量子力学新讲

现代导引

陈童

Copyright © 2021 陈童

Published by Publisher

下载网址：http://newquanta.com/

作者本人 (即陈童) 对本讲义保有所有版权，目前读者可在上面的网站上自由下载，也可以私自传播，但仅限用于学习或教学等非商业用途，禁止私自以本讲义的部分或全部来获利，一经发现，作者将保留诉诸法律的权利。如果你有任何反馈，或者发现书中的任何错误，请给作者发邮件，tongchen@ecut.edu.cn

2021 年 *12* 月

Contents

0 最大的尺度和最小的尺度 13

0.1 最大的尺度 14

0.2 最小的尺度 15

0.2.1 习题 16

1 量子力学的建立 17

1.1 普朗克公式以及能量量子化 18

1.1.1 理想黑体模型 18

1.1.2 瑞利-金斯公式 19

1.1.3 普朗克公式与能量量子化 19

1.1.4 普朗克遗留的问题 20

1.1.5 习题 22

1.2 玻尔的氢原子模型 22

1.2.1 原子物理的难题 22

1.2.2 氢原子的量纲分析 23

1.2.3 玻尔氢原子模型 25

1.2.4 量子力学如何解决氢原子问题 27

1.3 波粒二象性以及波动力学简介 27

1.3.1 光的波粒二象性 27

1.3.2 微观粒子的波粒二象性 29

1.3.3 波动力学简介 31

1.3.4 双缝干涉实验再讨论 . . . 33
1.3.5 习题 . . . 35
1.4 电子自旋的发现 35
2 量子力学的基本原理 . . . 39
2.1 量子比特与薛定谔的猫 40
2.1.1 态叠加原理 . . . 40
2.1.2 量子比特与量子力学原理 . . . 42
2.1.3 多个量子比特 . . . 49
2.1.4 量子不可克隆定理与幺正性 . . . 51
2.1.5 薛定谔的猫 . . . 53
2.1.6 习题 . . . 56
2.2 算符与物理量 56
2.2.1 基础态 . . . 56
2.2.2 线性算符 . . . 57
2.2.3 为什么物理量用厄米算符表示 . . . 62
2.2.4 不确定原理与算符 . . . 64
2.2.5 幺正演化与薛定谔方程 . . . 66
2.2.6 算符的矩阵表示 . . . 68
2.2.7 * 一个数学附录 . . . 73
2.2.8 习题 . . . 75
2.3 量子纠缠 76
2.3.1 量子纠缠能实现超光速信息传递吗？ . . . 76
2.3.2 如何提取纠缠态中的信息 . . . 79
2.3.3 量子密集编码 . . . 80
2.3.4 量子隐形传态 . . . 81
2.3.5 GHZ 态以及为什么爱因斯坦错了 . . . 83
2.3.6 多体量子纠缠 . . . 84
2.4 补充材料：海森堡是怎么想到矩阵相乘的？ 85
2.4.1 原子光谱的一些事实和玻尔的新观念 . . . 85
2.4.2 海森堡的思路 . . . 86
2.4.3 海森堡思想的现代版本 . . . 89
3 坐标表象与波动力学 . . . 93
3.1 单个非相对论粒子 94
3.1.1 位置本征态和动量本征态 . . . 95
3.1.2 算符在坐标表象下的表示 . . . 96
3.1.3 位置算符，动量算符，哈密顿算符 . . . 96
3.1.4 概率流密度 . . . 98

3.1.5 习题 . . . 99
3.2 带电粒子与磁场的相互作用 99
3.2.1 磁场中带电粒子的薛定谔波动方程和概率流密度 . . . 99
3.2.2 超导体的磁通量子化 . . . 101
3.2.3 规范不变性 . . . 102
3.2.4 Aharonov-Bohm 效应 . . . 105
3.2.5 磁单极子 . . . 107
3.2.6 习题 . . . 109
3.3 量子隧穿效应 109
3.3.1 量子隧穿效应的半经典推导 . . . 110
3.3.2 约瑟夫森结 (Josephson junction) . . . 114
3.3.3 习题 . . . 117
3.4 自旋 1/2 的带电粒子 117
3.4.1 泡利算符 . . . 117
3.4.2 旋量波函数与自旋 1/2 薛定谔波动方程 . . . 119
3.4.3 习题 . . . 120

4 几个简单而重要的系统 . . . 121
4.1 两自旋耦合系统 122
4.1.1 泡利算符回顾 . . . 122
4.1.2 求解与讨论 . . . 124
4.1.3 习题 . . . 126
4.2 反射与透射 126
4.2.1 势散射的简单例子 . . . 127
4.2.2 散射态与束缚态 . . . 130
4.2.3 习题 . . . 131
4.3 单自由度线性谐振子 131
4.3.1 单自由度线性谐振子的量纲分析 . . . 132
4.3.2 求解与讨论 . . . 133
4.3.3 位力定理 . . . 136
4.3.4 习题 . . . 137
4.4 氢原子 138
4.4.1 角动量算符 . . . 140
4.4.2 能级和径向波函数 . . . 142
4.4.3 球谐函数 . . . 146
4.4.4 理解元素周期表的神秘规律 . . . 149
4.4.5 习题 . . . 150

4.5 朗道能级 151
4.5.1 朗道能级 . . . 151
4.5.2 习题 . . . 154
4.6 量子霍尔效应 154
4.6.1 经典霍尔效应 . . . 154
4.6.2 整数量子霍尔效应 . . . 156
4.6.3 量子力学的分析 . . . 158
4.6.4 无序—最终的解释 . . . 162
4.6.5 分数量子霍尔效应 . . . 165

5 定态微扰论和变分法 . . . 169
5.1 非简并定态微扰论 170
5.1.1 未扰动的系统以及投影算符 . . . 170
5.1.2 布里渊-维格纳微扰论 . . . 171
5.1.3 范德瓦尔斯力、万有引力、以及射电天文学 . . . 174
5.1.4 习题 . . . 179
5.2 * 有效哈密顿量、简并微扰论以及不稳定态 179
5.2.1 有效哈密顿量 . . . 180
5.2.2 简并微扰论 . . . 182
5.2.3 不稳定离散态 . . . 184
5.2.4 习题 . . . 187
5.3 变分法在量子力学中的应用 187
5.3.1 变分法求基态能量 . . . 187
5.3.2 定态作为能量泛函的极值状态 . . . 189
5.3.3 有关基态的几个定理 . . . 190
5.3.4 相变与对称性自发破缺 . . . 196
5.3.5 习题 . . . 199

6 * 量子力学中的对称性 . . . 201
6.1 量子变换 202
6.1.1 习题 . . . 206
6.2 时间平移、空间平移、空间旋转 207
6.2.1 时间平移 . . . 207
6.2.2 空间平移 . . . 208
6.2.3 空间旋转 . . . 210
6.2.4 习题 . . . 216

6.3 对称性 216
6.3.1 对称性与守恒定律 . . . 217
6.3.2 相位不变性与电荷守恒 . . . 218
6.3.3 对称性与能级简并 . . . 220
6.4 补充内容：伽利略推动 (Boost) 222
6.4.1 习题 . . . 226
6.5 空间反演对称性与时间反演对称性 226
6.5.1 空间反演对称性 . . . 226
6.5.2 时间反演对称性 . . . 231
6.5.3 习题 . . . 235

7 角动量理论与旋转群表示 . . . 237
7.1 角动量的本征值和本征态 238
7.1.1 角动量的本征值与本征态 . . . 238
7.1.2 * 旋转群的不可约表示 . . . 243
7.1.3 习题 . . . 244
7.2 角动量合成 245
7.2.1 角动量合成规则 . . . 245
7.2.2 * 旋转群不可约表示的直乘分解 . . . 250
7.2.3 习题 . . . 251
7.3 * 不可约张量算符与 Wigner-Eckart 定理 251
7.3.1 习题 . . . 255
7.4 2π 角旋转与超对称 255
7.4.1 2π 角旋转与超选择定则 . . . 255
7.4.2 超对称 . . . 256

8 全同性原理与多体量子力学 . . . 261
8.1 * 时间演化算符与路径积分 262
8.1.1 坐标表象中的时间演化算符 . . . 262
8.1.2 单粒子路径积分 . . . 265
8.2 * 路径的拓扑类和编织 268
8.3 * 全同粒子统计 272
8.3.1 全同粒子路径积分 . . . 272
8.3.2 $2+1$ 维以及 $3+1$ 维的讨论 . . . 274
8.3.3 作为规范对称性的置换对称性 . . . 278

8.4 * 自旋统计定理 283

8.5 玻色子和费米子 286

8.5.1 多体希尔伯特空间 . . . 286

8.5.2 置换对称性 . . . 287

8.5.3 应用举例 . . . 293

8.6 * 产生湮灭算符与全同粒子量子力学 295

8.7 * 附录：单粒子路径积分公式推导 301

9 * 再谈分数量子霍尔效应 . . . 305

9.1 从规范不变性到量子化霍尔电导 306

9.1.1 Byers-Yang 定理 . . . 306

9.1.2 量子化霍尔电导 . . . 307

9.1.3 分数量子霍尔效应的基态简并 . . . 309

9.2 Laughlin 波函数与分数统计 310

9.2.1 Laughlin 波函数 . . . 310

9.2.2 等离子体类比 . . . 310

9.2.3 准粒子激发 . . . 312

9.2.4 分数统计 . . . 313

9.3 附录：贝里相位 317

9.4 参考文献 318

10 含时问题与散射问题 . . . 319

10.1 量子力学含时微扰论 321

10.1.1 时间演化算符与含时微扰 . . . 321

10.1.2 恒定微扰 . . . 323

10.1.3 简谐振荡微扰 . . . 330

10.2 光的辐射与吸收 331

10.2.1 原子对光的辐射与吸收 . . . 331

10.2.2 爱因斯坦的受激辐射理论与激光 . . . 335

10.3 量子共振的一般性研究 337

10.3.1 共振近似 (旋波近似) . . . 337

10.3.2 Jaynes-Cummings 模型 . . . 338

10.4 * 散射的一般理论 341

10.4.1 散射的一般理论和 S 矩阵 . . . 341

10.4.2 S 矩阵的解析性和幺正性 . . . 344

10.4.3 哈密顿量 H 的散射定态 . . . 347

10.5 * 不稳定束缚态和散射共振态 349
10.5.1 含时问题与离散态衰变 . . . 350
10.5.2 自电离以及散射共振态 . . . 354
10.5.3 不稳定束缚态作为散射共振态 . . . 356
10.6 习题 361

11 * 再论势散射 . . . 363
11.1 一维势散射问题 364
11.1.1 直观的理论 . . . 365
11.1.2 与散射的一般理论的联系 . . . 367
11.1.3 空间反演对称性 . . . 372
11.1.4 一维势散射举例 . . . 374
11.1.5 共振态 . . . 376
11.2 三维势散射问题 378
11.2.1 直观理论 . . . 379
11.2.2 与散射的一般理论的联系 . . . 380
11.2.3 时间反演对称性和空间反演对称性 . . . 382
11.2.4 全同粒子散射 . . . 383
11.2.5 卢瑟福公式的非微扰精确求解 . . . 384
11.3 中心势场的分波法 387
11.3.1 对称性分析 . . . 387
11.3.2 分波法与散射振幅 . . . 389
11.3.3 如何确定相移 . . . 390
11.3.4 束缚态和共振态 . . . 392
11.3.5 关于相移的一般性讨论 . . . 393
11.3.6 零能束缚态与散射长度：一个例子 . . . 396
11.4 附注 399

12 * 密度算符与熵 . . . 401
12.1 密度算符与量子力学原理 402
12.1.1 期望值以及算符迹 . . . 402
12.1.2 从量子纠缠态到混态 . . . 403
12.1.3 Bloch 球 . . . 408
12.1.4 哥本哈根诠释和多世界诠释 . . . 409
12.2 再谈密度算符与系综 410
12.3 量子熵 412
12.3.1 冯诺伊曼熵 . . . 412
12.3.2 相对熵 . . . 415

12.4 开放系统以及热力学第二定律 417
12.4.1 开放系统的时间演化 . . . 417
12.4.2 热力学第二定律 . . . 420

13 * 狄拉克方程介绍 . . . 423
13.1 狄拉克方程 424
13.1.1 狄拉克方程的引入 . . . 424
13.1.2 矩阵表示 . . . 425
13.1.3 反粒子 . . . 427
13.1.4 协变形式 . . . 428
13.2 从狄拉克方程到自旋磁矩以及 LS 耦合 429
13.2.1 有效哈密顿量方法回顾 . . . 429
13.2.2 泡利哈密顿量 . . . 430
13.2.3 LS 耦合 . . . 430

14 附录: 线性代数概要 . . . 433
14.1 向量空间 433
14.2 有限维向量空间 435
14.3 线性算符 437
14.4 不变子空间、本征值与本征向量 439
14.5 内积空间 440



前言

《量子力学新讲》有两个基本目标，其一是，展现量子世界的奇妙和量子力学深层次的美。其二是，透彻地呈现量子力学的理论结构。将包括从量子力学入门到研究生量子力学的全部内容。

现代讲法，从一开始就糅合了量子信息科学的基本概念，深入讨论了量子力学的基本原理，量子力学中的对称性等内容，包括简介了超对称的基本思想。也会讲解量子力学在理论凝聚态物理中的一些基本应用，包括超导的有效理论和约瑟夫森结，量子霍尔效应，对称性自发破缺等等内容。同时对于量子散射理论也有较深入的讨论，并和粒子物理以及量子场论中的散射理论统一了起来。因此对于学习高能粒子物理的同学应该也会有些帮助。

数学处理力求简洁优美，不会有很繁琐枯燥的数学推导。对读者的预备知识要求并不高，主要是对线性代数要有较好的把握。但是要求读者在物理直观和数学能力上有一点点天分。

初学者在阅读第二章之前最好先阅读《附录：线性代数概要》。另外，所有章节内容中，打星号的部分肯定属于更深入的内容，初学者可以先跳过去。

习题还在逐步累积中，目前只能提供部分习题。



0. 最大的尺度和最小的尺度

本章非常简短，是让读者先闻闻量子力学的味。

你能想象的最大空间尺度是多少？你能想象的最小空间尺度又是多少？由于学过微积分，你可能觉得答案很简单，分别是无穷大和无穷小。但是，我们这个问题不是一个数学问题，而是一个物理问题。而物理是一门经验科学，它所涉及的量最终应该是可以测量的。所以我们实际上问的是，在物理上，我们所能测量到的最大空间距离是多少？所能测量到的最小空间距离又是多少？

你可能会想，这问题和量子力学有什么关系呢？关系很密切！实际上这问题和宇宙中两个最奇妙的东西密切相关，其一就是万有引力，其二就是我们这本书的主题，量子力学。

0.1 最大的尺度

先让我们来看最大的尺度。你可能会说，最大的尺度就是整个宇宙。但这个回答也是不准确的，宇宙到底有多大我们实际上是无法知道的。我们所能观测到的宇宙称之为可观测宇宙，可观测宇宙只是宇宙的一部分。之所以只能观测到宇宙的一部分，是因为信息的传播速度不能超过光速，而宇宙的年龄却是有限的，大约为 138 亿年，这样我们能接收到的最远信号就是来自 138 亿年前的光信号，它走过的距离就应该是我们能观测到的宇宙的半径，也就是 138 亿光年，约合 $1.3\times10^{26}m$。这个答案在量级上是对的，但具体来说却可能还差了好几倍。宇宙学家告诉我们，可观测宇宙的半径不是 138 亿光年，而是大约 465 亿光年。为什么会差这么多呢？

在回答这个问题之前，我们先要说说我们的宇宙。按照今天的看法，我们的宇宙起源于大爆炸，爆炸以后整个宇宙就在不断膨胀。这就好像我们吹气球，整个宇宙就好像是气球的表面。由于宇宙在膨胀，所以宇宙中两点之间的距离实际上是在增加，也就是说，实际上远方的星系是在远离我们，这称之为远方星系的退行，星系离我们越远则它退行的速度就越大。当然这样的效应要在宇宙学的尺度上才能观测到，在太阳系这样的尺度上是观测不到的。天文学观测也已经证实了这种远方星系的退行。而且值得说明的是，这种退行是由于空间的膨胀而引起的，而不是因为远方星系本身的运动而引起的，星系运动的速度当然不能超过光速，但退行的速度可以超过光速。实际上，离我们 138 亿光年远的星系相对于我们退行的速度已经大约是光速了。

另外，当我们说宇宙的年龄是 138 亿年的时候，这其实只是一个估算。对于宇宙的年龄，不同的宇宙学模型估算出来的数值会有所区别，在天文学上也会利用各种不同的观测结果来对宇宙学年龄进行估算，具体的数值依然是未定的，138 亿年只是比较公认的结果。

我们今天能接收到的最古老的光信号来自于 138 亿年前，但是由于发出这些信号的星系本身一直在离我们远去，即退行，而且它们退行的速度可以超过光速。所以这些星系今天离我们的距离就不只是 138 亿光年了，宇宙学家的估计是，这些星系今天距离我们大约为 465 亿光年。这就是我们今天的可观测宇宙的半径，也是我们能观测到的最大空间尺度。实际上，可观测宇宙是一个以我们为球心，直径约为 930 亿光年的球体。球体之外的宇宙是我们观测不到的。

0.2 最小的尺度

更有意思的问题是，我们能测量到的最小空间距离是多少？答案当然不是无穷小，实际上，按照量子力学的不确定关系，一个粒子的空间位置和它的动量不能同时被确定，我们测量一个粒子的空间位置所能达到的精度 δx, 和这个粒子的动量的精度 δp 之间必须满足一个关系，即

$$\delta x \delta p \sim \hbar, \tag{1}$$

式中的 $\hbar$ 就是著名的普朗克常数（更准确地说是约化的普朗克常数），它的量纲是长度乘以动量（和角动量的量纲一样），或者说是能量乘以时间的量纲，它的值很小，为 $1.05457266(63) \times 10^{-34} J \cdot s$，因此在经典世界里人们可以忽略它，但是在量子力学世界里 $\hbar$ 是不可忽略的，它的存在就标志着量子力学效应的存在。不确定关系 (1) 中的 δx 就是我们能测量到的空间尺度，我们也记为 r，而 δp 就是为了测到这一尺度所需要的探测粒子的典型动量。因此为了测量到更小的空间尺度，我们就需要提高探测粒子的动量。当探测粒子的动量足够高时，相对论效应就会是主要的，我们就可以忽略探测粒子的静止能量，这时候按照相对论，探测粒子的动量就是 E/c，E 就是探测粒子的典型能量。这样一来，不确定关系 (1) 就告诉我们，探测粒子能够探测到的空间尺度 r 和探测粒子的典型能量 E 之间就需要满足

$$rE \sim \hbar c. \tag{2}$$

也就是说，为了探测到更小的空间尺度，我们就需要用更高能的探测粒子。

那么我们可不可以不断提高探测粒子的能量，从而探测到越来越小的空间距离呢？答案是不能。为了说清楚这个问题，让我们考虑用两个探测粒子的对撞来探测空间距离，这时候 r 就是这两个探测粒子之间的最小距离。正如下面的分析将要告诉我们的，这个距离不可能是无穷小，原因在于这两个粒子之间必然会产生相互作用。

我们知道，自然界的基本相互作用有四种：电磁相互作用、强相互作用、弱相互作用、还有万有引力相互作用。粒子物理学家普遍认为，当能量足够高时，大约当粒子的能量超过 $10^{15} GeV$ 时（GeV 表示 10^9 电子伏特），这四种相互作用中的前三种就会统一成一种相互作用，而且这种相互作用有一个特点，即能量越高，相互作用的强度就越弱。

但是，万有引力相互作用恰恰相反，牛顿的万有引力公式告诉我们，两个质量为 M 距离为 r 的粒子之间的万有引力势能为

$$-G\frac{M^2}{r}. \tag{3}$$

利用爱因斯坦质能关系 $M = E/c^2$，我们可以将这个万有引力势能重新表达为

$$-G\frac{E^2}{rc^4}. \tag{4}$$

可见，对于万有引力来说，能量越高万有引力就越强。在通常的能量之下，两个探测粒子之间的万有引力是非常微弱的，可以忽略不计，但是当探测粒子的能量超过 $10^{15} GeV$ 甚至

更高时，万有引力相互作用就会变得很强。而我们前面也说过，这时候另三种相互作用的强度会开始变弱。因此，只要探测粒子的能量足够高，万有引力相互作用就会开始占主导地位。为了探测到更小的空间尺度，我们的探测粒子的能量当然是越高越好，因此我们只需要考虑探测粒子之间的万有引力相互作用。

但是，考虑到万有引力相互作用，我们的探测粒子的能量就不可能提高到无穷高，原因如下。首先，因为两个探测粒子之间的万有引力势能是负的，而相对论量子力学要求一个系统的总能量为正值，因此势能的绝对值不能超过静止能量和动能之和 E，最多和它相当。作为一个估算，这也就是说 $G\frac{E^2}{rc^4} \sim E$，换言之，当能量足够高时，两探测粒子之间的距离和探测粒子的能量之间满足如下关系，

$$r \sim G\frac{E}{c^4}. \tag{5}$$

实际上，这个关系就是著名的施瓦西半径，也就是说，当探测粒子能量高到一定的程度时，它们就会变成黑洞，施瓦西半径就是黑洞的半径。为了探测空间尺度，两个相互碰撞的探测粒子之间的距离必须大于它们的施瓦西半径，否则它们就合并成同一个黑洞了。因此，为了探测到更小的空间距离，探测粒子的能量也不能过高，否则按照 (5) 式，它们的施瓦西半径就会太大了。

现在，我们将公式 (2) 和公式 (5) 的左右两边对应地乘起来，就容易得到

$$r \sim \sqrt{\frac{G\hbar}{c^3}} = l_p. \tag{6}$$

l_p 就是我们能测量到的最小空间尺度，称之为普朗克长度，代入这几个基本物理常数的值可以算得，l_p 大约为 $10^{-35}m$。探测粒子的能量要多高才能探测到普朗克长度呢？由公式 (2) 我们可以得到这个能量为

$$\sqrt{\frac{\hbar c^5}{G}} = E_p. \tag{7}$$

E_p 称之为普朗克能量，它大约是 $10^{19}GeV$。

考虑到原子的能级为几个 eV，化学反应的能量为零点几个 eV，即使是当前最先进的粒子加速器 LHC 也只能把粒子的能量加速到 $14TeV$(TeV 就是 10^3GeV)，离普朗克能量也还差得很远。实际上，在实验室中我们肯定永远也无法把一个粒子的能量加速到普朗克能量。所以在实验室中我们可能永远也无法探测到普朗克长度这样小的距离。但是，宇宙大爆炸肯定是达到了普朗克能量的，并且它的效应可能在早期宇宙中有所余留，因此通过宇宙学观测，我们也许能观测到普朗克尺度上的物理。通过宇宙学观测探测普朗克尺度的物理本身是物理学家们正在研究的课题。在这里，最大的尺度，宇宙学尺度，和最小的尺度，普朗克尺度，密切联系起来了。

0.2.1 习题

1. 请用量子力学不确定关系估算氢原子的大小以及氢原子的最低能量。

2. 请讨论一下，如果普朗克常数变得非常大将会发生什么事情？

1. 量子力学的建立

本章讲述量子力学理论建立过程的若干关键点。内容包括普朗克的黑体辐射公式、玻尔的氢原子模型、波粒二象性以及波动力学简述。此外还包含一个对电子自旋的简单介绍。当然，我们不是要还原量子力学发展史，而是为了帮助读者更好地理解量子力学本身。因此我们这里关于量子力学建立过程的介绍是不全面的，最重要的缺失也许是，我们没有回顾海森堡建立矩阵力学的过程，在历史上，这实际上发生在薛定谔提出他的著名的薛定谔方程之前。不过，这个内容我们放在下一章正式开始讲述量子力学的基本原理中去了，因为我们觉得读者在对量子力学的基本原理有所了解之后再来看海森堡的思想会理解得更深入一些。

1.1 普朗克公式以及能量量子化

1.1.1 理想黑体模型

研究热辐射的物理学家发现，理想黑体的辐射本领是一个普适的函数 $e(\nu,T)$, 它表示温度为 T 时，理想黑体单位表面积单位时间之内所辐射出来的频率在 ν 附近的单位频率区间之内的电磁波的总能量，也叫理想黑体的单色辐出度。19 世纪末的实验物理学家通过测量画出了函数 $e(\nu,T)$ 的曲线，如图 (1.1) 所示，但是对于当时的物理学家来说，如何从

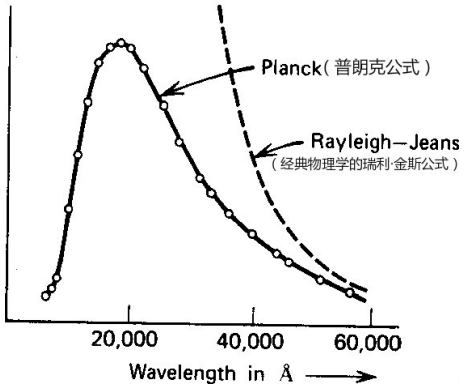

Figure 1.1: 黑体辐射谱的实验点，瑞利-金斯理论的拟合情况，以及普朗克公式的拟合情况。注意，横坐标是波长而不是频率。

理论上解释这些曲线是一个很大的挑战。

由于 $e(\nu,T)$ 是普适的，它不依赖于理想黑体的实现方式，因此我们只要就一种具体的理想黑体推导出 $e(\nu,T)$ 就可以了。一个这样的理想黑体可以这样来实现 (近似实现)，即取一个密闭的空腔，然后在这个空腔上挖一个很小的小孔，很显然，从小孔处往里看是漆黑的，因此这就是一个理想的黑体。将这样的空腔加热到温度 T 并让它达到热平衡，这时候测量小孔处辐射出来的电磁波，我们就能够得到函数 $e(\nu,T)$。假设我们记热平衡时空腔内单位频率区间上电磁波的能量密度为 $u(\nu,T)$，则由于小孔处的电磁波都是从空腔内部辐射出来的，所以很显然 $e(\nu,T)$ 一定正比于 $u(\nu,T)$，比例常数与频率 ν 和温度 T 都无关。因此要研究 $e(\nu,T)$，我们可以等价地研究能量密度 $u(\nu,T)$。$e(\nu,T)$ 的普适性意味着空腔内电磁波的能量密度 $u(\nu,T)$ 也是普适的。

研究 $u(\nu,T)$ 的方便之处在于，它和空腔的小孔无关，而仅仅是空腔达到热平衡时电磁场能量按照频率的分布情况，因此我们完全可以考虑一个完全密闭的空腔在热平衡时的能量分布情况。而且 $u(\nu,T)$ 的普适性告诉我们，只需要考虑下面这样一种更特殊的密闭空腔。那就是，设想一个密闭空腔，内壁是完全反射电磁波的镜子，空腔内充有热气体，它由一些带电谐振子组成，这些谐振子会不断地辐射电磁波，从而损失能量，但这些辐射出来的电磁波会在空腔内不停地反射，使得空腔内充满了电磁波，并最终又反过来作用在这些谐振子上，进而补充这些谐振子损失的能量。在热平衡时，空腔内的电磁波补充给谐振子的能量刚好等于谐振子辐射损失的能量。空腔内的电磁波能够补充给谐振子的能量显然正比于其能量密度 $u(\nu,T)$，而谐振子单位时间内辐射损失的能量则正比于谐振子本身

的平均能量 $\overline{E}$，因此我们有 $u(\nu,T)$ 正比于 $\overline{E}$。另一方面，电磁波每单位频率区间上的能量密度显然也正比于单位频率区间上的波模数 (即每单位频率区间上有多少个单色波)，这个波模数我们记为 $N(\nu)$。实际上，能量均分定理告诉我们，空腔内的电磁波与谐振子达到热平衡时，每一个单色波模的平均能量应该刚好和谐振子的平均能量 $\overline{E}$ 相等。上面的分析就告诉我们，$u(\nu,T)$ 正比于 $N(\nu)$ 与 $\overline{E}$ 的乘积，我们简单地写成

$$u(\nu,T)=N(\nu)\overline{E}. \tag{1.1}$$

1.1.2 瑞利-金斯公式

$N(\nu)$ 其实是可以直接计算的，计算方法也并不难。但这样的计算并不是我们关注的重点，因此我们将通过引用瑞利-金斯定律绕过这一计算。1900 年左右，瑞利和金斯提出了一个关于 $u(\nu,T)$ 的经典理论公式，

$$u(\nu,T)_{RJ}=\frac{8\pi\nu^2}{c^3}k_BT, \tag{1.2}$$

$u(\nu,T)_{RJ}$ 的下标表示瑞利-金斯的结果，k_B 是玻尔兹曼常数。注意这个公式中不出现普朗克常数 h，因此完全是一个经典的结果。但是人们发现，瑞利-金斯的这个结果只在 ν 很小的长波区域才能与实验测量数据相拟合，当 ν 比较大时就与实验数据相差很远。尤其是，瑞利-金斯的结果告诉我们，在 $\nu\to\infty$ 的紫外区域，能量密度将趋于无穷大，这当然是完全不可接受的。当时的物理学家称瑞利-金斯的这一不可接受的结论为紫外灾难。它表明对于黑体辐射问题，经典物理失效了！

为了使用瑞利-金斯的结果，下面我们按照经典物理的办法来计算谐振子的平均能量 $\overline{E}$。经典统计物理告诉我们，在温度为 T 的热平衡时，一个谐振子能量取值为 E 的概率 $\rho(E)$ 由玻尔兹曼分布给出，为

$$\rho(E)=\frac{1}{\mathscr{Z}}e^{-\frac{E}{k_BT}}, \tag{1.3}$$

式中 $\mathscr{Z}=\int_0^{+\infty}dEe^{-\frac{E}{k_BT}}$ 为归一化因子。因此，经典谐振子的平均能量为

$$\overline{E}=\int_0^{+\infty}dEE\rho(E)=\frac{\int_0^{+\infty}dEEe^{-\frac{E}{k_BT}}}{\int_0^{+\infty}dEe^{-\frac{E}{k_BT}}}=k_BT. \tag{1.4}$$

将这个结果代入公式 (1.1) 并与瑞利-金斯公式进行比较，我们就可以得到

$$N(\nu)=\frac{8\pi\nu^2}{c^3}. \tag{1.5}$$

1.1.3 普朗克公式与能量量子化

瑞利-金斯公式只能部分地拟合黑体辐射的实验数据，能完全拟合 $u(\nu,T)$ 的测量数据的公式是普朗克给出来的，它是

$$u(\nu,T)=\frac{8\pi h\nu^3}{c^3}\frac{1}{e^{\frac{h\nu}{k_BT}}-1}. \tag{1.6}$$

这就是著名的普朗克公式，普朗克最初是通过将瑞利-金斯的公式和维恩的公式进行一个适当的内插而得到的，普朗克发现这一公式能够在所有的频率上拟合 $u(\nu,T)$ 的测量数据。但问题是，理论上如何推导出这一公式呢？当然，普朗克公式中出现了一个新的参数 h，它就是后来著名的普朗克常数。

通过研究，普朗克发现，下面的“奇怪”假设可以导出公式 (1.6)。即假设空腔内的谐振子辐射电磁波的时候不是连续的，而是一份一份的，我们可以把每一份电磁波的能量记为 ε。由于谐振子辐射也要满足能量守恒，因此这一假设也就是说，谐振子的能量 E 不能取连续的值，而只能取 ε 的整数倍 $n\varepsilon, n=0,1,2,...,\infty$。按照这一假设，$\overline{E}$ 就不能通过积分来计算，而应该像下面这样计算，

$$\overline{E}=\frac{\sum_{n=0}^{\infty} n\varepsilon e^{-\beta n\varepsilon}}{\sum_{n=0}^{\infty} e^{-\beta n\varepsilon}}. \tag{1.7}$$

这里 $\beta=1/(k_BT)$。下面这个数学技巧可以让我们很容易地把结果计算出来。很容易验证上面的计算公式可以重写成

$$\overline{E}=-\frac{\partial}{\partial\beta}\log\left(\sum_{n=0}^{\infty} e^{-\beta n\varepsilon}\right). \tag{1.8}$$

对数函数括号里的那个无穷求和是一个等比数列的和，利用等比数列的求和公式我们容易计算出，

$$\overline{E}=\frac{\varepsilon}{e^{\frac{\varepsilon}{k_BT}}-1}. \tag{1.9}$$

将这个结果以及式 (1.5) 代入公式 (1.1) 我们就得到

$$u(\nu,T)=\frac{8\pi\nu^2\varepsilon}{c^3}\frac{1}{e^{\frac{\varepsilon}{k_BT}}-1}. \tag{1.10}$$

很显然，只要进一步假设

$$\varepsilon=h\nu, \tag{1.11}$$

我们就可以得到普朗克公式 (1.6)。

总之，普朗克发现，如果假设电磁波的辐射和吸收过程不是连续的，而是一份一份的，并且每一份的能量 $\varepsilon=h\nu$，那么就可以用统计物理的方法推导出黑体辐射的普朗克公式。普朗克做出的这一假设就是能量量子化假设，基本的能量单位 $h\nu$ 称为能量子。这种能量量子化的现象当然完全超出了经典物理的范围，因为在经典物理里面能量一定是连续变量。普朗克的这一假设也许是为了推导出黑体辐射的正确公式而做出的无奈之举，但它却开启了量子力学的大门。

1.1.4 普朗克遗留的问题

我们想强调一下，普朗克只是认为物质辐射或者吸收电磁波的时候是一份一份的，或者等价的说，普朗克只是认为空腔里面辐射电磁波的振子能量是量子化的，即这些振子能

量只能取 $nh\nu$ 这样的形式。至于电磁波本身，那个时候由于麦克斯韦电磁场理论的巨大成功，关于光是一种电磁波的观念已经深入人心，所以就连普朗克也根本不敢想象空腔里的电磁波本身就是量子化的，相反，普朗克肯定地认为，空腔里的电磁场是连续的。据说普朗克本人有一个比喻，用来理顺光辐射和吸收的时候不连续而以电磁场的形式存在的时候却连续这两者之间的矛盾。普朗克说，这就像水龙头滴水，水滴下来的时候都是一滴一滴的，但滴到池子里以后，池子里的水总是连续，黑体辐射的空腔就是那个水池，空腔里的电磁场总是连续的。当然，没有人会对这样的比喻真正满意，处在当时的时代，人们最可能做的事情就是，用电磁波的连续性来攻击普朗克能量量子化的假设。

观念的下一次突破来源于爱因斯坦，在 1905 年的《关于光的产生和转化的一个启发性观点》(网上可以找到这篇论文，有中文翻译，大家可以找来看一下) 这篇革命性的论文中，爱因斯坦坚持从普朗克公式出发，而不是从传统的电磁波理论出发，进而研究了空腔里的电磁波的熵，爱因斯坦发现，在高频时，电磁波的熵公式和粒子气的熵公式很类似。进而爱因斯坦提出了著名的光量子假设，即认为电磁波本身是量子化的，是由光量子组成的，每一个光量子的能量由 $h\nu$ 给出，并且他还用这个光量子假设解释了光电效应。爱因斯坦本人非常看重他的这篇论文，他认为这是他 1905 年 (所谓的物理学奇迹年) 所发表的 5 篇伟大论文中最具有革命性的一篇，注意，爱因斯坦并不认为狭义相对论是最革命的。这当然是因为狭义相对论和麦克斯韦方程依然属于同一个框架，而光量子的这篇论文突破了麦克斯韦的理论，提出了全新的观念！正是爱因斯坦的这篇论文把人们引向了微观粒子的波粒二象性。

然而无论是普朗克还是爱因斯坦，他们都没有解释空腔里振子的能量或者电磁波的能量为什么量子化，尤其是为什么量子化为 $nh\nu$ 的形式？而且，一旦人们接受光量子的新观念，那么普朗克原来的振子和电磁波之间的区别就没有那么大了，毕竟电磁波就是一种电磁场的振动，本身就可以看成是许许多多的谐振子。所以人们需要解释的其实是，为什么一个谐振子的能量会量子化为 $nh\nu$？这个问题直到海森堡提出他的矩阵力学才得以解决，其实，对谐振子的考虑和对氢原子能级问题的考虑本身就是促使海森堡提出矩阵力学的原因。当然我们今天的量子力学教材上通常是用与矩阵力学相等价的薛定谔方程来求解谐振子问题的。

不光电磁场是一个多自由度的谐振子系统，固体里面的声波当然也是一种振动形式，也可以看成是多自由度的谐振子系统，所以类似的普朗克公式和爱因斯坦的光量子观念应该也适用于声波。其实这就是爱因斯坦早期的重要贡献之一，爱因斯坦首先把量子化的观念应用于固体物理，类比于光波的光量子，他提出了声波的声子，并进而用一个类似于普朗克公式的公式解释了固体在低温下的比热容问题。

其实，从后来的观点看，电磁波和声波都是一种场，电磁波和声波的量子化实际上是场的量子化。一维线性谐振子是单自由度的谐振子，而在不考虑相互作用的时候，任何物理的场都可以看成是多自由度 (无穷自由度) 的谐振子系统。因此理解单自由度的一维线性谐振子也是理解场量子化的一个基础。

当然，我们还只谈了普朗克遗留的一些具体问题，而普朗克遗留的最大问题，当然就是建立一个关于量子化的自恰理论，这就是我们今天所说的量子力学。

1.1.5 习题

1. 斯特藩-玻尔兹曼定律（Stefan-Boltzmann law）告诉我们，黑体单位表面积的辐射总功率正比于温度的四次方，请证明这个结果。

2. 维恩位移定律 (Wien displacement law) 告诉我们，理想黑体的温度与其辐射本领最大值相对应的波长 λ_m 的乘积为一常数，即 $\lambda_m T = b$。式中，$b = 0.002897 m \cdot K$ 为维恩常量。请证明这个结果。在宇宙中，不同恒星随表面温度的不同会呈现出不同的颜色，温度较高的显蓝色，次之显白色，而红巨星表面温度只有 2000-3000K，显红色，试简单地解释这是什么原因?

1.2 玻尔的氢原子模型

1.2.1 原子物理的难题

19 世纪末 20 世纪初，物理学家对原子的发光行为进行了大量研究。他们发现原子的光谱是分立的，也就是每个原子都只能发出和吸收一些特定波长的光。从经典理论的角度看来，这是不可思议的，因为按照经典的电磁学理论，核外电子围绕原子核运动就会发出电磁辐射，但是这些电磁辐射的频率 (或者等价地，波长) 是可以连续取值的，而不是仅仅只能取一些特定的分立值。因此，如何解释原子的光谱就成了当时的物理学家面临的一大难题。

但在理论上，经典物理的问题还不仅仅是解释不了原子光谱，它在逻辑上更大的问题是，按照经典物理，绕核运动的电子会不断地发出电磁辐射，电磁辐射本身是有能量的，因此核外电子就会不断地损失能量，这样它就会一边绕核运动一边逐步往原子核上掉，很快所有的核外电子都会完全掉落到原子核上。原子的大小是 $10^{-10} m$ 的量级，但是原子核的大小只有 $10^{-15} m$ 左右，因此很快所有原子的尺寸都会坍缩 5 个数量级。我们的世界是由原子构成的，因此这也就意味着整个世界都会坍缩。这当然是不对的，可以说当时的物理学家面临的更大难题是，如何拯救世界。

这个拯救世界的人就是玻尔。但在讨论玻尔如何拯救了世界之前，我们还是先回到当时关于原子光谱的观测。当时关于各种原子的光谱有大量的观测数据，这些数据杂乱无章，没有人能理出头绪。但是不久之后，通过巴尔末等人的工作，就最简单的原子，氢原子的光谱，人们逐渐理出了头绪。人们发现，所有氢原子的光谱线都可以用两个整数来刻画，每一条光谱线对应一个整数对 (n,m), 对应规则是下面这个漂亮的数学公式

$$\nu = Rc\left(\frac{1}{n^2} - \frac{1}{m^2}\right), \tag{1.12}$$

式中 ν 是光谱线的频率，c 是光速，R 是里德伯（Rydberg）常数，用氢原子光谱的观测数据可以定出它的值是 $1.097373177 \times 10^7 m^{-1}$。这么多光谱线竟然可以归纳出如此简单漂亮的一个公式，简直就是一个奇迹。但是，如何解释这个奇迹呢?

正因为氢原子是最简单的原子，而且它的光谱满足上面那个奇迹般的公式。所以当时玻尔等物理学家想到，要从理论上解释原子光谱的数据，首先就应该研究氢原子，并推导出公式 (4.26)。这就是玻尔的氢原子模型解决的问题。

1.2.2 氢原子的量纲分析

在讨论玻尔的氢原子模型之前，我们先用量纲分析的办法来估计一下氢原子的一些相关物理量，比如说，氢原子的能量 E，核外电子绕核运动的轨道半径 r，以及核外电子的速度 v 等等。我们先作两个假设，第一，我们假设核外电子绕核运动的速度远小于光速 (后面我们会验证这个假设)，因此我们要处理的物理是非相对论物理，也就是说，我们的问题中将不涉及光速 c。第二，正如我们上面所讨论的，经典物理无法解决氢原子的问题，因此我们假设氢原子的物理需要涉及到普朗克所引入的量子物理，也就是说要涉及普朗克常数 $\hbar$。

由于原子核远比核外电子重，我们可以认为原子核是静止不动的，处于坐标原点。因此，氢原子的拉格朗日量 L 就等于核外电子的动能减去库仑势能 (注意这里的库仑势能是负的)，即 $L=\frac{1}{2}m_e\dot{\vec{x}}^2+\frac{e^2}{4\pi\varepsilon_0|\vec{x}|}$，这里 m_e 表示电子质量，$\vec{x}$ 表示核外电子的位置矢量。从这个拉格朗日量中我们很容易看到，$\frac{e^2}{4\pi\varepsilon_0}$ 总是作为一个整体出现的，因此我们将把这个量作为一个整体来处理，库伦势能当然是能量量纲 (用 $[E]$ 来表示)，因此 $\frac{e^2}{4\pi\varepsilon_0}$ 就是 $[E][L]$ 量纲，这里 $[L]$ 表示长度量纲。另外，普朗克常数是角动量的量纲，或者说能量乘以时间的量纲，因此 $\hbar c$ 也是能量乘以长度的量纲，即 $[\hbar c]=[E][L]$。这样一来我们就能构造一个无量纲的常数 α,

$$\alpha=\frac{e^2}{4\pi\varepsilon_0\hbar c}. \tag{1.13}$$

但是，正如我们前面假设的，我们要处理的问题是非相对论的，不涉及 c，因此 αc 必然作为一个整体出现 (由 α 的定义式显然可以看出 αc 与光速 c 无关)。反而是单独的 α 或者单独的 c 由于都涉及到了光速 c 因此不能出现在我们的非相对论分析中。

α 就是著名的精细结构常数，可以算得它的值近似是 $1/137$。按照我们的假设，我们要处理的问题涉及 $\hbar$, 因此上一段的分析就告诉我们，所有涉及物理量 $\frac{e^2}{4\pi\varepsilon_0}$ 的地方都可以用 αc 以及 $\hbar$ 的函数 $\alpha c\hbar$ 来替换。

总结一下，上面的分析告诉我们，氢原子的物理只涉及三个物理量，电子质量 m_e, 以及 αc 和 $\hbar$。因此氢原子的能量 E, 轨道半径 r, 以及电子速度 v 都必然能够表达成这三个物理量的函数，我们分别记为 $E=E(m_e,\alpha c,\hbar)$, $r=r(m_e,\alpha c,\hbar)$, $v=v(m_e,\alpha c,\hbar)$。由于 E、r、v 分别是能量量纲、长度量纲、以及长度除以时间量纲。因此，所谓的量纲分析就是用 m_e、αc、$\hbar$ 这三个量分别构造一个能量量纲的物理量，一个长度量纲的物理量，以及一个长度除以时间量纲 (即速度的量纲) 的物理量。注意到 αc 是速度量纲，$[m_e\alpha^2c^2]$ 是能量量纲，$[\hbar]$ 是能量乘以时间量纲，我们就很容易验证，$m_e\alpha^2c^2$，$\alpha c\hbar/(m_e\alpha^2c^2)=\hbar/(m_e\alpha c)$，以及 αc 这三个物理量正好满足我们的要求。又由于 m_e、αc、$\hbar$ 这三个量无法构造出任何无量纲的量，因此我们刚刚给出的构造就是唯一的。因此我们就必然有，

$$E\sim m_e\alpha^2c^2,\ v\sim\alpha c,\ r\sim\frac{\hbar}{m_e\alpha c}. \tag{1.14}$$

此外还有一些物理量也值得我们讨论一下，按照我们的量纲分析，αc 是核外电子速度，因此 $m_e\alpha c$ 就是核外电子的动量 p，而轨道半径 $r\sim\frac{\hbar}{m_e\alpha c}$，因此核外电子的角动量 $J=rp\sim\hbar$。其实更常见的是反过来分析，由于 $[\hbar]$ 也是动量量纲乘以长度量纲，即是角动

量的量纲，因此有 $J = rp \sim \hbar$, 又由于 $p \sim m_e \alpha c$, 所以就有 $r = J/p \sim \frac{\hbar}{m_e \alpha c}$, 正是 (1.14) 中给出来的结果。

当然，由于氢原子是一个束缚态，它的能量一定是负的，因此关于能量的表达式我们最好写成 $E \sim -m_e \alpha^2 c^2$, 用 $\sim$ 而不是用等于号的原因在于，量纲分析所确定的物理量可以相差一个数值系数，当然这个数值系数纯粹是一个数字，而不是任何物理量的表达式。要确定精确的数值系数那就需要一个更具体的氢原子模型了。另外，由 $v \sim \alpha c$ 以及精细结构常数的数值，我们可以估计到核外电子的速度要比光速小两个数量级，这就验证了我们的非相对论假设是自洽的。而将 αc、$\hbar$、m_e 的数值代入量纲分析的表达式 (1.14)，我们就能估计出，氢原子的能量大概在 $-10eV$ 的量级，电子的轨道半径大致在 $10^{-10}m$ 的量级。这样估计出来的原子大小和人们用其他的办法估计出来的原子大小是一致的，氢原子能量的估计也和电离能相当，这就进一步说明了，我们这一节一开始所作的两个假设都是合理的。尤其是，这一估计说明了，要解决氢原子的问题，我们必然需要用到量子物理!

读者可能想到，非相对论只不过是相对论在低速下的一种极限情况。因此人们完全可以坚持一开始就使用完整的相对论性的分析，而不仅仅是只分析非相对论极限。如果是这样，我们就应该在我们的量纲分析中多加一个物理量 c, 以氢原子的能量为例，考虑到相对论，能量 E 就应该是四个物理量 m_e、αc、c、以及 $\hbar$ 的函数，可以写成 $E = E(m_e, \alpha c, c, \hbar)$。与非相对论时的情形类似，我们可以由这四个量构造出一个能量量纲的量 $m_e c^2$, 但现在我们同时还可以由这四个量构造一个无量纲的量，即 $\alpha = \alpha c/c$, 注意，在之前的非相对论性的分析中 α 是不能单独出现的，它只能以 αc 的整体出现。由于 α 无量纲，因此就可以引入它的任意函数 $f(\alpha)$，它当然也是无量纲的，因此 $f(\alpha) m_e c^2$ 依然具有能量量纲，它才是由 m_e、αc、c、以及 $\hbar$ 这四个量构造出来的具有能量量纲的量的一般表达式。这也就是说，在相对论性的分析中，氢原子能量必定有如下形式的表达式

$$E = f(\alpha) m_e c^2. \tag{1.15}$$

函数 $f(\alpha)$ 仅仅用量纲分析是无法具体确定的，必须用真正的相对论量子力学的理论才能算出来。计算的办法大体是这样的，首先我们注意到 $\alpha = 1/137$ 是一个很小的数，因此我们可以设想把 $f(\alpha)$ 进行级数展开，即 $f(\alpha) = f_0 + f_1 \alpha + f_2 \alpha^2 + ...$, 这里 $f_0, f_1, f_2, f_3, f_4, ...$ 是与 α 无关的展开系数, 计算 $f(\alpha)$ 的关键技巧就在于用关于无量纲参数 α 的微扰展开的方式逐级地将级数展开系数 $f_0, f_1, f_2, ...$ 计算出来。实际上，在相对论量子力学中，$f(\alpha)$ 其实是 α 的偶函数，因此 $f_1, f_3, ...$ 这些系数必定都为零。f_0 对应的是我们取 $\alpha = 0$(也就是假设没有原子核产生的库伦场) 时的结果，这时候核外电子就是一个自由电子，它的能量当然就应该是爱因斯坦著名的 $m_e c^2$, 因此我们必定有 $f_0 = 1$，注意这一项的能量是自由电子的静止质量所贡献的能量，在前面的非相对论性的分析中是没有的。与非相对论的分析相对应的是 f_2 这一项，至于 f_4 以及更高阶的项，那就是对非相对论的结果的相对论修正，其中 $f_4 \alpha^4 m_e c^2$ 这一项通常被称之为氢原子能谱的精细结构。但是如果进一步考虑到相对论量子场论，那情况就会变得更复杂，这时候由于量子场论的重整化，$f(\alpha)$ 里面甚至可能包含 $\alpha^5 \log \alpha$ 这样的项。

1.2.3 玻尔氢原子模型

量纲分析的估计只能让我们确定量子物理有可能解决氢原子问题。但要具体的解决本文第一节提到的两大难题，尤其是推导出巴尔末的神奇公式 (4.26) 那就需要玻尔出马了。

由于当时还没有发现量子物理的基本原理和基本方程，玻尔提出了三条假设，这三条假设就构成了玻尔的氢原子模型。首先，针对核外电子由于辐射损失能量而不能稳定运动，要坍缩到原子核上的问题。玻尔提出了定态假设，玻尔说，考虑到量子物理以后，核外电子绕核运动的轨道不能是任意的，而是只能取某些特定的轨道，当电子在一个这样的特定轨道，比方说第 n 个特定轨道上运动时，它是不辐射电磁波的，因此它的能量就不会损失，而是会保持在一个确定的值 E_n，核外电子在特定轨道上运动的这种状态就称之为定态。

那么为什么我们能观测到氢原子的光谱呢？为此玻尔引入了第二条假设，跃迁假设。玻尔说，核外电子处在定态上的时候虽然不辐射，但是，它可能从一个定态跃迁到另一个定态，比方说从第 m 个定态跃迁到第 n 个定态 (假定 $m>n$)，这时候电子就会辐射出特定频率 ν 的电磁波，按照爱因斯坦的光量子假设，这个辐射出来的光子的能量就是 $h\nu$，因此由能量守恒，玻尔写出了如下方程，

$$h\nu = E_m - E_n. \tag{1.16}$$

玻尔的这两条假设自然地解释了氢原子光谱的分立性，原因就在于定态是分立的，而且公式 (1.16) 进一步告诉我们，每一条光谱线的确都对应两个整数 m,n，这正是巴尔末发现的规律。

但是，如果想进一步推导出巴尔末公式 (4.26)，玻尔就必须进一步找到定态能量 E_n 的表达式。类比于行星绕太阳运动，玻尔知道决定核外电子轨道的是能量和角动量这两个守恒量，量纲分析又告诉他，考虑到量子效应，必然有 $J\sim\hbar$。因此玻尔又引入了第三条假设，称之为轨道角动量量子化假设，玻尔说，处在定态 n 上的电子的轨道角动量 J 是量子化的，它满足

$$J = rp = n\hbar. \tag{1.17}$$

有了这三条假设，玻尔就能导出定态能量 E_n 了。玻尔的推导大致如下。首先为了简单起见，我们不妨假定核外电子绕核运动的轨道是圆轨道。因此按照牛顿定律，我们有

$$m_e\frac{v^2}{r} = \frac{e^2}{4\pi\varepsilon_0 r^2}, \tag{1.18}$$

很显然，这个公式也可以重写成，

$$m_e v^2 = \frac{\alpha c\hbar}{r}. \tag{1.19}$$

另外，由角动量量子化假设我们有

$$rm_e v = n\hbar. \tag{1.20}$$

由这两个结果我们容易推导出，定态的轨道半径 r 必然满足

$$r = n^2\frac{\hbar}{m_e\alpha c}. \tag{1.21}$$

很显然关于轨道半径的这个结果和我们前面量纲分析的结果一致。特别的，第 1 个定态的轨道半径为 $\frac{\hbar}{m_e\alpha c}=a_0$, a_0 称之为玻尔半径，可以算得它的数值是 $0.52917721067(12)\times10^{-10}m$。

另一方面，氢原子的能量是动能和势能的和，即 $E=\frac{1}{2}m_ev^2-\frac{e^2}{4\pi\varepsilon_0 r}=\frac{1}{2}m_ev^2-\frac{\alpha c\hbar}{r}$，由式 (1.19) 就可以得到，$E=-\frac{1}{2}\frac{\alpha c\hbar}{r}$(当然它也等于 $-\frac{1}{2}m_ev^2$)。将定态轨道半径的结果 (1.21) 代入这个能量的式子，我们就能得到第 n 个定态能量 E_n 的表达式

$$E_n=-\frac{1}{2}m_e\alpha^2c^2\frac{1}{n^2}=\frac{E_1}{n^2}. \tag{1.22}$$

显然这个能量的表达式和我们前面量纲分析的结果是一致的，式中 $E_1=-\frac{1}{2}m_e\alpha^2c^2$ 就是氢原子第 1 个定态 (称之为基态) 的能量，它也是氢原子的最低能量，代入相应物理量的值就可以算得 $E_1=-13.6eV$。$E_1=-\frac{1}{2}m_e\alpha^2c^2$ 的这个结果很好记，由于氢原子的总能量等于 $-\frac{1}{2}m_ev^2$，而在基态的时候电子的速度 $v=\alpha c$, 因此就有 $E_1=-\frac{1}{2}m_e\alpha^2c^2$。

更令人高兴的是，将公式 (1.22) 代入跃迁假设的公式 (1.16), 玻尔就能导出氢原子的光谱，

$$\nu=\frac{m_e\alpha^2c^2}{4\pi\hbar}\left(\frac{1}{n^2}-\frac{1}{m^2}\right), \tag{1.23}$$

这和巴尔末的神奇公式 (4.26) 形式完全一样，将它和巴尔末公式比较我们就可以得到，如果里德堡常数可以由下式给出

$$R=\frac{m_e\alpha^2c}{4\pi\hbar}, \tag{1.24}$$

那么玻尔就正确地推导出了氢原子的光谱。代入物理常数的值人们容易算出公式 (1.24) 给出的里德堡常数和由光谱数据定出来的里德堡常数吻合得非常好，这充分说明了玻尔的模型的确解释了氢原子的光谱。

正如我们已经看到的，为了解决氢原子的难题，玻尔引入了一些革命性的新概念，包括定态的概念，定态跃迁的概念，还有角动量量子化的概念。玻尔完全是凭其物理直觉引入这些概念的，但后来完整的量子力学理论表明，玻尔所引入的这些概念在完整的量子力学理论中也依然是核心概念，这当然说明了玻尔的物理直觉有多么强大。但从这里也可以看出玻尔理论的缺陷，比方说玻尔的定态概念完全是作为一个假设提出的，玻尔并没有解释清楚定态的本质，也没有解释为什么定态上的电子不辐射。同样玻尔也不能用一个基本原理来解释他的定态跃迁假设和角动量量子化假设。更令人不满意的是，在玻尔推导定态能量 E_n 的表达式的过程中，他还需要求助于经典力学的牛顿方程，而他也没有用一个新的基本原理将牛顿方程和他新引入的三条假设统一起来。换言之，玻尔并没有证明牛顿方程和他的三条假设是相容的，相反，正如本文一开始所说的，牛顿方程和经典的电磁学理论结合起来会导致完全错误的结论，而玻尔的三条假设正是为了挽救这样的错误而引入的，因此这些假设实际上天然地与牛顿方程不相容。因此，玻尔的理论虽然取得了巨大的成功，但在理论结构上依然是不令人满意的，它遗留的问题几乎和它解决的问题一样大。

玻尔之后，人们依然在寻找一个基本的量子理论，它将以牛顿的经典理论为其极限，即当标志量子效应的普朗克常数可以忽略时 (等价的就是让 $\hbar\to0$), 它要回到牛顿力学。并且它还要能正确地导出氢原子能级的公式 (1.22), 而无需求助于牛顿方程。首先取得这

一成功的是海森堡，1925 年海森堡沿着玻尔氢原子模型所开辟的道路提出了量子力学的正确方程，通常叫做矩阵力学。但是当时海森堡没能用他的矩阵力学推导出公式 (1.22)(半年以后泡利完成了这一工作)，首先用正确的量子力学方程推导出氢原子能级的就是泡利和薛定谔。但是薛定谔是沿着与玻尔模型不同的思路完成这一工作的，下一节我们会简单地介绍薛定谔的思路。

1.2.4 量子力学如何解决氢原子问题

正如我们已经看到的，氢原子问题的一个核心特征就在于，光谱的分立取值，或者等价地说，在于氢原子能量取值的分立性。但是频率取特定的分立值并不是量子物理的独有现象。在经典物理中，比方说鼓声，它就有一些特征的频率。这是由于鼓声来源于鼓膜的振动，或者说来源于鼓膜上的机械波。由于鼓膜的边缘是固定的，鼓膜上的这些机械波在边缘上反射以后最终就会在鼓膜上形成二维的驻波。驻波条件就要求了它们的频率只能取一些特定的分立的频率，每一种驻波模式对应一个特征频率，这就是鼓的特征频率。在数学上，这些特征频率由定义在鼓膜这样一个有限区域上的两维拉普拉斯算符的本征值决定。这个例子给当时的物理学家两点启发，第一，要得到氢原子能量的分立值，一个自然的办法是引入波动，特别是引入驻波。正好，当时德布罗意已经对所有的微观粒子引入物质波的概念了。第二，在数学上得到能量分立值的一个自然办法是，让这些分立值对应某个算符的本征值。

什么算符的本征值可以取成 E_1/n^2 这种形式呢？当时有很多数学家都在猜，但没有人得到结果。最终解决这个问题的是物理学家薛定谔，薛定谔给出的这个算符就是后来量子力学中的哈密顿算符。薛定谔提出量子力学方程的那篇著名论文标题就叫作《Quantisation As A Problem of Proper Value》，Proper Value 就是本征值。薛定谔是如何想到哈密顿算符的呢？经典力学中我们有哈密顿量，但是经典力学的哈密顿量是实数值的，它不是一个算符，薛定谔是通过什么操作从经典的哈密顿量走向量子的哈密顿算符的呢？这就是薛定谔的论文开篇就在讨论的问题，感兴趣的读者可以在网上找到这篇论文读一下。当然，需要强调的是，量子力学是全新的物理规律，哈密顿算符是无法通过严格的数学推导从经典物理的哈密顿量中推导出来的，薛定谔所给出的讨论也只是他如何猜测的一个思考过程，而不是一个严格的数学推导，在本质上，哈密顿算符还是猜出来的，只不过薛定谔是有物理地猜，而不是像当时的数学家那样瞎猜。哈密顿算符猜得对不对，关键就看它能不能给出氢原子能级 E_1/n^2 这样的本征值，薛定谔的原始论文更大的篇幅就是在证明这件事情。

1.3 波粒二象性以及波动力学简介

1.3.1 光的波粒二象性

前面我们说过，1905 年爱因斯坦提出，电磁场本身是量子化的，电磁场由光量子组成，光量子也就是我们今天常说的光子，它是微观粒子。爱因斯坦提出，频率为 ν 的电磁波，其能量是 $h\nu$ 的整数倍，可以写成 $\varepsilon = nh\nu$, n 就是组成这个电磁波的光子数目。也即是说，每一个光子的能量为 $h\nu$。

爱因斯坦的这个概念是彻底革命性的，因为当时麦克斯韦的电磁场理论已经深入人心了，按照麦克斯韦的理论，光就是一种电磁波，是一种波动。现在爱因斯坦却提出光是粒子，是光子，这当然完全超越了麦克斯韦的理论范围。但是，爱因斯坦指出光子的概念可以解释光电效应，用光子的概念，爱因斯坦预言：除非入射光的频率超过某个最小值 ν_{min}，否则当光照射在金属表面上的时候将不会有光电子出射。而当入射光的频率超过 ν_{min} 时，金属里面的自由电子有一定的概率吸收一个光子从而变成光电子出射，出射的光电子的动能将由 $h\nu - h\nu_{min}$ 给出，$h\nu_{min}$ 相应于电子脱离金属表面所需要的最小功，也叫逸出功。也就是说，根据光子的概念，爱因斯坦预言光电效应中出射光电子的动能将和入射光的频率成线性关系，这是经典的电磁波理论完全给不出的预言。但是实验很快证实了爱因斯坦的这个预言，并且还据此测出了普朗克常数 h，人们发现它和由黑体辐射曲线算出来的结果相当一致。因此爱因斯坦的光量子假设的确是有道理的。

即使如此，爱因斯坦的光量子概念也没有立即被当时的物理学家广为接受，这种情况一直持续到 1922 年康普顿完成著名的康普顿散射实验为止。康普顿散射实验即是用 X 射线和电子发生散射，实验结果表明，入射 X 射线的行为表现得完全类似于一个粒子，而且定量的结果也与根据爱因斯坦的光量子假说推导出来的结果完全一致。这就直接证实了光是微观粒子这一假说。

当光子的概念被实验确立之后，问题就反过来变成，光子还有没有波动性，还能不能发生干涉？因为根据经典物理，粒子和波是不相容的，一个东西要么是粒子，要么是波，光子如果是粒子，那它就不应该干涉了。当然，人们也知道大量光子的总体必定是可以干涉的，因为经典的电磁波就是由大量光子组成的，电磁波当然可以干涉。问题是，单光子能不能发生干涉呢？

为了确定这一点，英国物理学家杰弗里 · 泰勒重做了光学里面的双缝干涉实验。不过这一次，他设法让光源一次只发射一个光子，这一个光子通过双缝打在屏幕上之后，下一个光子再接着发射。泰勒发现，当屏幕上累积的光子数足够多以后，同样出现了干涉条纹，如图 (1.2) 所示。由于一次只通过一个光子，所以这种干涉不可能发生在不同的光子之间，

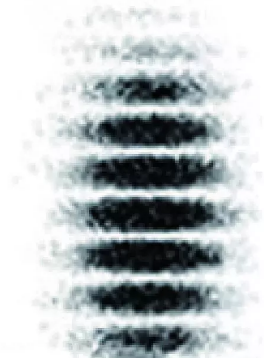

Figure 1.2: 泰勒的单光子双缝干涉实验。

只可能是单个光子和它自身产生了干涉。只不过干涉的结果要等到屏幕上累积了大量的光子以后才能显现出来。如果只看一个光子的话，那它在屏幕上出现的位置看起来就是随机

的。也即是说，屏幕上的干涉条纹直接反映的是单个光子在屏幕上各点出现的概率！

因此，光既是微观粒子，同时也是波，一个光子只会打在屏幕上的一个点，这和粒子的行为一样，但是单个光子也会和它自身干涉，因此单个光子同时也是波，这个波与经典的电磁波不同，它不是描写许多光子的总体，而是描写单个光子。屏幕上的干涉条纹反映的就是单个光子通过双缝以后在屏幕上的概率分布。当然，干涉条纹的这种概率解释最初人们并不清楚，它最早是由玻恩提出来的。光的这种即是粒子又是波的性质通常被称为波粒二象性。

1.3.2 微观粒子的波粒二象性

1924 年，德布罗意产生了一个大胆的关键性想法。他想到，对于光来说，之前人们都认为是波 (电磁波)，但是爱因斯坦却发现它也是粒子，是光子。德布罗意想，那么反过来，对于电子这样的微观粒子而言，原来人们都认为它们是粒子，那么有没有可能它们同时也是波呢?

基于这个猜测，德布罗意就把爱因斯坦对于光量子的关系式 $\varepsilon = h\nu$ 推广了，它认为一个能量为 E 的微观粒子，将同时也是一个角频率为 ω 的波，ω 和 E 之间满足关系式，

$$E = \hbar\omega, \tag{1.25}$$

式中 $\hbar = h/(2\pi)$ 为约化的普朗克常数，也常常简称为普朗克常数。

不仅如此，德布罗意进一步想到，根据狭义相对论，一个平面波的角频率 ω 和波矢量 $\mathbf{k}$ 共同构成了一个四维矢量 $(\omega, \mathbf{k})$。另一方面，一个粒子的能量 E 和动量 $\mathbf{p}$ 也构成了一个四矢量 $(E, \mathbf{p})$。因此如果 ω 和 E 之间满足 (1.25) 式，那动量 $\mathbf{p}$ 和波矢量 $\mathbf{k}$ 之间也应该类似地满足

$$\mathbf{p} = \hbar\mathbf{k}. \tag{1.26}$$

也即是说，一个动量为 $\mathbf{p}$ 的微观粒子将同时是一个波矢量为 $\mathbf{k}$ 的平面波。注意到波矢量的模长和波长 λ 之间的关系式 $|\mathbf{k}| = \frac{2\pi}{\lambda}$，德布罗意又由 (1.26) 式得到动量大小为 $p = |\mathbf{p}|$ 的微观粒子的波长 (称之为德布罗意波长)，

$$\lambda = h/p. \tag{1.27}$$

对于原子的核外电子而言，它的德布罗意波长大约在 $10^{-10}m$ 的量级。对于我们日常的宏观尺寸而言，这个波长无疑太小了，德布罗意认为这就是电子这样的微观粒子的波动性长期未被发现的原因。

德布罗意关于微观粒子同时也是波的这个猜想就叫做微观粒子的波粒二象性。它是光的波粒二象性的大胆推广。德布罗意提出微观粒子的波粒二象性时，关于光的波粒二象性基本上已经被实验所证实了。然而当时的人们认为波粒二象性是光的特殊本性，因此人们并不能立即肯定德布罗意的想法是对的 (除了爱因斯坦，爱因斯坦当时对德布罗意工作的评价是说它，“揭开了伟大序幕的一角”)。不过，德布罗意当时也从物理直观上表明了他关于微观粒子的波粒二象性并非瞎想，因为它可以非常自然而又漂亮地解释玻尔氢原子模

型的第三条假设，即核外电子的轨道角动量量子化为 $L = n\hbar$ 的假设。当时人们普遍觉得玻尔的这第三条假设很神秘，然而却不可或缺。

德布罗意的解释非常简单，根据波粒二象性，核外电子也是波，因此当核外电子处在轨道半径为 r 的定态上时，它的波必定要绕着半径为 r 的圆周首尾相连，形成一个驻波，如图 (1.3) 所示。这也就是说，这个圆周的周长 $2\pi r$ 必定是核外电子德布罗意波长的整数

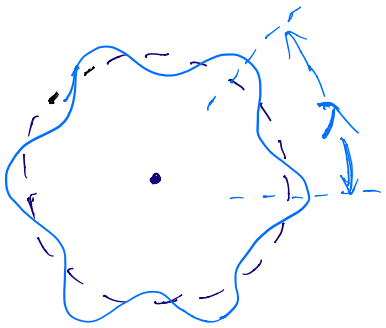

Figure 1.3: 德布罗意对玻尔的量子化条件的解释。

倍，即满足驻波条件 $2\pi r = n\lambda$。利用德布罗意关系 (1.27), 这个驻波条件就等价于玻尔的第三条假设，$L = rp = n\hbar$。

然而电子的德布罗意波长太短了，因此想用双缝干涉实验来检验它的波动性实现起来会比较困难。但是晶体中原子之间的间距却可以和电子德布罗意波长的量级相当。因此当时的人们想到可以用电子束在晶体上的衍射实验来检验电子是否有波动性。1927 年戴维孙和革末完成了这样的实验。实验结果表明，电子的确有波动性，会在晶体上发生衍射，而且根据衍射图案计算出来的入射电子波长也正好与德布罗意给出来的公式 (1.27) 吻合。

与光的波粒二象性一样，电子的波粒二象性也是适用于单个电子而不是适用于大量电子之间的，波粒二象性的波描写的是单个电子，而不是大量电子。也即是说，单电子就能发生衍射和干涉。电子的双缝干涉实验的确不容易实现，但是随着技术的进展，后来人们的确实现了这样的实验，不仅如此，和单光子的双缝干涉实验一样，人们后来也实现了单电子的双缝干涉实验，证实了单个电子就有波粒二象性。不仅是电子，人们后来也直接检验了诸如中子等等其它微观粒子的波粒二象性。到今天我们可以说，微观粒子的波粒二象性已经是量子力学中被实验检验得最多的概念之一。

当然，波的干涉来源于多个波的叠加，单个微观粒子可以和其自身干涉当然就意味着描写单个粒子的波可以是多个波的叠加。因此，单粒子波粒二象性的波是可以叠加的。当然，单个微观粒子和它自身干涉的干涉条纹要等到屏幕上累积的粒子数足够多以后才能看得出来，仅仅看一个粒子的话，它在屏幕上出现的位置看起来将是随机的。正是基于这个观察，玻恩提出，屏幕上的干涉条纹分布反映的是单个粒子落在相应点的概率。

1.3.3 波动力学简介

根据德布罗意的波粒二象性，一个动量为 $\mathbf{p}$ 能量为 E 的自由粒子将同时是一个频率为 $\omega = E/\hbar$, 波矢量为 $\mathbf{k} = \mathbf{p}/\hbar$ 的平面波，因此相应的波函数 $\psi(\mathbf{x},t)$ 可以写成

$$\psi(\mathbf{x},t) \propto e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}} = e^{-iEt/\hbar + i\mathbf{p}\cdot\mathbf{x}/\hbar}. \tag{1.28}$$

当然在经典物理中，将一个平面波的波函数写成这种复数形式只是一个数学技巧，真正物理的波函数可以通过取实部来得到。但是，量子力学后来的发展表明，复数形式对于量子力学里的波函数是实质性的，人们通常不能只取量子力学波函数的实部。当然，量子力学波函数是一个复数也意味着波函数本身不可以直接测量。因此在量子力学发展的早期，物理学家们对量子力学波函数在物理上代表什么争论不休，直到波恩提出波函数的统计解释这样的争论才开始停息。

我们不妨先来看一下公式 (3.5) 给出的自由粒子平面波满足什么样的波动方程。很容易看出，这样的波函数满足下面的关系式

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t) = E\psi(\mathbf{x},t), \quad -i\hbar\nabla\psi(\mathbf{x},t) = \mathbf{p}\psi(\mathbf{x},t). \tag{1.29}$$

由此又可以得到 $\mathbf{p}^2\psi(\mathbf{x},t) = -\hbar^2\nabla^2\psi(\mathbf{x},t)$，进一步注意到对于非相对论自由粒子，我们有 $E = \mathbf{p}^2/(2m)$，进而我们立即就可以得到如下方程

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t) = \left[-\frac{\hbar^2}{2m}\nabla^2\right]\psi(\mathbf{x},t). \tag{1.30}$$

这就是公式 (3.5) 给出的自由粒子平面波所满足的波动方程。注意，这个方程是一个线性方程，也即是说，不仅对于确定动量的自由粒子平面波这个方程成立，而且对于多个不同动量的自由粒子平面波的叠加，这个方程依然成立。当然，一般来说，由于波的相干叠加性，自由粒子的波函数不一定是 (3.5) 式这样的平面波，而更可能是多个不同波矢量 (也就是不同动量) 的平面波的叠加，而方程 (1.30) 的线性性就意味着它对于任何自由粒子波函数都成立。

考察微观粒子波函数所满足的波动方程是薛定谔的重要贡献。不仅如此，薛定谔还把 (1.30) 式这样的自由粒子波动方程推广到一般性的非自由粒子情形，也就是推广到粒子受一个势场 $V(\mathbf{x})$ 作用的情形。推广的办法是注意到根据 (1.29) 式，方程 (1.30) 右边可以理解为 $\frac{\mathbf{p}^2}{2m}\psi(\mathbf{x},t)$, 而方程左边可以理解为 $E\psi(\mathbf{x},t)$, 对于自由粒子 $E = \mathbf{p}^2/(2m)$，因此有方程 (1.30)。但是，当存在势场 $V(\mathbf{x})$ 时，粒子的能量将是动能与势能 $V(\mathbf{x})$ 之和，因此这时候就需要把方程 (1.30) 右边修正为 $\left[\frac{\mathbf{p}^2}{2m} + V(\mathbf{x})\right]\psi(\mathbf{x},t)$。最后，依然按照 (1.29) 式 (虽然严格来说这个式子仅对自由粒子平面波才成立) 将动量和能量都分别替换成对 $\mathbf{x}$ 和对 t 的导数，我们就会得到

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right]\psi(\mathbf{x},t). \tag{1.31}$$

这个波动方程就是著名的单粒子薛定谔方程。

由于波的叠加性，一个任意的波函数可以是多个不同角频率的波函数的叠加，而根据德布罗意的关系，角频率 $\omega = E/\hbar$，因此这也就是说，粒子的一个任意波函数可以是多个

具有不同能量值的波函数的叠加。这和经典物理完全不同，在经典物理里面，任何时候粒子都有唯一确定的能量，但是在量子力学里面，对于一个这种由多个有不同能量值的波函数叠加出来的波函数，其相应粒子的能量将无法确定。

但是，如果我们考察的是一个具有确定能量 E 的粒子，那么它的波函数将有确定的角频率 $E/\hbar$, 从而必定可以写成如下形式

$$\psi(\mathbf{x},t) = e^{-iEt/\hbar}\psi(\mathbf{x}). \tag{1.32}$$

对于这样的一个有确定能量的粒子，我们就说它处在定态。这里的定态其实就是玻尔最早在其氢原子模型中所引入的定态概念的一般化，因为玻尔的定态指的同样也是一个确定能级。将 (1.32) 式代入薛定谔方程 (1.31), 我们就可以得到定态波函数的时间无关部分 $\psi(\mathbf{x})$ 必须满足的方程

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right]\psi(\mathbf{x}) = E\psi(\mathbf{x}), \tag{1.33}$$

这个方程就是所谓的定态薛定谔方程，其中的 $-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})$ 就是单粒子的哈密顿算符，记作 H。因此定态薛定谔方程也可以写成

$$H\psi(\mathbf{x}) = E\psi(\mathbf{x}), \tag{1.34}$$

这个方程也叫做算符 H 的本征方程，其中的 E 也叫 H 的本征值。也即是说，定态能级其实是哈密顿算符的本征值。这也就是薛定谔的那篇著名论文《Quantisation As A Problem of Proper Value》标题的来源。

很显然，薛定谔方程 (1.31) 和人们之前在物理学中见过的所有方程都不一样，它包含了一个虚数单位 i。因此由薛定谔方程解出来的波函数 $\psi(\mathbf{x},t)$ 也必然是一个复函数。这样的一个复函数当然不可以直接测量，那么波函数在物理上代表什么呢？这个问题直到玻恩提出他的统计解释人们才搞清楚。前面我们说过，基于对单粒子干涉实验结果的观察，玻恩提出，屏幕上的干涉条纹分布反映的是单个微观粒子在相应点出现的概率。另一方面，人们在经典物理中就已经知道，干涉条纹可以由波的强度分布来计算，人们也知道，如果把波函数写成复数形式，那么波的强度将正比于波函数的模方。基于此，玻恩又进一步提出，量子力学波函数的模方代表的就是粒子在相应点出现的概率！当然，由于波函数的自变量 $\mathbf{x}$ 是一个连续变量，所以更严格地说，波函数的模方 $|\psi(\mathbf{x},t)|^2$ 代表的是粒子出现在 $\mathbf{x}$ 点的概率密度！如果将这个概率密度记为 $\rho(\mathbf{x},t)$，那么玻恩对波函数的解释即是

$$\rho(\mathbf{x},t) = |\psi(\mathbf{x},t)|^2. \tag{1.35}$$

很显然，将波函数乘上一个任意的非 0 复系数 c 不会改变粒子在空间不同点的相对概率分布，因此在物理上，$c\psi(\mathbf{x},t)$ 和 $\psi(\mathbf{x},t)$ 其实是等价的。但是，根据玻恩的解释，在体积元 $d\tau = d^3\mathbf{x}$ 内找到粒子的概率 dP 是 $dP = |\psi(\mathbf{x},t)|^2 d\tau$，而在全空间找到粒子的总概率当然就是 dP 对全空间的积分。由此人们就可以对波函数加一个自然的要求，即要求全空间的总概率等于 1，也即是

$$\int |\psi(\mathbf{x},t)|^2 d\tau = 1. \tag{1.36}$$

人们可以通过适当地调节波函数前面的复系数 c 来满足这样的要求，而这样的过程就称为波函数的归一化。

以上即是对波动力学的简单概括。当然波动力学不仅仅适用于单个微观粒子，它也可以推广到多个微观粒子相互作用的多粒子体系，这时候波函数就应该推广成 $\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N,t)$，这里 $\mathbf{x}_i$ 表示第 i 个粒子的位置坐标。而薛定谔方程 (1.31) 就应该推广成

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x}_1,...\mathbf{x}_N,t)=\left[-\sum_{i=1}^{N}\frac{\hbar^2}{2m_i}\nabla_i^2+V(\mathbf{x}_1,...,\mathbf{x}_N)\right]\psi(\mathbf{x}_1,...\mathbf{x}_N,t). \tag{1.37}$$

式中 $V(\mathbf{x}_1,...,\mathbf{x}_N)$ 表示这 N 个粒子的相互作用势能。

1.3.4 双缝干涉实验再讨论

为了让读者初步看到量子力学的不可思议。我们不妨对单电子的双缝干涉实验进行进一步的讨论。物理学家费曼最早在他著名的《费曼物理学讲义》里强调了双缝干涉实验对于理解量子力学的重要性。不过，在费曼讨论单电子的双缝干涉时，它还是一个想象中的实验，然而在今天，我们将要讨论的一切都已经是在真实的实验中证实了的。

电子的双缝干涉实验本质很简单，就是设计一把“电子枪”，让它一个接着一个地发射电子，发射出来的电子经过一个双缝，最后再打在远处的屏幕上，如图 (1.4) 所示。电

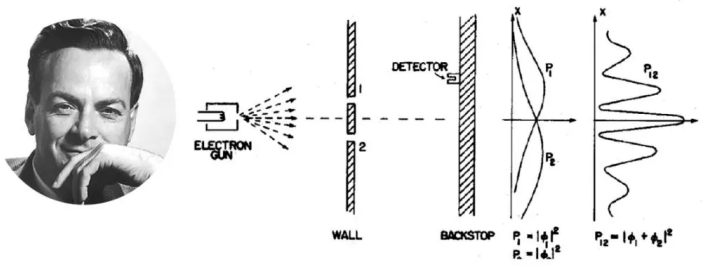

Figure 1.4: 费曼和电子双缝干涉实验。

子枪每次只发射一个电子，而且上一个电子已经打到屏幕上以后下一个电子才接着发射。等屏幕上累积了足够多的电子时，我们观察电子数在屏幕上的分布状况。我们也已经知道，这个分布曲线也可以解释为单个电子最终在屏幕上的概率分布曲线。

如果只打开第 1 个缝，而将缝 2 关闭，那当然不会有干涉，屏幕上的概率分布曲线将如图 (1.4) 中所示的曲线 P_1，类似的，只开缝 2 时概率分布曲线将是图 (1.4) 中所示的 P_2。但是，根据德布罗意的波粒二象性，电子也是波，所以当缝 1 和缝 2 同时打开时，将会发生波的干涉现象，最终屏幕上电子的概率分布曲线将是图 (1.4) 中所示的 P_{12}。很显然，由于存在干涉

$$P_{12}\neq P_1+P_2. \tag{1.38}$$

由于电子是波，假设我们记通过缝 1 到达屏幕的波为 ϕ_1，记通过缝 2 到达屏幕的波为 ϕ_2。那么根据玻恩的波函数的统计解释，显然我们将有

$$P_1 = |\phi_1|^2, \quad P_2 = |\phi_2|^2. \tag{1.39}$$

而当两个缝同时打开时，ϕ_1 波和 ϕ_2 波会相干叠加，屏幕上总的波将是 $\phi = \phi_1 + \phi_2$。根据玻恩的统计解释，这时候概率分布 P_{12} 将为 $P_{12} = |\phi|^2$，也即是

$$\begin{aligned} P_{12} &= |\phi_1 + \phi_2|^2 = |\phi_1|^2 + |\phi_2|^2 + (\phi_1\phi_2^* + \phi_1^*\phi_2) \\ &= P_1 + P_2 + (\phi_1\phi_2^* + \phi_1^*\phi_2). \end{aligned} \tag{1.40}$$

最终的表达式中多出来的交叉项 $(\phi_1\phi_2^* + \phi_1^*\phi_2)$ 正好可以用来解释干涉，称之为干涉项。

但是，电子也是粒子。作为粒子它是怎么通过双缝的呢？作为基本粒子，电子当然不可能分裂成两半，所以似乎它无法同时通过缝 1 和缝 2。也即是说，屏幕上的每一个电子似乎要么是通过缝 1 而到达屏幕的，要么就是通过缝 2 到达屏幕的。但是，不可思议的是，这种看起来很显然的观点其实是有问题的！因为如果真是这样的话，我们就可以把屏幕上的所有电子分成两类，一类是通过缝 1 来的，我们称它们为第 1 类电子，另一类是通过缝 2 来的，我们称它们为第 2 类电子。第 1 类电子由于只通过缝 1，对于它们来说，缝 2 存在或者不存在应该都一样，因此，第 1 类电子在屏幕上的分布曲线将为 P_1 曲线。类似的，第 2 类电子的分布曲线将为 P_2 曲线。但是，屏幕上的电子只有这两类，因此当双缝都打开时屏幕上所有电子的总分布曲线就应该是 $P_1 + P_2$，而不是 P_{12}，也就是说，应该根本就没有干涉！但这显然与实验结果相矛盾，也与波动的观点相矛盾！

也即是说，当双缝都打开时，屏幕上的所有电子可以分成 1、2 两类的观点必定是错的！这也即是说，当双缝都开屏幕上有干涉条纹时，我们其实根本就无法确定电子是怎么经过两个缝的。我们认为每一个电子要么经过缝 1 要么经过缝 2 的直观看法其实有问题！我们的日常生活经验误导了我们。这很不可思议，但却并不违背科学原理，因为物理学是一门实验科学，物理学是基于实验观测的。在日常生活中我们也许观察过子弹是如何经过双缝的，我们发现每一颗子弹要么经过缝 1 要么经过缝 2。但是，我们并没有观察过微观世界的电子是如何经过双缝的，因此当我们说电子要么经过缝 1 要么经过缝 2 时其实是想当然。

要想确定每一个电子是如何经过双缝的，最直接的方法就是用一个很灵敏的仪器“盯着”它看，也就是观测。假设我们的确这样做了，我们的确盯着电子看并进而可以确定每一个电子经过哪条缝，那屏幕上的电子就真能分成 1、2 两类了，按照我们刚才的逻辑推理，这时候屏幕上的干涉条纹就应该消失，电子的分布就应该是 $P_1 + P_2$。但这又是不可思议的，难道我们看一下电子就会影响最终的实验结果吗？真实的实验是怎么样的呢？真实的实验发现，的确如此，我们看一下电子真会影响最终的实验结果，如果我们通过观测能确定每一个电子经过哪条缝，那最终屏幕上真的不会出现干涉条纹。我们的日常生活经验又一次欺骗了我们，但好在这样的实验结果和我们的逻辑推理倒是一致的。

我们常常这样来概括上面的实验结果。我们说：当双缝同时打开并且我们没有看电子如何经过双缝时，电子的波函数由 $\phi_1 + \phi_2$ 这样的叠加态来描述，所以最终的概率由

$|\phi_1+\phi_2|^2$ 来描述，有干涉。但是，一旦我们观测到电子如何通过双缝，那电子的波函数就变了，它要么由 ϕ_1 描述 (当我们观测到电子经过缝 1 时)，要么由 ϕ_2 描述 (当我们观测到电子经过缝 2 时)，最终的概率就由这两种可能性的和 $|\phi_1|^2+|\phi_2|^2$ 给出，没有干涉。物理学家常常这样来解释这个结果，他们说观测导致了波函数的塌缩，当我们观测到电子经过缝 1 时，它的波函数就由原来的 $\phi_1+\phi_2$ 塌缩到了 ϕ_1，相反，如果观测到电子经过缝 2，那它的波函数就由 $\phi_1+\phi_2$ 塌缩到了 ϕ_2。

1.3.5 习题

1. 量子力学的波函数是复数，因此不可以直接测量，那么量子力学波函数可不可以测量呢？请查文献资料讨论这个问题。

2. 在电子的晶体衍射实验中，如果增加入射电子的动能，衍射环将如何变化？

1.4 电子自旋的发现

1922 年施特恩和格拉赫 (Stern and Gerlach) 做了一个著名的实验，这个实验最终表明电子有一个内禀的角动量，也就是电子自旋。在经典物理中，当一个粒子围绕另一个粒子作轨道运动是，它就有一个轨道角动量 $\mathbf{L}=\mathbf{x}\times\mathbf{p}$。而且我们知道，对于一个带电量为 q 的带电粒子，它的轨道角动量会产生一个磁矩，我们可以记作 μ_L，电动力学的知识告诉我们

$$\mu_L=\frac{q}{2m}\mathbf{L}. \tag{1.41}$$

当然，我们也知道，磁矩会和磁场耦合，如果存在一个磁场 $\mathbf{B}$，那磁矩 μ 和磁场相互作用的势能 V 将是

$$V=-\mu\cdot\mathbf{B}. \tag{1.42}$$

但是电子的自旋角动量完全是内禀的，它和轨道运动没有任何关系。人们可以粗略地将它想象成是电子围绕着自身的转轴自转，但是严格来说，这幅经典物理的图像是错的，电子的自旋纯粹是一个相对论量子力学的效应。实际上，由于电子有自旋，所以一个电子要转 720 度 (而不是 360 度) 才能回到原来的状态。所以在我们经典世界的人看来，电子自旋的存在是不可思议的，在理论上，要等到狄拉克提出著名的相对论量子力学方程，也就是狄拉克方程，人们才算对电子自旋有了比较深入的理解。

施特恩和格拉赫的实验装置如图 (1.5)a 所示。在这个实验中，施特恩和格拉赫让一束经过准直的银原子束通过一个非均匀磁场，然后再打在一个屏幕上。这个磁场如图 (1.5)b 所示，它在竖直的 z 方向有一个梯度 $\frac{\partial B_z}{\partial z}$。因此如果银原子有一个磁矩 μ 的话，它在 z 方向将受到一个力 F_z, $F_z=-\frac{\partial V}{\partial z}=\mu\cdot\frac{\partial\mathbf{B}}{\partial z}\simeq\mu_z\frac{\partial B_z}{\partial z}$。从图 (1.5)$b$ 可以看到，在这里磁场的梯度 $\frac{\partial B_z}{\partial z}$ 其实是负的 (下方的磁场更强，上方的磁场更弱)，因此如果 μ_z 为负的话，那银原子的受力将沿着 z 轴的正方向。

按照经典物理，$\mu_z=|\mu|\cos\theta$，θ 是银原子的磁矩与 z 方向的夹角。在斯特恩-格拉赫实验中，入射的银原子束是完全非极化的，也就是说，要么它的磁矩为 0，要么它磁矩的

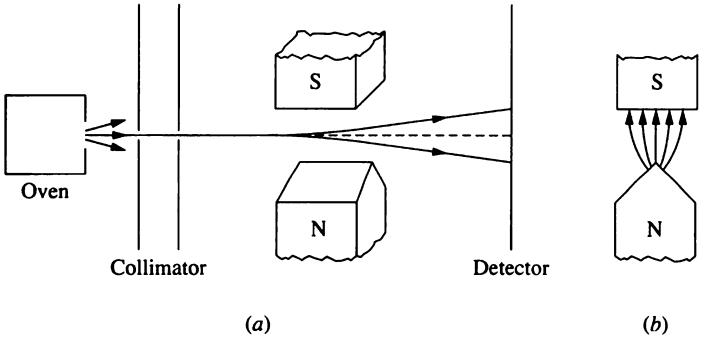

Figure 1.5: 施特恩-格拉赫实验。

方向就是随机的，即经典物理告诉我们的 θ 将有连续的不同取值。如果银原子束的磁矩为 0，那银原子将不受磁场的作用，因此最终打在屏幕上的银原子将会是一束。如果银原子有磁矩，那按照经典物理，θ 连续可变，因此银原子的受力 $F_z \simeq \mu_z \frac{\partial B_z}{\partial z} = \frac{\partial B_z}{\partial z}|\mu|\cos\theta$ 将在一个范围内连续可变，磁矩取向不同的银原子将会受到不同的力，从而最终会打在屏幕的不同位置，那最终打在屏幕上的银原子将展宽成连续的一片。

但是，实验发现，通过非均匀磁场以后，银原子束分裂了，但是仅仅分裂成了上下两束！也就是说，银原子的确有磁矩，但是 μ_z 的取值不是连续可变的，而是量子化的，μ_z 只有两个不同的分立取值。

实际上，银原子有 47 个核外电子和一个很重的原子核。斯特恩-格拉赫实验发现的银原子磁矩不可能来自于银原子核，因为磁矩的大小反比于粒子的质量，银原子核很重，其磁矩将会很小从而可以忽略。另外，我们知道带电粒子的磁矩总是来源于其角动量，而银原子核外的 47 个电子中，46 个电子的状态完全球对称，因此总角动量是 0，因此银原子的磁矩必定来源于额外的那第 47 个电子。但是最后的这个电子的空间波函数也是球对称的，因此其轨道角动量也是 0。因此，斯特恩-格拉赫的实验结果就意味着，电子除了轨道角动量以外，必定还有一个内禀的角动量，也就是自旋！它也是一种角动量，通常记作 $\mathbf{S}$。而 μ_z 的量子化就意味着，电子自旋的 z 分量 S_z 必定是量子化的，而且只能有两个不同的分立取值。当然，z 分量相对于 x 分量和 y 分量来说并没有任何特殊的地方，因此 S_x 和 S_y 必定也都是量子化的，也只能有两个不同的分立取值。

进一步的研究发现，电子自旋分量的这两个不同分立取值是 $\pm\hbar/2$。而电子自旋所产生的磁矩为

$$\mu_S = g\frac{q}{2m}\mathbf{S}. \tag{1.43}$$

式中 $q = -e$ 是电子的电荷，而 $g = 2$，也即是说与轨道磁矩 (1.41) 相比，自旋磁矩的公式多了一个 2 倍因子。这个 $g = 2$ 的因子是一个相对论效应，它可以从狄拉克方程中很漂亮地得到。当然，更仔细的实验测量也发现，电子自旋磁矩的这个 g 因子实际上和 2 有很细微的差别，实际上它近似是 $g \simeq 2(1 + \alpha/(2\pi))$, 这里的 α 就是精细结构常数，当然，这个细微的差别要用量子电动力学的计算才能解释。

其实，人们是事后才认识到斯特恩-格拉赫实验的正确物理解释的。斯特恩和格拉赫是1922 年做的实验，而直到 1926 年，乌伦贝克和古德斯密特 (Uhlenbeck and Goudsmit) 才提出电子自旋的概念。而且一开始人们是按照经典物理的图像将电子自旋理解为电子绕着自身的转轴自转。实际上，当乌伦贝克和古德斯密特提出电子自旋的时候，他们的导师埃伦费斯特建议他们去和洛伦兹讨论。洛伦兹很快就告诉他们，电子自旋的概念肯定是错的，因为如果将电子看成是一个小球，而自旋是这个小球绕着自身的转轴转动，那人们不难论证，这个小球表面的转动速度将超过光速，而这就违反了狭义相对论。尽管如此，埃伦费斯特还是帮助乌伦贝克和古德斯密特将论文发表了，并且安慰他们说，“不要担心，年轻犯点错误非常正常”。

2. 量子力学的基本原理

本章将会深入地探讨量子力学最基本最核心的原理，同时也会讲述量子力学的基本语言。这些原理和语言都是普遍适用的，它适用于单粒子体系、多粒子体系，适用于原子分子，也适用于许许多多原子分子构成的一整块材料，适用于量子计算机和量子通信，适用于量子场、甚至也适用于整个宇宙！

本章的第一节我们也讨论了量子比特、量子货币、量子不可克隆定理、和薛定谔的猫等等内容，当然，我们主要是想借这些内容阐明量子力学的基本原理。本章的第二节会比较抽象一些，是从算符的角度概括量子力学的原理，搭建量子力学的基本理论框架。第三节将探讨量子力学的核心本质之一，量子纠缠。我们将解释为什么量子纠缠不能实现超光速信息传递，我们还将简要讨论量子纠缠在量子通信中的重要应用，并且会说服大家，为什么爱因斯坦错了。在本章的补充材料中，我们将回顾历史，讨论海森堡是如何想到矩阵和矩阵相乘的，并进而把海森堡的思想和量子力学的基本原理联系起来。

2.1 量子比特与薛定谔的猫

2.1.1 态叠加原理

物理系统的量子状态可以由波函数来描述，对于一个微观粒子来说，波函数就是形如 $\psi(x,t)$ 的一个复函数。玻恩告诉我们，在量子力学里我们只能讨论概率，波函数的模方 $|\psi(x)|^2$ 就给出微观粒子在 x 处出现的概率密度。波函数包含了量子系统的所有信息，但是我们不能直接测量波函数，我们能直接测量的是诸如波函数模方这样的概率。

在经典物理里面我们有时候也使用概率，但在那里我们之所以使用概率都是因为我们没有掌握足够的信息，比如由于我们没有收集到所有大气分子的运动信息，我们就无法准确地计算明天的天气情况，因此我们才说明天下雨的概率是多少多少，但只要我们足够努力，原则上我们可以收集到所有大气分子的信息。在经典物理里，原则上我们能够掌握所有信息，并进而用这些信息确定地计算出诸如明天是否会下雨这样的问题。但是量子力学的概率是根本不同的，它是量子力学本身的内在性质，或者说是我们的世界本身的内在性质，也就是说，即使你知道所有的信息，也就是知道系统的波函数，你也只能计算概率，因为我们的世界本身就是如此，就是不确定的。何况，在量子力学里面，对于系统的一个任意状态，我们实际上常常无法通过测量收集到关于这个状态的所有信息，量子力学本身限制了我们的信息提取能力。

量子力学的规律是普适的，不仅仅微观粒子可以用一个波函数来描写，整个宇宙都能用一个波函数来描写，我们可以称之为宇宙波函数，著名物理学家霍金的重要工作之一就是研究这个宇宙波函数。当然，宇宙波函数比单个微观粒子的波函数 $\psi(x,t)$ 要复杂好多，它大概是下面那样的，

$$\Psi[a(t),h_{ij}(x),A_i(x),\psi_e(x),\psi_q(x),....] \tag{2.1}$$

其中 $a(t)$ 是宇宙的膨胀因子，$h_{ij}(x,t)$ 是引力波，$A_i(x)$ 是电磁场，$\psi_e(x)$ 是电子场，$\psi_q(x)$ 代表夸克场，此外还需要写上胶子场等等，总之要把一切基本粒子的场都作为宇宙波函数 Ψ 的自变量包括进来。

但是不管多么复杂，所有的波函数都满足态叠加原理，可以说态叠加原理是量子力学里面最基本的原理。根据态叠加原理，波函数 ψ 和波函数 $c\psi$(c 是一个任意的非零复数) 描写的是同一个量子态。更重要的是，根据态叠加原理，如果 ψ_1 是系统的一个可能态，ψ_2 也是系统的一个可能态，那么它们的任意线性叠加也将是一个可能态，即 $\psi=c_1\psi_1+c_2\psi_2$(c_1,c_2 是两个任意的复数) 也是一个可能态。

由于这个线性叠加原理，一个量子系统所有可能态的集合就构成了一个多维线性空间，而一个量子态 ψ 就是这个线性空间里的一个矢量，常常也记作 $|\psi\rangle$。$|\psi\rangle$ 这样的记号和波函数是一一对应的，但是，相比于波函数的描述，$|\psi\rangle$ 这样的量子态记号有许多优点：首先，它是通用的，无论是对于单粒子，还是多粒子，甚至是整个宇宙，我们都可以用一样的这种记号来表示其量子力学状态。其次，它是抽象的，它抽象地对应于量子态，而与具体怎么定量描述这个量子态无关。而波函数是一个具体的函数，它只是对量子态的一种具体描述方式，而后面我们将会看到，一个量子态可以有无穷多种相互等价的具体描述方式。比如，对于单个微观粒子而言，符号 $\psi(x)$ 总是一个数值，而 $|\psi\rangle$ 不是用来表示这个

函数值，而是用来代表函数映射本身，因此它不是一个复数，而是一个抽象的记号，当然，对于单个微观粒子而言，这个记号和具体的函数 $\psi(x)$ 是一一对应的。这样的记号称作狄拉克记号，不管是微观粒子的波函数还是更复杂的宇宙波函数，它们所描述的量子态都可以用同样的狄拉克记号来表示。用狄拉克记号，量子态的线性叠加就可以写成如下形式，

$$|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle. \tag{2.2}$$

作为线性空间里的矢量，人们完全可以按照线性代数里的方法取一个合适的矢量基，然后在这个矢量基中将态矢量 $|\psi\rangle$ 展开成分量形式，并把它的所有分量排列成一个列矢量，所以，如果你觉得 $|\psi\rangle$ 这样的记号过于抽象，你也完全可以将它想象成一个普通的列矢量，这两者是可以相互对应的。你可能想到，线性代数里除了有列矢量，还有行矢量，$|\psi\rangle$ 可以看作列矢量，那么行矢量是什么呢？很简单，行矢量就是列矢量的转置，但是，由于我们考虑的线性空间不是一个实线性空间，而是复线性空间，所以在量子力学里我们要多加上一个复数共轭 (为什么要加复数共轭的原因我们稍后会解释)，将行矢量规定为列矢量的共轭转置，也就是先将一个列矢量的每一个分量取复数共轭，然后再将结果转置，$|\psi\rangle$ 的共轭转置就记作 $\langle\psi|$，即

$$\langle\psi| = |\psi\rangle^\dagger, \tag{2.3}$$

记号 † 就表示共轭转置，也叫厄米共轭。因此在量子力学里，一个量子态如果用列矢量来表示，我们就用 $|\psi\rangle$ 这样的记号，如果用行矢量来表示我们就用 $\langle\psi|$ 这样的记号。

我们知道，一个行矢量乘以一个列矢量就会得到一个数。所以在量子力学里，由任意两个量子态 $|\psi\rangle, |\phi\rangle$，我们都可以计算出一个复数，记作 $\langle\phi|\psi\rangle$，它就是先将 $|\phi\rangle$ 态表示成行矢量 $\langle\phi|$，然后再将行矢量 $\langle\phi|$ 乘以列矢量 $|\psi\rangle$，$\langle\phi|\psi\rangle$ 就表示这个乘积。你已经想到了，由刚才的两个量子态，我们还可以计算出另一个复数 $\langle\psi|\phi\rangle$，这两个复数是什么关系呢？很显然，$\langle\psi|\phi\rangle$ 就是 $\langle\phi|\psi\rangle$ 的共轭转置，即 $\langle\phi|\psi\rangle^\dagger = |\psi\rangle^\dagger \cdot \langle\phi|^\dagger = \langle\psi|\phi\rangle$。这里请回想一下转置的规则 $(AB)^T = B^T A^T$，而复数共轭的规则是普通的，因此共轭转置的规则就是 $(AB)^\dagger = B^\dagger A^\dagger$。当然，一个复数的共轭转置其实就是它的复共轭（复数的转置还是它本身），因此我们有

$$\langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle. \tag{2.4}$$

当然狄拉克记号和波函数是一一对应的，你可能想知道 $\langle\phi|\psi\rangle$ 用波函数的形式来写是什么，对于单个微观粒子的情形，结果如下

$$\langle\phi|\psi\rangle = \int dx \phi^*(x)\psi(x). \tag{2.5}$$

为了帮助你理解这个结果和我们上一段说的有什么联系，我们不妨将 $\psi(x)$ 记为 ψ_x，将 $\phi(x)$ 记为 ϕ_x，并且将积分写成求和 (定积分的定义本来就是求和以后取极限)，这样一来上面这个结果就可以重写成

$$\langle\phi|\psi\rangle = \sum_x \phi_x^* \psi_x, \tag{2.6}$$

你想一下，这是不是行矢量乘以列矢量的标准形式（当然，别忘了对行矢量我们多了一个复数共轭）？ψ_x 就是列矢量 $|\psi\rangle$ 的第 x 个分量嘛！

特别的，我们考虑 $|\psi\rangle$ 与它自身的乘积，即 $\langle\psi|\psi\rangle$，对于单个微观粒子情形，按照公式 (2.5)，它就是

$$\langle\psi|\psi\rangle = \int dx \psi^*(x)\psi(x) = \int dx |\psi(x)|^2, \tag{2.7}$$

玻恩告诉我们，这就是处在 $|\psi\rangle$ 态上的粒子在空间各点出现的概率的总和，通常要求这个总概率等于 1，称之为量子态的归一化。现在我们就能理解为什么在定义行矢量的时候要多加一个复数共轭了，因为只有这样 $\langle\psi|\psi\rangle$ 才是概率嘛，尤其是，只有这样 $\langle\psi|\psi\rangle$ 才会是一个正实数嘛，即对于任意 $|\psi\rangle$

$$\langle\psi|\psi\rangle \geq 0, \tag{2.8}$$

式中等于 0 仅当 $|\psi\rangle$ 本身为 0 时才可能。$\langle\psi|\psi\rangle$ 也称作态矢量 $|\psi\rangle$ 模长的平方，因此归一化的态矢量就是模长为 1 的态矢量，即满足

$$\langle\psi|\psi\rangle = 1. \tag{2.9}$$

由于 $|\psi\rangle$ 和 $c|\psi\rangle$ 描述的是同一个量子态，因此态矢量的模长没有绝对的物理意义，这就是为什么我们可以将它们归一化的原因。

在数学上 $\langle\phi|\psi\rangle$ 有一个专门的名称，它叫作态矢量 $|\psi\rangle$ 和态矢量 $|\phi\rangle$ 的内积，因为它其实就是我们熟悉的三维空间矢量内积 (点乘) 的推广。和通常两个三维空间的单位矢量的内积大小反映的是这两个矢量在方向上的靠近程度一样，两个模长为 1 的态矢量的内积反映的其实是这两个态矢量之间的相似程度。特别的，两个完全不相似的态矢量的内积必定为零，这时候我们称这两个态矢量正交。很显然，这里的正交概念也是普通的三维空间矢量正交概念的一个推广。数学家常常把一个定义了内积的线性空间称作希尔伯特空间。因此很明显，态叠加原理以及上面几段的定义告诉我们，一个量子系统所有可能量子态的集合构成了一个希尔伯特空间，常常记作 $\mathscr{H}$。

2.1.2 量子比特与量子力学原理

你可能听说过量子比特，量子比特就是量子信息的基本单位，是量子计算机技术的基础。经典信息的单位是比特，一个比特就是或者为 0 或者为 1 的两种可能取值，实现一个经典比特需要两个不同的状态，比如 0 用某个系统的低电压状态来表示，1 则用高电压状态来表示。一个量子比特则是一个量子系统，它可以说是最简单的量子力学系统，它有两个完全不同的量子态，其中一个叫 $|0\rangle$ 态，另一个叫 $|1\rangle$ 态。量子比特有多种实现方式，比方说我们可以把一个两能级量子系统的低能级称作 $|0\rangle$ 态，高能级称作 $|1\rangle$ 态。再比方说，我们知道电子有自旋，电子自旋角动量的 z 分量 S_z 只有两个不同的量子化取值 $\pm\hbar/2$，这对应于电子的两个不同自旋状态，$S_z = \hbar/2$ 时我们就说电子处在自旋向上态，记为 $|\uparrow\rangle$，相反 $S_z = -\hbar/2$ 时我们就说电子处在自旋向下态，记作 $|\downarrow\rangle$。我们可以把一个电子的自旋量子态当作一个量子比特，自旋向下态 $|\downarrow\rangle$ 称作 $|0\rangle$ 态，自旋向上态 $|\uparrow\rangle$ 称作 $|1\rangle$ 态。当

然，我们也可以把一个光子的两个不同偏振态分别当作 $|0\rangle$ 态和 $|1\rangle$ 态，比如 $|0\rangle$ 对应 x 方向偏振的偏振态 $|x\rangle$，$|1\rangle$ 对应 y 方向偏振的偏振态 $|y\rangle$。

量子态的区分性

总之，实现量子比特 $|0\rangle$ 态和 $|1\rangle$ 态的方式很多。但是所有这些实现方式必须满足一个共同的要求，那就是 $|0\rangle$ 态和 $|1\rangle$ 必须可以确定地区分。比方说，光子的两个偏振态 $|x\rangle$ 和 $|y\rangle$ 就可以确定地区分，如果有两个光子，一个处在 $|x\rangle$ 态，另一个处在 $|y\rangle$ 态，那你只要取一个偏振化方向为 x 方向的偏振片 (即这个偏振片会让 x 方向偏振的偏振光顺利通过，而完全吸收偏振方向与 x 垂直的 y 方向偏振的偏振光)，然后分别让这两个光子通向这个偏振片，则能通过的光子就一定处于 $|x\rangle$ 态，通不过的光子就一定处于 $|y\rangle$，这样你就能以 100% 的把握将这两个态区分开来，这就是可以确定地区分的含义。相反，光子沿着 45 度角方向偏振的偏振态 (我们记为 $|+\rangle$) 和 $|x\rangle$ 偏振态之间就不可以确定地相互区分，这是因为，$|x\rangle$ 偏振态的光子当然一定能通过偏振化方向为 x 方向的偏振片，但是 $|+\rangle$ 偏振的光子也有一定的概率 (按照马吕斯定律，这个概率是 $\cos^2(\pi/4)$) 通过这个偏振化方向为 x 方向的偏振片，因此，假设你只有一个光子只能做一次实验，如果你观测到光子通过了偏振片，你就无从判断你的光子原来是处在 $|+\rangle$ 态还是处在 $|x\rangle$ 态, 因此你就不能 100% 地将这两个量子态区分开来。类似的，光子沿着 135 度角方向偏振的偏振态 (我们记为 $|-\rangle$) 和 $|x\rangle,|y\rangle$ 之间也都不可以确定地相互区分。当然，由于 135 度角和 45 度角成正交关系，所以 $|+\rangle$ 和 $|-\rangle$ 这两个偏振态是可以确定地区分的。

量子态的区分性是量子力学的一个很本质的性质，它在前沿的量子信息科学中有很重要的应用。比方说，有人提出可以利用这种区分性制作量子货币。通常货币的一个重大缺陷是人们总可以制造假币，即使电子货币也有假币。但是，量子货币绝对无法造假。

所谓的量子货币，是指的这样一种货币，每一张量子货币上都有一个编码，和一个确定偏振态的光子，比方说，这个偏振态可以是 $|x\rangle,|y\rangle,|+\rangle,|-\rangle$ 这四个偏振态中的某一个，我们不妨假设这四个偏振态有相同的概率被选用。量子货币上的编码和相应的光子偏振态之间是相对应的，可以用来检验光子属于四个偏振态中的哪一个。但是，编码和偏振态之间的对应关系只有银行的系统知道，是绝对不公开的。比方说，如果光子的偏振态是 $|x\rangle$，就编码为 00，如果是 $|y\rangle$ 就编码为 01，而如果光子的偏振态是 $|+\rangle$，就编码为 10，如果是 $|-\rangle$，就编码为 11。对于真币，银行的系统看到编码的第一位是 0，就知道偏振态一定是 $|x\rangle,|y\rangle$ 中的某一个，由于这两个偏振态可以确定地区分，所以具体是哪一个，银行的系统可以简单地通过使用一个偏振化方向为 x 方向的偏振片来确定，然后再把确定的结果和编码的第二位进行比较。类似的，看到编码的第一位为 1，银行的系统就知道光子的偏振态是 $|+\rangle,|-\rangle$ 中的某一个，通过使用一个偏振化方向为 45 度角方向的偏振片，银行就能进一步确定偏振态到底是哪一个。总之，对于真币，根据量子货币上的编码，银行的系统就能确定地检验它上面的光子的偏振态。但是，如果编码和光子偏振态的对应关系出错，那光子能通过银行的检测系统的概率就小于 1，比方说，在编码第一位，$|+\rangle,|-\rangle$ 态错误地对应到了 0，那根据马吕斯定律，这时候光子通过银行检测系统的偏振化方向为 x 方向的偏振片的概率就是 $\cos^2(\pi/4) < 1$。

伪造者想要伪造量子货币就必须复制光子的偏振态，为此他就得先知道他想伪造的这张真币上的光子的偏振态，因此就必须测量这个光子的量子态。伪造者当然也可以看到货币上的编码，他们甚至可能知道光子的偏振态是 $|x\rangle, |y\rangle, |+\rangle, |-\rangle$ 中的某一个，但是，他们不知道编码和偏振态之间的对应关系 (这个信息是不公开的，量子货币本身也没有这个信息，所以伪造者无从知道)。因此在测量光子的偏振态的时候，他就不知道是该使用偏振化方向为 x 方向的偏振片还是使用偏振化方向为 45 度角方向的偏振片，他就只能瞎蒙。比方说，他使用了偏振化方向为 x 方向的偏振片来测量光子的偏振态，由于光子处在四个偏振态上的概率均等，因此被测量的这张量子货币上的光子有可能通过伪造者的这个偏振片，也有可能通不过，如果光子通过了，伪造者就知道货币上的光子不可能处在 $|y\rangle$ 态，但是由于 $|+\rangle$ 态和 $|-\rangle$ 态都不能与 $|x\rangle$ 态确定地区分，所以剩下的三个偏振态 $|x\rangle, |+\rangle, |-\rangle$ 都有可能通过伪造者的偏振片，因此伪造者无从判断货币上的光子到底处在哪个偏振态，他只能从剩下的这三个态中瞎蒙一个，当然他有可能蒙对，但他蒙对的概率 p 一定小于 1 (实际上，在我们分析的这个例子中 $p = 3/4$)。对于伪造者测量的光子没有通过偏振片的情况，分析也是类似的。总之，由于这四个偏振态之间有一定的不可区分性，伪造者就不可能确定地知道光子到底处在哪个偏振态，他一定要靠蒙，而他蒙对的概率 p 一定小于 1，即 $p < 1$。

由于仅在伪造者蒙对的时候，他复制的假币上的编码和光子偏振态之间的对应才可能完全正确，才能 100% 地通过银行的检测，因此很显然，伪造者伪造的这张假币能通过银行检测的概率一定小于 p。看起来这个概率好像也不算小，但是，到此为止我们仅仅分析了量子货币上只有一个光子的情形。现在，假设货币上有 N 个光子 (货币上对应的编码当然也要扩大成一个 $2N$ 位的二进制数，每两位对应一个光子)，那很显然，伪造者伪造的假币能通过银行检测的概率就一定小于 p^N。但是，$p < 1$，因此只要 N 足够大，p^N 就会无限接近于零，也就是说，只要量子货币上的光子数目够多，假币能通过银行检测的概率就可以忽略不计。因此，量子货币无法造假!

以上只讨论了光子偏振态的可区分性。同样的，电子的自旋向上态 $|\uparrow\rangle$ 和自旋向下态 $|\downarrow\rangle$ 也可以确定地区分，因为如果一个电子处于这两个态中的某一个，你就可以通过测量它 z 方向的自旋 S_z 的值来区分它到底处在哪个态，如果测到 $S_z = \hbar/2$, 那它就处在 $|\uparrow\rangle$ 态，如果测到 $S_z = -\hbar/2$，那它就处在 $|\downarrow\rangle$ 态。一般地，如果一个量子系统的某一个物理性质 (物理量) 有两个不同的取值 λ_0 和 λ_1, 则这个物理性质的值为 λ_0 的量子态 (不妨记作 $|0\rangle$ 态)，和物理性质的值为 λ_1 的量子态 (不妨记作 $|1\rangle$ 态), 这两个量子态就必定是两个可以确定地区分的量子态。因为我们可以通过测量这个物理性质的值来 100% 地将这两个状态区分开来。这个结果当然可以推广到更一般的量子系统，如果一个量子系统的某一个物理性质 (物理量) 有多个不同的取值，记为 $\lambda_i, i = 1, 2, 3...$, 则这些不同取值对应的量子态, 分别记为 $|i\rangle, i = 1, 2, 3....$, 相互必定都可以确定地区分。

那么，在数学上，两个量子态 $|\psi_1\rangle$ 和 $|\psi_2\rangle$ 可以确定地区分的充要条件是什么呢? 很简单，充要条件就是这两个量子态要正交，即 $\langle\psi_1|\psi_2\rangle = 0$。这是因为量子态的内积反映的是两个量子态的相似程度，可以确定地区分的两个量子态必定是完全不相似的，因此也就必定正交。

从上面的分析中，我们可以得到一个重要的推论，即一个量子系统的某一个物理性质（物理量）A 不同取值的量子态 $|i\rangle, i = 1,2,3...$ 必定是两两正交的，即满足

$$\langle i|j\rangle = 0, i \neq j. \tag{2.10}$$

特别的，量子比特的 $|0\rangle$ 态和 $|1\rangle$ 态必定是相互正交的。通常我们还要求它们都是归一的，因此就有

$$\langle 0|0\rangle = \langle 1|1\rangle = 1, \quad \langle 0|1\rangle = \langle 1|0\rangle = 0. \tag{2.11}$$

测量与量子态塌缩

但是量子比特和经典比特的不同之处就在于，经典比特要么取 0，要么取 1，二者必居其一，但是对于量子比特来说，由于 $|0\rangle$ 和 $|1\rangle$ 都是它的可能态，则根据态叠加原理，$|0\rangle$ 和 $|1\rangle$ 的任意线性叠加也是这个量子比特的一个可能量子态，也就是说，任意形如下式的量子态 $|\psi\rangle$ 都是量子比特的可能态，

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle. \tag{2.12}$$

当系数 c_0 和 c_1 都不等于零时，我们就称这样的 $|\psi\rangle$ 态为 $|0\rangle$ 态和 $|1\rangle$ 态的相干叠加态。问题是，如果量子比特处在一个相干叠加态 $|\psi\rangle$，我们想问，它到底是取 0 呢？还是取 1 呢？(如果是用电子自旋态来实现量子比特，那么这就是在问，在相干叠加态上，自旋 S_z 的值是 $\hbar/2$ 还是 $-\hbar/2$？$-\hbar/2$ 就记为 0，$\hbar/2$ 就记为 1) 如果你说取 0，那么根据 $|0\rangle$ 态的定义，量子比特就不能处在叠加态 $|\psi\rangle$，而是应该处在 $|0\rangle$ 态，因为 $|0\rangle$ 态的定义就是量子比特取 0 时对应的态，所以这个答案肯定是不对的。同样，回答取 1 也是不对的。你可能会说，既不是 0，也不是 1，而是有一定的概率取 0，也有一定的概率取 1。但问题是，我们现在只有一个量子比特，而不是好多个相同的量子比特，而且这个量子比特的量子态还是确定的 $|\psi\rangle$，因此这里其实并没有使用概率的余地。如果按照我们的直观想法，每一个确定状态的量子比特总会有一个确定的值，但问题是，刚刚我们已经分析过了，如果这个状态是一个相干叠加态，那这个值既不能是 0，也不能是 1。但问题是，量子比特的值就只有 0 和 1 这两种可能性 (在自旋态实现中即是 S_z 的值只有 $\hbar/2$ 和 $-\hbar/2$ 两种可能性)，因此你也不能回答说取 1/2 或者诸如此类的其它值。

哪个答案都不对，那问题到底出在哪呢？只可能是我们的问题本身错了，或者说我们的直观想法错了！处在确定的叠加态 $|\psi\rangle$ 上的量子比特没有确定的取值，甚至我们也不能说它有一定的概率取 0 一定的概率取 1，而是它的取值根本无从确定，讨论它是没有意义的。有时候我们也说在相干叠加态上的量子比特的值是不确定的，这里所谓的不确定实际就是说问这个问题没有意义。

这就是最难理解的地方，我们在日常经典世界里形成的直观是，一个比特总会有一个值，不管我们看没看它，作没作测量，它总会取一个值。但就是这个直觉误导了我们，物理学是一门测量的科学，当我们说一个物理量取某个值时，这个值一定是指测量的时候会得到的值。严格来说，离开测量来讨论物理量的值是没有意义的。那为什么经典比特总有

一个确定的值？即使我们没有测量。回答是，经典比特是处在一个经典环境之中，即使我们没有直接对它进行测量，但其实经典比特所处环境中的空气分子等等东西都起到了测量它的作用。如果你把经典比特的所有环境都屏蔽掉，那它其实就会变成一个量子比特。

总之，量子比特到底取什么值这一问题，只有在我们测量量子比特的时候才有意义，任何这样的值一定是指的测量值。现在，假设量子比特处在一个相干叠加态 $|\psi\rangle$, $|\psi\rangle$ 由公式 (2.12) 给出，我们来测量它到底取什么，我们会测到什么值呢？答案是，0 或者 1，这也就是说，可能在某次测量实验中我们会得到 0，但如果我们用相同状态的量子比特重复测量，我们也可能得到 1。是得到 0 还是得到 1，是完全随机的，是概率性的。注意，在量子力学里，概率是和测量密切联系在一起的，没有测量就无所谓概率，而只有确定的量子态。

现在，假设在某次测量实验中，对于一个原来处在 $|\psi\rangle$ 态的量子比特，我们测到它的值是 0，那么根据 $|0\rangle$ 态的定义，测量完成以后这个量子比特就不可能还是处在 $|\psi\rangle$ 态，由于这时已经确定了它的值是 0，那测量以后它就一定处在 $|0\rangle$ 态上。这也就是说，测量过程必然会干扰被测对象，在我们讨论的这个例子中就是使这个量子比特从原来的 $|\psi\rangle$ 态坍缩到了 $|0\rangle$ 态。同样的，如果我们测到的值是 1，那测量完成以后量子比特就将塌缩到 $|1\rangle$ 态。由于测到 0 还是测到 1 完全是随机的，所以测量引起的量子态的塌缩也是随机的，可能坍缩到 $|0\rangle$，也可能塌缩到 $|1\rangle$。(这里请读者回想一下第一章中我们在 “双缝干涉实验再讨论” 这一小节中的讨论以及这一小节最后的总结)

在量子力学中，某个物理量的取值确定的态通常叫做这个物理量的本征态，相应的物理量的值称之为本征值。比如在量子比特的自旋态实现中，$|\uparrow\rangle$ 态和 $|\downarrow\rangle$ 态就是自旋 S_z 的两个本征态，因为它们都有确定的 S_z 值，分别为 $+\hbar/2$ 和 $-\hbar/2$, 在量子比特的这一物理实现中，所谓测量量子比特的值其实就是测量物理量 S_z 的值。比方说，假定记某个量子系统的某个物理量 A 的第 i 个本征态为 $|i\rangle, i = 1,2,3,...$(即是说当系统处于 $|i\rangle$ 态时，物理量 A 有确定取值 λ_i)。假定系统处在物理量 A 的某个相干叠加态 $|\psi\rangle$，它可以写成

$$|\psi\rangle = \sum_i c_i|i\rangle, \tag{2.13}$$

式中 c_i 为复叠加系数。那么根据我们前面对量子比特的讨论可以知道，系统在 $|\psi\rangle$ 态上其物理量 A 的值无法确定，没有定义！而且，量子力学的逻辑自恰性要求，如果对物理量 A 进行测量，假设测量开始之前系统处在这个多个不同本征态的相干叠加态 $|\psi\rangle$，如果我们测得了这个物理量的某个值 λ_i，那测量完成之后这个系统的量子态一定会变成与这个值相应的本征态 $|i\rangle$（而不再是处在原来的相干叠加态 $|\psi\rangle$）。通常我们称这个过程为测量引起的量子态的塌缩，但这种说法完全是出于物理学家们的用语习惯，在实际中，任何测量都不是瞬时的，而需要一个过程，因此其实并没有量子态的突变。研究量子态的塌缩这一过程如何进行往往并不容易，当前主流的观点叫做退相干，后面我们会简单解释退相干的基本思想。

测量会造成态的塌缩实际上就是量子货币最主要的缺陷，因为即使是一张真的量子货币，只要有人 (比方说伪造者) 去测量它的光子的偏振态了，那这测量就可能对光子的量子态造成干扰，比方说使得它从 $|+\rangle$ 态塌缩到 $|x\rangle$ 态，那这张真币也无法通过银行的检测

了，因此真币被别人测量以后也有可能被银行当成假币。所以，量子货币只是一个理论构想，它可以帮助我们领悟量子力学的奇妙，但却不一定能实用。

回到我们的量子比特。你可能会想，是不是量子比特的所有量子态都一定能写成公式 (2.12) 这样的形式呢？量子比特有没有可能处在某个不能写成形如 (2.12) 式的量子态呢？答案是没有，因为对于处在任何量子态 $|\psi\rangle$ 的量子比特，你都可以去测量它的值，而量子比特只有两个可能值，要么你测到 0，要么你测到 1，测到 0 那 $|\psi\rangle$ 就塌缩到 $|0\rangle$，测到 1 那 $|\psi\rangle$ 就塌缩到 $|1\rangle$，只要你去测量，任何 $|\psi\rangle$ 都必然要塌缩到 $|0\rangle$ 态和 $|1\rangle$ 态中的某一个，因此量子比特的任何量子态 $|\psi\rangle$ 都必然是 $|0\rangle,|1\rangle$ 的线性叠加。物理学家常常把这称为量子比特的 $\{|0\rangle,|1\rangle\}$ 态的完备性。很显然，完备性的概念也可以推广到任何量子系统的任何物理量的本征态集合。

测量概率的玻恩规则

我们已经说了，对于 $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ 态的量子比特，你去测量它的值，有可能你会得到 0，也有可能你会得到 1，得到 0 还是得到 1 完全是随机的。一个重要的问题是，测量得到 0 的概率是多少？得到 1 的概率又是多少？量子力学的基本原理说，对于归一化的 $|\psi\rangle$，测量得到 0 的概率 p_0 是 $|c_0|^2$，测量得到 1 的概率 p_1 是 $|c_1|^2$。而量子态 $|\psi\rangle$ 的归一化说的就是概率总和等于 1

$$p_0 + p_1 = |c_0|^2 + |c_1|^2 = 1. \tag{2.14}$$

利用前面的 $|0\rangle,|1\rangle$ 的正交归一性 (2.11)，我们很容易得到 $\langle 0|\psi\rangle = \langle 0|\big(c_0|0\rangle + c_1|1\rangle\big) = c_0\langle 0|0\rangle + c_1\langle 0|1\rangle = c_0$，类似的 $\langle 1|\psi\rangle = c_1$，因此概率 p_0 和 p_1 就可以写成，

$$p_0 = |\langle 0|\psi\rangle|^2, \quad p_1 = |\langle 1|\psi\rangle|^2. \tag{2.15}$$

由此我们就可以知道，如果 p_0 和 p_1 都不等于 0，那 $|0\rangle$ 态和 $|\psi\rangle$ 态就不可能是同一个量子态，也不可能正交，反过来也一样。

这就告诉我们，不正交的两个不同量子态一定不可以确定地区分。以上面的 $|0\rangle$ 态和 $|\psi\rangle$ 态为例，假设一个量子比特处在 $|0\rangle$ 态和 $|\psi\rangle$ 态中的某一个，而且这两个态不同并且也不正交，那由上面的公式 (2.15) 就可以知道，p_0 不等于 0(当然也不等于 1)，因此当你测量这个量子比特的值并且得到 0 的时候 (注意你只有一个量子比特)，你就无从判断你的量子比特原来是处在 $|0\rangle$ 态还是 $|\psi\rangle$ 态，因为 $|\psi\rangle$ 也有 p_0 的概率测得 0。你可能会说：没关系，虽然我只有一个量子比特，但是我可以把测量过程重复多遍，然后看统计结果。但在量子力学中，这是不可能的，因为测量会干扰原来的量子态，会导致态的塌缩，也就是说，无论原来的量子比特是处在 $|0\rangle$ 态还是 $|\psi\rangle$ 态，你只要测到 0，它就一定塌缩到 $|0\rangle$ 态了，以后再重复做测量就没有意义了。总之，你没有 100% 的把握将 $|0\rangle$ 态和 $|\psi\rangle$ 态区分开来。这就说明了不正交的两个量子态一定不能确定地区分，即不能 100% 地区分，因此这也就证明了为什么我们前面说，正交是两个量子态可以确定地区分的充要条件。

公式 (2.15) 是量子力学的核心结论之一，有时候人们也称之为玻恩规则 (因为它和玻恩的波函数的统计解释密切相关)，它当然是被实验验证了的。实际上，光学里的马吕斯定

律就验证了这个结果。由于光的偏振方向也是普通三维空间里的矢量，所以沿 θ 角方向偏振的光子，其偏振态 $|\theta\rangle$ 可以写成，$|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle$。公式 (2.15) 告诉我们，一个沿 θ 方向偏振的光子通过偏振化方向为 x 方向的偏振片的概率为 $\cos^2\theta$，而这正是马吕斯定律。当然，光子除了有线偏振态，还有其它的偏振态，比方说圆偏振态，圆偏振态可以写成 $|\circlearrowleft\rangle = \sqrt{\frac{1}{2}}(|x\rangle + i|y\rangle)$ 和 $|\circlearrowright\rangle = \sqrt{\frac{1}{2}}(|x\rangle - i|y\rangle)$，分别对应左旋光和右旋光。应用公式 (2.15)，人们很容易算出它们通过 x 方向的偏振片的概率。公式 (2.15) 当然可以推广，比方说，一个左旋光子通过一个偏振化方向为 θ 方向的偏振片的概率就是 $|\langle\theta|\circlearrowleft\rangle|^2$。

类似的，假定某个量子系统处在某个 $|\psi\rangle$ 态 (当然，我们默认 $|\psi\rangle$ 已经归一化了)，假定 $|\psi\rangle$ 是某个物理量 A 的叠加态，它可以写成

$$|\psi\rangle = \sum_i \psi_i |i\rangle, \tag{2.16}$$

式中 ψ_i 为复叠加系数，$|i\rangle$ 是物理量 A 的第 i 个本征态 (即是说当系统处于 $|i\rangle$ 态时，物理量 A 有确定取值 λ_i)。假定所有这些本征态 $|i\rangle$ 都已经归一化了，则由于这些本征态两两正交 (2.10)，我们很容易有

$$\langle i|\psi\rangle = \psi_i. \tag{2.17}$$

而测量系统在 $|\psi\rangle$ 态上物理量 A 的值，得到某个 λ_i 的概率 p_i 就是

$$p_i = |\langle i|\psi\rangle|^2. \tag{2.18}$$

很显然，这个式子是公式 (2.15) 的直接推广。而且，我们可以证明 $\langle\psi|\psi\rangle = \sum_i p_i$ (这是由于 $\langle\psi|\psi\rangle = \langle\psi|\left(\sum_i \psi_i|i\rangle\right) = \sum_i \psi_i\langle\psi|i\rangle = \sum_i \psi_i\langle i|\psi\rangle^* = \sum_i |\psi_i|^2 = \sum_i p_i$)，所以量子态的归一化 $\langle\psi|\psi\rangle = 1$ 其实就是要求测量的总概率等于 1。

不确定原理

下面，我们对量子比特的电子自旋态实现稍微展开一点讨论。在这种实现中，量子比特的 $|0\rangle$ 态对应电子的自旋向下态 $|\downarrow\rangle$, $|1\rangle$ 态对应电子的自旋向上态 $|\uparrow\rangle$。按照我们之前的讨论，电子在 $|\uparrow\rangle$ 态，其自旋 z 分量 S_z 有确定的值 $\hbar/2$, 而在 $|\downarrow\rangle$ 态，S_z 为 $-\hbar/2$，但是在下面的 $|\rightarrow\rangle$ 态和 $|\leftarrow\rangle$ 态，S_z 的值没有定义无从确定，

$$|\rightarrow\rangle = \sqrt{\frac{1}{2}}(|\uparrow\rangle + |\downarrow\rangle), |\leftarrow\rangle = \sqrt{\frac{1}{2}}(|\uparrow\rangle - |\downarrow\rangle). \tag{2.19}$$

如果对 $|\rightarrow\rangle$ 态或 $|\leftarrow\rangle$ 态的电子的 S_z 进行测量，则我们有 1/2 的概率得到 $\hbar/2$, 也有 1/2 的概率得到 $-\hbar/2$, 总之，即使进行测量，$|\rightarrow\rangle$ 态和 $|\leftarrow\rangle$ 态的 S_z 也是不确定的。但是，可以证明 $|\rightarrow\rangle$ 态的自旋 x 分量 S_x 的值却是确定的，它是 $\hbar/2$，类似的，$|\leftarrow\rangle$ 态的 S_x 为 $-\hbar/2$，因此这两个态都是物理量 S_x 的本征态。反过来，由于

$$|\uparrow\rangle = \sqrt{\frac{1}{2}}(|\rightarrow\rangle + |\leftarrow\rangle), |\downarrow\rangle = \sqrt{\frac{1}{2}}(|\rightarrow\rangle - |\leftarrow\rangle), \tag{2.20}$$

所以对于 S_x 来说，$|\uparrow\rangle$ 态和 $|\downarrow\rangle$ 态反而是相干叠加态，因此它们的 S_x 反而是不确定的。这也就是说，电子自旋的 z 分量 S_z 和 x 分量 S_x 不能同时确定！这与经典物理完全不同，在经典物理中，物体角动量的 x 分量和 z 分量当然同时有确定的值，在经典物理中，当我们说一个物体的角动量是多少多少的时候，我们实际上就同时给出了它的 x 分量和 z 分量。但我们已经看到了，在量子力学中，这通常是办不到的。类似的，在量子力学中，粒子的坐标和动量也不能同时确定。人们通常把这样的结果叫做量子力学的不确定原理，它是量子力学区别于经典力学的一个核心本质。

根据不确定原理，在量子力学中，人们不能谈论类似于电子自旋 S_x, S_z 同时取 $\hbar/2$ 的概率 $p(S_x = \hbar/2, S_z = \hbar/2)$ 这样的量。注意，我们不是说 $p(S_x = \hbar/2, S_z = \hbar/2) = 0$，而是说讨论这个量是没有意义的，在量子力学中，它根本无法定义！因为 S_x 和 S_z 同时取值就意味着这两者可以同时确定，而这是不可能的，根据我们前面的讨论，当 S_x 取确定值时，电子必定处在 S_z 的叠加态 (公式 (2.19))，因此 S_z 的值没有定义，除非你测量 S_z，但只要你测量成功，原来的态就塌缩了到 S_z 的本征态了，而 S_z 的本征态反而是 S_x 的叠加态 (公式 (2.20))，因此这时候 S_x 的值反而就没有定义了。总之，S_x 和 S_z 的值是不可能同时有定义的。因此，在量子力学中，$p(S_x, S_z)$ 这样的量是没有意义的，它无法定义，不能讨论！但是，$p(S_z = \hbar/2 | S_x = \hbar/2)$ 这样的条件概率却是定义良好的，根据条件概率的标准含义，$p(S_z = \hbar/2 | S_x = \hbar/2)$ 表示已知电子自旋 S_x 为 $\hbar/2$ 时 (因此这时候电子必定处在 $|\rightarrow\rangle$ 态) 测得电子自旋 S_z 取值为 $\hbar/2$ 的概率，根据 (2.19) 式，结果就是 $1/2$，即 $p(S_z = \hbar/2 | S_x = \hbar/2) = 1/2$。这就是量子力学的概率与经典概率的本质区别，在经典概率里面，$p(S_x, S_z)$ 和 $p(S_z | S_x)$ 总是同时有定义的，实际上，经典概率的公理告诉我们 $p(S_x, S_z) = p(S_z | S_x) p(S_x)$，但在量子力学里这样的公理是不成立的。当然，如果你考察的都是一些经典事件，而不是量子物理量取值的概率，那经典概率当然就是完全适用的。

根据完备性，电子自旋的任意一个量子态一定可以写成 $|\psi\rangle = c_\downarrow |\downarrow\rangle + c_\uparrow |\uparrow\rangle$ 的形式。但是，由公式 (2.20) 我们可以知道，$|\psi\rangle$ 也可以重写成 $|\psi\rangle = c_\leftarrow |\leftarrow\rangle + c_\rightarrow |\rightarrow\rangle$ 的形式。所以，一个态可以表示成不同物理量的本征态的不同叠加，量子态的表示形式不是唯一的。在数学上，这种现象我们在线性代数里早就熟悉了，从数学角度来看，$\{|\downarrow\rangle, |\uparrow\rangle\}$ 不过是电子自旋态的希尔伯特空间这样一个线性空间里的一组矢量基，而 $\{|\leftarrow\rangle, |\rightarrow\rangle\}$ 则是另一组矢量基，线性空间里的任意一个矢量 $|\psi\rangle$ 当然可以在不同的矢量基下进行展开。从一组矢量基变换到另一组矢量基的变换过程，物理学家常常称之为表象变换。

关于一个量子比特，我们就讨论到这里，我们讨论一个量子比特主要是想用它来讲清楚量子力学的基本原理。因此上面的所有核心结论都可以推广到一般情况。

2.1.3 多个量子比特

但是，量子计算机当然不只一个量子比特，多个量子比特该怎么处理呢？以两个量子比特为例，很显然，这时候量子比特的值有四种可能性 00、01、10、11，分别对应 $|00\rangle$ 态、$|01\rangle$ 态、$|10\rangle$ 态和 $|11\rangle$ 态，有时候人们也把这些态写成 $|00\rangle = |0\rangle_1 |0\rangle_2$ 这样的形式，表示是 1、2 两个量子比特的量子态乘起来。态叠加原理告诉我们，对于两个量子比特的量子

系统，其任意量子态 $|\psi\rangle$ 可以写成

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle, \tag{2.21}$$

所有这样的量子态的集合就构成了两量子比特系统的希尔伯特空间 $\mathscr{H}_{2^2}$。因此，一个量子比特的希尔伯特空间有两个基矢量 $|0\rangle$ 和 $|1\rangle$，而两个量子比特的系统的希尔伯特空间有四个基矢量，所以记作 $\mathscr{H}_{2^2}$，表示它是一个 $2^2 = 4$ 维的线性空间。类似这样的结论很容易推广到 n 个量子比特，这时候系统的可能值就是从 $x = 000...00$ 到 $x = 111...11$ 的所有 n 位二进制数，而系统的任意一个量子态 $|\psi\rangle$ 必定可以写成

$$|\psi\rangle = \sum_{x=000...00}^{111...11} \psi_x |x\rangle. \tag{2.22}$$

这里我们已经改用 ψ_x 来表示叠加系数了。当然，所有的基矢量 $|x\rangle$(注意，这里 $|x\rangle$ 不是表示 x 方向的偏振态) 都是归一化的，而且不同的基矢量之间是两两正交的 (因为它们都可以确定地区分)，因此我们容易得到 $\psi_x = \langle x|\psi\rangle$，而测量这个 n 量子比特系统的值 (在所谓的计算基下) 得到一个 n 位 2 进制数 x 的概率 p_x 为

$$p_x = |\langle x|\psi\rangle|^2. \tag{2.23}$$

如果我们离开量子比特系统，回到单个微观粒子系统，那么上一段中的 ψ_x 就可以对应到单个微观粒子的波函数 $\psi(x)$，而 $|x\rangle$ 就对应于单个微观粒子处于 x 位置的量子态，因此我们刚才的结果就可以重写成

$$\psi(x) = \langle x|\psi\rangle, \tag{2.24}$$

这就是波函数和狄拉克记号之间的精确联系。而根据公式 (2.18), $|\langle x|\psi\rangle|^2 = |\psi(x)|^2$ 就是当微观粒子处在 $|\psi\rangle$ 态时我们在 x 位置上测到它的概率 (即测得它的位置物理量是 x 的概率)，当然由于这时候 x 是连续变量，所以更准确的说法是概率密度，而这就是玻恩对波函数的统计解释。因此，玻恩的统计解释其实是关于量子态的测量原理的一个特殊情况。

回到我们的 n 个量子比特的系统。公式 (2.22) 告诉我们，这一系统的希尔伯特空间有 2^n 个基矢量，因此是 2^n 维线性空间，通常记作 $\mathscr{H}_{2^n}$。总之，量子计算机的希尔伯特空间总是有限维的。但是，通常的微观粒子，比方说氢原子，由于有无穷多个能级，因此它的希尔伯特空间往往是无穷维的。这实际上就是模拟量和数字量之间的区别，模拟量由于可以连续取值，所以是无穷多的，而数字量只可能有限多。但是，我们处理任何问题其实都只需要达到一定的精度即可，因此在精度范围之内我们总可以将模拟量数字化。同样的道理，在一定的精度范围之内我们也可以用量子计算机来处理任何微观粒子系统，比方说对于氢原子，很多时候我们其实可以忽略主量子数 $n \to \infty$ 的高能级，这时候我们处理的实际上就是一个有限维的量子系统。而且，基本物理定律也告诉我们，我们对空间和时间的测量也有一个最小的精度，即普朗克长度和普朗克时间，因此没准我们的世界本来就是“数字化”的，也许我们的世界本来就是一台量子计算机。当然，这些都还只是人们的猜测，有可能对，但错的可能性也同样大。

我们已经看到，n 个量子比特能构建一个 2^n 维的希尔伯特空间 $\mathscr{H}_{2^n}$，也就是一个 2^n 维的复线性空间，其中的每一个态矢量对应 2^n 个复的叠加系数。因此，如果用经典计算机来模拟这样的系统的话，由于每一个复系数都需要若干个经典比特才能刻画，因此总共需要的经典比特数将是 2^n 的某个倍数，也就是说，是随着 n 指数增长的。这也就是说，只要量子比特的数目够多 (也不需要太多，100 左右足够)，则相应的量子系统就无法用经典计算机来模拟了 (毕竟 2^{100} 可是个天文数字)。但是，对于未来的量子计算机来说，100 个量子比特当然将会是家常便饭，这也就是量子计算机比经典计算机强大的一个基本原因。

量子态的希尔伯特空间是一个抽象的空间，它和我们生存的时空没有关系 (而且它很可能比时空还要更基本)。这个看起来平凡的论断有时候能导出一些神奇的结论，比如说按照这个量子力学的逻辑，在我们的宇宙大爆炸产生时间和空间之前，在时间的概念可能还不成立的时候，宇宙的量子态应该就已经存在了，你可以把它记作 $|U\rangle$，但是 $|U\rangle$ 是什么却还没有人真正搞明白。

再举一个和我们关系更密切而且实验上早就实现了的神奇例子。这个例子只涉及到两个量子比特，我们知道，按照态叠加原理，下面的量子态 $|\Phi^+\rangle$ 是两量子比特系统的一个可能态，

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \tag{2.25}$$

$|\Phi^+\rangle$ 态的神奇之处就在于，它是一个整体，它不能因式分解成第一个量子比特的某个态 $|\phi_1\rangle$ 和第二个量子比特的某个态 $|\phi_2\rangle$ 的乘积 $|\phi_1\rangle|\phi_2\rangle$ 这样的形式 (读者不妨试着分解一下)。由于量子态和时空没有直接关系，这也即是说，这两个量子比特即使一个在南昌，另一个远在天边，它们也还是一个整体，它们整体性地处在 $|\Phi^+\rangle$ 态中，而不是南昌的量子比特处在某个 $|\phi_1\rangle$，天边的量子比特处在某个 $|\phi_2\rangle$。而在量子力学中，类似这样的量子态有很多，它们叫做量子纠缠态。量子纠缠态不但在量子信息和量子计算机技术中处于核心地位，而且近年来人们发现，时空本身可能就起源于量子纠缠。

2.1.4 量子不可克隆定理与幺正性

科幻小说里经常有这样的情节：一个人进入一个透明的玻璃罩里面，然后出现一束“光”，它从头到脚缓缓地扫过这个人全身。同时，在一个遥远的星球也有一个类似的玻璃罩子，当地球上的这个人被扫描以后，遥远星球的这个玻璃罩子里面也出现一束“光”，随着这束光从下扫向上，玻璃罩里面先是出现一双脚，接着是身子，最后是脑袋... 最后出现了一个和地球上的那个人一模一样的人，地球上的这个人被复制到这颗遥远的星球了！

这可能吗？我问的是，基本物理原理允许这样的事情吗？回答是，如果是复制，也就是说，最后有两个一模一样的人，那这事不可能。但如果是传送，也就是说，从头到尾只有一个人，只不过这个人从地球传送到了遥远的星球，那这是可能的。我们这里将仅仅解释前一种情况为什么不可能，至于后一种情况为什么可能我们将在后面的小节中进一步讨论。

有人说，复制为什么就不可能呢？我把这个人的每一个细胞、每一个分子甚至每一个原子的所有信息都测量出来，然后按照这些信息在遥远的星球上把这个人复制一份，这在

原理上有什么不可能的呢。但是别忘了，量子力学有一个不确定原理，它限制了你的测量能够收集到的信息，比方说你不能同时确定一个电子两个不同分量的自旋，你也不能同时确定一个原子的位置和动量。测量会导致态的塌缩，我们永远也无法通过测量得到关于这个人的量子态的完整信息。

那你可能会想，不测量行不行呢？量子力学对测量能够收集的信息有基本的限制，但如果不用测量呢？如果有一个量子演化过程在不经过测量的情况下就把这个人的量子态 $|\psi\rangle$ 演化到了 $|\psi\rangle|\psi\rangle$，那这不就成功地复制这个人的量子态了吗？

然而，这也是不可能的。简单的解释是，因为量子力学所满足的最基本的原理态叠加原理是一个线性叠加原理，而从 $|\psi\rangle$ 到 $|\psi\rangle|\psi\rangle$ 相当于把 $|\psi\rangle$ 做了一个平方，而平方是一个非线性操作，因此是和态叠加原理相矛盾的，因此不可能。这个简单的解释实际上就构成了量子不可克隆定理的实质。下面我们以两个量子比特的系统为例来将这个解释精确化。

设想有两个量子比特，一个原来处在 $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ 态，另一个原来处在 $|0\rangle$ 态，所谓的一个复制操作 $\mathscr{C}$，即是这个两量子比特系统的一个演化过程，它使得

$$|\psi,0\rangle \to \mathscr{C}(|\psi,0\rangle) = |\psi,\psi\rangle = |\psi\rangle|\psi\rangle. \tag{2.26}$$

即在复制演化之下，第一个量子比特的量子态要复制到第二个量子比特上。由于 $|\psi\rangle$ 是任意的，因此如果存在这样的 $\mathscr{C}$，那就必然有 $\mathscr{C}(|00\rangle) = |00\rangle = |0\rangle|0\rangle$，$\mathscr{C}(|10\rangle) = |11\rangle = |1\rangle|1\rangle$。另一方面，态叠加原理要求系统的任何演化都要保持线性叠加性，因此 $\mathscr{C}$ 当然也要保持线性叠加性，但这就意味着 $\mathscr{C}(|\psi,0\rangle) = c_0\mathscr{C}(|00\rangle) + c_1\mathscr{C}(|10\rangle) = c_0|00\rangle + c_1|11\rangle$。而克隆方程要求的是 $\mathscr{C}(|\psi,0\rangle) = |\psi\rangle|\psi\rangle = c_0^2|00\rangle + c_0c_1(|01\rangle + |10\rangle) + c_1^2|11\rangle$。显然，除非 c_0, c_1 中一个系数取 0 另一个系数取 1，否则克隆方程一定是无法成立的。这也就是说，克隆任意未知量子态是不可能的！但是当然，我们可以将 $|0\rangle$ 态和 $|1\rangle$ 态这两个正交态分别克隆一份。正交态可以克隆实际上就是我们可以把一份文件这样的经典信息克隆一份的根本原因，因为所谓的经典信息其实就是损失了量子相干叠加性的量子信息，经典信息当然可以确定地区分，因此用量子态来描述的话，这些信息当然就是正交的。

实际上，我们也可以由不正交的两个量子态一定不可确定地区分这一结论导出量子不可克隆定理。因为假如任意量子态都可以克隆的话，对于一个量子比特，我们就可以考虑它的这样两个量子态，一个是 $|0\rangle$ 态，另一个是 $|0\rangle$ 态和 $|1\rangle$ 态的相干叠加态 $|\psi\rangle$，这两个态当然是不正交的，实际上由公式 (2.15) 可以知道，$0 \neq p_0 = |\langle 0|\psi\rangle|^2 < 1$。如果量子不可克隆定理不成立的话，我们就可以把 $|\psi\rangle$ 和 $|0\rangle$ 都克隆任意多份，也就是说我们可以制造下面两个态

$$|\mathbf{0}\rangle = |0\rangle|0\rangle...|0\rangle, |\Psi\rangle = |\psi\rangle|\psi\rangle...|\psi\rangle, \tag{2.27}$$

式中的省略号表示 N 份连乘。如此一来，$|\langle\mathbf{0}|\Psi\rangle|^2 = |\langle 0|\psi\rangle|^{2N} = p_0^N$。由于 $p_0 < 1$，因此只要 N 足够大，$|\langle\mathbf{0}|\Psi\rangle|^2$ 就可以无限接近于 0，因此在 $N \to \infty$ 的极限下，$|\mathbf{0}\rangle$ 就和 $|\Psi\rangle$ 正交，因此这两个态就可以确定地区分。但是，$|\mathbf{0}\rangle$ 和 $|\Psi\rangle$ 不过是 $|0\rangle$ 和 $|\psi\rangle$ 的复制，因此这就意味着我们可以通过确定地区分 $|\mathbf{0}\rangle$ 和 $|\Psi\rangle$ 来 100% 地将 $|0\rangle$ 和 $|\psi\rangle$ 区分开来，而这就与不正交的两个量子态不可确定地区分相矛盾。因此，只要不正交的两个量子态不可确定地区分，那么量子不可克隆定理就必须成立。

幺正性

前面我们提到，量子态的时间演化一定要保持线性叠加原理 (态叠加原理)。实际上，这是量子力学基本原理施加给量子态演化的限制之一，量子态的演化必须满足的另一条限制是，任意两个可以确定地区分的量子态在演化之下必须始终保持可以确定地区分。由于可确定地区分的充要条件是两个量子态正交，所以这实际上就是要求，任意两个相互正交的量子态在时间演化下要保持正交。当然，由于一个态和它自身的内积是总概率，在物理上应该始终归一化为 1。因此概况起来说，第二条限制实际上是，量子态的演化应该始终保持态的正交归一性。

满足这两个要求的量子态演化就叫做幺正演化，而这一性质就叫做量子力学的幺正性。其实，之所以量子态的演化要满足薛定谔方程，就是因为薛定谔方程是满足幺正性的。下面我们就来证明这一点。

先让我们回想一下，对于单粒子情形，所谓的薛定谔方程就是形如 $i\hbar\partial_t\psi(x,t)=H\psi(x,t)$ 这样的波动方程，H 就是所谓的哈密顿算符。这当然是波函数表达下的写法，如果用更一般性的狄拉克记号，那么薛定谔方程就可以写成

$$i\hbar\partial_t|\psi(t)\rangle = H|\psi\rangle. \tag{2.28}$$

前文已经说过，所谓的态矢量 $|\psi\rangle$，其实可以简单地想象为一个列矢量，而哈密顿算符是乘在这个列矢量前面的，所以应该想象为一个矩阵。实际上，这个矩阵是一个厄米矩阵，即满足 $H^\dagger = H$。很显然，薛定谔方程 (4.83) 是一个线性方程，所以当然是满足态叠加原理的。另一方面，假如我们将方程 (4.83) 两边进行共轭转置，我们就可以得到

$$-i\hbar\partial_t\langle\phi(t)| = \langle\phi|H, \tag{2.29}$$

这里我们已经将 $|\psi\rangle$ 换成了 $|\phi\rangle$，这个等式的右边显然是一个行矢量乘以矩阵 H。由方程 (4.83) 和方程 (2.29)，我们很容易有

$$i\hbar\partial_t\left(\langle\phi(t)|\psi(t)\rangle\right) = i\hbar\left(\partial_t\langle\phi(t)|\right)|\psi(t)\rangle + i\hbar\langle\phi(t)|\partial_t\left(|\psi(t)\rangle\right) \tag{2.30}$$

$$= -\langle\phi|H|\psi\rangle + \langle\phi|H|\psi\rangle = 0. \tag{2.31}$$

这就意味着两个态的内积 $\langle\phi|\psi\rangle$ 在时间演化之下是不变的，如果将这两个态取成正交的两个态，那这个结果就意味着薛定谔方程是保持正交性的。而如果将 $|\phi\rangle$ 取成和 $|\psi\rangle$ 一样，那内积不变的这个结果也叫做总概率守恒，因为这时候 $\langle\psi|\psi\rangle$ 就是粒子在全空间出现的总概率。这就完成了我们对幺正性的证明。回顾这个证明过程，我们会发现薛定谔方程里出现的虚数单位 i 有实质性的作用，如果没有这个 i，那薛定谔方程是无法满足幺正性的。

2.1.5 薛定谔的猫

至此，我们已经讨论了所有量子力学最基本最核心的原理。你可能觉得这些原理虽然不容易理解，但却并非不可接受，毕竟我们通常总是觉得量子力学原理是适用于微观世界的，和我们的宏观世界没有什么关系，因此即使它奇怪一点也没有什么不可接受的。但是，

本文一开始就已经说过了，量子力学并非仅适用于微观世界，不，作为最基本的物理学规律，它甚至适用于全宇宙，更不要说我们日常的宏观世界了。

为了反映量子力学原理在宏观世界中有多么不可思议，薛定谔提出了一个著名的思想实验，那就是物理学四大神兽之一的薛定谔的猫。这只著名的神兽大致是这么一回事：在一个与世隔绝的密闭盒子里放一只猫，一个放射性的原子，这个原子衰变的概率为 1/2。再放一个粒子探测器，一把锤子，一瓶剧毒的气体。如果原子衰变了，那衰变放出来的粒子就会被粒子探测器探测到，然后粒子探测器就会发出信号让锤子掉下来，砸破毒气瓶子，毒气就会跑出来将猫毒死。用量子力学的语言来说即是，这时候猫会处在 $|Dead\rangle$ 态。但是，如果原子没有衰变，那刚才描述的一切都不会发生，猫将活得好好的，也就是说，它将处于 $|Live\rangle$ 态。但问题是，原子衰变是一个量子力学过程，而这个原子只有 1/2 的概率衰变，因此实际上盒子里的放射性原子是处在衰变和不衰变的叠加态。因此盒子里的猫也将处于死和活的叠加态，可以表示为

$$|Cat\rangle = \frac{1}{\sqrt{2}}\left(|Dead\rangle + |Live\rangle\right). \tag{2.32}$$



Figure 2.1: The Schrodinger's Cat. 图片来源：thelifeofpsi.com

人们当然可以把 $|Dead\rangle$ 态和 $|Live\rangle$ 态对应到量子比特的 $|0\rangle$ 态和 $|1\rangle$ 态，这样薛定谔的这只猫就是一个宏观的量子比特。根据我们对量子比特的讨论，叠加态 $|Dead\rangle + |Live\rangle$ 的猫既不是死，也不是活。实际上，根据我们对量子比特的讨论，在我们没有打开盒子看 (看就是一种测量) 的时候，猫的死活是没有意义的，这只处在叠加态的猫超越了生死！这里已经发生我们很难接受的事情了，因为猫完全是一个宏观物体，是我们日常很熟悉的生命，因此根据我们日常的经验，猫要么是死了，要么就是活着，不可能有什么超越生死的猫，根据我们日常生活的经验，生死的概念对于猫总是成立的，生死不可能没有意义。

更奇怪的是，根据我们前面对量子比特的讨论。当我们打开盒子往里看的时候，猫的量子态就塌缩了，我们有 1/2 的概率看到一只活猫，同样也有 1/2 的概率看到一只死猫。请注意，根据我们前面的讨论，猫的死活完全是我们看的结果。的确，我们要么看到猫死了，要么看到猫活着，只要我们打开盒子看了，那么猫就或者是死或者是活，再没有第三

种可能性。但，只要我们不看，那猫就没有死活，生死的概念就没有意义。爱因斯坦完全接受不了这个，他说：难道我们不看月亮，月亮就不存在吗？对，我们不看猫，猫就没有死活吗？

这些都令人很难以接受，不可思议！那，这是不是说明量子力学对于猫这样的宏观物体不适用呢？不是的，虽然我们还没有实现让一只猫处在相干叠加态，但是，让一个宏观物体 (比方说一个宏观的场) 处在相干叠加态，甚至让一个生命体 (细菌) 处于相干叠加态，都是实验室中已经实现了的事情。从这个意义上来说，薛定谔的猫今天已经不再是一个思想实验了，而是实验室中真正的实验。而实验的结果是发现，量子力学同样适用于宏观客体，薛定谔的猫和前面我们讨论过的量子比特满足一样的量子力学规律。不看猫，猫就是没有死活超越了生死，猫的死活的确就是因为我们打开盒子看了！作为一个物理性质的两种不同取值，生死是两种可以确定区分的状态，但是一只保持了量子相干性的猫可以超越生死。

那为什么我们日常生活中的猫不是这样呢？为什么我们日常生活中的猫不能保持量子相干叠加性呢？这就涉及到薛定谔的猫的一个隐含假设，他假设一个与世隔绝的盒子，这里所谓的与世隔绝，就是与一切环境隔绝，也就是说，外界的一切，光声热电，全都不能透过这个盒子，自然界的四大相互作用力全都不能透过这个盒子。极端一点来说，这需要把这个盒子孤立在宇宙之外。然而，我们日常生活中的猫总是处在环境之中的，作为一种宏观物体，猫有大量的原子分子，这些原子分子和整个环境有非常复杂的相互作用。因此虽然量子力学同样适用于日常生活中的猫，但这时候猫的量子态和环境的量子态纠缠在一起了，它们形成了一个整体的量子态，而猫原来的量子态的信息现在就会由这个整体来承载。而我们无法将整个环境 (甚至整个宇宙) 的量子信息都收集起来，当我们忽视环境的信息，而仅仅关心猫的时候，我们实际上就忽视掉了猫和环境的整体的大部分信息，因此也就无法重现猫原来的量子态了，我们将丢失掉原来猫态的相干叠加信息。这时候，我们就会发现猫的量子相干性消失了，猫要么是生，要么是死，而不能是生和死的量子叠加。这样的一只猫就变成了我们通常所熟悉的，满足经典物理学定律的猫。这样一个从量子到经典的过程就叫做退相干，正是退相干使得我们的宏观世界由量子力学描写的世界变成了一个由经典物理描写的世界。

这也说明了为什么微观世界常常需要量子力学规律，而不能使用经典物理规律。原因就在于，微观物体的德布罗意波长相对比较长，它常常超过我们关心的微观尺寸，这样一来在这样的微观尺寸之内系统就能保持量子相干性，因此就必须用量子力学来描述。但即使这样，退相干在我们对电子自旋这样的量子系统作测量的时候也同样要起作用。实际上，当我们测量一个量子系统的时候，我们的测量仪器就是量子系统所处环境的一部分，测量过程就是使得量子系统由相干叠加态退相干，进而坍缩到本征态的过程。

这也就是量子计算机为什么这么难实现的基本原因，因为为了造出一台实用的量子计算机，我们就必须在一个很大的尺寸上保持量子相干性。而尺寸越大，量子比特的数目越多，它和环境的耦合就越复杂，量子态的信息就越容易泄露到环境中去，因此系统也就越容易退相干。制造量子计算机其实就是要千方百计地对抗退相干，让整个量子计算机处在某种薛定谔的猫态。极端简化地说，造量子计算机就相当是要造出一只薛定谔的猫。

2.1.6 习题

1. 对于一个量子比特，请将量子态 $|\psi\rangle = |0\rangle + e^{i\alpha}|1\rangle$ 和量子态 $|\phi\rangle = e^{i\beta}|0\rangle + 2|1\rangle$ 分别归一化 (式中 α,β) 为两个实数，然后请计算内积 $\langle\phi|\psi\rangle$。

2. 对于一个量子比特，请验证量子态 $|+\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)$、$|-\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big)$ 正交归一。

3. 当电子处在量子态 $|\psi\rangle = 2|\uparrow\rangle + e^{i\alpha}|\downarrow\rangle$ 上时，请问测到它的 $S_z = -\hbar/2$ 的概率是多少?

