矩 阵 光 学

阎吉祥　魏光辉　哈流柱　林永昌　江先进　编著

兵器工业出版社

(京)新登字 049 号

内容简介

本书以线性矩阵变换为数学工具，系统介绍工程光学和光电子学中的主要光学原理与应用。全书共 6 章，内容包括矩阵与张量运算，几何光学基础，激光束的光学变换与光学谐振腔，偏振光学，反射棱镜调整理论以及薄膜光学原理。采用矩阵方法进行描述，便于揭示和理解上述内容的物理实质和规律，对初学的光学工作者和希望涉及基础光学的电子学工作者来说，本书是一本很好的入门书。

本书内容丰富，叙述简明，反映了各位作者从事专题的科学研究成果，可参考性强，可作为理科光学专业、工科光学仪器、光电子学及电子学、光通信等有关专业研究生和高年级大学生的教学参考书与相关专业科技人员参考。

图书在版编目(CIP)数据

矩阵光学/魏光辉等编著. —北京：兵器工业出版社，1995.8

ISBN 7-80038-850-6

Ⅰ.矩…　Ⅱ.魏…　Ⅲ.矩阵-光学　Ⅳ.O439

中国版本图书馆 CIP 数据核字(94)第 15224 号

兵器工业出版社出版发行

(北京市海淀区车道沟 10 号)

各地新华书店经销

北京育才印刷厂印装

*

开本：787×1092　1/16　印张：18.875　字数：463.32 千字

1995 年 8 月第 1 版　1995 年 8 月第 1 次印刷

印数：1—1000　定价：22.00 元

前　言

矩阵光学，确切地说，是指光学中的矩阵方法，也就是用矩阵这一数学工具把光学中的典型线性变换问题统一地联系起来。光学中能够用矩阵描述的内容很多，本书选择了工程光学中最基本的、应用较多的几何光学、激光束光学、偏振光学、薄膜光学、反射棱镜调整理论等作为主要内容。矩阵方法不但能够十分简洁地对这些内容进行描述，而且便于揭示和理解其基本规律。当然，矩阵方法并不是处理这些光学问题的唯一方法。

本书涉及的内容大多是光学中的重要学科分支，内涵浩繁，不乏专著。我们的宗旨是力图给予读者关于矩阵光学的系统基本知识，在叙述方法上着重用矩阵方法描述物理概念，然后列举典型实例说明基本规律，不过多涉及专深内容。这样做也许更有利于读者阅读和理解。

本书各章内容相对独立，因此，读者可根据需要选读其中一部分内容，这样并不影响连贯性。

编写工作由北京理工大学工程光学系魏光辉等五位同志集体完成。第 1 章由阎吉祥编写，第 2、3 章由魏光辉编写，第 4 章由哈流柱编写，第 5 章由林永昌编写，第 6 章由江先进编写。全书由魏光辉统稿。第 3 章至第 6 章所编入的内容中相当一部分是作者多年从事该学科教学和科研成果的总结，具有一定的实用参考价值。

于美文先生倡议集体编写本书。她要求作者用自己的研究心得和成果向读者介绍矩阵光学原理。作者谨向于先生致以学生的敬意，感谢她对编写本书的指导和长期的教诲。

限于作者水平，书中不当之处，请读者指正。

编著者

1993 年 7 月

目　录

1 矩阵及其运算

本章是以后各章的基础。鉴于读者已经具备线性代数的基本知识，本章对读者比较熟悉的内容只做简单介绍，而对读者比较生疏的问题则给以较详细的讨论。

在 1.1 中首先回顾矩阵概念，并介绍矢量和张量的矩阵表示。紧接着，1.2 讨论矩阵的两种最基本的运算，即加法运算和乘法运算。

在数学物理中经常要用到变换，用矩阵表示变换，将使问题的处理大为简化，本章 1.3 用较多的篇幅描述变换的矩阵表示。接下来的两节介绍矩阵的另外两种运算，即转置和求逆。而 1.6 给出两类特殊矩阵：厄米特矩阵和酉矩阵，它们在本章最后一节求矩阵的特征值(亦称本征值)和特征向量的过程中起重要作用。

