

定义. 用这个术语, (4.1.6)变为

**【4.1.7】** 零化度 + 秩 =  $V$  的维数

**定理 4.1.6 的证明** 假设  $V$  是有限维的, 比如说  $n$  维. 令  $k$  是  $\ker T$  的维数, 并设  $(u_1, \dots, u_k)$  是核的基. 将该集合扩充为  $V$  的一个基:

**【4.1.8】**  $(u_1, \dots, u_k; v_1, \dots, v_{n-k})$

(见(3.4.15)). 对于  $i=1, \dots, n-k$ , 令  $w_i = T(v_i)$ . 如果证明  $C = (w_1, \dots, w_{n-k})$  是像的基, 则由此可得像  $\text{im} T$  的维数为  $n-k$ . 这样将证明定理.

我们需证  $C$  张成且它是一个线性无关集. 设  $w$  为像中的元素. 则对  $v \in V$  有  $w = T(v)$ . 用基写出  $v$ :

$$v = a_1 u_1 + \dots + a_k u_k + b_1 v_1 + \dots + b_{n-k} v_{n-k}$$

应用  $T$ , 注意到  $T(u_i) = 0$ :

$$w = T(v) = b_1 w_1 + \dots + b_{n-k} w_{n-k}$$

这样  $w$  属于  $C$  的张成.

下面证明  $C$  是无关的. 假设有线性关系

**【4.1.9】**  $c_1 w_1 + \dots + c_{n-k} w_{n-k} = 0$

令  $v = c_1 v_1 + \dots + c_{n-k} v_{n-k}$ , 其中  $v_i$  是基(4.1.8)中的向量. 于是,

$$T(v) = c_1 w_1 + \dots + c_{n-k} w_{n-k} = 0$$

因此,  $v$  属于零空间. 于是, 可用零空间的基  $(u_1, \dots, u_k)$  表出  $v$ , 比如说  $v = a_1 u_1 + \dots + a_k u_k$ . 则有  $-a_1 u_1 - \dots - a_k u_k + c_1 v_1 + \dots + c_{n-k} v_{n-k} = -v + v = 0$ . 但基(4.1.8)是无关的. 于是,  $-a_1 = 0, \dots, -a_k = 0$ , 且  $c_1 = 0, \dots, c_{n-k} = 0$ . 因此, 关系(4.1.9)是平凡的. 所以,  $C$  是线性无关集. ■

103

当  $T$  是矩阵  $A$  (4.1.3) 的左乘(变换)时,  $T$  的核(即  $A$  的零空间)是齐次方程  $AX=0$  的解集.  $T$  的像是列向量空间, 它是由  $A$  的列张成的空间, 也是  $F^n$  中使得线性方程  $AX=B$  有解(3.4.6)的向量  $B$  的集.

众所周知, 齐次方程  $AX=0$  的解加到非齐次方程  $AX=B$  的一个特解  $X_0$  上, 就得到非齐次方程的所有解. 关于此结论的另一个说法为  $AX=B$  的解集是  $F^n$  中零空间  $N$  的加法陪集  $X_0 + N$ .

其行列式不为零的  $n \times n$  矩阵  $A$  是可逆的, 对每个  $B$  方程  $AX=B$  有唯一解. 在此情形下, 零空间为  $\{0\}$ , 列空间是整个空间  $F^n$ . 另一方面, 如果行列式为零, 则零空间  $N$  有正维数, 而像(即列空间)的维数比  $n$  小. 不是所有的方程  $AX=B$  都有解, 但有解的那些方程的解不止一个, 因为解集是  $N$  的陪集.

## 第二节 线性变换的矩阵

每个从一个列向量空间到另一个列向量空间的线性变换是用一个矩阵左乘.

**【4.2.1】引理** 令  $T: F^n \rightarrow F^m$  是列向量空间之间的线性变换, 并设  $T(e_j)$  的坐标向量为  $A_j =$

$(a_{1j}, \dots, a_{2j}, a_{mj})'$ . 令  $A$  是一个以  $A_1, A_2, \dots, A_n$  为列向量的  $m \times n$  矩阵. 则  $T$  作用在  $F^n$  的向量上相当于左乘矩阵  $A$ .

**证明**  $T(X) = T(\sum_j e_j x_j) = \sum_j T(e_j) x_j = \sum_j A_j x_j = AX$ . ■

例如, 令  $c = \cos\theta$ ,  $s = \sin\theta$ . 令  $\rho: \mathbf{R}^2 \rightarrow \mathbf{R}^2$  是关于原点将平面逆时针旋转  $\theta$  角的线性变换. 其矩阵为

$$\text{【4.2.2】} \quad R = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

我们证明这个矩阵乘法使平面旋转. 将向量  $X$  记为  $r(\cos\alpha, \sin\alpha)'$  的形式, 此处  $r$  是向量  $X$  的长度. 令  $c' = \cos\alpha$ ,  $s' = \sin\alpha$ , 由两角和的正弦与余弦公式可以证明

$$RX = r \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c' \\ s' \end{bmatrix} = r \begin{bmatrix} cc' - ss' \\ sc' + cs' \end{bmatrix} = r \begin{bmatrix} \cos(\theta + \alpha) \\ \sin(\theta + \alpha) \end{bmatrix}$$

所以,  $RX$  是由向量  $X$  经过旋转  $\theta$  角得到的.

一旦给定两个空间的基, 就可以对任意线性变换  $T: V \rightarrow W$  作类似于引理 4.2.1 中的计算.

设  $B = (v_1, \dots, v_n)$  是空间  $V$  的一组基, 用简短记号  $T(B)$  表示超向量

$$\text{【4.2.3】} \quad T(B) = (T(v_1), \dots, T(v_n))$$

104

若  $v = BX = v_1 x_1 + \dots + v_n x_n$ , 则

$$\text{【4.2.4】} \quad T(v) = T(v_1)x_1 + \dots + T(v_n)x_n = T(B)X$$

**【4.2.5】命题** 令  $T: V \rightarrow W$  是一个线性变换, 且  $B = (v_1, \dots, v_n)$  和  $C = (w_1, \dots, w_m)$  分别是空间  $V$  和  $W$  的一组基. 令  $X$  是一个任意向量  $v$  在基  $B$  下的坐标向量, 令  $Y$  是其变换后的像  $T(v)$  的坐标向量. 于是  $v = BX$  且  $T(v) = CY$ . 存在一个  $m \times n$  矩阵  $A$  具有两个对偶的性质:

$$\text{【4.2.6】} \quad T(B) = CA, \quad AX = Y$$

矩阵  $A$  叫做线性变换  $T$  在这两组基下的变换矩阵. (4.2.6) 的任何一个公式都刻画了这个矩阵.

**证明** 记  $T(v_j)$  为基  $C$  的一个线性组合, 比如

$$\text{【4.2.7】} \quad T(v_j) = w_1 a_{1j} + \dots + w_m a_{mj}$$

把系数  $a_{ij}$  写成列向量  $A_j = (a_{1j}, \dots, a_{mj})'$ , 则  $T(v_j) = CA_j$ . 若  $A$  是以  $A_1, A_2, \dots, A_n$  为列向量的矩阵, 则

$$\text{【4.2.8】} \quad T(B) = (T(v_1), \dots, T(v_n)) = (w_1, \dots, w_m) \begin{bmatrix} A \end{bmatrix} = CA$$

其次, 若  $v = BX$ , 则

$$T(v) = T(B)X = CAX$$

因此,  $T(v)$  的坐标向量(记为  $Y$ )等于  $AX$ . ■

$T$  与  $A$  间的关系可以用由两个基(3.5.3)所确定的同构  $\phi: F^n \rightarrow V$  和  $\phi': F^m \rightarrow W$  加以解释. 如果用这些同构将  $V$  和  $W$  等同于  $F^n$  和  $F^m$ , 则  $T$  对应于用  $A$  左乘, 如下图所示:

## 【4.2.9】图

$$\begin{array}{ccc}
 F^n & \xrightarrow{A} & F^m \\
 B \downarrow & & \downarrow C \\
 V & \xrightarrow{T} & W
 \end{array}
 \qquad
 \begin{array}{ccc}
 X & \rightsquigarrow & AX \\
 \downarrow & & \downarrow \\
 BX & \rightsquigarrow & T(B)X = CAX
 \end{array}$$

从  $F^n$  到  $W$  沿两个方向行进得到相同的答案. 具有这种性质的图叫交换图. 本书里所有的图都是交换的.

因此, 一旦两个空间的基取定之后, 有限维向量空间  $V$  与  $W$  间的任意线性变换就可与矩阵乘法对应起来. 这是好的结果, 但如果我们变换基, 则可以做得更好.

105

## 【4.2.10】定理

(a) 向量空间形式: 设  $T: V \rightarrow W$  为有限维向量空间之间的线性变换. 则  $V$  与  $W$  分别有基  $B, C$  使得  $T$  的矩阵具有形式

## 【4.2.11】

$$A' = \begin{array}{|c|c|} \hline I_r & \\ \hline & 0 \\ \hline \end{array}$$

其中  $I_r$  为  $r \times r$  单位矩阵, 且  $r$  为  $T$  的秩.

(b) 矩阵形式: 给定任意  $m \times n$  矩阵  $A$ , 存在可逆矩阵  $Q$  和  $P$  使  $A' = Q^{-1}AP$  具有上面所示形式.

证明

(a) 设  $(u_1, \dots, u_k)$  为  $T$  的核的基. 将该基扩张为  $V$  的一个基  $B$ , 把增加的向量列在前面, 比如说  $(v_1, \dots, v_r; u_1, \dots, u_k)$ , 其中  $r+k=n$ . 设  $w_i = T(v_i)$ . 则如 (4.1.6) 的证明中所指出的,  $(w_1, \dots, w_r)$  是  $T$  的像的基. 将其扩张成  $W$  的基  $C$ , 把增加的向量列在后面, 比如说  $(w_1, \dots, w_r; z_1, \dots, z_s)$ .  $T$  关于这些基的矩阵具有形式 (4.2.11).

定理的 (b) 部分可用行变换和列变换证得. 该证明是练习 2.4. ■

这个定理是我们后面将证明的许多结果的原型. 因为任意线性变换的结构可以用一个非常简单的矩阵 (4.2.11) 来描述, 所以它展示了在向量空间中不用固定的基 (或坐标) 的优势. 但为什么所考虑的 (a) 与 (b) 是同一个定理的两种形式? 为回答这个问题, 我们需要分析当选择另一个基时线性变换矩阵变化的方式.

令  $A$  是如同 (4.2.6) 里  $T$  的关于  $V$  和  $W$  的基  $B$  和  $C$  的矩阵, 令  $B' = (v'_1, \dots, v'_n)$  与  $C' = (w'_1, \dots, w'_m)$  是  $V$  和  $W$  的新基. 如同 (3.5.11) 所示, 可用可逆  $n \times n$  矩阵  $P$  把新基  $B'$  与旧基  $B$  联系起来. 类似地, 用可逆  $m \times m$  矩阵  $Q$  把新基  $C'$  与旧基  $C$  联系起来. 这些矩阵有性质

$$\text{【4.2.12】} \quad B' = BP, PX' = X \quad \text{和} \quad C' = CQ, QY' = Y$$

【4.2.13】命题 令  $A$  是线性变换  $T$  关于给定基  $B$  和  $C$  的矩阵.

(a) 设新基  $B'$  与  $C'$  由矩阵  $P$  与  $Q$  同旧基  $B$  和  $C$  联系起来, 如上所示, 则  $T$  关于新基的矩阵是  $A' = Q^{-1}AP$ .

(b) 表示  $T$  关于其他基的矩阵  $A'$  均有形式  $A' = Q^{-1}AP$ , 其中  $Q$  和  $P$  为任意行列数适当的可逆矩阵.

## 证明

(a) 在方程  $Y=AX$  (4.2.6) 里做替换  $X=PX'$  与  $Y=QY'$ , 得  $QY'=APX'$ . 于是,  $Y'=(Q^{-1}AP)X'$ . 因为  $A'$  是使得  $A'X'=Y'$  的矩阵, 这表明  $A'=Q^{-1}AP$ .

(b) 因为基变换矩阵可以是任一个可逆矩阵 (3.5.9), 故 (b) 部分结论得证. ■

106

由该命题可知定理的两部分相当于同一回事. 为由 (b) 推证 (a), 假设已知线性变换  $T$ , 并且我们从  $V$  与  $W$  的任意基开始, 得到矩阵  $A$ . (b) 部分告诉我们存在可逆矩阵  $P$  与  $Q$  使得  $A'=Q^{-1}AP$  有形式 (4.2.11). 当我们用这些矩阵变换  $V$  与  $W$  的基时, 矩阵  $A$  变成矩阵  $A'$ .

为由 (a) 推证 (b), 我们视任一矩阵  $A$  为关于列向量“由  $A$  左乘”线性变换的矩阵. 那么,  $A$  为  $T$  关于  $F^n$  与  $F^m$  的标准基的矩阵, 并且 (a) 保证  $P$  与  $Q$  的存在使得  $Q^{-1}AP$  有形式 (4.2.11).

由于矩阵左乘是线性变换, 因此关于矩阵乘法我们这里也学到了一些重要东西. 任一矩阵  $A$  的左乘与 (4.2.11) 形式矩阵的左乘是相同的, 但它们是关于不同坐标的.

在将来, 我们常常用两种等价方式叙述一个结果, 即用向量空间形式与矩阵形式, 而不再证明两种形式是等价的. 这样, 我们将提供看上去更简单的一种证明.

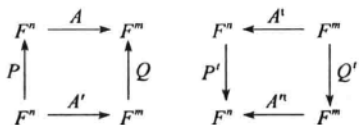
可以用定理 4.2.10 推导出矩阵乘法的另一个有趣的性质. 令  $N$  与  $U$  表示变换  $A:F^n \rightarrow F^m$  的零空间与列空间. 于是,  $N$  是  $F^n$  的子空间,  $U$  是  $F^m$  的子空间. 令  $k$  与  $r$  分别表示  $N$  与  $U$  的维数. 因此,  $k$  是  $A$  的零度,  $r$  是它的秩.

转置矩阵  $A'$  的左乘定义了反方向的变换  $A':F^m \rightarrow F^n$  以及另两个子空间:  $A'$  的零子空间  $N_1$  与列空间  $U_1$ . 这里  $U_1$  是  $F^n$  的子空间,  $N_1$  是  $F^m$  的子空间. 令  $k_1$  与  $r_1$  分别表示  $N_1$  与  $U_1$  的维数. 定理 4.1.6 告诉我们  $k+r=n$ , 且还有  $k_1+r_1=m$ . 下面的定理 4.2.14 给出这些整数的另外一个关系.

**【4.2.14】定理** 使用上面的记号,  $r_1=r$ : 矩阵的秩等于其转置矩阵的秩.

**证明** 令  $P$  与  $Q$  是可逆矩阵使得  $A'=Q^{-1}AP$  有形式 (4.2.11). 首先注意对于矩阵  $A'$ , 结论是显然成立的. 其次, 我们检验图

**【4.2.15】图**



球极平面射影

竖直箭头是双射. 所以, 在左边的图里,  $Q$  把  $A'$  的列空间 ( $A'$  乘法的像) 双射地映到  $A$  的列空间上. 这两个列空间的维数 (即  $A$  的秩与  $A'$  的秩) 是相等的. 类似地,  $A'$  的秩与  $A$  的秩也是相等的. 所以, 为证明本定理, 我们可用  $A'$  替换矩阵  $A$ . 这就把证明简化到了矩阵的平凡情形 (4.2.11). ■

107

我们能够重新解释转置矩阵  $A'$  的秩  $r_1$ . 由定义, 它是由  $A'$  的列张成的空间的维数, 而这可以看成是由  $A$  的行张成的行向量空间的维数. 因此,  $r_1$  常常称为  $A$  的行秩, 且  $r$  常



常称为  $A$  的列秩. 行秩是矩阵无关行的最大数, 列秩是矩阵无关列的最大数. 定理 4.2.14 可叙述如下:

**【4.2.16】推论**  $m \times n$  矩阵的行秩与列秩是相等的.

### 第三节 线性算子

本节讨论一个向量空间到其自身的线性变换  $T: V \rightarrow V$ . 这样的线性变换称为线性算子. 用元素属于  $F$  的  $n \times n$  矩阵左乘定义了列向量空间  $F^n$  的一个线性算子.

例如, 令  $c = \cos\theta$ ,  $s = \sin\theta$ , 旋转矩阵(4.2.2)

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

是平面  $\mathbf{R}^2$  上的一个线性算子.

对于线性算子, 维数公式  $\dim(\ker T) + \dim(\operatorname{im} T) = \dim V$  是成立的. 但这里, 由于定义域和值域是相等的, 因此我们有关于维数公式更多的信息.  $T$  的核与像都是  $V$  的子空间.

**【4.3.1】命题** 令  $K$  与  $W$  分别表示有限维向量空间  $V$  上线性算子  $T$  的核与像.

(a) 下列条件是等价的:

- $T$  是双射,
- $K = \{0\}$ ,
- $W = V$ .

(b) 下列条件是等价的:

- $V$  是直和  $K \oplus W$ ,
- $K \cap W = \{0\}$ ,
- $K + W = V$ .

**证明**

(a)  $T$  是双射当且仅当核  $W$  是零且像  $W$  是整个空间  $V$ . 如果核是零, 由维数公式知  $\dim W = \dim V$ . 所以,  $W = V$ . 类似地, 如果  $W = V$ , 维数公式表明  $\dim K = 0$ . 所以,  $K = \{0\}$ . 在这两种情形下,  $T$  是双射.

(b)  $V$  是直和  $K \oplus W$  当且仅当条件  $K \cap W = \{0\}$  与  $K + W = V$  均成立. 如果  $K \cap W = \{0\}$ , 那么,  $K$  与  $W$  是无关的. 于是, 和  $U = K + W$  是直和  $K \oplus W$ , 且  $\dim U = \dim K + \dim W$  (3.6.6)(a). 维数公式表明  $\dim U = \dim V$ . 所以,  $U = V$ , 这表明  $K \oplus W = V$ . 如果  $K + W = V$ , 则维数公式与命题 3.6.6(a) 表明  $K$  与  $W$  是无关的, 且  $V$  是直和. ■

108

**注** 满足条件(4.3.1)(a)的线性算子称为可逆算子. 它的逆函数也是线性算子. 不是可逆的算子称为奇异算子.

当  $V$  是无限维时, 命题 4.3.1(a)的条件不是等价的. 例如, 令  $V = \mathbf{R}^\infty$  是无限行向量  $(a_1, a_2, \dots)$  的空间(见第三章第七节). 如下定义

**【4.3.2】**  $S^+(a_1, a_2, \dots) = (0, a_1, a_2, \dots)$

的右移位算子  $S^+$  的核是零空间, 而它的像是  $V$  的真子空间. 由

$$S^-(a_1, a_2, a_3, \dots) = (a_2, a_3, \dots)$$

定义的左移位算子  $S^-$  的核是  $V$  真子空间, 而它的像是整个空间.

当讨论线性算子时, 上节关于基的讨论需作少许改动. 显然, 我们希望在  $V$  中只取一个基  $B = (v_1, \dots, v_n)$ , 用它代替(4.2.6)中的基  $B$  和  $C$ . 换言之, 为定义  $T$  关于  $B$  的矩阵  $A$ , 我们应该写出

**[4.3.3]**  $T(B) = BA$ , 且  $AX = Y$  如前所述.

像任一个线性变换(4.2.7)那样,  $A$  的列是基向量的像  $T(v_j)$  的坐标向量:

**[4.3.4]**  $T(v_j) = v_1 a_{1j} + \dots + v_n a_{nj}$

一个线性算子是可逆的当且仅当它关于任一个基的矩阵是可逆矩阵.

当说到空间  $F^n$  上线性算子的矩阵时, 除非特别声明, 总假设基是标准基  $E$ . 因此, 算子那个矩阵的乘法.

当我们研究基变换的作用时, 一个新的特征就出现了. 假设  $B$  由新基  $B'$  替换.

**[4.3.5] 命题** 令  $A$  是线性算子  $T$  关于  $B$  的矩阵.

(a) 假设新基  $B'$  由  $B' = BP$  给出.  $T$  关于这个基的矩阵是  $A' = P^{-1}AP$ .

(b) 算子  $T$  对于不同基的矩阵  $A'$  都是形如  $A' = P^{-1}AP$  的矩阵, 其中  $P$  可以是任意的可逆矩阵.

换言之, 是由共轭所进行的矩阵变换. 这是一个易混淆的事实. 所以, 尽管它可从(4.2.13)得到, 我们还是要重新推证它. 因为  $B' = BP$ , 且  $T(B) = BA$ , 我们有

$$T(B') = T(B)P = BAP$$

还没完全结束. 我们所得的公式用旧基  $B$  表示  $T(B')$ . 为得到新矩阵, 必须根据新基  $B'$  写出  $T(B')$ . 所以, 我们把  $B = B'P^{-1}$  替换进方程, 于是, 得到  $T(B') = B'P^{-1}AP$ .

109

一般地, 我们说方阵  $A$  相似于另一个矩阵  $A'$ , 如果对某个可逆矩阵  $P$ , 有  $A' = P^{-1}AP$ . 这样的矩阵  $A'$  是从  $A$  由  $P^{-1}$  共轭得到的. 因为  $P$  是任意可逆矩阵, 故  $P^{-1}$  也是任意的. 用术语共轭代替相似是正确的.

现在, 如果已知矩阵  $A$ , 很自然地要寻找特别简单的相似矩阵  $A'$ . 我们希望得到有点儿类似定理 4.2.10 的结果. 但这里允许的变换是非常苛刻的, 因为仅有一个基, 从而仅有一个矩阵供使用. 线性变换的定义域和值域是相等的, 这初看上去似乎会简单化问题, 但实际上, 这会使问题更难.

通过把假定的基变换矩阵写成初等矩阵的乘积, 比如说,  $P = E_1 \dots E_r$ , 会对问题有所洞悉. 于是,

$$P^{-1}AP = E_r^{-1} \dots E_1^{-1}AE_1 \dots E_r$$

由初等变换, 允许对  $A$  进行若干步变换:  $A \rightsquigarrow E^{-1}AE$ . 换言之, 我们可对  $A$  进行任意的列变换  $E$ , 但也必须做与逆矩阵  $E^{-1}$  对应的行变换. 不幸的是, 这些行列变换相互干扰, 这使得不能直接分析这些作用的效果.

## 第四节 特征向量

分析线性算子  $T: V \rightarrow V$  的主要工具是特征向量和不变子空间.

注  $V$  的一个子空间  $W$  称为不变的或  $T$ -不变的, 如果它在算子  $T$  的作用下变到自身:

$$\text{【4.4.1】} \quad TW \subset W$$

换言之, 如果对所有  $w \in W$ , 有  $T(w) \in W$ , 则  $W$  是  $T$ -不变的. 当  $W$  为  $T$ -不变的时,  $T$  在  $W$  上定义一个线性算子, 称为  $T$  在  $W$  上的限制. 我们常常记这个限制为  $T|_W$ .

如果  $W$  为  $T$ -不变子空间, 则可构造  $V$  的一个基  $B$ , 它由在  $W$  的一个基  $(w_1, \dots, w_k)$  上添加向量而得到:

$$\text{【4.4.2】} \quad B = (w_1, \dots, w_k, v_1, \dots, v_{n-k})$$

这样,  $W$  是不变子空间这一事实可以从  $T$  的矩阵中看出来. 这个矩阵的列为像向量的坐标向量(见(4.2.3)), 称这个矩阵为  $M$ . 但  $T(w_j)$  属于子空间  $W$ , 从而它是基  $(w_1, \dots, w_k)$  的线性组合. 因此, 当我们把  $T(w_j)$  用基  $B$  表出时, 向量  $v_1, \dots, v_{n-k}$  的系数为零. 由此, 矩阵  $M$  具有分块形式

$$\text{【4.4.3】} \quad M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

110 其中  $A$  是  $k \times k$  矩阵, 为  $T$  在  $W$  上的限制的矩阵.

如果  $V$  恰好为两个  $T$ -不变子空间的直和  $W_1 \oplus W_2$ , 并且把  $W_1$  与  $W_2$  的基顺序排起来构成  $V$  的一个基  $B = (B_1, B_2)$ . 这时,  $T$  的矩阵为分块对角形式

$$\text{【4.4.4】} \quad M = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

其中  $A_i$  是  $T$  在  $W_i$  上的限制的矩阵.

特征向量的概念与不变子空间的概念是紧密联系的.

注 线性算子  $T$  的特征向量  $v$  是对某个标量  $\lambda$ , 亦即  $F$  里的某个元素, 使得

$$\text{【4.4.5】} \quad T(v) = \lambda v$$

的非零向量. 非零列向量是方阵  $A$  的特征向量, 如果它是由  $A$  左乘变换的特征向量.

出现在(4.4.5)里的标量  $\lambda$  称为相伴于特征向量  $v$  的特征值. 当我们说到线性算子  $T$  或未指定特征向量的矩阵  $A$  的特征值时, 指的是一个标量  $\lambda \in F$ , 它是相伴于某个特征向量的特征值. 特征值可以是  $F$  的任意元素, 包括零, 但特征向量不能为零. 这里, 特征值常用希腊字母  $\lambda$  (lambda) 来记<sup>⊖</sup>.

相伴于特征值 1 的特征向量是固定不动的向量:  $T(v) = v$ . 相伴于特征值 0 的特征向

⊖ 德语“eigen”大意是“characteristic”(特征的), Eigenvector 和 eigenvalue 有时也叫 characteristic vector 和 characteristic value.

eigenvector 与 eigenvalue 应译成“本征向量”与“本征值”. 但在国内代数教材里, 一般都叫特征向量与特征值. 所以, 在翻译中, 我们把 eigenvector 与 eigenvalue 译成“特征向量”与“特征值”. ——译者注

量属于零空间:  $T(v)=0$ . 当  $V=\mathbf{R}^n$  时, 非零向量  $v$  是特征向量, 如果  $v$  与  $T(v)$  平行.

如果  $v$  是线性算子  $T$  的相伴于特征值  $\lambda$  的特征向量, 由  $v$  张成的子空间  $W$  是  $T$ -不变的, 因为在  $W$  中  $T(cv)=c\lambda v$  对所有标量  $c$  成立. 反之, 如果由  $v$  张成的 1 维子空间是不变的, 则  $v$  是一个特征向量. 于是, 一个特征向量可以描述为 1 维不变子空间的基.

容易判别一个给定向量  $X$  是否为矩阵  $A$  的特征向量. 可简单验证  $AX$  是否为  $X$  的倍数. 如果  $A$  是  $T$  关于基  $B$  的矩阵, 并且如果  $X$  是向量  $v$  的坐标向量, 则  $X$  是  $A$  的特征向量当且仅当  $v$  是  $T$  的特征向量.

标准基向量  $e_1=(1, 0)^t$  是矩阵

$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

的相伴于特征值 3 的特征向量, 向量  $(1, -1)^t$  是相伴于特征值 2 的另一特征向量. 向量  $(0, 1, 1)^t$  是矩阵

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ 3 & 0 & 2 \end{bmatrix}$$

的相伴于特征值 2 的特征向量.

111

如果  $(v_1, \dots, v_n)$  是  $V$  的一个基, 并且如果  $v_1$  是线性算子  $T$  的特征向量, 则  $T$  的矩阵有分块形式

$$\text{【4.4.6】} \quad \begin{bmatrix} \lambda & B \\ 0 & D \end{bmatrix} = \left[ \begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & * \end{array} \right]$$

其中  $\lambda$  是  $v_1$  相伴的特征值. 这是 1 维不变子空间情形下的分块形式(4.4.3).

**【4.4.7】命题** 相似矩阵( $A'=P^{-1}AP$ )有相同的特征值.

因为相似矩阵表示同一线性变换, 故结论成立.

**【4.4.8】命题**

(a) 令  $T$  是向量空间  $V$  上的线性算子.  $T$  关于基  $B=(v_1, \dots, v_n)$  的矩阵是对角矩阵当且仅当每个基向量  $v_j$  是特征向量.

(b)  $n \times n$  矩阵  $A$  与对角矩阵相似当且仅当  $F^n$  里有一个基是由特征向量构成的.

由矩阵  $A$  的定义(见(4.3.4))可得证. 如果  $T(v_j)=\lambda_j v_j$ , 则

$$\text{【4.4.9】} \quad T(B) = (v_1 \lambda_1, \dots, v_n \lambda_n) = (v_1, \dots, v_n) \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

这个命题表明我们可简单地用对角矩阵表示线性算子, 倘若它有足够的特征向量. 在本章第五节我们将看到, 复向量空间上每个线性算子至少有一个特征向量, 在本章第六节

将看到大多数情形下存在一个特征向量基. 但实向量空间上的线性算子不一定有特征向量. 例如, 平面上通过一个角  $\theta$  的旋转, 除非  $\theta=0$  或  $\pi$ , 否则不会把任何一个向量变换到与其平行的向量. 因此, 除了  $\theta=0$  或  $\pi$  的情形外, 旋转矩阵(4.2.2)没有实特征向量.

**注** 至少有一个实特征值的实矩阵的一般情形是具有正元素的实矩阵. 这样的矩阵称为正矩阵. 它们在实际应用中经常出现, 其最重要的性质之一是总有一个坐标为正数的特征向量(正特征向量).

我们不证明这个事实, 而是通过考察在  $\mathbf{R}^2$  上正  $2 \times 2$  矩阵  $A$  的乘法的作用来加以说明. 设  $w_i = Ae_i$  是  $A$  的列向量. 向量加法的平行四边形法则指出,  $A$  将第一象限  $S$  映到由向量  $w_1$  和  $w_2$  所界定的扇形. 而  $w_i$  的坐标向量是  $A$  的第  $i$  列. 因为  $A$  的元素都是正的, 故向量  $w_i$  都在第一象限中. 从而  $A$  把第一象限映到第一象限:  $S \supset AS$ . 再用  $A$  作用, 得  $AS \supset A^2S$ , 继续下去, 有

$$\text{【4.4.10】} \quad S \supset AS \supset A^2S \supset A^3S \supset \dots$$

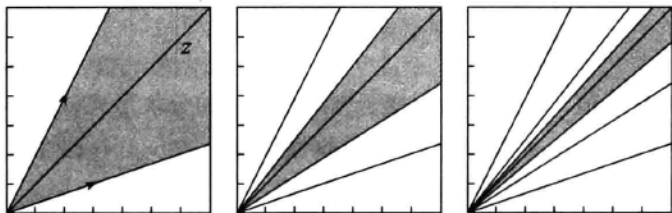
如下面矩阵  $A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$  所示.

现在, 扇形嵌套集  $Z = \bigcap A^n S$  的交或为一个扇形或为一条半直线. 这里,  $Z = \bigcap A^n S$  为半直线. 从直观上看这是合理的, 也可以用各种方法来证明, 但我们省略证明. 在关系  $Z = \bigcap A^n S$  的两边用  $A$  乘, 得到

$$AZ = A\left(\bigcap_0^\infty A^n S\right) = \bigcap_1^\infty A^n S = Z$$

因此,  $Z = AZ$ . 所以,  $Z$  中的非零向量是特征向量.

**【4.4.11】图**



第一象限在正矩阵重复乘积之下的像

## 第五节 特征多项式

本节确定任意线性算子  $T$  的特征向量. 我们先回顾一下,  $T$  的特征向量是使得

$$\text{【4.5.1】} \quad T(v) = \lambda v$$

对某个  $\lambda \in F$  成立的非零向量  $v$ . 如果不知道  $\lambda$ , 当线性算子相应的矩阵很复杂时似乎很难求出其特征向量. 诀窍是转而解决另一个问题, 即先求特征值. 当特征值  $\lambda$  确定后, 方程(4.5.1)成为  $v$  的坐标的线性方程组, 对其求解是没有问题的.

首先, 将(4.5.1)写为形式

$$\text{【4.5.2】} \quad [\lambda I - T](v) = 0$$

其中  $I$  为恒等算子, 而  $\lambda I - T$  是由

$$\text{【4.5.3】} \quad [\lambda I - T](v) = \lambda v - T(v)$$

定义的线性算子. 容易验证  $\lambda I - T$  的确是线性算子. (4.5.2) 可复述为

**【4.5.4】** 非零向量  $v$  是伴随于特征值  $\lambda$  的特征向量  
当且仅当它在  $\lambda I - T$  的核中.

**【4.5.5】推论** 令  $T$  是有限维向量空间  $V$  上的线性算子.

(a)  $T$  的特征值为  $F$  中使得  $\lambda I - T$  是奇异算子的标量  $\lambda$ , 亦即它的零空间是非零的.

(b) 下列条件是等价的:

- $T$  是奇异算子.
- $0$  为  $T$  的特征值.
- 如果  $A$  是  $T$  关于任意基的矩阵, 则  $\det A = 0$ .

若  $T$  关于某个基的矩阵为  $A$ , 则  $\lambda I - T$  的矩阵是  $\lambda I - A$ . 于是,  $\lambda I - T$  是奇异的当且仅当  $\det(\lambda I - A) = 0$ . 这个带有未定元  $\lambda$  的行列式可以计算出来, 并且这至少在原理上给我们提供了一个确定特征值和特征向量的方法.

例如, 假设  $A$  为矩阵  $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ , 它在  $\mathbf{R}^2$  上的作用如图 4.4.11 所示. 则

$$\lambda I - A = \begin{bmatrix} \lambda - 3 & -2 \\ -1 & \lambda - 4 \end{bmatrix}$$

且

$$\det(\lambda I - A) = \lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2)$$

当  $\lambda = 5$  或  $2$  时, 行列式为零, 于是  $A$  的特征值为  $5$  或  $2$ . 要求特征向量, 解两个线性方程组  $(5I - A)X = 0$  和  $(2I - A)X = 0$ . 其解在不计标量因子时是确定的:

$$\text{【4.5.6】} \quad v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

现在我们对任意未定元矩阵作同样的计算. 改变符号会方便一些. 习惯上用变量  $t$  代替符号  $\lambda$ . 构造矩阵  $tI - A$ :

$$\text{【4.5.7】} \quad tI - A = \begin{bmatrix} (t - a_{11}) & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & (t - a_{22}) & \cdots & -a_{2n} \\ \vdots & & & \\ -a_{n1} & \cdots & \cdots & (t - a_{nn}) \end{bmatrix}$$

于是, 行列式的完全展开[第一章(1.6.4)]指出  $\det(tI - A)$  是  $t$  的  $n$  次多项式, 其系数为标量, 它是  $F$  中的元素.

**【4.5.8】定义** 线性算子  $T$  的特征多项式是

$$p(t) = \det(tI - A)$$

其中  $A$  是  $T$  关于某个基的矩阵.

结合(4.5.5)与(4.5.8),  $T$  的特征值确定下来:

**【4.5.9】推论** 线性算子的特征值是其特征多项式的根.

**【4.5.10】推论** 令  $A$  是  $n \times n$  上三角矩阵或下三角矩阵, 其对角元为  $a_{11}, \dots, a_{nn}$ , 则  $A$  的特征多项式是  $(t-a_{11})(t-a_{22})\cdots(t-a_{nn})$ ,  $A$  的对角元为其特征值.

**证明** 如果  $A$  是上三角矩阵, 则  $tI-A$  也是, 并且  $tI-A$  的对角元为  $t-a_{ii}$ . 三角矩阵的行列式为其对角线元的积. ■

**【4.5.11】命题** 算子  $T$  的特征多项式与基的选择无关.

**证明** 第二个基相应的矩阵为  $A' = P^{-1}AP$  (命题 4.3.5), 且

$$tI - A' = tI - P^{-1}AP = P^{-1}(tI - A)P$$

于是,

$$\det(tI - A') = \det P^{-1} \det(tI - A) \det P = \det(tI - A) \quad \blacksquare$$

$2 \times 2$  矩阵  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  的特征多项式是

$$\text{【4.5.12】 } p(t) = \det(tI - A) = \det \begin{bmatrix} t-a & -b \\ -c & t-d \end{bmatrix} = t^2 - (\text{trace}A)t + (\det A)$$

其中迹  $\text{trace}A = a+d$ .

下一个命题给出了  $n \times n$  矩阵的特征多项式的不完整描述, 该命题可通过计算证明. 不难确定余下的系数, 但它们的简明公式不常用.

**【4.5.13】命题**  $n \times n$  矩阵  $A$  的特征多项式具有

$$p(t) = t^n - (\text{trace}A)t^{n-1} + (\text{中间项}) + (-1)^n(\det A)$$

的形式, 其中,  $A$  的迹  $\text{trace}A$  是对角元的和

$$\text{trace}A = a_{11} + a_{22} + \cdots + a_{nn}$$

115

命题 4.5.11 表明特征多项式的所有系数与基无关. 例如,  $\text{trace}(P^{-1}AP) = \text{trace}A$ .

因为特征多项式、迹和行列式都是与基无关的, 所以它们仅依赖于算子  $T$ , 因而可以定义线性算子  $T$  的特征多项式、迹和行列式. 它们是由  $T$  关于任意基的矩阵所得到的.

**【4.5.14】命题** 设  $T$  是有限维向量空间  $V$  上的线性算子.

(a) 若  $V$  的维数为  $n$ , 则  $T$  最多有  $n$  个特征值.

(b) 若  $F$  为复数域且  $V \neq \{0\}$ , 则  $T$  至少有一个特征值, 因而它至少有一个特征向量.

**证明**

(a) 特征值是次数为  $n$  的特征多项式的根.  $n$  次多项式最多有  $n$  个不同的根. 这对系数属于任意域  $F$  里的多项式都是成立的(见(12.2.20)).

(b) 代数基本定理表明每一复系数的正次数多项式至少有一个复根. 在第十五章(15.10.1)中有代数基本定理的一个证明. ■

例如, 设  $R_\theta$  为表示实平面  $\mathbf{R}^2$  上转过角度  $\theta$  的逆时针旋转的矩阵(4.2.2). 其特征多项式为  $p(t) = t^2 - (2\cos\theta)t + 1$ , 除了  $\theta = 0, \pi$  外它没有实根, 从而它没有实特征值时. 前面我们观察到了这个结果. 但是, 由  $R_\theta$  在  $\mathbf{C}^2$  上定义的算子有复特征值  $e^{\pm i\theta}$ .

**注意** 当我们说到多项式  $p(t)$  的根或矩阵或线性算子的特征值时, 对应重根的重数是要假设包括在内的. 这个术语尽管不准确, 但是很方便.

**【4.5.15】推论** 如果  $\lambda_1, \dots, \lambda_n$  是  $n \times n$  复矩阵  $A$  的特征值, 则  $\det A$  是积  $\lambda_1 \cdots \lambda_n$ , 且迹  $\text{trace} A$  是和  $\lambda_1 + \cdots + \lambda_n$ .

**证明** 令  $p(t)$  是  $T$  的特征多项式. 则

$$(t - \lambda_1) \cdots (t - \lambda_n) = p(t) = t^n - (\text{trace} A)t^{n-1} + \cdots \pm (\det A) \quad \blacksquare$$

## 第六节 三角形与对角形

本节证明对“大多数”复向量空间的线性算子, 存在一个基, 使得算子的矩阵是对角的. 其关键事实我们在第五节的结尾处已注意到, 即每一个正次数的复多项式都有一个根. 这表明每个线性算子都至少有一个特征向量.

### 【4.6.1】命题

(a) 向量空间形式: 设  $T$  是有限维复向量空间  $V$  上的线性算子. 存在  $V$  的基  $B$ , 使得  $T$  关于这个基的矩阵为上三角的.

(b) 矩阵形式: 每一个  $n \times n$  复矩阵  $A$  相似于一个上三角矩阵. 换言之, 存在矩阵  $P \in GL_n(\mathbf{C})$ , 使得  $P^{-1}AP$  为上三角的. 116

**证明** 由(4.3.5), 两个断言是等价的. 我们将用矩阵讨论. 令  $V = \mathbf{C}^n$ . 命题 4.5.14 (b) 表明  $V$  含有  $A$  的一个特征向量, 称之为  $v_1$ . 设  $\lambda$  是它的特征值. 把  $v_1$  扩张为  $V$  的一个基  $B = (v_1, \dots, v_n)$ . 新矩阵  $A' = P^{-1}AP$  有分块形式

$$\text{【4.6.2】} \quad A' = \left[ \begin{array}{c|c} \lambda & * \\ \hline 0 & D \end{array} \right]$$

其中  $D$  是一个  $(n-1) \times (n-1)$  矩阵(见(4.4.6)). 对  $n$  应用归纳法, 可假设已证明存在某个  $Q \in GL_{n-1}(\mathbf{C})$ , 使得  $Q^{-1}DQ$  是上三角的. 令

$$Q_1 = \left[ \begin{array}{c|c} 1 & 0 \\ \hline 0 & Q \end{array} \right]$$

则  $A'' = Q_1^{-1}A'Q_1 = \left[ \begin{array}{c|c} \lambda & * \\ \hline 0 & Q_1^{-1}DQ_1 \end{array} \right]$  为上三角的, 且  $A'' = (PQ_1)^{-1}A'(PQ_1)$ . \blacksquare

**【4.6.3】推论** 当把短语“上三角的”替换为“下三角的”时, 命题 4.6.1 仍然成立.

把命题 4.6.1(a) 的基  $B$  反序列就得到下三角形式.

证明命题 4.6.1 的要点是每个复多项式都有根. 同样的证明对任意域  $F$  都成立, 倘若特征多项式的根都在这个域里.

### 【4.6.4】推论

(a) 向量空间形式: 设  $T$  是域  $F$  上有限维向量空间  $V$  上的线性算子, 且假设  $T$  的特征



多项式在域  $F$  中为线性因子之积. 则存在  $V$  的基  $B$ , 使得  $T$  的矩阵  $A$  为上(或下)三角的.

(b) 矩阵形式: 设  $A$  是元素在  $F$  中的  $n \times n$  矩阵, 其特征多项式在域  $F$  中为线性因子之积. 则存在矩阵  $P \in GL_n(F)$ , 使得  $P^{-1}AP$  为上(或下)三角的.

证明是相同的, 除了在归纳步骤中需要验证出现在(4.6.2)里的矩阵  $D$  的特征多项式为  $p(t)/(t-\lambda)$ , 其中  $p(t)$  是  $A$  的特征多项式. 这样, 我们对  $A$  的特征多项式分解为线性因子乘积的假设对于  $D$  也成立.

我们现在问: 哪些矩阵相似于对角矩阵. 它们称为可对角化矩阵. 如在(4.4.8)(b)中所见, 它们是以特征向量为基的矩阵  $A$ . 类似地, 有特征向量基的线性算子称为可对角化算子. 对角元素除了其顺序外是由线性算子  $T$  决定的. 它们是特征根.

117

下面的定理 4.6.6 得出问题的部分解答; 更完整的解答将在下节给出.

**【4.6.5】命题** 设  $v_1, \dots, v_r$  为线性算子  $T$  的特征向量, 伴随有不同的特征值  $\lambda_1, \dots, \lambda_r$ . 则集合  $(v_1, \dots, v_r)$  无关.

**证明** 对  $r$  作数学归纳. 当  $r=1$  时, 结论成立, 因为特征值不为零. 设给定相关关系

$$0 = a_1 v_1 + \dots + a_r v_r$$

我们要证对所有  $i$  有  $a_i = 0$ . 应用线性算子  $T$ :

$$0 = T(0) = a_1 T(v_1) + \dots + a_r T(v_r) = a_1 \lambda_1 v_1 + \dots + a_r \lambda_r v_r$$

这是  $(v_1, \dots, v_r)$  中的第二个相关关系. 我们从两个关系中消去  $v_r$ , 将第一个关系乘上  $\lambda_r$  并减去第二个:

$$0 = a_1 (\lambda_r - \lambda_1) v_1 + \dots + a_{r-1} (\lambda_r - \lambda_{r-1}) v_{r-1}$$

应用归纳法, 假设  $(v_1, \dots, v_{r-1})$  是无关的. 于是, 系数  $a_i (\lambda_r - \lambda_i) (i < r)$  全为零. 因为  $\lambda_i$  互不相同, 若  $i < r$ , 则  $\lambda_r - \lambda_i$  不为 0. 这样  $a_1 = \dots = a_{r-1} = 0$ , 而原来的关系化简为  $a_r v_r = 0$ . 因为特征向量不为零, 故亦有  $a_r = 0$ . ■

下面的定理由(4.4.8)与(4.6.5)结合起来得到.

**【4.6.6】定理** 设  $T$  是域  $F$  上  $n$  维向量空间  $V$  的线性算子. 如果其特征多项式在  $F$  中有  $n$  个不同的根, 则存在  $V$  的基使得  $T$  关于它的矩阵为对角形.

**注意** 对角化是个有力的工具. 当给出可对角化算子时, 应自动反应用特征向量基分析讨论.

作为对角化的例子, 考虑矩阵

**【4.6.7】** 
$$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$$

其特征向量在(4.5.6)中已计算出. 这些特征向量构成  $\mathbf{R}^2$  的基  $B = (v_1, v_2)$ . 根据(3.5.13), 联系标准基  $E$  与这个基  $B$  的矩阵为

**【4.6.8】** 
$$P = [B] = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \quad P^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$$

$$\text{【4.6.9】} \quad P^{-1}AP = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & \\ & 2 \end{bmatrix} = \Lambda$$

下个命题是命题 4.4.8 的变化形式. 我们略去它的证明.

118

**【4.6.10】命题** 令  $F$  为域.

(a) 设  $T$  是  $F^n$  上的线性算子. 如果  $\mathbf{B}=(v_1, \dots, v_n)$  是由  $T$  的特征向量构成的基, 并且若  $P=[\mathbf{B}]$ , 则  $\Lambda=P^{-1}AP=[\mathbf{B}]^{-1}A[\mathbf{B}]$  是对角的.