2.2 算符与物理量

2.2.1 基础态

我们已经知道：某物理量有确定取值的量子态称之为这个物理量的本征态，相应的物理量的值称之为本征值。由于可以根据物理量值的不同将本征态确定地区分开来，所以任何一个物理量的两个不同本征值的本征态必定正交。但是，当系统处在不同本征态的叠加态上时，其相应物理量的值是无从确定的，这时候如果不进行测量，这个物理量的值就没有定义。

在量子力学中，不是所有物理量的值都可以同时确定，比如电子自旋的 x 分量 S_x 和 z 分量 S_z 就不能同时确定。但是，我们可以根据所有可以同时确定的物理量的值的不同，来形成一组可以确定地区分的量子态，记作 $\{|i\rangle, i = 1,2,3....\}$($|i\rangle$ 只是一个记号，表示第 i 个态，当然你也可以改成 $|\psi_i\rangle$ 之类的记号)。比方说，对于一个三维微观粒子，这组态可以是 $\{|x,y,z\rangle\}$，它表示粒子处在位置 $\mathbf{x} = (x,y,z)$ 的量子态，相应的可以用来区分这组物理态的物理量就是粒子的三个坐标分量 X,Y,Z。也即是说，在这组量子态 $\{|i\rangle, i = 1,2,3....\}$ 中，任何两个不同态总有某些物理量的值是不同的，因此这组态中的任何两个都可以确定地区分，也即是满足正交归一性，

$$\langle i|j\rangle = \delta_{ij}, \tag{2.33}$$

这里 δ_{ij} 的定义是，当 i,j 相等时取 1，当 i,j 不同时取 0。

对于系统的任何一个量子态 $|\psi\rangle$，我们的任何一次测量总能得到这组物理量的某一组值，而这组值唯一地对应某一个本征态 $|i\rangle$，因此，任何量子态 $|\psi\rangle$ 必然都可以写成 $\{|i\rangle, i = 1,2,3....\}$ 的某个线性叠加，

$$|\psi\rangle = \sum_i \psi_i |i\rangle, \tag{2.34}$$

这里 ψ_i 表示叠加系数。这个结果称之为 $\{|i\rangle, i = 1,2,3....\}$ 的完备性。同时满足正交归一性和完备性的一组态矢量就称之为希尔伯特空间的一组正交矢量基，相应的基矢量称作基础态。另外，从正交归一公式 (2.33) 我们容易得到 (2.34) 式中的叠加系数 ψ_i 总是可以表示成

$$\psi_i = \langle i|\psi\rangle. \tag{2.35}$$

由于 ψ_i 就是列矢量 $|\psi\rangle$ 在矢量基 $\{|i\rangle, i=1,2,3....\}$ 中的第 i 个分量，因此公式 (2.35) 告诉我们，任意态矢量 $|\psi\rangle$ 在基 $\{|i\rangle, i=1,2,3....\}$ 中的第 i 个分量都可以方便地由 $\langle i|\psi\rangle$ 求出，反过来，知道了所有的分量 $\langle i|\psi\rangle$，那态矢量本身也就确定了。

量子力学的基本原理告诉我们，假如在本征态 $|i\rangle$ 上某个物理量 $\mathscr{O}$ 的本征值为 λ_i，则测量 $|\psi\rangle$ 态的系统，得到物理量 $\mathscr{O}$ 的值为 λ_i 的概率 p_i 等于

$$p_i = |\langle i|\psi\rangle|^2. \tag{2.36}$$

而一旦我们测到了 λ_i，原来的量子态就必然塌缩到相应的本征态 $|i\rangle$。

2.2.2 线性算符

这一小节我们来讨论一下希尔伯特空间上的线性算符，之所以只考虑线性算符，是因为态叠加原理告诉我们，希尔伯特空间是一个线性空间，在线性空间上最自然的算符就是线性算符。这一节的内容完全是数学，是为了后面的量子力学应用作准备的。

所谓希尔伯特空间上的一个线性算符 A, 就是某个对量子态的操作，这个操作把任意一个量子态 $|\psi\rangle$ 变换成某个新的量子态 $|\phi\rangle$，记作 $A|\psi\rangle = |\phi\rangle$。而且这个操作应该保持态矢量之间的线性叠加关系。可见，线性算符其实就是通常线性代数里的线性变换。你也可以将算符 A 想象成某个“设备”，经过这个“设备”的作用以后系统的量子态就从 $|\psi\rangle$ 变成了 $|\phi\rangle$。不妨举一个线性算符的例子，比方说给定两个量子态 $|u\rangle, |v\rangle$，我们可以构造一个线性算符 T，它的定义如下，

$$T = |u\rangle\langle v|. \tag{2.37}$$

注意，这里列矢量 $|u\rangle$ 是在左边，行矢量 $\langle v|$ 在右边，我们是用列矢量乘以行矢量，所以这个 T 可以想象成一个矩阵。正如线性代数里线性变换和矩阵是一一对应的一样，在量子力学里，算符和矩阵也基本上是同一回事。T 为什么是一个算符呢？原因在于，任给一个态矢量 $|\psi\rangle$，我们可以定义 T 在 $|\psi\rangle$ 上的作用为

$$T|\psi\rangle = |u\rangle\langle v|\psi\rangle. \tag{2.38}$$

注意这里 $\langle v|\psi\rangle$ 是一个数，所以等式 (4.27) 的右边其实就是一个正比于 $|u\rangle$ 的态矢量，因此 T 的确将 $|\psi\rangle$ 变换到了另一个态矢量，它的确是一个算符。而且由于两个态矢量的内积对于右边的列矢量来说是线性的，所以很容易验证 T 的确保持了态矢量的线性叠加关系，因此是一个线性算符。

实际上，有一个数学定理 (奇异值分解定理) 说，希尔伯特空间的任何一个线性算符 A 必定能写成如下形式

$$A = \sum_i \lambda_i |u_i\rangle\langle v_i|. \tag{2.39}$$

这里 $\lambda_i \neq 0$，$\{|u_i\rangle, i=1,2,3...\}$ 和 $\{|v_i\rangle, i=1,2,3...\}$ 分别从属于希尔伯特空间上的两组正交归一矢量基。复数 λ_i 通常称作线性算符 A 的奇异值。从这里也能看出来，希尔伯特空间的线性算符总是能分解成列矢量乘以行矢量的形式，因此任何这样的算符都能想象成是

一个矩阵。后面我们会进一步给出算符和矩阵之间的更精确的联系。为了方便读者参考，在本文的附录中，我们给出了奇异值分解定理的数学证明。

由于 $A|\psi\rangle$ 结果仍然是一个态矢量，因此我们可以将它和另一个态矢量 $|\phi\rangle$ 作内积，得到如下表达式

$$\langle\phi|A|\psi\rangle. \tag{2.40}$$

这个表达式当然是一个数，实际上它可以看成是行矢量 $\langle\phi|$ 乘以矩阵 A 然后再乘以列矢量 $|\psi\rangle$, 其结果当然是一个复数。现在我们将这个表达式进行共轭转置，所谓共轭转置就是首先将一切都复数共轭，然后再转置，因此共轭转置的规则和转置的规则是类似的，都满足 $(AB)^\dagger = B^\dagger A^\dagger$，$\dagger$ 就表示共轭转置，有时候也叫厄米共轭。利用共轭转置的规则，$\big(\langle\phi|A|\psi\rangle\big)^\dagger = |\psi\rangle^\dagger A^\dagger \langle\phi|^\dagger = \langle\psi|A^\dagger|\phi\rangle$，因此 $\langle\phi|A|\psi\rangle$ 共轭转置以后就是 $\langle\psi|A^\dagger|\phi\rangle$。但是由于 $\langle\phi|A|\psi\rangle$ 是一个复数，而一个复数的共轭转置当然其实就只是复共轭，因此我们就有

$$\langle\phi|A|\psi\rangle^* = \langle\psi|A^\dagger|\phi\rangle. \tag{2.41}$$

这里出现的算符 A 的共轭转置 $A^\dagger$ 就叫做算符 A 的厄米共轭算符。公式 (3.2) 有时候也被看作是厄米共轭算符的定义式。如果我们将算符 A 写成奇异值分解的形式 (2.39), 那我们也很容易利用共轭转置的规则得到 $A^\dagger$ 的奇异值分解

$$A^\dagger = \sum_i \lambda_i^* |v_i\rangle\langle u_i|. \tag{2.42}$$

两个线性算符当然可以相加，算符也可以相乘。由于算符可以想象为矩阵，所以线性算符的加法与乘法就可以理解为矩阵的加法与乘法。比方说，假设有两个线性算符，T_1, T_2，$T_1 = |u_1\rangle\langle v_1|$, $T_2 = |u_2\rangle\langle v_2|$, 那么 $T_1T_2 = |u_1\rangle\langle v_1|u_2\rangle\langle v_2| = \langle v_1|u_2\rangle|u_1\rangle\langle v_2|$(我们利用了乘法结合律，且在最后一个等号中我们把复数 $\langle v_1|u_2\rangle$ 提到了式子的前面), 结果显然也是一个线性算符。还是这两个算符，但如果你要算的是 T_2T_1，则结果就是 $T_2T_1 = |u_2\rangle\langle v_2|u_1\rangle\langle v_1| = \langle v_2|u_1\rangle|u_2\rangle\langle v_1|$。很显然，一般来说 $T_2T_1 \neq T_1T_2$，也即是说，算符的乘法通常是不满足交换律的。这其实是由于算符可以当作矩阵来理解，而矩阵的乘法一般来说当然是不可交换的。

如果希尔伯特空间的某个线性算符的集合包含恒等算符 (即将任何态矢量都变换到它本身的恒等操作) 以及它的常数倍，而且这个集合在算符加法、乘法、以及厄米共轭之下是封闭的，那这个线性算符的集合就定义了一个算符代数，数学家也称之为冯诺依曼代数。比如说单个量子比特系统的冯诺依曼代数可以由算符 $c\cdot 1$(表示给任何态矢量乘上复数 c) 以及下面两个算符 Z 和 X 生成，

$$\mathrm{Z}|0\rangle = |0\rangle, \mathrm{Z}|1\rangle = -|1\rangle, \mathrm{X}|0\rangle = |1\rangle, \mathrm{X}|1\rangle = |0\rangle. \tag{2.43}$$

这里 $\{|0\rangle, |1\rangle\}$ 组成了量子比特的矢量基，由于线性算符总是保持线性叠加，所以线性算符在任意态上的作用都可以由其在基矢量上的作用决定，正因为如此，给出一个线性算符对所有基矢量的作用其实也就决定了这个线性算符本身。很容易验证，量子比特的 Z 和 X 算符有一些很显著的性质，比如满足 $\mathrm{Z}^2 = \mathrm{X}^2 = 1, \mathrm{ZX} + \mathrm{XZ} = 0$。

如果一个态矢量 $|u_n\rangle$ 满足如下方程

$$A|u_n\rangle = \lambda_n|u_n\rangle, \tag{2.44}$$

则 $|u_n\rangle$ 就称之为算符 A 的一个本征态，相应的 λ_n 就称之为算符 A 的本征值，而这个方程就称作算符 A 的本征方程。算符 A 很可能有多个本征态，因此我们加上下标 n 以示区分，$|u_n\rangle$ 就叫做 A 的第 n 个本征态，λ_n 就叫做 A 的第 n 个本征值。这里的本征态和本征值当然是一个数学概念，但我们前面说过的物理量的本征态和本征值的说法其实就是从这里来的，因为在量子力学中，物理量和算符是密切相关的，至于为什么是这样，我们后文会给予解释。

有两类算符对于量子力学尤其重要。第一类就是所谓的幺正算符，常常记作 U, 它满足

$$UU^\dagger = U^\dagger U = 1, \tag{2.45}$$

这里的 1 表示恒等操作，也叫做单位算符。将幺正算符作用在量子态上也叫作对量子态进行幺正变换。比方说我们有一个幺正算符 U, 还有两个任意的量子态 $|\psi\rangle$ 和 $|\phi\rangle$, 那么 $|\psi'\rangle = U|\psi\rangle$ 就是 $|\psi\rangle$ 幺正变换以后的结果，同样 $|\phi'\rangle = U|\phi\rangle$ 是 $|\phi\rangle$ 幺正变换的结果。幺正变换有一个很重要的性质，那就是变换以后的两个态的内积 $\langle\phi'|\psi'\rangle$ 和变换之前是一样的，即

$$\langle\phi'|\psi'\rangle = \langle\phi|\psi\rangle. \tag{2.46}$$

利用 $|\psi'\rangle = U|\psi\rangle$ 和 $\langle\phi'| = \langle\phi|U^\dagger$(即将等式 $|\phi'\rangle = U|\phi\rangle$ 两边分别共轭转置)，以及幺正算符的定义 (3.4)，人们很容易证明这一结果。这也就是说，希尔伯特空间的内积在幺正变换下总是不变的。

幺正算符的定义 (3.4) 实际上告诉我们算符 U 是可逆的，$U^\dagger$ 就是它的逆算符，即 $U^{-1} = U^\dagger$。因此如果将上面的证明反过来，假设希尔伯特空间的内积在某个可逆算符 U 的变换下保持不变，那刚才的证明过程就会反过来告诉我们 $U^\dagger U = 1$。由于 U 可逆，因此我们将这个等式两边从右方乘以 U^{-1} 就得到，$U^\dagger = U^{-1}$, 但我们又可以将刚得到的等式从左方乘以 U，从而进一步得到 $UU^\dagger = 1$。也就是说，$U^\dagger U = 1$ 和 $UU^\dagger = 1$ 同时成立。按照幺正算符的定义，这就反过来证明了任何这样的保持希尔伯特空间内积不变的可逆变换一定是一个幺正变换。

另一类重要的算符就是所谓的厄米算符，也就是满足 $A^\dagger = A$ 的算符。根据 (3.2) 厄米算符满足

$$\langle\phi|A|\psi\rangle^* = \langle\psi|A|\phi\rangle. \tag{2.47}$$

厄米算符有一些很重要的特性，首先，厄米算符的本征值一定是实数，其次，厄米算符不同本征值的本征态一定正交。为了证明这两个性质，我们记厄米算符 A 的第 i 个本征态为 $|i\rangle$，相应的本征值记为 λ_i。因此，

$$\lambda_j\langle i|j\rangle = \langle i|A|j\rangle = \langle j|A^\dagger|i\rangle^* = \langle j|A|i\rangle^* = \lambda_i^*\langle j|i\rangle^* = \lambda_i^*\langle i|j\rangle, \tag{2.48}$$

这里第一个等号和第四个等号是利用本征态的定义，即 $A|i\rangle = \lambda_i|i\rangle$，第二个等号是利用方程 (3.2)。推导过程 (2.48) 告诉我们 $\lambda_j\langle i|j\rangle = \lambda_i^*\langle i|j\rangle$, 取 $i=j$，就可以知道 $\lambda_i^* = \lambda_i$, 因此厄米算符的本征值必为实数，由此我们进一步得到，$(\lambda_j - \lambda_i)\langle i|j\rangle = 0$, 因此当 $\lambda_i \neq \lambda_j$ 时，必有 $\langle i|j\rangle = 0$, 即相应的两个本征态必定正交。这其实就已经完成了我们需要的证明。但有一个问题是，如果某两个本征态 $|i\rangle, |j\rangle$ 有相同的本征值，即如果 $\lambda_i = \lambda_j$，这时候怎么办，这时候我们就不能证明这两个态正交了。这种情况，我们就叫做这两个本征态简并，如果只有 $|i\rangle, |j\rangle$ 两个本征态相简并，我们就叫做本征值 λ_i 有二重简并 (当然，如果有 N 个本征态都简并，我们就叫做 N 重简并)。碰到这种二重简并的情况，我们注意到

$$A(c_1|i\rangle + c_2|j\rangle) = c_1A|i\rangle + c_2A|j\rangle = \lambda_i c_1|i\rangle + \lambda_j c_2|j\rangle = \lambda_i(c_1|i\rangle + c_2|j\rangle), \tag{2.49}$$

最后一个等号我们用到了简并条件 $\lambda_i = \lambda_j$。这个推导过程告诉我们，如果 $|i\rangle, |j\rangle$ 简并，则它们的任意线性组合将依然是算符 A 的本征值为 λ_i 的本征态，也就是说，$\{|i\rangle, |j\rangle\}$ 实际上张成了一个两维的线性子空间，这个子空间里的任何态矢量都是 A 的本征值为 λ_i 的本征态，这样一个线性子空间就称之为简并子空间。而在一个两维的线性子空间 (可以将这样的两维矢量空间想象成一个两维平面) 上我们总是可以重新选择两个相互正交的基矢量，我们可以将这两个正交的基矢量重新定义为新的 $\{|i\rangle, |j\rangle\}$ 态。类似的操作当然也可以推广到多重简并的情况。因此，这也就是说，碰到本征态简并的情况，我们总是可以对这些简并的本征态进行重新选择，使得重新选择以后的本征态正交。因此，对于厄米算符，我们总是可以选取一组两两正交的本征态，加上归一化条件，就有

$$\langle i|j\rangle = \delta_{ij}. \tag{2.50}$$

对于在量子力学中出现的厄米算符，由于它们总是可以作用在希尔伯特空间的任意态上，因此这组正交归一的本征态的线性叠加一定能构造出希尔伯特空间里的任何态矢量，也就是说，它们是完备的。同时满足正交归一性和完备性，这就意味着厄米算符的这一组本征态 $\{|i\rangle, i = 1,2,3....\}$ 可以构成整个希尔伯特空间的一组正交矢量基。

有时候，我们所考察的厄米算符 A 可能依赖于一个连续的控制参数 μ，记作 $A(\mu)$，这时候本征值 λ_i 也将是 μ 的函数，记为 $\lambda_i(\mu)$，相应的本征态 $|i\rangle$ 当然也将依赖于 μ，不妨重新记为 $|u_i(\mu)\rangle$。随着控制参数 μ 的连续变化，本征态矢量 $|u_i(\mu)\rangle$ 将在希尔伯特空间中连续变化。但是，按照我们在上一段中的证明，如果两个本征态 $|u_i(\mu)\rangle, |u_j(\mu)\rangle$ 有不同的本征值，那么它们一定正交，也就是说，虽然在控制参数 μ 的调节之下，这两个本征态都将在希尔伯特空间中连续变化，但是无论怎么变，它们将始终保持正交。什么时候它们才可能变得不再正交呢? 上一段的分析告诉我们，仅当在某个参数 μ_c 上，$\lambda_i(\mu_c) = \lambda_j(\mu_c)$ 时，也就是这两个态变成简并时，这种情况才可能发生。可见，从不简并到简并往往会带来系统性质的改变，这在量子相变的分析中是有用的。

最简单的厄米算符是什么呢? 当然是单位算符，或者说恒等算符，我们记作 1(读者很容易根据上下文确定任何公式中的 1 是代表数字 1，还是代表单位算符)。单位算符有一个非常有用的分解定理 (常常也称之为基础态的封闭性关系)，即

$$1 = \sum_i |i\rangle\langle i|, \tag{2.51}$$

式中 $\{|i\rangle, i=1,2,3....\}$ 是希尔伯特空间的一组正交归一矢量基。这个定理可以看作是奇异值分解定理的特例，但是，由于它在量子力学中极其有用，所以我们给出一个单独的推导。首先，由公式 (2.34) 和公式 (2.35) 可以知道，任意一个态矢量 $|\psi\rangle$ 都可以分解成 $|\psi\rangle = \sum_i \langle i|\psi\rangle |i\rangle$，很显然这个表达式可以重写成

$$|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle = \big(\sum_i |i\rangle\langle i|\big)|\psi\rangle, \tag{2.52}$$

上面所有的步骤都是行矢量、列矢量、以及矩阵的乘法规则，特别是用到了乘法结合律。公式 (2.52) 最右边的 $\sum_i |i\rangle\langle i|$ 显然是一个算符，公式 (2.52) 告诉我们，这个算符作用在任意态矢量 $|\psi\rangle$ 上都等于 $|\psi\rangle$ 本身，按照定义，这样的算符当然就是单位算符，这样我们就得到了重要结果 (2.51)。

下面我们就用公式 (2.51) 来推导出一些有意思的结论。首先，我们用它来推导厄米算符的谱分解定理 (这个名称是数学家的叫法)。我们从厄米算符 A 的本征方程 $A|i\rangle = \lambda_i|i\rangle$ 出发，将这个方程左右两边从右方乘以行矢量 $\langle i|$ 就得到，$A|i\rangle\langle i| = \lambda_i|i\rangle\langle i|$，两边对 i 求和，就有 $A\sum_i |i\rangle\langle i| = \sum_i \lambda_i|i\rangle\langle i|$, 前面说了，厄米算符的本征态可以取成希尔伯特空间的一组正交归一基矢量，因此我们可以利用单位算符的分解定理 (2.51)，进而得到

$$A = \sum_i \lambda_i |i\rangle\langle i|. \tag{2.53}$$

这就是厄米算符的谱分解定理，本征值集合 $\{\lambda_i, i=1,2,3....\}$ 又称为厄米算符 A 的谱，很显然，这个定理其实也是奇异值分解定理的特例。

如果希尔伯特空间的基矢量是连续的，那公式 (2.51) 中的对 i 求和当然就应该改写成某个积分。比方说，微观粒子的位置为确定的 x 值的状态 $|x\rangle$ 的集合就可以构成单粒子希尔伯特空间的一组正交基，因此与公式 (2.35) 相应的，我们有

$$\psi(x) = \langle x|\psi\rangle. \tag{2.54}$$

而与公式 (2.51) 相应的则有

$$\int dx|x\rangle\langle x| = 1. \tag{2.55}$$

$|x\rangle$ 也称作位置本征态，因为 $|x\rangle$ 态的粒子有确定的位置坐标 x。不同坐标的 $|x\rangle$ 态由于可以根据粒子坐标取值的不同而可以确定地区分，因此它们是相互正交的，数学上我们可以将这个正交归一化关系写成 $\langle x|y\rangle = \delta(x-y)$，用 δ 函数而不用克隆内克符号 δ_{ij} 的原因当然是因为 x,y 现在是连续变量。同样，微观粒子动量确定的态 $|p\rangle$ 的集合也构成了单粒子希尔伯特空间的一组正交基，因此

$$\int dp|p\rangle\langle p| = 1. \tag{2.56}$$

由上面这三个式子我们就有

$$\psi(x) = \langle x|\psi\rangle = \langle x|\int dp|p\rangle\langle p|\psi\rangle = \int dp\langle x|p\rangle\langle p|\psi\rangle, \tag{2.57}$$

其中第二个等号就是在行矢量 $\langle x|$ 和列矢量 $|\psi\rangle$ 的中间插入一个单位矩阵 (单位算符)1。类似于 (4.61), $\langle p|\psi\rangle$ 就是动量空间波函数，我们记作 $\psi(p)$ (当然从数学语言的角度来看，这一记号是不合法的，因为动量空间波函数和坐标空间波函数当然不是同一个函数，因此不应该用同一个函数记号 ψ, 我们这里都用 ψ 是为了强调它们是同一个量子态的不同表示，但是当然要记住 $\psi(p)$ 和 $\psi(x)$ 不仅仅是自变量符号不同，它们还应该理解成两个不同的函数)。另外，按照公式 (4.61)，$\langle x|p\rangle$ 当然就应该是动量为 p 的粒子的波函数，德布罗意告诉我们这个波函数是平面波 $\frac{1}{(2\pi\hbar)^{\frac{1}{2}}}e^{ipx/\hbar}$, 因此我们有

$$\langle x|p\rangle = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}e^{ipx/\hbar}. \tag{2.58}$$

将 (3.12) 代入刚才得到的 (2.57)，就有

$$\psi(x) = \int dp \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}e^{ipx/\hbar}\psi(p), \tag{2.59}$$

这就是傅里叶变换。类似的，我们也容易得到傅里叶逆变换，

$$\psi(p) = \langle p|\psi\rangle = \int dx\langle p|x\rangle\langle x|\psi\rangle = \int dx\langle p|x\rangle\psi(x), \tag{2.60}$$

再代入 $\langle p|x\rangle = \langle x|p\rangle^* = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}e^{-ipx/\hbar}$ 就得到傅里叶逆变换的公式。所以傅里叶变换我们也不需要记了，只需要记住单位算符的分解定理就可以了。当然，我们这里只处理了一维空间，但你也很容易将上面的推导推广到三维空间。

关于单位算符的分解定理及其应用我们暂时就谈到这里，后面我们还会不断地用到这个定理。

以上我们只考虑了希尔伯特空间上的线性算符，这是因为我们设想在经过这些算符所代表的操作之后，量子态之间的线性叠加关系应该得以保持。但实际上，在量子力学中也出现了所谓的反线性算符，那就是时间反演算符，以及尤其是量子场论中会出现的 CPT 算符，但它们都是所谓的反线性反幺正算符，是作为量子系统的某种对称变换而出现的 (分别对应时间反演不变性以及 CPT 不变性)，并不是作用在量子态上的一个通常操作，也就是说，不能设想通过某个“设备”来实现这样的算符，因此它们暂时和我们关系不大，我们将忽视它们。

2.2.3 为什么物理量用厄米算符表示

任何物理量或者说可观测量，它的取值当然是一个实数，但是我们也说过，在量子力学中，不能谈论处系统处在叠加态上时相应物理量的值，因为这时候这个物理量的值是没有定义的，而当某个物理量有确定取值时，系统必定处于这个物理量的某个本征态。这就说明，在量子力学中，物理量本身不可能像在经典物理中那样用一个实数值的量来表示，因为否则这个物理量的值就总是有定义的，它就总是有确定取值。实际上，人们发现，在量子力学中，物理量本身应该用厄米算符来表示。那这是为什么呢?

原因在于，首先，根据量子力学的基本原理，某个物理量有不同确定取值的量子态之间，由于可以根据这个物理量取值的不同来确定地区分，因此这些态必定是两两正交的。

而厄米算符刚好有类似的性质，根据我们在前面的数学证明，厄米算符不同本征值的本征态必定是相互正交的。而且，厄米算符的本征值当然只在本征态上有定义，如果系统处在某个厄米算符本征态的叠加态上，那当然就谈不上本征值，而厄米算符本身是一个算符而不是一个实数，因此也就是说，这时候厄米算符的值是没有定义。这些都刚好吻合物理量所必须遵循的基本原理。这就使得我们想到，可以把物理量的本征态当作某个厄米算符的本征态，把物理量的值当成某个厄米算符的本征值，也就是说，把物理量用厄米算符来表示！而且，我们前面也证明过，厄米算符的本征值刚好是实数，因此完全可以当成物理量的值，如果厄米算符的本征值可以是复数的话，那将物理量表示成厄米算符就不可能成立了，因为任何物理量的值当然都不可能是复数，毕竟复数纯粹是数学构想，是没有测量意义的。

不光如此，我们还可以进一步计算一下系统在任意 $|\psi\rangle$ 态时某物理量 A 的期望值 $\overline{A}$。假设这个物理量的可能取值为 $\{\lambda_i, i=1,2,3...\}$(前面我们已经定义过，这些值叫做物理量的本征值)，假定测量 $|\psi\rangle$ 态的系统得到 λ_i 值的概率为 p_i，则很显然，A 的期望值等于 $\overline{A}=\sum_i \lambda_i p_i$。而根据前面的公式 (2.36)，$p_i=|\langle i|\psi\rangle|^2$，因此我们有

$$\overline{A}=\sum_i \lambda_i |\langle i|\psi\rangle|^2=\sum_i \langle\psi|i\rangle\lambda_i\langle i|\psi\rangle, \tag{2.61}$$

式中最后一个等号我们利用了 $\langle\psi|i\rangle=\langle i|\psi\rangle^*$。观察一下这个结果最右边的等式，我们就会发现最后出现了一个 $\sum_i |i\rangle\lambda_i\langle i|$，而根据前面我们证明的厄米算符的谱分解定理，这刚好是一个厄米算符，这个厄米算符的本征值正是 $\{\lambda_i, i=1,2,3...\}$，本征态正是 $\{|i\rangle, i=1,2,3...\}$，因此它正是我们上一段猜想的与物理量 A 对应的厄米算符，我们不妨还是用同样的符号 A 来表示它。这样一来，刚才的推导结果就告诉我们，物理量 A 在 $|\psi\rangle$ 态上的期望值 $\overline{A}$ 可以表示成

$$\overline{A}=\langle\psi|A|\psi\rangle, \tag{2.62}$$

等式右边的 A 正是我们猜想的与物理量 A 对应的厄米算符。由于 $|\psi\rangle$ 是任意的，而且结论 (2.62) 对于任何物理量都成立，它告诉我们，对于任何物理量，其在任意量子态上的期望值都可以通过一个相应的厄米算符计算出来。这就已经足以证明我们上一段建立起来的物理量与厄米算符之间的联系成立了！

有了物理量与厄米算符之间的这个联系以后，剩下的问题就是找到这些厄米算符了。在实际应用中，我们往往是根据各种物理直观和物理推理过程写出某个物理量相应的厄米算符，然后再根据这个厄米算符的本征方程，求出物理量的本征态和本征值 (现在物理量的本征态与本征值与相应厄米算符的本征态和本征值已经是一回事了)，然后将这些本征值与测量所得的物理量的值进行比较，如果两者吻合得很好，那我们就知道我们根据物理直观所写出来的这个厄米算符已经正确地表示出这个物理量了，如果吻合得不好，我们就在厄米算符中加上一些项或者减去一些项以使得最后的结果更加吻合，当然如果一开始差得太远的话我们甚至需要完全重写这个厄米算符。当然，我们往往还需要解释我们加上的这些项代表什么物理含义，我们减去某些项的物理理由又是什么。但是，不管怎么样，从本质上来说，写出一个物理量所对应的厄米算符本质上是一个不断猜测然后再实验验证的

过程。此外还有一种情况也很常见，尤其在理论物理学中是最常见的做法，那就是有时候我们可以根据一些基本的物理原理导出一个物理量的厄米算符表示。

关于这一如何寻找一个物理量所对应的厄米算符的基本逻辑，最清楚的演示莫过于我们如何找到一个系统的哈密顿算符 (即与能量相对应的厄米算符)。但这里我们不妨以角动量算符的历史演化为例来加以说明。对于角动量算符，人们先是根据经典力学的基本原理导出 $\vec{J}=\vec{x}\times\vec{p}$，再根据量子力学的量子化规则得出与 $\vec{x}$ 和 $\vec{p}$ 对应的位置算符和动量算符，进而也就得到了角动量算符。但是后来我们发现这不能解释斯特恩-盖拉赫实验，因此我们又引入了电子自旋的概念，因此电子的自旋量子态以及相应的自旋算符最初就是猜测的，是为了解释实验。同时，由于自旋-轨道耦合，我们发现在原子物理中，轨道角动量和自旋角动量分别都不守恒，仅当我们把自旋角动量算符加进角动量算符的表达式以后，原子的总角动量才总是符合角动量守恒的，这里既有物理原理的引导也有根据实验数据而来的猜测。只到最后，狄拉克才从相对论量子力学的基本原理出发推导出了电子的自旋算符，同时也发现自旋角动量算符自动和轨道角动量算符加在一起构成了总角动量算符。但即使是狄拉克从相对论和量子力学的基本原理出发得到的方程，也还是需要实验检验的。

2.2.4 不确定原理与算符

前面我们已经说过，在量子力学中，并非所有物理量都可以同时有确定值，比方说，电子自旋角动量的两个不同分量就不能同时有确定值，再比方说位置坐标和相应的动量也不能同时取确定值。以电子自旋的 x 分量 S_x 和 z 分量 S_z 为例，这两者不能同时确定的原因在于，S_x 的本征态其实是 S_z 的相干叠加态，因此 S_x 有确定值的时候，S_z 的值没有定义，当然也就不确定，反过来也一样，S_z 的本征态其实是 S_x 的叠加态，因此 S_z 有确定值的时候，S_x 的值也没有定义，总之，S_x 和 S_z 这两者不能同时有确定值。类似的，利用德布罗意告诉我们的平面波波函数，微观粒子的动量本征态 $|p\rangle$ 可以写成

$$|p\rangle=\int dx|x\rangle\langle x|p\rangle=\int dx\frac{1}{(2\pi\hbar)^{\frac{1}{2}}}e^{ipx/\hbar}|x\rangle, \tag{2.63}$$

可见它是位置本征态 $|x\rangle$ 的相干叠加态，因此粒子动量有确定值的时候其位置坐标的值不确定。反过来，位置本征态 $|x\rangle$ 也可以写成动量本征态 $|p\rangle$ 的叠加态，

$$|x\rangle=\int dp|p\rangle\langle p|x\rangle=\int dp\frac{1}{(2\pi\hbar)^{\frac{1}{2}}}e^{-ipx/\hbar}|p\rangle, \tag{2.64}$$

因此，当微观粒子位置坐标有确定值的时候，其动量也是不确定的。总之，微观粒子的位置坐标和其动量不能同时确定。

由此可见，如果两个物理量的值总是可以同时测定 (请注意总是这个词)，那就意味着，不管系统处在任何量子态，每次测量时所塌缩到的量子态必然同时是这两个物理量的本征态，由于我们假设对任何量子态进行测量总能同时得到这两个物理量的某个值，因此这也意味着这些共同的本征态必然可以张成希尔伯特空间的一组矢量基 (即任何量子态都可以写成这组共同本征态的线性叠加)，因此，如果两个物理量的值总是可以同时测定，那就意味着它们必然有一组共同的本征态矢量基。但是，我们已经看到，物理量是用厄米算符来

表示的，那么从厄米算符的角度来看，两个物理量的值总是可以同时测定的充要条件是什么呢？答案很简单，是这两个算符的乘积要可以交换顺序，简称这两个算符可对易。

我们首先证明必要性。即假设两个物理量 A,B 的值总是可以同时测定，也就是说它们有共同的本征态矢量基 (不妨记作 $\{|i\rangle, i=1,2,3...\}$，并记 A 的相应本征值为 α_i，B 的相应本征值为 β_i)，证明相应的厄米算符 A 和 B 必定可对易，即 $AB=BA$ 或者说 $AB-BA=0$。证明很容易，首先由于 $\{|i\rangle, i=1,2,3...\}$ 是希尔伯特空间的矢量基，因此任何量子态 $|\psi\rangle$ 都必定可以写成 $|\psi\rangle=\sum_i c_i|i\rangle$ 的形式，如此一来就有

$$\begin{aligned}(AB-BA)|\psi\rangle &= \sum_i c_i(AB-BA)|i\rangle \\ =\sum_i c_i(\beta_i A|i\rangle - \alpha_i B|i\rangle) &= \sum_i c_i(\beta_i\alpha_i - \alpha_i\beta_i)|i\rangle = 0.\end{aligned} \tag{2.65}$$

即 $(AB-BA)$ 在任意态上的作用都等于零，因此即有 $AB-BA=0$。这就完成了必要性的证明。

下面证明充分性 (充分性的证明略微有一点复杂，并且纯粹是一个数学定理，读者第一遍读的时候可以先略过)。即要证明如果两个物理量的厄米算符满足 $AB=BA$，则它们必有一组共同的本征态矢量基。首先，我们前文已经说过，对于量子力学中考察的厄米算符而言，其本征态必定可以取作希尔伯特空间的正交矢量基。所以真正需要证明的是，我们总是可以把这些本征态取成这两个算符共同的本征态。为此我们假设 $\{|i\rangle, i=1,2,3...\}$ 为算符 A 的一组本征态，本征值分别为 α_i(注意，α_i 可以有简并)，则由于

$$A(B|i\rangle) = AB|i\rangle = BA|i\rangle = \alpha_i B|i\rangle, \tag{2.66}$$

因此我们可以知道 $B|i\rangle$ 必然也是算符 A 的本征值为 α_i 的本征态。如果算符 A 的 α_i 这个本征值对应的本征态没有简并，那就必然有 $B|i\rangle$ 依然正比于 $|i\rangle$ 这个态，假如把比例系数记为 β_i，从而也就有 $B|i\rangle=\beta_i|i\rangle$，这也就是说 $|i\rangle$ 同时也是算符 B 的本征态。因此在这种情况下 $|i\rangle$ 就已经是 A,B 共同的本征态了。

但是，如果算符 A 的本征值为 α_i 的本征态有简并，那这时候由 (2.66) 式我们就不能得出 $B|i\rangle$ 正比于 $|i\rangle$ 的结论了。比方说假设有二重简并 (多重简并的证明是类似的)，也即是说有某个 $\alpha_j=\alpha_i$，从而相应的 $|j\rangle$ 态和 $|i\rangle$ 态有相同的本征值 α_i。这就和我们前面证明总是可以将厄米算符的本征态取成正交的时候碰到的简并情况是一样的。$|i\rangle$ 和 $|j\rangle$ 简并的一个显然结果是，它们的任意线性叠加 $c_1|i\rangle+c_2|j\rangle$ 都将是 A 的本征态，本征值都是 α_i，因此这就张成了一个两维的简并子空间，我们可以把这个简并子空间记为 $\mathscr{H}_{\alpha_i}$。那这时候我们由 (2.66) 式就只能得出 $B|i\rangle$ 一定是这个两维的简并子空间里的某个态矢量。不仅如此，由于

$$\begin{aligned}A[B(c_1|i\rangle+c_2|j\rangle)] &= B(c_1A|i\rangle + c_2A|j\rangle) \\ =B(c_1\alpha_i|i\rangle+c_2\alpha_j|j\rangle) &= \alpha_i[B(c_1|i\rangle+c_2|j\rangle)],\end{aligned} \tag{2.67}$$

这里最后一个等号我们利用了简并条件 $\alpha_j=\alpha_i$。这个推导过程 (2.67) 就告诉我们，$[B(c_1|i\rangle+c_2|j\rangle)]$ 依然是 A 的某个本征值为 α_i 的本征态，这也就是说，B 作用在简并子空间 $\mathscr{H}_{\alpha_i}$ 里

的任意态上，结果都依然还是简并子空间里的态。因此我们就总可以在这个两维的简并子空间 $\mathscr{H}_{\alpha_i}$ 里求解算符 B 的本征方程，从而得到 B 的两个本征态 $|i'\rangle$ 和 $|j'\rangle$。前文引入厄米算符本征态时的相关证明告诉我们，无论这两个态对应的 B 本征值相同还是不同，我们总是可以让它们正交，从而成为二维简并子空间 $\mathscr{H}_{\alpha_i}$ 新的正交基 (因此原来的 $|i\rangle$ 和 $|j\rangle$ 也都可以反过来写成这两个新的基矢量的线性叠加)。由于 $|i'\rangle$ 和 $|j'\rangle$ 都属于 A 的简并子空间 $\mathscr{H}_{\alpha_i}$，所以它们当然同时也是 A 的本征态，这样它们就是 A 和 B 共同的本征态。也就是说，碰到这种二重简并的情况，我们只需要把原来的 $|i\rangle,|j\rangle$ 态替换成新的 $|i'\rangle$ 和 $|j'\rangle$ 态就找到了这时候两个算符的共同本征态。多重简并情形的推理和结论都是类似的。因此这就证明了，A 和 B 的共同本征态总是可以找到的。这样就完成了我们对充分性的证明。

刚才的充分性证明告诉我们，如果两个算符可对易，那么它们相应的物理量的值必可同时测定。这个命题的逆否命题就是，不可同时测定的两个物理量，它们相应的算符必定不对易。由于电子自旋的 S_x 分量和 S_z 分量不可同时测定，因此我们必有，算符 S_x 和 S_z 必定不对易，即

$$[S_x,S_z]\neq 0. \tag{2.68}$$

同样的，我们也必定有

$$[X,P]\neq 0. \tag{2.69}$$

这里 X 表示微观粒子的位置算符，P 表示微观粒子的动量算符，$[A,B]$ 称作两个算符的对易子，它的定义是 $[A,B]=AB-BA$。那么这两个例子中的这些不等于零的对易子到底应该是多少呢？我们后面将继续讨论。

另外，必要性命题的逆否命题是，如果两个算符不对易，那么它们相应物理量的值就不能总是可以同时测定。这里就出现了另一种可能性，即算符 A,B 虽然对易子不等于零，但是可能偶尔可以同时测定。这时候就说明，这两个算符必定存在某些共同的本征态，但这些本征态不足以构成希尔伯特空间的矢量基。一个例子就是角动量算符，角动量算符的任何两个分量都不可对易，但是对于特殊的角动量为 0 的量子态 $|\vec{J}^2=0\rangle$ 而言，它就是 J_x,J_y,J_z 的一个共同本征态，本征值都是 0。因此如果系统处在 $|\vec{J}^2=0\rangle$ 态上，那我们就可以同时测得 J_x,J_y,J_z 的值都为 0。当然，这时候，由于 $|\vec{J}^2=0\rangle$ 是 J_x,J_y,J_z 的一个共同本征态，所以虽然 $[J_x,J_y]\neq 0$，但是 $[J_x,J_y]|\vec{J}^2=0\rangle=0$。反过来，如果某两个算符 A,B, 它们的对易子在任何态 $|\psi\rangle$ 上的期望值都不为 0，即 $\langle\psi|[A,B]|\psi\rangle\neq 0$ 对所有的 $|\psi\rangle$ 都成立，那 A,B 就不可能有任何共同的本征态，因此这时候 A,B 的值必定不可同时测定。一个典型的例子就是位置算符和动量算符，即 X,P，这两个算符的对易子就是，$[X,P]=i\hbar$，即对易子是个常数，因此当然在任何态上的期望值都不为 0，因此我们就必然有，在任何情况下 X 和 P 都不可能同时测定。当然，这个结论我们在本小节最开始的时候就已经知道了。

2.2.5 么正演化与薛定谔方程

现在让我们来考虑量子态随时间的演化，态叠加原理作为量子力学最为基本的原理，它当然应该在量子态的演化之下得以保持，因此如果我们将系统从 t_1 时刻到 t_2 时刻的时

间演化看是由 t_1 时刻的量子态 $|\psi(t_1)\rangle$ 到 t_2 时刻的量子态 $|\psi(t_2)\rangle$ 的一个映射的话，这个映射应该是一个线性映射，也就是说，要用线性算符来刻画，不妨将这个线性算符记为 $U(t_2,t_1)$，称之为时间演化算符，则

$$|\psi(t_2)\rangle = U(t_2,t_1)|\psi(t_1)\rangle. \tag{2.70}$$

但是，我们也可以让时间反演，让系统反过来从 t_2 时刻演化到 t_1 时刻，从而有

$$|\psi(t_1)\rangle = U(t_1,t_2)|\psi(t_2)\rangle. \tag{2.71}$$

由这两个式子我们就有 $|\psi(t_2)\rangle = U(t_2,t_1)U(t_1,t_2)|\psi(t_2)\rangle$，以及 $|\psi(t_1)\rangle = U(t_1,t_2)U(t_2,t_1)|\psi(t_1)\rangle$。而由于 $|\psi(t)\rangle$ 是希尔伯特空间的任意量子态，因此这就意味着

$$U(t_1,t_2)U(t_2,t_1) = U(t_2,t_1)U(t_1,t_2) = 1. \tag{2.72}$$

也就是说，时间演化算符 $U(t_2,t_1)$ 必然是可逆的，$U^{-1}(t_2,t_1) = U(t_1,t_2)$。

另外在物理上，任意态 $|\psi\rangle$ 与其自身的内积代表总概率，概率守恒要求它恒等于 1。这就要求时间演化算符必须保持

$$\langle\psi(t_2)|\psi(t_2)\rangle = \langle\psi(t_1)|U^\dagger(t_2,t_1)U(t_2,t_1)|\psi(t_1)\rangle = \langle\psi(t_1)|\psi(t_1)\rangle, \tag{2.73}$$

由于 $|\psi(t)\rangle$ 是任意的量子态，因此这一式子的后一个等号仅在下式得以满足时才能成立，

$$U^\dagger(t_2,t_1)U(t_2,t_1) = 1. \tag{2.74}$$

正如我们前面在线性算符这一节中已经证明过了的，这时候 $U(t_2,t_1)$ 的可逆性会进一步告诉我们，$U(t_2,t_1)$ 必然是一个幺正算符。这也就是说，量子态随着时间的演化必然是一种幺正演化，当然它也必然保持希尔伯特空间任意两个态的内积，这就是所谓的量子力学的幺正性。特别的，它告诉我们任意两个相互正交的量子态在时间演化之下将会始终保持正交，由于正交性是两个量子态可以确定地区分的充要条件，因此幺正性也告诉我们，任意两个可以确定地区分的量子态在时间演化之下将会始终都可以确定地区分。

很明显，如果时间演化算符中的 $t_2 = t_1 = t$，也就是说 t 的量子态演化到 t 时刻，那当然还是它本身，因此必然有 $U(t,t) = 1$。进一步，如果我们取 t_1 和 t_2 为两个相隔无穷小的时间，即 $t_1 = t, t_2 = t + dt$，那么我们就可以将时间演化算符作无穷小展开

$$U(t+dt,t) = 1 - \frac{i}{\hbar}Hdt. \tag{2.75}$$

式中我们已经忽略了高阶无穷小量，式中的 H 当然也是一个算符，而引入系数 $\frac{i}{\hbar}$ 的目的我们后面可以看到 (纯粹从数学的角度来说这么做是任意的)。将这个无穷小展开式代入幺正性的方程 (2.74)，我们就可以得到

$$1 = (1 + \frac{i}{\hbar}H^\dagger dt)(1 - \frac{i}{\hbar}Hdt) = 1 + \frac{i}{\hbar}(H^\dagger - H)dt + ... \tag{2.76}$$

同样的，高阶无穷小量被忽略了，比较这个式子的最左边和最右边，我们发现一阶无穷小部分必须等于零，从而

$$H^\dagger = H. \tag{2.77}$$

也就是说，幺正性告诉我们，之前引入的算符 H 必定是一个厄米算符。

另一方面，由时间演化算符的定义，我们有

$$\begin{aligned}|\psi(t+dt)\rangle &= U(t+dt,t)|\psi(t)\rangle \\ &= \left(1-\frac{i}{\hbar}Hdt\right)|\psi(t)\rangle = |\psi(t)\rangle - \frac{i}{\hbar}Hdt|\psi(t)\rangle.\end{aligned} \tag{2.78}$$

由于 $\partial_t|\psi(t)\rangle = (|\psi(t+dt)\rangle - |\psi(t)\rangle)/dt$，因此由上面的结果就可以得到

$$i\hbar\partial_t|\psi(t)\rangle = H|\psi(t)\rangle. \tag{2.79}$$

这就是薛定谔方程最一般的形式，而我们之前引入的厄米算符 H 就是哈密顿算符。在最一般的情况下，哈密顿算符 H 可以是显含时间 t 的，但如果 H 不显含 t，那我们就很容易验证 $|\psi(t)\rangle = \exp(-iH(t-t_0)/\hbar)|\psi(t_0)\rangle$ 是薛定谔方程 (10.2) 在形式上的通解，将这个通解与时间演化算符的定义进行比较，就可以知道这时候

$$U(t,t_0) = \exp\left(-i\frac{H}{\hbar}(t-t_0)\right). \tag{2.80}$$

从时间演化算符的定义还能够导出它的一些一般性质，比方说 $U(t_3,t_1) = U(t_3,t_2)U(t_2,t_1)$，在数学上，这一乘法性质再加上幺正性就完全刻画了时间演化算符的数学结构。

当然我们还需要确定哈密顿算符代表的是什么物理量。为此，我们注意到德布罗意关系告诉我们，一个具有确定能量 E 的粒子其波函数对时间的依赖关系为 $\exp(-iEt/\hbar)$，将这个结果与哈密顿算符不显含时间情形下的表达式 $|\psi(t)\rangle = \exp(-iHt/\hbar)|\psi(0)\rangle$ 进行比较，我们就可以知道，哈密顿算符是与能量相对应的，是能量这个物理量的算符表示。

到此为止，我们就已经从量子力学的幺正性出发，证明了哈密顿算符的存在性，并确定了它的物理含义是能量算符。下一个问题就是，给定一个量子力学系统，我们怎么具体地写出它的哈密顿算符呢？哈密顿算符对应能量是一个一般性的指引，也就是说，我们要把对系统能量有贡献的每一项都包括在哈密顿算符里面。但有时候，我们也许不能预先确定哪些项对系统能量有贡献，这时候就要用到我们在讨论如何构造物理量的厄米算符的时候谈到的一般性原则了，也就是说，我们可以先猜测哪些项对能量有贡献，然后写出相应的哈密顿算符，再求出这个哈密顿算符的本征值 (当然这就是系统能量的取值)，再将算出来的这些值和能量的实际测量值进行比较，或者我们也可以根据写出来的这个哈密顿算符求出系统的时间演化，进而算出在这种时间演化之下某些物理量的期望值随时间的变化，然后再将之与实验进行比较。如此不断调整我们的猜测，直到理论和实验相吻合。

2.2.6 算符的矩阵表示

到此为止我们所有的推导都是使用狄拉克符号和相应的线性算符来进行的。这样做的好处很明显，因为它们不光是所有的物理学家使用的通用语言，更重要的是它们最容易说清楚量子力学的一般性理论结构。因此，到此为止我们几乎所有的讨论都是普遍适用的，它适用于单个微观粒子，也适用于多粒子体系，同样也适用于量子场，甚至适用于整个宇宙。但这样的表述有时候也有它不方便的地方，那就是过于抽象了，我们是在比较抽象的

意义上讨论量子态和算符的，我们使用的狄拉克符号也是一种抽象的记号，从某种意义上来说，我们前面所有涉及到狄拉克符号和算符的推导实际上都是在做某种抽象代数的运算。当然，前面谈到态矢量 $|\psi\rangle$ 的时候，我们常常让大家将它想象成列矢量，说到 $\langle\psi|$ 的时候我们会让大家想象列矢量 $|\psi\rangle$ 的厄米共轭，也就是一个行矢量，而谈到算符的时候，我们也总是让大家将它想象成矩阵。但严格来说，这些都是不必要的，我们这么说的时候其实是为了帮助大家理解这些抽象的东西的运算规则。

然而，在某种意义上，态矢量 $|\psi\rangle$ 和列矢量的确本质是一回事，算符和矩阵也的确是一回事。更准确一点来说，列矢量是态矢量 $|\psi\rangle$ 在希尔伯特空间的某个矢量基下的表示，同样，矩阵也是算符的具体表示。下面我们将进一步建立这两者间的密切联系。

首先，我们可以写出任意一个算符 A 的定义方程，即 $A|\psi\rangle = |\phi\rangle$。然后我们在希尔伯特空间中任选一组正交归一的矢量基 $\{|i\rangle, i = 1,2,3...\}$。我们把算符的定义方程和某个基矢量 $|j\rangle$ 做内积，从而就得到 $\langle j|A|\psi\rangle = \langle j|\phi\rangle$。最后，我们在这个表达式的 A 和 $|\psi\rangle$ 之间插入一个单位算符 1，并使用单位算符 1 的分解定理，从而就可以得到

$$\sum_i \langle j|A|i\rangle\langle i|\psi\rangle = \langle j|\phi\rangle. \tag{2.81}$$

前面我们说过 $\langle i|\psi\rangle$ 就是态矢量 $|\psi\rangle$ 在矢量基 $\{|i\rangle, i = 1,2,3...\}$ 中的第 i 个分量，同样 $\langle i|\phi\rangle$ 是态矢量 $|\phi\rangle$ 的第 i 个分量，这些分量都是数，因此我们当然可以将它们分别排成一个列矢量 ψ 和 ϕ (现在这两个符号表示的是通常的列矢量，希望不会引起大家的混淆) $\psi = (\langle 1|\psi\rangle, \langle 2|\psi\rangle, \langle 3|\psi\rangle, ...)^T$, $\phi = (\langle 1|\phi\rangle, \langle 2|\phi\rangle, \langle 3|\phi\rangle, ...)^T$(由于我们这里写的是行矢量，所以加一个转置符号 T 表示将之变成列矢量)。同时，我们引入一个矩阵 $\widehat{A}$, 它的第 j 行第 i 列 $\widehat{A}_{ji}$ 定义成 $\widehat{A}_{ji} = \langle j|A|i\rangle$。现在，我们就可以看出来，方程 (2.81) 实际上是一个标准的线性方程，$\widehat{A}\psi = \phi$。类似的，你也很容易得到 $\langle\phi|\psi\rangle = \sum_i \langle\phi|i\rangle\langle i|\psi\rangle = \phi^\dagger\psi$。类似的，如果你有两个算符 A 和 B, 它们的矩阵表示分别为 $\widehat{A}$ 和 $\widehat{B}$, 那么由于 $\langle k|AB|i\rangle = \sum_j \langle k|A|j\rangle\langle j|B|i\rangle$ (请注意，等式右边的表达式实际上就是矩阵 $\widehat{A}$ 和 $\widehat{B}$ 的乘积)，因此你可以看到，两个算符的乘积 AB 将表示成两个矩阵的乘积 $\widehat{A}\widehat{B}$，数学家常常称这样的过程为算符代数的矩阵表示。

特别的，对于厄米算符 A，由于它满足 (2.47), 所以它在任意矢量基 $\{|i\rangle, i = 1,2,3...\}$ 中的矩阵表示必然满足

$$\langle i|A|j\rangle^* = \langle j|A|i\rangle. \tag{2.82}$$

也即是说，矩阵 $\widehat{A}$ 的共轭转置必然等于它本身 $\widehat{A}^\dagger = \widehat{A}$，这样的矩阵称为厄米矩阵。所以，厄米算符的表示矩阵必定为厄米矩阵。

前面的分析告诉我们，只要选定了一组矢量基，那么任何量子态都可以在这个基中表示成一个列矢量，而任何算符都可以在这个基中表示成一个矩阵。反过来，给出某组矢量基中某个量子态的列矢量，和某个算符的矩阵，我们也可以唯一性地得到原来的抽象的态矢量和抽象的算符。这是因为，只要给出了这些列矢量和矩阵，我们其实就有诸如 (2.81) 式这样的方程，而利用单位算符 1 的分解定理，我们就能从这个方程反过来得到，$\langle j|A|\psi\rangle = \langle j|\phi\rangle$, 其中的 $|\psi\rangle$ 反过来由列矢量 ψ 给出，即 $|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle$, 进一步我们又

有 $|j\rangle\langle j|A|\psi\rangle = |j\rangle\langle j|\phi\rangle$，将这个式子两边分别对 j 进行求和，并再一次利用单位算符 1 的分解定理，我们就回到了原来的算符方程 $A|\psi\rangle = |\phi\rangle$, 而其中算符 A 现在反过来由矩阵 $\widehat{A}$ 给出，即 $A = \sum_{i,j}|j\rangle\widehat{A}_{ji}\langle i|$。由此可见，只要选定了一组矢量基，那么态矢量及其列矢量表示，算符及其矩阵表示，实际上是完全等价的。列矢量和矩阵表示的好处是方便于具体的计算，坏处是必须先选取一组矢量基。而抽象的狄拉克符号和抽象的算符不依赖于矢量基的选取，因此往往在推导理论公式的时候更加简洁方便。人们通常称算符 A 在矢量基 $\{|i\rangle, i = 1,2,3...\}$ 中的表示矩阵 $\widehat{A}$ 为算符 A 在 $\{|i\rangle, i = 1,2,3...\}$ 表象中的矩阵表示。

值得强调的是，矢量基的选取是任意的，因此如果你换一组不同的矢量基，那你将得到同一个态矢量的不同列矢量表示，同样，你也将得到同一个算符的不同矩阵表示。利用单位算符 1 在不同矢量基之下的不同分解，你很容易推导出这些不同表示之间的联系，人们有时候称这种联系为表象变换。既然矢量基的选取可以是任意的，一个重要的问题将是，选取哪个矢量基是最好的呢？回答是，这要看具体情况，在不同的情形中，选取不同的矢量基有不同的方便之处。当然，有一些矢量基是人们常用的，下面我们就举一个具体的例子。

电子自旋算符的矩阵表示

斯特恩-格拉赫实验发现电子有一个内部的自由度，即自旋。实验发现电子自旋状态要用量子力学来描述，在任何一个方向上，电子的自旋角动量只可能有两个不同的取值，即 $\pm\hbar/2$。作为物理量，电子的自旋角动量当然也要用厄米算符来表示，因此我们用 S_x, S_y, S_z 分别表示 x, y, z 三个分量的自旋角动量算符。由于自旋的任何分量只有两个取值 $\pm\hbar/2$，这就告诉我们算符 S_x, S_y, S_z 分别都只有这两个本征值。这也就是说算符 S_x^2, S_y^2, S_z^2 都只有一个取值即 $\hbar^2/4$, 这样的算符当然只能是常数算符，也就是说，

$$S_x^2 = S_y^2 = S_z^2 = \hbar^2/4. \tag{2.83}$$

另外，作为一种角动量，自旋算符也要满足角动量算符的一般代数关系，这里我们直接给出这些代数关系，

$$[S_x, S_y] = i\hbar S_z, [S_y, S_z] = i\hbar S_x, [S_z, S_x] = i\hbar S_y. \tag{2.84}$$

为了数学上的方便，习惯上人们通常引入所谓的泡利算符 $\sigma_x, \sigma_y, \sigma_z$，它们的定义是 $S_x = (\hbar/2)\sigma_x, S_y = (\hbar/2)\sigma_y, S_z = (\hbar/2)\sigma_z$。因此，要将电子自旋算符表示成矩阵，我们只需要等价地将泡利算符表示成矩阵就可以了。泡利算符当然也是厄米算符，而且它们的定义告诉我们，其本征值总是 ± 1，从而方程 (2.83) 就可以重写成

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1. \tag{2.85}$$

而方程 (2.84) 就将变成 $[\sigma_x, \sigma_y] = 2i\sigma_z, [\sigma_y, \sigma_z] = 2i\sigma_x, [\sigma_z, \sigma_x] = 2i\sigma_y$。因此就有 $\sigma_x\sigma_y + \sigma_y\sigma_x = (\sigma_x[\sigma_z, \sigma_x] + [\sigma_z, \sigma_x]\sigma_x)/(2i) = (\sigma_x\sigma_z\sigma_x - \sigma_x^2\sigma_z + \sigma_z\sigma_x^2 - \sigma_x\sigma_z\sigma_x)/(2i) = 0$，式中第一个等号我们利用了 $[\sigma_z, \sigma_x] = 2i\sigma_y$，第二个等号是直接从算符对易子的定义得来的，第三个等号利用了 $\sigma_x^2 = 1$，因此我们得到了 $\sigma_x\sigma_y + \sigma_y\sigma_x = 0$。类似的推导告诉我们

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0, \sigma_y\sigma_z + \sigma_z\sigma_y = 0, \sigma_z\sigma_x + \sigma_x\sigma_z = 0. \tag{2.86}$$

也就是说，这三个泡利算符是两两反交换的，或者说反对易的，所谓两个算符反对易即是指 $AB=-BA$。不光如此，而且由 $[\sigma_x,\sigma_y]=2i\sigma_z$ 我们还有 $\sigma_x\sigma_y-\sigma_y\sigma_x=2i\sigma_z$, 由于 σ_x 和 σ_y 反对易，这其实就是

$$\sigma_x\sigma_y=i\sigma_z. \tag{2.87}$$

我们注意到，泡利算符同时满足方程 (3.91) 和 (2.86), 由这两个方程其实就能推导出必有 $\sigma_x\sigma_y=\pm i\sigma_z$, 方程 (2.87) 只是在这两种可能性中选定了一种。同时满足类似这样的方程的抽象代数就叫 Clifford 代数，数学家已经系统地研究过这类代数的矩阵表示了，下面我们将演示给大家这样的表示可以如何进行。

下面的推导在数学上是完全严格的。它完全是从泡利算符的抽象代数出发，也即是从方程 (3.91)，(2.86) 以及方程 (2.87) 出发，然后构造出一切。尤其是，我们没有隐含地假设任何算符的本征态的存在性，我们是直接构造出了这些本征态。

下面开始我们的推导。首先我们注意到 σ_z 可以由 σ_x 和 σ_y 的乘积给出，所以我们的注意力将主要放在 σ_x 和 σ_y 上。我们定义两个新的算符 σ 和 $\sigma^\dagger$,

$$\sigma=(\sigma_x-i\sigma_y)/2,\ \sigma^\dagger=(\sigma_x+i\sigma_y)/2. \tag{2.88}$$

则利用方程 (3.91) 和 (2.86)，我们很容易得到

$$\sigma^2=\sigma^{\dagger 2}=0. \tag{2.89}$$

而且，由于 $\sigma^\dagger\sigma=(\sigma_x+i\sigma_y)(\sigma_x-i\sigma_y)/4=(\sigma_x^2+\sigma_y^2+i\sigma_y\sigma_x-i\sigma_x\sigma_y)/4=(1+\sigma_z)/2$, 这里最后一个等号我们利用了 σ_x 和 σ_y 的反对易以及 $\sigma_x\sigma_y=i\sigma_z$, 类似的，我们也可以算得 $\sigma\sigma^\dagger=(1-\sigma_z)/2$，归纳一下即有

$$\sigma^\dagger\sigma=(1+\sigma_z)/2,\ \sigma\sigma^\dagger=(1-\sigma_z)/2. \tag{2.90}$$

由后一个式子我们可以得到 $0=\sigma\sigma^{\dagger 2}=(\sigma^\dagger-\sigma_z\sigma^\dagger)/2$(因为 $\sigma^{\dagger 2}=0$), 即

$$\sigma_z\sigma^\dagger=\sigma^\dagger. \tag{2.91}$$

由于 $\sigma^2=0$，所以我们必定可以在电子的所有自旋量子态中找到某个量子态，我们记作 $|\downarrow\rangle$, 它满足

$$\sigma|\downarrow\rangle=0. \tag{2.92}$$

这是因为，任取一个自旋量子态 $|\psi\rangle$，如果它满足 $\sigma|\psi\rangle=0$, 那这个 $|\psi\rangle$ 就是我们要找的态，如果 $\sigma|\psi\rangle\neq 0$，那我们就可以令 $\sigma|\psi\rangle=|\phi\rangle$, 这时候将有 $\sigma|\phi\rangle=\sigma^2|\psi\rangle=0$(后一个等号是由于 $\sigma^2=0$), 因此 $|\phi\rangle$ 就将是我们要找的态。总之，满足方程 (2.92) 的量子态 $|\downarrow\rangle$ 总能找到。对于这个态我们必有 $0=\langle\downarrow|\sigma^\dagger\sigma|\downarrow\rangle=\langle\downarrow|(1+\sigma_z)|\downarrow\rangle/2$, 即有 $\langle\downarrow|(1+\sigma_z)|\downarrow\rangle=0$, 由于 σ_z 的本征值为 ±1, 因此这个结果就告诉我们 $|\downarrow\rangle$ 必定是 σ_z 的本征值为 -1 的本征态，即

$$\sigma_z|\downarrow\rangle=-|\downarrow\rangle. \tag{2.93}$$

这就是为什么我们记这个态为自旋向下态的原因，因为 σ_z 在这个态上的本征值为 -1 这就意味着它是 S_z 的本征值为 $-\hbar/2$ 的本征态，我们称这样的量子态为自旋向下态。

下面我们引入一个新的自旋量子态，我们记作 $|\uparrow\rangle$, 它的定义是

$$|\uparrow\rangle = \sigma^\dagger|\downarrow\rangle. \tag{2.94}$$

利用方程 (2.91), 我们容易得到 $\sigma_z|\uparrow\rangle = \sigma_z\sigma^\dagger|\downarrow\rangle = \sigma^\dagger|\downarrow\rangle = |\uparrow\rangle$, 也就是说，$|\uparrow\rangle$ 必为 σ_z 的本征值为 $+1$ 的本征态，即

$$\sigma_z|\uparrow\rangle = |\uparrow\rangle. \tag{2.95}$$

这就是我们为什么将这个态记作自旋向上态的基本原因，因为它作为 σ_z 的本征值为 $+1$ 的本征态就必然也是 S_z 的本征值为 $+\hbar/2$ 的本征态。

将方程 (2.94) 共轭转置，就有 $\langle\uparrow| = \langle\downarrow|\sigma$, 因此，$\langle\uparrow|\downarrow\rangle = \langle\downarrow|\sigma|\downarrow\rangle = 0$(后一个等号是由于方程 (2.92))，也即

$$\langle\uparrow|\downarrow\rangle = 0. \tag{2.96}$$

因此自旋向上态和自旋向下态必定是两个正交的量子态，这正符合这两个态可以确定地区分的物理要求。而且我们也可以算得 $\langle\uparrow|\uparrow\rangle = \langle\downarrow|\sigma\sigma^\dagger|\downarrow\rangle = \langle\downarrow|(1-\sigma_z)/2|\downarrow\rangle = \langle\downarrow|\downarrow\rangle$, 因此如果我们将最开始时选的 $|\downarrow\rangle$ 态归一化，那由方程 (2.94) 定义的 $|\uparrow\rangle$ 也将自动是归一化的, 即

$$\langle\downarrow|\downarrow\rangle = \langle\uparrow|\uparrow\rangle = 1. \tag{2.97}$$

由于同时满足正交归一性，因此我们构造出来的 $\{|\downarrow\rangle, |\uparrow\rangle\}$ 可以作为电子自旋量子态的两个基矢量。从物理上来说，这两个基矢量当然应该是完备的。但是，我们能够从泡利算符的代数出发从数学上证明这一点吗？回答很简单，可以。比方说我们观察刚才的两个基矢量 $|\downarrow\rangle, |\uparrow\rangle = \sigma^\dagger|\downarrow\rangle$, 我们发现 $|\uparrow\rangle$ 是用 $\sigma^\dagger$ 作用在 $|\downarrow\rangle$ 上得到的，那我们能用 $\sigma^\dagger$ 再重复作用一次从而得到第三个态吗？答案是不能，因为 $\sigma^{\dagger 2} = 0$, 因此 $\sigma^\dagger$ 的重复作用只能得到 0，即

$$\sigma^\dagger|\uparrow\rangle = 0. \tag{2.98}$$

那如果用 σ 作用在 $|\uparrow\rangle$ 上能得到新的态吗？答案是，也不能，因为 $\sigma|\uparrow\rangle = \sigma\sigma^\dagger|\downarrow\rangle = (1-\sigma_z)/2|\downarrow\rangle = |\downarrow\rangle$, 也即

$$\sigma|\uparrow\rangle = |\downarrow\rangle. \tag{2.99}$$

因此我们就证明了电子自旋的基矢量 $\{|\downarrow\rangle, |\uparrow\rangle\}$ 不能再进一步扩大，从而它们就是完备的。因此它们就可以作为电子自旋态空间的矢量基。

以 $\{|\uparrow\rangle,|\downarrow\rangle\}$ 为矢量基 (现在要注意这两个基矢量的排列顺序)，我们就能够分别计算出泡利算符 $\sigma_z,\sigma,\sigma^\dagger$ 在这个基中的表示矩阵 $\widehat{\sigma}_z,\widehat{\sigma},\widehat{\sigma}^\dagger$.

$$\widehat{\sigma}_z=\begin{pmatrix}\langle\uparrow|\sigma_z|\uparrow\rangle & \langle\uparrow|\sigma_z|\downarrow\rangle\\ \langle\downarrow|\sigma_z|\uparrow\rangle & \langle\downarrow|\sigma_z|\downarrow\rangle\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \tag{2.100}$$

$$\widehat{\sigma}=\begin{pmatrix}\langle\uparrow|\sigma|\uparrow\rangle & \langle\uparrow|\sigma|\downarrow\rangle\\ \langle\downarrow|\sigma|\uparrow\rangle & \langle\downarrow|\sigma|\downarrow\rangle\end{pmatrix}=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \tag{2.101}$$

$$\widehat{\sigma}^\dagger=\begin{pmatrix}\langle\uparrow|\sigma^\dagger|\uparrow\rangle & \langle\uparrow|\sigma^\dagger|\downarrow\rangle\\ \langle\downarrow|\sigma^\dagger|\uparrow\rangle & \langle\downarrow|\sigma^\dagger|\downarrow\rangle\end{pmatrix}=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}. \tag{2.102}$$

在计算的过程中，我们需要使用方程 (2.92)，(2.93)，(2.94)，(2.95)，(2.96)，(2.97)，(4.9)，(2.99)，当然，这些方程中的每一个其实都很简单。

而由 $\sigma,\sigma^\dagger$ 的定义式 (4.45)，我们又可以进一步得到 σ_x 的表示矩阵 $\widehat{\sigma}_x=\widehat{\sigma}+\widehat{\sigma}^\dagger$ 以及 σ_y 的表示矩阵 $\widehat{\sigma}_y=i(\widehat{\sigma}-\widehat{\sigma}^\dagger)$。这三个矩阵 $\widehat{\sigma}_x$，$\widehat{\sigma}_y$，$\widehat{\sigma}_z$ 就是著名的泡利矩阵，我们将它们归纳在下面，

$$\widehat{\sigma}_x=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\widehat{\sigma}_y=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\widehat{\sigma}_z=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{2.103}$$

2.2.7 * 一个数学附录

在这个附录中我们将给出希尔伯特空间线性算符奇异值分解定理的证明，我们将假定所要考察的这个希尔伯特空间是有限维的，并且为了简单起见，我们常常默认本征值非零的本征态没有简并，当然将我们的证明扩展到有简并的情况并不是一件很难的事情。我们要证明的即是，希尔伯特空间的任何线性算符 A 都可以写成如下形式

$$A=\sum_i|u_i\rangle\lambda_i\langle v_i|, \tag{2.104}$$

其中 $\lambda_i\neq 0$，$\{|u_i\rangle,i=1,2,3....\}$ 和 $\{|v_i\rangle,i=1,2,3....\}$ 分别从属于希尔伯特空间的两组正交归一的矢量基，因此有

$$\langle u_i|u_j\rangle=\langle v_i|v_j\rangle=\delta_{ij}. \tag{2.105}$$

首先构造厄米算符 $A^\dagger A$ 和 $AA^\dagger$。设 $|v_i\rangle$ 为 $A^\dagger A$ 的本征态，相应的本征值为 a_i，同样设 $|u_i\rangle$ 为 $AA^\dagger$ 的本征态，相应的本征值为 b_i。即，

$$A^\dagger A|v_i\rangle=a_i|v_i\rangle,AA^\dagger|u_i\rangle=b_i|u_i\rangle. \tag{2.106}$$

作为厄米算符的本征态，$\{|u_i\rangle,i=1,2,3....\}$ 和 $\{|v_i\rangle,i=1,2,3....\}$ 当然分别都是正交归一的。

又由于 $(AA^\dagger)A|v_i\rangle=A(A^\dagger A)|v_i\rangle=a_iA|v_i\rangle$，可见 $A|v_i\rangle$ 是 $AA^\dagger$ 的一个本征态，不妨取

$$A|v_i\rangle=\lambda_i|u_i\rangle. \tag{2.107}$$

同样，由于 $(A^\dagger A)A^\dagger|u_i\rangle = b_i A^\dagger|u_i\rangle$，所以 $A^\dagger|u_i\rangle$ 也是 $A^\dagger A$ 的一个本征态。又由于 $\lambda_i\delta_{ji} = \lambda_i\langle u_j|u_i\rangle = \langle u_j|A|v_i\rangle = \langle v_i|A^\dagger|u_j\rangle^*$，即 $\langle v_i|A^\dagger|u_j\rangle = \lambda_i^*\delta_{ji}$。可见必有

$$A^\dagger|u_j\rangle = \lambda_j^*|v_j\rangle. \tag{2.108}$$

由方程 (2.107) 和方程 (2.108) 可知，

$$A = \sum_i |u_i\rangle\lambda_i\langle v_i|, A^\dagger = \sum_i |v_i\rangle\lambda_i^*\langle u_i|. \tag{2.109}$$

这就证明了奇异值分解定理。

上面最后一步的详细推理过程是这样的：由方程 (2.107) 和方程 (2.108) 可知，$|u_i\rangle$ 和 $|v_i\rangle$ 是一一对应的。进而可以得到 $A|v_i\rangle\langle v_i| = \lambda_i|u_i\rangle\langle v_i|$，两边对 i 求和，就得到 $A\sum_i|v_i\rangle\langle v_i| = \sum_i\lambda_i|u_i\rangle\langle v_i|$，由于 $\{|v_i\rangle, i = 1,2,3....\}$ 是希尔伯特空间的矢量基，因此 $\sum_i|v_i\rangle\langle v_i| = 1$，这样就得到了最终的结果。而且由方程 (2.107) 和方程 (2.108) 也很容易知道 $a_i = |\lambda_i|^2 = b_i$。

一个细节的问题是，仔细观察我们的证明过程，你会发现过程中建立起来的 $|u_i\rangle$ 和 $|v_i\rangle$ 之间的一一对应仅在 $a_i = b_i > 0$(也就是 $\lambda_i \neq 0$) 时才成立。至于 $a_i = 0$ 的那些态，它和 $b_j = 0$ 的那些态之间其实并没有对应关系，而是相互独立的，好在这并不影响我们最后的结论，只是需要我们在最后的表达式中加上 $\lambda_i \neq 0$ 的限制，并将分别与 $a_i = 0$ 和 $b_j = 0$ 相应的那些本征态排除在求和范围之外。这些与 $a_i = 0$ 对应的本征态 (假设共有 N_a 个，即 N_a 重简并)，以及与 $b_j = 0$ 对应的本征态 (假设共有 N_b 个，即 N_b 重简并)，它们其实是数学家很关心的东西，因为对于很大一类线性算符，数学家证明了一个著名的数学定理，它说，$N_a - N_b$ 其实是一个拓扑不变量，这就是阿蒂亚-辛格指标定理。

$N_a - N_b$ 有某种拓扑不变性其实是不难理解的。所谓的拓扑不变性就是指在连续变化下的某种不变性，也就是说，设想线性算符 A 依赖于某些连续参数 μ，当 μ 连续变化时，$N_a(\mu) - N_b(\mu)$ 将保持不变。这里的原因其实很简单，设想随着 μ 的连续变化，某个 $a_i = 0$ 的本征态 $|v_i\rangle$ 变成了一个 $a_i \neq 0$ 的态，因此 N_a 就减少了 1。但是奇异值分解定理的证明过程告诉我们，$a_i \neq 0$ 的本征态一定是和 $b_i \neq 0$ 的本征态一一对应的，现在，$a_i \neq 0$ 的本征态增加了一个，那这就必然意味着 $b_i \neq 0$ 的本征态也同时新增了一个，这个新增的态是从哪来的呢？必然是因为某个原来 $b_i = 0$ 的本征态 $|u_i\rangle$ 随着参数 u 的连续调节变成了 $b_i \neq 0$。那这就意味着，N_b 也同时减少了 1。也就是说，随着参数 μ 的连续调节，N_a 和 N_b 必然是同步减少的，同样，我们也可以论证它们必然是同步增加的。既然 N_a 和 N_b 总是同步的增加或者同步的减少，那它们的差值 $N_a(\mu) - N_b(\mu)$ 就将保持不变。所以 $N_a - N_b$ 具有拓扑不变性的直观论证并不难，难的是找到相应的拓扑不变量的表达式，并证明它等于 $N_a - N_b$。阿蒂亚和辛格实际上是对一大类线性算符完成了这一困难的工作。

回到我们的奇异值分解定理。如果算符 A 满足 $AA^\dagger = A^\dagger A$(数学家称这种算符为正规算符)，则由奇异值分解定理的证明过程可知，这时候总可以取 $|u_i\rangle = |v_i\rangle$，因此 $|u_i\rangle$ 也是 A 的本征态，而 A 必然能够分解成

$$A = \sum_i |u_i\rangle\lambda_i\langle u_i|. \tag{2.110}$$

如果把算符表示成矩阵，与正规算子对应的就叫作正规矩阵，正规算子的分解定理实际上告诉我们，正规矩阵一定是可对角化的。反过来，如果一个希尔伯特空间的线性算符可对角化，也就是能分解成如 (2.110) 这样的形式，则由 $|u_i\rangle$ 的正交归一性很容易验证这个算符必为正规算符。

特别的，幺正算符 U 就是一个正规算符，而且由于 $U^\dagger U = UU^\dagger = 1$，所以 $|\lambda_i|^2 = 1$，所以必有 $\lambda_n = e^{i\theta_n}$。即幺正算符一定能分解成

$$U = \sum_n e^{i\theta_n} |u_n\rangle\langle u_n|. \tag{2.111}$$

比方说时间演化算符 $U = e^{-iHt/\hbar}$ 就可以分解成

$$e^{-iHt/\hbar} = \sum_n e^{-iE_n t/\hbar} |E_n\rangle\langle E_n|. \tag{2.112}$$

厄米算符也是正规算符，因此也必定能分解成形如 $\sum_i \lambda_i |u_i\rangle\langle u_i|$ 的形式，而且这时候本征值 λ_i 必定是实数。

2.2.8 习题

1. 请证明施瓦兹 (Schwarz) 不等式，

$$\langle\psi|\psi\rangle\langle\phi|\phi\rangle \geq |\langle\phi|\psi\rangle|^2.$$

提示：利用 $\big(\langle\psi| + z^*\langle\phi|\big) \cdot \big(|\psi\rangle + z|\phi\rangle\big) \geq 0$($z$ 为任意复数)。

2. 对于任意一个厄米算符 A 和任意一个量子态 $|\psi\rangle$，我们可以定义厄米算符 $\Delta A = A - \overline{A}$。$(\Delta A)^2$ 在 $|\psi\rangle$ 态上的期望值 $\overline{(\Delta A)^2} = \overline{A^2} - \overline{A}^2$ 称为物理量 A 的均方差。假设有两个厄米算符 A, B，以及某个任意的量子态 $|\psi\rangle$，请证明下面的不等式

$$\overline{(\Delta A)^2} \cdot \overline{(\Delta B)^2} \geq \frac{1}{4} \big|\overline{[A,B]}\big|^2.$$

提示：首先，对量子态 $|\psi_1\rangle = \Delta A|\psi\rangle$、$|\psi_2\rangle = \Delta B|\psi\rangle$ 用施瓦兹不等式。其次，注意到 $\Delta A \Delta B = \frac{1}{2}[A,B] + \frac{1}{2}\{\Delta A, \Delta B\}$，其中 $[A,B]^\dagger = -[A,B]$ 为一个反厄米算符，其期望值为纯虚数，而 $\{\Delta A, \Delta B\} = \Delta A \Delta B + \Delta B \Delta A$ 为厄米算符，其期望值为实数。

3. 证明：任意 2×2 的厄米矩阵 X 都可以写称 $X = x_0 \cdot 1 + x_1\hat{\sigma}_x + x_2\hat{\sigma}_y + x_3\hat{\sigma}_z$ 的形式，式中 x_0, x_1, x_2, x_3 是四个实数，并请证明 X 的行列式为 $\det(X) = x_0^2 - x_1^2 - x_2^2 - x_3^2$。

4. 注意上一题的 $\det(X)$。如果将 x_0, x_1, x_2, x_3 理解为狭义相对论的四维矢量的四个分量，那么 $\det(X)$ 刚好是这个四维矢量的“长度”平方。这意味着，任何四维矢量都对应于一个 2×2 的厄米矩阵。狭义相对论的间隔不变性告诉我们，任何四矢量的“长度”平方在洛伦兹变换下都是不变的，请据此证明：任意洛伦兹变换都可以表示为一个行列式为 1 的 2×2 复矩阵 (通常称这样的复矩阵的集合为 $SL(2,\mathbb{C})$)，并由此证明洛伦兹变换有 6 个独立的实参数。

5. 请证明算符恒等式: (1), $[A,BC]=[A,B]C+B[A,C]$。(2), $[A,[B,C]]+[B,[C,A]]+[C,[A,B]]=0$。

6. 有两个算符 A,B，记它们的对易子 $[A,B]=C$。如果 $[C,A]=[C,B]=0$。请证明算符恒等式

$$e^{A+B}=e^{A}e^{B}e^{-\frac{1}{2}[A,B]}.$$

7. 已知量子比特的两组正交归一矢量基 $\{|0\rangle,|1\rangle\}$ 以及 $\{|+\rangle=\frac{1}{\sqrt{2}}\big(|0\rangle+|1\rangle\big),\quad |-\rangle=\frac{1}{\sqrt{2}}\big(|0\rangle-|1\rangle\big)\}$, 且已知算符 F 在 $\{|0\rangle,|1\rangle\}$ 表象中的表示矩阵为

$$F=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}. \tag{2.113}$$

求算符 F 在表象 $\{|+\rangle,|-\rangle\}$ 中的表示矩阵。

2.3 量子纠缠

我们来考察双电子系统的自旋量子态，由于态叠加原理，下面这个量子态 $|\phi^+\rangle$ 显然是双电子系统的一个可能自旋态，

$$|\phi^+\rangle=\frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle+|\downarrow\downarrow\rangle). \tag{2.114}$$

这个量子态有一个特别的性质，即它不能因式分解成第一个电子的某个自旋态 $|\phi_1\rangle$ 与第二个电子的某个自旋态 $|\phi_2\rangle$ 的乘积 $|\phi_1\rangle|\phi_2\rangle=|\phi_1,\phi_2\rangle$ 这样的形式。这就意味着处在 $|\phi^+\rangle$ 态上的两个电子共同形成了一个不可分解的整体，而这个不可分割的整体与两个电子之间的距离没有关系，也即是说，即使你将其中一个电子放在南昌，另一个放到天边，这两者遥遥相隔，但是它们却依然处于同一个不可分解的整体之中。

现在，假设你沿着 z 轴测量南昌的这个电子的自旋，那你会得到两种可能的结果，$1/2$ 的概率你将测到南昌的这个电子自旋向上 $\uparrow$, $1/2$ 的可能性你将测到它自旋向下 $\downarrow$。但是，奇妙的是，由于南昌的这个电子和天边的那个电子形成了不可分解的整体 $|\phi^+\rangle$，因此当你测到南昌的电子自旋向上时就意味着原来的 $|\phi^+\rangle$ 态塌缩到了 $|\uparrow\uparrow\rangle$ 态，而这又意味着天边的电子立即处在 $|\uparrow\rangle$ 态，同样的，当你测到南昌的电子自旋向下时，原来的 $|\phi^+\rangle$ 态就塌缩到了 $|\downarrow\downarrow\rangle$ 态，因此天边的电子立即就处在 $|\downarrow\rangle$ 态。总之，无论你测到南昌的电子自旋向上还是自旋向下，两者的整体就塌缩了，相应的就立即决定了天边电子的量子态。由于这种你在南昌的测量对天边电子的影响是立即的，所以很多人会说：这意味着在量子力学里信息可以超光速传播，利用量子纠缠态可以超光速地传递信息。情况果真如此吗?

2.3.1 量子纠缠能实现超光速信息传递吗？

为了分析量子纠缠态能否实现超光速传递信息的问题，让我们假设某个实验室中制备了 N 对电子，每一对都处在纠缠态 $|\phi^+\rangle$, 你和你的她分别持有每一对电子中的一个，你待在南昌，而她去了天边，因此你们分享着 N 个纠缠对，但是你们之间不能通信。在这种情

况下，你能用你们之间分享的纠缠对来给她瞬时传递信息吗？比方说，你和她约定，如果你们在南昌养的那只猫死了，你就会对你的电子进行测量，而你的测量立即就会影响她持有的另一个电子，因此看起来只要她接着对自己持有的那个电子进行测量就能获知猫死的信息。情况真是这样的吗？

现在，假设南昌的猫死了，因此你沿着 z 轴测量了你们共享的 N 个纠缠对中你所持有的那些电子，但是纠缠对的塌缩是随机的，因此你有 $1/2$ 的可能性测到某个电子自旋向上，$1/2$ 的可能性测到自旋向下，而你无法决定自己的测量结果，因此完成测量之后，你的 N 个电子大约会有一半自旋向上，另一半自旋向下，哪些自旋向上，哪些自旋向下是完全随机的。接着，她也对自己的电子进行了测量，当然，你的测量立即影响到了她的电子，因此所有你测到你的电子自旋向上的那些纠缠对，她也会测到自己的另一个电子自旋向上，所有你测到自己的电子自旋向下的纠缠对，她也会测到自己的另一个电子自旋向下。但问题是，你们之间不能通信，因此她无从得知你的测量结果，也就是说，她不知道你测到的哪些电子自旋向上，哪些电子自旋向下。对她来说，她唯一能知道的就是，她自己测量的结果是，大约有一半的电子自旋向上，另一半的电子自旋向下，而且在她看来，哪些电子自旋向上，哪些电子自旋向下是完全随机的。而根据纠缠态 $|\phi^+\rangle$ 塌缩的随机性，即使你根本没有作任何测量，只有她一个人在天边对她的那些电子进行测量，她也会得到完全类似的结论。也就是说，你的测量根本就不能增加任何她从自己的测量中获取的信息，她甚至根本就无从判断你有没有测量。因此她根本就无从得知你们的猫死了。

看来，你根本就无法利用纠缠对超光速地传递信息，甚至你根本就无法利用这些纠缠对来传递信息。但是，你说且慢，以上只考虑了你沿着 z 轴测量的情况，假设你有两种不同的测量选择，要么你沿 z 轴测量你的所有电子，要么你沿着 x 轴测量你的电子，那能不能通过你的测量对她的即时影响将你的这两种不同选择传递给她呢？如果能的话，那你就可以用你的不同选择来代表猫的两种不同状态，从而就能将猫死了的信息超光速地传递给她。

为了下一步的分析，我们首先来看一下 x 方向的测量和 z 方向的测量有什么不同。量子力学的基本原理告诉我们，如果我们沿着 x 轴测量电子的自旋态，那么被测的电子就会塌缩到 x 方向上的两个自旋本征态 $|\uparrow_x\rangle$ 和 $|\downarrow_x\rangle$ 中的某一个，这两个本征态和 z 方向本征态的关系是

$$|\uparrow\rangle=\frac{1}{\sqrt{2}}(|\uparrow_x\rangle+|\downarrow_x\rangle),|\downarrow\rangle=\frac{1}{\sqrt{2}}(|\uparrow_x\rangle-|\downarrow_x\rangle). \tag{2.115}$$

当然，反过来也有

$$|\uparrow_x\rangle=\frac{1}{\sqrt{2}}(|\uparrow\rangle+|\downarrow\rangle),|\downarrow_x\rangle=\frac{1}{\sqrt{2}}(|\uparrow\rangle-|\downarrow\rangle). \tag{2.116}$$

另外，由于 $|\uparrow\uparrow\rangle=|\uparrow\rangle|\uparrow\rangle=\frac{1}{2}(|\uparrow_x\rangle+|\downarrow_x\rangle)(|\uparrow_x\rangle+|\downarrow_x\rangle)=\frac{1}{2}(|\uparrow_x\rangle|\uparrow_x\rangle+|\downarrow_x\rangle|\downarrow_x\rangle+|\uparrow_x\rangle|\downarrow_x\rangle+|\downarrow_x\rangle|\uparrow_x\rangle)=\frac{1}{2}(|\uparrow_x\uparrow_x\rangle+|\downarrow_x\downarrow_x\rangle+|\uparrow_x\downarrow_x\rangle+|\downarrow_x\uparrow_x\rangle)$，类似的 $|\downarrow\downarrow\rangle=\frac{1}{2}(|\uparrow_x\uparrow_x\rangle+|\downarrow_x\downarrow_x\rangle-|\uparrow_x\downarrow_x\rangle-|\downarrow_x\uparrow_x\rangle)$，所以我们可以知道，原来的纠缠态 $|\phi^+\rangle$ 也可以写成

$$|\phi^+\rangle=\frac{1}{\sqrt{2}}(|\uparrow_x\uparrow_x\rangle+|\downarrow_x\downarrow_x\rangle). \tag{2.117}$$

假设你有沿着 z 轴和沿着 x 轴两种测量选择，她也知道你有这两种选择，并且你们约定，如果你沿着 z 轴测量，那就代表猫还活着，如果你沿 x 轴测量那就代表猫死了。由于你们无法正常通信，所以你当然不能直接告诉她你的测量选择是什么，她只能从她自己随后的测量结果中对你的测量方式进行推断，那么她能推断出你的测量方式进而得知猫的死活吗？由于她无法预先知道你的测量方式，所以她自己只能从两种不同测量方式中随机选取一种，比如说，假设她总是选择沿着 z 轴测量她的电子 (由于 z 轴和 x 轴对于我们的纠缠态 $|\phi^+\rangle$ 完全对称，所以对于她选择沿 x 轴测量的所有分析将是完全类似的)。当然，我们总是假定她的测量在你的测量完成之后进行，而你有两种测量方式的选择，你沿着 z 轴测量她也沿着 z 轴测量的情况我们已经分析过了，结论是，她将得到大约一半电子自旋向上，另一半电子自旋向下的结论，并且哪些电子自旋向上哪些自旋向下对于她来说是完全随机的。

现在假设你沿着 x 轴测量，之后她再沿着 z 轴进行测量，我们来看她是否能得到不同的结果。由于纠缠态 $|\phi^+\rangle$ 也可以写成 (2.117) 的形式，所以你沿着 x 轴的测量结果将是，大约有一半你所持有的电子会塌缩到 $|\uparrow_x\rangle$ 态，当然由于纠缠态的性质 (2.117)，这时候她相应的另一个电子也会立即塌缩到 $|\uparrow_x\rangle$ 态，同样，你另半数电子会塌缩到 $|\downarrow_x\rangle$ 态, 这时候她的相应电子也会立即塌缩到 $|\downarrow_x\rangle$ 态，当然这种塌缩是完全随机的。因此在你的测量完成以后，她所持有的每一个电子有 $1/2$ 的概率处在 $|\uparrow_x\rangle$ 态，$1/2$ 的概率处在 $|\downarrow_x\rangle$ 态，当然由于她还不知道你的测量方式，所以对于这个结果她是并不知情的。她只是选择沿着 z 轴进行她的测量，由 (2.116) 式可以知道，这时候如果她测的电子处在 $|\uparrow_x\rangle$ 态，那她的测量将会使得这个电子以 $1/2$ 的可能性塌缩到 $|\uparrow\rangle$，还有 $1/2$ 的可能性塌缩到 $|\downarrow\rangle$, 如果她测的电子处在 $|\downarrow_x\rangle$ 态，结论也是一样的，因此总的来说，对于她所持有的每一个电子，她都有 $1/2$ 的概率测到它自旋向上，$1/2$ 的概率测到它自旋向下。因此，当她完成所有的测量以后，她同样发现，在她所持有的电子中，大约有一半自旋向上，有一半自旋向下，而且这个结果是完全随机的。你已经看到了，她所得到的这个结果和你沿着 z 轴进行测量时她所得到的结果完全一样。也即是说，远在天边的她根本无从推断出你的测量方式，因此当然也就无法得知猫的死活。因此，用这种方式同样无法实现信息的超光速传递。

那么这是不是意味着只要你不通过正常的通信直接告诉她你的测量方式，她就根本无从得知呢？能不能说明，量子纠缠态不仅无法超光速传递信息，甚至根本就无法用来传递信息呢？答案是不能。因为以上的结论都是在假定你和她之间无法进行通常的通信的情形下得到的。如果你们之间能够进行通常的通信，那即使你不直接告诉她你的测量方式，她也有可能推断出这个信息，即是说，如果你们之间能够正常通信，那你们之间共享的量子纠缠对是能够用来传递信息的。我们来看一下这是怎么回事。

假设你们之间能够正常通信，那么即使你不直接告诉她你的测量方式，你也还可以告诉她你的测量结果，也就是说，你也还可以告诉她你所测的每一个电子自旋是沿着你测量轴的正方向，还是沿着你测量轴的反方向，而她得知你的这些信息就能进一步推断出你是沿着 z 轴进行测量，还是沿着 x 轴进行测量。不妨假设你沿着 x 轴进行测量，因此在你的测量完成之后所有电子都塌缩到了某个 x 方向的本征态，这时候如果她也沿着 x 轴进行测量，那么由于纠缠的性质，她会发现她所测的每一个电子结果都和你的测量结果一样 (指

电子自旋是沿着测量轴还是反着测量轴的结果。注意，你已经告诉她这个测量结果了)。但是，如果她是沿着 z 轴进行测量，这时候由于她的电子已经处在 x 方向的某个本征态，而 x 方向的本征态塌缩到 z 方向的本征态是完全随机的，所以她会发现她所测的电子到底是沿着测量轴的正方向还是反方向，这与你告诉她的测量结果之间完全没有关联。所以，通过这种将她的测量结果和你的测量结果相比较的方式，她就能推断出你的测量方式与她的是否一样，从而也就知道了你是沿着 z 轴测量，还是沿着 x 轴进行测量。这样，关于你的测量方式的信息就成功传递给她了。

可见，利用你和她之间共享的量子纠缠对，你的确有可能向她传递额外的信息。但，前提是，你们之间必须可以进行通常的经典的通信，她必须先得到你用经典方式传递过来的信息，才能进一步获知你用纠缠对传递过来的信息。而经典通信肯定是无法超光速的，因此量子纠缠对也无法用来超光速地传递信息。

2.3.2 如何提取纠缠态中的信息

从现在开始，为了和量子信息的语言对接起来，假设我们利用电子的自旋态来实现量子比特，自旋向上态 $|\uparrow\rangle$ 代表 $|0\rangle$ 态，自旋向下态 $|\downarrow\rangle$ 代表 $|1\rangle$ 态。因此之前我们所讨论的两电子自旋的纠缠对 $|\phi^+\rangle$ 就可以重写成两个量子比特的纠缠对 $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$。但是两个量子比特的希尔伯特空间是 $2\times 2 = 4$ 维的，它应该有 4 个正交归一的基矢量，假设我们将 $|\phi^+\rangle$ 选作其中一个基矢量，那么其余三个基矢量可以怎么选择呢？事实上我们可以将其余三个基矢量也选作纠缠态，

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \tag{2.118}$$

$$|\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \tag{2.119}$$

很容易验证这 4 个态的确是正交归一的。为了纪念物理学家贝尔在量子纠缠上所做的开创性工作 (贝尔不等式)，人们通常称这四个态所构成的两量子比特希尔伯特空间矢量基为贝尔基，有时候也称这些量子态为贝尔态。

现在，假定你和你远在天边的她是同一个导师的学生，你们的导师将两比特的重要信息 (这就有四种可能性) 用这四个贝尔态来编码，每种可能性对应一个贝尔态，并且导师将这个贝尔纠缠态中的第一个量子比特发给了南昌的你，而将与之纠缠的另一个量子比特发给了远在天边的她。即是说，如果你们想获知导师给出的具体是什么信息，你们就得确定共享的是四个贝尔态中的哪一个。

你们如何能确定这一点呢？由于相隔遥远，所以你们首先想到的可能就是各自独立地对手中的那个量子比特进行测量。但是，显然的是，只要你们不相互通信，那你们就不可能从这种测量中获取任何信息，因为从这四个贝尔态的表达式 (2.119) 可以知道，不管你们共享的是哪一个贝尔态，你的测量将总是以 1/2 的可能性得到 0，以 1/2 的可能性得到 1，结果完全是随机的，你完全无法从中得知贝尔态的任何具体信息，当然她也一样。这是贝尔态和通常的单个量子比特量子态的根本性不同，对于通常的量子比特，如果你存进一比特的信息，比如将它存为 $|0\rangle$ 态，那你只要测量这个量子比特的值，就一定能把这个信息读取出来。但是，现在你导师明明存进了两个比特的信息，但你和她只要相互保持独立

不进行沟通，就什么信息也读取不出来。也即是说，信息是隐藏的，是存储在你们共享的纠缠对的整体之中。

你可能会问，那这四个贝尔态到底如何区分呢？答案很简单，只要同时测量算符 Z_1Z_2 和算符 X_1X_2 的值，这里写在算符上的下标表示这个算符仅对相应的量子比特进行作用，下标 1 就表示这个算符仅作用在纠缠对的第 1 个量子比特上，下标 2 就表示这个算符仅作用在纠缠对的第 2 个量子比特上，Z 算符其实就是泡利算符 σ^z，它的作用规则是 $Z|0\rangle = |0\rangle, Z|1\rangle = -|1\rangle$，$X$ 算符其实就是泡利算符 σ^x，它的作用规则是 $X|0\rangle = |1\rangle, X|1\rangle = |0\rangle$，当然，作用在不同量子比特上的算符是相互对易的。利用泡利算符的反对易关系，人们很容易验证 Z_1Z_2 和 X_1X_2 是对易的，因此可以有共同的本征矢量，实际上，贝尔基的四个量子态就是它们共同的本征矢量。从贝尔态的表达式 (2.119) 可以很容易验证，Z_1Z_2 的两个不同本征值可以用来区分四个贝尔态是属于 $|\phi\rangle$ 类型，还是属于 $|\psi\rangle$ 类型，而 X_1X_2 的本征值可以用来进一步区分它们在这两个类型中的 $\pm$ 号。但是，正如你已经看到的，要完成这种区分，就需要对纠缠对中的两个量子比特进行联合测量，比如说，对算符 Z_1Z_2 的值进行测量。注意，这和你测 Z_1，她同时测 Z_2 有根本性的不同，因为在联合测量中可以仅仅测得 Z_1Z_2 的值而不必同时知道 Z_1、Z_2 分别是多少。实际上，要同时对 X_1X_2 和 Z_1Z_2 进行这种联合测量，就必须首先将你和她分别持有的量子比特放到一起来。也即是说，只要你们依然分别持有纠缠对中的一个量子比特，那就无法进行这样的联合测量，从而也就无法完整地读出导师存储的两比特信息。

当然，你们依然可以通过交流各自的测量结果，从而获取部分信息，比方说，你和她分别测量了 Z_1 和 Z_2，并将结果进行了比较，由于 Z_1, Z_2 都与 Z_1Z_2 对易，所以根据测量结果你们依然可以推断出你们共享的贝尔态是属于 $|\phi\rangle$ 类型还是 $|\psi\rangle$ 类型，但是，由于 Z_1，Z_2 都与 X_1X_2 反对易，因此你们的测量必定会干扰 X_1X_2 的本征态信息，因此你们也就不可能进一步获知被测贝尔态的 $\pm$ 号。总之，通过这样的方式你们只能获知两比特信息中的 1 比特。

2.3.3 量子密集编码

到此为止，实际上我们还没有看到量子纠缠对有什么神奇的用处。实际上，量子纠缠对能大幅度提高我们的通信能力，当然，代价是要额外消耗掉纠缠对，因此在量子信息中，人们通常将量子纠缠态看成是一种会被消耗的资源。下面就让我们来看一下这种资源的一种神奇应用。

还是假定你在南昌，你的她远在天边，假设你要发送两比特的经典信息给她，如果用经典的方式，通过经典信道进行发送，那你得占用两比特的信道。但是，如果你和她之间建立了量子信道，你就可以通过量子信道发送量子比特，并且，假定你和她在前年就分别持有了某个量子纠缠对中的一个量子比特，不妨假定这个纠缠对处在 $|\phi^+\rangle$ 态吧。如果你们有这些资源，那么你只要占用一比特的量子信道，就能完成对两比特经典信息的发送，这就是所谓的量子密集编码。

你如何做到这一点呢？首先，你们之间得事先约定好一种如何用四个贝尔态对应两比特经典信息的编码方式。其次，你注意到可以对自己所持有的量子比特进行四种不同的幺

正变换，分别为恒等变换 1，变换 Z_1，变换 X_1，以及变换 Z_1X_1(利用泡利算符的性质，你很容易验证这四个算符的幺正性)，变换的结果你也很容易算出来，

$$1|\phi^+\rangle = |\phi^+\rangle, Z_1|\phi^+\rangle = |\phi^-\rangle, X_1|\phi^+\rangle = |\psi^+\rangle, Z_1X_1|\phi^+\rangle = |\psi^-\rangle. \tag{2.120}$$

然后你再根据你们事先约定好的编码方式，将你要发送的信息对应成四个贝尔态中的某一个，并对你的量子比特进行相应的幺正变换。最后，完成了合适的幺正变换以后，你将你手中的那个量子比特通过量子信道发送给天边的她。她本来就持有纠缠对中的另一个量子比特，再接收到你的量子比特以后就拥有整个纠缠对了。为了确定这个纠缠对是四个贝尔态中的哪一个，她只需要同时进行 Z_1Z_2 和 X_1X_2 两种联合测量就可以了。再根据约定好的编码方式她就能得知你所发送的信息。这样，你们就完成了仅用 1 比特的量子信道就传送 2 比特的经典信息了。这就是量子密集编码的基本思想。注意，虽然你们前年就共同持有了这对纠缠对，但那时候的纠缠对中完全不含有你现在要发送的信息，你的信息发送的确是通过现在传送这一个量子比特完成的。

2.3.4 量子隐形传态

量子纠缠对真正神奇的地方在于，只要有足够的量子纠缠对，那么原则上就可能做到将你超空间传送到一个遥远的星球。当然，这只是原则上，实际上这个目标对于我们来说可能永远都遥不可及。但原则上并没有什么物理定律禁止我们实现这种科幻场景。因为，今天在实验室里早就做到将一个量子比特的量子态超空间传送到非常远的地方，这就是量子隐形传态。

还是假设你在南昌，你的她远在天边，你们共同持有一对纠缠的量子比特，不妨称你持有的那个为量子比特 1，她持有的那个为量子比特 2，这两个量子比特处在纠缠态 $|\phi^+\rangle_{12}$(式中我们加上了下标 12，这是因为我们即将引入第 3 个量子比特)。现在，假设你还有另一个量子比特 3, 它处在某个未知的状态 $|\Psi\rangle_3 = \alpha|0\rangle_3 + \beta|1\rangle_3$, 而你想把这第三个量子比特发送给天边的她。你当然可以通过你们之间的量子信道直接进行发送，但你发送的量子比特很有可能被别人拦截，从而带来不可预料的结果。当然，你真正要发送的其实是第三个量子比特的量子态，因为它包含了某些你想发送给她的重要未知信息。

那么你们有没有什么绝对安全的量子信息发送方法呢？有。答案就在于充分利用你们之间共享的纠缠对 $|\phi^+\rangle_{12}$。为了说清楚这一点，首先由贝尔态的定义式 (2.119), 我们很容易得到

$$|00\rangle = \frac{1}{\sqrt{2}}(|\phi^+\rangle + |\phi^-\rangle), |11\rangle = \frac{1}{\sqrt{2}}(|\phi^+\rangle - |\phi^-\rangle) \tag{2.121}$$

$$|01\rangle = \frac{1}{\sqrt{2}}(|\psi^+\rangle + |\psi^-\rangle), |10\rangle = \frac{1}{\sqrt{2}}(|\psi^+\rangle - |\psi^-\rangle). \tag{2.122}$$

其次，我们注意到整个系统的量子态可以重写成

$$
\begin{aligned}
|\Psi\rangle_3|\phi^+\rangle_{12} &= (\alpha|0\rangle_3+\beta|1\rangle_3)\frac{1}{\sqrt{2}}(|00\rangle_{12}+|11\rangle_{12}) \\
&= \frac{1}{\sqrt{2}}[\alpha|00\rangle_{31}|0\rangle_2+\alpha|01\rangle_{31}|1\rangle_2+\beta|10\rangle_{31}|0\rangle_2+\beta|11\rangle_{31}|1\rangle_2] \\
&= \frac{1}{2}[(\alpha|0\rangle_2+\beta|1\rangle_2)|\phi^+\rangle_{31}+(\alpha|0\rangle_2-\beta|1\rangle_2)|\phi^-\rangle_{31}] \\
&+ \frac{1}{2}[(\alpha|1\rangle_2+\beta|0\rangle_2)|\psi^+\rangle_{31}+(\alpha|1\rangle_2-\beta|0\rangle_2)|\psi^-\rangle_{31}],
\end{aligned}
\tag{2.123}
$$

式中第三个等号我们利用了公式 (2.122)，并且由贝尔态的正交归一性可以知道，第三个等号右边的最终表达式中的 4 个态相互正交。因此，你只需要将你的第 3 个量子比特和你所持有的第 1 个量子比特放在一起，并同时对它们进行 Z_3Z_1 和 X_3X_1 的联合测量，测量的结果是，第 3 个量子比特和第 1 个量子比特所构成的系统将会塌缩到它们相应的四个贝尔态中的某一个。然后，你再把你的测量结果用经典的方式发送给天边的她，她如果得知你的结果是 $|\phi^+\rangle_{31}$，那就什么也不需要做，因为从式 (2.123) 可以知道，这时候她的量子比特 2 的量子态已经变成 $\alpha|0\rangle_2+\beta|1\rangle_2=|\Psi\rangle_2$ 态了，如果她得知你的结果是 $|\phi^-\rangle_{31}$，根据 (2.123) 式，那她就对她的量子比特进行 Z_2 的幺正变换，由于 $Z_2(\alpha|0\rangle_2-\beta|1\rangle_2)=(\alpha|0\rangle_2+\beta|1\rangle_2)=|\Psi\rangle_2$，因此变换以后她也将得到正确的量子态，同样的，如果你的结果是 $|\psi^+\rangle_{31}$，那她就对自己的量子比特进行 X_2 的幺正变换，根据 (2.123) 式，结果也将是正确的量子态，而如果你的结果是 $|\psi^-\rangle_{31}$，根据 (2.123) 式，那她就进行 Z_2X_2 的幺正变换，同样会得到正确的量子态。总之，在得知你的测量结果以后，她总可以让自己的量子比特变成 $|\Psi\rangle_2$ 态。这样一来，你的第三个量子比特的未知量子态就成功传送给她了。

这里有几点值得进一步讨论，第一，$|\Psi\rangle$ 态的信息从来也没有在某个量子信道中进行传送！$|\Psi\rangle$ 态的成功传送完全是超空间的，所以称之为隐形传态。第二，你告知给她的测量结果中也完全不包含 $|\Psi\rangle$ 态的信息，否则根据量子力学的基本原理，你的测量就已经对 $|\Psi\rangle$ 态造成了不可逆的扰动，那此后她也就不可能得到 $|\Psi\rangle$ 态了。同时，正因为你发送给她的测量结果中不包含任何 $|\Psi\rangle$ 态的信息，因此即使有人窃听了你们的通信，他也无法获得 $|\Psi\rangle$ 态。第三，整个过程并没有违反量子不可克隆定理，这是因为，在你进行你的测量之后，你就已经摧毁了第三个量子比特原来的态，因此天边的她后来所做的并不是把你的量子比特 3 的态复制一份。第四，直到获知你的测量结果之前，天边的她都还无法得到正确的 $|\Psi\rangle$，并且由于你的测量结果是完全随机的，因此从概率上来说，这时候她只能得到与 $|\Psi\rangle$ 不相关的态。因此，这里也没有信息的超光速传递，因为你告诉她测量结果时采用的经典通信方式当然是无法超光速的。

如果回顾我们上一节所讨论的量子密集编码和这一节讨论的量子隐形传态，人们就会发现，这两个过程的实现都需要消耗纠缠对。这两个过程无论哪一个，当它成功完成之后，原来由两个人共同持有的纠缠对就被消耗掉了，因此对于你和你远在天边的她来说，你们共享的那些纠缠对是一种稀缺资源。

2.3.5 GHZ 态以及为什么爱因斯坦错了

以上我们只讨论了两个量子比特的纠缠，而且实际上我们还只讨论了两个量子比特的那些所谓最大纠缠态，也就是我们所说的贝尔态。人们自然会想到，三个量子比特的纠缠会怎么样呢？这就是我们这一小节想要讨论的问题。当然我们讨论三个量子比特的纠缠，不是因为它有多特殊，而是因为借助于它人们可以了结量子力学发展史上的一段著名公案。

在量子力学刚刚发展起来的时候，爱因斯坦认为量子力学理论是不完备的。因为他认为一个物理量的值总是存在的，或者说任何时候任何物理量总会有一个确定的值，虽然可能因为种种原因你不能测到这个值，但它总存在，爱因斯坦所谓的“即使你没有看月亮，月亮也存在”就是这个意思，由此他指出量子力学里面物理量的值之所以不确定，之所以我们只能测得物理量取值的一个概率分布，是因为我们还缺失了一些信息，爱因斯坦称之为有一些隐变量，如果我们能进一步掌握这些缺失的信息，那量子力学将和经典力学一样，没有任何不确定性。总之，爱因斯坦认为物理量的值总是存在的，量子力学中的不确定性和概率的起源与我们日常概念中的概率起源一样，都是因为我们缺失了一些信息，因此他强烈反对玻尔和海森堡等人的不确定性和概率是世界的内在属性的观念，认为不确定性在量子力学理论中的存在只不过反映了量子力学理论的不完备性，也即是说量子力学理论没有把所有隐变量都包括进来。

那爱因斯坦的观点到底对不对呢？为此人们曾经长期争论不休，直到物理学家贝尔从爱因斯坦的观念出发推导出了著名的贝尔不等式，贝尔说，你只要用实验检验贝尔不等式是否成立，就能判定爱因斯坦到底对不对。只要爱因斯坦对，那贝尔不等式就一定成立，相反，如果玻尔和海森堡等人的观点对，那贝尔不等式就可以被破坏。后来的实验证明，爱因斯坦的确错了，不过虽然爱因斯坦错了，但是他为了否定量子力学的完备性却提出了今天非常重要的量子纠缠的概念。我们前面对量子纠缠的讨论，以及贝尔的开创性工作都是在量子纠缠的概念上进一步发展而来的。下面我们将要描述的，就是另外一个更为直接了当，比贝尔不等式更简单的可以用于判定爱因斯坦是否正确的情形。

我们要讨论的就是由物理学家 Greenberg, Horne, 以及 Zeilinger 提出来并在实验上实现的一种特殊的三量子比特纠缠态，通常称作 GHZ 态。假设我们把这三个量子比特分别标记为 1、2、3，那么 GHZ 态可以写成

$$|GHZ\rangle_{123} = \frac{1}{\sqrt{2}}(|000\rangle_{123} + |111\rangle_{123}). \tag{2.124}$$

利用泡利算符的作用规则，人们很容易验证 $|GHZ\rangle_{123}$ 同时是下面三个相互对易的算符的本征值为 1 的本征态，这三个算符是

$$Z_1Z_2,\ Z_2Z_3,\ X_1X_2X_3, \tag{2.125}$$

利用泡利算符的反对易性，人们很容易验证这三个算符两两对易。现在我们进一步引入泡利算符 σ^y，在这里写作 Y，根据泡利算符的乘法规则我们有 $ZX = iY$。由于 $Y_1Y_2X_3 = -Z_1X_1Z_2X_2X_3 = -(Z_1Z_2)(X_1X_2X_3)$(式中我们已经利用了不同量子比特的算符相互对易的性质)，因此我们可以知道，$|GHZ\rangle_{123}$ 也是 $Y_1Y_2X_3$ 的本征值为 -1 的本征态。类似的，我们

可以得到，当作用在 $|GHZ\rangle_{123}$ 上时，下面的等式必定同时成立

$$Y_1Y_2X_3=-1, X_1Y_2Y_3=-1, Y_1X_2Y_3=-1, X_1X_2X_3=1. \tag{2.126}$$

下面，让我们来假设爱因斯坦的观点也成立，即任何情况下，每一个物理量的值总是存在，我们将三个泡利算符的值分别记作 $v(X), v(Y), v(Z)$，当然由于泡利算符的平方等于 1，所以它们的值只能取 ± 1。请注意，作为物理量的值，$v(X), v(Y), v(Z)$ 都是普通的数，因此它们相互之间当然都相互对易。如果 $v(X), v(Y), v(Z)$ 总是存在，虽然可能由于我们不了解隐变量，缺失了信息，导致我们不能完全确定这些取值到底是多少，但按照爱因斯坦的观点，它们总是存在的，总是有定义的。如此一来，根据 (2.126)，对于 $|GHZ\rangle_{123}$ 态，我们就必定有

$$v(Y_1)v(Y_2)v(X_3)=-1, v(X_1)v(Y_2)v(Y_3)=-1, \tag{2.127}$$

$$v(Y_1)v(X_2)v(Y_3)=-1, v(X_1)v(X_2)v(X_3)=1. \tag{2.128}$$

但是，对于普通的数来说，这些式子是自相矛盾的，因为将前三个式子乘起来并利用 $(v(Y))^2=1$，你很容易得到 $v(X_1)v(X_2)v(X_3)=-1$，这和第四个式子矛盾。

因此，这也就是说，要么量子力学是错的，从而 (2.126) 式不对，要么爱因斯坦就是错的 (也就是说我们不能假定 $v(X)$ 这样的量总是存在)，两者必居其一。到底谁错了呢？实验发现，爱因斯坦错了，(2.126) 式是成立的。因此，这就从实验上否定了物理量的取值必定存在的观点。也从侧面证明了为什么在量子力学中物理量只能表示成算符，因为算符的值当然只在本征态上有定义，对于叠加态，谈算符的取值是没有意义的，因为它根本就不存在。

2.3.6 多体量子纠缠

以上我们只讨论了几个最为重要的两体 (两个量子比特) 和三体 (三个量子比特) 量子纠缠态及其应用。但是读者可以想见，如果我们的系统是一个多体量子系统，比方说量子计算机的多量子比特，再比方说凝聚态物理里面的多体系统 (多原子，多自旋等等)，那么量子纠缠就可能出现在多体之间。可以说多体量子纠缠是整个量子计算和量子信息技术的核心，比如说，在量子信息中，人们总是通过合适的量子编码将量子信息储存在许多量子比特的纠缠态中，这时候由于量子纠缠态的整体性，即使存储信息的某些量子比特出现差错，原来的量子信息也依然能从整体的量子纠缠中得到恢复。不仅如此，近年来多体量子纠缠也被广泛应用于理论凝聚态甚至量子引力的研究。但是，如何一般性地刻画多体之间的量子纠缠现在还是一个没有完全解决的问题，凝聚态物理中常常采用所谓的张量网络来表示多体的量子纠缠态。凝聚态物理学家关心多体量子纠缠的原因在于，近些年来的研究发现，有一些量子多体系统的基态其实是一个量子纠缠态，甚至可能是一个多体长程量子纠缠态，比方说分数量子霍尔效应的基态就是一个这样的长程量子纠缠态。而处在这种长程量子纠缠态的多体系统常常会有所谓的拓扑序，会处在某种拓扑相。这种长程量子纠缠和拓扑序近年来引起了理论物理学家们的极大兴趣。总之，对量子纠缠的深入研究无疑将处于整个物理学研究的核心。

2.4 补充材料：海森堡是怎么想到矩阵相乘的？

只要学习量子力学，你就知道物理量要用算符来表示，但通常的量子力学书只是直接把这作为一个基本假设而不会告诉人们为什么？在历史上，最早产生这个思想的，就是海森堡，是海森堡首先想到物理量要用一张数据表来表示，而更为革命性的是，海森堡首先想到物理量的乘法是一种矩阵乘法。当时的海森堡甚至不知道矩阵的数学概念，所以促使他想到矩阵相乘的肯定不是数学，其背后隐含的是深刻的物理新思想，我们想要讨论的就是这一思想。而矩阵和线性算符当然本质是一回事，物理量满足矩阵乘法从而也就成了线性算符。那么海森堡想到矩阵和矩阵相乘的基本思路是什么呢？理解海森堡的思路将有助于我们深入理解为什么物理量应该表示成线性厄米算符。不仅如此，它更有助于我们深入把握量子力学的基本原理。因此，在阐述完海森堡的原始思想以后，我们还会给出其原创性思想的现代版本，并以此建立起量子力学的基本原理。

不过，这里要说一下，我们不是要还原历史，所以当然会有些贡献在历史上并非来自于海森堡或者不完全属于海森堡，有时候我们把一些结论称作是海森堡的其实只是一种简化说法，想弄清楚历史事实的读者应该去读量子力学发展史方面的专著。但这里要谈的最核心想法无疑是海森堡的，而我们想讨论的，其实也只是海森堡的原创性物理思想。在物理学史上，这样的思想常常会以各种面目出现，可谓影响深远 (有时候甚至会超出量子力学的范围)。比如说，超弦的 M 理论有一个版本就叫做矩阵理论 (所以 M 理论的这个 M 也表示 Matrix)。再比如说，如果细细探究的话，人们可以在费曼的量子力学路径积分表述中察觉到海森堡思想的某种变种 (当然费曼有可能是独立想到这一思想的)，其实，我们这里将要阐述的海森堡思想已经是包含费曼贡献的版本了，至于海森堡在写出他的那些矩阵相乘的时候真正想的是什么，那当然是人们不可能搞清楚的。

2.4.1 原子光谱的一些事实和玻尔的新观念

在玻尔提出他的原子模型之前，人们已经注意到原子光谱的一些惊人事实，比方说呈现出分立的谱线，这是当时的经典物理学完全解释不了的。而那时候人们还发现了原子光谱的另一个令人吃惊的规律，这个规律我们今天已经很少提到，但它对于物理学的发展也许是极其重要的，这规律就是里兹组合规则，这规则是说，人们总是可以合适地将原子光谱的每一条分立谱线都对应到一个正整数对 (m,n)，并且会发现相应的谱线频率 ω_{mn} 满足如下这条规则，

$$\omega_{mn} = \omega_{ml} + \omega_{ln}. \tag{2.129}$$

也就是说，假设有一条谱线是 ω_{ln}，另外还有一条谱线是 ω_{ml}，那么这两条谱线的组合 $\omega_{ml} + \omega_{ln}$ 也必定是一条谱线，对应于整数对 (m,n)。这一规律就是所谓的里兹组合规则。如果人们进一步定义 $\omega_{nm} = -\omega_{mn}$，那么就会发现里兹组合规则对于所有的正整数指标都成立。

当然，在玻尔提出它的原子模型之后，人们就已经完全理解里兹组合规则了，之所以有这样的规则成立，原因很简单，因为根据玻尔的定态跃迁假设，原子从第 m 个定态跃迁到第 n 个定态，辐射 (或者吸收) 的光谱频率为 $\omega_{mn} = (E_m - E_n)/\hbar$，其中 E_n 就是第 n 个

定态的能级，如此一来当然就有里兹组合规则了。对里兹组合规则的解释很简单，那就是核外电子从 n 态跃迁到 m 态所放出的能量，等于它从 n 态跃迁到 l 态所放出的能量加上它从 l 态跃迁到 m 态放出的能量。问题是，给定 n 态和 m 态，里兹组合规则中的 l 态可以是任意的，里兹组合规则对 l 没有任何限制！

这就谈到了玻尔在研究原子光谱尤其是氢原子光谱问题时所引入的革命性新观念了，这就是玻尔的定态概念，以及定态跃迁概念。对于当时的海森堡来说，玻尔氢原子模型中所涉及到的核外电子轨道概念是没有道理的，因为当时的海森堡在物理学哲学上深受玻尔的影响，他们坚持物理学应该用可观测量建立起来，或者换句话说，他们认为物理学理论中所用到的量应该都是可观测的，我们这里不是要探讨这一哲学有没有道理，但是当时年轻的海森堡深信这一哲学，因此他觉得核外电子轨道的概念没有道理，因为不可观测。对于原子物理来说，可观测的是什么呢，是光谱，是谱线频率和谱线的强度。不是电子轨道，电子轨道不过是对太阳系模型的一个类比，而不是原子物理的观测事实。因此海森堡觉得关键是要研究谱线和谱线强度。

但是海森堡也知道，原子光谱的分立性就意味着，虽然核外电子轨道的概念没有道理，但是玻尔的定态和定态跃迁假设必定是对的。没有定态假设就无法解释原子稳定性的问题(因为否则原子就会不断发出电磁辐射从而掉到原子核上去)，但是光有定态假设而没有定态跃迁假设就无法解释光谱线的存在，更何况这两个假设放在一起，的确很好地解释了原子光谱的分立性。

以上也许就是导致海森堡产生他的原创性新思想的一些基本事实，以及他从玻尔那里继承的革命性新观念。当然，还有一个具体的物理学结论对于海森堡验证他的想法非常重要，那就是通过普朗克对黑体辐射公式的推导，海森堡知道，一个线性谐振子的能量是量子化的，并且量子化为 $n\hbar\omega$ 的形式，n 就是线性谐振子的第 n 个能级。海森堡当然也知道，玻尔的定态和定态跃迁概念不只是适用于氢原子，也同样适用于一个可以发出电磁辐射的线性谐振子。其实，很可能在海森堡看来，定态和定态跃迁是普适性的物理学观念。

2.4.2 海森堡的思路

前面说过，海森堡坚持原子物理应该回到可观测的量，也就是谱线频率和谱线强度。玻尔的原子模型当然对谱线频率的解释很不错，尤其是完全推导出了氢原子的所有光谱线，但是玻尔氢原子模型并没有涉及谱线的强度。因此海森堡想推进当时的物理学，那就要研究谱线的强度，核外电子从 m 态跃迁到 n 态时所发出的这条谱线的强度当然取决于单位时间之内它从 m 态自发跃迁到 n 态的概率 A_{nm}，而这是一个当时的物理学家已经提出来了的物理量，只是当时还没有人会计算这个量，海森堡要做的，就是研究如何计算出这个量。但是这个量涉及到原子和电磁场的相互作用，而当时对电磁场的量子理论还只是知道一个光量子假设再加上爱因斯坦的受激辐射理论，对于原子和电磁场如何相互作用的量子理论还一无所知，毕竟，量子力学的正确理论还正有赖于海森堡目前正在进行的这一工作呢。因此海森堡只能转而去看经典理论是如何计算电磁波辐射的，通过经典电动力学我们知道，一个加速运动的电子会辐射电磁波，电动力学也推导出了其辐射功率为

$$P = \frac{e^2}{3\pi\varepsilon_0 c^3}\ddot{\mathbf{x}}^2. \tag{2.130}$$

但是，海森堡知道这个经典公式 (2.130) 肯定是不对的，不仅因为它的表达式中涉及到了核外电子轨道 $\mathbf{x}(t)$ 这样一个在海森堡看来无法观测的量，更重要的是，它和定态以及定态跃迁的概念是相矛盾的。当然，玻尔引入定态和定态跃迁的原因之一本就是为了避免由于 (2.130) 这一经典公式而导致的原子稳定性问题。因此海森堡首先要做的就是改造这个经典公式 (2.130)，使得它和定态以及定态跃迁的概念相容。

根据定态跃迁假设，核外电子只在从一个定态 n 跃迁到另一个定态 m 时才辐射，注意，跃迁总是涉及到两个定态，n 和 m。因此海森堡想到，经典意义上的 $\mathbf{x}(t)$ 是不可观测的，是没有意义的，取而代之的应该是一个类似于 $[\mathbf{x}]_{mn}e^{i\omega_{mn}t}$ 这样的涉及两个定态的量。其中海森堡将这个量对时间的依赖因子写成 $e^{i\omega_{mn}t}$ 的原因在于，通过普朗克对黑体辐射公式的推导，海森堡知道，原子定态跃迁发出电磁辐射的过程可以看成是一种简谐振动，写成指数形式而不是 cos 形式的原因可能是因为当时德布罗意引入物质波的时候就是这样做的，更重要的原因是，后面我们会看到，当将两个物理量相乘时，如果这两个量对时间的依赖因子是 cos 形式，那是不可能满足里兹组合规则的，写成指数形式则很轻易就可以满足这一点。总之，海森堡将 $[\mathbf{x}]_{mn}e^{i\omega_{mn}t}$ 代入公式 (2.130) 就得到

$$P(n\to m)=\frac{e^2\omega_{mn}^4}{3\pi\varepsilon_0 c^3}|[\mathbf{x}]_{mn}|^2. \tag{2.131}$$

注意，现在的这个新公式 (2.131) 不光用到了定态跃迁的概念，它和定态假设也是相容的，由于 $\omega_{nn}=0$, 因此就有 $P(n\to n)=0$，这当然与定态假设一致。经典的轨道 $\mathbf{x}(t)$ 不可观测，但是现在的 $[\mathbf{x}]_{mn}$ 是可以观测的，这是因为，从 n 态跃迁到 m 态的辐射功率等于单位时间辐射概率 A_{mn} 乘以辐射出来的光子能量 $\hbar\omega_{nm}$, 因此由 (2.131) 我们就可以得到 A_{mn} 的表达式，$\mathrm{A}_{mn}\propto|[\mathbf{x}]_{mn}|^2$。$\mathrm{A}_{mn}$ 当然可以直接观测，因此 $[\mathbf{x}]_{mn}$(严格来说是其模 $|[\mathbf{x}]_{mn}|$) 就可以观测。不仅如此，注意到 $e^{i\omega_{nm}t}=e^{-i\omega_{mn}t}=[e^{i\omega_{mn}t}]^*$, 海森堡还进一步规定

$$[\mathbf{x}]_{nm}=[\mathbf{x}]_{mn}^*. \tag{2.132}$$

这样一来，对应经典物理里面的位置坐标，海森堡就引入了一个新的物理量 $[\mathbf{x}]_{nm}$。用我们今天的话来说，这个量是一个厄米矩阵，但除非我们能论证它满足矩阵的乘法，否则我们说它是矩阵就并没有多大意义。更何况，当时的海森堡根本不知道存在矩阵这样的数学概念，对他来说 $[\mathbf{x}]_{nm}$ 不过是一张有无穷行无穷列的数据表，将位置坐标这样的概念变成一张这样的数据表，这本身就蕴含了最高的革命性。由于海森堡是在定态跃迁思想的指引下引入 $[\mathbf{x}]_{nm}$ 的，我们不妨称之为跃迁元。

当我们有了位置物理量 $[\mathbf{x}]_{mn}e^{i\omega_{mn}t}$ 以后，将它对时间求导当然就得到速度 $[\mathbf{v}]_{mn}e^{i\omega_{mn}t}=i\omega_{mn}[\mathbf{x}]_{mn}e^{i\omega_{mn}t}$, 进而就有动量 $[\mathbf{p}]_{mn}e^{i\omega_{mn}t}=im_e\omega_{mn}[\mathbf{x}]_{mn}e^{i\omega_{mn}t}$, 这里 m_e 表示电子质量，而且很容易验证

$$[\mathbf{p}]_{nm}^*=[\mathbf{p}]_{mn}. \tag{2.133}$$

现在，动量也变成了一张跃迁元数据表。因此在海森堡想来，一切物理量都应该变成跃迁元数据表，只有跃迁元才是可观测的。有了位置的跃迁元数据表和动量的跃迁元数据表以后，原则上应该就可以构造出其它一切物理量的跃迁元数据表。由于定态跃迁可以看成是

一种谐振动，海森堡知道所有这些跃迁元应该都具有 $[A]_{mn}e^{i\omega_{mn}t}$ 这样的形式。但是这里有一个重要的问题，比方说动能 T, 它在经典物理里面是这样的 $T=\frac{\mathbf{p}^2}{2m_e}$，这里涉及到动量的平方。更一般地，假设物理量 C 等于两个物理量 A 和 B 的乘积, 即 $C=AB$。那么根据海森堡的跃迁元思想，就需要将 C 的跃迁元 $[C]_{mn}e^{i\omega_{mn}t}$，用 A 的跃迁元 $[A]_{mn}e^{i\omega_{mn}t}$ 和 B 的跃迁元 $[B]_{mn}e^{i\omega_{mn}t}$ 的乘积的形式表达出来。但现在 A 和 B 的所有跃迁元都分别构成了一张无穷行无穷列的数据表，两张这样的数据表该怎么乘，就是海森堡需要解决的大难题。

正是在这里，里兹组合规则，或者等价地说玻尔的公式 $\omega_{mn}=(E_m-E_n)/\hbar$ 给了海森堡必要的指引。由于按照定义 $[C]_{mn}e^{i\omega_{mn}t}=[AB]_{mn}e^{i\omega_{mn}t}$，而按照里兹组合规则，$\omega_{mn}$ 总是可以分裂成两部分的组合，$\omega_{mn}=\omega_{ml}+\omega_{ln}$，所以海森堡想到可以令 $[AB]_{mn}e^{i\omega_{mn}t}=[A]_{ml}e^{i\omega_{ml}t}[B]_{ln}e^{i\omega_{ln}t}$，也就是说，应该有 $[C]_{mn}=[A]_{ml}[B]_{ln}$, 看来问题就解决了。但是海森堡最天才的地方就在这里，他注意到里兹组合规则对中间的 l 态没有任何限制，也就是说 l 有无穷多种可能性，怎么对付这无穷多种可能性呢？海森堡的处理办法很简单，他说，把所有的可能性都加起来！因此海森堡给出

$$[C]_{mn}=\sum_l [A]_{ml}[B]_{ln}. \tag{2.134}$$

当然这就是我们熟悉的矩阵乘法，但是海森堡并没有学过矩阵乘法，对他来说，之所以要这么乘，是因为有无穷多种中间可能性都满足要求，我们需要把所有的中间可能性都加起来！今天看来，这是一个极其原创性的想法，因为它把握住了量子力学相干叠加的本质！

这就是为什么在量子力学里面物理量要变成线性算符的根本原因，因为线性算符的乘法是一种矩阵乘法，而只有矩阵乘法才蕴含着把中间的所有可能性都加起来这一量子相干叠加的本质！当然，今天的量子力学书通常是反过来处理这个问题的，我们是先引入量子态的线性叠加原理，然后再要求物理量保持这一原理，从而只能是线性算符。但是正如我们已经阐述过的，海森堡的革命性是双重的，他不只是想到了要把所有中间可能性都加起来，他更引入了跃迁元的思想 (我们今天常常称之为跃迁矩阵元)，这一思想的好处就是它和观测联系很密切，非常有物理内涵，不像抽象的线性算符。对于所有刚开始学习量子力学的人而言，最难以理解的事情之一无疑就是，物理量怎么就变成算符了呢？从量子态跃迁以及跃迁元入手无疑要好理解得多。当然，跃迁元思想也有其不便之处，人们最好是同时掌握几种不同的思想。

下一小节我们将采用今天的现代观点来对海森堡的这两个原创性想法进行更具一般性的阐述，从而将之与我们通常更熟悉的量子力学原理更密切地联系起来。

但在我们进一步做这样的事情之前，我们首先得像海森堡一样说服自己，跃迁元和矩阵相乘这两个想法的确是 Work 的。海森堡是怎么说服自己的呢？他想到首先将他的思想用到线性谐振子上。也即是说，假设发出电磁辐射的是一个带电线性谐振子 (而不是更复杂的氢原子)，一个线性谐振子从一个定态能级跃迁到另外一个定态能级从而辐射出电磁波。根据普朗克对黑体辐射公式的推导，线性谐振子的能量应该是量子化的，而且其量子化能级具有 $n\hbar\omega$ 这样的形式。由于线性谐振子比氢原子简单得多，所以海森堡可以用线性谐振子的结果来检验他的新思想。结合其他人对线性谐振子的一些研究，海森堡可以用他的新理论推导出两个重要结论：1. 一维线性谐振子的定态能级由 $(n+\frac{1}{2})\hbar\omega$ 给出，正吻

合普朗克推导普朗克公式的时候作出来的假设，而普朗克公式当然和黑体辐射的观测完全吻合，因此海森堡知道他的跃迁矩阵和矩阵乘法应该是对的。2. 通过对一维线性谐振子的研究，海森堡还导出了坐标跃迁矩阵 X 和动量跃迁矩阵 P 之间必定满足

$$[X,P]=XP-PX=i\hbar. \tag{2.135}$$

这当然是量子力学中最著名的结论之一，也是最重要的结论之一。

2.4.3 海森堡思想的现代版本

这一节我们将要给出海森堡思想的现代版本，值得说明的是，这个现代版本当然是很多物理学家工作的结果，比如，其中有一些贡献要归功于费曼，相关的阐述读者可以在费曼的经典论文《Space-time approach to non-relativistic quantum mechanics》以及著作《费曼物理学讲义第三卷》和《量子力学与路径积分》中找到。

首先，在量子力学的世界，一个系统有多种可能性，我们将系统的一组可以相互确定地区分的可能性完备集记作 $\mathscr{S}=\{i,i=1,2,3....\}$。所谓可能性的完备集，我们是指，在确保能收集到这些可能性所列举的信息的同时，你不可能再额外收集到系统更多的信息，只要你采用适当的方式观测这个系统，你总会发现它一定处在这些可能性中的某一个，从某种意义上来说，这一完备集穷尽了你的观测所能收集到的所有可能性。而所谓可以确定地区分，是指你总是可以 100% 地将系统这些可能性中的任何两个区分开来，也就是说，每次你看到系统处于这组可能性中的某一个，那怕只看一次，你也能 100% 地将它和其它可能性区分开来。比方说一个束缚态的氢原子，它的所有用 n,l,m,m_s 这四个量子数标记的可能性集合就构成了这样一组可能性完备集。比方说，当我们仅关心电子自旋的时候，它的向上向下两种自旋可能性 $\{\uparrow,\downarrow\}$ 也构成了这样一组可能性完备集。当然，一个量子比特的 $\{0,1\}$ 两种可能性也构成了这样的可确定区分可能性完备集。而 n 个量子比特的所有从 000....0 到 111...1 的 n 位二进制数也是这样的可能性完备集。但是，一个系统的可确定区分可能性完备集不是唯一的，而是有无穷多组，比方说，你只关心电子的自旋，那么 $\{\uparrow,\downarrow\}$ 两种自旋可能性是这样的一个可能性完备集，但是 $\{\rightarrow,\leftarrow\}$ 同样是一个这样的可能性完备集。

在我们的海森堡思想的现代版本中，这样的一组可能性完备集就相当于海森堡原始思想中的定态。因为很显然，从现在的观点来看，对于氢原子系统，海森堡思想中的定态就是一组用 n,l,m,m_s 这四个量子数标记的不同可能性，这当然属于我们这个可能性完备集的一个例子。

仿照海森堡的跃迁元思想，我们可以说，每个物理量都应该表示成两种可能性之间的跃迁元，不过这两种可能性都必须是从一个预先选定的可确定区分可能性完备集 $\mathscr{S}$ 中选出来的，我们可以将这两种可能性分别标示在跃迁元的左边和右边，也即是说，对于某个物理量 A, 我们可以将它的跃迁元记为 $\langle j|A|i\rangle$，它就相应于我们之前的记号 $[A]_{ji}$。

请大家暂时忘掉算符，忘掉量子态的希尔伯特空间，忘记单位算符 1 的分解定理，也不必把我们这里的 $|i\rangle$ 和 $\langle j|$ 看成是量子态的狄拉克记号，甚至你可以像海森堡一样不知道矩阵的概念，这些都不需要。我们下面也会避免使用诸如希尔伯特空间内积的性质之类

的东西。现在的这个跃迁元记号 $\langle j|A|i\rangle$ 中的 A 就代表物理量 A 本身，左边的 j 和右边的 i 只是表示这个可能性完备集中的两种可能性，之所以把 j 放在左边，而把 i 放在右边，仅仅只是在强调这个跃迁元 $\langle j|A|i\rangle$ 代表的是，在物理量 A 的影响下系统从可能性 i 跃迁到可能性 j 所对应的跃迁元。

同样，假设另有一个物理量 B，那么它的跃迁元就可以记为 $\langle j|B|i\rangle$。海森堡告诉我们，当我们将两个物理量 A 和 B 相乘得出物理量 AB 时，我们是把 B 的从可能性 i 到可能性 j 的跃迁元 $\langle j|B|i\rangle$ 乘上 A 的从可能性 j 到可能性 k 的跃迁元 $\langle k|A|j\rangle$，当然因为我们考虑的是 AB，因此要把 A 的跃迁元放到左边，B 的跃迁元放到右边，进而得到 $\langle k|A|j\rangle\langle j|B|i\rangle$，注意，$A$ 右边的可能性必须要和 B 左边的可能性一样，只有这样我们才可以将这两个跃迁元乘起来。可以直观地将 $\langle k|A|j\rangle\langle j|B|i\rangle$ 理解成，系统在物理量 B 的作用下先从可能性 i 跃迁到了中间可能性 j，然后再在物理量 A 的作用下接着从可能性 j 跃迁到可能性 k，因此总的来说 $\langle k|A|j\rangle\langle j|B|i\rangle$ 相应于系统在 AB 的联合作用下从可能性 i 到可能性 k 的一个跃迁。但它只是对 AB 的跃迁元 $\langle k|AB|i\rangle$ 的贡献之一，海森堡告诉我们，为了得到完整的跃迁元 $\langle k|AB|i\rangle$，我们需要把所有的中间可能性 j 都加起来，因此就有

$$\langle k|AB|i\rangle = \sum_j \langle k|A|j\rangle\langle j|B|i\rangle. \tag{2.136}$$

这就是海森堡告诉我们的一条量子力学基本原理，它有两个要点：第一，物理量要用左右两种可能性之间的跃迁元来表示，物理量相乘就是将相应的跃迁元首尾相乘。第二，为了得到正确的联合作用物理量的跃迁元，我们需要对所有的中间可能性进行求和。这就是海森堡给出来的跃迁元概念及其乘法规则。

那么物理上可观测的是什么呢？一般来说，并不是跃迁元本身，而是跃迁元的模方。因此，由于海森堡的跃迁元乘法规则 (2.136)，如果你测量的是 $|\langle k|AB|i\rangle|^2$，那你就会发现不同的中间可能性之间会产生干涉，因此，海森堡的对所有中间可能性进行求和，实际上就是在量子力学中处于最本质地位的量子力学的相干叠加。

另外，仔细研究海森堡的跃迁元思想及其乘法规则，你会发现，实际上跃迁元左右两边的可能性根本不必属于同一个可确定区分可能性完备集。只需要注意在相乘的时候让它们首尾相乘 (即乘积前面一个跃迁元右边的可能性要和后一个跃迁元左边的可能性相对应) 就可以了。比方说，我们完全可以设想有三个不同的可确定区分可能性完备集，$\mathscr{S} = \{i, i = 1,2,3...\}$,$\mathscr{S}' = \{j', j' = 1,2,3...\}$, 以及 $\mathscr{S}'' = \{k'', k'' = 1,2,3...\}$, 我们同样可以写出类似于 (2.136) 的跃迁元乘法规则

$$\langle k''|AB|i\rangle = \sum_{j'\in\mathscr{S}'} \langle k''|A|j'\rangle\langle j'|B|i\rangle. \tag{2.137}$$

上面这种推广虽然看起来很显然，但其实非常重要，因为它使得我们可以定义不同时刻的可确定区分可能性完备集 (而不是原来的相同时刻的可确定区分可能性完备集)，比方说，我们可以让上面的 $\mathscr{S}$ 定义在某个初始时刻 t_0，$\mathscr{S}'$ 定义在某个中间时刻 t_1，而 $\mathscr{S}''$ 则定义在某个末尾时刻 t_2。相应的，我们可以设想物理量 A 和物理量 B 分别为某个实验装置，其中 B 只在初始时刻 t_0 和中间时刻 t_1 的期间起作用，而实验装置 A 只在 t_1 和 t_2 的

期间起作用。这种推广以后的海森堡跃迁元乘法规则就不是在定义通常的物理量乘积，而是将不同时刻的物理量乘起来。

当然，海森堡的跃迁元乘法规则并没有包含所有的量子力学基本原理。除了还需要给出具体如何计算跃迁元的办法以外，至少还有一条重要的原理我们还没有涉及到。那就是，当海森堡说到将所有的中间可能性加起来的时候，是默认我们没有对这些中间可能性进行测量，我们不确定系统经过 B 和 A 的连续作用从 i 最后跃迁到 k'' 时中间都经过了些什么，也就是说，我们不确定在 B 起作用之后，A 起作用之前，系统到底经过的是什么中间可能性，这就是我们所谓的没有对中间可能性 j' 进行测量。如果我们在这中间，对中间可能性 j' 进行了测量，那我们定义物理量乘积 AB 的时候就不能将所有的中间可能性都加起来了，因为我们已经知道了（假定我们的测量是完全有效的）系统都经过了什么中间可能性了。相反，这时候我们需要对每一种中间可能性 j' 分别定义一个 $(AB)_{j'}$，它简单地由下式给出

$$\langle k''|(AB)_{j'}|i\rangle = \langle k''|A|j'\rangle\langle j'|B|i\rangle. \tag{2.138}$$

这就是量子力学另一条最基本的原理，它说的就是，一旦我们可以确定系统经过的中间可能性具体是什么，我们就不能再把所有的中间可能性都加起来了。为了看清楚这一条原理的含义，假设我们计算跃迁元的模方测量值 $|\langle k''|(AB)_{j'}|i\rangle|^2$，我们就会发现，中间可能性的相干叠加消失了。事实上这时候 $|\langle k''|(AB)_{j'}|i\rangle|^2 = |\langle k''|A|j'\rangle|^2|\langle j'|B|i\rangle|^2$。由于我们可以对任意的物理量 A 和 B 进行这样的考察，因此我们不妨记 $B|i\rangle = |\psi\rangle$，$\langle k''|A = \langle\phi|$，那刚才的结果就变成了 $|\langle\phi|j'\rangle|^2|\langle j'|\psi\rangle|^2$，这就是我们通常更熟悉的语言所谓的测量导致量子态的塌缩，这一结果就可以解释成，已知系统初末两个时刻的状态 $|\psi\rangle$ 和 $|\phi\rangle$，测量导致系统在中间时刻塌缩到 j' 可能性的概率。这当然是一条关于测量的量子力学基本原理。

一般来说，两个物理量相乘，结果当然是一个新的物理量。但是，有一种特别重要的物理量例外，那就是单位 1，1 和 1 相乘结果当然还是 1。这也就是说，1 的跃迁元乘以 1 的跃迁元，结果还是 1 的跃迁元。注意，由于跃迁元左右两边可以选择不同的可确定区分可能性完备集，所以 1 的跃迁元一般来说并不是平凡的。假设我们有两个不同的可能性完备集，$\mathscr{S} = \{i, i = 1,2,3...\}$ 和 $\mathscr{S}' = \{j', j' = 1,2,3...\}$，那么 $\langle j'|i\rangle = \langle j'|1|i\rangle$ 就是一个 1 的跃迁元。通常我们会把 1 的跃迁元叫做跃迁幅，$\langle j'|i\rangle$ 就是一个跃迁幅。当然，我们也可以让 1 的跃迁元左右两边的可能性属于同一个可能性完备集，那这时候的跃迁幅是什么呢？很显然，根据可确定区分的概念，属于同一可能性完备集的两个不同可能性之间仅在 1 的作用下是不可能相互跃迁的，因此即有 $\langle j|i\rangle = 0, j \neq i$。另外，很显然可能性 i 到其自身的跃迁幅应该为 1，因此我们即有

$$\langle j|i\rangle = \delta_{ji}, \tag{2.139}$$

这就是通常人们所说的正交归一性。

由于单位 1 自乘结果还是 1，所以跃迁幅乘以跃迁幅结果还是跃迁幅，只不过要将所有中间可能性都加起来。这就是跃迁幅特有的美妙性质，费曼的量子力学路径积分表述就是利用跃迁幅的这一美妙性质建立起来的。我们可以在前文的公式 (2.137) 中将物理量 A

和 B 都取作 1，从而将这一性质写成，

$$\langle k''|i\rangle = \sum_{j'\in\mathscr{S}'} \langle k''|j'\rangle\langle j'|i\rangle. \tag{2.140}$$

另外，我们前面已经说过了，跃迁幅的模方，比如说 $|\langle j'|i\rangle|^2$，就是测量导致系统从可能性 i 塌缩到可能性 j' 的概率，由于所有不同的 j' 可能性都可以确定地区分并且完备，因此总概率等于 1 的概率守恒就要求

$$1 = \sum_{j'\in\mathscr{S}'} |\langle j'|i\rangle|^2 = \sum_{j'\in\mathscr{S}'} \langle j'|i\rangle^*\langle j'|i\rangle. \tag{2.141}$$

另外，利用正交归一性和 (2.140), 我们又有

$$1 = \langle i|i\rangle = \sum_{j'\in\mathscr{S}'} \langle i|j'\rangle\langle j'|i\rangle. \tag{2.142}$$

比较这两个式子我们就有

$$\langle i|j'\rangle = \langle j'|i\rangle^*, \tag{2.143}$$

这就是我们熟悉的希尔伯特空间内积的基本性质，因此终于我们知道了，跃迁幅代表的是某种希尔伯特空间内积。

现在，我们来考虑另外一个有意思的情形，即任何物理量 A 与单位 1 相乘，结果当然还是 A，即 $A = A\cdot 1$。假设我们取两组不同的可确定区分可能性完备集 $\{i, i = 1,2,3....\}$ 和 $\{m, m = 1,2,3...\}$，那这就告诉我们

$$\langle i|A|m\rangle = \sum_j \langle i|A|j\rangle\langle j|m\rangle. \tag{2.144}$$

其中 $\langle i|A|j\rangle$ 左右两边的可能性既然是属于同一可能性完备集，那它就是通常的矩阵的矩阵元 (因此这时候跃迁元的概念就变成了我们通常所说的跃迁矩阵元)。上面这个式子就告诉我们，可以把 $|m\rangle$ 看成是某个线性空间里的列矢量，其第 j 分量就是 $\langle j|m\rangle$，而物理量 A 就可以看成是线性算符，其在 $|m\rangle$ 上的作用就是线性的矩阵相乘。算符 A 所对应的这个矩阵的第 i 行第 j 列就是 $\langle i|A|j\rangle$(我们有时候也记作 $\widehat{A}_{ij}$)。因此我们就可以将跃迁元的语言翻译成线性叠加的量子态和线性算符的语言。这也就是为什么我们会采用狄拉克符号这样的记号来标记跃迁元的基本原因，因为最终我们会发现跃迁元及其乘法规则和通常的量子态及线性算符，它们只不过是表述同一套量子力学原理的不同语言。



3. 坐标表象与波动力学

本章主要是从量子力学的基本原理出发建立起薛定谔方程在坐标表象下的表示形式，也就是波动力学。我们主要讨论单个微观粒子的情形，但是类似的处理当然也可以推广到多粒子情形。在本章中，我们讨论了不考虑自旋的非相对论薛定谔波动方程，讨论了磁场中带电粒子的薛定谔波动方程，也讨论了自旋 1/2 带电粒子的薛定谔波动方程。

作为波动力学的应用，本章中我们还会讨论超导物理的有效理论，讨论超导的磁通量子化，以及超导约瑟夫森结的基本物理规律。我们也将讨论带电粒子与磁场耦合的薛定谔波动方程的规范不变性，讨论 Aharonov-Bohm 效应以及磁单极子。此外，我们还会用波动力学的半经典近似来讨论量子隧穿效应，超导约瑟夫森结就是作为量子隧穿效应的一个例子而讨论的。

上一章我们已经详细地阐明了量子力学的基本原理，而在前面的第一章中我们也简单介绍了非相对论粒子的波动力学。实际上，后者可以看成是前者对于非相对论粒子在坐标表象下的表示。如何应用坐标表象的概念建立波动力学是我们这一章所要讨论的核心内容之一。

让我们先概括一下这一章将要用到的一些普遍公式。首先，任何一个量子态 $|\psi\rangle$ 随时间的演化，可以用如下一般形式的薛定谔方程描述，

$$i\hbar\partial_t|\psi(t)\rangle = H|\psi(t)\rangle. \tag{3.1}$$

H 就是所谓的哈密顿算符，它对应的是能量这个物理量。

其次，由共轭转置的基本规则我们很容易知道，一个任意的算符 B 和它的厄米共轭算符 $B^\dagger$ 之间满足，

$$\langle\phi|B|\psi\rangle^* = \langle\psi|B^\dagger|\phi\rangle. \tag{3.2}$$

最后，人们总是可以将任意可观测量 A 的线性独立本征态 $\{|i\rangle, i = 1,2,3...\}$ 取成相互正交归一的，即满足

$$\langle i|j\rangle = \delta_{ij}. \tag{3.3}$$

不仅如此，由于对任何量子态进行测量，总是可以得到物理量 A 的某个值，因此 A 的这一组正交归一本征态还必然是完备的，也就是说，任意态 $|\psi\rangle$ 都可以写成 $\{|i\rangle, i = 1,2,3...\}$ 的线性叠加的形式。正因为如此，这一组本征态 $\{|i\rangle, i = 1,2,3...\}$ 必然可以作为希尔伯特空间的一组正交归一矢量基，因此有

$$\sum_i |i\rangle\langle i| = 1, \tag{3.4}$$

这里的 1 就是单位算符，或者说恒等算符。

3.1 单个非相对论粒子

这一节我们将主要以一维运动为例推导单个非对论粒子的波动力学方程。为此，除了量子力学的一般原理以外，我们还需要一个额外的基本假设，即著名的德布罗意关系式，也就是，一个动量为 p 的微观粒子可以用一个波数 $k = p/\hbar$ 的平面波来描述。或者说，动量为 p 的微观粒子，其波函数为

$$\psi_p(x) = c\exp(ipx/\hbar). \tag{3.5}$$

当然有些量子力学书是用一些其它假设来取代这个假设的，但是采用这个假设作为出发点有如下一些好处：首先，德布罗意的这个波粒二象性关系有直接的实验验证，也可以说它是量子力学中被实验检验得最多的一个基本事实，这不像其它的一些代替性假设，它们和实验之间的关系往往是间接的。第二，这么处理也是最符合量子力学的历史发展过程的。第三，它也许是最不抽象最直观的。

3.1.1 位置本征态和动量本征态

对于一个非相对论粒子来说，它的基本物理量就是坐标和动量，但是德布罗意关系 (3.5) 告诉我们，一个动量为 p 的粒子，其位置是完全不确定的，也就是说坐标和动量不可同时测定。因此为了选定希尔伯特空间的一组矢量基，我们要么取一组位置本征态 $|x\rangle$，要么取一组动量本征态 $|p\rangle$。所谓的坐标表象，就是将微观粒子的量子态 $|\psi\rangle$ 在位置本征态 $|x\rangle$ 中进行展开。根据一般性公式 (3.4)，位置本征态必然要满足

$$\int dx|x\rangle\langle x| = 1, \tag{3.6}$$

这里因为位置是一个连续变量，所以我们将公式 (3.4) 中离散的求和改成了积分。任意量子态 $|\psi\rangle$ 必然可以展开成，$|\psi\rangle = \int dx|x\rangle\langle x|\psi\rangle$，展开系数 $\langle x|\psi\rangle$ 就叫做微观粒子的波函数 $\psi(x)$，即

$$\psi(x) = \langle x|\psi\rangle. \tag{3.7}$$

值得说明一下的是，与方程 (3.6) 相应的，位置本征态必然要满足 $\langle x|y\rangle = \delta(x-y)$ 这样的正交归一关系。这是由于 $\delta(z-x) = \langle z|x\rangle = \int dy\langle z|y\rangle\langle y|x\rangle = \int dy\delta(z-y)\delta(y-x)$ 是一个恒等式，因此只有这样的正交归一方式才能与 (3.6) 相容。

利用波函数的概念，我们就可以计算两个量子态之间的内积

$$\langle\phi|\psi\rangle = \int dx\langle\phi|x\rangle\langle x|\psi\rangle = \int dx\phi^*(x)\psi(x) = (\phi,\psi), \tag{3.8}$$

这里的第一个等号是在行矢量 $\langle\phi|$ 与列矢量 $|\psi\rangle$ 中间插入了一个单位算符 1，并利用了等式 (3.6)，而第二个等号是利用了 $\langle\phi|x\rangle = \langle x|\phi\rangle^* = \phi^*(x)$。另外，$(\phi,\psi)$ 是坐标表象下的一个常用记号，它就定义成 $(\phi,\psi) = \int dx\phi^*(x)\psi(x)$，通常称作两个波函数的内积，很显然，它也满足

$$(\phi,\psi)^* = (\psi,\phi). \tag{3.9}$$

由于德布罗意告诉我们的波函数 (3.5) 是一个动量确定的粒子的波函数，也就是动量本征波函数，因此如果将波函数定义式 (4.61) 中的 $|\psi\rangle$ 态取作动量本征态 $|p\rangle$ 即有，

$$\langle x|p\rangle = c\exp(ipx/\hbar). \tag{3.10}$$

同样，作为希尔伯特空间的一组矢量基，动量本征态 $|p\rangle$ 当然也要满足类似的

$$\int dp|p\rangle\langle p| = 1. \tag{3.11}$$

由 $\delta(x-y) = \langle x|y\rangle = \int dp\langle x|p\rangle\langle p|y\rangle = \int dp\langle x|p\rangle\langle y|p\rangle^* = \int dp|c|^2\exp(ip(x-y)/\hbar)$
$= |c|^2 2\pi\hbar\delta(x-y)$(式中第四个等号我们代入了 (3.10) 式，最后一个等号则利用了 $\int dke^{ikx} = 2\pi\delta(x)$)，我们可以知道，$c$ 应该取成 $1/\sqrt{2\pi\hbar}$，因此，

$$\langle x|p\rangle = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}\exp(ipx/\hbar). \tag{3.12}$$

从这个式子我们可以很容易注意到

$$-i\hbar\partial_x\langle x|p\rangle = p\langle x|p\rangle. \tag{3.13}$$

3.1.2 算符在坐标表象下的表示

我们已经将抽象的量子态 $|\psi\rangle$ 表示成了坐标表象下的波函数 $\psi(x)$，现在来考察算符在坐标表象下的表示。对于一维的非相对论粒子，它最基本的算符就是位置算符 X 和动量算符 P，其它物理量的算符都可以由这两个算符通过算符乘法、加法、以及乘上一个常数倍数这样的操作生成，也就是说，都具有 $\mathscr{O}(X,P)$ 这样的形式。这里 $\mathscr{O}(X,P)$ 表示某个 X 和 P 的函数，当然我们还要求它有厄米性。

对于任何一个希尔伯特空间的抽象算符 A，我们定义它在坐标表象中的表示为 $\widehat{A}$, $\widehat{A}$ 由下式定义

$$\widehat{A}\psi(x) = \langle x|A|\psi\rangle, \tag{3.14}$$

式中 $|\psi\rangle$ 是任意一个量子态，$\psi(x)$ 是它的波函数。很显然，如果抽象算符 A 作用在态 $|\psi\rangle$ 上得到 $|\phi\rangle$，即 $A|\psi\rangle = |\phi\rangle$。那么根据我们的定义 (3.14)，就有 $\widehat{A}\psi(x) = \langle x|A|\psi\rangle = \langle x|\phi\rangle = \phi(x)$，即有 $\widehat{A}\psi(x) = \phi(x)$。也就是说，$\widehat{A}$ 在波函数上的作用与抽象算符 A 在相应狄拉克符号上的作用完全对应，因此它的确够成了抽象算符 A 的表示，也因此，只要 A 是一个线性算符，那么相应的 $\widehat{A}$ 也必然是一个线性算符。假设 A 的本征方程为 $A|u_n\rangle = \lambda_n|u_n\rangle$，那么相应就有 $\widehat{A}$ 的本征方程，$\widehat{A}u_n(x) = \lambda_n u_n(x)$，两者的本征值和本征矢量都完全对应。

不仅如此，我们还可以计算任意算符 A 在任意两个态 $|\psi\rangle$ 和 $|\phi\rangle$ 上的矩阵元 $\langle\phi|A|\psi\rangle$，$\langle\phi|A|\psi\rangle = \int dx\langle\phi|x\rangle\langle x|A|\psi\rangle = \int dx\phi^*(x)\widehat{A}\psi(x) = (\phi,\widehat{A}\psi)$，即有

$$\langle\phi|A|\psi\rangle = (\phi,\widehat{A}\psi). \tag{3.15}$$

那么，在坐标表象中我们怎么定义一个算符 $\widehat{A}$ 的厄米共轭算符 $\widehat{A}^\dagger$ 呢？很简单，由于 $\widehat{A}$ 是抽象算符 A 在坐标表象下的表示，我们就相应地定义 $\widehat{A}^\dagger$ 为 $A^\dagger$ 在坐标表象下的表示，即，

$$\widehat{A}^\dagger\psi(x) = \langle x|A^\dagger|\psi\rangle. \tag{3.16}$$

因此我们同样有 $\langle\psi|A^\dagger|\phi\rangle = (\psi,\widehat{A}^\dagger\phi)$，而由公式 (3.2) 我们可以知道 $\langle\psi|A^\dagger|\phi\rangle^* = \langle\phi|A|\psi\rangle$，因此 $(\psi,\widehat{A}^\dagger\phi)^* = (\phi,\widehat{A}\psi)$。又由于根据波函数内积所满足的 (3.9)，$(\psi,\widehat{A}^\dagger\phi)^* = (\widehat{A}^\dagger\phi,\psi)$，因此我们就有

$$(\widehat{A}^\dagger\phi,\psi) = (\phi,\widehat{A}\psi). \tag{3.17}$$

这就是我们定义的 $\widehat{A}^\dagger$ 所满足的基本关系式，它和人们在波函数空间中对一个算符的厄米共轭算符的通常定义完全一样。根据这个结果我们可以知道，任何一个抽象的厄米算符，其在坐标表象中的表示也必然是一个厄米算符！特别的，位置算符 X 的表示 $\widehat{X}$ 必然是一个厄米算符，动量算符 P 的表示 $\widehat{P}$ 也必然是一个厄米算符。我们下面要做的，就是根据我们对算符表示的定义 (3.14)，求出 $\widehat{X}$ 和 $\widehat{P}$ 的具体形式。

3.1.3 位置算符，动量算符，哈密顿算符

我们先来看位置算符。由于 $|x\rangle$ 是位置本征态，相应的坐标值是 x，所以 $X|x\rangle = x|x\rangle$，也即 $\langle x|X = x\langle x|$。因此根据我们的定义 (3.14)，$\widehat{X}\psi(x) = \langle x|X|\psi\rangle = x\langle x|\psi\rangle = x\psi(x)$，这也就是说，在坐标表象中，位置算符 $\widehat{X}$ 其实就是给波函数乘上 x，你也可以简记为 $\widehat{X} = x$。

