$ 是集合。若已知全部输出都在某个给定集合 Y 中，也可以直接在 Y 内用分离得到输出集合。替换的陈述不要求事先提供这个共同范围。

例 2.12 (带指标的集合族与成员集合)**.** 设 I 是一个集合，称为指标集。对每个 $i \in I$，某个固定公式唯一规定一个集合 A_i。记 $(A_i)_{i\in I}$，读作“以 $i \in I$ 为指标的集合族”。这里的下标 i 记录来源，不是乘法；整串圆括号表示这一族数据，不表示一个有序对。

替换保证所有输出组成的集合

$$\mathcal{F} = \{A_i : i \in I\}$$

存在。于是并集公理给出

$$\bigcup_{i\in I} A_i = \bigcup \mathcal{F}.$$

左边读作“对所有 $i \in I$ 取 A_i 的并”。若 $I = \varnothing$，它等于空集。

例如 $I = \{\alpha, \beta\}$，$\alpha \neq \beta$，且 $A_\alpha = A_\beta = \{u\}$。指标族仍记录两个不同指标各自对应哪一个集合，但 $\mathcal{F} = \{\{u\}\}$ 只含一个不同的成员集合。因此，指标族与“出现过哪些集合所成的集合”记录的信息不同。若要把指标族本身编码成集合，可用替换收集所有 (i, A_i)，下一章的映射定义将使用这种图像。

命题 2.13 (不交并的存在)**.** 给定指标族 $(A_i)_{i\in I}$，所有带来源标签的元素构成一个集合，记为

$$\coprod_{i\in I} A_i = \{(i, a) : i \in I,\ a \in A_i\}.$$

符号 $\coprod$ 读作“不交并”。

证明. 先取 $U = \bigcup_{i\in I} A_i$，再构造已经证明存在的笛卡尔积 $I \times U$。在其中使用分离，保留那些第一坐标为 i、第二坐标属于 A_i 的有序对。每个保留下来的对象恰为 $i \in I, a \in A_i$ 时的 (i, a)，得到所需集合。当 $I = \varnothing$ 时，候选笛卡尔积为空，结果也为空。 □

在刚才两个指标都对应 $\{u\}$ 的例子中，普通并集和不交并分别为

$$\bigcup_{i\in I} A_i = \{u\}, \qquad \coprod_{i\in I} A_i = \{(\alpha, u), (\beta, u)\}.$$

因为 $\alpha \neq \beta$，有序对相等判别说明右边的两个元素不同。不交并所做的事，是给来自不同成员集合的元素附上标签；原来的 A_i 不需要互不相交。

补充：为什么还同时列出分离与替换？

前面把分离和替换都作为可直接使用的规则。它们各自的陈述清楚，但并非彼此独立。下面引用一个标准结果；它不作为后续构造的先修证明。

命题 2.14 (替换给出分离，引用标准结果)**.** 在外延公理下，配对、并集、空集及替换公理已足以推出分离公理模式。

来源：Weiss[2, pp. 27–28，定理 4 及其证明]。其构造是：满足筛选条件的输入 x 对应 $\{x\}$，不满足的输入对应空集；用替换收集输出，再取并，就得到需要的筛选结果。本章仍把分离单列，后文可直接使用它的成员条件。

2.8 正则性、无穷与选择

正则公理：成员关系不能形成某些循环

正则公理也称基础公理，标准陈述见 Weiss[2, p. 24]：每个非空集合 A 都有一个元素 $a \in A$，满足 $a \cap A = \varnothing$。展开交集的意思，就是 a 自己没有任何元素同时属于 A。这里断言的是 A 中*至少有一个*这样的成员，不是要求每个成员都满足。

命题 2.15 (没有自身成员与二元成员循环)**.** *不存在 x 满足 $x \in x$；也不存在 x, y 满足 $x \in y$ 且 $y \in x$。*

证明. 假设 $x \in x$。考察非空集合 $\{x\}$。正则公理要求它有一个成员与它不交，而它唯一的成员就是 x，所以必须有 $x \cap \{x\} = \varnothing$。可是 $x \in x$ 且 $x \in \{x\}$，使 x 成为这个交集的元素，矛盾。

假设 $x \in y$ 且 $y \in x$，考察非空集合 $A = \{x, y\}$。它的成员只能是 x 或 y。若正则公理给出的成员是 x，则 $y \in x$ 且 $y \in A$，所以 $y \in x \cap A$；若给出的成员是 y，则 $x \in y$ 且 $x \in A$，所以 $x \in y \cap A$。两种情形都使交集非空，与正则公理矛盾。 □

成员关系 $x \in x$ 被排除，与子集关系 $A \subseteq A$ 总成立并不冲突。前者把整个 x 当作一个候选元素，后者逐一检查 A 已有的元素。

无穷公理：一个集合同时容纳不断延续的后继

先对任意集合 x 定义

$$S(x) = x \cup \{x\}.$$

S 是本章的后继记号，$S(x)$ 称为 x 的后继集合。这里先给 x 添入“整个 x 作为一个元素”，再与原来已有的元素合并。例如

$$S(\varnothing) = \{\varnothing\}, \qquad S(\{\varnothing\}) = \{\varnothing, \{\varnothing\}\}.$$

配对与并集保证每个给定 x 的后继存在。

集合 I 称为*归纳集*，若它满足两个条件：

1. $\varnothing \in I$；
2. 对每个 $x \in I$，都有 $S(x) \in I$。

无穷公理断言：存在归纳集。逐次构造一个后继，只说明每次操作都能完成；无穷公理进一步给出一个集合，使这些不断构造出的对象同时属于其中。第四章将从一个归纳集构造最小归纳集，称为自然数集，再证明归纳法与递归原理。无穷公理在此没有直接把归纳法当作未经证明的结论。

选择公理：把各次选择共同记录下来

这里使用与下一章衔接的集合形式。设 $\mathcal{F}$ 是一个集合族，其中每个成员 T 都非空，且不同成员两两不交：

$$T, U \in \mathcal{F}, \quad T \neq U \quad \Rightarrow \quad T \cap U = \varnothing.$$

选择公理断言 [2, p. 28]：存在 $C \subseteq \bigcup \mathcal{F}$，使 C 与每个 $T \in \mathcal{F}$ 的交恰有一个元素。这样的 C 记录了从每个成员集合里选出的一个元素。

对某一个固定的非空 T，“存在 $t \in T$”就是非空性的含义。公理要求的是把针对全族的选择同时组成一个集合 C。对于有限多个非空集合，这一点可以由逐次使用存在性来证明；对于任意集合族，这项统一要求由选择公理保证。若已给出一个明确规则，并证明它对每个成员都唯一指定一个元素，可以用替换收集结果；选择公理的陈述不要求提供这样的规则。

若 $\mathcal{F}$ 为空，取 $C = \varnothing$ 就满足要求。当原有集合族的成员可能重叠时，可先用标签把它们改为不交副本。下一章会用这种做法构造选择函数，并证明选择与满射右逆之间的关系。

2.9 这些构造怎样用于下一章

下一章将用 $A \times B$ 的子集表示关系，用满足唯一输出条件的关系表示映射，并把等价关系中的每一个类构造成一个集合。本章的公理已为这些步骤准备好相应对象：配对与有序对记录两个位置，笛卡尔积提供关系的范围，分离筛出符合条件的元素，幂集和替换用于收集子集或等价类。

无穷公理将在自然数一章中使用。选择公理将在下一章构造选择函数时使用；每个引用它的论证都要说明需要怎样的同时选择。外延与正则则分别控制相等判别与成员关系的某些行为。

这些公理构成通常采用的 Zermelo–Fraenkel 集合论加选择公理，简称 ZFC；去掉选择公理的部分称为 ZF。这里把空集与分离单列，便于引用；列出的条目不要求相互独立。

进一步阅读：集合记号、包含、幂集和集合运算可读 Hammack[3, 第 1 章，§1.1–§1.5]；本章用到的集合论语言、类记号与公理可对照 Weiss[2, pp. 19–29]。本章按代数学所需的构造重新安排顺序，不逐段翻译原书。

3 关系、映射与商集

3.1 二元关系与函数图像

集合 A, B 之间的*二元关系*是 $A \times B$ 的一个子集 R。写 $a\,R\,b$ 表示 $(a,b) \in R$。因此给出关系，就是指定哪些有序对满足它；同一个 a 可以与多个 b 相关，也可以不与任何 b 相关。

若关系 $\Gamma \subseteq A \times B$ 满足

$$\forall a \in A\ \exists! b \in B\ ((a,b) \in \Gamma),$$

则它可以表示把每个输入送到唯一输出的规则。这里存在性要求没有输入被遗漏，唯一性要求一个输入不能得到两个不同输出。

3.2 映射、复合与可逆性

定义 3.1 (映射)**.** 映射 $f : A \to B$ 给每个 $a \in A$ 指定唯一的元素 $f(a) \in B$。A 称为定义域，B 称为陪域，$f(a)$ 为 a 的像。两个映射相等，要求它们定义域相同、陪域相同，并且在每个输入处取值相同。

把映射精确写成集合论对象时，可以取数据 (A, B, Γ_f)，其中 $\Gamma_f = \{(a, f(a)) : a \in A\}$ 为它的图像。三项数据可用嵌套有序对编码。因此，本书在说两个映射相等时，也要求它们的陪域相同。

所有从 A 到 B 的映射组成一个集合，记为 B^A。它的存在性可以这样检查：所有候选图像都属于 $\mathcal{P}(A \times B)$，再用分离取出满足唯一输出条件的图像，便得到图像的集合；附上固定的定义域和陪域，由替换得到相应映射的集合。

若 $A = \varnothing$，恰有一个从 A 到 B 的映射，它的图像为空。若 $A \neq \varnothing$ 且 $B = \varnothing$，则不存在这样的映射，因为任取一个输入，都没有目标元素可以给它作为输出。

陪域是定义中的数据。例如同样的规则 $x \mapsto x^2$，视为 $\mathbb{R} \to \mathbb{R}$ 时不能取到负数；视为 $\mathbb{R} \to [0, \infty)$ 时可以取到陪域内每个元素。区间 $[0, \infty)$ 表示所有非负实数。

定义 3.2 (复合与恒等映射)**.** 若 $f : A \to B, g : B \to C$，定义 $g \circ f : A \to C$ 为 $(g \circ f)(a) = g(f(a))$。恒等映射 $\mathrm{id}_A : A \to A$ 满足 $\mathrm{id}_A(a) = a$。

复合从右向左作用。对第三个映射 $h : C \to D$，两种复合在 a 处都给出 $h(g(f(a)))$，因此 $h \circ (g \circ f) = (h \circ g) \circ f$。又有 $f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f$。

定义 3.3 (单射、满射与双射)**.** 映射 $f : A \to B$ 称为单射，若 $f(a) = f(a')$ 总能推出 $a = a'$；称为满射，若对每个 $b \in B$ 均存在 $a \in A$ 使 $f(a) = b$；同时满足二者时称为双射。

满射提供每个目标元素的原像的存在性；单射保证一个目标元素至多有一个原像。因而双射恰好让我们可以把每个 b 唯一地送回它的原像。

定理 3.4 (逆映射)**.** *$f : A \to B$ 为双射，当且仅当存在 $g : B \to A$ 满足*

$$g \circ f = \mathrm{id}_A, \qquad f \circ g = \mathrm{id}_B.$$

此时 g 唯一，记为 f^{-1}，称为逆映射。

证明. 若 f 双射，对每个 $b \in B$，满射性给出原像，单射性保证其唯一。把此原像定义为 $g(b)$。它的图像可在 $B \times A$ 内由条件 $f(a) = b$ 分离得到，因而确实是一个集合。定义直接给出两个恒等式。

反之，若这两个恒等式成立，则 $f(a) = f(a')$ 时在等式两边作用 g，得到 $a = a'$；故 f 单射。对任意 $b \in B$，$a = g(b)$ 满足 $f(a) = b$，故 f 满射。

若 h 也满足逆映射的两个恒等式，则

$$h = h \circ \mathrm{id}_B = h \circ (f \circ g) = (h \circ f) \circ g = \mathrm{id}_A \circ g = g.$$

这证明唯一性。 □

只有 $g \circ f = \mathrm{id}_A$ 就足以推出 f 单射，只有 $f \circ g = \mathrm{id}_B$ 就足以推出 f 满射；证明中两个方向分别使用了不同的恒等式。

3.3 像、原像与集合运算

定义 3.5 (像、原像与限制)**.** 对映射 $f : A \to B$，子集 $S \subseteq A, T \subseteq B$，定义

$$f(S) = \{f(s) : s \in S\}, \qquad f^{-1}(T) = \{a \in A : f(a) \in T\}.$$

它们分别称为 S 的像和 T 的原像。$\mathrm{im}\, f = f(A)$ 称为映射的像。限制映射 $f|_S : S \to B$ 仍按 $s \mapsto f(s)$ 取值。

原像记号 $f^{-1}(T)$ 对每个映射都有意义。对 $f(x) = x^2$，集合 $\{1\}$ 的原像为 $\{-1, 1\}$；这没有给 f 构造逆映射。两者的输入分别是一个子集和一个元素。

命题 3.6 (像与原像的基本关系)**.** 对任意 $S \subseteq A, T \subseteq B$，有

$$S \subseteq f^{-1}(f(S)), \qquad f(f^{-1}(T)) = T \cap \mathrm{im}\, f.$$

若 f 单射，第一式对所有 S 为等号；若 f 满射，第二式对所有 T 等于 T。

证明. 若 $a \in S$，则 $f(a) \in f(S)$，所以 $a \in f^{-1}(f(S))$。单射时，若 $f(a) \in f(S)$，存在 $s \in S$ 使 $f(a) = f(s)$，从而 $a = s \in S$，得到反向包含。

若 $b \in f(f^{-1}(T))$，则 $b = f(a)$，其中 $f(a) \in T$，故 $b \in T \cap \mathrm{im}\, f$。反之，若 $b \in T \cap \mathrm{im}\, f$，可写为 $b = f(a)$，且 $a \in f^{-1}(T)$，所以 b 属于左边。满射时 $\mathrm{im}\, f = B$，交集就是 T。 □

命题 3.7 (原像保持集合运算)**.** 对 B 的子集族 $(T_i)_{i \in I}$，有

$$f^{-1}\left(\bigcup_{i \in I} T_i\right) = \bigcup_{i \in I} f^{-1}(T_i),$$
$$f^{-1}\left(\bigcap_{i \in I} T_i\right) = \bigcap_{i \in I} f^{-1}(T_i),$$
$$f^{-1}(B \setminus T) = A \setminus f^{-1}(T).$$

这里 $T \subseteq B$。交集在左边以 B 为共同范围，在右边以 A 为共同范围；因此 $I = \varnothing$ 时，两处空交分别为 B 与 A，交集公式成为 $f^{-1}(B) = A$。空并均为空集。

证明. 以交集为例，对 $a \in A$，有

$$a \in f^{-1}\left(\bigcap_{i\in I} T_i\right) \iff \forall i \in I,\ f(a) \in T_i \iff \forall i \in I,\ a \in f^{-1}(T_i).$$

这正是右边的成员条件。把全称量词换成存在量词得到并集公式；使用否定得到补集公式。 □

像也保持任意并，但对交一般只有

$$f(S \cap S') \subseteq f(S) \cap f(S').$$

例如 $f(x) = x^2, S = \{-1\}, S' = \{1\}$ 时，左边为空集，右边为 $\{1\}$。右边的两个存在性条件可以由不同原像满足，左边却要求同一个原像同时属于两个集合。若 f 单射，两份原像必相同，等号便恢复。

3.4 等价关系：表示不同与对象相同

分数的写法 $1/2$ 与 $2/4$ 不同，却表示同一个有理数。要把这种情况写成一个构造，先把“表示”作为对象：取

$$D = \{(a, b) : a, b \in \mathbb{Z},\ b \neq 0\}.$$

$(1, 2)$ 与 $(2, 4)$ 是 D 的不同元素。我们希望把交叉乘积相等的两对归为同一类，定义

$$(a, b) \sim (c, d) \quad \iff \quad ad = bc.$$

这条规则只用整数运算，尚未把 a/b 当作已经构造好的有理数。它必须足以使“同一类”一致：每个表示与自身同类，两个表示之间没有方向差别，并且经过第三个表示的比较不能造成冲突。

定义 3.8 (二元关系与等价关系)**.** 集合 A 上的二元关系为 $A \times A$ 的一个子集，记 $a \sim b$ 表示 (a, b) 属于该子集。若对任意 $a, b, c \in A$ 都满足

1. $a \sim a$，称为自反性；
2. $a \sim b$ 蕴含 $b \sim a$，称为对称性；
3. $a \sim b$ 且 $b \sim c$ 蕴含 $a \sim c$，称为传递性，

则称它为等价关系。

例 3.9 (分数关系的验证)**.** 自反性来自 $ab = ba$，对称性来自等式的对称性。为验证传递性，设 $(a, b) \sim (c, d)$，且 $(c, d) \sim (e, h)$。则 $ad = bc, ch = de$。于是

$$adh = bch = bde.$$

因 $d \neq 0$，整数乘法消去 d 得 $ah = be$，即 $(a, b) \sim (e, h)$。分母不为零正是在这一步使用的。

例 3.10 (少一条条件会怎样)**.** 在 $\mathbb{R}$ 上定义 $x\ R\ y$ 当且仅当 $|x - y| < 1$。这个关系自反且对称，但 $0\ R\ 3/4$、$3/4\ R\ 3/2$，而 0 与 $3/2$ 不相关。因此按“与某点相关”形成的集合会相互重叠却不相同，无法形成稳定的分类。传递性排除了这种现象。

定义 3.11 (等价类与代表元)**.** 对 A 上的等价关系，元素 a 的等价类为 $[a]=\{x\in A : x\sim a\}$。称 a 为该类的一个代表元。

代表元是原集合的元素，等价类是原集合的子集。由自反性 $a\in[a]$。若换用同类的另一个代表元，得到的是不是同一个子集，需要证明。

定理 3.12 (等价类的相等判别)**.** 对 $a,b\in A$，有 $[a]=[b]$ 当且仅当 $a\sim b$。

证明. 先设 $a\sim b$。若 $x\in[a]$，则 $x\sim a$，与 $a\sim b$ 合用传递性得 $x\sim b$，所以 $x\in[b]$。这证明 $[a]\subseteq[b]$。由对称性 $b\sim a$，以同一论证得 $[b]\subseteq[a]$，于是两集合相等。

反之，若 $[a]=[b]$，由自反性 $a\in[a]$，故 $a\in[b]$。展开后一个成员条件即为 $a\sim b$。 □

推论 3.13 (相交的等价类相同)**.** 两个等价类或者相等，或者不相交。

证明. 若 $c\in[a]\cap[b]$，则 $c\sim a, c\sim b$。相等判别给出 $[c]=[a]$ 且 $[c]=[b]$，于是 $[a]=[b]$。 □

3.5 划分与商集：把一类作为一个元素

定义 3.14 (划分)**.** A 的一个划分是由非空子集组成的集合 $\mathcal{P}$，满足每个 $a\in A$ 恰好属于其中一块。等价地，各块两两不交，且并为 A。空集的划分取空族。

每个类 $[a]$ 由分离公理在 A 内构造。它由 a 唯一确定，所以替换公理保证所有类组成集合。也可以在 $\mathcal{P}(A)$ 内分离出恰好等于某个 $[a]$ 的子集，得到同一个集合。

定理 3.15 (等价关系与划分)**.** 集合 A 上的每个等价关系给出由其等价类组成的划分。反过来，每个划分都唯一确定一个等价关系：两个元素相关，当且仅当它们属于同一块。这两项构造互为逆构造。

证明. 给定等价关系。由自反性，每个 a 都属于 $[a]$，所以所有类非空且覆盖 A。相交的类相等，故不同的类不相交。这证明它们构成划分。

反过来，设 $\mathcal{D}$ 为 A 的划分。在 $A\times A$ 内定义关系 R：$a\,R\,b$ 当且仅当存在 $D\in\mathcal{D}$ 同时包含 a,b。每个 a 属于某一块，所以 $a\,R\,a$。若一块同时含 a,b，它也同时含 b,a，所以对称性成立。若 a,b 在块 D 中，b,c 在块 D' 中，由于两块共享元素 b，划分条件给出 $D=D'$。因此同一块含 a,c，得到传递性。

对 $a\in A$，令 D_a 为含 a 的唯一一块。若 $b\in[a]_R$，则某一块同时含 a,b，它必须是 D_a，所以 $b\in D_a$。若 $b\in D_a$，则 a,b 同块，故 $b\in[a]_R$。于是 $[a]_R=D_a$。每块非空，因而都是某个元素的类；这证明从划分出发再取等价类，返回原划分。

最后，从原等价关系出发，其划分中的同块关系等价于 $[a]=[b]$，又等价于 $a\sim b$，所以也返回原关系。若 $A=\varnothing$，唯一的关系与唯一的划分均为空集，结论仍成立。 □

定义 3.16 (商集与自然投影)**.** 全体等价类组成的集合记为 $A/\!\sim\;=\{[a] : a\in A\}$，称为商集。映射

$$q : A\longrightarrow A/\!\sim,\qquad a\longmapsto[a]$$

称为自然投影。每个类均有代表元，所以 q 是满射。

例如 $A=\{0,1,2,3,4,5\}$，按奇偶分类时，商集为

$$\big\{\{0,2,4\},\{1,3,5\}\big\}.$$

它有两个元素，每个元素本身是一个集合。于是 $0\in[0]$，$[0]\in A/\!\sim$，$[0]\subseteq A$，这三个式子表达三个不同的关系。

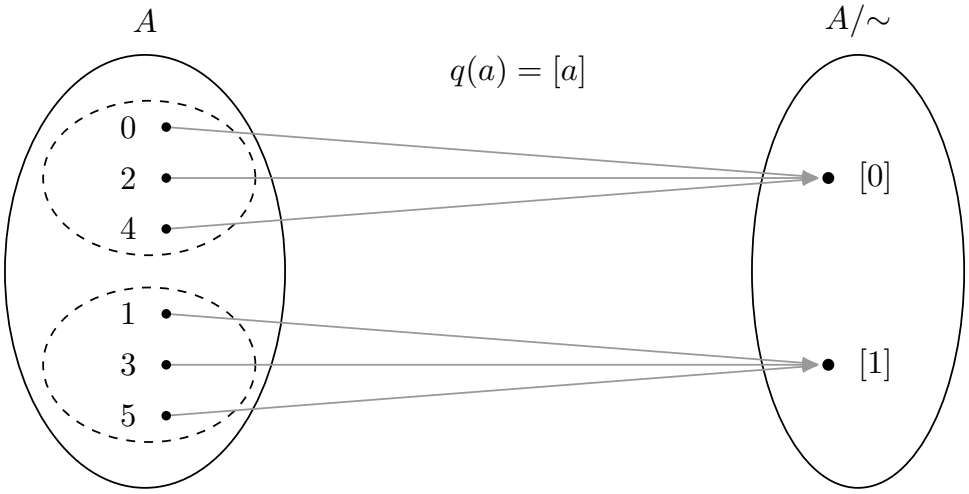


图 3.1: 自然投影与等价类

在分数的构造中，商集 $D/\!\sim$ 的一个元素是所有等价分数表示组成的一类。把类 $[(a,b)]$ 记为 a/b，得到有理数的集合模型。此时 $1/2=2/4$ 是等价类相等，而不是有序对 $(1,2)$ 与 $(2,4)$ 相等。

3.6 良定义性：计算能否只依赖对象

在商集上计算，通常先选一个代表元，再用原集合上的规则处理它。这样的公式要成为映射，同一个商集元素必须只有一个输出。

例 3.17 (一个成功的规则与一个失败的规则)**.** 仍在分数商集上，规则

$$[(a,b)]\longmapsto[(-a,b)]$$

与代表元无关：若 $ad=bc$，则 $(-a)d=b(-c)$，输出仍同类。因此它确实定义了商集到自身的映射。

规则 $[(a,b)]\mapsto a\in\mathbb{Z}$ 则失败，因为 $[(1,2)]=[(2,4)]$，而两个代表元给出整数 1 与 2。即使事先决定总选约分后的分数，也只是另行制定一个带选择步骤的新规则，不能证明原公式对任意代表元都成立。

定理 3.18 (商集上的映射与因子化)**.** *设 $f:A\to B$，$q:A\to A/\!\sim$ 为自然投影。存在映射 $\bar f:A/\!\sim\to B$ 满足 $f=\bar f\circ q$，当且仅当*

$$a\sim a'\quad\Longrightarrow\quad f(a)=f(a').$$

存在时 $\bar f$ 唯一，并由 $\bar f([a])=f(a)$ 给出。

证明. 必要性：若 $f = \bar{f} \circ q$，$a \sim a'$ 时有 $q(a) = q(a')$，故

$$f(a) = \bar{f}(q(a)) = \bar{f}(q(a')) = f(a').$$

充分性：对一个类 C，考虑满足“存在 $a \in C$ 使 $b = f(a)$”的输出 b。类的非空性保证输出存在。若 a, a' 是同一类中的两个代表元，则 $a \sim a'$，假设保证输出不变。因此每个类的输出唯一；在 $(A/\!\sim) \times B$ 内分离得到相应图像，便定义了映射 $\bar{f}$。对每个 a，有 $\bar{f}(q(a)) = \bar{f}([a]) = f(a)$，所以复合相等。

若 h 为另一个满足 $h \circ q = f$ 的映射，则对商集每个元素 $[a]$，

$$h([a]) = h(q(a)) = f(a) = \bar{f}([a]).$$

故 $h = \bar{f}$。这里自然投影的满射性保证已经检查了商集的每个元素。 □

定理中“f 在每一类上为常值”正是代表元公式的良定义性条件。“因子化”表示原映射可以分两步完成：先只记住等价类，再从类得到原来的输出。用图可写为

$$\begin{array}{ccc} A & \xrightarrow{\ f\ } & B \\ {\scriptstyle q}\downarrow & \nearrow_{\bar{f}} & \\ A/\!\sim & & \end{array}$$

两条从 A 到 B 的路径给出同一映射，称这张图交换。虚线箭头表示由定理得到的映射。一般的交换方形

$$\begin{array}{ccc} A & \xrightarrow{\ f\ } & B \\ {\scriptstyle u}\downarrow & & \downarrow{\scriptstyle v} \\ C & \xrightarrow[\ g\ ]{} & D \end{array}$$

表示 $v \circ f = g \circ u$；它把一个复合等式连同各映射的定义域、陪域同时写出。

3.7 每个映射都带来一种分类

设 $f : A \to B$。若只观察输出，两个输入无法区分的条件就是 $f(a) = f(a')$。定义

$$a \sim_f a' \quad \Longleftrightarrow \quad f(a) = f(a').$$

等号的自反、对称与传递保证这是等价关系。对 $b \in B$，集合 $f^{-1}(\{b\})$ 称为 b 上方的纤维；非空纤维恰好是此关系的各个等价类。

定理 3.19 (映射的集合因子分解)**.** 映射 $f : A \to B$ 给出双射

$$\theta : A/\!\sim_f \longrightarrow \operatorname{im} f, \qquad [a] \longmapsto f(a).$$

从而 f 是如下三步的复合：

$$A \xrightarrow{q} A/\!\sim_f \xrightarrow{\theta} \operatorname{im} f \xrightarrow{\iota} B,$$

其中 q 满射，θ 双射，ι 是子集的包含映射。

证明. 同一类中的输入有相同输出，由定理 3.18 得到良定义的 θ。若 $\theta([a]) = \theta([a'])$，则 $f(a) = f(a')$，故 $a \sim_f a'$，即 $[a] = [a']$，所以 θ 单射。任意 $b \in \operatorname{im} f$ 都可写成 $f(a)$，因而是 $\theta([a])$，所以 θ 满射。包含映射 ι 将元素视为 B 的元素，不改变其值，故三步复合在每个 a 处等于 $f(a)$。 □

这个分解在集合层面已经说明了商构造与像的关系。以后研究群同态和线性映射时，还要检验这些集合上的映射保持相应运算；其同构定理会再次出现这项构造。

3.8 选择函数与满射的右逆

设 $(A_i)_{i\in I}$ 是一族非空集合。一个映射

$$c : I \longrightarrow \bigcup_{i\in I} A_i$$

若满足 $c(i) \in A_i$，就称为这族集合的*选择函数*。这里每个 i 处的输出必须属于相应的 A_i；不是任取一个到并集的映射就能满足要求。

命题 3.20 (选择公理的函数形式)**.** *选择公理等价于：对任意集合 I 和集合族 $(A_i)_{i\in I}$，若每个 A_i 均非空，则该族有选择函数。指标集 I 可以为空。*

证明. 先假设上一章所述的选择公理。对每个 $i \in I$，令

$$D_i = \{(i, a) : a \in A_i\}.$$

D_i 非空；有序对的相等判别保证不同指标对应的 D_i 两两不交。因此选择公理给出 $C \subseteq \bigcup_i D_i$，使 C 与每个 D_i 的交恰有一个元素。C 的元素都是 (i, a)，每个指标 i 恰出现一次，且对应的 a 属于 A_i。所以 C 本身就是所需选择函数的图像。若 I 为空，空图像同样给出所需函数。

反过来，设每族非空集合都有选择函数，给定两两不交的非空集合族 $\mathcal{A}$。以 $\mathcal{A}$ 为指标集，取选择函数 c，其中 $c(A) \in A$。令 $C = \operatorname{im} c$。对每个 A，$c(A) \in C \cap A$；若 $x \in C \cap A$，则 $x = c(B) \in B$，其中 $B \in \mathcal{A}$。不交性迫使 $B = A$，故 $x = c(A)$。所以每个 $C \cap A$ 恰含一个元素，得到选择公理的原形式。 □

定理 3.21 (满射的右逆与选择公理)**.** *选择公理等价于：每个满射 $f : A \to B$ 都存在映射 $s : B \to A$，满足 $f \circ s = \mathrm{id}_B$。这样的 s 称为 f 的右逆。*

证明. 假设选择公理成立。满射性保证每个纤维 $A_b = f^{-1}(\{b\})$ 非空。对族 $(A_b)_{b\in B}$ 取选择函数 s，则 $s(b) \in A_b$，所以 $f(s(b)) = b$，即所需等式。若 $B = \varnothing$，存在 $A \to B$ 的映射已迫使 $A = \varnothing$，两个空映射满足等式。

反过来，设每个满射都有右逆。给定两两不交的非空集合族 $\mathcal{A}$，定义

$$p : \bigcup \mathcal{A} \longrightarrow \mathcal{A}, \qquad p(x) = \text{含有 } x \text{ 的那一个集合}.$$

并集定义保证这样的集合存在，不交性保证唯一，因此 p 良定义。每个 $A \in \mathcal{A}$ 非空，取其中一个 x 就有 $p(x) = A$，所以 p 满射。由假设取右逆 s。等式 $p(s(A)) = A$ 说明 $s(A) \in A$，故 s 是选择函数。上一命题给出选择公理。 □

右逆允许从每个纤维中指定一个代表。若纤维含多个元素，这个代表一般没有唯一性；相比之下，双射的纤维恰有一个元素，其逆映射不需要额外选择。

3.9 幂集比原集合有更多元素

有限集合的幂集很快变大。即使不先定义无限集合的基数，也能精确表达原集合与幂集之间的差别。

定理 3.22 (Cantor)**.** *对任意集合 A，不存在满射 $A \to \mathcal{P}(A)$。*

证明. 给定任意映射 $f : A \to \mathcal{P}(A)$。用分离构造

$$D = \{a \in A : a \notin f(a)\}.$$

这是 A 的子集，故 $D \in \mathcal{P}(A)$。若存在 $d \in A$ 使 $f(d) = D$，则

$$d \in D \quad \Longleftrightarrow \quad d \notin f(d) \quad \Longleftrightarrow \quad d \notin D,$$

矛盾。因此 D 不在 f 的像中，f 不是满射。 □

另一方面，$a \mapsto \{a\}$ 定义单射 $A \to \mathcal{P}(A)$：若 $\{a\} = \{b\}$，由于 $a \in \{a\}$，有 $a \in \{b\}$，从而 $a = b$。所以原集合可以一一放进它的幂集，但不能与幂集建立双射。这是一个同时适用于有限与无限集合的结论。

对角集合 D 对每个 a 都在位置 a 上与 $f(a)$ 作相反的成员决定。这项逐位置构造能够同时排除所有 $f(a)$，证明无需对 A 的元素作枚举。

3.10 偏序与幂集中的包含关系

等价关系把相关对象归为同类。另一种常见关系用于比较对象。对子集而言，包含关系保留了比较方向。

定义 3.23 (偏序、全序与良序)**.** 集合 P 上的关系 $\leq$ 若满足自反性 $x \leq x$、反对称性 “$x \leq y$ 且 $y \leq x$ 蕴含 $x = y$”、以及传递性，则称为*偏序*。称 $(P, \leq)$ 为偏序集。若任意 $x, y \in P$ 均有 $x \leq y$ 或 $y \leq x$，则称为全序。记 $x < y$ 表示 $x \leq y$ 且 $x \neq y$。若全序集中每个非空子集都有最小元，即有一个属于该子集且小于或等于其所有元素的元素，则称此全序为良序。

命题 3.24 (幂集的偏序)**.** *对任意集合 A，包含关系是 $\mathcal{P}(A)$ 上的偏序。*

证明. 每个子集 S 都有 $S \subseteq S$，得到自反性。若 $S \subseteq T$ 且 $T \subseteq S$，由外延公理 $S = T$，得到反对称性。若 $S \subseteq T \subseteq U$，任取 $x \in S$，先由第一个包含得 $x \in T$，再由第二个得 $x \in U$，所以 $S \subseteq U$，得到传递性。 □

若 $a \neq b$，子集 $\{a\}$ 与 $\{b\}$ 互不包含。因此偏序允许有不可比较的元素。反对称性也与对称性不同：对称性要求把每次相关反向后仍相关；反对称性则要求当双向相关确实同时出现时，两个对象相等。

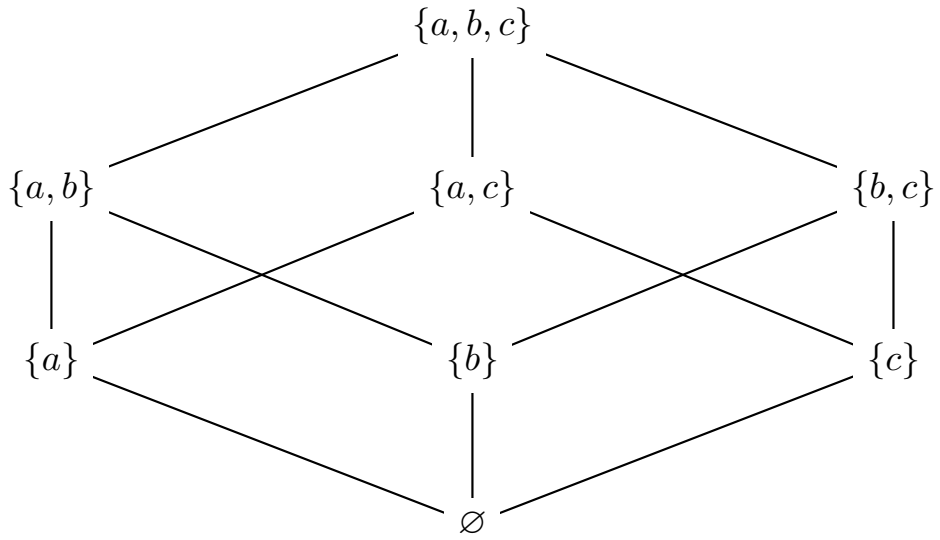


图 3.2: 幂集的包含关系

图中取 $A=\{a,b,c\}$，其中 a,b,c 互不相同。相邻层之间用线连接，表示较大的子集恰比小的多一个元素。沿线向上走表示包含；$\{a\}$ 与 $\{b,c\}$ 之间没有这样的向上路径，二者不可比较。图中线段的交叉处不另表示元素。这种只画出直接相邻比较的图称为 Hasse 图；线段的上下位置已经表达方向，因而无须再加箭头。

对子集族 $\mathcal{S}\subseteq\mathcal{P}(A)$，集合 $\bigcup\mathcal{S}$ 是包含每个 $S\in\mathcal{S}$ 的最小子集。因为若 x 属于并集，它属于某个 S，所以属于任何包含全部 S 的子集。相对于 A 的交集则是被每个 S 包含的最大子集：任何同时包含于各个 S 的集合，其元素都满足交集的成员条件。这样，并和交也可通过它们在偏序中的位置来刻画。

文献：映射、等价关系与商集的基础叙述参照 [3] 第 11、12 章。商集上的因子化在本章中由代表元公式直接证明。

4　自然数、归纳与递归

4.1　最小归纳集

归纳法用来证明一个性质对所有自然数成立，递归则用来定义一个在每个自然数处取值的对象。两者都涉及从一项到下一项，却承担不同任务。要知道它们为何有效，先要说明自然数集本身的构造。

继续用 $S(x) = x \cup \{x\}$ 表示后继集合。由无穷公理，取一个归纳集 I。考虑 I 的所有归纳子集：

$$\mathcal{I} = \{J \in \mathcal{P}(I) : \varnothing \in J \text{ 且} \forall x \in J\ (S(x) \in J)\}.$$

幂集与分离保证 $\mathcal{I}$ 是集合；$I \in \mathcal{I}$，所以它非空。定义

$$\omega = \bigcap \mathcal{I}.$$

定理 4.1 (自然数集的最小性质)**.** *ω 是归纳集，且包含于每个归纳集。特别地，它不依赖最初所取的归纳集 I。*

证明. 每个 $J \in \mathcal{I}$ 都含 $\varnothing$，故 $\varnothing \in \omega$。若 $x \in \omega$，则对每个 $J \in \mathcal{I}$ 有 $x \in J$，再由 J 归纳得到 $S(x) \in J$。于是 $S(x) \in \omega$，证明 ω 归纳。

设 K 为任意归纳集。$I \cap K$ 含空集；若 $x \in I \cap K$，两个集合的归纳性都给出 $S(x)$ 属于其中，所以 $I \cap K$ 也是归纳集。它还是 I 的子集，故属于 $\mathcal{I}$。由交集的定义，$\omega \subseteq I \cap K \subseteq K$。

若从另一个归纳集出发构造出 ω'，它也具有上述最小性质。于是 $\omega \subseteq \omega'$ 且 $\omega' \subseteq \omega$，两者相等。 □

定义 4.2 (自然数)**.** ω 的元素称为自然数。记 $\mathbb{N} = \omega$，$0 = \varnothing$，$1 = S(0)$，$2 = S(1)$，$3 = S(2)$，依此命名。通常把 $S(n)$ 记为 $n + 1$；在加法构造以前，这个写法只表示后继。

最初几个对象为

$$0 = \varnothing, \qquad 1 = \{0\}, \qquad 2 = \{0, 1\}, \qquad 3 = \{0, 1, 2\}.$$

这里 2 有元素 $0, 1$；$\{2\}$ 则只有一个元素 2。数字的这项集合编码使大小关系也能用成员关系表示，但这需要证明。

4.2　归纳法来自什么

定理 4.3 (数学归纳法)**.** *若 $E \subseteq \mathbb{N}$，$0 \in E$，且对每个 $n \in E$ 有 $S(n) \in E$，则 $E = \mathbb{N}$。*

证明. 这些条件说明 E 是归纳集。定理 4.1 给出 $\mathbb{N} \subseteq E$；与假设的反向包含合用即得等式。 □

若要证明性质 $P(n)$ 对所有自然数成立，用分离取出

$$E = \{n \in \mathbb{N} : P(n)\}.$$

验证 $P(0)$，就是验证 $0 \in E$；证明 $P(n)$ 蕴含 $P(S(n))$，就是验证 E 对后继封闭。上一定理于是给出 $E = \mathbb{N}$。归纳步骤中的 n 是任意自然数，因此这一步证明的是一个统一的蕴含，而不是只检查某几个数字。

定义 4.4 (传递集合)**.** 集合 T 称为传递的，若 $x \in T$ 总蕴含 $x \subseteq T$。等价地，若 $y \in x \in T$，则 $y \in T$。

引理 4.5 (自然数的传递性)**.** 每个 $n \in \mathbb{N}$ 都是传递集合，且 $n \subseteq \mathbb{N}$。因此 $\mathbb{N}$ 自身也传递。

证明. 先对 $n \subseteq \mathbb{N}$ 使用归纳法。$0 = \varnothing \subseteq \mathbb{N}$。若 $n \subseteq \mathbb{N}$，由于 $n \in \mathbb{N}$，有 $S(n) = n \cup \{n\} \subseteq \mathbb{N}$，所以归纳步骤成立。

再证明每个 n 传递。空集传递，因为它没有需要检查的元素。假设 n 传递，取 $x \in S(n)$。若 $x = n$，有 $x \subseteq S(n)$；若 $x \in n$，由传递性有 $x \subseteq n \subseteq S(n)$。所以 $S(n)$ 传递。最后，每个 $n \in \mathbb{N}$ 都是 $\mathbb{N}$ 的子集，正是 $\mathbb{N}$ 传递的定义。 □

命题 4.6 (零与后继)**.** 0 不是任何自然数的后继；$S(m) = S(n)$ 蕴含 $m = n$；每个非零自然数都是唯一自然数的后继。

证明. $S(n)$ 含元素 n，而 0 没有元素，故 $S(n) \neq 0$。

对于传递集合 n，有 $\bigcup S(n) = n$。一方面，$n \in S(n)$，故 $n \subseteq \bigcup S(n)$。另一方面，若 $t \in \bigcup S(n)$，存在 $x \in S(n)$ 使 $t \in x$。$x = n$ 时 $t \in n$；$x \in n$ 时，由 n 传递仍有 $t \in n$。因此等式成立。在 $S(m) = S(n)$ 两边取并，便得 $m = n$。

用替换与并集构造

$$D = \{0\} \cup \{S(n) : n \in \mathbb{N}\}.$$

它是 $\mathbb{N}$ 的子集，含 0，且若 $x \in D$，则 $x \in \mathbb{N}$，所以 $S(x) \in D$。归纳法给出 $D = \mathbb{N}$。因此每个非零自然数都是某个后继；唯一性来自刚证明的后继单射性。 □

这些结果给出了自然数的归纳结构：一个起始元素、一个不返回起点的单射后继，以及归纳原理。代数中的递推公式和有限次构造，正是建立在这项结构上。

4.3 成员关系成为自然数的次序

对自然数 m, n，定义

$$m < n \iff m \in n, \qquad m \leq n \iff (m < n \text{ 或} m = n).$$

这时还需要验证两个不同自然数一定可比较。

引理 4.7 (每个自然数内部的良序)**.** 对任意 $n \in \mathbb{N}$，成员关系在 n 上是严格全序，且 n 的每个非空子集都有最小元。

证明. 对 n 归纳。空集上没有要比较的元素，也没有非空子集，结论成立。假设结论对 n 成立。$S(n) = n \cup \{n\}$ 比 n 多出元素 n；$n \notin n$ 已由正则公理证明。每个旧元素 $x \in n$ 都满足 $x < n$，而 $n \in x$ 不可能成立，否则形成二元成员循环。旧元素之间保持原先的严格全序，所以新元素是最大的元素。

传递性也可直接核对：若 $x \in y \in z$，其中 $z \in S(n)$，则 z 是自然数，故由引理 4.5 传递，得到 $x \in z$。

若 $E \subseteq S(n)$ 非空且 $E \cap n \neq \varnothing$，由归纳假设，$E \cap n$ 有最小元 e。它小于所有其他旧元素，也小于新元素 n，故为 E 的最小元。若 $E \cap n = \varnothing$，则 $E = \{n\}$，其最小元为 n。 □

引理 4.8 (传递子集是初始部分)**.** *若 $T \subsetneq n$ 且 T 传递，则 $T \in n$。*

证明. 由严格包含，$n \setminus T$ 非空，取它在 n 的良序下的最小元 k。若 $x \in k$，由 n 传递有 $x \in n$；k 的最小性排除 $x \in n \setminus T$，故 $x \in T$。所以 $k \subseteq T$。

反过来，取 $x \in T$。x, k 都在 n 中，严格全序给出三种情形。$x = k$ 与 $k \notin T$ 矛盾；若 $k \in x$，由 T 传递与 $x \in T$ 得 $k \in T$，也矛盾。故只能 $x \in k$。因此 $T \subseteq k$，最终得到 $T = k \in n$。 □

定理 4.9 (自然数的比较与良序)**.** *对任意 $m, n \in \mathbb{N}$，$m < n, m = n, n < m$ 恰有一个成立；且 $\mathbb{N}$ 的每个非空子集都有最小元。*

证明. 令 $T = m \cap n$。由于 m, n 传递，T 也传递：若 $x \in T$ 且 $y \in x$，则 $y \in m, n$，所以 $y \in T$。若 T 同时是真包含于 m, n 的子集，引理 4.8 给出 $T \in m$ 和 $T \in n$，从而 $T \in m \cap n = T$，与没有自身成员矛盾。因此 $T = m$ 或 $T = n$。

这说明 $m \subseteq n$ 或 $n \subseteq m$。若包含严格，再次用引理 4.8，得到较小的集合是较大的集合的元素。这证明三种情形至少有一种。自身成员与二元成员循环均不可能，故三种情形两两排斥。

现在设 $E \subseteq \mathbb{N}$ 非空，取 $a \in E$。若 $E \cap a = \varnothing$，就没有 E 的元素小于 a，故 a 最小。若 $E \cap a \neq \varnothing$，由引理 4.7，取此交的最小元 b。假设有 $c \in E$ 小于 b。由于 $b \in a$ 且 a 传递，$c \in a$，于是 $c \in E \cap a$，与 b 最小矛盾。因此 b 是整个 E 的最小元。 □

推论 4.10 (强归纳法)**.** *若一个性质满足：对每个 $n \in \mathbb{N}$，假定所有 $m < n$ 都有此性质，就能证明 n 也有此性质，则所有自然数都有此性质。*

证明. 若反例集非空，取其最小元 n。所有 $m < n$ 都不是反例，由假设可得 n 也不是反例，矛盾。 □

强归纳法允许同时使用所有更小的情形。在分解一个对象、把它化为几个较小对象时，这比只写相邻两项的归纳假设更贴合证明的实际需要。

4.4　递归定义为什么存在

设 X 是集合，$x_0 \in X$，并给定映射 $F : \mathbb{N} \times X \to X$。我们希望规定一个函数 $u : \mathbb{N} \to X$，满足

$$u(0) = x_0, \qquad u(S(n)) = F(n, u(n)).$$

等式告诉我们已有一项后如何得到下一项，但还没有直接给出包含所有函数值的集合。以下证明把有限片段相容地接起来。

定理 4.11 (自然数上的递归)**.** *在上述条件下，存在唯一映射 $u : \mathbb{N} \to X$ 满足这两个递归等式。*

证明. 称映射 $v : S(n) \to X$ 为一个合法片段，若 $v(0) = x_0$，且对每个 $k < n$，有

$$v(S(k)) = F(k, v(k)).$$

$S(n)$ 的元素正是 0 到 n，所以这个片段给出了截至第 n 项的值。

先证明每个 n 都有合法片段。$n = 0$ 时，图像 $\{(0, x_0)\}$ 给出所需映射。假设 $v : S(n) \to X$ 合法。在新位置 $S(n)$ 赋值 $F(n, v(n))$，把图像扩为

$$\Gamma_v \cup \{(S(n), F(n, v(n)))\}.$$

$S(n) \notin S(n)$，故新输入原先不在定义域中，不会与已有值冲突。所得映射的定义域为 $S(S(n))$。原有递归等式保持不变，新增的最后一个等式由定义满足，故得到下一合法片段。归纳法证明所需存在性。

再证明任意两个合法片段在共同定义域上一致。设定义域分别为 $S(m), S(n)$。由自然数比较，不妨设 $m \le n$，从而 $S(m) \subseteq S(n)$。在 0 处两个片段都取值 x_0；若已知在 $k < m$ 处取值相同，两个递归等式给出它们在 $S(k)$ 处的值也相同。归纳法使它们在 $S(m)$ 的每个元素处一致。

所有合法片段的图像是 $\mathcal{P}(\mathbb{N} \times X)$ 中满足上述条件的元素，故由分离组成一个集合 $\mathcal{V}$。取

$$\Gamma = \bigcup \mathcal{V}.$$

相容性保证同一输入不会对应两个不同输出；每个 n 都有合法片段，保证每个输入至少有一个输出。因此 Γ 是某个映射 $u : \mathbb{N} \to X$ 的图像。

它在 0 处取值 x_0。对任意 n，取定义域为 $S(S(n))$ 的合法片段；它同时给出位置 $n, S(n)$ 的值，并与 u 一致，因此 $u(S(n)) = F(n, u(n))$。

最后，若 w 也满足递归等式，则 $u(0) = w(0)$；若 $u(n) = w(n)$，应用 $F(n, -)$ 得 $u(S(n)) = w(S(n))$。归纳法给出每个位置相等，故 $u = w$。 □

存在性的证明同时解释了三个问题：有限片段确实能延长，不同片段不会冲突，把所有片段接起来仍是集合。唯一性的证明则说明，满足同一递归条件的任何构造都给出同一个函数。

例 4.12 (重复施行一个变换)**.** 给定 $T : X \to X$ 和 $x \in X$，取 $F(n, y) = T(y)$，递归定理给出唯一序列 $u(0) = x$，$u(S(n)) = T(u(n))$。这使“对 x 施行 n 次 T”具有精确含义。若把函数本身作为对象，在集合 X^X 中递归，还可以定义 $T^0 = \mathrm{id}_X$，$T^{S(n)} = T \circ T^n$。

4.5 加法的构造与运算律

对每个固定的 $a \in \mathbb{N}$，将递归定理用于 $X = \mathbb{N}$，初值 a，以及变换 $y \mapsto S(y)$，得到唯一映射。把它在 b 处的值记为 $a + b$。于是

$$a + 0 = a, \qquad a + S(b) = S(a + b).$$

由于 $1 = S(0)$，有 $a + 1 = a + S(0) = S(a + 0) = S(a)$，所以加一确实与后继一致。每个 (a, b) 都有唯一输出，替换公理保证这项规则给出 $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ 的映射。计算 $2 + 3$，就是从 2 开始连续取三次后继。

引理 4.13 (零与左侧后继)**.** 对任意自然数 a, b，有 $0 + b = b$，以及 $S(a) + b = S(a + b)$。

证明. 对第一个等式按 b 归纳。$0+0=0$；若 $0+b=b$，则

$$0+S(b)=S(0+b)=S(b).$$

对第二个等式也按 b 归纳。$S(a)+0=S(a)=S(a+0)$。若等式对 b 成立，则

$$S(a)+S(b)=S(S(a)+b)=S(S(a+b))=S(a+S(b)).$$

其中第一个、最后一个等号来自加法的递归定义，中间使用归纳假设。 □

定理 4.14 (加法结合律与交换律). *对任意 $a,b,c\in\mathbb{N}$，有*

$$(a+b)+c=a+(b+c),\qquad a+b=b+a.$$

证明. 结合律按 c 归纳。$c=0$ 时两边都是 $a+b$。若它对 c 成立，则

$$\begin{aligned}(a+b)+S(c)&=S((a+b)+c)\\&=S(a+(b+c))\\&=a+S(b+c)=a+(b+S(c)).\end{aligned}$$

这证明结合律。

交换律按 b 归纳。$a+0=a=0+a$，其中第二个等号由引理 4.13 得到。若 $a+b=b+a$，则

$$a+S(b)=S(a+b)=S(b+a)=S(b)+a,$$

最后一步再次使用该引理。因此交换律成立。 □

命题 4.15 (加法消去律). *若 $a+c=b+c$，则 $a=b$。*

证明. 对 c 归纳。$c=0$ 时由加法定义直接得到结论。若结论对 c 成立，而 $a+S(c)=b+S(c)$，则 $S(a+c)=S(b+c)$。后继的单射性给出 $a+c=b+c$，再用归纳假设得 $a=b$。 □

这里没有使用减法。自然数中并非每个减法都能进行；消去律是由后继单射性得到的运算性质。

命题 4.16 (次序与加法的关系). *对 $a,b\in\mathbb{N}$，有*

$$a\le b\quad\Longleftrightarrow\quad\exists c\in\mathbb{N}\ (a+c=b).$$

证明. 先对 c 归纳证明 $a\le a+c$。$c=0$ 时为相等。若 $a\le a+c$，则 $a\in S(a+c)$：相等时由后继定义得到；严格小于时由 $a\in a+c\subseteq S(a+c)$ 得到。故 $a\le a+S(c)$。

反过来，对 b 归纳证明：每当 $a\le b$，便存在所需的 c。$b=0$ 时，因 0 没有元素，只有 $a=0$，取 $c=0$。若结论对 b 成立，考察 $a\le S(b)$。若 $a=S(b)$，取 $c=0$；否则 $a\in S(b)=b\cup\{b\}$，所以 $a\le b$。由归纳假设写成 $a+c=b$，于是 $a+S(c)=S(b)$。 □

4.6　乘法、分配律与幂

对固定 $a \in \mathbb{N}$，再次应用递归定理，以 0 为初值，以下一步加上 a 为递归规则，定义 ab：

$$a0 = 0, \qquad aS(b) = ab + a.$$

例如 $a2 = (a0 + a) + a = 0 + a + a = a + a$。乘法表示有限次累加；只有证明相应运算律后，才能自由交换和组合乘积。

引理 4.17 (乘法的零与左侧后继)**.** $0b = 0$，且 $S(a)b = ab + b$。

证明. $00 = 0$；若 $0b = 0$，则 $0S(b) = 0b + 0 = 0$，所以第一式成立。

第二式对 b 归纳。$b = 0$ 时两边为零。假设 $S(a)b = ab + b$，则

$$\begin{aligned} S(a)S(b) &= S(a)b + S(a) \\ &= (ab + b) + S(a) \\ &= S(ab + b + a), \\ aS(b) + S(b) &= (ab + a) + S(b) \\ &= S(ab + a + b). \end{aligned}$$

加法结合律与交换律给出 $ab + b + a = ab + a + b$，故两式相等，完成归纳。 □

定理 4.18 (乘法的基本运算律)**.** 对 $a, b, c \in \mathbb{N}$，有

$$ab = ba, \qquad a(b + c) = ab + ac, \qquad (ab)c = a(bc),$$

且 $a1 = 1a = a$。

证明. 先按 b 归纳证明交换律。$a0 = 0 = 0a$。若 $ab = ba$，则

$$aS(b) = ab + a = ba + a = S(b)a,$$

最后一步使用引理 4.17。

分配律按 c 归纳。$c = 0$ 时 $a(b + 0) = ab = ab + a0$。若等式对 c 成立，则

$$\begin{aligned} a(b + S(c)) &= aS(b + c) = a(b + c) + a \\ &= (ab + ac) + a = ab + (ac + a) = ab + aS(c). \end{aligned}$$

这证明所写的一侧分配律；另一侧由乘法交换律得到：$(a + b)c = c(a + b) = ca + cb = ac + bc$。

结合律按 c 归纳。$c = 0$ 时两边为零。若 $(ab)c = a(bc)$，则

$$\begin{aligned} (ab)S(c) &= (ab)c + ab = a(bc) + ab \\ &= a(bc + b) = a(bS(c)). \end{aligned}$$

其中倒数第二个等号使用刚证明的分配律。

最后 $a1 = aS(0) = a0 + a = 0 + a = a$，再由交换律得 $1a = a$。 □

若 $a \neq 0$，由命题 4.6 可写为 $S(r)$。因此当 a, b 均非零时，写 $b = S(s)$，就有

$$ab = aS(s) = as + a = as + S(r) = S(as + r) \neq 0.$$

这证明自然数乘法没有两个非零因子的零乘积，最后一步使用的是零不是后继。

对固定 a，递归定理还给出幂运算

$$a^0 = 1, \qquad a^{S(n)} = a^n a.$$

此约定也给出 $0^0 = 1$，它表示空乘积。每一种递归定义都先确定初值和下一步规则，再由递归定理保证整个函数存在且唯一。

例 4.19 (等差求和的逐步证明)**.** 用递归定义 $s(0) = 0$，$s(S(n)) = s(n) + S(n)$，即从 1 加到 n 的和。我们证明

$$2s(n) = n(n+1).$$

$n = 0$ 时两边为零。若结论对 n 成立，则使用分配律与加法结合律，得到

$$\begin{aligned} 2s(n+1) &= 2s(n) + 2(n+1) \\ &= n(n+1) + 2(n+1) \\ &= (n+2)(n+1) = (n+1)((n+1)+1). \end{aligned}$$

这正是下一项的等式。由归纳法，等式对所有自然数成立。将它写成 $s(n) = n(n+1)/2$，还涉及整数或有理数中的除法；原式本身已经在自然数运算内完成了证明。

文献：Weiss[2] 第 4 章通过相容的有限函数片段建立递归，并据此定义算术。本章采用最小归纳集构造自然数，对次序、递归和运算律分别给出证明。

5 整数与整除

5.1 从自然数到整数：让差也成为数

自然数中，方程 $5+x=3$ 没有解。我们希望加入一个数来记录“3 减去 5”，但不能把尚未定义的减法拿来作定义。可以先用有序对 $(3,5)$ 记录这项要求。$(4,6)$ 应当记录同一个差，因为两边同时增加 1 不应改变结果。一般地，用交叉相加来比较两种表示。

定义 5.1 (整数的差对模型)**.** 在 $\mathbb{N}\times\mathbb{N}$ 上定义

$$(a,b)\sim(c,d)\quad\Longleftrightarrow\quad a+d=c+b.$$

令 $\mathbb{Z}=(\mathbb{N}\times\mathbb{N})/\!\sim$，把类 $[(a,b)]$ 理解为由 a,b 记录的差。

这个关系确为等价关系。自反性、对称性由加法交换律和等式的对称性得到。若 $a+d=c+b$、$c+f=e+d$，则

$$a+d+f=c+b+f=e+d+b.$$

消去共同加数 d，得 $a+f=e+b$，即传递性。这里全程只用了自然数加法。

命题 5.2 (整数运算与自然数的嵌入)**.** *下列公式与代表元无关：*

$$\begin{aligned}[(a,b)]+[(c,d)]&=[(a+c,b+d)],\\ [(a,b)]\,[(c,d)]&=[(ac+bd,ad+bc)],\\ -[(a,b)]&=[(b,a)].\end{aligned}$$

它们使 $\mathbb{Z}$ 成为交换有单位环，零元为 $[(0,0)]$，单位元为 $[(1,0)]$。映射 $\iota:\mathbb{N}\to\mathbb{Z}$，$n\mapsto[(n,0)]$，为单射且保持加法、乘法和单位元。

证明. 先验证良定义性。设 $(a,b)\sim(a',b')$，即 $a+b'=a'+b$。与固定 (c,d) 相加后，所需等式为

$$(a+c)+(b'+d)=(a'+c)+(b+d),$$

它由原等式在两边加 $c+d$ 得到。交换 a,b 后，原等式仍保证 $(b,a)\sim(b',a')$，故取负良定义。对于乘法，将原等式分别乘 c,d，再重排得

$$\begin{aligned}(ac+bd)+(a'd+b'c)&=(ac+b'c)+(bd+a'd)\\ &=(a'c+bc)+(ad+b'd)\\ &=(a'c+b'd)+(ad+bc).\end{aligned}$$

这正是两个乘积表示等价的条件。乘法公式对两组输入对称，因此更换第二个代表元也不改变结果。

加法结合律、交换律由逐坐标的自然数加法得到。并且

$$[(a,b)]+[(b,a)]=[(a+b,a+b)]=[(0,0)],$$

所以取负给出加法逆元。乘法交换律和单位元公式直接由定义得到。验证乘法结合律时，两种括号得到的代表元均为

$$(ace + adf + bcf + bde,\ acf + ade + bce + bdf),$$

其中第三个因子为 $[(e,f)]$。分配律也由自然数分配律逐坐标展开得到；例如把 $[(a,b)]$ 乘以 $[(c+e,d+f)]$，两坐标分别为 $ac+ae+bd+bf$ 和 $ad+af+bc+be$，与先乘后加相同。因此全部环运算律成立。