(b) 令  $\mathbf{B}=(v_1, \dots, v_n)$  是  $F^n$  的基, 并且设  $\Lambda$  是对角矩阵, 其对角元  $\lambda_1, \dots, \lambda_n$  不一定互不相同, 则对  $i=1, \dots, n$ , 存在唯一的矩阵  $A$  使得  $v_i$  是  $A$  的 (伴随于) 特征值  $\lambda_i$  的特征向量, 也就是矩阵  $[\mathbf{B}]\Lambda[\mathbf{B}]^{-1}$ .

书写  $[\mathbf{B}]^{-1}A[\mathbf{B}]=\Lambda$  的好方式为

$$\text{【4.6.11】} \quad A[\mathbf{B}] = [\mathbf{B}]\Lambda$$

定理 4.6.6 的一个应用是计算可对角化矩阵的幂. 下一个引理需要指出, 尽管当展开方程的左边并消去  $PP^{-1}$  时它是平凡的.

**【4.6.12】引理** 令  $A, B$ , 与  $P$  是  $n \times n$  矩阵. 如果  $P$  是可逆的, 则  $(P^{-1}AP)(P^{-1}BP) = P^{-1}(AB)P$ , 并且对所有  $k \geq 1$ ,  $(P^{-1}AP)^k = P^{-1}A^kP$ .

因此, 如果  $A, P$  与  $\Lambda$  如同 (4.6.9) 中所述, 则

$$A^k = P\Lambda^kP^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 5 & \\ & 2 \end{bmatrix}^k \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 5^k + 2 \cdot 2^k & 2(5^k - 2^k) \\ 5^k - 2^k & 2 \cdot 5^k + 2^k \end{bmatrix}$$

如果  $f(t)=a_0+a_1t+\dots+a_nt^n$  是  $t$  的系数在  $F$  里的多项式, 并且若  $A$  为元素属于  $F$  的  $n \times n$  矩阵, 则  $f(A)$  表示形式上用  $A$  替换  $t$  所得的矩阵.

$$\text{【4.6.13】} \quad f(A) = a_0I + a_1A + \dots + a_nA^n$$

常数项  $a_0$  替换为  $a_0I$ . 于是, 如果  $A=P\Lambda P^{-1}$ , 则

$$\text{【4.6.14】} \quad f(A) = f(P\Lambda P^{-1}) = a_0I + a_1P\Lambda P^{-1} + \dots + a_nP\Lambda^n P^{-1} = Pf(\Lambda)P^{-1}$$

类似的记号应用于线性算子: 如果  $T$  是域  $F$  上向量空间  $V$  的线性算子, 则  $V$  上算子  $f(T)$  定义为

$$\text{【4.6.15】} \quad f(T) = a_0I + a_1T + \dots + a_nT^n$$

其中  $I$  表示恒等算子. 算子  $f(T)$  的作用向量定义为  $f(T)v = a_0v + a_1Tv + \dots + a_nT^nv$ . (为了避免太多的括号, 把  $T(v)$  写为  $Tv$  以略去一些括号.)

119

## 第七节 若尔当形

假设已知有限维复向量空间  $V$  上的线性算子  $T$ . 我们已经看到, 如果它的特征多项式的根是不同的, 则存在特征向量基, 并且  $T$  关于这个基的矩阵是对角的. 在这儿我们要问: 如果不假设特征值是不同的, 我们能做什么. 当特征多项式有重根时, 大多情形是没有特征向量基的, 但我们将看到, 不管怎样, 矩阵能够被相当地简单化.

线性算子  $T$  的伴随于特征值  $\lambda$  的特征向量是一个非零向量  $v$ , 满足  $(T-\lambda)v=0$ . (这里

我们将  $T - \lambda I$  写成  $T - \lambda$ .) 因为算子  $T$  不一定有足够的特征向量, 所以我们用广义特征向量分析讨论.

**注** 线性算子  $T$  的伴随于特征值  $\lambda$  的广义特征向量是一个非零向量  $x$ , 使得  $(T - \lambda)^k x = 0$  对某个  $k > 0$  成立. 它的幂指数是使得  $(T - \lambda)^d x = 0$  的最小整数  $d$ .

**【4.7.1】命题** 令  $x$  是  $T$  的伴随于特征值  $\lambda$  和幂指数  $d$  的广义特征向量, 并且对  $j \geq 0$ , 设  $u_j = (T - \lambda)^j x$ . 令  $\mathbf{B} = (u_0, \dots, u_{d-1})$ , 并设  $X = \text{Span} \mathbf{B}$ . 则  $X$  是  $T$ -不变子空间, 并且  $\mathbf{B}$  是  $X$  的一个基.

在证明中我们用到下面一个引理.

**【4.7.2】引理** 符号  $u_j$  同上, 线性组合  $y = c_j u_j + \dots + c_{d-1} u_{d-1}$  是伴随于特征值  $\lambda$  和幂指数  $d - j$  的广义特征向量, 其中  $j \leq d - 1$  且  $c_j \neq 0$ .

**证明** 因为  $x$  的幂指数是  $d$ , 故  $(T - \lambda)^{d-1} x = u_{d-1} \neq 0$ . 所以,  $(T - \lambda)^{d-j-1} y = c_j u_{d-1}$  不为 0, 但  $(T - \lambda)^{d-j} y = 0$ . 于是, 如所断言的,  $y$  是伴随于特征值  $\lambda$  和幂指数  $d - j$  的广义特征向量. ■

**命题的证明** 我们注意到

$$\mathbf{【4.7.3】} \quad Tu_j = \begin{cases} \lambda u_j + u_{j+1} & \text{若 } j < d - 1 \\ \lambda u_j & \text{若 } j = d - 1 \\ 0 & \text{若 } j > d - 1 \end{cases}$$

所以, 对所有  $j$ ,  $Tu_j$  在子空间  $X$  中. 这表明  $X$  是不变的. 其次由定义,  $\mathbf{B}$  生成  $X$ . 引理表明  $\mathbf{B}$  的每个非平凡线性组合是广义特征向量, 于是, 它不为 0. 所以,  $\mathbf{B}$  是无关集. ■

**【4.7.4】推论** 令  $x$  是  $T$  的伴随于特征值  $\lambda$  的广义特征向量, 则  $\lambda$  是一般的特征值—— $T$  的特征多项式的一个根.

**120** **证明** 如果  $x$  的幂指数是  $d$ , 符号同上, 则  $u_{d-1}$  伴随于特征值  $\lambda$  的特征向量. ■

公式(4.7.3)确定了描述  $T$  作用在命题 4.7.1 的基  $\mathbf{B}$  上的矩阵. 它是  $d \times d$  若尔当块  $J_\lambda$ . 对较小数值  $d$ , 若尔当块如下所示:

$$\mathbf{【4.7.5】} \quad J_\lambda = [\lambda], \quad \begin{bmatrix} \lambda & \\ & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & & \\ 1 & \lambda & \\ & & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & & & \\ 1 & \lambda & & \\ & & \lambda & \\ & & & \lambda \end{bmatrix}, \dots$$

当  $\lambda = 0$  时, 若尔当块的运算非常简单.  $d \times d$  块  $J_0$  在  $\mathbb{C}^d$  的标准基上的运算为

$$\mathbf{【4.7.6】} \quad e_1 \rightsquigarrow e_2 \rightsquigarrow \dots \rightsquigarrow e_d \rightsquigarrow 0$$

$1 \times 1$  若尔当块  $J_0$  是 0.

下面的若尔当分解定理断言: 任一个  $n \times n$  复矩阵都相似于由对角若尔当块(4.7.5)构成的矩阵  $J$ ——它有若尔当形

$$\text{【4.7.7】} \quad J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{bmatrix}$$

其中对某个  $\lambda_i$ ,  $J_i = J_{\lambda_i}$ . 块  $J_i$  有各种各样的行列数  $d_i$  且  $\sum d_i = n$ , 对角元素  $\lambda_i$  不一定不同. 矩阵  $J$  的特征多项式是

$$\text{【4.7.8】} \quad p(t) = (t - \lambda_1)^{d_1} (t - \lambda_2)^{d_2} \cdots (t - \lambda_t)^{d_t}$$

$2 \times 2$  与  $3 \times 3$  若尔当形是

【4.7.9】

$$\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}, \begin{bmatrix} \lambda_1 & \\ 1 & \lambda_1 \end{bmatrix}; \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix}, \begin{bmatrix} \lambda_1 & & \\ 1 & \lambda_1 & \\ & & \lambda_2 \end{bmatrix}, \begin{bmatrix} \lambda_1 & & \\ 1 & \lambda_1 & \\ & 1 & \lambda_1 \end{bmatrix}$$

其中诸标量  $\lambda_i$  可以相等, 也可不相等, 在第四个矩阵里, 可以用另一个顺序列出块.

【4.7.10】定理(若尔当分解)

(a) 向量空间形式: 令  $T$  是有限维复向量空间  $V$  上的线性算子, 则存在  $V$  的一个基  $B$ , 使得  $T$  关于  $B$  的矩阵有若尔当形(4.7.7).

(b) 矩阵形式: 令  $A$  是  $n \times n$  复矩阵, 则存在可逆复矩阵  $P$  使得  $P^{-1}AP$  有若尔当形. 算子  $T$  或矩阵  $A$  的若尔当形除若尔当块的顺序外是唯一的. 121

证明 这个证明是 Filippov[Filippov]给出的. 对  $V$  的维数应用归纳法允许我们假设本定理对  $T$  在任一个真子空间上的限制成立. 所以, 如果  $V$  是  $T$ -不变子空间的直和, 比如说,  $V = V_1 \oplus \cdots \oplus V_r$ , 其中  $r > 1$ , 则定理对  $T$  成立.

假设对  $i = 1, \dots, r$  有广义特征向量  $v_i$ . 令  $V_i$  是如命题 4.7.1 里所定义的子空间, 其中  $x = v_i$ . 如果  $V$  是直和  $V_1 \oplus \cdots \oplus V_r$ , 则定理将对  $V$  成立, 并且我们说  $v_1, \dots, v_r$  是  $T$  的若尔当生成元. 我们将证明若尔当生成元的集合存在.

第一步: 选取  $T$  的特征值  $\lambda$ , 用  $T - \lambda I$  替换算子  $T$ . 如果  $A$  是  $T$  关于一个基的矩阵, 则  $T - \lambda I$  关于同一个基的矩阵是  $A - \lambda I$ . 并且, 若矩阵  $A$  或  $A - \lambda I$  之一是若尔当形, 则另一个也是若尔当形. 所以, 用  $T - \lambda I$  替换算子  $T$  是可允许的. 经此步骤后, 我们的算子(仍叫做  $T$ )将有特征值 0. 这将简化记号.

第二步: 设 0 是  $T$  的特征值. 令  $K_i$  与  $U_i$  分别表示  $i$  次幂  $T$  的核与像, 则  $K_1 \subset K_2 \subset \cdots$ ,  $U_1 \supset U_2 \supset \cdots$ . 因  $V$  是有限维的, 故对足够大的  $m$ , 这些子空间的链会终止于常值, 比如  $K_m = K_{m+1} = \cdots$ ,  $U_m = U_{m+1} = \cdots$ . 令  $K = K_m$ ,  $U = U_m$ . 我们证明  $K$  与  $U$  是不变子空间, 并且  $V$  是直和  $K \oplus U$ .

因为  $TK_m \subset K_{m-1} \subset K_m$  与  $TU_m = U_{m+1} = U_m$ , 故这些子空间是不变的. 要证明  $V = K \oplus U$ , 只要证明  $K \cap U = \{0\}$  就够了(见命题 4.3.1(b)). 令  $z$  是  $K \cap U$  的元素, 那么  $T^m z = 0$ , 并且对  $V$  中的某个  $v$  还有  $z = T^m v$ . 所以,  $T^{2m} v = 0$ . 于是,  $v$  是  $K_{2m}$  中的元素. 但

$K_{2m} = K_m$ , 所以,  $T^m v = 0$ , 亦即,  $z = 0$ .

因为  $T$  有特征值 0, 故  $K$  不是零子空间. 所以,  $U$  的维数比  $V$  的小, 并且由归纳假设, 定理对  $T|_U$  成立. 不幸的是, 我们对  $K$  不能使用这样的推理, 因为  $U$  可能是零. 所以, 还需对  $T|_K$  证明存在若尔当形. 我们用  $K$  替换  $V$ , 用  $T|_K$  替换  $T$ .

注 向量空间  $V$  上的线性算子  $T$  称为幂零的, 如果对某个正整数  $r$ , 算子  $T^r$  为零. 我们把证明简化到幂零算子的情形.

第三步: 假设算子  $T$  是幂零的. 每个非零向量都是具有特征值 0 的广义特征向量. 令  $N$  与  $W$  分别表示  $T$  的核与像. 由于  $T$  是幂零的, 故  $N \neq \{0\}$ . 所以,  $W$  的维数比  $V$  的小, 并且由归纳法, 定理对算子在  $W$  上的限制是成立的. 于是, 存在  $T|_W$  的若尔当生成元  $w_1, \dots, w_r$ . 令  $e_i$  表示  $w_i$  的幂指数, 并且设  $W_i$  表示如同命题 4.7.1 中的利用广义特征向量  $w_i$  构成的子空间. 所以,  $W = W_1 \oplus \dots \oplus W_r$ .

对每个  $i$ , 选取  $V$  的元素  $v_i$  使得  $Tv_i = w_i$ .  $v_i$  的幂指数  $d_i$  将等于  $e_i + 1$ . 设  $V_i$  表示如同命题 4.7.1 中的利用向量  $v_i$  构成的子空间. 那么,  $TV_i = W_i$ . 令  $U$  表示和  $V_1 + \dots + V_r$ . 因为每个  $V_i$  都是不变子空间, 故  $U$  也是. 现在证明  $v_1, \dots, v_r$  是限制  $T|_U$  的若尔当生成元; 亦即, 诸子空间  $V_i$  是无关的.

122

注意到两点: 首先, 因为  $TV_i = W_i$ , 故  $TU = W$ . 其次  $V_i \cap N \subset W_i$ . 这可由引理 4.7.2 得到, 说明  $V_i \cap N$  是最后一个基向量  $T^{d_i-1} v_i$  的张成. 由于  $d_i - 1 = e_i$  且是正的, 故  $T^{d_i-1} v_i$  在像  $W_i$  里.

假设已给关系  $\tilde{v}_1 + \dots + \tilde{v}_r = 0$ , 其中  $\tilde{v}_i \in V_i$ . 需证对所有  $i$ , 有  $\tilde{v}_i = 0$ . 令  $\tilde{w}_i = T\tilde{v}_i$ . 则  $\tilde{w}_1 + \dots + \tilde{w}_r = 0$ , 且  $\tilde{w}_i \in W_i$ . 由于诸子空间  $W_i$  是无关的, 故  $\tilde{w} = 0$  对所有  $i$  成立. 于是  $T\tilde{v}_i = 0$ , 这意味着  $\tilde{v}_i \in V_i \cap N$ . 所以,  $\tilde{v}_i \in W_i$ . 再次利用子空间  $W_i$  是无关的事实, 我们得出结论  $\tilde{v}_i = 0$  对所有  $i$  成立.

第四步: 证明  $T$  的若尔当生成元集可由把  $N$  的某些元素加到  $T|_U$  的若尔当生成元集  $\{v_1, \dots, v_r\}$  得到.

令  $v$  是  $V$  的任意元素, 且设  $Tv = w$ . 由于  $TV = W$ , 故在  $U$  中存在向量  $u$  使得  $Tu = w = Tv$ . 这样  $z = v - u$  在  $N$  里, 且  $v = u + z$ . 所以,  $U + N = V$ . 于是, 我们把  $U$  的基通过添加  $N$  的元素(比如说,  $z_1, \dots, z_r$ )扩充成  $V$  的基(见命题 3.4.16(a)). 令  $N'$  是  $(z_1, \dots, z_r)$  的张成, 则  $U \cap N' = \{0\}$  且  $U + N' = V$ . 所以,  $V$  是直和  $U \oplus N'$ .

算子  $T$  在  $N'$  上是 0, 于是  $N'$  是不变子空间, 且  $T|_{N'}$  的矩阵是零矩阵, 其有若尔当形. 它的若尔当块是  $1 \times 1$  零矩阵. 所以,  $\{v_1, \dots, v_r; z_1, \dots, z_r\}$  是  $T$  的若尔当生成元. ■

倘若已给特征值, 不难确定算子  $T$  的若尔当形, 并且分析还证明了若尔当形的唯一性. 然而, 求  $V$  的一个适当的基是痛苦的, 最好是避免掉.

要确定若尔当形, 可选取一个特征值  $\lambda$ , 用  $T - \lambda I$  替换  $T$  以简化到  $\lambda = 0$  的情形. 令  $K_i$  是  $T^i$  的核, 且令  $k_i$  是  $K_i$  的维数. 在单个  $d \times d$  若尔当块且  $\lambda = 0$  的情形, 这些维数是:

$$k_i^* = \begin{cases} i & \text{若 } i \leq d \\ d & \text{若 } i \geq d \end{cases}$$

对一般算子  $T$ , 维数  $k_i$  通过对  $\lambda=0$  的每个块的数  $k_i^k$  相加得到. 所以,  $k_1$  是具有  $\lambda=0$  的块的个数,  $k_2-k_1$  是具有  $\lambda=0$  且  $d \geq 2$  的块的个数, 等等.

两个例子:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \quad \text{和} \quad B = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ -1 & 1 & -1 \end{bmatrix}$$

这里  $A^3=0$ , 但  $A^2 \neq 0$ . 如果  $v$  是一个向量, 使得  $A^2 v \neq 0$ , 例如  $v=e_1$ , 则  $(v, Tv, T^2 v)$  是一个基. 若尔当形由一个  $3 \times 3$  块组成.

另一方面,  $B^2=0$ . 再次取  $v=e_1$ , 集合  $(v, Tv)$  是无关系的, 并且这给出一个  $2 \times 2$  块. 要想得到若尔当形, 必须在  $N$  里添加一个向量, 例如  $v'=e_2+e_3$ , 这个向量将给出一个  $1 \times 1$  块(等于 0). 所需要的基是  $(v, Tv, v')$ .

123

将若尔当形写为  $J=D+N$  常常是有用的, 其中  $D$  是矩阵的对角部分, 而  $N$  是对角线下面的部分. 对于单个若尔当块, 有  $D=\lambda I$  与  $N=J_0$ , 如下面对  $3 \times 3$  块所展示的:

$$J_\lambda = \begin{bmatrix} \lambda & & \\ 1 & \lambda & \\ & 1 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda & & \\ 0 & \lambda & \\ & 0 & \lambda \end{bmatrix} + \begin{bmatrix} 0 & & \\ 1 & 0 & \\ & 1 & 0 \end{bmatrix} = \lambda I + J_0 = D + N$$

因为  $D$  与  $N$  是交换的, 故写  $J=D+N$  是方便的.  $J$  的幂可通过二项式展开计算:

$$[4.7.11] \quad J^r = (D+N)^r = D^r + \binom{r}{1} D^{r-1} N + \binom{r}{2} D^{r-2} N^2 + \dots$$

当  $J$  是  $n \times n$  矩阵时,  $N^n=0$ , 且这个展开式至多有  $n$  项. 在单个若尔当块情形下, 公式为

$$[4.7.12] \quad J^r = (\lambda I + J_0)^r = \lambda^r I + \binom{r}{1} \lambda^{r-1} J_0 + \binom{r}{2} \lambda^{r-2} J_0^2 + \dots$$

[4.7.13] 推论 令  $T$  是有限维复向量空间  $V$  上的线性算子, 下列条件是等价的:

- (a)  $T$  是可对角化算子.
- (b) 每个广义特征向量是特征向量.
- (c)  $T$  的若尔当形里的所有若尔当块都是  $1 \times 1$  块.

类似叙述对复方矩阵  $A$  也是对的.

证明 (a)  $\Rightarrow$  (b): 设  $T$  是可对角化的, 比如说,  $T$  关于基  $B=(v_1, \dots, v_n)$  的矩阵是对角矩阵  $\Lambda$ , 其对角元为  $\lambda_1, \dots, \lambda_n$ . 令  $v$  是  $V$  中广义特征向量, 比如说,  $(T-\lambda)^k v=0$  对某个  $\lambda$  与某个  $k > 0$  成立. 用  $T-\lambda$  替换  $T$  简化到情形  $T^k v=0$ . 令  $X=(x_1, \dots, x_n)^t$  是  $v$  的坐标向量.  $T^k v$  的坐标将是  $\lambda^k x_i$ , 因为  $T^k v=0$ , 故要么  $\lambda_i=0$ , 要么  $x_i=0$ , 不论哪种情形都有  $\lambda^k x_i=0$ . 所以,  $Tv=0$ .

(b)  $\Rightarrow$  (c): 我们对照地证明. 如果  $T$  的若尔当形有一个  $k \times k$  若尔当块且  $k > 1$ , 则回顾(4.7.6)  $J_1 - \lambda I$  的作用, 我们看到存在不是特征向量的广义特征向量. 所以, 若(c)是错的, 则(b)也是错的. 最后, 显然有(c)  $\Rightarrow$  (a). ■

这里有若尔当形一个很好的应用.

**【4.7.14】定理** 令  $T$  是有限维复向量空间  $V$  上的线性算子. 如果  $T$  的某个正幂是恒等的, 比如说,  $T^r = I$ , 则  $T$  是可对角化的.

**证明** 只要证明每个广义特征向量是特征向量就够了. 为此, 设  $(T - \lambda)^2 v = 0$  且  $v \neq 0$ , 我们证明  $(T - \lambda)v = 0$ . 因为  $\lambda$  是特征值且  $T^r = I$ , 故  $\lambda^r = 1$ . 用  $t - \lambda$  去除  $t^r - 1$ :

$$t^r - 1 = (t^{-1} + \lambda t^{-2} + \cdots + \lambda^{r-2} t + \lambda^{-1})(t - \lambda)$$

用  $T$  替换  $t$ , 且应用算子到  $v$ . 设  $w = (T - \lambda)v$ . 因为  $T^r - I = 0$ , 故

$$\begin{aligned} 0 &= (T^r - I)v = (T^{r-1} + \lambda T^{r-2} + \cdots + \lambda^{r-2} T + \lambda^{-1})(T - \lambda)v \\ &= (T^{r-1} + \lambda T^{r-2} + \cdots + \lambda^{r-2} T + \lambda^{-1})w \\ &= r\lambda^{-1}w. \end{aligned}$$

(对于最后一个等式, 用到了  $Tw = \lambda w$  的事实.) 因  $r\lambda^{-1}w = 0$ , 故  $w = 0$ . ■

我们回顾一下本节的结果. 哪里用到了  $V$  是复数上的向量空间的假设? 答案是它的应用处仅是保证特征多项式有足够的根.

**【4.7.15】推论** 令  $V$  是域  $F$  上的有限维向量空间, 且设  $T$  是  $V$  上的线性算子, 其特征多项式在  $F$  上分解线性因子的乘积, 则若尔当分解定理 4.7.10 对  $T$  是成立的.

**证明** 与  $F = \mathbb{C}$  情形的证明一样.

**【4.7.16】推论** 令  $T$  是特征为零的域上的有限维向量空间  $V$  上的线性算子. 假设  $T^r = I$  对某个  $r \geq 1$  成立, 并且多项式  $t^r - 1$  在  $F$  上分解为线性因子的乘积. 则  $T$  是可对角化的.

在定理 4.7.14 证明的最后一步需要特征为零的假设, 在此我们想从关系  $r\lambda^{-1}w = 0$  推出  $w = 0$ . 在不是零特征的情形下, 定理是不成立的.



## 练 习

### 第一节 维数公式

- 1.1 设  $A$  是  $\ell \times m$  矩阵并设  $B$  是  $n \times p$  矩阵. 证明法则  $M \mapsto AMB$  定义一个由  $m \times n$  矩阵空间  $F^{m \times n}$  到空间  $F^{\ell \times p}$  的线性变换.
- 1.2 设  $v_1, \dots, v_n$  是向量空间  $V$  的元素. 证明由  $\varphi(X) = v_1 x_1 + \cdots + v_n x_n$  定义的映射  $\varphi: F^n \rightarrow V$  是一个线性变换.
- 1.3 设  $A$  是  $m \times n$  矩阵. 用维数公式证明线性方程组  $AX = 0$  的解空间的维数至少是  $n - m$ .
- 1.4 证明每一个秩为 1 的  $m \times n$  矩阵具有  $A = XY'$  的形式, 其中  $X, Y$  为  $m$  维和  $n$  维列向量. 如何唯一确

⊖ 我收到许多询问这个画谜的电子邮件. Yvonne 是一位人类学家, 她和她的丈夫 Jean-Louis (一位数学家) 是亲密朋友, 悲惨地死于 1989 年. 为纪念他们, 我把他们包括在被引用人里. 圣瓦伦丁的历史是 Yvonne 的许多兴趣之一, 她发送这个画谜作为纪念.

定这些向量?

- 1.5 (a) 令  $U$  与  $W$  是域  $F$  上的向量空间. 证明关于向量对的两个运算  $(u, w) + (u', w') = (u + u', w + w')$  与  $c(u, w) = (cu, cw)$  使得  $U \times W$  构成一个向量空间. 该空间称为积空间.
- (b) 令  $U$  与  $W$  是向量空间  $V$  的子空间. 证明由  $T(u, w) = u + w$  定义的映射  $T: U \times W \rightarrow V$  是一个线性变换.
- (c) 用  $V$  的子空间的维数表示  $T$  的维数公式.

125

## 第二节 线性变换的矩阵

- 2.1 设  $A$  与  $B$  为  $2 \times 2$  矩阵. 确定  $2 \times 2$  矩阵空间  $F^{2 \times 2}$  上的线性算子  $T: W \rightsquigarrow AMB$  关于  $F^{2 \times 2}$  的基  $(e_{11}, e_{12}, e_{21}, e_{22})$  的矩阵.
- 2.2 设  $A$  是  $n \times n$  矩阵, 设  $V = F^n$  表示  $n$  维行向量空间. 线性算子“用  $A$  右乘”关于  $V$  的标准基的矩阵是什么?
- 2.3 求将直线  $y = x$  变到直线  $y = 3x$  的所有  $2 \times 2$  矩阵.
- 2.4 用行和列变换证明定理 4.2.10(b).
- 2.5 设  $A$  是秩为  $r$  的  $m \times n$  矩阵, 设  $I$  是  $r$  个行指标的集合使得与  $A$  对应的  $r$  个行是无关的, 设  $J$  是  $r$  个列指标的集合使得与  $A$  对应的  $r$  个列是无关的. 设  $M$  是由取自  $I$  的行与  $J$  的列得到的  $A$  的  $r \times r$  子矩阵. 证明  $M$  是可逆的.

## 第三节 线性算子

- 3.1 确定空间  $\mathbf{R}^n$  上的线性算子  $T$  的核与像的维数, 其中  $T$  的定义为

$$T(x_1, \dots, x_n)^t = (x_1 + x_n, x_2 + x_{n-1}, \dots, x_n + x_1)^t$$

- 3.2 (a) 设  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  是实  $2 \times 2$  矩阵, 其中  $c \neq 0$ . 证明利用初等矩阵的共轭可消去元素“ $a$ ”.
- (b)  $c = 0$  的哪个矩阵与其中元素“ $a$ ”为 0 的矩阵相似?
- 3.3 设  $T: V \rightarrow V$  是 2 维向量空间的线性算子. 设  $T$  不是标量乘法算子. 证明存在向量  $v \in V$  使得  $(v, T(v))$  是  $V$  的基, 并描述  $T$  关于这个基的矩阵.
- 3.4 令  $B$  是复  $n \times n$  矩阵. 证明或反证: 由  $T(A) = AB - BA$  定义的  $n \times n$  矩阵空间上的线性算子  $T$  是奇异的.

## 第四节 特征向量

- 4.1 设  $T$  是向量空间  $V$  的线性算子, 并设  $\lambda$  是标量, 设  $V^{(\lambda)}$  是  $T$  的伴随于特征值  $\lambda$  的特征向量加上 0 的集合. 证明  $V^{(\lambda)}$  是一个  $T$ -不变子空间.
- 4.2 (a) 设  $T$  是有限维向量空间  $V$  的线性算子, 满足  $T^2$  是恒等算子. 证明: 对  $V$  中任意向量  $v$ ,  $v - Tv$  要么是伴随于特征值  $-1$  的特征向量, 要么是零向量. 用习题 4.1 中的记号, 证明  $V$  是特征子空间  $V^{(1)}$  与  $V^{(-1)}$  的直和.
- (b) 推广这个方法, 证明使得  $T^n = I$  的线性算子  $T$  分解复向量空间为四个特征子空间的直和.
- 4.3 设  $T$  是向量空间  $V$  的线性算子. 证明: 如果  $W_1$  与  $W_2$  是  $V$  的  $T$ -不变子空间, 则  $W_1 + W_2$  与  $W_1 \cap W_2$  是  $T$ -不变的.
- 4.4  $2 \times 2$  矩阵  $A$  有伴随于特征值 2 的特征向量  $v_1 = (1, 1)^t$  和伴随于特征值 3 的特征向量  $v_2 = (1, 2)^t$ .

126



确定  $A$ .

4.5 求矩阵为 (a)  $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$  和 (b)  $\begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}$  的实线性算子的所有不变子空间.

4.6 设  $P$  是次数  $\leq n$  的多项式  $p(x) = a_0 + a_1x + \cdots + a_nx^n$  的实向量空间, 设  $D$  表示导数  $\frac{d}{dx}$ , 把它视为  $P$  上的线性算子.

(a) 证明  $D$  是幂零算子, 即  $D^k = 0$  对充分大的  $k$  成立.

(b) 求  $D$  关于一个方便的基的矩阵.

(c) 确定  $P$  的所有  $D$ -不变子空间.

4.7 令  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  是任意一个实  $2 \times 2$  矩阵. 列向量  $X$  是用  $A$  左乘的特征向量的条件是  $AX = Y$  是  $X$  的

标量倍数, 这表示斜率  $s = \frac{y_2}{x_1}$  与  $s' = \frac{y_2}{y_1}$  相等.

(a) 求  $s$  中表示这一等式的方程.

(b) 如果  $A$  的元素是正实数, 证明在第一象限有一个特征向量, 在第二象限也有一个特征向量.

4.8 令  $T$  是有限维向量空间的线性算子, 且每个非零向量都是其特征向量. 证明  $T$  是标量乘法算子.

### 第五节 特征多项式

5.1 对下列复矩阵求特征多项式、特征值和特征向量.

(a)  $\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$       (b)  $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$       (c)  $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

5.2 下面矩阵的特征多项式是  $t^3 - 4t - 1$ . 确定缺失的元素.

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 1 & * & * \end{bmatrix}$$

5.3 哪些复数是使得

(a)  $T^* = I$     (b)  $T^2 - 5T + 6I = 0$

的线性算子  $T$  的特征值?

127

5.4 求  $k \times k$  矩阵

$$\begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & & & \ddots & \\ & 1 & & & \ddots \\ & & & \ddots & & 1 \\ & & & & & & 1 & 0 \end{bmatrix}$$

的特征多项式的递归关系并计算  $k \leq 5$  的特征多项式.

5.5 哪个实  $2 \times 2$  矩阵有实特征值? 证明特征值是实的, 如果对角元以外的元素有相同的符号.

5.6 设  $V$  是具有基  $(v_0, \dots, v_n)$  的向量空间, 且设  $a_0, \dots, a_n$  为标量.  $V$  上由规则  $T(v_i) = v_{i+1}$  ( $i < n$ )

和  $T(v_n) = a_0 v_0 + a_1 v_1 + \cdots + a_n v_n$  定义一个线性算子  $T$ . 确定  $T$  关于所给定的基的矩阵与  $T$  的特征多项式.

5.7  $A$  与  $A'$  有相同的特征值吗? 有相同的特征向量吗?

5.8 设  $A = (a_{ij})$  是  $3 \times 3$  矩阵. 证明特征多项式里  $t$  的系数是  $2 \times 2$  对称子行列式的和:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \det \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} + \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

5.9 考虑在所有  $m \times m$  矩阵的空间  $F^{m \times m}$  上的  $m \times m$  矩阵  $A$  的左乘定义的线性算子. 确定这个算子的迹和行列式.

5.10 令  $A$  和  $B$  是  $n \times n$  矩阵, 确定空间  $F^{n \times n}$  上由  $M \rightsquigarrow AMB$  所定义的算子的迹和行列式.

## 第六节 三角形与对角形

6.1 令  $A$  是  $n \times n$  矩阵, 其特征多项式分解成线性因子乘积:  $p(t) = (t - \lambda_1) \cdots (t - \lambda_n)$ . 证明  $\text{trace} A = \lambda_1 + \cdots + \lambda_n$ ,  $\det A = \lambda_1 \cdots \lambda_n$ .

6.2 设复  $n \times n$  矩阵  $A$  有不同特征值  $\lambda_1, \dots, \lambda_n$ , 令  $v_1, \dots, v_n$  是伴随于这些特征值的特征向量.

(a) 证明每个特征向量都是诸向量  $v_i$  之一的倍数.

(b) 说明如何从特征值与特征向量恢复矩阵.

6.3 设  $T$  是有两个具有同一特征值  $\lambda$  的线性无关的特征向量的线性算子. 证明  $\lambda$  是  $T$  的特征多项式的重根.

6.4 令  $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ , 求矩阵  $P$  使得  $P^{-1}AP$  是对角的, 并求矩阵  $A^{30}$  的公式.

6.5 在每一情形里, 求复矩阵  $P$  使得  $P^{-1}AP$  是对角的.

$$(a) \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \quad (b) \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (c) \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

6.6 设  $A$  是可对角化的. 在特殊线性群里能够用矩阵  $P$  进行对角化吗?

6.7 证明: 如果  $A$  与  $B$  是  $n \times n$  矩阵且如果  $A$  是非奇异的, 则  $AB$  与  $BA$  相似.

6.8 线性算子  $T$  是幂零的, 如果某个正数幂  $T^k$  是零. 证明线性算子  $T$  是幂零的当且仅当存在  $V$  的基使得  $T$  的矩阵为上三角形的, 且对角元素都为零.

6.9 求使得  $A^2 = I$  的所有实  $2 \times 2$  矩阵, 并用几何方法描述它们的左乘在  $\mathbf{R}^2$  上的作用.

6.10 设  $M$  是由两个对角块组成的矩阵:  $M = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$ . 证明  $M$  是可对角化的当且仅当  $A$  和  $D$  都是可对角化的.

6.11 设  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  是有特征值  $\lambda$  的  $2 \times 2$  矩阵.

(a) 证明  $(b, \lambda - a)^t$  是一个特征向量, 除非它为 0.

(b) 假设  $b \neq 0$  并且  $A$  有不同的特征值, 求矩阵  $P$  使得  $P^{-1}AP$  是对角的.

## 第七节 若尔当形

7.1 确定矩阵的若尔当形:  $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ .

7.2 证明  $A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$  是幂等矩阵, 亦即,  $A^2 = A$ , 求它的若尔当形.

7.3 令  $V$  是维数为 5 的复向量空间, 且  $T$  是  $V$  上的线性算子, 其特征多项式为  $(t-\lambda)^5$ . 假设算子  $T-\lambda I$  的秩为 2.  $T$  可能的若尔当形是什么?

7.4 (a) 确定矩阵的所有可能的若尔当形, 该矩阵的特征多项式为  $(t+2)^2(t-5)^3$ .

(b) 当伴随于特征值 2 的特征向量的空间是 1 维的, 而伴随于特征值 5 的特征向量的空间是 2 维的时, 其特征多项式为  $(t+2)^2(t-5)^3$  的矩阵的所有可能的若尔当形是什么?

7.5 所有特征向量都是单个向量的倍数的矩阵  $A$  的若尔当形是什么?

7.6 确定其若尔当形由一个块构成的线性算子的所有不变子空间.

7.7 使得  $A^2 = A$  的复方阵是可对角化的吗?

7.8 每个复方阵  $A$  都相似于它的转置吗?

7.9 求一个  $2 \times 2$  矩阵: 元素属于  $F_p$ , 有一个幂等于恒等元, 且在  $F_p$  中有一个特征值, 但不是可对角化的.

129

### 杂题

M.1 令  $v = (a_1, \dots, a_n)$  是一个实行向量. 我们可构造  $n! \times n$  矩阵  $M$ , 其行由  $v$  的元素的所有可能置换得到. 行可以按任意序排列. 因此, 如果  $n=3$ , 则  $M$  可以是

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ a_1 & a_3 & a_2 \\ a_2 & a_3 & a_1 \\ a_2 & a_1 & a_3 \\ a_3 & a_1 & a_2 \\ a_3 & a_2 & a_1 \end{bmatrix}$$

确定这样矩阵所有可能的秩.

M.2 令  $A$  是  $n \times n$  复矩阵, 具有  $n$  个不同特征值  $\lambda_1, \dots, \lambda_n$ . 假设  $\lambda_1$  是最大特征值, 亦即,  $|\lambda_1| > |\lambda_i|$  对所有  $i > 1$  成立.

(a) 证明对于大多数向量  $X$ , 序列  $X_k = \lambda_1^{-k} A^k X$  收敛于伴随于特征值  $\lambda_1$  的一个特征向量  $Y$ , 并精确描述  $X$  使其成立的条件.

(b) 不假设  $\lambda_1, \dots, \lambda_n$  是不同的, 证明同一结论.

M.3 用练习 M.2 的方法计算矩阵  $\begin{bmatrix} 3 & 1 \\ 3 & 4 \end{bmatrix}$  的最大特征值, 精确到小数点后 3 位.

M.4 如果  $X = (x_1, x_2, \dots)$  是无限实行向量, 并且  $A = (a_{ij})$ ,  $0 < i, j < \infty$  是无限实矩阵, 可能也许不能定义矩阵乘积  $XA$ . 对每一矩阵  $A$ , 在所有无限行向量的空间  $\mathbf{R}^\infty$  (3.7.1) 上能够定义右乘法? 在空间  $Z$  (3.7.2) 上呢?

M.5 设  $\varphi: F^n \rightarrow F^m$  是用  $m \times n$  矩阵  $A$  左乘.

(a) 证明下列叙述等价:

- $A$  有右逆, 即存在一个矩阵  $B$  使得  $AB = I$ .
- $\varphi$  是满射.

- $A$  的秩是  $m$ .

(b) 证明下列叙述等价:

- $A$  有左逆, 即存在一个矩阵  $B$  使得  $BA = I$ .
- $\varphi$  是单射.
- $A$  的秩是  $n$ .

130

M. 6 不用特征多项式, 证明  $n$  维向量空间的一个线性算子至多有  $n$  个不同的特征值.

M. 7 (算子幂) 令  $T$  是向量空间  $V$  上的线性算子. 设  $K_r$  与  $W_r$  分别表示  $T^r$  的核与像.

(a) 证明  $K_1 \subset K_2 \subset \dots$  与  $W_1 \supset W_2 \supset \dots$ .

(b) 下列条件对  $r$  的特殊值可能成立也可能不成立.

$$(1) K_r = K_{r+1} \quad (2) W_r = W_{r+1} \quad (3) W_r \cap K_1 = \{0\} \quad (4) W_1 + K_r = V$$

当  $V$  是有限维时, 求条件(1)~(4)的所有蕴含关系.

(c) 当  $V$  是无限维时, 做同样的事情.

M. 8 令  $T$  是有限维复向量空间  $V$  上的线性算子.

(a) 令  $\lambda$  是  $T$  的特征值, 并且设  $V_\lambda$  是广义特征向量与零向量的集合. 证明  $V_\lambda$  是  $T$ -不变子空间.

(这个子空间称为广义特征子空间.)

(b) 证明  $V$  是它的广义特征子空间的直和.

M. 9 令  $V$  是有限维向量空间. 线性算子  $T: V \rightarrow V$  称为射影: 如果  $T^2 = T$  (不一定是“正交射影”). 设  $K$  与  $W$  表示线性算子  $T$  的核与像. 证明:

(a)  $T$  是  $W$  上的射影当且仅当  $T$  在  $W$  上的限制是恒等映射.

(b) 如果  $T$  是一个射影, 则  $V$  是直和  $K \oplus W$ .

(c) 射影  $T$  的迹等于它的秩.

M. 10 令  $A$  与  $B$  表示是  $m \times n$  与  $n \times m$  实矩阵.

(a) 证明: 如果  $\lambda$  是  $m \times m$  矩阵  $AB$  的非零特征值, 则它也是  $n \times n$  矩阵  $BA$  的特征值. 用例子说明如果  $\lambda = 0$ , 这不一定成立.

(b) 证明  $I_m - AB$  可逆当且仅当  $I_n - AB$  可逆.

131

## 第五章 线性算子的应用

好的记号能免除大脑做无用功，  
从而使它集中精力解决  
更高深的问题。

— Alfred North Whitehead

### 第一节 正交矩阵与旋转

在这一节，标量域是实数域。

我们假设熟悉  $\mathbf{R}^2$  中向量的点积。  $\mathbf{R}^2$  中列向量  $X = (x_1, \dots, x_n)'$ ,  $Y = (y_1, \dots, y_n)'$  的点积定义为

$$\text{【5.1.1】} \quad (X \cdot Y) = x_1 y_1 + \dots + x_n y_n$$

为了方便，将点积写为行向量与列向量的矩阵积：

$$\text{【5.1.2】} \quad (X \cdot Y) = X'Y$$

对于  $\mathbf{R}^2$  中的向量，有公式

$$\text{【5.1.3】} \quad (X \cdot Y) = |X| |Y| \cos \theta$$

其中  $\theta$  是向量之间的夹角。这个公式可由余弦定理

$$\text{【5.1.4】} \quad c^2 = a^2 + b^2 - 2ab \cos \theta$$

得到，其中  $a, b, c$  是三角形的三边长， $\theta$  是边  $a$  与边  $b$  的夹角。为了推导(5.1.3)，在以点  $O$ ,  $X, Y$  为顶点的三角形上应用余弦定理。它的边长是  $|X|$ ,  $|Y|$  和  $|X - Y|$ ，于是余弦定理可写为

$$((X - Y) \cdot (X - Y)) = (X \cdot X) + (Y \cdot Y) - 2|X| |Y| \cos \theta$$

左边展开为  $(X \cdot X) - 2(X \cdot Y) + (Y \cdot Y)$ ，将此与右边比较就得到公式(5.1.3)。这个公式对于  $\mathbf{R}^n$  中的向量也成立，但需要理解夹角的含义，我们现在不花时间讨论它(见(8.5.2))。

$\mathbf{R}^2$  与  $\mathbf{R}^3$  中向量的最重要的知识点为：

- 向量  $X$  的长度的平方  $|X|^2$  是  $(X, X) = X'X$ ,
- 向量  $X$  与另一个向量  $Y$  正交，写为  $X \perp Y$ ，当且仅当  $X'Y = 0$ 。

我们用这些作为  $\mathbf{R}^n$  中向量的长度  $|X|$  与向量的正交的定义。注意长度  $|X|$  是正的，除非  $X$  是零向量，因为  $|X|^2 = X'X = x_1^2 + \dots + x_n^2$  是平方和。

**【5.1.5】定理(毕达哥拉斯定理)** 如果  $X \perp Y$  且  $Z = X + Y$ ，则  $|Z|^2 = |X|^2 + |Y|^2$ 。

展开  $Z'Z$  可证得该定理。如果  $X \perp Y$ ，则  $X'Y = Y'X = 0$ ，于是，

$$Z'Z = (X + Y)'(X + Y) = X'X + X'Y + Y'X + Y'Y = X'X + Y'Y.$$

我们转换到小写的向量记号. 如果  $v_1, \dots, v_k$  是  $\mathbf{R}^n$  中正交向量, 并且  $w = v_1 + \dots + v_k$ , 则由归纳法, 用毕达哥拉斯定理可证明

$$\text{【5.1.6】} \quad |w|^2 = |v_1|^2 + \dots + |v_k|^2$$

**【5.1.7】引理**  $\mathbf{R}^n$  中正交非零向量的任意集合  $(v_1, \dots, v_k)$  是无关的.

**证明** 令  $w = c_1 v_1 + \dots + c_k v_k$  是一个线性组合, 其中诸  $c_i$  不全为 0, 设  $w_i = c_i v_i$ . 则  $w$  是正交向量的和  $w_1 + \dots + w_k$ , 它们不全为 0. 由毕达哥拉斯定理,  $|w|^2 = |w_1|^2 + \dots + |w_k|^2 > 0$ , 故  $w \neq 0$ . ■

**注**  $\mathbf{R}^n$  中标准正交基  $B = (v_1, \dots, v_n)$  是正交单位向量(长度为 1 的向量)的基. 另一叙述方式为,  $B$  是标准正交基, 如果

$$\text{【5.1.8】} \quad (v_i \cdot v_j) = \delta_{ij}$$

其中  $\delta_{ij}$  (Kronecker 符号) 是恒等矩阵的第  $i$  行第  $j$  列元素, 若  $i=j$ , 其值为 1; 若  $i \neq j$ , 其值为 0.

**【5.1.9】定义** 实  $n \times n$  矩阵  $A$  是正交的, 如果  $A'A = I$ , 也就是说,  $A$  是可逆的, 它的逆是  $A'$ .

**【5.1.10】引理**  $n \times n$  矩阵  $A$  是正交的当且仅当它的列构成  $\mathbf{R}^n$  的标准正交基.

**证明** 令  $A_i$  表示  $A$  的第  $i$  列, 那么,  $A'_i$  表示  $A'$  的第  $i$  行.  $A'A$  的第  $i$  行第  $j$  列元素是  $A'_i A_j$ , 于是,  $A'A = I$  当且仅当  $A'_i A_j = \delta_{ij}$  对所有  $i$  与  $j$  成立. ■

容易证明正交矩阵的下一性质:

133

**【5.1.11】命题**

(a) 正交矩阵的积是正交的, 并且正交矩阵的逆(即它的转置)是正交的. 正交矩阵构成  $GL_n$  的子群  $O_n$ , 即正交群.