对于动量算符，由定义 (3.14)，$\widehat{P}\psi(x)=\langle x|P|\psi\rangle=\int dp\langle x|P|p\rangle\langle p|\psi\rangle=\int dpp\langle x|p\rangle\langle p|\psi\rangle=\int dp-i\hbar\partial_x\langle x|p\rangle\langle p|\psi\rangle=-i\hbar\partial_x\int dp\langle x|p\rangle\langle p|\psi\rangle=-i\hbar\partial_x\psi(x)$ (式中第三个等号我们利用了 $|p\rangle$ 作为动量本征态的定义，即 $P|p\rangle=p|p\rangle$，第四个等号我们利用了等式 (3.13), 而且整个式子我们还利用了 (3.11)), 即 $\widehat{P}\psi(x)=-i\hbar\partial_x\psi(x)$, 由此可见，坐标表象下的动量算符 $\widehat{P}$ 为

$$\widehat{P}=-i\hbar\partial_x. \tag{3.18}$$

由 $\widehat{X}$ 和 $\widehat{P}$ 的表达式，我们容易得到 $[\widehat{X},\widehat{P}]=(-i\hbar)(x\cdot\partial_x-\partial_x\cdot x)=(-i\hbar)(x\cdot\partial_x-1-x\cdot\partial_x)=i\hbar$, 即 $[\widehat{X},\widehat{P}]=i\hbar$。由于 $\widehat{X}$ 和 $\widehat{P}$ 分别是抽象算符 X 和 P 的表示，因此这也就说明了

$$[X,P]=i\hbar. \tag{3.19}$$

抽象算符写法的好处是，它不依赖于具体的表象。

前面在第二章的习题中我们证明了，如果两个可观测物理算符 A,B 不对易，那么在任意量子态 $|\psi\rangle$ 上将有，

$$\overline{(\Delta A)^2}\cdot\overline{(\Delta B)^2}\geq\frac{1}{4}\left|\overline{[A,B]}\right|^2. \tag{3.20}$$

我们将这个结果应用于 $A=X,B=P$，并记 $\delta X=\sqrt{\overline{(\Delta X)^2}}$, $\delta P=\sqrt{\overline{(\Delta P)^2}}$ 分别为 X,P 的测量不确定度，那么利用 $[X,P]=i\hbar$，我们即有

$$\delta X\delta P\geq\frac{\hbar}{2}. \tag{3.21}$$

这就是著名的海森堡不确定关系。它告诉我们粒子的位置和动量不可同时测定，两者测量不确定度的乘积必须大于等于 $\hbar/2$。

假设记哈密顿算符 H 在坐标表象下的表示为 $\widehat{H}$, 那么由于一维非相对论粒子的总能量等于动能与势能之和，即为 $\frac{p^2}{2m}+V(x)$，所以很容易猜到相应的哈密顿算符应该是 $H=\frac{P^2}{2m}+V(X)$, 也就有 $\widehat{H}=\frac{\widehat{P}^2}{2m}+V(\widehat{X})$, 代入上面我们找到的 $\widehat{X}$ 和 $\widehat{P}$ 的表达式，即有 $\widehat{H}=-\frac{\hbar^2}{2m}\partial_x^2+V(x)$。又由 $\mathrm{i}\hbar\partial_t\psi(x,t)=\mathrm{i}\hbar\partial_t\langle x|\psi(t)\rangle=\langle x|\mathrm{i}\hbar\partial_t|\psi(t)\rangle=\langle x|H|\psi(t)\rangle=\widehat{H}\psi(x,t)$(式中第三个等号我们使用了一般形式的薛定谔方程 (3.1)), 我们就有 $\mathrm{i}\hbar\partial_t\psi(x,t)=\widehat{H}\psi(x,t)$。由此就可以知道一维非相对论粒子在坐标表象下的薛定谔方程为

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t)=\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)\right]\psi(x,t). \tag{3.22}$$

这就是最广为人知的薛定谔波动方程。这样我们就从作为量子力学基本原理之一的最一般形式的薛定谔方程 (3.1) 推导出了薛定谔波动方程，当然，在这个过程中，为了得到动量算符的具体表示形式，我们需要额外用到德布罗意关系 (3.5)。

人们很容易把上面的推理过程推广到三维空间，这时候动量算符 $\mathbf{P}$ 的坐标表象表示就是 $\widehat{\mathbf{P}}=-i\hbar\nabla$。进一步写出相应的坐标表象薛定谔方程是直接了当的，结果是，

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t)=\left[-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})\right]\psi(\mathbf{x},t). \tag{3.23}$$

当然，人们也很容易得到三维空间位置算符和动量算符之间的对易关系，我们可以用分量形式将它们概括如下

$$[X_i, X_j] = [P_i, P_j] = 0, \quad [X_i, P_j] = i\hbar\delta_{ij}. \tag{3.24}$$

式中指标 $i, j = 1,2,3$，分别对应 x, y, z 分量。

3.1.4 概率流密度

在第二章中我们已经看到，薛定谔方程有幺正性，会保证概率守恒。而且我们也知道，对于单粒子波动力学，我们还可以引入局域的概率密度概念，即 $\rho(x,t) = |\psi(x,t)|^2$。因此一个自然的问题是，在单粒子波动力学中有没有局域的概率守恒？换言之，我们想知道存不存在一个概率流密度 $J(x,t)$，使得下面的连续性方程成立

$$\partial_t \rho + \partial_x J = 0. \tag{3.25}$$

$J(x,t)$ 代表的就是单位时间沿着 x 轴正方向流过 x 点的概率。如果能找到这样的 $J(x,t)$，我们就说单粒子波动力学是局域概率守恒的。

实际上，$J(x,t)$ 的确存在，为此我们将单粒子薛定谔波动方程及其复共轭方程重写如下

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t) &= \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi(x,t). \\ -\mathrm{i}\hbar\frac{\partial}{\partial t}\psi^*(x,t) &= \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi^*(x,t). \end{aligned} \tag{3.26}$$

然后我们用 ψ^* 乘以上式的第一个方程，用 ψ 乘以上式的第二个方程，再将结果相减，就可以得到

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial}{\partial t}(\psi^*\psi) &= \left(-\frac{\hbar^2}{2m}\right)\left[\psi^*\frac{\partial^2}{\partial x^2}\psi - \psi\frac{\partial^2}{\partial x^2}\psi^*\right] \\ &= \frac{\partial}{\partial x}\left[-\frac{\hbar^2}{2m}(\psi^*\partial_x\psi - \psi\partial_x\psi^*)\right]. \end{aligned} \tag{3.27}$$

很显然，最终的这个方程的确能够写成连续性方程 (3.25) 的形式，只需要我们取

$$\begin{aligned} J(x,t) &= \frac{-i\hbar}{2m}\left(\psi^*\partial_x\psi - \psi\partial_x\psi^*\right) \\ &= \frac{1}{2m}\left(\psi^*\widehat{P}\psi - \psi\widehat{P}\psi^*\right). \end{aligned} \tag{3.28}$$

这就是一维非相对论粒子概率流密度的表达式。

我们也很容易将上面的结果推广到三维情形，这时候概率流密度将是一个矢量 $\mathbf{J}(\mathbf{x},t)$，代表的是在这个矢量的横截面上，单位时间之内通过单位横截面积的概率。这时候相应的局域概率守恒方程就是

$$\partial_t \rho + \nabla\cdot\mathbf{J} = 0. \tag{3.29}$$

为了得到 $\mathbf{J}(\mathbf{x},t)$ 的表达式，人们只需将 (3.28) 式中的一维动量算符替换成三维动量算符 $\widehat{\mathbf{P}}=-i\hbar\nabla$, 从而即有

$$\mathbf{J}(\mathbf{x},t)=\frac{-i\hbar}{2m}\left(\psi^*\nabla\psi-\psi\nabla\psi^*\right). \tag{3.30}$$

读者可以直接写出三维的薛定谔波动方程，然后按照和一维情形完全一样的推导来证明这一表达式的确是对的。

3.1.5 习题

1. 假设某一维运动的自由粒子 (即势能等于 0) 在 $t=0$ 时刻的波函数为高斯型波包 $\psi(x,0)=Ce^{-\frac{1}{4}x^2/\sigma^2}$，这里 σ 为某个非 0 实数。(1). 请将这个波函数归一化，定出常数 C。(2). 假设定义动量空间波函数为 $\psi(p,t)=\langle p|\psi(t)\rangle$，请计算 $\psi(p,0)$。(3). 请导出 $\psi(p,t)$ 所满足的薛定谔波动方程，并根据上一问求出来的 $\psi(p,0)$ 解出 $\psi(p,t)$。(4). 请根据上一问求出来的 $\psi(p,t)$ 求出这个粒子在 t 时刻的波函数 $\psi(x,t)$。(5). 请计算 $\psi(x,t)$ 态上粒子位置坐标 X 以及 X^2 的期望值，再计算粒子动量 P 和 P^2 的期望值，进而验证海森堡不确定关系。

2. 这道题是两个基本的计算。(1). 假设粒子的波函数为 $\psi(\mathbf{x})=Ce^{i\mathbf{k}\cdot\mathbf{x}}$($C$ 为某个非 0 复数)，请计算概率流密度。(2). 假设粒子的波函数为球面波 $\psi(\mathbf{x})=Ce^{ikr}/r$，请计算相应的概率流密度。

3. 假设我们定义能量密度 $w(\mathbf{x},t)=\frac{\hbar^2}{2m}(\nabla\psi)\cdot(\nabla\psi^*)+\psi^*V(\mathbf{x})\psi$, 以及能量流密度 $\mathbf{S}=-\frac{\hbar^2}{2m}\left(\frac{\partial\psi^*}{\partial t}\nabla\psi+\frac{\partial\psi}{\partial t}\nabla\psi^*\right)$(注意这里的 $\mathbf{S}$ 不是表示自旋算符)，这里 $\psi(\mathbf{x},t)$ 为粒子的波函数，$V(\mathbf{x})$ 为势能。请证明，它们满足局域的能量守恒方程

$$\frac{\partial w}{\partial t}+\nabla\cdot\mathbf{S}=0. \tag{3.31}$$

4. 假设粒子处于形如 $\psi(\mathbf{x},t)=e^{-iEt/\hbar}\psi(\mathbf{x})$ 的定态上，满足定态薛定谔方程 $\left[-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})\right]\psi(\mathbf{x})=E\psi(\mathbf{x})$。请证明，在这样的定态上，粒子的概率密度和概率流密度均与时间无关。同样，请证明粒子的能量密度和能量流密度也都与时间无关，并请证明 $\mathbf{S}=E\mathbf{J}$。

3.2 带电粒子与磁场的相互作用

3.2.1 磁场中带电粒子的薛定谔波动方程和概率流密度

我们这里还想讨论另一个也很常见的薛定谔波动方程，即一个带电粒子 (假设电荷为 e) 在磁场中的薛定谔方程。在电动力学中我们知道，磁场总可以用矢量势 $\mathbf{A}(\mathbf{x})$ 来描述。而且从理论力学中我们知道，处于磁场中的一个非相对论带电粒子的哈密顿量可以写成 $\frac{(\mathbf{p}-e\mathbf{A})^2}{2m}+V(\mathbf{x})$, $\mathbf{p}$ 就是带电粒子的动量，或者更严格地说是正则动量。因此很容易猜到，带电粒子的哈密顿算符为

$$H=\frac{(\mathbf{P}-e\mathbf{A})^2}{2m}+V(\mathbf{X}). \tag{3.32}$$

由于动量算符 $\mathbf{P}$ 在坐标表象下的表示为 $\widehat{\mathbf{P}}=-i\hbar\nabla$，因此我们很容易得到相应的坐标表象下的哈密顿算符，进而就可以得到如下薛定谔波动方程，

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t)=\left[-\frac{\hbar^2}{2m}\left(\nabla-\frac{ie}{\hbar}\mathbf{A}(\mathbf{x})\right)^2+V(\mathbf{x})\right]\psi(\mathbf{x},t). \tag{3.33}$$

从这个方程 (3.33) 可以看到，磁场的引入就相当于在薛定谔波动方程中将梯度算符 ∇ 替换成 $\nabla-\frac{ie}{\hbar}\mathbf{A}$，而这个表达式是复的。注意到这一点，我们就可以完全按照前面推导概率流密度的类似办法来推导磁场中带电粒子的概率流密度，最终的结果是

$$\mathbf{J}(\mathbf{x},t)=\frac{-i\hbar}{2m}\left(\psi^*\nabla\psi-\psi\nabla\psi^*\right)-\frac{e}{m}\mathbf{A}\rho. \tag{3.34}$$

我们看到，加了磁场以后，带电粒子的概率流密度多了一个正比于 $\mathbf{A}\rho$ 的项。特别的，假设我们把波函数 $\psi(\mathbf{x},t)$ 写成 $\psi(\mathbf{x},t)=\sqrt{\rho}e^{i\theta(\mathbf{x},t)}$ 的形式 (式中 $\theta(\mathbf{x},t)$ 为波函数的相位，为实数)，则上面的概率流密度就可以重写成

$$\mathbf{J}(\mathbf{x},t)=\left(\hbar\nabla\theta-e\mathbf{A}\right)\rho/m. \tag{3.35}$$

对于一块导体而言，假设我们近似认为其传导电子是相互独立的，总粒子数为 N，那么，如果它的所有传导电子都可以处于相同的量子态 $\psi(\mathbf{x},t)$ 的话，那它的电荷密度将是 $\rho_e=eN\rho=en(\mathbf{x},t)$，$n(\mathbf{x},t)=N\rho$ 为传导电子数密度，而电流密度将是 $\mathbf{J}_e=eN\mathbf{J}$。这时候，如果导体内传导电子的电荷密度不为零，那磁场的存在就会引起一个额外的电流 $-\frac{e^2}{m}\mathbf{A}n$。但是，正如我们将会在后面的章节中进一步讨论的，电子是费米子，由于泡利不相容原理，不同的传导电子不可能处于相同的量子态，所以正常来说，这一段的分析并不能成立。

但是，在传统的超导体中，电子成双成对，通过声子传递相互作用，两个电子配成了所谓的库珀对，妙就妙在，由于库珀对由两个费米子组成，所以它其实是一个玻色子。费米子总是孤独的，但是就好比大量光子能凑在一起形成宏观的电磁波，在温度足够低时，作为玻色子，所有库珀对会凝聚到相同的最低能态上去，因此对于库珀对上一段对电荷密度以及电流密度的讨论就可以成立了。这时候前面的波函数 $\psi(\mathbf{x},t)$ 可以看成是单个库珀对的波函数，当然，由于所有库珀对都处于相同的状态，$\psi(\mathbf{x},t)$ 其实同时是任何库珀对的波函数，而前面的 e 现在就应该理解成库珀对的电荷 (也即是电子电荷的两倍)。

由于超导材料背景正离子的电荷分布几乎是均匀的，所以电荷平衡要求库珀对的电荷密度在超导内部也是一个不为零的常数，即库珀对在超导体内部会是一个均匀分布 (从而 ρ 是一个常数)。一般来说这也意味着，超导体的基态将具有空间平移不变性，从而库珀对波函数的相位 θ 也将是一个常数，因此根据 (3.35) 式，对于这种情形的超导体将有

$$\mathbf{J}(\mathbf{x},t)=-e\mathbf{A}\rho/m. \tag{3.36}$$

这时候根据 (3.36) 式，磁场所满足的电动力学方程 $\nabla^2\mathbf{A}=-\mu_0\mathbf{J}_e$ 就会变成，$\nabla^2\mathbf{A}=\mu_0\frac{e^2n}{m}\mathbf{A}$。这个方程的解是指数衰减的 (指数增长解非物理)，$\mathbf{A}\sim e^{-x/\lambda}$，式中 $(1/\lambda)^2=\mu_0\frac{e^2n}{m}$，这也就是说，磁场线只能穿透到超导材料表面很薄的一层 (穿透深度大概为 λ)，也即是说，无论初始状态为何，超导体总会将磁场线排斥出去，在超导体内部磁场总是 0，超导体有完全抗磁性，这就是著名的迈斯纳效应。

3.2.2 超导体的磁通量子化

上一小节最后的分析也告诉我们，超导电流只分布在超导体表面穿透深度约为 λ 的一个小薄层里面，超导体内部是没有电流的。实际上，我们可以将整个论证过程反过来，根据超导体内部没有电流的事实，进而由 (3.35) 式得出，在超导体内部有

$$\hbar\nabla\theta = e\mathbf{A}, \tag{3.37}$$

这个式子是一个更为一般性的结果，因为它允许我们讨论系统没有空间平移不变性的 θ 不是常数的情形。根据这个结果 (3.37)，我们马上就有，在超导体内部 $\mathbf{B} = \nabla\times\mathbf{A} = \hbar/e\nabla\times\nabla\theta = 0$，即超导体内部磁场恒为 0，这也就是迈斯纳效应。

但是，为什么超导体内部通常没有超导电流呢？为了看清楚这一点，首先由局域概率守恒方程 $\partial_t\rho + \nabla\cdot\mathbf{J} = 0$ 可知，当超导材料达到稳定状态，从而库珀对的密度不随时间变化时，ρ 将与时间无关，从而必有 $\nabla\cdot\mathbf{J} = 0$，也即是说，这时候超导体内部的超导电流必为稳恒电流。进一步取 $\psi(\mathbf{x},t) = \sqrt{\rho}e^{i\theta(\mathbf{x},t)}$（假设库珀对均匀分布，从而 ρ 是一个常数），代入薛定谔波动方程 (3.33)，再利用 (3.35) 式，就有

$$-\hbar\frac{\partial\theta}{\partial t} = \frac{m}{2\rho^2}\mathbf{J}^2 + V(\mathbf{x}). \tag{3.38}$$

假设系统达到稳定状态，从而方程 (3.38) 右边与时间无关，因此可以记 $-\hbar\frac{\partial\theta}{\partial t} = \varepsilon(\mathbf{x})$，从而 $\theta(\mathbf{x},t) = \theta_0(\mathbf{x}) - \varepsilon(\mathbf{x})t/\hbar$，也即是说，库珀对的波函数将具有 $\psi(\mathbf{x},t) = \psi(\mathbf{x})e^{-i\varepsilon(\mathbf{x})t/\hbar}$ 的形式，因此这时候，$\varepsilon(\mathbf{x}) = -\hbar\frac{\partial\theta}{\partial t}$ 可以看成是 $\mathbf{x}$ 点的库珀对的能量。另一方面，当温度非常低时，库珀对将处于最低能态，而从 (3.38) 式可以清楚地看到，要让这个能量 ε 最低就必定要求 $\mathbf{J} = 0$，这就是超导体内部通常没有超导电流的根本原因。

迈斯纳效应告诉我们，超导体内部磁场必定为 0。但是，有可能出现这样一种情况，即在超导材料中间的某个很小的管状区域内，材料可能不处在超导态，而是处在通常的物态，这个通常态的管状小区域被超导态包围着。因为这个小区域是通常态，所以它允许磁场通过，因此我们可以设想有磁通从这个通常态小管子内通过，如图 (3.1) 所示。我们知

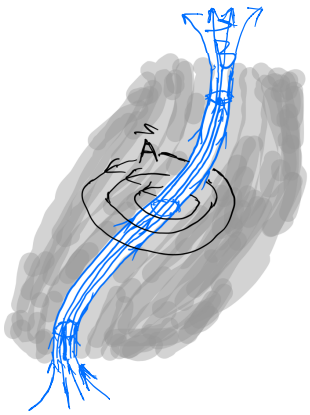

Figure 3.1: 超导体的磁通量子化。

道，在通常态管状区域与周围超导态交界的一个厚度为 λ 的小薄层内会有超导电流。当通

常态小管子内的磁场发生变化时，它就会在这个薄层内产生感生电场，由于超导电流的电阻为 0，所以即使一点点感生电场也可以产生足够大的感生电流，因此，当通常态小管子内的磁场发生变化时，在这个交界薄层内将会有足够大的超导电流阻碍这个磁通的变化。因此，可以设想，被超导态包围的通常态小管子内的磁通将保持恒定不变。实际上，正如我们将要证明的，这个磁通不仅恒定不变，而且它还是量子化的。

由于中间有磁场通过，所以在如图 (3.1) 所示的这种情况下，超导态内部任何一条包围通常态小管子的闭合回路上，矢量势 $\mathbf{A}$ 将不为 0，这是因为，沿着这样的闭合回路 $\oint_L \mathbf{A}\cdot\mathbf{dl} = \int_S \mathbf{B}\cdot\mathbf{dS} = \Phi$，式中 S 表示回路 L 所包围的面积，Φ 就是通常态小区域上的磁通。

另一方面，我们又知道，在超导态内部，电流必定为 0，从而有方程 (3.37)，根据这个方程，$\oint_L \mathbf{A}\cdot\mathbf{dl} = \hbar/e\oint_L \nabla\theta\cdot\mathbf{dl} = \hbar/e\oint_L d\theta$，但是由于 θ 是波函数的相位，因此围绕着闭合回路一周，θ 的改变量只可能取 $2n\pi(n\in\mathbb{Z})$，从而必有

$$\Phi = \oint_L \mathbf{A}\cdot\mathbf{dl} = n2\pi\hbar/e. \tag{3.39}$$

因此，超导体允许的磁通必然是量子化的，其量子化单位为

$$\Phi_0 = 2\pi\hbar/e, \tag{3.40}$$

称为一个磁通量子。

人们通常称超导体内部出现的这种带量子化磁通的通常态小管子为超导体的涡旋线激发，因为从远处看，它就是一根内部有磁力线通过的细线，而上一段中的整数 n 就称为相应涡旋线的拓扑量子数。这样的涡旋线激发实际上可以在超导体里面运动变形，不仅如此，两个不同拓扑量子数的涡旋线还可以合并成一个涡旋线，而一个高拓扑量子数的涡旋线也可能分裂成多个低拓扑量子数的涡旋线，但是，在整个这样的合并或分裂过程前后，总拓扑量子数是守恒的。

拓扑量子数的守恒可以这样来论证：假设在超导体内取一个很大的闭合回路 L，则很显然，$\frac{1}{2\pi}\oint_L d\theta = w\in\mathbb{Z}$, 而这个结果显然不依赖于 $\theta(\mathbf{x},t)$ 在回路 L 所包围的区域内的具体细节，只要 $\theta(\mathbf{x},t)$ 在区域的边界 L 上渐近不变，那无论 $\theta(\mathbf{x},t)$ 在区域内部如何变化，$\frac{1}{2\pi}\oint_L d\theta = w$ 都将保持不变。另一方面，由于 θ 是波函数的相位，$\psi\sim e^{i\theta}$，因此对于任何闭合回路 C，$\frac{1}{2\pi}\oint_C d\theta$ 其实就是当我们围绕 C 逆时针走一周时，波函数 ψ(是一个复数) 在复平面内围绕着原点转动的圈数。从而涡旋线的拓扑量子数其实就是当我们围绕着涡旋线走一周时，波函数在复平面内转动的圈数。而 $\frac{1}{2\pi}\oint_L d\theta = w$ 就是当我们围绕着大闭合回路 L 走一圈时波函数在复平面内转动的总圈数，很显然，它必然等于围绕 L 包围的区域内每一个涡旋线激发走一圈时波函数在复平面内转动的圈数的总和，也就说，w 必然等于 L 所包围的区域内所有涡旋线的拓扑量子数之和，如图 (3.2) 所示。从而 w 的恒定不变性就意味着拓扑量子数守恒。

3.2.3 规范不变性

在电动力学中我们知道麦克斯韦方程组有所谓的规范不变性，即在如下规范变换

$$\phi \to \phi - \partial_t\varepsilon, \quad \mathbf{A}\to\mathbf{A}+\nabla\varepsilon, \tag{3.41}$$

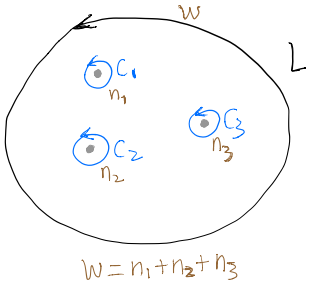

Figure 3.2: 超导体的横截面图，回路 L 之内包围了三个涡旋线。

的作用下，电场强度 $\mathbf{E}$ 和磁场强度 $\mathbf{B}$ 都保持不变。式中 $\varepsilon(\mathbf{x},t)$ 为一个任意的实函数，ϕ 为电磁场标量势，$\mathbf{A}$ 为矢量势，我们这里主要关心矢量势以及与之相关的磁场，因此下面我们将仅仅考察 $\varepsilon(\mathbf{x})$ 与 t 无关的规范变换。

下面我们想要考察的是，带电粒子的薛定谔波动方程 (3.33) 如何才能保持规范不变性。答案其实非常简单，只要在矢量势作规范变换 $\mathbf{A}\to\mathbf{A}+\nabla\varepsilon$ 的同时，让带电粒子的波函数 $\psi(\mathbf{x},t)$ 同时作如下变换

$$\psi(\mathbf{x},t)\to e^{ie\varepsilon(\mathbf{x})/\hbar}\psi(\mathbf{x},t)\quad with\quad \mathbf{A}\to\mathbf{A}+\nabla\varepsilon. \tag{3.42}$$

因为很容易验证，在这样的变换下将有，

$$\left(\nabla-\frac{ie}{\hbar}\mathbf{A}(\mathbf{x})\right)\psi(\mathbf{x},t)\to e^{ie\varepsilon(\mathbf{x})/\hbar}\left(\nabla-\frac{ie}{\hbar}\mathbf{A}(\mathbf{x})\right)\psi(\mathbf{x},t). \tag{3.43}$$

进而就容易知道，带电粒子的薛定谔波动方程 (3.33) 将在规范变换 (3.42) 下保持形式不变。

正因为算符 $\nabla-\frac{ie}{\hbar}\mathbf{A}$ 在规范变换 (3.42) 下有形如 (3.43) 这样美好的数学性质，即 $\left(\nabla-\frac{ie}{\hbar}\mathbf{A}\right)\psi(\mathbf{x},t)$ 和 $\psi(\mathbf{x},t)$ 的变换规律是一样的。所以，数学家常常称 $\nabla-\frac{ie}{\hbar}\mathbf{A}$ 为协变导数，用分量形式可以记作 D_i, 即

$$D_i=\partial_i-\frac{ie}{\hbar}A_i(\mathbf{x}). \tag{3.44}$$

与协变导数密切相关，数学家也常常称矢量势 A_i 为 $U(1)$ 向量丛上的联络。读者不要被 $U(1)$ 向量丛这样的数学术语吓住，其实它就是在空间每一点 $\mathbf{x}$ 上指定一个复数 (向量)$\psi(\mathbf{x})$, 并且这些复数可以按照规则 (3.42) 进行变换。所谓的 $U(1)$ 其实指的就是 $\psi(\mathbf{x})$ 的变换规则是乘上相因子 $e^{ie\varepsilon(\mathbf{x})/\hbar}$。

人们常常类比于动量算符 $\widehat{P}_i=-i\hbar\partial_i$(也称作正则动量算符), 但是将普通的偏导替换成协变导数，进而引入力学动量算符 $\vec{\Pi}$, 其分量形式可以定义成

$$\Pi_i=-i\hbar D_i=\widehat{P}_i-eA_i(\mathbf{x}). \tag{3.45}$$

与通常的正则动量算符不同，力学动量算符两两不对易，读者可以计算出它们的对易关系是

$$[-i\hbar D_i,-i\hbar D_j]=(i\hbar e)F_{ij}, \tag{3.46}$$

式中 $F_{ij}=\partial_i A_j-\partial_j A_i$，人们很容易看出，它实际上就是对矢量场 $\mathbf{A}$ 求旋度的分量式写法，比如 $\partial_x A_y-\partial_y A_x$ 其实就是 $(\nabla\times\mathbf{A})_z$。因此，$F_{ij}$ 和磁场强度 $\mathbf{B}$ 的分量形式是一一对应的，人们可以通过使用所谓的列维-席维塔符号 ε_{ijk} 并结合求和约定，从而将这两者之间的对应关系写成 $F_{ij}=\varepsilon_{ijk}B_k$。这里 ε_{ijk} 关于三个指标是全反对称的，即任意两个指标交换顺序都会出一个负号 (因此当三个指标中有两个取值一样时就会得到 0)，比如说 $\varepsilon_{jik}=-\varepsilon_{ijk}$，并且 $\varepsilon_{123}=1$。而所谓的求和约定是指，当一个表达式中出现两个重复指标时，就默认对这个指标进行求和，比方说刚才的表达式 $\varepsilon_{ijk}B_k$ 我们就默认对指标 k 的三个不同取值进行求和。

在数学上，求和约定并不是本质性的东西，但是它有时候可以省略掉一大堆求和符号，从而使得式子更加简洁明了，据说爱因斯坦在研究广义相对论的时候最早引入了这样的求和约定，所以有时候也称之为爱因斯坦求和约定。利用 $F_{ij}=\varepsilon_{ijk}B_k$，我们又可以将 (3.46) 式重写成

$$[\Pi_i,\Pi_j]=(i\hbar e)\varepsilon_{ijk}B_k. \tag{3.47}$$

之所以称 $\vec{\Pi}=\widehat{\mathbf{P}}-e\mathbf{A}$ 为力学动量的原因是，利用它我们可以将磁场中带电粒子的哈密顿量写成标准的动能加势能的形式

$$\widehat{H}=\frac{\vec{\Pi}^2}{2m}+V(\mathbf{x}). \tag{3.48}$$

列维-席维塔符号 ε_{ijk} 结合求和约定的一个很常见的应用是用来表示三维矢量的叉乘，比方说假设 $\mathbf{c}=\mathbf{a}\times\mathbf{b}$，那么我们很容易验证 $c_k=\varepsilon_{ijk}a_ib_j$，注意这里对指标 i,j 都是要求和的。因此，我们也可以将 F_{ij} 和磁场强度 $\mathbf{B}$ 的分量形式之间的对应关系反过来写成，$B_k=\frac{1}{2}\varepsilon_{ijk}F_{ij}$。

如果将波函数 $\psi(\mathbf{x})$ 推广成一个多分量的列矢量波函数，同时将变换的相因子 $e^{ie\varepsilon(\mathbf{x})/\hbar}$ 推广成一个多行多列的幺正矩阵，矢量势 $A_i(\mathbf{x})$ 也作相应的推广 (也要变成矩阵)，那我们就把电磁场推广到了所谓的杨-米尔斯场 (或者说非阿贝尔规范场)。你可能知道，杨-米尔斯场可以描写夸克和胶子的相互作用，也是粒子物理标准模型的基础。但杨-米尔斯场其实是电磁场的推广，而杨-米尔斯场的杨就是指的杨振宁，因为他和米尔斯两个人最早做出了这一理论推广。不过，在非阿贝尔规范场的推广中，公式 (3.46) 依然成立，实际上这时候它可以看成是 F_{ij} 的定义式。但是，由于 $A_i(\mathbf{x})$ 变成了矩阵，所以 $[-i\hbar D_i,-i\hbar D_j]=(i\hbar e)\{\partial_iA_j-\partial_jA_i+(e/i\hbar)[A_i,A_j]\}$，也即是说，这时候 $F_{ij}=\partial_iA_j-\partial_jA_i+(e/i\hbar)[A_i,A_j]$，在 F_{ij} 的表达式中额外多加一个矩阵对易子项 $(e/i\hbar)[A_i,A_j]$ 从而使得公式 (3.46) 普遍成立，就是杨振宁和米尔斯在规范场数学上的关键性突破。

规范变换告诉我们，当我们用波函数 $\psi(\mathbf{x})$ 来描写带电粒子，同时用矢量势 $\mathbf{A}$ 来描写电磁场时，我们的描述并不唯一，两个可以通过规范变换 (3.42) 联系起来的数学描述在物理上其实相互等价。也就是说，真正物理上可以测量的量必须是规范不变的，$\psi(\mathbf{x})$ 和 $\mathbf{A}$ 本身不是物理可观测量，而只是我们描述物理规律的时候用到的数学变量，这样的数学描述可以相差一个规范变换，因此在物理上它们是一种冗余的描述。人们之所以采用这种冗余的描述，是因为它可以使得方程很简洁。

总之，真正物理上的可观测量必须规范不变。什么样的量在规范变换 (3.42) 下会保持不变呢？首先我们很容易发现，带电粒子的概率密度 $\rho(\mathbf{x},t)=|\psi(\mathbf{x},t)|^2$ 是规范不变的，因此可观测。其次，我们注意到，在规范变换 (3.42) 下，从空间某点 a 到空间某点 b 的线积分 $\exp\left(i\frac{e}{\hbar}\int_a^b \mathbf{A}\cdot\mathbf{dl}\right)$(如图 (3.3) 所示) 将变换为 $\exp(ie\varepsilon(b)/\hbar)\exp\left(i\frac{e}{\hbar}\int_a^b \mathbf{A}\cdot\mathbf{dl}\right)\exp(-ie\varepsilon(a)/\hbar)$。因此，下面的表达式是规范不变的

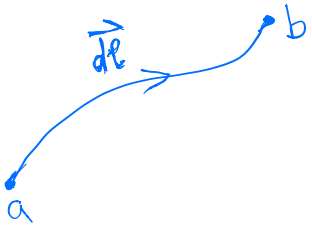

Figure 3.3: 矢势场 $\mathbf{A}$ 中的一条路径。

$$\psi^*(b)\exp\left(i\frac{e}{\hbar}\int_a^b \mathbf{A}\cdot\mathbf{dl}\right)\psi(a). \tag{3.49}$$

在物理上，这个表达式正比于一个概率幅，即 ψ 态的带电粒子，我们测到它从 a 点出发沿着如图 (3.3) 所示的积分路线运动到 b 点的概率幅，式中的 $\psi^*(b)\psi(a)$ 就正比于矢量势 $\mathbf{A}=0$ 时的相应概率幅。

从上面的表达式 (3.49) 可以看到，当带电粒子在矢量势 $\mathbf{A}$ 中沿着某条路径 L 运动时，它的量子力学概率幅要额外多乘一个因子

$$\exp\left(i\frac{e}{\hbar}\int_L \mathbf{A}\cdot\mathbf{dl}\right), \tag{3.50}$$

这是规范不变性所要求的。

另外，如果我们将波函数写成 $\psi(\mathbf{x},t)=\sqrt{\rho(\mathbf{x},t)}e^{i\theta(\mathbf{x},t)}$ 的形式，那 (3.49) 式的规范不变性也等价于下面的相位差 $\Delta\theta_A$ 的规范不变性，

$$\Delta\theta_A=\theta(b)-\theta(a)-\frac{e}{\hbar}\int_a^b \mathbf{A}\cdot\mathbf{dl}. \tag{3.51}$$

3.2.4 Aharonov-Bohm 效应

我们知道，矢量势 $\mathbf{A}$ 本身不物理，磁场强度 $\mathbf{B}$ 才是物理的。因此人们可以会想，在磁场区之外运动的带电粒子应该不受磁场影响。在经典物理的层次上，这个结论无疑是对的，因为磁场区之外的带电粒子受到的洛伦兹力当然是 0。但是，在量子力学的层次上，我们从薛定谔波动方程 (3.33) 可以看到，与带电粒子直接耦合的是矢量势 $\mathbf{A}$，而不是磁场强度 $\mathbf{B}$，因此这个结论是否依然成立就值得进一步研究了。

实际上，Aharonov-Bohm 指出，在量子力学的层次上，上面所说的结论并不对，磁场可以对带电粒子产生非局域的影响。为了看清楚这一点，Aharonov-Bohm 重新考察了著名的电子双缝干涉实验，只不过 Aharonov-Bohm 在双缝的背后添加了一个细小的螺线管，它内部携带一个垂直于纸面的磁通 Φ，如图 (3.4) 所示。电子从源 S 发出，经过双缝，

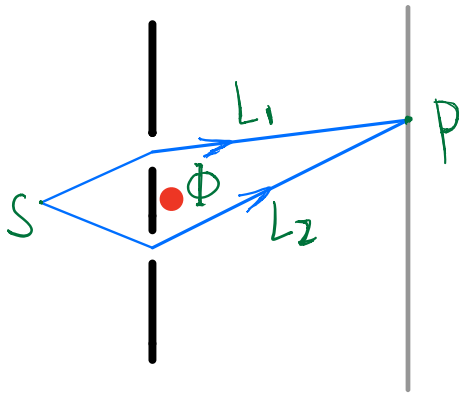

Figure 3.4: Aharonov-Bohm 双缝干涉实验。

然后在屏幕上干涉。由于螺线管基本上在电子的运动路径之外，而螺线管外面并没有磁场，所以如果按照经典理论告诉我们的直觉，螺线管的存在将对屏幕上的电子干涉条纹没有影响，特别的，即使我们改变螺线管内的磁通 Φ，干涉条纹也不会移动。

但是实验发现，情况并非如此，螺线管的存在会影响到干涉条纹，改变螺线管内的磁通 Φ，电子的干涉条纹就会发生移动。看起来，螺线管内磁场的存在可以对电子产生非局域的影响。这是为什么呢?

正确的解释是，在量子力学的层次上，与电子直接耦合的是矢量势 $\mathbf{A}$，而不是磁场强度 $\mathbf{B}$。虽然在螺线管之外，$\mathbf{B}=0$，但是由于管内磁场的存在，螺线管之外的 $\mathbf{A}\neq 0$。虽然 $\mathbf{A}$ 本身不是物理的，但是由 $\mathbf{A}$ 可以定义一个规范不变的物理量 $\exp\left(i\frac{e}{\hbar}\oint_L \mathbf{A}\cdot\mathbf{dl}\right)$，正是这个量会影响到干涉条纹。

具体来说：在没有加螺线管时，假设我们记电子从 S 出发经过上面的第 1 条缝到达屏幕上 P 点的概率幅为 ϕ_1，记经过第 2 条缝到达 P 点的概率幅为 ϕ_2，那么由于双缝同时打开，电子到达 P 点的概率将由下式给出

$$|\phi_1+\phi_2|^2=|\phi_1|^2+|\phi_2|^2+(\phi_1^*\phi_2+\phi_1\phi_2^*). \tag{3.52}$$

上面这个式子最后的交叉项就是所谓的干涉项，它解释了屏幕上的干涉条纹。实际上，假设将 ϕ_1、ϕ_2 重写成 $\phi_1=\sqrt{\rho_1}e^{i\theta_1}$、$\phi_2=\sqrt{\rho_2}e^{i\theta_2}$，那么人们很容易看到，这个交叉项 $(\phi_1^*\phi_2+\phi_1\phi_2^*)=2\mathrm{Re}(\phi_1^*\phi_2)=2\sqrt{\rho_1\rho_2}\cos(\theta_2-\theta_1)$，即取决于两个概率幅的相位差。当然，这些都属于我们已经在第一章中讨论过了的通常的电子双缝干涉内容。

加上螺线管以后，根据我们在上一小节的讨论，我们要给电子的概率幅额外乘上因子 $\exp\left(i\frac{e}{\hbar}\int_L \mathbf{A}\cdot\mathbf{dl}\right)$。因此，电子经过缝 1 到达 P 点的概率幅将是 $\exp\left(i\frac{e}{\hbar}\int_{L_1} \mathbf{A}\cdot\mathbf{dl}\right)\cdot\phi_1$，经过缝 2 到达 P 点的概率幅将是 $\exp\left(i\frac{e}{\hbar}\int_{L_2} \mathbf{A}\cdot\mathbf{dl}\right)\cdot\phi_2$ 。从而，由于螺线管的存在，电子到达 P 点的概率现在将是

$$\begin{aligned}&\left|\exp\left(i\frac{e}{\hbar}\int_{L_1}\mathbf{A}\cdot\mathbf{dl}\right)\phi_1+\exp\left(i\frac{e}{\hbar}\int_{L_2}\mathbf{A}\cdot\mathbf{dl}\right)\phi_2\right|^2\\ =\;&|\phi_1|^2+|\phi_2|^2+2\mathrm{Re}\{\phi_1^*\phi_2\exp\left[i\frac{e}{\hbar}\left(\int_{L_2}\mathbf{A}\cdot\mathbf{dl}-\int_{L_1}\mathbf{A}\cdot\mathbf{dl}\right)\right]\}\end{aligned} \tag{3.53}$$

注意到积分路径 L_1 的方向是 $S\to P$, 假设我们改变路径的方向为 $P\to S$, 并记变向后的路径为 $\overline{L_1}$, 则很显然 $\exp\left(i\frac{e}{\hbar}\int_{L_1}\mathbf{A}\cdot\mathbf{dl}\right)=\exp\left(-i\frac{e}{\hbar}\int_{\overline{L_1}}\mathbf{A}\cdot\mathbf{dl}\right)$。从而, $\exp\left[i\frac{e}{\hbar}\left(\int_{L_2}\mathbf{A}\cdot\mathbf{dl}-\int_{L_1}\mathbf{A}\cdot\mathbf{dl}\right)\right]=\exp\left(i\frac{e}{\hbar}\int_{L_2\cup\overline{L_1}}\mathbf{A}\cdot\mathbf{dl}\right)$。很明显, $L_2\cup\overline{L_1}$ 是围绕螺线管的一条闭合路径, 因此, $\exp\left(i\frac{e}{\hbar}\int_{L_2\cup\overline{L_1}}\mathbf{A}\cdot\mathbf{dl}\right)=\exp\left(i\frac{e}{\hbar}\Phi\right)=\exp\left(i2\pi\Phi/\Phi_0\right)$。因此，我们最后得到，加了螺线管以后电子到达 P 点的概率将为

$$|\phi_1|^2+|\phi_2|^2+2\mathrm{Re}\{\exp\left(i2\pi\Phi/\Phi_0\right)\phi_1^*\phi_2\}. \tag{3.54}$$

从这个结果我们清楚地看到，螺线管的引入会在双缝干涉项中引入一个额外的相位差 $2\pi\Phi/\Phi_0$, 因此只要螺线管的磁通不是磁通量子 Φ_0 的整数倍，那它就会影响干涉条纹，尤其是，当我们改变螺线管内的磁通时，电子的干涉条纹会发生移动。

3.2.5 磁单极子

麦克斯韦的电磁场方程关于电场和磁场其实非常对称，然而电场可以从单个带电粒子发出，电场有源，而磁场却只能由电流产生，无源。人们很早就想到，既然电和磁这么对称，那类比于电子，可不可能存在可以直接发出磁场的磁单极子呢？虽然实验上到今天为止都还没有观测到磁单极子，但是这不妨碍人们进行理论上的探讨。

所谓的一个磁单极子，就是一个能产生下面这种磁场的粒子，

$$\mathbf{B}=\frac{g}{4\pi r^2}\mathbf{e}_r, \tag{3.55}$$

式中，$\mathbf{e}_r$ 为径向方向的单位矢量。很显然，在包围这个磁单极子的闭合曲面 S 上磁通量为

$$\oint_S\mathbf{B}\cdot d\mathbf{S}=g. \tag{3.56}$$

所以 g 又称为磁荷。

很显然，如果存在磁单极子，那磁场 $\mathbf{B}$ 与矢量势 $\mathbf{A}$ 之间的关系式 $\mathbf{B}=\nabla\times\mathbf{A}$ 就不能在空间整体上成立了，因为否则的话，闭合曲面 S 上的磁通将为 $\oint_S\mathbf{B}\cdot d\mathbf{S}=\oint_S(\nabla\times\mathbf{A})\cdot d\mathbf{S}$, 利用斯托克斯公式并注意到闭合曲面没有边界线，从而必有 $\oint_S(\nabla\times\mathbf{A})\cdot d\mathbf{S}=0$，这当然和上一段的结果矛盾。不过，在每一个空间局部上 (而不是象闭合曲面这样的整体上)，我们都依然有 $\mathbf{B}=\nabla\times\mathbf{A}$。也即是说，如果存在磁单极子，那矢量势 $\mathbf{A}$ 就不能整体上定义，而只能在每一个空间局部上分别定义。这其实并不会有任何问题，因为我们知道 $\mathbf{A}$ 本身并不是物理的，也不是唯一的。也因为如此，不同空间局部上分别定义的矢量势之间其实可以相差一个规范变换。

总之，在经典物理的层次上，磁单极子的存在完全可能。但是考虑到量子力学，情况会变得有些微妙的不同。正如我们即将看到的，考虑到电荷与磁单极子相互作用的量子力学，磁单极子仅在一个电荷量子化的世界里才有可能，换言之，如果存在磁单极子，那么电荷必须是量子化的。所谓电荷量子化，也就是所有带电体的电荷总量都是某个最小电荷单元的整数倍的现象。显然，我们生活的这个世界是电荷量子化的，所以原则上可以有磁单极子存在，实际上，狄拉克最早就是用磁单极来解释电荷量子化现象的。不过，实际上电荷量子化可以有其它的解释 (关于这一点我们将在关于对称性与守恒定律的相关章节中

再来讨论), 因而更合适的看法应该是，在量子力学的层次上，只有电荷量子化才能允许磁单极子存在。下面我们来解释其原因。

假设有一个磁荷为 g 的磁单极子位于坐标原点，一个电荷为 e 的带电粒子在这个磁单极子的磁场中沿着闭合路径 C 运动一圈，如图 (3.5) 所示。则根据我们在第 (3.2.3) 小节

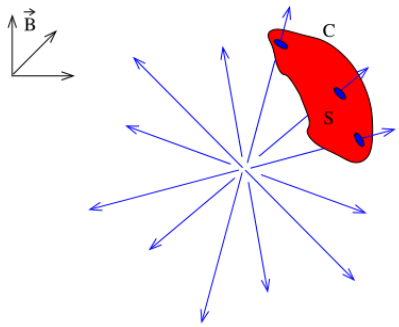

Figure 3.5: C 包围面积 S。图片来源于 D. Tong 的讲义。

中的讨论，磁场的存在将在带电粒子的概率幅中额外引入一个相因子

$$\exp\left(i\frac{e}{\hbar}\oint_C \mathbf{A}\cdot\mathbf{dl}\right). \tag{3.57}$$

然而关于这个相因子我们有两种不同的计算方法。第一种方法是，我们取闭合路径包围的曲面为 S, 如图 (3.5) 所示，则应用斯托克斯公式我们有 $\oint_C \mathbf{A}\cdot\mathbf{dl} = \int_S (\nabla\times\mathbf{A})\cdot\mathbf{dS} = \int_S \mathbf{B}\cdot\mathbf{dS}$, 假设曲面 S 对应的立体角是 Ω, 则由于 4π 立体角上的总磁通为 g，所以 $\int_S \mathbf{B}\cdot\mathbf{dS} = \frac{\Omega}{4\pi}g$, 因此按照这第一种计算方法，单极子的磁场在带电粒子概率幅上额外引入的相因子为

$$\exp\left(i\frac{e}{\hbar}\oint_C \mathbf{A}\cdot\mathbf{dl}\right) = \exp\left(i\frac{eg}{\hbar}\frac{\Omega}{4\pi}\right). \tag{3.58}$$

但是，我们还有另一种不同的计算相因子 $\exp\left(i\frac{e}{\hbar}\oint_C \mathbf{A}\cdot\mathbf{dl}\right)$ 的方法，那就是取闭合路径 C 包围的曲面为 S'，如图 (3.6) 所示。这时候由于 S' 对应的立体角为 $4\pi-\Omega$, 而且，假设

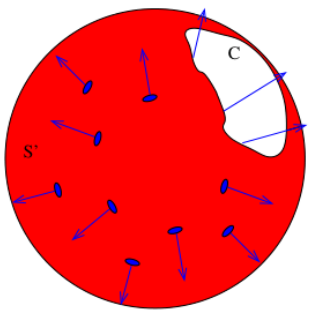

Figure 3.6: C 包围面积 S'。图片来源于 D. Tong 的讲义。

回路 C 的绕行方向与之前的曲面 S 的外法线方向成右手关系的话，那 C 的绕行方向就会与曲面 S' 的外法线成左手关系，从而在应用斯托克斯公式的时候得额外加一个负号，所以最终我们将得出

$$\exp\left(i\frac{e}{\hbar}\oint_C \mathbf{A}\cdot\mathbf{dl}\right)=\exp\left(i\frac{eg}{\hbar}(-)\frac{4\pi-\Omega}{4\pi}\right). \tag{3.59}$$

很显然，要让结果 (3.58) 和结果 (3.59) 不相互矛盾，除非

$$eg=n2\pi\hbar,\quad n\in\mathbb{Z}. \tag{3.60}$$

这个结果就是著名的狄拉克量子化条件，它是理论中任意带磁荷 g 的粒子和任意带电荷 e 的粒子之间必须满足的关系式。

给定一个磁荷为 g 的粒子，那么方程 (3.60) 将告诉我们，任何带电粒子的电荷都必须取 $n2\pi\hbar/g$ 的形式，也就是电荷必须量子化，否则就不可能有任何磁单极子存在。反过来，给定一个最小的电荷单元 e，方程 (3.60) 将告诉我们，所有磁单极子的磁荷也必须量子化为 $n2\pi\hbar/e=n\Phi_0$。

理论上，不只可能有磁单极子，甚至还可能有所谓的双荷子 (dyon)，即既携带磁荷同时又携带电荷的粒子，假设我们考虑一个荷为 (e_1,g_1) 的双荷子在另一个荷为 (e_2,g_2) 的双荷子的电磁场中运动，那么上面的狄拉克量子化条件将要推广成

$$e_1g_2-e_2g_1=n2\pi\hbar. \tag{3.61}$$

3.2.6 习题

1. 对于一个均匀磁场 $\mathbf{B}$, 我们可以将相应的矢量势取成 $\mathbf{A}=\frac{1}{2}\mathbf{B}\times\mathbf{x}$。请证明，在这样的一个均匀磁场中，电荷为 e 的带电粒子的哈密顿算符 $\widehat{H}$ 可以写成,

$$\widehat{H}=-\frac{\hbar^2}{2m}\nabla^2-\frac{e}{2m}\mathbf{B}\cdot\mathbf{L}+\frac{e^2}{8m}\left(\mathbf{B}\times\mathbf{x}\right)^2. \tag{3.62}$$

式中 $\mathbf{L}=\mathbf{x}\times\widehat{\mathbf{P}}=-i\hbar\mathbf{x}\times\nabla$ 是粒子的角动量算符，因此这个带电粒子的轨道运动磁矩为 $\vec{\mu}_L=\frac{e}{2m}\mathbf{L}$，因此上面哈密顿算符表达式的第二项正是标准的磁矩与磁场的相互作用势能项。

3.3 量子隧穿效应

量子隧穿效应是量子世界有别于经典世界的一个典型效应。它不光是很多量子现象的重要物理机制，同时也在量子技术上非常重要。这一节我们将主要以一维运动为例讨论量子隧穿效应。首先，我们将用波动力学的半经典近似证明存在非 0 的量子隧穿概率，并给出其近似公式。其次，我们将用波动力学讨论一种重要的量子隧穿现象，那就是超导物理中的约瑟夫森结。

3.3.1 量子隧穿效应的半经典推导

前面我们通过取坐标表象，推导了一维薛定谔波动方程。这是一个关于波函数如何随着时间演化的方程。但是，有很多时候，我们更关心的是一个有确定能量的量子态 (称之为定态)，这时候，波函数就可以写成 $\psi(x,t)=e^{-iEt/\hbar}\psi(x)$, 代入时间演化的薛定谔波动方程，我们就可以得到波函数的时间无关部分 $\psi(x)$ 所满足的方程

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)\right]\psi(x)=E\psi(x). \tag{3.63}$$

这就是我们在第一章中介绍过的定态薛定谔方程，使得这个方程有非 0 解的 E 就是定态的能量，相应的 $\psi(x)$ 就是定态波函数。

给定一个如图 (3.7) 所示的势能曲线，假设粒子最初处在中间的势阱之内，假设粒子的能量 E 不太高。那么，如果是一个经典的粒子，我们知道它必定会被约束在势阱的 $[a,b]$ 区间之内，如图 (3.7) 所示。但是，如果这个粒子是一个量子的粒子，那情况就可能有很

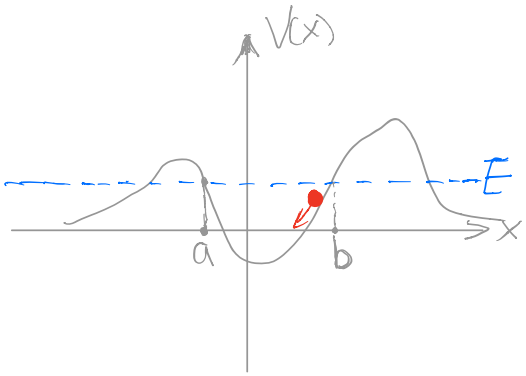

Figure 3.7: 经典粒子在势阱中运动。

大的不同，因为根据波动力学，微观粒子同时是波，如图 (3.8) 所示。在势阱之内，微观粒子的动量 p 大于 0，相应的波矢量 $k=p/\hbar$ 也大于 0，因此这个波当然是振荡的。但是作为波，虽然它会被两边的势垒反射，但在势垒之外依然有可能有透射波，如图 (3.8) 所示。而这就意味着，微观粒子有一定的概率隧穿到势阱外面去。真的会有这种隧穿效应吗？一般来说，这需要通过求解一维定态薛定谔方程 (3.63) 才能知道。

WKB 近似

对于求解一维定态薛定谔方程 (3.63), 物理学家 Wentzel, Kramers 以及 Brillouin 发展了一种重要的半经典近似方法，常常称之为 WKB 近似。这一近似仅在势能 $V(x)$ 变化得很缓慢时才成立。

首先，我们做一个变换，令

$$\psi(x)=e^{iS(x)/\hbar}, \tag{3.64}$$

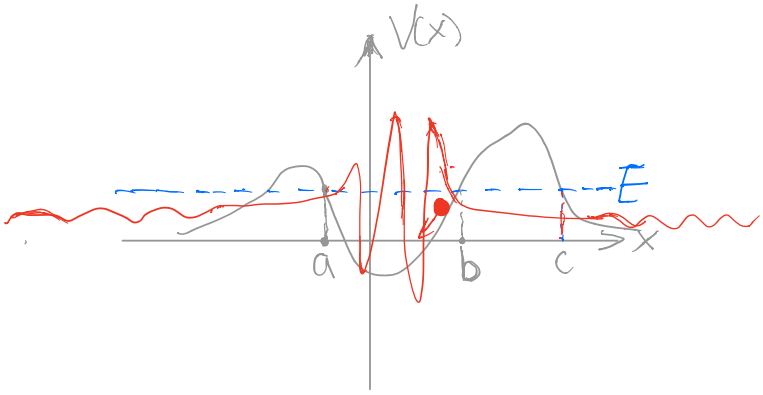

Figure 3.8: 量子粒子有一定的概率隧穿到势阱外面去。

由于波函数通常不只是一个相位因子，因此一般来说，$S(x)$ 当然是一个复函数。将这个变换代入定态薛定谔方程 (3.63)，我们就能得到 $\frac{1}{2m}\left(\frac{dS}{dx}\right)^2 - i\hbar\frac{1}{2m}\frac{d^2S}{dx^2} + V(x) = E$，整理一下，引入经典动量函数 $p(x) = \sqrt{2m(E-V(x))}$，则有

$$\left(\frac{dS}{dx}\right)^2 = p^2(x) + i\hbar\frac{d^2S}{dx^2}. \tag{3.65}$$

下面，我们假定

$$\hbar\left|\frac{d^2S}{dx^2}\right| \ll \left|\frac{dS}{dx}\right|^2. \tag{3.66}$$

因此，我们可以将方程 (3.65) 中的 $\hbar\frac{d^2S}{dx^2}$ 项当成一个微扰小量来处理，形式上即是将 $\hbar$ 当成一个小量[1]，然后将我们要求解的 $S(x)$ 按照 $\hbar$ 的级数进行展开。

在 $\hbar$ 的 0 阶近似上，我们可以忽略方程 (3.65) 中的 $\hbar\frac{d^2S}{dx^2}$，从而得到

$$\frac{dS_0}{dx} = \pm p(x). \tag{3.67}$$

进一步，将这个 0 阶近似结果代入方程 (3.65) 的右边，我们又可以得到直到 $\hbar$ 的一阶展开为止的结果 $S_1(x)$，为

$$\frac{dS_1}{dx} = \pm p(x) + i\hbar\frac{1}{2}\frac{d\log p}{dx}. \tag{3.68}$$

或者也可以将这个结果等价地写成 $S_1(x) = \pm\int^x p(x)dx + i\hbar\frac{1}{2}\log p(x) + C_1$，代入 $\psi(x) = e^{iS/\hbar}$ 即有

$$\psi(x) = C\frac{1}{\sqrt{p(x)}}e^{\pm i\int^x p(x)dx/\hbar}. \tag{3.69}$$

现在，假设我们把空间坐标 x 解析延拓成 x 复平面，从而波函数 $\psi(x)$ 变成这个复平面上的解析函数。$p(x) = 0$ 的经典转向点就变成解析函数 $p(x) = \sqrt{2m(E-V(x))}$ 的支

[1]严格来说，这一说法有点问题，因为 $\hbar$ 是一个有量纲的量，有量纲量的大小并没有绝对意义，因为它和单位的选取有关。因此当我们这样处理的时候其实是说，$\hbar$ 相对于问题中涉及的同量纲物理量而言非常小。

点，而如图 (3.8) 所示的两个经典转向点之间的区间 $[a,b]$ 就变成解析函数 $p(x)$ 在复平面上的一条割线。我们可以考虑围道积分 $\frac{1}{2\pi}\oint_C p(x)dx$, 其中围道 C 紧紧贴着割线 [a, b] 的上下沿，如图 (3.9) 所示。在经典物理的层次上，我们当然可以直接计算这个围道积分，

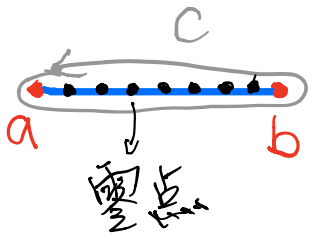

Figure 3.9: Bohr-Sommerfeld 量子化条件。

在割线的下沿我们取 $p(x)=\sqrt{2m(E-V(x))}$，注意到多值函数每穿过割线一次都要改变多值支，从而在割线的上沿就应该取 $p(x)=-\sqrt{2m(E-V(x))}$，因此最终的积分结果即是 $\frac{1}{2\pi}\oint_C p(x)dx=\frac{1}{\pi}\int_a^b\sqrt{2m(E-V(x))}dx$。但是现在，我们可以进一步考察在考虑到量子修正以后的半经典近似上，这个积分是多少。为此我们需要利用 WKB 近似的结果 (3.68)(对于式中的 $\pm p(x)$，和上面一样，在割线下沿取正号，上沿取负号), 并取 $\psi(x)=e^{iS_1/\hbar}$, 从而就可以得到，

$$\frac{1}{2\pi}\oint_C p(x)dx=\hbar\frac{1}{2\pi i}\oint_C\frac{d\psi}{\psi}-i\hbar\frac{1}{2}\frac{1}{2\pi}\oint_C d\log p. \tag{3.70}$$

由于 $\psi(x)$ 是 x 复平面上的解析函数，它在 x 复平面上会有一些零点，尤其是在 $[a,b]$ 这条实线段上会有一些如图 (3.9) 所示零点 (注意 $\psi(x)$ 在这条实线段上不会有极点，因为这时候它是薛定谔方程的波函数，而波函数必定是有限的)，复变函数的知识告诉我们，围道积分 $\frac{1}{2\pi i}\oint_C\frac{d\psi}{\psi}$ 的结果就是解析函数 $\psi(x)$ 在围道 C 之内的零点数目，从而也就是 $\psi(x)$ 在 $[a,b]$ 这条实线段上的零点数目，或者说是波函数的节点数，不妨假设为 n(值得注意的是，考虑到量子修正以后，$[a,b]$ 这条实线段就不是波函数 $\psi(x)$ 的割线了，而是量子化为一些分立的零点了)。另一方面，由于 a 和 b 都是 $p(x)$ 的支点，$[a,b]$ 这条实线段是 $p(x)$ 的割线，因此围绕着 C 转一圈，$p(x)$ 的相位必定改变 2π，也即是说，必有 $\oint_C d\log p=2\pi i$。将这些结果代入 (3.70) 式，就有

$$\frac{1}{2\pi}\oint_C p(x)dx=\left(n+\frac{1}{2}\right)\hbar. \tag{3.71}$$

这就是在量子物理发展史上曾经很著名的 Bohr-Sommerfeld 量子化条件。熟悉分析力学的读者会知道 (3.71) 式左边的式子很容易一般化为分析力学里的作用变量 (它和角变量是一对正则变量，可以看成是轨道角动量的推广。参见朗道《力学》第七章)，因此 Bohr-Sommerfeld 量子化条件是对玻尔氢原子模型第三条假设 (轨道角动量量子化假设) 的推广。

对于 Bohr-Sommerfeld 量子化条件还有一种物理解释，即将 $\oint_C p(x)dx$ 看成是粒子在经典相空间 (即 $x-p$ 所构成的两维空间) 的封闭轨迹线[2]之内所包围的面积。把这个面积

[2]注意，在经典物理层次上，粒子在如图 (3.8) 所示的 $[a,b]$ 区间作周期运动，从而相轨迹是封闭的。

分成若干格，每格的面积为 $(2\pi\hbar)$，进而 (3.71) 告诉我们，我们将共得到 n 格，而 n 就是粒子第 n 能级波函数的节点数。另一方面，由于波函数每增加一个节点，就相应于势阱之内出现一个新的束缚态，因此 n 也就是能量不超过 E 的量子态的总数目。从而也就是说，在半经典近似下，相空间每一个面积为 $2\pi\hbar$ 的相格刚好对应一个量子态。从而相空间体积元 $\Delta x\Delta p$ 之内的量子态数目为

$$\frac{\Delta x\Delta p}{2\pi\hbar}. \tag{3.72}$$

这是用单个运动自由度所得出来的结果，将这个结果推广到任意的 s 个自由度 (记相应的正则坐标为 $q_i, i=1,2,...,s$，正则动量为 $p_i, i=1,2,...,s$)，则相空间体积元内的量子态数就应该是

$$\Delta N = \frac{\Delta q_1...\Delta q_s\Delta p_1...\Delta p_s}{(2\pi\hbar)^s}. \tag{3.73}$$

最后，利用 $S(x)$ 的 0 阶近似，并注意到 $\frac{dp}{dx}=\frac{d}{dx}[\sqrt{2m(E-V(x))}]=-\frac{dV}{dx}\frac{m}{p}$，我们可以将近似条件 (3.66) 重写为

$$\left|\lambda\frac{dV}{dx}/E_k\right| \ll 4\pi. \tag{3.74}$$

式中 λ 为粒子的德布罗意波长，$E_k=p^2/(2m)$ 为粒子的动能，因此这个条件说的无非是，在一个德布罗意波长的尺寸上势能的变化量要远小于动能。特别的，这个近似条件告诉我们，在 $p(x)=0$ 的经典转向点，即图 (3.8) 中所示的 a、b、c 点，我们的 WKB 近似解 (3.69) 是不成立的。所以真正用 WKB 近似来求解一维问题的时候，我们要分区域分别应用 WKB 近似解 (3.69)，然后在经典转向点附近将不同区域的解拼接起来。

量子隧穿的 WKB 近似

到此为止我们还没有看到 WKB 近似任何神奇的地方，看起来它只不过略微修正了一下德布罗意波。这是因为，我们心中默认了 $E>V(x)$。然而真正神奇的是，作为一种数学方法，并没有任何理由限制我们将上面的推导限制于 $E>V(x)$ 的区域，同样的推导也完全适用于 $E<V(x)$ 的在经典物理上不被允许的区域。只不过这时候 $p(x)=\sqrt{2m(E-V(x))}=-i\sqrt{2m(V(x)-E)}$ 是一个纯虚数[3]，因此相应的 WKB 近似解 (3.69) 是指数衰减或者指数增长的，

$$\psi(x)=C\frac{1}{\sqrt{p(x)}}e^{\pm\int^x\sqrt{2m(V(x)-E)}dx/\hbar}. \tag{3.75}$$

因此，根据 WKB 近似，我们可以大致理解图 (3.8) 中所示的波函数解。首先，在势阱之内的 $[a,b]$ 区间，$E>V(x)$，从而 WKB 波函数 (3.69) 是指数振荡的。其次，在势垒

[3]注意，这并不意味着粒子的动量可以是虚数，因为严格来说我们这一节所引入的 $p(x)$ 函数并不是测量粒子动量所得到的值，测量粒子动量所得的值当然要是动量算符的本征值，是一个数，而不是一个函数。当然，当 $E>V(x)$ 时我们可以将 $p(x)$ 理解为经典动量，而 $E<V(x)$ 时，可以将 $h/|p(x)|$ 看成是隧穿距离，而 $p(x)$ 本身则并没有很清楚的物理解释。

区，比如图 (3.8) 中的 $[b,c]$ 区间，$E<V(x)$，粒子的确会被势垒反射，但是，根据我们上一段的分析，$[b,c]$ 区间的波函数并不是 0，而是可以有一个非 0 的指数衰减的 WKB 波函数 (3.75) (指数增长解显然不符合物理要求)。最后出了势垒区，比方说在图 (3.8) 中的 $[c,+\infty)$ 区间, 波函数再次开始增荡，代表隧穿到势垒之外的粒子。

因此，根据 WKB 近似的分析我们可以知道，在量子力学的层次上，即使粒子的能量并不足够大，它也有一定的概率隧穿过势垒，从而挣脱势阱的束缚跑到外面去。以图 (3.8) 中所示的情形为例，我们可以定义粒子从 $[a,b]$ 区间隧穿到 $[c,+\infty)$ 区间的概率幅 T 为波函数在外内两个经典转向点的比值，也即是在 c 点测到粒子的概率幅比上在 b 点测到粒子的概率幅，即

$$T=\psi(c)/\psi(b). \tag{3.76}$$

根据波函数的连续性，我们可以在 $[b,c]$ 势垒区求解这个比值。严格求解当然会比较复杂，但是我们可以设想在 $[b+\varepsilon,c-\varepsilon]$ 的区域近似地使用 WKB 近似，从而根据 (3.75)(取指数衰减解)，我们就有

$$T\approx T_0 e^{-\frac{1}{\hbar}\int_b^c dx\sqrt{2m(V(x)-E)}}. \tag{3.77}$$

或者说，隧穿概率 $D=|T|^2$ 近似为

$$D\approx |T_0|^2 e^{-\frac{2}{\hbar}\int_b^c dx\sqrt{2m(V(x)-E)}}. \tag{3.78}$$

从这个结果可以看到，势垒越高隧穿的概率越小，势垒越厚，隧穿的概率也越小。

伽莫夫最早用量子隧穿效应解释了 α 粒子衰变。伽莫夫将放射性原子核看成一个口袋，这个口袋里面束缚了一些 α 粒子，但是，由于量子隧穿效应，口袋中的 α 粒子有一定的概率能穿过“口袋壁”的势垒跑到原子核外面去，这就是 α 粒子衰变。

1981 年，IBM 公司的宾尼希 (G. Binnig) 和罗勒 (H. Rohrer) 发明了基于量子隧穿效应的扫描隧穿显微镜 (scanning tunneling microscope, 简称 STM), 它使得科学家可以直接观察和定位单个原子，甚至可以精确操纵原子，从而极大地推动了纳米科技的发展。

当然，作为一种有别于经典物理的典型量子效应，量子隧穿效应起作用的地方远远不止这些，下面我们将介绍量子隧穿效应在超导物理的约瑟夫森结中如何起作用。

3.3.2 约瑟夫森结 (Josephson junction)

假设我们取两块超导体，中间通过一个绝缘体薄层将两者联结起来，绝缘层的厚度大概在 $1nm$ 量级，如图 (3.10) 所示。这就构成了一个约瑟夫森结，结两边的超导体我们分别标记为 1、2。我们可以将绝缘层看成是一个势垒，但是，只要绝缘层足够薄，那么对于库珀对而言，它完全有可能通过量子隧穿效应从超导体 1 隧穿到超导体 2，也有可能从超导体 2 隧穿到超导体 1。不妨分别记超导体 1、2 内的库珀对波函数为 $\psi_1=\sqrt{\rho_1}e^{i\theta_1}$ 和 $\psi_2=\sqrt{\rho_2}e^{i\theta_2}$，可以设想，如果 ψ_1 与 ψ_2 完全相同，那么从 1 隧穿到 2 的库珀对与从 2 隧穿到 1 的库珀对将会刚好平衡，从而结上并没有超导电流流过。

但是，如果结两边库珀对的相位不同，那 1 到 2 的隧穿和 2 到 1 的隧穿就可能不能完全平衡，从而在结上产生一个非 0 的隧穿超导电流 $\mathbf{J}_e$。由于 $\mathbf{J}_e=eN\mathbf{J}$($N$ 为材料中库珀

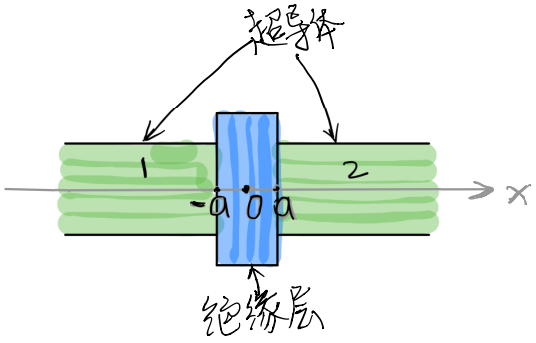

Figure 3.10: 约瑟夫森结。

对的总数目，e 为库珀对的电荷)，因此这也就是说，这时候单个库珀对在结上可能有非 0 的概率流密度。同时，根据局域概率守恒，流过结的超导概率流密度 $\mathbf{J}$ 必定来源于左右两边的超导体，因此，为了维持结上的超导电流，我们需要把两个超导区域 1 和 2 联接起来形成一个闭合回路。

由于超导体内部没有电流，所以如果将如图 (3.10) 所示的两个超导区域 1 和 2 看成是两个沿着 x 轴放置的超导圆柱体的话，那超导电流当然分布在圆柱体的表面并沿着 x 轴方向流动。但是，超导电流在超导区域上具体如何分布并不是我们这里关心的重点，由于我们主要关心电流沿 x 轴的流动，所以为了简化问题，我们不妨将整个约瑟夫森结看成是沿着 x 轴的一条线，这条线的左边是超导线，中间是很短的一段绝缘线，右边又是超导线。作这样的一维简化以后，电流密度和电流强度当然就是一回事了。

我们不妨记约瑟夫森结左边的超导概率流密度为 $\mathbf{J}_1$，右边的概率流密度为 $\mathbf{J}_2$。根据对 (3.38) 式的讨论，库珀对在约瑟夫森结左边运动时，能量为 $\frac{m}{2\rho_1^2}\mathbf{J}_1^2$，而在结右边运动时，能量为 $\frac{m}{2\rho_2^2}\mathbf{J}_2^2$，但是，由于库珀对可以通过量子隧穿效应从 1 跑到 2 或者从 2 跑到 1，而库珀对隧穿前后的能量应该是守恒的，从而 $\frac{m}{2\rho_1^2}\mathbf{J}_1^2 = \frac{m}{2\rho_2^2}\mathbf{J}_2^2$，我们不妨将这个能量记为 ε。在实际应用中，约瑟夫森结两边总是会取同一种超导材料，从而实际上 $\rho_1 = \rho_2 = \rho$，是一个常数。而且当超导电流达到稳恒时，对于一维电流流动，稳恒电流条件告诉我们，结两边的超导电流密度以及结上的电流密度必定都相等，从而 $\mathbf{J}_1 = \mathbf{J}_2 = \mathbf{J}$，因此隧穿前后，库珀对的能量守恒条件总能满足，相应的能量可以写作 $\varepsilon = \frac{m}{2\rho^2}\mathbf{J}^2$。下面我们要做的就是找到流过结的电流密度 $\mathbf{J}$ 的基本规律。

为此，假设绝缘层的势垒为 $V(x)$，$V(x) > \varepsilon$，且绝缘区为从 $x = -a$ 到 $x = a$ 这样一个厚度为 $2a$ 的薄层内，如图 (3.10) 所示。我们记 $\kappa(x) = \sqrt{2m(V(x) - \varepsilon)}$，则根据一维定态薛定谔方程的 WKB 近似 (3.75)，绝缘区的波函数必定为 WKB 指数衰减解和指数增长解的线性组合，因此，我们总是可以将绝缘区的 WKB 波函数写成

$$\psi(x) = Ae^{-\frac{1}{\hbar}\int_{-a}^{x}\kappa(x)dx} + Be^{-\frac{1}{\hbar}\int_{x}^{a}\kappa(x)dx}. \tag{3.79}$$

记 $d = e^{-\frac{1}{\hbar}\int_{-a}^{a}\kappa(x)dx}$，则有 $\psi(-a) = A + Bd$，$\psi(a) = Ad + B$。另外，由于 $\psi(-a)$ 在超导体 1 内，从而 $\psi(-a) = \sqrt{\rho_1}e^{i\theta_1}$，同样 $\psi(+a) = \sqrt{\rho_2}e^{i\theta_2}$(当然，正如我们上面说过的，实际上

$\rho_2=\rho_1$)。进而我们可以解得

$$A=\frac{\sqrt{\rho_1}e^{i\theta_1}-d\sqrt{\rho_2}e^{i\theta_2}}{1-d^2},\quad B=\frac{\sqrt{\rho_2}e^{i\theta_2}-d\sqrt{\rho_1}e^{i\theta_1}}{1-d^2}. \tag{3.80}$$

将绝缘区的 WKB 波函数 (3.79) 代入概率流密度的表达式 (3.34), 就可以得到

$$\begin{aligned}\mathbf{J}&=\mathbf{e}_x\frac{\kappa(x)}{m}(-i)(A^*B-B^*A)d=2\mathbf{e}_x\frac{\kappa(x)}{m}\mathrm{Im}(A^*B)d\\&=\mathbf{e}_x\frac{2d}{1-d^2}\frac{\kappa(x)}{m}\sqrt{\rho_1\rho_2}\sin(\theta_2-\theta_1)\\&=\mathbf{J}_C\sin(\theta_2-\theta_1).\end{aligned} \tag{3.81}$$

式中 $\mathbf{J}_C=\mathbf{e}_x\frac{2d}{1-d^2}\frac{\kappa(x)}{m}\sqrt{\rho_1\rho_2}$，$\mathbf{e}_x$ 为 x 方向的单位矢量。可见，穿过约瑟夫森结的超导电流取决于结两边的相位差，只要两边有相位差，通常就有电流通过。如果不只有相位差，还有矢量势穿过约瑟夫森结，那么根据规范不变性 (3.51), 可将相位差推广为，

$$\Delta\theta_A=\theta_2-\theta_1-\frac{e}{\hbar}\int_{-a}^{a}\mathbf{A}\cdot\mathbf{dl}. \tag{3.82}$$

这时候穿过结的超导概率流密度就应该是

$$\mathbf{J}=\mathbf{J}_C\sin(\Delta\theta_A). \tag{3.83}$$

但是，如果我们不只是把两个超导区域 1 和 2 联接起来形成闭合回路。而是更一般地，将两个超导区域分别联至电池的两端，从而给结两边加上一个电压 U(U 可以为 0)，那么情况将发生什么变化呢？当然，正如我们将要看到的，这样一来，库珀对就不再能维持在能量确定的定态了，穿过结的超导概率流密度也将随着时间演化。不过，可以论证，(3.83) 式依然成立。

但是，有了电压以后，相位差 $\Delta\theta_A$ 将随着时间演化。这是因为，首先，对于任何波函数为 $\psi=\sqrt{\rho}e^{i\theta}$ 的超导体，如果它的外电势为 ϕ(从而势能 $V=e\phi$), 则由 (3.38) 式可知，

$$-\hbar\partial_t\theta=\varepsilon+e\phi, \tag{3.84}$$

式中 ε 就是前面的 $\frac{m}{2\rho^2}\mathbf{J}^2$。进而我们就可以知道，对于约瑟夫森结，设结两边的电势分别为 ϕ_1 和 ϕ_2，则

$$\begin{aligned}\partial_t(\Delta\theta_A)&=\big(\partial_t\theta_2-\partial_t\theta_1\big)-\frac{e}{\hbar}\int_{-a}^{a}\partial_t\mathbf{A}\cdot\mathbf{dl}\\&=-\frac{e}{\hbar}\big(\phi_2-\phi_1\big)-\frac{e}{\hbar}\int_{-a}^{a}\partial_t\mathbf{A}\cdot\mathbf{dl}\\&=\frac{2\pi}{\Phi_0}\int_{-a}^{a}(-\nabla\phi-\partial_t\mathbf{A})\cdot\mathbf{dl}\\&=\frac{2\pi}{\Phi_0}\int_{-a}^{a}\mathbf{E}\cdot\mathbf{dl}=\frac{2\pi}{\Phi_0}U.\end{aligned} \tag{3.85}$$

我们可以将结果 (3.83) 和结果 (3.85) 概括在一起，并把概率流密度替换成通过结的电流强度，从而就有基本方程组

$$I=I_C\sin(\Delta\theta_A),\quad U=\frac{\Phi_0}{2\pi}\partial_t(\Delta\theta_A). \tag{3.86}$$

这组方程就是所谓的 Josephson 关系式。从这组方程可以看到，如果结两端的电压 $U=0$，那么一般来说流过结的电流强度并不为 0！而可以是从 $-I_C$ 到 I_C 之间的任何值。在 0 电压之下可以有有限的电流，当然就意味着 0 电阻，这正是超导的特性之一。相反，如果给 Josephson 结两边加一个有限的直流电压 U_0，那么根据 (3.86) 式，将有 $(\Delta\theta_A)=(\Delta\theta_A)_0+\frac{2\pi U_0}{\Phi_0}t$。进而，$I=I_C\sin\left((\Delta\theta_A)_0+\frac{2\pi U_0}{\Phi_0}t\right)$，从而穿过结的超导电流将以角频率 $\frac{2\pi U_0}{\Phi_0}$ 快速振荡 (因为磁通量子 Φ_0 很小)，从而平均来说，结上的超导电流反而为 0。当然，这种电流的快速振荡一定会辐射电磁波，实际上这种电磁辐射就来源于库珀对从 Josephson 结的一边隧穿到另一边时，由于电势能的落差而辐射出来的光子。

值得说明的是，Josephson 结在超导量子计算中有很重要的应用，它可以用来实现超导量子比特，我们在后面的章节中还会回到这个问题。

3.3.3 习题

1. 沿着 x 轴运动的一维线性谐振子的势能为 $V(x)=\frac{1}{2}m\omega^2x^2$，式中 m 是谐振子的质量，ω 是线性谐振子的角频率。请用 Bohr-Sommerfeld 量子化条件导出一维线性谐振子的量子化能级，为 $E_n=(n+\frac{1}{2})\hbar\omega$。

2. 对于薛定谔波动方程 $\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\mathbf{x},t)=\left[-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})\right]\psi(\mathbf{x},t)$，假设令 $\psi(\mathbf{x},t)=\sqrt{\rho(\mathbf{x},t)}e^{iS(\mathbf{x},t)/\hbar}$，式中 $S(\mathbf{x},t)$ 为一个实函数，$\rho(\mathbf{x},t)$ 为概率密度。请证明，在 $\hbar\to 0$ 的极限下，$S(\mathbf{x},t)$ 满足经典力学的哈密顿-雅可比方程，即满足

$$\frac{\partial S}{\partial t}+\frac{1}{2m}(\nabla S)^2+V(\mathbf{x})=0. \tag{3.87}$$

因此这就证明了，经典力学可以作为量子力学在 $\hbar\to 0$ 时的极限。

3.4 自旋 1/2 的带电粒子

3.4.1 泡利算符

前面我们已经研究了磁场中的一个带电粒子，比如说电子，的薛定谔波动方程。但是斯特恩-格拉赫实验已经告诉我们了，电子还有一个内禀自由度，即自旋，电子的自旋会产生一个自旋磁矩，当电子处在磁场中的时候，这个自旋磁矩就会和磁场相耦合。为了反映电子自旋这个内禀自由度，除了位置和动量以外，人们还需要额外引入了一个新的基本物理量，即电子的自旋角动量。电子的自旋角动量纯粹是一个量子力学的物理量，它并没有经典对应，因此它只能用算符来描述，我们记为 $\mathbf{S}=(S_x,S_y,S_z)$，$\mathbf{S}$ 的三个分量我们记作 $S_i,i=1,2,3$ 分别对应 S_x,S_y,S_z。电子的自旋只有两个取值 $\pm\hbar/2$，这意味着自旋算符的每一个分量都只有 $\pm\hbar/2$ 的本征值。

等价的，人们也常用泡利算符 $\vec{\sigma}$ 来描述电子的自旋，它和自旋算符的关系是

$$\mathbf{S}=\frac{\hbar}{2}\vec{\sigma}. \tag{3.88}$$

在第二章中我们已经证明，泡利算符的三个分量是两两反对易的，比如说 $\sigma_x\sigma_y=-\sigma_y\sigma_x$。另外，泡利算符每一个分量的平方都等于 1，即 $\sigma_x^2=\sigma_y^2=\sigma_z^2=1$。不光如此，泡利算符

还满足 $\sigma_x\sigma_y\sigma_z = i$。实际上，我们可以将泡利算符的这些性质归纳为如下这个式子

$$\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k, \tag{3.89}$$

这里，我们应用了爱因斯坦求和约定，因此在公式 (3.89) 中我们要对 k 指标进行求和，从 1 加到 3。后面我们将对所有将要出现的公式都使用这样的求和约定。当然，也请大家注意公式 (3.89) 中出现了两个含义不同的 i，不同于作为指标的 i，作为系数出现的那个是表示虚数单位 i。

而泡利算符有一个很巧妙的应用，那就是可以用来将诸如 $\sqrt{a_x^2+a_y^2+a_z^2}$ 这样的开方表达出来 (这里 a_x, a_y, a_z 是矢量 $\mathbf{a}$ 的三个分量)。我们先把这个答案写出来，然后再来证明它

$$\sqrt{a_x^2+a_y^2+a_z^2} = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z. \tag{3.90}$$

证明实际上非常简单，人们直接将这个式子两边平方，然后利用泡利算符两两反对易的性质以及平方等于 1 的性质就能得到结果。用矢量形式来说，证明上面这个式子其实就是要证明

$$\mathbf{a}^2 = (\mathbf{a}\cdot\vec{\sigma})^2. \tag{3.91}$$

利用 $(\mathbf{a}\cdot\vec{\sigma})^2 = a_i\sigma_i a_j\sigma_j = a_i a_j(\delta_{ij} + i\varepsilon_{ijk}a_k) = a_i a_i$ 就得到了证明，推导过程中我们利用了方程 (3.89), 并且利用了 $a_i a_j\varepsilon_{ijk} = 0$(其实就是 $(\mathbf{a}\times\mathbf{a})_k = 0$)。

由于自旋和空间变量是相互独立的自由度，所以泡利算符与位置算符，泡利算符与动量算符，更一般的，泡利算符与任何只涉及位置和动量的算符都是对易的。与这相关，有一个有用的公式我们后面会用到，即假设某个矢量算符 $\mathbf{K}$ 只涉及位置和动量 (因此与泡利算符对易)，那么

$$(\vec{\sigma}\cdot\mathbf{K})^2 = \mathbf{K}^2 + i\vec{\sigma}\cdot(\mathbf{K}\times\mathbf{K}). \tag{3.92}$$

利用泡利算符的代数关系式 (3.89) 人们容易证明这一结果。值得注意的是，因为 $\mathbf{K}$ 是算符，它的不同分量之间一般来说不对易，因此 $\mathbf{K}\times\mathbf{K} \neq 0$(比方说 $(\mathbf{K}\times\mathbf{K})_z = K_xK_y - K_yK_x$ 一般来说就不为零)，当然，如果某个 $\mathbf{K}$ 的不同分量相互对易，那 $\mathbf{K}\times\mathbf{K}$ 就和普通的矢量叉乘一样等于 0 了。

考虑到自旋以后，我们该怎么写磁场中的电子的哈密顿算符呢？最简洁的一个写法是泡利给出来的，

$$H = \frac{[\vec{\sigma}\cdot(\mathbf{P} - e\mathbf{A})]^2}{2m} + V(\mathbf{X}). \tag{3.93}$$

利用刚才我们得到的 (3.91) 式，人们容易证明，当磁场等于零时，即 $\mathbf{A} = 0$ 时，这个哈密顿算符就变成标准的不考虑自旋的非相对论粒子的哈密顿算符。那么有磁场时这个结果对不对呢？这要看它能不能和实验相吻合，我们将在下一小节进一步讨论这一点。另一个问题是，泡利为什么不再加上 $\vec{\sigma}\cdot\mathbf{P}$ 这样的项呢？原因在于，这样的项不满足空间反演不变性。因为在空间反演之下，$\mathbf{P}\to -\mathbf{P}$, 但是自旋是一种角动量，而角动量是一种赝矢量，也就是说，在空间反演之下 $\vec{\sigma}\to\vec{\sigma}$, 因此 $\vec{\sigma}\cdot\mathbf{P}$ 就不是空间反演不变的。

3.4.2 旋量波函数与自旋 1/2 薛定谔波动方程

考虑到电子的自旋以后，希尔伯特空间的矢量基就不能只取位置本征态 $|\mathbf{x}\rangle$ 了，还必须同时把自旋态包括进来。人们通常还选择自旋算符 S_z 的两个本征态，即本征值为 $\hbar/2$ 的自旋向上态 $|\uparrow\rangle$ 以及本征值为 $-\hbar/2$ 的自旋向下态 $|\downarrow\rangle$。因此完整的矢量基要选 $\{|\mathbf{x},\uparrow\rangle,|\mathbf{x},\downarrow\rangle\}$ 这样的态构成，它们是位置算符和 S_z 的共同本征态。矢量基的完备性关系 (3.4) 现在就应该写成

$$\int d\tau\big(|\mathbf{x},\uparrow\rangle\langle\mathbf{x},\uparrow|+|\mathbf{x},\downarrow\rangle\langle\mathbf{x},\downarrow|\big)=1, \tag{3.94}$$

也就是说既要对向上向下自旋求和，也要对空间坐标进行积分，式中 $d\tau=d^3\mathbf{x}$。

如此一来，电子的任意一个态矢量 $|\psi\rangle$ 就应该展开成 $|\psi\rangle=\int d\tau|\mathbf{x},\uparrow\rangle\langle\mathbf{x},\uparrow|\psi\rangle+\int d\tau|\mathbf{x},\downarrow\rangle\langle\mathbf{x},\downarrow|\psi\rangle$。引入波函数 $\psi_\uparrow(\mathbf{x})=\langle\mathbf{x},\uparrow|\psi\rangle$，和波函数 $\psi_\downarrow(\mathbf{x})=\langle\mathbf{x},\downarrow|\psi\rangle$，态矢量的展开就可以写成 $|\psi\rangle=\int d\tau\psi_\uparrow(\mathbf{x})|\mathbf{x},\uparrow\rangle+\int d\tau\psi_\downarrow(\mathbf{x})|\mathbf{x},\downarrow\rangle$，这也就是说，在坐标表象中，电子的态矢量 $|\psi\rangle$ 应该用一对波函数 $(\psi_\uparrow(\mathbf{x}),\psi_\downarrow(\mathbf{x}))$ 来表示，我们把这一对波函数排成一个两分量旋量 $\Psi(\mathbf{x})$，称作旋量波函数，即

$$\Psi(\mathbf{x})=\begin{pmatrix}\psi_\uparrow(\mathbf{x})\\ \psi_\downarrow(\mathbf{x})\end{pmatrix}. \tag{3.95}$$

类似的，我们可以将电子的另一个任意态矢量 $|\phi\rangle$ 所对应的旋量波函数记为 $\Phi(\mathbf{x})$，那么完全类似于前面没有自旋的情形，我们将有

$$\langle\phi|\psi\rangle=\int d\tau\Phi^\dagger(\mathbf{x})\Psi(\mathbf{x}). \tag{3.96}$$

当然，由于取了坐标表象，所以这时候动量算符就应该表示成 $-i\hbar\nabla$。又由于也取了自旋向上和自旋向下两个自旋基矢量，所以泡利算符 $\vec{\sigma}$ 就要表示成 2×2 的厄米矩阵 $\widehat{\vec{\sigma}}$，这就是所谓的泡利矩阵，它们的具体形式可以由下式给出，

$$\widehat{\sigma}_x=\begin{pmatrix}0&1\\1&0\end{pmatrix},\widehat{\sigma}_y=\begin{pmatrix}0&-i\\i&0\end{pmatrix},\widehat{\sigma}_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{3.97}$$

当然，泡利矩阵作为泡利算符的矩阵表示，它们同样也满足泡利算符的代数关系 (3.89)。一般地，我们可以把电子的任意一个抽象算符 $\mathscr{O}$ 表示成 $\widehat{\mathscr{O}}$，一般来说 $\widehat{\mathscr{O}}$ 将既是一个微分算符，同时又是作用在二分量旋量上的 2×2 矩阵。完全类似于没有自旋的情况，我们将有

$$\langle\phi|\mathscr{O}|\psi\rangle=\int d\tau\Phi^\dagger(\mathbf{x})\widehat{\mathscr{O}}\Psi(\mathbf{x}). \tag{3.98}$$

如此一来，根据泡利给出的哈密顿算符 (3.93)，我们就可以把磁场中电子的薛定谔方程写成

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t)=\left[-\frac{\hbar^2}{2m}\left(\widehat{\vec{\sigma}}\cdot(\nabla-\frac{ie}{\hbar}\mathbf{A})\right)^2+V(\mathbf{x})\right]\Psi(\mathbf{x},t). \tag{3.99}$$

下面我们就来检验一下 (3.99) 这个方程到底能不能解释实验所观测到的事实。为此我们主要想把式中的 $\left(\widehat{\vec{\sigma}}\cdot(\nabla-\frac{ie}{\hbar}\mathbf{A})\right)^2$ 再进一步计算出来。首先，根据公式 (3.92)，$\left(\widehat{\vec{\sigma}}\cdot(\nabla-\frac{ie}{\hbar}\mathbf{A})\right)^2=(\nabla-\frac{ie}{\hbar}\mathbf{A})^2+i\widehat{\vec{\sigma}}\cdot[(\nabla-\frac{ie}{\hbar}\mathbf{A})\times(\nabla-\frac{ie}{\hbar}\mathbf{A})]$。很显然 $\nabla\times\nabla=0$, $\mathbf{A}\times\mathbf{A}=0$，因此 $(\nabla-\frac{ie}{\hbar}\mathbf{A})\times(\nabla-\frac{ie}{\hbar}\mathbf{A})=-\frac{ie}{\hbar}(\nabla\times\mathbf{A}+\mathbf{A}\times\nabla)$。注意 $\nabla\times\mathbf{A}$ 是算符的相乘，因此 $\nabla\times\mathbf{A}=(\nabla\times\mathbf{A})-\mathbf{A}\times\nabla$，这里的 $(\nabla\times\mathbf{A})$ 就是普通的旋度，因此就等于磁场强度 $\mathbf{B}$, 而等式右边第二项的负号来自于矢量叉乘交换顺序会出一个负号。因此 $(\nabla-\frac{ie}{\hbar}\mathbf{A})\times(\nabla-\frac{ie}{\hbar}\mathbf{A})=-\frac{ie}{\hbar}\mathbf{B}$。因此 $\left(\widehat{\vec{\sigma}}\cdot(\nabla-\frac{ie}{\hbar}\mathbf{A})\right)^2=(\nabla-\frac{ie}{\hbar}\mathbf{A})^2+\frac{e}{\hbar}\mathbf{B}\cdot\widehat{\vec{\sigma}}$，代入 (3.99) 就可以得到，

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t)=\left[-\frac{\hbar^2}{2m}(\nabla-\frac{ie}{\hbar}\mathbf{A})^2-\frac{e\hbar}{2m}\widehat{\vec{\sigma}}\cdot\mathbf{B}+V(\mathbf{x})\right]\Psi(\mathbf{x},t). \tag{3.100}$$

我们看到，多出来的这一项 $-\frac{e\hbar}{2m}\widehat{\vec{\sigma}}\cdot\mathbf{B}$ 完全吻合实验中发现的电子自旋磁矩和磁场的耦合，如果将这一项写成标准的 $-\vec{\mu}_S\cdot\mathbf{B}$ 的形式，我们就会发现，自旋磁矩 $\vec{\mu}_S=\frac{e\hbar}{2m}\widehat{\vec{\sigma}}=\frac{e}{m}\widehat{\mathbf{S}}$，这正是完全正确的电子自旋磁矩的形式 (当然我们要注意到电子电荷 e 为负)，尤其是正确地得到了比通常的轨道角动量磁矩公式多出来的 2 倍因子。由此可见，方程 (3.99) 与实验是吻合的。

3.4.3 习题

1. 请证明 $(\mathbf{a}\cdot\vec{\sigma})(\mathbf{b}\cdot\vec{\sigma})=(\mathbf{a}\cdot\mathbf{b})+i(\mathbf{a}\times\mathbf{b})\cdot\vec{\sigma}$, 式中 $\mathbf{a}$、$\mathbf{b}$ 为两个普通的矢量，不是算符。

2. 请证明 $e^{i\frac{\theta}{2}(\mathbf{n}\cdot\vec{\sigma})}=\cos\frac{\theta}{2}+i(\mathbf{n}\cdot\vec{\sigma})\sin\frac{\theta}{2}$，式中 $\mathbf{n}$ 为一个任意的单位矢量，θ 是一个角度。

3. 我们已经知道了旋量波函数的概念。在数学上，旋量也可以定义成这样的一个两分量的量 $\phi=\binom{\phi_1}{\phi_2}$，它在绕 $\mathbf{n}$ 轴旋转 θ 角时按照 $\phi\to\exp\left(i\frac{\theta}{2}\mathbf{n}\cdot\widehat{\vec{\sigma}}\right)\phi$ 的方式变化。现在，请证明由旋量 ϕ 构造出来的量 $\phi^\dagger\widehat{\vec{\sigma}}\phi$ 是一个通常的三维矢量，也即是说，在空间旋转之下，这个量像通常的矢量那样变化。



4. 几个简单而重要的系统

本章主要讨论几个简单而又重要的量子力学系统。它们包括，两自旋耦合系统、一维势散射的一个简单例子、单自由度线性谐振子、氢原子、以及均匀磁场中的电子，最后这个系统也就是著名的朗道能级。这些系统都精确可解。我们介绍的求解方法是多样的: 对于两自旋耦合系统我们介绍了矩阵对角化的求解方法; 对于一维势散射和氢原子问题，我们介绍了解波动力学的定态薛定谔方程; 对于单自由度线性谐振子，我们介绍了算符代数的方法; 而对于均匀磁场中的电子，我们结合了代数法和求解波函数的方法。

当然，每一个例子的求解方法往往都不唯一，比如单自由度线性谐振子也可以通过定态薛定谔波动方程来求解，而氢原子问题也可以利用它隐藏的 $SO(4)$ 动力学对称性进而用算符代数的方法求解。在每一个例子中，我们都仅仅只是介绍更简洁同时相对来说更重要的求解方法。

作为朗道能级的一个重要物理应用，在本章的最后一节我们讨论了在凝聚态物理领域极其重要的量子霍尔效应，主要是讨论了整数量子霍尔效应的物理机制。当然，量子霍尔效应本身是一个很大的课题，我们只可能作一个初步的介绍。

4.1 两自旋耦合系统

铁，钴，镍等物质在很小的外磁场影响下，就能产生远大于其他物质的磁化效应，这就是铁磁性。这些物质之所以有铁磁性，是因为在它们的每一个小区域之内都会自发磁化形成磁畴。而自发磁化的原因，是因为铁磁物质不同原子的磁矩产生了相互平行的指向，就像这样 (↑↑↑↑↑↑↑↑↑↑↑↑↑↑↑)。这些原子的磁矩从哪儿来呢？来自于铁磁原子未配对的核外电子的自旋，已经配对的那些电子由于泡利不相容原理，它们的自旋一定是相反的，因而磁矩总是抵消为 0。那么不同原子的这些未配对电子的自旋为什么会平行地指向呢？为什么只有铁磁性物质会这样平行指向呢？为了对这一问题进行理论研究，海森堡提出，这是因为铁磁性物质邻近原子的未配对电子存在自旋与自旋间的相互作用，为了反映这种自旋相互作用是如何导致自发磁化的，海森堡提出了一个简单的模型，这就是今天依然有大量研究者在研究的著名的海森堡模型，这个模型对于铁磁性的研究，甚至对于整个凝聚态物理来说都非常重要，因此我们想介绍一下。海森堡模型的哈密顿算符是下面这样的

$$H = J \sum_{<i,j>} \vec{S}_i \cdot \vec{S}_j, \tag{4.1}$$

式中的 $< i, j >$ 表示两个邻近格点位置 i 和 j, $\vec{S}_i, \vec{S}_j$ 分别表示格点位置 i, j 上的原子的自旋，式中的求和表示将每一对相互邻近的原子的贡献都加起来，另外，J 的大小表示两自旋之间的耦合强度，更常见的是将之写成 $-J$，不过后面我们对 J 取正和取负两种情况都会进行讨论。

对于一个任意的两维或三维格点，求解海森堡模型不是一件容易的事情，因此我们当然不可能在这里讨论这样的课题。以上所说主要是给我们提供一个物理背景，我们真正将要研究的，是只考虑两个原子，每个原子只有一个未配对电子，这两个电子的自旋之间按照海森堡模型的形式进行耦合。这时候，由于电子的自旋算符可以用泡利算符来进行表达，所以我们也可以将相应的模型写成下面的形式，

$$H = J(\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y + \sigma_1^z \sigma_2^z). \tag{4.2}$$

(4.2) 才是我们这一节将要进行求解的问题。我们将会看到，即使这样简单的一个模型也能告诉我们，什么时候系统会出现铁磁性，也就是两电子的自旋相平行，什么时候系统又会出现反铁磁性，也就是两电子的自旋反平行。不过，首先我们想说一下，什么是求解出这个系统？我们这里所谓的求解出哈密顿算符 (4.2)，实际上指的是求解出它的本征值，也就是系统的本征能量 $\{E_n\}$，以及相应的本征态 $\{|\psi_n\rangle\}$

$$H|\psi_n\rangle = E_n|\psi_n\rangle. \tag{4.3}$$

正如我们已经知道的，由于在哈密顿算符的本征态 $|\psi_n\rangle$ 上，系统有确定的能量取值 E_n，所以这样的本征态也叫做定态，相应的本征能量的集合也称作系统的能谱或者能级。

4.1.1 泡利算符回顾

为了求解出上面的两电子自旋耦合系统的能谱，这一小节让我们先回顾一下关于泡利算符的一些相关知识。首先，泡利算符具有下面的基本代数关系，

$$[\sigma^x, \sigma^y] = 2i\sigma^z, [\sigma^y, \sigma^z] = 2i\sigma^x, [\sigma^z, \sigma^x] = 2i\sigma^y. \tag{4.4}$$

并且三个泡利算符的平方均等于 1，即 $(\sigma^x)^2=(\sigma^y)^2=(\sigma^z)^2=1$。由这两组基本代数关系可以进一步导出三个泡利算符是两两反对易的。

在第二章中我们还曾经引入 $\sigma=\frac{1}{2}(\sigma^x-i\sigma^y),\sigma^\dagger=\frac{1}{2}(\sigma^x+i\sigma^y)$, 并且我们也知道 $\sigma^2=(\sigma^\dagger)^2=0$。另外，利用对易子代数关系 (4.4)，我们可以得到

$$[\sigma^z,\sigma]=-2\sigma,\quad [\sigma^z,\sigma^\dagger]=2\sigma^\dagger,\quad [\sigma^\dagger,\sigma]=\sigma^z. \tag{4.5}$$

设 $|s\rangle$ 为 σ^z 的本征值为 s 的本征态，满足本征方程 $\sigma^z|s\rangle=s|s\rangle$。由 $[\sigma^z,\sigma^\dagger]=2\sigma^\dagger$ 可以知道 $\sigma^z\sigma^\dagger=\sigma^\dagger(\sigma^z+2)$, 因此 $\sigma^z(\sigma^\dagger|s\rangle)=\sigma^\dagger(\sigma^z+2)|s\rangle=(s+2)(\sigma^\dagger|s\rangle)$，也就是说，$\sigma^\dagger|s\rangle$ 如果非 0，则必为 σ^z 的本征值为 $s+2$ 的本征态，即必有

$$\sigma^\dagger|s\rangle=a_s|s+2\rangle, \tag{4.6}$$

其中 a_s 为一个待定复数。类似的，由 $[\sigma^z,\sigma]=-2\sigma$ 可以知道，必有

$$\sigma|s\rangle=b_s|s-2\rangle, \tag{4.7}$$

其中 b_s 为待定复数。正因为如此，σ 通常被称作自旋降算符，因为当它作用在 σ^z 的本征态上时，会把本征值 s 降 2，变成 $s-2$，相反 $\sigma^\dagger$ 称为自旋升算符，因为当它作用在 σ^z 的本征态上时，会把本征值升 2。

由于 $(\sigma^z)^2=1$, 所以其本征值必为 ±1, 即 $s=\pm1$。设相应的本征态分别为 $|\downarrow\rangle,|\uparrow\rangle$, 其中 $|\downarrow\rangle$ 的本征值为 -1，$|\uparrow\rangle$ 的本征值为 $+1$, 并且我们假定它们都已经归一化了，因而是正交归一的。由于 $|\downarrow\rangle$ 具有最低的本征值 -1，所以它不能再降了，再降就必定为 0，因此我们必有

$$\sigma|\downarrow\rangle=0. \tag{4.8}$$

类似的，$|\uparrow\rangle$ 具有最高的本征值 $+1$，所以它不能再升了，再升必定等于 0，即有

$$\sigma^\dagger|\uparrow\rangle=0. \tag{4.9}$$

另外, 我们已经知道 $\sigma^\dagger|\downarrow\rangle=a_{-1}|\uparrow\rangle$, $\sigma|\uparrow\rangle=b_{+1}|\downarrow\rangle$。为了求出 a_{-1} 和 b_{+1}, 我们将这两个结果厄米共轭就得到 $\langle\downarrow|\sigma=\langle\uparrow|a_{-1}^*$, $\langle\uparrow|\sigma^\dagger=\langle\downarrow|b_{+1}^*$。利用 $|a_{-1}|^2\langle\uparrow|\uparrow\rangle=\langle\downarrow|\sigma\sigma^\dagger|\downarrow\rangle=\langle\downarrow|[\sigma,\sigma^\dagger]|\downarrow\rangle=\langle\downarrow|-\sigma^z|\downarrow\rangle=\langle\downarrow|\downarrow\rangle=1$(其中第二个等号利用了 $\sigma|\downarrow\rangle=0$)，因此就有 $|a_{-1}|^2=1$，类似的我们也可以得到 $|b_{+1}|^2=1$。由于现在我们有两个本征态 $|\downarrow\rangle,|\uparrow\rangle$，同时有两个不为零的复系数 a_{-1},b_{+1}, 因此我们总是可以通过分别将这两个本征态乘上合适的复相位，从而使得 a_{-1},b_{+1} 都成为正实数。如此一来 $|a_{-1}|^2=1$ 和 $|b_{+1}|^2=1$，就意味着 $a_{-1}=b_{+1}=1$，从而我们就得到

$$\sigma^\dagger|\downarrow\rangle=|\uparrow\rangle,\quad \sigma|\uparrow\rangle=|\downarrow\rangle. \tag{4.10}$$

公式 (4.8)、(4.9) 以及公式 (4.10)，再加上下面的两个本征方程，就是我们后面会用到的一些基本结论，

$$\sigma^z|\uparrow\rangle=|\uparrow\rangle,\quad \sigma^z|\downarrow\rangle=-|\downarrow\rangle. \tag{4.11}$$

4.1.2 求解与讨论

下面我们就可以开始求解 (4.2) 给出来的两自旋耦合系统了。首先，两个不同电子的自旋当然是相互独立的，因此任何两个属于不同电子的泡利算符都必定相互对易。其次，类似于上一小节的做法，我们分别引入两个不同电子的自旋降算符和自旋升算符，$\sigma_1=\frac{1}{2}(\sigma_1^x-i\sigma_1^y),\sigma_2=\frac{1}{2}(\sigma_2^x-i\sigma_2^y)$, $\sigma_1^\dagger=\frac{1}{2}(\sigma_1^x+i\sigma_1^y),\sigma_2^\dagger=\frac{1}{2}(\sigma_2^x+i\sigma_2^y)$。则很容易验证，原来的哈密顿量 (4.2) 可以重写成

$$H=2J(\sigma_1\sigma_2^\dagger+\sigma_1^\dagger\sigma_2)+J(\sigma_1^z\sigma_2^z). \tag{4.12}$$

现在，我们给系统选定四个正交归一基矢量 $\{|\uparrow\uparrow\rangle,|\downarrow\uparrow\rangle,|\uparrow\downarrow\rangle,|\downarrow\downarrow\rangle\}$(请注意这四个基矢量的排列顺序，后面的处理要与这个顺序一致)，以基矢量 $|\downarrow\uparrow\rangle$ 为例，它两个标记符号的前一个 $\downarrow$ 表示第 1 个电子的自旋状态，后一个符号 $\uparrow$ 表示第 2 个电子的自旋状态，其余三个基矢量也是用类似办法进行标记的。

为了求解我们的系统，我们需要将哈密顿算符在这四个基矢量构成的矢量基中表示出来。为了做到这一点，让我们首先注意一个简单的数学事实，即，如果有归一化的量子态 $|u\rangle,|v_1\rangle,|v_2\rangle$，并且 $|v_1\rangle$ 和 $|v_2\rangle$ 正交，若某算符 A 将 $|u\rangle$ 映射到 $|v_1\rangle$ 和 $|v_2\rangle$ 的某个叠加态，比如 $A|u\rangle=\alpha_1|v_1\rangle+\alpha_2|v_2\rangle(\alpha_1,\alpha_2$ 为叠加系数)，那么利用 $|v_1\rangle$ 和 $|v_2\rangle$ 的正交关系可知必有 $\langle v_1|A|u\rangle=\alpha_1,\langle v_2|A|u\rangle=\alpha_2$。因此为了求出哈密顿算符在我们的矢量基中的表示矩阵，我们只需要求出它对四个基矢量的作用。为了求出这样的作用，我们只需要反复利用上一节中的简单公式 (4.8)、(4.9)、(4.10) 以及 (13.35)。比如，$H|\uparrow\uparrow\rangle=2J(\sigma_1\sigma_2^\dagger+\sigma_1^\dagger\sigma_2)|\uparrow\uparrow\rangle+J(\sigma_1^z\sigma_2^z)|\uparrow\uparrow\rangle=J|\uparrow\uparrow\rangle$ (注意，$\sigma^\dagger|\uparrow\rangle=0$)，即 $H|\uparrow\uparrow\rangle=J|\uparrow\uparrow\rangle$。类似的可以求出，$H|\downarrow\uparrow\rangle=2J|\uparrow\downarrow\rangle-J|\downarrow\uparrow\rangle$, $H|\uparrow\downarrow\rangle=2J|\downarrow\uparrow\rangle-J|\uparrow\downarrow\rangle$, 以及 $H|\downarrow\downarrow\rangle=J|\downarrow\downarrow\rangle$。因此就容易写出 H 在矢量基 $\{|\uparrow\uparrow\rangle,|\downarrow\uparrow\rangle,|\uparrow\downarrow\rangle,|\downarrow\downarrow\rangle\}$ 中的表示矩阵 $\widehat{H}$，

$$\widehat{H}=J\begin{pmatrix}1&0&0&0\\0&-1&2&0\\0&2&-1&0\\0&0&0&1\end{pmatrix}. \tag{4.13}$$

很容易求出这个矩阵的四个本征值，分别为 $J,J,J,-3J$, 分别对应下面的四个归一化本征矢量

$$\psi_{1,+1}=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\psi_{1,0}=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\\0\end{pmatrix},\psi_{1,-1}=\begin{pmatrix}0\\0\\0\\1\end{pmatrix},\psi_{0,0}=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\1\\0\end{pmatrix}. \tag{4.14}$$

分别以这四个本征矢量作为叠加系数，相应的就有原来的哈密顿算符 H 的四个本征态

$$|1,+1\rangle=|\uparrow\uparrow\rangle,|1,0\rangle=\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle+|\downarrow\uparrow\rangle),|1,-1\rangle=|\downarrow\downarrow\rangle, \tag{4.15}$$

以及

$$|0,0\rangle=\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle), \tag{4.16}$$

其中前面三个本征态是简并的，对应的本征值都为 J，这三个简并态称作两 $1/2$ 自旋耦合系统的自旋三重态，最后一个本征态 $|0,0\rangle$ 对应的本征能量是 $-3J$，它称作两自旋耦合系统的自旋单态。

人们很容易看到，自旋三重态的三个态有一个共同点，那就是它们对于两个电子来说是对称的，即如果把电子 1 和电子 2 进行交换，那么这三个态都将保持不变，而且下一节我们将看到它们的总自旋大小是两电子贡献相加的，是 1，因此我们常常称三重态为自旋平行态。相反，自旋单态关于两个电子是反对称的，如果将电子 1 和电子 2 进行交换，那相应的自旋单态就会出一个负号，而且它的总自旋两电子抵消为 0 了，因此我们常常称单态为自旋反平行态。

在量子力学中，一个系统的能量最低的本征态又称作系统的基态，基态的能量就是量子系统可能具有的最低能量，比基态能量更高的定态就称作激发态。基态 (有时候也包括低激发态) 对于研究一个系统的性质而言尤其重要，原因在于，对于一个宏观系统而言，只要环境的温度足够低，那么它就会处在基态附近，因此基态和低激发态往往决定了一个系统的宏观行为。

那么我们的两电子自旋耦合系统的基态是什么呢? 很显然，答案依赖于 J 是正还是负。如果 J 大于 0，那么由于自旋单态的能量为 $-3J$，为最低能量，因此系统的基态就将是这样一个自旋反平行态，这时我们称系统处在反铁磁相。想反，如果 $J<0$，那么能量为 J 的自旋三重态就将是系统的基态，注意，这时候系统的基态是简并的，$|1,+1\rangle, |1,0\rangle, |1,-1\rangle$ 张成了一个简并子空间，任何一个给定的系统当然不可能同时处在这三个态，所以它就必须从这三个简并基态所张成的简并子空间中随机选取一个态。另一方面，从我们原来的哈密顿算符 (4.2) 来看，三个空间方向地位完全平等，因此我们的系统当然是空间旋转不变的。但是，这个简并子空间中的任意一个态都有一个特定的总自旋方向，比方说，$|1,+1\rangle$ 态总自旋沿着 z 轴向上，$|1,-1\rangle$ 态总自旋反着 z 轴向下，因此不管系统从简并子空间中选择了哪个态，都意味着它选定了一个特定的空间方向，因此就破坏了空间旋转不变性，由于这种对空间旋转不变性的破坏完全是系统自发的，而不是因为我们在系统的哈密顿算符中引入了某个特定指向的磁场之类的东西，所以我们常常称这种旋转对称性的自发破坏为自发破缺 (注意我们不用破坏这个词，而是用自发破缺)！但是，不管系统破缺到简并子空间中的哪一个态，它都是一个自旋平行态 (比如，$|1,0\rangle$ 态的两电子也是自旋平行的，只不过平行的方向不是在 z 轴上而已)，这时候我们就称系统处在铁磁相。

从这个例子我们可以想见，更具一般性的海森堡模型 (4.1) 到底是铁磁的还是反铁磁的，这取决于 J 的符号，要从理论上解释某个材料的铁磁性，那我们就必须从理论上计算出一个小于 0 的 J。但这是一件相当困难的事情，很少有人能从理论上计算出某个材料的 J。因此人们通常采用的理论研究方案是，从某一个对系统的更基本描述出发，然后看看有没有一个机制能够让这个系统的有效描述可以是一个 J 为负的海森堡铁磁模型。如果理论上找到了一个这样的机制，那么人们就会设计实验来实现这样的机制，并看它是否能和理论推导一样使得系统处于铁磁相。总之，即使从海森堡提出他的模型到现在已经有很长时间了，关于铁磁性的研究依然一直是理论凝聚态的前沿领域之一，就是因为这种研究并不容易。

两自旋耦合系统的总自旋

让我们再次回到两自旋耦合的哈密顿算符 (4.2), 很显然，我们可以用泡利算符的矢量形式将它重写成 $H=J\vec{\sigma}_1\cdot\vec{\sigma}_2$。而且，我们可以引入一个总泡利算符 $\vec{\sigma}=\vec{\sigma}_1+\vec{\sigma}_2$, 它和系统总自旋 $\vec{S}$ 之间的关系是 $\vec{S}=\hbar\vec{\sigma}/2$，所以研究总泡利算符等价于研究系统总自旋。让我们先来考察一下 $(\vec{\sigma})^2=(\vec{\sigma}_1)^2+(\vec{\sigma}_2)^2+2\vec{\sigma}_1\cdot\vec{\sigma}_2$，而由于 $(\vec{\sigma}_1)^2=(\sigma_1^x)^2+(\sigma_1^y)^2+(\sigma_1^z)^2=3$, 同样 $(\vec{\sigma}_2)^2=3$, 所以我们可以知道 $\vec{\sigma}_1\cdot\vec{\sigma}_2=(\vec{\sigma})^2/2-3$，因此原来的哈密顿算符就可以重写成 $H=J[(\vec{\sigma})^2/2-3]$。

显然，算符 $(\vec{\sigma})^2$ 和哈密顿算符 H 的本征态是一样的，由前面求出来的 H 的本征值和本征态我们容易知道，自旋三重态 $|1,+1\rangle,|1,0\rangle,|1,-1\rangle$ 和自旋单态 $|0,0\rangle$ 分别是 $(\vec{\sigma})^2$ 的本征值为 8 和本征值为 0 的本征态。如果用总自旋 $(\vec{S})^2=(\vec{\sigma})^2\hbar^2/4$ 来考虑的话，那么自旋三重态就是 $(\vec{S})^2$ 的本征值为 $2\hbar^2$ 的本征态，通常我们说这种态的总自旋为 1，这就是三重态的记号中前一个 1 所代表的含义，不过，我们暂时不用管为什么把这样的量子态称之为自旋为 1，暂时我们只需要将它理解成一个约定，其含义就是，这种态对应于 $(\vec{S})^2=2\hbar^2$, 也就是说，这三个态的总自旋大小是一样的，都是 $\sqrt{2}\hbar$。当然，用 $(\vec{S})^2$ 来考虑，自旋单态就对应于 $(\vec{S})^2=0$，也即是说，自旋单态是总自旋为 0 的态，这就是 $|0,0\rangle$ 这个记号中前一个 0 的含义。

很明显，$(\vec{S})^2$ 的本征值不足以将两自旋耦合系统的四个本征态都区分出来，因为自旋三重态的那三个态对于 $(\vec{S})^2$ 来说是简并的。因此，人们通常进一步考察总自旋的 z 分量 $S_z=\sigma^z\hbar/2$。由于 $(\vec{\sigma})^2=6+2\vec{\sigma}_1\cdot\vec{\sigma}_2=6+2(\sigma_1^x\sigma_2^x+\sigma_1^y\sigma_2^y+\sigma_1^z\sigma_2^z)$, 我们很容易利用泡利算符的代数关系 (4.4) 验证 $[\sigma^z,(\vec{\sigma})^2]=0$, 这也即是说，$S_z$ 和 $(\vec{S})^2$ 对易，

$$[(\vec{S})^2,S_z]=0. \tag{4.17}$$

因此，S_z 和 $(\vec{S})^2$ 可以有共同的本征态，实际上，由于 $S_z=\frac{\hbar}{2}\sigma^z=\frac{\hbar}{2}(\sigma_1^z+\sigma_2^z)$，人们很容易验证：$S_z|1,+1\rangle=\hbar|1,+1\rangle$(即 $|1,+1\rangle$ 是 S_z 的本征值为 1 倍 $\hbar$ 的本征态，这就是 $|1,+1\rangle$ 记号中后一个 $+1$ 的含义)，$S_z|1,0\rangle=0|1,0\rangle$(这就是本征态记号中的那个 0 的含义)，$S_z|1,-1\rangle=-\hbar|1,-1\rangle$(即本征值为 -1 倍的 $\hbar$, 这就是记号 $|1,-1\rangle$ 中那个 -1 的含义)，类似的，$S_z|0,0\rangle=0|0,0\rangle$(这就是本征态记号中第二个 0 的含义)。简单归纳一下，即，自旋单态总自旋大小为 0，当然 S_z 也为 0，自旋三重态总自旋为 1，其 S_z 的值分别为 $+1,0,-1$(以 $\hbar$ 为单位)。

4.1.3 习题

1. 请求解 $H=J(\sigma_1^x\sigma_2^x+\sigma_1^y\sigma_2^y+\Delta\sigma_1^z\sigma_2^z)$ 这样一个 XXZ 模型并对结果进行讨论，式中 Δ 是一个实参数。

4.2 反射与透射

根据波动力学的观点，微观粒子也是波，而从波动光学的研究我们了解到波有一些非常典型的现象，比方说反射和透射。让一束波从真空向一块材料入射，我们可以预期有一

部分波会被材料反射，产生反射波，同时也可能有一部分波会进入材料内部，成为透射波。既然微观粒子也是波，那假如我们将一束粒子向一块材料入射，也将会有一些粒子被材料反射，而另一些粒子则进入材料内部。那么，对于单个入射粒子而言，它被材料反射的概率是多少呢？它透射的概率又是多少呢？这一节我们将通过一个最简单的数学模型来研究这些问题。

4.2.1 势散射的简单例子

为了简单起见，我们假设所考察的是垂直入射，因此只需要考察粒子在垂直于材料表面的方向上的运动，问题将简化为一个一维问题。进一步，我们用一个常数势垒 V_0 来模拟整块材料。如此一来，问题就变成粒子在如图 (4.1) 所示的一维势场 $V(x)$ 中运动的问题。图中的区域 I 就对应真空，区域 II 对应材料，势能函数 $V(x)$ 可以写成

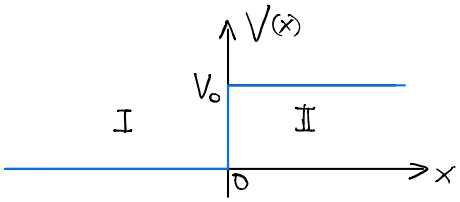

Figure 4.1: 一维势能台阶。

$$V(x)=\begin{cases}0, & x<0\\ V_0, & x>0\end{cases}. \tag{4.18}$$

而我们需要求解的就是一维定态薛定谔方程

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right]\psi(x)=E\psi(x), \tag{4.19}$$

式中的 E 也就是入射粒子能量。

由于势能函数 $V(x)$ 是一个分段函数。所以我们要分 I、II 两个区域分别求解方程 (4.19)，然后再将两个区域的解拼接起来。我们记 I、II 两个区域的解分别为 $\psi_I(x)$ 和 $\psi_{II}(x)$。稍微整理一下，人们就能看出，$\psi_I(x)$ 和 $\psi_{II}(x)$ 分别满足如下方程

$$\begin{aligned}&\frac{d^2}{dx^2}\psi_I(x)+\frac{2mE}{\hbar^2}\psi_I(x)=0.\\ &\frac{d^2}{dx^2}\psi_{II}(x)+\frac{2m(E-V_0)}{\hbar^2}\psi_{II}(x)=0.\end{aligned} \tag{4.20}$$

$E>V_0$ 情形

首先，假设 $E>V_0$。这时候可以令 $\frac{2mE}{\hbar^2}=k_1^2$、$\frac{2m(E-V_0)}{\hbar^2}=k_2^2$。从而方程 (4.20) 变成

$$\begin{aligned}&\frac{d^2}{dx^2}\psi_I(x)+k_1^2\psi_I(x)=0.\\ &\frac{d^2}{dx^2}\psi_{II}(x)+k_2^2\psi_{II}(x)=0.\end{aligned} \tag{4.21}$$

很显然，这两个方程的解可以写成

$$\psi_I(x) = A_1 e^{ik_1 x} + A_1' e^{-ik_1 x}.$$
$$\psi_{II}(x) = A_2 e^{ik_2 x} + A_2' e^{-ik_2 x}. \tag{4.22}$$

注意到 e^{ikx} 这样的波函数描述的是一个右行波，而 e^{-ikx} 这样的波函数描写的是左行波。如果粒子是从 $x=-\infty$ 处入射，那 I 区域的 $A_1 e^{ik_1 x}$ 描写的就是入射波，而 $A_1' e^{-ik_1 x}$ 描写的就是左行的反射波。类似的，II 区域的 $A_2 e^{ik_2 x}$ 描写的就是右行的透射波。但是，由于 II 区域上没有粒子入射，所以 II 区域上的左行波 $A_2' e^{-ik_2 x}$ 必定为零，即

$$A_2' = 0. \tag{4.23}$$

我们可以分别计算 I、II 两个区域上的概率流密度 J_I、J_{II}，并得到

$$J_I = |A_1|^2 \frac{\hbar k_1}{m} - |A_1'|^2 \frac{\hbar k_1}{m}$$
$$J_{II} = |A_2|^2 \frac{\hbar k_2}{m}. \tag{4.24}$$

局域概率守恒意味着，通过 I 区域的概率流必定要和通过 II 区域的概率流相等，从而必定有

$$\left(|A_1|^2 - |A_1'|^2\right) k_1 = |A_2|^2 k_2. \tag{4.25}$$

待会儿我们可以验证这个结果是否成立。

很显然，J_I 表达式中的 $|A_1|^2 \frac{\hbar k_1}{m}$ 代表的是入射粒子概率流密度，$-|A_1'|^2 \frac{\hbar k_1}{m}$ 代表的是反射粒子概率流密度。我们可以定义反射概率 R 为 $|A_1'|^2 \frac{\hbar k_1}{m}$ 比上 $|A_1|^2 \frac{\hbar k_1}{m}$，从而

$$R = \left|\frac{A_1'}{A_1}\right|^2. \tag{4.26}$$

从这个定义可以看出 R 其实就是反射粒子流强与入射粒子流强的比值，也就是单位时间之内反射粒子数在单位时间入射粒子总数中的比例。类似的，我们可以定义透射概率 T 为透射粒子概率流密度与入射粒子概率流密度的比值，从而

$$T = \frac{k_2}{k_1}\left|\frac{A_2}{A_1}\right|^2. \tag{4.27}$$

R 和 T 就是我们最终要计算的物理可观测量。很显然，概率守恒的要求 (4.25) 其实就等价于

$$R + T = 1. \tag{4.28}$$

待会儿我们可以验证这个结果。

下面，我们将两个不同区域的解 (4.22) 拼接起来。为此我们需要注意到，在任意空间点 x，波函数必然都连续。这是因为，波函数的模方反映的是粒子在空间的概率分布，它

没有理由不连续，而要在所有情况下都保证波函数模方连续，那只有波函数本身连续。特别的，在 $x=0$ 点，波函数连续，即 $\psi_I(0)=\psi_{II}(0)$，利用解 (4.22)，我们即有

$$A_1+A_1'=A_2. \tag{4.29}$$

但是，我们有三个未定常数 A_1,A_1',A_2，而我们要决定两个物理量 R,T。所以这一个拼接条件还不够，我们还得找一个拼接条件。

这另一个解的拼接条件就是波函数的一阶导数要连续。在我们这一例子中，这其实可以证明。为此，我们任取一点 x，然后在 x 附近对定态薛定谔方程 (4.19) 的表达式进行积分，从 $x^-=x-\varepsilon$ 积到 $x^+=x+\varepsilon$，ε 为一个正的无穷小量。从而就可以得到

$$-\frac{\hbar^2}{2m}[\psi'(x^+)-\psi'(x^-)]=\int_{x^-}^{x^+}dx(E-V(x))\psi(x), \tag{4.30}$$

式中 $\psi'(x)=\partial_x\psi(x)$。注意到，波函数 $\psi(x)$ 连续，因此只要 $V(x)$ 不太奇异，那么 $(E-V(x))\psi(x)$ 也必定连续。特别的，对于我们这里涉及到的阶跃势能，$(E-V(x))\psi(x)$ 虽然不连续，但是 $\int_{x^-}^{x^+}dx(E-V(x))\psi(x)\to 0$。从而由 (4.30) 可知，在 $\varepsilon\to 0$ 时必有

$$\psi'(x^+)=\psi'(x^-). \tag{4.31}$$

这就意味着波函数的一阶导数在任意 x 都连续。

根据 $\psi_I'(0)=\psi_{II}'(0)$，由分段解 (4.22)，我们可以得到

$$k_1(A_1-A_1')=k_2A_2. \tag{4.32}$$

结合拼接条件 (4.29) 和拼接条件 (4.32)，我们就可以得到

$$\frac{A_1'}{A_1}=\frac{k_1-k_2}{k_1+k_2},\quad \frac{A_2}{A_1}=\frac{2k_1}{k_1+k_2}. \tag{4.33}$$

代入 (4.26) 式和 (4.27) 式，我们就可以得到

$$R=1-\frac{4k_1k_2}{(k_1+k_2)^2},\quad T=\frac{4k_1k_2}{(k_1+k_2)^2}. \tag{4.34}$$

显然满足 $R+T=1$，因此的确概率守恒。

$E<V_0$ 情形

假设入射粒子能量不够，$E<V_0$。由于入射粒子能量 E 总大于 0，所以这时候 I 区域内的解形式上还和之前的情形一样。但是对于 II 区域，由于这时候 $\frac{2m(E-V_0)}{\hbar^2}<0$，我们不妨令 $\frac{2m(E-V_0)}{\hbar^2}=-\kappa_2^2$。因此这时候 II 区域的解就将变成

$$\psi_{II}(x)=B_2e^{-\kappa_2x}+B_2'e^{\kappa_2x}. \tag{4.35}$$

但是，由于 $x\to+\infty$ 时波函数得保持有限，从而这个解的指数增长部分必定要为零，即

$$B_2'=0. \tag{4.36}$$

这时候我们很容易算得 II 区域的概率流密度 $J_{II}=0$。从而由局域概率守恒，也有 $J_I=J_{II}=0$。根据 (4.24) 式，我们必有 $|A_1'|^2=|A_1|^2$，或者说

$$R=1. \tag{4.37}$$

也即是说，这时候粒子必定全反射，稍后我们会验证这个结果。

对 $E<V_0$ 这种情形的求解其实和 $E>V_0$ 情形一样。实际上，我们只需要将 $E>V_0$ 情形的 k_2 替换成 $i\kappa_2$ 就会过渡到 $E<V_0$ 情形。所以现在的解其实很简单，为

$$\frac{A_1'}{A_1}=\frac{k_1-i\kappa_2}{k_1+i\kappa_2},\quad \frac{B_2}{A_1}=\frac{2k_1}{k_1+i\kappa_2}. \tag{4.38}$$

很显然，我们的确有 $R=\left|\frac{A_1'}{A_1}\right|^2=1$。而且，注意到 $\frac{A_1'}{A_1}$ 是一个相位因子，我们不妨将之记为 $\frac{A_1'}{A_1}=e^{i\delta}$，

$$e^{i\delta}=\frac{k_1-i\kappa_2}{k_1+i\kappa_2}. \tag{4.39}$$

因此这时候 I 区域的波函数可以写成

$$\psi_I(x)=A_1(e^{ik_1x}+e^{i\delta}e^{-ik_1x}). \tag{4.40}$$

可见，反射波相对于入射波发生了 δ 的相位移动。特别的，如果势垒趋于无穷高，即 $V_0\to+\infty$, 从而 $\kappa_2\to+\infty$, 那由 (4.39) 式可知，这时候的相移为

$$\delta=\pm\pi. \tag{4.41}$$

这就是半波损失。

4.2.2 散射态与束缚态

在上一小节的例子中，对于任何大于零的能量 E，我们都可以找到相应定态薛定谔方程的物理解。并没有出现能量量子化，系统的能谱 E 依然连续可变。并且我们还看到，无论是 $E>V_0$ 还是 $E<V_0$，相应的定态波函数 $\psi(x)$ 在 $x\to-\infty$ 时都不趋于零，在物理上，它代表粒子可以从 $x=-\infty$ 处入射，也可以反射到 $x=-\infty$ 处去，总之，粒子可以出现在 $x=-\infty$ 处。像这样的，能谱连续，粒子可以出现在空间无穷远处的定态，我们通常称之为散射态。我们可以将散射态的能量解释成粒子在无穷远处的能量，在物理上，无穷远处的粒子可以认为远离了一切相互作用势，从而是自由粒子，从而散射态的能量就等于无穷远处的这个自由粒子的能量，它当然是一个大于 0 的连续谱。

之所以称这样的态为散射态，是因为粒子在无穷远处是自由的，而入射到势场区以后就会被势场散射。一般来说，散射态正是描述了粒子从无穷远处入射，最后又被散射到无穷远处去的这样一个量子态。就好比上一小节的例子中，粒子从 $-\infty$ 处入射最终被 $V(x)$ 散射到 $-\infty$(对应反射波) 或者 $+\infty$ 处 (对应透射波)。

与散射态相对的，定态薛定谔方程还可能出现另一类不同的解，称之为束缚态。它描述的是粒子被束缚在一个有限的区域内，无法到达空间无穷远处的情形。因此，与散射态

不同，束缚态要求波函数 $\psi(\mathbf{x})$ 在空间无穷远处以足够快的速度趋于零。如果无穷远处势能为 0 的话，那通常来说束缚态的能量本征值必然要小于 0，因为否则的话粒子就可能有一定的概率隧穿到无穷远处去。由于束缚态的粒子总是被束缚在一个区域之内，因此它就会在这个区域内形成某种“驻波”，但并非任意 $E<0$ 都能满足“驻波”条件，因此束缚态的能量通常是量子化的，对应分立的能量本征值。本章后面几节将研究一些束缚态的具体例子。下面的图 (4.2) 示意性地画出了束缚在一个势阱中的粒子的三个束缚态能级，E_0, E_1, E_2，其中，E_0 是基态。

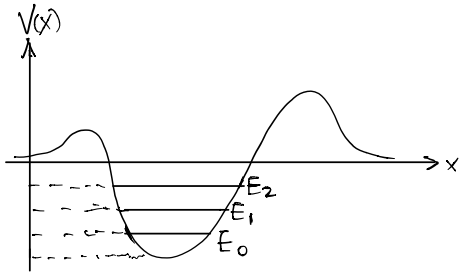

Figure 4.2: 束缚态能级

4.2.3 习题

1. 粒子在一维势场 $V(x)=-A\delta(x)$ 中运动，假定 $A>0$。(1). 请问粒子波函数的一阶导数是否连续，为什么？(2). 请问这个系统有没有束缚态，如果有，请求出束缚态能级。

2. 某一维系统的势能函数 $V(x)$ 如下，

$$V(x)=\begin{cases}0, & x<0,\quad x>a\\ V_0, & 0<x<a\end{cases}. \tag{4.42}$$

其中 $V_0>0$ 为势垒的高度。假设粒子从势垒左边 $x\to-\infty$ 的地方入射，能量 $0<E<V_0$。请求出粒子隧穿过势垒的概率，也即粒子的透射概率。并验证在 $\hbar$ 很小的半经典近似下，结果和第三章中用 WKB 近似得到的结果吻合。

4.3 单自由度线性谐振子

线性谐振子的研究在量子物理中的重要性怎么强调都不过分。首先，最早普朗克在推导他的黑体辐射公式的时候，就假定辐射电磁波的是谐振子 (这是由于黑体辐射结果的普适性，它并不依赖于具体是什么物质在辐射电磁波)，为了推导出黑体辐射公式，普朗克假设这些谐振子的能量是量子化的，只能取 $n\hbar\omega$ 这样的量子化值。随后，爱因斯坦的光量子理论紧接着说，其实电磁场本身就是量子化的，电磁场的量子叫做光子，一个角频率为 ω 的光子其能量是 $\hbar\omega$，n 个相同的光子能量就是 $n\hbar\omega$。从我们今天的观点来看，爱因斯坦的电磁场量子化假设和普朗克的谐振子量子化假设其实密切相关，原因在于，我们可

以把任何电磁波分解成平面波 (或者空腔里的驻波) 的线性叠加，任何电磁场都由这种平面波叠加的叠加系数唯一性地决定，代入麦克斯韦方程人们就能发现，这些叠加系数所满足的方程就是标准的线性谐振子方程，所以，电磁场可以看成是一个多自由度的谐振子系统，每一个波模式对应一个谐振自由度。但是，无论普朗克也好还是爱因斯坦也好，他们都没有从理论上解释为什么线性谐振子的能量会量子化为 $n\hbar\omega$ 这样的形式，首先用量子力学理论推导出这个结论的是海森堡。用量子力学推导出这个结论就是我们这一节的主要内容，我们将要给出来的推导方法是相当标准的，而且类似的方法在量子理论中应用非常广泛。

爱因斯坦不仅提出了电磁场的光量子，他还将类似的观念应用于固体里面的声波，因为声波的每一个模式同样是一种简谐振动，声波的量子就叫声子，正是因为考虑了声波的量子化，爱因斯坦才能成功地解释固体在低温下的比热容，而这是此前的经典理论无法解释的。类似的，在铁磁体和反铁磁体这些自旋相互作用的体系中还存在自旋波，自旋波的量子通常叫做磁振子 (Magnon)，等等。总之，谐振子的概念无论是在量子场论中还是在凝聚态物理中都是基础性的。

不仅如此，任何力学系统在其稳定平衡位置附近的运动都是一种简谐振动，而组成一个分子的若干原子在通常的温度下就是在其平衡位置附近来回振动，因此谐振子的研究同样可以应用于诸如分子振动光谱之类的原子分子物理问题。关于线性谐振子的更多讨论，我们推荐读者参考 Cohen-Tannoudji 的《量子力学》第一卷。

4.3.1 单自由度线性谐振子的量纲分析

所谓的单自由度线性谐振子，我们指的是问题中只涉及一个力学变量 X(以及和它相应的正则动量 P)，并且哈密顿量为 $H=\frac{P^2}{2m}+\frac{1}{2}m\omega^2X^2$ 的系统，这里 ω 就是这个线性谐振子的角频率，$\frac{1}{2}m\omega^2X^2$ 当然就是线性谐振子的弹性势能。从这个哈密顿量我们可以看到，单自由度线性谐振子只涉及 m,ω 这两个物理常量，当然由于我们将要进行的是一个量子力学分析，所以我们还需要加上普朗克常量 $\hbar$。因此整个问题只涉及 $m,\omega,\hbar$ 这三个有量纲的物理常量，也即是说，所有我们将要分析的物理量，能级也好，X 的方差也好，最终都应该是仅涉及这三个物理常量的某个表达式。在具体求解问题之前，我们不妨先对这样的表达式进行一下量纲分析。首先我们注意到，$m,\omega,\hbar$ 这三个量无法构造出任何无量纲的量 (因为它们实际上对应于三个相互独立的基本量纲，比方说 $1/\omega$ 有时间量纲，m 有质量量纲)，这就告诉我们用量纲分析的方法得出来的表达式将是唯一的 (除了未定的数值系数以外)，因为没有任何未定的无量纲因子可以再乘上去。

很显然，$m,\omega,\hbar$ 能够构造出来的有能量量纲的表达式是 $\hbar\omega$，也就是说，线性谐振子的能级将有 $E\sim\hbar\omega$ 的形式。其次我们注意到 $[\hbar\omega]=[P^2/m]$(符号 [] 表示取一个表达式的量纲) 具有能量量纲，因此 $[P^2]=[m\hbar\omega]$，这也即是说，$\sqrt{m\hbar\omega}$ 有动量量纲。类似的，$[\hbar\omega]=[m\omega^2X^2]$ 也具有能量量纲，因此 $[X^2]=[\hbar/(m\omega)]$，这也即是说，$\sqrt{\hbar/(m\omega)}$ 具有长度量纲。有了这些量纲表达式以后，我们就可以将所有的力学变量都无量纲化，完成这一步的最简单办法是直接令 $\hbar=1,m=1,\omega=1$，这样我们的哈密顿算符就变成，

$$H=\frac{1}{2}(P^2+X^2), \tag{4.43}$$

现在 X,P,H 都已经无量纲了，因此我们求解这个无量纲问题 (4.43) 最终得出来的就是一些数值。而恢复量纲的办法也很简单，比如，对于 (4.43) 的本征值你直接乘上量纲表达式 $\hbar\omega$ 就得到原来谐振子的能级了，再比如对于 (4.43) 的本征波函数 $\psi(x)$ 里面的 x，你直接替换成 $\psi(\frac{x}{\sqrt{\hbar/(m\omega)}})$，那替换以后的 x 就不再是无量纲，而是有通常的长度量纲了。事先进行量纲分析的好处在于，作了无量纲化处理以后，我们就只需要求解一个数学上更简洁的无量纲哈密顿量 (4.43)，最后再将得到的结果恢复量纲就可以了。

4.3.2 求解与讨论

下面我们来具体求解算符 (4.43) 的本征值和本征态。我们用到的关键算符代数关系就是 $[X,P]=i\hbar$ 这样一个基本对易关系，当然由于作了无量纲化处理 $\hbar=1$，所以现在这个代数关系应该写成

$$[X,P]=i. \tag{4.44}$$

观察一下我们要求解的哈密顿量 $H=\frac{1}{2}(X^2+P^2)$，你会发现，如果我们处理的不是算符，而是普通的数，那很显然我们可以引入复变量，将这个二次型的哈密顿量写成某个复数的模方的形式，即 $H=\frac{1}{2}(X-iP)(X+iP)$。这就启发我们引入两个新的算符 a 和 $a^\dagger$，它们的定义是

$$a=\frac{1}{\sqrt{2}}(X+iP), a^\dagger=\frac{1}{\sqrt{2}}(X-iP), \tag{4.45}$$

注意这两个算符都不是厄米算符，而是互为厄米共轭的。利用代数关系 (4.44)，我们很容易得到 $[a,a^\dagger]=\frac{1}{2}[X+iP,X-iP]=\frac{1}{2}([X,-iP]+[iP,X])=1$，即

$$[a,a^\dagger]=1. \tag{4.46}$$

另外，由于 $a^\dagger a=\frac{1}{2}(X-iP)(X+iP)=\frac{1}{2}(X^2+P^2+iXP-iPX)=\frac{1}{2}(X^2+P^2)-\frac{1}{2}$，所以我们知道原来的哈密顿算符 (4.43) 可以写成

$$H=a^\dagger a+\frac{1}{2}, \tag{4.47}$$

式中多出来的这个 $1/2$ 就是因为我们处理的是相互不对易的算符，而不是普通的数。算符 a 和 $a^\dagger$ 通常分别称作湮灭算符和产生算符，它们的典型特征就是满足代数关系 (4.46)。

不妨记 $a^\dagger a=N$，很显然，算符 N 和哈密顿算符 H 有完全一样的本征态，而且 N 的本征值加上 $1/2$ 就是 H 的本征值。我们很容易论证 N 的本征值一定大于或者等于 0。这是因为，任给一个量子态 $|\psi\rangle$，假设我们记 $|\phi\rangle=a|\psi\rangle$，则有 $\langle\phi|=\langle\psi|a^\dagger$，因此 $\langle\phi|\phi\rangle=\langle\psi|a^\dagger a|\psi\rangle=\langle\psi|N|\psi\rangle$，而很显然 $\langle\phi|\phi\rangle\geq 0$，之所以这里允许 $\langle\phi|\phi\rangle$ 等于 0，是因为可能出现 $a|\psi\rangle=0$ 的情况，这时候 $|\phi\rangle$ 本身就是 0。这就告诉我们，对于任意量子态 $|\psi\rangle$，我们总有 $\langle\psi|N|\psi\rangle\geq 0$，因此 N 的本征值总是大于等于 0 的，并且等于 0 的情况仅在 $a|\psi\rangle=0$ 时才会出现。

为了下面的推导，我们介绍一个算符恒等式，假设有三个算符 A,B,C，那么我们将有下面的恒等式

$$[A,BC]=[A,B]C+B[A,C], \tag{4.48}$$

把两边按照对易子的定义展开，人们就能很容易地验证这个恒等式。在计算算符对易子的时候这样的恒等式非常有用。

利用上面的恒等式和 (4.46), 人们很容易算出 $[a,N]=a,[a^\dagger,N]=-a^\dagger$, 也常常写作

$$[N,a]=-a,[N,a^\dagger]=a^\dagger. \tag{4.49}$$

假设我们有厄米算符 N 的某个本征值为 λ 的本征态 $|\lambda\rangle$, 即

$$N|\lambda\rangle=\lambda|\lambda\rangle. \tag{4.50}$$

则由于 (4.49) 式，我们可以知道，$N(a|\lambda\rangle)=(Na)|\lambda\rangle=(aN-a)|\lambda\rangle=(\lambda-1)(a|\lambda\rangle)$(注意由 $[N,a]=-a\Rightarrow Na=aN-a$), 即，如果 $a|\lambda\rangle\neq 0$, 那它必定是 N 的本征值为 $\lambda-1$ 的本征态，即有

$$a|\lambda\rangle=\alpha_\lambda|\lambda-1\rangle, \tag{4.51}$$

式中 α_λ 为某个复系数。也即是说，a 作用在 N 的本征态上，可以得到一个本征值降低 1 的新本征态。为了看清楚什么情况下 $a|\lambda\rangle$ 等于 0，我们不妨记 $a|\lambda\rangle=|\varphi\rangle$，则有 $\langle\varphi|\varphi\rangle=\langle\lambda|a^\dagger a|\lambda\rangle=\langle\lambda|N|\lambda\rangle=\lambda\langle\lambda|\lambda\rangle$，很显然，只要 $\lambda\neq 0$, 那么 $a|\lambda\rangle$ 就不等于 0，反过来 $\lambda=0$ 就一定意味着 $a|\lambda\rangle=0$。

上面的分析有一个很重要的推论，即 N 的任何本征值 λ 都必须是非负整数。因为否则的话我们就可以不断重复地用 a 作用在相应的本征态上，得到一串新的本征态，它们的本征值分别为 $\lambda,\lambda-1,\lambda-2,\lambda-3,....$。如果 λ 不是非负整数，那这一串本征值都将不为 0，那么由上一段的分析就有，a 作用在这一串本征态的任何一个上面都将不是 0。因此用 a 进行作用的这个过程就可以无限地重复下去，这样一来，给定任何一个初始的 λ，我们总能作用任意 m 个 a 在其本征态上，得到一个本征值为 $\lambda-m$ 的本征态，只要 m 足够大，$\lambda-m$ 就必定为负数。但这是不可能的，因为前面我们已经证明过了，N 的本征值必须大于等于 0。因此，λ 必须为非负整数，我们记为 $\lambda=n$。

类似的，利用 (4.49) 式我们还可以得到，$a^\dagger$ 作用在 N 的某个本征态 $|n\rangle$ 上将会得到一个本征值为 $n+1$ 的本征态。归纳一下即有

$$N|n\rangle=n|n\rangle,\ a|n\rangle=\alpha_n|n-1\rangle,\ a^\dagger|n\rangle=\beta_{n+1}|n+1\rangle, \tag{4.52}$$

这里 $n\geq 0$, α_n,β_{n+1} 为待定复系数。同时，由于 $n|n\rangle=a^\dagger a|n\rangle=\alpha_n a^\dagger|n-1\rangle=\alpha_n\beta_n|n\rangle$，所以必有

$$\alpha_n\beta_n=n. \tag{4.53}$$

我们已经得到了第一个重要结论，即哈密顿算符 (4.47) 的本征谱为 $n+1/2$，如果恢复量纲，那么我们就有，单自由度线性谐振子的本征谱为

$$(n+\frac{1}{2})\hbar\omega, \tag{4.54}$$

可见，除了多出来一个 $(1/2)\hbar\omega$ 之外，这个结果正符合普朗克最初的假设。

多出来的这个 $(1/2)\hbar\omega$ 其实有重要的效应: 首先, 它意味着即使对于 $n=0$ 的基态, 线性谐振子的能量也不为 0，所以这个能量通常叫做零点能。之所以有零点能，是因为量子力学的不确定关系 $\delta x\delta p\sim\hbar$, 所以即使在基态，谐振子的坐标和动量也不能同时确定为 0，而是有涨落，零点能就是这个涨落的能量。其次，按照当前的观点，一切物质最终都是由基本粒子组成的，而所有的基本粒子都可以用场来描述，在忽略相互作用的时候，场可以看成是许许多多的谐振子，因此，即使所有的场都处在基态，它们也还是有零点能。基本粒子场活动的舞台当然就是我们的时空，所有基本粒子都处在基态所对应的就是真空态，因此场的零点能的存在就告诉我们，真空其实并不真正空，而是总有基本粒子场的涨落，这些涨落表现出来的零点能就叫作真空能。所以，真空其实充满能量，并且人们已经测到了真空的能量，这就是所谓的卡西米尔效应。另一方面，宇宙的大部分区域当然是真空，因此按照道理来说，宇宙中应该充满了真空能，它的一个效应就是会使得宇宙加速膨胀，今天的天文学观测的确发现我们的宇宙在加速膨胀，但是，宇宙学观测计算出来的使得宇宙加速膨胀的这个能量虽然占了宇宙总能量的大部分，但却还是比量子场论计算出来的真空能小太多太多，这到底是怎么回事现在还没有人搞清楚，这个问题就是著名的暗能量问题，它问的就是，第一，使得宇宙加速膨胀的暗能量的来源到底是什么，第二，为什么量子场论计算出来的宇宙真空能不对。

下面我们接着来研究线性谐振子的本征态 $|n\rangle$, 我们还是采用无量纲化的处理，只在最后才进行量纲恢复。假定所有的这些本征态 $|n\rangle$ 都已经正交归一了，因此我们首先要做的就是确定 (4.52) 式中的待定系数 α_n,β_n。为此，我们注意到由 $a^\dagger|n\rangle=\beta_{n+1}|n+1\rangle$ 可以得到，$|\beta_{n+1}|^2\langle n+1|n+1\rangle=\langle n|aa^\dagger|n\rangle=\langle n|a^\dagger a+1|n\rangle=(n+1)\langle n|n\rangle$(注意由 $[a,a^\dagger]=1\Rightarrow aa^\dagger=a^\dagger a+1$), 进一步利用本征态的归一性就有 $|\beta_{n+1}|^2=(n+1)$。同时, 在假定 $|n\rangle$ 态的相位已经设定的前提下, 我们总是可以合适地调节 $|n+1\rangle$ 态的相位，使得由方程 $a^\dagger|n\rangle=\beta_{n+1}|n+1\rangle$ 定义出来的 β_{n+1} 变成一个正实数。如此一来，我们就有 $\beta_{n+1}=\sqrt{n+1}$, 代入 (4.53) 式就可以进一步得到 $\alpha_n=\sqrt{n}$。因此我们有

$$a|n\rangle=\sqrt{n}|n-1\rangle,\ a^\dagger|n\rangle=\sqrt{n+1}|n+1\rangle. \tag{4.55}$$

特别的，对于 $n=0$ 的基态，我们有

$$a|0\rangle=0. \tag{4.56}$$

另外，将 (4.55) 式进行递推，我们可以得到

$$|n\rangle=\frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle. \tag{4.57}$$

到此为止实际上我们已经完成了所有的求解过程，但有时候人们还是希望得到坐标表象下的本征波函数。不妨让我们先来求坐标表象下的基态波函数，首先由 (4.56) 式，我们有 $\langle x|a|0\rangle=0$, 进一步由产生湮灭算符的定义式 (4.45) 可知 $\langle x|X+iP|0\rangle=0$。最后，将坐标表象下的基态波函数记为 $\psi_0(x)=\langle x|0\rangle$，并利用动量算符在坐标表象下的表示 $\widehat{P}=-i\partial_x$(已经作了 $\hbar=1$ 的无量纲化)，我们就有 $0=\langle x|X+iP|0\rangle=(x+\partial_x)\langle x|0\rangle$, 即

$$(\frac{\partial}{\partial x}+x)\psi_0(x)=0, \tag{4.58}$$

求解这个微分方程并将结果进行归一化就可以得到

$$\psi_0(x) = \left(\frac{1}{\pi^{\frac{1}{2}}}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2}x^2\right). \tag{4.59}$$

为了求激发态波函数，我们记 $\psi_n(x) = \langle x|n\rangle$。由 (4.57) 式我们可以知道 $\psi_n(x) = \frac{1}{\sqrt{n!}}\langle x|(a^\dagger)^n|0\rangle = \frac{1}{\sqrt{2^n n!}}\langle x|(X-iP)^n|0\rangle = \frac{1}{\sqrt{2^n n!}}(x-\partial_x)^n\langle x|0\rangle$，即

$$\psi_n(x) = \left(\frac{1}{\pi^{\frac{1}{2}}2^n n!}\right)^{\frac{1}{2}} (x-\partial_x)^n \exp\left(-\frac{1}{2}x^2\right). \tag{4.60}$$

如果人们将这个式子中的 $(x-\partial_x)^n$ 的作用结果算出来，就会发现它对应于一个 n 次多项式 $H_n(x)$, 通常称作厄米多项式。最后，让我们恢复量纲，从而就有

$$\psi_n(x) = \left(\frac{\kappa}{\pi^{\frac{1}{2}}2^n n!}\right)^{\frac{1}{2}} H_n(\kappa x)\exp\left(-\frac{1}{2}\kappa^2 x^2\right), \tag{4.61}$$

式中 $\kappa = 1/\sqrt{\hbar/(m\omega)} = \sqrt{\frac{m\omega}{\hbar}}$ 是长度的负一次方量纲。值得注意的是，由波函数归一化 $\int dx|\psi(x)|^2 = 1$ 可以知道，波函数本身也是有量纲的，在一维是长度的 $-1/2$ 次方量纲，这就是为什么在最后恢复量纲的时候 (4.61) 式的前面会多出一个 $\kappa^{\frac{1}{2}}$ 因子的原因。下面的图 (4.3) 示意性地画出了一些线性谐振子的定态波函数，以及相应的概率分布。从这幅图中可

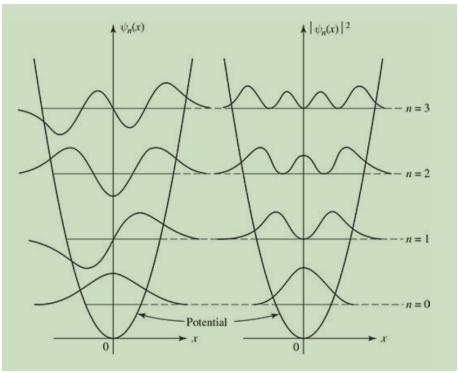

Figure 4.3: 线性谐振子波函数以及概率分布

以清楚地看到，线性谐振子的定态 $\psi_n(x)$ 都是一些束缚态。实际上，由于线性谐振子的势能在无穷远处趋于无穷大，因此任何情况下粒子都不可能出现在无穷远处，所以线性谐振子系统没有散射态。

4.3.3 位力定理

利用上一小节的求解结果，让我们来作一些简单的计算，我们的计算都将采用先在无量纲化下处理，最后再恢复量纲的办法。

先让我们计算在定态 $|n\rangle$ 上，谐振子坐标的平均值 $\langle n|X|n\rangle$。利用产生湮灭算符的定义式我们显然有 $\langle n|X|n\rangle = \frac{1}{\sqrt{2}}\langle n|a+a^\dagger|n\rangle$, 进一步利用 (4.55) 式以及不同本征态的正交性很容易得到最终结果为 0。类似的，我们也可以得到动量的平均值也为 0，即有

$$\langle n|X|n\rangle = \langle n|P|n\rangle = 0. \tag{4.62}$$

下面我们来计算 X^2 的平均值。由 $\langle n|X^2|n\rangle = \frac{1}{2}\langle n|(a+a^\dagger)^2|n\rangle = \frac{1}{2}\langle n|a^2+(a^\dagger)^2+aa^\dagger+a^\dagger a|n\rangle = \frac{1}{2}\langle n|2a^\dagger a+1|n\rangle = (n+\frac{1}{2})$, 式中我们利用了 (4.55) 式以及不同本征态的正交性。类似的我们也可以计算出 $\langle n|P^2|n\rangle = (n+\frac{1}{2})$。由此我们就可以得到一个重要的结论，即在定态上，谐振子动能 $\frac{1}{2}P^2$ 的平均值和势能 $\frac{1}{2}X^2$ 的平均值相等，都等于相应能级的一半。恢复量纲即有

$$\langle n|\frac{P^2}{2m}|n\rangle = \langle n|\frac{1}{2}m\omega^2X^2|n\rangle = \frac{1}{2}(n+\frac{1}{2})\hbar\omega. \tag{4.63}$$

上面的这个结论其实是所谓的位力定理的一个特殊情况，下面我们就以一维问题为例来推导一下位力定理。假设我们考虑的系统哈密顿算符为 $H = \frac{P^2}{2m}+V(X)$, $|n\rangle$ 为它的第 n 个定态，相应的能级为 E_n，即

$$H|n\rangle = E_n|n\rangle. \tag{4.64}$$

那么我们容易有 $\langle n|[PX,H]|n\rangle = \langle n|PXH-HPX|n\rangle = E_n\langle n|PX|n\rangle - E_n\langle n|PX|n\rangle = 0$，式中最后一个等号我们利用了 $\langle n|H = E_n\langle n|$。然而在另一方面，$[PX,H] = [PX,\frac{P^2}{2m}]+[PX,V(X)] = P[X,\frac{P^2}{2m}]+[P,V(X)]X = i\hbar\left(\frac{P^2}{m}-X\frac{\partial}{\partial X}V(X)\right)$, 式中我们利用了公式 (4.48) 以及基本对易关系 $[X,P] = i\hbar$。将这两个计算结合起来，即有

$$2\langle n|T|n\rangle = \langle n|\frac{P^2}{m}|n\rangle = \langle n|X\frac{\partial}{\partial X}V(X)|n\rangle. \tag{4.65}$$

这就是位力定理，式中的 T 就表示动能算符。特别的，对于线性谐振子，$V(X) = \frac{1}{2}m\omega^2X^2$, 因此 $X\frac{\partial}{\partial X}V(X) = 2V(X)$, 如此一来即有

$$\langle n|T|n\rangle = \langle n|V|n\rangle. \tag{4.66}$$

这正是我们前面通过具体计算发现的结论。

4.3.4 习题

1. 对于线性谐振子，人们可以定义一种特殊的量子态，叫做相干态。任给复数 z，没有归一化的相干态 $|z)$(我们用右括号而不是 $\rangle$ 来强调这个量子态没有归一化) 可以定义为 $|z) = e^{za^\dagger}|0\rangle$。

(1). 请证明 $|z) = \sum_{n=0}^{+\infty}\frac{z^n}{\sqrt{n!}}|n\rangle$。

(2). 请证明 $|z)$ 是湮灭算符 a 的本征态，满足 $a|z) = z|z)$。请注意，由于 a 不是一个厄米算符，因此其本征值可以是复数 z。

(3). 请将 $|z)$ 归一化，并计算相干态 $|z)$ 与相干态 $|w)$ 的内积为 $(w|z)$。

(4). 假设记归一化以后的相干态为 $|z\rangle$。请证明 $|z\rangle$ 可以取为 $|z\rangle = D(z)|0\rangle$, 式中 $D(z) = e^{za^\dagger - z^*a}$ 为幺正算符。提示：可以应用第二章习题中证明的算符恒等式 $e^{A+B} = e^Ae^Be^{-\frac{1}{2}[A,B]}$。

(5). 假设记复数 $z = u+iv$，记 $d^2z = dudv$。请证明 $\frac{1}{\pi}\int d^2z e^{-|z|^2}|z)(z| = 1$，这里的积分是对整个复平面积分，最后的结果 1 表示的是恒等算符。

(6). 请计算相干态 $|z\rangle$ 上，位置算符 X 以及动量算符 P 的不确定度 δX 和 δP，并证明 $\delta X\delta P$ 取最小值，即 $\delta X\delta P = \hbar/2$。

(7). 请求出相干态 $|z\rangle$ 在坐标空间的波函数 $\psi_z(x)=\langle x|z\rangle$, 从而证明相干态在坐标空间的概率分布 $|\psi_z(x)|^2$ 是高斯分布。

(8). 假设 $t=0$ 时刻，线性谐振子处在相干态 $|z\rangle$，请证明 t 时刻它将处在相干态 $e^{-iwt/2}|ze^{-i\omega t}\rangle$。

(9). 请研究处在相干态上的谐振子其在坐标空间的概率分布如何随时间演化，即研究概率分布 $|\psi_{ze^{-iwt}}(x)|^2$。请将结果与谐振子的经典振动行为进行比较，进而领会为什么相干态是一种准经典态。

2. 超导 LC 振荡电路和普通的经典 LC 振荡电路不同，它是量子的，也即是说超导 LC 电路电容器上的电荷 Q，以及电路中的电流都是量子算符。超导 LC 振荡电路的哈密顿量可以写成

$$H=\frac{Q^2}{2C}+\frac{1}{2}LI^2. \tag{4.67}$$

但是可以论证，电流算符 I 和电荷算符 Q 必定不对易。

论证方式有两种：第一种方式是通过利用 $I=-\dot{Q}$, 进而将哈密顿量写成 $H=\frac{1}{2}L\dot{Q}^2+\frac{Q^2}{2C}$, 很显然这样写以后，我们可以将 Q 类比成粒子坐标，$\dot{Q}$ 类比成速度，L 类比成质量，从而磁场能量 $\frac{1}{2}L\dot{Q}^2$ 就可以类比成动能，而电场能量 $\frac{Q^2}{2C}$ 就可以类比成势能。当然，在这样的类比之下 $L\dot{Q}=-LI=-\Phi$ 就应该类比成粒子动量 (这里 $\Phi=LI$ 表示电感线圈的磁通，它同样是算符)。那么根据位置坐标和动量之间的基本对易关系我们就应该有 $[Q,-\Phi]=i\hbar$。