最后，$[(n,0)] = [(m,0)]$ 等价于 $n = m$，证明单射；把 $b = d = 0$ 代入运算公式，即见 ι 保持加法与乘法，并保持 $0, 1$。 □

以下把 $\iota(n)$ 简记为 n，即把自然数与它在整数中的副本认同。自然数次序的全序性和命题 4.16 说明：若 $b \le a$，存在唯一 $n \in \mathbb{N}$ 使 $a = b + n$，于是 $[(a,b)] = [(n,0)] = n$；若 $a \le b$，同理 $[(a,b)] = [(0,n)] = -n$。除 $0 = -0$ 外，这些表示互不相同：两个同号表示的相等归结为自然数相等，而 $[(n,0)] = [(0,m)]$ 要求 $n + m = 0$，只能有 $n = m = 0$。

定义 5.3 (整数的次序与绝对值)**.** 对 $x, y \in \mathbb{Z}$，定义 $x \le y$ 当且仅当 $y - x$ 属于上述自然数副本。若 $x \neq y$ 且 $x \le y$，写 $x < y$。定义 $|x| = x$（$x \ge 0$ 时），$|x| = -x$（$x < 0$ 时）。

这里 $y - x$ 已有意义，它是环中的 $y + (-x)$。每个整数差恰为非负或非正，且两者同时成立只在差为零时发生，这给出全序性和反对称性。若 $y - x$、$z - y$ 均为自然数，则它们的和 $z - x$ 也是自然数，给出传递性。加上同一个整数不改变差，故次序与加法相容。两个正整数的乘积仍为正，因为它们来自两个非零自然数，而自然数的乘积非零。因此 $x \le y$、$z \ge 0$ 时有 $xz \le yz$。在自然数副本上，这个次序与原次序一致，仍满足良序性。又因正整数至少为 1，任意两个相邻整数之间没有别的整数。

这些事实提供了下文使用的有序整数运算法则。例如 $[(3,5)] = -2$，而 $-(-2) = 2$；所有整数的加减乘法现在都可以从自然数运算和商集构造中得到。

5.2 整除如何转化为等式

整数方程 $ax = b$ 若允许实数解，在 $a \neq 0$ 时可直接除以 a；若要求 x 是整数，问题就变为 b 是否为 a 的整数倍。整除因此记录了一种受限制的可解性。

定义 5.4 (整除)**.** 对 $a, b \in \mathbb{Z}$，若存在 $k \in \mathbb{Z}$ 满足 $b = ak$，称 a 整除 b，记为 $a \mid b$。a 是 b 的因子，b 是 a 的倍数。

整除式提供一个存在性条件。要使用 $a \mid b$，通常把它写为 $b = ak$；要证明 $a \mid c$，则要找到 $c = a\ell$ 的整数系数 ℓ。

命题 5.5 (整除与整数线性组合)**.** *若 $d \mid a, d \mid b$，则对任意 $u, v \in \mathbb{Z}$，有 $d \mid ua + vb$。*

证明. 写 $a = dr, b = ds$，其中 $r, s \in \mathbb{Z}$。则 $ua + vb = d(ur + vs)$，且 $ur + vs$ 是整数。 □

以后把 $ua + vb$ 称为 a, b 的*整数线性组合*。公因数必须整除每一个这样的组合。一个关键问题是：能否反过来从这些组合中找出控制全部公因数的整数？

5.3 带余除法：存在性与唯一性

定理 5.6 (带余除法)**.** *若 $a \in \mathbb{Z}, b > 0$，则存在唯一整数 q, r 满足*

$$a = bq + r, \qquad 0 \le r < b.$$

q 称为商，r 称为余数。

证明. 考虑非负差值的集合

$$S = \{a - bk : k \in \mathbb{Z},\ a - bk \ge 0\}.$$

取 $k = -|a|$，因 $b \ge 1$，有 $a - bk = a + b|a| \ge a + |a| \ge 0$，故 S 非空。由自然数的良序性，S 有最小元 $r = a - bq$。若 $r \ge b$，则 $r - b = a - b(q+1) \ge 0$ 仍在 S，又小于 r，矛盾。因此 $0 \le r < b$，证明存在性。

设还有 $a = bq' + r'$，其中 $0 \le r' < b$。两式相减得 $b(q - q') = r' - r$。由余数范围，$-b < r' - r < b$，故

$$-b < b(q - q') < b, \qquad -1 < q - q' < 1.$$

整数 $q - q'$ 只能为 0，故 $q = q'$；再代回相减等式得到 $r = r'$。 □

例 5.7 (负整数的余数)**.** $-7 = 3(-3) + 2$，所以除以 3 的余数是 2。表达式 $-7 = 3(-2) - 1$ 虽然成立，却不满足规定的余数范围。区间 $0 \le r < b$ 同时决定存在性问题中的目标和唯一性问题中的约束。

5.4 欧几里得算法与 Bézout 等式

定义 5.8 (最大公因数与互素)**.** 若 a, b 不全为零，其正公因数中的最大者称为最大公因数，记为 $\gcd(a, b)$。最大公因数为 1 时称 a, b 互素。

最大公因数存在：1 是正公因数，任取 a, b 中非零的一项，正公因数均不超过这一项的绝对值，因此只有有限个候选。

命题 5.9 (更换一对整数而保留公因数)**.** *若 $a = bq + r$，则 a, b 的公因数与 b, r 的公因数完全相同。因此在二者不全为零时，$\gcd(a, b) = \gcd(b, r)$。*

证明. 若 d 整除 a, b，它也整除 $r = a - bq$。若 d 整除 b, r，它也整除 $a = bq + r$。这分别证明两组公因数之间的两个包含关系。 □

将较大的正整数除以较小的，再用除数和余数继续计算，就逐步把问题转到更小的整数上。非零余数严格递减，因而不能无限继续。这一过程称为*欧几里得算法*。

例 5.10 (计算与回代)**.** 取 $a = 252, b = 198$，有

$$\begin{aligned} 252 &= 198 + 54, \\ 198 &= 3 \cdot 54 + 36, \\ 54 &= 36 + 18, \\ 36 &= 2 \cdot 18. \end{aligned}$$

每一步保留公因数，最后得到 $\gcd(252,198)=\gcd(18,0)=18$。回代则给出另一种信息：

$$\begin{aligned}18 &= 54-36\\ &= 54-(198-3\cdot 54)=4\cdot 54-198\\ &= 4(252-198)-198=4\cdot 252-5\cdot 198.\end{aligned}$$

所以最大公因数本身也是原来两数的整数线性组合。

定理 5.11 (Bézout 等式及其整除意义)**.** *若 a,b 不全为零，令 $d=\gcd(a,b)$，则存在 $u,v\in\mathbb{Z}$ 使 $d=ua+vb$。每个公因数都整除 d，并且*

$$\{ax+by : x,y\in\mathbb{Z}\}=d\mathbb{Z},$$

其中 $d\mathbb{Z}=\{dk : k\in\mathbb{Z}\}$。

证明. 考虑所有正的整数线性组合，取其中最小者 $m=ua+vb$。它们存在，因为 a,b 中至少一个非零，取它或它的相反数即可得到正组合。对 a 除以 m，写 $a=mq+r$，其中 $0\le r<m$。则

$$r=a-q(ua+vb)=(1-qu)a-qvb$$

仍是整数线性组合。若 $r>0$，与 m 的最小性矛盾，所以 $r=0$，即 $m\mid a$。对 b 除以 m，写 $b=mt+s$，$0\le s<m$，则

$$s=b-t(ua+vb)=-tua+(1-tv)b.$$

若 $s>0$，又得到小于 m 的正组合，矛盾。因此 $s=0$，即 $m\mid b$。

任意公因数都整除 $ua+vb=m$。因此 m 是正公因数，且其他正公因数不超过它，故 $m=d$。这既证明 Bézout 等式，也证明公因数都整除 d。

因 $d\mid a,b$，每个 $ax+by$ 都属于 $d\mathbb{Z}$。反向由 $d=ua+vb$，对任意 $k\in\mathbb{Z}$ 有 $dk=a(uk)+b(vk)$，也属于左边。这证明集合相等。 □

欧几里得算法给出了实际求系数的方法；上面的最小正组合证明解释了为何所有组合最终由一个整数控制。

推论 5.12 (整数线性方程的可解性)**.** *设 $a,b,c\in\mathbb{Z}$，且 a,b 不全为零。方程 $ax+by=c$ 有整数解，当且仅当 $\gcd(a,b)\mid c$。*

证明. 必要性来自公因数整除每个线性组合。若 $c=dk$，将 Bézout 等式乘以 k，得到一组解 $x=uk,y=vk$。 □

例如 $252x+198y=36$ 有解 $(8,-10)$，但等于 1 时无整数解。这是最大公因数对整数方程可解性的直接控制。

5.5 素数与唯一分解

定义 5.13 (单位与素数)**.** 具有整数乘法逆元的整数称为单位，恰为 $1,-1$。整数 $p>1$ 称为素数，若其正因子只有 $1,p$。

素数定义说它自身不能分成两个更小正整数的乘积；唯一分解证明还需要它在一个乘积中怎样出现。

引理 5.14 (Euclid 引理)**.** *若 $\gcd(a,b)=1$ 且 $a \mid bc$，则 $a \mid c$。特别地，素数 p 若整除 bc，则 $p \mid b$ 或 $p \mid c$。*

证明. 由 Bézout 等式取 $ua+vb=1$；由 $a \mid bc$ 写 $bc=ak$，$k\in\mathbb{Z}$。于是

$$c=(ua+vb)c=uac+vbc=a(uc+vk).$$

括号内为整数，故 $a \mid c$。若 p 为素数且不整除 b，则 p,b 的正公因数只能为 1，从而 $\gcd(p,b)=1$，应用前句即可。 □

互素条件确有作用：$6 \mid 2\cdot 3$，但 6 既不整除 2，也不整除 3。所以不能对任意整数直接使用素数的乘积性质。

定理 5.15 (整数唯一分解)**.** *每个整数 $n>1$ 均可写为有限个素数的乘积；这样的表达式只相差因子的排列次序。*

证明. 存在性用强归纳。若 n 是素数，已经得到分解；否则 $n=ab$，其中 $1<a,b<n$。由归纳假设分解 a,b，合并得到 n 的分解。

唯一性中，设 $p_1\cdots p_r=q_1\cdots q_s$，所有因子均为素数。Euclid 引理反复应用，说明 p_1 整除某个 q_j，故 $p_1=q_j$。将该因子移到首位并消去，得到更短的两个分解。继续消去，直到两边均无因子。若有一边先耗尽，另一边剩下的素数乘积至少为 2，不能等于空积 1。因此因子数相同，且可逐一配对。 □

存在性使用整数大小的下降，唯一性使用素数在乘积中的整除性质。这是证明中的两个不同任务。

定义 5.16 (最小公倍数)**.** 对非零整数 a,b，其最小正公倍数记为 $\operatorname{lcm}(a,b)$。

写 $|a|=\prod_p p^{\alpha_p}, |b|=\prod_p p^{\beta_p}$，其中 p 遍历素数，指数为非负整数且仅有限个非零。公因数在每个素数处的指数至多为 $\min(\alpha_p,\beta_p)$；公倍数的指数至少为 $\max(\alpha_p,\beta_p)$。于是

$$\gcd(a,b)=\prod_p p^{\min(\alpha_p,\beta_p)},\qquad \operatorname{lcm}(a,b)=\prod_p p^{\max(\alpha_p,\beta_p)}.$$

由于最小值与最大值之和为 $\alpha_p+\beta_p$，得到 $\gcd(a,b)\operatorname{lcm}(a,b)=|ab|$。

5.6 同余：只保留除法的余数

固定正整数 n。若只关心除以 n 的余数，相差 n 的整数倍的两个整数就应当视为同类。

定义 5.17 (同余与剩余类)**.** 定义 $a\equiv b \pmod n$ 当且仅当 $n \mid a-b$。它为等价关系：差为 0 给出自反性，取负给出对称性，两个差相加给出传递性。商集记为 $\mathbb{Z}/n\mathbb{Z}$，a 的类记为 $[a]_n$。

带余除法给出每个类在 $\{0,1,\dots,n-1\}$ 中的一个代表元。若范围内两个代表元同余，其差是绝对值小于 n 的 n 的倍数，只能为 0。所以每个类在此范围内恰有一个代表元，商集恰有 n 个元素。

命题 5.18 (运算与代表元)**.** 规则 $[a]_n+[b]_n=[a+b]_n$，$[a]_n[b]_n=[ab]_n$ 在商集上良定义。

证明. 设 $a-a'=nr, b-b'=ns$。则

$$(a+b)-(a'+b')=n(r+s),\qquad ab-a'b'=a(b-b')+b'(a-a')=n(as+b'r).$$

两种结果都与更换代表元之前同余。因此输出只依赖输入的类。 □

计算时可以先在整数中运算，再改用规定范围内的代表元。例如 $[8]_{10}+[7]_{10}=[15]_{10}=[5]_{10}$。这里 $[15]_{10}$ 与 $[5]_{10}$ 本来就是同一个元素。

命题 5.19 (哪些类可以相除)**.** $[a]_n$ 有乘法逆元，当且仅当 $\gcd(a,n)=1$。

证明. 若 $[a]_n[b]_n=[1]_n$，则存在整数 k 使 $ab-kn=1$。a,n 的每个公因数因而都整除 1，所以它们互素。

若 $\gcd(a,n)=1$，由 Bézout 等式取 u,v 使 $ua+vn=1$，转为剩余类等式即得 $[u]_n[a]_n=[1]_n$。 □

所以求模逆元又回到了同一个整数线性方程。$7\cdot 3=21\equiv 1 \pmod{10}$，故 $[7]_{10}^{-1}=[3]_{10}$。而 $[2]_{10}$ 无逆元，且 $[2]_{10}[5]_{10}=[0]_{10}$，两因子却均非零类。在模运算中使用消去或除法，必须先检验相应条件。

文献：整数算法、Bézout 等式与唯一分解见 [8, 9]。

6 群论

群的运算可以表现为整数相加，也可以表现为可逆变换的复合。本章从二元运算开始，先证明群公理所允许的运算规则，再构造子群、陪集、商群和同态。随后用群作用建立有限群的计数方法，最后通过字与关系构造自由群及自由积。群中乘法一般不交换；每次改变因子次序都须有相应依据。

6.1 二元运算与群公理

设 G 为非空集合。二元运算是一个映射 $\mu : G \times G \to G$。记 $\mu(a,b) = ab$；由于陪域为 G，任意 $a, b \in G$ 的乘积已经要求属于 G。同一集合可以配上不同运算，因此群的对象是集合与运算组成的结构。

定义 6.1 (群)**.** 带二元运算的非空集合 G 称为群，若存在元素 $e \in G$，使下列条件成立：

$$(ab)c = a(bc) \qquad (a,b,c \in G), \tag{6.1}$$

$$ea = a = ae \qquad (a \in G), \tag{6.2}$$

$$\text{对每个} a \in G\text{，存在} b \in G \text{ 使} ab = e = ba. \tag{6.3}$$

e 称为单位元，式 (6.3) 中的 b 称为 a 的逆元。若还满足 $ab = ba$ 对所有 $a, b \in G$ 成立，则称 G 为交换群或阿贝尔群。

以下先从三条公理推出基本计算规则。特别是逆元的记号只有在证明唯一性之后才固定。

命题 6.2 (单位元与逆元的唯一性)**.** *群中单位元唯一；每个元素的逆元唯一。*

证明. 若 e, e' 均为单位元，则

$$e = ee' \quad (e' \text{ 为右单位元}), \qquad ee' = e' \quad (e \text{ 为左单位元}),$$

故 $e = e'$。固定 $a \in G$，若 u, v 均为其逆元，则

$$\begin{aligned} u &= ue && \text{单位元性质} \\ &= u(av) && av = e \\ &= (ua)v && \text{结合律} \\ &= ev && ua = e \\ &= v && \text{单位元性质.} \end{aligned}$$

于是逆元唯一，记作 a^{-1}。 □

命题 6.3 (消去、逆元与方程)**.** *对 $a, b, c \in G$，有*

$$ab = ac \iff b = c, \qquad ba = ca \iff b = c, \qquad (ab)^{-1} = b^{-1}a^{-1}, \quad (a^{-1})^{-1} = a.$$

方程 $ax = b$ 与 $ya = b$ 分别有唯一解 $a^{-1}b$ 与 ba^{-1}。

证明. 若 $ab = ac$，则

$$b = eb = (a^{-1}a)b = a^{-1}(ab) = a^{-1}(ac) = (a^{-1}a)c = ec = c.$$

若 $ba = ca$，则

$$b = be = b(aa^{-1}) = (ba)a^{-1} = (ca)a^{-1} = c(aa^{-1}) = ce = c.$$

两个逆向蕴含只需在 $b = c$ 两边同乘 a。对乘积逆元，分别计算

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = e,$$
$$(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}eb = e.$$

由逆元唯一性得公式。又 $a^{-1}a = aa^{-1} = e$，所以 a 是 a^{-1} 的逆元。

最后

$$ax = b \implies x = a^{-1}(ax) = a^{-1}b, \qquad a(a^{-1}b) = (aa^{-1})b = b.$$

前式证明任何解都被迫等于 $a^{-1}b$，后式证明它确实是解。右侧方程同样由

$$ya = b \implies y = (ya)a^{-1} = ba^{-1}, \qquad (ba^{-1})a = b$$

得到存在性与唯一性。 □

命题 6.4 (有限乘积与括号). 固定元素序列 $a_1, \dots, a_n$，保持次序的所有合法加括号方式给出相同乘积。

证明. 对 n 作强归纳。$n \le 2$ 时无须比较。任意一个 n 因子括号式，其最后一步都在某个 i 处分开，形如

$$(a_1 \cdots a_i)(a_{i+1} \cdots a_n), \qquad 1 \le i < n.$$

每个括号内的值由归纳假设唯一确定。比较两个分点 $i < j$，记三个非空连续段的乘积分别为 A, B, C。归纳假设把两种值分别写成 $A(BC)$ 与 $(AB)C$，而 (6.1) 给出 $A(BC) = (AB)C$。相同分点的情形直接由归纳假设相等。因此任意两种括号式相等。 □

6.2 整数次幂与具体群

定义 6.5 (整数次幂与阶). 固定 $g \in G$。递归定义 $g^0 = e$、$g^{n+1} = g^n g$ $(n \ge 0)$，并对 $n > 0$ 定义 $g^{-n} = (g^{-1})^n$。若存在正整数 n 使 $g^n = e$，称其中最小者为 g 的阶，记 $\operatorname{ord}(g)$；否则称 g 为无限阶。群的阶 $|G|$ 指其集合的基数。

命题 6.6 (幂的运算法则). 对任意 $m, n \in \mathbb{Z}$，有

$$g^{m+n} = g^m g^n, \qquad (g^m)^n = g^{mn}, \qquad (g^m)^{-1} = g^{-m}.$$

证明. 先将递推公式扩展到负指标。若 $k = -r < 0$，$r \ge 1$，则

$$g^k g = (g^{-1})^r g = (g^{-1})^{r-1}(g^{-1}g) = (g^{-1})^{r-1} = g^{k+1}.$$

因此 $g^{k+1} = g^k g$ 对每个整数 k 成立，并且右乘 g^{-1} 得 $g^{k-1} = g^k g^{-1}$。

固定任意整数 m。当 $n=0$ 时，$g^{m+0}=g^m=g^mg^0$。若等式在 n 成立，则

$$g^{m+n+1}=g^{m+n}g=(g^mg^n)g=g^mg^{n+1},$$
$$g^{m+n-1}=g^{m+n}g^{-1}=(g^mg^n)g^{-1}=g^mg^{n-1}.$$

由从零分别向正、负方向归纳，第一式对所有 n 成立。取 $n=-m$，得到

$$g^mg^{-m}=g^0=e=g^{-m}g^m,$$

由逆元唯一性得第三式。对 $n\geq 0$，第二式按 n 归纳：

$$(g^m)^0=e=g^{m\cdot 0},\qquad (g^m)^{n+1}=(g^m)^ng^m=g^{mn}g^m=g^{m(n+1)}.$$

若 $n=-r<0$，则

$$(g^m)^n=((g^m)^{-1})^r=(g^{-m})^r=g^{-mr}=g^{mn}.$$

□

例 6.7 (整数加法与剩余类)**.** 在 $\mathbb{Z}$ 中取加法为运算，则结合律为 $(a+b)+c=a+(b+c)$，单位元为 0，逆元为 $-a$。交换律给出这是交换群。对 $n\geq 1$，加法

$$[a]_n+[b]_n=[a+b]_n$$

也使整数剩余类组成群。若 $a'=a+rn, b'=b+sn$，则

$$(a'+b')-(a+b)=(r+s)n,$$

所以加法与代表元无关。其单位为 $[0]_n$，逆元为 $[-a]_n$；结合律与交换律由整数加法下降。

例 6.8 (置换群)**.** 对集合 X，令 $\mathrm{Sym}(X)$ 为全体双射 $X\to X$，以映射复合为乘法，约定右边的映射先作用。双射 σ,τ 的复合仍为双射，因为

$$(\sigma\tau)(\tau^{-1}\sigma^{-1})=\mathrm{id}_X=(\tau^{-1}\sigma^{-1})(\sigma\tau).$$

对任意 $x\in X$，

$$((\sigma\tau)\rho)(x)=\sigma(\tau(\rho(x)))=(\sigma(\tau\rho))(x),$$

故复合结合，id_X 为单位，逆映射为逆元，因而构成群。$X=\{1,\dots,n\}$ 时记作 S_n。依次选择 $\sigma(1),\dots,\sigma(n)$，可选数为 $n,n-1,\dots,1$，故 $|S_n|=n!$。

循环记号 $(i_1\ i_2\ \cdots\ i_r)$ 表示把 i_j 送到 i_{j+1}、把 i_r 送到 i_1，并固定其余元素的置换。两字母循环称为对换。在 S_3 中，若 $s=(12), t=(23)$，则

$$(st)(1)=s(1)=2,\qquad (ts)(1)=t(2)=3,$$

故 $st\neq ts$。逐点计算还有 $st=(123)$、$ts=(132)$。

6.3 子群、交与生成

定义 6.9 (子群)**.** 子集 $H \subseteq G$ 若在 G 的运算限制下本身为群，称为子群，记作 $H \leq G$。

命题 6.10 (子群判别)**.** *非空子集 $H \subseteq G$ 为子群，当且仅当*

$$a, b \in H \implies ab^{-1} \in H.$$

证明. 先设 H 在限制运算下为群，其单位元暂记 e_H。因为 $e_H e_H = e_H$，在 G 中左乘 e_H^{-1} 得 $e_H = e_G$。每个 $b \in H$ 在 H 中的逆元也满足 G 的逆元等式，由唯一性等于 b^{-1}。因此 $a, b \in H$ 蕴含 $ab^{-1} \in H$。

反过来，设条件成立。取 $a \in H$，依次使用条件得到

$$\begin{aligned} a, a \in H &\implies aa^{-1} = e \in H, \\ e, b \in H &\implies eb^{-1} = b^{-1} \in H, \\ a, b^{-1} \in H &\implies a(b^{-1})^{-1} = ab \in H. \end{aligned}$$

于是 H 中有单位元、每个元素的逆元，且乘法封闭。对 $a, b, c \in H$，结合律 $(ab)c = a(bc)$ 是 G 中同一等式的限制，故 H 为群。 □

命题 6.11 (交与生成子群)**.** *任意子群族 $(H_i)_{i \in I}$ 的交是子群，空族之交取为 G。对 $S \subseteq G$，定义*

$$\langle S \rangle = \bigcap_{S \subseteq H \leq G} H.$$

则

$$\langle S \rangle = \{ s_1^{\varepsilon_1} \cdots s_r^{\varepsilon_r} : r \geq 0,\ s_i \in S,\ \varepsilon_i \in \{1, -1\} \}.$$

$r = 0$ 时取空积 e。

证明. 每个 H_i 含 e，所以交非空；若 $a, b \in \bigcap_i H_i$，则

$$\forall i \in I, \quad ab^{-1} \in H_i, \quad \text{故} ab^{-1} \in \bigcap_i H_i.$$

由子群判别，交为子群。包含 S 的子群至少有 G，所以所定义的交存在且为子群。

记右侧的有限乘积集合为 W。它含空积；两个词连接仍为词，而且

$$(s_1^{\varepsilon_1} \cdots s_r^{\varepsilon_r})^{-1} = s_r^{-\varepsilon_r} \cdots s_1^{-\varepsilon_1} \in W.$$

因此 W 为含 S 的子群，故 $\langle S \rangle \subseteq W$。另一方面，对每个含 S 的子群 H，由逆元和乘法封闭有 $W \subseteq H$，从而 $W \subseteq \bigcap_{S \subseteq H \leq G} H = \langle S \rangle$。两包含给出相等。 □

例 6.12 (整数的子群)**.** $\mathbb{Z}$ 的每个加法子群都形如 $d\mathbb{Z}$，其中 $d \geq 0$。若子群 $H = \{0\}$，取 $d = 0$。否则 H 含正整数，令 d 为其中最小者。对 $a \in H$ 作带余除法 $a = qd + r$，$0 \leq r < d$，则

$$r = a - qd \in H.$$

若 $r > 0$，与 d 的最小性矛盾，故 $r = 0$，即 $H \subseteq d\mathbb{Z}$。由 $d \in H$ 及整数倍封闭，$d\mathbb{Z} \subseteq H$，所以 $H = d\mathbb{Z}$。正生成元也唯一：$d\mathbb{Z} = d'\mathbb{Z}$ 且 $d, d' > 0$ 时，$d \mid d'$ 且 $d' \mid d$，故 $d = d'$。

6.4 同态、核与纤维

定义 6.13 (群同态)**.** 设 G, K 为群。映射 $f : G \to K$ 若满足 $f(ab) = f(a)f(b)$，称为群同态。双射同态称为同构，记为 $G \cong K$。

命题 6.14 (同态的基本性质)**.** 同态 $f : G \to K$ 满足

$$f(e_G) = e_K, \qquad f(a^{-1}) = f(a)^{-1}, \qquad f(a^n) = f(a)^n \quad (n \in \mathbb{Z}).$$

双射同态的逆映射仍为同态；恒等映射和同态的复合都是同态。

证明. 设 $u = f(e_G)$。由 $u = f(e_G e_G) = u^2$，得

$$e_K = u^{-1}u = u^{-1}u^2 = u.$$

又

$$f(a)f(a^{-1}) = f(aa^{-1}) = e_K, \qquad f(a^{-1})f(a) = f(a^{-1}a) = e_K,$$

所以由逆元唯一性得第二式。$n \geq 0$ 时第三式由递归定义归纳得到：

$$f(a^0) = e_K = f(a)^0, \qquad f(a^{n+1}) = f(a^n a) = f(a)^n f(a) = f(a)^{n+1}.$$

负指数由第二式及正指数情形得到。

若 f 双射，对 $x, y \in K$，令 $a = f^{-1}(x), b = f^{-1}(y)$，则

$$f^{-1}(xy) = f^{-1}(f(a)f(b)) = f^{-1}(f(ab)) = ab = f^{-1}(x)f^{-1}(y).$$

若 $h : K \to L$ 为同态，则 $(hf)(ab) = h(f(a)f(b)) = (hf)(a)(hf)(b)$。恒等映射的等式为 $\mathrm{id}_G(ab) = \mathrm{id}_G(a)\mathrm{id}_G(b)$。 □

定义 6.15 (核与像)**.** 对同态 $f : G \to K$，定义

$$\ker f = \{a \in G : f(a) = e_K\}, \qquad \operatorname{im} f = \{f(a) : a \in G\}.$$

命题 6.16 (核、像与单射判别)**.** $\ker f \leq G$，$\operatorname{im} f \leq K$，且

$$f(a) = f(b) \iff b^{-1}a \in \ker f.$$

因此 f 为单射当且仅当 $\ker f = \{e_G\}$。

证明. 由 $f(e_G) = e_K$，核非空；若 $a, b \in \ker f$，则

$$f(ab^{-1}) = f(a)f(b)^{-1} = e_K e_K^{-1} = e_K,$$

故核为子群。像含 e_K，且

$$f(a)f(b)^{-1} = f(ab^{-1}) \in \operatorname{im} f,$$

故像为子群。对于任意 $a, b \in G$，

$$\begin{aligned} f(a) = f(b) &\iff f(b)^{-1}f(a) = e_K \\ &\iff f(b^{-1}a) = e_K \\ &\iff b^{-1}a \in \ker f. \end{aligned}$$

若 f 单射，则 $a \in \ker f$ 蕴含 $f(a) = f(e_G)$，故 $a = e_G$。若核平凡，则 $f(a) = f(b)$ 蕴含 $b^{-1}a = e_G$，即 $a = b$。 □

6.5 循环群的分类与子群

定义 6.17 (循环群)**.** 若 $G=\langle g\rangle$，称 G 为循环群，g 为生成元。由幂的加法公式，$\langle g\rangle=\{g^k: k\in\mathbb{Z}\}$。

定理 6.18 (循环群的分类)**.** *若 g 无限阶，则 $\langle g\rangle\cong\mathbb{Z}$。若 g 的阶为 n，则*

$$g^k=e \iff n\mid k, \qquad \langle g\rangle\cong\mathbb{Z}/n\mathbb{Z}.$$

证明. 映射 $\phi:\mathbb{Z}\to\langle g\rangle$，$\phi(k)=g^k$，由 $g^{k+\ell}=g^kg^\ell$ 为满同态。若 g 无限阶且 $g^k=e$，则 $k>0$ 与无限阶矛盾；$k<0$ 时取逆得到 $g^{-k}=e$ 也矛盾，故 $k=0$。于是核为零，ϕ 为同构。

若 $\operatorname{ord}(g)=n$，对 $k\in\mathbb{Z}$ 写 $k=qn+r$，$0\le r<n$。则

$$g^k=(g^n)^qg^r=g^r.$$

若 $g^k=e$，就有 $g^r=e$，由 n 的最小性得 $r=0$；反之 $n\mid k$ 时上式显然为 e。进而

$$g^k=g^\ell \iff g^{k-\ell}=e \iff n\mid(k-\ell) \iff [k]_n=[\ell]_n.$$

所以 $[k]_n\mapsto g^k$ 良定义且单射，满射来自生成性，保持加法到乘法来自幂公式，故为同构。 □

定理 6.19 (循环群的子群与元素的阶)**.** *循环群的每个子群仍为循环群。若 $\operatorname{ord}(g)=n<\infty$，则*

$$\operatorname{ord}(g^k)=\frac{n}{\gcd(n,k)} \quad (k\in\mathbb{Z}),$$

其中 $\gcd(n,k)$ 取正值。阶为 n 的循环群对每个正因子 $d\mid n$ 恰有一个阶为 d 的子群，为 $\langle g^{n/d}\rangle$。

证明. 设 $H\le\langle g\rangle$。取前面的满同态 $\phi:\mathbb{Z}\to\langle g\rangle$。集合

$$A=\phi^{-1}(H)=\{k: g^k\in H\}$$

含零；若 $k,\ell\in A$，则 $g^{k-\ell}=g^k(g^\ell)^{-1}\in H$，故 $k-\ell\in A$。于是 $A=d\mathbb{Z}$，某个 $d\ge0$。由于 ϕ 满射，$H=\phi(A)=\langle g^d\rangle$。

令 $c=\gcd(n,k)$，写 $n=cn_0, k=ck_0$，则 $\gcd(n_0,k_0)=1$。对正整数 t，

$$(g^k)^t=e \iff n\mid kt \iff n_0\mid k_0t \iff n_0\mid t.$$

最后一步使用互素数的整除性质。因此最小的这样的 t 为 $n_0=n/c$。

若 H 为有限循环群的子群，其逆像 $A=d\mathbb{Z}$ 包含 $n\mathbb{Z}$，故 $d\mid n$；这里可取 $1\le d\le n$，包括平凡子群对应的 $d=n$。上式给出 $|H|=n/d$。所以若指定 $|H|=r$，则必须 $d=n/r$，存在性与唯一性同时得到。 □

6.6 陪集、等价关系与 Lagrange 定理

定义 6.20 (陪集与指数)**.** 设 $H\le G$。对 $g\in G$，左陪集、右陪集分别为

$$gH=\{gh: h\in H\}, \qquad Hg=\{hg: h\in H\}.$$

左陪集组成的集合记为 G/H，此处只赋予它集合结构。其基数称为指数，记 $[G:H]$。

命题 6.21 (陪集判别与划分)**.** 对 $g,k\in G$，有

$$gH=kH \iff k^{-1}g\in H.$$

任意两个左陪集或相等或不交；所有左陪集覆盖 G。每个 gH 均与 H 等势。

证明. 若 $gH=kH$，则 $g=ge\in kH$，所以存在 $h_0\in H$ 使 $g=kh_0$，从而 $k^{-1}g=h_0\in H$。反过来，若 $g=kh_0$，则

$$gH=kh_0H=kH.$$

其中 $h_0H=H$ 的两个包含分别由 $h_0h\in H$ 及 $h=h_0(h_0^{-1}h)$ 得到。

若 $gH\cap kH\neq\varnothing$，取 $gh=kh'$，$h,h'\in H$。则

$$k^{-1}g=h'h^{-1}\in H,$$

故 $gH=kH$。又 $g=ge\in gH$ 对每个 g 成立，所以这些集合覆盖 G。

定义 $L_g:H\to gH$，$L_g(h)=gh$，以及 $R_g:gH\to H$，$R_g(x)=g^{-1}x$。后者确实取值于 H，因为 $x=gh$ 时 $g^{-1}x=h$。并且

$$R_gL_g(h)=g^{-1}gh=h,\qquad L_gR_g(x)=gg^{-1}x=x.$$

所以它们互逆，L_g 为双射。 □

定义 $a\sim b\iff b^{-1}a\in H$，则它的等价类正是左陪集：

$$[a]=\{b:a^{-1}b\in H\}=aH.$$

关系的自反、对称、传递分别由 $a^{-1}a=e\in H$、$(b^{-1}a)^{-1}=a^{-1}b\in H$、$c^{-1}a=(c^{-1}b)(b^{-1}a)\in H$ 给出。这把陪集划分与前面等价关系的商集构造连接起来。

定理 6.22 (Lagrange 定理)**.** 若 G 有限且 $H\le G$，则

$$|G|=[G:H]|H|.$$

因此 $|H|\mid|G|$，且对每个 $g\in G$ 有 $\operatorname{ord}(g)\mid|G|$。

证明. 令 $g_1H,\dots,g_rH$ 为互不相同的全部左陪集，则 $r=[G:H]$。划分性质说明这些子集两两不交，且

$$G=\bigcup_{i=1}^{r}g_iH.$$

每块通过 $h\mapsto g_ih$ 与 H 双射，故

$$|G|=\sum_{i=1}^{r}|g_iH|=\sum_{i=1}^{r}|H|=r|H|.$$

取 $H=\langle g\rangle$，循环群分类给出 $|H|=\operatorname{ord}(g)$，得到最后结论。 □

推论 6.23. 阶为素数 p 的群为循环群。若子群 H 的指数为 2，则 $gH=Hg$ 对所有 $g\in G$ 成立。

证明. 若 $|G|=p$ 且 $g\neq e$，则 $1<\operatorname{ord}(g)\mid p$，所以 $\operatorname{ord}(g)=p$，从而 $\langle g\rangle=G$。

对于指数为 2 的子群，若 $g\in H$，则 $gH=H=Hg$。若 $g\notin H$，左陪集恰为 H,gH，故 $gH=G\setminus H$。右陪集与左陪集等势对应，具体为 $aH\mapsto Ha^{-1}$；逆向同样取逆，所以右陪集也恰有两个。因此 $Hg=G\setminus H=gH$。 □

6.7 正规子群与商群的乘法

对子群 $N \le G$，我们希望在陪集集合上定义 $(gN)(hN) = ghN$。若将 g, h 换成 gn, hm，其中 $n, m \in N$，则

$$(gn)(hm) = gnhm = gh(h^{-1}nh)m.$$

因此需要 $h^{-1}nh \in N$，才能保证两个乘积属于同一陪集。正规性的条件正是在这里出现的。

定义 6.24 (正规子群)**.** 子群 $N \le G$ 称为正规子群，记为 $N \trianglelefteq G$，若

$$gNg^{-1} = N \qquad (g \in G).$$

其中 $gNg^{-1} = \{gng^{-1} : n \in N\}$。

命题 6.25 (正规性的等价条件)**.** *对于子群 $N \le G$，以下三个条件等价：对所有 $g \in G$，有 $gNg^{-1} = N$；对所有 $g \in G$，有 $gNg^{-1} \subseteq N$；对所有 $g \in G$，有 $gN = Ng$。任意群同态的核正规。*

证明. 相等显然蕴含包含。若包含对每个 g 成立，将 g 换为 g^{-1} 得 $g^{-1}Ng \subseteq N$，所以

$$N = g(g^{-1}Ng)g^{-1} \subseteq gNg^{-1} \subseteq N.$$

于是相等。左右乘集合是双射给出的集合操作，故

$$gNg^{-1} = N \iff (gNg^{-1})g = Ng \iff gN = Ng.$$

若 $N = \ker f$，则对 $n \in N, g \in G$，

$$f(gng^{-1}) = f(g)f(n)f(g)^{-1} = f(g)ef(g)^{-1} = e,$$

所以 $gNg^{-1} \subseteq N$。由刚才证明的等价性，核正规。 □

定理 6.26 (商群的构造)**.** *规则 $(gN)(hN) = ghN$ 与代表元无关，当且仅当 $N \trianglelefteq G$。在此条件下，它使 G/N 成为群，其单位元为 N，逆元为 $(gN)^{-1} = g^{-1}N$。投影*

$$q : G \to G/N, \qquad q(g) = gN$$

是满同态，且 $\ker q = N$。

证明. 设 N 正规。若 $g'N = gN$ 且 $h'N = hN$，由陪集判别存在 $n, m \in N$ 使 $g' = gn, h' = hm$。令 $u = (h^{-1}nh)m$，正规性与子群乘法封闭给出 $u \in N$。于是

$$g'h'N = gnhmN = ghuN = ghN.$$

所以乘法良定义。对任意 $g, h, k \in G$，

$$\begin{aligned}((gN)(hN))(kN) = ((gh)k)N = (g(hk))N = (gN)((hN)(kN)),\\ N(gN) = (eg)N = gN = (ge)N = (gN)N,\\ (gN)(g^{-1}N) = (gg^{-1})N = N = (g^{-1}g)N = (g^{-1}N)(gN).\end{aligned}$$

这逐项验证了群公理。投影满足

$$q(gh) = ghN = (gN)(hN) = q(g)q(h).$$

每个陪集 gN 都等于 $q(g)$，故 q 满射，而且

$$g \in \ker q \iff gN = N \iff g \in N.$$

反过来，若规则良定义，则 $nN = N = eN$ 对 $n \in N$ 成立，因而

$$ngN = (nN)(gN) = (eN)(gN) = gN.$$

陪集判别给出 $g^{-1}ng \in N$。由于 g 任意，$g^{-1}Ng \subseteq N$ 对所有 g 成立，所以上一命题给出 N 正规。 □

例 6.27 (不正规时的计算)**.** 取 $G = S_3$、$H = \{e, (12)\}$。记 $s = (12), t = (23)$，逐点复合得

$$t^{-1}st = (13) \notin H.$$

由于 $sH = H$，若陪集乘法良定义，就应有 $(sH)(tH) = H(tH) = tH$；但按照候选规则左边为 stH，而

$$stH = tH \iff t^{-1}st \in H$$

不成立。因此同一输入陪集 H 的两个代表元 e, s 给出了不同的输出。

6.8 商群的因子化与第一同构定理

定理 6.28 (因子化定理)**.** 设 $N \trianglelefteq G$、$f : G \to K$ 为同态，$q : G \to G/N$ 为投影。存在同态 $\bar{f} : G/N \to K$ 使 $f = \bar{f}q$，当且仅当 $N \subseteq \ker f$。存在时它唯一，并且

$$\bar{f}(gN) = f(g).$$

证明. 若 $f = \bar{f}q$，则对 $n \in N$，

$$f(n) = \bar{f}(q(n)) = \bar{f}(N) = e_K,$$

所以 $N \subseteq \ker f$。

反过来，假设此包含成立。若 $gN = g'N$，可写 $g' = gn$，其中 $n \in N$，故

$$f(g') = f(gn) = f(g)f(n) = f(g)e_K = f(g).$$

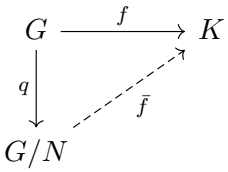
因此公式 $\bar{f}(gN) = f(g)$ 良定义。并且

$$\bar{f}((gN)(hN)) = \bar{f}(ghN) = f(gh) = f(g)f(h) = \bar{f}(gN)\bar{f}(hN),$$

故为同态，且 $\bar{f}q(g) = \bar{f}(gN) = f(g)$。若 $u : G/N \to K$ 也满足 $uq = f$，则对每个陪集 gN，

$$u(gN) = u(q(g)) = f(g) = \bar{f}(gN).$$

所以 $u = \bar{f}$，得到唯一性。 □



图中 $\bar f$ 在商类上的值不是任意指定的；等式 $\bar f q = f$ 强迫它取 $f(g)$。条件 $N \subseteq \ker f$ 则保证这个被迫给出的公式与代表元无关。

定理 6.29 (第一同构定理)**.** *每个同态 $f : G \to K$ 给出同构*

$$\Phi : G/\ker f \longrightarrow \operatorname{im} f, \qquad \Phi(g \ker f) = f(g).$$

证明. 核正规，由因子化定理，Φ 是良定义同态。对每个 $y \in \operatorname{im} f$，存在 $g \in G$ 使 $y = f(g) = \Phi(g \ker f)$，故 Φ 满射。对任意 $g, h \in G$，

$$\begin{aligned} \Phi(g \ker f) = \Phi(h \ker f) &\iff f(g) = f(h) \\ &\iff h^{-1}g \in \ker f \\ &\iff g \ker f = h \ker f. \end{aligned}$$

因此 Φ 单射，故为同构。 □

例 6.30 (同态的纤维是核的陪集)**.** 取 $f : \mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$，$f(k) = [k]_6$。核为 $6\mathbb{Z}$，每个纤维为 $a + 6\mathbb{Z}$。例如

$$f^{-1}(\{[2]_6\}) = \{\ldots, -10, -4, 2, 8, 14, \ldots\} = 2 + 6\mathbb{Z}.$$

一般地，若 $f : G \to K$ 为同态、$y = f(g)$，则

$$f^{-1}(\{y\}) = \{h \in G : f(h) = f(g)\} = \{h : g^{-1}h \in \ker f\} = g \ker f.$$

所以第一同构定理将同态的每个纤维替换成一个商群元素，随后保留它实际取到的输出。

6.9 子群乘积、第二与第三同构定理

命题 6.31 (子群乘积)**.** *设 $H, N \le G$ 且 $N \trianglelefteq G$。定义 $HN = \{hn : h \in H, n \in N\}$，则 $HN \le G$，且 $N \trianglelefteq HN$。若 G 有限，则*

$$|HN| = \frac{|H||N|}{|H \cap N|}.$$

证明. $e = ee \in HN$。对 $h_1n_1, h_2n_2 \in HN$，

$$(h_1n_1)(h_2n_2) = h_1h_2(h_2^{-1}n_1h_2)n_2 \in HN,$$

因为 $h_1h_2 \in H$，$(h_2^{-1}n_1h_2)n_2 \in N$。又

$$(hn)^{-1} = n^{-1}h^{-1} = h^{-1}(hn^{-1}h^{-1}) \in HN.$$

因此 HN 为子群。对 $x \in HN \subseteq G$ 已有 $xNx^{-1} = N$，故 $N \trianglelefteq HN$。

计数时考虑满射 $\mu : H \times N \to HN$，$\mu(h,n) = hn$。固定 $h_0 n_0 \in HN$，则

$$hn = h_0 n_0 \iff h_0^{-1} h = n_0 n^{-1} \in H \cap N.$$

令此公共元素为 c，得到 $h = h_0 c, n = c^{-1} n_0$。反过来每个 $c \in H \cap N$ 都给出一个这样的表示。因此纤维与 $H \cap N$ 双射，所有纤维均有 $|H \cap N|$ 个元素，于是

$$|H||N| = \sum_{x \in HN} |\mu^{-1}(x)| = |HN||H \cap N|.$$

□

定理 6.32 (第二同构定理)**.** *若 $H \le G, N \trianglelefteq G$，则*

$$H \cap N \trianglelefteq H, \qquad H/(H \cap N) \cong HN/N.$$

证明. 定义 $f : H \to HN/N$，$f(h) = hN$。对 $h, h' \in H$，

$$f(hh') = hh'N = (hN)(h'N) = f(h)f(h').$$

每个陪集 $(hn)N = hN$，故 f 满射；并且

$$\ker f = \{h \in H : hN = N\} = \{h \in H : h \in N\} = H \cap N.$$

核正规，且第一同构定理给出 $h(H \cap N) \mapsto hN$ 的同构。 □

定理 6.33 (第三同构定理)**.** *设 $N, K \trianglelefteq G$ 且 $N \subseteq K$，则 $K/N \trianglelefteq G/N$，并且*

$$(G/N)/(K/N) \cong G/K.$$

证明. 定义 $f : G/N \to G/K$，$f(gN) = gK$。若 $gN = hN$，则 $h^{-1}g \in N \subseteq K$，所以 $gK = hK$；因此良定义。其乘法与满射性质分别由

$$f((gN)(hN)) = ghK = (gK)(hK), \qquad gK = f(gN)$$

得到。核满足

$$gN \in \ker f \iff gK = K \iff g \in K,$$

故 $\ker f = \{gN : g \in K\} = K/N$。由核正规性及第一同构定理得到结论。 □

定理 6.34 (子群对应定理)**.** *设 $q : G \to G/N$ 为正规子群的投影。映射*

$$H \longmapsto q(H), \qquad L \longmapsto q^{-1}(L)$$

给出包含 N 的子群 $H \le G$ 与子群 $L \le G/N$ 之间的互逆对应。此对应保持包含关系，并满足

$$H \trianglelefteq G \iff q(H) \trianglelefteq G/N.$$

证明. 同态像是子群，故 $q(H)$ 为子群。若 $a, b \in q^{-1}(L)$，则 $q(ab^{-1}) = q(a)q(b)^{-1} \in L$，故 $q^{-1}(L)$ 为子群；它含 N，因为 $q(N) = \{N\} \subseteq L$。

对于 $N \subseteq H$，逐个元素计算

$$\begin{aligned} g \in q^{-1}(q(H)) &\iff \exists h \in H, \quad gN = hN \\ &\iff \exists h \in H, \quad h^{-1}g \in N \\ &\iff g \in H. \end{aligned}$$

最后一个方向用 $g = hn \in H$；反方向取 $h = g, n = e$。因此 $q^{-1}(q(H)) = H$。对 $y \in G/N$，由 q 满射，

$$y \in q(q^{-1}(L)) \iff y \in L.$$

所以两操作互逆。包含关系由像与原像的定义保持。

若 H 正规，对 $g \in G, h \in H$，

$$q(g)q(h)q(g)^{-1} = q(ghg^{-1}) \in q(H),$$

所以 $q(H)$ 正规。若 $q(H)$ 正规，则对于 $g \in G, h \in H$，同一等式表明 $ghg^{-1} \in q^{-1}(q(H)) = H$，故 H 正规。 □

6.10 群作用：用群元素规定置换

定义 6.35 (左作用)**.** 群 G 在集合 X 上的左作用是映射 $G \times X \to X$，记作 $(g, x) \mapsto g \cdot x$，满足

$$e \cdot x = x, \qquad (gh) \cdot x = g \cdot (h \cdot x).$$

命题 6.36 (作用与置换同态)**.** 一个左作用等价于一个同态 $\rho : G \to \mathrm{Sym}(X)$，对应公式为 $\rho(g)(x) = g \cdot x$。

证明. 给定作用，先证明每个 $\rho(g)$ 是双射。对 $x \in X$，

$$\rho(g^{-1})\rho(g)(x) = g^{-1} \cdot (g \cdot x) = (g^{-1}g) \cdot x = x;$$

将 g 换成 g^{-1} 得另一复合也为恒等。因此 $\rho(g) \in \mathrm{Sym}(X)$。又

$$\rho(gh)(x) = (gh) \cdot x = g \cdot (h \cdot x) = (\rho(g)\rho(h))(x),$$

故 $\rho(gh) = \rho(g)\rho(h)$。反过来，若 ρ 为同态，定义 $g \cdot x = \rho(g)(x)$，则

$$e \cdot x = \rho(e)(x) = x, \qquad (gh) \cdot x = \rho(g)\rho(h)(x) = g \cdot (h \cdot x).$$

两种构造逐点互逆。 □

定义 6.37 (轨道、稳定子与不动点)**.** 对 $x \in X$，定义轨道 $Gx = \{g \cdot x : g \in G\}$，稳定子 $G_x = \{g \in G : g \cdot x = x\}$。全体被所有群元素固定的点记为

$$X^G = \{x \in X : \forall g \in G,\ g \cdot x = x\}.$$

若 $Gx = X$ 对某个 x 成立，称作用为传递作用。若作用同态 ρ 单射，称作用忠实。

命题 6.38 (轨道划分与稳定子)**.** 各轨道构成 X 的划分，且 $G_x \leq G$。对任意 $g \in G$，有

$$G_{g \cdot x} = gG_xg^{-1}.$$

证明. 关系 $x \sim y \iff y = g \cdot x$ 对某个 g 成立，是等价关系，因为

$$x = e \cdot x, \qquad y = g \cdot x \implies x = g^{-1} \cdot y,$$

且 $y = g \cdot x, z = h \cdot y$ 蕴含 $z = (hg) \cdot x$。等价类恰为轨道，故划分成立。

$e \in G_x$。若 $a, b \in G_x$，则 $b^{-1} \cdot x = b^{-1} \cdot (b \cdot x) = x$，从而

$$(ab^{-1}) \cdot x = a \cdot (b^{-1} \cdot x) = a \cdot x = x.$$

由子群判别，$G_x \leq G$。最后对 $h \in G$，

$$\begin{aligned} h \in G_{g \cdot x} &\iff h \cdot (g \cdot x) = g \cdot x \\ &\iff (g^{-1} h g) \cdot x = x \\ &\iff g^{-1} h g \in G_x \\ &\iff h \in g G_x g^{-1}. \end{aligned}$$

这逐元素证明所需集合相等。 □

定理 6.39 (轨道与陪集)**.** 映射

$$\Phi : G/G_x \to Gx, \qquad \Phi(gG_x) = g \cdot x$$

为双射。若 G 有限，则 $|Gx| = [G : G_x] = |G|/|G_x|$。

证明. 关键等价式为

$$\begin{aligned} g \cdot x = h \cdot x &\iff (h^{-1} g) \cdot x = x \\ &\iff h^{-1} g \in G_x \\ &\iff g G_x = h G_x. \end{aligned}$$

从右向左读，证明不同代表元给出同一点，所以 Φ 良定义；从左向右读，证明相同像必来自同一陪集，所以 Φ 单射。任意轨道元素均为 $g \cdot x = \Phi(gG_x)$，所以满射。有限时再用 Lagrange 定理得到计数式。 □

定理 6.40 (Cayley 定理)**.** 每个群同构于其元素集合的置换群的一个子群。

证明. 定义左平移 $L_g : G \to G$，$L_g(a) = ga$。则

$$L_{g^{-1}} L_g(a) = g^{-1} g a = a, \qquad L_g L_{g^{-1}}(a) = g g^{-1} a = a,$$

所以 L_g 为双射。又

$$L_{gh}(a) = (gh)a = g(ha) = (L_g L_h)(a),$$

故 $L : g \mapsto L_g$ 是同态 $G \to \mathrm{Sym}(G)$。若 $L_g = \mathrm{id}_G$，在 e 处取值给出 $g = L_g(e) = e$，所以 $\ker L = \{e\}$。于是 L 单射，并给出 $G \cong \operatorname{im} L \leq \mathrm{Sym}(G)$。 □

6.11 共轭、中心、中心化子与正规化子

定义 $g\cdot x=gxg^{-1}$，其中 $g,x\in G$。这是共轭作用，因为

$$exe^{-1}=x,\qquad (gh)x(gh)^{-1}=g(hxh^{-1})g^{-1}.$$

它的轨道称为共轭类，记 $\mathrm{Cl}_G(x)=\{gxg^{-1}:g\in G\}$。由轨道关系，共轭是等价关系。

定义 6.41 (中心与中心化子)**.** 群 G 的中心与元素 x 的中心化子分别为

$$Z(G)=\{z\in G:\forall g\in G,\ zg=gz\},\qquad C_G(x)=\{g\in G:gx=xg\}.$$

对子群 $H\le G$，其正规化子为 $N_G(H)=\{g\in G:gHg^{-1}=H\}$。

命题 6.42. *$C_G(x)$ 为共轭作用中 x 的稳定子；$Z(G)=\bigcap_{x\in G}C_G(x)$ 为正规子群；$N_G(H)$ 为子群且 $H\trianglelefteq N_G(H)$。有限群中有*

$$|\mathrm{Cl}_G(x)|=[G:C_G(x)].$$

证明. 共轭作用的稳定条件满足

$$gxg^{-1}=x\iff gx=xg,$$

故稳定子恰为 $C_G(x)$，并且它是子群。交的元素正是与所有 x 交换的元素，因此 $Z(G)=\bigcap_x C_G(x)$ 为子群。若 $z\in Z(G)$，则 $gzg^{-1}=z\in Z(G)$，所以中心正规。

正规化子的单位条件为 $eHe^{-1}=H$。若 $a,b\in N_G(H)$，则 $bHb^{-1}=H$ 给出 $b^{-1}Hb=H$，从而

$$(ab^{-1})H(ab^{-1})^{-1}=a(b^{-1}Hb)a^{-1}=aHa^{-1}=H.$$

故 $N_G(H)$ 为子群。对 $h\in H$，由子群封闭性 $hHh^{-1}\subseteq H$，再将 h 换为 h^{-1} 得反向包含，故 $H\subseteq N_G(H)$；H 在其中正规则直接来自正规化子的定义。最后一个计数式为轨道公式。 □

正规化子也有作用解释：G 在全体子群组成的集合上以 $g\cdot H=gHg^{-1}$ 作用，结合律的计算与元素共轭相同。故 $N_G(H)$ 是这个作用的稳定子，而 H 的不同共轭子群的数目为 $[G:N_G(H)]$。

例 6.43 (S_3 中的共轭类)**.** 对置换 σ，逐点计算有 $\sigma(ij)\sigma^{-1}=(\sigma(i)\ \sigma(j))$。因此 (12) 的共轭类为 $\{(12),(13),(23)\}$。轨道大小为 3，所以 $|C_{S_3}((12))|=6/3=2$；其中已含 $e,(12)$，故中心化子正是这两元素组成的子群。

对三循环，$\sigma(123)\sigma^{-1}=(\sigma(1)\ \sigma(2)\ \sigma(3))$，所以 (123) 与 (132) 共轭。全部共轭类因此为

$$\{e\},\qquad \{(12),(13),(23)\},\qquad \{(123),(132)\}.$$

它们大小为 $1,3,2$，总和为 6，且只有 e 的共轭类为单元素，故 $Z(S_3)=\{e\}$。

6.12 自同构、内自同构与交换化

定义 6.44 (自同构)**.** 群到自身的同构称为自同构。全体自同构记为 $\mathrm{Aut}(G)$，运算为复合。映射 $c_g:G\to G$，$c_g(x)=gxg^{-1}$，称为由 g 给出的内自同构，全体内自同构记为 $\mathrm{Inn}(G)$。

命题 6.45. $\mathrm{Aut}(G)$ 为群，$\mathrm{Inn}(G) \trianglelefteq \mathrm{Aut}(G)$，且 $\mathrm{Inn}(G) \cong G/Z(G)$。因此可定义外自同构群 $\mathrm{Out}(G) = \mathrm{Aut}(G)/\mathrm{Inn}(G)$。

证明. 复合保持同态且保持双射，恒等映射为单位，逆映射仍为同态，结合律继承自映射复合，所以 $\mathrm{Aut}(G)$ 为群。对于 c_g，

$$c_g(xy) = gxyg^{-1} = (gxg^{-1})(gyg^{-1}) = c_g(x)c_g(y),$$

且 $c_{g^{-1}}c_g = \mathrm{id}_G = c_g c_{g^{-1}}$，故 $c_g \in \mathrm{Aut}(G)$。又

$$c_g c_h(x) = g(hxh^{-1})g^{-1} = (gh)x(gh)^{-1} = c_{gh}(x).$$

所以 $c : g \mapsto c_g$ 是同态，其像为 $\mathrm{Inn}(G)$，核满足

$$g \in \ker c \iff \forall x,\ gxg^{-1} = x \iff \forall x,\ gx = xg \iff g \in Z(G).$$

第一同构定理给出 $G/Z(G) \cong \mathrm{Inn}(G)$。

最后，对 $\alpha \in \mathrm{Aut}(G)$、$g, x \in G$，

$$\begin{aligned}(\alpha c_g \alpha^{-1})(x) &= \alpha(g\alpha^{-1}(x)g^{-1}) \\ &= \alpha(g)x\alpha(g)^{-1} \\ &= c_{\alpha(g)}(x).\end{aligned}$$

所以 $\alpha\,\mathrm{Inn}(G)\alpha^{-1} \subseteq \mathrm{Inn}(G)$；对 α^{-1} 使用相同包含得到相等，故正规，商群可定义。 □

定义 6.46 (交换子与交换子子群)**.** 定义 $[g,h] = ghg^{-1}h^{-1}$，称为 g,h 的交换子。全体交换子生成的子群记为 $[G,G]$。

命题 6.47 (交换化的泛性质)**.** $[G,G] \trianglelefteq G$，商群 $G^{\mathrm{ab}} = G/[G,G]$ 为交换群。每个到交换群 A 的同态 $f : G \to A$ 唯一经过投影 $G \to G^{\mathrm{ab}}$ 因子化。

证明. 对 $x, g, h \in G$，

$$x[g,h]x^{-1} = [xgx^{-1}, xhx^{-1}].$$

右侧由展开四个因子并消去相邻 $x^{-1}x$ 得到。于是共轭把每个交换子及其逆元送到交换子或其逆元，也把它们的有限乘积送到这样的乘积，故 $x[G,G]x^{-1} \subseteq [G,G]$；对所有 x 使用正规性判别得正规。

记投影为 q。由于 $q([g,h]) = e$，

$$q(g)q(h)q(g)^{-1}q(h)^{-1} = e \implies q(g)q(h) = q(h)q(g).$$

q 满射，所以商群所有元素都交换。若 A 交换，则

$$f([g,h]) = f(g)f(h)f(g)^{-1}f(h)^{-1} = e_A.$$

核为子群且含所有交换子，故 $[G,G] \subseteq \ker f$。商群因子化定理给出所需唯一同态。 □

6.13 类方程与素数幂阶群

定义 6.48 (p-群)**.** 设 p 为素数。有限群若阶为 p^a，其中 $a \geq 0$，称为有限 p-群。以下涉及 p-群的计数结论均在有限情形下使用。

定理 6.49 (类方程)**.** *设 G 为有限群，取全部非单元素共轭类的代表 $x_1, \ldots, x_r$，则*

$$|G| = |Z(G)| + \sum_{i=1}^{r} [G : C_G(x_i)].$$

证明. 元素 x 的共轭类为单元素，当且仅当

$$\forall g \in G,\ gxg^{-1} = x \iff \forall g \in G,\ gx = xg \iff x \in Z(G).$$