(b) 正交矩阵的行列式是  $\pm 1$ . 具有行列式是 1 的正交矩阵构成  $O_n$  的指标为 2 的子群  $SO_n$ , 即特殊正交群.

**【5.1.12】定义**  $\mathbf{R}^n$  上的正交算子  $T$  是保持点积的线性算子: 对每对向量  $X, Y$ ,

$$(TX \cdot TY) = (X \cdot Y)$$

**【5.1.13】命题**  $\mathbf{R}^n$  上的线性算子  $T$  是正交的当且仅当它保持向量的长度, 或者, 当且仅当对每个向量  $X$ ,  $(TX \cdot TX) = (X \cdot X)$ .

**证明** 假设保持长度, 令  $X$  与  $Y$  是  $\mathbf{R}^n$  里的任意向量, 则

$$(T(X+Y) \cdot T(X+Y)) = ((X+Y) \cdot (X+Y))$$

$(TX \cdot TY) = (X \cdot Y)$  的事实由展开等式的两边并消去等量得到. ■

**【5.1.14】命题**  $\mathbf{R}^n$  上的线性算子  $T$  是正交的当且仅当它的关于标准基的矩阵  $A$  是正交矩阵.

**证明** 如果  $A$  是  $T$  的矩阵, 则

$$(TX \cdot TY) = (AX)'(AY) = X'(A'A)Y$$

算子是正交的当且仅当右边等于  $X'Y$  对于所有  $X$  与  $Y$  成立. 我们把这个条件写为  $X'(A'A - I)Y = 0$ . 下一引理证明这是成立的当且仅当  $A'A - I = 0$ . 所以,  $A$  是正交的. ■

**【5.1.15】引理** 令  $M$  是  $n \times n$  矩阵. 如果  $X'MY = 0$  对所有列向量  $X$  与  $Y$  成立, 则  $M = 0$ .

**证明** 积  $e_i^t M e_j$  的值是  $M$  的第  $i$  行第  $j$  列元素. 例如,

$$\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = m_{21}$$

如果  $e_i^t M e_j = 0$  对所有列向量  $i$  与  $j$  成立, 则  $M=0$ . ■

我们现在描述正交  $2 \times 2$  矩阵.

**注**  $\mathbf{R}^2$  上的线性算子  $T$  是反射的, 如果它有分别伴随于特征值  $1$  与  $-1$  的正交向量  $v_1$  与  $v_2$ .

134

因为它固定  $v_1$ , 并且改变正交向量  $v_2$  的符号, 使得算子关于由  $v_1$  张成的  $1$  维子空间的反射. 关于  $e_1$  轴的反射由矩阵

**【5. 1. 16】** 
$$S_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

给出.

**【5. 1. 17】定理**

(a) 具有行列式值为  $1$  的正交  $2 \times 2$  矩阵是矩阵

**【5. 1. 18】** 
$$R = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

其中对某个角  $\theta$ ,  $c = \cos\theta$ ,  $s = \sin\theta$ . 矩阵  $R$  表示平面  $\mathbf{R}^2$  上关于原点转过角  $\theta$  的逆时针旋转.

(b) 具有行列式值为  $-1$  的正交  $2 \times 2$  矩阵是矩阵

**【5. 1. 19】** 
$$S = \begin{bmatrix} c & s \\ s & -c \end{bmatrix} = RS_0$$

其中  $c$  与  $s$  同上. 矩阵  $S$  关于  $\mathbf{R}^2$  的与  $e_1$  轴成  $\frac{1}{2}\theta$  夹角的  $1$  维子空间反射平面.

**证明** 比如说

$$A = \begin{bmatrix} c & * \\ s & * \end{bmatrix}$$

是正交的. 那么, 它的列是单位向量(5. 1. 10). 于是, 点  $(c, s)^t$  位于单位圆上, 且对某个角  $\theta$ ,  $c = \cos\theta$ ,  $s = \sin\theta$ . 检查积  $P = R^t A$ , 其中  $R$  是矩阵(5. 1. 18):

**【5. 1. 20】** 
$$P = R^t A = \begin{bmatrix} 1 & * \\ 0 & * \end{bmatrix}$$

由于  $R^t$  与  $A$  是正交矩阵, 所以  $P$  也是正交矩阵. 由引理 5. 1. 10 知, 第二列是与第一列正交的单位向量. 于是,

**【5. 1. 21】** 
$$P = \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$$

回过头来,  $A=RP$ , 如果  $\det A=1$ ,  $A=P$ ; 如果  $\det A=-1$ ,  $A=S=RS_0$ .

我们已经看到  $R$  表示旋转(4.2.2), 但需恒等化由矩阵  $S$  定义的算子.  $S$  的特征多项式是  $t^2-1$ , 所以, 它的特征值是 1 与  $-1$ . 令  $X_1$  与  $X_2$  是伴随于这些特征值的单位长度的特征向量. 因为  $S$  是正交的, 故

$$(X_1 \cdot X_2) = (SX_1 \cdot SX_2) = (X_1 \cdot -X_2) = -(X_1 \cdot X_2)$$

135

于是,  $(X_1 \cdot X_2)=0$ . 特征向量是正交的.  $X_1$  的张成是反射直线. 为确定这条直线, 把单位向量  $X$  写为  $(c', s')^t$ , 其中  $c' = \cos \alpha$ ,  $s' = \sin \alpha$ . 这样,

$$SX = \begin{bmatrix} cc' + ss' \\ -sc' - cs' \end{bmatrix} = \begin{bmatrix} \cos(\theta - \alpha) \\ \sin(\theta - \alpha) \end{bmatrix}$$

当  $\alpha = \frac{1}{2}\theta$  时,  $X$  是伴随于特征值 1 的特征向量, 它是一个固定的向量. ■

下面我们描述  $3 \times 3$  旋转矩阵.

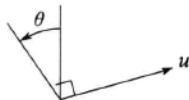
**[5.1.22] 定义**  $\mathbf{R}^3$  的关于原点的旋转是一个具有如下性质的线性算子  $\rho$ :

- $\rho$  固定单位向量  $u$ , 称为  $\rho$  的极点.
- $\rho$  旋转与  $u$  正交的 2 维子空间  $W$ .

旋转轴是由  $u$  张成的直线. 我们也称恒等算子为一个旋转, 尽管它的轴是不确定的.

如果由一个  $3 \times 3$  矩阵  $R$  所作的乘法(变换)是  $\mathbf{R}^3$  的一个旋转, 就称  $R$  是一个旋转矩阵.

**[5.1.23] 图**



$\mathbf{R}^3$  的一个旋转

旋转角的符号取决于子空间  $W$  是如何规定方向的. 我们从箭头  $u$  的头看子空间  $W$  来确定其方向. 图中所示的角  $\theta$  是正的. (这是“右手法则”.)

当  $u$  是向量  $e_1$  时, 集合  $(e_2, e_3)$  是  $W$  的一个基, 并且  $\rho$  的矩阵具有形式

$$\mathbf{[5.1.24]} \quad M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}$$

其中右下的  $2 \times 2$  子式是旋转矩阵(5.1.18).

**注意** 非恒等的旋转用偶对  $(u, \theta)$  描述, 称为自旋, 它由极点  $u$  和非零旋转角  $\theta$  组成.

具有自旋  $(u, \theta)$  的旋转记为  $\rho_{(u, \theta)}$ . 不同于恒等的每个旋转有两个极点, 即旋转轴  $\ell$  与  $\mathbf{R}^3$  里单位球的交. 这些是  $\rho$  的伴随于特征值 1 的单位长特征向量. 极点  $u$  的选取定义了  $\ell$  上的方向, 方向的变化引起旋转角符号的变化. 如果  $(u, \theta)$  是  $\rho$  的自旋, 则  $(-u, -\theta)$  也是. 因此, 每个旋转有两个自旋, 并且  $\rho_{(u, \theta)} = \rho_{(-u, -\theta)}$ .

136

**[5.1.25] 定理(欧拉定理)**  $3 \times 3$  旋转矩阵是行列式为 1 的正交  $3 \times 3$  矩阵, 它是特殊正交



群  $SO_3$  中的元素.

欧拉定理有个著名的推论, 可由  $SO_3$  是群的事实得到. 在代数上或者几何上, 它不是显然的.

**【5. 1. 26】推论** 绕任两个轴的旋转的合成是绕某个其他轴的旋转.

因为它们的元素表示旋转, 故群  $SO_2$  与  $SO_3$  称为 2 维旋转群与 3 维旋转群. 维数大于 3 时, 情况变得很复杂.

$$\text{【5. 1. 27】} \quad \begin{bmatrix} \cos\alpha & -\sin\alpha & & \\ \sin\alpha & \cos\alpha & & \\ & & \cos\beta & -\sin\beta \\ & & \sin\beta & \cos\beta \end{bmatrix}$$

是  $SO_4$  的一个元素. 由这个矩阵的左乘旋转由  $(e_1, e_2)$  张成的 2 维子空间, 其旋转的角是  $\alpha$ ; 它旋转由  $(e_3, e_4)$  张成的子空间, 其旋转的角是  $\beta$ .

在证明欧拉定理之前, 我们注意另外两个结果:

**【5. 1. 28】推论** 令  $M$  是  $SO_3$  里表示具有自旋  $(u, \theta)$  的旋转  $\rho_{(u, \theta)}$  的矩阵.

(a)  $M$  的迹是  $1+2\cos\alpha$ .

(b) 令  $B$  是  $SO_3$  的另一个元素, 且令  $u' = Bu$ . 共轭  $M' = BMB^{-1}$  表示带有自旋  $(u', \alpha)$  的旋转  $\rho(u', \alpha)$ .

**证明**

(a) 选择  $\mathbf{R}^3$  中的正交基  $(v_1, v_2, v_3)$  使得  $v_1 = u$ .  $\rho$  关于这个新基的矩阵有形式 (5. 1. 24), 它的迹是  $1+2\cos\alpha$ . 由于迹不依赖于基, 故  $M$  的迹也是  $1+2\cos\alpha$ .

(b) 因为  $SO_3$  是群, 故  $M'$  是  $SO_3$  的元素. 由欧拉定理知,  $M'$  是旋转矩阵. 而且  $u'$  是这个旋转的极点: 因  $B$  是正交的, 故  $u' = Bu$  长度为 1, 并且

$$M'u' = BMB^{-1}u' = BMu = Bu = u'.$$

令  $\alpha'$  是  $M'$  绕极点  $u'$  的旋转角.  $M$  的迹与它的共轭  $M'$  的迹是相等的, 于是,  $\cos\alpha = \cos\alpha'$ . 这蕴含着  $\alpha' = \pm\alpha$ . 由欧拉定理知, 矩阵  $B$  也表示旋转, 比如说, 绕某个极点旋转角  $\beta$ . 因为  $B$  与  $M'$  连续依赖  $\beta$ , 所以对  $\alpha'$  仅两个值  $\pm\alpha$  能出现. 当  $\beta=0$  时,  $B=I$ ,  $M'=M$ ,  $\alpha'=\alpha$ . 所以, 对所有  $\beta$ ,  $\alpha'=\alpha$ . ■

**【5. 1. 29】引理** 行列式为 1 的  $3 \times 3$  正交矩阵有特征值为 1.

**证明** 为证 1 是特征值, 我们证明矩阵  $M-I$  的行列式是 0. 如果  $B$  是  $n \times n$  矩阵, 则  $\det(-B) = (-1)^n \det B$ . 因处理  $3 \times 3$  矩阵, 故  $\det(M-I) = -\det(I-M)$ . 而且,  $\det(M-I)' = \det(M-I)$ ,  $\det M = 1$ . 这样,

$$\det(M-I) = \det(M-I)' = \det M \det(M-I)' = \det(M(M' - I)) \det(I-M)$$

关系  $\det(M-I) = \det(I-M)$  表明  $\det(M-I) = 0$ . ■

**欧拉定理的证明** 假设  $M$  表示具有自旋  $(u, \alpha)$  的旋转  $\rho$ . 由  $u$  添加到与  $u$  正交的子空间  $W$  的标准正交基构成  $V$  的标准正交基  $B$ .  $\rho$  关于这个基的矩阵  $M'$  有形式 (5. 1. 24), 它是正交的且其行列式为 1. 而且,  $M = PM'P^{-1}$ , 其中矩阵  $P$  等于  $[B]$  (3. 5. 13). 因为它的列

是正交的, 故 $[B]$ 是正交的. 所以,  $M$ 也是正交的, 且它的行列式等于1.

反过来, 令 $M$ 是行列式为1的正交矩阵, 设 $T$ 表示 $M$ 的左乘. 令 $u$ 为伴随于特征值为1的单位长度的特征向量, 并且 $W$ 是与 $u$ 正交的2维空间. 因为 $T$ 是固定 $u$ 的正交算子, 所以它把 $W$ 映为自身. 于是,  $W$ 是 $T$ -不变子空间, 并且我们可 $T$ 把限制到 $W$ .

因为 $T$ 是正交的, 故保持长度(5.1.13). 于是, 它到 $W$ 的限制也是正交的. 现在,  $W$ 有维数2, 且我们知道正交算子分成两类: 旋转和反射(5.1.17). 反射是其行列式为-1的算子. 如果算子 $T$ 作为反射作用在 $W$ 上, 且固定正交向量 $u$ , 则它的行列式也是-1. 由于情况不是这样, 故 $T|_W$ 是旋转. 这就证明了定义5.1.22的第二个条件, 从而 $T$ 是旋转. ■

## 第二节 连续性的使用

利用基于我们这里解释的连续性的推理, 关于复矩阵各式各样的事实都可通过对角化化简.

$n \times n$ 矩阵序列 $A_k$ 收敛于 $n \times n$ 矩阵 $A$ , 如果对每个 $i$ 与 $j$ ,  $A_k$ 的 $(i, j)$ -<sup>⊖</sup>元素收敛于 $A$ 的 $(i, j)$ -元素. 类似地, 具有复系数的次数为 $n$ 的多项式序列 $p_k(t)$  ( $k=1, 2, \dots$ )收敛于次数为 $n$ 的多项式 $p(t)$ , 如果对每个 $j$ ,  $p_k(t)$ 中 $t^j$ 的系数收敛于 $p$ 的对应系数. 我们指出, 用 $S_k \rightarrow S$ 表示的复数、矩阵或多项式的序列 $S_k$ 收敛于 $S$ .

**【5.2.1】命题(根的连续性)** 令 $p_k(t)$ 是次数 $\leq n$ 的首项系数为1的多项式序列, 设 $p(t)$ 是次数为 $n$ 的首项系数为1的另一个多项式. 令 $\alpha_{k,1}, \dots, \alpha_{k,n}$ 与 $\alpha_1, \dots, \alpha_n$ 表示这些多项式的根.

138

(a) 如果 $\alpha_{k,\nu} \rightarrow \alpha_\nu$ 对 $\nu=1, \dots, n$ 成立, 则 $p_k \rightarrow p$ .

(b) 反之, 如果 $p_k \rightarrow p$ , 则 $p_k$ 的根 $\alpha_{k,\nu}$ 可排序使得 $\alpha_{k,\nu} \rightarrow \alpha_\nu$ 对 $\nu=1, \dots, n$ 成立.

在(b)部分, 每个多项式 $p_k$ 的诸根必须单独重新排序.

**证明** 注意 $p_k(t) = (t - \alpha_{k,1}) \cdots (t - \alpha_{k,n})$ 与 $p(t) = (t - \alpha_1) \cdots (t - \alpha_n)$ . (a)部分由这样的事实得到:  $p(t)$ 的系数是根的多项式函数, 从而是连续函数. 但(b)部分不那么显然.

第一步: 令 $\alpha_{k,1}$ 是 $p_k$ 的最接近 $\alpha_1$ 的根, 亦即, 使得 $|\alpha_{k,1} - \alpha_1|$ 是极小的根. 我们重新排序 $p_k$ 的根使得这个根变为 $\alpha_{k,1}$ . 这样,

$$|\alpha_1 - \alpha_{k,1}|^n \leq |(\alpha_1 - \alpha_{k,1}) \cdots (\alpha_1 - \alpha_{k,1})| = |p_k(\alpha_1)|$$

右边收敛于 $|p(\alpha_1)| = 0$ . 所以, 左边也是. 这就证明了 $\alpha_{k,1} \rightarrow \alpha_1$ .

第二步: 写 $p_k(t) = (t - \alpha_{k,1})q_k(t)$ 与 $p(t) = (t - \alpha_1)q(t)$ . 于是,  $q_k$ 与 $q$ 都是首项系数为1的多项式, 并且它们的诸根分别为 $\alpha_{k,2}, \dots, \alpha_{k,n}$ 与 $\alpha_2, \dots, \alpha_n$ . 如果我们证明 $q_k \rightarrow q$ , 那么对次数 $n$ 用归纳法, 能够重排 $q_k$ 诸根的顺序, 使得它们收敛于 $q$ 的诸根, 这样就完成了证明.

为证 $q_k \rightarrow q$ , 我们做简单的除法. 为简化记号, 从 $\alpha_1$ 里去掉下标1. 比如说,  $p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0$ ,  $q(t) = t^{n-1} + b_{n-2}t^{n-2} + \cdots + b_1t + b_0$ ,  $p_k$ 与 $q_k$ 的记号是类似的. 方程 $p(t) = (t - \alpha)q(t)$ 蕴含着

$$\begin{aligned} b_{n-2} &= \alpha + \alpha_{n-1}, \\ b_{n-3} &= \alpha^3 + \alpha + \alpha_{n-2}, \\ &\vdots \\ b_0 &= \alpha^{n-1} + \alpha^{n-2}\alpha_{n-1} + \cdots + \alpha a_2 + a_1 \end{aligned}$$

因为  $\alpha_{k,i} \rightarrow \alpha_i$  与  $a_{k,i} \rightarrow a_i$ , 故  $b_{k,i} \rightarrow b_i$ . ■

**【5.2.2】命题** 令  $A$  是一个  $n \times n$  复矩阵.

(a) 存在收敛于  $A$  的矩阵序列  $A_k$ , 使得对所有  $k$ ,  $A_k$  的特征多项式  $p_k(t)$  有不同的根.

(b) 如果矩阵序列  $A_k$  收敛于  $A$ , 则它的特征多项式序列  $p_k(t)$  收敛于  $A$  的特征多项式  $p(t)$ .

(c) 令  $\lambda_i$  是特征多项式  $p$  的根. 如果  $A_k \rightarrow A$ , 则  $p_k$  的根可重新排序使得  $\lambda_{k,i} \rightarrow \lambda_i$  对每个  $i$  成立.

**证明**

(a) 命题 4.6.1 告诉我们, 存在可逆  $n \times n$  矩阵  $P$  使得  $A' = P^{-1}AP$  是上三角的. 它的特征值是这个上三角矩阵的对角元. 令  $A'_k$  是收敛于  $A'$  的矩阵序列, 其非对角元与  $A'$  的一样, 其对角元是不同的. 则  $A'_k$  是上三角的, 并且特征多项式有不同的根. 因为矩阵乘法是连续的, 故  $A_k \rightarrow A$ .  $A_k$  的特征多项式与  $A'_k$  的相同, 所以, 它有不同的根.

(b) 部分由 (a) 部分得到, 因为特征多项式的系数连续依赖于矩阵的元素. 这样, (c) 部分由命题 5.2.1 得到. ■

用连续性可证明著名的凯莱-哈密顿定理. 我们以矩阵形式叙述定理.

**【5.2.3】定理(凯莱-哈密顿定理)** 令  $p(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1t + c_0$  是  $n \times n$  复矩阵  $A$  的特征多项式. 则  $p(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I$  是零矩阵.

例如,  $2 \times 2$  矩阵  $A$  的特征多项式(例如  $a, b, c, d$  为其元素)是  $t^2 - (a+d)t + (ad-bc)$  (4.5.12). 定理断言

**【5.2.4】** 
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 - (a+d) \begin{bmatrix} a & b \\ c & d \end{bmatrix} + (ad-bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

很容易验证这个等式.

**凯莱-哈密顿定理的证明**

第一步:  $A$  是对角矩阵的情形. 令对角元素是  $\lambda_1, \dots, \lambda_n$ , 特征多项式是

$$p(t) = (t - \lambda_1) \cdots (t - \lambda_n)$$

这里  $p(A)$  也是对角矩阵, 且它的对角元素为  $p(\lambda_i)$ . 由于  $\lambda_i$  是  $p$  的根, 故  $p(\lambda_i) = 0$ ,  $p(A) = 0$ .

第二步:  $A$  的特征值不同的情形.

在此情形里,  $A$  是可对角化的; 比如说,  $A' = P^{-1}AP$  是对角的. 这样,  $A'$  的特征多项式与  $A$  的特征多项式相同, 而且,

$$p(A) = Pp(A')P^{-1}$$

(见(4.6.14)). 由第一步,  $p(A')=0$ , 所以  $p(A)=0$ .

第三步: 一般情形.

应用命题 5.2.2. 令  $A_k$  是其特征值不同且收敛于  $A$  的矩阵序列. 令  $p_k$  是  $A_k$  的特征多项式. 因为序列  $p_k$  收敛于  $A$  的特征多项式  $p$ , 故  $p_k(A_k) \rightarrow p(A)$ . 由第二步知,  $p_k(A_k)=0$  对所有  $k$  成立. 所以,  $p(A)=0$ . ■

140

### 第三节 微分方程组

在微积分中我们学过一阶线性微分方程

$$\text{【5.3.1】} \quad \frac{dx}{dt} = ax$$

的解为  $x(t) = ce^{at}$ , 其中  $c$  为任意常数. 回顾一下证明, 因为我们想再次使用其讨论. 首先,  $ce^{at}$  是该方程的解. 要证每一个解都有这样的形式, 令  $x(t)$  是任一个解. 利用乘积法则求微分  $e^{-at}x(t)$ :

$$\text{【5.3.2】} \quad \frac{d}{dt}(e^{-at}x(t)) = -ae^{-at}x(t) + e^{-at}ax(t) = 0$$

这样,  $e^{-at}x(t)$  是常数  $c$ , 并且  $x(t) = ce^{at}$ .

为将这个解拓广到常系数微分方程组, 我们使用下面的术语. 一个向量值函数或矩阵值函数是其元素为  $t$  的函数的一个向量或矩阵:

$$\text{【5.3.3】} \quad X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad A(t) = \begin{bmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{m1}(t) & \cdots & a_{mn}(t) \end{bmatrix}$$

取极限、微分等微积分运算通过分别对每一个元素进行运算拓广到向量值和矩阵值函数. 向量值或矩阵值函数的导数是对每个元素求导得到的函数:

$$\text{【5.3.4】} \quad \frac{dX}{dt} = \begin{bmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{bmatrix}, \quad \frac{dA}{dt} = \begin{bmatrix} a'_{11}(t) & \cdots & a'_{1n}(t) \\ \vdots & & \vdots \\ a'_{m1}(t) & \cdots & a'_{mn}(t) \end{bmatrix}$$

其中  $x'_i(t)$  是  $x_i(t)$  的导数, 等等. 于是,  $\frac{dX}{dt}$  有定义当且仅当每一个函数  $x_i(t)$  可微. 导数也可用向量记号描述为

$$\text{【5.3.5】} \quad \frac{dX}{dt} = \lim_{h \rightarrow 0} \frac{X(t+h) - X(t)}{h}$$

这里  $X(t+h) - X(t)$  用向量加法计算, 分母上的  $h$  是指用  $h^{-1}$  乘的标量. 极限是分别在每个元素上取极限得到的. 于是, (5.3.5) 的元素为导数  $x'_i(t)$ . 同样的结论对矩阵值函数也是正确的.

141

微分的许多性质可移到矩阵值函数上. 乘积法则就是一个例子, 其证明留作练习:

**【5.3.6】引理(乘积法则)**

(a) 令  $A(t)$  与  $B(t)$  是  $t$  的可微的矩阵值函数, 且它们可以做乘积运算, 则矩阵积  $A(t)B(t)$  是可微的, 并且它的导数是

$$\frac{d(AB)}{dt} = \frac{dA}{dt}B + A \frac{dB}{dt}$$

(b) 令  $A_1, \dots, A_k$  是  $t$  的可微的矩阵值函数, 且它们可以做乘积运算, 则矩阵积  $A_1 \cdots A_k$  是可微的, 并且它的导数是

$$\frac{d(A_1 \cdots A_k)}{dt} = \sum_{i=1}^k A_1 \cdots A_{i-1} \frac{dA_i}{dt} A_{i+1} \cdots A_k$$

齐次一阶线性常系数微分方程组是形如

$$\text{【5.3.7】} \quad \frac{dX}{dt} = AX$$

的矩阵方程, 其中  $A$  是一个  $n \times n$  的常数矩阵, 而  $X(t)$  为  $n$  维向量值函数. 写出这样的方程组, 我们得到形如

$$\frac{dx_1}{dt} = a_{11}x_1(t) + \cdots + a_{1n}x_n(t)$$

**【5.3.8】**

$$\vdots$$

$$\frac{dx_n}{dt} = a_{n1}x_1(t) + \cdots + a_{nn}x_n(t)$$

的  $n$  个微分方程的方程组.  $x_i(t)$  是未知函数,  $a_{ij}$  是标量. 例如, 如果

$$\text{【5.3.9】} \quad A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix},$$

(5.3.7) 便成为含有两个未知量的两个方程的方程组:

$$\frac{dx_1}{dt} = 3x_1 + 2x_2$$

**【5.3.10】**

$$\frac{dx_2}{dt} = x_1 + 4x_2$$

最简单的方程组是其中的  $A$  为对角矩阵的那些, 其对角元素为  $\lambda_i$ , 则方程(5.3.8) 写成

$$\text{【5.3.11】} \quad \frac{dx_i}{dt} = \lambda_i x_i(t), \quad i = 1, \dots, n$$

这里未知函数  $x_i$  没有被方程混合起来, 因而对某个常数  $c_i$ , 我们可以分别解出每一个

$$\text{【5.3.12】} \quad x_i = c_i e^{\lambda_i t}.$$

在大多数情形下, 解方程(5.3.7)的实际上是: 若  $V$  是  $A$  的一个伴随于特征值为  $\lambda$  的特征向量, 亦即, 若  $AV = \lambda V$ , 则

$$\text{【5.3.13】} \quad X = e^{\lambda t} V$$

是(5.3.7)的一个特解. 这里  $e^{\lambda t} V$  解释为变量标量  $e^{\lambda t}$  与常数向量  $V$  的积. 微分作用于标量

函数上, 固定向量  $V$ , 而用  $A$  左乘作用于向量  $V$ , 固定标量函数  $e^{\lambda t}$ . 这样  $\frac{d}{dt} e^{\lambda t} V = \lambda e^{\lambda t} V$ , 并且有  $A e^{\lambda t} V = \lambda e^{\lambda t} V$ . 例如,

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{与} \quad \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

是矩阵(5.3.9)的特征向量, 其伴随特征值分别为 5 与 2, 并且

$$\text{【5.3.14】} \quad \begin{bmatrix} e^{5t} \\ e^{5t} \end{bmatrix} \quad \text{与} \quad \begin{bmatrix} 2e^{2t} \\ -e^{2t} \end{bmatrix}$$

是微分方程组(5.3.10)的解.

这一事实使我们能够在矩阵  $A$  有不同的实特征值时解方程组(5.3.7). 这一情形中, 每一解将是特解(5.3.13)的线性组合. 通过对角化可方便地将解求出来.

**【5.3.15】命题** 令  $A$  是一个  $n \times n$  矩阵, 设  $P$  是一个可逆矩阵, 使得  $\Lambda = P^{-1}AP$  是对角矩阵, 其对角元为  $\lambda_1, \dots, \lambda_n$ . 方程组  $\frac{dX}{dt} = AX$  的一般解是  $X = P \tilde{X}$ , 其中  $\tilde{X} = (c_1 e^{\lambda_1 t}, \dots, c_n e^{\lambda_n t})'$  是方程  $\frac{d\tilde{X}}{dt} = \Lambda \tilde{X}$  的解.

常数  $c_i$  是任意的. 它们常由指定的初始条件—— $X$  在某一特殊  $t_0$  的值确定.

**证明** 用  $P$  乘方程  $\frac{d\tilde{X}}{dt} = \Lambda \tilde{X}$ :  $P \frac{d\tilde{X}}{dt} = P\Lambda \tilde{X} = AP \tilde{X}$ . 但因  $P$  是一个常量, 故  $P \frac{d\tilde{X}}{dt} =$

$\frac{d(P\tilde{X})}{dt} = \frac{dX}{dt}$ . 这个推导可逆推回去, 所以  $\tilde{X}$  是带有  $\Lambda$  的方程的解当且仅当  $X$  是带有  $A$  的方程的解. ■

对角化矩阵(5.3.10)的矩阵在(4.6.8)之前计算出来:

$$\text{【5.3.16】} \quad A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 5 & \\ & 2 \end{bmatrix}$$

143

因此,

$$\text{【5.3.17】} \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = P \tilde{X} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 e^{5t} \\ c_2 e^{2t} \end{bmatrix} = \begin{bmatrix} c_1 e^{5t} + 2c_2 e^{2t} \\ c_1 e^{5t} - c_2 e^{2t} \end{bmatrix}$$

换言之, 每个解是两个基本解(5.3.14)的线性组合.

我们现在考虑系数矩阵  $A$  有不同的特征值但是不全为实数的情形. 为了重复上面所用的方法, 必须先考虑形如(5.3.1)的微分方程, 其中  $a$  为复数. 通过适当的解释, 这样的微分方程的解仍为  $ce^{at}$  的形式. 唯一要记住的是  $e^{at}$  现在是实变量  $t$  的复值函数.

倘若极限(5.3.5)存在, 则复值函数导数的定义与实值函数是一样的. 不会出现新的特征. 可将这样的函数  $x(t)$  用实值函数, 即其实部和虚部写出来, 比如说

$$\text{【5.3.18】} \quad x(t) = p(t) + iq(t)$$

则  $x$  是可微的当且仅当  $p$  和  $q$  都是可微的, 且当它们可微时,  $x$  的导数为  $x' = p' + iq'$ . 这由定义就可直接得到. 通常的微分法则(比如乘法法则)对复值函数成立. 这些法则可在  $p$  和  $q$  上应用实函数相应的定理得到, 也可将实函数的证明搬到复函数的情形.

复数  $a = r + si$  的幂指数定义为

$$\text{【5.3.19】} \quad e^a = e^{r+si} = e^r(\cos s + i\sin s)$$

这个公式的微分表明  $de^a/dt = ae^a$ . 所以,  $ce^a$  是微分方程(5.3.1)的解, 且本节开头的证明表明这是仅有的解.

将一个方程的情形推广到复系数的情形后, 当  $A$  是具有不同特征值的任意复矩阵时, 可以用对角化方法解方程组(5.3.7).

例如, 设  $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ . 向量  $v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$  与  $v_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$  分别是伴随于特征值  $1+i$  与  $1-i$  的特征向量. 设基  $B = (v_1, v_2)$ . 则  $A$  可以由矩阵  $P = [B]$  对角化:

$$\text{【5.3.20】} \quad P^{-1}AP = \frac{1}{2} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} = \begin{bmatrix} 1+i & \\ & 1-i \end{bmatrix} = \Lambda$$

于是,  $\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} c_1 e^{(1+i)t} \\ c_2 e^{(1-i)t} \end{bmatrix}$ . (5.3.7)的解是

$$\text{【5.3.21】} \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = P \tilde{X} = \begin{bmatrix} c_1 e^{(1+i)t} + ic_2 e^{(1-i)t} \\ ic_1 e^{(1-i)t} + c_2 e^{(1-i)t} \end{bmatrix}$$

其中  $c_1, c_2$  是任意复数. 因而每个解都是两个基本解

$$\text{【5.3.22】} \quad \begin{bmatrix} e^{(1+i)t} \\ ie^{(1+i)t} \end{bmatrix} \text{ 与 } \begin{bmatrix} ie^{(1-i)t} \\ e^{(1-i)t} \end{bmatrix}$$

的线性组合. 然而, 这些解并不是完全令人满意的, 因为我们从实系数微分方程开始, 得到的答案却是复的. 当原来的矩阵为实矩阵时, 我们想要的解为实解. 注意下面的引理:

**【5.3.23】引理** 设  $A$  是实  $n \times n$  矩阵, 并设为  $X(t)$  是微分方程  $\frac{dX}{dt} = AX$  的复值解. 则  $X(t)$  的实部和虚部也是同一方程的解.

现在原方程(5.3.7)的每个解无论是实的还是复的, 对某个复数  $c_i$  都具有(5.3.21)的形式. 于是, 实解亦在我们所得的解之中. 为了把它们具体写出, 可取复解的实部和虚部.

基本解(5.3.22)的实部和虚部由等式(5.3.19)确定. 它们是

$$\text{【5.3.24】} \quad \begin{bmatrix} e^t \cos t \\ -e^t \sin t \end{bmatrix} \text{ 和 } \begin{bmatrix} e^t \sin t \\ -e^t \cos t \end{bmatrix}.$$

每个实解都是这些特解的实线性组合.

## 第四节 矩阵指数

一阶线性常系数微分方程组亦可用矩阵指数形式求解.

$n \times n$  实或复矩阵  $A$  的指数由将矩阵代入  $e^x$  的泰勒展开式

$$\text{【5.4.1】} \quad e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad \boxed{145}$$

得到, 其中将  $x$  替换为  $A$ , 将  $1$  替换为  $I$ . 这样由定义,

$$\text{【5.4.2】} \quad e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$$

我们主要感兴趣的是标量变量  $t$  的矩阵值函数  $e^{tA}$ . 于是, 用  $tA$  替换  $A$ :

$$\text{【5.4.3】} \quad e^{tA} = I + \frac{tA}{1!} + \frac{t^2 A^2}{2!} + \frac{t^3 A^3}{3!} + \dots$$

### 【5.4.4】定理

(a) 序列 (5.4.2) 对复矩阵的有界集合绝对收敛和一致收敛.

(b)  $e^{tA}$  是  $t$  的可微函数, 并且它的导数是矩阵积  $Ae^{tA}$ .

(c) 令  $A$  与  $B$  是交换的复  $n \times n$  矩阵:  $AB=BA$ . 则  $e^{A+B} = e^A e^B$ .

为不打断讨论, 我们把这个定理的证明放到本节末尾.

$A$  与  $B$  是交换的假设对于矩阵保持基本性质  $e^{x+y} = e^x e^y$  是必需的. 不管怎样, (c) 是很有用的.

**【5.4.5】推论** 对任意  $n \times n$  复矩阵  $A$ , 矩阵指数  $e^A$  是可逆的, 并且它的逆为  $e^{-A}$ .

**证明** 因为  $A$  与  $-A$  是交换的, 故  $e^A e^{-A} = e^{A-A} = e^0 = I$ . ■

因为矩阵乘法相对较为复杂, 所以直接写出  $e^A$  的矩阵元素并不容易. 特别是  $e^A$  的元素通常不是由  $A$  的元素取幂得到, 除非  $A$  是对角矩阵. 如果  $A$  是对角矩阵, 其对角元素为  $\lambda_1, \dots, \lambda_n$ , 则考察该级数表明  $e^A$  也是对角的, 并且其对角元素为  $e^{\lambda_i}$ .

对于  $2 \times 2$  三角矩阵, 指数的计算也相对比较容易. 例如, 设

$$A = \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix}$$

则

$$\text{【5.4.6】} \quad e^A = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \frac{1}{1!} \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} 1 & 3 \\ & 4 \end{bmatrix} + \dots = \begin{bmatrix} e & * \\ & e^2 \end{bmatrix}$$

由序列直接算出缺失的元素  $*$  是很好的练习. □

只要知道有矩阵  $P$  使  $\Lambda = P^{-1}AP$  是对角的, 就可以确定矩阵  $A$  的指数. 应用法则  $P^{-1}A^kP = (P^{-1}AP)^k$  (4.6.12) 和矩阵乘法的分配律,

$$\text{【5.4.7】} \quad P^{-1}e^AP = (P^{-1}IP) + \frac{(P^{-1}AP)}{1!} + \frac{(P^{-1}AP)^2}{2!} + \dots = e^{P^{-1}AP} = e^\Lambda$$

设  $\Lambda$  是对角的, 其对角元素为  $\lambda_i$ . 则  $e^\Lambda$  也是对角的, 其对角元素为  $e^{\lambda_i}$ . 因此可以具体地



算出  $e^A$  :

$$\text{【5.4.8】} \quad e^A = Pe^{\Lambda}P^{-1}$$

例如, 如果  $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$  且  $P = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ , 则  $P^{-1}AP = \Lambda = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$ . 所以

$$e^A = Pe^{\Lambda}P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e & \\ & e^2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix} = \begin{bmatrix} e & e^2 - e \\ & e^2 \end{bmatrix}$$

下一个定理联系矩阵指数与微分方程组:

**【5.4.9】定理** 令  $A$  是实或复  $n \times n$  矩阵, 矩阵  $e^{tA}$  的列构成微分方程  $\frac{dX}{dt} = AX$  解空间的基.

**证明** 定理 5.4.4(b) 表明  $e^{tA}$  的列是微分方程的解. 要证每个解是列的线性组合, 我们只需复制第三节开始给出的证明. 设  $X(t)$  是任意一个解. 利用乘积法则 (5.3.6), 对矩阵乘积  $e^{-tA}X(t)$  进行微分, 得

$$\text{【5.4.10】} \quad \frac{d}{dt}(e^{-tA}X(t)) = (-Ae^{-tA})X(t) + e^{-tA}(AX(t))$$

幸运的是,  $A$  与  $e^{-tA}$  可交换. 这可直接由指数的定义得到. 所以导数为零. 由此得  $e^{-tA}X(t)$  是一个常量列向量, 比如设为  $C = (c_1, \dots, c_n)^t$ , 则  $X(t) = e^{tA}C$ . 这就把  $X(t)$  表示为  $e^{tA}$  列的线性组合, 而  $c_i$  为其系数. 由于  $e^{tA}$  是可逆矩阵, 故这个表达式唯一的. ■

尽管矩阵指数总是微分方程 (5.3.7) 的解, 但因为指数的直接计算非常困难, 所以这个定理在具体情形的应用并不容易. 但如果  $A$  是可对角化矩阵, 则指数可如 (5.4.8) 中那样进行计算. 我们可用计算  $e^{tA}$  的方法解方程 (5.3.7), 当然会得到与前面相同的结果. 这样, 如果  $A, P$  和  $\Lambda$  是如 (5.3.16) 中所用到的, 则

$$\text{【147】} \quad e^A = Pe^{\Lambda}P^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{5t} & \\ & e^{2t} \end{bmatrix} \begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix} = \left(\frac{1}{3}\right) \begin{bmatrix} (e^{5t} + 2e^{2t}) & (2e^{5t} - 2e^{2t}) \\ (e^{5t} - e^{2t}) & (2e^{5t} + e^{2t}) \end{bmatrix}$$

矩阵右边的列构成 (5.3.7) 里得到的解空间的另一个基.

也可用若尔当形解微分方程. 对任意  $k \times k$  若尔当块  $J_\lambda$  (4.7.5), 可通过计算矩阵指数确定其解. 如同 (4.7.12) 中那样, 记  $J_\lambda = \lambda I + N$ , 其中  $N$  是  $k \times k$  若尔当块  $J_0$ , 且  $\lambda = 0$ . 于是,  $N^k = 0$ , 所以

$$e^{tN} = I + \frac{tN}{1!} + \dots + \frac{t^{k-1}N^{k-1}}{(k-1)!}$$

由于  $N$  与  $\lambda I$  是交换的, 故

$$e^{tJ} = e^{t\lambda I} e^{tN} = e^{t\lambda} \left( I + \frac{tN}{1!} + \dots + \frac{t^{k-1}N^{k-1}}{(k-1)!} \right)$$

因此, 如果  $J$  是  $3 \times 3$  矩阵块

$$J = \begin{pmatrix} 3 & & \\ 1 & 3 & \\ & 1 & 3 \end{pmatrix}$$

则

$$e^{tJ} = \begin{bmatrix} e^{3t} & & \\ & e^{3t} & \\ & & e^{3t} \end{bmatrix} \begin{bmatrix} 1 & & \\ t & 1 & \\ \frac{1}{2!}t^2 & t & 1 \end{bmatrix} = \begin{bmatrix} e^{3t} & & \\ te^{3t} & e^{3t} & \\ \frac{1}{2!}t^2e^{3t} & te^{3t} & e^{3t} \end{bmatrix}$$

这个矩阵的列构成微分方程  $\frac{dX}{dt} = JX$  解空间的一个基.

现在我们回过头来证明定理 5.4.4. 下面给出将要用到的关于序列极限的主要事实, 并参阅文献 [Mattuck] 与 [Rudin]. 这些作者仅考虑实值函数, 但证明可以搬到复值函数上, 因为复值函数的极限与导数可分别通过计算实部和虚部的极限与导数给出定义.

如果  $r$  与  $s$  是实数, 且  $r < s$ , 记号  $[r, s]$  表示区间  $r \leq t \leq s$ .

**【5.4.11】定理** ([Mattuck] 定理 22.2B, [Rudin] 定理 7.9) 令  $m_k$  是正实数序列, 使得  $\sum m_k$  收敛. 如果  $u^{(k)}(t)$  是区间  $[r, s]$  上的函数, 且对所有  $k$  和区间里的所有  $t$ , 有  $|u^{(k)}(t)| \leq m_k$ , 则级数  $\sum u^{(k)}(t)$  在区间上一致收敛.

**【5.4.12】定理** ([Mattuck] 定理 11.5B, [Rudin] 定理 7.17) 令  $u^{(k)}(t)$  是在区间  $[r, s]$  上有连续导函数的函数列. 假设级数  $\sum u^{(k)}(t)$  在区间上收敛于函数  $f(t)$ , 且导数级数  $\sum u'^{(k)}(t)$  在区间上一致收敛于函数  $g(t)$ . 则  $f$  在区间上可微, 且它的导数是  $g$ .

148

**定理 5.4.4(a)的证明** 这里我们将矩阵  $A$  的  $i, j$  元素记为  $(A)_{ij}$ . 这样,  $(AB)_{ij}$  表示乘积矩阵  $AB$  的元素, 而  $(A^k)_{ij}$  表示  $A^k$  的元素. 借助这个记号,  $e^A$  的  $i, j$  元素是级数的和

$$\text{【5.4.13】} \quad (e^A)_{ij} = (I)_{ij} + \frac{(A)_{ij}}{1!} + \frac{(A^2)_{ij}}{2!} + \frac{(A^3)_{ij}}{3!} + \dots$$

为了证明指数级数绝对收敛和一致收敛, 我们需要证明给定矩阵的幂  $A^k$  的元素不会增长得太快.

我们用记号  $\|A\|$  表示矩阵  $A$  的元素的绝对值的最大绝对值, 即使得

$$\text{【5.4.14】} \quad |(A)_{ij}| \leq \|A\| \quad \text{对所有 } i, j$$

成立的最小实数. 它的基本性质是:

**【5.4.15】引理** 设  $A, B$  是复  $n \times n$  矩阵. 则  $\|AB\| \leq n\|A\|\|B\|$ , 且对所有  $k > 0$ , 有  $\|A^k\| \leq n^{k-1}\|A\|^k$ .

**证明** 我们估计  $AB$  的  $i, j$  元素的大小:

$$|(AB)_{ij}| = \left| \sum_{v=1}^n (A)_{iv}(B)_{vj} \right| \leq \sum_{v=1}^n |(A)_{iv}| |(B)_{vj}| \leq n\|A\|\|B\|$$

从第一个不等式出发, 作数学归纳法可得第二个不等式. ■

我们现在估计指数级数: 令  $a$  是正实数, 使得  $n\|A\| \leq a$ . 由引理 5.4.15 知,  $|(A^k)_{ij}| \leq a^k$  (省略一个  $n$ ). 所以

$$\text{【5.4.16】} \quad |(e^A)_{ij}| \leq |(I)_{ij}| + |(A)_{ij}| + \frac{1}{2!} |(A^2)_{ij}| + \frac{1}{3!} |(A^3)_{ij}| + \dots$$

$$\leq 1 + \frac{a}{1!} + \frac{a^2}{2!} + \frac{a^3}{3!} + \dots$$

比率计算表明最后级数当然收敛于  $e^a$ . 定理 5.4.11 表明对所有满足  $n\|A\| \leq a$  的  $A$ ,  $e^A$  的级数一致收敛与绝对收敛. ■

**定理 5.4.4(b)和(c)的证明** 我们使用一个技巧以缩短证明. 首先, 假设  $A$  与  $B$  是交换的  $n \times n$  矩阵, 对  $e^{tA+B}$  的级数进行微分.  $tA+B$  的导数是  $A$ , 且

$$\text{【5.4.17】} \quad e^{tA+B} = I + \frac{(tA+B)}{1!} + \frac{(tA+B)^2}{2!} + \dots$$

用乘积法则(5.3.6), 对  $k > 0$ , 这个级数的  $k$  次项的导数为

149

$$\frac{d}{dt} \left( \frac{(tA+B)^k}{k!} \right) = \left( \frac{1}{k!} \sum_{i=1}^k (tA+B)^{i-1} A (tA+B)^{k-i} \right)$$

因为  $AB=BA$ , 故可抽出中间项里的  $A$  到左边:

$$\text{【5.4.18】} \quad \frac{d}{dt} \left( \frac{(tA+B)^k}{k!} \right) = kA \frac{(tA+B)^{k-1}}{k!} = A \frac{(tA+B)^{k-1}}{(k-1)!}$$

这是矩阵  $A$  与矩阵指数级数  $k-1$  次项的乘积. 所以, 对(5.4.17)逐项微分得到  $Ae^{tA+B}$  的级数.

要证逐项微分是合理的, 我们应用定理 5.4.4(a). 这个定理表明对给定的  $A$  与  $B$ , 指数级数  $e^{tA+B}$  在任意区间  $r \leq t \leq s$  上一致收敛. 而且, 导数级数一致收敛于  $Ae^{tA+B}$ . 由定理 5.4.12,  $e^{tA+B}$  的导数可逐项计算, 所以, 对任意交换矩阵对  $A, B$ ,

$$\frac{d}{dt} e^{tA+B} = Ae^{tA+B}$$

取  $B=0$  就证明了定理 5.4.4(b).