1.1 矩阵、矢量和张量

1.1.1 矩阵的概念

由 $m\times n$ 个元素按照一定的规则排列成的矩形表

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{pmatrix} \tag{1.1-1}$$

称为 m 行 n 列矩阵，或 $m\times n$ 矩阵；$a_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 称为第 i 行，第 j 列的矩阵元，它可以是实数，也可以是复数，还可以是多项式。当 $m=1$ 时得到单行矩阵；$n=1$ 时给出单列矩阵。如果矩阵的行数和列数相等，即 $m=n$，则称该矩阵为 n 阶方阵。在 n 阶方阵中，$a_{ii}(i=1,2,\cdots,n)$ 称为方阵的对角元素。除对角元素外，其余元素都为零的方阵称为对角方阵，而对角元素均为 1 的对角方阵称为单位方阵。也就是说，如果方阵的矩阵元 a_{ij} 可写为

$$a_{ij}=a_{ii}\delta_{ij}$$

那么我们得到对角方阵，其中

$$\delta_{ij}=\begin{cases}1, i=j\\ 0, i\neq j\end{cases}$$

是熟知的克罗内克尔符号，而当

$$a_{ij}=\delta_{ij}$$

时，便得到单位方阵。

所有矩阵元均为零的矩阵称为零矩阵。

1.1.2 矢量和张量的矩阵描述

我们知道，在三维空间中，只需一个量就可以描述某一点的光强，一般来说，用三个量才能完全描述某一点的电场强度，而为了描述一点的应力，则需要 9 个分量。于是，我们分别称光

强、电场强度和应力为标量、矢量和张量。此处所谓张量确切地说应称为二阶张量。

更一般地，三维空间的 m 阶张量可以有 3^m 个独立分量，而 n 维空间的 m 阶张量则可以有 n^m 个独立分量。因此，矢量可视为一阶张量，而标量可以认为是零阶张量。

这里有两点值得读者注意：

i)在数学和理论物理学的严格表述中，标量、矢量和张量主要不是根据其分量的多少，而是根据这些分量在一定的坐标变换中所遵从的变换关系来定义的。这将在本章"变换的矩阵表示"以后加以介绍。

ii)有些量并非是某一真实矢量的分量，但为了处理方便，我们可以认为它们构成一个向量。例如，在线性方程组中，将诸变元视为一个抽象向量的分量，而由这些向量构成的空间就是抽象的向量空间。又譬如在研究光线的传播时，设光线与某一选定轴的夹角为 θ，光线与其路径上某一横断面的交点到该轴的距离为 r，显然，这里的 r 与 θ 并非一个真正矢量的两个分量，但如果我们设想它们组成一个列向量，处理起来就很方便，这一点读者在以后的章节中将会看到。

前面讲到，矩阵是若干元素按一定规则排列成的矩形表，特殊情况下它可以是单行表、单列表或方形数表。我们又知道，n 维空间的矢量可以用 n 个分量表示，二阶张量可以用 n^2 个分量表示，……，m 阶张量可以用 $n^m=n\times n^{m-1}$ 个分量表示，因此，我们可以用单行矩阵或单列矩阵表示一个矢量，而用方阵表示一个二阶张量。至于 m 阶张量，它的分量似乎可有多种排列方式，例如 n 行 n^{m-1} 列，n^2 行 n^{m-2} 列，……，n^{m-1} 行 n 列。然而，由于空间的维数为 n，故 m 阶张量当用 n 行 n^{m-1} 列矩阵表示，这一点在讲完"变换的矩阵表示"后会更加明显。