第二种论证方式是，利用 $-\dot{\Phi}$ 是电感线圈两端的电压 (电磁感应定律)，从而 $\dot{\Phi}$ 等于电容器两端的电压，从而 $C\dot{\Phi}=Q$, 进而将电路的哈密顿量重写为 $H=\frac{1}{2}C\dot{\Phi}^2+\frac{1}{2}\frac{\Phi^2}{L}$。很显然，这样重写哈密顿量以后就可以将 Φ 类比成粒子坐标，$\dot{\Phi}$ 类比成速度，从而电场能量 $\frac{1}{2}C\dot{\Phi}^2$ 就类比成动能，磁场能量 $\frac{1}{2}\frac{\Phi^2}{L}$ 就应该类比成势能。当然，这时候 $C\dot{\Phi}=Q$ 就应该类比成粒子动量，因此同样应该有

$$[\Phi,Q]=i\hbar. \tag{4.68}$$

很明显，在这两种不同的论证方式中，电场能量和磁场能量的角色刚好互换了，但是，结果是一样的，即都有

$$[I,Q]=i\hbar/L. \tag{4.69}$$

这种结果的一致性反映了电和磁的相互对偶。我们的问题是，请根据 Q,I 之间的对易关系求出超导 LC 振荡电路的量子化能级。

4.4 氢原子

玻尔的氢原子模型成功地解释了氢原子的光谱。但是玻尔的模型还只是一个半经典的理论，因此从物理理论的逻辑来说依然需要发展一个关于氢原子的更基本的理论。薛定谔正是通过解决这个问题证明了其薛定谔方程的威力。我们这里正是要仿照当年的薛定谔，通过求解薛定谔方程来得到一个氢原子的量子力学理论。

其实玻尔的氢原子模型更令人不满意的地方在于，它只适用于氢原子，对于其它原子光谱的解释都很不成功。尤其是，玻尔的理论作为一个关于原子的理论，它却完全无法解释元素周期表。比方说，我们知道元素周期表的排布规律中出现了几个神秘的数字 $2,8,18,32....$。玻尔等人的原子模型完全无法解释这几个神秘的数字。但是，正如我们将要看到的，对氢原子的量子力学研究不仅能让我们得到正确的氢原子能级，而且只要作一个很平凡的推广就几乎能解释这些数字。这里说几乎是因为，正如我们即将看到的，通过求解薛定谔方程，我们的确能得到一串神奇的数字，但它们是 $1,4,9,16,...$。你已经发现了，这和元素周期表的数字规律正好差 2 倍。正是为了解释这个两倍，泡利提出电子还有一个内部自由度，这内部自由度只有两种可能的取值，对应电子的两个不同内部状态，这刚好带来一个 2 倍，这样从数字上来说就完全对得上了。电子的这个内部自由度当然就是人们发现的电子自旋。所以你看，通过用薛定谔方程来研究氢原子我们不光能得到氢原子能级，还能解释元素周期表，不仅如此，我们还能预言电子有自旋。

不仅如此，如果我们足够小心，其实还能发现其它伟大的隐藏秘密。正如我们将会看到的，加上电子自旋以后，我们的确得到了正确的数字 $2,8,18,32....$，但它是能级简并度，也即是说，是每一个能级可以有多少个完全不同的量子态。但是，元素周期表的 $2,8,18,32....$ 似乎是核外电子数目呈现出来的排布规律。你想一想，这是不是意味着每一个量子态里面只能排一个电子。如果你这么想，你就和当年的泡利一样发现了自然界的一个伟大的秘密。为什么每一个量子态最多只能排一个电子呢？对于当年的泡利来说，这也许是完全不可思议的，因为你想光子就不是这样，n 个光子可以处在一个完全相同的量子态，给出 $nh\nu$ 的能级，这几乎是爱因斯坦提出光量子的时候就已经知道的事情 (当然这本身也是一件神奇的事情，第一个研究光子这个神奇特性的人是印度物理学家玻色，所谓的玻色-爱因斯坦凝聚就是爱因斯坦在玻色工作的基础上提出来的)。所以，泡利为此专门提出了一条物理学原理，就是泡利不相容原理，它即是说，任何两个电子都不能处在完全相同的量子态上。泡利不相容原理是我们这个世界的深层秘密之一，可以说没有这条原理，就不可能有我们的世界。当然，这不是我们这里要讨论的话题，这里提到这个只是想表明我们即将进行的数学推导的背后隐藏着多少自然界的秘密。

要求解的方程

对于氢原子来说，我们要考虑的就是一个核外电子围绕原子核运动的问题。由于原子核比核外电子重 1 千多倍，因此为了使得分析简单一点，我们可以假设原子核处在坐标原点不动，要考虑的只是核外电子在原子核所产生的库伦场中的量子力学。当然，核外电子和原子核之间有库伦势能 $V(r)$, 它由下面的式子给出

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0 r} = -\frac{\alpha c\hbar}{r}, \tag{4.70}$$

式中我们已经引入了无量纲的精细结构常数 $\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \simeq 1/137$, 其中 c 是光速，但是我们即将要求解的是非相对论量子力学，所以从根本上来说应该是和 c 无关的，正因为如此 αc 必须作为一个整体出现，因为很显然 αc 中的 c 其实消掉了。如果我们记核外电子的质量为 m_e，那简单的量纲分析就可以使得我们引入一个长度量纲的量，$a_0 = \frac{\hbar}{m_e\alpha c}$, 这就是所

谓的玻尔半径，另外，我们还常常构造一个能量量纲的量 $\frac{1}{2}m_e\alpha^2c^2$，很显然只要将 αc 理解成衡量了电子的速度我们很快就能理解这两个量。

我们要求解的就是核外电子的定态薛定谔方程

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2+V(r)\right)\psi(\mathbf{x})=E\psi(\mathbf{x}). \tag{4.71}$$

使得这个方程有符合物理要求的非零解的 E 就是氢原子的能级。当然，我们关心的是核外电子被原子束缚围绕原子核运动的情形，也即是所谓的束缚态。束缚态上的电子当然不可能出现在空间无穷远处，因此我们的定态波函数必须满足下面的物理要求

$$r=|\mathbf{x}|\to+\infty:\psi(\mathbf{x})\to 0. \tag{4.72}$$

4.4.1 角动量算符

由于我们要求解的是电子绕核运动，这种问题根据我们在经典力学里的经验，角动量将是一个至关重要的物理量，因为在这种问题中角动量是守恒的。而且，开普勒三定律告诉我们，行星绕太阳运动只取决于两个物理量，那就是角动量和能量，开普勒第二定律告诉了我们角动量的重要性，第三定律实际上进一步告诉了我们能量的重要性，给定了角动量和能量，行星绕太阳运动的轨道就定了。当然，我们这里的电子绕核运动不是一个经典运动，而是一个量子力学问题，所以当然不可能照搬经典物理的分析，但是经典物理的分析至少告诉我们，在这样的问题中，真正重要的物理量就是角动量和能量。能量的算符当然就是哈密顿算符，这已经反映在我们要求解的定态薛定谔方程中了。我们这一小节就是要对角动量算符进行一些研究。

我们用 $\mathbf{x}$ 标记三维坐标矢量，它的三个分量 x_1,x_2,x_3 分别对应通常的 x,y,z 分量。假设我们记 $\partial_i=\frac{\partial}{\partial x_i}$，那么角动量算符 $\mathbf{L}=\mathbf{x}\times\mathbf{P}=-i\hbar\mathbf{x}\times\nabla$ 的三个分量就可以写成

$$L_1=-i\hbar(x_2\partial_3-x_3\partial_2),L_2=-i\hbar(x_3\partial_1-x_1\partial_3),L_3=-i\hbar(x_1\partial_2-x_2\partial_1). \tag{4.73}$$

或者，我们也可以利用列维-西维塔符号 ε_{ijk} 将这些表达式统一写成

$$L_i=-i\hbar\varepsilon_{ijk}x_j\partial_k, \tag{4.74}$$

式中 i,j,k 的取值范围都是 $\{1,2,3\}$，并且在这里和本节后面我们都约定对表达式中重复出现两次的指标默认求和。

为了研究角动量的大小，我们常常需要考察角动量算符的平方 $\mathbf{L}^2=L_iL_i$。为此我们注意到 $L_i^2=-\hbar^2(\varepsilon_{ijk}x_j\partial_k)(\varepsilon_{ij'k'}x_{j'}\partial_{k'})=-\hbar^2(\delta_{jj'}\delta_{kk'}-\delta_{jk'}\delta_{kj'})x_j\partial_kx_{j'}\partial_{k'}=-\hbar^2(x_j\partial_kx_j\partial_k-x_j\partial_kx_k\partial_j)=-\hbar^2(x_jx_j\partial_k\partial_k+\delta_{kj}x_j\partial_k-x_jx_k\partial_k\partial_j-3x_j\partial_j)=-\hbar^2(x_j^2\partial_k^2-x_kx_j\partial_k\partial_j-2x_j\partial_j)=-\hbar^2(x_j^2\partial_k^2-x_k\partial_kx_j\partial_j-x_j\partial_j)$(式中，第二个等号我们利用了 $\varepsilon_{ijk}\varepsilon_{ij'k'}=\delta_{jj'}\delta_{kk'}-\delta_{jk'}\delta_{kj'}$，这个等式读者可以直接验证，注意到列维-西维塔符号的全反对称性，以及所有指标只能在 $1,2,3$ 中进行取值，这个等式其实不难验证。前面推导中的第四个等号我们利用了 $\partial_kx_j=x_j\partial_k+\delta_{kj}$ 和 $\partial_kx_k=x_k\partial_k+3$，之所以出现 3 是因为要对指标 k 的三个取值进行求和。推导中的最后一个

等号我们再次利用了 $x_j\partial_k = \partial_k x_j - \delta_{kj}$)，因此我们最终得到

$$\mathbf{L}^2 = -\hbar^2(x_j^2\partial_k^2 - x_k\partial_k x_j\partial_j - x_j\partial_j) \tag{4.75}$$

$$= -\hbar^2[r^2\nabla^2 - (r\partial_r)^2 - r\partial_r] \tag{4.76}$$

$$= -\hbar^2[r^2\nabla^2 - r\partial_r^2 r], \tag{4.77}$$

这里我们进一步利用了 $x_j^2 = r^2, \partial_k^2 = \nabla^2$, 以及 $x_j\partial_j = \mathbf{x}\cdot\nabla = r\partial_r$，$r$ 表示位置矢量的长度，在球坐标系中它就是径向坐标。有时候我们也会将上面这个式子重写成

$$\nabla^2 = \frac{1}{r}\partial_r^2 r - \frac{\mathbf{L}^2}{\hbar^2 r^2}, \tag{4.78}$$

后面我们会用这个表达式将球坐标中的动能算符表示成径向的动能算符和由角动量平方提供的转动动能算符之和。

但是，除了角动量算符的平方以外，我们还常常需要关心角动量算符的各个分量在不同坐标系中的形式。比方说，引入下面的复坐标常常是很有用的，

$$x_\pm = x_1 \pm ix_2. \tag{4.79}$$

假设我们记 $\partial_\pm = \frac{\partial}{\partial x_\pm}$, 则很显然 $\partial_1 = \partial_+ + \partial_-$, $\partial_2 = i(\partial_+ - \partial_-)$，因此 $L_z = L_3 = -i\hbar(x_1\partial_2 - x_2\partial_1) = -i\hbar[i(x_+ + x_-)(\partial_+ - \partial_-) + i(x_+ - x_-)(\partial_+ + \partial_-)]/2 = \hbar[x_+\partial_+ - x_-\partial_-]$, 即

$$L_z = \hbar[x_+\partial_+ - x_-\partial_-]. \tag{4.80}$$

另外，注意到 ∂_+ 和 ∂_- 的相互独立性，也就是两者的相互对易性，我们还可以很容易地得到

$$\partial_1^2 + \partial_2^2 = 4\partial_+\partial_-. \tag{4.81}$$

另外，由于我们要求解的问题 (11.9) 是球对称的，即 $V(r)$ 是球对称的，因此球坐标也将会非常有用，它的定义是 $x_1 = r\sin\theta\cos\phi, x_2 = r\sin\theta\sin\phi, x_3 = r\cos\theta$。很显然我们有 $x_\pm = x_1 \pm ix_2 = r\sin\theta e^{\pm i\phi}$。这些就是我们将要用到的几种坐标之间的联系。

下面让我们来考察因为 ϕ 的无穷小改变 $\delta\phi$ 所引起来的坐标改变 $\delta x_\pm$, 由 $x_\pm$ 的球坐标表达式，很显然结果是 $\delta x_+ = ix_+\delta\phi$, $\delta x_- = -ix_-\delta\phi$, 也即是说 $\delta\phi$ 既会引起 x_+ 的无穷小改变，又会引起 x_- 的无穷小改变 (但显然 $\delta\phi$ 与 x_3 无关)。因此假设有一个函数 $f(x_+, x_-, x_3)$, 我们考虑它在 $\delta\phi$ 的改变下所产生的无穷小改变 δf，那就必然有 $\frac{\partial}{\partial\phi}f = \frac{\delta}{\delta\phi}f = \frac{\delta x_+}{\delta\phi}\frac{\partial f}{\partial x_+} + \frac{\delta x_-}{\delta\phi}\frac{\partial f}{\partial x_-} = i(x_+\frac{\partial}{\partial x_+} - x_-\frac{\partial}{\partial x_-})f$，注意到函数 f 是任意的，因此我们可以将这个结果应用到 (4.80) 式，就会得到

$$L_z = -i\hbar\frac{\partial}{\partial\phi}. \tag{4.82}$$

从这个式子我们可以看到，在球坐标中 L_z 只和角度坐标有关，而和径向坐标 r 无关。然而，在角动量的三个分量中，L_z 在物理上并没有任何特殊的地方，因此这也就告诉我们，在球坐标中，另外两个分量的角动量算符 L_x, L_y 同样也只会和角度坐标 θ, ϕ 有关，而和

径向坐标 r 必然是无关的。当然由于球坐标的具体定义式对于 x,y,z 三个分量的不对称性，所以算符 L_x 和 L_y 在球坐标中的具体表达式可能会比 L_z 的表达式复杂一些，但是这并不意味着它们在物理上有什么不平等性，正因为如此，我们只需要选择 L_z 来研究通常就够了。

4.4.2 能级和径向波函数

下面我们开始具体的求解过程。首先利用 (4.78) 式，我们可以将要求解的定态薛定谔方程 (11.9) 重写为，

$$\left(-\frac{\hbar^2}{2m_e}\frac{1}{r}\partial_r^2 r+\frac{\mathbf{L}^2}{2m_e r^2}+V(r)\right)\psi(\mathbf{x})=E\psi(\mathbf{x}). \tag{4.83}$$

这个方程的下面这一等价形式也将是有用的，

$$\mathbf{L}^2\psi(\mathbf{x})=2m_e r^2(E-V(r))\psi(\mathbf{x})+\hbar^2 r\partial_r^2 r\psi(\mathbf{x}). \tag{4.84}$$

让我们首先来考察波函数在坐标原点附近的行为。很显然，核外电子出现在坐标原点处的概率必定是有限的，而且由于 $V(r)\sim 1/r$，从而 $r^2V(r)\sim r$，因此对于核外电子来说，坐标原点并不存在什么根本的奇异性，所以波函数 $\psi(\mathbf{x})$ 在坐标原点附近必然是有限且光滑的。因此，我们可以在直角坐标中将波函数 $\psi(\mathbf{x})$ 在坐标原点处进行泰勒展开，假设领头阶是 l 阶小量 ($l\geq 0$ 且是整数)，这也即是说，我们假设在泰勒展开的领头阶，波函数是三个直角坐标分量 $x_i,i=1,2,3$ 的 l 阶齐次函数，之所以得是齐次当然是因为所有的领头阶小量都得有相同的阶 (比方说 $ax_1^2+bx_1x_2+cx_2x_3+dx_3^2$ 就是一个 2 阶齐次函数)，在球坐标中由于每一个 x_i 的球坐标表达式中都含有一个 r，因此我们可以将这个 l 阶齐次函数记为 $Cr^lY(\theta,\phi)$(C 是一个常数)。

为了考察方程 (4.84) 在坐标原点附近，即当 $r\to 0$ 时的行为。我们注意到 $V(r)\sim 1/r$，而能量 E 是一个常数，因此当 $r\to 0$ 时，$2m_e r^2(E-V(r))\psi(\mathbf{x})$ 将是一个 $l+1$ 阶小量，而 $r\partial_r^2 r\psi(\mathbf{x})\sim(r\partial_r^2 r)(Cr^lY(\theta,\phi))=C\cdot l(l+1)r^lY(\theta,\phi)+O(r^{l+1})$($O(r^{l+1})$ 表示更高阶的小量)。因此，保留到领头的 l 阶而忽略所有的更高阶小量，我们可以看到方程 (4.84) 在原点附近将变成 $\mathbf{L}^2(r^lY(\theta,\phi))=l(l+1)\hbar^2 r^lY(\theta,\phi)$，同时，由于角动量算符只和角度坐标有关，而和径向坐标 r 无关，所以我们可以将这个等式两边的 r^l 消去，从而得到

$$\mathbf{L}^2Y(\theta,\phi)=l(l+1)\hbar^2Y(\theta,\phi). \tag{4.85}$$

这个方程意味着，$\mathbf{L}^2$ 的本征值为 $l(l+1)\hbar^2$，l 是非负整数，而且前面通过直角坐标分量的 l 阶齐次函数所引入的 $Y(\theta,\phi)$ 正是 $\mathbf{L}^2$ 的本征函数。由于 l 是非负整数，所以这里隐含着一个重要的物理结论，即角动量的大小是量子化的，但是具体的量子化形式和玻尔在其氢原子模型中所假设的形式有所不同。而这里的 l，就称为角量子数，因为它决定了角动量的大小。

现在我们考察方程 (4.84) 在任意位置的一般行为。注意到方程 (4.84) 等号左边的算符只涉及到角度坐标而与径向坐标无关，而等号右边的算符只涉及到径向坐标而与角度坐标无关，因此我们可以知道，这个方程必定可以分离变量，即必有 $\psi(\mathbf{x})=R(r)A(\theta,\phi)$ 的

形式。但是由于 $r\to 0$ 时，$\psi(\mathbf{x})=R(r)A(\theta,\phi)\to Cr^l Y(\theta,\phi)$, 因此必有 $A(\theta,\phi)=Y(\theta,\phi)$, 即

$$\psi(\mathbf{x})=R(r)Y(\theta,\phi), \tag{4.86}$$

而且径向波函数 $R(r)$ 在 $r\to 0$ 时的渐进行为是

$$R(r)\to Cr^l+.... \tag{4.87}$$

另外,由于我们考察的是束缚态,因此由通常的波函数归一化条件 $\int_{S^2}\sin\theta d\theta d\phi\int_0^\infty r^2 dr R^2|Y|^2=1$ 可以知道，径向波函数和角度部分波函数分别都得归一化，而且在径向波函数的归一化中，由于 $\int^\infty drr^2R^2(r)$ 在无穷远处的积分贡献必须有限，这就意味着当 $r\to+\infty$ 时，必有

$$rR(r)\to 0. \tag{4.88}$$

将波函数的表达式 (4.86) 代入定态的薛定谔方程 (4.83), 并利用 $\mathbf{L}^2$ 的本征方程 (4.85), 进而将方程 (4.83) 两边的角度部分波函数 $Y(\theta,\phi)$ 消掉，我们就可以得到径向波函数 $R(r)$ 所满足的方程,

$$\left(-\frac{\hbar^2}{2m_e}\frac{1}{r}\partial_r^2 r+\frac{l(l+1)\hbar^2}{2m_e r^2}+V(r)\right)R(r)=ER(r). \tag{4.89}$$

下面我们令 $rR(r)=u(r)$, 则方程 (4.89) 可以转化为一个关于 $u(r)$ 的方程,

$$-\frac{\hbar^2}{2m_e}\partial_r^2 u(r)+\left(\frac{l(l+1)\hbar^2}{2m_e r^2}+V(r)\right)u(r)=Eu(r). \tag{4.90}$$

如果我们将 r 看作是定义在 $[0,+\infty)$ 区间上的一维坐标变量，那么这个方程 (4.90) 就是一个标准的一维定态薛定谔方程，其中有效的势能为 $\frac{l(l+1)\hbar^2}{2m_e r^2}-\frac{\alpha c\hbar}{r}$。而且由原来的径向波函数渐进条件 (4.87) 和 (4.88) 可以知道，$u(r)$ 必定满足如下渐进条件

$$r\to 0 \quad : \quad u(r)\to Cr^{l+1} \tag{4.91}$$

$$r\to+\infty \quad : \quad u(r)\to 0. \tag{4.92}$$

我们可以引入无量纲的坐标 ρ 以及无量纲的能量 ε,

$$\rho=r/a_0,\ \varepsilon=E/(\frac{1}{2}m_e\alpha^2c^2). \tag{4.93}$$

从而可以将方程 (4.90) 变成,

$$\partial_\rho^2 u(\rho)+\left(\varepsilon+\frac{2}{\rho}-\frac{l(l+1)}{\rho^2}\right)u(\rho)=0. \tag{4.94}$$

考虑这个方程在 $\rho\to+\infty$ 时的渐进形式，会得到 $\partial_\rho^2 u(\rho)+\varepsilon u(\rho)=0$, 如果 $\varepsilon>0$，那这个方程的解是 cos 和 sin 形式的振荡解，因此不可能满足无穷远处的衰减型边界条件 (4.92)。这也即是说，仅当 $\varepsilon<0$ 时才可能有束缚态的存在。为此我们令 $\varepsilon=-\gamma^2(\gamma>0)$, 因此渐进方程就是 $\partial_\rho^2 u(\rho)-\gamma^2 u(\rho)=0$, 渐进解就是 $u(\rho)\to e^{-\gamma\rho}$ (还有一个指数增长的特解由于不

满足无穷远处的衰减边界条件而必须抛弃)。因此，令 $u(\rho)=e^{-\gamma\rho}f(\rho)$, 代入方程 (4.94), 就可以得到一个关于函数 $f(\rho)$ 的微分方程

$$\left(\partial_\rho^2-2\gamma\partial_\rho+\frac{2}{\rho}-\frac{l(l+1)}{\rho^2}\right)f(\rho)=0. \tag{4.95}$$

由于在 $\rho\to 0$ 处的边界条件 (4.91), 我们可以知道 $f(\rho)$ 在原点附近泰勒展开的领头阶一定有 ρ^{l+1} 这样的形式。不妨设整个 $f(\rho)$ 可以写成级数形式

$$f(\rho)=\sum_{k=l+1}^{+\infty}a_k\rho^k, \tag{4.96}$$

注意由于渐进条件，这个级数必然是从 $l+1$ 阶开始的。将这个级数代入方程 (4.95), 并比较所得结果每一个 ρ^k 项的系数，我们可以得到，对于领头阶 $k=l+1$，方程给出来的是一个恒等式，这是不出意外的，因为在我们求解整个问题的一开始我们就已经考察过方程在坐标原点处的领头阶了，我们的边界条件 (4.91) 就是这样导出来的。对于 $k\geq l+2$ 阶，方程给出来的是一个递推关系

$$a_k=2\frac{\gamma(k-1)-1}{k(k-1)-l(l+1)}a_{k-1}. \tag{4.97}$$

因此给定 a_{l+1} 我们就能递推出后面所有的 a_k，这也就是说，我们能够将整个函数 $f(\rho)$ 确定到只剩下一个整体的未定常数 a_{l+1}。到此为止好像我们已经完成所有的求解了。

但，还是让我们来检验一下这个解符不符合物理要求，为此我们取 $k\to\infty$, 则根据递推关系 (4.97) 就有

$$\frac{a_k}{a_{k-1}}\to\frac{2\gamma}{k}. \tag{4.98}$$

注意到 $e^{2\gamma\rho}=\sum_k\frac{(2\gamma)^k}{k!}\rho^k$, 其相邻两个展开项系数之比正好是 $\frac{2\gamma}{k}$, 因此这就意味着当 $\rho\to+\infty$ 时 (这时候级数展开的高阶项是绝对主导的)，我们解出来的 $f(\rho)\to e^{2\gamma\rho}$。而这又进一步意味着原来的 $u(\rho)=e^{-\gamma\rho}f(\rho)\to e^{\gamma\rho}$, 显然不满足无穷远处的衰减性边界条件。因此，我们找到的这个解实际上不符合物理要求。

那这是不是意味着这个问题没有解呢？不是的，问题在于我们刚才的分析中其实有一个漏洞，如果 γ 是一个任意的大于 0 的实数，那我们刚才的分析没有任何问题，这时候的确不存在满足物理要求的解。但是，如果

$$\gamma=\frac{1}{n}, \tag{4.99}$$

这里 n 是某个正整数 (即 $n\geq 1$)，那我们的递推关系 (4.97) 就会在 $k=n+1$ 的地方截断。也就是说我们将有 $a_{n+1}=0$, 那么根据递推公式 (4.97)，此后所有 $k>n+1$ 的系数 a_k 也都会等于 0。因此这时候我们解出来的 $f(\rho)$ 其实是一个 n 次多项式 (这时候我们不能再任意取 $k\to+\infty$ 了，(4.98) 式的分析也就不再适用了)，它在无穷远处当然不会指数增长，那这时候关于 $u(\rho)=e^{-\gamma\rho}f(\rho)$ 在无穷远处的衰减性边界条件 (4.92) 就能够得到满足，相

应的解就是符合物理要求的。当然，需要注意的是，由于 n 次多项式 $f(\rho)$ 的所有级数阶都必须大于等于 $l+1$, 所以我们有

$$l+1 \leq n. \tag{4.100}$$

也就是说，对于任意正整数 n, 任意非负整数 l, 只要它们满足 $l \leq n-1$, 那就有一个符合物理要求的解, 我们记为 $f_{nl}(\rho)$。利用 $f_{nl}(\rho)$ 我们最终就能得到相应的径向波函数，为了强调它依赖于 n 和 l, 现在我们将之记为 $R_{nl}(r)$,

$$R_{nl}(r) = \frac{u_{nl}(r)}{r} = N_{nl} e^{-\frac{r}{na_0}} F_{nl}(r/a_0), \tag{4.101}$$

其中 $F_{nl}(\rho) = f_{nl}(\rho)/\rho$ 是一个 $n-1$ 次多项式，其最低次项为 ρ^l 项，式中 N_{nl} 为归一化系数。对于这种符合物理要求的解，我们的无量纲能量 $\varepsilon = -\gamma^2 = -1/n^2$，因此也就意味着最初的能量本征值 E 必须满足

$$E = -\frac{1}{n^2}(\frac{1}{2} m_e \alpha^2 c^2). \tag{4.102}$$

这正是完全正确的能够成功解释氢原子光谱数据的能级公式。

但是，现在我们已经比玻尔多得到一些东西了，注意到能级公式 (4.102) 只依赖于 n, 而和角量子数 l 无关，因此通常将 n 称为主量子数。也就是说，给定主量子数 n, 我们不只有一个本征态的解，而是所有满足 $l \leq n-1$ 的解 $R_{nl}(r)$ 都对应相同的能量本征值，这些不同 l 的解是简并的。能级简并这是玻尔的模型不可能描述的。

实际上，径向波函数可以用标准的特殊函数表达出来，为了方便读者查阅，我们将结果写在下面

$$\begin{aligned} R_{nl}(r) = & \frac{1}{(2l+1)!}\left(\frac{2r}{na_0}\right)^l e^{-r/na_0}\left[(\frac{2}{na_0})^3 \frac{(n+l)!}{2n(n-l-1)!}\right]^{1/2} \\ & \times F(-n+l+1; 2l+2; 2r/na_0). \end{aligned} \tag{4.103}$$

式中 $F(a;b;x)$ 为合流超几何函数。这个径向波函数已经归一化了，满足归一化条件 $\int_0^{+\infty} u_{nl}^2(r)dr = \int_0^{+\infty} R_{nl}^2(r) r^2 dr = 1$。

散射态

以上我们关心的都是束缚态，因为我们是在研究氢原子，假设认为原子核近似不动，那氢原子的定义就是一个电子在库伦势场 $V(r) = -\frac{\alpha c\hbar}{r}$ 中的束缚态。但是，电子在库伦势场中除了有束缚态，其实还有散射态，对应于电子被库伦势场 $V(r) = -\frac{\alpha c\hbar}{r}$ 散射的问题。假设电子距离原子核无穷远时的动量大小为 $\hbar k$, 那么这种散射态的能量就应该是连续谱 $E = \frac{\hbar^2 k^2}{2m_e}$，因为距离无穷远时，电子当然可以看成是自由粒子。

我们可以这样来得到散射态的径向波函数: 首先，我们注意到氢原子的能级为 $E = -\frac{1}{2} m_e (\alpha c)^2/n^2$, 因此 $\frac{1}{n} = \sqrt{\frac{-2E}{m_e(\alpha c)^2}}$ 当然，这里 E 是小于 0 的。现在，假设我们将 E 解析延拓到 $E = \frac{\hbar^2 k^2}{2m_e} > 0$ 的散射态情形，那这时候 $\frac{1}{n}$ 就应该延拓成

$$\frac{1}{n} = ik\frac{\hbar}{m_e \alpha c} = ika_0. \tag{4.104}$$

最后，我们将这个新的解析延拓后的 n 值代入径向波函数 (4.103), 从而就可以得到散射态的径向波函数 $R_{kl}(r)$

$$R_{kl}(r) = C_{kl}(2kr)^l e^{-ikr} F(\frac{i}{ka_0} + l + 1; 2l + 2; 2ikr). \tag{4.105}$$

这里我们已经注意到，由于散射态波函数在无穷远处不趋于零，无法进行通常的归一化，所以我们干脆保留了一个整体常数 C_{kl}。

4.4.3 球谐函数

上一节的求解过程虽然得到了一些最重要的物理结论，但是我们仅仅只求出了径向波函数的具体形式，对于角度部分的波函数 $Y(\theta,\phi)$, 我们只有两个一般性的结论：第一，当把 $r^l Y(\theta,\phi)$ 用直角坐标来表达时，它是三个直角坐标分量的齐次度为 l 的齐次函数，也就是我们上一节所谓的 l 阶齐次函数，这里我们将把它记为 $G_l(\mathbf{x})$, 当然这个记号和原来的 $r^l Y(\theta,\phi)$ 仅仅只有坐标系的不同，因此

$$G_l(\mathbf{x}) = r^l Y(\theta,\phi), \tag{4.106}$$

第二，$r^l Y(\theta,\phi)$ 满足 $\mathbf{L}^2$ 的本征方程 (4.85)。但是，我们还没有解出 $Y(\theta,\phi)$ 的具体表达式，这一节我们就来完成这件事情。

为了具体解出 $r^l Y(\theta,\phi)$，关键就是要求解本征方程 (4.85)。最简单地做到这一点的办法是，注意到 $\mathbf{L}^2$ 和拉普拉斯算符 ∇^2 之间的关系式 (4.77) 和 (4.78), 我们很容易验证，$\mathbf{L}^2$ 的本征方程 (4.85) 其实等价于下面的方程

$$\nabla^2(r^l Y(\theta,\phi)) = \nabla^2 G_l(\mathbf{x}) = 0. \tag{4.107}$$

这个方程在直角坐标中是最容易求解的，或者等价地说在 $\{x_+,x_-,x_3\}$ 坐标中是最容易求解的，原因在于，首先，拉普拉斯算符在直角坐标中的形式远比球坐标简单，更重要的原因是，我们已经知道了 $r^l Y(\theta,\phi)$ 在直角坐标，或者等价的在 $\{x_+,x_-,x_3\}$ 坐标中是一个齐次函数 $G_l(\mathbf{x})$, 而这样的函数在数学上处理起来比其它函数远为简单。

下面我们将主要采用 $\{x_+,x_-,x_3\}$ 坐标，这其实是因为一会儿我们还要接着研究 L_z 的本征波函数, 而 L_z 在 $\{x_+,x_-,x_3\}$ 坐标中的表达形式 (4.80) 比它在更通常的直角坐标中的表达形式还要容易处理一些。

由于公式 (4.81)，我们容易知道 $\nabla^2 = \partial_1^2 + \partial_2^2 + \partial_3^2 = 4\partial_+\partial_- + \partial_3^2$, 所以我们可以将方程 (4.107) 改写成

$$(4\partial_+\partial_- + \partial_3^2) G_l(x_+,x_-,x_3) = 0. \tag{4.108}$$

注意这个方程是一个线性方程，所以所有可能解所构成的解空间一定是一个线性空间，再加上 G_l 是齐次函数，那这个方程其实就很容易求解。比方说，对于 $l = 2$ 的情况，$G_2(x_+,x_-,x_3)$ 有如下五个线性无关的特解

$$x_+^2,\ x_+x_3,\ x_+x_- - 2x_3^2,\ x_-x_3,\ x_-^2. \tag{4.109}$$

人们很容易验证这五个解都满足方程 (4.108), 因此它们的任意线性叠加也都是 (4.108) 的解。找到这些解的方法很简单,比方说,由于 G_2 的齐次度是 2,你可以假设 $G_2 = ax_+x_- + bx_3^2$, 代进方程 (4.108), 你很容易发现仅当 $b = -2a$ 时，方程 (4.108) 才能满足，这就得到了特解 $x_+x_- - 2x_3^2$, 你作其它的不同假设就会得到不同的特解。如果你一开始假设一个最一般性的 2 阶齐次函数，那代入方程 (4.108)，你将得到所有上面 5 个特解的线性组合，这也就是 $l = 2$ 时的通解。这种方法显然可以推广到任意 l 的情形。

(4.109) 式给出来的这个例子足以说明，给定一个 l，G_l 有多个线性无关的可能解，由于方程 (4.108) 与 $\mathbf{L}^2$ 的本征方程等价，这也就是说，$\mathbf{L}^2$ 对应同一个本征值 $l(l+1)\hbar^2$ 的本征态是简并的，有多个可能性。那么如何进一步区分这些简并的解呢? 这就要用到角动量算符的 z 分量 L_z，尤其是用到它的 $L_z = \hbar(x_+\partial_+ - x_-\partial_-)$ 的形式, 简单来说就是，我们可以取方程 (4.108) 的每一个线性独立的特解同时为 L_z 的本征波函数。比方说，由 $L_z = \hbar(x_+\partial_+ - x_-\partial_-)$ 人们很容易验证，上面式子 (4.109) 给出来的 5 个 $l = 2$ 的特解在 L_z 的作用下分别具有如下本征值

$$2\hbar, \hbar, 0, -\hbar, -2\hbar. \tag{4.110}$$

所以我们常将这 5 个解分别记作

$$G_{22} \sim x_+^2,\ G_{21} \sim x_+x_3,\ G_{20} \sim x_+x_- - 2x_3^2,\ G_{2-1} \sim x_-x_3,\ G_{2-2} \sim x_-^2, \tag{4.111}$$

很容易看出来，这里我们给齐次函数 G 添上的第 2 个指标是用来标记 L_z 的不同本征值相对应的量子数的。

其实人们很容易验证，对于任何变量 x，如果将算符 $x\partial_x$ 作用在 x 的某个 p 次齐次函数上 x^p 上，那么得到的就是 px^p, 即 $x\partial_x(x^p) = px^p$. 也即是说，$x\partial_x$ 这样的算符的作用效果其实就相当于数一个表达式关于 x 的齐次度。

很显然，算符 $L_z/\hbar = x_+\partial_+ - x_-\partial_-$ 就是由两个这样的数齐次度的算符相减而构成的，$x_+\partial_+$ 数的是一个表达式中含有 x_+ 因子的多少次方，$x_-\partial_-$ 数的是同一个表达式中含有 x_- 因子的多少次方，假设某一个表达式为 $x_+^{\nu_+}x_-^{\nu_-}$, 那么这个表达式对应的 $L_z/\hbar$ 的本征值就是 $\nu_+ - \nu_-$, 通常记为 m, 即

$$m = \nu_+ - \nu_-. \tag{4.112}$$

而所谓我们取方程 (4.108) 的解同时是 L_z 的本征态，也就是要求任何一个给定解中的每一项都有相同的 $\nu_+ - \nu_- = m$，这时候这个 $\nu_+ - \nu_- = m$ 就是这个解在 $L_z/\hbar$ 作用下的本征值，我们就是利用不同的 m 来区分方程 (4.108) 的那些相互简并的解。

那么给定一个 l, 有多少个不同的 $m = \nu_+ - \nu_-$ 呢? 由于 l 是 G_l 总的齐次度数，其中除了 x_+,x_- 的贡献之外还要包括 x_3 的贡献，所以很显然，任何解中的所有项都得满足

$$\nu_+ + \nu_- \le l. \tag{4.113}$$

满足这个约束的 $\nu_+ - \nu_-$ 有多少可能取值呢? 稍微想一下就能发现 (你可以注意 $\nu_+ - \nu_-$ 必然是整数，而且 $-\nu_+ - \nu_- \le \nu_+ - \nu_- \le \nu_+ + \nu_-$)，只有下面这些可能取值

$$l,\ l-1,\ l-2,\ .....-l+1,\ -l, \tag{4.114}$$

即一共有 $2l+1$ 个不同的可能性。

(9.15) 式给出来的就是在给定 l 的前提之下，$L_z/\hbar$ 的本征值 m 的所有可能性，通常我们把每一个可能性所对应的本征波函数记为 G_{lm}，因此通过引入量子数 m，我们就可以将 $\mathbf{L}^2$ 的所有简并本征态一一区分开来。也就是说，G_{lm} 同时满足如下本征方程

$$\mathbf{L}^2 G_{lm} = l(l+1)\hbar^2 G_{lm},\ L_z G_{lm} = m\hbar G_{lm}, \tag{4.115}$$

其中，m 的 $2l+1$ 个不同取值由 (9.15) 式给出。

解出来所有的 G_{lm} 以后，利用 $x_\pm = r\sin\theta e^{\pm i\phi}, x_3 = r\cos\theta$ 变换到球坐标中，再根据定义 (4.106), 我们就能得到角度部分的波函数 $Y_{lm}(\theta,\phi)$,

$$G_{lm} = r^l Y_{lm}(\theta,\phi). \tag{4.116}$$

通常人们称 $Y_{lm}(\theta,\phi)$ 为球谐函数，它显然满足

$$\mathbf{L}^2 Y_{lm} = l(l+1)\hbar^2 Y_{lm},\ L_z Y_{lm} = m\hbar Y_{lm}. \tag{4.117}$$

比方说，对于前面给出来的 5 个 $l=2$ 的本征函数 (4.109), 我们很容易变换到球坐标中得到 5 个相应的球谐函数

$$Y_{22} \sim \sin^2\theta e^{i2\phi},\ Y_{21} \sim \sin\theta\cos\theta e^{i\phi},\ Y_{20} \sim \sin^2\theta - 2\cos^2\theta, \tag{4.118}$$

$$Y_{2-1} \sim \sin\theta\cos\theta e^{-i\phi},\ Y_{2-2} \sim \sin^2\theta e^{-i2\phi}, \tag{4.119}$$

当然，我们没有对这些球谐函数进行归一化，不过人们很容易将它们在单位球面上进行归一化。

由于球谐函数是角度部分的波函数，它所对应的量子态由量子数 l,m 标记，因此可以记为 $|l,m\rangle$，根据 (4.117), 它将同时是 $\mathbf{L}^2$ 和 L_z 的本征态，满足

$$\mathbf{L}^2|l,m\rangle = l(l+1)\hbar^2|l,m\rangle,\quad L_z|l,m\rangle = m\hbar|l,m\rangle. \tag{4.120}$$

作为厄米算符的不同本征态，$|l,m\rangle$ 当然是正交归一的，满足

$$\langle l',m'|l,m\rangle = \delta_{l'l}\delta_{m'm}. \tag{4.121}$$

另外，利用波函数与狄拉克符号的关系，我们也可以将球谐函数写成

$$\langle\theta,\phi|l,m\rangle = Y_{lm}(\theta,\phi). \tag{4.122}$$

作为球面上的位置本征态，$|\theta,\phi\rangle$ 当然满足

$$\int d\Omega|\theta,\phi\rangle\langle\theta,\phi| = 1. \tag{4.123}$$

式中的积分是在整个单位球面上积分，Ω 为立体角，$d\Omega = \sin\theta d\theta d\phi$，因此通过插入 (4.123) 式，我们就可以将正交归一关系 (4.121) 转化为球谐函数的正交归一关系

$$\int d\Omega Y^*_{l'm'} Y_{lm} = \delta_{l'l}\delta_{m'm}. \tag{4.124}$$

我们正是利用这个式子来将球谐函数归一化的。

有时候我们也会将球坐标中的 θ,ϕ 角的方向用某个单位矢量 $\mathbf{n}$ 来表示，这时 $|\theta,\phi\rangle$ 态就可以记为 $|\mathbf{n}\rangle$，球谐函数 $Y_{lm}(\theta,\phi)$ 就可以记为 $Y_{lm}(\mathbf{n})$。而球谐函数与相应狄拉克符号之间的关系就可以重写成

$$\langle \mathbf{n}|l,m\rangle = Y_{lm}(\mathbf{n}). \tag{4.125}$$

当然，人们容易按照我们这一小节的方法求出任意球谐函数的表达式，并将它们归一化。不过，其实数学家们早就得到球谐函数的统一表达式了，对于 $m\geq 0$，它可以由下式给出

$$Y_{lm}(\theta,\phi) = \frac{(-)^l}{2^l l!}\sqrt{\frac{(2l+1)}{4\pi}\frac{(l+m)!}{(l-m)!}}e^{im\phi}\frac{1}{\sin^m\theta}\frac{d^{l-m}}{d(\cos\theta)^{l-m}}(\sin\theta)^{2l}, \tag{4.126}$$

另外还有

$$Y_{l-m}(\theta,\phi) = (-)^m Y^*_{lm}(\theta,\phi). \tag{4.127}$$

这两个式子就统一给出了所有的球谐函数。特别的，从这些式子我们可以知道，对于 $m=0$ 的球谐函数 $Y_{l0}(\theta,\phi)$，我们有

$$Y_{l0}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta), \tag{4.128}$$

式中 $P_l(x)$ 为勒让德多项式。

4.4.4 理解元素周期表的神秘规律

到此为止就已经完成了我们对氢原子的量子力学研究。但是，前面我们说过，关于氢原子的这个研究只要作一个很平凡的推广，就能用来理解元素周期表所呈现出来的 $2,8,18,32...$ 这样的规律。为此，让我们来考察一个原子序数为 Z 的原子，暂时让我们假设只有一个电子在围绕它的原子核运动，我们来看一下这样一个系统会有什么样的能级和什么样的定态波函数。很显然，现在的情况和氢原子其实是完全类似的，只不过现在的库伦势能为

$$V(r) = -\frac{Ze^2}{4\pi\varepsilon_0 r} = -\frac{Z\alpha c\hbar}{r}. \tag{4.129}$$

很显然，关于氢原子的所有推导和求解都可以照搬，只需要将氢原子问题中出现的 αc 替换成 $Z\alpha c$ 就可以了。因此这也就意味着，氢原子解中的 a_0 要替换成 a_0/Z，而所得到的单电子能级现在应该是

$$-\frac{1}{n^2}\left(\frac{1}{2}m_e Z^2\alpha^2c^2\right). \tag{4.130}$$

相应的定态波函数 $\psi_{nlm}(r,\theta,\phi)$ 也完全类似于氢原子，都可以写成径向波函数 $R_{nl}(r)$ 和球谐函数 $Y_{lm}(\theta,\phi)$ 的乘积，

$$\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi). \tag{4.131}$$

也即是说，每一个量子态要由三个量子数 (n,l,m) 来刻画，当然 $0\le l\le n-1$, $m=0,\pm1,\pm2,...,\pm l$，而相同 n 不同 l,m 的所有量子态都是简并的。我们可以计算一下，给定 n 的简并度是多少，首先 m 有 $2l+1$ 种可能性，其次 l 可以从 0 取到 $n-1$，所以这个简并度就是

$$\sum_{l=0}^{n-1}(2l+1)=n^2. \tag{4.132}$$

对于 $n=1,2,3,4...$, n^2 的简并度将给出 $1,4,9,16...$, 正如泡利所注意到的，这串数字正好是元素周期表的 $2,8,18,32...$ 的一半。如果我们考虑到电子自旋，由于自旋有自旋向上和自旋向下两种可能性，那最终的简并度将正好是 $2n^2$，与元素周期表的规律完全一致。

之所以单电子量子态的简并度能与元素周期表相吻合，是因为我们可以设想，核外电子是一个接着一个地加入到原子核外面的。作为一个初步的近似，我们可以忽略不同电子之间的相互作用，这样，每一个新加进来的电子的量子力学处理都可以看是上面所论述的单电子问题。先加进来的电子会占据能量较低的那些量子态，而由于泡利不相容原理，一个量子态上只能占据一个电子，后加进来的电子就只能占据其它的未被占据的量子态。电子就这样一个接着一个地从低能级填充到高能级，直至填满 Z 个电子。根据这样的理解，元素周期表的电子排布规律和单电子的量子态之间就能一一对应，因此 $2n^2$ 的简并度正好对应元素周期表的规律也就不难理解了。

当然，上面这幅物理图像在细节上有很多不对的地方。问题之一在于我们忽略了电子与电子之间的相互作用，基于此人们不难想见，即使我们只关心每一个核外电子所带来的能量增量，由公式 (4.130) 所给出来的结果肯定也是不对的，因此根据这个结果而把每一个电子的贡献加起来得到的整个原子的能级也不会对。更何况，除了电子间的相互作用之外，还有一个严重的物理问题我们也没有考虑，那就是全同性原理，也就是说，所有的电子是完全全同完全不可区分的。因此，一个原子所有的核外电子之间也完全不可区分，这一点也会给原子最终的能级带来不小的修正。当然，怎么正确地算出任意一个原子的能级是原子物理学的专门研究课题，所以我们还是到此为止吧。

4.4.5 习题

1. 请按照本节的方法计算出 $l=3$ 时的 7 个球谐函数 $Y_{3\pm3}$、$Y_{3\pm2}$、$Y_{3\pm1}$、$Y_{3\pm0}$，并请将它们归一化。

2. 设 $|\psi_n\rangle$ 为某个系统 H 的束缚态，满足定态薛定谔方程 $H|\psi_n\rangle=E_n|\psi_n\rangle$。请证明

$$\frac{\partial E_n}{\partial g}=\langle\psi_n|\frac{\partial H}{\partial g}|\psi_n\rangle. \tag{4.133}$$

其中 g 为系统的某个任意参数，也就是哈密顿量 H 中的某个任意参数。这个结果有时候也被称作 Hellmann-Feynman 定理。

3. 请利用上一题证明的 Hellmann-Feynman 定理，(1). 证明对于线性谐振子，必有 $\langle n|T|n\rangle=\langle n|V|n\rangle$。(2). 对于氢原子的任何一个本征态 $|nlm\rangle$(波函数为 ψ_{nlm})，请计算 $\langle nlm|\frac{1}{r}|nlm\rangle$。(3). 对于氢原子定态 $|nlm\rangle$, 证明 $\langle nlm|T|nlm\rangle=-\frac{1}{2}\langle nlm|V|nlm\rangle$, 式中 $T=\frac{\mathbf{P}^2}{2m_e}$ 为动能算符。

4. 一量子的刚性转子，转动惯量为 I，哈密顿量为 $H=\frac{\mathbf{L}^2}{2I}$。求下列情况下的定态能量和定态波函数。(1). 转子绕一固定转轴转动。(2). 转子绕一固定点转动。

5. 假设用一半径为 R 的球形空腔将粒子束缚起来，系统的势能为

$$V(r)=\begin{cases}0, & r<R\\ +\infty, & r\geq R\end{cases}. \tag{4.134}$$

求粒子的能级和定态波函数。

4.5 朗道能级

4.5.1 朗道能级

让我们来考察一个在两维 $x-y$ 平面内运动的电子，电荷为 e(对于电子 e 当然为负，但为了公式简单，下面我们不妨默认 $e>0$)，假设我们在垂直的 z 方向加上一个均匀磁场 B，朗道发现，虽然并没有施加一个势阱将电子约束在有限的区域内，这个系统依然出现了能量量子化现象，这就是所谓的朗道能级。另外，这个系统的每一个量子化能级都有很高的简并度，包括基态能级。而且，这样一个两维系统并不是理论家们的虚构之物，实际上，后面我们将看到，它和所谓的量子霍尔效应密切相关。

根据我们在第三章中所学的知识，上面所说的这样一个系统的哈密顿量可以写成

$$H=\frac{\vec{\Pi}^2}{2m}=\frac{\Pi_x^2}{2m}+\frac{\Pi_y^2}{2m}. \tag{4.135}$$

式中 $\vec{\Pi}=\mathbf{P}-e\mathbf{A}$ 就是所谓的力学动量，现在它当然只有 x,y 两个分量。根据第三章的知识我们知道，Π_x 和 Π_y 不对易，它们的对易子为

$$[\Pi_x,\Pi_y]=i\hbar eB_z=i\hbar eB. \tag{4.136}$$

不妨引入一个特征长度 l_B, 称之为磁长度，其定义是

$$l_B=\sqrt{\frac{\hbar}{eB}}. \tag{4.137}$$

对于处于强度为 1 特斯拉的磁场中的电子，其磁长度约为 $2.5\times10^{-8}m$。因此，在 $2\pi l_B^2$ 面积上的磁通为 $B(2\pi l_B^2)=2\pi\hbar/e=\Phi_0$，即刚好为一个磁通量子。

这个二次型的哈密顿量 (4.135) 和相应的对易关系 (4.136) 都很容易使我们想起线性谐振子。因此，下面我们引入产生湮灭算符 a 和 $a^\dagger$，其定义是 (注意到根据德布罗意关系，$\hbar$ 除以长度将具有动量量纲)

$$(\frac{\hbar}{l_B})a=\frac{1}{\sqrt{2}}(\Pi_x+i\Pi_y),\quad (\frac{\hbar}{l_B})a^\dagger=\frac{1}{\sqrt{2}}(\Pi_x-i\Pi_y). \tag{4.138}$$

则根据 (4.136) 式可以验证，$[a,a^\dagger]=1$，因此它们的确是产生湮灭算符。而且由这个产生湮灭算符的定义式 (4.138) 可以知道

$$(\frac{\hbar}{l_B})^2\left(a^\dagger a+\frac{1}{2}\right)=(\frac{\hbar}{l_B})^2\frac{1}{2}\left(a^\dagger a+aa^\dagger\right)=\frac{1}{2}\left(\Pi_x^2+\Pi_y^2\right). \tag{4.139}$$

从而我们就可以把系统的哈密顿量 (4.135) 重写成标准的谐振子形式

$$H=\hbar\omega_B\left(a^\dagger a+\frac{1}{2}\right), \tag{4.140}$$

式中 $\hbar\omega_B=(\frac{\hbar}{l_B})^2/m$，从而 $\omega_B=eB/m$, 称为回旋频率，因为它就是经典物理里电荷受磁场洛伦兹力而作回旋运动的角频率。

因此，由单自由度线性谐振子一节的推导可以知道，产生湮灭算符的代数关系将意味着哈密顿量 (4.140) 的本征值是量子化的，从而系统的能量是量子化的，为

$$E_n=\hbar\omega_B\left(n+\frac{1}{2}\right). \tag{4.141}$$

这就是朗道能级。很显然，朗道能级的能级差 $\hbar\omega_B$ 正比于磁场强度，磁场不强朗道能级的量子化效应就无法体现出来，这就是量子霍尔效应为什么需要强磁场的原因之一。

为了进一步看清楚朗道能级的简并，下面我们来求解一下定态波函数。我们首先关心的是基态波函数 $\psi_0=\langle x,y|0\rangle$。由基态的定义，它得满足微分方程

$$a\psi_0=0. \tag{4.142}$$

为了求解上面这个方程,我们需要选择一个确定的规范,我们将选取 $(A_x,A_y)=\frac{B}{2}(-y,x)$ 的对称规范。从而 $\Pi_x=-i\hbar D_x=-i\hbar(\partial_x-i\frac{e}{\hbar}A_x)=-i\hbar(\partial_x+i\frac{1}{2}y/l_B^2)$。类似的，$\Pi_y=-i\hbar D_y=-i\hbar(\partial_y-i\frac{1}{2}x/l_B^2)$。则，根据产生湮灭算符的定义式 (4.138)，我们有

$$\frac{1}{\sqrt{2}}(\frac{\hbar}{l_B})a=-i\hbar\left[\frac{1}{2}(\partial_x+i\partial_y)+(x+iy)\frac{1}{4l_B^2}\right]. \tag{4.143}$$

在 $x-y$ 平面内引入复坐标 $w=x+iy$、$\overline{w}=x-iy$。则有

$$a=-i\sqrt{2}l_B\left(\overline{\partial}+\frac{w}{4l_B^2}\right), \tag{4.144}$$

式中 $\overline{\partial}=\frac{\partial}{\partial\overline{w}}=\frac{1}{2}(\partial_x+i\partial_y)$。同样的，我们也有

$$a^\dagger=-i\sqrt{2}l_B\left(\partial-\frac{\overline{w}}{4l_B^2}\right), \tag{4.145}$$

式中 $\partial=\frac{\partial}{\partial w}=\frac{1}{2}(\partial_x-i\partial_y)$。

因此，我们要求解的基态波函数方程 (4.142) 就等价于

$$\left(\overline{\partial}+\frac{w}{4l_B^2}\right)\psi_0(w,\overline{w})=0. \tag{4.146}$$

令 $\psi_0(w,\overline{w})=e^{-\frac{|w|^2}{4l_B^2}}f(w,\overline{w})$，则由这个 ψ_0 的方程可以得到

$$\overline{\partial}f(w,\overline{w})=0, \tag{4.147}$$

从而 $f(w,\overline{w})$ 是一个任意的全纯函数，记为 $f(w)$。因此，朗道能级的基态波函数为

$$\psi_0(w,\overline{w})=f(w)e^{-\frac{|w|^2}{4l_B^2}}. \tag{4.148}$$

由于任意全纯函数都可以给出一个基态，所以朗道能级的基态高度简并。

将产生算子 $a^\dagger$ 不断作用在基态 ψ_0 上就可以得到其它的激发态。由产生算子的表达式 (4.145) 可知，朗道第 n 能级的波函数 $\psi_n(w,\overline{w})$ 将为

$$\begin{aligned}\psi_n(w,\overline{w}) &\sim \left(\partial - \frac{\overline{w}}{4l_B^2}\right)^n \left(e^{-\frac{|w|^2}{4l_B^2}} f(w)\right) \\ &= e^{-\frac{|w|^2}{4l_B^2}} \left[e^{+\frac{|w|^2}{4l_B^2}} \left(\partial - \frac{\overline{w}}{4l_B^2}\right) e^{-\frac{|w|^2}{4l_B^2}}\right]^n f(w) \\ &= e^{-\frac{|w|^2}{4l_B^2}} \left(\partial - \frac{\overline{w}}{2l_B^2}\right)^n f(w). \end{aligned} \tag{4.149}$$

可见任意朗道能级波函数中都包含一个和基态波函数一样的未定全纯函数 $f(w)$，因此任意朗道能级都高度简并，并且简并度和基态一模一样。

那么，基态朗道能级的简并度是多少呢？为了确定这一点，我们可以将全纯函数 $f(w)$ 展开成级数形式 $f(w)=\sum_{m=0}^{+\infty} c_m w^m$，很显然，这个级数的每一项都给出一个基态，因此任意基态应该记成

$$\psi_{0,m} \sim w^m e^{-\frac{|w|^2}{4l_B^2}}. \tag{4.150}$$

实际上，$\psi_{0,m}$ 同时是角动量算符 L_z 的本征态，本征值为 $m\hbar$。这是因为，在 2 维复坐标中，L_z 可以写成 $L_z=\hbar(w\partial - \overline{w}\overline{\partial})$(前面在处理氢原子时，我们实际上已经知道这个结果了，只不过在那里，w 记作 x_+，$\overline{w}$ 记作 x_-)。很容易验证 $(w\partial - \overline{w}\overline{\partial})\psi_{0,m} = m\psi_{0,m}$，所以 $\psi_{0,m}$ 的角动量量子数为 m。

波函数 $\psi_{0,m}$ 呈环状，其概率密度 $|\psi_{0,m}|^2$ 最大的地方为一个半径 r_m 的圆周，

$$r_m = \sqrt{2m} l_B. \tag{4.151}$$

因此 $\psi_{0,m}$ 包围的面积为 $\pi r_m^2 = 2\pi l_B^2 m$。当然，$\psi_{0,m-1}$ 的半径要小一些，它所包围的面积是 $2\pi l_B^2(m-1)$。这相邻两个态的面积差为 $2\pi l_B^2$, 这个面积上的磁通正好是一个磁通量子 Φ_0。因此，这也就是说，面积每增加一个磁通量子就增加了一个量子态，即每一个磁通量子对应一个量子态。所以简并基态的总数目其实就等于 $x-y$ 平面上能通过的磁通量子数 N_Φ，假设整个材料平面的面积是 S，由于它的每一个 $2\pi l_B^2$ 面积的磁通为一个磁通量子，所以

$$N_\Phi = \frac{S}{2\pi l_B^2} = \frac{eBS}{2\pi\hbar}. \tag{4.152}$$

显然，每一个朗道能级都高度简并！而电子的量子态要用两个量子数 n 和 m 来标记，其中 n 标记朗道能级，m 标记轨道角动量量子数，这样的量子态可以记作 $|n,m\rangle$，相应的波函数可以记为 $\psi_{n,m}$，n 相同 m 不同的所有量子态都是简并的。

值得注意的是，朗道能级是指单个电子的能级，而不是整个两维材料的能级，前面所说的基态简并也是指单电子的基态简并，而不是整个两维材料的基态简并。整个两维材料里面有大量的电子，它的最低能态不见得简并。比方说假如这些电子间没有相互作用，因此每一个电子都独立的在填充朗道能级，考虑到泡利不相容原理，一个单电子量子态上只

能填充一个电子，假如所有的电子刚好填充完某一个朗道能级，那这就没有任何自由度留下了，填充方法唯一，因此整个材料系统的基态 (最低能态) 反而没有任何简并。一块材料的基态是否有简并是一种很重要的特性，因为如果一个没有对称性保护的系统基态有简并，那它就很可能有长程量子纠缠，有拓扑序 (否则的话这种简并早就被相互作用微扰破坏掉了)，这是人们正在研究的前沿课题之一。

4.5.2 习题

1. 请在 $A_x = -By, A_y = 0$ 的规范下重新求解朗道能级的本征波函数。由于这一规范有明显的沿着 x 轴的平移不变性，所以用来讨论长方形的两维材料薄片更为方便一些。现在假设一个这样的薄片长为 L_x 宽为 L_y，请用这个新规范下求解出来的波函数重新讨论朗道能级的简并度，并和文中对称规范下的结果相比较。

4.6 量子霍尔效应

4.6.1 经典霍尔效应

1879 年 Edwin Hall 做了一个经典的实验，它让电流沿着 x 方向通过一个两维薄片 (即片的厚度相对可以忽略)，然后在垂直于薄片平面的 z 方向加上一个均匀磁场 B，如图 (4.4) 所示。霍尔发现，在薄片的 y 方向出现了电势差 V_H，而且 V_H 与电流强度 I 成正比。

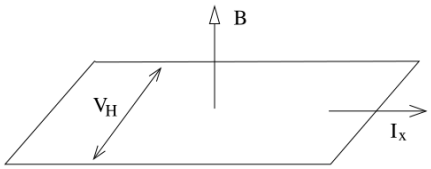

Figure 4.4: 经典霍尔效应

这个实验的不寻常之处在于电压和电流不在一个方向上，因此意味着出现了一种新的横向电阻 ρ_{xy}，它把 x 方向的电流和 y 方向的压降 V_H 联系起来，$V_H = \rho_{xy} I$。或者用电导率 $\sigma_{xy} = 1/\rho_{xy}$ 来写即是

$$I = \sigma_{xy} V_H. \tag{4.153}$$

这就是经典霍尔效应，它的解释并不难。实际实验中载流子往往是电子，它们的电荷 e 当然是负的，不过下文的推导中我们都默认 $e > 0$ 来进行分析，$e < 0$ 情形的推导我们留给读者。经典霍尔效应的原因是，电子在磁场中受洛伦兹力的作用会发生偏转，从而在材料薄片 y 方向两个平行于 x 轴的侧面上就会积累正负电荷 (注意整个材料是电中性的，载流子的电荷要与背景离子的电荷相平衡)，从而在 y 方向产生一个电场 E。当达到稳定时，电子受到的洛伦兹力会和这个电场力相平衡，从而有

$$vB = E, \tag{4.154}$$

式中 v 是电子沿着 x 轴的漂移速度。另外，我们知道 x 方向的电流密度 J_x 为

$$J_x = nev, \tag{4.155}$$

式中 n 为电子数密度，即每单位体积内的电子数。因此我们即有

$$J_x = \frac{ne}{B}E, \tag{4.156}$$

将这个式子沿着 y 轴积分 (注意 x 方向的电流密度按照定义是每单位横截线上的电流强度，因此 J_x 沿着横向的 y 轴积分就得到电流强度 I)，就会得到 $I = \frac{ne}{B}V_H$，这也即是说

$$\sigma_{xy} = \frac{ne}{B}. \tag{4.157}$$

从而我们知道横向电阻 $\rho_{xy} \propto B$。

我们知道，通常的电阻会引起能量的耗散，也就是发热。这种能量耗散主要是因为传导电子不断被背景离子散射而引起的，因为这种散射的结果是使得传导电子的运动方向发生随机的改变，从而将定向运动的能量转化为无规热运动的能量。很显然，这样的电阻将反比于电子在两次散射之间的平均飞行时间 τ(也称为散射时间)，或者说通常的电导应该正比于 τ。但是从公式 (4.157) 我们清楚地看到，霍尔效应的横向电导 σ_{xy} 与散射时间 τ 无关。这就意味着，霍尔电阻根本不会带来能量的耗散！因此如果一块材料通常的电阻 (我们称之为纵向电阻，电阻率记为 ρ_{xx}) 为零，而只有霍尔电阻的话，那它根本就不会消耗能量，因此也就不需要持续的能量输入！

在霍尔效应中，洛伦兹力不做功，因此当然不会输入能量，而横向的电场由于和电子运动的 x 方向相垂直，因此做功也为零，也不会输入能量。不过，在经典霍尔效应中，纵向电阻实际上并不为零，沿着霍尔薄片的纵向 x 方向实际上也要加一个电场，这个纵向电场的电能输入最终转化为了纵向电阻的能量耗散。

而且，似乎我们可以论证，在霍尔效应中纵向电阻不可能为零。这是因为，考虑到纵向和横向两种不同的电阻以后，电阻率就应该推广成一个 2×2 的矩阵，同样电导率也应该推广成一个 2×2 的矩阵，而电流密度和电场强度之间的关系式就应该推广成

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} \rho_{xx} & \rho_{xy} \\ -\rho_{xy} & \rho_{yy} \end{pmatrix} \begin{pmatrix} J_x \\ J_y \end{pmatrix}, \quad \begin{pmatrix} J_x \\ J_y \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{yy} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}. \tag{4.158}$$

很显然，电导率矩阵和电阻率矩阵互为逆矩阵。通过求逆矩阵，我们可以得到

$$\sigma_{xx} = \frac{\rho_{xx}}{\rho_{xx}^2 + \rho_{xy}^2}. \tag{4.159}$$

很显然，如果横向电阻率是零，那我们就有通常的 $\sigma_{xx} = 1/\rho_{xx}$。但是我们已经看到了，在霍尔效应中，横向电阻 $\rho_{xy} \neq 0$。根据 (4.159) 式，如果这时候 $\rho_{xx} = 0$, 那么将有 $\sigma_{xx} = 0$，即纵向电导也等于 0。因此这时候沿着纵向材料实际上是绝缘的，根本不会有纵向电流。而在霍尔效应中，沿着纵向 x 轴当然有电流，这就说明 $\sigma_{xx} \neq 0$，从而 $\rho_{xx} \neq 0$。

然而，下面我们将会看到，在量子霍尔效应中，一切都将变得不同！

4.6.2 整数量子霍尔效应

前面我们看到，在经典霍尔效应中，横向电阻 $\rho_{xy} \propto B$。然而 1980 von Klitzing 做了一个实验，它在很低的温度 ($\sim 1K$) 下测量了一个比较干净的样品的霍尔电阻 ρ_{xy}，实验结果如图 (4.5) 所示。从图中可以看到，当磁场比较弱时，的确有 $\rho_{xy} \propto B$ 的线性关系。但是，

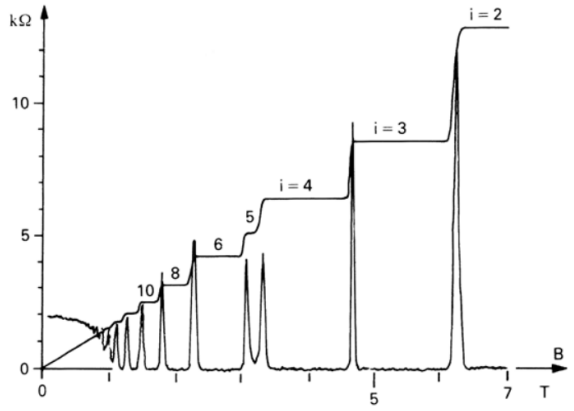

Figure 4.5: 整数量子霍尔效应。图中的两条曲线分别对应横向电阻率 ρ_{xy} 和纵向电阻率 ρ_{xx}。

当磁场变得很强时，ρ_{xy} 出现了量子化的台阶，也即是说，当磁场在一些合适的区域内变动时，ρ_{xy} 可以保持在一个平台值不变，然后随着磁场进一步增加到一定时候，它再迅速地跳到下一个平台。而且仔细的研究发现，ρ_{xy} 的各个平台值刚好与正整数 $\nu \in \mathbb{Z}$ 成反比，或者，用霍尔电导 σ_{xy} 来说即是

$$\sigma_{xy} \propto \nu. \tag{4.160}$$

这是一种典型的量子化现象，但是，霍尔电阻或者霍尔电导的量子化是非常不同寻常的，因为它们并不是一个微观系统的物理量，而是一个宏观系统的物理量，其中涉及到了大量电子的行为，所以这是一种非常独特的现象，问题是这一现象的物理解释是什么呢?

很快人们就意识到，霍尔电导 σ_{xy} 的量子化可以用朗道能级来初步解释。首先，在很低的温度下，量子力学效应将变得重要，因此通过霍尔薄片的电子气要用量子力学来描述。其次，假设忽略电子间的相互作用，那就可以近似将薄片中的每一个电子看成是独立的，而单个电子在磁场中的量子力学行为就是我们上一小节研究过的朗道能级。因此我们可以将霍尔薄片中的电子气的量子态看成是在填充朗道能级，从低能级逐渐往高能级填充，由于泡利不相容原理，每一个单电子量子态上只能填充一个电子，因此每一个朗道能级都可以填充 $N_\Phi = \frac{eBS}{2\pi\hbar}$ 个电子，假设每一个填充满的朗道能级都会贡献一个单位的霍尔电导的话，那如果有 ν 个朗道能级被填充，最后的霍尔电导就将是 ν 单位，这样就可以解释 $\sigma_{xy} \propto \nu$ 的量子化现象。

而且以上并不是一种牵强的解释。为了看清楚这一点，我们可以将经典理论中的霍尔

电导公式 (4.157) 作如下改写

$$\sigma_{xy} = \frac{neS}{BS} = \frac{eN}{\Phi} = \frac{eN}{N_\Phi \Phi_0} = \frac{e^2}{2\pi\hbar}\frac{N}{N_\Phi}. \tag{4.161}$$

式中 S 是霍尔薄片的面积，因此 $nS = N$ 是霍尔片内的电子总数，Φ 是霍尔片上的总磁通，它当然等于 $N_\Phi\Phi_0$。由于每个朗道能级有 N_Φ 个量子态，所以上面公式最终的 N/N_Φ 就是霍尔片内的电子总共能够填充的朗道能级数，也就是 ν，从而我们就得到

$$\sigma_{xy} = \frac{e^2}{2\pi\hbar}\nu. \tag{4.162}$$

正好解释了 von Klitzing 发现的霍尔电导量子化现象。而且，仔细检验人们还发现，公式 (4.162) 给出的电导量子单位 $\frac{e^2}{2\pi\hbar} = \frac{e^2}{h}$ 也和实验值完全吻合。

由于朗道能级的简并度 $N_\Phi \propto B$, 所以当磁场较小时，每个朗道能级的态数目就比较小，这样电子就可以填充较多的朗道能级，每一个被填充的朗道能级都贡献一个 $\frac{e^2}{2\pi\hbar}$ 的量子化霍尔电导，从而磁场较小时的霍尔电导就较大，霍尔电阻 ρ_{xy} 就较小。随着磁场的增加，每个朗道能级上的态数目 N_Φ 会逐步增加，从而原来占据在更高朗道能级的电子就会跃迁到低朗道能级多出来的态上，这样被电子占据的朗道能级数就会变少，霍尔电导就变小，而霍尔电阻 ρ_{xy} 就会跳跃到更高的台阶。这和实验 (4.5) 也完全吻合。

但是，von Klitzing 的实验除了发现霍尔电导的量子化以外，它还发现，在霍尔电阻的每一个平台区，纵向电阻 $\rho_{xx} = 0$, 如图 (4.5) 中的另一条实验曲线所显示。但，之前在讨论经典霍尔效应时我们已经论证过了，这似乎是不可能的！因为这将意味着纵向电导也是零，即 $\sigma_{xx} = 0$, 从而意味着当磁场处在平台区的时候，霍尔片沿着纵向实际上是绝缘的，因此根本就不会有纵向电流。但纵向电流在实验中当然存在，我们就是利用纵向电流强度与横向电压的比值来测量霍尔电导 σ_{xy} 的。看起来我们的结果自相矛盾了。

对上面这个矛盾的解答揭示了量子霍尔效应的另一个重要特性。结果人们发现，的确，霍尔薄片的体内是绝缘的，没有纵向电流，纵向电流实际上沿着霍尔片的边缘流动，是一种新奇的边缘态。我们可以用如图 (4.6) 所示的经典图像来直观地理解这个物理机制。如

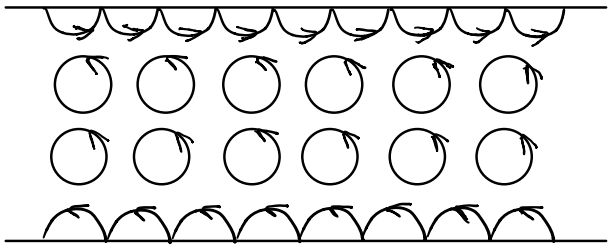

Figure 4.6: 量子霍尔效应的手性边缘态。

这幅图所显示的，在磁场中，电子将作回旋运动，在霍尔薄片的体内，相邻电子的回旋运动刚好抵消，所以宏观上并没有电流，因此是绝缘的。但是，在靠近霍尔薄片的边缘处，电子的回旋运动会被边缘挡住，从而不能完成一个完整的回旋，这样的电子碰到边缘以后只能反弹，但是由于电子的回旋方向固定，所以最终靠近边缘的这个电子就会像图 (4.6) 中

所示的那样运动。从这幅图中我们可以看得很清楚，从宏观上看，上边缘处的电子只能向右行，而下边缘处的电子只能向左行！这些右行和左行的电子就分别构成上下边缘上的边缘电流。所以在量子霍尔效应中，纵向电流沿着边缘流动。从图 (4.6) 中我们可以看得很清楚，上下边缘的电流流向相反。如果这两个边缘没有电势差，那这两个流向相反的电流强度将刚好相等，从而纵向的总电流就是零，但是如果上下边缘有一个电势差，那它们的电流强度就不能完全抵消，从而会有一个净的纵向电流，它的强度正比于横向的这个电势差，比例系数就是霍尔电导 σ_{xy}。

边缘电子只能沿着一个方向运动是一个很奇妙的性质。因为通常导线中的电子不是这样的，通常电子可以前进也可以后退，它的运动方向可以向正方向，也可以向负方向。这样的电子碰到障碍物 (比方说碰到背景离子) 的时候就可以被散射回来，从而失去原来的运动方向，因此它的能量就被耗散掉了，这种能量耗散的宏观表现就是一个非零的纵向电阻。但是，由于霍尔薄片边缘上的电子只能沿着一个方向运动，这样，即使它碰到障碍物，它也不能被散射回反方向，它只能绕过去，如图 (4.7) 所示。因此，量子霍尔效应的这种

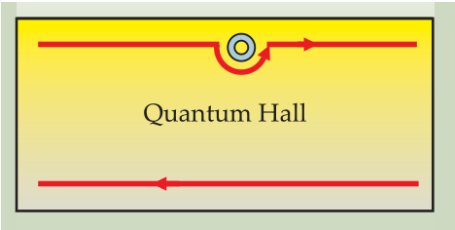

Figure 4.7: 手性电流不会有能量耗散。图片来自论文 “The quantum spin Hall effect and topological insulators” , Xiao-Liang Qi and Shou-Cheng Zhang

边缘态在输运电流的时候就没有能量的耗散，从而纵向电阻就为零。这就解释了实验中观测到的 $\rho_{xx} = 0$。

由于量子霍尔效应的边缘态只能沿一个方向运动，所以人们常常称之为手性边缘态。手性边缘态有不会耗散能量的奇妙性质，然而它并不能在通常的导线中实现。因为有一个定理叫 Nielsen-Ninomiya theorem 告诉我们，通常导线中的电子一定既有左行波态又有右行波态，无法限制它只取一种手性 (用量子场论的话来说，是因为否则就会有手征反常，它会破坏规范对称性)。不过，这与量子霍尔效应中出现手性边缘态并不矛盾，因为在量子霍尔效应中这种一维的手性边缘态是出现在一个两维系统的边缘上，而这是允许的 (因为霍尔薄片体内的物理也是反常的，而这种反常刚好可以和边缘态的手征反常抵消掉，这也是为什么量子霍尔效应中必须有手性边缘态的原因之一，因为否则体内的规范不变性就会因为反常而破坏掉)。

4.6.3 量子力学的分析

然而以上对整数量子霍尔效应的分析应用了很多经典物理的东西。比方说，我们在解释霍尔电导的量子化时其实是从经典的霍尔电导公式出发的。而我们在解释手性边缘态时也用了经典物理的图像。完全用量子力学该如何分析上面的问题呢?

为了更好地应用我们在朗道能级那一节中的量子力学结果来分析整数量子霍尔效应。现在假设我们的样品霍尔薄片是如图 (4.8)(a) 所示的圆环 (这种具有旋转对称性的情形能够使得我们更好地利用朗道能级一节中求出的角动量本征态 $\psi_{n,m}$)。除了和磁场相互作用以外，样品中的电子还会感受到一个由背景离子和其它电子共同产生的势场，尤其是在样品边缘处，为了阻止电子从边缘逸出样品，那里一定有一个很高的势垒，最终势能作为径向坐标 r 的函数 $V(r)$ 大致会如图 (4.8)(b) 所示。考虑到这个势能 $V(r)$ 以后朗道能级就需

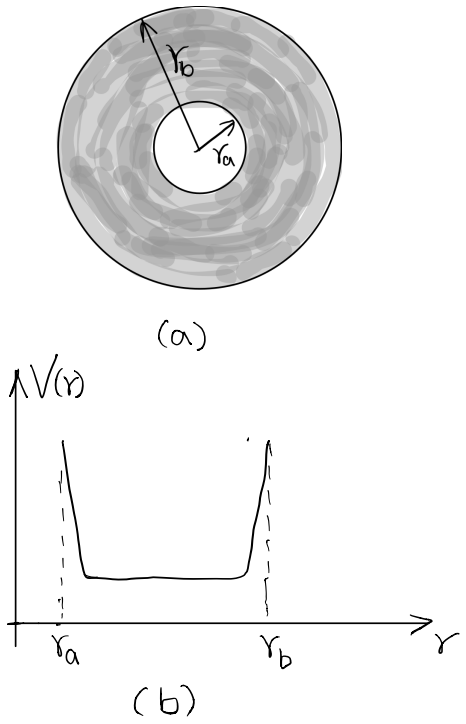

Figure 4.8: (a) 圆环状霍尔片，磁场垂直于这个薄片。(b) 样品内的势能曲线。

要作适当的修正。不过我们也不必重新求解带势能 $V(r)$ 的带电粒子定态薛定谔方程，简单的分析就能告诉我们大致应该怎样修正。首先，由 (4.151) 式可以知道，一个绕圆心的轨道角动量为 $m\hbar$ 的电子大体上会分布在半径 $r_m = \sqrt{2m}l_B$ 处，这个地方的势能为 $V(r_m)$，它要加到原来的朗道能级上去，所以一个 $|n,m\rangle$ 态的电子，其能量 $E_{n,m}$ 大约为

$$E_{n,m} = \hbar\omega_B(n+\frac{1}{2}) + V(r_m). \tag{4.163}$$

我们可以将 $E_{n,m}$ 重写成 $E_n(r_m)$，并将之看成是不同朗道能级的能量曲线，大致如图 (4.9) 所示。我们可以看到，这些能量曲线在样品的内外两个边缘处相互趋近。而且，由于 m 和 r_m 之间的一一对应关系，同一个朗道能级，不同 m 的那些量子态就落在不同的径向坐标处。

由于能量最低原理，电子当然是按从低能到高能的顺序填充这些能量曲线的。填充进来的电子的最高能量就叫费米能，记作 E_F。也即是说，能量比费米能更低的量子态都已经填充了电子，如图 (4.9) 所示。在这幅图中，我们假定费米能是落在两个朗道能级之间，

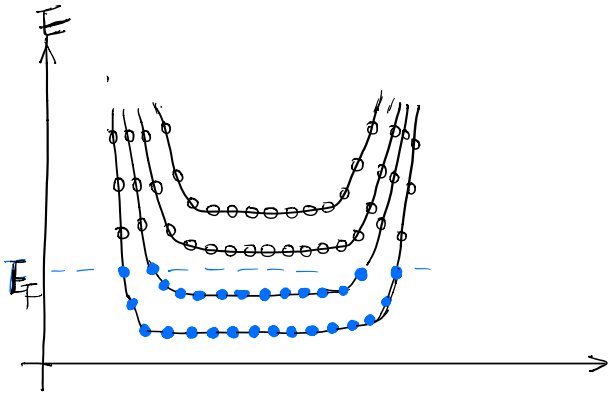

Figure 4.9: 不同朗道能级的能量曲线，水平坐标是 r，每一条曲线代表一个朗道能级。每一个点都代表一个量子态，同一朗道能级上不同的点有不同的 m 量子数，相应即有不同的径向坐标 r_m。蓝色的实心点代表填充了电子的量子态，空心点代表未被占据的量子态。

下面我们的推导也是基于这种情况。但是，如果费米能落在某个朗道能级的底部，那情况就会比较复杂，因为它可能意味着最上面的这个朗道能级只被填充了一部分，如图 (4.10) 所示，这种情况我们后面再来讨论。

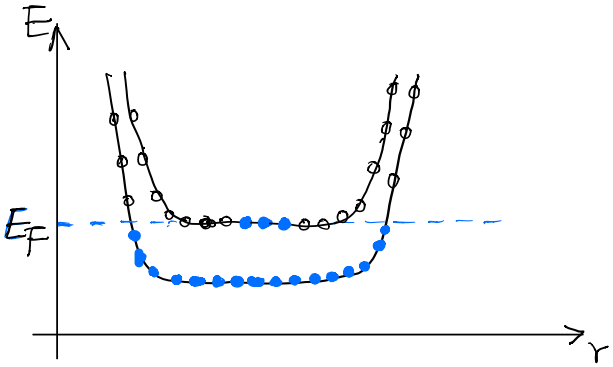

Figure 4.10: 最上面的朗道能级只部分填充的情况。当然这朗道能级上面还有更多完全未填充的朗道能级。

现在，在样品圆环上取极坐标，以 ϕ 表示极角，我们来计算整个样品围绕圆环环流的电流强度 I_ϕ。首先，样品内一个在 T 秒内围着圆心转一圈的电子贡献的电流强度将是 $I_\phi = e/T$，用电子的旋转角速度 ω 来表达即是 $I_\phi = \frac{e}{2\pi}\omega$。不过在量子力学层次，角速度 ω 将是一个算符，我们计算的其实是它的期望值。因此，一个填充在 $|n,m\rangle$ 态上的电子贡献的电流强度将是 $I_\phi = \frac{e}{2\pi}\langle n,m|\omega|n,m\rangle$。由于角速度是 ϕ 角对时间的导数，因此我们可以记 $\langle n,m|\omega|n,m\rangle = (\dot{\phi})_{n,m}$。另外，在经典物理中角动量 L_z 和角度 ϕ 是一对正则变量，满足正则方程 $\dot{\phi} = \frac{\partial H}{\partial L_z}$($H$ 为哈密顿量)。而哈密顿量就是能量，而且一个 $|n,m\rangle$ 态的电子，其角动量 $L_z = m\hbar$，能量为 $E_{n,m}$，因此可以想见，在量子力学层次，近似地应该有 $(\dot{\phi})_{n,m} = \frac{1}{\hbar}\frac{\partial E_{n,m}}{\partial m}$。所以，一个填充在 $|n,m\rangle$ 态上的电子贡献的电流强度为 $I_\phi = \frac{e}{2\pi}\frac{1}{\hbar}\frac{\partial E_{n,m}}{\partial m} = \frac{e}{2\pi\hbar}\frac{\partial E_{n,m}}{\partial m}$。而每一

个已经填充了电子的量子态都会贡献类似的电流强度，所以最终总的电流强度就是

$$I_\phi = \frac{e}{2\pi\hbar}\sum_{(n,m)\in F}\frac{\partial E_{n,m}}{\partial m}. \tag{4.164}$$

式中 F 表达已经填充了电子的量子态集合。

为了计算出 (4.164) 式，现在我们来考察如上面图 (4.9) 所示的情形，假设费米能落在第 ν 个朗道能级和第 $\nu+1$ 个朗道能级之间。公式 (4.164) 实际上是对 m 的差分进行求和，我们近似地可以替换成积分，从而将公式 (4.164) 改写成

$$\begin{aligned} I_\phi &= \frac{e}{2\pi\hbar}\sum\int_{(n,m)\in F} dm\frac{\partial E_{n,m}}{\partial m} \\ &= \frac{e}{2\pi\hbar}\sum\int_{(n,r_m)\in F} dr_m\frac{\partial E_n(r_m)}{\partial r_m}. \end{aligned} \tag{4.165}$$

式中的第二个等号是从 m 到 r_m 的一个变量代换。

显然，对于每一个被填充的朗道能级，公式 (4.165) 关于 r_m 的积分结果都是内外两个边缘上的能量差，当然也就是内外两个边缘上的费米能之差 ΔE_F，它当然和 m 没什么关系，所以 (4.165) 式其实就是

$$I_\phi = \frac{e}{2\pi\hbar}\sum_{n=0}^{\nu-1}\Delta E_F = \frac{e}{2\pi\hbar}\nu(\Delta E_F). \tag{4.166}$$

对于如图 (4.9) 所示的这种情况，最后结果实际上是零，因为两个边缘上的费米能是一样的。不过，如果内外两个边缘上有电势差，那由于不同的电势能，填充在不同边缘上的电子最高能量也会不同，也就是说，这时候内外两个边缘上的费米能就会有差别，如图 (4.11) 所示，这个费米能的差别当然就等于电子的电势能之差 eV_H，V_H 就是内外两个边缘上的电势差。因此这时候我们就有

$$I_\phi = \big(\frac{e^2}{2\pi\hbar}\nu\big)V_H. \tag{4.167}$$

这正好是我们想要推导出来的霍尔电导公式。不过这一次，我们基本上是用量子力学推导出来的。

不光如此，公式 (4.165) 还告诉我们一些额外的信息。由于电流强度是电流密度沿着横截方向的积分，即 $I_\phi = \int dr J_\phi$，与 (4.165) 式比较，我们可以看到，第 n 朗道能级在半径 r 处贡献的电流密度 $J_{\phi,n}(r)$ 其实应该是

$$J_{\phi,n}(r) = \frac{e}{2\pi\hbar}\frac{\partial E_n(r)}{\partial r}. \tag{4.168}$$

但是，在前面的图形中我们清楚地看到，能量曲线 $E_n(r)$ 在样品体内是平坦的，对 r 的导数是零，因此，在样品体内将没有环绕样品圆环的电流！但是，在样品边缘附近，能量曲线 $E_n(r)$ 对 r 的导数不等于零，所以样品边缘才有电流！而且，从能量曲线可以清楚地看到，在外边缘，$\frac{\partial E_n(r)}{\partial r} > 0$, 而在内边缘 $\frac{\partial E_n(r)}{\partial r} < 0$, 因此公式 (4.168) 就告诉我们，内外两个边缘的电流方向相反，一个顺时针，一个逆时针，如图 (4.12) 所示，这当然就是我们上一小节通过经典图像发现的手性边缘态。

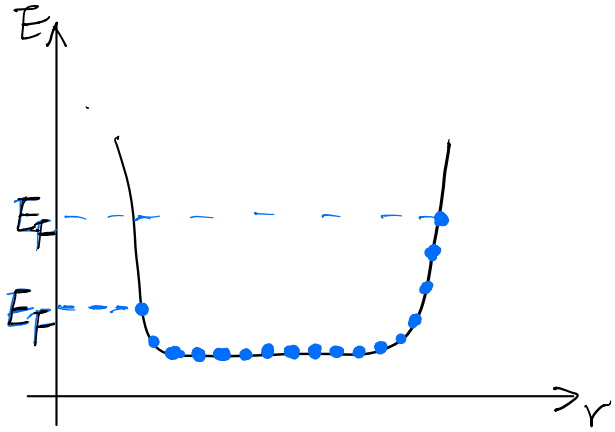

Figure 4.11: 电势差导致内外两个边缘的费米能有差别。注意我们只示意性地画了一条朗道能级，并且我们只画了占据态，没画空态。

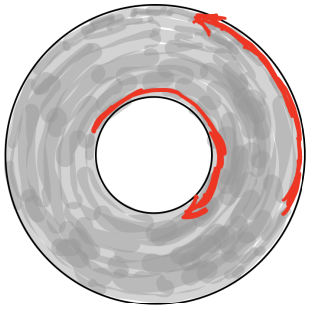

Figure 4.12: 量子霍尔效应的手性边缘态

4.6.4 无序—最终的解释

但是，以上解释依然有问题。问题就在于朗道能级的能量曲线在霍尔片体内其实并不是平坦的。比方说，如果内外两个边缘有电势差，那在霍尔片体内沿着径向方向就会有一个横向电场，这个电场会改变体内电子的电势能，考虑到这个电势能以后，朗道能级的能量曲线就会是倾斜的，如图 (4.13) 所示。这时候如果体内填充有电子的话，那它们的 $\frac{\partial E_n(r)}{\partial r}$ 就不等于零，根据 (4.168) 式，这些电子就会在体内产生电流，这在解释霍尔电阻的平台区时就会遇到困难。

为了解释清楚问题，我们不妨假设磁场从一个非常强的磁场逐渐降低，也就是说，我们要从右往左读整数量子霍尔效应的实验结果 (4.5)。在霍尔电阻的 ν 平台上，前 ν 个朗道能级都已经填充满了，这时随着磁场的减小，每个朗道能级的态数目 N_Φ 就在减少，这样原来填充在前 ν 个朗道能级上的电子就会有部分多出来，这些多出来的电子就会开始填充第 $\nu+1$ 个朗道能级，当然，这时候这第 $\nu+1$ 个朗道能级是部分填充的，填充情况大体如图 (4.13) 所示，随着磁场的连续减少，它会逐渐填充得越来越多。但是，霍尔电阻实验测量的平台区意味着，这第 $\nu+1$ 个部分填充的朗道能级并没有参与导电，否则由于导电的电子数逐渐变多，霍尔电导就会连续变大，霍尔电阻就会连续变小，那就根本不会

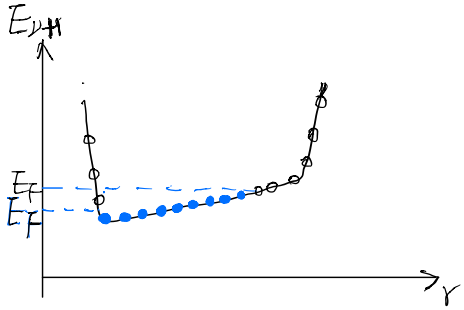

Figure 4.13: 电场中的朗道能级以及部分填充

保持在一个平台值不变。但是，根据刚才我们的理论分析，图 (4.13) 中所示的部分填充的第 $\nu+1$ 个朗道能级是要参与导电的，而且是在霍尔薄片的体内导电。由于这些电子是在体内导电，而不是处于手性边缘态，那它们当然就要和背景离子散射从而损耗能量，因此理论上纵向电阻也不能为零。很显然，这些理论分析结果都和实验观测相矛盾。

解决问题的关键在于，实际上，朗道能级的能量曲线在霍尔片体内不仅仅不平坦，而且是相当不平坦，相当崎岖。这是因为，实际的霍尔片不可能完全纯净 (如果太纯净那反而观察不到霍尔电阻的平台)，它里面有各种杂质，我们统称为无序。在效果上，这些无序就相当于一个各点势能取值随机分布的随机势，这样的势能当然会非常崎岖，从而使得整个霍尔片内部布满了各种坑坑洼洼，各种势能的“山峰”和“山谷”。这样的系统当然没有绕 z 轴的旋转对称性，因此前面的角动量量子数 m 就不再是一个好量子数，不同的量子态不再对应不同的径向坐标，我们也就不再能将分析简化为沿着径向坐标的一维分析。不过，由于每一个磁通量子对应一个量子态，我们很容易想到，薄片上每一个面积为 $2\pi l_B^2$ 的极小的局部小圆盘，都可以用来标记一个量子态。由于这些小圆盘极小，可以近似看成一个圆点，因此 $x-y$ 平面的每一个圆点都相应于一个量子态，它的能量是该点处的随机势能加上标准的 $\hbar\omega_B(n+\frac{1}{2})$，由于势能很崎岖，所以不同点的量子态通常并不简并。这样，加上这个崎岖的随机势能以后，前面朗道能级能量曲线的概念就应该推广成两维的能量曲面，一个朗道能级就对应一个这样的崎岖能量曲面。当然，在霍尔片边缘处，分析基本上还和前面一样，因为边缘一定会有一个很高的势垒 (否则电子就从样品边缘逃逸出去了)。

在图 (4.14) 中，我们示意性地画出了一个长方形霍尔薄片某一个朗道能级的等能量曲线 (类似于地图中的等高线)，等能量线的密度反映能量曲面的梯度，等能量线越密的地方，能量梯度就越大。图中的 + 号标记的是能量的“山峰”，− 号标记能量的“山谷”。同一个朗道能级的不同量子态就以一个个面积为 $(2\pi l_B^2)$ 的非常小的小圆点的形式分布在这个能量“地形”中。当然，这个图形是示意性的，真实霍尔片中的能量地形可能要崎岖得多，“山峰”和“山谷”都会很多。

以上的图像并不完全。因为占据在一个局部小圆点上的电子实际上还会跳到邻近的小圆点上去，从而产生运动。这些电子如何运动呢？为了搞清楚这个问题，我们再来仔细观察一下前面在圆周对称的情形中得到的电流密度公式 (4.168)，我们发现公式中的 $\frac{\partial E_n(r)}{\partial r}$ 实际上就是能量的梯度，而电流密度的方向 ϕ 实际上就是和这个梯度垂直的方向。因此，我

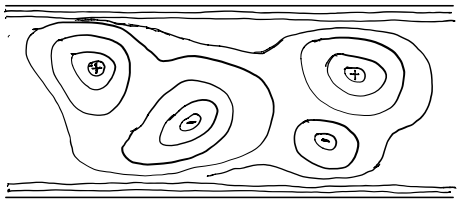

Figure 4.14: 霍尔薄片某一个朗道能级的能量“地形”图

们可以这样来推广电流密度 J_n 的公式，我们说: 在量子霍尔效应的一般情形中，n 朗道能级上占据在某小圆点上的电子贡献的电流密度 J_n 大小等于 $\frac{e}{2\pi\hbar}$ 乘以这点的能量梯度，并且电流沿着与梯度方向相垂直的方向流，也就是沿着等能量线流，而电流的正负号则由能量梯度的正负号决定。

当电子填充朗道能级的时候，如果它们填充到霍尔片的边缘，那由于边缘处有很高的能量梯度，所以就会产生较大的电流密度，又由于边缘处的等能量线都沿着霍尔薄片的边缘走，所以这些电流也沿着边缘流，而且，边缘处能量梯度的正负号是确定的，因此这些电流只有一个流向，这就是手性边缘电流。

当然，实际电子填充朗道能级的时候是按照能量最低原理进行的，也就是会先填充体内那些能量“山谷”。最先填在谷底的电子，由于谷底的能量梯度为零，所以它就待在谷底不动了。接着填进来的电子就会填在围绕谷底的能量稍高一些的等能量线上，上面我们已经分析过了，这些电子就会沿着这些等能量线围绕着谷底打转。所以这些电子实际上已经被局域在某个能量“山谷”的附近了，我们称之为局域态。局域态的电子运动范围很小，所以当然不能产生宏观的电流，因此它们实际上是不参与导电的。

类似的，在最后填充进霍尔片体内的电子已经没有地方可占据了，它就只能填充到能量“山峰”附近，同样，这些电子也只能围绕着某个能量“山峰”打转，从而实际上也被局域化了，也无法参与导电。

能参与导电的是在中间的时候填充进来的电子，这时候它将刚好填在能量的“海平面”附近，它也只能沿着等能量线运动，但是“海平面”附近的等能量线范围会很大，如图 (4.14) 所示，因此这些电子就可以在一个宏观的范围内运动，产生宏观的电流。这样的态就称为延展态。在霍尔片体内，只有延展态可以参与导电。局域态、延展态的讨论通常主要是指霍尔片体内的电子态，但是当然，我们也可以把手性边缘态看成是一种延展态。

现在我们就可以解释霍尔电阻的平台是怎么回事了。当电子填充满前 ν 个朗道能级开始填充第 $\nu+1$ 朗道能级的时候，由于最先填进来的那些电子都被能量“山谷”局域化了，根本不导电，所以它们对霍尔电导根本就没有贡献，从而虽然电子在持续填充，但是霍尔电导会保持不变，从而就出现一个平台。而且由于这些局域化的电子根本不导电，所以这时候纵向电导就是零，纵向电阻也就是零。霍尔片上的电流都来自于前 ν 个朗道能级的手性边缘态。

但是，如果电子填充到了第 $\nu+1$ 个朗道能级的“海平面”附近，那体内就会出现延展态，这些延展态有可能连接到手性边缘态上，从而使得手性边缘态可以被散射，产生能

量耗散，进而使得纵向电阻出现非零的跳变，这就是图 (4.5) 中所显示的纵向电阻在两个相邻平台过渡的时候出现的跳变。

当 $\nu+1$ 能级的“海平面”附近被填满以后，霍尔电阻实际上就已经跳到 $\nu+1$ 平台区去了，这时候接着填充进来的电子又是被能量“山峰”局域化的局域态，同样不导电，霍尔电阻就会保持在新的 $\nu+1$ 平台上。

至此，我们才算真正解释了量子霍尔效应的实验观测结果。

虽然有这么多复杂性，但每一个填充满的朗道能级贡献的霍尔电导依然是 $e^2/(2\pi\hbar)$。为了看清楚这一点，我们来计算第 n 朗道能级沿着纵向 x 方向的电流强度 $I_{x,n}$。正如我们分析过的，x 方向的电流密度 $J_{x,n}$ 将由能量曲面沿着 y 方向的梯度给出，$J_{x,n}=\frac{e}{2\pi\hbar}\frac{\partial E_n(x,y)}{\partial y}$，从而

$$I_{x,n}=\frac{e}{2\pi\hbar}\int dy\frac{\partial E_n(x,y)}{\partial y}. \tag{4.169}$$

很显然，对于每一个填满的朗道能级，这个式子对 y 坐标积分的结果都只依赖于霍尔片上下边缘的费米能之差，从而也就只依赖于上下边缘的电势差，而和能量曲线在霍尔片体内的具体细节没有关系，如图 (4.15) 所示。不管能量曲线在霍尔片体内多么崎岖，最终的结

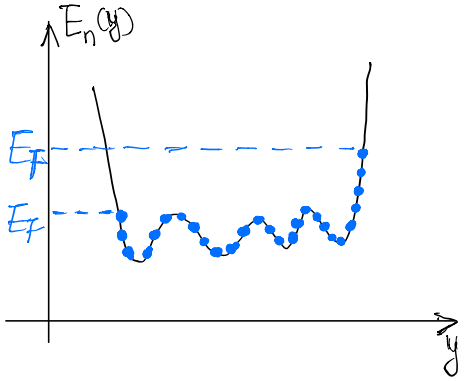

Figure 4.15: 能量曲面沿着 y 方向的横截图

果都是一样的，从而每一个填满的朗道能级都会贡献恒定不变的霍尔电导 $e^2/(2\pi\hbar)$。当然，从宏观上看，局域态根本不导电，对霍尔电导有贡献的是朗道能级中的延展态，其实主要是手性边缘态。

4.6.5 分数量子霍尔效应

整数量子霍尔效应已经够令人吃惊了，但是，1982 年 Tsui 和 Stormer 发现了更为令人吃惊的事情。它们也在低温和强磁场的条件下测量了霍尔电阻 ρ_{xy}，结果发现，不仅是在 ν 等于整数的时候会出现量子化平台，而且在 ν 等于分数时也会出现量子化平台，如图 (4.16) 所示。比方这幅图中最右边的那个霍尔电阻平台就对应 $\nu=1/3$(因此 $\rho_{xy}=3h/e^2$)。不仅如此，图中 $\nu=2/5$ 的地方也有一个很明显的量子化平台 (对应 $\rho_{xy}=2.5h/e^2$)。前面我们已经理解了 ν 等于整数时的量子化现象，它相应于前 ν 个朗道能级被填充满的情形。

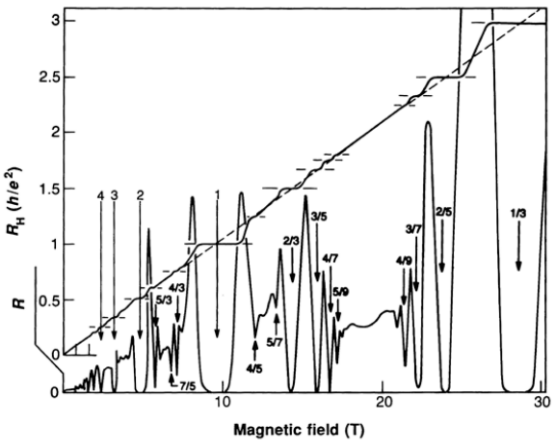

Figure 4.16: 分数量子霍尔效应

那么 $\nu = 1/3$ 是什么意思呢？它不应该对应最低朗道能级仅仅被填充 1/3 这种部分填充的情况吗？为什么也会出现量子化平台呢？

不仅如此，此后人们又不断观测到了更多的量子化平台，它们对应的 ν 都是有理数。这就带来两个问题：第一，为什么朗道能级部分填充也会出现量子化平台。第二，为什么刚好是有理数，既然部分填充能出现量子化平台，为什么没有 $\nu = 1/\sqrt{2}$ 的量子化平台 [1]。到今天为止，对这两个问题人们其实已经有了很多了解，尤其后一个问题，一定意义上其实可以从数学上证明，不过这种数学证明的深度超出我们这本书的范围了。

显然，为了理解这种分数量子霍尔效应，我们不再能忽略电子间的相互作用。否则的话前面我们理解整数量子霍尔效应时使用的单电子填充朗道能级的图像就依然完全成立，那我们就只能得到整数量子霍尔效应。为了解释朗道能级部分填充时出现的分数量子化现象，电子间的库伦排斥力将非常重要。实际上，如果人们用 $\frac{e^2}{4\pi\varepsilon_0 l_B} = \frac{\hbar\alpha c}{l_B}$ 来估算这个库伦排斥势能的话，就会发现它并不比朗道能级差 $\hbar\omega_B$ 小。所以某种意义上，人们反而是要解释为什么在讨论整数量子霍尔效应的时候可以忽略电子间的库伦排斥力。不过，由于单电子近似是凝聚态物理中的传统近似，它在固体能带论中取得了巨大成功，而且它的成立性可以用朗道的费米液体理论来解释 (后面的章节中我们会简单地讨论这个问题)。所以分数量子霍尔效应的发现其实是使得人们开始意识到，单电子近似有时候是不成立的，电子间的相互作用可能非常重要，这也就是所谓的强耦合强关联问题。

在理解整数量子霍尔效应时，我们之所以可以忽略电子间的相互作用，还因为这时候最重要的是讨论朗道能级被填满的情形，而对于这种情形，电子间的相互作用并不会实质性地改变我们的物理图像，因为我们可以设想将电子间的相互作用从零连续调节到实际值，这时候朗道能级可能会发生一些“形变”，但它被填满的实质并不会改变。

但是，对于分数量子霍尔效应，它的朗道能级是部分填充的，如果忽略相互作用，那就会有巨量的不同填充方式。比方说对于 $\nu = 1/3$ 的情形，这时候是将 $N = N_\Phi/3$ 个电子

[1] 当然，$\nu = 1/\sqrt{2}$ 可以用两个较大的整数之比来近似，但这种难以分辨的有理数也没有观测到

填充到最低朗道能级的 N_Φ 个态上去，填充方式将有 $\binom{N_\Phi}{N_\Phi/3}$ 种，考虑到 N_Φ 非常大，所以这将是一个非常大的组合数。而这些不同填充方式都是简并的，对应整个材料的简并量子态 (注意，不是单电子的简并态)。对于这样的简并态空间，任何相互作用微扰都可能非常重要，都会极大地改变系统的行为。关于这一点，我们后面学习了量子力学微扰论以后再回头看会更加清楚。

神奇的是，在分数量子霍尔效应的系统基态 (整个材料的基态) 之上，可以出现奇异的准粒子激发。所谓的准粒子，意思就是，它们并不是自然界中普遍存在的粒子，因此不是组成霍尔片材料本身的电子和离子，而是在霍尔片中出现的行为和一个粒子的行为完全类似的一种局部状态。它们本质上是分数量子霍尔效应中的电子通过相互作用产生的东西。但是，分数量子霍尔效应中出现的这些准粒子非常奇怪，它们可以有分数电荷！比方说在 $\nu = 1/3$ 的霍尔态中，可以出现 $e/3$ 电荷的准粒子激发。而且，我们知道，自然界中的粒子可以分成两大类，玻色子和费米子，但是，这些准粒子既不是玻色子，也不是费米子，它们是一种被称为任意子 (anyon) 的东西。而且，我们知道，电子可以有两种不同的自旋状态，自旋向上和自旋向下，它们是电子的内禀状态，但是，在某些分数量子霍尔态中出现的准粒子甚至可以有诸如 $\sqrt{2}$ 这样的无理数个不同的内禀状态！

对这些奇异任意子的研究是当前的前沿之一，因为，原则上它们有实现大规模量子计算的能力。当前凝聚态研究中的拓扑序和量子计算研究中的拓扑量子计算有很大一部分都是在从不同的角度研究这样的任意子。

5. 定态微扰论和变分法

像上一章那样可以精确求解而又重要的量子系统并不是随处可见。在量子力学中，更多实际的系统都不可精确求解。这时候就需要发展近似计算的方法。本章讨论两种重要的求解量子力学系统的方法，定态微扰论和变分法。值得强调的是，它们不仅仅是数学上的方法，同时也是我们建立物理图像的重要方法。而且，通过应用变分法，我们还介绍了相变和对称性自发破缺的物理机制。

此外，本章还介绍了一种有效哈密顿量的方法，并用它讨论了与连续谱耦合的离散态的衰变问题。

5.1 非简并定态微扰论

对于哈密顿量不显含时间的体系，最重要的就是求解定态薛定谔方程，进而得到系统的能谱和能量本征态。但是，像线性谐振子和氢原子这样可以精确求解的例子其实并不多见。大量的量子力学体系可能都要求助于数值计算，比方说即使像氦原子这样的两个核外电子的体系都没有精确解，而需要求助于计算机的帮助。但是数值求解常常有一个缺陷，那就是虽然我们能求出一些可以和实验数据进行对比的值，但是背后隐藏的物理图像却常常被掩盖了，这会使得我们对出现这些数值背后的物理机制缺乏了解，这样我们就没有真正地解释实验现象，也没有真正地理解物理现象的本质。当然，数值计算有时候也可以引导我们发现物理机制。但是在量子力学上，还存在一大类物理图像比较清晰的求解定态薛定谔方程的方法。这就是我们这里要介绍的定态微扰理论。

定态微扰理论之所以有比较清晰的物理图像，是因为它用来处理的是这样一些问题。在这类问题中，我们可以把系统看成是一个相对简单的主要部分加上一个比较复杂的微扰。其中这个相对简单的部分提供物理图像的基础，也就是说，我们对于这个简单部分的了解是清楚的，我们已经通过各种办法得到了这个部分的能谱和定态波函数，并在这个基础上发展了我们的物理图像。复杂部分往往是阻碍我们发展物理图像的原因，但现在，它只是对主要部分的一个微扰，因此我们就可以用微扰展开的办法逐级分析这个微扰如何修正原来的能谱和定态波函数。由此我们也能进一步搞清楚这个微扰的引入在微扰展开的每一级上给我们原来的物理图像带来了些什么新的东西。这样我们就能以简单的主要部分的物理图像为基础得到整个问题的一个相对清晰的物理图像。正因为如此，在量子力学中，定态微扰论不仅仅是一个近似计算的方法，它同时也是一个建立物理图像的方法。幸运的是，自然界中有大量物理体系都可以用类似这样的微扰论方法来得到物理理解。不幸的是，微扰论的方法并不能处理所有的物理体系，的确有一些系统是强耦合强关联的，这时候非微扰的方法将会变得很重要。然而，在量子的世界中，目前人们还没有发展出一套像定态微扰论这样系统的非微扰理论框架。可以说如何处理非微扰系统是许多物理学领域共同面临的难题。

5.1.1 未扰动的系统以及投影算符

假设我们有一个了解得很清楚的量子系统，它的哈密顿算符我们记作 H_0。所谓了解得很清楚，我们指的是，我们已经求出了 H_0 的所有本征值 ε_m 和相应的本征态 $|m\rangle$，

$$H_0|m\rangle = \varepsilon_m|m\rangle, \tag{5.1}$$

不妨假定 $m = 1,2,3,.....$。当然，我们可以将这些本征态取成是正交归一的，即满足

$$\langle m|n\rangle = \delta_{mn}. \tag{5.2}$$

用这些正交归一的本征态我们可以构造一些被称为投影算符的特殊算符，比方说我们可以定义

$$P_m = |m\rangle\langle m|. \tag{5.3}$$

很显然，P_m 是一个厄米算符，而且 $P_m^2 = |m\rangle\langle m|m\rangle\langle m| = |m\rangle\langle m| = P_m$，即 P_m 满足

$$P_m^2 = P_m, \tag{5.4}$$

满足这种幂等关系的厄米算符就是投影算符。由于 $P_m|\phi\rangle = |m\rangle\langle m|\phi\rangle = \langle m|\phi\rangle|m\rangle$，所以投影算符 P_m 的作用是，当它作用在任意态 $|\phi\rangle$ 上时，都会将这个态投影到 $|m\rangle$ 态上。由于正交性，我们很容易看出，当 $m \neq n$ 时，$P_m P_n = P_n P_m = 0$，这时候我们称这两个投影算符正交，并且这时候很容易验证 $P_m + P_n$ 也是一个投影算符。

一般的，对于正交归一本征态的任何一个子集 $\mathscr{S}$，我们可以定义

$$P_{\mathscr{S}} = \sum_{m\in\mathscr{S}} P_m = \sum_{m\in\mathscr{S}} |m\rangle\langle m|, \tag{5.5}$$

很容易验证 $P_{\mathscr{S}}$ 是一个投影算符（即满足 $P_{\mathscr{S}}^2 = P_{\mathscr{S}}$），它的作用就是将任意量子态投影到由集合 $\mathscr{S}$ 里的所有本征态所张成的希尔伯特子空间里面，我们常常简称这个子空间为 $P_{\mathscr{S}}$ 子空间。假设我们记集合 $\mathscr{S}^\perp$ 为集合 $\mathscr{S}$ 在所有正交归一本征态全集中的补集，则很显然 $P_{\mathscr{S}^\perp}$ 也是一个投影算符，而且由于两个集合中的态相互正交，所以 $P_{\mathscr{S}^\perp} P_{\mathscr{S}} = P_{\mathscr{S}} P_{\mathscr{S}^\perp} = 0$，并且由于 $\mathscr{S} \cup \mathscr{S}^\perp$ 就是所有本征态的全集，所以

$$P_{\mathscr{S}^\perp} + P_{\mathscr{S}} = 1, \tag{5.6}$$

这里的 $1 = \sum_m |m\rangle\langle m|$ 表示单位算符。正因为如此，我们常常将 $P_{\mathscr{S}^\perp}$ 记作 $P_{\mathscr{S}}^\perp$，即 $P_{\mathscr{S}}^\perp = P_{\mathscr{S}^\perp}$，它的作用就是将量子态投影到 $P_{\mathscr{S}}$ 希尔伯特子空间的正交补空间。特别的，我们有投影算符 $P_n^\perp$，

$$P_n^\perp = \sum_{m\neq n} |m\rangle\langle m|, \tag{5.7}$$

它就是投影到除 $|n\rangle$ 态之外所有其余本征态所张成的希尔伯特子空间的投影算符。由于涉及到的这些态都是 H_0 的本征态，所以很显然

$$P_n H_0 = H_0 P_n,\ P_n^\perp H_0 = H_0 P_n^\perp. \tag{5.8}$$

5.1.2 布里渊-维格纳微扰论

现在，设想用一个所有矩阵元的取值都比较小的算符 V 扰动原来那个系统 H_0，进而得到一个哈密顿算符为 $H = H_0 + V$ 的新系统。由于在很多情况下 V 虽然是一个微扰，但是它可能是一个很复杂的算符，会在原来的系统中引入很复杂的相互作用，因此精确求解 H 的本征谱和本征态可能就会变得很困难。但鉴于我们已经了解清楚了原来的系统 H_0，同时也鉴于 V 是一个微扰，这时候我们往往可以用微扰展开的办法来处理新的系统 H。

具体来说，假设在扰动之后，原来 H_0 的本征态 $|n\rangle$ 变成了新系统的某个相应本征态 $|\psi_n\rangle$，相应的本征值也变成 E_n，即

$$H|\psi_n\rangle = E_n|\psi_n\rangle. \tag{5.9}$$

假定原来的能量本征值 ε_n 和 H_0 的其余本征谱之间存在着一个有限的谱隙，即对于任何 $m \neq n$，$|\varepsilon_n - \varepsilon_m| \geq \Delta > 0$。则，我们总是能够将 $|\psi_n\rangle$ 在分别由 P_n 和 $P_n^\perp$ 投影出来的两个正交且互补的空间中进行正交分解，通过合适地调整量子态整体的未定系数，我们可以设

$$|\psi_n\rangle = |n\rangle + P_n^\perp |\psi_n\rangle. \tag{5.10}$$

这个式子有很清楚的物理解释，$|n\rangle$ 就是原来 H_0 的第 n 个本征态，$P_n^\perp|\psi_n\rangle$ 表示由于 V 的引入而对这个态产生的扰动，$|\psi_n\rangle$ 就是扰动之后新的 H 的相应本征态，它由原来的 $|n\rangle$ 态加上扰动部分组成，当然，这种形式的 $|\psi_n\rangle$ 态是没有归一化的。

为了从 (5.10) 式出发得到关于 $|\psi_n\rangle$ 态的微扰展开式，我们将本征态 $|\psi_n\rangle$ 所满足的本征方程 (5.9) 重写为 $V|\psi_n\rangle = (E_n - H_0)|\psi_n\rangle$, 则 $P_n^\perp V|\psi_n\rangle = P_n^\perp(E_n - H_0)|\psi_n\rangle = (E_n - H_0)P_n^\perp|\psi_n\rangle$, 由于 H_0 在 $P_n^\perp$ 子空间上的谱和 E_n 之间有一个有限的谱隙，因此我们可以在这个子空间上求算符 $E_n - H_0$ 的逆算符 [1]，则

$$P_n^\perp|\psi_n\rangle = \frac{1}{E_n - H_0} P_n^\perp V|\psi_n\rangle. \tag{5.11}$$

定义 $G_0^\perp(E_n) = \frac{1}{E_n - H_0}P_n^\perp = \sum_{m\neq n}\frac{|m\rangle\langle m|}{E_n - \varepsilon_m} = P_n^\perp \frac{1}{E_n - H_0}P_n^\perp$, 则将 (5.11) 代入 (5.10) 式，我们可以得到关于 $|\psi_n\rangle$ 的一个方程

$$|\psi_n\rangle = |n\rangle + G_0^\perp(E_n)V|\psi_n\rangle. \tag{5.12}$$

将这个方程进行迭代[2]，就可以得到

$$\begin{aligned}|\psi_n\rangle = |n\rangle &+ G_0^\perp(E_n)V|n\rangle + G_0^\perp(E_n)VG_0^\perp(E_n)V|n\rangle \\ &+ G_0^\perp(E_n)VG_0^\perp(E_n)VG_0^\perp(E_n)V|n\rangle + ...\end{aligned} \tag{5.13}$$

这就是关于 $|\psi_n\rangle$ 按照微扰 V 进行级数展开的展开式。但是这个展开式依赖于 E_n，而到目前为止 E_n 的值还是未知的，所以下一步我们就是要给出计算 E_n 的方程。

为此，将 $V|\psi_n\rangle = (E_n - H_0)|\psi_n\rangle$ 左右两边分别和 $|n\rangle$ 作内积，利用 $\langle n|\psi_n\rangle = \langle n|n\rangle = 1$, 就可以得到 $E_n - \varepsilon_n = \langle n|V|\psi_n\rangle$, 进而代入 $|\psi_n\rangle$ 的微扰展开式 (5.13) 就可以得到

$$E_n - \varepsilon_n = \langle n|V|n\rangle + \langle n|VG_0^\perp(E_n)V|n\rangle + \langle n|VG_0^\perp(E_n)VG_0^\perp(E_n)V|n\rangle + ... \tag{5.14}$$

这当然是一个关于 E_n 的方程，通常可以通过将这个方程截断到有限项进而求出 E_n 的一个近似解。

但是，我们也可以按照微扰迭代的办法不断逼近方程 (5.14) 的精确解，具体来说即是：首先，每个 V 都是一个一阶小量。其次，我们设想可以按照微扰小量的阶数逐渐逼近 E_n 的精确值。也就是，我们先算出 E_n 的 0 阶近似 $E_n^{(0)}$，当然 $E_n^{(0)} = \varepsilon_n$，由 $E_n^{(0)}$ 加上一阶微扰，我们就能算出 E_n 的包含一阶微扰的近似 $E_n^{(1)}$，忽略掉 (5.14) 式中的高阶微扰，很容易看出 $E_n^{(1)} = \varepsilon_n + \langle n|V|n\rangle$。由 $E_n^{(0)}$ 和 $E_n^{(1)}$ 可以算出 E_n 的包含前二阶微

[1] 在整个希尔伯特空间上，算符 $E_n - H_0$ 可能有 0 本征值，从而不一定有逆算符

[2] 即将公式 (5.12) 理解成 $|\psi_n\rangle_{i+1} = |n\rangle + G_0^\perp(E_n)V|\psi_n\rangle_i$, $i = 0,1,2,3.....$, 并令 $|\psi_n\rangle_0 = |n\rangle$

扰的近似 $E_n^{(2)}$，这时候我们要计算 (5.14) 式的二阶项 $\langle n|VG_0^{\perp}(E_n)V|n\rangle$，但由于这一项本身已经是微扰的二阶小量了，所以其中的 E_n 只需用 0 阶近似 $E_n^{(0)}$ 代入即可，因此由 (5.14) 式忽略掉三阶以上的微扰，就有 $E_n^{(2)}=\varepsilon_n+\langle n|V|n\rangle+\langle n|VG_0^{\perp}(E_n^{(0)})V|n\rangle$。同样的，由 $E_n^{(0)},E_n^{(1)},E_n^{(2)}$ 又可以进一步算出 E_n 的包含前三阶微扰的近似 $E_n^{(3)}$，为 $E_n^{(3)}=\varepsilon_n+\langle n|V|n\rangle+\langle n|VG_0^{\perp}(E_n^{(1)})V|n\rangle+\langle n|VG_0^{\perp}(E_n^{(0)})VG_0^{\perp}(E_n^{(0)})V|n\rangle$, 等等[3]，依此类推。很显然，假设这个过程收敛的话，那么必有

$$\lim_{i\to+\infty}E_n^{(i)}=E_n, \tag{5.15}$$

E_n 就是能量本征值的精确值。

即使没有算出 E_n 的精确值，得出这些 E_n 的近似值 $E_n^{(i)}$ 以后，代入 $|\psi_n\rangle$ 的微扰展开式 (5.13), 我们也可以算出 $|\psi_n\rangle$ 的近似解，比如说 $|\psi_n\rangle$ 的包含前二阶微扰的近似解 $|\psi_n^{(2)}\rangle$ 为，$|\psi_n^{(2)}\rangle=|n\rangle+G_0^{\perp}(E_n^{(1)})V|n\rangle+G_0^{\perp}(E_n^{(0)})VG_0^{\perp}(E_n^{(0)})V|n\rangle$。

当然，在实际中我们用得最多的是，能量本征值的直到二阶微扰为止的近似 $E_n^{(2)}$，以及量子态的到一阶微扰为止的近似 $|\psi_n^{(1)}\rangle$, 由上面给出的微扰展开方程以及 $G_0^{\perp}(E_n)$ 的定义，我们可以得到

$$\begin{aligned}E_n^{(2)}&=\varepsilon_n+\langle n|V|n\rangle+\sum_{m\neq n}\frac{\langle n|V|m\rangle\langle m|V|n\rangle}{E_n^{(0)}-\varepsilon_m}\\&=\varepsilon_n+\langle n|V|n\rangle+\sum_{m\neq n}\frac{|\langle n|V|m\rangle|^2}{\varepsilon_n-\varepsilon_m}.\end{aligned} \tag{5.16}$$

$$|\psi_n^{(1)}\rangle=|n\rangle+\sum_{m\neq n}|m\rangle\frac{\langle m|V|n\rangle}{\varepsilon_n-\varepsilon_m}. \tag{5.17}$$

人们通常用瑞利-薛定谔 (Rayleigh–Schrodinger) 微扰理论得到这两个结论，但是瑞利-薛定谔微扰理论推广到更高阶会比较复杂，这里我们采用布里渊-维格纳 (Brillouin Wigner) 的办法来处理，正如我们已经看到的，这种处理办法的好处是，可以很容易地得到任意阶微扰论的一般公式。

为了进一步理解二阶微扰修正，我们注意到 ε_n 和其余所有能级之间存在一个有限的能隙 Δ, 即对于 $m\neq n$, $|\varepsilon_n-\varepsilon_m|\geq\Delta$。因此 $|\sum_{m\neq n}\frac{|\langle n|V|m\rangle|^2}{\varepsilon_n-\varepsilon_m}|\leq\sum_{m\neq n}\frac{|\langle n|V|m\rangle|^2}{|\varepsilon_n-\varepsilon_m|}\leq\frac{1}{\Delta}(\sum_{m\neq n}\langle n|V|m\rangle\langle m|V|n\rangle)=\frac{1}{\Delta}[\langle n|V|(1-|n\rangle\langle n|)|V|n\rangle]=\frac{1}{\Delta}[\langle n|V^2|n\rangle-(\langle n|V|n\rangle)^2]=\frac{1}{\Delta}\langle n|(\Delta V)^2|n\rangle$(注意分母中的 Δ 表示能隙，而分子中的 $\Delta V=V-\langle n|V|n\rangle$)。从上面推导过程最后的式子 $\frac{1}{\Delta}\langle n|(\Delta V)^2|n\rangle$ 可以看出，二阶微扰修正实际上和 V 的涨落 $\langle n|(\Delta V)^2|n\rangle$ 密切相关，如果没有这个 V 的涨落那就不会有二阶微扰修正。

另外，从上面关于波函数的一阶微扰修正中我们很容易看出来微扰论适用的条件，那就是

$$|\frac{\langle m|V|n\rangle}{\varepsilon_n-\varepsilon_m}|\ll 1. \tag{5.18}$$

[3]注意 $E^{(1)}$ 中包含了一个因为微扰而来的一阶小量，但是 $E^{(1)}$ 出现在表达式 $G_0^{\perp}(E_n^{(1)})$ 的分母上，因此式中的 $G_0^{\perp}(E_n^{(1)})$ 也应该按照这个微扰一阶小量的级数展开来理解，而且如果我们还要继续计算更高阶的 $E_n^{(i)},i>3$ 的话，那 $G_0^{\perp}(E_n^{(1)})$ 只能取到微扰级数的 1 阶为止，更高阶小量应该舍去，因为否则的话我们就将本应归入更高阶微扰的贡献算进 $E_n^{(3)}$ 中了。类似的，所有 $G_0^{\perp}(E_n^{(i)})$ 这样的表达式都应该理解成只取到微扰级数展开的第 i 阶小量为止，更高阶小量要舍去。

5.1.3 范德瓦尔斯力、万有引力、以及射电天文学

这一节我们举三个应用二阶微扰论的例子。第一个例子是用二阶微扰论解释非极性分子之间的范德瓦尔斯力，顺带也解释了一下为什么万有引力是一个吸引力，其道理其实和非极性分子之间的范德瓦尔斯力是一个吸引力 (当分子间的距离足够大时) 有相通之处。在这个例子之后，利用一个类似的偶极相互作用微扰项，我们还用一阶微扰论估算了氢原子超精细能级分裂，进而简单解释了射电天文学中著名的 $21cm$ 波长谱线。最后，我们要举的第三个例子是有关非线性振子能级的一个二阶微扰论计算。

非极性分子之间的范德瓦尔斯力

当分子间距离合适时，范德瓦尔斯力是两个中性分子之间的吸引力。两个中性分子之间存在吸引力有时候并不难理解，比方说，如果这两个分子是极性分子，那么它们的电偶极矩之间就可能存在相互作用，这时候这两个分子之间就会存在范德瓦尔斯力。不过极性分子之间相互作用势能和它们的电偶极矩的相对取向有关系，而由于热运动，分子电偶极矩的空间取向有一定的随机性，特别的，当温度无限高时，分子电偶极矩的空间取向将会完全随机，因此两个分子各自电偶极矩的空间取向将完全无关，这时候，这种由非 0 的电偶极矩间的相互作用所产生的范德瓦尔斯力就会趋于 0。

但是，令人觉得不可思议的是，即使是非极性分子之间也存在范德瓦尔斯力，甚至大小并不比极性分子之间的范德瓦尔斯力弱。非极性分子之间的范德瓦尔斯力是怎么来的呢？为什么在分子间的距离足够大时它是一个吸引力呢？甚至为什么在距离足够大时，这个力也和距离的 7 次方成反比呢？后面这两个问题可以合并成一个问题，即为什么非极性分子之间在距离足够大时，它们的相互作用能趋向于 $-\frac{A}{R^6}$ 呢？这里 R 是两个分子之间的距离，$A>0$ 是一个常数。要解释这里的物理规律就需要用到量子力学二阶微扰论。

首先，在量子力学里面，电偶极矩当然也要表示成算符，对于任何多个带电粒子的体系，这个算符总是能够定义的，所有的分子 (即使是非极性分子) 都是有多个带电粒子的体系，因此当然都有一个电偶极矩算符。当我们说一个分子是非极性分子时，指的是在分子所处的量子态上，其电偶极矩算符的期望值等于 0。但是这个算符总是存在的。因此在算符的意义上，任何两个分子的电偶极矩之间都会有相互作用，并且给系统贡献一个电偶极相互作用项 V，很显然，V 应该和每个分子的电偶极矩算符成正比。假设我们记所考察的这两个分子的电偶极矩算符分别为 $\mathbf{d}^A$ 和 $\mathbf{d}^B$, 则它们的电偶极相互作用算符 V 显然将具有 $V=C_{ij}d_i^Ad_j^B$ 这样的形式，这里 C_{ij} 依赖于两个分子之间的距离，并且显然只要分子间的距离足够大，它应该就是随着距离衰减的, 同时这里的 $i,j=1,2,3$ 表示三个直角坐标分量, 并且我们默认对重复的指标进行求和。

这样一来，当分子间的距离足够大时 (也就是当两个分子间的距离大于每个分子内部电子的德布罗意波长时)，它们的电偶极相互作用 V 就可以看成是对这个双分子系统的微扰。如果忽略掉这个微扰相互作用，那么两个分子将相互独立 (注意我们假设了分子间距离足够大，从而两个分子内部电子的波函数不会出现交叠)，各自有其哈密顿量 H_0^A 和 H_0^B。由于能量最低原理，这两个分子将分别处在 H_0^A 的基态 $|\psi_0^A\rangle$, 和 H_0^B 的基态 $|\psi_0^B\rangle$, 它们的基态能量分别记作 ε_0^A 和 ε_0^B。如果忽略 V 的扰动，这个双分子系统的哈密顿量将

是 $H_0 = H_0^A + H_0^B$。很显然 H_0 的基态将是 $|\psi_0\rangle = |\psi_0^A, \psi_0^B\rangle = |\psi_0^A\rangle|\psi_0^B\rangle$, 基态能量 ε_0 将是 $\varepsilon_0 = \varepsilon_0^A + \varepsilon_0^B$。

考虑到微扰 V，那这个双分子系统的完整哈密顿量应该是 $H = H_0 + V$。对基态能量的一阶微扰修正将是 $\langle\psi_0|V|\psi_0\rangle = \langle\psi_0^B|\langle\psi_0^A|C_{ij}d_i^A d_j^B|\psi_0^A\rangle|\psi_0^B\rangle = C_{ij}\langle\psi_0^A|d_i^A|\psi_0^A\rangle\langle\psi_0^B|d_j^B|\psi_0^B\rangle$。但是我们前面已经说过，非极性分子电偶极矩算符的期望值为 0，因此 $\langle\psi_0^A|d_i^A|\psi_0^A\rangle = \langle\psi_0^B|d_j^B|\psi_0^B\rangle = 0$，从而 $\langle\psi_0|V|\psi_0\rangle = 0$。因此对于非极性分子，$V$ 的一阶微扰对系统的能量没有贡献，要解释范德瓦尔斯力我们需要进一步考虑二阶微扰。

根据二阶微扰论，对基态能量的二阶微扰将是 (我们将把二阶微扰对基态能量的修正记为 ε')

$$\varepsilon' = \sum_{n\neq 0, m\neq 0} \frac{|\langle\psi_0^A, \psi_0^B|V|\psi_n^A, \psi_m^B\rangle|^2}{\varepsilon_0 - (\varepsilon_n^A + \varepsilon_m^B)}, \tag{5.19}$$

式中对于 $n \neq 0, m \neq 0$, ε_n^A 表示 H_0^A 的激发态能级，$|\psi_n^A\rangle$ 是相应的激发态，ε_m^B 表示 H_0^B 的激发态能级，$|\psi_m^B\rangle$ 同样是相应的激发态。因此 $|\psi_n^A, \psi_m^B\rangle = |\psi_n^A\rangle|\psi_m^B\rangle$ 就是整个 H_0 的激发态，$\varepsilon_n^A + \varepsilon_m^B$ 是相应的激发态能量本征值。

很显然 $\langle\psi_0^A, \psi_0^B|V|\psi_n^A, \psi_m^B\rangle = C_{ij}\langle\psi_0^A|d_i^A|\psi_n^A\rangle\langle\psi_0^B|d_j^B|\psi_m^B\rangle$。对于非极性分子，虽然 d_i^A 和 d_j^B 在各自基态上的期望值都等于 0，但是，由于每个分子都有可能形变，从而都存在正负电荷中心不完全重合的激发态，因此，一般来说，总存在一些激发态 $|\psi_n^A\rangle$ 和 $|\psi_m^B\rangle$ 使得 $\langle\psi_0^A|d_i^A|\psi_n^A\rangle \neq 0$, 同时 $\langle\psi_0^B|d_j^B|\psi_m^B\rangle \neq 0$, 这时候就会出现 $\langle\psi_0^A, \psi_0^B|V|\psi_n^A, \psi_m^B\rangle \neq 0$。这也就是说，二阶微扰修正总是非 0 的。并且由于 $|\langle\psi_0^A, \psi_0^B|V|\psi_n^A, \psi_m^B\rangle|^2 \propto |C_{ij}|^2$，所以很显然 $\varepsilon' \propto |C_{ij}|^2$。在物理上，这是因为非极性分子的偶极矩在基态上的期望值虽然为 0，但是根据不确定原理，偶极矩的涨落依然存在，由于这样的涨落，每个分子可能发生瞬时的变形，产生瞬时的偶极矩，这些瞬时的偶极矩会产生一个瞬时的电场作用在另一个分子上，并使得另一个分子出现瞬时的极化，而极化以后的另一个分子的瞬时偶极矩又会产生一个瞬时的电场反作用在第一个分子上。整个过程可以形象地说成是两个分子交换了一个瞬时的偶极子电场，最后就表现为一个非 0 的二阶微扰相互作用。并且正是因为这种相互作用涉及到偶极电场的交换，所以它的效应不是正比于与偶极子电场密切相关的 C_{ij}，而是正比于它的模方。

但是，公式 (5.19) 最重要的一个结论是，$\varepsilon' < 0$，即对这个两分子系统基态能量的二阶微扰修正是负的。原因其实很简单，因为在公式 (5.19) 中，分母上的 $\varepsilon_n^A + \varepsilon_m^B$ 代表的是激发态的能量，而 $\varepsilon_0 = \varepsilon_0^A + \varepsilon_0^B$ 代表的是未微扰的基态能量，因此很显然 $\varepsilon_n^A + \varepsilon_m^B > \varepsilon_0$，从而公式 (5.19) 的分母一定是负的，这显然意味着整个表达式必然是负的，即 $\varepsilon' < 0$。实际上，这是一个普遍的结果，在量子力学中，不管什么系统，对基态的二阶微扰修正总是小于 0 的，人们可以很容易地用二阶微扰论的公式一般性地证明这一点。

这里我们插入一个简短的其它话题。正因为二阶微扰对基态能量的修正是负的，所以在粒子物理中任何两个粒子如果交换一个标量粒子，所产生的相互作用力将一定是一个吸引力。这是因为，两个粒子交换一个标量粒子的过程可以看成是对这个两粒子系统的一个二阶微扰，微扰之前系统处在这个被交换的标量粒子的真空态 (也就是基态) 上，因此根据刚才的推导，交换标量粒子的这个二阶微扰给系统能量带来的修正 ε' 一定小于 0，即

$\varepsilon' < 0$。而另一方面，交换标量粒子所产生的相互作用强度当然随着两个粒子的距离 R 单调递减，即相互作用能的绝对值 $|\varepsilon'|$ 必然是 R 的减函数，因此 $\frac{\partial}{\partial R}|\varepsilon'(R)| < 0$。同时，由于 $\varepsilon' < 0$, 所以 $-\varepsilon' = |\varepsilon'|$, 因此根据保守力的公式，交换标量粒子所产生的相互作用力 $F = -\frac{\partial}{\partial R}\varepsilon'(R) = \frac{\partial}{\partial R}|\varepsilon'(R)| < 0$，从而必然是一个吸引力。在一定的意义上，万有引力之所以是一个吸引力也是因为类似的原因，当然引力子并不是一个标量粒子，但是物质通过交换引力子进行相互作用其道理和交换标量粒子其实有些类似。因此你也可以说，二阶微扰修正的负号同样可以解释为什么万用引力势能是负的。

上一段的讨论当然也完全适用于范德瓦尔斯力的情形。我们已经看到，这时候 $\varepsilon' < 0$, 并且由于在分子间距离足够大时，C_{ij} 随着距离衰减，所以二阶微扰带来的非极性分子之间的相互作用能 $|\varepsilon'(R)| \propto |C_{ij}(R)|^2$ 必然也随着距离衰减，因此根据上一段的讨论，这时候范德瓦尔斯力必定是一个吸引力。

下面我们进一步来理解为什么 ε' 反比于分子间距离的 6 次方。为此我们只需要研究 ε' 随着分子间距离是如何变化的 (R 足够大时)，由于 $|\varepsilon'(R)| \propto |C_{ij}(R)|^2$，所以我们又只需要确定 $C_{ij}(R)$ 的函数形式。下面的步骤其实就和量子力学没有什么关系了，完全是电动力学的内容。如果你还记得两个偶极子之间的相互作用势能反比于距离 3 次方的话，那你马上就知道，$C_{ij}(R) \propto 1/R^3$，从而 $|\varepsilon'(R)| \propto |C_{ij}(R)|^2 \propto 1/R^6$, 注意到 $\varepsilon' < 0$，那你马上就能得到答案

$$\varepsilon'(R) = -\frac{A}{R^6}, \tag{5.20}$$

在原则上，A 是某个可以根据上面描述的二阶微扰论算出来的大于 0 的常数，不过，对于两个比较复杂的分子，具体计算出 A 当然并不是一件容易的事情，但这是原子分子物理学家要做的事情，我们这里暂时忽略它。

假设你想知道得更具体一些，那么我们可以将两个分子的偶极相互作用描述成，一个分子的偶极子产生一个电场，这个电场作用在另一个分子的偶极矩上。电动力学里是这样描述的：首先，一个偶极子 d_i^A 在距离为 R 处 (R 足够大) 产生的电势为

$$\varphi = -\frac{1}{4\pi\varepsilon_0} d_i^A \partial_i \left(\frac{1}{R}\right), \tag{5.21}$$

这里 $\partial_i = \frac{\partial}{\partial x_i}$($x_i$ 是离分子 A 的位置为 $\mathbf{x}$ 处的坐标分量，因此 $R^2 = x_i x_i$), 并且我们默认对重复的指标进行求和。其次，$\mathbf{x}$ 位置的偶极子 d_j^B 处在 A 的电场 φ 中，与电场 φ 的相互作用势能为

$$V = d_j^B \partial_j \varphi. \tag{5.22}$$

将结果 (5.21) 和结果 (5.22) 结合起来，我们就能够算出偶极子 A 与偶极子 B 之间的偶极相互作用势能为

$$V = -\frac{1}{4\pi\varepsilon_0} d_i^A d_j^B \partial_i \partial_j \left(\frac{1}{R}\right) = \frac{1}{4\pi\varepsilon_0} d_i^A d_j^B \left(\frac{\delta_{ij} - 3e_i e_j}{R^3}\right), \tag{5.23}$$

式中 $e_i = \frac{x_i}{R}$ 是从分子 A 指向分子 B 的单位矢量。从这个偶极相互作用表达式，我们可以读出

$$C_{ij} = \frac{\delta_{ij} - 3e_i e_j}{4\pi\varepsilon_0 R^3}, \tag{5.24}$$

显然，它的确反比于距离 3 次方。(这一段的讨论请读者参考电动力学里关于电多极矩展开的相关内容)

超精细能级

完全类似的，读者很容易明白，两个磁偶极子也会有非常类似的偶极相互作用，它也可以写成 $V = C_{ij}\mu_i^A\mu_j^B$ 的形式，式中 μ_i^A 和 μ_j^B 分别是 A,B 两个磁偶极子的磁矩，不过由于是磁场，所以这时候

$$C_{ij} = \frac{\mu_0}{4\pi}\frac{\delta_{ij} - 3e_ie_j}{R^3}. \tag{5.25}$$

比如对于氢原子来说，电子是在质子所产生的场中运动，但是电子本身有自旋从而有磁矩，同样质子也有自旋因此也有磁矩，所以电子和质子的磁矩之间就会发生这种偶极相互作用 (实际上电子与质子的磁矩相互作用比这要略微复杂一些，由于电子可能有非 0 的概率出现在质子所在的坐标原点，所以两者的磁矩相互作用中其实还要包含一个接触项)。不过这时候这种偶极相互作用常常会有非 0 的一阶微扰效应，实际上，它是造成氢原子能级超精细分裂的原因之一。比方说，它会引起氢原子的基态产生超精细分裂。注意到这是一个一阶微扰效应，所以我们很容易估算这种超精细分裂的大小。首先，它正比于电子磁矩和质子磁矩的乘积，根据玻尔磁子的数值我们很容易估算，电子磁矩大约是 10^{-23}(单位 $A\cdot m^2$，即安培乘以米的平方)，由于质子质量比电子重 1000 多倍，所以质子磁矩大约是电子磁矩的 $1/1000$, 也就是大约 $10^{-26}A\cdot m^2$，将这两者乘起来再乘以 $\mu_0/(4\pi) = 10^{-7}N/A^2$，就得到 $10^{-56}N\cdot m^4$。由于 C_{ij} 反比于电子与质子距离的 3 次方，而对于基态上的氢原子 $R \sim a_0 \sim 0.5\times 10^{-10}m$($a_0$ 是玻尔半径)，所以 $1/R^3 \sim 10^{31}m^{-3}$。将 $10^{-56}N\cdot m^4$ 和 $1/R^3 \sim 10^{31}m^{-3}$ 乘起来就是能级超精细分裂的量级，显然结果是

$$10^{-25}N\cdot m \sim 10^{-6}eV. \tag{5.26}$$

将这个结果除以普朗克常数 $h \sim 6.62607015\times 10^{-34}J\cdot s$, 得到 $10^9Hz = 1000MHz$，这就是对氢原子超精细能级跃迁发出来的光频率的估算。更仔细的计算可以得出氢原子基态的超精细分裂所发出来的光频率为 $1420MHz$, 波长在 $21cm$。

这个 $21cm$ 谱线在射电天文学中有很重要的应用。因为星系中的氢原子会发射或吸收这一波长的辐射，所以用调谐到 $1420MHz$ 的射电望远镜，我们就能观察星系中氢原子气体密集处的位置和速度，进而就能了解星系的位置和运动情况。射电望远镜利用的这个 $21cm$ 谱线，波长在微波波段，这个波长的好处在于，它远比无线电波要短，因此有很好的定向性。同时它的波长又远比可见光的几百个纳米要长，因此也不像可见光那样容易被散射，因此射电望远镜能够看到很多光学望远镜看不到的天文事物。例如，通过射电望远镜，人们发现了天鹅座 A 的射电星系，它每秒钟发出的射电能量要比太阳每秒钟发出的能量强 1 亿亿倍以上，是迄今发现的最大射电星系，而光学望远镜对它却是一无所知。尤其是，宇宙微波背景辐射也是用射电望远镜发现的，因为微波背景辐射的波长当然在微波波段。目前世界上最大的射电望远镜，就是中国的位于贵州平塘的 500 米口径球面射电望远镜，简称 FAST，俗称中国天眼。

非线性振子的能级

我们要考察的第二个例子是一个非线性振子。具体来说就是给单自由度线性谐振子 $H_0=\frac{P^2}{2m}+\frac{1}{2}m\omega^2X^2$ 加上一个非线性微扰

$$V=\lambda\hbar\omega(\frac{\hbar}{m\omega})^{-2}X^4, \tag{5.27}$$

式中我们已经注意到 $(\frac{\hbar}{m\omega})^{\frac{1}{2}}$ 具有长度量纲，$\hbar\omega$ 具有能量量纲，因此将 V 写成这种形式是为了使得式中的 $\lambda\ll 1$ 是一个无量纲的常数。因此，在作了 $\hbar=m=\omega=1$ 的无量纲化处理以后，我们的微扰相互作用就是简单的

$$V=\lambda X^4. \tag{5.28}$$

下面我们就是要用二阶微扰论来计算这个微扰给谐振子的能级造成的影响。

为此我们引入产生湮灭算符 $a=\frac{1}{\sqrt{2}}(X+iP)$, $a^\dagger=\frac{1}{\sqrt{2}}(X-iP)$。则无量纲化的 H_0 就可以表示成 $H_0=a^\dagger a+\frac{1}{2}$, 其本征值为 $n+\frac{1}{2}(n\geq 0)$, 相应的本征态我们记作 $|n\rangle$。并且我们有代数关系 $a^\dagger|n\rangle=\sqrt{n+1}|n+1\rangle$, $a|n\rangle=\sqrt{n}|n-1\rangle$, 以及 $[a,a^\dagger]=1$。

很显然，由产生湮灭算符的定义我们有 $X=\frac{1}{\sqrt{2}}(a+a^\dagger)$, 代入 (5.28) 并不断地利用产生湮灭算符的代数关系 $[a,a^\dagger]=1$，我们就可以将微扰相互作用重写成

$$\begin{aligned}V&=\frac{\lambda}{4}(a+a^\dagger)^4=\lambda[a^\dagger a^\dagger(a^\dagger a)+(a^\dagger a)aa]+\frac{3\lambda}{2}[a^\dagger a^\dagger+aa]\\&+\frac{\lambda}{4}[a^4+(a^\dagger)^4]+\frac{3\lambda}{2}[(a^\dagger a)^2+(a^\dagger a)+\frac{1}{2}].\end{aligned} \tag{5.29}$$

从这个表达式我们很容易得到 V 的非 0 矩阵元，

$$\begin{aligned}\langle n+4|V|n\rangle&=\langle n|V|n+4\rangle=\lambda\frac{1}{4}\sqrt{(n+1)(n+2)(n+3)(n+4)}\\\langle n+2|V|n\rangle&=\langle n|V|n+2\rangle=\lambda(n+\frac{3}{2})\sqrt{(n+1)(n+2)}\\\langle n|V|n\rangle&=\lambda\frac{3}{2}(n^2+n+\frac{1}{2}).\end{aligned} \tag{5.30}$$

为了得到这些结果，我们要不断地使用 $(a^\dagger a)|n\rangle=n|n\rangle$ 以及 $a^\dagger|n\rangle=\sqrt{n+1}|n+1\rangle$, $a|n\rangle=\sqrt{n}|n-1\rangle$ 这样的关系式。除了这些矩阵元之外，V 的其它矩阵元都是 0。

利用二阶微扰论的公式 (5.16), 并代入 (5.30) 式给出来的 V 的矩阵元，注意到对于第 n 能级的二阶微扰修正只有 $n\pm 2$ 能级和 $n\pm 4$ 能级有贡献，进而我们就能够得到能级 E_n 直到二阶微扰的结果，

$$\begin{aligned}E_n/(\hbar\omega)&=\lambda^2\frac{1}{64}[(n-3)(n-2)(n-1)n-(n+1)(n+2)(n+3)(n+4)]\\&+\lambda^2\frac{1}{2}[(n-\frac{1}{2})^2(n-1)n-(n+\frac{3}{2})^2(n+1)(n+2)]\\&+n+\frac{1}{2}+\lambda\frac{3}{2}(n^2+n+\frac{1}{2})\\&=n+\frac{1}{2}+\lambda\frac{3}{2}(n(n+1)+\frac{1}{2})-\lambda^2\frac{1}{8}(2n+1)(17n(n+1)+21).\end{aligned}$$

这里我们已经恢复了量纲。很显然，这个结果有其适用范围，它只对比较低的能级适用，具体来说就是要求 $n^2\lambda \ll 1$。也即是说，对于非常高的激发态，微扰论的计算其实是失效的。原因在于对于高激发态，振子的能量很高，因此振幅就会比较大，也即是说波函数在 X 取值很大的地方也会有非 0 值。而当 X 很大的时候，λX^4 就会变得很重要，这时候它就不再能当作微扰来看了，事实上，这时候这一项比标准的 X^2 项更重要。因此，前面的微扰论处理当然就失效了。

5.1.4 习题

1. 假设某两态系统的哈密顿量可以写成

$$H = \begin{pmatrix} \varepsilon_1 + a\varepsilon & b\varepsilon \\ b\varepsilon & \varepsilon_2 \end{pmatrix}. \tag{5.31}$$

式中 $\varepsilon_1 \neq \varepsilon_2$，$a,b$ 为实数，ε 非常小。(1). 请用二阶微扰算出系统的两个能量本征值。(2). 请求出以上哈密顿量的精确本征值，然后将它们按照 ε 进行级数展开，并将结果与二阶微扰的结果进行比较。

2. 两个全同的零自旋玻色子放在一个无限深的方势阱中，阱内势能为零，阱外势能为正无穷，势阱的宽度为 a。假设这两个粒子之间存在相互作用 $V(x_1,x_2) = -V_0 a\delta(x_1 - x_2)$。(1). 忽略两粒子间的相互作用，请求出系统的基态和第一激发态 (包括能量本征值和相应的本征波函数)。(2). 请用一阶微扰论计算相互作用对基态和对第一激发态的能量修正。

3. 一个刚性转子绕着固定的 z 轴转动，转动惯量为 I。现在用一个 x 方向的均匀电场 E 来扰动此系统，假设转子的电偶极矩大小为 D，请用二阶微扰论计算转动的能量本征值和相应的本征波函数。

5.2 * 有效哈密顿量、简并微扰论以及不稳定态

上一节我们推导定态微扰论的相关公式时有一个关键性的假设，即假设我们考察的能级 n 非简并，我们假设 $|n\rangle$ 态和其余所有能态之间有一个有限的能隙 Δ。但其实有时候我们也会碰到这种情况，即作为 H_0 的本征态，我们所关心的这个能态 $|n\rangle$ 其实是简并的，那这个时候微扰论该怎么处理呢？这就是我们这一节要探讨的内容之一，通常称作简并定态微扰理论。

简并定态微扰理论的核心思想是，消去量子态在简并子空间之外的分量，将问题约化到简并子空间里来处理。这种降低希尔伯特空间的维数，将物理问题约化到物理上更相关的希尔伯特子空间里来处理的思想其实很重要，它本身有很广泛的用途。在这一节中我们将首先用有效哈密顿量的概念来介绍这一思想。然后将简并定态微扰理论作为有效哈密顿量方法的一个应用。最后，作为有效哈密顿量方法的另一个应用，我们还将用它来讨论与连续谱耦合的态的不稳定性。

5.2.1 有效哈密顿量

在讨论简并微扰论之前，我们先来讨论一下所谓的有效哈密顿量方法。所谓的有效哈密顿量就是在求解量子力学问题时，我们可以先把一部分我们不关心的量子态信息消去，进而将系统约化到某个希尔伯特子空间上所得到的对系统的有效描述。

具体来说，假设我们要求解定态薛定谔方程

$$H|\psi\rangle = E|\psi\rangle. \tag{5.32}$$

假设系统的整个希尔伯特空间可以分解成两个正交补空间，分别对应投影算子 P 和 $P^\perp$，它们满足

$$P + P^\perp = 1, \tag{5.33}$$

此外还满足 $PP^\perp = P^\perp P = 0$。因此我们可以将 $|\psi\rangle$ 态正交分解成 $|\psi\rangle = P|\psi\rangle + P^\perp|\psi\rangle$，其中 $P^\perp|\psi\rangle$ 的信息是我们不关心的，我们想推导的就是，消去 $P^\perp|\psi\rangle$ 以后，$P|\psi\rangle$ 满足的方程是什么?

