高等量子力学

（理论原理）

Advanced Quantum Mechanics

(Theoretical Foundation)

◆ 吴兆颜

吉林大学出版社
JILIN UNIVERSITY PRESS

高等量子力学（理论原理）

Advanced Quantum Mechanics

(Theoretical Foundation)

责任编辑／安　斌
封面设计／孙　群



定价：42.00元

0413.1/76

2008

高等量子力学

（理论原理）

Advanced Quantum Mechanics

（Theoretical Foundation）

吴兆颜

吉林大学出版社

图书在版编目（CIP）数据

高等量子力学 / 吴兆颜著. —长春：吉林大学出版社，2008
ISBN 978－7－5601－3820－6

Ⅰ. 高… Ⅱ. 关… Ⅲ. 量子力学－研究生－教材
Ⅳ.0413.1
中国版本图书馆 CIP 数据核字（2008）第 041446 号

书　名：高等量子力学
作　者：吴兆颜　著
责任编辑、责任校对：安　斌　　封面设计：孙　群
吉林大学出版社出版、发行　　吉林科华印刷厂　印刷
开本：787×960 毫米　1/16　　2008 年 4 月第 1 版
印张：20.625　字数：330 千字　　2008 年 4 月第 1 次印刷
ISBN 978－7－5601－3820－6　　定价：42.00 元

版权所有　翻印必究
社址：长春市明德路 421 号　邮编：130021
发行部电话：0431－88499826
网址：http: //www.jlup.com.cn
E-mail：jlup@mail.jlu.edu.cn

前　言

整个现代物理学是建立在量子力学和相对论这两大理论支柱之上的。特别是量子力学，它是从基本粒子物理、核物理、原子分子物理、纳米物理、光学、等离子体物理、凝聚态物理，一直到天体物理、宇宙学，物理科学各个分支的理论基础。此外，量子力学在自然科学的其他分支，如材料科学、化学、生命科学等，和工程技术，如计算机技术、电子工程、通讯工程等，也有重要的应用。在当今世界上各个高等学校中，量子力学已是物理系大学本科生必修的基础理论课之一。而高等量子力学则是物理学科各专业研究生必修的基础理论课之一。八十年代末以来，作者曾在吉林大学多次讲授高等量子力学课，形成了一份讲义。本书是在此讲义的基础之上，考虑教学和科研实践的经验，并参考其他高等量子力学教材，经过多次修改写就的。

上面所说的物理学理论支柱之一的（广义的）量子力学，既包括讨论微观物理体系运动规律的非相对论性理论（狭义的量子力学），又包括讨论微观物理体系运动规律的相对论性理论（量子场论）。作为物理学科各专业研究生高等量子力学课的教材，本书只讨论非相对论性的量子力学。

微观物理体系的行为和宏观物理体系的行为是如此地不同，这告诉我们不得把我们熟悉的经典物理图象强加到微观物理体系上。因而建立正确的量子概念，不能过多地依赖直观。作者认为，对 Hilbert 空间及其上线性算符的理论有一个很好的了解，是建立正确的量子概念，学好量子力学理论并掌握应用量子力学处理物理问题的方法的前提条件之一。在本书的第 0 章，作者将和读者一起回顾有关的数学预备知识。这些数学知识，原则上包含在各种数学教科书中。但是数学书很少是专为学习量子力学而写的。学生要花很多时间去翻阅许多书，还不一定能找到所需要的知识。作者希望第 0 章能帮助读者，花最少时间而能较系统地掌握学习量子力学所

需要的数学知识。

读者在大学本科的量子力学课程中，已学习过量子概念和量子力学发生和发展的历史，并对量子力学的基本理论和应用有了一个初步的认识。本书的讨论将从重述量子力学的基本原理开始。作者将不涉及自量子力学建立以来就有的，关于量子力学基础的那些争论。这不是因为作者认为它们不重要，而是因为它们还没有达成公认的结论。在以后的各章，作者向读者介绍了对称性分析，角动量理论，二次量子化形式和相对论单粒子波动方程。这些内容，在我们看来，是在篇幅有限的情况下，最值得首先介绍给读者的。

尽管本书主要是为物理学各专业研究生写的教材，作者希望物理学工作者和大学物理教师也能从中找到有用的东西。

目　　录

第0章　一些数学预备知识 …… (1)
§0.0　集合论中的一些概念和记法 …… (1)
Ⅰ.集合论的一些记法 …… (2)
Ⅱ.映射 …… (3)
§0.1　线性空间 …… (5)
Ⅰ.基本概念 …… (5)
Ⅱ.线性相关性与维数 …… (6)
Ⅲ.子线性空间 …… (8)
§0.2　内积空间、希尔伯特空间 …… (9)
Ⅰ.内积空间 …… (9)
Ⅱ.希尔伯特空间 …… (11)
§0.3　线性算符 …… (12)
Ⅰ.基本概念 …… (12)
Ⅱ.线性算符的矩阵 …… (13)
Ⅲ.特征子空间 …… (16)
Ⅳ.不可约线性算符组 …… (19)
Ⅴ.两两对易的线性算符组 …… (20)
Ⅵ.投影算符 …… (21)
* §0.4　线性算符的指数函数 …… (23)
Ⅰ.内积空间上线性算符的范数与算符序列的极限 …… (23)
Ⅱ.矩阵的指数函数 …… (25)
Ⅲ.矩阵的幂级数 …… (28)

Ⅳ. Baker-Hausdorff 公式 …… (30)
§0.5 对偶空间、线性空间的张量积 …… (34)
Ⅰ. F – 线性空间的对偶空间 …… (34)
Ⅱ. F – 内积空间的对偶空间 …… (34)
Ⅲ. Dirac 符号 …… (35)
Ⅳ. 线性空间的张量积 …… (36)
* §0.6 非对易分析与超算符
(Non-commutative Calculus and Hyper-operators) …… (37)
Ⅰ. 向量的范数(norm) …… (38)
Ⅱ. 算符函数的微商 …… (40)
Ⅲ. 超(线性)算符 …… (41)
Ⅳ. 用超算符表达算符函数的微商的示例 …… (42)
Ⅴ. 高阶微商与 Taylor 展开 …… (43)
Ⅵ. 超算符应用示例 …… (45)
第 1 章 量子力学的基本原理 …… (47)
§1.1 量子力学的基本假定 …… (47)
Ⅰ. 波函数公设 …… (47)
Ⅱ. 算符公设 …… (49)
Ⅲ. 力学量平均值公设 …… (50)
Ⅳ. 动力学演化方程 …… (52)
Ⅴ. 全同性原理 …… (53)
§1.2 演化算符 …… (54)
Ⅰ. 演化算符的引入和一般性质 …… (54)
Ⅱ. 演化算符的方程、形式解 …… (55)
Ⅲ. 一个求解含时 Schrödinger 方程的方法 …… (57)
§1.3 表象与绘景(Representations and Pictures) …… (61)
Ⅰ. 量子力学的矩阵形式 …… (61)
Ⅱ. 量子力学的各种绘景 …… (64)
Ⅲ. 对绘景的另一种理解 …… (68)

§1.4 密度算符 …… (71)
Ⅰ.纯态系综与混态系综 …… (71)
Ⅱ.混态系综的描述 …… (73)
Ⅲ.密度算符的性质 …… (76)
Ⅳ.密度算符的演化 …… (77)
Ⅴ.约化密度算符 …… (78)
第2章 量子力学中的对称性分析 …… (81)
§2.1 物理学中的对称性 …… (82)
§2.2 空间平移 …… (83)
Ⅰ.空间平移群 …… (83)
Ⅱ.空间平移算符(态矢量在平移下的变换) …… (83)
Ⅲ.力学量算符在平移下的变换 …… (85)
Ⅳ.体系的平移不变性与动量守恒 …… (86)
§2.3 空间绕固定点的转动 …… (87)
Ⅰ.三维空间旋转群$SO(3)$ …… (87)
Ⅱ.单个无旋粒子体系态矢量的转动 …… (92)
Ⅲ.单个无旋粒子体系力学量算符的转动 …… (95)
Ⅳ.任意一微观体系的转动 …… (98)
* §2.4 态矢量空间的幺正变换和反幺正变换 …… (102)
附录 …… (109)
§2.5 空间反射 …… (112)
Ⅰ.空间反射群 …… (112)
Ⅱ.态矢量和力学量算符在空间反射下的变换性质 …… (113)
Ⅲ.宇称守恒与破缺 …… (116)
§2.6 时间平移与时间反演 …… (118)
Ⅰ.时间均匀性与能量守恒 …… (119)
Ⅱ.经典力学中时间反演 …… (120)
Ⅲ.量子力学中时间反演 …… (121)
Ⅳ.角动量本征态的时间反演性质 …… (123)

Ⅴ.微观运动的可逆性 …… (125)
§2.7 力学量算符对易关系的几何根源 …… (126)
Ⅰ.三维欧几里德群$E(3)$ …… (128)
Ⅱ.$E(3)$群在态矢量空间 S 上的表示 …… (130)
§2.8 量子体系的对称性群(symmetry group)与能级简并 …… (131)
Ⅰ.量子体系的对称变换 …… (131)
Ⅱ.能级简并与对称性 …… (133)
Ⅲ.氢原子的动力学对称性 …… (134)
Ⅳ.再论量子体系的对称性群 …… (144)
附录 …… (145)
第3章 角动量理论 …… (147)
§3.1 角动量算符的形式定义 …… (147)
Ⅰ.旋转群在态矢量空间上的酉表示的无穷小算符 …… (147)
Ⅱ.绕固定点转动的刚体的角动量算符 …… (152)
§3.2 单个角动量的理论 …… (156)
Ⅰ.$\boldsymbol{J}$ – 不可约不变子空间的转动性质 …… (157)
Ⅱ.$SO(3)$群表示空间的分解 …… (163)
§3.3 转动算符在标准基下的矩阵 …… (165)
Ⅰ.欧拉角与旋转群的不变体积元 …… (165)
Ⅱ.转动矩阵的一般性质 …… (167)
Ⅲ.D – 函数的解析表达式 …… (170)
Ⅳ.D – 函数的对称性 …… (175)
§3.4 两个角动量的耦合 …… (177)
Ⅰ.和角动量的多重态子空间 …… (178)
Ⅱ.Clebsch-Gordan 系数 …… (181)
Ⅲ.$C.-G.$系数的对称性关系与 3 – j 符号 …… (188)
Ⅳ.$C.-G.$系数与 D – 函数 …… (191)
Ⅴ.两个对易的角动量的相加 …… (193)
附录 …… (197)

§3.5 不可约张量算符 …… (198)
Ⅰ.不可约张量算符的定义 …… (198)
Ⅱ.Wigner-Eckart 定理 …… (202)
Ⅲ.约化矩阵元的例子 …… (204)
Ⅳ.约化矩阵元的耦合定律 …… (205)
§6 多个角动量的耦合与 Racah 系数 …… (208)
Ⅰ.三个角动量相加的再耦合系数 …… (209)
Ⅱ.6-j 符号的性质 …… (213)
Ⅲ.9-j 符号 …… (215)
第 4 章 二次量子化 …… (217)
§4.1 全同粒子体系的态空间与力学量 …… (218)
§4.2 粒子产生与消灭算符 …… (225)
§4.3 用产生湮灭算子表达力学量算符 …… (231)
Ⅰ.全同玻色子体系的 k-体算符 …… (232)
Ⅱ.全同费米子体系的 k-体算符 …… (235)
Ⅲ.写下全同粒子体系的 k-体算符的示例 …… (238)
§4.4 二次量子化量子力学的运动方程 …… (241)
Ⅰ.二次量子化对传统形式量子力学的继承与发展 …… (242)
Ⅱ.二次量子化一词的来源 …… (242)
Ⅲ.求解二次量子化形式的薛定格方程的示例
(原子与光场相互作用的 Jaynes-Commings 模型) …… (246)
第 5 章 相对论性单粒子波动方程 …… (251)
§5.1 Klein-Gordon 方程 …… (252)
§5.2 自由电子的 Dirac 方程 …… (256)
Ⅰ.Dirac 方程的建立 …… (256)
Ⅱ.几率守恒,电子自旋 …… (259)
Ⅲ.Dirac 方程的平面波解系 …… (260)
Ⅳ.Dirac 方程的 Lorentz 协变性 …… (264)
附录 Dirac 矩阵 γ^{μ} …… (268)

§5.3 电磁场中无旋粒子的 klein-Gordon 方程 …… (271)
Ⅰ.方程的建立 …… (271)
Ⅱ.π 介子原子 …… (272)
§5.4 电磁场中电子的 Dirac 方程 …… (276)
Ⅰ.方程的建立 …… (276)
Ⅱ.泡利方程—非相对论极限与电子内禀磁矩 …… (276)
Ⅲ.中心力场下的非相对论近似,旋轨耦合 …… (278)
Ⅳ.氢原子光谱的精细结构 …… (279)
附录 群表示论简介 …… (281)
§1. 群的概念 …… (281)
Ⅰ.抽象群的定义 …… (281)
Ⅱ.子群与陪集 …… (283)
Ⅲ.同态的基本定理 …… (284)
§2 群表示理论 …… (285)
Ⅰ.群表示的概念 …… (285)
Ⅱ.等价表示 …… (287)
Ⅲ.既约表示与可约表示 …… (289)
Ⅳ.正交性定理与完备性定理 …… (290)
Ⅴ.特征标 …… (292)
§3 李群表示的几个定理 …… (294)
Ⅰ.r 维李群 …… (294)
Ⅱ.不变体积元 …… (296)
Ⅲ.李群表示与无穷小算符 …… (297)
§4 李群与李代数 …… (303)
Ⅰ.李代数的概念 …… (303)
Ⅱ.矩阵李群的李代数 …… (304)
Ⅲ.常见的矩阵李群与它们的李代数 …… (305)
§5 $SO(3)$群与 $SU(2)$群 …… (310)
§6 Lorentz 群 …… (311)

Ⅰ.狭义相对论的时空变换 …… (311)
Ⅱ.Lorentz 群$O(3,1)$ …… (313)
Ⅲ.无穷小 Lorentz 变换,李代数$o(3,1)$ …… (316)

第0章 一些数学预备知识

数学一向是理论物理学的语言。为学好一门物理理论，你必须对相关的数学有一个很好的了解。与经典物理学相比，量子理论要抽象得多。这就要求学习量子理论的人在相关的数学知识方面有一个更坚实的基础。因此本书第0章将介绍学习高等量子力学所需要的一些数学预备知识。

读者在大学本科量子力学课程的学习中已经知道，一个微观物理体系的状态是由波函数来描述的。表象理论告诉我们，这些波函数是某个希尔伯特空间中的向量。体系的力学量是由定义在此希尔伯特空间上的厄米算符所表示的。在这个意义上我们说，量子力学的数学框架是希尔伯特空间及其上线性算符的理论。而希尔伯特空间的概念是由线性空间和内积空间的概念发展而来的，熟悉线性空间和内积空间上线性算符的理论，是学好希尔伯特空间上线性算符理论的前提条件。现代数学，包括代数学在内，是建立在集合论基础之上的。因此，本章将首先介绍集合论的一些概念和记法，接着讨论线性空间、内积空间及其上线性算符的理论，最后介绍希尔伯特空间及其上线性算符的理论。我们的讲述将以物理需要为取舍准则，并尽量照顾数学理论本身的系统性。然而本章毕竟带有复习性质，希望要求自己严格的读者，能参考相应的数学书，系统地掌握本章所介绍的数学预备知识。群论已经是物理专业研究生的必修课，掌握本章内容也是学习群表示论的前提条件。

§0.0 集合论中的一些概念和记法

现代数学的特征之一是使用集合论的语言。它是最精确的语言。花一定时间去掌握它，并养成使用这种语言思考的习惯，对一个物理学工作者是必要的，也是明智

的。建议读者一定要非常认真地学习这些基本概念。你在这上面所花的每一分努力,都将会得到数倍的回报。

Ⅰ. 集合论的一些记法

我们假定读者知道集合论的最基本的知识,在这里只列出常用的一些集合论的记法。

$x\in S$ 读作 x 是集合 S 的元素,或者 x 属于集合 S。

$A\subset S$ 读作集合 A 是集合 S 的子集,或者集合 A 包含在集合 S 内。意思是凡属于 A 的事物一定属于 S。若 $A\subset S\&A\neq S$,则称 A 是 S 的真子集。

$\varnothing$ 代表空集,即一个元素也没有的集。

显然对任一集 S 有 $\varnothing\subset S$ 和 $S\subset S$。称 $\varnothing$ 和 S 为 S 的平庸子集。

$A\cup B$ 读作集 A 与集 B 的并集。它是由所有至少属于 A、B 之一的事物构成的。

$A\cap B$ 读作集 A 与集 B 的交集。它是由所有同时属于 A 和 B 的事物构成的。

$A\setminus B$ 读作集 A 与集 B 的差集。它是由所有属于 A 而不属于 B 的事物构成的。

2^S 读作集 S 的幂集。它是由集 S 的所有子集构成的集合。

$A\times B$ 读作集 A 与集 B 的笛卡儿积集。它是由全体这样的元素对 (a,b),其中 $a\in A$ 而 $b\in B$,构成的集合。用符号可以记作 $A\times B\equiv\{(a,b)\mid a\in A,b\in B\}$。

$|A|$ 读作集合 A 的基数(cardinal),表示 A 中元素的个数。基数可以是 0,1,2,…和各种超穷数,例如,可数集的基数 $\aleph_0$(分离无穷),连续统的基数 $\aleph$(连续无穷),等等。

设 S 是非空集,$\mathscr{M}$ 是 S 的一些非空子集的集合。若 $\mathscr{M}$ 的不同元交空,且 $\mathscr{M}$ 的所有元的并等于 S,则称 $\mathscr{M}$ 为 S 的一个分拆。又设 $\mathscr{N}$ 也是 S 的一个分拆。若 $\mathscr{M}$ 的每一个元都包含在 $\mathscr{N}$ 的一个元中,则称分拆 $\mathscr{M}$ 细于分拆 $\mathscr{N}$。

练习 0-0-1 试证:对任意集合 A、B、C 都有(并对于交,交对于并的分配律)

$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$,$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$。

练习 0-0-2 试证:对任意集合 A、B、C 都有(de-Morgan 公式)

$A\setminus(B\cup C)=(A\setminus B)\cap(A\setminus C)$,$A\setminus(B\cap C)=(A\setminus B)\cup(A\setminus C)$。

练习 0-0-3 设 A 是有限集，以 $|A|$ 记 A 的元素的个数。试证：$|2^A| = 2^{|A|}$。

练习 0-0-4 试证：对任意集合 A_1, A_2, B_1, B_2，都有

$(A_1 \times A_2) \cap (B_1 \times B_2) = (A_1 \cap B_1) \times (A_2 \cap B_2)$，$(A_1 \times A_2) \cup (B_1 \times B_2) \subset (A_1 \cup B_1) \times (A_2 \cup B_2)$。

Ⅱ. 映射

映射是最基本的数学概念之一。它是读者在数学分析中所学到的函数概念的推广。通常这一概念是这样定义的：设 A 和 X 是非空集，若按某一规则 f，对集 A 的每一元素 a，都有集 X 的一个确定的元素 x 与之对应，则称 f 是集 A 到集 X 的一个映射，记作 $f: A \to X$。称与 $a \in A$ 对应的集 X 的元 x 为 a 在映射 f 下的象，记作 $x = f(a)$。

映射的这一定义可以用集合论的语言表达得更确切。

定义 设 A 和 X 是非空集，$f \subset A \times X$。若

(1) 对每一 $a \in A$，都有 $x \in X$，使 $(a, x) \in f$，

(2) 由 $(a, x) \in f$ 和 $(a, y) \in f$，可推出 $x = y$，

则称 f 是集 A 到集 X 的一个映射，记作 $f: A \to X$。若 $(a, x) \in f$，则称 x 为 a 在映射 f 下的象，记作 $x = f(a)$，或者 $f: a \mapsto x$。集 A 称为映射 f 的定义域，集 X 称为映射 f 的值域。

在映射的这个定义中，定义域集 A 与值域集 X 的地位是不平等的：对集 A 的每一元素，与之对应的集 X 的元素的个数必为 1；而对集 X 的元素，与之对应的集 A 的元素的个数可以是任一基数(cardinal)，即可以是 0、1、2 …和各种超穷数。

设 f 是集 A 到集 X 的一个映射，$B \subset A$，$Y \subset X$。称 X 的子集 $\{x \in X \mid \exists\, b \in B$，使 $x = f(b)\}$ 为 B 在 f 下的象，记作 $f(B)$。特别是，称 $f(A)$ 为映射 f 的象，记作 $\mathrm{Im}f$。称 A 的子集 $\{a \in A \mid f(a) \in Y\}$ 为 Y 在 f 下的完全逆象，记作 $f^{-1}(Y)$。特别是，当 Y 是独点集 $\{y\}$ 时，$f^{-1}(Y)$ 也称为 y 的完全逆象，记作 $f^{-1}(Y)$。考虑下面的定义域 A 的子集的集合 $\{f^{-1}(x) \in 2^A \mid x \in \mathrm{Im}f\}$。显然它是 A 的一个分拆，我们称之为映射 f 确定的分拆。

设有两个映射 $f: A \to X$，$g: B \to Y$。若对每一 $(a, b) \in A \times B$，令 $(f(a), g(b)) \in X \times Y$ 与之对应，则我们得到一个 $A \times B$ 到 $X \times Y$ 的映射，称之为映射 f 与 g 的笛卡

尔积，记作 $f\times g: A\times B\to X\times Y$。

设 $f: A\to X$ 是映射，$S\subset A$。若对每一 $s\in S$，令 $f(s)\in X$ 与之对应，则我们得到一个 S 到 X 的映射，称之为 f 到 S 上的限制，记作 $f|_S: S\to X$。

设 $f: A\to X$ 是映射。当 $f(A)=X$ 时，称 f 为满射(surjection)。当 A 中的不同元素在 f 下的象一定不同时，即当 $f(a)=f(b)\Rightarrow a=b$ 成立时，称 f 为单射(injection)。一个既是单射又是满射的映射称为双射(bijection)，也就是平常所说的双方一一对应。

设 $f: A\to X$ 是双射，则 $X\times A$ 的子集 $\{(f(a),a)\mid a\in A\}$ 是集 X 到集 A 的一个映射，称之为 f 的逆映射，记作 $f^{-1}: X\to A$。显然 f^{-1} 也是双射，而且 $(f^{-1})^{-1}=f$。

设 A，B 是非空集，以 B^A 记全体 A 到 B 的映射的集合，这是因为 $|B^A|=|B|^{|A|}$。

设有两个映射 $f: A\to M$ 和 $g: M\to X$。不难看出 $A\times X$ 的子集 $\{(a,g(f(a)))\mid a\in A\}$ 是集 A 到集 X 的一个映射，称之为 g 与 f 的乘积或复合，记作 $g\circ f: A\to X$。又设 $h: X\to Z$，则不难验证 $h\circ(g\circ f)=(h\circ g)\circ f$。这就是说映射的乘法满足结合律。由于结合律，我们可以写 $h\circ g\circ f$，而不引起歧义。

设 A 是非空集。$A\times A$ 的子集 $\{(a,a)\mid a\in A\}$ 是 A 到自身的一个映射，它把 A 的每一个元素映到自身，称作集 A 的恒等映射，记作 $1_A: A\to A$。若 $f: A\to X$ 是双射，则显然有 $f\circ f^{-1}=1_X$，$f^{-1}\circ f=1_A$。

设有映射 $f: A\to B$，$g: B\to C$，$h: C\to D$，$j: A\to M$ 和 $k: M\to D$。若有 $h\circ g\circ f=k\circ j$ 则称下图交换：

$$\begin{array}{ccccc} A & \xrightarrow{\ f\ } & B & \xrightarrow{\ g\ } & C \\ \downarrow j & & & & h\downarrow \\ M & & \xrightarrow{\quad k\quad} & & D \end{array}$$

练习 0-0-5 设有映射 $f: A\to X$，以及 $B_i\subset A$，$Y_i\subset X(i=1,2)$。试证：

$f(B_1\cup B_2)=f(B_1)\cup f(B_2)$，$f(B_1\cap B_2)\subset f(B_1)\cap f(B_2)$；

$f^{-1}(Y_1\cup Y_2)=f^{-1}(Y_1)\cup f^{-1}(Y_2)$，$f^{-1}(Y_1\cap Y_2)=f^{-1}(Y_1)\cap f^{-1}(Y_2)$。

练习 0-0-6 设有映射 $f: A\to X$ 和 $g: A\to M$。试证存在映射 $h: M\to X$ 使 $f=h\circ g$ 的充分必要条件是 g 确定的分拆细于 f 确定的分拆。

练习 0-0-7 设 A, B 是非空有限集。试证:集 A 到集 B 的映射共有 $|B|^{|A|}$ 个,即 $|B^A| = |B|^{|A|}$。

§0.1 线性空间

线性空间或许是读者在高等代数课程中所遇到的第一个用公理定义的代数系。我们希望读者通过本章的复习,能系统地掌握线性空间的理论,而不是仅仅知道有关它的一些零散的知识。由于是复习,我们只简要地介绍基本定义和定理,很多定理我们以练习的形式留给读者自己证明。如果读者能够顺利地做出这些练习,则可以越过相应的正文。否则,你应该参考线性代数书,仔细地阅读本节内容。掌握高等代数中的线性代数方程组的理论是学好本节内容的前提条件之一。

Ⅰ. 基本概念

定义 设 V 是非空集,称它的元为向量,而 F 是数域(复数域、实数域或有理数域)。若定义了两个映射 $a: V \times V \to V$ 和 $m: F \times V \to V$,对所有 $x, y, z \in V$ 和 $\lambda, \mu \in F$,称 $a(x, y)$ 为向量 x 和向量 y 的和,记作 $x + y$ 称 $m(\lambda, x)$ 为数 λ 与向量 x 的积,记作 λx(或 $x\lambda$),满足下面条件:

(1) $(x + y) + z = x + (y + z)$,(结合律)

(2) 存在 $0 \in V$,使 $0 + x = x$,称 0 为零向量,

(3) 存在 $-x \in V$,使 $(-x) + x = 0$,称 $-x$ 为 x 的负向量,

(4) $x + y = y + x$,(交换律)

(5) $\lambda(\mu x) = (\lambda\mu)x$,

(6) $1x = x$,

(7) $(\lambda + \mu)x = \lambda x + \mu x$,(向量的数乘关于数的加法的分配律)

(8) $\lambda(x + y) = \lambda x + \lambda y$,(向量的数乘关于向量的加法的分配律) (0.1.1)

则称 V 为复数域(实数域或有理数域) F 上的线性空间,或复(实或有理)线性空间。

例 1 所有 $m \times n$ 复矩阵的集合,在矩阵加法和复数与复矩阵的乘法下是一个

复线性空间。

例 2 所有定义在开区间(a,b)上的实值函数的集合,在函数加法和实数与实值函数的乘法下是一个实线性空间。

例 3 全体复系数一元多项式的集合,在多项式加法和复数与复多项式的乘法下是一个复线性空间。

例 4 实数域在实数加法和实数乘法下是一实线性空间。复数域在复数加法和复(实)数与复数的乘法下是一个复(实)线性空间。

设 V 是数域 F 上的线性空间,由定义可以推出

(1) V 的零向量是唯一的。

事实上,若 0_1 和 0_2 是 V 的零向量,则有 $0_2 = 0_1 + 0_2 = 0_2 + 0_1 = 0_1$。

(2)对 V 的每一向量 x,其负向量是唯一的。

事实上,若 y,z 是 x 的负向量,则有

$$y = 0 + y = (z + x) + y = z + (x + y) = (y + x) + z = 0 + z = z$$

(3) $0x = 0,(-1)x = -x,\lambda 0 = 0$。(其证明留作练习)

(4)若 $\lambda x = 0$,则 $\lambda = 0$,或者 $x = 0$。

事实上,设 $\lambda x = 0$,若 $\lambda \neq 0$,以 λ^{-1}乘等式两边得 $\lambda^{-1}(\lambda x) = \lambda^{-1}0$,此即$(\lambda^{-1}\lambda)x = 0, 1x = 0, x = 0$。

练习 0-1-1 试由线性空间定义中的其他条件推出条件(4)来。

练习 0-1-2 试证:线性空间定义中的条件(6)可以等价地换成向量的数乘 $m: F \times V \to V$ 是满射。

练习 0-1-3 设 V 是数域 F 上的线性空间,试证:

$0x = 0,(-1)x = -x,\lambda 0 = 0, \forall x \in V, \lambda \in F$。

Ⅱ. 线性相关性与维数

设 $x_1, x_2 \cdots, x_s$,是 F - 线性空间 V 的一组向量,$\alpha^1, \alpha^2, \cdots, \alpha^s$ 是 F 中的一组数(纯量)。称向量 $\alpha^1 x_1 + \alpha^2 x_2 + \cdots + \alpha^s x_s$ 为 $x_1, x_2, \cdots, x_s$ 的一个**线性组合**,称 $\alpha^1, \alpha^2, \cdots, \alpha^s$ 为该线性组合的系数。

定义 设$\{x_1, x_2, \cdots, x_s\}$是 F - 线性空间 V 的一个有限向量组,若存在 F 中的一组不全为零的数$\alpha^1, \alpha^2, \cdots, \alpha^s$,使 $\alpha^1 x_1 + \alpha^2 x_2 + \cdots + \alpha^s x_s = 0$,则称设$\{x_1, x_2, \cdots, x_s\}$

是**线性相关**的。称 V 的任一向量组(无限或有限)**是线性相关**的,当它包含线性相关的有限子组。

我们称线性空间 V 中的某个向量组是完备的或张成 V,若 V 的每一个向量都可以表作该向量组中向量的线性组合。完备的向量组也叫线性空间 V 的生成组。

练习 0-1-4 试证:若 $x_1, x_2, \cdots, x_s$ 是线性相关的,则任一包含 $x_1, x_2, \cdots, x_s$ 的向量组也是线性相关的(亦即,若向量组 $x_1, x_2, \cdots, x_s$ 是线性无关的,则它的任一部分组也是线性无关的)。

练习 0-1-5 试证:若向量组 $x_1, x_2, \cdots, x_s (s \geq 2)$ 是线性相关的,当且仅当向量组中有一个向量等于其他向量的一个线性组合。

练习 0-1-6 试证:任一包含零向量的向量组是线性相关的。

练习 0-1-7 试证:设 $x_1, x_2, \cdots, x_s$ 和 $y_1, y_2, \cdots, y_t$ 是 F - 线性空间 V 的两个向量组。若每一个 y_i 都可以表作第一组向量的线性组合,且 $t > s$,则向量组 $y_1, y_2, \cdots, y_t$ 是线性相关的。

定义 设 $x_1, x_2, \cdots$ 和 $y_1, y_2, \cdots$ 是 F - 线性空间 V 的两个向量组。若每一个 y_i 都可以表作第一组向量的线性组合,而每一个 x_i 都可以表作第二组向量的线性组合,则称这两个向量组**等价**。

定义 设 $x_1, x_2, \cdots$ 是 F - 线性空间 V 的向量组,$x_{i(1)}, x_{i(2)}, \cdots, x_{i(t)}$ 是它的一个部分组。若后者是线性无关的,且每一个 $x_j (1 \leq j)$ 都可表作该部分组中向量的线性组合,则称此部分组是原向量组的**极大线性无关组**。不难证明,相互等价的向量组的极大线性无关组包含相同个数的向量(请读者证明之)。特别是,V 的各个极大线性无关组包含相同个数的向量。

定义 设 V 是数域 F 上的线性空间。若存在 V 的线性无关的向量组 $e_1, e_2, \cdots, e_n$,使 V 的每一个向量都可以表作它们的线性组合,则称 V 的维数是 n,记作 $\dim V = n$。由向量组 $e_1, e_2, \cdots, e_n$ 的线性无关性可推出,任一向量 $v \in V$ 表成向量组 $e_1, e_2, \cdots, e_n$ 的线性组合的方式 $v = \lambda^1 e_1 + \cdots + \lambda^n e_n = \sum_{\alpha=1}^{n} e_\alpha \lambda^\alpha \equiv e_\alpha \lambda^\alpha$(我们将采用爱因斯坦求和约定)是唯一的。我们可以用线性组合的系数构成的 $n \times 1$ 矩阵 $\begin{bmatrix} \lambda^1 \\ \vdots \\ \lambda^n \end{bmatrix} \equiv [v]$ 来

刻画 V 中的向量。称向量组 $e_1, e_2, \cdots, e_n$ 为 V 的一个**基**。

设 V 是 n 维 $-F$ 线性空间，$x_1, x_2, \cdots, x_n$ 是 V 中的线性无关的向量组，则 V 中的任一向量都可唯一地表作它们的线性组合。(请读者证明之)

由于这一事实，我们可选 n 维 $F-$线性空间 V 的任一线性无关的 n 向量组为 V 的一个基。

练习 0-1-8 试证：相互等价的向量组的极大线性无关组包含相同个数的向量。

练习 0-1-9 试证：设 V 是 n 维 $F-$线性空间，$x_1, x_2, \cdots, x_n$ 是 V 中的线性无关的向量组，则 V 中的任一向量都可唯一地表作它们的线性组合。

练习 0-1-10 试证：设 V 是 n 维 $F-$线性空间，$x_1, x_2, \cdots x_m (m < n)$ 是 V 中的线性无关的向量组，则总可以找到 $n-m$ 个 V 中的向量 $x_{m+1}, \cdots, x_n$，使 $x_1, \cdots, x_n$ 构成 V 的一个基。

Ⅲ. 子线性空间

定义 设 V 是数域 F 上的线性空间，$S \subset V$ 是 V 的非空子集。若 S 中向量的和仍在 S 中，F 中任意数与 S 中向量的乘积仍在 S 中，则 S 在向量加法和向量的 $F-$数乘下是一个数域 F 上的线性空间。称为 V 的子线性空间。

显然，任一线性空间 V 本身是 V 的子空间；由零向量构成的独点集 $\{0\}$ 也是 V 的子空间。称它们为 V 的平庸子空间。

定义 设 $W_1, \cdots, W_s$ 是 $F-$线性空间 V 的子空间，则全体可表作各 W_i 中向量和的向量的集合 $\{v \in V \mid \exists\, w_i \in W_i (i = 1, \cdots, s)$ 使 $v = w_1 + \cdots + w_s\}$ 是 V 的子线性空间，称之为 $W_1, \cdots, W_s$ 的和空间，记作 $\sum_{i=1}^{s} W_i (\equiv W_1 + \cdots + W_s)$。显然子线性空间的求和满足交换律与结合律。

设 $W_1, \cdots, W_s$ 是 $F-$线性空间 V 的子空间，若下面的蕴涵命题成立，$w_i \in W_i, \forall\, 1 \leq i \leq s \;\&\; w_1 + \cdots + w_s = 0 \Rightarrow w_i = 0, \forall\, 1 \leq i \leq s$，则和空间 $W_1 + , \cdots, + W_s \left(\sum_{i=1}^{s} W_i \right)$ 中每一向量 x 的分解 $x = y_1 + \cdots + y_s, (y_i \in W_i)$ 是唯一的。这时和空间 $W_1 + \cdots + W_s \left(\sum_{i=1}^{s} W_i \right)$ 称为直和空间，并记作 $\bigoplus_{i=1}^{s} W_i (\equiv W_i \oplus \cdots \oplus W_s)$。显然有 $\dim \bigoplus_{i=1}^{s} W_i$

$= \sum_{i=1}^{s} \dim W_i$。如果 $V = W_1 \oplus W_2$，则称 W_1, W_2 互为补空间 。

练习 0-1-11 试证：F – 线性空间 V 的任意多个子空间的交仍是 V 的子线性空间；而 F – 线性空间 V 的两个子空间的并却不一定是 V 的子线性空间。

练习 0-1-12 试证：维数公式 $\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$。

练习 0-1-13 设 W 是 F – 线性空间 V 的子空间，试问在什么情况下它有唯一的补空间？

§0.2 内积空间、希尔伯特空间

Ⅰ. 内积空间

在上一节我们复习了线性空间的概念，可以看到在线性空间中只有两种运算：向量加法和数乘（称它们为线性运算）。我们还没有权利谈论向量的长度、两个向量是否正交，等等。这些是内积空间才有的度量性质。现在让我们给出内积空间的定义。

定义 设 V 是 F – 线性空间，若定义了一个映射 $f: V \times V \to F$，对所有的 $x, y, z, \in V$ 和 $\mu, v \in F$，称 $f(x, y)$ 为向量 x 与 y 的内积，记作 $\langle x | y \rangle$，满足下面的条件：

(1) $\langle x | \mu y + vz \rangle = \mu \langle x | y \rangle + v \langle x | z \rangle$ (2) $\langle x | y \rangle = \overline{\langle y | x \rangle}$（$\overline{\alpha}$ 表示 α 的共轭复数）

(3) $\langle x | x \rangle \geq 0$，当且仅当 $x = 0$ 时等号成立 (0.2.1)

则称 V 为 F – 内积空间。

称 $\sqrt{\langle x | x \rangle} \equiv \| x \|$ 为向量 x 的长度（范数，*norm*）。当 $\langle x | y \rangle = 0$ 时，称向量 x 与 y 正交。我们看到，内积空间首先是一个 F – 线性空间，因而 F – 线性空间的性质它都有。内积空间的维数即它作为线性空间的维数，内积空间的基即它作为线性空间的基。

设 $(e_1, \cdots, e_n)$ 是 F – 内积空间 V 的一个基，则以 $\langle e_j | e_k \rangle \equiv g_{jk}$ 为 j 行 k 列元素的矩阵 G 是可逆矩阵。因为，不然的话，齐次线性代数方程组 $\sum_{k=1}^{n} \langle e_j | e_k \rangle \xi^k = 0, \forall 1 \leq j \leq n$，有非零解。非零向量 $\sum_{k=1}^{n} e_k \xi^k$ 与所有基矢 $e_j (1 \leq j \leq n)$ 正交，从而与自己正交。

这与内积的正定性矛盾。

利用这一结果不难证明，对 F – 内积空间 V 的任意一个基 $\{e_1,\cdots,e_n\}$，存在 V 的唯一的一个基 $\{e^1,\cdots,e^n\}$，满足条件 $\langle e_\alpha \mid e^\beta\rangle = \delta_\alpha^\beta$。（留作练习）称作 $\{e_1,\cdots,e_n\}$ 的对偶基。显然当且仅当 $e_\alpha = e^\alpha$，$\forall\, 1\leq\alpha\leq n$ 时，$\{e_1,\cdots,e_n\}$ 是 V 的标准正交基（正交归一基）。

设 $x\in V$，$\{e_1,\cdots,e_n\}$ 是 V 的一个基。x 可以向 $\{e_1,\cdots,e_n\}$ 唯一地展开：$x = e_\alpha\xi^\alpha$。取对偶基中的基矢 e^β 与等式两边的内积得 $\langle e^\beta \mid x\rangle = \langle e^\beta \mid e_\alpha\rangle\xi^\alpha = \xi^\beta$。于是我们有

$$x = e_\alpha\langle e^\alpha \mid x\rangle \tag{0.2.2}$$

这对内积空间 V 的任一向量 x 和任一基 $\{e_1,\cdots,e_n\}$ 都成立。特别是，当 $\{e_1,\cdots,e_n\}$ 是 V 的标准正交基时，我们有

$$x = \sum_{\alpha=1}^{n} e_\alpha\langle e_\alpha \mid x\rangle \tag{0.2.3}$$

两个向量 $x = e_\alpha\xi^\alpha$ 和 $y = e_\beta\eta^\beta$ 的内积为 $\langle x \mid y\rangle = \bar{\xi}^\alpha\langle e_\alpha \mid e_\beta\rangle\eta^\beta = [x]^+[G][y]$，式中 $[G]$ 是以 $\langle e_\alpha \mid e_\beta\rangle$ 为 α 行 β 列元素的 $n\times n$ 矩阵。特别是，当 $\{e_1,\cdots,e_n\}$ 是 V 的标准正交基时，$[G]$ 是 $n\times n$ 单位矩阵，我们有 $\langle x \mid y\rangle = \sum_{\alpha=1}^{n}\bar{\xi}^\alpha\eta^\alpha = [x]^+[y]$。

设 W 是内积空间 V 作为 F – 线性空间的子线性空间，把 V 上内积 $f: V\times V\to F$ 的定义域限制到 $W\times W$，$f|_{W\times W}: W\times W\to F$ 可以当作 W 上的内积，使 W 成为一个内积空间，称为内积空间 V 的子空间。

设 W 是内积空间 V 的子空间，则 V 中与 W 的所有向量都正交的向量的集合 $\{x\in V \mid \langle x \mid w\rangle = 0,\ \forall\, w\in W\} \equiv W^\perp$，是 V 的子空间，称作 W 的正交补空间。不难证明 $V = W\oplus W^\perp$。

练习 0-2-1 试证：$|\langle x \mid y\rangle| \leq \|x\|\cdot\|y\|$，$\forall\, x,y\in V$。（Schwartz 不等式）

练习 0-2-2 试证：$\|x+y\| \leq \|x\| + \|y\|$，$\forall\, x,y\in V$。（三角不等式）

练习 0-2-3 试证：内积空间中由两两正交的非零向量构成的向量组线性无关。

练习 0-2-4 试证：设 $\{e_1,\cdots,e_n\}$ 是内积空间 V 的一个基，则存在 V 的唯一的一组向量 $\{e^1,\cdots,e^n\}$，满足条件 $\langle e_\alpha \mid e^\beta\rangle = \delta_\alpha^\beta$，而且这组向量是线性独立的。（对偶基的存在唯一性）

练习 0-2-5 试证：在 n 维内积空间中存在这样的基 $\{e_1,\cdots,e_n\}$ 它由两两正交

的单位向量构成：$\langle e_j | e_k \rangle = \delta_{jk}, \forall 1 \leq j,k \leq n$。（标准正交基的存在性）

练习 0-2-6 试证：设 V 是内积空间，$W_1, \cdots, W_s$ 是 V 的一组两两正交的子空间。则 $W_1, \cdots, W_s$ 的和空间是直和空间。

Ⅱ. 希尔伯特空间

向量的加法是二元运算。我们可以简单地把它推广到任意有限多个向量的求和。但是为了定义无穷多个向量的和(部分和序列的极限)，就必须利用向量的范数，先定义向量序列的极限。

让我们首先在此回顾一下数学分析中有关数列极限的一些知识。设 a 是复数(实数、有理数)，$\{x_1, x_2, \cdots\}$ 是复数(实数、有理数)序列。若对任给的正数 $\varepsilon > 0$，都有自然数 N，使得 $|x_n - a| < \varepsilon, \forall n > N$，则称数列序列 $\{x_1, x_2, \cdots\}$ 的极限为 a。一个数列最多有一个极限。称数列 $\{x_1, x_2, \cdots\}$ 为 Cauchy 序列，如果对任给的 $\varepsilon > 0$，都有自然数 N，使得 $|x_n - x_m| < \varepsilon, \forall n, m > N$。实数(复数)序列存在实数(复数)极限的充分必要条件为它是 Cauchy 序列。而有理数序列存在有理数极限的必要条件为它是 Cauchy 序列。但有理数 Cauchy 序列不一定有(有理数)极限。例如，有理数序列 $\left[\left(1+\frac{1}{1}\right)^1, \left(1+\frac{1}{2}\right)^2, \cdots, \left(1+\frac{1}{n}\right)^n, \cdots\right]$，把它看作实数序列，其极限是 e，不是有理数。因而该有理数序列在有理数域中没有极限。

称内积空间 V 中的向量序列 $\{x_1, x_2, \cdots\}$ 的极限为 a，并记作 $a = \lim\limits_{n\to\infty} x_n$，如果对任给的 $\varepsilon > 0$，都有自然数 N，使得 $\|x_n - a\| < \varepsilon, \forall n > N$。称向量序列 $\{x_1, x_2, \cdots\}$ 为 Cauchy 序列，如果对任给的 $\varepsilon > 0$，都有自然数 N，使得 $\|x_n - x_m\| < \varepsilon, \forall n, m > N$。如果内积空间 V 中任一 Cauchy 序列都有极限存在，则称 V 是完备的。称完备的内积空间为希尔伯特空间。

上面的讨论与基的选择无关。若选定了内积空间 V 的一个基，以 $[x]$ 记向量 x 在这个基下的矩阵，则易知 $n \times 1$ 复矩阵序列 $\{[x_1], [x_2], \cdots\}$ 以 $[a]$ 为极限，当且仅当向量序列 $\{x_1, x_2, \cdots\}$ 的极限为 a。

因而，在希尔伯特空间中，我们可以谈论任一 Cauchy 序列的极限。而在一个非完备的内积空间，谈论 Cauchy 序列的极限可能没有意义。更不能在一个没有定义范数的线性空间中，谈论向量序列的极限。相反，关于一般线性空间和内积空间的结

论,对希尔伯特空间都成立。

在量子力学中,体系的态矢量空间是希尔伯特空间,而且常常是无穷维的希尔伯特空间。读者要注意,对任意有限维线性空间成立的命题,对无限维线性空间不一定成立。

§0.3 线性算符

Ⅰ. 基本概念

定义 设 V, W 是两个 F – 线性空间,如果映射 $L: V \to W$ 满足下面条件:

$$L(\mu x + \nu y) = \mu L(x) + \nu L(y), \forall x, y \in V; \mu, \nu \in F \tag{0.3.1}$$

则称 L 是线性映射。

从上面定义可知,线性映射 L 是保持线性运算的映射,因而是 F – 线性空间 V 到 W 的同态。有时,也采用记号 $Lx \equiv L(x)$。当 $W = F$ 时,称 L 是线性函数。当 $W = V$ 时,L 也叫做 V 上的线性变换或线性算符。显然 V 的零向量在线性映射 L 下的像是 W 的零向量:$L(0) = 0$。W 的零向量在线性映射 L 下的完全原像 $L^{-1}(0) \equiv KerL$ 是 V 的线性子空间,叫做线性映射 L 的核。V 在线性映射 L 下的像 $L(V) \equiv \mathrm{Im}L$ 是 W 的线性子空间。不难证明:

$$\dim V = \dim L^{-1}(0) + \dim L(V) \tag{0.3.2}$$

两个线性算符的和,线性算符与数的乘积,两个线性算符的乘积分别由下列各式定义:

(1)

$$(A + B)x = Ax + Bx, \forall x \in V; \tag{0.3.3}$$

(2)

$$(\lambda A)x = \lambda(Ax), \forall \lambda \in F, x \in V; \tag{0.3.4}$$

(3)

$$(AB)x = A(Bx), \forall x \in V。 \tag{0.3.5}$$

F – 线性空间 V 上的线性算符的全体,记作 $EndV$,在算符加法(0.3.3)与数乘(0.3.

4)下是一个线性空间。而算符乘法是一个双线性运算,亦即对所有的线性算符 A, $B, C \in EndV$ 都有 $(\alpha A + \beta B)C = \alpha(AC) + \beta(BC)$, $C(\alpha A + \beta B) = \alpha(CA) + \beta(CB)$, $\forall \alpha, \beta \in F$。

在代数学中,在一个 F – 线性空间上,如果定义了向量的一个双线性的乘法运算,则称此线性空间为一个 F – 代数。F – 代数的维数是它作为线性空间的维数。当乘法满足结合律时,称之为结合的 F – 代数。若乘法还有单位元,称之为有 1 的结合的 F – 代数。n 维 F – 线性空间 V 上的线性算符的全体 $EndV$,是一个有 1 的结合的 n^2 维 F – 代数。全体 $n \times n$ 复矩阵的集合 $M_n(C)$,也是一个有 1 的结合的 n^2 维 F – 代数。

练习 0-3-1 试证:$\dim V = \dim L^{-1}(0) + \dim L(V)$,其中 L 是 F – 线性空间 V 到 W 的同态。

Ⅱ. 线性算符的矩阵

让我们考察下面的事实。设$\{a_1 \quad a_2 \quad \cdots \quad a_m\}$是线性空间 V 的一个向量组(不一定是 V 的基),A, B 是 V 上线性算符。若 $Aa_\alpha = \sum_{\beta=1}^{m} \alpha_\beta A_{\beta\alpha}$, $Ba_\alpha = \sum_{\beta=1}^{m} \alpha_\beta B_{\beta\alpha}$, $\forall \alpha = 1,2,\cdots,m$ 或写作矩阵形式 $[Aa_1 \quad Aa_2 \quad \cdots \quad Aa_m] = [a_1 \quad a_2 \quad \cdots \quad a_m][A]$, $[Ba_1 \quad Ba_2 \quad \cdots \quad Ba_m] = [a_1 \quad a_2 \quad \cdots \quad a_m][B]$,其中$[A]([B])$是以 $A_{\beta\alpha}(B_{\beta\alpha})$ 为 β 行 α 列矩阵元的 $m \times m$ 矩阵,则有 $[ABa_1 \quad ABa_2 \quad \cdots \quad ABa_m] = [a_1 \quad a_2 \quad \cdots \quad a_m][A][B]$。一般地 $[f(A,B)a_1 \quad f(A,B)a_2 \quad \cdots \quad f(A,B)a_m] = [a_1 \quad a_2 \quad \cdots \quad a_m]f([A],[B])$,其中 f 是任一解析函数。上式的分量表达式为 $f(A,B)a_\alpha = \sum_{\beta=1}^{m} \alpha_\beta f([A],[B])_{\beta\alpha}$。

设 A 是 F – 线性空间 V 上的线性算符。当选定 V 的一个基$\{e_1, e_2, \cdots, e_n\}$,我们考察向量 $x = \sum_{j=1}^{n} e_j \xi^j$ 的矩阵 $[x] = \begin{bmatrix} \xi^1 \\ \vdots \\ \xi^n \end{bmatrix}$ 与向量 $Ax \equiv y = \sum_{k=1}^{n} e_k \eta^k$ 的矩阵 $[y] \equiv [Ax] = \begin{bmatrix} \eta^1 \\ \vdots \\ \eta^n \end{bmatrix}$ 间的关系。以线性算符 A 作用于 $x = \sum_{j=1}^{n} e_j \xi^j$ 的两边得

$$Ax = \sum_{j=1}^{n} (Ae_j)\xi^j \tag{0.3.6}$$

由于 Ae_j 仍是 V 中向量，总可唯一地向基矢展开

$$Ae_j = \sum_{k=1}^{n} e_k A_j^k \tag{0.3.7}$$

代入式(0.3.6)得 $\sum_{k=1}^{n} e_k \eta^k = \sum_{k=1}^{n} e_k \sum_{j=1}^{n} A_j^k \xi^j$，于是我们有

$$\eta^k = \sum_{j=1}^{n} A_j^k \xi^j \equiv A_j^k \xi^j \text{(我们采用爱因斯坦求和约定)} \tag{0.3.8}$$

从以上的讨论看到，当选定 V 的一个基 $e_1, \cdots, e_n$，可以用一个由等式(0.3.7)确定的 $n \times n$ 矩阵 $[A]$(A_j^k 是其 k 行 j 列元素)来表达线性算符 A。向量 Ax 的矩阵 $[Ax]$ 简单地等于线性算符 A 的矩阵 $[A]$ 左乘向量 x 的矩阵 $[x]$：

$$[y] = [Ax] = [A][x] \tag{0.3.9}$$

由等式(0.3.7)可知恒等算符 E 的矩阵是 $n \times n$ 单位矩阵。不难验证(留作练习)

$$[A + B] = [A] + [B], [\lambda A] = \lambda [A], [AB] = [A][B] \tag{0.3.10}$$

可见，若 A 是可逆算符，则有 $[A^{-1}] = [A]^{-1}$。矩阵元为数域 F 中的数的 $n \times n$ 矩阵的全体是一个 n^2 维的有 1 的结合的 F - 代数，记作 $M_n(F)$。若 V 是 n 维 F - 线性空间，则 $EndV$ 也是一个 n^2 维的有 1 的结合的 F - 代数。选定 V 的一个基 $e_1, \cdots, e_n$，V 上的每一个线性算符 A，都可以按等式(0.3.7)确定一个 $n \times n$ 的 F - 矩阵 $[A]$；反之每一个 $n \times n$ 的 F - 矩阵 $[A] \in M_n(F)$，都可以按等式(0.3.7)确定 V 上的一个线性算符 A。由(0.3.10)可知，对每一 $A \in EndV$，令 $A \mapsto [A]$，这样确定的映射，是 F - 代数 $EndV$ 到 $M_n(F)$ 的同构。因而，当问题仅涉及线性算符的线性运算和算符乘法时，我们完全可以代之以讨论相应的矩阵线性运算和矩阵乘法。这样做有时(特别是进行具体计算时)会带来很大方便。

显然 A 和 x 的矩阵都与基有关。那么，当基变换时，向量与线性算符的矩阵如何变换呢？设 $\{e_1', \cdots, e_n'\}$ 是 V 的另一个基，它与基 $\{e_1, \cdots, e_n\}$ 间的变换关系是

$$e_j' = e_k T_j^k, [e_1' \quad \cdots \quad e_n'] = [e_1 \quad \cdots \quad e_n] \begin{bmatrix} T_1^1 & \cdots & T_n^1 \\ \vdots & \vdots & \vdots \\ T_1^n & \cdots & T_n^n \end{bmatrix} \tag{0.3.11}$$

$$e_k = e'_j (T^{-1})^j_k, [e_1 \quad \cdots \quad e_n] = [e'_1 \quad \cdots \quad e'_n] \begin{bmatrix} (T^{-1})^1_1 & \cdots & (T^{-1})^1_n \\ \vdots & \vdots & \vdots \\ (T^{-1})^n_1 & \cdots & (T^{-1})^n_n \end{bmatrix}$$

式中$[T]$是 $n \times n$ 非奇异矩阵。若向量 $x = e_j \xi^j = e'_k \xi'^k$，不难看到

$$\begin{bmatrix} \xi'^1 \\ \vdots \\ \xi'^n \end{bmatrix} = \begin{bmatrix} (T^{-1})^1_1 & \cdots & (T^{-1})^1_n \\ \vdots & \vdots & \vdots \\ (T^{-1})^n_1 & \cdots & (T^{-1})^n_n \end{bmatrix} \begin{bmatrix} \xi^1 \\ \vdots \\ \xi^n \end{bmatrix}, [X]' = [T^{-1}][X] \tag{0.3.12}$$

以及

$$[A]' = [T]^{-1}[A][T] \tag{0.3.13}$$

这表明当基经受变换(0.3.11)时，同一向量的矩阵要经受变换(0.3.12)，而同一线性算符的矩阵要经受变换(0.3.13)。后者称为相似变换。

以上讨论的是 F – 线性空间上的线性算符，所有结论对内积空间(它是一类特殊的线性空间)上的线性算符当然成立。下面讨论内积空间 V 上的线性算符所特有的一些性质。

选定 V 的一个基 $e_1, \cdots e_n$(不一定是标准正交的)，以 $\langle e_j | e_k \rangle \equiv g_{kj}$ 为 j 行 k 列元素的 $n \times n$ 矩阵$[G]$是非奇异的，以 g^{ij} 记$[G]^{-1}$的 i 行 j 列元素。V 上的线性算符 A 的矩阵的定义式是(0.3.7)式。取 e_i 与(0.3.7)式两端的内积得

$$\langle e_i \mid A \mid e_j \rangle = \langle e_i \mid e_k \rangle A^k_j = g_{ik} A^k_j$$

从而得出

$$g^{li} \langle e_i \mid A \mid e_j \rangle = \delta^l_k A^k_j = A^l_j \tag{0.3.14}$$

当基 $e_1, \cdots, e_n$ 是标准正交基时，我们得到

$$A^l_j = \delta^{li} \langle e_i \mid A \mid e_j \rangle = \langle e_l \mid A \mid e_j \rangle \tag{0.3.15}$$

设 V 是 n 维内积空间，A 是 V 上的线性算符。则存在唯一的 V 上的线性算符 B，使得

$$\langle x \mid A \mid y \rangle = \langle Bx \mid y \rangle, \forall x, y \in V \tag{0.3.16}$$

事实上，任选 V 的一个标准正交基$\{e_1, \cdots, e_n\}$。则条件(0.3.16)等价于

$$\langle e_i \mid A \mid e_j \rangle = \langle Be_i \mid e_j \rangle, \forall 1 \leq i, j \leq n \tag{0.3.17}$$

我们定义 V 上线性算符 B，使 $Be_i = \sum_{j=1}^{n} e_j \langle Ae_j \mid e_i \rangle$。直接验证可知，这样定义的 $B \in EndV$ 满足条件(0.3.17)。至此我们用构造法证明了算符 B 的存在性。算符 B 的唯一性是显然的。我们称 B 为 A 的厄米共轭，记作 $B = A^+$。于是

$$\langle x \mid A \mid y \rangle = \langle A^+ x \mid y \rangle, \forall x, y \in V \tag{0.3.18}$$

可见，在没有定义内积的线性空间，讲线性算符的厄米共轭是没有意义的。显然有 $(A^+)^+ = A$。如果 $A = A^+$，则称 A 为厄米算符(自轭算符)。如果 $A = -A^+$，则称之为反厄米算符。若 $\langle Ax \mid Ay \rangle = \langle x \mid y \rangle$，$\forall x, y \in V$，则称 A 是酉算符(幺正算符)。A 是酉算符的充要条件是 $A^+A = E$(恒等算符)，即 A^+ 是 A 的左逆。当 V 是有限维内积空间时，这等价于 $A^{-1} = A^+$，即 A^+ 是 A 的逆。

若 $AA^+ = A^+A$，则称 A 是正规(proper)算符。由于 $(A^+)^+ = A$，这时 A^+ 也是正规(proper)算符。显然厄米算符、反厄米算符、酉算符都是正规算符。

矩阵[A]的厄米共轭[A]$^+$是[A]的转置复共轭。如果[A]$^+$ = [A]，则称[A]为厄米矩阵。如果[A]$^+$ = [A]$^{-1}$，则称[A]为幺正矩阵。在标准正交基下，我们有[A^+] = [A]$^+$。特别是，在标准正交基下，厄米(幺正)算符的矩阵是厄米(幺正)矩阵。(留作练习)

练习 0-3-2 试证：设 $A, B \in EndV$。当 V 是有限维的，若 $AB = E$，则 $BA = E$；当 V 是无限维的，若 $AB = E$，则 $BA = E$ 不一定成立(举出反例)。

练习 0-3-3 试证：设 $A \in EndV$，V 是内积空间。则在标准正交基下，我们有[A^+] = [A]$^+$。(因而在标准正交基下，厄米算符的矩阵是厄米矩阵，幺正算符的矩阵是幺正矩阵)

Ⅲ. 特征子空间

设 A 是 n 维 F - 线性空间 V 上的线性算符，$\lambda \in F$。若齐次线性方程

$$Ax = \lambda x \tag{0.3.19}$$

有非零解，则称 λ 是 A 的特征值，非零解称作 A 的属于特征值 λ 的特征向量。全体解的集合构成线性空间 V 的一个子空间，称为线性算符 A 的属于特征值 λ 的特征子空间，记作 V_λ^A。$\dim V_\lambda^A$ 称作特征值 λ 的简并度。方程(0.3.18)称为线性算符 A 的特征方程。当选定线性空间 V 的一个基时它可以写成矩阵形式[A][x] = λ[x]，此即

$(\lambda[E]-[A])[x]=0$,其中$[E]$是 $n\times n$ 单位矩阵。这一齐次线性方程有非零解的充分必要条件是其系数行列式为零$|\lambda[E]-[A]|=0$。这是一个关于未知数 λ 的 n 次代数方程,称为 A 的特征多项式方程(久期方程)。它的解就是 A 的特征值。当 F 是复数域时,由代数基本定理,A 至少有一个本征值。当 F 是实数域或有理数域时,A 可能没有本征值(请举出例子)。

定理 1 线性算符 A 的属于不同特征值的特征向量线性无关。

证明 设 $\alpha^1x_1+\cdots\alpha^jx_j+\cdots+\alpha^sx_s=0$,其中 $Ax_j=\lambda_jx_j$, $x_j\neq0$, $\forall\,1\leq j\leq s$; $\lambda_j\neq\lambda_k$, $\forall\,j\neq k$。l 是 1 与 s 间的一个任意给定的数。以 $\prod\limits_{i(\neq l)}(A-\lambda_i)$ 作用于上式得 $\alpha^l\prod\limits_{i(\neq l)}(\lambda_l-\lambda_i)x_l=0$。从而 $\alpha^l=0$。

由式(0.3.7)可知,当且仅当每个基矢都是线性算符 A 特征向量时,

$$Ae_j=\lambda_je_j=\lambda_j\sum_{k=1}^{n}\delta_j^ke_k(\text{对 }j\text{ 不求和})$$

A 的矩阵是对角的,对角元是相应的特征值

$$[A]=\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}$$

由上面的定理可知,若 n 维线性空间 V 上的线性算符 A 有 n 个不同的特征值时,它有 n 个线性无关的特征向量,选它们为基时,$[A]$是对角矩阵。

考察 2×2 矩阵 $\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$,它只有一个特征值 0,相应的特征子空间 $\left\{\begin{bmatrix}\alpha\\ 0\end{bmatrix}\middle|\,\alpha\in F\right\}$是一维的,即只有一个线性独立的特征向量。这表明它不能通过相似变换对角化。

一般而言,一个 $n\times n$ 复矩阵能通过相似变换对角化的充分必要条件是它的最小化零多项式没有重根。用线性算符的语言,上述命题可以表作,n 维复线性空间 V 上的线性算符 A 具有完备的特征向量组的充分必要条件是它的最小化零多项式没有重根。这种线性算符称为半单线性算符。读者可以从任何一本线性代数书中找到它的证明。限于篇幅,本书将不讨论这个定理。但我们将向读者介绍一个非常有用的

定理,而它的证明要简单得多。

定理 2 设 V 是 n 维复内积空间,A 是 V 上的线性算符。则 A 的特征子空间两两正交,且张成 V(A 具有完备的正交的特征向量组)的充分必要条件是 $A^+A = AA^+$,即 A 是正规算符。

为证明这个定理我们将介绍线性算符的不变子空间的概念和几个引理。

设 A 是 F – 线性空间 V 上的线性算符,W 是 V 的子空间。若 $Aw \in W$, $\forall w \in W$,则称 W 是 A – 不变子空间(或者 A 的不变子空间)。V 和零向量空间 $\{0\}$ 显然是 A – 不变子空间,称之为平庸的不变子空间。A 的特征子空间当然是 A – 不变子空间。

当 W 是 A – 不变子空间,而我们只考虑 A 对 W 中的向量的作用时,可以把 A 看作 W 上的线性算符,记作 $A|_W$,称之为 A 到 W 上的限制。

引理 1 设 A、B 是 F – 线性空间 V 上的线性算符,若 $AB = BA$,则 A 的特征子空间是 B – 不变子空间。

事实上,若 $x \in V_\lambda^A$,则有 $A(BX) = B(AX) = B(\lambda x) = \lambda(Bx)$。这表明 $Bx \in V_\lambda^A$。

引理 2 设 A 是有限维复内积空间 V 上的线性算符,W 是 A – 不变子空间,$W^\perp$ 是其正交补空间,则 $W^\perp$ 是 A^+ – 不变子空间。

事实上,若 $y \in W^\perp$,则有 $\langle A^+ y | x\rangle = \langle y | A | x\rangle = 0$, $\forall x \in W$。这表明 $A^+ y \in W^\perp$。

现在我们来证明定理本身。先证明条件的必要性。若 $V_\lambda^A \perp V_\mu^A$, $\forall \lambda \neq \mu$,而且 $V = \sum_\lambda V_\lambda^A = \bigoplus_\lambda V_\lambda^A$(对 A 的所有特征值求和)。我们可以在每个特征子空间中选一个标准。

正交基,它们合在一起构成 V 的一个标准正交基。A 在这个基下的矩阵 $[A]$ 是对角的,而 $[A^+] = [A]^+$ 也是对角的。两个对角矩阵当然交换,它们代表的算符当然也交换,$A^+A = AA^+$。

现在证明充分性。对 $\dim V$ 作数学归纳。当 $\dim V = 1$,定理显然成立。今设定理对 $\dim V \leq n$ 的情况成立。我们讨论 $\dim V = n$ 的情况。如果 A 是恒等算符的纯量倍,定理显然成立。我们讨论 A 不是恒等算符纯量倍的情况。A 至少有一个特征值 λ。相应的特征子空间 $V_\lambda^A \neq \{0\}$。以 V' 记 V_λ^A 的正交补空间,$V = V_\lambda^A \oplus V'$,我们有 $\dim V' < n$。由引理 1,因 $A^+A = AA^+$,可知 V_λ^A 是 A^+ 的不变子空间。由引理 2,V' 是 A

和 A^+ 的不变子空间。不难看出，A 到不变子空间 V' 上的限制 $A|_{V'} \equiv A'$ 仍是正规算符，由归纳假定，$V' = \bigoplus_{\lambda'} V'^{A'}_{\lambda'}$；$V'^{A'}_{\lambda'} \perp V'^{A'}_{\mu'}$，$\forall \lambda' \neq \mu'$。显然所有 λ' 不等于 λ，而且 $V'^{A'}_{\lambda'} = V^{A}_{\lambda'}$，$V^{A}_{\lambda} \perp V^{A}_{\lambda'}$。于是有 $V = V^{A}_{\lambda} \oplus V' = V^{A}_{\lambda} \oplus (\bigoplus_{\lambda'} V^{A}_{\lambda'})$，$V$ 是 A 的两两正交的特征子空间的直和，定理得证。

于是，正规线性算符是半单的。

练习 0-3-4 试证：相似矩阵有相同的特征多项式方程（线性算符的特征多项式方程与基的选择无关）。

练习 0-3-5 试证：上三角矩阵的特征值就是它的对角矩阵元。

Ⅳ. 不可约线性算符组

设 $\{A_\sigma \in EndV | \sigma \in S\}$ 是 F – 线性空间 V 上的线性算符组，式中 S 是下标的集合。若只有 V 和零空间是所有 A_σ 的不变子空间，则称 $\{A_\sigma \in EndV | \sigma \in S\}$ 是不可约算符组。

Schur 引理 1 设 V 是复线性空间，$\{A_\sigma \in EndV | \sigma \in S\}$ 是不可约算符组，$B \in EndV$。若 $A_\sigma B = BA_\sigma$，$\forall \sigma \in S$，则 B 是恒等算符 E 的纯量倍。

事实上，由于 V 是复线性空间，B 至少有一个特征子空间 $V^B_\lambda \neq \{0\}$。由上一节的引理 1，V^B_λ 是所有 A_σ 的不变子空间，只能 $V^B_\lambda = V$。于是 $B = \lambda E$。

Schur 引理 2 设 $\{A_\sigma \in EndV | \sigma \in S\}$ 是 F – 线性空间 V 上的不可约线性算符组，$\{B_\sigma \in EndW | \sigma \in S\}$ 是 F – 线性空间 W 上的不可约线性算符组，$L: V \to W$ 是线性映射。若 $LA_\sigma = B_\sigma L$，$\forall \sigma \in S$，则 $L = 0$ 或者 L 是双射（L 是 F – 线性空间的同构）。

事实上，$L(V)$ 是所有 B_σ 的不变子空间。（$\because B_\sigma(Lv) = L(A_\sigma v) \in L(V)$，$\forall v \in V$，$\sigma \in S$）因而 $L(V) = \{0\}$ 或者 $L(V) = W$。$L^{-1}(0)$ 是所有 A_σ 的不变子空间。（$\because L(A_\sigma x) = B_\sigma(Lx) = B_\sigma 0 = 0$，$\forall x \in L^{-1}(0)$，$\sigma \in S$）因而 $L^{-1}(0) = V$ 或者 $L^{-1}(0) = \{0\}$。由此得到四种逻辑上可能的组合：

(1) $L(V) = \{0\}$ & $L^{-1}(0) = V$，此时 $L = 0$；

(2) $L(V) = \{0\}$ & $L^{-1}(0) = \{0\}$，此时 $V = \{0\}$，$L = 0$；

(3) $L(V) = W$ & $L^{-1}(0) = V$，此时 $W = \{0\}$，$L = 0$；

(4) $L(V) = W$ & $L^{-1}(0) = \{0\}$，此时 L 是双射。

可见,只能有 $L=0$,或者 L 是 F – 线性空间的同构。定理得证。

考虑 F – 线性空间 V 上的线性算符组$\{A_1, A_2, \cdots\}$。**W 是每一个A_σ 的不变子空间。若$\{A_1|_W, A_2|_W, \cdots\}$是不可约线性算符组,则称 W 是$\{A_1, A_2, \cdots\}$的不可约不变子空间。**

由下面的练习,可看出不可约算符组的意义。

练习 0-3-6 试证:设 V 是复线性空间,$\{A_\sigma \in EndV | \sigma \in S\}$是不可约算符组。若在基$\{e'_1, \cdots, e'_n\}$ 与基$\{e_1, \cdots, e_n\}$下,所有 A_σ 的矩阵相同:$[A_\sigma]' = [A_\sigma]$, $\forall \sigma \in S$。则 $e'_j = \lambda e_j, 0 \neq \lambda \in F$, $\forall 1 \leq j \leq n$。(这表明不可约算符组的矩阵可以把基确定到只差一个非零纯量倍。)

练习 0-3-7 试证:设 V 是内积空间。若$\{A_\sigma \in EdnV | \sigma \in S\}$是不可约算符组,则$\{A_\sigma^+ \in EdnV | \sigma \in S\}$也是不可约算符组。

练习 0-3-8 试证:设 V 是内积空间,$\{A_\sigma \in EdnV | \sigma \in S\}$是不可约算符组。若在基$\{e_1, \cdots, e_n\}$下,$[A_\sigma]^+ = [A_\sigma^+]$, $\forall \sigma \in S$,则$\langle e_j | e_k \rangle = \lambda \sigma_{jk}$, $(\lambda > 0)$。($\{e_1, \cdots, e_n\}$是正交等长基)

练习 0-3-9 试证:设 V 是线性空间。若$\{A_\sigma \in EdnV | \sigma \in S\}$是 $EndV$ 的生成组(即 V 上的任一线性算符都可以表作若干A_σ 的乘积的线性组合),则$\{A_\sigma \in EdnV | \sigma \in S\}$是不可约的。

Ⅴ. 两两对易的线性算符组

考虑 F – 线性空间 V 上的线性算符组$\{A_1, \cdots, A_S\}$。若存在 F 中的数组$\{\lambda_1, \cdots, \lambda_S\}$使齐次线性方程组 $A_\sigma x = \lambda_\sigma x$, $\forall 1 \leq \sigma \leq S$ 有非零解。则其解空间$\{x \in V | A_\sigma x = \lambda_\sigma x, \forall 1 \leq \sigma \leq S\}$称为该线性算符组的属于特征值$\{\lambda_1, \cdots, \lambda_S\}$的共同的特征子空间,记作 $V_{\lambda_1 \cdots \lambda_S}^{A_1 \cdots A_S}$,其中的非零解称为该线性算符组的属于特征值$\{\lambda_1 \cdots \lambda_S\}$的共同的特征向量。不难证明 $V_{\lambda_1}^{A_1} \cap \cdots \cap V_{\lambda_S}^{A_S} = V_{\lambda_1 \cdots \lambda_S}^{A_1 \cdots A_S}$。(留作练习)

上一小节我们说过,内积空间上的线性算符有完备的正交的特征向量组,当且仅当它是正规线性算符。现在我们来讨论一组正规线性算符有共同的完备的正交的特征向量组的条件。

定理 3 F – 内积空间 V 上的正规线性算符组$\{A, A_1, \cdots, A_S\}$有完备的共同的特

征向量组(它们的共同的特征子空间张成 V)的充要条件是组中线性算符两两对易。

证明 先证明充分性。对 $\dim V$ 作数学归纳。定理当 $\dim V=1$ 时显然成立。今假定当 $\dim V<n$ 时定理成立,我们讨论 $\dim V=n$ 的情况。当所有的 A_σ 都是恒等算符的纯量倍时,定理显然成立。不然,半单线性算符组中至少有一个算符,不妨设为 A,不是恒等算符的纯量倍。V 是 A 的两两正交的特征子空间的直和:$V=\bigoplus_\lambda V_\lambda^A$。每个 V_λ^A 都是所有 A_σ 的不变子空间。所有 A_σ 到 V_λ^A 上的限制 $A_\sigma|_{V_\lambda^A}\equiv A_\sigma^\lambda$ 是 V_λ^A 上的一组两两对易的正规线性算符。由归纳假定,和 $\dim V_\lambda^A<n$,可知 $V_\lambda^A=\bigoplus_{\lambda_1\cdots\lambda_S}(V_\lambda^A)_{\lambda_1\cdots\lambda_S}^{\lambda\cdots\lambda}$。而 $(V_\lambda^A)_{\lambda_1\cdots\lambda_S}^{\lambda\cdots\lambda}=V_{\lambda\lambda_1\cdots\lambda_S}^{AA_1\cdots A_S}$。于是 $V=\bigoplus_{\lambda\lambda_1\cdots\lambda_S}V_{\lambda\lambda_1\cdots\lambda_S}^{AA_1\cdots A_S}$。

现在证明必要性。若 $V=\bigoplus_{\lambda\lambda_1\cdots\lambda_S}V_{\lambda\lambda_1\cdots\lambda_S}^{AA_1\cdots A_S}$。。选定每个 $V_{\lambda\lambda_1\cdots\lambda_S}^{AA_1\cdots A_S}$。的一个基。这些基的并构成 V 的一个基。在这个基下所有 A_σ 的矩阵$[A_\sigma]$都是对角的,因而两两对易。它们

所表示的算符也是两两对易的。定理证完。

特别是,一组正规算符的共同的特征子空间张成整个内积空间的充分必要条件是它们两两对易。在这种情况下共同的特征子空间两两正交。

在量子力学中,一组可同时测量的力学量的算符,是一组两两对易的厄米算符。

当一组两两对易的厄米算符的共同的特征子空间全是一维子空间时,对应的力学量组是一完全力学量组。通过完全力学量组的同时测量能制备出处于纯态的量子体系。这个刚制备出来的量子体系所处的态是该完全力学量组的共同的本征态。一般而言,非完全力学量组的同时测量制备出处于混态的量子体系。

练习 0-3-10 试证:$V_{\lambda_0}^{A_0}\cap\cdots\cap V_{\lambda_S}^{A_S}=V_{\lambda_0\cdots\lambda_S}^{A_0\cdots A_S}$。

Ⅵ. 投影算符

定义 设 V 是 F – 线性空间,P 是 V 上的线性算符。若 $P^2=P$,则称 P 是 V 上的投影算符。

零算符和恒等算符都是投影算符的平庸的例子。考虑某个平面上由固定点 o 引出的所有向量的集合,这是一个二维实线性空间,记作 R^2。任选两个线性独立的向量$\{e_1,e_2\}$为基(不一定是正交归一基)。对任一向量 $x=e_1\xi^1+e_2\xi^2\in V$,令 $Px=$

$e_1\xi^1$。则易知 $P\in EndV, P^2=P$。从而 P 是 R^2 上的投影算符。

由这个定义可知,投影算符的特征值不外乎 1 和 0。事实上,以 P 作用于其特征方程 $Px=\lambda x$ 得 $P(Px)=\lambda(Px)=\lambda^2 x$。由于 $P^2=P$,我们有 $\lambda(\lambda-1)x=0$。又由于 $x\neq 0$,我们得 $\lambda(\lambda-1)=0$。

投影算符 P 的象 $\mathrm{Im}P=P(V)$中的向量在 P 的作用下不变。投影算符 P 的核 $KerP=P^{-1}(0)$中的向量在 P 的作用下变作 O。显然有 $\mathrm{Im}P=V_1^P, KerP=V_0^P$。

$E=(E-P)+P$。易知 $E-P$ 也是投影算符:$(E-P)^2=E-P$,投影算符 P 与 $(E-P)$对易,且 $P(E-P)=(E-P)P=0$。对每一 $x\in V, x=Ex=(E-P)x+Px$。这表明 $V=(E-P)(V)+P(V)=\mathrm{Im}(E-P)+\mathrm{Im}P$。又由 $P(E-P)=(E-P)P=0$,可知 $V=(E-P)(V)\oplus P(V)=\mathrm{Im}(E-P)\oplus\mathrm{Im}P=KerP\oplus\mathrm{Im}P$。以上的讨论表明,给了 F – 线性空间 V 上的一个投影算符 P,则 V 可以分解成 P 的核与 P 的像的直和。可见,投影算符的特征向量张成整个空间。

反之,给了 V 的一个直和分解:$V=W\oplus W'$,存在唯一的投影算符 P,使 $W=\mathrm{Im}P, W'=KerP$。

由于当 W 不是平庸的子空间时,其补空间不是唯一的,故仅由 $W=\mathrm{Im}P$,一般不能唯一地确定投影算符 P。换句话说,"P 是把 V 投到子空间 W 的投影算符"并不能唯一地界定 P。

以上讨论了 F – 线性空间上的投影算符。现在让我们转而讨论内积空间上的投影算符。

定理 4 设 V 是内积空间,P 是 V 上的投影算符。则 $P=P^+\Leftrightarrow\mathrm{Im}P=(KerP)^\perp$。(正投影算符就是厄米投影算符)

证明 充分性。已经证明了 $V=\mathrm{Im}P\oplus KerV$,现设 $P=P^+$,我们来证明 $\mathrm{Im}P\perp KerP$。设 $x\in\mathrm{Im}P, y\in KerP$,于是$\langle x|y\rangle=\langle Px|y\rangle=\langle x|P^+y\rangle=\langle x|Py\rangle=\langle x|0\rangle=0$。

必要性。设 $\mathrm{Im}P\perp KerP$。每一 $v\in V$ 都可唯一地表作 $v=x+y$,其中 $x\in\mathrm{Im}P, y\in KerP$。$v, v'$是任意两个向量。应有$\langle v|Pv'\rangle=\langle P^+v|v'\rangle$。左端等于$\langle x+y|P(x'+y')\rangle=\langle x+y|x'\rangle=\langle x|x'\rangle=\langle x|x'+y'\rangle=\langle Px|v'\rangle$,因而 $P=P^+$。

练习 0-3-11 试证:设 P, Q 是 F – 线性空间上的投影算符。则下面各命题等价:

(1) $(P+Q)^2 = P+Q$;

(2) $PQ = QP = 0$;(亦即 $\mathrm{Im}Q \subset KerP \& \mathrm{Im}Q \subset KerP$)

(3) $PQ = -QP$。

* §0.4 线性算符的指数函数

Ⅰ. 内积空间上线性算符的范数与算符序列的极限

为了能定义算符序列的极限,我们引入线性算符的范数的概念。设 V 是 n 维 F - 内积空间。对 $x \in V$,我们曾定义向量 x 的范数为 $\|x\| = \sqrt{\langle x|x\rangle}$。对 V 上线性算符 $A \in EndV$,定义线性算符的范数为

$$\|A\| = \max_{x\in V, \|x\|=1} \{\|Ax\|\} \tag{0.4.1}$$

这个定义与内积空间的基的选择无关。由此定义可知,对任一数(纯量)$\alpha \in F$,都有 $\|\alpha x\| = |\alpha| \cdot \|x\|$, $\|\alpha A\| = |\alpha| \cdot \|A\|$。对每一 $x \in V$,易见

$$\|Ax\| \leq \|A\| \cdot \|x\| \tag{0.4.2}$$

事实上,若 $x = 0$,则上式显然成立。若 $x \neq 0$,则 $\frac{x}{\|x\|}$ 是单位向量。从而 $\|A\| \geq \left\|A\frac{x}{\|x\|}\right\| = \frac{\|Ax\|}{\|x\|}$,上式也成立。

命题 1 设 V 是 n 维内积空间,$A, B \in EndV$,则 $\|AB\| \leq \|A\| \cdot \|B\|$。

事实上,对所有的使 $\|v\| = 1$ 的 $v \in V$,都有 $\|ABv\| \leq \|A\| \cdot \|Bv\| \leq \|A\| \cdot \|B\| \cdot \|v\| = \|A\| \cdot \|B\|$。

推论 $\|A^m\| \leq \|A\|^m$, $\forall\, m \geq 1$。

命题 2 设 V 是 n 维内积空间,$A, B \in EndV$, $\alpha \geq 0$,则 $\|\alpha A\| = \alpha\|A\|$, $\|A+B\| \leq \|A\| + \|B\|$。

事实上,$\|(\alpha A)v\| = \|\alpha(Av)\| = \alpha\|Av\|$, $\forall \|v\| = 1$。故使 $\|(\alpha A)v\|$ 最大的 v,也使 $\|Av\|$ 最大。从而 $\|\alpha A\| = \alpha\|A\|$。又,$\|(A+B)v\| = \|Av + Bv\| \leq \|Av\| + \|Bv\| \leq \|A\| + \|B\|$, $\forall \|v\| = 1$。故有 $\|A+B\| \leq \|A\| + \|B\|$。

现在我们来定义线性算符序列的极限。设 $A_1, A_2, \cdots$ 是线性算符序列。若对任给的 $\varepsilon > 0$,都存在自然数 N,使 $\| A - A_m \| < \varepsilon$, $\forall\, m > N$,则称 A 是线性算符序列 $A_1, A_2, \cdots$ 的极限,记作 $A = \lim\limits_{m\to\infty} A_m$。

设 $A_1, A_2, \cdots$ 是线性算符序列。若对任给的 $\varepsilon > 0$,都存在自然数 N,使 $\| A_k - A_m \| < \varepsilon$, $\forall\, k, m > N$,,则称 $A_1, A_2, \cdots$ 是 Cauchy 序列。

让我们来定义 $n \times n$ 复矩阵 $[A]$ 的范数。把 $[A]$ 看作由所有 $n \times 1$ 复矩阵构成的内积空间(其中两个向量内积的定义为 $\langle [x] | [y] \rangle \equiv [x]^+ [y]$)上的线性算符,其范数称作 $[A]$ 的范数。于是

$$\| [A] \| = \max_{\sum\limits_{k=1}^{n} |v^k|^2 = 1} \sqrt{\sum_{j=1}^{n} \left| \sum_{k=1}^{n} A_k^j v^k \right|^2} \tag{0.4.3}$$

由上式可知,$[A]$ 的每个列向量的范数小于 $[A]$ 的范数 $\sqrt{\sum\limits_{j=1}^{n} \left| A_k^j \right|^2} \leq \| [A] \|$, $\forall\, 1 \leq k \leq n$。特别是,$[A]$ 的每个阵元的模小于 $[A]$ 的范数 $| A_k^j | \leq \| [A] \|$, $\forall\, 1 \leq j, k \leq n$。

不难证明,n 维内积空间 V 上的线性算符 A 的范数,等于它在任一标准正交基下的矩阵 $[A]$ 的范数。(留作练习)

上面已给出了线性算符序列极限的定义,把 $n \times n$ 复矩阵看成作用于所有 $n \times 1$ 复矩阵构成的内积空间上的线性算符,就得到 $n \times n$ 复矩阵序列的极限的定义:

设 $A_1, A_2, \cdots$ 是 $n \times n$ 复矩阵序列(省掉方括号),A 是 $n \times n$ 复矩阵。若对任给的 $\varepsilon > 0$,都存在自然数 N,使 $\| A - A_m \| < \varepsilon$, $\forall\, m > N$,则称 A 是 $n \times n$ 复矩阵序列 $A_1, A_2, \cdots$ 的极限,记作 $A = \lim\limits_{m\to\infty} A_m$。不难证明 $A = \lim\limits_{m\to\infty} A_m$ 的充要条件是复数序列 $(A_1)_{ij}$, $(A_2)_{ij}, \cdots$ 以 A_{ij} 为极限:$A_{ij} = \lim\limits_{m\to\infty} (A_m)_{ij}$, $\forall\, 1 \leq i, j \leq n$。

设 $A_1, A_2, \cdots$ 是 $n \times n$ 复矩阵序列。若对任给的 $\varepsilon > 0$,都存在自然数 N,使 $\| A_k - A_m \| < \varepsilon$, $\forall\, k, m > N$,则称 $A_1, A_2, \cdots$ 是 Cauchy 序列。不难证明,$A_1, A_2, \cdots$ 是 Cauchy 序列的充要条件为复数序列 $(A_1)_{ij}, (A_2)_{ij}, \cdots$ 是 Cauchy 序列, $\forall\, 1 \leq i, j \leq n$。

数学分析告诉我们,复数(实数)序列有极限当且仅当它是 Cauchy 序列。因此 $n \times n$ 复矩阵序列有极限当且仅当它是 Cauchy 序列。

n 维内积空间 V 上的线性算符 A,在任给的标准正交基 $\{ e_1, \cdots, e_n \}$ 下按式(0.3.7)确定一个 $n \times n$ 矩阵 $[A]$。n 维内积空间 V 上的线性算符的序列 $\{ A_1, A_2, \cdots \}$,在任

给的标准正交基$\{e_1,\cdots,e_n\}$下按式(0.3.7)确定一个 $n\times n$ 矩阵序列$[A_1],[A_2],\cdots$。练习(0-4-2)告诉我们,$\|A\|=\|[A]\|$,$\|A_\alpha\|=\|[A_\alpha]\|$,$\forall\alpha\leq 1$。可见$\lim\limits_{n\to\infty}A_n=A\Leftrightarrow\lim\limits_{n\to\infty}[A_n]=[A]$。于是 n 维内积空间 V 上的线性算符的序列$A_1,A_2,\cdots$有极限,当且仅当它是 Cauchy 序列。

复数的无穷级数 $\sum\limits_{n=1}^{\infty}a_n$,是指其部分和序列$\{s_1,s_2,\cdots\}\left(s_n=\sum\limits_{i=1}^{n}a_i\right)$的极限:

$$\sum_{n=1}^{\infty}a_n=\lim_{m\to\infty}s_m=\lim_{m\to\infty}\sum_{n=1}^{m}a_n$$

练习 0-4-1 试证:设 A,B,U 是 $n\times n$ 复矩阵,$U^{-1}=U^{+}$,$A=UBU^{-1}$则$\|A\|=\|B\|$。

练习 0-4-2 试证:n 维内积空间 V 上的线性算符 A 的范数,等于它在任一标准正交基下的矩阵$[A]$的范数。

Ⅱ.矩阵的指数函数

由于 n 维内积空间 V 上的线性算符 A 的范数,等于它在任一标准正交基下的矩阵$[A]$的范数。所有我们对$(n\times n)$矩阵的幂级数的讨论,都适用于(n 维内积空间上的)线性

算符的幂级数。

定义 $n\times n$ 矩阵A(省去方括号)的指数函数为

$$\exp A\equiv e^{A}\equiv\sum_{j=0}^{\infty}\frac{A^j}{j!}\tag{0.4.4}$$

式中 A^0 是$(n\times n)$单位矩阵 E_n。为表明这一定义的合理性,考察部分和序列$\{S_0,S_1,S_2,\cdots\}$,其中 $S_m=\sum\limits_{j=0}^{m}\frac{A^j}{j!}$。对于 $k>m$,

$$\|S_k-S_m\|=\left\|\sum_{j=m+1}^{k}\frac{A^j}{j!}\right\|\leq\sum_{j=m+1}^{k}\frac{\|A\|^j}{j!}\tag{0.4.5}$$

由数学分析知,对任意$\|A\|\in(-\infty,\infty)$,$\exp\|A\|\equiv e^{\|A\|}=\sum\limits_{j=0}^{\infty}\frac{\|A\|^j}{j!}$的部分和序列有极限。这就是说,对任意$\|A\|\in(-\infty,\infty)$,$\sum\limits_{j=0}^{\infty}\frac{\|A\|^j}{j!}$的部分和序列是 Cauchy 序列。

由式(0.4.5)可知，对任意 $n\times n$ 矩阵 A，$\sum_{j=0}^{\infty}\frac{A^j}{j!}$ 的部分和序列是 Cauchy 序列，从而部分和序列的极限 $\lim_{m\to\infty}S_m$ 存在。因此 exp 是 M_n（全体 $n\times n$ 矩阵的集合）到自身的映射。

容易看到 $n\times n$ 矩阵 X 的指数函数 $Y=\exp X$ 当自变量 $X=0$ 时，函数值为 $Y=E$（$n\times n$ 单位矩阵）。当 $X=\begin{bmatrix}\xi_1 & & 0\\ & \xi_2 & \\ & & \ddots \\ 0 & & \xi_n\end{bmatrix}$（对角矩阵）时，$Y=\begin{bmatrix}e^{\xi_1} & & 0\\ & e^{\xi_2} & \\ & & \ddots \\ 0 & & e^{\xi_n}\end{bmatrix}$。当 $X=i2m\pi E$ 时，$Y=E$。可见 exp 不是 M_n 的单射。但是如果把定义域限制到零矩阵的一个充分小的邻域 U，我们可以把 exp 看作 U 到单位矩阵 E 的邻域 $\exp U$ 的双射。

命题 3 设 A,B 是 $n\times n$ 矩阵，B 是非奇异的，则

$$B(\exp A)B^{-1}=\exp(BAB^{-1})\equiv\sum_{j=0}^{\infty}\frac{(BAB^{-1})^j}{j!} \tag{0.4.6}$$

事实上，上式右边的部分和与上式左边的差的范数为

$$\left\|\sum_{j=0}^{m}\frac{(BAB^{-1})^j}{j!}-B(\exp A)B^{-1}\right\|$$

$$=\left\|B\left(\sum_{j=0}^{m}\frac{A^j}{j!}-\exp A\right)B^{-1}\right\|\leq\|B\|\left\|\sum_{j=0}^{m}\frac{A^j}{j!}-\exp A\right\|\|B^{-1}\|$$

由于 $\sum_{j=0}^{\infty}\frac{A^j}{j!}$ 的部分和序列极限存在，而且是 $\exp A$。故当 m 充分大时，上式右边可以任意小。

在进一步讨论之前，让我们回忆一下 $n\times n$ 复矩阵的一些性质。方阵的迹(trace)是它的所有对角元的和. 相似矩阵的特征值方程相同，从而特征值相同。相似矩阵的行列式相同，迹相同。上三角矩阵的特征值就是它的对角元。它的行列式就

是它的 n 个对角元的乘积,也就是它的所有特征值的乘积(m 重特征值在乘积中出现 m 次)。而任一 $n\times n$ 复矩阵都可以通过相似变换上三角化。故 $n\times n$ 复矩阵的行列式是它的所有特征值的乘积。我们用归纳法来证明。

命题 4 任一 $n\times n$ 复矩阵都可以通过相似变换上三角化。

当 $n=1$ 时,命题显然成立。设命题对 $n=m$ 成立。现讨论 $n=m+1$ 的情况。设 A 是 $n\times n$ 复矩阵,则它至少有一个特征值 λ。$Av_1=\lambda v_1, v_1\neq 0$。($v_1$ 是 $(m+1)\times 1$ 矩阵)存在非异 $(m+1)\times(m+1)$ 矩阵 B,使 $Bv_1=e_1$,($e_1=[1\quad 0\quad \cdots\quad 0]^T$)。易见 $(m+1)\times(m+1)$ 矩阵 BAB^{-1} 的第一列为 $BAB^{-1}e_1=BAv_1=\lambda Bv_1=\lambda e_1=[\lambda\quad 0\quad \cdots\quad 0]^T$。这意味着 $BAB^{-1}=\begin{bmatrix}\lambda & * & \cdots & *\\ 0 & * & \cdots & *\\ \vdots & \vdots & \ddots & \vdots\\ 0 & * & \cdots & *\end{bmatrix}=\begin{bmatrix}\lambda & * & \cdots & *\\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{bmatrix}$。由归纳假定,存在非异的 $m\times m$ 的矩阵 B' 使 $B'A'(B')^{-1}=\begin{bmatrix}k & * & \cdots & *\\ 0 & \mu & \cdots & *\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & v\end{bmatrix}$。令 $C=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{bmatrix}B$,则 C 是非异的,且 $CAC^{-1}=\begin{bmatrix}\lambda & * & \cdots & *\\ 0 & & & \\ \vdots & & B'A'(B')^{-1} & \\ 0 & & & \end{bmatrix}=\begin{bmatrix}\lambda & * & \cdots & *\\ 0 & k & \cdots & *\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & v\end{bmatrix}$。命题证完。

注意到 $\begin{bmatrix}\lambda_1 & * & \cdots & *\\ 0 & \lambda_2 & \cdots & *\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}\begin{bmatrix}\mu_1 & * & \cdots & *\\ 0 & \mu_2 & \cdots & *\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \mu_n\end{bmatrix}=\begin{bmatrix}\lambda_1\mu_1 & * & \cdots & *\\ 0 & \lambda_2\mu_2 & \cdots & *\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\mu_n\end{bmatrix}$,可知

$$\exp\begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}=\begin{bmatrix}\exp\lambda_1 & * & \cdots & * \\ 0 & \exp\lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \exp\lambda_n\end{bmatrix} \tag{0.4.7}$$

由以上的讨论可知

命题 5 对任一 $n\times n$ 复矩阵 A 都有 $\det(\exp A)=\exp(trA)$。

事实上，存在非异的 $n\times n$ 复矩阵 B，使 $BAB^{-1}=\begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}$ 是上三角矩阵。

$$\det(\exp A)=\det[B(\exp A)B^{-1}]=\det[\exp(BAB^{-1})]$$

$$=\det\left(\exp\begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}\right)$$

$$=\det\begin{bmatrix}e^{\lambda_1} & * & \cdots & * \\ 0 & e^{\lambda_2} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n}\end{bmatrix}=\prod_{j=1}^{n}e^{\lambda_j}=e^{\sum_{j=1}^{n}\lambda_j}=e^{trA}$$

Ⅲ. 矩阵的幂级数

矩阵的指数函数是矩阵的幂级数中的一种。为了更好地理解矩阵的幂级数，让我们在此回顾一下复变量的幂级数的性质。（读者可参考相关资料，例如，L. Ahlfors: "Complex Analysis"，郭敦仁：《数学物理方法》）

复变量 z 的幂级数是指 $\sum_{n=0}^{\infty}c_n(z-a)^n$。关于它的收敛性，我们有 Abel 的第一定理：若幂级数 $\sum_{n=0}^{\infty}c_n(z-a)^n$ 在 z_0 点收敛。则它在开圆 $|z-a|<|z_0-a|$ 上绝对

收敛,在任一闭圆$|z-a|\leq r(<|z_0-a|)$上一致收敛。

由这个定理可知,复变量 z 的幂级数 $\sum_{n=0}^{\infty}c_n(z-a)^n$ 的收敛行为分三类:

(1) 仅当 $z=a$ 时幂级数收敛,这时称幂级数的收敛半径 $r=0$;

(2) 在整个复平面上幂级数收敛,这时称幂级数的收敛半径 $r=\infty$;

(3) 存在有限大的正数 $r>0$,当 $|z-a|>r$ 时幂级数发散,当 $|z-a|<r$ 时幂级数收敛于一个解析函数。

考察下面的例子:$\sum_{n=0}^{\infty}n!(z-a)^n$,$e^z=\sum_{n=0}^{\infty}\frac{z^n}{n!}$,$\frac{1}{1-z}=\sum_{n=0}^{\infty}z^n$,$\ln(1+z)=\sum_{j=0}^{\infty}(-1)^{j+1}\frac{z^j}{j}$。

它们的收敛半径分别是 $0,\infty,1,1$。

Abel 第一定理的证明,用到复数的模的下列性质:对所有复数 a,b 和非负实数 α,都有

$$|a+b|\leq|a|+|b|,|\alpha a|=\alpha|a|,|ab|=|a|\cdot|b| \tag{0.4.8}$$

矩阵的范数对应的性质稍有不同:对所有 $n\times n$ 复矩阵 A,B 和非负实数 α,都有

$$\|A+B\|\leq\|A\|+\|B\|,\|\alpha A\|=\alpha\|A\|,\|AB\|\leq\|A\|\cdot\|B\| \tag{0.4.9}$$

特别是,两个非零复数之积非零,但两个范数非零的矩阵之积可以为零。这使得对矩阵变量的幂级数,Abel 第一定理不成立。相应地,我们有:

定理 1 设复变量 z 的幂级数 $\sum_{n=0}^{\infty}c_nz^n$ 的收敛半径为 r。若 $\|A\|<r$,矩阵变量 A 的幂级数 $\sum_{n=0}^{\infty}c_nA^n$ (绝对)收敛,特别是,幂级数 $\sum_{n=0}^{\infty}c_nA^n$ 在闭域 $\|A\|\leq r'(<r)$ 上一致地(绝对)收敛;但若 $\|A\|>r$,矩阵变量 A 的幂级数 $\sum_{n=0}^{\infty}c_nA^n$ 不一定发散。

证明 第二部分成立只须考察反例:$\sum_{n=0}^{\infty}z^n$ 的收敛半径 $r=1$。$A=\begin{bmatrix}0&2\\0&0\end{bmatrix}$ 的范数为 $\|A\|=2$。但 $\sum_{n=0}^{\infty}A^n=E_2+A=\begin{bmatrix}1&2\\0&1\end{bmatrix}$。

第一部分证明:选 $z_0\in C$,使 $\|A\|<|z_0|<r$。幂级数 $\sum_{n=0}^{\infty}c_nz_0^n$ 收敛。故有正整数

g,使$|c_n z_0^n| < g$, $\forall n \geq 0$。考察$\|c_n A^n + \cdots + c_{n+p} A^{n+p}\| \leq |c_n| \|A\|^n + \cdots + |c_{n+p}| \|A\|^{n+p} \leq |c_n| |z_0|^n \frac{\|A\|^n}{|z_0|^n} + \cdots + |c_{n+p}| |z_0|^{n+p} \frac{\|A\|^{n+p}}{|z_0|^{n+p}} \leq g\left(\frac{\|A\|^n}{|z_0|^n} + \cdots + \frac{\|A\|^{n+p}}{|z_0|^{n+p}}\right) \leq gq^n\left(\frac{1-q^{p+1}}{1-q}\right) \leq \frac{gq^n}{1-q}$,式中 $q = \frac{\|A\|}{|z_0|} < 1$。可见幂级数 $\sum_{n=0}^{\infty} |c_n| \|A\|^n$ 的部分和序列是 Cauchy 序列,从而幂级数 $\sum_{n=0}^{\infty} c_n A^n$(绝对)收敛。现讨论一致收敛性。对任一 $r' < r$,总有 $z_0 \in C$,使 $r' < |z_0| < r$。幂级数 $\sum_{n=0}^{\infty} c_n z_0^n$ 收敛。故有正整数 g,使$|c_n z_0^n| < g$, $\forall n \geq 0$。对所有的$\|A\| \leq r'$,幂级数 $\sum_{n=0}^{\infty} c_n A^n$ 的极限与部分和的差 $\left\| \sum_{m=1}^{\infty} c_m A^m - \sum_{m=1}^{n} c_m A^m \right\| = \|c_{n+1} A^{n+1} + c_{n+2} A^{n+2} + \cdots\| \leq |c_{n+1}| \cdot \|A\|^{n+1} + |c_{n+2}| \cdot \|A\|^{n+2} + \cdots \leq |c_{n+1}| r'^{n+1} + |c_{n+2}| r'^{n+2} + \cdots \leq g\left(\frac{r'^{n+1}}{|z_0|^{n+1}} + \frac{r'^{n+2}}{|z_0|^{n+2}} + \cdots\right) = \frac{gq'^{n+1}}{1-q'}$,式中 $q' = \frac{r'}{|z_0|} < 1$。可见对任给的 $\varepsilon > 0$,都存在自然数 N,使$\left\| \sum_{m=1}^{\infty} c_m A^m - \sum_{m=1}^{\infty} c_m A^m \right\| < \varepsilon$, $\forall n > N$, $\|A\| \leq r'$。此即幂级数$\sum_{n=0}^{\infty} c_n A^n$ 在闭域$\|A\| \leq r'(< r)$ 上一致地(绝对) 收敛。

这一定理使我们可以讨论两个幂级数的乘积。设$\sum_{n=0}^{\infty} a_n z^n$, $\sum_{n=0}^{\infty} b_n z^n$ 的收敛半径分别是r_1, r_2。则当$\|A\| < r = \min(r_1, r_2)$时,$\sum_{m=0}^{\infty} a_m A^m \sum_{n=0}^{\infty} b_n A^n = \sum_{k=0}^{\infty} c_k A^k$(绝对)收敛,式中 $c_n = \sum_{m=0}^{n} a_m b_{n-m}$。

例如,$\exp A = \sum_{m=0}^{\infty} \frac{A^m}{m!}$, $\exp(-A) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} A^n$,收敛半径都是 $r = \infty$。它们的乘积$(\exp A)[\exp(-A)] = \sum_{l=0}^{\infty} c_l A^l$ 的系数为$c_l = \sum_{m=0}^{l} \frac{1}{m!} \frac{(-1)^{l-m}}{(l-m)!} = (1-1)^l = \delta_{l0}$。故有 $\exp A[\exp(-A)] = E_n$。

Ⅳ. Baker-Hausdorff 公式

设 t 是一个实(复)参数,我们用 $A(t)$表示一个 $n \times n$ 复矩阵,其矩阵元 $A_{ij}(t)$是

t 的函数。称 $A(t)$是可微的，当且仅当所有矩阵元 $A_{ij}(t)$是可微的。$\frac{\mathrm{d}}{\mathrm{d}t}A(t)=\dot{A}(t)$ 是以 $\dot{A}_{ij}(t)$为矩阵元的矩阵。读者不难证明，如果 $A(t)$，$B(t)$是可微的 $n\times n$ 复矩阵，则下列等式成立：

$$\frac{\mathrm{d}}{\mathrm{d}t}[\alpha A(t)+\beta B(t)]=\alpha\dot{A}(t)+\beta\dot{B}(t)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[A(t)B(t)]=\dot{A}(t)B(t)+A(t)\dot{B}(t)$$

若 A 是一个 $n\times n$ 常数复矩阵，则有

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA)=\sum_{n=1}^{\infty}\frac{t^{n-1}A^n}{(n-1)!}=A\exp(tA)=\exp(tA)A$$

微分方程$\frac{\mathrm{d}}{\mathrm{d}t}A(t)=0$ 的通解是任意常数矩阵。

利用这些结果我们来讨论微分方程$\frac{\mathrm{d}}{\mathrm{d}t}A(t)=KA(t)$（式中 K 是一个 $n\times n$ 常数复矩阵）的通解。作变量变换 $B(t)=\exp(-Kt)A(t)$，则 $B(t)$满足微分方程。$\dot{B}(t)=-K\exp(-tK)A(t)+\exp(-tK)\dot{A}(t)=0$ 故 $B(t)=L$（任意常数矩阵），从而 $A(t)=\exp(tK)L$。类似的讨论可得，微分方程$\frac{\mathrm{d}}{\mathrm{d}t}A(t)=A(t)K$ 的通解为 $A(t)=L\exp(tK)$。

定理 2 设 A，B 是 $n\times n$ 矩阵。若$[A,B]=AB-BA=0$，则 $\exp(A+B)=\exp A\exp B$。

证明 我们用两种办法证明之。首先用两个绝对收敛的矩阵幂级数的乘法公式。上式右边等于 $\sum_{N=0}^{\infty}\frac{1}{N!}\left(\sum_{m=0}^{N}\frac{N!A^mB^{N-m}}{m!(N-m)!}\right)=\sum_{N=0}^{\infty}\frac{1}{N!}(A+B)^N=\exp(A+B)$，我们用到了条件 $AB=BA$。第二种证法：考虑矩阵函数 $C(t)=\exp(-t(A+B))\exp(tA)\exp(tB)$对参数 t 的微商。$\dot{C}(t)=-(A+B)C(t)+AC(t)+BC(t)=0$。故 $C(t)=C(0)=E_n$。以 $\exp(t(A+B))$左乘之得 $\exp(tA)\exp(tB)=\exp(t(A+B))$。当 $t=1$ 时，我们得到 $\exp(A+B)=\exp A\exp B$。

当 A，B 不对易时，上式不成立。但是当 A，B 充分接近 0 时，在 E_n 的充分小的邻域中有唯一的依赖 A，B 的 $n\times n$ 复矩阵 $C(A,B)$，使 $\exp A\exp B=\exp C(A,B)$。当 A，B 对易时，$C=A+B$。在一般情况下，我们有 Baker-Hausdorff 公式：

定理 3 对于充分接近零的 $n \times n$ 复矩阵 A, B 和充分接近 E_n 的 $n \times n$ 复矩阵，$C = \ln(\exp A \exp B)$，有

$$C = B + \int_0^1 g[\exp(t\, adA)\exp(adB)](A)\mathrm{d}t \tag{0.4.10}$$

式中 g 是由下式定义的幂级数：

$$g(z) = \frac{\ln z}{z-1} = \sum_{j=0}^{\infty} \frac{(1-z)^j}{j+1} = 1 + \frac{1}{2}(1-z) + \frac{1}{3}(1-z)^2 + \cdots, |1-z| < 1 \tag{0.4.11}$$

为证明这一定理，需要首先建立一些引理。这些引理本身用处就很大。

我们以 $M_n(C)$ 记所有 $n \times n$ 复矩阵的集合。在矩阵加法和复数与矩阵乘法下 $M_n(C)$ 是一 n^2 维复线性空间。加上矩阵乘法 $M_n(C)$ 是一 n^2 维复代数。设 $A \in M_n(C)$，定义线性空间 $M_n(C)$ 上的线性算符 $adA: M_n(C) \to M_n(C)$，使 $adA(B) = AB - BA$，$\forall B \in M_n(C)$。设 V 是 n 维复线性空间，V 上线性算符的全体 $EndV$ 是一 n^2 维复代数。$EndV$ 与 $M_n(C)$ 同构。与上面关于矩阵代数的讨论平行，对线性算符代数有：设 $A \in EndV$，定义 $EndV$ 上的线性算符 $adA: EndV \to EndV$，使 $abA(B) = AB - BA$，$\forall B \in EndV$。

引理 1 设 A, B 是 $n \times n$ 复矩阵（或者说，n 维复线性空间 V 上的线性算符），则

$$e^A B e^{-A} = (\exp(adA))(B) = \sum_{j=0}^{\infty} \frac{(adA)^j}{j!} B = B + \frac{1}{1!}[A, B] + \frac{1}{2!}[A, [A, B]] + \cdots \tag{0.4.12}$$

证明 令 $F(t) = e^{tA} B e^{-tA}$，显然有 $F(0) = B$，而且 $F(t)$ 的矩阵元是 t 的整函数。于是存在矩阵序列 $C_0, C_1, C_2, \cdots$ 使 $F(t) = \sum_{j=0}^{\infty} C_j t^j$，$C_0 = B$，$C_j = \frac{1}{j!} \frac{\mathrm{d}^j}{\mathrm{d}t^j} F(t)|_{t=0}$，$\forall j \geq 1$。$\dot{F}(t) = AF(t) - F(t)A = adA(F(t))$。对左式微商得 $\frac{\mathrm{d}^j}{\mathrm{d}t^j} F(t) = (adA)^j (F(t))$。$C_j = \frac{(adA)^j}{j!} B$，$F(t) = \sum_{j=0}^{\infty} \frac{(adA)^j}{j!} B\, t^j$。当 $t = 1$ 时得 $e^A B e^{-A} = [\exp(adA)](B)$。

引理 2 设 $A(t)$ 是 $n \times n$ 复矩阵（或者说，n 维复线性空间 V 上的线性算符），它的每一个矩阵元 $A_{jk}(t)$ 都是参数 t 的解析函数，而 $f(z) = \frac{e^z - 1}{z} = 1 + \frac{z}{2!} + \frac{z^2}{3!} + \cdots (z$

的一个整函数)则

$$e^{A(t)}\frac{\mathrm{d}}{\mathrm{d}t}e^{-A(t)} = -f(adA(t))(\dot{A}(t)) \tag{0.4.13}$$

证明 令 $B(s,t)=e^{sA(t)}\frac{\mathrm{d}}{\mathrm{d}t}e^{-sA(t)}$。它是 s 的整函数,故有 $B(s,t)=\sum_{j=0}^{\infty}C_j(t)s^j$, $C_j=\frac{1}{j!}\frac{\mathrm{d}^j}{\mathrm{d}s^j}B(s,t)|_{s=0}$。显然有 $B(0,t)=0$。计算 $B(s,t)$对 s 的偏微商$\frac{\partial}{\partial s}B(s,t)=A(t)B(s,t)-e^{sA(t)}\frac{\mathrm{d}}{\mathrm{d}t}(e^{-sA(t)}A(t))=(adA(t))B(s,t)-\dot{A}(t)$,此式两边对 s 的偏微商给出$\frac{\partial^j}{\partial s^j}B(s,t)=(adA(t))^jB(s,t)-(adA(t))^{j-1}\dot{A}(t)$, $\forall j\geq 1$。因而我们有 $C_0=0$, $C_j=\frac{-1}{j!}(adA(t))^{j-1}\dot{A}(t)$, $B(s,t)=\sum_{j=1}^{\infty}\frac{-1}{j!}(adA(t))^{j-1}\dot{A}(t)s^j$。当 $s=1$ 时,此式给出 $e^{A(t)}\frac{\mathrm{d}}{\mathrm{d}t}e^{-A(t)}=-f(adA(t))\dot{A}(t)$。

现在我们来证明 Baker-Hausdorff 定理。令 $C(t)=\ln(e^{tA}e^{B})$, $0\leq t\leq 1$。由引理 1 和 $e^{tA}e^{B}=e^{C(t)}$,可知对任一 $n\times n$ 复矩阵 H 有 $\exp(adC(t))H=e^{C(t)}He^{-C(t)}=e^{tA}(e^{B}He^{-B})e^{-tA}=\exp(ad(tA))\exp(adB)H$。从而 $\exp(adC(t))=\exp(ad(tA))\exp(adB)$, $adC(t)=\ln\{\exp[ad(tA)]\exp(adB)\}$。由 $C(t)$的定义得 $e^{C(t)}\frac{\mathrm{d}}{\mathrm{d}t}e^{-C(t)}=-e^{tA}e^{B}e^{-B}e^{-tA}A=-A$。而由引理 2 知

$$f(adC(t))\dot{C}(t)=A \tag{0.4.14}$$

此处的 $f(z)=\frac{\exp z-1}{z}$。注意到 $g(z)=\frac{\ln z}{z-1}$,以及 $g(e^z)f(z)=1$。我们有 $g(e^{\Omega})f(\Omega)=E_{n\times n}$,对任一接近零的 $M_n(C)$上的线性算符 Ω。以 $g\{\exp[adC(t)]\}$作用于式(0.4.14)两边得

$$\dot{C}(t)=g\{\exp[adC(t)]\}A=g\{\exp[ad(tA)]\exp(adB)\}A$$

积分上式得

$$C(t)=\int_0^t \mathrm{d}t' g\{\exp[ad(t'A)]\exp(adB)\}A+B \tag{0.4.15}$$

当 $t=1$ 时,上式给出式(0.4.10)。将其被积函数展开幂级数,并逐项积分得

$$C=A+B+\frac{1}{2}[A,B]+\frac{1}{12}[A,[A,B]]-\frac{1}{12}[B,[B,A]]+\cdots \tag{0.4.16}$$

当$[A,B]$与 A,B 都对易时,我们有

推论 若$[A,[A,B]]=[B,[A,B]]=0$,则 $e^Ae^B=e^{A+B}e^{\frac{1}{2}[A,B]}$。

练习 0-4-3 试证: $ad[A,B]=[adA,adB]$。

§0.5 对偶空间、线性空间的张量积

Ⅰ. F – 线性空间的对偶空间

设 V 是 F – 线性空间。考虑全体 V 上线性函数的集合。它在函数加法以及函数与数的乘法下构成一个 F – 线性空间。我们称之为 V 的对偶空间,记作 V^*。

若 V 是有限维的,则不难证明 $\dim V=\dim V^*$。事实上,若 $\dim V=n$,选定 V 的一个基 $e_1,\cdots e_n$。V 中每一向量 x 都可以唯一地展成基矢的线性组合 $x=e_\alpha\xi^\alpha$。定义 n 个 V 上函数 $\tilde{e}^1,\cdots,\tilde{e}^n$,使 $\tilde{e}^\alpha(x)=\xi^\alpha$。易知,它们是线性函数。由定义可知,$\tilde{e}^\alpha(e_\beta)=\delta^\alpha_\beta$。我们来证明 $\tilde{e}^1,\cdots,\tilde{e}^n$ 构成对偶空间 V^* 的一个基。先证明任一 V 上线性函数 f 都可以写作 $\tilde{e}^1,\cdots,\tilde{e}^n$ 的线性组合。事实上,对 V 中每一向量 x 都有 $f(x)=f(e_\alpha\xi^\alpha)=f(e_\alpha)\xi^\alpha=f(e_\alpha)(\tilde{e}^\alpha(x))=(f(e_\alpha)\tilde{e}^\alpha)(x)$。从而 $f=f(e_\alpha)\tilde{e}^\alpha$,这就是说 $\tilde{e}^1,\cdots,\tilde{e}^n$ 是完备的。现证明 $\tilde{e}^1,\cdots,\tilde{e}^n$ 是线性独立的。若 $\lambda_\alpha\tilde{e}^\alpha=0$,则 $\lambda_\alpha\tilde{e}^\alpha(e_\beta)=0,\forall 1\leq\beta\leq n$。此即 $\lambda_\alpha\delta^\alpha_\beta=\lambda_\beta=0,\forall 1\leq\beta\leq n$。从而 $\tilde{e}^1,\cdots,\tilde{e}^n$ 是线性独立的。称对偶空间 V^* 的基 $\{\tilde{e}_1,\cdots,\tilde{e}_n\}$ 是 F – 线性空间 V 的基 $\{e^1,\cdots,e^n\}$ 的对偶基。请注意这儿定义的对偶基与 §0.2 定义的 F – 内积空间中,基 $\{e_1,\cdots,e_n\}$ 的对偶基 $\{e^1,\cdots,e^n\}$ 是不同的。稍后我们将讨论二者之间的关系。

Ⅱ. F – 内积空间的对偶空间

F – 内积空间是定义了向量内积的 F – 线性空间。设 V 是 F – 内积空间,$v\in V$。则对应 $V\ni x\mapsto\langle v|x\rangle\in F$,确定了一个 V 上的线性函数,记作 $\langle v|:V\to F$。由内积

定义不难证明$\langle \alpha u+\beta v|=\overline{\alpha}\langle u|+\overline{\beta}\langle v|$。此外，若两个向量$v,u\in V$确定的线性函数$\langle v|,\langle u|$相同，则必有$v=u$。我们还可以证明，对任一$V$上的线性函数$f$，都存在$w\in V$，使得$f=\langle w|$。（留作练习）因此$F$－内积空间$V$的对偶空间就是$V^*=\{\langle v||v\in V\}$。

若$e_1,\cdots,e_n$是F－内积空间V的一个基，则$\langle e_1|,\cdots,\langle e_n|$是对偶空间$V^*$的一个基。事实上，由于已知$\dim V=\dim V^*$，只要证明$\langle e_1|,\cdots,\langle e_n|$线性独立就行。若$\alpha^1\langle e_1|+\cdots+\alpha^n\langle e_n|=0$，则$\alpha^1\langle e_1|x\rangle+\cdots+\alpha^n\langle e_n|x\rangle=0,\forall x\in V$此即$\langle\overline{\alpha^1}e_1+\cdots+\overline{\alpha^n}e_n|x\rangle=0,\forall x\in V$。从而$\overline{\alpha^1}e_1+\cdots+\overline{\alpha^n}e_n=0$。由$e_1,\cdots,e_n$是线性独立的，可知所有$\overline{\alpha^i}=0$，于是所有$\alpha^i=0$，$\langle e_1|,\cdots,\langle e_n|$是线性独立的，构成$V^*$的一个基。

设V是F－内积空间，$e_1,\cdots,e_n$是V的一个基。若$e^1,\cdots,e^n$是§0.2定义的对偶基，则$\langle e^1|,\cdots,\langle e^n|$是对偶空间$V^*$的基，而且是$F$－内积空间$V$的基$e_1,\cdots,e_n$的对偶基，即$\langle e^\alpha|=\tilde{e}^\alpha,\forall 1\leq\alpha\leq n$。

特别是，当$e_1,\cdots,e_n$是F－内积空间V的一个标准正交基时，$\langle e_1|,\cdots,\langle e_n|$不仅是对偶空间$V^*$的基，而且是$F$－内积空间$V$的基$e_1,\cdots,e_n$的对偶基。

练习0-5-1 试证：设V是F－内积空间，则对任一V上的线性函数f，都存在唯一的$w\in V$，使得$f=\langle w|$。

Ⅲ. Dirac符号

在量子力学中，体系的态矢量空间V是复Hilbert空间，当然是复内积空间。仿效Dirac，把其中向量ψ,φ记作$|\psi\rangle,|\varphi\rangle$，把其中向量$\psi,\varphi$的内积记作$\langle\psi|\varphi\rangle$，称$V$为右矢空间，而称其对偶空间$V^*=\left\{\langle\psi|\,\middle|\,|\varphi\rangle\in V\right\}$为左矢空间。每一个左矢量$\langle\lambda|\in V^*$，是一个$V$上的线性函数。

Dirac用记号$|\psi\rangle\langle\varphi|$表示这样一个$V$上的线性算符，$|\psi\rangle\langle\varphi|:V\to V$，它把任一态矢量$|\lambda\rangle$变成态矢量$|\psi\rangle\langle\varphi|\lambda\rangle$。注意到左矢是$V$上的线性函数（$V$到$F$的线性映射），$\langle\mu|:V\to F$。复合映射$(\langle\mu|)\circ(|\psi\rangle\langle\varphi|):V\to F$是一个$V$上的线性函数，它把任一右矢$|\lambda\rangle$映成数$\langle\mu|\psi\rangle\langle\varphi|\lambda\rangle$。Dirac把这一线性函数记作$\langle\mu|\psi\rangle\langle\varphi|:V\to F$。我们看到，映射$|\psi\rangle\langle\varphi|:V\to V$右乘映射$\langle\mu|:V\to F$的乘积是映射$\langle\mu|\psi\rangle\langle\varphi|:V\to$

F。在这个意义下,物理学家说,线性算符$|\psi\rangle\langle\varphi|$从右边作用于一个左矢$\langle\mu|$得到一个新的左矢$\langle\mu|\psi\rangle\langle\varphi|$。他们把$|\psi\rangle\langle\varphi|$也看作左矢空间上的线性算符。

由内积空间 V 上线性算符A 的厄米共轭A^+的定义$\langle x|A|y\rangle=\langle A^+x|y\rangle$, $\forall x,y\in V$,可知$(|\psi\rangle\langle\varphi|)^+=|\varphi\rangle\langle\psi|$。事实上$\langle(|\psi\rangle\langle\varphi|)^+x|y\rangle=\langle x|(|\psi\rangle\langle\varphi|)y\rangle=\langle x|\psi\rangle\langle\varphi|y\rangle=\langle(\varphi\langle\psi|x\rangle)|y\rangle=\langle(|\varphi\rangle\langle\psi|)x|y\rangle$, $\forall x,y\in V$。

设$\{|u_1\rangle,|u_2\rangle,\cdots\}$是态矢量空间 V 的一个基。由式(0.2.2)可知,$\sum\limits_{\alpha\geq1}|\widetilde{u}_\alpha\rangle\langle u_\alpha|=1$。特别是$\{|u_1\rangle,|u_2\rangle,\cdots\}$是态矢量空间 V 的一个标准正交基时,我们有$\sum\limits_{\alpha\geq1}|u_\alpha\rangle\langle u_\alpha|=1$。

练习 0-5-2 设$\{|u_1\rangle,|u_2\rangle,\cdots\}$是态矢量空间 V 的一个标准正交基,则任一 V 上的线性算符A 可以唯一地写作$A=\sum\limits_{\alpha,\beta\geq1}|u_\alpha\rangle\langle u_\beta|A_{\alpha\beta}$,式中 $A_{\alpha\beta}=\langle u_\alpha|A|u_\beta\rangle$,是线性算符 A 在基$\{|u_1\rangle,|u_2\rangle,\cdots\}$下的矩阵元。

Ⅳ. 线性空间的张量积

让我们通过考察下面的例子,来引入线性空间张量积的概念。设 V_e 是单个电子体系的态矢量空间,它是由所有模平方可积的复值函数 $\psi(\boldsymbol{r}_e,s_{ze})$构成的,$\{u_n(\boldsymbol{r}_e,s_{ze})|n\geq1\}$是 V_e 的一个基;而 V_p 是单个质子体系的态矢量空间,它是由所有模平方可积的复值函数 $\varphi(\boldsymbol{r}_p,s_{zp})$构成的,$\{w_m(\boldsymbol{r}_p,s_{zp})|m\geq1\}$是 V_p 的一个基。现考虑氢原子体系(一个电子加一个质子)的态矢量空间 V_{e+p},它是由所有模平方可积的复值函数 $\Psi(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})$构成的。我们来证明,$\{u_n(\boldsymbol{r}_e,s_{ze})w_m(\boldsymbol{r}_p,s_{zp})|n\geq1\&m\geq1\}$构成 V_{e+p}的一个基。事实上,当固定氢原子任一波函数 $\Psi(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})$的自变量 $\boldsymbol{r}_e,s_{ze}$ 时,可以把它看作是以 $\boldsymbol{r}_p,s_{zp}$为自变量的模平方可积的复值函数,从而可以唯一地展成$\{w_m(\boldsymbol{r}_p,s_{zp})|m\geq1\}$的线性组合

$$\Psi(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})=\sum_{m\geq1}c_m(\boldsymbol{r}_e,s_{ze})w_m(\boldsymbol{r}_p,s_{zp})$$

式中系数 $c_m(\boldsymbol{r}_e,s_{ze})$又可以唯一地展成$\{u_n(\boldsymbol{r}_e,s_{ze})|n=1,2,\cdots N\}$的线性组合

$$c_m(\boldsymbol{r}_e,s_{ze})=\sum_{n\geq1}a_{nm}u_n(\boldsymbol{r}_e,s_{ze})$$

从而有

$$\Psi(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})=\sum_{m\geq1}\sum_{n\geq1}a_{nm}u_n(\boldsymbol{r}_e,s_{ze})w_m(\boldsymbol{r}_p,s_{zp})$$

此即，V_{e+p}中任一矢量$\Psi(\boldsymbol{r}_e, s_{ze}, \boldsymbol{r}_p, s_{zp})$可唯一地展成$\{u_n(\boldsymbol{r}_e, s_{ze})w_m(\boldsymbol{r}_p, s_{zp}) \mid n, m \geq 1\}$的线性组合，后者构成$V_{e+p}$的一个基。我们把$V_{e+p}$叫做是$V_e$与$V_p$的张量积，记作$V_{e+p} = V_e \otimes V_p$。

定义 设V和U是F-线性空间，$\dim V = n$，$\dim U = m$。若W是F-线性空间，且定义了一个双线性映射$f: V \times U \to W$，使得对V的任意一组基$\{e_1, \cdots, e_n\}$和U的任意一组基$\{c_1, \cdots, c_m\}$，$\{f(e_j, c_i) \mid 1 \leq j \leq n, 1 \leq i \leq m\}$都构成$W$的一组基。则称$\{W, f\}$(或简称$W$)是$F$-线性空间$V$和$U$的张量积，记作$W = V \otimes U$。对任一$v \in V$和$u \in U$，称$f(v, u)$为向量$v$和$u$的张量积，记作$f(v, u) = v \otimes u$。

这一定义，可方便地推广到任意有限个F-线性空间的张量积的定义。

定义 设$V_\alpha(\alpha = 1, \cdots, s)$是$F$-线性空间，$\dim V_\alpha = n_\alpha$，$(\alpha = 1, \cdots, s)$。若$W$是$F$-线性空间，且定义有一个$s$重线性映射$f: V_1 \times \cdots \times V_s \to W$，使得对$V_1$的任意一组基$\{e_1^1, \cdots, e_{n_1}^1\}$，$V_2$的任意一组基$\{e_1^2, \cdots, e_{n_2}^2\}$，$\cdots V_s$的任意一组基$\{e_1^s, \cdots, e_{n_s}^s\}$，$\{f(e_{i_1}^1, \cdots, e_{i_s}^s) \mid 1 \leq i_\alpha \leq n_\alpha, (1 \leq \alpha \leq s)\}$都构成$W$的一组基。则称$\{W, f\}$(或简称$W$)是$F$-线性空间$V_1, V_2, \cdots, V_s$的张量积，记作$W = V_1 \otimes \cdots \otimes V_s$。对任一组$v_\alpha \in V_\alpha(\alpha = 1, \cdots, s)$，称$f(v_1, v_2, \cdots, v_s)$为向量$v_1, \cdots, v_s$的张量积，记作$f(v_1, \cdots, v_s) = v_1 \otimes \cdots \otimes v_s$。

在量子力学中，若某个体系是由若干子体系构成的，则它的位形空间是其子体系位形空间的笛卡儿积，而它的态空间就是其子体系态空间的张量积。

练习 0-5-3 (思考题)是否张量积空间$V \otimes U$中的每一向量都是V中某个向量v与U中某个向量u的张量积$v \otimes u$？证明你的断言。

练习 0-5-4 设X, Y是非空有限集，它们的基数(元素的个数)分别是$|X| = m$，$|Y| = n$。C是复数域。试证：全体由集X(集Y)到C的映射的集合，按通常方式定义加法和数乘后，构成一个m维(n维)的复线性空间，全体由笛卡尔积集$X \times Y$到C的映射的集合，按通常方式定义加法和数乘后，构成一个$m \cdot n$维的复线性空间，而且是前两者的张量积。

§0.6 非对易分析与超算符
(Non-commutative Calculus and Hyper-operators)

在量子力学中，体系的状态由一个定义在其位形空间上的模平方可积的复值函

数描写。定义在这个位形空间上的模平方可积的复值函数的全体构成一个 Hilbert 空间,称之为该体系的态矢量空间;我们把它记作 S。体系的力学量由 S 上的厄米算符描写。力学量算符间的关系基本上就是对应的经典力学量间的关系。例如,动能算符与动量算符间的关系为 $\hat{T}=\frac{\hat{\boldsymbol{p}}^2}{2m}$,轨道角动量算符与坐标算符和动量算符间的关系为 $\hat{\boldsymbol{l}}=\hat{\boldsymbol{r}}\times\hat{\boldsymbol{p}}$,等等。所以在量子力学中,你会经常遇到这种函数,其自变量是 S 上的线性算符,函数值也是 S 上的线性算符,例如 e^X,$\ln\rho$(X,ρ 是 S 上的线性算符)。这类函数我们将简称为算符函数。算符函数的微商与大学微积分中普通函数的微商有很大的不同。亲爱的读者,你能回答这个问题吗?$(e^X)'=?$ 如果你回答不出,请不要太泄气,因为也许很多物理学家也不知道它的答案。但这并不表明这个问题太深奥或不重要。之所以在这里介绍非对易分析和超算符的概念,是因为作者觉得,只花几个小时就可以在观念上提高一大步并受益后半生,对物理系学生而言,是一件非常值得做的事。

e^X,$\ln\rho$ 等等都是通过幂级数定义的。有人认为自变量是线性算符,函数值也是线性算符的"解析函数"的最一般的形式就是自变量的系数为复数的幂级数。我们将指出,这是一种误解。"解析的算符函数"要远比自变量的系数为复数的幂级数广泛得多。

Ⅰ. 向量的范数(norm)

向量的范数就是向量的大小。向量 x 的范数记作$\|x\|$,应该满足下面的条件:$\|\alpha x\|=|\alpha|\|x\|$($\alpha$ 是数);$\|x\|=0\Leftrightarrow x=0$;$\|x+y\|\leq\|x\|+\|y\|$(三角形不等式)。在内积空间中,总是把向量的范数定义作$\|x\|=\langle x|x\rangle^{\frac{1}{2}}$。

设 A 是 m 维复内积空间。X 是 n 维复内积空间。我们以 $\Lambda(A;X)$表示 A 到 X 的线性映射的全体所成的 $m\times n$ 维复线性空间。对 $J\in\Lambda(A;X)$,可以证明存在唯一 $P\in\Lambda(A;X)$使

$$\langle x\mid Ja\rangle_X=\langle Px\mid a\rangle_A,\forall\, a\in A,x\in X \tag{0.6.1}$$