1.2 矩阵的加法和乘法

1.2.1 矩阵的加法

两个行数与列数都相同的矩阵

$$A=(a_{ij})\quad i=1,2,\cdots,m;j=1,2,\cdots,n;$$

$$B=(b_{ij})\quad i=1,2,\cdots,m;j=1,2,\cdots,n,$$

如果它们的对应元素都相等，即

$$a_{ij}=b_{ij}$$

则称这两个矩阵相等，并记为

$$A=B$$

两个行数与列数分别相等的矩阵 A 和 B 可以相加，并得到一个具有相同行、列数的矩阵，如果将和矩阵记为 C，则

$$C=A+B$$

而 C 的矩阵元为

$$c_{ij}=a_{ij}+b_{ij}$$

由定义可知，矩阵加法适合交换律和结合律，即

$$A+B=B+A$$

$$A+B+C=A+(B+C)$$

1.2.2 矩阵的乘法

一个 $m\times p$ 的矩阵

$$A=(a_{ij}),i=1,2,\cdots,m;j=1,2,\cdots,p;$$

可以左乘一个 $p\times n$ 的矩阵

$$B=(b_{qr}),q=1,2,\cdots,p;r=1,2,\cdots,n,$$

并得到一个 $m\times n$ 矩阵

$$C=(c_{ir}),i=1,2,\cdots,m;r=1,2,\cdots,n$$

这一运算可以用

$$C=AB$$

来表示，而矩阵 C 的元素 c_{ir} 等于矩阵 A 的第 i 行元素与矩阵 B 的第 r 列对应元素之积之和，即

$$c_{ir}=\sum_{k=1}^{p}a_{ik}b_{kr},i=1,2,\cdots,m;r=1,2,\cdots,n \tag{1.2-1}$$

当 $n=p$，或 $n=p^{\alpha-1}$，α 为正整数时，如前所述，矩阵 B 可以描述 p 维空间的二阶或 α 阶张量，如果又有 $m=p$，则 A 为 p 阶方阵，所得的 C 亦为 $p\times n$ 矩阵，它可以描述 p 维空间的一个二阶或 α 阶张量。这就是说，一个张量矩阵可以被列数与其维数相同的方阵左乘，并得到另一个具有相同维数的张量矩阵。

特别是，当 $\alpha=1$ 时，有 $n=1$，即 B 为一个单列矩阵，于是，上述结果表明，一个用单列矩阵表示的矢量，被列数与其维数相同的方阵左乘，仍得一个具有相同维数的矢量。

矩阵相乘的一个特例是数乘矩阵。数 K 与 $m\times n$ 矩阵 $A=(a_{ij})$ 相乘，得到另一个 $m\times n$ 矩阵 $C=(c_{ij})$，其元素 $c_{ij}=Ka_{ij}$。

至此，我们可以把矩阵加法的概念推广到矩阵减法上；矩阵 A 减矩阵 B 可用下法得到，即首先用 -1 乘矩阵 B，然后将结果与矩阵 A 相加。

矩阵的元素可以是实数，也可以是复数，还可以是某一未知数的多项式。用矩阵相加和数乘矩阵的知识，可以将以某一未知数的多项式为元素的矩阵表示为同一未知数的一个多项式，其最高幂等于原来矩阵元中的最高次幂，而每一项的系数都是一个与原矩阵同行同列的矩阵，其元素为原矩阵中相应的未知数幂的系数。例如矩阵

$$A=\begin{bmatrix}x^4-x^2+1 & x^2-3\\ x^3+2x^2-4 & x\end{bmatrix}$$

按上述方法可表示为

$$A'=\begin{bmatrix}1&0\\0&0\end{bmatrix}x^4+\begin{bmatrix}0&0\\1&0\end{bmatrix}x^3+\begin{bmatrix}-1&1\\2&0\end{bmatrix}x^2+\begin{bmatrix}0&0\\0&1\end{bmatrix}x+\begin{bmatrix}1&-3\\-4&0\end{bmatrix}$$

行数和列数适当的矩阵可以连乘，例如，设 A 为 $m\times p$ 矩阵，B 为 $p\times q$ 矩阵，C 为 $q\times n$ 矩阵，则 $R=AB$ 为 $m\times q$ 矩阵，第 i 行第 j 列的矩阵元为

$$r_{ij}=\sum_{k=1}^{p}a_{ik}b_{kj}$$

而 $\Pi=RC=ABC$ 为 $m\times n$ 矩阵，其矩阵元为

$$\pi_{ij}=\sum_{h=1}^{q}r_{ih}c_{hj}$$

$$= \sum_{h=1}^{q}\Big(\sum_{k=1}^{p} a_{ik}b_{kh}\Big)c_{hj}$$

一个方阵自身可以连乘，其结果为阶数不变的新方阵，并可写成幂的形式。例如，设 A 为方阵，则

$$\underbrace{AA\cdots A}_{m个} = A^m$$

矩阵乘法法则导致一个有趣的结果，即如果矩阵 A 和 B 中至少有一个是零矩阵，则它们的乘积 $C=AB$ 必为零矩阵；但是反过来，如果已知 $C=AB$ 是零矩阵，则不能断言 A、B 中至少有一个是零矩阵。这一结论可以由式(1.2-1)清楚地看出。如果 A、B 中有一个为零矩阵，为确定起见，且不失一般性，不妨设 A 是零矩阵，则 A 的所有元素均为零，即

$$a_{ik} = 0, \quad i = 1,2,\cdots,m; k = 1,2,\cdots,p$$

于是，由式(1.2-1)可知，C 的所有元素为零，即

$$c_{ij} = 0, \quad i = 1,2,\cdots,m; j = 1,2,\cdots,n$$

亦即 C 是零矩阵。反过来，设 C 为零矩阵，则 $c_{ij}=0$，但这只能说明所有求和

$$\sum_{k=1}^{p} a_{ik}b_{kj} = 0, \quad i = 1,2,\cdots,m; j = 1,2,\cdots,n$$

并不能说明每一项 $a_{ik}b_{kj}=0$；退而言之，纵然每一项 $a_{ik}b_{kj}$都为零，也可能有些是 $a_{ik}=0$ 的结果，有些则是 b_{kj}所致，而不一定是所有的 a_{ik}或全部 b_{kj}为零。以下两个简单的例子可作为这两种情况的佐证。