其次, 我们复制定理 5.4.9 的证明所用过的方法. 仍假设  $A$  与  $B$  是交换的, 对乘积  $e^{-tA} e^{tA+B}$  进行微分. 如同(5.4.10)中那样, 我们发现

$$\frac{d}{dt} (e^{-tA} e^{tA+B}) = (-Ae^{-tA})(e^{tA+B}) + (e^{-tA})(Ae^{tA+B}) = 0$$

所以,  $e^{-tA} e^{tA+B} = C$ , 其中  $C$  是常数矩阵. 置  $t=0$  表明  $e^B = C$ . 置  $B=0$  表明  $e^{-tA} = (e^{tA})^{-1}$ . 则  $(e^{tA})^{-1} e^{tA+B} = e^B$ . 置  $t=1$  表明  $e^{A+B} = e^A e^B$ . 这就证明了定理 5.4.4(c). ■

在第九章我们将再次使用矩阵指数的著名性质.

对一般情形中定理的证明工作

我认为没有必要.

— Arthur Cayley <sup>⊖</sup>

⊖ 亚瑟·凯莱, 以其名命名凯莱-哈密顿定理的数学家之一, 他的论文叙述的是  $n \times n$  矩阵的一般情形, 而验证的是  $2 \times 2$  的情形(见(5.2.4)). 他用这里所引述的话来结束该定理的讨论.

## 练 习

### 第一节 正交矩阵与旋转

- 1.1 确定表示  $\mathbf{R}^3$  的下列旋转的矩阵:  
 (a) 角  $\theta$ , 轴  $e_2$  (b) 角  $2\pi/3$ , 包含向量  $(1, 1, 1)'$  的轴 (c) 角  $\pi/2$ , 包含向量  $(1, 1, 0)'$  的轴
- 1.2 表示绕轴  $u$  转过  $\theta$  角  $\mathbf{R}^3$  的旋转的矩阵  $A$  的复特征值是什么?
- 1.3  $O_n$  是否同构于积群  $SO_n \times \{\pm I\}$ ?
- 1.4 给出行列式为  $-1$  的正交  $3 \times 3$  矩阵作用的几何描述.
- 1.5 令  $A$  是  $3 \times 3$  正交矩阵,  $\det A = 1$ , 其旋转角不同于  $0$  或  $\pi$ , 且设  $M = A - A'$ .  
 (a) 证明  $M$  的秩为  $2$ , 且  $M$  的零空间里的非零向量  $X$  是  $A$  的伴随于特征值  $1$  的特征向量.  
 (b) 求这样一个用  $A$  的元素简洁表出的特征向量.

150

### 第二节 连续性的使用

- 2.1 使用凯莱-哈密顿定理给出由  $A$ ,  $(\det A)^{-1}$  和特征多项式的系数表示的  $A^{-1}$  的表达式. 在  $2 \times 2$  情形里验证你给出的表达式.
- 2.2 令  $A$  是  $m \times m$  复矩阵,  $B$  是  $n \times n$  复矩阵, 考虑在所有复矩阵的空间  $\mathbf{C}^{m \times n}$  上由  $T(M) = AMB$  定义的线性算子  $T$ .  
 (a) 指出如何由一对列向量  $X, Y$  构造的特征向量, 其中  $X$  是  $A$  的特征向量而  $Y$  是  $B'$  的特征向量.  
 (b) 用  $A$  和  $B$  的特征值确定  $T$  的特征值.  
 (c) 确定这个算子的迹.
- 2.3 令  $A$  是  $n \times n$  复矩阵.  
 (a) 考虑在所有复  $n \times n$  矩阵的空间  $\mathbf{C}^{n \times n}$  上由法则  $T(M) = AM - MA$  定义的线性算子. 证明这个算子的秩至多是  $n^2 - n$ .  
 (b) 用  $T$  的特征值  $\lambda_1, \dots, \lambda_n$  确定  $A$  的特征值.
- 2.4 令  $A$  和  $B$  是可对角化的复矩阵. 证明存在可逆矩阵  $P$  使得  $P^{-1}AP$  与  $P^{-1}BP$  均是对角的当且仅当  $AB = BA$ .

### 第三节 微分方程组

- 3.1 证明矩阵值函数微分的乘积法则.
- 3.2 令  $A(t)$  与  $B(t)$  是  $t$  的矩阵值可微函数. 计算  
 (a)  $\frac{d}{dt}(A(t)^3)$  (b)  $\frac{d}{dt}(A(t)^{-1})$  (c)  $\frac{d}{dt}(A(t)^{-1}B(t))$
- 3.3 对下列矩阵  $A$  解微分方程  $\frac{dX}{dt} = AX$ :  
 (a)  $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$  (b)  $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$  (c)  $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & -1 \end{bmatrix}$  (d)  $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$
- 3.4 令  $A$  和  $B$  是常数矩阵, 且  $A$  是可逆的. 用方程  $\frac{dX}{dt} = AX$  的解解非齐次微分方程  $\frac{dX}{dt} = AX + B$ .

### 第四节 矩阵指数

- 4.1 对下列矩阵  $A$  计算  $e^A$ :

151

$$(a) \begin{bmatrix} a & b \end{bmatrix} \quad (b) \begin{bmatrix} 2\pi i & 2\pi i \\ & 2\pi i \end{bmatrix} \quad (c) \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix} \quad (d) \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad (e) \begin{bmatrix} 0 & & \\ 1 & 0 & \\ & 1 & 0 \end{bmatrix}$$

4.2 证明公式  $e^{\text{trace}A} = \det(e^A)$ .4.3 设  $X$  是  $n \times n$  矩阵  $A$  的特征向量, 特征值为  $\lambda$ .(a) 证明: 如果  $A$  可逆, 则  $X$  也是  $A^{-1}$  的特征向量, 特征值是  $\lambda^{-1}$ .(b) 证明  $X$  是  $e^A$  的特征向量, 特征值是  $e^\lambda$ .4.4 令  $A$  和  $B$  是可交换矩阵. 为证  $e^{A+B} = e^A e^B$ , 可把两边展开为双和, 其诸项为  $A^i B^j$  的倍数. 证明所得的两个双和是相同的.4.5 当  $A$  是已知矩阵, 解微分方程  $\frac{dX}{dt} = AX$ :

$$(a) \begin{bmatrix} 2 & \\ 1 & 2 \end{bmatrix} \quad (b) \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad (c) \begin{bmatrix} 1 & & \\ 1 & 1 & \\ & 1 & 1 \end{bmatrix}$$

4.6 对  $n \times n$  矩阵  $A$ , 用  $\sin x$  和  $\cos x$  的泰勒级数展开式定义  $\sin A$  和  $\cos A$ .(a) 证明: 对所有  $A$ , 这些级数收敛.(b) 证明  $\sin(tA)$  是  $t$  的可微函数, 且  $\frac{d\sin(tA)}{dt} = A\cos(tA)$ .

4.7 讨论下列恒等式成立的范围:

(a)  $\cos^2 A + \sin^2 A = I$ ,(b)  $e^{iA} = \cos A + i\sin A$ ,(c)  $\sin(A+B) = \sin A \cos B + \cos A \sin B$ ,(d)  $e^{2\pi i A} = I$ ,(e)  $\frac{d(e^{A(t)})}{dt} = e^{A(t)} \frac{dA}{dt}$ , 其中  $A(t)$  是  $t$  的可微的矩阵值函数.4.8 令  $P, B_k$  与  $B$  是  $n \times n$  矩阵, 其中  $P$  是可逆的. 证明: 如果  $B_k$  收敛于  $B$ , 则  $P^{-1} B_k P$  收敛于  $P^{-1} B P$ .**杂题**M.1 确定具有整数元素的正交矩阵群  $O_n(\mathbf{Z})$ .

M.2 用若尔当形证明凯莱-哈密顿定理.

M.3 设  $A$  是一个  $n \times n$  复矩阵. 证明: 如果对所有  $k > 0$  有迹  $\text{trace} A^k = 0$ , 则  $A$  是幂零的.M.4 令  $A$  是  $n \times n$  复矩阵, 其所有特征值的绝对值均小于 1, 证明级数  $I + A + A^2 + \cdots$  收敛于  $(I - A)^{-1}$ .M.5 斐波那契数  $0, 1, 1, 2, 3, 5, 8, \cdots$  是在初始条件  $f_0 = 0, f_1 = 1$  之下由迭代关系  $f_n = f_{n-1} + f_{n-2}$  定义的. 迭代关系可用矩阵形式写为

152

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} f_{n-2} \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} f_{n-1} \\ f_n \end{bmatrix}$$

(a) 证明公式

$$f_n = \frac{1}{\alpha} \left[ \left( \frac{1+\alpha}{\alpha} \right)^n - \left( \frac{1-\alpha}{2} \right)^2 \right]$$

其中  $\alpha = \sqrt{5}$ .

(b) 设序列  $a_n$  由关系  $a_n = \frac{1}{2}(a_{n-1} + a_{n-2})$  定义, 用  $a_0, a_1$  计算  $\lim a_n$ .

M. 6 (积分算子) 区间  $[0, 1]$  上连续函数  $f(u)$  的空间  $C$  是类似  $\mathbf{R}^n$  的无限维空间之一, 且方形区域  $0 \leq u, v \leq 1$  上的连续函数  $A(u, v)$  是无限维的, 类似于矩阵. 积分

$$A \cdot f = \int_0^1 A(u, v) f(v) dv$$

类似于矩阵与向量乘法. (为使其形象化, 在  $u, v$  平面按顺时针方向旋转单位方形和区间  $[0, 1]$   $90^\circ$ .) 在适当的假设下, 桥对变化的载重的反应由这个积分表示. 为此,  $f$  表示沿桥的载重, 于是  $A \cdot f$  将把由载重引起的桥的垂度计算出来.

这个问题把积分作为线性算子处理. 对函数  $A = u + v$ , 确定算子的像. 确定它的非零特征值, 用某些为零的积分描述它的核. 对函数  $A = u^2 + v^2$  做同样的事.

M. 7 令  $A$  是  $2 \times 2$  复矩阵, 其特征值互不相同, 设  $X$  是  $2 \times 2$  未定矩阵. 矩阵方程  $X^2 = A$  有多少个解?

M. 8 用几何方法确定两个三维旋转的合成的旋转轴.

## 第六章 对 称

代数不是写出的几何，  
几何也不是画出的代数。

——Sophie Germain

对称为群论提供了最引人入胜的应用。群最早是为了分析称为扩域的代数结构——域扩张(第十六章)——的对称性而发明的。因为对称性是所有科学中一个共同的现象，所以它仍是群论应用的两个主要方式之一。另一个应用方式是通过群的表示，我们将在第十章加以讨论。第一节所研究的平面图形的对称称为第七节引入的群作用的一般概念提供了丰富的实例和背景。

我们将允许自由使用几何推理，而回溯到几何公理的讨论将留到其他场合。

### 第一节 平面图形的对称

平面图形的对称通常可分为图 6.1.1~图 6.1.3 所示的几种主要类型。

【6.1.1】图



双侧对称

【6.1.2】图



旋转对称

154

【6.1.3】图



平移对称

这样的图形可假设在两个方向上任意延伸。还存在第四类对称，尽管对它的名称(滑动对称)可能不太熟悉。

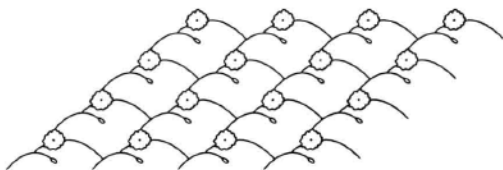
【6.1.4】图



滑动对称

像墙纸图案这样的图形可以有两个独立的平移对称。

【6.1.5】图



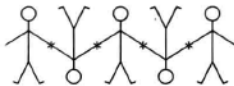
其他对称的组合也可能出现. 例如, 星形图案具有双侧对称和旋转对称. 在下面的图形里, 平移对称与旋转对称组合起来:

【6.1.6】图



另一个例子:

【6.1.7】图



平面的刚体运动称为一个等距, 并且如果一个等距把平面的子集  $F$  映为自身, 就称为  $F$  的对称.  $F$  的全体对称的集合构成平面的所有等距群的一个子群: 如果  $m$  与  $m'$  映  $F$  到  $F$ , 则合成映射  $mm'$  也映  $F$  到  $F$ , 等等. 这是  $F$  的对称群.

图 6.1.3 的对称群是由使图形向左移动一个单位的平移  $t$  生成的一个无限循环群.

$$G = \{\dots, t^{-2}, t^{-1}, 1, t, t^2, \dots\}$$

图 6.1.7 除了平移外还有对称.

## 第二节 等 距

$\mathbf{R}^n$  的两点间的距离是向量  $u-v$  的长度  $|u-v|$ .  $n$  维空间  $\mathbf{R}^n$  的等距是从  $\mathbf{R}^n$  到其自身的一个保持距离的映射  $f$ , 它是一个对  $\mathbf{R}^n$  的所有  $u$  与  $v$ , 使得

$$|f(u) - f(v)| = |u - v|$$

成立的映射. 等距映射把一个图形映为一个全等图形.

【6.2.2】例

(a) 正交线性算子是等距.

因为正交算子  $\varphi$  是线性的, 故  $\varphi(u) - \varphi(v) = \varphi(u-v)$ , 从而  $|\varphi(u) - \varphi(v)| = |\varphi(u-v)|$ . 由于  $\varphi$  是正交的, 它保持点积, 故保持长度, 于是  $|\varphi(u-v)| = |u-v|$ .

(b) 对向量  $a$ , 由  $t_a(x) = x+a$  定义的平移映射  $t_a$  是等距.

平移不是线性算子, 因为它不能把  $0$  映为  $0$ , 除非由零向量所定义的平移, 当然这时它是恒等映射.

(c) 等距的合成是等距. ■

【6.2.3】定理 关于映射  $\varphi: \mathbf{R}^n \rightarrow \mathbf{R}^n$ , 下列条件是等价的:

(a)  $\varphi$  是固定原点的等距:  $\varphi(0) = 0$ .



(b)  $\varphi$  保持点积: 对所有  $v$  和  $w$ ,  $(\varphi(v) \cdot \varphi(w)) = (v \cdot w)$ .

(c)  $\varphi$  是正交线性算子.

我们已经看到(c)蕴含(a). 下面给出的(b) $\Rightarrow$ (c)的证明是几年前由 Sharon Hollander 提供的, 当时她是 MIT 代数班里的学生.

**[6.2.4] 引理** 令  $x$  与  $y$  是  $\mathbf{R}^n$  的点, 如果三个点积  $(x \cdot x)$ ,  $(x \cdot y)$  与  $(y \cdot y)$  相等, 则  $x=y$ .

**证明** 假设  $(x \cdot x) = (x \cdot y) = (y \cdot y)$ . 则

$$((x-y) \cdot (x-y)) = (x \cdot x) - 2(x \cdot y) + (y \cdot y) = 0$$

$x-y$  的长度为零, 所以,  $x=y$ . ■

**定理 6.2.3 中(b) $\Rightarrow$ (c)的证明** 令  $\varphi$  是保持点积的映射. 则它是正交的, 倘若它是线性算子(5.1.12). 要证  $\varphi$  是线性算子, 需证  $\varphi(u+v) = \varphi(u) + \varphi(v)$  与  $\varphi(cv) = c\varphi(v)$  对所有的  $u$  与  $v$  和所有的标量  $c$  成立.

已知  $x \in \mathbf{R}^n$ , 用符号  $x'$  代表  $\varphi(x)$ . 也用符号  $w$  表示和, 写成  $w = u+v$ . 于是, 要证的关系  $\varphi(u+v) = \varphi(u) + \varphi(v)$  变成了  $w' = u' + v'$ .

用  $x=w'$  与  $y=u'+v'$  在引理 6.2.4 里做替换. 要证  $w' = u' + v'$ , 只需证明三个点积

$$(w' \cdot w'), \quad (w' \cdot (u' + v')) \quad \text{与} \quad ((u' + v') \cdot (u' + v'))$$

相等. 展开第二个和第三个点积. 只需证明

$$(w' \cdot w') = (w' \cdot u') + (w' \cdot v') = (u' \cdot u') + 2(u' \cdot v') + (v' \cdot v')$$

由假设,  $\varphi$  保持点积. 所以, 我们可丢掉撇:  $(w' \cdot w') = (w \cdot w)$ , 等等. 这样, 只需证明

$$\mathbf{[6.2.5]} \quad (w \cdot w) = (w \cdot u) + (w \cdot v) = (u \cdot u) = 2(u \cdot v) + (v \cdot v)$$

现在, 要证  $w' = u' + v'$ , 而由定义,  $w = u+v$  是成立的. 于是, 可用  $u+v$  替换  $w$ . 这样, (6.2.5)就成立了.

要证  $\varphi(cv) = c\varphi(v)$ , 写  $u=cv$ , 我们需证  $u' = cv'$ . 证明与我们刚给出的证明类似. ■

**定理 6.2.3 中(a) $\Rightarrow$ (b)的证明** 令  $\varphi$  是固定原点的等距. 利用撇号,  $\varphi$  的保距性质为

$$\mathbf{[6.2.6]} \quad ((u' - v') \cdot (u' - v')) = ((u - v) \cdot (u - v))$$

对  $\mathbf{R}^n$  的所有  $u$  与  $v$  成立. 替换  $v=0$ . 由于  $0'=0$ , 故  $(u' \cdot u') = (u, u)$ . 类似地,  $(v' \cdot v') = (v, v)$ . 现在, 展开(6.2.6)式并且从方程两边消去  $(u, u)$  和  $(v, v)$  就得(b). ■

**[6.2.7] 推论**  $\mathbf{R}^n$  的每个等距  $f$  是一个正交线性算子与一个平移的合成. 更确切地, 如果  $f$  是一个等距, 并且  $f(0)=a$ , 则  $f=t_a\varphi$ , 其中  $t_a$  是平移,  $\varphi$  是正交线性算子.  $f$  的这个表达式是唯一的.

**证明** 令  $f$  是一个等距, 设  $a=f(0)$ ,  $\varphi=t_{-a}f$ . 推论等于断言:  $\varphi$  是正交线性算子.

因为  $\varphi$  是等距  $t_{-a}$  与  $f$  的合成, 故它是一个等距. 因为  $\varphi(0)=t_{-a}f(0)=t_{-a}(a)=0$ , 所以  $\varphi$  固定原点. 定理 6.2.3 表明  $\varphi$  是正交线性算子. 表达式  $f=t_a\varphi$  是唯一的, 这是因为由于  $\varphi(0)=0$ , 我们必须有  $a=f(0)$ , 从而  $\varphi=t_{-a}f$ . ■

为利用等距表达式  $t_a\varphi$ , 我们需要确定两个这样的表达式的(合成)积. 我们知道, 正交算子的合成是正交算子. 其他规则是:

**[6.2.8]**  $t_a t_b = t_{a+b}$ ,  $\varphi t_a = t_{a'}\varphi$ , 其中  $a' = \varphi(a)$

我们证明最后一个关系:  $\varphi t_a(x) = \varphi(x+a) = \varphi(x) + \varphi(a) = \varphi(x) + a' = t_{a'}\varphi(x)$ .

**[6.2.9] 推论**  $\mathbf{R}^n$  的所有等距的集合构成一个群, 记为  $M_n$ , 函数的合成作为它的合成法则.

**证明** 等距的合成是等距, 并且, 等距的逆也是等距, 因为正交算子与平移是可逆的, 且若  $f = t_a\varphi$ , 则  $f^{-1} = \varphi^{-1}t_a^{-1} = \varphi^{-1}t_{-a}$ . 这是等距的合成. ■

**注意** 直接从定义上不是很容易证明等距是可逆的.

同态  $M_n \rightarrow O_n$

存在一个重要映射  $\pi: M_n \rightarrow O_n$ , 它由去掉等距  $f$  的平移部分所定义. 把  $f$  写为(唯一)形式  $f = t_a\varphi$ , 定义  $\pi(f) = \varphi$ .

**[6.2.10] 命题** 映射  $\pi$  是满同态, 它的核是平移的集合  $T = \{t_v\}$ , 它是  $M_n$  的正规子群.

**证明** 显然,  $\pi$  是满射, 且我们一旦证明  $\pi$  是同态, 那么  $T$  显然是它的核, 因此,  $T$  是正规子群. 我们必须证明如果  $f$  与  $g$  是等距, 则  $\pi(fg) = \pi(f)\pi(g)$ . 比如说,  $f = t_a\varphi$ ,  $g = t_b\psi$ , 所以  $\pi(f) = \varphi$  且  $\pi(g) = \psi$ . 这样,  $\varphi t_b = t_{b'}\varphi$ , 其中  $b' = \varphi(b)$  并且,  $fg = t_a\varphi t_b\psi = t_{a+b'}\varphi\psi$ . 所以,  $\pi(fg) = \varphi\psi = \pi(f)\pi(g)$ . ■

## 坐标的变换

令  $P$  表示  $n$  维空间. 等距公式  $t_a\varphi$  依赖于坐标的选取, 所以, 我们想知道当坐标变化时, 公式如何变化. 我们将允许由正交矩阵引起的变化, 也允许通过平移所做的原点的平移. 换言之, 我们可用等距变换坐标.

要分析这样变换的效果, 我们从等距  $f$  (即  $P$  的点  $p$ ) 和它的像  $q = f(p)$  开始, 不涉及坐标. 当引进坐标系时, 空间  $P$  与  $\mathbf{R}^n$  就变得一致了, 点  $p$  与  $q$  有坐标, 比如说,  $x = (x_1, \dots, x_n)'$ ,  $y = (y_1, \dots, y_n)'$ . 还有, 等距  $f$  有坐标公式  $t_a\varphi$ ; 称这个公式为  $m$ . 方程  $q = f(p)$  平移到  $y = m(x) (= t_a\varphi(x))$ . 当坐标变化时, 我们想确定坐标向量与公式发生什么变化. 基在线性算子下变化的类似计算给出线索:  $m$  由共轭变换.

坐标变换由某个等距给出, 记之为  $\eta$  (等等). 令  $p$  与  $q$  的新坐标向量为  $x'$  与  $y'$ .  $f$  的新公式  $m'$  使得  $m'(x') = y'$ . 我们还有类似于基变化公式  $PX' = X$  (3.5.11) 的公式  $\eta(x') = x$ .

在方程  $m(x) = y$  里做替换  $\eta(x') = x$  与  $\eta(y') = y$ , 得  $m\eta(x') = \eta(y')$ , 或  $\eta^{-1}m\eta(x') = y'$ . 新公式是共轭的, 如所期望的那样:

**[6.2.11]** 
$$m' = \eta^{-1}m\eta$$

**[6.2.12] 推论** 当原点由平移变换时, 同态  $\pi: M_n \rightarrow O_n$  (6.2.10) 不变化.

当原点由平移  $t_v = \eta$  变换时, (6.2.11) 为  $m' = t_{-v}m t_v$ , 由于平移是  $\pi$  的核, 而  $\pi$  是同态, 故  $\pi(m') = \pi(m)$ .

## 方向

$\mathbf{R}^n$  上正交算子  $\varphi$  的行列式是  $\pm 1$ . 算子称为是保向的, 如果它的行列式是 1; 称为反向的, 如果它的行列式为  $-1$ . 类似地, 保向(或反向)的等距  $f$  的定义为, 当它写为形式  $f = t_a \varphi$  时, 算子  $\varphi$  是保向(或反向)的. 平面的等距是反向的, 如果它把平面的前后互换; 是保向的, 如果它把前面映为前面.

映射

$$\text{【6.2.13】} \quad \sigma: M_n \rightarrow \{\pm 1\}$$

是一个群同态, 它映保向等距为 1 且映反向等距为  $-1$ .

## 第三节 平面的等距

本节我们在代数和几何两方面描述平面的等距.

用  $M$  记平面的等距的群. 为计算这个群, 选取某些特殊的等距作为生成元, 并且得到它们间的关系. 这些关系有点儿类似于定义对称群  $S_3$  的那些关系, 但因  $M$  是无限的, 所以这其间关系更多.

选取坐标系, 用其将平面  $P$  与空间  $\mathbf{R}^2$  一致起来. 这样, 我们选取平移、绕原点的旋转与绕  $e_1$  轴的反射为生成元. 记转过角  $\theta$  的旋转为  $\rho_\theta$ , 关于  $e_1$  轴的反射为  $r$ . 这些都是线性算子, 它们的矩阵  $R$  与  $S_0$  在前面表示过(见(5.1.17)与(5.1.16)).

159

### 【6.3.1】

1. 由向量  $a$  确定的平移  $t_a$ :  $t_a(x) = x + a = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ .
2. 绕原点旋转角  $\theta$  的旋转  $\rho_\theta$ :  $\rho_\theta(x) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ .
3. 关于  $e_1$  轴的反射:  $r(x) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ .

我们没有列出所有的等距. 绕非原点的点旋转没有被列出, 关于其他直线的反射或者滑动反射也未被列出. 然而,  $M$  的每个元素都是这些等距的乘积, 于是, 它们生成这个群.

**【6.3.2】定理** 令  $m$  是平面的一个等距, 对唯一确定的向量  $v$  与角  $\theta$  ( $v$  与  $\theta$  可能为零), 有  $m = t_v \rho_\theta$ , 或  $m = t_v \rho_\theta r$ .

**证明** 推论 6.2.7 断言任意等距  $m$  可唯一地写成形式  $m = t_v \varphi$ , 其中  $\varphi$  是正交算子.  $\mathbf{R}^2$  上的正交线性算子是绕原点的旋转  $\rho_\theta$  以及关于过原点的直线的反射. 反射有形式  $\rho_\theta r$  (见(5.1.17)).

形如  $t_v \rho_\theta$  的等距保向, 而  $t_v \rho_\theta r$  反向.

$M$  里的计算可用符号  $t_v$ ,  $\rho_\theta$  和  $r$  进行, 使用下列规则合成它们. 这些规则可用

公式 6.3.1 来证明(见(6.2.8)).

$$\rho_\theta t_v = t_v \rho_\theta, \quad \text{其中 } v' = \rho_\theta(v),$$

$$r t_v = t_v r, \quad \text{其中 } v' = r(v),$$

**【6.3.3】**

$$r \rho_\theta = \rho_{-\theta} r,$$

$$t_v t_w = t_{v+w}, \quad \rho_\theta \rho_\gamma = \rho_{\theta+\gamma}, \quad r r = 1.$$

下一个定理从几何上刻画平面的等距.

**【6.3.4】定理** 平面的每个等距都有下面的形式之一:

(a) 保向等距:

(i) 平移: 映  $p \rightsquigarrow p+v$  的映射  $t_v$ .

(ii) 旋转: 绕某点转过角度  $\theta$  的平面旋转.

(b) 反向等距:

(i) 反射: 关于直线  $\ell$  双侧对称.

(ii) 滑动反射(或简称滑动): 关于直线  $\ell$  的反射, 然后经由平行于直线  $\ell$  的非零向量确定的平移.

这个著名定理的证明如下所述. 它的推论之一是, 绕两不同点的旋转的合成是绕第三个点的旋转, 除非它是平移. 这不是显然的, 但由这个定理可得, 因为合成会保持方向.

160

一些合成容易可视化. 绕同一点转过两个角  $\alpha$  与  $\beta$  的旋转的合成是绕该点的旋转, 转角为  $\alpha+\beta$ . 由向量  $a$  与  $b$  确定的平移的合成是这两个向量和  $a+b$  所确定的平移.

关于两条非平行线  $\ell_1$  与  $\ell_2$  的反射是绕交点  $\ell_1 \cap \ell_2$  的旋转. 这也可由这个定理得到, 因为合成是保向的, 且它固定点  $p$ . 关于平行线的反射的合成是由一个与这两条直线正交的向量所确定的平移.

**定理 6.3.4 的证明** 首先考虑保向等距. 令  $f$  是一个保向的但不是平移的等距. 我们必须证明  $f$  是绕某个点的旋转. 选取坐标把  $f$  的公式写为如(6.3.3)里的  $m = t_v \rho_\theta$ . 因为  $m$  不是平移, 故  $\theta \neq 0$ .

**【6.3.5】引理** 具有形式  $m = t_v \rho_\theta$  的等距  $f$  (其中  $\theta \neq 0$ ) 是围绕平面上一点转过角度  $\theta$  的旋转.

**证明** 为简化记号, 记  $\rho_\theta$  为  $\rho$ . 要证  $f$  表示围绕平面上某一点  $p$  转过角度  $\theta$  的旋转, 我们用平移  $t_p$  变换坐标. 希望选取  $p$  使得等距  $f$  的新公式变为  $m' = \rho$ . 如果这样, 则  $f$  是围绕平面上点  $p$  转过角度  $\theta$  的旋转.

坐标变换规则是  $t_p(x') = x$ , 所以,  $f$  的新公式变为  $m' = t_p^{-1} m t_p = t_{-p} t_v \rho t_p$  (6.2.11). 我们使用规则(6.3.3):  $\rho t_p = t_p \rho$ , 其中  $p' = \rho(p)$ . 这样, 如果  $b = -p + a + p' = a + \rho(p) - p$ , 则有  $m' = t_b \rho$ . 我们希望选取  $p$  使得  $b=0$ .

令  $I$  表示恒等算子, 设  $c = \cos\theta$  与  $s = \sin\theta$ . 线性算子  $I - \rho$  的矩阵是

**【6.3.6】**

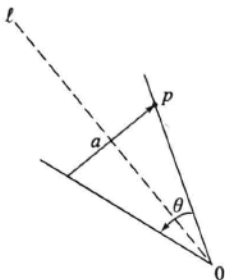
$$\begin{bmatrix} 1-c & s \\ -s & 1-c \end{bmatrix}$$

它的行列式是  $2-2c=2-2\cos\theta$ . 这个行列式不为零, 除非  $\cos\theta=1$ , 而这仅当  $\theta=0$  时才发生. 由于  $\theta\neq 0$ , 方程  $(I-\rho)p=a$  对于  $p$  有唯一的解. 当需要时, 方程可有明确解. ■

点  $p$  是等距  $t_a\rho_\theta$  的固定点, 可用如下所示的几何方法找到它. 直线  $\ell$  通过原点, 并与向量  $a$  垂直. 取适当位置使夹角为  $\theta$  的扇形能被  $\ell$  平分, 固定点  $p$  可以如图所示通过将向量  $a$  插入扇形得到.

161

【6.3.7】图

等距  $t_a\rho_\theta$  的固定点

要完成定理 6.3.4 的证明, 我们证明反向等距  $m=t_a\rho_\theta$  是滑动或反射. 为此, 我们变换坐标. 等距  $\rho_\theta$  是关于过原点的直线  $\ell_0$  的反射. 我们可旋转坐标使得  $\ell_0$  是水平轴. 在新的坐标系里, 反射变成了我们的标准反射  $r$ , 并且平移  $t_a$  仍是平移, 尽管向量  $a$  的坐标变了. 用同一符号  $a$  表示这个新向量. 在新坐标系里, 等距变为  $m=t_a r$ . 它的作用为

$$m \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = t_a \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix} = \begin{bmatrix} x_1 + a_1 \\ -x_2 + a_2 \end{bmatrix}$$

这个等距是关于直线  $\ell: \left\{ x_2 = \frac{1}{2}a_2 \right\}$  反射所得的滑动, 而该直线是由向量  $a_1 e_1$  的平移得到的. 如果  $a_1=0$ , 则  $m$  是反射.

这就完成了定理 6.3.4 的证明. ■

**【6.3.8】推论** 等距  $t_a\rho_\theta$  的滑动直线与反射  $\rho_\theta$  的直线平行.

固定原点的等距是正交线性算子, 所以, 当选取坐标时, 正交群  $O_2$  成为等距群  $M$  的子群. 我们也可考虑固定平面上原点以外的点的等距群  $M$  的子群. 这个群与正交群的关系由下一个命题给出.

**【6.3.9】命题** 假设平面上已选取坐标, 使得正交群  $O_2$  成为固定原点的等距群  $M$  的子群. 则固定平面上的点  $p$  的等距群是共轭子群  $t_p O_2 t_p^{-1}$ .

**证明** 如果等距  $m$  固定  $p$ , 则  $t_p^{-1} m t_p$  固定原点:  $t_p^{-1} m t_p o = t_p^{-1} m p = t_p^{-1} p = o$ . 反之, 如果  $m$  固定  $o$ , 则  $t_p m t_p^{-1}$  固定  $p$ . ■

162

可以以下面的方式可视化绕点  $p$  的旋转: 首先由平移  $t_{-p}$  移动  $p$  到原点, 然后绕原点旋转, 最后再平移到  $p$ .

我们回到在 (6.2.10) 里定义的同态  $\pi: M \rightarrow O_2$ . 上面的讨论证明了:

**【6.3.10】命题** 令  $p$  是平面的一个点, 且设  $\rho_{\theta,p}$  表示围绕  $p$  转过角度  $\theta$  的旋转, 则  $\pi(\rho_{\theta,p}) = \rho_{\theta}$ . 类似地, 如果  $r_{\ell}$  是关于直线  $\ell$  的反射或具有与  $x$  轴平行的滑动直线  $\ell$  的滑动, 则  $\pi(r_{\ell}) = r$ .

## 点与向量

在本书大部分篇幅里, 没有必要区分平面  $P = \mathbf{R}^2$  的点  $p$  与从原点  $o$  到  $p$  的向量, 这个向量在微积分书里常常写为  $\vec{op}$ . 然而, 当利用等距时, 最好是保持这种区分. 于是, 我们引进平面的另一个复制, 称之为  $V$ , 并且把它的元素看作平移向量. 由  $V$  中的向量  $v$  确定的平移作用  $P$  的点  $p$  为  $t_v(p) = p + v$ . 它用  $v$  平移平面上的每个点.

$V$  与  $P$  都是平面. 它们之间的不同仅当变换坐标时变得明显. 假设在  $P$  中用平移变换  $\eta = t_w$  平移坐标. 则变换坐标的规则是  $\eta(p') = p$ , 或  $p' + w = p$ . 同时, 等距  $m$  变为  $m' = \eta^{-1} m \eta = t_{-w} m t_w$  (6.2.11). 如果用  $m = t_v$  应用这个规则, 则  $m' = t_{-w} t_v t_w = t_v$ .  $P$  的点得到新的坐标, 但平移向量未变.

另一方面, 如果用正交算子  $\varphi$  变换坐标, 则  $\varphi(p') = p$ , 且如果  $m = t_v$ , 则  $m' = \varphi^{-1} t_v \varphi = t_{v'}$ , 其中  $v' = \varphi^{-1} v$ . 所以,  $\varphi v' = v$ . 用正交算子变换坐标的效果在  $P$  上及  $V$  上是一样的.

$P$  与  $V$  的仅有差别是  $P$  中的原点不一定是固定不动的, 而零向量在  $V$  被选作原点. 正交算子作用于  $V$ , 但它们不作用于  $P$ , 除非在  $P$  上选择了原点.

## 第四节 平面上正交算子的有限群

**【6.4.1】定理** 令  $G$  是正交群  $O_2$  的有限子群. 存在整数  $n$  使得  $G$  是下列群之一:

(a)  $C_n$ : 由旋转  $\rho_{\theta}$  生成的秩为  $n$  的循环群, 其中  $\theta = 2\pi/n$ .

(b)  $D_n$ : 由两个元素(旋转  $\rho_{\theta}$  与关于过原点的直线  $\ell$  的反射  $r'$ )生成的秩为  $2n$  的二面体群, 其中  $\theta = 2\pi/n$ .

在证明定理前我们用点时间描述下二面体群  $D_n$ . 这个群依赖于反射的直线, 但如果选取坐标使得  $\ell$  成为水平轴, 则该群将含有标准反射  $r$ , 其矩阵为

$$\mathbf{【6.4.2】} \quad \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$$

163

这样, 如果把  $\rho_{\theta}$  写成  $\rho$ , 则该群的  $2n$  个元素将是  $\rho$  的  $n$  个幂  $\rho^i$  与  $n$  个积  $\rho^i r$ .  $\rho$  与  $r$  的交换规则是

$$r_{\rho} = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} = \rho^{-1} r$$

其中  $c = \cos\theta$ ,  $s = \sin\theta$ ,  $\theta = 2\pi/n$ .

为与群的习惯记号一致, 记旋转  $\rho_{2\pi/n}$  为  $x$ , 反射  $r$  为  $y$ .

**【6.4.3】命题** 二面体群  $D_n$  有秩  $2n$ . 它由两个满足下列关系

$$x^n = 1, \quad y^2 = 1, \quad yx = x^{-1}y$$

的元素  $x$  与  $y$  生成.  $D_n$  的元素为

$$1, x, x^2, \dots, x^{n-1}; \quad y, xy, x^2y, \dots, x^{n-1}y$$

利用(6.4.3)里的前两个关系,第三个关系可写成各式各样的形式.它等价于

$$\text{【6.4.4】} \quad xyxy = 1, \quad \text{也等价于} \quad yx = x^{n-1}y.$$

当  $n=3$  时,关系与对称群  $S_3$ (2.2.6)的一样.

**【6.4.5】推论** 二面体群  $D_3$  与对称群  $S_3$  是同构的.

对  $n>3$ ,二面体群与对称群是不同构的,因为  $D_n$  秩为  $2n$ ,而  $S_n$  秩为  $n!$ .

当  $n \geq 3$  时,二面体群  $D_n$  的元素是把正  $n$  边形  $\Delta$  映为自身的正交算子—— $\Delta$  的对称群.这很容易看出,易可由定理得出:绕中心转过角度  $2\pi/n$  的旋转与一些反射把正  $n$  边形映为自身.定理 6.4.1 把所有对称的群与  $D_n$  等同起来.

二面体群  $D_1, D_2$  太小以至于不是通常意义下  $n$  边形的对称群.  $D_1$  是两个元素的群  $\{1, r\}$ . 所以,它是循环群,像  $C_2$  一样. 但  $D_1$  的元素  $r$  是反射,而  $C_2$  中不同于恒等元的元素是具有角度  $\pi$  的旋转. 群  $D_2$  含有 4 个元素  $\{1, \rho, r, \rho r\}$ , 其中  $\rho$  是具有角度  $\pi$  的旋转,  $\rho r$  是关于竖直轴的反射. 这个群同构于克莱因四元群.

如果我们喜欢,可把  $D_1$  与  $D_2$  看作 1 边与 2 边的对称群:



164

我们现在开始证明定理 6.4.1. 实数加群  $\mathbf{R}^+$  的子群  $\Gamma$  称为离散的, 如果存在(小)正实数  $\epsilon$  使得  $\Gamma$  的每个非零元素有绝对值  $\geq \epsilon$ .

**【6.4.6】引理** 令  $\Gamma$  是  $\mathbf{R}^+$  的离散子群. 则或者  $\Gamma = \{0\}$  或者  $\Gamma$  是正实数  $a$  的整数倍集合  $\mathbf{Z}a$ .

**证明** 与定理 2.3.3 的证明很相似,  $\mathbf{Z}^+$  的非零子群有形式  $\mathbf{Z}n$ .

如果  $a$  与  $b$  是  $\Gamma$  的不同元素, 那么由于  $\Gamma$  是群, 故  $a-b$  在  $\Gamma$  里, 且  $|a-b| \geq \epsilon$ .  $\Gamma$  的不同元由至少为  $\epsilon$  的距离分开. 因为只有有限多个由  $\epsilon$  分开的元素可放进任一有界区间, 故有界区间含有  $\Gamma$  的有限多个元素.

假设  $\Gamma \neq \{0\}$ . 则  $\Gamma$  含有非零元素  $b$ , 且因为它是一个群, 故  $\Gamma$  也含有  $-b$ . 所以, 它含有正元素, 比如说  $a'$ . 在  $\Gamma$  中选取最小正元素  $a$ . 能够这样做是因为只需在区间  $0 \leq x \leq a'$  里选取  $\Gamma$  的有限子集的最小元素.

我们证明  $\Gamma = \mathbf{Z}a$ . 因为  $a$  在  $\Gamma$  中, 且  $\Gamma$  是群, 故  $\mathbf{Z}a \subset \Gamma$ . 令  $b$  是  $\Gamma$  的元素, 则对某个实数  $r$ , 有  $b = ra$ . 取出  $r$  的整数部分, 写为  $r = m + r_0$ , 其中  $m$  是整数且  $0 \leq r_0 < 1$ . 因为  $\Gamma$  是群, 故  $b' = b - ma$  在  $\Gamma$  中且  $b' = r_0 a$ . 于是,  $0 \leq b' < a$ . 因为  $a$  是  $\Gamma$  中最小正元素, 故  $b'$  一定为零. 所以,  $b = ma$ , 它包含在  $\mathbf{Z}a$  里. 这就证明了  $\Gamma \subset \mathbf{Z}a$ , 所以  $\Gamma = \mathbf{Z}a$ . ■

**定理 6.4.1 的证明** 令  $G$  是  $O_2$  的有限群. 我们要证  $G$  是  $C_n$  或  $D_n$ . 记住  $O_2$  的元素是旋转  $\rho_\theta$  和反射  $\rho_\theta r$ .

情形 1:  $G$  的所有元素是旋转.

我们必须证明  $G$  是循环的. 令  $\Gamma$  是使得  $\rho_\alpha$  在  $G$  里的实数  $\alpha$  的集合, 则  $\Gamma$  是加群  $\mathbf{R}^+$  的子群, 且它含有  $2\pi$ . 因为  $G$  是有限的, 故  $\Gamma$  是离散的. 所以,  $\Gamma$  有形式  $\mathbf{Z}a$ . 这样,  $G$  由转过角  $\alpha$  的整数倍角的旋转组成. 由于  $2\pi$  在  $\Gamma$  中, 故它是  $a$  的整数倍. 所以, 对某个整数

$n, \alpha=2\pi/n$ , 且  $G=C_n$ .

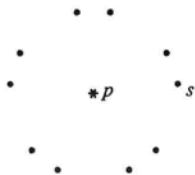
情形 2:  $G$  含有反射.

我们调整坐标使得标准反射  $r$  在  $G$  中. 令  $H$  表示由属于  $G$  的旋转组成的子群. 应用在情形 1 里所证明的结论得到  $H$  是由  $\rho_\theta$  生成的循环群, 其中  $\theta=2\pi/n$ . 这样, 对  $0 \leq k < n-1$ ,  $2n$  个积  $\rho_\theta^k$  和反射  $\rho_\theta^k r$  在  $G$  中, 故  $G$  含有二面体群  $D_n$ . 我们断言  $G=D_n$ . 为证明这个断言, 取  $G$  的任意元素  $g$ , 则  $g$  或是旋转, 或是反射. 如果  $g$  是旋转, 则由  $H$  的定义,  $g \in H$ .  $H$  的元素也在  $D_n$  中, 从而  $g \in D_n$ . 如果  $g$  是反射, 则对某个  $\rho_\alpha$ , 把它写为形式  $\rho_\alpha r$ . 因为  $r \in G$ , 故有乘积  $gr = \rho_\alpha$ . 所以,  $\rho_\alpha$  是  $\rho_\theta$  的幂, 于是  $g \in D_n$ . ■

165

**【6.4.7】定理(不动点定理)** 设  $G$  是平面的等距的有限群, 则平面上存在一个点, 它在  $G$  的每个元素作用之下不动, 即存在点  $p$  使得对所有  $g$  属于  $G$  有  $g(p)=p$ .

**证明** 这是一个漂亮的几何证明. 设  $s$  是平面上的任意点, 并设  $S$  是在  $G$  中各个等距作用下  $s$  的像点的集合. 因对  $G$  中的某个  $g$ ,  $S$  的每个元素  $s'$  有形式  $s'=g(s)$ . 这个集合称为  $s$  在  $G$  作用下的轨道. 元素  $s$  属于轨道, 因为单位元  $1$  在  $G$  中, 且  $s=1(s)$ . 一个典型的轨道与运算的不动点  $p$  如下图所示, 此时  $G$  是正五边形对称群.



群  $G$  的任意元素都将置换轨道  $S$ . 换言之, 若  $s' \in S$  且  $h \in G$ , 则  $h(s') \in S$ . 比如  $s'=g(s)$ , 其中  $g \in G$ . 因为  $G$  是群, 故  $hg \in G$ . 于是  $hg(s) \in S$ , 且等于  $h(s')$ . ■

我们任意排列  $S$  的元素, 记  $S = \{s_1, \dots, s_n\}$ . 所求的不动点是轨道的重心或轨道的重力中心, 定义为

$$\text{【6.4.8】} \quad p = \frac{1}{n}(s_1 + \dots + s_n)$$

其中右边可以在平面上的任一坐标系下用向量加法计算.

**【6.4.9】引理** 等距把重心映为重心: 令  $S = \{s_1, \dots, s_n\}$  是平面的有限点集, 设  $p$  是它的重心, 如同(6.4.8)定义的. 设  $m$  是等距. 设  $m(p) = p'$  与  $m(s_i) = s'_i$ , 则  $p'$  是集合  $S' = \{s'_1, \dots, s'_n\}$  的重心.

集合  $S$  的重心是不动点的事实由此即得.  $G$  的元素  $g$  置换轨道  $S$ . 它映  $S$  到  $S$ , 所以, 它映  $p$  到  $p$ .

**引理 6.4.9 的证明** 可由物理上的推理得以证明, 也可通过几何方法证明. 为此, 只要分别处理  $m=t_a$  与  $m=\varphi$  的情形即可, 其中  $\varphi$  是正交算子. 任何一个等距可以由这些等距合得到.