我们把 P 称作 J 的厄米共轭,并记作 $P=J^+$;于是

$$\langle x\mid Ja\rangle_X=\langle J^+x\mid a\rangle_A,\forall\, a\in A,x\in X \tag{0.6.2}$$

注意,当 A 不同于 X 时,$J\in\Lambda(A;X)$和 $J^+\in\Lambda(X;A)$是不同性质的线性映射。对

$J, K \in \Lambda(A;X)$,定义 J 与 K 的内积

$$\langle J \mid K \rangle = tr_A(J^+ K) \tag{0.6.3}$$

注意,$J^+ K$ 是 A 上的线性算符,$tr_A(J^+ K)$与 A 的基的选择无关。这样 $\Lambda(A;x)$就成了一个复内积空间。利用内积,定义 $J \in \Lambda(A;X)$的范数

$$\| J \| = \langle J \mid J \rangle^{1/2} \tag{0.6.4}$$

容易证明这样定义的范数满足下面的条件:

$$\| \alpha J \| = |\alpha| \| J \| (\alpha \text{ 是任一复数})$$

$$\| J \| = 0 \Leftrightarrow J = 0 \tag{0.6.5}$$

$$\| J + K \| \leq \| J \| + \| K \|$$

有了范数,我们可以利用它,给向量序列的极限下定义。

现在让我们从一般的数学理论回到量子力学。设体系的态矢量空间 S 是 N 维复 Hilbert 空间。于是,S 上线性算符的全体 $\Lambda(S;S) \equiv \mathbf{M}$ 是 N^2 维复线性空间。对 $A, B \in \mathbf{M}$,定义 A, B 的内积为

$$\langle A \mid B \rangle = tr(A^+ B) \tag{0.6.6}$$

这样 $\mathbf{M}$ 就成了一个复内积空间。利用内积,定义 $A \in \mathbf{M}$ 的范数

$$\| A \| = \langle A \mid A \rangle^{1/2} \tag{0.6.7}$$

现在我们可以谈论 S 上线性算符 A 的范数,也就是 A 的大小了。

同理考虑 $\mathbf{M}$ 到 $\mathbf{M}$ 的线性映射的全体 $\Lambda(\mathbf{M};\mathbf{M})$,它的元素叫超(线性)算符。$\Lambda(\mathbf{M};\mathbf{M})$是一个 $N^2 \times N^2$ 维的复线性空间。超算符 $\chi \in \Lambda(\mathbf{M};\mathbf{M})$的厄米共轭 $\chi^+ \in \Lambda(\mathbf{M};\mathbf{M})$由下式确定

$$\langle B \mid \chi A \rangle = \langle B \mid \chi A \rangle = \langle \chi^+ B \mid A \rangle \quad \forall A, B \in \mathbf{M} \tag{0.6.8}$$

定义两个超算符 $\chi, y \in \Lambda(\mathbf{M};\mathbf{M})$的内积

$$\langle \chi \mid y \rangle = tr(\chi^+ y) \tag{0.6.9}$$

这样 $\Lambda(\mathbf{M};\mathbf{M})$就是一个 $N^2 \times N^2$ 维的复内积空间。利用内积,可定义超算符 $\chi \in \Lambda(\mathbf{M};\mathbf{M})$的范数

$$\| \chi \| = \langle \chi \mid \chi \rangle^{1/2} \tag{0.6.10}$$

按同样的方法,我们还可以定义 $\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$中向量(超超算符)的厄米共轭,两个向量的内积和向量的范数。这些是讨论算符函数微商的必要的预备知识。

练习 0-6-1 试证:若在内积空间中定义向量的范数为 $\| x \| = \langle x \mid x \rangle^{1/2}$,则有

(ⅰ)$\|\alpha x\| = |\alpha|\|x\|$($\alpha$ 是数);(ⅱ)$\|x\| = 0 \Leftrightarrow x = 0$;(ⅲ)$\|x + y\| \leq \|x\| + \|y\|$(三角形不等式)。

练习 0-6-2 试证:对所有 $J \in \Lambda(A;X)$都有$(J^+)^+ = J$。

练习 0-6-3 试证:设 $A, B \in \mathbf{M} \equiv \Lambda(S;S)$,$[A]$,$[B]$分别是它们在 S 的某个正交归一基下的矩阵;则按(0.6.6),(0.6.7),$\langle A|B\rangle = \sum\limits_{\alpha,\beta=1}^{N} \overline{A^{\alpha}_{\beta}} B^{\alpha}_{\beta}$,$\langle A|A\rangle = \sum\limits_{\alpha,\beta=1}^{N} |A^{\alpha}_{\beta}|^2$。

Ⅱ. 算符函数的微商

设 $\mathbf{U}$ 是 $\mathbf{M} \equiv \Lambda(S;S)$的一个开子集。算符函数指的就是映射

$$F : \mathbf{U} \to \mathbf{M} \tag{0.6.11}$$

为了定义 F 的微商,让我们回忆大学微积分教程中普通函数 $f:(a,b) \to \mathbf{R}$ 微商的定义:对 $x \in (a,b)$

$$f'(x) \equiv \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{0.6.12}$$

如果我们模仿它,对 $X \in U$,令

$$F'(x) \equiv \lim_{\Delta x \to 0} \frac{F(X + \Delta X) - F(X)}{\Delta X} \tag{0.6.13}$$

则会发现这是行不通的。首先,Δx 是数,当 $\Delta X \neq 0$ 时,$\frac{1}{\Delta x}$总是存在的;而 ΔX 是 S 上的线性算符,非零线性算符 ΔX 不一定有逆。其次,$\frac{1}{\Delta x}$与$f(x + \Delta x) - f(x)$对易,而线性算符的乘法是非交换的。即便$\frac{1}{\Delta X}$存在,$\frac{1}{\Delta X}(F(X + \Delta X) - F(X))$与$(F(X + \Delta X) - F(X))\frac{1}{\Delta X}$不一定相等。

幸好,$f'(x)$有另一种(等价的)定义方法:$x \in (a,b)$,若对所有满足条件 $x + \Delta x \in (a,b)$的 Δx 都有

$$f(x + \Delta x) - f(x) = \varphi(x)\Delta x + o(\Delta x) \tag{0.6.14}$$

其中 $\varphi(x)$与 Δx 无关,而 $o(\Delta x)$是 Δx 的高阶无穷小

$$\frac{|o(\Delta x)|}{|\Delta x|} \xrightarrow[\Delta x \to 0]{} 0 \tag{0.6.15}$$

则称 f 在点 x 可微,称 $\varphi(x)$为 f 在点 x 的微商,记作 $\varphi(x) = f'(x)$。于是

$$f(x + \Delta x) - f(x) = f'(x)\Delta x + o(\Delta x) \tag{0.6.16}$$

注意,上式右边的第一项是函数改变量中线性地依赖于的 Δx 部分。微分的实质就是把函数局部地线性化。仿此,我们给出下面的定义:

定义 $\mathbf{U}$ 是 $\mathbf{M}$ 的开子集,$F:\mathbf{U}\to\mathbf{M}$, $X\in\mathbf{U}$。若对所有满足条件 $X+\Delta X\in\mathbf{U}$ 的 ΔX 都有

$$F(X+\Delta X)-F(X)=\varphi(X)\Delta X+o(\Delta X) \tag{0.6.17}$$

其中 $\Phi(X)\in\Lambda(\mathbf{M};\mathbf{M})$是与 ΔX 无关的超算符,线性算符 $o(\Delta X)$是 ΔX 的高阶小

$$\frac{\|o(\Delta X)\|}{\|\Delta X\|}\xrightarrow[\Delta X\to 0]{}0 \tag{0.6.18}$$

则 F 称在点 X 可微,称 $\Phi(X)$为 F 在点 X 的微商,记作 $\Phi(X)=F'(X)$。于是

$$F(X+\Delta X)-F(X)=F'(X)\Delta X+o(\Delta X) \tag{0.6.19}$$

注意,上式右边第一项是函数增量中线性地依赖于 ΔX 的部分。线性算符 $\varphi(X)$左乘线性算符 ΔX 的乘积 $\varphi(X)\Delta X$ 是线性地依赖于 ΔX 的线性算符,但远远不能穷尽后者,故 $\Phi(X)$应是超线性算符。$F'(X)\Delta X$ 不应读作两个线性算符 $F'(X)$与 ΔX 的乘积,而应读作超算符 $F'(X)$作用在线性算符 ΔX 上。

当 $F:\mathbf{U}\to\mathbf{M}$ 在 $\mathbf{U}$ 的每点 X 都可微,则称 F 是 $\mathbf{U}$ 上的可微函数。这时对每一 $X\in\mathbf{U}$,令 $X\mapsto F'(X)\in\Lambda(\mathbf{M};\mathbf{M})$,我们得到 F 的导函数

$$F':\mathbf{U}\mapsto\Lambda(\mathbf{M};\mathbf{M}) \tag{0.6.20}$$

可见函数 F 和它的一阶导函数 F'是性质不同的映射。

Ⅲ.超(线性)算符

命题 设 U,V 是复内积空间。$\{|u_\alpha\rangle|1\le\alpha\le m\}\{|v_\rho\rangle|1\le\rho\le n\}$,分别是它们的正交归一基。则$\{|v_\rho\rangle\langle u_\alpha\|1\le\rho\le n,1\le\alpha\le m\}$构成 $\Lambda(U;V)$的一个正交归一基。

把这一普遍成立的数学命题应用到量子力学。

设$\{|\varphi_\alpha\rangle|1\le\alpha\le N\}$是态矢量空间 S 的一个正交归一基,则$\{|\varphi_\alpha\rangle\langle\varphi_\beta\|1\le\alpha,\beta\le N\}$是 $\mathbf{M}\equiv\Lambda(S;S)$的正交归一基。

对 $A\in\mathbf{M}$,定义超算符 $\mathscr{L}mA,\mathscr{R}mA,\mathscr{A}dA\in\Lambda(\mathbf{M};\mathbf{M})$, 使

$$\mathscr{L}mAX=AX,\mathscr{R}mAX=XA,\mathscr{A}dAX=AX-XA,\ \forall X\in\mathbf{M} \tag{0.6.21}$$

由此定义可知,对所有 $A,B\in\mathbf{M},\alpha\in\mathbf{C}$ 都有

$$\mathscr{L}m(A+B)=\mathscr{L}mA+\mathscr{L}mB,\mathscr{L}m(\alpha A)=\alpha\mathscr{L}mA,\mathscr{L}m(AB)=\mathscr{L}mA\mathscr{L}mB \tag{0.6.22}$$

$$\mathscr{R}m(A+B)=\mathscr{R}mA+\mathscr{R}mB,\mathscr{R}m(\alpha A)=\alpha\mathscr{R}mA,\mathscr{R}m(AB)=\mathscr{R}mB\mathscr{R}mA \tag{0.6.23}$$

$$\mathscr{A}dA=\mathscr{L}mA+\mathscr{R}m(-A),\ \mathscr{L}mA\mathscr{R}mB=\mathscr{R}mB\mathscr{L}mA \tag{0.6.24}$$

易知 $\mathscr{L}m\mathbf{M}\equiv\{\mathscr{L}mA\mid A\in\mathbf{M}\}$,$\mathscr{R}m\mathbf{M}\equiv\{\mathscr{R}mA\mid A\in\mathbf{M}\}$是 $\Lambda(\mathbf{M};\mathbf{M})$的子代数。式(0.6.22)表明是 $\mathscr{L}m:\mathbf{M}\to\Lambda(\mathbf{M};\mathbf{M})$是 $\mathbf{M}$ 到 $\Lambda(\mathbf{M};\mathbf{M})$的代数单同态。式(0.6.23)表明 $\mathscr{R}m:\mathbf{M}\to\Lambda(\mathbf{M};\mathbf{M})$是 $\mathbf{M}$ 到 $\Lambda(\mathbf{M};\mathbf{M})$的代数反单同态。

下面我们证明

命题:$\{\mathscr{L}m\mid\varphi_\alpha\rangle\langle\varphi_\rho\mid\mathscr{R}m\mid\varphi_\sigma\rangle\langle\varphi_\beta\|1\le\alpha,\beta,\sigma,\rho\le N\}$构成 $\Lambda(\mathbf{M};\mathbf{M})$的正交归一基。

事实上,$\{\mathscr{L}m\mid\varphi_\alpha\rangle\langle\varphi_\rho\mid\mathscr{R}m\mid\varphi_\sigma\rangle\langle\varphi_\beta\mid\}\mid\varphi_\mu\rangle\langle\varphi_\nu\mid$

$$=\mid\varphi_\alpha\rangle\langle\varphi_\rho\mid\varphi_\mu\rangle\langle\varphi_\nu\mid\varphi_\sigma\rangle\langle\varphi_\beta\mid=\mid\varphi_\alpha\rangle\langle\varphi_\beta\mid\delta_{\rho\mu}\delta_{\nu\sigma} \tag{0.6.25}$$

可见超算符 $\mathscr{L}m\mid\varphi_\alpha\rangle\langle\varphi_\rho\mid\mathscr{R}m\mid\varphi_\sigma\rangle\langle\varphi_\beta\mid$ 把 $\mathbf{M}$ 的基矢 $\mid\varphi_\rho\rangle\langle\varphi_\sigma\mid$ 变成 $\mathbf{M}$ 的基矢 $\mid\varphi_\alpha\rangle\langle\varphi_\beta\mid$,而把其他基矢变成 0。由本小节开始时介绍的命题,$\{\mathscr{L}m\mid\varphi_\alpha\rangle\langle\varphi_\rho\mid\mathscr{R}m\mid\varphi_\sigma\rangle\langle\varphi_\beta\|1\le\alpha,\beta,\sigma,\rho\le N\}$构成 $\Lambda(\mathbf{M};\mathbf{M})$的正交归一基。由这个命题可知,$\mathscr{L}m\mathbf{M}\cup\mathscr{R}m\mathbf{M}$ 代数地生成 $\Lambda(\mathbf{M};\mathbf{M})$。由于 $\mathscr{L}m(\mathscr{R}m)$是代数同态(代数反同态),这使得可以把超算符的运算,归为普通线性算符的运算。

Ⅳ.用超算符表达算符函数的微商的示例

(i)考察算符函数 $F(X)=X^3$,这是一个定义在整个 $\mathbf{M}$ 上的映射 $F:\mathbf{M}\to\mathbf{M}$

$$\begin{aligned}F(X+\Delta X)-F(X)=&X^2(\Delta X)+X(\Delta X)X+(\Delta X)X^2+X(\Delta X)^2\\&+(\Delta X)X(\Delta X)+(\Delta X)^2X+(\Delta X)^3\\=&(\mathscr{L}mX^2+\mathscr{L}mX\mathscr{R}mX+\mathscr{R}mX^2)\Delta X+o(\Delta X)\end{aligned} \tag{0.6.26}$$

可见

$$F'(X)=\mathscr{L}mX^2+\mathscr{L}mX\mathscr{R}mX+\mathscr{R}mX^2 \tag{0.6.27}$$

(ii)考察算符函数 $F(X)=AX^nB$,(A,B 是 S 上的两个给定的算符),这也是定义在整个 $\mathbf{M}$ 上的映射。

$$F'(X)=\mathscr{L}mA\mathscr{R}mB\Big(\sum_{j=0}^{n-1}\mathscr{L}mX^{n-1-j}\mathscr{R}mX^j\Big) \tag{0.6.28}$$

(iii) $\exp X \equiv e^X \equiv \sum_{n\geq 0} \frac{X^n}{n!}$也是定义在整个 **M** 上的算符函数。

$$(e^X)' = \sum_{n\geq 1} \frac{1}{n!} \sum_{j=0}^{n-1} \mathscr{L}mX^{n-1-j} \mathscr{R}mX^j \tag{0.6.29}$$

以上这些微商都是超算符。如果企图在 **M** 中寻求答案,恐怕永远不会找到正确的答案。

Ⅴ. 高阶微商与 *Taylor* 展开

定义 设 **U** 是 **M** 的开子集,$F:\mathbf{U}\to\mathbf{M}$ 是 **U** 上的可微函数。考察导函数 $F':\mathbf{U}\to\Lambda(\mathbf{M};\mathbf{M})$。$X\in\mathbf{U}$,若对所有满足条件 $X+\Delta X\in\mathbf{U}$ 的 ΔX 都有

$$F'(X+\Delta X) - F'(X) = \Psi(X)\Delta X + o_1(\Delta X)$$

其中 $\Psi(X)\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$与 ΔX 无关,而超算符 $o_1(\Delta X)$是 ΔX 的高阶小

$$\frac{\| o_1(\Delta X)\|}{\|\Delta X\|} \xrightarrow[\Delta X\to 0]{} 0$$

则称 F'在点 X 处可微(F 在点 X 处有二阶导数),称 $\Psi(X)$为 F 在点 X 处的二阶微商,记作 $\Psi(X)=F''(X)$,这样

$$F'(X+\Delta X) - F'(X) = F''(X)\Delta X + o_1(\Delta X) \tag{0.6.30}$$

若 F'在 **U** 的每个点处都可微,则称 F 是 **U** 上的两次可微的算符函数。注意,这时二次导函数

$$F'':\mathbf{U}\to\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M})) \tag{0.6.31}$$

与 F,F'是性质各不相同的映射。

以同样的方式,我们可以定义 $F:\mathbf{U}\to\mathbf{M}$ 的高阶导数。在点 $X\in\mathbf{U}$ 处,函数 F,函数 F 的一次导函数 F',两次导函数 F'',…的值分别是线性算符,超算符,超超算符,…

$$F(X)\in\mathbf{M}, F'(X)\in\Lambda(\mathbf{M};\mathbf{M}), F''(X)\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M})),\cdots \tag{0.6.32}$$

让我们考察

$$\Lambda(\mathbf{M};\Lambda(\mathbf{M};\cdots\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))\cdots)) \equiv [\Lambda(\mathbf{M};]^k\mathbf{M})^k \tag{0.6.33}$$

我们不妨称其元素为 k 次超算符。为表达这种 k 次超算符,让我们引入下面的符号(记法):设 $X_0, X_1, \cdots, X_n\in\mathbf{M}$,定义 $X_0\underline{1}X_1\in\Lambda(\mathbf{M};\mathbf{M})$,使 $X_0\underline{1}X_1(Y)=X_0YX_1\in\mathbf{M}$,

$\forall Y\in\mathbf{M}$。定义 $X_0\underline{1}X_1\underline{2}X_2, X_0\underline{2}X_1\underline{1}X_2\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$,使 $X_0\underline{1}X_1\underline{2}X_2(Y)=X_0\underline{1}X_1YX_2\in\Lambda(\mathbf{M};\mathbf{M})$, $X_0\underline{2}X_1\underline{1}X_2(Y)=X_0YX_1\underline{1}X_2\in\Lambda(\mathbf{M};\mathbf{M})$, $\forall Y\in\mathbf{M}$。一般地,定义

$$X_0\underline{1}X_1\underline{2}X_2\cdots\underline{k}X_k,\cdots,X_0\underline{P(1)}X_1\cdots\underline{k}X_{p^{-1}(k)}\cdots\underline{P(k)}X_k\cdots.\in)[\Lambda(\mathbf{M};]^k\mathbf{M})^k,\text{使}$$

$$X_0\underline{1}X_1\underline{2}X_2\cdots\underline{k}X_k(Y)=X_0\underline{1}X_1\underline{2}X_2\cdots\underline{k-1}X_{k-1}YX_k\in[\Lambda(\mathbf{M};]^{k-1}\mathbf{M})^{k-1}\cdots,$$

$$X_0\underline{P(1)}X_1\cdots\underline{k}X_{p^{-1}(k)}\cdots\underline{P(k)}X_k(Y)=X_0\underline{P(1)}X_1\cdots YX_{p^{-1}(k)}\cdots\underline{P(k)}X_k\in[\Lambda(\mathbf{M};]^{k-1}\mathbf{M})^{k-1},\cdots \tag{0.6.34}$$

上式中的 $P\in S_k$ 是 k – 次置换。

这种记法与 $\mathscr{L}m,\mathscr{R}m$ 的关系:

$$X_0\underline{1}X_1=\mathscr{L}mX_0\mathscr{R}mX_1\underline{1}\equiv\mathscr{L}mX_0\mathscr{R}mX_1\in\Lambda(\mathbf{M};\mathbf{M})$$

$$X_0\underline{1}X_1\underline{2}X_2=\mathscr{L}mX_0\mathscr{R}m(LmX_1RmX_2\underline{2})\underline{1}\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M})),\cdots \tag{0.6.35}$$

我们看到记法(0.6.34)要比 $\mathscr{L}m,\mathscr{R}m$ 更系统,在表达算符函数高阶导数时更方便。但用 $\mathscr{L}mA,\mathscr{R}mB$ 表达一阶导函数值(超算符)更接近通常算符的表达方式。

命题:

$\{|m\rangle\langle m_k|\underline{k}|n_k\rangle\langle m_{k-1}|\underline{k-1}|n_{k-1}\rangle\langle m_{k-2}|\cdots|n_3\rangle\langle m_2|\underline{2}|n_2\rangle\langle m_1|\underline{1}|n_1\rangle\langle n\|\underline{1}\leq m,n,m_1,n_1,m_2,n_2,\cdots m_k,n_k\leq N\}$构成$[\Lambda(\mathbf{M};]^k\mathbf{M})^k(k\geq1)$的正交归一基。其中$\{|m\rangle|1\leq m\leq N\}$是 S 的正交归一基。

证明 由 S 的正交归一基$\{|m\rangle|1\leq m\leq N\}$出发,可以构造 $\Lambda(S;S)\equiv\mathbf{M}$ 的正交归一基$\{|m\rangle\langle n\|1\leq m,n\leq N\}\equiv B_0$(参见前面Ⅲ小节开头的命题)。在$|m\rangle\langle n|$的中间插入$\langle m_1|\underline{1}|n_1\rangle$得$|m\rangle\langle m_1|\underline{1}|n_1\rangle\langle n|\in\Lambda(\mathbf{M};\mathbf{M})$; 它把 $\mathbf{M}$ 的正交归一基 B_0 中的基矢$|m_1\rangle\langle n_1|$变成 $\mathbf{M}$ 的基矢$|m\rangle\langle n|$,而把其它基矢消灭掉,故$\{|m\rangle\langle m_1|\underline{1}|n_1\rangle\langle n\|1\leq m,n,m_1,n_1\leq N\}\equiv B_1$ 是 $\Lambda(\mathbf{M};\mathbf{M})$的正交归一基。在$|m\rangle\langle m_1|\underline{1}|n_1\rangle\langle n|$的$|m\rangle$和$\langle m_1|$中间或$|n_1\rangle$和$\langle n|$中间插入$\langle m_2|\underline{2}|n_2\rangle$得$|m\rangle\langle m_2|\underline{2}|n_2\rangle\langle m_1|\underline{1}|n_1\rangle\langle n|\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$ 或$|m\rangle\langle m_1|\underline{1}|n_1\rangle\langle m_2|\underline{2}|n_2\rangle\langle n|\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$。它们都把$|m_2\rangle\langle n_2\in B_0$变成$|m\rangle\langle m_1|\underline{1}|n_1\rangle\langle n|\in B_1$,故$\{|m\rangle\langle m_2|\underline{2}|n_2\rangle\langle m_1|\underline{1}|n_1\rangle\langle n\|1\leq m,n,m_1,n_1,m_2,n_2\leq N\}$构成 $\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$的正交归一基。依次类推,我们可以确信命题成立。

现在让我们考察如何用各次超算符表达算符函数的高阶微商值。

例 $F(X)=X^4$。F 定义在整个 $\mathbf{M}$ 上，而且在 $\mathbf{M}$ 上有任意阶导数。

$$F'(X)=\mathscr{L}mX^3+\mathscr{L}mX^2\mathscr{R}mX+\mathscr{L}mX\mathscr{R}mX^2+\mathscr{R}mx^3=X^3\underline{1}+X^2\underline{1}X+X\underline{1}X^2+\underline{1}X^3\in\Lambda(\mathbf{M};\mathbf{M})$$

$$F''(X)=X^2\underline{2}\,\underline{1}+X\underline{2}X\underline{1}+\underline{2}X^2\underline{1}+X^2\underline{1}\,\underline{2}+X\underline{2}\,\underline{1}X+\underline{2}X\underline{1}X+X\underline{1}X\underline{2}+X\underline{1}\,\underline{2}X+\underline{2}\,\underline{1}X^2+\underline{1}X^2\underline{2}+\underline{1}X\underline{2}X+\underline{1}\,\underline{2}X^2\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$$

$$F'''(X)=X\underline{3}\,\underline{2}\,\underline{1}+\underline{3}X\underline{2}\,\underline{1}+X\underline{2}\,\underline{3}\,\underline{1}+\underline{3}\,\underline{2}X\underline{1}+\cdots+\underline{1}\,\underline{2}X\underline{3}+\underline{1}\,\underline{2}\,\underline{3}X\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M})))$$

$$F''''(X)=\sum_{P\in S_4}\underline{P(1)}\,\underline{P(2)}\,\underline{P(3)}\,\underline{P(4)}\in[\Lambda(\mathbf{M};]^4\mathbf{M})^4$$

$$F^5(X)=0\in[\Lambda(\mathbf{M};]^5\mathbf{M})^5$$

注意$\underline{1}\,\underline{2}\,\underline{3}\,\underline{4}\cong 1_s\,\underline{1}1_s\,\underline{2}1_s\,\underline{3}1_s\,\underline{4}1_s$，其中 $1_s\in\mathbf{M}$ 是 S 上的恒等算符($\mathbf{M}$ 中乘法的单位元)。对算符函数也有 *Taylor* 展开。

$$F(X)=F(0)+\frac{1}{1!}F'(0)X+\frac{1}{2!}(F''(0)X)X+\frac{1}{3!}[(F'''(0)X)X]X+\cdots \tag{0.6.36}$$

与普通函数的 *Taylor* 展开不同，展开“系数” $F(0)\in\mathbf{M}$，$F'(0)\in\Lambda(\mathbf{M};\mathbf{M})$，$F''(0)\in\Lambda(\mathbf{M};\Lambda(\mathbf{M};\mathbf{M}))$，…一般不是数，而是算符和各种超算符。所以展开系数为数的 X 的幂级数只是“解析的”算符函数的一小部分。连很简单的算符函数 $F(X)=AX$(A 是常算符，但不是恒等算符的纯量倍)都不能写作以复数为系数的 X 的幂级数。AX 本身就是它的 *Taylor* 展开。

Ⅵ. 超算符应用示例

(i) §0.4 中的引理 1：设 A，B 是 n 维线性空间上的线性算符，则 $e^ABe^{-A}=e^{\mathscr{A}dA}B$。在那里我们用了一个很巧妙的方法证明了它。现在利用超算符，给它另一个证明。

$$e^ABe^{-A}=\mathscr{L}me^A\mathscr{R}me^{-A}B=e^{\mathscr{L}mA}e^{\mathscr{R}m(-A)}B=e^{\mathscr{L}mA+\mathscr{R}m(-A)}B=e^{\mathscr{A}dA}B, \tag{0.6.37}$$

最后第二步利用了 $\mathscr{L}mA$，$\mathscr{R}m(-A)$ 的可对易性。可以看出，利用超算符概念，证明要简单直接。

(ii)利用超算符概念，我们还容易的得出以前没有的结果。例如，设 A，B，C 是 n

维线性空间 V 上的线性算符,则

$$e^{A}Be^{C} = e^{\mathscr{L}mA+\mathscr{R}mC}B \tag{0.6.38}$$

这一结果,在讨论李群与李代数的关系时非常有用(见附录)。

(iii)大家都知道,若 A 是 n 维线性空间 V 上的线性算符,$x(t)\in V$ 满足微分方程 $\dot{x}(t)=Ax(t)$,则有

$$x(t) = e^{At}x(0) \tag{0.6.39}$$

在 *Heisenberg* 绘景下,力学量算符满足 *Heisenberg* 方程

$$\dot{A}(t) = \frac{i}{\hbar}[H,A(t)] \tag{0.6.40}$$

利用超算符概念,这个方程可以重新写作

$$\dot{A}(t) = \mathscr{A}d\left(\frac{i}{\hbar}H\right)A(t) \tag{0.6.41}$$

由(0.6.39)可以马上写下

$$A(t) = e^{\mathscr{A}d\left(\frac{i}{\hbar}H\right)t}A(0) \tag{0.6.42}$$

我们希望超算符与非对易分析概念的建立,会对读者以后的学习与科研有所帮助。

第 1 章 量子力学的基本原理

读者在大学本科的量子力学课程中，已学习过量子力学发生和发展的历史，并对量子概念和量子力学的基本理论和应用有了一个初步的认识。我们将从重述量子力学的基本原理出发，展开我们的讨论。

§1.1 量子力学的基本假定

量子力学的理论体系是建立在几个基本假设(公理)之上的。

Ⅰ. 波函数公设

假设Ⅰ. 对于每个物理体系，定义在该体系位形空间(configuration space)上的模平方可积的复值函数的全体构成一个 Hilbert 空间。量子力学假定，体系的每一个状态都由该 Hilbert 空间中的一个非零向量来描写。反之，该 Hilbert 空间中每一个非零向量都描写体系的一个状态。我们称此 Hilbert 空间为该体系的态矢量空间，或简称为体系的态空间。

上面所说的体系的位形空间是指其广义坐标的取值范围。

例 1 由一个自旋为零的粒子所构成的体系，其位形空间就是通常的三维实欧氏空间 R^3。它的态矢量空间是由定义在 R^3 上的模平方可积的复值函数 $\psi\left(\int_{R^3}\mathrm{d}x\mathrm{d}y\mathrm{d}z\bar{\psi}(\boldsymbol{r})\psi(\boldsymbol{r})\leq\infty\right)$ 的全体所构成的。两个态矢量的内积由下式定义 $\langle\varphi\mid\psi\rangle=\int_{R^3}\mathrm{d}x\mathrm{d}y\mathrm{d}z\,\bar{\varphi}(\boldsymbol{r})\psi(\boldsymbol{r})$。

例 2 由一个电子$\left(\text{其自旋为}\frac{1}{2}\right)$所构成的体系，其位形空间就是三维实欧氏空间 R^3 与集合$\left\{\frac{\hbar}{2}, \frac{-\hbar}{2}\right\}$的笛卡儿积(参阅 §0.0)。它的态矢量空间是由满足条件 $\sum_{s_z=\frac{-\hbar}{2}}^{\frac{\hbar}{2}}\int_{R^3}\mathrm{d}x\mathrm{d}y\mathrm{d}z\mid\varphi(\boldsymbol{r},s_z)\mid^2\leq\infty$ 的 φ 所构成的。两个态矢量的内积由下式定义 $\langle\varphi\mid\psi\rangle=\sum_{s_z=\frac{-\hbar}{2}}^{\frac{\hbar}{2}}\int_{R^3}\mathrm{d}x\mathrm{d}y\mathrm{d}z\,\overline{\varphi}(\boldsymbol{r},s_z)\psi(\boldsymbol{r},s_z)$。

例 3 一个氢原子是由一个电子和一个质子构成的体系。它的位形空间是电子的位形空间与质子的位形空间的笛卡儿积。它的态矢量空间是由满足下列条件的 $\Phi(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})$所构成的：$\sum_{s_{ze}=\frac{-\hbar}{2}}^{\frac{\hbar}{2}}\int_{R^3}\mathrm{d}x_e\mathrm{d}y_e\mathrm{d}z_e\sum_{s_{zp}=\frac{-\hbar}{2}}^{\frac{\hbar}{2}}\int_{R^3}\mathrm{d}x_p\mathrm{d}y_p\mathrm{d}z_p\mid\Phi(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})\mid^2\leq\infty$。氢原子的态矢量空间是电子的态矢量空间与质子的态矢量空间的张量积(参阅 §0.5)。两个态矢量的内积由下式定义

$$\langle\Phi\mid\Psi\rangle=\sum_{s_{ze}=\frac{-\hbar}{2}}^{\frac{\hbar}{2}}\int_{R^3}\mathrm{d}x_e\mathrm{d}y_e\mathrm{d}z_e\sum_{s_{zp}=\frac{-\hbar}{2}}^{\frac{\hbar}{2}}\int_{R^3}\mathrm{d}x_p\mathrm{d}y_p\mathrm{d}z_p\overline{\Phi}(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})\Psi(\boldsymbol{r}_e,s_{ze},\boldsymbol{r}_p,s_{zp})。$$

例 4 绕固定轴转动的刚体的位形空间是一个圆周，即闭区间$[-\pi,\pi]$，但两端点应视为同一点。量子化后，它的态矢量空间是由定义在$[-\pi,\pi]$上的连续函数 $\Phi(\vartheta)\{\Phi(-\pi)=\Phi(\pi)\}$所构成的。两个态矢量的内积由下式定义：$\langle\Phi\mid\Psi\rangle=\int_{-\pi}^{\pi}\mathrm{d}\vartheta\,\overline{\Phi}(\vartheta)\Psi(\vartheta)$。

注意，两端点视为同一点的闭区间$[-\pi,\pi]$正好是二维空间转动群 $SO(2)$的群流形。

* **例 5** 绕固定点转动的刚体的位形空间是三维空间转动群 $SO(3)$的群流形，也就是三维空间中的一个实心球，其半径为 π，但同一个直径的两个端点要看成同一个点。量子化后，它的态矢量空间是由定义在这个位形空间上的连续函数所构成的。若采用欧拉角(α,β,γ)(即先绕 z 轴转 γ 角，再绕 y 轴转 β 角，最后绕 z 轴转 α 角)为广义坐标，欧拉角的取值范围是$(-\pi\leq\alpha\leq\pi,0\leq\beta\leq\pi,-\pi\leq\gamma\leq\pi)$，则体系的态空间

由欧拉角的连续函数

$$\Psi(\alpha,\beta,\gamma)\{\Psi(-\pi,\beta,\gamma)=\Psi(\pi,\beta,\gamma),\Psi(\alpha,0,\gamma)=\Psi(\alpha-\Delta,0,\gamma+\Delta),$$
$$\Psi(\alpha,\pi,\gamma)=\Psi(\alpha+\Delta,\pi,\gamma+\Delta),\Psi(\alpha,\beta,\pi)=\Psi(\alpha,\beta,-\pi)\}$$

的全体所构成。两个态矢量的内积由下式定义

$$\langle\Phi\mid\Psi\rangle=\int_{-\pi}^{\pi}\mathrm{d}\alpha\int_{0}^{\pi}\mathrm{d}\beta\int_{-\pi}^{\pi}\mathrm{d}\gamma\sin\beta\overline{\Phi}(\alpha,\beta,\gamma)\Psi(\alpha,\beta,\gamma)$$

上式中的 $\sin\beta\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma$ 是 $SO(3)$群的群流形的不变体积元。

若一个量子体系是由两个较小的量子体系 A 和 B 所构成的，则体系 $A+B$ 的位形空间是其两个子体系的位形空间的笛卡尔积：$C_{A+B}=C_A\times C_B$，体系 $A+B$ 的态矢量空间是其两个子体系的态矢量空间的张量积：$S_{A+B}=S_A\otimes S_B$（参阅 §0.5）。体系 $A+B$ 的态矢量，当 $\dim A>1$ 且 $\dim B>1$ 时，不一定能够写作一个体系 A 的态矢量与一个体系 B 的态矢量的张量积（请读者证明之），这就是平时所说的，体系 $A+B$ 的波函数不一定能够分离变量。不能够分离变量的波函数所描述的的态称作体系 $A+B$ 的纠缠态（entangled state）；而可分离变量的波函数，即子体系 A 的态矢量与子体系 B 的态矢量的张量积，所描述的体系 $A+B$ 的态叫做非纠缠态。纠缠态的概念在量子信息学和量子测量理论中扮演着一个非常重要的角色。

Ⅱ. 算符公设

假设Ⅱ. 微观体系的每一个力学量 A，都由该体系态空间上的一个厄米算符（记作 $\hat{A}$，或者当不会引起误解时，也记作 A）来描写。特别是，体系广义坐标的算符为 $\hat{q}_j=q_j$，相应的广义动量的算符为 $\hat{p}_j=-i\hbar\dfrac{\partial}{\partial q_j}$。体系的任一力学量 $F(q,p)$的算符由下式确定 $\hat{F}(q,p)=F(\hat{q},\hat{p})$（$F(q,p,t)$在不同时刻可能是体系的不同的力学量，由在不同时刻可能是不同的厄米算符 $F(\hat{q},\hat{p},t)$描写）。

由上面 $\hat{q}_j,\hat{p}_k$ 的具体形式可知，$\hat{p}_j\hat{q}_k=\hat{q}_k\hat{p}_j-i\hbar\delta_{jk}$，即算符 $\hat{q}_j,\hat{p}_j$ 不对易。而在经典力学中，总是有 $p_jq_k=q_kp_j$。这使得对经典力学量而言相同的表达式，例如 $x^3p_x=p_xx^3=\dfrac{1}{2}(x^3p_x+p_xx^3)$，代换成量子力学算符后，等式可能不复成立。从而，如何从一个力学量的经典表达式 $F(q,p)$，写下它的量子力学算符 $\hat{F}(q,p)=F(\hat{q},\hat{p})$，方式不是唯一的。即使加上厄米性要求 $\hat{F}(p,q)^+=\hat{F}(p,q)$，方式仍可能不唯一。哪一种

写法正确，最后只能由实验判断。详细的讨论，有兴趣的读者可参考喀兴林教授所著的《高等量子力学》(北京，高教出版社，1999)。

Ⅲ. 力学量平均值公设

假设Ⅲ. 当微观体系处于状态 ψ 时，测量其任一力学量 A，一般而言，测得的值不是确定的。每次测量 A 得到的可能值是 $\hat{A}$ 的特征值。对处于状态 ψ 的体系，多次测量 A 的平均值为

$$\langle A\rangle = \frac{\langle \psi \mid \hat{A} \mid \psi\rangle}{\langle \psi \mid \psi\rangle} \tag{1.1.1}$$

由此可见，两个只差一复数倍的非零态矢量，描写微观体系的同一个状态。为方便起见，我们常常采用归一化的态矢量 ψ，即满足条件 $\langle \psi \mid \psi\rangle = 1$ 的 ψ，来描述体系的状态。此时有 $\langle A\rangle = \langle \psi \mid \hat{A} \mid \psi\rangle$。

$\hat{A}$ 的厄米性保证了测量力学量 A 的可能值和平均值是实数。$\hat{A}$ 的厄米性还保证了它有完备的正交归一化的特征向量组：$\hat{A}u_n = a_n u_n$，$\langle u_m \mid u_n\rangle = \delta_{mn}$，任一态矢量 ψ 都可写作 $\hat{A}$ 的特征向量的线性组合 $\psi = \sum_n u_n \langle u_n \mid \psi\rangle$。从而，我们可以得到特征向量组的完备性关系 $\sum_n \mid u_n\rangle\langle u_n \mid = 1$。对归一化的 ψ 有 $\langle A\rangle = \sum_n a_n \mid \langle u_n \mid \psi\rangle \mid^2$。可见 $\mid \langle u_n \mid \psi\rangle \mid^2$ 是测得 A 为可能值 a_n 的几率。从而当且仅当体系处在 $\hat{A}$ 的特征态 u_n，测量 A 得确定值 a_n。(作为练习，请读者证明以上各断言)

在下面一小段的讨论中，我们将沿用上一小段的各个记号并以 S 记体系的态矢量空间。为了刻画对处于某个态 ψ 下的体系测量力学量 A 的不确定性，我们引入均方差

$$\begin{aligned}\langle (\Delta A)^2\rangle &= \langle \psi \mid (\hat{A} - \langle A\rangle)^2 \mid \psi\rangle = \langle \psi \mid \hat{A}^2 \mid \psi\rangle - \langle A\rangle^2 \\ &= \langle (\hat{A} - \langle A\rangle)\psi \mid (\hat{A} - \langle A\rangle)\psi\rangle\end{aligned}$$

易见，当且仅当 $(\hat{A} - \langle A\rangle)\psi\rangle = 0$，均方差为零即在态 ψ 下力学量 A 取确定值。我们又得到了前面的断言：当且仅当体系处在 $\hat{A}$ 的特征态 u_n，测量 A 得确定值 a_n。

如果对处在态 $\psi = \sum_n u_n \langle u_n \mid \psi\rangle$ 下的体系，测量其力学量 A，测得的值是算符 $\hat{A}$ 的某个特征值 a_m。紧接着再对该体系做一次力学量 A 的测量(两次测量相隔无穷小瞬间)，则测得的值必然仍是 a_m。这表明第一次测量使体系的态由 ψ 突变为 u_m。这种测量导致的态矢量的突变称之为波包塌缩(wave packet collapse) 或波函数约化

(wave function reduction)。当力学量 A 的特征值 a_m 是简并的，$Au_{mi}=a_m u_{mi}$，$\langle u_{mi},u_{mj}\rangle=\delta_{ij}$，$\forall\,1\le i,j\le d_m>1$，若测量其力学量 A，测得的值是算符 $\hat{A}$ 的特征值 a_m，则波函数约化作 $\sum_{i=1}^{d_m}u_{mi}\langle u_{mi}|\psi\rangle$。与非简并情况不同，约化后的波函数有某种不确定性。

前面提到，当且仅当体系处在 $\hat{A}$ 的特征态 u_n，测量 A 得确定值 α_n。测量体系力学量 B 得确定值 β_m，要求体系处在 $\hat{B}$ 的属于特征值 β_m 的特征态 v_m。因此 A,B 同时取确定值要求体系处在 $\hat{A},\hat{B}$ 共同的特征态。而厄米算符 $\hat{A},\hat{B}$ 有完备的共同的特征向量组的充要条件是 $\hat{A}\hat{B}=\hat{B}\hat{A}$（参考 §0.3）。这里要提醒注意的是，当 $\hat{A},\hat{B}$ 不对易时，它们仍可有共同的特征向量。譬如 $Y_{00}(\vartheta,\varphi)$ 是 $\hat{l}_z,\hat{l}_x$ 的共同的特征向量，尽管它们不对易。而当 $\hat{A},\hat{B}$ 对易时，并非 $\hat{A}$ 的特征向量就一定是 $\hat{B}$ 的特征向量。譬如 $Y_{00}(\vartheta,\varphi)+Y_{10}(\vartheta,\varphi)$ 是 $\hat{l}_z$ 的特征向量，而非 $\hat{l}^2$ 的特征向量，尽管它们对易。

在 §0.3 我们证明过一组厄米算符 $\{\hat{A},\hat{B},\cdots,\hat{G}\}$ 有完备的共同的特征向量组的充要条件是它们两两对易。这也就是我们所说的一组力学量可以同时测量的充要条件。当同时测量这样一组力学量 $\{A,B,\cdots,G\}$，体系的态矢量被垂直地投影到对应于测量值 $\{a,b,\cdots,g\}$ 的公共的特征子空间 $\{u\in S\,|\,\hat{A}u=au,\hat{B}u=bu,\cdots,\hat{G}u=gu\}$。特别是，当这组力学量的算符的公共的特征子空间全是一维的，则塌缩后的波函数由各个力学量的测量值所确定。我们称这种力学量组为完备的力学量组。例如 $\{p_x,p_y,p_z\}$，是单个无旋粒子体系的完备力学量组，但不是单个电子体系的完备力学量组。$\left\{\hat{T}-\frac{e^2}{r},\hat{l}^2,\hat{l}_z,\hat{s}_z\right\}$ 是单个电子体系的完备力学量组。

让我们给测不准关系一个定量的描述：在任意态 ψ 下测量力学量 A,B 的结果必须满足下面的不等式：$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle\ge\frac{1}{4}|\langle[\hat{A},\hat{B}]_-\rangle|^2$。事实上，利用关于内积的 Schwarz 不等式 $\langle x|x\rangle\langle y|y\rangle\ge|\langle x|y\rangle|^2$，$\forall\,x,y\in S$，我们有

$$\langle(\Delta\hat{A})\psi\,|\,(\Delta\hat{A})\psi\rangle\langle(\Delta\hat{B})\psi\,|\,(\Delta\hat{B})\psi\rangle\ge|\langle(\Delta\hat{A})\psi\,|\,(\Delta\hat{B})\psi\rangle|^2$$

上式左边即 $\langle(\Delta\hat{A})^2\rangle\langle(\Delta\hat{B})^2\rangle$，而右边等于 $|\langle\psi|(\Delta\hat{A})(\Delta\hat{B})\psi|^2=|\langle\psi|\hat{A}\hat{B}|\psi\rangle-\langle A\rangle\langle B\rangle|^2$。把 $\hat{A}\hat{B}$ 写作反厄米部分与厄米部分之和 $\hat{A}\hat{B}=\frac{1}{2}[\hat{A},\hat{B}]_-+\frac{1}{2}\{\hat{A},\hat{B}\}_+$，代入上式得 $|\langle\psi|\hat{A}\hat{B}|\psi\rangle-\langle A\rangle\langle B\rangle|^2=\left|\langle\psi|\frac{1}{2}[\hat{A},\hat{B}]_-|\psi\rangle+\langle\psi|\frac{1}{2}\{\hat{A},\hat{B}\}_+|\psi\rangle-\langle A\rangle\langle B\rangle\right|^2\ge\left|\langle\psi|\frac{1}{2}[\hat{A},\hat{B}]_-|\psi\rangle\right|^2$ 最后一步利用了复数性质：$|\alpha+i\beta|\ge|\beta|$。于是我

们证明了测不准关系。当 $A=q, B=p$ 时，我们得到$\langle(\Delta q)^2\rangle\langle(\Delta p)^2\rangle \geq \frac{1}{4}\hbar^2$。

Ⅳ. 动力学演化方程

假设Ⅳ. 微观体系状态随时间的变化遵守 Schrödinger 方程

$$i\hbar\frac{\partial}{\partial t}\psi(t) = \hat{H}(t)\psi(t) \tag{1.1.2}$$

式中的 $\hat{H}(t)$是 t 时刻体系的哈密顿量的算符。

Schrödinger 方程的未知函数是态矢量空间中的随时间 t 变化的矢量。Schrödinger 方程是关于自变量 t 的一阶微分方程。它的解由其初始值 $\psi(t_0)$唯一地确定。从物理上讲，这表示，在给定的外界条件下，每一体系初始时刻的状态唯一地确定它在以后各时刻的状态。而这正是因果律所要求的。

Schrödinger 方程中出现的哈密顿算符 $\hat{H}(t)$是线性算符。这使得方程的解的任一常系数

线性组合仍是方程的解。而这正是态的迭加原理所要求的。$\hat{H}(t)$的厄米性，导致 Schrödinger 方程的解 $\psi(t), \varphi(t)$的内积不随时间变化：$\langle\psi(t)|\varphi(t)\rangle = \langle\psi(t_0)|\varphi(t_0)\rangle$。特别是，当 $\psi(t)=\varphi(t)$时，$\langle\psi(t)|\varphi(t)\rangle = \langle\psi(t_0)|\varphi(t_0)\rangle$。这代表总几率守恒，在初始时刻 t_0 归一化的态矢量，到任一时刻都保持归一化。（作为练习，请读者证明以上各断言）

当体系的某个力学量 F 不随时间变化，并与体系的哈密顿算符 $H(t)$对易时，即当$\frac{\partial F}{\partial t}=0, [F, H(t)]=0$ 时，称 F 为运动积分（F 不显含时间，而 $H(t)$可以显含时间）。由 Schrödinger 方程可推出，运动积分 F 有如下性质：(1)运动积分 F 的平均值不随时间变化；(2)若初始时刻体系处在 F 的属于本征值 f 的本征态，则它在任一时刻都处在 F 的属于同一本征值 f 的本征态；(3)运动积分 F 取各个可能值的几率不随时间变化。

事实上，(1)$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|F|\psi(t)\rangle = \langle\dot{\psi}(t)|F|\psi(t)\rangle + \langle\psi(t)|\frac{\partial}{\partial t}F|\psi(t)\rangle + + \langle\psi(t)|F|\dot{\psi}(t)\rangle = \frac{-1}{i\hbar}\langle H(t)\psi(t)|F|\psi(t)\rangle + \frac{1}{i\hbar}\langle\psi(t)|F|H(t)\psi(t)\rangle = \frac{1}{i\hbar}\langle\psi(t)|[F, H(t)]|\psi(t)\rangle = 0$。(2)以 F 作用于 Schrödinger 方程两边得，$i\hbar\frac{\partial}{\partial t}\{F\psi(t)\} = H$

$(t)\{F\psi(t)\}$。这表明若 $\psi(t)$是 Schrödinger 方程的解,则 $F\psi(t)$也是 Schrödinger 方程的解。我们知道,若 $\psi(t)$是 Schrödinger 方程的解,则 $f\psi(t)$也是 Schrödinger 方程的解。如果 $F\psi(t_0)=f\psi(t_0)$,即解 $F\psi(t)$,$f\psi(t)$满足同样的初始条件,则应有 $F\psi(t)=f\psi(t)$。(3)设 $Fu_n=f_nu_n$,$\langle u_m\mid u_n\rangle=\delta_{mn}$,$\{u_n\mid n\geq 0\}$是 F 的一组完备的正交归一化的特征向量组。以 $u_n(t)$记 Schrödinger 方程的满足初始条件 $u_n(t_0)=u_n$ 的解,由上面的讨论可知,对任给的时刻 t,$Fu_n(t)=f_nu_n(t)$,$\langle u_m(t)\mid u_n(t)\rangle=\delta_{mn}$,$\{u_n(t)\mid n\geq 0\}$是 F 的一组完备的正交归一化的特征向量组。Schrödinger 方程的满足初始条件 $\psi(t_0)=\sum\limits_{n\geq 0}c_nu_n$ 解一定是 $\psi(t)=\sum\limits_{n\geq 0}c_nu_n(t)$。可见,对态 $\psi(t_0)$测量 F,得到可能值 f_n 的几率为 $|c_n|^2$,对态 $\psi(t)$测量 F,得到可能值 f_n 的几率也是 $|c_n|^2$。

Schrödinger 方程在量子力学中的地位恰如牛顿第二定律在经典力学中的地位。研究微观物理体系的动力学性质,就要求解 Schrödinger 方程。

一类重要的特殊情况是,体系的哈密顿量不随时间变化。孤立体系,或者处于恒定外场中的体系的哈密顿算符就不依赖于时间。这时,求解 Schrödinger 方程的问题,可归为求解哈密顿算符的特征方程 $\hat{H}u=\varepsilon u$。事实上,直接验证可知,$\psi(t)=ue^{-i\frac{\varepsilon}{\hbar}t}$ 是 Schrödinger 方程的解。在这种态中,体系的任一力学量的测量结果,都与时间无关。我们称之为定态。不显含时间的哈密顿算符的特征方程也被称作定态 Schrödinger 方程。设$\{u_1,u_2,\cdots\}$是 $\hat{H}$ 的一组完备的正交归一化的特征向量组,分别属于特征值$\{\varepsilon_1,\varepsilon_2,\cdots\}$,则 Schrödinger 方程的任意一个解都可以写作各定态解的迭加:$\psi(t)=\sum\limits_{n\geq 1}c_nu_ne^{-i\frac{\varepsilon_n}{\hbar}t}$,其中系数 c_n 由初值确定:$c_n=\langle u_n\mid\psi(0)\rangle$,$\forall\, n\geq 1$。

Ⅴ.全同性原理

假设Ⅴ. 描写微观体系状态的波函数的自变量是构成体系的各个粒子的坐标和自旋。当交换体系中的两个全同粒子的坐标和自旋时,体系的波函数不变(交换一对全同玻色子时)或变号(交换一对全同费米子时)。这就是全同粒子不可分辨原理。

在相对论性量子理论中可以证明,自旋为整数的粒子为玻色子,自旋为半整数的粒子为费米子。

§1.2 演化算符

Ⅰ. 演化算符的引入和一般性质

由上面对 Schrödinger 方程一般性质的讨论可知,对任给的两个时刻 t_0、t,体系在 t_0 时刻的态矢量 $\psi(t_0)$,唯一地确定体系在 t 时刻的态矢量 $\psi(t)$。从而我们可以认为,态矢量 $\psi(t)$是某个算符 $U(t,t_0): S\to S$(S 是体系的态空间)作用在态矢量 $\psi(t_0)$上得到的:

$$\psi(t) = U(t,t_0)\psi(t_0) \tag{1.2.1}$$

我们称 $U(t,t_0)$为演化算符。当 $t=t_0$ 时,显然有 $\psi(t_0)=U(t_0,t_0)\psi(t_0)$, $\forall\,\psi(t_0)\in S$。从而

$$U(t_0,t_0) = 1 \tag{1.2.2}$$

上面的讨论对任意两个时刻 t_0,t 都成立。因而,对时刻 t,t_1 我们有 $\psi(t_1)=U(t_1,t)\psi(t)=U(t_1,t)U(t,t_0)\psi(t_0)=U(t_1,t_0)\psi(t_0)$, $\forall\,\psi(t_0)\in V$。从而

$$U(t_1,t)U(t,t_0) = U(t_1,t_0) \tag{1.2.3}$$

当 $t_1=t_0$ 时,上式变作 $U(t_0,t)U(t,t_0)=U(t_0,t_0)=1$。对调 t_0,t 位置得 $U(t,t_0)U(t_0,t)=U(t,t)=1$。可见

$$U(t,t_0) = U(t_0,t)^{-1} \tag{1.2.4}$$

由于 Schrödinger 方程任意多个解的常系数线性组合仍是它的解,这使得 $U(t,t_0)$是线性算符。由于 Schrödinger 方程任意两个解的内积与时间无关:

$$\langle U(t,t_0)\phi(t_0)\,|\,U(t,t_0)\psi(t_0)\rangle=\langle\phi(t_0)\,|\,\psi(t_0)\rangle,\ \forall\,\phi(t_0),\psi(t_0)\in V,$$

我们可以推出

$$U(t,t_0)^{+}\,U(t,t_0) = 1 \tag{1.2.5}$$

由于体系的态空间不一定是有限维的,单由上式还不能断言 $U(t,t_0)$是酉算符。但式(1.2.4)告诉我们 $U(t,t_0)$有逆算符。在这个条件下,它的任一左逆或右逆算符都等于这个逆算符。于是

$$U(t,t_0)^+ = U(t,t_0)^{-1} \tag{1.2.6}$$

这表明 $U(t,t_0)$是酉算符(幺正算符)。

Ⅱ. 演化算符的方程、形式解

把演化算符的定义式(1.2.1)代入 Schrödinger 方程

$$i\hbar\frac{\partial}{\partial t}U(t,t_0)\psi(t_0) = \hat{H}(t)U(t,t_0)\psi(t_0), \forall\,\psi(t_0)\in V$$

从而得到演化算符所满足的微分方程

$$i\hbar\frac{\partial}{\partial t}U(t,t_0) = \hat{H}(t)U(t,t_0) \tag{1.2.7}$$

演化算符就是上述微分方程的满足初始条件

$$U(t_0,t_0) = 1 \tag{1.2.8}$$

的解。

一类重要的特殊情况是哈密顿算符不依赖时间的情况。此时方程(1.2.7)变作

$$\frac{\partial}{\partial t}U(t,t_0) = \frac{\hat{H}}{i\hbar}U(t,t_0) \tag{1.2.9}$$

令 $V(t)=\exp\left(\frac{i\hat{H}}{\hbar}(t-t_0)\right)U(t,t_0)$。易见,$V(t_0)=1$。考察 $V(t)$满足的微分方程。不难看到 $\dot{V}(t)=\frac{i\hat{H}}{h}V(t)+\frac{\hat{H}}{i\hbar}V(t)=0$,从而 $V(t)=V(t_0)=1$。我们得到,当哈密顿算符不含时间时,演化算符为

$$U(t,t_0) = \exp\left(\frac{\hat{H}}{i\hbar}(t-t_0)\right) \tag{1.2.10}$$

在最一般的情况下,哈密顿算符可以随时间变化。此时为求解演化算符,我们对微分方程(1.2.7)积分,从 t_0 积到 t,并考虑到初值条件(1.2.8)可得

$$U(t,t_0) = 1+\int_{t_0}^{t}\mathrm{d}t_1\,\frac{H(t_1)}{i\hbar}U(t_1,t_0) \tag{1.2.11}$$

这是一个关于演化算符 $U(t,t_0)$的积分方程,它与微分方程(1.2.7)的初值问题(1.2.8)等价。

我们将用所谓逐次迭代法(iteration)来求解此积分方程,反复地把出现在积分号下的未知算符用表达式(1.2.11)代入。

$$
\begin{aligned}
U(t,t_0) &= 1+\int_{t_0}^{t}\mathrm{d}t_1\,\frac{H(t_1)}{i\hbar}\Big[1+\int_{t_0}^{t_1}\mathrm{d}t_2\,\frac{H(t_2)}{i\hbar}U(t_2,t_0)\Big]\\
&= 1+\int_{t_0}^{t}\mathrm{d}t_1\,\frac{H(t_1)}{i\hbar}+\int_{t_0}^{t}\mathrm{d}t_1\int_{t_0}^{t_1}\mathrm{d}t_2\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}\\
&\quad+\int_{t_0}^{t}\mathrm{d}t_1\int_{t_0}^{t_1}\mathrm{d}t_2\int_{t_0}^{t_2}\mathrm{d}t_3\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}\,\frac{H(t_3)}{i\hbar}+\cdots\\
&\quad+\int_{t_0}^{t}\mathrm{d}t_1\int_{t_0}^{t_1}\mathrm{d}t_2\cdots\int_{t_0}^{t_{n-1}}\mathrm{d}t_n\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}\cdots\frac{H(t_n)}{i\hbar}+\cdots
\end{aligned}
\tag{1.2.12}
$$

利用亥维赛阶梯函数 $\vartheta(t)=\begin{cases}1, & if\quad t\geq 0\\ 0, & if\quad t<0\end{cases}$，上式右端的第三项可以写作

$$
\int_{t_0\leq t_2\leq t_1\leq t}\mathrm{d}t_1\mathrm{d}t_2\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}
$$

$$
=\int_{t_0}^{t}\int_{t_0}^{t}\mathrm{d}t_1\mathrm{d}t_2\,\vartheta(t_1-t_2)\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}
$$

$$
=\frac{1}{2}\int_{t_0}^{t}\int_{t_0}^{t}\mathrm{d}t_1\mathrm{d}t_2\Big[\vartheta(t_1-t_2)\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}+\vartheta(t_2-t_1)\,\frac{H(t_2)}{i\hbar}\,\frac{H(t_1)}{i\hbar}\Big]
$$

而上式右端的一般项可以写作

$$
\int_{t_0\leq t_n\leq t_{n-1}\cdots\leq t_1\leq t}\mathrm{d}t_1\cdots\mathrm{d}t_n\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}\cdots\frac{H(t_n)}{i\hbar}
$$

$$
=\int_{t_0}^{t}\cdots\int_{t_0}^{t}\mathrm{d}t_1\cdots\mathrm{d}t_n\vartheta(t_1-t_2)\vartheta(t_2-t_3)\cdots\vartheta(t_{n-1}-t_n)\,\frac{H(t_1)}{i\hbar}\,\frac{H(t_2)}{i\hbar}\cdots\frac{H(t_n)}{i\hbar}
$$

$$
=\frac{1}{n!}\int_{t_0}^{t}\cdots\int_{t_0}^{t}\mathrm{d}t_1\cdots\mathrm{d}t_n\sum_{P\in S_n}\vartheta(t_{P(1)}-t_{P(2)})\vartheta(t_{P(2)}-t_{P(3)})\cdots\vartheta(t_{P(n-1)}-t_{p(n)})
$$

$$
\frac{H(t_{P(1)})}{i\hbar}\,\frac{H(t_{P(2)})}{i\hbar}\cdots\frac{H(t_{P(n)})}{i\hbar}
$$

把最后一式右边的被积函数叫做算符式$\frac{H(t_1)}{i\hbar}\frac{H(t_2)}{i\hbar}\cdots\frac{H(t_n)}{i\hbar}$的编时积，记作

$$C\left\{\frac{H(t_1)}{i\hbar}\ \frac{H(t_2)}{i\hbar}\cdots\ \frac{H(t_n)}{i\hbar}\right\}$$

$$\equiv\sum_{P\in S_n}\vartheta(t_{P(1)}-t_{P(2)})\vartheta(t_{P(2)}-t_{P(3)})\cdots\vartheta(t_{P(n-1)}-t_{p(n)})$$

$$\frac{H(t_{P(1)})}{i\hbar}\ \frac{H(t_{P(2)})}{i\hbar}\cdots\ \frac{H(t_{P(n)})}{i\hbar}$$

$$=\frac{H(t_{Q(1)})}{i\hbar}\ \frac{H(t_{Q(2)})}{i\hbar}\cdots\ \frac{H(t_{Q(n)})}{i\hbar},\text{当 } t_{Q(1)}\geq t_{Q(2)}\geq\cdots\geq t_{Q(n)}$$

可见编时积代表把算符式中各个依赖时间的因子按时间增加的顺序由右向左排列。这样(1.2.12)式可以写作

$$U(t,t_0)=1+\int_{t_0}^{t}\mathrm{d}t_1\frac{H(t_1)}{i\hbar}+\frac{1}{2!}\int_{t_0}^{t}\mathrm{d}t_1\int_{t_0}^{t}\mathrm{d}t_2C\left\{\frac{H(t_1)}{i\hbar}\ \frac{H(t_2)}{i\hbar}\right\}$$

$$+\frac{1}{3!}\int_{t_0}^{t}\mathrm{d}t_1\int_{t_0}^{t}\mathrm{d}t_2\int_{t_0}^{t}\mathrm{d}t_3C\left\{\frac{H(t_1)}{i\hbar}\ \frac{H(t_2)}{i\hbar}\ \frac{H(t_3)}{i\hbar}\right\}\qquad(1.2.13)$$

$$+\cdots+\frac{1}{n!}\int_{t_0}^{t}\mathrm{d}t_1\cdots\int_{t_0}^{t}\mathrm{d}t_nC\left\{\frac{H(t_1)}{i\hbar}\ \frac{H(t_2)}{i\hbar}\cdots\ \frac{H(t_n)}{i\hbar}\right\}+\cdots$$

考虑到算符的指数函数的定义，最后一个等式可以形式地记作

$$U(t,t_0)=C\exp\left\{\int_{t_0}^{t}\mathrm{d}t'\frac{H(t')}{i\hbar}\right\}\qquad(1.2.14)$$

此式并不代表先求定积分，再做指数映射，最后取编时积。它只是式(1.2.13)的一个约定的简单记法。式(1.2.13)是否收敛，依赖哈密顿算符的具体形状。我们不能一般而论地断言它是收敛的。在此意义下，我们称式(1.2.13)是演化算符的形式解。

Ⅲ. 一个求解含时 Schrödinger 方程的方法

运用量子力学理论，研究微观物理体系的性质和运动，通常就要求解含时 Schrödinger 方程。然而，精确求解含时 Schrödinger 方程总是很困难的。为求解含时 Schrödinger 方程，人们提出了各种各样的近似方法。一般而言，每种近似方法都有它

的适用范围。在此,我们向读者介绍一种适用性较广的近似方法。

在很多情况下,哈密顿算符可以写作这样两部分之和 $H(t)=H_0(t)+H_1(t)$,其中 $H_1(t)$很小,而 $H_0(t)$驱动的含时 Schrödinger 方程能够求解。这时含时 Schrödinger 方程可写作

$$i\hbar\frac{\partial}{\partial t}U(t,t_0)=[H_0(t)+H_1(t)]U(t,t_0),U(t_0,t_0)=1 \tag{1.2.15}$$

读者不难验证,若

$$i\hbar\frac{\partial}{\partial t}U^{(0)}(t,t_0)=H_0(t)U^{(0)}(t,t_0),U^{(0)}(t_0,t_0)=1 \tag{1.2.16}$$

$$\begin{aligned}i\hbar\frac{\partial}{\partial t}U^{(n)}(t,t_0)&=H_1(t)U^{(n-1)}(t,t_0)+H_0(t)U^{(n)}(t,t_0),U^{(n)}(t,t_0)\\&=0,\forall\, n\geq 1\end{aligned} \tag{1.2.17}$$

则 $U(t,t_0)\equiv\sum_{n=0}^{\infty}U^{(n)}(t,t_0)$ 满足(1.2.15)。这只要把(1.2.16)与所有的(1.2.17)相加即可。因而求解方程(1.2.15)的问题化为求解方程组(1.2.16)和(1.2.17)。我们已经知道方程(1.2.16)的形式解为

$$U^{(0)}(t,t_0)=C\exp\int_{t_0}^{t}\frac{H_0(t')}{i\hbar}\mathrm{d}t' \tag{1.2.18}$$

如果我们能够足够精确地计算出它,则利用(1.2.17),我们可以逐个地求出 $U^{(1)}(t,t_0),U^{(2)}(t,t_0)$直到我们对近似程度感到满意。

事实上,对(1.2.17)的第一式积分并考虑(1.2.17)的第二式得

$$U^{(n)}(t,t_0)=\int_{t_0}^{t}\mathrm{d}t_1\frac{H_1(t_1)}{i\hbar}U^{(n-1)}(t_1,t_0)+\int_{t_0}^{t}\mathrm{d}t_1\frac{H_0(t_1)}{i\hbar}U^{(n)}(t_1,t_0),\forall\, n\geq 1 \tag{1.2.19}$$

积分方程组(1.2.19)等价于微分方程组的初值问题(1.2.17)。我们用逐次迭代法求解它得

$$\begin{aligned}U^{(n)}(t,t_0)=&\int_{t_0}^{t}\mathrm{d}t_1\frac{H_1(t_1)}{i\hbar}U^{(n-1)}(t_1,t_0)+\int_{t_0}^{t}\mathrm{d}t_1\frac{H_0(t_1)}{i\hbar}\int_{t_0}^{t_1}\mathrm{d}t_2\frac{H_1(t_2)}{i\hbar}U^{(n-1)}(t_2,t_0)\\&+\int_{t_0}^{t}\mathrm{d}t_1\frac{H_0(t_1)}{i\hbar}\int_{t_0}^{t_1}\mathrm{d}t_2\frac{H_0(t_2)}{i\hbar}\int_{t_0}^{t_2}\mathrm{d}t_3\frac{H_1(t_3)}{i\hbar}U^{(n-1)}(t_3,t_0)+\cdots\end{aligned}$$

$$+\int_{t_0}^{t}\mathrm{d}t_1\,\frac{H_0(t_1)}{i\hbar}\int_{t_0}^{t_1}\mathrm{d}t_2\,\frac{H_0(t_2)}{i\hbar}\cdots$$
$$\int_{t_0}^{t_{m-2}}\mathrm{d}t_{m-1}\,\frac{H_0(t_{m-1})}{i\hbar}\int_{t_0}^{t_{m-1}}\mathrm{d}t_m\,\frac{H_1(t_m)}{i\hbar}U^{(n-1)}(t_m,t_0)+\cdots \tag{1.2.20}$$

我们来分析上式右边的典型项(最后一项)。它的积分区域是 $t_0\leq t_m\leq\cdots\leq t_2\leq t_1\leq t$。它的积分顺序是:先积 t_m,再积 t_{m-1},…,最后积 t_1。改变其中一个积分变量的名称:$t_m\mapsto t'$,并改变积分顺序:先积 t_{m-1},再积 t_{m-2},…,再积 t_1,最后积 t'。这样典型项可写作

$$\int_{t_0}^{t}\mathrm{d}t'\Big[\int_{t'}^{t}\mathrm{d}t_1\,\frac{H_0(t_1)}{i\hbar}\int_{t'}^{t_1}\mathrm{d}t_2\,\frac{H_0(t_2)}{i\hbar}\cdots\int_{t'}^{t_{m-2}}\mathrm{d}t_{m-1}\,\frac{H_0(t_{m-1})}{i\hbar}\Big]\frac{H_1(t')}{i\hbar}U^{(n-1)}(t',t_0) \tag{1.2.21}$$

于是(1.2.20)可以写作

$$\begin{aligned}U^{(n)}(t,t_0)=&\int_{t_0}^{t}\mathrm{d}t'\,\frac{H_1(t')}{i\hbar}U^{(n-1)}(t',t_0)+\int_{t_0}^{t}\mathrm{d}t'\Big[\int_{t'}^{t}\mathrm{d}t_1\,\frac{H_0(t_1)}{i\hbar}\Big]\frac{H_1(t')}{i\hbar}U^{(n-1)}(t',t_0)\\&+\int_{t_0}^{t}\mathrm{d}t'\Big[\int_{t'}^{t}\mathrm{d}t_1\,\frac{H_0(t_1)}{i\hbar}\int_{t'}^{t_1}\mathrm{d}t_2\,\frac{H_0(t_2)}{i\hbar}\Big]\frac{H_1(t')}{i\hbar}U^{(n-1)}(t',t_0)+\cdots\\&+\int_{t_0}^{t}\mathrm{d}t'\Big[\int_{t'}^{t}\mathrm{d}t_1\,\frac{H_0(t_1)}{i\hbar}\int_{t'}^{t_1}\mathrm{d}t_2\,\frac{H_0(t_2)}{i\hbar}\cdots\int_{t'}^{t_{m-2}}\mathrm{d}t_{m-1}\,\frac{H_0(t_{m-1})}{i\hbar}\Big]\\&\frac{H_1(t')}{i\hbar}U^{(n-1)}(t',t_0)+\cdots\end{aligned} \tag{1.2.22}$$

注意到

$$U^{(0)}(t,t')=1+\int_{t'}^{t}\mathrm{d}t_1\,\frac{H_0(t_1)}{i\hbar}+\int_{t'}^{t}\mathrm{d}t_1\,\frac{H_0(t_1)}{i\hbar}\int_{t'}^{t_1}\mathrm{d}t_2\,\frac{H_0(t_2)}{i\hbar}+\cdots \tag{1.2.23}$$

我们得到

$$U^{(n)}(t,t_0)=\int_{t_0}^{t}\mathrm{d}t'U^{(0)}(t,t')\,\frac{H_1(t')}{i\hbar}U^{(n-1)}(t',t_0),\ \forall\, n\geq 1 \tag{1.2.24}$$

反复利用这一递推公式还可以得到

$$U^{(n)}(t,t_0)=\int_{t_0}^{t}\mathrm{d}t_nU^{(0)}(t,t_n)\,\frac{H_1(t_n)}{i\hbar}\int_{t_0}^{t_n}\mathrm{d}t_{n-1}U^{(0)}(t_n,t_{n-1})=\frac{H_1(t_{n-1})}{i\hbar}\cdots$$

$$\int_{t_0}^{t_2} \mathrm{d}t_1 U^{(0)}(t_2,t_1)\frac{H_1(t_1)}{i\hbar}U^{(0)}(t_1,t_0) = \int_{t_0}^{t}\mathrm{d}t_n\int_{t_0}^{t_n}\mathrm{d}t_{n-1}\cdots\int_{t_0}^{t_2}\mathrm{d}t_1$$

$$U^{(0)}(t,t_n)\frac{H_1(t_n)}{i\hbar}U^{(0)}(t_n,t_{n-1})\frac{H_1(t_{n-1})}{i\hbar}\cdots U^{(0)}(t_2,t_1)\frac{H_1(t_1)}{i\hbar}U^{(0)}(t_1,t_0) \tag{1.2.25}$$

我们看到，知道了 $U^{(0)}(t,t_0)$，由公式(1.2.24)或(1.2.25)就可以计算出 $U^{(1)}(t,t_0)$，$U^{(2)}(t,t_0)$，…如果 $H_1(t)$不大，可期望 $U^{(n)}(t,t_0)\to 0$，当 $n\to\infty$ 时。在实际计算中，我们计算有限次修正，使精确度达到满意程度就成。在推导过程中，我们没有作任何近似。因而，只要 $H_0(t)$驱动的含时 Schrödinger 方程能够求解，而 $H_1(t)$较小，使级数 $\sum_{n=0}^{\infty}U^{(n)}(t,t_0)$收敛，我们就能使用这一方法，求解含时 Schrödinger 方程。可见它具有较广的适用范围。

让我们讨论 $U^{(n)}(t,t_0)$的物理意义。以$(i\hbar)^{-1}\mathrm{d}t$ 乘等式(1.2.15)得 $\mathrm{d}U(t,t_0)=\left(\frac{H_0(t)}{i\hbar}+\frac{H_1(t)}{i\hbar}\right)U(t,t_0)\mathrm{d}t$。此式告诉我们，在$(t,t+\mathrm{d}t)$时间间隔内，参与驱动的哈密顿算符 $H(t)$造成的演化算符的改变为$\frac{H(t)}{i\hbar}U(t,t_0)\mathrm{d}t$。体系的演化是由 $H_0(t)$，$H_1(t)$共同驱动的。在由 t_0 到 t 的时间间隔内，完全由 $H_0(t')(t_0<t'<t)$驱动的演化是 $U^{(0)}(t,t_0)$。在由 t_0 到 t 的时间间隔内，演化始终由 $H_0(t')$驱动，$H_1(t')$只在$(t_1,t_1+\mathrm{d}t_1)$内参与驱动。$H_1(t')$的这一参与驱动引起的 $U^{(0)}(t_1+\mathrm{d}t_1,t_0)$的变化为$\frac{H_1(t_1)}{i\hbar}U^{(0)}(t_1,t_0)\mathrm{d}t_1$，而它引起的 $U^{(0)}(t,t_0)$的变化为 $U^{(0)}(t,t_1)\frac{H_1(t_1)}{i\hbar}U^{(0)}(t_1,t_0)\mathrm{d}t_1$。所有这类参与驱动引起的对 $U^{(0)}(t,t_0)$的修正为 $U^{(1)}(t,t_0)$。$U(t,t_0)$的零级近似值 $U^{(0)}(t,t_0)$是 $H_0(t)$驱动的演化算符，是酉算符。$U(t,t_0)$的一级修正值 $U^{(1)}(t,t_0)$不是酉算符，故不对应于任一哈密顿算符驱动的演化算符。这种质的区别使 $U^{(n)}(t,t_0)(n\geq 2)$没有类似于 $U^{(1)}(t,t_0)$的物理意义。

作为例子，我们来利用上面公式推导含时微扰理论。设体系的哈密顿算符可以写作 $H(t)=H_0+H_1(t)$，式中 H_0 与时间无关，而 $H_1(t)$很小。H_0 的特征方程 $H_0u_n=\varepsilon_n u_n$ 可以求解，$u_1,u_2,\cdots$是 H_0 的一组完备的正交归一化的特征向量组。若在 $t_0=0$ 时刻，体系处于态 u_m，求在 t 时刻体系处在态 $u_l(l\neq m)$的几率 $P_{lm}(t)$。

事实上，t 时刻体系的态矢量为 $\psi(t)=U(t,0)u_m$，故在 t 时刻体系处在 $u_l(l\neq$

m)的几率

$$P_{lm}(t) = |\langle u_l | U(t,0) | u_m\rangle|^2$$
$$= |\langle u_l | U^{(0)}(t,0) | u_m\rangle + \langle u_l | U^{(1)}(t,0) | u_m\rangle + \cdots|^2 \quad (1.2.26)$$

由不含时间的 H_0 所驱动的演化算符为

$$U^{(0)}(t,0) = \exp\left(\frac{H_0}{i\hbar}t\right) \quad (1.2.27)$$

在 H_0 表象中，H_0 的矩阵是对角的。由矩阵乘法的定义可知，$(H_0)^n$ 的矩阵也是对角的。从而 $U^{(0)}(t,0)$的矩阵也是对角的

$$\langle u_l | U^{(0)}(t,0) | u_m\rangle = \exp\left(\frac{\varepsilon_m t}{i\hbar}\right)\delta_{lm} \quad (1.2.28)$$

取式(1.2.25)两边的 l 行 m 列($l \neq m$)矩阵元并令式中的 $n=1$ 得

$$\langle u_l | U^{(1)}(t,0) | u_m\rangle = \sum_{k\geq 1}\sum_{n\geq 1}\int_0^t \mathrm{d}t' \langle u_l | U^{(0)}(t,t') | u_k\rangle$$
$$\langle u_k | \frac{H_1(t')}{i\hbar} | u_n\rangle\langle u_n | U^{(0)}(t',0) | u_m\rangle$$
$$= \int_0^t \mathrm{d}t' \exp\left[\frac{\varepsilon_l}{i\hbar}(t-t')\right]\langle u_l | \frac{H_1(t')}{i\hbar} | u_m\rangle \exp\left[\frac{\varepsilon_m}{i\hbar}t'\right]$$
$$= \exp\left[\frac{\varepsilon_l t}{i\hbar}\right]\int_0^t \mathrm{d}t' \langle u_l | \frac{H_1(t')}{i\hbar} | u_m\rangle \exp\left(\frac{\varepsilon_m - \varepsilon_l}{i\hbar}t'\right)$$

从而在一级近似下跃迁几率

$$P_{lm}(t) = \left|\int_0^t \mathrm{d}t' \langle u_l \left| \frac{H_1(t')}{i\hbar} \right| u_m\rangle \exp\left(\frac{\varepsilon_m - \varepsilon_l}{i\hbar}t'\right)\right|^2 \quad (1.2.29)$$

这就是含时微扰理论的结果。

§1.3 表象与绘景(Representations and Pictures)

Ⅰ.量子力学的矩阵形式

量子力学中，体系在某一时刻的状态是由其态矢量 Hilbert 空间 S 中的非零向量

所刻画的。体系的任一力学量由该 Hilbert 空间上的一个厄米算符所刻画。而读者从第 0 章可知，当选定了 Hilbert 空间 S 的一个基$\{u_1, u_2, \cdots\}$，S 中的任一向量ψ 都可以唯一地表作基矢的线性组合

$$\psi = \sum_j u_j C^j \equiv u_j C^j \tag{1.3.1}$$

我们将把由所有展开系数排列而成的列矩阵$\begin{bmatrix} C^1 \\ C^2 \\ \vdots \end{bmatrix}$记作$[\psi]$。这样我们可以用列矩阵$[\psi]$代表 ψ。读者不难证明$[\alpha\psi + \beta\varphi] = \alpha[\psi] + \beta[\varphi]$，也就是说向量线性组合的列矩阵等于向量列矩阵的同样系数的线性组合。从第 0 章的内容中读者还可以知道，当选定了 Hilbert 空间 S 的一个基$\{u_1, u_2, \cdots\}$，S 上的任一线性算符 A，按下列方式确定一个方阵$[A]$：其 j 行 k 列矩阵元为下式中的 A_k^j

$$Au_k = \sum_j u_j A_k^j \equiv u_j A_k^j \tag{1.3.2}$$

不难证明(见 §0.3)向量 $A\psi$ 的列矩阵等于线性算符 A 的方阵左乘向量 ψ 的列矩阵

$$[A\psi] = [A][\psi] \tag{1.3.3}$$

同样不难证明，对任意线性算符 A，B 和复数 α

$$\begin{aligned} [A + B] &= [A] + [B] \\ [A\alpha] &= \alpha[A] \\ [AB] &= [A][B] \end{aligned} \tag{1.3.4}$$

以上的讨论表明，当选定体系态矢量空间 S 的一个基$\{u_1, u_2, \cdots\}$时，我们可以用由式(1.3.1)所确定的列矩阵$[\psi]$代表态矢量 ψ，用由式(1.3.2)确定的方阵$[A]$代表线性算符 A。态矢量的线性运算可归为列向量的线性运算(指向量加法和向量的数乘)，线性算符对向量的作用可归为线性算符的方阵左乘向量的列矩阵，而线性算符的加法，数乘和乘法可归为方阵的加法，数乘和乘法。

特别是，当基$\{u_1, u_2, \cdots\}$是正交归一基

$$\langle u_j \mid u_k \rangle = \delta_{jk} \tag{1.3.5}$$

时，由式(1.3.1)所确定的态矢量 ψ 的列矩阵$[\psi]$的矩阵元为

$$C^j = \langle u_j \mid \psi \rangle \tag{1.3.6}$$

(由于对任一态矢量 $\psi \in S$ 都有$\mid \psi \rangle = \sum_j \mid u_j \rangle \langle u_j \mid \psi \rangle$，可得正交归一基的完备性$\sum_j \mid$

$u_j\rangle\langle u_j| = 1$,这是一个非常有用的性质。)而由式(1.3.2)所确定的线性算符 A 的方阵 $[A]$的矩阵元为

$$A_k^j = \langle u_j \mid Au_k\rangle \equiv \langle u_j \mid A \mid u_k\rangle \tag{1.3.7}$$

由上式可得

$$A_k^j = \langle A^+ u_j \mid u_k\rangle \equiv \langle u_k \mid A^+ u_j\rangle^* = (A_j^{+k})^*$$

此即

$$[A^+] = [A]^+ \tag{1.3.8}$$

也就是说,在正交归一基下,线性算符 A 的厄米共轭的矩阵等于线性算符 A 的矩阵的厄米共轭。读者不难证明,对于正交归一基,两个态矢量的内积可以写作

$$\langle \psi \mid \varphi\rangle = [\psi]^+ [\varphi] \tag{1.3.9}$$

从而态矢量的归一化条件可以写作

$$1 = \langle \psi \mid \psi\rangle = [\psi]^+ [\psi] \tag{1.3.10}$$

而力学量的平均值可以写作

$$\langle A\rangle = \langle \psi \mid A \mid \psi\rangle = [\psi]^+ [A][\psi] \tag{1.3.11}$$

由于当两个向量相等时,它们的列矩阵相等,线性算符 A 的特征方程 $Av_n = a_n v_n$ 可以写作

$$[A][v_n] = a_n[v_n] \tag{1.3.12}$$

而 Schrödinger 方程 $i\hbar \dfrac{\partial}{\partial t}\psi(t) = H(t)\psi(t)$可以写作

$$i\hbar \frac{\partial}{\partial t}[\psi(t)] = [H(t)][\psi(t)] \tag{1.3.13}$$

我们看到量子力学有不同矩阵表象,但它们给出同样的测量结果(物理量的平均值,可能值,各可能值出现的几率等),在处理具体问题时,采用适当的表象可以使问题大大的简化。当我们在作理论分析时,往往谈到态矢量 ψ(或$|\psi\rangle$)和线性算符 A,它们是与主观采用的表象无关的概念。量子力学的不同矩阵表象对应态矢量空间不同正交归一基的选择。由第 0 章知,每一个正交归一基都可以看作是某个厄米算符(某个力学量)的特征向量组。相应的矩阵表象称为该力学量的表象。

练习 1-3-1 试证:等式(1.3.6)、(1.3.7)和(1.3.8)对非正交归一基不一定成立。

练习 1-3-2 把本小节的讨论推广到连续谱的情况。

练习 1-3-3 把坐标表象,动量表象纳入(广义)矩阵形式。

Ⅱ. 量子力学的各种绘景

上面介绍的是量子力学的各种矩阵表象,有一个共同的特点,即所选用的态矢量空间 S 的正交归一基都不随时间变化。正是这个特点保证了(不显含时间的)力学算符的矩阵不随时间变化,而态矢量 $\psi(t)$ 的矩阵 $[\psi(t)]$ 满足矩阵形式的 Schrödinger 方程(1.3.13)。对应于态矢量空间 S 的不随时间变化的正交归一基的这一类矩阵表象,统称为 Schrödinger 绘景。

在本章第一节中我们讨论过,由于哈密顿算符的厄米性,Schrödinger 方程任意两个解的内积都不随时间变化 $\langle\varphi(t)\mid\psi(t)\rangle=\langle\varphi(0)\mid\psi(0)\rangle$。这使得我们可以谈论态矢量空间 S 的由 Schrödinger 方程的解构成的正交归一基。事实上,设 $\{u_1,u_2,\cdots\}$ 是 S 的任一正交归一基。令 $u_j(t)$ 是 Schrödinger 方程的由初始条件 $u_j(0)=u_j$ 确定的解,则

$$\{u_1(t),u_2(t)\cdots\} \tag{1.3.14}$$

就是态矢量空间 S 的由 Schrödinger 方程的解构成的正交归一基。采用这种基,把体系的态矢量 $\psi(t)$ 向基(1.3.14)展开,

$$\psi(t)=\sum_j u_j(t)\langle u_j(t)\mid\psi(t)\rangle \tag{1.3.15}$$

由于 $u_j(t),\psi(t)$ 都是 Schrödinger 方程的解,故 $\langle u_j(t)\mid\psi(t)\rangle$ 与时间 t 无关。从而 $\psi(t)$ 在基(1.3.14)下的列矩阵 $[\psi(t)]^H$ 与时间无关:

$$[\psi(t)]^H=[\psi(0)]^H \tag{1.3.16}$$

体系任一力学量 $A(t)$ 在基(1.3.14)下的矩阵 $[A(t)]^H$ 的矩阵元为 $\langle u_j(t)\mid A(t)\mid u_k(t)\rangle$。对 t 微商得

$$\begin{aligned}i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle u_j(t)\mid A(t)\mid u_k(t)\rangle&=\langle u_j(t)\mid A(t)H(t)-H(t)A(t)\mid u_k(t)\rangle\\&\quad+\langle u_j(t)\mid i\hbar\frac{\partial A(t)}{\partial(t)}\mid u_k(t)\rangle\\&=\sum_l(\langle u_j(t)\mid A(t)\mid u_l(t)\rangle\langle u_l(t)\mid H(t)\mid u_k(t)\rangle\end{aligned}$$

$$-\langle u_j(t)\mid H(t)\mid u_l(t)\rangle\langle u_l(t)\mid A(t)\mid u_k(t)\rangle$$
$$+i\hbar\langle u_j(t)\left|\frac{\partial A(t)}{\partial(t)}\right|u_k(t)\rangle$$

写成矩阵等式

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}[A(t)]^H=\left[[A(t)]^H,[H(t)]^H\right]_-+i\hbar\left[\frac{\partial A(t)}{\partial(t)}\right]^H \tag{1.3.17}$$

我们得到了在基(1.3.14)下,力学量 $A(t)$的矩阵$[A(t)]^H$ 随时间变化所遵循的方程(海森堡方程)。特别是,当$\frac{\partial A(t)}{\partial t}=0$,即力学量不显含时间时,我们有

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}[A]^H=\left[[A]^H,[H(t)]^H\right]_- \tag{1.3.18}$$

在这种矩阵表象下,两个态矢量的内积为

$$\langle\varphi(t)\mid\psi(t)\rangle=[\psi(0)]^{H+}[\psi(0)]^H \tag{1.3.19}$$

体系任一力学量 $A(t)$的平均值为

$$\langle A(t)\rangle=\langle\psi(t)\mid A(t)\mid\psi(t)\rangle=[\psi(0)]^{H+}[A(t)]^H[\psi(0)]^H \tag{1.3.20}$$

我们把这类矩阵表象统称为 Heisenberg 绘景。它们对应于态矢量空间 S 的由 Schrödinger 方程的解构成的正交归一基。在这类矩阵表象中,态矢量 $\psi(t)$的列矩阵$[\psi(t)]^H=[\psi(0)]^H$ 与时间无关,而力学量 $A(t)$的方阵$[A(t)]^H$ 按海森堡方程(1.3.17)随时间变化。

除了上面介绍的两种绘景外,还有一类矩阵表象,它们对应的态矢量空间 S 的正交归一基$\{u_1(t),u_2(t)\cdots\}$随时间变化(至少有一个基矢随时间变化),但并非所有基矢都是 Schrödinger 方程的解。这类矩阵表象统称相互作用绘景。

把体系的哈密顿量写成两个厄米算符之和,

$$H(t)=H_0(t)+H_1(t) \tag{1.3.21}$$

由于 $H_0(t)$是厄米算符,故由 $H_0(t)$驱动的 Schrödinger 方程

$$i\hbar\frac{\partial}{\partial t}\varphi(t)=H_0(t)\varphi(t) \tag{1.3.22}$$

的任意两个解的内积不随时间变化:$\langle\varphi_1(t)\mid\varphi_2(t)\rangle=\langle\varphi_1(0)\mid\varphi_2(0)\rangle$。任给 S 的一个正交归一基$\{v_1,v_2,\cdots\}$,令 $v_j(t)$是方程(1.3.22)的满足初始条件 $v_j(0)=v_j$ 的解,

$$i\hbar\frac{\partial}{\partial t}v_j(t)=H_0(t)v_j(t) \tag{1.3.23}$$

显然

$$\{v_1(t), v_2(t), \cdots\} \tag{1.3.24}$$

是态矢量空间 S 的一个正交归一基，其中每个基矢都满足 $H_0(t)$ 驱动的 Schrödinger 方程(1.3.22)。选定 S 的正交归一基(1.3.24)，把体系态矢量 $\psi(t)$ 向基(1.3.24)展开，则所得列矩阵 $[\psi(t)]^I$ 的矩阵元为 $\langle v_j(t) | \psi(t)\rangle$。体系的任一力学量 $A(t)$ 的方阵 $[A(t)]^I$ 的矩阵元为 $\langle v_j(t) | A(t) | v_k(t)\rangle$。把 $\langle v_j(t) | \psi(t)\rangle$ 对时间微商得

$$\begin{aligned} i\hbar \frac{\partial}{\partial t}\langle v_j(t) \mid \psi(t)\rangle &= -\langle H_0(t) v_j(t) \mid \psi(t)\rangle + \langle v_j(t) \mid H(t)\psi(t)\rangle \\ &= \langle v_j(t) \mid H_1(t)\psi(t)\rangle \\ &= \sum_k \langle v_j(t) \mid H_1(t) \mid v_k(t)\rangle\langle v_k(t) \mid \psi(t)\rangle \end{aligned}$$

写成矩阵的形式就是

$$i\hbar \frac{\partial}{\partial t}[\psi(t)]^I = [H_1(t)]^I[\psi(t)]^I \tag{1.3.25}$$

而把 $\langle v_j(t) | A(t) | v_k(t)\rangle$ 对 t 微商得

$$\begin{aligned} i\hbar \frac{\partial}{\partial t}\langle v_j(t) \mid A(t) \mid v_k(t)\rangle &= -\langle H_0(t) v_j(t) \mid A(t) \mid v_k(t)\rangle \\ &+ \langle v_j(t) \mid A(t)H(t) \mid v_k(t)\rangle + i\hbar\langle v_j(t) \mid \frac{\partial A(t)}{\partial t} \mid v_k(t)\rangle \\ = i\hbar\langle v_j(t) \mid \frac{\partial A(t)}{\partial t} \mid v_k(t)\rangle &+ \sum_l \langle v_j(t) \mid A(t) \mid v_l(t)\rangle\langle v_l(t) \mid H_0(t) \mid v_k(t)\rangle \\ &- \sum_l \langle v_j(t) \mid H_0(t) \mid v_l(t)\rangle\langle v_l(t) \mid A(t) \mid v_k(t)\rangle \end{aligned}$$

写成矩阵形式就是

$$i\hbar \frac{\partial}{\partial t}[A(t)]^I = i\hbar\left[\frac{\partial A(t)}{\partial t}\right]^I + [[A(t)]^I, [H_0(t)]^I]_- \tag{1.3.26}$$

体系任意两个态矢量 $\psi(t)$ 和 $\varphi(t)$ 的内积为

$$\langle \psi(t) \mid \varphi(t)\rangle = [\psi(t)]^{I+}[\varphi(t)]^I \tag{1.3.27}$$

体系任一力学量 $A(t)$ 在态 $\psi(t)$ 下的平均值为

$$\langle A(t)\rangle = \langle \psi(t) \mid A(t) \mid \psi(t)\rangle = [\psi(t)]^{I+}[A(t)]^I[\varphi(t)]^I \tag{1.3.28}$$

力学量 $A(t)$的特征方程 $A(t)w_n(t)=\alpha_n(t)w_n(t)$的矩阵形式为

$$[A(t)]^I[w_n(t)]^I=\alpha_n(t)[w_n(t)]^I \tag{1.3.29}$$

我们看到在这类矩阵表象中,由于基矢满足 $H_0(t)$驱动的 Schrödinger 方程,体系态矢量 $\psi(t)$的列矩阵$[\psi(t)]^I$ 满足$H_1(t)$驱动的 Schrödinger 方程(1.3.25),而力学量 $A(t)$的方阵$[A(t)]^I$ 满足 $H_0(t)$驱动的 Heisenberg 方程(1.3.26)。在实际问题中,$H_1(t)$往往取作相互作用能,故把这类矩阵表象统称相互作用绘景。

一般而言,设$\{v_1(t),v_2(t),\cdots\}$是态矢量空间 S 的任一含时正交归一基。我们可以证明存在唯一确定的厄米算子 $H_0(t)$,使

$$i\hbar\frac{\partial}{\partial t}v_j(t)=H_0(t)v_j(t),\forall j=1,2,\cdots \tag{1.3.30}$$

事实上,对任给的时间 t,方程组

$$v_j(t)=V_0(t,0)v_j(0),\forall j=1,2,\cdots$$

可以唯一确定线性算符 $V(t,0)$。由于$\{v_1(0),v_2(0),\cdots\}$和$\{v_1(t),v_2(t),\cdots\}$都是正交归一基,$V(t,0)$一定是酉算符。显然有 $V(0,0)=I$(S 上的恒等算符)。令

$$H_0(t)=i\hbar\frac{\partial V(t,0)}{\partial t}V^+(t,0) \tag{1.3.31}$$

则这样定义的 $H_0(t)$一定使(1.3.30)成立,此外不难看出(1.3.31)定义的 $H_0(t)$是厄米的(请读者证明之)。至此我们用构造法证明了满足条件(1.3.30)的 $H_0(t)$的存在性。现在考虑唯一性。若有另一厄米算符 $G_0(t)$满足(1.3.30),则我们有

$$H_0(t)v_j(t)=G_0(t)v_j(t),\forall j=1,2,\cdots$$

由于$\{v_1(t),v_2(t),\cdots\}$是完备向量组,我们得到线性算符等式 $H_0(t)=G_0(t)$。唯一性证完。

因此在一般情况下,态矢量空间 S 的任一含时正交归一基$\{v_1(t),v_2(t),\cdots\}$都可以看作由某个厄米算符 $H_0(t)$所驱动的 Schrödinger 方程的解构成的。体系的真正的哈密顿 $H(t)$总可以写作两个厄米算子之和 $H(t)=H_0(t)+(H(t)-H_0(t))\equiv H_0(t)+H_1(t)$。在极端情况下,当

$$\frac{\partial v_j(t)}{\partial t}=0,\forall j=1,2,\cdots$$

我们有 $H_0(t)=0,H_1(t)=H(t)$。我们得到 Schrödinger 绘景;当 $v_j(t)(\forall j=1,2,\cdots)$

是 $H(t)$驱动的 Schrödinger 方程的解时，我们有 $H_0(t)=H(t)$，$H_1(t)=0$，我们得到 Heisenberg 绘景。一般情况下，$H_0(t)\neq 0$ 同时 $H_0(t)\neq H(t)$（体系实际的哈密顿），我们得到相互作用表象。

练习 1-3-4 试证：若$\{v_1(0),v_2(0),\cdots\}$和$\{v_1(t),v_2(t),\cdots\}$都是正交归一基，则由方程组 $v_j(t)=V(t,0)v_j(0)$，$\forall j=1,2,\cdots$确定线性算符 $V(t,0)$为 $V(t,0)=\sum_{j\geq 1}|v_j(t)\rangle\langle v_j(0)|$，而式 (1.3.31) 中的 $H_0(t)=i\hbar\frac{\partial V(t,0)}{\partial t}V^{+}(t,0)=i\hbar\sum_{j\geq 1}\left(\frac{\partial}{\partial t}|v_j(t)\rangle\right)\langle v_j(t)|$。

Ⅲ. 对绘景的另一种理解

在量子力学中我们用态矢量描述体系的状态，用厄米算符描述力学量。但我们物理上可测量的量却不是态矢量或力学量本身，而是由它们给出的力学量的可能值，平均值，各可能值出现的几率等。这就使我们有可能用不同的方式赋予体系状态以态矢量，赋予体系力学量以厄米算符，只要不同方法给出的物理测量结果相同就行。

在本章第一节我们介绍量子力学基本假定时，赋予体系的每一个状态一个态矢量，而赋予体系的每一个确定的力学量 $F(q,p)$一个厄米算符 $F(\hat{q},\hat{p})$。体系的状态随时间变化，描述体系状态的态矢量也随时间变化。我们假设体系状态的变化遵从 Schrödinger 方程 $i\hbar\frac{\partial}{\partial t}\psi(t)=H(t)\psi(t)$。而描写确定力学量（特别是坐标 q 和动量 p）的厄米算符不随时间变化。（显含时间 t 的力学量 $F(q,p,t)$在不同时刻 t 是不同的力学量，故算符 $F(\hat{q},\hat{p},t)$随时间变化与我们的断言并不矛盾）。在 $\psi(t)$态下，力学量 $F(\hat{q},\hat{p},t)$的平均值是$\langle\psi(t)|F(\hat{q},\hat{p},t)|\psi(t)$，力学量 $F(\hat{q},\hat{p},t)$的可能值是厄米算符 $F(\hat{q},\hat{p},t)$的本征值。设 $F(\hat{q},\hat{p},t)$的本征方程为 $F(\hat{q},\hat{p},t)u_n(t)=f_n(t)u_n(t)$，$\{u_1(t),u_2(t),\cdots\}$是完备的正交归一的特征向量组。则在态 $\psi(t)$下测量 $F(q,p,t)$，可能值 $f_n(t)$出现的几率为$|\langle u_n(t)|\psi(t)\rangle|^2$。

我们把量子力学的这种表述方式称之为 Schrödinger 绘景。我们将以此为出发点，引入量子力学的另一种表述方式 Heisenberg 绘景。

从 §1.2 我们知道演化算符 $U(t,t_0)$是幺正算符。若体系在时刻 t 的状态在 Schrödinger 绘景中由态矢量 $\psi(t)$刻画，令

$$\psi^H(t) = U(0,t)\psi(t) \tag{1.3.32}$$

我们改由态矢量 $\psi^H(t)(=\psi(0))$刻画体系在 t 时刻的状态。相应地,若体系的力学量 $F(q,p,t)$在 Schrödinger 绘景下对应于厄米算子 $F(\hat{q},\hat{p},t)$,令

$$F^H(\hat{q},\hat{p},t) = U(0,t)F(\hat{q},\hat{p},t)U(0,t)^+ \tag{1.3.33}$$

我们改由算符 $F^H(\hat{q},\hat{p},t)$刻画体系的力学量。由于(1.3.32)是幺正变换,(1.3.33)是幺正相似变换,读者不难证明

$$\langle\psi^H(t)\mid\varphi^H(t)\rangle = \langle\psi(t)\mid\varphi(t)\rangle \tag{1.3.34}$$

$$\langle\psi^H(t)\mid F^H(\hat{q},\hat{p},t)\mid\psi^H(t)\rangle = \langle\psi(t)\mid F(\hat{q},\hat{p},t)\mid\psi(t)\rangle \tag{1.3.35}$$

此外 Schrödinger 由绘景下力学量 $F(q,p,t)$算符的本征方程 $F(\hat{q},\hat{p},t)u_n(t) = f_n(t)u_n(t)$,我们容易得出

$$F^H(\hat{q},\hat{p},t)U(0,t)u_n(t) = f_n(t)U(0,t)u_n(t) \tag{1.3.36}$$

这表明两种表述方式给出两个态的同样的内积(1.3.34),给出力学量同样的测量平均值(1.3.35)和可能值(1.3.36)。联合(1.3.34)和(1.3.36),我们有

$$\langle u_n(t)\mid\psi(t)\rangle = \langle U(0,t)u_n(t)\mid U(0,t)\psi(t)\rangle = \langle u_n^H(t)\mid\psi^H(t)\rangle \tag{1.3.37}$$

这表明两种表述方式给出力学量可能值 $f_n(t)$相同的几率。补充一句,幺正相似变换保持算符的厄米性。由(1.3.32)我们看到 $\psi^H(t)=\psi(0)$不随时间变化。而不显含时间 t 的力学量 $F(q,p)$的算符 $F^H(\hat{q},\hat{p})=U(0,t)F(\hat{q},\hat{p})U(0,t)^+$也随时间 t 变化。由(1.3.33)可得

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}F^H(\hat{q},\hat{p},t) = i\hbar\left(\frac{\partial F(\hat{q},\hat{p},t)}{\partial t}\right)^H + [F(\hat{q},\hat{p},t),H^H(t)]_- \tag{1.3.38}$$

量子力学的这种表述方式称作 Heisenberg 绘景。在这种绘景下,体系随时间演化所经历的一系列的状态由一个态矢量刻划 $\psi^H(t)=\psi(0)=\psi^H(0)$。而力学量算符按 Heisenberg 方程(1.3.38)随时间变化。

把体系的哈密顿量写成两个厄米算符之和 $H(t)=H_0(t)+H_1(t)$。令 $V(t,t_0)$代表 $H_0(t)$驱动的演化算符:

$$i\hbar\frac{\partial}{\partial t}V(t,t_0) = H_0(t)V(t,t_0), V(t_0,t_0) = 1 \tag{1.3.39}$$

$V(t,t_0)$显然是幺正算符。设体系在 t 时刻的状态在 Schrödinger 绘景下由态矢量 ψ

(t)刻划。

令

$$\psi^I(t) = V(0,t)\psi(t) \tag{1.3.40}$$

我们改用 $\psi^I(t)$刻划体系在 t 时刻的状态。相应地,令

$$F^I(\hat{q},\hat{p},t) = V(0,t)F(\hat{q},\hat{p})V(0,t)^+ \tag{1.3.41}$$

我们改用 $F^I(\hat{q},\hat{p},t)$描写体系的力学量 $F(q,p,t)$,它在 Schrödinger 绘景下是由厄米算符 $F(\hat{q},\hat{p},t)$刻划的。由于(1.3.40)是幺正变换,(1.3.41)是幺正相似变换,不难看出

$$\langle\psi(t)\mid\varphi(t)\rangle = \langle\psi^I(t)\mid\varphi^I(t)\rangle \tag{1.3.42}$$

$$\langle\psi(t)\mid F(\hat{q},\hat{p},t)\mid\psi(t)\rangle = \langle\psi^I(t)\mid F^I(\hat{q},\hat{p},t)\mid\psi^I(t)\rangle \tag{1.3.43}$$

另外由 $F(\hat{q},\hat{p},t)$的特征方程 $F(\hat{q},\hat{p},t)u_n(t)=f_n(t)u_n(t)$可得

$$F^I(\hat{q},\hat{p},t)V(0,t)u_n(t) = f_n(t)V(0,t)u_n(t) \tag{1.3.44}$$

从而有

$$\langle V(0,t)u_n(t)\mid\psi^I(t)\rangle = \langle u_n^I(t)\mid\psi^I(t)\rangle = \langle u_n(t)\mid\psi(t)\rangle \tag{1.3.45}$$

这表明用 $\psi^I(t)$, $F^I(\hat{q},\hat{p},t)$,给出的物理上可测量的结果与用 $\psi(t)$, $F(\hat{q},\hat{p},t)$给出的结果相同。让我们观察 $\psi^I(t)$如何随时间变化。取式(1.3.40)对时间 t 的微商得

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\psi^I(t) &= i\hbar\frac{\partial}{\partial t}V(t,0)^+\psi(t) + V(0,t)H(t)\psi(t) \\ &= -(H_0(t)V(t,0))^+\psi(t) + V(0,t)H(t)\psi(t) \\ &= V(0,t)H_1(t)V(0,t)^+V(0,t)\psi(t) = H_1^I(t)\psi^I(t) \end{aligned} \tag{1.3.46}$$

为观察 $F^I(\hat{q},\hat{p},t)$如何随时间 t 变化,取式(1.3.41)对 t 的微商得

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}F^I(\hat{q},\hat{p},t) = i\hbar\left(\frac{\partial F(\hat{q},\hat{p},t)}{\partial t}\right)^I + [F^I(\hat{q},\hat{p},t),H_0^I(t)]_- \tag{1.3.47}$$

我们看到,在这种表述方式下,描写体系状态的态矢量 $\psi^I(t)$满足由 $H_1^I(t)$驱动的 Schrödinger 方程(1.3.46),而刻划体系力学量 $F(q,p,t)$的算符 $F^I(\hat{q},\hat{p},t)$满足由 $H_0^I(t)$驱动的 Heisenberg 方程(1.3.47)。量子力学的这种表述方式我们称之为相互作用绘景。

§1.4 密度算符

Ⅰ.纯态系综与混态系综

量子力学对处于态$|\psi\rangle$的体系的各种力学量测量结果的预言是统计性的。这意味着要对大量处在同一量子态的体系测量力学量才能把测量结果和理论预言作比较。然而由同一装置在同样条件下制备出的各个量子体系,有时处于同一量子态,有时并不处于同一量子态。为搞清楚这一点,让我们观察下面的例子。图 1.4-1 所示的装置由电子射束源、准直仪、滤速器和产生非均匀磁场的磁铁构成。

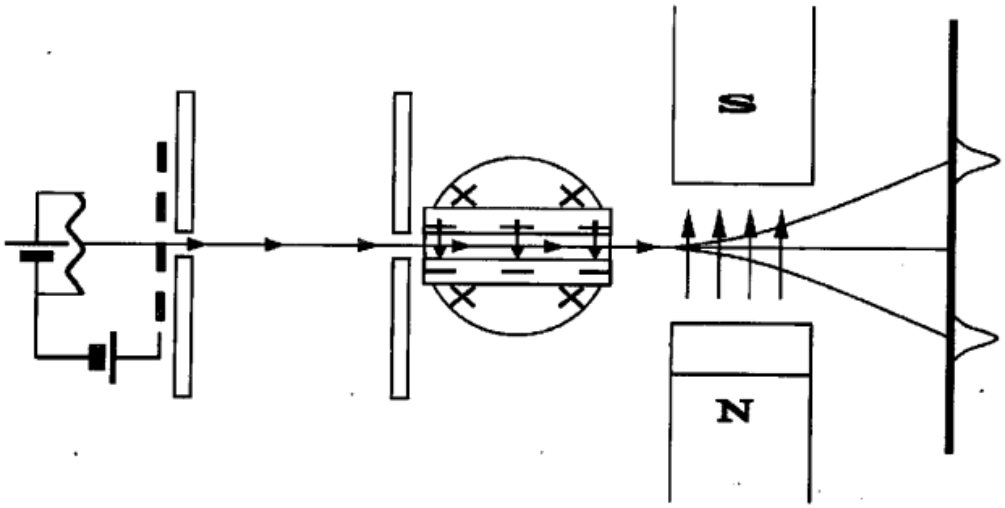

图 1.41

滤速器由互相垂直的均匀电场 $\boldsymbol{E}$ 和均匀磁场 $\boldsymbol{B}$ 构成。当带电粒子以速度 v 按 $\boldsymbol{E}\times\boldsymbol{B}$ 的方向入射,点电荷 q 受到的 Lorentz 力为 $\boldsymbol{f}=q\boldsymbol{E}+\frac{q}{c}\boldsymbol{v}\times\boldsymbol{B}$,其中电场力与磁场力方向相反。在图 1.4-1 的情况下

$$\boldsymbol{f}=q\left(E-\frac{v}{c}B\right)\hat{k}=q\left(E-\frac{p}{mc}B\right)\hat{k} \tag{1.4.1}$$

可见只有动量满足条件 $E=\frac{p}{mc}B$ 的粒子受到的力为零,能够飞出滤速器。

现有热电子束飞向准直仪,只有动量满足下面条件的电子能通过准直仪和滤速

器：

$$\boldsymbol{p} = \frac{Emc}{B}\hat{j} \tag{1.4.2}$$

它们继续飞经非均匀磁场。电子磁矩和自旋的关系为 $\boldsymbol{\mu} = \frac{-e}{mc}\boldsymbol{S}$，而磁矩与磁场的相互作用能为 $-\boldsymbol{\mu}\cdot\boldsymbol{B}(\boldsymbol{r})$。在非均匀磁场中电子磁矩受到一个力 $\boldsymbol{F}(\boldsymbol{r}) = \nabla(\boldsymbol{\mu}\cdot\boldsymbol{B}(\boldsymbol{r}))$。在图 1.4-1 所示的情况下

$$\boldsymbol{F}(\boldsymbol{r}) = \mu_z \frac{\partial B_z(\boldsymbol{r})}{\partial z}\hat{k} = -S_z \frac{e}{mc}\frac{\partial B_z(\boldsymbol{r})}{\partial z}\hat{k} \tag{1.4.3}$$

式中 $\frac{\partial B_z(\boldsymbol{r})}{\partial z} < 0$。在这个力的作用下，电子束一分为二。向上偏转的一束中的电子自旋向上，向下偏转的一束中的电子自旋向下。在整个实验中电子束的流密度非常小，可以忽略电子间的相互作用。在这个实验中，如果我们只收集向上偏转的电子，则所收集的每个电子都处在同一个量子态 $\Psi_+(\boldsymbol{r}) = (2\pi\hbar)^{-3/2}\exp\left(i\frac{\boldsymbol{p}_+\cdot\boldsymbol{r}}{\hbar}\right)\begin{pmatrix}1\\0\end{pmatrix}$。它们构成所谓“纯态系综”，由态矢量 $\Psi_+(\boldsymbol{r})$ 描述。如果我们收集全部由非均匀磁场出射的电子，则在所收集的电子中，50% 处在量子态 $\Psi_+(\boldsymbol{r})$，而另一半处在量子态 $\Psi_-(\boldsymbol{r}) = (2\pi\hbar)^{-3/2}\exp\left(i\frac{\boldsymbol{p}_-\cdot\boldsymbol{r}}{\hbar}\right)\begin{pmatrix}0\\1\end{pmatrix}$。如果我们在滤速器出口处收集全部出射的电子，则我们所收集的电子，50% 处在量子态 $(2\pi\hbar)^{-3/2}\exp\left(i\frac{\boldsymbol{p}\cdot\boldsymbol{r}}{\hbar}\right)\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}$，50% 处在量子态 $(2\pi\hbar)^{-3/2}\exp\left(i\frac{\boldsymbol{p}\cdot\boldsymbol{r}}{\hbar}\right)\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}}\end{bmatrix}$。其中 $\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}$、$\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}}\end{bmatrix}$ 是 $\hat{S}_x$ 的本征态。在后两种情况下，制备出的样品处在不同的量子态，它们在不同量子态间的分布不随时间变化，这种系综称为混合系综。

一般而言，对大量的某种微观体系进行力学量完全集的测量，并且只收集特定测量结果的样品，这时所收集的每个样品都处在同一个量子态 $|\psi\rangle$，它们的集合称为纯态系综，用态矢量 $|\psi\rangle$ 描述。在其它情况下，例如非完备测量，或者完备测量但收集

多种测量结果的样品等，我们所收集的样品按一定相对丰度处于不同的量子态。大量这种样品的集合就是混态系综。无论是纯态系综或者混态系综，同一系综中所有样品必须是处在同样外界条件下的同种量子体系，它们的哈密顿算符相同。

我们使用了系综一词，但请读者注意，它和经典统计中系综一词的涵义不同。在经典情况下，两次测量同一力学量如果测量结果不同，意味着两次被测量的样品一定处在不同的状态。而在量子纯态系综情况下，所有样品处在同一量子态，两次测量同一力学量，如果测量结果不同并不意味着两次被测样品处在不同的量子态。在量子混态系综情况下，对同一力学量进行测量，测量结果的统计分布是由两方面的因素确定的，一是纯量子的，即在同一量子态$|\psi_i\rangle$中力学量的分布，另一可以称之为“经典的”，即样品在不同量子态间的分布。

Ⅱ. 混态系综的描述

对一个由处于量子态$|\psi\rangle$($\langle\psi|\psi\rangle=1$)的样品构成的纯态系综测量任一力学量$A$，平均值由态矢量$|\psi\rangle$给出

$$\overline{A}=\langle\psi|A|\psi\rangle \tag{1.4.4}$$

由于全部力学量的平均值涵盖了微观体系状态的全部运动学信息，故我们可用态矢量$|\psi\rangle$描述一个纯态系综。

注意到态矢量空间S有如下的性质：对每一个归一化的态矢量$|\psi\rangle\in S$，都有S的正交归一基$\{|u_1\rangle,|u_2\rangle,\cdots\}$使$|u_1\rangle=|\psi\rangle$。在这个基下，线性算符$|\psi\rangle\langle\psi|$的矩阵元为$(|\psi\rangle\langle\psi|)_{\alpha\beta}=\langle u_\alpha|\psi\rangle\langle\psi|u_\beta\rangle=\delta_{\alpha 1}\delta_{\beta 1}$。于是

$$\langle\psi|A|\psi\rangle=A_{11}=\sum_{\alpha,\beta}A_{\alpha\beta}\delta_{\beta 1}\delta_{\alpha 1}\sum_{\alpha,\beta}A_{\alpha\beta}(|\psi\rangle\langle\psi|)_{\beta\alpha}=tr([A][|\psi\rangle\langle\psi|])$$

其中$[A]$、$[|\psi\rangle\langle\psi|]$为算符$A$、$|\psi\rangle\langle\psi|$在正交归一基$\{|u_1\rangle,|u_2\rangle,\cdots\}$下的矩阵，$tr$代表取方阵对角元的和。由§0.3知，线性算符的矩阵与基有关，但线性算符矩阵的迹(trace)与基无关，故我们可以谈论线性算符的迹，它等于算符在任一基下矩阵的迹，这是一个与基无关的概念。于是对一个由处于量子态$|\psi\rangle$($\langle\psi|\psi\rangle=1$)的样品构成的纯态系综

$$\overline{A}=\langle\psi|A|\psi\rangle=tr(A|\psi\rangle\langle\psi|) \tag{1.4.5}$$

现在有一个混态系综，它的样品处在量子态$|\psi_i\rangle$($i=1,2,\cdots,s$)的几率为$p_i>0$。

这里$|\psi_1\rangle,\cdots,|\psi_s\rangle$是 $s(\geq 2)$个不同的量子态，$\sum_{i=1}^{s} p_i = 1$，$\langle\psi_i|\psi_i\rangle = 1$，除此之外，不要求这 s 个不同的量子态$|\psi_i\rangle(1,2,\cdots,s)$满足其它条件，譬如正交性、线性无关性、完备性等。让我们看一看，对这样一个混态系综测量任一个力学量 A 的平均值是什么？

对处于量子态$|\psi_i\rangle$的样品，力学量 A 的平均值是$\langle\psi_i|A|\psi_i\rangle$。而样品处在量子态$|\psi_i\rangle$的几率是 p_i，故对混态系综测量 A 的平均值应为二次平均值

$$\overline{A} = \sum_{i=1}^{s} p_i\langle\psi_i|A|\psi_i\rangle \tag{1.4.6}$$

考虑到(1.4.5)的第二个等式，我们可以把上式重新写作

$$\overline{A} = \sum_{i=1}^{s} p_i tr(A|\psi_i\rangle\langle\psi_i|) = tr\Big[A\sum_{i=1}^{s} p_i(|\psi_i\rangle\langle\psi_i|)\Big]$$

引入密度算符(由于常被写成矩阵形式，也叫密度矩阵)

$$\rho = \sum_{i=1}^{s} p_i(|\psi_i\rangle\langle\psi_i|) \tag{1.4.7}$$

我们得到

$$\overline{A} = tr(A\rho) \tag{1.4.8}$$

我们看到密度算符可以给出混态系综所有力学量的平均值，而这涵盖了混态系综的全部运动学信息。所以我们用密度算符(1.4.7)来描述混态系综。由式(1.4.7)可知，ρ 由系综的制备历史$\{p_i,|\psi_i\rangle|i=1,2,\cdots,s\}$所确定。下面我们指出，反过来由 ρ 并不能唯一地确定系综的制备历史$\{p_i,|\psi_i\rangle|1\leq i\leq s\}$。考察下面三个混态系综。系综 1 的样品处在 S_x 的两个本征态$|+\rangle_x$和$|-\rangle_x$的几率都是 0.5，系综 2 的样品处在 S_y 的两个本征态$|+\rangle_y$和$|-\rangle_y$的几率都是 0.5，而系综 3 的样品处在 S_z 的两个本征态$|+\rangle_z$和$|-\rangle_z$的几率都是 0.5。三个系综的样品都是单个电子(只考虑自旋运动，不考虑空间运动)。让我们按式(1.4.7)写下它们的密度算符。

$$\rho_1 = \frac{1}{2}|+\rangle_{xx}\langle+| + \frac{1}{2}|-\rangle_{xx}\langle-| = \frac{1}{2}I$$

$$\rho_2 = \frac{1}{2}|+\rangle_{yy}\langle+| + \frac{1}{2}|-\rangle_{yy}\langle-| = \frac{1}{2}I \tag{1.4.9}$$

$$\rho_3 = \frac{1}{2}|+\rangle_{xx}\langle+| + \frac{1}{2}|-\rangle_{xx}\langle-| = \frac{1}{2}I$$

式中 I 是二维自旋空间的单位算符。最后一步的得出利用了正交归一基的完备性。我们看到尽管三种情况下系综的制备历史不同,但密度算符相同。这说明 ρ 不能唯一地确定系综的制备历史$\{p_i, |\psi_i\rangle | 1 \leq i \leq s\}$。

让我们再考察一个例子。两个二电子体系(只考虑自旋运动)的混态系综。系综 1 中的样品处在自旋单重态$|00\rangle = \frac{1}{\sqrt{2}}(|+\rangle\otimes|-\rangle - |-\rangle\otimes|+\rangle)$的几率是 1/2,处在自旋三重态$|10\rangle = \frac{1}{\sqrt{2}}(|+\rangle\otimes|-\rangle + |-\rangle\otimes|+\rangle)$的几率也是 1/2。按照式(1.4.7),它的密度算符为 $\rho_1 = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|10\rangle\langle 10| = \frac{1}{2}|+\rangle\langle+|\otimes|-\rangle\langle-| + \frac{1}{2}|-\rangle\langle-|\otimes|+\rangle\langle+|$。系综 2 中的样品处在态$|+\rangle\otimes|-\rangle$的几率为 1/2,处在态$|-\rangle\otimes|+\rangle$的几率也是 1/2。它的密度算符为 $\rho_2 = \frac{1}{2}|+\rangle\langle+|\otimes|-\rangle\langle-| + \frac{1}{2}|-\rangle\langle-|\otimes|+\rangle\langle+|$。注意到$|00\rangle$和$|10\rangle$都是纠缠态,而$|+\rangle\otimes|-\rangle$和$|-\rangle\otimes|+\rangle$都是非纠缠态,这两个系综竟有相同的密度算符,这是非常耐人寻味的。

上面的例子告诉我们,有时尽管两个混态系综的制备历史$\{p_i, |\psi_i\rangle | 1 \leq i \leq s\}$不一样,但它们仍有可能具有相同的密度算符。由于密度算符 ρ 可以给出一切力学量的测量分布,这包含了系综的全部运动学信息,而制备历史是无法测量的,我们在以后将把密度算符相同的系综看作同一系综,而不去计较它们的制备历史的同异。回想在上一小节我们是利用制备历史定义纯态系综($s=1$)与混态系综($s \geq 2$)的。若同一密度算符 ρ 可以有不同的制备历史,这种定义方法会不会引起歧义?答案是否定的:不会。事实上,若$|\phi\rangle$,$|\psi\rangle$和$|\psi_1\rangle, \cdots, |\psi_s\rangle$是归一化的态矢量,而$|\psi_1\rangle, \cdots, |\psi_s\rangle$是 $s(\geq 2)$个不同的量子态,$|\phi\rangle$和$|\psi\rangle$是不同的量子态,则利用 *Schwarz* 不等式可以证明

(i)$|\phi\rangle\langle\phi| \neq |\psi\rangle\langle\psi|$ ($|\phi\rangle, |\psi\rangle$线性无关,即它们是不同的量子态)

(ii)$|\phi\rangle\langle\phi| \neq \sum_{i=1}^{s} p_i |\psi_i\rangle\langle\psi_i|$,$(p_i > 0, \sum_{i=1}^{s} p_i = 1, |\psi_1\rangle, \cdots, \langle\psi_s|$是 s 个不同的量子态)

(上面两个算符不等式的左边的算符在态$|\phi\rangle$下的期待值等于 1,而右边的算符在态$|\phi\rangle$下的期待值严格小于 1,故两边算符不等)不等式(ii)告诉我们一个密度算符

不可能同时有混态制备历史和纯态制备历史。不等式(i)告诉我们纯态密度矩阵的制备历史是唯一的。

练习 1-4-1 设$\{|u_1\rangle,\cdots,|u_n\rangle\}$和$\{|v_1\rangle,\cdots,|v_n\rangle\}$是体系态矢量空间 S 的两个正交归一基。系综 1 的样品处在各量子态$|u_i\rangle$的几率都是$\frac{1}{n}$,而系综 2 的样品处在各量子态$|v_i\rangle$的几率都是$\frac{1}{n}$。试证 $\rho_1=\rho_2=\frac{1}{n}I$,其中 I 是 S 上的恒等算符。(这是混乱程度最大的混态)

练习 1-4-2 设$|\psi\rangle$是态矢量空间中归一化的态矢量。试证 $tr(|\psi\rangle\langle\psi|)=1$。

练习 1-4-3 求电子自旋方向的极角为(θ,φ)的态矢量,并计算电子自旋取向各向均等的混态系综的密度算符。

Ⅲ. 密度算符的性质

密度算符定义式(1.4.7)中的 s 是指系综中样品可能处于的不同量子态的个数,当 $s\geq 2$ 时我们得到混态系综,当 $s=1$ 时我们得到纯态系综。密度算符是作用在样品的态矢量空间 S 上的线性算符。下面让我们来研究它的性质。由式(1.4.7)可知,

(i)ρ 是厄米算符:$\rho^+=\rho$,

(ii)ρ 是半正定的:$\langle\psi|\rho|\psi\rangle\geq 0,\forall|\psi\rangle\in S$, (1.4.10)

(iii)$tr(\rho)=1$。

半正定的厄米算符的所有本征值为非负实数。反之,所有本征值为非负实数的厄米算符为半正定的。上面的条件(i)、(ii)、(iii)不仅是密度算符的必要条件而且还是密度算符的充分条件。事实上,若 ρ 是厄米算符,它一定有完备的正交归一化的特征向量组$|u_1\rangle,|u_2\rangle,\cdots$,$\rho$ 的谱分解为$\rho=\sum_i\lambda_i|u_i\rangle\langle u_i|$。若 ρ 是半正定的,则所有 $\lambda_i\geq 0$。若 $tr(\rho)=1$,则$\sum_i\lambda_i=1$。于是 ρ 可以看作这样一个系综的密度算符:其样品处在量子态$|u_i\rangle$的几率为 λ_i。

上一小节我们指出混态系综的密度算符 ρ_m 不能写作$\rho_m=|\psi\rangle\langle\psi|$的形式。因而 ρ_m 在自身表象中的矩阵至少有 2 个非零对角元:$\rho_m=\sum_i\lambda_i|u_i\rangle\langle u_i|$(至少有两项)。从而 $\rho_m^2=\sum_{i,j}\lambda_i\lambda_j|u_i\rangle\langle u_i|u_j\rangle\langle u_j|=\sum_i\lambda_i^2\langle u_i|u_i\rangle$。由于求和多于 1 项,所以 $0<\lambda_i<1,\lambda_i^2<\lambda_i$。注意到$|u_i\rangle\langle u_j|$构成 S 上线性算符的全体所成的线性空间 $EndS$ 的

基,我们断言 $\rho_m \neq \rho_m^2$。而对纯态系综的密度算符 $\rho_p = |\psi\rangle\langle\psi|$,不难看出 $\rho_p^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho_p$。于是系综密度算符具有幂等性 $\rho^2 = \rho$,是纯态的充要条件。

设 A 是量子体系的任一力学量,$|u_1\rangle, |u_2\rangle, \cdots$ 是 A 的一个完备的正交归一化的特征向量组:$A|u_\alpha\rangle = a_\alpha|u_\alpha\rangle$,$\langle u_\alpha|u_\beta\rangle = \delta_{\alpha\beta}$,$\sum_\alpha |u_\alpha\rangle\langle u_\alpha| = 1$。现考察对系综 ρ 测量 A 的平均值(1.4.8),我们采用 A – 表象。

$$\overline{A} = tr(A\rho) = \sum_{\alpha,\beta}\langle u_\alpha|A|u_\beta\rangle\langle u_\beta|\rho|u_\alpha\rangle = \sum_\alpha a_\alpha\langle u_\alpha|\rho|u_\alpha\rangle \tag{1.4.11}$$

可见$\langle u_\alpha|\rho|u_\alpha\rangle$代表对系综 ρ 测量力学量 A 时本征值 a_α 出现的几率 $P(a_\alpha)$。而$\langle u_\alpha|\rho|u_\alpha\rangle$可以写作 $tr(|u_\alpha\rangle\langle u_\alpha|\rho)$,于是

$$P(a_\alpha) = tr(|u_\alpha\rangle\langle u_\alpha|\rho) \tag{1.4.12}$$

把 $P(a_\alpha)$(测量 A 得到本征值 a_α 的几率)看成一个力学量,则它的算符是$|u_\alpha\rangle\langle u_\alpha|$。

练习 1-4-4 试证:系综 ρ 是纯态系综的另一充要条件为$[\rho]$的秩为 1。

Ⅳ.密度算符的演化

在Ⅰ、Ⅱ、Ⅲ小节我们介绍了如何引入密度算符 ρ 刻画纯态和混态系综(在某一时刻)的状态。这些内容属于运动学范畴。下面我们将考虑理论的动力学部分,即密度算符

$$\rho(t) = \sum_{i=1}^{s} p_i |\psi_i(t)\rangle\langle\psi_i(t)| \tag{1.4.13}$$

如何随时间演化(系综的状态如何随时间变化)。我们说过,一个系综的各个样品必须是处在相同外界条件下的同种量子体系,它们有相同的哈密顿量 $H(t)$,样品在 s 个不同量子态$|\psi_1(t)\rangle, \cdots, |\psi_s(t)\rangle$间的分布 $p_1, \cdots, p_s$ 不随时间变化。取式(1.4.13)两边对 t 的微商得

$$i\hbar\dot{\rho}(t) = \sum p_i(H(t)|\psi_i(t)\rangle\langle\psi_i(t)| - |\psi_i(t)\rangle\langle\psi_i(t)|H(t))$$

这里利用了哈密顿算符的厄米性 $H(t)^+ = H(t)$。上式可以写作

$$i\hbar\dot{\rho}(t) = H(t)\rho(t) - \rho(t)H(t) = [H(t), \rho(t)] \tag{1.4.14}$$

这就是密度算符 $\rho(t)$的演化方程。注意,这个方程与 $\rho(t)$的制备历史$\{p_i, |\psi_i\rangle|1 \leq$

$i \leq s\}$无关。它是关于变量 t 的一阶微分方程,其解由初值 $\rho(t_0)$唯一确定。我们说过,$\rho(t_0)$完全刻画了系综在 t_0 时刻的状态。按因果律,在给定的外界条件下,系综在 t_0 时刻的状态应唯一地确定它在以后各时刻的状态。所以 $\rho(t)$满足的运动方程(1.4.14)是一阶微分方程,这正是因果律的要求。