因此，把中心中的单元素共轭类合在一起，可写为

$$G = Z(G) \cup \bigcup_{i=1}^{r} \mathrm{Cl}_G(x_i).$$

右边各部分两两不交。分别计数，再代入 $|\mathrm{Cl}_G(x_i)| = [G : C_G(x_i)]$，便得等式。 □

命题 6.50 (p-群作用的不动点计数)**.** *若有限 p-群 P 作用在有限集合 X 上，则*

$$|X| \equiv |X^P| \pmod{p}.$$

特别地，非平凡有限 p-群的中心至少有 p 个元素。

证明. 对每个 $x \in X$，轨道大小为 $[P : P_x]$，由 Lagrange 定理，它是 p 的幂。单元素轨道恰好对应 X^P 中的点，其余轨道的大小均被 p 整除。若非单元素轨道为 $O_1, \ldots, O_s$，则

$$|X| = |X^P| + \sum_{i=1}^{s} |O_i| \equiv |X^P| \pmod{p}.$$

取 $X = P$、作用为共轭，则 $X^P = Z(P)$。若 $|P| = p^a$ 且 $a \geq 1$，上式给出 $|Z(P)| \equiv 0 \pmod{p}$。中心含单位元，所以非空；其大小是正的 p 的倍数，故至少为 p。 □

命题 6.51. *若 $G/Z(G)$ 为循环群，则 G 交换。因此阶为 p^2 的群必交换。*

证明. 设 $G/Z(G) = \langle gZ(G) \rangle$。对 $a, b \in G$，存在整数 m, n 及 $z, w \in Z(G)$ 使 $a = g^m z, b = g^n w$。于是

$$ab = g^m z g^n w = g^{m+n} zw = g^{n+m} wz = g^n w g^m z = ba.$$

所以 G 交换。

若 $|G| = p^2$，中心的阶整除 p^2 且至少为 p，故只能为 p 或 p^2。若为 p^2，则 $G = Z(G)$，已交换。若为 p，则 $|G/Z(G)| = p$，商群为循环群，由第一部分得 G 交换，进而又有 $Z(G) = G$。因此两种情况下均交换。 □

6.14 Cauchy 定理：对元素组作循环置换

定理 6.52 (Cauchy)**.** *若素数 p 整除有限群 G 的阶，则 G 中有阶为 p 的元素。*

证明. 考虑有限集合

$$X = \{(g_1, \dots, g_p) \in G^p : g_1 \cdots g_p = e\}.$$

前 $p-1$ 个元素可任取，末项被唯一确定为 $(g_1 \cdots g_{p-1})^{-1}$，所以 $|X| = |G|^{p-1}$。

定义循环移动 $\sigma(g_1, \dots, g_p) = (g_2, \dots, g_p, g_1)$。它将 X 映入自身，因为

$$g_2 \cdots g_p g_1 = g_1^{-1}(g_1 g_2 \cdots g_p) g_1 = e.$$

且 $\sigma^p = \mathrm{id}_X$，所以 $\mathbb{Z}/p\mathbb{Z}$ 通过 $[k] \cdot x = \sigma^k(x)$ 作用在 X 上；$[k]$ 的不同代表相差 p 的倍数，不改变 σ^k，故作用良定义。

一个点被此作用固定，当且仅当它被 σ 固定，即

$$g_1 = g_2 = \cdots = g_p = g, \qquad g^p = e.$$

因此不动点集与 $Y = \{g \in G : g^p = e\}$ 双射。不动点计数给出

$$|Y| = |X^{\mathbb{Z}/p\mathbb{Z}}| \equiv |X| = |G|^{p-1} \equiv 0 \pmod{p}.$$

$e \in Y$，所以 $|Y| \geq 1$；作为正的 p 的倍数，它至少为 p。故存在 $g \neq e$ 满足 $g^p = e$。元素阶的整除判别给出 $\mathrm{ord}(g) \mid p$，又 $\mathrm{ord}(g) > 1$，所以 $\mathrm{ord}(g) = p$。 □

6.15 Sylow 子群的存在性

定义 6.53 (Sylow 子群)**.** 设 $|G| = p^a m$，其中 $p \nmid m$。阶为 p^a 的子群称为 Sylow p-子群。所有这些子群组成的集合记为 $\mathrm{Syl}_p(G)$，其元素数记为 n_p。

对非负整数 N，记 $N! = 1 \cdot 2 \cdots N$，并令 $0! = 1$。从一个 N 元集合中依次选取 k 个不同元素，有 $N(N-1)\cdots(N-k+1)$ 种有序选择。同一个 k 元子集对应 $k!$ 种排列，因此 k 元子集的数目为整数

$$\binom{N}{k} = \frac{N(N-1)\cdots(N-k+1)}{k!} = \frac{N!}{k!(N-k)!} \qquad (0 \leq k \leq N).$$

引理 6.54 (一个二项系数的整除性质)**.** *若 $a \geq 1$、$m \geq 1$ 且 $p \nmid m$，则*

$$p \nmid \binom{p^a m}{p^a}.$$

证明. 对正整数 r，令 $v_p(r)$ 为整除 r 的 p 的最高幂的指数。由整数唯一分解，$v_p(rs) = v_p(r) + v_p(s)$。

对 $1 \leq i < p^a$，写 $i = p^t u$，其中 $p \nmid u$、$t < a$。则

$$p^a m - i = p^t(p^{a-t} m - u), \qquad p^a - i = p^t(p^{a-t} - u).$$

两个括号内的整数模 p 均等于 $-u$，所以均不被 p 整除，因而

$$v_p(p^a m - i) = t = v_p(p^a - i).$$

令 $B = \binom{p^a m}{p^a}$。由二项系数的整数乘积公式，

$$B\,p^a \prod_{i=1}^{p^a-1}(p^a - i) = p^a m \prod_{i=1}^{p^a-1}(p^a m - i).$$

两边取 v_p，逐项使用上式，得

$$v_p(B) + a + \sum_{i=1}^{p^a-1} v_p(p^a - i) = a + \sum_{i=1}^{p^a-1} v_p(p^a m - i).$$

相等的求和项消去后得到 $v_p(B) = 0$，即结论。 □

定理 6.55 (Sylow 存在定理)**.** *每个有限群 G 对每个素数 $p \mid |G|$ 都有 Sylow p-子群。*

证明. 写 $|G| = p^a m$，$p \nmid m$。令 X 为 G 的所有含 p^a 个元素的子集组成的集合。左平移

$$g \cdot A = \{ga : a \in A\}$$

保持子集大小，并满足作用公理，故给出 G 对 X 的作用。由上一引理，$|X| = \binom{p^a m}{p^a}$ 不被 p 整除。将 X 分成轨道后，至少有一个轨道 O 的大小不被 p 整除，否则各轨道大小之和也被 p 整除。

取 $A \in O$，令 $H = G_A$ 为其稳定子。轨道公式给出

$$|G| = |O||H|.$$

因此 $v_p(|H|) = v_p(|G|) - v_p(|O|) = a$，即 $p^a \mid |H|$。另一方面，固定 $a_0 \in A$，映射

$$H \longrightarrow A, \qquad h \longmapsto ha_0$$

良定义，因为 $hA = A$；它单射，因为 $ha_0 = h'a_0$ 蕴含 $h = h'$。所以 $|H| \le |A| = p^a$。结合正整数 $|H|$ 被 p^a 整除，得到 $|H| = p^a$，于是 H 是所需子群。 □

6.16 Sylow 子群的包含、共轭与数目

定理 6.56 (包含与共轭)**.** *设 P 为 G 的 Sylow p-子群，$H \le G$ 为任意 p-子群。存在 $g \in G$ 使 $g^{-1}Hg \subseteq P$。特别地，任意两个 Sylow p-子群共轭。*

证明. 令 H 在左陪集集合 G/P 上以左乘作用：$h \cdot (gP) = hgP$。若 $gP = g'P$，则 $g' = gp$，$p \in P$，从而 $hg'P = hgpP = hgP$，故作用良定义。

集合大小为 $[G : P] = m$，不被 p 整除。对 H 使用不动点计数得到

$$|(G/P)^H| \equiv |G/P| = m \not\equiv 0 \pmod{p},$$

故存在不动陪集 gP。其条件为

$$\begin{aligned} \forall h \in H,\quad hgP = gP &\iff \forall h \in H,\quad g^{-1}hg \in P \\ &\iff g^{-1}Hg \subseteq P. \end{aligned}$$

若 H 本身也为 Sylow 子群，则 $|g^{-1}Hg| = |H| = |P|$，有限集合的包含与等势给出 $g^{-1}Hg = P$。 □

定理 6.57 (Sylow 子群数目)**.** *若 $|G| = p^a m$、$p \nmid m$，则*

$$n_p \mid m, \qquad n_p \equiv 1 \pmod{p}.$$

Sylow p-子群正规，当且仅当它是唯一的 Sylow p-子群。

证明. 取 Sylow 子群 P。由共轭定理，$\mathrm{Syl}_p(G)$ 是子群共轭作用的一个轨道，稳定子为 $N_G(P)$，故

$$n_p = [G : N_G(P)].$$

因为 $P \le N_G(P)$，有限阶计数给出

$$m = [G : P] = [G : N_G(P)][N_G(P) : P] = n_p[N_G(P) : P],$$

所以 $n_p \mid m$。

再令 P 通过共轭作用在 $\mathrm{Syl}_p(G)$ 上。若 Q 是不动点，则 $P \subseteq N_G(Q)$。Q 在 $N_G(Q)$ 中正规，因此 PQ 为子群，并且乘积计数给出

$$|PQ| = \frac{|P||Q|}{|P \cap Q|}.$$

分子、分母都是 p 的幂，所以 $|PQ|$ 是 p 的幂；它整除 $|G|$，故不超过 $p^a = |Q|$。又 $Q \subseteq PQ$，只能 $PQ = Q$，从而 $P \subseteq Q$；两者阶相同，故 $P = Q$。反过来 P 的确固定自己。因此该作用唯一不动点为 P。由不动点计数，

$$n_p = |\mathrm{Syl}_p(G)| \equiv 1 \pmod{p}.$$

若 P 正规，则每个共轭 $gPg^{-1} = P$，由共轭定理所有 Sylow 子群均为 P。若只有一个，则每个共轭也是 Sylow 子群，只能等于 P，所以正规。 □

例 6.58 (阶为十五的群)**.** 设 $|G| = 15$。Sylow 定理给出

$$n_5 \mid 3, \quad n_5 \equiv 1 \pmod{5} \quad \Longrightarrow \quad n_5 = 1,$$

$$n_3 \mid 5, \quad n_3 \equiv 1 \pmod{3} \quad \Longrightarrow \quad n_3 = 1.$$

因此唯一的 Sylow 子群 P, Q 阶分别为 $5, 3$，都正规。由 Lagrange 定理，$|P \cap Q|$ 同时整除 $5, 3$，故 $P \cap Q = \{e\}$。

对 $x \in P, y \in Q$，交换子 $xyx^{-1}y^{-1}$ 同时属于 P, Q：

$$xyx^{-1}y^{-1} = x(yx^{-1}y^{-1}) \in P, \qquad xyx^{-1}y^{-1} = (xyx^{-1})y^{-1} \in Q.$$

所以交换子为 e，即 $xy = yx$。乘积计数给出 $|PQ| = 5 \cdot 3 = 15$，故 $PQ = G$。

取生成元 $x \in P, y \in Q$。若 $(xy)^k = e$，由交换性有 $x^k = y^{-k} \in P \cap Q$，故 $5 \mid k$ 且 $3 \mid k$，从而 $15 \mid k$；而 $(xy)^{15} = e$。因此 xy 阶为 15，得到 $G \cong \mathbb{Z}/15\mathbb{Z}$。

6.17 直积与交换群的直和

定义 6.59 (群的直积)**.** 设 $(G_i)_{i\in I}$ 为群族。直积 $\prod_{i\in I} G_i$ 是所有元组 $(g_i)_{i\in I}$ 的集合，定义逐坐标乘法 $(g_i)(h_i)=(g_ih_i)$。

群公理逐坐标验证。例如

$$((g_i)(h_i))(k_i)=((g_ih_i)k_i)=(g_i(h_ik_i))=(g_i)((h_i)(k_i)).$$

单位为 (e_i)，逆元为 (g_i^{-1})，因为其两种乘积的每个坐标均为 e_i。坐标投影 $\pi_i:\prod_j G_j\to G_i$ 为同态。

命题 6.60 (直积的映射性质)**.** *给定群 K 及同态族 $f_i:K\to G_i$，存在唯一同态 $f:K\to\prod_i G_i$ 满足 $\pi_i f=f_i$。*

证明. 被要求的坐标强迫 $f(k)=(f_i(k))_i$。这个公式给出映射，且

$$f(kk')=(f_i(kk'))_i=(f_i(k)f_i(k'))_i=f(k)f(k').$$

故为同态，并满足每个坐标条件。如果 u 也满足条件，则 $u(k)$ 的每个坐标都为 $f_i(k)$，所以 $u(k)=f(k)$，得到唯一性。 □

定义 6.61 (交换群的直和)**.** 对交换群族 $(A_i)_{i\in I}$，采用加法记号。元组 $a=(a_i)$ 的支撑是 $\{i:a_i\neq 0\}$。直和 $\bigoplus_i A_i$ 为直积中有限支撑元组组成的子集。

零元支撑为空；若 a,b 有限支撑，则 $a-b$ 的支撑包含于两个支撑的并，因而仍有限。所以直和为子群。记 $\iota_i:A_i\to\bigoplus_j A_j$ 为第 i 个坐标的包含。

命题 6.62 (直和的映射性质)**.** *给定交换群 B 及同态族 $u_i:A_i\to B$，存在唯一同态 $u:\bigoplus_i A_i\to B$ 满足 $u\iota_i=u_i$。*

证明. 对有限支撑元组 $a=(a_i)$，定义 $u(a)=\sum_i u_i(a_i)$。只有有限个非零项，故和有定义。对有限支撑 a,b，取包含两者支撑的有限集 J，由 B 的交换律可重排求和，得到

$$u(a+b)=\sum_{i\in J}u_i(a_i+b_i)=\sum_{i\in J}u_i(a_i)+\sum_{i\in J}u_i(b_i)=u(a)+u(b).$$

所以它为同态，且 $u\iota_i=u_i$。任何同态 v 满足这些条件时，因 $a=\sum_i\iota_i(a_i)$，必有 $v(a)=\sum_i u_i(a_i)=u(a)$，故唯一。 □

6.18 自由群的字与约简

设 S 为集合。为严格区分字母和逆字母，取带标记集合 $\mathcal{A}=S\times\{1,-1\}$，记 $(s,1)$ 为 s、$(s,-1)$ 为 s^{-1}；此处上标只表示形式记号。定义字母的反转 $a\mapsto\bar a$，使 $\bar s=s^{-1}$、$\overline{s^{-1}}=s$，于是 $\bar{\bar a}=a$。

定义 6.63 (字与约简字)**.** 字是 $\mathcal{A}$ 中字母的有限序列 $a_1\cdots a_r$，长度为 r。长度为零的字记为 ε。相邻两字母为 $a\bar a$ 时，可以删除它们；也允许在任意位置插入这样的相邻对。由有限次插入或删除连接的字称为等价字，记为 $w\sim v$。不含相邻反转字母的字称为约简字。

空步、逆向操作以及拼接操作序列，分别保证 $\sim$ 的自反、对称和传递性，所以它是等价关系。任一删除使长度减少 2，故连续删除必在有限步后停在某个约简字。还须证明结果与删除顺序无关。

定理 6.64 (约简字的唯一性)**.** *每个等价类恰含一个约简字。*

证明. 构造确定的读字过程。维护一个约简字 r，初始为 ε。读入字母 a 时，若 r 的末字母为 $\bar{a}$，就删除该末字母；否则把 a 接在末尾。记这个操作为 $D_a(r)$。两种情形的结果都仍约简。

我们证明 $D_{\bar{a}}D_a(r) = r$。若 r 不以 $\bar{a}$ 结束，则 $D_a(r) = ra$，再读 $\bar{a}$ 时删掉刚加上的 a，所以返回 r。若 $r = t\bar{a}$，因为 r 约简，t 为空或不以 a 结束；于是

$$D_a(t\bar{a}) = t, \qquad D_{\bar{a}}(t) = t\bar{a} = r.$$

所以两种情况都成立；将 a 换成 $\bar{a}$ 也得到 $D_aD_{\bar{a}} = \mathrm{id}$。

将字 w 读完的结果记为 $R(w)$。在任何前缀读完后的约简状态插入 $a\bar{a}$，连续操作 $D_a, D_{\bar{a}}$ 不改变状态，后续同一后缀的读取也就相同。因此

$$R(ua\bar{a}v) = R(uv)$$

对任意前后缀 u, v 成立，故 $w \sim v$ 蕴含 $R(w) = R(v)$。

读字过程的每一步或者接入一个字母，或者删除一对相邻反转字母。对前缀长度归纳，所得约简状态与该前缀等价，因而 $w \sim R(w)$。若 w 本来约简，读取时每一步都只接入字母，故 $R(w) = w$。所以每个等价类有约简代表 $R(w)$；两个约简字若等价，就分别等于自身的 R 值，而等价字的 R 值相同，故两者相等。 □

定理 6.65 (自由群的构造)**.** *令 $F(S)$ 为所有字的等价类组成的集合。在其上定义 $[u][v] = [uv]$，则这是群，称为以 S 为自由生成集的自由群。单位为 $[\varepsilon]$，并且*

$$[a_1 \cdots a_r]^{-1} = [\bar{a}_r \cdots \bar{a}_1].$$

证明. 若 $u \sim u'$，对相同后缀 v 保留每次插入、删除的位置，得到 $uv \sim u'v$。若 $v \sim v'$，对相同前缀 u' 同样得到 $u'v \sim u'v'$。由传递性 $uv \sim u'v'$，所以乘法与代表元无关。

有限序列的连接满足 $(uv)w = u(vw)$，故

$$([u][v])[w] = [(uv)w] = [u(vw)] = [u]([v][w]).$$

又 $u\varepsilon = \varepsilon u = u$，所以空字给出单位。依次删去中间的 $a_r\bar{a}_r$、$a_{r-1}\bar{a}_{r-1}$ 等相邻对，可得

$$a_1 \cdots a_r\bar{a}_r \cdots \bar{a}_1 \sim \varepsilon.$$

反向次序的乘积同样约到空字，故所给公式为两侧逆元。 □

定理 6.66 (自由群的泛性质)**.** *每个集合映射 $u : S \to K$，其中 K 为群，都唯一延拓为同态 $\widetilde{u} : F(S) \to K$，满足 $\widetilde{u}([s]) = u(s)$。*

证明. 先对形式字母定义 $u(s^{-1}) = u(s)^{-1}$，再令

$$\widetilde{u}([a_1 \cdots a_r]) = u(a_1) \cdots u(a_r), \qquad \widetilde{u}([\varepsilon]) = e_K.$$

若在词中插入 $a\bar{a}$，右侧相应插入 $u(a)u(a)^{-1} = e_K$，不改变乘积；删除同样不改变。因此它与代表元无关。连接两个字时右侧乘积也连接，故 $\widetilde{u}([v][w]) = \widetilde{u}([v])\widetilde{u}([w])$。任意延拓同态必须将 $[s]^{-1}$ 送到 $u(s)^{-1}$，再由保持有限乘积，其每个字类上的值均被上述公式确定，所以唯一。 □

例 6.67 (两个字母时的计算)**.** 在 $F(\{a,b\})$ 中，

$$(ab^{-1}a^{-1})(ab) = ab^{-1}(a^{-1}a)b = ab^{-1}b = a.$$

这里每一步只消去相邻逆字母。字 $aba^{-1}b^{-1}$ 没有相邻可消去对，且非空；由约简唯一性，它不代表单位元，所以 $ab \neq ba$。单字母自由群的约简字只能为 a^k，$k \in \mathbb{Z}$，乘法为指数相加，因而 $F(\{a\}) \cong \mathbb{Z}$。当 $S = \varnothing$ 时只有空字，故自由群为平凡群。

6.19 正规闭包与生成元关系

定义 6.68 (正规闭包)**.** 对 $R \subseteq G$，包含 R 的全部正规子群的交记为 $\langle\langle R\rangle\rangle$，称为 R 在 G 中的正规闭包。

命题 6.69 (正规闭包的乘积描述)**.** $\langle\langle R\rangle\rangle$ 恰为所有形如

$$\prod_{i=1}^{r} x_i r_i^{\epsilon_i} x_i^{-1}, \qquad x_i \in G, \quad r_i \in R, \quad \epsilon_i \in \{1,-1\}, \quad r \geq 0$$

的有限乘积组成的集合。

证明. 记这个集合为 W。空积给出 e，连接两个有限乘积仍在 W，而

$$\left(\prod_{i=1}^{r} x_i r_i^{\epsilon_i} x_i^{-1}\right)^{-1} = \prod_{i=r}^{1} x_i r_i^{-\epsilon_i} x_i^{-1} \in W,$$

右侧按 $i = r, r-1, \dots, 1$ 的顺序相乘。所以 W 为子群。对 $g \in G$，

$$g\left(\prod_{i=1}^{r} x_i r_i^{\epsilon_i} x_i^{-1}\right) g^{-1} = \prod_{i=1}^{r} (gx_i) r_i^{\epsilon_i} (gx_i)^{-1} \in W,$$

所以 W 正规。又 $r = ere^{-1} \in W$，故 $R \subseteq W$，从而 $\langle\langle R\rangle\rangle \subseteq W$。

若 N 为任意含 R 的正规子群，则 $r_i^{\epsilon_i} \in N$，正规性给出 $x_i r_i^{\epsilon_i} x_i^{-1} \in N$，乘法封闭给出 $W \subseteq N$。对所有这样的 N 取交得 $W \subseteq \langle\langle R\rangle\rangle$，所以相等。 □

定义 6.70 (群的表示)**.** 设 S 为集合，$R \subseteq F(S)$ 为一组字类，定义

$$\langle S \mid R\rangle = F(S)/\langle\langle R\rangle\rangle.$$

称之为由生成元 S、关系字 R 给出的群。关系 $u = v$ 表示把 uv^{-1} 列入关系字。

命题 6.71 (表示所规定的同态)**.** 给定群 K，映射 $u : S \to K$ 能延拓为 $\langle S \mid R\rangle \to K$ 的同态，当且仅当把各字母代入 u 后，每个关系字的值均为 e_K。存在时延拓唯一。

证明. 自由群泛性质先给出唯一同态 $f : F(S) \to K$。若全部关系字被消去，则 $R \subseteq \ker f$；核正规，故 $\langle\langle R\rangle\rangle \subseteq \ker f$。商群因子化定理给出唯一下降同态。反过来，若 f 能下降，则 $\langle\langle R\rangle\rangle \subseteq \ker f$，特别每个 $r \in R$ 映为单位元。 □

例 6.72 (一个循环群与一个自由交换群)**.** 单字母自由群同构于 $\mathbb{Z}$, 其子群都正规。因此 $\langle a \mid a^n\rangle$ 的关系子群对应 $n\mathbb{Z}$，从而此群同构于 $\mathbb{Z}/n\mathbb{Z}$。

再取 $P=\langle a,b \mid aba^{-1}b^{-1}\rangle$。关系等价于 $ab=ba$，因此 P 的每个元素可整理为 a^mb^n。映射 $a\mapsto(1,0)$、$b\mapsto(0,1)$ 消去关系，给出同态 $f:P\to\mathbb{Z}^2$。反向定义 $h:\mathbb{Z}^2\to P$，$h(m,n)=a^mb^n$，它为同态，因为

$$h(m,n)h(r,s)=a^mb^na^rb^s=a^{m+r}b^{n+s}=h(m+r,n+s).$$

于是 $fh(m,n)=(m,n)$；而 hf 固定生成元 a,b，所以固定它们的所有乘积，得到 $hf=\mathrm{id}_P$。因此 $P\cong\mathbb{Z}^2$，这也证明整理后没有额外关系。

每个群都有表示。具体地, 对群 G 的每个元素取一个标记字母, 集合记为 S, 映射 $S\to G$ 把标记送回相应元素。其自由延拓 $f:F(S)\to G$ 为满射，因为每个 g 已是某个字母的像。因此第一同构定理给出 $G\cong F(S)/\ker f$。取关系集为 $\ker f$ 即得一个表示。

6.20 自由积与融合自由积的商构造

定理 6.73 (自由积的存在性与泛性质)**.** *对群 G,H，存在群 $G*H$ 及同态 $\iota_G:G\to G*H$、$\iota_H:H\to G*H$，使任意同态 $u:G\to K, v:H\to K$ 唯一延拓为同态 $w:G*H\to K$。两个指定同态 ι_G,ι_H 均为单射。*

证明. 取 G,H 的带标记不交副本为字母集 S，相应字母记为 $[g]_G,[h]_H$。在 $F(S)$ 中，对所有 $g,g'\in G, h,h'\in H$ 取关系

$$[e_G]_G,\quad [e_H]_H,\quad [g]_G[g']_G[gg']_G^{-1},\quad [h]_H[h']_H[hh']_H^{-1}.$$

令 Q 为按这些关系取商的群。字母 $[g]_G$ 在 Q 中的像定义 $\iota_G(g)$；关系保证

$$\iota_G(e_G)=e_Q,\qquad \iota_G(g)\iota_G(g')=\iota_G(gg'),$$

故 ι_G 为同态。H 的情形相同。

将 $[g]_G$ 代为 $u(g)$，$[h]_H$ 代为 $v(h)$。由于 u,v 为同态，所有关系的值为单位元，所以表示的泛性质给出 $w:Q\to K$。它在全部字母上的值已经固定，这些字母生成 Q，故 w 唯一。取 $G*H=Q$。

为了证明 ι_G 单射，应用刚证明的泛性质于 $u=\mathrm{id}_G$ 和平凡同态 $v:H\to G$，得到 $r_G:Q\to G$，满足 $r_G\iota_G=\mathrm{id}_G$。若 $\iota_G(g)=\iota_G(g')$，作用 r_G 即得 $g=g'$。交换两群同样得到 ι_H 单射。 □

因此可将两个群视为自由积中的子群。这里“自由”由同态的唯一延拓性质表达；构造中的关系只要求各因子内的乘法保持。

定义 6.74 (融合自由积)**.** 给定同态 $i:C\to G, j:C\to H$，在 $G*H$ 中取正规子群

$$N=\langle\!\langle \iota_G(i(c))\iota_H(j(c))^{-1} : c\in C\rangle\!\rangle,$$

定义 $G*_C H=(G*H)/N$。当 i,j 为单射时，称为沿共同子群 C 的融合自由积；构造数据包含这两条指定映射。

命题 6.75 (融合自由积的映射性质)**.** *若 $u : G \to K, v : H \to K$ 为同态且 $ui = vj$，则存在唯一同态 $w : G *_C H \to K$ 延拓两者的指定映射。*

证明. 自由积泛性质先给出 $f : G * H \to K$。对每个 $c \in C$，

$$f(\iota_G(i(c))\iota_H(j(c))^{-1}) = u(i(c))v(j(c))^{-1} = e_K.$$

核正规，故 $N \subseteq \ker f$；因子化定理给出所需下降同态 w。若另有 w'，将其与商投影复合，便是同一对 u, v 的自由积延拓，由唯一性得 $w'q = wq$；投影满射，再得 $w' = w$。 □

例 6.76 (两种自由构造的计算)**.** 两个无限循环群的自由积同构于 $F(\{a, b\})$。因为从它们出发到任意群的两个同态分别由 a, b 的任意像决定，这正是两生成元自由群的泛性质。分别延拓两个标准生成元的包含得到双向同态，两复合固定生成元，因而为恒等。

取 $C = \langle c \rangle \cong \mathbb{Z}$，指定 $c \mapsto a^m$、$c \mapsto b^n$，其中 $m, n \geq 1$。由无限循环群的幂判别，两映射单射。融合关系为所有 $a^{mk}b^{-nk}$，$k \in \mathbb{Z}$。关系 $a^m b^{-n} = e$ 已蕴含 $a^m = b^n$，从而对所有 k 有 $a^{mk} = b^{nk}$，因此其正规闭包与全部融合关系的正规闭包相同。故得到表示

$$\langle a, b \mid a^m b^{-n} \rangle.$$

当 $m = n = 1$ 时 $a = b$，此群是无限循环群；仅给出两个生成元并不能据此判断群的秩，关系同样决定对象。

文献：基本定义、陪集、正规性和同构定理参见 [4] 第 1 章；自由群与表示参见其第 2 章；自同构与群作用参见其第 3–4 章；Sylow 定理参见其第 5 章。综合教材参见 [8, 9]。本章的 Cauchy 定理采用循环移动元素组的证明，约简唯一性采用逐字读取的构造。

7 线性代数

7.1 从坐标运算到向量空间

实数对 (x,y) 可以相加，可以乘以一个实数。例如

$$(1,2)+(3,-1)=(4,1),\qquad 2(1,2)=(2,4).$$

同样的运算也可用于函数：把两个函数在每一点的值相加，或将每个值乘以一个实数。坐标对与函数是不同的对象，但运算遵循相同规律。向量空间公理提取这些规律，让线性方程和线性组合的推理同时适用于它们。

定义 7.1 (域与标量)**.** 集合 $\mathbb{F}$ 连同加法与乘法称为域，若：其加法构成交换群，单位元记为 0；非零元素在乘法下构成交换群，单位元记为 1；有 $1\neq 0$；并且乘法对加法满足分配律。域的元素称为标量。

例如 $\mathbb{Q},\mathbb{R},\mathbb{C}$ 为域。$\mathbb{Z}$ 不是域，因为整数 2 没有整数乘法逆元。若 p 为素数，$\mathbb{Z}/p\mathbb{Z}$ 为域，因为整数一章证明了每个非零类都可逆。域公理保证遇到非零系数时可以相除；后面的交换引理正要使用这一点。

定义 7.2 (向量空间)**.** 域 $\mathbb{F}$ 上的向量空间为集合 V，连同加法 $V\times V\to V$ 和数乘 $\mathbb{F}\times V\to V$，使 $(V,+)$ 为交换群，并满足对任意 $a,b\in\mathbb{F},u,v\in V$，

$$a(u+v)=au+av,\quad (a+b)v=av+bv,\quad (ab)v=a(bv),\quad 1v=v.$$

V 的元素称为向量，其加法单位元记为 0_V，通常简记为 0。

例 7.3 (坐标空间与函数空间)**.** $\mathbb{F}^n$ 表示所有有序 n 元组 $(x_1,\dots,x_n)$，按坐标相加与数乘。它是向量空间，因为每条公理可逐坐标归结到域的运算规律。

对任意集合 X，记 $\mathbb{F}^X$ 为所有函数 $X\to\mathbb{F}$ 的集合，定义 $(f+g)(x)=f(x)+g(x)$，$(af)(x)=af(x)$。逐点验证同样说明它是向量空间。其零向量是恒为 0 的函数。于是“向量”可以是一个函数，定义关注的是运算性质。

群论已经保证零向量与加法逆元唯一。数乘还需满足一些可以从公理推出的公式。

命题 7.4 (零标量与负向量)**.** 对每个 $v\in V,a\in\mathbb{F}$，有 $0v=0_V$，$a0_V=0_V$，$(-1)v=-v$。

证明. 由 $0v=(0+0)v=0v+0v$，在加法群中消去 $0v$ 得 $0v=0_V$。由 $a0_V=a(0_V+0_V)=a0_V+a0_V$，同样得到第二式。最后

$$v+(-1)v=(1+(-1))v=0v=0_V,$$

所以 $(-1)v$ 为 v 的加法逆元，由唯一性得第三式。 □

第一式左边的 0 属于标量域，右边的 0_V 属于向量空间；当两种对象都简记为 0 时，仍应根据运算位置区分。

注 7.5 (从坐标空间到函数空间)**.** Axler 在 [10] 第 1B 节、第 13–14 页，将坐标空间、序列空间与函数空间并列。其联系可以直接写出：元组 $(x_1,\dots,x_n)$ 对应有限集合 $\{1,\dots,n\}$ 上的函数 $k\mapsto x_k$，按坐标运算就是按函数值运算。把定义域换成无限集合，会改变可用的函数，却不会改变加法与数乘的公理。

7.2 子空间与最小封闭集合

定义 7.6 (子空间)**.** 子集 $U \subseteq V$ 若在原有加法与数乘下本身为向量空间，称为 V 的子空间，记为 $U \le V$。

命题 7.7 (子空间判别)**.** *U 为子空间，当且仅当 $0 \in U$，且对任意 $u, v \in U, a, b \in \mathbb{F}$，有 $au + bv \in U$。*

证明. 若 U 是向量空间，则其零向量与 V 的零向量相同：由群的子群判别证明可知子群与原群的单位相同。数乘和加法封闭给出 $au + bv \in U$。

反过来，设所列条件成立。对 $u, v \in U$、$a \in \mathbb{F}$，分别取系数得到

$$u + v = 1u + 1v \in U, \qquad -u = (-1)u + 0v \in U, \qquad au = au + 0v \in U.$$

因此加法和数乘在 U 上封闭，且 $0_V \in U$。在限制后的运算中，

$$(u + v) + w = u + (v + w), \quad u + v = v + u, \quad u + 0_V = u, \quad u + (-u) = 0_V,$$

以及

$$a(u + v) = au + av, \quad (a + b)u = au + bu, \quad (ab)u = a(bu), \quad 1u = u$$

均是 V 中已成立的等式，且各项都在 U。所以这些等式验证全部向量空间公理。 □

例 7.8 (齐次条件与非齐次条件)**.** $\mathbb{R}^2$ 中的集合 $U = \{(x, y) : x + y = 0\}$ 为子空间：两个满足等式的向量的线性组合仍满足等式。集合 $L = \{(x, y) : x + y = 1\}$ 不包含零向量，故不是子空间。相同的线性左端，右端是否为零会改变集合的结构。

给定向量 $v_1, \dots, v_m$，任何包含它们的子空间都必须包含它们的数乘，再包含这些数乘的有限和。这给出从指定向量得到最小子空间的方法。

定义 7.9 (线性组合与张成)**.** 表达式 $a_1 v_1 + \cdots + a_m v_m$，其中 $a_i \in \mathbb{F}$，称为给定向量组的线性组合。所有这些组合组成的集合称为张成空间，记为 $\operatorname{span}(v_1, \dots, v_m)$。空向量组的张成为 $\{0\}$。

命题 7.10 (张成的最小性)**.** *$\operatorname{span}(v_1, \dots, v_m)$ 是包含这些向量的最小子空间。*

证明. 记 $S = \{\sum_i a_i v_i : a_i \in \mathbb{F}\}$。取全零系数得 $0 \in S$。若 $x = \sum_i a_i v_i$、$y = \sum_i b_i v_i$，则对 $c, d \in \mathbb{F}$，

$$cx + dy = \sum_i (ca_i + db_i) v_i \in S.$$

由子空间判别得 $S \le V$。取第 j 个系数为 1、其余为零，得到 $v_j \in S$。若 $U \le V$ 含所有 v_j，则数乘、有限加法封闭保证 $\sum_i a_i v_i \in U$，所以 $S \subseteq U$。这证明其最小性。 □

定义 7.11 (子空间的和与交)**.** 若 $U, W \le V$，定义 $U + W = \{u + w : u \in U, w \in W\}$；$U \cap W$ 为同时属于二者的向量集合。

$U + W$ 是同时包含 U, W 的最小子空间；$U \cap W$ 是同时包含在二者中的最大子空间。以和为例，加法与数乘可分别在 U, W 中完成，故它是子空间；任何包含二者的子空间都包含所有 $u + w$。交则因两个封闭条件同时成立而封闭。

普通并集通常不满足这样的封闭性。在 $\mathbb{R}^2$ 中，x 轴与 y 轴的并包含 $(1, 0)$ 和 $(0, 1)$，却不包含它们的和 $(1, 1)$；二者的和才是整个平面。

7.3 线性无关来自表示的唯一性

设 v 已属于 $\operatorname{span}(v_1,\dots,v_m)$，所以至少存在一份表示 $v=\sum_i a_i v_i$。若还有一份 $v=\sum_i b_i v_i$，相减得到

$$0=\sum_{i=1}^{m}(a_i-b_i)v_i.$$

因此两份表示是否相同，由零向量能否以非零系数表示决定。

定义 7.12 (线性无关与线性相关)**.** 向量组 $v_1,\dots,v_m$ 线性无关，若 $\sum_i c_i v_i=0$ 蕴含所有 $c_i=0$。否则称为线性相关，即存在不全为零的系数使该组合为零。空向量组约定为线性无关。

命题 7.13 (无关性与表示唯一性)**.** 一个向量组线性无关，当且仅当其张成空间中的每个向量都有唯一的线性组合表示。

证明. 若向量组无关，而 $\sum_i a_i v_i=\sum_i b_i v_i$，则

$$\sum_i (a_i-b_i)v_i=0 \quad\Longrightarrow\quad a_i-b_i=0\ (\forall i),$$

所以两份表示系数相同。反过来，若表示唯一，任意 $\sum_i c_i v_i=0$ 都可与 $0=\sum_i 0v_i$ 比较，由唯一性得 $c_i=0$ 对每个 i 成立，即线性无关。 □

例 7.14. 在 $\mathbb{R}^2$ 中，$v_1=(1,0), v_2=(0,1), v_3=(1,1)$ 张成整个平面，但 $v_3=v_1+v_2$。所以 $v_1+v_2-v_3=0$ 是一份非全零系数的关系。向量 $(1,1)$ 既能写成 v_3，也能写成 v_1+v_2，表示不唯一。

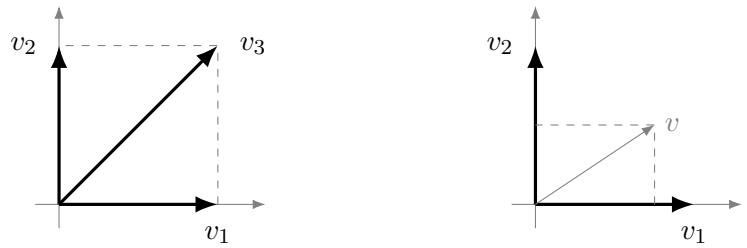


图 7.1: 删去冗余向量 v_3 后仍张成平面，并恢复每个向量的唯一表示。

这还说明相关性意味着冗余。若某个关系中 $c_k\neq 0$，利用标量域中的除法，得到

$$v_k=-\sum_{i\neq k}\frac{c_i}{c_k}v_i.$$

所以删掉 v_k 不会改变张成空间：凡是原来用到 v_k 的组合，均可用右边代入。这一删除操作是构造基的依据。

7.4 基、交换引理与维数

定义 7.15 (基与有限维)**.** 张成 V 且线性无关的向量组称为 V 的基。若 V 有有限的张成组，称 V 有限维；否则称无限维。对无限向量族，线性组合只允许有限个非零系数，无关性要求每个有限子族无关。

由上一节，基恰好同时保证每个向量表示的存在性和唯一性。但还需回答：基是否存在？不同基是否有相同长度？

引理 7.16 (无关组的长度受张成组控制)**.** *若 $w_1,\dots,w_m$ 张成 V，$u_1,\dots,u_r$ 线性无关，则 $r \le m$。*

证明. 从张成组 $(w_1,\dots,w_m)$ 出发逐次替换。设已得到张成组

$$u_1,\dots,u_k,z_{k+1},\dots,z_m.$$

若 $k<r$，其张成性给出

$$u_{k+1}=\sum_{i=1}^{k}a_iu_i+\sum_{j=k+1}^{m}b_jz_j.$$

若所有 $b_j=0$，则 $\sum_{i=1}^{k}a_iu_i-u_{k+1}=0$，其中 u_{k+1} 的系数为 $-1\neq0$，违背无关性。因此某个 $b_\ell\neq0$。域中可除以它，得到

$$z_\ell=b_\ell^{-1}u_{k+1}-\sum_{i=1}^{k}b_\ell^{-1}a_iu_i-\sum_{\substack{j>k\\ j\neq\ell}}b_\ell^{-1}b_jz_j.$$

所以用 u_{k+1} 替换 z_ℓ 后，原张成组的每个元素仍能由新组表示，故新组仍张成 V。重排剩余 z_j，可重复操作。

若 $r>m$，重复 m 次得到 $(u_1,\dots,u_m)$ 张成 V。于是 $u_{m+1}=\sum_{i=1}^{m}c_iu_i$，再次得到一个系数不全为零的无关组关系，矛盾。因此 $r\le m$。 □

定理 7.17 (基的存在、删减与扩充)**.** *有限维空间有基；其任意有限张成组可删减为基，任意线性无关组可扩充为基。有限维空间的子空间也有限维。*

证明. 从有限张成组出发，若相关，就按上一节的关系删除一个冗余向量；每步保持张成，有限次后得到无关的张成组，即基。

给定无关组，将一个有限张成组接在它后面，依次检查新加向量是否在此前向量的张成中：若在，就删除；若不在，就保留。保留步骤不会破坏无关性，因为若有新关系且新向量系数非零，就能解出它属于此前张成。全部检查完仍张成 V，得到所需基。

对 $U\le V$，从空组开始；只要还未张成 U，就取一个不在已有张成中的 $u\in U$ 加入。每步仍无关。引理 7.16 用 V 的某个有限张成组限制了步骤数，故有限步后必张成 U。 □

定理 7.18 (维数的良定义性)**.** *有限维空间的任意两组基有相同长度。*

证明. 一组基无关，另一组基张成，由引理得第一组长度不超过第二组。交换二者角色得到反向不等式，故相等。 □

现在才将这一共同长度定义为 $\dim V$，称为维数。零空间 $\{0\}$ 的基为空组，维数为 0。证明也说明，有限维子空间的维数不超过原空间的维数。

定义 7.19 (坐标)**.** 固定有序基 $\mathcal{B}=(v_1,\dots,v_n)$。由基的表示唯一性，任意 $v\in V$ 唯一写为 $v=\sum_i a_iv_i$，称列向量 $[v]_{\mathcal{B}}=(a_1,\dots,a_n)^{t}$ 为 v 关于该基的坐标；上标 t 表示将横排的数写成竖列。

坐标同时依赖向量和所选基。在 $\mathbb{R}^2$ 的标准基 $e_1=(1,0), e_2=(0,1)$ 下，$(2,1)$ 的坐标为 $(2,1)^t$；在 $b_1=(1,1), b_2=(1,-1)$ 下，因 $(2,1)=\frac{3}{2}b_1+\frac{1}{2}b_2$，坐标为 $(3/2,1/2)^t$。空间中的向量相同，两份坐标描述不同。

7.4.1 坐标映射与抽象空间的具体模型

给定有序基 $\mathcal{B}=(v_1,\dots,v_n)$，定义

$$\kappa_{\mathcal{B}}: V \longrightarrow \mathbb{F}^n, \qquad \kappa_{\mathcal{B}}(v)=[v]_{\mathcal{B}}.$$

它的逆规则是 $(a_1,\dots,a_n)^t \mapsto \sum_i a_i v_i$。基的张成性保证每个 v 都能送到一份坐标，无关性保证这份坐标唯一。再由

$$\left[\sum_i a_i v_i + \sum_i b_i v_i\right]_{\mathcal{B}} = (a_1+b_1,\dots,a_n+b_n)^t$$

可知，坐标保留加法；标量 $c\in\mathbb{F}$ 的作用满足

$$\left[c\sum_i a_i v_i\right]_{\mathcal{B}} = (ca_1,\dots,ca_n)^t = c\left[\sum_i a_i v_i\right]_{\mathcal{B}},$$

所以也保留数乘。下一节将把这种保留运算的映射称为线性映射。

因此，选定一组基后，V 中的运算可以完全转到 $\mathbb{F}^n$ 中完成。但基的选取仍是额外数据：向量空间公理并没有指定哪组向量必须充当坐标方向。例如在一个平面中，任何两条不平行的过原点方向都能给出基，不要求两条方向垂直，也不要求作为基的向量长度相同。

例 7.20 (多项式也是向量)**.** 固定非负整数 d，记 $\mathcal{P}_d(\mathbb{R})$ 为所有形如

$$p(t)=a_0+a_1t+\cdots+a_dt^d, \qquad a_i\in\mathbb{R}$$

的实多项式函数，按函数的加法和数乘运算。这里 t 为实变量；允许最高的若干系数为零，所以零函数也在其中。相加、数乘不会产生次数大于 d 的项，故这是函数空间的子空间。

函数 $1,t,\dots,t^d$ 张成它。它们也无关：若一个系数不全为零的多项式在所有实数处为零，则它有无限多个根，与非零 k 次实多项式至多有 k 个不同实根矛盾。后一个结论可由

$$t^j-a^j=(t-a)(t^{j-1}+t^{j-2}a+\cdots+a^{j-1})$$

推出。具体地，设 $p(t)=\sum_{j=0}^{k}a_jt^j$，$a_k\neq 0$，且 $p(a)=0$，则

$$p(t)=p(t)-p(a)=(t-a)\sum_{j=1}^{k}a_j\sum_{\ell=0}^{j-1}t^{j-1-\ell}a^{\ell}.$$

右端的第二因子记为 $q(t)$；它的最高次项为 a_kt^{k-1}，所以次数为 $k-1$。对 k 归纳：非零常数没有根；若 p 没有根，结论成立；若有一个根 a，则任意其他根 $b\neq a$ 满足

$$0=p(b)=(b-a)q(b), \qquad b-a\neq 0,$$

所以 $q(b)=0$。由归纳假设，其他根至多有 $k-1$ 个，加上 a 共至多 k 个。这就证明了根数限制。因此 $(1,t,\dots,t^d)$ 为基，$\dim\mathcal{P}_d(\mathbb{R})=d+1$。

在 $\mathcal{P}_2(\mathbb{R})$ 中，$p(t)=2-t+3t^2$ 的坐标是 $(2,-1,3)^t$。坐标列有三个分量，而向量本身是一个函数。相同的系数运算同时描述这两个对象。

注 7.21 (Axler 对表示问题的处理)**.** Axler 在 [10] 第 2A、2B 节先研究线性表示是否唯一，再用张成与无关刻画基。由这一安排可以把两项条件分开理解：张成保证每个向量都有表达式，无关保证没有两份不同的系数表达同一向量。坐标正是这两项结论合在一起的产物。这里概述的是原书的论证次序。

7.4.2 改变标量域会改变维数

把 $\mathbb{C}$ 看成 $\mathbb{C}$ 上的向量空间，(1) 是基，因为每个 $z \in \mathbb{C}$ 都唯一写成 $z \cdot 1$。把同一个集合看成 $\mathbb{R}$ 上的向量空间，则 $(1, i)$ 是基：每个复数唯一写成 $a + bi$，其中 $a, b \in \mathbb{R}$，$i^2 = -1$。所以

$$\dim_{\mathbb{C}} \mathbb{C} = 1, \qquad \dim_{\mathbb{R}} \mathbb{C} = 2.$$

下标说明允许使用哪些系数。在实标量下，i 不能由 1 的实数倍得到；在复标量下，系数本身可以取 i，故 $i = i \cdot 1$。无关性、张成和维数都依赖标量域。

无限维也可由已经建立的定义检验。全体实多项式函数的空间记为 $\mathcal{P}(\mathbb{R})$。任给有限多个多项式，取它们次数的最大值为 d，则它们的任意线性组合仍在 $\mathcal{P}_d(\mathbb{R})$，不能表示 t^{d+1}。故 $\mathcal{P}(\mathbb{R})$ 没有有限张成组。这里无限维的原因是任何有限组都遗漏了更高次的方向。

7.5 线性映射怎样由一组基确定

定义 7.22 (线性映射与同构)**.** 映射 $T : V \to W$ 线性，若对任意 $u, v \in V, a, b \in \mathbb{F}$，有 $T(au + bv) = aT(u) + bT(v)$。线性双射称为线性同构。

令 $a = b = 0$ 可得 $T(0) = 0$。线性映射还保持任意有限线性组合。因此一旦给定基向量的像，就有

$$T\left(\sum_i a_i v_i\right) = \sum_i a_i T(v_i),$$

其余值全部被确定。反过来，任意指定 $w_i \in W$，规则 $\sum_i a_i v_i \mapsto \sum_i a_i w_i$ 因坐标唯一而良定义，并且线性。若改从一个相关的张成组出发，就不能任意指定像：例如 $v_3 = v_1 + v_2$ 强制 $T(v_3) = T(v_1) + T(v_2)$。

7.5.1 线性条件控制整条直线

由 $T(av) = aT(v)$，过原点的直线 $\{av : a \in \mathbb{F}\}$ 被送到 $\{aT(v) : a \in \mathbb{F}\}$；若 $T(v) = 0$，整条直线的像只有零向量。一般仿射直线上的点为 $p + av$，则

$$T(p + av) = T(p) + aT(v).$$

所以一条直线的像仍是一条直线或一个点。实空间中，对 $0 \le s \le 1$，

$$T((1 - s)u + sv) = (1 - s)T(u) + sT(v),$$

表示线段上的分点连同参数 s 一起被保留。线性映射可以改变长度和夹角，但其对线性组合的作用由同一份系数控制。

单独检验 $T(0) = 0$ 不够。例如 $F : \mathbb{R} \to \mathbb{R}$，$F(x) = x^2$，虽有 $F(0) = 0$，却有 $F(1+1) = 4 \neq F(1) + F(1) = 2$。平移 $G(v) = v + b$ 在 $b \neq 0$ 时甚至不保持零向量，故也不是线性的。

7.5.2 伸缩、反射、剪切与旋转

在 $\mathbb{R}^2$ 中，令 $e_1=(1,0)$、$e_2=(0,1)$。因为 $(x,y)=xe_1+ye_2$，每个线性自映射均满足

$$T(x,y)=xT(e_1)+yT(e_2).$$

以下几种作用可以直接从这两个像读出。

伸缩 $D(x,y)=(ax,by)$ 把 e_1,e_2 分别送到 ae_1,be_2。当 a,b 都非零时，沿两轴分别除以相应系数即可恢复输入；当 $a=0$ 时，输入的第一坐标完全不影响输出。反射 $F(x,y)=(x,-y)$ 是 $a=1,b=-1$ 的情形，作用两次恢复原向量。

对实参数 s，剪切

$$H_s(x,y)=(x+sy,y)$$

固定 x 轴，保持每个点的高度，把高度为 y 的整条水平线沿水平方向移动 sy。由于 $H_s(e_1)=e_1$、$H_s(e_2)=se_1+e_2$，单位正方形被送到平行四边形。计算

$$H_{-s}(H_s(x,y))=(x+sy-sy,y)=(x,y)$$

说明它可逆。把一个图形倾斜并不必然造成信息丢失。

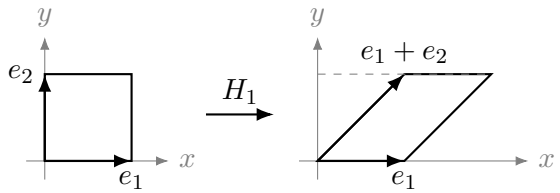


图 7.2: 剪切由两个基向量的像决定。每条水平线整体移动，平行四边形仍保留两条独立方向。

逆时针旋转角 θ 的映射记为 R_θ。几何旋转保留向量的合成和实数倍，故线性。两个单位方向旋转后的坐标为

$$R_\theta(e_1)=(\cos\theta,\sin\theta),\qquad R_\theta(e_2)=(-\sin\theta,\cos\theta).$$

展开任意向量便得

$$R_\theta(x,y)=(x\cos\theta-y\sin\theta,\,x\sin\theta+y\cos\theta).$$

旋转公式由基向量的像推出；其中的正负号记录旋转后两个基方向的实际位置。

7.5.3 函数的变换也服从同一规则

定义 $D:\mathcal{P}_2(\mathbb{R})\to\mathcal{P}_2(\mathbb{R})$ 与 $E:\mathcal{P}_2(\mathbb{R})\to\mathbb{R}$，作用规则为

$$D(p)=p',\qquad E(p)=p(0).$$

这里 p' 为多项式逐项求导得到的函数，$E:\mathcal{P}_2(\mathbb{R})\to\mathbb{R}$ 为零点取值。若 $p(t)=a+bt+ct^2$，则 $D(p)(t)=b+2ct$，$E(p)=a$。逐项相加与数乘说明两者线性。

例如 $p(t)=1+2t+3t^2$ 先求导得到 $2+6t$，再取零点值得到 2。复合 $E\circ D$ 提取原多项式的一次项系数。这与从坐标列中提取某个分量是同一种线性操作，只是输入被写成函数。

与之不同，规则 $Q(p)=p^2$ 一般不线性，因为 $(p+q)^2=p^2+2pq+q^2$ 含有交叉项；而且若固定陪域为 $\mathcal{P}_2(\mathbb{R})$，取 $p(t)=t^2$ 后 p^2 已不在陪域中。给出一个变换时，须同时确定输入对象、输出对象和作用规则。

7.5.4 核与像描述消失和保留的方向

定义 7.23 (核、像、秩与零化度)**.** 对线性映射 $T: V \to W$，定义

$$\ker T = \{v \in V : T(v) = 0\}, \qquad \operatorname{im} T = \{T(v) : v \in V\}.$$

在相关空间有限维时，记 $\operatorname{rank} T = \dim \operatorname{im} T$，$\operatorname{nullity} T = \dim \ker T$，分别称秩与零化度。

命题 7.24 (核与像是子空间)**.** $\ker T \le V$，$\operatorname{im} T \le W$。$T$ 单射当且仅当 $\ker T = \{0\}$。

证明. 核包含零向量；若 $u, v \in \ker T$，则 $T(au + bv) = a0 + b0 = 0$，所以核满足子空间判别。像也含零；若 $x = T(u), y = T(v)$，则 $ax + by = T(au + bv)$ 仍在像中。最后

$$T(u) = T(v) \quad \Longleftrightarrow \quad T(u - v) = 0.$$

若核为零，便有 $u = v$；若 T 单射，$T(u) = 0 = T(0)$ 强制 $u = 0$。 □

这与群同态的核具有同一作用：把任意两个输入的合并问题转成它们的差是否在核中。

定理 7.25 (秩—零化度定理)**.** *若 V 有限维，$T: V \to W$ 线性，则 $\operatorname{im} T$ 有限维且*

$$\dim V = \dim \ker T + \dim \operatorname{im} T.$$

证明. 取 $\ker T$ 的基 $u_1, \dots, u_r$，扩充为 V 的基 $u_1, \dots, u_r, v_1, \dots, v_s$。对任意 $y \in \operatorname{im} T$，选 $x \in V$ 使 $y = T(x)$，写

$$x = \sum_{i=1}^{r} a_i u_i + \sum_{j=1}^{s} b_j v_j.$$

则

$$y = T(x) = \sum_{i=1}^{r} a_i T(u_i) + \sum_{j=1}^{s} b_j T(v_j) = \sum_{j=1}^{s} b_j T(v_j).$$

因此 $T(v_1), \dots, T(v_s)$ 张成像。

若 $\sum_{j=1}^{s} c_j T(v_j) = 0$，则 $T(\sum_j c_j v_j) = 0$，故 $\sum_j c_j v_j \in \ker T$。按核的基展开为 $\sum_j c_j v_j = \sum_i d_i u_i$，得到

$$\sum_{j=1}^{s} c_j v_j - \sum_{i=1}^{r} d_i u_i = 0.$$

这是一组基的线性关系，故 $c_j = 0$、$d_i = 0$。所以像向量组无关，因而为像的基。于是

$$\dim \ker T = r, \quad \dim \operatorname{im} T = s, \quad \dim V = r + s,$$

合并得到所需等式。 □

证明选出的 u_i 负责映射后消失的方向，补入的 v_j 则在像中留下无关的方向。等式来自这组基的划分；不同映射的秩为何会变，也可以由它们合并了哪些方向来理解。

推论 7.26 (同维空间之间的单射与满射)**.** *若 V, W 有限维且 $\dim V = \dim W = n$，则线性映射 $T: V \to W$ 单射、满射、为同构三者等价。*

证明. 单射等价于核维数为零，由秩—零化度等价于像维数为 n。像为 W 的子空间且同维，故必等于 W：若为真子空间，将其基扩充为 W 的基时至少增加一个向量，与同维矛盾。因此单射等价于满射；两者同时成立恰为同构。 □

有限维的条件实际用于这项计数。取全体实多项式的空间 $\mathcal{P}(\mathbb{R})$，映射 $S(p)(t) = tp(t)$ 单射，却不能输出常数多项式 1，所以不满射。逐项求导 D 则满射，因为 $\sum_{j=0}^{d} a_j t^j$ 是 $\sum_{j=0}^{d} a_j t^{j+1}/(j+1)$ 的导数；但所有常数都被送到零，所以不单射。

7.6 商空间、纤维与线性方程的解

设 $U \le V$。如果只关心两个向量相差的部分是否属于 U，定义 $v \sim v'$ 当且仅当 $v - v' \in U$。零向量、取负与加法的封闭性分别给出自反、对称与传递。

定义 7.27 (商向量空间)**.** 此关系的等价类为 $v + U = \{v + u : u \in U\}$，商集记为 V/U，定义

$$(v + U) + (w + U) = (v + w) + U, \qquad a(v + U) = av + U.$$

等价类公式来自 $v' \sim v$ 当且仅当 $v' = v + u$。若 $v' = v + u, w' = w + u'$，则 $(v' + w') - (v + w) = u + u' \in U$，$av' - av = au \in U$，故两种运算都与代表元无关。零类为 U，加法逆类为 $-v + U$。例如，结合律与两个分配律逐式成为

$$\begin{aligned}((u + U) + (v + U)) + (w + U) &= ((u + v) + w) + U = (u + (v + w)) + U,\\ a((v + U) + (w + U)) &= a(v + w) + U = (av + aw) + U,\\ (a + b)(v + U) &= ((a + b)v) + U = (av + bv) + U.\end{aligned}$$

加法交换律来自 $v + w = w + v$；$a(b(v + U)) = (abv) + U = (ab)(v + U)$，且 $1(v + U) = v + U$，所以全部向量空间公理成立。自然投影 $q(v) = v + U$ 满足

$$q(av + bw) = a(v + U) + b(w + U), \qquad q(v) = U \iff v \in U,$$

因而线性且核为 U。

例 7.28 (商空间中的一个元素)**.** 令 $V = \mathbb{R}^2$，$U = \{(x, 0) : x \in \mathbb{R}\}$。则 $(a, b) + U = \{(x, b) : x \in \mathbb{R}\}$，是一整条水平线。同一水平线上的所有点都代表商空间的同一元素。规则 $(a, b) + U \mapsto b$ 与代表元无关，给出 $V/U \to \mathbb{R}$ 的线性同构。

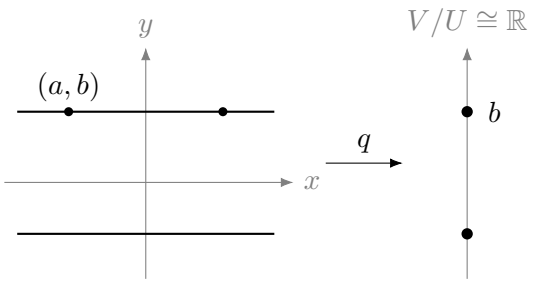


图 7.3: 每条水平线成为商空间的一个元素；纵坐标区分这些元素。

定理 7.29 (线性因子化与第一同构定理)**.** 若 $T : V \to W$ 线性，则它可唯一经过 $q : V \to V/U$ 因子化为线性映射，当且仅当 $U \subseteq \ker T$。特别地，

$$V/\ker T \cong \operatorname{im} T, \qquad v + \ker T \longmapsto T(v).$$

证明. 设 $U \subseteq \ker T$。若 $v + U = v' + U$，则

$$v - v' \in U \quad \Longrightarrow \quad T(v) - T(v') = T(v - v') = 0.$$

所以 $\bar{T}(v + U) = T(v)$ 良定义。并且对 $a, b \in \mathbb{F}$，

$$\begin{aligned} \bar{T}\big(a(v+U) + b(w+U)\big) &= \bar{T}((av + bw) + U) \\ &= T(av + bw) = aT(v) + bT(w) \\ &= a\bar{T}(v+U) + b\bar{T}(w+U). \end{aligned}$$

因此 $\bar{T}$ 线性，且 $\bar{T}q = T$。若另有 $Lq = T$，则 $L(v + U) = Lq(v) = T(v) = \bar{T}(v + U)$，故唯一。反过来，若 $T = Lq$，则对 $u \in U$，$T(u) = L(0 + U) = 0$，故 $U \subseteq \ker T$。

现在取 $U = \ker T$，并将陪域取为 $\operatorname{im} T$。每个 $T(v)$ 都等于 $\bar{T}(v + \ker T)$，故满射。又

$$\bar{T}(v+\ker T) = \bar{T}(w+\ker T) \iff T(v-w) = 0 \iff v-w \in \ker T \iff v+\ker T = w+\ker T,$$