为此，我们注意到，我们所要求解的定态薛定谔方程 (5.32) 可以重写为 $H(P|\psi\rangle) + H(P^\perp|\psi\rangle) = E|\psi\rangle$, 分别用投影算符 P 和 $P^\perp$ 作用在这个方程上，就可以得到

$$\begin{aligned} & H_{PP}(P|\psi\rangle) + H_{PP^\perp}(P^\perp|\psi\rangle) = E(P|\psi\rangle) \\ , \qquad & H_{P^\perp P}(P|\psi\rangle) + H_{P^\perp P^\perp}(P^\perp|\psi\rangle) = E(P^\perp|\psi\rangle), \end{aligned} \tag{5.34}$$

式中 $H_{PP} = PHP$ 为哈密顿算符在 P 子空间里的作用，$H_{P^\perp P^\perp} = P^\perp H P^\perp$ 为哈密顿算符在 $P^\perp$ 子空间里的作用，而 $H_{PP^\perp} = (H_{P^\perp P})^\dagger = PHP^\perp$ 将两个正交的子空间耦合了起来。

由 (5.34) 的第二个方程可以得到 $H_{P^\perp P}(P|\psi\rangle) = (E - H_{P^\perp P^\perp})(P^\perp|\psi\rangle)$, 由于我们关心的是 P 子空间里的物理，所以可以假设能量 E 和 $H_{P^\perp P^\perp}$ 的所有本征值都不同，则我们就有 $\frac{1}{E - H_{P^\perp P^\perp}} H_{P^\perp P}(P|\psi\rangle) = (P^\perp|\psi\rangle)$。将这个结果代入 (5.34) 的第一个方程消去 $(P^\perp|\psi\rangle)$ 就可以得到

$$\left(H_{PP} + H_{PP^\perp}\frac{1}{E - H_{P^\perp P^\perp}} H_{P^\perp P}\right)(P|\psi\rangle) = E(P|\psi\rangle). \tag{5.35}$$

换言之，假设对于任意探测能量 $z = \hbar\omega$ 我们定义作用在 P 子空间里的有效哈密顿量 $H_{\text{eff}}(z)$

$$H_{\text{eff}}(z) = H_{PP} + H_{PP^\perp}\frac{1}{z - H_{P^\perp P^\perp}} H_{P^\perp P}. \tag{5.36}$$

那么满足这个有效哈密顿量本征方程 $H_{\text{eff}}(E)|\varphi\rangle = E|\varphi\rangle$ 的本征值 E 就是原哈密顿量 H 的本征值，而 $|\varphi\rangle$ 就是 H 的相应本征态 $|\psi\rangle$ 在 P 子空间里的投影，即 $|\varphi\rangle = P|\psi\rangle$。换言之，对于探测能量 $z = \hbar\omega$，P 子空间里所有量子态的动力学都完全由有效哈密顿量 $H_{\text{eff}}(z)$ 决定。

实际上，可以证明，

$$P\frac{1}{z - H}P = \frac{1}{z - H_{\text{eff}}(z)}. \tag{5.37}$$

证明方法是: 在 P 和 $P^{\perp}$ 两个正交子空间中，分别将哈密顿量 H 和投影算符 P 表示成如下分块矩阵

$$\begin{pmatrix} H_{PP} & H_{PP^{\perp}} \\ H_{P^{\perp}P} & H_{P^{\perp}P^{\perp}} \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.38}$$

则 $P\frac{1}{z-H}P$ 就相当于

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} z-H_{PP} & -H_{PP^{\perp}} \\ -H_{P^{\perp}P} & z-H_{P^{\perp}P^{\perp}} \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.39}$$

最后，利用矩阵代数里分块矩阵求逆矩阵的方法 (见本小节最后) 就可以得到上式的结果是

$$\begin{pmatrix} \frac{1}{z-H_{\text{eff}}(z)} & 0 \\ 0 & 0 \end{pmatrix}. \tag{5.40}$$

这就已经证明了我们需要的结果 (5.37)。

一个简单的例子

让我们来考察一个最简单的例子，假设某系统的哈密顿量 H 可以写成一个 2×2 的矩阵，

$$H = \begin{pmatrix} U & t \\ t & 0 \end{pmatrix}. \tag{5.41}$$

我们可以消去这个系统的第 1 行和第 1 列，从而得到一个作用在第 2 行第 2 列的子空间里的有效哈密顿量 $H_{\text{eff}}(z)$。由 (5.36) 式，我们有

$$H_{\text{eff}}(z) = \frac{t^2}{z-U}. \tag{5.42}$$

如果我们要求原哈密顿量 H 的本征值 E，那根据有效哈密顿量方法，我们可以在约化后的子空间里解方程 $H_{\text{eff}}(E)|\psi\rangle = E|\psi\rangle$, 由于现在只有一行一列，所以本征值 E 必定满足

$$E = H_{\text{eff}}(E) = \frac{t^2}{E-U}, \tag{5.43}$$

很显然，这个方程的确是原哈密顿量 H 的本征值方程。

就我们这个简单例子来说，方程 (5.43) 很容易精确求解。但这里我们想描述一种更具一般性的求解方法，为此我们假设 $t \ll U$。这时候我们可以这样来求解 (5.43), 简单来说即是迭代，首先，我们把 t 看成是一个微扰，很显然如果忽略这个微扰，那第 2 行第 2 列子空间里的本征值将是 $E=0$。因此我们将 $E=0$ 代入方程 (5.43) 右边，这样就会得到 $E=-\frac{t^2}{U}$，接着再将 $E=-\frac{t^2}{U}$ 代入方程 (5.43) 右边, 近似就会得到 $E=-\frac{t^2}{U}+\frac{t^4}{U^3}$，如此不断迭代下去，就可以得到一个关于 t/U 的无穷级数，如果这个级数收敛，那它最终就会收敛到真正的能量本征值。

分块矩阵求逆公式

本小节用到分块矩阵求逆矩阵，这里给出公式，

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & -(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} & A_{22}^{-1}+A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \end{pmatrix}$$

式中 $A_{11},A_{22},A_{12},A_{21}$ 均为子矩阵。这个分块矩阵求逆矩阵的公式推导没有那么简单，但是证明却很简单，直接验证就可以了。另外，如果多观察一下，比如观察指标的规律之类，人们将会发现，这个公式其实是挺有规律的，不难记住。

5.2.2 简并微扰论

下面我们将上一小节的有效哈密顿量理论用于简并微扰论。为此我们取投影算符 P 对应于 H_0 的某个简并度为 d 的简并能级所对应的简并子空间。不妨将这 d 个简并的能态记为 $|1\rangle,|2\rangle,...|d\rangle$, 我们以 $\mathscr{D}=\{i=1,2,3,...,d\}$ 来表示这 d 个简并态的集合，我们记这 d 个态所对应的 H_0 本征值为 ε。那么投影算符 P 就可以写成

$$P=\sum_{i=1}^{d}|i\rangle\langle i|. \tag{5.44}$$

同样，我们以 $P^{\perp}$ 来表示 H_0 的所有正交归一本征态中除了我们正在关心的这 d 个简并态之外，其余所有本征态 (即 $\{|m\rangle,m\in\mathscr{D}^{\perp}\}$) 共同张成的希尔伯特子空间的投影算符，我们记这些非简并本征态的 H_0 本征值为 ε_m。如果按照我们在正文中所引入的记号，那显然有 $P=P_{\mathscr{D}}$，$P^{\perp}=P_{\mathscr{D}^{\perp}}$。当然，$P^{\perp}$ 和 P 之间就不再有任何简并了，而是有有限的能隙 Δ。另外，由于 P 和 $P^{\perp}$ 都是用 H_0 的本征态定义的，所以很显然，P 和 $P^{\perp}$ 均与 H_0 对易。

现在我们加入微扰 V，因此完整的哈密顿量应该是 $H=H_0+V$。假设考虑到微扰以后，原来 H_0 的那个 d 重简并子空间分裂成了 H 的 d 个不同本征态 $\{|\psi_l\rangle,l=1,2,3...,d\}$, 它们分别满足本征方程

$$H|\psi_l\rangle=E_l|\psi_l\rangle. \tag{5.45}$$

我们将要做的，就是要研究如何用微扰展开的办法计算 E_l 和 $|\psi_l\rangle$。

很显然，我们可以利用上一小节的有效哈密顿量方法先求解出 $P|\psi_l\rangle$, 为此我们需要先计算有效哈密顿量 $H_{\text{eff}}(z)$。由于 $H_{PP}=PH_0P+PVP=\varepsilon+V_{PP}$, $H_{PP^{\perp}}=PH_0P^{\perp}+PVP^{\perp}=H_0PP^{\perp}+V_{PP^{\perp}}=V_{PP^{\perp}}$, 同理 $H_{P^{\perp}P}=P^{\perp}VP=V_{P^{\perp}P}$。因此我们很容易得到有效哈密顿量 $H_{\text{eff}}(z)$ 为

$$H_{\text{eff}}(z)=\varepsilon+V_{PP}+V_{PP^{\perp}}\frac{1}{z-H_{P^{\perp}P^{\perp}}}V_{P^{\perp}P}. \tag{5.46}$$

式中 $H_{P^{\perp}P^{\perp}}=H_0^{\perp}+V_{P^{\perp}P^{\perp}}$，其中 $H_0^{\perp}=P^{\perp}H_0P^{\perp}$。从而 $P|\psi_l\rangle$ 满足的方程即是

$$H_{\text{eff}}(E_l)(P|\psi_l\rangle)=E_l(P|\psi_l\rangle). \tag{5.47}$$

假设我们定义一个有效的“未微扰”哈密顿量 $H_{\text{eff},0}$, 和一个有效的微扰 $V_{\text{eff}}(E_l)$, 它们的定义分别是

$$H_{\text{eff},0}=\varepsilon+V_{PP},\ V_{\text{eff}}(E_l)=V_{PP^\perp}\frac{1}{E_l-H_{P^\perp P^\perp}}V_{P^\perp P}. \tag{5.48}$$

从而，$H_{\text{eff}}(E_l)=H_{\text{eff},0}+V_{\text{eff}}(E_l)$。那么关于 $P|\psi_l\rangle$ 的方程 (5.47) 显然就可以根据标准的非简并定态微扰论来求解。也即是说，我们可以首先在 P 子空间中求出 $H_{\text{eff},0}$ 的 d 个本征态 $|\phi_l\rangle$,

$$H_{\text{eff},0}|\phi_l\rangle=\widetilde{\varepsilon}_l|\phi_l\rangle. \tag{5.49}$$

一般来说，因为定义中包含了 V, 所以 $H_{\text{eff},0}$ 的 d 个本征态通常就不再简并了, 那么下面我们就可以将 $V_{\text{eff}}(E_l)$ 当成微扰，然后照搬上一节中发展的非简并微扰论的办法来求解方程 (5.47) 了。

比方说，如果保留到微扰小量 V 的二阶 (由于 $V_{\text{eff}}(E_l)$ 本身是 V 的二阶项，因此当然就只需要算到 $V_{\text{eff}}(E_l)$ 的一阶微扰修正 $\langle\phi_l|V_{\text{eff}}(E_l)|\phi_l\rangle$)，按照上面的方法我们可以得到

$$E_l=\widetilde{\varepsilon}_l+\sum_{i,j\in\mathscr{D},\ m\in\mathscr{D}^\perp}\frac{\langle\phi_l|i\rangle\langle i|V|m\rangle\langle m|V|j\rangle\langle j|\phi_l\rangle}{\varepsilon-\varepsilon_m}. \tag{5.50}$$

注意, 上式右边的第二项我们代入了 E_l 的零阶近似, 而 $\widetilde{\varepsilon}_l$ 中是包含了 V 的一阶修正的, E_l 的 0 阶近似应该是 ε。另外, 这个式子中的 $\langle i|\phi_l\rangle$ 是 $|\phi_l\rangle$ 在原来的简并基 $\{|i\rangle, i=1,2,3,...,d\}$ 中的分量形式。其实在关于简并微扰论的实际应用中，最有用也是用得最多的就是这个结果。

以上就是一般性的简并微扰理论，我们最终是把它转变为非简并微扰的情形来处理的。上面讨论的简并微扰理论有一种特别简单的情形，那就是一开始的时候 H_0 的所有本征态都简并，那这种情况很简单，即有 $P=1$，$P^\perp=0$, 由 (5.48) 式可以看出，这时候 $V_{\text{eff}}(E_l)=0$, $H_{\text{eff},0}=\varepsilon+V$, 所以这时候只需要直接求 V 的本征态和本征值就可以了。

当然，前面还留有一个小问题，那就是逆算符 $\frac{1}{E_l-H_0^\perp-V_{P^\perp P^\perp}}$ 如何处理。一般来说，我们可以将它按照 $V_{P^\perp P^\perp}$ 进行级数展开，从而得到

$$\begin{aligned}&\frac{1}{E_l-H_0^\perp-V_{P^\perp P^\perp}}\\&=\frac{1}{E_l-H_0^\perp}+\frac{1}{E_l-H_0^\perp}V_{P^\perp P^\perp}\frac{1}{E_l-H_0^\perp}\\&+\frac{1}{E_l-H_0^\perp}V_{P^\perp P^\perp}\frac{1}{E_l-H_0^\perp}V_{P^\perp P^\perp}\frac{1}{E_l-H_0^\perp}+....\end{aligned}$$

这里我们利用了关于可逆算符 A 的恒等式

$$\begin{aligned}&\left(\frac{1}{A}+\frac{1}{A}B\frac{1}{A}+\frac{1}{A}B\frac{1}{A}B\frac{1}{A}+...\right)(A-B)=1\\&(A-B)\left(\frac{1}{A}+\frac{1}{A}B\frac{1}{A}+\frac{1}{A}B\frac{1}{A}B\frac{1}{A}+...\right)=1.\end{aligned}$$

在形式上，读者可以通过直接将这两个等式左边乘出来进而验证它们。不过，严格一点来说，这两个等式仅在括号里的那个无穷级数收敛时才有意义，因此通常只在 B 是一个微扰时才能这么做。很显然，这两个恒等式意味着

$$\frac{1}{A-B}=\frac{1}{A}+\frac{1}{A}B\frac{1}{A}+\frac{1}{A}B\frac{1}{A}B\frac{1}{A}+... \tag{5.51}$$

5.2.3 不稳定离散态

这一节我们将利用有效哈密顿量的方法处理一类普遍存在而又重要的物理现象，那就是不稳定离散态。这类现象之所以普遍，首先是因为，由于和环境耦合，基本上所有束缚系统的激发态都不稳定，都会衰变，从而都属于不稳定离散态。比方说，原子的激发态能级就不稳定，会向低能级跃迁。其次，无论是粒子物理中，还是凝聚态物理中，都有大量不稳定的粒子，比方说著名的上帝粒子 (Higgs) 粒子就是不稳定粒子，比方说凝聚态物理里面大量的准粒子其实也都是不稳定粒子。这些不稳定的，会衰变成其它粒子的粒子其实也都属于不稳定离散态。不过，对于这种不稳定粒子的情形，我们得稍微解释一下离散的含义。虽然不稳定粒子的动量是连续的，但是，在单个这样的不稳定粒子的衰变过程中，动量是守恒的，因此我们完全可以在一个总动量固定不变的子空间里来考察问题，而在保持动量不变的前提下，单个不稳定粒子当然也是离散态。

离散态的衰变不稳定性其实有一个共同的本质，那就是它们和一个较宽的连续谱有耦合。比方说原子向低能级跃迁放出光子，由于能量守恒，放出的这个光子的频率当然是分立的，但是它从属于一个连续谱。一个静止的不稳定粒子衰变成多个粒子，比方说衰变成两个粒子，这两个粒子的总动量保持是 0，但是它们各自的动量和能量分配都连续可变，因此衰变的末态也是一个连续谱。总之，离散态之所以不稳定，都是因为它们可以衰变到一个较宽的连续谱。

为了用有效哈密顿量的方法分析与连续谱耦合的离散态的不稳定性，我们记这些正交归一离散态为 $\{|i\rangle, i=1,2,3,..,N\}$，记这些离散态张成的希尔伯特子空间为投影算符 P，$P=\sum_{i=1}^{N}|i\rangle\langle i|$。假设与这些离散态耦合的连续谱态张成了 P 的正交补空间 $P^{\perp}$。假设在 $P^{\perp}$ 子空间上，有效哈密顿量中的 $H_{P^{\perp}P^{\perp}}$ 可以表示成

$$H_{P^{\perp}P^{\perp}}=\int d\alpha E_{\alpha}|\alpha\rangle\langle\alpha|. \tag{5.52}$$

从而 $|\alpha\rangle$ 构成 $P^{\perp}$ 的正交基，满足正交归一关系

$$\langle\alpha|\alpha'\rangle=\delta(\alpha-\alpha'). \tag{5.53}$$

这里指标 α 是用来标记这组连续谱态的抽象记号，它必须包含一个或多个连续指标，当然，α 里面也可以包含离散指标，对于 α 里面有离散指标的情形，人们只需将正交归一关系中相应的狄拉克 δ 函数理解成克龙内克 δ 符号，同时将相应的指标积分理解成离散求和就可以了！

由于 $|\alpha\rangle$ 是连续谱态，E_{α} 当然是连续变量。很多时候，我们常常选能量本身作为 $|\alpha\rangle$ 态的指标之一，这时候我们重新标记 $|\alpha\rangle$ 态为，$|\alpha\rangle=|E,\chi\rangle$, E 在数值上就等于 E_{α}，χ 是

用来进一步区分 $|\alpha\rangle$ 态的其它指标。作了这样的指标变换以后，态空间微元 $d\alpha$ 一般就会变换成

$$d\alpha = \rho(E,\chi)dEd\chi, \tag{5.54}$$

式中 $\rho(E,\chi)$ 是指标变换的雅可比行列式，也叫做态密度，因为它表示单位能量区间单位 χ 微元之内 $|\alpha\rangle$ 态的数目。

由于离散态和连续谱态之间有耦合，所以 $(H_{P^\perp P})^\dagger = H_{PP^\perp} \neq 0$。不妨记 $\langle i|H_{PP^\perp}|\alpha\rangle = \langle i|H|\alpha\rangle = A_{i\alpha}$, 式中我们用到了 $|\alpha\rangle$ 在 $P^\perp$ 中的投影是它本身, 同样 $|i\rangle$ 在 P 中的投影也是它本身的事实。并且, 我们常常记 $\sum_{i=1}^N |i\rangle A_{i\alpha} = \mathrm{A}_\alpha$。因此 $\sum_{i=1}^N \langle\alpha|H_{P^\perp P}|i\rangle\langle i| = \sum_{i=1}^N A_{i\alpha}^*\langle i| = \mathrm{A}_\alpha^\dagger$。如此一来，根据 (5.36) 式，离散态子空间 P 上的有效哈密顿量 $H_{\mathrm{eff}}(z)$ 就可以写成

$$\begin{aligned} H_{\mathrm{eff}}(z) &= H_{PP} + \int d\alpha \frac{\mathrm{A}_\alpha \mathrm{A}_\alpha^\dagger}{z - E_\alpha} \\ &= H_{PP} + \int dEd\chi\rho(E,\chi)\frac{\mathrm{A}_{E,\chi}\mathrm{A}_{E,\chi}^\dagger}{z-E}. \end{aligned} \tag{5.55}$$

很显然，如果探测能量 $z = \hbar\omega$ 落在连续能谱之外，那上面的 (5.55) 式定义良好。但是，如果探测能量 $z = \hbar\omega$ 落在连续能谱的区间之内，那在 $E = z$ 处，上面 (5.55) 的两个式子右边的第二项都将没有定义。为了避免这种没有定义的困难，我们可以给探测能量 $\hbar\omega$ 加一个无穷小的正虚部，即取 $z = \hbar\omega^+ = \hbar\omega + i\varepsilon$ (至于为什么要这样处理我们稍后再解释)，最后再取 $\varepsilon \to 0$。利用数学公式 (式中 $\mathscr{P}$ 表示取主值部分)

$$\lim_{\varepsilon\to 0}\frac{1}{x+i\varepsilon} = \mathscr{P}\frac{1}{x} - i\pi\delta(x), \tag{5.56}$$

我们就可以将 $H_{\mathrm{eff}}(\hbar\omega)$ 重写成

$$H_{\mathrm{eff}}(\hbar\omega) = H_{PP} + \Delta(\hbar\omega) - \hbar\frac{i}{2}\Gamma(\hbar\omega), \tag{5.57}$$

式中，$\Delta(\hbar\omega)$ 和 $\Gamma(\hbar\omega)$ 均为作用在离散态子空间 P 上的厄米算符，它们分别是

$$\Delta(\hbar\omega) = \mathscr{P}\int dEd\chi\rho(E,\chi)\frac{\mathrm{A}_{E,\chi}\mathrm{A}_{E,\chi}^\dagger}{\hbar\omega - E} = \mathscr{P}\int d\alpha\frac{\mathrm{A}_\alpha\mathrm{A}_\alpha^\dagger}{\hbar\omega - E_\alpha}, \tag{5.58}$$

$$\Gamma(\hbar\omega) = \frac{2\pi}{\hbar}\int d\chi \mathrm{A}_{E,\chi}\mathrm{A}_{E,\chi}^\dagger\rho(E,\chi)|_{E=\hbar\omega}. \tag{5.59}$$

可见，由于和连续谱的耦合，有效哈密顿量 $H_{\mathrm{eff}}(\hbar\omega)$ 不再是一个厄米算符！尤其是，它有了一个负虚部 $-\hbar\frac{i}{2}\Gamma(\hbar\omega)$, 下面我们将看到，这个负虚部的存在将意味着 P 子空间里的态是不稳定的，会随着时间指数衰减，从而表现出衰变行为。

值得提及的是，在一些场合中，人们也常常将结果 (5.57) 中的 $\Delta(\hbar\omega) - \hbar\frac{i}{2}\Gamma(\hbar\omega)$ 记作 $\Sigma(\hbar\omega)$, 即 $\Sigma(\hbar\omega) = \Delta(\hbar\omega) - \hbar\frac{i}{2}\Gamma(\hbar\omega)$，称为自能。

在清楚地看到自能的负虚部会导致量子态的指数衰变之前，让我们先来理解前面碰到 (5.55) 式时，为什么要作 $z = \hbar\omega + i\varepsilon$ 的处理，这个无穷小的正虚部从何而来。为此，我们

注意到，对于任何不显含时间的哈密顿量 H，当时间 $t>0$ 时，时间演化算符 $e^{-iHt/\hbar}$ 可以表示成，

$$e^{-iHt/\hbar}=\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega+i\varepsilon-H}. \tag{5.60}$$

为了证明这个式子，我们将它作用在 H 的任意一个本征值为 E 的本征态 $|E\rangle$ 上，这时候方程 (5.60) 左边将给出时间演化因子 $e^{-iEt/\hbar}$，而方程 (5.60) 右边的时间演化因子是积分 $\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega+i\varepsilon-E}$。我们将这个积分的被积函数延拓到整个 ω 复平面上，注意到 $t>0$，所以被积函数在 ω 的下半平面指数衰减，这是因为在复下半平面 ω 有一个负虚数，从而指数因子 $e^{-i\omega t}$ 呈现出指数衰减，特别是，在下半平面无穷远处的半圆上，$e^{-i\omega t}$ 将衰减为 0。因此，沿着实轴的积分 $\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega+i\varepsilon-E}$ 其实等于沿着实轴然后再加下半平面无穷远处的半圆这样一个闭合围道 C 上的积分，如图 (5.1) 所示。从图 (5.1) 也可以清楚地看到，

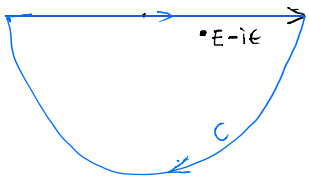

Figure 5.1: ω 复平面上的围道积分，最终要取极点 $E-i\varepsilon$ 的虚部 $\varepsilon\to 0$。

当将 ω 加上一个无穷小正虚部以后，围道积分 $\frac{i}{2\pi}\oint_C\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega+i\varepsilon-E}$ 将包含极点 $E-i\varepsilon$，从而根据留数定理我们可以得到 $\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega+i\varepsilon-E}=e^{-iEt/\hbar}$，刚好与方程 (5.60) 左边的结果相同。由于方程 (5.60) 作用在任意一个能量本征态 $|E\rangle$ 上都成立，而任意量子态一定能够写成能量本征态的线性叠加，从而方程 (5.60) 作用在任意量子态上都将成立，从而方程本身成立。

从上面对方程 (5.60) 的证明过程中我们清楚地看到，为了从 $\frac{1}{\hbar\omega-E}$ 这样的表达式中得到正确的量子态随时间的演化，我们必须给 $\hbar\omega$ 加上一个无穷小的正虚部。这也就是我们前面碰到 (5.55) 式时那样处理的原因。

现在我们来考察离散态子空间 P 之内的量子态如何随时间演化，这时候决定时间演化的当然是有效哈密顿量 $H_{\text{eff}}(\hbar\omega)$。实际上，将量子态的时间演化方程 (5.60) 投影到 P 子空间上，那么等式 (5.60) 左边就变成 $Pe^{-iHt/\hbar}P$，而等式右边就是要算 $\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega P\frac{e^{-i\omega t}}{\hbar\omega+i\varepsilon-H}P$，利用算符恒等式 (5.37)，这个式子的最终结果就是将哈密顿量 H 替换成有效哈密顿量 $H_{\text{eff}}(\hbar\omega+i\varepsilon)$，从而也就是计算积分

$$\begin{aligned}
&\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega-H_{\text{eff}}(\hbar\omega+i\varepsilon)+i\varepsilon}\\
&=\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega-H_{PP}-\Sigma(\hbar\omega)+i\varepsilon}\\
&=\frac{i}{2\pi}\int_{-\infty}^{+\infty}\hbar d\omega\frac{e^{-i\omega t}}{\hbar\omega-(H_{PP}+\Delta(\hbar\omega))+\hbar\frac{i}{2}\Gamma(\hbar\omega)+i\varepsilon}.
\end{aligned} \tag{5.61}$$

对于这个积分，我们可以采用和上面一样的围道来处理。但是现在，由于自能 $\Sigma(\hbar\omega)$ 有了一个负虚部 $-\hbar\frac{i}{2}\Gamma(\hbar\omega)$，情况就有了根本性的不同。为了看清楚这个不同，我们不妨假设

$\Gamma(\hbar\omega)$ 随着探测能量 $\hbar\omega$ 的变化足够平缓，以致于近似可以看成与 ω 无关，记作 Γ。这时候对 ω 的围道积分所包围的极点将有一个有限大的负虚部 $-\frac{i}{2}\Gamma$, 根据留数定理，这个负虚部最终将在积分 (5.61) 中贡献出如下因子

$$e^{-\frac{1}{2}\Gamma t}. \tag{5.62}$$

很显然，这个因子就意味着，P 子空间里的这些态将随着时间指数衰减，典型的衰减时间(称之为态的寿命)τ 为

$$\tau = \frac{1}{|\Gamma|}. \tag{5.63}$$

式中 $|\Gamma|$ 表示算符 Γ 的某个本征值的大小。

以上就清楚地说明了，与连续谱的耦合的确将导致离散态发生衰变。在上面结果的实际应用中，系统 H 往往可以分解成微扰之前的部分 H_0 和微扰部分 V 之和。这时候我们可以选 P 由 H_0 的离散本征态张成，而 $P^\perp$ 由 H_0 的连续谱本征态张成，因此 $H_{PP^\perp} = PH_0P^\perp + PVP^\perp = H_0PP^\perp + PVP^\perp = V_{PP^\perp}$，即 $H_{PP^\perp} = V_{PP^\perp}$。因此，正是微扰 V 产生了 H_0 的两种不同谱态之间的耦合。同时，在这种情况下，为了保证 (5.52) 式依然成立，我们还常常假定 $V_{P^\perp P^\perp} = 0$。

5.2.4 习题

1. 假设某系统的哈密顿量 H 可以写成如下矩阵，

$$H = \begin{pmatrix} U & 0 & t & t \\ 0 & U & t & t \\ t & t & 0 & 0 \\ t & t & 0 & 0 \end{pmatrix}. \tag{5.64}$$

(1). 请求出这个系统约化到右下角的 2×2 子空间里的有效哈密顿量 $H_{\text{eff}}(z)$。(2). 假设 $t \ll U$, 请求出系统在 0 附近的两个本征值 (计算到 t^2/U^2 阶为止)。(这个题目来自 Physics StackExchange 上的提问。提示：不能直接代简并微扰论一节最终给出的二阶微扰公式，得回到有效哈密顿量上想办法。)

2. 一个质量为 m 的粒子在势场 $V(x,y) = \frac{1}{2}m\omega^2(x^2+y^2)$ 中运动，其中 ω 是常数。假设给粒子加上一个小的微扰 $V' = \lambda xy$。请证明这个微扰会使得系统第一激发态的两个简并能态产生能级劈裂，请用二阶微扰论计算劈裂以后的能级分别为多少。

5.3 变分法在量子力学中的应用

5.3.1 变分法求基态能量

在量子力学中，求解一个系统的基态尤其是求解基态能量，有最为重要的作用。因为一个系统的基态和低激发态常常决定了系统的宏观行为。但是，能够精确求解的量子力学

系统少之又少。因此人们发展了一些对基态能量和基态波函数的近似计算方法，变分法就是这些近似计算方法中最为重要的方法之一。

用变分法计算基态能量基于下面这个简单的定理。即对于任意量子态 $|\psi\rangle$，我们有

$$E[\psi]=\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}\geq E_0, \tag{5.65}$$

式中，H 是系统的哈密顿量 (要求这个哈密顿量有下界，否则将不存在基态)，E_0 就是系统的基态能量，$E[\psi]$ 表示依赖于量子态 $|\psi\rangle$ 的一个能量泛函，式中的等号当且仅当 $|\psi\rangle$ 是系统基态时才能取到。

这个定理的证明非常简单。假设我们记 $|n\rangle(n\geq 0)$ 为 H 的本征态，本征值为 E_n，$n=0$ 对应的就是基态。则我们可以将 $|\psi\rangle$ 展开为 $|\psi\rangle=\sum_n\psi_n|n\rangle$，$\psi_n$ 为展开系数。很显然 $E[\psi]=\frac{\sum_n E_n|\psi_n|^2}{\sum_m|\psi_m|^2}$，注意到 $E_n\geq E_0$ 就有，$E[\psi]\geq E_0\frac{\sum_n|\psi_n|^2}{\sum_m|\psi_m|^2}=E_0$，定理因此成立。很明显，等于号成立的充要条件是 $\psi_n=0,n\neq 0$，这时候 $|\psi\rangle$ 本身就是基态。

变分法的基本思想是，合理地猜测一组试探波函数，然后选取这组试探波函数中使得能量泛函 $E[\psi]$ 取值最低的一个，将之作为对基态波函数的近似，相应的能量泛函的值就作为对基态能量的近似。很显然，根据刚才的定理，这样算出来的基态能量一定是偏高的，因此为了得到更好的近似结果，人们通常是在一个逐步扩大的试探波函数集合中选取对基态的近似，因此得到的基态能量近似值也将是逐步降低的，当这种降低趋向于收敛的时候，我们就可以认为已经得到对基态和基态能量的很好近似了。在实际操作中，我们常常是让试探波函数依赖于一些参数 $\lambda_1,\lambda_2,...$，用这个带参数的试探波函数计算能量泛函 $E[\psi]$，算出来的当然就是一个依赖于这些参数的能量表达式 $E(\lambda_1,\lambda_2,...)$，求这个表达式的最小值，并得出最小值所对应的参数，则与最小值参数相应的试探波函数就可以看成是近似的基态波函数，能量泛函的最小值就可以看成是近似的基态能量。同样，为了保证得到足够好的近似，我们可以不断地增加参数，然后接着求新参数所对应的能量泛函最小值，直到这些最小值趋于收敛。

求出足够好的近似基态以后，我们可以接着在与这个近似基态正交的子空间里按照同样的办法求第一激发态波函数以及能量的近似。因为很显然，在与基态正交的子空间里有一个和上面的定理完全类似的定理，只不过这时候不等号右边的下界是第一激发态的能量 E_1。当然，按照这样的办法人们还可以逐步地对各个激发态都进行近似求解。不过，相比较来说，近似求解基态是变分法用得最多的地方。

下面我们用变分法来近似求解氦原子的基态能级。我们假设氦原子核是不动的，位于坐标原点，因此氦原子的哈密顿量是

$$H=-\frac{\hbar^2}{2m}(\nabla_1^2+\nabla_2^2)-\frac{2e^2}{4\pi\varepsilon_0 r_1}-\frac{2e^2}{4\pi\varepsilon_0 r_2}+\frac{e^2}{4\pi\varepsilon_0 r_{12}}, \tag{5.66}$$

式中 $\mathbf{x}_1,\mathbf{x}_2$ 是两个电子的位置矢量，$r_1=|\mathbf{x}_1|,r_2=|\mathbf{x}_2|$，$r_{12}=|\mathbf{x}_1-\mathbf{x}_2|$。很显然氦原子的哈密顿量可以看成是在两个相互独立的 He^+ 离子的哈密顿量的基础上加上一个电子与电子的相互作用项 $\frac{e^2}{4\pi\varepsilon_0 r_{12}}$。$He^+$ 离子是一个类氢系统，其基态波函数在归一化以后是 $\left(\frac{Z^3}{\pi a_0^3}\right)^{1/2}e^{-Zr/a_0}$，式中 Z 就是原子序数，对于 He^+ 离子当然 $Z=2$，a_0 就是玻尔半径。

因此，两个 He^+ 离子 (共用同一个核) 的基态波函数当然就是

$$\psi(r_1,r_2)=\frac{Z^3}{\pi a_0^3}e^{-Z(r_1+r_2)/a_0}. \tag{5.67}$$

但是现在，由于电子与电子之间的库伦排斥相互作用，上面的 $\psi(r_1,r_2)$ 当然不可能是氦原子真正的基态波函数，但是我们不妨将它作为对氦原子基态的一个试探波函数。同时，考虑到由于第二个电子的屏蔽效应，每个电子不再能感受到一个裸的氦原子核的库伦场了，我们可以设想每个电子将感受到一个有效的原子序数 $Z\neq 2$, 并通过变分法来确定这个有效的 Z 值。也就是说，我们不妨将公式 (5.67) 中的 Z 变成一个参数，从而得到一组试探波函数 $\psi(r_1,r_2,Z)$。计算相应的能量泛函

$$\begin{aligned}E(Z)=&\int d\tau_1 d\tau_2\psi(r_1,r_2,Z)\left[-\frac{\hbar^2}{2m}(\nabla_1^2+\nabla_2^2)-\frac{2e^2}{4\pi\varepsilon_0 r_1}-\frac{2e^2}{4\pi\varepsilon_0 r_2}\right]\psi(r_1,r_2,Z)\\&+\int d\tau_1 d\tau_2\psi(r_1,r_2,Z)\left[\frac{e^2}{4\pi\varepsilon_0 r_{12}}\right]\psi(r_1,r_2,Z)\\=&-2(\frac{1}{2}m\alpha^2c^2)[4Z-Z^2-\frac{5}{8}Z].\end{aligned} \tag{5.68}$$

式中 $d\tau$ 表示三维体积元，$\alpha=\frac{e^2}{4\pi\varepsilon_0\hbar c}\simeq 1/137$ 是精细结构常数，因此 $\frac{1}{2}m\alpha^2c^2=13.6eV$ 是里德堡能量。式中的计算只有 $\int d\tau_1 d\tau_2\psi(r_1,r_2,Z)\left[\frac{e^2}{4\pi\varepsilon_0 r_{12}}\right]\psi(r_1,r_2,Z)$ 稍微难一点点，不过也无需用到任何特殊函数之类的东西，算是一道高数习题吧。

求表达式 $E(Z)$ 的最小值得到 $Z=\frac{27}{16}$, 代入得到 $E(Z)$ 的最小值约为 $-77.5eV$，这就是我们用变分法求出来的氦原子基态能量。这个值比真实值 $-78.6eV$ 略大，但是误差很小。考虑到氦原子基态并没有精确解，因此我们能如此迅速地得到这样的近似结果，这已经足以说明变分法的威力了。

5.3.2 定态作为能量泛函的极值状态

从上一节关于变分法的讨论中其实我们已经能够想见，能量本征态一定是能量泛函 $E[\psi]$ 的极值状态，基态更是整体的极小值状态。这里我们首先用数学推导来确认一下这一点。假设 $|\psi\rangle$ 态发生了一个无穷小变动，$|\psi\rangle\to|\psi\rangle+|\delta\psi\rangle$, 式中 $|\delta\psi\rangle$ 表示一个正比于无穷小量的任意状态。则能量泛函的改变量为

$$\begin{aligned}\delta(E[\psi])=&\frac{\langle\psi|H|\psi\rangle+\langle\delta\psi|H|\psi\rangle+\langle\psi|H|\delta\psi\rangle+O((\delta\psi)^2)}{\langle\psi|\psi\rangle+\langle\delta\psi|\psi\rangle+\langle\psi|\delta\psi\rangle+O((\delta\psi)^2)}-\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}\\=&\frac{\langle\delta\psi|H|\psi\rangle+\langle\psi|H|\delta\psi\rangle-E[\psi](\langle\delta\psi|\psi\rangle+\langle\psi|\delta\psi\rangle)}{\langle\psi|\psi\rangle}+O((\delta\psi)^2)\\=&\frac{\langle\delta\psi|(H-E[\psi])|\psi\rangle+\langle\psi|(H-E[\psi])|\delta\psi\rangle}{\langle\psi|\psi\rangle}+O((\delta\psi)^2).\end{aligned} \tag{5.69}$$

式中 $O((\delta\psi)^2)$ 代表高阶无穷小量。能量泛函的极值条件等价于 $\delta(E[\psi])$ 表达式中的一阶小量等于 0, 从上面的表达式容易看出这又等价于 $\langle\delta\psi|(H-E[\psi])|\psi\rangle+\langle\psi|(H-E[\psi])|\delta\psi\rangle=0$ 对于任意的无穷小变动 $|\delta\psi\rangle$ 都成立。不妨令 $|\varphi\rangle=(H-E[\psi])|\psi\rangle$, 因此一阶小量等于 0 的充要条件就变成 $\langle\delta\psi|\varphi\rangle+\langle\varphi|\delta\psi\rangle=0$ 对于任意的 $|\delta\psi\rangle$ 成立。取 $|\delta\psi\rangle=\varepsilon|\varphi\rangle$, 其中 ε

为一个实无穷小量，则这个一阶小量等于 0 的条件又意味着，$\varepsilon\langle\varphi|\varphi\rangle=0$，而这又意味着 $\langle\varphi|\varphi\rangle=0$, 从而 $|\varphi\rangle=0$。也即是说，$|\psi\rangle$ 是能量泛函的极值状态的充要条件是下式成立

$$(H-E[\psi])|\psi\rangle=0, \tag{5.70}$$

由于 $E[\psi]$ 是一个数，所以这一方程其实就是哈密顿算符的本征方程。因此，能量泛函取极值的充要条件是 $|\psi\rangle$ 是哈密顿算符的本征态，这时候能量泛函的值 $E[\psi]$ 显然就等于相应的本征值。

上面的这个结论通常也称之为里兹定理。它当然是对变分法的一个确认和补充。有时候用拉格朗日乘子法来表述里兹定理也是很方便的，即能量本征态是下面的能量泛函 $E[\psi]$ 在归一化约束 $\langle\psi|\psi\rangle=1$ 下的极值状态，

$$E[\psi]=\langle\psi|H|\psi\rangle. \tag{5.71}$$

这样表述以后人们就可以用标准的约束下取极值的拉格朗日乘子法来处理这一极值问题。当然，这样的处理和我们之前的处理完全等价。

有时候将能量泛函的表达式在坐标表象中用波函数的形式具体写出来也非常有用。我们以哈密顿量 $H=-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})$ 为例来进行说明。很显然这时候 $E[\psi]=\int d\tau\psi^*(\mathbf{x})\left[-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})\right]\psi(\mathbf{x})/(\int d$ 对这个式子进行分部积分就可以得到 $E[\psi]=\{\int d\tau\left[\frac{\hbar^2}{2m}|\nabla\psi(\mathbf{x})|^2+V(\mathbf{x})|\psi(\mathbf{x})|^2\right]-\frac{\hbar^2}{2m}\int_{S(\infty)}d\vec{S}\cdot$ $\psi^*(\mathbf{x})\nabla\psi(\mathbf{x})\}/(\int d\tau|\psi(\mathbf{x})|^2)$(式中 $S(\infty)$ 表示空间的边界)。对于表达式中的边界项 $\frac{\hbar^2}{2m}\int_{S(\infty)}d\vec{S}\cdot$ $\psi^*(\mathbf{x})\nabla\psi(\mathbf{x})$，假如我们考虑的是束缚态情形，或者是周期性边界条件的情形 (这相当于没有边界)，那这个边界项其实就等于 0。在这种情况下我们就有

$$E[\psi]=\frac{\int d\tau\left[\frac{\hbar^2}{2m}|\nabla\psi(\mathbf{x})|^2+V(\mathbf{x})|\psi(\mathbf{x})|^2\right]}{\int d\tau|\psi(\mathbf{x})|^2}. \tag{5.72}$$

当然，人们也可以将这个能量泛函用拉格朗日乘子法来表述，这时候根据里兹定理，将能量泛函在归一化约束之下对波函数 $\psi(\mathbf{x})$ 进行变分求极值，我们就会反过来得到坐标表象下的定态薛定谔方程。实际上，在某些比较复杂的坐标系中 (比如球坐标)，人们常常可以先在这种坐标系下写出能量泛函 (5.72)(取波函数归一化约束)，然后再用拉格朗日乘子法对波函数变分求出这种坐标下相应的定态薛定谔方程。相比于直接写这样的复杂坐标下的定态薛定谔方程，这样处理的简捷之处在于，能量泛函 (5.72) 中只涉及到一阶偏导，而一阶偏导在复杂坐标系中的表达形式远比直接写定态薛定谔方程时涉及的拉普拉斯算符 ∇^2 这样的二阶偏导要好处理。

5.3.3 有关基态的几个定理

这一节，我们利用变分原理讨论几个有关于系统基态的一般性定理。首先我们证明对于一大类量子力学系统，定态波函数总能取成实函数。这些系统就是具有如下形式定态薛定谔方程的系统，

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)\right]\psi(x)=E\psi(x). \tag{5.73}$$

很显然对于任意满足上面方程的 $\psi(x)$，其复共轭 $\psi^*(x)$ 也满足同样的方程，所以对于这样的系统我们总是能重新将定态波函数选成 $\psi(x)+\psi^*(x)$ 或者 $-i(\psi(x)-\psi^*(x))$，显然，这样重新选择以后的定态波函数是实函数。特别的，这一类系统的基态波函数总能取成实函数。

因此结合变分原理，系统 (5.73) 的基态就是使得下面的能量泛函取最小值的态。

$$E[\phi]=\frac{\int dx\left[\frac{\hbar^2}{2m}\left(\frac{\partial\phi}{\partial x}\right)^2+V\phi^2\right]}{\int\phi^2dx}, \tag{5.74}$$

其中 $\phi(x)$ 是一个任意的实波函数。

从这里我们能够得出一个定理，那就是系统 (5.73) 的基态波函数不能有节点，也就是说，除了在空间边界处之外，基态波函数必定恒大于 0。因为假设基态波函数有像图 (10.1) 所示的节点的话，那就必定有一部分波函数的值是负的，根据能量泛函的表达式，这时候我们完全可以把波函数的负值部分沿着自变量的 x 轴对折为正，就像图 (10.3) 所示的那样，由于能量泛函 $E[\phi]$ 对波函数以及波函数的导数的依赖都是平方型的，所以很显然这样做对 $E[\phi]$ 的值不会有任何影响。但是，一个类似于图 (10.3) 那样的波函数不可能对应能量泛函的最小值，因为很显然图 (10.3) 中波函数的导数在节点处有突变，而这就意味着我们可以将波函数作一个微小的改变，从而使得它在节点附近变得光滑，就像图 (5.4) 所示的那样，很显然这种光滑化的小手续会明显地降低能量泛函中与导数相关的部分 $\left(\frac{\partial\phi}{\partial x}\right)^2$ 在节点附近的值，并保持对能量泛函的所有其它贡献不变，因此通过这样的光滑化小手续我们就可以进一步降低能量泛函的值。这也即是说，图 (10.3) 那样的波函数不可能对应能量泛函的最小值，根据变分原理，从而它也就不可能是系统的基态，当然这也意味着，任何像图 (10.1) 那样的波函数也不可能是系统的基态。因此，系统 (5.73) 的基态波函数必定不能有节点。

上一段的定理有一个重要的推论。那就是系统 (5.73) 的基态不可能简并。因为否则的话，假设 $\phi_0(x)$ 和 $\phi_0'(x)$ 都是系统 (5.73) 的基态，则根据线性叠加原理 $a\phi_0(x)+b\phi_0'(x)$ 也将是系统 (5.73) 的基态，这里 a,b 是两个任意的非 0 实数。很显然，由于 $\phi_0(x)$ 和 $\phi_0'(x)$ 是两个不同的态，因此我们只要合适地选取叠加系数 a,b 就能让 $a\phi_0(x)+b\phi_0'(x)$ 出现节点。但根据上一段的定理，这将与 $a\phi_0(x)+b\phi_0'(x)$ 是系统 (5.73) 的基态相矛盾。从而这就证明了系统 (5.73) 的基态不可能简并。

另外，假设 $\phi_n(x)$ 是系统 (5.73) 的某个取成实函数的激发态波函数。则由于激发态和基态 $\phi_0(x)$ 必定要正交，从而我们可以知道，激发态波函数 $\phi_n(x)$ 必定存在节点。因为否则的话 $\int dx\phi_n(x)\phi_0(x)$ 的被积函数将恒大于 0 或者恒小于 0，那积分结果将不可能为 0，从而与正交性相矛盾。

此外还有一个推论也值得一提，那就是假设我们所考虑的系统 (5.73) 是宇称守恒的，即假设在空间反射 $x\to-x$ 之下有 $V(-x)=V(x)$，那系统的基态波函数 $\phi_0(x)$ 将必定是偶宇称的，即必定满足 $\phi_0(-x)=\phi_0(x)$。这是因为，如果 $V(-x)=V(x)$，那 $\phi_0(-x)$ 也必定满足基态的本征方程，但由于基态不简并，所以必有 $\phi_0(-x)=\alpha\phi_0(x)$。由于空间反射两次等于不做任何操作，所以 $\alpha^2=1$，$\alpha=\pm1$。但是由于 $\phi_0(x)$ 没有节点，所以它在空间反射之

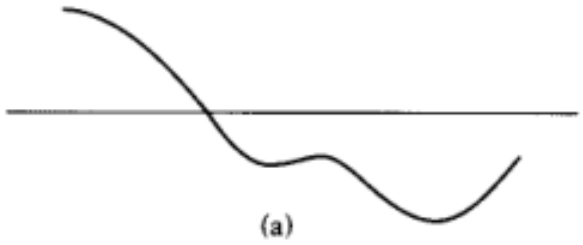

Figure 5.2: 假设波函数有节点. 图片来源：费曼统计力学

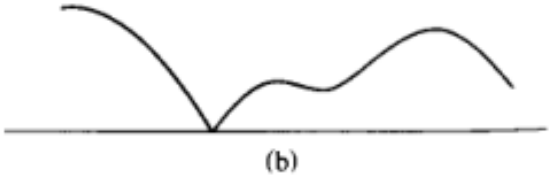

Figure 5.3: 把波函数的负部分反射成正. 图片来源：费曼统计力学

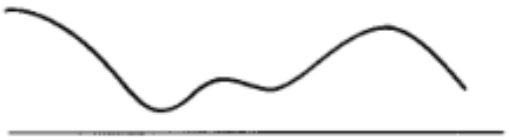

Figure 5.4: 把反射后的波函数作微小的光滑化. 图片来源：费曼统计力学

下也不能改变符号，因此只可能是 $\alpha = 1$，也即是说，$\phi_0(-x) = \phi_0(x)$。

以上所有的结论我们都是以单自由度系统为例来得到的，但显然这些结论全都可以推广到类似下面那样的多自由度多粒子体系。

$$\left[-\sum_{i=1}^{N}\frac{\hbar^2}{2m_i}\nabla_i^2+\sum_{i<j}V(r_{ij})\right]\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)=E\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N), \tag{5.75}$$

这里我们假设的是一个 N 粒子系统，$\mathbf{x}_i$ 是其中第 i 个粒子的位置矢量，$V(r_{ij})$ 是两体相互作用势能，其中 $r_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$ 是两个粒子的空间距离。

另外，如果我们考虑的是 N 个完全相同粒子的全同多粒子体系，如果这时候系统的哈密顿量依然具有 (5.75) 这样的形式的话 (当然这时候所有全同粒子的质量都是相同的，即 $m_i = m$), 那我们还可以得出这种哈密顿量的基态波函数必定是一个全对称波函数，即在不同粒子的位置作任意置换之下，波函数都保持不变。这是因为对于全同粒子，其哈密顿量在不同粒子位置矢量的任意置换 σ 下都必定是不变的。因此假设 $\psi_0(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$ 是哈密顿量 (5.75) 的基态波函数，则 $\psi_0(\mathbf{x}_{\sigma(1)},\mathbf{x}_{\sigma(2)},...,\mathbf{x}_{\sigma(N)})$(式中 $\sigma(i)$ 表示 i 在置换 σ 作用之后的结果) 必然也是哈密顿量 (5.75) 的一个基态波函数。但是，由于哈密顿量 (5.75) 的基

态必定不简并，所以必有 $\psi_0(\mathbf{x}_{\sigma(1)},\mathbf{x}_{\sigma(2)},...,\mathbf{x}_{\sigma(N)})=\alpha_\sigma\psi_0(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$，本来根据全同性原理[4]这样的 α_σ 可能等于 ±1，但是由于基态波函数不能有节点，因此在坐标置换前后不能改变符号，从而必有 $\alpha_\sigma=1$ (这也告诉我们，反对称的空间波函数必定是有节点的)。也即是说，对于任意置换 σ，哈密顿量 (5.75) 的基态波函数必定满足

$$\psi_0(\mathbf{x}_{\sigma(1)},\mathbf{x}_{\sigma(2)},...,\mathbf{x}_{\sigma(N)})=\psi_0(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N), \tag{5.76}$$

从而必定是一个全对称波函数。

强调一下以上这些定理的适用范围：1. 它们只适用于空间波函数，因此不能包括自旋自由度。比方说，考虑到自旋自由度，一个理想的非相对论性氢原子的基态波函数也是二重简并的。2. 一般来说不能有磁场，因为否则哈密顿量 (5.75) 中出现的空间梯度 ∇ 就应该替换成 $\nabla-\frac{iq}{\hbar}\vec{A}(\mathbf{x})$(式中 $\vec{A}(\mathbf{x})$ 是矢量势，q 是粒子的电荷)，因此相应的定态波函数就无法取成实函数。3. 不适用于粒子数目超过一定限度的全同费米子体系。比方说对于三个 $1/2$ 自旋的全同费米子体系，由于它们的自旋波函数无法完全反对称化，从而为了满足全同性原理，这种体系的空间波函数也不能做到全对称化，即使对于基态波函数也一样，因此这种体系当然是不满足以上定理的。一般的，对于多费米子体系，即使有时候可以完全将自旋自由度和空间自由度分离开来，但由于半整数自旋的自旋波函数一般来说不能做到完全反对称化 (当粒子数目超过一定数量时)，因此它的空间波函数也不能是全对称的，从而这就违反了上面证明的定理。或者你也可以说，对于这样的体系，人们的确能求出哈密顿量 (5.75) 的一个全对称能量最低态波函数，但是由于这一波函数不能满足全同性原理，从而实际上是非物理的。4. 以上定理只适用于有限自由度的系统，在 $N\to+\infty$ 时可能不成立。5. 最后，严格来说，由于我们使用的能量泛函的表达形式没有包含边界项，因此这些定理只适用于束缚态，或者通过周期性边界条件将空间边界去除的情形。

两个例子

氢分子的基态波函数能很清楚地说明上面的这些定理。假设我们分别称氢分子的两个质子为 a 和 b，由于质子比电子重很多，我们不妨假设这两个质子是不动的，两者之间的距离为 R。两个电子，电子 1 和 2，绕着这两个质子运动，这就构成了一个氢分子系统。很明显，这个氢分子系统的哈密顿量就属于 (5.75) 这样的类型，现在我们来分析这个系统的基态。由于质子对电子的库伦吸引力 (质子与电子间的库伦势能为负)，为了保持最低的能量，电子 1 和 2 应该尽量靠近质子 a 和 b。但是，如果两个电子同时靠近一个质子的话，那这两个电子之间就离得太近了，这时候由于电子与电子之间的库伦排斥力 (电子与电子间的库伦势能为正)，系统的能量反而会比较高。所以为了尽量降低系统的能量，最后的情

[4]根据全同性原理，所有的电子都完全不可区分，因而是全同的。类似的，所有的光子也是全同的，所有的质子同样是全同的。全同粒子可以分成两类，像光子这样的自旋为整数的粒子是全同玻色子，像电子这样的自旋为半整数的粒子是全同费米子。对于全同玻色子体系，全同性原理要求系统的总波函数 (同时包括空间变量和自旋变量的波函数) 在这些全同粒子的任意置换下保持不变，人们称这样的波函数为全对称波函数。对于全同费米子体系，全同性原理要求系统的总波函数在这些全同费米子的奇置换下出负号，在偶置换下保持不变，人们称这样的波函数为全反对称波函数。有时候人们可以用全反对称的空间波函数乘上全对称的自旋波函数，或者用全对称的空间波函数乘上全反对称的自旋波函数来构造全同费米子的总波函数。

况一定是，每一个电子分别围绕一个质子。但这样一来，系统好像就变成两个独立的氢原子放在一起了，问题就在于，为什么这两个氢原子能进一步降低能量，进而结合成氢分子呢?

原因在于，两个电子在两个质子上的配置有两个可能性：电子 1 围绕质子 a 电子 2 围绕质子 b，以及电子 1 围绕质子 b 电子 2 围绕质子 a(如图 (5.5) 所示)，相应的波函数我们分别记为 $|A\rangle = \psi_{ab}(\mathbf{x}_1,\mathbf{x}_2)$ 和 $|B\rangle = \psi_{ab}(\mathbf{x}_2,\mathbf{x}_1)$，式中 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 分别是电子 1 和电子 2 的位置矢量。很显然，由于两个质子之间以及两个电子之间都没有什么区别，所以这两种可能



Figure 5.5: 氢分子的两种可能性

性的地位完全平等，它们对应的系统能量可以记为 E_0，当两个质子之间的距离足够大时，E_0 就等于两个独立的氢原子的能量之和。但是，$|A\rangle$ 和 $|B\rangle$ 都是一种过于简化的描述，它们都不是氢分子系统的基态，E_0 也不是基态的能量，不过系统的基态可以用一个以 $|A\rangle$ 和 $|B\rangle$ 为基的有效哈密顿量来描述。在这个有效描述中，$H_{AA} = H_{BB} = E_0$ 可以看作有效哈密顿量的对角项，但是，这个有效哈密顿量还有两个非 0 的非对角项 H_{AB} 和 H_{BA}。这是由于，每个电子除了要受到其围绕的质子的库伦吸引之外，它还要受到另一个距离较远的质子的库伦吸引，因此尽管电子与电子之间的库伦排斥力造成了一个不小的势垒，$|A\rangle$ 和 $|B\rangle$ 这两种可能性之间依然有一定的概率幅通过量子隧穿效应相互转变。和任何典型的隧穿效应一样，这种隧穿效应也会产生一个负的非对角有效哈密顿量，这就是 H_{AB} 和 H_{BA}，我们不妨记 $H_{AB} = H_{BA} = -J(J > 0)$。因此氢分子系统真正的基态可以通过求解下面的有效哈密顿量来得到，

$$\begin{pmatrix} E_0 & -J \\ -J & E_0 \end{pmatrix}, \tag{5.77}$$

这个哈密顿量的两个能量本征值分别为 $E_g = E_0 - J, E_e = E_0 + J$，相应的两个本征态分别为 $|g\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$，$|e\rangle = \frac{1}{\sqrt{2}}(|A\rangle - |B\rangle)$。或者也可以写成

$$|g\rangle = \frac{1}{\sqrt{2}}[\psi_{ab}(\mathbf{x}_1,\mathbf{x}_2) + \psi_{ab}(\mathbf{x}_2,\mathbf{x}_1)], \quad |e\rangle = \frac{1}{\sqrt{2}}[\psi_{ab}(\mathbf{x}_1,\mathbf{x}_2) - \psi_{ab}(\mathbf{x}_2,\mathbf{x}_1)]. \tag{5.78}$$

上面的分析告诉我们两个结论：第一，两个氢原子可以通过交换电子 (即上面描述的量子隧穿) 进一步将能量降低 $-J$，从而结合成一个氢分子。第二，氢分子的基态波函数 $|g\rangle = \frac{1}{\sqrt{2}}[\psi_{ab}(\mathbf{x}_1,\mathbf{x}_2) + \psi_{ab}(\mathbf{x}_2,\mathbf{x}_1)]$，很显然是一个对称波函数。正是在这里，这个结论和我们前面证明的一般性定理完全吻合。并且从这个基态波函数很容易看出，基态上的这两个电子实际上都是被两个质子共享的。

但是，由于电子是费米子，所以我们还得验证刚才得到的基态空间波函数 $|g\rangle$ 是否能满足全同性原理。答案是可以，因为现在只涉及到两个电子，所以对于对称的空间波函数，

我们完全可以通过将自旋波函数构造成反对称的自旋单态来满足全同性原理。因此上面的分析其实是告诉我们，氢分子基态上的两个电子其实处于自旋单态，并且由于它们被两个质子共享，所以就在两个质子间形成了一个典型的共价键 (如图 (5.6) 所示)，前面的 $-J$ 就是这个共价键的键能。而对于另外一个能量为 $E_0 + J$ 的态 $|e\rangle$，很显然它的空间波函数是反对称的，因此为了满足全同性原理，相应的电子自旋波函数就只能取自旋平行的对称三重态。因此，如果从电子自旋的视角来看待之前的有效哈密顿量，那它告诉我们的显然是，两电子自旋反平行 (自旋单态) 时能量较低，而自旋平行时能量较高。换言之，这个有效哈密顿量描述的是两电子自旋的反铁磁相互作用。

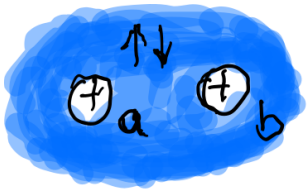

Figure 5.6: 氢分子

与氢分子这个例子略微有些相关的，我们举一个前面的那些定理不适用的例子，那就是铁磁体的基态。对于铁磁体而言，当它处于基态时，任意两个相邻格点上的电子自旋都是平行的 (为简单起见我们假设每一个格点上只有一个电子围绕着格点上的这个原子运动)，但这就是铁磁体难以理解的地方。因为根据我们上面对氢分子的分析，通常而言相邻两个电子自旋反平行才会使得系统能量比较低，为什么铁磁体的自旋反而会平行呢? 有一种物理上的解释是这样的：注意到铁磁体都是导体，所以除了围绕各自原子运动的电子之外，整个铁磁体中一定还存在传导电子，由于传导电子在整个铁磁体中运动，所以应该看成是被所有格点共享的。现在，由于泡利不相容原理，每个格点上的电子自旋会趋向于与传导电子的自旋反平行，但传导电子被所有格点共享了，因此这就意味着所有格点上的电子要自旋平行 (如图 (5.7) 所示)，这就导致了铁磁体的出现。这种自旋平行化的出现显然意味着相邻格点上的电子的空间波函数不可能是对称的，这正好是不满足我们前面证明的那些定理的一个例子。但这完全是因为我们现在处理的体系涉及到的电子 (费米子) 数目太多 (多于三个电子)，超出了定理的适用范围，属于我们前面已经注明过了的定理不适用的情形。

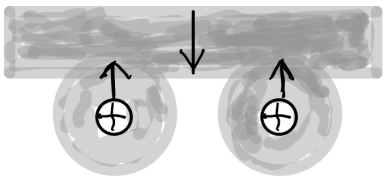

Figure 5.7: 铁磁体通过共享的传导电子而自旋平行

5.3.4 相变与对称性自发破缺

这一节我们将在变分法的近似下发展一个关于量子相变的物理图像。这一节的处理方式和例子基本上来自于文小刚老师关于拓扑序的一门在线课程的 PPT。

通常来说，一个量子系统总会依赖于一些外界参数，不妨象征性地将这些参数记为 g，比方说它可能是某个与系统耦合的外场等等。因此，这个量子系统的哈密顿量就将是 g 的光滑函数，不妨记为 H_g。之所以强调是光滑函数是因为任何物理上合理的哈密顿量都会有比较良好的数学性质，因此它们对物理参数的依赖都应该是光滑的。

现在假设我们用一组合理的试探波函数 $|\Psi_\phi\rangle$ 来近似求解 H_g 的基态能量，这里 ϕ 表示试探波函数里的未定参数。假设 $|\Psi_\phi\rangle$ 已经归一化好了，那么我们将得到一组能量泛函 $E_g(\phi)$，

$$E_g(\phi) = \langle\Psi_\phi|H_g|\Psi_\phi\rangle. \tag{5.79}$$

作为合理的试探波函数，它们的数学性质当然也应该比较良好，因此 $|\Psi_\phi\rangle$ 对参数 ϕ 的依赖也将是光滑的。因此，$E_g(\phi)$ 将同时是系统参数 g 和试探参数 ϕ 的光滑函数。

但是，$E_g(\phi)$ 还不是系统的基态能量。按照变分法，为了得到系统基态能量的近似值，我们还需要求 $E_g(\phi)$ 关于 ϕ 的最小值，假设这个最小值在 $\phi = \phi_m(g)$ 时取到，那基态能量的近似值 $E(g)$ 实际上是

$$E(g) = E_g(\phi_m(g)). \tag{5.80}$$

注意到 $\phi_m(g)$ 一般来说也依赖于系统参数 g(在物理上这就相当于在不同物理条件 g 之下系统的基态会不一样), 所以最后我们得到的基态能量近似值 $E(g)$ 作为 g 的函数就可能完全不同于原来的 $E_g(\phi)$。特别的，经过变分法的最终手续以后会产生这样一种可能性，即 $E(g)$ 作为 g 的函数不再是光滑的了，比方说它的某阶导数可能在某个参数点 g_c 变得不连续。在物理上，这就相当于说，随着外界物理参数的变化，系统的基态能量在 g_c 处出现了某种奇异性 (即使系统的哈密顿量 H_g 关于 g 光滑)，物理学家常常称这种奇异性的出现为系统发生了相变。

为了帮助读者看清这种可能性是如何发生的，我们不妨举一个简单的例子。假设有某个量子系统，我们用一组合理的试探波函数 $|\Psi_\phi\rangle$ 计算得出能量泛函为 $E_g(\phi) = (\phi^2 + g)^2$。很显然这个函数对于系统参数 g 和试探参数 ϕ 来说都光滑。现在让我们求 $E_g(\phi)$ 关于 ϕ 的最小值，很显然当 $g \geq 0$ 时，$\phi_m = 0$, 因此这时候有 $E(g) = g^2$，而当 $g < 0$ 时，$\phi_m = \pm\sqrt{-g}$, 很明显这时候 $E(g) = 0$。因此最终我们得到的基态能量函数 $E(g)$ 是关于 g 的一个分段光滑函数，很显然这个分段函数整体上是不光滑的，它的二阶导数在 $g = g_c = 0$ 处不连续。对于这样一个系统，我们就说它在 $g_c = 0$ 处发生了相变，由于是基态能量的二阶导数出现了不连续，所以这样的相变也被称为二级相变。

仔细观察上面这个例子出现奇异性 (即相变) 的基本机制。我们就会发现，关键点就在于随着参数 g 由大于 0 变为小于 0，$E_g(\phi)$ 关于 ϕ 的极小结构出现了突变。具体来说就是，当 $g \geq 0$ 时，$E_g(\phi)$ 关于 ϕ 只有一个极小点，即 $\phi_m = 0$, 但是当 $g < 0$ 时，$E_g(\phi)$ 关于

ϕ 的极小点劈裂成了两个，即 $\phi_m = +\sqrt{-g}$ 和 $\phi_m = -\sqrt{-g}$。正是这种极小结构的突变导致了二级相变。

不妨再举一个一级相变的例子，看看和二级相变有什么不同。比方说对于某个量子系统我们算出来 $E_g(\phi) = \phi^4/4 + g\phi^3/3 - \phi^2/2$, 很显然这个函数既是 g 的光滑函数又是 ϕ 的光滑函数。但其实这个系统在 $g = g_c = 0$ 处也有相变。为了看清楚这一点，我们不妨考察 g 在 0 附近时的情况，这时候，$E_g(\phi)$ 的两个极小点近似为 (保留到一阶小量) $\phi_{m1} \approx -\frac{g}{2} - 1$ 和 $\phi_{m2} \approx -\frac{g}{2} + 1$, 相应的两个极小值约为 (保留到一阶小量)$E_g(\phi_{m1}) \approx -\frac{1}{4} - \frac{1}{3}g$, $E_g(\phi_{m2}) \approx -\frac{1}{4} + \frac{1}{3}g$。显然，当 $g > 0$ 时，ϕ_{m1} 对应真正的最小值，因此这时 $E(g) \approx -\frac{1}{4} - \frac{1}{3}g$, 反过来，当 $g < 0$ 时，ϕ_{m2} 对应真正的最小值，因此这时 $E(g) \approx -\frac{1}{4} + \frac{1}{3}g$。很显然，这样算出来的基态能量 $E(g)$ 也是一个分段光滑函数，它在 $g = 0$ 处不光滑，因为它的一阶导数在 $g = 0$ 处不连续。人们常常把这样的一阶导数不连续的相变称作一级相变。观察这种相变发生的原因，我们会发现是因为 $E_g(\phi)$ 关于 ϕ 有多个极小值点，而随着系统参数 g 的变化 (由 $g > 0$ 变到 $g < 0$)，整体的最小值点在不同的极小之间发生了切换 (由 ϕ_{m1} 切换到了 ϕ_{m2})。这种极小切换就会导致一级相变。

但是，我们刚才给出来的两个例子都不是真正算出来的例子。也即是说，我们并没有给出实现这些例子的量子系统，而只是直接给出了两个能够说明相变现象的函数 $E_g(\phi)$。对真正量子系统的具体计算研究发现，对于有限的系统，比方说自由度数目有限的系统，只要试探波函数真正合理，最后得到的 $E(g)$ 实际上通常都是整体光滑的，通常不会有真正的量子相变。人们真正感兴趣的 $E(g)$ 的奇异性和量子相变其实都是发生在无限大的量子系统中。当然，对于无限大的量子系统，其基态能量 $E(g)$ 会正比于系统的体积 V，因此也是趋于无穷的，所以这时候更合适的研究对象其实应该是能量密度, 因为即使系统体积趋于无穷，其能量密度总还是有限的。也就是说，我们要定义一个能量密度泛函 $\varepsilon_g(\phi)$,

$$\varepsilon_g(\phi) = E_g(\phi)/V, \tag{5.81}$$

然后用 $\varepsilon_g(\phi)$ 来进行我们刚才所做过的那些分析。相应的基态能量密度 $\varepsilon(g)$ 就是

$$\varepsilon(g) = \varepsilon_g(\phi_m(g)), \tag{5.82}$$

$\varepsilon(g)$ 关于参数 g 的奇异性就意味着相变。当然，我们之前通过例子发现的相变机制依然可以照搬过来，只要将能量换成能量密度就可以了。值得说明的是，无限大系统并不是物理学家凭空构想出来的，实际上，任何一块宏观材料由于包含了巨量的原子分子，因此都可以近似地看成是无限大系统，要研究材料的宏观性质就必须研究无限大系统，这其实就是物理学家为什么对无限大系统感兴趣的基本原因。

下面我们举一个真正量子系统的例子。假设有 N 个 $1/2$ 自旋等间距地排布在一个圆周上 (因此第 $N+1$ 个自旋其实就是第 1 个自旋)。假设相邻两个自旋之间的相互作用为 $-\sigma_i^x\sigma_{i+1}^x$, 其中 $\sigma_i^x, \sigma_i^y, \sigma_i^z$ 是第 i 个自旋的三个泡利算符。假设整个系统处在一个 z 方向的磁场中，因此这个磁场和每一个自旋的相互作用能可以写成 $-g\sigma_i^z$，参数 g 正比于磁场强度，不妨假设 $g > 0$。因此，整个系统的哈密顿量 H_g 可以写成

$$H_g = -\sum_i (\sigma_i^x\sigma_{i+1}^x + g\sigma_i^z). \tag{5.83}$$

这样一个系统通常被称为横场伊辛模型，它实际上是一个可以精确求解的模型，不过我们这里是想用它来讨论变分法的应用，而不是要讨论它的精确求解。注意到在绕着 z 轴的 180 度旋转之下，$\sigma_i^z \to \sigma_i^z$, $\sigma_i^x \to -\sigma_i^x$, $\sigma_i^y \to -\sigma_i^y$。很显然 (5.83) 式给出的哈密顿量在这个 180 度旋转之下是不变的，因此这个操作是系统的一个对称性。实际上，这个绕 z 轴 180 度旋转的对称性可以用幺正变换 $U = \prod_i \exp(i\pi \frac{1}{2}\sigma_i^z)$ 来表示。

注意到 $g \to \infty$ 时系统的状态完全由 z 方向的磁场决定，因此这时候的基态为 $\prod_i(|\uparrow_i\rangle)$，而 $g = 0$ 时，系统有两个简并的基态 $\prod_i(|\rightarrow_i\rangle)$ 和 $\prod_i(|\leftarrow_i\rangle)$ (这里 $|\rightarrow_i\rangle = \frac{1}{\sqrt{2}}(|\uparrow_i\rangle + |\downarrow_i\rangle)$，$|\leftarrow_i\rangle = \frac{1}{\sqrt{2}}(|\uparrow_i\rangle - |\downarrow_i\rangle)$ 为 σ_i^x 的两个本征态)。所以对于一般情形的基态，我们可以取归一化的试探波函数 $|\Psi_\phi\rangle = \prod_i |\psi_i\rangle$, $|\psi_i\rangle = \cos\frac{\phi}{2}|\uparrow_i\rangle + \sin\frac{\phi}{2}|\downarrow_i\rangle$。注意到 $\sigma^z|\uparrow\rangle = |\uparrow\rangle, \sigma^z|\downarrow\rangle = -|\downarrow\rangle$, $\sigma^x|\uparrow\rangle = |\downarrow\rangle, \sigma^x|\downarrow\rangle = |\uparrow\rangle$，易有 $\langle\psi_i|\sigma_i^z|\psi_i\rangle = \cos^2\frac{\phi}{2} - \sin^2\frac{\phi}{2} = \cos\phi$, $\langle\psi_i|\sigma_i^x|\psi_i\rangle = 2\sin\frac{\phi}{2}\cos\frac{\phi}{2} = \sin\phi$。进而可以得到能量密度泛函，

$$\begin{aligned}\varepsilon_g(\phi) &= \langle\Psi_\phi|H_g|\Psi_\phi\rangle/N \\ &= -\sum_i \left[\langle\psi_i|\sigma_i^x|\psi_i\rangle\langle\psi_{i+1}|\sigma_{i+1}^x|\psi_{i+1}\rangle + g\langle\psi_i|\sigma_i^z|\psi_i\rangle\right]/N \\ &= -(\sin^2\phi + g\cos\phi). \end{aligned} \tag{5.84}$$

我们按照 g 从大到小的顺序画了几幅 $\varepsilon_g(\phi) = -(\sin^2\phi + g\cos\phi)$ 作为 ϕ 的函数图像 (如图 (5.8)、(5.9)、(5.10) 所示)。很显然，当 $g \geq 2$ 时，$\varepsilon_g(\phi)$ 只有一个极小值点，即 $\phi = 0$，因此这时候基态能量密度 $\varepsilon(g) = -g$。但是当 $g < 2$ 时，$\varepsilon_g(\phi)$ 的极小值点劈裂成了两个 $\phi = \phi_m = \arccos(\frac{g}{2})$，以及 $\phi = -\phi_m = -\arccos(\frac{g}{2})$，这时候 $\varepsilon(g) = -(1 + \frac{g^2}{4})$。这种极小劈裂的情况正符合我们前面关于二级相变机制的讨论。从基态能量密度的表达式也可以直接看出 $\varepsilon(g)$ 在 $g = 2$ 处的二阶导数不连续，因此正是一个二级相变。

这个例子其实揭示了由朗道发现的关于连续相变 (二级以上的相变) 的一个一般性机制，那就是，连续相变所需要的能量泛函极小结构的突变常常都是由于对称性自发破缺带来的。在我们这个例子中，在 $g > 2$ 的相，变分法求出来的系统基态为 $|\Psi_0\rangle = \prod_i(|\uparrow_i\rangle)$，这个态在绕着 z 轴的 180 度旋转下变换为 $|\Psi_0\rangle \to U|\Psi_0\rangle = \prod_i(i|\uparrow_i\rangle) = i^N|\Psi_0\rangle$(非指标的那个 i 为虚数单位)，可见变换前后的基态只差一个相位，在物理上它们描述的是同一个量子态，因此这也就是说，在 $g > 2$ 相中，基态保持了绕 z 轴的 180 度旋转对称性。但是，对于 $g < 2$ 的相，变分法求出来的基态是简并的两个 (这里严格来说要取 $N \to \infty$ 的无限大系统极限)，分别是 $|\Psi_{\phi_m}\rangle$ 和 $|\Psi_{-\phi_m}\rangle$，而且人们很容易计算出这两个态在绕 z 轴的 180 度旋转之下是如何变换的，$|\Psi_{\phi_m}\rangle \to U|\Psi_{\phi_m}\rangle = i^N|\Psi_{-\phi_m}\rangle$, $|\Psi_{-\phi_m}\rangle \to U|\Psi_{-\phi_m}\rangle = i^N|\Psi_{\phi_m}\rangle$。可见，这两个基态单独都不保持绕 z 轴的 180 度旋转对称性，而是在这个旋转操作下互相变为对方。但是，一个真实的系统只可能处在两个简并基态中的某一个，但是不管真实系统是处在 $|\Psi_{\phi_m}\rangle$ 还是 $|\Psi_{-\phi_m}\rangle$，它都破缺了这个绕 z 轴的 180 度旋转对称性。因此这就是一个典型的，哈密顿量有一个对称性，但是系统基态却破缺了这个对称性的例子，这就叫做对称性自发破缺。我们这个例子正好演示了二级相变和对称性自发破缺的密切联系，这种联系正是朗道首先指出的，朗道的这个理论后来成为了理解凝聚态系统不同相和相变的基础，因此成为凝聚态物理的一块基石，并且还给粒子物理学家带来了极大的启发，间接促成了基本粒子标准模型中的希格斯机制的提出。

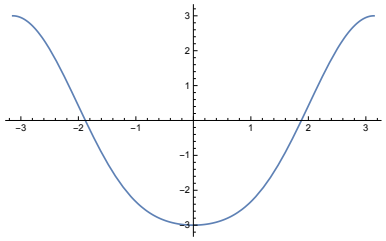

Figure 5.8: Fig of $-(\sin^2\phi+3\cos\phi)$

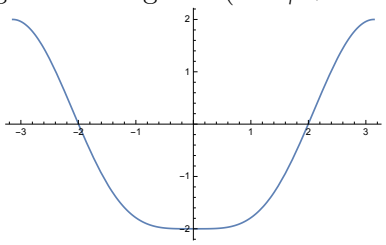

Figure 5.9: Fig of $-(\sin^2\phi+2\cos\phi)$

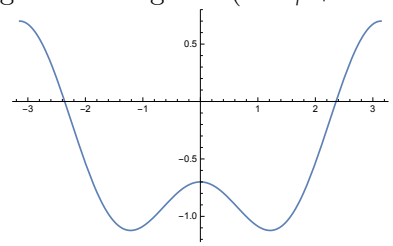

Figure 5.10: Fig of $-(\sin^2\phi+0.7\cos\phi)$

上面例子中对称相和对称破缺相的不同还可以通过一个物理量直接显示出来，那就是通过 σ_i^x 在基态上的期望值。利用 $\langle\Psi_\phi|\sigma_i^x|\Psi_\phi\rangle=\langle\psi_i|\sigma_i^x|\psi_i\rangle=\sin\phi$ 人们很容易得到，在对称相，$\langle\Psi_0|\sigma_i^x|\Psi_0\rangle=0$，而在对称破缺相 $\langle\Psi_{\pm\phi_m}|\sigma_i^x|\Psi_{\pm\phi_m}\rangle=\pm\sin\phi_m=\pm\sqrt{1-\frac{g^2}{4}}\neq 0$。在凝聚态物理中，通常把这样的能够区分对称相和对称破缺相的物理量称之为序参量。

最后我们简单提一下，横场伊辛模型的精确求解发现，真正的相变点其实在 $g=1$ 处，而不是变分法发现的 $g=2$ 处，而且这实际上是一个量子临界点。很显然，我们的变分法近似没有发现这一点。这主要是因为，在我们用变分法求横场伊辛模型基态的时候，我们是将所有的量子效应用一个整体的参数 ϕ 来平均了，而并没有将各种量子涨落包括进来，但是对于 $g=1$ 的量子临界点来说，各个尺度上的量子涨落都很重要，因此我们的变分法近似不能发现它也并不奇怪。这也清楚地说明了变分法毕竟只是一种近似方法，它并不能解释所有的量子现象。

5.3.5 习题

1. 请用试探波函数 $\psi_\lambda=e^{-\lambda r}$ 估算氢原子基态的能量，并和精确结果比较。

2. 假设在中心力场中运动的粒子受汤川势的相互作用，势能为

$$V(r) = -A\frac{e^{-\mu r}}{r}, \tag{5.85}$$

式中 $A>0$，从而这是一个吸引势。请通过取试探波函数 $\psi_\lambda = e^{-\lambda r}$ 确定这系统存在束缚态的一个充分条件。

3. 两个离子可以通过共享一个电子进而降低能量形成共价键。最简单的例子是氢分子离子 H_2^+，它是两个质子共享一个电子。假设质子静止不动，则系统的哈密顿量可以写成

$$H = \frac{\mathbf{p}^2}{2m} - \frac{e^2}{4\pi\varepsilon_0}\Big[\frac{1}{|\mathbf{x}|} + \frac{1}{|\mathbf{x}-\mathbf{R}|} - \frac{1}{R}\Big]. \tag{5.86}$$

式中 $\mathbf{p}$ 是电子的动量，$\mathbf{x}$ 是电子的位置矢量，$\mathbf{R}$ 是两个质子之间的距离。假设取试探波函数

$$\Psi_{\mathbf{R}} = \psi(\mathbf{x}) + \psi(\mathbf{x}-\mathbf{R}), \quad with \quad \psi(\mathbf{x}) = \sqrt{\frac{1}{\pi a_0^3}}e^{-r/a_0}. \tag{5.87}$$

并注意到下面的积分

$$\begin{aligned} u(R) &= \int d^3\mathbf{x}\psi(\mathbf{x})\psi(\mathbf{x}-\mathbf{R}) = \big(1+\frac{R}{a_0}+\frac{R^2}{3a_0^2}\big)e^{-R/a_0} \\ v(R) &= \int d^3\mathbf{x}\frac{\psi(\mathbf{x})\psi(\mathbf{x}-\mathbf{R})}{|\mathbf{x}|} = \frac{1}{a_0}\big(1+\frac{R}{a_0}\big)e^{-R/a_0} \\ w(R) &= \int d^3\mathbf{x}\frac{\psi^2(\mathbf{x})}{|\mathbf{x}-\mathbf{R}|} = \frac{1}{R} - \frac{1}{R}\big(1+\frac{R}{a_0}\big)e^{-2R/a_0}. \end{aligned} \tag{5.88}$$

请利用这些结果证明，哈密顿量在试探波函数上的平均能量 $E(R)$ 与氢原子基态的能量 E_0 之间的差可以写成

$$E(R) - E_0 = \frac{e^2}{4\pi\varepsilon_0}\big(\frac{1}{R} - \frac{v(R)+w(R)}{1+u(R)}\big). \tag{5.89}$$

请画出 $E(R)-E_0$ 作为横坐标 R 的函数的数值曲线，并以此估算氢分子离子的键能。[这个题目来自 David Tong 的课程习题]

6. * 量子力学中的对称性

本章我们首先利用维格纳的定理定义了量子变换的概念，它可以看成是经典力学中正则变换的量子力学对应物。然后我们研究了时间平移变换、空间平移变换、以及空间旋转变换，并以此一般性地定义了任何量子系统的总动量算符和总角动量算符。进而一般性地推导了这些算符之间的代数关系，特别是一般性地得到了角动量算符的代数关系。

本章的核心之一是介绍连续对称性与守恒定律之间的深刻联系。也就是量子力学版本的诺特定理。根据这个定理，能量守恒定律源自于时间平移对称性，动量守恒定律源自于空间平移对称性，而角动量守恒定律源自于空间旋转对称性。另外，电荷守恒定律和电荷量子化则源自于一个抽象的 $U(1)$ 相位对称性。当然，除此之外，我们也深入讨论了伽利略协变性。

我们也讨论了对称性与能级简并之间的深刻联系。

最后，本章还讨论了空间反演对称性以及相应的宇称守恒。特别的，我们介绍了李政道和杨振宁发现的弱相互作用下宇称不守恒。当然，我们也讨论了时间反演对称性，证明了 Kramer 定理。

描述对称性的基本数学语言是群论。本章我们也简要性地介绍了群论和群表示论的基本思想。其中群表示是本章的一个核心数学概念。

6.1 量子变换

一个量子系统所有可能量子态的集合构成希尔伯特空间 $\mathscr{H}$。在物理上，我们常常默认量子态已经归一化了，但即使对于已经归一化的量子态 $|\psi\rangle$，它其实也并不唯一确定，因为根据态叠加原理，$e^{i\theta}|\psi\rangle$ 和 $|\psi\rangle$ 描述的是同一个量子态！所以，严格来说，一个量子态其实是希尔伯特空间归一化矢量的一个等价类，满足 $e^{i\theta}|\psi\rangle \sim |\psi\rangle$，$\sim$ 表示等价关系。

根据量子力学的玻恩规则，对于处在 $|\psi\rangle$ 态上的系统，我们在 $|\phi\rangle$ 态上测到它的概率为 $|\langle\phi|\psi\rangle|^2$，其中两个量子态的内积 $\langle\phi|\psi\rangle$ 也称之为概率幅，因为它的模方给出了概率。在量子力学中，概率 (而不是量子态) 才是我们在物理上真正测量的东西。

量子变换

一个重要的问题是，在什么变换下，作为物理可观测量的概率将保持不变呢？对这个问题的回答就引出了我们所谓量子变换的概念[1]。所谓的量子变换我们指的是希尔伯特空间 $\mathscr{H}$ 到其自身的一个映射，假设在这个映射的作用下，$|\psi\rangle \to |\psi'\rangle$，$|\phi\rangle \to |\phi'\rangle$，那么我们要求映射前后的玻恩概率保持不变，即

$$|\langle\phi'|\psi'\rangle|^2 = |\langle\phi|\psi\rangle|^2. \tag{6.1}$$

维格纳 (Wigner) 有一个著名的定理[2]说，量子变换要么是希尔伯特空间的幺正变换，要么就是反线性反幺正变换，此外没有其它的可能性。幺正变换我们在第二章中就已经很熟悉了，最重要的例子比如有时间演化算符 $U(t_2,t_1)$。那么什么是反线性反幺正变换呢？首先，作用在希尔伯特空间的算符 Θ, 如果满足下面的关系，它就是一个**反线性算符**，

$$\Theta(\alpha|\psi\rangle + \beta|\phi\rangle) = \alpha^*\Theta|\psi\rangle + \beta^*\Theta|\phi\rangle, \tag{6.2}$$

式中 α,β 为两个复叠加系数。任何反线性反幺正变换 Θ 首先是一个反线性算符，其次，它还满足

$$\langle\Theta\phi|\Theta\psi\rangle = \langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*. \tag{6.3}$$

注意，对于反线性算符，我们不能将之理解为一个矩阵 (因为矩阵一定是线性的)！因此反线性算符只能向右边的量子态作用，诸如 $\langle\phi|\Theta$ 这样的表达式是没有意义的，因为它不能理解为行矢量乘上矩阵。这就是式 (6.3) 中我们要把第一个表达式写成 $\langle\Theta\phi|\Theta\psi\rangle$ 的原因，它表示 $\Theta|\psi\rangle$ 与 $\Theta|\phi\rangle$ 的内积。由于反线性算符不能理解为矩阵，所以它的厄密共轭的定义也与线性算符不同，而且反线性算符的厄密共轭也不能理解为复共轭加转置，因为它根本不能理解为矩阵，所以当然谈不上转置。反线性算符 Θ 的厄密共轭算符 $\Theta^\dagger$ 定义

[1] 值得提醒读者的是，量子变换这个名称还不是通用的，更多的书称之为对称变换。我们这里称之为量子变换是想稍微拓展一下这个概念的应用范围。

[2] 关于维格纳 (Wigner) 这个定理的严格证明，我们推荐读者参考：Steven Weinberg,《The Quantum Theory of Fields》，Volume I，第二章的 Appendix A。

为 (形式上和线性算符的厄密共轭定义相差一个整体的复数共轭，因为它的反线性本身包含了一个复数共轭)

$$\langle\phi|\Theta\psi\rangle = \langle\Theta^\dagger\phi|\psi\rangle^*. \tag{6.4}$$

根据这个定义 $\langle\Theta\phi|\Theta\psi\rangle = \langle\phi|\psi\rangle^* \Leftrightarrow \langle\phi|\Theta^\dagger\Theta\psi\rangle^* = \langle\phi|\psi\rangle^*$，即 $\Theta^\dagger\Theta = 1$，所以所谓的反幺正算符其实就是满足和幺正算符类似的 $\Theta^\dagger = \Theta^{-1}$ 的反线性算符。

很显然，幺正变换乘以幺正变换，结果必定还是幺正变换，而反幺正变换乘以反幺正变换结果也是幺正变换，只有幺正变换乘以反幺正变换结果才可能是反幺正变换。关于这几点的证明我们留给读者。正因为如此，任何反幺正变换总是能写成某个幺正变换和另一个反幺正变换的乘积。

操作诱导的量子变换

量子变换常常由对系统的操作诱导产生。比方说，我们将系统作一个旋转 R, 这就会诱导一个幺正的量子变换 $U(R)$，它描述系统旋转之后的量子态与旋转之前的量子态之间的联系。对于考察量子系统的对称性而言，我们所要考察的对系统的操作都是可逆的，而且满足一种乘法合成规则。也即是说，如果先对系统进行操作 T_1，接着再对它进行操作 T_2，最终的效果就等价于直接进行某个操作 T_3，我们将 T_3 写成 $T_3 = T_2T_1$。因此，两个先后进行的操作的合成就定义了一种操作间的乘法，其中先进行的操作写在乘号右边，后进行的操作写在左边。通常来说，这种乘法并不满足交换律 (类似于矩阵乘法)，比方说，先将系统绕 x 轴旋转 90 度 (记作 $R_x(90)$)，接着再将系统绕 y 轴旋转 90 度 (记作 $R_y(90)$)，其结果 $R_y(90)R_x(90)$ 显然与相反操作顺序的结果 $R_x(90)R_y(90)$ 不同。我们可以将一个操作 T 的逆操作记作 T^{-1}，按照操作乘法，它显然满足 $TT^{-1} = T^{-1}T = 1$。这里 1 就表示恒等操作，也就是不进行任何操作 (很显然，恒等操作所诱导的量子变换一定是恒等变换，也就是恒等算符)。这样的定义了乘法运算的可逆操作的集合就构成了数学上所谓的群。

抽象一点来说，一个群是一个集合，它的每一个元素就称为一个群元。群的要义有三点：其一，群元之间定义了某种乘法; 其二，每一个群元都有一个逆元；其三，群本身在乘法运算下是封闭的，即任何两个群元相乘结果必定依然是一个群元。根据这个定义，我们将限于考察那些能构成一个群的对量子系统的操作，以后我们都会默认这一点。

对于我们考察的这类操作而言，每一个对系统的操作都会诱导一个幺正或者反幺正的量子变换。假设记 T_1 操作的量子变换为 $U(T_1)$, T_2 操作的量子变换为 $U(T_2)$, T_1 和 T_2 的合成操作 T_2T_1 诱导的量子变换为 $U(T_2T_1)$。那么根据定义, 在 T_1 操作之下, 原来的量子态 $|\psi\rangle$ 将变换为 $|\psi\rangle \to U(T_1)|\psi\rangle$, 紧接着再进行操作 T_2，这个量子态就会接着变换为 $U(T_1)|\psi\rangle \to U(T_2)U(T_1)|\psi\rangle$, 所以先后进行 T_1、T_2 操作总的效果是将 $|\psi\rangle$ 变换为 $|\psi\rangle \to U(T_2)U(T_1)|\psi\rangle$。另一方面，根据定义，合成操作 T_2T_1 将会把 $|\psi\rangle$ 变换为 $|\psi\rangle \to U(T_2T_1)|\psi\rangle$。根据操作合成的定义，这两个结果应该是一样的，也即是说，我们应该有 $U(T_2T_1)|\psi\rangle = U(T_2)U(T_1)|\psi\rangle$。但是，这里有一个微妙的细节，由于量子态对应的希尔伯特空间矢量不唯一确定，而是可以相差一个相位，所以严格来说，我们不能说 $U(T_2T_1)|\psi\rangle$ 必须和 $U(T_2)U(T_1)|\psi\rangle$ 相等，而只能说，这两者最多相差一个相位，即 $U(T_2T_1)|\psi\rangle = e^{i\omega(T_2,T_1)}U(T_2)U(T_1)|\psi\rangle$, 由于 $|\psi\rangle$ 是

一个任意的态，因此我们就有

$$U(T_2T_1) = e^{i\omega(T_2,T_1)}U(T_2)U(T_1). \tag{6.5}$$

方程 (6.5) 最关键的地方就在于，它告诉我们，操作诱导的量子变换将保持操作的乘法结构！数学上人们称这种保持一个群的乘法结构的诱导关系为群的表示，因此操作所诱导的量子变换构成了操作群在希尔伯特空间的表示。同样，微妙点在于，这种对乘法结构的保持可以相差一个相位，所以人们通常称方程 (6.5) 这样的群表示关系为投影表示。读者无需追究投影这一名称的来源，只需记住投影表示满足 (6.5) 这样的方程就可以了。不过，很多情况下，投影表示的相因子 $\omega(T_2,T_1)$ 其实都可以取为零，我们将在本章的正文部分默认这一点，不过，在本章中间的补充内容中，我们考察了一个 $\omega(T_2,T_1)$ 不能取为零的例子[3]。

假设我们的操作 $T(\theta)$ 依赖于某个连续参数 θ, $\theta = 0$ 表示恒等操作。比方说，θ 可以是绕某个固定轴的旋转角度。如此一来我们就可以把这些操作所诱导的量子变换记作 $U(T(\theta))$, 又时候也简记为 $U(\theta)$。由于 $U(0) = 1$ 是一个幺正变换，而随着参数的连续调节，$U(\theta)$ 可以连续地变化为 $U(0) = 1$，所以，$U(\theta)$ 也必定为一个幺正变换 (反幺正变换涉及复共轭，而幺正变换不涉及，因此反幺正变换不可能连续地变化为幺正变换)。因此，依赖于连续参数并且包含恒等变换的量子变换必定为幺正变换，我们常常称这样的幺正变换为连续的幺正变换。而任何依赖于连续参数的反幺正变换往往总能写成一个连续的幺正变换乘以一个分立的反幺正变换。因此，在物理上真正独立存在的反幺正变换其实都是分立的，或者说离散的，在量子力学中，它其实就是所谓的时间反演变换，在量子场论中更著名的就是所谓的 CPT 变换。当然，反幺正变换是特殊的，因为它们并不是我们真正可以对量子系统实施的操作，我们定义它们完全是因为我们可以从中引出时间反演对称性或者 CPT 对称性的概念。

现在，假设连续参数 θ 为一个无穷小量，$\theta = \varepsilon$。因此幺正变换 $U(\varepsilon)$ 就可以泰勒展开为

$$U(\varepsilon) = 1 + i\varepsilon G + ... \tag{6.6}$$

其中算符 G 就称作连续幺正变换 $U(\theta)$ 的生成元。很显然，$U(\varepsilon)$ 也可以近似写成 $U(\varepsilon) = \exp(i\varepsilon G)$。利用幺正性 $U^\dagger(\varepsilon)U(\varepsilon) = 1$，我们容易有 $(1 - i\varepsilon G^\dagger + ...)(1 + i\varepsilon G + ...) = 1$, 比较等式左右两边的 ε 一次项，就有

$$G^\dagger = G. \tag{6.7}$$

也即是说，连续幺正变换的生成元必定为厄米算符，因此也就是一个物理量。

量子变换在算符上的作用

[3]关于 $\omega(T_2,T_1)$ 何时可以取为零的讨论，读者可以参考 Steven Weinberg,《The Quantum Theory of Fields》, Volume I，第二章的 2.7 节。

前面我们定义量子变换的时候，是将它作用在系统的量子态上。但其实，我们也可以等价地认为量子变换对系统量子态没有作用，而是对物理量的算符有作用。具体来说，考察一个量子变换 U，再考察一个任意的算符 A。在物理上，我们关心的其实总是算符的矩阵元，为此我们考察算符 A 在系统量子态 $|\psi\rangle$ 和 $|\phi\rangle$ 之间的矩阵元 $\langle\phi|A|\psi\rangle$。由于在量子变换之后 $|\psi\rangle \to U|\psi\rangle$, $|\phi\rangle \to U|\phi\rangle$，所以在量子变换的作用下矩阵元 $\langle\phi|A|\psi\rangle$ 将变换为

$$\langle\phi|A|\psi\rangle \to \langle\phi|U^{-1}AU|\psi\rangle. \tag{6.8}$$

很显然，为了满足这个物理矩阵元的变换关系，我们完全可以等价地认为量子变换对系统量子态没有作用，而是会将算符 A 变换为

$$A \to U^{-1}AU. \tag{6.9}$$

当然，严格来说，上面处理的只是幺正量子变换，对于反幺正量子变换 Θ, $|\psi\rangle \to \Theta|\psi\rangle$, $|\phi\rangle \to \Theta|\phi\rangle$, 情况略微有些不同。这时候，在变换之下 $\langle\phi|A|\psi\rangle \to \langle\Theta\phi|A|\Theta\psi\rangle$。由于 A 是一个普通的算符，所以我们有 $\langle\Theta\phi|A|\Theta\psi\rangle = \langle\Theta\psi|A^\dagger|\Theta\phi\rangle^*$。而根据 (6.4) 式，我们又有 $\langle\Theta\psi|A^\dagger|\Theta\phi\rangle = \langle\psi|(\Theta^\dagger A^\dagger\Theta)|\phi\rangle^*$。所以最终我们得到，在反幺正变换 Θ 之下，

$$\langle\phi|A|\psi\rangle \to \langle\psi|(\Theta^\dagger A^\dagger\Theta)|\phi\rangle. \tag{6.10}$$

我们发现，反幺正变换把左矢 $\langle\ |$ 变成了右矢 $|\ \rangle$，把右矢变成了左矢，同时它在算符上的等价作用为

$$A \to \Theta^{-1}A^\dagger\Theta, \tag{6.11}$$

式中我们利用了 $\Theta^\dagger = \Theta^{-1}$。很显然，如果 A 是厄密算符，那 (6.11) 式在形式上和幺正变换的 (6.9) 式就是一样的。

根据量子变换作用在系统量子态但是不作用在物理量算符上的观点，我们实际上是认为，我们的操作是“转动”了系统，但是没有“转动”实验室里测量系统的仪器，从而物理量算符不变。而根据量子变换作用在物理量算符而不作用在系统量子态上的观点，我们是等价地认为，我们的操作是将实验室里的仪器作了一个反方向的“旋转”，同时保持待测系统原地不动，从而算符按照 (6.9) 式变换，而系统的量子态不变[4]。

根据算符 A 的本征方程 $A|a\rangle = a|a\rangle$，结合 (6.9) 式容易看出来，在量子变换的作用下，算符 A 的本征值 a 不会变，但是本征态 $|a\rangle$ 将变为 $U^{-1}|a\rangle$(物理上这是因为仪器反方向“旋转”了)。如果对于一个原来处于 $|\psi\rangle$ 态的系统，我们计算测到 A 的值为 a 的概率 $p_a = |\langle a|\psi\rangle|^2$，那么在量子变换之下对于这个概率将如何变化我们有两种等价的观点：其一，是认为算符 A 不变 (从而它的本征态 $|a\rangle$ 也不变)，但是系统的量子态 $|\psi\rangle$ 变换为 $|\psi\rangle \to U|\psi\rangle$，根据这一观点 p_a 将变换为 $|\langle a|\psi\rangle|^2 \to |\langle a|U|\psi\rangle|^2$。其二，我们也可以认为系统的量子态 $|\psi\rangle$ 没变，但是作为物理量的算符 A 按照 (6.9) 式变换了，从而本征态 $|a\rangle$ 将

[4]其实还有一种与这两种观点并不等价但同样成立的观点，也是人们常常采用的，即认为我们同时将系统和仪器作了一个同方向的同步“转动”。当然很多书在采用这种观点的时候忽略了对它背后的这一物理本质的强调。读者可以自行写出在这种观点下，系统量子态和算符该如何变换的相应公式。

变为 $U^{-1}|a\rangle$，即 $\langle a|$ 将变为 $\langle a|U$，根据这一观点，p_a 依然将变换为 $|\langle a|\psi\rangle|^2 \to |\langle a|U|\psi\rangle|^2$。很显然，两种不同观点完全等价。

方程 (6.9) 可以让我们看到，量子变换其实是经典力学中的正则变换在量子力学中的对应物。为了看清楚这一点，我们假设 U 是一个连续的幺正变换 $U(\theta)$，我们来考察无穷小变换 $U(\varepsilon) = \exp(i\varepsilon G/\hbar)$。记某物理量 A 在这一无穷小量子变换的作用下变换为 $A \to A' = U^{-1}(\varepsilon)AU(\varepsilon)$, 记变换前后物理量 A 的无穷小改变为 $\delta A = A' - A$。则根据 (6.9) 式，我们将有 $\delta A = (1 - i\varepsilon G/\hbar + ...)A(1 + i\varepsilon G^\dagger/\hbar + ...) - A = \varepsilon[G,A]/(i\hbar)$, 即

$$\delta A = \varepsilon[G,A]/(i\hbar). \tag{6.12}$$

熟悉分析力学的读者很快就能明白，只要我们将算符对易子 $[G,A]/(i\hbar)$ 对应于分析力学中的泊松括号 $[G,A]_{PB}$, 那么方程 (6.12) 就正好是分析力学中的无穷小正则变换的方程[5]

$$\delta A = \varepsilon[G,A]_{PB}. \tag{6.13}$$

其中 G 就叫做无穷小正则变换的生成元。

在分析力学中，如果一个正则变换保持系统的哈密顿量不变，我们就称它为力学系统的一个对称性。如果力学系统有一个连续的对称性从而使得哈密顿量在无穷小正则变换 G 下保持不变 (为简单起见，假设 G 不显含 t，即 $\frac{\partial G}{\partial t} = 0$)，即 $\delta H = \varepsilon[G,H]_{PB} = 0$，从而 $[G,H]_{PB} = 0$，则根据力学量 G 的哈密顿运动方程 $\frac{dG}{dt} = [G,H]_{PB}$, 我们必有 $\frac{dG}{dt} = 0$，也即是说连续对称性的生成元 G 必定是一个守恒量。这就是分析力学中连续对称性与守恒量之间的密切联系，通常人们是在拉格朗日量的框架下讨论它的，称之为诺特定理。后文中我们将看到连续对称性与守恒量之间的这一普遍联系在量子力学中依然成立。

最后，我们来初步验证一下算符对易子 $[G,A]/(i\hbar)$ 与泊松括号 $[G,A]_{PB}$ 之间的对应关系 (如果是从经典泊松括号出发再对应到量子的算符对易子，人们就称这样的过程为正则量子化)。我们以一维运动的粒子为例，这时候在量子力学中我们有算符对易关系 $[X,P] = i\hbar$, 而在分析力学中我们有泊松括号 $[X,P]_{PB} = 1$，很显然 $[X,P]/(i\hbar) = [X,P]_{PB}$。所以，对于这一简单的 $G = X$, $A = P$ 的情形我们就验证了量子与经典之间的这一对应关系，更一般地，注意到分析力学中的所有物理量都是正则坐标和正则动量的函数，人们可以进一步证明，对于任何这样的物理量 G 和 A，在 $\hbar \to 0$ 的经典极限下均有 $[G,A]/(i\hbar) = [G,A]_{PB}$。由于本书并不想强调正则量子化的观点，也不想深入探讨量子力学与经典力学之间的对应关系，所以我们将忽略这个数学证明。

6.1.1 习题

1. 设 A、B 为两个算符，λ 为一个参数，请证明

$$e^{\lambda A}Be^{-\lambda A} = B + [A,B]\lambda + \frac{1}{2!}[A,[A,B]]\lambda^2 + \frac{1}{3!}[A,[A,[A,B]]]\lambda^3 + ... \tag{6.14}$$

[5]参见 Goldstein H. Classical Mechanics. Third ed. 第 9 章-9.6 节.

6.2 时间平移、空间平移、空间旋转

这一节我们考察对量子系统的三种常见操作，时间平移、空间平移，以及空间旋转。我们主要考察这三种时空操作所诱导的连续幺正量子变换。

6.2.1 时间平移

我们来考察对量子系统的时间平移操作 T_τ，假设在这样的时间平移操作之下，t 时刻的系统被平移到 $t+\tau$ 时刻，即系统的时间坐标从 t 平移到了 $t+\tau$，$T_\tau : t \to T_\tau t = t+\tau$。我们记表示这个时间平移操作的量子变换为 $U(\tau)$。很显然时间平移操作具有可加性，因此

$$U(\tau_1)U(\tau_2) = U(\tau_1+\tau_2). \tag{6.15}$$

具有这种性质的算符 $U(\tau)$ 必定具有如下形式

$$U(\tau) = \exp\big(iH\tau/\hbar\big), \tag{6.16}$$

其中厄米算符 H 就是时间平移的生成元，其物理意义我们马上就会清楚。

现在，我们将时间平移算符 $U(\tau)$ 作用在系统 t 时刻的某个态 $|\psi(t)\rangle$ 上，并记这种量子变换的结果为一个新的态 $|\phi(t)\rangle$, $|\phi(t)\rangle = U(\tau)|\psi(t)\rangle$。由于 $U(\tau)$ 相应于把 t 时刻的系统平移到 $t+\tau$ 时刻，因此平移以后 t 时刻的态 $|\phi(t)\rangle$ 必定就是平移前 $t-\tau$ 时刻的态 (此前 $t-\tau$ 时刻的系统刚好平移到 t 时刻)，也就是 $|\psi(t-\tau)\rangle$，因此我们有

$$U(\tau)|\psi(t)\rangle = |\psi(t-\tau)\rangle = |\psi(T_\tau^{-1}t)\rangle. \tag{6.17}$$

在上式中取 $\tau = t$, 我们就可以得到

$$|\psi(t)\rangle = \exp\big(-iHt/\hbar\big)|\psi(0)\rangle, \tag{6.18}$$

很显然，这正是量子态的时间演化方程，从而前面引入的时间平移生成元 H 就是系统的哈密顿算符!

我们把通常的算符称之为薛定谔绘景中的算符，比如本书前面碰到的位置算符、动量算符等等都是薛定谔绘景中的算符。薛定谔绘景中的算符通常不随时间演化 (除非这个算符的表达式本身显含时间 t), 随时间演化的是量子态。但是，考虑一个薛定谔绘景中的算符 A 在 t 时刻的某个矩阵元 $\langle\psi_2(t)|A|\psi_1(t)\rangle$, 利用量子态的时间演化关系我们发现这个矩阵元可以重写为

$$\begin{aligned}\langle\psi_2(t)|A|\psi_1(t)\rangle &= \langle\psi_2(0)|\exp\big(iHt/\hbar\big)A\exp\big(-iHt/\hbar\big)|\psi_1(0)\rangle \\ &= \langle\psi_2(0)|A_H(t)|\psi_1(0)\rangle.\end{aligned} \tag{6.19}$$

式中我们引入了一个新的算符 $A_H(t)$, 其定义为

$$A_H(t) = \exp\big(iHt/\hbar\big)A\exp\big(-iHt/\hbar\big). \tag{6.20}$$

方程 (6.19) 告诉我们，对于计算真正可测量的物理量的矩阵元而言，我们既可以用通常的薛定谔绘景中的算符 A，然后认为是量子态在随时间演化，也可以等价地认为量子态永远

都固定在 $t=0$ 时刻不变，但是物理量要重新表示成一个新的随时间演化的算符 $A_H(t)$。如果将系统所有的算符都作 (6.20) 式这样的变换，那我们就进入了所谓的海森堡绘景，在海森堡绘景中，量子态不随时间演化，但是算符要按照 (6.20) 式这样随时间演化。$A_H(t)$ 这样的算符就称为海森堡绘景中的算符。

值得强调的是，薛定谔绘景和海森堡绘景是量子力学的两种等价描述。应用的时候我们选取其中之一就可以了。

有了海森堡绘景，我们就可以将时间平移的量子变换 $U(\tau)$ 等价地作用在算符上了。事实上，在海森堡绘景中，由于量子态不随时间演化，所以再将时间平移算符作用上去就没有意义了。相反，这时候随时间演化的是 $A_H(t)$ 这样的算符。所以我们应该等价地将时间平移的量子变换作用在 $A_H(t)$ 上，变换以后的新算符 $B_H(t)$ 就是 $B_H(t)=U^{-1}(\tau)A_H(t)U(\tau)$。完全类似于前面关于量子态的讨论，由于量子变换 $U(\tau)$ 是将 $t-\tau$ 时刻的系统平移到 t 时刻，所以变换之后 t 时刻的算符 $B_H(t)$ 应该等于变换之前 $t-\tau$ 时刻的算符，也就是 $A_H(t-\tau)$，因此即有

$$U^{-1}(\tau)A_H(t)U(\tau)=A_H(t-\tau)=A_H(T_\tau^{-1}t). \tag{6.21}$$

在上式中取 $\tau=t$，我们马上即有

$$A_H(t)=\exp\big(iHt/\hbar\big)A_H(0)\exp\big(-iHt/\hbar\big). \tag{6.22}$$

这正是海森堡绘景的算符随时间演化的关系。

6.2.2 空间平移

下面我们考察对系统的空间平移操作 $T_{\mathbf{a}}$，$T_{\mathbf{a}}$ 的定义是将系统的空间坐标平移 $\mathbf{a}$，具体来说就是，在 $T_{\mathbf{a}}$ 的作用之下，系统中任意一个位于 $\mathbf{x}$ 的粒子将被平移到 $\mathbf{x}+\mathbf{a}$，$T_{\mathbf{a}}:\mathbf{x}\to T_{\mathbf{a}}\mathbf{x}=\mathbf{x}+\mathbf{a}$。我们记空间平移操作诱导的量子变换为 $U(\mathbf{a})$。根据狭义相对论，时间坐标和空间坐标会构成一个四维矢量，注意到两个四矢量的内积具有 $a_0b_0-\mathbf{a}\cdot\mathbf{b}$ 的形式，而由于时间平移操作 $U(\tau)$ 具有 $\exp(i\tau H/\hbar)$ 的形式，所以空间平移 $U(\mathbf{a})$ 必定具有 $\exp(-i\mathbf{a}\cdot\mathbf{P}/\hbar)$ 的形式，即

$$U(\mathbf{a})=\exp(-i\mathbf{a}\cdot\mathbf{P}/\hbar). \tag{6.23}$$

其中厄米算符 $\mathbf{P}$ 就是空间平移的生成元。而且，由于时间平移的生成元 H 是系统的总能量算符，所以空间平移生成元 $\mathbf{P}$ 必定是系统的总动量算符，这是因为根据狭义相对论，总能量和总动量会构成一个四维矢量。一个系统的总动量算符是空间平移的生成元，这就是量子物理中对系统动量的一般性定义，它不仅适用于非相对论粒子的量子力学，而且在量子场论中也依然成立。

根据矢量的合成规则我们知道 $U(\mathbf{a})U(\mathbf{b})=U(\mathbf{a}+\mathbf{b})=U(\mathbf{b})U(\mathbf{a})$，也即

$$U(\mathbf{a})U(\mathbf{b})=U(\mathbf{b})U(\mathbf{a}). \tag{6.24}$$

式中 $\mathbf{a}$、$\mathbf{b}$ 为任意矢量。分别取 $\mathbf{a}=\varepsilon_1$、$\mathbf{b}=\varepsilon_2$ 为两个无穷小矢量，从而有 $U(\varepsilon_1)U(\varepsilon_2)=U(\varepsilon_2)U(\varepsilon_1)\Rightarrow(1-i\varepsilon_1\cdot\mathbf{P}/\hbar)(1-i\varepsilon_2\cdot\mathbf{P}/\hbar)=(1-i\varepsilon_2\cdot\mathbf{P}/\hbar)(1-i\varepsilon_1\cdot\mathbf{P}/\hbar)$, 从而即有 $(\varepsilon_1\cdot\mathbf{P})(\varepsilon_2\cdot\mathbf{P})=(\varepsilon_2\cdot\mathbf{P})(\varepsilon_1\cdot\mathbf{P})$, 由于 ε_1、ε_2 为两个任意的无穷小矢量，从而即有

$$[P_i,P_j]=0. \tag{6.25}$$

式中 $P_i,i=1,2,3$ 代表总动量算符 $\mathbf{P}$ 的三个直角坐标分量。对于点粒子情形，方程 (6.25) 是我们早就知道的，但是，上面给出的是一个最具一般性的推导，即使在量子场论中它依然成立。

下面我们回到非相对论点粒子情形，这时候粒子的位置坐标就成了一个定义良好的算符 (在量子场论中，由于粒子数不守恒，粒子可以产生湮灭，一个粒子可以转变成多个不同粒子，因此粒子的位置坐标不再是定义良好的算符)，为了简单起见，不妨假设整个系统只有一个粒子，记它的位置算符为 $\mathbf{X}$，位置坐标本征值为 $\mathbf{x}$，相应的坐标表象波函数为 $\psi(\mathbf{x})=\langle\mathbf{x}|\psi\rangle$。则根据空间平移的定义 $\phi(\mathbf{x})=U(\mathbf{a})\psi(\mathbf{x})$ 为空间平移以后的粒子波函数，由于平移以后 $\mathbf{x}$ 点的粒子必定来自于平移之前 $\mathbf{x}-\mathbf{a}$ 点的粒子，所以 $\phi(\mathbf{x})$ 必定等于平移之前 $\mathbf{x}-\mathbf{a}$ 点的波函数 $\psi(\mathbf{x}-\mathbf{a})$，从而即有

$$U(\mathbf{a})\psi(\mathbf{x})=\psi(\mathbf{x}-\mathbf{a})=\psi(T_{\mathbf{a}}^{-1}\mathbf{x}). \tag{6.26}$$

等价的，我们也可以认为空间平移算符 $U(\mathbf{a})$ 作用在粒子的坐标算符 $\mathbf{X}$ 上，作用以后 $\mathbf{X}$ 将变换为 $\mathbf{X}\to U^{-1}(\mathbf{a})\mathbf{X}U(\mathbf{a})$。我们知道，这种作用相当于将测量粒子坐标的坐标系 (仪器的一部分) 向反方向平移 $-\mathbf{a}$，这时候粒子坐标将变换为 $\mathbf{x}\to\mathbf{x}+\mathbf{a}$，因此位置算符的变换关系实际上必然为 $\mathbf{X}\to\mathbf{X}+\mathbf{a}$。也即是说，我们必有

$$U^{-1}(\mathbf{a})\mathbf{X}U(\mathbf{a})=\mathbf{X}+\mathbf{a}. \tag{6.27}$$

同时，由于空间平移对于粒子的动量不会有任何影响，所以我们还必然有

$$U^{-1}(\mathbf{a})\mathbf{P}U(\mathbf{a})=\mathbf{P}. \tag{6.28}$$

由于动量算符都相互对易，所以 $\mathbf{P}$ 与 $U(\mathbf{a})=\exp(-i\mathbf{a}\cdot\mathbf{P}/\hbar)$ 必然相互对易，所以 (6.28) 式将自动得到满足。为了看清 (6.27) 式会告诉我们什么结果，我们取 $\mathbf{a}=\varepsilon$ 为一个无穷小矢量，从而 (6.27) 式等价于 $(1+i\varepsilon\cdot\mathbf{P}/\hbar)\mathbf{X}(1-i\varepsilon\cdot\mathbf{P}/\hbar)=\mathbf{X}+\varepsilon$，比较这个式子左右两边的一阶无穷小项，就有 $i[\varepsilon\cdot\mathbf{P}/\hbar,\mathbf{X}]=\varepsilon$, 注意到 ε 为任意无穷小量，所以这个结果等价于

$$[X_i,P_j]=i\hbar\delta_{ij}. \tag{6.29}$$

这当然是我们熟知的量子力学对易关系，我们只是用一种更具一般性的方式再一次导出它而已。在坐标表象中，为了满足这样的对易关系，我们可以取 $\mathbf{P}=-i\hbar\nabla$(注意我们这里假定只有一个粒子，因此系统的总动量也就是这个粒子的动量)，从而 $U(\mathbf{a})=\exp(-\mathbf{a}\cdot\nabla)$，因此如果方程 (6.26) 能够满足，那我们将有

$$\exp(-\mathbf{a}\cdot\nabla)\psi(\mathbf{x})=\psi(\mathbf{x}-\mathbf{a}). \tag{6.30}$$

但这其实是一个数学恒等式，为了看清楚这一点，我们将 $\exp(-\mathbf{a}\cdot\nabla)$ 泰勒展开为 $\exp(-\mathbf{a}\cdot\nabla)=\sum_{n=0}^{+\infty}\frac{1}{n!}(-\mathbf{a}\cdot\nabla)^n$，从而上面等式的左边就成为 $\exp(-\mathbf{a}\cdot\nabla)\psi(\mathbf{x})=\sum_{n=0}^{+\infty}\frac{1}{n!}(-\mathbf{a}\cdot\nabla)^n\psi(\mathbf{x})$，根据函数的泰勒展开公式，这个结果正好等于右边的 $\psi(\mathbf{x}-\mathbf{a})$。从而方程 (6.26) 的确能够成立。

6.2.3 空间旋转

角动量算符代数关系

现在我们来考察对系统的空间旋转操作 R，它将会把系统内的某点 $\mathbf{x}$ 旋转到 $\mathbf{x}'=R\mathbf{x}$，$\mathbf{x}'$ 和 $\mathbf{x}$ 之间满足如下关系

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=\begin{pmatrix}&R&\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}\tag{6.31}$$

式中 R 代表 3×3 的正交旋转矩阵，满足 $R^TR=RR^T=1$，$(x,y,z)^T$ 代表 $\mathbf{x}$ 的三个坐标分量，$(x',y',z')^T$ 代表 $\mathbf{x}'$ 的三个坐标分量。也即是说，旋转操作和 3×3 的正交旋转矩阵一一对应。先后两个旋转操作的合成就对应相应两个正交旋转矩阵的矩阵相乘，而所有旋转操作所构成的旋转群，就对应于所有正交旋转矩阵按照矩阵乘法所构成的群，通常记这个群为 $SO(3)$ 群，称为三维空间旋转群。

我们记旋转操作 R 所诱导的幺正量子变换为 $U(R)$。首先考察绕 z 轴的 θ 角旋转 $R_z(\theta)$。假设先绕 z 轴旋转 θ_1，接着再绕 z 轴旋转 θ_2，显然总的效果相当于绕 z 轴旋转 $\theta_1+\theta_2$，因此我们有 $R_z(\theta_2)R_z(\theta_1)=R_z(\theta_1+\theta_2)$。类似的，对于相应的希尔伯特空间幺正量子变换而言，必有

$$U\big(R_z(\theta_2)\big)U\big(R_z(\theta_1)\big)=U\big(R_z(\theta_1+\theta_2)\big).\tag{6.32}$$

这种可加性意味着，$U\big(R_z(\theta)\big)$ 必定可以写成

$$U\big(R_z(\theta)\big)=\exp\big(-i\theta J_z/\hbar\big),\tag{6.33}$$

式中厄密算符 J_z 为绕 z 轴旋转的生成元，为了看清楚 J_z 的物理含义，我们注意到这里的绕 z 轴的转角 θ 可以看成是一个广义坐标，$R_z(\theta)$ 就可以看成是沿着这个坐标的“平移”，如此一来 $U\big(R_z(\theta)\big)$ 就是这个“平移”所诱导的量子变换，从而根据上一节对空间平移的讨论，$U\big(R_z(\theta)\big)$ 必定具有 (6.33) 式这样的形式，式中的 J_z 就是与绕 z 轴的这个坐标 θ 相对应的系统“总动量”，那也就是绕 z 轴的总角动量，因为与角度坐标对应的“动量”其实就是角动量。所以，绕 z 轴旋转的量子变换的生成元 J_z 就是系统绕 z 轴的总角动量！

类似的，我们也可以分别考察系统绕 x 轴和绕 y 轴的旋转，并得到

$$U\big(R_x(\theta)\big)=\exp\big(-i\theta J_x/\hbar\big),\quad U\big(R_y(\theta)\big)=\exp\big(-i\theta J_y/\hbar\big).\tag{6.34}$$

式中生成元 J_x、J_y 分别是系统绕 x 轴和绕 y 轴的总角动量。这其实就是关于一个量子系统的总角动量算符的一般性定义，它不仅适应于单个粒子，也适应于多粒子体系，甚至适用于量子场论。

我们很容易写出 $R_x(\theta)$、$R_y(\theta)$ 和 $R_z(\theta)$ 的具体表达式。比方说，$R_z(\theta)$ 的表达式为

$$R_z(\theta)=\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}. \tag{6.35}$$

类似的，结合 x,y,z 坐标轴的右手关系，人们也很容易写出 $R_x(\theta)$ 和 $R_y(\theta)$。

为了导出角动量算符所满足的代数关系式，我们现在来考察无穷小转动，即假设 $\theta=\varepsilon$ 为一无穷小量。则保留到 ε 的二阶项为止，我们有

$$R_z(\varepsilon)=1+\varepsilon\begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}+\varepsilon^2 C_z. \tag{6.36}$$

等式右边第一项的 1 代表 3×3 的单位矩阵，等式右边第三项的 C_z 是一个与 ε 无关的 3×3 矩阵，其具体形式与我们将要进行的推导无关。类似的，我们有

$$R_x(\varepsilon)=1+\varepsilon\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{pmatrix}+\varepsilon^2 C_x. \tag{6.37}$$

以及

$$R_y(\varepsilon)=1+\varepsilon\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{pmatrix}+\varepsilon^2 C_y. \tag{6.38}$$

利用这些式子我们容易将表达式 $R_y(-\varepsilon)R_x(-\varepsilon)R_y(\varepsilon)R_x(\varepsilon)$ 计算到 ε 的平方阶（更高阶全部忽略），结果为

$$R_y(-\varepsilon)R_x(-\varepsilon)R_y(\varepsilon)R_x(\varepsilon)=R_z(-\varepsilon^2). \tag{6.39}$$

我们知道旋转操作所诱导的幺正量子变换必定构成旋转群的表示，也即是说，幺正量子变换 $U(R)$ 将保持矩阵 R 之间的乘法关系。从而根据上面的 (6.39) 式，我们必定有

$$U\big(R_y(-\varepsilon)\big)U\big(R_x(-\varepsilon)\big)U\big(R_y(\varepsilon)\big)U\big(R_x(\varepsilon)\big)=U\big(R_z(-\varepsilon^2)\big). \tag{6.40}$$

注意到 $\exp\big(-i\varepsilon J/\hbar\big)=1-i\varepsilon J/\hbar-\frac{1}{2}\varepsilon^2J^2/\hbar^2+\ldots$，代入上面的 (6.40) 式，并比较等号两边关于 ε^2 的项，我们就可以得到

$$[J_x,J_y]=i\hbar J_z. \tag{6.41}$$

利用三个直角坐标间的轮换关系 $x\to y,y\to z,z\to x$，我们可以类似地得到

$$[J_y,J_z]=i\hbar J_x,\quad [J_z,J_x]=i\hbar J_y. \tag{6.42}$$

利用列维-西维塔符号 ε_{ijk} 以及求和约定 (后文也将一直使用求和约定)，我们也可以将这几个式子统一写成

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k. \tag{6.43}$$

这就是**角动量算符的基本代数关系**。注意，和总动量算符不同，总角动量算符不同分量之间是不可对易的。

矢量算符

我们知道，旋转操作 R 所诱导的量子变换 $U(R)$ 可以等价地认为是作用在物理量算符上。假设我们有一个矢量算符 $\mathbf{V}$, 其三个直角坐标分量为 $V_i, i = 1,2,3$，则在量子变换 $U(R)$ 的作用下这些分量将变换为 $V_i \to U^{-1}(R)V_iU(R)$。在物理上，这也就是保持量子系统不动，而将我们观测系统的仪器反向旋转，特别是将测量系统各物理矢量分量的直角坐标系反向旋转。由于这样的坐标系反向旋转，矢量分量 $V_i, i = 1,2,3$ 将变换为，$V_i \to R_{ij}V_j$(式中 R_{ij} 是 3×3 的旋转矩阵 R 的矩阵元), 从而我们必有

$$U^{-1}(R)V_iU(R) = R_{ij}V_j. \tag{6.44}$$

由于各种书的定义略有区别, 初学者往往搞不清, (6.44) 这样的式子是应该写成 (6.44) 式这种形式还是应该写成 $U^{-1}(R)V_iU(R) = (R^{-1})_{ij}V_j$ 的形式 (或者等价地 $U(R)V_iU^{-1}(R) = R_{ij}V_j$ 的形式)。下面我们提供一个实质性的数学判断方式。

假设我们将表达式 (6.44) 错误地写成了

$$U^{-1}(R)V_iU(R) = (R^{-1})_{ij}V_j, \tag{6.45}$$

我们该如何发现错误呢? 方法很简单，我们可以考虑接连进行的两个旋转操作 $U(R_1)$ 和 $U(R_2)$，根据错误表达式 (6.45)，在 $U(R_1)$ 作用下我们将有

$$U^{-1}(R_1)V_iU(R_1) = (R_1^{-1})_{ij}V_j, \tag{6.46}$$

接着再对这个式子进行 $U(R_2)$ 作用，从而就有

$$\begin{aligned} &U^{-1}(R_2)U^{-1}(R_1)V_iU(R_1)U(R_2)\\ =&(R_1^{-1})_{ij}U^{-1}(R_2)V_jU(R_2), \end{aligned} \tag{6.47}$$

再次利用 (6.45) 式就有

$$\begin{aligned} &U^{-1}(R_2)U^{-1}(R_1)V_iU(R_1)U(R_2)\\ =&(R_1^{-1})_{ij}U^{-1}(R_2)V_jU(R_2)\\ =&(R_1^{-1})_{ij}(R_2^{-1})_{jk}V_k, \end{aligned} \tag{6.48}$$

整理一下，我们可以将这个式子重写为

$$\big(U(R_1)U(R_2)\big)^{-1}V_i\big(U(R_1)U(R_2)\big) = \big((R_2R_1)^{-1}\big)_{ik}V_k, \tag{6.49}$$

利用 $U(R_1)U(R_2)=U(R_1R_2)$, 这也就是

$$\big(U(R_1R_2)\big)^{-1}V_i\big(U(R_1R_2)\big)=\big((R_2R_1)^{-1}\big)_{ik}V_k. \tag{6.50}$$

但这个最终的式子肯定是错的！原因在于 $U(R)$ 构成了旋转群的表示，而表示的要义就在于保持乘法关系，但是从 (6.50) 中我们清楚地看到，等式左边两个旋转矩阵的乘法关系是 R_1R_2，但右边的乘法关系却是 R_2R_1，两边的乘法关系反过来了，这就不符合群表示的数学定义，因此 (6.50) 式是错的！从而作为出发点的 (6.45) 式也必定是错的！

相反，完全类似的推导过程，但是假如我们的出发点是正确的 (6.44) 式，那最终我们将得到

$$\big(U(R_1R_2)\big)^{-1}V_i\big(U(R_1R_2)\big)=(R_1R_2)_{ik}V_k. \tag{6.51}$$

很显然，这里乘法关系完全对得上，没有任何错误。总之，在我们确立 (6.44) 式这一类方程时，实质性的要点是要确保 $U(R)$ 构成一个群表示！

回到 (6.44) 式。我们取 R 为绕 z 轴的无穷小旋转 $R_z(\varepsilon)$, 利用 $U\big(R_z(\varepsilon)\big)=1-i\varepsilon J_z/\hbar+\ldots$, 我们就可以由 (6.44) 式得到

$$(1+i\varepsilon J_z/\hbar)V_i(1-i\varepsilon J_z/\hbar)=R_z(\varepsilon)_{ij}V_j. \tag{6.52}$$

代入 (6.36) 式，并比较上面方程左右两边关于 ε 的项，就可以得到

$$[J_z,V_x]=i\hbar V_y,\quad [J_z,V_y]=-i\hbar V_x,\quad [J_z,V_z]=0. \tag{6.53}$$

同理，我们也可以进一步考虑绕 x 轴和绕 y 轴作无穷小旋转的情形，最终的所有结果可以归纳如下

$$[J_i,V_j]=i\hbar\varepsilon_{ijk}V_k. \tag{6.54}$$

(6.54) 式就是矢量算符与角动量算符之间的代数关系式，反过来，满足这样的代数关系的算符 V_i 就是矢量算符。比较 (6.54) 式和 (6.43) 式，我们发现，角动量算符 J_i 本身其实就是一种特殊的矢量算符，这正好印证了角动量是一个矢量。从而我们可以对角动量算符使用矢量形式 $\mathbf{J}$。进而我们又可以一般性地把绕 $\mathbf{n}$ 轴 ($\mathbf{n}$ 为单位矢量) 旋转的量子变换写成

$$U(\mathbf{n},\theta)=\exp(-i\theta\mathbf{n}\cdot\mathbf{J}/\hbar). \tag{6.55}$$

另外，我们也知道，一个系统的总动量算符 $\mathbf{P}$ 必定是矢量算符，从而我们必有

$$[J_i,P_j]=i\hbar\varepsilon_{ijk}P_k. \tag{6.56}$$

轨道角动量

以上所有代数关系都是一般性的，我们前面谈论的始终是一般性的抽象的角动量算符。现在我们开始来研究一些具体的角动量算符。首先是轨道角动量算符。我们将以单粒

子体系为例，这时候我们其实已经知道轨道角动量算符的表达式为 $\mathbf{L}=\mathbf{X}\times\mathbf{P}$。但这是沿用了经典力学中对轨道角动量的定义。现在我们有了关于角动量的一般性量子力学定义，我们想试试能不能从这些一般性的量子力学结果中推导出这个轨道角动量算符的表达式。为此，我们将具体地选取坐标表象来处理问题。

在坐标表象中，当我们将系统旋转 R 时，系统的波函数将变换为 $\psi(\mathbf{x})\to\psi'(\mathbf{x})=U(R)\psi(\mathbf{x})$。旋转以后 $\mathbf{x}$ 点的波函数 $\psi'(\mathbf{x})$ 必定来源于旋转之前 $R^{-1}\mathbf{x}$ 点的波函数 $\psi(R^{-1}\mathbf{x})$ (因为原来的 $R^{-1}\mathbf{x}$ 点刚好旋转到 $\mathbf{x}$ 点)，即 $\psi'(\mathbf{x})=\psi(R^{-1}\mathbf{x})$，从而我们必有

$$U(R)\psi(\mathbf{x})=\psi(R^{-1}\mathbf{x}). \tag{6.57}$$

人们可以验证这一方程的确符合 $U(R)$ 是旋转群的表示。相反，如果人们将这样的方程写成 $U(R)\psi(\mathbf{x})=\psi(R\mathbf{x})$, 那就会和群表示相矛盾!

现在，我们考虑将系统绕 $\mathbf{n}$ 轴旋转一个无穷小角度 ε，则粒子的坐标将旋转为 $\mathbf{x}\to\mathbf{x}+\varepsilon\mathbf{n}\times\mathbf{x}$。根据 (6.55) 式，这时候相应的幺正变换为 $U(\mathbf{n},\varepsilon)=\exp(-i\varepsilon\mathbf{n}\cdot\mathbf{J}/\hbar)=1-i\varepsilon\mathbf{n}\cdot\mathbf{J}/\hbar+...$, 从而方程 (6.57) 将变成

$$(1-i\varepsilon\mathbf{n}\cdot\mathbf{J}/\hbar)\psi(\mathbf{x})=\psi(\mathbf{x}-\varepsilon\mathbf{n}\times\mathbf{x}). \tag{6.58}$$

根据泰勒展开，这个等式右边是 $\psi(\mathbf{x}-\varepsilon\mathbf{n}\times\mathbf{x})=\psi(\mathbf{x})-(\varepsilon\mathbf{n}\times\mathbf{x})\cdot\nabla\psi(\mathbf{x})=\psi(\mathbf{x})-\varepsilon\mathbf{n}\cdot(\mathbf{x}\times\nabla)\psi(\mathbf{x})$, 与等式 (6.58) 左边相比较我们就能得到，$\mathbf{J}=-i\hbar\mathbf{x}\times\nabla$, 这正是我们想要推导的轨道角动量算符的表达式，不过为了强调它是轨道角动量算符，我们常常将之记为 $\mathbf{L}$，从而我们就用量子力学的办法推导出了

$$\mathbf{L}=-i\hbar\mathbf{x}\times\nabla. \tag{6.59}$$

算符 $\mathbf{L}$ 既然是角动量算符，那它当然满足角动量算符的一般代数关系 (6.43)，我们可以将这个代数关系重写如下

$$[L_i,L_j]=i\hbar\varepsilon_{ijk}L_k. \tag{6.60}$$

不过，在这一特殊情况下，这一代数关系也可以根据 $\mathbf{L}=\mathbf{X}\times\mathbf{P}$ 的构造直接验证。

自旋角动量

我们知道除了轨道角动量以外，电子还有一个内禀的自旋角动量，相应的角动量算符为 $\mathbf{S}$，它当然也满足代数关系 (6.43)，通常写作 $[S_i,S_j]=i\hbar\varepsilon_{ijk}S_k$。实际上我们在第二章中就是从这一代数关系出发开始讨论电子自旋算符和泡利算符的。不过，当时我们无法给出这一代数关系的推导，这个缺陷到现在才算补上。

假设我们仅仅关心电子的自旋，那么量子态就一定能写成 $|\psi\rangle=\psi_\uparrow|\uparrow\rangle+\psi_\downarrow|\downarrow\rangle$ 的形式，式中 $\psi_\uparrow,\psi_\downarrow$ 是两个复的叠加系数。根据 (6.55) 式，当我们将系统绕 $\mathbf{n}$ 轴旋转 θ 角时，自旋量子态 $|\psi\rangle$ 将变换为

$$|\psi\rangle\to\exp\big(-i\theta\mathbf{n}\cdot\mathbf{S}/\hbar\big)|\psi\rangle=\exp\big(-i\frac{1}{2}\theta\mathbf{n}\cdot\vec{\sigma}\big)|\psi\rangle. \tag{6.61}$$

当然，我们也可以在 $\{|\uparrow\rangle,|\downarrow\rangle\}$ 表象中将电子的自旋态表示成列矢量 $\binom{\psi_\uparrow}{\psi_\downarrow}$，人们通常称这一两分量列矢量为旋量，或者自旋波函数。很显然，当我们将系统绕 $\mathbf{n}$ 轴旋转 θ 角时，旋量将变换为

$$\begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix}\to \exp\big(-i\theta\mathbf{n}\cdot\mathbf{S}/\hbar\big)\begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix}=\exp\big(-i\frac{1}{2}\theta\mathbf{n}\cdot\vec{\sigma}\big)\begin{pmatrix}\psi_\uparrow\\ \psi_\downarrow\end{pmatrix}. \tag{6.62}$$

当然，现在式中的 $\mathbf{S}$ 应该理解为 2×2 的表示矩阵，式中的 $\vec{\sigma}$ 应该理解为泡利矩阵。我们之所以没有更换记号是因为人们很容易根据上下文确定这些东西是抽象的算符还是矩阵。

根据我们在第三章习题中的一个结果，

$$\exp\big(-i\frac{1}{2}\theta\mathbf{n}\cdot\vec{\sigma}\big)=\cos(\frac{\theta}{2})-i(\mathbf{n}\cdot\vec{\sigma})\sin(\frac{\theta}{2}). \tag{6.63}$$

因此，$\exp\big(-i\frac{1}{2}2\pi\mathbf{n}\cdot\vec{\sigma}\big)=-1!$, 也就是说，对于电子自旋态而言，旋转 2π 角会出一个负号，而不是不变。只有旋转 4π 角才能保证自旋态不变，因为 $\exp\big(-i\frac{1}{2}4\pi\mathbf{n}\cdot\vec{\sigma}\big)=1$。也即是说，考虑到自旋，电子要旋转 720 度 (而不是 360 度) 才能回到原初的状态！从经典物理的角度来看，这当然是不可思议的。

电子总角动量

现在，我们同时考虑电子的自旋自由度和空间自由度，根据第 3 章中所学的知识我们知道，这时候电子的波函数应该推广成两分量的旋量波函 $\Psi(\mathbf{x})=\binom{\psi_\uparrow(\mathbf{x})}{\psi_\downarrow(\mathbf{x})}$。注意，旋量波函数 $\Psi(\mathbf{x})$ 有双重身份，它既是一个旋量，同时又是空间坐标的波函数。因此，结合 (6.57) 式和 (6.62) 式我们可以知道，当我们将系统绕 $\mathbf{n}$ 轴旋转 θ 角时，旋量波函数将变换为

$$\begin{aligned}\Psi(\mathbf{x})\to U(\mathbf{n},\theta)\Psi(\mathbf{x})&=\exp(-i\theta\mathbf{n}\cdot\mathbf{J}/\hbar)\Psi(\mathbf{x})\\&=\exp\big(-i\theta\mathbf{n}\cdot\mathbf{S}/\hbar\big)\Psi\big(R^{-1}(\mathbf{n},\theta)\mathbf{x}\big).\end{aligned} \tag{6.64}$$

式中 $R(\mathbf{n},\theta)$ 表示绕 $\mathbf{n}$ 轴旋转 θ 角的 3×3 正交旋转矩阵。

如果我们取 $\theta=\varepsilon$ 为一个无穷小旋转，那根据上面的 (6.64) 式，我们就有

$$(1-i\varepsilon\mathbf{n}\cdot\mathbf{J}/\hbar)\Psi(\mathbf{x})=(1-i\varepsilon\mathbf{n}\cdot\mathbf{S}/\hbar)\Psi(\mathbf{x}-\varepsilon\mathbf{n}\times\mathbf{x}). \tag{6.65}$$

将 $\Psi(\mathbf{x}-\varepsilon\mathbf{n}\times\mathbf{x})$ 泰勒展开，并比较上面式子左右两边的 ε 相关项，我们就可以得到

$$\mathbf{J}=\mathbf{S}+(-i\hbar\mathbf{x}\times\nabla)=\mathbf{S}+\mathbf{L}. \tag{6.66}$$

这个式子告诉我们，电子的总角动量其实是轨道角动量与自旋角动量的和！当然，其实自旋角动量 $\mathbf{S}$ 和轨道角动量 $\mathbf{L}$ 作用的对象并不相同，自旋角动量是作用在两分量旋量空间，而轨道角动量是作用在坐标空间。

6.2.4 习题

1. 空间标度变换可以定义为 $\mathbf{x}\to e^{\lambda}\mathbf{x}$，其中 λ 为一个实参数。(1) 请证明相应的希尔伯特空间量子变换可以写成 $e^{-i\lambda D}$ 的形式，其中 D 称为尺度变换生成元。(2) 请导出对易子 $[D,\mathbf{P}]$ 和对易子 $[D,\mathbf{J}]$。

2. 在球坐标中，波函数可以写成 $\psi(r,\theta,\phi)$ 的形式，根据绕 z 轴的旋转变换在波函数上的作用

$$e^{-i\varphi L_z/\hbar}\psi(r,\theta,\phi)=\psi(r,\theta,\phi-\varphi), \tag{6.67}$$

请导出球坐标中轨道角动量算符 $L_z=-i\hbar\frac{\partial}{\partial\phi}$。

3. 假设我们定义空间平移算符 $U(\mathbf{a})$ 在位置本征态上的作用为 $U(\mathbf{a})|\mathbf{x}\rangle=|\mathbf{x}+\mathbf{a}\rangle$。并定义 $U(\mathbf{a})\psi(\mathbf{x})=\langle\mathbf{x}|U(\mathbf{a})|\psi\rangle$。请据此从数学上证明：(1) $U(\mathbf{a})\psi(\mathbf{x})=\psi(\mathbf{x}-\mathbf{a})$。(2) $U^{-1}(\mathbf{a})\mathbf{X}U(\mathbf{a})=\mathbf{X}+\mathbf{a}$。

4. 我们知道空间旋转变换在电子自旋量子态上的作用为 $\exp\left(-i\frac{1}{2}\theta\mathbf{n}\cdot\vec{\sigma}\right)$，但是我们也可以利用三个欧拉角将任意一个旋转变换写成，$\exp\left(-i\frac{1}{2}\alpha\sigma_z\right)\exp\left(-i\frac{1}{2}\beta\sigma_y\right)\exp\left(-i\frac{1}{2}\gamma\sigma_z\right)$，请在数学上将这两种不同的写法联系起来。

5. 对于轨道角动量算符 $\mathbf{L}=\mathbf{X}\times\mathbf{P}$，请利用对易关系 $[X_i,P_j]=i\hbar\delta_{ij}$ 验证角动量的代数关系 $[L_i,L_j]=i\hbar\varepsilon_{ijk}L_k$。

6. 我们可以将 3×3 的正交旋转矩阵 $R(\mathbf{n},\theta)$ 写成 $R(\mathbf{n},\theta)=\exp(-i\theta\mathbf{n}\cdot\mathscr{L}/\hbar)$。请证明：(1)$(\mathscr{L}_i)_{jk}=-i\hbar\varepsilon_{ijk}$。(2) 请验证矩阵 $\mathscr{L}$ 满足角动量代数关系 $[\mathscr{L}_i,\mathscr{L}_j]=i\hbar\mathscr{L}_k$。(3) 请证明坐标的无穷小旋转 $\mathbf{x}\to\exp(-i\varepsilon\mathbf{n}\cdot\mathscr{L}/\hbar)\mathbf{x}$ 可以写成 $\mathbf{x}\to\mathbf{x}+\varepsilon\mathbf{n}\times\mathbf{x}$。

6.3 对称性

到现在为止，虽然我们已经讲述了空间平移和空间旋转等操作，以及它们诱导的量子变换。但是我们并没有处理空间平移对称性和空间旋转对称性，实际上，我们还没有涉及对称性。那么，什么是量子系统的对称性呢？首先，对称性当然相应于对量子系统进行一个变换，因此对称性必须对应于量子变换。其次，这样的量子变换必须保持系统的动力学规律不变，我们称之为对称变换。如果我们讨论的是包含时间平移对称性在内的时空对称性，那这第二点要求常常可以放松为，这样的量子变换构成时空对称群的表示。这一要求的含义是，我们的量子系统的动力学规律在时空对称操作下是协变的 (而不一定不变)。但是，在这一节的正文中，我们讨论的实际上都是动力学规律不变的情形，不过在后面的补充内容中，我们讨论了一个动力学规律协变的例子，即非相对论系统的伽利略推动对称性。

6.3.1 对称性与守恒定律

在量子力学中，所谓的动力学规律，其实就是指量子态的演化规律。这里我们假设系统的哈密顿量 H 不显含 t，从而量子态从 0 时刻到 t 时刻的演化规律就是

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle. \tag{6.68}$$

假设 $U(T)$ 为系统的对称变换 (这一对称变换由对系统的对称操作 T 诱导)，在它的作用下 $|\psi(0)\rangle$ 将变换为 $|\psi'(0)\rangle$，$|\psi'(0)\rangle = U(T)|\psi(0)\rangle$，类似的，$|\psi(t)\rangle$ 将变换为 $|\psi'(t)\rangle$ (从而 $U(T)$ 不包含时间反演，关于时间反演对称性我们后面会单独讨论)，$|\psi'(t)\rangle = U(T)|\psi(t)\rangle$。所谓的对称变换保持系统的动力学规律不变，指的就是变换以后的 $|\psi'(0)\rangle$ 和 $|\psi'(t)\rangle$ 之间同样按照下式演化

$$|\psi'(t)\rangle = e^{-iHt/\hbar}|\psi'(0)\rangle. \tag{6.69}$$

显然这个方程意味着 $U(T)|\psi(t)\rangle = e^{-iHt/\hbar}U(T)|\psi(0)\rangle$，代入 $|\psi(t)\rangle$ 的演化规律即有，$U(T)e^{-iHt/\hbar}|\psi(0)\rangle = e^{-iHt/\hbar}U(T)|\psi(0)\rangle$，也即

$$U(T)e^{-iHt/\hbar} = e^{-iHt/\hbar}U(T). \tag{6.70}$$

也即是说，对称变换需要和时间演化算符对易。将 $e^{-iHt/\hbar}$ 泰勒展开到 t 的一阶项并代入 (6.70) 式就有

$$U(T)H = HU(T). \tag{6.71}$$

即对称变换必定和哈密顿算符对易，或者等价地说哈密顿算符在对称变换下不变，即 $U^{-1}(T)HU(T) = H$。

很明显，时间平移算符 $U(\tau)$ 与 $e^{-iHt/\hbar}$ 显然对易。因此这就告诉我们，哈密顿量不显含 t 的系统必定有时间平移对称性。而且很显然，对于这样的系统，如果我们计算哈密顿算符 H(也即能量算符) 在任意两个态 $|\psi(t)\rangle$ 和 $|\phi(t)\rangle$ 上的矩阵元，那么显然有 $\langle\phi(t)|H|\psi(t)\rangle = \langle\phi(0)|H|\psi(0)\rangle$ 不依赖于时间 t，也即是说，系统的能量是守恒的。看出这一点的更快方式是注意到 H 在海森堡绘景里面依然是 H，从而海森堡绘景的能量算符 H 不随时间演化，这就说明了系统的能量守恒。总之，有时间平移对称性的系统，就必定有能量守恒。

时间平移对称性当然是连续对称性。现在，假设我们考察的对称性是一个任意的连续对称性，我们考虑无穷小对称变换 $U(\varepsilon) = 1 + i\varepsilon G + ...$，式中 G 为这个连续对称变换的生成元。代入 (6.71) 式，并比较方程两边有关 ε 的项，我们马上就能得到

$$[G, H] = 0. \tag{6.72}$$

即连续对称性的生成元必定与系统的哈密顿算符对易。当然，正如我们会在补充内容中看到的，对于对称变换使得动力学规律协变的情形，情况可能会有些不同。(6.72) 式就告诉我们，算符 G 在海森堡绘景中依然为 G，从而海森堡绘景中的连续对称性生成元不随时间

演化，这就意味着，G 是一个守恒量。因此，任何连续对称性的生成元都是守恒量。连续对称性与守恒量之间有对应关系，这当然就是量子力学版本的诺特定理！

因此，如果系统有空间平移对称性，那么必定有

$$[\mathbf{P},H]=0, \tag{6.73}$$

同时系统总动量 $\mathbf{P}$ 必定是守恒量。反过来也一样，如果 $[\mathbf{P},H]=0$，那系统就必定有空间平移对称性，这是因为这时候空间平移变换 $U(\mathbf{a})=e^{-i\mathbf{a}\cdot\mathbf{P}/\hbar}$ 必定和时间演化算符对易。同样的，如果系统有空间旋转对称性，那就必定有

$$[\mathbf{J},H]=0, \tag{6.74}$$

同时系统的总角动量 $\mathbf{J}$ 必定守恒。反过来，如果 $[\mathbf{J},H]=0$，那系统就必定有空间旋转对称性。

下面我们给出一个具体的具有空间平移对称性的例子。考虑一个多粒子体系，假设不存在外场，并且粒子间的相互作用势能仅仅依赖于它们的相对位置，也就是仅仅依赖于粒子之间的坐标差。一个具体的例子是具有如下哈密顿量的多粒子体系

$$H=\sum_n\frac{\mathbf{P}_n^2}{2m_n}+\sum_{n<m}V(\mathbf{X}_n-\mathbf{X}_m), \tag{6.75}$$

式中 $\mathbf{X}_n$ 为第 n 个粒子的位置算符，$\mathbf{P}_n$ 为第 n 个粒子的动量算符。由于空间平移算符 $U(\mathbf{a})=e^{-i\mathbf{a}\cdot\mathbf{P}/\hbar}$($\mathbf{P}=\sum_n\mathbf{P}_n$ 为系统总动量算符) 有如下代数关系

$$U^{-1}(\mathbf{a})\mathbf{X}_nU(\mathbf{a})=\mathbf{X}_n+\mathbf{a},\quad U^{-1}(\mathbf{a})\mathbf{P}_nU(\mathbf{a})=\mathbf{P}_n. \tag{6.76}$$

所以很显然，$U^{-1}(\mathbf{a})HU(\mathbf{a})=\sum_n\frac{\mathbf{P}_n^2}{2m_n}+\sum_{n<m}V\big((\mathbf{X}_n-\mathbf{a})-(\mathbf{X}_m-\mathbf{a})\big)=H$，从而这样的系统具有空间平移不变性，其总动量 $\mathbf{P}$ 必定守恒。

至于具有空间旋转对称性的例子，氢原子就是一个很好的例子。这是因为氢原子的库伦势场是一个中心力场，它是球对称的，因此当然在空间旋转之下保持不变。从而对于氢原子，必定有 $[\mathbf{L},H]=0$, 角动量必定守恒。这其实就是为什么我们在求解氢原子的时候用角动量量子数 l 和 m 来标记系统量子态的原因，因为角动量守恒意味着这些量子数是所谓的“好量子数”。

实际上，任何一个孤立系统都有时间平移对称性、空间平移对称性以及空间旋转对称性，从而其总能量、总动量以及总角动量都必然是守恒的。能量守恒、动量守恒、以及角动量守恒其实源于时空的性质，因此它们才是自然界中最为基本的规律。

6.3.2 相位不变性与电荷守恒

除了能量守恒、动量守恒以及角动量守恒之外，自然界的另一条基本守恒定律是电荷守恒。为了从基本的微观层次讨论电荷守恒的起源，我们需要从微观粒子谈起，物质世界是由微观粒子组成的，而每一个微观粒子都具有一个确定的电荷。用量子力学的话来说，微观粒子是电荷的本征态。假设我们考虑一个任意的多粒子态，它由带电量分别为 $q_1,q_2,...q_N$ 的 N 个粒子组成[6]，现在我们忽略每一个粒子的其它量子数，将这个多粒子态

[6]在量子场论中，由于真空极化所产生的屏蔽效应，我们在不同的距离上测量粒子的电荷结果其实稍有区别，这里的电荷是指在离粒子足够远时测得的。

简记为 $|q_1,q_2,...q_N\rangle$(粒子的种类和数目 N 都可变), 它当然是电荷的一个本征态。实际上，任何多粒子态都是电荷的本征态，电荷本身已经成了我们对微观粒子定义的一部分。

现在，对于任意多粒子态 $|q_1,q_2,...q_N\rangle$，我们可以定义一个相位变换

$$|q_1,q_2,...q_N\rangle \to e^{iq_1\alpha}e^{iq_2\alpha}...e^{iq_N\alpha}|q_1,q_2,...q_N\rangle. \tag{6.77}$$

我们将会看到，如果这样的相位变换是世界的对称性，那就必然有电荷守恒。为了看清楚这一点，我们定义电荷算符在任意多粒子态 $|q_1,q_2,...,q_N\rangle$ 上的作用为

$$Q|q_1,q_2,...q_N\rangle = (\sum_{n=1}^{N} q_n)|q_1,q_2,...q_N\rangle, \tag{6.78}$$

这样的式子其实就给出了电荷算符 Q 的定义。根据 (6.78) 式，我们可以将相位变换 (6.77) 重写成 $|q_1,q_2,...q_N\rangle \to e^{i\alpha Q}|q_1,q_2,...q_N\rangle$, 这样重写以后的方程显然对于所有的多粒子态都成立，因此它也对希尔伯特空间的任意量子态 $|\psi\rangle$ 成立 (因为世界由微观粒子组成，任意量子态都必定是一些多粒子态的线性叠加)，从而原来的相位变换 (6.77) 即相应于希尔伯特空间的幺正变换

$$|\psi\rangle \to e^{i\alpha Q}|\psi\rangle. \tag{6.79}$$

如果我们的世界在形如 (6.77) 式的相位变换下对称，那就等价于说，幺正变换 (6.79) 是世界的一个对称性。由于相位参数 α 连续可变，所以这当然是一个连续对称性。根据我们在上一小节中关于对称性与守恒定律的讨论。这一连续对称性就意味着

$$[Q,H] = 0. \tag{6.80}$$

从而电荷算符 Q 是一个守恒量。

由于 Q 是守恒量，在时间演化之下，它的本征值必定要保持不变。假设某个时间演化过程使得多粒子态 $|q_1,q_2,...q_N\rangle$ 演化成了多粒子态 $|q'_1,q'_2,...,q'_M\rangle$。那 Q 的本征值保持不变就意味着 $\sum_{n=1}^{N} q_n = \sum_{m=1}^{M} q'_m$。这就是我们通常所说的电荷守恒。总之，我们对电荷守恒起源的解释是，因为我们的世界具有 $e^{i\alpha Q}$ 的对称性，或者说我们的世界具有某种相位对称性。

那们，为什么我们世界的电荷是量子化的呢? 也就是说，为什么所有带电粒子的电荷都是某个最小电荷单元的整数倍呢? 解释也很简单，取决于对称变换 $e^{i\alpha Q}$ 的相位参数，如果相位参数 α 的取值范围是整个实轴，那就不会有什么电荷量子化。这时候我们称相应的相位对称性为具有对称群 $\mathbb{R}^*$，$\mathbb{R}^*$ 就是非零实数的集合，它在实数乘法下构成一个群，$\mathbb{R}^*$ 的群元就是 e^{α}。但是，如果相位参数 α 具有某种周期性，比方说周期为 $2\pi/e$, 那就必然有电荷量子化。因为这时候我们将有 $e^{i(2\pi/e)Q} = e^{i0\cdot Q} = 1$。从而 Q 的本征值只可能取 $ne, n \in \mathbb{Z}$, 这就是电荷量子化，e 就是最小的电荷单元! 这时候我们称相应的相位对称性为具有对称群 $U(1)$，它的群元是 $e^{ie\alpha}$。

所以，之所以我们有电荷守恒和电荷量子化，是因为我们的世界具有某种 $U(1)$ 对称性!

6.3.3 对称性与能级简并

前面我们学习氢原子的时候曾经发现，不考虑自旋，氢原子的能级简并度为 n^2。这些能级简并从何而来呢？回答是，从对称性中来。

为了理解这一点，假设我们考虑一个量子系统的某个能级 E, E 就是这个能级的能量本征值，本征方程为

$$H|\psi_E\rangle = E|\psi_E\rangle, \tag{6.81}$$

式中 $|\psi_E\rangle$ 为任意一个本征值为 E 的本征态。假设系统有对称性 $U(T)$, T 表示对系统的对称操作，它们构成一个对称群 $\mathscr{G}$，$U(T)$ 就是这个对称群在希尔伯特空间的表示。当然，$HU(T) = U(T)H$。

现在，我们将 $U(T)$ 作用在 $|\psi_E\rangle$ 上，得到一个新的态 $U(T)|\psi_E\rangle$，由于我们有 $H\big(U(T)|\psi_E\rangle\big) = U(T)H|\psi_E\rangle = E\big(U(T)|\psi_E\rangle\big)$，所以很显然，$(U(T)|\psi_E\rangle)$ 也是系统的一个本征值为 E 的能量本征态。由于这里的 T 是任意的，只要是对称操作群 $\mathscr{G}$ 中的一个群元即可，所以一般来说，总有某个 T 会使得 $\big(U(T)|\psi_E\rangle\big)$ 与 $|\psi_E\rangle$ 是两个不同的态。因此，一般来说，由于对称群 $\mathscr{G}$ 的存在，能级 E 将会是简并的，我们记所有简并态所张成的希尔伯特子空间为 $\mathscr{H}_E$。

很显然，上文的 $|\psi_E\rangle$ 属于 $\mathscr{H}_E$ 中的某个任意矢量，而且对于任意的 $T \in \mathscr{G}$, $\big(U(T)|\psi_E\rangle\big)$ 也是 $\mathscr{H}_E$ 中的矢量。因此也就是说，$\mathscr{H}_E$ 在所有可能的对称变换 $U(T)$ 的作用下封闭！这就意味着，我们可以将任意对称变换 $U(T)$ 限制在 $\mathscr{H}_E$ 之内，限制以后的 $U(T)$ 依然是 $\mathscr{H}_E$ 空间上的一个对称变换。数学家称这样的 $\mathscr{H}_E$ 空间为对称群 $\mathscr{G}$ 的表示空间。因为当我们将 $\mathscr{G}$ 在整个希尔伯特空间的表示 $U(\mathscr{G})$ 限制在这个子空间 $\mathscr{H}_E$ 以后，我们依然会得到 $\mathscr{G}$ 的一个表示 (因为限制以后的 $U(\mathscr{G})$ 在 $\mathscr{H}_E$ 上的作用封闭)，虽然这个表示要小一些[7]。所以对于一个有对称性的系统，**简并子空间将构成对称群的表示空间**。

上面我们发现，将整个希尔伯特空间限制到简并子空间 $\mathscr{H}_E$ 以后，我们会得到对称群 $\mathscr{G}$ 的一个小一些的表示。但是，表示空间 $\mathscr{H}_E$ 可能依然不是最小的，也就是说，我们能够把 $\mathscr{H}_E$ 正交分解成多个 (比方说 N 个) 更小的子空间，并且使得每一个这样的子空间都依然是对称群 $\mathscr{G}$ 的表示空间，也即是说，可以让 $\mathscr{H}_E$ 正交分解以后的这些子空间在任意 $U(T)$ 的作用下自身都保持封闭。我们记分解以后的第 i 个子空间为 $\mathscr{H}_{E,i}$, 并要求它同时是对称群 $\mathscr{G}$ 的表示空间 (在任意 $U(T)$ 作用下具有封闭性)，则我们可以把这样的正交分解表示成如下方程

$$\mathscr{H}_E = \mathscr{H}_{E,1} \oplus \mathscr{H}_{E,2} \oplus .... \oplus \mathscr{H}_{E,N}, \tag{6.82}$$

人们称 $\mathscr{H}_E$ 可以分解为 N 个子空间 $\mathscr{H}_{E,i}, i = 1,2,...,N$ 的直和。可以想象，我们可以将这种正交分解过程进行到底，使得最终分解出来的每一个子空间 $\mathscr{H}_{E,i}$ 本身是“最小”的，也就是说，$\mathscr{H}_{E,i}$ 不再能进一步正交分解成更小的子空间的直和，并同时让这些更小的子空间在任意 $U(T)$ 作用下都具有封闭性。这时候，最终的这些表示空间 $\mathscr{H}_{E,i}$ 就称为不可约表示空间，可以说不可约表示空间就是构成群表示空间的“原子”。而我们将 $U(\mathscr{G})$ 限制在这些不可约表示空间上所得到的群表示就称之为对称群 $\mathscr{G}$ 的不可约表示！总之，不可约

[7] 指表示空间维数小一些。

表示空间的要点在于：第一，它在群作用下具有封闭性，是表示空间。第二，它不能进一步分解成更小的表示空间。

比方说，我们已经知道，氢原子具有旋转对称性，因此它任意第 n 能级的简并空间 $\mathscr{H}_n$ 都必须是旋转对称群 (也就是 $SO(3)$ 群) 的表示空间，这里 n 表示氢原子的主量子数，相应的能量本征值为 $E_n = E_1/n^2$。但是，$\mathscr{H}_n$ 本身不是旋转对称性的不可约表示空间。旋转对称性的不可约表示空间的角动量大小得保持恒常，也就是说，旋转对称性的不可约表示空间得是 $\mathbf{L}^2$ 算符的简并空间。从前面对氢原子的学习中我们知道，$\mathbf{L}^2$ 的本征值为 $l(l+1)\hbar^2$，每一个不同的量子数 l(l 称作角动量量子数) 对应一个不同的本征值。而且我们还知道，$\mathbf{L}^2$ 的本征态空间是简并的，所有 l 量子数相同，但是 m 量子数 (m 称作磁量子数，它由 L_z 的本征值 $m\hbar$ 确定) 不同的态都具有相同的 $\mathbf{L}^2$ 本征值 $l(l+1)\hbar^2$，我们通常记这些简并态为 $|lm\rangle$ 态。$\mathbf{L}^2$ 的简并空间简并度为 $2l+1$，因为给定 l，m 有 $2l+1$ 个不同取值。所有这 $2l+1$ 个正交归一的 $\mathbf{L}^2$ 简并态 $|lm\rangle, m = 0, \pm 1, ..., \pm l$ 就张成了旋转对称性的一个不可约表示空间，因此，旋转对称性的不可约表示空间可以用量子数 l 来标记。从对氢原子的学习中我们还知道，对于每一个氢原子能级 n，l 可以有 n 个不同取值，$l = 0, 1, ..., n-1$, 它们的能量本征值都是简并的。因此我们知道，氢原子能级的简并子空间 $\mathscr{H}_n$ 可以分解成 n 个旋转群的不可约表示空间 $\mathscr{H}_{n,l}$ 的直和，即

$$\mathscr{H}_n = \mathscr{H}_{n,0} \oplus \mathscr{H}_{n,1} \oplus ... \oplus \mathscr{H}_{n,l} \oplus ... \oplus \mathscr{H}_{n,n-1}. \tag{6.83}$$

对于每一个其简并能级可以作形如 (6.82) 式这样分解的量子系统，我们可以设想给它加上一些微扰 V，如果这些微扰保持系统原来的对称性，也就是说，微扰 V 与任意对称变换 $U(T)$ 对易，则微扰以后的系统哈密顿量 $H+V$ 将依然与 $U(T)$ 对易。从前面第五章的学习中我们知道，微扰可以使得简并能级产生分裂，一般来说，这将导致原来相互简并的 N 个不可约表示空间 $\mathscr{H}_{E,i}, i = 1, 2, ..., N$ 的能量本征值产生分裂。但是，假如 $H+V$ 依然与所有的 $U(T)$ 对易，那微扰以后系统的能级一般来说依然会存在简并，其简并子空间依然能够分解成不可约表示空间的直和，从而这也就是说原来的不可约表示空间 $\mathscr{H}_{E,i}$ 在微扰以后将会依然简并 (因为它无法再分解)。可以设想，在一些相当一般性的对称微扰 (这些微扰都与任意对称变换 $U(T)$ 对易) 之下，所有可能分裂的能级都将分裂开来，最终系统的每一个能级简并空间将正好包含对称群 $\mathscr{G}$ 的仅仅一个不可约表示空间。只要微扰不破坏系统原来的对称性，那这些不可约表示空间的能级简并就将无法被分裂。

对于氢原子来说，我们可以设想给它加上一个相当一般性的球对称微扰 $\varepsilon V(r)$, 由于球对称，这样的微扰当然保持系统原来具有的旋转对称性。在这样的微扰作用下，氢原子的能级将变成 $E_{n,l}$, 也就是说，不同 l 量子数的态将不再简并。因此，这时候每一个系统能级的简并空间将对应唯一一个旋转对称性的不可约表示空间 $\mathscr{H}_{n,l}$。也就是说，在一般性的球对称微扰 $\varepsilon V(r)$ 之下，原来 (6.83) 式中氢原子的 n 个相互简并的不可约表示空间现在会完全分裂开来。

那么，是不是任意的球对称微扰都会使得氢原子相互简并的不可约表示空间分裂开来呢？答案是否定的，所有 $\varepsilon \frac{1}{r}$ 类型的微扰 (也就是所有库伦势型的微扰) 都不能使得 (6.83) 式的 n 个简并不可约表示空间分裂开来！这是因为，对于任意库伦势型的相互作用势能，

系统的对称性将不止是 $SO(3)$ 的旋转对称性，而是有一个更大的隐藏对称性，即所谓的 $SO(4)$ 对称性，因为对于库伦势型的相互作用势能，系统不止角动量会守恒，而且它的龙格-楞茨矢量 (Runge-Lenz vector)[8]也将会守恒，加上龙格-楞茨矢量所对应的算符，系统的对称性就会从原来的 $SO(3)$ 扩大为 $SO(4)$。而对于这个更大的 $SO(4)$ 对称性来说，$\mathscr{H}_{n,l}$ 空间太小了无法保证封闭性，只有整个 $\mathscr{H}_n$ 才是一个不可约表示空间。因此，只要不破坏 $SO(4)$ 对称性，我们就无法将 $\mathscr{H}_n$ 空间里的简并分裂开来。

当然，我们也可以进一步破坏氢原子的球对称性，比方说，我们给氢原子加上一个 z 方向的均匀磁场，这时候系统就没有整个三维空间的旋转对称性了，而仅仅只有绕 z 轴的旋转对称性，这种轴对称性我们称之为 $SO(2)$ 对称性。加上一个 z 方向的均匀磁场破坏原来的 $SO(3)$ 对称性以后，原来的 $SO(3)$ 不可约表示空间 $\mathscr{H}_{n,l}$ 中的 $2l+1$ 个简并态就会分裂开来，氢原子在均匀磁场中的这种能级分裂就是著名的塞曼效应。所以，对于氢原子来说，原来我们有一个很大的 $SO(4)$ 对称性，它给出了氢原子能级的 n^2 重简并，但是，加上一个非库伦势型的球对称微扰以后，$SO(4)$ 对称性就被破坏了，系统的对称性将变成 $SO(3)$ 空间旋转对称性，相应的能级简并度就是 $2l+1$，进一步加上一个 z 方向的均匀磁场将 $SO(3)$ 对称性破坏到 $SO(2)$ 轴对称性以后，$2l+1$ 重的能级简并也没有了，所有的 $|nlm\rangle$ 态的能量都将分裂开来，这就是塞曼效应。整个逐次破坏对称性的过程可以表示为如下数学方程

$$SO(4) \supset SO(3) \supset SO(2), \tag{6.84}$$

其中 $SO(4) \supset SO(3)$ 这样的表达式表示 $SO(4)$ 对称性包含 $SO(3)$ 对称性，所以这个方程从左到右就是一个不断破坏和降低对称性的过程。

因此，对称性总会导致能级简并，而一般来说只有破坏对称性才能降低这种能级简并。

6.4 补充内容：伽利略推动 (Boost)

在正文中我们集中处理的是保持动力学规律不变的对称性。如果这样的对称性是一个连续对称性，那这就要求它相应的量子变换的生成元和哈密顿量对易。但是我们也说了，对于包含时间平移对称性的时空对称群，我们实际上只需要求物理规律协变，也就是只需要求这些对称性的量子变换构成时空对称群的表示，严格来说实际上是只需构成投影表示。非相对论的伽利略对称群就是这样的一个时空对称群。这个对称群由时间平移、空间平移、空间旋转、以及伽利略推动 (Boost) 生成。其中，时间平移、空间平移、空间旋转我们其实都已经在正文中研究过了。这个里主要是要讨论伽利略推动对称性，当然，我们将仅仅只考虑孤立系统的伽利略推动对称性。

在非相对论力学中，我们知道，对于一个孤立系统，假设它由 n 个粒子组成，我们可以对它进行一个伽利略推动 (Boost)，推动之后，系统中任意一个粒子 a 的坐标将变换为 $\mathbf{x}_a \to \mathbf{x}_a + \mathbf{v}t$，因此这个粒子的动量将变换为 $\mathbf{p}_a \to \mathbf{p}'_a = \mathbf{p}_a + m_a\mathbf{v}$，式中 $\mathbf{v}$ 为伽利略推动的速度。我们知道，对于孤立系统，这样的伽利略推动是一个对称操作，它就是著名的伽利略协变性。

[8]假设系统的势能为 $-\frac{c}{r}$，则龙格-楞茨矢量的定义为 $\mathbf{M} = \frac{\mathbf{p}\times\mathbf{L}}{m} - \frac{c}{r}\mathbf{x}$。

伽利略推动的特殊之处就在于，它一定会影响系统的哈密顿量 (或者说系统的能量)。为了考察这种影响，我们将哈密顿量中的动能项写成 $\frac{\mathbf{p}^2}{2m}+\sum_a \frac{\widetilde{\mathbf{p}}_a^2}{2m_a}$，式中 $\mathbf{p}$ 为系统的总动量，m 为系统的总质量，而 $\widetilde{\mathbf{p}}_a=\mathbf{p}_a-m_a\mathbf{v}_c$($\mathbf{v}_c$ 为质心的速度) 为粒子 a 在质心系中的动量，因此动能算符分解出来的这两项分别代表质心的动能和系统在质心系中的动能。很显然，由于 $\mathbf{p}_a$ 的变换与 $m_a\mathbf{v}_c$ 的变换刚好抵消，因此在伽利略推动之下，$\widetilde{\mathbf{p}}_a$ 将保持不变，从而系统在质心系中的动能也将保持不变。同时，对于孤立系统，系统的势能只依赖于不同粒子之间的相对位置，因此显然也是伽利略推动不变的。因此，在伽利略推动之下，系统的哈密顿量中仅仅只有质心的动能是会变的，我们不妨将哈密顿量 H 重写为

$$H=\frac{\mathbf{p}^2}{2m}+H_c, \tag{6.85}$$

式中 H_c 表示系统在质心系中的能量，对于孤立系统，它在伽利略推动之下保持不变，在物理上，这一点非常好理解，因为孤立系统的质心系当然与伽利略推动无关。

在量子力学上，我们记任意一个伽利略推动为幺正变换 $U(\mathbf{v})=\exp\left(i\mathbf{v}\cdot\mathbf{K}/\hbar\right)$，式中 $\mathbf{K}$ 为这个伽利略推动的生成元，由于伽利略推动与时间坐标 t 有关，因此一般来说，$\mathbf{K}$ 可能依赖于 t。注意到在伽利略推动之下，系统总动量将变换为 $\mathbf{p}\to\mathbf{p}+m\mathbf{v}$，所以我们有

$$U^{-1}(\mathbf{v})\mathbf{P}U(\mathbf{v})=\mathbf{P}+m\mathbf{v}, \tag{6.86}$$

式中 $\mathbf{P}$ 表示系统的总动量算符。为了满足这个方程，我们必有

$$[K_i,P_j]=i\hbar m\delta_{ij}. \tag{6.87}$$

为了考察 $\mathbf{K}$ 与哈密顿量 H 的对易关系，我们首先注意到，由于孤立系统质心系的哈密顿量 H_c 与伽利略推动无关，因此很显然 $\mathbf{K}$ 与 H_c 对易，$[\mathbf{K},H_c]=0$。注意到哈密顿算符 $H=\frac{\mathbf{P}^2}{2m}+H_c$，所以由 (6.87) 式我们可以得到

$$[\mathbf{K},H]=i\hbar\mathbf{P}. \tag{6.88}$$

(6.87) 式和 (6.88) 式告诉我们，伽利略推动虽然是一个对称操作，但是它的生成元与哈密顿量 H 并不对易。虽然如此，但正如我们后面将会表明的，这并不会破坏对称性与守恒量之间的普遍性联系。

除了 (6.87) 式和 (6.88) 式这两个代数关系之外，注意到伽利略推动速度 $\mathbf{v}$ 的矢量叠加性，我们有 $U(\mathbf{v}_1)U(\mathbf{v}_2)=U(\mathbf{v}_1+\mathbf{v}_2)=U(\mathbf{v}_2)U(\mathbf{v}_1)$。也即是说，任意两个伽利略推动都可对易，从而我们有

$$[K_i,K_j]=0. \tag{6.89}$$

另外，注意到 $\mathbf{K}$ 是一个矢量算符，所以我们当然也有代数关系

$$[J_i,K_j]=i\hbar\varepsilon_{ijk}K_k, \tag{6.90}$$

这里我们已经默认了求和约定。

代数关系 (6.87)、(6.88) 以及 (6.89)、(6.90), 再加上我们在本章正文中推导的 $\mathbf{J}$、$\mathbf{P}$ 以及 H 这些算符之间的代数关系，就构成了所谓的伽利略对称性的代数关系。而伽利略对称性就是孤立的非相对论量子系统最一般的连续的时空对称性。

伽利略对称性可以看成是狭义相对论的庞加莱对称性在低速下的极限形式，因此它的这些生成元都可以推广到相对论情形，比方说伽利略推动在相对论中就相应于洛伦兹推动(它生成的就是著名的洛伦兹变换)。不过，相对论情形的代数关系需要作一些修正，比方说，在相对论情形，$[K_i,K_j]\neq 0$, 而大致是 $[K_i,K_j]\sim -i\hbar\varepsilon_{ijk}J_k/c^2$, 其中 c 为光速。实际上在相对论情形我们通常会选取不同的单位重新定义 $\mathbf{K}$，大体上，相对论中的洛伦兹推动生成元相当于非相对论中的 $c\mathbf{K}$。而相对论中的 K_i、K_j 之间的这个非零对易子所带来的物理效应就是所谓的托马斯进动。

回到我们的伽利略推动。什么样的生成元 $\mathbf{K}$ 能满足代数 (6.87)、(6.88) 以及 (6.89)、(6.90) 呢? 由于系统的总动量 $\mathbf{P}$ 也即是质心的动量，所以很显然，系统质心的位置算符 $\mathbf{X}_c$ 与质心动量算符 $\mathbf{P}$ 之间有代数关系，$[X_{c,i},P_j]=i\hbar\delta_{ij}$。由此我们很容易验证，如果取 $\mathbf{K}=m\mathbf{X}_c$, 那么代数关系 (6.87)、(6.88) 以及 (6.89)、(6.90) 都能得到满足, 其中证明 (6.88) 式能满足需要用到质心位置算符 $\mathbf{X}_c$ 与质心系的哈密顿量 H_c 对易 (这是因为在质心系中，质心位置其实确定为零)。

但是，能满足代数关系 (6.87)、(6.88) 以及 (6.89)、(6.90) 的 $\mathbf{K}$ 并不唯一，注意到孤立系统的总动量守恒 $[\mathbf{P},H]=0$，很容易验证 $m\mathbf{X}_c+f(t)\mathbf{P}$ 同样满足这组代数关系。这里 $f(t)$ 可以是关于 t 的一个任意函数。那么我们应该选哪个算符作为我们的伽利略推动生成元 $\mathbf{K}$ 呢?