例 1. 设

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

则

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1-1 & 1-1 \\ 1-1 & 1-1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

例 2. 令

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

则

$$C = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1\times 0+0\times 1 & 1\times 0+0\times 1 \\ 1\times 0+0\times 1 & 1\times 0+0\times 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

在第一个例子中，乘积矩阵的矩阵元由两项组成，这两项的代数和为零，但每一项并不为零。在第二个例子中，乘积矩阵的每一个矩阵元亦由两项组成，且这两项都为零，但一项是因 B 的矩阵元为零引起，另一项则由 A 的矩阵元为零所决定。

下面讨论矩阵乘法的性质

(1)适合结合律　如果 ABC 存在，则 $A(BC)$也存在，且

$$ABC = A(BC)$$

证：设 A 为 $m\times p$ 矩阵，B 为 $p\times q$ 矩阵，C 为 $q\times n$ 矩阵，则由前面的讨论可知，

$$\Pi = ABC$$

其矩阵元为

$$\pi_{ij} = \sum_{k=1}^{p}\sum_{h=1}^{q} a_{ik}b_{kh}c_{hj}$$

另一方面，由矩阵乘法法则可知，BC 是 $p\times n$ 矩阵，且位于第 k 行第 j 列的矩阵元为 $\sum_{h=1}^{q} b_{kh}c_{hj}$。再次应用矩阵乘法法则，得

$$\Pi' = A(BC)$$

为 $m\times n$ 矩阵，其一般元素为

$$\pi_{ij}' = \sum_{k=1}^{p} a_{ik}\Big(\sum_{h=1}^{q} b_{kh}c_{hj}\Big)$$

$$= \sum_{k=1}^{p}\sum_{h=1}^{q} a_{ik}b_{kh}c_{hj}$$

因而

$$\Pi' = \Pi,$$

亦即

$$ABC = A(BC)$$ 证毕。

(2)不适合交换律　如果 AB 和 BA 都存在的话，一般来说它们是不相等的。例如，设

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

则 AB 第一行第一列的元素为

$$(AB)_{11} = a_{11}b_{11} + a_{12}b_{21};$$

而矩阵 BA 的对应元素为

$$(BA)_{11} = a_{11}b_{11} + a_{21}b_{12}$$

二者显然是不恒等的。

更有甚者，经常会遇到这样的情况，即 AB 是一个好端端的矩阵，而 BA 却根本不存在，更谈不上与 AB 相等了。例如，设 A 为 $m\times p$ 矩阵，B 为 $p\times n$ 矩阵，则矩阵 AB 总是存在的，而矩阵 B 左乘矩阵 A 的运算则是无意义的，除非 $m=n$。

矩阵乘法中几个常见的可对易特例是

i)数乘矩阵时位置可易

ii)单位方阵

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

与任一同阶方阵对易。

以上两个结论的正确性都是不言而喻的。

iii)任一方阵与其本身对易，且与它自己的任意次幂对易。前一结论是显然的，下面证明后一结论。

设 A 为 n 阶方阵，由结合律可得

$$AA^2 = A(AA) = (AA)A = A^2A$$

即 A 与 A^2 可对易。

设方阵 A 与其 k 次幂 A^k(前面已指出,它也是 n 阶方阵)对易,即

$$AA^k = A^kA,$$

则由结合律,

$$AA^{k+1} = A(A^kA) = (AA^k)A = A^{k+1}A$$

即 A 与其 $k+1$ 次幂也对易。根据数学归纳法,原结论是正确的。

iv)阶数相同的对角方阵彼此可对易。

即如果 A、B 是两个 n 阶对角方阵,则

$$AB = BA$$

证:设 A 与 B 的矩阵元分别为 $a_{ij}\delta_{ij}$和 $b_{ij}\delta_{ij}$,则

$C=AB$ 的矩阵元为

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\delta_{ik}b_{kj}\delta_{kj}$$

$$= a_{ii}b_{ij}\delta_{ij} = \begin{cases} a_{ii}b_{ii}, i = j; \\ 0, \quad i \neq j. \end{cases}$$

而

$C' = BA$ 的矩阵元为

$$c_{ij}' = \sum_{k=1}^{n} b_{ik}\delta_{ik}a_{kj}\delta_{kj}$$

$$= a_{jj}b_{ij}\delta_{ij} = \begin{cases} a_{ii}b_{ii}, i = j; \\ 0, \quad i \neq j. \end{cases}$$