所以单射。其逆的线性性由下式验证：对于 $y = \bar{T}(x), z = \bar{T}(x')$，

$$\bar{T}^{-1}(ay + bz) = \bar{T}^{-1}(\bar{T}(ax + bx')) = ax + bx' = a\bar{T}^{-1}(y) + b\bar{T}^{-1}(z).$$

□

对 $w \in \operatorname{im} T$，取一个原像 v_0，则

$$T(v) = w \quad \Longleftrightarrow \quad T(v - v_0) = 0 \quad \Longleftrightarrow \quad v \in v_0 + \ker T.$$

因此一个非空纤维是核的平移。在线性方程中，这就是“全部解等于一个特解加上全部齐次解”的原因。

7.7 直和：分解是否唯一

对 $U, W \leq V$，属于 $U + W$ 的向量至少能写成一份 $u + w$。若有 $u + w = u' + w'$，则

$$u - u' = w' - w \in U \cap W.$$

所以分解的歧义恰好来自两子空间的交。

定义 7.30 (内部直和)**.** 若 $U \cap W = \{0\}$，称 $U + W$ 为 U, W 的直和，记为 $U \oplus W$。等价地，$U + W$ 中每个向量的分解 $u + w$ 唯一。

等价性的一方向由上式成立；反向若有非零 $z \in U \cap W$，就有 $z = z + 0 = 0 + z$ 两份不同的分解。

对两个任意向量空间，笛卡尔积 $U \times W$ 按分量运算也为向量空间，称为外部直和，亦记 $U \oplus W$。当 U, W 是同一空间的子空间且交为零时，$(u, w) \mapsto u + w$ 给出外部直和到内部直和的线性同构。无限族的直积允许任意坐标，直和则只允许有限个非零坐标；这与线性组合只能取有限和的约定一致。

7.7.1 直和对应一对投影

若 $V = U \oplus W$，每个 v 唯一写成 $v = u + w$，因此可定义

$$P(v) = u, \qquad Q(v) = w.$$

由两份分解相加、数乘仍为相应向量的分解，以及分解的唯一性，可知 P, Q 线性。它们满足

$$P + Q = \mathrm{id}_V, \quad P^2 = P, \quad Q^2 = Q, \quad PQ = QP = 0.$$

其中 $P^2 = P$ 表示已经取出 U 分量后，再取一次不改变结果；$PQ = 0$ 表示纯粹的 W 分量没有 U 分量。

命题 7.31 (投影与直和分解)**.** *若线性自映射 P 满足 $P^2 = P$，则*

$$V = \operatorname{im} P \oplus \ker P.$$

证明. 对任意 v，写成 $v = P(v) + (v - P(v))$。前一项属于像，后一项因为 $P(v - P(v)) = P(v) - P^2(v) = 0$ 而属于核。若 $x \in \operatorname{im} P$，则 $x = P(y)$，从而 $P(x) = P^2(y) = x$；若它同时属于核，就有 $x = P(x) = 0$。故两子空间交为零，和为整个空间。 □

这样的映射称为投影。子空间 U 本身尚不能唯一指定投影，还要决定沿哪个补空间丢弃分量。例如在 $\mathbb{R}^2$ 中令 U 为 x 轴：

$$P_1(x, y) = (x, 0), \qquad P_2(x, y) = (x - y, 0).$$

两者均满足 $P_i^2 = P_i$，且像都为 U。但 $\ker P_1$ 为 y 轴，$\ker P_2$ 为直线 $x = y$；对 $(1, 1)$，前者输出 $(1, 0)$，后者输出零。后面加入内积后，沿 $U^{\perp}$ 的投影会由垂直条件唯一确定。

7.8 矩阵：把映射写成坐标运算

选定基以后，一个向量可以由坐标列表示；线性映射则由基向量的像决定。把这些像的坐标列排在一起，就得到矩阵。

定义 7.32 (矩阵表示)**.** 设 $\mathcal{B} = (v_1, \dots, v_n)$、$\mathcal{C} = (w_1, \dots, w_m)$ 分别为 V, W 的有序基。对线性映射 $T : V \to W$，写

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

标量的矩形数组 $A = (a_{ij})$ 称为 T 关于两组基的矩阵，记 $[T]_{\mathcal{C} \leftarrow \mathcal{B}}$。全体 m 行 n 列矩阵组成的集合记为 $M_{m \times n}(\mathbb{F})$。

若 $v = \sum_j x_j v_j$，则

$$T(v) = \sum_j x_j \sum_i a_{ij} w_i = \sum_i \left(\sum_j a_{ij} x_j \right) w_i.$$

因此像的第 i 个坐标为 $\sum_j a_{ij} x_j$。这给出矩阵作用于坐标列的规则，简记为 $[T(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}}$。矩阵第 j 列记录第 j 个输入基向量被送到哪里。

若另有 $S:U\to V$，矩阵为 $B=(b_{jk})$，将 S 的输出坐标代入上式，得到复合的第 i,k 项为 $\sum_j a_{ij}b_{jk}$。所以定义矩阵乘法

$$(AB)_{ik}=\sum_j a_{ij}b_{jk},$$

便得到 $[T\circ S]=[T][S]$，其中中间空间使用同一组基。乘法的求和公式因此来自复合，且乘法顺序与映射复合顺序一致。

7.8.1 同一乘法的分量、列与复合解释

若 A 的列为 $a_1,\dots,a_n$，则矩阵作用可以写成

$$Ax=x_1a_1+\cdots+x_na_n.$$

以

$$A=\begin{pmatrix}2&1\\1&-1\end{pmatrix},\qquad x=\begin{pmatrix}3\\2\end{pmatrix}$$

为例，按行计算为 $(2\cdot3+1\cdot2,\,1\cdot3-1\cdot2)^t=(8,1)^t$；按列计算为

$$3\begin{pmatrix}2\\1\end{pmatrix}+2\begin{pmatrix}1\\-1\end{pmatrix}=\begin{pmatrix}8\\1\end{pmatrix}.$$

按行计算逐个输出坐标，按列计算则把基向量的像按输入系数组合起来。因此 $Ax=b$ 也可理解为：列向量能否用系数 $x_1,\dots,x_n$ 组合出 b？有解等价于 b 在列的张成中；解唯一等价于这些列没有非平凡的零组合。

设 B 为 n 行 p 列，列向量为 $b_1,\dots,b_p$。逐列作用便有

$$AB=\begin{pmatrix}Ab_1&\cdots&Ab_p\end{pmatrix}.$$

这同时解释尺寸要求：B 输出 n 个坐标，恰好供 A 接收。若中间尺寸不一致，两个坐标映射无法复合。

注 7.33 (Strang 的按列理解)**.** Strang 在 [11] 第 1.4 节，尤其第 27–29 页，并列使用逐项的行乘列计算与整列的线性组合。他用前者计算各个数，用后者理解整个乘法。本例中的两种算法只是求和次序不同；列的解释还直接说明列空间、解的存在性和矩阵复合。

7.8.2 顺序怎样影响变换

取剪切 $H(x,y)=(x+y,y)$ 与伸缩 $D(x,y)=(2x,y)$。在标准基中其矩阵分别为

$$H=\begin{pmatrix}1&1\\0&1\end{pmatrix},\qquad D=\begin{pmatrix}2&0\\0&1\end{pmatrix}.$$

这里以同一字母表示变换和其标准基矩阵。先 H 后 D 得

$$DH=\begin{pmatrix}2&2\\0&1\end{pmatrix},\qquad (D\circ H)(x,y)=(2x+2y,y);$$

先 D 后 H 得

$$HD=\begin{pmatrix}2&1\\0&1\end{pmatrix},\qquad (H\circ D)(x,y)=(2x+y,y).$$

第一种顺序中，剪切增加的水平位移也随后被放大；第二种顺序中，增加位移发生在伸缩之后。矩阵乘法不交换在此记录了两种不同的作用过程。

结合律仍然成立，因为 $(T\circ S)\circ R=T\circ(S\circ R)$，两边都将输入依次经过 R,S,T。于是矩阵满足 $(AB)C=A(BC)$；加法分配律则来自映射对输出向量之和的线性性。

7.8.3 函数变换的矩阵计算

仍取 $\mathcal{P}_2(\mathbb{R})$ 的基 $\mathcal{B}=(1,t,t^2)$，把求导视为自映射 $D:\mathcal{P}_2(\mathbb{R})\to\mathcal{P}_2(\mathbb{R})$。由

$$D(1)=0,\qquad D(t)=1,\qquad D(t^2)=2t$$

逐列得到

$$[D]_{\mathcal{B}\leftarrow\mathcal{B}}=\begin{pmatrix}0&1&0\\0&0&2\\0&0&0\end{pmatrix}.$$

取零点值的矩阵为 $[E]=(1\ \ 0\ \ 0)$，因为 $E(1)=1$、$E(t)=E(t^2)=0$。所以

$$[E\circ D]=(1\ \ 0\ \ 0)\begin{pmatrix}0&1&0\\0&0&2\\0&0&0\end{pmatrix}=(0\ \ 1\ \ 0).$$

这正是提取一次项系数的规则。再算

$$[D]^2=\begin{pmatrix}0&0&2\\0&0&0\\0&0&0\end{pmatrix},\qquad [D]^3=0.$$

任意二次多项式求导三次都为零，所以最后的零矩阵也可在计算之前由变换的作用预见。

7.8.4 秩给出独立输入输出基下的最简形式

命题 7.34 (矩形矩阵的秩标准形)**.** *设 $T:V\to W$ 线性，$\dim V=n$、$\dim W=m$，且 $\operatorname{rank}T=r$。可分别选择输入与输出的基，使 T 的矩阵为*

$$\begin{pmatrix}I_r&0\\0&0\end{pmatrix},$$

其中前 r 个输入坐标被原样送到前 r 个输出坐标，其余输入坐标被送到零；I_r 为 r 阶单位矩阵。零尺寸的块允许省去。

证明. 按秩—零化度定理的证明，先取 $\ker T$ 的基，再补入向量 $v_1,\dots,v_r$，使这些补入向量的像构成 $\operatorname{im}T$ 的基。将输入基重新排序，把 $v_1,\dots,v_r$ 放在前面。输出基则先放 $T(v_1),\dots,T(v_r)$，再将它们扩充为 W 的基。前 r 个基向量的像就是前 r 个输出基向量；剩下的输入基向量在核中。逐列记录即得矩阵。 □

此处允许输入、输出独立选基。若 $V=W$ 且要求两端使用同一组基，就不能任意实施这项化简；同一个自映射在反复作用时，输出还要作为下一次输入，必须保留两端坐标之间的关系。

命题 7.35 (可逆矩阵)**.** *方阵 A 可逆，指存在方阵 B 使 $AB=BA=I$，其中 I 为对角项为 1、其余项为 0 的单位矩阵。A 可逆当且仅当其线性映射可逆；此时 $B=A^{-1}$ 为逆映射的矩阵。*

证明. I 表示恒等映射，矩阵乘法表示复合，因此两个矩阵等式恰好表示映射互为逆。逆映射唯一，故逆矩阵也唯一。线性双射的逆线性已在上一节证明。 □

7.9 换基：输入与输出坐标的转换

定义 7.36 (过渡矩阵)**.** 同一空间的两组有序基 $\mathcal{B},\mathcal{B}'$ 之间，过渡矩阵 P 的各列为新基向量在旧基中的坐标，因而 $[v]_{\mathcal{B}}=P[v]_{\mathcal{B}'}$。反向转换提供它的逆矩阵。

若值域的过渡矩阵为 $[w]_{\mathcal{C}}=Q[w]_{\mathcal{C}'}$，则新坐标先经 P 转成旧输入坐标，经过旧矩阵 A，再经 Q^{-1} 转成新输出坐标。因此

$$[T]_{\mathcal{C}'\leftarrow\mathcal{B}'}=Q^{-1}[T]_{\mathcal{C}\leftarrow\mathcal{B}}P.$$

当 $T:V\to V$ 且输入输出同时更换同一组基时，新矩阵为 $P^{-1}AP$。这个公式表达的是同一个映射在不同坐标下的描述。

7.9.1 把换基公式逐步用于一个向量

取标准基 $\mathcal{E}=(e_1,e_2)$ 与 $\mathcal{B}=(b_1,b_2)$，其中 $b_1=(1,1)$、$b_2=(1,-1)$。过渡矩阵和逆为

$$P=\begin{pmatrix}1&1\\1&-1\end{pmatrix},\qquad P^{-1}=\frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}.$$

直接相乘得 $P^2=2I$，所以逆矩阵公式成立。令 $T(x,y)=(x,2y)$，标准基矩阵为 $A=\begin{pmatrix}1&0\\0&2\end{pmatrix}$。在新基下，

$$AP=\begin{pmatrix}1&1\\2&-2\end{pmatrix},\qquad A_{\mathcal{B}}=P^{-1}AP=\begin{pmatrix}3/2&-1/2\\-1/2&3/2\end{pmatrix}.$$

可逐列核对：

$$T(b_1)=(1,2)=\tfrac{3}{2}b_1-\tfrac{1}{2}b_2,\qquad T(b_2)=(1,-2)=-\tfrac{1}{2}b_1+\tfrac{3}{2}b_2.$$

于是新矩阵的列确实记录同一个 T 对新基向量的作用。

再取 $v=(2,1)$，其新坐标为 $(3/2,1/2)^{t}$。计算

$$A_{\mathcal{B}}\begin{pmatrix}3/2\\1/2\end{pmatrix}=\begin{pmatrix}2\\0\end{pmatrix},\qquad P\begin{pmatrix}2\\0\end{pmatrix}=\begin{pmatrix}2\\2\end{pmatrix}=[T(v)]_{\mathcal{E}}.$$

新坐标中的输出 $(2,0)^{t}$ 表示 $2b_1$，恢复到标准坐标才是 $(2,2)^{t}$。同样的两数字写在不同基的下标后面，可以指代不同的向量。

整个过程由交换图表示：

$$\begin{array}{ccc} [v]_{\mathcal{B}} & \xrightarrow{A_{\mathcal{B}}} & [T(v)]_{\mathcal{B}} \\ {\scriptstyle P}\downarrow & & \downarrow{\scriptstyle P} \\ [v]_{\mathcal{E}} & \xrightarrow[A]{} & [T(v)]_{\mathcal{E}}. \end{array}$$

沿两条路到达同一坐标列，给出 $PA_{\mathcal{B}} = AP$，再左乘 P^{-1} 就得到换基公式。

7.9.2 向量的移动与坐标的改变

固定标准基时，规则 $x \mapsto Px$ 可以被看成一个真正的线性变换，它把 e_1 送到 b_1、e_2 送到 b_2。而在等式 $[v]_{\mathcal{E}} = P[v]_{\mathcal{B}}$ 中，P 用来转换同一个向量的两份坐标，向量 v 本身没有移动。矩阵相同，其在论证中的角色由输入和输出的含义决定。

对于自映射，关系 $A_{\mathcal{B}} = P^{-1}AP$ 称为矩阵相似。它依次完成新坐标译为旧坐标、执行变换、输出译回新坐标。若输入、输出分别采用不同的坐标变换 P, Q，则是 $Q^{-1}AP$。在选择矩阵前先写清两端的坐标关系，便能确定每个逆矩阵的位置。

与换基不同，单独左乘 E 得 EA，在固定坐标下表示原变换之后又作用一次 E；它通常改变原自映射的特征值。因而解方程所用的行消元与研究同一变换所用的相似换基，需要依据各自保持的对象来使用。

7.10 消元、主元与线性方程的计算

定义 7.37 (线性方程组与初等变换)**.** 矩阵方程 $Ax = b$ 表示每行给出的线性方程。初等行变换包括：交换两行；以非零标量乘一行；将一行的标量倍加到另一行。每一种都有逆操作，故对应左乘一个可逆矩阵，称为初等矩阵。

要保持方程的解，应同时对系数矩阵与右端列进行同一行变换：$Ax = b$ 等价于 $EAx = Eb$。它的逆变换来自 E^{-1}。列变换则改变输入坐标；若 $x = Py$ 且 P 可逆，原方程成为 $APy = b$，最后需换回 x。

例 7.38 (从计算回到核与纤维)**.** 取 $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$，令 $x = (x_1, x_2, x_3)^t$，$b = (b_1, b_2)^t$。第二行给出 $x_2 = b_2 - x_3$，代入第一行得 $x_1 = b_1 - 2b_2 + x_3$。所以全部解为

$$x = (b_1 - 2b_2, b_2, 0)^t + t(1, -1, 1)^t, \qquad t \in \mathbb{R}.$$

任意 $b \in \mathbb{R}^2$ 均有解，所以像为 $\mathbb{R}^2$；令 $b = 0$ 得核为 $\operatorname{span}((1, -1, 1)^t)$。于是秩为 2，零化度为 1，与定义域维数 3 一致。非齐次解集是一条核的平移直线。

7.10.1 消元为什么终止，主元如何记录自由度

对一行不全为零的矩阵，最左非零项称为该行的首项。行阶梯形要求零行都在底部，且每一非零行的首项位置严格位于上一行首项的右侧。这些首项称为主元，所在列称为主元列；其余列称为自由列。

从最左列开始，若当前列在尚未处理的行中有非零项，就将相应行交换到当前首行，以非零数缩放使主元为 1，再减去该行的倍数，把主元下方各项消为零。随后只处理右下方尚未处

理的子矩阵。每形成一个主元，可用的行数和列数各减少至少一，故有限步后得到阶梯形。再从下向上消去主元上方的项，就得到简化行阶梯形：每个主元为 1，主元列其余项为零。

这里每一步除法都只发生在已选定的非零主元上，因此在任意域上均可执行。若某列没有可选的非零项，算法直接跳到右边的列；它不要求每一列都出现主元。

例 7.39 (相容条件从最后一行出现)**.** 考虑任意右端 $b=(b_1,b_2,b_3)^t$ 的方程

$$\begin{pmatrix}1&1&1\\2&2&2\\1&2&3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}.$$

记 R_i 为增广矩阵的第 i 行，竖线右侧为右端列。先作 $R_2\leftarrow R_2-2R_1$、$R_3\leftarrow R_3-R_1$，再交换后两行，得到

$$\left(\begin{array}{ccc|c}1&1&1&b_1\\0&1&2&b_3-b_1\\0&0&0&b_2-2b_1\end{array}\right).$$

因此有解的充要条件为 $b_2=2b_1$。右端若不满足此条件，最后一行要求 0 等于一个非零数；它明确指出原来第二个方程必须是第一个的两倍。

条件成立时，再将第二行从第一行减去：

$$\left(\begin{array}{ccc|c}1&0&-1&2b_1-b_3\\0&1&2&b_3-b_1\\0&0&0&0\end{array}\right).$$

主元列是第 $1,2$ 列。令自由变量 $z=s\in\mathbb{R}$，得到

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}2b_1-b_3\\b_3-b_1\\0\end{pmatrix}+s\begin{pmatrix}1\\-2\\1\end{pmatrix}.$$

每个自由变量都可任取，每个主元变量随后由本行唯一确定，所以本例解集有一个自由参数。取 $b=0$ 得到核的一组基 $(1,-2,1)^t$。

7.10.2 行操作保持哪些结构

若 $R=EA$ 且 E 可逆，则 $Rx=0$ 等价于 $Ax=0$，故核不变。每一新行是原行的线性组合，所以新行的张成包含在旧行的张成中；利用 $A=E^{-1}R$ 又得反向包含，故行空间也不变。这里行空间指所有行向量的线性组合所成的空间。

列空间却被 E 映为另一子空间，满足 $\operatorname{im}R=E(\operatorname{im}A)$，通常不与 $\operatorname{im}A$ 相等。例如上例原矩阵所有输出都满足第二坐标等于第一坐标的两倍，化简后所有输出都满足第三坐标为零。这两个平面不同，但经 E 线性同构，维数相同。

命题 7.40 (主元与列空间的基)**.** *若行消元将 A 化为简化行阶梯形 R，有 r 个主元，则 A 中对应的 r 个主元列构成 $\operatorname{im}A$ 的一组基；$\operatorname{rank}A=r$。非主元列的个数等于 $\dim\ker A$。*

证明. R 的主元列为前 r 个标准坐标向量，每个其余列都是它们的线性组合。又因 $R=EA$ 且 E 可逆，一份列间关系

$$c_1A_1+\cdots+c_nA_n=0$$

成立当且仅当相同系数的关系 $\sum_j c_j R_j = 0$ 成立，其中 A_j, R_j 为第 j 列。因此 R 中主元列的无关性与对其余列的张成关系都传回 A，得到第一项。对齐次方程，每个非主元变量可自由指定，主元变量由化简方程唯一决定。这给出解空间到 $\mathbb{F}^{n-r}$ 的线性双射，所以核的维数为 $n-r$。 □

寻找原列空间的基时，消元确定列的编号，向量仍须从原矩阵取出。上例主元列编号为 $1,2$，故可取原列 $(1,2,1)^t$、$(1,2,2)^t$ 作为基。

7.10.3 把原矩阵分解为独立列与组合系数

上例第三列满足 $A_3 = -A_1 + 2A_2$。所以

$$A = \underbrace{\begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 1 & 2 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}}_{R}.$$

矩阵 C 收集选出的列空间基，R 的第 j 列记录原第 j 列相对于这组基的坐标。这就是分解 $A = CR$。任意向量的作用被拆为

$$\mathbb{R}^3 \xrightarrow{R} \mathbb{R}^2 \xrightarrow{C} \mathbb{R}^3.$$

前一步先计算两个有效系数，后一步再用两个独立输出方向重建结果。C 单射，故 $\ker A = \ker R$；R 满射，故 $\operatorname{im} A = \operatorname{im} C$。两步分别把核和像的结构显示出来。

Strang 在 [11] 第 1.4 节用一般的 $A = CR$ 组织这项观察：挑出独立列后，其余列只需记录组合系数。它把列的几何解释与行消元的计算联系起来。

7.11 对偶空间：用标量测量向量

固定一组基时，可以从向量取出第一个坐标，也可以取几个坐标的线性组合。这些取值规则本身是线性映射，它们共同组成一个新的向量空间。

定义 7.41 (线性映射空间与对偶空间)**.** 所有线性映射 $V \to W$ 的集合记为 $\operatorname{Hom}_{\mathbb{F}}(V,W)$，按逐点相加和数乘成为向量空间。记 $\operatorname{End}(V) = \operatorname{Hom}_{\mathbb{F}}(V,V)$。对偶空间定义为 $V^* = \operatorname{Hom}_{\mathbb{F}}(V,\mathbb{F})$，其元素称为线性泛函。

线性泛函的输出是一个标量。例如 $\lambda : \mathbb{R}^2 \to \mathbb{R}, \lambda(x,y) = 2x - y$ 是线性的。在未给定额外结构时，一个向量与一个泛函是不同类型的对象。

命题 7.42 (对偶基)**.** *设 $\mathcal{B} = (v_1, \dots, v_n)$ 为 V 的基。定义 $v_i^*(\sum_j a_j v_j) = a_i$，则 $(v_1^*, \dots, v_n^*)$ 为 V^* 的基。因此 $\dim V^* = \dim V$。*

证明. 坐标唯一性保证 v_i^* 良定义，线性组合的坐标逐项相加保证它线性。对任意泛函 λ 以及 $v = \sum_j a_j v_j$，有

$$\lambda(v) = \sum_j a_j \lambda(v_j) = \left(\sum_j \lambda(v_j) v_j^* \right)(v).$$

因此这些泛函张成 V^*。若 $\sum_i c_i v_i^* = 0$，将等式作用于 v_j 得 $c_j = 0$，故无关。 □

定义 7.43 (对偶映射)**.** 对线性映射 $T : V \to W$，定义 $T^* : W^* \to V^*$ 为 $T^*(\lambda) = \lambda \circ T$。

箭头方向反转，因为先从 V 映到 W，才能使用 W 上的泛函测量。对 $S : U \to V$，结合律给出 $(T \circ S)^* = S^* \circ T^*$。

若 $T(v_j) = \sum_i a_{ij} w_i$，在对偶基中

$$T^*(w_i^*)(v_j) = w_i^*(T(v_j)) = a_{ij},$$

所以 $T^*(w_i^*) = \sum_j a_{ij} v_j^*$。其矩阵的第 j, i 项等于原矩阵第 i, j 项，即转置矩阵 A^t。转置的指标交换来自这项配对计算。

7.11.1 对偶基怎样实际求出

仍取 $\mathbb{R}^2$ 中的 $b_1 = (1, 1)$、$b_2 = (1, -1)$。由

$$(x, y) = \frac{x + y}{2} b_1 + \frac{x - y}{2} b_2$$

读出两个对偶基泛函：

$$b_1^*(x, y) = \frac{x + y}{2}, \qquad b_2^*(x, y) = \frac{x - y}{2}.$$

逐项代入得 $b_1^*(b_1) = b_2^*(b_2) = 1$、$b_1^*(b_2) = b_2^*(b_1) = 0$。它们分别提取新基下的第一、第二坐标。

一般地，若 P 的列为某组基在标准基中的坐标，则新坐标为 $P^{-1}x$；因此 P^{-1} 的各行就是对偶基泛函在标准坐标中的系数。基向量按列编码，对偶基按逆矩阵的行编码，两者由“先组合，再提取系数”必须恢复原坐标这一要求联系。

7.11.2 转置是把输出端的测量拉回输入端

取 $T : \mathbb{R}^2 \to \mathbb{R}^3$ 为

$$T(x, y) = (x + 2y,\, 3x - y,\, y), \qquad A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \\ 0 & 1 \end{pmatrix}.$$

在输出端取泛函 $\lambda(u, v, w) = 2u - v + 4w$。先作用 T 再测量，得到

$$T^*(\lambda)(x, y) = 2(x + 2y) - (3x - y) + 4y = -x + 9y.$$

用行向量表示泛函系数时，是 $(2, -1, 4)A = (-1, 9)$；若把同一份系数竖写为列，就是

$$A^t \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 9 \end{pmatrix}.$$

原映射把输入向量推到输出空间，对偶映射把输出空间的一项线性测量变成输入空间的测量。维数从 $2 \to 3$ 改为 $3 \to 2$，也解释了为什么转置后尺寸反转。

7.11.3 用泛函描述子空间

对 $U \le V$，定义它的零化子

$$U^\circ = \{\lambda \in V^* : \lambda(u) = 0 \text{ 对所有} u \in U\}.$$

它是 V^* 的子空间，因为对 U 恒为零的泛函的任意线性组合仍对 U 恒为零。

命题 7.44 (零化子的维数)**.** 若 $\dim V = n$、$\dim U = r$，则 $\dim U^\circ = n - r$。此外，对 $v \in V$，有

$$v \in U \quad \Longleftrightarrow \quad \lambda(v) = 0 \text{ 对所有} \lambda \in U^\circ.$$

证明. 取 U 的基 $v_1, \dots, v_r$，扩充为 V 的基 $v_1, \dots, v_n$。泛函 $\lambda = \sum_i c_i v_i^*$ 在 U 上为零当且仅当前 r 个系数为零，所以 $v_{r+1}^*, \dots, v_n^*$ 为 U° 的基。若 $v \in U$，右侧由定义成立；若 $v \notin U$，它至少有一个编号 $j > r$ 的坐标非零，此时 $v_j^* \in U^\circ$ 却有 $v_j^*(v) \neq 0$，故右侧不成立。 □

于是一个 r 维子空间可以由 $n - r$ 个独立线性条件刻画。用基描述它，是列出所有可组合的方向；用零化子描述它，是列出所有必须满足的齐次条件。

对线性映射 $T : V \to W$，定义直接给出

$$\ker T^* = (\operatorname{im} T)^\circ,$$

因为 $\lambda \circ T = 0$ 恰表示 λ 在所有输出上为零。有限维时，由零化子的维数和秩—零化度得到

$$\operatorname{rank} T^* = \dim W^* - \dim(\operatorname{im} T)^\circ = \dim W - (\dim W - \operatorname{rank} T) = \operatorname{rank} T.$$

在坐标中，T^* 的矩阵为 A^t，所以矩阵的行秩与列秩相等。这个相等来自像与其线性约束之间的关系。

7.12 行列式：交替多线性与可逆性

给定一个 n 维坐标空间中的 n 个列向量，我们希望有一个标量表达它们的整体退化程度。要求对每列分别线性，使每列的线性组合能够直接展开；要求两列相同时结果为零，体现重复方向的退化；最后规定标准基的值为 1 来固定尺度。

定义 7.45 (多线性与交替性)**.** 函数 $D : (\mathbb{F}^n)^n \to \mathbb{F}$ 称为多线性，若固定其余变量时，它对每一个变量都线性。若任意两个变量相等时函数值为 0，称它为交替的。

交替多线性函数交换两个变量会变号。固定其余变量，把 $u + v$ 同时放入这两个位置，函数值为零；展开四项，其中同为 u 或同为 v 的两项为零，剩下两项之和为零，即得变号规律。

定义 7.46 (排列符号与行列式)**.** 对置换 $\sigma \in S_n$，令 $N(\sigma)$ 为满足 $i < j$ 且 $\sigma(i) > \sigma(j)$ 的指标对数，定义 $\operatorname{sgn}(\sigma) = (-1)^{N(\sigma)}$。对方阵 $A = (a_{ij})$ 定义

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)},$$

称为行列式。

为使用这个符号，先验证它与置换复合相容。

引理 7.47 (排列符号的乘法性)**.** 对 $\sigma,\tau\in S_n$，有

$$\operatorname{sgn}(\sigma\tau)=\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau),\qquad \operatorname{sgn}(\sigma^{-1})=\operatorname{sgn}(\sigma).$$

每个对换的符号为 -1。

证明. 记 $s_i=(i\ i+1)$。右复合 s_i 交换排列中相邻位置的值 $\sigma(i),\sigma(i+1)$。这两个值与其他位置形成的逆序对总数不变：位于二者左侧或右侧的位置，分别仍与这两个值比较一次。二者之间的逆序对则从有变无或从无变有，因此

$$N(\sigma s_i)=N(\sigma)\pm 1,\qquad \operatorname{sgn}(\sigma s_i)=-\operatorname{sgn}(\sigma).$$

若一个排列不是递增排列，则至少有一对相邻位置逆序；否则每对相邻值递增，传递性使整个排列递增，只能是恒等排列。每次交换一个相邻逆序使逆序数减少 1，故经过恰 $N(\tau)$ 次这类交换可把 τ 化为恒等排列。将这些交换逆序排列，得到 $\tau=s_{i_1}\cdots s_{i_m}$，其中 $m=N(\tau)$。于是

$$\operatorname{sgn}(\sigma\tau)=(-1)^m\operatorname{sgn}(\sigma)=\operatorname{sgn}(\tau)\operatorname{sgn}(\sigma).$$

取 $\tau=\sigma^{-1}$，得两者符号的积为 1；它们均为 ±1，故相等。若 $i<j$，则

$$(i\ j)=s_is_{i+1}\cdots s_{j-2}s_{j-1}s_{j-2}\cdots s_i$$

是 $2(j-i)-1$ 个相邻对换的乘积，因子数为奇数，所以符号为 -1。 □

上述公式给出 $\det\begin{pmatrix}a&b\\c&d\end{pmatrix}=ad-bc$；零阶空矩阵的行列式约定为 1。

定理 7.48 (行列式的刻画与乘法性)**.** 行列式是唯一的关于各列交替多线性、且在单位矩阵上取值 1 的函数。它满足 $\det(AB)=\det A\det B$，且 A 可逆当且仅当 $\det A\neq0$。

证明. 记标准基为 $e_1,\dots,e_n$，矩阵 A 的第 j 列为 $c_j=\sum_i a_{ij}e_i$。若 D 是任意交替多线性函数，令 $d=D(e_1,\dots,e_n)$，则逐列展开得到

$$D(c_1,\dots,c_n)=\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}\left(\prod_{j=1}^{n}a_{i_j,j}\right)D(e_{i_1},\dots,e_{i_n}).$$

有重复指标的项由交替性为零。其余指标唯一构成某个排列 τ，即 $i_j=\tau(j)$。按上一引理的相邻交换方法把这些基向量重排，每交换一次 D 变号，故

$$D(e_{\tau(1)},\dots,e_{\tau(n)})=\operatorname{sgn}(\tau)d.$$

因此

$$D(c_1,\dots,c_n)=d\sum_{\tau\in S_n}\operatorname{sgn}(\tau)\prod_j a_{\tau(j),j}=d\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_i a_{i,\sigma(i)},$$

最后一个等号将 τ 换为 σ^{-1}，并用逆排列符号相同及乘法交换律。所以规定 $d=1$ 后，D 必等于排列公式，证明唯一性。

再验证存在性。排列公式的每项对每列恰取一个系数，所以对该列线性。若第 j,k 两列相同，令 $t=(j\ k)$，将 σ 项与 $t\sigma$ 项配对。二者无固定配对项，因为 $t\sigma=\sigma$ 将迫使 t 为恒等；其系数乘积相同，而

$$\operatorname{sgn}(t\sigma)=-\operatorname{sgn}(\sigma).$$

故每对之和为零，在特征 2 时也成立。单位矩阵中，若 $\sigma \neq \mathrm{id}$，至少一个 $a_{i,\sigma(i)} = 0$；只有恒等排列项为 1，所以值为 1。

固定矩阵 A，令

$$D_A(b_1, \dots, b_n) = \det(Ab_1, \dots, Ab_n).$$

A 的线性性和行列式的多线性性保证 D_A 多线性；两输入相同则两输出列相同，故它交替。刚才对任意 D 的展开公式给出

$$\det(AB) = D_A(b_1, \dots, b_n) = D_A(e_1, \dots, e_n) \det B = \det A \det B.$$

若 A 可逆，则 $1 = \det I = \det A \det A^{-1}$，所以 $\det A \neq 0$。若 A 不可逆，秩—零化度说明 $Ax = 0$ 有非零解，即列向量相关。取一个非零系数解出相应列是其余列的组合，按该列展开，每一项都有重复列，行列式为零。取逆否即得非零行列式保证可逆。 □

命题 7.49 (Laplace 展开)**.** *若 A_{ij} 是删除 A 的第 i 行、第 j 列所得矩阵，则*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}.$$

证明. 在排列公式中按 $\sigma(i) = j$ 将项分类。固定该选择后，其余行列的排列给出 $\det A_{ij}$。将第 i 行与第 j 列移到首位分别用 $i-1, j-1$ 次交换，故相对符号为 $(-1)^{i+j-2} = (-1)^{i+j}$。对各 j 求和得式。 □

7.12.1 平面面积怎样进入行列式

在实平面中，令 $u = (a, c)$、$v = (b, d)$ 为矩阵 A 的两列。它们张成的平行四边形面积为

$$\mathrm{Area}(u, v) = |ad - bc| = |\det A|.$$

可以用底乘高核对：若 $u \neq 0$，底长为 $\sqrt{a^2 + c^2}$，垂直于底边的单位向量为 $(-c, a)/\sqrt{a^2 + c^2}$。向量 v 在这个方向上的有向分量为 $(ad - bc)/\sqrt{a^2 + c^2}$；取绝对值得到高，再乘底长。若 $u = 0$，面积与行列式都为零。

$\det A$ 的符号还区分了有序两列的取向：交换两列时，面积不变而符号反转。对任意两列 u, v，将它们排成矩阵 B，则变换后排成 AB，故

$$\det(Au, Av) = \det A \ \det(u, v).$$

这里 $\det(u, v)$ 表示以这两列组成的行列式。因此同一个线性变换对每个平行四边形的面积都乘以 $|\det A|$，与原平行四边形的位置、边长无关。

剪切 H_s 的行列式为 1。它虽然改变斜边，却不改变对应底边的高度，所以面积保持。投影 $(x, y) \mapsto (x, 0)$ 的行列式为零，整个正方形被压到一条线；反射 $(x, y) \mapsto (x, -y)$ 的行列式为 -1，面积保持而取向反转。

7.12.2 从交替多线性回到实际计算

对列向量的三种操作可直接求得：交换两列，行列式变号；一列乘 c，行列式乘 c；把一列的 c 倍加到另一列，行列式不变。最后一项因为

$$\det(\dots, u + cv, \dots, v, \dots) = \det(\dots, u, \dots, v, \dots) + c \det(\dots, v, \dots, v, \dots),$$

第二项含重复列而为零。

排列公式还给出 $\det A^t = \det A$：将求和指标 σ 改为 σ^{-1}，乘积各因子仅重排，且二者符号相同。所以相同三条规则也适用于行。上三角矩阵的行列式为对角线元素的积：在排列公式中，任何非零项必须逐行选择不低于对角线的位置，这迫使各行依次选择自己的对角项。

例如令

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 3 \end{pmatrix}.$$

先将第二行减去第一行的两倍，再将第三行减去新的第二行，得到

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

两步均不改变行列式，故 $\det A = 1 \cdot 1 \cdot 2 = 2$。若中间交换行或缩放行，须同时记录它们引入的符号与倍数。消元求行列式使用的是这些已由定义证明的变换规律。

7.13 特征向量：映射保留哪些方向

一个自映射 $T : V \to V$ 可以反复作用，记 $T^0 = \mathrm{id}_V$、$T^{k+1} = T \circ T^k$。若某个子空间 U 满足 $T(U) \subseteq U$，就称 U 为 T 的不变子空间。此时从 U 出发的所有迭代都留在 U 中，限制 $T|_U$ 本身成为 U 上的自映射。

最小的非零子空间是一条直线 $U = \mathrm{span}(v)$，其中 $v \neq 0$。它不变当且仅当 $T(v) \in \mathrm{span}(v)$，也就是存在标量 λ 使 $T(v) = \lambda v$。因而特征向量的问题可以从寻找一维不变子空间自然产生。

注 7.50 (Axler 与 Strang 的两个出发点)**.** Axler 在 [10] 第 5A 节、第 133–134 页，先提出把自映射限制到不变子空间，再由一维情形引入特征值。Strang 在 [11] 第 6.1 节、第 216–218 页，则从反复作用 $A, A^2, \ldots$ 出发：若某个方向只发生数乘，反复作用就变成反复乘同一个标量。前者解释寻找哪些子空间，后者解释找到以后怎样计算迭代。

定义 7.51 (特征值与特征向量)**.** 设 $T : V \to V$ 线性。若 $v \neq 0$ 且 $T(v) = \lambda v$，称 v 为属于特征值 $\lambda \in \mathbb{F}$ 的特征向量。等价地，非零向量 v 属于 $\ker(T - \lambda \mathrm{id}_V)$。

特征向量所在的直线在映射下不会离开自身；$\lambda = 0$ 时，这条直线被送到零。零向量必须排除，因为它对所有 λ 都满足等式，不能区分任何方向或数值。

命题 7.52 (不同特征值的特征向量无关)**.** *若 $v_1, \ldots, v_r$ 分别属于两两不同的特征值 $\lambda_1, \ldots, \lambda_r$，则它们线性无关。*

证明. 对 r 归纳。$r = 1$ 时由 $v_1 \neq 0$ 成立。设结论对 $r-1$ 成立，并有 $\sum_i c_i v_i = 0$。作用 $T - \lambda_r \mathrm{id}_V$ 得

$$\sum_{i=1}^{r-1} c_i (\lambda_i - \lambda_r) v_i = 0.$$

由归纳假设，各系数为零；因 $\lambda_i - \lambda_r \neq 0$，故 $c_i = 0$ 对 $i < r$ 成立。代回原式得 $c_r v_r = 0$，故 $c_r = 0$。 □

定义 7.53 (特征多项式)**.** 有限维时，取 T 的矩阵 A，定义 $\chi_T(t)=\det(tI-A)$。

换基后矩阵为 $P^{-1}AP$，且 $tI-P^{-1}AP=P^{-1}(tI-A)P$；乘法性给出相同行列式，所以此多项式与所选基无关。λ 为特征值当且仅当 $T-\lambda\mathrm{id}_V$ 的核非零，当且仅当 $\chi_T(\lambda)=0$。

例 7.54 (标量域影响特征值)**.** 在 $\mathbb{R}^2$ 上取 $T(x,y)=(-y,x)$。若 $T(v)=\lambda v$ 且 $v\neq 0$，再作用 T 得 $-v=\lambda^2 v$，强制 $\lambda^2=-1$，无实数解。因此它没有实特征值。把相同矩阵视为 $\mathbb{C}^2$ 上的映射，则特征多项式 t^2+1 有根 $i,-i$。特征值须属于指定的标量域。

7.13.1 对角化把迭代拆成独立的标量计算

定义 7.55 (特征子空间与可对角化)**.** 对标量 λ，令 $E_\lambda=\ker(T-\lambda\mathrm{id}_V)$，称为相应特征子空间。其非零元素恰为属于 λ 的特征向量。若 V 有一组完全由特征向量组成的基，称 T 可对角化。

这时取基 $\mathcal{B}=(v_1,\dots,v_n)$ 且 $T(v_i)=\lambda_i v_i$，矩阵为对角阵 $D=\mathrm{diag}(\lambda_1,\dots,\lambda_n)$，其中 diag 表示给定数为对角项，其余项为零。若原矩阵为 A，以特征向量为列的过渡矩阵为 P，则

$$AP=PD,\qquad A=PDP^{-1}.$$

反复相乘时，相邻的 $P^{-1}P$ 消去，故

$$A^k=PD^kP^{-1},\qquad T^k\left(\sum_i c_i v_i\right)=\sum_i c_i\lambda_i^k v_i.$$

最后一式也可直接对 k 归纳。选特征基后，每个坐标沿自己的方向独立演化。

例 7.56 (一次完整的对角化)**.** 取

$$A=\begin{pmatrix}2&1\\1&2\end{pmatrix},\qquad \chi_A(t)=(t-2)^2-1=(t-3)(t-1).$$

当 $\lambda=3$，方程 $(A-3I)v=0$ 等价于 $x=y$；当 $\lambda=1$，等价于 $x=-y$。故取 $v_1=(1,1)^t$、$v_2=(1,-1)^t$，它们无关，构成基。于是

$$P=\begin{pmatrix}1&1\\1&-1\end{pmatrix},\quad D=\begin{pmatrix}3&0\\0&1\end{pmatrix},\quad A^k=\frac{1}{2}\begin{pmatrix}3^k+1&3^k-1\\3^k-1&3^k+1\end{pmatrix}.$$

例如 $e_1=\frac{1}{2}v_1+\frac{1}{2}v_2$，因此

$$A^k e_1=\tfrac{1}{2}\,3^k v_1+\tfrac{1}{2}v_2.$$

沿 v_1 的分量每次放大三倍，沿 v_2 的分量保持。矩阵幂公式中的各项来自这两个分量的合成。

7.13.2 特征值存在为何还不够

剪切矩阵 $H=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ 满足 $\chi_H(t)=(t-1)^2$。但

$$(H-I)\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\0\end{pmatrix}=0$$

强制 $y=0$，所以 E_1 只有一维。无法从这条直线上取出两个无关向量，故 H 不可对角化。两个计重数的特征根，并不自动提供两个独立的特征方向。

仍可从作用规则求幂：$H(x,y)=(x+y,y)$，每次将不变的 y 加到第一坐标，故

$$H^k(x,y)=(x+ky,y),\qquad H^k=\begin{pmatrix}1&k\\0&1\end{pmatrix}.$$

这个例子也可写成 $H=I+N$，其中 $N(x,y)=(y,0)$ 且 $N^2=0$；直接归纳得 $(I+N)^k=I+kN$。所有特征值都为 1，但某些向量仍产生随 k 增长的位移。

若 $\dim V=n$ 且有 n 个不同的特征值，先前的无关性定理保证对应特征向量构成基，因此一定可对角化。重复特征值的情形则须实际计算各特征子空间，判断能否凑成一组基。

7.14 双线性型、内积与正交分解

对偶中的一个泛函接受一个向量并输出标量。如果要比较两个向量，可以使用两个变量的标量函数；要求它对每个变量分别线性，就得到双线性型。

定义 7.57 (双线性型与二次型)**.** 实向量空间 V 上的双线性型为映射 $B:V\times V\to\mathbb{R}$，对两个变量分别线性。若 $B(u,v)=B(v,u)$，称为对称双线性型。函数 $q(v)=B(v,v)$ 称为由它确定的二次型。

固定基 $v_1,\dots,v_n$，令 $a_{ij}=B(v_i,v_j)$，按两个变量分别展开，得到 $B(u,v)=[u]^tA[v]$，其中坐标均关于该基。对称性等价于 $a_{ij}=a_{ji}$。若改用过渡矩阵 P，由代入可得新矩阵为 P^tAP；它与线性自映射的换基公式不同，因为两个输入都需要变换。

命题 7.58 (实对称型的极化)**.** *实对称双线性型由其二次型唯一确定，且*

$$B(u,v)=\frac{1}{2}\big(q(u+v)-q(u)-q(v)\big).$$

证明. 展开 $q(u+v)=B(u,u)+B(u,v)+B(v,u)+B(v,v)$，再用对称性将中间两项合为 $2B(u,v)$，移项并除以 2 即得。 □

条件中“对称”与“实数”各有作用：反对称部分在 $B(v,v)$ 中可能完全消失；在特征为 2 的域中也不能除以 2。

定义 7.59 (内积、范数与正交)**.** 实向量空间上的内积是正定的对称双线性型，记为 $\langle u,v\rangle$；正定指 $v\neq0$ 时 $\langle v,v\rangle>0$。定义范数 $\|v\|=\sqrt{\langle v,v\rangle}$。若 $\langle u,v\rangle=0$，称二者正交，记 $u\perp v$。对子空间 W，定义正交补

$$W^{\perp}=\{v\in V:\forall w\in W,\ \langle v,w\rangle=0\}.$$

标准例为 $\mathbb{R}^n$ 上的 $\langle x,y\rangle=\sum_i x_iy_i$。双线性型 $B(x,y)=x_1y_1-x_2y_2$ 也对称，却不正定；非零向量 $(1,1)$ 满足 $B(v,v)=0$，因此不能作为内积。

命题 7.60 (Cauchy–Schwarz 不等式)**.** $|\langle u,v\rangle|^2\le\langle u,u\rangle\langle v,v\rangle$。

证明. 若 $v=0$，两边为零。否则取 $t=\langle u,v\rangle/\langle v,v\rangle$，由正定性及展开得

$$0\le\langle u-tv,u-tv\rangle=\langle u,u\rangle-\frac{\langle u,v\rangle^2}{\langle v,v\rangle}.$$

乘以正数 $\langle v,v\rangle$ 得结论。 □

选取这个 t，就是从 u 中减去沿 v 的部分，使余项与 v 正交。这项计算在构造正交基时会反复出现。

定理 7.61 (有限维正交分解)**.** *若 V 为有限维实内积空间，$W \le V$，则 $V = W \oplus W^{\perp}$。*

证明. 首先 $W^{\perp}$ 是子空间：零向量显然在内；若 $x, y \in W^{\perp}$、$a, b \in \mathbb{R}$，则对 $w \in W$，

$$\langle ax + by, w\rangle = a\langle x, w\rangle + b\langle y, w\rangle = 0,$$

所以 $ax + by \in W^{\perp}$。若 $W = 0$，则 $W^{\perp} = V$，结论成立。

否则取 W 的基 $v_1, \dots, v_r$。递归定义

$$z_k = v_k - \sum_{i<k} \langle v_k, u_i\rangle u_i, \qquad u_k = z_k / \|z_k\|.$$

用归纳证明每步可行，并且 $u_1, \dots, u_k$ 正交归一，与 $v_1, \dots, v_k$ 张成相同子空间。$k = 1$ 时 $z_1 = v_1 \neq 0$。若前 $k-1$ 步已成立，则对 $j < k$，

$$\begin{aligned}\langle z_k, u_j\rangle &= \langle v_k, u_j\rangle - \sum_{i<k} \langle v_k, u_i\rangle \langle u_i, u_j\rangle \\ &= \langle v_k, u_j\rangle - \langle v_k, u_j\rangle = 0.\end{aligned}$$

若 $z_k = 0$，则 $v_k \in \operatorname{span}(u_1, \dots, u_{k-1}) = \operatorname{span}(v_1, \dots, v_{k-1})$，违背原基无关，所以 $z_k \neq 0$。正定性给出 $\|z_k\| > 0$，归一化合法，且 $\langle u_k, u_k\rangle = 1$。两条表示式

$$u_k = \|z_k\|^{-1}\left(v_k - \sum_{i<k}\langle v_k, u_i\rangle u_i\right), \qquad v_k = \|z_k\| u_k + \sum_{i<k}\langle v_k, u_i\rangle u_i$$

给出两组张成空间的双向包含，完成归纳。这就是 Gram–Schmidt 正交化。

对 $v \in V$，令 $w = \sum_{i=1}^{r} \langle v, u_i\rangle u_i \in W$。对每个 j，

$$\langle v - w, u_j\rangle = \langle v, u_j\rangle - \sum_{i} \langle v, u_i\rangle \langle u_i, u_j\rangle = 0.$$

任意 $x \in W$ 可写为 $x = \sum_j c_j u_j$，故 $\langle v - w, x\rangle = \sum_j c_j \langle v - w, u_j\rangle = 0$。于是 $v - w \in W^{\perp}$，得到 $V = W + W^{\perp}$。

若 $x \in W \cap W^{\perp}$，取正交条件中的 $w = x$，得 $\langle x, x\rangle = 0$，由正定性 $x = 0$。因此 $W \cap W^{\perp} = 0$，和为直和。 □

映射 $v \mapsto w$ 称为到 W 的正交投影。对任意 $w' \in W$，$v - w$ 与 $w - w'$ 正交，因而

$$\|v - w'\|^2 = \|v - w\|^2 + \|w - w'\|^2 \ge \|v - w\|^2.$$

故投影同时给出了距离 v 最近的 W 中元素；等号只有在 $w' = w$ 时成立。

7.14.1 内积在非正交基下怎样计算

固定一组基 $\mathcal{B} = (b_1, \dots, b_n)$，令

$$G = (g_{ij}), \qquad g_{ij} = \langle b_i, b_j\rangle,$$

称 G 为这组基的 Gram 矩阵。若 $u=\sum_i x_i b_i$、$v=\sum_j y_j b_j$，由双线性展开得

$$\langle u,v\rangle=\sum_{i,j} x_i g_{ij} y_j = x^t G y.$$

只有在基正交归一时 $G=I$，坐标列的普通点积才直接等于原空间的内积。

例如在标准实平面中取 $b_1=(1,0)$、$b_2=(1,1)$，则

$$G=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}.$$

向量 $u=b_2$ 的新坐标为 $(0,1)^t$，其长度平方为 $(0,1)G(0,1)^t=2$，并非 $0^2+1^2=1$。同时 b_1,b_2 的坐标列普通点积为零，但原向量内积为 1，所以它们并不正交。

若标准坐标下内积矩阵为 I，且新基过渡矩阵为 P，则

$$\langle Px,Py\rangle = x^t P^t P y,$$

故新 Gram 矩阵为 P^tP。这也具体解释了双线性型的换基公式：同一测量规则同时接收两个向量，两份坐标都要代入。

7.14.2 正交化的一次完整计算

在标准内积的 $\mathbb{R}^3$ 中取

$$v_1=(1,1,0),\qquad v_2=(1,0,1).$$

它们无关，但内积为 1，没有正交。先取 $u_1=v_1/\sqrt{2}$，再从 v_2 减去沿 u_1 的投影：

$$z_2=v_2-\langle v_2,u_1\rangle u_1=(1,0,1)-\tfrac{1}{2}(1,1,0)=(1/2,-1/2,1).$$

因 $\|z_2\|^2=3/2$，归一化得

$$u_2=\frac{(1,-1,2)}{\sqrt{6}}.$$

直接计算可核对 $\langle u_1,u_2\rangle=0$、$\|u_1\|=\|u_2\|=1$。操作中只减去先前方向的倍数，所以

$$\operatorname{span}(u_1,u_2)=\operatorname{span}(v_1,v_2).$$

正交化改变了基，没有改变所描述的平面。

7.14.3 正交矩阵是保持内积的坐标变换

定义 7.62 (正交矩阵)**.** 实方阵 Q 若满足 $Q^tQ=I$，称为正交矩阵。

矩阵 Q^tQ 的第 i,j 项为 Q 的第 i 列与第 j 列的内积，所以此条件等价于列构成正交归一基。并且对所有 x,y，

$$\langle Qx,Qy\rangle = x^tQ^tQy = x^ty=\langle x,y\rangle.$$

反向若保持所有内积，代入标准基向量就得 $Q^tQ=I$。因此正交矩阵准确描述保持标准内积的线性变换，既保持长度也保持夹角。

由 $Q^tQ=I$，若 $Qx=0$ 则 $x=0$；方阵的秩—零化度保证其可逆，于是 $Q^{-1}=Q^t$。旋转矩阵满足此条件，反射矩阵也满足；两者都不改变长度。行列式满足 $(\det Q)^2=1$，所以 $\det Q=1$ 或 -1，分别对应保持或反转取向。

7.14.4 从投影到最小二乘方程

设实矩阵 A 有 m 行 n 列，列线性无关。对给定 $b\in\mathbb{R}^m$，即使 $Ax=b$ 无解，仍可寻找使误差长度 $\|Ax-b\|$ 最小的 x。因为所有 Ax 恰组成 $\operatorname{im}A$，这个问题等价于求 b 在列空间上的正交投影。

若 $\widehat{x}$ 是候选解，残差记为 $r=b-A\widehat{x}$。正交投影的条件为 r 与每列 A_j 都正交，即

$$A^t(b-A\widehat{x})=0.$$

因此须解正规方程

$$A^tA\widehat{x}=A^tb.$$

矩阵 A^tA 可逆：若 $A^tAx=0$，左乘 x^t 得

$$0=x^tA^tAx=\|Ax\|^2,$$

所以 $Ax=0$；由列无关得 $x=0$。方阵零核蕴含可逆，故

$$\widehat{x}=(A^tA)^{-1}A^tb,\qquad \Pi=A(A^tA)^{-1}A^t$$

分别给出唯一最小二乘解与到列空间的正交投影矩阵。这里唯一性还用了 A 单射：唯一投影点只能由一个输入得到。

为直接检验最小性，任取 $h\in\mathbb{R}^n$。因 $Ah\in\operatorname{im}A$ 且 $r\perp\operatorname{im}A$，

$$\|A(\widehat{x}+h)-b\|^2=\|Ah-r\|^2=\|Ah\|^2+\|r\|^2\geq\|r\|^2.$$

等号要求 $Ah=0$，由单射性要求 $h=0$。

例 7.63 (一条直线拟合三个点)**.** 用直线 $y=c+dt$ 拟合平面上的三点 $(0,1),(1,2),(2,2)$。精确通过三个点等价于

$$\underbrace{\begin{pmatrix}1&0\\1&1\\1&2\end{pmatrix}}_{A}\begin{pmatrix}c\\d\end{pmatrix}=\underbrace{\begin{pmatrix}1\\2\\2\end{pmatrix}}_{b}.$$

第一点要求 $c=1$，第二点再要求 $d=1$，却使第三点的预测为 3，所以无精确解。正规方程为

$$\begin{pmatrix}3&3\\3&5\end{pmatrix}\begin{pmatrix}c\\d\end{pmatrix}=\begin{pmatrix}5\\6\end{pmatrix}.$$

第二行减第一行得 $2d=1$，故 $d=1/2$、$c=7/6$。预测向量与残差分别为

$$A\widehat{x}=(7/6,5/3,13/6)^t,\qquad r=(-1/6,1/3,-1/6)^t.$$

残差的坐标和为零，与 $(0,1,2)^t$ 的内积也为零，所以它确与 A 的两列正交。平方误差之和为 $1/6$；上面的正交分解证明任何其他直线都不会给出更小值。

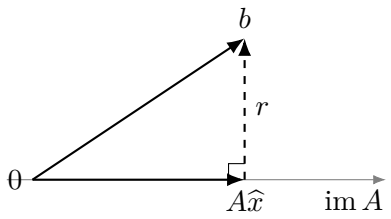


图 7.4: 在列空间与残差方向张成的平面内看投影：$b = A\hat{x} + r$，其中 $r \perp \operatorname{im} A$。

7.15 四个基本子空间与解方程的整体结构

对实 $m \times n$ 矩阵 A，有四个相互联系的子空间：

子空间	所在空间	维数
$\operatorname{im} A^t$（行空间）	$\mathbb{R}^n$	r
$\ker A$	$\mathbb{R}^n$	$n-r$
$\operatorname{im} A$（列空间）	$\mathbb{R}^m$	r
$\ker A^t$（左零空间）	$\mathbb{R}^m$	$m-r$

其中 $r = \operatorname{rank} A = \operatorname{rank} A^t$，后一等式已由对偶映射证明。“左零空间”的名称来自 $A^t y = 0$ 等价于 $y^t A = 0$。

定理 7.64 (行列空间的正交分解)**.** *在标准实内积下，*

$$(\operatorname{im} A^t)^\perp = \ker A, \qquad (\operatorname{im} A)^\perp = \ker A^t.$$

因此

$$\mathbb{R}^n = \operatorname{im} A^t \oplus \ker A, \qquad \mathbb{R}^m = \operatorname{im} A \oplus \ker A^t,$$

两式均为正交直和。

证明. $x \perp \operatorname{im} A^t$ 当且仅当对每个 $y \in \mathbb{R}^m$，

$$0 = \langle x, A^t y\rangle = \langle Ax, y\rangle.$$

若 $Ax = 0$，此条件成立；若对所有 y 成立，取 $y = Ax$ 得 $\|Ax\|^2 = 0$，所以 $Ax = 0$。这证明第一式。将 A 换为 A^t 得第二式。有限维正交分解定理随即给出两项直和。 □

输入向量因此可唯一写成 $x = x_{\text{row}} + x_{\text{null}}$，分别属于行空间与核。作用 A 后，后一项消失：

$$Ax = Ax_{\text{row}}.$$

而 A 在行空间上的限制给出 $\operatorname{im} A^t \to \operatorname{im} A$ 的线性同构。它单射，因为行空间与核的交为零；它满射，因为任意原像去掉核分量后仍有相同输出。这样，四个空间说明哪些输入差别被消去、哪些方向保留为独立输出，以及哪些右端分量无法由输入产生。

命题 7.65 (相容性的对偶判别)**.** *方程 $Ax = b$ 有解，当且仅当对每个满足 $A^t y = 0$ 的 y，都有 $y^t b = 0$。*

证明. 有解当且仅当 $b \in \operatorname{im} A$。正交分解说明，这又等价于 b 与 $\ker A^t$ 正交，正是所述条件。 □

这个判别把所有右端限制集中到左零空间中。若某个 y 满足 $y^tA=0$ 却有 $y^tb\neq 0$，用 y_i 作为系数把原方程的各行相加，就得到 $0=y^tb$ 的矛盾。

例 7.66 (回看消元中的相容条件)**.** 对前面的矩阵

$$A=\begin{pmatrix}1&1&1\\2&2&2\\1&2&3\end{pmatrix}$$

已知列空间基为 $(1,2,1)^t,(1,2,2)^t$，核基为 $(1,-2,1)^t$。行空间可取原第 $1,3$ 行作为基；左零空间基为 $(-2,1,0)^t$。分别计算内积可验证两个正交关系：

$$(1,1,1)\cdot(1,-2,1)=0,\qquad (1,2,3)\cdot(1,-2,1)=0;$$

列空间中的每个向量也都与 $(-2,1,0)^t$ 正交。因而相容性恰为

$$(-2,1,0)\cdot(b_1,b_2,b_3)=0,\quad \text{即}b_2=2b_1.$$

这正是消元所得最后一行。计算算法发现的限制，与子空间之间的正交关系相同。

注 7.67 (Strang 组织四个空间的方法)**.** Strang 在 [11] 第 3.5 节把列空间、行空间和两个零空间放在同一张维数表中。这样研究一个矩阵时，解方程与理解变换可以相互回看：核给出解的不唯一性，列空间给出可实现的右端，左零空间给出右端必须满足的约束；行空间则为每一份可实现输出提供唯一的正交于核的输入。