情形 1:  $m=t_a$  是平移. 这样,  $s'_i = s_i + a$ ,  $p' = p + a$ . 于是, 下式成立:



$$p' = p + a = \frac{1}{n}((s_1 + a) + \cdots + (s_n + a)) = \frac{1}{n}(s'_1 + \cdots + s'_n)$$

情形 2:  $m = \varphi$  是线性算子. 这样

166

$$p' = \varphi(p) = \varphi\left(\frac{1}{n}(s_1 + \cdots + s_n)\right) = \frac{1}{n}(\varphi(s_1) + \cdots + \varphi(s_n)) = \frac{1}{n}(s'_1 + \cdots + s'_n) \quad \blacksquare$$

把定理 6.4.1 与定理 6.4.7 合起来, 得到平面上有界图形对称群的描述.

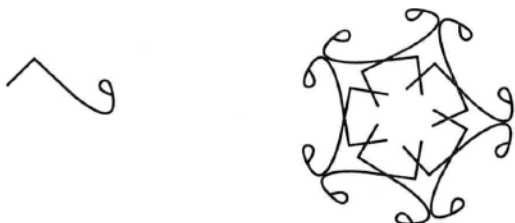
**【6.4.10】推论** 令  $G$  是平面的等距群  $M$  的有限子群. 如果坐标选取得合适, 则  $G$  为定理 6.4.1 中描述的群  $C_n$  或  $D_n$  之一.

## 第五节 离散等距群

本节讨论诸如图 6.1.5 中所示的无界图形的对称群. 我所称的万花筒原理可用来构造具有已知对称群的图形. 你也许看过万花筒. 人们在万花筒的底端看见一个扇形, 其边由两个镜片分界, 镜片以角度  $\theta$  (如  $\theta = \pi/6$ ) 放置. 人们还看见每个镜片里扇形的反射, 然后看见反射的反射, 等等. 通常在扇形里有一点玻璃的颜色, 其反射形成图案.

这其中涉及一个群. 在万花筒底端的平面上, 设  $\ell_1$  与  $\ell_2$  是由镜片构成的扇形的分界直线. 这个群是二面体群, 由关于  $\ell_i$  的反射  $r_i$  生成. 这些反射之积  $r_1 r_2$  保向, 且固定两直线交点不动, 于是它是一个旋转. 它的旋转角是  $\pm 2\theta$ .

可应用同一原理到  $M$  的任意子群  $G$ . 我们不准给出精确的推理来证明这一点, 但讨论的方法可以演变成一个证明. 从平面上一个随机的图形  $R$  开始. 群  $G$  中的每一个元素  $g$  都会将  $R$  移到一个不同的位置, 称为  $gR$ . 图形  $F$  是所有图形  $gR$  的并集.  $G$  的一个元素  $h$  将  $gR$  映到  $hgR$ , 它也是  $F$  的一部分, 于是, 它映  $F$  到自身. 如果  $R$  充分地随机, 则  $G$  将是  $F$  的对称群. 正如我们从万花筒所知道的, 图形  $F$  常常是非常引人入胜的. 下面是当  $G$  是正五边形的对称群时应用这一方法所得的结果.



当然, 许多图形都会有相同或类似的对称群. 尽管如此, 描述这样的群仍是很有意思并具有指导意义的. 我们将讨论群的大致分类, 在练习中将会对它们加以改进.

167

$M$  的一些子群没有合理的几何意义. 例如, 如果在万花筒里的镜片置放角度  $\theta$  不是  $2\pi$  的有理数倍, 则扇形有无限多不同的反射. 我们需要把这种可能排除在外.

**【6.5.1】定义** 平面  $P$  的等距群  $G$  是离散的, 如果它不包含任意小的平移或旋转. 更准确地说,  $G$  是离散的, 如果存在一个正实数  $\epsilon$  使得

(i) 如果  $G$  中元素是由非零向量  $a$  产生的平移, 则  $a$  的长度至少为  $\epsilon$ :  $|a| \geq \epsilon$ ;

(ii) 如果  $G$  中元素是围绕平面上某点转过非零角度  $\theta$  的旋转, 则角度  $\theta$  的绝对值至少为  $\theta: |\theta| \geq \epsilon$ .

**注意** 因为平移向量和旋转角度构成不同的集合, 所以对它们分别设置下界也许更合适. 然而, 在这个定义里, 我们不关心向量和角度的最佳下界, 所以选取  $\epsilon$  足够小以期同时照顾到它们.

平移和旋转都是保向等距(6.3.4), 这些条件亦应用于它们的所有元素. 我们没对反向等距附加条件. 如果  $m$  是具有非零滑动向量  $v$  的滑动, 则  $m^2$  是平移  $t_{2v}$ . 所以, 平移向量的下界也决定了滑动向量的下界.

分析离散群  $G$  有三个主要工具:

**【6.5.2】**

- 平移群  $L$ , 平移向量的群  $V$  的子群.
- 点群  $\bar{G}$ , 正交群  $O_2$  的子群.
- $\bar{G}$  在  $L$  上的作用.

**平移群**

$G$  的平移群  $L$  是向量  $v$  的集合, 使得平移  $t_v$  属于  $G$ .

**【6.5.3】**

$$L = \{v \in V \mid t_v \in G\}$$

因为  $t_v t_w = t_{v+w}$ ,  $t_v^{-1} = t_{-v}$ , 故  $L$  是所有平移向量的加群  $V^+$  的子群.  $G$  中平移的下界  $\epsilon$  界定了  $L$  中向量的长度:

**【6.5.4】**

$L$  中每个非零向量  $v$  的长度  $|v| \geq \epsilon$ .

**注** 加群  $V^+$  或  $\mathbf{R}^n$  的对某个  $\epsilon > 0$  满足条件(6.5.4)的子群  $L$  称为离散子群. (这个定义以前对  $\mathbf{R}^+$  定义过.)

子群  $L$  是离散的当且仅当  $L$  的不同向量  $a$  和  $b$  的间距至少是  $\epsilon$ . 距离是  $b-a$  的长度, 因为  $L$  是子群, 有  $b-a$  属于  $L$ . 所以, 这个结论是正确的. 如果(6.5.4)成立, 则  $|b-a| \geq \epsilon$ .

**【6.5.5】定理**  $V^+$  或  $\mathbf{R}^{2+}$  的每个离散子群具有下列形式之一:

(a) 零群:  $L = \{0\}$ .

(b) 一个非零向量  $a$  的整数倍的集合:

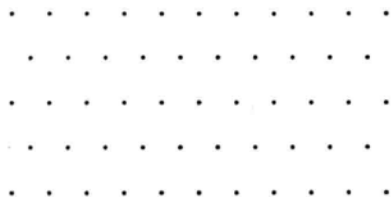
$$L = \mathbf{Z}a = \{ma \mid m \in \mathbf{Z}\}, \text{ 或}$$

(c) 两个线性无关的向量  $a$  和  $b$  的整数组组合的集合:

$$L = \mathbf{Z}a + \mathbf{Z}b = \{ma + nb \mid m, n \in \mathbf{Z}\}$$

上面所列的第三类群称为格, 生成集合  $(a, b)$  称为格基.

**【6.5.6】图**



格

**【6.5.7】引理** 令  $L$  是  $V^+$  或  $\mathbf{R}^{2+}$  的离散子群.

(a) 平面的有界区域仅含有  $L$  的有限多个点.

(b) 如果  $L$  不是平凡群, 则它含有极小长度的非零向量.

**证明**

(a) 因为  $L$  的元素由至少为  $\epsilon$  的长度分离, 故小方形区域至多可含  $L$  的一个点. 平面区域是有界的, 如果它包含在某个大的矩形内. 用有限多个小方形可覆盖任意矩形, 每个小方形至多含有  $L$  的一个点.

(b) 我们说一个向量  $v$  是  $L$  的极小长度的非零向量, 如果  $L$  不含有更短长度的非零向量. 要证这样的向量存在, 我们利用  $L$  不是平凡群的假设.  $L$  中存在某个非零向量  $a$ , 则关于原点半径为  $|a|$  的圆盘是含有  $a$  和  $L$  的有限多个非零点的有界区域. 这些点中的某一个将有极小长度. ■

已知  $\mathbf{R}^2$  的基  $B=(u, w)$ , 令  $\Pi(B)$  表示具有顶点  $o, u, w, u+w$  的平行四边形. 它由线性组合  $ru+sw$  组成, 其中  $0 \leq r \leq 1$  与  $0 \leq s \leq 1$ . 用  $\Pi'(B)$  表示从  $\Pi(B)$  去掉两条边  $[u, u+w]$  与  $[w, u+w]$  得到的区域. 它由线性组合  $ru+sw$  组成, 其中  $0 \leq r < 1, 0 \leq s < 1$ .

**【6.5.8】引理** 令  $B=(u, w)$  是  $\mathbf{R}^2$  的基, 设  $L$  是  $B$  的整数组合的格.  $\mathbf{R}^2$  的每个向量  $v$  可唯一写成  $v=x+v_0$  的形式, 其中  $x$  属于  $L$ ,  $v_0$  属于  $\Pi'(B)$ .

169

**证明** 因为  $B$  是基, 故每个向量是线性组合  $ru+sw$ , 其中  $r$  与  $s$  为实系数. 取出它们的整数部分, 写作  $r=m+r_0$  与  $s=n+s_0$ ,  $m$  与  $n$  为整数, 且  $0 \leq r_0, s_0 < 1$ . 这样,  $v=x+v_0$ , 其中  $x=mu+nv$  属于  $L$ ,  $v_0=r_0u+s_0w$  属于  $\Pi'(B)$ . 只有一种方法写成这样. ■

**定理 6.5.5 的证明** 考虑  $\mathbf{R}^{2+}$  的离散子群  $L$  就够了.  $L$  为零群的情形包括在列表里. 如果  $L \neq \{0\}$ , 则有两种可能情形:

情形 1:  $L$  中所有向量位于一条过原点的直线  $\ell$  之上.

这样,  $L$  是与  $\mathbf{R}^+$  同构的加群  $\ell^+$  的子群. 引理 6.4.6 表明  $L$  有形式  $\mathbf{Z}a$ .

情形 2:  $L$  的元素不在一条直线上.

在这种情形下,  $L$  含有无关向量  $a'$  与  $b'$ . 于是,  $B'=(a', b')$  是  $\mathbf{R}^2$  的基. 我们必须证明  $L$  存在格基.

首先, 考虑由  $a'$  所张成的直线  $\ell$ .  $\ell^+$  的子群  $L \cap \ell$  是离散的, 且  $a'$  不是 0. 所以, 由情形 1 所证的, 对某个向量  $a$ ,  $L$  有形式  $\mathbf{Z}a$ . 适当调整坐标使  $a$  成为向量  $(1, 0)'$ .

其次, 如果有必要, 我们用  $-b'$  替换  $b'=(b'_1, b'_2)'$  使得  $b'_2$  为正的. 在  $L$  中寻找向量  $b=(b_1, b_2)'$  使得  $b_2$  为正的, 否则, 使  $b_2$  尽可能地小. 于是, 我们有无限多个元素要检查. 然而, 由于  $b'$  属于  $L$ , 故仅需检查使得  $0 < b_2 \leq b'_2$  的元素  $b$ . 而且, 可以加  $a$  的倍数到  $b$ , 所以, 可假设  $0 < b_1 \leq 1$ . 当完成这步以后,  $b$  将在含有  $L$  的有限多个点的有界区域里. 通过这个有限集寻找所求元素  $b$ , 我们证明  $B=(a, b)$  是  $L$  的格基.

令  $\tilde{L}=\mathbf{Z}a+\mathbf{Z}b$ . 于是  $\tilde{L} \subset L$ . 我们必须证明  $L$  的某个元素属于  $\tilde{L}$ . 根据引理 6.5.8, 应用于格  $\tilde{L}$ , 只要证明  $L$  在  $\Pi'(B)$  中仅有的元素是零向量就够了. 令  $c=(c_1, c_2)'$  是  $L$  在那个区域中的点, 使得  $0 \leq c_1 < 1$  且  $0 \leq c_2 < b_2$ . 因为  $b_2$  选为极小的, 故  $c_2=0$ , 且  $c$  在直线  $\ell$  上. 这样,  $c$  是  $a$  的倍数. 由于  $0 \leq c_1 < 1$ , 所以,  $c=0$ . ■

## 点群

我们现在转到分析等距的离散群的第二个工具. 选取坐标, 回到同态  $\pi: M \rightarrow O_2$ , 其核为平移群  $T(6.3.10)$ . 当限制这个同态到离散子群  $G$  时, 我们得到同态

$$\text{【6.5.9】} \quad \pi|_G: G \rightarrow O_2$$

点群  $\bar{G}$  是  $G$  在正交群  $O_2$  里的像.

在群  $G$  的元素与它的点群  $\bar{G}$  的元素之间做个清晰的区分是重要的. 所以, 为避免混淆, 当符号代表  $\bar{G}$  的元素时, 在其上面加一个横杠. 对于中  $G$  的  $g$ ,  $\bar{g}$  表示正交算子.

由定义, 如果  $G$  中含有某个形如  $t_a \rho_\theta$  的元素, 则旋转  $\bar{\rho}_\theta$  属于  $\bar{G}$ , 并且这是围绕平面上某个点转过角度  $\theta$  的旋转(6.3.5).  $\bar{G}$  的元素  $\bar{\rho}_\theta$  在  $G$  中的原像由所有  $G$  的这样的元素组成, 即围绕平面上某个点转过角度  $\theta$  的旋转. 170

类似地, 设  $\ell$  表示  $\rho_\theta r$  的反射轴的直线. 如我们前面已注意到的, 它与  $e_1$  轴的夹角是  $\frac{1}{2}\theta$ (5.1.17). 如果  $G$  中包含某个元素  $t_a \rho_\theta r$ , 则点群  $\bar{G}$  包含  $\bar{\rho}_\theta r$ , 而  $t_a \rho_\theta r$  是沿某条与  $\ell$  平行的直线的反射或滑动反射(6.3.5).  $\bar{\rho}_\theta r$  的原像由  $G$  中所有这样的元素组成, 即它们是沿某条与  $\ell$  平行的直线的反射或滑动反射. 总结如下:

点群  $\bar{G}$  记录  $G$  中元素的旋转的角度、滑动直线的斜率与反射的直线轴.

**【6.5.10】命题**  $O_2$  的离散子群  $\bar{G}$  是有限的, 所以, 要么是循环群要么是二面体群.

**证明** 因为  $\bar{G}$  不含有小旋转, 所以使得  $\bar{\rho}_\theta$  属于  $\bar{G}$  的实数  $\theta$  的集合  $\Gamma$  是加群  $\mathbf{R}^+$  的含有  $2\pi$  的离散子群. 由引理 6.4.6 知,  $\Gamma$  有形式  $\mathbf{Z}\theta$ , 其中对某个整数  $n$ , 有  $\theta = 2\pi/n$ . 在此, 可照搬定理 6.4.1 的证明. ■

## 晶体限制

如果等距的离散群的平移群  $G$  是平凡群, 则  $\pi$  到  $G$  的限制是单射. 在此情形下,  $G$  与它的点群  $\bar{G}$  同构, 为循环群或二面体群. 下一个命题是分析无限离散群的第三种工具. 它将点群与平移群联系起来.

除非已经选定原点, 否则正交矩阵  $Q_2$  不作用在平面  $P$  上. 但它确实作用在平移向量空间  $V$  上.

**【6.5.11】命题** 令  $G$  是  $M$  的离散子群. 设  $a$  是它的平移群  $L$  的元素, 且设  $\bar{g}$  是它的点群  $\bar{G}$  的元素, 则  $\bar{g}(a)$  属于  $L$ .

我们可重述命题为:  $\bar{G}$  的元素映  $L$  到自身. 所以, 当视  $L$  为平面  $V$  的一个图形时,  $\bar{G}$  包含在  $L$  的对称群里.

**命题 6.5.11 的证明** 令  $a$  和  $g$  分别是  $L$  与  $G$  的元素, 设  $\bar{g}$  是  $g$  在  $\bar{G}$  中的像, 且设  $a' = \bar{g}(a)$ . 我们将证明  $t_{a'}$  是共轭  $gt_a g^{-1}$ . 这将表明  $t_{a'}$  属于  $G$ , 从而  $a'$  属于  $L$ . 我们写  $g = t_b \varphi$ . 这样,  $\varphi$  属于  $O_2$ , 且  $\bar{g} = \bar{\varphi}$ . 于是,  $a' = \bar{\varphi}(a)$ . 利用公式(6.2.8), 我们有:

$$gt_a g^{-1} = (t_b \varphi) t_a (\varphi^{-1} t_{-b}) = t_b t_{a'} \varphi \varphi^{-1} t_{-b} = t_{a'}$$
■

注意 理解群  $G$  不作用在它的平移群  $L$  上是重要的. 的确, 问  $G$  是否作用在  $L$  上是没有意义的, 因为  $G$  的元素是平面  $P$  的等距, 而  $L$  是  $V$  的子集. 除非原点固定不动, 否则  $P$  没有意义. 我们也许会问:  $P$  中是否存在一点使之成为原点,  $G$  的元素映  $L$  到自身吗? 答案为有时对, 有时不对. 结论依赖于群.

171

下一定理描述了当平移群  $L$  不是平凡群时所出现的点群.

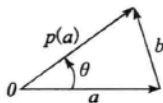
**[6.5.12] 定理 (晶体限制)** 令  $L$  是  $V^+$  或  $\mathbf{R}^{2+}$  的离散子群, 且设  $H \subset O_2$  是  $L$  的对称群的子群. 假设  $L$  不是平凡群, 则

(a)  $H$  的每个旋转的阶是 1, 2, 3, 4 或 6;

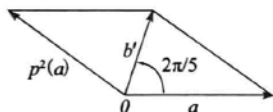
(b)  $H$  为群  $C_n$  或  $D_n$  之一, 其中  $n=1, 2, 3, 4$  或 6.

特别地, 阶为 5 的旋转排除在外. 不存在五重旋转对称的墙纸图案. (然而, 的确存在有五重对称的“拟周期”的图案. 例如, 见 [Senechal].)

**晶体限制的证明** 我们证(a). (b)由(a)与定理 6.4.1 可得. 设  $\rho$  是  $H$  中具有角度  $\theta$  的旋转, 并设  $a$  是  $L$  中长度极小的非零向量. 由于  $H$  在  $L$  上作用, 故  $\rho(a)$  属于  $L$ . 于是  $b = \rho(a) - a$  亦属于  $L$ ; 因为  $a$  具有极小长度, 故  $|b| \geq |a|$ . 如下图所示, 当  $\theta < 2\pi/6$  时,  $|b| < |a|$ , 由此必有  $\theta \geq 2\pi/6$ . 于是, 群  $H$  是离散的, 从而是有限的. 因此,  $\rho$  的阶  $\leq 6$ .



$\theta = 2\pi/5$  的情形也可排除在外, 因为对于这个角度, 元素  $b' = \rho^2(a) + a$  比  $a$  短:



## 第六节 平面晶体群

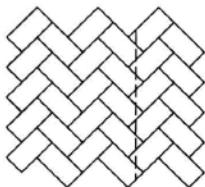
回到等距的离散群  $G \subset M$ . 我们已经见到, 当  $L$  是平凡群时,  $G$  是循环群或二面体群. 使得  $L$  是无限循环群(6.5.5)(b)的离散群  $G$  是如图 6.1.3 及图 6.1.4 所示的带状图案的对称群. 这些群的分类留作练习.

当  $L$  是格时,  $G$  称为二维晶体群. 这些晶体群是二维晶体(诸如石墨)的对称群. 我们想象一个晶体是无限大的. 这样, 分子规则排列的事实反映了其对称群总是包含两个无关平移. 墙纸图案在两个不同的方向上重复——一次是沿着纸条方向, 因为图案是用滚筒印刷的, 另一次因为纸条是边对着边粘在墙上的. 晶体条件限制了可能性, 把晶体群分成了 17 种类型. 不同类型的对称的代表图案见图 6.6.2.

172

点群  $\bar{G}$  与平移群  $L$  不能完全确定群  $G$ .  $\bar{G}$  中的反射不必是  $G$  中一个反射的像这一事实使事情变得复杂. 如下面砖形图案所示, 它在  $G$  中可能只由一个滑动代表. 这个图案(我所喜爱的)是相对微妙的, 因为它的对称群不含有反射. 它有绕每个砖的中心转过角度  $\pi$  的旋转对称. 所有这些旋转代表点群  $\bar{G}$  的同一元素  $\bar{\rho}_\pi$ . 除了旋转角度 0 与  $\pi$  外, 没有其他非平凡

旋转对称. 图案也有沿图形所画虚线的滑动对称, 于是  $\bar{G} = D_2 = \{\bar{1}, \bar{\rho}_x, \bar{r}, \bar{\rho}_x \bar{r}\}$ .



分两步可很容易地确定一个图案的点群: 首先寻找旋转对称. 通常较容易找到. 点群  $\bar{G}$  中的旋转  $\bar{\rho}_\theta$  由图案对称群  $G$  中具有同一转动角度的一个旋转代表. 当所有旋转对称都找到时, 将得到一个整数  $n$  使得点群  $G$  为  $C_n$  或  $D_n$ . 这样, 要区分  $C_n$  与  $D_n$ , 可以看图案是否有反射或滑动对称. 如果有, 则  $\bar{G} = D_n$ ; 如果没有, 则  $\bar{G} = C_n$ .

### 在点群里具有四重旋转的平面晶体群

作为等距的离散群分类方法的范例, 我们分析其点群为  $C_4$  或  $D_4$  的群.

令  $G$  是一个这样的群, 设  $\bar{\rho}$  为  $\bar{G}$  中转过角度  $\pi/2$  的旋转, 并设  $L$  是  $G$  的格, 即使得  $t_v$  属于  $G$  的向量  $v$  的集合.

**【6.6.1】引理** 格  $L$  是方形.

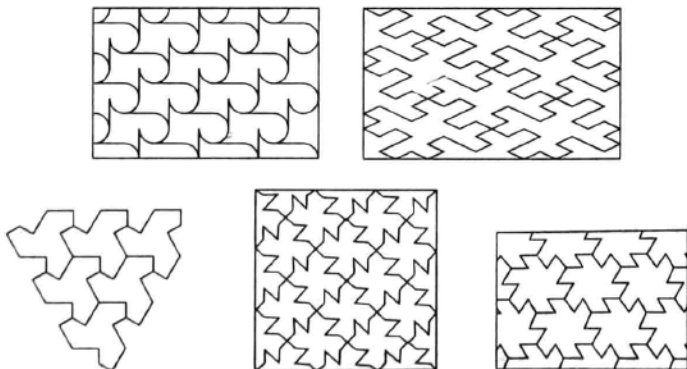
**证明** 在  $L$  里选取极小长度的非零向量  $a$ . 点群作用于  $L$  上, 所以  $\bar{\rho}(a) = b$  属于  $L$  且与  $a$  正交. 我们断言:  $(a, b)$  是  $L$  的格基.

假设不是. 根据引理 6.5.8, 将存在  $L$  的点属于由点  $r_1 a + r_2 b$  组成的区域  $\Pi'$  里, 其中  $0 \leq r_i < 1$ . 这样的点  $w$  与方形四个顶点  $0, a, b, a+b$  之一的距离将小于  $|a|$ . 称这个顶点为  $v$ . 这样,  $v-w$  也属于  $L$ , 且  $|v-w| < |a|$ . 这与  $a$  的选取矛盾. ■

适当选取坐标使得  $a$  与  $b$  是标准基向量  $e_1$  与  $e_2$ . 这样,  $L$  成为一个具有整数坐标的向量格,  $\Pi'$  成为向量  $(s, t)'$  的集合, 其中  $0 \leq s < 1, 0 \leq t < 1$ . 这在差一个平移的意义下确定了平面  $P$  中的坐标.

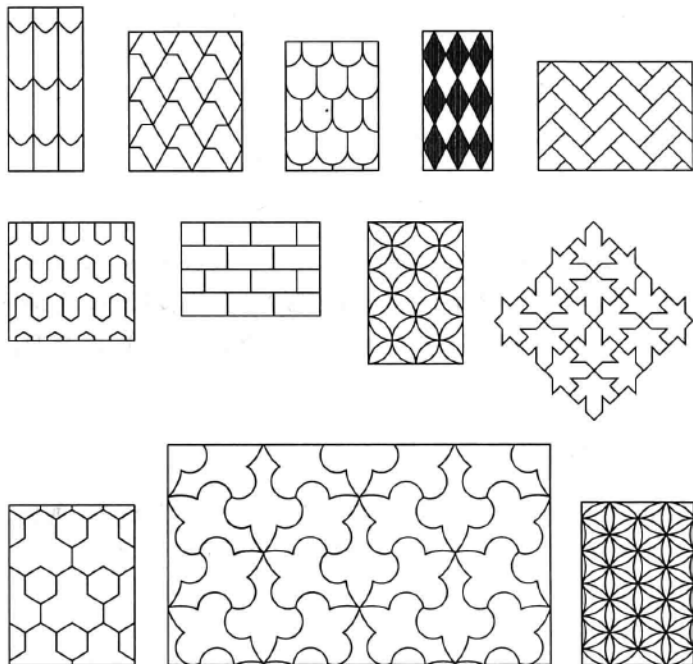
173

**【6.6.2】图**



17 个平面晶体群的图案范例

【6.6.2】图(续)



174

$V$  上映  $L$  到自身的正交算子构成二面体群  $D_4$ , 该二面体群由转过角度为  $\pi/2$  的旋转  $\bar{\rho}$  与标准反射  $\bar{\tau}$  生成. 我们的假设是  $\bar{\rho}$  属于  $\bar{G}$ . 如果  $\bar{\tau}$  也属于  $\bar{G}$ , 则  $\bar{G}$  是二面体群  $D_4$ . 如果不属于, 则  $\bar{G}$  是循环群  $C_4$ . 首先, 当  $\bar{G}$  是  $C_4$  时, 我们刻画群  $G$ . 令  $g$  是  $G$  的元素, 其在  $\bar{G}$  里的像为旋转  $\bar{\rho}$ . 这样,  $g$  是平面上围绕某个点  $p$  转过角度  $\pi/2$  的旋转. 我们在平面  $P$  里平移坐标使得点  $p$  为原点. 在这个坐标系里,  $G$  含有绕原点的旋转  $\rho = \rho_{\pi/2}$ .

**【6.6.3】命题** 令  $G$  是平面晶体群, 其点群  $\bar{G}$  是循环群  $C_4$ . 选取坐标使得  $L$  为具有整数坐标的点格, 且  $\rho = \rho_{\pi/2}$  是  $G$  的元素, 则群  $G$  由积  $t_v \rho^i$  组成, 其中  $v$  属于  $L$ ,  $0 \leq i < 4$ :

$$G = \{t_v \rho^i \mid v \in L\}$$

**证明** 令  $G'$  表示形如  $t_v \rho^i$  的元素集合, 其中  $v$  属于  $L$ . 必须证明  $G' = G$ . 由  $L$  的定义,  $t_v$  属于  $G$ ,  $\rho$  也属于  $G$ . 于是,  $t_v \rho^i$  属于  $G$ , 所以  $G'$  是  $G$  的子集.

要证反包含结论, 令  $g$  是  $G$  的元素. 由于点群  $\bar{G}$  是  $C_4$ , 故  $G$  的每个元素保持方向. 所以, 对某个平移向量  $u$  与某个角度  $\alpha$ ,  $g$  有形式  $g = t_u \rho_\alpha$ . 在点群里这个元素的像是  $\bar{\rho}_\alpha$ , 所以  $\alpha$  是  $\pi/2$  的倍数, 且对某个  $i$ ,  $\rho_\alpha = \rho^i$ . 由于  $\rho$  属于  $G$ , 故  $g \rho^{-i} = t_u$  属于  $G$ , 且  $u$  属于  $L$ . 所以,  $g$  属于  $G'$ . ■

现在考虑点群  $\bar{G}$  是  $D_4$  的情形.

**【6.6.4】命题** 令  $G$  是平面晶体群, 其点群  $\bar{G}$  是二面体群  $D_4$ . 选取坐标使得  $L$  为具有整数坐标的点格, 且  $\rho = \rho_{\pi/2}$  是  $G$  的元素. 再令  $c$  表示向量  $(\frac{1}{2}, \frac{1}{2})^t$ . 有两种可能:

(a)  $G$  的元素是积  $t_v\varphi$ , 其中  $v$  属于  $L$ ,  $\varphi$  属于  $D_4$ ,

$$G = \{t_v\rho^i | v \in L\} \cup \{t_v\rho^i r | v \in L\}, \text{ 或}$$

(b)  $G$  的元素是积  $t_v\varphi$ , 其中  $\varphi$  属于  $D_4$ . 如果  $\varphi$  是旋转, 则  $x$  属于  $L$ ; 如果  $\varphi$  是反射, 则  $x$  属于陪集  $c+L$ :

$$G = \{t_v\rho^i | v \in L\} \cup \{t_u\rho^i r | u \in c+L\}$$

**证明** 令  $H$  是  $G$  中保向等距的子集. 这是  $G$  的一个子群, 其平移格是  $L$ , 且含有  $\rho$ . 所以, 它的点群是  $C_4$ . 由命题 6.6.3 知,  $H$  由元素  $t_v\rho^i$  组成, 其中  $v$  属于  $L$ .

点群也含有反射  $\bar{r}$ . 选取  $G$  的元素  $g$  使得  $\bar{g} = \bar{r}$ . 对某个向量  $u$ , 它有形式  $g = t_u r$ , 但我们不知道  $u$  是否属于  $L$ . 分析这个情形需要浪费一点儿时间. 比如说  $u = (p, q)'$ .

我们用  $G$  里的平移  $t_v$  (亦即,  $v$  属于  $L$ ) 左乘  $g$ , 移动  $u$  到区域  $\Pi'$  里, 其中  $0 \leq p, q < 1$ . 假设这已经完成.

175

利用公式(6.3.3), 使用  $g = t_u r$  计算:

$$g^2 = t_u r t_u r = t_{u+ru}, \quad (g\rho)^2 = (t_u r \rho)^2 = t_{u+r\rho u}$$

这些是  $G$  的元素, 所以,  $u+ru = (2p, 0)'$ , 且  $u+r\rho u = (p-q, q-p)'$  属于格  $L$ . 它们是具有整数坐标的向量. 因为  $0 \leq p, q < 1$ , 且  $2p$  是整数, 故  $p$  或者为 0, 或者为  $\frac{1}{2}$ . 因为

$p-q$  也是整数, 故如果  $p=0, q$  为 0; 如果  $p=\frac{1}{2}, q$  为  $\frac{1}{2}$ . 所以, 对  $u$  仅有两种可能:

或者  $u = (0, 0)'$ , 或者  $u = c = (\frac{1}{2}, \frac{1}{2})'$ . 在前一情形,  $g = r$ , 于是,  $G$  含有一个反射.

这是命题里的情形(a). 第二种可能是情形(b). ■

## 第七节 抽象对称: 群作用

对称的概念可以应用到除几何图形之外的其他对象上. 例如, 复共轭  $(a+bi) \rightsquigarrow (a-bi)$  可以认为是复数的对称. 由于复数共轭与加法和乘法相容, 故称之为域  $\mathbf{C}$  的自同构. 几何上, 它是复平面关于实轴的双边对称, 但是, 它是自同构这一说法涉及其代数结构. 域  $F = \mathbf{Q}[\sqrt{2}]$  (其元素为形如  $(a+b\sqrt{2})$  的实数, 其中  $a$  与  $b$  是有理数) 也有自同构, 映  $(a+b\sqrt{2}) \rightsquigarrow (a-b\sqrt{2})$ . 这不是几何对称. 抽象“双边”对称的另一个例子由秩 3 的循环群  $H$  给出. 这个群有一个自同构, 它交换  $H$  中不等于单位元的两个元素.

代数结构  $X$  (诸如群或域) 的自同构的集合构成一个群, 合成法则为映射的合成. 在下面的意义下每个自同构都可以看做是  $X$  的对称: 它是与代数结构相容的  $X$  中元素的一个置换. 但在这种情形下, 结构是代数的而不是几何的.

因此, “自同构”和“对称”这两个词的意义或多或少是相同的, 只是“自同构”用于描述保持某个代数结构的集合的置换, 而“对称”常常(尽管不总是)指保持几何结构的置换.

自同构和对称都是群作用这个更一般的概念的特殊情形. 群  $G$  在集合  $S$  的一个作用是



组合  $G$  的元素  $g$  和  $S$  的元素  $s$  而得到  $S$  的另一个元素  $gs$  的法则. 换言之, 它是一个映射  $G \times S \rightarrow S$ . 我们暂时把应用于这个法则到元素  $g$  和元素  $s$  的结果记为  $g * s$ . 一个作用需要满足下列公理:

**【6.7.1】例**

(a) 对所有  $s \in S$ , 有  $1 * s = s$  ( $1$  是  $G$  的单位元).

(b) 结合律: 对所有  $g, g' \in G$  和所有  $s \in S$ , 有  $(gg') * s = g * (g' * s)$ . ■

通常省略星号, 把这个作用以乘法形式写为  $g, s \rightsquigarrow gs$ . 用乘法符号, 公理为  $1s = s$  与  $(gg')s = g(g's)$ .

176

群作用集合的例子在许多地方都能找到<sup>⊖</sup>, 且作用公理成立通常是很清楚的. 平面的等距的群  $M$  作用于平面上的点集. 它在平面上的直线的集合上作用与在平面上的三角形的集合上作用. 对称群  $S_n$  作用于指标集  $\{1, 2, \dots, n\}$ .

把这样的合成法则称为作用的原因在于: 如果固定  $G$  的一个元素  $g$  而让元素  $s$  在  $S$  中变动, 则用  $g$  左乘(或  $g$  的作用)定义一个  $S$  到其自身的映射. 用  $m_g$  记这个映射, 其刻画了元素  $g$  作用的方式:

**【6.7.2】** 
$$m_g : S \rightarrow S$$

是由  $m_g(s) = gs$  所定义的映射. 它是  $S$  的一个置换, 即它是一个双射, 因为它有逆函数  $m_g^{-1}$ : 由  $g^{-1}$  所做的乘法.

**注** 已知群  $G$  在集合  $S$  上的作用,  $S$  的一个元素  $s$  由群作用变为其他各式各样元素, 把这些元素收集在一起, 得到一个子集, 称之为  $s$  的轨道  $O_s$ :

**【6.7.3】** 
$$O_s = \{s' \in S \mid s' = gs, g \in G\}$$

当平面的等距群  $M$  作用于平面中三角形集合  $S$  上时, 已知的三角形  $\Delta$  的轨道  $O_\Delta$  是所有与  $\Delta$  全等的三角形的集合. 当证明有限群在平面作用(6.4.7)的不动点存在时, 引进了另一个轨道.

群作用轨道是等价关系的等价类:

**【6.7.4】** 
$$s \sim s', \text{ 如果对某个 } g \in G, \text{ 有 } s' = gs$$

所以, 如果  $s \sim s'$ , 也就是对某个  $g \in G$ , 如果  $s' = gs$ , 则  $s$  的轨道与  $s'$  的轨道是相同的. 因为它们都是等价类:

**【6.7.5】** 轨道分划集合  $S$

群  $G$  独立作用每个轨道. 例如, 平面上的三角形的集合可以划分为全等类, 且一个等距分别置换每一个全等类.

如果  $S$  恰由一个轨道组成, 则  $G$  的作用称为是可迁的. 这意味着  $S$  的每一个元素可通过群中的某个元素映为任意其他元素. 对称群  $S_n$  可迁地作用于指标集  $\{1, \dots, n\}$ . 平面的

<sup>⊖</sup> 在写书时, 数学家 Masayoshi Nagata 决定英语语言需要这个词 (manywhere); 后来他实际上在词典里找到了这个词.

等距群  $M$  可迁地作用于平面上的点的集合, 同时也可迁地作用于平面上的直线的集合. 但它不是可迁地作用于平面上的三角形的集合.

注  $S$  的元素  $s$  的稳定子是  $G$  中保持  $s$  不动的元素的集合. 它是  $G$  的子群, 常记为  $G_s$ :

$$[6.7.6] \quad G_s = \{g \in G \mid gs = s\}$$

177

例如, 群  $M$  对平面的点集作用时, 原点的稳定子同构于正交算子群  $O_2$ . 对称群  $S_n$  的作用指标  $n$  的稳定子同构于  $\{1, \dots, n-1\}$  的置换子群  $S_{n-1}$ . 或者, 如果  $S$  是平面中的三角形集合, 则一个特定的等边三角形  $\Delta$  的稳定子是它的对称群, 即  $M$  的同构于二面体群  $D_3$  的子群.

注意 下列的清晰区分是重要的: 当我们说等距  $m$  稳定一个三角形  $\Delta$  时, 不是意味着  $m$  使三角形  $\Delta$  的点固定不动. 使一个三角形的每个点都不动的唯一等距是恒等映射. 我们的意思是, 在置换三角形集合时,  $m$  把  $\Delta$  映为自身.

就像群同态  $\varphi: G \rightarrow G'$  的核  $K$  告诉我们什么时候  $G$  的两个元素  $x$  与  $y$  有同样的像一样, 也就是说, 如果  $x^{-1}y$  属于  $K$ ,  $S$  的元素  $s$  的稳定子  $G_s$  告诉我们什么时候  $G$  的两个元素  $x$  与  $y$  用相同的方式作用于元素  $s$ .

**[6.7.7] 命题** 令  $S$  是群  $G$  作用的集合, 设  $s$  是  $S$  的元素, 且  $H$  是  $s$  的稳定子.

(a) 如果  $a$  与  $b$  是  $G$  的元素, 则  $as = bs$  当且仅当  $a^{-1}b$  属于  $H$ ; 该命题为真当且仅当  $b$  属于陪集  $aH$ .

(b) 假设  $as = s'$ . 则  $s'$  的稳定子  $H'$  是共轭子群:

$$H' = aHa^{-1} = \{g \in G \mid g = aha^{-1}, h \in H\}$$

证明

(a)  $as = bs$  当且仅当  $s = a^{-1}bs$ .

(b) 如果  $g$  属于  $aHa^{-1}$ , 比如说  $g = aha^{-1}$ , 其中  $h \in H$ , 则  $gs' = (aha^{-1})(as) = ahs = as = s'$ . 于是,  $g$  稳定  $s'$ . 这表明  $aHa^{-1} \subset H'$ . 因为  $s = a^{-1}s'$ , 故可以交换  $s$  与  $s'$  的角色, 得到  $a^{-1}H'a \subset H$ , 从而  $H' \subset aHa^{-1}$ . 所以,  $H' = aHa^{-1}$ . ■

注意 命题的 (b) 部分解释了我们前面已经几次看见的现象: 当  $as = s'$ , 群元素  $g$  使  $s$  不动当且仅当  $aga^{-1}$  使  $s'$  不动.

## 第八节 对陪集的作用

设  $H$  是群  $G$  的一个子群. 就像我们所知道的那样, 左陪集  $aH$  划分  $G$ . 我们常将  $H$  在  $G$  中左陪集的集合记为  $G/H$ , 这是当子群为正规子群时用来表示商群的记号 (2.12.1). 当考虑作为集合  $G/H$  的元素时, 用括号  $[C]$  表示陪集  $C$ .

陪集的集合  $G/H$  不是群, 除非  $H$  是正规子群.

注  $G$  以一种自然的方式在  $G/H$  上作用.

作用是相当明显的: 设  $g$  是群  $G$  的元素, 并设  $C$  是一个陪集. 则  $g[C]$  定义为陪集  $[gC]$ , 其中  $gC = \{gc \mid c \in C\}$ . 因此, 如果  $[C] = [aH]$ , 则  $g[C] = [gaH]$ . 下面是一个基本命题.

178

**【6.8.1】命题** 令  $H$  是群  $G$  的一个子群.

(a)  $G$  在陪集的集合  $G/H$  上的作用是可迁的.

(b) 陪集  $H$  的稳定子是子群  $H$ .

再一次提请注意区别: 由  $H$  的元素  $h$  确定的乘法不是平凡地作用在陪集  $H$  的元素上, 但它将陪集  $[H]$  映到自身.

请仔细地推敲下面的例子. 设  $G$  是对称群  $S_3$ , 带有通常的表示, 令  $H$  是循环子群  $\{1, y\}$ . 它的左陪集是

**【6.8.2】**  $C_1 = H = \{1, y\}$ ,  $C_2 = xH = \{x, xy\}$ ,  $C_3 = x^2H = \{x^2, x^2y\}$

(见(2.8.4)), 且  $G$  作用在陪集集合  $G/H = \{[C_1], [C_2], [C_3]\}$ . 元素  $x$  与  $y$  用相同的方式作用于指标集  $\{1, 2, 3\}$ :

**【6.8.3】**  $m_x \leftrightarrow (1\ 2\ 3)$ ,  $m_y \leftrightarrow (2\ 3)$

例如,  $yC_2 = \{yx, yxy\} = \{x^2y, x^2\} = C_3$ .

下面的命题(有时叫轨道-稳定子定理)表明如何通过陪集上的作用来刻画任意群作用.

**【6.8.4】命题** 令  $S$  是群  $G$  作用的集合, 且  $s$  是  $S$  的元素. 设  $H$  与  $O_s$  分别是  $s$  的稳定子与轨道. 则存在一个由  $[aH] \rightsquigarrow as$  定义的双射  $\epsilon: G/H \rightarrow O_s$ . 这个映射与群的作用相容:  $\epsilon(g[C]) = g\epsilon([C])$  对每一个陪集  $C$  与  $G$  的每个元素  $g$  成立.

例如, 二面体群  $D_5$  作用于正五边形顶点. 令  $\mathcal{V}$  表示顶点集合, 令  $H$  是一个特殊顶点的稳定子. 存在双射  $D_5/H \rightarrow \mathcal{V}$ . 在平面  $P$  的等距群  $M$  作用下, 一个点的轨道是  $P$  的所有点的集合. 原点的稳定子是正交算子群  $O_2$ , 且存在一个双射  $M/O_2 \rightarrow P$ . 类似地, 如果  $H$  表示一条直线的稳定子, 且  $\mathcal{L}$  表示平面上所有点的集合, 则存在双射  $M/H \rightarrow \mathcal{L}$ .

**命题 6.8.4 的证明** 显然, 命题叙述中定义的映射  $\epsilon$  与群的作用相容, 如果它存在的话. 在符号上,  $\epsilon$  简单地用符号  $s$  替换了  $H$ . 究竟法则  $[gH] \rightsquigarrow gs$  能否定义一个映射并不清楚. 因为许多符号  $gH$  代表的是同一个陪集, 所以我们必须证明如果  $a$  和  $b$  是群元素且如果陪集  $aH = bH$ , 则也有  $as = bs$ . 假设  $aH = bH$ . 则  $a^{-1}b$  属于  $H$  (2.8.5). 因为  $H$  是  $s$  的稳定子, 故  $a^{-1}bs = s$ , 所以  $as = bs$ . 我们的定义是合理的, 且由反向推理知,  $\epsilon$  是单射. 由于  $\epsilon$  映  $[gH]$  为  $gs$  且  $gs$  可以为  $O_s$  的任意元素, 所以  $\epsilon$  是满射也是单射. ■

179

**注意** 定义映射  $\epsilon$  的推理经常出现. 假设集合  $\bar{S}$  是集合  $S$  上一个等价关系的等价类集, 令  $\pi: S \rightarrow \bar{S}$  是映它的元素  $s$  到它的等价类  $\bar{s}$  的映射. 从  $\bar{S}$  到另一个集合  $T$  定义一个映射的通常方法是: 给定  $\bar{S}$  的元素  $x$ , 在  $S$  中选取一个元素  $s$  使得  $x = \bar{s}$ , 且用  $s$  定义  $\epsilon(x)$ . 然后就像我们上面所做的, 必须证明定义不依赖于等价类为  $x$  的元素  $s$  的选取, 而仅依赖于等价类  $x$ . 这个过程用来证明映射是定义良好的.

## 第九节 计数公式

设  $H$  是有限群  $G$  的子群. 我们知道,  $H$  在  $G$  中的所有陪集有同样数量的元素, 用记号  $G/H$  表示陪集的集合, 阶  $|G/H|$  就是所说的  $H$  在  $G$  中的指标  $[G:H]$ . 计数公式

(2.8.8)变成了

$$\text{【6.9.1】} \quad |G| = |H| |G/H|$$

对任意群作用的轨道有类似的公式:

**【6.9.2】命题(计数公式)** 令  $S$  是群  $G$  作用的有限集合, 设  $G_s$  与  $O_s$  分别是  $S$  的元素  $s$  的稳定子与轨道, 则

$$|G| = |G_s| |O_s|$$

或

$$(G \text{ 的阶}) = (\text{稳定子的阶}) \cdot (\text{轨道的阶})$$

由(6.9.1)和命题 6.8.4 可得.

因此, 轨道的阶等于稳定子的指标,

$$\text{【6.9.3】} \quad |O_s| = [G:G_s]$$

且它整除群的阶. 对  $S$  的每个元素  $s$  都有一个这样的公式.

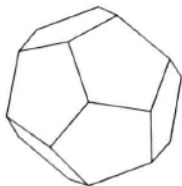
另一个公式利用集合  $S$  划分成轨道以对其元素进行计数. 将组成  $S$  的不同轨道以任意方式标号, 如  $O_1, \dots, O_k$ . 则

$$\text{【6.9.4】} \quad |S| = |O_1| + |O_2| + \dots + |O_k|$$

公式(6.9.2)与(6.9.4)有许多应用.

### 【6.9.5】例

(a) 正十二面体的旋转对称群  $G$  可迁地作用在它的面的集合  $F$  上. 一个特殊面  $f$  的稳定子  $G_f$  是围绕过  $f$  的中心转过角度  $2\pi/5$  的倍数的旋转群, 它的阶是 5. 十二面体有 12 个面. 由公式 6.9.2 得  $60 = 5 \cdot 12$ , 所以,  $G$  的阶是 60. 或者,  $G$  可迁地作用在顶点集  $V$  上. 一个顶点  $v$  的稳定子  $G_v$  是围绕过该顶点转过角度  $2\pi/3$  的倍数的 3 阶旋转群. 十二面体有 20 个顶点, 于是验证得  $60 = 3 \cdot 20$ . 对边有类似的计算:  $G$  可迁地作用在边的集合上, 且边  $e$  的稳定子含有恒等映射与围绕过  $e$  的中心转过角度  $\pi$  的旋转. 所以,  $|G_e| = 2$ . 因为  $60 = 2 \cdot 30$ , 故十二面体有 30 条边.