为了确定 $\mathbf{K}$, 我们需要进一步考察伽利略推动在位置算符上的作用。由于在伽利略推动之下，质心位置将变换为 $\mathbf{x}_c\to\mathbf{x}_c+\mathbf{v}t$，所以我们还有算符方程

$$U^{-1}(\mathbf{v})\mathbf{X}_cU(\mathbf{v})=\mathbf{X}_c+\mathbf{v}t. \tag{6.91}$$

为了满足这个方程，必有 $[K_i,X_{c,j}]=i\hbar\delta_{ij}t$，因此我们只能取

$$\mathbf{K}=m\mathbf{X}_c-\mathbf{P}t. \tag{6.92}$$

这个算符的特殊之处在于，它显含时间坐标 t。

为了进一步验证对称性与守恒量之间的关系，我们可以变换到海森堡绘景，

$$\mathbf{K}_H=m\mathbf{X}_c^H-\mathbf{P}t. \tag{6.93}$$

式中 $\mathbf{X}_c^H$ 表示海森堡绘景中的质心位置算符[9]。利用定义式 $\mathbf{K}_H=e^{iHt/\hbar}\mathbf{K}e^{-iHt/\hbar}$，以及代数关系 $[m\mathbf{X}_c,H]=i\hbar\mathbf{P}$，以及孤立系统的 $[\mathbf{P},H]=0$, 我们也可以将海森堡绘景的 $\mathbf{K}_H$ 重写为，

$$\mathbf{K}_H=m\mathbf{X}_c. \tag{6.94}$$

注意,$\mathbf{X}_c$ 其实是一个在薛定谔绘景中定义的算符，因此它肯定不随时间演化，这里我们只是发现它刚好可以用作海森堡绘景中的 $\mathbf{K}_H$ 而已。可见，$\mathbf{K}_H$ 不随时间演化，因此它是一

[9]注意，由于 $[\mathbf{P},H]=[\mathbf{J},H]=0$，所以海森堡绘景中的总动量算符和总角动量算符均与薛定谔绘景中的相应算符相同。

个守恒量！由于伽利略推动是系统的一个对称性，那么其生成元 $\mathbf{K}_H$ 当然应该是一个守恒量。我们只不过是再一次验证了对称性与守恒量之间的普遍联系。虽然我们所考察的这个对称性有些特殊，它的生成元与哈密顿算符并不对易，但是，对称性与守恒量之间的普遍联系在这里依然成立。

我们已经得到了伽利略对称性的代数关系。但是，到现在为止我们还没有在量子力学的层次上严格验证它们的确能生成孤立系统的伽利略协变性。也就是说，我们还没有验证幺正算符 $U(\mathbf{v})$ 能和时间平移算符 $U(\tau)$, 空间平移算符 $U(\mathbf{a})$(其实更一般的还要包括空间旋转幺正算符 $U(R)$) 一起构成伽利略对称群的表示。为此，我们首先来考察一下在孤立系统情形下，伽利略协变性对伽利略推动算符与时间平移算符的对易关系有何要求。然后我们再验证前面得到的代数关系的确能满足这样的要求。

我们考察系统的任意一个动量本征态 $|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle$，然后将伽利略推动 $U(\mathbf{v})$ 和时间平移算符 $U(\tau)=\exp\left(iH\tau/\hbar\right)$ 先后作用到这个态上。首先，考察 $U(\tau)U(\mathbf{v})|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle$, 它等于 $U(\tau)|\mathbf{p}'_1,\mathbf{p}'_2,...,\mathbf{p}'_n\rangle=\exp\{i(\frac{\mathbf{p}'^2}{2m}+H_c)\tau/\hbar\}|\mathbf{p}'_1,\mathbf{p}'_2,...,\mathbf{p}'_n\rangle$, 式中 $\mathbf{p}'$ 为伽利略推动之后的系统总动量，显然 $\mathbf{p}'=\mathbf{p}+m\mathbf{v}$, 因此即有

$$U(\tau)U(\mathbf{v})|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle=\exp\{i(\frac{\mathbf{p}'^2}{2m}+H_c)\tau/\hbar\}|\mathbf{p}'_1,\mathbf{p}'_2,...,\mathbf{p}'_n\rangle. \tag{6.95}$$

再考察 $U(\mathbf{v})U(\tau)|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle$,对于孤立系统,由于 H_c 与 $\mathbf{K}$ 对易,因此即有 $U(\mathbf{v})U(\tau)|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle=U(\mathbf{v})\exp\{i(\frac{\mathbf{p}^2}{2m}+H_c)\tau/\hbar\}|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle=\exp\{i(\frac{\mathbf{p}^2}{2m}+H_c)\tau/\hbar\}|\mathbf{p}'_1,\mathbf{p}'_2,...,\mathbf{p}'_n\rangle$, 即

$$U(\mathbf{v})U(\tau)|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle=\exp\{i(\frac{\mathbf{p}^2}{2m}+H_c)\tau/\hbar\}|\mathbf{p}'_1,\mathbf{p}'_2,...,\mathbf{p}'_n\rangle. \tag{6.96}$$

比较 (6.95) 式和 (6.96) 式，我们可以得到

$$\begin{aligned}
&U(\tau)U(\mathbf{v})|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle\\
=&\exp\{i(\frac{\mathbf{p}'^2}{2m}-\frac{\mathbf{p}^2}{2m})\tau/\hbar\}U(\mathbf{v})U(\tau)|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle\\
=&U(\mathbf{v})U(\tau)\exp\left(i\mathbf{P}\cdot\mathbf{v}\tau/\hbar+i\frac{1}{2}m\mathbf{v}^2\tau/\hbar\right)|\mathbf{p}_1,\mathbf{p}_2,...,\mathbf{p}_n\rangle.
\end{aligned} \tag{6.97}$$

式中总动量算符 $\mathbf{P}$ 作用在动量本征态上的结果就是总动量 $\mathbf{p}$。由于 (6.97) 式对于任意动量本征态都成立，也就是在整个动量表象中都成立，因此如果我们的孤立系统有伽利略变换协变性，那就必须满足算符等式

$$\begin{aligned}
U(\tau)U(\mathbf{v})&=U(\mathbf{v})U(\tau)\exp\left(i\mathbf{P}\cdot\mathbf{v}\tau/\hbar+i\frac{1}{2}m\mathbf{v}^2\tau/\hbar\right)\\
&=\exp\left(i\frac{1}{2}m\mathbf{v}^2\tau/\hbar\right)U(\mathbf{v})U(\tau)U(-\mathbf{v}\tau).
\end{aligned} \tag{6.98}$$

可见，作为一种时空对称性，伽利略推动与时间平移算符不对易，(6.98) 式其实就是在这种不对易的情形下，对于孤立系统，伽利略协变性强加给我们的一个具体要求。等式 (6.98) 右边的 $U(-\mathbf{v}\tau)$ 表示空间平移操作，其定义为 $U(\mathbf{a})=\exp(-i\mathbf{a}\cdot\mathbf{P}/\hbar)$, 它的作用是将系统平移到 $\mathbf{x}\to\mathbf{x}+\mathbf{a}$ 处。很容易验证，等式 (6.98) 左右两边的作用都是将系统变换到 $t\to t+\tau$、$\mathbf{x}\to\mathbf{x}+\mathbf{v}t$ 时空位置，因此这个等式正是伽利略协变性的要求。当然，在这里我们发现等

式中还多出了一个相因子 $\exp\left(i\frac{1}{2}m\mathbf{v}^2\tau/\hbar\right)$，它正是我们在正文中提到的对称群的投影表示相因子。

另一方面，根据 Baker-Campbell-Hausdorff 公式，我们有

$$e^A e^B = e^B e^A e^{[A,B]+\frac{1}{2}[[A,B],A+B]+\cdots}. \tag{6.99}$$

由此我们容易得到算符恒等式，

$$\begin{aligned} &U(\tau)U(\mathbf{v}) = U(\mathbf{v})U(\tau)\times \\ &\exp\left(-[H,\mathbf{K}\cdot\mathbf{v}]\tau/\hbar^2 - i\frac{1}{2}\tau[[H,\mathbf{K}\cdot\mathbf{v}],H\tau+\mathbf{K}\cdot\mathbf{v}]/\hbar^3+\ldots.\right). \end{aligned} \tag{6.100}$$

在这个算符恒等式中代入代数关系式 (6.87) 和 (6.88)，我们容易验证它正好给出 (6.98) 式，这正好符合孤立系统伽利略协变性的要求。所以我们得到的代数关系 (6.87) 和 (6.88) 的确生成了孤立系统的伽利略协变性。

6.4.1 习题

1. 我们可以定义一个任意的非惯性坐标变换 $\mathbf{x}\to\mathbf{x}+\mathbf{d}(t)$，式中 $\mathbf{d}(t)$ 可以是 t 的任意函数。对于这样的坐标变换，其诱导的希尔伯特空间量子变换 $U[\mathbf{d}(t)]$ 可以写成

$$U[\mathbf{d}(t)] = e^{im\dot{\mathbf{d}}(t)\cdot\mathbf{X}_c/\hbar}e^{-i\mathbf{d}(t)\cdot\mathbf{P}/\hbar}e^{-i\frac{1}{2}m\int_0^t ds\dot{\mathbf{d}}^2(s)/\hbar}. \tag{6.101}$$

请验证：(1)$U[\mathbf{d}(t)]$ 在位置算符和动量算符上的作用能给出正确的变换关系。(2) 对于空间平移和伽利略推动的特殊情形，$U[\mathbf{d}(t)]$ 分别和 $U(\mathbf{a})$ 以及 $U(\mathbf{v})$ 一致。

6.5 空间反演对称性与时间反演对称性

前面我们讨论的都是连续对称性。这一节我们主要讨论两个重要的离散对称性，即空间反演对称性和时间反演对称性。这两个对称性都是时空对称性，我们可以将它们包含在时间平移对称性里一起来考虑，而我们所谓的系统具有空间反演对称性或时间反演对称性是指系统的动力学规律在空间反演或时间反演之下协变。比方说，我们可以将对系统的时间平移操作 T_τ 与空间反演操作 $\mathscr{P}$ 放在一起构成一个时空对称群，所谓一个系统具有空间反演对称性，即是要求对系统的空间反演操作所诱导的量子变换 $\Pi = U(\mathscr{P})$ 与时间平移算符 $U(\tau)$ 一起共同构成这个时空对称群的表示。类似的，我们也可以将时间平移操作 T_τ 与时间反演操作 $\mathscr{T}$ 放在一起构成一个时空对称群，而时间反演对称性就意味着，时间反演操作所诱导的量子变换 $\Theta = U(\mathscr{T})$ 与 $U(\tau)$ 一起共同构成这个时空对称群的表示。

6.5.1 空间反演对称性

空间反演变换

我们先来讨论空间反演对称性。但在讨论对称性之前，我们先来讨论对任何系统都可以定义的空间反演操作 $\mathscr{P}$，以及它所诱导的量子变换 $\Pi = U(\mathscr{P})$。当然，在空间反演操作 $\mathscr{P}$ 的作用下，空间坐标变换为 $\mathbf{x}\to\mathscr{P}\mathbf{x} = -\mathbf{x}$。另一方面，通过前面的学习我们知道，任

何量子系统都可以通过空间平移操作和空间旋转操作定义一个总动量算符 $\mathbf{P}$ 和总角动量算符 $\mathbf{J}$ (它们不一定要守恒)。我们首先要确立的是，这些算符和空间反演算符 Π 之间的代数关系。不过，在下面将要进行的讨论中，有一个微妙的问题值得预先指出，即对于连续操作的量子变换，我们能够确定它们是幺正变换，但是，空间反演是一个离散操作，因此我们并不能预先确立 Π 是幺正算符还是反幺正算符。

为了确立 Π 和 $\mathbf{P}$ 之间的代数关系，我们注意到，在空间平移操作 $T_{\mathbf{a}}$ 的作用下，空间坐标变换为 $\mathbf{x}\to\mathbf{x}+\mathbf{a}$，而在空间反演操作 $\mathscr{P}$ 的作用下，$\mathbf{x}\to-\mathbf{x}$。因此 $\mathscr{P}T_{\mathbf{a}}$ 的联合作用将会把 $\mathbf{x}$ 变换为 $\mathbf{x}\to-(\mathbf{x}+\mathbf{a})=-\mathbf{x}-\mathbf{a}$。类似的，我们发现 $T_{-\mathbf{a}}\mathscr{P}$ 的联合作用也会将 $\mathbf{x}$ 变换为 $\mathbf{x}\to=-\mathbf{x}-\mathbf{a}$，即 $\mathscr{P}T_{\mathbf{a}}=T_{-\mathbf{a}}\mathscr{P}$。由于量子变换是时空操作群的表示，所以这就意味着，$U(-\mathbf{a})\Pi$ 和 $\Pi U(\mathbf{a})$ 实际上是一样的，即

$$\Pi U(\mathbf{a})=U(-\mathbf{a})\Pi. \tag{6.102}$$

取 $\mathbf{a}=\varepsilon$ 为一无穷小平移，并将上面方程左右两边展开到 ε 的一阶项，就可以得到

$$\Pi(i\mathbf{P})=-(i\mathbf{P})\Pi. \tag{6.103}$$

类似的,我们可以考虑绕 $\mathbf{n}$ 轴的旋转操作和空间反演操作的联合作用。很显然,$\Pi U(\mathbf{n},\theta)$ 的效果是将空间坐标变换为 $\mathbf{x}\to-R(\mathbf{n},\theta)\mathbf{x}$, 这和 $U(\mathbf{n},\theta)\Pi$ 的作用效果一样。因此我们有

$$\Pi U(\mathbf{n},\theta)=U(\mathbf{n},\theta)\Pi. \tag{6.104}$$

取 $\theta=\varepsilon$ 为一无穷小旋转就可以得到

$$\Pi(i\mathbf{J})=(i\mathbf{J})\Pi. \tag{6.105}$$

但是，到现在为止我们还不能确定 Π 是幺正变换还是反幺正变换。为了确定这一点，我们可以考虑一个非相对论单粒子系统，这时候，粒子的坐标算符 $\mathbf{X}$ 是一个定义良好的算符。而且我们也知道，当我们考虑对系统的空间反演操作时，可以等价地将相应的量子变换作用在 $\mathbf{X}$ 算符上，形如 $\mathbf{X}\to\Pi^{-1}\mathbf{X}\Pi$。另一方面我们又知道，空间反演操作会将空间坐标变换为 $\mathbf{x}\to-\mathbf{x}$，因此坐标算符的变换关系必定是 $\mathbf{X}\to-\mathbf{X}$。从而我们即有

$$\Pi^{-1}\mathbf{X}\Pi=-\mathbf{X}. \tag{6.106}$$

另外我们也知道，动量正比于速度，也即正比于坐标对时间的导数，而空间反演当然不影响时间，所以进一步我们即有

$$\Pi^{-1}\mathbf{P}\Pi=-\mathbf{P}. \tag{6.107}$$

将这个式子和前面的 (6.103) 式比较，我们可以知道，Π 和虚数单位 i 一定可交换，即 $\Pi i=i\Pi$，从而 Π 必定是一个线性算符而不可能是反线性算符，进而也就必定是一个**幺正算符**。将这个结果代入 (6.105) 式，我们又可以得到

$$\Pi^{-1}\mathbf{J}\Pi=\mathbf{J}. \tag{6.108}$$

读者很容易看出来，最终我们得到的结果 (6.106)、(6.107)、(6.108) 与我们对空间反演的物理直观结论完全一致。这些结果中，(6.107)、(6.108) 对于任何量子系统都普遍成立，因为只要利用 Π 是一个线性算符 (而不是反线性算符)，那我们立马就能从前面的一般性结果 (6.103) 和 (6.105) 中得到它们。当然，(6.106) 式只适用于非相对论量子力学，在量子场论系统中是不适用的。

在非相对论量子力学中，根据 (6.106) 式和 (6.107) 式，我们很容易有

$$\Pi^{-1}\mathbf{L}\Pi = \mathbf{L}. \tag{6.109}$$

又由于 $\mathbf{J} = \mathbf{L} + \mathbf{S}$，进而根据 (6.108) 我们又可以进一步得到

$$\Pi^{-1}\mathbf{S}\Pi = \mathbf{S}. \tag{6.110}$$

虽然都是矢量算符，但由于在空间反演之下，这些角动量算符的变换关系与通常的矢量算符，比如动量算符、位置算符，相差一个负号，所以我们通常称角动量为赝矢量。除了角动量以外，物理学中常见的赝矢量还有磁场强度 $\mathbf{B}$。

对于非相对论量子系统，与前面讨论时间平移、空间平移、空间旋转时的推理一样，我们也容易知道，在坐标表象中，Π 对波函数 $\psi(\mathbf{x})$ 的作用必为

$$\Pi\psi(\mathbf{x}) = \psi(\mathscr{P}^{-1}\mathbf{x}) = \psi(-\mathbf{x}). \tag{6.111}$$

一般地，我们注意到 Π^2 与空间平移以及空间旋转等等所有这些时空操作都对易，而且空间反演两次效果上相当于不作任何操作，所以通常来说，Π^2 必定为 1，即

$$\Pi^2 = 1, \tag{6.112}$$

也即 $\Pi^{-1} = \Pi$。对于非相对论量子系统，我们其实很容易利用 (6.111) 式证明这个结果。由于 $\Pi^2 = 1$, 所以 Π 的本征值只能为 ± 1, 这个本征值称之为宇称，$+1$ 称为偶宇称，-1 称为奇宇称。而 Π 的本征态即是一个有确定宇称的量子态。

比方说，我们知道对于球谐函数而言将 $r^l Y_{lm}$ 在直角坐标中表达出来就是一个 l 阶齐次函数，在 $\mathbf{x} \to -\mathbf{x}$ 的空间反演之下，这样的齐次函数当然会出一个 $(-)^l$ 因子，因此 $\Pi(r^l Y_{lm}) = (-)^l r^l Y_{lm}$。但是，径向坐标 $r = |\mathbf{x}|$ 是空间反演不变的，因此这就意味着

$$\Pi Y_{lm}(\theta,\phi) = (-)^l Y_{lm}(\theta,\phi). \tag{6.113}$$

这是坐标表象的写法，如果用抽象的写法那就是

$$\Pi|lm\rangle = (-)^l|lm\rangle. \tag{6.114}$$

可见，角动量算符的本征态 $|lm\rangle$ 具有确定的宇称 $(-)^l$。

再比方说，在单自由度线性谐振子中，由于 $a \sim (X + iP/m\omega)$, $a^\dagger \sim (X + iP/m\omega)$，所以很显然 $\Pi^{-1}a\Pi = -a$, $\Pi^{-1}a^\dagger\Pi = -a^\dagger$。又于线性谐振子的基态波函数是一个高斯型函数，是偶函数，所以为偶宇称，即

$$\Pi|0\rangle = |0\rangle. \tag{6.115}$$

又由于 $|n\rangle \sim (a^\dagger)^n|0\rangle$, 所以 $\Pi|n\rangle \sim \Pi(a^\dagger)^n\Pi^{-1}|0\rangle = (-)^n(a^\dagger)^n|0\rangle$，即

$$\Pi|n\rangle = (-)^n|n\rangle. \tag{6.116}$$

所以，线性谐振子的能量本征态也都有确定的宇称。

空间反演对称性和宇称守恒

前面我们讨论的仅仅只是空间反演变换，对任何量子系统我们都可以定义这样的变换，那么什么情况下一个系统具有空间反演对称性呢?

为了弄清楚这个问题，我们可以将空间反演操作 $\mathscr{P}$ 和时间平移操作 T_τ 放在一起构成一个时空对称群。所谓一个系统具有空间反演对称性即是幺正变换 Π 和 $U(\tau)$ 一起共同构成这个时空对称群的表示。我们注意到 $\mathscr{P}T_\tau$ 的联合作用会产生时空坐标变换 $t \to t+\tau, \mathbf{x} \to -\mathbf{x}$，同样 $T_\tau\mathscr{P}$ 的联合作用也会产生同样的时空坐标变换, 即 $\mathscr{P}T_\tau = T_\tau\mathscr{P}$。因此，如果 Π 和 $U(\tau)$ 的确构成了这个时空对称群的表示，那 $U(\tau)\Pi$ 和 $\Pi U(\tau)$ 就应该相等，即

$$\Pi U(\tau) = U(\tau)\Pi. \tag{6.117}$$

取 $\tau = \varepsilon$ 为一个无穷小时间平移，并将上式左右两边展开到 ε 的一阶项，同时注意到 Π 是一个线性幺正算符，我们就可以得到

$$\Pi H = H\Pi. \tag{6.118}$$

也即是说，如果系统具有空间反演对称性，则空间反演变换 Π 必定和 H 对易。反过来其实也一样，如果幺正算符 Π 和 H 对易，则系统就具有空间反演对称性。

前面我们讨论连续对称性的时候曾经证明，系统的每一个连续对称性都会对应一条守恒定律，这也就是著名的诺特定理。但是，现在空间反演变换 Π 是一个离散对称性。它有没有相应的守恒定律呢? 其实也是有的，因为很显然，这时候海森堡绘景中的空间反演算符和薛定谔绘景的 Π 一样，也就是说，海森堡绘景中的空间反演算符不随时间演化，而这就意味着 Π 本身是一个守恒量！注意，这里的情形和连续对称性情形有一个本质性的不同，连续对称性情形的守恒量是对称变换的生成元，但是，这里的守恒量是对称变换本身。对于连续对称性的生成元来说，它的本征值有一种可加性，比如能量算符 H 是时间平移对称性的生成元，而对于一个多粒子态而言，不同粒子的能量当然是加在一起的，同样，量子系统的动量算符是空间平移的生成元，而不同粒子的动量当然也是加在一起的。但是，对于对称变换 Π 来说，它的本征值 (也就是宇称) 具有可乘性，而不是可加性。下面我们会具体地说明这个问题。

首先，算符 Π 是一个守恒量意味着，如果系统初始时处于 Π 的本征态，有一个确定的宇称，那么无论这个量子态如何演化，它将始终是 Π 的本征态，而且本征值 (也就是宇称) 保持不变！这就是所谓的宇称守恒。因此，空间反演对称性意味着宇称守恒。

在粒子物理中，如果系统具有空间反演对称性，那我们就可以要求每一个粒子都是空间反演算符的本征态，有一个确定的宇称，这就是粒子的内禀宇称。对于任意粒子 A, 它

的内禀宇称 η_A 就是在 A 静止的参考系中，粒子量子态 $|A\rangle$ 在 Π 作用下的本征值，即

$$\Pi|A\rangle = \eta_A|A\rangle, \tag{6.119}$$

当然，通常来说 η_A 只能等于 $+1$ 或者 -1。

假设我们有两个粒子 A 和 B，它们的内禀宇称分别为 η_A, η_B, 在这个两粒子系统的质心系中，两粒子的相对轨道角动量量子数为 l。则在质心系中，我们可以标记这个两粒子态为 $|A,B,l\rangle$。$|A,B,l\rangle$ 其实也是 Π 的本征态，其本征值我们先记为 η。下面我们将要说明如何计算这个 η, 首先，Π 在 $|A,B,l\rangle$ 态上作用时，它其实是同时对三个东西进行作用，第一，它作用在 A 粒子上，得到内禀宇称 η_A，第二，它同样作用在 B 上，得到内禀宇称 η_B，第三，但同时，Π 还作用在两粒子相对运动的空间波函数上，也就是相对运动的球谐函数 Y_{lm} 上，得到宇称 $(-)^l$。注意，由于 Π 是同时作用在这三个东西上，所以我们要把这三个宇称乘起来，得到整个量子态 $|A,B,l\rangle$ 的宇称 η, 即

$$\eta = \eta_A\eta_B(-)^l. \tag{6.120}$$

这个例子就清楚地说明了什么是宇称的可乘性。

理论分析发现电子 e、质子 p、中子 n 的内禀宇称都是 $+1$, 即 $\eta_e = \eta_p = \eta_n = +1$。当然，由于反粒子的量子数必定和粒子相反，所以它们的反粒子的内禀宇称就都是 -1。人们知道，强相互作用有空间反演对称性，因此强相互作用过程必定宇称守恒。据此人们就可以给各种强子 (包括质子、中子、π 介子等等，总之强子可以简单理解为多个夸克的束缚态) 标定内禀宇称，比如说人们发现三种 π 介子 (π^0、π^+、π^-) 的内禀宇称都是 -1，记为 $\eta_\pi = -1$。

宇称不守恒

在 1956 年之前，人们普遍相信，空间反演对称性是我们世界的一个基本对称性，就和时间平移对称性、空间平移对称性以及空间旋转对称性一样基本。因此，那时候人们认为宇称守恒是普遍成立的。但是，在 1940 年代，粒子物理学家发现了一件很困惑的事情，他们发现有两个自旋为 0 的标量粒子，当时人们称其中之一为 τ 粒子，称另一个为 θ 粒子，人们发现这两个粒子几乎在所有的方面都一样，有同样的质量同样的寿命等等，几乎就像同一个粒子一样。但是，奇怪的是，人们发现 τ 粒子可以衰变成 3 个 π 介子[10]，而 θ 粒子只能衰变成两个 π 介子。当时人们认为，这就意味着两者必定是两个不同的粒子，因为那时候人们相信宇称守恒普遍成立，因此如果一个标量粒子会衰变成 3 个 π 介子，那它的内禀宇称就必定为 $\eta_\pi^3 = -1$[11]，而如果一个标量粒子衰变成 2 个 π 介子，那它的内禀

[10] π 介子都是自旋为 0 的标量粒子。

[11] 这时候 3 个 π 介子的轨道运动部分对宇称没有贡献。因为在衰变的标量粒子静止的参考系中，由于初态总角动量为 0，角动量守恒意味着，3 个 π 介子的动量必定满足 $\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3=0$，衰变矩阵元只能依赖于 $\mathbf{p}_1\cdot\mathbf{p}_2$ 等等诸如此类 3 者动量的标量积，而且不可能是 $\mathbf{p}_1\cdot(\mathbf{p}_2\times\mathbf{p}_3)$ 这样的赝标量，由于 $\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3=0$，这种赝标量实际上必定等于 0。空间反演变换虽然会将所有的动量 $\mathbf{p}$ 变成 $-\mathbf{p}$，但是对于 $\mathbf{p}_1\cdot\mathbf{p}_2$ 这样的动量标量积，在空间反演之下它们实际上是不变的。因此，3 个 π 介子的轨道运动对宇称没有贡献。

宇称就必定为 $\eta_\pi^2=+1$ [12]。$+1$ 和 -1 当然是两个不同的量子数，因此，τ 粒子和 θ 粒子必定不同。但为什么这两个不同粒子的其它方面都完全一样呢？物理学家们可不太相信巧合。

1956 年李政道和杨振宁注意到，τ 粒子和 θ 粒子衰变成 π 介子的过程是属于所谓的弱相互作用过程。而且他们还注意到，的确，电磁相互作用和强相互作用过程都宇称守恒，但此前人们并没有用实验检验过弱相互作用过程中宇称是否守恒。所以，李政道和杨振宁大胆地提出：τ 粒子和 θ 粒子可能根本就是同一个粒子，只是弱相互作用过程没有空间反演对称性，在弱相互作用过程中宇称可能不守恒！因此前面关于这个粒子内禀宇称的那些分析其实根本不成立。

李政道和杨振宁提出用 $\mathrm{Co}^{60}\to\mathrm{Ni}^{60}+e^-+\overline{\nu}$ 这样一个弱相互作用导致的衰变过程来检验他们的理论假设。在这样一个衰变过程中，初态的 Co^{60} 静止但是有确定的自旋，不妨设初态 Co^{60} 的自旋沿 z 轴向上，记为 S_z，因此我们可以将初态标记为 $|\mathrm{Co}^{60},S_z\rangle$。实验主要关心末态放出来的电子，假设其自旋为 s_z，动量为 $\mathbf{p}$，则我们可以将末态标记为 $|e^-,\mathbf{p},s_z,....\rangle$，省略号表示末态的其它粒子。假设弱相互作用算符为 H'，则衰变概率将正比于

$$\left|\langle e^-,\mathbf{p},s_z,....|H'|\mathrm{Co}^{60},S_z\rangle\right|^2. \tag{6.121}$$

由于 $\Pi^2=1$，而且如果弱相互作用有空间反演对称性，则有 $\Pi H'=H'\Pi$。因此

$$\begin{aligned}&\langle e^-,\mathbf{p},s_z,....|H'|\mathrm{Co}^{60},S_z\rangle\\=&\langle e^-,\mathbf{p},s_z,....|H'\Pi^2|\mathrm{Co}^{60},S_z\rangle\\=&\langle e^-,\mathbf{p},s_z,....|\Pi H'\Pi|\mathrm{Co}^{60},S_z\rangle.\end{aligned}$$

由于空间反演不改变粒子自旋，但是会将 $\mathbf{p}$ 变成 $-\mathbf{p}$，即 $\langle\mathbf{p}|\Pi=\langle-\mathbf{p}|$，所以上面的推导告诉我们 $\langle e^-,\mathbf{p},s_z,....|H'|\mathrm{Co}^{60},S_z\rangle=\langle e^-,-\mathbf{p},s_z,....|H'|\mathrm{Co}^{60},S_z\rangle$，即有

$$\left|\langle e^-,\mathbf{p},s_z,....|H'|\mathrm{Co}^{60},S_z\rangle\right|^2=\left|\langle e^-,-\mathbf{p},s_z,....|H'|\mathrm{Co}^{60},S_z\rangle\right|^2. \tag{6.122}$$

因此，如果弱相互作用过程有空间反演对称性，则电子从 $+\mathbf{p}$ 和 $-\mathbf{p}$ 两个方向衰变出来的概率将相等。如果我们在这两个相反方向上测到的电子数不同，那就不可能有空间反演对称性。

同一年，吴健雄领导团队用高超的实验技巧证明，两个相反方向上衰变出来的电子数真的不同，这就证实了李政道和杨振宁提出来的，弱相互作用没有空间反演对称性的理论假设！

6.5.2 时间反演对称性

时间反演变换

现在我们来讨论时间反演变换 Θ。按照定义它是对系统的时间反演操作 $\mathscr{T}$ 所诱导的量子变换，其中 $\mathscr{T}:t\to\mathscr{T}t=-t$。当然，任何量子系统都可以定义 $\mathbf{P}$ 和 $\mathbf{J}$，我们首先关心

[12]和前一个注解一样，2 个 π 介子的轨道运动对于宇称也没有贡献。

的就是，Θ 与它们之间的代数关系。不过，我们也要指出，由于 Θ 是一个离散的变换，所以我们不能预先确定它是幺正的还是反幺正的 (当然实际上我们使用的符号已经暗示了答案)。

首先，我们把时间反演操作 $\mathscr{T}$ 和空间平移操作 $T_{\mathbf{a}}$ 放在一起构成一个对时空坐标的操作群，很容易验证 $\mathscr{T}T_{\mathbf{a}} = T_{\mathbf{a}}\mathscr{T}$。因此，如果 Θ 和 $U(\mathbf{a})$ 的确构成这些时空对称群的表示，那我们将有

$$\Theta U(\mathbf{a}) = U(\mathbf{a})\Theta. \tag{6.123}$$

取 $\mathbf{a} = \varepsilon$ 为无穷小空间平移，然后将上式左右两边展开到 ε 的一阶项，我们就有

$$\Theta(i\mathbf{P}) = (i\mathbf{P})\Theta. \tag{6.124}$$

类似的，我们可以把时间反演操作和空间旋转操作放在一起，从而最终会有

$$\Theta U(\mathbf{n}, \theta) = U(\mathbf{n}, \theta)\Theta. \tag{6.125}$$

取 $\theta = \varepsilon$ 为无穷小旋转，我们又可以得到

$$\Theta(i\mathbf{J}) = (i\mathbf{J})\Theta. \tag{6.126}$$

为了确定 Θ 是幺正变换还是反幺正变换，我们可以取一个非相对论单粒子系统，由于在时间反演操作下，粒子的时空坐标变换为 $t \to -t, \mathbf{x} \to \mathbf{x}$, 因此在时间反演之下，粒子的动量将变换为 $\mathbf{p} \to -\mathbf{p}$。因此，当我们等价地将 Θ 作用在厄密算符 $\mathbf{P}$ 上式，我们将有

$$\mathbf{P} \to \Theta^{-1}\mathbf{P}\Theta = -\mathbf{P}. \tag{6.127}$$

将这个结果与 (6.124) 式比较，我们就能发现，Θ 必定为一个**反幺正算符**，即必有 $\Theta i = -i\Theta$。将这个反线性关系代入 (6.126) 式，我们又可以得到

$$\Theta^{-1}\mathbf{J}\Theta = -\mathbf{J}. \tag{6.128}$$

这个式子告诉我们，在时间反演之下，粒子的角动量将会反转，这和我们的直观完全一致。

利用时间反演不改变粒子位置坐标，我们当然有 $\Theta^{-1}\mathbf{X}\Theta = \mathbf{X}$, 结合 (6.127) 式即有

$$\Theta^{-1}\mathbf{L}\Theta = -\mathbf{L}. \tag{6.129}$$

由于对于带自旋的粒子 $\mathbf{J} = \mathbf{L} + \mathbf{S}$，结合 (6.128) 式, 我们也将有

$$\Theta^{-1}\mathbf{S}\Theta = -\mathbf{S}. \tag{6.130}$$

这个式子告诉我们，在时间反演之下，粒子的自旋也会反转。

对于无自旋的非相对论粒子，其量子态可以在坐标表象中表示为

$$|\psi\rangle = \int d^3\mathbf{x}\psi(\mathbf{x})|\mathbf{x}\rangle. \tag{6.131}$$

式中 $\psi(\mathbf{x})=\langle\mathbf{x}|\psi\rangle$ 为波函数。将 Θ 作用在这个式子两边，并注意到它的反线性性，以及粒子位置坐标时间反演不变的 $\Theta|\mathbf{x}\rangle=|\mathbf{x}\rangle$ 式，我们就有

$$\Theta|\psi\rangle=\int d^3\mathbf{x}\psi^*(\mathbf{x})|\mathbf{x}\rangle. \tag{6.132}$$

这个式子告诉我们，对于无自旋粒子，在坐标表象中，时间反演算符 Θ 其实就是复数共轭算符 K (注意这里 K 不是伽利略推动生成元，而是表示对波函数取复数共轭)，即 $\Theta=K$。从这里我们很容易证明，对于无自旋的非相对论粒子

$$\Theta^2=1. \tag{6.133}$$

这个结果很好理解，因为时间反演两次效果上看起来应该相当于不作任何操作。

粗略一看，(6.133) 式好像允许 Θ 算符以任何相因子 ζ 为本征值。因为如果 $\Theta|\zeta\rangle=\zeta|\zeta\rangle$, 则 $\Theta^2|\zeta\rangle=\zeta^*\Theta|\zeta\rangle=|\zeta|^2|\zeta\rangle=|\zeta\rangle$, 完全符合 (6.133) 式。但实际上，这样的本征值 ζ 并没有任何物理意义，因为我们可以重新定义 $|\zeta'\rangle=\zeta^{1/2}|\zeta\rangle$, 对于 $|\zeta'\rangle$ 而言，$\Theta|\zeta'\rangle=(\zeta^*)^{1/2}\Theta|\zeta\rangle=(\zeta^*)^{1/2}\zeta|\zeta\rangle=\zeta^{1/2}|\zeta\rangle=|\zeta'\rangle$。总之，对于时间反演算符，我们无需考虑类似于“内禀宇称”那样的反演相因子。

但是，对于有自旋的情形，(6.133) 式实际上不成立。为了说清楚这个问题，我们首先取坐标表象，从而可以将粒子的波函数表示成旋量波函数，我们假设 Θ 在旋量波函数上的作用可以分解为

$$\Theta=YK. \tag{6.134}$$

式中 K 就是反幺正的复数共轭算符，从而 Y 必然是一个幺正算符。注意到对于电子，在这样的旋量波函数表象中，自旋算符都表示成了 2×2 的矩阵，具体来说，$\mathbf{S}=\frac{\hbar}{2}\vec{\sigma}$，这里 $\vec{\sigma}$ 是泡利矩阵。注意到 σ_x,σ_z 都是实矩阵，而 σ_y 是一个虚矩阵，它满足 $\sigma_y^*=-\sigma_y$。由此可知

$$K^{-1}S_xK=S_x,\quad K^{-1}S_yK=-S_y,\quad K^{-1}S_zK=S_z. \tag{6.135}$$

而另一方面我们也知道，在时间反演变换的作用下，$\Theta^{-1}\mathbf{S}\Theta=-\mathbf{S}$，即

$$\Theta^{-1}S_x\Theta=-S_x,\quad \Theta^{-1}S_y\Theta=-S_y,\quad \Theta^{-1}S_z\Theta=-S_z. \tag{6.136}$$

比较 (6.135) 式和 (6.136) 式可以知道，幺正算符 Y 必须满足，

$$Y^{-1}S_xY=-S_x,\quad Y^{-1}S_yY=S_y,\quad Y^{-1}S_zY=-S_z.. \tag{6.137}$$

另一方面，我们注意到自旋是一个矢量算符，因此，它们必然满足

$$\exp(i\theta\mathbf{n}\cdot\mathbf{S}/\hbar)S_i\exp(-i\theta\mathbf{n}\cdot\mathbf{S}/\hbar)=R(\mathbf{n},\theta)_{ij}S_j. \tag{6.138}$$

特别的，注意到绕 y 轴旋转 π 角会把一个 x 方向和 z 方向的矢量转到 $-x$ 和 $-z$ 方向，从而我们有

$$\begin{aligned}&\exp(i\pi S_y/\hbar)S_z\exp(-i\pi S_y/\hbar)=-S_z\\&\exp(i\pi S_y/\hbar)S_x\exp(-i\pi S_y/\hbar)=-S_x.\end{aligned} \tag{6.139}$$

所以很显然，为了满足 (6.137) 式，我们可以取 $Y=\exp(-i\pi S_y/\hbar)$。这也即是说，对于有自旋的粒子，

$$\Theta=\exp(-i\pi S_y/\hbar)K. \tag{6.140}$$

注意到 S_y 是纯虚的，所以 K 与 $\exp(-i\pi S_y/\hbar)$ 对易，从而对于有自旋的粒子易得 $\Theta^2=\exp(-i2\pi S_y/\hbar)$。由于 $J_y=L_y+S_y$，而对于轨道角动量而言，旋转 2π 角相当于不变，从而我们也可以把刚才的结果写成

$$\Theta^2=\exp(-i2\pi J_y/\hbar)=U(2\pi). \tag{6.141}$$

也即是说，时间反演两次并不是等于 1，而是等价于 2π 角旋转！在非相对论量子力学中 (6.141) 是普遍成立的，它不仅适用于单粒子情形，也适用于多粒子情形。在相对论量子力学中也有类似的式子，不过这时候 Θ 应该理解成 CT 变换，T 指的当然就是时间反演，而 C 则是电荷共轭，它的作用是把粒子变换成反粒子。而在量子场论中，同样有一个类似的 (6.141) 式，不过这时候 Θ 应该理解成 CPT 变换，即在相对论量子力学的 CT 变换基础上再把空间反演变换 P 联合进来。

回到非相对论量子力学。对于无自旋的系统，旋转 2π 角当然和不作任何操作是一样的，因此这时候 (6.141) 式就退回到之前我们得到的 $\Theta^2=1$。但是，对于单个电子这样的自旋 1/2 粒子的系统，由于旋转 2π 角会多出一个负号, 所以 (6.141) 式告诉我们，

$$\Theta^2=-1. \tag{6.142}$$

很明显，这个结论不仅对单个电子的系统成立，对任意奇数个电子的系统也都成立 (因为奇数个 -1 乘起来还是 -1)。

时间反演对称性

什么情况下，时间反演变换是量子系统的一个对称性呢？为了看清楚这个问题，我们将时间反演操作 $\mathscr{T}$ 和时间平移操作 T_τ 放在一起构成一个时空对称群，所谓一个系统有时间反演对称性，即是要求 Θ 和时间平移算符 $U(\tau)$ 一起构成这个时空对称群的表示。

我们注意到，在 $\mathscr{T}T_\tau$ 的联合作用下，时间坐标将变换为 $t\to-(t+\tau)=-t-\tau$。类似的，我们也很容易发现 $T_{-\tau}\mathscr{T}$ 的联合作用有同样的效果。因此

$$\mathscr{T}T_\tau=T_{-\tau}\mathscr{T}. \tag{6.143}$$

如果 Θ 和 $U(\tau)$ 一起构成这个时空对称群的表示，那就必定有

$$\Theta U(\tau)=U(-\tau)\Theta. \tag{6.144}$$

取 $\tau=\varepsilon$ 为无穷小时间平移，将上式两边展开到 ε 的一阶项，即有

$$\Theta(iH)=-(iH)\Theta. \tag{6.145}$$

注意到，Θ 为反线性算符，从而即有

$$\Theta H = H\Theta. \tag{6.146}$$

也即是说，如果一个系统有时间反演对称性，那 Θ 必定与哈密顿量对易，反过来也一样，如果时间反演变换 Θ 与哈密顿量对易，那系统就必定有时间反演对称性。

前面我们说了，Θ^2 可能等于 1，也可能等于 -1。如果一个量子系统的 $\Theta^2=-1$，而且它有时间反演对称性，那这个系统的每一个能级都将有偶数重简并。这是因为，假设 $|\psi_E\rangle$ 为系统的一个能量为 E 的哈密顿量本征态，满足 $H|\psi_E\rangle = E|\psi_E\rangle$。则由于 $\Theta H = H\Theta$，$|\psi'_E\rangle = \Theta|\psi_E\rangle$ 也必定是 H 的一个能量为 E 的本征态。可以证明 $|\psi'_E\rangle$ 必定和 $|\psi_E\rangle$ 正交，这是因为，由于 $\Theta^2=-1$，即 $\Theta^{-1}=-\Theta$，所以 $\langle\psi_E|\psi'_E\rangle = \langle\psi_E|\Theta\psi_E\rangle = \langle\Theta^\dagger\psi_E|\psi_E\rangle^* = \langle\psi_E|\Theta^\dagger\psi_E\rangle = \langle\psi_E|\Theta^{-1}\psi_E\rangle = -\langle\psi_E|\Theta\psi_E\rangle = -\langle\psi_E|\psi'_E\rangle$，即 $\langle\psi_E|\psi'_E\rangle = -\langle\psi_E|\psi'_E\rangle$，从而必有 $\langle\psi_E|\psi'_E\rangle = 0$。因此 $|\psi_E\rangle$ 和 $\Theta|\psi_E\rangle$ 就构成了二重简并。通常称这种偶数重简并为 **Kramer 简并**。

6.5.3 习题

1. 假设一个量子系统有空间反演对称性，再设 $t=0$ 时系统的初态 $|\psi(0)\rangle$ 有确定的宇称，请证明，系统任意时刻的量子态 $|\psi(t)\rangle$ 都有确定的宇称，而且这个宇称恒定不变。

2. 我们可以定义一个沿着与 z 轴垂直的 $x-y$ 平面作镜像反射的操作 Γ_z，它的定义是 $\Gamma_z : x\to x, y\to y, z\to -z$。记 Γ_z 诱导的量子变换为 $\mathrm{M}_z = U(\Gamma_z)$，请证明：(1) M_z 是一个幺正变换。(2) 请导出 M_z 与 $\mathbf{P}$ 以及 $\mathbf{J}$ 之间的代数关系。(3) 系统具有 M_z 反射对称性的充要条件是 $[\mathrm{M}_z, H]=0$。(4) 如果系统具有空间旋转对称性，请证明系统同时具有这种镜像反射对称性 M_z 的充要条件是它有空间反演对称性。

3. 假设系统具有时间反演对称性，请证明，时间反演算符 Θ 的本征值不是守恒量。

4. 请证明处于外磁场中的带电粒子系统必定没有时间反演对称性。



7. 角动量理论与旋转群表示

这一章我们主要讲述一般性的角动量表示理论。内容包括角动量算符的本征值和本征态，角动量的耦合规则，不可约张量算符等等。与此同时，我们也讲述了旋转群的不可约表示，以及它们的直乘分解。其中，我们关于角动量耦合规则的证明很可能是文献中没有出现过的。我们这个证明的好处是比较直观，也比较清晰。

除此之外我们还讨论了规范不变性与无质量粒子自旋之间的关系，解释了为什么光子只有左旋和右旋两种偏振。我们还将这种讨论延伸到了引力子等类似情形。

最后，我们讨论了量子力学中的 2π 角旋转，介绍了超选择定则的概念。进一步，我们也简单介绍了超对称的基本思想。

7.1 角动量的本征值和本征态

7.1.1 角动量的本征值与本征态

上一章我们得到了角动量的基本代数关系，它可以写成

$$[J_x,J_y]=i\hbar J_z,\quad [J_y,J_z]=i\hbar J_x,\quad [J_z,J_x]=i\hbar J_y. \tag{7.1}$$

下面我们定义角动量矢量的平方 $\mathbf{J}^2=J_x^2+J_y^2+J_z^2$。则我们有 $[\mathbf{J}^2,J_z]=[J_x^2,J_z]+[J_y^2,J_z]=J_x[J_x,J_z]+[J_x,J_z]J_x+J_y[J_y,J_z]+[J_y,J_z]J_y=-i\hbar(J_xJ_y+J_yJ_x)+i\hbar(J_yJ_x+J_xJ_y)=0$，即 $[\mathbf{J}^2,J_z]=0$。当然，z 分量在角动量的 3 个分量中并没有任何特殊的地方，所以这个结果也意味着 $\mathbf{J}^2$ 与角动量的任何一个分量均对易，

$$[\mathbf{J}^2,J_x]=[\mathbf{J}^2,J_y]=[\mathbf{J}^2,J_z]=0. \tag{7.2}$$

由于 $\mathbf{J}^2$ 和 J_z 对易，所以我们可以求这两个算符共同的本征态。但是，由于 J_z 和 J_x、J_y 均不对易，所以一般来说这个 J_z 的本征态无法同时是 J_x 或 J_y 的本征态。在物理上，这就是说，我们最多可以同时测量 $\mathbf{J}^2$ 和 J_z 的值，并以它们的值来标记一个系统的角动量量子态。

假设归一化的量子态 $|a,b\rangle$ 为 $\mathbf{J}^2$ 和 J_z 共同的本征态，本征值分别为 $a\hbar^2$ 和 $b\hbar$，即

$$\mathbf{J}^2|a,b\rangle=a\hbar^2|a,b\rangle,\quad J_z|a,b\rangle=b\hbar|a,b\rangle. \tag{7.3}$$

由于角动量 $\mathbf{J}$ 的量纲就是 $\hbar$ 的量纲，而 $\mathbf{J}^2$ 的量纲就是 $\hbar^2$ 的量纲，因此式中的 a,b 均无量纲。由于

$$(a-b^2)\hbar^2=\langle a,b|(\mathbf{J}^2-J_z^2)|a,b\rangle=\langle a,b|(J_x^2+J_y^2)|a,b\rangle\geq 0. \tag{7.4}$$

所以很显然

$$a\geq 0,\quad b^2\leq a. \tag{7.5}$$

即，给定一个 $\mathbf{J}^2$ 的本征值，J_z 的本征值有上下限。

下面，引入两个重要的算符 J_+ 和 J_-，其定义是

$$J_\pm=J_x\pm iJ_y. \tag{7.6}$$

很明显，J_+ 和 J_- 都不是厄密算符，而是互为厄密共轭关系。由于 (7.2) 式，很显然，$[\mathbf{J}^2,J_\pm]=0$。由此很容易验证，$J_\pm|a,b\rangle$ 仍然是 $\mathbf{J}^2$ 的本征态，本征值依然为 $a\hbar^2$。另外，利用基本代数关系 (7.1)，很容易验证

$$[J_z,J_\pm]=\pm\hbar J_\pm. \tag{7.7}$$

由此即有 $J_z(J_+|a,b\rangle)=(J_+J_z+[J_z,J_+])|a,b\rangle=(J_+J_z+\hbar J_+)|a,b\rangle=(b+1)\hbar(J_+|a,b\rangle)$，即 $J_+|a,b\rangle$ 依然为 J_z 的本征态，但是本征值变成了 $(b+1)\hbar$。由此可知 $J_+|a,b\rangle$ 必定正比于 $|a,b+1\rangle$，即 (式中 $c_{a,b}^+$ 为复系数)

$$J_+|a,b\rangle=c_{a,b}^+|a,b+1\rangle. \tag{7.8}$$

正因为如此，人们常常说 J_+ 是角动量的升算符，它在 J_z 本征态上的作用会将相应的量子数升 1。完全类似的，人们也很容易验证，J_- 是角动量的降算符，它在 J_z 本征态上的作用会将相应的量子数降 1，即有

$$J_-|a,b\rangle = c^-_{a,b}|a,b-1\rangle. \tag{7.9}$$

另一方面，根据 $J_\pm$ 的定义我们很容易验证

$$J_-J_+ = \mathbf{J}^2 - J_z(J_z+\hbar), \quad J_+J_- = \mathbf{J}^2 - J_z(J_z-\hbar). \tag{7.10}$$

我们有 (7.8) 式，并且还可以得到它的厄密共轭方程 $\langle a,b|J_- = (c^+_{a,b})^*\langle a,b+1|$, 将这个共轭方程和原方程 (7.8) 作内积即有 $\langle a,b|J_-J_+|a,b\rangle = |c^+_{a,b}|^2$。代入 (7.10) 式，可知 $|c^+_{a,b}|^2 = \langle a,b|\mathbf{J}^2 - J_z(J_z+\hbar)|a,b\rangle = [a-b(b+1)]\hbar^2$, 即

$$|c^+_{a,b}|^2 = [a-b(b+1)]\hbar^2. \tag{7.11}$$

完全类似的，由方程 (7.9) 与其共轭方程的内积，我们可以得到

$$|c^-_{a,b}|^2 = [a-b(b-1)]\hbar^2. \tag{7.12}$$

前面我们说了，给定 a, 则 b 的取值有上限，相应的 b 值我们记为 $b_{\max}$。因此，当我们不断用 J_+ 作用在 $|a,b\rangle$ 上，不断升高 b 时，我们最终一定会到达这个 $b_{\max}$，这时候相应的 $|a,b_{\max}\rangle$ 态就不可再升了，也即是说，

$$J_+|a,b_{\max}\rangle = 0. \tag{7.13}$$

将这个方程厄密共轭一下即有 $\langle a,b_{\max}|J_- = 0$, 从而 $\langle a,b_{\max}|J_-J_+|a,b_{\max}\rangle = 0$。代入 (7.10) 式即有

$$a - b_{\max}(b_{\max}+1) = 0. \tag{7.14}$$

同样我们也知道，b 有一个下限 $b_{\min}$, 当我们不断用 J_- 作用在 $|a,b\rangle$ 上，不断降低 b 时，我们也一定会到达 $b_{\min}$，这时候必有

$$J_-|a,b_{\min}\rangle = 0. \tag{7.15}$$

相应的共轭方程为 $\langle a,b_{\min}|J_+ = 0$，从而 $\langle a,b_{\min}|J_+J_-|a,b_{\min}\rangle = 0$。代入 (7.10) 式即有

$$a - b_{\min}(b_{\min}-1) = 0. \tag{7.16}$$

比较 (7.14) 式和 (7.16) 式，即有 $b_{\max}(b_{\max}+1) = b_{\min}(b_{\min}-1)$, 这个式子等价于 $(b_{\max}+b_{\min})(b_{\max}-b_{\min}+1) = 0$, 从而即有

$$b_{\min} = -b_{\max}. \tag{7.17}$$

另一方面，我们又可以从 $|a,b_{\min}\rangle$ 态出发，不断用升算符 J_+ 作用上去，将 b 从 $b_{\min}$ 开始逐步升高，每次作用升高 1，最终一直升高到 $b_{\max}$，由此可见，$b_{\max}$ 与 $b_{\min}$ 之间只能相差一个非负整数，设为 $n\in\mathbb{Z}^+$，即

$$b_{\max}-b_{\min}=n. \tag{7.18}$$

由此结合前面得到的 $b_{\min}=-b_{\max}$，我们很容易解出 $b_{\max}=-b_{\min}=n/2$。记 $n/2=j$ 为一个非负的**整数或半整数** (即 $j=0,1/2,1,3/2,2,...$)，则我们得到

$$b_{\max}=j,\quad b_{\min}=-j. \tag{7.19}$$

代入 (7.14) 式马上得到

$$a=j(j+1). \tag{7.20}$$

另外，对于一般性的 $|a,b\rangle$ 态，由于 b 的值只能与 $b_{\max}$ 以及 $b_{\min}$ 相差一个整数，所以所有可能的 b 都得是整数或半整数，记 $b=m$, 从角动量的升降过程可知，其可能的取值有 $2j+1$ 个，分别为

$$b=m=-j,-j+1,-j+2,...,j-1,j. \tag{7.21}$$

综上可知，我们完全可以将标记 $\mathbf{J}^2$ 与 J_z 共同本征态的两个量子数 a,b 替换成量子数 j,m, 从而可以将 $|a,b\rangle$ 态重记为 $|jm\rangle$，它满足

$$\mathbf{J}^2|jm\rangle=j(j+1)\hbar^2|jm\rangle,\quad J_z|jm\rangle=m\hbar|jm\rangle. \tag{7.22}$$

当然，作为厄密算符不同本征值的本征态，我们必有正交归一关系

$$\langle j'm'|jm\rangle=\delta_{j'j}\delta_{m'm}. \tag{7.23}$$

给定 j 但是 m 不同的 $2j+1$ 个正交态对于 $\mathbf{J}^2$ 来说都是简并的，这 $2j+1$ 个态通常称作角动量的 (j) **多重态** (或者称为角动量的 $2j+1$ 重态), 它们的所有可能线性叠加就张成了 $\mathbf{J}^2$ 的简并子空间 $\mathscr{H}_j$，$\mathscr{H}_j$ 内的所有态都具有同样的角动量大小 (即 $\mathbf{J}^2$ 的值均为 $j(j+1)\hbar^2$)。

在前面第 4 章中，对于轨道角动量 $\mathbf{L}$ 的本征值和本征态，我们得到过与 (7.22) 式和 (7.23) 式完全类似的式子，只不过在轨道角动量的情形，j 要替换成 l，并且只能取非负整数，不能取 $1/2,3/2,...$ 这样的半整数。所以，第 4 章关于轨道角动量本征值本征态的结论其实是 (7.22) 式的一个特例。(7.22) 式的另外一个特例是电子的自旋角动量，我们知道电子自旋角动量 S_z 只能取 $-\frac{1}{2}\hbar$ 和 $+\frac{1}{2}\hbar$, 相应的 $\mathbf{S}^2=S_x^2+S_y^2+S_z^2=3(\frac{1}{2}\hbar)^2=\frac{3}{4}\hbar^2=\frac{1}{2}(\frac{1}{2}+1)\hbar^2$, 所以很显然，自旋角动量相当于这里 $j=\frac{1}{2}$ 的情形，这就是为什么我们常常说电子的自旋是 $\frac{1}{2}$，指的就是它的量子数 j 的值，不过，为了强调它是自旋量子数，人们常常将这个量子数记为 s，对于电子自旋，$s=\frac{1}{2}$。奇妙的是，电子的自旋是半整数我们之前完全是从斯特恩-格拉赫实验的分析中得出的，但是现在，我们通过研究自旋角动量的一般性理论发现，理论上果然可以允许这样的半整数量子化现象出现!

电子自旋是角动量量子理论的一个特例这一事实使得人们想到，微观粒子的自旋除了可以是 $\frac{1}{2}$ 以外，也许还可能取其它值。的确，粒子物理的研究发现，有大量的粒子自旋量子数 $s=0$，称之为标量粒子，比方说 π 介子就是这样的粒子，再比方说所谓的“上帝粒子” Higgs 粒子也是这样的标量粒子。而且，也存在矢量粒子，比方说光子就是一个矢量粒子，它的自旋是 1，不过，根据我们刚才的分析，按照道理来说，$s=1$ 的话，那 m 的取值应该有 3 个可能性，$-1,0,+1$，但是光子只有两种线性独立的偏振态，即左旋光和右旋光，相应于 $m=\pm 1$，光子并没有 $m=0$ 的偏振态！从物理上说也即是，光波是横波，只有两个横向分量，没有纵向分量。

回到我们的角动量理论。将 $|a,b\rangle$ 态重记为 $|jm\rangle$ 以后，前面的叠加系数 $c^{+}_{a,b}$ 和 $c^{-}_{a,b}$ 当然也可以分别重记为 $c^{+}_{j,m}$ 和 $c^{-}_{j,m}$, 而方程 (7.11) 和 (7.12) 就变成

$$\begin{aligned} |c^{+}_{j,m}|^2 &= [j(j+1)-m(m+1)]\hbar^2, \\ |c^{-}_{j,m}|^2 &= [j(j+1)-m(m-1)]\hbar^2. \end{aligned} \tag{7.24}$$

另外，由方程 (7.8) 和方程 (7.9) 我们可以得到

$$J_-J_+|jm\rangle = c^{+}_{j,m}c^{-}_{j,m+1}|jm\rangle. \tag{7.25}$$

代入 $J_-J_+=\mathbf{J}^2-J_z(J_z+\hbar)$ 即可得到

$$[j(j+1)-m(m+1)]\hbar^2 = c^{+}_{j,m}c^{-}_{j,m+1} \tag{7.26}$$

现在，我们可以通过调节 j 相同 m 不同的各 $|jm\rangle$ 态之间的相对相位，使得叠加系数 $c^{+}_{j,m}$ 为非负实数。具体来说，我们首先调节 $|j,-j+1\rangle$ 的相位，使得 $c^{+}_{j,-j}>0$, 接着再进一步调节 $|j,-j+2\rangle$ 的相位，使得 $c^{+}_{j,-j+1}>0$, 如此逐步进行下去，直到 $c^{+}_{j,j-1}>0$，而由 (7.24) 式 $c^{+}_{j,j}$ 显然等于 0，从而就完成了对 $c^{+}_{j,m}$ 的相位调节。进一步，根据 (7.26) 式可知，如果 $c^{+}_{j,m}>0$，则必有 $c^{-}_{j,m+1}>0$，而由 (7.24) 式 $c^{-}_{j,-j}$ 显然等于 0，这样所有的 $c^{-}_{j,m}$ 也成了非负实数。从而根据 (7.24) 式，我们立即有

$$\begin{aligned} c^{+}_{j,m} &= \hbar\sqrt{[j(j+1)-m(m+1)]}, \\ c^{-}_{j,m} &= \hbar\sqrt{[j(j+1)-m(m-1)]}. \end{aligned} \tag{7.27}$$

也即

$$\begin{aligned} J_+|j,m\rangle &= \hbar\sqrt{[j(j+1)-m(m+1)]}|j,m+1\rangle, \\ J_-|j,m\rangle &= \hbar\sqrt{[j(j+1)-m(m-1)]}|j,m-1\rangle. \end{aligned} \tag{7.28}$$

根据 (7.22) 式和 (7.28) 式，很容易求出 $\mathbf{J}$ 任何分量的矩阵元 $\langle j',m'|\mathbf{J}|j,m\rangle$, 由正交归一关系 (7.23) 我们容易看出

$$\langle j',m'|\mathbf{J}|j,m\rangle \propto \delta_{j'j}. \tag{7.29}$$

由于量子数 j 决定了 $\mathbf{J}^2$ 的本征值，或者说决定了角动量的大小 (j 当然不能决定角动量的“方向”，这个“方向”由角动量在 z 轴上的投影 $m\hbar$ 决定)，因此这也就是说，角动量算符

在不同角动量大小的量子态之间的矩阵元必定为 0。因此，我们通常可以取定一个固定的 j 来考察角动量算符的表示矩阵，也即可以取定一个固定的 (j) 多重态来考察角动量算符的表示矩阵。这是因为，根据 (7.22) 式和 (7.28) 式，$\mathbf{J}$ 在 $|jm\rangle$ 态上的作用不会改变量子数 j。并且，由于简并子空间 $\mathscr{H}_j$ 中的任意态都可以写成 (j) 多重态的线性叠加，即对于任意 $|\psi\rangle \in \mathscr{H}_j$, 我们有 $|\psi\rangle = \sum_{m=-j}^{+j} \psi_m |jm\rangle$, 式中 ψ_m 为叠加系数。所以很显然，$\mathbf{J}$ 作用在任何这样的态上，结果将依然是 $\mathscr{H}_j$ 中的一个态。即任给 $|\psi\rangle \in \mathscr{H}_j$, 必有 $\mathbf{J}|\psi\rangle \in \mathscr{H}_j$，或者说简并子空间 $\mathscr{H}_j$ 对于角动量算符的作用来说是封闭的。

另一方面，从 (7.22) 式可以看出，J_z 的矩阵元是对角的 (当然是实数)，因为我们的表象就是 $\mathbf{J}^2, J_z$ 的共同本征态表象。此外，从 (7.28) 式可以看出，$J_\pm$ 的矩阵元都是实数。进而由于 $J_x = (J_+ + J_-)/2$，所以 J_x 的矩阵元也都是实数。而由于 $J_y = -i(J_+ - J_-)/2$, 所以 J_y 的矩阵元必定为纯虚数。

比方说，对于 $j = 1/2$ 的情形，它的 (1/2) 多重态由 $|\frac{1}{2}, -\frac{1}{2}\rangle, |\frac{1}{2}, \frac{1}{2}\rangle$ 两个态组成 (也叫角动量 2 重态)，容易算得角动量算符的相应表示矩阵为

$$
\begin{aligned}
J_x^{(\frac{1}{2})} &= \begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2}|J_x|\frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, \frac{1}{2}|J_x|\frac{1}{2}, -\frac{1}{2}\rangle \\ \langle \frac{1}{2}, -\frac{1}{2}|J_x|\frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, -\frac{1}{2}|J_x|\frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix} = \frac{\hbar}{2}\sigma_x \\
J_y^{(\frac{1}{2})} &= \begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2}|J_y|\frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, \frac{1}{2}|J_y|\frac{1}{2}, -\frac{1}{2}\rangle \\ \langle \frac{1}{2}, -\frac{1}{2}|J_y|\frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, -\frac{1}{2}|J_y|\frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix} = \frac{\hbar}{2}\sigma_y \\
J_z^{(\frac{1}{2})} &= \begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2}|J_z|\frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, \frac{1}{2}|J_z|\frac{1}{2}, -\frac{1}{2}\rangle \\ \langle \frac{1}{2}, -\frac{1}{2}|J_z|\frac{1}{2}, \frac{1}{2}\rangle & \langle \frac{1}{2}, -\frac{1}{2}|J_z|\frac{1}{2}, -\frac{1}{2}\rangle \end{pmatrix} = \frac{\hbar}{2}\sigma_z.
\end{aligned} \tag{7.30}
$$

式中 $\sigma_x, \sigma_y, \sigma_z$ 为 3 个泡利矩阵。这个结果清楚地显示了，电子自旋的确是角动量一般理论的一个特例。对于电子自旋来说，$|\frac{1}{2}, -\frac{1}{2}\rangle$ 态也就是 $|\downarrow\rangle$ 态，而 $|\frac{1}{2}, \frac{1}{2}\rangle$ 态也就是 $|\uparrow\rangle$ 态。

规范不变性与自旋

自旋为 1 的光子为什么只有 2 个偏振分量 (左旋光和右旋光) 而不是 3 个分量呢? 为什么没有 $m = 0$ 的态呢? 这不是和我们前面关于角动量的一般性结果矛盾吗? 回答是, 对于光子, 它的内禀角动量代数其实和我们上面分析用到的角动量代数有所不同。$[S_i, S_j] = i\hbar\varepsilon_{ijk}S_k$ 这样的内禀角动量代数是对于有质量的，能够静止下来的粒子而言的。这时候我们可以取粒子静止的参考系，在这个参考系里粒子的内禀角动量和通常的角动量一样，可以作三维旋转，所以代数关系为 $[S_i, S_j] = i\hbar\varepsilon_{ijk}S_k$。但是，光永远不能静止下来，不存在一个光子静止的参考系。因此对于任何观察者来说，光都有一个运动方向，光子内禀的运动只能绕着其运动方向的旋转，不能作任意三维旋转。由于光子的内禀角动量算符只能沿着光子的运动方向，所以不可能有自旋 0(因为自旋零一定是角动量指向了三维的其它方向，它在这个方向上的投影才为 0)，因此就只有 $m = \pm 1$ 的左旋光和右旋光。并且，由于电磁场的规范不变性，即使考虑到光与带电粒子的相互作用，这个结论也不会被改变。

类似的，引力子 (引力波的量子) 自旋是 2，但是由于引力子同样是光速运动，没有一个静止参考系，因此它的 m 只能取 ± 2，分别对应左旋引力子和右旋引力子。而且由于爱

因斯坦等效原理导致的广义协变性 (这也是一种规范不变性)，即使考虑到引力子的自相互作用以及引力与物质的相互作用，这个结论也不会被改变。

光子的自旋是 1，引力子的自旋是 2，它们都无质量，分别导致电磁场和爱因斯坦广义相对论所描述的引力场。人们自然想问，有没有自旋为 3 的无质量粒子呢？如果有的话岂不说明自然界中还存在一个更不可思议的力场。同样，有没有自旋为 4 甚至自旋更高的无质量粒子呢？它们会不会对应一些更奇妙的力场呢？为什么实验中没有观测到这些力场呢？回答是，这样的力场统统不存在，理论上可以证明自然界中不存在自旋为 3 或者更高自旋的无质量粒子[1]，证明的基本思想其实很简单，即对于这样的粒子，找不到相应的规范不变性可以让它们的质量在相互作用下保持为零。

7.1.2 * 旋转群的不可约表示

现在我们来考察旋转变换 $U(\mathbf{n},\theta)=\exp\left(-i\theta\mathbf{n}\cdot\mathbf{J}/\hbar\right)$ 在 $\mathbf{J}^2$ 的简并子空间 $\mathscr{H}_j$ 中的作用。由于这样的简并子空间在 $\mathbf{J}$ 的作用下封闭，所以我们很容易想到，$\mathscr{H}_j$ 在任意旋转变换 $U(\mathbf{n},\theta)$ 的作用下应该也封闭。证明其实也很简单，任取 $|\psi\rangle\in\mathscr{H}_j$，根据简并子空间的定义显然有 $\mathbf{J}^2|\psi\rangle=j(j+1)\hbar^2|\psi\rangle$，又由于 $[\mathbf{J}^2,\mathbf{J}]=0$，所以有 $\mathbf{J}^2U(\mathbf{n},\theta)=U(\mathbf{n},\theta)\mathbf{J}^2$，从而 $\mathbf{J}^2\big(U(\mathbf{n},\theta)|\psi\rangle\big)=U(\mathbf{n},\theta)\mathbf{J}^2|\psi\rangle=j(j+1)\hbar^2\big(U(\mathbf{n},\theta)|\psi\rangle\big)$，从而可知 $U(\mathbf{n},\theta)|\psi\rangle\in\mathscr{H}_j$。这就证明了 $\mathscr{H}_j$ 在任意旋转变换下都封闭。根据我们在上一章所学的知识，这也即是说，$\mathscr{H}_j$ 构成了空间旋转群的表示空间。

而且，这样的表示空间 $\mathscr{H}_j$ 必定是不可约表示空间。这是因为角动量升降算符的作用可以将 $\mathscr{H}_j$ 上的任何两个正交基矢 $|jm\rangle$ 和 $|jm'\rangle$ 联系起来。这就告诉我们，必定有一个适当的旋转变换可以将 $|jm\rangle$ 变换到 $|jm'\rangle$，反过来也一样。因此，任何 $\mathscr{H}_j$ 内部的更小的子空间都无法在所有旋转变换的作用下保持封闭。也即是说，表示空间 $\mathscr{H}_j$ 无法再分解成“更小”的表示空间了，从而 $\mathscr{H}_j$ 就是旋转群的不可约表示空间。

将旋转变换 $U(R)$ 限制在不可约表示空间 $\mathscr{H}_j$ 上，我们就可以得到旋转群的一个不可约表示，记为 (j) 不可约表示。取 (j) 多重态为 $\mathscr{H}_j$ 的矢量基，我们就可以将 (j) 不可约表示所对应的矩阵形式写出来，它就是如下 $(2j+1)\times(2j+1)$ 的不可约表示矩阵

$$D^{(j)}_{m_1m_2}(R)=\langle jm_1|U(R)|jm_2\rangle. \tag{7.31}$$

而且由于旋转变换本身是旋转群的表示，即 $U(R_1)U(R_2)=U(R_1R_2)$，我们有 $\langle jm_1|U(R_1)U(R_2)|jm_3\rangle=\langle jm_1|U(R_1R_2)|jm_3\rangle$，在 $U(R_1)$ 和 $U(R_2)$ 的中间插入 $\mathscr{H}_j$ 空间上的封闭性关系 $\sum_{m_2}|jm_2\rangle\langle jm_2|=1$，即有

$$\sum_{m_2}\langle jm_1|U(R_1)|jm_2\rangle\langle jm_2|U(R_2)|jm_3\rangle=\langle jm_1|U(R_1R_2)|jm_3\rangle, \tag{7.32}$$

用不可约表示矩阵来写即是

$$\sum_{m_2}D^{(j)}_{m_1m_2}(R_1)D^{(j)}_{m_2m_3}(R_2)=D^{(j)}_{m_1m_3}(R_1R_2). \tag{7.33}$$

[1]除非这些粒子根本不和任何东西 (包括它们自身) 相互作用，如果这样的话那它的存在对我们毫无意义，我们也无法观测它们。

很显然，旋转群的群乘法正好对应于不可约表示矩阵的矩阵乘法。实际上，由于旋转变换 $U(R)$ 是幺正变换，相应的不可约表示矩阵一定是幺正矩阵，也即有

$$D_{m'm}(R^{-1}) = D^*_{mm'}(R). \tag{7.34}$$

我们也有 $U(R)|jm\rangle = \sum_{m'} |jm'\rangle\langle jm'|U(R)|jm\rangle = \sum_{m'} |jm'\rangle D^{(j)}_{m'm}(R)$，即

$$U(R)|jm\rangle = \sum_{m'} |jm'\rangle D^{(j)}_{m'm}(R). \tag{7.35}$$

这正好反应了旋转群的不可约表示矩阵就是旋转变换在 $\mathscr{H}_j$ 空间的 (j) 多重态表象中的表示矩阵。

以上我们只谈到任意一个单独的不可约表示空间 $\mathscr{H}_j$，以及相应的不可约表示，而整个希尔伯特空间当然可以正交分解成许多不可约表示空间的直和，即

$$\mathscr{H} = ... \oplus \mathscr{H}_j \oplus ... \oplus \mathscr{H}_{j'} \oplus ... \tag{7.36}$$

而旋转变换 $U(R)$ 在整个希尔伯特空间上的作用就可以表示成一个很大的块对角化矩阵，其中每一个对角块相应于一个不可约表示矩阵。块对角化的原因是，每一个不可约表示空间在旋转变换的作用下都封闭，所以不同的不可约表示空间无法在旋转变换的作用下相互过渡，因此就造成了块对角化，等价地说这也就是 $\langle j'm'|U(R)|jm\rangle \propto \delta_{j'j}$。

7.1.3 习题

1. 请证明 $\langle jm|J_x|jm\rangle = \langle jm|J_y|jm\rangle = 0$，并计算 $\langle jm|J_x^2|jm\rangle$ 和 $\langle jm|J_y^2|jm\rangle$。进而验证角动量算符满足海森堡不确定关系。

2. (本题来源于 J.J.Sakurai, Jim Napolitano, Modern Quantum Mechanics, 习题) 对于角动量的 $j=1$ 不可约表示。

1) 请显示计算如下 3×3 矩阵，

$$\langle j=1,m'|J_y|j=1,m\rangle.$$

2) 请证明在 $j=1$ 情形，

$$\exp(-iJ_y\beta/\hbar) = 1 - i\left(\frac{J_y}{\hbar}\right)\sin\beta - \left(\frac{J_y}{\hbar}\right)^2(1-\cos\beta).$$

3) 请利用上面结果证明，不可约表示矩阵 $\langle j=1,m'|\exp(-iJ_y\beta/\hbar)|j=1,m\rangle$ 为

$$\begin{pmatrix} \frac{1}{2}(1+\cos\beta) & -\frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ \frac{1}{\sqrt{2}}\sin\beta & \cos\beta & -\frac{1}{\sqrt{2}}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & \frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{pmatrix}.$$

3. 试证如果 $\mathbf{n}$, $\mathbf{n}'$ 为三维空间的两个普通矢量 (不是算符), 则有

$$[\mathbf{J}\cdot\mathbf{n}, \mathbf{J}\cdot\mathbf{n}'] = i\hbar(\mathbf{n}\times\mathbf{n}')\cdot\mathbf{J}.$$

但如果 $\mathbf{n}$, $\mathbf{n}'$ 为两个矢量算符，请证明

$$[\mathbf{J}\cdot\mathbf{n}, \mathbf{J}\cdot\mathbf{n}'] = -i\hbar(\mathbf{n}\times\mathbf{n}')\cdot\mathbf{J}.$$

7.2 角动量合成

我们经常会碰到将两个电子的自旋耦合在一起，或者将一个电子的轨道角动量与自旋角动量耦合在一起的问题。这一节就是要研究如何一般性地处理这样的问题。为此，我们假设一个系统的物理自由度可以分成两部分，这样的两部分可以是一个双电子系统的两个电子，也可以是一个绕核运动的电子的自旋部分和轨道运动部分。为了表达清楚起见，我们分别称这两部分为“粒子 1”和“粒子 2”。假设“粒子 1”的希尔伯特空间为 $\mathscr{H}_1$，“粒子 2”的希尔伯特空间为 $\mathscr{H}_2$。再假设在“粒子 1”的 $\mathscr{H}_1$ 上取矢量基 $\{|n_1\rangle, n_1 = 1,2,3...\}$，在“粒子 2”的 $\mathscr{H}_2$ 上取矢量基 $\{|n_2\rangle, n_2 = 1,2,3...\}$，则整个系统的任意量子态 $|\psi\rangle$ 当然可以写成矢量基 $|n_1\rangle|n_2\rangle$ 的线性叠加，从而具有下面的形式

$$|\psi\rangle = \sum_{n_1,n_2} \psi_{n_1,n_2}|n_1\rangle|n_2\rangle. \tag{7.37}$$

人们通常把所有这种形式的量子态所构成的希尔伯特空间称作 $\mathscr{H}_1$ 与 $\mathscr{H}_2$ 的张量积，记为 $\mathscr{H}_1 \otimes \mathscr{H}_2$。因此，我们这里整个系统的希尔伯特空间 $\mathscr{H}$ 就是，

$$\mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2. \tag{7.38}$$

7.2.1 角动量合成规则

当然，我们要处理的是角动量问题。因此假设“粒子 1”的角动量为 $\mathbf{J}_1$，“粒子 2”的角动量为 $\mathbf{J}_2$，则系统的总角动量为

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2. \tag{7.39}$$

由于 $\mathbf{J}_1$ 和 $\mathbf{J}_2$ 分别作用在两个不同的希尔伯特空间 $\mathscr{H}_1$ 和 $\mathscr{H}_2$ 上，所以这两个角动量算符显然相互对易。

由于 $\mathbf{J}_1$、$\mathbf{J}_2$ 相互对易，人们很容易看出 $\mathbf{J}_1^2$、$\mathbf{J}_2^2$、J_{1z}、J_{2z} 这四个算符两两对易。所以对于描述整个系统的角动量量子态而言，我们可以取这四个算符共同的本征态，$|j_1 j_2; m_1 m_2\rangle$。当然，我们可以将这四个算符分成“粒子 1”的算符和“粒子 2”的算符两部分，对于“粒子 1”的算符 $\mathbf{J}_1^2$ 和 J_{1z}，其本征态为 $|j_1 m_1\rangle$。同样，对于“粒子 2”的算符 $\mathbf{J}_2^2$ 和 J_{2z}，其本征态为 $|j_2 m_2\rangle$。这样整个系统的角动量本征态又可以写成 $|j_1 m_1\rangle|j_2 m_2\rangle$，也即是说

$$|j_1 j_2; m_1 m_2\rangle = |j_1 m_1\rangle|j_2 m_2\rangle. \tag{7.40}$$

当然，给定 j_1, m_1 的取值范围是 $-j_1, -j_1+1, ..., j_1-1, j_1$，类似的，给定 j_2, m_2 的取值范围是 $-j_2, -j_2+1, ..., j_2-1, j_2$。

但是，我们显然也可以从整个系统总的希尔伯特空间 $\mathscr{H}$ 的角度来考察问题。这时候可以证明，我们同样可以将描述整个系统角动量量子态的四个算符选为 $\mathbf{J}_1^2$、$\mathbf{J}_2^2$、$\mathbf{J}^2$、J_z，这里 J_z 是总角动量 $\mathbf{J}$ 的 z 分量。要证明这种选择同样合适，只需证明这四个算符两两对易，从而有共同的本征态。很显然，我们真正需要证明的是，$\mathbf{J}_1^2$ 与 $\mathbf{J}^2$ 对易，与 J_z 也对易，同样 $\mathbf{J}_2^2$ 也与 $\mathbf{J}^2$ 对易，同时与 J_z 对易。

由于 $J_z = J_{1z} + J_{2z}$，其中角动量代数关系告诉我们 $[J_{1z}, \mathbf{J}_1^2] = 0$，而 J_{2z} 与 $\mathbf{J}_1^2$ 是相互独立的自由度，所以显然也有 $[J_{2z}, \mathbf{J}_1^2] = 0$。从而我们马上有 $[J_z, \mathbf{J}_1^2] = 0$。同理也有 $[J_z, \mathbf{J}_2^2] = 0$。

为了证明 $\mathbf{J}^2$ 与 $\mathbf{J}_1^2$ 对易，我们利用 $\mathbf{J}^2 = (\mathbf{J}_1 + \mathbf{J}_2)^2 = \mathbf{J}_1^2 + \mathbf{J}_2^2 + 2\mathbf{J}_1 \cdot \mathbf{J}_2$。很显然，要证明 $\mathbf{J}^2$ 与 $\mathbf{J}_1^2$ 对易，其实只需证明 $\mathbf{J}_1 \cdot \mathbf{J}_2$ 与 $\mathbf{J}_1^2$ 对易，因为其它部分的对易子为 0 是显然的。而由于角动量代数告诉我们 $[\mathbf{J}_1, \mathbf{J}_1^2] = 0$，而且 $\mathbf{J}_2$ 与 $\mathbf{J}_1^2$ 作用在两个不同的希尔伯特空间上，从而也有 $[\mathbf{J}_2, \mathbf{J}_1^2] = 0$，因此 $\mathbf{J}_1 \cdot \mathbf{J}_2$ 的确与 $\mathbf{J}_1^2$ 对易。这就证明了 $[\mathbf{J}^2, \mathbf{J}_1^2] = 0$，类似的也有 $[\mathbf{J}^2, \mathbf{J}_2^2] = 0$。

这样我们就证明了 $\mathbf{J}_1^2$、$\mathbf{J}_2^2$、$\mathbf{J}^2$、J_z 两两对易，从而有共同的本征态，记为 $|j_1 j_2; jm\rangle$，它满足

$$\begin{aligned}
\mathbf{J}_1^2 |j_1 j_2; jm\rangle &= j_1(j_1+1)\hbar^2 |j_1 j_2; jm\rangle, \\
\mathbf{J}_2^2 |j_1 j_2; jm\rangle &= j_2(j_2+1)\hbar^2 |j_1 j_2; jm\rangle \\
\mathbf{J}^2 |j_1 j_2; jm\rangle &= j(j+1)\hbar^2 |j_1 j_2; jm\rangle, \\
J_z |j_1 j_2; jm\rangle &= m\hbar |j_1 j_2; jm\rangle.
\end{aligned} \tag{7.41}$$

因此为了表示整个系统的角动量量子态，我们有两种选择，其一是使用 $|j_1 j_2; m_1 m_2\rangle$，其二是使用 $|j_1 j_2; jm\rangle$。很显然，这两种选择有共同的 j_1 和 j_2，但是第一种选择 $|j_1 j_2; m_1 m_2\rangle$ 是用两个 “粒子” 各自的 J_{1z} 和 J_{2z} 的本征态来描述系统，是角动量的无耦合表象。而第二种选择 $|j_1 j_2; jm\rangle$ 是用系统总的 $\mathbf{J}^2$ 和 J_z 的本征态来描述角动量量子态，是角动量的耦合表象。给定 j_1, j_2，我们有时候也将无耦合表象的 $|j_1 j_2; m_1 m_2\rangle$ 记为 $|m_1 m_2\rangle_{j_1 j_2}$，将耦合表象的 $|j_1 j_2; jm\rangle$ 记为 $|jm\rangle_{j_1 j_2}$。给定 j_1, j_2 和一个可能的 j，不同 m 的 $2j+1$ 个 $|jm\rangle_{j_1 j_2}$ 态就构成了总角动量的一个 (j) 多重态。

既然只是不同的表象，$|j_1 j_2; m_1 m_2\rangle$ 和 $|j_1 j_2; jm\rangle$ 当然是互为对方的线性组合的关系。比方说，我们有

$$|j_1 j_2; jm\rangle = \sum_{m_1, m_2} |j_1 j_2; m_1 m_2\rangle \langle j_1 j_2; m_1 m_2 | j_1 j_2; jm\rangle. \tag{7.42}$$

这里我们利用了无耦合表象的封闭性关系 $\sum_{m_1, m_2} |j_1 j_2; m_1 m_2\rangle\langle j_1 j_2; m_1 m_2| = 1$。式 (7.42) 中的表象变换矩阵 $\langle j_1 j_2; m_1 m_2 | j_1 j_2; jm\rangle$ 就是 Clebsch-Gordan 系数，它常常也被记为 $C_{j_1 j_2}(jm; m_1 m_2)$，或者利用**维格纳的 $3-j$ 符号**来标记，

$$\langle j_1 j_2; m_1 m_2 | j_1 j_2; j_3, -m_3\rangle = (-1)^{j_1 - j_2 - m_3} \sqrt{2j_3 + 1} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{7.43}$$

将 J_z 作用在 (7.42) 式左边，可以得到

$$J_z |jm\rangle_{j_1 j_2} = m\hbar |jm\rangle_{j_1 j_2} = \sum_{m_1, m_2} m\hbar |m_1 m_2\rangle_{j_1 j_2} C_{j_1 j_2}(jm; m_1 m_2). \tag{7.44}$$

但是由于 $J_z = J_{1z} + J_{2z}$，所以我们可以等价地用 $J_{1z} + J_{2z}$ 作用在等式 (7.42) 的右边，从而得到

$$J_z |jm\rangle_{j_1 j_2} = \sum_{m_1, m_2} (m_1 + m_2)\hbar |m_1 m_2\rangle_{j_1 j_2} C_{j_1 j_2}(jm; m_1 m_2). \tag{7.45}$$

将以上的两个结果相减就有

$$0=\sum_{m_1,m_2}(m_1+m_2-m)\hbar|m_1m_2\rangle_{j_1j_2}C_{j_1j_2}(jm;m_1m_2). \tag{7.46}$$

由于基矢 $|j_1j_2;m_1m_2\rangle$ 相互之间线性无关，因此这个式子要成立除非

$$(m_1+m_2-m)C_{j_1j_2}(jm;m_1m_2)=0. \tag{7.47}$$

这也即是说，仅当 $m_1+m_2=m$ 时，Clebsch-Gordan 系数才可能非 0。

下一个我们要解决的问题是，给定 j_1,j_2，耦合表象的总角动量量子数 j 可以取哪些可能值？答案很简单，利用经典物理矢量合成的图像我们可以知道，当 $\mathbf{J}_1$ 和 $\mathbf{J}_2$ 反平行时，它们合成出来的总角动量最小，这个最小的 $j=|j_1-j_2|$，相反，当 $\mathbf{J}_1$ 和 $\mathbf{J}_2$ 平行时，它们合成出来的总角动量最大，这个最大的 $j=j_1+j_2$。所以，在量子力学的层次上，j 的可能取值为 (对于其它的 j 值，相应的 Clebsch-Gordan 系数为 0)

$$|j_1-j_2|,|j_1-j_2|+1,....,j_1+j_2-1,j_1+j_2. \tag{7.48}$$

而且，当将 j_1 和 j_2 进行合成时，以上每一个可能的 j 都仅仅只能出现一次，不能重复出现。以上就是**角动量合成规则**。下面我们所要做的，就是证明这一角动量合成规则。

角动量合成规则的证明

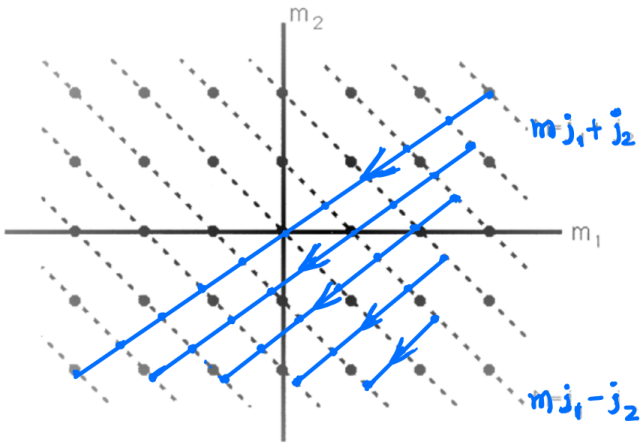

Figure 7.1: 角动量合成规则的证明。其中所有 (m_1,m_2) 点构成了一个沿水平方向有 $2j_1+1$ 个点，沿竖直方向有 $2j_2+1$ 个点的矩形。

图 (7.1) 是我们证明角动量合成规则的关键。图中每一个黑点代表无耦合表象中的一个正交归一态，黑点 (m_1,m_2) 就代表 $|m_1m_2\rangle_{j_1j_2}$ 态。图中每一根斜着的虚线有固定的 $m=m_1+m_2$，由于给定一个 m，有多个可能的 (m_1,m_2) 对，所以每一根虚线上通常有多个黑点，这样的虚线代表总 J_z 的简并子空间，我们称之为 m 虚线或 m 层虚线。m 虚线上的那些黑点构成了相应 J_z 简并子空间的一个正交归一矢量基。

总角动量 **J** 的每一个 (j) 多重态对于图 (7.1) 中的一根蓝线，我们称之为 j 蓝线，它总是从右上方的 $m=j$ 虚线开始，在总角动量降算符 J_- 的作用下沿着箭头逐层降到左下方的 $m=-j$ 虚线。而每一个 j 蓝线与 m 虚线的交点 (j,m) 都代表耦合表象中的一个正交归一态 $|jm\rangle_{j_1j_2}$。一根 j 蓝线必定要和 $2j+1$ 根虚线相交 (包括起点和终点在内)，这 $2j+1$ 个交点刚好对应耦合表象 (j) 多重态的 $2j+1$ 个态。一根固定虚线与所有蓝线的交点所代表的态和这根虚线上黑点代表的态当然不同，但它们都张成了 J_z 的简并子空间，因此互为线性组合关系。

值得强调的是，图 (7.1) 中的蓝线，包括它们与虚线的交点，都仅仅只是示意性的。所以图形看起来对不对称其实都没有关系。

对于每一层虚线，其上面的黑点数与这根虚线和所有蓝线的交点数必须相等。因为它们只是用两个不同的表象来表示同一个 m 值固定的 J_z 简并子空间。因此，两种不同表示线性独立的态数目当然要相等。比方说，在我们给出的示意性图中，黑点数最多的虚线上只有 5 个黑点，所以我们只能画 5 根蓝线。

画这些蓝线多重态时，我们是从最右上角 $m=j_1+j_2$ 虚线仅有的一个态开始，把它降到最左下角，从而就画出最长的那根蓝线，显然，它对应 $j=j_1+j_2$。接着，我们再从下一层虚线，即 $m=j_1+j_2-1$ 虚线 (因为这一层虚线上有两个黑点，因此可以再和一根蓝线相交) 开始画次长的蓝线，显然它对应 $j=j_1+j_2-1$，以此类推。很显然，对于每一个可能的 (j) 多重态，相应的蓝线都只能画出唯一一条。这就告诉我们，j_1 和 j_2 耦合出来的**每一个可能的 j 都只能出现一次**。在图 (7.1) 所示 $j_1>j_2$ 的情形中，这些蓝线中最后画的那根，也就是最短的那根蓝线一定对应 $j=j_1-j_2$。这是因为，从图形右上角往左下角走，虚线上的黑点数逐层增加 1，一直增加到 $m=j_1-j_2$ 虚线为止，而最短的蓝线就起于这根虚线。读者很容易明白，对于一般性的情况，最短的这根蓝线必定为 $|j_1-j_2|$。从图 (7.1) 中很容易看出来，相邻蓝线的 j 值刚好相差 1，因此从最短的蓝线到最长的蓝线，所有可能的 j 值必定为

$$|j_1-j_2|,|j_1-j_2|+1,....,j_1+j_2-1,j_1+j_2. \tag{7.49}$$

这其实就已经**完成了我们的证明**。

值得指出的是，由于每一根虚线和所有蓝线的交点数刚好等于这根虚线上的黑点数，所以整个图形蓝线与虚线的交点总数必定等于图形中的黑点总数。这也即是说，耦合表象的正交归一基矢总数刚好等于无耦合表象的正交归一基矢总数。这是因为，不同的表象之间只是基矢变换，当然不会改变基矢数目，也就是不会改变希尔伯特空间维数。

Clebsch-Gordan 系数

作为两个不同表象之间的变换矩阵，Clebsch-Gordan 系数 $C_{j_1j_2}(jm;m_1m_2)$ 当然是一个幺正矩阵。而且事实上，我们可以将所有 Clebsch-Gordan 系数取为**实数**，因此 $\langle j_1j_2;jm|j_1j_2;m_1m_2\rangle=\langle j_1j_2;m_1m_2|j_1j_2;jm\rangle$，即表象反变换的矩阵 $\langle j_1j_2;jm|j_1j_2;m_1m_2\rangle$ 和之前我们研究的正变换 $\langle j_1j_2;m_1m_2|j_1j_2;jm\rangle$ 一样，从而我们只需要同一组 Clebsch-Gordan 系数 $C_{j_1j_2}(jm;m_1m_2)$。

另外，实的幺正矩阵当然也就是正交矩阵，从而满足正交关系

$$\sum_{m_1,m_2} C_{j_1 j_2}(jm;m_1m_2)C_{j_1 j_2}(j'm';m_1m_2)=\delta_{jj'}\delta_{mm'}$$
$$\sum_{j,m} C_{j_1 j_2}(jm;m_1m_2)C_{j_1 j_2}(jm;m_1'm_2')=\delta_{m_1m_1'}\delta_{m_2m_2'}. \tag{7.50}$$

特别的，在上式的第一个式子中令 $j'=j,m'=m$，就有

$$\sum_{m_1,m_2}\left|C_{j_1 j_2}(jm;m_1m_2)\right|^2=1. \tag{7.51}$$

由于 $C_{j_1 j_2}(jm;m_1m_2)=\langle j_1 j_2;m_1m_2|j_1 j_2;jm\rangle$，所以上面的这个式子实际上是量子态 $|jm\rangle_{j_1 j_2}$ 在无耦合表象中的归一化关系。

为了计算 Clebsch-Gordan 系数，有一个递推关系非常有用。为了推导这个递推关系，我们将从表象变换的 (7.42) 式出发，我们将它重写如下

$$|jm\rangle_{j_1 j_2}=\sum_{m_1,m_2}|m_1m_2\rangle_{j_1 j_2}C_{j_1 j_2}(jm;m_1m_2). \tag{7.52}$$

我们用 $J_\pm=J_{1\pm}+J_{2\pm}$ 分别作用在上式左右两边，并利用角动量的升降关系式 (7.28)，就可以得到

$$\begin{aligned}&\sqrt{j(j+1)-m(m\pm1)}|j,m\pm1\rangle_{j_1 j_2}\\ =&\sum_{m_1,m_2}\sqrt{j_1(j_1+1)-m_1(m_1\pm1)}|m_1\pm1,m_2\rangle_{j_1 j_2}C_{j_1 j_2}(jm;m_1m_2)\\ +&\sum_{m_1,m_2}\sqrt{j_2(j_2+1)-m_2(m_2\pm1)}|m_1,m_2\pm1\rangle_{j_1 j_2}C_{j_1 j_2}(jm;m_1m_2).\end{aligned}$$

用 $|m_1m_2\rangle_{j_1 j_2}$ 去和这个式子做内积，就能得到

$$\begin{aligned}&\sqrt{j(j+1)-m(m\pm1)}C_{j_1 j_2}(j,m\pm1;m_1m_2)\\ =&\sqrt{j_1(j_1+1)-(m_1\mp1)m_1}C_{j_1 j_2}(jm;m_1\mp1,m_2)\\ +&\sqrt{j_2(j_2+1)-(m_2\mp1)m_2}C_{j_1 j_2}(jm;m_1,m_2\mp1).\end{aligned} \tag{7.53}$$

这就是 Clebsch-Gordan 系数的递推关系，可以证明，根据这个递推关系式和前面的归一化关系 (7.51)，原则上我们就可以计算出所有的 Clebsch-Gordan 系数。

关于如何一般性地利用上面的递推关系计算 Clebsch-Gordan 系数，我们推荐读者参考 J.J. Sakurai, Jim Napolitano, Modern Quantum Mechanics, Chapter 3, section 3.8. 我们这里只直接给出两种情形的 Clebsch-Gordan 系数。

第一种情形就是轨道角动量与自旋耦合的情形，这时候 $j_1=l$ 为整数 (不妨进一步假定 l 非零)，$j_2=s=\frac{1}{2}$。很显然，根据角动量合成规则，这时候总角动量 j 仅有两种可能性，$j=l\pm\frac{1}{2}$。相应的 Clebsch-Gordan 系数为

$$\begin{aligned}|j=l\pm\frac{1}{2},m\rangle_{l\frac{1}{2}}=&\pm\sqrt{\frac{l\pm m+\frac{1}{2}}{2l+1}}|m-\frac{1}{2},\frac{1}{2}\rangle_{l\frac{1}{2}}\\ &+\sqrt{\frac{l\mp m+\frac{1}{2}}{2l+1}}|m+\frac{1}{2},-\frac{1}{2}\rangle_{l\frac{1}{2}}.\end{aligned} \tag{7.54}$$

第二种情形就是两个电子自旋耦合情形。这种情形是我们在第 4 章中仔细研究过的。这时候，$j_1=j_2=\frac{1}{2}$。耦合的总角动量 j 只有两种取值，$j=0$ 和 $j=1$，分别对应自旋单态和自旋 3 重态。相应的 Clebsch-Gordan 系数我们在第 4 章中已经给出过了，结果为

$$|0,0\rangle=\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle), \tag{7.55}$$

以及

$$|1,+1\rangle=|\uparrow\uparrow\rangle,|1,0\rangle=\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle+|\downarrow\uparrow\rangle),|1,-1\rangle=|\downarrow\downarrow\rangle. \tag{7.56}$$

式中我们已经将无耦合表象的 $|\frac{1}{2}\frac{1}{2}\rangle_{\frac{1}{2}\frac{1}{2}}$ 态简记为 $|\uparrow\uparrow\rangle$，将 $|\frac{1}{2}-\frac{1}{2}\rangle_{\frac{1}{2}\frac{1}{2}}$ 记为 $|\uparrow\downarrow\rangle$, $|-\frac{1}{2}\frac{1}{2}\rangle_{\frac{1}{2}\frac{1}{2}}$ 记为 $|\downarrow\uparrow\rangle$, $|-\frac{1}{2}-\frac{1}{2}\rangle_{\frac{1}{2}\frac{1}{2}}$ 记为 $|\downarrow\downarrow\rangle$。

7.2.2 * 旋转群不可约表示的直乘分解

很显然，旋转变换 $U(\mathbf{n},\theta)$ 在不可约表示空间 $\mathscr{H}_{j_1}$ 上的作用为，$\exp\big(-i\theta\mathbf{n}\cdot\mathbf{J}_1/\hbar\big)\mathscr{H}_{j_1}$，同样，$U(\mathbf{n},\theta)$ 在不可约表示空间 $\mathscr{H}_{j_2}$ 上的作用为 $\exp\big(-i\theta\mathbf{n}\cdot\mathbf{J}_2/\hbar\big)\mathscr{H}_{j_2}$。因此，$U(\mathbf{n},\theta)$ 在张量积空间 $\mathscr{H}_{j_1}\otimes\mathscr{H}_{j_2}$ 上的作用即为 $\exp\big(-i\theta\mathbf{n}\cdot\mathbf{J}_1/\hbar\big)\exp\big(-i\theta\mathbf{n}\cdot\mathbf{J}_2/\hbar\big)\mathscr{H}_{j_1}\otimes\mathscr{H}_{j_2}=\exp\big(-i\theta\mathbf{n}\cdot\mathbf{J}/\hbar\big)\mathscr{H}_{j_1}\otimes\mathscr{H}_{j_2}$，式中 $\mathbf{J}=\mathbf{J}_1+\mathbf{J}_2$。

由此可见，两个不可约表示的张量积 $\mathscr{H}_{j_1}\otimes\mathscr{H}_{j_2}$ 依然是旋转群的表示空间。这时候旋转变换的生成元是 $\mathbf{J}_1$ 和 $\mathbf{J}_2$ 的合成，相应的表示空间通常是可约的。根据前面证明的角动量合成规则，我们可以将旋转群的这个张量积表示空间 $\mathscr{H}_{j_1}\otimes\mathscr{H}_{j_2}$ 按照合成后的生成元 $\mathbf{J}$ 的不可约表示空间进行正交分解，角动量合成规则告诉我们，分解的结果必定为

$$\mathscr{H}_{j_1}\otimes\mathscr{H}_{j_2}=\mathscr{H}_{|j_1-j_2|}\oplus\mathscr{H}_{|j_1-j_2|+1}\oplus\ldots\oplus\mathscr{H}_{j_1+j_2-1}\oplus\mathscr{H}_{j_1+j_2}. \tag{7.57}$$

将旋转变换限制在不可约表示空间 $\mathscr{H}_j$ 上，我们就有不可约表示 (j)，进而也常常把上面的张量积分解方程写成

$$(j_1)\otimes(j_2)=(|j_1-j_2|)\oplus(|j_1-j_2|+1)\oplus\ldots\oplus(j_1+j_2-1)\oplus(j_1+j_2). \tag{7.58}$$

通常称 $(j_1)\otimes(j_2)$ 为两个不可约表示的直乘，而上面的方程就称之为旋转群不可约表示的直乘分解。

现在，在 $\mathscr{H}_{j_1}\otimes\mathscr{H}_{j_2}$ 上选取基矢量 $|j_1m_1\rangle|j_2m_2\rangle$，并利用 Clebsch-Gordan 系数将之重新写成

$$|j_1m_1\rangle|j_2m_2\rangle=\sum_{j,m}|jm\rangle_{j_1j_2}C_{j_1j_2}(jm;m_1m_2). \tag{7.59}$$

将旋转变换 $U(R)$ 作用在这个式子左右两边，并利用 (7.35) 式，我们即有

$$\begin{aligned}&\sum_{m_1',m_2'}|j_1m_1'\rangle|j_2m_2'\rangle D^{(j_1)}_{m_1'm_1}(R)D^{(j_2)}_{m_2'm_2}(R)\\=&\sum_{j,m}\sum_{m'}|jm'\rangle_{j_1j_2}D^{(j)}_{m'm}(R)C_{j_1j_2}(jm;m_1m_2).\end{aligned} \tag{7.60}$$

将这个式子左右两边分别和下式左右两边做内积，

$$|j_1 m_1'\rangle|j_2 m_2'\rangle = \sum_{j,m'} |jm'\rangle_{j_1 j_2} C_{j_1 j_2}(jm';m_1' m_2'). \tag{7.61}$$

则我们就可以得到

$$D^{(j_1)}_{m_1' m_1}(R) D^{(j_2)}_{m_2' m_2}(R) = \sum_{j,m',m} D^{(j)}_{m'm}(R) C_{j_1 j_2}(jm';m_1' m_2') C_{j_1 j_2}(jm;m_1 m_2). \tag{7.62}$$

这个结果就是不可约表示矩阵的直乘分解关系。

7.2.3 习题

1. 请证明维格纳 $3-j$ 符号 $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ 仅在同时满足下面这些条件时才可能非零: $m_1+m_2+m_3=0$, $j_1+j_2+j_3$ 为整数，且 (j_1,j_2,j_3) 可以构成一个三角形的三个边长。

7.3 * 不可约张量算符与 Wigner-Eckart 定理

不可约张量算符

在上一章中，我们定义了矢量算符 $V_i (i=1,2,3$ 为三维直角坐标分量)，我们说它就是在旋转变换下按照下式变换的一组算符

$$U^{-1}(R) V_i U(R) = R_{ij} V_j. \tag{7.63}$$

这里 R_{ij} 是 3×3 的空间旋转矩阵的矩阵元。在上式中，将 R 替换成 R^{-1}，并利用 $U(R^{-1}) = U^{-1}(R)$，以及对于 3×3 的空间旋转矩阵，$R^{-1} = R^T$(因为它是正交矩阵)，我们就可以把上面这个式子改写成如下等价形式

$$U(R) V_i U^{-1}(R) = V_j R_{ji}. \tag{7.64}$$

注意，(7.63) 式和 (7.64) 式都使用了求和约定，(7.63) 式右边是 R_{ij} 的列指标和 V_j 的指标进行求和，而 (7.64) 式右边是 R_{ji} 的行指标和 V_j 的指标进行求和。之所以做这样的改写，是因为在上一章中，算符的变换关系和量子态的变换关系是相反的 (在上一章中我们说，这是因为我们可以认为量子态不变，是实验仪器在“反向旋转”)，而做了 (7.64) 式这样的等价改写之后，算符的变换关系就很类似于量子态的变换关系了，因为现在它们都是“正向旋转”了。后文我们将会看到，算符和量子态“同向变换”的这一处理会带来一些方便。

(7.64) 式只适用于 3 维空间矢量算符。其中的旋转矩阵 R_{ij} 实际上可以看成是旋转群 $j=1$ 的不可约表示矩阵，只不过我们不是在 $|j=1,m\rangle$ 这样的基中写这一表示矩阵而已。但这样看待 (7.64) 式的话，我们就很容易想到可以推广它。因为现在我们可以有更一般的不可约表示矩阵 $D^{(j)}_{m'm}(R)$，这使得我们想到，可以将 (7.64) 式推广成

$$U(R) T^{(k)}_q U^{-1}(R) = \sum_{q'} T^{(k)}_{q'} D^{(k)}_{q'q}(R). \tag{7.65}$$

我们称按照这样的式子变换的一组算符 $T_q^{(k)}$ 为不可约张量算符。比较上面的 (7.65) 式和之前的 (7.35) 式，我们发现，其实 $T_q^{(k)}$ 的变换关系和 $\mathbf{J}^2, J_z$ 共同本征态 $|kq\rangle$ 的变换关系完全一样，只不过由于 $T_q^{(k)}$ 是算符，所以旋转变换 $U(R)$ 在它上面的作用方式有所不同而已。

不仅如此，我们还可以进一步考察 $T_q^{(k)}|jm\rangle$ 的变换关系。利用 (7.35) 式和 (7.65) 式，我们有

$$\begin{aligned} U(R)\big(|T_q^{(k)}|jm\rangle\rangle\big) &= U(R)T_q^{(k)}U^{-1}(R)U(R)|jm\rangle \\ &= \sum_{q',m'} T_{q'}^{(k)}|jm'\rangle D_{q'q}^{(k)}(R)D_{m'm}^{(j)}(R). \end{aligned} \tag{7.66}$$

可见 $T_q^{(k)}|jm\rangle$ 的变换关系和角动量无耦合表象态 $|kq\rangle|jm\rangle = |kj;qm\rangle$ 的变换关系一样。因此它当然也可以按照耦合表象来进行重新组合，并且会满足同样的角动量合成规则，只不过现在我们合成的是旋转群的 (k) 不可约表示和 (j) 不可约表示 (分别相当于前面角动量合成规则中的 j_1 和 j_2)。由此我们立即可以得到下面的**选择定则** (它的严格证明可以从后面将要讨论的 Wigner-Eckart 定理中推论得到)，即

$$\langle \alpha' j'm'|T_q^{(k)}|\alpha jm\rangle = 0, \quad \text{unless} \quad |k-j| \le j' \le k+j, \quad m' = m+q. \tag{7.67}$$

式中 α, α' 是系统除角动量算符之外其它算符的量子数。

为了具体构造不可约张量算符，我们注意到第 4 章学过的与球谐函数密切相关的 l 阶齐次函数 $G_{lm}(\mathbf{x}) = r^l Y_{lm}$ 是角动量算符 $\mathbf{J}^2$ 和 J_z 共同的本征函数，因此在空间旋转之下，$G_{lm}(\mathbf{x})$ 和 (l) 不可约表示的基矢 $|lm\rangle$ 一样变换。现在我们只要把普通矢量 $\mathbf{x}$ 替换成一般的矢量算符 $\mathbf{V}$，就可以得到不可约张量算符 $T_m^{(l)}$，

$$T_m^{(l)} = G_{lm}(\mathbf{V}). \tag{7.68}$$

比方说，由 $G_{1,\pm1} \sim \mp\frac{x \pm iy}{\sqrt{2}}$, $G_{1,\pm1} \sim z$，我们容易将矢量算符重新组合成标准的 1 阶不可约张量算符

$$T_{\pm1}^{(1)} = \mp\frac{V_x \pm iV_y}{\sqrt{2}}, \quad T_0^{(1)} = V_z. \tag{7.69}$$

从而根据上面的一般性选择定则，我们就有选择定则 (假定 $j > 0$)

$$\begin{aligned} \langle \alpha' j'm'|V_x|\alpha jm\rangle &= \langle \alpha' j'm'|T_{-1}^{(1)} - T_1^{(1)}|\alpha jm\rangle/\sqrt{2} \\ &= 0 \quad \text{unless} \quad j' = j \pm 1, \quad m' = m \pm 1. \\ \langle \alpha' j'm'|V_z|\alpha jm\rangle &= \langle \alpha' j'm'|T_0^{(1)}|\alpha jm\rangle \\ &= 0 \quad \text{unless} \quad j' = j \pm 1, \quad m' = m. \end{aligned} \tag{7.70}$$

这个选择定则对于决定原子的偶极辐射来说非常重要。

当然，由于不可约张量算符 $T_m^{(j)}$ 与 $|jm\rangle$ 态的变换关系一样，所以我们也可以从任意两组不可约张量算符 $A_{m_1}^{(j_1)}$ 和 $B_{m_2}^{(j_2)}$ 出发，利用 Clebsch-Gordan 系数合成出一些新的不可约张量算符 $T_m^{(j)}$。

Wigner-Eckart 定理

利用 $D^{(k)}_{q'q}(R)=\langle kq'|U(R)|kq\rangle$, 并取绕 $\mathbf{n}$ 轴的 θ 角旋转，我们可以将不可约张量算符的定义式 (7.65) 重写为

$$U(\mathbf{n},\theta)T^{(k)}_q U^{-1}(\mathbf{n},\theta)=\sum_{q'}T^{(k)}_{q'}\langle kq'|U(\mathbf{n},\theta)|kq\rangle. \tag{7.71}$$

取 $\theta=\varepsilon$ 为无穷小旋转，代入上式并将等式两边展开到 ε 的一阶项，就可以得到

$$[\mathbf{J}\cdot\mathbf{n},T^{(k)}_q]=\sum_{q'}T^{(k)}_{q'}\langle kq'|\mathbf{J}\cdot\mathbf{n}|kq\rangle. \tag{7.72}$$

将上式中的 $\mathbf{J}\cdot\mathbf{n}$ 替换成 $J_\pm$，就可以得到

$$[J_\pm,T^{(k)}_q]=\hbar\sqrt{k(k+1)-q(q\pm1)}T^{(k)}_{q\pm1}. \tag{7.73}$$

将 (7.72) 式中的 $\mathbf{J}\cdot\mathbf{n}$ 替换成 J_z，就可以得到

$$[J_z,T^{(k)}_q]=q\hbar T^{(k)}_q. \tag{7.74}$$

下面我们来证明关于不可约张量算符矩阵元的 Wigner-Eckart 定理。我们首先写出定理的表述，然后再给出证明。

Wigner-Eckart 定理: 不可约张量算符在角动量本征态上的矩阵元满足

$$\langle\alpha',jm|T^{(j_1)}_{m_1}|\alpha,j_2m_2\rangle=C_{j_1j_2}(jm;m_1m_2)\cdot\langle\alpha',j\|T^{(j_1)}\|\alpha,j_2\rangle. \tag{7.75}$$

式中用双竖线表达的矩阵元不依赖于量子数 m_1,m_2,m。

为了证明 Wigner-Eckart 定理,我们实际上只需要证明所关心的矩阵元 $\langle\alpha',jm|T^{(j_1)}_{m_1}|\alpha,j_2m_2\rangle$ 满足和 Clebsch-Gordan 系数一样的递推关系式 (7.53), 这是因为对于任何给定的 j_1,j_2 和 j，这样的递推关系式可以将 Clebsch-Gordan 系数确定到仅仅相差一个整体的常数，这个常数就对应 (7.75) 中用双竖线表达的待定矩阵元。

为此，我们利用 $[J_\mp,T^{(j_1)}_{m_1}]=\hbar\sqrt{j_1(j_1+1)-m_1(m_1\mp1)}T^{(j_1)}_{m_1\mp1}$, 从而就可以得到

$$\begin{aligned}
&\langle\alpha',jm|[J_\mp,T^{(j_1)}_{m_1}]|\alpha,j_2m_2\rangle\\
=&\hbar\sqrt{j_1(j_1+1)-m_1(m_1\mp1)}\langle\alpha',jm|T^{(j_1)}_{m_1\mp1}|\alpha,j_2m_2\rangle\Longleftrightarrow\\
&\langle\alpha',jm|J_\mp T^{(j_1)}_{m_1}|\alpha,j_2m_2\rangle=\langle\alpha',jm|T^{(j_1)}_{m_1}J_\mp|\alpha,j_2m_2\rangle\\
+&\hbar\sqrt{j_1(j_1+1)-m_1(m_1\mp1)}\langle\alpha',jm|T^{(j_1)}_{m_1\mp1}|\alpha,j_2m_2\rangle.
\end{aligned} \tag{7.76}$$

代入角动量量子态的升降关系 $J_\mp|j_2m_2\rangle=\hbar\sqrt{j_2(j_2+1)-m_2(m_2\mp1)}|j_2,m_2\pm1\rangle$, 以及它的厄密共轭式 $\langle jm|J_\mp=\hbar\sqrt{j(j+1)-m(m\pm1)}\langle j,m\pm1|$, 我们就可以得到

$$\begin{aligned}
&\sqrt{j(j+1)-m(m\pm1)}\langle\alpha',j,m\pm1|T^{(j_1)}_{m_1}|\alpha,j_2m_2\rangle\\
=&\sqrt{j_1(j_1+1)-m_1(m_1\mp1)}\langle\alpha',jm|T^{(j_1)}_{m_1\mp1}|\alpha,j_2m_2\rangle\\
+&\sqrt{j_2(j_2+1)-m_2(m_2\mp1)}\langle\alpha',jm|T^{(j_1)}_{m_1}|\alpha,j_2,m_2\mp1\rangle.
\end{aligned} \tag{7.77}$$

很显然这个递推关系和 Clebsch-Gordan 系数的递推关系 (7.53) 其实是一样的。这就证明，给定 j_1, j_2 和 j，递推的结果 $\langle \alpha', jm|T^{(j_1)}_{m_1}|\alpha, j_2 m_2\rangle$ 只能与 Clebsch-Gordan 系数 $C_{j_1 j_2}(jm; m_1 m_2)$ 相差一个不依赖于 m_1, m_2, m 的整体常数。这就完成了对 Wigner-Eckart 定理的证明。

利用 Wigner-Eckart 定理，我们还可以证明如下**定理**：对于矢量算符 $\mathbf{V}$，我们有

$$\langle \alpha', jm'|V_q|\alpha, jm\rangle = \frac{\langle \alpha', jm|\mathbf{J}\cdot\mathbf{V}|\alpha, jm\rangle}{j(j+1)\hbar^2}\langle jm'|J_q|jm\rangle. \tag{7.78}$$

为了证明这一定理，我们首先注意到根据 Wigner-Eckart 定理，对于标量算符 S，必有 $\langle \alpha', j'm'|S|\alpha, jm\rangle = \delta_{j'j}\delta_{m'm}\langle \alpha', j\|S\|\alpha, j\rangle$。$\mathbf{J}\cdot\mathbf{V}$ 当然是一个标量算符，所以 $\langle \alpha', jm|\mathbf{J}\cdot\mathbf{V}|\alpha, jm\rangle$ 必定不依赖于 m。另一方面，$\mathbf{J}\cdot\mathbf{V} = J_z V_z + \frac{1}{2}(J_- V_+ + J_+ V_-)$(式中 $V_\pm = V_x \pm iV_y$)，因此

$$\begin{aligned}
&\langle \alpha', jm|\mathbf{J}\cdot\mathbf{V}|\alpha, jm\rangle = m\hbar\langle \alpha', jm|V_z|\alpha, jm\rangle \\
&+\frac{1}{2}\hbar\sqrt{j(j+1)-m(m+1)}\langle \alpha', j, m+1|V_+|\alpha, jm\rangle \\
&+\frac{1}{2}\hbar\sqrt{j(j+1)-m(m-1)}\langle \alpha', j, m-1|V_-|\alpha, jm\rangle.
\end{aligned} \tag{7.79}$$

利用关于 V_q 的 Wigner-Eckart 定理，很容易看出上式最终的结果必定具有 $c_{jm}\langle \alpha', j\|V\|\alpha, j\rangle$ 的形式，式中 c_{jm} 与 α, α' 以及与算符 $\mathbf{V}$ 无关。但是，我们刚刚又论证过 $\langle \alpha', jm|\mathbf{J}\cdot\mathbf{V}|\alpha, jm\rangle$ 不依赖于 m，所以 $c_{jm} = c_j$，即

$$\langle \alpha', jm|\mathbf{J}\cdot\mathbf{V}|\alpha, jm\rangle = c_j\langle \alpha', j\|V\|\alpha, j\rangle. \tag{7.80}$$

在上面的式子 (7.80) 中取 $\mathbf{V} = \mathbf{J}$，并取 $\alpha' = \alpha$，就可以得到

$$j(j+1)\hbar^2 = c_j\langle \alpha, j\|J\|\alpha, j\rangle. \tag{7.81}$$

代入 (7.80) 式消去 c_j，即有

$$\langle \alpha', jm|\mathbf{J}\cdot\mathbf{V}|\alpha, jm\rangle = j(j+1)\hbar^2\frac{\langle \alpha', j\|V\|\alpha, j\rangle}{\langle \alpha, j\|J\|\alpha, j\rangle}. \tag{7.82}$$

又根据 V_q 和 J_q 的 Wigner-Eckart 定理，可以得到

$$\frac{\langle \alpha', jm'|V_q|\alpha, jm\rangle}{\langle \alpha, jm'|J_q|\alpha, jm\rangle} = \frac{\langle \alpha', j\|V\|\alpha, j\rangle}{\langle \alpha, j\|J\|\alpha, j\rangle}. \tag{7.83}$$

将这个结果代入 (7.82) 式，就可以得到

$$\langle \alpha', jm|\mathbf{J}\cdot\mathbf{V}|\alpha, jm\rangle = j(j+1)\hbar^2\frac{\langle \alpha', jm'|V_q|\alpha, jm\rangle}{\langle \alpha, jm'|J_q|\alpha, jm\rangle}. \tag{7.84}$$

将这个式子整理一下就是我们要证明的结果。

7.3.1 习题

1. 原子的磁矩 $\vec{\mu}$ 具有如下形式

$$\vec{\mu} = \frac{-e}{2m_e}\left(g_L\mathbf{L} + g_s\mathbf{S}\right). \tag{7.85}$$

式中 $\mathbf{L}$ 为原子的总轨道角动量，$\mathbf{S}$ 为原子的总自旋角动量。根据 Wigner-Eckart 定理的推论，

$$\langle \alpha, jm'|\vec{\mu}|\alpha, jm\rangle = \frac{-e}{2m_e} g_{\text{eff}} \langle jm'|\mathbf{J}|jm\rangle, \tag{7.86}$$

请导出其中 g_{eff} 的表达式，并证明对于 $g_L = 1, g_s = 2$ 的情形

$$g_{\text{eff}} = 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)}, \tag{7.87}$$

这就是所谓的兰德 (Lande)g 因子。

2. 请用 Wigner-Eckart 定理将下面三个矩阵元联系起来

$$\langle j'm'| \mp \frac{1}{\sqrt{2}}(V_x \pm iV_y)|jm\rangle, \quad \langle j'm'|V_z|jm\rangle.$$

3. 氢原子的电四极矩为 $D_{ij} = e(x_i x_j - r^2\delta_{ij})$。请写出相应的电四极矩算符，并将它们用 2 阶不可约张量算符表示出来。最后，请导出氢原子电四极跃迁的选择定则。

7.4 2π 角旋转与超对称

7.4.1 2π 角旋转与超选择定则

这一节我们考察绕任意轴 $\mathbf{n}$ 的 2π 角旋转。首先我们取 $\mathbf{n}$ 为 z 轴，考察 $\exp\left(-i2\pi J_z/\hbar\right)|jm\rangle$，显然结果为 $\exp\left(-i2\pi J_z/\hbar\right)|jm\rangle = (-1)^{2m}|jm\rangle$，但是，由于任意 m 与 j 之间只相差一个整数，所以这个结果也可以写成 $\exp\left(-i2\pi J_z/\hbar\right)|jm\rangle = (-1)^{2j}|jm\rangle$，它对整个 (j) 多重态都成立，因此这就意味着，$\exp\left(-i2\pi J_z/\hbar\right)$ 在整个不可约表示空间 $\mathscr{H}_j$ 任意态上的作用结果均为乘上 $(-1)^{2j}$ 因子。但是，z 方向并不是一个特殊的方向，任意方向都可以叫做 z 方向，这就意味着，对于绕任意轴 $\mathbf{n}$ 的 2π 角旋转 $U(\mathbf{n}, 2\pi) = \exp\left(-i2\pi\mathbf{n}\cdot\mathbf{J}/\hbar\right)$，我们均有

$$U(\mathbf{n}, 2\pi)|\psi_j\rangle = (-1)^{2j}|\psi_j\rangle. \tag{7.88}$$

式中 $|\psi_j\rangle$ 为 $\mathscr{H}_j$ 中的任意量子态。可见，在量子力学中，2π 角旋转虽然不等于不转，但是，2π 角旋转的结果不依赖于转轴，因此我们可以一般性地将 2π 角旋转记作 $U(2\pi)$。

(7.88) 式告诉我们，2π 角旋转的本征值是 ± 1，对于整数角动量的态，2π 角旋转是一个恒等操作，而对于半整数角动量的态，2π 角旋转要多出一个负号！因此根据 $U(2\pi)$ 的作用结果，我们可以把整个希尔伯特空间 $\mathscr{H}$ 正交分解成两个子空间 $\mathscr{H}_+$ 和 $\mathscr{H}_-$，在 $\mathscr{H}_+$ 上，$U(2\pi)$ 是恒等操作，在 $\mathscr{H}_-$ 上，$U(2\pi)$ 要多出一个负号。我们称这种正交分解为，整个希尔伯特空间 $\mathscr{H}$ 是 $\mathscr{H}_+$ 和 $\mathscr{H}_-$ 的直和，记为

$$\mathscr{H} = \mathscr{H}_+ \oplus \mathscr{H}_-. \tag{7.89}$$

那么，在物理上，2π 角旋转的结果是否可观测呢？显然，通常的物理直观告诉我们，2π 角旋转的结果必定不可观测。这就意味着，系统的任何物理可观测量 $\mathscr{O}$，在 2π 角旋转之下都必须保持不变，即有

$$U(2\pi)\mathscr{O}U^{-1}(2\pi)=\mathscr{O}\Leftrightarrow U(2\pi)\mathscr{O}=\mathscr{O}U(2\pi). \tag{7.90}$$

注意，我们并没有说物理理论中的任何量都在 2π 角旋转之下不变，我们只是说实验上可测量的量必得不变，而物理理论中往往有很多量实际上是不可测量的，比如电磁学里面的矢量势 $\mathbf{A}$ 就不可测量。

(7.90) 式就意味着，对于任何可观测量 $\mathscr{O}$，对于 $\mathscr{H}_+$ 中的任意量子态 $|\psi_+\rangle$ 和 $\mathscr{H}_-$ 中的任意量子态 $|\psi_-\rangle$，我们必有

$$\langle\psi_-|\mathscr{O}|\psi_+\rangle=\langle\psi_+|\mathscr{O}|\psi_-\rangle=0. \tag{7.91}$$

这是因为 $\langle\psi_-|\mathscr{O}|\psi_+\rangle=\langle\psi_-|\mathscr{O}U(2\pi)|\psi_+\rangle=\langle\psi_-|U(2\pi)\mathscr{O}|\psi_+\rangle=-\langle\psi_-|\mathscr{O}|\psi_+\rangle$，从而 $\langle\psi_-|\mathscr{O}|\psi_+\rangle=0$，同样，$\langle\psi_+|\mathscr{O}|\psi_-\rangle=0$。(7.91) 式意味着，$\mathscr{H}_+$ 空间里的态与 $\mathscr{H}_-$ 空间里的态之间的干涉无法观测！它们之间可能有干涉，但在物理上这个干涉是永远测量不到的。(7.91) 式也意味，没有任何物理过程可以导致 $\mathscr{H}_+$ 空间里的态与 $\mathscr{H}_-$ 空间里的态之间相互跃迁！因此，我们常常称这一结果为超选择定则。

超选择定则和对称性导致的选择定则有所不同，选择定则只是说，某个或者某些物理量在两个给定态之间的矩阵元为 0，换一个任意的物理量，相应的矩阵元通常就不是 0。但是，超选择定则说的是，所有物理量的某一类矩阵元都为 0。

物理学中，类似的超选择定则还有，一个量子系统任何物理量在不同电荷本征值的电荷本征态之间的矩阵元也必定为 0。这是因为，上一章我们讲过，电荷守恒来源于一个 $e^{i\alpha Q}$ 的相位变换，实际上，这个相位变换不仅仅只是一个普通的相位变换，它还是我们以前讲过的规范变换的一部分，是一种空间各点同步进行的整体规范变换。既然是规范变换，那物理量的规范不变性就告诉我们，任何可观测量 $\mathscr{O}$ 在 $e^{i\alpha Q}$ 变换下必定不变，即

$$\mathscr{O}e^{i\alpha Q}=e^{i\alpha Q}\mathscr{O}. \tag{7.92}$$

从而这就会导致刚才所说的电荷超选择定则。所以，电荷守恒不仅仅是一条普通的守恒定律，它还是电荷超选择定则的必然结果！

7.4.2 超对称

[2]

前面我们讲到，2π 角旋转可以把希尔伯特空间分成两个超选择子空间，$\mathscr{H}_+$ 和 $\mathscr{H}_-$。一个美妙的想法是，可不可以给系统强加一种“超对称性”，让这两个子空间里的量子态一一对应起来呢？

[2]本节关于超对称的讨论取材于 Edward Witten 1982 年的经典论文，CONSTRAINTS ON SUPERSYMMETRY BREAKING，Nuclear Physics B202 (1982) 253-316。

实际上，这就是超对称的基本想法。具体来说，超对称是设想给系统引入一类特殊的对称变换，它可以将 $\mathscr{H}_+$ 空间的态变换到 $\mathscr{H}_-$ 空间里，反过来也一样。通常记这种特殊的对称变换的生成元为 Q，称作超荷。注意，在这一小节中，Q 不代表电荷算符，而是超荷算符，是超对称变换的生成元。

也即是说，对于任何 $|\psi_+\rangle \in \mathscr{H}_+$, $Q|\psi_+\rangle$ 要么等于 0，要么就属于 $\mathscr{H}_-$。同样，对于任何 $|\psi_-\rangle \in \mathscr{H}_-$, $Q|\psi_-\rangle$ 要么等于 0，要么就属于 $\mathscr{H}_+$。这就意味着，超荷 Q 满足

$$U(2\pi)Q = -QU(2\pi). \tag{7.93}$$

当然，前面的超选择定则意味着，超荷 Q 本身必定不是一个物理可观测量。但是，这并不意味着超对称仅仅只是数学游戏，Q 本身虽然不可观测，但是它生成的超对称变换会极大地限制物理系统的可能相互作用。这其实和任何对称性的基本原理是一样的，在现代物理中，对称性的作用就是限制相互作用，杨振宁甚至说得更极端更清楚一点，他说，**对称性决定相互作用**！

当然，作为对称变换的生成元，我们要求

$$[H, Q] = 0. \tag{7.94}$$

实际上，超对称会给系统的哈密顿量强加更高的要求，比方说，超对称要求, 在系统的质心系中 (或者在任何惯性系的非相对论极限下)

$$H = Q^2. \tag{7.95}$$

取哈密顿量 H 的本征值为 E 的所有可能本征态构成简并子空间 $\mathscr{H}_E$。由于 $HU(2\pi) = U(2\pi)H$，所以 $U(2\pi)$ 与 H 可以有共同的本征态，也即是说，我们可以进一步按照 $U(2\pi)$ 的本征值将 $\mathscr{H}_E$ 正交分解成

$$\mathscr{H}_E = \mathscr{H}_{E,+} \oplus \mathscr{H}_{E,-}. \tag{7.96}$$

而且，注意到超荷 Q 是一个厄密算符，我们能够证明超对称系统的能量本征值必定大于等于 0。为了证明这一点，我们任取 $|\psi_E\rangle \in \mathscr{H}_E$, 定义 $|\phi_E\rangle = Q|\psi_E\rangle$($\phi_E$ 可能为零)，其厄密共轭式为 $\langle\phi_E| = \langle\psi_E|Q$, 从而 $E = \langle\psi_E|H|\psi_E\rangle = \langle\psi_E|Q^2|\psi_E\rangle = \langle\phi_E|\phi_E\rangle \geq 0$，等号当且仅当 $|\phi_E\rangle = 0$ 时成立，这时 $Q|\psi_E\rangle = 0$。也即是说，H 的零本征值态必定也是 Q 的零化态，我们称这样的态为超对称不变的。

对于 $E > 0$ 的能量本征态，容易证明 $\mathscr{H}_{E,+}$ 和 $\mathscr{H}_{E,-}$ 的态之间必定严格地一一对应。具体来说，任取 $|\psi_{E,+}\rangle \in \mathscr{H}_{E,+}$, 则必定可以定义一个

$$|\psi_{E,-}\rangle = \frac{1}{\sqrt{E}} Q|\psi_{E,+}\rangle \in \mathscr{H}_{E,-}, \tag{7.97}$$

很显然 (利用能量本征方程)，反过来也有

$$|\psi_{E,+}\rangle = \frac{1}{\sqrt{E}} Q|\psi_{E,-}\rangle. \tag{7.98}$$

所以，$\mathscr{H}_{E,+}$ 和 $\mathscr{H}_{E,-}$ 的确严格一一对应，我们说这两个空间的态一一匹配，特别的，这两个空间的维数必定一样。

但是，对于 $E=0$ 的能量本征态，由于 $Q|\psi_{E=0,+}\rangle = Q|\psi_{E=0,-}\rangle = 0$, 所以上一段所建立的 $\mathscr{H}_{E,+}$ 和 $\mathscr{H}_{E,-}$ 之间的匹配关系对于 $E=0$ 情形并不成立。(图 (7.2) 中示意性地画出了超对称系统能量本征态的分布情况。) 这时候，$\mathscr{H}_{E=0,+}$ 和 $\mathscr{H}_{E=0,-}$ 的维数就有可能不同，

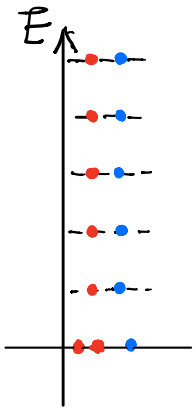

Figure 7.2: 一个超对称系统的能量本征态分布示意图，红点表示一个 $+$ 态，蓝点表示一个 $-$ 态，对于 $E>0$ 能级，两者一一匹配，但是对于 $E=0$ 能级，两者之间没有匹配关系。

我们记 n_+ 为 $\mathscr{H}_{E=0,+}$ 的维数，n_- 为 $\mathscr{H}_{E=0,-}$ 的维数，则下面的表达式就有可能不等于 0，

$$n_+ - n_-. \tag{7.99}$$

可以论证，$n_+ - n_-$ 是一个拓扑不变量。也就是说，我们可以连续地调节系统的某些参数，随着参数的调节，n_+ 和 n_- 都可能发生变化，但是它们的差值将保持不变！这是因为，比方说，随着参数的调节某些 $E>0$ 的能量本征态连续地变成了零能本征态，从而使得 n_+ 和 n_- 的数目增加了。但是，请注意，只要 $E>0$，那么 $|\psi_{E,+}\rangle$ 就和 $|\psi_{E,-}\rangle$ 一一匹配，所以如果有一个 $|\psi_{E,+}\rangle$ 态连续地变成了一个零能态，那必定同步地也有一个 $|\psi_{E,-}\rangle$ 态连续地变成零能态，n_+ 和 n_- 都增加 1，从而差值保持不变，如图 (7.3) 所示。总之，随

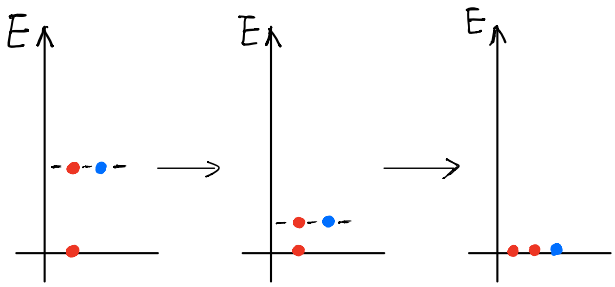

Figure 7.3: 随着系统参数的连续调节，一对 $E>0$ 的能量本征态连续地转变成了零能本征态，$n_+ - n_-$ 保持不变。

着参数的连续调节，n_+ 和 n_- 总是同步增加或者同步减少，使得最后的差值保持不变，从而 $n_+ - n_-$ 是超对称系统的一个拓扑不变量。

$n_+ - n_-$ 这样的拓扑不变量有什么用呢？当然，它在数学上很有用，通过计算它可以得出一些阿蒂亚-辛格指标定理。那么在物理上，$n_+ - n_-$ 有什么用呢？如果这个拓扑不变量等于 0，那的确不能告诉我们太多有用的信息，但是，如果 $n_+ - n_- \neq 0$，那它就意味着，不管系统怎么变化，不管系统的参数怎么增强，零能本征态必定存在！因为否则就必有 $n_+ = n_- = 0 \Rightarrow n_+ - n_- = 0$。零能本征态是超对称系统可能的最低能态，因此，对于这种情形，其基态 (在量子场论中也就是真空态) 必定为零能本征态。另一方面，前面我们说过，零能本征态必定满足 $Q|\psi_{E=0,\pm}\rangle = 0$，是超对称不变的。那这就意味着，对于 $n_+ - n_- \neq 0$ 情形，系统的基态必定是超对称不变的。

前面章节讲对称性自发破缺的时候我们了解到，在物理上，如果基态在某个对称变换下不是不变的，我们就说这个对称性自发破缺了，如果基态在对称变换下保持不变，我们就说对称性没有自发破缺。因此，以上的讨论告诉我们，如果计算出一个超对称系统的 $n_+ - n_- \neq 0$，那不管系统参数如何变化，它的超对称性将始终不会自发破缺！

推广

以上我们只讨论了系统只有一个超荷 Q 的情形，实际上，在大量的超对称模型中，系统总是具有多个超荷，比方说 K 个超荷，分别记为 $Q_1, Q_2, ..., Q_K$。则这时候系统的哈密顿量将为

$$H = Q_1^2 = Q_2^2 ... = Q_K^2. \tag{7.100}$$

不同的超荷之间会满足如下代数关系

$$Q_i Q_j + Q_j Q_i = 0, \quad \text{for} \quad i \neq j. \tag{7.101}$$

很容易验证，所有这些超荷都与哈密顿量对易。

超对称是一个很美妙的理论结构，大量的例子表明，如果一个量子系统具有超对称性，那它将具有许多美妙的性质。我们的世界是不是超对称的呢？很明显不是的，因为我们并没有观察到半整数角动量的量子态 (比方说半整数自旋的粒子) 和整数角动量的量子态 (比方说整数自旋的粒子) 之间的简并和一一匹配。但这很可能是因为我们生活的是一个低能世界，在这个世界中，超对称性自发破缺了！那在多高的能量量级上世界的超对称性才可能“恢复”，从而被观察到呢？目前的答案是，我们并不知道。当前的粒子加速器能够将碰撞粒子的能量加速到十几个 TeV，此前人们曾经幻想在这个能标上能观测到超对称粒子，然而实验结果并没有给出正面的支持。受此影响，当前大量粒子物理学家对粒子物理的超对称性持有怀疑态度。然而超对称的想法是很美妙的，它无疑会在许多其它的物理领域产生应用。



8. 全同性原理与多体量子力学

通常教科书在处理全同性原理时作了一些默认的假定：第一，全同粒子体系的任何物理状态在相差一个相因子的意义上由唯一一个数学上的态矢量来描述。第二，全同粒子体系满足置换对称性。第三，在全同粒子置换下所有物理可观测量都保持不变。这三条假设意味着，全同粒子的置换对称性其实是一种特殊的规范对称性 (而不是通常的对称性)，或者文小刚老师所谓的规范结构。而要从逻辑上解释这种规范对称性的来源，最自然的办法是使用路径积分。

然而，令人吃惊的是，路径积分的推导表明，在 $2+1$ 维时空，全同粒子的置换对称性可以被推广成编织操作。当然，这时候它就不再是一个对称性概念了，但是编织操作允许 $2+1$ 维时空出现既非玻色子又非费米子的任意子。理论研究表明，有一些任意子可以帮助我们实现通用量子计算机，从而在最近一些年引起了人们广泛的兴趣。而且任意子已经不再是一个理论构想，最近的实验已经在两维系统中观测到了任意子。

本章除了讨论 $2+1$ 维的任意子统计和 $3+1$ 的玻色-费米统计之外，我们还将一般性地解释自旋-统计定理。而且我们也会讨论在理论上如何更方便地处理多个全同玻色子的体系或者多个全同费米子的体系的量子力学。但是，本章的很多讨论在理论上偏深，对于初学者来说，**我们建议仅仅阅读本章的 (8.5) 节玻色子和费米子**，等到日后想要更深入地理解全同性原理和多体量子力学时再阅读其它各节。

有一个哲学观点说，世界上没有两件完全一样的事物，莱布尼茨将这个观点简化为，“世界上没有完全相同的两片树叶”。但在量子世界里，这个观点是错的！量子客体可以完全相同，比方说，世界上所有的电子都完全相同，世界上所有的光子也完全相同。更具体一点说，假设有两个电子，分别在 A, B 两地，那么除非你一直持续不断地追踪测量这两个电子的运动轨迹，否则就无从区分 A, B 两地的电子分别是哪一个。但是，由于位置和动量的不确定关系，在量子力学中，精确追踪电子的运动轨迹实际上是不可能的。因此，从根本上来说，人们无法区分两个电子中哪个是哪个。这一章我们要学习的就是这种量子全同性，以及相应的多体量子力学。

比方说，如图 (8.1) 所示，(1)、(2) 两个全同粒子分别从左右两侧入射，相互碰撞后其中一个粒子进入探测器 D，但是探测器探测到的有可能是粒子 (1)(如图 (8.1)(a))，也有可能是粒子 (2)(如图 (8.1)(b))，由于 (1)、(2) 两个粒子全同，(a)、(b) 这两种可能性无法分辨，最终探测器 D 探测到一个粒子的概率幅应该是 (a)、(b) 两种情形的概率幅的叠加。然而两个概率幅的线性叠加有无穷多种可能性，最终计算出来的概率通常都不相同，对于

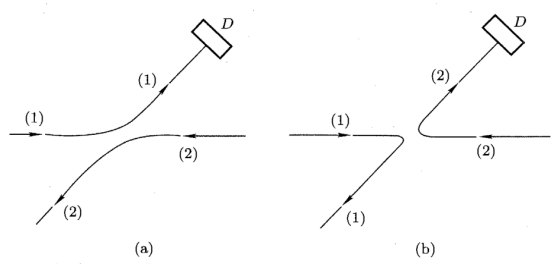

Figure 8.1: 从左右两侧入射的两个全同粒子碰撞以后，探测器 D 探测到粒子的概率幅应该是 (a) 和 (b) 两种情况的叠加。

量子全同粒子我们该使用哪种叠加方式呢？这就是全同性原理所要解决的问题。

另一方面，虽然有时候研究单个微观粒子的量子力学就能得到关于世界的一些重要规律。但是，从根本上说，决定我们这个世界的绝非单个粒子，甚至也不是几个粒子，而是涉及到大量的全同粒子。比方说，分数量子霍尔效应就涉及到大量电子的相互作用，如果我们仅仅研究单个电子在磁场中的规律，那就无法解释分数量子霍尔效应。因此我们也有必有研究如何才能更方便地处理多个全同粒子的体系的量子力学。

8.1 * 时间演化算符与路径积分

8.1.1 坐标表象中的时间演化算符

在前面的第二章中我们已经知道，一个量子系统的动力学演化规律完全由时间演化算符 $U(t_f,t_i)$ 决定。$U(t_f,t_i)$ 是一个幺正算符，当将它作用在 t_i 时刻的初态 $|\psi(t_i)\rangle$ 上，我们就可以得到系统在 t_f 时刻的量子态 $|\psi(t_f)\rangle$，即 $|\psi(t_f)\rangle = U(t_f,t_i)|\psi(t_i)\rangle$。很显然，$U(t,t)=1$，并且我们也知道 $U(t_f,t_i)=\exp\big(-iH(t_f-t_i)/\hbar\big)$。一般地，时间演化算符满足如下基本方程

$$U(t_f,t_i)=U(t_f,t)U(t,t_i), \tag{8.1}$$

这个方程的含义非常简单，即，先将系统从 t_i 时刻演化到某个中间 t 时刻，接着再从 t 时刻演化到 t_f 时刻，其效果就相当于将系统从 t_i 时刻演化到 t_f 时刻。

所有量子系统的时间演化算符都必须满足方程 (8.1)，同时还得是一个幺正算符。在理论上，给出一个量子系统就是要给出它的时间演化算符表达式，而幺正性和方程 (8.1) 就是对这些时间演化算符的基本限制。换言之，只有满足方程 (8.1) 的时间演化算符才能从理论上定义一个量子系统。

有哪些可能的时间演化算符可以满足方程 (8.1) 呢？$U(t_f,t_i)=\exp\big(-iH(t_f-t_i)/\hbar\big)$ 显然是一个回答，这个回答是用哈密顿量来给出的，从第二章中我们已经知道，这个回答可以方便地给出薛定谔方程，而且在这一回答中，由于哈密顿算符是厄密算符，所以相应的时间演化算符必然是幺正的[1]。但实际上，人们也可以在坐标表象中利用拉格朗日量来回答这个问题，这就是所谓的量子力学的路径积分表述。当然，对于路径积分给出来的坐标表象时间演化概率幅，我们需要额外证明其幺正性。对于单粒子情形，在本章附录中我们通过直接从哈密顿量形式的时间演化算符 (它必然幺正) 推导出路径积分形式的时间演化公式完成了这一证明。

具体来说，为了得出路径积分形式的时间演化公式，我们首先要取坐标表象 (更一般的，也就是取量子系统基本动力学变量的本征态表象)。对于一个 N 个全同粒子的体系，我们以有序坐标组 $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N\}$ 表示这 N 个粒子的坐标，并将这一组坐标笼统地简记为 $\{\mathbf{x}\}$，称之为系统的一个位形[2]。当然，一个粒子的动力学变量除了有坐标之外，还可能携带有自旋等等内部自由度，为了符号简洁起见，我们将标记粒子局域内部自由度的量子数也包括在坐标记号 $\mathbf{x}$ 里面，比方说，对于一个自旋量子数为 m_s 的粒子，我们的 $\mathbf{x}$ 其实是代表 $(\mathbf{x},m_s)$ 这一对变量。推而广之，我们这里的 $\mathbf{x}$ 实际上刻画了粒子所携带的所有局域信息，这些局域信息是和粒子一起运动的。总之，系统的坐标本征态可以记为 $|\{\mathbf{x}\}\rangle$，它满足如下完备性关系

$$\sum_{\{\mathbf{x}\}}|\{\mathbf{x}\}\rangle\langle\{\mathbf{x}\}|=1. \tag{8.2}$$

式中的求和号 $\sum_{\{\mathbf{x}\}}$ 代表对每一个粒子的坐标进行全空间积分，同时对所有离散的局域内部自由度量子数求和。

注意，我们会在两个略有区别的意义上使用自由度这个词，一是用来指系统的动力学变量，另一方面我们也会将粒子本身称作自由度，当我们在后一种意义上使用自由度这个词时，我们通常的说法是称作物理自由度。

局域性与非局域性

坐标表象的好处之一是，我们很容易利用它来定义局域性 (Locality) 的概念。物理学中关于局域性有各种不同的定义，为了帮助读者直观理解局域性，我们这里直接将局域定义成时空局部或者说时空邻域[3]。因此所谓的局域物理自由度，就是坐标在同一个时空邻

[1]在第二章中，我们实际上是反过来利用时间演化算符的幺正性一般性地证明哈密顿算符是厄密算符。

[2]这只是一种方便的说法，由于粒子全同性，真正的位形应该是 $\{\mathbf{x}\}$ 在所有可能置换下的等价类。

[3]考虑到相对论的话，时空邻域的概念其实需要进一步精确化。

域内的那些物理自由度，比如 A 点附近的一个点粒子就是 A 点附近的一个局域物理自由度。如果 A 点附近有好几个靠得很近的粒子，那么它们都属于 A 附近的局域物理自由度。当然，有时候我们也直接将点粒子本身称作局域物理自由度，因为在坐标表象中，每个时刻点粒子都只能待在一个局域时空点附近。

另外，我们还要引入**局域信息和非局域信息**的概念。所谓的局域信息，就是单独一个时空邻域内的局域物理自由度所携带的信息。但是，前面第二章中我们学习过量子纠缠，我们知道，两个不同空间点的粒子可以通过纠缠处于一个整体的纠缠态，这时候每一个粒子本身都不携带这个纠缠态的信息，两个不同空间点的粒子作为一个整体才携带了纠缠态的信息。这种由类空相间的两个不同时空邻域内的局域物理自由度作为一个整体携带的纠缠态信息就是一种非局域信息。类空相间当然是一个比较精确的说法，它即是指这样的两个邻域之间无法传递信息。在非相对论量子力学中，由于光速非常大，这时候与类空相间对应的概念其实就是给定时刻，两个距离比较远的空间邻域。两体纠缠能够产生非局域信息，多体纠缠当然也能产生非局域信息。当然，如果我们将处于纠缠态的一对粒子放到一起，使得它们处于同一个时空邻域内，那这时候这个纠缠态的信息就变成了局域信息，这时候我们就可以通过对邻域内的这一对电子进行联合测量来提取这个局域信息。注意，如果纠缠的这一对粒子离得很远，类空相间，那它们的纠缠态信息就是非局域信息，这时候没有仪器可以实现对两者的联合测量，这其实是一种更深层次的局域性原理。

总之，对量子纠缠的学习告诉我们，即使是在局域的坐标表象中，量子系统依然可能产生非局域信息。当然，量子纠缠的非局域信息是将多个局域物理自由度纠缠起来产生的，比方说你让一对电子处于某个贝尔态中，然后将这两个电子分别局域在不相邻近的 A、B 两个不同点，这时候，这个贝尔态的量子信息当然是非局域的，但作为一个整体携带这种非局域信息的两个电子各自都是局域的。正因为如此，在最基本的层次上，量子纠缠态的信息和局域信息并不完全相互独立，这种不相互独立就是可以利用量子纠缠来进行量子隐形传态将局域信息从 A 地隐形传送到 B 地的关键。

但是，可以设想我们考察的不是一个量子系统最基本的物理自由度，而是系统的有效物理自由度，比方说在分数量子霍尔效应中我们不考察最基本的电子，而考察系统的任意子激发。这时候系统的某些非局域信息就完全有可能和这些有效物理自由度的局域信息相互独立。

回到我们前面研究的 N 个全同粒子的系统。上面的分析告诉我们，如果我们考察的这 N 个全同粒子是有效物理自由度的话，那就有可能还有一些与局域信息 $\{\mathbf{x}\}$ 完全独立的非局域信息。我们用一个独立的量子数 n 来区分这些非局域信息，$n=1,2,3...M$。如此一来，我们的坐标表象本征态就应该进一步推广成 $|\{\mathbf{x}\},n\rangle$，而原来的完备性关系 (8.2) 就应该推广成，

$$\sum_{\{\mathbf{x}\},n} |\{\mathbf{x}\},n\rangle\langle\{\mathbf{x}\},n| = 1. \tag{8.3}$$

在这个坐标表象下，我们可以把时间演化算符 $U(t_b,t_a)$ 表示成

$$\langle\{\mathbf{x}\}_b,m|U(t_b,t_a)|\{\mathbf{x}\}_a,n\rangle = \langle\{\mathbf{x}\}_b,m,t_b|\{\mathbf{x}\}_a,n,t_a\rangle. \tag{8.4}$$

等式右边 $|\{\mathbf{x}\},n,t\rangle$ 这样的记号表示 $|\{\mathbf{x}\},n,t\rangle = \exp(iHt/\hbar)|\{\mathbf{x}\},n\rangle$, 它其实是海森堡绘景中坐标算符的本征态 (海森堡绘景中坐标算符要随着时间演化，因此其本征态也依赖于时间)[4]。通过利用完备性关系 (8.3), 现在我们就可以把时间演化算符的基本方程 (8.1) 重新表示成

$$\begin{aligned}&\langle\{\mathbf{x}\}_f,m|U(t_f,t_i)|\{\mathbf{x}\}_i,n\rangle\\ =&\sum_{\{\mathbf{x}\}_b,k}\langle\{\mathbf{x}\}_f,m|U(t_f,t_b)|\{\mathbf{x}\}_b,k\rangle\langle\{\mathbf{x}\}_b,k|U(t_b,t_i)|\{\mathbf{x}\}_i,n\rangle.\end{aligned}\tag{8.5}$$

或者也可以用等式 (8.4) 右边的记号重写为 [5]

$$\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_i,n,t_i\rangle=\sum_{\{\mathbf{x}\}_b,k}\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_b,k,t_b\rangle\langle\{\mathbf{x}\}_b,k,t_b|\{\mathbf{x}\}_i,n,t_i\rangle.\tag{8.6}$$

只要知道了 $\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_i,n,t_i\rangle$, 我们就很容易求出任何波函数的时间演化。这是因为，根据 $|\psi(t_f)\rangle = U(t_f,t_i)|\psi(t_i)\rangle$, 我们可以得到

$$\begin{aligned}\langle\{\mathbf{x}\}_f,m|\psi(t_f)\rangle&=\sum_{\{\mathbf{x}\}_i,n}\langle\{\mathbf{x}\}_f,m|U(t_f,t_i)|\{\mathbf{x}\}_i,n\rangle\langle\{\mathbf{x}\}_i,n|\psi(t_i)\rangle\\&=\sum_{\{\mathbf{x}\}_i,n}\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_i,n,t_i\rangle\langle\{\mathbf{x}\}_i,n|\psi(t_i)\rangle.\end{aligned}\tag{8.7}$$

式中 $\langle\{\mathbf{x}\},n|\psi(t)\rangle = \psi(\{\mathbf{x}\},n,t)$ 就是系统的波函数，而上面这个方程给出的就是波函数的时间演化规律。

8.1.2 单粒子路径积分

路径积分不是直接告诉我们量子态如何随时间演化，而是利用作用量给出时间演化算符 $U(t_b,t_a)$ 在坐标表象下的矩阵元 $\langle\{\mathbf{x}\}_b,m,t_b|\{\mathbf{x}\}_a,n,t_a\rangle$ 的计算公式。下面我们先就单粒子情形阐述费曼得出路径积分公式的基本思想，然后给出单粒子情形的费曼路径积分公式，并解释这个公式的由来，最后再具体验证它满足基本方程 (8.6)。至于对路径积分幺正性的证明，我们放在附录中进行。

如果整个系统只有一个粒子，那所有的信息都只能由这个粒子携带，从而都必定是局域信息，因此标记非局域信息的量子数这时候实际上不存在。这时候坐标表象本征态可以简单记为 $|\mathbf{x}\rangle$, $\mathbf{x}$ 就是这个粒子的坐标 (包括一些可能的局域内部自由度量子数)。而时间演化算符在坐标表象下的矩阵元就是 $\langle\mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$, 它描述的就是，$t_a$ 时刻粒子从位置 $\mathbf{x}_a$ 出发，在 t_b 时刻演化到 $\mathbf{x}_b$ 位置的概率幅。根据 (8.6)，$\langle\mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 满足如下基本方程

$$\langle\mathbf{x}_f,t_f|\mathbf{x}_i,t_i\rangle=\sum_{\mathbf{x}_b}\langle\mathbf{x}_f,t_f|\mathbf{x}_b,t_b\rangle\langle\mathbf{x}_b,t_b|\mathbf{x}_i,t_i\rangle.\tag{8.8}$$

[4]如果用我们在第二章第 4 节《补充材料：海森堡是怎么想到矩阵相乘的》的语言来说，那么 $\{\{\mathbf{x}\},n,t\}$ 就代表系统 t 时刻的一组特定的可确定区分可能性完备集，而等式 (8.4) 右边的 $\langle\{\mathbf{x}\}_b,m,t_b|\{\mathbf{x}\}_a,n,t_a\rangle$ 其实就是 t_a 时刻到 t_b 时刻的一个跃迁幅。因此，第二章第 4 节中跃迁幅的概念其实就是时间演化算符的矩阵表示。

[5]这个方程其实就是我们在第二章第 4 节《补充材料：海森堡是怎么想到矩阵相乘的》中通过 1 的自身等于 1 以及海森堡和费曼的跃迁元乘法规则所导出来的跃迁幅的基本性质。

如果这个粒子是一个经典粒子，我们知道它将遵循最小作用量原理，在 t_a 时刻从 $\mathbf{x}_a$ 出发，t_b 时刻到达 $\mathbf{x}_b$ 位置的所有可能路径中，它将走使得作用量泛函 $S[\mathbf{x}]$ 取极值的那条路径，也就是经典路径。注意，这里出现了作用量这个量，这其实是一个提示，它告诉我们当我们在量子力学层次上计算 $\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 时，作用量 S 也可能很重要。最早注意到这个提示的是狄拉克，但直到费曼才将这个提示发展成一个关于量子力学的自洽理论，也就是路径积分。费曼注意到，对于一个量子粒子而言，由于其位置和动量不可同时确定，所以在量子的层次上，粒子的运动路径任何情况下都无法确定 (因为确定了一条运动路径就意味着位置和速度，或者说动量，的同时确定)。海森堡同样知道这一点，但海森堡的看法是，这说明粒子的运动路径不可观测，这个概念应该从量子力学理论中抛弃掉。海森堡当然并没有错，在量子力学中，运动路径的确不再是一种可观测的物理实在，在这个意义上，确定的路径确实是不存在的!