若 $Ax=b$ 相容，从全部解中取行空间内的唯一解 x_{row}，其他解均为 $x_{\mathrm{row}}+z$，其中 $z\in\ker A$。正交性给出

$$\|x_{\mathrm{row}}+z\|^2=\|x_{\mathrm{row}}\|^2+\|z\|^2.$$

所以它还恰好是长度最小的解。若原方程不相容，先把 b 正交投影到列空间，再取这一投影的行空间原像，就得到全部最小二乘解中长度最小的一个。

7.16 仿射子空间与凸组合

线性方程 $T(v)=w$ 的非空解集为 $v_0+\ker T$，它通常不包含零向量。要描述这样的平移，需要保留线性结构而允许改变原点。

定义 7.68 (仿射子空间与仿射组合)**.** 实向量空间中，子空间 U 的平移 $p+U=\{p+u:u\in U\}$ 称为仿射子空间。若 $a_0+\cdots+a_k=1$，则 $\sum_{i=0}^{k}a_ix_i$ 称为点 $x_0,\dots,x_k$ 的仿射组合。

系数和为 1 保证平移的相容性：同时将所有点平移 p，得到

$$\sum_i a_i(x_i+p)=\sum_i a_ix_i+\left(\sum_i a_i\right)p=\sum_i a_ix_i+p.$$

因此输出也恰好平移 p。若 $x_i=p+u_i$ 同属 $p+U$，仿射组合仍为 $p+\sum_i a_iu_i$，所以留在该仿射子空间中。

定义 7.69 (凸组合与凸集)**.** 若仿射组合还满足各 $a_i\geq 0$，称为凸组合。集合 C 称为凸集，若任意 $x,y\in C$ 以及 $t\in[0,1]$ 均满足 $(1-t)x+ty\in C$。这要求两点之间的整条线段留在集合中。

命题 7.70 (有限凸组合与凸包)**.** 凸集包含其任意有限点组的所有凸组合。对集合 S，全体有限凸组合组成的集合记为 $\operatorname{conv}(S)$，称为凸包；它是包含 S 的最小凸集。空集的凸包为空集。

证明. 对点数归纳。设 $\sum_{i=0}^{k} a_i x_i$ 为凸组合。若 $a_k = 1$，结果为 x_k；否则 $1 - a_k > 0$，将表达式写为

$$(1-a_k)\left(\sum_{i=0}^{k-1} \frac{a_i}{1-a_k} x_i\right) + a_k x_k.$$

括号内为较少点的凸组合，按归纳假设在凸集中，再用二点凸性得到结论。

S 的每个点都是自身系数为 1 的组合，故 $S \subseteq \operatorname{conv}(S)$。将两个有限凸组合再作二点凸组合，展开后系数仍非负且和为 1，所以凸包为凸集。任何包含 S 的凸集都由前一结论包含这些组合，故得到最小性。 □

7.17 子空间的维数公式

命题 7.71. *若 $U, W \le V$ 均有限维，则*

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W).$$

若 V 有限维且 $U \le V$，则 $\dim(V/U) = \dim V - \dim U$。

证明. 定义 $S : U \times W \to U + W$ 为 $S(u, w) = u + w$。它为满线性映射。核的元素满足 $w = -u$，所以核恰为 $\{(x, -x) : x \in U \cap W\}$；规则 $x \mapsto (x, -x)$ 是交到核的线性同构。

将 U 的基写成 $(u_i, 0)$、W 的基写成 $(0, w_j)$，合起来就是 $U \times W$ 的基，故其维数为 $\dim U + \dim W$。对 S 使用秩—零化度得第一式。第二式对满线性投影 $q : V \to V/U$ 应用同一定理，其核为 U。 □

第一式扣除了两个子空间共同的方向；第二式则计算把 U 中的差别合并以后还保留多少独立方向。两式均来自已经建立的核与像的维数关系。

7.18 综合算例

由线性方程构造基与补空间

设 $U = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$。求它的一组基及其在 $\mathbb{R}^3$ 中的一个补空间。由 $x + y + z = 0$ 得 $z = -x - y$，因而

$$(x, y, z) = x(1, 0, -1) + y(0, 1, -1).$$

令 $u_1 = (1, 0, -1)$、$u_2 = (0, 1, -1)$。上式说明二者张成 U；若 $a u_1 + b u_2 = 0$，比较前两个坐标得 $a = b = 0$，所以它们也是无关的，构成 U 的一组基。

取 $n = (1, 1, 1)$。若 $a u_1 + b u_2 + c n = 0$，将三个坐标相加，得到 $3c = 0$，故 $c = 0$，再由无关性得 $a = b = 0$。这三个向量构成 $\mathbb{R}^3$ 的基，故 $W = \operatorname{span}\{n\}$ 满足 $\mathbb{R}^3 = U \oplus W$。还可以直接写出分解：对 $v = (x, y, z)$，令 $c = (x + y + z)/3$，则

$$v = (v - cn) + cn, \qquad v - cn \in U, \quad cn \in W.$$

若有另一个分解，两式相减后，一边属于 U，另一边属于 W；而 $tn \in U$ 蕴含 $3t = 0$，所以交仅为零，分解唯一。标准内积下 n 与 U 中每个向量正交，因此这次选出的补空间恰好为 $U^{\perp}$。

差分映射与商空间

设 $T:\mathbb{R}^3\to\mathbb{R}^2$ 为 $T(x,y,z)=(x-y,y-z)$。计算其核、像及对应的商空间同构。等式 $T(x,y,z)=0$ 等价于 $x=y=z$，所以 $\ker T=\operatorname{span}\{(1,1,1)\}$。给定任意 $(a,b)\in\mathbb{R}^2$，向量 $(a+b,b,0)$ 被送到 (a,b)，故 T 为满射。定义

$$\overline{T}:\mathbb{R}^3/\ker T\longrightarrow\mathbb{R}^2,\qquad [(x,y,z)]\longmapsto(x-y,y-z).$$

若两个代表元相差 (t,t,t)，它们的相邻坐标差相同，所以此规则良定义。反过来，相邻坐标差相同就意味着两向量的三个坐标差彼此相等，因此它们属于同一等价类。这证明了单射性；满射性由上面的原像构造得到。映射的线性由 T 的线性直接推出。

逆映射为 $(a,b)\mapsto[(a+b,b,0)]$。为核对两个复合，令 $a=x-y$、$b=y-z$，则该规则得到代表元 $(x-z,y-z,0)$，它与 (x,y,z) 相差 $(z,z,z)\in\ker T$，因而确实返回原来的等价类。商空间中的一个向量可以有许多代表元，逆映射只需返回同一等价类。

多项式的平移

在次数至多为 2 的实多项式空间 $\mathcal{P}_2(\mathbb{R})$ 上，设 $S(p)(t)=p(t+1)$。以下在基 $(1,t,t^2)$ 下计算矩阵、迭代与逆，并判断可对角化性。由 $S(1)=1$、$S(t)=1+t$、$S(t^2)=1+2t+t^2$，逐列得到

$$[S]=\begin{pmatrix}1&1&1\\0&1&2\\0&0&1\end{pmatrix}.$$

迭代 k 次就是把变量平移 k，可由归纳证明 $S^k(p)(t)=p(t+k)$。因此

$$[S]^k=\begin{pmatrix}1&k&k^2\\0&1&2k\\0&0&1\end{pmatrix},\qquad [S]^{-1}=\begin{pmatrix}1&-1&1\\0&1&-2\\0&0&1\end{pmatrix}.$$

逆规则为 $p(t)\mapsto p(t-1)$。上三角矩阵的特征多项式为 $(\lambda-1)^3$；解 $([S]-I)(a,b,c)^t=0$ 得 $2c=0$、$b+c=0$，所以 $b=c=0$。其特征子空间只有常数多项式构成的一维空间，故不可对角化。二次项的平移会向一次项、常数项贡献新系数，矩阵中的 k 与 k^2 正记录这些贡献。

秩亏方程组的一般解与最短解

令 $A=\begin{pmatrix}1&1&0\\0&0&1\\1&1&1\end{pmatrix}$，$b=(b_1,b_2,b_3)^t$。计算四个基本子空间，再讨论 $Ax=b$ 的相容性及解。矩阵的前两列相等，第三列与第一列无关，故列空间的基为 $(1,0,1)^t$、$(0,1,1)^t$，秩为 2。核的基为 $(1,-1,0)^t$，行空间的基为 $(1,1,0)$、$(0,0,1)$。左零空间的基为 $(-1,-1,1)^t$，所以相容条件是 $b_3=b_1+b_2$。

条件成立时，方程只要求 $x_1+x_2=b_1$、$x_3=b_2$，一般解为

$$x=(s,b_1-s,b_2)^t,\qquad s\in\mathbb{R}.$$

它的长度平方可以配方为

$$s^2+(b_1-s)^2+b_2^2=2(s-b_1/2)^2+b_1^2/2+b_2^2.$$

唯一最小值在 $s = b_1/2$ 处取得，最短解为 $(b_1/2, b_1/2, b_2)^t$。它恰与核向量 $(1, -1, 0)^t$ 正交，因此也正是位于行空间中的那个解。

文献：向量空间、基、线性映射、对偶、内积与特征值见 [10] 第 1–3、5–7 章；矩阵乘法、四个基本子空间与迭代见 [11] 第 1.4、3.5、6.1 节；行列式、多线性与双线性型另见 [8]。正文中署名的注概述所引节段的讲法，具体例子与推导使用本文的记号。

8 环、理想与多项式

8.1 为什么需要环

整数可以相加、相减、相乘，却不能对任意非零整数作除法。向量空间恰好相反：非零标量都有逆元，因此线性方程可以除以系数。很多代数问题既需要加法结构，又需要乘法结构，却不能假定所有元素可逆。环把这两种运算放在同一对象中。

本章的环总是结合且带单位元；除非特别说明，环同态保持单位元。允许零环 $1=0$，因为商环可能把原环的单位与零合并。环的定义先保留非交换情形；从理想一节起，固定讨论交换环。

定义 8.1 (环)**.** 环是一个集合 R，带有加法 $+$、乘法 $\cdot$ 和指定元素 1_R，满足：$(R,+)$ 是交换群；乘法结合；对任意 $a,b,c\in R$，

$$a(b+c)=ab+ac,\qquad (a+b)c=ac+bc;$$

并且 $1_Ra=a=a1_R$。若 $ab=ba$ 对所有 a,b 成立，称 R 为交换环。

加法群的单位记为 0。由分配律，$0a=a0=0$：例如 $0a=(0+0)a=0a+0a$，在加法群中消去一个 $0a$ 即得。又有 $(-a)b=-(ab)=a(-b)$。这些基本恒等式不是额外公理，而是由加法群与分配律推出的。

定义 8.2 (单位与零因子)**.** 若 $u\in R$ 存在 $v\in R$ 使 $uv=vu=1_R$，称 u 为单位，v 为 u 的逆，记全体单位为 $R^\times$。若 $a\neq 0, b\neq 0$ 且 $ab=0$，称 a,b 为一对零因子。满足 $1_R\neq 0_R$ 的交换环若没有非零零因子，称为整环；整环中每个非零元素都有乘法逆元时称为域。

命题 8.3 (单位构成群)**.** $R^\times$ *在环的乘法下是群。*

证明. 1_R 是单位。若 u,v 为单位，则 $(uv)^{-1}=v^{-1}u^{-1}$，因为

$$(uv)(v^{-1}u^{-1})=u(vv^{-1})u^{-1}=1_R$$

且反向乘积同样为 1_R。若 u 为单位，逆元由 $uv=vu=1_R$ 唯一确定：若 w 也为逆，则 $w=w(uv)=(wu)v=v$。结合律由环继承。 □

例 8.4 (三个基本环)**.** $\mathbb{Z}$ 是交换整环但不是域，因 2 没有整数逆元。$\mathbb{Q},\mathbb{R},\mathbb{C}$ 都是域。若 $n\geq 2$，整数模 $\mathbb{Z}/n\mathbb{Z}$ 是交换环；当且仅当 n 为素数时它是域。$\mathbb{Z}/6\mathbb{Z}$ 中 $[2]\neq[0]$、$[3]\neq[0]$，但 $[2][3]=[0]$，所以它有零因子。

上述素数判别可直接由整数的整除理论证明。若 $n=p$ 为素数，非零类 $[a]$ 满足 $p\nmid a$，故 $\gcd(a,p)=1$。取 Bézout 等式 $ua+vp=1$，在模 p 下得到 $[u][a]=[1]$，所以每个非零类可逆。若 n 合成，写 $n=ab$，其中 $1<a,b<n$，则 $[a],[b]$ 均非零而乘积为零；域中非零元素可逆，不可能出现这样的乘积。因此 $\mathbb{Z}/n\mathbb{Z}$ 为域当且仅当 n 为素数。

8.2 子环、理想与由生成元产生的结构

定义 8.5 (子环)**.** 环 R 的子环是子集 $S \subseteq R$，含有 1_R，并且对加法、加法逆元和乘法封闭。若不要求含有 1_R，则称为非单位子环；本章的“子环”遵循含单位的约定。

命题 8.6 (子环判别)**.** *非空子集 $S \subseteq R$ 是子环，当且仅当对任意 $a, b \in S$ 有 $a-b \in S$、$ab \in S$，并且 $1_R \in S$。*

证明. 若 S 是子环，条件直接成立。反过来，取 $a \in S$，则 $0 = a - a \in S$；再取 $0 - a = -a \in S$。差封闭给出加法封闭，乘法封闭给出乘法封闭，单位元已假定，其他环公理从 R 继承。 □

以下至本章结束，R, S 均为交换有单位环。

定义 8.7 (理想)**.** 交换环 R 的理想是一个加法子群 $I \leq (R, +)$，满足对任意 $r \in R, i \in I$ 有 $ri \in I$。在非交换环中，若只要求左乘吸收，称左理想；若只要求右乘吸收，称右理想。

理想通常不含 1_R。若 $1_R \in I$，则任意 $r = r1_R \in I$，所以 $I = R$。这正是理想能够作为“被压成零的元素集合”的原因：它必须在任意外部系数乘法下保持不变，而不应被迫包含整个环的单位。

命题 8.8 (理想的判别与交)**.** *非空子集 $I \subseteq R$ 是理想，当且仅当对 $i, j \in I, r \in R$ 有 $i - j \in I$ 且 $ri \in I$。任意一族理想的交仍是理想。*

证明. 第一句的证明与子环判别相同，差封闭给出加法单位和逆元。设 $(I_\lambda)_{\lambda \in \Lambda}$ 是理想族。若 x, y 属于交集，则它们属于每个 I_λ，所以 $x - y$ 也属于每个 I_λ；对任意 $r \in R$，rx 同样属于每个 I_λ。因此交集是理想。空族的交在 R 中取为 R，它也是理想。 □

定义 8.9 (生成的理想)**.** 给定 $S \subseteq R$，包含 S 的所有理想的交称为 S 生成的理想，记为 (S)。单元素生成的理想记为 (a)。

在交换环中，

$$(S) = \left\{ \sum_{k=1}^{m} r_k s_k : \ m \in \mathbb{N},\ r_k \in R,\ s_k \in S \right\},$$

其中 $m = 0$ 时取空和 0。右侧是理想：两个有限和相减仍是有限和，乘以 r 只需把每个系数 r_k 换为 rr_k；它含 S，因为 $s = 1_R s$。任何包含 S 的理想都含这些有限和，因此右侧正是交定义的最小理想。

例 8.10 (整数中的理想)**.** $\mathbb{Z}$ 的理想 (6) 是所有 $6k$ 的集合。它包含 6，并在加法和任意整数倍下封闭。$(6) \subseteq (2)$，因为每个 $6k$ 都是 $2(3k)$；包含关系记录整除关系的反向：$(a) \subseteq (b)$ 当且仅当 $b \mid a$。

定义 8.11 (环同态)**.** 设 R, S 为环。映射 $\varphi : R \to S$ 称为环同态，若对 $a, b \in R$ 满足

$$\varphi(a + b) = \varphi(a) + \varphi(b), \quad \varphi(ab) = \varphi(a)\varphi(b), \quad \varphi(1_R) = 1_S.$$

由加法群同态的性质，有 $\varphi(0_R) = 0_S$、$\varphi(-a) = -\varphi(a)$。其核为 $\ker \varphi = \{a : \varphi(a) = 0_S\}$，像为 $\operatorname{im} \varphi = \{\varphi(a) : a \in R\}$。

命题 8.12 (理想的原像与像)**.** *对交换环之间的同态 $\varphi : R \to S$，若 J 为 S 的理想，则 $\varphi^{-1}(J)$ 为 R 的理想；若 I 为 R 的理想，则 $\varphi(I)$ 为 $\operatorname{im} \varphi$ 的理想。*

证明. 因为 $\varphi(0)=0\in J$，原像非空。若 $a,b\in\varphi^{-1}(J)$、$r\in R$，则

$$\varphi(a-b)=\varphi(a)-\varphi(b)\in J,\qquad \varphi(ra)=\varphi(r)\varphi(a)\in J.$$

由理想判别，原像为理想。对像，若 $x=\varphi(a),y=\varphi(b)$，其中 $a,b\in I$，则

$$x-y=\varphi(a-b)\in\varphi(I).$$

若 $t=\varphi(r)\in\operatorname{im}\varphi$，则 $tx=\varphi(ra)\in\varphi(I)$，所以它为像环的理想。对于一般 $t\in S$，未必有原像 r，故不能由此推出它为 S 的理想。例如包含 $\mathbb{Z}\to\mathbb{Q}$ 将理想 $2\mathbb{Z}$ 送到同一子集，但 $\frac{1}{2}\cdot 2=1\notin 2\mathbb{Z}$，所以其像不是 $\mathbb{Q}$ 的理想。 □

8.3 商环与同态

群的商构造引出一个问题：换代表元以后，运算结果是否不变。环中必须同时检查加法和乘法。

定义 8.13 (商环)**.** 设 I 是交换环 R 的理想。在 R 上定义

$$r\sim_I s\quad\Longleftrightarrow\quad r-s\in I.$$

等价类记为 $r+I$，全体等价类记为 R/I。定义

$$(r+I)+(s+I)=(r+s)+I,\qquad (r+I)(s+I)=rs+I.$$

这里的 $r+I$ 是加法陪集，不是尚未定义的商环加法。

定理 8.14 (商环运算良定义)**.** 上述两个运算与代表元的选择无关，且使 R/I 成为交换环。自然投影

$$q:R\longrightarrow R/I,\qquad r\longmapsto r+I$$

是满射环同态，且 $\ker q=I$。

证明. 先证 $\sim_I$ 是等价关系。自反性由 $r-r=0\in I$；对称性由 I 对取负封闭；传递性由理想的加法封闭。

若 $r+I=r'+I$、$s+I=s'+I$，则 $r-r',s-s'\in I$。加法结果之差

$$(r+s)-(r'+s')=(r-r')+(s-s')\in I.$$

乘法结果之差为

$$rs-r's'=r(s-s')+s'(r-r')\in I,$$

其中使用了理想对任意环元素的吸收性。因此两种公式良定义。

将 $r+I$ 简记为 $[r]$。加法的结合律、交换律、零元和逆元为

$$([r]+[s])+[t]=[(r+s)+t]=[r+(s+t)]=[r]+([s]+[t]),$$

$$[r]+[s]=[r+s]=[s+r]=[s]+[r],\qquad [r]+[0]=[r],\quad [r]+[-r]=[0].$$

乘法及其与加法的相容性由

$$\begin{aligned}([r][s])[t] = [(rs)t] = [r(st)] = [r]([s][t]),\\ [r][s] = [rs] = [sr] = [s][r], \qquad [1][r] = [r],\\ [r]([s]+[t]) = [r(s+t)] = [rs+rt] = [r][s] + [r][t]\end{aligned}$$

给出；交换律再给出另一侧的分配律。因此 R/I 为交换有单位环。并且

$$q(r+s) = [r+s] = q(r) + q(s), \quad q(rs) = [rs] = q(r)q(s), \quad q(1) = [1].$$

任意商类 $[r]$ 都等于 $q(r)$，故 q 是满射环同态。最后

$$q(r) = 0 + I \iff r + I = I \iff r \in I,$$

故核为 I。 □

命题 8.15 (核是理想)**.** *环同态的核是理想；若 $I \subseteq \ker\varphi$，则存在唯一环同态 $\bar\varphi : R/I \to S$ 满足 $\varphi = \bar\varphi q$。*

证明. 核是零理想的原像，因而是理想；也可逐式计算 $\varphi(a-b) = 0$、$\varphi(ra) = \varphi(r)0 = 0$ 来验证。设 $I \subseteq \ker\varphi$，定义 $\bar\varphi(r+I) = \varphi(r)$。若 $r + I = r' + I$，则

$$r - r' \in I \subseteq \ker\varphi \implies 0 = \varphi(r - r') = \varphi(r) - \varphi(r'),$$

所以公式良定义。对任意商类，分别计算

$$\begin{aligned}\bar\varphi((r+I)+(s+I)) = \varphi(r+s) = \varphi(r) + \varphi(s),\\ \bar\varphi((r+I)(s+I)) = \varphi(rs) = \varphi(r)\varphi(s),\\ \bar\varphi(1+I) = \varphi(1) = 1_S.\end{aligned}$$

故它为环同态，且 $\bar\varphi q(r) = \varphi(r)$。若另一同态 h 满足 $hq = \varphi$，则对每个商类有 $h(r+I) = h(q(r)) = \varphi(r) = \bar\varphi(r+I)$，所以唯一。 □

定理 8.16 (环的第一同构定理)**.** *对环同态 $\varphi : R \to S$，存在唯一环同构*

$$R/\ker\varphi \;\cong\; \operatorname{im}\varphi, \qquad r + \ker\varphi \longmapsto \varphi(r).$$

证明. 前一命题给出良定义环同态 $\Phi(r + \ker\varphi) = \varphi(r)$，其值落在 $\operatorname{im}\varphi$。若 $y \in \operatorname{im}\varphi$，取 r 使 $y = \varphi(r)$，就有 $y = \Phi(r + \ker\varphi)$，故满射。又

$$\begin{aligned}\Phi(r+\ker\varphi) = \Phi(s+\ker\varphi) &\iff \varphi(r) = \varphi(s)\\ &\iff \varphi(r-s) = 0\\ &\iff r - s \in \ker\varphi\\ &\iff r + \ker\varphi = s + \ker\varphi.\end{aligned}$$

所以单射。双射环同态的逆保持运算：若 $x = \Phi(a), y = \Phi(b)$，则

$$\Phi^{-1}(x+y) = a + b = \Phi^{-1}(x) + \Phi^{-1}(y), \qquad \Phi^{-1}(xy) = ab = \Phi^{-1}(x)\Phi^{-1}(y),$$

且 $\Phi^{-1}(1) = 1$，故确为环同构。 □

8.4 中国剩余定理

能否找到一个整数，除以 2 余 1，除以 3 余 2？5 就是一个答案。更一般地，任意指定这两个余数时，是否总能找到同时满足要求的整数？换成环与理想的语言，就是问自然映射 $R \to R/I \times R/J$ 何时为满射。中国剩余定理给出的条件是 $I + J = R$；在整数的这个例子中，$1 = 3 - 2 \in (2) + (3)$，所以条件成立。映射的核则记录两个答案的差可以是什么。

对两个理想 I, J，先定义它们的和与积：

$$I + J = \{i + j : i \in I, j \in J\}, \qquad IJ = \left\{ \sum_{k=1}^{r} i_k j_k : i_k \in I, j_k \in J, r \geq 0 \right\}.$$

它们都是理想。对于和，

$$(i + j) - (i' + j') = (i - i') + (j - j') \in I + J, \quad r(i + j) = ri + rj \in I + J.$$

对于积，两个有限和的差仍为这种形式，将负号吸收到 i_k；而

$$r \sum_k i_k j_k = \sum_k (r i_k) j_k \in IJ.$$

两者均含零，由理想判别得结论。和是包含 I, J 的最小理想：任何含二者的理想都对加法封闭。并且 $IJ \subseteq I \cap J$，因为每项 $i_k j_k$ 同时属于两个理想。

若 A, B 为环，直积环 $A \times B$ 定义为逐坐标加法与乘法，单位为 $(1_A, 1_B)$。例如

$$((a, b)(a', b'))(a'', b'') = ((aa')a'', (bb')b'') = (a(a'a''), b(b'b'')),$$

从而乘法结合，其余环公理也由两个坐标的相应等式给出。

定义 8.17 (互素理想)**.** 交换环 R 中的理想 I, J 称为互素，若 $I + J = R$。等价地，存在 $i \in I, j \in J$ 使 $i + j = 1$。

引理 8.18 (互素理想的交与积)**.** *若 $I + J = R$，则 $I \cap J = IJ$，其中 IJ 是有限和 $\sum i_k j_k$ $(i_k \in I, j_k \in J)$ 组成的理想。*

证明. 每个有限和 $\sum i_k j_k$ 同时属于 I 和 J，故 $IJ \subseteq I \cap J$。反过来，取 $x \in I \cap J$，写 $1 = i + j$，其中 $i \in I, j \in J$。则

$$x = x(i + j) = xi + xj.$$

因 $x \in J, i \in I$，有 $xi \in IJ$；因 $x \in I, j \in J$，有 $xj \in IJ$。故 $x \in IJ$。 □

定理 8.19 (两个理想的中国剩余定理)**.** *若 $I + J = R$，则映射*

$$\Phi : R \longrightarrow R/I \times R/J, \qquad r \longmapsto (r + I, r + J)$$

诱导环同构

$$R/(I \cap J) \cong R/I \times R/J.$$

证明. 逐坐标计算有

$$\begin{aligned} \Phi(r + s) &= ((r + s) + I, (r + s) + J) = \Phi(r) + \Phi(s), \\ \Phi(rs) &= (rs + I, rs + J) = \Phi(r)\Phi(s), \\ \Phi(1) &= (1 + I, 1 + J). \end{aligned}$$

所以 Φ 是环同态。其核为

$$\{r : r + I = I,\ r + J = J\} = I \cap J.$$

由 $I + J = R$ 取 $i \in I, j \in J$ 使 $i + j = 1$。给定目标 $(a + I, b + J)$，令 $r = aj + bi$，则

$$r - a = a(j - 1) + bi = -ai + bi = (b - a)i \in I,$$

$$r - b = aj + b(i - 1) = aj - bj = (a - b)j \in J.$$

所以 $\Phi(r) = (a + I, b + J)$，证明满射。第一同构定理给出所需同构。 □

例 8.20 (整数的剩余条件)**.** $\mathbb{Z}/(6) \cong \mathbb{Z}/(2) \times \mathbb{Z}/(3)$，因为 $(2) + (3) = \mathbb{Z}$。具体地，$1 = (-1) \cdot 2 + 1 \cdot 3$，而一个整数模 6 的类由它模 2 和模 3 的类唯一确定。若两个模数不互素，例如 $2, 4$，则不能直接写成模 2 与模 4 的乘积：两个条件必须在公共模 2 上相容。

8.5 多项式环与代入

令 R 为交换环。多项式不是先验给定的函数，而是有限形式和

$$f = a_0 + a_1 x + \cdots + a_n x^n, \qquad a_i \in R,$$

其中 x 是形式符号。两个形式相等，当且仅当每个次数的系数相等。加法逐项进行，乘法按卷积定义：

$$\left(\sum_{i} a_i x^i\right)\left(\sum_{j} b_j x^j\right) = \sum_{k} \left(\sum_{i+j=k} a_i b_j\right) x^k.$$

有限支撑保证每个系数的内层和是有限和。

这些公式确实构成环。若 f, g, h 的系数分别为 a_i, b_j, c_ℓ，则乘积的第 k 个系数满足

$$((fg)h)_k = \sum_{u+\ell=k} \left(\sum_{i+j=u} a_i b_j\right) c_\ell = \sum_{i+j+\ell=k} (a_i b_j) c_\ell,$$

$$(f(gh))_k = \sum_{i+v=k} a_i \left(\sum_{j+\ell=v} b_j c_\ell\right) = \sum_{i+j+\ell=k} a_i (b_j c_\ell).$$

R 中的结合律使两者相等。分配律则由

$$(f(g+h))_k = \sum_{i+j=k} a_i (b_j + c_j) = \sum_{i+j=k} a_i b_j + \sum_{i+j=k} a_i c_j = (fg + fh)_k$$

得到。乘法交换律来自 $a_i b_j = b_j a_i$ 并交换求和指标。加法群公理逐系数继承自 R；常数多项式 1 的系数为 $1, 0, 0, \ldots$，卷积后不改变任意系数，故为单位。

定义 8.21 (多项式环与次数)**.** 上述形式的集合记为 $R[x]$。非零多项式的次数 $\deg f$ 是最高的非零系数下标，零多项式的次数约定为 $-\infty$。首项系数称为首系数；首系数为 1 的多项式称为首一。

命题 8.22 (多项式环的泛性质)**.** 设 S 为交换环，$\varphi: R \to S$ 为环同态，$s \in S$。存在唯一环同态

$$\widetilde{\varphi}: R[x] \to S$$

满足 $\widetilde{\varphi}(a) = \varphi(a)$ $(a \in R)$ 且 $\widetilde{\varphi}(x) = s$。它由

$$\widetilde{\varphi}\left(\sum_i a_i x^i\right) = \sum_i \varphi(a_i) s^i$$

给出。

证明. 公式给出有限和。若 $f = \sum_i a_i x^i, g = \sum_i b_i x^i$，补零系数使两个求和范围相同，则

$$\widetilde{\varphi}(f+g) = \sum_i \varphi(a_i + b_i) s^i = \sum_i \varphi(a_i) s^i + \sum_i \varphi(b_i) s^i = \widetilde{\varphi}(f) + \widetilde{\varphi}(g).$$

对乘法，

$$\widetilde{\varphi}(fg) = \sum_k \varphi\left(\sum_{i+j=k} a_i b_j\right) s^k = \sum_{i,j} \varphi(a_i)\varphi(b_j) s^{i+j} = \widetilde{\varphi}(f)\widetilde{\varphi}(g).$$

它保持单位。任意满足条件的同态必须把 $a_i x^i$ 送到 $\varphi(a_i) s^i$，再由有限和得到同一公式，故唯一。 □

命题 8.23 (乘积的次数)**.** 若 R 为整环，$f, g \in R[x]$ 均非零，则 $fg \neq 0$ 且 $\deg(fg) = \deg f + \deg g$。

证明. 令 $m = \deg f, n = \deg g$，首系数分别为 a_m, b_n。若 $i + j > m + n$，则必有 $i > m$ 或 $j > n$，故对应乘积系数为零。对于次数 $m+n$，只有 $i = m, j = n$ 可能贡献非零项，因而

$$(fg)_{m+n} = a_m b_n \neq 0,$$

其中最后一步使用整环无零因子。因此最高非零次数正好为 $m + n$。 □

例 8.24 (形式多项式与函数)**.** 在 $\mathbb{F}_2[x]$ 中，$x^2 - x$ 的二次系数为 1，所以它不是零多项式；但代入 0 与 1，都有 $a^2 - a = 0$，所以它给出零函数 $\mathbb{F}_2 \to \mathbb{F}_2$。由此可见，按系数定义的多项式相等与在一个有限域上取值相等，并不等价。

定理 8.25 (域上的带余除法)**.** 设 F 为域，$f, g \in F[x]$ 且 $g \neq 0$。存在唯一 $q, r \in F[x]$ 使

$$f = qg + r, \qquad r = 0 \text{ 或 } \deg r < \deg g.$$

证明. 先证明存在性。若 $f = 0$，取 $q = r = 0$。否则对非负整数 $\deg f$ 作强归纳，固定非零除式 g。当 $\deg f < \deg g$ 时取 $q = 0, r = f$。当 $m = \deg f \geq n = \deg g$，记首系数为 a, b，定义

$$f_1 = f - a b^{-1} x^{m-n} g.$$

被减项的 m 次系数为 $ab^{-1}b = a$，所以它与 f 的首项抵消；因此 $f_1 = 0$ 或 $\deg f_1 < m$。对 f_1 使用归纳假设，写 $f_1 = q_1 g + r$，其中 $r = 0$ 或 $\deg r < n$，则

$$f = (q_1 + a b^{-1} x^{m-n}) g + r.$$

所以递归每次严格降低非零被除式的次数，必有限步终止。

若 $f = qg + r = q'g + r'$，则 $(q - q')g = r' - r$。假设 $q \neq q'$，次数命题给出

$$\deg((q-q')g) = \deg(q-q') + \deg g \geq \deg g.$$

而右侧若非零，其次数小于 $\deg g$；若为零，则与左侧非零矛盾。因此 $q = q'$，代回等式得 $r = r'$。 □

命题 8.26 (因子定理与根数上界)**.** *设 F 为域，$a \in F$，$f \in F[x]$。则*

$$f(a) = 0 \quad \Longleftrightarrow \quad (x - a) \mid f.$$

若 $f \neq 0$ 且 $\deg f = d$，则 f 在 F 中至多有 d 个不同的根。在任意扩域 $E \supseteq F$ 中，同一根数上界仍成立。

证明. 把 f 除以 $x - a$，余式为一个常数 r。代入 a 得 $f(a) = r$，所以余式为零恰好等价于 $f(a) = 0$。

对非零多项式的次数 d 归纳证明根数上界。$d = 0$ 时非零常数没有根。若 $d \geq 1$ 且 f 有根 a，写 $f = (x - a)g$，其中 $g \neq 0$、$\deg g = d - 1$。对任何不同于 a 的根 b，有 $0 = (b - a)g(b)$；因 $b - a \neq 0$ 且 F 是域，必有 $g(b) = 0$。归纳假设说明这样的 b 至多有 $d - 1$ 个，加上 a 共至多 d 个。若 f 没有根，结论直接成立。

将 f 视为 $E[x]$ 的多项式时，非零首系数仍非零，次数不变。在域 E 中应用同一证明即可。 □

这一结果也见 Milne[15] 第 10 页命题 1.7。根数上界中不能随意删去域的条件：在 $\mathbb{Z}/8\mathbb{Z}$ 上，二次多项式 $x^2 - 1$ 有 $[1], [3], [5], [7]$ 四个不同的根。上述证明在消去非零的 $b - a$ 时使用了无零因子性质；整环同样满足这个条件，但一般交换环未必满足。

例 8.27 (由一个关系得到复数)**.** 在 $\mathbb{R}[x]$ 中，每个多项式唯一写成

$$f(x) = q(x)(x^2 + 1) + a + bx.$$

因此商环 $\mathbb{R}[x]/(x^2 + 1)$ 的每个元素唯一写为 $a + b\xi$，其中 $\xi = x + (x^2 + 1)$ 且 $\xi^2 = -1$。映射

$$a + b\xi \longmapsto a + bi$$

是环同构。若 $(a, b) \neq (0, 0)$，则

$$(a + b\xi)^{-1} = \frac{a - b\xi}{a^2 + b^2},$$

所以这个商环是域。商环不是把 x 当作一个具体实数，而是把关系 $x^2 + 1 = 0$ 写进代数结构。

8.6 整环中的不可约分解

定义 8.28 (整除、伴随、不可约与素元)**.** 整环 R 中，$a \mid b$ 表示存在 $c \in R$ 使 $b = ac$。若存在单位 u 使 $a = ub$，称 a, b 伴随。

1. 非零非单位 p 称为不可约元，若每当 $p = ab$，必有 a 或 b 为单位。
2. 非零非单位 p 称为素元，若对任意 $a, b \in R$，$p \mid ab$ 蕴含 $p \mid a$ 或 $p \mid b$。

定义 8.29 (欧几里得整环、PID 与 UFD)**.** 整环 R 是欧几里得整环，若存在 $\delta : R\setminus\{0\}\to\mathbb{N}$，使对 $a\in R, b\in R\setminus\{0\}$ 存在 q,r 满足 $a=bq+r$ 且 $r=0$ 或 $\delta(r)<\delta(b)$。若每个理想均为主理想，称 R 为主理想整环（PID）。若每个非零非单位都能分解为不可约元，且分解除去单位和顺序后唯一，称 R 为唯一分解整环（UFD）。

定理 8.30 (欧几里得整环是 PID)**.** *每个欧几里得整环都是主理想整环。*

证明. 设 I 为非零理想。在 I 的非零元素中取 d 使 $\delta(d)$ 最小。对任意 $a\in I$，作除法 $a=dq+r$，其中 $r=0$ 或 $\delta(r)<\delta(d)$。因 $a,d\in I$，有 $r=a-dq\in I$；最小性排除非零余式，故 $r=0$，即 $a\in(d)$。反向 $(d)\subseteq I$ 由 $d\in I$ 和理想吸收性得到。因此 $I=(d)$。零理想由 0 生成。 □

引理 8.31 (PID 中的 Bézout 性质)**.** *若 R 为 PID，$a,b\in R$，则存在 $d\in R$ 使 $(a,b)=(d)$；并且 $d=ua+vb$ 对某些 $u,v\in R$。*

证明. 由 PID 定义，理想 (a,b) 为某个 (d)。因 $d\in(a,b)$，按生成理想的具体形式可写成 $d=ua+vb$。另一方面 $a,b\in(d)$，所以 d 同时整除 a,b；这说明 d 是一个最大公因子，差一个单位的选择不影响主理想。 □

定理 8.32 (PID 中的 Euclid 引理)**.** *若 p 是 PID 中的不可约元，且 $p\mid ab$，则 $p\mid a$ 或 $p\mid b$。*

证明. 若 $p\mid a$，结论已成立。设 $p\nmid a$。取 $(p,a)=(d)$，则 $p=dc$、$a=dt$ 对某些 $c,t\in R$ 成立。由于 p 不可约，d 或 c 为单位。若 c 为单位，则

$$a=dt=pc^{-1}t,$$

与 $p\nmid a$ 矛盾。因此 d 为单位，$(p,a)=R$。于是存在 $u,v\in R$ 使 $up+va=1$。又 $p\mid ab$，写 $ab=pw$，则

$$b=(up+va)b=p(ub)+v(pw)=p(ub+vw),$$

故 $p\mid b$。 □

定理 8.33 (PID 是 UFD)**.** *每个 PID 都是 UFD。*

证明. 先证明理想的上升链会稳定。若 $I_1\subseteq I_2\subseteq\cdots$，令 $I=\bigcup_n I_n$。若 $a,b\in I$，二者属于某个共同的 I_N，于是 $a-b\in I_N\subseteq I$；若 $r\in R$，则 $ra\in I_N\subseteq I$。所以 I 为理想。由 PID 性质，$I=(d)$。取 N 使 $d\in I_N$，则

$$I=(d)\subseteq I_N\subseteq I,$$

故 $I_N=I$，之后每个 I_n 均等于 I。

假设存在非零非单位 a_1 没有不可约分解。它必可写为 $a_1=a_2b_1$，其中两个因子均非单位，且至少一个没有不可约分解，否则合并两者的分解即得到 a_1 的分解。选取其中没有分解的因子记为 a_2，重复得到 $a_n=a_{n+1}b_n$，其中 a_{n+1},b_n 均非单位。于是 $(a_n)\subsetneq(a_{n+1})$。若相等，则 $a_{n+1}=a_nc$，从而

$$a_n=a_{n+1}b_n=a_ncb_n.$$

整环中 $a_n\neq0$，可消去得到 $cb_n=1$，与 b_n 非单位矛盾。因此每一步包含均严格，违背刚证明的上升链稳定性。所以不可约分解存在。

对于唯一性，允许把单位单列，设

$$a = up_1 \cdots p_r = vq_1 \cdots q_s,$$

其中 u, v 为单位，各 p_i, q_j 不可约。由 Euclid 引理反复作用于右侧乘积，p_1 整除某个 q_j：它不能整除单位 v，否则 $1 = v^{-1}p_1 c$，使 p_1 成为单位。重排后设 $q_1 = p_1 c$。不可约性和 p_1 非单位强迫 c 为单位。消去非零的 p_1，得到

$$up_2 \cdots p_r = vc\, q_2 \cdots q_s.$$

对不可约因子数归纳。如果一边先没有因子，则另一边的不可约元乘积为单位；但乘积 xy 为单位时，交换环中 $x(y(xy)^{-1}) = 1$，所以 x 也为单位，矛盾。因此 $r = s$，逐次配对的不可约元只差单位，即伴随。 □

例 8.34 (为什么多项式环的系数环很重要)**.** 若 F 是域，则 $F[x]$ 为欧几里得整环，欧几里得函数为次数。若系数环不是域，首系数可能不可逆，带余除法不一定成立。例如，试图在 $\mathbb{Z}[x]$ 中写

$$x = q(x)(2x+1) + r(x), \qquad r = 0 \text{ 或 } \deg r < 1.$$

若 $q = 0$，则 $r = x$ 不符合次数要求；若 $q \neq 0$，比较次数可知 q 必为常数整数。再比较一次项系数，得到 $2q = 1$，在 $\mathbb{Z}$ 中无解。因此这一次除法无法按规定完成。

8.7 极大理想与素理想

定义 8.35 (极大理想与素理想)**.** 交换环 R 的真理想 M 称为极大理想，若不存在真理想 I 满足 $M \subsetneq I \subsetneq R$。真理想 P 称为素理想，若 $ab \in P$ 蕴含 $a \in P$ 或 $b \in P$。

命题 8.36 (商环判别)**.** *交换环 R 的真理想 M 极大，当且仅当 R/M 是域；真理想 P 素，当且仅当 R/P 是整环。*

证明. 设 M 极大，取非零类 $a + M$，即 $a \notin M$。理想 $M + (a)$ 含 M 且含 a，所以严格大于 M，从而 $M + (a) = R$。因此存在 $m \in M, r \in R$ 使 $1 = m + ra$，取商得

$$(r + M)(a + M) = ra + M = 1 + M.$$

每个非零类均有逆元。又 M 为真理想，$1 \notin M$，故 $1 + M \neq 0 + M$，因此 R/M 为域。

反过来，设 R/M 为域，且 $M \subsetneq I \subseteq R$。取 $a \in I \setminus M$，其非零类有逆元 $r + M$。于是 $ra - 1 \in M \subseteq I$，而 $ra \in I$，所以

$$1 = ra - (ra - 1) \in I \quad \Longrightarrow \quad I = R.$$

这证明不存在严格夹在 M, R 之间的理想，故 M 极大。

对于真理想 P，R/P 的单位类非零。而

$$(a + P)(b + P) = 0 + P \iff ab \in P,$$
$$a + P = 0 + P \text{ 或} b + P = 0 + P \iff a \in P \text{ 或} b \in P.$$

所以商环无零因子的条件，逐字就是素理想的定义。这证明第二个等价命题。 □

例 8.37 (整数中的两类理想)**.** 在 $\mathbb{Z}$ 中，(p) 对素数 p 是极大理想，因为 $\mathbb{Z}/(p) \cong \mathbb{F}_p$ 是域；(0) 是素理想，因为 $\mathbb{Z}$ 是整环，但不是极大理想，因为 $\mathbb{Z}/(0) = \mathbb{Z}$ 不是域。极大理想一定是素理想，而反向需要额外条件。

文献：理想、商环、中国剩余定理和多项式环见 [8, 9, 12]。本章中的商构造沿用第 3 章对商集良定义性的证明方式；PID 与 UFD 的证明把 Euclid 引理、主理想上升链和不可约分解的依赖逐项写出。

9 交换群与模论

9.1 从向量空间到模

向量空间的标量来自域，因此每个非零标量都能除。若把标量环换成 $\mathbb{Z}$ 或多项式环，线性组合仍然有意义，但系数未必可逆，向量之间也可能出现被非零标量消去的现象。模论保留线性组合、核、像和商构造，同时记录系数环不能提供除法的部分。

本章固定一个交换有单位环 R。除非特别说明，R-模的标量乘法写作 rm，模同态保持加法与标量乘法。取 $R=\mathbb{Z}$ 时，模正好是交换群：整数的标量乘法由重复加法和取负唯一确定。

定义 9.1 (R-模)**.** R-模 M 是一个交换群 $(M,+)$，连同映射 $R\times M\to M$，满足

$$(r+s)m=rm+sm,\quad r(m+n)=rm+rn,\quad (rs)m=r(sm),\quad 1_Rm=m.$$

由公理可推出 $0_Rm=0_M$、$r0_M=0_M$ 和 $(-r)m=-(rm)$。例如

$$0_Rm=(0_R+0_R)m=0_Rm+0_Rm$$

在加法群中消去一个 0_Rm 即得第一式。另外两式由

$$r0_M=r(0_M+0_M)=r0_M+r0_M,\qquad rm+(-r)m=(r+(-r))m=0_Rm=0_M$$

分别推出；第一式消去 $r0_M$，第二式使用加法逆元的唯一性。模的加法结构和标量结构不能混为一谈：M 是交换群并不意味着它是一个环，rm 也不是环乘法。

例 9.2 (三个标量环)**.** R^n 是自由 R-模，标准向量构成基。$\mathbb{Z}/n\mathbb{Z}$ 是 $\mathbb{Z}$-模，也是一个循环交换群。若 F 为域，F-模就是通常的向量空间；因此线性代数是模论在系数为域时的特殊情形。

例 9.3 (一个生成元与一条关系)**.** 在 $\mathbb{Z}$-模 $\mathbb{Z}/6\mathbb{Z}$ 中，每个元素都能写成 $k[1]$，所以一个元素 $[1]$ 就能生成整个模。但它不能作为基，因为 $0[1]=6[1]=[0]$，表示并不唯一。这里 6 是非零整数，不能像域中的非零标量那样消去。

也可以从自由模 $\mathbb{Z}e$ 出发，再要求 $6e=0$。为使这条等式与加法、整数倍相容，所有 $6ke$ 都必须变成零；因此应取商 $\mathbb{Z}e/(6\mathbb{Z}e)$。映射 $ke\mapsto[k]_6$ 把这个商与 $\mathbb{Z}/6\mathbb{Z}$ 对应起来。这就是“生成元与关系”的一个实例：生成元说明元素怎样写出，关系说明哪些写法应当视为同一个元素。下面先建立子模与商模，随后用自由模和同态把这种描述写成一般形式。

定义 9.4 (模同态)**.** 模同态是保持加法和标量乘法的映射。其核与像分别为

$$\ker f=\{m:f(m)=0\},\qquad \operatorname{im} f=\{f(m):m\in M\}.$$

全体从 M 到 N 的 R-模同态组成的集合记为 $\operatorname{Hom}_R(M,N)$；下标标明必须保持的标量环。

9.2 子模、商模与良定义性

定义 9.5 (子模)**.** 子集 $N \subseteq M$ 若在 M 的加法与标量乘法的限制下成为 R-模，称为 M 的子模，记作 $N \leq M$。

命题 9.6 (子模判别)**.** 非空子集 N 是子模，当且仅当对任意 $x, y \in N, r \in R$ 有 $x - y \in N$、$rx \in N$。

证明. 若 N 是子模，$-y \in N$ 且 $x + (-y) \in N$，所以 $x - y \in N$；标量乘法的封闭性给出 $rx \in N$。

反过来，由非空性选 $a \in N$。依次使用差封闭，得到

$$0_M = a - a \in N, \qquad -x = 0_M - x \in N, \qquad x + y = x - (-y) \in N.$$

故 N 是加法子群。对 $x, y \in N, r, s \in R$，封闭性使下列等式的两边均在 N 内，而等式本身在 M 中成立：

$$(r+s)x = rx + sx, \qquad r(x+y) = rx + ry, \qquad (rs)x = r(sx), \qquad 1_R x = x.$$

因此限制后的运算满足全部模公理。 □

定义 9.7 (商模)**.** 设 $N \leq M$。定义

$$x \sim_N y \quad \Longleftrightarrow \quad x - y \in N.$$

等价类记为 $x + N$，商集记为 M/N。定义

$$(x + N) + (y + N) = (x + y) + N, \qquad r(x + N) = rx + N.$$

定理 9.8 (商模的良定义性)**.** 上述运算与代表元的选择无关，并使 M/N 成为 R-模。自然投影

$$q : M \to M/N, \qquad x \mapsto x + N$$

是满模同态，且 $\ker q = N$。

证明. 等价关系的自反、对称和传递性分别来自 $x - x = 0 \in N$、N 对取负封闭及加法封闭。

设 $x + N = x' + N$ 与 $y + N = y' + N$。则 $x - x', y - y' \in N$，故

$$(x + y) - (x' + y') = (x - x') + (y - y') \in N,$$

从而加法结果相等。对于标量，$rx - rx' = r(x - x') \in N$，所以标量作用也与代表元无关。

将 $x + N$ 简记为 $[x]$。商模的加法满足

$$\begin{aligned} ([x] + [y]) + [z] &= [(x + y) + z] = [x + (y + z)] = [x] + ([y] + [z]), \\ [x] + [y] &= [x + y] = [y + x] = [y] + [x], \\ [x] + [0] &= [x], \qquad [x] + [-x] = [0]. \end{aligned}$$

标量乘法的四条公理分别由下式给出：

$$\begin{aligned}(r+s)[x] &= [(r+s)x] = [rx+sx] = r[x]+s[x],\\ r([x]+[y]) &= [r(x+y)] = [rx+ry] = r[x]+r[y],\\ (rs)[x] &= [(rs)x] = [r(sx)] = r(s[x]),\\ 1_R[x] &= [1_R x] = [x].\end{aligned}$$

因此 M/N 是 R-模。自然投影满足

$$q(x+y) = [x+y] = [x]+[y], \qquad q(rx) = [rx] = r[x],$$

所以为模同态；任意 $[x] \in M/N$ 都等于 $q(x)$，故满射。最后

$$q(x) = 0+N \iff x+N = N \iff x \in N,$$

故核恰为 N。 □

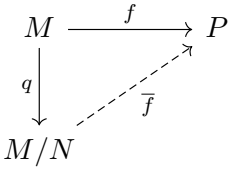


图 9.1: 包含于核的子模使同态下降到商模

定理 9.9 (模的因子化)**.** *若 $f : M \to P$ 为模同态且 $N \le \ker f$，则存在唯一模同态 $\overline{f} : M/N \to P$ 使*

$$f = \overline{f} \circ q, \qquad \overline{f}(x+N) = f(x).$$

证明. 设 $x+N = y+N$。由陪集相等的判别，$x-y \in N \subseteq \ker f$，所以

$$0_P = f(x-y) = f(x)-f(y), \qquad f(x) = f(y).$$

因此 $\overline{f}$ 良定义。对 $x, y \in M, r \in R$，

$$\begin{aligned}\overline{f}((x+N)+(y+N)) &= f(x+y) = f(x)+f(y),\\ \overline{f}(r(x+N)) &= f(rx) = rf(x).\end{aligned}$$

右侧分别是 $\overline{f}(x+N)+\overline{f}(y+N)$ 与 $r\overline{f}(x+N)$，故为模同态。而

$$(\overline{f} \circ q)(x) = \overline{f}(x+N) = f(x).$$

若 $gq = f$，则对任意商类 $x+N = q(x)$，

$$g(x+N) = g(q(x)) = f(x) = \overline{f}(x+N).$$

故 $g = \overline{f}$，证明唯一性。 □

命题 9.10 (核与像)**.** *模同态的核是定义域的子模，像是陪域的子模；双射模同态的逆也是模同态。*

证明. 先有 $f(0)=f(0+0)=f(0)+f(0)$，消去后得 $f(0)=0$。故核含 0，像也含 $0=f(0)$。若 $x,y\in\ker f, r\in R$，则

$$f(x-y)=f(x)-f(y)=0-0=0,\qquad f(rx)=rf(x)=r0=0,$$

所以 $x-y, rx\in\ker f$。若 $u=f(x), v=f(y)\in\operatorname{im} f$，则

$$u-v=f(x-y)\in\operatorname{im} f,\qquad ru=rf(x)=f(rx)\in\operatorname{im} f.$$

对核与像分别应用子模判别，得到两项结论。

若 f 双射，令 $g=f^{-1}$。对 $u=f(x), v=f(y)$，有

$$g(u-v)=g(f(x-y))=x-y=g(u)-g(v),\qquad g(ru)=g(f(rx))=rx=rg(u).$$

所以逆保持模结构。 □

定理 9.11 (模的第一同构定理). 对模同态 $f:M\to P$，有自然同构

$$M/\ker f\cong\operatorname{im} f,\qquad m+\ker f\longmapsto f(m).$$

证明. 记 $K=\ker f$。因子化定理给出良定义模同态

$$\overline{f}:M/K\to\operatorname{im} f,\qquad m+K\longmapsto f(m).$$

对任意 $u\in\operatorname{im} f$，按像的定义选 $m\in M$ 使 $u=f(m)$，则 $u=\overline{f}(m+K)$，所以满射。对任意两个商类，

$$\begin{aligned}\overline{f}(m+K)=\overline{f}(n+K)&\iff f(m)=f(n)\\&\iff f(m-n)=0\\&\iff m-n\in K\\&\iff m+K=n+K.\end{aligned}$$

这证明单射。前一命题保证双射模同态的逆也保持模运算，故 $\overline{f}$ 是模同构。 □

9.3 生成、自由模与直和

定义 9.12 (生成子模). 给定 $S\subseteq M$，由 S 生成的子模为所有有限和

$$\left\{\sum_{k=1}^{m} r_k s_k : m\in\mathbb{N},\ r_k\in R,\ s_k\in S\right\}$$

. 若该子模等于 M，称 S 生成 M。有限生成指存在有限生成集。

定义 9.13 (基与自由模). 元素族 $(e_i)_{i\in I}$ 是 M 的基，若每个 $m\in M$ 都能唯一写成有限和

$$m=\sum_{i\in I} r_i e_i,$$

其中只有有限多个 r_i 非零。具有基的模称为自由模。

定义 9.14 (直积与直和)**.** 模族 $(M_i)_{i\in I}$ 的直积 $\prod_i M_i$ 是所有元组 (m_i)，逐坐标运算。直和 $\bigoplus_i M_i$ 是其中仅有有限多个 $m_i \neq 0$ 的元组组成的子模。有限指标集时二者相同；无限指标集时直积允许无限支撑，而直和不允许。

定理 9.15 (自由模的泛性质)**.** 令 $R^{(I)} = \bigoplus_{i\in I} R$，$e_i$ 为第 i 个标准向量。对任意 R-模 M 和任意族 $(m_i)_{i\in I}$，存在唯一模同态

$$f : R^{(I)} \to M, \qquad f(e_i) = m_i.$$

当且仅当 (m_i) 是 M 的基时，f 是同构。

证明. 对有限支撑向量 $r = (r_i)$ 定义 $f(r) = \sum_i r_i m_i$。和只在有限集合 $\operatorname{supp}(r) = \{i : r_i \neq 0\}$ 上计算，增加零项不改变和，因此公式有意义。若 $r = (r_i), s = (s_i)$，在两支撑的有限并集上计算，得

$$f(r+s) = \sum_i (r_i + s_i)m_i = \sum_i r_i m_i + \sum_i s_i m_i = f(r) + f(s),$$
$$f(ar) = \sum_i (ar_i)m_i = a\sum_i r_i m_i = af(r) \qquad (a \in R).$$

又 $f(e_i) = 1_R m_i = m_i$。若 h 为另一个具有这些基向量像的模同态，则

$$h((r_i)) = h\left(\sum_i r_i e_i\right) = \sum_i r_i h(e_i) = \sum_i r_i m_i = f((r_i)),$$

所以 $h = f$。

若 (m_i) 是基，表示的存在性说明 f 满射；表示的唯一性说明 f 的核为零，故为同构。反过来，若 f 是同构，则每个 $m \in M$ 有原像，给出有限线性组合；若两个组合给出同一元素，其差在核中，因核为零，系数相同。 □

命题 9.16 (有限生成模的商表示)**.** M 由 n 个元素生成，当且仅当存在满同态 $R^n \twoheadrightarrow M$。因此任意有限生成模同构于 R^n/K，其中 $K \leq R^n$。

证明. 给定生成元 $m_1, \dots, m_n$，把标准基 e_i 送到 m_i，自由模泛性质给出满同态。反向，标准基的像生成陪域。取核 K，对该满同态应用第一同构定理即可。 □

9.4 挠与无挠

从现在起，令 R 为整环。非零标量能够消去某个元素，是模中区别于向量空间的现象。

定义 9.17 (挠元与挠子模)**.** $m \in M$ 称为挠元，若存在 $0 \neq r \in R$ 使 $rm = 0$。全体挠元记为 M_{tor}。若 $M_{\mathrm{tor}} = 0$，称 M 无挠。

命题 9.18. M_{tor} 是子模。

证明. 若 $rm = 0$、$sn = 0$，其中 $r, s \neq 0$，则

$$rs(m+n) = s(rm) + r(sn) = 0.$$

整环保证 $rs \neq 0$，故 $m+n$ 挠。对任意 $a \in R$，$r(am) = a(rm) = 0$，所以 am 挠；加法逆元由 $r(-m) = 0$ 得到。 □

例 9.19 (无挠不等于自由)**.** $\mathbb{Q}$ 作为 $\mathbb{Z}$-模无挠，但不是自由模。若它有基，取基中 e。因 $e/2 \in \mathbb{Q}$，可写成有限整数线性组合；将等式 $e = 2(e/2)$ 展开，右侧在该基中的每个系数都是偶数，而左侧对应 e 的系数为 1，矛盾。

9.5 PID 上的子模

为了分类有限生成模，需要知道自由模的子模仍然足够规整。这个结论对一般环错误；主理想整环的主理想性质正是证明能够进行的原因。

定理 9.20 (PID 的自由子模定理)**.** *若 R 是主理想整环，则 R^n 的每个子模 K 都有一个元素数不超过 n 的基。*

证明. 对 n 作归纳。$n = 0$ 时 R^0 只有零向量，唯一子模是零模；$n = 1$ 时 K 是理想，故 $K = (d)$。若 $d = 0$，它是零模；若 $d \neq 0$，它由 d 自由生成。

设 $n > 1$，令 $\pi : R^n \to R$ 为第一坐标投影。$\pi(K)$ 是 R 的理想，故 $\pi(K) = (d)$。若 $d = 0$，则 $K \subseteq 0 \oplus R^{n-1}$，由归纳假设结论成立。

设 $d \neq 0$。取 $u \in K$ 使 $\pi(u) = d$，记 $L = \ker(\pi|_K)$。对 $k \in K$，存在 $r \in R$ 使 $\pi(k) = rd$。于是

$$\pi(k - ru) = \pi(k) - r\pi(u) = rd - rd = 0, \qquad k = ru + (k - ru) \in Ru + L.$$

反向，$u \in K$ 且 $L \subseteq K$，故 $Ru + L \subseteq K$。因此 $K = Ru + L$。考虑模同态

$$\Phi : R \oplus L \to K, \qquad (r, v) \longmapsto ru + v.$$

分解等式说明它满射；若 $\Phi(r, v) = 0$，则

$$0 = \pi(ru + v) = rd + 0 = rd.$$

整环中 $d \neq 0$，故 $r = 0$，再代回得 $v = 0$，所以 Φ 单射。

删去首坐标使 $0 \oplus R^{n-1}$ 与 R^{n-1} 同构；它将 L 送到子模，故由归纳假设可取 L 的基 $v_1, \dots, v_t$，其中 $t \leq n - 1$。每个 $k \in K$ 都可写成

$$k = ru + \sum_{i=1}^{t} a_i v_i.$$

若 $ru + \sum_i a_i v_i = 0$，取首坐标得 $rd = 0$，故 $r = 0$；再由 v_i 的无关性得全部 $a_i = 0$。因此 $u, v_1, \dots, v_t$ 是 K 的基，其元素数为 $1 + t \leq n$。 □

9.6 整数矩阵与 Smith 标准形

取 $R = \mathbb{Z}$。有限生成交换群是有限生成 $\mathbb{Z}$-模，所以可以用整数矩阵记录它的生成元和关系。

定义 9.21 (模的有限表示)**.** 对同态 $A : R^m \to R^n$，其余核为 $\operatorname{coker} A = R^n / \operatorname{im} A$。若 $M \cong \operatorname{coker} A$，称 A 是 M 的一个有限表示。R^n 的标准基是生成元，A 的每一列是一个关系。

若对 R^n 的值域换基，等于对矩阵做可逆行变换；若对 R^m 的定义域换基，等于做可逆列变换。这两种变化只改变表示，不改变余核的同构类。乘以 2 不是可逆的整数变换，因为 2 在 $\mathbb{Z}$ 中没有单位逆元。

定义 9.22 (整数初等变换)**.** 整数矩阵允许交换两行或两列、将一行或一列乘以 -1、以及把一行或一列的整数倍加到另一行或另一列。每一步都由可逆整数矩阵实现。