(b) 我们也可把群  $G$  的作用限制到子群  $H$  上. 由限制定义,  $G$  在一个集合  $S$  上的作用定义  $G$  在  $S$  上的作用, 且这个作用导致更多数量关系. 元素  $s$  的  $H$ -轨道将含在  $s$  的  $G$ -轨道内, 于是, 单个  $G$ -轨道将划分成若干  $H$ -轨道.

例如, 设  $F$  是十二面体的面的集合, 而  $H$  是某个特定面  $f$  的稳定子, 它是一个 5 阶循环群. 任意  $H$ -轨道的阶或者是 1, 或者是 5. 所以, 当把 12 个面的集合  $F$  划分成  $H$ -轨道

时, 必须找出阶为 1 的两个轨道. 我们发现:  $H$  固定  $f$  不动, 且它也使与  $f$  相对的面不动. 剩下其他面组成两条阶为 5 的轨道. 对于群  $H$  在面的集合上的作用, 公式 6.9.4 为  $12=1+1+5+5$ . 或者, 令  $K$  表示一个顶点的稳定子, 它是阶为 3 的循环群. 我们也可把集合  $F$  划分成  $K$ -轨道. 在这种情形下, 公式 6.9.4 成为  $12=3+3+3+3$ . ■

## 第十节 在子集上的作用

假设群  $G$  作用在子集  $S$  上. 如果  $U$  为  $S$  的阶为  $r$  的子集, 则

$$\text{【6. 10. 1】} \quad gU = \{gu \mid u \in U\}$$

是另一个阶为  $r$  的子集. 这就允许我们定义  $G$  在  $S$  的阶为  $r$  的子集的集合上的作用. 作用的公理很容易验证.

例如, 令  $O$  是一个立方体的 24 个旋转的八面体群, 且  $F$  是该立方体的 6 个面的集合. 于是,  $O$  也作用在  $F$  的阶为 2 的子集上, 也就是说, 作用在面的无序偶对上. 有 15 个偶对, 它们构成两个轨道:  $F = \{\text{相对面的偶对}\} \cup \{\text{邻面的偶对}\}$ . 这些轨道分别为 3 阶和 5 阶的.

子集  $U$  的稳定子是群元素  $g$  的集合, 使得  $[gU] = [U]$ , 就是说,  $gU = U$ . 相对面的偶对的稳定子的阶为 8.

再次提醒注意这点:  $U$  的稳定子由使得  $gU = U$  的群元素组成. 这意味着  $g$  在  $U$  里置换元素, 即只要  $u$  属于  $U$ , 则  $gu$  也属于  $U$ .

## 第十一节 置换表示

181 在本节里, 我们分析群  $G$  作用在集  $S$  上的各种方式.

注 群  $G$  的置换表示是从该群到一个对称群的同态:

$$\text{【6. 11. 1】} \quad \varphi: G \rightarrow S_n$$

【6. 11. 2】命题 令  $G$  是一个群, 则在  $G$  对集合  $S = \{1, \dots, n\}$  的作用与置换表示  $G \rightarrow S_n$  之间存在双射对应:

$$[G \text{ 在 } S \text{ 上的作用}] \leftrightarrow [\text{置换表示}]$$

证明 这是很简单的, 尽管当第一次看时可能产生混淆. 若已知  $G$  在  $S$  上的作用, 我们通过用  $\varphi(g) = m_g$ , 即由  $g$  确定的乘法 (6.7.2) 定义置换表示  $\varphi$ . 结合性质  $g(hi) = (gh)i$  表明

$$m_g(m_h i) = g(hi) = m_{gh} i$$

因此,  $\varphi$  是一个同态. 反之, 如果  $\varphi$  是置换表示, 则同样的公式定义了  $G$  在  $S$  上的作用. ■

例如, 二面体群  $D_n$  在正  $n$  边形的顶点集  $(v_1, \dots, v_n)$  上的作用定义了同态  $\varphi: D_n \rightarrow S_n$ .

命题 6.11.2 与它所讨论的指标集没有任何关系. 如果  $\text{Perm}(S)$  是任意集合  $S$  的置换群, 则我们也称同态  $\varphi: G \rightarrow \text{Perm}(S)$  为  $G$  的一个置换表示.

【6. 11. 3】推论 令  $\text{Perm}(S)$  表示集合  $S$  的置换群, 且  $G$  是一个群, 则在  $G$  对集合  $S$  的作用与置换表示  $\varphi: G \rightarrow \text{Perm}(S)$  之间存在一个双射对应:

$$[G \text{ 在 } S \text{ 上的作用}] \leftrightarrow [\text{同态 } G \rightarrow \text{Perm}(S)]$$

置换表示  $G \rightarrow \text{Perm}(S)$  不一定是单射. 如果它碰巧是单射, 就称对应的作用是忠实的. 要成为忠实的, 作用必须有如下性质:  $m_g$  (即由  $g$  确定的乘法) 不是恒等映射, 除非  $g=1$ :

**【6. 11. 4】** 一个作用是忠实的, 如果它有如下性质:

$G$  的使得对每个  $s \in S$  有  $gs = s$  的仅有元素是恒等元

等距群  $M$  在平面上等边三角形的集合上的作用是忠实的, 因为等距是仅有的使所有等边三角形映为自身的恒等映射.

182

置换表示  $\varphi: G \rightarrow \text{Perm}(S)$  很少是满射, 因为  $\text{Perm}(S)$  的阶往往很大. 但下面的例子给出了一种情形.

**【6. 11. 5】例** 具有系数 mod 2 的可逆矩阵的群  $GL_2(\mathbf{F}_2)$  同构于对称群  $S_3$ .

用  $F$  表示域  $\mathbf{F}_2$ , 用  $G$  表示群  $GL_2(\mathbf{F}_2)$ . 列向量空间  $F^2$  由下面 4 个向量构成:

$$0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad e_1 + e_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

群  $G$  作用在 3 个非零向量的集合  $S = \{e_1, e_2, e_1 + e_2\}$ , 这给出一个置换表示  $\varphi: G \rightarrow S_3$ . 恒等矩阵是固定  $e_1$  与  $e_2$  不动的仅有的矩阵, 所以,  $G$  在  $S$  上的作用是忠实的, 且  $\varphi$  是单射. 可逆矩阵的列一定是  $S$  的不同元素的有序偶对. 有 6 个这样的偶对, 所以,  $|G| = 6$ . 因为  $S_3$  的阶为 6, 故  $\varphi$  是同构. ■

## 第十二节 旋转群的有限子群

本节我们应用计数公式对  $\mathbf{R}^3$  的旋转群  $SO_3$  的有限子群进行分类. 如同平面的有限等距群一样, 它们全部都是我们熟悉的图形的对称群.

**【6. 12. 1】定理**  $SO_3$  的有限子群是下列群之一:

$C_k$ : 绕一直线转过角度  $2\pi/k$  的倍数的旋转的循环群, 其中  $k$  是任意的;

$D_k$ : 正  $k$  边形的对称的二面体群, 其中  $k$  是任意的;

$T$ : 四面体的 12 个旋转对称的四面体群;

$O$ : 立方体或八面体的 24 个旋转对称的八面体群;

$I$ : 十二面体或二十面体的 60 个旋转对称的二十面体群.



**注意** 二面体群通常作为平面中正多边形的对称群表现出来, 其中反射逆向. 然而, 平面的反射可由三维空间中转过角度  $\pi$  的旋转得到, 用这种方法正多边形的对称可以作为  $\mathbf{R}^3$  的旋转来实现. 二面体群  $D_n$  可由围绕  $e_1$  轴转过角度  $2\pi/n$  的旋转  $x$  和围绕  $e_2$  轴转过角度  $\pi$  的旋转  $y$  生成. 设  $c = \cos 2\pi/n$ ,  $s = \sin 2\pi/n$ , 表示这些旋转的矩阵为

183

$$\text{【6. 12. 2】} \quad x = \begin{bmatrix} 1 & & \\ & c & -s \\ & s & c \end{bmatrix}, \quad y = \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix}$$

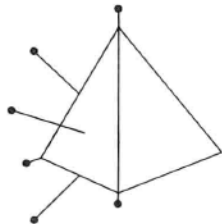
令  $G$  是  $SO_3$  的有限子群, 其阶  $N > 1$ . 我们称  $G$  的元素  $g \neq 1$  的极点为群的极点. 除了恒等映射外,  $\mathbf{R}^3$  的任意旋转有两个极点——旋转轴与单位球面  $S^2$  的交点. 所以,  $G$  的极点是由不同于 1 的群元素  $g$  固定不动的 2 维球面上的点.

**【6. 12. 3】例** 四面体  $\Delta$  的旋转对称群  $T$  的阶是 12, 它的极点是  $S^2$  上位于顶点上面、面的中心上面或边的中点上面的点. 因为有 4 个面、4 个顶点和 6 个边, 故有 14 个极点.

$$|\text{极点}| = 14 = |\text{面}| + |\text{顶点}| + |\text{边}|$$

$T$  有 11 个元素  $g \neq 1$ , 每个这样的元素有两个自旋——两个偶对  $(g, p)$ , 其中  $p$  是  $g$  的极点. 所以, 一共有 22 个自旋. 一个面的稳定子的阶是 3. 它的不等于 1 的两个元素共有面的中心上面的一个极点. 类似地, 有两个元素共有一个顶点上面的极点, 有一个元素有边的中心上面的顶点.

$$|\text{自旋}| = 22 = 2|\text{面}| + 2|\text{顶点}| + |\text{边}| \quad \blacksquare$$



用  $\mathcal{P}$  表示有限子群  $G$  的所有极点的集合. 我们将通过数极点的个数得到关于群的信息. 就像例子所展示的, 计数可能比较混乱.

**【6. 12. 4】引理**  $G$  的极点集  $\mathcal{P}$  是  $G$ -轨道的并. 所以,  $G$  作用在  $\mathcal{P}$  上.

**证明** 设  $p$  是一个极点, 比如它是  $G$  的元素  $g \neq 1$  的极点. 设  $h$  是  $G$  的另一个元素且  $q = hp$ . 我们要证明  $q$  是一个极点, 也就是说证明  $q$  在  $G$  的某个不是单位元的元素  $g'$  的作用下固定不动. 所需的元素为  $hgh^{-1}$ . 因为  $g \neq 1$ , 故这个元素不等于 1, 且  $hgh^{-1}q = hgp = hp = q$ .  $\blacksquare$

极点  $p$  的稳定子是在  $G_p$  中所有围绕  $p \in G$  的旋转的群. 这个群是循环群, 它由在  $G$  中转过最小正角度  $\theta$  的旋转生成. 记它的阶为  $r_p$ , 则  $\theta = 2\pi/r_p$ .

由于  $p$  是极点, 故稳定子  $G_p$  除单位元外还含有其他元素, 所以  $r_p > 1$ .  $G$  中具有给定极点  $p$  的元素的集合是稳定子  $G_p$ , 其中不包含单位元. 于是有  $r_p - 1$  个以  $p$  为极点的群元素. 除了 1 之外的每个群元素有两个极点. 因为  $|G| = N$ , 故有  $2N - 2$  个自旋. 这给出关系

$$\text{【6. 12. 5】} \quad \sum_{p \in \mathcal{P}} (r_p - 1) = 2(N - 1)$$

合并相关项简化这个方程的左边: 令  $n_p$  表示  $p$  的轨道  $O_p$  的阶数. 由计数公式 (6. 9. 2),

$$\text{【6. 12. 6】} \quad r_p n_p = N$$

如果两个极点  $p$  与  $p'$  属于同一轨道, 则它们的轨道是相等的, 于是,  $n_p = n_{p'}$ , 所以  $r_p = r_{p'}$ . 我们将各种轨道以任意方式编号, 比如说  $O_1, O_2, \dots, O_k$ , 并且令  $n_i = n_p$  与  $r_i = r_p$ , 其中  $p \in O_i$ , 从而  $n_i r_i = N$ . 因为轨道  $O_i$  含有  $n_i$  个元素, 故在(6. 12. 5)的左边有  $n_i$  项等于  $r_i - 1$ . 合并这些项, 得到方程

$$\sum_{i=1}^k n_i (r_i - 1) = 2N - 2$$

我们用  $N$  整除两边得到著名公式:

$$\text{【6. 12. 7】} \quad \sum_i \left(1 - \frac{1}{r_i}\right) = 2 - \frac{2}{N}$$

一眼看上去, 对这个公式可能没有什么太多的期望, 但实际上它告诉了我们许多东西. 右边大于 1 小于 2, 而左边的每一项至少是  $\frac{1}{2}$ . 由此可以得到最多有三条轨道.

剩下的分类工作可通过列出不同的可能情形进行:

一条轨道:  $1 - \frac{1}{r_1} = 2 - \frac{2}{N}$ . 这是不可能的, 因为  $1 - \frac{1}{r_1} < 1$ , 而  $2 - \frac{2}{N} \geq 1$ .

两条轨道:  $\left(1 - \frac{1}{r_1}\right) + \left(1 - \frac{1}{r_2}\right) = 2 - \frac{2}{N}$ . 即  $\frac{1}{r_1} + \frac{1}{r_2} = \frac{2}{N}$ .

因为  $r_i$  整除  $N$ , 所以这个方程仅当  $r_1 = r_2 = N$  时成立. 这样,  $n_1 = n_2 = 1$ . 有两个极点  $p_1$  和  $p_2$ , 它们都为群的每一个元素所固定不动. 于是,  $G$  是围绕过  $p_1$  和  $p_2$  的直线  $\ell$  旋转的循环群  $C_N$ .

三条轨道:  $\left(1 - \frac{1}{r_1}\right) + \left(1 - \frac{1}{r_2}\right) + \left(1 - \frac{1}{r_3}\right) = 2 - \frac{2}{N}$ .

这是最有趣的情形. 因为  $\frac{2}{N}$  是正的, 故公式蕴含着

$$\text{【6. 12. 8】} \quad \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} > 1$$

将  $r_i$  按增序排列. 则  $r_1 = 2$ ; 如果所有的  $r_i$  都至少为 3, 则左边将  $\leq 1$ .

情形 1:  $r_1 = r_2 = 2$ . 第三个阶  $r_3 = k$  可以是任意的, 且  $N = 2k$ :

$$r_i = 2, 2, k; \quad n_i = k, k, 2; \quad N = 2k$$

有一对极点  $\{p, p'\}$  形成轨道  $O_3$ . 所以,  $G$  的一半元素固定  $p$  不动, 另一半元素使  $p$  和  $p'$  互换. 因而  $G$  的元素要么是绕过  $p$  和  $p'$  的直线  $\ell$  的旋转, 要么是绕与  $\ell$  垂直的直线转过角度  $\pi$  的旋转. 群  $G$  是使一个正  $k$  边形  $\triangle$  固定不动的旋转群, 是一个二面体群  $D_k$ . 多边形  $\triangle$  位于与  $\ell$  垂直的平面上,  $\triangle$  的顶点和面的中心对应于剩下的极点.  $\mathbf{R}^2$  中多边形的双侧(反射)对称成为了  $\mathbf{R}^3$  中转过角度  $\pi$  的旋转.

情形 2:  $r_1 = 2, 2 < r_2 \leq r_3$ . 方程  $1/2 + 1/4 + 1/4 = 1$  排除了  $r_2 \geq 4$  的可能性. 所以  $r_2 = 3$ . 这样, 方程  $1/2 + 1/3 + 1/6 = 1$  排除了  $r_3 \geq 6$ . 剩下的只有三种可能:



## 【6. 12. 9】

(i)  $r_i=2, 3, 3; n_i=6, 4, 4; N=12$ .

轨道  $O_3$  里的极点是正四面体的顶点, 且  $G$  是它的 12 个旋转对称的四面体群  $T$ .

(ii)  $r_i=2, 3, 4; n_i=12, 8, 6; N=24$ .

轨道  $O_3$  里的极点是正八面体的顶点, 且  $G$  是它的 24 个旋转对称的四面体群  $O$ .

(iii)  $r_i=2, 3, 5; n_i=30, 20, 12; N=60$ .

轨道  $O_3$  里的极点是正二十面体的顶点, 且  $G$  是它的 60 个旋转对称的二十面体群  $I$ . 在每一种情形中, 整数  $n_i$  分别是边数、面数和顶点数.

直观上, 一条轨道上的极点应该是正多面体的顶点, 因为它们一定均衡地分布在球面上. 然而这并不太精确, 例如, 像立方体的边的中点构成一个轨道, 但却不能张成一个正多面体. 它们张成的图案称为截多面体.

我们将证明(iii)的结论. 令  $V$  是阶为 12 的轨道  $O_3$ . 我们想证明这个轨道里的极点是正二十面体的顶点. 设  $p$  是  $V$  里的极点之一. 把  $p$  想象成单位球面的北极, 这样就有赤道与南极. 令  $H$  是  $p$  的稳定子. 因为  $r_3=5$ , 故这是一个循环群, 由绕  $p$  转过角度  $2\pi/5$  的旋转  $x$  生成. 当把  $V$  分解成  $H$ -轨道时, 我们一定得到两个阶为 1 的  $H$ -轨道. 这些是北极与南极. 其他 10 个极点构成两个阶为 5 的  $H$ -轨道. 我们记它们为  $\{q_0, \dots, q_4\}$  与  $\{q'_0, \dots, q'_4\}$ , 其中  $q_i = x^i q_0$ ,  $q'_i = x^i q'_0$ . 由于北极与南极的对称性, 这些轨道中, 一个在北半球, 一个在南半球, 或者都在赤道上. 比如说, 轨道  $\{q_i\}$  在北半球或在赤道上.

186

令  $|x, y|$  表示单位球面上点  $x$  与  $y$  之间的球面距离. 注意  $d = |p, q_i|$  与  $i=0, \dots, 4$  无关, 因为存在  $H$  的元素映  $q_0 \rightsquigarrow q_i$ , 而固定  $p$  不动. 类似地,  $d' = |p, q'_i|$  与  $i$  无关, 所以, 当  $p'$  遍历轨道  $V$  时, 距离  $|p, p'|$  仅取四个值  $0, d, d'$  和  $\pi$ . 值  $d$  与  $d'$  取到 5 次,  $0$  与  $\pi$  取到 1 次. 因为  $G$  可迁地作用在  $V$  上, 故当用  $V$  里任意其他极点替换  $p$  时, 我们将得到同样四个值.

注意  $d \leq \pi/2$ , 而  $d' \geq \pi/2$ . 因为在轨道  $\{q_i\}$  里有五个极点, 所以球面距离  $|q_i, q_{i+1}|$  小于  $\pi/2$ , 于是, 它等于  $d$ , 且  $d < \pi/2$ . 所以, 那个轨道不是赤道. 三个极点  $p, q_i, q_{i+1}$  构成等边三角形. 有五个全等的等边三角形在  $p$  处相交, 所以, 五个全等三角形在极点相交. 它们构成二十面体的面.

**注意** 恰好存在五个正多面体. 这可通过计数方法证明, 即把全等正多边形放在一个顶点上构造正多面体. 可构造三个、四个或五个等边三角形、三个正方形或三个正五边形. (六个三角形、四个正方形或三个六边形粘在一起成为平坦面.) 于是, 恰有五种可能. 但这个分析略去了有趣的存在性问题. 二十面体存在吗? 当然, 我们可用纸板做一个. 但当我们做时, 三角形从来没有精确合适地粘在一起, 我确信其原因是我们的不精确. 如果关于乐谱里的五分之一圆周我们下类似的结论, 那就错了: 五分之一圆周几乎闭合, 但不是特别闭合. 保证二十面体存在的最后办法也许是写下它的顶点坐标, 验证距离. 这是练习 12.7.

平面等距的讨论对于 3 维空间的等距群有类似结果. 可以定义晶体群的概念, 这是其平移群是三维格的离散子群. 晶体群与二维格群类似, 三维结构中的晶体形状以这样的群为其对称结构的例子. 与存在 17 个格群(6.6.2)类似, 可以证明存在 230 类晶体群. 这些群的列表太长而不太有用, 因而晶体被粗分为七个晶体系. 对于这方面更多的内容以及对于 32 个晶体点群的讨论, 请参看有关晶体的书, 例如[Schwarzenbach].

一个好的传统比最有趣的几何问题更有价值,  
因为它保持了一般方法,  
并且有助于很好地解决问题.

—Gottfried Wilhelm Leibnitz<sup>①</sup>

187

## 练 习

### 第一节 平面图形的对称

1.1 确定图 6.1.4、图 6.1.6 和图 6.1.7 的所有对称.

### 第三节 平面的等距

3.1 验证规则(6.3.3).

3.2 令  $m$  是反向等距. 用代数方法证明  $m^2$  是一个平移.

3.3 证明  $\mathbf{R}^2$  上的线性算子是一个反射当且仅当它的特征值是 1 与  $-1$ , 且伴随于这些特征值的特征向量是正交的.

3.4 证明  $M$  里的滑动反射的共轭是滑动反射, 且滑动向量有相同长度.

3.5 用复变量  $z = x + iy$  的形式写出等距(6.3.1)的公式.

3.6 (a) 令  $s$  是平面上围绕点  $(1, 1)$  转过角度  $\pi/2$  的旋转. 写出  $s$  作为乘积  $t_{a\rho\theta}$  的公式.

(b) 令  $s$  是平面上围绕竖轴  $x=1$  的反射. 求一个等距  $g$  使得  $grg^{-1} = s$ , 并把  $s$  写成  $t_{a\rho\theta r}$  的形式.

### 第四节 平面上正交算子的有限群

4.1 在二面体群  $D_n$  中, 以  $x^i y^j$  的形式写出乘积  $x^2 y x^{-1} y^{-1} x^3 y^3$ .

4.2 (a) 列出二面体群  $D_4$  的所有子群, 并确定哪些是正规的.

(b) 列出二面体群  $D_{15}$  的真正正规子群  $N$ , 并确定其商群  $D_{15}/N$ .

(c) 列出  $D_6$  的不含有  $x^3$  的子群.

4.3 (a) 在二面体群  $D_{10}$  中计算子群  $H = \{1, x^5\}$  的左陪集.

(b) 证明  $H$  是正规的, 且  $D_{10}/H$  同构于  $D_5$ .

(c)  $D_{10}$  同构于  $D_5 \times H$  吗?

### 第五节 离散等距群

5.1 令  $\ell_1$  与  $\ell_2$  是  $\mathbf{R}^2$  中过原点的直线, 且它们相交的角度为  $\pi/n$ , 设  $r_i$  是关于  $\ell_i$  的反射. 证明  $r_1$  与  $r_2$  生成二面体群  $D_n$ .

① 我从 V. I. Arnold 那里获悉该引文. 洛必达写信给莱布尼兹, 道歉长时间没有回信, 并说他在乡下照顾遗产. 回信中, 莱布尼兹告诉他不用担心, 并用该引文继续.

- 5.2 什么是其平移群  $L$  有形式  $\mathbf{Z}a$  且  $a \neq 0$  的等距离散群的晶体限制?
- 5.3  $\mathbf{R}^2$  中格  $L$  里含有多少指标为 3 的子格?
- 5.4 设  $(a, b)$  是  $\mathbf{R}^2$  中一个格  $L$  的格基. 证明每个其他格基具有  $(a', b') = (a, b)P$  的形式, 其中  $P$  是一个行列式为  $\pm 1$  的  $2 \times 2$  整数矩阵.
- 5.5 证明装饰图案  $\triangleleft \triangleleft \triangleleft \triangleleft \triangleleft \triangleleft$  的对称群同构于二阶循环群和无限循环群的直积  $C_2 \times C_\infty$ .
- 188 5.6 令  $G$  是装饰图案  $\sqcup \sqcup \sqcup \sqcup \sqcup \sqcup \sqcup \sqcup$  的对称群. 确定  $G$  的点群  $\bar{G}$  以及它的平移子群在  $G$  里的指标.
- 5.7 令  $N$  表示直线  $\mathbf{R}^1$  的等距群. 对  $N$  的离散子群分类, 将  $\mathbf{R}^1$  上原点与单位长度的选择有差别的离散子群等同起来.
- 5.8 设  $N'$  是一条无限长的带子

$$R = \{(x, y) \mid -1 \leq y \leq 1\}$$

的等距群. 它可视为群  $M$  的子群. 下列元素属于  $N'$ :

$$t_a: (x, y) \rightarrow (x + a, y)$$

$$s: (x, y) \rightarrow (-x, y)$$

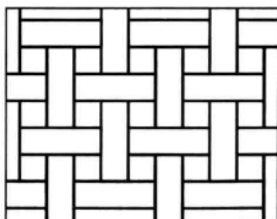
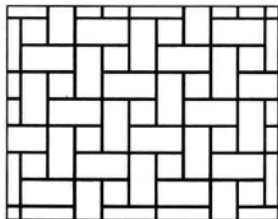
$$r: (x, y) \rightarrow (x, -y)$$

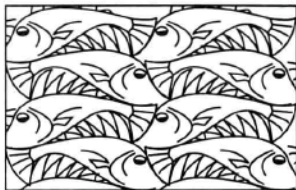
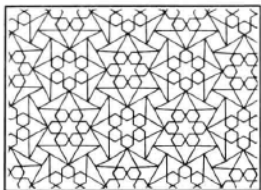
$$\rho: (x, y) \rightarrow (-x, -y)$$

- (a) 对这些等距叙述并证明 (6.3.3) 的类似结果.
- (b) 在其对称群离散的意义下, 装饰图案为带子上周期的图案. 对出现的对称群进行分类, 将那些仅在带子上原点和单位长度的选择上有差别的群等同起来. 先试着做具有不同种类的对称的图案. 当证明你的结论时请做细致的情形分析.
- 5.9 令  $G$  是  $M$  的离散子群, 其平移群是非平凡的. 证明平面上存在点  $p_0$  不为  $G$  的除恒等元以外的任意元素所固定不动.
- 5.10 令  $f$  与  $g$  是平面上围绕两个不同点转过任意非零旋转角度  $\theta$  与  $\varphi$  的旋转. 证明由  $f$  与  $g$  生成的群含有一个平移.
- 5.11 如果  $S$  与  $S'$  是  $\mathbf{R}^n$  的子集且  $S \subset S'$ , 则  $S$  在  $S'$  中是稠密的, 如果对  $S'$  的每个元素  $s'$ , 存在  $S$  的元素任意接近  $s'$ .
- (a) 证明  $\mathbf{R}^+$  的子群  $\Gamma$  或者在  $\mathbf{R}$  里稠密, 或者是离散的.
- (b) 证明由 1 和  $\sqrt{2}$  生成的  $\mathbf{R}^+$  的子群在  $\mathbf{R}^+$  里稠密.
- (c) 令  $H$  是角度群  $G$  的子群. 证明  $H$  或者是循环群, 或者它在  $G$  里稠密.
- 5.12 对加群  $\mathbf{R}^{3+}$  的离散子群进行分类.

## 第六节 平面晶体群

- 6.1 (a) 确定在图 6.6.2 中描绘的每个图案的点群  $\bar{G}$ .
- (b) 对哪个图案可选取坐标使得群  $G$  作用于格  $L$ ?
- 6.2 令  $G$  是等边三角格  $L$  的对称群. 确定  $G$  的平移子群在  $G$  里的指标.
- 189 6.3 确定下面所显示的图案的点群, 并找出一个与图 6.6.2 中有同一类型的对称的图案.





- 6.4 将具有点群  $D_4 = \{\bar{1}, \bar{4}\}$  的平面晶体群进行分类.
- 6.5 (a) 证明: 如果 2 维晶体群  $G$  的点群是  $C_6$  或  $D_6$ , 则平移群  $L$  是等边三角格.  
(b) 对这些群加以分类.
- 6.6 证明图 6.6.2 中所给图的对称群穷尽了所有可能情形.

### 第七节 抽象对称: 群作用

- 7.1 令  $G = D_4$  是正方形的对称二面体群.  
(a) 顶点的稳定子是什么? 边的呢?  
(b)  $G$  在由对角线组成的二元集合上作用, 对角线的稳定子是什么?
- 7.2 平面的等距群  $M$  作用于平面上直线的集合, 确定一条直线的稳定子.
- 7.3 对称群  $S_3$  作用于阶为 3 的两个集合  $U$  与  $V$ . 在下列两种情形下分解积集  $U \times V$  为“对角作用” $g(u, v) = (gu, gv)$  的轨道:  
(a) 在  $U$  与  $V$  上的作用是可迁的.  
(b) 在  $U$  上的作用是可迁的, 在  $V$  上的作用轨道是  $\{v_1\}$  与  $\{v_2, v_3\}$ .
- 7.4 在练习 6.3 的每个图中求有非平凡稳定子的点, 并且确定其稳定子.
- 7.5 令  $G$  是一个正方体的含有反向对称的对称群, 用几何方法描述  $G$  的元素.
- 7.6 令  $G$  是一个等边三角棱柱体  $P$  的含有反向对称的对称群, 确定  $P$  的一个矩形面的稳定子和群的阶.
- 7.7 设  $G = GL_n(\mathbf{R})$  通过左乘作用于集合  $V = \mathbf{R}^n$ .  
(a) 描述  $V$  在这个作用下的轨道分解.  
(b)  $e_1$  的稳定子是什么?
- 7.8 对于  $GL_2(\mathbf{C})$  的下列作用, 分解  $2 \times 2$  复矩阵集合  $\mathbf{C}^{2 \times 2}$  为轨道.  
(a) 左乘 (b) 共轭.
- 7.9 (a) 设  $S$  是实  $m \times n$  矩阵的集合  $\mathbf{R}^{m \times n}$ , 并设  $G = GL_m(\mathbf{R}) \times GL_n(\mathbf{R})$ . 证明规则  $(P, Q) * A = PAQ^{-1}$  定义  $G$  在  $S$  上的作用.  
(b) 刻画  $S$  的  $G$ -轨道分解.  
(c) 设  $m \leq n$ . 矩阵  $[I | 0]$  的稳定子是什么?
- 7.10 (a) 刻画矩阵  $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$  在一般线性群  $GL_n(\mathbf{R})$  共轭作用之下的轨道和稳定子.  
(b) 在  $GL_2(\mathbf{F}_5)$  中解释该矩阵, 求轨道的阶(元素的个数).
- 7.11 证明对称群  $S_4$  的仅有的 12 阶子群是交错群  $A_4$ .

### 第八节 对陪集的作用

- 8.1 规则  $P * A = PAP^{-1}$  定义  $GL_n$  在  $n \times n$  矩阵的集合上的作用吗?
- 8.2 对于  $G$  在  $G/H$  上的作用, 陪集  $aH$  的稳定子是什么?
- 8.3 当  $G$  是二面体群  $D_4$  而  $S$  是正方形顶点的集合时, 具体写出双射(6.8.4).
- 8.4 对于对称群  $G = S_n$  在指标集  $\{1, \dots, n\}$  上的作用, 令  $H$  是指标 1 的稳定子, 刻画  $G$  中  $H$  的左陪集

并在此情形下刻画映射(6.8.4).

### 第九节 计数公式

- 9.1 利用计数公式确定立方体和四面体的旋转对称群的阶.
- 9.2 设  $G$  是立方体的旋转对称群, 设  $G_v, G_e, G_f$  为顶点  $v$ 、边  $e$  和面  $f$  的稳定子, 并设  $V, E, F$  分别是立方体顶点、边和面的集合. 求代表三个集合中的每一个对于其每一个子群分解为轨道的公式.
- 9.3 当允许像平面反射这样的反面对称时, 确定十二面体对称群的阶.
- 191 9.4 确定正四面体的包含反面对称的旋转对称群  $T'$ .
- 9.5 令  $F$  是一段  $I$ -光束, 可看作是字母  $I$  与单位区间的积集合, 确定它的包含反面对称的对称群.
- 9.6 确定全球的对称群, 考虑缝口(但不考虑针线口), 允许反面对称.

### 第十节 在子集上的作用

- 10.1 确定  $D_3$  的阶为 3 的子集的集合上左乘轨道的阶.
- 10.2 令  $S$  是一个群  $G$  可迁地作用的有限集合, 并设  $U$  是  $S$  的子集. 证明诸子集  $gU$  均匀地覆盖  $S$ , 亦即,  $S$  的每个元素属于同一数量的集合  $gU$ .
- 10.3 考虑  $G$  在它的子集的集合上的左乘作用. 设  $U$  是子集使得集合  $gU$  划分  $G$ . 令  $H$  是在这个轨道里含有 1 的唯一子集. 证明  $H$  是  $G$  的子群.

### 第十一节 置换表示

- 11.1 刻画  $S_3$  在四元素集合上的所有作用方式.
- 11.2 刻画四面体群  $T$  在二元素集合上的所有作用方式.
- 11.3 设  $S$  是集合, 群  $G$  在其上作用, 且设  $H$  是使得对所有  $s \in S$  有  $gs = s$  的元素  $g$  的子集. 证明  $H$  是  $G$  的正规子群.
- 11.4 设  $G$  是正方形的对称二面体群  $D_4$ .  $G$  在顶点集合上的作用是忠实的吗? 在对角线上呢?
- 11.5 群  $G$  忠实地作用在五元素集合  $S$  上, 且有两个轨道, 一个轨道阶为 3, 一个轨道阶为 2.  $G$  可能是什么群?
- 11.6 设  $F = \mathbb{F}_3$ . 列向量空间  $F^2$  有四个一维子空间. 列出这些子空间. 用可逆矩阵左乘可置换这些子空间. 证明这个作用定义一个同态  $\varphi: GL_2(F) \rightarrow S_4$ . 确定这个同态的核和像.
- 11.7 对下面每个群, 求最小整数  $n$ , 使得群在  $n$  元集上有忠实作用: (a)  $D_4$ , (b)  $D_6$ , (c) 四元数群  $H$ .
- 11.8 在乘法群  $F_p^\times$  与  $p$  阶循环群的自同构集合之间求双射对应.
- 11.9 三张矩形纸片  $S_1, S_2, S_3$  累成一摞. 令  $G$  是这个构形的所有对称的群, 包括单张纸片的对称以及三张矩形纸片集合的置换. 确定  $G$  的阶和由集合  $\{S_1, S_2, S_3\}$  的置换定义的映射  $G \rightarrow S_3$  的核.

### 第十二节 旋转群的有限子群

- 192 12.1 解释为什么十二面体与二十面体的对称群是同构的.
- 12.2 刻画八面体旋转群的极点的轨道.
- 12.3 设  $O$  是立方体的旋转群, 且设  $S$  是连接对角顶点的四个对角线的集合. 求连接对角线的稳定子.
- 12.4 设  $G=O$  是立方体的旋转群, 并设  $H$  是将其两个内接四面体之一映到自身的子群. 证明  $H=T$ .
- 12.5 证明二十面体群有 10 阶子群.
- 12.6 确定下列群的所有子群: (a) 四面体群, (b) 二十面体群.
- 12.7 如果适当选择  $\alpha > 1$ , 则 12 个点  $(\pm 1, \pm \alpha, 0)'$ ,  $(0, \pm 1, \pm \alpha)'$ ,  $(\pm \alpha, 0, \pm 1)'$  构成正二十面体的顶点. 验证这一点, 并确定  $\alpha$ .
- 12.8 证明三维晶体群的晶体限制: 晶体的旋转对称的阶为 2, 3, 4 或 6.

## 杂题

- M.1 令  $G$  是 2 维晶体群使得没有元素  $g \neq 1$  固定平面任意点不动. 证明  $G$  由两个平移生成, 或者由一个平移与一个滑动平移生成.
- M.2 (a) 证明群  $G$  的自同构集合  $\text{Aut}G$  构成群, 合成法则是函数的合成.  
 (b) 证明由  $g \rightsquigarrow$  (由  $g$  确定的共轭) 定义的映射  $\varphi: G \rightarrow \text{Aut}G$  是同态, 并确定它的核.  
 (c) 由群元素确定的共轭所得到的自同构叫做内自同构. 证明内自同构集合 (即  $\varphi$  的像) 是群  $\text{Aut}G$  的正规子群.
- M.3 确定群的自同构群 (见练习 M.2).  
 (a)  $C_4$  (b)  $C_6$  (c)  $C_2 \times C_2$  (d)  $D_4$  (e) 四元数群  $H$
- M.4  $\mathbb{R}^n$  中的坐标  $x_1, \dots, x_n$  如通常, 由不等式  $-1 \leq x_i \leq +1 (i=1, \dots, n)$  定义的点集是  $n$  维超立方  $C_n$ . 1 维超立方是一线段, 2 维超立方是一正方形. 4 维超立方有八个面立方, 它是由  $\{x_i=1\}$  与  $\{x_i=-1\} (i=1, \dots, 4)$  定义的 3 维立方, 它有 16 个顶点  $(\pm 1, \pm 1, \pm 1, \pm 1)$ .  
 令  $G_n$  表示正交群  $O_n$  的映超立方到自身的元素的子群, 即  $C_n$  的对称群, 包含有反向对称. 坐标与符号的置换在  $G_n$  的元素间变化.  
 (a) 用计数公式与归纳法确定群  $G_n$  的阶.  
 (b) 简洁地描述  $G_n$ , 并确定顶点  $(1, \dots, 1)$  的稳定子. 通过证明  $G_2$  同构于二面体群  $D_4$  来验证你的答案.
- M.5 (a) 找出一方法确定图 6.6.2 中第一个图案构成的一个河马头的面积. 对该图底部图案里的一个鸢尾花形纹章做同样的事情.  
 (b) 平面晶体群的基本区域  $D$  是平面的一个有界区域, 使得像  $gD (g \in G)$  恰好覆盖平面一次, 而没有重叠. 对河马图案对称群找出两个不全等的基本区域. 对鸢尾花形纹章做同样的事情.  
 (c) 证明: 如果  $D$  与  $D'$  是同一图案的基本区域, 则  $D$  可切割成有限多个小块, 并组合构成  $D'$ .  
 (d) 求一个联系基本区域与图案点群的阶的公式.
- M.6 令  $G$  是  $M$  的离散子群. 在平面中选取一点  $p$ , 其在  $G$  中的稳定子是平凡的, 设  $S$  是  $p$  的轨道. 对  $S$  的除  $p$  外的每个点  $q$ , 令  $\ell_q$  为线段  $\ell[p, q]$  的垂直平分线, 且  $H_q$  是含有  $p$  的以  $\ell_q$  为界的半平面. 证明  $D = \bigcap H_q$  是  $G$  的基本区域 (见 M.5).
- M.7 令  $G$  是作用在有限集合  $S$  上的有限群. 对  $G$  的每个元素  $g$ , 设  $S^g$  表示由  $g$  固定不动的  $S$  的元素的子集:  $S^g = \{s \in S \mid gs = s\}$ , 令  $G_s$  是  $s$  的稳定子.  
 (a) 我们可以想象断言  $gs = s$  的真-伪值表, 比如行以  $G$  中元素为指标而列以  $S$  中元素为指标. 对二面体群  $D_3$  在一个三角形的顶点上的作用构造这样的表.  
 (b) 证明公式  $\sum_{s \in S} |G_s| = \sum_{g \in G} |S^g|$ .  
 (c) 证明伯恩赛德公式:  $|G| \cdot (\text{轨道个数}) = \sum_{g \in G} |S^g|$ .
- M.8 存在  $70 = \binom{8}{4}$  种对八边形的边着色的方法, 使之有四条黑边四条白边. 群  $D_8$  作用在这 70 个元素的集合上, 轨道代表等价的颜色. 用伯恩赛德公式 (见练习 M.7) 计算等价类的个数.

## 第七章 群论的进一步讨论

多做或少证明的越多，做起来或证明起来就越容易。

——James Joseph Sylvester

本章我们讨论三个主题：共轭——最重要的一种群运算；西罗定理——有限群里刻画阶为素数幂的子群的定理；群的生成元和关系。

### 第一节 凯莱定理

每个群都以各种方式对自身作用，左乘是其中之一：

$$\begin{aligned} \text{【7.1.1】} \quad & G \times G \rightarrow G \\ & g, x \rightsquigarrow gx \end{aligned}$$

这是个可迁作用——仅有一条轨道。任意元素的稳定子都是平凡子群 $\langle 1 \rangle$ ，因而作用是忠实的，并且由这个作用定义的置换表示(见第六章第十一节)

$$\begin{aligned} \text{【7.1.2】} \quad & G \rightarrow \text{Perm}(G) \\ & g \rightsquigarrow m_g \text{——} g \text{ 的左乘} \end{aligned}$$

是单射。

**【7.1.3】定理(凯莱定理)** 每一个有限群都同构于某个置换群的子群。如果 $G$ 的阶为 $n$ ，则它同构于对称群 $S_n$ 的子群。

**证明** 因为左乘作用是忠实的，所以 $G$ 同构于它在 $\text{Perm}(G)$ 中的像。如果 $G$ 的阶为 $n$ ，则 $\text{Perm}(G)$ 同构于 $S_n$ 。 ■

虽然凯莱定理本身是很有意思的，但它难于使用，因为 $S_n$ 的阶与 $n$ 相比太大了。

### 第二节 类方程

共轭，即由

$$\text{【7.2.1】} \quad (g, x) \rightsquigarrow gxg^{-1}$$

定义的群 $G$ 到自身的作用比左乘更为微妙和重要。显然，我们不应用乘法记号表示这个作用。我们将证明这个作用的结合律(6.7.1)，使用 $g * x$ 作为共轭 $gxg^{-1}$ 的暂时记号：

$$(gh) * x = (gh)x(gh)^{-1} = ghxh^{-1}g^{-1} = g(h * x)g^{-1} = g * (h * x)$$

验证完公理，我们回到通常记号 $gxg^{-1}$ 。

**注**  $G$ 的元素 $x$ 关于共轭作用的稳定子叫做 $x$ 的中心化子。常记为 $Z(x)$ ：

$$\text{【7.2.2】} \quad Z(x) = \{g \in G \mid gxg^{-1} = x\} = \{g \in G \mid gx = xg\}$$

$x$ 的中心化子是与 $x$ 可交换的群元素的集合。

**注**  $x$ 关于共轭的轨道叫做 $x$ 的共轭类，常记为 $C(x)$ 。它由所有共轭 $gxg^{-1}$ 组成：

$$【7.2.3】 \quad C(x) = \{x' \in G \mid x' = gxg^{-1}, g \in G\}$$

计数公式(6.9.2)告诉我们:

$$【7.2.4】 \quad |G| = |Z(x)| \cdot |C(x)|$$

$$|G| = |\text{中心化子}| \cdot |\text{共轭类}|$$

群  $G$  的中心  $Z$  在第二章定义过. 它是与群的每个元素交换的元素的集合:  $Z = \{z \in G \mid zy = yz, y \in G\}$ .

**【7.2.5】命题**

(a)  $G$  的元素  $x$  的中心化子  $Z(x)$  含有  $x$ , 并且它含有中心  $Z$ .

(b)  $G$  的元素  $x$  属于中心当且仅当它的中心化子  $Z(x)$  是整个群  $G$ , 并且  $Z(x) = G$  当且仅当共轭类  $C(x)$  由单个元素  $x$  组成.

因为共轭类是群作用的轨道, 所以它们划分了群  $G$ . 这个事实给出了有限群的类方程:

$$【7.2.6】 \quad |G| = \sum_{\text{共轭类 } C} |C|$$

如果对共轭类编号, 比如记为  $C_1, \dots, C_k$ , 则这个公式成为

$$【7.2.7】 \quad |G| = |C_1| + \dots + |C_k|$$

恒等元 1 的共轭类由元素 1 单独组成. 首先列出这个类似乎是自然的, 于是,  $|C_1| = 1$ . 在类方程右边, 1 的再出现对应着  $G$  的中心  $Z$  的元素. 还要注意右边的每一项整除左边, 因为它是轨道的阶.

**【7.2.8】** 类方程右边的数整除群的阶, 且其中至少有一个为 1

196

这对于可能出现在这样方程中的整数组合是一个很强的限制.

对称群  $S_3$  的阶是 6. 用通常的记号, 元素  $x$  的阶是 3. 它的中心化子  $Z(x)$  含有  $x$ , 所以它的阶是 3 或 6. 因为  $yx = x^2y$ , 故  $x$  不属于群的中心, 且  $|Z(x)| = 3$ . 于是,  $Z(x) = \langle x \rangle$ , 计数公式(7.2.4)表明共轭类  $C(x)$  的阶是 2. 类似的推理表明元素  $y$  的共轭类  $C(y)$  的阶是 3. 对称群  $S_3$  的类方程为

$$【7.2.9】 \quad 6 = 1 + 2 + 3$$

如我们所见, 计数公式有助于确定类方程. 可直接确定共轭类的阶, 或者, 计算它的中心化子的阶. 中心化子是子群, 它有更多结构, 计算它的阶常常是较好的途径. 在下一节我们将看到一种容易确定共轭类的情形, 但让我们看一下使用中心化子的另一种情形.

令  $G$  是域  $F_3$  上行列式为 1 的矩阵的特殊线性群  $SL_2(F_3)$ . 这个群的阶是 24 (见练习 4.4). 通过罗列  $G$  的元素计算类方程是相当乏味的, 而从计算一些矩阵  $A$  的中心化子开始则比较好. 这通过对矩阵  $P$  求解方程  $PA = AP$  完成. 用这个方程比用  $PAP^{-1} = A$  更容易. 例如, 令

$$A = \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}, \quad P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

方程  $PA = AP$  迫使条件  $b = -c$  与  $a = d$  成立, 从而方程  $\det P = 1$  变成了  $a^2 + c^2 = 1$ . 这个



方程在  $\mathbf{F}_3$  里有四个解:  $a = \pm 1, c = 0$  与  $a = 0, c = \pm 1$ . 所以  $|Z(A)| = 4, |C(A)| = 6$ . 这就给出了类方程:  $24 = 1 + 6 + \dots$ . 要完成计算, 需要计算更多一些矩阵的中心化子. 因为共轭元素有相同的特征多项式, 故可从选取具有不同特征多项式的元素开始.