Ⅴ.约化密度算符

设有一个体系,处在量子态$|\Psi\rangle\!\rangle$($\langle\!\langle\Psi|\Psi\rangle\!\rangle = 1$)。我们只对它的一部分感兴趣,对体系的其余部分不感兴趣。把体系的这两部分分别叫做子体系Ⅰ和子体系Ⅱ。体系的位形空间显然是两个子体系位形空间的笛卡儿积 $C = C^{I} \times C^{II}$,从而体系的态矢量空间是两个子体系态矢量空间的张量积 $S = S^{I} \otimes S^{II}$。S 中两个向量的内积是这样定义的:使对所有非纠缠态$|u \otimes x\rangle\!\rangle$和$|v \otimes y\rangle\!\rangle$都有$\langle\!\langle u \otimes x | v \otimes y\rangle\!\rangle = \langle u|v\rangle\langle x|y\rangle$,其中$\langle u|v\rangle$是 S^{I} 中两个向量$|u\rangle$,$|v\rangle$的内积,$\langle x|y\rangle$是 S^{II}中两个向量$|x\rangle$,$|y\rangle$的内积。容易证明,当$\{|u_\alpha\rangle|1 \leq \alpha \leq m\}$和$\{|w_\rho\rangle|1 \leq \rho \leq n\}$分别是 S^{I} 和S^{II}的正交归一基时,$\{|u_\alpha\rangle \otimes |w_\rho\rangle (\equiv |u_\alpha\rangle \otimes |w_\rho\rangle\!\rangle)|1 \leq \alpha \leq m \& 1 \leq \rho \leq n\}$是 S 的正交归一基。记 $S(S^{I}, S^{II})$上线性算符的全体所构成的复线性空间为 $\mathbf{A}(\mathbf{A}^{I}, \mathbf{A}^{II})$。则有 $\mathbf{A} = \mathbf{A}^{I} \otimes \mathbf{A}^{II}$。$S(S^{I}, S^{II})$上两个线性算符 A, B 的内积是这样定义的

$$(A, B) = tr(AB^{+}) \tag{1.4.15}$$

这个定义满足内积空间的公理,使 $\mathbf{A}(\mathbf{A}^{I}, \mathbf{A}^{II})$成为复内积空间。由 §0.3 知,在态矢量空间的正交归一基下,线性算符 B 的厄米共轭 B^{+}的矩阵等于 B 的矩阵$[B]$的厄米共轭$[B]^{+}$:$[B^{+}] = [B]^{+}$。式(1.4.15)可以写作

$$(A, B) = \sum_{k,l} A_{kl} B^{*}_{kl} \tag{1.4.16}$$

可见$\{|u_\alpha\rangle\langle u_\beta||1 \leq \alpha, \beta \leq m\}$是 $\mathbf{A}^{I}$ 的正交归一基,$\{|w_\rho\rangle\langle w_\sigma||1 \leq \rho, \sigma \leq n\}$是 $\mathbf{A}^{II}$的正交归一基,而$\{|u_\alpha \otimes w_\rho\rangle\!\rangle\langle\!\langle u_\beta \otimes w_\sigma||1 \leq \alpha, \beta \leq m \& 1 \leq \rho, \sigma \leq n\}$是 $\mathbf{A}$ 的正交归一基。

现在让我们来考察当大体系处在量子态$|\Psi\rangle\!\rangle$时,子体系Ⅰ的力学量 A 的期待值。这时力学量 A 的算符应写作 $A \otimes 1^{II}$(1^{II}是 S^{II}上的恒等算符)。于是

$$\overline{A} = \langle\!\langle \Psi | A \otimes 1^{II} | \Psi \rangle\!\rangle = \sum_{\alpha\beta\rho\sigma} \langle\!\langle \Psi | u_\alpha \otimes w_\rho \rangle\!\rangle \langle\!\langle u_\alpha \otimes w_\rho | A \otimes 1^{II} | u_\beta \otimes w_\sigma \rangle\!\rangle \langle\!\langle u_\beta \otimes w_\sigma | \Psi \rangle\!\rangle$$

$$= \sum_{\alpha\beta} \langle u_\alpha \mid A \mid u_\beta \rangle \sum_{\rho} \langle\!\langle u_\beta \otimes w_\rho \mid \Psi \rangle\!\rangle \langle\!\langle \Psi \mid u_\alpha \otimes w_\rho \rangle\!\rangle \tag{1.4.17}$$

上式中的$\sum_{\rho}\langle\!\langle u_\beta \otimes w_\rho \mid \Psi \rangle\!\rangle \langle\!\langle \Psi \mid u_\alpha \otimes w_\rho \rangle\!\rangle$与 S^{II} 的正交归一基$\{\mid w_\rho\rangle 1 \leq \rho \leq n\}$的选择无关。而当 S^I 的正交归一基变换时，它与 S^I 上线性算符的β行α列矩阵元按同样的方式变换。所以，我们可以认为它是 S^I 上某个线性算符ρ^I的β行α列矩阵元

$$\langle u_\beta \mid \rho^I \mid u_\alpha \rangle \equiv \sum_{\rho} \langle\!\langle u_\beta \otimes w_\rho \mid \Psi \rangle\!\rangle \langle\!\langle \Psi \mid u_\alpha \otimes w_\rho \rangle\!\rangle \tag{1.4.18}$$

我们可以形式地把由上式定义的 ρ^I 写作

$$\rho^I = \sum_{\rho} \langle w_\rho \mid \Psi \rangle\!\rangle \langle\!\langle \Psi \mid w_\rho \rangle = tr^{II}(\mid \Psi \rangle\!\rangle \langle\!\langle \Psi \mid) \tag{1.4.19}$$

式中$\langle w_\rho \mid \Psi \rangle\!\rangle$是 S^I 中的向量，它与$\mid u_\beta\rangle$的内积为$\langle\!\langle u_\beta \otimes w_\rho \mid \Psi \rangle\!\rangle$。而$\langle\!\langle \Psi \mid w_\rho\rangle$是 S^I 上的线性函数，自变量为$\mid u_\alpha\rangle$时的函数值为$\langle\!\langle \Psi \mid u_\alpha \otimes w_\rho \rangle\!\rangle$。而式(1.4.17)可以重新写作

$$\overline{A} = \sum_{\alpha\beta} \langle u_\alpha \mid A \mid u_\beta \rangle \langle u_\beta \mid \rho^I \mid u_\alpha \rangle = tr^I(A\rho^I) \tag{1.4.20}$$

由算符 ρ^I 的定义式(1.4.18)可知

$$\langle u_\beta \mid \rho^I \mid u_\alpha \rangle = \langle u_\alpha \mid \rho^I \mid u_\beta \rangle^* \tag{1.4.21}$$

这表明 ρ^I 是厄米算符：$(\rho^I)^+ = \rho^I$。对任一$\mid \psi\rangle \in S^I$，以$\langle \psi \mid u_\beta\rangle\langle u_\alpha \mid \psi\rangle$乘式(1.4.18)并对 α, β 求和得

$$\langle \psi \mid \rho^I \mid \psi \rangle = \sum_{\rho} \langle\!\langle \psi \otimes w_\rho \mid \Psi \rangle\!\rangle \langle\!\langle \Psi \mid \psi \otimes w_\rho \rangle\!\rangle \geq 0 \tag{1.4.22}$$

这表明 ρ^I 是半正定的。令式(1.4.18)中的 $\alpha = \beta$ 并对 α 求和得

$$tr^I \rho^I = \langle\!\langle \Psi \mid \Psi \rangle\!\rangle = 1 \tag{1.4.23}$$

于是 ρ^{I} 是子体系 I 的密度算符，称之为约化密度算符。由 ρ^{I} 可以得到子体系 I 的所有力学量的平均值(见式(1.4.20))，这涵盖了子体系 I 的全部运动学信息。可以证明，尽管$\mid \Psi\rangle\!\rangle$唯一地确定 ρ^{I}，但 ρ^{I} 不能反过来确定$\mid \Psi\rangle\!\rangle$，后者包含了整个体系的运动学信息。当我们只对子体系 I 感兴趣时，ρ^{I} 提供了最恰当的描述手段。

作为一个例子，考察处在自旋单重态$\mid \Psi\rangle\!\rangle = \mid 00\rangle\!\rangle = \frac{1}{\sqrt{2}}(\mid + \rangle_1 \otimes \mid - \rangle_2 - \mid - \rangle_1 \otimes \mid + \rangle_2)$的二电子体系(不考虑空间运动，只考虑自旋运动)。让我们写下第一个电子的约化密度算符 ρ^{I}，并写下它在 s_{z1}—表象中的矩阵。按照等式(1.4.19)第一个电子的约化密度算符为 $\rho^{\mathrm{I}} = {}_2\langle + \mid \Psi\rangle\!\rangle \langle\!\langle \Psi \mid + \rangle_2 + {}_2\langle - \mid \Psi\rangle\!\rangle \langle\!\langle \Psi \mid - \rangle_2 = \frac{1}{2} \mid - \rangle_{11} \langle - \mid + \frac{1}{2} \mid$

$+\rangle_{11}\langle+|$。它在 s_{z1}—表象中的矩阵为$[\rho^{\mathrm{I}}]=\begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix}$。

下面让我们来考虑动力学问题。设子体系Ⅰ与子体系Ⅱ间的相互作用可以忽略

$$H(t)=H^{\mathrm{I}}(t)+H^{\mathrm{II}}(t) \tag{1.4.24}$$

我们要考察子体系Ⅰ的约化密度算符

$$\langle u_\beta|\rho^{\mathrm{I}}(t)|u_\alpha\rangle=\sum_\rho\langle\!\langle u_\beta\otimes w_\rho|\Psi(t)\rangle\!\rangle\langle\!\langle\Psi(t)|u_\alpha\otimes w_\rho\rangle\!\rangle \tag{1.4.25}$$

或者

$$\rho^{\mathrm{I}}(t)=\sum_\rho\langle w_\rho|\Psi(t)\rangle\!\rangle\langle\!\langle\Psi(t)|w_\rho\rangle=tr^{\mathrm{II}}(|\Psi(t)\rangle\!\rangle\langle\!\langle\Psi(t)|) \tag{1.4.26}$$

如何随时间演化。取式(1.4.25)对 t 的微商得

$$\begin{aligned}i\hbar\langle u_\beta|\dot{\rho}^{\mathrm{I}}(t)|u_\alpha\rangle=&\sum_\rho(\langle\!\langle u_\beta\otimes w_\rho|H(t)|\Psi(t)\rangle\!\rangle\langle\!\langle\Psi(t)|u_\alpha\otimes w_\rho\rangle\!\rangle)\\&-\langle\!\langle u_\beta\otimes w_\rho|\Psi(t)\rangle\!\rangle\langle\!\langle\Psi(t)|H(t)|u_\alpha\otimes w_\rho\rangle\!\rangle)\\=&\sum_{\beta'}\langle u_\beta|H^{\mathrm{I}}(t)|u_{\beta'}\rangle\langle u_{\beta'}|\rho^{\mathrm{I}}(t)|u_\alpha\rangle\\&-\sum_{\alpha'}\langle u_\beta|\rho^{\mathrm{I}}(t)|u_{\alpha'}\rangle\langle u_{\alpha'}|H^{\mathrm{I}}(t)|u_\alpha\rangle\end{aligned} \tag{1.4.27}$$

把这个矩阵元的等式写成算符的等式

$$i\hbar\dot{\rho}^{\mathrm{I}}(t)=[H^{\mathrm{I}}(t),\rho^{\mathrm{I}}(t)] \tag{1.4.28}$$

这个演化方程表明，t_0 时刻子体系的密度算符 $\rho^{\mathrm{I}}(t_0)$唯一地确定了 $t>t_0$ 时刻子体系Ⅰ的密度算符 $\rho^{\mathrm{I}}(t)$。要确定 $\rho^{\mathrm{I}}(t)$，除 $\rho^{\mathrm{I}}(t_0)$之外，不需要 $\Psi(t_0)$所提供的更多的信息。

上面的讨论提供了关于混态密度算符的另一种解释：即要考察的体系是一个处在量子态$|\Psi(t)\rangle\!\rangle$的大的孤立体系的子体系，这个子体系与大体系的其余部分间的相互作用很弱。在这种条件下，约化密度算符(1.4.25)或(1.4.26)是对感兴趣的体系的最好描述。

第 2 章　量子力学中的对称性分析

宇宙浩瀚无垠，世界变化万千。自然科学的任务，就是透过错综复杂的自然现象，探求支配现象的自然规律。只有掌握了自然规律，才能更深刻地了解自然现象，预见事物的发展，能动地改造世界。大千世界的各种事物为什么如此千变万化？为什么人们相信在这些千变万化的事物的背后，存在着支配它们的（相对于现象的复杂而言的）简单规律？准确地回答这类问题是很困难的。但我们和读者在这里作一些讨论，相信会是有益的。考察下面的事例：参观一个有 N 个景点（例如熊猫馆、海豚馆、狮虎山等）的动物园，若每个景点都看一遍，则有 $N!$ 个不同的路线。路线的个数随景点个数 N 急剧地增大。当 $N=3$ 时，有 6 个不同的路线。而当 $N=100$ 时，就有 $100!$ 个不同的路线。这是一个天文数字！读者都知道围棋、象棋、国际象棋等棋类运动，并为它们的变化无穷而赞叹。然而，支配它们的规则，只消几分钟就可以对任何一个初学者讲明白。围棋的复杂性来自其棋盘之大；象棋、国际象棋的复杂性来自其棋盘之大和棋子之多。棋类的棋盘格数和棋子个数的数量级才不过是 10^2，对局就已如此复杂。世界之大，世上事物错综复杂就是理所当然的了。下棋对局是由对弈双方行子序列构成，而下棋规则本质上是对每步行子方式的限定。物质运动，事物发展也是由基本过程的序列构成的，而自然规律，特别是物理学规律，本质上就是对基本过程的限定。经典电磁现象千变万化，但万变不离其宗——支配电磁场运动变化的 Maxwell 方程组。后者限定了相邻时空点上电磁场变化运动之间的关联。而相邻时空点上电磁场变化运动之间的关联可以看作是电磁运动的基本过程。这些事例使我们相信，尽管世界错综复杂，并不排除存在支配它们的简单规律。所谓规律，一定有或大或小的普适性。而普适性本身就意味着对称性。

事实上，对称性广泛地存在于自然界，而且被人们自觉或不自觉地用来认识自然现象和探索自然规律。例如，如果我们在某时某地观察到把浓硫酸加到白糖中会引

起激烈的化学反应,那么我们会期待,在另一个时刻另一地点把浓硫酸加到白糖中也会引起激烈的化学反应。实际上这就是我们在利用时间和空间的均匀性。Pierre Curie(居里夫人,Marie Curie 的丈夫)是最早认识到对称性分析对物理学的重要性的人之一。但是在物理学中大量地应用群表示论(对称性分析的数学)还是在量子力学建立之后。这点与量子体系的状态是用某个 Hilbert 空间中的非零向量描述的,满足态的叠加原理有关。由于 Weyl,Wigner 和 Racah 等人的开创性的工作,群论的语言,最终成了物理学家通用语言的重要组成部分。现在早已无人怀疑群论与对称性分析在量子力学中的重要性了。

阅读本章,要求读者了解群表示论的基本概念。当讨论要求较高深的群论知识时,读者可在本书的附录中找到它们。

§2.1 物理学中的对称性

在经典力学中,若体系的哈密顿量 H——它包含了体系的全部动力学信息——不依赖某个广义坐标 q_α,由正则运动方程

$$\begin{cases} \dfrac{\partial H}{\partial p_i} = \dot{q}_i \\ -\dfrac{\partial H}{\partial q_i} = \dot{p}_i \end{cases} \qquad \forall\, i = 1,\cdots,N \tag{2.1.1}$$

可知,与 q_α 共轭的正则动量 p_α 守恒:$p_\alpha = \mathrm{const}$。此外若

$$\begin{cases} p_i = f_i(t) \\ q_i = g_i(t) \end{cases} \qquad \forall\, i = 1,\cdots,N \tag{2.1.2}$$

是方程组(2.1.1)的一个解,把其中 q_α 的换成 $q_\alpha = g_\alpha(t) + c$($c$ 是任一实数),

$$\begin{cases} p_i = f_i(t) \\ q_i = g_i(t) + c\delta_{i\alpha} \end{cases} \qquad \forall\, i = 1,\cdots,N$$

仍然是正则方程的一个解。

如果哈密顿量不显含时间,由(2.1.1)可知 $\dfrac{\mathrm{d}}{\mathrm{d}t}H(p,q) = \dot{p}_i\dfrac{\partial H}{\partial p_i} + \dot{q}_i\dfrac{\partial H}{\partial q_i} = 0$,我们有

能量守恒。在上面的讨论中，我们没有利用哈密顿量的具体形式，没有具体求解正则运动方程，就得出了有用的结论。这些讨论显示了经典物理学中体系对称性与守恒量的关系，它们都是对称性分析的例子。

在大学的量子力学课程中，我们都学过，在中心力场中粒子的角动量守恒，而且粒子的能级与磁量子数无关。这些都与体系是各向同性的有关，但我们可能并没从对称性分析的高度来理解这一事实。

本章将专门讨论量子力学中的对称性分析。为了避免表述过于抽象，我们将从具体的例子谈起。

§2.2 空间平移

Ⅰ. 空间平移群

我们所说的空间平移 $\boldsymbol{\rho}$，实际是指由如下形式的对应

$$\boldsymbol{r} \mapsto \boldsymbol{r} + \boldsymbol{\rho}, \forall \boldsymbol{r} \in R^3 \tag{2.2.1}$$

所确定的三维实欧氏空间 R^3 到自身的映射。为避免混淆，我们以后也用 $T(\boldsymbol{\rho}): R^3 \to R^3$ 来表示这个映射：$T(\boldsymbol{\rho})\boldsymbol{r} \mapsto \boldsymbol{r} + \boldsymbol{\rho}$。显然空间平移 $\boldsymbol{\rho}$ 或 $T(\boldsymbol{\rho})$ 是 R^3 到自身的双射，但它并不是 R^3 上的线性算符。让空间相继经受平移 $\boldsymbol{\rho}$ 和平移 $\boldsymbol{\sigma}$ 的结果，等于让空间经受一个平移 $\boldsymbol{\sigma} + \boldsymbol{\rho}$，我们把后者定义为前两者的和。可见空间平移的二元合成是位移向量加法，显然是可换的。平移 $\mathbf{0}$ 是加法运算零元，平移 $-\boldsymbol{\rho}$ 显然是平移 $\boldsymbol{\rho}$ 的逆元。这表明全体空间平移的集合在平移加法运算下构成一个阿贝尔群。事实上它是一个三维李群，称作三维实欧氏空间的平移群，记作 $T_3(R)$。

Ⅱ. 空间平移算符(态矢量在平移下的变换)

让我们考察由一个无自旋粒子构成的体系。采用坐标表象。设体系处于由 $\psi(\boldsymbol{r})$ 描述的状态。现在我们把实验室做一空间平移 $\boldsymbol{\rho}$，体系将处于一个新的、由 $\psi'(\boldsymbol{r})$ 描写的状态。显然应该有

$$\psi'(\boldsymbol{r} + \boldsymbol{\rho}) = \psi(\boldsymbol{r}), \forall \boldsymbol{r} \in R^3 \tag{2.2.2}$$

这可以重新写作

$$\psi'(\boldsymbol{r}) = \psi(\boldsymbol{r} - \boldsymbol{\rho}), \forall\, \boldsymbol{r} \in R^3 \tag{2.2.2}$$

以 **S** 记体系的态矢量空间,则对应 $\psi \mapsto \psi'$, $\forall\, \psi \in \mathbf{S}$,在 **S** 上确定了一个算符 $D(\boldsymbol{\rho})$:

$$[D(\boldsymbol{\rho})\psi](\boldsymbol{r}) = \psi'(\boldsymbol{r}) = \psi'(\boldsymbol{r} - \boldsymbol{\rho}), \forall\, \boldsymbol{r} \in R^3 \tag{2.2.3}$$

不难证明算符 $D(\boldsymbol{\rho})$是线性的,而且是幺正的,此外,

$$D(\boldsymbol{\sigma} + \boldsymbol{\rho}) = D(\boldsymbol{\sigma}) + D(\boldsymbol{\rho}), \forall\, \boldsymbol{\rho}, \boldsymbol{\sigma} \in R^3 \tag{2.2.4}$$

这表明 D 是空间平移群 $T_3(R)$在态矢量空间 **S** 上的一个幺正表示。把(2.2.3)式的右边作 *Taylor* 展开,

$$[D(\boldsymbol{\rho})\psi](\boldsymbol{r}) = \psi(\boldsymbol{r} - \boldsymbol{\rho}) = \psi(\boldsymbol{r}) + (-\boldsymbol{\rho}) \cdot \nabla \psi(\boldsymbol{r}) + \frac{[(-\boldsymbol{\rho}) \cdot \nabla]^2}{2!}\psi(\boldsymbol{r}) + \cdots,$$

$$= \left\{1 + (-\boldsymbol{\rho}) \cdot \nabla + \frac{[(-\boldsymbol{\rho}) \cdot \nabla]^2}{2!} + \cdots\right\}\psi(\boldsymbol{r})$$

我们得到平移算符的具体形式

$$D(\boldsymbol{\rho}) = \exp(-\boldsymbol{\rho} \cdot \nabla) \tag{2.2.5}$$

三维李群 $T_3(R)$的表示 D 的无穷小算符为:

$$I_\alpha \equiv \frac{\partial}{\partial \rho_\alpha} D(\boldsymbol{\rho})\,|_{\boldsymbol{\rho}=0} = -\frac{\partial}{\partial x_\alpha} = -ip_\alpha \tag{2.2.6}$$

(我们看到,在坐标表象中,无穷小算符乘上 $i\hbar$ 正好是粒子的动量算符。这决不是偶然的巧合,而且与具体表象无关。稍后我们将阐明其原因)于是平移算符还可以表作

$$D(\boldsymbol{\rho}) = \exp(\rho_1 I_1 + \rho_2 I_2 + \rho_3 I_3) = \exp\left[-\frac{i}{\hbar}(\rho_1 p_1 + \rho_2 p_2 + \rho_3 p_3)\right] \tag{2.2.7}$$

从无穷小算符的具体形式(2.2.6)可以直接验证它们是反厄米的:

$$I_\alpha^+ = -I_\alpha, \forall\, 1 \leq \alpha \leq 3 \tag{2.2.8}$$

但更一般地,利用(2.2.7)式,由平移表示 D 的幺正性,也可得出其无穷小算符的反厄米性。反之,无穷小算符的反厄米性也保证了平移算符的幺正性。事实上,$D(\boldsymbol{\rho})^+ = D^{-1}(\boldsymbol{\rho}) \Leftrightarrow \exp(\rho_1 I_1^+ + \rho_2 I_2^+ + \rho_3 I_3^+) = \exp[-(\rho_1 I_1 + \rho_2 I_2 + \rho_3 I_3)] \Leftrightarrow I_\alpha^+ = -I_\alpha$, $\forall\, \alpha = 1,2,3$。由(2.2.6)式还可以直接看出无穷小算符两两可换:

$$[I_\alpha, I_\beta] = 0, \forall\, 1 \leq \alpha, \beta \leq 3 \tag{2.2.9}$$

但更一般地,空间平移群表示的无穷小算符两两可换与空间平移群是阿贝尔群有内

在联系。事实上，由

$$D(\rho_1,0,0)D(0,\rho_2,0)D(-\rho_1,0,0) = D(0,\rho_2,0)$$

对两边取混合偏导数，并令 $\rho_1=\rho_2=0$，就可得到

$$\left\{\begin{array}{l}\dfrac{\partial D(\rho_1,0,0)}{\partial\rho_1}\dfrac{\partial D(0,\rho_2,0)}{\partial\rho_2}D(-\rho_1,0,0)\\ +D(\rho_1,0,0)\dfrac{\partial D(0,\rho_2,0)}{\partial\rho_2}\dfrac{\partial D(-\rho_1,0,0)}{\partial\rho_1}\end{array}\right\}_{\rho_1=\rho_2=0}=0$$

此即$[I_1,I_2]=0$。

Ⅲ. 力学量算符在平移下的变换

在前面讨论态矢量在平移下如何变换时，我们采用了主动观点。下面为了考虑力学量算符在平移下如何变换，让我们暂时改用被动观点。

把实验室作一空间平移 $\boldsymbol{\rho}$ 而保持坐标系Σ不变(主动观点)，相当于保持实验室不动而把坐标系作一空间平移 $-\boldsymbol{\rho}$(被动观点)。以Σ'记新的坐标系。量子体系的同一个客观的态，若从坐标系Σ看是 ψ，则从坐标系Σ'看应是$\psi'=D(\boldsymbol{\rho})\psi$。让我们从这两个不同的坐标系观察体系的同一客观的力学量算符把体系的一个状态变成另一个状态的作用。设体系的同一客观的力学量在两个坐标系中对应的算符分别是 A 与A'。这一力学量对态的作用在两个坐标系中分别被表作 $\varphi=A\psi$ 和$\varphi'=A'\psi'$。考虑到 $\psi'=D(\boldsymbol{\rho})\psi$ 和$\varphi'=D(\boldsymbol{\rho})\varphi$，可得 $D(\boldsymbol{\rho})A\psi=A'D(\boldsymbol{\rho})\psi$。此式对所有的 $\psi\in\mathbf{S}$ 都成立，故有算符等式

$$A' = D(\boldsymbol{\rho})AD^{-1}(\boldsymbol{\rho}) \tag{2.2.10}$$

回到主动观点，上式可以理解作，当把实验室作一平移时，体系的力学量变成一个新的力学量，新老力学量算符所满足的变换关系。

例如，当 A 是粒子相对于老坐标系Σ原点的位置矢量时，则 A'是粒子相对于新坐标系Σ'原点的位置矢量：(读者可以验证下式)

$$D(\boldsymbol{\rho})\boldsymbol{r}D(-\boldsymbol{\rho}) = \boldsymbol{r}-\boldsymbol{\rho} \tag{2.2.11}$$

当 A 是粒子相对于老坐标系Σ的动量 $\boldsymbol{p}$，则 A'是粒子相对于新坐标系Σ'的动量，由于两个坐标系相对静止，且对应的坐标轴平行，A 与A'二者应一致。直接计算可得

$$D(\boldsymbol{\rho})(-i\hbar\nabla)D(-\boldsymbol{\rho}) = -i\hbar\nabla \tag{2.2.12}$$

我们愿意指出,变换关系(2.2.10)不仅对体系的力学量算符成立,而且对体系态矢量空间 **S** 上的任一算符都成立。其理由是:当 A 把 ψ 变成 $A\psi$ 时,则 A' 应把 $D(\boldsymbol{\rho})\psi$ 变成 $A'D(\boldsymbol{\rho})\psi = D(\boldsymbol{\rho})A\psi$, $\forall\ \psi\in\mathbf{S}$。

把平移算符作用于 A 的本征方程,$A|\alpha\rangle = \alpha|\alpha\rangle$,我们得到 $A'D(\boldsymbol{\rho})|\alpha\rangle = \alpha D(\boldsymbol{\rho})|\alpha\rangle$。这表明 A 的本征态 $|\alpha\rangle$,经受平移后,变作 $D(\boldsymbol{\rho})|\alpha\rangle$,它是 A' 的本征态,对应的本征值仍是 α。

Ⅳ. 体系的平移不变性与动量守恒

现在让我们考虑体系的哈密顿量——它包含了体系的全部动力学信息——在空间平移下变换。在非相对量子力学中,设粒子的哈密顿算符为

$$H = \frac{\boldsymbol{p}^2}{2\mu} + \boldsymbol{U}(\boldsymbol{r})$$

若把实验室作一空间平移 $\boldsymbol{\rho}$,则平移后的体系的哈密顿算符为

$$H' = D(\boldsymbol{\rho})HD^{-1}(\boldsymbol{\rho}) = \frac{\boldsymbol{p}^2}{2\mu} + \boldsymbol{U}(\boldsymbol{r}-\boldsymbol{\rho}) \tag{2.2.13}$$

一般而言,$\hat{H}'$ 与 $\hat{H}$ 不同。当粒子不受外场作用时(当体系是孤立体系时),$U(\boldsymbol{r})$ 为常数。这时我们有

$$D(\boldsymbol{\rho})\hat{H}D^{-1}(\boldsymbol{\rho}) = \hat{H},\ \forall\ \boldsymbol{\rho}\in T^3 \tag{2.2.14}$$

或者

$$D(\boldsymbol{\rho})\hat{H} = \hat{H}D(\boldsymbol{\rho}),\ \forall\ \boldsymbol{\rho}\in T^3 \tag{2.2.15}$$

这就是体系具有平移不变性。

考虑到表达式(2.2.7),式(2.2.15)等价于

$$\boldsymbol{p}H = H\boldsymbol{p} \tag{2.2.16}$$

我们看到,当且仅当体系具有平移不变性时,体系的动量守恒。

当体系是由 N 个粒子构成时,类似的讨论可得,作用在 N 个粒子体系态矢量空间 **S** 上的平移算符为

$$D(\boldsymbol{\rho}) = \exp\Big(-\boldsymbol{\rho}\cdot\sum_{\alpha=1}^{N}\nabla_\alpha\Big) \tag{2.2.17}$$

它们构成空间平移群 $T_3(R)$ 在 **S** 上的一个酉表示。其无穷小算符的 $i\hbar$ 倍恰好是体

系总动量算符。**S** 中的算符在空间平移 $\boldsymbol{\rho}$ 下按下面方式变化：

$$A' = D(\boldsymbol{\rho})AD^{-1}(\boldsymbol{\rho}) \tag{2.2.18}$$

特别是体系的哈密顿算符 $H = \sum_{\alpha=1}^{N} \frac{-\hbar}{2m_\alpha}\nabla_\alpha^2 + U(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_N)$ 按下面方式变化：

$$H' = D(\boldsymbol{\rho})HD(-\boldsymbol{\rho}) = \sum_{\alpha=1}^{N} \frac{-\hbar^2}{2m_\alpha}\nabla_\alpha^2 + U(\boldsymbol{r}_1-\boldsymbol{\rho},\cdots,\boldsymbol{r}_N-\boldsymbol{\rho}) \tag{2.2.19}$$

如果体系是孤立体系，势能只与体系内各粒子间的相对位置有关，因而在体系平移下势能不变：$U(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_N) = U(\boldsymbol{r}_1-\boldsymbol{\rho},\cdots,\boldsymbol{r}_N-\boldsymbol{\rho})$，从而体系的哈密顿算符在空间平移下不变。

当且仅当体系是空间平移不变的，体系的总动量 $\boldsymbol{P} = -i\hbar\sum_{\alpha=1}^{N}\boldsymbol{\nabla}_\alpha$ 与哈密顿算符 H 对易，我们有总动量守恒。

§2.3 空间绕固定点的转动

Ⅰ. 三维空间旋转群 $SO(3)$

在刚体运动学中我们证明过，具有固定点 O 的刚体的每一位移都可以通过绕某一过 O 点的固定轴的旋转来实现。如果不区分绕同一轴转角差为 2π 整数倍的两个旋转，则刚体绕固定点 O 的转动和刚体保持 O 点不动的位移之间是双方一一对应的。

我们用由 O 点引出的向量 $\boldsymbol{\vartheta}$ 来表示这样一个绕 O 点的旋转：绕转动轴 $\boldsymbol{\vartheta}$，按右手螺旋法则转 $|\boldsymbol{\vartheta}| \equiv \vartheta$ 角。有时也用 $\boldsymbol{\vartheta}$ 的三个分量 $(\vartheta_x,\vartheta_y,\vartheta_z)$ 来表示旋转。

让刚体相继经受两个绕 O 点转动，先 $\boldsymbol{\vartheta}_1$ 后 $\boldsymbol{\vartheta}_2$，其结果仍是刚体绕 O 点的一个转动（刚体的保持 O 点不动的一个位移），称之为旋转 $\boldsymbol{\vartheta}_1$ 与 $\boldsymbol{\vartheta}_2$ 的乘积，记作 $\boldsymbol{\vartheta}_2 * \boldsymbol{\vartheta}_1$。这里两个相继旋转的顺序是重要的。不难验证上面定义的旋转的乘法 $*$ 是满足结合律的：$\vartheta_3 * (\vartheta_2 * \vartheta_1) = (\vartheta_3 * \vartheta_2) * \vartheta_1$。因为等式两边都代表相继作三个转动 $\vartheta_1,\vartheta_2,\vartheta_3$。零旋转（不动）$\boldsymbol{0}$ 是乘法的单位元，旋转 $-\boldsymbol{\vartheta}$ 是旋转 $\boldsymbol{\vartheta}$ 的逆元。因而刚体

绕固定点 O 的旋转的全体，在上面定义的乘法运算下构成一个群，这是一个三维李群，称作三维欧氏空间旋转群，记作 $SO(3)$（我们将在下面说明采用这种记法的理由）。

以 $\mathbf{R}^3$ 记全体由固定点 O 引出的向量所构成的三维实欧氏空间。$\mathbf{R}^3$ 中的每一向量 $\boldsymbol{r}$，在绕 O 点的旋转 $\varphi \in SO(3)$ 作用下，变成 $\mathbf{R}^3$ 中的一个新的向量 $\boldsymbol{r}'$。对应

$$\boldsymbol{r} \mapsto \boldsymbol{r}', \forall\, \boldsymbol{r} \in \mathbf{R}^3$$

确定了 $\mathbf{R}^3$ 上的一个算符 $R(\varphi):\mathbf{R}^3 \to \mathbf{R}^3$，$R(\varphi)\boldsymbol{r} = \boldsymbol{r}'$。即使不利用 $R(\varphi)$ 的具体表达式，我们也可以判断它是线性算符：$R(\varphi)(\alpha\boldsymbol{a} + \beta\boldsymbol{b}) = \alpha R(\varphi)\boldsymbol{a} + \beta R(\varphi)\boldsymbol{b}$，而且它保持 $\mathbf{R}^3$ 中任意两个向量的标量积不变：$R(\varphi)\boldsymbol{a} \cdot R(\varphi)\boldsymbol{b} = \boldsymbol{a} \cdot \boldsymbol{b}$。数学上称保持实内积空间中任意两个向量的标量积不变的算符为实正交算符。$R(\varphi)$ 还保持向量间的矢积关系不变：$\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{c} \Rightarrow R(\varphi)\boldsymbol{a} \times R(\varphi)\boldsymbol{b} = R(\varphi)\boldsymbol{c}$（不是保持两个向量的矢量积不变：$\boldsymbol{a} \times \boldsymbol{b} = R(\varphi)\boldsymbol{a} \times R(\varphi)\boldsymbol{b}$）。特别是，保持任意坐标系的手征不变（右手坐标系转动后仍为右手坐标系，左手坐标系转动后仍为左手坐标系）。读者不难看出 $R(0) = 1$，$\mathbf{R}(\varphi) * R(\psi) = R(\varphi * \psi)$。这表明，若令三维空间旋转群的每一个元素，旋转 $\varphi \in SO(3)$，与 $\mathbf{R}^3$ 上的转动算符 $R(\varphi)$ 对应：$\varphi \mapsto R(\varphi)$，$\forall\, \varphi \in SO(3)$，则我们得到了三维空间旋转群 $SO(3)$ 在三维实欧氏空间 $\mathbf{R}^3$ 上的一个实正交表示。由于这是一个忠实表示，人们有时把旋转 φ 和三维空间 $\mathbf{R}^3$ 上的转动算符 $R(\varphi)$ 看成同一个事物。当选取 $\mathbf{R}^3$ 的正交归一基 $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\} \equiv \{\hat{i}, \hat{j}, \hat{k}\}$ 时，由等式

$$R(\varphi)\hat{e}_\alpha = \hat{e}_\rho R(\varphi)_{\rho\alpha} \tag{2.3.1}$$

所确定的转动算符 $R(\varphi)$ 的矩阵 $[R(\varphi)]$ 为实正交矩阵（即转置等于逆的实矩阵）：$[R(\varphi)]^T = [R(\varphi)]^{-1}$，且行列式为 1。事实上，由两个向量的标积在转动下不变可得，

$$\begin{aligned}\delta_{\alpha\beta} &= \langle \hat{e}_\alpha \mid \hat{e}_\beta \rangle = \langle R(\varphi)\hat{e}_\alpha \mid R(\varphi)\hat{e}_\beta \rangle = R(\varphi)_{\rho\alpha} R(\varphi)_{\sigma\beta} \langle \hat{e}_\rho \mid \hat{e}_\sigma \rangle \\ &= R(\varphi)_{\sigma\alpha} R(\varphi)_{\sigma\beta}\end{aligned}$$

而这就是矩阵等式 $[R(\varphi)]^T[R(\varphi)] = E_3$（$3 \times 3$ 单位矩阵）的分量表达式。再考虑到向量间的矢积关系在转动下不变：$R(\varphi)\boldsymbol{x} \times R(\varphi)\boldsymbol{y} = R(\varphi)(\boldsymbol{x} \times \boldsymbol{y})$，和向量混和积公式 $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$，我们有 $1 = \hat{e}_1 \cdot (\hat{e}_2 \times \hat{e}_3) = R(\varphi)\hat{e}_1 \cdot (R(\varphi)\hat{e}_2 \times R$

$(\boldsymbol{\varphi})\hat{e}_3) = R(\boldsymbol{\varphi})_{\alpha 1}\varepsilon_{\alpha\beta\gamma}R(\boldsymbol{\varphi})_{\beta 2}R(\boldsymbol{\varphi})_{\gamma 3} = \det[R(\boldsymbol{\varphi})]$，式中 $\varepsilon_{\alpha\beta\gamma}$是三维 Civita 符号。* 这说明了为什么把三维空间旋转群记作 $SO(3)$，读作三维特殊正交群(3-dimenssional special orthogonal group)，特殊是指行列式为 1。

特别是，旋转算符 $R(\theta_x,0,0)$的矩阵为$\begin{bmatrix}1 & 0 & 0\\ 1 & \cos\theta_x & -\sin\theta_x\\ 0 & \sin\theta_x & \cos\theta_x\end{bmatrix}$，$R(0,\theta_z,0)$的矩阵为$\begin{bmatrix}\cos\theta_y & 0 & \sin\theta_y\\ 0 & 1 & 0\\ -\sin\theta_y & 0 & \cos\theta_y\end{bmatrix}$，$R(0,0,\theta_z)$的矩阵为$\begin{bmatrix}\cos\theta_z & -\sin\theta_z & 0\\ \sin\theta_z & \cos\theta_z & 0\\ 0 & 0 & 1\end{bmatrix}$。相应的无穷小算符为

$$I_1^R = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{bmatrix}, I_2^R = \begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{bmatrix}, I_3^R = \begin{bmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \tag{2.3.2}$$

利用 Civita 符号，它们可以统一地表作

$$(I_\alpha^R)_{\beta\gamma} = -\varepsilon_{\alpha\beta\gamma} \tag{2.3.3}$$

直接验证可知，它们满足如下的对易关系

$$[I_\alpha^R, I_\beta^R] = \sum_\gamma \varepsilon_{\alpha\beta\gamma}I_\gamma^R \tag{2.3.4}$$

利用下图读者不难证明，$\boldsymbol{r}\in\mathbf{R}^3$ 经受旋转 $\boldsymbol{\varphi} = \hat{n}\varphi \in SO(3)$($\hat{n}$ 式单位向量)后变作

$$R(\boldsymbol{\varphi})\boldsymbol{r} = \boldsymbol{r}\cos\varphi + \hat{n}(\hat{n}\cdot\boldsymbol{r})(1-\cos\varphi) + (\hat{n}\times\boldsymbol{r})\sin\varphi \tag{2.3.5}$$

事实上，可以把向量 $\boldsymbol{r}$ 写成平行于转轴 $\hat{n}$ 的分量 $\hat{n}(\hat{n}\cdot\boldsymbol{r})$与垂直于转轴 $\hat{n}$ 的分量 $\boldsymbol{r}-\hat{n}(\hat{n}\cdot\boldsymbol{r})$之和：$\boldsymbol{r} = \hat{n}(\hat{n}\cdot\boldsymbol{r}) + [\boldsymbol{r}-\hat{n}(\hat{n}\cdot\boldsymbol{r})]$。在绕 $\hat{n}$ 轴逆时针转动 φ 角时，平行分量不动，而垂直分量变作$[\boldsymbol{r}-\hat{n}(\hat{n}\cdot\boldsymbol{r})]\cos\varphi + \hat{n}\times[\boldsymbol{r}-\hat{n}(\hat{n}\cdot\boldsymbol{r})]\sin\varphi = [\boldsymbol{r}-\hat{n}(\hat{n}\cdot\boldsymbol{r})]\cos\varphi + \hat{n}\times\boldsymbol{r}\sin\varphi$。

当$|\boldsymbol{\varphi}|\ll 1$时，上式亦可以近似地写作

$$R(\boldsymbol{\varphi})\boldsymbol{r} \approx \boldsymbol{r} + \boldsymbol{\varphi}\times\boldsymbol{r} \tag{2.3.6}$$

---

* Civita 符号 $\varepsilon_{\alpha\beta\gamma}$由下面等式所确定：$\varepsilon_{123}=1$；$\varepsilon_{\alpha\beta\gamma} = \delta_P\varepsilon_{P(\alpha)P(\beta)P(\gamma)}$，$\forall P\in S_3$，式中 $\delta_P = +1(-1)$，当 P 是偶(奇)置换。利用它，三阶行列式$|A|$可以表作$|A| = \varepsilon_{\alpha\beta\gamma}A_1^\alpha A_2^\beta A_3^\gamma$。

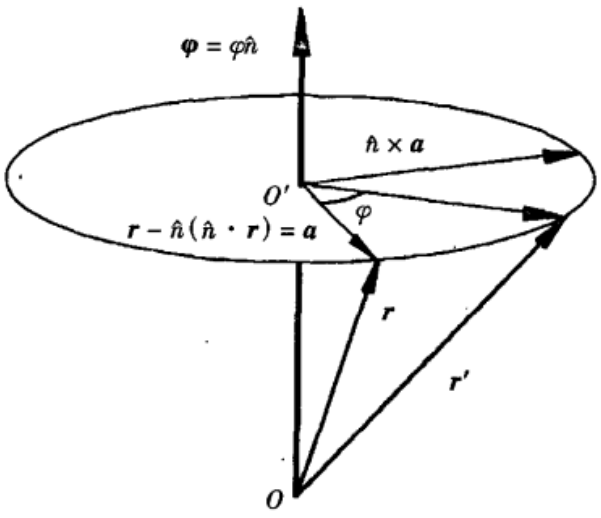

图 3.2-1

式中 $\boldsymbol{\varphi}$ 是无穷小角位移矢量。利用这个结果可以得到，对两个无穷小旋转 $\boldsymbol{\varphi}$ 和 $\boldsymbol{\psi}$，在忽视高阶小时有 $R(\boldsymbol{\varphi} * \boldsymbol{\psi})\boldsymbol{r} = R(\boldsymbol{\varphi})R(\boldsymbol{\psi})\boldsymbol{r} \approx R(\boldsymbol{\varphi})(\boldsymbol{r} + \boldsymbol{\psi} \times \boldsymbol{r}) \approx \boldsymbol{r} + \boldsymbol{\psi} \times \boldsymbol{r} + \boldsymbol{\varphi} \times \boldsymbol{r} \approx R(\boldsymbol{\varphi} + \boldsymbol{\psi})\boldsymbol{r}$

$$R(\boldsymbol{\varphi} * \boldsymbol{\psi}) \approx R(\boldsymbol{\varphi} + \boldsymbol{\psi}), \text{或者 } \boldsymbol{\varphi} * \boldsymbol{\psi} \approx \boldsymbol{\varphi} + \boldsymbol{\psi} \tag{2.3.7}$$

绕过 O 点的同一轴 $\hat{n}$ 的旋转的全体构成 $SO(3)$ 的一个一维子李群 $SO(2)$，它是一个阿贝尔群。其中乘法为

$$(\hat{n}\psi) * (\hat{n}\varphi) = \hat{n}(\psi + \varphi) \tag{2.3.8}$$

上式是精确的等式，对任意的 φ, ψ 都成立。

关于 $SO(3)$ 群，有一个很重要的事实是

$$\boldsymbol{\varphi} * \boldsymbol{\psi} * (-\boldsymbol{\varphi}) = R(\boldsymbol{\varphi})\boldsymbol{\psi} \tag{2.3.9}$$

即先绕 $\boldsymbol{\varphi}$ 轴转 $-\varphi$ 角，接着绕 $\boldsymbol{\psi}$ 轴转 ψ 角，最后绕 $\boldsymbol{\varphi}$ 轴转 φ 角，相当于绕 $R(\boldsymbol{\varphi})\boldsymbol{\psi}$ 轴转 ψ 角。让我们来证明这一事实。请参考下图，它是沿 $-\boldsymbol{\varphi}$ 方向的正投影图。

把有固定点 O 的刚体绕 $\boldsymbol{\varphi}$ 轴转 $-\varphi$ 角，本体坐标系从与空间坐标系 $\sum$ 重合的位置，变到 $\sum_1$，而向量 $R(\boldsymbol{\varphi})\boldsymbol{\psi}$ 变到 $\boldsymbol{\psi}$：$\sum \overset{R(-\boldsymbol{\varphi})}{\mapsto} \sum_1$，$R(-\boldsymbol{\varphi})\boldsymbol{\psi} \overset{R(-\boldsymbol{\varphi})}{\mapsto} \boldsymbol{\psi}$。从坐标系 $\sum$ 看向量 $R(\boldsymbol{\varphi})\boldsymbol{\psi}$ 与从坐标系 $\sum_1$ 看向量 $\boldsymbol{\psi}$ 是一样的，即向量 $R(\boldsymbol{\varphi})\boldsymbol{\psi}$ 在坐标系 $\sum$ 下的坐标，与向量 $\boldsymbol{\psi}$ 在坐标系 $\sum_1$ 下的坐标是一样的。接着把刚体绕 $\boldsymbol{\psi}$ 轴转 ψ 角，本体坐标系从 $\sum_1$ 变到 $\sum_2$，而向量 $\boldsymbol{\psi}$ 保持不变：$\sum_1 \overset{R(\boldsymbol{\psi})}{\mapsto} \sum_2$，$\boldsymbol{\psi} \overset{R(\boldsymbol{\psi})}{\mapsto} \boldsymbol{\psi}$。最后把刚体绕 $\boldsymbol{\varphi}$ 轴转 φ 角，坐标系 $\sum_1$ 变回到与空间坐标系 $\sum$ 重合的位置，本体坐标系从 $\sum_2$ 变到 $\sum'$，而向

量 $\boldsymbol{\psi}$ 变回到 $R(\boldsymbol{\varphi})\boldsymbol{\psi}$：$\sum_1 \overset{\boldsymbol{\varphi}}{\mapsto} \sum$，$\sum_2 \overset{\boldsymbol{\varphi}}{\mapsto} \sum'$，$\boldsymbol{\psi} \overset{\boldsymbol{\varphi}}{\mapsto} R(\boldsymbol{\varphi})\boldsymbol{\psi}$。从坐标系 $\sum_1$ 看坐标系 $\sum_2$ 和向量 $\boldsymbol{\psi}$，与从坐标系 $\sum$ 看坐标系 $\sum'$ 和向量 $R(\boldsymbol{\varphi})\boldsymbol{\psi}$ 完全一样。由坐标系 $\sum_1$ 绕 $\boldsymbol{\psi}$ 轴转 ψ 角得到坐标系 $\sum_2$，所以由坐标系 $\sum$ 绕 $R(\boldsymbol{\varphi})\boldsymbol{\psi}$ 轴转 ψ 角得到坐标系 $\sum'$。

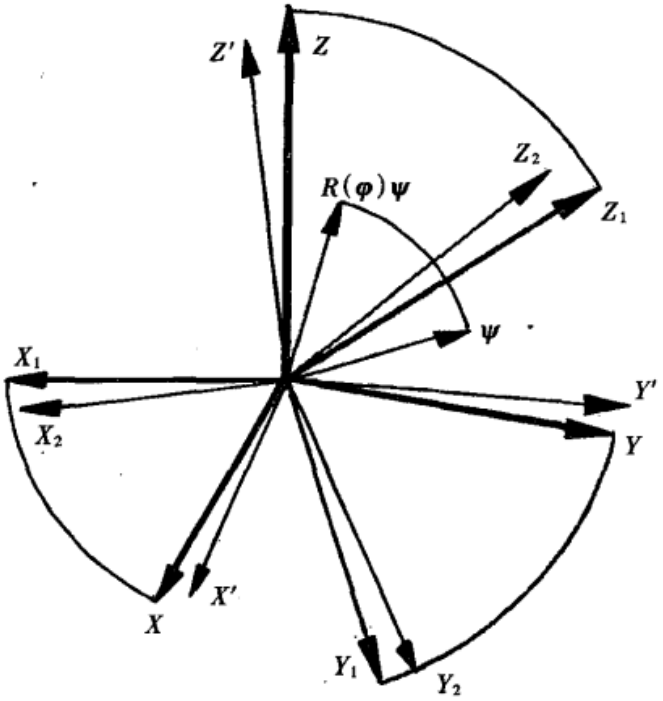

图 3.2-2

式(2.3.9)表明，$SO(3)$ 中的两个元素(两旋转)，若共轭则转角相等。反之，若 $SO(3)$ 中两旋转 $\boldsymbol{\psi}, \boldsymbol{\theta}$ 的转角相等：$|\boldsymbol{\psi}| = |\boldsymbol{\theta}|$，则因在 $SO(3)$ 中，总有旋转 $\boldsymbol{\varphi}$，使 $R(\boldsymbol{\varphi})\boldsymbol{\psi} = \boldsymbol{\theta}$，由(2.3.9)可知 $\boldsymbol{\theta} = R(\boldsymbol{\varphi})\boldsymbol{\psi} = \boldsymbol{\varphi} * \boldsymbol{\psi} * (-\boldsymbol{\varphi})$，$\boldsymbol{\psi}$ 与 $\boldsymbol{\theta}$ 共轭。这表明群 $SO(3)$ 中的共轭类是由全体转角相等的旋转构成的。

作为式(2.3.9)的一个特例，我们有

$$(\theta_x,0,0) * (0,\theta_y,0) * (-\theta_x,0,0) = (0,\theta_y\cos\theta_x,\theta_y\sin\theta_x) \qquad (2.3.10)$$

式(2.3.10)所表达的旋转间的关系是角动量算符对易关系的基础。很快我们就会看到这一点。

练习 2-3-1 试证：$R(\boldsymbol{\varphi})\boldsymbol{r} = \boldsymbol{r}\cos\varphi + \hat{n}(\hat{n}\cdot\boldsymbol{r})(1-\cos\varphi) + (\hat{n}\times\boldsymbol{r})\sin\varphi$

练习 2-3-2 试证：在正交归一基下 $\{\hat{i},\hat{j},\hat{k}\}$，$[R(\boldsymbol{\varphi})]=[R(\varphi_x,\varphi_y,\varphi_z]=$

$$\begin{bmatrix} \cos\varphi+\frac{\varphi_x\varphi_x}{\varphi^2}(1-\cos\varphi) & -\frac{\varphi_z}{\varphi}\sin\varphi+\frac{\varphi_x\varphi_Y}{\varphi^2}(1-\cos\varphi) & \frac{\varphi_Y}{\varphi}\sin\varphi+\frac{\varphi_x\varphi_z}{\varphi^2}(1-\cos\varphi) \\ \frac{\varphi_z}{\varphi}\sin\varphi+\frac{\varphi_Y\varphi_x}{\varphi^2}(1-\cos\varphi) & \cos\varphi+\frac{\varphi_Y\varphi_Y}{\varphi^2}(1-\cos\varphi) & -\frac{\varphi_x}{\varphi}\sin\varphi+\frac{\varphi_Y\varphi_z}{\varphi^2}(1-\cos\varphi) \\ -\frac{\varphi_Y}{\varphi}\sin\varphi+\frac{\varphi_z\varphi_x}{\varphi^2}(1-\cos\varphi) & \frac{\varphi_x}{\varphi}\sin\varphi+\frac{\varphi_z\varphi_Y}{\varphi^2}(1-\cos\varphi) & \cos\varphi+\frac{\varphi_z\varphi_z}{\varphi^2}(1-\cos\varphi) \end{bmatrix}$$

Ⅱ. 单个无旋粒子体系态矢量的转动

让我们考虑由一个无旋粒子构成的体系，其态矢量空间记作 **S**。采用坐标表象。设体系处于由 $\varphi(\boldsymbol{r})$所描述的状态。现在把实验室绕坐标原点 O 作一旋转 $\boldsymbol{\theta}$，显然描述体系新状态的波函数应是

$$\varphi'(\boldsymbol{r})=\varphi[R^{-1}(\boldsymbol{\theta})\boldsymbol{r}] \tag{2.3.11}$$

也就是说，新函数 φ' 在任一 $\boldsymbol{r}$ 点的函数值应等于老函数 φ 在 $R(-\boldsymbol{\theta})\boldsymbol{r}$ 点的函数值。对应 $\varphi\mapsto\varphi'$，$\forall\,\varphi\in\mathbf{S}$，确定了体系态矢量空间 S 上的一个线性算符 $D(\boldsymbol{\theta})$：

$$D(\boldsymbol{\theta})\varphi(\boldsymbol{r})=\varphi[R^{-1}(\boldsymbol{\theta})\boldsymbol{r}]\quad\forall\,\varphi\in\boldsymbol{S} \tag{2.3.12}$$

由此定义式可知

$$D(\boldsymbol{0})=1,D(\boldsymbol{\theta}_1)D(\boldsymbol{\theta}_2)=D(\boldsymbol{\theta}_1*\boldsymbol{\theta}_2),D^{-1}(\boldsymbol{\theta})=D^{+}(\boldsymbol{\theta}) \tag{2.3.13}$$

事实上，由 $D(\boldsymbol{0})\varphi(\boldsymbol{r})=\varphi(R^{-1}(\boldsymbol{0})\boldsymbol{r})=\varphi(\boldsymbol{r})$，可见 $D(\boldsymbol{0})=1$；而由

$$\varphi(\boldsymbol{r})\overset{D(\boldsymbol{\theta}_2)}{\longmapsto}\varphi(R^{-1}(\boldsymbol{\theta}_2)\boldsymbol{r})\overset{D(\boldsymbol{\theta}_1)}{\longmapsto}\varphi(R^{-1}(\boldsymbol{\theta}_2)R^{-1}(\boldsymbol{\theta}_1)\boldsymbol{r})$$

$$=\varphi(R^{-1}(\boldsymbol{\theta}_1*\boldsymbol{\theta}_2)\boldsymbol{r})=D(\boldsymbol{\theta}_1*\boldsymbol{\theta}_2)\varphi(\boldsymbol{r})$$

可知 $D(\boldsymbol{\theta}_1)D(\boldsymbol{\theta}_2)=D(\boldsymbol{\theta}_1*\boldsymbol{\theta}_2)$。为证明 $D(\boldsymbol{\theta})$的幺正性，考察内积 $\langle D(\boldsymbol{\theta})\varphi\mid D(\boldsymbol{\theta})\psi\rangle=\int_{R^3}d^3x\overline{\varphi}(R^{-1}(\boldsymbol{\theta})\boldsymbol{r})\psi(R^{-1}(\boldsymbol{\theta})\boldsymbol{r})$。作积分变量变换 $\boldsymbol{r}\mapsto\boldsymbol{r}'=\boldsymbol{R}^{-1}(\boldsymbol{\theta})\boldsymbol{r}$，Jacobi 行列式 $\left|\frac{\partial x}{\partial x'}\right|=\det[R(\boldsymbol{\theta})]=1$（转动下体积元大小不变），积分区域也不变。故上面的积分等于 $\int_{R^3}d^3x'\overline{\varphi}(\boldsymbol{r}')\psi(\boldsymbol{r}')=\langle\varphi\mid\psi\rangle$，从而 $D^{+}(\boldsymbol{\theta})D(\boldsymbol{\theta})=1$，即 $D^{+}(\boldsymbol{\theta})$是 $D(\boldsymbol{\theta})$的左逆。无穷维线性空间 **S** 上线性算符的左逆不一定是右逆，但是由于 $D(\boldsymbol{\theta})$有逆算符：$D(-\boldsymbol{\theta})=D^{-1}(\boldsymbol{\theta})$。以 $D^{-1}(\boldsymbol{\theta})$右乘 $D^{+}(\boldsymbol{\theta})D(\boldsymbol{\theta})=1$ 的两边可得 D^{-1}

$(\boldsymbol{\theta}) = D^{+}(\boldsymbol{\theta})$。这表明 D 是 $SO(3)$群在体系态矢量空间 $\mathbf{S}$ 上的一个幺正表示。

让我们来确定空间转动算符 $D(\boldsymbol{\theta})$的具体形式。由(2.3.12)知

$$D(\theta_x,0,0)\varphi(x,y,z) = \varphi(x,\cos\theta_x y + \sin\theta_x z, -\sin\theta_x y + \cos\theta_x z)$$

对上式对 θ_x 的微商,并令 $\theta_x = 0$,我们得到

$$I_x\varphi(x,y,z) = \left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right)\varphi(x,y,z) \tag{2.3.14}$$

式中

$$I_x = \frac{\partial}{\partial\theta_x}D(\theta_x,0,0)\mid_{\theta_x=0} \tag{2.3.15}$$

是旋转群 $SO(3)$的表示 D 的关于参数θ_x 无穷小算符。从式(2.3.14)得 I_x 的具体形式为

$$I_x = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z} = -il_x \tag{2.3.16}$$

式中 I_x 是粒子轨道角动量的 x 分量。可见表示 D 的无穷小算符 I_x 是反厄米的。由(2.3.8)和(2.3.13)可知

$$D(\delta,0,0)D(\theta_x,0,0) = D(\delta+\theta_x,0,0)$$

双方对 δ 求导,并令 $\delta = 0$ 得

$$I_xD(\theta_x,0,0) = \frac{\partial}{\partial\theta_x}D(\theta_x,0,0) \tag{2.3.17}$$

这个方程满足初始条件(零转动对应恒等算符 1)

$$D(0,0,0) = 1 \tag{2.3.18}$$

的解为

$$D(\theta_x,0,0) = \exp(I_x\theta_x) \tag{2.3.19}$$

类似地可以得到绕 y, z 轴旋转的算符:$D(0,\theta_y,0) = \exp(I_y\theta_y)$, $D(0,0,\theta_z) = \exp(I_z\theta_z)$,式中 $I_y = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x} = -il_y$, $I_z = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} = -il_z$ 是表示 D 的无穷小算符。

现在让我们来求绕任一固定轴 $\hat{n}(|\hat{n}| = 1)$旋转的算符 $D(\hat{n}\theta)$。把等式

$$D(\hat{n}\delta)D(\hat{n}\theta) = D[\hat{n}(\delta+\theta)] \tag{2.3.20}$$

的两边对 δ 求导,并令 $\delta = 0$ 得

$$I_nD(\hat{n}\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}D(\hat{n}\theta) \tag{2.3.21}$$

式中

$$I_n = \frac{\mathrm{d}}{\mathrm{d}\theta} D(\hat{n}\theta) \mid_{\theta=0} \tag{2.3.22}$$

方程(2.3.21)满足初始条件

$$D(\hat{n}0) = 1 \tag{2.3.23}$$

的解是

$$D(\hat{n}\theta) = \exp(I_n\theta) \tag{2.3.24}$$

下面让我们来确定 I_n 的具体形式。由(2.3.12)、(2.3.5)知

$$\begin{aligned} D(\boldsymbol{n}\theta)\varphi(\boldsymbol{r}) &= \varphi[R(-\hat{n}\theta)\boldsymbol{r}] \\ &= \varphi[\cos\theta\boldsymbol{r} + (1-\cos\theta)\hat{n}(\hat{n}\cdot\boldsymbol{r}) - \sin\theta(\hat{n}\times\boldsymbol{r})] \end{aligned}$$

对上面等式两边就 θ 求导,并令 $\theta=0$ 得

$$I_n\varphi(\boldsymbol{r}) = -(\hat{n}\times\boldsymbol{r})\cdot\nabla\varphi(\boldsymbol{r}) = -\hat{n}\cdot(\boldsymbol{r}\times\nabla)\varphi(\boldsymbol{r}) = -i\hat{n}\cdot\boldsymbol{l}\varphi(\boldsymbol{r})$$

上式对任意 $\varphi\in\mathbf{S}$ 都成立,故有

$$I_n = -\hat{n}\cdot(\boldsymbol{r}\times\nabla) = -i\hat{n}\cdot\boldsymbol{l} = -il_n \tag{2.3.25}$$

从而

$$D(\hat{n}\theta) = \exp[-\theta\hat{n}\cdot(\boldsymbol{r}\times\nabla)] = \exp[-i\theta\hat{n}\cdot\boldsymbol{l}] = \exp(-i\theta l_n)$$

或者

$$D(\theta_x,\theta_y,\theta_z) = \exp(\theta_xI_x + \theta_yI_y + \theta_zI_z) = \exp[-i(\theta_xl_x + \theta_yl_y + \theta_zl_z)] \tag{2.3.26}$$

这就是单个无旋粒子态矢量空间中转动算符的具体形式。

为确定 I_n 的具体形式,还有一种方法。当 D 是 $SO(3)$群的任一表示时都有

$$I_n \equiv \frac{\mathrm{d}}{\mathrm{d}\theta} D(\hat{n}\theta) \mid_{\theta=0} = \frac{\mathrm{d}}{\mathrm{d}\theta} D(n_x\theta, n_y\theta, n_z\theta) \mid_{\theta=0} = n_xI_x + n_yI_y + n_zI_z = \hat{n}\cdot\boldsymbol{I} \tag{2.3.27}$$

对我们现在讨论的表示 D,三个无穷小算符为 $I_\alpha = -il_\alpha(\alpha = x,y,z)$。于是我们又得到式(2.3.25)和(2.3.26)。

注意,式(2.3.20)对 $SO(3)$群的任一表示 D 都成立。由它可以得到式(2.3.24)。而式(2.3.27)也对 $SO(3)$群的任一表示 D 都成立。从而由式(2.3.24)和式(2.3.27)得到的等式

$$D(\boldsymbol{\theta}) = D(\theta_x,\theta_y,\theta_z) = \exp(\theta_x I_x + \theta_y I_y + \theta_z I_z) = \exp(\boldsymbol{\theta}\cdot\boldsymbol{I}) \quad (2.3.28)$$

对 $SO(3)$群的任一表示 D 都成立,它揭示了 $SO(3)$群任一表示的无穷小算符与有限转动算符间的关系。由此可见 $SO(3)$群的任一表示由它的无穷小算符所完全决定。

由 $SO(3)$群表示 D 的无穷小算符 I_x, I_y, I_z 的具体形式(2.3.16),可以验证

$$[I_\alpha, I_\beta] = \sum_\gamma \epsilon_{\alpha\beta\gamma} I_\gamma \quad (2.3.29)$$

我们再次强调,无穷小算符的对易关系(2.3.29)与旋转间的关系(2.3.10)有着内在的联系。事实上,对等式

$$D(\theta_x,0,0)D(0,\theta_y,0)D(-\theta_x,0,0) = D(0,\theta_y\cos\theta_x,\theta_y\sin\theta_x)$$

两边取混合偏导数$\frac{\partial^2}{\partial\theta_x\partial\theta_y}$,并令 $\theta_x=\theta_y=0$,就给出$[I_x, I_y] = I_z$。

Ⅲ.单个无旋粒子体系力学量算符的转动

现在让我们看一下单个无旋粒子的力学量算符,在旋转实验室时如何变化。从上一节的讨论我们知道,体系态矢量空间 $\mathbf{S}$ 上的每一线性算符 $A\in End\mathbf{S}$,应按下述方式变换

$$A \mapsto A' = D(\boldsymbol{\theta})AD(-\boldsymbol{\theta}) \quad (2.3.30)$$

当 A 是粒子的位置矢量$\boldsymbol{r}$ 在某一方向$\hat{e}$($\hat{e}$ 是单位向量)上的垂直投影 $\hat{e}\cdot\boldsymbol{r}$ 时,我们有

$$D(\boldsymbol{\theta})(\hat{e}\cdot\boldsymbol{r})D(-\boldsymbol{\theta}) = \hat{e}\cdot(R(-\boldsymbol{\theta})\vec{r}) = (\boldsymbol{R}(\boldsymbol{\theta})\hat{e})\cdot\boldsymbol{r} \quad (2.3.31)$$

事实上,对每一 $\varphi\in\mathbf{S}$,

$$\varphi(\boldsymbol{r}) \overset{D(-\boldsymbol{\theta})}{\mapsto} \varphi(R(\boldsymbol{\theta})\boldsymbol{r}) \overset{\hat{e}\cdot\boldsymbol{r}}{\mapsto} (\hat{e}\cdot\boldsymbol{r})\varphi(R(\boldsymbol{\theta})\boldsymbol{r}) \overset{D(\boldsymbol{\theta})}{\mapsto} (\hat{e}\cdot R(-\boldsymbol{\theta})\boldsymbol{r})\varphi(\boldsymbol{r})$$
$$= (R(\boldsymbol{\theta})\hat{e}\cdot\boldsymbol{r})\varphi(\boldsymbol{r})$$

最后一步利用了 $R(\boldsymbol{\theta})$是正交算符($R(\boldsymbol{\theta})\boldsymbol{a}\cdot R(\boldsymbol{\theta})\boldsymbol{b} = \boldsymbol{a}\cdot\boldsymbol{b}$)这一事实。上面的讨论表明,测量粒子位置矢量在方向 $\hat{e}$ 上的垂直投影的实验装置,在把实验室作一转动 $\boldsymbol{\theta}$ 后,将变成测量粒子位置矢量在方向 $R(\boldsymbol{\theta})\hat{e}$ 上的垂直投影的实验装置。当式(2.3.31)中的单位向量 $\hat{e}$ 是笛卡尔基矢 $\hat{e}_\alpha$ 时,我们有

$$D(\boldsymbol{\theta})x_\alpha D(-\boldsymbol{\theta}) = R(-\boldsymbol{\theta})_{\alpha\beta}x_\beta = x_\beta R(\boldsymbol{\theta})_{\beta\alpha} \quad (2.3.32)$$

以 $\hat{e}_\alpha$ 乘等式两边并对 α 求和得 $\sum_{\alpha=1}^{3}\hat{e}_\alpha D(\boldsymbol{\theta})x_\alpha D(-\boldsymbol{\theta}) = \sum_{\alpha=1}^{3}\hat{e}_\alpha(R(-\boldsymbol{\theta})_{\alpha\beta}x_\beta)$,

$D(\boldsymbol{\theta})\left(\sum_{\alpha=1}^{3}\hat{e}_\alpha x_\alpha\right)D(-\boldsymbol{\theta}) = \sum_{\alpha=1}^{3}\hat{e}_\alpha(R(-\boldsymbol{\theta})\boldsymbol{r})_\alpha$，此即

$$D(\boldsymbol{\theta})\boldsymbol{r}D(-\boldsymbol{\theta}) = R(-\boldsymbol{\theta})\boldsymbol{r} \tag{2.3.33}$$

在解读上式时请注意，态矢量空间 **S** 上的转动算符 $D(\boldsymbol{\theta}), D(-\boldsymbol{\theta})$从左右两边乘 **S** 上的算符 x, y, z。而三维实欧氏空间 R^3 上的转动算符 $R(\boldsymbol{\theta})$作用在 $\hat{\boldsymbol{i}}, \hat{\boldsymbol{j}}, \hat{\boldsymbol{k}}$ 上。取等式(2.3.33)两边与单位向量 $\hat{e}$ 的标量积得$\hat{e}\cdot D(\boldsymbol{\theta})\boldsymbol{r}D(-\boldsymbol{\theta}) = \hat{e}\cdot R(-\boldsymbol{\theta})\boldsymbol{r}$，此即式(2.3.31)。我们看到，等式(2.3.31)、(2.3.32)、(2.3.33)等价。

当 A 是粒子动量 $\boldsymbol{p}$ 在某一方向 $\hat{e}$（$\hat{e}$ 是单位向量）上的垂直投影 $\hat{e}\cdot\boldsymbol{p}$ 时，把实验室做一转动 $\boldsymbol{\theta}$ 后，它将变作动量 $\boldsymbol{p}$ 在 $R(\boldsymbol{\theta})\hat{e}$ 方向上的垂直投影。事实上，对每一 $\varphi \in \mathbf{S}$ 都有

$$\varphi(\boldsymbol{r}) \overset{D(-\boldsymbol{\theta})}{\longmapsto} \varphi(R(\boldsymbol{\theta})\boldsymbol{r}) \overset{\frac{\partial}{\partial x_\alpha}}{\longmapsto} \frac{\partial}{\partial x_\alpha}\varphi(R(\boldsymbol{\theta})\boldsymbol{r}) = \sum_{\beta=1}^{3}\frac{\partial(R(\boldsymbol{\theta})\boldsymbol{r})_\beta}{\partial x_\alpha}\frac{\partial\varphi(R(\boldsymbol{\theta})\boldsymbol{r})}{\partial(R(\boldsymbol{\theta})\boldsymbol{r})_\beta} =$$

$$\sum_{\beta=1}^{3}R(\boldsymbol{\theta})_{\beta\alpha}\frac{\partial\varphi(R(\boldsymbol{\theta})\boldsymbol{r})}{\partial(R(\boldsymbol{\theta})\boldsymbol{r})_\beta} \overset{D(-\boldsymbol{\theta})}{\longmapsto} \sum_{\beta=1}^{3}R(\boldsymbol{\theta})_{\beta\alpha}\frac{\partial\varphi(\boldsymbol{r})}{\partial x_\beta} =$$

$$\sum_{\beta=1}^{3}(R(\boldsymbol{\theta})\hat{e}_\alpha)_\beta\frac{\partial\varphi(\boldsymbol{r})}{\partial x_\beta} = (R(\boldsymbol{\theta})\hat{e}_\alpha)\cdot\nabla\varphi(\boldsymbol{r})$$

故有 $D(\boldsymbol{\theta})(\hat{e}_\alpha\cdot\nabla)D^{-1}(\boldsymbol{\theta}) = (R(\boldsymbol{\theta})\hat{e}_\alpha)\cdot\nabla = \hat{e}_\alpha\cdot(R(-\boldsymbol{\theta})\nabla)$，以 $-i\hbar$ 乘两边得

$$D(\boldsymbol{\theta})p_\alpha D^{-1}(\boldsymbol{\theta}) = p_\beta R(\boldsymbol{\theta})_{\beta\alpha} = R(-\boldsymbol{\theta})_{\alpha\beta}p_\beta \tag{2.3.34}$$

这就是单粒子动量分量算符在实验室转动下的变换方式。以 $\hat{e}_\alpha$ 乘上式两边并对 α 求和即得

$$D(\boldsymbol{\theta})\boldsymbol{p}D^{-1}(\boldsymbol{\theta}) = R(-\boldsymbol{\theta})\boldsymbol{p} \tag{2.3.35}$$

取任一单位向量 $\hat{e}$（不一定是笛卡尔基矢 $\hat{e}_\alpha$）与上式两边的标积即得

$$D(\boldsymbol{\theta})\hat{e}\cdot\boldsymbol{p}D(-\boldsymbol{\theta}) = (R(\boldsymbol{\theta})\hat{e})\cdot\boldsymbol{p} \tag{2.3.36}$$

而式(2.3.34)是式(2.3.36)的特例。可见式(2.3.36)与式(2.3.34)和式(2.3.35)都等价。

一般而言，当 A 是$f(\boldsymbol{r},\boldsymbol{p})$时，把实验室转动 $\boldsymbol{\theta}$ 后，它变作

$$\begin{aligned} D(\boldsymbol{\theta})f(\boldsymbol{r},\boldsymbol{p})D^{-1}(\boldsymbol{\theta}) &= f(D(\boldsymbol{\theta})\boldsymbol{r}D(-\boldsymbol{\theta}), D(\boldsymbol{\theta})\boldsymbol{p}D(-\boldsymbol{\theta})) \\ &= f(R^{-1}(\boldsymbol{\theta})\boldsymbol{r}, R^{-1}(\boldsymbol{\theta})\boldsymbol{p}) \end{aligned} \tag{2.3.37}$$

特别是，当 A 为体系的轨道角动量 $\boldsymbol{l} = \boldsymbol{r}\times\boldsymbol{p}$ 时，把实验室转动 $\boldsymbol{\theta}$ 时，它变作

$$D(\boldsymbol{\theta})(\boldsymbol{r}\times\boldsymbol{p})D^{-1}(\boldsymbol{\theta}) = D(\boldsymbol{\theta})\boldsymbol{r}D(-\boldsymbol{\theta})\times D(\boldsymbol{\theta})\boldsymbol{p}D(-\boldsymbol{\theta})$$
$$= R^{-1}(\boldsymbol{\theta})\boldsymbol{r}\times R^{-1}(\boldsymbol{\theta})\boldsymbol{p} = R^{-1}(\boldsymbol{\theta})(\boldsymbol{r}\times\boldsymbol{p})$$

最后一个等号成立是因为 $R(\boldsymbol{\theta})\boldsymbol{a}\times R(\boldsymbol{\theta})\boldsymbol{b} = R(\boldsymbol{\theta})(\boldsymbol{a}\times\boldsymbol{b})$。于是

$$D(\boldsymbol{\theta})\boldsymbol{l}D^{-1}(\boldsymbol{\theta}) = R(-\boldsymbol{\theta})\boldsymbol{l}$$

此即

$$D(\boldsymbol{\theta})\hat{e}\cdot\boldsymbol{l}D(-\boldsymbol{\theta}) = \hat{e}\cdot(R(-\boldsymbol{\theta})\boldsymbol{l}) = (R(\boldsymbol{\theta})\hat{e})\cdot\boldsymbol{l} \tag{2.3.38}$$

我们将把 $\boldsymbol{r},\boldsymbol{p},\boldsymbol{l}$ 这类力学量称作向量，它们在转动下按同样的方式变换(严格定义见下一小节的讨论)。

当 A 为体系的哈密顿量 H 时，把实验室做转动 $\boldsymbol{\theta}$，H 变作

$$H' = U(\boldsymbol{\theta})HD^{-1}(\boldsymbol{\theta}) = \frac{(R^{-1}(\boldsymbol{\vartheta})\boldsymbol{p})^2}{2\mu} + U(R^{-1}(\boldsymbol{\theta})\boldsymbol{r}) = \frac{\boldsymbol{p}^2}{2\mu} + U(R^{-1}(\boldsymbol{\theta})\boldsymbol{r}) \tag{2.3.39}$$

一般而言，$U(R^{-1}(\boldsymbol{\theta})\boldsymbol{r})$不同于 $U(\boldsymbol{r})$。但如果粒子处在以 O 点为中心的中心力场中，这时对每一绕 O 点的旋转 $\boldsymbol{\theta}$，都有 $U[R(\boldsymbol{\theta})r] = U(\boldsymbol{r})$，体系的哈密顿量在任意绕 O 点的转动 $\boldsymbol{\theta}$ 下都不变：$D(\boldsymbol{\vartheta})HD^{-1}(\boldsymbol{\vartheta}) = H$，亦即

$$\exp(\theta_x I_x + \theta_y I_y + \theta_z I_z)H = H\exp(\theta_x I_x + \theta_y I_y + \theta_z I_z)$$

易知此等式对每一旋转 $\boldsymbol{\theta}$ 都成立的充要条件是

$$[H, I_\alpha] = 0, \forall\,\alpha = x,y,z \tag{2.3.40}$$

这表明厄米算符 $l_\alpha = iI_\alpha$ 代表的物理量，当且仅当体系的哈密顿量具有旋转不变性时守恒。这个当且仅当体系具有旋转不变性时守恒的向量，应该被认同为无旋粒子的(轨道)角动量。这与以前我们从经典对应得到的结果是一致的。

现在从 $SO(3)$群表示的观点，考察体系力学量算符在转动下的变换。考虑体系态矢量空间 $\mathbf{S}$ 上的线性算符的全体所成的线性空间 $End\mathbf{S}$。对每一旋转 $\boldsymbol{\theta}\in SO(3)$，定义作用在线性空间 $End\mathbf{S}$ 上的线性算符 $\widetilde{D}(\boldsymbol{\theta}): End\mathbf{S}\to End\mathbf{S}$，使得

$$\widetilde{D}(\boldsymbol{\theta})A = D(\boldsymbol{\theta})AD(-\boldsymbol{\theta}), \forall\, A\in End\mathbf{S} \tag{2.3.41}$$

式中 D 是 $SO(3)$群在线性空间 $\mathbf{S}$ 上的一个表示。对应 $\boldsymbol{\theta}\mapsto\widetilde{D}(\boldsymbol{\theta})$，给出 $SO(3)$群在线性空间 $End\mathbf{S}$ 上的一个表示 $\widetilde{D}$。因为，利用 D 是旋转群 $SO(3)$在体系态矢量空间 $\mathbf{S}$ 上的一个表示，不难验证(留作练习)：$\widetilde{D}(\boldsymbol{0}) = 1$，$\widetilde{D}(\boldsymbol{\theta}*\boldsymbol{\varphi}) = \widetilde{D}(\boldsymbol{\theta})\widetilde{D}(\boldsymbol{\varphi})$。由式(2.

3.41)可得,表示 $\tilde{D}$ 的无穷小算符为

$$\tilde{I}_\alpha = \left.\frac{\partial \tilde{D}(\boldsymbol{\theta})}{\partial \theta^\alpha}\right|_{\vartheta=0} = \mathrm{adI}_\alpha, \forall\, \alpha = \mathrm{x,y,z} \tag{2.3.42}$$

式中 $I_\alpha = -il_\alpha$ 是表示 D 的无穷小算符。它们满足如下的对易关系:

$$[\tilde{I}_\alpha, \tilde{I}_\beta] = [adI_\alpha, adI_\beta] = ad[I_\alpha, I_\beta] = ad\sum_{\gamma=1}^{3}\varepsilon_{\alpha\beta\gamma}I_\gamma = \sum_{\gamma=1}^{3}\varepsilon_{\alpha\beta\gamma}adI_\gamma = \sum_{\gamma=1}^{3}\varepsilon_{\alpha\beta\gamma}\tilde{I}_\gamma$$

由式(2.3.28)可知

$$\tilde{D}(\boldsymbol{\theta}) = \exp((\theta_x adI_x + \theta_y adI_y + \theta_z adI_z)) \tag{2.3.43}$$

注意粒子的坐标算符 x_1, x_2, x_3 张成线性空间 $End\mathbf{S}$ 的一个三维子空间,记作 $\mathbf{X}$。由式(2.3.32)可得 $\tilde{D}(\boldsymbol{\theta})x_\alpha = \tilde{D}(\boldsymbol{\theta})x_\alpha D(-\boldsymbol{\theta}) = (R(\boldsymbol{\theta})\hat{e}_\alpha)\cdot \boldsymbol{r} = \sum_{\beta=1}^{3}R_{\beta\alpha}(\boldsymbol{\theta})x_\beta$.,表明这个由 x_1, x_2, x_3 张成的三维子空间 $\mathbf{X}$ 是表示 $\tilde{D}$ 的不变子空间;比较上式与等式 $R(\boldsymbol{\theta})\hat{e}_\alpha = \sum_{\beta=1}^{3}\hat{e}_\beta R_{\beta\alpha}(\boldsymbol{\theta})$,可知算符 $\tilde{D}(\boldsymbol{\theta})|_x$($\tilde{D}(\boldsymbol{\theta})$限制到 $\mathbf{X}$ 上)在基 x_1, x_2, x_3 下的矩阵等于 R^3 上的算符 $R(\boldsymbol{\theta})$在基 $\hat{i}, \hat{j}, \hat{k}$ 下的矩阵:

$$[\tilde{D}(\boldsymbol{\theta})|_X] = [R(\boldsymbol{\theta})] \tag{2.3.44}$$

$\tilde{D}|_X$ 与 R 是 $SO(3)$群的两个等价的三维表示。显然,表示 $\tilde{D}|_X$ 的无穷小算符在基 x_1, x_2, x_3 下的矩阵等于表示 R 的无穷小算符在基 $\hat{i}, \hat{j}, \hat{k}$ 下的矩阵:

$$[adI_\alpha] = \left.\frac{\partial[\tilde{D}(\boldsymbol{\theta})]}{\partial\theta^\alpha}\right|_{\vartheta=0} = \left.\frac{\partial[R(\boldsymbol{\theta})]}{\partial\theta^\alpha}\right|_{\vartheta=0} = [I_\alpha^R], \forall\, \alpha = 1,2,3 \tag{2.3.45}$$

有限转动算符的矩阵与无穷小转动算符的矩阵满足下面关系:

$$[R(\boldsymbol{\theta})] = \exp\Big(\sum_{\alpha=1}^{3}\theta^\alpha[I_\alpha^R]\Big) = [\tilde{D}(\boldsymbol{\theta})|_x] = \exp\Big(\sum_{\alpha=1}^{3}\theta^\alpha[I_\alpha^R]\Big) \tag{2.3.46}$$

练习 2-3-3 试证:$\tilde{D}(0) = 1, \tilde{D}(\boldsymbol{\theta} * \boldsymbol{\varphi}) = \tilde{D}(\boldsymbol{\theta})\tilde{D}(\boldsymbol{\varphi})$(利用 D 是旋转群 $SO(3)$在体系态矢量空间 $\mathbf{S}$ 上的一个表示)。

Ⅳ. 任意一微观体系的转动

上面讨论单个无旋粒子体系转动的方法,可以被推广用来讨论任一微观体系的

转动。以 $\mathbf{M}$ 记体系的态矢量空间。把实验室作一转动 $\boldsymbol{\theta}$，体系的每一态矢量 $\psi\in\mathbf{M}$，都应经受这一变换 $\psi\mapsto\psi'\in\mathbf{M}$。以 $D(\boldsymbol{\theta})$ 记 $\mathbf{M}$ 中的这样一个变换所确定的算符。显然把实验室相继做两次转动 $\boldsymbol{\theta}$ 和 $\boldsymbol{\theta}'$，使体系态矢量经受的变换 $\psi\overset{D(\boldsymbol{\theta})}{\mapsto}D(\boldsymbol{\theta})\psi\overset{D(\boldsymbol{\theta}')}{\mapsto}D(\boldsymbol{\theta}')D(\boldsymbol{\theta})\psi$，与把实验室做一次转动 $\boldsymbol{\theta}'*\boldsymbol{\theta}$，体系态矢量所经受的变换 $\psi\overset{D(\boldsymbol{\theta}'*\boldsymbol{\theta})}{\mapsto}D(\boldsymbol{\theta}'*\theta)\psi$ 结果相同：

$$D(\boldsymbol{\theta}'*\boldsymbol{\theta})=D(\boldsymbol{\theta}')D(\boldsymbol{\theta})$$

对处于任一状态下的体系进行任一观测的结果(测量仪器的读数)，不应受坐标系变换的影响。这表明对任意 $\psi,\varphi\in\mathbf{M}$，应有

$$|\langle\psi\mid\varphi\rangle|=|\langle D(\boldsymbol{\theta})\psi\mid D(\boldsymbol{\theta})\varphi\rangle|$$

Wigner 证明了(见下一节)微观体系态矢量空间中的保持向量内积模不变的变换，或者是幺正变换，或者是反幺正变换。按照 Wigner 的结果，$D(\boldsymbol{\theta})$ 只能是幺正的，或反幺正的。而对零转动 $\mathbf{0}$，$D(\mathbf{0})=1$ 是幺正的，而非反幺正的。如果要求 $D(\boldsymbol{\theta})$ 连续地依赖于 $\boldsymbol{\theta}$，则对任意旋转 $\boldsymbol{\theta}$，$D(\boldsymbol{\theta})$ 应是幺正的。于是，D 给出了旋转群 $SO(3)$ 在体系态矢量空间 $\mathbf{M}$ 上的一个幺正表示。它的三个无穷小算符

$$I_\alpha=\frac{\partial}{\partial\theta_\alpha}D(\boldsymbol{\theta})\mid_{\boldsymbol{\theta}=0}\tag{2.3.47}$$

应是反厄米的。代表任一有限转动 $\boldsymbol{\theta}$ 的算符是(参照式(2.3.28))$D(\boldsymbol{\theta})=\exp(\theta_xI_x+\theta_yI_y+\theta_zI_z)$。利用旋转间的关系(2.3.10)，重复上一小节中的讨论，我们可以得到无穷小算符间的对易关系：

$$[I_\alpha,I_\beta]=\sum_\gamma\varepsilon_{\alpha\beta\gamma}I_\gamma\tag{2.3.48}$$

体系的任一力学量算符 $A\in End\mathbf{M}$，在实验室转动下应按下面的方式变换 $A\mapsto D(\boldsymbol{\theta})AD^{-1}(\boldsymbol{\theta})\equiv\widetilde{D}(\boldsymbol{\theta})A$。态空间 $\mathbf{M}$ 上的全体线性算符所成的线性空间 $End\mathbf{M}$ 按 $SO(3)$ 群的表示 $\widetilde{D}$ 变换。表示 $\widetilde{D}$ 的无穷小算符为 $\tilde{I}_\alpha=adI_\alpha$，满足对易关系

$$[\tilde{I}_\alpha,\tilde{I}_\beta]=\sum_\gamma\varepsilon_{\alpha\beta\gamma}\tilde{I}_\gamma\tag{2.3.49}$$

下面我们将证明，厄米算符 $J_\alpha=iI_\alpha(\alpha=x,y,z)$，所代表的力学量，在实验室旋转下与坐标 x_α，动量 $p_\alpha=-i\hbar\frac{\partial}{\partial x_\alpha}$ 这类矢量算符变换方式一样：

$$\widetilde{D}(\boldsymbol{\theta})J_\alpha = D(\boldsymbol{\theta})J_\alpha D^{-1}(\boldsymbol{\theta}) = J_\beta R(\boldsymbol{\theta})_{\beta\alpha} \tag{2.3.50}$$

这等价于算符

$$\boldsymbol{J} = \hat{i}J_x + \hat{j}J_y + \hat{k}J_z \tag{2.3.51}$$

在转动下按如下方式变换：

$$D(\boldsymbol{\theta})\boldsymbol{J}D^{-1}(\boldsymbol{\theta}) = R^{-1}(\boldsymbol{\theta})\boldsymbol{J} \tag{2.3.52}$$

现在让我们来给出一个更一般的结果。先给出微观体系的向量算符的定义。

定义 设 **M** 是某个微观体系的态矢量空间，在实验室转动下，体系的态矢量按照 $SO(3)$群的幺正表示 $D: SO(3)\to GL(\mathbf{M})$变换；$V_x, V_y, V_z$ 是 **M** 上的厄米算符。若表示 D 的无穷小算符 I_x, I_y, I_z 与算符 V_x, V_y, V_z 满足对易关系

$$[I_\alpha, V_\beta] = \sum_{\gamma=1}^{3}\varepsilon_{\alpha\beta\gamma}V_\gamma, \forall\,\alpha,\beta = 1,2,3 \tag{2.3.53}$$

则称 $\boldsymbol{V} = \hat{i}V_x + \hat{j}V_y + \hat{k}V_z$ 为该体系的向量算符。

注意，这个定义并未要求 $V_x, V_y. V_z$ 是三个线性独立的线性算符，$\boldsymbol{\theta} = \hat{e}0 + \hat{j}0 + \hat{k}0$ 是向量算符。我们看到，$\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{l}$ 都是单粒子体系的向量算符。而旋转群 $SO(3)$在体系态矢量空间 **M** 上的幺正表示 D 的三个无穷小算符的 i 倍 $J_\alpha = iI_\alpha(\alpha = x, y, z)$，正好是该体系的向量算符(见式(2.3.48))。让我们来证明，下面两个条件等价，因而二者都可以当作向量算符的定义。

命题 $[I_\alpha, V_\beta] = \sum_{\gamma=1}^{3}\varepsilon_{\alpha\beta\gamma}V_\gamma, \forall\,\alpha,\beta = 1,2,3 \Leftrightarrow D(\boldsymbol{\theta})V_\alpha D^{-1}(\boldsymbol{\theta}) = \sum_{\beta=1}^{3}V_\beta R_{\beta\alpha}(\boldsymbol{\theta}), \forall\,\boldsymbol{\theta}\in SO(3)$

事实上，这两个条件可以重写为 $adI_\alpha(V_\beta) = \sum_{\gamma=1}^{3}\in_{\alpha\beta\gamma}V_\gamma, \forall\,\alpha,\beta = 1,2,3$ 和 $\widetilde{D}(\boldsymbol{\theta})V_\gamma = \exp(\sum_{\alpha=1}^{3}\theta_\alpha adI_\alpha)V_\gamma = \sum_{\alpha=1}^{3}V_\beta R_{\beta\gamma}(\boldsymbol{\theta}), \forall\,\boldsymbol{\theta}\in SO(3)$，或者(参考式(0.3.14))

$$adI_x[V_x, V_y, V_z] = [V_x, V_y, V_z]\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{bmatrix} = [V_x, V_y, V_z][adI_x], \cdots \tag{2.3.54}$$

和

$$\widetilde{D}(\boldsymbol{\theta})[V_x, V_y, V_z] = \exp\Big(\sum_{\alpha=1}^{3}\theta_\alpha adI_\alpha\Big)[V_x, V_y, V_z] = [V_x, V_y, V_z][R(\boldsymbol{\theta})] \tag{2.3.55}$$

注意,式(2.3.54)中出现的 3×3 矩阵$[adI_\alpha]$(由于未要求 V_x, V_y, V_z 是三个线性独立的线性算符,我们不说它是 adI_α 在基 V_x, V_y, V_z 下的矩阵)与 $SO(3)$群的表示 R 的无穷小算符在基 $\hat{i},\hat{j},\hat{k}$ 下的矩阵$[I_\alpha^R]$相同:$[I_\alpha^R]=[adI_\alpha]$,尽管它们代表不同的算符。

今假定式(2.3.54)成立。由式(0.3.15)知,$\exp\left(\sum_{\alpha=1}^{3}\theta_\alpha adI_\alpha\right)[V_x,V_y,V_z]=[V_x,V_y,V_z]\exp\left(\sum_{\alpha=1}^{3}\theta_\alpha[adI_\alpha]\right)$。利用式(2.3.46)得,$\exp\left(\sum_{\alpha=1}^{3}\theta_\alpha adI_\alpha\right)[V_x,V_y,V_z]=[V_x,V_y,V_z][R(\boldsymbol{\theta})]$。此即式(2.3.55)。

反之,若式(2.3.55)成立,取其两边对 θ_α 的偏导数,并令 $\boldsymbol{\theta}=0$,就得到式(2.3.54)。

我们看到,微观体系的向量算符有两种等价的定义。在讨论具体问题时,有时使用这种定义方便,有时使用另一种定义方便。读者应学会灵活使用两种定义。

以上的讨论指出了,当转动实验室时,微观体系的态矢量空间 $\mathbf{M}$ 按旋转群的一个幺正表示 D 变换。这个表示的无穷小算符 I_α 的 i 倍 $J_\alpha=iI_\alpha$ 是厄米的,它们代表的物理量在旋转下与位置矢量,动量矢量等向量算符的变换方式相同。

决定体系动力学性质的哈密顿量一般不一定是旋转不变量。它是旋转不变量的充要条件为

$$[H,J_\alpha]=0 \tag{2.3.56}$$

由此可见,在旋转下按向量方式变换的厄米算符(2.3.49),当且仅当体系具有旋转不变性时守恒。故应把它们认同为体系的总角动量算符。

下面我们将讨论由若干个自由度较小的子体系构成的体系的转动性质。后者的态矢量空间应是各子体系态矢量空间的张量积。

$$\mathbf{M}=\mathbf{M}_1\otimes\cdots\otimes\mathbf{M}_s \tag{2.3.57}$$

例如,电子的态矢量空间是它的轨道运动态矢量空间与自旋运动态矢量空间的张量积

$$\mathbf{M}_e=\mathbf{M}_e^o\otimes\mathbf{M}_e^s \tag{2.3.58}$$

而氢原子态矢量空间是电子、质子态矢量空间的张量积

$$\mathbf{M}_H=\mathbf{M}_e\otimes\mathbf{M}_p \tag{2.3.59}$$

在实验室(整个体系)旋转时,若子体系态矢量空间 $\mathbf{M}_\alpha$ 按旋转群的幺正的表示

D_α 变换，则体系的态矢量空间 $\mathbf{M}$ 应按直积表示

$$D = D_1 \otimes \cdots \otimes D_N \tag{2.3.60}$$

变换，也就是说，作用在体系态矢量空间 $\mathbf{M}$ 上的旋转算符 $D(\boldsymbol{\theta})$ 是各子体系旋转算符 $D_\alpha(\boldsymbol{\theta})$ 的直积

$$D(\boldsymbol{\theta}) = D_1(\boldsymbol{\theta}) \otimes \cdots \otimes D_S(\boldsymbol{\theta}) \tag{2.3.61}$$

D 是幺正表示。它的无穷小算符 I_j 是反厄米的，易见

$$I_\alpha = \left.\frac{\partial D(\boldsymbol{\theta})}{\partial \theta^\alpha}\right|_{\boldsymbol{\theta}=0} = I_{1\alpha} + I_{2\alpha} + \cdots + I_{S\alpha} \tag{2.3.62}$$

从而体系的角动量算符是各子体系角动量算符之和。

在下一章我们将讨论角动量的一般理论。

§2.4 态矢量空间的幺正变换和反幺正变换

用测量装置对微观体系的某个力学量进行测量，测量结果（即测量到的力学量的可能值和相应的几率）应与观测者（采用的参考坐标系或表象）无关。

设我们从两个参考坐标系 S 和 S' 来观察同一客观测量过程。体系所处的同一客观状态在 S 和 S' 中分别用态矢量 φ 和 φ' 表示。要测量的体系的力学量在 S 和 S' 中分别用算符 A 和 A' 表示。测量任一力学量的结果与观察者所采用的参考坐标系无关，意味着（ⅰ）A 与 A' 的本征值集合相同；（ⅱ）若 $Au_\alpha = \alpha u_\alpha$，则有 $A'u'_\alpha = \alpha u'_\alpha$；（ⅲ）对体系的每一客观状态都有 $|\langle u_\alpha | \varphi\rangle| = |\langle u'_\alpha | \varphi'\rangle|$。

由于上面所说的 A 可以是任一力学量，故 u_α 可以是任意的，于是，对所有 $\psi, \varphi \in \mathscr{M}$，都有

$$|\langle \varphi | \psi\rangle| = |\langle \varphi' | \psi'\rangle| \tag{2.4.1}$$

前面我们讨论过，在空间平移 $\boldsymbol{\rho}$ 或转动 $\boldsymbol{\theta}$ 下，体系的态矢量要经受一个幺正变换：$\psi \mapsto \psi' = T(\boldsymbol{\delta})\psi$，或 $\psi \mapsto \psi' = D(\boldsymbol{\theta})\psi$。态矢量空间上的任一算符 A 要经受相应的变换 $A \mapsto A' = T(\boldsymbol{\delta})AT(-\boldsymbol{\delta})$，或 $A \mapsto A' = D(\boldsymbol{\theta})AD(-\boldsymbol{\theta})$。不难看出，这些变换满足测量结果与观察者所采用的参考坐标系无关所要求的条件。但这并不是说，只有幺正变换满足这些条件。Wigner（1932，1959）证明了，Hilbert 空间 $\mathscr{M}$ 的满足条件

(2.4.1)的变换,总能写成幺正线性算符(幺正算符)与幺正的反线性算符(反幺正算符)与一个相因子的乘积。

在证明 Wigner 的定理之前,让我们首先解释一下什么叫幺正的反线性算符。

定义 设 $\mathscr{M}$ 是复线性空间,A 是 $\mathscr{M}$ 上的算符 $A:\mathscr{M}\to\mathscr{M}$。若对所有的 $\varphi,\psi\in\mathscr{M}$ 和 $\alpha,\beta\in C$ 都有

$$A(\alpha\varphi+\beta\psi)=\alpha^{*}A\varphi+\beta^{*}A\psi \tag{2.4.2}$$

则称 A 是 $\mathscr{M}$ 中的反线性算符。

在单个无旋粒子的态矢量空间(坐标表象)中,取波函数复共轭的算符 $K_o:K_o\varphi(\boldsymbol{r})=\varphi^{*}(\boldsymbol{r})$就是一个反线性算符。设 $u_1,u_2,$是复线性空间 $\mathscr{M}$ 的一个基,则 $\mathscr{M}$ 上的反线性算符 A 完全取决于 $Au_1,Au_2,\cdots$。又若 $f_1,f_2,\cdots$是复线性空间 $\mathscr{M}$ 中的一组向量,则存在唯一的反线性算符 B 使 $Bu_1=f_1,Bu_2=f_2,\cdots$。读者不难看到,若干个线性算符和偶数个反线性算符的乘积是线性算符;若干个线性算符和奇数个反线性算符的乘积是反线性算符。只有零算符既是线性算符又是反线性算符。

现在进一步假定,$\mathscr{M}$ 是复内积空间。设 A 是 $\mathscr{M}$ 上的反线性算符,$\psi,\varphi\in\mathscr{M}$。对固定的 φ,A 而言,$\langle A\psi\mid\varphi\rangle$是 ψ 的线性泛函(数$\langle A\psi\mid\varphi\rangle$线性地依赖于 ψ)。因而存在唯一 $\xi\in\mathscr{M}$,使$\langle\xi\mid\psi\rangle=\langle A\psi\mid\varphi\rangle$。显然 ξ 决定于 A 和 φ。对固定的 A 而言,向量 ξ 反线性地依赖于 φ。于是,对应 $\varphi\mapsto\xi,\forall\varphi\in\mathscr{M}$,确定了 $\mathscr{M}$ 上一个反线性算符,叫作反线性算符 A 的厄米共轭算符,记作 A^{+}。于是$\langle A^{+}\varphi\mid\psi\rangle=\langle A\psi\mid\varphi\rangle$,亦即

$$\langle\psi\mid A^{+}\mid\varphi\rangle=\langle\varphi\mid A\mid\psi\rangle \tag{2.4.3}$$

上面所举的反线性算符的例子 K_o 的厄米共轭算符是 K_o 自己:$K_o^{+}=K_o$。

若反线性算符 A 是自己的厄米共轭算符 $A^{+}=A$(或是自己的厄米共轭算符变号 $A^{+}=-A$),则称 A 是厄米的(或反厄米的)反线性算符。K_o 就是厄米的反线性算符。

若反线性算符 A,使 $A^{-1}=A^{+}$则称 A 为幺正的反线性算符(通常称作反幺正算符)。因而幺正的反线性算符 A 有如下性质

$$\langle A\varphi\mid A\psi\rangle=\langle\psi\mid\varphi\rangle,\forall\varphi,\psi\in\mathscr{M} \tag{2.4.4}$$

由于 $K_o^2=1$,可知 K_o 是幺正的反线性算符(反幺正算符)。

定理(Wigner) 设 $\mathscr{M}$ 是 Hilbert 空间,$T:\mathscr{M}\to\mathscr{M}$ 是一双射。若对所有 $\varphi,\psi\in\mathscr{M}$,都有

$$|\langle T\varphi \mid T\psi\rangle| = |\langle \psi \mid \varphi\rangle| \tag{2.4.5}$$

则存在幺正算符或反幺正算符 U;使 T 与 U 只差一个相因子:

$$T\varphi = C_\varphi U\varphi, \forall \varphi \in \mathscr{M} \tag{2.4.6}$$

式中 $|C_\varphi| = 1$。

证明:分以下几个步骤进行

(ⅰ)首先证明,若 $u_1, u_2, \cdots$ 是 $\mathscr{M}$ 的标准正交基,则 $Tu_1, Tu_2, \cdots$ 亦然。

$Tu_1, Tu_2, \cdots$ 的正交归一性由条件(2.4.5)和内积定义可以立即得到。事实上,由 $|\langle Tu_\alpha \mid Tu_\beta\rangle| = |\langle u_\alpha \mid u_\beta\rangle| = |\delta_{\alpha\beta}| = \delta_{\alpha\beta}$ 可知,当 $\alpha \neq \beta$ 时,$\langle Tu_\alpha \mid Tu_\beta\rangle = 0$,而 $\langle Tu_\alpha \mid Tu_\alpha\rangle = |\langle Tu_\alpha \mid Tu_\alpha\rangle| = 1$。合并写作 $\langle Tu_\alpha \mid Tu_\beta\rangle = \delta_{\alpha\beta}$(正交归一性)。现证其完备性:

由 T 是满射,$\mathscr{M}$ 中每一矢量 ψ 都可以写作 $\psi = T\varphi(\varphi \in \mathscr{M})$,可以证明矢量 $T\varphi - \sum_\alpha Tu_\alpha\langle Tu_\alpha \mid T\varphi\rangle$ 与自身的内积为零,事实上,

$$\langle T\varphi - \sum_\alpha Tu_\alpha\langle Tu_\alpha \mid T\varphi\rangle \mid T\varphi - \sum_\beta Tu_\beta\langle Tu_\beta \mid T\varphi\rangle\rangle = \langle T\varphi \mid T\varphi\rangle - \sum_\beta \langle T\varphi \mid Tu_\beta\rangle\langle Tu_\beta \mid T\varphi\rangle - \sum_\alpha \langle Tu_\alpha \mid T\varphi\rangle\langle T\varphi \mid Tu_\alpha\rangle + \sum_\alpha\sum_\beta \langle Tu_\alpha \mid Tu_\beta\rangle\langle T\varphi \mid Tu_\alpha\rangle\langle Tu_\beta \mid T\varphi\rangle$$

$$= \langle \varphi \mid \varphi\rangle - \sum_\beta \langle \varphi \mid u_\beta\rangle\langle u_\beta \mid \varphi\rangle - \sum_\alpha \langle u_\alpha \mid \varphi\rangle\langle \varphi \mid u_\alpha\rangle + \sum_\alpha \langle \varphi \mid u_\alpha\rangle\langle u_\alpha \mid \varphi\rangle$$

$$= \langle \varphi \mid \varphi - \sum_\beta u_\beta\langle u_\beta \mid \varphi\rangle\rangle = \langle \varphi \mid 0\rangle = 0,(\text{利用了 } \varphi = \sum_\alpha u_\alpha\langle u_\alpha \mid \varphi\rangle) \text{ 此即}$$

$$T\varphi = \sum_\alpha Tu_\alpha\langle Tu_\alpha \mid T\varphi\rangle \tag{2.4.7}$$

故知 $\mathscr{M}$ 中的每一矢量 Ψ 都可以写成 $Tu_1, Tu_2, \cdots$ 的线性组合:

$$\psi = \sum_\alpha Tu_\alpha\langle Tu_\alpha \mid \psi\rangle \tag{2.4.8}$$

即 $Tu_1, Tu_2, \cdots$ 是完备的,所以是 $\mathscr{M}$ 的标准正交基。

把 $T\varphi$ 向 $Tu_1, Tu_2, \cdots$ 的线性展开式(2.4.7)与 φ 向 $u_1, u_2, \cdots$ 的线性展开式

$$\varphi = \sum_\alpha u_\alpha\langle u_\alpha \mid \varphi\rangle \tag{2.4.9}$$

比较,可见对应的展开系数模相同。

(ⅱ)利用这一结果,对所有 $\alpha \geqslant 2$,考虑

$$T(u_1 + u_\alpha) = a_{1\alpha}Tu_1 + b_{1\alpha}Tu_2 = a_{1\alpha}\left(Tu_1 + \frac{b_{1\alpha}}{a_{1\alpha}}Tu_\alpha\right)$$

注意，$|a_{1\alpha}| = |b_{1\alpha}| = 1$。令

$$u'_1 = Tu_1, u'_\alpha = \frac{b_{1\alpha}}{a_{1\alpha}} Tu_\alpha, \forall \alpha \geq 2 \tag{2.4.10}$$

于是对上面式(2.4.10)定义的标准正交基 $u'_1, u'_2, \cdots$ 有

$$T(u_1 + u_\alpha) = \alpha_{1\alpha}(u'_1 + u'_\alpha) \tag{2.4.11}$$

现在我们来证明，对所有 $\alpha < \beta$

$$T(u_\alpha + u_\beta) = \alpha_{\alpha\beta}(u'_\alpha + u'_\beta) \tag{2.4.12}$$

式中 $|\alpha_{\alpha\beta}| = 1$。（当 $\alpha = 1$ 时，此即式(2.4.11)，故只须证明 $\alpha > 1$ 的情况）

事实上，

$$T(u_1 + u_\alpha + u_\beta) = au'_1 + bu'_\alpha + cu'_\beta \tag{2.4.13}$$

式(2.4.11)与式(2.4.13)内积的模为 $2 = |a + b|$

考虑到 $|a| = |b| = 1$，把上式写作

$$2 = \left|1 + \frac{b}{a}\right| \tag{2.4.14}$$

可得 $\frac{b}{a} = 1$ 即 $a = b$。

同理，取 $T(u_1 + u_\beta) = a_{1\beta}(u'_1 + u'_\beta)$ 与式(2.4.13)的内积的模，可得 $2 = |a + c| = \left|1 + \frac{c}{a}\right|$

可知 $a = c$，故(2.4.13)式可写作

$$T(u_1 + u_\alpha + u_\beta) = a(u'_1 + u'_\alpha + u'_\beta) \tag{2.4.15}$$

取 $T(u_\alpha + u_\beta) = d_{\alpha\beta}(u'_\alpha + hu'_\beta)$ 与上式内积的模得 $2 = |1 + h|$。考虑到 $|h| = 1$，故有式 $h = 1$。从而式(2.4.12)得证。

(ⅲ)我们来证明，对任一

$$\psi = a_\alpha u_\alpha + a_\beta u_\beta + a_\gamma u_\gamma + \cdots \in \mathscr{M}, \tag{2.4.16}$$

都有

$$T\psi = A_\psi(a_\alpha u'_\alpha + a_\beta u'_\beta + a_\gamma u'_\gamma + \cdots) \tag{2.4.17}$$

或者

$$T\psi = B_\psi(a^*_\alpha u'_\alpha + a^*_\beta u'_\beta + a^*_\gamma u'_\gamma + \ldots) \tag{2.4.18}$$

式中 $|A_\psi| = |B_\psi| = 1$。

当 ψ 为零时,命题显然成立。现讨论 $\psi \neq 0$ 的情况。设式(2.4.16)右边只包含系数非零的项。右边只包含一项时,命题显然成立。我们讨论多于一项的情况。把 $T\psi$ 展开

$$T\psi = T(a_\alpha u_\alpha + a_\beta u_\beta + a_\gamma u_\gamma + \cdots) = a'_\alpha u'_\alpha + a'_\beta u'_\beta + a'_\gamma u'_\gamma + \cdots \quad (2.4.19)$$

注意

$$| a_\alpha | = | a'_\alpha | \neq 0, | a_\beta | = | a'_\beta | \neq 0 \quad (2.4.20)$$

取式(2.4.12)与(2.4.19)内积模得

$$| a_\alpha + a_\beta | = | a'_\alpha + a'_\beta |$$

即

$$\left| 1 + \frac{a_\beta}{a_\alpha} \right| = \left| 1 + \frac{a'_\beta}{a'_\alpha} \right| \quad (2.4.21)$$

注意 $\left| \frac{a_\beta}{a_\alpha} \right| = \left| \frac{a'_\beta}{a'_\alpha} \right|$,式(2.4.21)要求

$$\frac{a'_\beta}{a'_\alpha} = \frac{a_\beta}{a_\alpha} \qquad \left(\frac{a'_\alpha}{a_\alpha} = \frac{a'_\beta}{a_\beta} \right) \quad (2.4.22)$$

或者

$$\frac{a'_\beta}{a'_\alpha} = \frac{a^*_\beta}{a^*_\alpha} \qquad \left(\frac{a'_\alpha}{a^*_\alpha} = \frac{a'_\beta}{a^*_\beta} \right) \quad (2.4.23)$$

注意等式(2.4.22)和(2.4.23)有时可以同时成立。出现在展开式(2.4.16)和(2.4.19)中的所有(非零)系数 $a_\alpha, a_\beta, a'_\alpha, a'_\beta (\alpha \neq \beta)$ 至少满足等式(2.4.22)、(2.4.23)之一。

为证在等式(2.4.22)和(2.4.23)二者之中,至少有一个对所有的 $\alpha \neq \beta$ 都成立,只须证明同时有 $\frac{a'_\mu}{a_\mu} \neq \frac{a'_v}{a_v} (\mu < v)$ 和 $\frac{a'_\rho}{a^*_\rho} \neq \frac{a'_\sigma}{a^*_\sigma} (\rho < \sigma)$ 将导致矛盾。由 $\frac{a'_\mu}{a_\mu} \neq \frac{a'_v}{a_v}$ 推出 $\frac{a'_\mu}{a^*_\mu} = \frac{a'_v}{a^*_v}$,由 $\frac{a'_\rho}{a^*_\rho} \neq \frac{a'_\sigma}{a^*_\sigma}$ 推出 $\frac{a'_\rho}{a_\rho} = \frac{a'_\sigma}{a_\sigma}$。分以下各种情况讨论:(a)当 $\mu = \rho, v = \sigma$ 时,等式 $\frac{a'_\rho}{a_\rho} = \frac{a'_\sigma}{a_\sigma}$ 就是 $\frac{a'_\mu}{a_\mu} = \frac{a'_v}{a_v}$,与不等式 $\frac{a'_\mu}{a_\mu} \neq \frac{a'_v}{a_v}$ 矛盾;(b)当 $\mu = \rho, v \neq \sigma$ 时,有 $\frac{a'_\sigma}{a^*_\sigma} \neq \frac{a'_v}{a^*_v}$ 和 $\frac{a'_\sigma}{a_\sigma} \neq \frac{a'_v}{a_v}$,这与式(2.4.21)矛盾;四个下标取三个不同值的其他情况,证明方法类似;(c)当四个

下标无重复时：(1)若$\frac{a'_\rho}{a_\rho}=\frac{a'_v}{a_v}$，则$\frac{a'_\rho}{a_\rho}\neq\frac{a'_\mu}{a_\mu}\neq\frac{a'_\sigma}{a_\sigma}$，从而$\frac{a'_\mu}{a^*_\mu}=\frac{a'_\rho}{a^*_\rho}$，$\frac{a'_\mu}{a^*_\mu}=\frac{a'_\sigma}{a^*_\sigma}$，导致$\frac{a'_\rho}{a^*_\rho}=\frac{a'_\sigma}{a^*_\sigma}$，与不等式$\frac{a'_\rho}{a^*_\rho}\neq\frac{a'_\sigma}{a^*_\sigma}$矛盾；(2)若$\frac{a'_\rho}{a_\rho}\neq\frac{a'_v}{a_v}$，则$\frac{a'_\rho}{a^*_\rho}=\frac{a'_v}{a^*_v}$，从而$\frac{a'_v}{a^*_v}\neq\frac{a'_\sigma}{a^*_\sigma}$，导致$\frac{a'_\sigma}{a_\sigma}=\frac{a'_v}{a_v}$，于是有$\frac{a'_\rho}{a_\rho}\neq\frac{a'_\sigma}{a_\sigma}$，与等式$\frac{a'_\rho}{a_\rho}=\frac{a'_v}{a_v}$矛盾。

由以上讨论可知，对任给 $\psi\in\mathscr{M}$，式(2.4.16)和(2.4.19)中的所有系数，或者全满足式(2.4.22)，或者全满足式(2.4.23)(也不排斥同时全满足式(2.4.22)和(2.4.23)这种可能)。在前一情况下，

$$\frac{a'_\alpha}{a_\alpha}=\frac{a'_\beta}{a_\beta}=\frac{a'_\gamma}{a_\gamma}=\cdots=A_\psi \tag{2.4.24}$$

式中$|A_\psi|=1$，我们有 $T\psi=A_\psi(a_\alpha u_\alpha+a_\beta u_\beta+\cdots)$，式(2.4.17)成立。

在后一情况下，

$$\frac{a'_\alpha}{a^*_\alpha}=\frac{a'_\beta}{a^*_\beta}=\frac{a'_\gamma}{a^*_\gamma}=\cdots=B_\psi \tag{2.4.25}$$

式中$|B_\psi|=1$，我们有 $T\psi=B_\psi(a^*_\alpha u_\alpha+a^*_\beta u_\beta+\cdots)$，式(2.4.18)成立。

(ⅳ)现在让我们证明，对给定的变换 $T:\mathscr{M}\to\mathscr{M}$，或者全体矢量 $\psi\in\mathscr{M}$ 满足(2.4.17)，或者全体矢量 $\psi\in\mathscr{M}$ 满足(2.4.18)。(两者不相容)

事实上，若对给定的变换 T，某一个态矢量 $\psi=a_\alpha u_\alpha+a_\beta u_\beta+a_\gamma u_\gamma+\cdots$对应的 $T\psi$ 不满足(2.4.18)，则它必然满足(2.4.17)。$T\psi$ 的展开式 $T\psi=a'_\alpha u_\alpha+a'_\beta u_\beta+a'_\gamma u_\gamma+\cdots$ 中必有两个(非零)系数使

$$\left(\frac{a'_\sigma}{a_\rho}\right)^*\neq\frac{a'_\sigma}{a'_\rho}=\frac{a_\sigma}{a_\rho} \tag{2.4.26}$$

由此可得

$$T(u_\sigma+iu_\rho)=Y_{\sigma\rho}(u'_\sigma+iu'_\rho)\neq Z_{\sigma\rho}(u'_\sigma-iu'_\rho) \tag{2.4.27}$$

这是因为$|a_\sigma-ia_\rho|=|a_\sigma-ia_\rho|\neq|a_\sigma+ia_\rho|$。

如还存在

$$\varphi=b_\alpha u_\alpha+b_\beta u_\beta+b_\gamma u_\gamma+\cdots\in\mathscr{M} \tag{2.4.28}$$

使其对应的

$$T\varphi = b'_\alpha u_\alpha + b'_\beta u_\beta + b'_\gamma u_\gamma + \cdots \tag{2.4.29}$$

不满足式(2.4.17),则必满足式(2.4.18)。式(2.4.28)和(2.4.29)中必有两系数使

$$\left(\frac{b_v}{b_\mu}\right)^* = \frac{b'_v}{b'_\mu} \neq \frac{b_v}{b_\mu} \tag{2.4.30}$$

由此可得

$$T(u_\mu + iu_v) = Z_{\mu v}(u'_\mu - iu'_v) \neq Y_{\mu v}(u'_\mu + iu'_v) \tag{2.4.31}$$

这是因为$|b_\mu - ib_v| = |b^*_\mu + ib^*_v| \neq |b^*_\mu - ib^*_v|$。

而考虑矢量

$$\xi = \mu_\sigma + i\mu_\rho + u_\mu + iu_v \tag{2.4.32}$$

的变换行为(为了简单,这里只考虑 σ,ρ,μ,v 各不相同的情况,其它情况,由读者完成)。由式(2.4.27)可知 ξ 按式(2.4.17)变换,而由式(2.4.31)可知 ξ 不能按式(2.4.17)变换,我们得到矛盾。于是当存在不满足式(2.4.17)的 φ 时,不能有不满足式(2.4.18)的 ψ。

(∨)当对变换 T,全体 $\psi \in \mathscr{M}$ 满足式(2.4.17)时,由下式确定的线性算符 U

$$U(a_1u_1 + a_2u_2 + \cdots) = a_1u'_1 + a_2u'_2 + \cdots \tag{2.4.33}$$

是幺正算符:

$$\langle U\psi \mid U\varphi\rangle = \langle \psi \mid \varphi\rangle, \forall\, \psi,\varphi \in \mathscr{M} \tag{2.4.34}$$

我们有

$$T\varphi = A_\varphi U\varphi, (|A_\varphi| = 1) \qquad \forall\, \varphi \in \mathscr{M} \tag{2.4.35}$$

当对变换 T,全体 $\psi \in \mathscr{M}$ 满足式(2.4.18)时,由下式确定的反线性算符 U

$$U(a_1u_1 + a_2u_2 + \cdots) = a^*_1u'_1 + a^*_2u'_2 + \cdots \tag{2.4.36}$$

是反幺正算符:

$$\langle U\psi \mid U\varphi\rangle = \langle \varphi \mid \psi\rangle, \forall\, \psi,\varphi \in \mathscr{M} \tag{2.4.37}$$

我们有

$$T\varphi = B_\varphi U\varphi, \forall\, \varphi \in \mathscr{M} \tag{2.4.38}$$

定理全部证完。

我们看到,保持内积的模不变这一要求把态矢量空间 $\mathscr{M}$ 中的变换(双射)T,限制到只能与一个幺正算符或反幺正算符相差一个相因子。这个相因子一般依赖于被

T 作用的矢量。

可以证明，当且仅当式(2.4.6)中的相因子 C_φ 与 φ 无关时，T 本身是幺正算符或反幺正算符。

当态矢量空间 $\mathscr{M}$ 中的向量经受变换

$$\psi \mapsto T\psi, \forall\, \psi \in \mathscr{M} \tag{2.4.39}$$

时，$\mathscr{M}$ 中的算符 $A:\mathscr{M}\to\mathscr{M}$，经受如下变换

$$A \longrightarrow TAT^{-1} \tag{2.4.40}$$

可以证明(见附录)$\mathscr{M}$ 中的每一线性算符 $A\in EndH$ 在变换(2.4.40)下仍为线性算符的充要条件是式(2.4.6)中的 C_φ 与 φ 无关(亦即 T 是幺正算符或反幺正算符)。于是在量子力学中，保持态矢量空间 $\mathscr{M}$ 中内积模不变，而且保持 $\mathscr{M}$ 中的线性算符仍然是线性算符的变换(2.4.40)，只能是幺正变换或反幺正变换。

我们在此强调。只有零算符既是线性的又是反线性的。恒等算符是线性的，而不是反线性的。前面讨论的空间平移变换和空间转动变换都对应于态矢量空间 $\mathscr{M}$ 中的幺正算符(平移算符，转动算符)这是因为空间平移变换或空间转动变换都与恒等变换属于连续变换群的同一连通分支(由恒等变换出发，可以连续地变到任一平移或任一转动)。而恒等变换对应于恒等算符(线性、幺正算符)，线性幺正算符不可能连续地变成反幺正算符。

如果某一变换与恒等变换不属于连续变换群的同一连通分支(或者变换群是分立的)，我们就无法用上面的论证，断言它一定对应于幺正算符。这时需要作其它的考虑，才可能断定变换对应的算符是幺正的还是反幺正的。

练习 2-4-1 试证：设 A，B 是复内积空间 V 上的线性算符或反线性算符，则 $(AB)^+ = B^+A^+$。

附录

Wingner 证明了，Hilbert 空间 H 中保持内积模不变的变换 T，只能与一个幺正变换或反幺正变换 U 相差一个相因子，即对所有 $\psi\in H$，都有

$$T\varphi = C_{U\varphi}U\varphi \tag{1}$$

式中 $|C_{U\varphi}| = 1$。

对应 $\varphi\mapsto C_\varphi$，$\forall\,\varphi\in H'$ 确定了一个映射(H' 指 H 中全体非零向量的集合)

$$c: H' \longrightarrow \{z \in C \| z| = 1\} \tag{2}$$

定义 H 中的算子 $\hat{C}: H' \longrightarrow H'$ 使

$$\hat{C}\varphi = C_\varphi \varphi \tag{3}$$

则式(1)中的算符 T 可以表作

$$T = \hat{C}U \tag{4}$$

为保证 T 是 H 到自身上的双射,要求 $\hat{C}$ 是双射,亦即,对 H 中所有非零向量 φ,对应

$$\zeta \mapsto C_{\zeta\varphi}\zeta \tag{5}$$

是复平面上单位圆 $\{z \in C \| z| = 1\}$ 到自身的双射。

引理:式(4)中的算符 $T = \hat{C}U$ 是线性算符(或反线性算符)的充要条件为 $C\varphi$ 与 φ 无关($\hat{C}$ 是乘常相因子)

证明:充分性显见。现证明必要性。由 T 是线性(反线性)的可知,对 H 中所有线性无关的向量 φ, ψ,都有

$$T(U^{-1}\varphi + U^{-1}\psi) = TU^{-1}\varphi + TU^{-1}\psi$$

$$\hat{C}(\varphi + \psi) = \hat{C}\varphi + \hat{C}\psi$$

$$C_{\varphi+\psi}\varphi + C_{\varphi+\psi}\psi = C_\varphi\varphi + C_\psi\psi$$

故知

$$C_\varphi = C_{\varphi+\psi} = C_\psi \tag{6}$$

对 $\alpha \neq 0$,

$$C_{\alpha\varphi} = C_\psi = C_\varphi \tag{7}$$

我们证明了对 H 中所有非零向量(不管线性相关否) C_φ 是一个常相因子。引理证完。

这个证明利用了 $dimH > 1$,对体系的态矢量空间这个条件显然是成立的。

定理:对 H 中所有线性算符 $A \in EndH$, TAT^{-1} 仍是线性算符的充要条件为 C_φ 与 φ 无关。

证明:充分性由引理,当 $-C_\varphi$ 与 φ 无关, T 是幺正或反幺正的,易知 TAT^{-1} 是线性的。

必要性,记 $\varphi \xrightarrow{\hat{C}^{-1}} D_\varphi\varphi, \forall \varphi \in H'$

设对所有 $A \in EndH$, TAT^{-1} 是线性的。 $\psi, \varphi \in H$ 是线性独立的。

$$TAT^{-1}(\varphi + \psi) = TAT^{-1}\varphi + TAT^{-1}\psi$$

以 $TAT^{-1}=\hat{C}UAU^{-1}\hat{C}^{-1}$代入上式得

$$\frac{C[D_{\varphi+\psi}UAU^{-1}(\varphi+\psi)]D_{\varphi+\psi}}{C[UAU^{-1}\hat{C}^{-1}\psi]D_{\psi}}TAT^{-1}\varphi+\frac{C[D_{\varphi+\psi}UAU^{-1}(\varphi+\psi)]D_{\varphi+\psi}}{C[UAU^{-1}\hat{C}^{-1}\psi]D_{\psi}}TAT^{-1}\psi$$

$$=TAT^{-1}\varphi+TAT^{-1}\psi \tag{8}$$

由 A 的任意性和 ψ,φ 的线性独立性,对任意线性独立的 $\xi,\eta\in H$

可使 $UAU^{-1}\hat{C}^{-1}\varphi=\xi, UAU^{-1}\hat{C}^{-1}\psi=\eta$

由于 ξ 与 $TAT^{-1}\varphi$, η 与 $TAT^{-1}\psi$ 差相因子,故 $TAT^{-1}\varphi, TAT^{-1}\psi$ 线性独立。由式(8)得

$$\frac{C[D_{\varphi+\psi}UAU^{-1}(\varphi+\psi)]D_{\varphi+\psi}}{C_{\xi}}=1 \tag{9}$$

保持 φ,ψ 不变,选 $B\in EndH$ 不变,使

$$\begin{cases} UBU^{-1}\hat{C}^{-1}\varphi=\eta & (10)\\ UBU^{-1}(\varphi+\psi)=UAU^{-1}(\varphi+\psi) & (11)\end{cases}$$

这是可以实现的,只须选

$$UBU^{-1}\hat{C}^{-1}=\left(1-\frac{D_{\psi}}{D_{\varphi}}\right)\eta+\frac{D_{\psi}}{D_{\varphi}}\xi \tag{12}$$

与上面同样的讨论可得

$$\frac{C[D_{\varphi+\psi}UBU^{-1}(\varphi+\psi)]D_{\varphi+\psi}}{C_{\eta}D_{\varphi}}=1 \tag{13}$$

与式(9)比较得到

$$C_{\xi}=C_{\eta} \tag{14}$$

可见对所有线性独立的 $\xi,\eta\in H$,上式成立。

$\because \alpha\xi(\alpha\neq 0)$与 η 也线性独立,故有

$$C_{\alpha\xi}=C_{\eta}=C_{\xi}\qquad(\alpha\neq 0) \tag{15}$$

于是 H 中所有非零矢量对应于一个共同的相因子:C_{φ} 与 φ 无关,$\hat{C}^{-1}$代表乘积以常相因子的运算。定理证明完毕。

§2.5 空间反射

Ⅰ. 空间反射群

三维实欧氏空间关于固定点 O 的反射是指对应 $\boldsymbol{r}\mapsto-\boldsymbol{r}$, $\forall\,\boldsymbol{r}\in R^3$ 所确定的映射

$$I: R^3\to R^3, I\boldsymbol{r}=-\boldsymbol{r}, \forall\,\boldsymbol{r}\in R^3 \tag{2.5.1}$$

显然对所有的 $\boldsymbol{x},\boldsymbol{y}\in R^3$ 都有

$$I\boldsymbol{x}\cdot I\boldsymbol{y}=\boldsymbol{x}\cdot\boldsymbol{y} \tag{2.5.2}$$

即向量的标积在空间反射下不变,也就是说向量的长度和两向量的夹角在空间反射下不变。在这一点上,它与空间绕 O 点的旋转是一样的:

$$R(\boldsymbol{\theta})\boldsymbol{x}\cdot R(\boldsymbol{\theta})\boldsymbol{y}=\boldsymbol{x}\cdot\boldsymbol{y} \tag{2.5.3}$$

但是,空间绕 O 点的旋转保持矢积关系:

$$R(\boldsymbol{\theta})\boldsymbol{x}\times R(\boldsymbol{\theta})\boldsymbol{y}=R(\boldsymbol{\theta})(\boldsymbol{x}\times\boldsymbol{y}) \tag{2.5.4}$$

而在空间反演下矢积关系要变号,但矢积本身不变:

$$I\boldsymbol{x}\times I\boldsymbol{y}=-I(\boldsymbol{x}\times\boldsymbol{y})=\boldsymbol{x}\times\boldsymbol{y} \tag{2.5.5}$$

这表明右手坐标系在空间反射下变成左手坐标系。此外,显然有

$$I^2=e \tag{2.5.6}$$

(e 是 R^3 上的恒等变换)于是 $\{e,I\}$ 构成一个二阶群,这当然是一个阿贝尔群。

空间反射一个重要的性质是它与空间绕 O 点的转动是可换的:

$$R(\boldsymbol{\theta})I=IR(\boldsymbol{\theta}), \forall\,\boldsymbol{\theta}\in SO(3) \tag{2.5.7}$$

因而我们有直积群

$$SO(3)\otimes\{e,I\}=O(3) \tag{2.5.8}$$

这是三维实欧氏空间 R^3 上,保持固定点 O 不变,又保持任意两个向量标积不变(保持任意两点距离不变)的变换的全体所构成的群,三维实正交变换群,也就是全体 3×3 实正交矩阵构成的群。

空间反射与空间非零平移 $T(\boldsymbol{\rho})\in T_3(R)(\boldsymbol{\rho}\neq\boldsymbol{0})$ 是不对易的:

$$(T(\boldsymbol{\rho}) * I)\boldsymbol{r} = -\boldsymbol{r} + \boldsymbol{\rho}, (I * T(\boldsymbol{\rho}))\boldsymbol{r} = -\boldsymbol{r} - \boldsymbol{\rho} \tag{2.5.9}$$

代替对易关系，由式(2.5.9)我们可得

$$I * T(-\boldsymbol{\rho}) = T(\boldsymbol{\rho}) * I \tag{2.5.10}$$

在三维实欧氏空间中，关于固定点 O 的空间反射，绕 O 点的空间旋转和空间平移的这些关系，在很大程度上决定了态矢量空间中相应算符的对易关系。

练习 2-5-1 试证：空间关于固定点 O 的反射与空间绕 O 点的转动对易(式(2.5.7))。

Ⅱ. 态矢量和力学量算符在空间反射下的变换性质

先考察单个经典质点。若质点的某一运动由方程 $\boldsymbol{r} = \boldsymbol{f}(t)$ 刻画，在空间反射下，这一运动将变为由方程 $\boldsymbol{r} = -\boldsymbol{f}(t)$ 刻画的运动。质点在任一时刻 t_0 的状态在空间反射下，将从$(\boldsymbol{r} = \boldsymbol{f}(t_0), \boldsymbol{p} = m\boldsymbol{f}'(t_0))$变作$(\boldsymbol{r} = -\boldsymbol{f}(t_0), \boldsymbol{p} = -m\boldsymbol{f}'(t_0))$。质点在任一时刻 t_0 的轨道角动量 $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p} = \boldsymbol{f}(t_0) \times m\boldsymbol{f}'(t_0)$在空间反射下不变。

过渡到量子力学，考察在空间反射下，单个粒子的态矢量$|\varphi\rangle$如何变换。以$|\varphi'\rangle$记$|\varphi\rangle$在空间反射下变成的态矢量，按照 Wigner 定理，显然应该有

$$|\varphi'\rangle = \Pi |\varphi\rangle \tag{2.5.11}$$

式中的 Π 是态矢量空间 S 上的幺正算符或者反幺正算符$(\Pi^+ = \Pi^{-1})$，称作该粒子的宇称算符。让我们来确定宇称算符 Π 的具体形式。由于空间反射两次是恒等变换$(I^2 = e)$，故应有 $\Pi^2 = 1$(Π 是对合的)，从而

$$\Pi = \Pi^{-1} = \Pi^+ \tag{2.5.12}$$

Π 是厄米的。以$|\boldsymbol{r}m_s\rangle$记 $\hat{\boldsymbol{r}}, \hat{s}_z$ 的共同的本征矢；以$|\boldsymbol{p}m_s\rangle$记 $\hat{\boldsymbol{p}}, \hat{s}_z$ 的共同的本征矢：

$$\begin{aligned}&\hat{\boldsymbol{r}} |\boldsymbol{r}m_s\rangle = \boldsymbol{r} |\boldsymbol{r}m_s\rangle, \hat{s}_z |\boldsymbol{r}m_s\rangle = m_s |\boldsymbol{r}m_s\rangle;\\ &\hat{\boldsymbol{p}} |\boldsymbol{p}m_s\rangle = \boldsymbol{p} |\boldsymbol{p}m_s\rangle, \hat{s}_z |\boldsymbol{p}m_s\rangle = m_s |\boldsymbol{p}m_s\rangle\end{aligned} \tag{2.5.13}$$

由对应原理，显然应有(注意，此处的 $\boldsymbol{r}, \boldsymbol{p}$ 是本征值)

$$\hat{\boldsymbol{r}}\Pi |\boldsymbol{r}m_s\rangle = -\boldsymbol{r}\Pi |\boldsymbol{r}m_s\rangle = -\Pi\boldsymbol{r} |\boldsymbol{r}m_s\rangle = -\Pi\hat{\boldsymbol{r}} |\boldsymbol{r}m_s\rangle$$

$$\hat{\boldsymbol{p}}\Pi |\boldsymbol{p}m_s\rangle = -\boldsymbol{p}\Pi |\boldsymbol{p}m_s\rangle = -\Pi\boldsymbol{p} |\boldsymbol{r}m_s\rangle = -\Pi\hat{\boldsymbol{p}} |\boldsymbol{r}m_s\rangle$$

因为 $\hat{\boldsymbol{r}}, \hat{s}_z(\hat{\boldsymbol{p}}, \hat{s}_z)$有完备的共同的本征向量组，由此可得算符等式

$$\hat{\boldsymbol{r}}\Pi = -\Pi\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}\Pi = -\Pi\hat{\boldsymbol{p}} \tag{2.5.14}$$

$\hat{\boldsymbol{r}},\hat{\boldsymbol{p}}$ 与宇称算符 Π 反对易。由上式易知轨道角动量算符 $\hat{\boldsymbol{l}}=\hat{\boldsymbol{r}}\times\hat{\boldsymbol{p}}$ 与宇称算符 Π 对易：

$$\hat{\boldsymbol{l}}\Pi = \Pi\hat{\boldsymbol{l}} \tag{2.5.15}$$

式(2.5.14)可以重新写作

$$\Pi\hat{\boldsymbol{r}}\Pi^{-1} = -\hat{\boldsymbol{r}},\Pi\hat{\boldsymbol{p}}\Pi^{-1} = -\hat{\boldsymbol{p}} \tag{2.5.16}$$

计算对易关系 $\Pi[\hat{x},\hat{p}_x]\Pi^{-1}=[\Pi\hat{x}\Pi^{-1},\Pi\hat{p}_x\Pi^{-1}]=[-\hat{x},-\hat{p}_x]$，可以得到 $\Pi i\Pi^{-1}=i$，这表明 Π 不是反线性算符，故只能是幺正算符。

于是，当空间经受关于固定点 O 的反射变换 $I\in\{e,I\}$ 时，单粒子的态矢量经受幺正变换 Π。回忆 §2.3 的讨论，当空间经受绕固定点 O 的转动 $\boldsymbol{\theta}\in SO(3)$ 时，单粒子的态矢量经受幺正变换 $D(\boldsymbol{\theta})$。因为空间关于固定点 O 的反射与空间绕 O 点的转动是可换的，故应有 $D(\boldsymbol{\theta})\Pi=\Pi D(\boldsymbol{\theta})$，$\forall\,\boldsymbol{\theta}\in SO(3)$，此即

$$\hat{\boldsymbol{j}}\Pi = \Pi\hat{\boldsymbol{j}} \tag{2.5.17}$$

单粒子的总角动量与宇称算符对易。鉴于式(2.5.15)，单粒子的自旋角动量与宇称算符对易：

$$\Pi\hat{\boldsymbol{s}} = \hat{\boldsymbol{s}}\Pi \tag{2.5.18}$$

上式亦可以写作

$$\Pi\hat{\boldsymbol{s}}\Pi^{-1} = \hat{\boldsymbol{s}} \tag{2.5.19}$$

命题 式(2.5.16)和(2.5.19)把可逆线性算符 Π 确定到一个非零复数倍。

事实上，若有线性算符 Λ 也满足式(2.5.16)及(2.5.19)：$\Lambda\hat{\boldsymbol{r}}\Lambda^{-1}=-\hat{\boldsymbol{r}}$，$\Lambda\hat{\boldsymbol{p}}\Lambda^{-1}=-\hat{\boldsymbol{p}}$，$\Lambda\hat{\boldsymbol{s}}\Lambda^{-1}=\hat{\boldsymbol{s}}$，与式(2.5.16)及(2.5.19)合并可得：$(\Lambda^{-1}\Pi)\hat{\boldsymbol{r}}=\hat{\boldsymbol{r}}(\Lambda^{-1}\Pi)$，$(\Lambda^{-1}\Pi)\hat{\boldsymbol{p}}=\hat{\boldsymbol{p}}(\Lambda^{-1}\Pi)$，$(\Lambda^{-1}\Pi)\hat{\boldsymbol{s}}=\hat{\boldsymbol{s}}(\Lambda^{-1}\Pi)$。考虑到态矢量空间 S 上的任一线性算符都可以由 $\hat{\boldsymbol{r}},\hat{\boldsymbol{p}},\hat{\boldsymbol{s}}$ 表出，可知它们是矢量空间 S 上的一个不可约线性算符组(参看练习 0-3-9)。由 Schur 引理可知，$\Lambda^{-1}\Pi=\omega 1$，ω 是复数。于是 $\Pi=\omega\Lambda$。由于 Π 是非奇异的，可知 $\omega\neq 0$。

在 $(\boldsymbol{r},s_z)$ - 表象中，由下式定义线性算符 $\Pi_o: S\to S$

$$\Pi_o\varphi(\boldsymbol{r},s_z) = \varphi(-\boldsymbol{r},s_z),\ \forall\,\varphi\in S \tag{2.5.20}$$

直接验算可知，$\Pi_o^2=1(\Pi_o=\Pi_o^{-1})$，$\Pi_o$ 是幺正算符($\Pi_o^{-1}=\Pi_o^{+}$)，从而 Π_o 是厄米算符($\Pi_o=\Pi_o^{+}$)。由定义式(2.5.20)还可验证 Π_o 满足条件(2.5.16)及(2.5.19)。我们称厄米算符 Π_o 为粒子的轨道宇称算符。而由上面的命题，粒子的宇称算符一般应为

$$\Pi = \omega\Pi_o \tag{2.5.21}$$

式中 ω 是非零复数。由 $\Pi^2 = 1$ 和 $\Pi_o^2 = 1$ 得，$\omega^2 = 1$，从而只有两种可能：

$$\omega = \pm 1 \tag{2.5.22}$$

式中的 ω 依赖于粒子的种类。而在自然界中，这两种可能都存在。称 $\omega = 1$ 的粒子具有偶内禀宇称，$\omega = -1$ 的粒子具有奇内禀宇称。令 $\Pi_i = \omega$，称之为粒子的内禀宇称算符。于是粒子的宇称算符为其轨道宇称算符和内禀宇称算符之积。

$$\Pi = \Pi_i\Pi_o = \Pi_o\Pi_i \tag{2.5.23}$$

我们看到，在空间反射下，单粒子体系的态矢量经受粒子宇称算符 Π 的作用。也就是说，单粒子态矢量空间 S 荷载着空间反射群 $\{e, I\}$ 的一个幺正表示 $\{1, \Pi\}$。又由于 $\Pi^2 = 1$（Π 是对合的），可知 Π 是厄米的。这个表示空间反射 I 的厄米算符 Π，就是粒子的宇称算符。

式(2.5.16) 及(2.5.19)给出了单粒子体系的力学量算符 $\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}, \hat{\boldsymbol{s}}$ 在空间反射下的变换方式。单粒子体系的任一力学量 $A = f(\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{s})$ 的算符在空间反射下按下式变换：

$$\hat{A} = f(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}, \hat{\boldsymbol{s}}) \mapsto \Pi f(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}, \hat{\boldsymbol{s}})\Pi^{-1} = f(-\hat{\boldsymbol{r}}, -\hat{\boldsymbol{p}}, \hat{\boldsymbol{s}}) \tag{2.5.24}$$

特别是，体系的哈密顿算符 $\hat{H}$ 在空间反射下变作 $\Pi\hat{H}\Pi^{-1} = \hat{H}'$。当且仅当 $\hat{H}' = \hat{H}$ 时，有

$$[\Pi, \hat{H}] = 0 \tag{2.5.25}$$

这时体系的宇称守恒。

由于 $\Pi_0, \hat{\boldsymbol{l}}$ 对易，可知 $\Pi_0, \hat{l}^2, \hat{l}_z$ 有共同的完备的本征函数。在空间反射下

$$(r, \theta, \varphi) \mapsto (r, \pi - \theta, \varphi + \pi) \tag{2.5.26}$$

回忆 $Y_{lm}(\theta, \varphi)$ 的具体形式：

$$Y_{lm}(\theta, \varphi) = i^l(-1)^m\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}P_l^m(\cos\theta)e^{im\varphi} \tag{2.5.27}$$

式中的连带 Legendre 多项式 $P_l^m(x) = \dfrac{1}{2^l \cdot l!}(1-x^2)^{m/2}\dfrac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(x^2-1)^l$。其中因子 $\dfrac{1}{2^l \cdot l!}(1-x^2)^{m/2}$ 是 x 的偶函数；而因子 $\dfrac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(x^2-1)^l$ 是 x 的 $l-m$ 次多项式，当 $l-m$ 是偶(奇)数时，只包含 x 的偶(奇)次方项。注意到数学恒等式 $\cos(\pi - \theta) = -$

$\cos\theta$, $e^{im(\varphi+\pi)} = (-1)^m e^{im\varphi}$。可得

$$\Pi_o Y_{lm}(\theta,\varphi) = Y_{lm}(\pi-\theta,\varphi+\pi) = (-1)^l Y_{lm}(\theta,\varphi) \tag{2.5.28}$$

上式表明球谐函数的轨道宇称为$(-1)^l$,与磁量子数 m 无关。这不是偶然的。由于 Π 与 $\hat{\boldsymbol{l}}$ 对易,从而 Π_o 与 $\hat{\boldsymbol{l}}$ 对易,与 $\hat{l}_\pm$ 对易。以 $\hat{l}_\pm$ 作用于等式 $\Pi_o Y_{lm}(\theta,\varphi) = c_{lm} Y_{lm}(\theta,\varphi)$ 两边就可证明球谐函数的轨道宇称与磁量子数 m 无关。

上面讨论了单个粒子体系的宇称算符。多个粒子体系的宇称算符是各个粒子宇称算符之积。这与多个粒子体系的动量算符,角动量算符是各个粒子动量算符,角动量算符之和不同。从群表示论看,这是很自然的。事实上,多个粒子体系的态矢量空间是各个粒子的态矢量空间的张量积:$\mathscr{M} = S_1 \otimes \cdots \otimes S_N$。在空间平移、空间转动或者空间反射下,$\mathscr{M}$ 所荷载的空间平移群、空间转动群或者空间反射群的表示,是各个单粒子态空间所荷载的表示的直积。多个粒子体系的态矢量空间上的平移算符(转动算符)是各个粒子态矢量空间上的平移算符(转动算符)的乘积;同样,多个粒子体系的宇称算符 Π 是各个粒子的宇称算符$\{\Pi^{(\alpha)} \mid \alpha = 1,\cdots,N\}$的乘积。

$$\Pi = \Pi^{(1)} \otimes \Pi^{(2)} \otimes \cdots \otimes \Pi^{(N)} = \Pi^{(1)}\Pi^{(2)}\cdots\Pi^{(N)} \tag{2.5.29}$$

但是动量算符,角动量算符是群表示的无穷小算符的 i 倍;而宇称算符本身是表示空间反射的算符。这说明了为什么多个粒子体系的动量算符,角动量算符是各个粒子动量算符,角动量算符之和,而多个粒子体系的宇称算符是各个粒子宇称算符之积。

在非相对论量子理论中,不考虑粒子数不守恒的过程,体系所包含的各种粒子的个数是不变的。内禀宇称算符 $\Pi_i = \Pi_i^{(1)}\Pi_i^{(2)}\cdots\Pi_i^{(N)}$ 对体系态矢量 $\Psi(t)$ 的作用,相当于乘上常相因子 $\Omega = \omega^{(1)}\omega^{(2)}\cdots\omega^{(N)} = \pm 1$,在物理上不起作用。与此相应的是,内禀宇称算符 Π_i,与所有的(不破坏粒子数守恒的)力学量算符对易。但在相对论性量子理论中,必须考虑粒子数不守恒的过程,体系不再是由固定的一些粒子所组成。体系在不同时刻的组成是不同的,甚至在同一时刻,体系的组成也可能是不确定的。这时必须考虑内禀宇称算符。

练习 2-5-1 试问:宇称算符 Π 与动量算符反对易与空间变换关系 $I*(-\boldsymbol{\rho}) = \boldsymbol{\rho} * I$ 之间有什么联系?