但是费曼认为，作为一个发展理论的概念，路径可以被保留。这就有点像电动力学中，矢量势也不可观测，但矢量势的概念对于电动力学理论依然很重要。费曼认识到量子力学的不确定性来自于叠加，如果电子自旋的 z 分量无法确定，那就意味着电子处在叠加态，需要将电子的自旋 $1/2$ 态和自旋 $-1/2$ 态叠加起来。类似的，费曼认为，粒子运动路径的无法确定意味着有许多运动路径都对概率幅 $\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 有贡献，而我们需要把这些不同路径的贡献叠加起来。这是因为，如果只有一条运动路径对 $\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 有贡献，就像经典力学里面的最小作用量路径一样，那粒子的运动路径就确定了，而这就和量子力学矛盾了。不过，与电子自旋叠加态的不同之处在于，电子的自旋 $1/2$ 态和 $-1/2$ 态本身都可观测，所以我们可以讨论这两个自旋态本身的叠加。但是，任何一条确定路径本身在物理观测上都不存在，所以当然也并没有不同路径本身相叠加的概念，我们叠加的是按不同路径计算出来的对 $\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 这个量的贡献。

进一步，费曼认为，如果我们完全不限制粒子的演化过程，那它的运动路径就会具有最大的量子不确定性，这时候所有可能路径都会对 $\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 有贡献，相应的我们就应该将所有这些贡献都加起来。

每一条运动路径对 $\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 的贡献是多少呢? 费曼在狄拉克的工作基础上提出，贡献是 $e^{iS[\mathbf{x}]/\hbar}$, $S[\mathbf{x}]$ 就表示这条路径的作用量。因此费曼提出他的路径积分公式

$$\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle = \sum_{\mathbf{x}(t)} e^{i\frac{S[\mathbf{x}(t)]}{\hbar}}. \tag{8.9}$$

式中的求和是对所有可能路径求和，$\mathbf{x}(t)$ 表示一条 t_a 时刻起于 $\mathbf{x}_a$, t_b 时刻到达 $\mathbf{x}_b$ 的运动路径，$S[\mathbf{x}(t)]$ 表示这条路径的作用量，它是拉格朗日量的时间积分

$$S[\mathbf{x}(t)] = \int_{t_a}^{t_b} dtL(\mathbf{x}(t),\dot{\mathbf{x}}(t),t). \tag{8.10}$$

至于如何对所有路径进行求和，这其实是一个重要的数学问题，我们将会在本章附录中进一步讨论。

为什么每条路径的贡献是 $e^{iS[\mathbf{x}]/\hbar}$ 呢? 大致的解释如下：首先，为了使得路径求和最终计算出来的结果 $\langle \mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$ 满足基本方程 (8.8), 每条路径的贡献就必须要取 $e^{iS/\hbar}$ 这样的指数形式，而且指数上的 S 应该是某个量沿着路径的时间积分。进一步，为了满足幺正性，

这个 S 只能是相应路径的作用量，它是拉格朗日量的时间积分。关于 S 为什么要是作用量，在本章的附录中我们将会看得很清楚，下面我们来具体验证 (8.9) 满足方程 (8.8)。

为了表达清晰，我们引入一些记号，我们将一条 t_a 时刻从 $\mathbf{x}_a$ 发出，t_b 时刻到达 $\mathbf{x}_b$ 的路径 $\mathbf{x}(t)$ 记为，$\mathbf{x}(t):a\to b$, 称之为 $\{t_a,\mathbf{x}_a\}\to\{t_b,\mathbf{x}_b\}$ 路径，相应的作用量记为 $S_{a\to b}[\mathbf{x}(t)]$。则根据费曼的路径积分公式 (8.9), 我们有

$$\begin{aligned}&\sum_{\mathbf{x}_b}\langle\mathbf{x}_f,t_f|\mathbf{x}_b,t_b\rangle\langle\mathbf{x}_b,t_b|\mathbf{x}_i,t_i\rangle\\=&\sum_{\mathbf{x}_b}\Big(\sum_{\mathbf{x}(t):b\to f}e^{i\frac{S_{b\to f}[\mathbf{x}(t)]}{\hbar}}\Big)\Big(\sum_{\mathbf{x}(t):i\to b}e^{i\frac{S_{i\to b}[\mathbf{x}(t)]}{\hbar}}\Big)\\=&\sum_{\mathbf{x}_b}\sum_{\mathbf{x}(t):i\to b\to f}e^{i\Big(S_{i\to b}[\mathbf{x}(t)]+S_{b\to f}[\mathbf{x}(t)]\Big)/\hbar}.\end{aligned}\tag{8.11}$$

式中求和下标 $\mathbf{x}(t):i\to b\to f$ 表示由一条 $\{t_i,\mathbf{x}_i\}\to\{t_b,\mathbf{x}_b\}$ 路径结合一条 $\{t_b,\mathbf{x}_b\}\to\{t_f,\mathbf{x}_f\}$ 路径所形成的路径，它其实是一条 $\{t_i,\mathbf{x}_i\}\to\{t_f,\mathbf{x}_f\}$ 路径，只不过这条路径中间的 $\{t_b,\mathbf{x}_b\}$ 点被固定了，式 (8.11) 中的求和号 $\sum_{\mathbf{x}(t):i\to b\to f}$ 就是表示对所有这种路径求和。但是，当我们如式 (8.11) 中那样进一步对所有 $\mathbf{x}_b$ 求和时，那这个中间 $\{t_b,\mathbf{x}_b\}$ 点就不再是固定的了，因此实际上 (8.11) 中的两个求和号 $\sum_{\mathbf{x}_b}\sum_{\mathbf{x}(t):i\to b\to f}$ 联合起来所代表的，就是对所有 $\{t_i,\mathbf{x}_i\}\to\{t_f,\mathbf{x}_f\}$ 路径求和，因此就是 $\sum_{\mathbf{x}(t):i\to f}$。另一方面，注意到

$$\begin{aligned}&S_{i\to b}[\mathbf{x}(t)]+S_{b\to f}[\mathbf{x}(t)]=\\=&\int_{t_i,\mathbf{x}_i}^{t_b,\mathbf{x}_b}dtL(\mathbf{x},\dot{\mathbf{x}},t)+\int_{t_b,\mathbf{x}_b}^{t_f,\mathbf{x}_f}dtL(\mathbf{x},\dot{\mathbf{x}},t)\\=&\int_{t_i,\mathbf{x}_i}^{t_f,\mathbf{x}_f}dtL(\mathbf{x},\dot{\mathbf{x}},t)=S_{i\to f}[\mathbf{x}(t)].\end{aligned}\tag{8.12}$$

所以我们可以将 (8.11) 式重写成

$$\begin{aligned}&\sum_{\mathbf{x}_b}\langle\mathbf{x}_f,t_f|\mathbf{x}_b,t_b\rangle\langle\mathbf{x}_b,t_b|\mathbf{x}_i,t_i\rangle\\=&\sum_{\mathbf{x}(t):i\to f}e^{iS_{i\to f}[\mathbf{x}(t)]/\hbar}=\langle\mathbf{x}_f,t_f|\mathbf{x}_i,t_i\rangle.\end{aligned}\tag{8.13}$$

很显然，结果正好给出 (8.8)。因此这就验证了费曼路径积分公式 (8.9) 的确满足时间演化的基本方程。

以上就是对于单粒子坐标表象下的时间演化算符的路径积分表述。这个表述和薛定谔方程其实是等价的，原则上，只要我们利用费曼给出来的公式 (8.9) 计算出任意 $\langle\mathbf{x}_b,t_b|\mathbf{x}_a,t_a\rangle$，那根据方程 (8.7), 我们就能得到波函数的时间演化

$$\psi(\mathbf{x},t)=\sum_{\mathbf{x}'}\langle\mathbf{x},t|\mathbf{x}',t'\rangle\psi(\mathbf{x}',t').\tag{8.14}$$

这和求解薛定谔波动方程所得到的波函数演化完全等价。

对于多个全同粒子情形，人们同样可以找到相应 $\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_i,n,t_i\rangle$ 的路径积分公式。不过这时候情况要稍微复杂一些，时间演化的基本方程 (8.6) 以及幺正性的要求允许 $\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_i,n,t_i\rangle$ 有更多的可能性。下一节我们将开始探讨这一问题。

8.2 * 路径的拓扑类和编织

首先**说明**一下，这一节标题中的编织英文是 Braid，在文献中它通常是特指 $2+1$ 维情形的，在中文里通常将相应的群 (后面会定义这个概念) 翻译成编织群或者辫子群，两者没有区别。但我们这里是要统一处理 $2+1$ 维情形和 $3+1$ 维情形，所以我们这几节中的编织具有更广一些的含义，为了以示区别，我们将把 $2+1$ 维情形的相应群称作辫子群，而将 $2+1$ 维情形和 $3+1$ 维情形的群统称为编织群 (虽然在 $3+1$ 维它其实是一个置换群)。

假设我们将 N 个全同粒子在 t_a 时刻的 N 个位置标记为 $\{\mathbf{x}_{1,a},\mathbf{x}_{2,a},...,\mathbf{x}_{N,a}\}$，同样将在 t_b 时刻的 N 个位置标记为 $\{\mathbf{x}_{1,b},\mathbf{x}_{2,b},...,\mathbf{x}_{N,b}\}$。并且对于这些粒子从 t_a 时刻的位形 $\{\mathbf{x}\}_a=\{\mathbf{x}_{1,a},\mathbf{x}_{2,a},...,\mathbf{x}_{N,a}\}$ 运动到 t_b 时刻的位形 $\{\mathbf{x}\}_b=\{\mathbf{x}_{1,b},\mathbf{x}_{2,b},...,\mathbf{x}_{N,b}\}$，假设我们在时空中画出每一个粒子的运动路径，很明显这些路径可能会如图 (8.2) 所示。但是由于全同粒

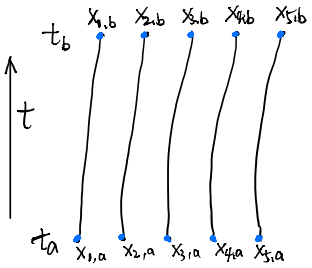

Figure 8.2: 多个全同粒子从 t_a 时刻到 t_b 时刻的可能路径。

子不可分辨，所以我们其实根本不可能知道 t_b 时刻的哪个位置来源于 t_a 时刻的哪个粒子，这些粒子完全有可能中途发生置换，因此图 (8.3) 所示的路径同样可能，而且我们实际上

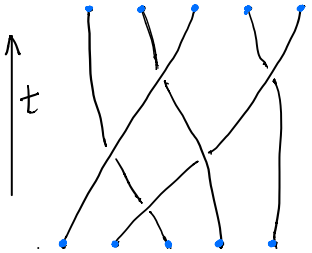

Figure 8.3: 多个全同粒子从 t_a 时刻到 t_b 时刻的另一种可能路径。

无法在物理上将它们和图 (8.2) 所示的路径区分开来。实际上，从 $\{\mathbf{x}\}_a$ 位形到 $\{\mathbf{x}\}_b$ 位形，图 (8.3) 这样的有比较复杂交叉关系的路径才是典型的。

为了方便下面的分析，现在我们将图 (8.3) 所示的典型路径进行连续形变，但是保持路径之间的交叉关系不变，读者很容易看出，我们总可以将图 (8.3) 这样的路径连续变形为图 (8.4) 所示的这种比较标准的形式。实际上，多个全同粒子从 t_a 时刻 $\{\mathbf{x}\}_a$ 位形到 t_b 时刻 $\{\mathbf{x}\}_b$ 位形的任何路径，我们总是可以在保持路径之间交叉关系的前提下，将它们连续变形为类似于图 (8.4) 所示的这种标准形式。在数学上，对于两组起末两端固定的路径，如果在保持交叉关系的前提下可以让它们通过路径的连续形变相互过渡，我们就说这**两组**

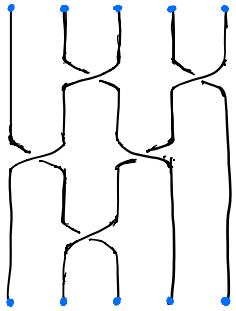

Figure 8.4: 全同粒子路径的标准形式。

路径在拓扑上等价，或者说它们属于全同粒子路径的同一个拓扑等价类！根据这个定义图 (8.3) 所示的路径与图 (8.4) 所示拓扑等价，属于同一拓扑等价类。相反，图 (8.2) 所示的路径与图 (8.3) 所示的路径就属于两个不同的拓扑等价类。因此，给定 t_a 时刻与 t_b 时刻的位形，全同粒子的路径可以分成多个不同的拓扑等价类，每一个拓扑等价类的路径都可以连续形变为类似于图 (8.4) 这样的标准形式。

从图 (8.4) 中我们可以发现，如果我们在纸面上表现全同粒子的路径，那这些标准路径非常有规律，实际上，你很容易发现它可以由对路径的一些基本编织操作生成。为了让读者看得更清楚一点，下面我们来定义这些基本编织操作，首先我们把 N 个空间坐标标记成 $1,2,...,N$，并按照从左到右的顺序排成一行表现在纸面上。我们定义编织操作 σ_i 为将第 i 个位置处的路径逆时针与相临的第 $i+1$ 位置的路径交换顺序，如图 (8.5) 左边这幅图所示。如果这种顺序交换是按照顺时针方式进行，如图 (8.5) 右边这幅图所示，那我们就将之定义成 σ_i^{-1}，称作 σ_i 的逆操作，至于为什么是逆操作我们很快就能看到。如果忘

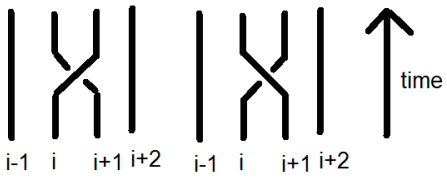

Figure 8.5: 左图是 σ_i，右图是 σ_i^{-1}.(图片来自于网络)

记时间坐标，只关心这两个编织操作在空间上形成的轨道，那从上方往下俯视，σ_i 和 σ_i^{-1} 将如图 (8.6) 所示。

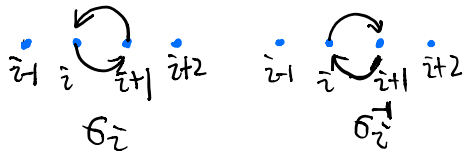

Figure 8.6: 左图是 σ_i，右图是 σ_i^{-1}.

我们将时间上先后进行的两个编织操作定义为这两个编织操作相乘，先进行的编织操作写在乘式的右边，后进行的编织操作写在乘式的左边。当然，有时候两个编织操作的先后顺序对结果并没有影响。我们定义任何两个编织操作相乘的结果依然是一个编织操作(通常是更复杂一点的编织操作)。利用这样的编织乘法，我们就可以把图 (8.4) 所对应的编织操作写成 $(\sigma_4\sigma_2)(\sigma_3^{-1}\sigma_1)(\sigma_2)$，如图 (8.7) 所示。根据这样的编织乘法，读者很容易验

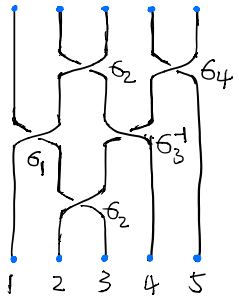

Figure 8.7: 编织操作 $(\sigma_4\sigma_2)(\sigma_3^{-1}\sigma_1)(\sigma_2)$.

证，$\sigma_i\sigma_i^{-1}$ 以及 $\sigma_i^{-1}\sigma_i$ 的结果在拓扑上都等价于不进行任何编织，如图 (8.8) 所示, 我们记为 $\sigma_i\sigma_i^{-1}=\sigma_i^{-1}\sigma_i=1$，这就是为什么我们称 σ_i^{-1} 为 σ_i 的逆操作的原因。

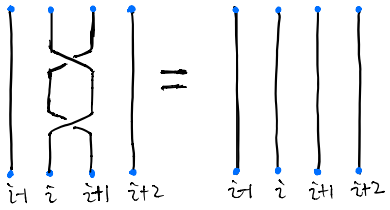

Figure 8.8: $\sigma_i^{-1}\sigma_i=1$.

通过将路径进行连续形变，我们很容易证明如图 (8.9) 所示的等价关系。我们可以将

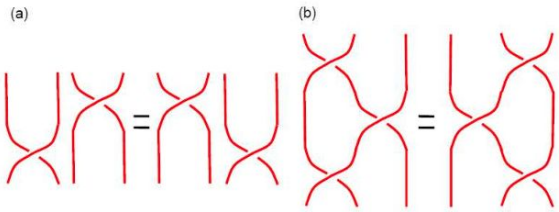

Figure 8.9: 左图表示当 $|i-j|\geq 2$ 时，$\sigma_i\sigma_j=\sigma_j\sigma_i$。右图表示 $\sigma_i\sigma_{i+1}\sigma_i=\sigma_{i+1}\sigma_i\sigma_{i+1}$。(图片来自于网络)

这两个等价关系写成如下方程

$$\begin{aligned}&\sigma_i\sigma_j=\sigma_j\sigma_i,\quad \text{for } |i-j|\geq 2\\&\sigma_i\sigma_{i+1}\sigma_i=\sigma_{i+1}\sigma_i\sigma_{i+1}.\end{aligned}\tag{8.15}$$

上面这个式子的后一个方程尤为基本和重要，虽然我们依然是用纸面上的基本关系来表达它，但正是因为有了这个方程，我们在纸面上表现的这些路径才能算是 $2+1$ 维时空或者 $3+1$ 维时空中的路径，而不仅仅只是纸面上的路径。

每一个编织操作我们称之为一个编织群元，很显然，所有的编织群元都可以通过将一些 σ_i 或者 σ_i^{-1} 这样的基本编织群元乘起来生成。比方说图 (8.4) 所对应的编织操作 $(\sigma_4\sigma_2)(\sigma_3^{-1}\sigma_1)(\sigma_2)$ 就是一个编织群元。两个编织群元 g_1 和 g_2 的乘积 g_2g_1 就定义为相应的两个编织操作相乘，g_2g_1 依然是对路径的一种编织，因此也是一个编织群元。也即是说，所有可能的编织群元的集合在乘积关系下是封闭的。不进行任何编织就称为单位群元，记为 1。读者想一下就能明白，对于任何一种复杂的编织操作 g，我们都可以找到一个“反向编织操作” g^{-1}，使得两者乘起来的结果拓扑等价于 1，我们称 g^{-1} 为 g 的逆。因此，任何一个编织群元都有一个逆元。在数学上，一个集合，如果其元素之间可以进行某种乘法运算 [6]，并且集合本身在这种乘法下保持封闭，而且集合中存在一个单位元，且任何元素都有一个乘法逆元，这样的集合数学家称之为群，集合中的元素就称之为群元。因此，对 N 个粒子的路径的所有可能编织操作构成了一个群，我们可以称之为**编织群**。σ_i 就称为编织群的生成元，它们满足基本代数关系 (8.15)。

但是，2 维空间 (加上一维时间就是 $2+1$ 维时空) 中粒子的编织群和 3 维空间粒子的编织群不一样！下面我们就来研究这两者的区别。

辫子群与置换群，$2+1$ 维与 $3+1$ 维的区别

编织群由一些基本的编织操作 σ_i 生成，σ_i 表示将一对相邻粒子逆时针交换位置，如果是顺时针交换位置我们就称之为 σ_i^{-1}，如图 (8.6) 所示。但是，逆时针编织和顺时针编织的区别是 $2+1$ 维时空的特殊情况，如果在 $3+1$ 维时空，那这两者实际上拓扑等价，即 $\sigma_i^{-1}=\sigma_i$，也即是说这时候图 (8.5) 中左右两种情况实际上拓扑等价 (之所以左右两边的图形看起来不等价，是因为我们无法在纸面上在表现出时间维的同时表现出 3 维空间)。换言之，对于 $3+1$ 维时空，编织群的生成元额外满足如下代数关系

$$\sigma_i^2=1. \tag{8.16}$$

即，将同一个基本编织操作连续进行两次等价于不编织，或者说，将两个粒子连续交换位置两次等价于不交换。

为了证明 (8.16) 式，我们注意到 σ_i^2 实际上就等价于将一个粒子围绕着另外一个粒子转一圈。如果我们忘记时间维，将注意力集中在这种编织操作的空间轨道上，我们就能发现，在 2 维空间，一个粒子围绕着另一个粒子转一圈不能拓扑等价于不转圈，如图 (8.10) 右图所示。但是，在 3 维以上空间，由于空间维度更多，给轨道的拓扑形变留下了足够的余地，所以一个粒子围绕着另外一个粒子转一圈完全可以连续变形为不转圈，如图 (8.10) 左图所示。因此，这就证明了在 3 维以及 3 维以上空间，$\sigma_i^2=1$。

正是因为在 $3+1$ 维时空基本编织操作额外满足代数关系式 (8.16)，即将两个粒子逆时针交换位置还是顺时针交换位置其实没有区别，所以这时候的路径编织在一定意义上是

[6] 注意，一般来说群的乘法不一定满足交换律。

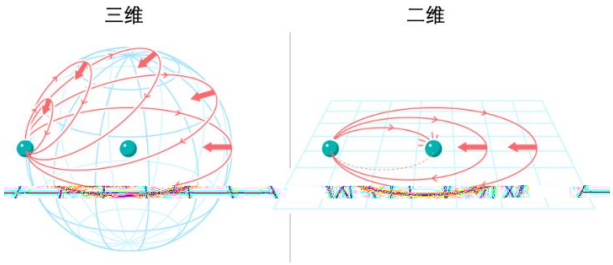

Figure 8.10: 图片来源：5W Infographics / Quanta Magazine

平凡的，我们只需要用 N 个粒子之间的位置置换来考虑问题就足够了。这时候的编织群实际上就是位置置换群，通常记作 S_N。只有在 $2+1$ 维时空，对路径的编织关系才是非平凡的，这时候对 N 个粒子路径的编织群我们称之为辫子群 (Braid Group)，记为 B_N, 因为这时候对路径编织的结果很像发辫。

很明显，N 个全同粒子从 t_a 时刻的 $\{\mathbf{x}\}_a$ 位形演化到 t_b 时刻的 $\{\mathbf{x}\}_b$，其所有可能演化路径可以分成许多个不同的拓扑等价类，每一个拓扑等价类的路径图都可以用类似于图 (8.4) 这样的标准形式来代表，因此与编织群的群元 g 一一对应。对于 $2+1$ 维时空，$g \in B_N$，而对于 $3+1$ 维时空，$g \in S_N$。粒子从 t_i 时刻经过 g_1 类路径演化到 t_b 时刻，接着再从 t_b 时刻经过 g_2 类路径演化到 t_f 时刻，其总效果就是粒子从 t_i 时刻经过 g_2g_1 类路径直接演化到 t_f 时刻，可以写成

$$i \xrightarrow{g_1} b \xrightarrow{g_2} f \Leftrightarrow i \xrightarrow{g_2 g_1} f. \tag{8.17}$$

这里我们定义 g_1 和 g_2 两类路径相乘为将它们的标准路径图在时间上先后对接起来，如此一来，**这种路径拓扑类的乘法与相应编织群元的乘法刚好一一对应**。

编织群作为一个群，它在数学上允许一种美妙的附加结构，称之为群的表示。具体来说，就是对于每一个群元 g，我们都将它对应到希尔伯特空间的一个幺正算符 $U(g)$，并且让这种对应保持群的乘法关系，即满足

$$U(g_2)U(g_1) = U(g_2 g_1). \tag{8.18}$$

这种满足 (8.18) 式的从一个群到希尔伯特空间幺正算符的映射关系就称为群的幺正表示。当然，这种幺正表示有可能是平凡的，比方说 $U(g) = 1$ 显然满足 (8.18) 式，从而构成编织群的幺正表示，但这个幺正表示显然是一个平凡的表示。然而编织群，无论是 $2+1$ 维的 B_N 还是 $3+1$ 维的 S_N，都可以有非平凡的幺正表示。正因为编织群可以有这些非平凡的幺正表示，就使得多个全同粒子的路径积分有了更多的可能性。

8.3 * 全同粒子统计

8.3.1 全同粒子路径积分

下面我们来考察 N 个全同粒子的路径积分公式。首先，假设我们将粒子的所有路径按照拓扑等价关系分类，每一类对应一个编织群元 g，并且假设我们可以引入一个非平凡

的 $U(g)$。由于 $U(g)$ 只依赖于路径的拓扑类，在路径的连续形变下将保持不变，而路径的连续形变是一种局域操作，它会改变系统的局域信息，这就说明 $U(g)$ 对系统的局域信息不能有作用，它只能作用在非局域信息的指标 $n=1,2,...,M$ 上。考虑到这一点我们就可以给出如下路径积分公式

$$\langle\{\mathbf{x}\}_b,m,t_b|\{\mathbf{x}\}_a,n,t_a\rangle=\frac{1}{|G|}\sum_g\Big[\langle m|U(g)|n\rangle\sum_{\{\mathbf{x}(t)\}\in g:a\to b}e^{iS_{a\to b}[\{\mathbf{x}(t)\}\in g]/\hbar}\Big]. \tag{8.19}$$

式中下标 $\{\mathbf{x}(t)\}\in g$ 表示属于拓扑类 g 的一组路径，式中最后的求和 $\sum_g$ 表示对所有的路径拓扑类求和，也即是对编织群的所有群元求和，式中的 $|G|$ 表示编织群的群元总数目。

相比于单粒子路径积分公式，(8.19) 的关键不同之处在于将多粒子路径按照拓扑等价类来处理了，一个等价类之内的每一条路径对路径积分的贡献和单粒子情形类似，都是 $e^{iS/\hbar}$，但这不是全部贡献，对于每一个等价类 g，我们还要额外乘以幺正表示 $U(g)$。

根据我们在附录中关于单粒子路径积分幺正性的证明可以类似地知道，$\sum_g\sum_{\{\mathbf{x}(t)\}\in g:a\to b}e^{iS_{a\to b}[\{\mathbf{x}(t)\}\in g]/\hbar}$ 是幺正的，又由于 $U(g)$ 是幺正算符，所以我们可以知道 (8.19) 给出来的 $\langle\{\mathbf{x}\}_b,m,t_b|\{\mathbf{x}\}_a,n,t_a\rangle$ 是幺正的。这就证明了路径积分公式 (8.19) 的幺正性。

下面我们来验证 (8.19) 式满足时间演化的基本方程 (8.6)。首先我们注意到由方程 (8.18) 容易有

$$\sum_k\langle m|U(g_2)|k\rangle\langle k|U(g_1)|n\rangle=\langle m|U(g_2g_1)|n\rangle. \tag{8.20}$$

其次，根据 (8.19) 式我们有

$$\begin{aligned}
&\sum_{\{\mathbf{x}\}_b,k}\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_b,k,t_b\rangle\langle\{\mathbf{x}\}_b,k,t_b|\{\mathbf{x}\}_i,n,t_i\rangle\\
=&\frac{1}{|G|^2}\sum_{g_1,g_2}\sum_k\Big[\langle m|U(g_2)|k\rangle\langle k|U(g_1)|n\rangle\\
&\times\sum_{\{\mathbf{x}\}_b}\Big(\sum_{\{\mathbf{x}(t)\}\in g_2:b\to f}e^{iS_{b\to f}[\{\mathbf{x}(t)\}\in g_2]/\hbar}\Big)\Big(\sum_{\{\mathbf{x}(t)\}\in g_1:i\to b}e^{iS_{i\to b}[\{\mathbf{x}(t)\}\in g_1]/\hbar}\Big)\Big]\\
=&\frac{1}{|G|^2}\sum_{g_1,g_2}\Big[\langle m|U(g_2g_1)|n\rangle\times\Big(\sum_{\{\mathbf{x}(t)\}\in g_2g_1:i\to f}e^{iS_{i\to f}[\{\mathbf{x}(t)\}\in g_2g_1]/\hbar}\Big)\Big].
\end{aligned}\tag{8.21}$$

式中的第二个等于号我们利用了路径拓扑类的乘法，以及我们在单粒子路径积分中相关推导的经验。现在，在 (8.21) 式中令 $g_2g_1=g$，并利用 $g_2=gg_1^{-1}$ 消去 g_2，则 $\sum_{g_1,g_2}=\sum_{g,g_1}$，而 (8.21) 式最后的式子就可以写成

$$\frac{1}{|G|^2}\sum_{g,g_1}\Big[\langle m|U(g)|n\rangle\times\Big(\sum_{\{\mathbf{x}(t)\}\in g:i\to f}e^{iS_{i\to f}[\{\mathbf{x}(t)\}\in g]/\hbar}\Big)\Big]. \tag{8.22}$$

注意到 (8.22) 式的被求和式子实际上与 g_1 无关，因此对所有 g_1 的求和就简单地等于群元数目 $|G|$，刚好将表达式中的 $\frac{1}{|G|^2}$ 抵消成 $\frac{1}{|G|}$。

综合以上这些结果，我们就能得到

$$\begin{aligned}&\sum_{\{\mathbf{x}\}_b,k}\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_b,k,t_b\rangle\langle\{\mathbf{x}\}_b,k,t_b|\{\mathbf{x}\}_i,n,t_i\rangle\\=&\frac{1}{|G|}\sum_g\Big[\langle m|U(g)|n\rangle\times\Big(\sum_{\{\mathbf{x}(t)\}\in g:i\to f}e^{iS_{i\to f}[\{\mathbf{x}(t)\}\in g]/\hbar}\Big)\Big]\\=&\langle\{\mathbf{x}\}_f,m,t_f|\{\mathbf{x}\}_i,n,t_i\rangle.\end{aligned}\tag{8.23}$$

这样就完成了对方程 (8.6) 的验证。

8.3.2 2+1 维以及 3+1 维的讨论

正如我们已经说过的，多粒子路径编织群的幺正表示 $U(g)$ 作用在与局域信息相互独立的非局域信息空间上，数学上称这样的希尔伯特空间为编织群的幺正表示空间，这个空间的基矢量是 $|n\rangle, n=1,2,...,M$，空间维数 M 称之为幺正表示的维数。由于假定这样的表示空间和系统的局域信息相互独立，因此系统的任何局域扰动对它都不会产生影响，这就意味着这样的表示空间不会在对系统的局域扰动之下分裂成多个更小的表示空间。数学上常常称具有这种性质的表示空间为不可约表示空间。矩阵 $\langle m|U(g)|n\rangle$ 就称为不可约表示矩阵，或者简称为不可约表示。M 就是这个不可约表示的维数。

数学家告诉我们，由于要满足方程 (8.20)，并非任意维度的希尔伯特空间都能构成编织群 B_N 或者 S_N 的不可约表示空间。最简单的不可约表示当然就是 1 维表示，即 $M=1$。由于 1 行 1 列的幺正矩阵必定是一个相位因子，因此这时候幺正算符 $U(g)$ 在这样的 1 维表示空间的作用只能是一个相位因子，可以记为 $e^{-i\varphi(g)}$。很显然的是，无论是 B_N 还是 S_N 都有一个平凡的一维表示，这时候 $U(g)$ 在这个 1 维表示空间上的作用恒等于 1，我们称这样的表示为恒等表示。

2+1 维情形

但是对于 2+1 维情形，这时候编织群是辫子群 B_N，而 B_N 有非常多的不平凡一维表示。不妨假设逆时针编织 σ_i 在某个一维表示中被表示成了 $e^{-i\theta_i}$，则很显然顺时针编织 σ_i^{-1} 将被表示成 $e^{i\theta_i}$。由于编织群的表示要保持乘法关系，特别的是要保持 (8.15) 式的乘法关系，则我们有 $e^{-i\theta_i}e^{-i\theta_{i+1}}e^{-i\theta_i}=e^{-i\theta_{i+1}}e^{-i\theta_i}e^{-i\theta_{i+1}}$，即有

$$e^{-i\theta_i}=e^{-i\theta_{i+1}}=...=e^{-i\theta}.\tag{8.24}$$

但是，除此之外，辫子群 B_N 无法对 θ 施加更多的约束，也就是说任何相位因子 $e^{-i\theta}$ 都是允许的。这就告诉我们，在一维表示情形中，当我们将两个邻近的 2+1 维全同粒子逆时针交换顺序时 (即逆时针编织)，波函数可以多出一个任意的相位因子 $e^{-i\theta}$，而顺时针交换顺序则会多出 $e^{i\theta}$。由于 θ 任意[7]，所以人们将这样的 2+1 维全同粒子称作任意子！严格一点说是阿贝尔任意子。所谓的阿贝尔是指，在这种一维表示情形下，B_N 的表示矩阵

[7]实际上，在编织张量范畴理论中可以证明，θ 只能取 $2\pi p/m(p,m$ 是两个互素的整数) 的形式。

是相位因子，它们满足乘法交换律，在群理论中，阿贝尔作为形容词指的就是满足乘法交换律！

对阿贝尔任意子的阐述很容易使我们想到，如果 B_N 的不可约表示空间不是一维，而是 2 维以上，那这时候表示矩阵 $\langle m|U(g)|n\rangle$ 就真是一个矩阵而不是可交换的复数了，矩阵不满足乘法交换律，所以对于这种情形人们就把相应的全同粒子称作非阿贝尔任意子。

任意子这种神奇的粒子最早完全是理论凝聚态物理学家的一个预言，但是最近的实验表明，阿贝尔任意子在特定的两维系统中真的存在，人们已经在实验中发现它们了。不仅阿贝尔任意子，许多凝聚态物理学家相信非阿贝尔任意子也能在某些特定的两维系统 (比如某些分数量子霍尔系统) 中存在。关于哪些系统中可能找到非阿贝尔任意子，理论学家已经有了不少预言，然而由于对实验要求很高，到目前为止，虽然有些观测迹象表明它们的确存在，但人们还没有找到很确凿的证据。然而非阿贝尔任意子比阿贝尔任意子更为神奇，也更为重要，其中的一个原因在于，理论研究表明，有一些非阿贝尔任意子可以用来实现通用的容错量子计算！从长远来看，这也许是最有希望的量子计算实现方案。

非阿贝尔任意子实现容错量子计算的原理其实不难理解。为了说清楚这一点，让我们设想对处于 $|\{\mathbf{x}\},n\rangle$ 态的这 N 个任意子进行一个编织操作 g，编织完了之后让这 N 个全同任意子的位形回到 $\{\mathbf{x}\}$，因此整个过程的效果类似于图 (8.4) 所示。很显然，这样的编织操作只会影响指标 n 标记的非局域信息，具体来说，这样的编织操作 g 其实就相当于在 $|\{\mathbf{x}\},n\rangle$ 上作用 $U(g)$，从而有

$$|\{\mathbf{x}\},n\rangle \to U(g)|\{\mathbf{x}\},n\rangle = \sum_m |\{\mathbf{x}\},m\rangle\langle m|U(g)|n\rangle. \tag{8.25}$$

现在，设想我们将要进行计算处理的量子态编码成非局域信息 $|\{\mathbf{x}\},\psi\rangle = \sum_n \psi_n|\{\mathbf{x}\},n\rangle$, 式中复叠加系数 ψ_n 所包含的就是我们要处理的量子信息。则在编织操作 g 之下我们将有

$$|\{\mathbf{x}\},\psi\rangle \to U(g)|\{\mathbf{x}\},\psi\rangle = \sum_m |\{\mathbf{x}\},m\rangle \sum_n \langle m|U(g)|n\rangle\psi_n. \tag{8.26}$$

很显然，编织操作 g 相当于对量子信息 ψ_n 进行了如下计算

$$\psi_m \to \sum_n \langle m|U(g)|n\rangle\psi_n. \tag{8.27}$$

因此，只要合适地利用对非阿贝尔任意子的编织操作，我们就能实现想要的量子计算。

这种计算的幺正矩阵 $\langle m|U(g)|n\rangle$ 是非局域的，系统的任何局域扰动都不能影响它。这一点正是利用非阿贝尔任意子进行量子计算的巨大优点，因为这样的计算过程将会天然对环境的退相干效应产生免疫。这是因为，退相干主要就是因为系统和环境之间存在大量的局域相互作用而产生的，现在既然这样的局域相互作用没法影响 $\langle m|U(g)|n\rangle$，那这样的量子计算当然就会对退相干免疫。寻找有合适非阿贝尔任意子的量子系统以及如何利用编织操作进行具体的量子计算，这些就是当前拓扑量子计算研究的核心课题。在拓扑量子计算中，所谓的计算其实就是一连串的编织操作，如图 (8.11) 所示 (注意图中水平方向才是时间方向)。

然而，一个存在任意子的两维系统通常不会只有一种类型的任意子，而是可以有多种不同的任意子类型，不同类型的任意子当然不是全同粒子，因此我们这里描述的全同粒

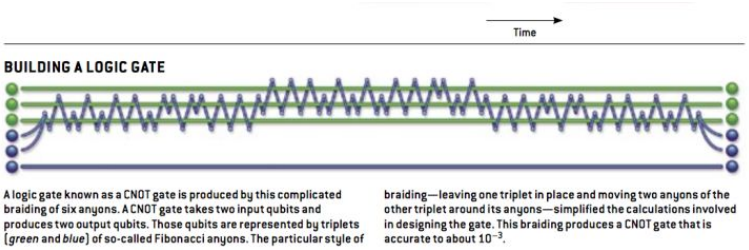

Figure 8.11: 摘自 Collins, Graham P. "Computing with quantum knots." Scientific American 294.4(2006):56-63

子理论其实不能完整地描写任意子系统。对任意子系统的理论描述需要新的数学，这就是所谓的编织张量范畴理论，这是一个非常抽象的数学理论，它当然大大超出了我们这本书的范围，一个相对比较容易入门的读物是 Steven H. Simon 的在线讲义 "Topological Quantum" [8]。

回到我们的 N 个全同粒子系统。由于张成辫子群不可约表示空间的这些非局域信息也是由这 N 个全同任意子作为一个整体携带的，因此可以想见，至少在 N 足够大时，不可约表示空间的维数 M 将随着 N 的数目指数增长，即有

$$M \sim d^N. \tag{8.28}$$

类似这种关系我们其实很熟悉，比方说，每个量子比特的希尔伯特空间是 2 维，因此 N 个量子比特的总希尔伯特空间就是 2^N 维。所以式 (8.28) 中的 d 可以理解为每个任意子的非局域信息的维数，称作非阿贝尔任意子的量子维数。但是，不可思议的是，非阿贝尔任意子的 d 可以不是整数，甚至可以是无理数，比方说有一种被称作斐波那契任意子的非阿贝尔任意子，它的 $d = \frac{1+\sqrt{5}}{2}$。

$3+1$ 维情形

3 维空间或者 3 维以上空间和 2 维情形有根本性的不同，这时候编织群是 S_N，即基本编织操作 σ_i 要额外满足代数关系 $\sigma_i = \sigma_i^{-1}$，或者说将两个粒子逆时针交换顺序和顺时针交换顺序结果是一样的，因此唯一体现路径编织的地方其实仅仅在于粒子的位置置换，S_N 其实就是 N 个粒子的置换群。这会带来一些根本性的不同，最大的不同在于，$2+1$ 维的编织由于逆时针编织和顺时针编织不同，所以必须跟随整个编织操作的时间演化，我们无法在一个固定的时刻定义 $2+1$ 维编织操作。但是，对于粒子的位置置换来说，我们无需追究具体是如何置换的，或者说将两个粒子的位置坐标交换一下这样的操作可以在任何一个固定时刻都有定义。这使得我们可以定义作用在固定时刻量子态空间的置换算符，正如后文我们将要看到的，这样的算符其实是对量子态的一个对称变换，从而使得位置置换成为 $3+1$ 维全同粒子系统的一种对称性，而置换群 S_N 就可以看成是一个对称群。而 $2+1$

[8] http://oxfordtopquantum.tiddlyspot.com/

维的辫子群 B_N 就不能简单理解为对称性，而是一个比对称性更抽象的概念，完整的数学描述要用编织张量范畴。

不过在具体定义 $3+1$ 维的置换对称性之前，我们是统一用路径编织的概念来看待 $2+1$ 维和 $3+1$ 维的，只不过 $3+1$ 维得额外满足 $\sigma_i^2=1$。对于编织群的一维表示 $e^{-i\theta}$ 来说，这意味着 $3+1$ 维要额外满足

$$\left(e^{-i\theta}\right)^2=1. \tag{8.29}$$

这意味着 $e^{-i\theta}$ 不能取任意值，而是仅仅只有 $e^{-i\theta}=\pm 1$ 这两种可能性。在数学上，这两种可能性分别对应置换群 S_N 仅有的两种 1 维不可约表示。

由于在 $3+1$ 维，σ_i 就是将两个粒子的位置对换，而 N 个粒子的任意位置置换总能通过一系列的两两对换来实现 (这里请读者自己用 3 个或者 4 个粒子的位置置换为例来看清这个结论的普遍性)，所以 $e^{-i\theta}=1$ 的可能性就意味着，对 N 个粒子进行任意置换所产生的影响都是不变的因子 1，即对于任何 g，相应的一维表示相位因子均为 $e^{-i\varphi(g)}=1$，这其实就是置换群的恒等表示。这种情形所描述的全同粒子就是所谓的全同玻色子。

而对于 $e^{-i\theta}=-1$ 这种情形，它必然意味着 N 个粒子的任何偶置换都产生因子 1(即相应的 $e^{-i\varphi(g)}=1$)，而任何奇置换都产生因子 -1(即相应的 $e^{-i\varphi(g)}=-1$)。所谓偶置换就是可以实现为偶数个两两对换相乘的置换，而所谓奇置换就是可以实现为奇数个对换相乘的置换。在我们现在的情形中，每个对换会贡献一个 -1，所以偶数个对换相乘产生的因子就是 1，而奇数个对换相乘产生的因子就是 -1。这种情形称作置换群 S_N 的交错表示 (alternating representation)。这种情形所描述的全同粒子就是所谓的全同费米子。

当然，$3+1$ 维编织群 S_N 也有 2 维以上的不可约表示。但是考虑到真实物理系统的一些附加限制，人们可以证明所有这些高于 1 维的不可约表示都不能描述真实的 $3+1$ 维全同粒子系统。证明中所需要的额外物理限制主要有两个：1. 任意粒子都有其反粒子，人们可以产生一对正反粒子，也可以让一对正反粒子相互湮灭。这一条物理限制当然早就被实验验证了，反粒子和反物质的概念可以说是现代物理学的常识。在历史上，反粒子的存在最早是狄拉克 1928 年在理论上预言的，1932 年被 C.D. 安德森用实验证实。2. 粒子间的相互作用具有局域性。这也就是说，在同一时刻北京产生或者湮灭一对正反粒子对于南昌的实验结果不应该有任何影响。具体的证明过程非常复杂，原始论文[9]过于长我也不推荐读者阅读。大体思想也许是证明在上述两个物理约束下，编织群为 S_N 的全同粒子体系携带的所有信息在一定意义上都不能与局域信息相互独立，因此独立非局域信息的维度 $M=1$，从而只能构成 S_N 的 1 维不可约表示，高于 1 维的不可约表示都不可行[10]。

一般来说，物理学和数学有很大的不同，物理学的不可行定理 (No go theorem) 的前提假设很容易出漏洞，人们不只一次地找到过各种物理学不可行定理的例外，因此这么复杂抽象的证明物理学家不一定会真正重视，这可能就是为什么曾经有一段时间有很多人在

[9] S. Doplicher, R. Haag, and J. E. Roberts. Local observables and particle statistics. I. Comm. Math. Phys., 23:199–230, 1971. S. Doplicher, R. Haag, and J. E. Roberts. Local observables and particle statistics. II. Comm. Math. Phys., 35:49–85, 1974.

[10] 一个等价的说法是，高于 1 维表示的准统计实际上等价于一些具有额外局域内部自由度的玻色或者费米子统计。

研究 2 维以上的不可约表示，称之为准统计 (parastatistics)。不过在后来的理论物理研究进展中，人们发现高于 1 维的不可约表示的确没有出现过，没有人发现过任何自洽的物理理论可以容得下它们，实验中更是没有发现有需要准统计的迹象。也许有极小的可能出现意外，但今天人们已经普遍接受 $3+1$ 维只有玻色子和费米子这两种一维表示的结论了。值得注意的是，$3+1$ 维的这种情况和 $2+1$ 维非阿贝尔任意子情形有根本性的不同，后者的确是有丰富物理内涵的。

8.3.3 作为规范对称性的置换对称性

这一节我们将深入探讨置换是如何成为 $3+1$ 维全同粒子体系的一种对称性的。这里主要涉及到如下几个问题：1. 如何定义置换算符。2. 置换算符在全同粒子波函数上如何作用。3. 为什么全同粒子置换是一种特殊的规范对称性。

置换算符

前面我们说过，在 $3+1$ 维中编织群是一个置换群 S_N。而对全同粒子的置换操作并不依赖于具体的置换过程，而是可以在任何一个瞬时定义。比方说，我们可以考察一个两粒子系统的位置本征态 $|\mathbf{x}_1,\mathbf{x}_2\rangle$，注意，由于上面讨论过的原因，对于 $3+1$ 维情形，我们可以假定独立非局域信息的维数 $M=1$，因此无需引入额外的量子数来区分不同的独立非局域信息。假设我们将这两个粒子分别标记为 1、2，并规定狄拉克符号里第一个位置总是用来标记第 1 个粒子的状态，第二个位置总是用来标记第 2 个粒子的状态，因此 $|\mathbf{x}_1,\mathbf{x}_2\rangle$ 就表示一个第 1 个粒子处在 $\mathbf{x}_1$ 位置，第 2 个粒子处在 $\mathbf{x}_2$ 位置的状态 (当然，如我们前面说过的，类似于 $\mathbf{x}$ 位置这样的说法是一个简化说法，实际上 $\mathbf{x}$ 这样的符号中还可以包含粒子的自旋等内部量子数)。从而我们就可以定义置换算符 P_{12}

$$P_{12}|\mathbf{x}_1,\mathbf{x}_2\rangle = |\mathbf{x}_2,\mathbf{x}_1\rangle. \tag{8.30}$$

P_{12} 的作用就是把两个粒子相互对换。我们也可以把这个对换理解为对这两个粒子瞬时进行的一个编织。注意，这和 $2+1$ 维中的编织操作有根本性的不同，后者依赖于我们具体是逆时针编织还是顺时针编织，因此必须考虑具体编织过程，像 (8.30) 这样的式子对于 $2+1$ 维的辫子群 B_N 是没有意义的。

根据 (8.30) 式的这个定义我们显然有 $P_{12}^2|\mathbf{x}_1,\mathbf{x}_2\rangle = P_{12}|\mathbf{x}_2,\mathbf{x}_1\rangle = |\mathbf{x}_1,\mathbf{x}_2\rangle$，即有算符方程

$$P_{12}^2 = 1. \tag{8.31}$$

这个结果相应于 $3+1$ 维中基本编织操作所满足的 $\sigma_i^2=1$。另一方面，根据 (8.30) 式我们也很容易看到

$$\begin{aligned}\langle \mathbf{x}_1',\mathbf{x}_2'|\mathbf{x}_1,\mathbf{x}_2\rangle &= \delta(\mathbf{x}_2'-\mathbf{x}_2)\delta(\mathbf{x}_1'-\mathbf{x}_1) = \langle \mathbf{x}_2',\mathbf{x}_1'|\mathbf{x}_2,\mathbf{x}_1\rangle \\ &= \langle \mathbf{x}_1',\mathbf{x}_2'|P_{12}^\dagger P_{12}|\mathbf{x}_1,\mathbf{x}_2\rangle.\end{aligned} \tag{8.32}$$

这就意味着 $P_{12}^\dagger P_{12}=1$，从而 (8.30) 式定义的置换算符 P_{12} 是一个幺正算符，结合 (8.31) 式我们就有

$$P_{12}^\dagger = P_{12} = P_{12}^{-1}. \tag{8.33}$$

(8.30) 式很容易推广到对 N 个粒子的任意置换。为此我们以 g 表示对 $\{1,2,...,N\}$ 的一个置换，记这个置换的作用方式为，$g: i \to g(i)$，即将 i 置换为 $g(i)$。则可以仿照 (8.30) 式定义对 N 个粒子的置换算符 P_g 如下 [11]

$$P_g|\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N\rangle = |\mathbf{x}_{g(1)},\mathbf{x}_{g(2)},...,\mathbf{x}_{g(N)}\rangle. \tag{8.34}$$

完全类似于两个粒子的情形，我们也可以证明 P_g 必定为幺正算符。

另一方面，正如我们前面提到过的，N 个粒子的任意置换必定可以分解成相继进行的一连串两两对换，因此对 N 个粒子的任意置换算符必定也可以写成一连串的两粒子对换算符的乘积。比方说，对于 3 个粒子情形，假设我们考虑如下置换

$$g: \{1,2,3\} \to \{2,3,1\}. \tag{8.35}$$

很容易看出来，这个置换操作和下面一串操作的效果是一样的

$$\{1,2,3\} \to \{2,1,3\} \to \{2,3,1\}. \tag{8.36}$$

按照群乘法的定义，相继进行的两个操作对应于两个群元相乘，根据我们前面的约定，先进行的操作在乘号的右边，后进行的操作在乘号的左边，因此 (8.36) 中的这一串操作就可以写成 $(13)(12)$。这里符号 (ij) 表示将 i,j 对换。也即是说，(8.35) 式的 $g=(13)(12)$，很明显这是一个偶置换。置换算符当然也要满足置换群乘法，因此相应的置换算符也必然可以写成 $P_g = P_{13}P_{12}$。由于前面证明过两个粒子的对换算符是幺正算符，从而这里的 P_g 也必定为幺正算符，这是另一种证明置换算符为幺正算符的方法。实际上，我们定义的置换算符构成了置换群的幺正表示。

将任意置换分解成一串两两对换相乘也有助于我们求出它的逆操作，为此我们只要注意到任何对换的逆操作是它本身。以上一段中的 $g=(13)(12)$ 为例，其逆元是 $g^{-1} = [(13)(12)]^{-1} = (12)^{-1}(13)^{-1} = (12)(13)$。对于相应的置换算符就有 $P_{g^{-1}} = P_{12}P_{13} = [P_{13}P_{12}]^{-1} = P_g^{-1} = P_g^\dagger$。很显然，偶置换的逆元将依然是一个偶置换，而奇置换的逆元也是一个奇置换。一般地，由于置换算符是置换群的幺正表示，所以我们有

$$P_{g^{-1}} = P_g^{-1} = P_g^\dagger. \tag{8.37}$$

假设置换算符 P_g 在任意量子态 $|\psi\rangle$ 上的作用结果为 $P_g|\psi\rangle = |\psi'\rangle$，记相应的坐标表象表达式为 $\widehat{P}_g\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) = \psi'(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$。则根据 (8.37) 式和 (8.34) 式，我们容易有 $\langle\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N|P_g = \langle\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N|P_{g^{-1}}^\dagger = \langle\mathbf{x}_{g^{-1}(1)},\mathbf{x}_{g^{-1}(2)},...,\mathbf{x}_{g^{-1}(N)}|$，从而可以得到置换算符在波函数上的作用结果为

$$\begin{aligned}\widehat{P}_g\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) &= \psi'(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) = \langle\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N|P_g|\psi\rangle \\ &= \langle\mathbf{x}_{g^{-1}(1)},\mathbf{x}_{g^{-1}(2)},...,\mathbf{x}_{g^{-1}(N)}|\psi\rangle = \psi(\mathbf{x}_{g^{-1}(1)},\mathbf{x}_{g^{-1}(2)},...,\mathbf{x}_{g^{-1}(N)}).\end{aligned} \tag{8.38}$$

[11] 非常细心的读者可能会注意到一个不重要的细节，即对粒子本身的置换和对局域变量 $\mathbf{x}$ 的置换其实是有区别的，但在我们的处理中故意混淆了这个细节性的区别，因为否则的话这里关于置换算符的定义以及后面关于置换算符在全同粒子波函数上如何作用的推理过程都会变得更复杂，这会不利于读者理解相关思路。更何况，这个细节其实并不会影响我们将要得到的关键性结论 (8.46) 式。

由于 g 置换将原来的 $\{\mathbf{x}_{g^{-1}(1)},\mathbf{x}_{g^{-1}(2)},...,\mathbf{x}_{g^{-1}(N)}\}$ 置换到了 $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N\}$，所以上面这个式子的含义其实就是，置换算符作用以后 $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N\}$ 位置的波函数其实就是原来 $\{\mathbf{x}_{g^{-1}(1)},\mathbf{x}_{g^{-1}(2)},...,\mathbf{x}_{g^{-1}(N)}\}$ 位置的波函数。记多粒子位形 $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N\}=\{\mathbf{x}\}$, 则有 $\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)=\psi(\{\mathbf{x}\})$。再示意性地记

$$\{\mathbf{x}_{g^{-1}(1)},\mathbf{x}_{g^{-1}(2)},...,\mathbf{x}_{g^{-1}(N)}\}=g^{-1}\{\mathbf{x}\}. \tag{8.39}$$

则可将 (8.38) 式重写成,

$$\widehat{P}_g\psi(\{\mathbf{x}\})=\psi(g^{-1}\{\mathbf{x}\}). \tag{8.40}$$

置换算符在全同粒子波函数上的作用

以上讨论的是置换算符在任意 N 粒子体系上的作用，这些粒子其实不必是全同粒子。对于全同粒子体系来说，情况会有些特殊，因为这时候体系的波函数并不是任意多变量函数，而是由全同粒子路径积分给出来的波函数。

具体来说，根据 (8.7) 式我们知道，任意 t 时刻的波函数 $\psi(\{\mathbf{x}\},t)$ 总是由某个初始 t_0 时刻的波函数按照下式演化而来

$$\psi(\{\mathbf{x}\},t)=\langle\{\mathbf{x}\}|\psi(t)\rangle=\sum_{\{\mathbf{x}\}_0}\langle\{\mathbf{x}\},t|\{\mathbf{x}\}_0,t_0\rangle\psi(\{\mathbf{x}\}_0,t_0). \tag{8.41}$$

代入全同粒子的路径积分公式 (8.19) 即有,

$$\psi(\{\mathbf{x}\},t)=\frac{1}{|G|}\sum_{g_1}\left[U(g_1)\sum_{\{\mathbf{x}\}_0}\sum_{\{\mathbf{x}\}_0\xrightarrow{g_1}\{\mathbf{x}\}}e^{iS[\{\mathbf{x}(t)\}\in g_1]/\hbar}\psi(\{\mathbf{x}\}_0,t_0)\right]. \tag{8.42}$$

式中 $\sum_{\{\mathbf{x}\}_0\xrightarrow{g_1}\{\mathbf{x}\}}$ 表示对从 $\{\{\mathbf{x}\}_0,t_0\}$ 到 $\{\{\mathbf{x}\},t\}$ 的所有属于拓扑类 g_1 的路径求和。另外，由于我们只需考虑置换群的 1 维表示，所以式中的 $U(g_1)$ 实际上是一个相位因子，即 $U(g_1)=e^{-i\varphi(g_1)}$。

下面我们考虑置换算符 P_g 在波函数上的作用，根据 (8.40) 式我们有, $\widehat{P}_g\psi(\{\mathbf{x}\},t)=\psi(g^{-1}\{\mathbf{x}\},t)$, 对等号右边应用全同粒子波函数的路径积分公式 (8.42) 则有

$$\widehat{P}_g\psi(\{\mathbf{x}\},t)=\frac{1}{|G|}\sum_{g_1}\left[U(g_1)\sum_{\{\mathbf{x}\}_0}\sum_{\{\mathbf{x}\}_0\xrightarrow{g_1}g^{-1}\{\mathbf{x}\}}e^{iS[\{\mathbf{x}(t)\}\in g_1]/\hbar}\psi(\{\mathbf{x}\}_0,t_0)\right]. \tag{8.43}$$

然而很显然，路径 $\{\mathbf{x}\}_0\xrightarrow{g_1}g^{-1}\{\mathbf{x}\}$ 和路径 $\{\mathbf{x}\}_0\xrightarrow{g^{-1}g_1}\{\mathbf{x}\}$ 是一回事，所以上面的表达式又可以写成

$$\widehat{P}_g\psi(\{\mathbf{x}\},t)=\frac{1}{|G|}\sum_{g_1}\left[U(g_1)\sum_{\{\mathbf{x}\}_0}\sum_{\{\mathbf{x}\}_0\xrightarrow{g^{-1}g_1}\{\mathbf{x}\}}e^{iS[\{\mathbf{x}(t)\}\in g^{-1}g_1]/\hbar}\psi(\{\mathbf{x}\}_0,t_0)\right]. \tag{8.44}$$

令 $g^{-1}g_1=g_2$，利用 $g_1=gg_2$ 将上式中的 g_1 代换成 g_2，进而将对 g_1 的求和代换成对 g_2 的求和，从而即有

$$\begin{aligned}\widehat{P}_g\psi(\{\mathbf{x}\},t)&=\frac{1}{|G|}\sum_{g_2}\Big[U(gg_2)\sum_{\{\mathbf{x}\}_0}\sum_{\{\mathbf{x}\}_0\xrightarrow{g_2}\{\mathbf{x}\}}e^{iS[\{\mathbf{x}(t)\}\in g_2]/\hbar}\psi(\{\mathbf{x}\}_0,t_0)\Big]\\&=U(g)\frac{1}{|G|}\sum_{g_2}\Big[U(g_2)\sum_{\{\mathbf{x}\}_0}\sum_{\{\mathbf{x}\}_0\xrightarrow{g_2}\{\mathbf{x}\}}e^{iS[\{\mathbf{x}(t)\}\in g_2]/\hbar}\psi(\{\mathbf{x}\}_0,t_0)\Big]\\&=U(g)\psi(\{\mathbf{x}\},t).\end{aligned}\tag{8.45}$$

式中第 2 个等号我们利用了群表示关系 $U(gg_2)=U(g)U(g_2)$，最后一个等号则是再次代入了波函数的路径积分公式 (8.42)。

小结一下，上面的推导结果告诉我们，在置换算符的作用下，$3+1$ 维中的全同粒子波函数必定满足

$$\widehat{P}_g\psi(\{\mathbf{x}\},t)=\psi(g^{-1}\{\mathbf{x}\},t)=U(g)\psi(\{\mathbf{x}\},t).\tag{8.46}$$

注意到这里的 $U(g)=e^{-i\varphi(g)}$ 是置换群的一维表示，因此这个结果告诉我们，$3+1$ 维全同粒子波函数仅有两种不同的可能性，对于全同玻色子，其波函数在任意置换下都将保持不变，我们称之为对称波函数。而对于全同费米子，其波函数在任何偶置换下将保持不变，但是在奇置换下都将出一个负号，特别的，对于任意两个粒子的对换，全同费米子波函数将相差一个负号，这样的波函数称之为反对称波函数。值得再次强调的是，这里诸如 $\mathbf{x}_i$ 这样的位置坐标只是一个示意性的记号，它同时指代第 i 个粒子的坐标和自旋量子数等一切局域信息指标，因此我们的置换操作不仅仅是作用在坐标变量上，而是同时也作用在自旋量子数等变量上。当我们说全同粒子波函数是对称波函数或者反对称波函数时，我们指的不是波函数对于坐标变量的置换对称或者反对称，而是指它对所有局域变量的同时置换对称或反对称。

(8.46) 式是一个坐标表象下的表达式，我们当然也可以写出与之相应的抽象希尔伯特空间表达式。利用 $\widehat{P}_g\psi(\{\mathbf{x}\})=\langle\{\mathbf{x}\}|P_g|\psi\rangle$，很容易看出这个抽象表达式为

$$P_g|\psi\rangle=U(g)|\psi\rangle.\tag{8.47}$$

这是 $3+1$ 维全同粒子体系的任何量子态 $|\psi\rangle$ 都必须满足的表达式。对于一个 N 粒子体系而言，并非所有数学上可能的量子态都满足 (8.47)，而只有满足 (8.47) 式的量子态才可能描写全同粒子体系的物理状态。假设记一个 N 粒子体系所有数学上可能的量子态集合为希尔伯特空间 $\mathscr{H}$，则很容易看出来，满足 (8.47) 式能描述全同粒子体系的量子态集合为 $\mathscr{H}$ 的一个子空间。对于全同玻色子体系而言，$U(g)=e^{-i\varphi(g)}=1$，因此这个子空间中的任何量子态在置换算符 P_g 的作用下都得保持不变，我们称这样的子空间为对称态子空间，记为 $\mathscr{H}_S$。而对于全同费米子体系而言，通常记相应的相位因子 $U(g)=e^{-i\varphi(g)}=\varepsilon(g)$，对于偶置换 $\varepsilon(g)=1$，对于奇置换 $\varepsilon(g)=-1$，因此相应希尔伯特子空间中的任何量子态 $|\psi\rangle$ 都得满足 $P_g|\psi\rangle=\varepsilon(g)|\psi\rangle$，这样的希尔伯特子空间称作反对称态子空间，记为 $\mathscr{H}_A$。

比方说，对于一个两粒子体系，$|\mathbf{x}_1,\mathbf{x}_2\rangle$ 态显然属于 $\mathscr{H}$，但它并不满足 (8.47) 式，从而既不属于 $\mathscr{H}_S$，也不属于 $\mathscr{H}_A$，因此它不能描写两个全同粒子的系统。但是，很容易验证 $|\mathbf{x}_1,\mathbf{x}_2\rangle+|\mathbf{x}_2,\mathbf{x}_1\rangle$ 态在两粒子置换算符的作用下是不变的，因此它属于 $\mathscr{H}_S$，能够描写两全同玻色子的某个物理状态。类似的也容易验证 $|\mathbf{x}_1,\mathbf{x}_2\rangle-|\mathbf{x}_2,\mathbf{x}_1\rangle$ 属于 $\mathscr{H}_A$，可以描写两全同费米子系统的某个物理状态。一个重要的问题是，给定一个 N 粒子系统的希尔伯特空间 $\mathscr{H}$，如何构建它的对称态子空间 $\mathscr{H}_S$ 以及反对称态子空间 $\mathscr{H}_A$，我们将在下一节中具体讨论这个问题。

置换对称性

置换其实是 $3+1$ 维全同粒子系统的一种对称性，而且是一种规范对称性。为了看清楚这一点，我们先来考察 $\langle\{\mathbf{x}\}_f|P_g^{-1}U(t_f,t_i)P_g|\{\mathbf{x}\}_i\rangle$，根据 (8.34) 式，我们显然有

$$\begin{aligned}&\langle\{\mathbf{x}\}_f|P_g^{-1}U(t_f,t_i)P_g|\{\mathbf{x}\}_i\rangle=\langle\{\mathbf{x}\}_f|P_g^{\dagger}U(t_f,t_i)P_g|\{\mathbf{x}\}_i\rangle\\=&\langle g\{\mathbf{x}\}_f|U(t_f,t_i)|g\{\mathbf{x}\}_i\rangle=\langle g\{\mathbf{x}\}_f,t_f|g\{\mathbf{x}\}_i,t_i\rangle.\end{aligned}\tag{8.48}$$

式中 $g\{\mathbf{x}\}=\{\mathbf{x}_{g(1)},\mathbf{x}_{g(2)},...,\mathbf{x}_{g(N)}\}$。很明显，$\langle g\{\mathbf{x}\}_f,t_f|g\{\mathbf{x}\}_i,t_i\rangle$ 其实就是同步地将初末态的 N 个坐标进行了一个 $1,2,...,N\to g(1),g(2),...,g(N)$ 的置换。仔细想一下人们就能明白，这样的初末态同步置换根本不会影响路径的拓扑分类，也不会影响路径的作用量 (或者说，由于全同粒子不可区分，作用量在粒子置换下将不变)，因此，根据路径积分公式 (8.19)，我们将有 $\langle g\{\mathbf{x}\}_f,t_f|g\{\mathbf{x}\}_i,t_i\rangle=\langle\{\mathbf{x}\}_f,t_f|\{\mathbf{x}\}_i,t_i\rangle=\langle\{\mathbf{x}\}_f|U(t_f,t_i)|\{\mathbf{x}\}_i\rangle$。代入 (8.48) 式即有

$$\begin{aligned}&\langle\{\mathbf{x}\}_f|P_g^{-1}U(t_f,t_i)P_g|\{\mathbf{x}\}_i\rangle=\langle\{\mathbf{x}\}_f|U(t_f,t_i)|\{\mathbf{x}\}_i\rangle\\\Leftrightarrow\ &P_g^{-1}U(t_f,t_i)P_g=U(t_f,t_i).\end{aligned}\tag{8.49}$$

即置换算符与时间演化算符对易。按照第六章中关于对称性的定义，这就说明，$3+1$ 维全同粒子体系具有置换对称性，置换算符 P_g 就是相应的幺正对称变换。特别的，由于 $U(t_f,t_i)=\exp(-iH(t_f-t_i)/\hbar)$，所以由 (8.49) 式我们必然还有 $[P_g,H]=0$。

类似的，对于海森堡绘景中的任何物理可观测量算符 $\mathscr{O}(t)$，我们可以考察 $\langle\{\mathbf{x}\}_f,t_f|P_g^{-1}\mathscr{O}(t)P_g|\{\mathbf{x}\}_i,t_i\rangle=\langle g\{\mathbf{x}\}_f,t_f|\mathscr{O}(t)|g\{\mathbf{x}\}_i,t_i\rangle$。这个结果同样可以用路径积分公式来计算，办法就是在 $\langle g\{\mathbf{x}\}_f,t_f|g\{\mathbf{x}\}_i,t_i\rangle$ 的路径积分计算中插入一个表达式 $\mathscr{O}(t)$。很明显，对于全同粒子体系，将初末态坐标进行同步置换对于这个中间时刻 t 的表达式 $\mathscr{O}(t)$ 不会有任何影响，因此和上一段的论证完全一样，我们将有 $\langle g\{\mathbf{x}\}_f,t_f|\mathscr{O}(t)|g\{\mathbf{x}\}_i,t_i\rangle=\langle\{\mathbf{x}\}_f,t_f|\mathscr{O}(t)|\{\mathbf{x}\}_i,t_i\rangle$，进一步就能得到

$$P_g^{-1}\mathscr{O}(t)P_g=\mathscr{O}(t).\tag{8.50}$$

即置换算符和所有物理可观测量均对易。换言之，全同粒子体系的任何物理可观测量在粒子置换下都得保持不变。这意味着，对于全同粒子体系，有一些很简单的算符其实是不可观测的，比方说位置算符 $\mathbf{X}_1$ 就不是可观测量，因为在位置置换下它无法不变，但是算符 $(\mathbf{X}_1+\mathbf{X}_2+...+\mathbf{X}_N)/N$ 却是可观测的，因为它在粒子置换下显然不变。这其实很好理解，由于全同粒子的不可区辨性，人们测量粒子坐标的时候实际上无法知道测量结果来自于 N

个粒子中的哪一个，因此结果是只能测到 $(\mathbf{X}_1+\mathbf{X}_2+...+\mathbf{X}_N)/N$ 这类算符的本征值，而根本无法测量到 $\mathbf{X}_1$ 的本征值。

由于置换算符与全同粒子体系的所有物理可观测量均对易，因此我们就可以把这些可观测量算符限制在对称态子空间 $\mathscr{H}_S$ 或者限制在反对称态子空间 $\mathscr{H}_A$ 上，也即是说，任意可观测量 $\mathscr{O}$ 作用在任何对称态上结果将依然是对称态，同样，可观测量 $\mathscr{O}$ 作用在任何反对称态上结果必然依然是反对称态。证明非常简单，不妨以反对称态的情况为例，假设 $|\psi\rangle\in\mathscr{H}_A$，即 $|\psi\rangle$ 满足 $P_g|\psi\rangle=\varepsilon(g)|\psi\rangle$，设 $\mathscr{O}$ 在 $|\psi\rangle$ 上的作用结果为 $|\psi'\rangle$，即 $|\psi'\rangle=\mathscr{O}|\psi\rangle$，则由于 (8.50)，我们有 $P_g|\psi'\rangle=P_g\mathscr{O}|\psi\rangle=\mathscr{O}P_g|\psi\rangle=\varepsilon(g)\mathscr{O}|\psi\rangle=\varepsilon(g)|\psi'\rangle$，可见 $|\psi'\rangle$ 依然是反对称态。我们说过，对称态描述的是全同玻色子，反对称态描述的是全同费米子，因此刚才的结果意味着，是玻色子还是费米子这是微观粒子的本质属性，任何物理过程和物理测量都无法改变这一属性。正是根据这种本质属性我们才可以将三维空间的微观粒子分成玻色子和费米子两大类。

(8.50) 也意味着，置换操作不仅是 $3+1$ 维全同粒子体系的对称性，而且是一种非常特殊的对称性。对于普通的对称变换来说，它只是不改变系统的动力学演化规律，也即是与时间演化算符对易，但与其它的物理量并不一定对易，因此一般来说，我们可以通过观测其它物理量的变化而观测到对称变换对系统的影响。但是，如果一个对称变换与所有的物理量都对易，那它的作用结果就是不可观测的，这样的对称性我们通常称之为规范对称性。这就是为什么我们说 $3+1$ 维全同粒子系统的置换对称性是一种规范对称性的原因。前面的章节中我们还碰到过另一个规范对称性的例子，即电荷守恒的 $U(1)$ 相位变换，而且我们也知道，电荷 $U(1)$ 相位变换同样也与所有的物理可观测量都对易。

8.4 * 自旋统计定理

这一节我们依然统一讨论 $2+1$ 维情形和 $3+1$ 维情形。我们的目标是建立两个全同粒子之间的基本编织操作和空间旋转操作之间的关系。一般来说，这两者之间并没有什么明确的关系，但是，在一种特殊的情形中，全同粒子编织和空间旋转的确密切相关，这种特殊情形允许我们将 $2+1$ 维全同阿贝尔任意子的编织与空间旋转等同起来，也允许我们建立 $3+1$ 维玻色-费米统计与粒子自旋之间的关系，所以叫自旋统计定理。

这种特殊情形其实非常简单，考虑一个荷为 a 的粒子在时空中的演化路径，如图 (8.12) 所示，图中左边的 a 粒子在某个时空局域上受到一个复杂的作用，使它在这个局域上“打了个圈”，虽然如此，局域性原理告诉我们，左边的态本质上依然是 a 粒子态，因此只能和右边的标准 a 粒子态相差一个相位 $e^{-i\theta_a}$。

另一方面，由于 a 粒子有内部自由度，所以我们不妨将它的时空路径画成一条带子，如图 (8.13) 所示。在这幅图中，我们把前面 a 粒子“打了个圈”的演化路径画成了一条时空带子，拉直这条带子就得到图 (8.13) 右边的那条带子，但是右边的这条带子显然是 a 粒子逆时针自转 2π 角的结果。这也即是说，a 粒子“打了个圈”所多出来的相因子应该刚好等于逆时针自转 2π 角的 $\exp(-2\pi iJ_a/\hbar)$，J_a 是 a 粒子自旋角动量。结合这两段的结果

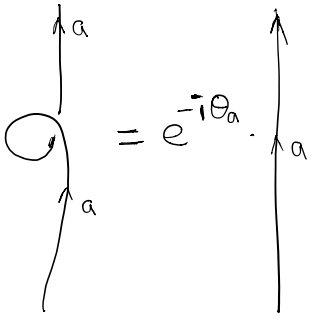

Figure 8.12: 左边的 a 粒子在某个时空局域上受到一个复杂的作用，使它在这个局域上“打了个圈”，虽然如此，局域性原理告诉我们，左边的态本质上依然是 a 粒子态，因此只能和右边的标准 a 粒子态相差一个相位 $e^{-i\theta_a}$。

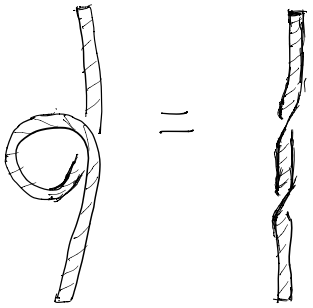

Figure 8.13: 左边的 a 粒子时空带子拉直后就变成右边的带子，但是右边的带子显然等于 a 粒子逆时针自转 2π 角。

我们就得到

$$e^{-i\theta_a} = \exp(-2\pi i J_a/\hbar). \tag{8.51}$$

现在只差一步我们就可以得到自旋-统计定理了，这一步涉及到对 $e^{-i\theta_a}$ 的解释。为此我们又得注意到图 (8.12) 和图 (8.13) 中 a 粒子“打了个圈”的演化路径其实非常奇特，奇特之处就在于，我们知道，竖直向上方向是时间轴，但是你看图 (8.12) 和图 (8.13) 中 a 粒子“打了个圈”的路径，它有一部分是逆着时间走的。对于微观粒子来说，这好像也没有什么问题，因为对于微观世界而言，时间正方向和负方向地位完全平等。但是，对于我们宏观的观察者而言，时间只能沿着正方向流逝，为此我们就得重新解释图 (8.12) 中 a 粒子演化路径的逆时间部分。注意到电荷的运动方向与反电荷的反向运动方向对于产生电流来说完全一样，因此我们可以将逆时间的那一部分 a 路径解释成 a 的反粒子 $\bar{a}$ 在顺着时间方向走。如此一来，我们就可以将图 (8.12) 中 a 粒子“打了个圈”的路径重新解释成 (如图 (8.14) 所示)：C 地有一个 a 粒子随着时间演化，突然 A 地产生了一对正反 a 粒子，其中正 a 粒子与 C 地运动过来的 a 粒子作了个逆时针编织，之后，A 地产生的粒子对中的反 a 粒子与 C 地过来的 a 粒子在 B 地湮灭，剩下一个 A 地粒子对中的正 a 粒子飞向未来。这样一分析我们就可以知道，前面的相因子 $e^{-i\theta_a}$ 其实正好来源于两个 a 粒子的逆时



Figure 8.14: a 粒子对产生，对湮灭，以及编织。图中的箭头方向可以重新解释为荷的流动方向。

针编织。如此一来，方程 (8.51) 就将一类特殊的全同粒子编织与粒子自旋联系了起来，因此可以称作自旋统计定理。

特别的，如果 a 是 $2+1$ 维阿贝尔任意子，那刚才的相因子 $e^{-i\theta_a}$ 必定是基本编织操作的阿贝尔表示，方程 (8.51) 就把这个阿贝尔表示与 2 维空间的 2π 角旋转联系了起来。值得注意的是，一般来说，2 维空间的 2π 角旋转并不一定等于 1，因此阿贝尔任意子可以有非平凡的 $e^{-i\theta_a}$。但是一般来说，我们要求，2 维空间虽然旋转一圈不等于恒等操作，但是旋转 n 圈总可以等于恒等操作，即要求对于某个 n，有 $\exp(-n2\pi iJ_a/\hbar)=1$, 或者利用自旋-统计定理即有，

$$e^{-in\theta_a}=1. \tag{8.52}$$

从而可见，阿贝尔任意子的统计角 θ_a 其实是 2π 乘以一个分数，所以又称之为分数统计。

进一步，如果 a 是一个 $3+1$ 维的粒子，那这时候空间旋转群就不是 2 维空间的 $SO(2)$ 群了，而是 3 维空间的 $SO(3)$ 群。在前面的章节中我们系统研究过这种旋转群的表示，我们发现，它有两种不同的表示，半整数表示和整数表示，对于整数表示 $\exp(-2\pi iJ_a/\hbar)=U(2\pi)=1$，对于半整数表示，$\exp(-2\pi iJ_a/\hbar)=U(2\pi)=-1$。根据自旋-统计定理 (8.51)，这就意味着，对于整数自旋粒子，必有 $e^{-i\theta_a}=1$，从而是玻色子，而对于半整数自旋粒子，则必有 $e^{-i\theta_a}=-1$，从而是费米子。这就是通常量子场论中证明的自旋-统计定理。通常人们将这个定理概括成如下等式

$$U(2\pi)=(-)^{\mathrm{F}}. \tag{8.53}$$

式中 $(-)^{\mathrm{F}}$ 称为费米宇称算符，它作用在费米量子态上为 -1，作用在玻色态上为 $+1$。

读者不妨回想一下前面章节中关于超对称的讨论，我们知道 $U(2\pi)$ 与超对称荷 Q 反对易，因此根据上面的自旋统计定理，就有

$$(-)^{\mathrm{F}}Q(-)^{\mathrm{F}}=-Q. \tag{8.54}$$

这个结果就说明，超对称荷 Q 是一个费米算符。

8.5 玻色子和费米子

本章前面几节我们利用全同粒子路径积分系统地讨论了 $2+1$ 维全同粒子体系和 $3+1$ 维全同粒子体系。这些讨论除了本身能够导出一些神奇而重要的结论之外，它对于深入理解全同性原理也非常有帮助，然而，对于初学者来说，这些讨论可能有些过深。这一节我们将回到量子力学初学者的角度，直接讨论 $3+1$ 维时空中的全同性原理，讨论玻色子统计和费米子统计，对于已经学过并理解了前几节内容的读者，本节的内容基本上可以看成前几节的自然发展。不过，为了照顾到跳过了前几节的初学者，本节也会用简化的处理重复前几节的少量必要内容。

从某个角度上说，$3+1$ 维时空的全同粒子体系当然更为基本，因为我们的真实时空就是 $3+1$ 维的。即使对所谓的两维材料，它们其实也都有厚度，从微观上追究起来它们也都是由 3 维空间中的原子-分子-电子等等组成，其两维特性其实是一种“有效”描述，仅当我们不使用足够高的能量去探测它的微观结构时才成立。正因为如此，这一节和后面的第 (8.6) 节，我们将专门地研究这种真实的 3 维空间中的全同粒子体系。

8.5.1 多体希尔伯特空间

全同粒子量子力学当然属于多体量子力学，本书前面部分也涉及过多体量子力学，比如量子纠缠就不是单体量子现象，它至少要涉及到两个粒子或者说两体。但是我们还没有专门地强调过多体量子力学的一些特性，现在，让我们暂时忘记全同粒子，先看一下任何一个多体体系的希尔伯特空间有何特性。

让我们从两体开始，假设一个量子系统由 A, B 两体组成，A, B 各自都可以有很复杂的内部结构，比方说各自本身可以是一个多体系统。假设单独对于 A 来说，其量子态的希尔伯特空间为 $\mathscr{H}_A$，单独对于 B 来说，希尔伯特空间为 $\mathscr{H}_B$，那么 A, B 作为一个整体其希尔伯特空间 $\mathscr{H}_{AB}$ 是怎么样的呢？虽然我们前面没有一般性地讨论过这个问题，但是我们已经见过很多两体的例子，两个自由粒子，两个量子比特等等。仔细回想一下我们怎么处理这些例子，就能明白下面的回答是非常自然的。首先，我们分别取 $\mathscr{H}_A$ 和 $\mathscr{H}_B$ 各自的一组正交归一基 $\{|i_A\rangle\}, \{|i_B\rangle\}$，那么整个系统的任何量子态 $|\psi\rangle$ 都必定能写成如下的叠加形式，

$$|\psi\rangle = \sum_{i_A,i_B} \psi_{i_A,i_B} |i_A, i_B\rangle = \sum_{i_A,i_B} \psi_{i_A,i_B} |i_A\rangle |i_B\rangle. \tag{8.55}$$

式中 ψ_{i_A,i_B} 是叠加系数，所有这类量子态的集合就构成了整个系统的希尔伯特空间 $\mathscr{H}_{AB}$。数学上，对于这样构造出来的希尔伯特空间有一个专门的名称，称之为 $\mathscr{H}_A$ 和 $\mathscr{H}_B$ 的张量积，记为 $\mathscr{H}_A \otimes \mathscr{H}_B$，因此

$$\mathscr{H}_{AB} = \mathscr{H}_A \otimes \mathscr{H}_B. \tag{8.56}$$

值得说明的是，张量积是一个重要的数学概念，它的基本含义就是将两个东西并置在一起，在数学上它并不是必须按照 (8.55) 式来定义，实际上，在所谓的张量范畴论中，(8.55) 式定义的张量积是一种“平凡”的张量积。但是在线性代数中以及我们的这本量子力学书中，我们只需关心这种“平凡”的张量积就足够了。之所以出现这种“平凡”的张

量积，是因为我们的 A、B 两个子系统有某种局域性，也即是说，仅仅对 A 的某个物理量进行测量，和仅仅对 B 的某个物理量进行测量，两者之间不会有任何相互影响。

张量积的概念当然可以推广到一般的多体情形。假设 $A,B,C,D....$ 等多个子系统构成一个大量子系统，则这个大量子系统的希尔伯特空间就是各个子系统希尔伯特空间的张量积，可以记为 $\mathscr{H}_A \otimes \mathscr{H}_B \otimes \mathscr{H}_C \otimes ....$，它是所有形如下式的量子态的集合

$$\begin{aligned} |\psi\rangle &= \sum_{i_A,i_B,i_C,\ldots} \psi_{i_A,i_B,i_C\ldots} |i_A, i_B, i_C \ldots\rangle \\ &= \sum_{i_A,i_B,i_C,\ldots} \psi_{i_A,i_B,i_C\ldots} |i_A\rangle |i_B\rangle |i_C\rangle \ldots \end{aligned} \tag{8.57}$$

对于 N 个全同粒子的体系，由于这些粒子不可区分，所以描写单个粒子的正交归一基矢量对于所有粒子都应该一样，不妨记这样的正交归一基矢量为 $|u_i\rangle$，指标 $i = 1,2,3,...$ 用来区分不同的基矢量。为了描写整个多粒子态，我们需要将这些全同粒子进行标记，设分别标记为 $1,2,3...,N$。当然，这样的标记是任意的，而且由于这些粒子严格不可区分，所以这些标记实际上也不是物理的，它仅仅是为了构造全同粒子量子态而引入的数学描述。作了这样的标记以后我们就可以将第 n 个粒子的正交归一基矢量记为 $|u_i(n)\rangle$。根据我们刚才的描述，整个 N 粒子体系的希尔伯特空间 $\mathscr{H}$ 将是各粒子希尔伯特空间的张量积，它的基矢量可以取为 $|u_i(1), u_j(2), ..., u_k(N)\rangle$，但是，正如我们将要看到的，这样的基矢量虽然可以直接描写非全同的多粒子体系，但却不能直接描写全同多粒子体系。实际上全同 N 粒子体系的希尔伯特空间将是 $\mathscr{H}$ 的某个特定子空间，下面我们将要讨论如何构造这样的子空间。

8.5.2 置换对称性

两全同粒子情形

我们先来讨论两个全同粒子的情形。将这两个粒子标记为 $1,2$，记相应的两粒子态希尔伯特空间为 $\mathscr{H}$，根据前面的讨论，$\mathscr{H}$ 的矢量基为 $|u_i(1), u_j(2)\rangle$。由于我们讨论的是两个全同粒子，我们现在非物理地将这两个粒子分别标记为 $1,2$，但实际在物理上，$1,2$ 是不可区分的，这就意味着，如果我们将这两个粒子相互置换一下，在物理上系统将不会发生任何改变。这有两层含义，第一，置换前后系统的物理状态不会改变。第二，这样的置换操作不会引起任何物理可观测量的改变。这两点实际上就意味着，这样的置换操作是这两个全同粒子体系的一种特殊对称性。

我们定义两粒子的置换算符 P_{12} 如下，

$$P_{12}|u_i(1), u_j(2)\rangle = |u_i(2), u_j(1)\rangle = |u_j(1), u_i(2)\rangle, \tag{8.58}$$

上式的最后一个等号来源于习惯上我们常常将标记为第 1 个粒子的量子态写在狄拉克符号的第 1 个位置。读者可以验证，这样定义的置换算符是一个幺正算符，即它保持态的内积不变。且对于希尔伯特空间 $\mathscr{H}$ 里的任意量子态 $|\psi\rangle = \sum_{i,j} \psi_{i,j} |u_i(1), u_j(2)\rangle$，我们有

$$P_{12}|\psi\rangle = \sum_{i,j} \psi_{i,j} |u_j(1), u_i(2)\rangle = \sum_{i,j} \psi_{j,i} |u_i(1), u_j(2)\rangle \tag{8.59}$$

上式的最后一个等号来自于求和指标的重命名。很显然，置换算符对两粒子量子态的作用相当于将叠加系数 $\psi_{i,j}$ 的两个下标互换，即 $\psi_{i,j}\to\psi_{j,i}$。很明显，连续进行两次置换相当于不进行任何操作，即有 $P_{12}^2=1$。

现在，如果这两个粒子是两个全同粒子，则由于置换操作不会改变系统的物理状态，所以必有

$$P_{12}|\psi\rangle=e^{-i\varphi}|\psi\rangle, \tag{8.60}$$

即置换算符作用之后的量子态只能和原来的量子态相差一个相位。又由于 $P_{12}^2=1$，从而有 $(e^{-i\varphi})^2=1$, 所以这个相位只有两种可能性，要么取 $+1$，要么取 -1。对于 $+1$ 的情形，我们就说这两个全同粒子为全同玻色子，对于 -1 的情形，我们就说这两个粒子为全同费米子。另外，根据上一段的讨论我们又知道，用量子态的展开系数来说，P_{12} 的作用相当于 $P_{12}:\psi_{i,j}\to\psi_{j,i}$。因此对于全同玻色子，由于量子态在置换算符的作用下要保持不变，从而我们必有 $\psi_{i,j}=\psi_{j,i}$，而对于全同费米子则必有，$\psi_{j,i}=-\psi_{i,j}$。人们常常称这个结果为：全同玻色子的波函数必定为对称波函数，全同费米子的波函数必定为反对称波函数。

两粒子希尔伯特空间 $\mathscr{H}$ 中，所有对称波函数所构成的希尔伯特子空间为全同玻色子的对称态空间，记为 $\mathscr{H}_S$，而所有反对称波函数所构成的希尔伯特子空间为全同费米子的反对称态空间，记为 $\mathscr{H}_A$。读者容易验证，$\mathscr{H}_S$ 的基矢量可以取为

$$|u_i,u_j\rangle_S=|u_i(1),u_j(2)\rangle+|u_i(2),u_j(1)\rangle=(1+P_{12})|u_i(1),u_j(2)\rangle. \tag{8.61}$$

这是因为，$|u_i,u_j\rangle_S$ 的任意线性叠加态 $\sum_{i,j}c_{i,j}|u_i,u_j\rangle_S=\sum_{i,j}c_{i,j}|u_i(1),u_j(2)\rangle+\sum_{i,j}c_{i,j}|u_i(2),u_j(1)\rangle=\sum_{i,j}(c_{i,j}+c_{j,i})|u_i(1),u_j(2)\rangle$，很明显 $(c_{i,j}+c_{j,i})=\psi_{i,j}$ 是一个对称波函数，因此在原来的希尔伯特空间 $\mathscr{H}$ 中来看，这样的叠加态当然都是对称态。类似的，两全同费米子反对称态空间 $\mathscr{H}_A$ 的基矢量可以取为

$$|u_i,u_j\rangle_A=(1-P_{12})|u_i(1),u_j(2)\rangle. \tag{8.62}$$

关于两全同费米子的反对称态，有一个特别重要的情形值得单独考察。这情形就是，假如我们在公式 (8.62) 中令 $i=j$，很显然，结果必然是 $|u_i,u_i\rangle_A=0$。也即是说，对于全同费米子系统，两个粒子不能处在同一个单粒子态上。这其实就是著名的泡利不相容原理。在前面的章节中我们讲过，正是因为有泡利不相容原理，我们才能正确地解释元素周期表。不仅如此，泡利不相容原理也是维持我们这个物质世界稳定性的根本原理之一，假如没有泡利不相容原理，那原子的核外电子都会占据到最低单电子能级上去，如此一来所有原子都会坍缩。两个费米子不能占据同一个单粒子态，这在效果上，就好像两个费米子之间有一个排斥力一样，对于大量全同费米子的系统，费米子之间的这种排斥力就称之为费米子的简并压。比方说，电子简并压正是使得白矮星能够抵抗自身的引力作用而稳定存在的原因，钱德拉塞卡甚至具体计算出，在恒星质量未超过 1.44 个太阳质量 (钱德拉塞卡极限) 时，电子简并压都能够阻止恒星在自身引力作用下的坍缩。

什么粒子是玻色子，什么粒子是费米子呢？关于这一点我们有所谓的自旋-统计定理。它告诉我们，整数自旋的粒子都是玻色子，而半整数自旋的粒子都是费米子。比方说光子

的自旋是 1，因此是玻色子，而电子的自旋是 1/2，因此是费米子。同样，质子中子也是费米子，而所谓的“上帝粒子” Higgs 玻色子，是一个标量粒子，自旋是零，因此是玻色子。

N 个全同粒子的情形

以上关于两全同粒子置换对称性的讨论当然可以推广到任意 N 个全同粒子情形，我们任意地将这 N 个全同粒子标记为 $1,2,3,...,N$。则对这 N 个粒子的置换就表现为对这 N 个数的置换，我们记在置换操作 g 的作用下，粒子 n 将被置换成粒子 $g(n)$, 则 g 本身就唯一由其在 N 个粒子上的作用 $g:\{1,2,3,....,N\}\to\{g(1),g(2),...,g(N)\}$ 决定。最简单的置换操作当然就是恒等置换，也就是不做任何置换，可以把相应的 g 简单记为 1。另外我们可以定义两个置换操作 g_1 和 g_2 的乘积 g_2g_1 为，先对 N 个粒子进行置换操作 g_1，之后再在 g_1 置换结果的基础上进行置换操作 g_2，很显然，按照这个置换乘法的定义，g_2g_1 仍然是一个置换操作。由此可见，N 个全同粒子的所有可能置换 (包括恒等置换) 的集合 S_N 在置换乘法下是封闭的。不仅如此，我们可以将任何置换过程逆转过来，从而得到任意置换操作 g 的逆操作 g^{-1}，它满足 $g^{-1}g=gg^{-1}=1$。也即是说，S_N 的每一个元素都可逆。满足这些性质的集合 S_N 就称之为一个群，元素 g 就是称作一个群元，S_N 就是 N 个粒子的置换群，由于 N 个粒子有 $N!$ 种置换，因此 S_N 共有 $N!$ 个群元。

置换操作中最简单的就是两粒子对换，通常记 i,j 两粒子的对换操作为 (ij)，很明显任何对换都满足 $(ij)^2=1$。数学上不难证明，N 粒子的任意置换群元 g 均可以写成一系列的两两对换的乘积。如果 g 对应于偶数个对换相乘，我们就称 g 为偶置换，如果 g 相应于奇数个对换相乘，我们就称之为奇置换。很容易证明偶置换的逆依然是偶置换，奇置换的逆依然是奇置换，比方说偶置换 $g=(13)(12)$ 的逆 $g^{-1}=(12)(13)$ 显然依旧是偶置换。

对于 N 个全同粒子的系统，我们可以首先忽略其全同性，得到 N 个粒子的希尔伯特空间 $\mathscr{H}$，其基矢量可以选为 $|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle$, 式中 $|u_i\rangle$ 是单粒子希尔伯特空间的正交归一基矢量。仿照两粒子情形，我们可以将对 N 个粒子的置换算符 P_g 定义成

$$\begin{aligned}&P_g|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle=|u_{i_1}(g(1)),u_{i_2}(g(2)),...,u_{i_N}(g(N))\rangle\\=&|u_{i_{g^{-1}(1)}}(1),u_{i_{g^{-1}(2)}}(2),...,u_{i_{g^{-1}(N)}}(N)\rangle.\end{aligned}\tag{8.63}$$

上式的最后一个等号是将多粒子态重新排列一下，使得第 1 个粒子的态排在狄拉克符号的第 1 个位置，以此类推。读者可以验证，这样定义出来的置换算符必定是一个幺正算符，即必定保持态的内积不变。另外，根据上面的定义，读者很容易验证

$$\begin{aligned}&P_{g_2}P_{g_1}|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle\\=&|u_{i_1}(g_2g_1(1)),u_{i_2}(g_2g_1(2)),...,u_{i_N}(g_2g_1(N))\rangle\\=&P_{g_2g_1}|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle.\end{aligned}\tag{8.64}$$

也即是说，置换算符满足如下关系

$$P_{g_2}P_{g_1}=P_{g_2g_1}.\tag{8.65}$$

这一代数关系的实质就在于，它告诉我们置换算符会保持置换操作的乘法关系，两个置换操作的乘积对应于两个相应置换算符的乘积。数学家称这种保持乘法关系的从置换群 S_N 到希尔伯特空间幺正算符的对应关系为置换群的幺正表示。也即是说，我们定义的置换算符刚好构成了相应置换群的幺正表示。

考虑到粒子的全同性，则有，置换算符的作用不会改变全同粒子系统的物理状态，即置换算符作用以后的量子态和原来的量子态只能相差一个相位。因此，N 个全同粒子的物理态空间不是 N 粒子希尔伯特空间 $\mathscr{H}$，而是它的一个子空间，这个子空间的态额外满足下面的约束条件

$$P_g|\psi\rangle = e^{-i\varphi(g)}|\psi\rangle. \tag{8.66}$$

虽然我们现在考虑的是 N 个全同粒子，但是由于 P_g 得构成置换群的幺正表示，所以这个约束方程的解其实和两全同粒子类似，即只有两个不同的解。其中一个解描述全同玻色子，这时候对于任意 g，恒有 $e^{-i\varphi(g)}=1$。另一个解描述全同费米子，这时候 $e^{-i\varphi(g)}=\varepsilon(g)$，记号 $\varepsilon(g)$ 对于偶置换取 1, 对于奇置换取 -1。这个一般性结论的证明其实并不难，关键点就在于要注意到：(1) 任意置换 g 总能写成一系列两粒子对换的乘积. (2) 由于 P_g 是置换群的表示，所以它也可以写成一系列两粒子对换算符的乘积。注意到这两点以后，人们就可以轻易地将 N 个全同粒子的分析约化为对两个全同粒子的分析，进而得出刚才所说的两个不同解。其中，玻色子的物理态空间称作对称态子空间，记为 $\mathscr{H}_S$，费米子的物理态空间称作反对称态子空间，记作 $\mathscr{H}_A$。

与两全同粒子类似，假如我们将 $|\psi\rangle$ 用基矢量 $|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle$ 展开成

$$|\psi\rangle = \sum_{i_1,i_2,...,i_N} \psi_{i_1,i_2,...,i_N}|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle. \tag{8.67}$$

则对于全同玻色子体系，叠加系数 (波函数) 将满足，$\psi_{i_1,i_2,...,i_N}$ 对于 N 个指标的任意置换都对称不变，这种波函数简称对称波函数，或者称为全对称波函数。而对于全同费米子体系，$\psi_{i_1,i_2,...,i_N}$ 对于任意两个指标的对换都反对称，比方说 $\psi_{i_2,i_1,...,i_N}=-\psi_{i_1,i_2,...,i_N}$, 这种波函数称作反对称波函数，或者称作全反对称波函数。因此，对称态子空间 $\mathscr{H}_S$ 其实就是所有对称波函数的空间，而反对称态子空间 $\mathscr{H}_A$ 则是所有反对称波函数的子空间。

与两全同粒子情形类似，对称态空间 $\mathscr{H}_S$ 的基矢量可以按照如下方法构造

$$|u_{i_1},u_{i_2},...,u_{i_N}\rangle_S = \sum_g P_g|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle. \tag{8.68}$$

证明其实非常简单，假如将某个置换算符 P_{g_1} 作用在这样的基矢量上，则由于 $P_{g_1}\sum_g P_g = \sum_g P_{g_1}P_g = \sum_g P_{g_1 g} = \sum_{g'} P_{g'}$(最后一个等号是令 $g_1 g = g'$，并将对 g 的求和替换成对 g' 的求和)，因此我们将有

$$\begin{aligned} P_{g_1}|u_{i_1},u_{i_2},...,u_{i_N}\rangle_S &= P_{g_1}\sum_g P_g|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle \\ &= \sum_{g'} P_{g'}|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle = |u_{i_1},u_{i_2},...,u_{i_N}\rangle_S. \end{aligned} \tag{8.69}$$

这就证明了这样的基矢量在任何置换算符的作用下都保持不变，因此是对称态空间的基矢量。

而反对称态空间 $\mathscr{H}_A$ 的基矢量的构造方法则是

$$|u_{i_1},u_{i_2},...,u_{i_N}\rangle_A=\sum_g \varepsilon(g)P_g|u_{i_1}(1),u_{i_2}(2),...,u_{i_N}(N)\rangle. \tag{8.70}$$

要证明它是反对称态空间的基矢量，只需要注意到根据置换的奇偶性，$\varepsilon(g)$ 满足如下关系

$$\varepsilon(g_2)\varepsilon(g_1)=\varepsilon(g_2g_1),\quad \varepsilon(g^{-1})=\varepsilon(g). \tag{8.71}$$

进而可以得出，

$$P_{g_1}\big(\sum_g \varepsilon(g)P_g\big)=\varepsilon(g_1)\big(\sum_{g'}\varepsilon(g')P_{g'}\big), \tag{8.72}$$

具体的证明细节我们留给读者自己练习。

全同费米子体系反对称态基矢量的构建公式 (8.70) 还可以写成一种更常见的行列式形式

$$\begin{aligned}|u_{i_1},u_{i_2},...,u_{i_N}\rangle_A&=\sum_g \varepsilon(g)|u_{i_1}(g(1)),u_{i_2}(g(2)),...,u_{i_N}(g(N))\rangle.\\&=\sum_g \varepsilon(g)|u_{i_1}(g(1))\rangle|u_{i_2}(g(2))\rangle...|u_{i_N}(g(N))\rangle.\\&=\det\begin{pmatrix}|u_{i_1}(1)\rangle & |u_{i_1}(2)\rangle & ... & |u_{i_1}(N)\rangle\\ |u_{i_2}(1)\rangle & |u_{i_2}(2)\rangle & ... & |u_{i_2}(N)\rangle\\ ... & ... & ... & ...\\ |u_{i_N}(1)\rangle & |u_{i_N}(2)\rangle & ... & |u_{i_N}(N)\rangle\end{pmatrix}\end{aligned} \tag{8.73}$$

在最后的行列式中，我们看到，每个单粒子态相应于行列式的一行，每个被标记的粒子相应于行列式的一列。很明显，将任意两个粒子对换相应于交换行列式的两列，因此当然是反对称的。这就再一次证明了，最终的行列式的确能构成反对称态空间的基矢量。另外，从最终的行列式中还可以轻易看出另一个重要结果，即一个单粒子态上不可能出现两个或两个以上的全同费米子，否则的话，这个行列式就会有两行或者更多行是完全一样的，而线性代数的知识告诉我们，这样的行列式必定为零，从而这种情况在物理上不可能出现。这个结果就是我们前面提到过的泡利不相容原理。

对于全同粒子体系，置换算符的作用不仅不能改变系统的物理状态，而且置换算符的作用效果应该是无法观测的，因为对于全同粒子体系而言，将粒子作一个任意置换对于物理结果不应该有任何影响。这就意味着，置换算符必须得和任意物理可观测量 $\mathscr{O}$ 均对易，即

$$P_g\mathscr{O}=\mathscr{O}P_g. \tag{8.74}$$

特别的，置换算符得和全同粒子哈密顿量对易。从而置换操作是全同粒子体系的一种特殊对称性。而且，由于置换算符与全同粒子哈密顿量对易，因此它当然也和时间演化算符

$U(t_f,t_i)$ 对易, 即

$$P_g U(t_f,t_i) = U(t_f,t_i)P_g. \tag{8.75}$$

在实际应用中，我们总是反过来应用公式 (8.74) 的，即我们总是将它看成是对全同粒子物理可观测量的一个约束，只有满足这个约束的量才可能是可观测量。比方说，根据这个约束，N 个全同粒子体系的位置算符 $\mathbf{X}_1+\mathbf{X}_2+...+\mathbf{X}_N$ 是可观测的，但是，$\mathbf{X}_1$ 是无法观测的。

与时间演化算符分析方法的关系

这一个小节我们主要想解决两个问题：第一，以两全同费米子的散射为例，分析全同粒子体系散射问题的特殊性。第二，将本节着重强调的置换算符和置换对称性分析方法与本章前几节中的时间演化算符和路径积分分析方法联系起来。

假设两全同费米子的散射初态为 $|u_{i_1},u_{i_2}\rangle_A$ 态，根据前面的构造方法，

$$|u_{i_1},u_{i_2}\rangle_A = \frac{1}{\sqrt{2}}(1-P_{12})|u_{i_1}(1),u_{i_2}(2)\rangle, \tag{8.76}$$

式中 $\frac{1}{\sqrt{2}}$ 为归一化因子。类似的，假设两全同费米子的散射末态为 $|v_{f_1},v_{f_2}\rangle_A$ 态，其构造为

$$|v_{f_1},v_{f_2}\rangle_A = \frac{1}{\sqrt{2}}(1-P_{12})|v_{f_1}(1),v_{f_2}(2)\rangle. \tag{8.77}$$

则两全同费米子的散射概率幅相应于时间演化算符的如下矩阵元

$${}_A\langle v_{f_1},v_{f_2}|U(t_f,t_i)|u_{i_1},u_{i_2}\rangle_A. \tag{8.78}$$

代入散射初态和散射末态的具体构造，我们就可以得到

$$\begin{aligned}&{}_A\langle v_{f_1},v_{f_2}|U(t_f,t_i)|u_{i_1},u_{i_2}\rangle_A\\ =&\frac{1}{2}\langle v_{f_1}(1),v_{f_2}(2)|(1-P_{12}^\dagger)U(t_f,t_i)(1-P_{12})|u_{i_1}(1),u_{i_2}(2)\rangle\end{aligned} \tag{8.79}$$

注意到 P_{12} 与 $U(t_f,t_i)$ 对易, 再进一步利用 $P_{12}^\dagger = P_{12}^{-1} = P_{12}$(首先由于 P_{12} 为幺正算符，所以 $P_{12}^\dagger = P_{12}^{-1}$，又由于 $P_{12}^2=1$，所以 $P_{12}^{-1}=P_{12}$)。我们很轻易就能推导出 (8.79) 等式右边的表达式等价于

$$\begin{aligned}&\langle v_{f_1}(1),v_{f_2}(2)|U(t_f,t_i)|u_{i_1}(1),u_{i_2}(2)\rangle\\ &-\langle v_{f_1}(1),v_{f_2}(2)|P_{12}^\dagger U(t_f,t_i)|u_{i_1}(1),u_{i_2}(2)\rangle\\ =&\langle v_{f_1}(1),v_{f_2}(2)|U(t_f,t_i)|u_{i_1}(1),u_{i_2}(2)\rangle\\ &-\langle v_{f_2}(1),v_{f_1}(2)|U(t_f,t_i)|u_{i_1}(1),u_{i_2}(2)\rangle.\end{aligned} \tag{8.80}$$

这就是两全同费米子散射幅的最终表达式，这个表达式有非常简单的解释，它的两项分别对应于 (8.15) 中的左图和右图。如果我们考察的是两个全同玻色子的散射，那推导过程是

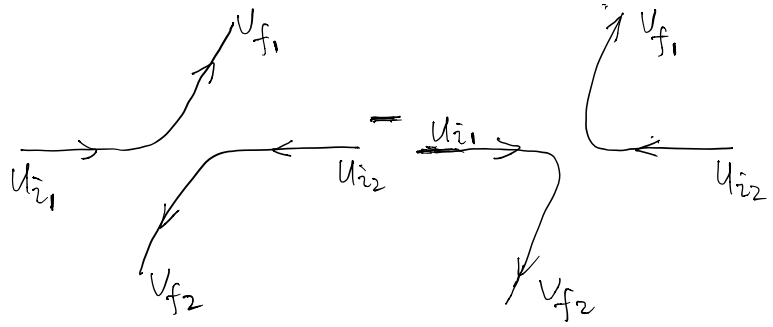

Figure 8.15: 两全同费米子散射，左图相应于散射幅 $\langle v_{f_1}(1), v_{f_2}(2)|U(t_f,t_i)|u_{i_1}(1), u_{i_2}(2)\rangle$，右图相应于散射幅 $\langle v_{f_2}(1), v_{f_1}(2)|U(t_f,t_i)|u_{i_1}(1), u_{i_2}(2)\rangle$.

完全类似的，只不过这时候，两幅图对应的散射幅应该是相加而不是相减。至此为止，实际上我们就完全解决了本章一开始提出来的问题。

总之，对于两全同费米子的散射幅，我们有

$$\begin{aligned} &{}_A\langle v_{f_1}, v_{f_2}|U(t_f,t_i)|u_{i_1}, u_{i_2}\rangle_A \\ =&\langle v_{f_1}(1), v_{f_2}(2)|U(t_f,t_i)|u_{i_1}(1), u_{i_2}(2)\rangle \\ -&\langle v_{f_2}(1), v_{f_1}(2)|U(t_f,t_i)|u_{i_1}(1), u_{i_2}(2)\rangle. \end{aligned} \tag{8.81}$$

对于全同玻色子情形，则应该将反对称态改为对称态，相应的公式右边的两项应该是相加而不是相减。但是，前面的图 (8.15) 画的是粒子的运动空间轨道图，如果我们画出粒子完整的时空路径图，那散射幅 ${}_A\langle v_{f_1}, v_{f_2}|U(t_f,t_i)|u_{i_1}, u_{i_2}\rangle_A$ 就应该示意性地表示为图 (8.16)。阅

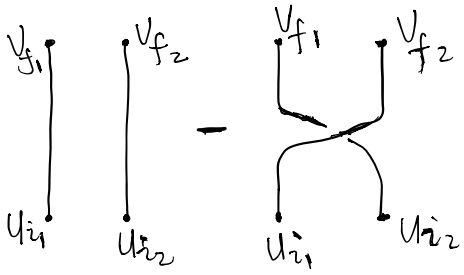

Figure 8.16: 左图相应于散射幅 $\langle v_{f_1}(1), v_{f_2}(2)|U(t_f,t_i)|u_{i_1}(1), u_{i_2}(2)\rangle$，右图相应于散射幅 $\langle v_{f_2}(1), v_{f_1}(2)|U(t_f,t_i)|u_{i_1}(1), u_{i_2}(2)\rangle$。竖直方向为时间轴。

读了本章前几节的读者一定已经发现，这个结果和我们前面用路径积分方法推导出来的结果完全一致。

8.5.3 应用举例

$\nu = 1$ 的整数量子霍尔效应基态波函数

在第四章中我们研究过朗道能级以及相应的量子霍尔效应。在那里我们看到，电子最低朗道能级的单粒子态为

$$u_i(w,\overline{w}) = w^i e^{-\frac{|w|^2}{4l_B^2}}, \tag{8.82}$$

式中 $w = x + iy$ 为两维平面的复坐标。当然这是一个 $2+1$ 维的系统，但是对于电子系统而言，它在 $2+1$ 维所满足的统计和 $3+1$ 维是一样的，都是全同费米子统计。因此对于填满最低朗道能级的 $N_\Phi = \frac{eBS}{2\pi\hbar}$ 个电子而言，它们的多体波函数必须是一个全反对称波函数，而且 N_Φ 个电子填充到 N_Φ 个态上，填充方法是唯一的，因此相应的反对称态基矢量也是唯一的，由下面的行列式给出 (式中 $N = N_\Phi$)

$$\Psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) = \det\begin{pmatrix} 1 & 1 & ... & 1 \\ w_1 & w_2 & ... & w_N \\ w_1^2 & w_2^2 & ... & w_N^2 \\ ... & ... & ... & ... \\ w_1^{N-1} & w_2^{N-1} & ... & w_N^{N-1} \end{pmatrix} \prod_{n=1}^{N} e^{-\frac{|w_n|^2}{4l_B^2}}. \tag{8.83}$$

这个行列式就是著名的范德蒙行列式，利用范德蒙行列式的标准计算结果，我们可以将上面的多体波函数重新表达成

$$\Psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) = \prod_{1\le n<m\le N}(w_n - w_m)\prod_{n=1}^{N} e^{-\frac{|w_n|^2}{4l_B^2}}. \tag{8.84}$$

显然，这个波函数在任意两个电子位置坐标的置换下都反对称。因而的确符合全同费米子波函数的要求。这个波函数就是填充分数 $\nu = 1$ 的整数量子霍尔效应的基态波函数。不仅如此，Laughlin 正是通过推广这个整数量子霍尔效应波函数，从而给出了分数量子霍尔效应的一个理论解释。

单费米子近似

下面我们考察一个特殊的 N 个全同费米子的系统，假设这些费米子之间没有相互作用，因此整个系统的哈密顿量 H 可以写成

$$H = h(1) + h(2) + ... + h(N). \tag{8.85}$$

式中 $h(n)$ 表示第 n 个费米子的哈密顿量，并且所有的 $h(n)$ 形式都一样，因此 H 与任意置换算符 P_g 均对易。

为了构造这个系统的态空间，我们可以首先考察单粒子态，设单粒子哈密顿量 h 的本征方程为

$$h|u_i\rangle = e_i|u_i\rangle, \tag{8.86}$$

式中 e_i 为单粒子能量本征值，不妨设 $e_1 < e_2 < ... < e_N < e_{N+1} < ....$。由于粒子间没有相互作用，则现在多体量子态的问题就转换为将 N 个全同费米子填充进这些单粒子能级的问

题。由于泡利不相容原理，每个费米子要占据一个不同的能级。显然，总能量最低的情况是，前 N 个单粒子能级被占据，系统的总能量为 $E_g = e_1 + e_2 + ... + e_N$。这就构成了整个系统的基态，其波函数是

$$|u_1, u_2, ..., u_N\rangle_A = \sum_g \varepsilon(g)|u_1(g(1)), u_2(g(2)), ..., u_N(g(N))\rangle. \tag{8.87}$$

全同费米子系统基态所对应的最高被占据单粒子能级就称作**费米能级**，在这里费米能级就是 e_N。

但是，如果我们考察的是全同玻色子体系，那情况将完全不同，这时候没有泡利不相容原理，所以 N 个全同玻色子可以全都占据最低单粒子能级 e_1，这就叫玻色-爱因斯坦凝聚。它告诉我们，当温度足够低时，玻色子气体可以全都凝聚到最低单粒子能级上去。

空间波函数与自旋波函数分离的情形

对于有些全同费米子系统，它的空间动力学变量和自旋变量之间没有耦合，简称没有自旋-轨道耦合。那这时候总的多体波函数就可以分离变量成空间波函数与自旋波函数的乘积。对于这种情况的全同费米子系统，由于全同性原理只要求总波函数是反对称波函数，所以我们既可以要求空间波函数对称而自旋波函数反对称，也可以要求空间波函数反对称而自旋波函数对称，这两种情况的总波函数均反对称，从而都可以满足全同性原理的要求。

比方说，对于一个类氢离子，其核外有两个电子（全同费米子），均处于 $n = 1, l = m = 0$ 的最低单粒子能级 $|\psi_{100}\rangle$，因此，系统的空间波函数必定是对称的 $|\psi_{100}, \psi_{100}\rangle$，从而这两个电子的自旋波函数必定反对称，也就是必定处于自旋单态 $\big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big)/\sqrt{2}$，从而系统的总波函数必定为

$$|\psi_{100}, \psi_{100}\rangle\big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big)/\sqrt{2}. \tag{8.88}$$

8.6 * 产生湮灭算符与全同粒子量子力学

前面我们描述全同 N 粒子系统时，我们是先把这 N 个粒子标记为 $1, 2, ..., N$ 再来处理的。这当然是一种表示 N 粒子置换对称性的方法，通过这种方法我们可以把全同粒子希尔伯特空间构建为 N 粒子希尔伯特空间 $\mathscr{H}$ 的子空间。相应的全同粒子体系的物理可观测量也可以表示为通常的 N 粒子体系的算符。这样的处理方式当然能够解决问题，但却不是最好的描述全同粒子体系的方式。这一节我们将看到，有一个巧妙的代数方法可以很方便地表示全同玻色子体系或者全同费米子体系的置换对称性，从而直接了当地建立它们的物理态空间和可观测量算符。顺便提一下，由于历史的原因，这种代数方法有一个很夸张的名称，称之为二次量子化。二次量子化方法当然没有其名字所暗示的那么夸张的物理内涵，它并不意味着我们要建立一个量子力学的量子力学，实际上量子力学只有一个，二次量子化方法只是一种描述全同粒子体系量子力学的巧妙代数方法而已，它的一部分重要性在于，它可以自然地将我们引导向量子场论。

为了说清楚这种代数方法，我们首先引入一些记号，我们将单粒子正交归一基矢量 $|u_i\rangle$ 简记为 $|i\rangle$，即 $|i\rangle = |u_i\rangle$。从而全同玻色子的对称态基矢量可以重记为 $|i_1, i_2, ..., i_N\rangle_S =$

$|u_{i_1},u_{i_2},...,u_{i_N}\rangle_S$, 它在任意一对指标交换下均保持对称不变。而全同费米子的反对称态基矢量可以重记为 $|i_1,i_2,...,i_N\rangle_A = |u_{i_1},u_{i_2},...,u_{i_N}\rangle_A$, 它在任意一对指标交换下均要出一个负号，从而反对称。我们将这两种情况统一在一起，记作 $|i_1,i_2,...,i_N\rangle_\zeta$, ζ 仅有两个取值 ±1, $\zeta=1$ 对应全同玻色子体系的对称态，$\zeta=-1$ 对应全同费米子体系的反对称态。

二次量子化方法中的多粒子态

二次量子化这种代数方法的关键点在于将不同粒子数目的全同粒子体系放在一起（也就是放在同一个希尔伯特空间里）来考察，这也是它能自然引导向量子场论的关键，因为在量子场论中粒子可以产生和湮灭，其数目通常无法固定。这个包含不同粒子数的希尔伯特空间通常称作 Fock 空间，记为 $\mathscr{F}$。

为了建立 Fock 空间，首先我们定义 0 个全同粒子对应 $\mathscr{F}$ 里的一个量子态 $|0\rangle$，称作真空态 (注意它不是量子比特里的 $|0\rangle$ 态)。下面我们引入关键的定义，我们定义单粒子态 $|i\rangle$ 为某个算符 $a_i^\dagger$ 作用在 $|0\rangle$ 态上产生的，即

$$a_i^\dagger|0\rangle = |i\rangle. \tag{8.89}$$

同理可以定义多粒子态，比如 2 粒子态 $|i_1,i_2\rangle_\zeta$ 和 3 粒子态 $|i_1,i_2,i_3\rangle_\zeta$ 可以简单定义如下

$$\begin{aligned} |i_1,i_2\rangle_\zeta &= a_{i_1}^\dagger a_{i_2}^\dagger|0\rangle, \\ |i_1,i_2,i_3\rangle_\zeta &= a_{i_1}^\dagger a_{i_2}^\dagger a_{i_3}^\dagger|0\rangle. \end{aligned} \tag{8.90}$$

另一方面，根据对称态和反对称态的定义我们又有

$$|i_1,i_2\rangle_\zeta = \zeta|i_2,i_1\rangle_\zeta = \zeta a_{i_2}^\dagger a_{i_1}^\dagger|0\rangle. \tag{8.91}$$

由此可见，要让我们构造出来的多粒子态自动是对称态或者反对称态，我们只需令算符 $a_i^\dagger$ 满足如下代数关系

$$a_{i_1}^\dagger a_{i_2}^\dagger = \zeta a_{i_2}^\dagger a_{i_1}^\dagger, \tag{8.92}$$

$\zeta=1$ 的情形我们称相应的算符为玻色算符，它对应于两个 $a^\dagger$ 算符对易，$\zeta=-1$ 的情形我们称相应的算符为费米算符，它对应于两个 $a^\dagger$ 算符反对易。因此通过令算符对易或者反对易，(8.92) 这个简单的代数办法就解决了全同粒子体系物理态的对称或反对称问题。

特别的，(8.92) 式这个代数关系告诉我们，对于全同费米子体系 $(a_i^\dagger)^2 = -(a_i^\dagger)^2 = 0$, 这其实就对应 $|i,i\rangle_A = 0$, 这也就是泡利不相容原理。

从前面的定义我们可以很清楚地看到，$a_i^\dagger$ 算符的作用会将全同粒子的数目增加 1，或者说 $a_i^\dagger$ 的物理含义就是产生一个 $|i\rangle$ 态的粒子，所以称作产生算符。当然，按照厄密共轭算符的定义，$a_i^\dagger$ 的厄密共轭算符 a_i 的作用就会将全同粒子数目减少 1，具体来说，a_i 会湮灭一个原本占据在 $|i\rangle$ 态上的粒子，故称作湮灭算符。很快读者就可以看到，这里的产生-湮灭算符和我们在前面的章节中求解线性谐振子时引入的产生-湮灭算符在数学上是一回事，但是物理解释有所不同，这里我们处理的是多粒子量子力学，产生或湮灭的是一个粒子，而在求解线性谐振子时产生或湮灭的是一个能级量子数。

和线性谐振子的这种类比启发我们按如下方法建立 Fock 空间的一组基矢量。我们要解决的问题就是如何更方便地标记多个全同粒子的量子态，或者说如何编码多粒子态。前面我们采用的方法是先把粒子标记为 $1,2,...,N$，然后按照粒子的顺序将多粒子态编码成 $|u_{i_1}(1),u_{i_2}(2),....\rangle$ 的形式，最后再要求对于全同粒子最终的物理态得全对称或者全反对称。现在我们作一个关键的改变，我们换成按照单粒子态的某个排列顺序编码多粒子态，也即是说，我们将单粒子态指标 i 排列成 $1,2,3,.....$ 的顺序，然后依次看第 1 个态上占据了多少个全同粒子，第 2 个态上又占据了多少个全同粒子，依此类推。也就是说，我们将多粒子态编码成 $|n_1,n_2,n_3,....\rangle$ 这一串非负整数，其中 n_i 表示 i 态上占据的全同粒子数目，称作 i 态上的占据数。

知道了单粒子态的占据情况，再加上交换两个粒子时对称或者反对称的要求，实际上就唯一确定了多个全同粒子的量子态，所以我们可以唯一性地把全同多粒子态编码成 $|n_1,n_2,n_3,....\rangle$ 的形式，这就是 Fock 空间的一组自然基矢量，相应的表象称作占据数表象。特别的，对于 $\zeta=-1$ 的全同费米子情形，由于泡利不相容原理，一个单粒子态上不能有两个或两个以上全同费米子占据，所以这时候，n_i 只有 0 或 1 这两种取值，即 $n_i=0,1$。当然，对于全同玻色子体系，n_i 取多少没有什么限制，只要大于等于 0 就可以了。

由于 $a_i^\dagger$ 是在单粒子态 i 上产生一个粒子，a_i 是将 i 态上原本占据的一个粒子湮灭掉，所以很显然我们将有

$$\begin{aligned} a_i^\dagger|n_1,n_2,...,n_i,....\rangle &\propto |n_1,n_2,...,n_i+1,....\rangle \\ a_i|n_1,n_2,...,n_i,....\rangle &\propto |n_1,n_2,...,n_i-1,....\rangle. \end{aligned} \tag{8.93}$$

在上面的式子中，所有的占据数都应该大于等于 0，如果某个占据数小于 0，则相应的表达式本身为 0。对于全同费米子体系，占据数还应该小于等于 1，如果某个占据数大于 1，则相应的表达式本身也为 0。

类比于线性谐振子问题，我们可以知道，占据数 n_i 应该是算符 $\widehat{n}_i=a_i^\dagger a_i$ 的本征值。由此类比于线性谐振子问题，我们可以知道

$$\begin{aligned} a_i^\dagger|n_1,n_2,...,n_i,....\rangle &= \sqrt{n_i+1}\,\zeta^{s_i}|n_1,n_2,...,n_i+1,....\rangle \\ a_i|n_1,n_2,...,n_i,....\rangle &= \sqrt{n_i}\,\zeta^{s_i}|n_1,n_2,...,n_i-1,....\rangle. \end{aligned} \tag{8.94}$$

对于玻色算符情形，这就是线性谐振子问题中告诉我们的标准的产生-湮灭算符作用关系。对于费米算符情形，这个公式多了一个因子 ζ^{s_i}，这里 $s_i=\sum_{j=1}^{i-1}n_j$, 为 i 态前面的 $i-1$ 个态上的总占据数。多出这个因子其实很好理解，因为这时候 a_i 或者 $a_i^\dagger$ 要正确地作用在 i 态上，它必须先和 i 前面的 s_i 个产生算符相交换，对于费米算符，每一次交换出一个因子 ζ，最后多出来的因子就是 ζ^{s_i}。

利用代数关系式 (8.94)，我们容易推导如下产生湮灭算符之间应该满足的标准代数关系式

$$[a_i,a_j^\dagger]_\zeta=\delta_{ij}.\quad [a_i^\dagger,a_j^\dagger]_\zeta=0,\quad [a_i,a_j]_\zeta=0. \tag{8.95}$$

这里 $[A,B]_\zeta=AB-\zeta BA$，即对于玻色算符它是普通的对易子，但是对于费米算符，它其实是反对易子。(8.95) 式的后两个等式没有任何新东西，它们显然就是将 (8.92) 式重写一下。

但 (8.95) 式的第一个等式是相当非平凡的，特别的，对于全同费米子情形，它告诉我们

$$a_i a_i^\dagger + a_i^\dagger a_i = 1. \tag{8.96}$$

这就是费米子产生湮灭算符所满足的反对易关系，它和玻色子的对易关系 $a_i a_i^\dagger - a_i^\dagger a_i = 1$ 有本质区别。

二次量子化方法中的算符

如此一来，利用产生-湮灭算子的对易或反对易性质我们就表示了全同玻色子或全同费米子的置换对称性，并且还建立了全同粒子体系的 Fock 空间。下面我们来讨论如何用产生-湮灭算子表示全同粒子体系的物理量算符。

首先我们来看单体算符，即不涉及多体相互作用的算符。这样的算符比方有动能算符等等。为简单起见，假设某个这样的单体算符 $\mathscr{O}_1$ 在单粒子希尔伯特空间中对应于算符 $\widehat{o}$, 假设在矢量基 $|i\rangle = |u_i\rangle$ 中 $\widehat{o}$ 是对角化的，本征值为 o_i，很明显 $o_i = \langle i|\widehat{o}|i\rangle$。则在全同粒子的多体希尔伯特空间中，$\mathscr{O}_1$ 的总本征值为每个粒子的本征值之和，也即是说，在 Fock 空间矢量基 $|n_1, n_2, n_3, ...\rangle$ 中，$\mathscr{O}_1$ 的本征值为

$$\sum_i o_i n_i. \tag{8.97}$$

显然，为了得到这样的本征值，我们只需将 $\mathscr{O}_1$ 表示成如下算符

$$\mathscr{O}_1 = \sum_i o_i \widehat{n}_i = \sum_i \langle i|\widehat{o}|i\rangle a_i^\dagger a_i. \tag{8.98}$$

为了得到 $\mathscr{O}_1$ 在任意单粒子矢量基中的表达式，我们利用单粒子希尔伯特空间的基矢变换 $|i\rangle = \sum_\alpha |\alpha\rangle\langle\alpha|i\rangle$, 并定义 $a_\alpha^\dagger|0\rangle = |\alpha\rangle$, $a_i^\dagger|0\rangle = |i\rangle$, 从而容易有

$$a_i^\dagger = \sum_\alpha a_\alpha^\dagger \langle\alpha|i\rangle, \quad a_i = \sum_\alpha \langle i|\alpha\rangle a_\alpha. \tag{8.99}$$

利用这样的基矢变换，我们就能将 (8.98) 式变换到任意单粒子矢量基中，从而有

$$\mathscr{O}_1 = \sum_{\alpha,\beta} \langle\alpha|\widehat{o}|\beta\rangle a_\alpha^\dagger a_\beta. \tag{8.100}$$

特别的，我们有总粒子数算符 $N = \sum_i a_i^\dagger a_i = \sum_\alpha a_\alpha^\dagger a_\alpha$。这是最简单的单体算符，其本征值 $\sum_i n_i$ 为系统的总粒子数目。很容易验证的是，以上的任何单体算符都与粒子数算符 N 对易，即 $[N, \mathscr{O}_1] = 0$。

下面我们来看两体相互作用算符 $\mathscr{O}_2$。假设它在两体希尔伯特空间相应于算符 $\widehat{v}$, 假设 $\widehat{v}$ 在两体矢量基 $|i,j\rangle$ 上的本征值为 v_{ij}，即 $\widehat{v}|i,j\rangle = v_{ij}|i,j\rangle$，很容易验证，这里无论 $|i,j\rangle$ 态是对称态还是反对称态，v_{ij} 关于指标 i,j 均是对称的。则在多粒子希尔伯特空间矢量基 $|n_1, n_2, n_3, ...\rangle$ 中，两体算符 $\mathscr{O}_2$ 的总本征值为所有可能两体本征值的和，即为

$$\frac{1}{2}\sum_{i\neq j} v_{ij} n_i n_j + \frac{1}{2}\sum_i v_{ii} n_i (n_i - 1), \tag{8.101}$$

式中第 1 项来自两个不同单粒子态上的粒子相互作用，其中因子 1/2 是因为 i,j 相互作用和 j,i 相互作用是一样的。式中第 2 项来自两个相同单粒子态上的粒子相互作用。很容易看出来，上式的结果也可以重写成

$$\frac{1}{2}\sum_{ij} v_{ij}(n_i n_j - n_i\delta_{ij}). \tag{8.102}$$

式中 $n_i n_j - n_i\delta_{ij}$ 是如下算符的本征值

$$\begin{aligned} &a_i^\dagger a_i a_j^\dagger a_j - a_i^\dagger a_i\delta_{ij} = a_i^\dagger\big(\delta_{ij}+\zeta a_j^\dagger a_i\big)a_j - a_i^\dagger a_i\delta_{ij} \\ =&\zeta a_i^\dagger a_j^\dagger a_i a_j = a_i^\dagger a_j^\dagger a_j a_i. \end{aligned} \tag{8.103}$$

所以为了在 Fock 空间上取到本征值 (8.102)，我们只需将两体算符 $\mathscr{O}_2$ 取成

$$\mathscr{O}_2 = \frac{1}{2}\sum_{ij} v_{ij} a_i^\dagger a_j^\dagger a_j a_i. \tag{8.104}$$

注意到 $v_{ij} = \langle i,j|\hat{v}|i,j\rangle$，从而可以进一步将上式变换到任意单粒子矢量基中，得到

$$\mathscr{O}_2 = \frac{1}{2}\sum_{\alpha\beta\gamma\delta}\langle\alpha\beta|\hat{v}|\gamma\delta\rangle a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma. \tag{8.105}$$

值得一提的是，人们很容易验证，任何以上形式的两体算符都和粒子数算符 $N=\sum_i a_i^\dagger a_i$ 对易，即有 $[N,\mathscr{O}_2]=0$。

二次量子化与场

下面假设我们忽略粒子自旋，考察单粒子希尔伯特空间的一组特殊矢量基，即坐标表象 $|\mathbf{x}\rangle$，即在上文中取 $i=\mathbf{x}$，相应的 $a_i = a_{\mathbf{x}}$ 现在我们记作 $a(\mathbf{x})$，对 i 的求和就变成了对 $\mathbf{x}$ 的积分，克隆内克符号 δ_{ij} 就变成了狄拉克 δ 函数。因此，(8.95) 式现在应该是

$$[a(\mathbf{x}), a^\dagger(\mathbf{x}')]_\zeta = \delta(\mathbf{x}-\mathbf{x}'). \tag{8.106}$$

很显然，粒子数算符 N 现在应该是

$$N = \int d^3\mathbf{x} a^\dagger(\mathbf{x})a(\mathbf{x}). \tag{8.107}$$

单粒子的能量算符 $\widehat{h}$ 现在应该是 $\widehat{h} = \frac{\widehat{\mathbf{p}}^2}{2m} + V(\mathbf{x})$, $V(\mathbf{x})$ 为粒子受到的外势场作用，代入单体算符 $\mathscr{O}_1$ 的公式 (8.100)，我们就能得到单体的哈密顿量

$$H_0 = \int d^3\mathbf{x} a^\dagger(\mathbf{x})\big[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\big]a(\mathbf{x}). \tag{8.108}$$

类似的，对于两体相互作用算符，注意到 i,j 应该分别替换成 $\mathbf{x},\mathbf{x}'$，从而前面的 v_{ij} 可以替换成相互作用势能 $v_{\mathbf{x},\mathbf{x}'}$，这里记为 $V(\mathbf{x},\mathbf{x}')$，进而可以将两体相互作用算符 (8.104) 写成

$$\widehat{V} = \frac{1}{2}\int d^3\mathbf{x}\int d^3\mathbf{x}' a^\dagger(\mathbf{x})a^\dagger(\mathbf{x}')V(\mathbf{x},\mathbf{x}')a(\mathbf{x}')a(\mathbf{x}). \tag{8.109}$$

假设仅仅只有这种两体相互作用，那么全同粒子体系完整的哈密顿量就应该是 $H=H_0+\widehat{V}$。前面我们说过，粒子数算符 N 与 H_0 和 $\widehat{V}$ 均对易，因此，$[N,H]=0$。这说明这样的全同粒子体系总粒子数是守恒的，我们可以限制在一个确定粒子数目 N 的子空间里来讨论问题。

以上这些关于坐标表象的讨论很容易推广到包括粒子自旋的情形，相应的推广我们留给读者。坐标表象的讨论自然将我们引导向一个结论，即我们可以将 $a(\mathbf{x})$ 和 $a^\dagger(\mathbf{x})$ 看成是空间分布的场，如此一来，上面关于坐标表象二次量子化方法的讨论就可以自然看成是某种量子场论。

Hubbard 模型

二次量子化方法应用非常广泛，比方说，在凝聚态物理中有大量的模型都是用二次量子化方法表述的。下面我们简单介绍一下这些模型中也许最著名的一个，即所谓的 Hubbard 模型。

我们的任务是要描述一个电子系统 (费米子)，只不过这个电子系统在一个晶格上运动。电子的单粒子态主要由电子所处的晶格位置描述，比如电子处在第 i 个格点，那就标记为 $|i\rangle$ 态，当然电子还有自旋，标记为可以取两个不同值的指标 σ, 即 $\sigma=\uparrow,\downarrow$。所以，我们用费米产生-湮灭算符 $a_{i\sigma},a_{i\sigma}^\dagger$ 来描述这些电子，不过，文献中习惯将电子的产生-湮灭算符记为 $c_{i\sigma},c_{i\sigma}^\dagger$。

假设单个电子在 $|i\rangle,|j\rangle$ 格点状态之间的哈密顿矩阵元为 $\langle i|\widehat{h}|j\rangle$, 假设两体相互作用的矩阵元为 $\langle lm|V|pn\rangle$, 式中 l,m,p,n 均为格点位置。则根据我们在本节前面几个小节的论述可以知道，整个全同电子系统的哈密顿算符可以写成

$$H=\sum_{i,j,\sigma}\langle i|\widehat{h}|j\rangle c_{i\sigma}^\dagger c_{j\sigma}+\frac{1}{2}\sum_{lmnp,\sigma,\sigma'}\langle lm|V|pn\rangle c_{l\sigma}^\dagger c_{m\sigma'}^\dagger c_{n\sigma'}c_{p\sigma}. \tag{8.110}$$

Hubbard 模型就是对上面这个哈密顿量 (8.110) 的进一步简化。为了完成这样的简化，我们得进一步假定：(1). 电子处在当前格点时本身具有能量 ε, 即 $\langle i|\widehat{h}|i\rangle=\varepsilon$。(2). 电子隧穿到相邻格点上的能量矩阵元为 $-t$。即当 i,j 相邻时，$\langle i|\widehat{h}|j\rangle=-t$。(3). 除上面两种情况之外，$\langle i|\widehat{h}|j\rangle=0$。即电子不能一次性隧穿到更远的格点上。最后我们还要假定, (4). 两电子间仅当处于同一格点位置时才会有相互作用，即 $\langle lm|V|pn\rangle$ 仅在 $l=p=m=n$ 时才非零，非零值设为 U。

根据上面假设，我们容易知道 (8.110) 中的相互作用项现在可以简化为

$$\frac{1}{2}U\sum_{i,\sigma,\sigma'}c_{i\sigma}^\dagger c_{i\sigma'}^\dagger c_{i\sigma'}c_{i\sigma}=U\sum_i\widehat{n}_{i\uparrow}\widehat{n}_{i\downarrow}. \tag{8.111}$$

式中 $\widehat{n}_{i\sigma}=c_{i\sigma}^\dagger c_{i\sigma}$, 而且在推导上式最终结果时我们已经用了泡利不相容原理 $c_{i\sigma}^2=0$。因此，结合上面所有的简化假设，最终哈密顿量 (8.110) 就可以被简化为

$$H=-t\sum_{<i,j>}\left[c_{i\sigma}^\dagger c_{j\sigma}+c_{j\sigma}^\dagger c_{i\sigma}\right]+\varepsilon\sum_{i\sigma}c_{i\sigma}^\dagger c_{i\sigma}+U\sum_i\widehat{n}_{i\uparrow}\widehat{n}_{i\downarrow}. \tag{8.112}$$

这就是著名的 Hubbard 模型，式中的 $\sum_{<i,j>}$ 表示对相邻的格点求和。

8.7 * 附录：单粒子路径积分公式推导

关于费曼路径积分，我们还有两个重要的问题有待解决，一是，对所有路径进行求和该怎么求，毕竟路径可不是一个离散变量，甚至一条路径也不是一个简单的实变量，而是一个关于时间 t 的函数，所以如何定义这个求和其实是一个很困难的数学问题。第二，我们还要证明路径积分给出来的结果是幺正的。本节我们就是要就单粒子情形解决这两个问题。我们的解决方案很简单，就是直接从时间演化算符的哈密顿量表述 $U(T,0)=\exp\{-iHT/\hbar\}$ 出发，导出其坐标表象下的路径积分公式。哈密顿量表述的时间演化算符当然是幺正的，因此其最终导出来的路径积分公式当然也满足幺正性。不仅如此，这种推导方式还会给出关于如何对路径进行求和的一个定义。

本附录的处理方式基本上是跟随 Polchinski，String Theory Volume I, Appdendix A, A short course on path integrals 的处理。更多的讨论请读者参阅这份原始材料。关于路径积分，我们推荐的另一份入门材料是，R. Shankar, Principles of Quantum Mechanics 中的相关内容。

下面我们以 $\hat{x}$ 表示一个粒子的正则坐标算符，以 $\hat{p}$ 表示相应的正则动量算符。路径积分其实就是坐标表象下时间演化算符的计算公式。假设粒子从 0 时刻的 $|x_i\rangle$ 位置演化到 T 时刻的 $|x_f\rangle$ 位置，则路径积分关心的就是如下概率幅，

$$\langle x_f|\exp\big(-iHT/\hbar\big)|x_i\rangle. \tag{8.113}$$

在前面的章节中我们引入过海森堡绘景，对于算符 $\mathscr{O}$，其相应海森堡绘景中的算符 $\mathscr{O}(t)$ 为

$$\mathscr{O}(t)=\exp(iHt/\hbar)\mathscr{O}\exp(-iHt/\hbar). \tag{8.114}$$

另外，引入海森堡绘景的位置本征态 $|x,t\rangle$ 对于我们来说是方便的，其定义是

$$|x,t\rangle=\exp(iHt/\hbar)|x\rangle. \tag{8.115}$$

很容易验证，$|x,t\rangle$ 满足 $\hat{x}(t)|x,t\rangle=x|x,t\rangle$，式中 $\hat{x}(t)$ 为海森堡绘景中的位置算符。很显然，路径积分所关心的坐标表象时间演化算符 (8.113) 又可以写成

$$\langle x_f,T|x_i,0\rangle. \tag{8.116}$$

通过在中间 t 时刻插入位置本征态的完备集，我们可以把路径积分关心的概率幅改写成

$$\langle x_f,T|x_i,0\rangle=\int dx\langle x_f,T|x,t\rangle\langle x,t|x_i,0\rangle. \tag{8.117}$$

显然，这个方程就是正文中所说的，时间演化算符所满足的基本方程。现在，我们如图 (8.17) 中所示的那样，将整个时间区间 T 等分成 N 个间距为 $\varepsilon=T/N$ 的小区间，其中各分割时刻分别为

$$t_m=m\varepsilon. \tag{8.118}$$

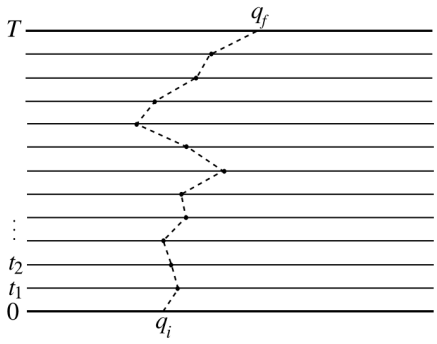

Figure 8.17: 图片来自于 Polchinski 的书，图中的 q_i, q_f 在我们这里分别是 x_i, x_f。

则通过在每一个中间分割时刻都插入相应位置本征态的完备集，我们有

$$\langle x_f, T|x_i, 0\rangle = \int dx_{N-1}...dx_1 \prod_{m=0}^{N-1} \langle x_{m+1}, t_{m+1}|x_m, t_m\rangle. \tag{8.119}$$

式中 $x_0 = x_i$, $x_N = x_f$。

根据我们的定义，有

$$\begin{aligned}\langle x_{m+1}, t_{m+1}|x_m, t_m\rangle &= \langle x_{m+1}|\exp\big(-iH\varepsilon/\hbar\big)|x_m\rangle \\ &= \int dp_m \langle x_{m+1}|p_m\rangle\langle p_m|\exp\big(-iH\varepsilon/\hbar\big)|x_m\rangle.\end{aligned} \tag{8.120}$$

式中哈密顿算符 H 是一个关于坐标算符和动量算符的函数，$H(\hat{p},\hat{x})$。通过算符对易关系我们总是可以将所有的 $\hat{p}$ 都放到 H 表达式的左边，而将所有的 $\hat{x}$ 都对易到右边，从而有

$$\langle p_m|H(\hat{p},\hat{x})|x_m\rangle = H(p_m, x_m)\langle p_m|x_m\rangle. \tag{8.121}$$

从而根据 (8.120) 式我们有 (精确到 ε 的一次方阶)

$$\begin{aligned}\langle x_{m+1}, t_{m+1}|x_m, t_m\rangle &= \int dp_m \exp\big[-iH(p_m, x_m)\varepsilon/\hbar\big]\langle x_{m+1}|p_m\rangle\langle p_m|x_m\rangle \\ &= \int \frac{dp_m}{2\pi\hbar}\exp\Big\{-i\big[H(p_m, x_m)\varepsilon - p_m(x_{m+1}-x_m)\big]/\hbar + O(\varepsilon^2)\Big\}.\end{aligned} \tag{8.122}$$

将 (8.122) 式代入 (8.119) 式，就可以得到

$$\begin{aligned}\langle x_f, T|x_i, 0\rangle &= \int \frac{dp_{N-1}dx_{N-1}}{2\pi\hbar}...\frac{dp_1dx_1}{2\pi\hbar}\frac{dp_0}{2\pi\hbar} \\ &\quad\times\exp\Big\{-i\sum_{m=0}^{N-1}\big[H(p_m, x_m)\varepsilon - p_m(x_{m+1}-x_m)\big]/\hbar + O(\varepsilon^2)\Big\} \\ &\quad\to\int[dpdx]\exp\Big\{\frac{i}{\hbar}\int_0^T dt\big[p\dot{x} - H(p,x)\big]\Big\}.\end{aligned} \tag{8.123}$$

在上式的最后一行中，我们取了 $N\to+\infty$，即 $\varepsilon\to 0$ 的极限。在这个极限下，上式中的复杂多重积分相当于对所有以 $x(0) = x_i, x(T) = x_f$ 为端点的相空间路径 $x(t), p(t)$ 进行积分。上式最后表达式中的 $\int[dpdx]$ 是对这个积分测度的简记符号。

通常来说，哈密顿量 $H(p,x)$ 关于正则动量 p 是一个二次型，所以 (8.123) 式中关于 $p(t)$ 的路径积分其实是一个高斯型积分，我们可以先把这个积分积掉，根据高斯积分的一般结果我们知道，积掉 $p(t)$ 首先是要将指数因子中的 p 替换成满足下式的 p，

$$0=\frac{\partial}{\partial p}[p\dot{x}-H(p,x)]=\dot{x}-\frac{\partial}{\partial p}H(p,x), \tag{8.124}$$

这样消去 p 以后，$p\dot{x}-H$ 的结果实际上就是拉格朗日量 $L(x,\dot{x})$, 这个过程其实就是所谓的勒让德变换。因此，由 (8.123) 式我们就可以进一步得到

$$\langle x_f,T|x_i,0\rangle=\int[dx]\exp\left\{\frac{i}{\hbar}\int_0^T dtL(x,\dot{x})\right\}=\sum_{x(t)}e^{\frac{i}{\hbar}S[x(t)]}. \tag{8.125}$$

这个式子就是坐标表象时间演化算符的路径积分公式，在这个式子的最后我们已经示意性地将对所有路径的积分写成了对所有路径的求和。当然，上面在对 $p(t)$ 进行高斯积分时我们会得到一些积分值，在写出 (8.125) 式这个结果时，我们已经把所有这些 $p(t)$ 的积分值都吸收到积分测度 $[dx]$ 的定义中去了。

我们可以利用 (8.125) 式这个路径积分公式讨论量子力学与经典力学间的对应关系。由于两条邻近路径的作用量之差近似为一阶变分 δS，而从 (8.125) 式的最后结果可以看出，当 $\hbar\to 0$ 时, 一般来说相邻路径的相位差 $\frac{\delta S}{\hbar}$ 是随着路径的微小变动快速振荡的，因此在 $\hbar\to 0$ 时，相邻路径通常总是干涉相消的，除非我们考虑的是 $\delta S=0$ 这条路径的邻近路径。对于这条 $\delta S=0$ 的路径，它和邻近路径的相位差近似为 0，从而是干涉加强的。很显然，这就是最小作用量原理，这条干涉加强的路径就是所谓的经典路径。从这个讨论我们再一次看到，经典物理是量子物理在 $\hbar\to 0$ 时的极限。

9. * 再谈分数量子霍尔效应

本章进一步讨论分数量子霍尔效应。我们将看到上一章谈到的任意子如何作为分数量子霍尔液体的准粒子激发而出现，这些准粒子又如何满足奇妙的分数统计。

本章我们还将讨论贝里相位，以及如何应用贝里相位计算准粒子的分数电荷和分数统计。

9.1 从规范不变性到量子化霍尔电导

这一章我们将进一步讨论分数量子霍尔效应。但还是让我们从整数量子霍尔效应开始，我们将首先利用规范不变性论证霍尔电导必须量子化为 e^2/h 的整数倍。然后我们再指出这个论证对于分数量子霍尔效应为什么失效了，又该作什么样的修正。由此引导出分数量子霍尔效应与整数量子霍尔效应的一个本质不同，即在分数量子霍尔效应中，整块霍尔薄片的基态将出现拓扑简并。从这种基态拓扑简并开始，我们将逐步探讨分数量子霍尔效应的一些奇妙性质。

9.1.1 Byers-Yang 定理

考虑一个任意的多电子系统，假设系统中心有一个洞，洞的中心落在坐标原点上，如图 (9.1) 所示。假设我们在洞中通过磁通 Φ。则根据第三章学过的 Aharnaov-Bohm 效应我们

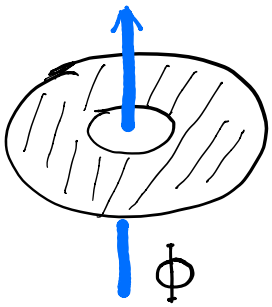

Figure 9.1: 中心有洞的多电子系统，洞中通过磁通 Φ.

知道，当电子围绕着洞转一圈再回到起点时，其波函数将会多出一个相位 $\exp\left(i2\pi\Phi/\Phi_0\right)$，式中 $\Phi_0 = 2\pi\hbar/e$ 称为磁通量子。很显然，如果洞中通过的磁通 Φ 是磁通量子 Φ_0 的整数倍，那这个多出来的相因子将为 1，从而 Aharnaov-Bohm 效应就无法探测到这样的磁通。

不仅如此，1961 年 Byers 和杨振宁还进一步证明，如果洞中通过的磁通改变 Φ_0 的整数倍，则改变前后整个系统的能量本征谱将保持不变。这就是 Byers-Yang 定理。

证明这个定理的关键在于利用规范不变性。根据定态薛定谔方程的规范不变性，我们可以知道，如果对于矢量势 $\mathbf{A}(\mathbf{x})$，多电子系统的哈密顿量有一个能量为 E 的本征波函数 $\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$。那么我们将容易找到另一个相应于矢量势 $\mathbf{A}'(\mathbf{x})$ 的有相同能量 E 的本征波函数 $\psi'(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$，$\mathbf{A}'(\mathbf{x})$ 和 $\psi'(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$ 分别是 $\mathbf{A}(\mathbf{x})$ 和 $\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$ 的规范变换

$$\begin{aligned}\mathbf{A}'(\mathbf{x}) &= \mathbf{A}(\mathbf{x}) + (\hbar/e)\nabla\varepsilon(\mathbf{x}) \\ \psi'(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) &= \Big[\prod_{j=1}^{N} e^{i\varepsilon(\mathbf{x}_j)}\Big]\psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N).\end{aligned} \tag{9.1}$$

对于如图 (9.1) 所示的这种中间有一个洞的系统，我们可以将规范变换函数 $\varepsilon(\mathbf{x})$ 取成下面这样的多值形式

$$\varepsilon(\mathbf{x}) = m\theta(\mathbf{x}), \tag{9.2}$$

式中 $\theta(\mathbf{x})$ 是围绕着洞中心分布的一个角度变量，当人们绕着洞走一圈时，$\theta(\mathbf{x})$ 的取值刚好从 0 变化到 2π。有时候人们称这样的规范变换为大规范变换。从上一段我们知道，系统的能量本征谱在这样的大规范变换下也是不变的。然而，变换以后通过洞的磁通却并非不变，而是会变化为

$$\begin{aligned}\Phi' &= \oint \mathbf{A}'(\mathbf{x})\cdot d\mathbf{x} = \oint \mathbf{A}(\mathbf{x})\cdot d\mathbf{x} + (\hbar/e)m\oint \nabla\theta(\mathbf{x})\cdot d\mathbf{x} \\ &= \Phi + 2\pi m\hbar/e = \Phi + m\Phi_0. \end{aligned} \tag{9.3}$$

即大规范变换前后通过洞的磁通将改变 Φ_0 的整数倍。这就说明，当洞中通过的磁通改变 Φ_0 的整数倍时，系统的能量本征谱将保持不变，这就证明了 Byers-Yang 定理。

9.1.2 量子化霍尔电导

我们知道，整数量子霍尔效应的霍尔电导是量子化的，在前面的第四章中我们也详细讨论过这种量子化的物理起源，本小节我们重新回到这个问题，我们将利用 Byers-Yang 定理给出霍尔电导必然量子化的一个新论证，这就是所谓的 Laughlin 论证。

考察一块圆环状的霍尔薄片，将它放在强磁场中让它量子化。我们把环绕圆环的方向称作纵向，与之垂直的方向称作径向，如图 (9.2) 所示。忽略电子间的相互作用以后，根

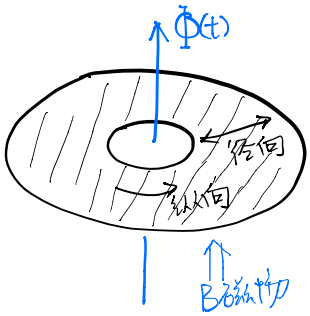

Figure 9.2: 圆环状霍尔薄片，洞中缓慢加上一个磁通量子。

据朗道能级的知识，每一个电子的量子态可以标记为 $|n,m\rangle$ 态，其中 $n = 0,1,2,\ldots$ 用来标记电子所处的朗道能级，m 则是电子的轨道角动量量子数，m 电子处在半径 $r_m = \sqrt{2m}l_B$ 的轨道上，相同朗道能级不同轨道角动量的那些量子态是简并的。现在假设系统的费米能 E_F 处在 $n = \nu - 1$ 和 $n = \nu$ 朗道能级之间，也即是说，系统的前 ν 个朗道能级是填满的。则根据前面第四章的知识我们知道，这时候系统的霍尔电导 σ_H 会量子化为 $\sigma_H = \nu\frac{e^2}{2\pi\hbar}$，这里 ν 是一个量子化整数。下面我们跟随 Laughlin，利用规范不变性来论证这种霍尔电导的量子化。

假设我们绝热地 (也就是足够缓慢地) 在圆环霍尔薄片中心的洞中通过磁通 $\Phi(t)$, $\Phi(t)$ 很缓慢地从 0 增加到一个磁通量子 Φ_0。根据法拉第电磁感应定律，这个变化的磁通会在纵向产生感应电场 $\mathbf{E}$ 以及感应电动势 $\mathscr{E}$,

$$\mathscr{E} = \oint \mathbf{E}\cdot d\mathbf{x} = -\frac{d\Phi(t)}{dt}. \tag{9.4}$$

由于霍尔效应，这个纵向感应电场会在与之垂直的径向产生霍尔电流 J

$$J = \sigma_H E. \tag{9.5}$$

J 从圆环的内边缘流向圆环的外边缘，我们可以将 J 环绕圆环积分，得到径向的电流强度 I,

$$I = \sigma_H \mathscr{E}. \tag{9.6}$$

再次对时间积分，就可以得到一段时间之内从圆环内边缘转移到外边缘的电荷量 ΔQ

$$\Delta Q = \int dt I = \sigma_H \int dt \mathscr{E} = -\sigma_H \int dt \frac{d\Phi(t)}{dt} = -\sigma_H \Delta\Phi. \tag{9.7}$$

式中 $\Delta\Phi$ 是过程前后洞中磁通量的改变量，在我们这里 $\Delta\Phi = \Phi_0$。

当然，整个霍尔薄片是一个多电子系统，从多电子的角度而不是从单电子朗道能级的角度来看，整个系统当然是处在多电子体系的最低能态，也就是基态。由于我们是绝热地增加洞中的磁通，磁通的改变过程足够缓慢，因此这个增加磁通的过程不会将系统激发到激发态上去，整个过程系统都将保持在基态。而根据 Byers-Yang 定理，当我们将洞中通过的磁通改变 Φ_0 时，系统的能量本征谱将保持不变。因此过程前后的基态能量应该是一样的，由于整个霍尔薄片的基态不简并，因此这实际上意味着过程前后霍尔薄片中的多电子系统处在完全一样的基态。

但是，前面我们说过，过程前后圆环内边缘将转移 $\Delta Q = -\sigma_H \Phi_0$ 的电荷到外边缘。如果过程前后薄片体内的多电子系统处于同样的量子态，又由于这个量子态包含的都是一些整数电子电荷，从而只能有整数个电子电荷从内边缘转移到外边缘，不妨设有 p 个电子电荷转移 (p 是一个整数)，即 $\Delta Q = -pe$(注意电子带负电)。从而即有

$$-pe = \Delta Q = -\sigma_H \Phi_0 = -\sigma_H \frac{2\pi\hbar}{e}, \tag{9.8}$$

由此我们就可以得到量子化的霍尔电导

$$\sigma_H = p\frac{e^2}{2\pi\hbar}. \tag{9.9}$$

不仅如此，实际上从单电子量子态 $|n,m\rangle$ 的角度来看，当洞中的磁通增加一个磁通量子时，相应的 $|n,m\rangle$ 态电子就会变成 $|n,m+1\rangle$ 态的电子，它在径向方向上就会往外移动一个轨道。给定一个填满的朗道能级 n，磁通增加一个 Φ_0 时，这个能级上的每一个电子都会往外挪一个轨道，最终的效果就是，内边缘有一个电子挪进了薄片体内，而同时最外面也有一个体内的电子挪到了外边缘上，总体就是，每一个被占据的朗道能级都有一个电子电荷从内边缘转移到了外边缘。由于总共有 ν 个朗道能级被占据，因此最终从内边缘转移到外边缘的电子电荷数目就是

$$p = \nu. \tag{9.10}$$

这就完成了我们对霍尔电导量子化关系的论证。

9.1.3 分数量子霍尔效应的基态简并

上一小节我们利用规范不变性论证了霍尔电导必然要量子化为 $e^2/(2\pi\hbar)$ 的整数倍，即填充分数 ν 要等于整数。然而，在分数量子霍尔效应中，ν 其实取下面这样的分数

$$\nu = p/q, \tag{9.11}$$

式中 p,q 为两个较小的整数。比方说第一个被观测到的分数量子霍尔平台就是 $\nu = 1/3$。问题是，分数量子霍尔效应是在哪个地方使得我们前面的论证不成立了呢?

另一方面，通过前面的论证我们知道，填充分数 ν 反映的是，在圆环霍尔薄片中间的洞中每增加一个磁通量子霍尔薄片内边缘转移到外边缘的电荷量，这个电荷量为 $-\nu e$。因此 $\nu = 1/q$ 这样的分数量子霍尔效应就意味着转移了 $e^* = e/q$ 的分数电荷。由于电子携带的是整数电荷 $-e$，因此这种分数电荷转移就意味着分数量子霍尔效应中必然要出现携带分数电荷的准粒子!