定理 9.23 (Smith 标准形)**.** *任意整数矩阵 A 存在可逆整数矩阵 P, Q，使*

$$PAQ = \operatorname{diag}(d_1, \dots, d_s, 0, \dots, 0), \qquad d_i > 0, \quad d_i \mid d_{i+1}.$$

证明. 对矩阵行数与列数的较小者作归纳。若行数或列数为零，或矩阵全为零，结论成立。否则先交换行列，把一个非零元素移到左上角，并用换号使其成为正整数 d。

以下归约中，每次发现不被 d 整除的元素，就把左上角的正整数严格减小。先检查第一列。若其中有 a 不被 d 整除，作带余除法 $a = qd + r$，其中 $0 < r < d$；相应行减去第一行的 q 倍，再交换两行。只看第一列的这两个位置，变换为

$$\begin{pmatrix} d \\ a \end{pmatrix} \longmapsto \begin{pmatrix} d \\ a - qd \end{pmatrix} = \begin{pmatrix} d \\ r \end{pmatrix} \longmapsto \begin{pmatrix} r \\ d \end{pmatrix}.$$

所以左上角变成 r。从这个更小的正整数重新开始。若第一列各项均被 d 整除，用行变换清除左上角以外各项。

再检查第一行。若其中有不被 d 整除的项 a，写 $a = qd + r$、$0 < r < d$。对对应的两列作

$$\begin{pmatrix} d & a \end{pmatrix} \longmapsto \begin{pmatrix} d & a - qd \end{pmatrix} = \begin{pmatrix} d & r \end{pmatrix} \longmapsto \begin{pmatrix} r & d \end{pmatrix},$$

便得到更小的正左上角，再重新开始。若各项都被整除，用列变换清除第一行其余项。由于此时第一列在下方已经为零，这些列变换不会破坏第一列的零元。于是得到

$$\begin{pmatrix} d & 0 \\ 0 & B \end{pmatrix}.$$

若 B 有元素 b_{ij} 不被 d 整除，把相应行加到第一行；第一行的左上角仍为 d，并出现 b_{ij}。写 $b_{ij} = qd + r$，其中 $0 < r < d$，再将对应列减去第一列的 q 倍并交换列。只看首行、该行与首列、该列形成的子块，整个操作为

$$\begin{pmatrix} d & 0 \\ 0 & b_{ij} \end{pmatrix} \longmapsto \begin{pmatrix} d & b_{ij} \\ 0 & b_{ij} \end{pmatrix} \longmapsto \begin{pmatrix} d & r \\ 0 & b_{ij} \end{pmatrix} \longmapsto \begin{pmatrix} r & d \\ b_{ij} & 0 \end{pmatrix}.$$

于是 r 位于左上角，随后重新开始整个过程。

每次重新开始都会严格减小一个正整数；而两次重新开始之间，只检查有限多项并进行有限次清除。所以过程终止，且终止时得到上式并满足 d 整除 B 的每个元素。对较小的矩阵 B 应用归纳假设。其行列变换均只在余下块内进行，不改变已经清零的第一行与第一列；其新元素仍是原来元素的整数线性组合，故仍被 d 整除。因此所得首个非零对角元被 d 整除，再结合余下块内的整除链，就得到整幅矩阵的整除链。所有步骤均为允许的可逆整数行列变换，合并它们得到 P, Q。 □

命题 9.24 (Smith 形与余核)**.** *若 $A : \mathbb{Z}^m \to \mathbb{Z}^n$ 的 Smith 形为 $D = \operatorname{diag}(d_1, \dots, d_s, 0, \dots)$，则*

$$\operatorname{coker} A \cong \mathbb{Z}^{n-s} \oplus \bigoplus_{i=1}^{s} \mathbb{Z}/d_i\mathbb{Z}.$$

当 $d_i = 1$ 时，相应的直和项为零群，可省略。

证明. 先验证换基前后的像满足

$$\operatorname{im}(PAQ) = \{PAQx : x \in \mathbb{Z}^m\} = \{PAy : y \in \mathbb{Z}^m\} = P(\operatorname{im} A),$$

第二个等号使用 Q 满射。故

$$\alpha : \mathbb{Z}^n/\operatorname{im} A \to \mathbb{Z}^n/\operatorname{im} D, \qquad x + \operatorname{im} A \longmapsto Px + \operatorname{im} D$$

良定义：$x - x' \in \operatorname{im} A$ 时 $P(x - x') \in \operatorname{im} D$。它保持加法和整数倍，逆映射由 P^{-1} 给出，因而是模同构。

再定义满同态

$$\rho : \mathbb{Z}^n \to \left(\bigoplus_{i=1}^{s} \mathbb{Z}/d_i\mathbb{Z}\right) \oplus \mathbb{Z}^{n-s}, \quad (x_1, \dots, x_n) \longmapsto ([x_1]_{d_1}, \dots, [x_s]_{d_s}, x_{s+1}, \dots, x_n).$$

其满射性由任取每个剩余类的整数代表得到。而

$$\begin{aligned} x \in \ker\rho &\iff d_i \mid x_i \ (1 \le i \le s), \quad x_j = 0 \ (j > s) \\ &\iff x = \sum_{i=1}^{s} d_i a_i e_i \quad (a_i \in \mathbb{Z}) \\ &\iff x \in \operatorname{im} D. \end{aligned}$$

对 ρ 应用第一同构定理，再与 α 复合，即得所需余核分解。 □

例 9.25 (完整的 Smith 计算)**.** 令

$$A = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}.$$

第二行减去第一行的三倍，再把第二列减去第一列的两倍，最后将第二行乘以 -1，得到

$$\begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 4 \\ 0 & -4 \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

所以

$$\operatorname{coker} A \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/4\mathbb{Z}.$$

由于行列式为 $-8 \neq 0$，$A : \mathbb{Z}^2 \to \mathbb{Z}^2$ 的核为零；余核记录的是像在整数格中的指数，而不是实数线性代数中的维数。

9.7 有限生成交换群结构定理

定理 9.26 (有限生成交换群)**.** 每个有限生成交换群 G 都同构于

$$\mathbb{Z}^r \oplus \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_k\mathbb{Z}, \qquad 2 \le d_1 \mid d_2 \mid \cdots \mid d_k.$$

这些 d_i 与 r 在同构意义下唯一。将各 d_i 分解为素数幂并使用互素因子的中国剩余定理，可得到初等因子分解。

证明. 取生成元 $g_1,\dots,g_n$，得到满同态 $\pi:\mathbb{Z}^n\to G$。由 PID 子模定理，$K=\ker\pi$ 是有限生成自由 $\mathbb{Z}$-模；选定其基后，包含 $K\hookrightarrow\mathbb{Z}^n$ 由整数矩阵 A 表示，并有 $G\cong\operatorname{coker}A$。Smith 标准形和上一命题给出所示分解，其中零对角元产生自由的 $\mathbb{Z}$ 因子，非单位对角元产生循环挠因子。

下面从群本身恢复这些参数。记 $T=G_{\mathrm{tor}}$。由已经证明的分解，T 正好为有限循环因子的直和，而 $H=G/T\cong\mathbb{Z}^r$。商群 $H/2H$ 具有自然的 $\mathbb{F}_2=\mathbb{Z}/2\mathbb{Z}$ 向量空间结构，标量 $[a]_2$ 作用为 $[h]\mapsto[ah]$；改变 a 或 h 的代表元只改变一个 $2H$ 中的元素，故良定义。由分解可知 $H/2H\cong\mathbb{F}_2^r$，所以它的维数唯一确定 r。

对每个素数 p，定义 $T_p=\{x\in T:p^ax=0\text{ 对某个}a\ge0\}$。这是由群本身确定的子群：若 $p^ax=p^by=0$，则 $p^{a+b}(x-y)=0$。对每个 $j\ge1$，商群

$$E_{p,j}=p^{j-1}T_p/p^jT_p$$

被 p 消去，因而是 $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$ 上的向量空间，标量同样由整数倍定义。对一个循环因子 $C=\mathbb{Z}/p^a\mathbb{Z}$，若 $a<j$，则 $p^{j-1}C=0$。若 $a\ge j$，映射

$$\mathbb{Z}\to p^{j-1}C/p^jC,\qquad t\longmapsto tp^{j-1}[1]+p^jC$$

满射，其核满足

$$\begin{aligned}tp^{j-1}[1]\in p^jC&\iff\exists u,v\in\mathbb{Z},\quad tp^{j-1}=up^j+vp^a\\&\iff\exists u,v\in\mathbb{Z},\quad t=up+vp^{a-j+1}\\&\iff p\mid t.\end{aligned}$$

最后的逆方向在 $t=up$ 时取 $v=0$。所以此商同构于 $\mathbb{Z}/p\mathbb{Z}$，贡献一维。直和中的运算逐坐标进行，故若 $T_p\cong\bigoplus_\ell\mathbb{Z}/p^{a_\ell}\mathbb{Z}$，则

$$\dim_{\mathbb{F}_p}E_{p,j}=\#\{\ell:a_\ell\ge j\},\qquad\#\{\ell:a_\ell=j\}=\dim E_{p,j}-\dim E_{p,j+1}.$$

这里 $\#$ 表示有限集合的元素数。因此每个指数的重数由群自身的这些商空间恢复，全部素数幂初等因子都由 G 唯一确定。

最后恢复整除链 $d_1\mid\cdots\mid d_k$。若 $T=0$，则 $k=0$。否则对每个 p，将恢复出的正指数非降序排列。k 等于这些序列长度的最大值：整除链中若 $p\mid d_1$，则 p 出现在全部 k 项中，故最大值确为 k。将每个较短序列在左侧补零到长度 k，第 i 项对应的各素数幂相乘，就唯一恢复 d_i。逐素数的中国剩余定理保证这一重组与原循环因子一致。这证明唯一性。 □

例 9.27 (整数模的两种分解)**.** $\mathbb{Z}/12\mathbb{Z}\cong\mathbb{Z}/4\mathbb{Z}\oplus\mathbb{Z}/3\mathbb{Z}$, 因为 $4,3$ 互素。另一方面, $\mathbb{Z}/4\mathbb{Z}\oplus\mathbb{Z}/2\mathbb{Z}$ 不是循环群：每个元素的阶至多为 4，而若它是 $\mathbb{Z}/8\mathbb{Z}$，应有阶为 8 的元素。前一个分解按互素素数拆开，后一个例子说明直和不总能合并成一个循环因子。

9.8 一般主理想整环上的结构定理

整数矩阵的归约使用了正整数严格下降。若系数环只是 PID，不能预先假定它带有这样的欧几里得函数。下面直接从主理想性质出发，寻找自由模和它的子模可以同时采用的坐标。

本节固定 PID R。先说明自由模的基元素数为何确定。若 R 是域，这就是向量空间维数的唯一性。若 R 不是域，取一个非零非单位，利用不可约分解取其不可约因子 p。由 PID 中

的 Bézout 性质，$R/(p)$ 是域：当 $p\nmid a$ 时，$(p,a)=R$，故 $up+va=1$，从而 $[v][a]=[1]$。若 $R^a\cong R^b$，该同构将 pR^a 双射到 pR^b，因为

$$f(px)=pf(x),\qquad f^{-1}(py)=pf^{-1}(y).$$

因此它诱导 $(R/(p))^a\cong(R/(p))^b$，向量空间维数唯一性给出 $a=b$。有限自由模的秩由此定义为任意一组基的元素数。

定义 9.28 (零化理想)**.** 对 R-模 M 和 $x\in M$，定义

$$\operatorname{Ann}_R(x)=\{a\in R:ax=0\},\qquad \operatorname{Ann}_R(M)=\{a\in R:ay=0\ \text{对每个}y\in M\}.$$

二者称为元素和模的零化理想。

它们确为理想：若 $ax=bx=0$，则 $(a-b)x=0$；若 $ax=0$，则 $(ra)x=r(ax)=0$。将这些等式对每个 $x\in M$ 使用，就得到第二个集合的理想性质。

命题 9.29 (循环子模)**.** *对 $x\in M$，有同构*

$$Rx\cong R/\operatorname{Ann}_R(x),\qquad ax\longleftrightarrow a+\operatorname{Ann}_R(x).$$

若 M 有限生成且每个元素都是挠元，则存在 $0\neq d\in R$ 使 $dM=0$。

证明. 映射 $\theta:R\to Rx$，$a\mapsto ax$，由模公理保持加法和标量乘法；按 Rx 的定义它满射，核恰为 $\operatorname{Ann}_R(x)$。第一同构定理给出所述同构。反向写出的对应也良定义，因为

$$ax=bx\iff(a-b)x=0\iff a-b\in\operatorname{Ann}_R(x).$$

若 $x_1,\dots,x_n$ 生成 M，为每个 x_i 选 $0\neq d_i$ 使 $d_ix_i=0$。令 $d=\prod_i d_i$；整环性质给出 $d\neq0$，且

$$d\sum_i a_ix_i=\sum_i a_i\Big(\prod_{j\neq i}d_j\Big)(d_ix_i)=0.$$

$n=0$ 时 $M=0$，取空积 $d=1$ 即可。 □

定理 9.30 (相容基定理)**.** *设 F 为秩 n 的有限自由 R-模，$K\le F$。存在 F 的基 $e_1,\dots,e_n$，以及非零元素 $a_1,\dots,a_s$，使*

$$K=Ra_1e_1\oplus\cdots\oplus Ra_se_s,\qquad s\le n,\qquad a_1\mid a_2\mid\cdots\mid a_s.$$

特别地，$a_1e_1,\dots,a_se_s$ 是 K 的基。

证明. 对 n 归纳。$n=0$ 或 $K=0$ 时取空基或空整除链。设 $n>0$ 且 $K\neq0$。考察理想族

$$\mathcal{S}=\{\varphi(K):\varphi\in\operatorname{Hom}_R(F,R)\}.$$

任取 F 的基 $v_1,\dots,v_n$，坐标函数 $v_i^*(\sum_j c_jv_j)=c_i$ 均为模同态。某个非零 $k\in K$ 必有一个非零坐标，故 $\mathcal{S}$ 含非零理想。

每个非零主理想 (c) 只有有限多个包含它的理想：若 $(c)\subseteq(b)$，则 $c=bd$；唯一分解说明，除伴随外 b 的不可约因子及指数只有有限种选择，而伴随元生成同一理想。从 $\mathcal{S}$ 中选非零理想，再在包含它的有限多个成员中选极大者，记为

$$(a)=\psi(K),\qquad a\neq0.$$

这里的极大性是指在族 $\mathcal{S}$ 内没有严格包含它的成员。取 $k_0 \in K$ 使 $\psi(k_0) = a$。

下面证明：对每个 $\varphi : F \to R$，有 $a \mid \varphi(k_0)$。将 $(a, \varphi(k_0))$ 写成 (b)，并取 $u, v \in R$ 使

$$b = ua + v\varphi(k_0) = (u\psi + v\varphi)(k_0).$$

于是

$$(a) \subseteq (b) \subseteq (u\psi + v\varphi)(K).$$

最右端属于 $\mathcal{S}$；极大性使两次包含均为等号。故 $\varphi(k_0) \in (a)$，所需整除成立。

对坐标函数逐一应用，写 $v_i^*(k_0) = ab_i$，令 $e_1 = \sum_i b_i v_i$，则

$$k_0 = ae_1, \qquad a = \psi(k_0) = a\psi(e_1), \qquad \psi(e_1) = 1.$$

最后一步使用 $a \neq 0$ 及整环中的消去律。令 $L = \ker \psi$。任取 $x \in F$，有

$$x = \psi(x)e_1 + \big(x - \psi(x)e_1\big), \qquad \psi\big(x - \psi(x)e_1\big) = 0.$$

若 $re_1 \in L$，则 $0 = \psi(re_1) = r$。因此

$$F = Re_1 \oplus L.$$

对于 $x \in K$，由 $\psi(K) = (a)$ 写 $\psi(x) = ar$，得到

$$x = rae_1 + (x - rae_1), \qquad x - rae_1 \in K \cap L.$$

反向 $Rae_1 \subseteq K$ 且 $K \cap L \subseteq K$，交为零由 $Re_1 \cap L = 0$ 得到，故

$$K = Rae_1 \oplus (K \cap L).$$

自由子模定理说明 L 有有限基。若基有 t 个元素，$F = Re_1 \oplus L \cong R^{1+t}$，由秩的唯一性得 $t = n - 1$。将归纳假设用于 $K \cap L \leq L$，取 L 的基 $e_2, \dots, e_n$ 使

$$K \cap L = Ra_2e_2 \oplus \cdots \oplus Ra_se_s, \qquad a_2 \mid \cdots \mid a_s.$$

令 $a_1 = a$。若 $s \geq 2$，在基 $e_1, \dots, e_n$ 下定义 $\eta = e_1^* + e_2^*$。因 $a_1e_1 \in K$，有 $a_1 = \eta(a_1e_1) \in \eta(K)$，所以 $(a_1) \subseteq \eta(K)$。由原来的极大性，$\eta(K) = (a_1)$。又

$$a_2 = \eta(a_2e_2) \in \eta(K) = (a_1),$$

故 $a_1 \mid a_2$。若 $s = 1$，无须再验证此式。直和等式同时证明所列元素生成 K；若 $\sum_i r_i a_i e_i = 0$，则基的无关性给出 $r_i a_i = 0$，再由 $a_i \neq 0$ 得 $r_i = 0$。因此它们确为基。 □

注 9.31 (坐标为何要同时选取)**.** Conrad 在 [13] 定理 2.14 的证明中，先寻找一个在子模上具有极大理想像的线性函数，再用它分裂出第一条坐标轴。上面的证明沿用这一构造。关键等式 $\psi(e_1) = 1$ 保证第一条轴在整个自由模中确实能分离出来；$\psi(K) = (a_1)$ 则测量子模沿这条轴只能走到哪些点。

例 9.32 (整格中的相容基)**.** 在 $F = \mathbb{Z}^2$ 中取

$$K = \{(x, y) : x \equiv y \pmod 3\}.$$

令 $e_1=(1,1)$，$e_2=(1,0)$。任意 (x,y) 有唯一表达

$$(x,y)=ye_1+(x-y)e_2,$$

故 e_1,e_2 是 F 的基；而 $x-y$ 被 3 整除恰好给出

$$K=\mathbb{Z}e_1\oplus\mathbb{Z}(3e_2),\qquad F/K\cong\mathbb{Z}/3\mathbb{Z}.$$

原来的两条标准坐标轴都没有一个标准基向量属于 K，相容基则将第一条完整保留的方向直接显现出来。

定理 9.33 (有限生成 PID 模的存在性)**.** *每个有限生成 R-模 M 都有分解*

$$M\cong R^r\oplus\bigoplus_{i=1}^{k}R/(d_i),\qquad 0\neq d_i\ \text{非单位},\qquad d_1\mid\cdots\mid d_k.$$

r,k 可为零。d_i 称为不变因子。

证明. 由有限生成性选满同态 $\pi:R^n\to M$，令 $K=\ker\pi$。对 $K\le R^n$ 应用相容基定理，得到基 e_i 及整除链 $a_1,\dots,a_s$。定义

$$\rho:R^n\to\left(\bigoplus_{i=1}^{s}R/(a_i)\right)\oplus R^{n-s},\quad \sum_{i=1}^{n}x_ie_i\mapsto([x_1],\dots,[x_s],x_{s+1},\dots,x_n).$$

它保持两种模运算，并由选取各剩余类的代表元可知满射。其核逐坐标计算为

$$\rho\Big(\sum_i x_ie_i\Big)=0\iff x_i\in(a_i)\ (i\le s),\quad x_i=0\ (i>s)$$

$$\iff\sum_i x_ie_i\in\bigoplus_{i=1}^{s}Ra_ie_i=K.$$

所以 $M\cong R^n/K\cong\operatorname{im}\rho$。单位 a_i 给出零模；因为 $a_i\mid a_{i+1}$，单位项只能出现在整除链的开头。删除它们，令 $r=n-s$，即得结论。 □

命题 9.34 (挠部分与自由部分)**.** *上述分解中，M_{tor} 恰好对应所有 $R/(d_i)$ 的直和，而 $M/M_{\mathrm{tor}}\cong R^r$。因此有限生成无挠 PID 模是自由模。*

证明. 若 $v=(v_1,\dots,v_r)\in R^r$，$av=0$ 且 $a\neq0$，则每个 $av_i=0$，整环性质给出 $v_i=0$。故自由模无挠。令 $d=\prod_i d_i$，则 d 消去全部循环因子，且 $d\neq0$。于是

$$(v,t)\text{ 是挠元}\iff v=0.$$

从左向右，用消去它的非零标量逐坐标计算；从右向左，用 d 消去 t。因此挠子模对应 $0\oplus\bigoplus_i R/(d_i)$。投影 $(v,t)\mapsto v$ 满射且核恰为此子模，第一同构定理给出商模。若 M 无挠，挠部分为零，故 $M\cong R^r$。 □

9.9 初等因子与不变因子的唯一性

本节仍固定 PID R。非零 a,b 若满足 $a = ub$，其中 u 是单位，称它们伴随；等价地 $(a) = (b)$。因此 $R/(d)$ 只能确定主理想 (d)，而不能区分相差单位的两个生成元。

命题 9.35 (主分解)**.** *设 T 是有限生成挠 R-模。对于不可约元 p，令*

$$T_p = \{x \in T : p^a x = 0 \text{ 对某个} a \in \mathbb{N}\}.$$

它是子模，只依赖 p 的伴随类。选择互不伴随的不可约元作代表，有

$$T = \bigoplus_p T_p, \qquad T_p \cong \bigoplus_{\ell=1}^{b_p} R/(p^{a_{p,\ell}}), \quad a_{p,\ell} \geq 1,$$

其中只有有限多个 T_p 非零。右侧的素幂称为初等因子。

证明. 若 $p^a x = p^b y = 0$，取 $c = \max(a,b)$，则

$$p^c(x-y) = p^{c-a}p^a x - p^{c-b}p^b y = 0, \qquad p^a(rx) = r(p^a x) = 0.$$

故为子模。若 $q = up$ 且 u 可逆，则 $q^a x = u^a p^a x$，所以 $T_q = T_p$。

由结构定理，T 为有限个 $R/(d_i)$ 的直和。将 d_i 作不可约分解。不同不可约元的幂生成互素理想：若 $(p^a, q^b) = (h) \neq R$，h 的不可约因子同时整除 p^a 和 q^b，唯一分解强制 p, q 伴随，矛盾。中国剩余定理因此给出

$$R/(d_i) \cong \bigoplus_{p \mid d_i} R/(p^{v_p(d_i)}),$$

这里 $v_p(d_i)$ 是 p 在 d_i 中的指数。该同构是模同构，因为剩余类映射保持加法和乘以 R 中元素。

在 $R/(p^a)$ 上，p^a 消去全部元素；若 q 与 p 不伴随，由 Bézout 等式选 u, v 使 $uq + vp^a = 1$，则乘以 q 的逆是乘以 u。故在此因子上 $q^j x = 0$ 强制 $x = 0$。于是上述分解中，被某个 p 的幂消去的元素恰好是全部 p 因子的坐标，其他坐标为零。这说明分解的对应子模正是由 T 自身定义的 T_p。它们的和等于 T，任一坐标分解唯一，故其和为内部直和。 □

引理 9.36 (逐层恢复素幂指数)**.** *设 $E_p = R/(p)$，并令*

$$L_{p,j} = p^{j-1}T_p / p^j T_p \qquad (j \geq 1).$$

这是有限维 E_p-向量空间，且

$$\dim_{E_p} L_{p,j} = \#\{\ell : a_{p,\ell} \geq j\}.$$

证明. 首先 p 消去 $L_{p,j}$，所以定义

$$[r]_p\,(x + p^j T_p) = rx + p^j T_p \qquad (x \in p^{j-1}T_p).$$

若 $r - r' = pu$ 且 $x - x' \in p^j T_p$，则

$$rx - r'x' = (r - r')x + r'(x - x') \in p^j T_p.$$

故作用与两种代表元都无关。向量空间公理从模公理下降得到，而 E_p 为域。

逐个计算 $C = R/(p^a)$ 的贡献。若 $a < j$，则 $p^{j-1}C = 0$。若 $a \geq j$，取满模同态

$$R \to p^{j-1}C/p^jC, \qquad r \mapsto rp^{j-1}[1] + p^jC.$$

它的核由下列等价式求得：

$$\begin{aligned} rp^{j-1}[1] \in p^jC &\iff \exists s,t \in R, \quad rp^{j-1} = sp^j + tp^a \\ &\iff r = sp + tp^{a-j+1} \\ &\iff r \in (p). \end{aligned}$$

第二步在整环中消去 $p^{j-1} \neq 0$；末步逆方向在 $r = sp$ 时取 $t = 0$。故这个商同构于 E_p，贡献一维。对有限直和，乘以 p、取像和作商都按坐标进行，因此各循环因子的贡献相加，得到所述维数。 □

定理 9.37 (PID 模结构的唯一性)**.** *定理 9.33 中的整数 r, k 唯一，且 d_i 在伴随意义下唯一。初等因子的伴随类及其重数也唯一。*

证明. 模同构保持挠子模：若 $ax = 0$，则 $af(x) = f(ax) = 0$，逆同构给出反向。因此 M/M_{tor} 的同构类由 M 确定。它同构于 R^r；先前自由秩的唯一性确定 r。

若 R 是域，每个非零标量可逆，因此挠子模为零，$k = 0$，证明结束。若 R 不是域，对每个不可约元的伴随类选择代表 p。模同构保持 T_p 和各 p^jT_p，因此保持层商 $L_{p,j}$ 的维数。由前引理，指数恰好等于 j 的循环因子个数为

$$c_{p,j} = \dim_{E_p} L_{p,j} - \dim_{E_p} L_{p,j+1}.$$

这些由 M 决定的非负整数恢复全部初等因子。

若挠部分为零，$k = 0$。否则，将同一 p 的全部正指数非降序排列，记其长度为 b_p。任何不变因子整除链中，$v_p(d_1), \dots, v_p(d_k)$ 非降序。又 d_1 非单位，取不可约因子 $p \mid d_1$，则 p 整除全部 d_i，所以该 p 的 $b_p = k$；对其他 p 均有 $b_p \leq k$。因此

$$k = \max_p b_p.$$

将每个 p 的正指数序列在左端补零到长度 k，记为 $b_{p,1}, \dots, b_{p,k}$。则

$$(d_i) = \left(\prod_p p^{b_{p,i}} \right), \qquad 1 \leq i \leq k.$$

乘积只含有限项，且等式由唯一分解得到。这就恢复了每个 (d_i)，即恢复 d_i 到单位倍。 □

推论 9.38 (何时可以延长子模的基)**.** *设 F 是有限自由 R-模，$K \leq F$。K 的某组基可以延长为 F 的基，当且仅当 F/K 无挠；若此条件成立，则 K 的每组基都可以延长。*

证明. 若 $u_1, \dots, u_s, w_1, \dots, w_t$ 是 F 的基，前 s 项是 K 的基，则映射

$$R^t \to F/K, \qquad (a_j) \mapsto \sum_j a_j w_j + K$$

满射。若像为零，存在 b_i 使 $\sum_j a_j w_j = \sum_i b_i u_i$，全体基向量的无关性给出 $a_j = b_i = 0$，故该映射单射。因此 $F/K \cong R^t$ 无挠。

反之，F/K 有限生成且无挠，由结构定理为自由模。取其基 $\overline{w}_1,\dots,\overline{w}_t$，并取提升 $w_j\in F$。对任意 K 的基 $u_1,\dots,u_s$ 和任意 $x\in F$，在商模中写

$$x+K=\sum_{j} a_j\overline{w}_j,\qquad x-\sum_{j} a_j w_j\in K.$$

将差写为 $\sum_i b_i u_i$，得到这些向量张成 F。若

$$\sum_{i} b_i u_i+\sum_{j} a_j w_j=0,$$

先取商得到 $\sum_j a_j\overline{w}_j=0$，所以全部 $a_j=0$；再用 K 的基无关性得到全部 $b_i=0$。故所列向量是一组基。 □

例 9.39 (子模自由与基可延长的区别)**.** $2\mathbb{Z}\le\mathbb{Z}$ 的基为单个元素 2，但 $\mathbb{Z}/2\mathbb{Z}$ 有非零挠元，所以该基不能延长为 $\mathbb{Z}$ 的基。直接计算也可看出：若 $1=2a$ 且 $a\in\mathbb{Z}$，则矛盾。相反，$K=\mathbb{Z}(1,1)\le\mathbb{Z}^2$ 的商无挠，取 $(1,0)$ 即可将其基延长，因为 $(x,y)=y(1,1)+(x-y)(1,0)$。

文献：模、有限生成交换群与 Smith 标准形见 [9, 12]；一般 PID 上的相容基构造参照 [13] 定理 2.14。本章把商模的良定义性、自由模泛性质、PID 子模定理和整数矩阵归约分别证明，再由它们推出结构定理。

10　多项式模与线性变换的标准形

前面研究一般线性映射时，定义域与陪域可以分别换基，矩阵因此能化成只由秩决定的标准形。现在研究的是同一空间上的变换 $T: V \to V$：我们还关心 T^2, T^3 的作用，所以输入、输出必须沿用同一组基。对应的矩阵关系是相似，秩已不足以区分这些变换。例如二维恒等变换与剪切变换都满秩，但只有前者满足 $T - I = 0$。

若找不到足够多的特征向量，可以从一个向量 v 出发，依次看 $v, Tv, T^2v, \dots$，寻找这些迭代之间的线性关系。把关系的系数写成多项式，就能使用上一章的模结构定理。本章沿这条路线得到两种形式：有理标准形在原标量域内用循环子空间组织变换；当最小多项式分裂为一次因子时，Jordan 标准形再把这些作用写成一条条向量链。下面先定义多项式怎样作用在向量上。

10.1　给向量规定多项式的作用

设 V 是域 $\mathbb{F}$ 上的有限维向量空间，$T: V \to V$ 是线性变换。记号 I 表示当前空间的恒等变换，或选定基后相应大小的单位矩阵。前面计算 T^k 时，已经把同一变换的反复复合当作一个整体。现在允许它们的线性组合。对

$$f(t) = \sum_{i=0}^{d} a_i t^i \in \mathbb{F}[t]$$

定义

$$f(T) = \sum_{i=0}^{d} a_i T^i, \qquad T^0 = \mathrm{id}_V.$$

这里 $f(T)$ 是 V 上的线性变换，$f(T)v$ 是它作用于向量 v 后的结果；多项式的变量 t 本身由 T 来实现。

命题 10.1 (代入的运算律)**.** *对 $f, g \in \mathbb{F}[t]$ 和 $a \in \mathbb{F}$，有*

$$(f+g)(T) = f(T) + g(T), \quad (af)(T) = af(T), \quad (fg)(T) = f(T)g(T), \quad 1(T) = \mathrm{id}_V.$$

特别地，所有 $f(T)$ 与 $g(T)$ 彼此交换。

证明. 在较短多项式的末尾补零，将加法写为

$$\sum_i (a_i + b_i) T^i = \sum_i a_i T^i + \sum_i b_i T^i.$$

数乘公式同理逐项成立。若 $f = \sum_i a_i t^i$，$g = \sum_j b_j t^j$，则

$$\begin{aligned} f(T)g(T) &= \Big(\sum_i a_i T^i\Big)\Big(\sum_j b_j T^j\Big) = \sum_{i,j} a_i b_j T^{i+j} \\ &= \sum_k \Big(\sum_{i+j=k} a_i b_j\Big) T^k = (fg)(T). \end{aligned}$$

各和有限，且 $T^i T^j = T^{i+j}$ 由复合的结合律及幂的递归定义得到。因 $fg = gf$，有 $f(T)g(T) = g(T)f(T)$。常数多项式 1 的代入按定义是 $1\mathrm{id}_V$。 □

定义 10.2 (由变换给出的多项式模)**.** 在 V 上定义 $\mathbb{F}[t]$ 的作用

$$f \cdot v = f(T)v.$$

保留 V 原有的加法，所得 $\mathbb{F}[t]$-模记为 V_T。

模公理可直接验证：代入的加法与乘法公式给出

$$(f+g)\cdot v = f\cdot v + g\cdot v, \qquad (fg)\cdot v = f\cdot(g\cdot v), \qquad 1\cdot v = v;$$

而 $f(T)$ 的线性性给出 $f\cdot(v+w) = f\cdot v + f\cdot w$。常数多项式的作用就是原有的 $\mathbb{F}$-标量乘法。

命题 10.3 (不变子空间与交织映射)**.** $U \subseteq V$ 是 V_T 的子模，当且仅当 U 是 T-不变的 $\mathbb{F}$-子空间。设 $S : W \to W$ 为另一线性变换，映射 $A : V_T \to W_S$ 为模同态，当且仅当它是 $\mathbb{F}$-线性映射且

$$AT = SA.$$

因此两个矩阵相似，当且仅当其对应的多项式模同构。

证明. 子模对常数多项式封闭，故是 $\mathbb{F}$-子空间；它对 t 的作用封闭，故 $T(U) \subseteq U$。反过来，若 U 是不变子空间，由归纳得 $T^i(U) \subseteq U$，再由线性组合封闭得

$$f(T)u = \sum_i a_i T^i u \in U \qquad (u \in U).$$

故为子模。

若 A 为模同态，限制于常数标量得 $\mathbb{F}$-线性，且

$$A(Tv) = A(t\cdot v) = t\cdot A(v) = SA(v).$$

反之，若 $AT = SA$，归纳给出 $AT^i = S^iA$，于是

$$A(f(T)v) = \sum_i a_i AT^i v = \sum_i a_i S^i A v = f(S)Av.$$

这连同 $\mathbb{F}$-线性说明 A 为模同态。若 A 可逆，$AT = SA$ 等价于 $S = ATA^{-1}$；在选定基后，这就是相似矩阵的关系。反向，相似所给出的可逆线性映射按上式是模同构。 □

10.2 最小多项式与循环子空间

本章以下设 $\dim_{\mathbb{F}} V = n \geq 1$。零空间的特征多项式与最小多项式均约定为 1，其标准形为空矩阵。

命题 10.4 (非零零化多项式的存在)**.** 存在非零 $f \in \mathbb{F}[t]$ 使 $f(T) = 0$。因此 V_T 是有限生成挠 $\mathbb{F}[t]$-模。

证明. 取 V 的基 $v_1, \ldots, v_n$。对每个 i，$n+1$ 个向量

$$v_i, Tv_i, \ldots, T^n v_i$$

线性相关，所以存在不全为零的系数 $a_{i,0},\dots,a_{i,n}$ 使

$$f_i(T)v_i=0,\qquad f_i(t)=\sum_{j=0}^{n}a_{i,j}t^j\neq 0.$$

令 $f=\prod_{i=1}^{n}f_i$。多项式环为整环，故 $f\neq 0$；代入后各因子可交换，故

$$f(T)v_i=\Big(\prod_{j\neq i}f_j(T)\Big)f_i(T)v_i=0.$$

由基张成性，$f(T)$ 消去每个向量。原来的 $\mathbb{F}$-基仍生成 V_T，因为常数多项式已经允许全部原来的线性组合。因此 V_T 有限生成；同一个非零 f 消去其每个元素，所以为挠模。　□

定义 10.5 (最小多项式)**.** 理想

$$I_T=\{f\in\mathbb{F}[t]:f(T)=0\}=\operatorname{Ann}_{\mathbb{F}[t]}(V_T)$$

非零。$\mathbb{F}[t]$ 是 PID，故 I_T 有唯一的首一生成元，记为 m_T，称为 T 的最小多项式。首一指最高次项的系数为 1。

存在性由任意非零理想生成元除以其首系数得到。若 a,b 都首一且 $(a)=(b)$，则 $a=ub$，其中 u 是非零常数；比较首系数得 $u=1$，故唯一。定义立即给出判别

$$f(T)=0\iff m_T\mid f.$$

特别地，m_T 在全部非零零化多项式中次数最小；若 $\deg f<\deg m_T$ 且 $f(T)=0$，整除关系强制 $f=0$。

定义 10.6 (循环子空间)**.** 对 $v\in V$ 定义

$$C_T(v)=\mathbb{F}[t]v=\operatorname{span}_{\mathbb{F}}\{v,Tv,T^2v,\dots\}.$$

理想 $\{f:f(T)v=0\}$ 的首一生成元记为 m_v，称为 v 相对于 T 的最小多项式。若 $C_T(v)=V$，称 v 是循环向量。

定理 10.7 (循环子空间的基)**.** 设 $v\neq 0$，$d=\deg m_v$。则

$$\mathbb{F}[t]/(m_v)\cong C_T(v),\qquad [f]\mapsto f(T)v,$$

且 $v,Tv,\dots,T^{d-1}v$ 是 $C_T(v)$ 的一组 $\mathbb{F}$-基。因此 $\dim C_T(v)=\deg m_v$，并且 $m_v\mid m_T$。

证明. 映射 $\theta_v:\mathbb{F}[t]\to C_T(v)$，$f\mapsto f(T)v$，由代入运算律是满模同态，核按定义为 (m_v)，故第一同构定理给出同构。

设 $m_v=t^d+c_{d-1}t^{d-1}+\cdots+c_0$。带余除法将任意 f 唯一写为 $f=qm_v+r$，$\deg r<d$，于是

$$f(T)v=q(T)m_v(T)v+r(T)v=r(T)v.$$

这证明所列 d 个向量张成。若

$$\sum_{i=0}^{d-1}b_iT^iv=0,$$

则 $r(t)=\sum_{i=0}^{d-1}b_it^i$ 属于 (m_v)，且次数小于 d，故 $r=0$，所有 $b_i=0$。因此线性无关。最后 $m_T(T)v=0$，所以 $m_T\in(m_v)$，即 $m_v\mid m_T$。　□

例 10.8 (剪切变换的循环向量)**.** 在 $\mathbb{F}^2$ 中令

$$T=\begin{pmatrix}1&1\\0&1\end{pmatrix},\qquad v=\begin{pmatrix}0\\1\end{pmatrix}.$$

有 $(T-I)v=(1,0)^{t}\neq 0$，$(T-I)^2v=0$，而 $v, Tv=(1,1)^{t}$ 无关。因此 $m_v=(t-1)^2$，且 v 是循环向量。T 在基 (v,Tv) 下的矩阵由

$$T(v)=Tv,\qquad T(Tv)=2Tv-v$$

给出，为 $\begin{pmatrix}0&-1\\1&2\end{pmatrix}$。这与原矩阵相似，但采用的是迭代生成的基。

10.3　伴随矩阵与有理标准形

定义 10.9 (伴随矩阵)**.** 对首一多项式 $f(t)=t^{d}+c_{d-1}t^{d-1}+\cdots+c_0$，$d\geq 1$，定义

$$C(f)=\begin{pmatrix}0&0&\cdots&0&-c_0\\1&0&\cdots&0&-c_1\\0&1&\cdots&0&-c_2\\\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&1&-c_{d-1}\end{pmatrix}.$$

$d=1$ 时它就是单项矩阵 $(-c_0)$。

矩阵的第 j 列给出第 j 个基向量的像。对循环子空间的基 $(v,Tv,\dots,T^{d-1}v)$，前 $d-1$ 列由 $T(T^{j}v)=T^{j+1}v$ 决定，末列则来自

$$m_v(T)v=0,\qquad T^{d}v=-c_0v-c_1Tv-\cdots-c_{d-1}T^{d-1}v.$$

因此其矩阵恰为 $C(m_v)$。

定理 10.10 (有理标准形)**.** *存在唯一的一列首一非常数多项式*

$$d_1\mid d_2\mid\cdots\mid d_k$$

使 T 的某个矩阵为

$$\operatorname{diag}\bigl(C(d_1),\dots,C(d_k)\bigr).$$

这称为 T 的有理标准形。这里的“有理”表示在原标量域 $\mathbb{F}$ 内完成，不要求 $\mathbb{F}=\mathbb{Q}$。

证明. 对有限生成挠 $\mathbb{F}[t]$-模 V_T 应用 PID 结构定理。自由部分必为零：若有非零自由直和项，其标准向量不会被任何非零多项式消去，与挠性矛盾。因此

$$V_T\cong\bigoplus_{i=1}^{k}\mathbb{F}[t]/(d_i).$$

取各 d_i 首一；非单位在 $\mathbb{F}[t]$ 中正好是非常数多项式。令 v_i 是第 i 项 $[1]$ 在此同构下的像，则其循环子空间构成内部直和。每个循环子空间的基为

$$v_i,Tv_i,\dots,T^{\deg d_i-1}v_i.$$

合并这些基得到 V 的基：生成性来自直和分解，若合并后有线性关系，先由直和性得每个循环子空间中的分量为零，再由各自的基无关性得全部系数为零。各空间不变，所以矩阵为所述块对角形。PID 结构定理的唯一性确定 (d_i)，首一条件进一步唯一确定 d_i。　□

引理 10.11 (伴随矩阵的特征多项式)**.** 对上述 f，有 $\det(xI - C(f)) = f(x)$，其中 x 是用于行列式计算的多项式变量。

证明. 对 d 归纳。$d = 1$ 时，$\det(x + c_0) = x + c_0$。设 $d \geq 2$。矩阵 $xI - C(f)$ 的第一行为 $(x, 0, \ldots, 0, c_0)$；次对角线各项为 -1，最后一项为 $x + c_{d-1}$。按第一行展开，删除第一行与第一列后的行列式记为 D。该矩阵是多项式

$$g(x) = x^{d-1} + c_{d-1}x^{d-2} + \cdots + c_1$$

的同型矩阵，故由归纳假设 $D = g(x)$。删除第一行与末列后的矩阵为上三角矩阵，对角线全为 -1，其行列式为 $(-1)^{d-1}$。所以

$$\begin{aligned} \det(xI - C(f)) &= xD + c_0(-1)^{1+d}(-1)^{d-1} \\ &= x\big(x^{d-1} + c_{d-1}x^{d-2} + \cdots + c_1\big) + c_0 = f(x). \end{aligned}$$

□

定理 10.12 (最小多项式、特征多项式与 Cayley–Hamilton)**.** 对有理标准形中的不变因子，有

$$m_T = d_k, \qquad \chi_T = \prod_{i=1}^{k} d_i.$$

特别地，$m_T \mid \chi_T$，且 $\chi_T(T) = 0$。

证明. 多项式 h 消去 $\mathbb{F}[t]/(d_i)$ 当且仅当它消去 $[1]$，等价于 $h \in (d_i)$。故

$$\begin{aligned} h(T) = 0 &\iff d_i \mid h \text{ 对每个} i \\ &\iff d_k \mid h. \end{aligned}$$

最后一步用整除链：若 $d_k \mid h$，则全部 $d_i \mid h$；反向只需取 $i = k$。于是 $I_T = (d_k)$，两者首一给出 $m_T = d_k$。

块对角矩阵的行列式等于各块行列式之积。这可从排列公式直接看出：非零乘积要求每一行所选列仍在同一个块中，否则所选矩阵元为零；所有保留的排列分别是各块内排列，其符号之积是整体排列的符号，因为不同块之间没有逆序。于是伴随矩阵引理给出

$$\chi_T(x) = \det(xI - [T]) = \prod_{i} \det(xI - C(d_i)) = \prod_{i} d_i(x).$$

该等式不依赖基，因为若 $B = P^{-1}AP$，则

$$xI - B = P^{-1}(xI - A)P, \quad \det(xI - B) = \det(P)^{-1}\det(xI - A)\det(P).$$

这些行列式运算在交换环 $\mathbb{F}[x]$ 中逐项成立。最后 d_k 是上述乘积的因子，故 $m_T \mid \chi_T$，由零化判别得 $\chi_T(T) = 0$。 □

推论 10.13 (相似的完整不变量与循环判别)**.** 两个同维线性变换相似，当且仅当它们的首一不变因子列相同。T 有循环向量，当且仅当 $m_T = \chi_T$，也当且仅当 $\deg m_T = n$。

证明. 相似等价于多项式模同构，由 PID 结构定理唯一性得到不变因子列相同；反过来，共同的标准形给出相似。

若存在循环向量，整个模是单个循环模，故 $k=1$，由上一定理 $m_T=\chi_T$。若 $\deg m_T=n$，因

$$\deg m_T=\deg d_k,\qquad n=\dim V=\sum_i \deg d_i,$$

而每个 $\deg d_i\ge 1$，故只能 $k=1$，第一个循环子空间即为整个 V。最后 $m_T\mid\chi_T$ 且两者首一、$\deg\chi_T=n$，所以 $\deg m_T=n$ 与 $m_T=\chi_T$ 等价。 □

例 10.14 (相同的两个多项式仍不足以分类)**.** 在 $\mathbb{F}^6$ 中考虑两个幂零变换，其循环模分解分别为

$$\mathbb{F}[t]/(t^3)\oplus\mathbb{F}[t]/(t^3),\qquad \mathbb{F}[t]/(t^3)\oplus\mathbb{F}[t]/(t^2)\oplus\mathbb{F}[t]/(t).$$

二者都有 $\chi_T=t^6$、$m_T=t^3$，但不变因子列不同，因而不相似。直接计算核也能区分它们：在 $\mathbb{F}[t]/(t^a)$ 上，乘以 t 的核由 $[t^{a-1}]$ 生成，故两者核的维数分别为 2 和 3。

10.4 从矩阵实际计算不变因子

有理标准形的存在性说明应当寻找哪些循环块。对于已经给出的矩阵，还需要把这个分解转成可执行的多项式运算。本节令 A 是 T 在一组基 $\mathcal{B}=(v_1,\dots,v_n)$ 下的矩阵，并令 $R=\mathbb{F}[t]$。

定理 10.15 ($tI-A$ 的表示)**.** 作为 $\mathbb{F}[t]$-模，有

$$V_T\cong\operatorname{coker}(tI-A)=\mathbb{F}[t]^n/(tI-A)\mathbb{F}[t]^n.$$

证明. 用 $\iota:\mathbb{F}^n\to V$ 表示坐标同构，即 $\iota(c_1,\dots,c_n)=\sum_i c_iv_i$，所以 $\iota(Ac)=T\iota(c)$。将多项式向量写成有限和 $q(t)=\sum_{k=0}^{d}t^kq_k$，其中 $q_k\in\mathbb{F}^n$，定义

$$E:\mathbb{F}[t]^n\to V_T,\qquad E(q)=\sum_{k=0}^{d}T^k\iota(q_k).$$

多项式的系数唯一，故定义无歧义。E 保持加法和 $\mathbb{F}$-数乘；又

$$E(tq)=\sum_k T^{k+1}\iota(q_k)=TE(q),$$

归纳并作有限线性组合可得 $E(fq)=f(T)E(q)$，故为 R-模同态。常数向量已映满 V，所以 E 满射。

对 $r(t)=\sum_k t^kr_k$，有

$$E(Ar)=\sum_k T^k\iota(Ar_k)=\sum_k T^{k+1}\iota(r_k)=E(tr).$$

故 $E((tI-A)r)=0$，于是 $(tI-A)R^n\subseteq\ker E$。

为证反向包含，在多项式矩阵环中直接计算望远镜和：

$$\begin{aligned}(tI-A)\sum_{j=0}^{k-1}t^{k-1-j}A^j&=\sum_{j=0}^{k-1}t^{k-j}A^j-\sum_{j=0}^{k-1}t^{k-1-j}A^{j+1}\\&=t^kI-A^k\qquad(k\ge1).\end{aligned}$$

因此

$$q(t)-\sum_{k=0}^{d}A^{k}q_{k}=(tI-A)\sum_{k=1}^{d}\sum_{j=0}^{k-1}t^{k-1-j}A^{j}q_{k}.$$

若 $q\in\ker E$，则

$$\iota\Big(\sum_{k}A^{k}q_{k}\Big)=\sum_{k}T^{k}\iota(q_{k})=E(q)=0.$$

ι 单射，故上式中的常数向量 $\sum_k A^k q_k=0$，从而 $q\in(tI-A)R^n$。这证明核的反向包含。对满同态 E 应用第一同构定理即得结论。 □

定理 10.16 (多项式矩阵的 Smith 形)**.** 对 $B\in M_{m\times n}(\mathbb{F}[t])$，可以通过交换行列、用非零常数乘一行或一列、把一行或一列的多项式倍加到另一行或列，将其化为

$$D=\operatorname{diag}(s_1,\dots,s_r,0,\dots),\qquad s_i\ \text{首一},\qquad s_1\mid\cdots\mid s_r.$$

所有这些操作均可逆。

证明. 每种操作的逆分别为再次交换、乘以非零常数的逆、以及加上相反的多项式倍，故均可逆。对 $\min(m,n)$ 归纳。矩阵全零时结束；否则交换行列将非零项移到左上角，并除以首系数使其首一，记为 s。

若第一列有元素 b 不被 s 整除，以多项式除法写 $b=qs+r$，$0\le\deg r<\deg s$。将对应行减去第一行的 q 倍，再交换两行，就将 $r\neq0$ 移到左上角。重新首一化后，其次数仍严格小于原来 s 的次数。第一行同样用列变换处理。若第一列或第一行中的某项被 s 整除，则减去相应倍数便可将其清零，而不改变左上角。

当第一行列除左上角外均已为零，矩阵成为 $\begin{pmatrix}s&0\\0&B_1\end{pmatrix}$。若余下块有元素 b 不被 s 整除，将它所在行加到第一行，使第一行出现 b。再用列除法 $b=qs+r$ 以及交换列，把非零余数 r 移到左上角，次数再次严格下降。每次次数下降后重新开始检查第一行列。

非负整数次数不可能无限严格下降；两次下降之间仅检查有限个元素，每个可整除元素都用一次操作清零。因此最终得到块对角形，且 s 整除余下块的每个元素。对余下块归纳。其内部初等变换所得每个元素仍为原来元素的多项式线性组合，故仍被 s 整除；特别地 s 整除归约后的第一个非零对角元。结合归纳得到的余下整除链，完成整个矩阵的归约。 □

对 $B=tI-A$，行列式为首一的 n 次多项式，因而非零。Smith 形中不可能有零对角元，否则它的行列式为零；可逆行列变换只使行列式乘以非零常数，矛盾。余核在这些变换下保持同构：若 $D=PBQ$，则 Q 满射给出 $\operatorname{im}D=P(\operatorname{im}B)$，于是 $[x]\mapsto[Px]$ 是余核间的同构。再逐坐标取商，得到

$$V_T\cong\bigoplus_{i=1}^{n}R/(s_i).$$

删除 $s_i=1$ 的零直和项后，余下的首一多项式就是 T 的不变因子。

定义 10.17 (子式理想与行列式因子)**.** 给定矩阵 B，其 j 阶子式是从 B 选出 j 行、j 列所得方阵的行列式。由全部 j 阶子式生成的理想记为 $J_j(B)$。在 $R=\mathbb{F}[t]$ 中，此理想的首一生成元记为 $\Delta_j(B)$；若理想为零，取 $\Delta_j(B)=0$，并约定 $\Delta_0(B)=1$。Δ_j 也就是这些子式的首一最大公因式。

引理 10.18 (子式理想不随初等变换改变)**.** 上述可逆初等行列变换均保持 $J_j(B)$。

证明. 先考虑行操作。交换行只使选择的行重新排列，并使相关行列式至多改变符号；用非零常数乘一行使包含该行的子式乘以单位，故这些变换不会改变生成的理想。

设第 a 行换成第 a 行加上 h 倍第 b 行，$a \neq b$。若一个子式不含第 a 行，它不变。若同时含第 a, b 行，由行列式对行的线性与重复行为零，得到

$$\det(\dots, \mathrm{row}_a + h\,\mathrm{row}_b, \dots, \mathrm{row}_b, \dots) = \det(\dots, \mathrm{row}_a, \dots, \mathrm{row}_b, \dots).$$

若包含第 a 行而不包含第 b 行，展开得到原子式加上 h 倍另一个子式；后一个子式使用第 b 行替换第 a 行，若需重排行序，只差一个符号。因此新矩阵的每个 j 阶子式都属于原来的 $J_j(B)$，即 $J_j(B_{\mathrm{new}}) \subseteq J_j(B)$。逆操作是加上 $-h$ 倍第 b 行，同样论证给出反向包含，故相等。

列操作使用行列式对列的相同性质，逐字作对应计算即可：目标列不被选中时子式不变；两列均选中时新增重复列项为零；仅选中目标列时新增项为另一子式的 $\pm h$ 倍。逆操作再次给出反向包含。 □

命题 10.19 (用各阶子式恢复不变因子)**.** *若 Smith 形非零项为 $s_1 \mid \cdots \mid s_r$，均首一，则*

$$\Delta_j(B) = s_1 \cdots s_j \quad (1 \le j \le r), \qquad s_j = \frac{\Delta_j(B)}{\Delta_{j-1}(B)}.$$

证明. 由前引理，只须对对角矩阵 D 计算。一个非零 j 阶子式必须选择相同的行列指标，且均在 $1, \dots, r$ 内，故形如

$$s_{i_1} \cdots s_{i_j}, \qquad 1 \le i_1 < \cdots < i_j \le r.$$

因 $i_\ell \ge \ell$，整除链给出 $s_\ell \mid s_{i_\ell}$，所以 $s_1 \cdots s_j$ 整除每个这样的乘积。另一方面，选前 j 行列就得到子式 $s_1 \cdots s_j$ 本身，故它也属于子式理想。因此该理想恰由此乘积生成，首一条件给出第一式，逐项相除得到第二式。 □

例 10.20 (由子式算出有理标准形)**.** 在 $\mathbb{Q}^3$ 上令 $A = \mathrm{diag}(1, 1, 2)$。其表示矩阵为

$$B = tI - A = \mathrm{diag}(t-1, t-1, t-2).$$

由 $(t-1) - (t-2) = 1$ 得 $\Delta_1 = 1$。非零二阶子式为

$$(t-1)^2, \qquad (t-1)(t-2), \qquad (t-1)(t-2),$$

它们的最大公因式为 $\Delta_2 = t - 1$。最后

$$\Delta_3 = \det B = (t-1)^2(t-2).$$

因此 $s_1 = 1$，$s_2 = t - 1$，$s_3 = (t-1)(t-2)$。删除单位因子后，不变因子为

$$d_1 = t - 1, \qquad d_2 = t^2 - 3t + 2,$$

有理标准形为

$$C(d_1) \oplus C(d_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}.$$

可以直接找到对应基核对计算。令 $u = e_2 + e_3$，则 $Au = e_2 + 2e_3$；e_1, u, Au 无关，因为后两列在 $\mathrm{span}(e_2, e_3)$ 中的坐标行列式为 1。再算

$$A^2 u = e_2 + 4e_3 = -2u + 3Au.$$

所以在基 (e_1, u, Au) 下，A 的矩阵正是上式。这里 $m_T = (t-1)(t-2)$，$\chi_T = (t-1)^2(t-2)$：其中重复的特征值并未强迫出现长度大于一的 Jordan 链。

10.5　互素因子给出的投影与主分解

假设最小多项式分解为

$$m_T = f_1 f_2 \cdots f_s, \qquad (f_i, f_j) = (1)\ (i \neq j),$$

其中各 f_i 首一且非常数。令 $M_i = m_T/f_i$。由唯一分解，$(M_i, f_i) = (1)$，选多项式 a_i, b_i 使

$$a_i M_i + b_i f_i = 1.$$

定义 $e_i = a_i M_i$，以及线性变换 $P_i = e_i(T)$。

定理 10.21 (多项式投影)**.** 上述变换满足

$$\sum_i P_i = I, \quad P_i^2 = P_i, \quad P_i P_j = 0\ (i \neq j), \quad \operatorname{im} P_i = \ker f_i(T).$$

因此

$$V = \bigoplus_{i=1}^{s} \ker f_i(T).$$

这些投影与 *Bézout* 系数的选取无关。

证明. 由定义，$e_i \equiv 1 \pmod{f_i}$，且 $e_i \equiv 0 \pmod{f_j}$ 对 $j \neq i$ 成立。故每个 f_j 都整除

$$\sum_i e_i - 1, \qquad e_i^2 - e_i, \qquad e_i e_\ell \quad (i \neq \ell).$$

互素多项式各自整除一个多项式时，其乘积也整除它：对两个因子，若 $h = fu$ 且 $g \mid h$，取 $af + bg = 1$ 得

$$u = afu + bgu = ah + bgu,$$

所以 $g \mid u$，从而 $fg \mid h$；对有限多个因子归纳即可。因此 m_T 整除上面三个多项式。代入 T，使用 $m_T(T) = 0$ 得三条投影等式。

又因 $f_i e_i = a_i m_T$，有 $f_i(T) P_i = 0$，故 $\operatorname{im} P_i \subseteq \ker f_i(T)$。若 $v \in \ker f_i(T)$，Bézout 等式给出

$$P_i v = (I - b_i(T) f_i(T))v = v,$$

所以 $v \in \operatorname{im} P_i$，反向包含成立。对于任意 $v \in V$，

$$v = \sum_i P_i v$$

给出所需分解。若 $\sum_i v_i = 0$，其中 $v_i \in \operatorname{im} P_i$，对该等式作用 P_j，由 $P_j v_i = 0$（$i \neq j$）与 $P_j v_j = v_j$ 得 $v_j = 0$。因此和为直和。

若换一组 Bézout 系数得到 e_i'，则 $e_i - e_i'$ 对每个 f_j 的余数均为零，故被 m_T 整除，代入后 $e_i(T) = e_i'(T)$。投影因而不依赖系数选择。 □

定义 10.22 (广义特征子空间)**.** 若 $\lambda \in \mathbb{F}$ 且 $(t - \lambda)^a$ 是 m_T 中对应的一次因子幂，定义

$$G_\lambda = \ker(T - \lambda I)^a.$$

称为 λ 的广义特征子空间。

若 $m_T=\prod_{i=1}^{s}(t-\lambda_i)^{a_i}$，其中 λ_i 两两不同，则这些因子互素，上一定理给出

$$V=\bigoplus_{i=1}^{s} G_{\lambda_i}.$$

因此每个向量能唯一分成属于不同广义特征值的分量；P_i 将各分量明确地提取出来。

例 10.23 (两条特征方向的投影)**.** 若 $m_T=(t-\alpha)(t-\beta)$ 且 $\alpha\neq\beta$，则

$$e_\alpha(t)=\frac{t-\beta}{\alpha-\beta},\qquad e_\beta(t)=\frac{t-\alpha}{\beta-\alpha}.$$

有 $e_\alpha+e_\beta=1$，所以

$$v=\frac{T-\beta I}{\alpha-\beta}v+\frac{T-\alpha I}{\beta-\alpha}v.$$

第一项属于 $\ker(T-\alpha I)$，因为

$$(T-\alpha I)\frac{T-\beta I}{\alpha-\beta}v=\frac{m_T(T)v}{\alpha-\beta}=0;$$

第二项同理属于 $\ker(T-\beta I)$。这给出了特征分量的计算式，而无须先解出特征向量再进行坐标分解。

10.6 从素幂循环模得到 Jordan 链

定义 10.24 (Jordan 块与 Jordan 链)**.** 对 $a\geq 1$，令 $J_a(\lambda)$ 为对角线全为 λ、紧邻对角线上方全为 1、其余项为零的 $a\times a$ 矩阵。相应的一列向量 $e_1,\dots,e_a$ 称为长度 a 的 Jordan 链，若

$$(T-\lambda I)e_1=0,\qquad (T-\lambda I)e_j=e_{j-1}\quad (2\leq j\leq a),$$

且这列向量线性无关。

引理 10.25 (一个素幂循环模中的链)**.** 在 $\mathbb{F}[t]/((t-\lambda)^a)$ 中令 $u=[1]$，并取

$$e_j=(t-\lambda)^{a-j}u\qquad (1\leq j\leq a).$$

这是一组 $\mathbb{F}$-基，乘以 t 的矩阵是 $J_a(\lambda)$。

证明. 任意次数小于 a 的多项式都能唯一写成

$$c_0+c_1(t-\lambda)+\cdots+c_{a-1}(t-\lambda)^{a-1}.$$

存在性可将 $t=(t-\lambda)+\lambda$ 代入通常的幂基展开；唯一性则因 $(t-\lambda)^r$ 首一且次数为 r，若非零线性组合的最高非零系数为 c_r，其最高次项为 $c_r t^r$，不可能为零。带余除法说明这些多项式的类张成整个商模。若其线性组合在商模中为零，该组合被次数 a 的 $(t-\lambda)^a$ 整除，而次数小于 a，故组合为零，再由上述唯一性得到全部系数为零。这证明所列基的张成与无关。