Ⅲ. 宇称守恒与破缺

自然界有一些基本的对称性。例如空间的均匀性,它意味着孤立体系的哈密顿

量在空间平移下是不变的,体系的总动量守恒;空间的各向同性,它意味着孤立体系的哈密顿量在空间转动下是不变的,体系的总角动量守恒;时间的均匀性,它意味着孤立体系的哈密顿量不显含时间,孤立体系的能量守恒。这些对称性为人类迄今为止的经验所证实。此外,一般还认为自然规律应对空间的镜面反射是不变的,即空间具有镜面反射对称性。空间关于固定点 O 的反射,相当于空间绕过 O 点的某固定轴 $\hat{n}$ 旋转 π 角,接着相对于过 O 点并垂直于轴 $\hat{n}$ 的平面做镜面反射。因此,自然规律对空间的镜面反射是不变的,就是自然规律对空间反射是不变的,孤立体系的哈密顿量在空间反射下是不变的。也就是说,一般认为,宇称守恒是自然界的基本规律之一。事实上,内禀宇称的引入和各种粒子内禀宇称的确定,都是根据宇称守恒这一准则而做的。作为例子,让我们看一下 π^- 介子内禀宇称是如何确定的。

已经知道质子 p、中子 n 的自旋为 1/2,π^- 介子的自旋为零;质子、中子具有偶的内禀宇称。由一个质子和一个中子构成的氘核,捕获一个 π^- 介子形成一个 π^- - 介原子(原子中的电子为带负电 π^- 的介子替代)。通过发射光子,很快跃迁到基态 s 态,然后衰变为两个中子。让我们来分析最后的这一衰变过程:$d\pi^- \to nn$。利用动量守恒、角动量守恒,宇称守恒和泡利原理,我们可以分析出 π^- 介子的内禀宇称为奇。事实上,我们采用质心坐标系,使衰变前后质心运动的动量为零:$\boldsymbol{P} = \boldsymbol{P}' = 0$。分别以 Ψ_i, Ψ_f 表示体系的初态和末态的态矢量。由 $\hat{\boldsymbol{p}}\Psi_i = 0\Psi_i, \hat{\boldsymbol{p}}\Psi_f = 0\Psi_f$,考虑到质心运动的角动量算符 $\hat{\boldsymbol{L}} = \hat{\boldsymbol{R}} \times \hat{\boldsymbol{P}}$,就不难得到 $\hat{\boldsymbol{L}}\Psi_i = 0\Psi_i, \hat{\boldsymbol{L}}\Psi_f = 0\Psi_f$。从而衰变前后质心运动的角动量为零:$\boldsymbol{L} = \boldsymbol{L}' = 0$(不对易的算符仍可有若干共同的本征矢)。因而初态的总轨道角动量(见第 2 章 § 1)是 π^- 介子相对于 d 的轨道角动量加上中子相对质子的轨道角动量,初态的总角动量就是质子、中子相对运动轨道角动量加 n, p 的自旋,也就是 d 的总自旋。实验告诉我们氘核的唯一结合态的总自旋(总角动量)为 1,而质子、中子相对运动波函数具有偶宇称。注意到当空间反射时,质心坐标和所有的相对坐标都变号,$\boldsymbol{P} = \boldsymbol{P}' = 0$ 的质心运动波函数不变。π^- 介子相对氘核 d 的相对运动是 s 态,具有偶宇称(见式(2.5.28))。初态的总宇称就是 π^- 介子的内禀宇称。终态必须满足泡利原理。若自旋反平行,那么自旋波函数是反对称的,空间波函数必须对称。交换两个中子的空间坐标等于把相对坐标 $\boldsymbol{r}$ 变号(质心坐标 $\boldsymbol{R}$ 不变)。故相对运动波函数必有偶轨道宇称,轨道角动量子数为偶数,考虑到质心运动轨道角动量为零,总角动量不可能是 1,从而破坏了角动量守恒,只可能终态的中子自旋平行。空

间波函数是反对称的。相对运动波函数具有奇宇称。终态的总宇称是奇的。可见 π^- 介子具有奇内禀宇称。

其它粒子的内禀宇称也都是利用实验，通过类似的分析确定的。

一直到 1956 年以前，人们对自然的空间反射对称性和宇称守恒律一直是深信不疑的。但是为了解释实验提出的 $\tau-\theta$ 疑难，这种看法才发生动摇。所谓 $\tau-\theta$ 疑难是指当时人们发现的按两种不同方式

$$\theta^+ = \pi^+ + \pi^0$$

$$\tau^+ = \pi^+ + \pi^+ + \pi^-$$

衰变的粒子 θ^+ 和 τ^+ 有着相同的寿命、自旋和质量。按宇称守恒律，θ^+ 应有偶内禀宇称，τ^+ 应有奇内禀宇称，不应被看作同一种粒子。

李政道和杨振宁分析检查已有的实验，发现虽然在电磁作用，强相互作用领域，实验表明宇称是守恒的，但并没有实验表明在弱相互作用中宇称守恒。他们提出 θ^+，τ^+ 是同一种粒子，只不过弱相互作用中没有宇称守恒，并且提出了判断这个理论的若干实验。吴健雄等人于 1957 年用钴 60 源在超低温下强磁场中极化，观察 β 衰变 $n\to p+e+\bar{\nu}_e$ 的角分布，发现与 ^{60}Co 自旋方向成 θ 角方向的 β 粒子流强度和与 ^{60}Co 自旋方向成 $\pi-\theta$ 角方向 β 的粒子流强度有很大的差别。从实验上证实了弱相互作用中宇称不守恒。

1957 年的一些发现，不仅建立了左右不对称性（P 破坏），也建立了正反粒子不对称性（C 破坏），同时还对时间反演和 CP 变换下的对称性产生了疑问。这至今仍是物理学家感兴趣的课题。

§2.6 时间平移与时间反演

在讨论牵涉到时间的变换之前，我们愿意提醒读者注意，体系在某一时刻的状态与体系在某一段时间内的运动这两个概念之间的区别。与此相应，在量子力学中单个无旋粒子的态矢量空间 S 是定义在三维空间 R^3 上的（满足波函数自然条件的）复值函数 $\psi(\boldsymbol{r})$ 的集合。而 $\{\varphi(\boldsymbol{r},t)\mid t_1<t<t_2\}$ 代表 S 中的一个单参数向量簇。若 $\varphi(\boldsymbol{r};t)$ 满足 Schrödinger 方程，则 $\{\varphi(\boldsymbol{r},t)\mid t_1<t<t_2\}$ 代表体系在 t_1 到 t_2 时段内一个可

能运动。当 t 取(t_1,t_2)中的某一固定值时，$\varphi(\boldsymbol{r},t)$代表体系在 t 时刻的状态。$\frac{\partial}{\partial x}$是 S 上的算符，但$\frac{\partial}{\partial t}$不能看作是$S$ 上的算符。

Ⅰ.时间均匀性与能量守恒

时间均匀性是自然的基本属性之一，它导致了能量守恒。

在经典力学中，一个孤立体系或一个处于恒定外场中的体系，其哈密顿量不依赖时间 t：

$$H = H(p,q), \frac{\partial}{\partial t}H = 0 \tag{2.6.1}$$

从而未知函数 $p_\alpha(t), q_\alpha(t)$的自变量 t 不出现在正则运动方程中

$$\begin{cases} \dot{p}_\alpha = -\dfrac{\partial H(p,q)}{\partial q_\alpha} \\ \dot{q}_\alpha = \dfrac{\partial H(p,q)}{\partial p_\alpha} \end{cases} \tag{2.6.2}$$

若$\{p_\alpha = f_\alpha(t), q_\alpha = g_\alpha(t)\}$是(2.6.2)的一个解，则$\{p_\alpha = f_\alpha(t-\tau), q_\alpha = g_\alpha(t-\tau)\}$($\tau$是常实数)也是方程(2.6.2)的一个解。后者代表的运动，是前者代表的运动推延时间 τ 后的重演。这时体系的哈密顿量是运动积分：

$$\frac{\mathrm{d}}{\mathrm{d}t}H(p,q) = 0 \tag{2.6.3}$$

$$H(p,q) = E \tag{2.6.4}$$

常数 E 是体系的能量。

处于变化外场中的体系，哈密顿量依赖时间，体系能量不守恒，这并不与时间均匀性相违背。如果我们把外场和场源也包含在体系之内，这种具体环境对时间均匀性的破坏就可以消除。能量守恒是自然的基本定律之一。

在量子力学中，一个孤立体系或一个处于恒定外场中的体系，哈密顿算符不依赖时间 t。由于自变量 t 不出现在薛定谔方程中

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \tag{2.6.5}$$

若$|\psi(t)\rangle$是(2.6.5)的解，则$|\psi(t-\tau)\rangle$也是(2.6.5)的一个解。后者代表的运动是

前者代表的运动推迟时间 τ 的重演。后者在 $t = t_0$ 时刻的状态正好是前者在 $t = t_0 - \tau$ 时刻的状态。

当体系的哈密顿量不依赖时间 t 时，第 1 章中引入的演化算符有简单的形式：

$$U(t, t_0) = \exp\left(\frac{t - t_0}{i\hbar}\hat{H}\right) \tag{2.6.6}$$

它完全由时间间隔 $t - t_0$ 所决定。故有

$$U(t - \tau, t_0 - \tau) = U(t, t_0) \tag{2.6.7}$$

设 $\psi(t)$和 $\varphi(t)$代表体系的两个可能的运动使

$$\psi(t_0 - \tau) = \varphi(t_0) \tag{2.6.8}$$

则因 $\psi(t - \tau) = U(t - \tau, t_0 - \tau)\psi(t_0 - \tau)$

$$\varphi(t) = U(t, t_0)\varphi(t_0)$$

我们有

$$\psi(t - \tau) = \varphi(t) \tag{2.6.9}$$

我们重新得到了前面讨论的结论。

任一不依赖时间，并与 $\hat{H}$ 对易的厄米算符代表一个守恒量。故不依赖时间的哈密顿量是守恒量——体系的能量。

如同在经典力学中一样，处于变化外场中体系的哈密顿量依赖时间 t，体系能量不守恒。但把外场、场源包含在体系之内，我们总有不依赖时间的哈密顿量。能量守恒是基本自然规律之一，时间的均匀性是自然的基本属性之一。

Ⅱ. 经典力学中时间反演

设想我们对某个质点在时间区间$(-\tau, 0)$内的运动进行三维立体录象，紧接着在时间区间$(0, \tau)$按录象速度倒放。若原质点的运动由

$$\boldsymbol{r} = \boldsymbol{R}(t), \quad -\tau < t < 0 \tag{2.6.10}$$

描写，则象质点的运动由

$$\boldsymbol{r} = \boldsymbol{R}(-t), 0 < t < \tau \tag{2.6.11}$$

描写。对任一 $t' \in (0, \tau)$，试比较象质点在 t' 时刻的状态和原质点在 $-t'$ 时刻的状态，易见前者是以位置矢量 $\boldsymbol{r} = \boldsymbol{R}(-t')$和动量 $\boldsymbol{p} = -m\dot{\boldsymbol{R}}(-t')$来标志的；而后者是以 $\boldsymbol{r} = \boldsymbol{R}(-t')$和动量 $\boldsymbol{p} = m\dot{\boldsymbol{R}}(-t')$来标志的。自然在这两个状态中质点的角动量

相反。我们称这样的两个状态互为时间反演态。我们称式(2.6.10)和式(2.6.11)所描写的运动互为时间反演运动。

现在我们要问,是否每个可能运动的时间反演运动都是可能运动?读者都知道,在重力场中,自由落体运动的时间反演便是上抛运动的上升阶段,斜抛运动的时间反演是斜抛运动。但若考虑阻力时,自由落体运动、斜抛运动的时间反演运动便不再是可能的运动了。一般而言,当体系的哈密顿量具有如下的对称性时

$$H(p,q) = H(-p,q) \tag{2.6.12}$$

若$\{p_\alpha = P_\alpha(t), q_\alpha = Q_\alpha(t) \mid \alpha = 1,2,\cdots,N\}$是正则方程的一个解,则

$\{p_\alpha = -P_\alpha(-t), q_\alpha = Q_\alpha(-t) \mid \alpha = 1,2,\cdots,N\}$也是正则方程的一个解。(读者证明之!)也就是说,要条件(2.6.12)成立时,可能运动的时间反演是可能运动。

Ⅲ.量子力学中时间反演

以 T 代表态矢量空间 $\mathscr{M}$ 中的时间反演算符,它把体系的任一态$|\varphi\rangle$变成其时间反演态$|\varphi\rangle$。显然 T 应该是幺正的或者是反幺正的,而且对所有 $\varphi\in\mathscr{M}$,都有

$$\langle T\varphi \mid \boldsymbol{r} \mid T\varphi\rangle = \langle\varphi \mid \boldsymbol{r} \mid \varphi\rangle \tag{2.6.13}$$

和

$$\langle T\varphi \mid \boldsymbol{p} \mid T\varphi\rangle = -\langle\varphi \mid \boldsymbol{p} \mid \varphi\rangle \tag{2.6.14}$$

于是对所有 $\varphi\in\mathscr{M}$,$\langle\varphi|T^+\boldsymbol{r}T|\varphi\rangle = \langle\varphi|\boldsymbol{r}|\varphi\rangle$,$\langle\varphi|T^+\boldsymbol{p}T|\varphi\rangle = -\langle\varphi|\boldsymbol{p}|\varphi\rangle$。从而 $T^+\boldsymbol{r}T = \boldsymbol{r}$,$T^+\boldsymbol{p}T = -\boldsymbol{p}$。此即

$$T\boldsymbol{r}T^+ = \boldsymbol{r} \tag{2.6.15}$$

$$T\boldsymbol{p}T^+ = -\boldsymbol{p} \tag{2.6.16}$$

由变换(2.6.15)和(2.6.16)可得轨道角动量在时间反演下的变换方式

$$T\boldsymbol{l}T^+ = -\boldsymbol{l} \tag{2.6.17}$$

自旋角动量与轨道角动量是同性质的量,故我们假定

$$T\boldsymbol{s}T^+ = -\boldsymbol{s} \tag{2.6.18}$$

由(2.6.15)和(2.6.16)还可得 $T(xp_x - p_x x)T^+ = -(xp_x - p_x x)$,$TiT^+ = -i$,可知 T 是反幺正算符。可以把式(2.6.15)、(2.6.16)、(2.6.18)和反幺正性看作确定时间反演算符的条件(式(2.6.17)可由式(2.6.15)、(2.6.16)推出,不是独立的)。

首先让我们考虑无旋粒子态矢量空间 S 中的时间反演算符。$\{\boldsymbol{r},\boldsymbol{p}\}$对此态矢量

空间 S 构成不可约算符组，故由 Schur 引理，关系(2.6.15)和(2.6.16)把 T 确定到只差一个常数倍，而反幺正性进一步把 T 确定到只差一个常相因子。

不难验证在通常坐标表象中的取复共轭算符 K_0：

$$K_0\varphi(\boldsymbol{r}) = \varphi^*(\boldsymbol{r}) \tag{2.6.19}$$

满足(2.6.15)和(2.6.16)，且是反幺正的，故可以把无旋粒子态矢量空间 S 中的时间反演算符 T 取作 K_0(约定 $T=\gamma K_0$ 中的常相因子 $\gamma=1$)：

$$T = K_0 \tag{2.6.20}$$

现在考虑自旋为 s 的粒子。仍以 T 表示态矢量空间 S' 中的时间反演算子。我们现在还必须考虑自旋算符在时间反演下的变换性质(2.6.18)。$\{\boldsymbol{r},\boldsymbol{p},\boldsymbol{s}\}$ 是 S' 中的不可约算符组，故由 Schur 引理，式(2.6.15)、(2.6.16)和(2.6.18)把 T 确定到一个常因子，而反幺正性进一步把 T 确定到只差一个常相因子。仍以 K_0 表示这个态矢量空间 S' 上的取复共轭算符(在通常的 $\{\boldsymbol{r},s_z\}$ 表象)

$$K_0\varphi(\boldsymbol{r},S_z) = \varphi^*(\boldsymbol{r},S_z) \tag{2.6.21}$$

由上式可知，K_0 是 S' 上的一个反幺正算符，满足(2.6.15)和(2.6.16)，而且是对合的：

$$K_0^2 = 1 \tag{2.6.22}$$

引入算符

$$K = TK_0 \tag{2.6.23}$$

K 显然是个幺正算符，利用(2.6.22)、(2.6.23)可得

$$T = KK_0 \tag{2.6.24}$$

由式(2.6.15)(2.6.16)和(2.6.18)，并考虑到在通常的 $\{\boldsymbol{r},s_z\}$ 表象中，s_y 是纯虚的，s_x,s_z 是纯实的(参考下一章)，可得

$$K\boldsymbol{r} = \boldsymbol{r}K, K\boldsymbol{p} = \boldsymbol{p}K \tag{2.6.25}$$

$$Ks_y = s_yK \tag{2.6.26}$$

$$Ks_{x,z} = -s_{x,z}K \tag{2.6.27}$$

不难验证，把自旋空间绕 y 轴转 π 角的算符

$$D_s(\pi\hat{j}) = \exp(-i\pi s_y) \tag{2.6.28}$$

满足(2.6.25)~(2.6.27)，而且是幺正的。在这儿，我们只证明它满足(2.6.27)。后

者等价于 $\exp(-i\pi s_y)s_{x,z}\exp(i\pi s_y)=-s_{x,z}$(把实验室绕 y 轴旋转 π 角,测量粒子自旋在 x,z 方向上的投影的装置,变成测量粒子自旋在 $-x,-z$ 方向上的投影的装置)。由式(0.4.12)可知,此即 $\exp(-i\pi\ ads_y)s_{x,z}=-s_{x,z}$。我们来证明把实验室绕 y 轴旋转任意角 α 时,$\exp(-i\alpha\ ads_y)s_x=\cos\alpha s_x-\sin\alpha s_z$。事实上,注意到 $(ads_y)s_x=[s_y,s_x]=-is_z$,$(ads_y)^2s_x=s_x$,和指数函数、正弦、余弦函数的 Taylor 展开式,可知等式左边等于 $\left(\sum_{n\geq0}\frac{(-i\alpha)^n}{n!}(ads_y)^n\right)s_x=\sum_{n\geq0}\frac{(-i\alpha)^{2n}}{(2n)!}s_x+\sum_{n\geq0}\frac{(-i\alpha)^{2n+1}}{(2n+1)!}(-is_z)=\cos\alpha s_x-\sin\alpha s_z$。当 $\alpha=\pi$ 时,我们得到 $\exp(-i\pi s_y)s_x\exp(i\pi s_y)=-s_x$。同理可得 $\exp(-i\pi s_y)s_z\exp(i\pi s_y)=-s_z$。

因此我们可以取(约定 $K=\eta\exp(-i\pi s_y)$中的常相因子 $\eta=1$)

$$K=D_s(\pi\hat{j})=\exp(-i\pi s_y) \tag{2.6.29}$$

于是对自旋为 s 的单粒体系,时间反演算符为

$$T=\exp(-i\pi s_y)K_0=K_0\exp(-i\pi s_y) \tag{2.6.30}$$

不难把这个结果推广到多粒子体系:

$$T=\exp(-i\pi s_y)K_0 \tag{2.6.31}$$

不过式中 K_0 是由下式定义的:

$$K_0\varphi(\boldsymbol{r}_1s_{z1},\cdots,\boldsymbol{r}_Ns_{zN})=\varphi^*(\boldsymbol{r}_1s_{z1},\cdots,\boldsymbol{r}_Ns_{zN}) \tag{2.6.32}$$

而 S_y 是所有粒子自旋之和:

$$S_y=\sum_{\alpha=1}^{N}s_{y\alpha} \tag{2.6.33}$$

由(2.6.31)可得(注意 K_0 与 $\exp(-i\pi S_y)$对易)

$$T^2=\exp(-i2\pi S_y)=\prod_{\alpha=1}^{N}\exp(-i2\pi s_{y\alpha}) \tag{2.6.34}$$

这是在体系自旋空间中绕 y 轴转 2π 角的算符,但它并不总是 1。由角动量理论知,当自旋角量子数 s 为半整数时,$\exp(-i2\pi s_y)=-1$(参看下一章角动量理论)。故当体系中费米子个数为奇数时,$T^2=-1$。以 N_F 记体系中费米子个数,则

$$T^2=(-1)^{N_F} \tag{2.6.35}$$

Ⅳ. 角动量本征态的时间反演性质

单个粒子轨道角动量的本征函数为 $Y_{lm}(\theta\varphi)=i^l(-1)^m\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}P_l^m$

$(\cos\theta)e^{im\varphi}$,在时间反演下它变作

$$TY_{lm}(\theta,\varphi) = K_oY_{lm}(\theta,\varphi) = Y_{lm}^*(\theta,\varphi) = (-1)^{l-m}Y_{l-m}(\theta,\varphi) \quad (2.6.36)$$

这里用到了连带 legendre 多项式的性质 $P_l^{-m}(x) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(x)$。在这里提请读者注意,上面采用的球函数的表达式与 Condon 和 Shortley 所采用的多一个相因子 i^l。对这两种球函数都有 $l_\pm Y_{lm} = \sqrt{(l\mp m)(l\pm m+1)}\,Y_{lm\pm1}$,轨道角动量算符在这两种球函数基下的矩阵相同。但是 Condon 和 Shortley 的球函数不满足式(2.6.36)。指出下面一个事实也很有意思:把实验室绕 y 轴旋转 π 角,空间点的球坐标 $(r,\theta,\varphi)\mapsto(r,\pi-\theta,\pi-\varphi)$,因而有

$$\begin{aligned} Y_{lm}(\theta,\varphi)\mapsto&\exp(-i\pi l_y)Y_{lm}(\theta,\varphi) = Y_{lm}(\pi-\theta,\pi-\varphi)\\ &= (-1)^{l-m}Y_{l-m}(\theta,\varphi) = K_oY_{lm}(\theta,\varphi)\end{aligned} \quad (2.6.37)$$

上式并不表示,在单个无旋粒子态空间上反线性算符 $T = K_o$ 等于线性算符 $\exp(-i\pi l_y)$。

下面我们来证明,当实验室绕 y 轴旋转 π 角时,粒子自旋角动量的本征函数 χ_{sm_s} 将变作

$$\exp(-i\pi s_y)\chi_{sm_s} = (-1)^{s-m_s}\chi_{s-m_s} \quad (2.6.38)$$

事实上,当粒子自旋角量子数 $s=0$ 时,上式成立。当粒子自旋角量子数 $s=\frac{1}{2}$ 时,$s=\frac{1}{2}\boldsymbol{\sigma}$(Pauli 矩阵),特别是 $s_y = \frac{1}{2}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}$。于是 $s_y^2 = \left(\frac{1}{2}\right)^2\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$,$\exp(-i\pi s_y) = \sum_{n\geq0}\frac{\left(\frac{-i\pi}{2}\right)^{2n}}{(2n)!}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \sum_{n\geq0}\frac{\left(\frac{-i\pi}{2}\right)^{2n+1}}{(2n+1)!}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} = \cos\frac{\pi}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \sin\frac{\pi}{2}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$。我们得出 $\exp(-i\pi s_y)\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$,$\exp(-i\pi s_y)\begin{bmatrix}0\\1\end{bmatrix} = -\begin{bmatrix}1\\0\end{bmatrix}$。

最后两式即 $\exp(-i\pi s_y)\chi_{\frac{1}{2}m_s} = (-1)^{\frac{1}{2}m_s}\chi_{\frac{1}{2}-m_s}$。因而式(2.6.37)当 $s=\frac{1}{2}$ 时也成立。下面的命题将保证,式(2.6.38)对自旋角量子数 s 为任意非负半整数的粒子都成立:

命题 设自旋角动量算符 $\boldsymbol{S}_1,\boldsymbol{S}_2$ 互相对易:$[S_{\alpha1},S_{\beta2}]=0$,$\forall\,\alpha,\beta=x,y,z$;$\boldsymbol{S}=$

$\boldsymbol{S}_1+\boldsymbol{S}_2$；$|S_1M_{s1}S_2M_{s2}\rangle$是$\{S_1^2,S_{z1},S_2^2,S_{z2}\}$共同的本征矢量(非耦合基矢)；$|S_1S_2SM_s\rangle$是$\{S_1^2,S_2^2,S^2,S_z\}$共同的本征矢量(耦合基矢)(参考下一章角动量理论)。若$\exp(-i\pi(S_{y1}+S_{y2}))|S_1M_{s1}S_2M_{s2}\rangle=(-1)^{S_1-M_{s1}}(-1)^{S_2-M_{s2}}|S_1-M_{s1}S_2-M_{s2}\rangle$则

$$\exp(-i\pi S_y)\,|S_1S_2SM_s\rangle=(-1)^{S-M_s}\,|S_1S_2S-M_s\rangle \tag{2.6.39}$$

事实上，$|S_1S_2SM_s\rangle=\sum|S_1M_{s1}S_2M_{s2}\rangle\langle S_1M_{s1}S_2M_{s2}|S_1S_2SM_s\rangle$，以$\exp(-i\pi S_y)=\exp(-i\pi(S_{y1}+S_{y2}))$作用于上式两边得，$\exp(-i\pi S_y)\,|S_1S_2SM_s\rangle=\sum_{M_{s1}M_{s2}}(-1)^{S_1-M_{s1}+S_2-M_{s2}}|S_1-M_{s1}S_2-M_{s2}\rangle\langle S_1M_{s1}S_2M_{s2}|S_1S_2SM_s\rangle$注意到$\langle S_1M_{s1}S_2M_{s2}|S_1S_2SM_s\rangle=(-1)^{S-S_1-S_2}\langle S_1-M_{s1}S_2-M_{s2}|S_1S_2S-M_s\rangle$，我们得到$\exp(-i\pi S_y)|S_1S_2SM_s\rangle=(-1)^{S-M_s}\sum_{M_{s1}M_{s2}}|S_1-M_{s1}S_2-M_{s2}\rangle\langle S_1-M_{s1}S_2-M_{s2}|S_1S_2S-M_s\rangle=(-1)^{S-M_s}|S_1S_2S-M_s\rangle$。

反复利用这个命题可得，式(2.6.37)对单个粒子自旋角量子数s为任意非负半整数的情况都成立；甚至对任意量子体系(不一定是单个粒子)的自旋角动量本征矢量都成立。

综上所述，对具有任意自旋s的单个粒子体系，其总角动量的本征函数都有

$$T\,|sljm\rangle=\exp(-i\pi s_y)K_0\,|sljm\rangle=(-1)^{j-m}\,|slj-m\rangle \tag{2.6.40}$$

事实上，上式对任意多粒子量子体系也成立。

Ⅴ.微观运动的可逆性

若$|\psi(t)\rangle(-\tau<t<0)$代表微观体系的一个可能运动，我们要问，$T|\psi(-t)\rangle(0<t<\tau)$是否也是体系的一个可能运动？

事实上，当H不依赖t，由$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle=H|\psi(t)\rangle$可得$-i\hbar\frac{\partial}{\partial t}|\psi(-t)\rangle=H|\psi(-t)\rangle$以反幺正算符$T$作用之

$$i\hbar\frac{\partial}{\partial t}T\,|\psi(-t)\rangle=THT^{-1}T\,|\psi(-t)\rangle \tag{2.6.41}$$

可见要使$T|\psi(-t)\rangle$满足薛定格方程，必须而且只要

$$THT^{-1}=H\qquad(TH=HT) \tag{2.6.42}$$

哈密顿算符H的时间反演不变性是运动可逆性的充要条件。由于T不是线性算符，

即使它与哈密顿算符对易，也不是守恒量。

哈密顿算符 H 的时间反演不变性(2.6.42)导致 Kramers 简并。当体系包含奇数个费米子时，时间反演算符的平方 $T^2 = -1$。设 $H\varphi = \varepsilon\varphi$，则由式(2.6.42)可得 $H(T\varphi) = \varepsilon(T\varphi)$。但是由 T 是反幺正算符，$\langle T\varphi | \varphi\rangle = \langle T\varphi | T^2\varphi\rangle = -\langle T\varphi | \varphi\rangle = 0$，可见 $\langle T\varphi | \varphi\rangle = 0$。能级 ε 至少包含 $\varphi, T\varphi$ 两个线性无关的本征态。这种与体系时间反演不变性相联系的简并叫作 Kramers 简并。

一般孤立体系的运动是可逆的。人们在宇称不一定守恒被发现之后，也开始对运动的可逆性是否自然的基本规律持谨慎态度。1964 年，Christenson 等人观测到长寿命的 K 介子也能衰变成 $\pi^+\pi^-$，从而证明了 CP(C 代表正反粒子变换，P 代表空间反演)不守恒。由于量子场理论证明了 CPT 守恒，这意味着 T 不守恒。

§2.7 力学量算符对易关系的几何根源

从前几节的讨论我们看到，把实验室作一任意平移 $\boldsymbol{\rho}$ 时，很自然地在量子体系态矢量空间 S 上诱导出一个平移算符 $D(\boldsymbol{\rho})$，在线性算符空间 $EndS$ 上也诱导出一个平移算符 $\widetilde{D}(\boldsymbol{\rho})$。$D: T_3(\mathbf{R}) \to GL(S)$ 是三维实欧氏空间 $\mathbf{R}^3$ 的平移群 $T_3(\mathbf{R})$ 在态矢量空间 S 上的一个酉表示。而 $\widetilde{D}: T_3(\mathbf{R}) \to GL(EndS)$ 是 $T_3(\mathbf{R})$ 群在线性算符空间 $EndS$ 上的一个酉表示。酉表示 D 的三个无穷小算符的 i 倍是三个厄米算符，它们代表的力学量当且仅当体系的哈密顿算符在所有平移下都不变时守恒。因此我们把它们认同为该体系的三个动量算符。由于平移群是阿贝尔群，它们任一表示的无穷小算符互相对易。所以任一量子体系的动量算符 P_x, P_y, P_z 两两对易。我们说平移的可换性

$$T(\boldsymbol{\sigma})T(\boldsymbol{\rho}) = T(\boldsymbol{\sigma}+\boldsymbol{\rho}) = T(\boldsymbol{\rho}+\boldsymbol{\sigma}) = T(\boldsymbol{\rho})T(\boldsymbol{\sigma}) \tag{2.7.1}$$

是动量算符 P_x, P_y, P_z 两两对易的几何根源。

类似的论证：把实验室绕空间固定点 O 作一任意旋转 $\boldsymbol{\theta}$ 时，很自然地在量子体系态矢量空间 S 上诱导出一个转动算符 $D(\boldsymbol{\theta})$，在体系的线性算符空间 $EndS$ 上诱导出一个转动算符 $\widetilde{D}(\boldsymbol{\theta})$。$D: SO(3) \to GL(S)$ 是空间绕固定点 O 的转动群 $SO(3)$ 在 S

上的一个酉表示。而 $\widetilde{D}: SO(3) \to GL(EndS)$ 是 $SO(3)$ 群在线性算符空间 $EndS$ 上的一个酉表示。酉表示 D 的三个无穷小算符的 i 倍是三个厄米算符。它们代表的力学量当且仅当体系的哈密顿算符在任一绕固定点 O 的转动下都不变时守恒。所以我们把它们认同为体系的角动量算符 J_x, J_y, J_z。$SO(3)$ 群所有表示的无穷小算符都满足相同的对易关系。由几何关系

$$\boldsymbol{\theta} * \boldsymbol{\varphi} * (-\boldsymbol{\theta}) = R(\boldsymbol{\theta})\boldsymbol{\varphi} \in SO(3) \tag{2.7.2}$$

可得任一量子体系的角动量算符 J_x, J_y, J_z 满足对易关系：

$$[J_\alpha, J_\beta] = i\sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma} J_\gamma \tag{2.7.3}$$

我们说(2.7.2)是角动量对易关系的几何根源。对易关系(2.7.3)是所有角动量算符都遵从的，不管其具体形式如何。

我们还看到空间反射 $I\boldsymbol{r} = -\boldsymbol{r}$ 在量子体系态矢量空间 S 上所诱导出的宇称算符 Π 与该体系的角动量算符对易。

$$[\Pi, \boldsymbol{J}] = 0 \tag{2.7.4}$$

它的几何根源在于关于 O 点的空间反射与绕 O 点的旋转交换

$$[I, R(\boldsymbol{\theta})] = 0, \forall\, \boldsymbol{\theta} \in SO(3) \tag{2.7.5}$$

我们还知道体系的角动量算符和动量算符之间满足对易关系

$$[J_\alpha, P_\beta] = i\sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma} P_\gamma \tag{2.7.6}$$

而同一体系的宇称算符与动量算符反对易

$$\Pi\boldsymbol{P} + \boldsymbol{P}\Pi = 0 \tag{2.7.7}$$

在本节我们将统一的探讨所有以上对易关系和反对易关系的几何根源。我们的讨论将不依赖具体的量子体系的信息，因而所得出的结论对所有量子体系都有效，是普适的。这样的探讨很有意义。譬如说，它使我们确信自由刚体在量子化后其角动量算符一定满足对易关系(2.7.3)，其动量算符一定互相对易，而角动量算符和动量算符之间的对易关系为(2.7.6)，即使我们不知道这些算符的具体形式。对无穷多自由度的体系场也一样，量子场的动量算符，角动量算符间的对易关系和任一量子体系的动量算符角动量算符满足同样的对易关系。

Ⅰ. 三维欧几里德群 $E(3)$

为了避免术语名词的滥用,我们在此重述"变换"一词的定义:设 A 是非空集,$\varphi: A\to A$ 是 A 到自身的映射。若 φ 是双射(双方一一对应),则称 φ 是 A 的一个变换。

现在考虑 A 的变换的全体所成的集合 $\mathbf{T}(A)$。对 $\psi,\varphi\in\mathbf{T}(A)$,定义 ψ 与 φ 的乘积为复合映射 $\varphi\circ\psi: A\to A$。由 §0.0 知,$\varphi\circ\psi$ 仍是 A 到 A 的双射,这样定义的 A 上的代数运算是结合的。A 的恒等映射 1_A 是这个乘法的单位元,每个 A 到 A 的双射的逆映射仍是 A 到 A 的双射,是它的逆元。于是 $\mathbf{T}(A)$在复合映射这个二元合成下构成一个群。它的任一子群被称为非空集 A 的一个变换群。

现在考虑我们生活在其中的空间,三维实欧氏空间 $\mathbf{R}^3$。我们感兴趣的并非 $\mathbf{R}^3$ 的所有变换,而是它的保持任意两点距离不变的变换。这类保距变换的全体构成群 $\mathbf{T}(\mathbf{R}^3)$的一个子群,称为三维欧几里德群,记作 $E(3)$。空间平移 $T(\boldsymbol{\rho})$,空间绕固定点 O 的转动 $R(\boldsymbol{\theta})$,空间关于固定点 O 的反射 I 都是 $\mathbf{R}^3$ 的保距变换。

$\mathbf{R}^3$ 的恒等变换 $e(e\boldsymbol{r}=\boldsymbol{r})$和关于点 O 的空间反射 $I(I\boldsymbol{r}=-\boldsymbol{r})$构成二阶群 $\{e,I\}$。空间绕 O 点的转动 $R(\boldsymbol{\theta})$的全体构成三维李群 $SO(3)$。注意到这两个群只有一个公共元 $e=R(\boldsymbol{0})$,而且关于点 O 的空间反射与绕点 O 的转动是可易的:

$$IR(\boldsymbol{\theta})=R(\boldsymbol{\theta})I,\ \forall R(\boldsymbol{\theta})\in SO(3) \tag{2.7.8}$$

于是可以由它们构造直积群 $SO(3)\otimes\{e,I\}$。这个群是由 $\mathbf{R}^3$ 的全体保持 O 点不动的保距变换构成的,正是三维实正交群,记作 $O(3)$。

$$SO(3)\otimes\{e,I\}=O(3) \tag{2.7.9}$$

直积群 $O(3)$中的任一元可以唯一地表作 $R(\boldsymbol{\theta})X$,其中 $R(\boldsymbol{\theta})\in SO(3)$,$X\in\{e,I\}$。$O(3)$群两个元的乘积为

$$R(\boldsymbol{\varphi})YR(\boldsymbol{\theta})X=R(\boldsymbol{\varphi})R(\boldsymbol{\theta})YX \tag{2.7.10}$$

$R(\boldsymbol{\theta})X$ 的逆元为

$$(R(\boldsymbol{\theta})X)^{-1}=R(\boldsymbol{\theta})^{-1}X^{-1} \tag{2.7.11}$$

由刚体运动学知,(三维自由)刚体的任一位移($\mathbf{R}^3$ 的保持笛卡尔坐标系手征不变的等距变换)总可以这样实现:先绕 O 点作一转动 $R(\boldsymbol{\theta})\in SO(3)$,接着把刚体作一平移 $T(\boldsymbol{\rho})\in T_3(\mathbf{R})$,而且这里的 $R(\boldsymbol{\theta})$和 $T(\boldsymbol{\rho})$由刚体的位移唯一地确定。把刚体的位移看成是 $\mathbf{R}^3$ 的一个变换(当然是保距的)则刚体位移的全体构成 $E(3)$群的一个

子群，称之为刚体运动群，记作 **M**。于是 $\mathbf{M}=\{T(\boldsymbol{\rho})R(\boldsymbol{\theta})\mid T(\boldsymbol{\rho})\in T_3(\mathbf{R}),R(\boldsymbol{\theta})\in SO(3)\}$。这是一个 6 维李群。

注意到这样一些事实：固定在刚体上的两个原点不同但取向一样（对应坐标轴正平行）的笛卡尔坐标系，在刚体转动下保持取向相同；平移是指保持本体坐标系取向不变的刚体位移。我们很容易看出

$$R(\boldsymbol{\theta})T(\boldsymbol{\rho})R(-\boldsymbol{\theta})=T(R(\boldsymbol{\theta})\boldsymbol{\rho}) \tag{2.7.12}$$

即先绕 O 点（空间的一个固定点）作旋转 $R(\boldsymbol{\theta})^{-1}=R(-\boldsymbol{\theta})$，接着把空间作一平移 $T(\boldsymbol{\rho})$，最后把空间绕 O 点作一旋转 $R(\boldsymbol{\theta})$，相当于把空间作一平移 $T(R(\boldsymbol{\theta})\boldsymbol{\rho})$。刚体运动群 **M** 中的两个元素的乘法为

$$\begin{aligned}(T(\boldsymbol{\rho}_2)R(\boldsymbol{\theta}_2))(T(\boldsymbol{\rho}_1)R(\boldsymbol{\theta}_1))&=T(\boldsymbol{\rho}_2)(R(\boldsymbol{\theta}_2)T(\boldsymbol{\rho}_1)R(-\boldsymbol{\theta}_2))R(\boldsymbol{\theta}_2)R(\boldsymbol{\theta}_1)\\&=T(\boldsymbol{\rho}_2)T(R(\boldsymbol{\theta}_2)\boldsymbol{\rho}_1)R(\boldsymbol{\theta}_2)R(\boldsymbol{\theta}_1)\\&=T(\boldsymbol{\rho}_2+R(\boldsymbol{\theta}_2)\boldsymbol{\rho}_1)R(\boldsymbol{\theta}_2*\boldsymbol{\theta}_1)\end{aligned} \tag{2.7.13}$$

由式(2.7.13)可知，(纯)平移的共轭元仍是(纯)平移：

$$(T(\boldsymbol{\rho})R(\boldsymbol{\theta}))T(\boldsymbol{\rho}')(T(\boldsymbol{\rho})R(\boldsymbol{\theta}))^{-1}=T(R(\boldsymbol{\theta})\boldsymbol{\rho}')\in T_3(\mathbf{R}) \tag{2.7.14}$$

故 $T_3(\mathbf{R})$ 是群 **M** 的正规子群。而刚体运动群 **M** 是它的正规子群 $T_3(\mathbf{R})$ 和子群 $SO(3)$ 的半直积：

$$\mathbf{M}=T_3(\mathbf{R})\otimes_S SO(3) \tag{2.7.15}$$

三维欧几里德群 $E(3)$ 中的元素或者是刚体位移，或者是先作关于 O 点的空间反射，再作一刚体位移。可见刚体运动群 **M** 是 $E(3)$ 的指数(index)为 2 的子群，其左陪集与右陪集相同，故一定是 $E(3)$ 的正规子群。$\{e,I\}$ 也是 $E(3)$ 的子群。而 $E(3)$ 是 M 与 $\{e,I\}$ 的半直积：

$$E(3)=M\otimes_S\{e,I\} \tag{2.7.16}$$

$\mathbf{R}^3$ 的任一保距变换 $g\in E(3)$ 都可以唯一地写作

$$g=T(\boldsymbol{\rho})R(\boldsymbol{\theta})X \tag{2.7.17}$$

---

* 设 G 是群，N 是 G 的正规子群，H 是 G 的子群。若 G 的每一元素 $g\in G$ 可以唯一地表作 $g=nh(n\in N,h\in H)$，则称 G 是 N 与 H 的半直积，记作 $G=N\otimes_S H$。

其中 $T(\boldsymbol{\rho})\in T_3(\mathbf{R}), R(\boldsymbol{\theta})\in SO(3), X\in\{e, I\}$。

练习 2-7-1 试证:三维欧几里德群 $E(3)$是平移群 $T_3(\mathbf{R})$和关于空间某一点 O 的三维正交群 $O(3)$的半直积:$E(3)=T_3(\mathbf{R})\otimes O(3)$。

Ⅱ. $E(3)$群在态矢量空间 S 上的表示

把空间 $\mathbf{R}^3$ 作一等距变换 $g\in E(3)$时,很自然地量子体系的态矢量要经受一个幺正变换 $\psi\mapsto D(g)\psi$。相继地作二次等距变换先 g 后 h,与作一次变换 hg 的结果是一样的,故应有

$$D(h)D(g)=D(hg) \tag{2.7.18}$$

对应于 $\mathbf{R}^3$ 的恒等变换 $e\in E(3)$,态矢量应该不变 $D(e)=1_s$,(1_s 是态矢量空间 S 上的恒等算符)。故 $D: E(3)\to GL(S)$是 $E(3)$群在态矢量空间 S 上的酉表示。由于 $\mathbf{R}^3$ 的任一保距变换 $g\in E(3)$都可以唯一地写作 $g=T(\boldsymbol{\rho})R(\boldsymbol{\theta})X$, $E(3)$群在态矢量空间 S 上的表示算符可以记作 $D(g)=D(\boldsymbol{\rho},\boldsymbol{\theta},X)$。特别是,$D(\boldsymbol{\rho},\boldsymbol{0},e)$, $D(\boldsymbol{0},\boldsymbol{\theta},e)$分别为前面所讲的态矢量空间 S 上的平移算符 $D(\boldsymbol{\rho})$和转动算符 $D(\boldsymbol{\theta})$。$E(3)$群的元素需要用 6 个连续变化的实参数 $P_x, P_y, P_z, \theta_x, \theta_y, \theta_z$ 和一个取两个值的参数 X 刻画。故 $E(3)$是 6 维李群。我们已经阐述过,为什么酉表示 D 关于这六个参数的无穷小算符的 i 倍分别是体系的动量和角动量算符。让我们考察与空间反演 $I=T(\boldsymbol{0})R(\boldsymbol{0})I$ 对应的酉算符 $D(I)=D(\boldsymbol{0},\boldsymbol{0},I)$。由于 $I^2=e$,于是 $D(I)D(I)=\mathbf{1}_s$, $D(I)=D(I)^{-1}$。考虑到幺正性 $D(I)^+=D(I^{-1})$,可知 $D(I)=D(I)^+$。我们称这个厄米算符是体系的宇称算符,并记作 $\Pi=D(I)$。

在前面我们已经利用 $T(\boldsymbol{\rho})T(\boldsymbol{\delta})=T(\boldsymbol{\delta})T(\boldsymbol{\rho})$,证明了动量算符间互相对易,利用式(**2.3.9**)证明了角动量算符之间满足对易关系 $[J_\alpha, J_\beta]=i\sum_{\gamma=1}^{3}\varepsilon_{\alpha\beta\gamma}J_\gamma$,现在让我们利用(**2.5.7**)证明了任一量子体系的宇称算符与角动量算符对易。事实上,由(**2.5.7**)可知,$D(\mathbf{0},\boldsymbol{\theta},e)D(I)=D(I)D(0,\boldsymbol{\theta},e)$。取等式两边对 θ_α 的微商,并令 $\boldsymbol{\theta}=0$ 便得 $\boldsymbol{J}\Pi=\Pi\boldsymbol{J}$。同样可以利用(2.5.10)来证明对任一量子体系

$$\Pi\boldsymbol{P}+\boldsymbol{P}\Pi=0 \tag{2.7.19}$$

事实上,由 $T(\boldsymbol{\rho})I=IT(-\boldsymbol{\rho})$可知

$$D(\boldsymbol{\rho},0,e)\Pi=\Pi D(-\boldsymbol{\rho},0,e) \tag{2.7.20}$$

取式(2.7.20)的两边对 ρ_α 的微商,并令 $\boldsymbol{\rho}=0$,便得

$$P_\alpha \Pi = -\Pi P_\alpha \tag{2.7.21}$$

为得到动量算符与角动量算符之间的对易关系,我们从几何关系(2.7.12)出发,利用表示的同态性质我们有

$$D(\boldsymbol{0},\boldsymbol{\theta})D(\boldsymbol{\rho},\boldsymbol{0})D(\boldsymbol{0},-\boldsymbol{\theta}) = D(R(\boldsymbol{\theta})\boldsymbol{\rho},\boldsymbol{0}) \tag{2.7.22}$$

取上式两边的偏导数$\frac{\partial}{\partial\theta_\alpha}$在$\boldsymbol{\theta}=0$ 处的值,我们就得到

$$J_\alpha D(\boldsymbol{\rho},\boldsymbol{0}) - D(\boldsymbol{\rho},\boldsymbol{0})J_\alpha = i\frac{\partial(R_{\beta\sigma}(\boldsymbol{\theta})\rho_\sigma)}{\partial\theta_\alpha}\frac{\partial(D(R(\boldsymbol{\theta})\boldsymbol{\rho},\boldsymbol{0}))}{\partial(R_{\beta\gamma}(\boldsymbol{\theta})\rho_\gamma)}\Bigg|_{\boldsymbol{\theta}=0}$$

$$= -i\varepsilon_{\alpha\beta\sigma}\rho_\sigma\frac{\partial D(\boldsymbol{\rho},\boldsymbol{0})}{\partial\rho_\beta}$$

再取上式两边的偏导数$\frac{\partial}{\partial\rho_\mu}$在$\boldsymbol{\rho}=0$ 处的值,利用 $\varepsilon_{\alpha\beta\gamma}$关于下标交换的反对称性可得

$$J_\alpha P_\mu - P_\mu J_\alpha = -\varepsilon_{\alpha\beta\mu}P_\beta$$

上式就是

$$[J_\alpha P_\beta] = i\sum_{\gamma=1}^{3}\varepsilon_{\alpha\beta\gamma}P_\gamma \tag{2.7.23}$$

注意,以上得到的关于 $\Pi,\boldsymbol{P},\boldsymbol{J}$ 之间的对易(反对易)关系,对任一量子体系都成立,但必须是同一量子体系的(总)宇称,(总)动量和(总)角动量算符。

练习 2-7-2 思考题:为什么量子体系的角动量、动量是各子体系角动量,动量之和,而体系的宇称却是个子体系宇称算符之积?(从群表示论的角度找原因)

§2.8 量子体系的对称性群(symmetry group)与能级简并

Ⅰ. 量子体系的对称变换

上一节我们讨论了三维实欧氏空间 $\mathbf{R}^3$ 的保距变换的全体所构成的群,三维实欧几里得群 $E(3)$。把量子体系所在的三维实欧氏空间 $\mathbf{R}^3$ 作一保距变换 $g=T(\boldsymbol{\rho})R(\boldsymbol{\theta})X(X\in\{e,I\})$,则在量子体系的态矢量空间 S 上自然地诱导出一个幺正线性算

符 $D(g)$:

$$D(g)\Psi(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_N)=\begin{cases}e^{-i\boldsymbol{\varphi}\cdot\boldsymbol{P}}e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}}\Psi(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_N) & \text{当 } X=e\\ e^{-i\boldsymbol{\varphi}\cdot\boldsymbol{P}}e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}}\Pi\Psi(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_N) & \text{当 } X=I\end{cases}$$

$$=\begin{cases}e^{-i\boldsymbol{\theta}\cdot\boldsymbol{S}}\Psi(R(-\boldsymbol{\theta})(\boldsymbol{r}_1-\boldsymbol{\rho}),\cdots,R(-\boldsymbol{\theta})(\boldsymbol{r}_N-\boldsymbol{\rho})) & \text{当 } X=e\\ e^{-i\boldsymbol{\theta}\cdot\boldsymbol{S}}\omega_1\cdots\omega_N\Psi(-R(-\boldsymbol{\theta})(\boldsymbol{r}_1-\boldsymbol{\rho}),\cdots,-R(-\boldsymbol{\theta})(\boldsymbol{r}_N-\boldsymbol{\rho})) & \text{当 } X=I\end{cases}\tag{2.8.1}$$

式中 $\Psi(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_N)$是列矩阵(对应于自旋表象),$\boldsymbol{S}$ 是 N 个粒子自旋之和,ω_j 是第 i 个粒子的内禀宇称。$\{D(g)\mid g\in E(3)\}$是 $E(3)$群在 S 上的一个酉表示。在 $\mathbf{R}^3$ 的保距变换 $g\in E(3)$下,量子体系的线性算符 $A\in EndS$ 也自然地经受一个酉变换

$$A\mid\rightarrow D(g)AD(g^{-1})\equiv\widetilde{D}(g)A\tag{2.8.2}$$

$\{\widetilde{D}(g)\mid g\in E(3)\}$构成 $E(3)$群在 $EndS$ 上的一个酉表示。

量子体系的全部动力学性质由体系的哈密顿算符 H 所确定。在 $\mathbf{R}^3$ 的保距变换 $g\in E(3)$下,H 按下式变换

$$H\rightarrow D(g)HD(g^{-1})\tag{2.8.3}$$

一般而言,$D(g)HD(g^{-1})$与 H 不一定相等。对给定的 $g\in E(3)$,等式 $H=D(g)HD(g^{-1})$是否成立,与具体量子体系有关。下面我们给出量子体系对称变换的定义:

定义:设 $g\in e(3)$,若对某个量子体系有 $D(g)HD(g^{-1})=H$,此即

$$D(g)H=HD(g)\tag{2.8.4}$$

则称 $g\in E(3)$是该量子体系的对称变换。

不难证明,量子体系的对称变换的全体构成 $E(3)$群的一个子群(请读者证明之),称之为该量子体系的对称性群(symmetry group)。注意,请勿与对称群(symmetric group)混淆。n 次对称群 S_n 是 $n!$ 个 n 次置换所构成的群,它的任一子群称之为置换群。这样定义的对称性群显然刻画的是量子体系的几何对称性。

单个自由粒子的哈密顿算符为 $H=\dfrac{\boldsymbol{p}^2}{2m}$,它的对称性群是 $E(3)$。

碱金属原子的哈密顿算符为 $H=\dfrac{\boldsymbol{p}^2}{2\mu}+U(r)+\dfrac{1}{2\mu^2c^2r}\dfrac{\mathrm{d}U(r)}{\mathrm{d}r}\boldsymbol{l}\cdot\boldsymbol{s}$,它的对称性群是 $O(3)$群。

练习 2-8-1 试证:量子体系所有对称变换构成 $E(3)$群的一个子群。

Ⅱ.能级简并与对称性

让我们考察量子体系的对称性是如何与能级简并相联系的。设 $u\in S$ 是体系哈密顿算符 H 的属于特征值 ε 的特征向量 $Hu=\varepsilon u$。以 G 记体系的对称性群。对体系的任一对称变换 $g\in G$,我们有 $D(g)H=HD(g)$,从而 $HD(g)u=\varepsilon D(g)u$。考虑到 $D(g)$是幺正算符,$D(g)u$ 不会是零向量,故知 $D(g)u$ 仍是 H 的属于特征值 ε 的特征向量。这样若 $D(g)u$ 不是 u 的复数倍(u 不是 $D(g)$的特征向量),则我们得到一个与 u 线性无关的算符 H 的属于特征值 ε 的特征向量。从 H 的属于特征值 ε 一个特征向量 u 出发,作用以所有对称变换 $g\in G$,我们得到一组属于特征值 ε 的能量本征函数 $\{D(g)u\mid g\in G\}$。当 u 不是所有 $D(g)(g\in G)$的共同的特征向量时,$\{D(g)u\mid g\in G\}$张成的 S 的子空间的维数 d 大于 1,能级 ε 就会简并。

上面的讨论勾勒出对称性和能级简并的大致关系。为了更精准、更深入地探讨这一问题,我们须使用群表示的语言。

体系的哈密顿算符 H 是体系态矢量空间 S 上的厄米算符,故 S 可以分解成 H 的两两正交的特征子空间的直和(参考 §0.3)$S=\bigoplus_n S_n$。由于对任一对称变换 $g\in G$ 都有 $D(g)H=HD(g)$,故 H 的每一个特征子空间 S_n 是所有 $D(g)(g\in G)$的不变子空间。因而 $D(g)$到不变子空间 S_n 上的限制构成对称性群 G 在 S_n 上的一个酉表示:$\{D(g)\mid_{S_n}\mid g\in G\}$。我们要问,"能级的简并全部是由对称性引起的吗?"这个问题相当于问,"对称性群在能量特征子空间 S_n 上的表示是不可约的吗?"

暂时我们不能回答这一问题。让我们考察下面的例子。

碱金属原子的对称性群是 $O(3)$群。能量与磁量子数无关。它的每个能级 ε_{nlj}荷载 $O(3)$群的一个 j 权($2j+1$ 维)不可约表示。

处在平面中心势场中的二维无旋粒子,在极坐标下,体系的哈密顿算符为

$$H=\frac{-\hbar^2}{2\mu}\left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho}+\frac{1}{\rho^2}\frac{\partial^2}{\partial\varphi^2}\right)+U(\rho) \tag{2.8.5}$$

容易验证 $l_z=-i\hbar\dfrac{\partial}{\partial\varphi}$是守恒量:

$$[l_z,H]=0 \tag{2.8.6}$$

故 H, l_z 有完备的共同的本征函数系。它们的共同的本征函数一定可以写作

$$R(\rho)e^{im\varphi}(m = 0, \pm 1, \pm 2, \cdots) \tag{2.8.7}$$

径向部分满足的特征方程为

$$\left[\frac{-\hbar^2}{2\mu}\left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho} - \frac{m^2}{\rho^2}\right) + U(\rho)\right]R(\rho) = ER(\rho) \tag{2.8.8}$$

由上式可以看出，当 $m \neq 0$ 时，能级是二度简并的。

如果我们认为这个体系的对称性群是 $SO(2)$群，考虑到 $SO(2)$是个阿贝尔群，所有不可约表示是一维的，我们将不能用对称性说明这种简并。但 $SO(2)$群并不能穷尽这个体系的对称性群。因为关于过原点 O 的任一直线的反射也是这个体系的对称变换。在这种对称变换下，$R(\rho)e^{im\varphi}$变为 $R(\rho)e^{-im\varphi}$。体系的对称性群是非阿贝尔群 $O(2)$。每个能级荷载 $O(2)$群的不可约表示。

上面的观察使我们期望对称性原则上可以解释全部能级简并，而把一些例外当作是巧合，所谓的“偶然简并”；它们是势能中的参数取某些特殊值时发生的。但是氢原子和三维各向同性谐振子的能级简并度要远比对称性分析所预期的要大。例如氢原子能级 ϵ_n 的简并度为 n^2(不考虑电子自旋)，能量与角量子数 l 和磁量子数 m 都无关。而 $SO(3)$对称性只能说明能量与磁量子数 m 无关。这种偏离似乎是系统的，不能被归为“偶然简并”。让我们进一步地考察氢原子。

Ⅲ. 氢原子的动力学对称性

类氢离子或氢原子的哈密顿量为(不考虑电子自旋)

$$H = \frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r} \qquad (\alpha = Ze^2 > 0) \tag{2.8.9}$$

显然它在绕力心 O 的转动下是不变的。$\{H, \boldsymbol{L}^2, L_z\}$构成一个好力学量的完全集。它们的共同的本征函数$\{|nlm\rangle | n \geq 1, 0 \leq l \leq n-1, -l \leq m \leq l\}$构成态矢量空间 S 的一个正交归一基。

$$\begin{aligned} H|nlm\rangle &= \epsilon_n |nlm\rangle \\ \boldsymbol{L}^2|nlm\rangle &= l(l+1)\hbar^2|nlm\rangle \\ L_z|nlm\rangle &= m\hbar|nlm\rangle \end{aligned} \tag{2.8.10}$$

式中

$$\varepsilon_n = -\frac{\mu\alpha^2}{2\hbar^2 n^2} = -\frac{\mu z^2 e^4}{2\hbar^2 n^2} \tag{2.8.11}$$

上面的结果表明,除基态($n=1$)外,每个能级都荷载 $SO(3)$群的一个可约表示。ε_n 能级荷载的 $SO(3)$群的表示是

$$D_n = \bigoplus_{l=0}^{n-1} D^{[l]} \tag{2.8.12}$$

当 $n>1$ 时,这是一个 n^2 维的可约表示。为了寻找能级简并度远大于预期的原因,让我们回忆守恒量与对称性的关系。由 §2.2 和 §2.3 知,当且仅当体系的动量守恒 $[p_\alpha, H]=0,(\alpha=1,2,3)$平移 $T(\boldsymbol{\rho})\in E(3)$是体系的对称变换;当且仅当体系的角动量守恒$[J_\alpha, H]=0,(\alpha=1,2,3)$,转动 $R(\boldsymbol{\theta})\in E(3)$是体系的对称变换。这启发我们,要寻找体系的更多的对称变换,应该去寻找体系的更多的运动积分(守恒量)。为此,让我们在此先回顾一下经典力学中质点在点电荷库仑场中的运动。

质点在势场 $U(\boldsymbol{r},t)$中所受到的力为 $-\nabla U(\boldsymbol{r},t)$。它的运动遵循正则运动方程

$$\begin{aligned}\dot{\boldsymbol{r}} &= \frac{\partial H(\boldsymbol{r},\boldsymbol{p})}{\partial \boldsymbol{p}}\\ \dot{\boldsymbol{p}} &= -\frac{\partial H(\boldsymbol{r},\boldsymbol{p})}{\partial \boldsymbol{r}}\end{aligned} \tag{2.8.13}$$

其中 H 是质点的哈密顿量 $H=\frac{\boldsymbol{p}^2}{2\mu}+U(\boldsymbol{r},t)$。一个重要的特殊情况是势场是恒定的:$U(\boldsymbol{r},t)=U(\boldsymbol{r})$。这时质点的能量守恒$\frac{\mathrm{d}H}{\mathrm{d}t}=\dot{\boldsymbol{r}}\cdot\frac{\partial H}{\partial \boldsymbol{r}}+\dot{\boldsymbol{p}}\cdot\frac{\partial H}{\partial \boldsymbol{p}}=0$。另一个重要的特殊情况是势场有球对称性(是中心力场):$U(\boldsymbol{r},t)=U(r,t)$。这时质点所受的力 $-\frac{\boldsymbol{r}}{r}\frac{\partial U(r,t)}{\partial r}$指向力心或沿矢径向外,关于力心的力矩为零。我们有角动量守恒:$\boldsymbol{l}=\boldsymbol{r}\times\boldsymbol{p}$ 是常矢量。注意从角动量的定义式可知 $\boldsymbol{l}$ 与 $\boldsymbol{r}$ 垂直,$\boldsymbol{l}$ 与 $\boldsymbol{p}$ 垂直,这与运动方程无关(与哈密顿量无关)。既然在中心力场中质点的角动量守恒,而,$\boldsymbol{r}$ 和 $\boldsymbol{p}$ 总是与 $\boldsymbol{l}$ 垂直,可知质点在与 $\boldsymbol{l}$ 垂直的过力心的平面内运动。如果中心力场是恒定的:$U(r,t)=U(r)$,我们将同时有能量守恒和角动量守恒。质点在过力心的以 $\boldsymbol{l}$ 为法向量的平面内运动。当 $E<0$ 时,运动轨道一般不一定闭合。点电荷的库仑场和点质量的引力场是一类重要的特殊中心力场

$$U(r) = -\frac{\alpha}{r} \qquad (\alpha>0) \tag{2.8.14}$$

能量小于零时,质点的运动轨道是以力心为焦点之一的椭圆(当然是闭合的)。我们将会看到这时除了能量和角动量外,还有其它的运动积分。考察 *Runge-Lenz* 矢量

$$\boldsymbol{R} = \frac{1}{\mu\alpha}\boldsymbol{p} \times \boldsymbol{l} - \frac{\boldsymbol{r}}{r} \tag{2.8.15}$$

从这个定义立即可得

$$\boldsymbol{R} \cdot \boldsymbol{l} = \boldsymbol{l} \cdot \boldsymbol{R} = 0 \tag{2.8.16}$$

$$\boldsymbol{R} \cdot \boldsymbol{R} = \frac{2}{\mu\alpha^2}\left(\frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r}\right)l^2 + 1 \tag{2.8.17}$$

事实上,由 $\boldsymbol{r}$ 与 $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$ 垂直,$\boldsymbol{l}$ 与 $\boldsymbol{p} \times \boldsymbol{l}$ 垂直,我们马上可得到式(2.8.16)。由单位向量长度平方为 1 知 $\frac{\boldsymbol{r}}{r} \cdot \frac{\boldsymbol{r}}{r} = 1$;由 $\boldsymbol{p}$ 与 $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$ 垂直知 $(\boldsymbol{p} \times \boldsymbol{l}) \cdot (\boldsymbol{p} \times \boldsymbol{l}) = \boldsymbol{p}^2\boldsymbol{l}^2$;由恒等式 $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = (\boldsymbol{c} \times \boldsymbol{a}) \cdot \boldsymbol{b}$ 知 $(\boldsymbol{p} \times \boldsymbol{l}) \cdot \boldsymbol{r} = (\boldsymbol{r} \times \boldsymbol{p}) \cdot \boldsymbol{l} = \boldsymbol{l}^2$。利用这些结果立即推出式(2.8.17)。注意(2.8.16)和(2.8.17)是两个恒等式,它们与运动方程无关(不依赖哈密顿量)。利用运动方程,我们可以证明在中心力场(2.8.14)中,$\boldsymbol{R}$ 是运动积分。事实上,

$$\dot{\boldsymbol{R}} = \frac{1}{\mu\alpha}(\dot{\boldsymbol{p}} \times \boldsymbol{l} + \boldsymbol{p} \times \dot{\boldsymbol{l}}) - \frac{\dot{\boldsymbol{r}}}{r} + \frac{(\boldsymbol{r} \cdot \dot{\boldsymbol{r}})\boldsymbol{r}}{r^3} = \frac{-\boldsymbol{r} \times \boldsymbol{l}}{\mu r^3} + \frac{(\boldsymbol{r} \cdot \boldsymbol{p})\boldsymbol{r} - (\boldsymbol{r} \cdot \boldsymbol{r}) \cdot \boldsymbol{p}}{\mu r^3} = 0 \tag{2.8.18}$$

最后一步利用了等式 $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$。与恒等式(2.8.16)和(2.8.17)不同,式(2.8.18)是运动方程的结果(体现在推导过程中利用了 $\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{r}} = -\frac{\alpha \boldsymbol{r}}{r^3}$),与体系哈密顿量 $H = \frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r}$ 有关。记住上面所讨论的经典力学量间的各种恒等关系式和在哈密顿量为 $H = \frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r}$ 时的各种运动积分,对我们寻找它们的量子对应(quantum counterparts)是非常有益的。

现在让我们回到量子力学。为书写简便起见,我们在下面仍以 A 表示力学量 A 的算符 $\hat{A}$,记略去表示算符的帽子"^"。应该注意的是,在公式推演中算符乘法不对易。

量子力学中氢原子(类氢离子)体系的相对运动的哈密顿算符为(不考虑电子自旋)

$$H = \frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r} \qquad (\alpha = Ze^2 > 0) \tag{2.8.19}$$

能量 H,轨道角动量 $\boldsymbol{l}$ 是运动积分。让我们考察 Runge-Lenz 矢量。它的算符应是厄米的,故经典力学量关系(2.8.15)应改为算符关系

$$\boldsymbol{R} = \frac{1}{2\mu\alpha}(\boldsymbol{p} \times \boldsymbol{l} - \boldsymbol{l} \times \boldsymbol{p}) - \frac{\boldsymbol{r}}{r} \tag{2.8.20}$$

上式中的

$$\begin{aligned} -\boldsymbol{l} \times \boldsymbol{p} &= -\hat{e}_\rho \varepsilon_{\rho\sigma\tau} l_\sigma p_\tau = -\hat{e}_\rho \varepsilon_{\rho\sigma\tau} p_\tau l_\sigma - \hat{e}_\rho \varepsilon_{\rho\sigma\tau}[l_\sigma, p_\tau] \\ &= \hat{e}_\rho \varepsilon_{\rho\sigma\tau} p_\sigma l_\tau - \hat{e}_\rho \varepsilon_{\rho\sigma\tau} i\hbar \varepsilon_{\sigma\tau\gamma} p_\gamma = \boldsymbol{p} \times \boldsymbol{l} - i\hbar 2\delta_{\rho\gamma} \hat{e}_\rho p_\gamma \\ &= \boldsymbol{p} \times \boldsymbol{l} - 2i\hbar \boldsymbol{p}(\text{参考本节附录}) \end{aligned}$$

故式(2.8.20)也可以写作

$$\boldsymbol{R} = \frac{1}{\mu\alpha}(\boldsymbol{p} \times \boldsymbol{l} - i\hbar \boldsymbol{p}) - \frac{\boldsymbol{r}}{r} \tag{2.8.21}$$

当 $\hbar \to 0$ 时,上式变作经典关系(2.8.15)。利用本节附录中的算符恒等式和三维 Civita 符号性质,我们还可以推出

$$\boldsymbol{l} \cdot \boldsymbol{R} = 0 \qquad \boldsymbol{R} \cdot \boldsymbol{l} = 0 \tag{2.8.22}$$

这与经典力学量关系(2.8.16)相对应;

$$\boldsymbol{R} \cdot \boldsymbol{R} = 1 + \frac{2}{\mu\alpha^2}\left(\frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r}\right)(\boldsymbol{l}^2 + \hbar^2) \tag{2.8.23}$$

这与经典力学量关系(2.8.17)相对应;以及对易关系

$$\left[\boldsymbol{R}, \frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r}\right] = 0 \tag{2.8.24}$$

这表明当 $H = \frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r}$ 时,$\boldsymbol{R}$ 是运动积分;

$$[l_\rho, R_\sigma] = i\hbar \varepsilon_{\rho\sigma\tau} R_\tau \tag{2.8.25}$$

这表明在绕力心转动时,$\boldsymbol{R}$ 按向量方式变换;

$$[R_\rho, R_\sigma] = -\frac{2}{\mu\alpha^2}\left(\frac{\boldsymbol{p}^2}{2\mu} - \frac{\alpha}{r}\right) i\hbar \varepsilon_{\rho\sigma\tau} l_\tau \tag{2.8.26}$$

对易关系(2.8.25)和(2.8.26),当 $\hbar \to 0$ 时其右边趋于零。考察 $l_\sigma, R_\sigma (\sigma = 1,2,3)$ 这六个运动积分间的对易关系(2.8.25)、(2.8.26)和角动量对易关系

$$[l_\rho, l_\sigma] = i\hbar \varepsilon_{\rho\sigma\tau} l_\tau \tag{2.8.27}$$

我们发现这六个与 $H=\frac{\boldsymbol{p}^2}{2\mu}-\frac{\alpha}{r}$ 对易的厄米算符所张成的 6 维线性空间在李乘法[,]下并不封闭(由于式(2.8.26)的右边出现氢原子的哈密顿算符 $H=\frac{\boldsymbol{p}^2}{2\mu}-\frac{\alpha}{r}$);它们并不张成某个 6 维李代数,并不对应某个 6 维李群。但是对易关系(2.8.24)、(2.8.25)、(2.8.26)和(2.8.27)给我们启示,如何引入态矢量空间 S 上的线性算符 $A_\sigma(\sigma=1,2,3)$,使它们是厄米的

$$A_\sigma^+ = A_\sigma \tag{2.8.28}$$

并且是运动积分

$$\left[\boldsymbol{A}, H=\frac{\boldsymbol{p}^2}{2\mu}-\frac{\alpha}{r}\right]=0 \tag{2.8.29}$$

在空间绕力心 O 转动时按向量方式变换

$$[l_\rho, A_\sigma] = i\hbar\varepsilon_{\rho\sigma\tau}A_\tau \tag{2.8.30}$$

互相间还满足如下的对易关系

$$[A_\rho, A_\sigma] = i\hbar\varepsilon_{\rho\sigma\tau}l_\tau \tag{2.8.31}$$

最后两个等式(2.8.30)和(2.8.31)加上(2.8.27)确保 $l_\sigma, A_\sigma(\sigma=1,2,3)$ 张成的 6 维线性空间在李乘法[,]下封闭,构成一个李代数,对应于一个由酉算符生成的 6 维李群,每个群元都与 H 对易。

由 §0.3 知,要定义线性空间 V 上的一个线性算符,只须规定它对 V 的某个基的每个基矢的作用。我们要定义氢原子态矢量空间 S 上的线性算符 $A_\sigma(\sigma=1,2,3)$,选取 S 的正交归一基(2.8.10),令

$$A_\rho \mid nlm\rangle = \sqrt{-\frac{\mu\alpha^2}{2\varepsilon_n}}R_\rho \mid nlm\rangle \tag{2.8.32}$$

由上式立即可以推出,若 $H\psi=\varepsilon_n\psi$,则有

$$A_\rho\psi = \sqrt{-\frac{\mu\alpha^2}{2\varepsilon_n}}R_\rho\psi \tag{2.8.33}$$

由 $\boldsymbol{R}, \boldsymbol{l}$ 是运动积分,可知

$$HR_\rho \mid nlm\rangle = \varepsilon_n R_\rho \mid nlm\rangle \tag{2.8.34}$$

$$Hl_\rho \mid nlm\rangle = \varepsilon_n l_\rho \mid nlm\rangle \tag{2.8.35}$$

由式(2.8.22),可知

$$\boldsymbol{l}\cdot\boldsymbol{A}=0,\boldsymbol{A}\cdot\boldsymbol{l}=0 \tag{2.8.36}$$

事实上，

$$l_\sigma A_\sigma \mid nlm\rangle=\sqrt{-\frac{\mu\alpha^2}{2\varepsilon_n}}l_\sigma R_\sigma \mid nlm\rangle=0$$

$$A_\sigma l_\sigma \mid nlm\rangle=\sqrt{-\frac{\mu\alpha^2}{2\varepsilon_n}}R_\sigma l_\sigma \mid nlm\rangle=0$$

由 $R_\sigma^+=R_\sigma$ 可知式(2.8.28)成立。事实上

$$\langle n'l'm' \mid A_\sigma \mid nlm\rangle=\sqrt{-\frac{\mu\alpha^2}{2\varepsilon_n}}\langle n'l'm' \mid R_\sigma \mid nlm\rangle$$

$$=\sqrt{-\frac{\mu\alpha^2}{2\varepsilon_{n'}}}\langle nlm \mid R_\sigma \mid n'l'm'\rangle^*=\langle nlm \mid A_\sigma \mid n'l'm'\rangle^*$$

由式(2.8.23)可以推出

$$\boldsymbol{A}\cdot\boldsymbol{A}=\frac{-\mu\alpha^2}{2H}-\boldsymbol{l}^2-\hbar^2 \tag{2.8.37}$$

式中$\frac{1}{H}$是由下式定义的线性算符

$$\frac{1}{H}\mid nlm\rangle=\frac{1}{\varepsilon_n}\mid nlm\rangle \tag{2.8.38}$$

事实上，

$$\begin{aligned}A_\rho A_\rho \mid nlm\rangle&=-\frac{\mu\alpha^2}{2\varepsilon_n}R_\rho R_\rho \mid nlm\rangle=\left(\frac{-\mu\alpha^2}{2\varepsilon_n}-\boldsymbol{l}^2-\hbar^2\right)\mid nlm\rangle\\&=\left(\frac{-\mu\alpha^2}{2H}-\boldsymbol{l}^2-\hbar^2\right)\mid nlm\rangle\end{aligned} \tag{2.8.39}$$

由 R_ρ 是运动积分，可以推出式(2.8.29)。事实上

$$[A_\rho,H]\mid nlm\rangle=\sqrt{\frac{-\mu\alpha^2}{2\varepsilon_n}}[R_\rho,H]\mid nlm\rangle=0$$

而由式(2.8.25)可以推出式(2.8.30)。事实上

$$[l_\rho,A_\sigma]\mid nlm\rangle=\sqrt{-\frac{\mu\alpha^2}{2\varepsilon_n}}[l_\rho,R_\sigma]\mid nlm\rangle=i\hbar\varepsilon_{\rho\sigma\tau}\sqrt{\frac{-\mu\alpha^2}{2\varepsilon_n}}R_\tau \mid nlm\rangle$$

$$=i\hbar\varepsilon_{\rho\sigma\tau}A_\tau \mid nlm\rangle$$

而由式(2.8.6)可以推出式(2.8.31)。事实上

$$[A_\rho, A_\sigma] \mid nlm\rangle = -\frac{\mu\alpha^2}{2\varepsilon_n}[R_\rho, R_\sigma] \mid nlm\rangle = i\hbar\varepsilon_{\rho\sigma\tau} l_\tau \mid nlm\rangle$$

由上面讨论的结果可知,式(2.8.32)所定义的 S 上的三个线性算符 $A_\rho(\rho=1,2,3)$ 是三个厄米算符,并且与氢原子的哈密顿算符 $H=\frac{\boldsymbol{p}^2}{2\mu}-\frac{\alpha}{r}$ 对易。由对易关系(2.8.27)、(2.8.30)和(2.8.31)我们看到 6 个运动积分(6 个与 $H=\frac{\boldsymbol{p}^2}{2\mu}-\frac{\alpha}{r}$ 对易的厄米算符)$\{l_\rho, A_\rho \mid \rho=1,2,3\}$ 张成的 6 维线性空间,在李乘法[,]下是封闭的,构成一个 6 维李代数。我们指出这个李代数正是 6 维李群 $SO(4)$ 的李代数(在同构的意义下)。

$SO(4)$ 群是所有行列式为 1 的 4×4 实正交矩阵构成的矩阵乘法群

$$SO(4) = \{Y \in M_4(\mathbf{R}) \mid Y^t = Y^{-1}, \det Y = 1\} \tag{2.8.40}$$

式中 $M_4(\mathbf{R})$ 是全体 4×4 实矩阵所成的代数。$SO(4)$ 群的李代数 $so(4)$ 是由全体反对称实矩阵构成的。以 $E_{ab}(a,b=0,1,2,3)$ 记 a 行 b 列阵元为 1,其余阵元为零的 4×4 矩阵,则

$$\{\tilde{A}_\rho \equiv E_{0\rho} - E_{\rho 0} \quad \tilde{l}_\rho \equiv -\varepsilon_{\rho\sigma\tau}E_{\sigma\tau} \mid \rho=1,2,3\} \tag{2.8.41}$$

构成李代数 $so(4)$ 的一个基。不难证明它们满足如下的对易关系(留作练习)

$$\begin{aligned}
[\tilde{l}_\rho, \tilde{l}_\sigma] &= \varepsilon_{\rho\sigma\tau}\tilde{l}_\tau \\
[\tilde{l}_\rho, \tilde{A}_\sigma] &= \varepsilon_{\rho\sigma\tau}\tilde{A}_\tau \\
[\tilde{A}_\rho, \tilde{A}_\sigma] &= \varepsilon_{\rho\sigma\tau}\tilde{l}_\tau
\end{aligned} \tag{2.8.42}$$

按照李群表示理论(参考附录),式(2.8.42)是 $SO(4)$ 群的每一个表示的无穷小算符所满足的对易关系;反之,任意一组满足对易关系(2.8.42)的 6 个线性算符,一定是 $SO(4)$ 群的某个表示的无穷小算符。对比式(2.8.42)和对易关系(2.8.27)、(2.8.30)、(2.8.31),我们看到 $\{l_\rho, A_\rho \mid \rho=1,2,3\}$ 是 $SO(4)$ 群在氢原子态矢量空间 S 上的某个幺正表示的无穷小算符的 $i\hbar$ 倍。

让我们进一步分析 $SO(4)$ 群的结构。代替(2.8.41),我们引入李代数 $so(4)$ 的另一个基

$$\tilde{B}_\rho \equiv \frac{1}{2}(\tilde{l}\rho + \tilde{A}_\sigma), \tilde{K}_\rho \equiv \frac{1}{2}(\tilde{l}\rho - \tilde{A}_\sigma) \quad (\rho=1,2,3) \tag{2.8.43}$$

$\tilde{B}_\rho, \tilde{K}_\rho$ 是 4×4 实反对称矩阵。由上式可解得

$$\tilde{l}_\rho = \tilde{B}_\rho + \tilde{K}_\rho, \tilde{A}_\rho = \tilde{B}_\rho - \tilde{K}_\rho \quad (\rho = 1,2,3) \tag{2.8.44}$$

由对易关系(2.8.27)、(2.8.30)、(2.8.31)可得

$$[\tilde{B}_\rho, \tilde{B}_\sigma] = \varepsilon_{\rho\sigma\tau}\tilde{B}_\tau$$

$$[\tilde{K}_\rho, \tilde{K}_\sigma] = \varepsilon_{\rho\sigma\tau}\tilde{K}_\tau$$

$$[\tilde{B}_\rho, \tilde{K}_\sigma] = 0 \tag{2.8.45}$$

这表明$\{\tilde{B}_1, \tilde{B}_2, \tilde{B}_3\}$张成李代数 $so(4)$的一个理想 $so(3)$，$\{\tilde{K}_1, \tilde{K}_2, \tilde{K}_3\}$张成李代数 $so(4)$的另一个理想 $so(3)$，而整个李代数 $so(4)$是这两个理想的直和

$$so(4) = so(3)_{\mathrm{I}} \oplus so(3)_{\mathrm{II}} \tag{2.8.46}$$

相应地，李群 $SO(4)$是两个正规子群的直积

$$SO(4) = SO(3)_{\mathrm{I}} \otimes SO(3)_{\mathrm{II}} \tag{2.8.47}$$

为区分两个同构的理想 $so(3)$(正规子群 $SO(3)$)，我们加上了下标Ⅰ、Ⅱ。由群理论知，直积群 $SO(4)$的元素可以唯一的表作

$$e^{\boldsymbol{\theta}_1 \cdot \mathcal{B}} e^{\boldsymbol{\theta}_2 \cdot \mathcal{K}} \equiv (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \in SO(4) \tag{2.8.48}$$

其中

$$e^{\boldsymbol{\theta}_1 \cdot \mathcal{B}} \equiv (\boldsymbol{\theta}_1, 0) \in SO(3)_{\mathrm{I}}$$

$$e^{\boldsymbol{\theta}_2 \cdot \mathcal{K}} \equiv (0, \boldsymbol{\theta}_2) \in SO(3)_{\mathrm{II}} \tag{2.8.49}$$

$SO(4)$群中元素的乘法可以归为 $SO(3)$群中元素的乘法：

$$(\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2) * (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = (\boldsymbol{\varphi}_1 * \boldsymbol{\theta}, \boldsymbol{\varphi}_2 * \boldsymbol{\theta}_2) \tag{2.8.50}$$

式中 $\boldsymbol{\varphi} * \boldsymbol{\theta}$ 是 $SO(3)$群中两个元素 $\boldsymbol{\theta}$ 和 $\boldsymbol{\varphi}$ 的乘积。$SO(4)$群(2.8.48)中的使 $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ 的元素的全体$\{(\boldsymbol{\theta}, \boldsymbol{\theta}) | \boldsymbol{\theta} \in SO(3)\}$，构成另一个子群 $SO(3)$，它不是正规子群，记作 $SO(3)_{\mathrm{III}}$。它的元素可以写作

$$(\boldsymbol{\theta}, \boldsymbol{\theta}) = e^{\boldsymbol{\theta} \cdot (\mathcal{B} + \mathcal{K})} = e^{\boldsymbol{\theta} \cdot \mathcal{l}} \tag{2.8.51}$$

由群表示理论知，由正规子群 $SO(3)_{\mathrm{I}}$ 的全体不可约表示 $D^{[j_1]}$和正规子群 $SO(3)_{\mathrm{II}}$ 的全体不可约表示 $D^{[j_2]}$按下面方式我们可以构造出直积群 $SO(4)$的全体不可约表示 $D^{[j_1, j_2]}$：

$$D^{[j_1,j_2]}(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) = D^{[j_1]}(\boldsymbol{\theta}_1) \otimes D^{[j_2]}(\boldsymbol{\theta}_2) \tag{2.8.52}$$

不可约表示 $D^{[j_1,j_2]}$的维数为$(2j_1+1)(2j_2+1)$。

现在让我们从 $SO(4)$群的一般理论回到它在氢原子态矢量空间 S 上的以 $\left\{\frac{l\rho}{i\hbar},\frac{A\rho}{i\hbar}\mid\rho=1,2,3\right\}$为无穷小算符的幺正表示 D：

$$D(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) = e^{-\frac{i}{\hbar}\boldsymbol{\theta}_1\cdot\boldsymbol{B}}e^{-\frac{i}{\hbar}\boldsymbol{\theta}_2\cdot\boldsymbol{K}} \tag{2.8.53}$$

其中

$$\boldsymbol{B} = \frac{1}{2}(\boldsymbol{l}+\boldsymbol{A})$$

$$\boldsymbol{K} \equiv \frac{1}{2}(\boldsymbol{l}-\boldsymbol{A}) \tag{2.8.54}$$

是两个互相对易的角动量向量(六个运动积分)：

$$[B_\rho,B_\sigma] = i\hbar\varepsilon_{\rho\sigma\tau}B_\tau$$

$$[K_\rho,K_\sigma] = i\hbar\varepsilon_{\rho\sigma\tau}K_\tau$$

$$[B_\rho,K_\sigma] = 0 \tag{2.8.55}$$

D 的表示空间 S 可以分解成氢原子哈密顿算符 $H=\frac{\boldsymbol{p}^2}{2\mu}-\frac{\alpha}{r}$的两两正交的特征子空间的直和

$$S = \bigoplus_{n\geq 1} S_n \tag{2.8.56}$$

由于

$$D(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)H = HD(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) \tag{2.8.57}$$

每个能级 S_n 都是 D – 不变子空间，荷载 $SO(4)$群一个低维表示。这些低维表示是群 $SO(4)$的不可约表示吗？

由式(2.8.36)可知

$$\boldsymbol{B}^2 = \frac{1}{4}(\boldsymbol{l}^2+\boldsymbol{A}^2) = \boldsymbol{K}^2 \tag{2.8.58}$$

可见 $\boldsymbol{B},\boldsymbol{K}$ 这两个互相对易的角动量的平方相等。式(2.8.58)是一个算符恒等式。由式(2.8.37)可得

$$\frac{-\mu\alpha^2}{2H} - \hbar^2 = \boldsymbol{A}^2+\boldsymbol{l}^2 = 4\boldsymbol{B}^2 = 4\boldsymbol{K}^2 \tag{2.8.59}$$

由上式可知，线性算符$\frac{1}{H}$、$\boldsymbol{B}^2$、$\boldsymbol{K}^2$有相同的特征子空间；由式(2.3.38)可知，$\frac{1}{H}$的特征子空间就是哈密顿算符$H=\frac{\boldsymbol{p}^2}{2\mu}-\frac{\alpha}{r}$的特征子空间。把(2.8.59)作用在$|nlm\rangle$上得

$$\left(\frac{-\mu\alpha^2}{2H}-\hbar^2\right)|nlm\rangle=4b_n(b_n+1)\hbar^2|nlm\rangle=4k_n(k_n+1)\hbar^2|nlm\rangle$$

式中b_n是$\boldsymbol{B}^2$的角量子数，k_n是$\boldsymbol{K}^2$的角量子数，显然$b_n=k_n$。由上式我们可以重新得到氢原子的能量

$$\varepsilon_n=\frac{-\mu\alpha^2}{2(2b_n+1)^2\hbar^2}=\frac{-\mu\alpha^2}{2(2k_n+1)^2\hbar^2} \tag{2.8.60}$$

与式(2.8.11)对比，可知

$$n=2b_n+1=2k_n+1,\ b_n=k_n=\frac{n-1}{2} \tag{2.8.61}$$

能级ε_n的简并度为n^2，正好是$(2b_n+1)(2k_n+1)$，$SO(4)$群不可约表示$D^{[b_n,k_n]}=D^{\left[\frac{n-1}{2},\frac{n-1}{2}\right]}$的维数。可见能级$\varepsilon_n$是$\boldsymbol{B},\boldsymbol{K}$一不可约不变子空间，按$SO(4)$群的不可约表示$D^{\left[\frac{n-1}{2},\frac{n-1}{2}\right]}$变换。

由式(2.8.53)和(2.8.54)可得

$$\boldsymbol{l}=\boldsymbol{B}+\boldsymbol{K} \tag{2.8.62}$$

$$\boldsymbol{A}=\boldsymbol{B}-\boldsymbol{K} \tag{2.8.63}$$

可见$\boldsymbol{l}$是两个互相对易的角动量之和。$SO(4)$群在S上的表示(2.8.53)可以重新写作

$$D(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)=e^{-\frac{i}{\hbar}\frac{\boldsymbol{\theta}_1+\boldsymbol{\theta}_2}{2}\cdot\boldsymbol{l}-\frac{i}{\hbar}\frac{\boldsymbol{\theta}_1-\boldsymbol{\theta}_2}{2}\cdot\boldsymbol{A}} \tag{2.8.64}$$

当限制到$SO(4)$的子群$SO(3)_{\mathrm{III}}$时(见式2.8.51)，我们得到$SO(3)_{\mathrm{III}}$在S上的酉表示

$$D(\boldsymbol{\theta},\boldsymbol{\theta})=\exp\left(-\frac{i}{\hbar}\boldsymbol{\theta}\cdot\boldsymbol{l}\right) \tag{2.8.65}$$

其无穷小算符的$i\hbar$倍是氢原子的(轨道)角动量，故知$SO(4)$群的这个$SO(3)_{\mathrm{III}}$子群就是实验室绕力心O的转动群，具有纯几何变换群的性质，是$E(3)$的子群。按照§3.4两个互相对易的角动量的相加的理论，和角动量$\boldsymbol{l}=\boldsymbol{B}+\boldsymbol{K}$的角量子数$l$在能级$\varepsilon_n$上的取值为

$$l_n = | b_n - k_n |, \cdots, b_n + k_n = 0, \cdots n - 1 \tag{2.8.66}$$

同一能级有不同的角量子数反映了氢原子对称性大于角动量 $\boldsymbol{l}$ 守恒对应的 $SO(3)_{\text{Ⅲ}}$ 群对称性。*Runge-Lenz* 矢量 $\boldsymbol{R}$ 或向量 $\boldsymbol{A}$ 守恒把对称性群扩大为 $SO(4)$。氢原子的能级简并可全部归因于 $SO(4)$对称性。但是细心的读者会注意到,向量 $\boldsymbol{A}$ 对应的 S 上的酉变换

$$D(\boldsymbol{\theta}, -\boldsymbol{\theta}) = \exp\left(-\frac{i}{\hbar}\boldsymbol{\theta}\cdot\boldsymbol{A}\right) \tag{2.8.67}$$

不对应于 $E(3)$群中的任一元素 $g \in E(3)$(见(2.8.1)式),不是纯几何的对称变换。变换(2.8.67)与 H 对易是氢原子体系动力学对称性的体现。

Ⅳ. 再论量子体系的对称性群

我们看到能级简并不仅可以由体系的几何对称性引起,还可以由体系的动力学对称性引起。氢原子和类氢离子并非唯一的例子,各向同性谐振子也有系统的由动力学对称性引起的能级简并。曾谨言教授所著的《量子力学》卷Ⅱ对此有详实、精辟的阐述,本书不再介绍。对氢原子的讨论显示,我们不应局限于只考虑体系的几何对称性,因而本节开始时所给出的体系对称变换和对称性群的定义应该修改。

定义:设量子体系的态矢量空间为 S,哈密顿算符为 H,U 为 S 上的酉算符。若

$$UH = HU \quad (\text{此即 } UHU - 1 = H) \tag{2.8.68}$$

则称 U 为体系的对称变换。体系的对称变换的全体构成一个群,称之为体系的对称性群。以 $\mathbf{U}(S)$记 S 上酉变换(幺正算符)全体所成的群 $\mathbf{U}(S) \equiv \{U \in GL(S) \mid U^{+} = U^{-1}\}$。显然体系的对称性群是 $\mathbf{U}(S)$的子群,而不是 $E(3)$的子群。我们指出,按照这个关于体系对称性群的定义,"偶然简并"不复存在。

以 G 记量子体系的对称性群。由于 H 是 S 上的厄米算符,S 可以写成两两正交的 H 的特征子空间的直和:

$$S = \bigoplus_n S_n \tag{2.8.69}$$

对每一 $U \in G$,由于 $UH = HU$,S_n 是 U 的不变子空间。U 到 S_n 上的限制 $U|_{S_n}$ 是 S_n 上的幺正算符(酉变换)。我们有

$$U = \bigoplus_n U|_{S_n} \tag{2.8.70}$$

反之,若 U_n 是 S_n 上的酉算符:$U_n \in \mathbf{U}(S_n)$,则 $U \equiv \bigoplus_n U_n$ 是与 H 对易的 S 上的酉算

符：

$$U^{-1} = U^{+}, [U \equiv \bigoplus_{n} U_n, H] = 0 \tag{2.8.71}$$

事实上，注意到 H 可以写作

$$H = \bigoplus_{n} \epsilon_n 1_{S_n} \tag{2.8.72}$$

式中 1_{S_n} 是 S_n 上的恒等算符。我们有

$$UH = \left(\bigoplus_{n} U_n\right)\left(\bigoplus_{n} \epsilon_n 1_{S_n}\right) = \bigoplus_{n} \epsilon_n U_n = \left(\bigoplus_{n} \epsilon_n 1_{S_n}\right)\left(\bigoplus_{n} U_n\right) = HU \tag{2.8.73}$$

$$U^{+} = \left(\bigoplus_{n} U_n\right)^{+} = \left(\bigoplus_{n} U_n^{+}\right) = \left(\bigoplus_{n} U_n^{-1}\right) = \left(\bigoplus_{n} U_n\right)^{-1} = U^{-1} \tag{2.8.74}$$

可见体系的对称性群是 H 的各个特征子空间 S_n 的酉算符群 $\mathbf{U}(S_n)$ 的直积

$$G = \bigotimes_{n} \mathbf{U}(S_n) \tag{2.8.75}$$

当把各个 S_n 的正交归一基顺序地合并成的 S 正交归一基时，体系的任一对称变换 $U \in G$ 具有下面形式的矩阵

$$U = \begin{bmatrix} U_1 & & & \\ & U_2 & & \\ & & U_2 & \\ & & & \ddots \end{bmatrix} \tag{2.8.76}$$

其中对角块 U_n 是 $d_n \times d_n$ 酉矩阵（$d_n = \dim S_n$，即能级 ϵ_n 的简并度）。

由于 $\mathbf{U}(S_n)$ 是 S_n 上的不可约算符组，故我们有下面定理。

定理：量子体系的每个能级荷载体系对称性群的一个不可约表示。

附录

在推导力学量算符间的恒等关系时，经常要用到三维 *Civita* 符号 $\epsilon_{\rho\sigma\tau}$（$\rho, \sigma, \tau \in \{1,2,3\}$），即三维单位全反对称三阶协变张量。它是由下面两个条件确定的：

$$\epsilon_{123} = 1$$

$$\epsilon_{\rho_1\rho_2\rho_3} = (-1)^{v(p)} \epsilon_{\rho p_{(1)} \rho p_{(2)} \rho p_{(3)}}, \forall P \in S_3$$

由上式可知，交换任意两个下标时 *Civita* 符号变号。因而有重复下标时 *Civita* 符号为零，例如 $\epsilon_{212} = \epsilon_{332} = \epsilon_{111} = 0$。由上面定义我们不难验证（采用爱因斯坦求和惯例）

$$\epsilon_{\rho\beta\gamma}\epsilon_{\sigma\beta\gamma} = 2\delta_{\rho\sigma}$$

$$\varepsilon_{\rho\xi\gamma}\varepsilon_{\sigma\eta\gamma} = \delta_{\rho\sigma}\delta_{\xi\eta} - \delta_{\rho\eta}\delta_{\xi\sigma}$$

利用 *Civita* 符号,粒子的轨道角动量算符 $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$ 的分量可以表作

$$l_\rho = \varepsilon_{\rho\beta\gamma}x_\beta p_\gamma, x_\beta p_\gamma - x_\gamma p_\beta = \varepsilon_{\beta\gamma\sigma}l_\sigma$$

角动量分量间的对易关系$[l_\rho, l_\sigma] = i\hbar\varepsilon_{\rho\sigma\tau}l_\tau$,角动量分量与坐标分量间的对易关系和角动量分量与动量分量间对易关系为

$$[l_\rho, x_\sigma] = i\hbar\varepsilon_{\rho\sigma\tau}x_\tau, [l_\rho, p_\sigma] = i\hbar\varepsilon_{\rho\sigma\tau}p_\tau$$

有基本对易关系$[x_\rho, x_\sigma] = 0, [p_\rho, p_\sigma] = 0, [x_\rho, p_\sigma] = i\hbar\delta_{\rho\sigma}$,我们可以推出以下有用的算符代数关系:

$$\boldsymbol{r} \cdot \boldsymbol{l} = x_\rho l_\rho = x_\rho\varepsilon_{\rho\beta\gamma}x_\beta p_\gamma = 0(\because x_\rho x_\beta = x_\beta x_\rho, \varepsilon_{\rho\beta\gamma} = -\varepsilon_{\beta\rho\gamma})$$

$$\boldsymbol{l} \cdot \boldsymbol{r} = l_\rho x_\rho = x_\rho l_\rho + [l_\rho, x_\rho] = 0$$

$$\boldsymbol{l} \cdot \boldsymbol{p} = l_\rho p_\rho = \varepsilon_{\rho\beta\gamma}x_\beta p_\gamma p_\rho = 0$$

$$\boldsymbol{p} \cdot \boldsymbol{l} = p_\rho l_\rho = l_\rho p_\rho + [p_\rho, l_\rho] = 0$$

$$\left[p_\rho, \frac{1}{r}\right] = \frac{i\hbar x_\rho}{r^3}$$

$$\left[p_\rho, \frac{1}{r^3}\right] = 3\left[p_\rho, \frac{1}{r}\right]\frac{1}{r^2} = \frac{3i\hbar x_\rho}{r^5}$$

第 3 章　角动量理论

§3.1　角动量算符的形式定义

Ⅰ. 旋转群在态矢量空间上的酉表示的无穷小算符

考察任一量子体系，将其态矢量空间记作 S。当我们把实验室绕空间的固定点 O 作任一旋转 $\boldsymbol{\vartheta} \in SO(3)$ 时（这相当于保持实验室不动，而把观察者的参考坐标系作一相反的旋转 $-\boldsymbol{\vartheta} \in SO(3)$），量子体系的态矢量 $\psi \in S$ 要经受一个酉变换：$\psi \mapsto D(\boldsymbol{\vartheta})\psi$。由于把实验室相继地做两次转动 $\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2$，与把实验室做一次转动 $\boldsymbol{\vartheta}_2 * \boldsymbol{\vartheta}_1$，实验室所到达的最后位置是一样的，所以量子体系的态矢量在这两种情况下所经受的变换应是一样的，

$$D(\boldsymbol{\vartheta}_2)D(\boldsymbol{\vartheta}_1)\psi = D(\boldsymbol{\vartheta}_2 * \boldsymbol{\vartheta}_1)\psi, \forall\, \psi \in S$$

此即

$$D(\boldsymbol{\vartheta}_2)D(\boldsymbol{\vartheta}_1) = D(\boldsymbol{\vartheta}_2 * \boldsymbol{\vartheta}_1) \tag{3.1.1}$$

实验室不旋转时态矢量保持不变：$D(\boldsymbol{0}) = 1$。可见对应 $\boldsymbol{\vartheta} \mapsto D(\boldsymbol{\vartheta})$，$\forall\, \boldsymbol{\vartheta} \in SO(3)$ 给出旋转群 $SO(3)$ 在体系态矢量空间 S 上一个酉表示。按照李群表示理论（参阅附录和 §2.3），$SO(3)$ 群的任意一个表示 D 的无穷小算符为 $I_\alpha = \dfrac{\partial}{\partial \theta_\alpha} D(\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=0}$（$\alpha = 1, 2, 3$），它们都满足如下的对易关系

$$[I_\alpha, I_\beta] = \sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma} I_\gamma \tag{3.1.2}$$

酉表示的无穷小算符是反厄米的：$I_\alpha = -I_\alpha^+$，$(\alpha = 1,2,3)$。令 $J_\alpha = i\hbar I_\alpha = iI_\alpha(\hbar = 1)$，则 J_α 是厄米算符。态矢量空间 S 上的转动算符可以表作

$$D(\boldsymbol{\theta}) = \exp(\theta_1 I_1 + \theta_2 I_2 + \theta_3 I_3) = \exp(\boldsymbol{\theta} \cdot \boldsymbol{I}) = \exp(-i\boldsymbol{\theta} \cdot \boldsymbol{J}) \quad (3.1.3)$$

测量体系某一力学量 A 的装置，当实验室转动 $\boldsymbol{\theta}$ 时，将变作测量体系另一力学量 A' 的装置，若算符 A 把任意一个态矢量 φ 变作 ψ：$A\varphi = \psi$，则算符 A' 应把任意态矢量 $D(\boldsymbol{\theta})\varphi$ 变作 $D(\boldsymbol{\theta})\psi$：$A'D(\boldsymbol{\theta})\varphi = D(\boldsymbol{\theta})\psi = D(\boldsymbol{\theta})A\varphi$，$\forall\, \varphi \in S$。从而我们得到算符等式

$$A' = D(\boldsymbol{\theta})AD^{-1}(\boldsymbol{\theta}) \quad (3.1.4)$$

事实上，上式是态矢量空间 S 上任一算符 A 在转动下的变换方式（参阅 §2.2 和 §2.3），特别地，是 S 上所有线性算符在转动下的变换方式。考察态矢量空间 S 上所有线性算符所成的线性空间 $EndS$（参阅 §0.4）。对应 $A \mapsto D(\boldsymbol{\theta})AD^{-1}(\boldsymbol{\theta})$，$\forall\, A \in EndS$，确定了线性空间 $EndS$ 上的一个线性算符 $\widetilde{D}(\boldsymbol{\theta})$：

$$\widetilde{D}(\boldsymbol{\theta})A = D(\boldsymbol{\theta})AD^{-1}(\boldsymbol{\theta}) \quad (3.1.5)$$

利用 D 是 $SO(3)$ 群在态矢量空间 S 上的表示，可证明 $\widetilde{D}$ 是 $SO(3)$ 群在线性算符空间 $EndS$ 上的表示。事实上，对任一 $A \in EndS$，$\widetilde{D}(\boldsymbol{\theta})\mathrm{A} = \mathrm{D}(\boldsymbol{\theta})\mathrm{AD}^{-1}(\boldsymbol{\theta}) = \mathrm{A}$，$\widetilde{\mathrm{D}}(\boldsymbol{\theta})(\widetilde{D}(\boldsymbol{\varphi})A) = D(\boldsymbol{\theta})(D(\boldsymbol{\varphi})AD^{-1}(\boldsymbol{\varphi}))D^{-1}(\boldsymbol{\theta}) = D(\boldsymbol{\theta} * \boldsymbol{\varphi})AD^{-1}(\boldsymbol{\theta} * \boldsymbol{\varphi}) = \widetilde{D}(\boldsymbol{\theta} * \boldsymbol{\varphi})A$

这表明 $\widetilde{D}(\boldsymbol{\theta}) = 1$，$\widetilde{D}(\boldsymbol{\theta})\widetilde{D}(\boldsymbol{\varphi}) = \widetilde{D}(\boldsymbol{\theta} * \boldsymbol{\varphi})$，$\widetilde{D}$ 是 $SO(3)$ 群在线性算符空间 $EndS$ 上的表示。对 $A, B \in EndS$，定义它们的内积 $(A, B) = tr(AB^+)$。易证：(A, B) 线性地依赖于 A；$(A, B) = (B, A)^*$；$(A, A) \geq 0$，当且仅当 $A = 0$ 时等号成立。现在我们可以把 $EndS$ 看成内积空间。利用 D 是 $SO(3)$ 群在态矢量空间 S 上的酉表示，可证明 $\widetilde{D}$ 是 $SO(3)$ 群在线性算符空间 $EndS$ 上的酉表示。表示 $\widetilde{D}$ 的无穷小算符 $\tilde{I}_\alpha = adI_\alpha$，$\alpha = 1,2,3$。它们满足如下的对易关系

$$[\tilde{I}_\alpha, \tilde{I}_\beta] = \sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma}\tilde{I}_\gamma \quad (3.1.6)$$

无穷小算符 $\tilde{I}_\alpha = -\tilde{I}_\alpha^+$，$(\alpha = 1,2,3)$ 是反厄米的。令 $\tilde{J}_\alpha = i\tilde{I}_\alpha$，则 $\tilde{J}_\alpha$ 是厄米算符。线性算符空间 $EndS$ 上的转动算符可以表作

$$\widetilde{D}(\boldsymbol{\theta}) = \exp(\theta_1\widetilde{I}_1 + \theta_2\widetilde{I}_2 + \theta_3\widetilde{I}_3) = \exp(\boldsymbol{\theta}\cdot\widetilde{I}) = \exp(-i\boldsymbol{\theta}\cdot\widetilde{J}) \quad (3.1.7)$$

包含了体系全部动力学信息的哈密顿算符 H,在转动下变作 $H' = D(\boldsymbol{\theta})HD^{-1}(\boldsymbol{\theta})$。一般而言,$H$,$H'$不一定相等。但若体系具有关于固定点 O 的旋转不变性,即哈密顿算符在转动下不变:$D(\boldsymbol{\theta})HD^{-1}(\boldsymbol{\theta}) = H$, $\forall\,\boldsymbol{\theta}\in SO(3)$,则 $J_\alpha H = HJ_\alpha$,$(\alpha = 1,2,3)$。反之,若 $J_\alpha H = HJ_\alpha$,$(\alpha = 1,2,3)$,则体系的哈密顿算符在转动下不变。我们看到,三个厄米算符 $J_\alpha(\alpha = 1,2,3)$代表的力学量,当且仅当体系具有关于固定点 O 的旋转不变性时守恒。在 §2.3 我们还指出过,它们在转动下和笛卡尔坐标系单位向量 $\hat{e}_\alpha$,坐标算符 x_α,动量算符 p_α 按同样的方式变换,即按向量方式变换:$D(\boldsymbol{\theta})J_\alpha D^{-1}(\boldsymbol{\theta}) = R_{\beta\alpha}(\boldsymbol{\theta})J_\beta$。很自然它们应被看作体系关于固定点 O 的角动量。因此,任一量子体系的角动量算符是体系态矢量空间所荷载的 $SO(3)$群的酉表示的无穷小算符的 i 倍,它们是厄米算符,满足对易关系

$$[J_\alpha, J_\beta] = i\sum_{\gamma=x,y,z}\varepsilon_{\alpha\beta\gamma}J_\gamma \quad (3.1.8)$$

但是荷载 $SO(3)$群表示的空间不一定是量子体系的态矢量空间。例如 $SO(3)$群的表示 $\widetilde{D}$ 的荷载空间是线性算符空间 $EndS$。$\widetilde{J}_\alpha = \widetilde{iI}_\alpha(\alpha = 1,2,3)$显然不是哪个量子体系的角动量算符。然而下面我们将会看到,对易关系(3.1.8)和厄米性 $J_\alpha^+ = J_\alpha$ 几乎完全确定了角动量算符的特性。所以我们把对易关系(3.1.8) 和厄米性 $J_\alpha^+ = J_\alpha$ 当作角动量算符的形式定义,即任意满足对易关系(3.1.8)的三个厄米线性算符 $J_\alpha(\alpha = 1,2,3)$叫做角动量算符。这样做的好处在于,不仅可以把态矢量按转动性质分类,还可以把力学量算符按转动性质分类。从前面的讨论和下面的定理知,这相当于说,$SO(3)$群的酉表示的无穷小算符的 i 倍叫做角动量算符。

定理 设 V 是有限维复内积空间,V 上反厄米线性算符 I_1, I_2, I_3 满足如下的对易关系:

$$[I_\alpha, I_\beta] = i\sum_{\gamma=1,2,3}\varepsilon_{\alpha\beta\gamma}I_\gamma \quad (3.1.9)$$

令

$$T(\boldsymbol{\theta}) \equiv \exp(\boldsymbol{\theta}\cdot\boldsymbol{I}) = \exp\Big(\sum_{\alpha=1}^{3}\theta_\alpha I_\alpha\Big) \quad \forall\,\boldsymbol{\theta}\in SO(3) \quad (3.1.10)$$

则 $T: SO(3)\to GL(V)$是 $SO(3)$群在 V 上的酉表示,并以 I_1, I_2, I_3 为无穷小算符。

证明 由式(3.1.10)得 $T(\boldsymbol{\theta})^{+} \equiv \exp(-\boldsymbol{\theta}\cdot\boldsymbol{I}) = T(\boldsymbol{\theta})^{-1}, T(\boldsymbol{\theta}) = 1_V$。故只要能证明

$$T(\boldsymbol{\theta}'')T(\boldsymbol{\theta}') = T(\boldsymbol{\theta}'' * \boldsymbol{\theta}'), \forall\, \boldsymbol{\theta}'', \boldsymbol{\theta}' \in SO(3) \tag{3.1.11}$$

就可断言 T 是 $SO(3)$群在 V 上的酉表示。反厄米线性算符组 I_1, I_2, I_3 的任一不变子空间的正交补仍是 I_1, I_2, I_3 的不变子空间。所以有限维内积空间 V 可写作两两正交的 $\{I_1, I_2, I_3\}$ - 不可约不变子空间的直和。故只要能证明对任一 $\{I_1, I_2, I_3\}$ - 不可约不变子空间 W 都有

$$T(\boldsymbol{\theta}'')\mid_W T(\boldsymbol{\theta}')\mid_W = T(\boldsymbol{\theta}'' * \boldsymbol{\theta}')\mid_W, \forall\, \boldsymbol{\theta}'', \boldsymbol{\theta}' \in SO(3) \tag{3.1.12}$$

就有式(3.1.11)。

回忆 $SO(3)$群在 R^3 上的表示 $R: \boldsymbol{\theta} \mapsto R(\boldsymbol{\theta}), \forall\, \boldsymbol{\theta} \in SO(3)$。$R(\boldsymbol{\theta})$在笛卡尔基 $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\} \equiv \{\hat{i}, \hat{j}, \hat{k}\}$下的矩阵$[R(\boldsymbol{\theta})]$由等式

$$R(\boldsymbol{\theta})\hat{e}_\alpha = \sum_{\beta=1}^{3} \hat{e}_\beta R(\boldsymbol{\theta})_{\beta\alpha} \text{ 或 } [R(\boldsymbol{\theta})\hat{e}_1, R(\boldsymbol{\theta})\hat{e}_2, R(\boldsymbol{\theta})\hat{e}_3] = [\hat{e}_1, \hat{e}_2, \hat{e}_3][R(\boldsymbol{\theta})] \tag{3.1.13}$$

确定。它的无穷小算符 $\{I_1^R, I_2^R, I_3^R\}$ 满足等式

$$I_\alpha^R \hat{e}_\beta = \sum_{\gamma=1,2,3} \varepsilon_{\alpha\beta\gamma}\hat{e}_\gamma \tag{3.1.14}$$

其在笛卡尔基 $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\} \equiv \{\hat{i}, \hat{j}, \hat{k}\}$ 下的矩阵为

$$[I_\alpha^R] = \begin{bmatrix} 0 & \varepsilon_{1\alpha 2} & \varepsilon_{1\alpha 3} \\ \varepsilon_{2\alpha 1} & 0 & \varepsilon_{2\alpha 3} \\ \varepsilon_{3\alpha 1} & \varepsilon_{3\alpha 2} & 0 \end{bmatrix} \tag{3.1.15}$$

满足等式

$$[I_\alpha^R\hat{e}_1, I_\alpha^R\hat{e}_2, I_\alpha^R\hat{e}_3] = [\hat{e}_1, \hat{e}_2, \hat{e}_3][I_\alpha^R] \tag{3.1.16}$$

由算符关系式 $R(\boldsymbol{\theta}) = \exp(\sum_{\alpha=1}^{3}\theta_\alpha I_\alpha^R)$ 可得矩阵等式

$$[R(\boldsymbol{\theta})] = \exp\Big(\sum_{\alpha=1}^{3}\theta_\alpha [I_\alpha^R]\Big) \tag{3.1.17}$$

另一方面,式(3.1.9)可以重新写作(参考 §0.4)

$$adI_\alpha I_\beta = \sum_{\gamma=x,y,z} \varepsilon_{\alpha\beta\gamma} I_\gamma \tag{3.1.18}$$

写成矩阵形式此即

$$[adI_\alpha\hat{e}_1, adI_\alpha\hat{e}_2, adI_\alpha\hat{e}_3] = [\hat{e}_1, \hat{e}_2, \hat{e}_3]\begin{bmatrix} 0 & \varepsilon_{1\alpha 2} & \varepsilon_{1\alpha 3} \\ \varepsilon_{2\alpha 1} & 0 & \varepsilon_{2\alpha 3} \\ \varepsilon_{3\alpha 1} & \varepsilon_{3\alpha 2} & 0 \end{bmatrix} \equiv [\hat{e}_1, \hat{e}_2, \hat{e}_3][adI_\alpha] \tag{3.1.19}$$

可见

$$[I_\alpha^R] = \begin{bmatrix} 0 & \varepsilon_{1\alpha 2} & \varepsilon_{1\alpha 3} \\ \varepsilon_{2\alpha 1} & 0 & \varepsilon_{2\alpha 3} \\ \varepsilon_{3\alpha 1} & \varepsilon_{3\alpha 2} & 0 \end{bmatrix} = [adI_\alpha] \tag{3.1.20}$$

所以我们有

$$[\exp(\theta_\alpha adI_\alpha)] = \exp(\theta_\alpha[adI_\alpha]) = [R(\boldsymbol{\theta})] \tag{3.1.21}$$

现考察

$$T(\boldsymbol{\theta}')I_\beta T(\boldsymbol{\theta}')^{-1} = \exp(\theta'_\alpha I_\alpha)I_\beta\exp(-\theta'_\alpha I_\alpha) = \exp(\theta'_\alpha adI_\alpha)I_\beta = I_\gamma[\exp(\theta'_\alpha adI_\alpha)]_{\gamma\beta}$$

(最后一步利用了练习 3-1-1 的结果)此即

$$T(\boldsymbol{\theta}')I_\beta T(\boldsymbol{\theta}')^{-1} = I_\gamma R(\boldsymbol{\theta}')_{\gamma\beta}, \forall\, \boldsymbol{\theta}' \in SO(3) \tag{3.1.22}$$

利用此式可得

$$T(\boldsymbol{\theta}'')T(\boldsymbol{\theta}')I_\beta T(\boldsymbol{\theta}')^{-1}T(\boldsymbol{\theta}'')^{-1} = T(\boldsymbol{\theta}'')I_\gamma T(\boldsymbol{\theta}'')^{-1}R(\boldsymbol{\theta}')_{\beta\gamma} =$$

$$I_\sigma R(\boldsymbol{\theta}'')_{\sigma\gamma}R(\boldsymbol{\theta}')_{\gamma\beta} = I_\sigma R(\boldsymbol{\theta}'' * \boldsymbol{\theta}')_{\sigma\beta} = T(\boldsymbol{\theta}'' * \boldsymbol{\theta}')I_\beta T(\boldsymbol{\theta}'' * \boldsymbol{\theta}')^{-1}$$

把上式限制到任一$\{I_1, I_2, I_3\}$-不可约不变子空间 W,利用 Schur 引理 1 可得

$$T(\boldsymbol{\theta}'')|_W T(\boldsymbol{\theta}')|_W = T(\boldsymbol{\theta}'' * \boldsymbol{\theta}')|_W C_W(\boldsymbol{\theta}'', \boldsymbol{\theta}') \tag{3.1.23}$$

其中 $C_W(\boldsymbol{\theta}'', \boldsymbol{\theta}')$是连续地依赖于 $\boldsymbol{\theta}'', \boldsymbol{\theta}'$的复值函数。利用

$$\det[T(\boldsymbol{\theta})] = \det(\exp\theta_\alpha[I_\alpha]) = \exp(\theta_\alpha tr[I_\alpha]) = 1, \forall\, \hat{\theta} \in SO(3) \tag{3.1.24}$$

由上式得

$$C_W(\boldsymbol{\theta}'', \boldsymbol{\theta}')^{\dim W} = 1 \tag{3.1.25}$$

此即 $C_W(\boldsymbol{\theta}'', \boldsymbol{\theta}')$必须是 1 的 $\dim W$ 次方根。由 $T(\hat{n}\theta'')T(\hat{n}\theta') = T(\hat{n}\theta'' * \hat{n}\theta')$知 $C_W(\hat{n}\theta'', \hat{n}\theta') = 1$。因为 $C_W(\boldsymbol{\theta}'', \boldsymbol{\theta}')$不能由 1 跳到 1 的其他的 $\dim W$ 次方根,故只能是

$$C_W(\boldsymbol{\theta}'', \boldsymbol{\theta}') = 1, T(\boldsymbol{\theta}'')|_W T(\boldsymbol{\theta}')|_W = T(\boldsymbol{\theta}'' * \boldsymbol{\theta}')|_W \tag{3.1.26}$$

从而式(3.1.11)成立,$T: SO(3) \to GL(V)$是 $SO(3)$群在 V 上的酉表示。易见它以 I_1, I_2, I_3 为无穷小算符。

练习 3-1-1. 试证:设$\{e_1 \quad e_2 \quad \cdots \quad e_n\}$是 m 维线性空间 V 的一个向量组(不一定是 V 的基),A,B 是 V 上线性算符。若 $Ae_\alpha = \sum_{\beta=1}^{n} e_\beta A_{\beta\alpha}$, $Be_\alpha = \sum_{\beta=1}^{n} e_\beta B_{\beta\alpha}$, $\forall \alpha = 1,2,\cdots,n$ 或写作 $[Ae_1 \quad Ae_2 \quad \cdots \quad Ae_n] = [e_1 \quad e_2 \quad \cdots \quad e_n][A]$, $[Be_1 \quad Be_2 \quad \cdots \quad Be_n] = [e_1 \quad e_2 \quad \cdots \quad e_n][B]$ 其中 $[A]$,$[B]$ 是 $n \times n$ 矩阵,则有 $[f(A,B)e_1 \quad f(A,B)e_2 \quad \cdots \quad f(A,B)e_n] = [e_1 \quad e_2 \quad \cdots \quad e_n]f([A],[B])$,其中 f 是任一解析函数。上式的分量表达式为 $f(A,B)e_\alpha = \sum_{\beta=1}^{n} e_\beta f([A],[B])_{\beta\alpha}$。