所以

$$AB = BA$$

(3)适合乘法对加法的分配律　即如果矩阵 A 可左乘矩阵 B 及 C,则

$$A(B + C) = AB + AC$$

证:设 A 为 $m\times p$ 矩阵,B、C 均为 $p\times n$ 矩阵,则$(B+C)$为 $p\times n$ 矩阵,矩阵元为$(b_{kj}+c_{kj})$,

而

$$V = A(B + C)$$

为 $m\times n$ 矩阵,元素为

$$v_{ij} = \sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj})$$

另一方面,AB 与 AC 都是 $m\times n$ 矩阵,且矩阵元分别为

$$\sum_{k=1}^{p} a_{ik}b_{kj} \text{ 和 } \sum_{k=1}^{p} a_{ik}c_{kj},$$

于是,由矩阵的加法可知,

$$U = AB + AC$$

也是一个 $m\times n$ 矩阵,且矩阵元为

$$u_{ij} = \sum_{k=1}^{p} (a_{ik}b_{kj} + a_{ik}c_{kj})$$

$$= \sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj}) = v_{ij}$$

因而

$$A(B+C)=AB+AC$$

(4)设 A、B 是行数与列数相同的两个矩阵,K 和 L 是两个数,则有

① $K(A+B)=KA+KB$

② $(K+L)A=KA+LA$

③ $K(LA)=(KL)A$

④ $K(AB)=(KA)B=A(KB)$

1.3 变换的矩阵表示

我们知道,正交变换的数学定义是:在欧几里得空间中,将一个标准正交基底($a_1,a_2,\cdots,a_n$)变成另一个标准正交基底($a_1',a_2',\cdots,a_n'$)的变换。而在物理学中,通常将长度为不变量的变换称为正交变换。例如,狭义相对论中的洛伦兹变换就是复四维空间中的正交变换。

在三维空间中,当直角坐标系有一转动时,坐标变换关系为

$$x_i'=\sum_{j=1}^{3}a_{ij}x_j,\qquad i=1,2,3 \tag{1.3-1}$$

由于矢量(x_1,x_2,x_3)的长度 $\sum_{i=1}^{3}x_ix_i$ 在转动中保持为不变量,即

$$\sum_{i=1}^{3}x_i'x_i'=\sum_{j=1}^{3}x_jx_j, \tag{1.3-2}$$

所以转动变换是三维空间的正交变换。

将式(1.3-1)展开,则有

$$\begin{cases} x_1'=a_{11}x_1+a_{12}x_2+a_{13}x_3 \\ x_2'=a_{21}x_1+a_{22}x_2+a_{23}x_3 \\ x_3'=a_{31}x_1+a_{32}x_2+a_{33}x_3 \end{cases}$$

或写成

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

这就是三维空间中的正交变换的矩阵表示。而

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

则称为正交矩阵。

将式(1.3-1)代入式(1.3-2)中,得到

$$\sum_{i=1}^{3}\Big(\sum_{j=1}^{3}a_{ij}x_j\Big)\Big(\sum_{k=1}^{3}a_{ik}x_k\Big)=\sum_{j=1}^{3}x_jx_j$$

即

$$\sum_{j=1}^{3}\Big[\sum_{k=1}^{3}\Big(\sum_{i=1}^{3}a_{ij}a_{ik}\Big)x_k\Big]x_j=\sum_{j=1}^{3}x_jx_j$$

比较两边系数,给出

$$\sum_{k=1}^{3}\left(\sum_{i=1}^{3} a_{ij}a_{ik}\right)x_k = x_j$$

由此可得 A 为正交矩阵的充要条件为

$$\sum_{i=1}^{3} a_{ij}a_{ik} = \delta_{kj} \tag{1.3-3}$$

假定相继施行两次转动，第一次转动后得到

$$\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{pmatrix} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

第二次转动后即得

$$\begin{aligned} \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(3)} \end{pmatrix} &= \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} \\ a_{31}^{(2)} & a_{32}^{(2)} & a_{33}^{(2)} \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} \\ a_{31}^{(2)} & a_{32}^{(2)} & a_{33}^{(2)} \end{pmatrix} \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

这就是说，表示总变换的矩阵，等于表示第二次变换的矩阵左乘表示第一次变换的矩阵。推而广之，如果相继施行 n 次转动，则表示总变换的矩阵等于表示各次变换的矩阵反转次序后相乘。换言之，设 $A^{(1)}$、$A^{(2)}$、…，$A^{(n)}$分别为第一次、第二次、…，第 n 次变换的矩阵，则有

$$\begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ x_3^{(n)} \end{pmatrix} = A^{(n)}A^{(n-1)}\cdots A^{(2)}A^{(1)} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{1.3-4}$$