令 N 为乘以 $t-\lambda$ 的作用，则

$$Ne_1=(t-\lambda)^a u=0,\qquad Ne_j=(t-\lambda)^{a-j+1}u=e_{j-1}\quad (j\geq 2).$$

因此 $te_1=\lambda e_1$，$te_j=\lambda e_j+e_{j-1}$。将这些像按列写出，得到 $J_a(\lambda)$。 □

定义 10.26 (代数闭域)**.** 域 $\mathbb{F}$ 称为代数闭域，若 $\mathbb{F}[t]$ 中每个正次数多项式都在 $\mathbb{F}$ 中有根。由因子定理（命题 8.26）反复取出一次因子，次数每次降低 1，可知这也等价于每个正次数多项式都在 $\mathbb{F}$ 上分裂为一次因子的乘积。

定理 10.27 (Jordan 标准形)**.** *若 m_T 在 $\mathbb{F}$ 上分裂为一次因子的乘积，则 V 有一组由 Jordan 链拼成的基，T 在此基下的矩阵为 Jordan 块的直和。给定每个 λ，相应块的大小及重数唯一。特别地，在代数闭域上每个有限维线性变换都有 Jordan 标准形。*

证明. V_T 的每个初等因子必须整除某个 d_i，而 $d_i \mid m_T$。分裂假设说明各初等因子均为 $(t-\lambda)^a$。由 PID 模的初等因子分解，

$$V_T \cong \bigoplus_{\lambda} \bigoplus_{j} \mathbb{F}[t]/((t-\lambda)^{a_{\lambda,j}}).$$

对每个直和项取前引理的基，并经模同构送回 V。模同构保持 t 的作用，因此送回的向量满足同样的 Jordan 链等式。合并后成为整个 V 的基，矩阵为相应 Jordan 块的直和。初等因子及其重数的唯一性确定块大小和重数。

若 $\mathbb{F}$ 代数闭，按定义每个正次数多项式有根；反复用一次因子除法并使次数下降，就将 m_T 分裂为一次因子的乘积，因而上述结论适用。 □

注 10.28 (Axler 的构造与多项式模的构造)**.** Axler 在 [10] 第 8C 节、定理 8.45 中直接研究幂零变换：先取一条最长的迭代链，再借助线性泛函构造不变补空间，最后对维数归纳。这里从 PID 模结构出发，每个 $\mathbb{F}[t]/((t-\lambda)^a)$ 自带一条长度 a 的链。两种证明都要完成同一个步骤：找到一条链之后，还须证明其余向量能安排在与它互补的不变空间中。模的直和分解给出了这个补空间及其唯一的块大小数据。

图 10.1 中，N 表示由三条链规定的幂零变换：$Ne_3 = e_2$，$Ne_2 = e_1$，$Nf_2 = f_1$，且 $Ne_1 = Nf_1 = Ng_1 = 0$。六个向量合为一组基。

$$0 \xleftarrow{N} e_1 \xleftarrow{N} e_2 \xleftarrow{N} e_3$$

$$0 \xleftarrow{N} f_1 \xleftarrow{N} f_2$$

$$0 \xleftarrow{N} g_1$$

$$\ker N \qquad \ker N^2 \qquad \ker N^3$$

图 10.1: 长度为 $3, 2, 1$ 的三条幂零链。每个核由其标签所在列及左侧全部向量张成。

10.7 怎样从核的维数读出块大小

命题 10.29 (核维数与 Jordan 块)**.** *固定特征值 λ，在 G_λ 上令 $N = (T-\lambda I)|_{G_\lambda}$，并定义*

$$b_j = \dim \ker N^j \quad (j \geq 0), \qquad N^0 = I, \quad b_0 = 0.$$

若相应块大小为 $a_1, \dots, a_r$，则

$$b_j = \sum_{i=1}^{r} \min(j, a_i), \qquad b_j - b_{j-1} = \#\{i : a_i \geq j\}.$$

大小恰为 j 的块数为 $2b_j - b_{j-1} - b_{j+1}$。

证明. 在一条长度为 a 的链中，反复使用 $Ne_i = e_{i-1}$ 得

$$N^j e_i = \begin{cases} 0, & i \le j, \\ e_{i-j}, & i > j. \end{cases}$$

因此对 $v = \sum_{i=1}^{a} c_i e_i$，

$$N^j v = \sum_{i=j+1}^{a} c_i e_{i-j} = 0 \iff c_{j+1} = \cdots = c_a = 0,$$

这里 $j \ge a$ 时右侧没有条件。该块中核为 $\operatorname{span}(e_1, \dots, e_{\min(j,a)})$，维数为 $\min(j, a)$。不同链构成不变直和，故 N^j 的核也是这些核的直和，得到第一式。

逐项相减，$\min(j,a) - \min(j-1,a)$ 在 $a \ge j$ 时等于 1，在 $a < j$ 时等于 0，故第二式成立。长度恰为 j 的链数是长度至少 j 的链数减去长度至少 $j+1$ 的链数，即

$$(b_j - b_{j-1}) - (b_{j+1} - b_j) = 2b_j - b_{j-1} - b_{j+1}.$$

□

推论 10.30 (三种重数数据). *λ 在 χ_T 中的重数是 $\dim G_\lambda$；特征子空间 $\ker(T - \lambda I)$ 的维数是 λ 的 Jordan 块数；λ 在 m_T 中的指数是最大 Jordan 块的大小。*

证明. 每个 $J_a(\lambda)$ 为上三角矩阵，其特征多项式为 $(t-\lambda)^a$；故所有 λ 块贡献的指数为大小之和，即 $\dim G_\lambda$。在每条 λ 链上，$T - \lambda I$ 的核恰为第一个向量张成的一维空间。

若 $\mu \ne \lambda$，在一个 μ 块上写 $T - \lambda I = (\mu - \lambda)I + N$，其中 $N^a = 0$。其逆显式为

$$\big((\mu-\lambda)I + N\big)^{-1} = \frac{1}{\mu - \lambda} \sum_{j=0}^{a-1} \left(-\frac{N}{\mu - \lambda} \right)^j,$$

因为乘开后有限几何和只剩 $I - (-N/(\mu-\lambda))^a = I$。所以其他块没有核的贡献。最后 m_T 是各初等因子的最小公倍式；这由“消去直和当且仅当消去各项”得到。因此其中 $(t-\lambda)$ 的指数是对应各块指数的最大值。 □

例 10.31 (只由核计算恢复六维幂零标准形). 设 $\dim V = 6$，$N^3 = 0$，并已算得

$$\dim \ker N = 3, \qquad \dim \ker N^2 = 5, \qquad \dim \ker N^3 = 6.$$

于是 $b_0 = 0, b_1 = 3, b_2 = 5, b_3 = b_4 = 6$。长度至少为 $1, 2, 3$ 的链数依次为

$$b_1 - b_0 = 3, \quad b_2 - b_1 = 2, \quad b_3 - b_2 = 1.$$

长度恰为 $1, 2, 3$ 的链数依次为 $3 - 2 = 1$、$2 - 1 = 1$、$1 - 0 = 1$。所以

$$[N] \sim J_3(0) \oplus J_2(0) \oplus J_1(0), \quad m_N = t^3, \quad \chi_N = t^6.$$

图 10.1 中每次新进入核的向量数分别为 $3, 2, 1$，正好给出这三个维数差。

10.8　实际换基、迭代与标量域

例 10.32 (平移算子的完整 Jordan 换基)**.** 在实向量空间 $V=\{a+bx+cx^2 : a,b,c\in\mathbb{R}\}$ 上，令

$$T(p)(x)=p(x+1),\qquad N=T-I.$$

在基 $\mathcal{B}=(1,x,x^2)$ 下，

$$[T]_{\mathcal{B}}=A=\begin{pmatrix}1&1&1\\0&1&2\\0&0&1\end{pmatrix},\qquad N(1)=0,\quad N(x)=1,\quad N(x^2)=2x+1.$$

从最高次向量 $u=x^2/2$ 开始，逐步计算

$$Nu=x+\tfrac12,\qquad N^2u=1,\qquad N^3u=0.$$

所以取

$$e_1=1,\quad e_2=x+\tfrac12,\quad e_3=\tfrac{x^2}{2},\qquad Ne_1=0,\quad Ne_2=e_1,\quad Ne_3=e_2.$$

三者最高次数分别为 $0,1,2$ 且首系数非零，故无关；也可从过渡矩阵直接验证：

$$P=\begin{pmatrix}1&1/2&0\\0&1&0\\0&0&1/2\end{pmatrix},\quad P^{-1}=\begin{pmatrix}1&-1/2&0\\0&1&0\\0&0&2\end{pmatrix}.$$

各列是新基在旧基中的坐标，且

$$AP=\begin{pmatrix}1&3/2&1/2\\0&1&1\\0&0&1/2\end{pmatrix}=P\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}.$$

故 $P^{-1}AP=J_3(1)$。令 $\binom{k}{2}=k(k-1)/2$。由于 $N^3=0$，可以归纳计算

$$T^k=(I+N)^k=I+kN+\binom{k}{2}N^2\qquad(k\in\mathbb{N}).$$

$k=0$ 时两边均为 I；若 k 时成立，则

$$\begin{aligned}T^{k+1}&=(I+N)\left(I+kN+\binom{k}{2}N^2\right)\\&=I+(k+1)N+\left(k+\binom{k}{2}\right)N^2+\binom{k}{2}N^3\\&=I+(k+1)N+\binom{k+1}{2}N^2.\end{aligned}$$

末步使用 $k+k(k-1)/2=k(k+1)/2$ 以及 $N^3=0$。在原基下，N 的矩阵为 $A-I$，其平方只有第 $(1,3)$ 项为 2，所以

$$A^k=\begin{pmatrix}1&k&k^2\\0&1&2k\\0&0&1\end{pmatrix}.$$

将其作用于坐标列 $(a,b,c)^t$，得到 $(a+bk+ck^2,\,b+2ck,\,c)^t$，恰好是 $p(x+k)$ 的系数。Jordan 链的计算与原来的平移规则逐项一致。

命题 10.33 (可对角化的多项式判别)**.** T 在 $\mathbb{F}$ 上可对角化，当且仅当 m_T 是若干互不相同的一次因子的乘积。

证明. 若 T 有特征向量基，设出现的不同特征值为 $\lambda_1,\dots,\lambda_s$。任意多项式 f 在一个属于 λ_i 的基向量 v 上满足

$$f(T)v=\sum_{j} a_j T^j v=\sum_{j} a_j \lambda_i^j v=f(\lambda_i)v.$$

所以 $f(T)=0$ 当且仅当所有 $f(\lambda_i)=0$。因一次因子互素，根与因子定理及互素乘积整除性给出

$$f(T)=0 \iff \prod_{i}(t-\lambda_i)\mid f.$$

比较首一生成元，得到 $m_T=\prod_i(t-\lambda_i)$。

反之，若 m_T 是此乘积，主分解定理给出

$$V=\bigoplus_{i}\ker(T-\lambda_i I).$$

在每个核中取基，合并后成为整个 V 的基，且每个基向量均为特征向量。因此 T 可对角化。 □

例 10.34 (实旋转与复特征方向)**.** 在 $\mathbb{R}^2$ 上取 $T(x,y)=(-y,x)$。有 $T^2=-I$，而 $e_1, Te_1=e_2$ 无关，故

$$V_T\cong\mathbb{R}[t]/(t^2+1),\qquad m_T=\chi_T=t^2+1,\qquad [T]=C(t^2+1)=\begin{pmatrix}0&-1\\1&0\end{pmatrix}.$$

由于 t^2+1 在 $\mathbb{R}$ 上不可约，这已经是实数域内的有理标准形，且没有实特征向量。

把同一矩阵的系数视为复数，并令其作用于 $\mathbb{C}^2$，则

$$t^2+1=(t-i)(t+i).$$

两个一次因子不同，故可对角化。具体地，

$$v_+=(1,-i)^{t},\quad v_-=(1,i)^{t},\qquad Tv_+=iv_+,\quad Tv_-=-iv_-.$$

过渡矩阵 $P=(v_+\ v_-)$ 的行列式为 $2i\neq 0$，所以

$$P^{-1}[T]P=\operatorname{diag}(i,-i).$$

改变标量域后，向量空间、可用的基以及多项式的不可约分解同时改变；原来的一块二次循环模于是分成两条一次特征方向。

文献：PID 模的相容基构造参照 [13] 第 2 节，有限生成模分解见其第 4 节。有理标准形与初等因子方法亦见 [9]。Axler 的幂零链构造见 [10] 第 8C 节。本文从模结构独立推导伴随矩阵行列式、Cayley–Hamilton、主分解投影及核维数公式。

11　多重线性代数

11.1　从双线性运算到张量积

本章固定交换有单位环 R，并假定 $1_R \neq 0_R$。所有模均为 R-模，所有线性映射均指 R-模同态。涉及域时另用 $\mathbb{F}$ 表示。本章的 R 允许有零因子。

线性映射保留一个输入中的线性组合。若一项运算接受两个输入，例如实数乘法 $(x,y)\mapsto xy$，它对每个输入分别线性，却不是从 $\mathbb{R}^2$ 到 $\mathbb{R}$ 的线性映射。事实上，把两个输入同时乘以 2，乘积变成 $4xy$；线性映射则应变成原值的 2 倍。多重线性代数研究分别线性的运算，以及能够记录这些运算的线性空间或模。

定义 11.1 (多重线性映射)**.** 设 $k\geq 1$，$M_1,\dots,M_k,P$ 为模。映射 $b: M_1\times\cdots\times M_k\to P$ 称为 k-线性映射，若固定其余变量后，它对每一个变量都是线性的。$k=2$ 时称为双线性映射。

例如，设 $b: M\times N\to P$ 为双线性映射，$a_1,a_2,c_1,c_2\in R$，$e_1,e_2\in M$，$f_1,f_2\in N$。则

$$\begin{aligned} b(a_1e_1+a_2e_2,\,c_1f_1+c_2f_2) &= a_1c_1\,b(e_1,f_1)+a_1c_2\,b(e_1,f_2)\\ &\quad + a_2c_1\,b(e_2,f_1)+a_2c_2\,b(e_2,f_2), \end{aligned}$$

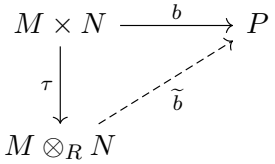
若两个模具有有限基，这样的展开表明，双线性映射由所有基元对上的取值确定。我们希望构造一个模，把每个输入对记录成一个元素，并把展开规则变成这个模内的等式。

Conrad 在 [14] 第 1–3 节先从基元对的形式乘积入手，再用泛性质消除对基的依赖。这里直接把上述要求写成定义，随后证明这样的模确实存在。

定义 11.2 (张量积)**.** 模 M,N 的张量积是模 $M\otimes_R N$，连同双线性映射

$$\tau: M\times N\longrightarrow M\otimes_R N,\qquad (m,n)\longmapsto m\otimes n,$$

满足：对任意模 P 和双线性映射 $b: M\times N\to P$，存在唯一线性映射 $\widetilde{b}: M\otimes_R N\to P$ 使 $\widetilde{b}\circ\tau = b$。



这个性质规定了从张量积出发的线性映射：给出它，等价于给出原来两个变量上的双线性映射。符号 $m\otimes n$ 表示 $\tau(m,n)$，不是原集合中的有序对 (m,n)。

由 τ 的双线性性，立即有

$$\begin{aligned} (m+m')\otimes n &= m\otimes n + m'\otimes n,\\ m\otimes(n+n') &= m\otimes n + m\otimes n',\\ (rm)\otimes n &= r(m\otimes n) = m\otimes(rn). \end{aligned}$$

对两个变量分别代入零，得到

$$0\otimes n=(0+0)\otimes n=0\otimes n+0\otimes n,\qquad m\otimes 0=m\otimes(0+0)=m\otimes 0+m\otimes 0.$$

在加法群中分别消去右侧的一项，便得 $0\otimes n=0$ 与 $m\otimes 0=0$。

11.2 存在性、唯一性与生成元

定理 11.3 (张量积的构造)**.** 令 F 为以集合 $M\times N$ 为基的自由模，基元记为 $e_{m,n}$。令 L 为下列元素生成的子模，其中 $m,m'\in M$，$n,n'\in N$，$r\in R$:

$$\begin{aligned}&e_{m+m',n}-e_{m,n}-e_{m',n},\\&e_{m,n+n'}-e_{m,n}-e_{m,n'},\\&e_{rm,n}-re_{m,n},\qquad e_{m,rn}-re_{m,n}.\end{aligned}$$

则 F/L 与映射 $(m,n)\mapsto e_{m,n}+L$ 满足张量积的泛性质。

证明. 四类关系在商模中为零，故映射 $\tau(m,n)=e_{m,n}+L$ 为双线性映射。给定双线性映射 $b:M\times N\to P$，由自由模的泛性质得到唯一线性映射 $\widehat{b}:F\to P$，满足 $\widehat{b}(e_{m,n})=b(m,n)$。分别将四类生成元代入，得到

$$\begin{aligned}\widehat{b}(e_{m+m',n}-e_{m,n}-e_{m',n})&=b(m+m',n)-b(m,n)-b(m',n)=0,\\\widehat{b}(e_{m,n+n'}-e_{m,n}-e_{m,n'})&=b(m,n+n')-b(m,n)-b(m,n')=0,\\\widehat{b}(e_{rm,n}-re_{m,n})&=b(rm,n)-rb(m,n)=0,\\\widehat{b}(e_{m,rn}-re_{m,n})&=b(m,rn)-rb(m,n)=0.\end{aligned}$$

因为核为子模，且包含全部生成元，所以 $L\subseteq\ker\widehat{b}$。模的因子化定理给出线性映射 $\widetilde{b}:F/L\to P$，且 $\widetilde{b}(e_{m,n}+L)=b(m,n)$。

商模 F/L 由所有 $e_{m,n}+L$ 生成。因此另一线性映射若在这些元素上同样取值为 $b(m,n)$，便在每个有限线性组合上与 $\widetilde{b}$ 相等。这证明所需唯一性。 □

定理 11.4 (泛性质所确定的唯一性)**.** 若 (T,τ) 和 (T',τ') 都是 M,N 的张量积，则存在唯一线性同构 $u:T\to T'$ 使 $u\tau=\tau'$。

证明. 将 τ' 作为取值于 T' 的双线性映射，由 T 的泛性质得到 $u:T\to T'$。交换两者得到 $v:T'\to T$。于是

$$vu\tau=v\tau'=\tau=\mathrm{id}_T\tau.$$

泛性质的唯一性给出 $vu=\mathrm{id}_T$。另一复合满足

$$uv\tau'=u\tau=\tau'=\mathrm{id}_{T'}\tau',$$

由 T' 的泛性质中的唯一性得到 $uv=\mathrm{id}_{T'}$。u 因而为同构，其唯一性仍由 T 的泛性质给出。 □

推论 11.5 (纯张量生成张量积)**.** 形如 $m\otimes n$ 的元素称为纯张量。每个张量均可写成有限和 $\sum_{i=1}^{s}m_i\otimes n_i$。两个从张量积出发的线性映射相等，当且仅当它们在每个纯张量上相等。

证明. 在 F/L 的构造中，基元的像生成整个商模；系数可吸收到第一个因子中。任一满足泛性质的对象都由上一同构与此构造相容地识别，所以也由纯张量生成。最后一句来自线性映射保持有限和。 □

在纯张量上指定公式时，仍须证明公式来自双线性映射。纯张量通常不是一组基，它们之间有关系。例如 $m\otimes 0=0$，所以任意指定 $m\otimes 0$ 的像为一个非零元素，都不可能延拓为线性映射。

11.3 映射的张量积与括号的处理

命题 11.6 (映射的张量积). *若 $f:M\to M'$、$g:N\to N'$ 线性，则存在唯一线性映射*

$$f\otimes g: M\otimes_R N\longrightarrow M'\otimes_R N',\qquad m\otimes n\longmapsto f(m)\otimes g(n).$$

若 $f':M'\to M''$、$g':N'\to N''$ 线性，则

$$(f'\otimes g')(f\otimes g)=(f'f)\otimes(g'g),\qquad \mathrm{id}_M\otimes\mathrm{id}_N=\mathrm{id}_{M\otimes_R N}.$$

证明. $(m,n)\mapsto f(m)\otimes g(n)$ 对每个变量分别线性，故可由泛性质延拓。两个复合公式的两边均为线性映射，并且在 $m\otimes n$ 上分别取相同的值 $f'f(m)\otimes g'g(n)$ 和 $m\otimes n$，所以相等。 □

命题 11.7 (对称、单位与结合). *以下公式给出线性同构：*

$$\begin{aligned} M\otimes_R N&\longrightarrow N\otimes_R M, & m\otimes n&\longmapsto n\otimes m,\\ R\otimes_R M&\longrightarrow M, & r\otimes m&\longmapsto rm,\\ (M\otimes_R N)\otimes_R P&\longrightarrow M\otimes_R(N\otimes_R P), & (m\otimes n)\otimes p&\longmapsto m\otimes(n\otimes p).\end{aligned}$$

证明. 第一式由双线性映射 $(m,n)\mapsto n\otimes m$ 得到；交换两次固定所有纯张量，故为恒等。第二式由标量乘法的双线性性得到。其逆为 $m\mapsto 1\otimes m$，因为 $1\otimes rm=r\otimes m$。

对第三式，先说明两边均使三线性映射线性化。给定三线性映射 $b:M\times N\times P\to Q$。固定 $p\in P$，先延拓前两个变量，得到线性映射 $b_p:M\otimes_R N\to Q$。若 $t=\sum_i m_i\otimes n_i$，则

$$b_p(t)=\sum_i b(m_i,n_i,p).$$

因此对固定的 t，这个值随 p 线性变化；它与 t 的表示无关，因为 b_p 已经是良定义映射。于是 $(t,p)\mapsto b_p(t)$ 为双线性映射，可再延拓到 $(M\otimes_R N)\otimes_R P$。唯一性由元素 $(m\otimes n)\otimes p$ 生成该模得到。先固定 m，对后两个变量作相同构造，可知 $M\otimes_R(N\otimes_R P)$ 具有同样的三线性泛性质。

把两边的标准三线性映射分别延拓，得到所示映射及反向映射；两复合固定所有三因子纯张量，所以互为逆。 □

逐次重复这个证明，就得到 k 个模的张量积：它使 k-线性映射唯一地线性化。任一括号排列都具有此性质，因为可按最外层括号把变量分成两组，先分别延拓，再延拓两组之间的双线性映射。所有重新加括号的映射均把同一输入组送到同一纯张量，故它们的复合彼此相容。以下通过这些指定同构省略括号，写作 $m_1\otimes\cdots\otimes m_k$。

11.4 张量的坐标、矩阵与对偶

定理 11.8 (自由模的张量基)**.** *若 M, N 的基分别为 $(e_i)_{i\in I}$、$(f_j)_{j\in J}$，则 $(e_i \otimes f_j)_{(i,j)\in I\times J}$ 是 $M \otimes_R N$ 的基。特别地，$R^m \otimes_R R^n \cong R^{mn}$。*

证明. 每个输入均为基元的有限线性组合，双线性展开说明这些张量张成。记 $\epsilon_i : M \to R$、$\eta_j : N \to R$ 为取相应基坐标的线性映射。对固定 (i,j)，函数 $(m,n) \mapsto \epsilon_i(m)\eta_j(n)$ 双线性，故给出线性泛函 $M \otimes_R N \to R$。它在 $e_a \otimes f_b$ 上的值为 $\delta_{ia}\delta_{jb}$，其中 δ_{ia} 在 $i = a$ 时为 1，否则为 0。对任一有限线性关系施加这些泛函，逐个系数都等于零，故线性无关。 □

例 11.9 (纯张量与矩阵)**.** 设 $V = \mathbb{F}^m, W = \mathbb{F}^n$。将 $\sum a_{ij} e_i \otimes f_j$ 对应于矩阵 (a_{ij})，便得到 $V \otimes_{\mathbb{F}} W$ 与 $m \times n$ 矩阵空间的线性同构。若 $v = \sum x_i e_i, w = \sum y_j f_j$，则 $v \otimes w$ 的矩阵为 $(x_i y_j) = x y^{\mathsf{T}}$。当 v, w 均非零时，所有列都是 x 的倍数，且至少一列非零，故秩为 1。反过来，秩为 1 的矩阵取一个非零列 x，其余各列都是 x 的倍数，从这些倍数构成 y，便写成 $x y^{\mathsf{T}}$。因此纯张量恰好对应秩至多为 1 的矩阵。

例如 $e_1 \otimes f_1 + e_2 \otimes f_2$ 的矩阵有一个二阶单位子块，秩为 2，所以它不是纯张量。这也解释了为何张量积必须允许纯张量的和。

例 11.10 (矩阵怎样作用在张量上)**.** 若 $f(e_j) = \sum_i a_{ij} e'_i$，$g(f_\ell) = \sum_k b_{k\ell} f'_k$，则

$$(f \otimes g)(e_j \otimes f_\ell) = \sum_{i,k} a_{ij} b_{k\ell}\, e'_i \otimes f'_k .$$

按先排列第一个指标、再排列第二个指标的顺序列基，矩阵的 $((i,k),(j,\ell))$ 项为 $a_{ij}b_{k\ell}$。这称为矩阵 $A = (a_{ij})$ 与 $B = (b_{k\ell})$ 的 Kronecker 积，记为 $A \otimes B$。例如

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 6 & 0 \\ 0 & 4 & 0 & 8 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

这里的块结构来自两个输入分别变换后的双线性展开。若只改变第一个因子，矩阵为 $A \otimes I$；只改变第二个因子则为 $I \otimes B$，其中 I 分别表示相应大小的单位矩阵。两种操作的复合等于 $A \otimes B$，因为它们在每个基张量上都给出 $f(e_j) \otimes g(f_\ell)$。

所有线性映射 $M \to N$ 组成的集合记为 $\operatorname{Hom}_R(M,N)$，逐点相加、数乘使它成为模；模公理逐点来自 N。双线性映射的集合 $\operatorname{Bil}_R(M,N;P)$ 同样逐点成为模。张量积的泛性质给出模同构

$$\operatorname{Hom}_R(M \otimes_R N, P) \cong \operatorname{Bil}_R(M,N;P).$$

正向把 h 送到 $h \circ \tau$，反向取唯一延拓。两者互逆由泛性质得到；它们保持加法和标量，因为等式可在每个输入对上逐点核对。

定理 11.11 (有限自由模的对偶与线性映射)**.** *定义对偶模 $V^* = \operatorname{Hom}_R(V, R)$。若 V 为有限自由模，W 为任意模，则有同构*

$$\Theta : V^* \otimes_R W \longrightarrow \operatorname{Hom}_R(V, W), \qquad \Theta(\varphi \otimes w)(v) = \varphi(v) w.$$

证明. $(\varphi, w) \mapsto (v \mapsto \varphi(v)w)$ 双线性，故得到 Θ。取 V 的基 $e_1, \dots, e_n$，坐标泛函为 $e_1^*, \dots, e_n^*$。定义

$$\Psi(T) = \sum_{i=1}^{n} e_i^* \otimes T(e_i).$$

它是线性映射，且对于 $v = \sum_i v_i e_i$，

$$\Theta\Psi(T)(v) = \sum_i v_i T(e_i) = T(v).$$

另一方面 $\varphi = \sum_i \varphi(e_i) e_i^*$，因为两边在每个 e_j 上相等。因此

$$\Psi\Theta(\varphi \otimes w) = \sum_i e_i^* \otimes \varphi(e_i)w = \varphi \otimes w.$$

纯张量生成，故另一复合也为恒等。虽然 Ψ 的公式使用基，它作为 Θ 的逆由 Θ 唯一确定。 □

11.5 直和、商模与循环模的计算

命题 11.12 (张量积保持直和)**.** *对任意模族 $(M_i)_{i\in I}$，有同构*

$$\left(\bigoplus_{i\in I} M_i\right) \otimes_R N \cong \bigoplus_{i\in I} (M_i \otimes_R N).$$

证明. 映射 $((m_i), n) \mapsto (m_i \otimes n)_i$ 双线性，且由于 (m_i) 有限支撑，其值在右侧直和中；延拓为线性映射 a。记 $\iota_i : M_i \to \bigoplus_j M_j$ 为坐标包含。对每个 i，映射 $\iota_i \otimes \mathrm{id}_N$ 已经良定义，因此可以定义

$$c((t_i)) = \sum_i (\iota_i \otimes \mathrm{id}_N)(t_i).$$

这也是有限和，给出反向线性映射。在右侧任意单坐标纯张量上，ac 是恒等。在左侧，

$$ca((m_i) \otimes n) = \sum_i \iota_i(m_i) \otimes n = (m_i) \otimes n.$$

两侧分别由这些元素生成，所以 a, c 互为逆。 □

命题 11.13 (张量积与商模)**.** *若 $U \le M$，令 $K \le M \otimes_R N$ 为所有 $u \otimes n$ $(u \in U, n \in N)$ 生成的子模，则*

$$(M/U) \otimes_R N \cong (M \otimes_R N)/K.$$

证明. $(m + U, n) \mapsto m \otimes n + K$ 与代表元无关，因为将 m 换成 $m + u$ 只增加 $u \otimes n \in K$。它双线性，给出正向映射。反向先由投影 $q : M \to M/U$ 得到 $q \otimes \mathrm{id}_N$，它将所有 $u \otimes n$ 送到零，因而下降到商模。两复合在所有纯张量或它们的商类上固定生成元，故为恒等。 □

命题 11.14 (循环模的张量积)**.** *若 I 为 R 的理想，IN 表示所有有限和 $\sum_j a_j n_j$ $(a_j \in I, n_j \in N)$ 的集合，则 IN 是子模，且*

$$(R/I) \otimes_R N \cong N/IN.$$

证明. 两个这样的有限和相减仍具此形式，乘以 $r\in R$ 后将 a_j 换成 $ra_j\in I$ 即可，故 IN 为子模。设 $r+I=r'+I$，则 $r-r'\in I$，故 $rn-r'n=(r-r')n\in IN$，所以

$$b(r+I,n)=rn+IN$$

与代表元无关。记商类为方括号，则

$$\begin{aligned} b([r]+[s],n)&=[(r+s)n]=[rn]+[sn],\\ b([r],n+n')&=[rn+rn']=[rn]+[rn'],\\ b(a[r],n)&=[arn]=a[rn],\qquad b([r],an)=[ran]=a[rn]. \end{aligned}$$

因此 b 双线性，诱导线性映射 $\alpha:(R/I)\otimes_R N\to N/IN$。线性映射 $n\mapsto(1+I)\otimes n$ 消去 IN，因为

$$(1+I)\otimes an=(a+I)\otimes n=0\qquad(a\in I).$$

于是它下降为线性映射 $\beta:N/IN\to(R/I)\otimes_R N$，$\beta(n+IN)=(1+I)\otimes n$。两复合分别满足

$$\begin{aligned} \alpha\beta(n+IN)&=\alpha((1+I)\otimes n)=n+IN,\\ \beta\alpha((r+I)\otimes n)&=\beta(rn+IN)=(1+I)\otimes rn=(r+I)\otimes n. \end{aligned}$$

第一式在所有商类上成立，第二式在生成整个张量积的所有纯张量上成立，所以两个复合均为恒等映射。 □

例 11.15 (整数系数的计算)**.** 设 $m,n\geq 1$，$d=\gcd(m,n)$。取 $N=\mathbb{Z}/m\mathbb{Z}$，则

$$(\mathbb{Z}/n\mathbb{Z})\otimes_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z})\cong(\mathbb{Z}/m\mathbb{Z})/n(\mathbb{Z}/m\mathbb{Z}).$$

由 Bézout 等式，$m\mathbb{Z}+n\mathbb{Z}=d\mathbb{Z}$，所以 $n(\mathbb{Z}/m\mathbb{Z})$ 正是 $d\mathbb{Z}/m\mathbb{Z}$。映射 $[a]_m\mapsto[a]_d$ 满射且核为该子群；第一同构定理给出

$$(\mathbb{Z}/n\mathbb{Z})\otimes_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z})\cong\mathbb{Z}/d\mathbb{Z}.$$

例如 $\mathbb{Z}/12\mathbb{Z}\otimes_{\mathbb{Z}}\mathbb{Z}/18\mathbb{Z}\cong\mathbb{Z}/6\mathbb{Z}$；而 $\mathbb{Z}/2\mathbb{Z}\otimes_{\mathbb{Z}}\mathbb{Z}/3\mathbb{Z}=0$。后一个例子中，若 $t=[1]_2\otimes[1]_3$，则 $2t=3t=0$，故 $t=3t-2t=0$；这个张量生成整个模，因而模为零。

11.6 扩张标量与平坦性

定义 11.16 (R-代数)**.** 结合有单位的 R-代数是一个结合有单位环 A，连同保持单位的环同态 $\iota:R\to A$，满足 $\iota(r)a=a\iota(r)$。以 $ra=\iota(r)a$ 定义标量乘法，A 成为 R-模，乘法对两个变量分别 R-线性。R-代数同态是满足 $h\iota_A=\iota_B$ 的保持单位的环同态；它因而也是 R-线性映射。

反过来，若一个 R-模具有结合的双线性乘法和单位元 1_A，则 $r\mapsto r1_A$ 是所需的中心环同态。乘法的双线性给出 $(r1_A)a=r(1_Aa)=ra=r(a1_A)=a(r1_A)$，也给出乘法保持性。

设 S 是交换 R-代数。张量积 $S\otimes_R M$ 上的 S-标量作用定义为

$$s'(s\otimes m)=(s's)\otimes m.$$

对固定 s'，右侧来自 R-线性映射 $s\mapsto s's$ 与 id_M 的张量积，故良定义。加法、结合律和单位律可在纯张量上核对，因而它使 $S\otimes_R M$ 成为 S-模，称为 M 的扩张标量。

命题 11.17 (扩张标量的泛性质)**.** *若 U 是 S-模，将其视为 R-模的方式为 $ru=\iota(r)u$。每个 R-线性映射 $f:M\to U$ 唯一延拓为 S-线性映射*

$$\widehat{f}:S\otimes_R M\to U,\qquad \widehat{f}(s\otimes m)=sf(m),$$

满足 $\widehat{f}(1\otimes m)=f(m)$。

证明. $(s,m)\mapsto sf(m)$ 为 R-双线性映射，故给出 R-线性延拓。公式表明它保持 S 的标量作用，因此 S-线性。任意满足条件的 S-线性映射都必须将 $s\otimes m=s(1\otimes m)$ 送到 $sf(m)$，所以唯一。 □

例 11.18 (从整数到有理数)**.** 由单位同构，$\mathbb{Q}\otimes_{\mathbb{Z}}\mathbb{Z}\cong\mathbb{Q}$。若 $x\in M$ 满足 $nx=0$，其中整数 $n\neq 0$，则对 $q\in\mathbb{Q}$ 有

$$q\otimes x=(n(q/n))\otimes x=(q/n)\otimes nx=0.$$

这里移过张量符号的标量是整数 n。因此对挠交换群 M 有 $\mathbb{Q}\otimes_{\mathbb{Z}}M=0$。对有限生成交换群，前一章的结构分解及直和公式给出

$$\mathbb{Q}\otimes_{\mathbb{Z}}(\mathbb{Z}^r\oplus M_{\mathrm{tor}})\cong\mathbb{Q}^r.$$

定义 11.19 (平坦模)**.** 模 F 称为平坦模，若对每个单射线性映射 $i:A\to B$，映射 $\mathrm{id}_F\otimes i:F\otimes_R A\to F\otimes_R B$ 仍为单射。

命题 11.20. *自由模是平坦模。对 $n\geq 2$，$\mathbb{Z}/n\mathbb{Z}$ 作为 $\mathbb{Z}$-模不平坦。*

证明. 若 $F\cong\bigoplus_{j\in J}R$，由直和与单位同构有 $F\otimes_R A\cong\bigoplus_j A$；这些同构将 $\mathrm{id}_F\otimes i$ 变成逐坐标映射 $(a_j)\mapsto(i(a_j))$。若所有像坐标为零，i 的单射性迫使所有 $a_j=0$，所以该映射单射。

另一方面，乘 n 的映射 $i:\mathbb{Z}\to\mathbb{Z}$ 为单射。用 $\mathbb{Z}/n\mathbb{Z}$ 张量后，两边均同构于 $\mathbb{Z}/n\mathbb{Z}$，而诱导映射为乘 n，即零映射。这个模非零，因此该映射不单射。 □

11.7 张量代数：把任意长度的乘积放在一起

定义 11.21 (分次代数)**.** R-代数 A 若具有模的直和分解 $A=\bigoplus_{k\geq 0}A^k$，满足 $1_A\in A^0$ 及 $A^pA^q\subseteq A^{p+q}$，称为非负分次代数。非零元素 $a\in A^p$ 称为次数为 p 的齐次元素，记 $|a|=p$。零元属于每个分量，不给它指定唯一的次数。

给定模 V，令 $V^{\otimes 0}=R$、$V^{\otimes 1}=V$，并用前面的多因子张量积定义 $V^{\otimes k}$。定义

$$T(V)=\bigoplus_{k\geq 0}V^{\otimes k}.$$

在两个纯张量上，用连接输入的方式定义乘法：

$$(v_1\otimes\cdots\otimes v_p)(w_1\otimes\cdots\otimes w_q)=v_1\otimes\cdots\otimes v_p\otimes w_1\otimes\cdots\otimes w_q.$$

多重线性泛性质保证这给出 $V^{\otimes p}\times V^{\otimes q}\to V^{\otimes(p+q)}$ 的双线性映射。对于零次分量，乘法取标量作用。对于一般元素 $x=\sum_p x_p$、$y=\sum_q y_q$，定义 $xy=\sum_{p,q}x_py_q$；由于两个和都有限，结果仍在直和中。

命题 11.22 (张量代数的自由性质)**.** *上述乘法使 $T(V)$ 成为分次 R-代数，单位为零次分量的 1_R，称为张量代数。对任意 R-代数 A 和 R-线性映射 $f : V \to A$，存在唯一 R-代数同态 $\widetilde{f} : T(V) \to A$ 延拓 f。*

证明. 连接三个纯张量时，无论先连接哪两个，得到同一个词。因而两种三线性运算在生成元上相等；逐项分配后，对所有元素都相等，这证明乘法结合。零次分量的 1_R 连接任意词都不改变它，故为单位。双线性性和次数相加来自构造。

对 $k \geq 1$, 映射 $(v_1, \ldots, v_k) \mapsto f(v_1) \cdots f(v_k)$ 为 k-线性映射, 故唯一延拓为 $f_k : V^{\otimes k} \to A$。在 $k = 0$ 时取 $f_0 = \iota_A : R \to A$。定义 $\widetilde{f}(\sum_k x_k) = \sum_k f_k(x_k)$。它线性且保持单位；在两个纯张量的乘积上，它保持乘法，因为 A 中乘法结合。双线性展开把这个结论推广到一般元素，故为代数同态。每个词都是 V 中元素的乘积，所以延拓的取值被迫由该公式给出，证明唯一性。 □

若 V 自由，张量代数中的一组模基是所有由基元组成的有限词，包含代表 1 的空词。这由各张量幂的基及直和构造得到。乘法是连接词；它通常不交换。例如在以 e_1, e_2 为基的模上，$e_1 \otimes e_2$ 与 $e_2 \otimes e_1$ 是两个不同的基元。

11.8 交替映射与外幂的构造

定义 11.23 (交替与反对称)**.** k-线性映射 $b : V^k \to W$ 称为交替映射，若任意两个输入相等时其值为零。$k = 1$ 时此条件为空，所有线性映射都称为交替。若交换任意两个相邻输入会使值乘以 -1，称为反对称映射。

引理 11.24 (交替性的相邻判别)**.** *k-线性映射只要在两个相邻输入相等时为零，就为交替映射。交替映射必反对称；若 $2 \in R$ 可逆，反对称映射也必交替。*

证明. 固定除相邻两处外的所有变量，把对应双线性映射记为 c。若 $c(x, x) = 0$，则

$$0 = c(x + y, x + y) = c(x, y) + c(y, x),$$

其中 $c(x, x)$、$c(y, y)$ 已为零。于是相邻交换使值变号。如果两个不相邻的输入相等，经过有限次相邻交换可将它们移到相邻位置；所得值为零，原值与它只差一个符号，也为零。这证明前两句。

反过来，反对称性给出 $c(x, x) = -c(x, x)$，故 $2c(x, x) = 0$。当 2 可逆时可消去 2，得到相邻相等时为零，再用第一句即可。 □

例如在 $\mathbb{F}_2$ 上，乘法 $(x, y) \mapsto xy$ 为双线性且反对称，因为 $-1 = 1$；但它在 $(1, 1)$ 上取值为 1，所以不交替。

定义 11.25 (外幂)**.** 对 $k \geq 2$，令 $K_k \leq V^{\otimes k}$ 为所有含相邻重复输入的纯张量生成的子模，即生成元具有形式

$$v_1 \otimes \cdots \otimes v_{i-1} \otimes v \otimes v \otimes v_{i+2} \otimes \cdots \otimes v_k.$$

定义 $\bigwedge^k V = V^{\otimes k}/K_k$，并令 $K_0 = K_1 = 0$、$\bigwedge^0 V = R$、$\bigwedge^1 V = V$。纯张量的商类记为 $v_1 \wedge \cdots \wedge v_k$。

定理 11.26 (外幂的泛性质)**.** 对 $k \geq 1$，映射 $V^k \to \bigwedge^k V$，$(v_1, \dots, v_k) \mapsto v_1 \wedge \cdots \wedge v_k$，为交替 k-线性映射。每个交替 k-线性映射 $b : V^k \to W$ 唯一地写成此映射与线性映射 $\widetilde{b} : \bigwedge^k V \to W$ 的复合。

证明. 标准映射为多重线性，并在相邻输入相等时为零，故由相邻判别知它交替。给定 b，先用张量幂的泛性质得到线性映射 $V^{\otimes k} \to W$。它消去 K_k 的每个生成元，故下降为 $\widetilde{b}$。外积词生成商模，因而延拓唯一。 □

外幂于是记录交替多重线性运算；它把重复输入变成零，把交换输入变成换号。Conrad 的 [16] 第 2–4 节依次证明交替性与反对称性的关系、构造外幂并求其基。这里以相邻重复关系构造商；相邻判别说明它与对所有重复输入取商的构造一致。

11.9 外积与外代数

令 $K = \bigoplus_{k \geq 0} K_k \subseteq T(V)$。在 K_k 的生成元前后连接任何纯张量，原有的相邻重复仍然存在，故乘积在相应的 K_{k+p} 中。线性展开说明 K 是双边理想。它恰为所有 $v \otimes v$ 生成的双边理想：一方面它含这些元素，另一方面 K_k 的每个生成元都由某个 $v \otimes v$ 在前后连接词得到。因此

$$\bigwedge V \stackrel{\text{def}}{=} T(V)/K \cong \bigoplus_{k \geq 0} \bigwedge^k V.$$

最后的同构把每个次数的商类送到总商中的同一类。它满射，因为每个张量只有有限个次数分量；若其像为零，则每个分量在 K_k 中，故各商类均为零，因而也单射。这给出了商代数的分次。

这个代数称为外代数，乘法记为 $\wedge$，称为外积。商代数乘法的良定义性可直接核对：改变 x, y 为 $x + a, y + b$，其中 $a, b \in K$，则乘积之差为 $xb + ay + ab \in K$。在纯外积上，乘法就是连接输入，因而结合且有单位 $1 \in \bigwedge^0 V$。

命题 11.27 (外积的符号规则)**.** 对 $u, v \in V$，有 $v \wedge v = 0$、$u \wedge v = -v \wedge u$。对 $\alpha \in \bigwedge^p V$、$\beta \in \bigwedge^q V$，有

$$\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha.$$

证明. 第一式来自定义。将 $(u + v) \wedge (u + v) = 0$ 展开，消去两个平方项得到第二式。若 α, β 都为纯外积词，将前 p 个因子逐一移过后 q 个因子，共需 pq 次相邻交换，故得公式。一般齐次元素是同一次数的纯外积词的有限线性组合，双线性展开给出同一公式。若其中一次数为零，公式就是标量与元素交换。 □

定理 11.28 (外代数的泛性质)**.** 若 A 是 R-代数，$f : V \to A$ 为 R-线性映射，且对每个 $v \in V$ 都有 $f(v)^2 = 0$，则 f 唯一延拓为 R-代数同态 $\bigwedge V \to A$。

证明. 由张量代数的自由性质，f 唯一延拓为 $T(V) \to A$。该同态消去每个 $v \otimes v$，故其核包含它们生成的双边理想 K；因而下降为所需代数同态。外代数由 V 及标量生成，故延拓唯一。 □

11.10 外幂的基、诱导映射与行列式

这里仍允许系数环 R 有零因子。行列式的排列求和公式只用加法、乘法和符号，所以下面的多重线性与交替性论证在任意交换环上都成立，不需要除法。域上“行列式非零则矩阵可逆”的判别却不能直接搬过来：例如整数矩阵 (2) 的行列式非零，但乘 2 的映射 $\mathbb{Z}\to\mathbb{Z}$ 取不到 1，所以没有整数矩阵作为它的逆。

定理 11.29 (有限自由模的外幂基)**.** *若 V 的基为 $e_1,\dots,e_n$，则对 $0\le k\le n$，$\bigwedge^k V$ 的基为*

$$e_{i_1}\wedge\cdots\wedge e_{i_k},\qquad 1\le i_1<\cdots<i_k\le n.$$

$k=0$ 时这指由 1 组成的基。因而其秩为 $\binom{n}{k}$；$k>n$ 时 $\bigwedge^k V=0$。

证明. $k=0$ 由定义成立。对 $k\ge 1$，先将任意输入按基展开，再通过相邻交换排列各指标。含重复指标的外积为零，其余项成为严格递增的词乘以符号。这些词因而张成；$k>n$ 时所有词都重复，故整个模为零。

还须证明张成词之间没有新的关系。对严格递增指标组 $I=(i_1,\dots,i_k)$，令 ϵ_i 为第 i 个坐标函数，定义

$$b_I(v_1,\dots,v_k)=\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\prod_{j=1}^{k}\epsilon_{i_{\sigma(j)}}(v_j).$$

这里 S_k 是 k 个字母的置换群，$\operatorname{sgn}(\sigma)$ 为其符号。这是坐标矩阵的行列式公式，且对每个变量线性。若第 a,b 个输入相等，将每个排列 σ 与 $\sigma\circ(a\ b)$ 配对；两项乘积相等、符号相反，所以总和为零。这在任意交换环中成立，不需要除以 2。

由外幂泛性质，b_I 给出线性泛函 $\bigwedge^k V\to R$。在递增基词 $e_{j_1}\wedge\cdots\wedge e_{j_k}$ 上，若 $I\neq J$，则所取坐标矩阵有零列，函数值为零；若 $I=J$，矩阵为单位矩阵，函数值为 1。对任何基词的线性关系逐个施加 b_I，所有系数都必须为零，故基词线性无关。 □

命题 11.30 (线性映射诱导外幂映射)**.** *每个线性映射 $f:V\to W$ 唯一诱导线性映射*

$$\bigwedge^k f:\bigwedge^k V\to\bigwedge^k W,\qquad v_1\wedge\cdots\wedge v_k\longmapsto f(v_1)\wedge\cdots\wedge f(v_k).$$

对 $k=0$ 取 $\bigwedge^0 f=\mathrm{id}_R$。这些映射保持复合与恒等映射。

证明. 所给输入组到右侧的映射为交替 k-线性映射，所以由外幂的泛性质延拓。若 $g:W\to U$，则 $\bigwedge^k(gf)$ 与 $(\bigwedge^k g)(\bigwedge^k f)$ 在所有纯外积上相同，故相等；恒等映射的情形也逐个固定纯外积。$k=0$ 时结论来自恒等映射的复合。 □

定理 11.31 (最高外幂与行列式)**.** *设 V 为秩 $n\ge 1$ 的自由模，$f:V\to V$ 在基 $e_1,\dots,e_n$ 中的矩阵为 $A=(a_{ij})$。则*

$$\Big(\bigwedge^n f\Big)(e_1\wedge\cdots\wedge e_n)=(\det A)(e_1\wedge\cdots\wedge e_n).$$

这个标量与所选基无关。

证明. 展开 $f(e_1)\wedge\cdots\wedge f(e_n)$，含重复指标的项消失，剩余项由排列索引。把每个外积按标准顺序排列得到

$$\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{\sigma(1),1}\cdots a_{\sigma(n),n}\,(e_1\wedge\cdots\wedge e_n),$$

其系数正是行列式。$\bigwedge^n V$ 为秩一自由模，因此这说明 $\bigwedge^n f$ 在整个模上就是乘 $\det A$。换另一组基后，得到另一个基元 $\omega' = u\omega$，其中 $u \in R$ 为单位：两个基元互相表示的系数乘积为 1。于是 $\bigwedge^n f(\omega') = u \det A\, \omega = \det A\, \omega'$，故作用标量不变。 □

由保持复合的性质立即得到 $\det(gf) = \det g \det f$：两边都是最高外幂上同一个复合映射的作用标量。这给出了行列式乘法公式的结构性证明。

例 11.32 (平面中的有向面积)**.** 取 $V = \mathbb{R}^2$，标准基为 e_1, e_2。若 $v = ae_1 + be_2$、$w = ce_1 + de_2$，则

$$v \wedge w = (ad - bc) e_1 \wedge e_2.$$

这个系数是两向量所张平行四边形的有向面积：当 v 位于水平轴时，它等于带符号的底乘高；一般情况下可用一个保向正交变换将 v 转到水平轴，正交变换的行列式为 1，故系数保持。若 $v = 0$，两种描述都为零。交换 v, w 改变取向，面积符号随之改变。

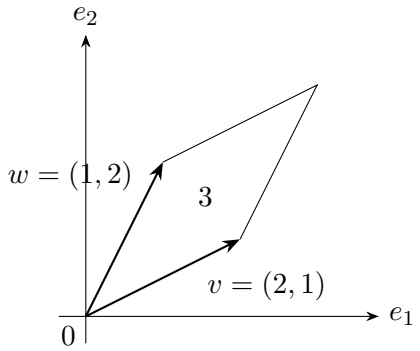


图 11.1: 以 v, w 为边的平行四边形；$v \wedge w = 3e_1 \wedge e_2$。

11.11 对称代数与多项式

交替关系消去重复输入，而对称关系只要求输入交换后相等。令 J 为 $T(V)$ 中所有 $u \otimes v - v \otimes u$ 生成的双边理想，定义对称代数

$$\operatorname{Sym}(V) = T(V)/J.$$

它由 V 的像及标量生成，且这些生成元彼此交换，故整个代数交换。J 由齐次的二次元素生成，因此每个元素均为有限个齐次元素之和，并且各齐次分量仍在 J 中：把左右相乘的因子按次数展开即可。于是令 $J_k = J \cap V^{\otimes k}$，有 $J = \bigoplus_k J_k$，从而

$$\operatorname{Sym}(V) = \bigoplus_{k \ge 0} \operatorname{Sym}^k(V), \qquad \operatorname{Sym}^k(V) = V^{\otimes k}/J_k.$$

同构的单射与满射可按外代数的逐次数论证核对。乘法仍为连接词再取商，记作通常的乘法符号。

命题 11.33 (对称代数的泛性质)**.** *每个从 V 到交换 R-代数 A 的 R-线性映射唯一延拓为 R-代数同态 $\operatorname{Sym}(V) \to A$。*

证明. 先由张量代数的自由性质延拓到 $T(V)$；由于 A 交换，每个 $u \otimes v - v \otimes u$ 都映为零，故映射下降。唯一性由生成元的取值确定。 □

定理 11.34 (有限自由模的对称代数)**.** *若 V 的基为 $e_1,\dots,e_n$，则 $e_i \mapsto x_i$ 给出代数同构*

$$\mathrm{Sym}(V) \cong R[x_1,\dots,x_n].$$

对 $n \ge 1$，$\mathrm{Sym}^k(V)$ 的基为 $e_1^{a_1}\cdots e_n^{a_n}$，其中 $a_i \ge 0$ 且 $\sum_i a_i = k$，故秩为 $\binom{n+k-1}{k}$。

证明. 多元多项式环可逐次构造为 $(R[x_1,\dots,x_{n-1}])[x_n]$，从 $R[x_1]$ 开始。对变量数归纳，由一元多项式系数的唯一性得到每个元素唯一写为有限和 $\sum_a r_a x_1^{a_1}\cdots x_n^{a_n}$。因此这些单项式为模基。对交换 R-代数 A，逐次代入可知：任意指定 x_i 的像，均唯一确定一个代数同态到 A。

将 e_i 送到 x_i，先由基确定线性映射，再用对称代数泛性质得到 $\mathrm{Sym}(V) \to R[x_1,\dots,x_n]$。反向用多项式的代入性质，将 x_i 送到 e_i 的商类。两复合固定所有生成元和标量，故为恒等。此同构把长度为 k 的词送到总次数为 k 的单项式，所以逐次数也为同构，基的结论由单项式表示的唯一性得到。

为了计数，把指数向量 $(a_1,\dots,a_n)$ 对应于一行 k 个圆点与 $n-1$ 条分隔线：第 i 组含 a_i 个圆点，允许相邻分隔线和空组。反过来，每种排列唯一恢复各 a_i。在 $k+n-1$ 个位置中选取 $n-1$ 个分隔线的位置，故数目为 $\binom{k+n-1}{n-1} = \binom{k+n-1}{k}$。

若 $n=0$，则 $V=0$，正次数的张量幂全为零，所以 $\mathrm{Sym}(0) = R$；零次分量为 R，正次数分量为零。 □

例 11.35 (同一个自由模上的三种乘积)**.** 当 $V = Re_1 \oplus Re_2$ 时，$T^2(V)$ 的基为四个有序词 $e_i \otimes e_j$；这里 $T^2(V)$ 表示张量代数的二次分量 $V^{\otimes 2}$。$\mathrm{Sym}^2(V)$ 的基为 e_1^2, e_1e_2, e_2^2，因为交换关系识别两种次序。$\bigwedge^2 V$ 的基只有 $e_1 \wedge e_2$，因为重复输入为零，而另一种次序等于它的负值。这些基分别由前面的张量基、对称幂基和外幂基定理保证独立，因而所列关系没有再消去其他元素。

11.12 分次交换与张量乘法的符号

定义 11.36 (分次交换)**.** 分次代数 A 称为分次交换代数，若对 $a \in A^p, b \in A^q$ 均有 $ab = (-1)^{pq}ba$。因此外代数是分次交换代数。对称代数具有通常的交换乘法，但以生成元为一次元的分次一般不满足这个符号规则。

例如在 $\mathbb{Z}[x]$ 中，x 的次数为 1，分次交换会要求 $x^2 = -x^2$；但 $2x^2 \neq 0$。因此“交换的分次代数”和“分次交换代数”是两个不同条件。

设 A, B 为分次代数。借助直和公式，把 $A \otimes_R B$ 的第 n 次分量定义为 $\bigoplus_{p+q=n} A^p \otimes_R B^q$。对齐次元素定义

$$(a \otimes b)(a' \otimes b') = (-1)^{|b||a'|}(aa' \otimes bb').$$

这个符号来自把 b 移过 a'。在每组固定次数上，右侧对四个输入分别线性；先延拓到相应张量积，再对次数求有限和，故得到良定义的双线性乘法。

命题 11.37 (分次张量积代数)**.** *上述乘法使 $A \otimes_R B$ 成为分次代数，单位为 $1_A \otimes 1_B$。若 A, B 都分次交换，则此代数也分次交换。*

证明. 设 $|a| = p, |b| = q, |a'| = r, |b'| = s, |a''| = t$。三个纯张量按左结合相乘时，符号指数为 $qr + (q+s)t$；按右结合时为 $st + q(r+t)$。两者相等，结合律由 A, B 的结合律得到。单位次数为零，代入公式得到单位律；次数相加也直接成立。一般情形由展开成齐次纯张量的有限和得到。

若两个因子分次交换，将 aa'、bb' 分别调换得到总符号 $(-1)^{qr+pr+qs}$。反向乘积本身带符号 $(-1)^{sp}$，而两张量次数分别为 $p+q, r+s$。因为

$$(p+q)(r+s)+sp \equiv pr+qr+qs \pmod 2,$$

故原乘积等于反向乘积的 $(-1)^{(p+q)(r+s)}$ 倍，证明分次交换律。 □

文献：张量积的构造、基本同构与扩张标量参见 [14] 第 3–6 节；交替映射、外幂、诱导映射、行列式和外代数参见 [16] 第 2–6、8、10 节。综合教材参见 [8, 9]。

12 域扩张与 Galois 理论

12.1 域中的标量与特征

在 $\mathbb{Q}$ 上，方程 $t^2-2=0$ 没有解；在 $\mathbb{R}$ 中加入它的一个根以后，便得到一个含有 $\mathbb{Q}$ 的新域。这里不仅多了一个数：它的和、积、负数、逆元也必须同时存在。域扩张研究这些运算共同产生的结构。

本章用 F, E, K, L 表示域，t 表示多项式的不定元。域同态均保持单位元。若 $\sigma: F \to E$ 是域同态，则它必单射。事实上，若 $a \neq 0$ 且 $\sigma(a)=0$，则

$$1=\sigma(1)=\sigma(aa^{-1})=\sigma(a)\sigma(a^{-1})=0,$$

矛盾。域之间的单同态也称为*嵌入*。

定义 12.1 (特征与素子域)**.** 对域 F，令 $\eta: \mathbb{Z} \to F$ 为 $n \mapsto n1_F$，负整数的像由加法逆元定义。若 $\ker\eta=0$，定义 $\operatorname{char} F=0$；若 $\ker\eta=p\mathbb{Z}$ 且 $p>0$，定义 $\operatorname{char} F=p$。F 中包含 1_F 的最小子域称为素子域。

η 是环同态：加法由重复相加的结合律得到，乘法由分配律得到；负号关系 $\eta(-n)=-\eta(n)$ 将等式延伸到全部整数。若核非零，其正生成元 p 必为素数。否则 $p=ab$，$1<a,b<p$，则

$$0=\eta(p)=\eta(a)\eta(b), \qquad \eta(a)\neq 0, \quad \eta(b)\neq 0,$$

与域无零因子矛盾。此时 $\eta(\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z}=\mathbb{F}_p$ 已是域；每个含 1_F 的子域都含这些整数倍，所以素子域正是 $\mathbb{F}_p$ 的此副本。

特征为零时，定义

$$\mathbb{Q} \to F, \qquad \frac{a}{b} \mapsto \eta(a)\eta(b)^{-1} \quad (b\neq 0).$$

若 $a/b=c/d$，则 $ad=bc$，所以 $\eta(a)\eta(d)=\eta(b)\eta(c)$；乘以 $\eta(b)^{-1}\eta(d)^{-1}$ 后，两种像相同。由通分公式验证加法与乘法，此映射是保持单位的域同态，因而单射。每个子域必须含 $\eta(a)$ 以及非零 $\eta(b)$ 的逆，所以素子域同构于 $\mathbb{Q}$。

命题 12.2 (Frobenius 映射)**.** *若 $\operatorname{char} F=p>0$，映射 $\varphi: F \to F$，$a \mapsto a^p$ 是单射域同态，称为 Frobenius 映射。*

证明. 对 $1 \le j \le p-1$，二项系数满足

$$j\binom{p}{j}=p\binom{p-1}{j-1}.$$

因 $p \nmid j$，Euclid 引理给出 $p \mid \binom{p}{j}$。在 F 中这些系数为零，故

$$(a+b)^p=a^p+b^p, \qquad (ab)^p=a^pb^p, \qquad 1^p=1.$$

因此它是域同态，单射性由上面的域同态判别得到。若 F 有限，集合上的单射也是满射，所以此时 φ 是自同构。迭代还给出 $(a+b)^{p^r}=a^{p^r}+b^{p^r}$。 □