Ⅱ. 绕固定点转动的刚体的角动量算符

作为例子我们来讨论绕固定点转动的刚体的量子化。经典力学中一个绕固定点转动的刚体的自由度为三,它的位形空间是三维空间转动群 $SO(3)$ 的群流形,也就是三维空间中的一个闭实心球,其半径为 π,但同一个直径的两个端点要看成同一个点。我们用一个矢量 $\boldsymbol{\theta}$ 标记旋转。但也往往选用欧拉角 (α,β,γ) 标记旋转。我们在本书中采用的欧拉角是按下述方式定义的:(见图 3.1-1)。

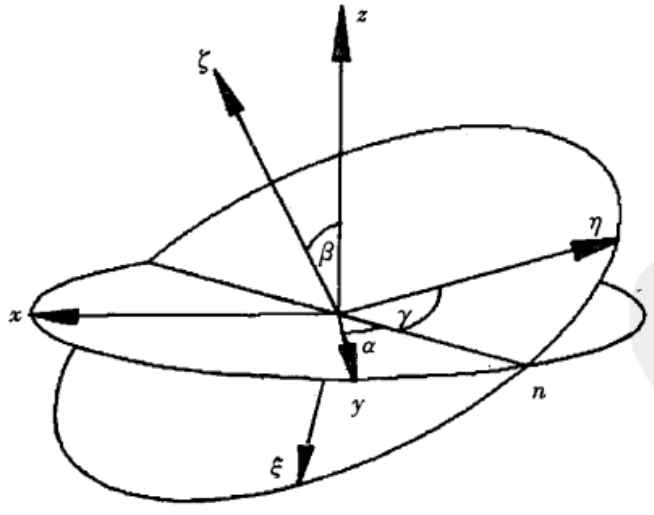

图 3.1-1

把本体坐标系统 z 轴转 γ 角$(0 \leq \gamma < 2\pi)$,再把它绕 y 轴转 β 角$(0 \leq \beta < \pi)$,最后把本体坐标系绕 z 轴转 α 角$(0 \leq \alpha < 2\pi)$。在这里我们每次旋转都是绕空间坐标系的坐标轴,采用主动观点时,这样表述欧拉角定义方便。上述定义也可以换一方式表

述:把本体坐标系绕 z 轴转 α 角,再绕新的 y 轴转 β 角,最后绕新 z 轴转 γ 角(请证明之)。在这里每一次旋转都是绕本体坐标系的坐标轴,采用被动观点时,这样表述欧拉角的定义方便。为讨论的方便,两种标记旋转的记法我们都采用。

把这样一个经典体系量子化,它的态矢量空间是由定义在这个位形空间上的连续函数 $\Psi(\boldsymbol{\theta})$或者 $\Psi(\alpha,\beta,\gamma)$所构成的。若采用$(\theta_x,\theta_y,\theta_z)\equiv\boldsymbol{\theta}$ 为广义坐标,其取值范围是$|\boldsymbol{\theta}|\leq\pi$,刚体波函数的连续性条件要求 $\Psi(\pi\hat{n})=\Psi(-\pi\hat{n})$。若采用欧拉角$(\alpha,\beta,\gamma)$为广义坐标,其取值范围是$\{-\pi\leq\alpha\leq\pi,0\leq\beta\leq\pi,-\pi\leq\gamma\leq\pi\}$,则体系的态空间由欧拉角的连续函数$\{\Psi(\alpha,\beta,\gamma)\mid\Psi(-\pi,\beta,\gamma)=\Psi(\alpha,\beta,\gamma),\Psi(\alpha,0,\gamma)=\Psi(\alpha-\Delta,0,\gamma+\Delta),\Psi(\alpha,\pi,\gamma)=\Psi(\alpha+\Delta,\pi,\gamma+\Delta),\Psi(\alpha,\beta,\gamma)=\Psi(\alpha,\beta,-\pi)\}$的全体所构成。两个态矢量的内积由下式定义

$$\langle\Phi\mid\Psi\rangle=\int_{-\pi}^{\pi}\mathrm{d}\alpha\int_{0}^{\pi}\mathrm{d}\beta\int_{-\pi}^{\pi}\mathrm{d}\gamma\sin\beta\overline{\boldsymbol{\Phi}}(\alpha,\beta,\gamma)\boldsymbol{\Psi}(\alpha,\beta,\gamma)\tag{3.1.27}$$

上式中的 $\sin\beta\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma$ 是 $SO(3)$群的群流形的不变体积元(请参见附录)。

选择刚体关于固定点 O 的惯量主轴坐标系为本体坐标系,则刚体的动能为

$$T=\sum_{\mu=\xi,\eta,\zeta}\frac{J_\mu^2}{2I_\mu}\tag{3.1.28}$$

式中 I_μ 是刚体关于 O 点的主转动惯量(不是无穷小算符 I_μ),$J_\mu=i\hbar I_\mu=iI_\mu$ 是刚体角动量在本体坐标系中的分量。

下面让我们写下角动量算符。按照前面的讨论,作用在刚体态空间上的转动算符 $D(\boldsymbol{\theta})$是这样确定的:把实验室作一转动 $\boldsymbol{\theta}$,本体坐标系的位置由位形空间中的点 $\boldsymbol{\varphi}\in SO(3)$变到点 $\boldsymbol{\theta}*\boldsymbol{\varphi}\in SO(3)$。转动后的波函数在 $\boldsymbol{\theta}*\boldsymbol{\varphi}$ 点的值应等于原来波函数在 $\boldsymbol{\varphi}$ 点的值,从而:

$$(D(\boldsymbol{\theta})\Psi)(\boldsymbol{\varphi})=\Psi((-\boldsymbol{\theta})*\boldsymbol{\varphi})\tag{3.1.29}$$

D 是$SO(3)$群在刚体态矢量空间上的酉表示,它的无穷小算符的 i 倍是刚体的角动量算符:$J_\alpha=iI_\alpha,(\alpha=x,y,z)$。角动量在某一方向 $\hat{a}$ 上的正投影 $\hat{a}\cdot\boldsymbol{J}$ 是绕轴 $\hat{a}$ 转单位无穷小角的转动算符的 i 倍。若采用欧拉角作广义坐标,把实验室绕 z 轴转 φ 角,本体坐标系的欧拉角从(α,β,γ)变到$(\alpha+\varphi,\beta,\gamma)$,则对应的转动算符应由下式确定:

$$D(\varphi\hat{k})\Psi(\alpha,\beta,\gamma)=\Psi(\alpha-\varphi,\beta,\gamma)=\exp\left(-\varphi\frac{\partial}{\partial\alpha}\right)\Psi(\alpha,\beta,\gamma)\tag{3.1.30}$$

于是角动量在 z 方向的正投影为

$$J_z = iI_z = i\frac{\partial}{\partial\varphi}D(\varphi\hat{k})\mid_{\varphi=0} = -i\frac{\partial}{\partial\alpha},(\hbar = 1) \tag{3.1.31}$$

把实验室绕 ζ 轴转 φ 角,本体坐标系的欧拉角从(α,β,γ)变到$(\alpha,\beta,\gamma+\varphi)$。对应的转动算符应由下式确定:

$$D(\varphi\hat{\zeta})\Psi(\alpha,\beta,\gamma) = \Psi(\alpha,\beta,\gamma-\varphi) = \exp\left(-\varphi\frac{\partial}{\partial\gamma}\right)\Psi(\alpha,\beta,\gamma) \tag{3.1.32}$$

于是角动量在 ζ 方向的正投影为

$$J_\zeta = iI_\zeta = i\frac{\partial}{\partial\varphi}D(\varphi\hat{\zeta})\mid_{\varphi=0} = -i\frac{\partial}{\partial\gamma},(\hbar = 1) \tag{3.1.33}$$

同理,把实验室绕 $\hat{n}$ 轴转 φ 角,本体坐标系的欧拉角从(α,β,γ)变到$(\alpha,\beta+\varphi,\gamma)$,则对应的转动算符应由下式确定:

$$D(\varphi\hat{n})\Psi(\alpha,\beta,\gamma) = \Psi(\alpha,\beta-\varphi,\gamma) = \exp\left(-\varphi\frac{\partial}{\partial\beta}\right)\Psi(\alpha,\beta,\gamma)$$

于是角动量在 $\hat{n}$ 方向的正投影为

$$J_n = iI_n = i\frac{\partial}{\partial\varphi}D(\varphi\hat{n})\mid_{\varphi=0} = i\frac{\partial}{\partial\varphi}\exp\left(-\varphi\frac{\partial}{\partial\beta}\right) = -i\frac{\partial}{\partial\beta} \tag{3.1.34}$$

由几何关系 $\hat{\zeta} = \sin\beta\cos\alpha\hat{i} + \sin\beta\cos\alpha\hat{j} + \cos\beta\hat{k}$, $\hat{n} = -\sin\alpha\hat{i} + \cos\alpha\hat{j}$ 可知 $J_\zeta = \sin\beta\cos\alpha J_x + \sin\beta\sin\alpha J_y + \cos\alpha J_z$, $J_n = -\sin\alpha J_x + \cos\alpha J_y$。由此求出

$$\begin{aligned} J_x &= i\cos\alpha\,\mathrm{ctg}\beta\frac{\partial}{\partial\alpha} + i\sin\alpha\frac{\partial}{\partial\beta} - i\frac{\cos\alpha}{\sin\beta}\frac{\partial}{\partial\gamma} \\ J_y &= i\sin\alpha\,\mathrm{ctg}\beta\frac{\partial}{\partial\alpha} - i\cos\alpha\frac{\partial}{\partial\beta} - i\frac{\sin\alpha}{\sin\beta}\frac{\partial}{\partial\gamma} \end{aligned} \tag{3.1.35}$$

由几何关系 $\hat{n} = \sin\gamma\hat{\xi} + \cos\gamma\hat{\eta}$, $\hat{k} = -\sin\beta\cos\gamma\hat{\xi} + \sin\beta\sin\gamma\hat{\eta} + \cos\beta\hat{\zeta}$,可知 $J_n = \sin\gamma J_\xi + \cos\gamma J_\eta$, $J_z = -\sin\beta\cos\gamma J_\xi + \sin\beta\sin\gamma J_\eta + \cos\beta J_\zeta$ 由此求出

$$\begin{aligned} J_\xi &= i\frac{\cos\gamma}{\sin\beta}\frac{\partial}{\partial\alpha} - i\sin\gamma\frac{\partial}{\partial\beta} - i\mathrm{ctg}\beta\cos\gamma\frac{\partial}{\partial\gamma} \\ J_\eta &= -i\frac{\sin\gamma}{\sin\beta}\frac{\partial}{\partial\alpha} - i\cos\gamma\frac{\partial}{\partial\beta} + i\mathrm{ctg}\beta\sin\gamma\frac{\partial}{\partial\gamma} \end{aligned} \tag{3.1.36}$$

直接验算可得

$$[J_x,J_y] = iJ_z,[J_y,J_z] = iJ_x,[J_z,J_x] = iJ_y \tag{3.1.37}$$

这是我们预期的。但直接验算还给出

$$[J_\xi, J_\eta] = -iJ_\zeta, [J_\eta, J_\zeta] = -iJ_\xi, [J_\zeta, J_\xi] = -iJ_\eta \tag{3.1.38}$$

出乎意料,这不是通常的角动量算符的对易关系。让我们回忆角动量算符之所以满足对易关系(3.1.19),是因为他们是 $SO(3)$群表示(关于群参数 $\theta_x, \theta_y, \theta_z$)的无穷小算符的 i 倍。而不允许选取 $\boldsymbol{\theta}$ 在本体坐标系(非固定坐标系)中的分量为 $SO(3)$群的参数。因而 J_ξ, J_η, J_ζ 不是$SO(3)$群表示的无穷小算符的 i 倍,不满足通常的角动量算符的对易关系就没有什么可奇怪的了。下面的两个事实将会帮助我们理解角动量向量在固定的笛卡尔坐标系中的分量,与角动量向量在随刚体一起运动的笛卡尔坐标系中的分量为什么满足不同的对易关系。

(ⅰ)设 $\boldsymbol{a}, \boldsymbol{b}$ 是三维空间 R^3 中两个固定的向量。考虑对易关系

$$[\boldsymbol{a}\cdot\boldsymbol{J}, \boldsymbol{b}\cdot\boldsymbol{J}] = a_\alpha b_\mu [J_\alpha, J_\mu] = a_\alpha b_\mu i\varepsilon_{\alpha\mu\sigma} J_\sigma = i(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{J} \tag{3.1.39}$$

由这个公式可得到通常的角动量对易关系(3.1.6)。

(ⅱ)设 $\boldsymbol{a}, \boldsymbol{b}$ 是三维空间 R^3 中两个固定的向量。故 $R(\boldsymbol{\theta})\boldsymbol{a}, R(\boldsymbol{\theta})\boldsymbol{b}$ 是随刚体一起运动的两个向量。考虑对易关系

$$\begin{aligned}
[(R(\boldsymbol{\theta})\boldsymbol{a})\cdot\boldsymbol{J}, (R(\boldsymbol{\theta})\boldsymbol{b})\cdot\boldsymbol{J}] &= [R_{\alpha\beta}(\boldsymbol{\theta})a_\beta J_\alpha, R_{\mu\nu}(\boldsymbol{\theta})b_\nu J_\mu] \\
&= a_\beta b_\nu [R_{\alpha\beta}(\boldsymbol{\theta})J_\alpha, R_{\mu\nu}(\boldsymbol{\theta})J_\mu] \\
&= a_\beta b_\nu R_{\mu\nu}(\boldsymbol{\theta})[R_{\alpha\beta}(\boldsymbol{\theta}), J_\mu]J_\alpha + a_\beta b_\nu R_{\alpha\beta}(\boldsymbol{\theta})[J_\alpha, R_{\mu\nu}(\boldsymbol{\theta})]J_\mu \\
&\quad + a_\beta b_\nu R_{\alpha\beta}(\boldsymbol{\theta})R_{\mu\nu}(\boldsymbol{\theta})[J_\alpha, J_\mu]
\end{aligned}$$

注意到$[J_\lambda, R_{\mu\nu}(\boldsymbol{\theta})] = -i\varepsilon_{\lambda\sigma\mu}R_{\sigma\nu}(\boldsymbol{\theta})$, $\forall\, \nu = 1,2,3$(参阅注),上式等于

$$\begin{aligned}
&i\varepsilon_{\mu\alpha\alpha}R_{\sigma\nu}(\boldsymbol{\theta})b_\nu R_{\sigma\beta}(\boldsymbol{\theta})a_\beta J_\alpha - i\varepsilon_{\alpha\sigma\mu}R_{\alpha\beta}(\boldsymbol{\theta})a_\beta R_{\sigma\nu}(\boldsymbol{\theta})b_\nu J_\mu + i\varepsilon_{\alpha\mu\sigma}R_{\alpha\beta}(\boldsymbol{\theta})a_\beta R_{\mu\nu}(\boldsymbol{\theta})b_\nu J_\sigma \\
&= i(R(\boldsymbol{\theta})\boldsymbol{b}\times R(\boldsymbol{\theta})\boldsymbol{a})\cdot\boldsymbol{J} - i(R(\boldsymbol{\theta})\boldsymbol{a}\times R(\boldsymbol{\theta})\boldsymbol{b})\cdot\boldsymbol{J} + i(R(\boldsymbol{\theta})\boldsymbol{a}\times R(\boldsymbol{\theta})\boldsymbol{b})\cdot\boldsymbol{J} \\
&= -i(R(\boldsymbol{\theta})(\boldsymbol{a}\times\boldsymbol{b}))\cdot\boldsymbol{J}
\end{aligned}$$

于是,

$$[(R(\boldsymbol{\theta})\boldsymbol{a})\cdot\boldsymbol{J}, (R(\boldsymbol{\theta})\boldsymbol{b})\cdot\boldsymbol{J}] = -i(R(\boldsymbol{\theta})(\boldsymbol{a}\times\boldsymbol{b}))\cdot\boldsymbol{J} \tag{3.1.40}$$

由这个公式可得到对易关系(3.1.38)。从推导过程看出 $\boldsymbol{a}\cdot\vec{\boldsymbol{J}} = a_\alpha J_\alpha$ 向 J_x, J_y, J_z 的展开系数 a_α 是常数,与刚体的角动量算符 $\boldsymbol{J}$ 对易,而$(R(\boldsymbol{\theta})\boldsymbol{a})\cdot\boldsymbol{J} = R_{\alpha\beta}(\boldsymbol{\theta})a_\beta J_\alpha$ 向 J_x, J_y, J_z 的展开系数$R_{\alpha\beta}(\boldsymbol{\theta})a_\beta$ 是欧拉角的函数,与刚体的角动量算符 $\boldsymbol{J}$ 不对易。这是导致不同对易关系原因。

由对易关系(3.1.6)可知,刚体角动量平方 $J^2 \equiv J_x^2 + J_y^2 + J_z^2$ 与 J_x, J_y, J_z 对易。

由(3.1.31)和(3.1.35)直接验算给出

$$J^2 = -\frac{1}{\sin^2\beta}\frac{\partial^2}{\partial\alpha^2} + \frac{2\cos\beta}{\sin^2\beta}\frac{\partial^2}{\partial\alpha\partial\gamma} - \frac{1}{\sin^2\beta}\frac{\partial^2}{\partial\gamma^2} - \frac{1}{\sin\beta}\frac{\partial}{\partial\beta}\left(\sin\beta\frac{\partial}{\partial\beta}\right) \quad (3.1.41)$$

由(3.1.33)和(3.1.36)直接验算给出,式(3.1.41)中的 $J^2 = J_\xi^2 + J_\eta^2 + J_\zeta^2$,这是我们所预期的。由对易关系(3.1.38)可知,它与 J_ξ, J_η, J_ζ 对易。

注 设 $J_\lambda(\lambda = x, y, z)$是绕固定点 O 运动的刚体的关于 O 点的角动量算符,考察

$$\begin{aligned}(J_\lambda R_{\mu\nu}(\boldsymbol{\theta}))\Psi(\boldsymbol{\theta}) &= J_\lambda(R_{\mu\nu}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta})) = i\frac{\partial}{\partial\varphi_\lambda}D(\boldsymbol{\varphi})\mid_{\boldsymbol{\varphi}=0}(R_{\mu\nu}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta}))\\ &= i\frac{\partial}{\partial\varphi_\lambda}(R_{\mu\nu}((-\boldsymbol{\varphi})*\boldsymbol{\theta})\Psi(-\boldsymbol{\varphi})*\boldsymbol{\theta}))\mid_{\boldsymbol{\varphi}=0}\\ &= i\frac{\partial}{\partial\varphi_\lambda}R_{\mu\nu}((-\boldsymbol{\varphi})*\boldsymbol{\theta})\mid_{\boldsymbol{\varphi}=0}\Psi(\boldsymbol{\theta}) + R_{\mu\nu}(\boldsymbol{\theta})i\frac{\partial}{\partial\varphi_\lambda}\Psi((-\boldsymbol{\varphi}*\boldsymbol{\theta})\mid_{\boldsymbol{\varphi}=0}\\ &= i\frac{\partial}{\partial\varphi_\lambda}R_{\mu\sigma}(-\boldsymbol{\varphi})\mid_{\boldsymbol{\varphi}=0}R_{\sigma\nu}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta}) + R_{\mu\nu}(\boldsymbol{\theta})i\frac{\partial}{\partial\varphi_\lambda}D(\boldsymbol{\varphi})\mid_{\boldsymbol{\varphi}=0}\Psi(\boldsymbol{\theta})\end{aligned}$$

(利用 $R(\boldsymbol{\theta})$的正交性和 (2.3.3)的结果$\frac{\partial}{\partial\varphi_\lambda}R_{\sigma\mu}(\boldsymbol{\varphi})\mid_{\boldsymbol{\varphi}=0} = -\varepsilon_{\lambda\sigma\mu}$)

$$= i\frac{\partial}{\partial\varphi_\lambda}R_{\sigma\mu}(\boldsymbol{\varphi})\mid_{\boldsymbol{\varphi}=0}R_{\sigma\nu}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta}) + R_{\mu\nu}(\boldsymbol{\theta})J_\lambda\Psi(\boldsymbol{\theta}) = -i\varepsilon_{\lambda\sigma\mu}R_{\sigma\nu}(\boldsymbol{\theta})\Psi(\boldsymbol{\theta}) + R_{\mu\nu}(\boldsymbol{\theta})J_\lambda\Psi(\boldsymbol{\theta})$$

于是

$$[J_\lambda, R_{\mu\nu}(\boldsymbol{\theta})] = -i\varepsilon_{\lambda\sigma\mu}R_{\sigma\nu}(\boldsymbol{\theta}), \forall\upsilon = 1,2,3 \quad (3.1.42)$$

§3.2 单个角动量的理论

本节的内容是整个角动量理论的基础。我们不打算在此重复读者在大学本科量子力学课程中学过的角动量理论,而是利用群表示论的知识,从一个新的视角探讨这一问题。

当把实验室绕固定点 O 作一旋转 $\boldsymbol{\theta}\forall \in SO(3)$时,量子体系的任一态矢量 $\psi \in S$(S 是体系的态矢量空间)要经受一个酉变换 $\psi \mapsto D(\boldsymbol{\theta})\psi$,量子体系的任一力学量算符 $A \in EndS$($EndS$ 是 S 上线性算符的全体所成的线性空间)也要经受一个酉变换 A

$\vdash\!\!\rightarrow D(\boldsymbol{\theta})AD(\boldsymbol{\theta})^{-1}\equiv\widetilde{D}(\boldsymbol{\theta})A$。$D(\widetilde{D})$是 $SO(3)$群在态矢量空间 S(线性算符空间 $EndS$)上的一个酉表示。$S(EndS)$按酉表示 $D(\widetilde{D})$分解成两两正交的不可约不变子空间的直和。因此,我们要研究量子体系的态矢量(力学量算符)在转动下的变换性质,只要研究各个不可约不变子空间在转动下如何变换。

Ⅰ. J – 不可约不变子空间的转动性质

群表示论的抽象性给我们的学习带来了困难,但正是这种抽象性使我们现在可以统一地观察态矢量和力学量算符的转动性质。设 T 是 $SO(3)$群在复 *Hilbert* 空间 V 上的一个酉表示(V 可以是态矢量空间 S,也可以是线性算符空间 $EndS$,还可以是其它复 *Hilbert* 空间)。它的无穷小算符 $I_\alpha=\left.\dfrac{\partial T(\boldsymbol{\theta})}{\partial\theta_\alpha}\right|_{\boldsymbol{\theta}=0}$($\alpha=1,2,3$ 或 $\alpha=x,y,z$)的 i 倍($i\hbar$ 倍,取 $\hbar=1$ 的单位制)$J_\alpha=iI_\alpha$ 是三个厄米算符,满足如下的对易关系

$$[J_\alpha,J_\beta]=i\hbar\sum_{r=1}^{3}\varepsilon_{\alpha\beta\gamma}J_r=i\sum_{r=1}^{3}\varepsilon_{\alpha\beta\gamma}J_r \tag{3.2.1}$$

因而 J_x,J_y,J_z 是角动量算符。由 §2.3 知,有限转动的算符为

$$T(\boldsymbol{\theta})=\exp(-i\boldsymbol{\theta}\cdot\boldsymbol{J}) \tag{3.2.2}$$

V 按照酉表示 T 分解成两两正交的不可约不变子空间的直和(事实上,若 W 是 V 的 $T(\boldsymbol{\theta})$ – 不变子空间,则 W 的正交补空间 $W^\perp$ 是 $T(\boldsymbol{\theta})^+=T(\boldsymbol{\theta})^{-1}=T(-\boldsymbol{\theta})$的不变子空间,当 $\boldsymbol{\theta}$ 取遍 $SO(3)$群,$-\boldsymbol{\theta}$ 也取遍 $SO(3)$群,所以酉表示 T 的不变子空间的正交补空间仍是 T – 不变子空间,表示空间 V 一定可以分解成两两正交的 T – 不可约不变子空间的直和。参考 §0.3 和附录)

$$V=W\oplus W'\oplus\cdots \tag{3.2.3}$$

让我们考察上式右边的任意一个不可约不变子空间 W 的转动性质,设 $\dim W<\infty$。

引入角动量平方算符

$$J^2=J_x^2+J_y^2+J_z^2 \tag{3.2.4}$$

和上升、下降算符

$$J_\pm=J_x\pm iJ_y \tag{3.2.5}$$

厄米算符的平方是半正定的厄米算符,半正定的厄米算符之和仍是半正定的厄米算符。故 J^2 是半正定的厄米算符,它在任意量子态中的期待值是非负实数,特别是,它

的所有本征值是非负实数。$J_\pm$ 不是厄米算符，它们互为厄米共轭

$$(J_\pm)^+ = J_\mp \tag{3.2.6}$$

由对易关系(3.2.1)可得

$$[J^2, J_\alpha] = 0 \quad (\alpha = 1,2,3), \quad [J^2, J_\pm] = 0 \tag{3.2.7}$$

$$[J_z, J_\pm] = \pm J_\pm, \text{此即 } J_z J_\pm = J_\pm (J_z \pm 1) \tag{3.2.8}$$

和

$$J_+ J_- = J^2 - J_z^2 + J_z, J_- J_+ = J^2 - J_z^2 - J_z, [J_+, J_-] = 2J_z \tag{3.2.9}$$

由于 W 是算符组 $\{T(\boldsymbol{\theta}) \mid \boldsymbol{\theta} \in SO(3)\}$ 的不可约不变子空间，W 一定是 $\{J_\alpha \mid \alpha = x, y, z\}$ 或 $\{J_+, J_-, J_z\}$ 的不可约不变子空间。我们可以把 $\{J_\alpha \mid \alpha = x, y, z\}$ 看成是 W 上的线性算符组，而且是 W 上的不可约线性算符组。把 $J^2, J_\pm$ 也看作 W 上的线性算符。由(3.2.7)和 Schur 引理 1(V，从而 W，是复线性空间，引理成立)可知

$$J^2 \mid_W = \lambda 1_W \tag{3.2.10}$$

式中 $J^2 \mid_W$ 是算符 J^2 到不变子空间 W 上的限制，1_W 是 W 上的恒等算符。式(3.2.10)表明 W 中的每一个非零向量都是 J^2 的属于特征值 λ 的特征向量。λ 是非负实数。$J_z \mid_W$ 是 W 上的厄米算符，以 μ 记它的最大的本征值。这个本征值是否简并现在暂时还不知道。以 $\mid \lambda\mu\rangle'$ 记 $J_z \mid_W$ 的某个属于特征值 μ 的归一化的特征向量。显然有

$$\begin{aligned} J^2 \mid \lambda\mu\rangle' &= \lambda \mid \lambda\mu\rangle' \\ J_z \mid \lambda\mu\rangle' &= \mu \mid \lambda\mu\rangle' \end{aligned} \tag{3.2.11}$$

由式(3.2.8)可得

$$J_z J_+ \mid \lambda\mu\rangle' = (\mu + 1) J_+ \mid \lambda\mu\rangle'$$

由于 $J_z \mid_W$ 的最大本征值是 μ，$\mu + 1$ 不可能是 $J_z \mid_W$ 的特征值，上式意味着

$$J_+ \mid \lambda\mu\rangle' = 0 \tag{3.2.12}$$

以 J_- 作用于上式并利用式(3.2.9)得

$$J_- J_+ \mid \lambda\mu\rangle' = (\lambda - \mu(\mu + 1)) \mid \lambda\mu\rangle' = 0$$

于是

$$\lambda = \mu(\mu + 1) \tag{3.2.13}$$

利用式(3.2.8)可得

$$J_z J_- \mid \lambda\mu\rangle' = (\mu - 1) J_- \mid \lambda\mu\rangle'$$

$$J_z J_-^n |\lambda\mu\rangle' = (\mu - n) J_-^n |\lambda\mu\rangle', \forall\, n \geq 1 \tag{3.2.14}$$

这表明若 $J_-^n |\lambda\mu\rangle' \neq 0$,则 $J_- |\lambda\mu\rangle', J_-^2 |\lambda\mu\rangle', \cdots, J_-^n |\lambda\mu\rangle'$ 是 J_z 的特征向量,分别属于特征值 $\mu-1, \mu-2, \cdots, \mu-n$。因为它们属于不变子空间 W,故它们还是 J^2 的属于特征值 λ 的特征向量。由于 $\dim W < \infty$ 是有限的,$J_z|_W$ 不可能有无穷多个特征值,故一定存在某个自然数 N,使

$$J_-^{N-1} |\lambda\mu\rangle' \neq 0, \text{而 } J_-^N |\lambda\mu\rangle' = 0 \tag{3.2.15}$$

以 J_+ 作用于最后一个等式的两边得

$$\begin{aligned} J_+ J_- J_-^{N-1} |\lambda\mu\rangle' &= (J^2 - J_z^2 + J_z) J_-^{N-1} |\lambda\mu\rangle' \\ &= [\lambda - (\mu - N + 1)(\mu - N)] J_-^{N-1} |\lambda\mu\rangle' = 0 \end{aligned}$$

从而有

$$\lambda = (\mu - N + 1)(\mu - N) \tag{3.2.16}$$

由方程(3.2.13)和(3.2.16)得

$$\mu = \frac{N-1}{2} \tag{3.2.17}$$

可见 μ 只能是非负的整数或半整数 $0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$。另一方面,由于 $J_-^{N-1} |\lambda\mu\rangle' \neq 0$,向量

$$|\lambda\mu\rangle', J_- |\lambda\mu\rangle', J_-^2 |\lambda\mu\rangle', \cdots, J_-^{N-1} |\lambda\mu\rangle', \tag{3.2.18}$$

是 $J^2|_W$ 和 $J_z|_W$ 共同的特征向量。由于它们属于 $J_z|_W$ 的不同特征值,故互相正交。让我们来考察它们所张成的 W 的 N 维子空间 W_1(即它们的线性组合的全体)。J_z、$J_\pm$ 对 W_1 的各生成向量的作用为

$$\left.\begin{aligned} &J_z J_-^f |\lambda\mu\rangle' = (\mu - f) J_-^f |\lambda\mu\rangle' \qquad (f = 0, 1, \cdots, N-1) \\ &J_- J_-^f |\lambda\mu\rangle' = J_-^{f+1} |\lambda\mu\rangle' \qquad (f = 0, 1, \cdots, N-2,), J_- J_-^{N-1} |\lambda\mu\rangle' = 0 \\ &J_+ |\lambda\mu\rangle' = 0, \qquad \textit{and for } f = 1, \cdots, N-1 \\ &J_+ J_-^f |\lambda\mu\rangle' = (J^2 - J_z^2 + J_z) J_-^{f-1} |\lambda\mu\rangle' = [\lambda - (\mu - f + 1)(\mu - f)] J_-^{f-1} |\lambda\mu\rangle \end{aligned}\right\} \tag{3.2.19}$$

这表明 W_1 是 $\{J_z, J_+, J_-\}$ 的不变子空间。由于 W 是 $\boldsymbol{J}$ – 不可约不变子空间,可知 $W_1 = W, \dim W = N$。

向量组(3.2.18)构成了 W 的一个正交基(不一定是正交归一基)。考虑到 $\lambda = \mu$

$(\mu+1)$，$\mu=\dfrac{N-1}{2}$和 $\dim W=N$，从式(3.2.19)可以看出 $J_z|_W$ 和 $J_\pm|_W$ 在此正交归一基下的矩阵完全由 $\dim W$ 确定。从而角动量算符 $\boldsymbol{J}|_W$ 和转动算符 $T(\boldsymbol{\theta})|_W=\exp(-i\boldsymbol{\theta}\cdot\boldsymbol{J}|_W)$在此正交基下的矩阵完全由 $\dim W$ 确定。用群表示论的语言说，$SO(3)$群的有限维不可约表示(在等价的意义上)完全由表示的维数确定。

我们看到 $J^2|_W$ 的本征值为 $\lambda=\dfrac{\dim W-1}{2}\left(\dfrac{\dim W-1}{2}+1\right)\hbar^2$，其中$\dfrac{\dim W-1}{2}$是非负整数(当 $\dim W$ 是奇数)或非负的半整数(当 $\dim W$ 是偶数)。这使我们可以引入角量子数 $j\equiv\dfrac{\dim W-1}{2}\in\left\{0,\dfrac{1}{2},1,\dfrac{3}{2},\cdots\right\}$来刻画 $J^2|_W$ 的本征值 $\lambda=j(j+1)\hbar^2$。$J_z|_W$ 有 $\dim W=2j+1$ 个不同的本征值$\{j,j-1,\cdots,-j\}$。引入磁量子数 $m\in\{j,j-1,\cdots,-j\}$用来刻画 $J_z|_W$ 的本征值 $m\hbar$。代替用 $J^2|_W$ 和 $J_z|_W$ 的本征值，我们将用角量子数 j 和磁量子数 m 来标记 $J^2|_W$ 和 $J_z|_W$ 的共同的特征向量(3.2.18)：

$$\{|jm\rangle'\equiv J_-^{j-m}|\lambda\mu\rangle'\mid m=j,j-1,\cdots,-j\} \tag{3.2.20}$$

于是

$$\begin{aligned}J^2|jm\rangle'&=j(j+1)|jm\rangle'\\ J_z|jm\rangle'&=m|jm\rangle'\end{aligned} \tag{3.2.21}$$

由式(3.2.19)可得

$$J_-|jm\rangle'=|jm-1\rangle',\forall m=j,j-1,\cdots,-j+1,J_-|j-j\rangle=0$$
$$J_+|jj\rangle'=0,J_+|jm\rangle'=(j-m)(j+m+1)|jm+1\rangle',\forall m=j-1,j-2,\cdots,-j \tag{3.2.22}$$

利用算符恒等式(作为练习，请读者证明之)

$$J_+^{\alpha}J_-^{\alpha}=\prod_{\sigma=1}^{\alpha}[J^2-(J_Z-\sigma)(J_Z-\sigma+1)] \tag{3.2.23}$$

可以证明

$$'\langle jm|jm\rangle'='\langle jj|J_+^{j-m}J_-^{j-m}|jj\rangle'=\prod_{\sigma=1}^{j-m}[j(j+1)-(j-\sigma)(j-\sigma+1)] \tag{3.2.24}$$

把向量组(3.2.20)归一化时，每一个向量仍有一任意相因子不能确定，必须人为地规定。我们按照 Condon-Shortley 约定取归一化因子：

$$\left\{\mid jm\rangle \equiv \left(\prod_{\sigma=1}^{j-m}[j(j+1)-(j-\sigma)(j-\sigma+1)]^{-\frac{1}{2}}\right)\mid jm\rangle' \mid m=j,j-1,\cdots,-j\right\} \tag{3.2.25}$$

向量组(3.2.25)是 $\boldsymbol{J}$ – 不可约不变子空间 W 的正交归一基。由(3.2.21)得

$$J^2\mid jm\rangle = j(j+1)\mid jm\rangle, \forall\, m=j,j-1,\cdots,-j \tag{3.2.26}$$

$$J_z\mid jm\rangle = m\mid jm\rangle, \forall\, m=j,j-1,\cdots,-j \tag{3.2.27}$$

而由(3.2.22)得

$$J_-\mid jm\rangle = \sqrt{(j+m)(j-m+1)}\mid jm-1\rangle, \forall\, m=j,j-1,\cdots,-j \tag{3.2.28}$$

$$J_+\mid jm\rangle = \sqrt{(j-m)(j+m+1)}\mid jm+1\rangle, \forall\, m=j,j-1,\cdots,-j \tag{3.2.29}$$

于是在正交归一基(3.2.25)下，$J^2\mid_W, J_z\mid_W, J_\pm\mid_W$，的矩阵为

$$[J^2\mid_W] = \begin{bmatrix} j(j+1) & 0 & \cdots & 0 \\ 0 & j(j+1) & \cdots & 0 \\ & \cdots & & \\ 0 & 0 & \cdots & j(j+1) \end{bmatrix} = [J^2]_j \tag{3.2.30}$$

$$[J_z\mid_W] = \begin{bmatrix} j & 0 & \cdots & 0 \\ 0 & j-1 & \cdots & 0 \\ & \cdots & & \\ 0 & 0 & \cdots & -j \end{bmatrix} \equiv [J_z]_j \tag{3.2.31}$$

$$[J_+|_W]=\begin{bmatrix}0&\sqrt{1(2j)}&0&0&\cdots&0&0&0\\0&0&\sqrt{2(2j-1)}&0&\cdots&0&0&0\\0&0&0&\sqrt{3(2j-2)}&\cdots&0&0&0\\&&&\cdots&&&&\\0&0&0&0&\cdots&0&\sqrt{(2j-1)2}&0\\0&0&0&0&\cdots&0&0&\sqrt{(2j)1}\\0&0&0&0&\cdots&0&0&0\end{bmatrix}$$

$$\equiv [J_+]_j \tag{3.2.32}$$

$$[J_-|_W]=\begin{bmatrix}0&0&0&\cdots&0&0&0\\\sqrt{1(2j)}&0&0&\cdots&0&0&0\\0&\sqrt{2(2j-1)}&0&\cdots&0&0&0\\0&0&\sqrt{3(2j-2)}&\cdots&0&0&0\\&&\cdots&&&&\\0&0&0&\cdots&0&0&0\\0&0&0&\cdots&\sqrt{(2j-1)2}&0&0\\0&0&0&\cdots&0&\sqrt{(2j)1}&0\end{bmatrix}$$

$$\equiv [J_-]_j \tag{3.2.33}$$

它们都是$(2j+1)\times(2j+1)$矩阵。

由于$\{J_z|_W, J_+|_W, J_-|_W\}$是不可约算符组，式(3.2.27)～(3.2.29)或矩阵形式(3.2.31)～(3.2.33)把 W 的基确定到一个公共的非零复数倍，把 W 的正交归一基确定到一个公共的相因子(参考练习 0-3-6)。我们把 $SO(3)$群不可约表示空间的 W 满足条件(3.2.27)～(3.2.29)的基称为 $SO(3)$群不可约表示 $D^{[j]}(\dim W=2j+1)$的标准基或 Condon-Shortley 基。由式(3.2.32)和(3.2.33)可得

$$[J_x|_W]=\frac{1}{2}\begin{bmatrix} 0 & \sqrt{1(2j)} & 0 & \cdots & 0 & 0 & 0 \\ \sqrt{1(2j)} & 0 & \sqrt{2(2j-1)} & \cdots & 0 & 0 & 0 \\ 0 & \sqrt{2(2j-1)} & 0 & \cdots & 0 & 0 & 0 \\ & & & \cdots & & & \\ 0 & 0 & 0 & \cdots & 0 & \sqrt{(2j-1)2} & 0 \\ 0 & 0 & 0 & \cdots & \sqrt{(2j-1)2} & 0 & \sqrt{(2j)1} \\ 0 & 0 & 0 & \cdots & 0 & \sqrt{(2j)1} & 0 \end{bmatrix}$$

$$\equiv [J_x]_j \tag{3.2.34}$$

$$[J_y|_W]=\frac{i}{2}\begin{bmatrix} 0 & -\sqrt{1(2j)} & 0 & \cdots & 0 & 0 & 0 \\ \sqrt{1(2j)} & 0 & -\sqrt{2(2j-1)} & \cdots & 0 & 0 & 0 \\ 0 & \sqrt{2(2j-1)} & 0 & \cdots & 0 & 0 & 0 \\ & & & \cdots & & & \\ 0 & 0 & 0 & \cdots & 0 & -\sqrt{(2j-1)2} & 0 \\ 0 & 0 & 0 & \cdots & \sqrt{(2j-1)2} & 0 & -\sqrt{(2j)1} \\ 0 & 0 & 0 & \cdots & 0 & \sqrt{(2j)1} & 0 \end{bmatrix}$$

$$\equiv [J_y]_j \tag{3.2.35}$$

可见在标准基或 Condon-Shortley 基下，$J_x|_W$ 的矩阵是对称的实矩阵，$J_y|_W$ 的矩阵是反对称的纯虚矩阵。

练习 3-2-1 试证：$|jm\rangle=\left(\frac{(j+m)!}{(2j)!\ (j-m)!}\right)^{1/2} J_-^{j-m}|jj\rangle$。

Ⅱ. $SO(3)$群表示空间的分解

从角动量的一般定义出发，上一小节的讨论告诉我们，每一 $\boldsymbol{J}$ – 不可约不变子空间 W 的维数 $\dim W=2j+1$ 完全决定了 W 的转动性质。在 W 中存在唯一的(不计公共相因子)正交归一基$\{|jm\rangle|m=j,j-1,\cdots,-j\}$使 $\boldsymbol{J}|_W$ 的矩阵为(3.2.31)，(3.2.34)和(3.2.35)。可见，对每一个自然数 N，$SO(3)$群至多有一个 N 维不可约表示(在等价的意义下)。但上面的讨论并没有告诉我们，对每一个自然数 N，$SO(3)$群确有

一个 N 维不可约表示。然而，对任一自然数 N，令 $N=2j+1$，按式 (3.2.31)，(3.2.34)和(3.2.35)构造 $N\times N$ 矩阵$[\boldsymbol{J}]$，直接验算可知，它们是满足对易关系(3.2.1)的厄米矩阵。由 §3.1 的定理，它们是 $SO(3)$群 N 维不可约酉矩阵表示的无穷小算符的 i 倍。于是，对每一个自然数 N，$SO(3)$群都有一个，且只有一个 N 维不可约表示。由此可知，对任一非负整数或半整数 $j=\dfrac{N-1}{2}$，都有以 j 为角量子数的角动量算符。事实上，单粒子轨道角动量的角量子数可取所有非负整数，刚体角动量算符的角量子数可取所有非负整数和半整数，而电子自旋角动量算符只有一个角量子数$\dfrac{1}{2}$。一个具体的角动量算符有多少不可约不变子空间，各对应于什么角量子数，不能单由对易关系(3.2.1)和厄米性得出。

角动量算符 $\boldsymbol{J}$ 所作用的线性空间 V 就是 $SO(3)$群的酉表示 $T(\boldsymbol{\theta})=\exp(-i\boldsymbol{\theta}\cdot\boldsymbol{J})$ 的表示空间。一般而言，T 是不可约表示的直和。以 r_j 记不可约表示 $D^{[j]}$在 T 中出现的次数，以 $d(J_z,m)$，$d(J^2,j)$分别记算符 J_z 本征值 $m\hbar$ 的简并度和算符 J^2 本征值 $j(j+1)\hbar^2$ 的简并度。由上面的讨论可知，

$$d(J^2,j)=(2j+1)r_j,\ d(J_z,m)=\sum_{j=|m|,|m|+1,\cdots} r_j \tag{3.2.36}$$

由最后一个等式可得

$$r_j=d(J_z,j)-d(J_z,j+1) \tag{3.2.37}$$

当知道了 J_z 各个本征值的简并度 $d(J_z,m)$，由上式可以求出各个不可约表示 $D^{[j]}$在 T 中出现的次数 r_j，从而求出算符 J^2 各个本征值 $j(j+1)\hbar^2$ 的简并度 $d(J^2,j)$。

把每个不可约不变子空间的标准基合在一起我们就得到整个空间 V 的一个标准正交基。当所有的 $r_j\leq 1$ 时，我们可以用$|jm\rangle$标记基矢，而不引起歧义。但当某些 $r_j>1$ 时，亦即对某些 j 值，T 中有多个 $D^{[j]}$时，j 不足以确定不可约不变子空间。我们必须引入附加的量子数 η 赋予各个不可约不变子空间，对 j 相同的各个不可约不变子空间 η 必须取不同值。以$|\eta jm\rangle$标记 V 的基矢。这就是当 J^2，J_z 不构成力学量完备组时，需要引入附加的与 $\boldsymbol{J}$ 对易的力学量(或与 $\boldsymbol{J}$ 对易，本身又互相对易的力学量组)N，使$\{N,J^2,j_z\}$构成力学量完备组，而$\{\eta jm\}$是它们的量子数。

§3.3 转动算符在标准基下的矩阵

Ⅰ.欧拉角与旋转群的不变体积元

三维空间 $SO(3)$绕固定点 O 的旋转,在前面我们是用一个向量 $\boldsymbol{\theta}$(大小等于转角,方向为由右手螺旋法则决定的转轴正方向)来标志的。有时也用沿转轴的单位向量 $\hat{n}\left(=\frac{\boldsymbol{\theta}}{\theta}\right)$和转角 θ 来标志旋转 $\boldsymbol{\theta}=\hat{n}\theta$。这种记法对抽象的理论分析很合用。但对写下转动算符的矩阵元用欧拉角(α,β,γ)标记旋转更方便。

由于群的重排定理,有限群 G 具有如下重要性质:设 f 是群上函数$f:G\to\mathbf{C}$($\mathbf{C}$ 是复数域),则对任给的 $h\in G$ 都有

$$\sum_{g\in G}f(g)=\sum_{g\in G}f(hg)=\sum_{g\in G}f(gh) \tag{3.3.1}$$

利用这个性质,可以证明:(1)有限群的每一表示等价于一个酉表示,由于酉表示的任一不变子空间的正交补空间仍是酉表示的不变子空间,因而酉表示总是不可约表示的直和;(2)有限群不可约酉矩阵表示的矩阵元正交归一。旋转群 $SO(3)$不是有限群,也不是分立的无限群,而是连续群。它有没有与有限群性质(3.3.1)类似的性质?首先分析一下所谓"类似的性质"的具体含义。有限群性质(3.3.1),对连续群很自然应改写作:

$$\int_G \mathrm{d}g f(g)=\int_G \mathrm{d}g f(hg)=\int_G \mathrm{d}g f(gh) \tag{3.3.2}$$

上式左端由群的重排定理可以作变量变换 $g\mapsto hg$,$g\mapsto gh$ 写成

$$\int_G f(g)\mathrm{d}g=\int_G \mathrm{d}(hg)f(hg)=\int_G \mathrm{d}(gh)f(gh) \tag{3.3.3}$$

与(3.3.2)式比较,我们看到连续群 G 具有性质(3.3.2)的充分必要条件是群流形有(左、右)不变的体积元:对任给的 $h\in G$。

$$d(hg)=\mathrm{d}g=\mathrm{d}(gh) \tag{3.3.4}$$

对连续群不变体积元(不变测度)存在性条件有兴趣的读者可参考 W.Miller.Jr 的《对

称性群及其应用》一书的第六章。我们在此只指出旋转群 $SO(3)$ 有不变体积元

$$\mathrm{d}g = \sin\beta \mathrm{d}\alpha \mathrm{d}\beta \mathrm{d}\gamma \tag{3.3.5}$$

它代表满足下列条件的所有旋转 g(其欧拉角记作 $\alpha_g, \beta_g, \gamma_g$)对应的体积元:

$$\begin{aligned} \alpha < \alpha_g < \alpha + \mathrm{d}\alpha \\ \beta < \beta_g < \beta + \mathrm{d}\beta \\ \gamma < \gamma_g < \gamma + \mathrm{d}\gamma \end{aligned} \tag{3.3.6}$$

对任给的 $h \in SO(3)$,以 $(\alpha', \beta', \gamma')$ 代表 hg 的欧拉角(它们是 g 的欧拉角的函数)。使(3.3.6)成立的旋转 hg(h 固定)对应的体积元为:

$$\mathrm{d}(hg) = \sin\beta' \mathrm{d}\alpha' \mathrm{d}\beta' \mathrm{d}\gamma' \tag{3.3.7}$$

要证明(3.3.5)与(3.3.7)相等(左不变性),原则上应计算从变量 $(\alpha_g, \beta_g, \gamma_g)$ 到 $(\alpha', \beta', \gamma')$ 的 Jacobi 行列式 $J \equiv \left| \frac{\partial(\alpha'\beta'\gamma')}{\partial(\alpha_g\beta_g\gamma_g)} \right|$ 看看是否使下式成立:

$$(\sin\beta') \mid J \mid = \sin\beta \tag{3.3.8}$$

但在此我们将利用几何考虑,直接证明体积元(3.3.5)的左不变性。事实上,(3.3.5)中的 $\sin\beta \mathrm{d}\alpha \mathrm{d}\beta$ 是满足(3.3.6)式(前二不等式)的旋转 g 的本体坐标系 ζ 轴所张的立体角元,这是一个旋转不变量(从各个右手笛卡尔坐标系看大小都是一样)。而 $\mathrm{d}\gamma$ 代表满足(3.3.6)式的旋转 g,头两个欧拉角 α_g, β_g 固定,γ_g 在 $(\gamma, \gamma + \mathrm{d}\gamma)$ 范围内变化时,本体坐标系 η 轴在 $\zeta\eta$ - 平面所张的角元,这也是个转动不变量。回忆旋转 $hg = hgh^{-1}h$,可以看作先转 h,再相对新坐标系 $(R(h)\boldsymbol{e}_1, R(h)\boldsymbol{e}_2, R(h)\boldsymbol{e}_3)$ 做转动 g(亦即相对 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ 做转动 hgh^{-1})。因而从坐标系 $(R(h)\boldsymbol{e}_1, R(h)\boldsymbol{e}_2, R(h)\boldsymbol{e}_3)$ 看体积元 $\mathrm{d}(hg)$ 和从坐标系 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ 看体积元 $\mathrm{d}g$ 完全一样。体积元 $\mathrm{d}(hg)$ 是立体角元 $\sin\beta' \mathrm{d}\alpha' \mathrm{d}\beta'$ 和角元 $\mathrm{d}\gamma'$ 的积,但后两者是转动不变量,故从 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ 和从 $(R(h)\boldsymbol{e}_1, R(h)\boldsymbol{e}_2, R(h)\boldsymbol{e}_3)$ 看 $\mathrm{d}(hg)$ 大小一样。于是从 $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ 看,$\mathrm{d}g$ 的大小与 $\mathrm{d}(hg)$ 的大小一样:体积元(3.3.5)是左移不变的。

体积元(3.3.5)的右移不变性留给读者证明。

注意到

$$\int_{B_3} \mathrm{d}g = \int_0^{2\pi} \mathrm{d}\alpha \int_0^{\pi} \mathrm{d}\beta \int_0^{2\pi} \mathrm{d}\gamma \sin\beta = 8\pi^2 \tag{3.3.9}$$

因而与有限群 G(其阶数为 $|G|$)的求和式

$$\frac{1}{|G|}\sum_{g\in G}f(g) \tag{3.3.10}$$

相应的应是旋转群 $SO(3)$的积分

$$\frac{1}{8\pi^2}\int_0^{2\pi}\mathrm{d}\alpha\int_0^{\pi}\mathrm{d}\beta\int_0^{2\pi}\mathrm{d}\gamma\sin\beta f(\alpha,\beta,\gamma) \tag{3.3.11}$$

与有限群性质(3.3.1)相应的应是旋转群 $SO(3)$的性质:

$$\begin{aligned}\int_0^{2\pi}\mathrm{d}\alpha\int_0^{\pi}\mathrm{d}\beta\int_0^{2\pi}\mathrm{d}\gamma\sin\beta f(\alpha,\beta,\gamma)&=\int_0^{2\pi}\mathrm{d}\alpha\int_0^{\pi}\mathrm{d}\beta\int_0^{2\pi}\mathrm{d}\gamma\sin\beta f(\alpha',\beta',\gamma')\\&=\int_0^{2\pi}\mathrm{d}\alpha\int_0^{\pi}\mathrm{d}\beta\int_0^{2\pi}\mathrm{d}\gamma\sin\beta f(\alpha'',\beta'',\gamma'')\end{aligned} \tag{3.3.12}$$

式中(α',β',γ'),$(\alpha'',\beta'',\gamma'')$是 hg 和 gh 的欧拉角,h 是$SO(3)$的任一指定的元素。

上面的讨论表明,有关有限群表示的一些定理(每个表示都等价于一个酉表示;每个表示都是不可约表示的直和;不可约酉矩阵表示的矩阵元正交归一)对 $SO(3)$群继续成立。

练习 3-3-1 证明 $SO(3)$群体积元 $\mathrm{d}g=\sin\beta\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma$ 是右移不变的:

$$\mathrm{d}g=\mathrm{d}(gh),\forall\, h\in SO(3)$$

Ⅱ. 转动矩阵的一般性质

我们已指出,按旋转群 $SO(3)$的幺正表示 D 变换的 Hilbert 空间 V,分解成两两正交的不可约不可变子空间(角动量多重态子空间)的直和 $V=W\oplus W'\oplus\cdots$。各个不可约不变子空间的标准基合起来构成整个 Hilbert 空间的一个标准正交基。对任一旋转 $g\in SO(3)$,转动算符 $D(g)=D(g)|_W\oplus D(g)|_{W'}\oplus\cdots$在这个基下的矩阵是幺正矩阵

$$[D(g)]=[D(g)|_W]\oplus[D(g)|_{W'}]\oplus\cdots \tag{3.3.13}$$

在前面我们指出过角动量算符在标准基下的矩阵由 j 决定,因而转动算符在标准基下的矩阵$[D(g)|_W]=[D^{[j]}(g)](\dim W=2j+1)$由 j 决定。现在来讨论$\{[D^{[j]}(g)]|g\in SO(3)\}\left(j=0,\frac{1}{2},1,\frac{3}{2},2\cdots\cdots\right)$的矩阵元。

$D^{[j]}(g)$在标准基$|\eta jm\rangle$下的矩阵元为

$$\langle \eta jm' \mid D(g) \mid \eta jm \rangle = D^j_{m'm}(g) \tag{3.3.14}$$

式中 η 是 V 中角动量多重态子空间($\boldsymbol{J}$ - 不可约不变子空间)的标记(附加量子数)。可以选与 $\boldsymbol{J}$ 对易,并与 J^2, J_z 一起构成力学量完全集的厄米算符的量子数作为 η。(例如选中心力场中无旋粒子的哈密顿量的本征值 E_n(或主量子数 n)为 η。

转动算符矩阵的幺正性

$$[D^{[j]}(g)][D^{[j]}(g)]^+ = [D^{[j]}(g)]^+ [D^{[j]}(g)] = E_{2j+1} \tag{3.3.15}$$

可以写作

$$\sum_{m=-j}^{j} D^j_{m'm}(g) D^j_{m''m}(g)^* = \sum_{m=-j}^{j} (D^j_{mm'}(g))^* D^j_{mm''}(g) = \delta_{m'm''} \tag{3.3.16}$$

用欧拉角表达转动算符

$$D(\alpha,\beta,\gamma) = \exp(-i\alpha J_z)\exp(-i\beta J_y)\exp(-i\gamma J_z) \tag{3.3.17}$$

它们在标准基下的矩阵元是所谓的 D - 函数:

$$\begin{aligned} D^j_{m'm}(\alpha,\beta,\gamma) &= \langle \eta jm' \mid \exp(-i\alpha J_z)\exp(-i\beta J_y)\exp(-i\gamma J_z) \mid \eta jm \rangle \\ &= \exp(-im'\alpha) d^j_{m'm} \exp(-im\gamma) \end{aligned} \tag{3.3.18}$$

式中

$$d^j_{m'm}(\beta) = \langle \eta jm' \mid \exp(-i\beta J_y) \mid \eta jm \rangle \tag{3.3.19}$$

是与 η 无关的数。注意到在标准基下的 $[J_y]$ 矩阵元是纯虚的,故转动算符 $\exp(-j\beta J_y)$ 在标准基下的矩阵 $[\exp(-j\beta J_y)] = \exp(-j\beta[J_y])$ 是实的:

$$d^j_{m'm}(\beta) = d^j_{m'm}(\beta)^* \tag{3.3.20}$$

考虑到幺正性,我们有

$$d^j_{m'm}(-\beta) = d^j_{mm'}(\beta) \tag{3.3.21}$$

转动算符与角动量算之间的关系与它们在标准基下矩阵之间的关系相同:

$$[D(\alpha,\beta,\gamma)] = e^{(-i\alpha[J_z])} e^{(-i\beta[J_y])} e^{(-i\gamma[J_z])} \tag{3.3.22}$$

或者

$$[D(\boldsymbol{\theta})] = e^{-i\boldsymbol{\theta}\cdot[\boldsymbol{J}]} \tag{3.3.23}$$

以及

$$[J_\alpha] = i \frac{\partial}{\partial \theta_\alpha} [D(\boldsymbol{\theta})] \mid_{\boldsymbol{\theta}=0} \tag{3.3.24}$$

下面介绍 D – 函数作为 $SO(3)$群的群上函数的一些性质。

考虑 $SO(3)$群上连续函数 $f: SO(3)\to \mathbf{C}$ 的全体所成的复线性空间,在其中定义两个函数的内积

$$\langle \psi \mid \varphi \rangle = \int \mathrm{d}g \psi^*(g)\psi(g) = \int_0^{2\pi}\mathrm{d}\alpha\int_0^{\pi}\mathrm{d}\beta\int_0^{2\pi}\mathrm{d}\gamma \sin\beta \psi^*(\alpha,\beta,\gamma)\psi(\alpha,\beta,\gamma) \tag{3.3.25}$$

我们来证明作为 $SO(3)$群不可约表示矩阵元的 D – 函数具有如下的正交归一化性质:

$$\int_0^{2\pi}\mathrm{d}\alpha\int_0^{\pi}\mathrm{d}\beta\int_0^{2\pi}\mathrm{d}\gamma \sin\beta [D^{j_1}_{m_1' m_1}(\alpha,\beta,\gamma)]^* D^{j_2}_{m_2' m_2}(\alpha,\beta,\gamma) =$$

$$\langle D^{j_1}_{m_1' m_1} \mid D^{j_2}_{m_2' m_2} \rangle = \frac{8\pi^2}{2j_1+1}\delta_{j_1 j_2}\delta_{m_1' m_1}\delta_{m_2' m_2} \tag{3.3.26}$$

事实上,上式左边可以写作

$$\begin{aligned}\int \mathrm{d}g [D^{j_1}_{m_1' m_1}(g)]^* D^{j_2}_{m_2' m_2}(g) &= \langle \eta_1 j_1 m_1 \mid [\int \mathrm{d}g D(g^{-1}) \mid \eta_1 j_1 m_1' \rangle \langle \eta_2 j_2 m_2' \mid D(g)] \eta_2 j_2 m_2 \rangle \\ &= \langle \eta_1 j_1 m_1 \mid P \mid \eta_2 j_2 m_2 \rangle\end{aligned} \tag{3.3.27}$$

这是一个与 η_1, η_2 无关的数,尽管式中算符

$$P = \int \mathrm{d}g D(g^{-1}) \mid \eta_1 j_1 m_1' \rangle \langle \eta_2 j_2 m_2' \mid D(g) \tag{3.3.28}$$

和刁矢$\langle \eta_1 j_1 m_1 \mid$,刃矢$\mid \eta_2 j_2 m_2 \rangle$分别都与 η_1, η_2 有关。式(3.3.26)的左边还可以写作

$$\int \mathrm{d}g D^{j_2}_{m_2' m_2}(g)[D^{j_1}_{m_1' m_1}(g)]^* = \langle \eta_2 j_2 m_2' \mid Q \mid \eta_1 j_1 m_1' \rangle \tag{3.3.29}$$

式中算符

$$Q = \int \mathrm{d}g D(g) \mid \eta_2 j_2 m_2 \rangle \langle \eta_1 j_1 m_1 \mid D(g^{-1}) \tag{3.3.30}$$

容易看出 P, Q 是转动不变的算符:

$$\begin{aligned} D(h)PD(h^{-1}) &= P \\ D(h)QD(h^{-1}) &= Q, \quad \forall h \in SO(3)\end{aligned} \tag{3.3.31}$$

此即 P,Q 与角动量算符对易：

$$[P,\boldsymbol{J}]=0,[Q,\boldsymbol{J}]=0 \tag{3.3.32}$$

因而式(3.3.27)的右边可以写作(请读者证明之)

$$\delta_{j_1j_2}\delta_{m_1m_2}F(j_1,m_1',m_2') \tag{3.3.33}$$

式中 $F(j_1,m_1',m_2')$是某个依赖于 j_1,m_1',m_2' 的数。同样,式(3.3.29)的右边可以写作

$$\delta_{j_1j_2}\delta_{m_1'm_2'}G(j_1,m_1,m_2) \tag{3.3.34}$$

式(3.3.33)和式(3.3.34)都是式(3.3.26)的左边,故有

$$\langle D^{j_1}_{m_1'm_1}\mid D^{j_2}_{m_2'm_2}\rangle=\delta_{j_1j_2}\delta_{m_1'm_2'}\delta_{m_1m_2}A(j_1) \tag{3.3.35}$$

式中 $A(j_1)$是与 $\eta_1,\eta_2,m_1,m_2,m_1',m_2',j_2$ 无关的常数：

$$\begin{aligned}A(j_1)&=\int \mathrm{d}g[D^{j_1}_{m_1j_1}(g)]^*D^{j_1}_{m_1j_1}(g)=\frac{1}{2j_1+1}\sum_{m_1=-j_1}^{j_1}\int \mathrm{d}(g)D^{j_1}_{j_1m_1}(g^{-1})D^{j_1}_{m_1j_1}(g)\\&=\frac{1}{2j_1+1}\int \mathrm{d}g\delta_{j_1j_1}=\frac{8\pi^2}{2j_1+1}\end{aligned} \tag{3.3.36}$$

于是

$$\langle D^{j_1}_{m_1'm_1}\mid D^{j_2}_{m_2'm_2}\rangle=\frac{8\pi^2}{2j_1+1}\delta_{j_1j_2}\delta_{m_1m_2}\delta_{m_2'm_2'} \tag{3.3.37}$$

练习 3-3-2 证明式(3.3.27)的右端$\langle\eta_1j_1m_1\mid P\mid\eta_2j_2m_2\rangle$可以写作

$$\delta_{j_1j_2}\delta_{m_1m_2}F(j_1,m_1',m_2')$$

Ⅲ. D – 函数的解析表达式

由等式(3.3.18)知,为求 D – 函数的解析表达式,只需求 $d^j_{m'm}(\beta)$的解析表达式。有若干种计算 $d^j_{m'm}(\beta)$的方法。其中最常用的是建立 $d^j_{m'm}(\beta)$满足的微分方程,并求解这微分方程的初值问题。但我们在此介绍的是玻色子算符方法(或者叫谐振子上升、下降算符方法)。

设 $a_\alpha,a_\alpha^+(\alpha=1,2)$是玻色子湮灭、产生算符(二维谐振子上升、下降算符),$|0\rangle$是真空态(谐振子基态)。

$$[a_\alpha, a_\beta^+] = \delta_{\alpha\beta}, [a_\alpha, a_\beta] = [a_\alpha^+, a_\beta^+] = 0 \tag{3.3.38}$$

$$a_\alpha \mid 0\rangle = 0, \forall\, \alpha = 1,2 \tag{3.3.39}$$

令

$$\begin{aligned} J_x &= \frac{1}{2}(a_1^+ a_2 + a_2^+ a_1) \\ J_y &= -\frac{i}{2}(a_1^+ a_2 - a_2^+ a_1) \\ J_z &= \frac{1}{2}(a_1^+ a_1 - a_2^+ a_2) \end{aligned} \tag{3.3.40}$$

易知(3.3.40)是厄米算符,且满足对易关系

$$[J_\alpha, J_\beta] = i\sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma} J_\gamma \tag{3.3.41}$$

因而它们是角动量算符。相应的上升、下降算符是

$$J_+ = a_1^+ a_2, J_- = a_2^+ a_1 \tag{3.3.42}$$

它们是玻色子产生、湮灭算符之间的对易关系是

$$\begin{aligned} &[J_+, a_1^+] = 0, \quad [J_+, a_2] = 0 \\ &[J_+, a_2^+] = a_1^+, \quad [J_+, a_1] = -a_2 \\ &[J_-, a_2^+] = 0, \quad [J_-, a_1] = 0 \\ &[J_-, a_1^+] = a_2^+, \quad [J_-, a_2] = -a_1 \end{aligned} \tag{3.3.43}$$

考虑在玻色子态矢量空间(二维谐振子态矢量空间)S 中的由下式定义的 $2j+1$ 个矢量:(j 是非负整数或半整数,$m = j, j-1, \cdots, -j$)

$$\mid jm\rangle = \frac{(a_1^+)^{j+m}(a_2^+)^{j-m}}{\sqrt{(j+m)!(j-m)!}} \mid 0\rangle \tag{3.3.44}$$

不难验证

$$\begin{aligned} J_z \mid jm\rangle &= m \mid jm\rangle \\ J_\pm \mid jm\rangle &= \sqrt{(j \mp m)(j \pm m + 1)} \mid jm \pm 1\rangle \end{aligned} \tag{3.3.45}$$

从而

$$J^2 \mid jm\rangle = j(j+1) \mid jm\rangle \tag{3.3.46}$$

以及

$$\langle jm \mid jm \rangle = 1 \tag{3.3.47}$$

这表明,这 $2j+1$ 个态矢量(3.3.44)构成 $SO(3)$群不可约酉表示 $D^{[j]}$的标准基。现在来计算转动算符 $\exp(-j\beta J_y)$在这个标准基下的矩阵元,它由下式确定:

$$\exp(-i\beta J_y) \mid jm \rangle = \sum_{m'=-j}^{j} \mid jm' \rangle d^j_{m'm}(\beta) \tag{3.3.48}$$

利用

$$\boldsymbol{J} \mid 0 \rangle = 0 \tag{3.3.49}$$

式(3.3.48)的左边可以改写

$$\frac{(\tilde{a}_1^+(\beta))^{j+m}(\tilde{a}_2^+(\beta))^{j-m}}{\sqrt{(j+m)!(j-m)!}} \mid 0 \rangle = \sum_{m'} d^j_{m'm}(\beta) \mid jm' \rangle \tag{3.3.50}$$

其中

$$\begin{aligned} \tilde{a}_1^+(\beta) &= \exp(-j\beta J_y) a_1^+ \exp(j\beta J_y) \\ \tilde{a}_2^+(\beta) &= \exp(-j\beta J_y) a_2^+ \exp(j\beta J_y) \end{aligned} \tag{3.3.51}$$

它们满足微分方程

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}\beta}\tilde{a}_1^+(\beta) = \dfrac{1}{2}\tilde{a}_2^+(\beta) \\ \dfrac{\mathrm{d}}{\mathrm{d}\beta}\tilde{a}_2^+(\beta) = -\dfrac{1}{2}\tilde{a}_1(\beta) \end{cases} \tag{3.3.52}$$

和初值条件 $\tilde{a}_\alpha^+(0) = \tilde{a}_\alpha^+ (\alpha = 1,2)$

求解上面微分方程的初值问题得

$$\begin{aligned} \tilde{a}_1^+(\beta) &= \cos\frac{\beta}{2} a_1^+ + \sin\frac{\beta}{2} a_2^+ \\ \tilde{a}_2^+(\beta) &= -\sin\frac{\beta}{2} a_1^+ + \cos\frac{\beta}{2} a_2^+ \end{aligned} \tag{3.3.53}$$

代回式(3.3.50)的左边,利用牛顿二项式公式得

$$\sum_{\mu=0}^{j+m}\sum_{v=0}^{j-m}(-1)^v \binom{j+m}{\mu}\binom{j-m}{v}\left(\cos\frac{\beta}{2}\right)^{2j-\mu-v}\left(\sin\frac{\beta}{2}\right)^{\mu+v} \frac{(a_1^+)^{j+m-\mu-v}(a_2^+)^{j-m+\mu-v}}{\sqrt{(j+m)!(j-m)!}} \mid 0 \rangle$$

$$= \sum_{m'=-j}^{j} \sqrt{\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!}}$$

$$\times \sum_v (-1)^v \binom{j+m}{m-m'+\mu}\binom{j-m}{v}\left(\cos\frac{\beta}{2}\right)^{2j-m+m'-2v}\left(\sin\frac{\beta}{2}\right)^{m-m'+2v} \mid jm'\rangle$$

$$= \sum_{m'=-j}^{j} d^j_{m'm}(\beta) \mid jm'\rangle$$

故有

$$d^j_{m'm}(\beta) = \left(\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!}\right)^{1/2} \sum_v (-1)^v \binom{j+m}{m-m'+v}\binom{j-m}{v}\left(\cos\frac{\beta}{2}\right)^{2j-m+m'-2v}\left(\sin\frac{\beta}{2}\right)^{m-m'+2v}$$

$$= [(j+m)!(j-m)!(j+m')!(j-m')!]^{1/2} \sum_v (-1)^v \frac{1}{(j+m'-v)!(m-m'+v)!(j-m-v)!v!} \left(\cos\frac{\beta}{2}\right)^{2j-m+m'-2v}\left(\sin\frac{\beta}{2}\right)^{m-m'+2v} \tag{3.3.54}$$

上式是关于变量 $\cos\frac{\beta}{2}$ 和 $\sin\frac{\beta}{2}$ 的一个 $2j$ 次齐次多项式。如果把各项中的 $\cos\frac{\beta}{2}$ 的最低幂和 $\sin\frac{\beta}{2}$ 的最低幂作为公因子提到求和号外，则剩下的部分将是 $\cos^2\frac{\beta}{2}$ 和 $\sin^2\frac{\beta}{2}$ 的齐次多项式。注意到 $\cos^2\frac{\beta}{2}=\frac{1}{2}(1+\cos\beta)$ 和 $\sin^2\frac{\beta}{2}=\frac{1}{2}(1-\cos\beta)$，可见 $d^j_{m'm}(\beta)$ 是 $\cos\beta$ 的一个多项式乘以 $\cos\frac{\beta}{2}$ 和 $\sin\frac{\beta}{2}$ 的幂。变元 $x=\cos\beta$ 的这种多项式是所谓的 *Jacobi* 多项式，兹简介如下：*Jacobi* 多项式是变元 x 的由下式定义的 n 次多项式：①

$$P_n^{(\alpha,\beta)}(x) = \sum_s \binom{n+\alpha}{s}\binom{n+\beta}{n-s}\left(\frac{x-1}{2}\right)^{n-s}\left(\frac{x+1}{2}\right)^s \tag{3.3.55}$$

式中 n 是非负整数，α,β 是复参数。而对任意复数 $z\in\mathbf{C}$，

① 另一等价定义是 $P_n^{(\alpha,\beta)}(x)=\frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{\mathrm{d}^n}{\mathrm{d}x^n}[(1-x)^{\alpha+n}(1+x)^{\beta+n}]$。

$$\binom{z}{a}=\begin{cases}z(z-1)\cdots(z-a-1)/a!, & \text{当 } a=1,2,\cdots\\ 1, & \text{当 } a=0\\ 0, & \text{当 } a=-1,-2,\cdots\end{cases}\tag{3.3.56}$$

由此定义立即可得

$$P_n^{(\alpha,\beta)}(1)=\binom{n+\beta}{n}$$

$$P_n^{(\alpha,\beta)}(-x)=(-1)^n P_n^{(\beta,\alpha)}(x)\tag{3.3.57}$$

对我们的需要，只须讨论 $n, n+\alpha, n+\beta, n+\alpha+\beta$ 全是非负整数的情况。此时上述定义可以重新写作

$$P_n^{(\alpha,\beta)}(x)$$

$$=(n+\alpha)!(n+\beta)!\sum_s \frac{1}{s!(n+\alpha-s)!(\beta+s)!(n-s)!}\left(\frac{x-1}{2}\right)^{n-s}\left(\frac{x+1}{2}\right)^s\tag{3.3.58}$$

式中 s 取遍所有使分母中各个阶乘都有意义的值。在限制条件“$n, n+\alpha, n+\beta, n+\alpha+\beta$ 全是非负整数”之下，参数 α, β 仍有可能取负整数值。由定义(3.3.58)可得：对整数 α, β

$$\begin{aligned}P_n^{(\alpha,\beta)}(x)&=\frac{(n+\alpha)!(n+\beta)!}{n!(n+\alpha+\beta)!}\left(\frac{x+1}{2}\right)^{-\beta}P_{n+\beta}^{(\alpha,-\beta)}(x)\\ P_n^{(\alpha,\beta)}(x)&=\frac{(n+\alpha)!(n+\beta)!}{n!(n+\alpha+\beta)!}\left(\frac{x-1}{2}\right)^{-\alpha}P_{n+\alpha}^{(-\alpha,\beta)}(x)\\ P_n^{(\alpha,\beta)}(x)&=\left(\frac{x-1}{2}\right)^{-\alpha}\left(\frac{x+1}{2}\right)^{-\beta}P_{n+\alpha+\beta}^{(-\alpha,-\beta)}(x)\end{aligned}\tag{3.3.59}$$

比较(3.3.54)和(3.3.58)我们得到

$$d^j_{m'm}(\beta)=\left(\frac{(j+m)!(j-m)!}{(j+m')!(j-m')!}^{1/2}\right)\left(\sin\frac{\beta}{2}\right)^{m-m'}\left(\cos\frac{\beta}{2}\right)^{m+m'}P_{j-m}^{(m-m',m+m')}(\cos\beta)\tag{3.3.60}$$

上式对 $m', m=-j, -j+1, \cdots, j; j$ 为任一非负整数或半整数全成立。因子 $\sin\frac{\beta}{2}$ 的指数 $m-m'$ 和 $\cos\frac{\beta}{2}$ 的指数 $m+m'$ 可正，可负。式(3.3.60)是 $\sin\frac{\beta}{2}$ 和 $\cos\frac{\beta}{2}$ 的多项式的性质不再是明显的了。为使之明显，令

$$k = \min\{j+m, j-m, j+m', j-m'\} \tag{3.3.61}$$

对非负整数 k 的四种可能，定义 μ, v 和 λ：

k	μ	v	λ
$j+m$	$m'-m$	$-m'-m$	$m'-m$
$j-m$	$m-m'$	$m+m'$	0
$j+m'$	$m-m'$	$-m'-m$	0
$j-m'$	$m'-m$	$m'+m$	$m'-m$

利用这些非负的参数可以把表作明显的多项式形式

$$d^j_{m'm}(\beta) = (-1)^\lambda \left(\frac{k!(2j-k)!}{(k+\mu)!(k+v)!}\right)^{1/2} \left(\sin\frac{\beta}{2}\right)^\mu \left(\cos\frac{\beta}{2}\right)^v P_k^{(\mu,v)}(\cos\theta) \tag{3.3.62}$$

Ⅳ. D – 函数的对称性

由 $d^j_{m'm}(\beta)$ 的解析表达式(3.3.54)，可以看出它有如下的对称性：

$$d^j_{m'm}(-\beta) = d^j_{mm'}(\beta) \quad (\text{由}[D(0,\beta,0)]\text{的实幺正性也可得此式})$$

$$d^j_{m'm}(-\beta) = (-1)^{m'-m} d^j_{m'm}(\beta) \tag{3.3.63}$$

$$d^j_{m'm}(\beta) = d^j_{-m-m'}(\beta)$$

由(3.3.63)的头两式可得

$$d^j_{m'm}(\beta) = (-1)^{m'-m} d^j_{mm'}(\beta) \tag{3.3.64}$$

再由(3.3.63)的第三式可得

$$d^j_{-m-m'}(\beta) = (-1)^{m'-m} d^j_{mm'}(\beta) \tag{3.3.65}$$

由式(3.3.60)和(3.3.57)可得

$$d^j_{m'-m}(\pi) = (-1)^{j-m}\delta_{m',-m} \tag{3.3.66}$$

再由(3.3.63)得

$$d^j_{m'-m}(-\pi) = (-1)^{j+m}\delta_{m',-m} \tag{3.3.67}$$

式(3.3.66)写成矩阵形式就是下面的$(2j+1)\times(2j+1)$方阵

$$d^j(\pi) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & \iddots & 0 & 0 & 0 & 0 \\ (-1)^{2j} & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.3.68}$$

由群性质有

$$[d^j(2\pi)] = [d^j(\pi)][d^j(\pi)] = (-1)^{2j}E \tag{3.3.69}$$

(E 是$(2j+1)\times(2j+1)$单位矩阵)和

$$[d^j(\beta + 2\pi)] = (-1)^{2j}[d^j(\beta)] \tag{3.3.70}$$

可见 j 为半整数时,$D^{[j]}$是 $SO(3)$群的双值表示。

由 $d^j_{m'm}(\beta)$的上述对称关系和特殊值,利用群性质可得

$$d^j_{m'm}(\pi + \beta) = (-1)^{j+m'}d^j_{-m'm}(\beta) \tag{3.3.71}$$

$$d^j_{m'm}(\pi - \beta) = (-1)^{j+m'}d^j_{m-m'}(\beta) \tag{3.3.72}$$

以上结果可用来讨论 $D^j_{m'm}(\alpha,\beta,\gamma)$的对称性。例如:

$$D^j_{m'm}(\alpha,\beta+\pi,\gamma) = (-1)^{j+m'}D^j_{-m'm}(-\alpha,\beta,\gamma) \tag{3.3.73}$$

$$\begin{aligned}[D^j_{m'm}(\alpha,\beta,\gamma)]^* &= e^{im'\alpha}d^j_{m'm}(\beta)e^{im\gamma} = (-1)^{m-m'}e^{im'\alpha}d^j_{-m'-m}(\beta)e^{im\gamma} \\ &= (-1)^{m'-m}D^j_{-m'-m}(\alpha,\beta,\gamma)\end{aligned} \tag{3.3.74}$$

最后我们愿意指出 D – 函数与球谐函数的关系。回忆伴随 *Legendre* 函数的定义:

$$P^m_l(x) = (1-x^2)^{m/2}\frac{d^m}{dx^m}P_l(x),(0 \le m \le l)$$

或者

$$P^m_l(x) = \frac{1}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l,(|m| \le l) \tag{3.3.75}$$

与 *Jacobi* 多项式的定义

$$P^{(\alpha,\beta)}_n(x) = \frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}[(1-x)^{\alpha+n}(1+x)^{\beta+n}] \tag{3.3.76}$$

比较它们并考虑(3.3.59)的第三个等式可得

$$P_l^m(x) = \frac{(l+m)!}{2^m l!}(1-x^2)^{m/2}P_{l-m}^{(m,m)}(x) \tag{3.3.77}$$

于是由式(3.3.60)得到

$$d_{0m}^l(\beta) = \left(\frac{(l-m)!}{(l+m)!}\right)^{1/2}P_l^m(\cos\beta) \tag{3.3.78}$$

特别是当 $m=0$ 时有

$$d_{00}^l(\beta) = P_l(\cos\beta) \tag{3.3.79}$$

把球函数的相因子取作(有不同的取法)

$$\begin{aligned} Y_{lm}(\theta,\varphi) &= i^l(-1)^{\frac{m+|m|}{2}}\sqrt{\frac{2l+1}{4\pi}\times\frac{(l-|m|)!}{(l+|m|)!}}P_l^{|m|}(\cos\theta)e^{im\varphi} \\ &= i^l(-1)^m\sqrt{\frac{2l+1}{4\pi}\times\frac{(l-|m|)!}{(l+|m|)!}}P_l^{|m|}(\cos\theta)e^{im\varphi} \end{aligned} \tag{3.3.80}$$

与常见的定义相差一个相因子 i^l,并不影响 *Condon-Shortley* 约定。但这样做可使时间反演态正好是对偶基

$$Y_{lm}^*(\theta,\varphi) = (-1)^{l-m}Y_{l-m}(\theta,\varphi) \tag{3.3.81}$$

在这种位相约定下,我们有

$$D_{0m}^l(\alpha\beta\gamma) = d_{0m}^l(\beta)e^{-im\gamma} = i^l(-1)^m\sqrt{\frac{4\pi}{2l+1}}Y_{lm}^*(\beta\gamma) \tag{3.3.82}$$

$$D_{m0}^l(\alpha\beta\gamma) = e^{-im\alpha}d_{m0}^l(\beta) = (-1)^m d_{0m}^l(\beta)e^{-im\alpha} = i^l\sqrt{\frac{4\pi}{2l+1}}Y_{lm}^*(\beta\alpha) \tag{3.3.83}$$

§3.4 两个角动量的耦合

考虑碱金属原子。价电子在原子实电场中的哈密顿算符为

$$H = H_0 + H' \tag{3.4.1}$$

式中

$$H_0 = \frac{-\hbar^2}{2\mu}\nabla^2 + U(r) \tag{3.4.2}$$

$$H' = \frac{1}{2\mu^2 c^2 r}\frac{\mathrm{d}U(r)}{\mathrm{d}r}\boldsymbol{L}\cdot\boldsymbol{S} \tag{3.4.3}$$

(参看相对论量子力学一章中电子的狄拉克方程) H' 与 H_0 大小之比约为 $\frac{v^2}{c^2}$,故在零级近似时,可以忽略。H_0 与 $\boldsymbol{L},\boldsymbol{S}$ 都对易。$\{H_0,\boldsymbol{L}^2,L_z,(\boldsymbol{S}^2),S_z\}$ 可以构成力学量完全集。它们的共同的本征函数可以记作 $|n,l,m_l,(s),m_s>$,对应于能级 ε_{nl},但是

$$H' = \frac{1}{2\mu^2 c^2 r}\frac{\mathrm{d}U(r)}{\mathrm{d}r}\frac{\boldsymbol{J}^2-\boldsymbol{L}^2-\boldsymbol{S}^2}{2} \tag{3.4.4}$$

H' 与 $\boldsymbol{J}$ 对易,不再与 $\boldsymbol{L}$ 和 $\boldsymbol{S}$ 对易,但与 $\boldsymbol{L}^2$ 和 $\boldsymbol{S}^2$ 对易。于是 $\{H,\boldsymbol{L}^2,(\boldsymbol{S}^2),\boldsymbol{J}^2,J_z\}$ 构成力学量完全组。按照与时间无关的微扰理论(H_0 有简并情况),$|n,l,m_l,(s),m_s\rangle$ 不是好零级近似本征函数。应把对应于能级 ε_{nl} 的各态矢 $|n,l,m_l,(s),m_s\rangle$ 线性组合成 $\boldsymbol{J}^2,J_z$ 的本征函数,即力学量完全组 $\{H_0,\boldsymbol{L}^2,(\boldsymbol{S}^2),\boldsymbol{J}^2,J_z\}$ 的共同的本征函数,才是 H 的好零级近似本征函数。于是物理向我们提出了两个角动量的耦合问题。

Ⅰ.和角动量的多重态子空间

设我们的体系是由两个子体系构成的,而且子体系的态矢量空间 $S(1)$ 和 $S(2)$,在空间转动下分别按旋转群 $SO(3)$ 的幺正表示 $D^{(1)}$ 和 $D^{(2)}$ 变换。复合体系的态矢量空间是张量积

$$S = S(1)\otimes S(2) \tag{3.4.5}$$

复合体系的态矢量总可以写作 $\sum_\alpha |x_\alpha\rangle\otimes y_\alpha\rangle\in S$,式中 $|x_\alpha\rangle\in S(1),|y_\alpha\rangle\in S(2)$ 分别是子体系 1 与子体系 2 的态矢量。体系的态矢量在空间转动 g 下,显然应按如下方式变换:

$$\sum_\alpha |x_\alpha\rangle\otimes y_\alpha\rangle \mapsto \sum_\alpha D^{(1)}(g)|x_\alpha\rangle\otimes D^{(2)}(g)|y_\alpha\rangle$$
$$= D^{(1)}(g)\otimes D^{(2)}(g)\sum_\alpha |x_\alpha\rangle\otimes y_\alpha\rangle$$

可见它们按旋转群 $SO(3)$ 的直积表示

$$D = D^{(1)} \otimes D^{(2)} \tag{3.4.6}$$

变换。若定义 S 中两个向量的内积，使得对 $z = |x\rangle \otimes |y\rangle \in S, w = |u\rangle \otimes |v\rangle \in S$，有

$$\langle z \mid w \rangle = \langle x \mid u \rangle \langle y \mid v \rangle \tag{3.4.7}$$

则 D 是 $SO(3)$ 在 S 上的幺正表示。由前面的讨论知，复合体系的角动量算符应为

$$\begin{aligned} J_\alpha &= i \frac{\partial}{\partial \theta_\alpha} D(\boldsymbol{\theta}) \mid_{\boldsymbol{\theta}=0} = i \frac{\partial}{\partial \theta_\alpha} [D^{(1)}(\boldsymbol{\theta}) \otimes D^{(2)}(\boldsymbol{\theta})] \Big|_{\boldsymbol{\theta}=0} \\ &= i \Big[\frac{\partial D^{(1)}(\boldsymbol{\theta})}{\partial \theta_\alpha} \otimes D^{(2)}(\boldsymbol{\theta}) \Big] \mid_{\boldsymbol{\theta}=0} + i \Big[D^{(1)}(\boldsymbol{\theta}) \otimes \frac{\partial D^{(2)}(\boldsymbol{\theta})}{\partial \theta_\alpha} \Big] \Big|_{\boldsymbol{\theta}=0} \\ &= J_\alpha(1) \otimes E(2) + E(2) \otimes J_\alpha(2) \equiv J_\alpha(1) + J_\alpha(2) \end{aligned} \tag{3.4.8}$$

式中 $E(\alpha)$ 是 $S(\alpha)$ 上的恒等算符。于是复合体系的角动量算符是各子体系角动量算符之和

$$\boldsymbol{J} = \boldsymbol{J}(1) + \boldsymbol{J}(2) \tag{3.4.9}$$

作为群 $SO(3)$ 幺正表示 D 的无穷小算符(的 i 倍)，显然应有

$$J_\alpha^+ = J_\alpha \qquad (\alpha = 1,2,3) \tag{3.4.10}$$

$$[J_\alpha, J_\beta]_- = i \sum_\gamma \in_{\alpha\beta\gamma} J_\gamma \tag{3.4.11}$$

由于作用在不同子体系态矢量空间上，显然有

$$[J_\alpha(1), J_\beta(2)]_- = 0 \tag{3.4.12}$$

从前面的讨论可知 $S(\alpha)(\alpha = 1,2)$ 分解成两两正交的 $\boldsymbol{J}(\alpha)$ - 不可约不变子空间的直和。设 $L(\alpha)(\alpha = 1,2)$ 是 $S(\alpha)$ 的 $\boldsymbol{J}(\alpha)$ - 不可约不变子空间($\boldsymbol{J}(\alpha)$ - 多重态子空间)。考察它们的张量积

$$L(1) \otimes L(2) \equiv L \tag{3.4.13}$$

L 是 $S = S(1) \otimes S(2)$ 的一个($\boldsymbol{J}(1), \boldsymbol{J}(2)$)不可约不变子空间，因而是 $\boldsymbol{J}$ - 不变子空间(不一定不可约)，总可以分解成 $\boldsymbol{J}$ - 多重态子空间的直和。

我们的问题是，当 $L(1)$ 是 $2j_1 + 1$ 维的 L_{j_1}，$L(2)$ 是 $2j_2 + 1$ 维的 L_{j_2}，则 L 将包含哪些 $\boldsymbol{J}$ - 多重态子空间。以 $\{|j_1 m_1\rangle \mid -j_1 \le m_1 \le j_1\}$ 和 $\{|j_2 m_2\rangle \mid -j_2 \le m_2 \le j_2\}$ 分别记 $L(1)$ 和 $L(2)$ 的标准基，则

$$\{|j_1 m_1\rangle \otimes |j_2 m_2\rangle \mid -j_1 \le m_1 \le j_1, -j_2 \le m_2 \le j_2\} \tag{3.4.14}$$

构成 L 的一个标准正交基，称作非耦合基。这$(2j_1+1)(2j_2+1)$个基矢都是 $J^2(1)$，$J_z(1)$，$J^2(2)$，$J_z(2)$的共同的本征矢量。为简便，把它们记作

$$|j_1m_1\rangle \otimes |j_2m_2\rangle = |j_1m_1j_2m_2\rangle \tag{3.4.15}$$

不难看出它们是 J_z 的本征矢量：

$$J_z|j_1m_1j_2m_2\rangle = (m_1+m_2)|j_1m_1j_2m_2\rangle \tag{3.4.16}$$

由下表可以看出它们对应于 J_z 的哪些本征值(为了确定起见，设 $j_1 \geq j_2$)：

m_1 \ m_2	j_2	j_2-1	j_2-2	$\cdots$	$-(j_2-2)$	$-(j_2-1)$	$-j_2$
j_1	j_1+j_2	j_1+j_2-1	j_1+j_2-2	$\cdots$	j_1-j_2+2	j_1-j_2+1	j_1-j_2
j_1-1	j_1+j_2-1	j_1+j_2-2	j_1+j_2-3	$\cdots$	j_1-j_2+1	j_1-j_2	j_1-j_2-1
j_1-2	j_1+j_2-2	j_1+j_2-3	j_1+j_2-4	$\cdots$	j_1-j_2	j_1-j_2-1	j_1-j_2-2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
$-(j_1-2)$	$-j_1+j_2+2$	$-j_1+j_2+1$	$-j_1+j_2$	$\cdots$	$-j_1-j_2+4$	$-j_1-j_2+3$	$-j_1-j_2+2$
$-(j_1-1)$	$-j_1+j_2+1$	$-j_1+j_2$	$-j_1+j_2-1$	$\cdots$	$-j_1-j_2+3$	$-j_1-j_2+2$	$-j_1-j_2+1$
$-j_1$	$-j_1+j_2$	$-j_1+j_2-1$	$-j_1+j_2-2$	$\cdots$	$-j_1-j_2+2$	$-j_1-j_2+1$	$-j_1-j_2$

非耦合基矢$|j_1m_1j_2m_2\rangle$所属的 J_z 的本征值出现在上表格的 m_1 – 行 m_2 – 列的交点。我们看到 J_z 到 L 上的限制 $J_z|_L$ 的各个本征值 m 的简度为

$$d(J_z|_L,m)=\begin{cases}0, & \text{当 } |m|>j_1+j_2,\\ j_1+j_2-|m|+1, & \text{当 } |j_1-j_2|\leq|m|\leq j_1+j_2,\\ 2j_2+1, & \text{当 } |m|\leq|j_1-j_2|,\end{cases} \tag{3.4.17}$$

由 $\boldsymbol{J}$ – 多重态空间的结构知恒有

$$d(J_z|_L,m)=d(J_z|_L,-m) \tag{3.4.18}$$

由(3.2.37)和(3.4.17)可得 $SO(3)$群的不可约表示 $D^{[j]}$在表示$\{D(g)|_L|g\in SO(3)\}$中出现的次数为

$$r_j=\begin{cases}0, j\geq j_1+j_2\\ 1, |j_1-j_2|\leq j\leq j_1+j_2\\ 0, j<|j_1-j_2|\end{cases} \tag{3.4.19}$$

不难相信，由于 j_1, j_2 在问题中的地位相同，式(3.4.19)对 $j_1 \leq j_2$ 的情况也成立。于是 L 包含权 j 等于 $j_1+j_2, j_1+j_2-1, \cdots, |j_1-j_2|$ 的 $\boldsymbol{J}$ - 多重态子空间各一次：(三角形关系选择定则)。用群论的语言说，就是群 $SO(3)$ 的不可约表示 $D^{[j_1]}$ 与 $D^{[j_2]}$ 的直积，可以约化成不可约表示 $D^{[j_1+j_2]}, \cdots, D^{[|j_1-j_2|]}$ 的直和

$$L_{j_1} \otimes L_{j_2} = L = \bigoplus_{j=-j}^{j} L_j, D^{[j_1]} \otimes D^{[j_2]} = \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} D^{[j]} \tag{3.4.20}$$

考虑到 $D^{[j]}$ 是 $2j+1$ 维的。上式右边各不可约表示维数之和

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} (2j+1) = (2j_1+1)(2j_2+1) \tag{3.4.21}$$

正好是张量积空间 L 的维数。式(3.4.20)右边各个不可约表示的标准基合在一起，便得到张量积空间 L 的一个标准正交基。每个基矢都是 $\boldsymbol{J}^2(1), \boldsymbol{J}^2(2), \boldsymbol{J}^2, J_z$ 的共同的本征矢，我们以相应的量子数标记之：

$$\{|j_1, j_2, j, m\rangle \,|\, |j_1-j_2| \leq j \leq j_1+j_2, -j \leq m \leq j\} \tag{3.4.22}$$

并称之为子空间 L 的耦合基。为避免混淆，在必要时分别把耦合基矢与非耦合基矢记作 $|j_1 j_2 j m\rangle_c$，$|j_1 m_1 j_2 m_2\rangle_u$。

Ⅱ. Clebsch-Gordan 系数

空间 L 的两个标准正交基(3.4.14)和(3.4.22)决定了一个幺正变换矩阵，其矩阵元由下式确定

$$|j_1 j_2 j m\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1 m_1 j_2 m_2\rangle\langle j_1 m_1 j_2 m_2 | j_1 j_2 j m\rangle \tag{3.4.23}$$

称为 Clebsch-Gordan 系数。(3.4.23)的逆变换为

$$|j_1 m_1 j_2 m_2\rangle = \sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{m=-j}^{j} |j_1 j_2 j m\rangle\langle j_1 j_2 j m | j_1 m_1 j_2 m_2\rangle \tag{3.4.24}$$

由非耦合基到耦合基的变换矩阵的幺正性可以写作：

$$\sum_{m_1 m_2} \langle j_1 m_1 j_2 m_2 | j_1 j_2 j' m'\rangle^* \langle j_1 m_1 j_2 m_2 | j_1 j_2 j m\rangle = \delta_{j'j}\delta_{m'm} \tag{3.4.25}$$

和

$$\sum_{jm}\langle j_1j_2jm \mid j_1m_1'j_2m_2'\rangle^*\langle j_1j_2jm \mid j_1m_1j_2m_2\rangle = \delta_{m_1'm_1}\delta_{m_2'm_2} \tag{3.4.26}$$

下面让我们来确定 $C.-G.$ 系数

首先我们指出，由于 $|j_1m_1j_2m_2\rangle$ 是厄米算符 J_z 的本征矢(属于本征值 m_1+m_2)，而 $|j_1j_2jm\rangle$ 也是 J_z 的本征矢(属于本征值 m)，故有选择定则

$$\langle j_1m_1j_2m_2 \mid j_1j_2jm\rangle = 0, \forall\, m \neq m_1+m_2 \tag{3.4.27}$$

故式(3.4.23)中的求和只需对使 $m_1+m_2=m$ 的(m_1,m_2)求和就行：

$$|j_1j_2jm\rangle = \sum_{m_1+m_2=m} |j_1m_1j_2m_2\rangle\langle j_1m_1j_2m_2 \mid j_1j_2jm\rangle \tag{3.4.28}$$

而式(3.4.24)可以写作

$$|j_1m_1j_2m_2\rangle = \sum_j |j_1j_2jm_1+m_2\rangle\langle j_1j_2jm_2'+m_2\rangle |j_1m_1j_2m_2\rangle \tag{3.4.29}$$

由式(34.28)，当 $j=j_1+j_2, m=j$ 时，右边的求和项只含一项

$$|j_1j_2j_1+j_2j_1+j_2\rangle = |j_1j_1j_2j_2\rangle\langle j_1j_1j_2j_2 \mid j_1j_2j_1+j_2j_1+j_2\rangle \tag{3.4.30}$$

由幺正性条件(3.4.25)，有

$$|\langle j_1j_1j_2j_2 \mid j_1j_2j_1+j_1j_1+j_2\rangle| = 1 \tag{3.4.31}$$

我们这样确定 $|j_1j_2j_1+j_1j_1+j_1\rangle$ 的相因子，使

$$\langle j_1j_1j_2j_2 \mid j_1j_2j_1+j_2j_1+j_2\rangle = 1 \tag{3.4.32}$$

亦即取

$$|j_1j_2j_1+j_2j_1+j_2\rangle = |j_1j_1j_2j_2\rangle \tag{3.4.33}$$

利用下降算符公式(3.2.28)，或练习 3-2-1 的结果

$$|j_1j_2jm\rangle = \left[\frac{(j+m)!}{(2j)!(j-m)!}\right]^{1/2}(J_-(1)+J_-(2))^{j-m}|j_1j_2jj\rangle \tag{3.4.34}$$

我们从(3.4.33)可以得出 $2(j_1+j_2)+1$ 个 $j=j_1+j_2$ 的耦合基矢。

一般而言，只要确定了 $|j_1j_2jj\rangle(j=j_1+j_2,j_1+j_2-1,\cdots,|j_1-j_2|)$，利用式(3.4.34)，就可以得到全部 $2j+1$ 个耦合基矢 $|j_1j_2jm\rangle(m=-j,-j+1,\cdots,j)$。

为了确定 $|j_1,j_2,j_1+j_2-1,j_1+j_2-1\rangle$ 注意它的非耦合基的展开式(3.4.28)只包含两项，由它与 $|j_1,j_2,j_1+j_2,j_1+j_2-1\rangle$ 的正交性就可以把它确定到只差一个位相因子。再要求(位相约定)

$$\langle j_1j_1j_2j_2-1 \mid j_1j_2j_1+j_2-1,j_1+j_2-1\rangle > 0 \tag{3.4.35}$$

就可以完全确定$|j_1j_2j_1+j_2-1,j_1+j_2-1\rangle$从而确定全部$|j_1j_2j_1+j_2-1,m\rangle$。

$|j_1j_2j_1+j_2-2,j_1+j_2-2\rangle$的展开式(3.4.28)只包含三项。利用正交性

$$0=\langle j_1j_2j_1+j_2,j_1+j_2-2\,|\,j_1j_2j_1+j_2-2,j_1+j_2-2\rangle$$

和

$$0=\langle j_1j_2j_1+j_2-1,j_1+j_2-2\,|\,j_1j_2j_1+j_2-2,j_1+j_2-2\rangle \tag{3.4.36}$$

以及位相约定

$$\langle j_1j_1j_2j_2-2\,|\,j_1j_2j_1+j_2-2,j_1+j_2-2\rangle>0 \tag{3.4.37}$$

就可以确定$|j_1j_2j_1+j_2-2,j_1+j_2-2\rangle$,从而全部$|j_1j_2j_1+j_2-2,m\rangle$。

继续下去,可以确定全部耦合基矢向非耦合矢的展开式。

以上的讨论是确定 $C.-G.$系数的大致方法。但这个方法要求

$$\langle j_1j_1j_2j-j_1\,|\,j_1j_2jj\rangle\neq 0 \tag{3.4.38}$$

位相规定$\langle j_1j_1j_2j-j_1|j_1j_2jj\rangle>0$才有意义。下面我们将会看到式(3.4.38)是成立的。

具体确定 $C.-G.$系数一般表达式的关键,是利用上升下降算符建立递推关系。

以 $J_+=J_+(1)+J_+(2)$作用于式(3.4.28)得

$$\begin{aligned}&\sqrt{(j-m)(j+m+1)}\,|\,j_1j_2jm+1\rangle\\&=\sum_{m_1+m_2=m}\{|\,j_1m_1+1j_2m_2\rangle\sqrt{(j_1-m_1)(j_1+m_1+1)}\langle j_1m_1j_2m_2\,|\,j_1j_2jm\rangle\\&\quad+|\,j_1m_1j_2m_2+1\rangle\sqrt{(j_2-m_2)(j_2+m_2+1)}\langle j_1m_1j_2m_2\,|\,j_1j_2jm\rangle\}\end{aligned}$$

上式可以改写作

$$\begin{aligned}&\sqrt{(j-m)(j+m+1)}\sum_{m_1+m_2=m+1}|\,j_1m_1j_2m_2\rangle\langle j_1m_1j_2m_2\,|\,j_1j_2jm+1\rangle\\&=\sum_{m_1+m_2=m+1}\{|\,j_1m_1j_2m_2\rangle[\sqrt{(j_1+m_1)(j_1-m_1+1)}\langle j_1m_1-1j_2m_2\,|\,j_1j_2jm\rangle\\&\quad+\sqrt{(j_2+m_2)(j_2-m_2+1)}\langle j_1m_1j_2m_2-1\,|\,j_1j_2jm\rangle]\}\end{aligned} \tag{3.4.39}$$

于是得到递推关系

$$\sqrt{(j-m)(j+m+1)}\langle j_1m_1j_2m_2\,|\,j_1j_2jm+1\rangle$$

$$= \sqrt{(j_1 + m_1)(j_1 - m_1 + 1)}\langle j_1 m_1 - 1 j_2 m_2 \mid j_1 j_2 j m\rangle$$
$$+ \sqrt{(j_2 + m_2)(j_2 - m_2 + 1)}\langle j_1 m_1 j_2 m_2 - 1 \mid j_1 j_2 j m\rangle \quad (m_1 + m_2 = m + 1) \tag{3.4.40}$$

同样以 $J_- = J_-(1) + J_-(2)$ 作用于式(3.4.28)得递推关系

$$\sqrt{(j + m)(j - m + 1)}\langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j m - 1\rangle$$
$$= \sqrt{(j_1 - m_1)(j_1 + m_1 + 1)}\langle j_1 m_1 + 1 j_2 m_2 \mid j_1 j_2 j m\rangle$$
$$+ \sqrt{(j_2 - m_2)(j_2 + m_2 + 1)}\langle j_1 m_1 j_2 m_2 + 1 \mid j_1 j_2 j m\rangle \quad (m_1 + m_2 = m - 1) \tag{3.4.41}$$

由递推关系(3.4.40)可知,当 $m = j$ 时有

$$\sqrt{(j_1 + m_1)(j_1 - m_1 + 1)}\langle j_1 m_1 - 1 j_2 m_2 \mid j_1 j_2 j j\rangle$$
$$+ \sqrt{(j_2 + m_2)(j_2 - m_2 + 1)}\langle j_1 m_1 j_2 m_2 - 1 \mid j_1 j_2 j j\rangle = 0 \quad (m_1 + m_2 = j + 1) \tag{3.4.42}$$

把上式中的 $m_1 - 1$ 换成 m_1 得

$$\langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j j\rangle$$
$$= \frac{-\sqrt{(j_2 + m_2)(j_2 - m_2 + 1)}}{\sqrt{(j_1 - m_1)(j_1 + m_1 + 1)}}\langle j_1 m_1 + 1 j_2 m_2 - 1 \mid j_1 j_2 j j\rangle, (m_1 + m_2 = j) \tag{3.4.43}$$

上式对所有满足 $j - j_1 \leq m_1 < j_1$(由 $-j_1 \leq m_1 < j_1$, $-j_2 \leq m_2 \leq j_2$, $m_1 + m_2 = j$ 推知)的 m_1 都成立。反复使用式(3.4.43)可得,对所有使 $j - j_2 \leq m_1 < j_1$ 成立的 m_1 都有

$$\langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j j\rangle = (-1)^{j_1 - m_1}\left[\frac{(j_1 + j_1 - j)!(j_1 + m_1)!(j_2 + m_2)!}{(2j_1)!(2j_2 + j - j_1)!(j_1 - m_1)!(j_2 - m_2)!}\right]^{1/2}$$
$$\times \langle j_1 j_1 j_2 j - j_1 \mid j_1 j_2 j j\rangle \quad (m_1 + m_2 = j) \tag{3.4.44}$$

这个式子告诉我们确有 $\langle j_1 j_1 j_2 j - j_1 \mid j_1 j_2 j j\rangle \neq 0$。因为,否则按照式(3.4.44),耦合基矢 $\mid j_1 j_2 j j\rangle$ 的全部向非耦合基的展开系数将都是零。这样我们可以规定 $\mid j_1 j_2 j j\rangle$ 的相因子,使

$$\langle j_1 j_1 j_2 j - j_1 \mid j_1 j_2 j j\rangle > 0 \tag{3.4.45}$$

正是这一位相约定，破坏了 $\boldsymbol{J}^{(1)}$ 与 $\boldsymbol{J}^{(2)}$ 地位的对称性，也破坏了 (m_1, m_2, m) 与 $(-m_1, -m_2, -m)$ 的对称性。

式(3.4.44)对所有使 $j-j_2 \leq m_1 \leq j_1$ 成立的 m_1 都成立。它给出了 $|j_1 j_2 jj\rangle$ 的所有非零展开系数与 $\langle j_1 j_1 j_2 j-j_1 | j_1 j_2 jj\rangle$ 的比值。而 $|j_1 j_2 jj\rangle$ 的所有展开系数的模平方和应为 1(幺正性条件(3.4.25))，这时我们可以确定 $|\langle j_1 j_1 j_2 j-j_1 | j_1 j_2 jj\rangle|$，由(3.4.45)，我们可以得到 $\langle j_1 j_1 j_2 j-j_1 | j_1 j_2 jj\rangle$，这是一个正实数。再由(3.4.44)我们得到全体 $\langle j_1 m_1 j_2 m_2 | j_1 j_2 jj\rangle$，它们都是实数。这样做的结果我们得到

$$\langle j_1 j_1 j_2 j-j_1 | j_1 j_2 jj\rangle = \left[\frac{(2j_1)!(-j_1+j_2+j)!}{(j_1-j_2-j)!}\right]^{1/2}\left[\sum_{m_1+m_2=j}\frac{(j_1+m_1)!(j_2+m_2)!}{(j_1-m_1)!(j_2-m_2)!}\right]^{1/2} \tag{3.4.46}$$

利用恒等式(见附录)

$$\sum_{\sigma}\frac{(a-\sigma)!(b+\sigma)!}{(c+d)!(d-\sigma)!} = \frac{(a+b+1)!(b-c)!(a-d)!}{(c+d)!(a+b-c-d+1)!} \tag{3.4.47}$$

得

$$\langle j_1 j_1 j_2 j-j_1 | j_1 j_2 jj\rangle = \left[\frac{(2j_1)!(2j+1)!}{(j_1+j_2+j+1)!(j_1-j_2+j)!}\right]^{1/2} \tag{3.4.48}$$

由(3.4.44)得

$$\langle j_1 m_1 j_2 m_2 | j_1 j_2 jj\rangle = (-1)^{j_1-m_1}\left[\frac{(2j+1)!(j_1+j_2-j)!}{(j_1+j_2+j+1)!(-j_1+j_2+j)!(j_1-j_2+j)!}\right]^{1/2} \times\left[\frac{(j_1+m_1)!(j_2+m_2)!}{(j_1-m_1)!(j_2-m_2)!}\right]^{1/2},(m_1+m_2=j) \tag{3.4.49}$$

至此我们得到了全体 $m=j$ 的 $C.-G.$ 系数，它们都是实数。利用递推关系(3.4.44)，我们可以求得全体 $C.-G.$ 系数(由式(3.4.44))可知它们全是实数)。为此，把式(3.4.41)重新写作

$$\begin{aligned}&\sqrt{(j-m)(j+m+1)}\langle j_1 m_1 j_2 m-m_1 | j_1 j_2 jm\rangle\\ =&\sqrt{(j_1-m_1)(j_1+m_1+1)}\langle j_1 m_1+1 j_2 m-m_1 | j_1 j_2 jm+1\rangle\\ &+\sqrt{(j_2-m+m_1)(j_2+m-m_1+1)}\langle j_1 m_1 j_2 m-m_1+1 | j_1 j_2 jm+1\rangle\end{aligned} \tag{3.4.50}$$

这个递推关系把 j_1, j_2, j 值相同各 $C.-G.$ 系数联系起来，由于 $m_1+m_2=m$ 可以把(3.4.50)看作 j_1, j_2, j 值固定，m_1, m 取各种可能的值的各 $C.-G.$ 系数间的关系。定义函数

$$g(m_1,m)=(-1)^{m_1+m}\left[\frac{(j_1+m_1)!(j_2+m-m_1)!(j-m)!}{(j_1-m_1)!(j_2-m+m_1)!(j+m)!}\right]^{1/2} \tag{3.4.51}$$

以 $(-1)^{m_1+m}\left[\frac{(j_1+m_1)!\ (j_2+m-m_1)!\ (j-m-1)!}{(j_1-m_1)!\ (j_2-m+m_1)!\ (j+m+1)!}\right]^{1/2}$ 乘(3.4.50)得

$$\begin{aligned} &g(m_1,m)\langle j_1m_1j_2m-m_1\mid j_1j_2jm\rangle \\ &=g(m_1+1,m+1)\langle j_1m_1+1j_2m-m_1\mid j_1j_2jm+1\rangle \\ &\quad -g(m_1,m+1)\langle j_1m_1j_2m+1-m_1\mid j_1j_2jm+1\rangle \end{aligned} \tag{3.4.52}$$

用差分记法可以写成

$$\begin{aligned} &g(m_1,m)\langle j_1m_1j_2m-m_1\mid j_1j_2jm\rangle \\ &=\underset{m_1}{\Delta}[g(m_1,m+1)\langle j_1m_1j_2m+1-m_1\mid j_1j_2jm+1\rangle] \\ &=\underset{m_1}{\Delta^{j-m}}[g(m_1,j)\langle j_1m_1j_2j-m_1\mid j_1j_2jj\rangle] \end{aligned} \tag{3.4.53}$$

利用差分公式

$$\underset{x}{\Delta^n}\mathrm{d}(x)=\sum_{v=0}^{n}(-1)^{n+v}\binom{n}{v}f(x+v) \tag{3.4.54}$$

归纳可以得到

$$\begin{aligned} &\langle j_1m_1j_2m-m_1\mid j_1j_2jm\rangle \\ &=\frac{(-1)^{j-m}}{g(m_1,M)}\sum_{v=0}^{j-m}(-1)^v g(m_1+v,j)\langle j_1m_1+vj_2j-m_1-v\mid j_1j_2jj\rangle \\ &=(-1)^{j_1-m_1}\left[\frac{(2j+1)(j_1+j_2-j)!(j-m)!(j+m)!}{(j_1+j_2+j+1)!(-j_1+j_2+j)!(j_1-j_2+j)!}\right. \\ &\quad\left.\times\frac{(j_1-m_1)!(j_2-m+m_1)!}{(j_1+m_1)!(j_2+m-m_1)!}\right]^{1/2} \\ &\quad\times\sum_v(-1)^v\frac{(j_1+m_1+v)!}{v!(j-m-v)!}\times\frac{(j_2+j-m_1-v)!}{(j_1-m_1-v)!(j_2-j+m_1+v)!} \end{aligned} \tag{3.4.55}$$

式中求和指标 v 取遍使各阶乘因子有意义的值。式(3.4.55)是 $C.-G.$ 系数的一般公

式。它可以化成更对称的形式,利用下面两个公式(见附录):

$$\sum_{\rho}[\rho!(m-v+\rho)!(n-\rho)!(v-\rho)!]^{-1}=\frac{(n+m)!}{n!m!v!(n+m-v)!},$$

$$\sum_{\rho}(-1)^{\rho}\frac{(p-\rho)!}{\rho!(n-\rho)!(v-\rho)!}=\frac{(p-n)!(p-v)!}{n!v!(p-n-v)!},\text{若 } n\geq 0,v\geq 0,p\geq n+v \tag{3.4.56}$$

(3.4.55)式右边的求和因子

$$\begin{aligned}&\sum_{v}(-1)^{j_1-m_1+v}\frac{(j_1+m_1+v)!(j_2+j-m_1-v)!}{v!(j-m-v)!(j_1-m_1-v)!(j_2-j+m_1+v)!}\\&=\sum_{v}\Big[(-1)^{j_1-m_1+v}\frac{(j_1+m_1+v)!}{v!(j_2-j+m_1+v)!}\sum_{\mu}\frac{(j_2+m-m_1)!(-j_1+j_2+j)!}{\mu!(j_1-j_2-m-v+\mu)!}\\&\times\frac{1}{(j_2+m-m_1-\mu)!(-j_1+j_2+j-\mu)!}\Big]\\&=\sum_{\mu}\frac{(-j_1+j_2+j)!(j_2+m-m_1)!}{\mu!(-j_1+j_2+j-\mu)!(j_2+m-m_1-\mu)!}\\&\times\sum_{v}(-1)^{j_1-m_1+v}\frac{(j_1+m_1+v)!}{v!(j_2-j+m_1+v)!(j_1-j_2-m-v+\mu)!}\end{aligned} \tag{3.4.57}$$

令

$$\rho=j-j_2-m+\mu-v \tag{3.4.58}$$

式(3.4.57)最右边的一个求和式可写作

$$\begin{aligned}&(-1)^{j_2+m-m_1-\mu}\sum_{\rho}(-1)^{\rho}\frac{(2j_1-j_2-m+m_1+\mu-\rho)!}{\rho!(j_1-j_2-m+\mu-\rho)!(j_1-j+m_1-m+\mu-\rho)!}\\&=(-1)^{j_2+m-m_1-\mu}\frac{(j_1+m_1)!(j_1-j_2+j)!}{(j_1-j_2-m+\mu)!(j_1+m_1-j-m+\mu)!(j+m-\mu)!}\end{aligned}$$

代入式(3.4.57),并令

$$\sigma=j_2+m-m_1-\mu \tag{3.4.59}$$

最后得

$$\begin{aligned}&\langle j_1m_1j_2m_2\mid j_1j_2jm\rangle\\&=\sigma_{m_1+m_2,m}\times\left[\frac{(2j+1)(j_1+j_2-j)!(-j_1+j_2+j)!(j_1-j_2+j)!}{(j_1+j_2+j+1)!}\right]^{1/2}\end{aligned}$$

$$\times[(j+m)!\times(j-m)!(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!]^{1/2}$$

$$\sum_{\sigma}(-1)^{\sigma}[\sigma!(j_1+j_2-j-\sigma)!(j_1-m_1-\sigma)!(j_2+m_2-\sigma)!$$

$$(j-j_1-m_2+\sigma)!(j-j_2+m_1+\sigma)!]^{-1} \tag{3.4.60}$$

$C.-G.$ 系数的这个一般表达式具有较高的对称性。例如，由式(3.4.60)可以直接看出下面的对称关系：

$$\langle j_1m_1j_2m_2\mid j_1j_2jm\rangle=\langle j_2-m_2j_1-m_1\mid j_1j_2j-m\rangle \tag{3.4.61}$$

在理论讨论时(相对于数值计算而言)，$C.-G.$ 系数的对称性所起的作用更大。

Ⅲ. $C.-G.$ 系数的对称性关系与 $3-j$ 符号

在两个角动量的耦合问题中，和角动量算符的定义是

$$\boldsymbol{J}=\boldsymbol{J}(1)+\boldsymbol{J}(2)=\boldsymbol{J}(2)+\boldsymbol{J}(1)$$

其中$[J_\alpha(1),J_\beta(2)]=0=[J_\beta(1),J_\alpha(2)]$。可以看到 $\boldsymbol{J}(1)$和 $\boldsymbol{J}(2)$处于完全相同的地位。我们可以用以上升下降算符得到$\langle j_2m_2j_1m_1\mid j_1j_2jm\rangle$满足$\langle j_1m_1j_2m_2\mid j_1j_2jm\rangle$所满足的同样的递推关系(3.4.40)和(3.4.41)。对固定的 j_1,j_2j 值，齐次线性代数方程组(3.4.40)和(3.4.41)把解确定到一个共同的常数倍

$$C(j_1j_2j):\langle j_2m_2j_1m_1\mid j_1j_2jm\rangle=C(j_1j_2j)\langle j_1m_1j_2jm_2\mid j_1j_2jm\rangle \tag{3.4.62}$$

由幺正性条件(3.4.25)可知

$$\mid C(j_1j_2j)\mid=1 \tag{3.4.63}$$

当 $m_1=j_1,m_2=j-j_1,m=j$ 时，(3.4.62)为

$$\langle j_2j-j_1j_1j_1\mid j_2j_1jj\rangle=C(j_1j_2j)\langle j_1j_1j_2j-j_1\mid j_1j_2jj\rangle \tag{3.4.64}$$

由(3.4.44)和(3.4.45)可知$\langle j_2j-j_1j_1j_1\mid j_2j_1jj\rangle$与$(-1)^{j_1+j_2-j}$同号。故

$$\langle j_2j-j_1j_1j_1\mid j_2j_1jj\rangle=(-1)^{j_1+j_2-j}\langle j_1j_1j_2j-j_1\mid j_1j_2jj\rangle \tag{3.4.65}$$

于是

$$\langle j_2m_2j_1m_1\mid j_1j_2jm\rangle=(-1)^{j_1+j_2-j}\langle j_1m_1j_2m_1\mid j_1j_2jm\rangle \tag{3.4.66}$$

$(j_1m_1),(j_2m_2)$的这种不对称是由位相约定(3.4.45)引入的。由(3.4.66)和(3.4.61)可知

$$\langle j_1 - m_1 j_2 - m_1 \mid j_1 j_2 j - m\rangle = (-1)^{j_1+j_2-j}\langle j_1 m_1 j_2 m_1 \mid j_1 j_2 j m\rangle \quad (3.4.67)$$

这反映了 z 方向与 $-z$ 方向有等同，不对称性也是由位相约定(3.4.45)引入的(它部分地破坏了 $m, -m$ 间的对称性)。

利用(3.4.60)式，我们来证明另一对称性：

$$\langle j_1 m_1 j - m \mid j_1 j j_2 m_1 - m\rangle = (-1)^{j_1-m_1}\left[\frac{2j_2+1}{2j+1}\right]^{1/2}\langle j_1 m_1 j_2 m - m_1 \mid j_1 j_2 j m\rangle \quad (3.4.68)$$

把(3.4.60)式用于上式右边，作变量变换

$$\sigma' = j_1 - m_1 - \sigma \quad (3.4.69)$$

所得的表示式正是(3.4.68)左边。从对称关系(3.4.66)，(3.4.67)和(3.4.68)，我们可以得到一系列的派生的对称关系。例如：

$$\langle j - m j_2 m_2 \mid j j_2 j_1 - m_1\rangle = (-1)^{j_2+m_2}\left[\frac{2j_1+1}{2j+1}\right]^{1/2}\langle j_1 m_1 j_2 m - m_1 \mid j_1 j_2 j m\rangle \quad (3.4.70)$$

$$\langle j m j_1 - m_1 \mid j j_1 j_2 m_2\rangle = (-1)^{j_1-m_1}\left[\frac{2j_2+1}{2j+1}\right]^{1/2}\langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j m\rangle \quad (3.4.71)$$

$$\langle j_2 - m_2 j m \mid j_2 j j_1 m_1\rangle = (-1)^{j_2-m_2}\left[\frac{2j_1+1}{2j+1}\right]^{1/2}\langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j m\rangle \quad (3.4.72)$$

以上这些对称关系请读者证明之。

由 $C.-G.$ 系数的对称关系，我们可以得出一些常遇到的 $C.-G.$ 系数的性质。例如：当 j_1, j_2 和 j 都是整数(记作 l_1, l_2 和 l)，m_1, m_2 和 m 全为零时，式(3.4.67)给出

$$\langle l_1 0 l_2 0 \mid l_1 l_2 l 0\rangle = (-1)^{l_1+l_2-l}\langle l_1 0 l_2 0 \mid l_1 l_2 l 0\rangle = (-1)^{l_1+l_2-l}\langle l_1 0 l_2 0 \mid l_1 l_2 l 0\rangle \quad (3.4.73)$$

可见

$$\langle l_1 0 l_2 0 \mid l_1 l_2 l 0\rangle = 0, \text{当 } l_1 + l_2 - l \text{ 为奇数} \quad (3.4.74)$$

Wigner 引入了具有更高对称性的 $3-j$ 符号，它与 $C.-G.$ 系数的关系为

$$\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{pmatrix} = (-1)^{j_1-j_2+m}(2j+1)^{1/2}\langle j_1 m_2 j_2 m_2 \mid j_1 j_2 j m\rangle \quad (3.4.75)$$

(请注意 3 – j 符号中各个元素的意义!)由 $C.-G.$ 系数的对称关系知,3 – j 符号满足如下对称关系:

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = \begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix} \tag{3.4.76}$$

此即在列的偶数次交换下,3 – j 符号值不变。

$$(-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = \begin{pmatrix} j_1 & j_3 & j_2 \\ m_1 & m_3 & m_2 \end{pmatrix} = \begin{pmatrix} j_3 & j_2 & j_1 \\ m_3 & m_2 & m_1 \end{pmatrix} \tag{3.4.77}$$

在列的奇数次交换下,3 – j 符号得一因子 $(-1)^{j_1+j_2+j}$ 而与(3.4.67)相应的是

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} \tag{3.4.78}$$

与 $C.-G.$ 系数幺正性质相应的是 3 – j 符号的正交归一化条件:对满足三角形关系的 j_1, j_2, j_3

$$\sum_{j_3 m_3}(2j_3+1)\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1' & m_2' & m_3' \end{pmatrix} = \delta_{m_1 m_1'}\delta_{m_2 m_2'} \tag{3.4.79}$$

和

$$\sum_{m_1 m_2}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3' \end{pmatrix} = (2j_3+1)\delta_{j_3 j_3'}\delta_{m_3 m_3'} \tag{3.4.80}$$

在量子物理问题的理论分析(与数值计算相对而言)中,除了常利用 $C.-G.$ 系数(3 – j 符号)对称关系外,还经常用到一些特殊的 $C.-G.$ 系数值。由下式

$$\langle j_1 j_1 j_2 j_2 \mid j_1 j_2 j_1 + j_2 j_1 + j_2 \rangle = 1 \tag{3.4.81}$$

(见式(3.4.32),(3.4.33))和对称关系(3.4.67)可得

$$\langle j_1 - j_1 j_2 - j_2 \mid j_1 j_2 j_1 + j_2 - j_1 - j_2 \rangle = 1 \tag{3.4.82}$$

在 $j_2 = 0$ 时,式(3.4.32)变作

$$\langle j_1 j_1 00 \mid j_1 0 j_1 j_1 \rangle = 1 \tag{3.4.83}$$

以 $J_- = J_-(1) \oplus J_-(2)$ 作用于

$$\mid j_1 0 j_1 j_1 \rangle_c = \mid j_1 j_1 00 \rangle_n \tag{3.4.84}$$

(下标 c 表示耦合基矢,n 表示非耦合基矢)

可得

$$|j_1 0 j_1 m_1\rangle_c = |j_1 m_1 00\rangle_n \tag{3.4.85}$$

从而

$$\langle j_1 m_1 00 | j_1 0 j_3 m_3\rangle = \delta_{j_1 j_3}\delta_{m_1 m_3} \tag{3.4.86}$$

利用对称关系(3.4.68),由上式可得:

$$\langle j_1 m_1 j_1 - m_1 | j_1 j_1 00\rangle = (-1)^{j_1 - m_1}\frac{1}{\sqrt{2j_1+1}} \tag{3.4.87}$$

上面的式(3.4.81),(3.4.82),(3.4.83),(3.4.86)和(3.4.87),是常用的 $C.-G.$ 系数数值,读者应熟练掌握。

Ⅳ. $C.-G.$ 系数与 $D-$ 函数

我们在数学准备一章中强调指出过,线性空间中线性算符的矩阵与基的选择有关。当基经受一非奇异线性变换:

$$[e_1, e_2, \cdots, e_n] \mapsto [e_1', \cdots, e_n'] = [e_1, e_2, \cdots, e_n]\begin{bmatrix} T_1^1 & T_2^1 & \cdots & T_n^1 \\ T_1^2 & T_2^2 & \cdots & T_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ T_1^n & T_2^n & \cdots & T_n^n \end{bmatrix}$$
$$\equiv [e_1, e_2, \cdots, e_n][T] \tag{3.4.88}$$

同一线性算符 A 的矩阵要相应地经受相似变换:

$$[A]' = [T]^{-1}[A][T]$$

当空间是内积空间,且 T 是把标准正交基变成标准正交基的变换是幺正变换时,上式可以写作

$$[A]' = [T]^{+}[A][T] \tag{3.4.89}$$

让我们考虑转动算符

$$D(\boldsymbol{\theta}) = e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}} = e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}^{(1)}}e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}^{(2)}} = D^{(1)}(\boldsymbol{\theta})D^{(2)}(\boldsymbol{\theta}) \tag{3.4.90}$$

到不变子空间(3.4.20)

$$L^{(1)} \otimes L^{(2)} = L = \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} L_j \tag{3.4.91}$$

的限制。它在非耦合基下的矩阵与在耦合基下的矩阵应满足关系(3.4.89)。这时幺正变换的矩阵元就是 $C.-G.$ 系数

$$T^{j}_{j_1 j_2 jm}{}^{m_1 j_2 m_2} = \langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j m \rangle \tag{3.4.92}$$

而转动算符 $D(\boldsymbol{\theta})$ 在两种基下的矩阵元分别是 $D-$ 函数和两个 $D-$ 函数的乘积：

$$\langle j_1 j_2 j' m' \mid D(\boldsymbol{\theta}) \mid j_1 j_2 j m \rangle = \delta_{j'j} D^{j}_{m'm}(\boldsymbol{\theta}) \tag{3.4.93}$$

$$\langle j_1 m'_1 j_2 m'_2 \mid D^{(1)}(\boldsymbol{\theta}) D^{(2)}(\boldsymbol{\theta}) \mid j_1 m_1 j_2 m_2 \rangle = D^{j}_{m'_1 m_1}(\boldsymbol{\theta}) D^{j_2}_{m'_2 m_2}(\boldsymbol{\theta}) \tag{3.4.94}$$

矩阵等式(3.4.89)可以写成

$$\delta_{j'j} D^{j}_{m'm}(\boldsymbol{\theta}) = \sum_{m_1 m_2 m'_1 m'_2} \langle j_1 m'_1 j_2 m'_2 \mid j_1 j_2 j' m' \rangle D^{j}_{m'_1 m_1}(\boldsymbol{\theta}) D^{j_2}_{m'_2 m_2}(\boldsymbol{\theta}) \langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j m \rangle \tag{3.4.95}$$

(3.4.89)的逆变换可以写成

$$\begin{aligned} D^{j}_{m'_1 m_1}(\boldsymbol{\theta}) D^{j_2}_{m'_2 m_2}(\boldsymbol{\theta}) &= \sum_{j' m' j m} \langle j_1 m'_1 j_2 m'_2 \mid j_1 j_2 j' m' \rangle \delta_{j'j} D^{j}_{m'm}(\boldsymbol{\theta}) \langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j m \rangle \\ &= \sum_{j} \langle j_1 m'_1 j_2 m'_2 \mid j_1 j_2 j m'_1 + m'_2 \rangle D^{j}_{m'_1 + m'_2, m_1 + m_2}(\boldsymbol{\theta}) \langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j m_1 + m_2 \rangle \end{aligned} \tag{3.4.96}$$

回忆 $D-$ 函数与球谐函数 Y_{lm} 的关系

$$D^{l}_{m0}(\alpha\beta\gamma) = i^{l} \sqrt{\frac{4\pi}{2l+1}} Y^{*}_{lm}(\beta, \alpha) \tag{3.4.97}$$

由(3.4.96)式可得

$$\begin{aligned} Y_{l_1 m_1}(\theta\varphi) Y_{l_2 m_2}(\theta\varphi) &= \sum_{l} i^{l_1 + l_2 - l} \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}} \\ &\times Y_{l m_1 + m_2}(\theta\varphi) \langle l_1 m_1 l_2 m_2 \mid l_1 l_2 l m_1 + m_2 \rangle \langle l_1 0 l_2 0 \mid l_1 l_2 l 0 \rangle \end{aligned} \tag{3.4.98}$$

回忆球谐函数的正交归一化性质。利用上式可得到三个球谐函数乘积的积分公式：

$$\begin{aligned} &\iint Y^{*}_{lm}(\theta\varphi) Y_{l_1 m_1}(\theta\varphi) Y_{l_2 m_2}(\theta\varphi) \sin\theta \mathrm{d}\theta \mathrm{d}\varphi \\ &= i^{l_1 + l_2 - l} \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}} \langle l_1 m_1 l_2 m_2 \mid l_1 l_2 l m \rangle \langle l_1 0 l_2 0 \mid l_1 l_2 l 0 \rangle \end{aligned} \tag{3.4.99}$$

Ⅴ. 两个对易的角动量的相加

上面讨论两个角动量的耦合时，我们假定量子体系是由两个子体系构成的。总体系的位形空间自然是两个子体系的位形空间的笛卡尔积 $C = C(1) \times C(2)$。从而总体系的态矢量空间是两个子体系的态矢量空间的张量积 $S = S(1) \otimes S(2)$；总体系的角动量是两个子体系的角动量之和 $\boldsymbol{J} = \boldsymbol{J}(1) \otimes E(2) + E(1) \otimes \boldsymbol{J}(2) = \boldsymbol{J}(1) + \boldsymbol{J}(2)$。由于作用在不同子体系的态矢量上，$\boldsymbol{J}(1)$，$\boldsymbol{J}(2)$互相对易$[J_\alpha(1,) J_\beta(2)] = 0$。关于两个角动量相加的理论是在这种前提下展开的。但是在量子力学中有时(例如在讨论氢原子动力学对称性时)会遇到这种情况：定义在量子体系态矢量空间 S 上有两个角动量算符，我们只知道它们互相对易，并不知道它们是否作用在不同的子体系的态矢量上。于是我们要问：'互相对易的两个角动量算符是否一定作用在不同子体系的态矢量上?'我们将会看到，答案是'不一定'。那么，'两个子体系角动量相加的理论，对两个不对易的角动量相加是否仍有效'?