这是容易证明的，事实上，设 $n=k$ 时上式成立，即设

$$\begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} = A^{(k)}A^{(k-1)}\cdots A^{(2)}A^{(1)} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

则

$$\begin{aligned} \begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} &= A^{(k+1)} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} \\ &= A^{(k+)}A^{(k)}\cdots A^{(2)}A^{(1)} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \end{aligned}$$

即 $n=k+1$ 时结论成立，而我们已经知道 $n=2$ 时结论是成立的。于是，根据数学归纳法，式(1.3-4)对任意自然数 n 都成立。

作为一系列变换相继作用的例子，我们研究矩阵

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

的 n 次方

$$\begin{bmatrix} A_{T_n} & B_{T_n} \\ C_{T_n} & D_{T_n} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^n$$

其中原矩阵的元素为

$$A = 1 - \frac{d}{f_2}$$

$$B = d\left(2 - \frac{d}{f_2}\right)$$

$$C = -\left[\frac{1}{f_1} + \frac{1}{f_2}\left(1 - \frac{d}{f_2}\right)\right]$$

$$D = -\left[\frac{d}{f_1} - \left(1 - \frac{d}{f_1}\right)\left(1 - \frac{d}{f_2}\right)\right]。$$

这个例子的结果对研究激光谐振腔是非常有用的，而参数 f_1、f_2 及 d 的物理意义也将在那里明确。

文献[1]给出，

$$\begin{aligned} A_{T_n} &= \frac{1}{\sin\theta}[A\sin(n\theta) - \sin(n-1)\theta] \\ B_{T_n} &= \frac{1}{\sin\theta}B\sin(n\theta) \\ C_{T_n} &= \frac{1}{\sin\theta}C\sin(n\theta) \\ D_{T_n} &= \frac{1}{\sin\theta}[D\sin(n\theta) - \sin(n-1)\theta] \end{aligned} \tag{1.3-5}$$

其中

$$\begin{aligned} \theta &= \cos^{-1}\left[\frac{1}{2}(A + D)\right] \\ &= \cos^{-1}\left(1 - \frac{d}{f_1} - \frac{d}{f_2} + \frac{d^2}{2f_1f_2}\right) \end{aligned} \tag{1.3-6}$$

现在我们来证明这一结论，所用的方法仍然是数学归纳法。

当 $n=1$ 时，由式(1.3-5)很容易算出

$$A_{T_1} = A,\ B_{T_1} = B,\ C_{T_1} = C,\ D_{T_1} = D$$

故知结论正确。

设当 $n=k$ 时结论亦真，则当 $n=k+1$ 时，

$$\begin{bmatrix} A_{T_{k+1}} & B_{T_{k+1}} \\ C_{T_{k+1}} & D_{T_{k+1}} \end{bmatrix} = \begin{bmatrix} A_{T_k} & B_{T_k} \\ C_{T_k} & D_{T_k} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

由此，

$$\begin{aligned} A_{T_{k+1}} &= A_{T_k}A + B_{T_k}C \\ &= \frac{1}{\sin\theta}\{(A^2 + BC)\sin(k\theta) - A\sin(k\theta)\cos\theta + A\sin\theta\cos(k\theta)\} \end{aligned}$$

很容易验证

$$AD - BC = 1$$

9610022

并利用式(1.3-6)，得

$$A^2 + BC = A(A + D) - 1 = 2A\cos\theta - 1$$

于是，

$$A_{T_{k+1}} = \frac{1}{\sin\theta}\{A\cos\theta\sin(k\theta) + A\sin\theta\cos(k\theta) - \sin(k\theta)\}$$

$$= \frac{1}{\sin\theta}\{A\sin[(k + 1)\theta] - \sin(k\theta)\};$$

而

$$B_{T_{k+1}} = A_{T_k}B + B_{T_k}D$$

$$= \frac{1}{\sin\theta}B[(A + D)\sin(k\theta) - \sin(k - 1)\theta]$$

因为

$$(A + D)\sin(k\theta) - \sin(k - 1)\theta$$

$$= 2\cos\theta\sin(k\theta) - \sin(k\theta)\cos\theta + \cos(k\theta)\sin\theta$$