12.2 扩张次数与生成子域

定义 12.3 (域扩张与次数)**.** 若 $F \subseteq E$ 是共享单位元的子域包含关系，称 E/F 为域扩张。E 的加法与 F 的数乘使它成为 F-向量空间，记

$$[E : F] = \dim_F E.$$

维数有限时称为有限扩张。若域同态 $E \to E'$ 在 F 上等于恒等映射，称为 F-嵌入；双射的 F-嵌入称为 F-同构。

例如 $\mathbb{C}/\mathbb{R}$ 有基 $(1, i)$，所以 $[\mathbb{C} : \mathbb{R}] = 2$。$\mathbb{Q}(i)/\mathbb{Q}$ 也有基 $(1, i)$。这两个二维向量空间使用的标量域不同。

定理 12.4 (塔式次数公式)**.** 设 $F \subseteq K \subseteq E$。$E/F$ 有限当且仅当 K/F 和 E/K 都有限；此时

$$[E : F] = [E : K][K : F].$$

证明. 先设 $u_1, \dots, u_m$ 为 K/F 的基，$v_1, \dots, v_n$ 为 E/K 的基。对任意 $x \in E$，依次展开

$$x = \sum_j b_j v_j, \qquad b_j = \sum_i a_{ij} u_i, \qquad x = \sum_{i,j} a_{ij} u_i v_j \quad (a_{ij} \in F).$$

故全部 $u_i v_j$ 张成 E。若 $\sum_{i,j} c_{ij} u_i v_j = 0$，则

$$\sum_j \left(\sum_i c_{ij} u_i \right) v_j = 0 \quad \Longrightarrow \quad \sum_i c_{ij} u_i = 0 \text{ 对每个} j \quad \Longrightarrow \quad c_{ij} = 0.$$

第一个推论使用 v_j 的 K-线性无关性，第二个使用 u_i 的 F-线性无关性。故有 mn 个基向量，得到乘法公式。

反过来，若 E 由 N 个元素在 F 上张成，同一批元素也在 K 上张成，所以 E/K 有限。又 K 是 E 的 F-子空间，故 K/F 有限。已证公式因而适用。 □

定义 12.5 (生成子环与生成子域)**.** 设 $S \subseteq E$。所有含 $F \cup S$ 的子环的交记为 $F[S]$，所有含 $F \cup S$ 的子域的交记为 $F(S)$。当 $S = \{\alpha_1, \dots, \alpha_r\}$ 时，分别写成 $F[\alpha_1, \dots, \alpha_r]$ 和 $F(\alpha_1, \dots, \alpha_r)$。由一个元素生成的域扩张称为简单扩张。

这两族子集都可在 E 的幂集内取出，并且都以 E 本身为一个成员，所以参与相交的族非空。若 x, y 属于交集，则在每个参与对象中都有 $x - y, xy$；子域情形还在 $x \neq 0$ 时含 x^{-1}。各参与对象也都含 $0, 1$，因此交集确实为子环或子域，且按包含关系最小。

$F[S]$ 正好由有限多项式表达式构成：有限和的差仍为有限和，两个有限和相乘由分配律仍给出有限和，所以这些表达式形成含 F, S 的子环；任何含 F, S 的子环都必须含这些表达式。两种包含给出相等。进一步，

$$F(S) = \left\{ \frac{a}{b} : a, b \in F[S],\ b \neq 0 \right\} \subseteq E.$$

右侧按 $a/b - c/d = (ad - bc)/(bd)$、$(a/b)(c/d) = ac/(bd)$ 及 $(a/b)^{-1} = b/a$ 封闭，是一个子域；所有含 F, S 的子域都含这些商，所以再次由最小性得到等号。$F[\alpha]$ 允许多项式表达式，$F(\alpha)$ 同时允许这些表达式的非零分母。

命题 12.6 (有限维子环中的逆元)**.** 若 $F \subseteq E$，且中间子环 A 含 F、在 F 上有限维，则 A 是域。

证明. 取 $0 \neq a \in A$。映射 $L_a : A \to A$，$x \mapsto ax$，满足

$$L_a(x+y) = L_a(x) + L_a(y), \qquad L_a(cx) = cL_a(x) \quad (c \in F).$$

若 $L_a(x) = 0$，E 无零因子给出 $x = 0$，所以它单射。有限维下单射蕴含满射，存在 $b \in A$ 使 $L_a(b) = 1$，即 $ab = 1$。因此每个非零元素在 A 内有逆。 □

12.2.1 用形式分数构造分式域

前面的生成子域是在已有的大域内部取商。一个尚未置于任何域中的整环，也可以直接构造出分母。以下设 A 是交换整环，令

$$P = \{(a,b) : a,b \in A,\ b \neq 0\}, \qquad (a,b) \sim (c,d) \iff ad = bc.$$

这是等价关系：自反、对称性来自等式；若 $ad = bc$ 且 $cf = de$，则

$$adf = bcf = bde, \qquad d(af - be) = 0.$$

因 $d \neq 0$，消去得到 $af = be$，即传递性。将 (a,b) 的类写成 a/b，类的集合记为 $\mathrm{Frac}(A)$。

命题 12.7 (分式域及其运算)**.** 定义

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}, \qquad \frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

这使 $\mathrm{Frac}(A)$ 成为一个域，且 $a \mapsto a/1$ 给出 A 到其中的单射环同态。

证明. 先验证代表元无关性。设 $ab' = a'b$、$cd' = c'd$。分母乘积非零，由上述两个等式有

$$\begin{aligned}(ad+bc)b'd' &= ab'dd' + bb'cd' = a'bdd' + bb'c'd \\ &= (a'd' + b'c')bd, \\ (ac)b'd' &= (ab')(cd') = (a'b)(c'd) = (a'c')bd.\end{aligned}$$

所以加法和乘法的等价类都不依赖代表元。

各条域公理由通分计算。对于 $x = a/b$，$y = c/d$，$z = e/f$，

$$\begin{aligned}(x+y)+z &= \frac{adf+bcf+bde}{bdf} = x + (y+z), \\ (xy)z &= \frac{ace}{bdf} = x(yz), \\ x + y &= \frac{ad+bc}{bd} = y + x, \qquad xy = \frac{ac}{bd} = yx.\end{aligned}$$

这里交换律等式用 A 的交换律将两个分子及分母重排。$0 = 0/1$、$1 = 1/1$ 满足

$$x + 0 = \frac{a}{b} = x, \quad x \cdot 1 = x, \quad x + \frac{-a}{b} = \frac{0}{b^2} = 0.$$

分配律的两端分别为

$$x(y+z) = \frac{a(cf+de)}{bdf}, \qquad xy + xz = \frac{acbf + abde}{b^2 df}.$$

右端的分子分母共同约去非零 b 后恰为左端；按等价关系，这两个分数相同。故得到交换有单位环。

等价关系说明 $a/b = 0/1$ 当且仅当 $a = 0$。若 $a/b \neq 0$，则 $a \neq 0$，b/a 有定义且

$$\frac{a}{b}\frac{b}{a} = \frac{ab}{ba} = 1.$$

又 $0/1 \neq 1/1$，所以这是域。最后 $a \mapsto a/1$ 保持加法、乘法与单位；若 $a/1 = c/1$，则 $a = c$，故单射。 □

若 A 已嵌入一个域 E，映射 $a/b \mapsto ab^{-1} \in E$ 良定义：$ad = bc$ 在 E 中乘以 $b^{-1}d^{-1}$ 后给出同一个值。反之，值相同也蕴含交叉乘积相同，所以它单射，并将上述加法、乘法送到 E 的运算。因此抽象分式域与先前在 E 内取商的构造一致。

特别地，对多项式环 $F[u]$，定义有理函数域 $F(u) = \mathrm{Frac}(F[u])$；其中 u 是不定元。$F[u]$ 嵌入此域，故不同多项式仍不同，u 在 F 上超越。后文的不可分扩张 $\mathbb{F}_p(u)/\mathbb{F}_p(u^p)$ 使用这一构造。

12.3 代数元与最小多项式

定义 12.8 (代数与超越)**.** 若 $\alpha \in E$ 满足某个非零 $f \in F[t]$ 的等式 $f(\alpha) = 0$，称 α 在 F 上代数；否则称为超越元。若 E 每个元素都在 F 上代数，称 E/F 为代数扩张。

代入映射 $\mathrm{ev}_\alpha : F[t] \to E$ 是环同态。有限和的展开给出

$$(f+g)(\alpha) = f(\alpha) + g(\alpha), \qquad (fg)(\alpha) = \sum_{i,j} a_i b_j \alpha^{i+j} = f(\alpha)g(\alpha).$$

它在常数上是恒等映射，像为 $F[\alpha]$。若 α 超越，核为零；若 α 代数，核是非零理想。

定理 12.9 (最小多项式与幂基)**.** *设 α 在 F 上代数。存在唯一首一多项式 $m_{\alpha,F}$ 生成 $\ker \mathrm{ev}_\alpha$。它不可约，并且*

$$f(\alpha) = 0 \iff m_{\alpha,F} \mid f, \qquad F(\alpha) = F[\alpha] \cong F[t]/(m_{\alpha,F}).$$

若 $d = \deg m_{\alpha,F}$，则 $(1, \alpha, \dots, \alpha^{d-1})$ 是 $F(\alpha)/F$ 的基，故 $[F(\alpha) : F] = d$。

证明. $F[t]$ 是 PID，故非零核为 (m)。将任意生成元除以其非零首系数，可令其首一。若首一 m, m' 生成同一个理想，则 $m = um'$，其中 u 是多项式环的单位，即非零常数。比较首系数得 $u = 1$，故 $m = m'$。

假设 $m = gh$ 且 g, h 均为正次数，则

$$0 = m(\alpha) = g(\alpha)h(\alpha).$$

E 无零因子，所以至少一项为零；于是 $g \in (m)$ 或 $h \in (m)$。但非零的 m 的倍式次数至少为 $\deg m$，与 $\deg g, \deg h < \deg m$ 矛盾。因此 m 不可约。核等式立即给出零化判别。

第一同构定理给出 $F[t]/(m) \cong F[\alpha]$。还可直接求出商中的逆：若 $0 \neq [r] \in F[t]/(m)$，取 $\deg r < d$ 的代表，则不可约性给出 $(r, m) = (1)$。选 $u, v \in F[t]$ 使

$$ur + vm = 1, \qquad [u][r] = [1].$$

所以商为域，$F[\alpha]$ 已是含 F,α 的子域，等于 $F(\alpha)$。

带余除法 $f = qm + r$，$\deg r < d$，给出 $f(\alpha) = r(\alpha)$，所以所列幂张成。若 $\sum_{i=0}^{d-1} c_i\alpha^i = 0$，则 $r = \sum_i c_i t^i$ 被 m 整除，次数限制迫使 $r = 0$，所以全部 $c_i = 0$。这证明幂基及次数公式。 □

命题 12.10 (添入一个根的构造)**.** *若 $m \in F[t]$ 首一不可约，$d = \deg m \geq 1$，则 $F[t]/(m)$ 是 F 的一个扩张，次数为 d；其中 $\alpha = [t]$ 满足 $m(\alpha) = 0$。*

证明. 前一证明的 Bézout 计算直接说明此商为域。常数映射 $F \to F[t]/(m)$，$a \mapsto [a]$，单射：若常数 $a \in (m)$，正次数的 m 不能整除非零常数，故 $a = 0$。据此将 F 与其像识别。对所有 $f \in F[t]$，$f([t]) = [f(t)]$，所以 $m([t]) = 0$；带余除法的同一基计算给出次数 d。 □

例 12.11 (三次扩张中的乘法与求逆)**.** 取 $\alpha = \sqrt[3]{2} \in \mathbb{R}$。若 $t^3 - 2$ 在 $\mathbb{Q}[t]$ 中可约，三次因式分解必有一个一次因子，因而有有理根 a/b，其中 $a, b \in \mathbb{Z}$ 互素、$b > 0$。等式 $a^3 = 2b^3$ 迫使 $b = 1$：若素数 $p \mid b$，则 $p \mid a^3$，继而 $p \mid a$，矛盾。随后 $a^3 = 2$ 也不可能，因为左端素因子 2 的指数为 3 的倍数，右端指数为 1。故 $t^3 - 2$ 不可约，$\mathbb{Q}(\alpha)$ 的基为 $(1, \alpha, \alpha^2)$。

用关系 $\alpha^3 = 2$，乘法化为

$$\begin{aligned}(a + b\alpha + c\alpha^2)(u + v\alpha + w\alpha^2) = {} & (au + 2bw + 2cv) \\ & + (av + bu + 2cw)\alpha + (aw + bv + cu)\alpha^2.\end{aligned}$$

例如 $(1 + \alpha)(1 - \alpha + \alpha^2) = 1 + \alpha^3 = 3$，所以

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

在基 $(1, \alpha, \alpha^2)$ 下，乘以 $a + b\alpha + c\alpha^2$ 的矩阵为

$$\begin{pmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{pmatrix}.$$

求其逆元也等价于用这个矩阵解右端为 $(1, 0, 0)^t$ 的线性方程组。

命题 12.12 (有限生成与代数性)**.** *扩张 E/F 有限，当且仅当 $E = F(\alpha_1, \dots, \alpha_r)$，其中每个 α_i 在 F 上代数。代数扩张具有传递性：若 E/K 与 K/F 代数，则 E/F 代数。*

证明. 若 $[E : F] = n$，任取 $\alpha \in E$，$n + 1$ 个向量 $1, \alpha, \dots, \alpha^n$ 相关；非零关系就是一个非零零化多项式。故 E/F 代数。取有限基 $u_1, \dots, u_n$，由线性张成知 $F(u_1, \dots, u_n) = E$。

反之，令 $K_0 = F$，$K_i = F(\alpha_1, \dots, \alpha_i)$。$\alpha_i$ 在 F 上的非零零化多项式仍在 $K_{i-1}[t]$ 中，故其在 K_{i-1} 上也代数，$[K_i : K_{i-1}] < \infty$。塔式公式给出 $[E : F] = \prod_i [K_i : K_{i-1}] < \infty$。

最后任取 $x \in E$，选其在 K 上的非零零化多项式 $\sum_{i=0}^{d} a_i t^i$。有限多个系数 a_i 均在 F 上代数，故 $K_0 = F(a_0, \dots, a_d)$ 在 F 上有限。x 在 K_0 上代数，所以 $K_0(x)/K_0$ 有限，进而 $K_0(x)/F$ 有限。刚证的第一部分说明 x 在 F 上代数。 □

12.4 嵌入由根的像决定

要比较两个添根的构造，需要确定：将系数域嵌入另一个域以后，新加入的根可以送到哪里？

若 $\sigma: K \to \Omega$ 是域嵌入，对 $f(t)=\sum_i a_i t^i \in K[t]$ 定义

$$\sigma_* f(t) = \sum_i \sigma(a_i) t^i \in \sigma(K)[t].$$

系数逐项作用给出一个环同构 $K[t] \to \sigma(K)[t]$；其逆由 $\sigma^{-1}: \sigma(K) \to K$ 逐系数作用给出。因此它保持次数、整除和不可约性。

定理 12.13 (简单代数扩张的嵌入)**.** 设 α 在 K 上代数，最小多项式为 $m = m_{\alpha,K}$，并给定域嵌入 $\sigma: K \to \Omega$。σ 到 $K(\alpha)$ 的延拓与 $\sigma_* m$ 在 Ω 中的根一一对应。给定根 β 后，唯一延拓为

$$\widetilde{\sigma}\left(\sum_{i=0}^{d-1} a_i \alpha^i\right) = \sum_{i=0}^{d-1} \sigma(a_i)\beta^i, \qquad d = \deg m.$$

证明. 若已有延拓，作用于 $m(\alpha)=0$ 得

$$0 = \widetilde{\sigma}(m(\alpha)) = (\sigma_* m)(\widetilde{\sigma}(\alpha)),$$

所以像必为根。因为 $K(\alpha) = K[\alpha]$，同态在系数和 α 上的像决定全部元素的像，故至多有一个这样的延拓。

反向取 $(\sigma_* m)(\beta)=0$，定义 $u: K[t] \to \Omega$ 为

$$u(f) = (\sigma_* f)(\beta).$$

逐系数运算和代入给出 $u(f+g)=u(f)+u(g)$、$u(fg)=u(f)u(g)$、$u(1)=1$。若 $f - g = mh$，则

$$u(f) - u(g) = (\sigma_* m)(\beta)(\sigma_* h)(\beta) = 0.$$

所以 u 下降为 $K[t]/(m) \to \Omega$ 的保持单位元的环同态；它从域出发，故单射。通过 $K[t]/(m) \cong K(\alpha)$ 即得延拓，并得到所写公式。不同的根给出不同的 α 的像，从而得到一一对应。 □

推论 12.14 (嵌入的数目上界)**.** 若 E/F 有限，$\sigma: F \to \Omega$ 是域嵌入，则将 σ 延拓到 E 的嵌入数目至多为 $[E:F]$。

证明. 写 $E = F(\alpha_1, \dots, \alpha_r)$，令 $K_i = F(\alpha_1, \dots, \alpha_i)$，$K_0 = F$，$d_i = [K_i : K_{i-1}]$。固定一个已定义在 K_{i-1} 上的嵌入，其延拓数目等于某个 d_i 次多项式在 Ω 中不同根的数目，故至多 d_i。每个最终嵌入唯一确定它在各层上的限制，因而不同的逐层选择不会重复计算最终嵌入。于是总数至多

$$\prod_{i=1}^{r} d_i = [E:F].$$

若某一步没有根，延拓数为零，仍符合上界。 □

12.5 分裂域的构造、延拓与唯一性

定义 12.15 (分裂域)**.** 对非常数多项式 $f \in F[t]$，若扩张 E/F 使

$$f(t) = c \prod_{i=1}^{d} (t - \alpha_i) \quad (c \in F^\times,\ \alpha_i \in E),$$

称 f 在 E 中分裂。若 $E = F(\alpha_1, \dots, \alpha_d)$，称其为 f 的分裂域。根可按重数重复列出。

定理 12.16 (分裂域的存在)**.** 每个非常数 $f \in F[t]$ 都有有限分裂域 E，且 $[E:F] \le (\deg f)!$。

证明. 对 $d = \deg f$ 归纳。$d = 1$ 时 F 本身即为分裂域。设 $d > 1$。选 f 的一个首一不可约因子 g，用命题 12.10 构造 $F_1 = F(\alpha)$，其中 $g(\alpha) = 0$。由于 $g \mid f$，也有 $f(\alpha) = 0$，因子定理（命题 8.26）给出

$$f = (t - \alpha) f_1, \qquad f_1 \in F_1[t], \quad \deg f_1 = d - 1.$$

归纳假设给出 f_1 在 F_1 上的有限分裂域 E。它由 α 与 f_1 的全部根在 F 上生成，故也是 f 的分裂域。并且

$$[E:F] = [E:F_1][F_1:F] \le (d-1)! \deg g \le d!.$$

每一步次数下降，故构造在有限次以后结束。 □

引理 12.17 (根生成扩张的嵌入延拓)**.** 设 $E = F(\alpha_1, \dots, \alpha_r)$，每个 α_i 是 $f \in F[t]$ 的根。给定 $\sigma : F \to \Omega$，并假设 $\sigma_* f$ 在 Ω 中分裂，则 σ 至少有一个到 E 的延拓。若 $\sigma_* f$ 的根均无重复，延拓数恰为 $[E:F]$。

证明. 令 $K_i = F(\alpha_1, \dots, \alpha_i)$。设已有 $\sigma_{i-1} : K_{i-1} \to \Omega$，并记 $m_i = m_{\alpha_i, K_{i-1}}$。因为 $f(\alpha_i) = 0$，有 $m_i \mid f$；作用于系数得

$$(\sigma_{i-1})_* m_i \mid \sigma_* f \quad \text{于} \Omega[t].$$

在多项式环的唯一分解中，一个分裂多项式的因子也是一次因子的乘积，因此 $(\sigma_{i-1})_* m_i$ 在 Ω 中分裂。它次数至少为一，所以有根，根像定理给出下一步延拓。

若 $\sigma_* f$ 无重复根，则它的每个因子也无重复根，因此这一步恰有 $\deg m_i = [K_i : K_{i-1}]$ 种根像。不同延拓的限制或新根像至少有一个不同，故互不相同。沿有限塔相乘，最终数目为

$$\prod_i [K_i : K_{i-1}] = [E:F].$$

□

定理 12.18 (分裂域唯一性)**.** 同一个 $f \in F[t]$ 的任意两个分裂域在 F 上同构。更一般地，若 E 是 f 的分裂域，$\sigma : F \to F'$ 是域同构，E' 是 $\sigma_* f$ 的分裂域，则 σ 可延拓为同构 $E \to E'$。

证明. 上一个引理给出嵌入 $\widetilde{\sigma} : E \to E'$。它将 f 的不同根送到 $\sigma_* f$ 的不同根，且把等式 $f = c \prod_i (t - \alpha_i)$ 送成

$$\sigma_* f = \sigma(c) \prod_i (t - \widetilde{\sigma}(\alpha_i)).$$

所以右边已经列出 $\sigma_* f$ 的全部根。$\widetilde{\sigma}(E)$ 是含 F' 及这些根的子域，而 E' 由它们生成，故 $E' \subseteq \widetilde{\sigma}(E)$。反向包含由映射的陪域得到，因此嵌入满射，为同构。取 $F' = F$、$\sigma = \mathrm{id}_F$ 得第一项结论。 □

12.6 形式导数与可分性

定义 12.19 (形式导数与可分多项式)**.** 对 $f(t) = \sum_{i=0}^{d} a_i t^i \in F[t]$，定义 $f'(t) = \sum_{i=1}^{d} i a_i t^{i-1}$。非零多项式若在分裂域中的全部根均无重复，称为可分多项式；非零常数也按此定义视为可分。

形式导数由有限和定义，不需要极限。对单项式有

$$(t^{i+j})' = (i+j)t^{i+j-1} = (t^i)'t^j + t^i(t^j)',$$

逐项乘开有限和，得到 $(fg)' = f'g + fg'$；加法与数乘性质也逐系数成立。

引理 12.20 (重复根与最大公因式)**.** *f 在任意分裂域中无重复根，当且仅当 $\gcd(f, f') = 1$ 于 $F[t]$。因此可分性不依赖分裂域的选择。*

证明. 在分裂域 E 中写 $f = (t-\alpha)^m h$，$m \geq 1$，$h(\alpha) \neq 0$。乘法法则给出

$$f' = m(t-\alpha)^{m-1}h + (t-\alpha)^m h'.$$

若 $m = 1$，$f'(\alpha) = h(\alpha) \neq 0$；若 $m \geq 2$，则 $f'(\alpha) = 0$。故重复根等价于 f, f' 的公共根。

若在 $F[t]$ 中 $uf + vf' = 1$，该等式也在 $E[t]$ 中成立，故不可能有公共根。反向，设首一 $d = \gcd(f, f')$ 非常数。因为 $d \mid f$ 且 f 在 E 中分裂，d 也在 E 中分裂；取其一个根，即为 f, f' 的公共根。这证明两个方向。 □

对扩张 K/F，若 f 在 F 上可分，Bézout 等式 $uf + vf' = 1$ 在 $K[t]$ 中仍成立，所以它在扩大系数域后依然可分。类似地，若 $d = \gcd(f, g)$ 在 $F[t]$ 中首一，则 $d \mid f, g$ 且 $d = uf + vg$，这些关系在 $K[t]$ 中仍成立；任何新的公因式仍整除 d，因此首一最大公因式不变。

命题 12.21 (不可约多项式的可分判别)**.** *设 $f \in F[t]$ 不可约。在特征零时 f 可分。在特征 $p > 0$ 时，f 不可分当且仅当 $f' = 0$，也当且仅当 $f(t) = g(t^p)$ 对某个 $g \in F[t]$ 成立。*

证明. 不可约性说明 $\gcd(f, f')$ 非单位当且仅当 $f \mid f'$；由于 $\deg f' < \deg f$（若 $f' \neq 0$），这等价于 $f' = 0$。特征零时，若 $f = \sum_i a_i t^i$ 的次数为 $d \geq 1$，则 $da_d \neq 0$，故 $f' \neq 0$。

特征 p 时，$f' = 0$ 等价于 $ia_i = 0$ 对每个 $i \geq 1$ 成立。若 $p \nmid i$，$i1_F \neq 0$ 在域中可逆，故必须 $a_i = 0$；若 $p \mid i$，该项导数为零。因此非零项仅能出现在 p 的倍数次幂，恰好写成 $g(t^p)$。 □

定义 12.22 (完美域)**.** 若域 F 上的每个不可约多项式都可分，称 F 为完美域。

特征零的域完美。特征 p 中，若 Frobenius 满射，取 $f(t) = g(t^p) = \sum_i a_i t^{pi}$，对每个系数选 $b_i^p = a_i$，得到

$$f(t) = \left(\sum_i b_i t^i \right)^p.$$

若 g 非常数，右端为非单位的 p 次幂，故 f 可约。结合前一判别，这证明 Frobenius 满射的域完美；有限域的 Frobenius 满射，因而有限域完美。

例 12.23 (一项不可分扩张)**.** 令 u 为不定元，在有理函数域 $E = \mathbb{F}_p(u)$ 内取 $F = \mathbb{F}_p(u^p)$。元素 $1, u, \dots, u^{p-1}$ 在 F 上无关。事实上，若

$$\sum_{i=0}^{p-1} r_i(u^p) u^i = 0,$$

清去有限多个非零多项式分母后，得 $\sum_i a_i(u^p)u^i = 0$。第 i 项所有幂的次数均模 p 余 i，不同项的次数不重合，所以多项式各系数为零，全部 $a_i = 0$，继而 $r_i = 0$。

又 u 满足 $t^p-u^p\in F[t]$，故最小多项式次数至多为 p。已证无关性迫使次数至少为 p，于是

$$m_{u,F}(t)=t^p-u^p=(t-u)^p \quad 于E[t], \qquad [E:F]=p.$$

它只有一个不同的根。任意 F-自同构必须将 u 送到该根，故只能固定 u，即 $\mathrm{Aut}(E/F)=\{\mathrm{id}\}$。扩张次数是 p，自同构数却为 1。

12.7 有限自同构群的固定域

定义 12.24 (自同构群与固定域)**.** 对扩张 E/F，记 $\mathrm{Aut}(E/F)$ 为固定 F 的域自同构所成的群。域 E 的全体自同构所成的群记为 $\mathrm{Aut}(E)$。若 G 是它的子群，定义

$$E^G=\{x\in E:\sigma(x)=x \text{ 对所有}\sigma\in G\}.$$

称为 G 的固定域。

自同构的复合仍保持两种运算；双射同态的逆保持两种运算，所以这里确实为群。若 $x,y\in E^G$，则

$$\sigma(x-y)=x-y,\qquad \sigma(xy)=xy,\qquad \sigma(x^{-1})=\sigma(x)^{-1}=x^{-1}\quad (x\neq 0).$$

故固定域确为子域。G 越大，要求固定的等式越多，这正是后来包含关系反转的原因。

例 12.25 (变换以后仍留在原处的数)**.** 在 $E=\mathbb{Q}(\sqrt{2})$ 中，每个元素唯一写为 $a+b\sqrt{2}$，其中 $a,b\in\mathbb{Q}$。规则

$$\sigma(a+b\sqrt{2})=a-b\sqrt{2}$$

保持加法与乘法：把两个这样的表达式展开即可检查，而乘法展开中使用的关系 $(\sqrt{2})^2=2$ 在换号后仍成立。又 $\sigma^2=\mathrm{id}$，故 σ 是自同构。任何固定 $\mathbb{Q}$ 的自同构都把 $\sqrt{2}$ 送到 t^2-2 的根，只能是 $\sqrt{2}$ 或 $-\sqrt{2}$，所以 $\mathrm{Aut}(E/\mathbb{Q})=\{\mathrm{id},\sigma\}$。

若一个元素被 σ 固定，则 $a+b\sqrt{2}=a-b\sqrt{2}$，得到 $2b\sqrt{2}=0$，故 $b=0$。因此 $E^{\{\mathrm{id},\sigma\}}=\mathbb{Q}$，而 $E^{\{\mathrm{id}\}}=E$。这个例子同时展示了自同构群、固定域，以及“要求越多自同构固定，留下的域越小”。下面的 Artin 定理把例子中的次数 2 与自同构数 2 相等推广到任意有限自同构群。

引理 12.26 (Artin 的维数上界)**.** *设 $G\le\mathrm{Aut}(E)$ 是有限群，$F=E^G$，则 $[E:F]\le|G|$。*

证明. 记 $|G|=m$，并列出 $\sigma_1=\mathrm{id},\sigma_2,\dots,\sigma_m$。任取 $n>m$ 个元素 $x_1,\dots,x_n\in E$，在域 E 上考虑齐次方程组

$$\sum_{j=1}^{n}\sigma_i(x_j)c_j=0\qquad (1\le i\le m).$$

m 个方程、n 个未知量的系数矩阵秩至多为 $m<n$，所以核非零。在所有非零解中，取非零坐标数最少的一个，重排 x_j 使 $c_1\neq 0$，并将解乘以 c_1^{-1}，使 $c_1=1$。

对 $\tau\in G$，应用 τ 得

$$\sum_{j}(\tau\sigma_i)(x_j)\tau(c_j)=0.$$

左乘 τ 置换 G 的全部元素，所以这些正是原方程组中各行的重排。因此 $\tau(c)=(\tau(c_j))$ 也是原方程组的解。若某个 $c_j\notin F$，选 τ 使 $\tau(c_j)\neq c_j$。解 $\tau(c)-c$ 非零，但第一坐标为 $1-1=0$，而原来为零的坐标仍为零，故非零坐标数更少，矛盾。

因此所有 $c_j \in F$。取 $\sigma_1 = \mathrm{id}$ 的一行，得

$$\sum_{j} x_j c_j = 0, \qquad c_1 = 1,$$

这是一个非平凡 F-线性关系。故任意 $m+1$ 个元素都相关，$\dim_F E \le m$。 □

定理 12.27 (Artin 固定域定理)**.** *若有限群 $G \le \mathrm{Aut}(E)$，则*

$$[E : E^G] = |G|, \qquad \mathrm{Aut}(E/E^G) = G.$$

而且对每个 $x \in E$，x 在 E^G 上的最小多项式为

$$m_{x,E^G}(t) = \prod_{y \in Gx} (t - y), \qquad Gx = \{\sigma(x) : \sigma \in G\},$$

其中轨道元素只列一次。

证明. 令 $F = E^G$，$H = \mathrm{Aut}(E/F)$。前引理使 E/F 有限，而推论 12.14 给出 $|H| \le [E:F]$。又 $G \subseteq H$，故

$$|G| \le |H| \le [E:F] \le |G|.$$

全部不等式为等号，且有限群的包含 $G \subseteq H$ 与相同阶数给出 $G = H$。

固定 $x \in E$，记 $g(t) = \prod_{y \in Gx}(t-y)$。任意 $\tau \in G$ 置换轨道，所以

$$\tau_* g(t) = \prod_{y \in Gx} (t - \tau(y)) = g(t).$$

比较系数说明全部系数在 F 中。因 $g(x) = 0$，$m = m_{x,F}$ 整除 g。另一方面，m 系数被每个 σ 固定，故

$$m(\sigma x) = \sigma(m(x)) = 0.$$

轨道中各元素不同，所以对应一次因子两两互素，因子定理给出 $g \mid m$ 于 $E[t]$。于是首一的 g, m 次数相同且互相整除，故 $m = g$。特别地，此最小多项式在 E 中分裂且无重复根。 □

注 12.28 (从线性方程中找回固定的系数)**.** Milne 在 [15] 定理 3.4 中，用自同构群写出上述方程组，再在非零解中取支撑最小者。决定性步骤是把 c_1 归一化为 1：这样 $\tau(c) - c$ 必少一个非零坐标。系数属于固定域由这个最小性推出，因此原来在大域 E 上解出的方程，最终给出小域 E^G 上的线性关系。

12.8 正规性、可分性与 Galois 扩张

定义 12.29 (正规、可分与 Galois)**.** 代数扩张 E/F 称为正规扩张，若每个在 E 中有根的 $F[t]$ 不可约多项式都在 E 中分裂；称为可分扩张，若每个 $x \in E$ 的最小多项式 $m_{x,F}$ 都可分。有限、正规且可分的扩张称为有限 Galois 扩张，其自同构群记为 $\mathrm{Gal}(E/F)$。

不可约多项式若有根 $x \in E$，其首一化就是 $m_{x,F}$，因为最小多项式整除它；所以正规性也等价于每个 $m_{x,F}$ 都在 E 中分裂。

定理 12.30 (有限 Galois 扩张的三个刻画)**.** *对扩张 E/F，以下条件等价：*

1. E/F *有限、正规且可分；*

2. E *是某个可分非常数多项式* $f \in F[t]$ *的分裂域；*

3. 存在有限群 $G \le \mathrm{Aut}(E)$，*使* $F = E^G$。

此时 $|\mathrm{Gal}(E/F)| = [E:F]$，*且* $E^{\mathrm{Gal}(E/F)} = F$。

证明. 先由第一项推出第二项。有限扩张有有限个生成元 $\alpha_1, \dots, \alpha_r$。取它们在 F 上的不同首一最小多项式之积 f。每项可分且两两不同、不可约；不同因子互素，故无公共根，所以 f 可分。正规性使每项在 E 中分裂，而这些根含所有生成元，因此 E 是 f 的分裂域。若 $E = F$，直接取 $f = t$。

再假设第二项，记 $G = \mathrm{Aut}(E/F)$。根生成扩张引理给出 $[E:F]$ 个 F-嵌入 $E \to E$；由于有限维 F-线性单射是满射，这些嵌入都是自同构，所以

$$|G| = [E:F].$$

令 $K = E^G$，则 $F \subseteq K \subseteq E$。Artin 定理给出 $[E:K] = |G|$；塔式公式因而给出

$$[E:F] = [E:K][K:F] = |G|[K:F], \qquad [K:F] = 1.$$

故 $K = F$，得到第三项。

最后由第三项推出第一项。Artin 定理保证 E/F 有限，并把任意 $x \in E$ 的最小多项式写成不同轨道元素对应的一次因子之积。因此所有这些最小多项式在 E 中分裂且可分，扩张正规、可分。次数和固定域的附加结论也已经在证明中得到。 □

推论 12.31 (生成可分元得到可分扩张)**.** *若* $E = F(\alpha_1, \dots, \alpha_r)$，*每个* α_i *在* F *上代数且可分，则* E/F *可分，并包含于某个有限 Galois 扩张* L/F。*若* L/F *有限 Galois 且* $F \subseteq K \subseteq L$，*则* L/K *也有限 Galois。*

证明. 取 α_i 的不同最小多项式之积 f。由前一定理证明中的因子计算，f 可分。在已给定的 E 上添入 f 的余下根，得到 L；它是 f 在 F 上的分裂域，所以 L/F Galois。L 的每个元素在 F 上可分，故其子域 E 的每个元素也可分。

对于第二项，取 L/F 的一个可分分裂多项式 $f \in F[t]$。Bézout 判别表明它在 $K[t]$ 中仍可分；又 $L = F(\text{全部根})$ 给出

$$F(\text{全部根}) \subseteq K(\text{全部根}) \subseteq L,$$

所以 $L = K(\text{全部根})$。它是 f 在 K 上的分裂域，仍为 Galois。 □

12.9 Galois 对应与正规子群

定理 12.32 (有限 Galois 基本定理)**.** *设* E/F *为有限 Galois 扩张，*$G = \mathrm{Gal}(E/F)$。*对应*

$$H \longmapsto E^H, \qquad K \longmapsto \mathrm{Gal}(E/K)$$

在 G *的子群与中间域* $F \subseteq K \subseteq E$ *之间互逆，且反转包含关系。若* $H \le G$，*则*

$$[E:E^H] = |H|, \qquad [E^H:F] = [G:H].$$

若 $K = E^H$，*则* K/F *正规当且仅当* H *在* G *中正规；此时* K/F *Galois，限制映射给出*

$$\mathrm{Gal}(K/F) \cong G/H.$$

证明. 对 $H \le G$，Artin 定理直接给出

$$\mathrm{Gal}(E/E^H) = H.$$

对中间域 K，前一推论保证 E/K Galois，再由固定域刻画得

$$E^{\mathrm{Gal}(E/K)} = K.$$

因此两个对应互逆。若 $H_1 \subseteq H_2$，每个被 H_2 固定的元素也被 H_1 固定，故 $E^{H_2} \subseteq E^{H_1}$；反向包含关系对自同构群的定义同样成立。次数公式由 Artin 定理及塔式公式得到：

$$[E^H : F] = \frac{[E:F]}{[E:E^H]} = \frac{|G|}{|H|}.$$

接着对 $\sigma \in G$ 计算共轭对应。对 $x \in E^H$、$h \in H$，

$$(\sigma h \sigma^{-1})(\sigma x) = \sigma(hx) = \sigma x,$$

所以 $\sigma(E^H) \subseteq E^{\sigma H \sigma^{-1}}$。反向，若 y 被 $\sigma H \sigma^{-1}$ 固定，则对每个 $h \in H$，

$$h(\sigma^{-1} y) = \sigma^{-1}(\sigma h \sigma^{-1} y) = \sigma^{-1} y,$$

故 $\sigma^{-1} y \in E^H$，即 $y \in \sigma(E^H)$。因此

$$\sigma(E^H) = E^{\sigma H \sigma^{-1}}.$$

若 H 正规，上式给出 $\sigma(K) = K$。于是限制映射

$$\rho : G \to \mathrm{Aut}(K/F), \quad \sigma \mapsto \sigma|_K$$

有定义，且 $\rho(\sigma\tau) = \rho(\sigma)\rho(\tau)$。其核为固定 K 的全体自同构，即 $\ker \rho = H$。令 $J = \operatorname{im} \rho$，则

$$K^J = \{x \in K : \sigma x = x \text{ 对所有} \sigma \in G\} = K \cap E^G = F.$$

J 有限，固定域刻画保证 K/F Galois；Artin 定理给出 $\mathrm{Gal}(K/F) = J$，第一同构定理给出 $J \cong G/H$。

反过来，若 K/F 正规，取 $\sigma \in G$，任意 $x \in K$ 的最小多项式在 K 中分裂，而

$$m_{x,F}(\sigma x) = \sigma(m_{x,F}(x)) = 0.$$

所以 $\sigma x \in K$，即 $\sigma(K) \subseteq K$。对 σ^{-1} 作同样推导得反向包含，故 $\sigma(K) = K$。共轭等式及对应的单射性于是给出 $\sigma H \sigma^{-1} = H$。这证明必要性。 □

命题 12.33 (交域与合成域)**.** *设 K_1, K_2 为有限 Galois 扩张 E/F 的中间域，$H_i = \mathrm{Gal}(E/K_i)$。在 E 内由 $K_1 \cup K_2$ 生成的域记为 K_1K_2，则*

$$\mathrm{Gal}(E/K_1K_2) = H_1 \cap H_2, \qquad K_1 \cap K_2 = E^{\langle H_1, H_2 \rangle}.$$

证明. 自同构固定 K_1K_2 必固定两子域；反过来，若它固定两者，则固定每个由两者元素通过有限次四则运算得到的元素，因而固定它们生成的域。所以第一项等式成立。

若 x 被 $\langle H_1, H_2 \rangle$ 固定，则被 H_1, H_2 分别固定，故 $x \in K_1 \cap K_2$。反过来，交域中的 x 被两个子群固定，也被它们的任意有限乘积及逆固定，故被生成子群固定。两种包含给出第二项等式。 □

12.10 有限域的构造与全部子域

有限域的元素数先由线性代数确定。若有限域 E 的特征为 p，取 E 在素子域 $\mathbb{F}_p$ 上的基 $u_1,\dots,u_n$。每个元素有唯一坐标列 $(a_1,\dots,a_n)\in\mathbb{F}_p^n$，所以

$$|E| = p^n.$$

$n\ge 1$，因为域至少有 $0,1$ 两个不同元素。

定理 12.34 (有限域的存在、唯一性与自同构)**.** *对每个素数 p 和整数 $n\ge 1$，存在元素数为 p^n 的域；任意两个这样的域同构，记为 $\mathbb{F}_{p^n}$。它是 $t^{p^n}-t$ 在 $\mathbb{F}_p$ 上的分裂域，且*

$$\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) = \langle \varphi \rangle, \quad \varphi(x)=x^p, \quad |\langle\varphi\rangle| = n.$$

证明. 令 $q=p^n$，取 $f=t^q-t$ 在 $\mathbb{F}_p$ 上的分裂域 L。导数 $f'=-1$，所以 f 恰有 q 个不同根。令

$$S=\{x\in L : x^q=x\}.$$

它含 $0,1$。对 $a,b\in S$，Frobenius 的迭代公式给出

$$(a-b)^q = a^q-b^q = a-b, \qquad (ab)^q = a^q b^q = ab;$$

若 $a\neq 0$，则 $(a^{-1})^q=(a^q)^{-1}=a^{-1}$。所以 S 是子域，包含全部根及 $\mathbb{F}_p$。由分裂域的生成条件，$L=\mathbb{F}_p(S)\subseteq S$；反向 $S\subseteq L$，故 $L=S$，其元素数为 q。这证明存在性。

若 E 是任意 q 元域，则 $E^\times$ 是 $q-1$ 阶群。Lagrange 定理给出 $a^{q-1}=1$ 对每个非零 a 成立；零也满足 $a^q=a$。因此 t^q-t 在 E 中已有 q 个不同根，次数限制说明它分裂且这些根就是 E 的全部元素。所以 E 同样是该多项式的分裂域，唯一性由定理 12.18 得到。

此分裂多项式可分，故 $E/\mathbb{F}_p$ Galois，群阶为 $[E:\mathbb{F}_p]=n$。Frobenius 是自同构，且 $\varphi^n(x)=x^q=x$。若它的阶为 $d<n$，则全部 q 个元素都满足 $x^{p^d}=x$，而非零多项式 $t^{p^d}-t$ 次数 $p^d<q$，不可能有 q 个根。故阶恰为 n，它的循环子群已与全群同阶，因此就是全群。 □

推论 12.35 (有限域的子域格)**.** *在 $E=\mathbb{F}_{p^n}$ 中，对每个正因子 $d\mid n$，存在唯一一个 p^d 元子域，且它为*

$$E^{\langle\varphi^d\rangle} = \{x\in E : x^{p^d}=x\}.$$

E 的每个子域都以这种方式出现。若 $d,e\mid n$，这些子域满足

$$\mathbb{F}_{p^d}\subseteq\mathbb{F}_{p^e} \quad\Longleftrightarrow\quad d\mid e.$$

证明. n 阶循环群的每个子群唯一为 $\langle\varphi^d\rangle$，其中 $d\mid n$，其阶为 n/d。Galois 对应给出相应固定域在 $\mathbb{F}_p$ 上的次数为 $n/(n/d)=d$，故元素数为 p^d。固定生成元等价于固定全子群，直接得到根集公式。每个子域包含素子域 $\mathbb{F}_p$，所以都在此对应范围内。

若 $d\mid e$，$\langle\varphi^e\rangle\subseteq\langle\varphi^d\rangle$，固定域包含反转，得到所需子域包含。反过来，子域包含及塔式公式给出 $e=[\mathbb{F}_{p^e}:\mathbb{F}_{p^d}]d$，所以 $d\mid e$。 □

例 12.36 (在 $\mathbb{F}_8$ 中逐项运算)**.** $t^3+t+1\in\mathbb{F}_2[t]$ 在 $0,1$ 处的值均为 1。三次多项式若可约则有一次因子，所以它不可约。令 $\alpha=[t]$，得到

$$\mathbb{F}_8 = \mathbb{F}_2[t]/(t^3+t+1), \qquad \alpha^3=\alpha+1.$$

幂基给出全部元素为 $a+b\alpha+c\alpha^2$，其中 $a,b,c\in\{0,1\}$。依次乘以 α 并用关系消去三次项，得到

$$\alpha^3=\alpha+1,\qquad \alpha^4=\alpha^2+\alpha,$$
$$\alpha^5=\alpha^2+\alpha+1,\qquad \alpha^6=\alpha^2+1,$$
$$\alpha^7=\alpha^3+\alpha=(\alpha+1)+\alpha=1.$$

$\alpha,\alpha^2,\dots,\alpha^6,1$ 的坐标列互不相同，正好列出七个非零元素。故 $\alpha^{-1}=\alpha^6=\alpha^2+1$。Frobenius 在 α 上的作用为

$$\alpha\longmapsto\alpha^2\longmapsto\alpha^4\longmapsto\alpha^8=\alpha.$$

这三个元素不同；它们同为 t^3+t+1 的根，因为 $f(\alpha)^2=f(\alpha^2)$。次数为三，故它们正是全部根。于是

$$t^3+t+1=(t-\alpha)(t-\alpha^2)(t-\alpha^4)\quad\text{于}\mathbb{F}_8[t].$$

命题 12.37 (有限乘法子群是循环群)**.** *域 F 的乘法群 $F^\times$ 的每个有限子群 A 都是循环群。特别地，有限域的非零元素构成循环群。*

证明. 令 m 为 A 中所有元素阶的最小公倍数。Lagrange 定理说明每个阶整除 $|A|$，所以 $m\mid|A|$。对每个素数幂 ℓ^a 恰好整除 m，按最小公倍数的定义，存在 $x\in A$ 的阶为 $\ell^a b$，其中 $\ell\nmid b$。元素 $y=x^b$ 的阶恰为 ℓ^a：$y^k=1$ 等价于 $\ell^a b\mid bk$，也即 $\ell^a\mid k$。

对 m 中各个不同素数这样选择元素 y_ℓ。它们在域的乘法群中彼此交换。两个可交换且阶互素的元素 u,v 的乘积阶是两阶之积。事实上，若 $(uv)^k=1$，则 $u^k=v^{-k}$ 属于 $\langle u\rangle\cap\langle v\rangle$；交中元素的阶同时整除两个互素数，只能为 1。所以 $u^k=v^k=1$，两个阶分别整除 k，其乘积也整除 k；反向取两阶之积为幂自然得到 1。对全部 y_ℓ 归纳，得到阶为 m 的元素 $y=\prod_\ell y_\ell$。

所有 $x\in A$ 均满足 $x^m=1$，因此 A 是非零多项式 t^m-1 的根集的子集，$|A|\le m$。结合 $m\mid|A|$，得 $|A|=m$。于是 y 的循环子群与 A 同阶，故 $A=\langle y\rangle$。 □

12.11 双二次扩张的基、固定域与图

例 12.38 (把四个自同构逐项写出)**.** 令 $a=\sqrt{2}$、$b=\sqrt{3}$ 为正实数，$E=\mathbb{Q}(a,b)$。先验证次数。t^2-2 没有有理根，所以 $\mathbb{Q}(a)/\mathbb{Q}$ 次数为二。若 $b=x+ya$，其中 $x,y\in\mathbb{Q}$，两边平方得到

$$3=x^2+2y^2+2xy\sqrt{2}.$$

由 $(1,\sqrt{2})$ 无关，$xy=0$。若 $y=0$，则 $x^2=3$ 不可能；若 $x=0$，则 $y^2=3/2$ 也不可能。这里可统一用唯一分解：有理数平方的每个素数指数均为偶数，而 3 和 $3/2$ 的素数 3 的指数均为奇数。因此 $b\notin\mathbb{Q}(a)$，t^2-3 在 $\mathbb{Q}(a)$ 上不可约。于是

$$[E:\mathbb{Q}]=[E:\mathbb{Q}(a)][\mathbb{Q}(a):\mathbb{Q}]=2\cdot2=4,\qquad (1,a,b,ab)\text{ 是一组基.}$$

它是 $(t^2-2)(t^2-3)$ 的分裂域，且特征为零，故为 Galois。

对 $\varepsilon,\delta\in\{1,-1\}$，定义

$$\sigma_{\varepsilon,\delta}(x_0+x_1a+x_2b+x_3ab)=x_0+\varepsilon x_1a+\delta x_2b+\varepsilon\delta x_3ab.$$

基表示唯一，故映射有定义；它保持加法及有理数倍。再验证乘法：每个单项式可写为 a^ib^j，$i,j\in\{0,1\}$，乘法在 $a^2=2,b^2=3$ 下归约。若乘积中有 a^2 或 b^2，对应符号平方为 $\varepsilon^2=\delta^2=1$，所以归约前后的符号相同，得到

$$\sigma_{\varepsilon,\delta}(a^ib^ja^kb^\ell)=\sigma_{\varepsilon,\delta}(a^ib^j)\sigma_{\varepsilon,\delta}(a^kb^\ell).$$

对任意两个元素分配展开，即证明保持乘法。每个映射自身平方是恒等，故可逆，确为自同构。

记 $\sigma=\sigma_{-1,1}$，$\tau=\sigma_{1,-1}$，则

$$G=\{1,\sigma,\tau,\sigma\tau\},\qquad \sigma^2=\tau^2=1,\quad \sigma\tau=\tau\sigma,\quad G\cong C_2\times C_2.$$

任意 $\mathbb{Q}$-自同构必须分别把 a,b 送到其最小多项式的根，因此上述四个已经穷尽。

求固定域时，令 $x=x_0+x_1a+x_2b+x_3ab$，利用基的无关性逐项比较：

$$\begin{aligned}
\sigma x=x &\iff x_1=x_3=0, & E^{\langle\sigma\rangle}&=\mathbb{Q}(b),\\
\tau x=x &\iff x_2=x_3=0, & E^{\langle\tau\rangle}&=\mathbb{Q}(a),\\
\sigma\tau x=x &\iff x_1=x_2=0, & E^{\langle\sigma\tau\rangle}&=\mathbb{Q}(ab).
\end{aligned}$$

G 的三个二阶子群及平凡子群、全群就是全部子群；所以图 12.1 已列出全部中间域，没有遗漏其他二次或非二次中间域。

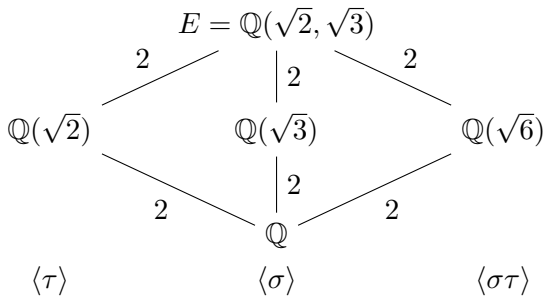


图 12.1: 双二次扩张的全部中间域。线段上数字是上方域对下方域的次数；图下方列出三列二次域各自对应的子群。

例 12.39 (两个平方根合并为一个生成元)**.** 仍取 $a=\sqrt{2},b=\sqrt{3}$，令 $x=a+b$。由 $a^2=2,b^2=3$，有

$$x^2=5+2ab,\qquad x^3=11a+9b.$$

将第二式分别减去 $9x$、$11x$，得

$$a=\frac{x^3-9x}{2},\qquad b=\frac{11x-x^3}{2}.$$

因此 $E=\mathbb{Q}(x)$。由 $(x^2-5)^2=4a^2b^2=24$，又得

$$x^4-10x^2+1=0.$$

由于 $[\mathbb{Q}(x):\mathbb{Q}]=4$，该首一四次零化多项式必为 $m_{x,\mathbb{Q}}$。其四个根是 $a+b,-a+b,a-b,-a-b$；它们是 x 的全部 G-轨道元素，四者不同。这也与 Artin 定理中的轨道乘积公式相符。

12.12 一个非交换的 Galois 群

例 12.40 (t^3-2 的分裂域与全部中间域)**.** 令 $a=\sqrt[3]{2}\in\mathbb{R}$，$\zeta=(-1+i\sqrt{3})/2\in\mathbb{C}$。直接计算得 $\zeta^2+\zeta+1=0$，所以 $\zeta^3=1$ 且 $\zeta\neq 1$。t^3-2 的三个根为

$$a,\quad \zeta a,\quad \zeta^2 a.$$

它的分裂域为 $E=\mathbb{Q}(a,\zeta)$。例 12.11 已证 $[\mathbb{Q}(a):\mathbb{Q}]=3$；又 $\mathbb{Q}(a)\subseteq\mathbb{R}$，而 $\zeta\notin\mathbb{R}$，所以 t^2+t+1 在 $\mathbb{Q}(a)$ 上无根、不可约。于是

$$[E:\mathbb{Q}]=2\cdot 3=6.$$

特征零保证可分，故 $E/\mathbb{Q}$ Galois，群阶为六。

$G=\mathrm{Gal}(E/\mathbb{Q})$ 置换三个根。若一个自同构固定三根，就固定它们生成的 E，所以此置换作用的核为零。因此 $G\hookrightarrow S_3$；两者阶均为六，故 $G\cong S_3$。

还可给出具体生成元。$[\mathbb{Q}(\zeta):\mathbb{Q}]=2$，塔式公式给出 $[E:\mathbb{Q}(\zeta)]=3$，所以 a 在 $\mathbb{Q}(\zeta)$ 上的最小多项式仍为 t^3-2。根像定理给出固定 ζ、将 a 送到 ζa 的自同构 σ。复共轭限制为 E 的自同构 τ，满足

$$\sigma(a)=\zeta a,\quad \sigma(\zeta)=\zeta,\qquad \tau(a)=a,\quad \tau(\zeta)=\zeta^2.$$

在两个生成元上计算得

$$\sigma^3=1,\qquad \tau^2=1,\qquad \tau\sigma\tau^{-1}=\sigma^{-1}.$$

例如最后一个等式作用于 a 得 $\tau\sigma(a)=\tau(\zeta a)=\zeta^2 a=\sigma^{-1}(a)$，作用于 ζ 的两边都等于 ζ。它们给出 S_3 中三循环与对换的关系。

$\langle\sigma\rangle$ 固定 $\mathbb{Q}(\zeta)$；其固定域次数为 $6/3=2$，与 $[\mathbb{Q}(\zeta):\mathbb{Q}]=2$ 相同，所以

$$E^{\langle\sigma\rangle}=\mathbb{Q}(\zeta).$$

$\langle\tau\rangle$ 固定 $\mathbb{Q}(a)$，固定域次数为三，所以 $E^{\langle\tau\rangle}=\mathbb{Q}(a)$。由共轭对应，其他两个二阶子群的固定域为

$$\mathbb{Q}(\zeta a),\qquad \mathbb{Q}(\zeta^2 a).$$

六阶群 S_3 的真非平凡子群只有唯一的三阶子群和三个二阶子群：子群阶由 Lagrange 定理只能为 2 或 3，它们分别由三个对换及两个互逆三循环生成。故上述四个域就是全部真非平凡中间域。

三阶子群正规，所以二次域 $\mathbb{Q}(\zeta)/\mathbb{Q}$ Galois，商群为 C_2；二阶子群不正规，所以三个三次中间域均不在 $\mathbb{Q}$ 上正规。特别地，$\mathbb{Q}(a)\subseteq\mathbb{R}$ 不包含另外两个非实根，这从根的角度直接说明了非正规性。

12.13 本原元：把有限生成改为单生成

定义 12.41 (本原元)**.** 若有限扩张 E/F 有元素 θ 使 $E=F(\theta)$，称 θ 为该扩张的本原元。

定理 12.42 (本原元定理)**.** 每个有限可分扩张都是简单扩张。若 F 无限且 $E=F(\alpha,\beta)$ 可分，则除有限多个 $c\in F$ 外，$\theta=\alpha+c\beta$ 都是本原元。

证明. 若 F 有限，有限次数使 E 也是有限集合。由命题 12.37，取 $E^\times$ 的生成元 θ，则每个非零元素是 θ 的幂，故 $E = F(\theta)$。

设 F 无限，先证二生成情形。取 α, β 的首一最小多项式 f, g，并在一个含 E 的有限扩张 L 中分裂它们。它们可分，列出不同根

$$\alpha = \alpha_1, \dots, \alpha_r, \qquad \beta = \beta_1, \dots, \beta_s.$$

若 $j \neq 1$，等式 $\alpha_i + c\beta_j = \alpha + c\beta$ 至多给出一个值

$$c = \frac{\alpha_i - \alpha}{\beta - \beta_j} \in L.$$

排除这些有限多个值中属于 F 的值，并排除 $c = 0$。由于 F 无限，仍有可选 $c \in F$；此时

$$\alpha_i + c\beta_j = \alpha + c\beta \quad \Longrightarrow \quad i = j = 1.$$

这里 $j = 1$ 时等式直接给出 $\alpha_i = \alpha$；$j \neq 1$ 已由选取排除。

令 $\theta = \alpha + c\beta$，$K = F(\theta)$。$g(t)$ 与 $f(\theta - ct)$ 属于 $K[t]$，共有根 β。若 β_j 是另一个公共根，则 $\theta - c\beta_j = \alpha_i$ 对某个 i 成立，与选取条件矛盾。由于 g 无重复根，这两个多项式在 $L[t]$ 中的首一最大公因式为 $t - \beta$。此前已证明扩大域不改变首一最大公因式，因此在 $K[t]$ 中算出的最大公因式同样是 $t - \beta$。其常数系数属于 K，故 $\beta \in K$；再由 $\alpha = \theta - c\beta$ 得 $\alpha \in K$。于是 $E \subseteq K$，反向 $\theta \in E$ 给出 $K \subseteq E$，故 $E = F(\theta)$。

一般有限可分扩张有有限生成元 $\alpha_1, \dots, \alpha_r$。先将前两个合并为一个，再与第三个合并，依次进行。每次的新元素仍在可分扩张 E 内，故在 F 上可分；二生成论证每次都适用。有限步后得到一个本原元。 □

文献：扩域与商域构造参见 [15] 第 1 章第 14–19 页；分裂域、嵌入及可分性参见其第 2 章；固定域与有限 Galois 对应参见第 3 章，尤其定理 3.4、3.10、3.17。有限域和本原元分别参见第 4 章第 53–54 页、第 5 章第 61–62 页。本文的最小支撑证明沿用 Milne 对 Artin 定理的组织方式，逐项展开方程、次数和限制映射。例子的基、逆元与中间域均按本文记号计算。

参考文献

[1] Herbert B. Enderton. *A Mathematical Introduction to Logic*. 2nd ed., Academic Press, 2001.

[2] William A. R. Weiss. *An Introduction to Set Theory*. University of Toronto, October 2, 2008. https://www.math.toronto.edu/weiss/set_theory.pdf

[3] Richard Hammack. *Book of Proof*. 3rd ed., 2018. https://richardhammack.github.io/BookOfProof/

[4] James S. Milne. *Group Theory*. Version 4.01, 2025. https://www.jmilne.org/math/CourseNotes/GT.pdf

[5] Paul R. Halmos. *Naive Set Theory*. Van Nostrand, 1960.

[6] Gottlob Frege. *Begriffsschrift*. Louis Nebert, 1879.

[7] David Hilbert and Wilhelm Ackermann. *Grundzüge der theoretischen Logik*. Springer, 1928.

[8] Michael Artin. *Algebra*. 2nd ed., Pearson, 2011.

[9] David S. Dummit and Richard M. Foote. *Abstract Algebra*. 3rd ed., Wiley, 2004.

[10] Sheldon Axler. *Linear Algebra Done Right*. 4th ed., Springer, 2024. https://linear.axler.net/

[11] Gilbert Strang. *Introduction to Linear Algebra*. 6th ed., Wellesley-Cambridge Press, 2023. 作者公开节选：https://math.mit.edu/~gs/linearalgebra/ila6/indexila6.html

[12] Michael F. Atiyah and Ian G. Macdonald. *Introduction to Commutative Algebra*. Addison-Wesley, 1969.

[13] Keith Conrad. *Modules over a PID*. 作者公开讲义, https://kconrad.math.uconn.edu/blurbs/linmultialg/modulesoverPID.pdf，访问于 2026-10-26。

[14] Keith Conrad. *Tensor Products*. 作者公开讲义，https://kconrad.math.uconn.edu/blurbs/linmultialg/tensorprod.pdf，访问于 2026-10-25。

[15] James S. Milne. *Fields and Galois Theory*. Version 5.10, 2022. https://www.jmilne.org/math/CourseNotes/FT.pdf

[16] Keith Conrad. *Exterior Powers*. 作者公开讲义，https://kconrad.math.uconn.edu/blurbs/linmultialg/extmod.pdf，访问于 2026-10-25。