为回答第一个问题，考虑电子－质子体系的自旋态空间 $\sum = \sum_e \otimes \sum_p$。电子自旋角动量算符 $\boldsymbol{S}_e$ 只作用在电子自旋波函数上，而质子自旋角动量算符 $\boldsymbol{S}_p$ 只作用在质子自旋波函数上。现构造两个线性算符 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，使它们在纠缠态基

$$\left(\frac{|++>+|-->}{\sqrt{2}}, \frac{|++>-|-->}{\sqrt{2}}, \frac{|+->+<-+|>}{\sqrt{2}}, \frac{|+->-|-+>}{\sqrt{2}}\right)$$

下的矩阵分别与 $\boldsymbol{S}_e$ 和 $\boldsymbol{S}_p$ 在非纠缠态基$(|++>, |+->, |-+>, |-->)$下的矩阵相同。$\boldsymbol{A}$、$\boldsymbol{B}$ 显然是两个互相对易的角动量算符，但它们并不分别作用在两个子体系的波函数上。

下面的命题将帮助我们回答第二个问题。

命题 设 $\boldsymbol{A}, \boldsymbol{B}$ 是 *Hilbert* 空间 S 上的两个相互对易的角动量算符。则 S 可以分解成两两正交的$\{\boldsymbol{A}, \boldsymbol{B}\}$－不可约不变子空间的直和 $S = \bigoplus_\sigma S(\sigma)$，其中的每个子空间都有正交归一基$\{u_{\alpha\beta}(\sigma) | -j_1 \leq \alpha \leq j_1, -j_2 \leq \beta \leq j_2\}$，使 $\boldsymbol{A}|_{S(\sigma)}, \boldsymbol{B}|_{S(\sigma)}$在此基下的矩阵分别是$[\boldsymbol{J}]_{j_1} \otimes E_{2j_2+1}, E_{2j_2+1} \otimes [\boldsymbol{J}]_{j_2}$，，其中$[\boldsymbol{J}]_j$ 为角动量在 j 权不可约表示空间标准基下的矩阵(3.2.31)、(3.2.34)和(3.2.35)。

事实上，由单个角动量的理论，S 可以分解为两两正交的 $\boldsymbol{B}$－不可约不变子空间的直和 $S = \bigoplus_\sigma S(\sigma)$。我们把这些不变子空间按维数由小到大的顺序编号。以 n_j 记

角量子数为 j 的 $\boldsymbol{B}$ – 不可约不变子空间的个数，则第 α 个角量子数为 j 的 $\boldsymbol{B}$ – 不可约不变子空间的编号为 $\sigma = \alpha + \sum_{k<j} n_k$。我们也可以用角量子数 j 和数 $\alpha(1 \leq \alpha \leq n_j)$ 来标记 $\boldsymbol{B}$ – 不可约不变子空间 $S(\sigma) = S_\alpha^j$。于是 $S = \bigoplus_j \bigoplus_{\alpha=1}^{n_j} S_\alpha^j$。选定每个 S_α^j 的标准基

$$\{\tilde{e}_{\alpha m}^j \mid -j \leq m \leq j\} \tag{3.4.100}$$

按 $\{j\alpha\}$ – 字典顺序把所有这些标准基合在一起，我们得到整个空间 S 的一个标准正交基。角动量算符 $\boldsymbol{B}$ 在这个基下的矩阵为拟对角的

$$\begin{aligned} \|\boldsymbol{B}\| &= \begin{bmatrix} [\boldsymbol{J}]_{j_1} & & & & & & \\ & \ddots & & 0 & & & \\ & & [\boldsymbol{J}]_{j_1} & & & & \\ & & & [\boldsymbol{J}]_{j_2} & & & \\ & & & & \ddots & & \\ & 0 & & & & [\boldsymbol{J}]_{j_2} & \\ & & & & & & [\boldsymbol{J}]_{j_3} \\ & & & & & & & \ddots \end{bmatrix} \\ &= E_{n_1} \otimes [\boldsymbol{J}]_{j_1} + E_{n_2} \otimes [\boldsymbol{J}]_{j_2} \oplus \cdots \end{aligned} \tag{3.4.101}$$

把角动量算符 $\boldsymbol{A}$ 的矩阵写成相应的分块矩阵形式

$$\|\boldsymbol{A}\| = \begin{bmatrix} [\boldsymbol{A}]_{11}^{j_1 j_1} & \cdots & [\boldsymbol{A}]_{1n_1}^{j_1 j_1} & [\boldsymbol{A}]_{11}^{j_1 j_2} & \cdots & [\boldsymbol{A}]_{1n_2}^{j_1 j_2} & [\boldsymbol{A}]_{11}^{j_1 j_3} & \cdots \\ \vdots & \ddots & \vdots & & & & & \\ [\boldsymbol{A}]_{n_1 1}^{j_1 j_1} & & [\boldsymbol{A}]_{n_1 n_1}^{j_1 j_1} & [\boldsymbol{A}]_{n_1 1}^{j_1 j_2} & & [\boldsymbol{A}]_{n_1 n_2}^{j_1 j_2} & [\boldsymbol{A}]_{n_1 1}^{j_1 j_3} & \cdots \\ [\boldsymbol{A}]_{11}^{j_2 j_1} & & [\boldsymbol{A}]_{1n_1}^{j_2 j_1} & [\boldsymbol{A}]_{11}^{j_2 j_2} & & [\boldsymbol{A}]_{1n_2}^{j_2 j_2} & [\boldsymbol{A}]_{11}^{j_2 j_3} & \\ \vdots & & & & \ddots & & & \\ [\boldsymbol{A}]_{n_2 1}^{j_2 j_1} & & [\boldsymbol{A}]_{n_2 n_1}^{j_2 j_1} & [\boldsymbol{A}]_{n_2 1}^{j_2 j_2} & & [\boldsymbol{A}]_{n_2 n_2}^{j_2 j_2} & [\boldsymbol{A}]_{n_2 1}^{j_2 j_3} & \cdots \\ [\boldsymbol{A}]_{11}^{j_3 j_1} & & [\boldsymbol{A}]_{1n_1}^{j_3 j_1} & [\boldsymbol{A}]_{11}^{j_3 j_2} & & [\boldsymbol{A}]_{1n_2}^{j_3 j_2} & [\boldsymbol{A}]_{11}^{j_3 j_3} & \\ \vdots & & \vdots & & & & & \ddots \end{bmatrix}$$

上面两个角动量算符的矩阵互相对易，利用 Schur 引理可知 $\boldsymbol{A}$ 的矩阵为

$$\|A'\| =$$

$$\begin{bmatrix}
(A^1)_{11} & 0 & \cdots & (A^1)_{1n_1} & 0 & & & & & & & \\
 & \ddots & & \cdots & & \ddots & & & & & & \\
0 & (A^1)_{11} & \cdots & 0 & (A^1)_{1n_1} & & & & & & & \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & & 0 & & \\
(A^1)_{n_1 1} & 0 & \cdots & (A^1)_{n_1 n_1} & 0 & & & & & & & \\
 & \ddots & & \cdots & & \ddots & & & & & & \\
0 & (A^1)_{n_1 1} & \cdots & 0 & (A^1)_{n_1 n_1} & & & & & & & \\
 & & & & & (A^2)_{11} & 0 & \cdots & (A^2)_{1n_2} & 0 & & \\
 & & & & & & \ddots & & \cdots & & \ddots & \\
 & & & & & 0 & (A^2)_{11} & \cdots & 0 & (A^2)_{1n_2} & & \\
 & & 0 & & & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
 & & & & & (A^2)_{n_2 1} & 0 & \cdots & (A^2)_{n_2 n_2} & 0 & & \\
 & & & & & & \ddots & & \cdots & & \ddots & \\
 & & & & & 0 & (A^2)_{n_2 1} & \cdots & 0 & (A^2)_{n_2 n_2} & & \\
 & & & & & & & & & & (A^3)_{11} & \\
 & & & & & & & & & & & \ddots
\end{bmatrix}$$

$$= [A^1] \otimes E_{2j_1+1} \oplus [A^2] \otimes E_{2j_2+1} \oplus \cdots \tag{3.4.102}$$

式中的$[A^p]$是 $n_p \times n_p$ 矩阵。利用 A 是角动量算符，可得$[[(A^p)_\alpha],[(A^p)_\beta]] = i\hbar\varepsilon_{\alpha\beta\gamma}[(A^p)_\gamma]$。由单个角动量理论知，存在 $n_p \times n_p$ 酉矩阵 U_p，使 $U_p[A^p]U_p^{-1}$ 是拟对角矩阵，其每个对角块是角动量算符在不可约不变子空间的标准基下的矩阵 $[J]_j$ (3.2.31)、(3.2.34)和(3.2.35)

$$U_p[A^p]U_p^{-1} = \begin{bmatrix} [J]_{j'_p} & & 0 \\ & [J]_{j''_p} & \\ & & \ddots \\ 0 & & \ddots \end{bmatrix} = [J]_{j'_p} \oplus [J]_{j''_p} \oplus \cdots \tag{3.4.103}$$

于是存在 $n_p(2j_p+1)\times n_p(2j_p+1)$酉矩阵

$$U_p \otimes E_{2j_p+1} \equiv W_p \tag{3.4.104}$$

使

$$W_p[\boldsymbol{A}^p]W_p^{-1} = U_p[\boldsymbol{A}^p]U_p^{-1}\otimes E_{2j_p+1} = [\boldsymbol{J}]_{j_p'}\otimes E_{2j_p+1}\oplus[\boldsymbol{J}]_{j_p''}\otimes E_{2j_p+1}\oplus\cdots \tag{3.4.105}$$

其中 $n_p=(2j_p'+1)+(2j_p''+1)+\cdots$而 $E_{n_p}\otimes[\boldsymbol{J}]_{j_p}$在此酉变换下不变

$$W_p(E_{n_p}\otimes[\boldsymbol{J}]_{j_p})W_p^{-1} = E_{n_p}\otimes[\boldsymbol{J}]_{j_p} = E_{2j_p'+1}\otimes[\boldsymbol{J}]_{j_p}\oplus E_{2j_p''+1}\otimes[\boldsymbol{J}]_{j_p}\oplus\cdots \tag{3.4.106}$$

这个酉变换对应的基的变换为

$$\tilde{e}^{j}_{dm} = \sum_{\beta=1}^{n_p} e^{j_p}_{\beta m}(U_p)^{\beta}_{\alpha} \tag{3.4.107}$$

于是在 *Hilbert* 空间 S 的正交归一基

$$\{\tilde{e}^{j}_{dm}\mid p\geq 1, 1\leq\alpha\leq n_p, -j_p\leq m\leq j_p\} \tag{3.4.108}$$

下,角动量 $\boldsymbol{A},\boldsymbol{B}$ 的矩阵为

$$\begin{aligned}[\boldsymbol{A}] = {} & [\boldsymbol{J}]_{j_1'}\otimes E_{2j_1+1}\oplus[\boldsymbol{J}]_{j_1''}\otimes E_{2j_1+1}\oplus\cdots\oplus[\boldsymbol{J}]_{j_2'}\otimes E_{2j_2+1}\oplus[\boldsymbol{J}]_{j_2''}\otimes E_{2j_2+1}\oplus\cdots\\ & \oplus[\boldsymbol{J}]_{j_p'}\otimes E_{2j_p+1}\oplus[\boldsymbol{J}]_{j_p''}\otimes E_{2j_p+1}\oplus\cdots\\ [\boldsymbol{B}] = {} & E_{2j_1'+1}\otimes[\boldsymbol{J}]_{j_1}\oplus E_{2j_1''+1}\otimes[\boldsymbol{J}]_{j_1}\oplus\cdots\oplus E_{2j_2'+1}\otimes[\boldsymbol{J}]_{j_2}\oplus E_{2j_2''}+1\otimes[\boldsymbol{J}]_{j_2}\oplus\cdots\\ & \oplus E_{2j_p'+1}\otimes[\boldsymbol{J}]_{j_p}\oplus E_{2j_p''+1}\otimes[\boldsymbol{J}]_{j_p}\oplus\cdots\end{aligned} \tag{3.4.109}$$

这两个等式右边的对应项是角动量 $\boldsymbol{A},\boldsymbol{B}$ 到同一个$\{\boldsymbol{A},\boldsymbol{B}\}$ – 不可约不变子空间上的限制。例如,两式右边第一项对应的$\{\boldsymbol{A},\boldsymbol{B}\}$ – 不可约不变子空间是由下面各基矢张成的:$\{e^{j}_{dm}\mid 1\leq\alpha\leq 2j_1'+1, -j_1\leq m\leq j_1\}$;而两式右边第二项对应的$\{\boldsymbol{A},\boldsymbol{B}\}$ – 不可约不变子空间是由下面各基矢张成的:$\{e^{j}_{dm}\mid 1+(2j_1'+1)\leq\alpha\leq(2j_1'+1)+(2j_1'+1), -j_1\leq m\leq j_1\}$,等等。命题证完。

由此命题,我们看到整个态矢量空间 S 可以分解成两两正交的$\{\boldsymbol{A},\boldsymbol{B}\}$ – 不可约不变子空间的直和。每个这样的子空间都有正交归一基$\{u_{\alpha\beta}(\sigma)\mid -j_1\leq\alpha\leq j_1, -j_2\leq\beta\leq j_2\}$,使 $\boldsymbol{A},\boldsymbol{B}$ 在此基下的矩阵分别是$[\boldsymbol{J}]_{j_1}\otimes E_{2j_2+1}, E_{2j_1+1}\otimes[\boldsymbol{J}]_{j_2}$。尽管 $u_{\alpha\beta}$

(σ)不一定是非纠缠态(两个子体系态矢量的张量积),但在推导 $C.-G.$ 系数过程中,只用到 $\boldsymbol{A},\boldsymbol{B}$ 对 $u_{\alpha\beta}(\sigma)$的作用,和 $\boldsymbol{A},\boldsymbol{B}$ 互相对易的性质。故我们可以断言,把 $\{u_{\alpha\beta}(\sigma)|-j_1\leq\alpha\leq j_1,-j_2\leq\beta\leq j_2\}$ 当非耦合基,用 $C.-G.$ 系数作线性组合一定可以得到和角动量 $\boldsymbol{A}+\boldsymbol{B}$ 的本征函数(耦合基)。

附录

牛顿二项式 $(x+y)^n=\sum_{r=0,1,2,\cdots}\binom{n}{r}x^{n-r}y^r$ 的系数

$$\binom{n}{r}=\frac{n(n-1)\cdots(n-r+1)}{r!},(r\text{ 是非负整数}) \tag{A-1}$$

当 n 是不小于 r 的整数时为

$$\binom{n}{r}=\frac{n!}{r!(n-r)!}, \tag{A-2}$$

当 n 是负整数时$(n=-v)$

$$\binom{n}{r}=(-1)^r\frac{(v+r-1)!}{r!(v-1)!}=(-1)^r\binom{v+r-1}{r} \tag{A-3}$$

比较等式

$$(x+y)^m(x+y)^n=(x+y)^{m+n} \tag{A-4}$$

两边同类项的系数可得二项式系数加法定理

$$\sum_{\mu}\binom{m}{\mu}\binom{n}{\sigma-\mu}\binom{m+n}{\sigma} \tag{A-5}$$

当 m,n 皆正整数时得

$$\sum_{\mu}[\mu!(m-\mu)!(\sigma-\mu)!(n-\sigma+\mu)!]^{-1}=\frac{(m+n)!}{m!n!\sigma!(m+n-\sigma)!} \tag{A-6}$$

当 m 是正整数,n 是负整数$(n=-v)(v>m)$有

$$\sum_{\mu}(-1)^{\mu}\binom{m}{\mu}=\binom{v+\sigma-\mu-1}{\sigma-\mu}\binom{v-m+\sigma-1}{\sigma} \tag{A-7}$$

引入新的参数 p 代替 $v(v>m)$

$$p=v+\sigma-1,(\therefore p\geq m+\sigma) \tag{A-8}$$

于是

$$\sum_{\mu}(-1)^{\mu}\frac{(p-\mu)!}{\mu!(m-\mu)!(\sigma-\mu)!}=\frac{(p-m)!(p-\sigma)!}{m!\sigma!(p-m-\sigma)!} \tag{A-9}$$

当 $m=-\mu, n=-v$(μ, v 为正整数)时,由(A-5)得

$$\sum_{\rho}\frac{(\mu+p-1)!(v+\sigma-p-1)!}{\rho!(\sigma-\rho)!}=\frac{(\mu-1)!(v-1)!(\mu+v+\sigma-1)!}{\sigma!(\mu+v-1)!} \tag{A-10}$$

引入新的参数和新求和变量 s,

$$\begin{aligned} s&=\rho-c,(\text{新求和变量})\\ a&=\mu+c-1\\ b&=v+\rho-c-1\\ d&=\sigma-c \end{aligned} \tag{A-11}$$

(A-10)变作

$$\sum_{s}\frac{(a+s)!(b-s)!}{(c+s)!(d-s)!}=\frac{(a+b+c)!(a-c)!(b-d)!}{(c+d)!(a+b-c-d+1)!} \tag{A-12}$$

(A-12)比(A-10)有更好的对称性。新参数 c 是多余的,可取任意值,问题是新求和变量 s 仅取遍使各阶乘因子有意义的值。

§3.5 不可约张量算符

Ⅰ.不可约张量算符的定义

在本章的§3.1中我们曾指出,当把实验室绕固定 O 点旋转时,体系的态矢量空间 S 按旋转群 $SO(3)$ 的一个幺正表示 $D: SO(3)\to GL(S)$ 变换,$\psi\mapsto D(g)\psi$, $\forall s\in SO(3)$;而 S 上线性算符的全体所成的线性空间 $EndS$,按 $SO(3)$ 群的酉表示 $\widetilde{D}: SO(3)\to GL(EndS)$ 变换

$$A\mapsto D(g)AD^{-1}(g)=\widetilde{D}(g)A \qquad \forall g\in SO(3) \tag{3.5.1}$$

S 按酉表示 D 分解成两两正交的 D – 不可约不变子空间的直和;而 $EndS$ 按酉表示 $\widetilde{D}$

分解成两两正交的 $\widetilde{D}$ – 不可约不变子空间的直和。每一个这种 $\widetilde{D}$ – 不可约不变子空间的标准基的基矢(是算符!)就是所谓的不可约张量算符。因而 S 上的每一线性算符 $A \in EndS$ 都可以写成不可约张量算符的线性组合。

不可约张量的定义如下:

设 k 是非负的整数或半整数, $\{T(k,q) \mid q = k, k-1, \cdots, -k\}$ 是量子体系的 $2k+1$ 个算符。若在空间转动下,它们按 $SO(3)$ 群的不可约表示 $D^{[k]}$ (k 权表示)变换:

$$\widetilde{D}(g)T(k,q) = D(g)T(k,g)D^{-1}(g) = \sum_{q'} T(k,q')D^{k}_{q'q}(g) \tag{3.5.2}$$

则称 $T(k,q)$ 为转动群 $SO(3)$ 的 k – 秩不可约张量算符(的 q 分量)。

让我们看下面的例子:

(ⅰ)粒子到固定旋转中心距离的算符

$$r = |\boldsymbol{r}| \tag{3.5.3}$$

作为作用在态矢量空间上的算符,它是个转动不变量:

$$D(g)rD(g^{-1}) = r \tag{3.5.4}$$

回忆

$$D^{0}_{00}(g) = 1 \tag{3.5.5}$$

是 0 权表示转动算符的唯一的矩阵元。故式(3.5.4) 表明 $r = |\boldsymbol{r}|$ 是 0 秩不可约张量算符。

(ⅱ)作为算符的球谐函数 $Y_{lm}(\theta,\varphi) = Y_{lm}(\boldsymbol{r})$ 对态矢量的作用如下:

$$Y_{lm}(\boldsymbol{r}):\psi(\boldsymbol{r}) \mid \rightarrow Y_{lm}(\boldsymbol{r})\psi(\boldsymbol{r}) \tag{3.5.6}$$

因而

$$D(g)Y_{lm}(\boldsymbol{r})D(g^{-1}) = Y_{lm}(R(g^{-1})\boldsymbol{r}) = \sum_{m'} Y_{lm'}D^{l}_{m'm}(g) \tag{3.5.7}$$

可见 $Y_{lm}(\theta,\varphi)$ 是 l 秩不可约张量。

(ⅲ)坐标算符 x, y, z 是位置矢量在笛尔坐标系中的分量。它们是一阶球谐函数的线性组合的 r 倍(r 是转动不变量):

$$x = -i\sqrt{\frac{2\pi}{3}}r(Y_{1-1} - Y_{11})$$

$$y = \sqrt{\frac{2\pi}{3}}r(Y_{1-1} + Y_{11})$$

$$z = -i\sqrt{\frac{4\pi}{3}}rY_{10} \tag{3.5.8}$$

把 Y_{lm}解出得

$$\begin{aligned} -i\sqrt{\frac{4\pi}{3}}rY_{11} &= \frac{-1}{\sqrt{2}}(x+iy) = \xi_1 \\ -i\sqrt{\frac{4\pi}{3}}rY_{1-1} &= \frac{1}{\sqrt{2}}(x-iy) = \xi_{-1} \\ -i\sqrt{\frac{4\pi}{3}}rY_{10} &= z = \xi_0 \end{aligned} \tag{3.5.9}$$

故可知 ξ_1, ξ_0, ξ_{-1}是一秩不可约张量。

(ⅳ)(L_x, L_y, L_z),(p_x, p_y, p_z)等与(x, y, z)在转动下变换方式一样。故

$$\begin{aligned} -\frac{1}{\sqrt{2}}(L_x + iL_y) &= -\frac{1}{\sqrt{2}}L_+ = L_1 \\ L_z &= L_0 \\ \frac{1}{\sqrt{2}}(L_x - iL_y) &= \frac{1}{\sqrt{2}}L_- = L_{-1} \end{aligned} \tag{3.5.10}$$

和

$$\begin{aligned} -\frac{1}{\sqrt{2}}(p_x + ip_y) &= p_1 \\ p_z &= p_0 \\ \frac{1}{\sqrt{2}}(p_x - ip_y) &= p_{-1} \end{aligned} \tag{3.5.11}$$

都是一秩不可约张量。

一般而言,向量算符(A_x, A_y, A_z)被定义作在转动下按下述方式变换的厄米算符:

$$D(g)A_\alpha D(g^{-1}) = \sum_\beta A_\beta R_{\beta\alpha}(g) \tag{3.5.12}$$

式中 $R_{\beta\alpha}(g)$是 $SO(3)$群在三维实欧氏空间 R^3 中的向量表示在笛卡尔基$(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$(也叫自然基)下的矩阵元。仿照(3.5.11)定义的 A_1, A_0, A_{-1}是 $SO(3)$群的一秩不可约张量。

把三维实欧氏空间 R^3 作复扩张，引入新的基

$$\hat{\xi}_1 = \frac{-1}{\sqrt{2}}(\boldsymbol{i} + i\boldsymbol{j})$$

$$\hat{\xi}_{-1} = \frac{1}{\sqrt{2}}(\boldsymbol{i} - i\boldsymbol{j})$$

$$\hat{\xi}_0 = \boldsymbol{k} \tag{3.5.13}$$

在空间旋转下，它们按标准基的方式变换

$$R(g)\hat{\xi}_\mu = \sum_{\mu'}\hat{\xi}_{\mu'}D^1_{\mu'\mu}(g) \tag{3.5.14}$$

态矢量空间上的向量算符

$$\boldsymbol{A} = A_x\hat{i} + A_y\hat{j} + A_z\hat{k} \tag{3.5.15}$$

向标准基 $\hat{\xi}_1, \hat{\xi}_0, \hat{\xi}_{-1}$ 作展开

$$\boldsymbol{A} = A_1\hat{\xi}_1 + A_0\hat{\xi}_0 + A_{-1}\hat{\xi}_{-1} \tag{3.5.16}$$

展开系数 (A_1, A_0, A_{-1}) 即一秩不可约张量算符。

一般而言，设 $R(k_1, q_1)$ 和 $S(k_2, q_2)$ 是不可约张量算符，则

$$T(k,q) = \sum R(k_1,q_1)S(k_2,q_2)\langle k_1q_1k_2q_2 \mid k_1k_2kq\rangle \equiv [R(k_1)\times S(k_2)]^k_q \tag{3.5.17}$$

是 k 秩不可约张量算符。(读者证明之!)特别是，

$$T(0,0) = \frac{1}{\sqrt{2k+1}}\sum_q(-1)^{k-q}R(k,q)S(k,-q) \tag{3.5.18}$$

是旋转不变量，0 秩不可约张量。当 k 是整数时，称旋转不变量

$$\sum_q(-1)^qR(kq)S(k-q) \equiv R(k)\cdot S(k) \tag{3.5.19}$$

为 $R(k)$ 与 $S(k)$ 的标积。

Racah 把不可约张量算符定义作 $2k+1$ 个满足如下对易关系的算符 $T(k,q)$：

$$[J_\pm, T(k,q)] = \sqrt{(k\mp q)(k\pm q+1)}\,T(k,q\pm 1)$$

$$[J_z, T(k,q)] = qT(k,q) \tag{3.5.20}$$

不难证明，这两个定义是等价的。Racah 的定义不涉及有限旋转，因此在某些情况下使用方便。

练习 3-5-1 试证不可约张量两种定义(3.5.2)和(3.5.20)是等价的。

Ⅱ. Wigner-Eckart 定理

我们已经指出，每一态矢量都可以展成各个角动量多重态空间的标准基 $|\gamma jm\rangle$ 的线性组合，这里的 γ 是角动量多重态子空间的标记，相当于一个力学量 Γ 的本征值，Γ 与 $\boldsymbol{J}$ 对易，且 Γ, J^2, J_z，构成力学量完全集合。例如 Γ 可以是电子在核库仑场中的哈密顿和轨道角动量平方 $\{H, L^2\}$，而态矢量空间中的每一线性算符都可以表示成不可约张量算符的线性组合。因而计算不可约张量算符在标准基下的矩阵元是量子力学中一件基本工作。

在讨论 Wigner-Eckart 定理之前，我们提请读者注意，当态矢量空间包含维数相同的角动量多重态子空间时，分解为角动量多重态子空间直和的方式不是唯一的。不同的分解对应于不同的完全力学量组。下面我们以 $|\gamma jm\rangle$，$|\gamma' j' m'\rangle$ 等记 $\{\Gamma, J^2, J_z\}$ 的共同的特征向量，而以 $|\lambda jm\rangle$，$|\lambda' j' m'\rangle$ 等记 $\{\Lambda, J^2, J_z\}$ 的共同的特征向量。因而 $\langle \lambda jm | \gamma jm\rangle$ 不一定为零，但 $\langle \gamma' j' m' | \gamma jm\rangle = \delta_{\gamma'\gamma}\delta_{j'j}\delta_{m'm}$，$\langle \lambda' j' m' | \lambda jm\rangle = \delta_{\lambda'\lambda}\delta_{j'j}\delta_{m'm}$。

Wigner-Eckart 定理：设 $\{T(kq) | q = k, k-1, \cdots, -k\}$ 是 k 秩不可约张量算符，$\{|\gamma jm\rangle | m = j, j-1, \cdots, -j\}$ 是权 j 的不可约表示空间的标准基，$\{|\lambda j' m'\rangle | m' = j', j'-1, \cdots, -j'\}$ 是群的权 j' 的不可约表示空间的标准基，则

$$\langle \lambda j' m' | T(kq) | \gamma jm\rangle = \langle \lambda j' \| T(k) \| \gamma j\rangle \langle kqjm | kjj' m'\rangle \tag{3.5.21}$$

式中 $\langle \lambda j' \| T(k) \| \gamma j\rangle$ 是一个与磁量子数 m', q, m 无关的数，叫约化矩阵元。

由这个定理，计算 $(2j'+1)(2j+1)(2k+1)$ 个矩阵元的任务归为计算约化矩阵元。事实上只要计算一个矩阵元 $\langle \lambda j' m' | T(kq) | \gamma jm\rangle$，若其对应的 $C.-G.$ 系数 $\langle kqjm | kjj' m'\rangle \neq 0$，就可以求出约化矩阵元 $\langle \lambda j' \| T(k) \| \gamma j\rangle$。然而 $W.-E.$ 定理的意义远不限于减少具体计算的工作量。它指出了对应于一定角量子数 j', j, k 的不可约张量算符在标准基间的矩阵元对磁量子数的依赖和 $C.-G.$ 系数对磁量子数的依赖一样。这一点对理论分析是非常重要的。总之，$W.-E.$ 定理揭示了对体系的旋转对称性分析能给我们哪些关于体系的知识。

定理的证明如下：在空间转动下，

$$\begin{aligned} D(g)T(kq)|\gamma jm\rangle &= D(g)T(kq)D^{-1}(g)D(g)|\gamma jm\rangle \\ &= \sum_{m_1 q_1} T(kq_1)|\gamma jm_1\rangle D^k_{q_1 q}(g) D^j_{m_1 m}(g) \end{aligned} \tag{3.5.22}$$

可见 $T(kq)|\gamma jm\rangle$按非耦合基$|kqjm\rangle$变换的方式变换。把它们用 $C.-G.$系数耦合起来

$$\sum_{mq} T(kq)\,|\,\gamma jm\rangle\langle kqjm\,|\,kjJM\rangle \equiv |\,(T\gamma)kjJM\rangle \tag{3.5.23}$$

利用§3.4的式(3.4.96),可以证明态矢(3.5.23)在空间转动下按耦合基$|kjJM\rangle$的变换方式变换,亦即在转动算符作用下与耦合基在转动算符作用下变换方式相同:

$$D(g)\,|\,(T\gamma)kjJM\rangle = \sum_{M'} |\,(T\gamma)kjJM'\rangle D^{J}_{M'M}(g) \tag{3.5.24}$$

从而在 $\boldsymbol{J}$ 的作用下与耦合基在 $\boldsymbol{J}$ 的作用下变换方式相同。因此可以把(3.5.23)看成一组满足惯例 $\boldsymbol{J}^2, J_z$ 的共同的本征函数(耦合基)。

解方程(3.5.23)得

$$T(kq)\,|\,\gamma jm\rangle = \sum_{JM} |\,(T\gamma)kjJM\rangle\langle kqjm\,|\,kjJM\rangle \tag{3.5.25}$$

代入(3.5.21)式的左边得

$$\begin{aligned}\langle \lambda j'm'\,|\,T(kq)\,|\,\gamma jm\rangle &= \sum_{JM}\langle \lambda j'm'\,|\,(T\gamma)kjJM\rangle\langle kqjm\,|\,kjJM\rangle \\ &= \sum_{JM}\delta_{j'j}\delta_{m'M}\langle \lambda j'm'\,|\,(T\gamma)kjj'm'\rangle\langle kqjm\,|\,kjJM\rangle \\ &= \langle \lambda j'm'\,|\,(T\gamma)kjj'm'\rangle\langle kqjm\,|\,kjj'm'\rangle\end{aligned} \tag{3.5.26}$$

剩下的要$\langle \lambda j'm'|(T\gamma)kjj'm'\rangle$证明与 m'无关。

设 $m' < j'$,则

$$(j'-m')(j'+m'+1) > 0 \tag{3.5.27}$$

$$\begin{aligned}\langle \lambda j'm'\,|\,(T\gamma)kjj'm'\rangle &= \langle \lambda j'm'\,|\,\frac{j'(j'+1)-m'^2-m'}{(j'-m')(j'+m'+1)}\,|\,(T\gamma)kjj'm'\rangle \\ &= \langle \gamma j'm'\,|\,\frac{J_-J_+}{(j'-m')(j'+m'+1)}\,|\,(T\gamma)kjj'm'\rangle \\ &= \frac{(j'-m')(j'+m'+1)}{(j'-m')(j'+m'+1)}\langle \lambda j'm'+1\,|\,(T\gamma)kjj'm'+1\rangle \\ &= \langle \lambda j'm'+1\,|\,(T\gamma)kjj'm'+1\rangle\end{aligned} \tag{3.5.28}$$

于是记

$$\langle \lambda j' m' \mid (T\gamma) kjj' m' \rangle = \langle \lambda j' \| T(k) \| \gamma j \rangle \tag{3.5.29}$$

称之为不可约张量算符 $\{T(kq) \mid q = -k, k-1, \cdots, k\}$ 的约化矩阵元。定理证完。

把(3.5.23)代入(3.5.29)并注意(3.5.29)与 m' 无关,得

$$\begin{aligned}\langle \lambda j' \| T(k) \| \gamma j \rangle &= \sum_{mq} \langle \lambda j' m' \mid T(kq) \mid \gamma jm \rangle \langle kqjm \mid kjj' m' \rangle \\ &= \frac{1}{2j'+1} \sum_{m'mq} \langle \lambda j' m' \mid T(kq) \mid \gamma jm \rangle \langle kqjm \mid kjj' m' \rangle\end{aligned} \tag{3.5.30}$$

Ⅲ. 约化矩阵元的例子

(ⅰ)由 §3.4 的(3.4.99)式知

$$\begin{aligned}&\langle Y_{lm} \mid Y_{l_1 m_1} \mid Y_{l_2 m_2} \rangle \\ &= \langle l_1 m_1 l_2 m_2 \mid l_1 l_2 lm \rangle i^{l_1+l_2-l} \langle l_1 0 l_2 0 \mid l_1 l_2 l0 \rangle \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}}\end{aligned} \tag{3.5.31}$$

可见约化矩阵元为

$$\langle Y_l \| Y_{l_1} \| Y_{l_2} \rangle = i^{l_1+l_2-l} \langle l_1 0 l_2 0 \mid l_1 l_2 l0 \rangle \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}} \tag{3.5.32}$$

(ⅱ)角动量向标准基 $\hat{\xi}_1, \hat{\xi}_0, \hat{\xi}_{-1}$ 作展开

$$\boldsymbol{J} = J_1 \hat{\xi}_1 + J_0 \hat{\xi}_0 + J_{-1} \hat{\xi}_{-1} \tag{3.5.33}$$

时的分量 (J_1, J_0, J_{-1}) 是一秩不可约张量。它们是

$$\begin{aligned}J_1 &= \frac{-1}{\sqrt{2}} J_+ \\ J_0 &= J_z \\ J_{-1} &= \frac{-1}{\sqrt{2}} J_-\end{aligned} \tag{3.5.34}$$

用 $\boldsymbol{J}$ 在标准基下的分量来表述不可约张量算符的 Racah 定义,我们有

$$\begin{aligned}[J_{\pm 1}, T(kq)]_- &= \mp \left[\frac{1}{2}(k \mp q)(k \pm q + 1)\right]^{1/2} T(kq \pm 1) \\ [J_0, T(kq)]_- &= qT(kq)\end{aligned} \tag{3.5.35}$$

另一方面，$j_1=1, j_2=j\neq 0$ 的系数 $\langle j_1 m_1 j_2 m_2 | j_1 j_2 jm\rangle$ 有如下的值：

$$\langle 11jm-1 \mid 1jjm\rangle = +\left[\frac{(j+m)(j-m+1)}{2j(j+1)}\right]^{1/2}$$

$$\langle 10jm-1 \mid 1jjm\rangle = \frac{-m}{[j(j+1)]^{1/2}}$$

$$\langle 1-1jm+1 \mid 1jjm\rangle = -\left[\frac{(j-m)(j+m+1)}{2j(j+1)}\right]^{1/2} \tag{3.5.36}$$

式(3.5.35)可以统一写作：对 $\mu=1,0,-1$，

$$[J_\mu, T(kq)]_- = (-1)^{1+\mu}[j(j+1)]^{1/2}\langle 1-\mu kq+\mu \mid 1kkq\rangle T(kq+\mu) \tag{3.5.37}$$

而 $J_\mu(\mu=1,0,-1)$ 对标准基的作用可以统一写作

$$J_\mu \mid \gamma jm\rangle = (-1)^{1+\mu}[j(j+1)]^{1/2}\langle 1-\mu jm+\mu \mid 1jjm\rangle \mid \gamma im+\mu\rangle \tag{3.5.38}$$

因而有

$$\begin{aligned}\langle \gamma' j' m' \mid J_\mu \mid \gamma jm\rangle &= \delta_{\gamma'\gamma}\delta_{j'j}\delta_{m',m+\mu}(-1)^{1+\mu}[j(j+1)]^{1/2}\langle 1-\mu jm+\mu \mid 1jjm\rangle \\ &= \delta_{\gamma'\gamma}\delta_{j'j}[j(j+1)]^{1/2}(-1)^{1+\mu}\langle 1-\mu jm' \mid 1jjm\rangle \\ &= -\delta_{\gamma'\gamma}\delta_{j'j}[j(j+1)]^{1/2}\langle 1\mu jm \mid 1jjm'\rangle\end{aligned}$$

最后一步利用了 $C.-C.$ 系数的对称性。考虑到因子 $\delta_{j'j}$ 的存在，把上式右边最后的 j 改成 j'，我们得到，对 $\mu=1,0,-1$，

$$\langle \gamma' j' m' \rangle \mid J_\mu \mid \gamma jm\rangle = -\delta_{\gamma'\gamma}\delta_{j'j}[j(j+1)]^{1/2}\langle 1\mu jm \mid 1jj'm'\rangle \tag{3.5.39}$$

于是约化矩阵元为

$$\langle \gamma' j' \| J(1) \| \gamma j\rangle = -\delta_{\gamma'\gamma}\delta_{j'j}[j(j+1)]^{1/2} \tag{3.5.40}$$

(ⅲ)0 秩不可约张量(旋转不变量)$T(00)$ 的约化矩阵元。按 $W.-E.$ 定理，有

$$\langle \lambda j' m' \mid T(00) \mid \gamma jm\rangle = \langle \lambda j' \| T(0) \| \gamma j\rangle\langle 00jm \mid j0j'm'\rangle = \delta_{j'j}\delta_{m'm}\langle \lambda j' \| T(0) \| \gamma j\rangle \tag{3.5.41}$$

这表明 0 秩不可约张量的约化矩阵元等于任一“对角元”(加引号表示 λ,γ 可能不同，不是真正对角元，只是 $j'=j, m'=m$)，$T(00)$ 的 $j'\neq j$ 或 $m'\neq m$ 的矩阵元为零。

Ⅳ. 约化矩阵元的耦合定律

让我们来考虑耦合的不可约张量算符(3.5.18)的约化矩阵元(3.5.29)：

$$
\begin{aligned}
&\langle \lambda j' \| [R(k_1) \times S(k_2)]^k \| \gamma j \rangle \\
&= \sum_{mq} \langle \lambda j' m' \mid [R(k_1) \times S(k_2)]^k_q \mid \gamma j m \rangle \langle kqjm \mid kjj'm' \rangle \\
&= \sum_{mqq_1q_2} \langle \lambda j' m' \mid R(k_1 q_1) S(k_2 q_2) \mid \gamma j m \rangle \\
&\quad \langle k_1 q_1 k_2 q_2 \mid k_1 k_2 kq \rangle \langle kqjm \mid kjj'm' \rangle \\
&= \sum_{\eta j'' m'' mqq_1q_2} \langle \lambda j' m' \mid R(k_1 q_1) \mid \eta j'' m'' \rangle \langle \eta j'' m'' \mid S(k_2 q_2) \mid \gamma j m \rangle \\
&\quad \langle k_1 q_1 k_2 q_2 \mid k_1 k_2 kq \rangle \langle kqjm \mid kjj'm' \rangle \\
&= \sum_{\eta j'' m'' mqq_1q_2} \langle \lambda j' \| R(k_1) \mid \eta j'' \rangle \langle \eta j'' \mid S(k_2) \| \gamma j \rangle \langle k_1 q_1 j'' m'' \mid k_1 j'' j' m' \rangle \langle k_2 q_2 jm \mid k_2 jj'' m'' \rangle \\
&\quad \times \langle k_1 q_1 k_2 q_2 \mid k_1 k_2 kq \rangle \langle kqjm \mid kjj'm' \rangle \\
&= \sum_{\eta j''} \langle \lambda j' \| R(k_1) \| \eta j'' \rangle \langle \eta j'' \| S(k_2) \| \gamma j \rangle (-1)^{j+k-j'} [(2j''+1)(2k+1)]^{1/2} \\
&\quad W(j' k_1 j k_2; j'' k)
\end{aligned}
\tag{3.5.42}
$$

式中 $W(j'k_1jk_2;j''k)$由下式定义

$$
\begin{aligned}
W(j'k_1jk_2;j''k) = &\sum_{m''mqq_1q_2} [(2j''+1)(2k+1)]^{-1/2} \langle k_1 q_1 k_2 q_2 \mid k_1 k_2 kq \rangle \langle kqjm \mid kjj'm' \rangle \\
&\times \langle k_2 q_2 jm \mid k_2 jj'' m'' \rangle \langle k_1 q_1 j'' m'' \mid k_1 j'' j' m' \rangle
\end{aligned}
\tag{3.5.43}
$$

是由 Racah 引入的,人们称之为 Racah 系数。式(3.5.42)表明,它是约化矩阵元的耦合系数。一个类似的系数,6 – j 符号由 Wigner 引入,与 Racah 系数的关系为

$$
\begin{Bmatrix} j' & k_1 & j'' \\ k_2 & j & k \end{Bmatrix} = (-1)^{j'+k_1+j+k_2} \overline{W}(j'k_1jk_2;j''k) \tag{3.5.44}
$$

让我们考虑约化矩阵元耦合定律的一个特殊情况:体系的态矢量空间是构成它的两个部分(子体系)的态矢量空间的张量积

$$
S = S^{(1)} \otimes S^{(2)}
$$

体系的角动量是子体系角动量之和

$$
\boldsymbol{J} = \boldsymbol{J}(1) + \boldsymbol{J}(2)
$$

而 $R^k_{q_1}(1)$是 $S^{(1)}$中的(相对于 $\boldsymbol{J}(1)$的)不可约张量算符;$S^{k_2}_{q_2}(2)$是 $S^{(2)}$中的(相对于 $\boldsymbol{J}$

(2)的)不可约张量算符。易知

$$R(k_1q_1) = R^{k_1}_{q_1}(1) \otimes 1(2)$$
$$S(k_2q_2) = 1(1) \otimes S^{k_2}_{q_2}(2) \tag{3.5.45}$$

是 S 中的(相对于 J 的)不可约张量算符。考虑耦合不可约张量算符$[R(k_1)S(k_2)]^k_q$在耦合角动量标准本征矢间的约化矩阵元。记

$$\sum_{m_1m_2} |\alpha_1j_1m_1\rangle \otimes |\alpha_2j_2m_2\rangle\langle j_1m_1j_2m_2 | j_1j_2jm\rangle$$
$$= |(\alpha_1\alpha_2)j_1j_2jm\rangle \equiv |\alpha jm\rangle \tag{3.5.46}$$

按式(3.5.42),我们应先计算 $R(k_1)$和 $S(k_2)$的约化矩阵元:

$$\langle \alpha'j' \| R(k_1) \| \alpha''j''\rangle = \sum_{m''q_1}\langle \alpha'j'm' | R(k_1q_1) | \alpha''j''m''\rangle\langle k_1q_1j''m'' | k_1j''j'm'\rangle$$
$$= \sum_{m''q_1}\langle (\alpha'_1\alpha'_2)j'_1j'_2j'm' | R^{k_1}_{q_1}(1) \otimes 1(2) | (\alpha''_1\alpha''_2)j''_1j''_2j''m''\rangle\langle k_1q_1j''m'' | k_1j''j'm'\rangle$$
$$= \sum_{q_1m''m'_1m'_2m''_1m''_2}\langle \alpha'_1j'_1m'_1 | R^{k_1}_{q_1}(1) | \alpha''_1j''_1m''_1\rangle\delta_{\alpha'_2\alpha''_2}\delta_{j'_2j''_2}\delta_{m'_2m''_2}\langle j'_1m'_1j'_2m'_2 | j'_1j'_2j'm'\rangle$$
$$\times \langle j''_1m''_1j''_2m''_2 | j''_1j''_2j''m''\rangle\langle k_1q_1j''m'' | k_1j''j'm'\rangle$$
$$= \delta_{\alpha'_2\alpha''_2}\delta_{j'_2j''_2}\langle \alpha'_1j'_1 \| R^{k_1} \| \alpha''_1j''_1\rangle \sum_{q_1m''m'_1m'_2m''_1}\langle k_1q_1j''_1m''_1 | k_1j''_1j'_1m'_1\rangle\langle j'_1m'_1j'_2m'_2 | j'_1j'_2j'm'\rangle$$
$$\langle j''_1m''_1j'_2m'_2 | j''_1j''_2j''m''\rangle\langle k_1q_1j''m'' | k_1j''j'm'\rangle \tag{3.5.47}$$

利用(3.5.43)和(3.5.44)上式可表作

$$\langle \alpha'j' \| R(k_1) \| \alpha''j''\rangle = \delta_{\alpha'_2\alpha''_2}\delta_{j'_2j''_2}\langle \alpha'_1j'_1 \| R^{k_1} \| \alpha''_1j''_1\rangle$$
$$\times (-1)^{k_1+j''_1+j'_2+j'}[(2j'_1+1)(2j''_1+1)]^{1/2}\begin{Bmatrix} j' & k_1 & j'' \\ j''_1 & j'_2 & j'_2 \end{Bmatrix} \tag{3.5.48}$$

用同样方法可得

$$\langle \alpha''j'' \| S(k_2) \| \alpha j\rangle = \delta_{\alpha''_1\alpha_1}\delta_{j''_1j_1}\langle \alpha''_2j''_2 \| S^{k_2} \| \alpha_2j_2\rangle$$
$$\times (-1)^{k_2+j+j_1+j''_2}[(2j''_2+1)(2j+1)]^{1/2}\begin{Bmatrix} j'' & k_2 & j \\ j_2 & j_1 & j''_2 \end{Bmatrix} \tag{3.5.49}$$

把(3.5.48)和(3.5.49)代入公式(3.5.42)得

$$\langle((\alpha_1'\alpha_2'j_1'j_2')j'\|[R^{k_1}(1)\times S^{k_2}(2)]^k\|(\alpha_1\alpha_2j_1j_2)j\rangle$$

$$=(-1)^{j_1+j_2-j+k_1+k_2-k+j_1'+j_2'-j'}[(2j_1'+1)(2j_2'+1)(2j+1)(2k+1)]^{1/2}\begin{Bmatrix}j_1 & j_2 & j\\ k_1 & k_2 & k\\ j_1' & j_2' & j'\end{Bmatrix}\cdot$$

$$\times\langle(\alpha_1')j_1'\|R^{k_1}\|(\alpha_1)j_1\rangle\langle(\alpha_2')j_2'\|S^{k_2}\|(\alpha_2)j_2\rangle \tag{3.5.50}$$

式中的 9 – j 符号定义如下：

$$\begin{Bmatrix}j_1 & j_2 & j\\ k_1 & k_2 & k\\ j_1' & j_2' & j'\end{Bmatrix}=(-1)^{j_1+j_2+j+k_1+k_2+k+j_1'+j_2'+j'}\sum_{j''}(-1)^{2j''}(2j''+1)\begin{Bmatrix}j' & k_2 & j''\\ & & \\ k_2 & j & k\end{Bmatrix}$$

$$\times\begin{Bmatrix}j' & j_2' & j_1'\\ & & \\ j_1 & k_1 & j''\end{Bmatrix}\begin{Bmatrix}j'' & j_1 & j_2'\\ & & \\ j_2 & k & j\end{Bmatrix} \tag{3.5.51}$$

事实上(3.5.48)和(3.5.49)可以作为(3.5.51)的特殊情形得到，分别对应于 $S^{k_2}_{q_2}(2)=1(2)(k_2=q_2=0)$和 $R^{k_1}_{q_1}(1)=1(1)(k_1=q_1=0)$。

从上面的讨论我们看到 Racah 系数(6 – j 符号)和 9 – j 符号怎样在不可约张量算符理论中被引入的。事实上，Racah 系数在多个角动量耦合理论中起着非常重要的作用。我们将在下一节中介绍它。

§6 多个角动量的耦合与 Racah 系数

从上节(§3.4，Ⅳ)的讨论，我们已经看到 Racah 系数是怎样被引入到不可约张量算符理论中的。它们是不可约张量算符约化矩阵元的耦合系数。在本节我们把 Racah 系数定义作三个角动量相加的再耦合系数，它们构成了多个角动量相加的再耦合系数的基础。

与 C. – G. 系数不同，Racah 系数不依赖于磁量子数(与基无关)。C. – G. 系数可

以作为 Racah 系数的极限得出，在角动量理论中出现的各种不变量都可由 Racah 系数得出。（参考 Biedenharn 等的《Angular momentum in Quantum Physics》和《The Racah - Wigner Algebra in Quantum Theory》）因而可以把它看作角动量理论中的基本量。

在多粒子体系物理（如原子物理、核物理）中，多个角动量理论起着举足轻重的作用。

Ⅰ. 三个角动量相加的再耦合系数

设体系由三部分（三个子体系）构成。把第 α 个子体系的态矢量空间记作 $S(\alpha)$。体系的态矢量空间为

$$S = S(1) \otimes S(2) \otimes S(3) \tag{3.6.1}$$

体系的总角动量为

$$\boldsymbol{J} = \boldsymbol{J}(1) + \boldsymbol{J}(2) + \boldsymbol{J}(3) \tag{3.6.2}$$

显然应有

$$[\boldsymbol{J}(\alpha), \boldsymbol{J}(\beta)] = 0 \qquad \forall \alpha \neq \beta \tag{3.6.3}$$

设 $L_{j\alpha}(\alpha) \subset S(\alpha)$，是 $\boldsymbol{J}(\alpha)$ 的 j_α - 权不可约不变子空间（角动量多重态子空间），则

$$L = L_{j_1}(1) \otimes L_{j_2}(2) \otimes L_{j_3}(3) \subset S \tag{3.6.4}$$

是 $\{\boldsymbol{J}(1), \boldsymbol{J}(2), \boldsymbol{J}(3)\}$ - 不可约不变子空间。以 $\{|j_\alpha m_\alpha\rangle | m_\alpha = j_\alpha, j_\alpha - 1, \cdots, -j_\alpha\}$ 记 $L_{j\alpha}(\alpha)$ 的标准基，则

$$\{|j_1 m_1\rangle \otimes |j_1 m_2\rangle \otimes |j_3 m_3\rangle \equiv |j_1 m_1 j_2 m_2 j_3 m_3\rangle | -j_\alpha \leq m_\alpha \leq j_\alpha\} \tag{3.6.5}$$

是张量积空间（3.6.4）的一个标准正交基，非耦合基。由两个角动量的耦合理论我们知道

$$L_{j_1}(1) \otimes L_{j_2}(2) = \bigotimes_{j_{12}=|j_1-j_2|}^{j_1+j_2} L_{j_{12}}(1,2) \tag{3.6.6}$$

$\boldsymbol{J}(1,2) = \boldsymbol{J}(1) + \boldsymbol{J}(2)$ 的不可约不变子空间 $L_{j_{12}}(12)$ 以耦合基

$$\left\{ \begin{aligned} &|j_1 j_2 j_{12} m_{12}\rangle = \sum_{m_1 m_2} |j_1 m_1 j_2 m_2\rangle \langle j_1 m_1 j_2 m_2 | j_1 j_2 j_{12} m_{12}\rangle \\ &| m_{12} = j_{12}, \cdots, -j_{12} \end{aligned} \right\} \tag{3.6.7}$$

为标准基。

张量积空间 $L_{j_{12}}(1,2)\otimes L_{j_3}(3)$是$\{\boldsymbol{J}(1,2),\boldsymbol{J}(3)\}$-不可约不变子空间,而

$$[\boldsymbol{J}(1,2),\boldsymbol{J}(3)]=0 \tag{3.6.8}$$

按照两个角动量耦合的理论,它们可以分解成$\boldsymbol{J}=\boldsymbol{J}(1,2)+\boldsymbol{J}(3)$的两两正交的不可约不变子空间的直和

$$L_{j_{12}}(1,2)\otimes L_{j_3}(3)=\bigoplus_{j=|j_{12}-j_3|}^{j_{12}+j_3} L_j^{j_{12}}(1,2;3) \tag{3.6.9}$$

而对固定的$j_1,j_2,j_{12}j_3,j$值,耦合基

$$\left\{\begin{array}{l}|(j_1j_2)j_{12}j_3jm\rangle=\sum\limits_{m_{12}m_3}|j_1j_2j_{12}m_{12}\rangle\otimes\langle j_3m_3\rangle\langle j_{12}m_{12}j_3m_3|j_{12}j_3jm\rangle \\ |-j\leq m\leq j\end{array}\right\} \tag{3.6.10}$$

是$\boldsymbol{J}$-不可约不变空间$L_j^{j_{12}}(1,2;3)$的一个标准基。而对固定的j_1,j_2,j_3值,

$$\left\{\begin{array}{l}|(j_1j_2)j_{12}j_3jm\rangle\ |\ j_{12}=|j_1-j_2|,\cdots,j_1+j_2; \\ j=|j_{12}-j_3|,\cdots,j_{12}+j_3;m=-j,-j+1,\cdots j\end{array}\right\} \tag{3.6.11}$$

是张量积空间(3.6.4)的一个标准正交基(耦合基)。(3.6.11)中各基矢是$\{J^2(1),J^2(2),J^2(1,2),J^2(3),J^2,J_z\}$的共同的本征矢。(3.6.11)与(3.6.5)两个标准正交基由一幺正变换矩阵相联系:

$$\begin{aligned}&|(j_1j_2)j_{12}j_3jm\rangle \\ &=\sum_{m_1m_2m_{12}m_3}|j_1m_1j_2m_2j_3m_3\rangle\langle j_1m_1j_2m_2|j_1j_2j_{12}m_{12}\rangle\langle j_{12}m_{12}j_3m_3|j_{12}j_3jm\rangle\end{aligned} \tag{3.6.12}$$

$\boldsymbol{J}$-不变的张量空间(3.6.4),还可以按其它方式分解成两两正交的$\boldsymbol{J}$-不可约不变子空间的直和。

例如,我们可以把

$$L_{j_2}(2)\otimes L_{j_3}(3)$$

分解成$\boldsymbol{J}(2,3)=\boldsymbol{J}(2)+\boldsymbol{J}(3)$的不可约不变子空间的直和。

$$L_{j_2}(2)\otimes L_{j_3}(3)=\bigoplus_{j_{23}=|j_2+j_3|}^{j_2+j_3} L_{j_{23}}(2,3) \tag{3.6.13}$$

再把张量积空间

$$L_{j_1}(1) \otimes L_{j_{23}}(2,3)$$

分解成 $\boldsymbol{J} = \boldsymbol{J}(1) + \boldsymbol{J}(2,3)$的不可约不变子空间的直和：

$$L_{j_1}(1) \otimes L_{j_{23}}(2,3) = \bigotimes_{j=|j_1-j_{23}|}^{j_1+j_{23}} L_j^{j_{23}}(1;2,3) \tag{3.6.14}$$

把各个 $L_j^{j_{23}}(1;2,3)$的标准基(另一种耦合基)

$$\begin{aligned}\{|\, j_1(j_2j_3)j_{23}jm\rangle = &\sum_{m_1m_2m_3m_{23}} |\, j_1m_1j_2m_2j_3m_3\rangle\langle j_2m_2j_3m_3 \,|\, j_2j_3j_{23}m_{23}\rangle \\ &\langle j_1m_1j_{23}m_{23} \,|\, j_1j_{23}jm\rangle \\ &|\, m = -j, -j+1, \cdots, j\}\end{aligned} \tag{3.6.15}$$

合在一起(j_{23}, j 取遍所有可能值)我们便得到张量积空间(3.6.4)的另一个标准正交基(也是一种耦合基,但不同于(3.6.11))。(3.6.16)中各基矢是$\{\boldsymbol{J}^2(1), \boldsymbol{J}^2(2), \boldsymbol{J}^2(3), \boldsymbol{J}^2(2,3), \boldsymbol{J}^2, J_z\}$的共同本征矢。

$$\left\{\begin{matrix}|\, j_1(j_2j_3)j_{23}jm\rangle \;|\; j_{23} = |\, j_2 - j_3\,|, \cdots, j_2 + j_3; \\ j = |\, j_1 - j_{23}\,|, \cdots, j_1 + j_{23}; m = -j, \cdots, j\end{matrix}\right\} \tag{3.6.16}$$

耦合基(3.6.11)和(3.6.16)对应于张量积空间(3.6.4)的两种分解方式

$$\begin{aligned}L &= L_{j_1}(1) \otimes L_{j_2}(2) \otimes L_{j_3}(3) \\ &= \bigoplus_{j_{12}=|j_1-j_2|}^{j_1+j_2} \bigoplus_{j=|j_{12}-j_3|}^{j_{12}+j_3} L_j^{j_{12}} = \bigoplus_{j_{23}=|j_2-j_3|}^{j_2+j_3} \bigoplus_{j=|j_1-j_{23}|}^{j_1+j_{23}} L_j^{j_{23}}\end{aligned} \tag{3.6.17}$$

由标准正交基(3.6.16)到(3.6.11)的变换矩阵是幺正矩阵,其矩阵元是两种基矢的内积。

$$\begin{aligned}|\,(j_1j_2)j_{12}j_3jm\rangle &= \sum_{j_{23}j'm'} |\, j_1(j_2j_3)j_{23}j'm'\rangle\langle j_1(j_2j_3)j_{23}j'm' \,|\, (j_1j_2)j_{12}j_3jm\rangle \\ &= \sum_{j_{23}} |\, j_1(j_2j_3)j_{23}jm\rangle\langle j_1(j_2j_3)j_{23}jm \,|\, (j_1j_2)j_{12}j_3jm\rangle\end{aligned} \tag{3.6.18}$$

最后一步利用了厄米算符 $\boldsymbol{J}^2, J_z$ 不同本征值的本征向量互相正交这一性质。注意变换系数$\langle j_1(j_2j_3)j_{23}jm \,|\, (j_1j_2)j_{12}j_3jm\rangle$与 m 无关(见式(3.5.28)),因而有时我们可以把其中的磁量子数 m 略去。由式(3.6.10)和(3.6.15)可知

$$\langle j_1(j_2j_3)j_{23}jm \mid (j_1j_2)j_{12}j_3jm\rangle = \sum_{m_1m_2m_3m_{12}m_{23}} \langle j_1m_1j_2m_2 \mid j_1j_2j_{12}m_{12}\rangle$$
$$\times \langle j_{12}m_{12}j_3m_3 \mid j_{12}j_3jm\rangle\langle j_2m_2j_3m_3 \mid j_2j_3j_{23}m_{23}\rangle\langle j_1m_1j_{23}m_{23} \mid j_1j_{23}jm\rangle \tag{3.6.19}$$

是实数。变换矩阵的幺正性为实正交性：

$$\sum_{j_{23}}\langle (j_1j_2)j_{12}j_3jm \mid j_1(j_2j_3)j_{23}jm\rangle\langle (j_1j_2)j'_{12}j_3jm \mid j_1(j_2j_3)j_{23}jm\rangle = \delta_{j_{12}j'_{12}} \tag{3.6.20}$$

$$\sum_{j_{12}}\langle j_1j_2j_{12}j_3jm \mid j_1j_2j_3j_{23}jm\rangle\langle j_1j_2j_{12}j_3jm \mid j_1j_2j_3j'_{23}jm\rangle = \delta_{j_{23}j'_{23}} \tag{3.6.21}$$

为了更好地表达对称性，Racah 引入了后来被人们称之为 Racah 系数的再耦合系数

$$W(j_1j_2jj_3;j_{12}j_{23}) = \frac{\langle (j_1j_2)j_{12}j_3j \mid j_1(j_2j_3)j_{23}j\rangle}{[(2j_{12}+1)(2j_{23}+1)]^{1/2}} \tag{3.6.22}$$

由式(3.6.19)可知

$$W(j_1j_2jj_3;j_{12}j_{23}) = [(2j_{12}+1)(2j_{23}+1)]^{1/2}$$
$$\times \sum_{m_1m_2m_{13}m_2m_{23}} \langle j_1m_1j_2m_2 \mid j_1j_1j_{12}m_{12}\rangle\langle j_{12}m_{12}j_3m_3 \mid j_{23}j_3jm\rangle$$
$$\times \langle j_2m_2j_3m_3 \mid j_2j_3j_{23}m_{23}\rangle\langle j_1m_1j_{23}m_{23} \mid j_1j_{23}jm\rangle \tag{3.6.23}$$

这个定义与 §3.5 的定义(3.5.43)是一致的，利用 $C.-G.$ 系数的对称性关系不难证明这一点。

Wigner 引入的 $6-j$ 符号比 Racah 系数有更高的对称性：

$$\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix} = (-1)^{j_1+j_2+j_3+j}W(j_1j_2jj_3;j_{12}j_{23}) \tag{3.6.24}$$

$6-j$ 符号可由 $C.-G.$ 系数表出：

$$\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix} = \frac{(-1)^{j_1+j_2+j_3+j}}{\sqrt{(2j_{12}+1)(2j_{23}+1)}}\frac{1}{2j+1}\sum_{m_1m_2m_3m_{12}m_{23}m} \langle j_1m_1j_2m_2 \mid j_1j_2j_{12}m_{12}\rangle$$
$$\langle j_{12}m_{12}j_3m_3 \mid j_{12}j_3jm\rangle\langle j_2m_2j_3m_3 \mid j_2j_3j_{23}m_{23}\rangle$$
$$\langle j_1m_1j_{23}m_{23} \mid j_1j_{23}jm\rangle \tag{3.6.25}$$

我们利用(3.6.23)式与 m 无关，对 m 取了平均值。

Ⅱ.6－j 符号的性质

6－j 符号的表达式(3.6.25)中包含四个 $C.-G.$ 系数的积,涉及 6 个角动量。它们的耦合关系可以用一个四面体表示(见图 3.6-1)

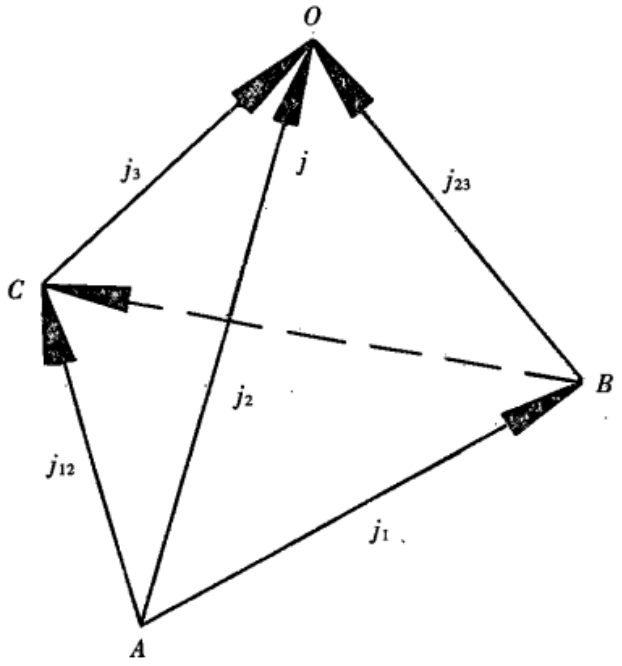

图 3.6-1

(ⅰ)非零必要条件:当四面体的四个三角形面对应的三角形关系不全满足时,6－j符号为零。这是 $C.-G.$ 系数三角形关系的直接结果。

(ⅱ)实正交性(实性＋幺正性):

$$\sum_{j_{23}}(2j_{12}+1)(2j_{23}+1)\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix}\begin{Bmatrix} j_1 & j_2 & j'_{12} \\ j_3 & j & j_{23} \end{Bmatrix}=\delta_{j_{12}j'_{12}} \tag{3.6.26}$$

$$\sum_{j_{12}}(2j_{12}+1)(2j_{23}+1)\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix}\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j'_{23} \end{Bmatrix}=\delta_{j_{23}j'_{23}} \tag{3.6.27}$$

(ⅲ)对称性

由式(3.6.25)和 $C.-G.$ 系数和对称关系可得:6－j 符号的值在任意两列的交换下不变

$$\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix} = \begin{Bmatrix} j_2 & j_1 & j_{12} \\ j & j_3 & j_{23} \end{Bmatrix} = \begin{Bmatrix} j_2 & j_{12} & j_1 \\ j & j_{23} & j_3 \end{Bmatrix} = \begin{Bmatrix} j_{12} & j_2 & j_1 \\ j_{23} & j & j_3 \end{Bmatrix}$$
$$= \begin{Bmatrix} j_{12} & j_1 & j_2 \\ j_{23} & j_3 & j \end{Bmatrix} = \begin{Bmatrix} j_1 & j_{12} & j_2 \\ j_3 & j_{23} & j \end{Bmatrix} \tag{3.6.28}$$

把图 3.6-1 中的四面体作成正四面体，则上面 $6-j$ 符号两列对换相当于正四面体保持顶角 O 不动的所有对称变换。(注意到通过 O 点的棱在 $6-j$ 符号的底行，不通过 O 点的棱在 $6-j$ 符号的顶行，而且同一列的两棱不相交，这个结论立刻可以得出。)

此外，$6-j$ 符号的值在同时倒置两列的变换下也是不变的。

$$\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix} = \begin{Bmatrix} j_3 & j & j_{12} \\ j_1 & j_2 & j_{23} \end{Bmatrix} = \begin{Bmatrix} j_3 & j_2 & j_{23} \\ j_1 & j & j_{12} \end{Bmatrix} = \begin{Bmatrix} j_1 & j & j_{23} \\ j_3 & j_2 & j_{12} \end{Bmatrix} \tag{3.6.29}$$

这相当于正四面体绕不相交的一对棱的中点联线转 π 角。

这两类变换生成了正四面体的对称性群 T_d，共 24 个元素。与此相应 $6-j$ 符号的值在此 24 个变换下保持不变。

(iv)$6-j$ 符号与 $C.-G.$ 系数的各种关系：

把(3.6.10)，(3.6.15)代入(3.6.18)，利用(3.6.22)和(3.6.24)得

$$\langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j_{12} m_1 + m_1 \rangle \langle j_{12} m_1 + m_2 j_3 m - m_1 - m_2 \mid j_{12} j_3 j m \rangle$$
$$= \sum_{j_{23}} (-1)^{j_1+j_2+j_3-j} [(2j_{12}+1)(2j_{23}+1)]^{1/2} \begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{Bmatrix}$$
$$\times \langle j_2 m_2 j_3 m - m_1 - m_2 \mid j_2 j_3 j_{23} m - m_1 \rangle \langle j_1 m_1 j_{23} m - m_1 \mid j_1 j_{23} j m \rangle \tag{3.6.30}$$

考虑由一次耦合基$\{\mid j_1 m_1 (j_2 j_3) j_{23} m - m_1 \rangle\}$到二次耦合基$\{\mid (j_1 j_2) j_{12} j_3 j m \rangle\}$的变换矩阵元

$$\langle (j_1 j_2) j_{12} j_3 j m \mid j_1 m_1 (j_2 j_3) j_{23} m - m_1 \rangle$$
$$= \langle (j_1 j_2) j_{12} j_3 j m \mid j_1 (j_2 j_3) j_{23} j m \rangle \langle j_1 (j_2 j_3) j_{23} j m \mid j_1 m_1 (j_2 j_3) j_{23} m - m_1 \rangle$$
$$= (-1)^{j_1+j_2+j_3+j} [(2j_{12}+1)(2j_{23}+1)]^{1/2} \begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{12} \end{Bmatrix} \langle j_1 m_1 j_{23} m - m_1 \mid j_1 j_{23} j m \rangle$$

另一方面在非耦合表象计算上式左边得

$$\sum_{m_2}\langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j_{12} m_1 + m_2\rangle\langle j_{12} m_1 + m_2 j_3 m - m_1 - m_2 \mid j_{12} j_3 j m\rangle$$
$$\times \langle j_2 m_2 j_3 m - m_1 - m_2 \mid j_2 j_3 j_{23} m - m_1\rangle$$

于是有

$$(-1)^{j_1+j_2+j_3+j}[(2j_{12}+1)(2j_{23}+1)]^{1/2}\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{12}\end{Bmatrix}\langle j_1 m_1 j_{23} m - m_1 \mid j_1 j_{23} j m\rangle$$
$$= \sum_m \langle j_1 m_1 j_2 m_2 \mid j_1 j_2 j_{12} m_1 + m_2\rangle\langle j_{12} m_1 + m_2 j_3 m - m_1 - m_2 \mid j_{12} j_3 j m\rangle$$
$$\times \langle j_2 m_2 j_3 m - m_1 - m_2 \mid j_2 j_3 j_{23} m - m_1\rangle \qquad (3.6.31)$$

式(3.6.30),(3.6.31)和(3.6.25)都是各种标准正交基间的实正交变换矩阵元。(各种标准正交基指非耦合基(3.6.5),二次耦合基(3.6.11),(3.6.16)以及各种一次耦合基)它们反映了相继变换的矩阵是各次变换矩阵之积。例如式(3.6.25)就代表由基(3.6.16)到基(3.6.11)的变换的矩阵是四次相继变换矩阵之积。

(Ⅴ)求和规则也是幺正变换的这种性质的表现。例如由

$$\langle (j_1 j_2) j_{12} j_3 j m \mid j_2 (j_3 j_1) j_{31} j m\rangle$$
$$= \sum_{j_{23}} \langle (j_1 j_2) j_{12} j_3 j m \mid j_1 (j_2 j_3) j_{23} j m\rangle\langle j_1 (j_2 j_3) j_{23} j m \mid j_2 (j_3 j_1) j_{31} j m\rangle$$

可得

$$\sum_{j_{23}} (-1)^{j_{23}+j_{31}+j_{12}}(2j_{23}+1)\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23}\end{Bmatrix}\begin{Bmatrix} j_2 & j_3 & j_{23} \\ j_1 & j & j_{21}\end{Bmatrix} = \begin{Bmatrix} j_3 & j_1 & j_{31} \\ j_2 & j & j_{12}\end{Bmatrix} \qquad (3.6.32)$$

考虑四个角动量耦合,Biedenharn 给出了另一个求和规则

$$\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j_{123} & j_{23}\end{Bmatrix}\begin{Bmatrix} j_{23} & j_1 & j_{123} \\ j_4 & j & j_{14}\end{Bmatrix}$$
$$= \sum_{j'} (-1)^{j_1+j_2+j_3+j_4+j_{12}+j_{23}+j_{14}+j_{123}+j+j'}(2j'+1)\begin{Bmatrix} j_3 & j_2 & j_{23} \\ j_{14} & j & j'\end{Bmatrix}\begin{Bmatrix} j_2 & j_1 & j_{13} \\ j_4 & j' & j_{14}\end{Bmatrix}\begin{Bmatrix} j_3 & j_{12} & j_{123} \\ j_4 & j & j'\end{Bmatrix} \qquad (3.6.33)$$

Ⅲ.9 – j 符号

让我们考虑四个角动量耦合的再耦合系数

$$\begin{aligned}
&\langle (j_1j_2)j_{12}(j_3j_4)j_{34}jm \mid (j_1j_3)j_{13}(j_2j_4)j_{24}jm\rangle \\
&= \sum_{j'}\langle (j_1j_2)j_{12}j_{34}j \mid j_1(j_2j_{34})j'j\rangle\langle j_2(j_3j_4)j_{34}j' \mid j_3(j_2j_4)j_{24}j'\rangle \\
&\quad \times \langle j_1(j_3j_{24})j'j \mid (j_1j_3)j_{13}j_{24}j\rangle \\
&= [(2j_{12}+1)(2j_{34}+1)(2j_{13}+1)(2j_{24}+1)]^{1/2} \\
&\quad \times \sum_{j'}(-1)^{2j'}(2j'+1)\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_{34} & j & j' \end{Bmatrix}\begin{Bmatrix} j_3 & j_4 & j_{34} \\ j_2 & j' & j_{34} \end{Bmatrix}\begin{Bmatrix} j_{13} & j_{24} & j \\ j' & j_1 & j_3 \end{Bmatrix}
\end{aligned} \tag{3.6.34}$$

这是一个与 m 无关的数。由下式定义 $9-j$ 符号：

$$\begin{aligned}
&\langle (j_1j_2)j_{12}(j_3j_4)j_{34}j \mid (j_1j_3)j_{13}(j_2j_4)j_{24}j\rangle \\
&= [(2j_{12}+1)(2j_{34}+1)(2j_{13}+1)(2j_{24}+1)]^{1/2}\begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & j_4 & j_{34} \\ j_{13} & j_{24} & j \end{Bmatrix}
\end{aligned} \tag{3.6.35}$$

于是我们有

$$\begin{Bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{Bmatrix} = \sum_k (-1)^{2k}(2k+1)\begin{Bmatrix} j_{11} & j_{21} & j_{31} \\ j_{32} & j_{33} & k \end{Bmatrix}\begin{Bmatrix} j_{12} & j_{22} & j_{32} \\ j_{21} & k & j_{23} \end{Bmatrix}\begin{Bmatrix} j_{13} & j_{23} & j_{33} \\ k & j_{11} & j_{12} \end{Bmatrix} \tag{3.6.36}$$

对 $9-j$ 符号我们不打算进一步讨论。有兴趣的读者可参考，例如，Edmonds: Angular momentum in Quantum Mechanics。

第 4 章　二次量子化

到目前为止我们向读者介绍的量子力学的理论形式，在两个方面有重大缺陷。首先，它不能描述伴有粒子产生、消灭的微观过程。按照传统形式量子力学的薛定格方程，若体系在初始时刻处在某个 N 粒子状态，则它在任意时刻都处在 N 粒子状态。而在现实的微观世界，特别是在高能领域中，粒子产生和消灭的过程时时处处都存在，例如 β - 衰变 $n \to p + e + \bar{v}_e$（一个中子衰变成一个质子、一个电子和一个电子型反中微子），双光子引起的正负电子对产生 $\gamma + \gamma \to e^+ + e^-$，正负电子对转化为正负 μ 子对 $e^+ + e^- \to \mu^+ + \mu^-$，等等。其次，全同粒子体系的波函数应具有粒子交换对称性（玻色子）或反对称性（费米子）。体系的力学量算符在粒子交换下是不变的，因而它们对波函数的作用不破坏其原有的粒子交换对称性或反对称性。例如，N 个电子体系的哈密顿算符是

$$H = \sum_{1 \le i \le N} T(i) + \sum_{1 \le i \le N} U(i) + \sum_{1 \le i < j \le N} \frac{e^2}{| \boldsymbol{r}_i - \boldsymbol{r}_j |}$$

它本身是交换不变的，但是在这种传统量子力学形式下，它被表作一些不具有交换不变性的项的和。要计算 H 对波函数的作用，须先计算每一项对波函数的作用，再对所有项求和。而个别项对波函数的作用，会破坏波函数原有的交换对称性或反对称性。也就是说，在计算力学量算符对波函数作用的中间过程中，波函数会跑出原来所在的态矢量空间（对称子空间或反对称子空间）。二次量子化理论形式的引入就是为了克服传统形式量子力学的上述缺陷。建议读者，在阅读了本章之后，再仔细考虑量子力学的两种理论形式的同异，二次量子化形式如何继承与发展了量子力学的传统形式。

§4.1 全同粒子体系的态空间与力学量

让我们考虑由 N 个某种微观粒子构成的体系。设单个这种粒子的态矢量空间是 S，而 $\{u_\alpha \equiv |\alpha\rangle \alpha \geq 0\}$ 是 S 的一个标准正交基：$\langle u_\alpha | u_\beta\rangle = \int \mathrm{d}q u_\alpha^*(q) u_\beta(q) = \delta_{\alpha\beta}$。暂时先不考虑全同粒子的不可分辨原理，则 N 粒子体系的态矢量空间是 N 个 S 的张量积 $S \otimes \cdots \otimes S \equiv S^{\otimes N}$。在 $S^{\otimes N}$ 上内积的定义是

$$\langle \Psi | \Phi\rangle = \int \cdots \int \mathrm{d}q_1 \cdots \mathrm{d}q_N \Psi^*(q_1, \cdots, q_N) \Phi(q_1, \cdots, q_N)$$

由这个定义可知

$$\{u_{\alpha_1} \otimes u_{\alpha_2} \otimes \cdots \otimes u_{\alpha_N} \equiv |\alpha_1 \cdots \alpha_N\rangle \mid \alpha_1, \cdots, \alpha_N \geq 0\} \tag{4.1.1}$$

构成 $S^{\otimes N}$ 的一个标准正交基。对每一个 N 次置换 $P \in S_N$（S_N 是 N 次对称群），在 N 粒子体系的态空间 $S^{\otimes N}$ 上定义线性算符 $D(P)$，

$$D(P)\Psi(1, \cdots, N) = \Psi(P_1, \cdots, P_N), \forall \Psi \in S^{\otimes N} \tag{4.1.2}$$

上式中 $\Psi(1, \cdots, N)$ 是 $\Psi(q_1, \cdots, q_N)$ 之略。对每个 N 次置换 $P \in S_N$，使 $D(P)$ 与 P 对应：$P \mapsto D(P)$，则我们在 N 粒态空间 $S^{\otimes N}$ 上，给出了 N 次对称群 S_N 的一个酉表示 D。让我们首先证明 $D(Q)D(P) = D(QP)$，$\forall P, Q \in S_N$。事实上，设 $\Psi(1, \cdots, N)$ 是任一 N 粒子态矢量，按照(4.1.2)，在 $D(P)$ 作用下变作 $D(P)\Psi(1, \cdots, N) = \Psi(P(1), \cdots, P(N))$，再受到 $D(Q)$ 的作用变为 $D(Q)D(P)\Psi(1, \cdots, N) = \Psi(QP(1), \cdots, QP(N))$。而这正是 $D(QP)\Psi(1, \cdots, N)$。从而 $D(Q)D(P) = D(QP)$，$\forall P, Q \in S_N$，$\{D(P) \mid P \in S_N\}$ 构成了 S_N 在 $S^{\otimes N}$ 上的一个表示。现在来证明这是一个酉表示。事实上，对每一 $P \in S_N$，$\langle D(P)\Psi | D(P)\Phi\rangle = \int \mathrm{d}q_1 \cdots \mathrm{d}q_N \Psi^*(P(1), \cdots, P(N)) \Phi(P(1), \cdots, P(N)) = \int \mathrm{d}q_1 \cdots \mathrm{d}q_N \Psi^*(1, \cdots, N)\Phi(1, \cdots, N) = \langle \Psi | \Phi\rangle$。这表明 $D(P)^+ D(P) = 1$，即 $D(P)^+$ 是 $D(P)$ 的左逆。无穷维线性空间上的线性算符的左逆不一定是右逆。但由于 $D(P^{-1}) = D(P)^{-1}$，即 $D(P)$ 有逆，逆为 $D(P^{-1})$，从而 $D(P)^+ = D(P)^{-1}$，即 $\{D(P) \mid P \in S_N\}$ 是 S_N 在 $S^{\otimes N}$ 上的一个酉表示。

让我们看一看算符 $D(P)$对基矢的作用：

$$(D(P)(u_{\alpha_1}\otimes\cdots\otimes u_{\alpha_N}))(1,2,\cdots,N)=u_{\alpha_1}(P_1)\cdots u_{\alpha_N}(P_N)=u_{\alpha_{P^{-1}(1)}}(1)\cdots u_{\alpha_{P^{-1}(N)}}(N)$$

此即

$$D(P)\mid\alpha_1,\cdots,\alpha_N\rangle=\mid\alpha_{P^{-1}(1)},\cdots,\alpha_{P^{-1}(N)}\rangle \tag{4.1.3}$$

值得读者注意的是，置换算符 $D(P)$对 N 粒态波函数的N 组自变量作置换P，相当于对基矢的 N 个单粒态下标作逆置换 P^{-1}。

按照全同粒子的不可分辨性原理，N 个全同玻色子体系的波函数 Φ 应该是对称的，即

$$D(P)\Phi(1,\cdots,N)=\Phi(P_1,\cdots,P_N)=\Phi(1,\cdots,N),\forall P\in S_N \tag{4.1.4}$$

而 N 个全同费米子体系的波函数 Ψ 应该是反对称的，即

$$D(P)\Psi(1,\cdots,N)=\Psi(P_1,\cdots,P_N)=(-1)^{v(P)}\Psi(1,\cdots,N),\forall P\in S_N \tag{4.1.5}$$

式中 $v(P)$是置换 P 的逆序数。当 P 是偶置换时，$(-1)^{v(P)}$为 $+1$；当 P 是奇置换时，$(-1)^{v(P)}$为 -1。不难证明 $v(P)=v(P^{-1})$。(留作练习)全体对称的 N 粒子波函数构成N 粒子态空间$S^{\otimes N}$的一个子空间，称之为对称子空间，记作$(S^{\otimes N})_B$，而全体反对称的 N 粒子波函数也构成N 粒子态空间$S^{\otimes N}$的一个子空间，称之为反对称子空间，记作$(S^{\otimes N})_F$。让我们来考虑如何由 $S^{\otimes N}$的基来构造这两个子空间的基。先考虑玻色子情况。

任一 N 粒子波函数Φ 都可以向$S^{\otimes N}$的基(4.1.1)展开

$$\Phi=\sum_{\alpha_1\cdots\alpha_N\geq 0}\mid\alpha_1\alpha_2\cdots\alpha_N\rangle c(\alpha_1\alpha_2\cdots\alpha_N) \tag{4.1.6}$$

由 Φ 的交换对称性可知，对所有 $P\in S_N$，

$$\begin{aligned}\Phi=D(P)\Phi&=\sum_{\alpha_1\cdots\alpha_N\geq 0}\mid\alpha_{P^{-1}(1)},\cdots,\alpha_{P^{-1}(N)}\rangle c(\alpha_1\alpha_2\cdots\alpha_N)\\&=\sum_{\beta_1\cdots\beta_N\geq 0}\mid\beta_1\beta_2\cdots\beta_N\rangle c(\beta_{P(1)}\beta_{P(2)}\cdots\beta_{P(N)})\\&=\sum_{\alpha_1\cdots\alpha_N\geq 0}\mid\alpha_1\alpha_2\cdots\alpha_N\rangle c(\alpha_{P(1)}\alpha_{P(2)}\cdots\alpha_{P(N)})\end{aligned} \tag{4.1.7}$$

最后两个等式成立是由于求和与求和指标的名称无关，这使我们可以改变求和指标的名称 $\alpha_{P^{-1}(j)}\mapsto\beta_j\mapsto\alpha_j$。由向量对同一基展开的唯一性可知

$$c(\alpha_1\cdots\alpha_N) = c(\alpha_{P(1)}\cdots\alpha_{P(N)}), \forall P \in S_N \tag{4.1.8}$$

这表明当对称的 N 粒子波函数向基展开时，由相同的 N 个单粒态按不同顺序相乘而得到的基矢具有相同的展开系数。例如，$|0022268\rangle$，$|8022026\rangle$，$|2226800\rangle$，…等 $420=\frac{7!}{2!\ 3!}$ 个基矢具有相同的系数。因而我们可以把展开式中由相同的 N 个单粒态按不同顺序相乘而得到的基矢合并在一块

$$\Phi = \sum_{0\le\alpha_1\le\cdots\le\alpha_N}\Big(\sum\nolimits'_P|\alpha_{P(1)}\cdots\alpha_{P(N)}\rangle\Big)c(\alpha_1\cdots\alpha_N) \tag{4.1.9}$$

式中第二个求和是对 N 个单粒态$(\alpha_1\cdots\alpha_N)$的所有不同的排列求和(不是对 N 个下标$(12\cdots N)$的所有 $N!$ 不同的排列求和!)。为了把这一问题表述得更清楚，让我们引入一个很有用的概念。

设$(\alpha_1\cdots\alpha_N)$是一 N 个非负整数的有序组，考察 N 次对称群 S_N 的子集

$$\{P\in S_N \mid (\alpha_1\cdots\alpha_N) = (\alpha_{P(1)}\cdots\alpha_{P(N)})\} \equiv S_N^{\alpha} \tag{4.1.10}$$

它是所有使 N 个非负整数的有序组$(\alpha_1\cdots\alpha_N)$保持不变的 N 次置换的集合。例如，当$(\alpha_1\cdots\alpha_N)=(333\cdots33)$时，$S_N^{\alpha}=S_N$；当$(\alpha_1\cdots\alpha_N)=(1234\cdots N)$时，$S_N^{\alpha}=\{E\}$，即只包含恒等置换 E 一个元素的子集；而当$(\alpha_1\cdots\alpha_N)=(2224567\cdots N)$时，$S_N^{\alpha}=S_3$，即关于头三个自然数的三次对称群。不难看出，$S_N^{\alpha}$ 是 S_N 的一个子群(作为练习，请读者证明之)。为简便起见，引入下面记法：$(\alpha_{P(1)}\cdots\alpha_{P(N)})\equiv\alpha_P$。于是上面的等式(4.1.10)可以写作

$$\{P\in S_N \mid \alpha_E = \alpha_P\} \equiv S_N^{\alpha} \tag{4.1.11}$$

若 N 个非负整数的有序组$(\alpha_1\cdots\alpha_N)$在两个置换 P 和 Q 下相等，即 $\alpha_{P(j)}=\alpha_{Q(j)}$，$\forall\,1\le j\le N$，把指标 j 改作 $Q^{-1}(k)$，左式变作 $\alpha_{PQ^{-1}(k)}=\alpha_k$，$\forall\,1\le k\le N$，，此即 $PQ^{-1}\in S_N^{\alpha}$。由关于子群的陪集的定理(设 H 是群 G 的子群，$a,b\in G$，则 $Ha=Hb\Leftrightarrow ab^{-1}\in H$)可知，$P$ 和 Q 属于同一个右陪集，亦即 $S_N^{\alpha}P=S_N^{\alpha}Q$。反之，当 P 和 Q 属于同一个右陪集，则有 $R\in S_N^{\alpha}$，使 $P=RQ$。从而 $\alpha_{P(i)}=\alpha_{RQ(i)}=\alpha_{R[Q(i)]}=\alpha_{Q(i)}$，$\forall\,1\le i\le N$。我们有 $\alpha_P=\alpha_Q$。利用这一结果，式(4.1.9)中的第二个求和 $\sum'_P$ 可以说成是让 P 跑遍子群 S_N^{α} 的所有右陪集，在每个右陪集中取一个任意指定的代表元。因而求和 $\sum'_P$ 包含 $\frac{N!}{|S_N^{\alpha}|}$ 个不同的项，式中 $|S_N^{\alpha}|$ 是子群 S_N^{α} 的阶。

对每一 N 个非负整数的有序组$(\alpha_1\cdots\alpha_N)$,可以唯一地确定非负整数序列$(n_0, n_1, n_2, \cdots)$,使 $n_\sigma = \sum_{s=1}^{N}\delta_{\sigma\alpha_s}$,代表 σ 在$(\alpha_1\cdots\alpha_N)$中出现的次数,也就是在 N 粒子态$|\alpha_1\cdots\alpha_N\rangle$中,处在单粒态 u_σ 的粒子数。显然有 $\sum_{\sigma\geq 0} n_\sigma = N$,非负整数序列$(n_0, n_1, n_2, \cdots)$只包含有限个非零项。反之,对每一个只包含有限个非零项的非负整数序列$(n_0, n_1, n_2, \cdots)$,也可以唯一地确定非负整数 $N = \sum_{\sigma\geq 0} n_\sigma$ 和非减的 N 个非负整数的有序组$(\alpha_1\cdots\alpha_N)$,$(0\leq\alpha_1\leq\cdots\leq\alpha_N)$,使 $n_\sigma = \sum_{s=1}\delta_{\sigma\alpha_s}$。利用粒子数 n_σ,不难写下子群 S_N^α 的阶 $|S_N^\alpha| = \prod_{\sigma\geq 0} n_\sigma!$(请读者证明之)。

让我们考察式(4.1.9)中的对称的 N 粒子波函数 $\sum_P{}' D(P^{-1})|\alpha_1\cdots\alpha_N\rangle$。由上面的讨论可知,它包含$\frac{N!}{\prod_{\sigma\geq 0} n_\sigma!}$个不同的项,每一项是归一化的,不同项互相正交。因而由式(4.1.9)可知

$$\left\{\sqrt{\frac{\prod_{\sigma\geq 0} n_\sigma!}{N!}}\sum_P{}' D(P)|\alpha_1\cdots\alpha_N\rangle \equiv |n_0 n_1 n_2\cdots\rangle_B \,\middle|\, 0\leq\alpha_1\leq\alpha_2\cdots\leq\alpha_N\right\} \tag{4.1.12}$$

构成 N 粒子态空间 $S^{\otimes N}$的对称子空间$(S^{\otimes N})_B$ 的一个标准正交基。(4.1.12)也可以写作

$$\left\{\frac{1}{\sqrt{N!\prod_{\sigma\geq 0} n_\sigma!}}\sum_{P\in S_N} D(P)|\alpha_1\alpha_2\cdots\alpha_N\rangle \equiv |n_0 n_1 n_2\cdots\rangle_B \,\middle|\, \sum_{\sigma\geq 0} n_\sigma = N, n_\sigma\geq 0\right\} \tag{4.1.13}$$

式(4.1.9)可以写作

$$\begin{aligned}\Phi &= \sum_{0\leq\alpha_1\leq\cdots\leq\alpha_N} |n_0 n_1 n_2\cdots\rangle_B \sqrt{\frac{N!}{\prod_{\sigma\geq 0} n_\sigma!}}\, c(\alpha_1\cdots\alpha_N) \\ &= \sum_{\sum_{\sigma\geq 0} n_\sigma = N, n_\sigma\geq 0} |n_0 n_1 n_2\cdots\rangle_B b(n_0 n_1 n_2\cdots)\end{aligned} \tag{4.1.14}$$

现在让我们来考虑如何由 $S^{\otimes N}$的基来构造其反对称子空间$(S^{\otimes N})_F$ 的基。把反

对称 N 粒子波函数 Ψ 向基(4.1.1)展开

$$\Psi = \sum_{\alpha_1\cdots\alpha_N\geq 0} |\alpha_1\alpha_2\cdots\alpha_N\rangle c(\alpha_1\alpha_2\cdots\alpha_N) \tag{4.1.15}$$

由 Ψ 的交换反对称性可知,对所有 $P\in S_N$,

$$\begin{aligned}\Psi &= (-1)^{v(P)}D(P)\Psi = \sum_{\alpha_1\cdots\alpha_N\geq 0} |\alpha_{P^{-1}(1)}\alpha_{P^{-1}(2)}\cdots\alpha_{P^{-1}(N)}\rangle c(\alpha_1\alpha_2\cdots\alpha_N)(-1)^{v(P)}\\ &= \sum_{\beta_1\cdots\beta_N\geq 0} |\beta_1\beta_2\cdots\beta_N\rangle c(\beta_{P(1)}\beta_{P(2)}\cdots\beta_{P(N)})(-1)^{v(P)}\\ &= \sum_{\alpha_1\cdots\alpha_N\geq 0} |\alpha_1\alpha_2\cdots\alpha_N\rangle c(\alpha_{P(1)}\alpha_{P(2)}\cdots\alpha_{P(N)})(-1)^{v(P)}\end{aligned} \tag{4.1.16}$$

由向量对同一基展开的唯一性可知

$$c(\alpha_1\cdots\alpha_N) = c(\alpha_{P(1)}\cdots\alpha_{P(N)})(-1)^{v(P)}, \forall P\in S_N \tag{4.1.17}$$

这表明当反对称的 N 粒子波函数按基(4.1.1)展开时,基矢 $|\alpha_1,\cdots,\alpha_N\rangle$ 的系数与基矢 $|\alpha_{P(1)},\cdots,\alpha_{P(N)}\rangle$ 的系数相差一个符号 $(-1)^{v(P)}$。特别是,含有两个相同单粒态的基矢的系数为零。因而我们可以把展开式中由相同的 N 个互异的单粒态按不同顺序相乘而得到的基矢合并在一块

$$\Psi = \sum_{0\leq\alpha_1<\cdots<\alpha_N}\Big(\sum_{P\in S_N}|\alpha_{P(1)}\cdots\alpha_{P(N)}\rangle(-1)^{v(P)}\Big)c(\alpha_1\cdots\alpha_N) \tag{4.1.18}$$

式中第二个求和是对 N 个互异单粒态 $(\alpha_1\cdots\alpha_N)$ 的所有不同的排列求和(也就是对 N 个下标(12…N)的所有不同的排列求和!)。由 N 个互异的单粒态 $(\alpha_1\cdots\alpha_N)$ 所确定的粒子数序列 $(n_0n_1n_2\cdots)$,显然也使 $\sum_{\sigma\geq 0}n_\sigma = N$。但与玻色子情况不同,在费米子情况下,$n_\sigma$ 只能取值 0 或 1。因而由式(4.1.18)可知

$$\left\{\frac{1}{\sqrt{N!}}\sum_{P\in S_N}|\alpha_{P(1)}\cdots\alpha_{P(N)}\rangle(-1)^{v(P)} \equiv |n_0n_1n_2\cdots\rangle_F \mid 0\leq\alpha_1<\alpha_2<\cdots<\alpha_N\right\}$$

或者

$$\left\{\frac{1}{\sqrt{N!}}\sum_{P\in S_N}|\alpha_{P(1)}\cdots\alpha_{P(N)}\rangle(-1)^{v(P)} \equiv |n_0n_1n_2\cdots\rangle_F \mid \sum_{\sigma\geq 0}n_\sigma = N, n_\sigma = 0,1\right\} \tag{4.1.19}$$

构成 N 粒子态空间 $S^{\otimes N}$ 的反对称子空间 $(S^{\otimes N})_F$ 的一个标准正交基。式(4.1.18)可

以重新写作

$$\Psi = \sum_{0\le a_1<\cdots<a_N} \frac{1}{\sqrt{N!}}\Big(\sum_{P\in S_N} | \alpha_{P(1)}\cdots\alpha_{P(N)}\rangle(-1)^{v(P)}\Big)\sqrt{N!}\,c(\alpha_1\cdots\alpha_N)$$

$$= \sum_{\sum_{\sigma\ge0} n_\sigma = N,\, n_\sigma = 0,1} | n_0 n_1 n_2\cdots\rangle_F f(n_0 n_1 n_2\cdots) \tag{4.1.20}$$

为了能够描述伴随有粒子产生和消灭的过程，在这些过程中微观体系的粒子组成不再是不变的，我们首先要把微观体系的概念作某些修正。我们不再说 N 个全同玻色子的体系，而是说某种全同玻色子的体系。后者的粒子数可以不确定。相应地，应该把体系的态矢量空间扩大。为此引入这种玻色子的真空态 $|\rangle_B$，即一个玻色子也没有的态，并令 ${}_B\langle|\rangle_B = 1$。以 $S^{\otimes 0}$ 记由 $|\rangle_B$ 张成的一维复内积空间 $C|\rangle_B$。我们假定，这种全同玻色子体系的态空间是

$$B = \bigoplus_{N\ge0}(S^{\otimes N})_B \tag{4.1.21}$$

即零粒态子空间，单粒态子空间，二粒对称态子空间，三粒对称态子空间，等等的直和。在这个复线性空间 B 中定义内积，使粒子数不同的态矢互相正交，而两个 N 粒态矢的内积就等于它们在 $(S^{\otimes N})_B$ 中的内积。不难看出

$$\left\{\frac{1}{\sqrt{N!\prod_{\sigma\ge0} n_\sigma!}}\sum_{P\in S_N} D(P) | \alpha_1\alpha_2\cdots\alpha_N\rangle \equiv | n_0 n_1 n_2\cdots\rangle_B \;\middle|\; N\ge0, \sum_{\sigma\ge0} n_\sigma = N, n_\sigma \ge 0\right\} \tag{4.1.22}$$

是 B 的一个标准正交基。

对某种全同费米子，同样引入这种费米子的真空态 $|\rangle_F$。假定这种全同费米子体系的态空间是

$$F = \bigoplus_{N\ge0}(S^{\otimes N})_F \tag{4.1.23}$$

即零粒态子空间，单粒态子空间，二粒反对称态子空间，三粒反对称态子空间，等等的直和。以类似的方式在 F 上定义内积，则

$$\left\{\frac{1}{\sqrt{N!}}\sum_{P\in S_N}(-1)^{v(P)} D(P) | \alpha_1\alpha_2\cdots\alpha_N\rangle \equiv | n_0 n_1 n_2\cdots\rangle_F \;\middle|\; N\ge0, \sum_{\sigma\ge0} n_\sigma = N, n_\sigma = 0,1\right\} \tag{4.1.24}$$

构成 F 的一个标准正交基。

数学上把 N 阶矩阵的行列式(determinant)记作

$$\sum_{P\in S_N}(-1)^{v(P)}a_{1P(1)}\cdots a_{NP(N)}=\begin{vmatrix}a_{11}&\cdots&a_{1N}\\ \vdots&\ddots&\vdots\\ a_{N1}&\cdots&a_{NN}\end{vmatrix} \tag{4.1.25}$$

仿此引入记法

$$\sum_{P\in S_N}a_{1P(1)}\cdots a_{NP(N)}=\begin{Vmatrix}a_{11}&\cdots&a_{1N}\\ \vdots&\ddots&\vdots\\ a_{N1}&\cdots&a_{NN}\end{Vmatrix} \tag{4.1.26}$$

并称之为 N 阶矩阵的积和式(permanent)。

现在让我们考虑 N 个全同粒子体系的力学量。由于全同粒子的不可分辨性,我们不能测量某个指定粒子的动能,而只能测量体系的总动能;不能测量某个指定粒子对间的库仑能,而只能测量体系所有粒子对间的总库仑能。因此,N 个全同粒子体系的可以测量的力学量是各种单体力学量 $T_N^1=\sum\limits_{1\le i\le N}t(i)$,各种二体力学量 $T_N^2=\sum\limits_{1\le i<j\le N}t(i,j)=\dfrac{1}{2}\sum\limits_{i\ne j}t(i,j)$,等等。或者说,$N$ 个全同粒子体系的可以测量的力学量是各种 k - 体力学量($k\geqslant 1$)

$$T_N^k=\sum_{1\le i_1<i_2<\cdots<i_k\le N}t(i_1,\cdots,i_k)=\frac{1}{k!}\sum_{i_b\ne i_c(\forall\, b\ne c)}t(i_1,\cdots,i_k),(k\ge 1) \tag{4.1.27}$$

上式右端的求和,是对从 N 个粒子中依次取出 k 个来的所有不同排列方式求和。注意这个表达式与粒子数 N 有关,当 $N<k$ 时,上式等于零。由粒子的全同性,应有

$$t(i_1,\cdots,i_k)=t(i_{P(1)},\cdots,i_{P(k)}),\forall\, P\in S_k \tag{4.1.28}$$

注意这里的置换 P 不是 N 级置换,而是 k 级置换。当 $N>k$ 时,上式中的 $t(i_1,\cdots,i_k)$不具有 N 级置换不变性。但是 k - 体力学量算符(4.1.27)(它们是定义在 N 粒态空间 $S^{\otimes N}$上的)具有 N 级置换不变性(请读者证明之)

$$D(P)\sum_{1\le i_1<i_2<\cdots<j_k\le N}t(i_1,\cdots,i_k)D^{-1}(P)=\sum_{1\le i_1<i_2<\cdots<j_k\le N}t(i_1,\cdots,i_k),\forall\, P\in S_N \tag{4.1.29}$$

由于这种不变性，$S^{\otimes N}$的对称子空间$(S^{\otimes N})_B$和反对称子空间$(S^{\otimes N})_F$都是k-体力学量算符(4.1.27)的不变子空间。N个全同玻色子(费米子)体系的力学量k-体算符(4.1.27)可看作定义在$(S^{\otimes N})_B((S^{\otimes N})_F)$上的算符。

全同玻色子(费米子)体系的态空间是$B=\bigoplus_{N\geq 0}(S^{\oplus N})_B(F=\bigoplus_{N\geq 0}(S^{\otimes N})_F)$，因而全同玻色子(费米子)体系的力学量算符应定义在$B(F)$上。全同玻色子(费米子)体系的k-体力学量算符T^k应这样确定，使得$(S^{\otimes N})_B((S^{\otimes N})_F)$是$T^k$的不变子空间，而$T^k$到$(S^{\otimes N})_B((S^{\otimes N})_F)$上的限制正好等于$T^k_N$。

练习 4-1-1 试证：$v(P)=v(P^{-1}),\forall P\in S_N$。

练习 4-1-2 试证：式(4.1.9)中的$\sum'_P D(P)\mid\alpha_1\cdots\alpha_N\rangle$是对称的$N$粒子波函数。

练习 4-1-3 试证：$\frac{1}{N!}\sum_{P\in S_N}D(P)$是把$N$粒子态空间$S^{\otimes N}$投射到其对称子空间$(S^{\otimes N})_B$的厄米投影算子。

练习 4-1-4 试证：$\frac{1}{N!}\sum_{P\in S_N}(-1)^{v(P)}D(P)$是把$N$粒子态空间$S^{\otimes N}$投射到其反对称子空间$(S^{\otimes N})_F$的厄米投影算子。

练习 4-1-5 试证：式(4.1.10)定义的S_N^α是N次对称群S_N的子群。

练习 4-1-6 试证：N次对称群S_N的子群S_N^α的阶$\mid S_N^\alpha\mid=\prod_{\sigma\geq 0}n_\sigma!$。

练习 4-1-7 试证：N个全同粒子体系的k-体力学量算符具有N次置换不变性。

$$D(P)\sum_{1\leq i_1<i_2<\cdots<i_k\leq N}t(i_1,\cdots,i_k)D^{-1}(P)=\sum_{1\leq i_1<i_2<\cdots<i_k\leq N}t(i_1,\cdots,i_k),\forall P\in S_N$$

练习 4-1-8 令$\boldsymbol{L}=\sum_{0\leq i\leq N}\boldsymbol{l}(i)$表示$N$个全同粒子体系的轨道角动量。问$L_z,L^2$是几体算符？

§4.2 粒子产生与消灭算符

二次量子化最关键的一步是引入粒子产生与消灭算符。它们是定义在扩大了的

全同粒子体系的态空间上的线性算符。对任一非负整数 N，N 粒态子空间不是这些算符的不变子空间。而在通常形式的量子力学中，N 个全同粒子体系的所有力学量算符都是定义在 N 粒态空间上的。这表明粒子产生与消灭算符是全新的一类算符。我们将证明全同粒子体系的任一力学量都可以通过它们来表达。

我们首先讨论全同玻色子体系的粒子产生与消灭算符。

在某个线性空间上定义一个线性算符，只要规定它对某一个基的所有基矢的作用结果就行了。不管如何规定，都能唯一地确定一个线性算符。从上一节的讨论我们知道，对单粒态空间 S 的一个标准正交基 $\{u_\alpha \equiv |\alpha\rangle | \alpha \geq 0\}$，(4.1.22)是全同玻色子体系态空间 B 的一个标准正交基。对每一个单粒态 $|\alpha\rangle$，在 B 上定义相应的产生算符 c_α，使得当粒子数 $N=0$ 时，

$c_\alpha |\rangle = |\alpha\rangle$，

当 $N \geq 1$ 时，

$$c_\alpha \frac{1}{\sqrt{N!}} \begin{Vmatrix} u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N) \\ u_{\alpha_2}(1) & u_{\alpha_2}(2) & \cdots & u_{\alpha_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N) \end{Vmatrix}$$

$$= \frac{1}{\sqrt{(N+1)!}} \begin{Vmatrix} u_\alpha(1) & u_\alpha(2) & \cdots & u_\alpha(N+1) \\ u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N+1) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N+1) \end{Vmatrix} \tag{4.2.1}$$

而乘上归一化因子后，这就是

$$c_\alpha \, | \, n_1 \cdots n_\sigma \cdots \rangle_B = \sqrt{n_\alpha + 1} \, | \, n_1 + \delta_{1\alpha} \cdots n_\sigma + \delta_{\sigma\alpha} \cdots \rangle_B \tag{4.2.2}$$

可见线性算符 c_α 在此标准正交基下不为零的矩阵元只有

$${}_B\langle n_1 + \delta_{1\alpha} \cdots n_\sigma + \delta_{\sigma\alpha} \cdots | \, c_\alpha \, | \, n_1 \cdots n_\sigma \cdots \rangle_B = \sqrt{n_\alpha + 1} \tag{4.2.3}$$

c_α 的厄米共轭 c_α^+ 在此标准正交基下不为零的矩阵元只有

$${}_B\langle n_1 \cdots n_\sigma \cdots | \, c_\alpha \, | \, n_1 + \delta_{1\alpha} \cdots n_\sigma + \delta_{\sigma\alpha} \cdots \rangle_B = \sqrt{n_\alpha + 1}$$

此即

$$ {}_B\langle n_1 - \delta_{1\alpha}\cdots n_\sigma - \delta_{\sigma\alpha}\cdots \mid c_\alpha^+ \mid n_1\cdots n_\sigma\cdots\rangle_B = \sqrt{n_\alpha} \tag{4.2.4} $$

从而

$$ c_\alpha^+ \mid n_1\cdots n_\sigma\cdots\rangle_B = \sqrt{n_\alpha} \mid n_1 - \delta_{1\alpha}\cdots n_\sigma - \delta_{\sigma\alpha}\cdots\rangle_B \tag{4.2.5} $$

为了与习惯一致,令 $b_\alpha = c_\alpha^+$,则有 $b_\alpha^+ = c_\alpha$,式(4.2.2)与(4.2.5)变作

$$ b_\alpha^+ \mid n_1\cdots n_\sigma\cdots\rangle_B = \sqrt{n_\alpha + 1} \mid n_1 + \delta_{1\alpha}\cdots n_\sigma + \delta_{\sigma\alpha}\cdots\rangle_B \tag{4.2.6} $$

$$ b_\alpha \mid n_1\cdots n_\sigma\cdots\rangle_B = \sqrt{n_\alpha} \mid n_1 - \delta_{1\alpha}\cdots n_\sigma - \delta_{\sigma\alpha}\cdots\rangle_B \tag{4.2.7} $$

特别是,当所有粒子数都为零时,我们有

$$ b_\alpha \mid 0,\cdots,0,\cdots\rangle_B = 0, {}_B\langle 0,\cdots,0,\cdots \mid b_\alpha^+ = 0, \forall\, \alpha \geq 0 $$

从等式(4.2.6)、(4.2.7)不难推出(请读者证明之)对易关系

$$ [b_\alpha, b_\beta^+] \equiv b_\alpha b_\beta^+ - b_\beta^+ b_\alpha = \delta_{\alpha\beta}, [b_\alpha, b_\beta] = [b_\alpha^+, b_\beta^+] = 0 \tag{4.2.8} $$

以及基矢表达式

$$ \mid n_1\cdots n_\sigma\cdots\rangle_B = \frac{(b_1^+)^{n_1}\cdots(b_\sigma^+)^{n_\sigma}}{\sqrt{n_1!\cdots n_\sigma!\cdots}} \mid 0\cdots 0\cdots\rangle_B \tag{4.2.9} $$

这表明由真空态 $\mid 0\cdots 0\cdots\rangle_B$ 出发,作用以各种产生算符,可以得到基(4.1.22)中的所有基矢(不计较归一化因子)。全同玻色子体系的态空间 B 可由 $\mid 0\cdots 0\cdots\rangle_B$ 和所有的 $(b_1^+)^{n_1}\cdots(b_\sigma^+)^{n_\sigma}\cdots\mid 0\cdots 0\cdots\rangle_B$ 张成。

对每一个单粒态 $\mid\alpha\rangle$,定义 B 上的线性算符

$$ N_\alpha = b_\alpha^+ b_\alpha \tag{4.2.10} $$

这是一个厄米算符,其本征值只能是非负实数。不难证明

$$ [N, b_\beta] = -\delta_{\alpha\beta} b_\beta, [N_\alpha, b_\beta^+] = \delta_{\alpha\beta} b_\beta^+, [N_\alpha, N_\beta] = 0 \tag{4.2.11} $$

以及

$$ N_\alpha \mid n_0\cdots n_\sigma\cdots\rangle_B = n_\alpha \mid n_0\cdots n_\sigma\cdots\rangle_B \tag{4.2.12} $$

因而我们称 N_α 为关于单粒态 $\mid\alpha\rangle$ 的粒子数算符。所有的粒子数算符 $\{N_\alpha \mid \alpha \geq 0\}$ 构成全同玻色子体系的一个完全力学量组。N_α 的本征值是全体非负整数。所有单粒态 $\mid\alpha\rangle(\alpha \geq 0)$ 的粒子数算符 N_α 之和,$\hat{N} = \sum\limits_{0\leq\alpha} N_\alpha$ 称作总粒子数算符。

现在让我们讨论全同费米子体系的粒子产生与消灭算符。对单粒态空间 S 的一个标准正交基 $\{u_\alpha \equiv \mid\alpha\rangle \mid \alpha \geq 0\}$,(4.1.24)是全同费米子体系态空间 F 的一个标准

正交基。仿照式(4.2.1),对每一个单粒态$|\alpha\rangle$,在 F 上定义相应的产生算符 a_α^+,使得

$$a_\alpha^+ |\rangle = |\alpha\rangle$$

$$a_\alpha^+ \frac{1}{\sqrt{N!}} \begin{vmatrix} u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N) \\ u_{\alpha_2}(1) & u_{\alpha_2}(2) & \cdots & u_{\alpha_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N) \end{vmatrix}$$

$$= \frac{1}{\sqrt{(N+1)!}} \begin{vmatrix} u_\alpha(1) & u_\alpha(2) & \cdots & u_\alpha(N+1) \\ u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N+1) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N+1) \end{vmatrix} \tag{4.2.13}$$

注意,$\frac{1}{\sqrt{N!}} \begin{vmatrix} u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N) \\ u_{\alpha_2}(1) & u_{\alpha_2}(2) & \cdots & u_{\alpha_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N) \end{vmatrix} = |n_0 \cdots n_\sigma \cdots\rangle_F$,其中的单粒态下标满足不等式 $\alpha_1 < \cdots < \alpha_N$,而粒子数 n_σ 只能等于 1 或 0。当 α 等于某个 α_k 时,即当 $n_\alpha = 1$ 时,等式右边的行列式为零。当 $\alpha_k < \alpha < \alpha_{k+1}$时,这时 $n_\alpha = 0$,把右边行列式的第一行移到第 $k+1$ 行,需要做 $k = \sum_{\sigma=0}^{\alpha-1} n_\sigma$ 次行交换。因此式(4.2.10)可以写作

$$a_\alpha^+ |n_0 \cdots n_\sigma \cdots\rangle_F = (-1)^{\sum_{\sigma=0}^{\alpha-1} n_\sigma (1-n_\alpha)} |n_0 + \delta_{0\alpha} \cdots n_\sigma + \delta_{\sigma\alpha} \cdots\rangle_F \tag{4.2.14}$$

可见线性算符 a_α^+ 在此标准正交基下不为零的矩阵元只有

$${}_F\langle n_1 + \delta_{1\alpha} \cdots n_\sigma + \delta_{\sigma\alpha} \cdots | a_\alpha^+ | n_1 \cdots n_\sigma \cdots\rangle_F = (-1)^{\sum_{\sigma=0}^{\alpha-1} n_\sigma (1-n_\alpha)} \tag{4.2.15}$$

a_α^+ 的厄米共轭 a_α 在此标准正交基下不为零的矩阵元只有

$${}_F\langle n_1 \cdots n_\sigma \cdots | a_\alpha | n_1 + \delta_{1\alpha} \cdots n_\sigma + \delta_{\sigma\alpha} \cdots\rangle_F = (-1)^{\sum_{\sigma=0}^{\alpha-1} n_\sigma (1-n_\alpha)}$$

此即

$$_F\langle n_1 - \delta_{1\alpha}\cdots n_\sigma - \delta_{\sigma\alpha}\cdots \mid a_\alpha \mid n_1\cdots n_\sigma\cdots\rangle_F = (-1)^{\sum_{\sigma=0}^{\alpha-1} n_\sigma n_\alpha} \tag{4.2.16}$$

从而

$$a_\alpha \mid n_0\cdots n_\sigma\cdots\rangle_F = (-1)^{\sum_{\sigma=0}^{\alpha-1} n_\sigma n_\alpha} \mid n_0 - \delta_{0\alpha}\cdots n_\sigma - \delta_{\sigma\alpha}\cdots\rangle_F \tag{4.2.17}$$

从上面(4.2.9)和(4.2.12)两个等式不难推出(请读者证明之)反对易关系

$$[a_\alpha, a_\beta^+]_+ \equiv a_\alpha a_\beta^+ + a_\beta^+ a_\alpha = \delta_{\alpha\beta}, [a_\alpha, a_\beta]_+ = [a_\alpha^+, a_\beta^+]_+ = 0 \tag{4.2.18}$$

以及基矢表达式

$$\mid n_0\cdots n_\sigma\cdots\rangle_F = (a_0^+)^{n_0}\cdots(a_\sigma^+)^{n_\sigma}\cdots \mid 0\cdots 0\cdots\rangle_F \tag{4.2.19}$$

这表明由真空态$|0\cdots 0\cdots\rangle_F$出发,作用以各种产生算符,可以得到基(4.1.24)中的所有基矢(不计较归一化因子)。全同费米子体系的态空间 F 可由$|0\cdots 0\cdots\rangle_F$和所有的$(a_0^+)^{n_0}\cdots(a_\sigma^+)^{n_\sigma}\cdots|0\cdots 0\cdots\rangle_F$张成。对每一个单粒态$|\alpha\rangle$,定义 F 上的线性算符

$$N_\alpha = a_\alpha^+ a_\alpha \tag{4.2.20}$$

这是一个厄米算符,其本征值只能是非负实数。不难证明

$$[N_\alpha, a_\beta] = -\delta_{\alpha\beta}a_\beta, [N_\alpha, a_\beta^+] = \delta_{\alpha\beta}a_\beta^+, [N_\alpha, N_\beta] = 0 \tag{4.2.21}$$

以及

$$N_\alpha \mid n_0\cdots n_\sigma\cdots\rangle_F = n_\alpha \mid n_0\cdots n_\sigma\cdots\rangle_F \tag{4.2.22}$$

因而我们称 N_α 为关于单粒态$|\alpha\rangle$的粒子数算符。所有的粒子数算符$\{N_\alpha \mid \alpha \geq 0\}$构成全同玻色子体系的一个完全力学量组。与玻色子情况不同,费米子的粒子数算符的本征值是 0 和 1,这表明费米子服从泡利不相容原理。我们称所有单粒态$|\alpha\rangle(\alpha \geq 0)$的粒子数算符 N_α 之和$\hat{N} = \sum_{0\leq\alpha} N_\alpha$,为总粒子数算符。

在上面的讨论中,不管是玻色子还是费米子的产生、消灭算子,都是针对单粒态空间 S 的某个选定的标准正交基的基矢定义的。所以我们要问,当选用不同的标准正交基时,产生、消灭算子将如何变换?设$\{v_\mu \mid \mu \geq 0\}$是单粒态空间 S 的另一标准正交基,与标准正交基$\{u_\alpha \equiv |\alpha\rangle \mid \alpha \geq 0\}$间的幺正变换关系为 $u_\alpha = \sum_{\mu\geq 0} v_\mu T_{\mu\alpha} = v_\mu T_{\mu\alpha}$,$\forall \alpha \geq 0$(采用爱因斯坦惯例,指标重复意味着求和)。以 d_μ^+, d_μ 记相应的产生消灭算子。为确定起见,先考虑费米子的情况。观察下面等式:

$$\begin{vmatrix} u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N) \\ u_{\alpha_2}(1) & u_{\alpha_2}(2) & \cdots & u_{\alpha_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N) \end{vmatrix} = T_{\mu_1\alpha_1}\cdots T_{\mu_N\alpha_N} \begin{vmatrix} v_{\mu_1}(1) & v_{\mu_1}(2) & \cdots & v_{\mu_1}(N) \\ v_{\mu_2}(1) & v_{\mu_2}(2) & \cdots & v_{\mu_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ v_{\mu_N}(1) & v_{\mu_N}(2) & \cdots & v_{\mu_N}(N) \end{vmatrix} \tag{4.2.23}$$

由产生算子的定义式(4.2.13)(注意,此式当 $\alpha_1 < \alpha_2 < \cdots < \alpha_N$ 不满足时也成立)可得

$$a_\alpha^+ \frac{1}{\sqrt{N!}} \begin{vmatrix} u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N) \\ u_{\alpha_2}(1) & u_{\alpha_2}(2) & \cdots & u_{\alpha_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N) \end{vmatrix}$$

$$= \frac{T_{\mu\alpha}T_{\mu_1\alpha_1}\cdots T_{\mu_N\alpha_N}}{\sqrt{(N+1)!}} \begin{vmatrix} v_\mu(1) & v_\mu(2) & \cdots & v_\mu(N+1) \\ v_{\mu_1}(1) & v_{\mu_1}(2) & \cdots & v_{\mu_1}(N+1) \\ \cdots & \cdots & \cdots & \cdots \\ v_{\mu_N}(1) & v_{\mu_N}(2) & \cdots & v_{\mu_N}(N+1) \end{vmatrix}$$

$$= \frac{T_{\mu\alpha}T_{\mu_1\alpha_1}\cdots T_{\mu_N\alpha_N}}{\sqrt{N!}} d_\mu^+ \begin{vmatrix} v_{\mu_1}(1) & v_{\mu_1}(2) & \cdots & v_{\mu_1}(N) \\ v_{\mu_2}(1) & v_{\mu_2}(2) & \cdots & v_{\mu_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ v_{\mu_N}(1) & v_{\mu_N}(2) & \cdots & v_{\mu_N}(N) \end{vmatrix}$$

$$= \frac{d_\mu^+ T_{\mu\alpha}}{\sqrt{N!}} \begin{vmatrix} u_{\alpha_1}(1) & u_{\alpha_1}(2) & \cdots & u_{\alpha_1}(N) \\ u_{\alpha_2}(1) & u_{\alpha_2}(2) & \cdots & u_{\alpha_2}(N) \\ \cdots & \cdots & \cdots & \cdots \\ u_{\alpha_N}(1) & u_{\alpha_N}(2) & \cdots & u_{\alpha_N}(N) \end{vmatrix}$$

于是

$$a_\alpha^+ = d_\mu^+ T_{\mu\alpha} \tag{4.2.24}$$

费米子产生算子与标准正交基的变换方式一样,经受同样的幺正变换(张量代数中称

作协变)。由式(4.2.24)可得费米子消灭算子按逆变方式变换

$$a_\alpha = (T^{-1})_{\alpha\mu} d_\mu \tag{4.2.25}$$

用类似方法可证玻色子产生算子和消灭算子在标准正交基的变换下,也分别按协变和逆变方式变换:

$$b_\alpha^+ = d_\mu^+ T_{\mu\alpha}, b_\alpha = (T^{-1})_{\alpha\mu} d_\mu \tag{4.2.26}$$

练习 4-2-1 由式(4.2.6)、(4.2.7)出发,试证对易关系$[b_\alpha, b_\beta^+] = b_\alpha b_\beta^+ - b_\beta^+ b_\alpha = \delta_{\alpha\beta}$, $[b_\alpha, b_\beta] = [b_\alpha^+, b_\beta^+] = 0$ 以及等式$|n_1 \cdots n_\sigma \cdots\rangle_B = \frac{(b_1^+)^{n_1} \cdots (b_\sigma^+)^{n_\sigma} \cdots}{\sqrt{n_1! \cdots n_\sigma! \cdots}} |0 \cdots 0 \cdots\rangle_B$ $(n_\sigma \geq 0)$

练习 4-2-2 试证:若 $N = N^+, [N, b] = -b, N|n\rangle = n|n\rangle$,则当 $b|n\rangle \neq 0 (b^+|n\rangle \neq 0)$时,它是 N 的属于特征值 $n-1(n+1)$的特征向量。

练习 4-2-3 由式(4.2.14)、(4.2.17)出发,证明反对易关系$[a_\alpha, a_\beta^+]_+ \equiv a_\alpha a_\beta^+ + a_\beta^+ a_\alpha = \delta_{\alpha\beta}$, $[a_\alpha, a_\beta]_+ = [a_\alpha^+, a_\beta^+]_+ = 0$ 以及等式$|n_1 \cdots n_\sigma \cdots\rangle_F = (a_1^+)^{n_1} \cdots (a_\sigma^+)^{n_\sigma} \cdots |0 \cdots 0 \cdots\rangle_F$ $(n_\sigma = 0, 1)$

练习 4-2-4 试求态空间 $B(F)$的真空态厄米投影算子,以产生、湮灭算子表达之。

练习 4-2-5 试证:$\{b_\sigma, b_\sigma^+ | \sigma \geq 0\}$($\{a_\sigma, a_\sigma^+ | \sigma \geq 0\}$)是不可约算符组,而且态矢量空间 $B(F)$上的任一线性算符 L 都可以表作它们的幂级数。

练习 4-2-6 试证:N 粒态子空间$(S^{\otimes N})_B$($(S^{\otimes N})_F$)是总粒子数算符 $\hat{N}$ 的不变子空间。

§4.3 用产生湮灭算子表达力学量算符

在本章开始时,我们曾提到量子力学传统的理论形式的缺陷:它不能描述粒子数变化的过程;在计算力学量算符对态矢量作用的中间过程中,不能始终保持态矢量的交换对称性(反对称性)。此外,在传统理论形式中力学量算符的表达式(4.1.27)与粒子数 N 有关。在本节,我们将证明,利用玻色子产生与湮灭算符可以把全同玻色

子体系的任一 k – 体力学量 T^k 统一地表成

$$T^k = \frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k,\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k\gamma_1\cdots\gamma_k} b^+_{\beta_k}\cdots b^+_{\beta_1} b_{\gamma_i}\cdots b_{\gamma_k} \tag{4.3.1}$$

式中

$$t_{\beta_1\cdots\beta_k\gamma_1\cdots\gamma_k} \equiv \langle \beta_1\cdots\beta_k \mid t \mid \gamma_1\cdots\gamma_k\rangle$$
$$= \int\cdots\int \mathrm{d}q_1\cdots\mathrm{d}q_k \bar{u}_{\beta_k}(q_k)\cdots\bar{u}_{\beta_1}(q_1)\, t(1,\cdots,k)\, u_{\gamma_1}(q_1)\cdots u_{\gamma_k}(q_k) \tag{4.3.2}$$

式(4.3.1)是一个与粒子数无关的表达式。玻色子产生与湮灭算符不会破坏态矢量的交换对称性。

我们还要证明,利用费米子产生与湮灭算符可以把全同费米子体系的任一 k – 体力学量 T^k 统一地表成与粒子数无关的形式

$$T^k = \frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k,\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k\gamma_1\cdots\gamma_k} a^+_{\beta_k}\cdots a^+_{\beta_1} a_{\gamma_i}\cdots a_{\gamma_k} \tag{4.3.3}$$

式中的 $t_{\beta_1\cdots\beta_k,\gamma_1\cdots\gamma_k}$ 为式(4.3.2)。费米子产生与湮灭算符不会破坏态矢量的交换反对称性。

Ⅰ. 全同玻色子体系的 k – 体算符

我们先证明全同玻色子体系的 k – 体算符表达式(4.3.1)的正确性。注意到它是定义在 Fock 空间 $B = \bigoplus_{N\ge 0}(S^{\otimes N})_B$ 上的线性算符。我们首先证明 N 粒态子空间 $(S^{\otimes N})_B$ 是算符(4.3.1) T^k 的不变子空间。再证明线性算符(4.3.1) T^k 对 $(S^{\otimes N})_B$ 的标准正交基(4.1.12)或(4.1.13)的作用与线性算符(4.1.27) T^k_N 对(4.1.12)或(4.1.13)的作用相同。

算符(4.3.1)的每一项所包含的产生算符的个数等于它所包含的消灭算符的个数,易证它与总粒子数算符对易。而 N 粒态子空间 $(S^{\otimes N})_B$ 是算符 $\hat{N}$ 的属于特征值 N 的特征子空间,从而是算符(4.3.1)的不变子空间。(线性算符 X 与 Y 对易,则 X 的特征子空间是 Y 的不变子空间,见 §0.2)算符(4.3.1)对标准正交基(4.1.12)或(4.1.13)中的任一基矢的作用为

$$\frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k,\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k\gamma_1\cdots\gamma_k} b^+_{\beta_k}\cdots b^+_{\beta_1} b_{\gamma_i}\cdots b_{\gamma_k} \mid n_0\cdots n_\sigma\cdots\rangle_B$$

$$= \frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k,\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k\gamma_1\cdots\gamma_k}\sqrt{n_{\gamma_k}}\sqrt{n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k}}\cdots\sqrt{n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\delta_{\gamma_1\gamma_3}-\cdots-\delta_{\gamma_1\gamma_k}}$$
$$\sqrt{1+n_{\beta_1}-\delta_{\beta_1\gamma_1}-\delta_{\beta_1\gamma_2}-\cdots-\delta_{\beta_1\gamma_k}}\sqrt{1+n_{\beta_2}-\delta_{\beta_2\gamma_1}-\delta_{\beta_2\gamma_2}-\cdots-\delta_{\beta_2\gamma_k}+\delta_{\beta_2\beta_1}}$$
$$\sqrt{1+n_{\beta_k}-\delta_{\beta_k\gamma_1}-\delta_{\beta_k\gamma_2}-\cdots-\delta_{\beta_k\gamma_k}+\delta_{\beta_k\beta_1}+\delta_{\beta_k\beta_2}+\cdots+\delta_{\beta_k\beta_{k-1}}}$$
$$|(n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k})\cdots$$
$$(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k})\cdots\rangle_B \tag{4.3.4}$$

而算符(4.1.27)对同一基矢的作用为

$$\frac{1}{k!}\sum_{\substack{1\le i_a\le N,\forall 1\le a\le k\\ i_b\ne i_c,\forall b\ne c}} t(i_1,\cdots,i_k)\,|\,n_0\cdots n_\sigma\cdots\rangle_B = \frac{1}{k!\sqrt{N!\prod_{a\ge0}(n_a!)}}$$
$$\times\sum_{P\in S_N}D(P)\sum_{\substack{1\le i_a\le N,\forall 1\le a\le k\\ i_b\ne i_c,\forall b\ne c}} t(i_1,\cdots,i_k)u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)$$
$$= \frac{1}{k!\sqrt{N!\prod_{a\ge0}(n_a!)}}\sum_{0\le\beta_1,\cdots,\beta_k}\sum_{\substack{1\le i_a\le N,\forall 1\le a\le k\\ i_r\ne i_s,(\forall r\ne s)}} t_{\beta_1\cdots\beta_k\alpha_{i_1}\cdots\alpha_{i_k}}$$
$$\sum_{P\in S_N}D(P)[u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)]\,|_{\alpha_{i_c}\mapsto\beta_c(\forall 1\le c\le k)} \tag{4.3.5}$$

上式右边的第二个求和是对从 N 个粒子中依次挑出 k 个来的所有不同排列方式求和，共有$\frac{N!}{(N-k)!}$项。其中使$\{\alpha_{i_1}\cdots\alpha_{i_k}\}=\{\gamma_1\cdots\gamma_k\}$的项的个数为 $n_{\gamma k}(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots(n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\cdots-\delta_{\gamma_1\gamma_k})$(参考本节的注)。这些项中的 $t_{\beta_1\cdots\beta_k\alpha_{i_1}\cdots\alpha_{i_k}}$显然相同，但这些项中$[u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)]|_{\alpha_{i_s}\mapsto\beta_c(\forall 1\le c\le k)}$一般而言不同，因为$\{i_1,\cdots,i_k\}$的取值不同。它们的 N 个单粒态构成一样，但排列顺序不同。然而对 N 个粒子所有不同的排列求和后，其和$\sum_{P\in S_N}D(P)[u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)]|_{\alpha_{i_c}=\gamma_c\mapsto\beta_c(\forall 1\le c\le k)}$相同。这个和可以写作

$$\sum_{P\in S_N}D(P)[u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)]\,|_{\alpha_{i_c}=\gamma_c\mapsto\beta_c(\forall 1\le c\le k)}$$
$$= \sqrt{N!\prod_{0\le\alpha}(n_\alpha-\delta_{\alpha\gamma_1}-\cdots-\delta_{\alpha\gamma_k}+\delta_{\alpha\beta_1}+\cdots+\delta_{\alpha\gamma_k})!}$$
$$|(n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k})\cdots$$
$$(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k})\cdots\rangle_B \tag{4.3.6}$$