$$= \sin(k + 1)\theta;$$

所以，

$$B_{T_{k+1}} = \frac{1}{\sin\theta}B\sin(k + 1)\theta$$

类似地，

$$C_{T_{k+1}} = C_{T_k}A + D_{T_k}C$$

$$= \frac{1}{\sin\theta}C[(A + D)\sin(k\theta) - \sin(k - 1)\theta]$$

$$= \frac{1}{\sin\theta}C\sin(k + 1)\theta$$

$$D_{T_{k+1}} = C_{T_k}B + D_{T_k}D$$

$$= \frac{1}{\sin\theta}[(CB + D^2)\sin(k\theta) - D\sin(k - 1)\theta]$$

$$= \frac{1}{\sin\theta}\{D[(A + D)\sin(k\theta) - \sin(k - 1)\theta] - \sin(k\theta)\}$$

$$= \frac{1}{\sin\theta}[D\sin(k + 1)\theta - \sin(k\theta)]$$

这就是说，如果 $n=k$ 时式(1.3-5)成立，则 $n=k+1$ 时式(1.3-5)亦成立。于是，由数学归纳法可知，式(1.3-5)对任意自然数 n 成立。

下面我们对标量、矢量和张量的概念作进一步的讨论。

(1)标量　标量就是在正交变换下数值不变的量。例如，式(1.3-2)表明，矢量的长度是正交变换下的不变量，因而是标量。

(2)矢量　如前所述，矢量的长度在坐标轴的正交变换下保持不变。而矢量的各分量则按一定规律变化，现在就来寻找这个规律。

设一直角坐标系转动前可以用 $O-x_1x_2x_3$ 表示，沿坐标轴的单位矢量为 $\boldsymbol{e}_i(i=1,2,3)$，且

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$$

该坐标系转动后表示为 $O-x_1'x_2'x_3'$，沿坐标轴的单位矢量为 $\boldsymbol{e}'_i(i=1,2,3)$，且

$$\boldsymbol{e}_i' \cdot \boldsymbol{e}_j' = \delta_{ij}$$

于是，坐标系中的某一矢量 $\boldsymbol{B}$ 可以表示为

$$\boldsymbol{B} = \sum_{j=1}^{3} B_j \boldsymbol{e}_j = \sum_{k=1}^{3} B_k' \boldsymbol{e}_k' \tag{1.3-7}$$

其中 B_k' 和 B_i 分别为矢量 $\boldsymbol{B}$ 在新、旧坐标系中的分量。用 $\boldsymbol{e}_i'$ 点乘方程(1.3-7)两边，并利用基底矢量的正交归一性，立即得到

$$B_i' = \sum_{j=1}^{3} (\boldsymbol{e}_i' \cdot \boldsymbol{e}_j) B_j \tag{1.3-8}$$

将此式与式(1.3-1)进行比较即可看出，矢量的分量在正交变换下遵循与坐标轴相同的变换规律。若令

$$a_{ij} = \boldsymbol{e}_i' \cdot \boldsymbol{e}_j,$$

则可将方程(1.3-8)写为矩阵形式

$$B' = AB$$

其中

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

为变换矩阵。

顺便指出，微分算符遵循与矢量相同的变换规律，即

$$\frac{\partial}{\partial x_i'} = \sum_{j=1}^{3} a_{ij} \frac{\partial}{\partial x_j}$$

所以可称为矢量算符。

(3)张量　在物理学中，有许多量要用张量来描述，这将在后面有关章节中讨论。本节的任务是从数学上引出张量概念，并介绍一些简单的运算法则，进而导出其变换规律。

我们已经熟悉，矢量

$$\boldsymbol{U} = U_1\boldsymbol{e}_1 + U_2\boldsymbol{e}_2 + U_3\boldsymbol{e}_3$$

和矢量

$$\boldsymbol{V} = V_1\boldsymbol{e}_1 + V_2\boldsymbol{e}_2 + V_3\boldsymbol{e}_3$$

可以求内积得到

$$\boldsymbol{U} \cdot \boldsymbol{V} = U_1V_1 + U_2V_2 + U_3V_3;$$

也可以求外积得到

$$\begin{aligned} \boldsymbol{U} \times \boldsymbol{V} = {} & (U_2V_3 - U_3V_2)\boldsymbol{e}_1 + (U_3V_1 - U_1V_3)\boldsymbol{e}_2 \\ & + (U_1V_2 - U_2V_1)\boldsymbol{e}_3 \end{aligned}$$