$SL_2(\mathbf{F}_3)$  的类方程是

$$\text{【7.2.10】} \quad 24 = 1 + 1 + 4 + 4 + 4 + 4 + 6$$

### 第三节 $p$ -群

类方程对阶是一个素数  $p$  的正幂的群  $G$  有若干应用. 这样的群称为  $p$ -群.

**【7.3.1】命题**  $p$ -群  $G$  的中心是非平凡群.

**证明** 比如说  $|G| = p^e$ , 其中  $e \geq 1$ . 类方程右边的每一项整除  $p^e$ , 于是, 它也是  $p$  的幂, 可能有  $p^0 = 1$ .  $p$  的正幂被  $p$  整除. 如果恒等元的类  $C_1$  是右边唯一给出 1 的项, 则类方程将为

$$p^e = 1 + \sum (\text{p 的倍数})$$

**197** 这是不可能的. 所以, 在右边一定有更多的 1. 中心是非平凡的. ■

类似的讨论可用来证明下面  $p$ -群作用的定理. 我们把它的证明留作练习.

**【7.3.2】定理(不动点定理)** 设  $G$  是一个  $p$ -群, 并设  $S$  是一个有限集合,  $G$  在它上面作用. 假设  $S$  的阶不被  $p$  整除, 则  $G$  在  $S$  上的作用有个不动点, 即稳定子为整个群的元素  $s$ .

**【7.3.3】命题** 每个阶为  $p^2$  的群是阿贝尔群.

**证明** 设  $G$  是阶为  $p^2$  的群. 根据前面的命题, 它的中心  $Z$  不是平凡群. 所以,  $Z$  的阶一定是  $p$  或  $p^2$ . 如果  $Z$  的阶是  $p^2$ , 则  $Z = G$ , 从而  $G$  如同命题所断言的那样是阿贝尔群. 假设  $Z$  的阶是  $p$ , 令  $x$  是  $G$  的但不属于  $Z$  的元素. 中心化子  $Z(x)$  包含  $x$  以及  $Z$ , 所以, 它严格大于  $Z$ . 因为  $|Z(x)|$  整除  $|G|$ , 故它一定等于  $p^2$ , 所以,  $Z(x) = G$ . 这意味着  $x$  与  $G$  的每个元素交换, 于是, 它属于中心. 矛盾. 所以, 中心不能为  $p$  阶. ■

**【7.3.4】推论** 阶为  $p^2$  的群或者是循环群, 或者是两个  $p$  阶循环群的积.

**证明** 设  $G$  是阶为  $p^2$  的群. 如果  $G$  包含  $p^2$  阶元素, 则它是循环群. 如果不是, 则  $G$  的每个不同于 1 的元素的阶为  $p$ . 选择阶为  $p$  的两个元素  $x$  与  $y$ , 使得  $y$  不属于子群  $\langle x \rangle$ . 命题 2.11.4 表明  $G$  同构于积  $\langle x \rangle \times \langle y \rangle$ . ■

$p^e$  阶群的同构类个数随  $e$  迅速增长. 有 5 个 8 阶群的同构类, 有 14 个 16 阶群的同构类和 51 个 32 阶群的同构类.

### 第四节 二十面体群 的类方程

本节我们确定十二面体的旋转对称的群——二十面体群  $I$  的共轭类, 以此来研究这个非常有趣的群. 当考虑这一点时, 你也许想参考十二面体的模型或者图解.

令  $\theta = 2\pi/3$ . 二十面体群含有围绕顶点  $v$  转过角度  $\theta$  的旋转. 这个旋转有自旋  $(v, \theta)$ , 于是, 记之为  $\rho_{(v, \theta)}$ . 20 个顶点构成  $I$ -轨道, 且如果  $v'$  是另一个顶点, 则  $\rho_{(v, \theta)}$  与  $\rho_{(v', \theta)}$  是  $I$

的共轭元素. 这由推论 5.1.28(b) 可得. 诸顶点构成 20 阶的轨道, 所以, 所有旋转  $\rho_{(v, \theta)}$  都是共轭的. 它们是不同的, 因为像  $(v, \theta)$  那样定义同一旋转的仅有自旋是  $(-v, -\theta)$  与  $-\theta \neq \theta$ . 于是, 这些旋转构成 12 阶的共轭类.

其次,  $I$  含有绕面的中心转过角度  $2\pi/5$  的旋转, 且 12 个面构成一个轨道. 像上面推导的那样, 我们找到 12 阶的共轭类. 类似地, 转过角度  $4\pi/5$  的旋转构成 12 阶的共轭类.

最后,  $I$  含有绕边的中心转过角度  $\pi$  的旋转. 有 30 条边, 它们给出 30 个自旋  $(e, \pi)$ . 但  $\pi = -\pi$ . 如果  $e$  是边的中心, 则  $-e$  也是, 从而自旋  $(e, \pi)$  与  $(-e, -\pi)$  代表同一旋转. 这个共轭类仅含有 15 个不同的旋转.

198

二十面体群的类方程是

$$\text{【7.4.1】} \quad 60 = 1 + 20 + 12 + 12 + 15$$

注意 称  $(v, \theta)$  与  $(e, \pi)$  为自旋是不准确的, 因为  $v$  与  $e$  都没有单位长度. 但这点显然是不重要的.

## 单群

群  $G$  是单群, 如果它不是平凡群并且它不包含真的正规子群——除  $\langle 1 \rangle$  和  $G$  外没有别的正规子群. (这个“单”字并不意味着“不复杂”. 它在这里大致的意思为“不是合成的”.) 素阶循环群不含有真子群; 所以, 它们是单群. 除平凡群外, 所有其他群均含有真子群, 尽管不一定是真正规子群.

下列引理的证明可直接得出.

**【7.4.2】引理** 令  $N$  是群  $G$  的正规子群.

- (a) 如果  $N$  含有元素  $x$ , 则它含有  $x$  的共轭类  $C(x)$ .
- (b)  $N$  是共轭类的并.
- (c)  $N$  的阶是它所包含的共轭类的阶的和.

现在用类方程证明下列定理.

**【7.4.3】定理** 二十面体群  $I$  是单群.

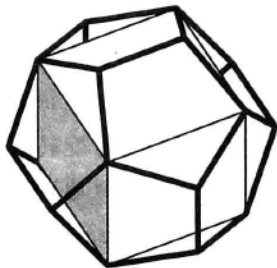
**证明** 二十面体群的真正规子群的阶是 60 的真因子, 根据引理, 它也是类方程 (7.4.1) 右边一些项的和, 包括项 1, 其为恒等元共轭类的阶. 没有整数满足这两个条件. 这就证明了定理. ■

单的性质是很有用的, 因为可能偶然碰到正规子群, 就像下面定理所展示的.

**【7.4.4】定理** 二十面体群与交错群  $A_5$  同构. 所以,  $A_5$  是单群.

**证明** 为描述这个同构, 我们需要找到一个  $I$  在其上面作用的有五个元素的集合  $S$ . 这是相当微妙的, 但这样一个集合由内接于十二面体的五个立方体组成, 其中一个的图示如下:

## 【7.4.5】图



十二面体的一个内接立方体.

二十面体群在这五个立方体的集合  $S$  上作用, 并且这个作用定义了一个相伴的置换表示的同态  $\varphi: I \rightarrow S_5$ . 我们证明  $\varphi$  定义了从  $I$  到交错群  $A_5$  的同构. 为证明它是一个同构, 我们将用到  $I$  是单群这个事实, 但需要的关于作用的仅有信息是它不是平凡的.

199

$\varphi$  的核是  $I$  的正规子群. 由于  $I$  是单群, 故  $\varphi$  的核或是平凡群  $\langle 1 \rangle$  或是整个群  $I$ . 如果核是整个群, 则  $I$  在五个立方体上的作用是平凡作用, 但这是不对的. 因此,  $\text{Ker}\varphi = \langle 1 \rangle$ . 这表明  $\varphi$  是单射. 于是, 它定义了一个从  $I$  到它在  $S_5$  中的像的一个同构.

接下来, 我们把同态  $\varphi$  与符号同态  $\sigma: S_5 \rightarrow \{\pm 1\}$  合成, 得到一个同态  $\sigma\varphi: I \rightarrow \{\pm 1\}$ . 如果这个同态是满射, 则其核将是  $I$  的真正子群. 由于  $I$  是单群, 故这是不可能的. 因而这个限制是个平凡同态, 这意味着  $\varphi$  的像包含在  $\sigma$  的核里, 这个核为交错群  $A_5$ .  $I$  与  $A_5$  的阶都是 60, 且  $\varphi$  是单射, 所以  $\varphi$  同构于  $I$  的像是  $A_5$ . ■

## 第五节 对称群里的共轭

在对称群里刻画共轭的最不会引起混淆的方法是考虑重新给指标标号. 如果给定指标是 1, 2, 3, 4, 5, 且如果把它们分别标为  $a, b, c, d, e$ , 则置换  $p = (134)(25)$  变为  $(acd)(be)$ .

要写这个过程的公式, 令  $\varphi: I \rightarrow L$  表示从指标集  $I$  到字母集  $L$  的重新标号映射:  $\varphi(1) = a, \varphi(2) = b$ , 等等. 这样, 重新标号的置换是  $\varphi \circ p \circ \varphi^{-1}$ . 其解释如下:

首先用  $\varphi^{-1}$  映射字母到指标.

其次, 用  $p$  置换指标.

最后, 用  $\varphi$  把指标映射回字母.

可用诸指标的置换  $q$  以同样方法重新标号. 结果(即共轭  $p' = qpq^{-1}$ )将是同一指标集新的置换. 例如, 如果用  $q = (1452)$  重新标号, 可得

$$qpq^{-1} = (1452) \circ (134)(25) \circ (2541) = (435)(12) = p'$$

200

有两点要注意. 第一, 重新标号将产生一个置换, 其循环与原来的有相同长度. 第二, 通过选取适当的置换  $q$ , 可得到任意其他具有相同长度的循环的置换. 如果把一个置换写在另一个置换上面, 并指定顺序使得循环对应, 则可用这个结果作为数表定义  $q$ . 例

如, 像我们上面所做的那样, 作为置换  $p = (1\ 3\ 4)(2\ 5)$  的共轭得到  $p' = (4\ 3\ 5)(1\ 2)$ , 我们可写

$$\frac{(1\ 3\ 4)(2\ 5)}{(4\ 3\ 5)(1\ 2)}$$

通过从上往下读这个数表:  $1 \rightsquigarrow 4$ , 等等, 可得到重新标号置换  $q$ .

因为循环可从它的任一指标开始, 所以常常是几个置换  $q$  产生同一共轭.

下一个命题总结了上面的讨论.

**【7.5.1】命题** 两个置换  $p$  与  $p'$  是对称群里的共轭元当且仅当它们的循环分解有相同的阶.

我们用命题 7.5.1 确定对称群  $S_4$  的类方程. 置换的循环分解划分集合  $\{1, 2, 3, 4\}$ . 划分 4 的子集的阶可能是

$$1, 1, 1, 1; \quad 2, 1, 1; \quad 2, 2; \quad 3, 1 \text{ 或 } 4$$

具有这些阶的循环的置换分别是恒等置换、对换、(不相交的)对换的积、3-循环与 4-循环.

有 6 个对换、3 个对换的积、8 个 3-循环与 6 个 4-循环. 由命题知, 这些集合的每一构成一个共轭类, 所以,  $S_4$  的类方程是

$$\text{【7.5.2】} \quad 24 = 1 + 3 + 6 + 6 + 8$$

类似的计算表明对称群  $S_5$  的类方程是

$$\text{【7.5.3】} \quad 120 = 1 + 10 + 15 + 20 + 20 + 30 + 24$$

在前一节(7.4.4)里我们看到交错群  $A_5$  是单群, 因为它同构于二十面体群  $I$ , 而  $I$  是单群. 我们现在证明大多数交错群都是单群.

**【7.5.4】定理** 对每个  $n \geq 5$ , 交错群  $A_n$  是单群.

为完整起见, 我们注意  $A_2$  是平凡群,  $A_3$  是 3 阶循环群,  $A_4$  不是单群. 由恒等置换和对换的 3 个积  $(1\ 2)(3\ 4)$ ,  $(1\ 3)(2\ 4)$ ,  $(1\ 4)(2\ 3)$  构成的 4 阶群是  $S_4$  与  $A_4$  的正规子群(见(2.5.13)(b)).

**【7.5.5】引理**

- (a) 对  $n \geq 3$ , 交错群  $A_n$  是由 3-循环生成的.  
 (b) 对  $n \geq 5$ , 3-循环在交错群  $A_n$  里构成唯一的共轭类.

201

**证明**

(a) 这与行约简方法类似. 比如说, 不是恒等置换的偶置换  $p$  固定  $m$  个指标不动, 我们证明如果用一个适当的 3-循环  $q$  左乘  $p$ , 则乘积  $qp$  至少固定  $m+1$  个指标不动. 对  $m$  进行归纳证明.

如果  $p$  不是恒等置换, 则它要么含有一个  $k$ -循环, 其中  $k \geq 3$ , 要么含有两个 2-循环的积. 由于如何给指标编号不碍事, 故可设  $p = (1\ 2\ 3 \cdots k) \cdots$  或  $p = (1\ 2)(3\ 4) \cdots$ . 令  $q = (3\ 2\ 1)$ . 乘积  $qp$  固定指标 1 不动以及固定由  $p$  固定不动的所有指标不动.

(b) 假设  $n \geq 5$ , 且令  $q = (1\ 2\ 3)$ . 根据命题 7.5.1, 3-循环在对称群  $S_n$  里是共轭的.

所以, 如果  $q'$  是另一个 3-循环, 则存在一个置换  $p$  使得  $pqp^{-1}=q'$ . 如果  $p$  是偶置换, 则  $q$  与  $q'$  在  $A_n$  中是共轭的. 假设  $p$  是奇置换. 因为  $n \geq 5$ , 故对换  $\tau = (4\ 5)$  属于  $S_n$ , 于是,  $\tau q \tau^{-1} = q$ . 这样,  $p\tau$  是偶的, 且  $(p\tau)q(p\tau)^{-1} = q'$ . ■

**定理 7.5.4 的证明** 我们现在进行定理的证明. 令  $N$  是交错群  $A_n$  的非平凡正规子群, 其中  $n \geq 5$ . 必须证明  $N$  是整个群  $A_n$ . 只要证明  $N$  含有 3-循环就够了. 如果证明了这个结论, 则由 (7.5.5)(b) 可知  $N$  含有每个 3-循环, 且由 (7.5.5)(a) 知  $N = A_n$ .

已知  $N$  是正规子群, 且含有不是恒等置换的置换  $x$ , 故在  $N$  中可做三种运算: 乘法、逆和共轭. 例如, 如果  $g$  是  $A_n$  的任一元素, 则  $gxg^{-1}$  与  $x^{-1}$  也属于  $N$ . 它们的乘积, 即交换子  $gxg^{-1}x^{-1}$  也属于  $N$ . 因为  $g$  是任意的, 故这些交换子给出许多属于  $N$  的元素.

我们的第一步是注意  $x$  的适当幂的阶是素数, 比如说, 阶  $\ell$ . 用这个幂替换  $x$ , 所以可假设  $x$  的阶是  $\ell$ . 这样,  $x$  的循环分解由  $\ell$ -循环和 1-循环组成.

不幸的是, 余下的证明需要分几个情形来进行. 在每一情形里, 计算交换子  $gxg^{-1}x^{-1}$ , 希望导出一个 3-循环. 适当的元素可通过实验找到.

情形 1:  $x$  的阶  $\ell \geq 5$ .

如何给指标编号是没关系的, 故可假设  $x$  含有  $\ell$ -循环  $(1\ 2\ 3\ 4\ 5 \cdots \ell)$ , 比如说  $x = (1\ 2\ 3\ 4\ 5 \cdots \ell)y$ , 其中  $y$  是剩余指标的置换. 设  $g = (4\ 3\ 2)$ . 于是

$gxg^{-1}x^{-1} = [(4\ 3\ 2)] \circ [(1\ 2\ 3\ 4\ 5 \cdots \ell)y] \circ [(2\ 3\ 4)]^{\circ} [y^{-1} (\overset{\text{先做这个}}{\ell \cdots 5\ 4\ 3\ 2\ 1})] = (2\ 4\ 5)$ . 因此, 这个交换子是 3-循环.

情形 2:  $x$  的阶为 3.

如果  $x$  是 3-循环, 就没有什么可证的了. 如果不是, 则  $x$  至少含有两个 3-循环, 比如说  $x = (1\ 2\ 3)(4\ 5\ 6)y$ . 设  $g = (4\ 3\ 2)$ . 则  $gxg^{-1}x^{-1} = (1\ 5\ 2\ 4\ 3)$ . 该交换子阶为 5. 我们回到情形 1.

202

情形 3a:  $x$  阶为 2, 且它含有 1-循环.

因为它是偶置换, 故  $x$  一定至少含有两个 2-循环, 比如说  $x = (1\ 2)(3\ 4)(5)y$ . 令  $g = (5\ 3\ 1)$ . 则  $gxg^{-1}x^{-1} = (1\ 5\ 2\ 4\ 3)$ . 该交换子阶为 5, 我们又回到情形 1.

情形 3b:  $x$  的阶  $\ell = 2$ , 且它不含有 1-循环.

因为  $n \geq 5$ , 故  $x$  含有多于两个的 2-循环. 比如说  $x = (1\ 2)(3\ 4)(5\ 6)y$ . 令  $g = (5\ 3\ 1)$ . 则  $gxg^{-1}x^{-1} = (1\ 5\ 3)(2\ 4\ 6)$ . 该交换子的阶为 3, 故我们回到情形 2.

这些是具有素数阶偶置换的所有可能, 所以, 定理的证明就完成了. ■

## 第六节 正规化子

考虑群  $G$  的子群  $H$  对于由  $G$  的共轭作用的轨道.  $[H]$  的轨道是共轭子群  $[gHg^{-1}]$  的集合, 其中  $g \in G$ .  $[H]$  关于这个作用的稳定子称为  $H$  的正规化子, 记为  $N(H)$ :

**【7.6.1】** 
$$N(H) = \{g \in G \mid gHg^{-1} = H\}$$

计数公式为

$$[7.6.2] \quad |G| = |N(H)| \cdot (\text{共轭子群数})$$

共轭子群数等于指标 $[G:N(H)]$ .

**[7.6.3] 命题** 令  $H$  是群  $G$  的子群, 设  $N$  是  $H$  的正规化子.

- (a)  $H$  是群  $N$  的正规子群.
- (b)  $H$  是群  $G$  的正规子群当且仅当  $N=G$ .
- (c)  $|H|$  整除  $|N|$  且  $|N|$  整除  $|G|$ .

例如, 令  $H$  是对称群  $S_5$  的由元素  $p=(1\ 2)(3\ 4)$  生成的 2 阶循环子群. 共轭类  $C(p)$  含有 15 对不相交对换, 每一个都生成  $H$  的共轭子群. 由计数公式知正规化子  $N(H)$  阶为 8, 即  $120=8 \cdot 15$ .

## 第七节 西罗定理

西罗定理描述了任意有限群的素数幂阶的子群. 它们用挪威数学家西罗命名. 西罗在 19 世纪发现了这些定理.

设  $G$  是阶为  $n$  的群, 设  $p$  是一个整除  $n$  的素整数. 令  $p^e$  表示  $p$  的整除  $n$  的最大的幂, 这样

$$[7.7.1] \quad n = p^e m$$

其中  $m$  是不能被  $p$  整除的整数.  $G$  的  $p^e$  阶的子群  $H$  称为  $G$  的西罗  $p$ -子群. 西罗  $p$ -子群是在群里其指标不能为  $p$  所整除的  $p$ -群.

203

**[7.7.2] 定理(西罗第一定理)** 其阶为素数  $p$  整除的有限群包含一个西罗  $p$ -子群.

西罗定理的证明放在本节最后.

**[7.7.3] 推论** 其阶为素数  $p$  整除的有限群包含一个  $p$  阶的元素.

**证明** 令  $G$  是这样一个群, 设  $H$  为  $G$  的西罗  $p$ -子群. 于是,  $H$  包含一个不同于 1 的元素  $x$ .  $x$  的阶整除  $H$  的阶, 所以, 它是  $p$  的正幂, 比如说  $p^k$ . 因此,  $x^{p^{k-1}}$  阶为  $p$ . ■

这个推论不是显然的. 我们已经知道任意元素的阶整除群的阶, 但可以想象, 比如说一个 6 阶群由单位元 1 和 5 个 2 阶元素组成. 这样的群是不存在的. 6 阶群必须含有一个 3 阶元素和一个 2 阶元素.

余下的西罗定理给出西罗子群的附加信息.

**[7.7.4] 定理(西罗第二定理)** 令  $G$  是其阶为素数  $p$  整除的有限群.

- (a)  $G$  的西罗  $p$ -子群是共轭子群.
- (b)  $G$  的每一个为  $p$ -群的子群都包含在一个西罗  $p$ -子群里.

西罗  $p$ -子群的共轭子群也是西罗  $p$ -子群.

**[7.7.5] 推论** 群  $G$  仅仅有一个西罗  $p$ -子群当且仅当这个子群是正规的.

**[7.7.6] 定理(西罗第三定理)** 令  $G$  是其阶  $n$  为素数  $p$  整除的有限群. 比如说  $n=p^e m$ , 其中  $p$  不能整除  $m$ , 设  $s$  表示西罗  $p$ -子群的个数. 则  $s$  整除  $m$ , 且  $s$  模  $p$  与 1 同余:  $s \equiv 1 \pmod{p}$

1 对某个整数  $k \geq 0$  成立.

在证明西罗定理之前, 我们将用它们对 6 阶群、15 阶群和 21 阶群进行分类. 这些例子展示这些定理的威力, 但当  $n$  有许多因子时,  $n$  阶群的分类是不容易的, 因为太多的可能.

### 【7.7.7】命题

(a) 每个 15 阶群都是循环群.

(b) 有两个 6 阶群的同构类, 即循环群  $C_6$  的类与对称群  $S_3$  的类.

(c) 有两个 21 阶群的同构类: 循环群  $C_{21}$  的类与由两个元素  $x$  和  $y$  生成的群  $G$  的类,

204 其中  $x$  和  $y$  满足关系  $x^7=1$ ,  $y^3=1$ ,  $yx=x^2y$ .

证明

(a) 令  $G$  是 15 阶群. 根据西罗第三定理, 它的西罗 3-子群的个数整除 5, 且模 3 与 1 同余. 仅有的这样的整数是 1. 所以, 有一个西罗 3-子群, 比如说  $H$ , 它是一个正规子群. 同样理由, 仅有一个西罗 5-子群, 比如说  $K$ , 它是正规的. 子群  $H$  是 3 阶循环子群,  $K$  是 5 阶循环子群. 交  $H \cap K$  是平凡群. 由命题 2.11.4(d) 可知  $G$  同构于积群  $H \times K$ . 所以, 所有 15 阶群都同构于循环群的积  $C_3 \times C_5$ , 且它们相互同构. 循环群  $C_{15}$  是一个这样的群, 所以, 所有 15 阶群都是循环群.

(b) 令  $G$  是 6 阶群. 由西罗第一定理,  $G$  包含一个西罗 3-子群  $H$  (为 3 阶循环群) 和一个西罗 2-子群  $K$  (为 2 阶循环子群). 由西罗第三定理可知西罗 3-子群的个数整除 2, 且模 3 与 1 同余. 仅有的这样的整数是 1. 所以, 有一个西罗 3-子群  $H$ , 它是一个正规子群. 同样的定理还告诉我们西罗 2-子群的个数整除 3, 且模 2 与 1 同余. 这个数或是 1 或是 3.

情形 1:  $H$  与  $K$  都是正规子群.

像前面的例子一样,  $G$  同构于积群  $H \times K$ , 这个积群是阿贝尔群. 所有 6 阶阿贝尔群都是循环群.

情形 2:  $G$  包含 3 个西罗 2-子群, 比如说  $K_1, K_2, K_3$ .

群  $G$  由共轭作用于阶为 3 的集合  $S = \{[K_1], [K_2], [K_3]\}$ , 这给出从  $G$  到对称群的一个同态  $\varphi: G \rightarrow S_3$ , 它是伴随的置换表示 (6.11.2). 由西罗第二定理知, 在  $S$  上的这个作用是可迁的, 所以, 元素  $[K_i]$  在  $G$  中的稳定子是正规化子  $N(K_i)$ , 阶为 2. 它等于  $K_i$ . 因为  $K_1 \cap K_2 = \{1\}$ , 所以恒等元是  $G$  的固定  $S$  的所有元素不动的唯一元素. 作用是忠实的, 且置换表示  $\varphi$  是单射. 因为  $G$  与  $S_3$  有相同的阶, 故  $\varphi$  是同构.

(c) 令  $G$  是 21 阶群. 西罗第三定理表明西罗 7-子群  $K$  一定是正规的, 且西罗 3-子群的个数是 1 或 7. 令  $x$  是  $K$  的生成元, 且设  $y$  是一个西罗 3-子群  $H$  的生成元. 这样,  $x^7=1$ ,  $y^3=1$ , 所以,  $H \cap K = \{1\}$ , 因此, 积映射  $H \times K \rightarrow G$  是单射 (2.11.4)(a). 由于  $G$  的阶为 21, 故积映射是双射.  $G$  的元素是乘积  $x^i y^j$ , 其中  $0 \leq i < 7$ ,  $0 \leq j < 3$ .

因为  $K$  是正规子群, 故  $yx y^{-1}$  是  $K$  的元素, 它为  $x$  的幂, 比如说  $x^i$ , 其中  $1 \leq i < 7$ . 所以, 元素  $x$  与  $y$  满足关系

【7.7.8】  $x^7 = 1, \quad y^3 = 1, \quad yx = x^i y$

这些关系足以确定群的乘法表. 然而, 关系  $y^3 = 1$  限制了可能的幂指数  $i$ , 因为它蕴含着  $y^3xy^{-3} = x$ :

$$x = y^3xy^{-3} = y^2x^iy^{-2} = yx^{i^2}y^{-1} = x^{i^3}$$

所以,  $i^3 \equiv 1 \pmod{7}$ . 这告诉我们  $i$  一定是 1, 2 或 4.

205

例如, 幂指数  $i=3$  蕴含着  $x = x^3 = x^6 = x^{-1}$ . 这样,  $x^2 = 1$ , 也有  $x^7 = 1$ , 因此,  $x = 1$ . 由关系 (7.7.8) 定义的群是由  $y$  生成的 3 阶循环群, 其中  $i=3$ .

情形 1:  $yxy^{-1} = x$ . 这样,  $x$  与  $y$  交换.  $H$  与  $K$  都是正规子群. 同前面一样,  $G$  同构于 3 阶循环群与 7 阶循环群的直积, 从而是循环群.

情形 2:  $yxy^{-1} = x^2$ . 像上面所注解的那样, 乘法表被确定. 但我们仍需证明这个群实际存在. 这就要证明关系坍缩成群, 像  $i=3$  时所发生的那样. 我们会学习一个系统方法来做事, 这个方法是 Todd-Coxeter 算法, 见本章第十一节. 另一个方法是简洁地展示群, 例如以矩阵群为例. 为此, 需要做个实验.

由于假设我们正在寻找的群包含 7 阶元, 因此自然试图寻找合适的矩阵, 其中元素都是模 7 的. 至少我们能写下元素在  $\mathbb{F}_7$  中阶为 7 的  $2 \times 2$  矩阵, 即下面这样的矩阵  $x$ . 这样,  $y$  可通过反复试验找到. 矩阵

$$x = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} \quad \text{与} \quad y = \begin{bmatrix} 2 & \\ & 1 \end{bmatrix}$$

的元素属于  $\mathbb{F}_7$ , 满足关系  $x^7 = 1$ ,  $y^3 = 1$ ,  $yx = x^2y$ , 且它们生成 21 阶群.

情形 3:  $yxy^{-1} = x^4$ . 这样,  $y^2xy^{-2} = x^2$ . 注意  $y^2$  也是 3 阶元素. 所以, 用  $y^2$  替换  $y$ , 这是  $H$  的另一个生成元. 结果是幂指数 4 由 2 替换, 这又回到前面的情形.

因此, 如同所断言的, 有两个 21 阶群的同构类. ■

在西罗第一定理的证明中用到两个引理.

**[7.7.9] 引理** 令  $U$  是群  $G$  的子集.  $[U]$  对于由  $G$  在它的子集的集合上的左乘作用的稳定子  $\text{Stab}([U])$  的阶整除阶  $|U|$  与  $|G|$ .

**证明** 如果  $H$  是  $G$  的子群, 则  $G$  的元素  $u$  对于由  $H$  左乘的  $H$ -轨道是右陪集  $Hu$ . 令  $H$  是  $[U]$  的稳定子. 则由  $H$  确定的乘法置换  $U$  的元素, 所以,  $U$  划分成  $H$ -轨道, 它们均是右陪集. 每个陪集的阶均为  $|H|$ , 所以,  $|H|$  整除  $|U|$ . 因为  $H$  是子群, 故  $|H|$  整除  $|G|$ . ■

**[7.7.10] 引理** 令  $n$  是形如  $p^e m$  的整数, 其中  $e > 0$  且  $p$  不整除  $m$ . 在阶为  $n$  的集合里  $p^e$  阶的子集个数  $N$  不被  $p$  整除.

**证明** 数  $N$  是二次项系数

$$\binom{n}{p^e} = \frac{n(n-1)\cdots(n-k)\cdots(n-p^e+1)}{p^e(p^e-1)\cdots(p^e-k)\cdots 1}$$

206

$N \not\equiv 0 \pmod{p}$  的理由是每当  $p$  整除  $N$  的分子上的项  $(n-k)$  时, 它也是整除分母上的项  $(p^e-k)$  恰好同样数量的次数: 如果将  $k$  写为  $k = p^i \ell$  的形式, 其中  $p$  不整除  $\ell$ , 则  $i < e$ . 所以,  $(m-k) = (p^e-k)$  和  $(n-k) = (p^e m - k)$  都被  $p^i$  整除而不能被  $p^{i+1}$  整除. ■



**西罗第一定理的证明** 设 $S$ 是 $G$ 的所有 $p^e$ 阶子集的集合. 这些子集中有一个是西罗子群, 但我们不是直接求出它, 而是考虑 $G$ 在 $S$ 上的左乘作用. 我们将证明这些 $p^e$ 阶子集 $[U]$ 中有一个 $p^e$ 阶稳定子. 这个稳定子就是所求的子群.

我们将 $S$ 分解为左乘作用的轨道, 得到公式

$$N = |S| = \sum_{\text{轨道 } O} |O|$$

根据引理 7.7.10,  $p$  不整除  $N$ . 于是, 至少有一个轨道具有不为  $p$  所整除的阶, 设它是子集  $[U]$  的轨道  $O_{[U]}$ . 令  $H$  是  $[U]$  的稳定子. 由引理 7.7.9 知,  $H$  的阶整除  $U$  的阶, 而  $U$  的阶是  $p^e$ . 于是,  $|H|$  是  $p$  的幂. 我们有  $|H| \cdot |O_{[U]}| = |G| = p^e m$ , 且  $|O_{[U]}|$  不为  $p$  整除. 所以,  $|O_{[U]}| = m$ ,  $|H| = p^e$ . 于是,  $H$  是西罗  $p$ -子群. ■

**西罗第二定理的证明** 给定一个群  $G$  的  $p$ -子群  $K$  和一个西罗  $p$ -子群  $H$ , 我们要证明  $H$  的某个共轭子群  $H'$  包含  $K$ , 这就证明了 (b). 如果  $K$  也是  $p$ -子群, 它将等于共轭子群  $H'$ , 这样 (a) 也得证.

选取一个群  $G$  作用的子集  $C$ , 它具有如下性质:  $p$  不整除阶  $|C|$ , 作用是可迁的, 且  $C$  包含一个元素  $c$ , 其稳定子是  $H$ .  $G$  中  $H$  的左陪集的集合具有这些性质, 于是, 这样的集合存在. (借助陪集我们没有过多使用记号.)

我们将  $G$  在  $C$  上的作用限制到  $p$ -群  $K$  上. 因为  $p$  不整除  $|C|$ , 故对  $K$  的作用存在不动点  $c'$ . 这是不动点定理 7.3.2. 因为  $G$  的作用是可迁的, 故对某个  $g \in G$  有  $c' = gc$ .  $c'$  的稳定子是  $H$  的共轭子群  $gHg^{-1}$  (命题 6.7.7). 因为  $K$  固定  $c'$  不动, 所以, 稳定子包含  $K$ . ■

**西罗第三定理的证明** 同以前一样, 写  $|G| = p^e m$ . 设  $s$  表示西罗  $p$ -子群的个数. 由西罗第二定理可知,  $G$  在西罗  $p$ -子群的集合  $S$  上的作用是可迁的. 一个特殊西罗  $p$ -子群  $[H]$  的稳定子是  $H$  的正规化子  $N = N(H)$ . 由计数公式知  $S$  的阶 (为  $s$ ) 等于指标  $[G:N]$ . 因为  $N$  包含  $H$  (7.6.3), 且  $[G:N]$  等于  $m$ , 故  $s$  整除  $m$ .

其次, 把集合  $S$  分解为由  $H$  的共轭作用的轨道.  $[H]$  的  $H$ -轨道的阶为 1. 由于  $H$  是  $p$ -群, 故任意  $H$ -轨道的阶是  $p$  的幂. 要证  $s \equiv 1 \pmod{p}$ , 我们证明除  $[H]$  外, 没有  $S$  的元素被  $H$  固定不动.

假设  $H'$  是西罗  $p$ -子群且由  $H$  的共轭固定  $[H']$  不动. 这样,  $H$  包含在  $H'$  的正规化子  $N'$  中, 所以,  $H$  与  $H'$  是  $N'$  的西罗  $p$ -子群. 由西罗第二定理知,  $N'$  的西罗  $p$ -子群是  $N'$  的共轭子群. 但  $H'$  是  $N'$  的正规子群 (命题 7.6.3(a)). 所以,  $H' = H$ . ■

207

## 第八节 12 阶群

我们用西罗定理对 12 阶群进行分类. 这个定理用来展示这一事实: 当阶有若干因子时, 对群进行分类变得复杂.

**[7.8.1] 定理** 存在 5 个 12 阶群的同构类. 它们的代表是:

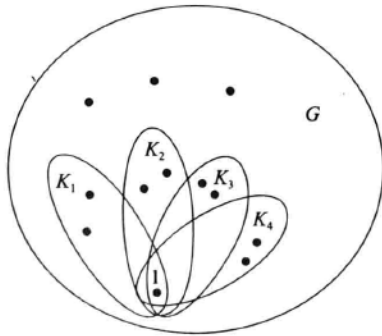
- 循环群的积  $C_4 \times C_3$ ,

- 循环群的积  $C_2 \times C_2 \times C_3$ ,
- 交错群  $A_4$ ,
- 二面体群  $D_6$ ,
- 由元素  $x$  与  $y$  生成的群, 满足关系  $x^4=1, y^3=1, xy=y^2x$ .

除最后一个外, 这些群都是众所熟知的. 积群  $C_4 \times C_3$  同构于  $C_{12}$ ,  $C_2 \times C_2 \times C_3$  同构于  $C_2 \times C_6$  (见命题 2.11.3).

**证明** 设  $G$  是 12 阶的群, 用  $H$  表示  $G$  的一个西罗 2-子群, 其阶为 4, 用  $K$  表示  $G$  的一个阶为 3 的西罗 3-子群. 由西罗第三定理知西罗 2-子群的个数或是 1 或是 3, 西罗 3-子群的个数或是 1 或是 4. 而且,  $H$  是 4 阶群, 所以, 它或者是循环群  $C_4$ , 或者是克莱因四元群  $C_2 \times C_2$  (命题 2.11.5). 当然,  $K$  是循环群.

尽管对于证明不是必要的, 但还是从证明两个子群至少有一个 ( $H$  或者  $K$ ) 是正规子群开始. 如果  $K$  不是正规子群, 则存在 4 个西罗 3-子群与  $K$  共轭, 比如说,  $K_1, \dots, K_4$ , 其中  $K_1=K$ . 这些群的阶是素数阶, 所以, 它们之中任意两个的交是平凡群  $\langle 1 \rangle$ . 这样, 仅有  $G$  的 3 个元素不属于任一个群  $K_i$ . 这个事实的图解如下.



西罗 2-子群  $H$  的阶为 4, 且  $H \cap K_i = \langle 1 \rangle$ . 所以,  $H$  由 3 个不属于任一个群  $K_i$  的元素与 1 组成. 这就为我们描述了任一个群  $H$ , 并证明了仅有一个西罗 2-子群. 因此,  $H$  是正规的.

其次, 我们注意到  $H \cap K = \langle 1 \rangle$ , 所以, 积映射  $H \times K \rightarrow G$  是集合的双射 (2.11.4).  $G$  的每个元素均有唯一的表示形式  $hk$ , 其中  $h$  属于  $H$ ,  $k$  属于  $K$ .

208

情形 1:  $H$  与  $K$  都是正规的.

这样,  $G$  同构于积群  $H \times K$  (2.11.4). 因为对  $H$  有两种可能, 对  $K$  有一种可能, 故对  $G$  有两种可能:

$$G \approx C_4 \times C_3 \quad \text{或} \quad G \approx C_2 \times C_2 \times C_3$$

这些是 12 阶阿贝尔群.

情形 2:  $K$  不是正规的. 则存在 4 个共轭的西罗 3-子群  $K_1, \dots, K_4$ , 且  $G$  通过共轭作用在这 4 个子群的集合  $S$  上. 这一作用确定一个置换表示, 即一个到对称群的同态  $\varphi: G \rightarrow S_4$ . 我们将证明  $\varphi$  将  $G$  同构地映射到交错群  $A_4$  上.

$K_i$  的正规化子  $N_i$  包含  $K_i$ , 由计数公式知  $|N_i| = 3$ . 所以,  $N_i = K_i$ . 因为诸子群  $K_i$  的唯一公共元素是恒等元, 故只有恒等元稳定所有这些子群. 因此,  $G$  的作用是忠实的,  $\varphi$  是单射, 且  $G$  同构于它在  $S_4$  中的像.

因为  $G$  有 4 个 3 阶子群, 故它含有 8 个 3 阶元素. 它们的像是  $S_4$  中的 3-循环, 且生成  $A_4$  (7.5.5). 所以,  $G$  的像包含  $A_4$ . 因为  $G$  与  $A_4$  有相同的阶, 故像等于  $A_4$ .

情形 3:  $K$  正规但  $H$  不正规.

这样,  $H$  通过共轭在  $K = \{1, y, y^2\}$  上作用. 因为  $H$  不是正规的, 故它含有不与  $y$  交换的元素  $x$ , 于是,  $xyx^{-1} = y^2$ .

情形 3a:  $K$  正规但  $H$  不正规, 且  $H$  是循环群.

元素  $x$  生成  $H$ , 所以,  $G$  由元素  $x$  与元素  $y$  生成, 且具有关系

$$\text{[7.8.2]} \quad x^4 = 1, y^3 = 1, xy = y^2x$$

这些关系确定了  $G$  的乘法表, 所以, 至多存在一个这样群的同构类. 但我们必须证明这些关系进一步构成群, 像 21 阶群一样 (见 7.7.8), 它是用矩阵表示群的最简便的方式. 这里我们使用复矩阵. 令  $\omega$  是单位复立方根  $e^{2\pi i/3}$ . 复矩阵

$$\text{[7.8.3]} \quad x = \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}, \quad y = \begin{bmatrix} \omega & \\ & \omega^2 \end{bmatrix}$$

满足这 3 个关系, 且它们生成一个 12 阶群.

情形 3b:  $K$  正规但  $H$  不正规, 且  $H \approx C_2 \times C_2$ .

$y$  对于  $H$  在集合  $\{y, y^2\}$  上通过共轭作用的稳定子的阶为 2. 于是,  $H$  含有元素  $z \neq 1$ , 使得  $zy = yz$ , 且  $H$  含有元素  $x$  使得  $xy = y^2x$ . 因为  $H$  是阿贝尔群, 故  $xz = zx$ . 这样,  $G$  由 3 个元素  $x, y, z$  生成, 且具有关系

$$\text{[209]} \quad x^2 = 1, y^3 = 1, z^2 = 1, yz = zy, xz = zx, xy = y^2x$$

这些关系确定了群的乘法表, 所以, 至多存在一个这样群的同构类. 二面体群  $D_6$  不是前面描述的 4 个群之一, 所以它一定是这个群. 所以,  $G$  同构于  $D_6$ . ■

## 第九节 自由群

我们看到可用通常的生成元  $x$  与  $y$  以及关系  $x^3 = 1, y^2 = 1, yx = x^2y$  来计算对称群  $S_3$ . 在本章的余下部分, 我们研究其他群里的生成元与关系.

我们首先考虑有生成元的群, 除了由群的公理所给出的那些关系 (诸如结合律) 外, 这些群不满足任何其他关系. 群的一个集合  $S$  称为是自由的, 如果其元素除了群的公理给出的关系外不满足任何别的关系. 具有自由的生成元集合的群称为自由群.

为了刻画自由群, 我们从任意一个集合  $S$  开始, 比如  $S = \{a, b, c, \dots\}$ . 我们称这些元素为“符号”, 并定义一个字为一个允许重复的有限符号串. 例如,  $a, aa, ba$  与  $aaba$  是字. 两个字可以通过并置合成, 即边靠边地放置它们:

$$aa, ba \rightsquigarrow aaba$$

这是字集合  $W$  上合成的结合律. 在  $W$  里“空字”作为恒等元素, 用符号  $1$  表示它. 这样, 集合  $W$  成为了集合  $S$  上的自由半群. 它不是群, 因为它没有逆元素, 而逆元素的加入是一件稍复杂的事情.

令  $S'$  是由符号  $a$  与  $a'$  ( $a \in S$ ) 构成的集合:

$$\text{【7.9.1】} \quad S' = \{a, a^{-1}, b, b^{-1}, c, c^{-1}, \dots\}$$

设  $W'$  是用  $S'$  里的符号构成的字的半群. 如果对于每个  $x \in S$ , 一个字看起来是

$$\dots xx^{-1} \dots \quad \text{或} \quad \dots x^{-1} x \dots$$

则可约定消去两个符号  $x$  和  $x^{-1}$  以缩短字的长度. 不能这样消去的字称为约化字. 从  $W'$  中任意字  $w$  开始, 可以执行一有限序列的消去而最终得到约化字  $w_0$ , 它也可能是空字  $1$ . 我们把这个字  $w_0$  称为  $w$  的约化型.

常常会有不止一种方式执行消去. 例如, 从  $w = abb^{-1}c^{-1}cb$  开始, 我们可用两种方式进行:

$$\begin{array}{ccc} a \underline{b} b^{-1} c^{-1} c \underline{b} & & \underline{a} b b^{-1} c^{-1} c b \\ \downarrow & & \downarrow \\ a c^{-1} c b & & \underline{a} b b^{-1} b \\ \downarrow & & \downarrow \\ ab & & ab \end{array}$$

最后得到相同的约化字, 尽管其符号来自原来的字的不同位置. (下划线的字母是最后剩下的字母.) 这总是对的.

**【7.9.2】命题** 一个给定的字  $w$  只有一个约化型.

210

**证明** 对字  $w$  的长度用归纳法. 如果  $w$  是约化字, 则没有什么需要证明的. 否则, 一定存在可以消去的某个符号对, 比如说, 划线字母对

$$w = \dots \underline{xx^{-1}} \dots$$

(我们用  $x$  表示  $S'$  中的任意元素, 并理解如果  $x = a^{-1}$ , 则  $x^{-1} = a$ .) 如果证明了通过消去符号对  $\underline{xx^{-1}}$  就可得到  $w$  的每个约化型, 则命题可用归纳法证明, 因为字  $\dots \underline{xx^{-1}} \dots$  较短.

设  $w_0$  是  $w$  的一个约化型, 它是由  $w$  通过一系列消去得到的. 第一种情形是我们的符号对  $\underline{xx^{-1}}$  在这个过程的某一步被消去. 如果这样, 就先消去  $\underline{xx^{-1}}$ . 于是, 这种情形就解决了. 另一方面, 因为  $w_0$  是约化的, 故符号对  $\underline{xx^{-1}}$  不会保留在  $w_0$  中. 因而两个符号中至少有一个在某个时候被消去. 如果这个符号对本身没有被消去, 则涉及该符号对的第一个消去一定是

$$\dots x^{-1} \underline{xx^{-1}} \quad \text{或} \quad \dots \underline{xx^{-1}} x^{-1} \dots$$

注意, 由这个消去得到的字与通过消去符号对  $\underline{xx^{-1}}$  得到的字是一样的. 于是, 可以在这一步用消去原来的符号对来代替. 这样, 我们回到了第一种情形, 于是命题得证. ■

我们称  $W'$  中的两个字  $w$  与  $w'$  是等价的, 并写作  $w \sim w'$ , 如果它们有相同的约化型. 这是一个等价关系.