于是(4.3.5)式中的第二个求和可以改为对 k 个单粒态$\{\gamma_1,\cdots,\gamma_k\}$求和

$$\frac{1}{k!}\sum_{\substack{1\le i_a\le N,\forall 1\le a\le k\\ i_r\neq i_s(\forall r\neq s)}} t(i_1,\cdots,i_k)\mid n_0\cdots n_\sigma\cdots\rangle_B$$

$$=\frac{1}{k!\sqrt{N!\prod_{\alpha\ge 0}(n_\alpha!)}}\sum_{0\le\beta_1,\cdots,\beta_k}\sum_{0\le\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k\alpha_{i_1}\cdots\alpha_{i_k}n_{\gamma_k}}$$

$$(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots(n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\cdots-\delta_{\gamma_1\gamma_k})$$

$$\sqrt{N!\prod_{0\le\alpha}(n_\alpha-\delta_{\alpha\gamma_1}-\cdots-\delta_{\alpha\gamma_k}+\delta_{\alpha\beta_1}+\cdots+\delta_{\alpha\gamma_k})!}$$

$$\mid(n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k})\cdots$$

$$(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k})\cdots\rangle_B$$

{考虑上式分母中的$\sqrt{\prod_{0\le\alpha}n_\alpha!}$与分子中的$n_{\gamma_k}(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots(n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\cdots-\delta_{\gamma_1\gamma_k})$如何相消。注意到$\{\gamma_1,\cdots,\gamma_k\}$中等于$\alpha$的下标的个数为$\delta_{\alpha\gamma_1}+\cdots+\delta_{\alpha\gamma_k}$，这个数就是乘积$n_{\gamma_k}(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots(n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\cdots-\delta_{\gamma_1\gamma_k})$中包含$n_\alpha$的因子的个数。这些因子的乘积是$n_\alpha(n_\alpha-1)\cdots(n_\alpha-\delta_{\alpha\gamma_1}-\cdots-\delta_{\alpha\gamma_k}+1)$。从而有

$$=\frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k}\sum_{0\le\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k\gamma_1\cdots\gamma_k}\sqrt{\frac{n_{\gamma_k}(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots(n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\cdots-\delta_{\gamma_1\gamma_k})}{\prod_{0\le\alpha}(n_\alpha-\delta_{\alpha\gamma_1}-\cdots-\delta_{\alpha\gamma_k})!}}$$

$$\sqrt{\prod_{0\le\alpha}(n_\alpha-\delta_{\alpha\gamma_1}-\cdots-\delta_{\alpha\gamma_k}+\delta_{\alpha\beta_1}+\cdots+\delta_{\alpha\beta_k})!}$$

$$\mid(n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k})\cdots$$

$$(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k})\cdots\rangle_B$$

现在考虑上式分母中的$\sqrt{\prod_{0\le\alpha}(n_\alpha-\delta_{\alpha\gamma_1}-\cdots-\delta_{\alpha\gamma_k})!}$如何与分子中的$\sqrt{\prod_{0\le\alpha}(n_\alpha-\delta_{\alpha\gamma_1}-\cdots-\delta_{\alpha\gamma_k}+\delta_{\alpha\beta_1}+\cdots+\delta_{\alpha\beta_k})!}$相消。注意到，若下标$\alpha$不出现在$\{\beta_1,\cdots,\beta_k\}$中，则分子分母中的包含$n_\alpha$的因子相消。消不掉的因子必须与$\{\beta_1,\cdots,\beta_k\}$有关。它们的乘积是$\sqrt{(n_{\beta_1}-\delta_{\beta_1\gamma_1}-\cdots-\delta_{\beta_1\gamma_k}+1)(n_{\beta_2}-\delta_{\beta_2\gamma_1}-\cdots-\delta_{\beta_2\gamma_k}+\delta_{\beta_1\beta_2}+1)\cdots(n_{\beta_k}-\delta_{\beta_k\gamma_1}-\cdots-\delta_{\beta_k\gamma_k}+\delta_{\beta_1\beta_k}+\cdots+\delta_{\beta_{k-1}\beta_k}+1)}$。这使我们得到

$$\frac{1}{k!}\sum_{i_r\neq i_s,(\forall r\neq s)} t(i_1,\cdots,i_k)\mid n_0\cdots n_\sigma\cdots\rangle_B$$

$$= \frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k,\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k,\gamma_1\cdots\gamma_k}\sqrt{n_{\gamma_k}}\sqrt{n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k}}\cdots$$
$$\sqrt{n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\delta_{\gamma_1\gamma_3}-\cdots-\delta_{\gamma_1\gamma_k}}\sqrt{1+n_{\beta_1}-\delta_{\beta_1\gamma_1}-\delta_{\beta_1\gamma_2}-\cdots-\delta_{\beta_1\gamma_k}}$$
$$\sqrt{1+n_{\beta_2}-\delta_{\beta_2\gamma_1}-\delta_{\beta_2\gamma_2}-\cdots-\delta_{\beta_2\gamma_k}+\delta_{\beta_2\beta_1}}\cdots$$
$$\sqrt{1+n_{\beta_k}-\delta_{\beta_k\gamma_1}-\delta_{\beta_k\gamma_2}-\cdots-\delta_{\beta_k\gamma_k}+\delta_{\beta_k\beta_1}+\delta_{\beta_k\beta_2}+\cdots+\delta_{\beta_k\beta_{k-1}}}$$
$$|(n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k})\cdots$$
$$(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k})\cdots\rangle_B \tag{4.3.7}$$

比较(4.3.4)与(4.3.7),我们看到 $T^k|_{(S^{\otimes N})_B}=T^k_N$,也就是说定义在 B 上的线性算符(4.3.1) T^k 到其不变子空间 $(S^{\otimes N})_B$ 上的限制,等于定义在 $(S^{\otimes N})_B$ 上的线性算符(4.1.27) T^k_N。

Ⅱ. 全同费米子体系的 k – 体算符

我们现在证明全同费米子体系的 k – 体算符表达式(4.3.3)的正确性。注意到它是定义在 Fock 空间 $F=\bigoplus_{N\ge0}(S^{\otimes N})_F$ 上的线性算符。我们首先证明 N 粒态子空间 $(S^{\otimes N})_F$ 是算符(4.3.3)的不变子空间。再证明算符(4.3.3)对 $(S^{\otimes N})_F$ 的标准正交基(4.1.19)作用与算符(4.1.27)对(4.1.19)的作用相同。

N 粒态子空间 $(S^{\otimes N})_F$ 是费米子总粒子数算符 $\hat{N}$ 的特征子空间,算符(4.3.3)与 $\hat{N}$ 交换,因而 $(S^{\otimes N})_F$ 是算符(4.3.3)的不变子空间。

算符(4.3.3)对标准正交基(4.1.19)中的任一基矢的作用为

$$\frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k,\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k,\gamma_1\cdots\gamma_k}a^+_{\beta_k}\cdots a^+_{\beta_1}a_{\gamma_i}\cdots a_{\gamma_k}|n_0\cdots n_\sigma\cdots\rangle_F$$
$$= \frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k,\gamma_1,\cdots,\gamma_k} t_{\beta_1\cdots\beta_k,\gamma_1\cdots\gamma_k}(-1)^{\sum_{\sigma=0}^{\gamma_k-1}n_\sigma}n_{\gamma_k}(-1)^{\sum_{\sigma=0}^{\gamma_{k-1}-1}(n_\sigma-\delta_{\sigma\gamma_k})}(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots$$
$$(-1)^{\sum_{\sigma=0}^{\gamma_1-1}(n_\sigma-\delta_{\sigma\gamma_2}-\cdots-\delta_{\sigma\gamma_k})}$$
$$(n_{\gamma_1}-\delta_{\gamma_1\gamma_k}-\cdots-\delta_{\gamma_1\gamma_2})(-1)^{\sum_{\sigma=0}^{\beta_1-1}(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k})}$$
$$[1-(n_{\beta_1}-\delta_{\beta_1\gamma_1}-\cdots-\delta_{\beta_1\gamma_k})](-1)^{\sum_{\sigma=0}^{\beta_2-1}(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\gamma_1})}$$

$$[1-(n_{\beta_2}-\delta_{\beta_2\gamma_k}-\cdots-\delta_{\beta_2\gamma_1}+\delta_{\beta_2\beta_1})]\cdots(-1)^{\sum_{\sigma=0}^{\beta_k-1}(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_{k-1}})}$$
$$[1-(n_{\beta_k}-\delta_{\beta_k\gamma_k}-\cdots-\delta_{\beta_k\gamma_1}+\delta_{\beta_k\beta_1}+\cdots+\delta_{\beta_k\beta_{k-1}})]$$
$$|(n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k})\cdots$$
$$(n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k})\cdots\rangle_F \tag{4.3.8}$$

而算符(4.1.27)对同一基矢的作用为

$$\frac{1}{k!}\sum_{\substack{1\le i_a\le N,\forall 1\le a\le k\\ i_r\ne i_s,(\forall r\ne s)}} t(i_1,\cdots,i_k)\,|\,n_0\cdots n_\sigma\cdots\rangle_F$$
$$=\frac{1}{k!\sqrt{N!}}\sum_{P\in S_N}D(P)(-1)^{v(P)}\sum_{\substack{1\le i_a\le N,(\forall 1\le a\le k)\\ i_r\ne i_s(\forall r\ne s)}} t(i_1,\cdots,i_k)u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)$$
$$=\frac{1}{k!\sqrt{N!}}\sum_{0\le\beta_1,\cdots,\beta_k}\sum_{i_r\ne i_s,(\forall r\ne s)} t_{\beta_1\cdots\beta_k\alpha_{i_1}\cdots\alpha_{i_k}}\sum_{P\in S_N}D(P)(-1)^{v(P)}[u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)]$$
$$|_{\alpha_{i_s}\mapsto\beta_c(\forall 1\le c\le k)} \tag{4.3.9}$$

上式中的第二个求和是对从 N 个粒子中依次挑出 k 个来的所有不同排列方式求和，共有$\frac{N!}{(N-k)!}$项。其中使$\{\alpha_{i_1}\cdots\alpha_{i_k}\}=\{\gamma_1\cdots\gamma_k\}$的项的个数为 $n_{\gamma_k}(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots(n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\cdots-\delta_{\gamma_1\gamma_k})$。然而与玻色子情况不同，在费米子情况有 $0\le\alpha_1<\alpha_2<\cdots<\alpha_N$，亦即 $\sum_{\sigma\ge0}n_\sigma=N, 0\le n_\sigma\le1, \forall\sigma\ge0$。因而其中使$\{\alpha_{i_1}\cdots\alpha_{i_k}\}=\{\gamma_1\cdots\gamma_k\}$的项的个数 $n_{\gamma_k}(n_{\gamma_{k-1}}-\delta_{\gamma_{k-1}\gamma_k})\cdots(n_{\gamma_1}-\delta_{\gamma_1\gamma_2}-\cdots-\delta_{\gamma_1\gamma_k})$，只能是 0 或 1。当它为 1 时，相应的 $\sum_{P\in S_N}D(P)(-1)^{v(P)}[u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)]\,|_{\alpha_{i_s}\mapsto\beta_c(\forall 1\le c\le k)}$ 与$\sqrt{N!}\,|\,n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k},\cdots,n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k},\cdots\rangle_F$ 相差一个正负号。让我们考察这一正负号。从行列式

$$\sum_{P\in S_N}(-1)^{v(P)}D(P)u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)$$
$$=\begin{vmatrix} u_{\alpha_1}(1) & \cdots & u_{\gamma_k}(1) & \cdots & u_{\gamma_1}(1) & \cdots & u_{\alpha_N}(1)\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ u_{\alpha_1}(N) & \cdots & u_{\gamma_k}(N) & \cdots & u_{\gamma_1}(N) & \cdots & u_{\alpha_N}(N)\end{vmatrix}$$

出发，注意$\{\alpha_{i_1}\cdots\alpha_{i_k}\}=\{\gamma_1\cdots\gamma_k\}$，以及$0\le\alpha_1<\alpha_2<\cdots<\alpha_N$，把 u_{γ_σ}换成u_{β_σ}后，一般而

言，单粒态下标不再随列数增加而增加。挪动行列式的列，使单粒态下标随列数增加而增加，须交换列的次数的奇偶性与交换的具体方式无关。我们这样挪动列：把 u_{γ_σ} 换成 u_{β_σ} 后（$\sigma=1,\cdots k$），先把 u_{β_k} 向左移到第一列，这需要交换列 $\sum_{\alpha=0}^{\gamma_k-1} n_\alpha$ 次；接着把 $u_{\beta_{k-1}}$ 向左移到第二列，这需要交换列 $\sum_{\alpha=0}^{\gamma_{k-1}-1}(n_\alpha-\delta_{\alpha\gamma_k})$ 次；依次类推，把 u_{β_1} 向左移到第 k 列，这需要交换列 $\sum_{\alpha=0}^{\gamma_1-1}(n_\alpha-\delta_{\alpha\gamma_k}-\cdots-\delta_{\alpha\gamma_2})$ 次；现在再把 u_{β_1} 反方向（向右）移到它应该占的列，这显然要交换 $\sum_{\alpha=0}^{\beta_1-1}(n_\alpha-\delta_{\alpha\gamma_k}-\cdots-\delta_{\alpha\gamma_1})$ 次；再把 u_{β_2} 反方向移到它应该占的列，显然要交换 $\sum_{\alpha=0}^{\beta_2-1}(n_\alpha-\delta_{\alpha\gamma_k}-\cdots-\delta_{\alpha\gamma_1}+\delta_{\alpha\beta_1})$ 次，依次类推，最后把 u_{β_k} 向右移到它应该占的列，这要交换 $\sum_{\alpha=0}^{\beta_k-1}(n_\alpha-\delta_{\alpha\gamma_k}-\cdots-\delta_{\alpha\gamma_1}+\delta_{\alpha\beta_1}+\cdots+\delta_{\alpha\beta_{k-1}})$ 次。以 X 记列的交换次数的总和，我们有

$$\begin{aligned}&\frac{1}{\sqrt{N!}}\sum_{P\in S_N}D(P)(-1)^{v(P)}[u_{\alpha_1}(1)\cdots u_{\alpha_N}(N)]\mid_{a_{i_s}\mapsto\beta_c(\forall 1\le c\le k)}\\&=(-1)^X\mid n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k},\cdots,\\&\quad n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k},\cdots\rangle_F\end{aligned}$$

从而可把(4.3.9)式的第二个求和改为对单粒态 $\gamma_1,\cdots,\gamma_k$ 求和

$$\begin{aligned}&\frac{1}{k!}\sum_{i_r\neq i_s,(\forall r\neq s)}t(i_1,\cdots,i_k)\mid n_0\cdots n_\sigma\cdots\rangle_F\\&=\frac{1}{k!}\sum_{0\le\beta_1,\cdots,\beta_k}\sum_{0\le\gamma_1,\cdots,\gamma_k}t_{\beta_1\cdots\beta_k a_{i_1}\cdots a_{i_k}}(-1)^{\sum_{\alpha=0}^{\gamma_k-1}n_\sigma}(-1)^{\sum_{\alpha=0}^{\gamma_{k-1}-1}(n_\alpha-\delta_{\alpha\gamma k})}\cdots\\&(-1)^{\sum_{\alpha=0}^{\gamma_1-1}(n_\alpha-\delta_{\alpha\gamma 2}-\cdots-\delta_{\alpha\gamma k})}(-1)^{\sum_{\alpha=0}^{\beta_1-1}(n_\alpha-\delta_{\alpha\gamma 1}-\cdots-\delta_{\alpha\gamma k})}\\&(-1)^{\sum_{\alpha=0}^{\beta_2-1}(n_\alpha-\delta_{\alpha\gamma 1}-\cdots-\delta_{\alpha\gamma k}+\delta_{\alpha\beta_1})}\cdots(-1)^{\sum_{\alpha=0}^{\beta_k}(n_\alpha-\delta_{\alpha\gamma 1}-\cdots-\delta_{\alpha\gamma k}+\delta_{\alpha\beta 1}+\cdots+\delta_{\alpha\beta_{k-1}})}\\&\mid n_0-\delta_{0\gamma_1}-\cdots-\delta_{0\gamma_k}+\delta_{0\beta_1}+\cdots+\delta_{0\beta_k},\cdots,\\&\quad n_\sigma-\delta_{\sigma\gamma_1}-\cdots-\delta_{\sigma\gamma_k}+\delta_{\sigma\beta_1}+\cdots+\delta_{\sigma\beta_k},\cdots\rangle_F\end{aligned}\tag{4.3.10}$$

比较式(4.3.8)与(4.3.10),我们看到 $T^k|_{(S^{\otimes N})_F}=T_N^k$,也就是说定义在 F 上的线性算符(4.3.3) T^k 到其不变子空间 $(S^{\otimes N})_F$ 上的限制,等于定义在 $(S^{\otimes N})_F$ 上的线性算符(4.1.27) T_N^k。

从上面的讨论可以看到,二次量子化形式的 k – 体算符(4.3.1)、(4.3.3)与传统形式的 k – 体算符(4.1.27) $T_N^k(N=0,1,2\cdots)$对 N 粒态基矢的作用完全相同。二者差别仅在于后者是定义在各自的 N 粒态子空间上的,而前者是定义在各个 N 粒态子空间的直和上的。

Ⅲ. 写下全同粒子体系的 k – 体算符的示例

作为例子,我们来讨论全同电子体系的角动量算符二次量子化形式。选取氢原子中电子的哈密顿算符 $H=T-\frac{Ze^2}{r}$,轨道角动量平方 L^2,轨道角动量在 z – 方向的投影 L_z,和自旋角动量在 z – 方向的投影 S_z 的共同的本征函数系 $\left\{\varphi_{nlm_lm_s}\middle|n\geq 1,l=0,1,\cdots,n-1;m_l=l,l-1,\cdots,-l;m_s=\frac{1}{2},-\frac{1}{2}\right\}$为单粒态空间 S 的标准正交基。相应的粒子产生、湮灭算符记作 $a^+_{nlm_lm_s}$, $a_{nlm_lm_s}$。全同费米子体系的轨道角动量算符 $\boldsymbol{L}$ 是单体算符。按照式(4.3.3)

$$\begin{aligned}L_z&=\sum_{n'l'm_l'm_s'nlm_lm_s}\langle n'l'm_l'm_s'\mid l_z\mid nlm_lm_s\rangle a^+_{n'l'm_l'm_s'}a_{nlm_lm_s}\\&=\sum_{n'l'm_l'm_s'nlm_lm_s}m_l\delta_{n'n}\delta_{l'l}\delta'_{m_l'm_l}\delta_{m_s'm_s}a^+_{n'l'm_l'm_s'}a_{nlm_lm_s}=\sum_{nlm_lm_s}m_la^+_{nlm_lm_s}a_{nlm_lm_s}\\&=\sum_{nlm_lm_s}m_l\hat{n}_{nlm_lm_s}\end{aligned}\tag{4.3.11}$$

上式告诉我们,当所选的单粒态空间的基矢是 $\hat{l}_z$ 的特征向量时,体系的角动量算符 L_z 是各个单粒态的粒子数算符 $\hat{n}_{nlm_lm_s}$ 与所属 $\hat{l}_z$ 的特征值 m_l 乘积的和。

$$\begin{aligned}L_\pm&=\sum_{n'l'm_l'm_s'nlm_lm_s}\langle n'l'm_l'm_s'\mid l_\pm\mid nlm_lm_s\rangle a^+_{n'l'm_l'm_s'}a_{nlm_lm_s}\\&=\sum_{n'l'm_l'm_s'nlm_lm_s}\sqrt{(l\mp m_l)(l\pm m_l+1)}\,\delta_{n'n}\delta_{l'l}\delta_{m_l'm_{l\pm1}}\delta_{m_s'm_s}a^+_{n'l'm_l'm_s'}a_{nlm_lm_s}\\&=\sum_{nlm_lm_s}\sqrt{(l\mp m_l)(l\pm m_l+1)}\,a^+_{nlm_l\pm1m_s}a_{nlm_lm_s}\end{aligned}\tag{4.3.12}$$

由 $L_x=\frac{1}{2}(L_+ + L_-)$, $L_y=\frac{1}{2i}(L_+ - L_-)$可得

$$L_x = \frac{1}{2}\Big(\sum_{nlm_lm_s}\sqrt{(l-m_l)(l+m_l+1)}\,a^+_{nlm_l+1m_s}a_{nlm_lm_s} + \sum_{nlm_lm_s}\sqrt{(l+m_l)(l-m_l+1)}\,a^+_{nlm_l-1m_s}a_{nlm_lm_s}\Big)$$

$$L_y = \frac{1}{2i}\Big(\sum_{nlm_lm_s}\sqrt{(l-m_l)(l+m_l+1)}\,a^+_{nlm_l+1m_s}a_{nlm_lm_s} - \sum_{nlm_lm_s}\sqrt{(l+m_l)(l-m_l+1)}\,a^+_{nlm_l-1m_s}a_{nlm_lm_s}\Big)$$

值得注意的是,体系轨道角动量平方 $\boldsymbol{L}^2$ 既不是单体算符,又不是二体算符。这由它的传统形式的表达式可清楚地看出。事实上,

$$\boldsymbol{L}^2 = \boldsymbol{L}\cdot\boldsymbol{L} = \sum_{i=1}^{N}\boldsymbol{l}(i)\cdot\sum_{j=1}^{N}\boldsymbol{l}(j) = \sum_{i=1}^{N}\boldsymbol{l}^2(i) + \sum_{i\neq j}\boldsymbol{l}(i)\cdot\boldsymbol{l}(j) \quad (4.3.13)$$

上式右边的第一项是单体算符,第二项是二体算符。按照式(4.3.3),第一项为

$$\sum_{n'l'm_l'm_s'nlm_lm_s}\langle n'l'm_l'm_s'\mid \boldsymbol{l}^2\mid nlm_lm_s\rangle a^+_{n'l'm_l'm_s'}a_{nlm_lm_s}$$

$$= \sum_{n'l'm_l'm_s'nlm_lm_s} l(l+1)\delta_{n'n}\delta_{l'l}\delta_{m_l'm_l}\delta_{m_s'm_s}a^+_{n'l'm_l'm_s'}a_{nlm_lm_s} = \sum_{nlm_lm_s} l(l+1)\,a^+_{nlm_lm_s}a_{nlm_lm_s} \quad (4.3.14)$$

为求第二项,注意到 $\sum_{i\neq j}\boldsymbol{l}(i)\cdot\boldsymbol{l}(j) = \frac{1}{2}\sum_{i\neq j}\{l_+(i)l_-(j) + l_-(i)l_+(j) + 2l_z(i)l_z(j)\}$,第二项为$\frac{1}{2}\sum_{\beta_1\beta_2\gamma_1\gamma_2}\langle n_1'l_1'm_{l1}'m_{s1}'n_2'l_2'm_{l2}'m_{s2}'\mid l_+(1)l_-(2) + l_-(1)l_+(2) + 2l_z(1)l_z(2)\mid n_1l_1m_{l1}m_{s1}n_2l_2m_{l2}m_{s2}\rangle a^+_{\beta_2}a^+_{\beta_1}a_{\gamma_1}a_{\gamma_2}$,其中 $\beta_i=\{n_i'l_i'm_{li}'m_{si}'\}$, $\gamma_i\equiv\{n_il_im_{li}m_{si}\}\ \forall\, i=1,2$。计算上式中的矩阵元得第二项为

$$\frac{1}{2}\sum_{\gamma_1\gamma_2}\sqrt{(l_1-m_{l1})(l_1+m_{l1}+1)(l_2+m_{l2})(l_2-m_{l2}+1)}\;a^+_{n_2l_2m_{l2}-1m_{s2}}a^+_{n_1l_1m_{l1}+1m_{s1}}a_{n_1l_1m_{l1}m_{s1}}a_{n_2l_2m_{l2}m_{s2}}$$

$$\frac{1}{2}\sum_{\gamma_1\gamma_2}\sqrt{(l_1+m_{l1})(l_1-m_{l1}+1)(l_2-m_{l2})(l_2+m_{l2}+1)}\;a^+_{n_2l_2m_{l2}+1m_{s2}}a^+_{n_1l_1m_{l1}-1m_{s1}}a_{n_1l_1m_{l1}m_{s1}}a_{n_2l_2m_{l2}m_{s2}}$$

$$+\frac{1}{2}\sum_{\gamma_1\gamma_2}2m_{l1}m_{l2}a^{+}_{n_2l_2m_{l2}m_{s2}}a^{+}_{n_1l_1m_{l1}m_{s1}}a_{n_1l_1m_{l1}m_{s1}}a_{n_2l_2m_{l2}m_{s2}} \tag{4.3.15}$$

为计算轨道角动量平方算符，也可以从关系式 $\boldsymbol{L}^2=L_+L_-+L_z^2-L_z(=L_-L_++L_z^2+L_z)$出发，把 $L_\pm$，L_z 的二次量子化表达式(4.3.11)和(4.3.12)代入即可。这是因为，传统形式量子力学中 N 个全同粒子体系的力学量算符间的关系，二次量子化后继续成立。不同方法都导致同一结果：

$$\begin{aligned}\boldsymbol{L}^2=&\sum_{n_1l_1m_{l1}m_{s1}n_2l_2m_{l2}m_{s2}}m_{l2}m_{l1}a^{+}_{n_2l_2m_{l2}m_{s2}}a_{n_2l_2m_{l2}m_{s2}}a^{+}_{n_1l_1m_{l1}m_{s1}}a_{n_1l_1m_{l1}m_{s1}}\\&-\sum_{nlm_lm_s}m_la^{+}_{nlm_lm_s}a_{nlm_lm_s}+\\&\sum_{n_2l_2m_{l2}m_{s2}}\sqrt{(l_2-m_{l2})(l_2+m_{l2}+1)}\,a^{+}_{n_2l_2m_{l2}+1m_{s2}}a_{n_2l_2m_{l2}m_{s2}}\\&\sum_{n_1l_1m_{l1}m_{s1}}\sqrt{(l_1+m_{l1})(l_1-m_{l1}+1)}\,a^{+}_{n_1l_1m_{l1}-1m_{s1}}a_{n_1l_1m_{l1}m_{s1}}\\=&\sum_{n_1l_1m_{l1}m_{s1}n_2l_2m_{l2}m_{s2}}\sqrt{(l_1+m_{l1})(l_1-m_{l1}+1)}\sqrt{(l_2-m_{l2})(l_2+m_{l2}+1)}\\&a^{+}_{n_2l_2m_{l2}+1m_{s2}}a^{+}_{n_1l_1m_{l1}-1m_{s1}}a_{n_1l_1m_{l1}m_{s1}}a_{n_2l_2m_{l2}m_{s2}}\\&+\sum_{n_1l_1m_{l1}m_{s1}n_2l_2m_{l2}m_{s2}}m_{l_2}m_{l_1}a^{+}_{n_2l_2m_{l2}m_{s2}}a^{+}_{n_1l_1m_{l1}m_{s1}}a_{n_1l_1m_{l1}m_{s1}}a_{n_2l_2m_{l2}m_{s2}}\\&+\sum_{nlm_lm_s}l(l+1)a^{+}_{nlm_lm_s}a_{nlm_lm_s}。\end{aligned}$$

从上面算符 L_z 的例子可以看出一个一般规律：任一单体算符 $T_N^1=\sum_{1\le i\le N}t(i)$，若选择算符 t 的正交归一化的特征向量组$\{u_\alpha\mid\alpha\ge0,tu_\alpha=\tau_\alpha u_\alpha,\langle u_\alpha\mid u_\beta\rangle=\delta_{\alpha\beta}\}$为单粒态空间的基，把相应的粒子产生、湮灭算符记作 a_α^+，a_α(b_α^+，b_α)，则有 $T^1=\sum_{\alpha\ge0}\tau_\alpha\hat{n}_\alpha$，式中的 $\hat{n}_\alpha$ 为粒子数算符 $\hat{n}_\alpha=a_\alpha^+a_\alpha(b_\alpha^+b_\alpha)$。

练习 4-3-1 选取氢原子中电子的哈密顿算符 $H=T-\frac{Ze^2}{r}$，轨道角动量平方 L^2，总角动量平方 J^2，和总角动量在 z - 方向的投影 J_z 的共同的本征函数系 $\left\{\varphi_{nljm}\middle|n\ge1,l=0,1,\cdots,n-1;j=l+\frac{1}{2},\cdots,\left|l-\frac{1}{2}\right|;m=j,\cdots,-j\right\}$为单粒态空间 S 的标准正交基。相应的粒子产生、湮灭算符记作 a^+_{nljm}，a_{nljm}。试用它们表出电子体系的总角动量算符 J^2，J_z；并证明$\{a^+_{n'l'j'm'}a^+_{nljm}\{|\rangle|-j'\le m'\le j',-j\le m\le j\}$构成 J^2，

J_z 本征函数非耦合基，$\left\{\sum_{m'm}\langle j'm'jm|j'jJM\rangle a^+_{n'l'j'm'}a^+_{nljm}|\rangle \| j-j'|\leq J\leq j+j', -J\leq M\leq J\right\}$ 构成耦合基。

注 设有 N 个箱子(编号为 $1,2,\cdots,N$)，在每个箱子里放置一个非负整数。这 N 个非负整数 $\alpha_1,\alpha_2,\cdots,\alpha_N$ 可以有重复。任一非负整数 σ 在这 N 个非负整数中出现的次数记作 $n_\sigma\left(=\sum_{j=1}^{N}\delta_{\sigma\alpha_j}\right)$。若从这 N 个箱子中先后挑出 k 个箱子(编号为 $i_1,i_2,\cdots,i_k$)，则我们得到放置在它们中的 k 个非负整数的有序组 $\alpha_{i_1},\alpha_{i_2}\cdots.,\alpha_{i_k}$。试问，当穷尽所有 $\frac{N!}{(N-k)}$ 个不同挑法时，k 个非负整数的有序组 $\gamma_1,\gamma_2,\cdots,\gamma_k$ 出现多少次?

挑第 1 个箱子有 N 种不同挑法：$i=1,2,\cdots,N$。其中使 $\alpha_{i_1}=\gamma_1$ 的挑法显然有 n_{γ_1} 种。现在假定已挑出第 1 个箱子，其中非负整数为 γ_1。在剩下的 $N-1$ 个箱子中，任一非负整数 σ 出现的次数变作 $n_\sigma-\delta_{\sigma\gamma_1}$。这时挑选第 2 个箱子有 $N-1$ 种挑法，使其中非负整数为 γ_2 的挑法显然有 $n_{\gamma_2}-\delta_{\gamma_1\gamma_2}$ 种。假定挑出了头两个箱子，其中两个非负整数为 $\gamma_1\gamma_2$，在剩下的 $N-2$ 个箱子中挑第 3 个箱子。显然使其中非负整数为 γ_3 的挑法有 $n_{\gamma_3}-\delta_{\gamma_1\gamma_3}-\delta_{\gamma_2\gamma_3}$ 种。继续挑选箱子，最后可得，在这 $\frac{N!}{(N-k)}$ 个挑出的有序组中，$\gamma_1,\gamma_2,\cdots,\gamma_k$ 出现的次数为 $n_{\gamma_1}(n_{\gamma_2}-\delta_{\gamma_1\gamma_2})(n_{\gamma_3}-\delta_{\gamma_1\gamma_3'}-\delta_{\gamma_2\gamma_3})\cdots(n_{\gamma_k}-\delta_{\gamma_1\gamma_k}-n_{\gamma_2\gamma_k}-\cdots-\delta_{\gamma_{k-1}\gamma_k})$。

§4.4 二次量子化量子力学的运动方程

在前几节我们给出了全同粒子体系的扩大了的态空间，在扩大了的态空间上引入了粒子产生湮灭算符，给出了全同粒子体系的任一 k - 体力学量算符的产生湮灭算符表达式，我们还指出所有由若干产生算符作用在真空态矢量得到的态矢量张成整个态空间。所有这些都属于运动学范畴。现在让我们来讨论动力学问题。在二次量子化理论形式的量子力学中，全同粒子体系的态矢量的演化仍遵守薛定格方程

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H(t)\,|\,\Psi(t)\rangle \tag{4.4.1}$$

式中的态矢量是扩大了的态空间中的向量，式中的哈密顿算符是扩大了的态空间上的线性算符，体系的动力学性质完全由它确定。如果我们不从根本上改变体系的哈密顿（也就是说仍采用传统形式量子力学中体系的哈密顿算符，只是用产生湮灭算符表达之），当体系的初态具有确定的粒子数时，则二次量子化形式与传统形式的量子力学完全等价。既然这样我们要问：（ⅰ）引入二次量子化又有什么必要和好处呢？（ⅱ）我们并没有把量子力学的物理进一步量子化，为什么把新形式称为二次量子化？

Ⅰ. 二次量子化对传统形式量子力学的继承与发展

事实上，不论是传统形式的或者是二次量子化形式的量子力学，体系状态的演化都遵守薛定格方程，体系的动力学性质都由其哈密顿算符决定。在前面我们已经证明了传统形式的 k – 体算符与二次量子化形式的 k – 体算符（用产生消灭算符表出的 k – 体算符）完全等价。自然，如果不改变哈密顿算符（不改变薛定格方程）和初始条件（初态仍是粒子数取确定值的态），二次量子化形式不会得出有别于传统形式的结果。特别是，由于哈密顿算符的每一项包含的产生算符的个数等于消灭算符的个数，这使得哈密顿算符与总粒子数算符对易，粒子数守恒，我们无法得到粒子数不守恒的过程。新的物理只能来自新的哈密顿。传统形式的量子力学只能容纳各种 k – 体算符，不能容纳使粒子数不守恒的哈密顿算符。然而，由于二次量子化理论形式中，态空间扩大了，允许引入使粒子数不守恒的哈密顿算符，即允许哈密顿算符的一些项包含的产生算符的个数不等于消灭算符的个数。这种项与所有的 k – 体算符或它们的和都不同。由于有这种项，哈密顿算符与总粒子数算符不再对易，粒子数不再守恒。

下面我们将介绍使粒子数不守恒的哈密顿算符的例子。

Ⅱ. 二次量子化一词的来源

为清楚起见，我们分别讨论玻色子、费米子两种情况。先讨论玻色子情况。在本章前几节讨论的基础之上，构作依赖粒子坐标和自旋 $(\boldsymbol{r}, s_z) \equiv q$ 的算符（我们将称之为波算符）

$$\sum_{\alpha \geq 0} b_\alpha u_\alpha(q) \equiv \Phi(q) \tag{4.4.2}$$

式中 b_α 是玻色子的湮灭算符。取式(4.4.2)的厄米共轭得

$$\sum_{\alpha\geq 0}\overline{u_\alpha}(q)b_\alpha^+ = \Phi^+(q) \tag{4.4.3}$$

利用函数系$\{u_\alpha(q)\mid\alpha\geq 0\}$的正交归一性,由式(4.4.2)和(4.4.3)可得

$$b_\alpha = \int\overline{u_\alpha}(q)\Phi(q)\mathrm{d}q = \langle u_\alpha\mid\Phi\rangle, b_\alpha^+ = \int\Phi^+(q)u_\alpha(q)\mathrm{d}q = \langle\Phi\mid u_\alpha\rangle \tag{4.4.4}$$

利用函数系$\{u_\alpha(q)\mid\alpha\geq 0\}$的完备性,由玻色子产生、湮灭算符的对易关系(4.2.8)可得

$$[\Phi(q),\Phi^+(q')] = \sum_{\alpha\geq 0}\sum_{\beta\geq 0}[b_\alpha,b_\beta^+]u_\alpha(q)\overline{u_\beta}(q') = \sum_{\alpha\geq 0}u_\alpha(q)\overline{u_\alpha}(q') = \delta(q-q')$$

$$[\Phi(q),\Phi(q')] = 0,[\Phi^+(q),\Phi^+(q')] = 0 \tag{4.4.5}$$

反之,由波算符的对易关系(4.4.5)和式(4.4.4)可得玻色子产生、湮灭算符的对易关系(4.2.8)。

由$b_\alpha|\rangle_B = 0, \forall\alpha\geq 0$,可知对所有$q$值都有

$$\Phi(q)|\rangle_B = 0 \tag{4.4.6}$$

不难看出,利用波算符$\Phi(q)$,全同玻色子体系的任一k-体力学量T^k(4.3.1)可以表作

$$T^k = \frac{1}{k!}\int\cdots\int\mathrm{d}q_1\cdots\mathrm{d}q_k\Phi^+(q_k)\cdots\Phi^+(q_1)t(1\cdots k)\Phi(q_1)\cdots\Phi(q_k) \tag{4.4.7}$$

这只要把(4.3.2)代入(4.3.1)即可。反之,把(4.4.2)、(4.4.3)代入(4.4.7)就能得到(4.3.1)。

前面我们曾指出过,全同玻色子体系的任一态矢量可以写作$|0\cdots 0\cdots\rangle_B$和所有的$(b_1^+)^{n_1}\cdots(b_\sigma^+)^{n_\sigma}\cdots|0\cdots 0\cdots\rangle_B$的线性组合。由$b_\alpha^+ = \int\Phi^+(q)u_\alpha(q)\mathrm{d}q$可知,全同玻色子体系的任一态矢量$\eta$可以写作$|0\cdots 0\cdots\rangle_B$和所有$\Phi^+(q_1)\cdots\Phi^+(q_s)\cdots|0\cdots 0\cdots\rangle_B$的线性组合:

$$\eta = c_0|\rangle_B + \sum_{s\geq 1}\int\cdots\int dq_1\cdots dq_s c_s(q_1\cdots q_s)\Phi^+(q_1)\cdots\Phi^+(q_s)|\rangle_B \tag{4.4.8}$$

反之,若全同玻色子体系的任一态矢量可以写作$|0\cdots 0\cdots\rangle_B$和所有$\Phi^+(q_1)\cdots\Phi^+(q_s)\cdots|0\cdots 0\cdots\rangle_B$的线性组合,则由$\sum_{\alpha\geq 1}\overline{u_\alpha}(q)b_\alpha^+ \equiv \Phi^+(q)$可知全同玻色子体系的任一态矢量$\eta$可以写作$|0\cdots 0\cdots\rangle_B$和所有的$(b_1^+)^{n_1}\cdots(b_\sigma^+)^{n_\sigma}\cdots|0\cdots 0\cdots\rangle_B$的线性组

合。

上面的讨论启示我们,全同玻色子体系的二次量子化也可以按相反的方向实现:首先把单粒子在坐标和自旋$(\boldsymbol{r}, s_z) \equiv q$表象中的波函数$\varphi(q)$变作作用在某个态矢量空间上的波算符$\Phi(q)$(正是因为这一点,历史上把这一做法称作二次量子化),令它们满足下面的对易关系

$$[\Phi(q), \Phi^+(q')] = \delta(q - q'), [\Phi(q), \Phi(q')] = 0, [\Phi^+(q), \Phi^+(q')] = 0, \tag{4.4.9}$$

引入真空态$|\rangle_B$,即一个粒子也没有的量子态,使${}_B\langle|\rangle_B = 1$。令对所有的$q$值

$$\Phi(q)|\rangle_B = 0 \tag{4.4.10}$$

并且$|\rangle_B$和所有的$\Phi^+(q_1)\cdots\Phi^+(q_s)|0\cdots0\cdots\rangle_B (s \geq 1)$张成全同玻色子体系的态空间。也就是说,体系的任一态矢量η总可以写作

$$\eta = c_0|\rangle_B + \sum_{s\geq 1}\int\cdots\int dq_1\cdots dq_s c_s(q_1\cdots q_s)\Phi^+(q_1)\cdots\Phi^+(q_s)|\rangle_B \tag{4.4.11}$$

体系的任一 k - 体力学量算符由波算符按下式给出:

$$T^k = \frac{1}{k!}\int\cdots\int \mathrm{d}q_1\cdots\mathrm{d}q_k\Phi^+(q_k)\cdots\Phi^+(q_1)t(1\cdots k)\Phi(q_1)\cdots\Phi(q_k) \tag{4.4.12}$$

由上面这些假定出发,我们可以得到前两节全同玻色子体系二次量子化的全部结果。事实上,设$\{u_\alpha(q)|\alpha \geq 0\}$是单粒态空间的任一正交归一基,把波算符$\Phi(q)$向它们展开

$$\Phi(q) = \sum_{\alpha\geq 0} b_\alpha u_\alpha(q) \tag{4.4.13}$$

式中的展开系数$b_\alpha (\alpha \geq 0)$是算符。(4.4.13)的厄米共轭为

$$\sum_{\alpha\geq 0}\overline{u}_\alpha(q)b_\alpha^+ \equiv \Phi^+(q) \tag{4.4.14}$$

利用$\{u_\alpha(q)|\alpha \geq 0\}$的正交归一性,由上两式可得

$$b_\alpha = \int\overline{u}_\alpha(q)\Phi(q)dq = \langle u_\alpha|\Phi\rangle, b_\alpha^+ = \int\Phi^+(q)u_\alpha(q)dq = \langle\Phi^+|u_\alpha\rangle \tag{4.4.15}$$

由(4.4.10)可知$b_\alpha|\rangle_B = 0, \forall\alpha \geq 0$。由波算符的对易关系(4.4.9)可得

$$[b_\alpha, b_\beta] = 0, [b_\alpha^+, b_\beta^+] = 0, [b_\alpha, b_\beta^+] = \left[\int \bar{u}_\alpha(q)\Phi(q)dq, \int \Phi^+(q')u_\beta(q')dq'\right]$$
$$= \iint dqdq'\bar{u}_\alpha(q')u_\beta(q')[\Phi(q), \Phi(q')] = \iint dqdq'\bar{u}_\alpha(q)u_\beta(q')\delta(q-q')$$
$$= \int dq\bar{u}_\alpha(q)u_\beta(q) = \langle u_\alpha \mid u_\beta\rangle = \delta_{\alpha\beta} \tag{4.4.16}$$

由 $|\rangle_B$ 和所有的 $\Phi^+(q_1)\cdots\Phi^+(q_s)|0\cdots0\cdots\rangle_B (s\geq 1)$ 张成全同玻色子体系的态空间，考虑到(4.4.14)，可知全同玻色子体系的任一态矢量可以写作 $|0\cdots0\cdots\rangle_B$ 和所有的 $(b_1^+)^{n_1}\cdots(b_\sigma^+)^{n_\sigma}\cdots|0\cdots0\cdots\rangle_B$ 的线性组合。

现在讨论费米子情况。把单粒子在坐标和自旋 $(\boldsymbol{r}, s_z)\equiv q$ 表象中的波函数 $\psi(q)$ 变作作用在某个态矢量空间上的波算符 $\Psi(q)$（正是因为这一点，历史上把这一做法称作二次量子化），令它们满足下面的反对易关系

$$\{\Psi(q), \Psi(q')\} = 0, \{\Psi^+(q), \Psi^+(q')\} = 0, \{\Psi(q), \Psi^+(q')\} = \delta(q-q') \tag{4.4.17}$$

引入真空态 $|\rangle_F$，即一个粒子也没有的量子态，使 ${}_F\langle|\rangle_F = 1$，并且令对所有的 q 值

$$\Psi(q)|\rangle_F = 0 \tag{4.4.18}$$

$|\rangle_F$ 和所有的 $\Psi^+(q_1)\cdots\Psi^+(q_s)|0\cdots0\cdots\rangle_F (s\geq 1)$ 张成全同费米子体系的态空间。也就是说，体系的任一态矢量 ξ 总可以写作

$$\xi = c_0|\rangle_F + \sum_{s\geq 1}\int\cdots\int dq_1\cdots dq_s c_s(q_1\cdots q_s)\Psi^+(q_1)\cdots\Psi^+(q_s|\rangle_F \tag{4.4.19}$$

体系的任一 k - 体力学量算符由波算符按下式给出：

$$T^k = \frac{1}{k!}\int\cdots\int dq_1\cdots dq_k\Psi^+(q_k)\cdots\Psi^+(q_1)t(1\cdots k)\Psi(q_1)\cdots\Psi(q_k) \tag{4.4.20}$$

由上面这些假定出发，我们可以得到前两节全同费米子体系二次量子化的全部结果。事实上，设 $\{u_\alpha(q)|\alpha\geq 0\}$ 是单粒态空间的任一正交归一基，把波算符 $\Psi(q)$ 向它们展开

$$\Psi(q) = \sum_{\alpha\geq 0} a_\alpha u_\alpha(q) \tag{4.4.21}$$

式中的展开系数 $a_\alpha (\alpha\geq 0)$ 是算符。(4.4.21)的厄米共轭为

$$\sum_{\alpha\geq 0}\bar{u}_\alpha a_\alpha^+ \equiv \Psi^+(q) \tag{4.4.22}$$

利用$\{u_\alpha(q)\mid\alpha\geq 0\}$的正交归一性,由上两式可得

$$a_\alpha=\int\overline{u}_\alpha(q)\Psi(q)dq=\langle u_\alpha\mid\Psi\rangle,a_\alpha^+=\int\Psi^+(q)u_\alpha(q)dq=\langle\Psi^+\mid u_\alpha\rangle \tag{4.4.23}$$

由(4.4.18)可知 $a_\alpha|\rangle_B=0,\forall\alpha\geq 0$。由波算符的反对易关系(4.4.17)可得

$$\begin{aligned}&\{b_\alpha,b_\beta\}=0,\{b_\alpha^+,b_\beta^+\}=0,\{b_\alpha,b_\beta^+\}=\left\{\int\overline{u}_\alpha(q)\Psi(q)dq,\int\Psi^+(q')u_\beta(q')dq'\right\}\\&=\iint dqdq'\overline{u}_\alpha(q)u_\beta(q')\{\Psi(q),\Psi^+(q')\}=\iint dqdq'\overline{u}_\alpha(q)u_\beta(q')\delta(q-q')\\&=\int dq\overline{u}_\alpha(q)u_\beta(q)=\langle u_\alpha\mid u_\beta\rangle=\delta_{\alpha\beta}\end{aligned} \tag{4.4.24}$$

由$|\rangle_F$和所有的$\Psi^+(q_1)\cdots\Psi^+(q_s)|0\cdots0\cdots\rangle_F(s\geq 1)$张成全同费米子体系的态空间,考虑到(4.4.25),可知全同费米子体系的任一态矢量可以写作$|0\cdots0\cdots\rangle_F$和所有的$(a_1^+)^{n_1}\cdots(a_\sigma^+)^{n_\sigma}|0\cdots0\cdots\rangle_F$的线性组合。

Ⅲ. 求解二次量子化形式的薛定格方程的示例(原子与光场相互作用的 Jaynes-Commings 模型)

通常的量子体系是由若干种全同粒子构成的。例如,当我们研究电子与辐射电磁场的相互作用时,体系是由全同费米子(电子)和全同玻色子(光子)构成的。体系的态空间是全同电子体系的态空间与全同光子体系的态空间的张量积,$B_{el}\otimes F_{ph}$。一般而言,当我们研究的微观过程涉及若干种全同粒子时,体系的态空间应选作这几种全同粒子态空间的张量积。这种态空间是很大的。但往往在实际问题中,对所讨论的具体过程,体系的态矢量 $\Psi(t)$只在这种态空间的一个子空间 W 中变化(只要$\Psi(0)\in W$,而且 W 是哈密顿算符$H(t)$的不变子空间,就能保证这一点)。如果 W 还是所有感兴趣的力学量算符的不变子空间,我们就可以把 W 当作体系的态空间,把所有感兴趣的力学量算符看作是 W 上的算符,从而把问题大大地简化。

在下面,我们将通过原子与光场相互作用的具体例子向读者介绍如何使用二次量子化形式的量子力学处理物理问题。

这个问题所讨论的量子体系是(处于核库仑场中的)电子体系和光子体系。它们

的自旋角量子数分别为$\frac{1}{2}$和1,故分别是费米子和玻色子。我们的态空间是$F_{el}\otimes B_{ph}$。光子的静质量为零,它们永远是相对论性的粒子。严格地处理光子的运动,需要用量子场论,超出了本教材的范围。按照量子场论,电子与电子间的电磁相互作用是通过交换光子(发射与吸收光子)来实现的。整个体系的哈密顿由三部分组成:处在原子核库仑场中的电子体系的能量(电子间没有相互作用,因而这是个单体算符),自由光子体系的能量和电子与光子的相互作用能(电子间的电磁相互作用由它描述)。由上一节最后一段的讨论可知,若选择核库仑场中单电子能量本征函数$\{u_\alpha \mid \alpha \geq 0\}$为电子单粒态空间的标准正交基,则处在原子核库仑场中的电子体系的哈密顿算符为$H_{el}=\sum_{\alpha\geq 0}\epsilon_\alpha a_\alpha^+ a_\alpha$(对核库仑场中电子的各个能级求和),自由光子体系的哈密顿算符为$H_{ph}=\sum_{k\sigma}\hbar\omega(\boldsymbol{k})b_{k\sigma}^+ b_{\sigma k}$(对各种动量$\boldsymbol{k}$和偏振$\sigma=1,2$求和),而电子与光子的相互哈密顿算符可以近似地写作

$$H_I=\hbar\sum_{\alpha\neq\beta}\sum_{k\sigma}a_\alpha^+ a_\beta(g_{\alpha\beta k\sigma}b_{k\sigma}+\bar{g}_{\beta\alpha k\sigma}b_{k\sigma}^+) \tag{4.4.25}$$

其中$g_{\alpha\beta k\sigma}$是耦合常数。当我们考虑二能级原子与单模光场的相互作用时,体系的哈密顿可以近似地写作(对详细推导感兴趣的读者可参考 D. F. Walls, G. J. Milburn: Quantum Optics)

$$H=H_0+H_1=(\epsilon_0 a_0^+ a_0+\epsilon_1 a_1^+ a_1+\hbar\omega b^+ b)+\hbar g(a_0^+ a_1 b^+ + a_1^+ a_0 b) \tag{4.4.26}$$

式中耦合常数$g=\bar{g}$是实数。上式称作 Jaynes-Commings 模型,其右边最后一项是二能级原子与单模光场相互作用哈密顿H_1。它包含的两项分别代表:电子从激发态(能级ϵ_1)跃迁到基态(能级ϵ_0)并放出一个光子,电子吸收一个光子并从基态(能级ϵ_0)跃迁到激发态(能级ϵ_1)。

让我们考察一下体系的态空间。从前面的讨论可知,$\{\hat{n}_0\equiv a_0^+ a_0, \hat{n}_1\equiv a_1^+ a_1, \hat{n}_1\equiv b^+ b\mid$构成体系的一个完备的力学量组。也就是说,它们的共同的特征向量组$\{\mid n_0,n_1\rangle_F\otimes\mid n\rangle_B \mid n_0,n_1=0,1;n\geq 0\}$构成体系态空间的一个正交归一基。设体系在$t=0$时刻处在态$\Psi(0)=\mid 0,1\rangle_F\otimes\mid N\rangle_B$,我们要问体系在$t$时刻的状态如何:$\Psi(t)=?$

注意到$\hat{n}_0+\hat{n}_1, \hat{n}_1+\hat{n}$是运动积分(请读者证明之),而且体系在初始时刻的态

矢量 $\Psi(0)$是这两个运动积分的本征态：$(\hat{n}_0+\hat{n}_1)\Psi(0)=\Psi(0)$，$(\hat{n}_1+\hat{n})\Psi(0)=(N+1)\Psi(0)$。因而对所有的 $t\geq 0$，$(\hat{n}_0+\hat{n}_1)\Psi(t)=1\Psi(t)$，$(\hat{n}_1+\hat{n})\Psi(t)=(N+1)\Psi(t)$。我们可以把问题限制在算符 $\hat{n}_0+\hat{n}_1$，$\hat{n}_1+\hat{n}$ 的属于特征值 1，$N+1$ 的共同的特征子空间来考虑。这是一个二维子空间，选 $|1,0\rangle_F\otimes|N+1\rangle_B\equiv u_1$，$|0,1\rangle_F\otimes|N\rangle_B\equiv u_2$ 为它的基。在这个基下

$$[H_0]=\begin{bmatrix}\varepsilon_0+\hbar\omega(N+1) & 0\\ 0 & \varepsilon_1+\hbar\omega N\end{bmatrix},[H_1]=\hbar g\sqrt{N+1}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}=\hbar g\sqrt{N+1}\sigma_x$$

$$[\Psi(0)]=\begin{bmatrix}0\\ 1\end{bmatrix},[\Psi(t)]=\begin{bmatrix}c_1(t)\\ c_2(t)\end{bmatrix} \tag{4.4.27}$$

先讨论一个重要的简单情况：当 $\varepsilon_1-\varepsilon_0\equiv\hbar\omega_0=\hbar\omega$，体系处于谐振。容易证明此时

$$[H_0]=(\varepsilon_0+\hbar\omega(N+1))\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{4.4.28}$$

于是演化算符的矩阵为

$$[U(t,0)]=\left[\exp\frac{-i}{\hbar}(H_0+H_I)t\right]=\exp\left[-i\left(\frac{\varepsilon_0}{\hbar}+\omega(N+1)\right)t\right]\exp(-i\Omega\sigma_x t)$$

式中的 $\Omega\equiv g\sqrt{N+1}$，注意到 $\sigma_x^2=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$，于是

$$\exp(-i\Omega\sigma_x t)=\sum_{n=0,2,\cdots}\frac{(-i\Omega t)^n}{n!}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+$$

$$\sum_{n=1,3,\cdots}\frac{(-i\Omega t)^n}{n!}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}=\begin{bmatrix}\cos\Omega t & -i\sin\Omega t\\ -i\sin\Omega t & \cos\Omega t\end{bmatrix}$$

$$[U(t,0)]=\exp\left[-i\left(\frac{\varepsilon_0}{\hbar}+\omega(N+1)\right)t\right]\begin{bmatrix}\cos\Omega t & -i\sin\Omega t\\ -i\sin\Omega t & \cos\Omega t\end{bmatrix} \tag{4.4.29}$$

我们得到

$$[\Psi(t)]=\exp\left[-i\left(\frac{\varepsilon_0}{\hbar}+\omega(N+1)\right)t\right]\begin{bmatrix}-i\sin\Omega t\\ \cos\Omega t\end{bmatrix} \tag{4.4.30}$$

可见二能级原子与单模光场相互作用达到谐振时，电子在基态和激发态之间来回振荡（*Rabi* 振荡），Ω 叫做 Rabi 振荡频率。

回到一般情况：$\varepsilon_1-\varepsilon_0\equiv\hbar\omega_0$ 不一定等于 $\hbar\omega$，这时

$$[H]=\left(\varepsilon_0+\hbar\omega N+\frac{\hbar(\omega+\omega_0)}{2}\right)\begin{bmatrix}1&0\\0&1\end{bmatrix}+\hbar\begin{bmatrix}\delta&\Omega\\\Omega&-\delta\end{bmatrix},\left(\delta\equiv\frac{\omega-\omega_0}{2}\right)\tag{4.4.31}$$

为计算$[H]$的本征值和本征向量，只须求解$\begin{bmatrix}\delta&\Omega\\\Omega&-\delta\end{bmatrix}$的本征方程。后者的本征值为$\pm\sqrt{\delta^2+\Omega^2}$，分别对应于本征向量

$$A_1\begin{bmatrix}\sqrt{\delta^2+\Omega^2}+\delta\\\Omega\end{bmatrix}\equiv[v_1],A_2\begin{bmatrix}\sqrt{\delta^2+\Omega^2}-\delta\\-\Omega\end{bmatrix}\equiv[v_2]\tag{4.4.32}$$

其中 $A_{1,2}\equiv[2(\delta^2+\Omega^2\pm\delta\sqrt{\delta^2+\Omega^2})]^{-\frac{1}{2}}$。$[v_1],[v_2]$也是$[H]$的正交归一化的特征向量，分别属于特征值

$$\left(\varepsilon_0+\hbar\omega N+\frac{\hbar(\omega+\omega_0)}{2}\right)\pm\hbar\sqrt{\delta^2+\Omega^2})]\equiv E_{1,2}\tag{4.4.33}$$

把初始时刻的态矢量$[\Psi(0)]$向$[H]$的特征向量$[v_1],[v_2]$展开：

$$[\Psi(0)]=\begin{bmatrix}0\\1\end{bmatrix}=\frac{\Omega}{2(\sqrt{\delta^2+\Omega^2}+\delta)\sqrt{\delta^2+\Omega^2}}\begin{bmatrix}\sqrt{\delta^2+\Omega^2}+\delta\\\Omega\end{bmatrix}$$
$$-\frac{\Omega}{2(\sqrt{\delta^2+\Omega^2}-\delta)\sqrt{\delta^2+\Omega^2}}\begin{bmatrix}\sqrt{\delta^2+\Omega^2}-\delta\\-\Omega\end{bmatrix}$$

于是

$$[\Psi(t)]=\frac{\Omega}{2(\sqrt{\delta^2+\Omega^2}+\delta)\sqrt{\delta^2+\Omega^2}}\begin{bmatrix}\sqrt{\delta^2+\Omega^2}+\delta\\\Omega\end{bmatrix}e^{-\frac{i}{\hbar}E_1t}$$
$$-\frac{\Omega}{2(\sqrt{\delta^2+\Omega^2}-\delta)\sqrt{\delta^2+\Omega^2}}\begin{bmatrix}\sqrt{\delta^2+\Omega^2}-\delta\\-\Omega\end{bmatrix}e^{-\frac{i}{\hbar}E_2t}$$
$$=\frac{\Omega}{2\sqrt{\delta^2+\Omega^2}}\begin{bmatrix}e^{-\frac{i}{\hbar}E_1t}-e^{-\frac{i}{\hbar}E_2t}\\\dfrac{\Omega e^{-\frac{i}{\hbar}E_1t}}{\sqrt{\delta^2+\Omega^2}+\delta}+\dfrac{\Omega e^{-\frac{i}{\hbar}E_2t}}{\sqrt{\delta^2+\Omega^2}-\delta}\end{bmatrix}$$
$$=\exp\left(-i\left(\frac{\varepsilon_0}{\hbar}+\omega N+\frac{\omega+\omega_0}{2}\right)t\right)\frac{1}{\sqrt{\delta^2+\Omega^2}}$$

$$\begin{bmatrix} -i\Omega\sin\sqrt{\Omega^2+\delta^2}\,t \\ \sqrt{\Omega^2+\delta^2}\cos\sqrt{\Omega^2+\delta^2}\,t + i\delta\sin\sqrt{\Omega^2+\delta^2}\,t \end{bmatrix} \tag{4.4.34}$$

当 $\omega=\omega_0$ 时，上式回到谐振情况(4.4.30)。当 $\omega\simeq\omega_0$，即稍有失谐时，由上式可知电子部分地在基态和激发态之间振荡，而且振荡的频率$\sqrt{\delta^2+\Omega^2}$比 Rabi 振荡频率 Ω 大。然而，当 ω 与ω_0 相差较大时，Jaynes-Commings 模型不再是好的近似，由它推出的(4.4.34)不复成立。(读者可参考 D.F.Walls, G.J.Milburn: Quantum Optics)

练习 4-4-1 利用 Jaynes-Commings 模型，讨论双能级原子与单模光场的相互作用。若体系在 $t=0$ 时刻处于态 $\Psi(0)=|1,0\rangle_F\otimes|N\rangle_B$，求 $\Psi(t)=?$（分 $N=0, N\geq 1$ 两种情况讨论）

第 5 章　相对论性单粒子波动方程

物理学对自然规律的描述不应停留在定性的层面上。但要定量地刻画物理体系的状态和运动就不得不引入参考坐标系。物理规律是客观的，而选择哪一个具体的参考坐标系是主观的。一个好的物理理论应具有协变性，就是说，对客观自然规律的正确表述(通常体现为运动方程)不应包含参考坐标系选择这样的主观因素。正是在这种思想的指引下，上一世纪初，爱因斯坦先后创建了狭义(1905)和广义(1916)相对论。相对论可以划入经典物理学的范畴，但它的关于时间和空间的观念，关于物理理论协变性的观念，远远超出了经典物理学。20 世纪 20 年代中期，几乎在非相对论量子力学创建的同时，人们就开始尝试建立相对论性的量子力学。事实上，Schrödinger 方程是 1926 年提出的，而描述无旋粒子的相对论性波动方程 Klein-Gordon 方程和描述电子的相对论性波动方程 Dirac 方程是 1926 至 1928 年间提出的。后两个方程克服了 Schrödinger 方程不具有相对论协变性的缺点，但也带来了负几率、负能量等困难。此外，它们并不能描述高能领域中普遍存在的粒子产生、湮灭过程。直到 1934 年，Pauli 和 Weisskopf 指出，不应把 Klein-Gordon 方程和 Dirac 方程当作单粒子的相对论性波动方程，而应把它们和经典电磁场的 Maxwell 方程同样看待，看成是经典场(无限大自由度体系)的方程。把经典场量子化就可以得到描述自旋为 $0,\frac{1}{2}\hbar,\hbar$ 的粒子体系的相对论性量子理论(量子场论)。量子场论克服了负几率、负能量等困难。成功地描述了各种高能过程，构成了基本粒子物理的理论基础。但它已超出了本教程的范围。本章将介绍作为单粒子相对论性波动方程的 Klein-Gordon 方程和 Dirac 方程。尽管它们是过渡性的理论，但它们在历史上曾引导人们认识电子自旋、电子内禀磁矩、自旋 - 轨道耦合和氢原子光谱的精细结构，并预言了正电子的存在。1932 年，Anderson 在宇宙线中找到了 Dirac 理论预言的这种电子的反粒子。我们介绍相对论

性单粒子波动方程,还因为它是学习量子场论不可或缺的准备。

§5.1 Klein-Gordon 方程

在牛顿力学(非相对论经典力学)中,自由质点的哈密顿量为

$$H = E = \frac{\boldsymbol{p}^2}{2m} \tag{5.1.1}$$

把上式中的哈密顿量(能量)和动量作如下的替换[①]

$$H = E \mapsto i\hbar \frac{\partial}{\partial t}, \quad \boldsymbol{p} = -i\hbar \nabla \tag{5.1.2}$$

并作用于波函数 $\psi(\boldsymbol{r}, t)$ 上我们便得到非相对论量子力学中自由无旋粒子的 *Schrödinger* 方程

$$i\hbar \frac{\partial}{\partial t}\psi(\boldsymbol{r}, t) = \frac{-\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r}, t) \tag{5.1.3}$$

由(5.1.3)我们可以得到关于几率的连续性方程

$$\frac{\partial}{\partial t}\rho(\boldsymbol{r}, t) + \nabla \cdot \boldsymbol{j}(\boldsymbol{r}, t) = 0 \tag{5.1.4}$$

其中

$$\begin{aligned} \rho(\boldsymbol{r}, t) &= \psi^*(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t), \boldsymbol{j}(\boldsymbol{r}, t) \\ &= \frac{i\hbar}{2m}[(\nabla\psi^*(\boldsymbol{r}, t))\psi(\boldsymbol{r}, t) - \psi^*(\boldsymbol{r}, t)\nabla\psi(\boldsymbol{r}, t)] \end{aligned} \tag{5.1.5}$$

Schrödinger 方程(5.1.3)包含波函数对 t 的一阶微商,而包含对 x, y, z 的二阶微商,时空坐标处在不平等的地位,故不可能是相对论协变的。这是因为它的出发点是非相对论质点的能量动量关系(5.1.1)。为了得到相对论性自由粒子的波动方程,我们应从相对论力学中自由质点的能量动量关系

$$E^2 = c^2\boldsymbol{p}^2 + m^2c^4 \tag{5.1.6}$$

出发,作替换(5.1.2)并作用于波函数 $\phi(\boldsymbol{r}, t)$ 上,我们便得到了自由无旋粒子的相对

---

① 与 $-i\hbar\nabla$ 不同,$i\hbar\frac{\partial}{\partial t}$ 不是体系态矢量空间 S 上的线性算符。

论性波动方程，Klein-Gordon 方程

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\phi(\boldsymbol{r},t) - \nabla^2\phi(\boldsymbol{r},t) + \frac{m^2c^2}{\hbar^2}\phi(\boldsymbol{r},t) = 0 \tag{5.1.7}$$

首先让我们来察看一下它的相对论不变性。为此我们采用狭义相对论中的记法（参看附录 §6 Lorentz 群）

$$x^0 = ct, x^1 = x, x^2 = y, x^3 = z$$

$$G = [g_{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

$$G^{-1} = [g^{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$x_\rho = g_{\rho\sigma}x^\sigma, x^\mu = g^{\mu\nu}x_\nu \tag{5.1.8}$$

$$x_0 = x^0 = ct, x_1 = -x^1 = -x, x_2 = -x^2 = -y, x_3 = -x^3 = -z$$

上式中的 G 称作度规张量。从惯性系 $\sum$ 变到惯性系 $\sum'$ 时，同一客观事件的时空坐标的变换关系为非齐次 Lorentz 变换

$$x'^\mu = L^\mu_\rho x^\rho + a^\mu \tag{5.1.9}$$

（L^μ_ρ, a^μ 与 x 无关是时空均匀性的要求）。光速不变原理要求

$$c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2 = c^2(\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2$$

此即 $g_{\mu\nu}\Delta x'^\mu \Delta x'^\nu = g_{\mu\nu}\Delta x^\mu \Delta x^\nu$。左式对所有 Δx 都成立，考虑到度规张量 G 的对称性，可得 $g_{\mu\nu}L^\mu_\rho L^\nu_\sigma = g_{\rho\sigma}$。此式可以写成矩阵形式 $L^TGL = G$。取两边的逆得

$$L^{-1}G^{-1}(L^{-1})^T = G^{-1} \tag{5.1.10}$$

记 $\phi(\boldsymbol{r},t) = \phi(x)$，Klein-Gordon 方程可以重新写作

$$g^{\mu\nu}\frac{\partial^2\phi(x)}{\partial x^\mu \partial x^\nu} + \frac{m^2c^2}{\hbar^2}\phi(x) = 0 \tag{5.1.11}$$

从惯性系 $\sum$ 变到惯性系 $\sum'$ 时，自由无旋粒子的波函数的值应保持不变 $\phi(x) = \phi'(x')$，或者说自由无旋粒子的波函数按 Lorentz 群的标量场表示变换。由时空坐标的

变换关系 $x'^{\mu} = L^{\mu}_{\rho}x^{\rho} + a^{\mu}$ 可得

$$\frac{\partial}{\partial x^{\rho}} = L^{\mu}_{\rho}\frac{\partial}{\partial x'^{\mu}} \tag{5.1.12}$$

其逆变换为

$$(L^{-1})^{\rho}_{\mu}\frac{\partial}{\partial x^{\rho}} = \frac{\partial}{\partial x'^{\mu}} \tag{5.1.13}$$

从而有

$$g^{\mu\nu}\frac{\partial^2}{\partial x'^{\mu}\partial x'^{\nu}} = g^{\mu\nu}(L^{-1})^{\rho}_{\mu}(L^{-1})^{\sigma}_{\nu}\frac{\partial^2}{\partial x^{\rho}\partial x^{\sigma}} = g^{\mu\nu}\frac{\partial^2}{\partial x^{\mu}\partial x^{\nu}}$$

最后一步利用了(5.1.10)。所以我们可由式(5.1.11)推得

$$g^{\mu\nu}\frac{\partial^2\phi'(x')}{\partial x'^{\mu}\partial x'^{\nu}} + \frac{m^2c^2}{\hbar^2}\phi'(x') = 0 \tag{5.1.14}$$

可见 Klein-Gordon 方程在不同惯性系中具有相同的形式。而这正是我们改造 Schrödinger 方程的初衷。在非相对论极限情况$\left(\frac{|\boldsymbol{v}|}{c} \ll 1\right)$下，式(6.1.6)给出 $E^2 = m^2c^2 + \frac{\boldsymbol{p}^2}{2m}$。令

$$\phi(\boldsymbol{r},t) = \psi(\boldsymbol{r},t)\exp\left(-i\frac{mc^2}{\hbar}t\right)$$

代入(5.1.7)得

$$-\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2}\phi(\boldsymbol{r},t) = [-\hbar^2\nabla^2\psi(\boldsymbol{r},t) + m^2c^2\psi(\boldsymbol{r},t)]\exp\left(-i\frac{mc^2}{\hbar}t\right)$$

而取式(5.1.2)两边对 t 的二阶微商得

$$-\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2}\phi(\boldsymbol{r},t) = \left[\frac{-\hbar^2}{c^2}\frac{\partial^2\psi(\boldsymbol{r},t)}{\partial t^2} + 2mi\hbar\frac{\partial\psi(\boldsymbol{r},t)}{\partial t} + m^2c^2\psi(\boldsymbol{r},t)\right]\exp\left(-i\frac{mc^2}{\hbar}t\right)$$

比较以上两式得

$$i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = \frac{-\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r},t) + \frac{\hbar^2}{2mc^2}\frac{\partial^2\psi(\boldsymbol{r},t)}{\partial t^2}$$

最后一项近似地等于$\frac{T^2}{2mc^2}\psi(\boldsymbol{r},t)$($T$ 是动能，远小于静止质量对应的能量 mc^2)，故可以略去，我们得到非相对论的 Schrödinger 方程(5.1.3)。尽管 Klein-Gordon 方程象我们期望地那样具有相对论协变性，又在非相对论近似下给出 Schrödinger 方程。但

是下面将会看到,它也带来了很大困难。首先由(5.1.7)也可以得到几率守恒定律(5.1.4),其中几率流密度向量与式(5.1.5)中的相同,但几率密度却是

$$\rho(\boldsymbol{r},t)=\frac{i\hbar}{2mc^2}\left[(\phi^*(\boldsymbol{r},t))\frac{\partial}{\partial t}\phi(\boldsymbol{r},t)-\frac{\partial\phi^*(\boldsymbol{r},t)}{\partial t}\phi(\boldsymbol{r},t)\right] \tag{5.1.15}$$

与 $\psi^*(\boldsymbol{r},t)\psi(\boldsymbol{r},t)$非负不同,它可以在四维时空的某些区域取负值。而这是与几率的概念不相容的。其次,式(5.1.7)是关于时间 t 的二阶偏微分方程。它的解要由初始时刻的波函数 $\phi(\boldsymbol{r},t)|_{t=t_0}$和波函数对时间的导数$\frac{\partial}{\partial t}\phi(\boldsymbol{r},t)|_{t=t_0}$来确定。如果我们坚持 t 时刻粒子状态完全由 $\phi(\boldsymbol{r},t)$刻画,按因果律自由粒子 t_0 时刻的状态 $\phi(\boldsymbol{r},t_0)$应能确定它以后 t 时间的状态。于是 Klein-Gordon 方程(5.1.7)与因果律矛盾。此外,Klein-Gordon 方程还有更严重的负能量困难。事实上,总可以把(5.1.7)的解 $\phi(\boldsymbol{r},t)$向动量算符的本征态展开

$$\phi(\boldsymbol{r},t)=\frac{1}{(2\pi\hbar)^{3/2}}\int d^3p f(\boldsymbol{p},t)e^{\frac{i\boldsymbol{p}\cdot\boldsymbol{r}}{\hbar}} \tag{5.1.16}$$

代入(5.1.7)可得

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}f(\boldsymbol{p},t)+\frac{\boldsymbol{p}^2+m^2c^2}{\hbar^2}f(\boldsymbol{p},t)=0 \tag{5.1.17}$$

这个常微分方程的通解为

$$f(\boldsymbol{p},t)=a(\boldsymbol{p})e^{-i\frac{E_+(\boldsymbol{p})t}{\hbar}}+b(\boldsymbol{p})e^{-i\frac{E_-(\boldsymbol{p})t}{\hbar}} \tag{5.1.18}$$

式中 $a(\boldsymbol{p}),b(\boldsymbol{p})$是积分常数,可以是 $\boldsymbol{p}$ 的任意函数,而

$$E_\pm(\boldsymbol{p})=\pm\sqrt{c^2\boldsymbol{p}^2+m^2c^4} \tag{5.1.19}$$

把式(5.1.18)代入(5.1.16)得到 Klein-Gordon 方程的一般解为

$$\phi(t,t)=\frac{1}{(2\pi\hbar)^{3/2}}\int d^3p\left[a(\boldsymbol{p})e^{i\frac{\boldsymbol{p}\cdot\boldsymbol{r}-E_+(\boldsymbol{p})t}{\hbar}}+b(\boldsymbol{p})e^{i\frac{\boldsymbol{p}\cdot\boldsymbol{r}-E_-(\boldsymbol{p})t}{\hbar}}\right] \tag{5.1.20}$$

特别是 Klein-Gordon 方程有平面波解$\frac{1}{(2\pi\hbar)^{3/2}}e^{i\frac{\boldsymbol{p}\cdot\boldsymbol{r}-E_\pm(\boldsymbol{p})t}{\hbar}}$,它们分别对应于能量 $E_\pm(\boldsymbol{p})$。注意到负能量没有下界。这样,在扰动下粒子将会不断地向更低能级跃迁,使物理体系很难处在稳定状态。而这与物理现实是不符合的。

§5.2 自由电子的 Dirac 方程

Ⅰ. Dirac 方程的建立

与非相对论自由质点的能量、动量关系(5.1.1)$H = E = \frac{\boldsymbol{p}^2}{2m}$相应,Schrödinger 方程包含波函数对 t 的一阶导数和对 x, y, z 的二阶导数。这注定了 Schrödinger 方程不可能是相对论不变的。利用相对论自由质点能量、动量关系(5.1.6)$E^2 = c^2\boldsymbol{p}^2 + m^2c^4$,Klein 和 Gordon 把波动方程中波函数对 t 的一阶导数提高到二阶,得到了协变的 Klein-Gordon 方程。但象我们在上一节指出的那样,它也带来了困难。这些困难是否与方程包含波函数对时、空坐标的二阶导数有关?我们可否通过把 Schrödinger 方程中的波函数对空间坐标的二阶导数降为一阶来建立相对论协变的波动方程?但是由上面相对论自由质点的能量、动量关系,我们得到 $E = \pm\sqrt{c^2\boldsymbol{p}^2 + m^2c^4}$,等式右边并不是 $\boldsymbol{p}$ 的线性表达式。1928 年 Dirac 大胆地写下了一个自由粒子能量、动量间的非齐次线性关系式

$$H = E = c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2 \tag{5.2.1}$$

由时间、空间均匀性,$\boldsymbol{\alpha}, \beta$ 与时空坐标无关。Dirac 希望(5.2.1)的平方能给出(5.1.6)。当 $\alpha_x, \alpha_y, \alpha_z$ 和 β 是满足乘法交换律的普通复数时,这显然是不可能的。但不妨先不承认乘法的交换律,看一下(5.2.1)的平方是什么。

$$E^2 = c^2\sum_{i,j=x,y,z}\alpha_i\alpha_j p_i p_j + \sum_{i=x,y,z}(\alpha_i\beta + \beta\alpha_i)mc^3 p_i + m^2c^4\beta \tag{5.2.2}$$

注意到 $p_ip_j = p_jp_i$,为使(5.2.2)与(5.1.6)一致,必须和只要有

$$\alpha_i\alpha_j + \alpha_j\alpha_i = 2\delta_{ij}, \alpha_i\beta + \beta\alpha_i = 0, \beta^2 = 1 \tag{5.2.3}$$

此即 $\alpha_x, \alpha_y, \alpha_z$ 和 β 各自平方为 1,两两反对易。Dirac 意识到这可以由矩阵来实现。当 $\alpha_x, \alpha_y, \alpha_z$ 和 β 是 N 阶方阵时,1 是 $N\times N$ 单位矩阵,相应的波函数 $\psi(\boldsymbol{r}, t)$ 应该是 $N\times 1$ 列矩阵。由(5.2.1)作代换(5.1.2),并作用于 $\psi(\boldsymbol{r}, t)$ 上,我们便得到了只包含波函数对时空坐标一阶微商的 Dirac 方程

$$i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = -i\hbar c\boldsymbol{\alpha}\cdot\nabla\psi(\boldsymbol{r},t) + \beta mc^2\psi(\boldsymbol{r},t) \tag{5.2.4}$$

其中哈密顿算符

$$H = -i\hbar c\boldsymbol{\alpha}\cdot\nabla + \beta mc^2 \tag{5.2.5}$$

应该是厄米的(作为一个力学量,它应该是厄米的,此外 $H = H^+$ 保证了总几率守恒)。从而应有

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}^+, \beta = \beta^+ \tag{5.2.6}$$

满足条件(5.2.3)和(5.2.6)的 $N\times N$ 矩阵 $\boldsymbol{\alpha},\beta$ 是否存在? 条件(5.2.3)和(5.2.6)把 $\boldsymbol{\alpha},\beta$ 确定到什么程度? 我们首先指出,当 $\boldsymbol{\alpha},\beta$ 满足(5.2.3)和(5.2.6)时,$U\boldsymbol{\alpha}U^+$,$U\beta U^+$ 也满足它们(U 是 $N\times N$ 酉矩阵)。也就是说,条件(5.2.3)和(5.2.6)在幺正相似变换下是不变的。

在下面的讨论中 $N\times N$ 矩阵的迹(trace)的概念扮演一个重要的角色,让我们在此简单地介绍一下它。设 A 是 $N\times N$ 矩阵,定义 A 的迹为 $trA = \sum_{\alpha=1}^{N} A_{\alpha\alpha}$。由这个定义我们立即可证 $tr(AB) = \sum_{\alpha=1}^{N}(AB)_{\alpha\alpha} = \sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} A_{\alpha\beta}B_{\beta\alpha} = tr(BA)$。特别是对非奇异矩阵 S,我们有 $tr(SAS^{-1}) = tr(A)$。这表明矩阵 A 的迹在相似变换下不变。由 §0.3 可知,N 维线性空间上的线性算符 $\hat{A}$,在基变换时,其矩阵要经受一个相似变换。故我们可以定义线性算符 $\hat{A}$ 的迹为它在任一基下的矩阵的迹,而不导致歧义。

(ⅰ)现在让我们来证明

$$tr\alpha_i = 0, tr\beta = 0, tr(\alpha_i\beta) = tr(\beta\alpha_i) = 0, tr(\alpha_i\alpha_j) = 0(i\neq j) \tag{5.2.7}$$

事实上 $tr\alpha_i = tr(\beta^2\alpha_i) = tr(\beta\alpha_i\beta) = tr(-\alpha_i\beta^2) = -tr\alpha_i$, $\therefore tr\alpha_i = 0$。

同样可得 $tr\beta = 0$。由反对易关系 $\alpha_i\beta = -\beta\alpha_i$ 得 $tr\alpha_i\beta = -tr\beta\alpha_i$, $2tr\beta\alpha_i = 0$,从而 $tr\beta\alpha_i = 0$。余类推。

(ⅱ)由这四个矩阵的对合性(平方为 1),可知它们的特征值的平方为 1,从而特征值不外乎 ± 1。

β 是厄米矩阵,其本征值不外乎 ± 1,变到自身表象 β 是对角矩阵,对角元不外乎 ± 1。$tr\beta = 0$ 表明本征值 $+1$ 出现的次数等于本征值 -1 出现的次数,故矩阵阶数 N 必须是一个偶数。

(ⅲ)最小的正偶数是 2,但 $\beta,\alpha_j(j=x,y,z)$不可能是 2×2 矩阵。让我们来证明这个断言。

以 $\mathbf{M}_2(\mathbf{C})$记全体 2×2 复矩阵所成的四维线性空间。则 2×2 单位矩阵和三个 *Pauli* 矩阵

$$I=\begin{bmatrix}1&0\\0&1\end{bmatrix},\sigma_x=\begin{bmatrix}0&1\\1&0\end{bmatrix},\sigma_y=\begin{bmatrix}0&-i\\i&0\end{bmatrix},\sigma_z=\begin{bmatrix}1&0\\0&-1\end{bmatrix} \tag{5.2.8}$$

构成 $\mathbf{M}_2(\mathbf{C})$的一个基。这四个 2×2 矩阵都是平方为 I 的厄米矩阵。泡利矩阵满足下面的关系

$$\sigma_x\sigma_y=\sigma_z=-\sigma_y\sigma_x,\sigma_y\sigma_z=\sigma_x=-\sigma_z\sigma_y,\sigma_z\sigma_x=\sigma_y=-\sigma_x\sigma_z \tag{5.2.9}$$

可见泡利矩阵两两反对易。泡利矩阵两两反对易,各自平方为 I 的性质可概括为

$$\sigma_j\sigma_k+\sigma_k\sigma_j=2\delta_{jk}I \tag{5.2.10}$$

若 $\beta,\alpha_j(j=x,y,z)$是 2×2 矩阵,则有

$$\beta=\beta_0I+\sum_{k=x,y,z}\beta^k\sigma_k,\alpha_j=\alpha_j^0I+\sum_{k=x,y,z}\alpha_j^k\sigma_k \tag{5.2.11}$$

厄米性要求展开系数是实数。由对合性(平方为 I)

$$\begin{aligned}\beta^2&=\left[(\beta_0)^2+\sum_{k=x,y,z}(\beta^k)^2\right]I+\sum_{k=x,y,z}2\beta^0\beta^k\sigma_k=I\\ \alpha_j^2&=\left[(\alpha_j^0)+\sum_{k=x,y,z}(\alpha_j^k)^2\right]I+2\sum_{k=x,y,z}\alpha_j^0\alpha_j^k\sigma_k=I\end{aligned} \tag{5.2.12}$$

可知,若 $\beta_0\neq0(\alpha_j\neq0)$,则 $\beta^k=0(\alpha_j^k=0)$,从而 $\beta=\pm I(\alpha_j=\pm1)$。但 $\pm I$ 不可能与任一非零的 2×2 矩阵反对易。于是只可能是 $\beta^0=0(\alpha_j^0=0)$:

$$\beta=\sum_{k=x,y,z}\beta^k\sigma_k\left(\alpha_j=\sum_{k=x,y,z}\alpha_j^k\sigma_k\right) \tag{5.2.13}$$

再次利用反对易性

$$\begin{aligned}&\beta\alpha_j+\alpha_j\beta=\sum_k\beta^k\sigma_k\sum_l\alpha_j^l\sigma_l+\sum_l\alpha_j^l\sigma_l\sum_k\beta^k\sigma_k=0,\\&\sum_k\sum_l\beta^k\alpha_j^l(\sigma_k\sigma_l+\sigma_l\sigma_k)=0,\\&\sum_k\sum_l2\delta_{kl}\beta^k\alpha_j^l=2\sum_{k=x,y,z}\beta^k\alpha_j^k=0\end{aligned} \tag{5.2.14}$$

可见$\begin{bmatrix}\beta^x\\\beta^y\\\beta^z\end{bmatrix}$与$\begin{bmatrix}\alpha_j^x\\\alpha_j^y\\\alpha_j^z\end{bmatrix}$必须正交。同理可得$\begin{bmatrix}\alpha_x^x\\\alpha_x^y\\\alpha_x^z\end{bmatrix},\begin{bmatrix}\alpha_y^x\\\alpha_y^y\\\alpha_y^z\end{bmatrix},\begin{bmatrix}\alpha_z^x\\\alpha_z^y\\\alpha_z^z\end{bmatrix}$两两正交。这四个两两正交

的非零向量一定线性无关。但不可能存在四个线性无关的三维向量。于是 $\beta, \alpha_j (j = x, y, z)$不可能是 2×2 矩阵。

(ⅳ)第二个最小的正偶数是 4。我们指出,利用泡利矩阵的性质可以直接验证。(留作练习)

$$\beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \boldsymbol{\alpha} = \begin{bmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{bmatrix} \tag{5.2.15}$$

这四个 4×4 矩阵满足条件(6.2.3)和(6.2.6)。于是我们用构造法证明了存在四个 4×4 厄米矩阵,两两反对易,各自平方为 1。$\beta, \boldsymbol{\alpha}$ 的矩阵表象(5.2.15)称作 *Pauli-Dirac* 表象,是一种 β - 表象。我们前面指出过,矩阵的厄米性,对合性和反对易关系在幺正相似变换下不变,故 $\beta, \boldsymbol{\alpha}$ 至少还有其它无穷多种 4×4 矩阵表示。

(ⅴ)在稍后我们将证明,满足条件(5.2.3)和(5.2.6)的 $N\times N$ 矩阵 β 和 $\boldsymbol{\alpha}$,只能是 4×4 矩阵(5.2.15)和它们的酉相似变换的直和(见本节附录)。

Ⅱ. 几率守恒,电子自旋

从以上的讨论可以看出,单由条件(5.2.3)和(5.2.6)只能推出 $\beta, \boldsymbol{\alpha}$ 的阶数 N 是 4 的整数倍。究竟 N 是 4 的几倍,还要看相应的 *Dirac* 方程能否很好地描述电子的行为。现在被普遍接受的看法是 $N=4$。在 $N=4$ 的限定下,$\beta, \boldsymbol{\alpha}$ 仍可以相差一个四维幺正相似变换,但这只是采用的表象不同,并不影响物理。在下面的讨论中我们将采用 *Pauli-Dirac* 表象(5.2.15)。

以 1×4 矩阵 $\psi^+(\boldsymbol{r}, t)$左乘(5.2.4),以 4×1 矩阵右乘(5.2.4)的厄米共轭矩阵,二者相减得

$$\frac{\partial}{\partial t}(\psi^+(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t)) + \nabla \cdot (\psi^+(\boldsymbol{r}, t) c\boldsymbol{\alpha}\psi(\boldsymbol{r}, t)) = 0 \tag{5.2.16}$$

这正是几率守恒定律,其中几率密度和几率流密度向量分别是

$$\rho(\boldsymbol{r}, t) = \psi^+(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t) \tag{5.2.17}$$

$$\boldsymbol{j}(\boldsymbol{r}, t) = \psi^+(\boldsymbol{r}, t) c\boldsymbol{\alpha}\psi(\boldsymbol{r}, t) \tag{5.2.18}$$

可以看到 *Klein-Gordon* 方程导致的负几率困难不复存在了。

不难验证动量算符 $\boldsymbol{P}$ 与哈密顿算符(5.2.5)对易,这是体系(自由电子)具有空间平移不变性的结果。自由电子体系自然应该有空间转动不变性。但是我们发现电

子的轨道角动量算符 $\boldsymbol{l}=\boldsymbol{r}\times\boldsymbol{p}$ 与哈密顿算符(5.2.5)并不对易:

$$[\boldsymbol{l},H=c\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2]=ci\hbar\boldsymbol{\alpha}\times\boldsymbol{p} \tag{5.2.19}$$

这启示我们,$\boldsymbol{l}$ 不是电子的总角动量,也就是说电子应该有自旋角动量。令

$$\boldsymbol{\Sigma}=\begin{bmatrix}\boldsymbol{\sigma} & 0\\ 0 & \boldsymbol{\sigma}\end{bmatrix} \tag{5.2.20}$$

不难验证

$$[\boldsymbol{\Sigma},H]=-2ci\boldsymbol{\alpha}\times\boldsymbol{p} \tag{5.2.21}$$

于是

$$\left[\boldsymbol{l}+\frac{\hbar}{2}\boldsymbol{\Sigma},H\right]=0 \tag{5.2.22}$$

可见应该把 $\boldsymbol{l}+\frac{\hbar}{2}\boldsymbol{\Sigma}$ 认同为电子的总角动量,而

$$\boldsymbol{s}=\frac{\hbar}{2}\boldsymbol{\Sigma} \tag{5.2.23}$$

是电子的自旋算符。在非相对论量子力学中,电子自旋是唯象的引入的。而现在我们由 Dirac 方程自然地得到了电子的自旋。

Ⅲ. Dirac 方程的平面波解系

让我们来研究自由电子 Dirac 方程的解。由于哈密顿算符不显含时间,问题可以归为求解定态 Schrödinger 方程。采用 Pauli-Dirac 表象,自由电子的哈密顿算符可以写作分块矩阵的形式

$$H=\begin{bmatrix}mc^2 & c\boldsymbol{\sigma}\cdot\boldsymbol{p}\\ c\boldsymbol{\sigma}\cdot\boldsymbol{p} & -mc^2\end{bmatrix} \tag{5.2.24}$$

由于动量 $\boldsymbol{p}$ 是运动积分,让我们来求 $H,\boldsymbol{p}$ 的共同的特征向量

$$u(\boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{x}}=\begin{bmatrix}\varphi(\boldsymbol{k})\\ \chi(\boldsymbol{k})\end{bmatrix}e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{5.2.25}$$

上式中 $u(\boldsymbol{k})$ 是 4×1 矩阵,$\varphi(\boldsymbol{k})$,$\chi(\boldsymbol{k})$ 是 2×1 矩阵。把(5.2.25)带入定态 Schrödinger 方程(H 的本征方程)得

$$\begin{bmatrix}mc^2-E(\boldsymbol{k}) & c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}\\ c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k} & -mc^2-E(\boldsymbol{k})\end{bmatrix}\begin{bmatrix}\varphi(\boldsymbol{k})\\ \chi(\boldsymbol{k})\end{bmatrix}=0 \tag{5.2.26}$$

这个齐次线性方程组有非零解的充要条件是系数行列式为零

$$[E(\boldsymbol{k})^2 - (m^2c^4 + c^2\hbar^2\boldsymbol{k}^2)]^2 = 0 \tag{5.2.27}$$

于是我们得到两个二重的特征根

$$E_{\pm}(\boldsymbol{k}) = \pm\sqrt{m^2c^4 + c^2\hbar^2\boldsymbol{k}^2} \tag{5.2.28}$$

我们看到动量取确定值 $\hbar\boldsymbol{k}$ 的自由电子,能量可取正负两个值。负能量随 $|\boldsymbol{p}|$ 增大是没有下限的。因此 Dirac 方程并没有解决负能量的困难。稍后我们将回到这个问题。把能量本征值(5.2.28)代回本征方程(6.2.26)

$$(mc^2 - E_{\pm}(\boldsymbol{k}))\varphi_{\pm}(\boldsymbol{k}) + c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}\chi_{\pm}(\boldsymbol{k}) = 0 \tag{5.2.29}$$

$$c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}\varphi_{\pm}(\boldsymbol{k}) - (mc^2 + E_{\pm}(\boldsymbol{k}))\chi_{\pm}(\boldsymbol{k}) = 0 \tag{5.2.30}$$

注意到$(\boldsymbol{\sigma}\cdot\boldsymbol{k})^2 = \boldsymbol{k}^2$,可知方程(5.2.29)与方程(5.2.30)是等价的。因而它们不能把 $u_{\pm}(\boldsymbol{k})$确定到只差一个常数倍。方程(5.2.26)对应于特征值 $E_{\pm}(\boldsymbol{k})$的解为

$$u_+(\boldsymbol{k}) = \begin{bmatrix} \varphi_+(\boldsymbol{k}) \\ \dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k}) + mc^2}\varphi_+(\boldsymbol{k}) \end{bmatrix}, u_-(\boldsymbol{k}) = \begin{bmatrix} -\dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k}) + mc^2}\chi_-(\boldsymbol{k}) \\ \chi_-(\boldsymbol{k}) \end{bmatrix} \tag{5.2.31}$$

其中 $\varphi_+(\boldsymbol{k})$和 $\chi_-(\boldsymbol{k})$可以是任一非零的 2×1 矩阵。这反映了自由电子的哈密顿量 H 和动量 $\boldsymbol{p}$ 并不构成力学量的完全集。自由电子的总角动量 $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s}$ 是运动积分。但$[j_\alpha, p_\beta] = [l_\alpha, p_\beta] = i\hbar\varepsilon_{\alpha\beta\gamma}p_\gamma$,当 $\alpha \neq \beta$ 时不为零。考虑运动积分 $\boldsymbol{j}\cdot\boldsymbol{p} = \boldsymbol{p}\cdot\boldsymbol{j}$。由算符恒等式 $\boldsymbol{l}\cdot\boldsymbol{p} = \boldsymbol{p}\cdot\boldsymbol{l} = 0$ 可知

$$\boldsymbol{j}\cdot\boldsymbol{p} = \boldsymbol{p}\cdot\boldsymbol{j} = \frac{\hbar}{2}\boldsymbol{\Sigma}\cdot\boldsymbol{p} = \boldsymbol{p}\cdot\frac{\hbar}{2}\boldsymbol{\Sigma} = \frac{\hbar}{2}\begin{bmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \end{bmatrix} \tag{5.2.32}$$

可见 $\boldsymbol{\Sigma}\cdot\boldsymbol{p}$ 与$H,\boldsymbol{p}$ 构成好力学量的完全集。要使 $H,\boldsymbol{p}$ 的共同的特征向量

$$\begin{bmatrix} \varphi_+(\boldsymbol{k}) \\ \dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k}) + mc^2}\varphi_+(\boldsymbol{k}) \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \begin{bmatrix} -\dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k}) + mc^2}\chi_-(\boldsymbol{k}) \\ \chi_-(\boldsymbol{k}) \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{5.2.33}$$

满足 $\boldsymbol{\Sigma}\cdot\boldsymbol{p}$ 的特征方程(注意,$\boldsymbol{p}$ 是算符,$\boldsymbol{k}\hbar$ 是特征值)

$$\begin{bmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \end{bmatrix} \begin{bmatrix} \varphi_+(\boldsymbol{k}) \\ \dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k}) + mc^2}\varphi_+(\boldsymbol{k}) \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}$$

$$= \lambda_+(\boldsymbol{k})\begin{bmatrix} \varphi_+(\boldsymbol{k}) \\ \dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k})+mc^2}\varphi_+(\boldsymbol{k}) \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}$$

$$\begin{bmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \end{bmatrix}\begin{bmatrix} -\dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k})+mc^2}\chi_-(\boldsymbol{k}) \\ \chi_-(\boldsymbol{k}) \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}$$

$$= \lambda_-(\boldsymbol{k})\begin{bmatrix} -\dfrac{c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}}{E_+(\boldsymbol{k})+mc^2}\chi_-(\boldsymbol{k}) \\ \chi_-(\boldsymbol{k}) \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{5.2.34}$$

则 $\varphi_+(\boldsymbol{k}),\chi_-(\boldsymbol{k})$必须满足方程

$$(\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}-\lambda_+(\boldsymbol{k}))\varphi_+(\boldsymbol{k})=0,(\hbar\boldsymbol{\sigma}\cdot\boldsymbol{k}-\lambda_-(\boldsymbol{k}))\chi_-(\boldsymbol{k})=0 \tag{5.2.35}$$

由$(\boldsymbol{\sigma}\cdot\boldsymbol{k})^2=\boldsymbol{k}^2$ 可知

$$\lambda_+(\boldsymbol{k})=\pm\hbar(\boldsymbol{k}),\lambda_-(\boldsymbol{k})=\pm\hbar(\boldsymbol{k}) \tag{5.2.36}$$

相应的特征向量为

$$\varphi_+(\boldsymbol{k})=\begin{pmatrix} |\boldsymbol{k}|+k_z \\ k_x+ik_y \end{pmatrix},\begin{pmatrix} -(k_x-ik_y) \\ |\boldsymbol{k}|+k_z \end{pmatrix},\chi_-(\boldsymbol{k})=\begin{pmatrix} |\boldsymbol{k}|+k_z \\ k_x+ik_y \end{pmatrix},\begin{pmatrix} -(k_x-ik_y) \\ |\boldsymbol{k}|+k_z \end{pmatrix}, \tag{5.2.37}$$

于是方程(5.2.4)的平面波解(动量取确定值的定态解)分成四类：

(ⅰ)$\boldsymbol{p}=\hbar\boldsymbol{k},E=\sqrt{m^2c^4+c^2\hbar^2\boldsymbol{k}^2},\boldsymbol{\Sigma}\cdot\boldsymbol{p}=\hbar|\boldsymbol{k}|$(自旋与运动方向一致)

$$\begin{bmatrix} |\boldsymbol{k}|+k_z \\ k_x+ik_y \\ \dfrac{(|\boldsymbol{k}|+k_z)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k})+mc^2} \\ \dfrac{(k_x+ik_y)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k})+mc^2} \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x}-\frac{i}{\hbar}E_+(\boldsymbol{k})t} \tag{5.2.38}$$

(ⅱ)$\boldsymbol{p}=\hbar\boldsymbol{k},E=\sqrt{m^2c^4+c^2\hbar^2\boldsymbol{k}^2},\boldsymbol{\Sigma}\cdot\boldsymbol{p}=-\hbar|\boldsymbol{k}|$(自旋与运动方向相反)

$$\begin{bmatrix} -(k_x - ik_y) \\ |\boldsymbol{k}| + k_z \\ \dfrac{(k_x - ik_y)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k}) + mc^2} \\ \dfrac{-(|\boldsymbol{k}| + k_z)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k}) + mc^2} \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x} - \frac{i}{\hbar}E_+(\boldsymbol{k})t} \tag{5.2.39}$$

(ⅲ) $\boldsymbol{p} = \hbar\boldsymbol{k}$, $E = -\sqrt{m^2c^4 + c^2\hbar^2\boldsymbol{k}^2}$, $\boldsymbol{\Sigma}\cdot\boldsymbol{p} = \hbar|\boldsymbol{k}|$(自旋与运动方向一致)

$$\begin{bmatrix} \dfrac{-(|\boldsymbol{k}| + k_z)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k}) + mc^2} \\ \dfrac{-(k_x + ik_y)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k}) + mc^2} \\ |\boldsymbol{k}| + k_z \\ k_x + ik_y \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x} + \frac{i}{\hbar}E_+(\boldsymbol{k})t} \tag{5.2.40}$$

(ⅳ) $\boldsymbol{p} = \hbar\boldsymbol{k}$, $E = -\sqrt{m^2c^4 + c^2\hbar^2\boldsymbol{k}^2}$, $\boldsymbol{\Sigma}\cdot\boldsymbol{p} = -\hbar|\boldsymbol{k}|$(自旋与运动方向相反)

$$\begin{bmatrix} \dfrac{-(k_x - ik_y)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k}) + mc^2} \\ \dfrac{(|\boldsymbol{k}| + k_z)c\hbar|\boldsymbol{k}|}{E_+(\boldsymbol{k}) + mc^2} \\ -(k_x + ik_y) \\ |\boldsymbol{k}| + k_z \end{bmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{x} + \frac{i}{\hbar}E_+(\boldsymbol{k})t} \tag{5.2.41}$$

Dirac 方程(5.2.4)的任一解都可以表作这四类解的线性叠加。(ⅰ),(ⅱ)类解是正能解,头两个分量比后两个分量大。(ⅲ),(ⅳ)类解是负能量解,后两个分量要比前两个分量大。

为了克服负能量困难,考虑到电子是费米子,服从泡利原理,Dirac 大胆的提出真空是全部负能态被充满,而全部正能态没有电子的状态。这样一个对应真空态的"负能海"是电荷密度的零点。当一个处在负能态的电子被激发到某个正能态,我们会测出这个处于正能态的电子和一个电荷与电子电荷相反,自旋质量与电子相同的"正电子"。这就是正负电子对的产生。当一个电子从正能态跃迁到一个本来未被占据的负能态(空穴)时,我们将有正负电子对的湮灭。几年之后,1932 年 Anderson 在宇宙

射线中找到了正电子。为此 Dirac 得到了 Nobel 物理奖。但是玻色子也有反粒子，负能海理论不能解释。1934 年，Pauli 和 Weisskonf 指出，不应该把 Klein-Gordon 和 Dirac 方程看作单粒子的相对论性的 Schrödinger 方程，而应该把它们和经典电磁场的 Maxwell 方程同样看待，它们是经典标量场(赝标量场)和旋量场的运动方程。把经典场(无限大自由度的体系)量子化，就可以得到描述自旋为 0，$\frac{1}{2}\hbar$，$\hbar$ 的粒子的相对论性量子理论(量子场论)。在量子场论中负几率、负能量等困难不复存在。它能很好地描述各种粒子产生、湮灭过程，构成了基本粒子物理的理论基础。

Ⅳ. Dirac 方程的 Lorentz 协变性

为了检验 Dirac 方程(5.2.4)的相对论协变性，以 $\frac{1}{c}\beta$ 左乘之，使时空坐标处在相同的地位

$$i\hbar\beta\frac{\partial\psi(\boldsymbol{r},t)}{c\partial t}+i\hbar\beta\boldsymbol{\alpha}\cdot\nabla\psi(\boldsymbol{r},t)-mc\psi(\boldsymbol{r},t)=0 \tag{5.2.42}$$

令 $\psi(x)=\psi(\boldsymbol{r},t)$，以及

$$\beta=\gamma^0=\gamma_0\quad\beta\alpha_x=\gamma^1=-\gamma_1\quad\beta\alpha_y=\gamma^2=-\gamma_2\quad\beta\alpha_z=\gamma^3=-\gamma_3 \tag{5.2.43}$$

Dirac 方程可以写作更对称的形式

$$i\hbar\gamma^\mu\frac{\partial}{\partial x^\mu}\psi(x)-mc\psi(x)=0 \tag{5.2.44}$$

由(5.2.3)和(5.2.6)可知

$$\gamma^\mu\gamma^\nu+\gamma^\mu\gamma^\nu=2g^{\mu\nu} \tag{5.2.45}$$

和

$$(\gamma^\mu)^+=\gamma_\mu \tag{5.2.46}$$

在 Pauli – Dirac 表象下

$$\gamma^0=\begin{pmatrix}I&0\\0&-I\end{pmatrix},\gamma^1=\begin{pmatrix}0&-\sigma_x\\-\sigma_x&0\end{pmatrix}$$
$$\gamma^2=\begin{pmatrix}0&-\sigma_y\\-\sigma_y&0\end{pmatrix},\gamma^3=\begin{pmatrix}0&\sigma_z\\-\sigma_z&0\end{pmatrix} \tag{5.2.47}$$

在四维时空平移下

$$x'^{\mu} = x^{\mu} + \rho^{\mu}(\mu = 0,1,2,3) \tag{5.2.48}$$

$$\psi'(x') = \psi(x) \tag{5.2.49}$$

容易由(5.2.44)得到

$$i\hbar\gamma^{\mu}\frac{\partial}{\partial x'^{\mu}}\psi'(x') - mc\psi'(x') = 0 \tag{5.2.50}$$

可见 Dirac 方程在四维时空平移下是不变的。下面只须讨论齐次 Lorentx 变换(四维时空转动)

$$x'^{\rho} = L^{\rho}_{\mu}x^{\mu} \text{ 或 } x' = Lx \tag{5.2.51}$$

其中

$$L^{T}GL = G \tag{5.2.52}$$

当时空坐标经受 Lorentz 变换(5.2.51)时,电子波函数 ψ 应按 $O(3,1)$群的某个表示 S 变换

$$S(L)\psi(x') = \psi'(x') = \Lambda(L)\psi(x) \text{ 或 } S(L)\psi(x) = \Lambda(L)\psi(L^{-1}x) \tag{5.2.53}$$

其中 Λ 是 $O(3,1)$群的四维表示,而 S 是 $O(3,1)$群的标量场表示与表示 Λ 的直积。

在惯性参考坐标系 $\sum'$ 中 Dirac 方程应为

$$i\hbar\gamma^{\rho}\frac{\partial}{\partial x'^{\rho}}\psi'(x') - mc\psi'(x') = 0 \tag{5.2.54}$$

以 $\Lambda(L)^{-1}$左乘上式得

$$i\hbar\Lambda(L)^{-1}\gamma^{\rho}\Lambda(L)(L^{-1})^{\mu}_{\rho}\frac{\partial}{\partial x^{\mu}}\psi(x) - mc\psi(x) = 0 \tag{5.2.55}$$

这应该与(5.2.44)一致,于是应有

$$\Lambda(L)^{-1}\gamma^{\rho}\Lambda(L)(L^{-1})^{\mu}_{\rho} = \gamma^{\mu}$$

此即

$$\Lambda(L)^{-1}\gamma^{\rho}\Lambda(L) = (L)^{\rho}_{\mu}\gamma^{\mu} \tag{5.2.56}$$

上式对任一 Lorentz 变换 $L \in O(3,1)$都成立。特别是对无穷小 Lorentz 变换(见附录§6)

$$L^{\rho}_{\mu} = \delta^{\rho}_{\mu} + \omega^{\rho}_{\mu} = \delta^{\rho}_{\mu} + \omega^{\rho\nu}g_{\mu\nu} \tag{5.2.57}$$

其中

$$\omega^{\rho v} + \omega^{v\rho} = 0, \mid \omega^{\rho v} \mid \ll 1 \tag{5.2.58}$$

$\Lambda(L)$可以用无穷小算符 $\lambda_{\rho v}$表作

$$\Lambda(L) = E + \sum_{\sigma < v} \omega^{\sigma v} \lambda_{\sigma v} = E + \frac{1}{2} \omega^{\sigma v} \lambda_{\sigma v} \tag{5.2.59}$$

其中 $\lambda_{\sigma v} = -\lambda_{v\sigma}$代入(5.2.56)得

$$-\frac{1}{2} \omega^{\sigma v} (\lambda_{\sigma v} \gamma^{\rho} - \gamma^{\rho} \lambda_{\sigma v}) = \omega_{\mu}^{\rho} \gamma^{\mu} = \frac{\omega^{\sigma v}}{2} (\delta_{\sigma}^{\rho} \gamma_{v} - \delta_{v}^{\rho} \gamma_{\sigma})$$

于是

$$\lambda_{\sigma v} \gamma^{\rho} - \gamma^{\rho} \lambda_{\sigma v} = \delta_{v}^{\rho} \gamma_{\sigma} - \gamma_{v} \delta_{\sigma}^{\rho}, \forall \rho = 0,1,2,3 \tag{5.2.60}$$

这是无穷小算符 $\lambda_{\sigma v}$与 $\gamma^0, \gamma^1, \gamma^2, \gamma^3$ 的对易关系。注意到由 4×4 矩阵 $\gamma^0, \gamma^1, \gamma^2, \gamma^3$ 的乘积的线性组合可以得到任一 4×4 矩阵,故它们是四维空间上的不可约线性算符组。由 Schur 引理可知,(5.2.60)把 $\lambda_{\sigma v}$确定到只差一个单位矩阵的纯量倍 $C_{\sigma v} E$。事实上,若 $\lambda'_{\sigma v}$也满足(5.2.60),则

$$(\lambda'_{\sigma v} - \lambda_{\sigma v}) \gamma^{\rho} = \gamma^{\rho} (\lambda'_{\sigma v} - \lambda_{\sigma v}), \forall \rho = 0,1,2,3 \tag{5.2.61}$$

由 Schur 引理得

$$\lambda'_{\sigma v} - \lambda_{\sigma v} = C_{\sigma v} E \tag{5.2.62}$$

直接验证可知

$$\lambda_{\sigma v} = \frac{1}{4} (\gamma_{\sigma} \gamma_{v} - \gamma_{v} \gamma_{\sigma}) \tag{5.2.63}$$

满足(5.2.60)。故不确定性只是 $C_{\sigma v} E$。注意到 $\lambda_{\sigma v}$之间的对易关系应与 $I_{\sigma v}$之间的对易关系相同(见附录李群表示定理),这样可以确定 $C_{\sigma v} = 0$。于是对应任一 Lorentz 变换

$$L = \exp\left(\frac{1}{2} \omega^{\sigma v} I_{\sigma v} \right) \in O(3,1) \tag{5.2.64}$$

$O(3,1)$群 4 维表示 Λ 的算符为

$$\Lambda(L) = \exp\left(\frac{1}{2} \omega^{\sigma v} \lambda_{\sigma v} \right) = \exp \frac{\omega^{\sigma v}}{8} (\lambda_{\sigma} \lambda_{v} - \lambda_{v} \lambda_{\sigma}) \tag{5.2.65}$$

对应于三维空间绕 z 轴转动的无穷小算符的 $i\hbar$ 倍正是

$$\frac{i\hbar}{4}(\gamma_1\gamma_2 - \gamma_2\gamma_1) = \frac{\hbar}{2}\begin{bmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{bmatrix} = \frac{\hbar}{2}\sum_3 \tag{5.2.66}$$

这正是电子自旋的 z 分量。空间绕坐标原点的转动群 $SO(3)$是齐次变换群 $O(3,1)$的子群。$SO(3)$群的四维表示 $\Lambda|_{SO(3)}$的无穷小算符的 $i\hbar$ 倍是相应自由度的角动量算符。上一小节得出的电子自旋算符与现在我们由对称性分析得到的结果完全一致。由式(5.2.65)和(5.2.66)可知,绕 z 轴转动 2π 角对应的 Λ 矩阵为 $-E$,故 Λ 是双值表示。按(5.2.53)$O(3,1)$的表示 S 变换的 $\psi(x)$称为旋量场。称 Λ 是 $O(3,1)$旋量表示。S 是 Λ 和 $O(3,1)$群标量场表示的直积。$O(3,1)$群的四维矩阵表示 Λ,并不是酉矩阵表示。事实上由(5.2.65)可得

$$\Lambda(L^+) = \exp\frac{\omega^{\sigma v}}{8}(\lambda_v^+\lambda_\sigma^+ - \lambda_\sigma^+\lambda_v^+) \tag{5.2.67}$$

并不恒等于

$$\Lambda(L^-) = \exp\frac{\omega^{\sigma v}}{8}(\lambda_v\lambda_\sigma - \lambda_\sigma\lambda_v) \tag{5.2.68}$$

这与 $O(3,1)$不是紧致李群有关(参考附录)。由于 $\Lambda(L)$不是酉矩阵,故 $\psi(x)^+\psi(x)\neq\psi'(x')^+\psi'(x')$不是 Lorentz 不变量。由

$$\gamma^0\gamma_v^+\gamma^0 = \gamma_v \tag{5.2.69}$$

可得

$$\gamma^0\Lambda(L)^+\gamma^0 = \Lambda(L)^{-1} \tag{5.2.70}$$

令

$$\overline{\psi}(x) \equiv \psi^+(x)\gamma^0 \tag{5.2.71}$$

称之为旋量波函数 $\psi(x)$的 Dirac 共轭。由(5.2.70)可得

$$\overline{\psi}'(x') \equiv \psi'^+(x')\gamma^0 = \psi(x)^+\Lambda(L)^+\gamma^0 = \overline{\psi}(x)\Lambda(L)^{-1} \tag{5.2.72}$$

于是

$$\overline{\psi}(x)\psi(x) = \overline{\psi}'(x')\psi'(x') \tag{5.2.73}$$

是 Lorentz 不变量。利用(5.2.53)、(5.2.56)和(5.2.72)可得

$$\overline{\psi}'(x')\gamma^\rho\psi'(x') = L_\mu^\rho\overline{\psi}(x)\gamma^\mu\psi(x) \tag{5.2.74}$$

这表明 $\overline{\psi}(x)\gamma^\mu\psi(x)$是四维矢量,而这正是四维几率流密度矢量应有的性质。作为练习,请读者证明

$$\bar{\psi}'(x')\frac{\gamma^\rho\gamma^\sigma-\gamma^\sigma\gamma^\rho}{2}\psi'(x')=L_\mu^\rho L_v^\sigma\bar{\psi}(x)\frac{\gamma^\mu\gamma^v-\gamma^v\gamma^\mu}{2}\psi(x) \tag{5.2.75}$$

$\bar{\psi}(x)\frac{\gamma^\mu\gamma^v-\gamma^v\gamma^\mu}{2}\psi(x)$在 Lorentz 变换下是一个二阶反对称逆变张量。

考虑空间反演这一特殊情况：

$$L=\mathrm{diag}(1,-1,-1,-1)=I \tag{5.2.76}$$

由(5.2.56)得

$$\Lambda(I)^{-1}\gamma^\rho\Lambda(I)=\gamma_\rho \text{ 或 } \gamma_\rho\Lambda(I)=\Lambda(I)\gamma_\rho \tag{5.2.77}$$

可见 $\Lambda(I)$与 γ^0 对易，与 γ^j 反对易。由于$\{E,I\}$是二阶群，我们可知 $\Lambda(I)$是酉矩阵，$\Lambda(I)^2=E$。于是 $\Lambda(I)$是厄米的。满足以上要求的 $\Lambda(I)$只能是(留作练习)

$$\Lambda(I)=\pm\gamma^0 \tag{5.2.78}$$

电子的内禀宇称是正的，故 $\Lambda(I)=\gamma^0$ 对于时间变号

$$L=\mathrm{diag}(-1,1,1,1)=-I \tag{5.2.79}$$

$$\Lambda(I)^{-1}\gamma^\rho\Lambda(-I)=-\gamma_\rho,\quad \gamma_\rho\Lambda(-I)=-\gamma_\rho\Lambda(-I) \tag{5.2.80}$$

可见 $\Lambda(-I)$与 γ^0 反对易，与 γ^j 对易。由于$\{E,I\}$是二阶群，$\Lambda(-I)^+=\Lambda(-I)^{-1}$，$\Lambda(-I)^2=E$。故 $\Lambda(-I)$是厄米阵。这些条件把 $\Lambda(-I)$确定到

$$\Lambda(-I)=\pm\gamma^1\gamma^2\gamma^3 \tag{5.2.81}$$

练习 5-2-1 试证：$\bar{\psi}(x)\frac{\gamma^\mu\gamma^v-\gamma^v\gamma^\mu}{2}\psi(x)$在 Lorentz 变换下是一个二阶反对称逆变张量。

附录 Dirac 矩阵 γ^μ

设 $\gamma^0,\gamma^1,\gamma^2,\gamma^3$ 是 $N\times N$ 矩阵，满足如下的反对易关系和厄米共轭关系

$$\gamma^\mu\gamma^v+\gamma^v\gamma^\mu=2g^{\mu v}\quad(\mu,v=0,1,2,3) \tag{1}$$

$$(\gamma^\mu)^+=\gamma_\mu=g_{\mu\nu}\gamma^\nu \tag{2}$$

称满足条件(1)和(2)的 $N\times N$ 矩阵为 Dirac 矩阵。

(ⅰ)由于这 4 个 Dirac 矩阵两两反对易，各自的平方是单位矩阵 E 或 $-E$，可以推知，有限个 Dirac 矩阵的乘积不外乎下面列举的 32 个 $N\times N$ 矩阵：

$$\begin{aligned}&\pm E,\ \pm\gamma^\mu(\mu=0,1,2,3),\ \pm\gamma^\mu\gamma^v(0\le\mu<v\le3)\\&\pm\gamma^\mu\gamma^v\gamma^\rho(0\le\mu<v<\rho\le3),\ \pm\gamma^0\gamma^1\gamma^2\gamma^3\end{aligned} \tag{3}$$

反过来，这 32 个 $N\times N$ 矩阵都等于有限个 Dirac 矩阵的乘积。事实上，$+E=\gamma^0\gamma^0$，其它带 + 号的 15 个 $N\times N$ 矩阵已经明显的写成了有限个 Dirac 矩阵的乘积。式(3)中带—号的 16 个 $N\times N$ 矩阵，每个都能写成相应的带 + 号的 $N\times N$ 矩阵左乘以 $\gamma^1\gamma^1$。这就证明了式(3)中的 32 个 $N\times N$ 矩阵是有限个 Dirac 矩阵乘积的全部。不难看出它们构成一个 32 阶群，称之为 Dirac 群。这是 $\gamma^0,\gamma^1,\gamma^2,\gamma^3$ 生成的群。

(ⅱ)在式(3)中的 32 个 $N\times N$ 矩阵中，除了 $\pm E$，$tr(\pm E)=\pm N\neq 0$ 外，其它 30 个矩阵的阵迹为零。事实上，利用 Dirac 矩阵的反对易性(1)和 $trAB=trBA$ 可得

$$tr\gamma^0=tr(-\gamma^1\gamma^1\gamma^0)=tr(\gamma^1\gamma^0\gamma^1)=tr(\gamma^1\gamma^1\gamma^0)=0$$

$$tr\gamma^j=tr(\gamma^0\gamma^0\gamma^j)=tr(-\gamma^0\gamma^j\gamma^0)=tr(-\gamma^0\gamma^0\gamma^j)=0\quad(j=1,2,3)$$

$$tr(\gamma^\mu\gamma^\nu)=tr(\gamma^\nu\gamma^\mu)=tr(-\gamma^\mu\gamma^\nu)=0\quad(\mu<\nu)$$

$$tr(\gamma^\mu\gamma^\nu\gamma^\rho)=tr(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_\sigma)=tr(-\gamma^\sigma\gamma^\mu\gamma^\nu\gamma^\rho\gamma_\sigma)$$

$$=tr(-\gamma^\mu\gamma^\nu\gamma^\rho)=0\quad(\sigma\text{ 不同于 }\mu<\nu<\rho)$$

$$tr(\gamma^0\gamma^1\gamma^2\gamma^3)=tr(-\gamma^1\gamma^2\gamma^3\gamma^0)=tr(-\gamma^0\gamma^1\gamma^2\gamma^3)=0 \tag{4}$$

(ⅲ)式(3)中带 + 号的 16 个 $N\times N$ 矩阵是线性独立的。以 $\Gamma_k(k=1,\cdots,16)$ 记这 16 个 $N\times N$ 矩阵。设

$$\sum_{k=1}^{16}C_k\Gamma_k=0 \tag{5}$$

取上式两边的阵迹得 $\sum_{k=1}^{16}C_k tr\Gamma_k=0$。因为只有 $trE=tr\Gamma_1=N\neq 0$，其它 15 个 Γ_k 的阵迹为零。故可得 $C_1=0$。从而(5)可以写作 $\sum_{k=2}^{16}C_k\Gamma_k=0$。以 Γ_2 右乘之得 $\sum_{k=2}^{16}C_k(\Gamma_k\Gamma_2)=0$，只有 $\Gamma_2\Gamma_2$ 是 $+E$ 或 $-E$，其它 $\Gamma_k\Gamma_2(k\geq 3)$ 的阵迹为零。故由 $\sum_{k=2}^{16}C_k tr(\Gamma_k\Gamma_2)=0$ 可得 $C_2=0$。以 Γ_3 右乘 $\sum_{k=3}^{16}C_3\Gamma_k=0$，再取阵迹可得 $C_3=0$。以此类推，可知式(5)中所有系数 $C_k=0$。可见式(3)中带 + 号的 16 个 $N\times N$ 矩阵是线性独立的。它们张成一个 16 维复线性空间。这个空间在矩阵乘法下是封闭的，因而构成一个 16 维的矩阵代数，称之为 Dirac 代数。这是 $\gamma^0,\gamma^1,\gamma^2,\gamma^3$ 生成的代数。

(ⅳ)Dirac 群的不可约表示

有限群的表示总可以写作不可约表示的直和。有限群的不等价的不可约表示的

个数等于有限群的共轭类的个数。我们看到 Dirac 群的元素在乘法下或对易或反对易：$\Gamma_j\Gamma_k = \pm \Gamma_k\Gamma_j$。只有 $+E$，$-E$ 与所有群元素对易。故 $+E$ 自成一共轭类，$-E$ 自成一共轭类。其它每一个元都与一些群元对易，与另一群元反对易。这些元与它们的变号构成一个共轭类。于是 Dirac 群有 $2+(32-2)/2=17$ 个共轭类，有 17 个不等价的不可约表示。所有不等价不可约表示矩阵元的个数等于群元素的个数。在我们的情况下，Dirac 群的元素的个数为 32，共轭类的个数为 17。以 d_j 记第 j 类不可约表示的维数，则应有

$$\sum_{j=1}^{17} d_j^2 = 32 \tag{6}$$

这个方程的使 $d_1 \leq d_2 \leq \cdots \leq d_{17}$ 的正整数解只有两个：

$$d_1 = d_2 = \cdots d_{12} = 1, d_{13} = d_{14} = \cdots = d_{17} = 2 \tag{7}$$

$$d_1 = d_2 = \cdots d_{16} = 1, d_{17} = 4 \tag{8}$$

考虑式(6.2.47)，直接验证可知，它们满足(6.2.45)或(1)，因而它们是 Dirac 群的一个四维矩阵表示。又由于它们的乘积的线性组合可得任一 4×4 矩阵。故不能有非平庸的不变子空间，一定构成 Dirac 群的不可约表示。从而 Dirac 群的不可约表示 15 个是一维的，1 个是四维的。

(Ⅴ)我们真正需要的是 Dirac 代数的表示。Dirac 群的元是 16 维 Dirac 代数的无穷多个元中的 32 个元，而且只考虑乘法，不考虑加法和数乘。满足群乘法同态要求，不一定能满足代数同态要求。群一维表示矩阵不可能反对易，不可能满足(1)，故只有 Dirac 群的四维不可约表示是 Dirac 代数的表示。

Dirac 代数的矩阵表示，当限制到其有限子集 Dirac 群，并只考虑乘法运算时，是 Dirac 群的矩阵表示。有限群 Dirac 群的矩阵表示是不可约表示的直和。这就是说，表示空间是不变子空间的直和。而作为 Dirac 群的表示的不变子空间，与作为 Dirac 代数的表示的不变子空间是相同的。故 Dirac 代数的矩阵表示也是不可约表示的直和；表示空间有相同的不变子空间分解。Dirac 群的一维不可约表示不能是 Dirac 代数的表示。故 Dirac 代数表示只能是若干个四维不可约表示的直和。

从上面的结果可知，Dirac 矩阵只能是 4×4 矩阵(6.2.47)和它们的酉相似变换的直和。

§5.3 电磁场中无旋粒子的 klein-Gordon 方程

Ⅰ.方程的建立

在电磁场中带电粒子的相对论性的经典运动方程为

$$\frac{d}{dt}\frac{m\dot{\boldsymbol{r}}}{\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}} = q\boldsymbol{E}(\boldsymbol{r},t)+\frac{q}{c}\dot{\boldsymbol{r}}\times\boldsymbol{B}$$

$$= q\left(-\nabla\varphi(\boldsymbol{r},t)-\frac{1}{c}\frac{\partial\boldsymbol{A}(\boldsymbol{r},t)}{\partial t}\right)+\frac{q}{c}\dot{\boldsymbol{r}}\times(\nabla\times\boldsymbol{A}(\boldsymbol{r},t))$$

$$= \nabla\left(-q\varphi(\boldsymbol{r},t)+\frac{q}{c}\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r},t)\right)-\frac{q}{c}\left(\frac{\partial\boldsymbol{A}(\boldsymbol{r},t)}{\partial t}+\dot{\boldsymbol{r}}\cdot\nabla\boldsymbol{A}(\boldsymbol{r},t)\right) \tag{5.3.1}$$

上式可以写作

$$\frac{d}{dt}\left(\frac{m\dot{\boldsymbol{r}}}{\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}}+\frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)=\nabla\left(-q\varphi(\boldsymbol{r},t)+\frac{q}{c}\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r},t)\right) \tag{5.3.2}$$

与 Euler-Lagrange 方程

$$\frac{d}{dt}\frac{\partial}{\partial\dot{\boldsymbol{r}}}L(\boldsymbol{r},\dot{\boldsymbol{r}},t)=\nabla L(\boldsymbol{r},\dot{\boldsymbol{r}},t) \tag{5.3.3}$$

对比,并注意到

$$\frac{\partial}{\partial\dot{\boldsymbol{r}}}\left(-mc^2\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}\right)=\frac{m\dot{\boldsymbol{r}}}{\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}} \tag{5.3.4}$$

式(5.3.2)可以写作

$$\frac{d}{dt}\frac{\partial}{\partial\dot{\boldsymbol{r}}}\left(-mc^2\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}+\frac{q}{c}\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r},t)-q\varphi(\boldsymbol{r},t)\right)$$

$$=\nabla\left(-mc^2\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}+\frac{q}{c}\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r},t)-q\varphi(\boldsymbol{r},t)\right) \tag{5.3.5}$$

可见电磁场中的荷电粒子的拉氏函数应为

$$L(\boldsymbol{r},\dot{\boldsymbol{r}},t) = -mc^2\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}} + \frac{q}{c}\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r},t) - q\varphi(\boldsymbol{r},t) \tag{5.3.6}$$

$\boldsymbol{r}$ 的正则共轭动量为

$$\boldsymbol{p} = \frac{\partial L(\boldsymbol{r},\dot{\boldsymbol{r}},t)}{\partial\dot{\boldsymbol{r}}} = \frac{m\dot{\boldsymbol{r}}}{\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}} + \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t) \tag{5.3.7}$$

$\boldsymbol{r}$ 的正则共轭动量不再等于荷电粒子的机械动量$\frac{m\dot{\boldsymbol{r}}}{\sqrt{1-\frac{\dot{\boldsymbol{r}}^2}{c^2}}}$。由上式可以把 $\dot{\boldsymbol{r}}$ 解出

$$\dot{\boldsymbol{r}} = C\left(\boldsymbol{p}-\frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)\left[\left(\boldsymbol{p}-\frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2 + m^2c^2\right]^{-1/2} \tag{5.3.8}$$

电磁场中荷电粒子的哈密顿量为

$$H(\boldsymbol{r},\boldsymbol{p},t) = \dot{\boldsymbol{r}}\cdot\boldsymbol{p} - L(\boldsymbol{r},\dot{\boldsymbol{r}},t) = c\sqrt{\left(\boldsymbol{p}-\frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2 + m^2c^2} + q\varphi(\boldsymbol{r},t) \tag{5.3.9}$$

这表明哈密顿量是机械动能与库仑能的和。由上式可得

$$(E - q\varphi(\boldsymbol{r},t))^2 = c^2\left(\boldsymbol{p}-\frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2 + m^2c^2 \tag{5.3.10}$$

把上式中的 E 和 $\boldsymbol{p}$ 作替换(5.1.2),并作用在波函数 $\phi(\boldsymbol{r},t)$ 上我们便得到电磁场中荷电的无旋粒子的 Klein-Gordon 方程

$$\left(i\hbar\frac{\partial}{\partial t} - q\varphi(\boldsymbol{r},t)\right)^2\phi(\boldsymbol{r},t) - c^2\left(-i\hbar\nabla - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2\phi(\boldsymbol{r},t) - m^2c^4\phi(\boldsymbol{r},t) = 0 \tag{5.3.11}$$

注意到

$$\varphi(\boldsymbol{r},t) = A^0(x) = A_0(x), A_x(\boldsymbol{r},t) = A^1(x) = -A_1(x),\cdots \tag{5.3.12}$$

式(5.3.11)可以写成明显协变的形式

$$\left(i\hbar\frac{\partial}{\partial x^\mu} - \frac{q}{c}A_\mu(x)\right)\left(i\hbar\frac{\partial}{\partial x_\mu} - \frac{q}{c}A^\mu(x)\right)\phi(x) = m^2c^2\phi(x) \tag{5.3.13}$$

Ⅱ. π 介子原子

π 介子自旋为零。π^- 介子与重核构成的束缚态体系叫 π 介子原子。此时的电磁

场为库仑场

$$\varphi(\boldsymbol{r},t)=\frac{Ze}{r},\boldsymbol{A}(\boldsymbol{r},t)=0 \tag{5.3.14}$$

Klein-Gordon 方程为

$$\left(i\hbar\frac{\partial}{\partial t}+\frac{Ze^2}{r}\right)^2\phi(\boldsymbol{r},t)=(c^2\boldsymbol{p}^2+m^2c)^4\phi(\boldsymbol{r},t) \tag{5.3.15}$$

尽管 Klein-Gordon 方程不能纳入 Schrödinger 方程 $i\hbar\frac{\partial}{\partial t}\psi=H\psi$ 的形式，但我们仍可以求方程(5.3.15)的“定态解”

$$\phi(\boldsymbol{r},t)=\Phi(\boldsymbol{r})e^{-iEt/\hbar} \tag{5.3.16}$$

代入(5.3.15)得

$$\left(E+\frac{Ze^2}{r}\right)^2\Phi(\boldsymbol{r})=(c^2\boldsymbol{p}^2+m^2c^4)\Phi(\boldsymbol{r}) \tag{5.3.17}$$

这并不是一个通常意义下的定态 Schrödinger 方程 $H\Phi=E\Phi$，我们不能说“$\boldsymbol{l}$ 与 H 对易，故有共同的完备的本征函数系”。但注意到

$$\boldsymbol{p}^2=-\hbar^2\nabla^2=-\frac{\hbar^2}{r}\frac{\partial^2}{\partial r^2}r+\frac{\boldsymbol{l}^2}{r^2} \tag{5.3.18}$$

我们仍可以求方程(5.3.17)的如下形式的解

$$\Phi(\boldsymbol{r})=\frac{1}{r}u_l(r)Y_{lm}(\theta,\varphi) \tag{5.3.19}$$

代入(5.3.17)得

$$\frac{d^2u_l(r)}{dr^2}+\left(\frac{E^2-m^2c^4}{\hbar^2c^2}+\frac{2EZe^2}{\hbar^2c^2r}+\frac{Z^2e^4}{\hbar^2c^2r^2}-\frac{l(l+1)}{r^2}\right)u_l(r)=0 \tag{5.3.20}$$

与非相对论性中心力场径向波函数的本征方程不同，方程(5.3.20)中不仅包含 E，而且还包含 E^2。我们要研究的是束缚态，在非相对论情况下对应 $E<0$，而在相对论情况下考虑到静能 mc^2，束缚态对应 $0<E<mc^2$。令

$$\beta^2=\frac{4(m^2c^4-E^2)}{\hbar^2c^2} \tag{5.3.21}$$

β 是一个量纲为 $[\beta]=L^{-1}$ 的参数。引入无量纲半径

$$\rho=\beta r \tag{5.3.22}$$

记

$$u_l(r) = v_l(\rho) \tag{5.3.23}$$

则方程(5.3.20)可以重新写作

$$\frac{d^2}{d\rho^2}v_l(\rho) + \left(\frac{\lambda}{\rho} - \frac{l(l+1) - Z^2\alpha^2}{\rho^2} - \frac{1}{4}\right)v_l(\rho) = 0 \tag{5.3.24}$$

其中 λ 是无量纲参数

$$\lambda = \frac{2Z\alpha E}{\hbar c\beta} = \frac{Z\alpha E}{\sqrt{m^2c^4 - E^2}} > 0 \tag{5.3.25}$$

α 是精细结构常数,也无量纲:

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137.0359895} \tag{5.3.26}$$

考察当 $\rho \to \infty$ 时方程(5.3.24)的渐进行为。当 $\rho \to \infty$ 时,方程可近似的写作

$$\frac{d^2v_l(\rho)}{d\rho^2} - \frac{1}{4}v_l(\rho) = 0 \tag{5.3.27}$$

可见 $\rho \to \infty$ 时,使 $v_1(\rho)$ 不趋于 ∞ 的解应为 $v_l(\rho) \sim e^{-\frac{1}{2}\rho}$。设当 $\rho \to 0$ 时,$v_l(\rho) \sim \rho^{s+1}$ 代入(5.3.24)的近似式得

$$s(s+1) - [l(l+1) - Z^2\alpha^2] = 0 \tag{5.3.28}$$

这是一个关于未知数 s 的一元二次方程。二根之和为 $\xi + \eta = -1$,二根之积为 $\xi\eta = Z^2\alpha^2 - l(l+1)$。由于 $Z\alpha \ll 1$,二根差的平方 $(\xi - \eta)^2 = (\xi + \eta)^2 - 4\xi\eta$ 不是整数;二根之差 $\xi - \eta$ 不是整数。故它们对应于微分方程(5.3.24)的线性独立的解。当 $l = 0$ 时,ξ, η 同号,它们对应的方程(5.3.24)的解在 $\rho = 0$ 的邻域模方可积。当 $l \geq 1$ 时,二根异号,只有正根 $\frac{-1 + \sqrt{1 + 4[l(l+1) - Z^2\alpha^2]}}{2}$ 对应的方程(5.3.24)的解在 $\rho = 0$ 的邻域模方可积。下面以 s 记代数方程(5.3.28)的根(当 $l = 0$)或正根(当 $l \geq 1$),令

$$v_l(\rho) = e^{-\frac{1}{2}\rho}\rho^{s+1}\sum_{\sigma \geq 0} a_\sigma \rho^\sigma \quad (a_\alpha \neq 0) \tag{5.3.29}$$

代入(5.3.24)得递推关系

$$(\sigma + 1)(2s + 2 + \sigma)a_{\sigma+1} + (\lambda - s - 1 - \sigma)a_\sigma = 0, \quad \forall \sigma \geq 0 \tag{5.3.30}$$

当 $\sigma \to \infty$ 时

$$\frac{a_{\sigma+1}}{a_\sigma} = \frac{-(\lambda - s - 1 - \sigma)}{(\sigma + 1)(2s + 2 + \sigma)} \to \frac{1}{\sigma} \tag{5.3.31}$$

可见当 $\rho\to\infty$，如果不中断，幂级数

$$\sum_{\sigma\geq 0} a_\sigma \rho^\sigma \sim e^\rho \tag{5.3.32}$$

为了得到模平方可积的波函数，我们必须让这个幂级数于某一项中断：

$$a_{n_r} \neq 0, a_{n_r+1} = 0 \quad (n_r = 0,1,\cdots)$$

于是

$$\lambda = s + 1 + n_r \quad (n_r = 0,1,\cdots) \tag{5.3.33}$$

由式(5.3.21)、(5.3.25)和(5.3.32)得

$$E = mc^2\left[1 + \frac{Z^2\alpha^2}{(n_r + s + 1)^2}\right]^{-1/2} \tag{5.3.34}$$

由代数方程(5.3.28)解出 S 的大根

$$s = -\frac{1}{2} + \sqrt{\left(l + \frac{1}{2}\right)^2 - Z^2\alpha^2} \tag{5.3.35}$$

代入式(5.3.34)得

$$E = mc^2\left[1 + \frac{Z^2\alpha^2}{\left(n_r + \frac{1}{2} + \sqrt{\left(l + \frac{1}{2}\right)^2 - Z^2\alpha^2}\right)^2}\right]^{-\frac{1}{2}} \tag{5.3.36}$$

展成 $Z^2\alpha^2 \ll 1$ 的幂级数：

$$E = mc^2\left[1 - \frac{Z^2\alpha^2}{2n^2} - \frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{l + \frac{1}{2}} - \frac{3}{4}\right) + \cdots\right] \tag{5.3.37}$$

其中 n 是主量子数

$$n \equiv n_r + l + 1 \tag{5.3.38}$$

当 $l=0$ 时，代数方程(5.3.28)的小根为

$$s = -\frac{1}{2} - \frac{1}{2}\sqrt{1 - Z^2\alpha^2} > -1 \tag{5.3.39}$$

代入式(5.3.34)得

$$E = mc^2\left[1 + \frac{Z^2\alpha^2}{\left(n_r + \frac{1}{2} - \sqrt{\frac{1}{4} - Z^2\alpha^2}\right)^2}\right]^{-\frac{1}{2}} \tag{5.3.40}$$

当 $n_r=0$ 时，上式右边方括号中第二项的分母 $\sim Z^4\alpha^4$，$E\to 0_+$。当 $n_r=1$ 时，上式右

边可展作 $Z^2\alpha^2$ 的幂级数

$$E = mc^2\left[1 - \frac{Z^2\alpha^2}{2n_r^2} - \frac{Z^4\alpha^4}{4n_r^4}\left(n_r + \frac{3}{2}\right) + \cdots\right] \tag{5.3.41}$$

由式(5.3.37)可看出,右边的第一项代表静质量对应的能量,第二项是非相对论量子力学给出的结合能。第三项和其后各项是相对论修正,使能量降低,而且降低量与 n, l 都有关。这说明相对论效应破坏了 $O(4)$动力学对称性。

练习 5-3-1 讨论 Klein-Gordon 方程(5.3.11)的 Lorentz 协变性。

§5.4 电磁场中电子的 Dirac 方程

Ⅰ.方程的建立

我们看到由自由粒子的 Klein-Gordon 方程出发,把 $i\hbar\frac{\partial}{\partial t}$ 换成 $i\hbar\frac{\partial}{\partial t} - q\phi(\boldsymbol{r},t)$,把 $-i\hbar\nabla$ 换成 $-i\hbar\nabla - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)$,就可以得到电磁场中荷电粒子的 Klein-Gordon 方程。现在从自由电子的 Dirac 方程出发,作同样的替换我们便得到电磁场中电子的 Dirac 方程

$$\left(i\hbar\frac{\partial}{\partial t} + e\phi(\boldsymbol{r},t)\right)^2\psi(\boldsymbol{r},t) = c\boldsymbol{\alpha}\cdot\left(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2\psi(\boldsymbol{r},t) - mc^2\beta\psi(\boldsymbol{r},t) \tag{5.4.1}$$

上式可以写成明显协变的形式

$$\left(i\hbar\frac{\partial}{\partial x^\mu} + \frac{e}{c}A_\mu(x)\right)\gamma^\mu\psi(x) - mc\psi(x) = 0 \tag{5.4.2}$$

练习 5-4-1 讨论方程(5.4.2)的 Lorentz 协变性。

Ⅱ.泡利方程—非相对论极限与电子内禀磁矩

为讨论 Dirac 方程(5.4.1)的非相对论近似,令

$$\psi(\boldsymbol{r},t) = \begin{pmatrix}\varphi(\boldsymbol{r},t)\\ \chi(\boldsymbol{r},t)\end{pmatrix}e^{-imc^2t/\hbar} \tag{5.4.3}$$

代入式(5.4.1)得

$$i\hbar\frac{\partial}{\partial t}\varphi(\boldsymbol{r},t)+e\phi(\boldsymbol{r},t)\varphi(\boldsymbol{r},t)=c\boldsymbol{\sigma}\cdot\left(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)\chi(\boldsymbol{r},t) \tag{5.4.4}$$

$$i\hbar\frac{\partial}{\partial t}\chi(\boldsymbol{r},t)+e\phi(\boldsymbol{r},t)\chi(\boldsymbol{r},t)=c\boldsymbol{\sigma}\cdot\left(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)\varphi(\boldsymbol{r},t)-2mc^2\chi(\boldsymbol{r},t) \tag{5.4.5}$$

在非相对论近似下,式(5.4.5)的左边相当于 $\chi(\boldsymbol{r},t)$ 乘以非相对论机械能,可以忽略。于是有

$$\chi(\boldsymbol{r},t)\cong\frac{1}{2mc}\boldsymbol{\sigma}\cdot\left(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)\varphi(\boldsymbol{r},t) \tag{5.4.6}$$

代入(5.4.4)式得

$$i\hbar\frac{\partial\varphi(\boldsymbol{r},t)}{\partial t}=\frac{1}{2m}\left[\boldsymbol{\sigma}\cdot\left(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)\right]^2\varphi(\boldsymbol{r},t)-e\phi(\boldsymbol{r},t)\varphi(\boldsymbol{r},t) \tag{5.4.7}$$

注意到(证明留作练习)

$$\left[\boldsymbol{\sigma}\cdot\left(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)\right]^2=\left(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2+\frac{e\hbar}{c}\boldsymbol{\sigma}\cdot\boldsymbol{B}(\boldsymbol{r},t) \tag{5.4.8}$$

式(5.4.7)可以写作

$$i\hbar\frac{\partial}{\partial t}\varphi(\boldsymbol{r},t)=\left[\frac{1}{2m}\left(\boldsymbol{p}+\frac{e}{c}\boldsymbol{A}(\boldsymbol{r},t)\right)^2+\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B}(\boldsymbol{r},t)-e\phi(\boldsymbol{r},t)\right]\varphi(\boldsymbol{r},t) \tag{5.4.9}$$

这正是 1927 年提出的 Pauli 方程。右边的第一项代表电子的机械动能,第三项是电子在库仑场中的势能,而第二项是电子内禀磁矩与磁场的相互作用能。后者在当时是 Pauli 唯象地引入的:

$$\boldsymbol{\mu}=-\frac{e\hbar}{2mc}\boldsymbol{\sigma}=-\mu_B\boldsymbol{\sigma}=-g_e\frac{e}{2mc}\boldsymbol{S} \tag{5.4.10}$$

μ_B 是 Bohr 磁子,$g_e=2$ 是电子的旋磁比。我们看到 Dirac 方程可以自然地得出电子的内禀磁矩和旋磁比。g_e 的精确的实验值要比 2 稍大,这种"反常磁矩"要由量子电动力学来解释。

练习 5-4-2 试证:式(5.4.8)成立。

Ⅲ. 中心力场下的非相对论近似，旋轨耦合

考虑中心力场情况：

$$\boldsymbol{A}(\boldsymbol{r},t)=0,\phi(\boldsymbol{r},t)=\phi(r) \tag{5.4.11}$$

由于哈密顿不显含时间，我们考虑定态解

$$\psi(\boldsymbol{r},t)=\begin{pmatrix}\varphi(\boldsymbol{r})\\ \chi(\boldsymbol{r})\end{pmatrix}e^{-iEt/\hbar} \tag{5.4.12}$$

定态方程为

$$\begin{pmatrix}mc^2-e\phi(r)-E & c\boldsymbol{\sigma}\cdot\boldsymbol{p}\\ c\boldsymbol{\sigma}\cdot\boldsymbol{p} & -mc^2-e\phi(r)-E\end{pmatrix}\begin{pmatrix}\varphi(\boldsymbol{r})\\ \chi(\boldsymbol{r})\end{pmatrix}=0 \tag{5.4.13}$$

令

$$E'=E-mc^2 \tag{5.4.14}$$

在非相对论近似下 $|E'|\ll mc^2$。式(5.4.13)可写作

$$[E'+e\phi(r)]\varphi(\boldsymbol{r})=c\boldsymbol{\sigma}\cdot\boldsymbol{p}\chi(\boldsymbol{r}) \tag{5.4.15}$$

$$c\boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi(\boldsymbol{r})=[E'+e\phi(r)+2mc^2]\chi(\boldsymbol{r}) \tag{5.4.16}$$

由最后一式得小分量

$$\chi(\boldsymbol{r})=\frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{2mc^2+E'+e\phi(r)}\varphi(\boldsymbol{r})=\frac{1}{2mc}\left(1+\frac{E'+e\phi(r)}{2mc}\right)^{-1}\boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi(\boldsymbol{r})$$

$$\approx\frac{1}{2mc}\left(1-\frac{E'+e\phi(r)}{2mc^2}\right)\boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi(\boldsymbol{r})$$

代入式(5.4.15)得大分量满足的方程

$$[E'+e\phi(r)]\varphi(\boldsymbol{r})=\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{2m}\left(1-\frac{E'+e\phi(r)}{2mc^2}\right)\boldsymbol{\sigma}\cdot\boldsymbol{p}\varphi(\boldsymbol{r}) \tag{5.4.17}$$

整理得

$$\left(\frac{\boldsymbol{p}}{2m}-\frac{E'\boldsymbol{p}^2}{4m^2c^2}-\frac{e}{4m^2c^2}(\boldsymbol{\sigma}\cdot\boldsymbol{p})\varphi(\boldsymbol{r})(\boldsymbol{\sigma}\cdot\boldsymbol{p})\right)\varphi(\boldsymbol{r})=[E'+e\phi(r)]\varphi(\boldsymbol{r}) \tag{5.4.18}$$

上式可以进一步改写作(留作练习)

$$\left(\frac{\boldsymbol{p}^2}{2m}-e\phi(r)-E'\right)\varphi(\boldsymbol{r})-\frac{1}{4m^2c^2}\boldsymbol{p}^2(e\phi(r)+E')\varphi(\boldsymbol{r})$$

$$-\frac{e}{4m^2c^2}\left[\frac{1}{r}\frac{\mathrm{d}\phi(r)}{\mathrm{d}r}\hbar(\boldsymbol{\sigma}\cdot\boldsymbol{l})+\hbar^2\nabla^2\phi(r)+\hbar^2\frac{\mathrm{d}\phi(r)}{\mathrm{d}r}\frac{\partial}{\partial r}\right]\varphi(\boldsymbol{r})=0 \tag{5.4.19}$$

上式第二项中的$(e\phi(r)+E')\varphi(\boldsymbol{r})$可以近似地代以$\frac{\boldsymbol{p}^2}{2m}\varphi(\boldsymbol{r})$,于是有

$$\left(\frac{\boldsymbol{p}^2}{2m}-e\phi(r)-\frac{1}{8m^3c^2}\boldsymbol{p}^4-\frac{1}{2m^2c^2}\frac{1}{r}\frac{\mathrm{d}\phi(r)}{\mathrm{d}r}(\boldsymbol{S}\cdot\boldsymbol{l})-\frac{e\hbar^2}{4m^2c^2}\nabla^2\phi(r)+\frac{e\hbar^2}{4m^2c^2}\frac{\mathrm{d}\phi(r)}{\mathrm{d}r}\frac{\partial}{\partial r}\right)\varphi(\boldsymbol{r})=E'\varphi(\boldsymbol{r}) \tag{5.4.20}$$

上式左边第三项是相对论动能修正(留作练习),第四项是 Thomas 修正项(旋轨耦合)。最后一项非厄米算符,这反映了$\langle\varphi|\varphi\rangle_+$不严格守恒。$H=H^+$保证的是$\langle\psi|\psi\rangle=\langle\psi|\varphi\rangle_++\langle\chi|\chi\rangle_-$守恒。

对类氢离子(氢原子)的情况

$$\varphi(r)=\frac{Ze}{r} \tag{5.4.21}$$

旋轨耦合为

$$\frac{Ze^2}{2m^2c^2r^3}\boldsymbol{S}\cdot\boldsymbol{l} \tag{5.4.22}$$

而第五项接触势(Darwin 项)为

$$\frac{\pi Z\hbar^2e^2}{2m^2c^2}\delta(\boldsymbol{r}) \tag{5.4.23}$$

练习 5-4-3 由式(5.4.18)推出式(5.4.19)。

思考题

既然旋轨耦合为$-\frac{1}{2m^2c^2}\frac{1}{r}\frac{\mathrm{d}\phi(r)}{\mathrm{d}r}(\boldsymbol{S}\cdot\boldsymbol{l})$,在自由电子的情况下应无旋轨耦合,为何轨道角动量或自旋角动量不分别守恒?