现在我们定义一种新的运算，即矢量的外乘(或称并矢)。$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 的外乘为

$$\begin{aligned} \boldsymbol{UV} &= (U_1\boldsymbol{e}_1 + U_2\boldsymbol{e}_2 + U_3\boldsymbol{e}_3)(V_1\boldsymbol{e}_1 + V_2\boldsymbol{e}_2 + V_3\boldsymbol{e}_3) \\ &= U_1V_1\boldsymbol{e}_1\boldsymbol{e}_1 + U_1V_2\boldsymbol{e}_1\boldsymbol{e}_2 + U_1V_3\boldsymbol{e}_1\boldsymbol{e}_3 \\ &\quad + U_2V_1\boldsymbol{e}_2\boldsymbol{e}_1 + U_2V_2\boldsymbol{e}_2\boldsymbol{e}_2 + U_2V_3\boldsymbol{e}_2\boldsymbol{e}_3 \\ &\quad + U_3V_1\boldsymbol{e}_3\boldsymbol{e}_1 + U_3V_2\boldsymbol{e}_3\boldsymbol{e}_2 + U_3V_3\boldsymbol{e}_3\boldsymbol{e}_3 \\ &= \sum_{i,j=1}^{3} U_iV_j\boldsymbol{e}_i\boldsymbol{e}_j \end{aligned}$$

两矢量求内积的结果产生标量，求外积的结果给出一个新的矢量，求外乘的结果则得到一个二阶张量。如果用顶上带两个箭头的字母表示二阶张量，则可记

$$\boldsymbol{UV} = \overleftrightarrow{\boldsymbol{T}}, \qquad U_i V_j = T_{ij},$$

和

$$\overleftrightarrow{\boldsymbol{T}} = \sum_{i,j=1}^{3} T_{ij}\boldsymbol{e}_i\boldsymbol{e}_j \tag{1.3-9}$$

其中

$$(T_{ij}) = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}$$

是张量 $\overleftrightarrow{\boldsymbol{T}}$ 的矩阵表示。

一个二阶张量可以继续与一个矢量求外乘并得到一个三阶张量。例如，

$$\overleftrightarrow{\boldsymbol{T}}\,\boldsymbol{W} = \sum_{i,j,k=1}^{3} T_{ij}W_k\boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k$$

两个二阶张量也可以求外乘，结果得到一个四阶张量。例如，

$$\overleftrightarrow{\boldsymbol{T}}^{(1)}\overleftrightarrow{\boldsymbol{T}}^{(2)} = \sum_{i,j,k,l=1}^{3} T_{ij}^{(1)}T_{kl}^{(2)}\boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k\boldsymbol{e}_l$$

循此继续下去，假定高阶张量之间也可进行外乘，则可通过这种办法得到任意阶张量。

为了寻找张量的分量在正交变换下的变换规律，先要定义有关张量的两条基本运算法则。

(1)张量与矢量点乘

张量与矢量的点乘归结为并矢$(\boldsymbol{e}_i\boldsymbol{e}_j)$与单位矢量 $\boldsymbol{e}_k$ 的点乘，若定义

$$(\boldsymbol{e}_i\boldsymbol{e}_j)\cdot\boldsymbol{e}_k = \boldsymbol{e}_i(\boldsymbol{e}_j\cdot\boldsymbol{e}_k)$$

则有

$$\overleftrightarrow{\boldsymbol{T}}\cdot\boldsymbol{U} = \sum_{i,j,k=1}^{3} T_{ij}U_k\boldsymbol{e}_i(\boldsymbol{e}_j\cdot\boldsymbol{e}_k)$$

注意到基矢的正交归一性，即

$$\boldsymbol{e}_j\cdot\boldsymbol{e}_k = \delta_{jk},$$

得到

$$\overleftrightarrow{\boldsymbol{T}}\cdot\boldsymbol{U} = \sum_{i,j=1}^{3} T_{ij}U_j\boldsymbol{e}_i,$$

很显然，这是一个矢量。也就是说，一个二阶张量与一个矢量点乘的结果得到一个新的矢量，其分量由参加点乘的张量与矢量的分量所决定。类似地，若定义

$$\boldsymbol{e}_k\cdot(\boldsymbol{e}_i\boldsymbol{e}_j) = (\boldsymbol{e}_k\cdot\boldsymbol{e}_i)\boldsymbol{e}_j,$$

则有

$$\boldsymbol{U}\cdot\overleftrightarrow{\boldsymbol{T}} = \sum_{i,j,k=1}^{3} T_{ij}U_k\delta_{ik}\boldsymbol{e}_j$$

$$= \sum_{i,j}^{3} T_{ij}U_i\boldsymbol{e}_j$$

这也是一个矢量，但它与前面得到的矢量并不相等，这是不难理解的，因为

$$(\boldsymbol{e}_i\boldsymbol{e}_j)\cdot\boldsymbol{e}_k = \boldsymbol{e}_i(\boldsymbol{e}_j\cdot\boldsymbol{e}_k) = \delta_{jk}\boldsymbol{e}_i$$

和