**【7.9.3】命题** 等价的字的乘积是等价的: 如果  $w \sim w'$  与  $v \sim v'$ , 则  $wv \sim w'v'$ .

**证明** 要得到等价于乘积  $wv$  的约化字, 可首先将  $w$  与  $v$  中的字母尽可能多地消去, 而将  $w$  约化为  $w_0$ ,  $v$  约化为  $v_0$ . 这样,  $wv$  约化为  $w_0v_0$ . 现在, 我们继续约化  $w_0v_0$ , 直到这个字是约化字. 如果  $w \sim w'$ ,  $v \sim v'$ , 则同样的过程用于  $w'v'$  时, 也需经过  $w_0v_0$ , 从而给出同样的约化字. ■

由该命题得到字的等价类可以做乘积:

**[7.9.4] 命题**  $W'$  中的字的等价类的集合  $\mathcal{F}$  是一个群, 其中合成法则为  $W'$  中的乘法(置并).

**证明** 乘法是结合的且空字 1 的类是单位元这两个事实由  $W'$  中的相应事实得到. 还需验证  $\mathcal{F}$  的所有元素可逆. 但显然, 如果  $w$  是  $S'$  中元素的乘积  $xy \cdots z$ , 则  $z^{-1} \cdots y^{-1} x^{-1}$  是  $w$  的类的逆. ■

$S'$  中的字的等价类的群  $\mathcal{F}$  称为集合  $S$  上的自由群.  $\mathcal{F}$  的一个元素恰好对应于  $W'$  中的一个约化字. 要将约化字相乘, 先组合然后再消去:  $(abc^{-1})(cb) \rightsquigarrow abc^{-1}cb = abb$ .

幂记号也可使用:  $aaab^{-1}b^{-1} = a^3b^{-2}$ .

**注意** 一个元素集合  $S = \{a\}$  上的自由群只是一个无限循环群. 与之相比, 两个或两个以上元素的集合上的自由群是非常复杂的.

[211]

## 第十节 生成元与关系

描述了自由群后, 我们现在考虑更为一般的情形, 即群的生成元的集合不是自由的——它们中存在一些非平凡的关系.

**[7.10.1] 定义** 群  $G$  的元素  $x_1, \dots, x_n$  之间的关系  $R$  是集合  $\{x_1, \dots, x_n\}$  上自由群里的字  $r$ , 其在  $G$  里的值为 1. 我们将把这样的关系写为  $r$ , 或为了强调写为  $r=1$ .

例如, 正  $n$  边形的对称的二面体群  $D_n$  是由绕过角度  $2\pi/n$  的旋转  $x$  与反射  $y$  生成的. 这些生成元满足在 (6.4.3) 中列出的关系:

**[7.10.2]** 
$$x^n = 1, y^2 = 1, xyxy = 1$$

(最后一个关系常写成  $yx = x^{-1}y$ , 但最好在这里写出每个形如  $r=1$  的关系.)

可用这些关系以形式  $x^i y^j$  (其中  $0 \leq i < n, 0 \leq j < 2$ ) 写出  $D_n$  的元素, 从而可算出群的乘法表. 因此, 关系确定群. 所以, 它们叫做定义关系. 当关系比较复杂时, 确定群的元素和简洁乘法表是困难的, 但利用自由群与下一个引理, 我们将定义由给定元素集合与给定关系集合生成的群概念.

**[7.10.3] 引理** 令  $R$  是群  $G$  的子集, 则唯一存在  $G$  的包含  $R$  的最小正规子群  $N$ , 称为由  $R$  生成的正规子群. 如果  $G$  的正规子群包含  $R$ , 则它包含  $N$ .  $N$  的元素可以用下列方式的某一个加以描述:

(a)  $G$  的元素属于  $N$  如果它由  $R$  的元素通过有限多次乘法、逆和共轭运算得到.

(b) 令  $R'$  是由元素  $r$  与  $r^{-1}$  组成的集合, 其中  $r$  属于  $R$ ,  $G$  的元素属于  $N$  如果它可写成某个任意长度的乘积  $y_1 \cdots y_r$ , 其中每个  $y_i$  是  $R'$  的一个元素的共轭.

**证明** 令  $N$  表示由 (a) 中提到的一系列运算得到的元素集合. 一个非空子集是正规子

群当且仅当它在这些运算下是封闭的. 因为  $N$  在这些运算下封闭, 故它是正规子群. 而且, 包含  $R$  的任一正规子群一定含有  $N$ . 所以, 包含  $R$  的最小正规子群存在, 且等于  $N$ . 类似的推导可确定  $N$  是(b)中所描述的子集. ■

像通常一样, 我们必须照顾到空集. 我们说空集生成了平凡子群  $\{1\}$ .

**【7. 10. 4】定义** 令  $\mathcal{F}$  是一个集合  $S = \{x_1, \dots, x_n\}$  上的自由群, 设  $R = \{r_1, \dots, r_k\}$  是  $\mathcal{F}$  的元素集合. 由  $S$  生成的子群(其中关系  $r_1 = 1, \dots, r_k = 1$ )是商群  $G = \mathcal{F}/\mathcal{R}$ , 其中  $\mathcal{R}$  是  $\mathcal{F}$  的由  $R$  生成的正规子群.

212

群  $G$  常常记为

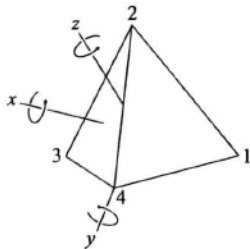
**【7. 10. 5】**  $\langle x_1, \dots, x_n \mid r_1, \dots, r_k \rangle$

因此, 二面体群  $D_n$  同构于群

**【7. 10. 6】**  $\langle x, y \mid x^n, y^2, xyxy \rangle$

**【7. 10. 7】例** 在正四面体的旋转对称的四面体群  $T$  里, 令  $x$  与  $y$  表示绕一个面的中心和一个顶点转过角度  $2\pi/3$  的旋转, 设  $z$  表示绕一个边的中心转过角度  $\pi$  的旋转, 如下图所示. 顶点标号如图所示,  $x$  在顶点上作用为置换  $(2\ 3\ 4)$ ,  $y$  作用为  $(1\ 2\ 3)$ ,  $z$  作用为  $(1\ 3)(2\ 4)$ . 计算这些置换的乘积表明  $xyz$  平凡地作用在顶点上. 因为唯一固定所有顶点不动的等距是恒等映射, 故  $xyz = 1$ .

**【7. 10. 8】图**



所以, 在四面体群里下列关系成立:

**【7. 10. 9】**  $x^3 = 1, y^3 = 1, z^2 = 1, xyz = 1$  ■

出现两个问题:

1. 这是  $T$  的定义关系的集合吗? 换言之, 群

**【7. 10. 10】**  $\langle x, y, z \mid x^3, y^3, z^2, xyz \rangle$

同构于  $T$  吗?

易证旋转  $x, y, z$  生成  $T$ , 但使用这些关系特别不容易. 没有重复的以生成元之积列出群的 12 个元素是很乱的. 下节证明答案是对的, 但我们不用写成群的元素来进行证明.

2. 在用生成元和关系表示的群  $G = \langle x_1, \dots, x_n \mid r_1, \dots, r_k \rangle$  里如何计算?

因为在自由群  $\mathcal{F}$  里容易计算, 所以唯一的问题是决定何时自由群的元素  $w$  表示  $G$  的恒等元素, 亦即何时  $w$  是子群  $\mathcal{R}$  的元素. 这是  $G$  的字问题. 如果我们能解决字问题, 则能够决定何时自由群的两个元素表示  $G$  的相等元素, 因为关系  $w_1 = w_2$  等价于  $w_1^{-1}w_2 = 1$ . 这使

213

我们能够计算.

字问题在任意有限群里可以解决, 但不是在每个群里都能解决. 然而, 我们将不讨论这一点, 因为某些工作需要给出字问题能或不能解决的确切意义. 如果你感兴趣, 可参见[Stillwell].

下一个例子表明 $\mathcal{R}$ 里的计算可能变得复杂, 甚至在相对简单的情形里亦如此.

**【7. 10. 11】例** 在群  $T$  里元素  $w = yxyx$  等于 1. 让我们证明  $w$  属于由 4 个关系(7. 10. 9)生成的正规子群 $\mathcal{R}$ . 采用你认为是标准的方法: 用允许的运算约化  $w$  为恒等元.

我们将使用的关系为  $z^2$  与  $xyz$ , 并分别用  $p$  与  $q$  记它们. 首先, 令  $w_1 = y^{-1}wy = xyxy$ . 因为 $\mathcal{R}$ 是正规子群, 故  $w_1$  属于 $\mathcal{R}$ 当且仅当  $w$  属于 $\mathcal{R}$ . 其次, 令  $w_2 = q^{-1}w_1 = z^{-1}xy$ . 因为  $q$  属于 $\mathcal{R}$ , 故  $w_2$  属 $\mathcal{R}$ 当且仅当  $w_1$  属于 $\mathcal{R}$ . 继续,  $w_3 = zw_2z^{-1} = xyz^{-1}$ ,  $w_4 = q^{-1}w_3 = z^{-1}z^{-1}$ ,  $pw_4 = 1$ . 往回求解,  $w = yqz^{-1}qp^{-1}zy^{-1}$  属于 $\mathcal{R}$ . 因此, 在群(7. 10. 10)里,  $w = 1$ . ■

我们回到由生成元和关系定义的群 $\mathcal{G}$ . 像任意商群一样, 我们有典范同态

$$\pi: \mathcal{F} \rightarrow \mathcal{F} / \mathcal{R} = \mathcal{G}$$

映字  $w$  为陪集  $\bar{w} = [w\mathcal{R}]$ ,  $\pi$  的核是 $\mathcal{R}$ (2. 12. 2). 为保持在所讨论的群里, 似乎在字母上放一个杠来记 $\mathcal{F}$ 的元素在 $\mathcal{G}$ 里的像比较妥当. 然而, 这不太令人习惯. 当在 $\mathcal{G}$ 里讨论时, 简单地记住自由群里的元素  $w_1$  与  $w_2$  在 $\mathcal{G}$ 里相等, 如果陪集  $w_1\mathcal{R}$  与  $w_2\mathcal{R}$  相等, 或者  $w_1^{-1}w_2$  属于 $\mathcal{R}$ .

因为定义关系  $r_i$  属于 $\mathcal{R}$ , 故  $r_i = 1$  在 $\mathcal{G}$ 里是对的. 如果把  $r_i$  写成字, 则因为  $\pi$  是同态, 故在 $\mathcal{G}$ 里对应的积将等于 1(见推论 2. 12. 3). 例如, 在群  $\langle x, y, z \mid x^3, y^3, z^2, xyz \rangle$  里,  $xyz = 1$  是对的.

我们再回到四面体群的例子和第一个问题. 群  $\langle x, y, z \mid x^3, y^3, z^2, xyz \rangle$  与  $T$  如何联系? 部分解释是基于自由群与商群的映射性质. 这两个性质是直观的. 它们的证明很简单, 将其留作练习.

**【7. 10. 12】命题(自由群的映射性质)** 令  $\mathcal{F}$  是一个集合  $S = \{a, b, \dots\}$  上的自由群, 并设  $G$  是一个群. 集合的任一映射  $f: S \rightarrow G$  以唯一的方式扩张成一个群同态  $\varphi: \mathcal{F} \rightarrow G$ . 如果把  $S$  的元素  $x$  的像  $f(x)$  记为  $\underline{x}$ , 则  $\varphi$  将  $S' = \{a, a^{-1}, b, b^{-1}, \dots\}$  的一个字映为  $G$  中元素  $\{\underline{a}, \underline{a}^{-1}, \underline{b}, \underline{b}^{-1}, \dots\}$  的对应的乘积.

这个性质反映了在 $\mathcal{F}$ 中除了群的公理推出的关系外  $S$  的元素不满足别的关系的事实. 它解释了形容词“自由”的原因.

**【7. 10. 13】命题(商群的映射性质)** 令  $\varphi: G' \rightarrow G$  是群同态,  $K$  为其核, 设  $N$  是  $G'$  的含于  $K$  的正规子群. 设  $\bar{G}' = G'/N$ , 且设  $\pi: G' \rightarrow \bar{G}'$  是典范映射  $a \mapsto \bar{a}$ . 规则  $\bar{\varphi}(\bar{a}) = \varphi(a)$  定义同态  $\bar{\varphi}: \bar{G}' \rightarrow G$ , 且  $\bar{\varphi} \circ \pi = \varphi$ .

$$\begin{array}{ccc} G' & \xrightarrow{\varphi} & G \\ & \searrow \pi & \nearrow \bar{\varphi} \\ & \bar{G}' & \end{array}$$

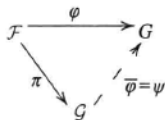
这个映射性质推广了第一同构定理. 当然,  $N$  含于核  $K$  里的假设是必要的.

下一个推论使用了之前引进的记号： $S = \{x_1, \dots, x_n\}$  是群  $G$  的子集， $R = \{r_1, \dots, r_k\}$  是  $G$  中  $S$  的元素之间关系的集合， $\mathcal{F}$  是  $S$  上的自由群，且  $\mathcal{R}$  是  $\mathcal{F}$  的由  $R$  生成的正规子群。最后， $\mathcal{G} = \langle x_1, \dots, x_n \mid r_1, \dots, r_k \rangle = \mathcal{F}/\mathcal{R}$ 。

**【7. 10. 14】推论**

- (i) 存在典型同态  $\psi: \mathcal{G} \rightarrow G$  映  $x_i \rightsquigarrow x_i$ 。
- (ii)  $\psi$  是满射当且仅当集合  $S$  生成  $G$ 。
- (iii)  $\psi$  是单射当且仅当  $S$  的元素之间的每个关系属于  $\mathcal{R}$ 。

**证明** 我们将证明 (i)，而略去 (ii) 与 (iii) 的证明。自由群的映射性质给出同态  $\varphi: \mathcal{F} \rightarrow G$  且  $\varphi(x_i) = x_i$ 。因为在  $G$  里关系  $r_i$  等于 1，故  $\mathcal{R}$  包含在  $\varphi$  的核  $K$  里。由于核是正规子群，故  $\mathcal{R}$  也包含在  $K$  里。这样，商群的映射性质给出映射  $\bar{\varphi}: \mathcal{G} \rightarrow G$ 。这是映射  $\psi$ ：



如果推论里刻画的映射  $\psi$  是双射，就说  $R$  构成生成元  $S$  之间关系的完备集。要确定这是否成立需要知道有关  $G$  的更多知识。回到四面体群，推论给出同态  $\psi: \mathcal{G} \rightarrow T$ ，其中  $\mathcal{G} = \langle x, y, z \mid x^3, y^3, z^2, xyz \rangle$ 。它是满射，因为  $x, y, z$  生成  $T$ 。在例 7. 10. 11 里我们看到  $T$  的元素之间成立的关系  $yxyx$  属于由集合  $\{x^3, y^3, z^2, xyz\}$  生成的正规子群  $\mathcal{R}$ 。  $x, y, z$  之间的每个关系都属于  $\mathcal{R}$  吗？如果不属于，我们想添加一些关系到列表里。也许是令人失望的，因为这个问题还没有答案，但在下节我们将看到  $\psi$  的确是双射。

扼要重述，当说到由生成元集  $S$  与关系集  $R$  定义的群时，我们指的是商群  $\mathcal{G} = \mathcal{F}/\mathcal{R}$ ，其中  $\mathcal{F}$  是  $S$  上的自由群， $\mathcal{R}$  是由  $R$  生成的  $\mathcal{F}$  的正规子群。任意关系集定义一个群。  $R$  越大， $\mathcal{R}$  就变得越大，从而同态  $\pi: \mathcal{F} \rightarrow \mathcal{G}$  发生的坍塌的可能性就越大。极端情形是  $\mathcal{R} = \mathcal{F}$ ，此时， $\mathcal{G}$  是平凡群。在平凡群里所有关系都是成立的。问题出现了，因为在  $\mathcal{F}/\mathcal{R}$  里进行计算可能是困难的。但在许多情形里生成元和关系允许有效的计算，故它们是有用的工具。

215

## 第十一节 托德-考克斯特算法

本节描述的托德-考克斯特算法是一个确定有限群  $G$  在子群  $H$  的陪集上作用的令人惊叹的方法。

要进行计算，需明确给出子集  $G$  与  $H$ 。这样我们考虑一个群

$$\text{【7. 11. 1】} \quad G = \langle x_1, \dots, x_m \mid r_1, \dots, r_k \rangle$$

如上节中那样由生成元和关系表出。

我们还假设  $G$  的子群  $H$  由一个在自由群  $\mathcal{F}$  中的字的集合

$$\text{【7. 11. 2】} \quad \{h_1, \dots, h_t\}$$

具体给出，这个集合在  $G$  中的像生成  $H$ 。



当用右陪集  $Hg$  讨论时, 算法通过构造易读取的表进行. 群  $G$  在右陪集的集合上用右乘作用, 这改变了作用合成的阶. 积  $gh$  通过“先用  $g$  乘, 再用  $h$  乘”的右乘作用. 类似地, 当我们想置换作用右边, 一定以这样的方式读取乘积:

$$\overset{\text{先算}}{(2\ 3\ 4)} \circ \overset{\text{后算}}{(1\ 2\ 3)} = (1\ 2)(3\ 4)$$

下面的规则足以确定  $G$  在右陪集上的作用.

### 【7.11.3】规则

1. 每个生成元的作用是一个置换.
2. 关系平凡地作用: 它们固定每一个陪集不动.
3.  $H$  的生成元固定陪集  $[H]$  不动.
4. 作用是可迁的.

第一个规则由群元素是可逆的事实得到. 第二个规则反映了关系代表  $G$  中的恒等元的事实. 规则 3 和 4 是在陪集上作用的特殊性质.

当应用这些规则时, 通常用指标  $1, 2, 3, \dots$  表示陪集, 用  $1$  代表陪集  $[H]$ . 开始时我们不知道需要多少指标, 故在必要时添加一些新指标.

我们从一个简单例子开始, 在关系 (7.10.9) 里用  $y^2$  替换  $y^3$ .

【7.11.4】例 令  $G$  是群  $\langle x, y, z \mid x^3, y^2, z^2, xyz \rangle$ , 设  $H$  是由  $z$  生成的循环子群  $\langle z \rangle$ .

首先, 由规则 3 知  $z$  映  $1$  到自身,  $1 \xrightarrow{z} 1$ . 这用尽了规则 3 的信息, 于是用规则 1 和 2. 规则 4 没有被直接用到.

我们不知道  $x$  如何在指标  $1$  上作用. 在这种情形里, 程序简单地指定一个新指标,

216  $1 \xrightarrow{x} 2$ . (由于  $1$  代表陪集  $[H]$ , 故指标  $2$  代表  $[Hx]$ , 但最好忽略这一点.) 继续, 我们不知道何处  $x$  映指标  $2$ , 于是指定第三个指标,  $2 \xrightarrow{x} 3$ . 因此,  $1 \xrightarrow{x^2} 3$ .

到现在为止都是部分作用, 这意味着一些生成元在一些指标上的作用被指定了. 沿着这一部分作用轨道一直进行下去是有帮助的. 迄今为止的部分作用是

$$z = (1) \dots \quad \text{与} \quad z = (1\ 2\ 3) \dots$$

$x$  的部分作用的右边括号没有封上, 因为我们没有确定  $x$  映哪个指标到  $3$ .

规则 2 现在开始起作用. 它告诉我们因  $x^3$  是个关系, 所以它固定每个指标不动. 因为  $x^2$  映  $1$  到  $3$ , 故  $x$  一定把  $3$  映回到  $1$ . 习惯上把这些信息综合在一个在指标上展示  $x$  作用的表里:

$$\begin{array}{cccc} x & x & x & \\ \hline 1 & 2 & 3 & 1 \end{array}$$

关系  $xxx$  出现在上面, 规则 2 反映了同一个指标  $1$  出现在其两端这一事实. 我们现在已经确定了部分作用

$$x = (1\ 2\ 3) \dots$$

只是还不知道指标  $1, 2, 3$  是否代表不同的陪集.

其次, 我们要求  $y$  在指标 1 上的作用. 我们也不知道它, 因而指定一个新的指标:  $1 \xrightarrow{y} 4$ . 再次应用规则 2. 因为  $y^2$  是一个关系, 故  $y$  一定把 4 映回到 1. 这展示在表

$$\begin{array}{ccc} y & y & \\ \hline 1 & 4 & 1 \end{array}$$

里, 于是,  $y = (14)\dots$ .

回顾一下, 我们现在已经确定了下面表里的元素. 4 个定义关系出现在上面.

$$\begin{array}{ccc} x & x & x \\ \hline 1 & 2 & 3 \end{array} \quad \begin{array}{ccc} y & y & \\ \hline 1 & 4 & 1 \end{array} \quad \begin{array}{ccc} z & z & \\ \hline 1 & 1 & 1 \end{array} \quad \begin{array}{ccc} x & y & z \\ \hline 1 & 2 & 1 \end{array}$$

表里对于  $xyz$  丢失的元素是 1. 这从  $z$  作为固定指标 1 不动的置换的事实可得. 把 1 加进表里, 我们看到  $2 \xrightarrow{y} 1$ . 但我们还有  $4 \xrightarrow{y} 1$ . 所以,  $4=2$ . 用 2 替换 4, 继续构造表.

现在下面表中的元素都已确定:

$$\begin{array}{ccc} x & x & x \\ \hline 1 & 2 & 3 \end{array} \quad \begin{array}{ccc} y & y & \\ \hline 1 & 2 & 1 \end{array} \quad \begin{array}{ccc} z & z & \\ \hline 1 & 1 & 1 \end{array} \quad \begin{array}{ccc} x & y & z \\ \hline 1 & 2 & 1 \end{array}$$

$xyz$  表的第三行表明  $2 \xrightarrow{z} 3$ , 并且这确定了表的余下部分. 共有 3 个指标, 而完整作用为

$$x = (123), y = (12), z = (23)$$

217

在本节末, 我们将证明这的确是由  $G$  在  $H$  的陪集上的作用定义的置换表示. ■

这样的表所告诉我们的依赖于特殊情形. 它给出陪集的个数, 即指标  $[G:H]$ , 等于不同指标的个数: 在我们的例子里是 3. 它也给出关于生成元的阶的一些信息. 在我们的例子里, 给定关系  $z^2=1$ , 所以  $z$  的阶一定为 1 或 2. 但  $z$  作用在指标上与对换  $(23)$  的作用一样, 这就告诉我们不能有  $z=1$ . 因此,  $z$  的阶为 2, 且  $|H|=2$ . 计数公式  $|G|=|H|[G:H]$  表明  $G$  的阶为  $2 \cdot 3=6$ . 上面列出的三个置换生成对称群  $S_3$ , 所以, 由这个作用定义的置换表示  $G \rightarrow S_3$  是一个同构.

如果取  $H$  为平凡群  $\{1\}$ , 则陪集与群元素一一对应, 且置换表示完全确定  $G$ . 这样做的代价是会有许多指标. 在其他情形里, 置换表示可能不足以确定  $G$  的阶.

我们将计算另外两个例子.

**[7.11.5] 例** 证明关系 (7.10.9) 构成四面体群的完全关系集. 如果用关系  $xyz=1$  消去生成元  $z$ , 则证明会简化一些. 因为  $z^2=1$ , 故这个关系蕴含着  $xy=z^{-1}=z$ . 余下的元素  $x, y$  就足以生成  $T$ . 所以, 用  $z=xy$  代入  $z^2$ , 用  $xyxy$  替换  $z^2$ . 关系变成

$$\mathbf{[7.11.6]} \quad x^3 = 1, y^3 = 1, xyxy = 1$$

$x$  与  $y$  之间的这些关系等价于  $x, y$  与  $z$  之间的关系 (7.10.9), 所以, 它们在  $T$  里成立.

令  $G$  表示群  $\langle x, y | x^3, y^3, xyxy \rangle$ . 推论 (7.10.14) 给出同态  $\psi: G \rightarrow T$ . 要证 (7.11.6) 定义  $T$  的关系, 我们证明  $\psi$  是双射. 因为  $x$  与  $y$  生成  $T$ , 故  $\psi$  是满射. 所以, 只要证明  $G$  的阶等于  $T$  的阶 (为 12) 就够了.

选取子群  $H = \langle x \rangle$ . 这个子群的阶为 1 或 3, 因为  $x^3$  是一个关系. 如果证明  $H$  阶为 3, 则  $H$  在  $G$  里的指标是 4, 则  $G$  的阶为 12, 证明就完成了. 下面是得到的表. 填表的时候从关系的两边开始.

$\begin{array}{cccc} x & x & x & \\ \hline 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 2 \\ 3 & 4 & 2 & 3 \\ 4 & 2 & 3 & 4 \end{array}$	$\begin{array}{cccc} y & y & y & \\ \hline 1 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \\ 3 & 1 & 2 & 3 \\ 4 & 4 & 4 & 4 \end{array}$	$\begin{array}{cccc} x & y & x & y \\ \hline 1 & 1 & 2 & 3 & 1 \\ 2 & 3 & 1 & 1 & 2 \\ 3 & 4 & 4 & 2 & 3 \\ 4 & 2 & 3 & 4 & 4 \end{array}$
--	--	--

置换表示为

$$\text{【7.11.7】} \quad x = (234), y = (123)$$

因为有 4 个指标, 故  $H$  的指标为 4. 还有,  $x$  的阶为 3, 不是 1, 因为伴随于  $x$  的置换的阶为 3.  $G$  的阶是 12, 正是所预计的.

附带地, 我们看到  $T$  同构于交错群  $A_4$ , 因为置换 (7.11.7) 生成这个群. ■

**【7.11.8】例** 我们对关系 (7.10.9) 稍微做些改动, 以说明“坏”关系如何坍塌群. 令  $G$  是群  $\langle x, y \mid x^3, y^3, yxyxy \rangle$ , 并设  $H$  是子群  $\langle y \rangle$ . 这里从做表开始:

$\begin{array}{cccc} x & x & x & \\ \hline 1 & 2 & 3 & 1 \\ 2 & & & 2 \end{array}$	$\begin{array}{cccc} y & y & y & \\ \hline 1 & 1 & 1 & 1 \\ 2 & & & 2 \end{array}$	$\begin{array}{cccc} y & x & y & x & y \\ \hline 1 & 1 & 2 & 3 & 1 & 1 \\ 2 & 3 & 1 & 1 & 2 & 2 \end{array}$
--	--	--

对表里的  $yxyxy$ , 第一行里的三个元素通过从左边计算确定, 后三个通过从右边计算确定.

这个行表明  $2 \xrightarrow{y} 3$ . 第二行通过从左边计算确定, 它表明  $2 \xrightarrow{y} 2$ . 所以,  $2=3$ . 看表中的  $xxx$ , 我们看到  $2=1$ . 仅有一个指标剩下, 从而剩下一个陪集, 所以  $H=G$ . 群  $G$  由  $y$  生成. 它是 3 阶循环群. ■

**注意** 当构造这样的表时必须小心谨慎, 任一失误都将引起作用的坍塌.

在我们的例子里, 取  $H$  为由  $G$  的一个生成元生成的子群. 如果  $H$  由一个字  $h$  生成, 则可引进新的生成元  $u$  与新的关系  $u^{-1}h=1$  (亦即,  $u=h$ ). 这样,  $G(7.11.1)$  同构于群

$$\langle x_1, \dots, x_m, u \mid r_1, \dots, r_k, u^{-1}h \rangle$$

且  $H$  变成由  $u$  生成的子群. 如果  $H$  有若干个生成元, 则我们对每个生成元都这样做.

现在考虑为什么我们所描述的过程的确给出了陪集上的作用这个问题. 在正式定义算法之前, 想要正式地证明这个事实是不可能的, 而我们还没有定义算法, 因而将非正式地讨论这个问题. (更完备的讨论见 [Todd-Coxeter].) 我们这样来描述计算的过程: 在计算的一个特定阶段, 有某个指标集合  $I$ 、在  $I$  上的部分作用以及某些生成元在一些指标上的作用已被确定. 一个部分作用不必符合规则 1, 2 和 3, 但应该是可迁的; 即每个指标都应该属于  $\mathbf{1}$  的“部分轨道”. 这里规则 4 起了作用, 它告诉我们不引入任何我们不需要的指标. 在开始的位置,  $I$  是一个元素的集合  $\{\mathbf{1}\}$ , 且没有指定作用.

在任何一个阶段都有两个可能的步骤:

**【7.11.9】(i)** 可以等同两个指标  $i$  与  $j$ , 如果规则告诉我们它们是相等的, 或者

(ii) 可以选择一个生成元  $x$  和一个指标  $i$  使得  $ix$  还没有被确定, 并且定义  $ix=j$ , 其中  $j$  是个新的指标.

我们从不让两个指标相等, 除非规则蕴含它们相等.

当作用已经确定并且它们符合我们的规则时, 就中止过程. 存在两个问题: 第一, 这个过程会中止吗? 第二, 如果过程中止, 作用是否是正解? 两个问题的答案都是肯定的. 可以证明如果群  $G$  有限且优先进行步骤(i), 则这个过程总是会中止的. 我们不去证明这一点. 对于应用来讲, 更为重要的事实是如果过程中止, 则得到的置换表示是正确的.

219

**【7. 11. 10】定理** 假设经过有限次反复使用步骤(i)和步骤(ii)得到一个与规则(7. 11. 3)相容的表, 则这个表定义一个置换表示, 并且通过适当地标号, 它是  $G$  中  $H$  的右陪集的代表.

**证明** 比如说, 群是  $G = \langle x_1, \dots, x_n | r_1, \dots, r_k \rangle$ , 令  $I^*$  表示最后得到的指标集合. 对每个生成元  $x_i$ , 表确定了指标的置换, 且关系平凡地作用. 推论 7. 10. 14 给出从  $G$  到  $I^*$  的置换群的同态, 从而给出  $G$  在  $I^*$  右边的作用(见命题 6. 11. 2). 倘若遵循规则, 该表表明  $G$  的作用是可迁的, 且子群  $H$  固定指标  $1$  不动.

令  $C$  表示  $H$  的右陪集的集合. 我们将通过定义一个从  $I^*$  到  $C$  的与群在这两个集合上的作用相容的一一映射  $\varphi^*: I^* \rightarrow C$  来证明命题. 我们归纳地定义  $\varphi^*$ : 在每一阶段定义一个由在该阶段所确定的指标集到  $C$  的映射  $\varphi: I \rightarrow C$ , 使得映射与  $I$  上所确定的部分作用相容. 开始时,  $\varphi: \{1\} \rightarrow C$  映  $1 \mapsto [H]$ . 假设  $\varphi: I \rightarrow C$  已经定义, 并设  $I'$  是在  $I$  上应用步骤(7. 11. 9)之一的结果.

在步骤(ii)的情形, 不难将  $\varphi$  拓广为一个映射  $\varphi': I' \rightarrow C$ . 比如说,  $\varphi(i)$  是陪集  $[Hg]$ , 生成元  $x$  在  $i$  上的作用定义为新的指标, 比如说  $ix=j$ . 我们定义  $\varphi'(j) = [Hgx]$ , 且对所有其他指标, 定义  $\varphi'(k) = \varphi(k)$ .

其次, 假设步骤(i)使两个指标  $i$  与  $j$  相同, 使得  $I$  被坍塌而构成一个新的指标集合  $I'$ . 下面的引理使我们能够定义映射  $\varphi': I' \rightarrow C$ .

**【7. 11. 11】引理** 假设给定映射  $\varphi: I \rightarrow C$ , 且与  $I$  上的部分作用相容. 令  $i$  与  $j$  是  $I$  里的指标, 且设规则之一使得  $i=j$ . 则  $\varphi(i) = \varphi(j)$ .

**证明** 这是成立的, 因为前面已经注意到, 陪集上的作用满足规则. ■

映射  $\varphi$  是满射从群在右陪集上作用是可迁的事实得到. 像我们现在所证明的, 单性由陪集  $[H]$  的稳定子是子群  $H$  的事实得到, 且指标  $1$  的稳定子包含  $H$ . 令  $i$  与  $j$  是指标. 因为  $I^*$  上的作用是可迁的, 故对某个群元素  $a$ , 有  $i=1a$ , 从而  $\varphi(i) = \varphi(1)a = [Ha]$ . 类似地, 如果  $j=1b$ , 则  $\varphi(j) = [Hb]$ . 假设  $\varphi(i) = \varphi(j)$ , 亦即,  $Ha = Hb$ , 则  $H = Hba^{-1}$ , 所以  $ba^{-1}$  是  $H$  的元素. 因为  $H$  稳定指标  $1$ , 故  $1 = 1ba^{-1}$  且  $i=1a=1b=j$ . ■

假定我们想要的方法有许多优势;

它们同诚实劳动的优势一样.

## 练 习

### 第一节 凯莱定理

- 1.1 规则  $g * x = xg^{-1}$  定义一个  $G$  在  $G$  上的作用吗?  
 1.2 令  $H$  是群  $G$  的子群. 描述  $H$  在  $G$  上通过左乘的作用的轨道.

### 第二节 类方程

#### 2.1 确定

(a)  $GL_2(\mathbb{F}_3)$  中的矩阵  $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$       (b)  $GL_2(\mathbb{F}_5)$  中的矩阵  $\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$

的中心化子与共轭类的阶.

- 2.2 21 阶群含有 3 阶共轭类  $C(x)$ .  $x$  在群里的阶是什么?  
 2.3 12 阶群  $G$  含有 4 阶共轭类. 证明  $G$  的中心是平凡的.  
 2.4 令  $G$  是群, 且设  $\varphi$  是  $n$  次幂映射:  $\varphi(x) = x^n$ .  $\varphi$  如何作用在共轭类上?  
 2.5 令  $G$  是形如  $\begin{bmatrix} x & y \\ & 1 \end{bmatrix}$  的矩阵群, 其中  $x, y \in \mathbb{R}$  且  $x > 0$ . 确定  $G$  中的共轭类, 在  $(x, y)$  平面概略地描述它们.

2.6 确定平面的等距群  $M$  里的共轭类.

2.7 在下列各式中尽可能多地划去那些 10 阶群的类方程.

$$1+1+1+2+5, \quad 1+2+2+5, \quad 1+2+3+4, \quad 1+1+2+2+2+2.$$

2.8 确定具有下面阶数的非阿贝尔群的可能的类方程: (a)8, (b)21.

2.9 确定下面每个群的类方程:

(a) 四元数群 (b)  $D_4$  (c)  $D_8$  (d)  $GL_2(\mathbb{F}_3)$  中的可逆上三角矩阵子群

2.10 (a) 令  $A$  是  $SO_3$  的表示转过角度  $\pi$  的旋转的元素. 用几何方法刻画  $A$  的中心化子.

(b) 确定平面等距群  $M$  里的关于  $e_1$  轴的反射  $r$  的中心化子.

2.11 确定下列每个矩阵在  $GL_3(\mathbb{R})$  中的中心化子.

$$\begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}, \begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 & \\ & 1 & \\ & & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}, \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}.$$

2.12 确定至多含有 3 个共轭类的所有有限群.

2.13 令  $N$  是群  $G$  的正规子群. 假设  $|N| = 5$  且  $|G|$  是奇数. 证明  $N$  包含在  $G$  的中心里.

2.14 群  $G$  的类方程是  $1+4+5+5+5$ .

(a)  $G$  有 5 阶子群吗? 如果有, 它是正规子群吗?

(b)  $G$  有 4 阶子群吗? 如果有, 它是正规子群吗?

2.15 证明  $SL_2(\mathbb{F}_3)$  的类方程 (7.2.10).

2.16 令  $\varphi: G \rightarrow G'$  一个满的群同态, 设  $C$  是  $G$  的元素  $x$  的共轭类, 设  $C'$  表示它的像  $\varphi(x)$  在  $G'$  中的共轭类. 证明  $\varphi$  将  $C$  满射地映射到  $C'$  上, 且  $|C'|$  整除  $|C|$ .

2.17 用类方程证明  $pq$  阶群含有  $p$  阶元素, 其中  $p$  与  $q$  是素数.

2.18 哪个矩阵对  $\begin{bmatrix} 0 & -1 \\ 1 & d \end{bmatrix}$ ,  $\begin{bmatrix} 0 & 1 \\ -1 & d \end{bmatrix}$  是下列群的共轭元?

(a)  $GL_n(\mathbf{R})$  (b)  $SL_n(\mathbf{R})$

### 第三节 $p$ -群

3.1 证明不动点定理 7.3.2.

3.2 设  $Z$  是群  $G$  的中心. 证明: 如果  $G/Z$  是循环群, 则  $G$  是阿贝尔群, 因而  $G=Z$ .

3.3 一个非阿贝尔群的阶为  $p^3$ , 其中  $p$  是素数.

(a) 中心  $Z$  的可能阶数是什么?

(b) 令  $x$  是  $G$  的不属于  $Z$  的元素. 它的中心化子  $Z(x)$  的阶是什么?

(c)  $G$  的可能的类方程是什么?

3.4 将 8 阶群分类.

### 第四节 二十面体群的类方程

4.1 二十面体群作用在十二面体的 5 个内接立方体的集合上. 确定一个立方体的稳定子.

4.2  $A_5$  是  $S_5$  仅有的真正规子群吗?

4.3 二十面体群  $I$  的 2 阶元素的稳定子是什么?

4.4 (a) 确定正四面体群  $T$  的类方程.

(b) 证明  $T$  有一个 4 阶正规子群, 没有 6 阶子群.

4.5 (a) 确定八面体群  $O$  的类方程.

(b) 这个群含有两个真正规子群. 求出它们, 证明它们是正规的, 并证明没有其他正规子群.

4.6 (a) 证明四面体群  $T$  同构于交错群  $A_4$ , 而八面体群  $O$  同构于对称群  $S_4$ .

提示: 从求这些群作用的一个四元素集合开始.

(b) 两个四面体可内接到一个正方体  $C$  中, 每个使用其一半顶点. 将这一事实与包含关系  $A_4 \subset S_4$  联系起来.

4.7 令  $G$  是非平凡地作用在  $r$  阶集合上的  $n$  阶群. 证明: 如果  $n > r!$ , 则  $G$  有一个真正规子群.

4.8 (a) 设群元素  $x$  的中心化子  $Z(x)$  的阶为 4. 关于群的中心有何结论?

(b) 设元素  $y$  的共轭类  $C(y)$  的阶为 4. 关于群的中心有何结论?

222

4.9 令  $x$  是群  $G$  的不为恒等元的元素, 其中心化子  $Z(x)$  的阶为  $pq$ , 其中  $p$  与  $q$  是素数. 证明  $Z(x)$  是阿贝尔的.

### 第五节 对称群里的共轭

5.1 (a) 证明置换  $(1\ 2), (2\ 3), \dots, (n-1\ n)$  生成对称群  $S_n$ .

(b) 需要多少置换来写出循环  $(1\ 2\ 3\ \dots\ n)$ ?

(c) 证明循环  $(1\ 2\ 3\ \dots\ n)$  与  $(1\ 2)$  生成对称群  $S_n$ .

5.2 在  $S_5$  里元素  $(1\ 2)$  的中心化子是什么?

5.3 确定对称群  $S_5$  中元素的阶.

5.4 在对称群  $S_6$  中描述置换  $\sigma = (1\ 5\ 3)(2\ 4\ 6)$  的中心化子  $Z(\sigma)$ , 并计算  $Z(\sigma)$  与  $C(\sigma)$  的阶.

- 5.5 设  $p$  与  $q$  是置换. 证明乘积  $pq$  与  $qp$  有同样大小的循环.
- 5.6 求对称群  $S_4$  的所有 4 阶子群, 并确定哪些是正规的.
- 5.7 证明  $A_n$  是  $S_n$  的唯一的指标为 2 的子群.
- ⊖ 5.8 确定整数  $n$  使得从对称群  $S_n$  到  $S_{n-1}$  存在满同态.
- 5.9 令  $q$  是  $S_n$  里的 3-循环. 存在多少偶置换  $p$  使得  $pqp^{-1}=q$ ?
- 5.10 对  $S_4$  与  $S_5$  的类方程证明公式(7.5.2)与(7.5.3), 并确定每个共轭类里代表元的中心化子.
- 5.11 (a) 令  $C$  是  $S_n$  里偶置换  $p$  的共轭类, 证明  $C$  或者是  $A_n$  中的共轭类, 或者是  $A_n$  中的两个等阶共轭类的并. 解释如何用  $p$  的中心化子确定出现哪种情况.  
 (b) 确定  $A_4$  与  $A_5$  的类方程.  
 (c) 也可把奇数阶置换的共轭类分解成  $A_n$ -轨道. 描述这个分解.
- 5.12 确定  $S_6$  与  $A_6$  的类方程.

## 第六节 正规化子

- 6.1 证明在  $GL_n(\mathbf{R})$  里可逆上三角矩阵子群  $B$  与可逆下三角矩阵子群  $L$  是共轭的.
- 6.2 令  $B$  是  $G=GL_n(\mathbf{R})$  的可逆上三角矩阵子群, 且设  $U \subset B$  是对角元为 1 的上三角矩阵的集合. 证明  $B=N(U)$  与  $B=N(B)$ .
- 6.3 令  $P$  表示  $GL_n(\mathbf{R})$  的由置换矩阵组成的子群. 确定正规化子  $N(P)$ .
- 6.4 令  $H$  是有限群  $G$  里的素数阶  $p$  的正规子群. 设  $p$  是整除  $G$  的阶的最小素数. 证明  $H$  包含在中心  $Z(G)$  里.
- 223 6.5 令  $p$  是素整数, 且设  $G$  是  $p$ -群. 令  $H$  是  $G$  的真子群. 证明  $H$  的正规化子  $N(H)$  严格大于  $H$ , 且  $H$  包含在指标  $p$  的一个正规子群里.
- 6.6 令  $H$  是有限群  $G$  的真子群. 证明:  
 (a) 群  $G$  不是  $H$  的共轭子群的并.  
 (b) 存在共轭类  $C$  与  $H$  不相交.

## 第七节 西罗定理

- 7.1 令  $n=p^a m$ , 如同(4.5.1)中那样, 且令  $N$  是  $n$  阶集合里阶  $p^a$  的子集的个数. 确定  $N$  模  $p$  的同余类.
- 7.2 令  $G_1 \subset G_2$  是其阶为  $p$  所整除的群, 且设  $H_1$  是  $G_1$  的西罗  $p$ -子群. 证明存在  $G_2$  的西罗  $p$ -子群  $H_2$  使得  $H_1=H_2 \cap G_1$ .
- 7.3 有多少个 5 阶元素包含在 20 阶群里?
- 7.4 (a) 证明不存在阶为  $pq$  的单群, 其中  $p$  与  $q$  是素数.  
 (b) 证明不存在阶为  $p^2 q$  的单群, 其中  $p$  与  $q$  是素数.
- 7.5 求下列群的西罗 2-子群: (a)  $D_{10}$  (b)  $T$  (c)  $O$  (d)  $I$
- 7.6 求出对称群  $S_7$  的 21 阶的非阿贝尔子群.
- 7.7 令  $n=pm$  是恰为  $p$  整除一次的整数, 且设  $G$  是  $n$  阶群. 设  $H$  是  $G$  的西罗  $p$ -子群, 且设  $S$  是所有西罗  $p$ -子群的集合. 解释  $S$  如何分解为  $H$ -轨道.

- 7.8 计算  $GL_n(\mathbb{F}_p)$  的阶. 求  $GL_n(\mathbb{F}_p)$  的西罗  $p$ -子群, 并确定西罗  $p$ -子群的个数.
- 7.9 将下列阶数的群分类: (a)33 (b)18 (c)20 (d)30
- 7.10 证明阶  $< 60$  的唯一单群是素数阶群.

### 第八节 12 阶群

- 8.1 定理 7.8.1 里所描述的哪个 12 阶群同构于  $S_3 \times C_2$ ?
- 8.2 (a) 确定最小整数  $n$  使得对称群  $S_n$  含有同构于群 (7.8.2) 的子群.  
(b) 求  $SL_2(\mathbb{F}_5)$  的同构于那个群的子群.
- 8.3 确定 12 阶群的类方程.
- 8.4 证明阶  $n=2p$  的群或是循环的, 或是二面体的, 其中  $p$  是素数.
- 8.5 令  $G$  是 28 阶的非阿贝尔群, 其西罗 2-子群是循环群.  
(a) 确定西罗 2-子群与西罗 7-子群的个数.  
(b) 证明至多存在一个这样群的同构类.  
(c) 确定每个阶的元素的个数与  $G$  的类方程.
- 8.6 令  $G$  是 55 阶群.  
(a) 证明  $G$  是由元素  $x$  与  $y$  生成的, 带有关系  $x^{11}=1, y^5=1, yxy^{-1}=x^r$ , 其中  $1 \leq r < 11$ .  
(b) 确定  $r$  的哪个值是可能的.  
(c) 证明存在两个 55 阶群的同构类.

224

### 第九节 自由群

- 9.1 令  $F$  是  $\langle x, y \rangle$  上的自由群. 证明三个元素  $u=x^2, v=y^2$  与  $z=xy$  生成同构于  $u, v$  与  $z$  上的自由群的子群.
- 9.2 可在  $S'$  里定义封闭字为连接字的首尾端得到的方向环. 反时针方向读,

$$\begin{array}{ccccc} & & c & a^{-1} & \\ & b & & & b \\ a & & & & b \\ & a & & & c \\ & & b & b & d \end{array}$$

是个封闭字. 在简约封闭字与自由群的共轭类之间建立一一对应.

### 第十节 生成元与关系

- 10.1 证明自由群和商群的映射性质.
- 10.2 令  $\varphi: G \rightarrow G'$  是群的满同态. 设  $S$  是  $G$  的子集, 其像  $\varphi(S)$  生成  $G'$ , 并设  $T$  是  $\ker \varphi$  的生成元集. 证明  $S \cup T$  生成  $G$ .
- 10.3 每个有限群  $G$  都能由有限生成元集与有限关系集表示出来吗?
- 10.4 群  $G = \langle x, y; xyx^{-1}y^{-1} \rangle$  叫做自由阿贝尔群. 证明这个群的映射性质: 如果  $u$  与  $v$  是阿贝尔群  $A$  的元素, 则存在唯一同态  $\varphi: G \rightarrow A$  使得  $\varphi(x) = u, \varphi(y) = v$ .
- 10.5 证明由  $x, y, z$  与单个关系  $yxyz^{-2} = 1$  生成的群实际上是自由群.
- 10.6 群  $G$  的子群  $H$  是特征的, 如果  $G$  的所有自同构把它映到它自身.  
(a) 证明每个特征子群是正规的, 中心  $Z$  是特征子群.  
(b) 确定四元数群的正规子群与特征子群.