Ⅳ. 氢原子光谱的精细结构

从 Dirac 方程出发,我们可以严格地导出氢原子光谱的精细结构。能级与主量子数 n 和总角量子数 j 都有关,理论与实验结果符合的较好。但是随着实验技术的改

进,1945 年 Lamb 和 Retherford 利用微波技术精确的测定了氢原子光谱的精细结构;发现同一量子数的能级有微小的劈裂。例如 $2S_{1/2}$能级比 $2P_{1/2}$能级高出(1057.8 ± 0.1) MHz。这种"Lamb 移动"在量子电动力学中考虑辐射修正后可以很好地解释。

本书不拟用 Dirac 方程详细讨论氢原子光谱的精细结构。有兴趣的读者可参考曾谨言教授的《量子力学Ⅱ》,其中有详实、精辟的论述。

附录 群表示论简介

远在量子理论之前，居里先生(Pierre Curie)已认识到对称性分析在(经典)物理学中的重要性。然而，群论(特别是其表示理论)被物理学家普遍接受和使用，还是在量子力学被发现之后。这与量子体系的状态由态矢量描述，并满足态的迭加原理有关。经过了曲折和反复，群论已被公认为对称性分析的最恰当的语言与工具。本附录头两节简述了群概念和表示理论的基本内容，目的在于使读者熟悉本书使用的术语和记法。后几节介绍了正文讨论所需要的李群表示的一些知识，并非多数群论书都包含这些内容。

§1 群的概念

Ⅰ. 抽象群的定义

群是最基本的代数系。学好群的概念是学好其他代数系的前提条件。

定义 设 G 是非空集。若在 G 上定义了一个代数运算，即定义了一个映射 m: $G\times G\to G$，称之为乘法，对 $g,h\in G$，记 $m(g,h)\equiv gh$，称之为元素 g 与 h 之积，使下面条件成立时：

(1)$f(gh)=(fg)h$，$\forall f,g,h\in G$(结合律)

(2)存在 $e\in G$，使 $eg=ge=g$，$\forall g\in G$(有单位元)

(3)对每一 $g\in G$ 存在 $g^{-1}\in G$，使 $g^{-1}g=gg^{-1}=e$(每一元都有逆)

则称 G 为群。若还有 $gh=hg$，$\forall g,h\in G$，则称 G 为交换群或 Abel 群。

当群 G 是有限集时，称之为有限群，所含元素个数称作群的阶，记作 $|G|$。当群

G 是无限集时，称之为无限群。

例 1 设 $A=\{1,2,\cdots,n\}$ 是头 n 个自然数的集合。称 A 到自身的双射(双方一一对应) $P:A\to A$ 为 n 次置换，以 S_n 记所有 n 次置换的集合。以映射的复合(乘法)为乘法，S_n 是一个群。称作 n 次对称群。易知，$|S_n|=n!$。

通常以下面的方式记 n 次置换 $P=\begin{pmatrix}1 & \cdots & j & \cdots & n\\ P(1) & \cdots & P(j) & \cdots & P(n)\end{pmatrix}$，其中 $P(j)$ 是 j 在双射 P 下的像。当然 n 次置换还可以记作 $P=\begin{pmatrix}j & \cdots & 1 & \cdots & n\\ P(j) & \cdots & P(1) & \cdots & P(n)\end{pmatrix}=\begin{pmatrix}n & \cdots & 1 & \cdots & j\\ P(n) & \cdots & P(1) & \cdots & P(j)\end{pmatrix}$，等等，而不致引起歧义，只要第一行是 n 个不同的数，每个数下面是它在双射 P 下的像。设 $Q,P\in S_n$，则有

$$QP=\begin{pmatrix}1 & \cdots & j & \cdots & n\\ Q(P(1)) & \cdots & Q(P(j)) & \cdots & Q(P(n))\end{pmatrix}$$

S_n 群在数学上和量子力学中都非常重要。

例 2 n 维 F - 线性空间 V 上的可逆线性算符的全体，在算符乘法下构成群，称为 V 的一般线性群，记作 $GL(V)$。这是一个无限群。

例 3 按照刚体运动学，有一个固定点 o 的刚体的任一位移，可以由一个确定的绕通过 o 的固定轴的转动来实现。所以，我们谈到刚体绕固定点 o 的一个转动时，指的是有固定点 o 的刚体的一个位移。先后把刚体做两次绕固定点 o 的转动 R_1，R_2，其结果仍是一个绕此固定点 o 的转动，称之为 R_2 与 R_1 的乘积，记作 R_2R_1。在这个乘法下，刚体绕固定点的转动的全体构成一个群，称为三维空间(绕固定点 o 的)转动群，记作 $SO(3)$。

例 4 全体非零有理数的集合，在乘法下构成一交换群。

练习 1-1 试证：上面群的定义中的条件(2)和(3)可以等价地改为

(2′)存在 $e\in G$，使 $eg=g$，$\forall\, g\in G$。(存在左单位元)

(3′)对每一 $g\in G$，存在 $g^{-1}\in G$，使 $g^{-1}g=e$，(每一元都有关于左单位元 e 的左逆)

练习 1-2 试证：设 G 是群，$g,h\in G$。则下面方程的解存在并唯一：

(1) $xg=h$，(2) $gy=h$。

练习 1-3 试证：设 G 是群，g 是 G 的任给的一个元素。当 x 取遍群 G 的所有元素一遍时，x^{-1}，xg 或 gx 也取遍群 G 的所有元素一遍。（重排定理）

Ⅱ. 子群与陪集

定义 设 G 是群，H 是其非空子集。若 H 对群乘法是封闭的，即 $hk \in H, \forall h, k \in H$，这样 G 上的代数运算自然地在 H 上诱导出一个代数运算。如果 H 在这个代数运算下构成一个群，则称 H 为 G 的子群。

显然，每个群自身是其子群，只包含单位元的子集 $\{e\}$ 也是其子群。这两种子群是平庸子群。当子群是群的真子集时，称之为真子群（proper subgroup）。

上面例 1 中，全体 n 级偶置换的集合构成 S_n 的一个子群。称之为 S_n 的交错子群，记作 A_n。S_n 的子群统称置换群。上面例 3 中，全体绕过 o 点的某一固定轴的转动，构成 $SO(3)$ 群的一个子群，记之为 $SO(2)$。上面例 4 中，全体自然数的集合对乘法是封闭的，但不构成子群。这是因为大于 1 的自然数的倒数不是自然数。

定理 设 G 是群，H 是 G 的非空子集。则 H 是 G 的子群，当切仅当

$$hg \in H, h^{-1} \in H, \forall h, f \in H \quad (\text{或者 } hf^{-1} \in H, \forall h, f \in H)$$

证明：第一个断言的充分性。在 H 上诱导出的代数运算显然是结合的。由 $H \neq \varnothing$，可知有 $h \in H$，从而 $h^{-1} \in H$，$hh^{-1} = e = \in H$。H 包含单位元。H 包含其每个元的逆元。故 H 构成 G 的子群。

第一个断言的必要性。显然。第二个断言的证明留作练习。

定义 设 G 是群，H 是其子群，$g \in G$。称 G 的子集 $\{x \in G \mid x = gh, \textit{for some } h \in H,\} \equiv gH$ 为以 g 为代表元的 H 的左陪集。

定理 设 G 是群，H 是 G 的子群。则 $|H|$ 可整除 $|G|$（拉格朗日定理）。

证明：首先，显然有 $g = ge \in gH, \forall g \in G$，即每个元属于以它为代表元的左陪集。其次，$H$ 的两个左陪集若不交空，则必重合。事实上，若 $gH \cap fH \neq \varnothing$，则有 h_1，$h_2 \in H$，使 $gh_1 = fh_2$，从而 $g = fh_2h_1^{-1}$。对每一 $h \in H$，$gh = f(h_2h_1^{-1}h) \in fH$，故有 $gH \subset fH$。同理有 $gH \supset fH$。所以 $gH = fH$。这表明 H 的所有不同的左陪集给出 G 的一个分拆。由重排定理可知，$|gH| = |H|$，$\forall g \in G$。于是，当 G 是有限群时，我们有 $|H|$ 可整除 $|G|$。

设 G 是有限群，$g \in G$。考察群元素序列 $g^1(=g), g^2(=gg), g^3, \cdots, g^j, \cdots$ 由于

群是有限阶的,出现在这个序列中的群元素必有重复。设 $g^{1+m}(m\geq 1)$是序列中最早重复出现的元素 $g^i=g^{m+1}$,$1\leq i\leq m$。$g^i=g^ig^{m+1-i}$,由重排定理知 $g^{m+1-i}=e$。以 g 乘等式两边得 $g^{1+m+1-i}=g^1$,从而 $1+m+1-i\geq 1+m$,只能是 $i=1$。于是 $g^m=e$。而且若 $g^l=e$,则 l 必为 m 的整数倍。群元素序列 $g,g^2,\cdots$ 是 m 个不同元素 $g,g^2,\cdots,g^m(=e)$的不断重复。这 m 个不同元素构成 G 的一个子群,g 生成的循环子群。称 g 为 m 阶元素。由拉格朗日定理 m 可整除 $|G|$。

可以同样地讨论子群的右陪集 Hg。一般而言,$gH=Hg$ 不一定成立。

定义 设 N 是群 G 的子群,若 $gN=Ng$,$\forall g\in G$,则称 N 是 G 的正规子群。

定义 设 G 是群,$g,h\in G$。若存在 $f\in G$,使 $g=fhf^{-1}$,则称 g 共轭于 h,记作 $g\sim h$。易知 $g\sim g$,$\forall g\in G$(自反性);若 $g\sim h$,则 $h\sim g$(对称性);若 $h\sim g$ 且 $g\sim f$,则 $h\sim f$(传递性)。这表明共轭 $\sim$ 是群 G 上的一种等价关系。因而所有共轭类给出群 G 的一个分拆。群的单位元总是自成一共轭类。交换群的共轭类都只包含一个元素。$SO(3)$的共轭类由绕过固定点 o 的不同轴而转角相同的旋转构成。(请读者证明之)

练习 1-4 试证:设 G 是群,H 是其子群,$a,b\in G$。则 $aH=bH\Leftrightarrow a^{-1}b\in H$;$Ha=Hb\Leftrightarrow ab^{-1}\in H$。

练习 1-5 试证:有限群每个共轭类所含元素的个数是群阶的因子。

练习 1-6 试证:群 G 的子群 N 是正规子群的充分必要条件是:N 由群 G 的若干个(整)共轭类构成。

Ⅲ.同态的基本定理

设 G,M 是群,$f:G\to M$ 是 G 到 M 的映射。若对所有的 $g,h\in G$,都有 $f(gh)=f(g)f(h)$,则称 f 是群 G 到 M 的同态。当 f 是满射时,称 f 是满同态,称 M 是 G 的同态象。当 f 是双射时,称 f 是同构,称 M 是 G 的同构象。

定理 设 $f:G\to M$ 是群 G 到群 M 的满同态。若 e' 是群 M 的单位元,则 e' 在同态 f 下的完全原象(同态核)$f^{-1}(e')\equiv N$ 是群 G 的正规子群。$g'\in M$,则 $f^{-1}(g')$ 是 N 的陪集。

证明:留作练习。

设 N 是群 G 的正规子群。若 $aN=bN$,$a'N=b'N$,则 $(aa')N=(bb')N$。这就是

说，若 a, b 属于同一个陪集，a', b' 属于同一个陪集，则 aa', bb' 属于同一个陪集。事实上，由 $a = bh, a' = b'h'$ 可得 $aa' = bhb'h' = bb'[(b')^{-1}hb']h'$。由于 N 是群 G 的正规子群，$(b')^{-1}hb' \in N$。从而 aa', bb' 属于同一个陪集。

考虑 N 的陪集的全体所成的集合 $\frac{G}{N}$。我们来定义它的两个元素（两个 N 的陪集）的乘法：$(aN)(a'N) = aa'N, \forall a, a' \in G$。也就是说，两个陪集的乘积等于，从两个陪集各取一个代表元相乘，乘积所在的陪集。由于上一段的讨论，两个陪集的乘积与代表元的选取无关。由群 G 的乘法是结合的，可知陪集乘法是结合的。正规子群是陪集乘法的单位元。$a^{-1}N, aN$ 互为逆元。可见正规子群 N 的陪集的全体所成的集合 $\frac{G}{N}$，在此乘法下构成一个群，我们称之为群 G 关于正规子群 N 的商群。令 $\pi(a) = aN, \forall a \in G$，则不难验证：$\pi: G \to \frac{G}{N}$ 是群 G 到商群 $\frac{G}{N}$ 的满同态。

练习 1-7 讨论正规子群，群同态和商群三者间的关系。

§2 群表示理论

Ⅰ. 群表示的概念

有固定点 O 的刚体的每一位移，都可以通过绕某一过 O 点的固定轴的旋转（简称为绕固定点 O 的转动）来实现。如果不区分绕同一轴的转角差为 2π 整数倍的两个旋转，则刚体的保持固定点 O 不动的位移和刚体绕固定点 O 的转动之间是双方一一对应的。让刚体相继作两个绕 O 点的转动，先 g 后 h，其结果仍是刚体绕 O 点的一个转动，称之为 h 与 g 的乘积，记作 hg。在这种乘法下，全体绕 O 点的转动的集合 G 构成一个群。

考虑由 O 点引出的实向量 $\boldsymbol{r}$ 的全体所成的三维实向量空间 $\mathbf{R}^3$。把空间绕 O 点作一转动 $g \in G$ 时，$\mathbf{R}^3$ 中的每个向量 $\boldsymbol{r}$，变成一个新的向量 $\boldsymbol{r}' \in \mathbf{R}^3$。对应 $\boldsymbol{r} \mapsto \boldsymbol{r}'$（$\forall \boldsymbol{r} \in \mathbf{R}^3$）确定了 $\mathbf{R}^3$ 上的一个可逆线性算符 $R(g)$：

$$R(g)\boldsymbol{r} = \boldsymbol{r}', \forall \boldsymbol{r} \in \mathbf{R}^3 \tag{2.1}$$

由于相继作两个绕 O 点的转动，先 g 后 h，和绕 O 点作一个转动 hg 相比，空间最后位置是一样的，故有

$$R(h)R(g)\boldsymbol{r} = R(hg)\boldsymbol{r}, \forall\, \boldsymbol{r} \in \mathbf{R}^3 \tag{2.2}$$

从而

$$R(h)R(g) = R(hg), \forall\, g, h \in G \tag{2.3}$$

对旋转群 G 的单位元 e，显然有

$$R(e) = E \tag{2.4}$$

上式中的 E 是 $\mathbf{R}^3$ 上的恒等算符。显然有

$$R(g^{-1}) = R(g)^{-1} \tag{2.5}$$

上面的事例表明，把空间绕 O 点作移动 $g \in G$ 时，很自然地在三维实线性空间 $\mathbf{R}^3$ 上得到了一个非异的线性算符 $R(g)$，使 $R(hg) = R(h)R(g), \forall\, g, h \in G$。

再考察单个无旋粒子的态矢量空间 S，它是由定义在体系位形空间 $\mathbf{R}^3$ 上的模方可积的复值函数的全体所构成的

$$S = \left\{ \varphi : \mathbf{R}^3 \to \mathbf{C} \,\middle|\, \int \mathrm{d}\tau \mid \varphi(\boldsymbol{r}) \mid^2 < \infty \right\} \tag{2.6}$$

把实验室绕 O 点作一转动 $g \in G$ 时，$\mathbf{R}^3$ 中的每个向量 $\boldsymbol{r}$ 变到 $\boldsymbol{r}' = R(g)\boldsymbol{r}$，体系的波函数也应经受一变换 $\psi \mapsto \psi'$，新波函数 ψ' 在 $\boldsymbol{r}'$ 点的值，应等于老波函数 ψ 在 $\boldsymbol{r}$ 点的值

$$\psi'(\boldsymbol{r}') = \psi(\boldsymbol{r}) = \psi(R(g)^{-1}\boldsymbol{r}'), \forall\, \boldsymbol{r}' \in \mathbf{R}^3 \tag{2.7}$$

容易验证，对应 $\psi \mapsto \psi'$，$\forall\, \psi \in S$，确定了 S 上的一个非异的线性算符 $D(g)$：

$$D(g)\psi(\boldsymbol{r}) = \psi'(\boldsymbol{r}) = \psi(R(g)^{-1}\boldsymbol{r}) \tag{2.8}$$

$$D(hg) = D(h)D(g), \forall\, g, h \in G; D(e) = 1, \tag{2.9}$$

这个例子说明，把实验室绕 O 点作移动 $g \in G$，在体系的态矢量空间 S 上自然地诱导出一个非异线性算符 $D(g)$。对应 $g \mapsto D(g)$ 满足条件(2.9)。

现在让我们来给出群表示的定义。

定义 1 设 G 是群，V 是线性空间，$GL(V)$ 是 V 上非异线性算符的全体所成的群。若映射 $T: G \to GL(V)$，使

$$T(h)T(g) = T(hg), \forall\, g, h \in G \tag{2.10}$$

也就是说 T 是群 G 到 $GL(V)$ 的同态，则称 T 是群 G 在线性空间 V 上的一个表示。V

称为表示空间，dimV 称为表示的维数。当 T 是单同态时，称 $\{T(g)\mid g\in G\}$ 为 G 的一个忠实表示。

由以上群表示的定义立即可得

(ⅰ) $T(e)=1_V$（$e\in G$ 是群 G 的单位元，1_V 是 V 上恒等算符）

(ⅱ) $T(g^{-1})=T(g)^{-1}$（留作练习）

以 L_N 记 $N\times 1$ 矩阵的全体所成的 N 维线性空间。$N\times N$ 矩阵可以看成是 L_N 上的线性算符。以 $GL_N=GL(L_n)$ 记全体 $N\times N$ 非异矩阵所成的群。则群 G 到 GL_N 的同态

$$T: G\to GL_N \tag{2.11}$$

是群 G 在 L_N 上的表示。这种特殊的表示称为群 G 的矩阵表示。不难看到，若 $T: G\to GL(V)$ 是 G 的表示。选定 V 的一个基，以 $[A]$ 记 V 上线性算符 A 的矩阵。则 $g\mapsto[T(g)]$，$\forall g\in G$，确定群 G 的一个矩阵表示。

练习 2-1 试证：若 $T: G\to GL(V)$ 是群 G 的表示。则(ⅰ) $T(e)=1_V$，(ⅱ) $T(g^{-1})=T(g)^{-1}$，$\forall g\in G$。

练习 2-2 试证：若 $T: G\to GL(V)$ 是群 G 的表示。则同态象 $T(G)=\{T(g)\mid g\in G\}$ 是 $GL(V)$ 的子群。

Ⅱ. 等价表示

定义 2 设 $T_1: G\to GL(V_1)$，$T_2: G\to GL(V_2)$，是群 G 的两个表示。若存在线性空间同构 $S: V_1\to V_2$，使

$$ST_1(g)S^{-1}=T_2(g),\ \forall g\in G \tag{2.12}$$

则称群 G 的表示 T_1 等价于 T_2，记作 $T_1\sim T_2$。

以 $R(G)$ 记群 G 表示的全体所成的集合。不难证明“$\sim$”是 $R(G)$ 中的一种等价关系（即满足下面条件的二元关系）：

(ⅰ) $T\sim T$，$\forall T\in R(G)$（自反性）

(ⅱ) 若 $T_1\sim T_2$，则 $T_2\sim T_1$。（对称性）

(ⅲ) 若 $T_1\sim T_2$，$T_2\sim T_3$，则 $T_1\sim T_3$。（传递性）

这样我们可以把群 G 的所有表示分成等价类。

命题 1:设 $T_1: G\to GL(V_1)$, $T_2: G\to GL(V_2)$是群 G 的两个等价表示;则对 V_1 每一个基$\{e_1,\cdots,e_N\}$,都存在 V_2 的基,使 $T_1(g)$和 $T_2(g)$的矩阵相同

$$[T_1(g)] = [T_2(g)], \forall g \in G \tag{2.14}$$

事实上,若 $T_1 \sim T_2$,则存在线性同构 $S: V_1\to V_2$,使

$$ST_1(g)S^{-1} = T_2(g), \forall g \in G$$

$\{e_1,\cdots,e_N\}$是 V_1 的基,则$\{Se_1,\cdots,Se_N\}$是 V_2 的基。按 §0.3, $T_1(g)$在基$\{e_1,\cdots,e_N\}$下的矩阵和 $T_2(g)$在基$\{Se_1,\cdots,Se_N\}$的矩阵分别由下面的式子确定

$$T_1(g)e_\beta = e_\alpha T_1(g)^\alpha_\beta \tag{2.15}$$

$$T_2(g)Se_\beta = Se_\alpha T_2(g)^\alpha_\beta \tag{2.16}$$

以 S 作用于(2.15)的两边得

$$ST_1(g)S^{-1}Se_\beta = Se_\alpha T_1(g)^\alpha_\beta \tag{2.17}$$

与(2.16)对比得 $T_2(g)^\alpha_\beta = T_1(g)^\alpha_\beta$,此即$[T_1(g)] = [T_2(g)]$。

可见等价表示的算符总可以有相同的矩阵。

定义 3:设 V 是内积空间, $T: G\to GL(V)$是群 G 在 V 上的表示。若 $T(g)^+ = T(g)^{-1}$, $\forall g\in G$,此即

$$\langle T(g)x \mid T(g)y\rangle = \langle x \mid y\rangle, \quad \forall g \in G, \quad x,y \in V \tag{2.18}$$

即表示 T 的每个算符 $T(g)$都是酉算符,则称 T 是群 G 在内积空间 V 上的酉表示。

由 §0.3 知,若选择 V 的正交归一基,则酉表示 T 的每个算符 $T(g)$的矩阵$[T(g)]$是酉矩阵:

$$[T(g)]^+ = [T(g)]^{-1}。$$

命题 2:有限群的有限维表示 T 总与一个酉表示等价。

事实上,设 G 是有限群, $\dim V = N$, $T: G\to GL(V)$是 G 在 N 维线性空间 V 上的表示。总可以定义 V 中向量的内积$\langle\mid\rangle$,使 V 成为一个内积空间。由 V 的内积$\langle\mid\rangle$,利用下式可定义 V 的另一种内积$\langle\parallel\rangle$:

对所有 $x,y\in V$,令

$$\langle x \parallel y\rangle = \frac{1}{|G|}\sum_{g\in G}\langle T(g)x \mid T(g)y\rangle \tag{2.19}$$

不难看出

$$\langle T(h)x \parallel T(h)y\rangle = \langle x \parallel y\rangle, \forall h \in G, x,y \in V \tag{2.20}$$

设$\{e_1,\cdots,e_N\}$是 V 的关于内积$\langle|\rangle$的正交归一基,$\{e'_1,\cdots,e'_N\}$是 V 的关于内积$\langle\|\rangle$的正交归一基。由条件

$$Se'_\alpha = e_\alpha \quad (1 \leq \alpha \leq N) \tag{2.21}$$

可确定 V 的自同构$S: V\to V$。显然有

$$\langle e'_\alpha \mid e'_\beta\rangle = \delta_{\alpha\beta} = \langle e_\alpha \mid e_\beta\rangle = \langle Se'_\alpha \mid Se'_\beta\rangle \tag{2.22}$$

从而

$$\langle x \| y\rangle = \langle Sx \mid Sy\rangle, \forall\, x,y \in V \tag{2.23}$$

$$\langle S^{-1}x \| S^{-1}y\rangle = \langle x \mid y\rangle, \forall\, x,y \in V \tag{2.24}$$

令 $T'(g) = ST(g)S^{-1}, \forall\, g\in G$,显然 T'是群 G 的与 T 等价的表示。不难看到,对所有 $x,y\in V, g\in G$ 都有

$$\begin{aligned}\langle x \mid y\rangle &= \langle S^{-1}x \| S^{-1}y\rangle = \langle T(g)S^{-1}x \| T(g)S^{-1}y\rangle \\ &= \langle ST(g)S^{-1}x \mid ST(g)S^{-1}y\rangle = \langle T'(g)x \mid T'(g)y\rangle\end{aligned} \tag{2.25}$$

可见 T'是群 G 在内积空间$\{V,\langle|\rangle\}$上的酉表示。

从上面的讨论可知,有限群的每个有限维表示所在的等价类都有幺正表示。互相等价的表示的算符在适当的基下矩阵相同。所以要研究有限群有限维表示的性质,只要在每个等价类中挑出一个表示来研究就行。由于幺正表示有一些"好的"性质(见下面的讨论),我们常选幺正表示为其所在等价类的代表。

Ⅲ. 既约表示与可约表示

设 A 是线性空间 V 上的线性算符,W 是 V 的子空间。若对所有的 $x\in W$ 都有 $Ax\in W$,则称 W 是 V 的 A – 不变子空间,或算符 A 的不变子空间。

显然 V 和零向量子空间$\{0\}$是 V 上任一线性算符的不变子空间,称之为平庸的不变子空间。

设$\{A_\sigma \mid \sigma\in\sum\}$($\sum$是下标的集合)是 V 上线性算符组,W 是 V 的子空间。若 W 是每个A_σ 的不变子空间,则称 W 是 V 的$\{A_\sigma \mid \sigma\in\sum\}$ – 不变子空间,或$\{A_\sigma \mid \sigma\in\sum\}$的不变子空间。

定义 4:设 $T: G\to GL(V)$是群 G 的表示。若不存在算符组$\{T(g)\mid g\in G\}$的非平庸不变子空间,则称 T 为群 G 的既约表示;否则称 T 为群 G 的可约表示。

设 $T: G\to GL(V)$是群 G 的表示,W 是$\{T(g)\mid g\in G\}$的不变子空间。令群 G 的

每个元 $g\in G$ 对应于 $T(g)$ 到不变子空间 W 上的限制 $T(g)|_W\equiv T_1(g)$，易知 $T_1: G\to GL(W)$，或 $g\overset{T_1}{\mapsto}T_1(g)\equiv T(g)|_W$，$\forall g\in G$ 是群 G 在 W 上的一个表示，称之为表示 T 到不变子空间 W 上的限制。

命题 3：设 V 是有限维内积空间，$T: G\to GL(V)$ 是群 G 的酉表示，W 是 V 的子空间。若 W 是 $\{T(g)|g\in G\}$ 的不变子空间，则 W 的正交补 W' 也是 $\{T(g)|g\in G\}$ 的不变子空间。

事实上，由 §0.3 知，$T(g)$ 的不变子空间 W 的正交补 W' 是 $T(g)^+$ 的不变子空间。由表示 T 的幺正性可知 $T(g)^+=T(g)^{-1}=T(g^{-1})$。故 W' 是 $T(g^{-1})$ 的不变子空间。当 g 跑遍 G 时，g^{-1} 也跑遍 G。故 W' 是 $\{T(g^{-1})|g\in G\}=\{T(g)|g\in G\}$ 的不变子空间。

这个命题告诉我们，群的酉表示 $T: G\to GL(V)$ 如果有非平庸不变子空间 W，则 W 的正交补空间 W' 也是 T - 不变子空间。T 的表示空间 V 可以写成两个互相正交的 T - 不变子空间 W 与 W' 的直和

$$V=W\oplus W'$$

$T(g)$ 到 W 上的限制 $\{T(g)|_W\equiv T_1(g)|g\in G\}$ 构成群 G 在 W 上的一个酉表示；$T(g)$ 到 W' 上的限制 $\{T(g)|_{W'}\equiv T_2(g)|g\in G\}$ 构成 G 在 W' 上的一个酉表示。我们有

$$T(g)=T_1(g)\oplus T_2(g),\forall g\in G \tag{2.26}$$

这时称群 G 的表示 T 是维数较低的表示 T_1 与 T_2 的直和。如果 T_1 和 T_2 都是既约的，则 T 是既约表示的直和；否则，例如 T_1 有非平庸不变子空间 $U\subset W$，则 U 在 W 中的正交补空间 U' 也是 T_1 - 不变子空间，我们有 $W=U\oplus U'$ 和 $T_1(g)=T_1(g)|_U\oplus T_1(g)|_{U'}$，$\forall g\in G$。从而 T_1 是两个维数更低的酉表示的直和。可见群的有限维酉表示是若干个既约表示的直和。有限群的有限维表示总等价于一个酉表示，故它们一定可以写成若干个既约表示的直和。

这样，研究有限群有限维表示的问题，可以归为研究有限群的有限维既约表示。

Ⅳ. 正交性定理与完备性定理

考虑有限群 G 到复数域 $\mathbf{C}$ 的映射

$$\varphi: G \to \mathbf{C} \tag{2.27}$$

这种映射是以群 G 的元素为自变量的复值函数,称之为群上函数。以 $\Phi(G)$记群 G 的全体群上函数的集合。在通常的函数加法和数乘的定义下,$\Phi(G)$显然是一个复线性空间。其维数等于群 G 的阶:

$$\dim\Phi(G) = |G| \tag{2.28}$$

对 $\varphi, \psi \in \Phi(G)$,令

$$\langle \varphi | \psi \rangle = \frac{1}{|G|}\sum_{g\in G}\overline{\varphi(g)}\psi(g) \tag{2.29}$$

易知$\langle | \rangle$满足内积的公理(见 §0.2),这样 $\Phi(G)$是$|G|$维内积空间。

定理 1:设 $\tau: G \to GL_d$ 是有限群 G 的 d 维酉矩阵表示。则作为群上函数的 $\tau_{\mu\nu}(g)$(矩阵 $\tau(g)$的 μ 行 ν 列矩阵元)满足下列正交归一化条件

$$\langle \tau_{\mu\nu} | \tau_{\rho\sigma} \rangle \equiv \frac{1}{|G|}\sum_{g\in G}\overline{\tau_{\mu\nu}(g)}\tau_{\rho\sigma}(g) = \frac{1}{d}\delta_{\mu\rho}\delta_{\nu\sigma} \tag{2.30}$$

证明从略。

定理 2:设 $\tau^{[\alpha]}$和 $\tau^{[\beta]}$是有限群 G 的两个不等价的不可约酉矩阵表示 ,则

$$\langle \tau^{[\alpha]}_{\mu\nu} | \tau^{[\beta]}_{\rho\sigma} \rangle = 0 \tag{2.31}$$

定理 1 和定理 2 可以合起来表作正交性定理:

$$\langle \tau^{[\alpha]}_{\mu\nu} | \tau^{[\beta]}_{\rho\sigma} \rangle = \frac{1}{d_\alpha}\delta_{\alpha\beta}\delta_{\mu\rho}\delta_{\nu\sigma} \tag{2.32}$$

由此定理可知,有限群 G 的不等价的不可约表示的维数的平方和小于等于群的阶$|G|$。下面我们将指出,实际情况是二者正好相等。

设 G 是有限群,$h \in G$。定义$|G|$ – 维复内积空间 $\Phi(G)$上的算符 $R(h)$:对 $\varphi \in \Phi(G)$

$$R(h)\varphi(g) = \varphi(gh), \forall\, g \in G \tag{2.33}$$

不难验证 $R(h)$是 $\Phi(G)$上的酉线性算符,而且

$$R(f)\varphi(h) = \varphi(fh), \forall\, f, h \in G \tag{2.34}$$

可见 $R: G \to GL[\Phi(G)]$是群 G 在$\Phi(G)$上的一个酉表示,称之为右正则表示。有限群的有限维酉表示 R 可以写作 G 的若干个不可约酉表示的直和

$$R = \bigoplus_\alpha \tau^{[\alpha]} \tag{2.35}$$

相应地,表示空间

$$\Phi(G) = \bigoplus_{\alpha} V^{[\alpha]} \tag{2.36}$$

选取 $V^{[\alpha]}$的正交归一基$\{e_V^{[\alpha]} \mid 1 \leq v \leq d_\alpha\}$，把它们合在一起便得到 $\Phi(G)$的正交归一基

$$\bigcup_{\alpha} \{e_V^{[\alpha]} \mid 1 \leq v \leq d_\alpha\} \tag{2.37}$$

对每一 $\varphi \in \Phi(G)$，都可唯一地写成(2.37)的线性组合

$$\varphi(g) = \sum_{\alpha, v} e_V^{[\alpha]}(g) C_v^{[\alpha]} \tag{2.38}$$

以 $R(h)$作用于上式两边

$$\varphi(gh) = R(h)\varphi(g) = \sum_{\alpha, \mu, v} e_\mu^{[\alpha]}(g) \tau_{\mu v}^{[\alpha]}(h) C_v^{[\alpha]} \tag{2.39}$$

当 $g = e$ 时，上式给出

$$\varphi(h) = \sum_{\alpha \mu v} \tau_{\mu v}^{[\alpha]}(h) e_\mu^{[\alpha]}(e) C_v^{[\alpha]}, \forall\, h \in G \tag{2.40}$$

这表明任一群上函数 φ 都可以写作为群上函数的不可约表示矩阵元 $\tau_{\mu v}^{[\alpha]}$的线性组合。因而有限群 G 的所有不等价不可约表示的矩阵元 $\tau_{\mu v}^{[\alpha]}$构成 $\Phi(G)$的一个正交基，而

$$\{\sqrt{d_\alpha}\tau_{\mu v}^{[\alpha]} \mid 1 \leq \mu, v \leq d_\alpha, \alpha \text{ 取遍 } G \text{ 的所有不等价的不可约表示}\} \tag{2.41}$$

是 $\Phi(G)$的正交归一基。考察下式

$$R(h)\tau_{\mu v}^{[\alpha]}(g) = \tau_{\mu v}^{[\alpha]}(gh) = \sum_{\sigma=1}^{d_\alpha} \tau_{\mu\sigma}^{[\alpha]}(g) \tau_{\sigma v}^{[\alpha]}(h) \tag{2.42}$$

可见$\{\tau_{\mu v}^{[\alpha]}(g) \mid 1 \leq v \leq d_\alpha\}$张成 $\Phi(G)$的 R – 不变子空间，按 G 的不可约表示 $\tau^{[\alpha]}$变换。d_α 维不可约表示在右正则表示中出现 d_α 次。

V. 特征标

设 $T: G \to GL(V)$是群 G 在线性空间 V 上的一个表示。选定 V 的一个基，则表示 T 的每个算符 $T(g)$对应一个方阵$[T(g)]$。定义表示 T 的特征标 χ 为这样一个群上函数：

$$\chi(g) = tr[T(g)] = \sum_{\mu} T_{\mu\mu}(g) \tag{2.43}$$

在 改换 V的基时，$T(g)$的矩阵经受相似变换，阵迹不变。因而特征标与 V的基的选

取无关。显然：

(ⅰ)等价表示的特征标一样。

(ⅱ)若 $T = T_1 \oplus \cdots \oplus T_S$，则 $\chi = \chi_1 \oplus \cdots \oplus \chi_S$。

(ⅲ)不可约表示的特征标满足正交归一化条件

$$\langle \chi^{[\alpha]} \mid \chi^{[\beta]} \rangle = \sum_{\rho,\sigma} \langle \tau_{\rho\rho}^{[\alpha]} \mid \tau_{\rho\rho}^{[\beta]} \rangle = \frac{\delta_{\alpha\beta}}{d_\alpha} \sum_{\rho\sigma} \delta_{\rho\sigma} = \delta_{\alpha\beta} \tag{2.44}$$

(ⅳ)可见：$\langle \chi \mid \chi \rangle = 1$，当且仅当 T 是不可约表示。对可约表示总有 $\langle \chi \mid \chi \rangle > 1$。

(ⅴ)作为群上函数，特征标的值只依赖自变量所在的共轭类。事实上，若 $hgh^{-1} = f$，即 g 与 f 共轭，则 $T(f) = T(h)T(g)T(h^{-1})$，

$$\chi(f) = tr(T(h)T(g)T(h^{-1})) = trT(g) = \chi(g) \tag{2.45}$$

称函数值只依赖自变量 g 所在的共轭类的群上函数为类函数。类函数的全体

$$\{\varphi \in \Phi(G) \mid \varphi(g) = \varphi(hgh^{-1}), \forall g, h \in G\} \equiv K(G) \tag{2.46}$$

构成 $\Phi(G)$ 的一个子空间。

(ⅵ)群 G 的不等价不可约表示的特征标 $\{\chi^{[\alpha]} \mid \alpha$ 取遍 G 的所有不等价不可约表示$\}$ 构成 $K(G)$ 的正交归一基。

正交归一性已证，现只须证明完备性。设 $\varphi \in K(G) \subset \Phi(G)$，此即

$$\varphi(hgh^{-1}) = \varphi(g,) \forall g, h \in G \tag{2.47}$$

φ 总可以向 G 的不等价不可约酉矩阵表示的阵元展开

$$\varphi(g) = \sum_{\alpha\mu\nu} \tau_{\mu\nu}^{[\alpha]} C_{\mu\nu}^{[\alpha]} \tag{2.48}$$

$\because \varphi$ 是类函数，故

$$\begin{aligned} \varphi(g) = \varphi(hgh^{-1}) &= \frac{1}{|G|} \sum_{h \in G} \varphi(hgh^{-1}) \\ &= \sum_{\alpha\mu\nu} \sum_{\rho\sigma} \frac{1}{|G|} \sum_{h \in G} \tau_{\mu\rho}^{[\alpha]}(h) \tau_{\rho\sigma}^{[\alpha]}(g) \tau_{\sigma\nu}^{[\alpha]}(h^{-1}) C_{\mu\nu}^{[\alpha]} \\ &= \sum_{\alpha\mu\nu} \sum_{\rho\sigma} \langle \tau_{\nu\sigma}^{[\alpha]} \mid \tau_{\mu\rho}^{[\alpha]} \rangle \tau_{\rho\sigma}^{[\alpha]}(g) C_{\mu\nu}^{[\alpha]} \\ &= \sum_{\alpha} \chi^{[\alpha]}(g) \sum_{\mu} \frac{C_{\mu\nu}^{[\alpha]}}{d_\alpha} \end{aligned} \tag{2.49}$$

可见任一类函数 φ 都可以向群 G 不可约表示特征标展开。

于是有限群不等价不可约表示的个数等于其元素共轭类的个数。

§3 李群表示的几个定理

Ⅰ. r 维李群

严格地定义连续群(李群)需要用到流形的概念。这超出了一般的物理专业硕士生的知识水平。我们将满足于下面的描述。

设 G 是不可数无限群。若须用 r 个连续变化的(和若干个取分离值的)实参数来标记它的元素,则称 G 是 r 维李群。这 r 个连续变化的(和若干个取分离值的)实参数的取值范围叫做 G 的群流形。

例 1 三维实欧氏空间 R^3 的平移群是三维李群,其群流形是 R^3。

例 2 绕固定轴转动的全体所构成的群 $SO(2)$是一维李群,其群流形是圆周。

例 3 全体非零实数的乘法群是一维李群。标记它的元素须用一个连续变化的实参数(取值范围是开区间$(0, +\infty)$)和一个取分离值$\{+1, -1\}$的实参数。群流形是去掉 0 点的数直线,它包含两个互相分开的连通分支。包含群的单位元的连通分支构成一个子群。

例 4 三维空间 R^3 绕固定点 O 转动的全体所构成的群 $SO(3)$是三维李群。在前面我们是用一个向量 $\boldsymbol{\theta}$(大小等于转角,方向为由右手螺旋法决定的转轴正方向)来标志的。有时也用沿转轴的单位向量 $\hat{n}(=\boldsymbol{\theta}/\theta)$和转角 θ 来标志旋转 $\boldsymbol{\theta}=\hat{n}\theta$。容易看出旋转群 $SO(3)$的群流形是一个半径为 π 的闭球,只不过把同一直径的两个端点(对径点)看成同一个点。在这个群流形中,连接球心(单位元)和任一给定点的道路可以分成两类,同类道路可经由连续变形互相转换,不同类道路则不可经由连续变形互相转换。图 3-1 中上边各条连接 O, P 两点的道路,各包含偶数对对径点,它们可以经由连续变形而相互转换;下边各条连接 O, P 两点的道路,各包含奇数对对径点,它们可以经由连续变形而相互转换。但是包含奇数对(偶数对)对径点的道路,在端点不动时,不能经连续变形变为包含偶数对(奇数对)对径点的道路。

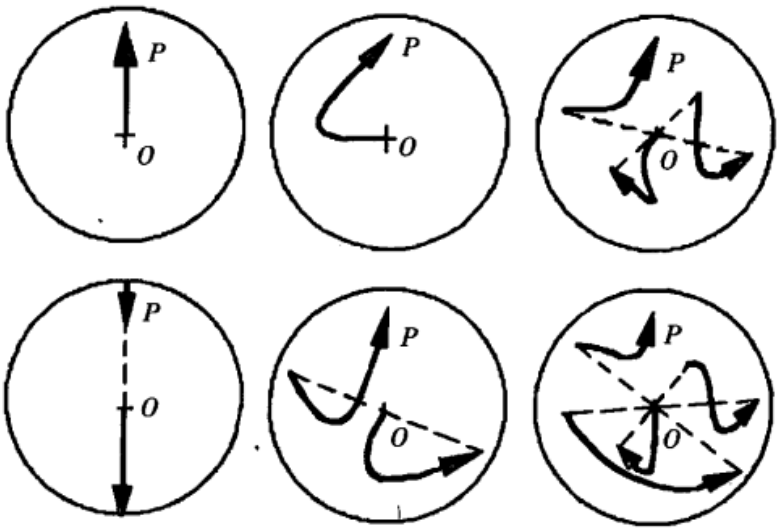

图 3-1

相应地，通过球心 O 的闭合道路可分成两类，一类可连续地变形缩成一点 O，另一类可连续地变形变为一个直径(对径点是同一点，故为闭合道路)(见图 3-2 的上排和下排)。



图 3-2

在以后的讨论中，我们将采用下面的记法。设 G 是 r 维李群。标记元素 $g \in G$ 的参数将被记作$(g^1, \cdots, g^\alpha, \cdots, g^r)$。我们约定单位元的参数全为零：$e^\alpha = 0(\alpha = 1, \cdots r)$。

两个群元素的乘积的参数是这两个群元素的参数的函数：

$$(hg)^\alpha = \varphi^\alpha(h^1,\cdots,h^r,g^1,\cdots,g^r) \equiv \varphi^\alpha(h,g) \tag{3.1}$$

乘法的结合律$(h(gf))^\alpha = ((hg)f)^\alpha$意味着函数$\varphi^\alpha(\alpha=1,\cdots,r)$满足方程

$$\varphi^\alpha(h^1,\cdots,h^r,\varphi^1(g,f),\cdots,\varphi^r(g,f)) = \varphi^\alpha(\varphi^1(h,g),\cdots,\varphi^r(h,g),f^1,\cdots,f^r) \tag{3.2}$$

单位元的性质$eg = ge = g, \forall g \in G$意味着

$$\varphi^\alpha(0,\cdots,0,g^1,\cdots,g^r) = \varphi^\alpha(g^1,\cdots,g^r,0,\cdots,0) = g^\alpha \tag{3.3}$$

设$h = g^{-1}$，则有

$$(hg)^\alpha = \varphi^\alpha(h^1,\cdots,h^r,g^1,\cdots,g^r) \equiv \varphi^\alpha(h,g) = 0, \forall 1 \leq \alpha \leq r$$

$$(gh)^\alpha = \varphi^\alpha(g^1,\cdots,g^r,h^1,\cdots,h^r) \equiv \varphi^\alpha(g,h) = 0, \forall 1 \leq \alpha \leq r$$

当李群及其参数的给法确定以后，函数φ^α的形式就确定了，函数φ^α的形式与群表示无关。

Ⅱ. 不变体积元

有限群的某些性质，例如每个有限维表示都是若干个不可约表示的直和，对无限群就不一定成立了。考察下面的例子：设I是全体整数的加群。对每个整数k，令2×2矩阵$\begin{bmatrix}1 & k\\ 0 & 1\end{bmatrix} \equiv T(k)$与之对应。不难验证$\begin{bmatrix}1 & k\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & l\\ 0 & 1\end{bmatrix} = \begin{bmatrix}1 & k+l\\ 0 & 1\end{bmatrix}$，$T$是全体整数的加群$I$的一个二维矩阵表示。它的表示空间是全体$2\times 1$复矩阵构成的二维复线性空间。$\begin{bmatrix}1\\ 0\end{bmatrix}$张成的一维子空间是表示$T$的不变子空间。因而$T$是可约的。当$k\neq 0$时，$\begin{bmatrix}1 & k\\ 0 & 1\end{bmatrix}$不能通过相似变换对角化(留作练习)。这说明群$I$的可约表示$T$不是两个一维不可约表示的直和。有限群的每个表示都是若干个不可约表示的直和，这一性质依赖于有限群的下述性质：对有限群G上的任一函数$\varphi: G\to C$，

$$\sum_{g\in G}\varphi(g) = \sum_{g\in G}\varphi(hg) = \sum_{g\in G}\varphi(gh), \forall h \in G \tag{3.4}$$

有一类李群，如$SO(2)$，$SO(3)$，其群流形是闭有界的，从拓扑学知，是紧致流形，故称作紧致李群。全体实数的加群不是紧致的，因为其群流形数直线$(-\infty,\infty)$不是有界闭的。紧致李群具有与有限群性质(3.4)类似的性质：存在群上非负实值连续函数$\rho: G\to R$，使$\int_G \rho(g)\mathrm{d}g > 0$，而且对任一群上复值连续函数$\varphi: G\to C$，都有

$$\int_G \varphi(g)\rho(g)\mathrm{d}g = \int_G \varphi(hg)\rho(g)\mathrm{d}g = \int_G \varphi(gh)\rho(g)\mathrm{d}g \qquad \forall\, h \in G \quad (3.5)$$

式中 $\rho(g)\mathrm{d}g$ 叫做群的不变体积元。其证明读者可参考 Miller 的书《对称性群及其应用》。

Ⅲ. 李群表示与无穷小算符

设 T 是 r 维李群 G 在复线性空间 V 上的表示。记 $T(g) \equiv T(g^1, \cdots, g^r)$。单位元对应于恒等算符：$T(e) = T(0, \cdots, 0) = 1_V$；同态 $T(h)T(g) = T(hg)$ 意味着

$$T(h^1, \cdots, h^r)T(g^1, \cdots, g^r) = T(\varphi^1(h, g), \cdots, \varphi^r(h, g)) \qquad (3.6)$$

表示 T 的无穷小算符是

$$I_\alpha = \frac{\partial}{\partial g^\alpha} T(g) \mid_{g=e} (\alpha = 1, \cdots, r) \qquad (3.7)$$

由上式可知，无穷小算符由表示 T 确定。下面将证明，李群 G 的表示 T 完全由它的无穷小算符确定。无穷小算符在李群表示理论中扮演着非常重要的角色。

现在介绍一个非常有用的方程，它给出李群表示 T 的无穷小算符 I_α 和表示群的任一元素 $g \in G$ 的算符 $T(g)$ 之间的联系，是我们讨论李群表示性质的出发点。设 T 是 r 维李群 G 在复线性空间 V 上的表示。显然有 $T(g) = T(gh^{-1})T(h)$，$\forall\, g, h \in G$。取等式两边对群参数 g^α 的偏导数 $\frac{\partial}{\partial g^\alpha} T(g) = \frac{\partial (gh^{-1})^\rho}{\partial g^\alpha} \frac{\partial T(gh^{-1})}{\partial (gh^{-1})^\rho} T(h)$，$\forall\, g, h \in G$。当 $h = g$ 时给出

$$\frac{\partial}{\partial g^\alpha} T(g) = \left. \frac{\partial (gh^{-1})^\rho}{\partial g^\alpha} \right|_{h=g} I_\rho T(g) \equiv S_\alpha^\beta(g) I_\beta T(g), \quad \forall\, g \in G \qquad (3.8)$$

上式是李群 G 在复线性空间 V 上的表示 T 所满足的微分方程。注意，并非一阶偏微分方程组(3.8)所有的解都是李群 G 的表示，李群 G 的表示 T 是它的满足下面初始条件的解

$$T(e) = 1_V \qquad (3.9)$$

微分方程理论告诉我们，一阶偏微分方程组(3.8)的满足初始条件(3.9)的解，如果存在的话，是唯一的。对给定的李群与其参数给法，偏微分方程组(3.8)完全由无穷小算符确定。因而，李群 G 的表示 T 完全由它的无穷小算符确定。

定理 1 r 维李群 G 的表示 T 的无穷小算符满足如下的对易关系：

$$[I_\alpha, I_\beta] = \sum_{\gamma=1}^{r} I_\gamma C_{\alpha\beta}^{\gamma} \tag{3.10}$$

式中系数 $C_{\alpha\beta}^{\gamma}$ 只与李群及其参数的给法有关，与表示无关，称作该李群的结构常数。

证明 r 维李群 G 的表示 T 显然满足微分方程组(3.8)和初始条件(3.9)。利用混合偏导数与顺序无关 $\dfrac{\partial^2 T(g)}{\partial g^\beta \partial g^\alpha} = \dfrac{\partial^2 T(g)}{\partial g^\alpha \partial g^\beta}$，由微分方程组(3.8)可得

$$\left.\frac{\partial (gh^{-1})^\rho}{\partial g^\alpha}\right|_{h=g} \left.\frac{\partial (gh^{-1})^\sigma}{\partial g^\alpha}\right|_{h=g} (I_\rho I_\sigma - I_\sigma I_\rho)$$

$$= \left\{\frac{\partial}{\partial g^\beta}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\alpha}\right|_{h=g}\right] - \frac{\partial}{\partial g^\alpha}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\beta}\right|_{h=g}\right]\right\} I_\tau \tag{3.11}$$

当 $g = e$ 时，由于 $\left.\dfrac{\partial (gh^{-1})^\rho}{\partial g^\alpha}\right|_{h=g=e} = \delta_\alpha^\rho$，上式给出式(3.10)，其中

$$C_{\alpha\beta}^{\tau} = \left\{\frac{\partial}{\partial g^\alpha}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\beta}\right|_{h=g}\right] - \frac{\partial}{\partial g^\beta}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\alpha}\right|_{h=g}\right]\right\}_{g=e} \tag{3.12}$$

定理证完。

把(3.10)代入(3.11)得，

$$\left.\frac{\partial (gh^{-1})^\rho}{\partial g^\alpha}\right|_{h=g} \left.\frac{\partial (gh^{-1})^\sigma}{\partial g^\beta}\right|_{h=g} C_{\rho\sigma}^{\tau} I_\tau$$

$$= \left\{\frac{\partial}{\partial g^\beta}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\alpha}\right|_{h=g}\right] - \frac{\partial}{\partial g^\alpha}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\beta}\right|_{h=g}\right]\right\} I_\tau$$

上式是无穷小算符的一个线性关系，系数与表示无关，只由李群及其参数的给法确定。如果李群 G 有这样的表示，其 r 个无穷小算符是线性独立的，则我们得到 r 个有关李群的恒等式

$$\left\{\frac{\partial}{\partial g^\beta}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\alpha}\right|_{h=g}\right] - \frac{\partial}{\partial g^\alpha}\left[\left.\frac{\partial (gh^{-1})^\tau}{\partial g^\beta}\right|_{h=g}\right]\right\}$$

$$+ \left.\frac{\partial (gh^{-1})^\rho}{\partial g^\alpha}\right|_{h=g} \left.\frac{\partial (gh^{-1})^\sigma}{\partial g^\beta}\right|_{h=g} C_{\rho\sigma}^{\tau} = 0 \quad (\tau = 1, \cdots, r) \tag{3.13}$$

而实际上，每一 r 维李群的左(右)正则表示的 r 个无穷小算符是线性独立的(留作练习)，故上式对所有李群成立。由(3.10)和算符恒等式

$$[I_\alpha, I_\beta] = -[I_\beta, I_\alpha], [I_\alpha, [I_\beta, I_\gamma]] + [I_\beta, [I_\gamma, I_\alpha]] + [I_\gamma, [I_\alpha, I_\beta]] = 0$$

可知

$$C^{\gamma}_{\alpha\beta} = -C^{\gamma}_{\beta\alpha}, C^{\sigma}_{\alpha\rho}C^{\rho}_{\beta\gamma} + C^{\sigma}_{\beta\rho}C^{\rho}_{\gamma\alpha} + C^{\sigma}_{\gamma\rho}C^{\rho}_{\alpha\beta} = 0 \tag{3.14}$$

命题 1 设 T, T' 是 r 维李群 G 的两个表示，其表示空间分别是 V, V'，而 $A: V \to V'$ 是可逆线性变换。则

$$T'(g) = AT(g)A^{-1}, \forall g \in G \tag{3.15}$$

的充分必要条件是

$$I'_{\alpha} = AI_{\alpha}A^{-1}, \forall \alpha = 1, \cdots, r \tag{3.16}$$

（可见，表示 T, T' 等价，当且仅当它们的无穷小算符组等价）

证明 必要性。只要取式(3.15)两边对群参数 g^{α} 的偏导数，再令 $g = e$ 就可得到式(3.16)。

充分性。表示 T, T' 都满足一阶偏微分方程组(3.8)和初始条件(3.9)。以 A，A^{-1} 分别从左右两侧乘 T 所满足的等式(3.8)和(3.9)的两边，

$$\frac{\partial}{\partial g^{\alpha}}(AT(g)A^{-1}) = \left.\frac{\partial (gh^{-1})^{\rho}}{\partial g^{\alpha}}\right|_{h=g} AI_{\rho}T(g)A_{-1} \equiv S^{\beta}_{\alpha}(g)I'_{\beta}(AT(g)A^{-1}), \forall g \in G$$

$$AT(e)A^{-1} = 1_{V}$$

与 T' 所满足的偏微分方程组和初始条件

$$\frac{\partial}{\partial g^{\alpha}}T'(g) = \left.\frac{\partial (gh^{-1})^{\rho}}{\partial g^{\alpha}}\right|_{h=g} I'_{\rho}T'(g), \forall g \in G$$

$$T'(e) = 1_{V'}$$

比较，可见 $AT(g)A^{-1}, T'(g)$ 都满足一阶偏微分方程组(3.8)和初始条件(3.9)。由解的唯一性得 $AT(g)A^{-1} = T'(g), \forall g \in G$。

命题 2 r 维李群 G 在复内积空间 V 上的表示 T 是酉表示，当且仅当其无穷小算符 $I_{\alpha}(\alpha = 1, \cdots, r)$ 都是反厄米的。

证明 必要性。若 T 是酉表示：$T(g)T^{+}(g) = T^{+}(g)T(g) = 1_{V}, \forall g \in G$，取等式两边对群参数 g^{α} 的偏导数 $\frac{\partial T(g)}{\partial g^{\alpha}}T^{+}(g) + T(g)\left(\frac{\partial T(g)}{\partial g^{\alpha}}\right)^{+} = 0, \forall g \in G$。当 $g = e$ 时，我们得 $I_{\alpha} + I^{+}_{\alpha} = 0$。可知无穷小算符 $I_{\alpha}(\alpha = 1, \cdots, r)$ 都是反厄米的。

充分性。设表示 T 的无穷小算符 $I_{\alpha}(\alpha = 1, \cdots, r)$ 都是反厄米的。以 $T^{+}(g)$ 左乘式(3.8)两边得 $T^{+}(g)\frac{\partial}{\partial g^{\alpha}}T(g) = \left.\frac{\partial (gh^{-1})^{\rho}}{\partial g^{\alpha}}\right|_{h=g} T^{+}(g)I_{\alpha}T(g), \forall g \in G$。取左式与

其厄米共轭式的和得$\frac{\partial}{\partial g^{\alpha}}(T^{+}(g)T(g))=0$(利用了条件 $I_{\alpha}^{+}=-I_{\alpha}$)。又由于 $T(e)=1_V$,可知 $T^{+}(g)T(g)=T^{+}(e)T(e)=1_V$, $T^{+}(g)=T^{-1}(g)$, $\forall g\in G$, T 是幺正表示。

定理 2 设 $A_{\alpha}(\alpha=1,\cdots,r)$是复线性空间 V 上的线性算符,满足条件

$$[A_{\alpha},A_{\beta}]=\sum_{\gamma=1}^{r}A_{\gamma}C_{\alpha\beta}^{\gamma} \tag{3.17}$$

式中 $C_{\alpha\beta}^{\gamma}$是 r 维李群 G 的结构常数。则存在 G 在 V 上的表示 T,以 $A_{\alpha}(\alpha=1,\cdots,r)$为其无穷小算符。

证明 考虑偏微分方程组

$$\frac{\partial}{\partial g^{\alpha}}T(g)=\left.\frac{\partial(gh^{-1})^{\rho}}{\partial g^{\alpha}}\right|_{h=g}A_{\rho}T(g),\forall g\in G,(\alpha=1,\cdots,r) \tag{3.18}$$

的满足初值条件 $T(e)=1_V$ 的解。由偏微分方程理论知,解存在的充分必要条件是,由式(3.18)计算出的混合偏导数$\frac{\partial^2 T(g)}{\partial g^{\alpha}\partial g^{\beta}}$的表达式与顺序无关:

$$\left\{\frac{\partial}{\partial g^{\beta}}\left[\left.\frac{\partial(gh^{-1})^{\tau}}{\partial g^{\alpha}}\right|_{h=g}\right]-\frac{\partial}{\partial g^{\alpha}}\left[\left.\frac{\partial(gh^{-1})^{\tau}}{\partial g^{\beta}}\right|_{h=g}\right]\right\}A_{\tau}+$$

$$\left.\frac{\partial(gh^{-1})^{\rho}}{\partial g^{\alpha}}\right|_{h=g}\left.\frac{\partial(gh^{-1})^{\sigma}}{\partial g^{\beta}}\right|_{h=g}[A_{\rho},A_{\sigma}]=0$$

以 A_{τ} 乘恒等式(3.13),并对 τ 求和,考虑到条件(3.23),我们就能得到上式。于是偏微分方程组(3.18)的满足初值条件 $T(e)=1_V$ 的解存在。显然解是唯一的。我们要证明这个解构成群 G 的表示,即满足条件 $T(gf)=T(g)T(f)$。以 $T(f)$右乘等式(3.18)得$(T(g)T(f))$满足微分方程

$$\frac{\partial}{\partial g^{\alpha}}(T(g)T(f))=\left.\frac{\partial(gh^{-1})^{\rho}}{\partial g^{\alpha}}\right|_{h=g}A_{\rho}(T(g)T(f)),\forall g\in G,\alpha=1,\cdots,r \tag{3.19}$$

易见它还满足初始条件

$$(T(e)T(f))=T(f) \tag{3.20}$$

另一方面,$T(gf)$满足微分方程(令(3.18)中的变量为 gf)

$$\frac{\partial}{\partial(gf)^{\alpha}}T(gf)=\frac{\partial(gfh^{-1})^{\rho}}{\partial(gf)^{\alpha}}\Big|_{h=gf}A_{\rho}T(gf),\forall g\in G,\alpha=1,\cdots,r \tag{3.21}$$

和初始条件

$$T(ef)=T(f) \tag{3.22}$$

以$\frac{\partial(gf)^{\alpha}}{\partial g^{\beta}}$乘式(3.21)并对 α 求和得$\frac{\partial}{\partial g^{\beta}}T(gf)=\frac{\partial(g(hf^{-1})^{-1})^{\rho}}{\partial g^{\beta}}\Big|_{hf^{-1}=g}A_{\rho}T(gf)$。而这就是(3.19)。$T(gf),T(g)T(f)$满足同样的微分方程和初始条件,由解的唯一性可知 $T(gf)=T(g)T(f)$。我们证明了(3.18)的满足初值条件 $T(e)=1_V$ 的解 T 确实构成群 G 在 V 上的一个表示。当(3.18)中的变量 $g=e$ 时,我们得到

$$\frac{\partial T(g)}{\partial g^{\alpha}}\Big|_{g=e}=A_{\alpha} \tag{3.23}$$

可见表示 T 的无穷小算符是$A_{\alpha}(\alpha=1,\cdots,r)$。定理证完。

命题 3 设 T 是 r 维李群 G 在复线性空间 V 上的表示,W 是 V 的子空间,则 W 是 T – 不变子空间当且仅当它是$\{I_{\alpha}\mid\alpha=1,\cdots,r\}$ – 不变子空间。

证明 必要性。取子空间 W 的一个基$\{e_1,\cdots,e_m\}$,并把它扩充成 V 的一个基$\{e_1,\cdots,e_m,e_{m+1},\cdots,e_n\}$。在这个基下,线性算符 $T(g)$的矩阵具有形式$[T(g)]=\begin{bmatrix}T_1(g) & X(g)\\ 0 & T_2(g)\end{bmatrix}$。因而无穷小算符 I_{α} 的矩阵具有形式$[I_{\alpha}]=\frac{\partial}{\partial g^{\alpha}}[T(g)]\Big|_{g=e}=\begin{bmatrix}I_{1\alpha} & J_{\alpha}\\ 0 & I_{2\alpha}\end{bmatrix}$。可见 W 是$\{I_{\alpha}\mid\alpha=1,\cdots,r\}$ – 不变子空间。

充分性。我们将从条件 W 是$\{I_{\alpha}\mid\alpha=1,\cdots,r\}$ – 不变子空间出发,证明对任一 $w\in W$,都有 $T(g)w\in W,\forall g\in G$。把李群表示 T 满足的微分方程组(3.8)和初始条件(3.9)作用于 w

$$\frac{\partial}{\partial g^{\alpha}}(T(g)w)\equiv S_{\alpha}^{\beta}(g)I_{\beta}(T(g)w),\forall g\in G \tag{3.24}$$

$$T(e)w=w \tag{3.25}$$

式(3.24)是复线性空间 V 中的向量 $T(g)w\equiv y(g)$所满足的微分方程。式(3.25)是 $T(g)w\equiv y(g)$满足的初始条件。由微分方程理论知,一阶微分方程(3.24)的解由初始条件(3.25)唯一地确定。以 I'_{β} 记 I_{β} 到不变子空间 W 上的限制

$$I'_{\beta}\equiv I_{\beta}\mid_{W} \tag{3.26}$$

显然$\{I'_\alpha \mid \alpha = 1, \cdots, r\}$和$\{I_\alpha \mid \alpha = 1, \cdots, r\}$满足同样的对易关系

$$[I'_\alpha, I'_\beta] = \sum_{\gamma=1}^{r} I'_\gamma C^\gamma_{\alpha\beta} \tag{3.27}$$

由定理2,存在群 G 在 W 上的表示 T',以$\{I'_\alpha \mid \alpha = 1, \cdots, r\}$为无穷小算符。于是我们有

$$\frac{\partial}{\partial g^\alpha} T'(g) = S^\beta_\alpha(g) I'_\beta T'(g), \forall\, g \in G \tag{3.28}$$

$$T'(e) = 1_W \tag{3.29}$$

把上面两式作用于 w 得

$$\frac{\partial}{\partial g^\alpha}(T'(g)w) \equiv S^\beta_\alpha(g) I'_\beta(T'(g)w), \forall\, g \in G \tag{3.30}$$

$$T'(e)w = w \tag{3.31}$$

由于 $T'(g)$是 W 上的线性算符,故 $T'(g)w \in W$,从而 $I'_\beta(T'(g)w) = I_\beta(T'(g)w)$。式(3.30)可以改写作

$$\frac{\partial}{\partial g^\alpha}(T'(g)w) \equiv S^\beta_\alpha(g) I_\beta(T'(g)w), \forall\, g \in G \tag{3.32}$$

对比微分方程初值问题(3.32) + (3.31)和(3.24) + (3.25),我们得到

$$T(g)w = T'(g)w \in W, \forall\, g \in G \tag{3.33}$$

可见 W 是 T - 不变子空间。

命题3表明,李群表示是不可约表示,当且仅当它的无穷小算符组是不可约算符组。

从上面的讨论可知,对任一给定的 r 维李群 G,若其群元素参数的给法确定了,$C^\gamma_{\alpha\beta}$是相应的结构常数,则 G 的所有表示的集合 $\mathbf{D}$,与所有满足对易关系$[I_\alpha, I_\beta] = I_\gamma C^\gamma_{\alpha\beta}$的 r - 线性算符组的集合 $\mathbf{J}$ 之间,存在双方一一对应关系:对 G 的每一个表示 $T \in \mathbf{D}$,都有一个确定的满足对易关系$[I_\alpha, I_\beta] = I_\gamma C^\gamma_{\alpha\beta}$的 r - 线性算符组$\{I_\alpha \mid \alpha = 1, \cdots, r\} \in \mathbf{J}$是表示 T 的无穷小算符组;反之,对每一个满足对易关系$[I_\alpha, I_\beta] = I_\gamma C^\gamma_{\alpha\beta}$的 r - 线性算符组$\{I_\alpha \mid \alpha = 1, \cdots, r\} \in \mathbf{J}$,都有 G 的一个确定的表示 $T \in \mathbf{D}$,以 r - 线性算符组$\{I_\alpha \mid \alpha = 1, \cdots, r\} \in \mathbf{J}$为其无穷小算符组。$G$ 在空间 V, V' 上的两个表示 T, T' 等价:存在可逆线性映射 $A: V \to V'$,使 $T'(g) = AT(g)A^{-1}, \forall\, g \in G$,当且仅当两个表示 T,

T' 的无穷小算符组等价，而且 $I'_\alpha = AI_\alpha A^{-1}$，$\forall\,\alpha = 1,\cdots,r$。由于 T 的不可约不变子空间是其无穷小算符组的不可约不变子空间，反之亦然，可知 G 的表示 T 可约，当且仅当其无穷小算符组可约。G 在内积空间上的表示 T 是幺正的，当且仅当其无穷小算符都是反厄米的。我们看到，研究李群 G 的表示，可归为研究满足对易关系 $[I_\alpha, I_\beta] = I_\gamma C^\gamma_{\alpha\beta}$ 的 r - 线性算符组 $\{I_\alpha \mid \alpha = 1,\cdots,r\}$。

§4 李群与李代数

Ⅰ. 李代数的概念

让我们考察下面的事例：设 G 是 r 维李群。$T: G \to GL(V)$ 是 G 在线性空间 V 上的表示。以 $EndV$ 记 V 上线性算符的全体所成的线性空间。表示 T 的无穷小算符组 $\left\{I_\alpha = \dfrac{\partial T(g)}{\partial g^\alpha}\Big|_{g=e} \,\Big|\, \alpha = 1,\cdots,r\right\}$ 张成 $EndV$ 的一个子线性空间 L。对所有的 $x, y \in EndV$，令

$$[x, y] = xy - yx \tag{4.1}$$

则 $[\quad,\quad]: EndV \times EndV \to EndV$ 是 $EndV$ 上的一个双线性的二元合成（双线性的代数运算），称之为李乘法。显然李乘法满足下面的条件：对所有 $x, y, z \in EndV$

$$(\text{ⅰ})\ [x, y] = -[y, x]\ (\text{或}\ [x, x] = 0)\ (\text{斜对称性}) \tag{4.2}$$

$$(\text{ⅱ})\ [[x, y], z] + [[y, z], x] + [[z, x], y] = 0\ (Jacobi\ \text{恒等式}) \tag{4.3}$$

由上一节李群表示理论的定理 1

$$[I_\alpha, I_\beta] = \sum_{\gamma=1}^{r} I_\gamma C^\gamma_{\alpha\beta} \tag{4.4}$$

可知 L 对 $EndV$ 上的二元合成 $[,]$ 是封闭的；故可以把 $[,]$ 看作是 L 上的代数运算。这样的 L 就是一个李代数。当 $\dim L = r$ 时，L 称为李群 G 的李代数。

在三维实欧式空间 $\mathbf{R}^3$ 上，两个向量 $\boldsymbol{x}, \boldsymbol{y}$ 的矢量积 $\boldsymbol{x} \times \boldsymbol{y}$ 是 $\mathbf{R}^3$ 上的一个双线性代数运算。读者容易验证，它是斜对称的，而且满足 Jacobi 恒等式。

现在我们给出李代数的一般定义。

定义 设 L 是 $\mathbf{F}$ – 线性空间($\mathbf{F}$ 是复数域 $\mathbf{C}$ 或实数域 $\mathbf{R}$)。若在 L 上定义了一个双线性的二元合成(双线性的代数运算)$[,]: L \times L \to L$ 满足下面的条件:对所有 $x, y, z \in L$

$$(\mathrm{i}) -[x,y]=[y,x](\text{或}[x,x]=0)\ (\text{斜对称性}) \tag{4.5}$$

$$(\mathrm{ii})[[x,y],z]+[[y,z],x]+[[z,x],y]=0(\textit{Jacobi}\ \text{恒等式}) \tag{4.6}$$

则 L 称是一个 $\mathbf{F}$ – 李代数,称代数运算$[,]$为李乘法。

从上述定义我们看到,李代数首先是一个线性空间,在这个线性空间上定义了一个双线性二元合成具有斜对称性并满足 Jacobi 恒等式,则线性空间就成了一个李代数。李代数的子李代数是这样定义的:设 M 是李代数 L 的非空子集,若 M 是线性空间 L 的子线性空间,而且对 L 的李乘法$[,]$是封闭的,则称 M 是 L 的子李代数。注意,并非李代数 L 的任一非空子集,对李乘法$[,]$是封闭的,就可以称为 L 的子李代数。

练习 4.1 试证:r 维李群 G 的表示 T 的 r 个无穷小算符$\{I_1, \cdots, I_r\}$线性无关时,T 是 G 的忠实表示。

Ⅱ. 矩阵李群的李代数

设 G 是 r 维矩阵李群。令 G 的每个元($N \times N$ 矩阵)g 与自身对应:$T(g)=g$,我们得到群 G 的一个忠实的矩阵表示 $T: G \to G$。为确保 r 个参数$(g^1, \cdots, g^r)$的不同值对应于 G 的不同元,总是这样选参数,使当 $\mathrm{d}T(g)=\sum_{\alpha=1}^{r}\frac{\partial T(g)}{\partial g^{\alpha}}\mathrm{d}g^{\alpha}=0$ 时,必有 $\mathrm{d}g^{\alpha}=0(\alpha=1,\cdots,r)$。

可见无穷小算符组

$$\left\{I_{\alpha}=\left.\frac{\partial T(g)}{\partial g^{\alpha}}\right|_{g=e}\middle|\alpha=1,\cdots,r\right\} \tag{4.7}$$

是线性独立的。它们张成的 r 维线性空间 L 在李乘法$[x,y]=xy-yx$ 下构成一个李代数。称为矩阵李群 G 的李代数。

$N \times N$ 矩阵李群 G 的李代数还可以这样给出。考虑群流形上过单位元 $e \in G$ 的曲线

$$f:(a,b) \to G,(a<0,b<0) \tag{4.8}$$

对 $t \in (a,b)$,$f(t) \in G$ 是 $N \times N$ 矩阵。我们总是这样选参数:使 $f(0)=e \in G$。$f'(0)$

是 $N\times N$ 矩阵。以 L 记所有这种群流形上过单位元的光滑曲线 f 在 $t=0$ 点的微商 $f'(0)$ 所成的集合。我们首先来证明 L 是由某些 $N\times N$ 矩阵构成的线性空间，即证明 L 在线性运算下是封闭的。

(ⅰ)若 $x\in L, a\in\mathbf{R}$。这表明存在群流形上过单位元 e 的曲线 $f(t)$，使 $f'(0)=x$。易见 $h(t)\equiv f(at)$ 也是群流形上过单位元 e 的曲线。

$$\therefore\quad h'(0)=af'(0)=ax\in L \tag{4.9}$$

(ⅱ)若 $x,y\in L$。这表明存在群流形上过单位元 e 的曲线 $f(t),g(t)$，使 $x=f'(0), y=g'(0)$。易见 $h(t)\equiv f(t)g(t)$ 也是群流形上过单位元 e 的单参数曲线。

$$\therefore\quad h'(0)=f'(0)g(0)+f(0)g'(0)=x+y\in L \tag{4.10}$$

现在进一步证明 L 在李乘法[,]下是封闭的。

(ⅲ)若 $x,y\in L$。这表明存在群流形上过单位元 e 的单参数曲线 $f(t),g(t)$，使 $x=f'(0), y=g'(0)$。易见 $f(s)g(t)f(s)^{-1}$ 对固定的 s，是以 t 为参数的群流形上过单位元 e 的曲线。

$$\therefore\quad f(s)g'(0)f(s)^{-1}=f(s)yf(s)^{-1}\in L,\ \forall\, s\in(a,b) \tag{4.11}$$

$f(s)ys(s)^{-1}$ 是线性空间 L 中的以 s 为参数的曲线。它对参数 s 的微商仍是 L 中的向量($N\times N$ 矩阵)。

$$f'(0)yf(0)^{-1}+f(0)y(-f'(0))=xy-yx=[x,y]\in L \tag{4.12}$$

可见 L 是一个李代数。当 G 是 r 维矩阵李群，L 是 r 维矩阵李代数，称之为 G 的李代数。

Ⅲ. 常见的矩阵李群与它们的李代数

在下面的讨论中，$N\times N$ 矩阵 X 的指数函数(见 §0.4)

$$Y=\exp X\equiv e^X\equiv\sum_{n\geq 0}\frac{X^n}{n!} \tag{4.13}$$

将扮演一个重要的角色。这是定义在整个 $\mathbf{M}_N(\mathbf{F})$ 上的映射。当 $X=0$ 时 $Y=E$($N\times N$ 单位矩阵)，当 X 是对角阵 $X=diag(x_{11},\cdots,x_{nn})$ 时，Y 也是对角矩阵 $Y=diag(e^{x_{11}},\cdots,e^{x_{nn}})$。当 $X=i2\pi nE$(n 是整数)时 $Y=E$，可见 $\exp:\mathbf{M}_N(\mathbf{F})\to\mathbf{M}_N(\mathbf{F})$ 不是单射。但是把函数的定义域限制在零矩阵的某些邻域 $\mathbf{U}$ 时，$\exp:\mathbf{U}\to\exp\mathbf{U}$ 可以是双射，$\exp\mathbf{U}$ 是 E 的一个邻域。

在 §0.4 中我们还证明过

$$\det Y = \det e^X = e^{trX} \tag{4.14}$$

让我们考察下面的矩阵李群和它们的李代数。

(ⅰ) $GL(N,\mathbf{F})$是全体非异的 $N\times N$ $\mathbf{F}$ - 矩阵所成的群:

$$GL(N,\mathbf{F}) = \{X \in \mathbf{M}_N(\mathbf{F}) \mid \det X \neq 0\} \tag{4.15}$$

它的李代数 $gl(N,\mathbf{F})$是全体 $N\times N$ $\mathbf{F}$ - 矩阵构成的。

以 $E_\alpha^{\ \beta}$ 记 α 行 β 列阵元为 1,其它阵元为零的 $N\times N$ 矩阵。$\{E_\alpha^{\ \beta} \mid 1\leq \alpha,\beta\leq N\}$构成 $gl(N,\mathbf{F})$的一个基。$gl(N,\mathbf{F})$的每一元 ξ 可以唯一地写作

$$\xi = \sum_{\alpha,\beta} \xi^\alpha_{\ \beta} E_\alpha^{\ \beta} (\xi^\alpha_{\ \beta} \in \mathbf{F}) \tag{4.16}$$

当 $\mathbf{F}=\mathbf{R}$ 时,$GL(N,\mathbf{R})$是 N^2 维李群。在单位元 $E\in GL(N,\mathbf{R})$的某个邻域中的每个群元 X 都可以唯一地写作

$$X = \exp\Big(\sum_{\alpha,\beta} \xi^\alpha_{\ \beta} E_\alpha^{\ \beta}\Big) \tag{4.17}$$

可以把$\{\xi^\alpha_{\ \beta} \mid 1\leq \alpha,\beta\leq N\}$看成 $GL(N,\mathbf{R})$群的局部坐标(群参数)。相应的无穷小算符是 $E_\alpha^{\ \beta}$。当 $\mathbf{F}=\mathbf{C}$ 时,$gl(N,\mathbf{C})$的每一元 ξ 可以唯一地写作

$$\xi = \sum_{\alpha,\beta} (\eta^\alpha_{\ \beta} + i\xi^\alpha_{\ \beta}) E_\alpha^{\ \beta} \tag{4.18}$$

$GL(N,\mathbf{C})$是 $2N^2$ - 维李群。在单位元 $E\in GL(N,\mathbf{C})$的某个邻域中的每个群元 X 都可以唯一地写作

$$X = \exp\sum_{\alpha,\beta} (\eta^\alpha_{\ \beta} E_\alpha^{\ \beta} + \xi^\alpha_{\ \beta}(iE_\alpha^{\ \beta})) \tag{4.19}$$

$\{\eta^\alpha_{\ \beta}, \zeta^\alpha_{\ \beta} \mid 1\leq \alpha,\beta\leq N\}$可以看作是群 $GL(N,\mathbf{C})$的局部坐标(群参数)。相应的无穷小算符是 $E_\alpha^{\ \beta}$ 和 $iE_\alpha^{\ \beta}$。(注意,$gl(N,\mathbf{C})$是 N^2 维复线性空间,可以看成 $2N^2$ 维实线性空间)

(ⅱ) $SL(N,\mathbf{F})$是全体行列式为 1 的 $N\times N$ $\mathbf{F}$ - 矩阵所成的群:

$$SL(N,\mathbf{F}) = \{X \in \mathbf{M}_N(\mathbf{F}) \mid \det X = 1\} \tag{4.20}$$

在单位元 $E\in SL(N,\mathbf{F})$的某个邻域中,每个群元 X 都可以唯一地表作 $X = e^\xi$。注意到(4.14),可知

$$\det X = 1 \Leftrightarrow tr\xi = 0 \tag{4.21}$$

对任一 $N\times N$ 迹零 $\mathbf{F}$ - 矩阵 ξ: $tr\xi = 0$,易知 $f(t) = e^{\xi t}$ 是 $SL(N,\mathbf{F})$群流形上过单位元

的曲线。$\therefore f'(0)=\xi\in sL(N,\mathbf{F})$。这表明任一迹零 $\mathbf{F}$ – 矩阵属于 $sl(N,\mathbf{F})$。

反过来,考虑 $SL(N,\mathbf{F})$群流形上的任一过单位元的曲线

$$g(t)=e^{\eta(t)}\in SL(N,\mathbf{F}),tr\eta(t)\equiv 0 \tag{4.22}$$

应有

$$g'(0)=\eta'(0)\in sl(N,\mathbf{F}) \tag{4.23}$$

由(4.22)可知,$SL(N,\mathbf{F})$群的李代数 $sl(N,\mathbf{F})$的每个元 $g'(0)$都是迹零的:

$$trg'(0)=tr\eta'(0)=0 \tag{4.24}$$

这样 $sl(N,\mathbf{F})$是由全体迹零的 $N\times N\mathbf{F}$ – 矩阵构成的。

$$\{E_\alpha^\beta\mid\alpha\neq\beta\}\cup\{E_\gamma^\gamma-E_{\gamma+1}{}^{\gamma+1}\equiv H_\gamma\mid 1\leq\gamma\leq N\} \tag{4.25}$$

构成 $sl(N,\mathbf{F})$的一个基。任一迹零的 $N\times N\mathbf{F}$ – 矩阵都可以唯一地写作

$$\xi=\sum_{\alpha\neq\beta}\xi^\alpha{}_\beta E_\alpha^\beta+\sum_{\gamma=1}^{N-1}\xi^\gamma H_\gamma \tag{4.26}$$

在单位元 $E\in SL(N,\mathbf{F})$的某个邻域中的每个群元 X 都可唯一地写作

$$X=\exp\Big(\sum_{\alpha\neq\beta}\xi^\alpha{}_\beta E_\alpha^\beta+\sum_{\gamma=1}^{N-1}\xi^\gamma H_\gamma\Big) \tag{4.27}$$

当 $\mathbf{F}=\mathbf{R}$ 时,$\{\xi^\alpha{}_\beta,\xi^\gamma\}$可以看作是 $SL(N,\mathbf{R})$群的局部坐标(群参数)。

当 $\mathbf{F}=\mathbf{C}$ 时,$\{\mathrm{Re}\xi^\alpha{}_\beta,\mathrm{Im}\xi^\alpha{}_\beta,\mathrm{Re}\xi^\gamma,\mathrm{Im}\xi^\gamma\}$可以看作是 $SL(N,\mathbf{C})$群的局部坐标(群参数)。

(ⅲ)设 S 是 $N\times N$ 非异矩阵,考察所有满足条件

$$X^TSX=S \tag{4.28}$$

的 $N\times N$ 矩阵的集合 G。单位矩阵 E 显然满足条件(4.28)。取(4.28)两边的行列式得$(\det X)^2=1$。这表明 G 的元的行列式不外乎 ±1,肯定是非异的。以$(X^T)^{-1}=(X^{-1})^T$ 左乘,以 X^{-1}右乘(4.28)式两边得$\cdot S=(X^{-1})^TSX^{-1}$,这表明若 $X\in G$,则 $X^{-1}\in G$。容易看到若 $X,Y\in G$,则 $XY\in G$。因而 G 是一个矩阵乘法群。设 $f(t)$是群 G 的群流形上过单位元的曲线:$f(t)\in G,f(0)=E$。由

$$f(t)^TSf(t)=S \tag{4.29}$$

可得

$$f'(0)^TS+Sf'(0)=0 \tag{4.30}$$

这说明矩阵李群 G 的李代数 g 的元 $f'(0)$满足条件(4.30)。反之设 $N\times N$ 矩阵 ξ 满

足条件

$$\xi^T S + S\xi = 0(\text{此即}(\mathscr{L}m\xi^T + \mathscr{R}m\xi)S = 0) \tag{4.31}$$

考察 $\mathbf{M}_N(\mathbf{F})$上的曲线

$$g(t) = \exp(\xi t) = e^{\xi t} \tag{4.32}$$

我们有(参见 §0.7)

$$g(t)^T S g(t) = e^{\xi^T t} S e^{\xi t} = e^{t(\mathscr{L}m\xi^T + \mathscr{R}m\xi)} S = S \tag{4.33}$$

可见 $g(t)$是群 G 的群流形上过单位元的曲线。于是

$$g'(0) = \xi \in g \tag{4.34}$$

这说明满足条件(4.31)的 $N \times N$ 矩阵一定是李代数 g 的元。于是由条件(4.28)定义的矩阵李群 G 的李代数 g 是全体满足条件(4.31)的 $N \times N$ 矩阵 ξ 所构成的。选定 g 的一个基$\{e^\alpha \mid 1 \leq \alpha \leq r\}$,则群 G 单位元的某个邻域中的每个群元 X 可以唯一地表作

$$X = e^{\sum_{\alpha=1}^{r} x^\alpha e_\alpha} \tag{4.35}$$

$\{x^\alpha \mid 1 \leq \alpha \leq r\}$可以看作是群 G 的局部坐标(群参数)。

考虑 $\mathbf{F} = \mathbf{R}$,实矩阵的情况。当

$$S = E \tag{4.36}$$

这时由条件(4.28)确定的群是 N 维实正交群 $O(N)$,其李代数 $o(N)$是由全体 $N \times N$ 反对称实矩阵构成。当

$$S = diag(\underbrace{-1, \cdots, -1}_{m \text{ copies}}, \underbrace{1, \cdots, 1}_{n \text{ copies}}) \quad (m + n = N) \tag{4.37}$$

这时由条件(4.28)确定的群是伪正交群 $O(m, n)$。当

$$S = \begin{pmatrix} 0 & E_l \\ -E_l & 0 \end{pmatrix} \tag{4.38}$$

时,条件(4.28)确定的群是辛群 $SP(2l, \mathbf{R})$。

类似地,我们有复正交群,复辛群。此外不再赘述。

(ⅳ)设 S 是 $N \times N$ 非异复矩阵,考察所有满足条件

$$X^+ SX = S \tag{4.39}$$

的 $N \times N$ 复阵所成的集合 G。单位阵 E 显然属于 G。取式(4.39)两边行列式得 $|\det X|^2 = 1$,这表明 G 的元肯定是非异的。由(4.39)可得 $S = (X^{-1})^+ SX^{-1}$,这表明

若 $X\in G$ 则 $X^{-1}\in G$。易见若 $X,Y\in G$,则 $XY\in G$。于是 G 是一个矩阵乘法群。设 $f(t)$是 G 的群流形上过单位元的曲线:$f(t)\in G, f(0)=E$。由

$$f(t)^{+}\, Sf(t) = S \tag{4.40}$$

可得

$$f'(0)^{+}\, S + Sf'(0) = 0 \tag{4.41}$$

这说明矩阵李群 G 的李代数 g 的元 $f'(0)$满足条件(4.41)。反之设 $N\times N$ 复矩阵满足条件

$$\xi^{+}\, S + S\xi = 0 \text{ (此即}(\mathscr{I}m\xi^{+} + \mathscr{R}m\xi)S = 0) \tag{4.42}$$

考察 $\mathbf{M}_N(\mathbf{C})$上的曲线

$$g(t) = \exp(\xi t) = e^{\xi t} \tag{4.43}$$

我们有

$$g(t)^{+}\, Sg(t) = e^{\xi^{+}t} S e^{\xi t} = e^{t(\mathscr{I}m\xi^{+} + \mathscr{R}m\xi)} S = S \tag{4.44}$$

可见 $g(t)$是群 G 的群流形上过单位元的曲线。于是

$$g'(0) = \xi \in g \tag{4.45}$$

这说明满足条件(4.42)的 $N\times N$ 复阵 ξ 一定是 g 的元。把 g 看成实线性空间,选定 g 的一个基$\{e_1,\cdots,e_r\}$则群 G 的单位元的某个邻域中的每个群元 X,都可唯一地表作

$$X = e^{\sum_{\alpha=1}^{r} x^{\alpha} e_{\alpha}} \qquad (x^{\alpha}\in \mathbf{R}) \tag{4.46}$$

$\{x^{\alpha}\mid 1\leq\alpha\leq r\}$可以看成群 G 的局部坐标(群参数)。

当 $S=E$ 时,由条件(4.39)确定的群 G 是 N 维酉群 $U(N)$,其李代数 $u(N)$是由全体 $N\times N$ 反厄米矩阵所构成。当

$$S = diag(\underbrace{-1,\cdots,-1}_{m\ copies},\underbrace{1,\cdots,1}_{n\ copies}) \quad (m+n=N) \tag{4.47}$$

对应的群 G 是伪酉群 $U(m,n)$群。

练习 4-2 讨论李代数 $o(m,n)$, $SP(2l,\mathbf{R})$和 $u(m,n)$。

§5 $SO(3)$群与 $SU(2)$群

2阶特殊酉矩阵群 $SU(2)=\{U\in GL(2,C)\mid \det U=1 \& U^{+}=U^{-1}\}$是全体行列式为1的$2\times 2$酉矩阵所构成的群，它是全体行列式不为0的$2\times 2$复矩阵所构成的群$GL(2,C)$(读作2阶一般复线性群)的子群。行列式为1的$2\times 2$酉矩阵的一般形式为$\begin{bmatrix}\alpha & \beta\\ -\bar{\beta} & \bar{\alpha}\end{bmatrix}=\begin{bmatrix}a+ib & c+id\\ -c+id & a-ib\end{bmatrix}$，其中$|\alpha|^2+|\beta|^2=a^2+b^2+c^2+d^2=1$。(证明留作练习) 可见$SU(2)$群的群流形是四维实欧氏空间$\mathbf{R}^4$中的三维单位超球面，是一单连通流形。

旋转群$SO(3)$(亦读作3阶特殊实正交矩阵群)是由全体行列式为1的3阶实正交矩阵所构成。我们现在将指出，存在一个由$SU(2)$群到$SO(3)$群的满同态。考虑全体2×2迹零厄米矩阵所成的实线性空间V。Pauli矩阵$\sigma_1,\sigma_2,\sigma_3$是$V$的一个基。全体$2\times 2$迹零厄米矩阵就是$\sigma_1,\sigma_2,\sigma_3$的全体实系数线性组合。因而在三维实线性空间$\mathbf{R}^3$与$V$之间有一个自然的同构：

$$\boldsymbol{x}\mapsto \boldsymbol{x}\cdot\boldsymbol{\sigma}\equiv x_\alpha\sigma_\alpha=\begin{bmatrix}x_3 & x_1-ix_2\\ x_1+ix_2 & -x_3\end{bmatrix}$$

对每一$U\in SU(2)$，定义V上的线性算符$R(U)$：$R(U)v=UvU^{+}$，$\forall v\in V$(注意，相似变换不改变方阵的阵迹，酉相似变换不改变方阵的厄米性)。对应$U\mapsto R(U)$，$\forall U\in SU(2)$给出了$SU(2)$群在V上的一个表示$R: SU(2)\to GL(V)$。V上线性算符$R(U)$在基$\sigma_1,\sigma_2,\sigma_3$下的矩阵$[R(U)]$由下式确定：$R(U)\sigma_\alpha=U\sigma_\alpha U^{+}=\sigma_\beta R_{\beta\alpha}(U)$，$R_{\beta\alpha}(U)$为实数，为研究$SU(2)$群的三维实矩阵表示$[R(U)]$的性质，考察$U\begin{bmatrix}x_3 & x_1-ix_2\\ x_1+ix_2 & -x_3\end{bmatrix}U^{+}=Ux_\alpha\sigma_\alpha U^{+}=\sigma_\beta R_{\beta\alpha}(U)x_\alpha\equiv\sigma_\beta y_\beta=\begin{bmatrix}y_3 & y_1-iy_2\\ y_1+iy_2 & -y_3\end{bmatrix}$，式中$y_\beta=R_{\beta\alpha}(U)x_\alpha$。由相似变换不改变方阵的行列式，我们得到$x_1^2+x_2^2+x_3^2=y_1^2+y_2^2+y_3^2$，即$[R(U)]$是实正交矩阵：$[R(U)]^T[R(U)]=E_3$，$E_n$为单位$n\times n$矩阵。从而$\det[R(U)]=\pm 1$。显然对$SU(2)$群的单位元$E_2\in SU(2)$有$[R(E_2)]=E_3$，从而det

$[R(E_2)]=1$。$\det[R(U)]$应连续地依赖于 $U\in SU(2)$。由于可以在 $SU(2)$群的群流形上画一条连接恒等元 $E_2\in SU(2)$与任一元 $U\in SU(2)$的曲线。所以 $\det[R(U)]=1,[R(U)]\in SO(3),\forall U\in SU(2)$。我们证明了 $SU(2)$群的矩阵表示 $U\mapsto[R(U)]$是 $SU(2)$群到 $SO(3)$群的同态。下面来证明它是一个满同态。设$[R]\in SO(3)$是一行列式为 1 的实正交矩阵。对任一实向量 $\boldsymbol{x}\in R^3$,令 $y_\beta=R_{\beta\alpha}x_\alpha$。考察 V 中的两个向量 $x_\alpha\sigma_\alpha,y_\beta\sigma_\beta$,它们是 2×2 迹零厄米矩阵。久期方程分别为 $\lambda^2+\det(x_\alpha\sigma_\alpha)=\lambda-x_\alpha\sigma_\alpha=0$ 和 $\lambda^2+\det(y_\alpha\sigma_\alpha)=\lambda^2-y_\alpha\sigma_\alpha=0$。(证明留作练习)由于$[R]$是实正交矩阵,可知 $x_\alpha x_\alpha=y_\beta y_\beta$,从而 2×2 厄米矩阵 $x_\alpha\sigma_\alpha,y_\beta\sigma_\beta$ 的特征值都是 $\lambda=\pm\sqrt{x_\alpha x_\alpha}=\pm\sqrt{y_\beta y_\beta}$,它们在各自的表象下矩阵相同。故存在 2×2 酉矩阵 $W\in U(2)$使 $y_\beta\sigma_\beta=W(x_\alpha\sigma_\alpha)W^+$。酉矩阵行列式的模为 1:$\det W=e^{i\theta}$,令 $U=e^{-i\frac{\theta}{2}}W$,则 $U\in SU(2)$,且 $R_{\beta\alpha}x_\alpha\sigma_\alpha=y_\beta\sigma_\beta=W(x_\alpha\sigma_\alpha)W^+=U(x_\alpha\sigma_\alpha)U^+$。可见,对任一行列式为 1 的 3×3 实正交矩阵$[R]\in SO(3)$,都有 $U\in SU(2)$,使$[R]=[R(U)]$。$U\mapsto R(U)$是 $SU(2)$群到 $SO(3)$群的满同态。让我们考察这个同态的核。设 $R(U)=E_3$,此即 $U\sigma_\alpha U^+=\sigma_\alpha,\forall\alpha=1,2,3$,或者 $U\sigma_\alpha=\sigma_\alpha U,\forall\alpha=1,2,3$。考虑到 $\sigma_1,\sigma_2,\sigma_3$ 是不可约线性算符组,由 Schur 引理,上面条件 U 把确定为 $U=\gamma E_2$。条件行列式为 1:$\det U=\gamma^2=1$,进一步把 U 的可能范围缩小为 $U=\pm E_2$。直接验证可知$(\pm E_2)\sigma_\alpha(\pm E_2)^+=\sigma_\alpha,\forall\alpha=1,2,3$。故同态的核为$\{E_2,-E_2\}$,它是 $SU(2)$群的 2 阶正规子群。$SU(2)$群的两个元 A,B 映到 $SO(3)$群同一元的充要条件是它们属于关于正规子群$\{E_2,-E_2\}$的同一陪集:$R(A)=R(B)\Leftrightarrow A=\pm B$。同态 $U\mapsto R(U)$是 2 对 1 的满同态。由此可见 $SO(3)$群的群流形是 $\mathbf{R}^4$ 中的三维单位超球面,只是要把每一对对径点视为同一点。

§6 Lorentz 群

Ⅰ.狭义相对论的时空变换

在狭义相对论中,当我们谈到某一“事件”,指的是该事件发生的时间和地点。这是一个不依赖具体参考坐标系的概念。同一客观事件在不同惯性参考坐标系Σ,Σ'

中的时空坐标

$$\begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} \equiv \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \equiv X \text{ 和} \begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} \equiv \begin{bmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{bmatrix} \equiv X' \tag{6.1}$$

应由一个可逆的非齐次线性变换相联系(这是时间空间均匀性的体现)

$$x'^{\mu} = L^{\mu}_{\ v}X^v + a^{\mu} \quad \text{或} \quad X' = LX + A \tag{6.2}$$

其中 L 是 4×4 非奇异实矩阵;而 A 是 4×1 实矩阵,代表坐标原点的平移(Σ的原点在Σ'中的时空坐标)。以 $\Delta X(\Delta X')$记两个客观事件在惯性参考坐标系$\Sigma(\Sigma')$中的时空坐标差,则由(6.2)可得

$$\Delta x'^{\mu} = L^{\mu}_{\ v}\Delta x^v \quad \text{或} \quad \Delta X' = L\Delta X \tag{6.3}$$

光速不变性(在所有惯性系中光在真空中的传播速度都是 c)要求

$$(\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 = 0 \Leftrightarrow (\Delta x'^0)^2 - (\Delta x'^1)^2 - (\Delta x'^2)^2 - (\Delta x'^3)^2 = 0 \tag{6.4}$$

最后一个等式可以写作

$$(L^0_{\ \mu}L^0_{\ v} - L^1_{\ \mu}L^1_{\ v} - L^2_{\ \mu}L^2_{\ v} - L^3_{\ \mu}L^3_{\ v})\Delta x^{\mu}\Delta x^v = 0 \tag{6.5}$$

可见(6.4)中的两个等式的左端都是四维实向量 Δx 的实对称二次型。由代数学知道,两个实对称二次型 $A_{\mu v}\xi^{\mu}\xi^v$ 和 $B_{\mu v}\xi^{\mu}\xi^v$,若总是同时取零值 $A_{\mu v}\xi^{\mu}\xi^v = 0 \Leftrightarrow B_{\mu v}\xi^{\mu}\xi^v = 0$,则必有非零实数 α 使 $A_{\mu v} = \alpha B_{\mu v}$。于是必有 $\alpha = \alpha^* \neq 0$ 使对任意 ΔX 都有

$$(\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 = \alpha[(\Delta x'^0)^2 - (\Delta x'^1)^2 - (\Delta x'^2)^2 - (\Delta x'^3)^2] \tag{6.6}$$

考虑到所有惯性参考坐标系应该是平等的,α 只能是 ±1。进一步考虑到:(ⅰ)对时间反演、空间反演,上式中的 α 等于 1;(ⅱ)若由惯性参考坐标系Σ可连续地变到惯性参考坐标系Σ',则应有 $\alpha = 1$。我们可以确信,对任意两个惯性参考坐标系Σ和Σ',式(6.6)中的 $\alpha = 1$。

$$(\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 = (\Delta x'^0)^2 - (\Delta x'^1)^2 - (\Delta x'^2)^2 - (\Delta x'^3)^2 \tag{6.7}$$

这就是狭义相对性原理对事件的时空坐标变换(6.2)的要求。现在让我们来讨论等

式(6.7)对线性变换 L 的限定。为此引入记号

$$G = [g_{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, G^{-1} = [g^{\mu\nu}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{6.8}$$

利用 G,式(6.7)可以写作

$$g_{\mu\nu}\Delta x^{\mu}\Delta x^{\nu} = g_{\mu\nu}\Delta x'^{\mu}\Delta x'^{\nu} = g_{\rho\sigma}L^{\rho}{}_{\mu}L^{\sigma}{}_{\nu}\Delta x^{\mu}\Delta x^{\nu} \tag{6.9}$$

或者写成矩阵形式

$$(\Delta X)^{T}G\Delta X = (\Delta X)^{T}L^{T}GL\Delta X \tag{6.10}$$

上面的实对称二次型对所有的 ΔX 的取值都相等,于是我们得到

$$L^{T}GL = G \quad \text{或者} \quad L^{\mu}{}_{\rho}g_{\mu\nu}L^{\nu}{}_{\sigma} = g_{\rho\sigma} \quad \forall \rho,\sigma = 0,1,2,3 \tag{6.11}$$

式(6.2)和(6.11)给出了满足狭义相对性原理的时空变换。(6.11)和(6.7)等价。

在以上的讨论中我们采用的是“被动观点”,即把式(6.2)看作同一客观事件在不同惯性参考坐标系中的时空坐标变换关系。当然我们也完全可以采用“主动观点”,把式(6.2)理解为在某一选定的惯性参考坐标系中观察到的四维时空的一个变换。满足条件(6.11)或(6.7)的四维时空变换(6.2)称为 Poincare 变换。齐次的 Poincare 变换 $x'^{\mu} = L^{\mu}{}_{\nu}x^{\nu}$ 或 $X' = LX$ 称作 Lorentz 变换。由于非齐次项代表坐标原点的平移,常常是平庸的,我们着重讨论 Lorentz 变换。

Ⅱ. Lorentz 群 $O(3,1)$

在上一小节所作的物理考虑的基础上,我们可以抽象出如下的数学定义:称 4×4 实矩阵 L 为 Lorentz 变换,若 $L^{T}GL = G$ 或 $L^{\mu}{}_{\rho}g_{\mu\nu}L^{\nu}{}_{\sigma} = g_{\rho\sigma}\ \forall \rho,\sigma = 0,1,2,3$ 四维时空的恒等变换 $E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$、空间反演 $P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$、时间反演 $T =$

$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = -P$ 和时空反演 $TP = PT = -E = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ 显然都是 Lorentz 变换。把全体 Lorentz 变换的集合记作

$$O(3,1) = \{L \in M_4(R) \mid L^T G L = G\} \tag{6.12}$$

其中 $M_4(R)$代表全体 4×4 实矩阵的集合。我们来证明非空集合 $O(3,1)$关于矩阵乘法是封闭的。事实上若 $K, L \in O(3,1)$,则$(LK)^T G(LK) = K^T(L^T GL)K = K^T GK = G, \therefore LK \in O(3,1)$。可见 4×4 矩阵的乘法在 $O(3,1)$上诱导出一个代数运算(不妨称之为 Lorentz 变换的乘法),它是结合的,E 是它的单位元。若 $L \in O(3,1)$,由 $L^T GL = G$ 知 L 是非奇异的,以$(L^{-1})^T = (L^T)^{-1}$左乘,以 L^{-1}右乘上式得$(L^{-1})^T GL^{-1} = G$,于是 $L^{-1} \in O(3,1)$。综上所述,$O(3,1)$在矩阵乘法下是一个群,称之为 Lorentz 群。

4×4 实矩阵有 16 个实阵元。条件(6.11)中的 16 个等式关于下标 ρ, σ 有交换对称性,故其中独立等式的个数为 $16 - C_4^2 = 10$,刻画 Lorentz 变换的独立的实参数的个数为 $16 - 10 = 6$。故 $O(3,1)$是 6 维李群。

由(6.11)可得

$$LG^{-1}L^T = G^{-1} \text{ 或者 } L^\mu{}_\rho g^{\rho\sigma} L^\nu{}_\sigma = g^{\mu\nu}, \forall \mu, \nu = 0,1,2,3 \tag{6.13}$$

在证明 Klein-Gordon 方程具有相对论不变性时将用到式(6.13)。由式(6.11)可得

$$(\det L)^2 = 1, \det L = \pm 1 \tag{6.14}$$

令式(6.11)中的 $\rho = \sigma = 0$,我们得到$(L^0{}_0)^2 - (L^1{}_0)^2 - (L^2{}_0)^2 - (L^3{}_0)^2 = 1$

$$L^0{}_0 = \pm\sqrt{1 + (L^1{}_0)^2 + (L^2{}_0)^2 + (L^3{}_0)^2} \tag{6.15}$$

可见 L_0^0 或者大于等于 1,或者小于等于 -1。令式(6.13)中的 $\mu = \nu = 0$,我们得到

$$L^0{}_0 = \pm\sqrt{1 + (L^0{}_1)^2 + (L^0{}_2)^2 + (L^0{}_3)^2} \tag{6.16}$$

按照 $\det L = \pm 1$ 和 $L^0{}_0 \geq 1$ 或者 $L^0{}_0 \leq -1$ 可以把 Lorentz 变换分成四类。它们构成群 $O(3,1)$的四个连通分支。

(ⅰ)使 $\det L = 1, L^0{}_0 \geq 1$ 的 Lorentz 变换(正常的正时 Lorentz 变换)。四维恒等变换 E 属于这一分支。

(ⅱ)使 $\det L = -1, L^0{}_0 \geq 1$ 的 Lorentz 变换(非正常的正时 Lorentz 变换)。空间反

演 P 属于这一分支。

(ⅲ)使 $\det L=-1, L^0{}_0\leq -1$ 的 Lorentz 变换(非正常的非正时 Lorentz 变换)。时间反演 T 属于这一分支。

(ⅳ)使 $\det L=1, L^0{}_0\leq -1$ 的 Lorentz 变换(正常的非正时 Lorentz 变换)。时空反演 $-E$ 属于这一分支。

让我们考察一下 Lorentz 变换的一些性质。由 $\det(LK)=\det L\det K$ 可知,两个正常(非正常)Lorentz 变换的乘积是正常的 Lorentz 变换。一个正常一个非正常 Lorentz 变换的乘积是非正常的 Lorentz 变换。设 $K, L\in O(3,1)$,则 $(LK)^0{}_0=L^0{}_0K^0{}_0+\sum_{i=1}^{3}L^0{}_iK^i{}_0\equiv L^0{}_0K^0{}_0+\boldsymbol{L}^0\cdot\boldsymbol{K}_0$。我们指出$(LK)^0{}_0$与 $L^0{}_0K^0{}_0$同号。为此只须证明 $|L^0{}_0K^0{}_0|>|\boldsymbol{L}^0\cdot\boldsymbol{K}_0|$。由式(6.15)和(6.16)知 $|L^0{}_0K^0{}_0|=\sqrt{1+\boldsymbol{L}^0\cdot\boldsymbol{L}^0}\sqrt{1+\boldsymbol{K}_0\cdot\boldsymbol{K}_0}>\sqrt{\boldsymbol{L}^0\cdot\boldsymbol{L}^0}\sqrt{\boldsymbol{K}_0\cdot\boldsymbol{K}_0}\geq|\boldsymbol{L}^0\cdot\boldsymbol{K}_0|$(最后一步是 Schwarz 不等式)。可见两个正时(非正时)Lorentz 变换的乘积是正时 Lorentz 变换,一个正时一个非正时 Lorentz 变换的乘积是非正时 Lorentz 变换。设 $L\in O(3,1)$,由式(6.11)可得

$$L^{-1}=G^{-1}L^{-1}G \quad 或 \quad (L^{-1})^{\rho}_{\mu}=g^{\rho\sigma}L^{\nu}{}_{\sigma}g_{\nu\mu}\equiv g^{\rho\sigma}L_{\sigma\mu}\equiv L_{\mu}{}^{\rho} \tag{6.17}$$

当 $\rho=\mu=0$ 时,上式给出

$$(L^{-1})^0{}_0=L_0{}^0=L^0{}_0 \tag{6.18}$$

这表明 L 和 L^{-1}同为正时或同为非正时 Lorentz 变换。

由矩阵理论知

$$(L^{-1})_{\mu}{}^{\rho}=\frac{A_{\mu}{}^{\rho}}{\det L}\ (A_{\mu}{}^{\rho}\text{是}L\text{的}\mu\text{行}\rho\text{列代数余子式}) \tag{6.19}$$

于是

$$A_0{}^0=L^0{}_0\det L \tag{6.20}$$

注意

$$A_0{}^0=\begin{bmatrix}L^1{}_1 & L^1{}_2 & L^1{}_3\\ L^2{}_1 & L^2{}_2 & L^2{}_3\\ L^3{}_1 & L^3{}_2 & L^3{}_3\end{bmatrix} \tag{6.21}$$

是空间坐标变换的 Jacobi 行列式。它的正负代表空间坐标系的手征不变(左手坐标系变到左手坐标系,右手坐标系变到右手坐标系)或改变。

由上面的讨论,我们看到 $O(3,1)$的连通分支(ⅰ)是它的一个正规子群,叫正常正时 Lorentz 变换群,记作 $SO^+(3,1)$。连通分支(ⅱ)和(ⅲ)和(ⅳ)是它的陪集。连通分支(ⅰ)和(ⅱ)的并,也是 $O(3,1)$的一个正规子群,叫正时 Lorentz 变换群,记作 $O^+(3,1)$。连通分支(ⅰ)和(ⅲ)的并也是 $O(3,1)$的一个正规子群,叫保手征 Lorentz 变换群。而连通分支(ⅰ)和(ⅳ)的并也是 $O(3,1)$的一个正规子群,叫正常 Lorentz 变换群,记作 $SO(3,1)$。

Ⅲ. 无穷小 Lorentz 变换,李代数 $o(3,1)$

齐次 Lorentz 群 $O(3,1)$是由全体满足下面条件的 4×4 实矩阵 X 构成的

$$X^TGX = X \tag{6.22}$$

它的李代数 $o(3,1)$是由全体满足下面条件的 4×4 实矩阵 ξ 所构成的

$$\xi^TG + G\xi = 0 \tag{6.23}$$

直接验证可知

$$I_{01} \equiv \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \equiv -I_{10}, I_{02} \equiv \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \equiv -I_{20}$$

$$I_{03} \equiv \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \equiv -I_{30}, I_{12} \equiv \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \equiv -I_{21}$$

$$I_{23} \equiv \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \equiv -I_{32}, I_{31} \equiv \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \equiv -I_{13} \tag{6.24}$$

这 6 个 4×4 矩阵满足等式(6.23),并构成 $o(3,1)$的一个基。在 $O(3,1)$群单位元 E 的某个邻域中,每个群元(Lorentz 变换)X 可以唯一地表作(拉丁字母指标取值

1,2,3,希腊字母指标取值0,1,2,3)

$$X = \exp\left(\sum_{\mu<\nu}\omega^{\mu\nu}I_{\mu\nu}\right) = \exp\left(\frac{1}{2}\sum_{\mu,\nu}\omega^{\mu\nu}I_{\mu\nu}\right) \tag{6.25}$$

其中 $\omega^{\mu\nu} = -\omega^{\nu\mu}$, $I_{\mu\nu} = -I_{\nu\mu}$。

这16个 $\omega^{\mu\nu}$ 实数,4个对角元为零,只有6个是独立的。$\{\omega^{01},\omega^{02},\omega^{03},\omega^{12},\omega^{23},\omega^{31}\}$ 常被选为群参数,对应的无穷小算符是(6.24)。I_{0j} 是惯性参考系 Σ 相对 Σ' 在 x'^j 方向的单位无穷小推动(boost),而 I_{12} 是 Σ 相对 Σ' 绕 x^3 轴单位无穷小转动,余类推。它们满足如下的对易关系:

$$[I_{01},I_{02}] = -I_{12},[I_{02},I_{03}] = -I_{23},[I_{03},I_{01}] = -I_{31} \tag{6.26}$$

$$\left.\begin{aligned}&[I_{01},I_{23}] = 0,[I_{01},I_{31}] = I_{03},[I_{01},I_{12}] = -I_{02}\\&[I_{02},I_{23}] = -I_{03},[I_{02},I_{31}] = 0,[I_{02},I_{12}] = I_{01}\\&[I_{03},I_{23}] = I_{01},[I_{03},I_{31}] = -I_{01},[I_{03},I_{12}] = 0\end{aligned}\right\} \tag{6.27}$$

$$[I_{23},I_{31}] = I_{12},[I_{31},I_{12}] = I_{23},[I_{12},I_{23}] = I_{31} \tag{6.28}$$

作复线性变换

$$A_x = \frac{1}{2}(I_{23} + iI_{01}),A_y = \frac{1}{2}(I_{31} + iI_{02}),A_z = \frac{1}{2}(I_{12} + iI_{03}) \tag{6.29}$$

$$B_x = \frac{1}{2}(I_{23} - iI_{01}),B_y = \frac{1}{2}(I_{31} - iI_{02}),B_z = \frac{1}{2}(I_{12} - iI_{03}) \tag{6.30}$$

可以验证

$$\boldsymbol{A}\times\boldsymbol{A} = \boldsymbol{A}$$

$$\boldsymbol{B}\times\boldsymbol{B} = \boldsymbol{B}$$

$$[\boldsymbol{A},\boldsymbol{B}] = 0$$

对应于两个互相对易的角动量。可见复化后的李代数 $o(3,1)$ 是两个 $so(3)$ 理想的直和:$o(3,1) = so(3)\oplus so(3)$ 而复化后的李群 $O(3,1)$ 是两个正规子群 $SO(3)$ 的直积:$O(3,1) = SO(3)\otimes SO(3)$。直积群 $G = A\otimes B$ 的所有不等价的不可约表示是正规子群 A 的所有不等价的不可约表示与正规子群 B 的所有不等价的不可约表示的“直积”。所以 $O(3,1)$ 的所有不等价的不可约表示为 $D^{[j,j']}(\boldsymbol{\theta},\boldsymbol{\theta}') = D^{[j]}(\boldsymbol{\theta})\otimes D^{[j']}(\boldsymbol{\theta}')$。