另外，通过上一小节的论证我们也知道，如果整个系统的基态没有简并，那增加一个磁通量子将只能转移整数个电子电荷，所以要使得转移分数电荷成为可能，就必然意味着分数量子霍尔效应整个系统的基态有多重简并。和上一小节的论证一样，由于在洞中增加磁通的过程是绝热地进行的，所以在整个过程中系统都将保持处于基态。结合 Byers-Yang 定理我们就可以论证，增加一个磁通量子的前后，基态能量不变。然而，如果基态有多重简并，我们就无法论证过程前后系统依然处于同一个基态，因为它完全可能变换到另一个简并基态上去。以 $\nu = 1/3$ 的情形为例，这时候系统实际上有 3 个简并基态，分别记为 $|\Omega_1\rangle$、$|\Omega_2\rangle$、$|\Omega_3\rangle$，随着洞中磁通量子的不断增加，每增加一个磁通量子系统就从一个基态变换到另一个基态，最终系统实际上是在这 3 个基态上循环，

$$\xrightarrow{\text{insert}\Phi_0} |\Omega_1\rangle \xrightarrow{\text{insert}\Phi_0} |\Omega_2\rangle \xrightarrow{\text{insert}\Phi_0} |\Omega_3\rangle \xrightarrow{\text{insert}\Phi_0} |\Omega_1\rangle \xrightarrow{\text{insert}\Phi_0} \tag{9.12}$$

在霍尔薄片的洞中每增加一个磁通量子，都会有 $e^* = e/3$ 的电荷从一个边缘转移到另外一个边缘，连续增加 3 个磁通量子以后，就正好转移一个完整的电子电荷，而这时候系统的基态也刚好循环到初始时的同一个基态，这样一来就与 Laughlin 论证完全不矛盾了。

所以分数量子霍尔效应中必然存在分数电荷的准粒子激发，同时也存在基态简并。而这两件事情都非常不可思议，分数电荷的不可思议无需多言，这里的基态简并同样神奇。这是因为，分数量子霍尔效应中的这种基态简并和对称性没有任何关系！相反，它与我们考察的霍尔薄片中间有一个洞这种拓扑非平凡性倒是密切相关，它是一个拓扑基态简并。正是由于注意到分数量子霍尔效应这种与对称性无关的拓扑基态简并，文小刚老师才提出了重要的拓扑序概念。

9.2 Laughlin 波函数与分数统计

9.2.1 Laughlin 波函数

在第 8 章中我们讨论过，填充分数 $\nu = 1$ 的整数量子霍尔效应的基态波函数可以写成如下形式，

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = \prod_{1 \le i < j \le N} (w_i - w_j) \prod_{i=1}^{N} e^{-\frac{|w_i|^2}{4l_B^2}}. \tag{9.13}$$

式中 $N = N_\Phi$ 为单电子最低朗道能级的简并度。起初人们不觉得这个波函数有多大的重要性，然而出人意料的是，Laughlin 用他的神来之笔将这个波函数作了一点“简单”推广就成了人们理解分数量子霍尔效应的基础。Laughlin 提出，对于 $\nu = 1/m$，m 是一个奇数的分数量子霍尔效应，其基态的试探波函数可以取成如下形式

$$\Psi^{(m)} = \prod_{1 \le i < j \le N} (w_i - w_j)^m \prod_{i=1}^{N} e^{-\frac{|w_i|^2}{4l_B^2}}. \tag{9.14}$$

这就是著名的 Laughlin 波函数，式中的电子数目 N 满足

$$N_\Phi = m(N-1). \tag{9.15}$$

从而，当 N 很大时填充分数 $\nu = N/N_\Phi$ 的确趋于 $1/m$。但公式 (9.15) 并不是为了凑对这个填充分数而强加的，之所以要求 N 满足公式 (9.15)，是因为 $m(N-1)$ 是 Langhlin 波函数全纯多项式部分的最高次幂，而从第 4 章关于朗道能级的研究中我们已经知道，这个最高幂次决定了霍尔液滴的半径，它乘上 $2\pi l_B^2$ 就是霍尔液滴在平面上的面积，由于每 $2\pi l_B^2$ 面积对应一个磁通量子，所以这个幂次本身应该等于朗道能级的总磁通量子数 N_Φ。正因为公式 (9.15) 的成立有自然的理由，所以我们反过来知道，Laughlin 波函数的确解释了填充分数 $\nu = 1/m$。

9.2.2 等离子体类比

我们已经有了一个看起来不错的分数量子霍尔效应的试探波函数，也就是 Laughlin 波函数，下一步当然就是用这个多体波函数来计算各种物理量的期望值，但这并不是一件简单的事情，因为 Laughlin 波函数的变量数目 N 太多了，处理起来会很困难。本节我们将介绍一种处理这一问题的统计物理方法。

我们首先要计算的就是 Laughlin 波函数的归一化，为此我们把这个 Laughlin 波函数重新取作下面的无量纲形式，

$$\Psi_L^{(m)} = \prod_{i<j} \frac{(w_i - w_j)^m}{l_B^m} e^{-\Sigma_i \frac{|w_i|^2}{4l_B^2}}. \tag{9.16}$$

为了计算归一化，我们要对波函数的模方进行积分，不妨将波函数的模方记为 $P[w]$，

$$P[w] = |\Psi_L^{(m)}|^2 = \prod_{i<j} \frac{|w_i - w_j|^{2m}}{l_B^{2m}} e^{-\Sigma_i \frac{|w_i|^2}{2l_B^2}}. \tag{9.17}$$

为了将波函数归一化，我们需要计算 $Z=\int\prod_{i=1}^{N}d^2w_iP[w]$, 注意到复坐标 $w=x+iy$，所以式中的 $d^2w_i=dx_idy_i$。

计算 Z 的一个直观办法是将之看成是一个统计物理问题的配分函数。为此我们将 $P[w]$ 重新写成玻尔兹曼因子的形式

$$P[w]=e^{-\beta U[w]}, \tag{9.18}$$

式中

$$\beta U[w]=-2m\sum_{i<j}\log\Big(\frac{|w_i-w_j|}{l_B}\Big)+\frac{1}{2l_B^2}\sum_i|w_i|^2. \tag{9.19}$$

为了看出这个统计物理系统到底是什么系统，我们进一步将逆温度 β 取成如下形式

$$\beta=\frac{2}{m}. \tag{9.20}$$

如此一来这个统计物理系统的相互作用能 $U[w]$ 就应该取成

$$U[w]=-m^2\sum_{i<j}\log\Big(\frac{|w_i-w_j|}{l_B}\Big)+\frac{m}{4l_B^2}\sum_i|w_i|^2. \tag{9.21}$$

下面我们将会看到，这刚好是 N 个“电荷” $q=m$ 的带电粒子在一个背景负“电荷”上的等离子体势能。也即是说，计算 Laughlin 波函数的归一化以及物理量期望值的问题，等价于一个逆温度 $\beta=2/m$ 的两维等离子体系统的统计物理问题，所以我们把这种方法称作等离子体类比。

看出公式 (9.21) 描述的是一个等离子体系统的关键有两个。第一个关键是要注意到表达式 (9.21) 的第一项描述的是两维平面内多个电荷为 $q=m$ 的带电粒子之间的库伦相互作用势能。为此我们需要求解两维空间点电荷 q 所产生的电势 ϕ 的泊松方程

$$-\nabla^2\phi=2\pi q\delta(\mathbf{x}). \tag{9.22}$$

人们可以验证，这个方程的解为

$$\phi=-q\log\Big(\frac{r}{l_B}\Big). \tag{9.23}$$

注意到两个点电荷之间的电势能 $U=q\phi$，我们容易知道 (9.21) 式的第一项的确是两维带电粒子系统的相互作用电势能。

等离子类比的第二个关键是要知道表达式 (9.21) 的第二项描述的是 $q=m$ 的带电粒子在电荷密度 $\rho_0=-1/(2\pi l_B^2)$ 的负电荷背景中的电势能。为了证明这一点，我们只需要验证 $\frac{1}{4l_B^2}|w|^2$ 是 ρ_0 的背景负电荷所产生的背景电势，也即是验证如下泊松方程

$$-\nabla^2\Big(\frac{1}{4l_B^2}|w|^2\Big)=2\pi\rho_0=-\frac{1}{l_B^2}. \tag{9.24}$$

而这个验证相当容易。这里值得注意的是，背景负电荷密度 ρ_0 的大小正好是磁通量子的密度。

现在，我们可以利用这幅等离子体图像帮助我们思考了。比方说，由于背景电荷均匀分布，因此为了最小化系统的能量 $U[w]$，这些电荷为 $q=m$ 的粒子也应该均匀分布，其粒子数密度 n 应该刚好能够平衡背景电荷密度，使得整个系统呈电中性，即满足 $nm=|\rho_0|$，由此得到

$$n=\frac{1}{2\pi l_B^2}\frac{1}{m}. \tag{9.25}$$

这刚好是填充分数 $\nu=1/m$ 的分数量子霍尔系统的电子数密度。这是因为，等离子体类比中的 $q=m$ 带电粒子正好相应于原来 Laughlin 波函数中的电子。所以我们这里实际上是利用等离子体类比得到了正确的填充分数 $\nu=1/m$。

9.2.3 准粒子激发

前面我们说过，分数量子霍尔效应的神奇之处之一是，它有一些带分数电荷的准粒子激发。实际上，有两种不同电荷符号的准粒子激发，分别称作准空穴和准电子，准空穴带正电，准电子带负电。这一小节我们讨论如何利用 Laughlin 波函数描述这些准粒子激发。

准空穴 (Quasiholes)

描写一个位于复坐标 η 的准空穴的波函数 $\Psi_{qh}(w,\eta)$ 为

$$\Psi_{qh}(w;\eta)=\prod_{i=1}^{N}(w_i-\eta)\Psi_L^{(m)}(w), \tag{9.26}$$

式中 $\Psi_L^{(m)}(w)$ 就是原来 (9.16) 式给出的 Laughlin 波函数。从这个波函数我们很容易看出，电子在 η 位置出现的概率密度为 0。也即是说，我们在原来的多电子形成的液体中创生了一个空穴，所以称之为准空穴。值得注意的是，在这个描述中，准空穴是系统在 η 位置的一个缺陷，因此 η 本身并不是动力学变量，而是准空穴波函数的一个参数，动力学变量是电子的坐标 w_i。

准空穴的奇异性质之一是，它携带 $e^*=e/m$ 的分数电荷。注意，电子携带的是 $-e$ 的电荷，因此准空穴携带的是正电荷。

为了解释为什么准空穴的电荷是 e/m，我们不妨在 η 位置放 m 个相同的准空穴，描述这 m 个准空穴的波函数当然应该是

$$\prod_{i=1}^{N}(w_i-\eta)^m\Psi_L^{(m)}(w). \tag{9.27}$$

很明显的是，如果 η 是一个动力学变量，那这个波函数就是标准的 Laughlin 波函数，只不过 η 位置也有一个电子。但现在 η 不是动力学变量，而只是一个参数，因此 η 位置的这个电子实际上是缺失的。也即是说，η 位置的 m 个准空穴等价于 η 位置的一个电子缺失，因此一个准空穴就相当于 $1/m$ 个电子缺失。由于电子带 $-e$ 的电荷，电子缺失就相当于 $+e$ 的正电荷，因此一个准空穴就相当于一个 e/m 的正电荷。

以上我们描述的是单个准空穴，人们很容易将这个描述推广到 M 个准空穴，相应的波函数是

$$\Psi_{qh}(w;\eta_1,\eta_2,...,\eta_M)=\prod_{a=1}^{M}\prod_{i=1}^{N}(w_i-\eta_a)\Psi_L^{(m)}(w). \tag{9.28}$$

准电子 (Quasielectrons)

与准空穴相反，准电子是一种携带负分数电荷的准粒子，其电荷为 $-e/m$。写出准电子的试探波函数相对更难一些。

首先，我们注意到准空穴的存在会增加多电子波函数全纯多项式部分的幂次，从而扩大了霍尔液滴所占的总面积，降低了液滴的密度。这是因为空穴的存在将电子排挤到了更远一点的地方。而准电子则相反，它相当于在霍尔液滴上多鼓出了一个小包 (而不是创生一个空穴)，因此准电子的存在会升高液滴的密度，从而也就是要缩小霍尔液滴所占的面积，这就意味着准电子的存在应该降低波函数全纯多项式部分的幂次。基于此，Laughlin 猜测一个位于坐标原点的准电子将有如下波函数

$$\Psi_{qe}(w;0)=\Big[\prod_i\Big(\frac{\partial}{\partial w_i}\Big)\prod_{k<l}(w_k-w_l)^m\Big]e^{-\Sigma_i\frac{|w_i|^2}{4l_B^2}}. \tag{9.29}$$

很显然，偏导项的存在正确地将每一个 w_i 的幂次降低了 1，从也就将每一个电子的角动量量子数降低了 1，因此每一个电子都会朝原点挪动一个轨道，由于所有的单电子轨道中，实际被电子占据的只有 $\nu=1/m$，所以总的来说，原点处就会累积出 $-\nu e=-e/m$ 的电荷，这就是原点处的准电子的电荷。

稍微推广一下，如果准电子不是位于坐标原点，而是位于复坐标 η 处，注意到轨道角动量量子数的吻合，人们应该将准电子波函数写成

$$\Psi_{qe}(w;\eta)=\Big[\prod_i\Big(2l_B^2\frac{\partial}{\partial w_i}-\overline{\eta}\Big)\prod_{k<l}(w_k-w_l)^m\Big]e^{-\Sigma_i\frac{|w_i|^2}{4l_B^2}}. \tag{9.30}$$

式中 $\overline{\eta}$ 是 η 的复共轭，式中偏导前面的因子 $2l_B^2$ 来自于 $\overline{w}_i$ 的作用投影到最低朗道能级正好是 $2l_B^2\frac{\partial}{\partial w_i}$。

9.2.4 分数统计

$\nu=1/m$ 分数量子霍尔液体的准粒子激发不仅携带分数电荷，而且它们还满足第 8 章中学过的分数统计，从而是一种任意子。这一小节我们就是要具体研究 $\nu=1/m$ 分数量子霍尔液体准空穴激发的分数统计。不仅如此，我们还将从一个不同的角度证明准空穴的确携带分数电荷。这一节的推导要用到贝里相位，相关介绍我们放在了本章附录中，所以在阅读本小节之前，请读者先阅读本章的附录。

我们考察 M 个准空穴激发，这些准空穴分别具有复坐标 $\eta_1,\eta_2,...,\eta_M$，这些准空穴的波函数是

$$\Psi(w;\eta_1,\eta_2,...,\eta_M)=\prod_{a=1}^{M}\prod_{i=1}^{N}(w_i-\eta_a)\prod_{k<l}(w_k-w_l)^m e^{-\Sigma_i\frac{|w_i|^2}{4l_B^2}}. \tag{9.31}$$

准空穴的坐标 $\eta_1,\eta_2,...,\eta_M$ 是这个波函数的参数，不是动力学变量。但这个波函数没有归一化，为了将这个波函数归一化，我们需要计算波函数模方的积分 $Z[\eta]$

$$Z[\eta]=\int\prod_{i=1}^{N}d^2w_i|\Psi(w;\eta_1,\eta_2,...,\eta_M)|^2. \tag{9.32}$$

重复我们前面关于等离子体类比的讨论可以知道，$Z[\eta]$ 可以看成是一个逆温度 $\beta=2/m$ 的统计物理系统的配分函数，这个系统的能量 $U[w]$ 如下

$$U[w]=-m\sum_{i,a}\log\Big(\frac{|w_i-\eta_a|}{l_B}\Big)-m^2\sum_{i<j}\log\Big(\frac{|w_i-w_j|}{l_B}\Big)+\frac{m}{4l_B^2}\sum_i|w_i|^2. \tag{9.33}$$

等号右边的第一项来自于准空穴，很显然这第一项可以看成是电荷为 m 的带电粒子和电荷为 1 的带电粒子之间的库伦相互作用能。由于在等离子类比中电子相当于统计物理系统的电荷为 m(与电子的真实电荷没有什么关系) 的带电粒子，因此这就启发我们将准空穴看成是位置固定的电荷为 1 的带电粒子。

但如果真将准空穴看成电荷为 1 的带电粒子的话，那整个统计物理系统的能量就应该修正为 $U[w;\eta]$,

$$\begin{aligned}U[w;\eta]=&\frac{m}{4l_B^2}\sum_i|w_i|^2+\frac{1}{4l_B^2}\sum_a|\eta_a|^2\\&-m^2\sum_{i<j}\log\Big(\frac{|w_i-w_j|}{l_B}\Big)-m\sum_{i,a}\log\Big(\frac{|w_i-\eta_a|}{l_B}\Big)\\&-\sum_{a<b}\log\Big(\frac{|\eta_a-\eta_b|}{l_B}\Big).\end{aligned} \tag{9.34}$$

我们说了，准空穴相当于电荷为 1 的带电粒子，因此上式第一行右边的第二项就来自于这种带电粒子与背景负电荷的相互作用势能，而上式第三行的那一项则来源于这种带电粒子相互之间的库伦相互作用能。

(9.34) 式的这个统计物理系统和 (9.33) 式的统计物理系统的关系是，两者的配分函数间满足如下关系

$$\int\prod_{i=1}^{N}d^2w_ie^{-\beta U[w;\eta]}=Z[\eta]\cdot\exp\Big(\frac{2}{m}\sum_{a<b}\log\frac{|\eta_a-\eta_b|}{l_B}-\frac{1}{2m}\sum_a\frac{|\eta_a|^2}{l_B^2}\Big). \tag{9.35}$$

这一关系使得我们可以通过计算 (9.34) 式这个系统的配分函数来得到我们想要计算的配分函数 $Z[\eta]$。

而 (9.34) 式这个系统的配分函数可以有物理的办法来得到。这是因为，很显然 (9.34) 式描述的依然是一个等离子体系统，只不过在等离子体中引入了 M 个电荷为 1 的杂质 (相应于准空穴)。然而，等离子体有一种著名的德拜屏蔽效应，它告诉我们，当在等离子体中加入带正电的杂质时，由于静电吸引，等离子体的正负电荷 (在这里就是 N 个电荷为 $+m$ 的带电粒子以及背景负电荷) 会重新分布，正电荷会被排斥到离杂质更远的地方，这会使得靠近杂质负电荷更多，进而在超过某个长度尺度 l_D 的距离上，杂质的电荷将完全被等离子的负电荷屏蔽掉。也即是说，只要杂质间的距离超过 l_D，那么它们相互就完全感觉不

到彼此的存在，杂质的坐标在这个意义上就是自由而无关紧要的。因此 (9.34) 式对应的配分函数 $\int\prod_{i=1}^{N}d^2w_ie^{-\beta U[w;\eta]}$ 实际上是一个不依赖于 η 坐标的常数 C。从而根据 (9.35) 式我们马上就有

$$Z[\eta]=C\cdot\exp\Big(-\frac{1}{m}\sum_{a<b}\log\frac{|\eta_a-\eta_b|^2}{l_B^2}+\frac{1}{2m}\sum_a\frac{|\eta_a|^2}{l_B^2}\Big). \tag{9.36}$$

利用 Z 我们就可以把准空穴的波函数 $\Psi(w;\eta_1,\eta_2,...,\eta_M)$ 归一化，记未归一化的波函数 $\Psi(w;\eta_1,\eta_2,...,\eta_M)$ 为 $\langle w,\overline{w}|\eta\rangle$，则有

$$\langle\eta|\eta\rangle=\int\prod_{i=1}^{N}d^2w_i|\Psi(w;\eta_1,\eta_2,...,\eta_M)|^2=Z[\eta]. \tag{9.37}$$

记归一化以后的量子态为 $|\psi\rangle$, 则有

$$|\psi\rangle=Z^{-\frac{1}{2}}|\eta\rangle. \tag{9.38}$$

注意 $\Psi(w;\eta_1,\eta_2,...,\eta_M)$ 关于所有的 η 参数都是全纯的，所以量子态 $|\eta\rangle$ 关于所有的 η 参数都全纯。

注意到这种全纯性，我们就可以计算贝里联络的全纯分量 $\mathscr{A}_\eta(\eta,\overline{\eta})$，

$$\mathscr{A}_\eta(\eta,\overline{\eta})=i\langle\psi|\frac{\partial}{\partial\eta}|\psi\rangle=-i\frac{1}{2}Z^{-1}\frac{\partial Z}{\partial\eta}+iZ^{-1}\langle\eta|\frac{\partial}{\partial\eta}|\eta\rangle. \tag{9.39}$$

由于 $|\eta\rangle$ 关于 η 全纯，所以 $\langle\eta|$ 关于 η 反全纯，所以 $\langle\eta|\frac{\partial}{\partial\eta}|\eta\rangle=\frac{\partial}{\partial\eta}\langle\eta|\eta\rangle=\frac{\partial Z}{\partial\eta}$，代入 (9.39) 式就可以得到

$$\mathscr{A}_\eta(\eta,\overline{\eta})=i\frac{1}{2}\frac{\partial\log Z}{\partial\eta}. \tag{9.40}$$

贝里联络的反全纯分量 $\mathscr{A}_{\overline{\eta}}(\eta,\overline{\eta})$ 当然是全纯分量的复共轭，所以我们也有

$$\mathscr{A}_{\overline{\eta}}(\eta,\overline{\eta})=i\langle\psi|\frac{\partial}{\partial\overline{\eta}}|\psi\rangle=-i\frac{1}{2}\frac{\partial\log Z}{\partial\overline{\eta}}. \tag{9.41}$$

进一步代入 Z 的计算结果 (9.36) 就可以得到

$$\begin{aligned}\mathscr{A}_{\eta_a}&=-i\frac{1}{2m}\sum_{b\neq a}\frac{1}{\eta_a-\eta_b}+i\frac{1}{4m}\frac{\overline{\eta}_a}{l_B^2}\\ \mathscr{A}_{\overline{\eta}_a}&=i\frac{1}{2m}\sum_{b\neq a}\frac{1}{\overline{\eta}_a-\overline{\eta}_b}-i\frac{1}{4m}\frac{\eta_a}{l_B^2}.\end{aligned} \tag{9.42}$$

当然这个结果仅当准空穴间距离不太短，大于德拜屏蔽长度 l_D 时才成立。下面我们就用这个结果来计算分数电荷与分数统计。计算的关键在于注意到，当将第 a 个准空穴沿着闭合路径 C 缓慢运动一圈以后，波函数将多出一个贝里相位 $e^{i\gamma_C}$，

$$\gamma_C=\oint_C\Big(\mathscr{A}_{\eta_a}d\eta_a+\mathscr{A}_{\overline{\eta}_a}d\overline{\eta}_a\Big). \tag{9.43}$$

分数电荷

为了计算准粒子的分数电荷，我们让第 a 个准空穴的环绕路径 C **不包围**任何其它准空穴。这时候根据复变函数的围道积分，(9.42) 各贝里联络表达式的第一项对于环路积分 (9.43) 均没有贡献。对环路积分 (9.43) 有贡献的是贝里联络表达式的第二项。很显然，对于这些项，我们有

$$\gamma_C = \frac{1}{4ml_B^2}\oint_C \left(i\overline{\eta}_a d\eta_a - i\eta_a d\overline{\eta}_a\right) = \frac{S}{ml_B^2}. \tag{9.44}$$

式中 S 是回路 C 所包围的面积，代入 $l_B^2 = \hbar/(eB)$, 就可以得到

$$\gamma_C = \frac{1}{m}(2\pi\Phi/\Phi_0). \tag{9.45}$$

Φ 为回路 C 所包围的磁通，Φ_0 为磁通量子。很显然，这个贝里相位来源于准空穴在磁场中环绕一圈的 Aharnaov-Bohm 效应，不妨将之重记为 γ_{AB}。如果准空穴的电荷为 e，那我们知道这个相位应该是 $(2\pi\Phi/\Phi_0)$，但现在算出来的相位是它的 $1/m$，这就说明准空穴的带电量为

$$e^* = e/m. \tag{9.46}$$

这正是准空穴的分数电荷。

分数统计

为了计算分数统计，我们让 a 准空穴的运动回路 C 中包围另一个准空穴 b，并假设 b 是回路 C 唯一包围的准空穴。进一步，我们假设回路 C 是围绕准空穴 b 逆时针转了一圈。现在，(9.42) 各贝里联络表达式的两项均对环路积分 (9.43) 有贡献，当然第二项的贡献是我们刚才已经计算过了的 Aharnaov-Bohm 相位 γ_{AB}, 现在的关键是要计算第一项的贡献 γ，即

$$\gamma_C = \gamma_{AB} - i\frac{1}{2m}\oint_C \left(\frac{d\eta_a}{\eta_a - \eta_b} - \frac{d\overline{\eta}_a}{\overline{\eta}_a - \overline{\eta}_b}\right) = \gamma_{AB} + \gamma. \tag{9.47}$$

由围道积分我们很容易计算出这个相位 γ 为

$$\gamma = 2\pi/m. \tag{9.48}$$

贝里相位 γ_C 中的 γ_{AB} 当然和全同粒子统计没有任何关系，但剩下的这个 γ 却没有任何其它解释，只能理解为起源于准空穴的全同粒子统计。即是说，当我们将一个准空穴逆时针围绕另一个准空穴转一圈时，波函数应该多出如下统计相因子

$$e^{i\gamma} = e^{i2\pi/m}. \tag{9.49}$$

这意味着当我们将两个全同准空穴逆时针交换位置时 (相当于一个准空穴围绕另一个转半圈)，应该多出一个统计相因子 $e^{i\pi/m}$。对于分数量子霍尔效应，这既不是玻色统计，也不是费米统计，而是任意子统计!

如果我们把 n 个准空穴绑在一起形成一个整体，并让一个这样的物体和另一个这样的物体逆时针交换位置，则由于每一对准空穴的位置交换都要贡献一个 $e^{i\pi/m}$，现在这两个物体中的准空穴可以有 n^2 种组对可能，所以总的统计相因子就应该是 $e^{in^2\pi/m}$。特别的，如果取 $n=m$，最终就会得到统计相因子 $e^{im\pi}$，由于 m 是一个奇数，所以最终的相位等于 -1，是费米统计。这就意味着 m 个准空穴绑在一起反而成了一个普通的费米子！

9.3 附录：贝里相位

贝里相位是现代量子力学的一个重要发现，在与量子物理相关的许多物理学方向上都有重要应用，在这个附录中，我们将对它进行一个简单介绍。

考虑一个量子系统，假定其哈密顿量 H 依赖于一些控制参数 $(\lambda_1,\lambda_2,...)$，记为 $H(\lambda)$，并记参数空间为 $\mathscr{M}$。现在假设我们很缓慢地调节各控制参数，使得它在参数空间画出一条路径 C。再假设系统处在某个能量本征态 $|\psi(\lambda)\rangle=|u_n(\lambda)\rangle$ 上 ($|u_n(\lambda)\rangle$ 表示第 n 个能量本征态)，并且假设在路径 C 所对应的各参数点上，相应的能级没有产生交叉。则，如果我们对参数的调节足够缓慢的话，那系统将会始终待在瞬时本征态 $|\psi(\lambda(t))\rangle$ 上。然而我们也知道，任何量子态都可以相差一个相位因子，因此瞬时本征态 $|\psi(\lambda(t))\rangle$ 依然可以相差一个待定的相位因子。贝里发现，在一些物理情形中，即使扣除动力学相位，这个相位因子也不能为零，称作贝里相位。

系统的瞬时能量本征方程是

$$H(\lambda)|\psi(\lambda)\rangle=E(\lambda)|\psi(\lambda)\rangle. \tag{9.50}$$

式中 $E(\lambda)=E_n(\lambda)$ 为这第 n 个能量本征态的本征能量。根据上一段所说，只要对参数的调节进行得足够缓慢，我们就可以将系统随着时间演化的量子态写成如下形式

$$|\Psi(t)\rangle=e^{i\gamma(t)}|\psi(\lambda(t))\rangle. \tag{9.51}$$

式中的 $e^{i\gamma(t)}$ 就是上一段所说的待定相位因子。

将时间演化态 $|\Psi(t)\rangle$ 代入时间演化的薛定谔方程 $i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle=H(\lambda(t))|\Psi(t)\rangle$，我们就可以得到

$$[-\hbar\dot{\gamma}+i\hbar\frac{\partial}{\partial t}]|\psi(\lambda(t))\rangle=E(\lambda(t))|\psi(\lambda(t))\rangle. \tag{9.52}$$

用左矢 $|\psi(\lambda(t))\rangle$ 去和这个方程作内积，就可以得到

$$\dot{\gamma}=-E(\lambda(t))/\hbar+i\langle\psi(\lambda(t))|\frac{\partial}{\partial t}|\psi(\lambda(t))\rangle. \tag{9.53}$$

将这个方程对时间积分，即有

$$\begin{aligned}\gamma(t_f)-\gamma(t_i)&=-\int_{t_i}^{t_f}E(\lambda(t))dt/\hbar+i\int_{t_i}^{t_f}dt\langle\psi(\lambda(t))|\frac{\partial}{\partial t}|\psi(\lambda(t))\rangle\\&=-\int_{t_i}^{t_f}E(\lambda(t))dt/\hbar+i\int_C d\lambda_i\langle\psi(\lambda)|\frac{\partial}{\partial\lambda_i}|\psi(\lambda)\rangle.\end{aligned} \tag{9.54}$$

这个最终表达式的第一项就是普通的动力学演化相位，第二项称之为几何相位，因为它只依赖于参数空间的几何路径 C。在写这个最终表达式时，我们应用了求和约定，即默认对重复出现的 i 指标求和，实际上，本附录我们对 i,j 指标始终默认求和约定。

特别的，我们可以调节参数在参数空间走一圈，然后又回到起始位置，那这时候上面的几何相位就可以写成

$$\gamma_C = i\oint_C d\lambda_i \langle\psi(\lambda)|\frac{\partial}{\partial\lambda_i}|\psi(\lambda)\rangle. \tag{9.55}$$

这就是贝里相位，贝里发现它不仅可能非零，而且甚至很重要。

习惯上，人们通常将 1 形式 $id\lambda_i\langle\psi(\lambda)|\frac{\partial}{\partial\lambda_i}|\psi(\lambda)\rangle$ 记为 $\mathscr{A}$，称之为贝里联络，

$$\begin{aligned}\mathscr{A} &= \mathscr{A}_i d\lambda_i = id\lambda_i\langle\psi(\lambda)|\frac{\partial}{\partial\lambda_i}|\psi(\lambda)\rangle. \\ &= i\langle\psi(\lambda)|d|\psi(\lambda)\rangle.\end{aligned} \tag{9.56}$$

式中 $\mathscr{A}_i = i\langle\psi|\frac{\partial}{\partial\lambda_i}|\psi\rangle$ 是贝里联络的第 i 分量。利用贝里联络我们就可以将贝里相位的计算公式重写为

$$\gamma_C = \oint_C \mathscr{A} = \int_S d\mathscr{A}. \tag{9.57}$$

式中 $d\mathscr{A}$ 表示在参数空间对 $\mathscr{A}$ 进行外微分，S 就是回路 C 在参数空间所包围的曲面，这个式子的第二个等号我们当然是用了微分形式的斯托克斯公式。

通常把 $\mathscr{F}_n = d\mathscr{A}$ 称作贝里曲率。利用公式 (9.56) 我们容易得到

$$\mathscr{F}_n = i\langle d\psi(\lambda)|\wedge|d\psi(\lambda)\rangle = i\langle du_n(\lambda)|\wedge|du_n(\lambda)\rangle. \tag{9.58}$$

式中 $|d\psi\rangle = d|\psi\rangle$, $\langle d\psi| = d\langle\psi| = |d\psi\rangle^\dagger$, 式中的第二个等号我们已经将 $|\psi(\lambda)\rangle$ 取作第 n 能量本征态 $|u_n(\lambda)\rangle$ 了。我们可以将哈密顿量对参数的微分 $dH = \partial_i H d\lambda_i$ 看作是一个微扰 (式中 $\partial_i H = \frac{\partial}{\partial\lambda_i}H(\lambda)$), 进而可以由定态微扰论中波函数的一阶微扰修正公式得到

$$|du_n(\lambda)\rangle = \sum_{m\neq n}\frac{\langle u_m|dH|u_n\rangle}{E_n - E_m}|u_m\rangle. \tag{9.59}$$

将之代入贝里曲率公式 (9.58)，就可以得到

$$\begin{aligned}\mathscr{F}_n &= i\sum_{m\neq n}\frac{\langle u_n|dH|u_m\rangle\wedge\langle u_m|dH|u_n\rangle}{(E_n - E_m)^2} \\ &= i\sum_{m\neq n}\frac{\langle u_n|\partial_i H|u_m\rangle\langle u_m|\partial_j H|u_n\rangle}{(E_n - E_m)^2}d\lambda_i\wedge d\lambda_j.\end{aligned} \tag{9.60}$$

9.4 参考文献

关于本章的内容，我们推荐两份扩展读物：

1. Steven H. Simon 的在线讲义 “Topological Quantum”，
http://oxfordtopquantum.tiddlyspot.com/

2. David Tong 的在线讲义“Lectures on the Quantum Hall Effect”, http://www.damtp.cam.ac.uk/user/t

10. 含时问题与散射问题

本章将系统地讨论量子力学系统的含时演化问题。并利用含时理论的框架建立散射的一般性理论。

薛定谔方程是一个关于量子态如何随着时间演化的方程，但是前面我们关心得更多的是定态问题，这一节我们将要系统的考察量子态的时间演化问题。如果系统的哈密顿算符 H 本身不显含时间 t，那么量子态的时间演化问题可以这样来求解，首先我们把初始 $t=0$ 时刻的量子态 $|\psi(0)\rangle$ 分解成定态 $|E_n\rangle$ 的叠加的形式，即 $|\psi(0)\rangle=\sum_n c_n|E_n\rangle$，式中 $|E_n\rangle$ 是 H 的本征态，本征值为 E_n，即 $H|E_n\rangle=E_n|E_n\rangle$。那么根据薛定谔方程，$t$ 时刻的量子态 $|\psi(t)\rangle$ 将由下式给出，

$$|\psi(t)\rangle=\exp\left(-i\frac{H}{\hbar}t\right)|\psi(0)\rangle=\sum_n c_n e^{-i\frac{H}{\hbar}t}|E_n\rangle=\sum_n c_n e^{-i\frac{E_n}{\hbar}t}|E_n\rangle. \tag{10.1}$$

由此我们就能求出 t 时刻的量子态，由这个 $|\psi(t)\rangle$ 态，我们就能求出 $|\psi(0)\rangle$ 态在 t 时刻跃迁到某个 $|\phi\rangle$ 态的概率 (也就是 t 时刻我们在 $|\phi\rangle$ 态上测量到系统的概率)，这个概率由 $|\langle\phi|\psi(t)\rangle|^2$ 给出。

但，在很多时候，H 的本征态不容易求解出来，或者甚至 H 本身就显含时间 t。这时候刚才说到的方法应用起来就不方便了。但如果这时候 H 能够分解成 H_0+V 的形式，其中 V 可以显含时间 t，但 H_0 不显含 t，并且其本征态 $|n\rangle$ 相对容易知道，而且我们很容易将初始时刻的系统准备在 H_0 的某个本征态 $|n\rangle$ 上，同时也很容易对末尾时刻 H_0 的本征态进行测量，那么在实验上我们就能测量系统从初始时刻的某个 $|n\rangle$ 态跃迁到末尾时刻的某个 $|m\rangle$ 态的概率 $p_{n\to m}$。这样的例子有很多，比如说原子的光谱就属于这种跃迁的例子，对于原子和电磁场相互作用的系统，整个包含了相互作用的哈密顿量 H 并不好求解，但忽略掉与电磁场的相互作用以后，原子本身的哈密顿量 H_0 是比较容易了解的，那就是原子的能级，而原子的光谱对应的就是原子从一个能级 $|n\rangle$ 到另一个能级 $|m\rangle$ 的跃迁，光谱的强度就和跃迁概率 $p_{n\to m}$ 密切相关。那么，在这些情况下，有没有什么方便的方法计算这个跃迁概率呢?

这一章我们将发展两种计算这一跃迁概率的重要方法。第一种方法就是微扰论的方法 (通常将这一方法称作含时微扰论)，也就是当相互作用 $V(t)$ 可以看成是一个微扰，并且这个微扰有效起作用的时间不是过于长的时候适用的方法。应用这种方法，我们研究了原子对光的辐射和吸收过程，并利用所得到的结果解释了激光的核心原理。第二种方法是时间尺度可以任意长，甚至 V 也不必是一个微扰，但是系统近似处于两个状态 $|n\rangle$ 态和 $|m\rangle$ 态共振情形下的求解方法，通过应用这种方法，我们解释了著名的拉比振荡现象。

另一方面，自从卢瑟福最早使用 α 粒子去轰击金箔，并进而推断出原子的内部结构以后，物理学家就开始学会了用一些入射粒子 (包括入射光) 去探测研究对象这一重要手段。比方说，在凝聚态物理中，人们常常用 X 射线在晶体上的散射来了解晶体结构，或者用中子散射去研究物质结构和微观动力学性质。再比方说，在高能粒子物理中，人们常常将多个粒子加速，并让它们相互碰撞 (也就是相互散射)，以此推断微观粒子相互作用的基本规律。因此，在这一章的最后两节，我们将利用含时微扰论发展一个关于多粒子散射的一般性量子理论。这是因为多粒子散射常常可以自然地满足含时微扰论的适用条件。我们假定有多个粒子 (只有一个入射粒子的情况当然也包含在内) 从不同的方向入射，在无穷远过去，即 $t\to-\infty$ 时，入射粒子离其目标的距离，以及它们相互之间的距离都比较遥远，因此所有的相互作用都可以忽略，整个系统可以看成由哈密顿量 H_0 描述，它是多个入射

自由粒子的哈密顿量之和。随着粒子离散射对象越来越近，或者不同粒子之间的距离越来越近，相互作用 V(V 不显含 t) 才开始变得不可忽略，从而就发生了粒子的散射，但是之后，出射粒子开始从散射区域飞离，随着出射粒子的相互飞离，它们之间的距离又开始越来越大，从而相互作用再次可以忽略，系统再次可以由自由粒子的哈密顿量之和 H_0 来描述。可见，只要相互作用 V 不太强，那么它就可以看成是对多个自由粒子系统的一种微扰，同时，只要相互作用力的力程不是一个长程力 (其实即使是长程力，比如库仑力，我们有时候也可以将之看成是短程力的一种极限)，那么相互作用 V 有效起作用的时间就不会过于长。因此，这时候多粒子的散射问题就可以用含时微扰论来处理。我们就可以用含时微扰论来计算无穷远过去的入射自由粒子初态，在无穷远将来 (即 $t\to\infty$) 跃迁为出射自由粒子末态的跃迁概率。

10.1 量子力学含时微扰论

10.1.1 时间演化算符与含时微扰

首先让我们来回顾一下时间演化算符的概念。对于一个哈密顿算符为 H 的量子系统，其量子态 $|\psi(t)\rangle$ 随时间的演化满足薛定谔方程

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle. \tag{10.2}$$

但很多时候用时间演化算符来描述量子态的演化是更加方便的，所谓的时间演化算符就是一组形如 $U(t,t_0)$ 这样的幺正算符，它描述的是量子系统从 t_0 时刻到 t 时刻的演化，也即是说，对于 t_0 时刻的任何一个量子态 $|\psi(t_0)\rangle$，它演化到 t 时刻将变成量子态 $|\psi(t)\rangle$，

$$|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle. \tag{10.3}$$

假设我们让系统先从 t_0 时刻演化到 t_1 时刻，接着再从 t_1 时刻演化到 t 时刻，那么这个两阶段演化的总效果当然就是使得系统从 t_0 时刻演化到了 t 时刻，因此容易得到时间演化算符的一条基本性质，即

$$U(t,t_0) = U(t,t_1)U(t_1,t_0). \tag{10.4}$$

另外，从 t 时刻到 t 时刻本身，系统当然没有发生任何演化，因此 $U(t,t)=1$。把时间演化算符的定义 (10.3) 代入薛定谔方程 (10.2)，我们就可以得到时间演化算符所满足的微分方程

$$i\hbar\frac{d}{dt}U(t,t_0) = HU(t,t_0), \tag{10.5}$$

这个方程的初始条件就是 $U(t_0,t_0)=1$。

假设我们原来有一个了解得很清楚的系统，其哈密顿算符为 H_0(假定 H_0 不显含时间)。所谓了解得很清楚意思就是我们完全知道 H_0 的所有本征态 $\{|n\rangle, n=1,2,3...\}$ 和相应的本征值 E_n(注意，从这里开始，在这一章中 E_n 都表示 H_0 的本征值，而不是 H 的本征值)，也即是说

$$H_0|n\rangle = E_n|n\rangle. \tag{10.6}$$

这里的 $|n\rangle$ 当然是不含时间的，完整的随时间演化的量子态 $|n,t\rangle$ 应该是 $|n,t\rangle = e^{-i\frac{E_n}{\hbar}t}|n\rangle$。注意到定态薛定谔方程 (10.6), 我们显然也可以将 $|n,t\rangle$ 重写成

$$|n,t\rangle = \exp\left(-i\frac{H_0}{\hbar}t\right)|n\rangle. \tag{10.7}$$

现在，假设在某个从 t_0 到 t 的时间段，系统 H_0 受到了一个比较复杂的微扰 $V(t)$, $V(t)$ 一般来说显含时间。因此，扰动作用以后的系统哈密顿量就应该是

$$H(t) = H_0 + V(t), \tag{10.8}$$

相应的时间演化算符就是我们上面定义的 $U(t,t_0)$。因此，假设在扰动作用之前我们有一个 H_0 的定态 $|n,t_0\rangle$，在扰动的作用下将它按照 $U(t,t_0)$ 演化到 t 时刻，我们就有一个 t 时刻的量子态 $|\psi_n(t)\rangle$, $|\psi_n(t)\rangle = U(t,t_0)|n,t_0\rangle$, 我们想问的是，在 t 时刻 (假设 t 时刻扰动已经结束了) 我们在 $|m,t\rangle$ 态上测到系统的概率是多少？由量子力学的基本原理，这个概率 $p_{n\to m}$ 显然由 $p_{n\to m} = |\langle m,t|\psi_n(t)\rangle|^2 = |\langle m,t|U(t,t_0)|n,t_0\rangle|^2$ 给出。在物理上，$p_{n\to m}$ 就是在微扰 V 的作用下，系统从 $|n,t_0\rangle$ 态跃迁到 $|m,t\rangle$ 态的概率，相应的跃迁概率幅 (简称跃迁幅) 就是 $\langle m,t|U(t,t_0)|n,t_0\rangle$。应用表达式 (10.7)，我们很容易将这个跃迁幅重新表达为

$$\langle m,t|U(t,t_0)|n,t_0\rangle = \langle m|\exp\left(i\frac{H_0}{\hbar}t\right)U(t,t_0)\exp\left(-i\frac{H_0}{\hbar}t_0\right)|n\rangle, \tag{10.9}$$

注意这个表达式右边的 $|n\rangle, |m\rangle$ 都不依赖于时间。习惯上，人们通常定义一个新的时间演化算符 $U_I(t,t_0)$, 它的定义是

$$U_I(t,t_0) = \exp\left(i\frac{H_0}{\hbar}t\right)U(t,t_0)\exp\left(-i\frac{H_0}{\hbar}t_0\right). \tag{10.10}$$

$U_I(t,t_0)$ 也是一个幺正算符，它常常被称为相互作用绘景下的时间演化算符，人们很容易验证它满足时间演化算符的基本性质 $U_I(t,t_0) = U_I(t,t_1)U_I(t_1,t_0)$, 以及 $U_I(t_0,t_0) = 1$。很显然，根据 (10.9) 式可以知道，跃迁概率 $p_{n\to m}$ 可以由下式给出

$$p_{n\to m} = |\langle m|U_I(t,t_0)|n\rangle|^2. \tag{10.11}$$

因此，为了计算跃迁概率，我们需要进一步研究时间演化算符 $U_I(t,t_0)$。利用 $U_I(t,t_0)$ 的定义式 (10.10)，以及 $U(t,t_0)$(注意不是 $U_I(t,t_0)$) 所满足的方程 (10.5), 我们容易得到微分方程

$$i\hbar\frac{d}{dt}U_I(t,t_0) = V_I(t)U_I(t,t_0), \tag{10.12}$$

式中 (注意 H_0 与 $V(t)$ 一般来说不对易)

$$V_I(t) = \exp\left(i\frac{H_0}{\hbar}t\right)V(t)\exp\left(-i\frac{H_0}{\hbar}t\right). \tag{10.13}$$

$V_I(t)$ 通常被称作相互作用绘景中的微扰算符。

将方程 (10.12) 对时间积分，并利用 $U_I(t_0,t_0)=1$，就可以得到

$$U_I(t,t_0)=1-\frac{i}{\hbar}\int_{t_0}^{t}V_I(t')U_I(t',t_0)dt', \tag{10.14}$$

将这个方程进行迭代就可以得到 $U_I(t,t_0)$ 关于微扰 V_I 的级数展开

$$\begin{aligned}U_I(t,t_0)=1+(-\frac{i}{\hbar})\int_{t_0}^{t}V_I(t_1)dt_1+(-\frac{i}{\hbar})^2\int_{t_0}^{t}dt_2\int_{t_0}^{t_2}dt_1V_I(t_2)V_I(t_1)\\+(-\frac{i}{\hbar})^3\int_{t_0}^{t}dt_3\int_{t_0}^{t_3}dt_2\int_{t_0}^{t_2}dt_1V_I(t_3)V_I(t_2)V_I(t_1)+\ldots\end{aligned} \tag{10.15}$$

这个级数 (10.15) 通常称作戴森级数。根据公式 (10.11)，为了用微扰展开的办法计算出跃迁概率 $p_{n\to m}$, 最主要的就是要逐级地将戴森级数在 $|m\rangle$ 态和 $|n\rangle$ 态上的矩阵元计算出来。

特别的，如果末态 $|m\rangle$ 与初态 $|n\rangle$ 不同，即 $m\neq n$, 则由正交性，戴森级数的 0 阶项 $\langle m|1|n\rangle=0$,为了计算跃迁概率,我们首先要计算戴森级数的一阶微扰项,$(-\frac{i}{\hbar})\int_{t_0}^{t}dt_1\langle m|V_I(t_1)|n\rangle=(-\frac{i}{\hbar})\int_{t_0}^{t}dt_1e^{i\omega_{mn}t_1}\langle m|V(t_1)|n\rangle$, 式中 $\omega_{mn}=(E_m-E_n)/\hbar$, 而且我们利用了 $V_I(t)$ 的定义式 (10.13)。因此这时候，在一阶微扰近似的意义上，跃迁概率 $p_{n\to m}$ 为

$$p_{n\to m}=\frac{1}{\hbar^2}\left|\int_{t_0}^{t}dt_1e^{i\omega_{mn}t_1}\langle m|V(t_1)|n\rangle\right|^2. \tag{10.16}$$

值得说明的是，微扰 $V(t)$ 具体是何时加上的，又是何时结束的，这些其实并不是要点，要点是在 t_0 时刻我们能准备一个 H_0 的本征态 $|n\rangle$ 作为初态, 并在最后的 t 时刻我们能测量 H_0 的本征态 $|m\rangle$，以此作为末态。这两个要求只需要我们在初末两个时刻都能测量 H_0 的本征态就能实现。前文论述中之所以假定微扰只在 $[t_0,t]$ 的时间范围之内起作用，是因为在这种情况下，系统 t_0 时刻 (微扰起作用之前) 和 t 时刻 (微扰结束以后) 的哈密顿量都是 H_0，因此我们只需要测量系统的能量就能实现测量 H_0 的本征态的要求。但也有很多情况，即使微扰一直都存在，但 H_0 本身却总是系统的一个很容易测量的物理量，在这种情况下，我们无论何时都能方便地测量 H_0 的本征态，因此就无需假定微扰只在 $[t_0,t]$ 的时间范围之内起作用了。因此后文我们将不再强调 $V(t)$ 起作用的时间范围，而是默认这两种情况 (V 只在 $[t_0,t]$ 之内起作用，或者我们总能方便地测量 H_0) 必定有一种得以满足。

10.1.2 恒定微扰

作为一个重要的例子，这一节我们将把上面发展的含时微扰论应用于恒定微扰的情形。所谓的恒定微扰，我们指的是微扰算符 V 不显含时间 t。我们将把系统初态记为 $|i\rangle$, 跃迁末态记为 $|f\rangle$, 当然 $|i\rangle,|f\rangle$ 都是 H_0 的本征态，本征值分别为 E_i,E_f。我们将区分两种情形，其一是 $|i\rangle$ 和 $|f\rangle$ 为两个离散能级的情形，另一种情形是，末态 $|f\rangle$ 属于一个连续谱的情形。

离散能级情形

对于离散能级的情况，为了方便，我们常常取初始时刻 t_0 为 0，即 $t_0=0$，取末尾时刻为 t 时刻。根据一阶含时微扰论的公式 (10.16), 为了计算从 $|i\rangle$ 态到 $|f\rangle$ 态的跃迁概率

$p_{i\to f}$, 我们需要计算 $\int_0^t dt_1 e^{i\omega_{fi}t_1}\langle f|V|i\rangle$(式中 $\omega_{fi}=(E_f-E_i)/\hbar$), 由于 V 不显含时间，很容易得到计算结果 $\langle f|V|i\rangle\frac{e^{i\omega_{fi}t}-1}{i\omega_{fi}}=\langle f|V|i\rangle e^{i\frac{1}{2}\omega_{fi}t}\frac{\sin(\omega_{fi}t/2)}{\omega_{fi}/2}$。因此跃迁概率 $p_{i\to f}$ 将是

$$p_{i\to f}=\frac{1}{\hbar^2}|\langle f|V|i\rangle|^2\left[\frac{\sin(\omega_{fi}t/2)}{\omega_{fi}/2}\right]^2. \tag{10.17}$$

其中对于给定的时间 t，$\left[\frac{\sin(\omega_{fi}t/2)}{\omega_{fi}/2}\right]^2$ 作为 ω_{fi} 的函数图像如图 (10.1) 所示。从图中可以看

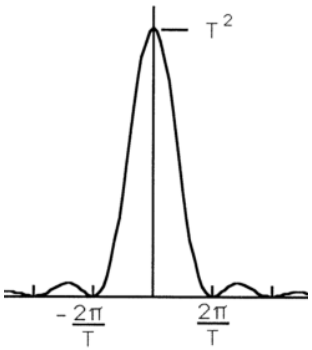

Figure 10.1: $F(\omega_{fi})=\left[\frac{\sin(\omega_{fi}T/2)}{\omega_{fi}/2}\right]^2$ 的函数图像.

出来，当 $\omega_{fi}=0$ 时 (即初末态能量相等时)，跃迁概率达到最大，为 $p_{i\to f}=\frac{1}{\hbar^2}|\langle f|V|i\rangle|^2t^2$。但是当然，不同态之间的跃迁概率一定是小于 1 的，并且一阶微扰近似要有效的话，这个跃迁概率还得远小于 1，因此我们可以得到离散能级时一阶含时微扰近似必须满足的条件，

$$|\langle f|V|i\rangle|t\ll\hbar. \tag{10.18}$$

末态为连续谱的情形

末态 $|f\rangle$ 属于一个连续谱情形的不同之处在于，这时候末态指标 f 实际上是一个连续指标，因此按照公式 (10.16) 求出来的 $p_{i\to f}$ 实际上是关于末态的一个概率密度，而不是概率，因此为了得到概率，我们需要对末态谱进行适当的积分。这个对末态谱的积分当然可以直接在公式 (10.17) 的基础上进行，只不过一般来说我们无法得到一个解析的积分结果。但是，对于实际应用中需要计算跃迁到末态连续谱的情形，情况可以简化。这是因为，末态连续谱和离散能级不同，正如后文我们将要研究的，离散能级即使在 $\omega_{fi}=0$ 时，它在一个充分长时间的微扰作用之下依然有可能发生初末态之间的振荡行为，这时候微扰论的分析是不成立的，这就是为什么对于离散能级我们的一阶微扰公式需要满足条件 (10.18) 的原因。但是，当末态是连续谱时，只要这个连续谱的宽度足够，那么系统一旦离开初态跃迁到末态连续谱以后，它将不会和原来的态发生振荡 (在本章的最后一节，我们将给出这一结论的证明)，这一结论的自然推论就是，只要连续谱足够宽，连续谱中的态不会和任何其它态发生振荡，因为任何态只要跃迁到连续谱就不能再跃迁回来了。形象地说，连续谱就好比一个泥潭，量子态跃迁进去就被“粘住”回不来了。因此，对于末态是连续谱的情形，在实际中我们总是可以测量一个充分长的时间段中的跃迁概率，也即是说，这时候我们真正关心的其实总是 $t\to\infty$ 的情形，这时候结果就可以进一步简化。

为了强调末态指标的连续性，我们将末态指标 f 改写成 α_f。α_f 只是一个示意性的指标，一般来说它可能包含多个连续指标，我们假定不同末态归一化到标准的 δ 函数上，即对于两个不同末态 $|\alpha_f\rangle$ 和 $|\alpha_f'\rangle$ 我们有

$$\langle\alpha_f'|\alpha_f\rangle = \delta(\alpha_f - \alpha_f'). \tag{10.19}$$

因此，为了得到跃迁概率，我们需要把按照公式 (10.16) 计算出来的概率密度 $p_{i\to\alpha_f}$ 对 α_f 进行一个合适的积分

$$\int_{D_f} p_{i\to\alpha_f} d\alpha_f, \tag{10.20}$$

式中 $d\alpha_f$ 可以理解成一个 α_f 指标微元内所包含的末态数目，而积分区域 D_f 的选取则和我们具体如何测量跃迁概率密切相关。

比方说，在实际测量中，我们的仪器对末态能量会有一个分辨率 Δ, 那么积分区域 D_f 中就必须包含一个以能量 E_f 为中心宽度为 Δ 的能谱 (按照假定这是一个连续谱) 区间 $\Delta(E_f)$，因此这就需要我们将末态指标从归一化到 δ 函数的指标 α_f 变换到以连续能量 E_f 为代表的指标。当然能量指标可能还不足以区分不同的末态，这时候我们就还需要一些额外的指标 χ_f。这种末态指标的变换会满足

$$d\alpha_f = \rho(E_f,\chi_f) dE_f d\chi_f, \tag{10.21}$$

其中指标变换的雅可比行列式 $\rho(E_f,\chi_f)$ 又称为末态态密度，因为它表示单位能量区间和单位 χ 指标区间之内的末态数目。

比方说，如果 H_0 是一个自由粒子哈密顿量，即 $H_0 = \frac{\mathbf{p}^2}{2m}$，因此末态是一个自由粒子，它可以用归一化到 δ 函数的动量本征态 $|\mathbf{p}\rangle$ 来标记 (它当然也是 H_0 的本征态), 这时候末态数目就是 $d^3\mathbf{p}$。变换到动量空间的球坐标，就有 $d^3\mathbf{p} = p^2 dp d\Omega = m\sqrt{2mE} dE d\Omega$, 式中 $p = |\mathbf{p}|$ 表示末态粒子动量大小，E 表示末态粒子能量，Ω 是立体角，$d\Omega$ 就相当于前面一般公式中的 $d\chi$。因此，在这种情况下末态态密度 $\rho(E)$ 就是

$$\rho(E) = m\sqrt{2mE}. \tag{10.22}$$

根据一阶含时微扰的公式 (10.16)，为了计算跃迁概率密度 $p_{i\to\alpha_f}$，我们需要计算 $\int_{t_0}^{t} dt_1 e^{i\omega_{fi}t_1}\langle\alpha_f|V|i\rangle = \langle\alpha_f|V|i\rangle\int_{t_0}^{t} dt_1 e^{i\omega_{fi}t_1}$。不过，为了后面的数学处理更方便，对于末态是连续谱的情形，我们常常取初始时刻 $t_0 = -t/2$，同时取末尾时刻为 $t/2$ 时刻 (因此整个过程持续的时间依然是 t)，因此我们要计算的实际上是 $\langle\alpha_f|V|i\rangle\int_{-t/2}^{t/2} dt_1 e^{i\omega_{fi}t_1}$。前面说过，对于末态是连续谱的情形，实际中关心的其实总是 $t\to\infty$ 时 (严格来说是 t 充分大时) 的结果，因此这时候我们就有 $\langle\alpha_f|V|i\rangle\lim_{t\to\infty}\int_{-t/2}^{t/2} dt_1 e^{i\omega_{fi}t_1} = \langle\alpha_f|V|i\rangle 2\pi\delta(\omega_{fi})$。这里我们利用了数学公式

$$\int_{-\infty}^{+\infty} dt e^{i\omega t} = 2\pi\delta(\omega). \tag{10.23}$$

但是为了计算跃迁概率密度，我们需要将跃迁概率幅进行模方，这就涉及到 δ 函数的平方该如何处理的问题。物理学家常常采用下面这种数学上不太严格，但是很有效很方便

的处理方法。那就是

$$(\delta(\omega))^2=\delta(\omega)\delta(\omega)=\delta(\omega)\lim_{t\to\infty}\int_{-t/2}^{t/2}\frac{dt_1}{2\pi}e^{i\omega t_1}=\delta(\omega)\lim_{t\to\infty}\int_{-t/2}^{t/2}\frac{dt_1}{2\pi}e^{i\cdot 0\cdot t_1}$$
$$=\delta(\omega)\lim_{t\to\infty}\int_{-t/2}^{t/2}\frac{dt_1}{2\pi}\cdot 1=\delta(\omega)\frac{t}{2\pi}|_{t\to\infty}, \tag{10.24}$$

式中第三个等号我们利用了 δ 函数的公式 $\delta(\omega)F(\omega)=\delta(\omega)F(0)$($F(\omega)$ 为 ω 的一个任意函数)。上面的推导告诉我们，当 t 充分大时，我们有

$$(\delta(\omega))^2=\delta(\omega)\frac{t}{2\pi}, \tag{10.25}$$

这里的 ω 当然是一个频率变量。

利用公式 (10.25) 以及之前对跃迁概率幅的计算结果 $-\frac{i}{\hbar}\langle\alpha_f|V|i\rangle 2\pi\delta(\omega_{fi})$, 我们马上就可以得到跃迁概率密度 $p_{i\to\alpha_f}$ 为

$$p_{i\to\alpha_f}=\frac{2\pi}{\hbar}|\langle\alpha_f|V|i\rangle|^2\delta(E_f-E_i)\cdot t, \tag{10.26}$$

为了得到这个式子我们还利用了 δ 函数的关系式 $\delta(x)/a=\delta(ax)$，其中 x 为变量，$a>0$ 为常数。将这个概率密度乘上末态数目 $d\alpha_f$，并对以 $E_f=E_i$ 为中心的末态能谱区间 $\Delta(E_f)$ 积分 (这也就是测量仪器的精度区间)，注意到 $\delta(E_f-E_i)$ 的积分是 1，我们就得到跃迁概率 (还是记作 $p_{i\to f}$)

$$p_{i\to f}=\frac{2\pi}{\hbar}|\langle\alpha_f|V|i\rangle|^2\rho(E_f,\chi_f)|_{E_f=E_i}d\chi_f\cdot t. \tag{10.27}$$

注意到这个跃迁概率正比于微扰作用的时间 t，因此，单位时间的跃迁概率 $w_{i\to f}=dp_{i\to f}/dt$ 为，

$$w_{i\to f}=\frac{2\pi}{\hbar}|\langle\alpha_f|V|i\rangle|^2\rho(E_f,\chi_f)|_{E_f=E_i}d\chi_f. \tag{10.28}$$

以上这个结果，常常被称为费米黄金规则。有时候人们也直接用单位时间的跃迁概率密度 $w_{i\to\alpha_f}$ 将这一规则表达成

$$w_{i\to\alpha_f}=\frac{2\pi}{\hbar}|\langle\alpha_f|V|i\rangle|^2\delta(E_f-E_i). \tag{10.29}$$

式中的 $\delta(E_f-E_i)$ 意味着，在恒定微扰的作用之下，末态的能量必然和初态的能量相等，因此这样的量子跃迁满足能量守恒。

我们也可以计算 t 时间之内，从初态 $|i\rangle$ 出发到所有可能末态 (当然不包括 $|i\rangle$ 态本身) 的跃迁总概率, 显然它是 $t\cdot\frac{2\pi}{\hbar}\int|\langle\alpha_f|V|i\rangle|^2\rho(E_f,\chi_f)|_{E_f=E_i}d\chi_f$, 式中的积分表示对所有可能的 $d\chi_f$(也就是整个末态相空间) 积分。这个总概率必定要小于 1(因为 $|i\rangle$ 有一定的概率不跃迁)，并且为了保证一阶微扰成立实际上应该远小于 1。注意到这个跃迁总概率正比于 t，因此这实际上说明了，前面我们取 $t\to\infty$ 其实只是一种近似，真实的 t 其实是微扰有效作用的时间，考虑到 $|\langle\alpha_f|V|i\rangle|^2$ 是一个二阶小量，因此这个时间可以充分大 (以使得前面 $t\to\infty$ 的近似处理成立)，但并不能真正趋于无穷，实际上跃迁总概率小于 1 的条件告诉我们，真实的微扰有效作用的时间并不能过于长，否则我们的含时微扰论将不能适用。

势散射理论初论

作为恒定微扰的一个例子，让我们来考虑一个粒子在一个固定势场中的散射问题，我们将把势能 $V(\mathbf{x})$ 看成是对自由粒子哈密顿量的微扰，我们假定 $V(\mathbf{x})$ 在空间无穷远处衰减得足够快，最重要的是 $V(\mathbf{x})$ 不显含时间，因此是一个恒定微扰。系统的哈密顿量是 $H = H_0 + V(\mathbf{x})$，其中 $H_0 = \frac{\mathbf{P}^2}{2m}$ 是自由粒子的哈密顿量。假设入射粒子处在 H_0 的本征态 $|\mathbf{p}_i\rangle$, $\mathbf{p}_i$ 表示入射粒子动量，因此入射粒子的能量是 $E_i = \frac{\mathbf{p}_i^2}{2m}$。类似的假设出射粒子处在 H_0 的本征态 $|\mathbf{p}_f\rangle$，$|\mathbf{p}_f\rangle$ 同时为动量本征态，满足归一化关系 $\langle \mathbf{p}_f|\mathbf{p}'_f\rangle = \delta(\mathbf{p}_f - \mathbf{p}'_f)$，出射粒子能量为 $E_f = \frac{\mathbf{p}_f^2}{2m}$。因此势散射问题其实就是粒子从初态 $|\mathbf{p}_i\rangle$ 跃迁到末态 $|\mathbf{p}_f\rangle$ 的跃迁问题。由于在这里末态显然是一个连续谱，所以可以认为入射粒子是在 $-t/2 \to -\infty$ 的无穷远过去入射，在 $t/2 \to \infty$ 的无穷远将来出射，因此在一阶微扰上，我们最终会有前面推导出来的费米黄金规则 (10.28)。

为了使用费米黄金规则 (10.28)，我们需要计算末态态密度，根据我们前面的计算 (10.22)，结果是 $\rho(E_f) = m\sqrt{2mE_f} = mp_f$, 式中 $p_f = |\mathbf{p}_f|$ 为末态动量的大小。另外，我们还需要计算矩阵元 $\langle \mathbf{p}_f|V(\mathbf{x})|\mathbf{p}_i\rangle$, 利用位置本征态的封闭性关系，我们有 $\langle \mathbf{p}_f|V(\mathbf{x})|\mathbf{p}_i\rangle = \int d^3\mathbf{x}\langle \mathbf{p}_f|V(\mathbf{x})|\mathbf{x}\rangle\langle \mathbf{x}|\mathbf{p}_i\rangle = \int d^3\mathbf{x}V(\mathbf{x})\langle \mathbf{p}_f|\mathbf{x}\rangle\langle \mathbf{x}|\mathbf{p}_i\rangle = \frac{1}{(2\pi\hbar)^3}\int d^3\mathbf{x}V(\mathbf{x})e^{i(\frac{\mathbf{p}_i-\mathbf{p}_f}{\hbar})\cdot\mathbf{x}}$, 式中我们利用了 $\langle \mathbf{x}|\mathbf{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}e^{i\frac{\mathbf{p}}{\hbar}\cdot\mathbf{x}}$。因此根据费米黄金规则 (10.28)，我们可以得到单位时间粒子从 $|\mathbf{p}_i\rangle$ 态散射到 $|\mathbf{p}_f\rangle$ 态的概率,

$$w_{i\to f} = \frac{2\pi}{\hbar}\frac{1}{(2\pi\hbar)^6}\left|\int d^3\mathbf{x}V(\mathbf{x})e^{i(\frac{\mathbf{p}_i-\mathbf{p}_f}{\hbar})\cdot\mathbf{x}}\right|^2 mp_f d\Omega, \tag{10.30}$$

式中 Ω 为出射粒子的立体角，并且整个散射过程要满足能量守恒，即

$$E_f = E_i. \tag{10.31}$$

但是实验上测量的其实是一个和 $w_{i\to f}$ 密切相关的量，叫做微分散射截面，记作 $d\sigma$, 它的定义是单位时间的散射概率 $w_{i\to f}$ 比上入射粒子概率流密度的大小 J_i，即

$$d\sigma = \frac{w_{i\to f}}{J_i}. \tag{10.32}$$

由于概率流密度的物理含义是单位时间在单位横截面积上粒子通过的概率，所以微分散射截面的物理含义实际上是：最后被散射到立体角 $d\Omega$ 之内的粒子数 (实验上入射粒子并不是一个，而是一束相同的粒子) 与单位时间单位截面积上入射的粒子数的比值。由于 $w_{i\to f}$ 是时间的负一次方量纲，而 J_i 的量纲是时间的负一次方乘上面积的负一次方，因此微分散射截面有面积的量纲，所以称作截面。

对于归一化到 δ 函数的入射平面波，其波函数是 $\frac{1}{(2\pi\hbar)^{3/2}}e^{i\frac{\mathbf{p}_i\cdot\mathbf{x}}{\hbar}}$，因此人们容易求得入射概率流密度的大小 J_i 为

$$J_i = \frac{1}{(2\pi\hbar)^3}\frac{p_i}{m}. \tag{10.33}$$

将这个结果和之前得到的 $w_{i\to f}$ 的结果 (10.30) 代入微分截面的定义式，我们就能得到

$$d\sigma = \left(\frac{m}{2\pi\hbar^2}\right)^2\left|\int d^3\mathbf{x}V(\mathbf{x})e^{i(\frac{\mathbf{p}_i-\mathbf{p}_f}{\hbar})\cdot\mathbf{x}}\right|^2 d\Omega, \tag{10.34}$$

其中我们利用了能量守恒关系 $p_f = p_i$。注意，这个结果 (10.34) 只是一阶微扰论的结果，通常称作玻恩近似的结果。在本章的最后两节我们将把这样的结论推广到微扰论的任意阶。

微分散射截面的概念不仅适用于单个入射粒子受一个固定势场散射的情形。也适用于两体散射情形，也就是两个粒子相互飞近，然后发生相互作用，再相互散射开来的情形。这时候入射粒子是两个粒子，散射以后出射的依然是这两个粒子。但是由于总动量守恒，两体问题总是可以化约为单体问题。这个结论我们在理论力学里就已经知道了，结论是，我们可以选取质心系 (从而使得我们可以完全忽略质心的运动)，然后考察一个粒子相对于另一个粒子的运动，即引入相对坐标 $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$，并将质量替换成约化质量 $m_R = m_1 m_2/(m_1 + m_2)$ 就可以了。因此对于两体散射问题，我们也可以选取质心系，并把其中的一个粒子称作靶粒子，把另外一个相对于靶粒子运动的粒子称作散射粒子，然后引入散射粒子相对于靶粒子的相对坐标 $\mathbf{x}$ 和约化质量，问题就依然归结为散射粒子受靶粒子势场 $V(\mathbf{x})$ 散射的问题。这时候前文中的入射粒子动量 $\mathbf{p}_i$ 以及出射粒子动量 $\mathbf{p}_f$ 就应该理解为散射粒子在质心系中散射前后的动量 (靶粒子的动量可以不用管，因为在质心系中，它一定和散射粒子的动量相反)。散射粒子入射时的概率流密度现在是 $J = \frac{1}{(2\pi\hbar)^3}\frac{p}{m_R} = \frac{1}{(2\pi\hbar)^3}|\frac{\mathbf{p}}{m_1} + \frac{\mathbf{p}}{m_2}| = \frac{1}{(2\pi\hbar)^3}|\frac{\mathbf{p}_1}{m_1} - \frac{\mathbf{p}_2}{m_2}| = \frac{1}{(2\pi\hbar)^3}|\mathbf{v}_1 - \mathbf{v}_2|$，这里我们利用了在质心系中，散射粒子动量 $\mathbf{p}_1 = \mathbf{p}$，靶粒子动量 $\mathbf{p}_2 = -\mathbf{p}$。也就是说，散射粒子相对于靶粒子的概率流密度由它们的相对速度决定。

但是，值得注意的是，散射截面的概念在 3 个粒子散射，或者更多粒子相互散射的情形中是不适用的。这时候人们往往需要直接测量跃迁概率。不过这样的情形在实验上相对难以实现，也相对没有那么重要。即使在量子场论中，也由于粒子可以湮灭和产生，散射末态的粒子可以和初态粒子不同，甚至可以有很多末态粒子，因此更多粒子的相互作用规律往往可以通过两个粒子碰撞产生很多末态粒子来观测，在这种情形下由于初态只有两个粒子，因此散射截面的概念依然适用。实际上，高能物理的粒子对撞机都是研究两粒子的对撞。

下面我们用玻恩近似 (10.34) 计算一个具体的散射截面，我们假设散射势能函数取如下形式

$$V(\mathbf{x}) = V_0 \frac{e^{-\frac{r}{a}}}{r}. \tag{10.35}$$

式中 r 是散射粒子的径向坐标，V_0 和 a 都是常数，并且 $a > 0$。V_0 表示这个势的强度，我们假设 $|V_0|$ 足够小，以致于玻恩近似可以成立。a 其实描写的是这个势相互作用的力程，如果散射粒子的距离 r 比 a 大很多，那实际上它所受到的相互作用力就可以忽略不计了，因此这个势描写的是一种短程力。但是如果我们取力程 $a \to +\infty$ 的极限，那这个势能显然就变成一个库伦势了。

(10.35) 式给出来的这个势称作汤川势。汤川首先想到用这样一个相互作用势描写力程约为一个费米的核力，为了解释这个势的来源，他不得不假设 π 介子的存在，后来，人们的确在实验上发现了 π 介子，汤川也为此得了诺贝尔奖。

为了计算出汤川势的散射截面，我们记入射粒子的波矢量为 $\mathbf{k}_i = \mathbf{p}_i/\hbar$，记这个粒子在

末态出射时的波矢量为 $\mathbf{k}_f = \mathbf{p}_f/\hbar$，记散射中的波矢量转移为 $\mathbf{k} = \mathbf{k}_f - \mathbf{k}_i = (\mathbf{p}_f - \mathbf{p}_i)/\hbar$, 因此根据公式 (10.34), 我们需要计算

$$\begin{aligned}
&V_0 \int d^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \frac{e^{-\frac{r}{a}}}{r} \\
=&V_0 2\pi \int_0^\infty r^2 dr \int_0^\pi \sin\theta d\theta e^{-i|\mathbf{k}|r\cos\theta} \frac{e^{-\frac{r}{a}}}{r} \\
=&V_0 \frac{4\pi}{|\mathbf{k}|} \int_0^\infty dr \sin(|\mathbf{k}|r) e^{-\frac{r}{a}} \\
=&V_0 \frac{4\pi}{|\mathbf{k}|} \mathrm{Im} \int_0^\infty dr e^{i|\mathbf{k}|r} e^{-\frac{r}{a}} \\
=&V_0 \frac{4\pi}{|\mathbf{k}|} \mathrm{Im} \left(\frac{1}{\frac{1}{a} - i|\mathbf{k}|} \right) = \frac{4\pi V_0}{(\frac{1}{a})^2 + |\mathbf{k}|^2}.
\end{aligned} \tag{10.36}$$

将这个结果代入微分截面的公式 (10.34)，就可以得到

$$d\sigma = \frac{4m^2V_0^2}{\hbar^4} \frac{1}{[(\frac{1}{a})^2 + |\mathbf{k}|^2]^2} d\Omega. \tag{10.37}$$

注意到 $|\mathbf{p}_i| = p_i = p_f = |\mathbf{p}_f|$, 进而由图 (10.2) 很容易看出 $|\mathbf{k}| = 2(p_i/\hbar)\sin(\frac{\theta}{2}) = 2(\sqrt{2mE}/\hbar)\sin(\frac{\theta}{2})$,

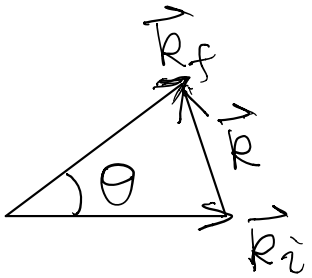

Figure 10.2: 散射前后波矢量的变化.

其中的 θ 不是刚才算积分的过程中涉及到的 θ, 而是散射角，即出射方向和入射方向的夹角, 而 E 则是入射粒子动能，当然由于能量守恒，也是出射粒子动能。将这些代入 (10.37), 就可以得到

$$d\sigma = \frac{4m^2V_0^2}{\hbar^4} \frac{1}{[(\frac{1}{a})^2 + \frac{8mE}{\hbar^2}\sin^2(\frac{\theta}{2})]^2} d\Omega. \tag{10.38}$$

如果在上面的结果中取 $a \to +\infty$ 的极限，并取 $V_0 = \frac{Z_1Z_2e^2}{4\pi\varepsilon_0} = Z_1Z_2e_s^2$(式中 $e_s^2 = \frac{e^2}{4\pi\varepsilon_0}$)，那么汤川势就变成了两个序数分别为 Z_1 和 Z_2 的原子核之间的库伦排斥势能，这时候上面的结果 (10.38) 就变成著名的卢瑟福公式

$$d\sigma = \frac{Z_1^2Z_2^2e_s^4}{16E^2\sin^4(\frac{\theta}{2})} d\Omega. \tag{10.39}$$

注意，在卢瑟福公式里，所有的普朗克常数 $\hbar$ 全都消去了，这就是为什么卢瑟福用经典力学的推导也能正确地得到这个公式的原因。

10.1.3 简谐振荡微扰

下面我们考察另外一种重要的微扰，简谐振荡微扰，即以固定频率 ω 随时间振荡的微扰，其微扰相互作用 $V(t)$ 可以写成

$$V(t)=We^{-i\omega t}+W^{\dagger}e^{i\omega t}, \tag{10.40}$$

其中算符 W 及其厄米共轭算符 $W^{\dagger}$ 都不显含时间。我们将首先处理末态是连续谱的情形，然后再讨论初末态都是离散能级的情形。

末态为连续谱的情形

对于末态是连续谱的情形，和恒定微扰的讨论一样，人们实际关心的总是一个充分长的时间，因此我们可以近似认为初始时刻是在 $-\infty$ 的无穷远过去，而末尾时刻是在 $+\infty$ 的无穷远将来，也即是说，整个微扰作用过程持续的时间 t 充分长。

首先我们要利用公式 (10.16) 计算跃迁概率密度 $p_{i\to\alpha_f}$。为此我们需要计算 $\int_{-\infty}^{+\infty}dt_1e^{i\omega_{fi}t_1}\langle\alpha_f|V(t_1)|i\rangle$，代入振荡微扰的一般形式 (10.40)，我们可以得到结果是 $\int_{-\infty}^{+\infty}dt_1\left[e^{i(\omega_{fi}-\omega)t_1}\langle\alpha_f|W|i\rangle+e^{i(\omega_{fi}+\omega)t_1}\langle\alpha_f|W^{\dagger}|i\rangle\right]=W_{fi}2\pi\delta(\omega_{fi}-\omega)+W_{if}^{*}2\pi\delta(\omega_{fi}+\omega)$，式中 $W_{fi}=\langle\alpha_f|W|i\rangle$，并且我们利用了 $(W^{\dagger})_{fi}=W_{if}^{*}$。类似于恒定微扰的情形，通过利用数学公式 (10.25)，并注意到 $\delta(\omega_{fi}-\omega)\delta(\omega_{fi}+\omega)=0$(因为无论 ω 取多少，这两个 δ 函数至少有一个会等于 0)，我们可以得到单位时间的跃迁概率密度 $w_{i\to\alpha_f}=\frac{dp_{i\to\alpha_f}}{dt}$，结果完全类似于费米黄金规则

$$w_{i\to\alpha_f}=\frac{2\pi}{\hbar}\left[|W_{fi}|^2\delta(E_f-E_i-\hbar\omega)+|W_{if}|^2\delta(E_f-E_i+\hbar\omega)\right]. \tag{10.41}$$

这个结论的物理意义很清楚，等式右边的第一项仅在末态能量 E_f 比初态能量 E_i 高 $\hbar\omega$ 时才非 0，因此表示的是系统从微扰中吸收能量 $\hbar\omega$ 进而跃迁到更高能级的过程，相反，等式右边的第二项表示的是系统释放能量 $\hbar\omega$，进而从最初的更高能级 E_i 跃迁到低能级 E_f 的过程，当然，等式 (10.41) 右边的两个 δ 函数不可能同时非 0，因此这两个过程不可能同时实现。和恒定微扰情形一样，为了得到单位时间的跃迁概率 $w_{i\to f}$，我们需要将概率密度 $w_{i\to\alpha_f}$ 对末态能谱作一个合适的积分，所有的处理都和恒定微扰情形完全一样，因此我们这里不再赘述。

离散能级情形

下面我们来考察初末态都是离散能级的情形，根据我们在恒定微扰中的讨论，这时候微扰作用的时间 t 不再能当作无穷大来处理，为此这时候我们不妨将初始时刻取为 $t_0=0$，末尾时刻取为有限时刻 t。并且这时候要计算的直接就是跃迁概率 $p_{i\to f}$ 而不再是概率密度。为了利用公式 (10.16) 计算 $p_{i\to f}$，我们首先需要计算 $\int_0^t dt_1e^{i\omega_{fi}t_1}\langle f|V(t_1)|i\rangle$。代入振荡微扰的一般形式 (10.40)，我们要计算的其实是

$$\int_0^t dt_1\left[e^{i(\omega_{fi}-\omega)t_1}W_{fi}+e^{i(\omega_{fi}+\omega)t_1}W_{if}^{*}\right], \tag{10.42}$$

式中 $W_{fi}=\langle f|W|i\rangle$。

由于算符 W 不显含时间，所以 (10.42) 式的积分当然可以很容易计算出来。但是对于这种一般情形，由于 (10.42) 式两项之间的复杂干涉，结果的物理图像其实并不是很清晰。

为了看清楚物理图像，人们通常考察所谓的共振近似 (也叫做旋波近似)，即我们分别考察 $\omega \sim \omega_{fi}$(由于假定 $\omega > 0$，因此这时候当然需要 $E_f > E_i$) 和 $\omega \sim \omega_{if} = -\omega_{fi}$(这时候当然需要 $E_f < E_i$) 两种特殊情形。并假定微扰作用的时间 t 比较长 (虽然不能近似成无穷长)，以至于

$$\omega t \sim |\omega_{fi}| t \gg 1. \tag{10.43}$$

不妨以 $\omega \sim \omega_{fi}$ 的情形为例。这时候公式 (10.42) 的两项中，第一项随时间的振荡频率为 $|\omega_{fi} - \omega| \ll \omega$, 因此是一个很缓慢的振荡，而第二项的振荡频率为 $\omega_{fi} + \omega \sim 2\omega$, 因此是一个相对比较快速的振荡，实际上，根据近似条件 (10.43), 在微扰的作用时间之内这个快速振荡项其实振荡了非常多次，因此实际上这一项对结果的贡献将会平均为 0(由于多次振荡会相互抵消)，也就是说，在 $\omega \sim \omega_{fi}$ 并满足近似条件 (10.43) 时，我们其实可以完全忽略公式 (10.42) 中第二项这个快速振荡项。因此近似的有 $\int_0^t dt_1 e^{i\omega_{fi}t_1} \langle f|V(t_1)|i\rangle \approx \int_0^t dt_1 e^{i(\omega_{fi}-\omega)t_1} W_{fi} = W_{fi} e^{i\frac{1}{2}(\omega_{fi}-\omega)t} \frac{\sin((\omega_{fi}-\omega)t/2)}{(\omega_{fi}-\omega)/2}$。因此，这时候跃迁概率 $p_{i\to f}$ 将近似为

$$p_{i\to f} \approx \frac{1}{\hbar^2} |W_{fi}|^2 \left[\frac{\sin((\omega_{fi}-\omega)t/2)}{(\omega_{fi}-\omega)/2}\right]^2. \tag{10.44}$$

这个结果完全类似于恒定微扰离散能级时的结果, 由于这时候 $E_f > E_i$，所以这个跃迁概率代表的是系统与微扰共振并吸收能量的情形。当 $\omega = \omega_{fi}$ 时，跃迁概率达到最大，为 $\frac{1}{\hbar^2}|W_{fi}|^2 t^2$，一阶微扰近似适用的条件要求这个概率要远小于 1，因此我们得到第二个近似条件,

$$|W_{fi}| t/\hbar \ll 1. \tag{10.45}$$

与共振近似条件 (10.43) 不同，这个近似条件来源于微扰论的适用条件。为了能同时满足近似条件 (10.43) 和 (10.45), 显然要求微扰算符 W 满足

$$|W_{fi}| \ll \hbar|\omega_{fi}| = |E_f - E_i|. \tag{10.46}$$

类似的，人们也可以讨论 $\omega \sim -\omega_{fi}$ 的情形。这时候可以得到，跃迁概率近似为

$$p_{i\to f} \approx \frac{1}{\hbar^2} |W_{if}|^2 \left[\frac{\sin((\omega_{fi}+\omega)t/2)}{(\omega_{fi}+\omega)/2}\right]^2. \tag{10.47}$$

由于这时候 $E_f < E_i$, 因此它代表的是系统与微扰共振但是释放能量的情形。

以上就是在一阶微扰的框架下对于离散能级与微扰共振的处理，值得再次强调的是，这种一阶微扰的处理办法需要满足微扰近似条件 (10.45)。因此如果微扰起作用的时间 t 过于长，以至于 (10.45) 不成立时，一阶微扰的处理将会失效，本章后面在关于量子共振的一般性研究中，我们将会进一步考察这时候该如何处理问题。

10.2 光的辐射与吸收

10.2.1 原子对光的辐射与吸收

这一节我们将应用前面发展的含时微扰论来研究原子与电磁场的相互作用，具体来说就是研究原子对光的辐射和吸收所导致的能级跃迁过程。研究光的辐射和吸收过程不仅可

以使得我们理解光和物质相互作用的规律，而且在技术上也很重要，比方说激光就是这一研究的重要产物。我们的激光打印机、商场的条形码扫描器、我们通过光纤网络浏览视频等等，都离不开激光，我们几乎每天都在使用激光。但要理解激光的工作原理就需要用量子力学来研究光的辐射和吸收过程。激光在自然界中原本是不存在的，完全是因为人们深入理解了光的辐射和吸收机理才人为制造了出来。

早在 1917 年爱因斯坦就迈出了理解光的辐射和吸收机理的关键一步，也正是爱因斯坦迈出的这一步才使得激光成为可能。爱因斯坦提出，光的辐射和吸收可以分成三个基本过程，如图 (10.3) 所示：自发辐射过程、吸收过程、以及受激辐射过程。吸收过程就是原

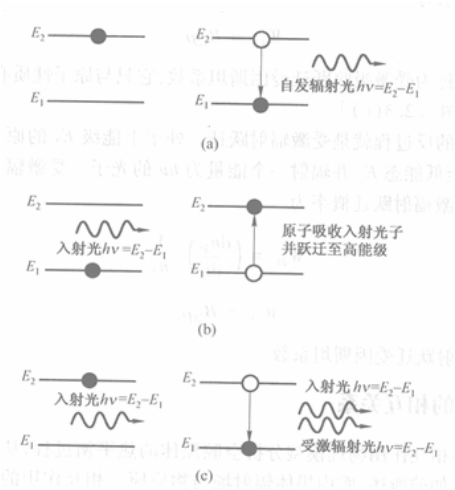

Figure 10.3: 光的辐射和吸收的三个基本过程：(a) 自发辐射过程. (b) 吸收过程. (c) 受激辐射过程.

子吸收一个光子进而跃迁到高能级的过程。自发辐射过程就是原子自发地从高能级跃迁到低能级并放出一个光子的过程。原子通过自发辐射过程发出来的光就是普通的光，如日光灯的光，霓虹灯的光等等，其特点是辐射光的频率、偏振方向、以及相位都很不一致。但是，如果原子受外来光子的诱导而从高能级跃迁到低能级并辐射出光子，那辐射出来的光子就会和诱导它的外来光子在频率、位相、以及传播方向等方面都完全一致，因此可以说，外来光被相干地放大了。这样一个有很好的相干性的发光过程就是受激辐射过程。

当然，上面所描述的受激辐射图像是后人在爱因斯坦的基础上进一步发展的结果，是一幅完全量子的图像。爱因斯坦关于光的辐射和吸收的原始模型里面并没有充分指出受激辐射的相干性。完整地处理这一量子图像需要量子光学或者量子电动力学的知识，也就是说，需要我们不仅用量子力学的方式处理原子，同时也要用量子力学的方式处理电磁场。然而将电磁场进行量子化超出了本书的范围 (不过在下一节中，我们将研究一个特殊的模型，在那个模型中，电磁场会被处理成量子谐振子)，因此我们这里将采用半经典的处理，也就是说我们将依然把电磁场看成是一个经典场，而仅仅将电磁场中的原子当作量子力学客体。

这种半经典的处理将使得我们仅仅只能直接研究原子对电磁波的吸收过程和受激辐

射过程，因为这两种过程都可以看成是在一个预先存在的电磁场背景下发生的，用经典的方式来处理这个背景电磁场并不会实质性地改变物理结论 (当然，要深入理解受激辐射的相干性也还是需要量子光学的处理)。然而，原子的自发辐射过程却有所不同，这时候并没有一个预先存在的背景电磁场，为了直接研究原子的自发辐射过程，我们必须同时将电磁场量子化。不过，爱因斯坦原始模型的一个经典结果将自发辐射系数和受激辐射联系了起来，这使得人们可以绕过电磁场的量子化得到自发辐射系数，结果与量子场论的直接推导完全吻合。值得指出的是，在历史上，第一个完全用量子场的方式推导出自发辐射系数的是狄拉克。

为了研究原子与背景电磁场的相互作用，我们假设电磁场的波长远大于原子的尺寸 (由于原子的能级差通常是 eV 量级，因此原子辐射和吸收的电磁波波长 λ 大致在几百个纳米到上千个纳米，即 $\lambda \sim 10^{-6}m$, 而原子的尺寸大致是 $10^{-10}m$，因此这个假设是成立的)。因此，在原子的尺寸上，我们可以忽略电磁场随着空间的变化，进而将原子看成是与一个均匀电磁场 (场强为 $\mathbf{E}(t),\mathbf{B}(t)$) 耦合的电偶极子和磁偶极子，我们将原子的电偶极矩算符记作 $\mathbf{d}$，将磁偶极矩算符记作 $\mathbf{m}$。因此原子与电磁场的相互作用项 $V(t)$ 为

$$V(t) = -\mathbf{d}\cdot\mathbf{E}(t) - \mathbf{m}\cdot\mathbf{B}(t). \tag{10.48}$$

这一近似通常称作偶极近似。

实际上，在低阶近似中，磁偶极矩和磁场的相互作用也可以忽略。不妨以氢原子为例来分析其原因，这时候，原子与电磁场的耦合实际上是核外电子与电磁场的耦合，因此 $\mathbf{d} = -e\mathbf{x}$($\mathbf{x}$ 为电子的位置矢量), 其大小在 $ea_0 = e\frac{\hbar}{m\alpha c}$ 的量级，这里 a_0 为玻尔半径，$\alpha = e^2/(4\pi\varepsilon_0\hbar c) \approx 1/137$ 为精细结构常数。类似的，这时候 $\mathbf{m} = -\frac{e}{2m}\mathbf{L}$($\mathbf{L}$ 为电子的角动量)，其大小在 $\frac{e\hbar}{2m}$ 的量级。又由于对于电磁波而言，其磁场强度 $\mathbf{B}$ 的大小等于电场强度 $\mathbf{E}$ 的大小除以 c，即 $|\mathbf{B}|/|\mathbf{E}| = 1/c$。因此我们有

$$\frac{|\mathbf{m}\cdot\mathbf{B}|}{|\mathbf{d}\cdot\mathbf{E}|} \sim \frac{1}{c}\cdot(\frac{e\hbar}{2m})/(\frac{e\hbar}{m\alpha c}) = \alpha \approx 1/137. \tag{10.49}$$

也就是说，磁偶极矩与磁场的相互作用大约比电偶极矩与电场的相互作用小两个数量级。因此，作为低阶微扰近似，忽略掉原子的磁偶极矩和磁场的相互作用是很合理的。

我们首先考虑一个单色电磁波与原子的耦合，假定 $\mathbf{E}(t) = \mathbf{E}_0\cos(\omega t)$, 式中 $\mathbf{E}_0$ 是个常数，表示场强的振幅。那么相应的相互作用 $V(t)$ 将是

$$V(t) = -\mathbf{d}\cdot\frac{\mathbf{E}_0}{2}(e^{-i\omega t} + e^{i\omega t}). \tag{10.50}$$

但是请注意，我们真正要讨论的其实并不是一个单色电磁波的情形，因为在通常的条件下，原子受激辐射或者吸收电磁波的时候，实际的背景电磁波其实不是一个单色波，而是形如 $\mathbf{E}(t) = \int d\omega\mathbf{E}(\omega)\cos(\omega t)$ 这样的一个单色波的连续叠加，因此频率 ω 并不是给定的，而是一个连续变量。在这种情况下，虽然原子的初末态可能都是离散能级，但是我们计算出来的单位时间跃迁概率 $w_{i\to f}$ 其实依然是一个概率密度，是关于微扰频率 ω 的概率密度，而最终的跃迁概率是要对 ω 进行一个合适积分的。因此，前面关于连续谱的那些讨论和公

式在这里依然适用。相反，如果实际的电磁场真的是一个单色波的话，那下面的讨论反而将不适用了。

我们以原子吸收电磁波的情形为例来进行讨论。这时候，根据上一节对于连续谱情形的讨论，单位时间之内原子从初态 $|i\rangle$ 跃迁到末态 $|f\rangle$ 的概率 $w_{i\to f}$(上面说过，实际上它是一个关于 ω 的概率密度) 为，

$$\begin{aligned} w_{i\to f} &= \frac{2\pi}{\hbar}|\langle f|(\mathbf{d}\cdot\mathbf{E}_0)/2|i\rangle|^2\delta(E_f-E_i-\hbar\omega) \\ &= \frac{\pi}{2\hbar^2}|\langle f|\mathbf{d}\cdot\mathbf{E}_0|i\rangle|^2\delta(\omega_{fi}-\omega). \end{aligned} \tag{10.51}$$

但是，这个结论处理的是一个有确定偏振方向 $\mathbf{E}_0$ 的电磁波，考虑到原子通常处在自然光条件下，而自然光是非极化的，因此我们需要将 $\mathbf{E}_0$ 的不同方向进行平均，也就是说要把上面结果中的 $|\langle f|\mathbf{d}\cdot\mathbf{E}_0|i\rangle|^2$ 替换成 $\frac{1}{3}|\mathbf{E}_0|^2|\langle f|\mathbf{d}|i\rangle|^2$, 注意 $|\mathbf{E}_0|^2$ 是和方向无关的，并且式中 $|\langle f|\mathbf{d}|i\rangle|^2=\langle f|\mathbf{d}|i\rangle\cdot\langle f|\mathbf{d}|i\rangle^*$。因此最终的结果是

$$w_{i\to f}=\frac{\pi}{6\hbar^2}|\mathbf{E}_0|^2|\langle f|\mathbf{d}|i\rangle|^2\delta(\omega_{fi}-\omega). \tag{10.52}$$

下面我们来考虑如何将上面的概率密度对 ω 进行一个适当的积分，也就是考虑如何从一个单色波的结果过渡到不同单色波的连续叠加的情形。我们注意到电场 $\mathbf{E}(t)=\mathbf{E}_0\cos(\omega t)$ 的能量密度对一个振动周期的平均值为 $\frac{1}{2}\varepsilon_0|\mathbf{E}_0|^2\cdot\frac{1}{2}$。由于电磁波的磁场能量密度对一个振动周期的平均值总是和电场能量密度的平均值相等 (电动力学的熟知结论)。所以 $\mathbf{E}(t)=\mathbf{E}_0\cos(\omega t)$ 的单色波所对应的电磁场能量密度平均值为 $\frac{1}{2}\varepsilon_0|\mathbf{E}_0|^2$。但这只是一个单色波的结果，要过渡到单色波的连续叠加情形的话，我们就需要把这个能量密度替换成 $u(\omega)d\omega$, 其中 $u(\omega)$ 表示电磁场在 ω 附近的单位频率区间里的能量密度。因此最终我们需要在结果 (10.52) 中进行 $\frac{1}{2}\varepsilon_0|\mathbf{E}_0|^2\to u(\omega)d\omega$ 的替换，从而就会得到

$$w_{i\to f}=\frac{\pi}{3\varepsilon_0\hbar^2}u(\omega)d\omega|\langle f|\mathbf{d}|i\rangle|^2\delta(\omega_{fi}-\omega). \tag{10.53}$$

现在，将这个结果对 ω 积分，就会得到 (我们依然使用了记号 $w_{i\to f}$, 希望不会引起读者的混淆)

$$w_{i\to f}=\frac{\pi}{3\varepsilon_0\hbar^2}|\langle f|\mathbf{d}|i\rangle|^2u(\omega_{fi}). \tag{10.54}$$

由于这个单位时间跃迁概率正比于电磁场的能量密度 u, 所以通常也写作 $w_{i\to f}=B_{fi}u(\omega_{fi})$, B_{fi} 称为从 $|i\rangle$ 态吸收一个光子跃迁到高能级 $|f\rangle$ 态的吸收系数，很显然它是

$$B_{fi}=\frac{\pi}{3\varepsilon_0\hbar^2}|\langle f|\mathbf{d}|i\rangle|^2. \tag{10.55}$$

特别的，对于氢原子，由于 $\mathbf{d}=-e\mathbf{x}$，所以

$$B_{fi}=\frac{(2\pi)^2\alpha c}{3\hbar}|\langle f|\mathbf{x}|i\rangle|^2, \tag{10.56}$$

这里 α 是精细结构常数。

类似的，假设原子原来处在高能级 E_f，我们也可以考察它在背景电磁场的作用下，在单位时间之内跃迁到低能级 E_i，并辐射出光子的概率 $w_{f\to i}$(即受激辐射的概率)，结果是与吸收光子的概率完全相等，即

$$w_{f\to i} = w_{i\to f}. \tag{10.57}$$

这是因为，根据公式 (10.41)，从 $|i\rangle$ 跃迁到 $|f\rangle$ 的辐射项正比于 $|W_{fi}|^2$，但现在我们考察的是从 $|f\rangle$ 跃迁到 $|i\rangle$ 的辐射项，那当然还是正比于 $|W_{fi}|^2$，因此一切都和吸收光子从 $|i\rangle$ 跃迁到 $|f\rangle$ 的情形完全一样。历史上，受激辐射的概念是 1917 年在量子力学的完整理论建立起来之前爱因斯坦首先提出来的。爱因斯坦假设，在单位时间之内，原子在高能级 $|f\rangle$ 上受电磁场的诱导跃迁到低能级 $|i\rangle$，进而辐射出一个光子的概率 $w_{f\to i}$ 正比于电磁场的能量密度 $u(\omega_{fi})$，爱因斯坦将之写作，$w_{f\to i} = B_{if}u(\omega_{fi})$，$B_{if}$ 称之为受激辐射系数。因此方程 (10.57) 告诉我们的实际上是，

$$B_{if} = B_{fi}, \tag{10.58}$$

即受激辐射系数和吸收系数相等。这个结果最早也是爱因斯坦通过其原子辐射和吸收模型得到的。

10.2.2 爱因斯坦的受激辐射理论与激光

下面，让我们简单地回顾一下爱因斯坦的原始模型，我们的着重点在于搞清楚原子的自发辐射如何与吸收以及受激辐射联系起来。爱因斯坦的推理要点就在于考察当系统达到热平衡时，原子的吸收过程和辐射过程如何相平衡。

设想在黑体辐射的空腔里有大量完全相同的原子在不断辐射和吸收电磁波，并达到热平衡。假设我们以 N_n 表示热平衡时 $|n\rangle$ 能级的原子数目。则热平衡时，单位时间之内通过吸收一个光子从 $|i\rangle$ 能级跃迁到高能级 $|f\rangle$ 的原子数将是 $w_{fi}N_i = B_{fi}u(\omega_{fi})N_i$。相反，单位时间之内，从高能级 $|f\rangle$ 通过辐射一个光子从而跃迁到低能级 $|i\rangle$ 的原子分成两类，一类是通过自发辐射跃迁到低能级的原子，自发辐射和空腔里的电磁场没有关系，因此单位时间的跃迁概率是个常数，记作 A_{if}，也称之为自发辐射系数，所以单位时间之内通过自发辐射从 $|f\rangle$ 能级跃迁到 $|i\rangle$ 能级的原子数为 $A_{if}N_f$。另一类从 $|f\rangle$ 能级跃迁到 $|i\rangle$ 能级的原子是由于受到了空腔里的电磁场的诱导而产生的辐射，这就是前面我们已经研究过了的受激辐射，单位时间之内通过受激辐射从 $|f\rangle$ 能级跃迁到 $|i\rangle$ 能级的原子数是 $B_{if}u(\omega_{fi})N_f$。因此，单位时间之内从 $|f\rangle$ 能级跃迁到 $|i\rangle$ 能级的总原子数是 $(A_{if} + B_{if}u(\omega_{fi}))N_f$。很显然，如果 $|i\rangle$ 到 $|f\rangle$ 的跃迁能和 $|f\rangle$ 到 $|i\rangle$ 的跃迁相平衡，即

$$B_{fi}u(\omega_{fi})N_i = (A_{if} + B_{if}u(\omega_{fi}))N_f, \tag{10.59}$$

那么系统就能维持在热平衡状态。

另一方面，根据统计物理的玻尔兹曼定律，在热平衡时，能量为 E 的原子数目将正比于 $\exp\left(-\frac{E}{kT}\right)$，式中 k 为玻尔兹曼常数，T 为热平衡温度。因此

$$N_f/N_i = \exp\left(-\frac{E_f - E_i}{kT}\right) = \exp\left(-\frac{\hbar\omega_{fi}}{kT}\right). \tag{10.60}$$

而且，根据普朗克黑体辐射公式，

$$u(\omega_{fi}) = \frac{\hbar\omega_{fi}^3}{\pi^2 c^3}\frac{1}{e^{\frac{\hbar\omega_{fi}}{kT}}-1}. \tag{10.61}$$

将这些结果代入 (10.59) 就可以得到

$$A_{if} = \frac{\hbar\omega_{fi}^3}{\pi^2 c^3}\frac{1}{e^{\frac{\hbar\omega_{fi}}{kT}}-1}\left[B_{fi}\exp\left(\frac{\hbar\omega_{fi}}{kT}\right) - B_{if}\right]. \tag{10.62}$$

由于这些系数 A_{if}、B_{fi} 以及 B_{if} 都与温度无关。因此为了让上式对于任何温度都成立，就必定有

$$B_{fi} = B_{if},\ A_{if} = \frac{\hbar\omega_{fi}^3}{\pi^2 c^3}B_{if}. \tag{10.63}$$

其中，第一个结果我们已经通过微扰论推导出来了。而更重要的是第二个结果，它使得我们可以避开对电磁场的量子化，直接从对 B_{fi} 的计算中得出自发辐射系数 A_{if}。

前面我们谈到受激辐射所发出的光有很好的相干性，这使得人们自然想到利用受激辐射，的确，激光就是这一想法的产物。然而要想制造出激光，我们还需要克服一些困难。第一个困难是，在通常的条件下，受激辐射其实很微弱，原子发光主要靠自发辐射。人们可以通过比较自发辐射的跃迁概率 A_{if} 和受激辐射的跃迁概率 $B_{if}u(\omega_{fi})$ 来理解这一点，根据爱因斯坦的模型导出来的结果，在热平衡时我们有

$$\frac{A_{if}}{B_{if}u(\omega_{fi})} = e^{\frac{\hbar\omega_{fi}}{kT}} - 1 = 2^{\omega_{fi}/\omega_c(T)} - 1. \tag{10.64}$$

式中，$\omega_c(T) = \frac{kT}{\hbar}\log 2$，当 $\omega_{fi} = \omega_c$ 时，这两个跃迁概率相等。当 ω_{fi} 大于 ω_c 时，自发辐射的概率与受激辐射概率的比值按照 $\omega_{fi}/\omega_c(T)$ 的指数关系增长。如果取 T 为室温的 $300K$, 则 $\omega_c = 2.74\times 10^{13}Hz$, 相应于波长 $\lambda_c = 6.9\times 10^{-5}m$。而可见光的波长在几百个纳米，即大约是 $10^{-2}\lambda_c$, 因此对于室温下的可见光辐射，自发辐射的概率约是受激辐射概率的 $2^{10^2} = 2^{100}$ 倍，因此相比来说，受激辐射完全可以忽略。

由于受激辐射的概率正比于电磁场的能量密度，因此要增加受激辐射的概率就需要增加辐射场的能量密度。在激光器中，这一点是通过一个谐振腔来实现的，就是让受激辐射出来的相干光不断地在谐振腔中谐振，一个光子通过受激辐射诱导出两个相干光子，两个相干光子再诱导出四个相干光子，如此不断谐振使得相干光子数不断倍增，从而实现一个强辐射场，其能量密度远大于热平衡时电磁场的能量密度。因此才能使得激光器中的受激辐射超过自发辐射占据主导地位。

但即使是这样，也还不能实现激光的输出，原因在于，工作物质中的原子不只会辐射光，它还会吸收光。前面我们已经知道，吸收的概率是 $B_{fi}u(\omega_{fi})$, 这个概率和受激辐射的概率是一样的。但是工作物质单位时间之内对光的吸收率等于这个概率乘以低能级的原子数 N_i，而辐射率则等于这个概率乘以高能级原子数 N_f。要实现激光的输出，就必须使得辐射率超过吸收率，也即要让 $N_f > N_i$。但是根据玻尔兹曼定律，在热平衡时，N_f 一定小于 N_i。因此要实现激光的输出也必须远离热平衡，实现高能级原子数 N_f 大于低能级原子

数 N_i 的粒子数反转，也叫做集居数反转 (population inversion)。用来实现集居数反转的特殊介质，叫做增益介质 (gain medium)，给增益介质提供能量使其反生集居数反转的能量源，叫做泵浦源 (pumping source)。但是在不同类型的激光器中，往往会有不同的泵浦方式来实现集居数反转，具体内容请读者参阅关于激光工作机理方面的专业书籍。

10.3 量子共振的一般性研究

这一节我们将在超越微扰论的意义上讨论离散能级之间的跃迁过程。对于离散能级，我们前面讨论过，微扰展开的处理办法适用的一个必要条件是 $|\langle m|V|n\rangle|t/\hbar \ll 1$, 其中 t 表示我们所考察的整个过程的时间尺度，或者也可以理解为微扰持续的时间，$|n\rangle$ 和 $|m\rangle$ 是两个相互跃迁的能级。这个必要条件意味着，仅在微扰相互作用比较微弱，并且时间尺度不过长的时候，我们才可以用前面发展的含时微扰论来处理问题。但是实际中有大量的含时演化问题超出了微扰论的这个适用范围，这时候我们又该如何处理问题呢?

我们还是假设系统 H 可以分成 H_0 和 V 两部分，并且 H_0 部分所描述的物理是我们已经了解清楚的，我们想讨论的是，含时相互作用 V 所引起的跃迁。因此，前面我们定义的相互作用绘景下的时间演化算符 $U_I(t,t_0)$ 依然适用，跃迁幅 $\langle m|U_I(t,t_0)|n\rangle$ 和跃迁概率的概念也依然成立，$U_I(t,t_0)$ 所满足的微分方程 (10.12) 也依然适用。但是，我们现在不再使用微扰展开的戴森级数 (10.15) 了。现在，我们直接以方程 (10.12) 为基本出发点。

10.3.1 共振近似 (旋波近似)

为了简单起见，我们不妨取初始时刻 $t_0 = 0$, 并将 $U_I(t,0)$ 简记为 $U_I(t) = U_I(t,0)$。假设系统的初态为 $|i\rangle$, 我们记 t 时刻系统跃迁到某个任意的 $|m\rangle$ 态的概率幅为 $b_m(t)$，因此

$$b_m(t) = \langle m|U_I(t)|i\rangle, \tag{10.65}$$

显然，$b_m(t)$ 满足初始条件 $b_m(0) = \delta_{mi}$。因此，t 时刻系统跃迁到 m 态的概率 $p_{i\to m}(t) = |b_m(t)|^2$。为了推导 $b_m(t)$ 所满足的演化方程，我们利用方程 (10.12)，从而即有，$\frac{db_m}{dt} = \langle m|\frac{d}{dt}U_I(t)|i\rangle = (-\frac{i}{\hbar})\langle m|V_I(t)U_I(t)|i\rangle = (-\frac{i}{\hbar})\sum_n\langle m|V_I(t)|n\rangle\langle n|U_I(t)|i\rangle = (-\frac{i}{\hbar})\sum_n\langle m|V_I(t)|n\rangle b_n(t)$, 即

$$\frac{db_m}{dt} = (-\frac{i}{\hbar})\sum_n\langle m|V_I(t)|n\rangle b_n(t) = (-\frac{i}{\hbar})\sum_n e^{i\omega_{mn}t}\langle m|V(t)|n\rangle b_n(t). \tag{10.66}$$

之后我们要做的就是直接求解这组方程。

下面我们以 $V = W^\dagger e^{i\omega t} + We^{-i\omega t}$ 为例来说明在不做微扰展开的情况下，如何作出必要的近似以简化方程组 (10.66), 当然我们假设算符 W 和 $W^\dagger$ 都不依赖于时间 t。对于这种类型的含时相互作用 V, 方程组 (10.66) 变成

$$\frac{db_m}{dt} = (-\frac{i}{\hbar})\sum_n\left[e^{i(\omega_{mn}+\omega)t}W^*_{nm}b_n(t) + e^{i(\omega_{mn}-\omega)t}W_{mn}b_n(t)\right], \tag{10.67}$$

式中 $W_{mn} = \langle m|W|n\rangle$, 并且我们利用了 $(W^\dagger)_{mn} = W^*_{nm}$。现在，我们将注意力集中在一种共振情形，即假设微扰的振荡频率 ω 接近于某个 $\omega_{fi} = (E_f - E_i)/\hbar$, 即 $\omega \sim \omega_{fi}$(具体来说，我们

假设 $|\omega-\omega_{fi}|\ll\omega$), 式中 i 态就是我们的初态, f 态表示的是某个将与其共振的态, 这里我们假设 $E_f>E_i$, 因此 $\omega_{fi}>0$, 对于 $E_f<E_i$ 的情形，人们可以类似地考虑 $\omega\sim\omega_{if}=-\omega_{fi}$。并且，我们进一步假设这样的共振能级 f 是唯一的，即对于所有 $(m,n)\neq(f,i)\cup(i,f)$ 都有 $|\omega\pm\omega_{mn}|\sim\omega$。

在上面的近似条件之下，人们很容易看出，方程组 (10.67) 所包含的所有项中，只有随时间的振荡频率为 $\omega_{if}+\omega$ 和 $\omega_{fi}-\omega$ 的项相对来说是振荡得很缓慢的，其余的所有项相对来说都随着时间振荡得比较快。因此假设我们考察一个很长的时间, $t\gg1/\omega$, 那么将只有那些缓慢振荡项是决定性的，而所有的快速振荡项都会因为多次振荡而抵消掉了。因此，我们可以直接忽略方程组 (10.67) 中的所有快速振荡项，进行得到下面的简化方程组，

$$\begin{aligned}\frac{db_i}{dt}&=-i\frac{W_{fi}^*}{\hbar}e^{i(\omega-\omega_{fi})t}b_f(t)\\ \frac{db_f}{dt}&=-i\frac{W_{fi}}{\hbar}e^{i(\omega_{fi}-\omega)t}b_i(t).\end{aligned}\tag{10.68}$$

可见，在我们的共振近似下，只有相互共振的两个能级，i 和 f 是相关的，问题简化为求解由方程 (10.68) 所描述的二能级系统，由于 i 是演化初态，所以初始条件是 $b_i(0)=1$, $b_f(0)=0$。

10.3.2 Jaynes-Cummings 模型

下面我们以一个更具体的物理例子来演示方程 (10.68) 的物理内涵，以及对它的求解过程如何进行。我们要研究的这个物理例子描述的是一个原子的二能级与激光场耦合共振的情形。我们把这两个能级中的低能级称作“基态” (ground state), 记为 $|g\rangle$，把二能级中的高能级称作激发态 (excited state)，记为 $|e\rangle$，假设 $|e\rangle$ 态和 $|g\rangle$ 态的能级差为 $\hbar\omega_0$。假定激光场的频率为 ω, 由于电磁场本身是一种简谐振动，所以激光场可以用一个谐振子来描写（人们可以严格地通过将共振腔内的电磁场量子化来得到类似的结果），这个谐振子的第 n 能级我们记为 $|n\rangle$, 它描述的是 n 个光子处在完全相同的量子态，每个光子的能量为 $\hbar\omega$, 所以 $|n\rangle$ 态的能量为 $n\hbar\omega$(由于零点能和我们将要进行的讨论无关，所以我们忽略它)，多个光子相干性地处于完全相同的量子态正是激光的典型特征，所以我们才说与原子耦合的是一个激光场。我们把描述激光场的这个谐振子的产生湮灭算符分别记为 $a^\dagger$,a，相应的粒子数算符 $N=a^\dagger a$ 描述的当然就是光子数，因此 $N|n\rangle=n|n\rangle$。在物理上，我们考察的这个原子和激光场的耦合系统需要把原子放置在量子光场的共振腔中才能实现，它属于量子光学的基本研究内容。

描述这个二能级原子与激光场耦合的最简单模型是所谓的 Jaynes-Cummings 模型，它是量子光学中一个极为重要也非常基本的模型。Jaynes-Cummings 模型的哈密顿量是，

$$H=\hbar\omega_0\frac{\sigma_z}{2}+\hbar\omega a^\dagger a+\lambda\hbar(a^\dagger\sigma_-+a\sigma_+).\tag{10.69}$$

式中 $\sigma_z=|e\rangle\langle e|-|g\rangle\langle g|$, 即满足 $\sigma_z|g\rangle=-|g\rangle$, $\sigma_z|e\rangle=|e\rangle$, 因此 $\hbar\omega_0\frac{\sigma_z}{2}$ 正好具有 $\hbar\omega_0$ 的能级差，而哈密顿量中的线性谐振子项 $\hbar\omega a^\dagger a$ 当然描述的就是激光场。另外，$\sigma_+=|e\rangle\langle g|$, $\sigma_-=|g\rangle\langle e|$, 即满足 $\sigma_+|g\rangle=|e\rangle$, $\sigma_-|e\rangle=|g\rangle$, 以及 $\sigma_+^2=\sigma_-^2=0$，也就是说，σ_+ 的作用是描述原子从“基态”到激发态的跃迁，而 σ_- 描述的是原子从激发态到“基态”的跃迁。而且，

从这几个算符的定义很容易得到 $\sigma_z^2 = |e\rangle\langle e| + |g\rangle\langle g| = 1$(这里我们利用了两能级系统的完备性关系), 以及 $\sigma_-\sigma_z = \sigma_-$, $\sigma_z\sigma_- = -\sigma_-$，从而 $[\sigma_z, \sigma_-] = -2\sigma_-$, 类似的，$[\sigma_z, \sigma_+] = 2\sigma_+$。总之，$\sigma_z, \sigma_+, \sigma_-$ 这几个算符所满足的代数关系和我们很熟悉的泡利算符 $\sigma_z, \sigma^\dagger, \sigma$ 其实是完全一样的，只不过这里我们将 σ 记成了 σ_-，将 $\sigma^\dagger$ 记成了 σ_+。

哈密顿量 (10.69) 中的 $\lambda\hbar(a^\dagger\sigma_- + a\sigma_+)$ 描述的就是原子二能级与激光场的相互作用，$\lambda\hbar$ 表示这种相互作用的强度，$a^\dagger\sigma_-$ 描述的是原子从激发态跃迁到 “基态”，同时放出一个光子 (因此光场的光子数增加 1)，产生算符 $a^\dagger$ 描述的正是这个光子数增加 1 的作用。相反，$a\sigma_+$ 描述的是原子从 “基态” 跃迁到激发态，同时吸收一个光子 (因此光子数减少 1)，湮灭算符 a 描述的正是这个光子数减少 1 的作用。如果我们假设整个系统最初处在 $|n, g\rangle$ 态，即初始时光场处在能级 $|n\rangle$，原子处在 “基态” $|g\rangle$, 那么通过原子与光场的相互作用，系统可以通过吸收一个光子跃迁到 $|n-1, e\rangle$ 态，当然，原子与光场的相互作用也可能反过来使得系统从 $|n-1, e\rangle$ 态跃迁到 $|n, g\rangle$ 态。因此，一般来说，Jaynes-Cummings 的哈密顿量 (10.69) 会描述一个 $|n, g\rangle$ 态与 $|n-1, e\rangle$ 态之间的振荡，这就是著名的拉比振荡，我们可以示意性地表示为

$$|n, g\rangle \rightleftharpoons |n-1, e\rangle. \tag{10.70}$$

下面我们就是要对这个拉比振荡过程进行精确的数学分析，我们假定系统的初态是 $|n, g\rangle$ 态。

值得说明的是，Jaynes-Cummings 模型可以从量子光场与原子相互作用的一般性理论中通过一个前面所说的共振近似推导出来，人们也常常称这样的近似为旋波近似 (rotating wave approximation)。

我们取 $H_0 = \hbar\omega_0\frac{\sigma_z}{2} + \hbar\omega a^\dagger a$, $V = \lambda\hbar(a^\dagger\sigma_- + a\sigma_+)$。很显然，$|n, g\rangle$ 态和 $|n-1, e\rangle$ 态都是 H_0 的本征态。为了应用基本方程 (10.66), 我们需要计算出相互作用绘景中的 $V_I(t) = e^{i\frac{H_0}{\hbar}t}Ve^{-i\frac{H_0}{\hbar}t}$。为了将 $V_I(t)$ 的具体表达式计算出来，我们首先注意到原子的算符和光场的算符是相互对易的，因为原子和光场是两个不同的子系统。另外，我们再注意到下面的恒等式

$$e^{i\theta N}ae^{-i\theta N} = e^{-i\theta}a,\ e^{i\theta N}a^\dagger e^{-i\theta N} = e^{i\theta}a^\dagger. \tag{10.71}$$

这两个恒等式互为厄米共轭关系，因此只需证明其中的一个就够了，在这里我们将给出第一个恒等式的直接证明。为此，我们定义一簇依赖于参数 θ 的算符 $F(\theta)$, $F(\theta) = e^{i\theta N}ae^{-i\theta N}$, 很显然 $F(0) = a$。另一方面，我们又有 $\frac{d}{d\theta}F(\theta) = ie^{i\theta N}Nae^{-i\theta N} - ie^{i\theta N}aNe^{-i\theta N} = ie^{i\theta N}[N, a]e^{-i\theta N} = -ie^{i\theta N}ae^{-i\theta N} = -iF(\theta)$, 式中倒数第二个等号我们利用了代数关系 $[N, a] = -a$。也就是说，$F(\theta)$ 满足微分方程 $\frac{d}{d\theta}F(\theta) = -iF(\theta)$, 而这一微分方程有唯一解 $F(\theta) = F(0)e^{-i\theta} = ae^{-i\theta}$，因此这就证明了我们需要的恒等式。类似的，我们也可以证明下面两个恒等式

$$e^{i\theta\frac{\sigma_z}{2}}\sigma_-e^{-i\theta\frac{\sigma_z}{2}} = e^{-i\theta}\sigma_-,\ e^{i\theta\frac{\sigma_z}{2}}\sigma_+e^{-i\theta\frac{\sigma_z}{2}} = e^{i\theta}\sigma_+, \tag{10.72}$$

不过为了证明这两个恒等式我们需要用到的代数关系是 $[\sigma_z, \sigma_-] = -2\sigma_-$。

根据 $V_I(t)$ 的定义，再利用上面的这些恒等式，我们可以算出

$$V_I(t) = \lambda\hbar(a\sigma_+ e^{i2\delta t} + a^\dagger\sigma_- e^{-i2\delta t}), \tag{10.73}$$

式中 $2\delta = \omega_0 - \omega$, 通常被称为失谐。注意到 $a|n\rangle = \sqrt{n}|n-1\rangle$, $a^\dagger|n-1\rangle = \sqrt{n}|n\rangle$, 人们很容易计算出 $V_I(t)$ 的两个非 0 矩阵元，为

$$\langle n-1,e|V_I|n,g\rangle = \sqrt{n}\lambda\hbar e^{i2\delta t},\ \langle n,g|V_I|n-1,e\rangle = \sqrt{n}\lambda\hbar e^{-i2\delta t}. \tag{10.74}$$

定义跃迁概率幅 $b_g(t) = \langle n,g|U_I(t)|n,g\rangle$, $b_e(t) = \langle n-1,e|U_I(t)|n,g\rangle$，它们满足初始条件 $b_e(0) = 0$, $b_g(0) = 1$。由一般性的方程 (10.66) 和刚才计算出来的 $V_I(t)$ 的非 0 矩阵元，我们有如下方程

$$\begin{aligned}\frac{db_g}{dt} &= -i\sqrt{n}\lambda e^{-i2\delta t}b_e(t),\\ \frac{db_e}{dt} &= -i\sqrt{n}\lambda e^{i2\delta t}b_g(t).\end{aligned} \tag{10.75}$$

很显然，这个方程组和我们前面在共振近似下得到的一般性方程组 (10.68) 有完全类似的数学结构，这表明 Jaynes-Cummings 模型所描写的光场和二能级系统共振的物理图像是普适的，并且我们只需要求解出 Jaynes-Cummings 模型的方程组 (10.75) 就能很容易地将解推广到一般情形 (10.68)。

为了求解上面的方程组 (10.75)，我们引入变量 $z_+(t)$ 和 $z_-(t)$，它们的定义是 $z_+(t) = e^{i\delta t}b_g(t) + ie^{-i\delta t}b_e(t)$, $z_-(t) = e^{i\delta t}b_g(t) - ie^{-i\delta t}b_e(t)$，很显然，它们满足初始条件 $z_+(0) = z_-(0) = 1$。根据方程 (10.75)，我们可以推导出新变量 $z_+(t)$ 和 $z_-(t)$ 所满足的微分方程，

$$\frac{dz_+}{dt} = (\sqrt{n}\lambda + i\delta)z_-(t),\ \frac{dz_-}{dt} = -(\sqrt{n}\lambda - i\delta)z_+(t). \tag{10.76}$$

由这两个方程我们很容易看出 $z_\pm(t)$ 均满足一个典型的谐振子方程，

$$\frac{d^2z_\pm}{dt^2} + \Omega_n^2 z_\pm(t) = 0,\ \Omega_n = \sqrt{n\lambda^2 + \delta^2}, \tag{10.77}$$

Ω_n 就是这个谐振子的振荡频率，在这里称为拉比频率。由于 $z_\pm(0) = 1$, 从而可知必有 $z_\pm(t) = \cos(\Omega_n t) + c_\pm\sin(\Omega_n t)$, 式中 $c_\pm$ 为两个待定常数。将这个通解代入方程 (10.76)，就可以得到 $c_+ = (\sqrt{n}\lambda + i\delta)/\Omega_n$, $c_- = -(\sqrt{n}\lambda - i\delta)/\Omega_n$。这样我们就完全确定了 $z_\pm(t)$，反过来我们就能得到 $b_g(t)$ 和 $b_e(t)$，它们为

$$e^{i\delta t}b_g(t) = \cos(\Omega_n t) + i\frac{\delta}{\Omega_n}\sin(\Omega_n t),\ e^{-i\delta t}b_e(t) = -i\frac{\sqrt{n}\lambda}{\Omega_n}\sin(\Omega_n t). \tag{10.78}$$

利用这个最终解我们就可以计算出 t 时刻，系统从 $|n,g\rangle$ 态跃迁到 $|n-1,e\rangle$ 态的概率 $p_{g\to e}(t) = |b_e(t)|^2$, 为

$$p_{g\to e}(t) = \frac{n\lambda^2}{n\lambda^2 + \delta^2}\sin^2(\Omega_n t), \tag{10.79}$$

从这个概率随时间 t 的振荡行为中，我们可以很清楚地看到，系统是在 $|n,g\rangle$ 态和 $|n-1,e\rangle$ 态之间振荡，这就是拉比振荡，$\Omega_n = \sqrt{n\lambda^2 + \delta^2} = \sqrt{n\lambda^2 + (\frac{\omega_0-\omega}{2})^2}$ 就是拉比振荡的频率。

为了推广拉比振荡的结果，我们将方程组 (10.68) 和方程组 (10.75) 进行类比。很显然，$|n,g\rangle$ 态对应于 $|i\rangle$ 态，$|n-1,e\rangle$ 态对应于 $|f\rangle$ 态，$\omega_0-\omega$ 对应于 $\omega_{fi}-\omega$, $n\lambda^2$ 对应于 $|\frac{W_{fi}}{\hbar}|^2$。因此我们可以将上面的跃迁概率 $p_{g\to e}(t)$ 一般性地推广到 $p_{i\to f}(t)$, 将公式 (10.79) 推广成

$$p_{i\to f}(t)=\frac{|\frac{W_{fi}}{\hbar}|^2}{|\frac{W_{fi}}{\hbar}|^2+(\frac{\omega_{fi}-\omega}{2})^2}\sin^2(\sqrt{|\frac{W_{fi}}{\hbar}|^2+(\frac{\omega_{fi}-\omega}{2})^2}\cdot t), \tag{10.80}$$

人们通常称这个一般性的公式为拉比公式。很容易看出，当 $|W_{fi}|\ll\hbar|\omega_{fi}-\omega|$ 时，拉比公式就近似成为一阶微扰论给出的公式 (10.44)。但是反过来，只要时间足够长，即使 $\omega=\omega_{fi}$，$|i\rangle$ 态和 $|f\rangle$ 态之间的正弦振荡也依然存在，这种振荡是一阶微扰论的结果 (10.44)(令 $\omega=\omega_{fi}$) 解释不了的。

10.4 * 散射的一般理论

10.4.1 散射的一般理论和 S 矩阵

在恒定微扰一节中，我们用一阶恒定微扰讨论了一个自由入射粒子在一个固定势场中的散射问题。其实可以用恒定微扰讨论更一般的散射问题，比方说多粒子散射问题。也可以不限于一阶微扰，可以将微扰散射理论发展到任意阶。发展这样的一个关于散射的一般理论就是我们这一节将要处理的内容。

由于入射粒子和出射粒子的能量和动量都可以连续取值，因此散射问题中所涉及的入射粒子量子态和出射粒子量子态都是连续谱，我们不妨称这些量子态为散射态。散射的确说明连续谱的量子态可以跃迁到连续谱，但是类似于本章最后一节的证明同样可以说明，只要末态是连续谱那么系统就不会振荡回初态。在物理直观上这是显然的，因为出射粒子当然不可能自己跃迁回入射状态。因此，根据前面关于末态连续谱的讨论，我们完全可以认为，散射是在从 $-\infty$ 到 $+\infty$ 的时间内发生的。当然，散射相互作用真正有效起作用的时间在宏观时间尺度来看其实很短，只是在微观时间尺度上看充分长而已。

对于无穷远过去的散射初态，以及无穷远将来的散射末态，粒子间的距离都很远，相互作用 V(这里假定 V 不显含时间) 可以忽略，因此作为初态的入射粒子 (多个粒子) 和作为末态的出射粒子 (也是多个粒子) 都可以看成是哈密顿量 H_0 的本征态，H_0 可以认为是多个自由粒子的哈密顿量之和。由于这些 H_0 的本征态是散射态连续谱，所以我们用 $|\alpha\rangle$ 这样的连续指标来标记，相应的本征值记为 E_α，也即是说

$$H_0|\alpha\rangle=E_\alpha|\alpha\rangle. \tag{10.81}$$

注意，$|\alpha\rangle$ 一般来说是多粒子态，描写多个自由粒子，并且它满足标准的 δ 函数归一化，

$$\langle\alpha|\beta\rangle=\delta(\alpha-\beta). \tag{10.82}$$

值得注意的是，α、β 其实只是抽象的记号，它们都可以是好多指标的总体，当然 α、β 所代表的指标中必须要包含一些连续指标，但是我们其实也可以允许 α、β 中包含有一些离

散指标，这些指标应该理解成是归一化到克龙内克 δ 符号，而对这些指标的积分就应该理解成求和。

我们想要计算的同样是 $-\infty$ 时刻的 $|\alpha\rangle$ 态在 $+\infty$ 时刻跃迁到 $|\beta\rangle$ 态的概率密度 $p_{\alpha\to\beta}$，为了强调指标的连续性，我们也常常将它写作 $p(\alpha\to\beta)$。根据本章第一节中的一般性讨论，$p(\alpha\to\beta)=|\langle\beta|U_I(+\infty,-\infty)|\alpha\rangle|^2$。其中幺正算符 $U_I(+\infty,-\infty)$ 习惯上记作 S，

$$S=U_I(+\infty,-\infty), \tag{10.83}$$

它可以按照戴森级数 (10.15) 的方法来进行微扰展开，我们将这个级数重写如下，

$$\begin{aligned}S=&1+(-\frac{i}{\hbar})\int_{-\infty}^{+\infty}V_I(t_1)dt_1\\&+(-\frac{i}{\hbar})^2\int_{-\infty}^{+\infty}dt_2\int_{-\infty}^{t_2}dt_1V_I(t_2)V_I(t_1)\\&+(-\frac{i}{\hbar})^3\int_{-\infty}^{+\infty}dt_3\int_{-\infty}^{t_3}dt_2\int_{-\infty}^{t_2}dt_1V_I(t_3)V_I(t_2)V_I(t_1)+\ldots\end{aligned} \tag{10.84}$$

根据上面这个戴森级数，我们来逐级地计算一下跃迁幅 $\langle\beta|U_I(+\infty,-\infty)|\alpha\rangle$。很显然在 V 的 0 阶近似上，我们有 $\langle\beta|1|\alpha\rangle=\delta(\beta-\alpha)$。注意到 $V_I(t)=e^{i\frac{H_0}{\hbar}t}Ve^{-i\frac{H_0}{\hbar}t}$，我们也很容易计算现在应该已经比较熟悉了的一阶微扰修正 $(-\frac{i}{\hbar})\int_{-\infty}^{+\infty}\langle\beta|V_I(t_1)|\alpha\rangle dt_1=(-\frac{i}{\hbar})\int_{-\infty}^{+\infty}e^{i\omega_{\beta\alpha}t}\langle\beta|V|\alpha\rangle dt_1=-2\pi i\delta(E_\beta-E_\alpha)\langle\beta|V|\alpha\rangle$，式中 $\omega_{\beta\alpha}=(E_\beta-E_\alpha)/\hbar$。即一阶微扰的结果是

$$-2\pi i\delta(E_\beta-E_\alpha)\langle\beta|V|\alpha\rangle. \tag{10.85}$$

对于微扰级数的二阶 $(-\frac{i}{\hbar})^2\int_{-\infty}^{+\infty}dt_2\int_{-\infty}^{t_2}dt_1\langle\beta|V_I(t_2)V_I(t_1)|\alpha\rangle$，利用散射态的封闭性关系 $\int d\alpha'|\alpha'\rangle\langle\alpha'|=1$，我们有

$$\begin{aligned}&(-\frac{i}{\hbar})^2\int_{-\infty}^{+\infty}dt_2\int_{-\infty}^{t_2}dt_1\langle\beta|V_I(t_2)V_I(t_1)|\alpha\rangle\\=&(-\frac{i}{\hbar})^2\int d\alpha'\int_{-\infty}^{+\infty}dt_2\int_{-\infty}^{t_2}dt_1\langle\beta|V_I(t_2)|\alpha'\rangle\langle\alpha'|V_I(t_1)|\alpha\rangle\\=&(-\frac{i}{\hbar})^2\int d\alpha'\int_{-\infty}^{+\infty}dt_2\int_{-\infty}^{t_2}dt_1e^{i\omega_{\beta\alpha't_2}}e^{i\omega_{\alpha'\alpha t_1}}\langle\beta|V|\alpha'\rangle\langle\alpha'|V|\alpha\rangle.\end{aligned} \tag{10.86}$$

下面将这个式子中对时间变量的积分都算出来，首先要计算 $\int_{-\infty}^{t_2}dt_1e^{i\omega_{\alpha'\alpha}t_1}$，但这个积分在 $t_1\to-\infty$ 时是振荡的，为了让积分在 $t_1\to-\infty$ 时收敛，我们可以给 $\omega_{\alpha'\alpha}$ 加一个很小的负虚部，即做替换 $\omega_{\alpha'\alpha}\to\omega_{\alpha'\alpha}-i\varepsilon$，式中 $\varepsilon>0$ 为一个无穷小量，并且我们将在对所有物理可观测量的计算都完成以后令 $\varepsilon\to0$。经过这样的手续以后，就有 $\int_{-\infty}^{t_2}dt_1e^{i(\omega_{\alpha'\alpha}-i\varepsilon)t_1}=\frac{1}{i(\omega_{\alpha'\alpha}-i\varepsilon)}e^{i(\omega_{\alpha'\alpha}-i\varepsilon)t_2}=i\hbar\frac{1}{E_\alpha-E_{\alpha'}+i\varepsilon}e^{i(\omega_{\alpha'\alpha}-i\varepsilon)t_2}$。将这个结果代入上面的 (10.86) 式，并利用 $\omega_{\beta\alpha'}+\omega_{\alpha'\alpha}=\omega_{\beta\alpha}$，就有

$$\begin{aligned}&(-\frac{i}{\hbar})\int d\alpha'\int_{-\infty}^{+\infty}dt_2e^{i(\omega_{\beta\alpha}-i\varepsilon)t_2}\langle\beta|V|\alpha'\rangle\frac{1}{E_\alpha-E_{\alpha'}+i\varepsilon}\langle\alpha'|V|\alpha\rangle\\=&-2\pi i\delta(E_\beta-E_\alpha)\int d\alpha'\langle\beta|V|\alpha'\rangle\frac{1}{E_\alpha-E_{\alpha'}+i\varepsilon}\langle\alpha'|V|\alpha\rangle\\=&-2\pi i\delta(E_\beta-E_\alpha)\langle\beta|V\frac{1}{E_\alpha-H_0+i\varepsilon}V|\alpha\rangle.\end{aligned} \tag{10.87}$$

完全类似的，我们也可以计算出三阶微扰，结果如下

$$
\begin{aligned}
&(-\frac{i}{\hbar})^3\int_{-\infty}^{+\infty}dt_3\int_{-\infty}^{t_3}dt_2\int_{-\infty}^{t_2}dt_1\langle\beta|V_I(t_3)V_I(t_2)V_I(t_1)|\alpha\rangle\\
=&-2\pi i\delta(E_\beta-E_\alpha)\langle\beta|V\frac{1}{E_\alpha-H_0+i\varepsilon}V\frac{1}{E_\alpha-H_0+i\varepsilon}V|\alpha\rangle.
\end{aligned}
\tag{10.88}
$$

做了这么多推导以后，人们应该很容易归纳出规律，进而可以写出微扰展开的任意阶了。将这些结果放在一起，即有

$$
\langle\beta|S|\alpha\rangle=\delta(\beta-\alpha)-2\pi i\delta(E_\beta-E_\alpha)\langle\beta|T|\alpha\rangle,
\tag{10.89}
$$

其中算符 T 称作 T 算符，它由下面的无穷级数定义

$$
T=V+V\frac{1}{E_\alpha-H_0+i\varepsilon}V+V\frac{1}{E_\alpha-H_0+i\varepsilon}V\frac{1}{E_\alpha-H_0+i\varepsilon}V+...
\tag{10.90}
$$

有时候，人们也常常记 $E_\alpha+i\varepsilon=E_\alpha^+$，很显然，$T$ 算符的定义依赖于 E_α^+，所以也常常记作 $T(E_\alpha^+)$。

幺正算符 S 的矩阵元 $\langle\beta|S|\alpha\rangle$ 通常称作 S 矩阵，记作 $S_{\beta\alpha}=\langle\beta|S|\alpha\rangle$。类似的，$T$ 算符的矩阵元通常叫做 T 矩阵，记作 $T_{\beta\alpha}=\langle\beta|T|\alpha\rangle$。有了 T 矩阵和 S 矩阵，我们就可以计算跃迁概率密度 $p(\alpha\to\beta)$，$p(\alpha\to\beta)=|\langle\beta|S|\alpha\rangle|^2$。如果我们考察初态 $|\alpha\rangle$ 跃迁到一个不同的态 $|\beta\rangle$ 的概率，那 S 矩阵的 0 阶项显然是 0，因此根据 (10.89) 式，跃迁概率实际上完全由 T 矩阵给出，$p(\alpha\to\beta)=(\frac{2\pi}{\hbar})^2(\delta(\omega_{\beta\alpha}))^2|\langle\beta|T|\alpha\rangle|^2$，这里我们利用了 $\delta(E_\beta-E_\alpha)=\delta(\omega_{\beta\alpha})/\hbar$。进一步利用数学公式 $(\delta(\omega_{\beta\alpha}))^2=\delta(\omega_{\beta\alpha})\cdot t/(2\pi)$，式中 t 是从初态跃迁到末态的整个时间，它趋于无穷大。容易有，$p(\alpha\to\beta)=(\frac{2\pi}{\hbar})\delta(E_\beta-E_\alpha)|\langle\beta|T|\alpha\rangle|^2\cdot t$，或者说，单位时间的跃迁概率密度 $w(\alpha\to\beta)=dp(\alpha\to\beta)/dt$ 是

$$
w(\alpha\to\beta)=\frac{2\pi}{\hbar}|\langle\beta|T|\alpha\rangle|^2\delta(E_\beta-E_\alpha).
\tag{10.91}
$$

很显然，这个结果实际上就是费米黄金规则的推广。由 T 算符的展开式 (10.90) 可以看到，如果只考虑微扰论的一阶，那 (10.91) 式就回到了我们前面得出的费米黄金规则。正因为 T 矩阵元 $T_{\beta\alpha}$ 的模方决定了散射概率，所以有时候也把它称为散射幅。

假如我们考察的是单个入射粒子在一个固定势场 (这个固定势场的作用可以和粒子自旋有关) 中的散射问题，那么我们依然可以定义微分散射截面 (11.76)，一切都和前面在一阶微扰中对势散射问题的讨论一样，乘上末态态密度以后，就有

$$
d\sigma=(2\pi\hbar)^6\left(\frac{m}{2\pi\hbar^2}\right)^2|\langle\mathbf{p}_f,s_f|T|\mathbf{p}_i,s_i\rangle|^2d\Omega.
\tag{10.92}
$$

其中的 s_i 表示入射粒子自旋，s_f 表示出射粒子自旋。公式前面的因子 $(2\pi\hbar)^6$ 完全是因为我们将入射和出射的自由粒子平面波取成了 $\frac{1}{(2\pi\hbar)^{3/2}}e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$，如果将平面波都取成 $e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$ 的形式，那这个因子就不存在了。

10.4.2 S 矩阵的解析性和幺正性

很明显，$T(E_\alpha^+)$ 的定义式 (10.90) 中还出现了一个算符，常常记作 $G_0(E^+)$，它的定义是

$$G_0(E^+)=\frac{1}{E-H_0+i\varepsilon}. \tag{10.93}$$

值得注意的是，由于在散射态连续谱的希尔伯特空间上，算符 $E-H_0$ 有 0 本征值 (对应 H_0 的本征值为 E 的那个散射态)，所以它实际上不可逆，也就是说 $1/(E-H_0)$ 没有定义，但是给 E 加上一个虚部 $+i\varepsilon$ 以后就避开了这个问题，所以 $G_0(E^+)$ 是有定义的。实际上，人们还常常将 $G_0(E^+)$ 的定义解析延拓到整个能量复平面上除了 H_0 的散射谱之外的区域 (在散射理论中，H_0 常常只有散射谱，也即只有散射态对应的连续谱)，记为 $G_0(z)$，

$$G_0(z)=\frac{1}{z-H_0}. \tag{10.94}$$

因此，H_0 的散射谱对应解析函数 $G_0(z)$ 在实轴上的一条割线 (即要从解析函数 $G_0(z)$ 的定义域中切除的线)，如图 (10.4) 所示。

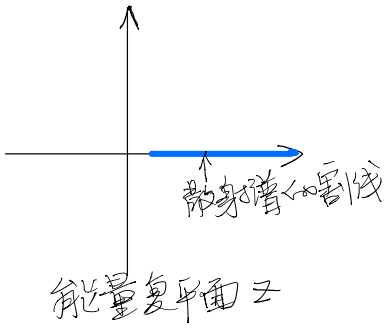

Figure 10.4: 将能量延拓成复平面，散射谱是实轴上的一条割线。

类似的，我们也可以将 $T(E^+)$ 延拓到能量复平面上得到 $T(z)$，

$$\begin{aligned}T(z)&=V+VG_0(z)V+VG_0(z)VG_0(z)V+....\\&=V+V[G_0(z)+G_0(z)VG_0(z)+G_0(z)VG_0(z)VG_0(z)....]V.\end{aligned} \tag{10.95}$$

显然，$G_0(z)$ 在实轴上的割线也是 $T(z)$ 的割线。不过当超越微扰论，将微扰级数的所有阶都加起来以后，$T(z)$ 在能量复平面上的解析结构比这要复杂一些。利用下面的算符微扰恒等式

$$\frac{1}{A-B}=\frac{1}{A}+\frac{1}{A}B\frac{1}{A}+\frac{1}{A}B\frac{1}{A}B\frac{1}{A}+... \tag{10.96}$$

我们容易有

$$\begin{aligned}&G_0(z)+G_0(z)VG_0(z)+G_0(z)VG_0(z)VG_0(z)....\\=&\frac{1}{z-H_0}+\frac{1}{z-H_0}V\frac{1}{z-H_0}+\frac{1}{z-H_0}V\frac{1}{z-H_0}V\frac{1}{z-H_0}+...\\=&\frac{1}{z-H_0-V}=\frac{1}{z-H}=G(z),\end{aligned} \tag{10.97}$$

上式的最后一个等号是 $G(z)$ 的定义。由于 $G(z)$ 是 $G_0(z)$ 的一个无穷级数，因此 $G_0(z)$ 在实轴上的割线也是 $G(z)$ 的割线，这实际上就保证了 H_0 的散射谱和 H 的散射谱 (散射态的连续谱) 是一一对应的。但是，现在包含相互作用以后的哈密顿算符 H 可能不只是有散射谱了，它还可能有束缚态，这些束缚态就对应 $G(z)=(z-H)^{-1}$ 的极点 (如图 (10.5) 所示)。另外，将上面的结果 (10.97) 代入 $T(z)$ 的定义式 (10.95), 就有

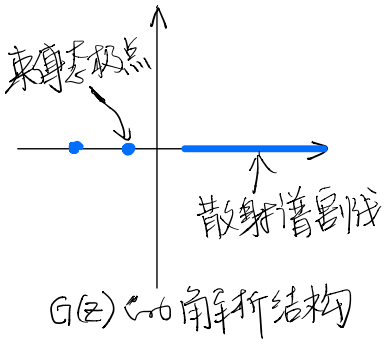

Figure 10.5: $G(z)$ 的解析结构，束缚态对应实轴上的极点。

$$T(z)=V+VG(z)V. \tag{10.98}$$

从这个表达式可以看出，$G(z)$ 的极点也是 $T(z)$ 的极点。前面我们已经看到，$T(E_\alpha^+)$ 的矩阵元决定了散射概率，因此这也就是说，我们可以从散射概率关于入射粒子总能量的极点结构中读出哈密顿量 H 的束缚态信息。因此，散射实验也可以探测到系统的束缚态信息。只不过，因为需要把所有阶的微扰级数都加起来，所以微扰计算实际上发现不了这些束缚态，为了得到这些束缚态的信息，我们需要用非微扰的方式计算 $T(z)$ 算符。

以上关于 $T(z)$ 解析性的讨论主要是在算符的意义上进行的，所有的割线和极点都作为哈密顿算符的本征值而出现，由于哈密顿算符的本征值一定是实数，因此这些割线和极点都一定在实轴上。但是，如果我们超出算符本征值的范围，讨论解析函数 $T_{\beta\alpha}(z)=\langle\beta|T(z)|\alpha\rangle$, 那么就还可能出现一种新的解析结构，即对于某些合适的初态 $|\alpha\rangle$ 和末态 $|\beta\rangle$, 函数 $T_{\beta\alpha}(z)$ 会在复平面的实轴之外出现极点 (如图 (10.6) 所示)，与这些极点相应的量子

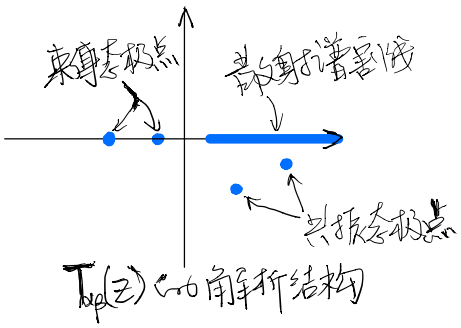

Figure 10.6: $T_{\alpha\beta}(z)$ 的解析结构。

态通常称作散射共振态，正如我们将要在本章最后一节中论证的，在物理上这些散射共振态相应于系统的一些不稳定的会衰变的束缚态。为了从一个基本理论出发得到这些散射共振态的信息，我们往往也需要用非微扰的方式计算 T 矩阵。

上面我们研究了 S 矩阵 (实际上我们研究的是和它密切相关的 T 矩阵) 作为入射粒子能量的函数的解析性。实际上 S 矩阵还有另外一个重要的性质，那就是幺正性。前面我们说过，由于 S 算符是一种特殊的时间演化算符，所以它是幺正算符，即满足 $S^\dagger S = SS^\dagger = 1$。由此我们就能推导出 S 矩阵必定是一个幺正矩阵，比方说 $\delta(\beta-\alpha) = \langle\beta|1|\alpha\rangle = \langle\beta|S^\dagger S|\alpha\rangle = \int d\gamma\langle\beta|S^\dagger|\gamma\rangle\langle\gamma|S|\alpha\rangle = \int d\gamma S^*_{\gamma\beta}S_{\gamma\alpha}$，式中利用了 $\langle\beta|S^\dagger|\gamma\rangle = \langle\gamma|S|\beta\rangle^* = S^*_{\gamma\beta}$，尤其是利用了 $S^\dagger S = 1$。我们也可以类似地利用 $SS^\dagger = 1$，最后就可以得到

$$\begin{aligned}\int d\gamma S^*_{\gamma\beta}S_{\gamma\alpha} &= \delta(\beta-\alpha)\\ \int d\gamma S_{\beta\gamma}S^*_{\alpha\gamma} &= \delta(\beta-\alpha).\end{aligned} \tag{10.99}$$

这就是 S 矩阵的幺正性。

利用方程 (10.89) 将 S 矩阵用 T 矩阵来表达，我们就可以由 S 矩阵的幺正性得到 T 矩阵的一些重要性质。比方说，根据 (10.99) 的第一个式子，我们可以得到 $(2\pi)\delta(E_\beta - E_\alpha)[-iT_{\beta\alpha}+iT^*_{\alpha\beta}]+(2\pi)^2\int d\gamma\delta(E_\gamma-E_\beta)\delta(E_\gamma-E_\alpha)T^*_{\gamma\beta}T_{\gamma\alpha}=0$，利用 $\delta(E_\gamma-E_\beta)\delta(E_\gamma-E_\alpha) = \delta(E_\gamma-E_\beta)\delta(E_\beta-E_\alpha)$，即有 $[-iT_{\beta\alpha}+iT^*_{\alpha\beta}]+(2\pi)\int d\gamma\delta(E_\gamma-E_\beta)T^*_{\gamma\beta}T_{\gamma\alpha}=0$，进一步取 $\beta=\alpha$，则有

$$\pi\int d\gamma\delta(E_\gamma-E_\alpha)|T_{\gamma\alpha}|^2 = -\mathrm{Im}T_{\alpha\alpha}. \tag{10.100}$$

由单位时间的散射概率公式 (10.91) 可以知道，这个方程的左边实际上就等于 $\frac{\hbar}{2}\int d\gamma w(\alpha\to\gamma)$，即正比于单位时间之内初态 α 到所有可能末态的总散射概率，而方程右边的 $T_{\alpha\alpha}$ 人们常常称为朝前散射幅。所以，方程 (10.100) 告诉我们的是，单位时间的总散射概率 Γ_α 由朝前散射幅的虚部决定，即

$$\Gamma_\alpha = \int d\gamma w(\alpha\to\gamma) = -\frac{2}{\hbar}\mathrm{Im}T_{\alpha\alpha}. \tag{10.101}$$

式中第一个等号涉及的是 Γ_α 的定义。这个结果通常称作光学定理。类似的，根据 (10.99) 的第二个式子，我们还可以得到

$$\pi\int d\gamma\delta(E_\alpha-E_\gamma)|T_{\alpha\gamma}|^2 = -\mathrm{Im}T_{\alpha\alpha}. \tag{10.102}$$

它告诉我们所有其它态散射到 $|\alpha\rangle$ 态的总概率也由 $|\alpha\rangle$ 的朝前散射幅决定。将这两个结果联合起来就有，α 态散射到其它态的总概率，等于其它态散射到 α 态的总概率，这正反映了概率守恒。这是因为，我们的这些结果都是根据幺正性得来的，而幺正性的本质正是概率守恒。

特别的，对于单个入射粒子在一个固定势场中的散射问题，我们有 $\sigma = \int d\sigma = \frac{1}{J_\alpha}\int d\beta w(\alpha\to\beta) = -\frac{2}{\hbar}\frac{\mathrm{Im}T_{\alpha\alpha}}{J_\alpha}$，式中 σ 称为总截面，它要对所有可能的散射末态积分，包括要对末态粒子

的自旋进行求和，式中的 J_α 当然就是入射粒子概率流密度，它等于 $\frac{1}{(2\pi\hbar)^3}\frac{\hbar k_\alpha}{m}$，$k_\alpha$ 表示入射粒子平面波的波数大小。将 J_α 代入总截面的表达式，就能得到

$$\frac{k_\alpha}{4\pi}\sigma = -(2\pi\hbar)^3\frac{m}{2\pi\hbar^2}\mathrm{Im}T_{\alpha\alpha}. \tag{10.103}$$

这就是散射截面的光学定理。表达式 (10.103) 前面的因子 $(2\pi\hbar)^3$ 当然是因为我们取入射粒子平面波为 $\frac{1}{(2\pi\hbar)^{3/2}}e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$ 了，如果将入射平面波取成 $e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$ 的形式，那这个因子就不存在了。

10.4.3 哈密顿量 H 的散射定态

从公式 (10.91) 可见，计算散射概率的关键就在于计算出 T 矩阵。最常用的计算办法就是按照 T 算符的微扰级数展开 (10.90) 逐级地计算。但是微扰论的方法并不是所有时候都好用的，比方说它往往不容易得到束缚态和共振态的信息。因此我们有时候需要寻找计算 T 矩阵的非微扰方法，或者至少是对 T 矩阵进行非微扰分析的方法。最重要的一种非微扰方法就是直接求解包含相互作用的完整哈密顿量 H 的散射定态。可以证明，从这些散射定态的解中可以直接地得到 T 矩阵 (当然，散射定态通常无法精确求解，但是很多时候我们依然可以借助于它进行非微扰的分析)。

为此，我们首先要定义什么是 H 的散射定态。前面在讨论解析性的时候我们已经提到，H_0 的散射谱可以对应到 H 的散射谱。H 的散射定态实际上就是这些散射谱所对应的本征态 (注意，不是束缚态)，有时候也称为 H 的散射态。这些散射态和 H_0 的散射态可以对应起来。实际上，根据不同的对应方法，我们可以定义 H 的两组不同散射定态。其中一组 H 的散射定态在无穷远过去即 $t\to-\infty$ 时 (这时候相互作用 V 可以忽略) 趋于 H_0 的散射态，因此在物理上，这一组散射定态是与入射自由粒子态相对应的，有时候也称为入射态或者 in 态。另一组 H 的散射定态在无穷远将来即 $t\to+\infty$ 时 (这时候相互作用也可以忽略) 趋于 H_0 的散射态，在物理上，这组散射定态当然就是与出射自由粒子态相对应的，因此有时候也称作出射态或者 out 态。不过值得强调的是，in 态并不是入射自由粒子态，out 态也不是出射自由粒子态，它们都是包含了相互作用 V 之后的，因为它们都是 $H=H_0+V$ 的本征态。

当然，为了得到 T 矩阵我们实际上并不需要 in 态和 out 态都同时研究，只研究其中之一就已经够了，因此下面我们将主要处理 in 态，除非特别说明，下面我们说到散射定态的时候都默认指的是 in 态，我们将直接从形式上构造出这些 in 态。与入射自由粒子态 $|\alpha\rangle$ 对应的 in 态通常被记作 $|\psi_\alpha^+\rangle$，我们将它构造为

$$|\psi_\alpha^+\rangle = U_I(0,-\infty)|\alpha\rangle, \tag{10.104}$$

这个构造的物理含义非常清楚，它说的无非是，$|\psi_\alpha^+\rangle$ 是用时间演化算符 U_I 将一个 H_0 的本征态 $|\alpha\rangle$ 从 $-\infty$ 演化到 0 时刻所得到的态，$-\infty$ 时的 $|\alpha\rangle$ 当然描述的是入射自由粒子，因此 $|\psi_\alpha^+\rangle$ 自然满足在无穷远过去趋于入射自由粒子态这一要求。在下文中，我们将阐明为什么从 $|\psi_\alpha^+\rangle$ 中可以得到 T 矩阵。另一方面，直接用 (10.104) 式这样的形式构造来计算 $|\psi_\alpha^+\rangle$ 实际上并不方便，所以下文我们将推导出 $|\psi_\alpha^+\rangle$ 所满足的一个基本方程 (Lippmann–

Schwinger 方程)，通过这个方程我们可以证明 $|\psi_\alpha^+\rangle$ 的确是 H 的本征态。如此一来，$|\psi_\alpha^+\rangle$ 就完全符合我们对散射定态的物理定义了。

但在具体进行上述推导之前，我们不妨先看一下 out 态可以如何构造，与 $|\alpha\rangle$ 态对应的 out 态 $|\psi_\alpha^-\rangle$ 可以从形式上构造为

$$|\psi_\alpha^-\rangle = U_I(0,+\infty)|\alpha\rangle = U_I^\dagger(+\infty,0)|\alpha\rangle. \tag{10.105}$$

将这个式子厄米共轭就有 $\langle\psi_\beta^-| = \langle\beta|U_I(+\infty,0)$，由此就可以进一步得到 $\langle\psi_\beta^-|\psi_\alpha^+\rangle = \langle\beta|U_I(+\infty,0)U_I(0,-\infty)|\alpha\rangle = \langle\beta|U_I(+\infty,-\infty)|\alpha\rangle = \langle\beta|S|\alpha\rangle$，即

$$\langle\beta|S|\alpha\rangle = \langle\psi_\beta^-|\psi_\alpha^+\rangle. \tag{10.106}$$

这个式子告诉我们，前面定义的散射 S 矩阵实际上是 in 态和 out 态之间的幺正变换矩阵，即是 H 的两组不同散射定态之间的幺正变换矩阵。

为了进一步确定散射定态 $|\psi_\alpha^+\rangle$，我们可以将它用自由粒子散射态 $\{|\beta\rangle\}$ 来进行展开，即计算 $\langle\beta|\psi_\alpha^+\rangle = \langle\beta|U_I(0,-\infty)|\alpha\rangle$。我们将 $U_I(0,-\infty)$ 展开成戴森级数 (10.15)，然后和前文计算 $\langle\beta|U_I(+\infty,-\infty)|\alpha\rangle$ 完全类似地逐级计算 $\langle\beta|U_I(0,-\infty)|\alpha\rangle$，这里将不再重复这个类似的计算过程了，具体推导留给读者作练习，结果是

$$\begin{aligned}
&\langle\beta|\psi_\alpha^+\rangle = \langle\beta|U_I(0,-\infty)|\alpha\rangle \\
=&\langle\beta|\alpha\rangle + \langle\beta|\frac{1}{E_\alpha - H_0 + i\varepsilon}V|\alpha\rangle \\
+&\langle\beta|\frac{1}{E_\alpha - H_0 + i\varepsilon}V\frac{1}{E_\alpha - H_0 + i\varepsilon}V|\alpha\rangle + ... \\
=&\langle\beta|\alpha\rangle + \langle\beta|\frac{1}{E_\alpha - H_0 + i\varepsilon}T|\alpha\rangle.
\end{aligned} \tag{10.107}$$

最后一个等式中的 T 就是 T 算符。从这里实际上已经可以看到 $|\psi_\alpha^+\rangle$ 中包含 T 矩阵的信息了，为了看得更清楚一点，我们可以将 (10.107) 改写成，

$$\begin{aligned}
|\psi_\alpha^+\rangle &= (1 + \frac{1}{E_\alpha - H_0 + i\varepsilon}V + \frac{1}{E_\alpha - H_0 + i\varepsilon}V\frac{1}{E_\alpha - H_0 + i\varepsilon}V + ...)|\alpha\rangle \\
&= |\alpha\rangle + \frac{1}{E_\alpha - H_0 + i\varepsilon}T|\alpha\rangle.
\end{aligned} \tag{10.108}$$

我们可以进一步看一下这个表达式中的 $T|\alpha\rangle$，根据 T 算符的级数形式 (10.90)，显然 $T|\alpha\rangle = (V + V\frac{1}{E_\alpha - H_0 + i\varepsilon}V + V\frac{1}{E_\alpha - H_0 + i\varepsilon}V\frac{1}{E_\alpha - H_0 + i\varepsilon}V + ...)|\alpha\rangle = V(1 + \frac{1}{E_\alpha - H_0 + i\varepsilon}V + \frac{1}{E_\alpha - H_0 + i\varepsilon}V\frac{1}{E_\alpha - H_0 + i\varepsilon}V + ...)|\alpha\rangle = V|\psi_\alpha^+\rangle$，这里最后一个等号用到了 (10.108) 中的第一个等号关系。因此我们有

$$T|\alpha\rangle = V|\psi_\alpha^+\rangle. \tag{10.109}$$

将这个结果代入 (10.108) 式的后一个等式，就可以得到 $|\psi_\alpha^+\rangle$ 所满足的基本方程，

$$|\psi_\alpha^+\rangle = |\alpha\rangle + \frac{1}{E_\alpha - H_0 + i\varepsilon}V|\psi_\alpha^+\rangle. \tag{10.110}$$

这个方程就是著名的 Lippmann–Schwinger 方程。

将 $E_\alpha - H_0$ 作用到 Lippmann–Schwinger 方程左右两边并取 $\varepsilon \to 0$，并注意到 $(E_\alpha - H_0)|\alpha\rangle = 0$, 我们就可以得到 $(E_\alpha - H_0)|\psi_\alpha^+\rangle = V|\psi_\alpha^+\rangle$, 也就是

$$H|\psi_\alpha^+\rangle = E_\alpha|\psi_\alpha^+\rangle. \tag{10.111}$$

这就证明了 $|\psi_\alpha^+\rangle$ 的确是 H 的本征态，并且本征值依然是 E_α, 后面这一点值得特别强调一下，因为按照我们最初的定义，E_α 只是 H_0 的本征值 (相应的本征态是 $|\alpha\rangle$), 也就是入射自由粒子的能量，现在我们又证明了 E_α 其实也是 H 的本征值。这当然是因为我们考察的是散射谱，而 H 的散射谱和 H_0 的散射谱是可以对应起来的，我们现在不过是建立了一个具体的对应方式，即将 $|\alpha\rangle$ 对应到 $|\psi_\alpha^+\rangle$。也正因为如此，Lippmann–Schwinger 方程不能直接用来求解 H 的束缚态，它只能用来求解散射定态，也即求解 $|\psi_\alpha^+\rangle$。

最后，从散射定态中读出 T 矩阵的办法非常简单，比方说根据 (10.109) 式，我们可以用下面的式子来得到 T 矩阵

$$\langle\beta|T|\alpha\rangle = \langle\beta|V|\psi_\alpha^+\rangle. \tag{10.112}$$

当然为此我们首先需要解出散射定态，这可以通过直接求解 H 的本征方程 (10.111) 来完成。当然，为了保证得到的解的确是散射定态，我们还要求这个本征态解可以写成如下形式

$$|\psi_\alpha^+\rangle = |\alpha\rangle + \frac{1}{E_\alpha - H_0 + i\varepsilon} T|\alpha\rangle. \tag{10.113}$$

注意, 这一形式中的 T 是解本征方程 (10.111) 得到的，其物理含义暂时未知，但从前面的讨论可以知道，如果 H 的本征态解有这种形式，那这个解就是散射定态，那么其中的 T 当然就是 T 算符。大体上我们就是这样从散射定态中读出 T 矩阵的。

而且散射概率的一般公式 (10.91) 清楚地告诉我们，散射初末态的能量是守恒的 (至少我们这里的理论只处理了这种最重要的情形)。因此实际上我们只需要知道 T 矩阵在初末态能量相等时的矩阵元。这其实相当于说，我们并不需要知道散射定态 $|\psi_\alpha^+\rangle$ 的完整解，只需要知道它的某种渐近形式就可以了。

从上面两段的简单描述中可以知道，这一得出 T 矩阵的办法是非微扰的，它的关键就是求解 H 本征方程 (10.111) 的某种特定类型的解，即散射定态解。但是我们也知道，对于大多数相互作用 V，没有办法得到本征方程 (10.111) 的精确解，因此也就无法得到精确的 T 矩阵。但是，我们依然可以用上面描述的过程来对 T 矩阵进行非微扰分析，而且，上一小节分析的 T 矩阵的解析性对我们的非微扰分析很有帮助，另外，T 矩阵和 S 矩阵是密切相关的，而 S 矩阵的幺正性也会极大地帮助我们进行非微扰分析。此外，非微扰的分析甚至精确求解真正能发挥威力的地方往往是系统有某些对称性的时候，这时候对称性会极大地简化对 T 矩阵和 S 矩阵的分析。在下一章中，我们将具体地示例整个分析过程可以如何进行。

10.5 * 不稳定束缚态和散射共振态

在量子力学中，末态是连续谱的情况非常常见，比方说一个原子受电场的作用电离成一个离子加上一个自由电子，由于末态中包含了一个自由电子，因此它当然是一个连续

谱。再比方说，一个处在激发态的原子通过与真空中的电磁涨落相互作用自发跃迁到较低能态，并辐射出一个光子，这时候由于末态中包含了一个自由的光子，因此当然也是一个连续谱 (当然，由于能量守恒，辐射出来的光子能量本身几乎是分立的，但是作为自由光子，这些辐射出来的光子从属于一个连续谱)。正如我们在第五章中用有效哈密顿量方法讨论过的，这种与末态连续谱的耦合将使得原来的离散初态变得不稳定，进而表现出一个有限的寿命。

从广义上来说，由于和环境或者真空中的量子涨落的耦合，除非受某一条守恒定律 (比方说非相对论极限下的粒子数守恒，比方说更精确的电荷守恒等等) 的保护，否则一切量子体系的激发态严格来说都不稳定，在一个足够长的时间尺度上都会衰变，这种衰变就是因为激发态跃迁到更低能态时向环境中释放的量子信号从属于一个连续谱 [1]。在第五章中我们已经证明过这个结论了，但是在这一节中我们还想仿照 Cohen-Tannoudji 的处理，通过直接求解含时问题给出一个与前面不同的证明。另外，在本节中我们还会将这种不稳定离散态与散射共振态联系起来，从而表明散射实验如何可以发现这样的不稳定离散态。实际上，在粒子物理中，大量的不稳定粒子都是通过这样的散射实验发现的，最近的重要例子是发现 Higgs 粒子，也就是所谓的上帝粒子。

另外，本章前面的小节中我们研究了在一个充分长的时间尺度上，微扰可以使得系统在两个离散态之间振荡，因此一阶含时微扰论的计算只在一个相对较短的时间内才成立。但是我们也提到，如果末态是一个连续谱的话，那么情况将会完全不同，这时候不会出现初末态之间的振荡，而是初态一旦跃迁到末态连续谱，就会被末态连续谱“粘住”，一去不复返。正因为如此，在我们对末态连续谱的含时微扰论处理中才可以近似地认为微扰的作用时间为无穷长。本节关于与末态连续谱耦合的离散初态的不稳定性的证明其实也是一个关于末态连续谱时不会出现振荡的证明。

当然，首先我们要注意到，对这种末态连续谱情形，含时微扰论的处理并不足以证明不会出现振荡，因为含时微扰论计算出来的跃迁概率正比于 $|T_{\beta\alpha}|^2 t/\hbar$, 作为概率它当然要小于 1，这就使得含时微扰论的处理仅在微扰的有效作用时间不过于长时才成立 (比方说不能是一个宏观的时间尺度)。当然，因为 $|T_{\beta\alpha}|^2$ 是一个二阶小量，这使得跃迁概率小于 1 这个要求可以在一个微观上充分长的时间内成立，因此在对末态连续谱的含时微扰论处理中取 $t \to \infty$ 的近似是可以成立的。但这对于证明末态连续谱时没有振荡依然不足够，为了完成这样的证明，我们必须要处理任意长的时间尺度，甚至要能够分析宏观的时间尺度。这时候含时微扰论的处理将不再成立。

10.5.1 含时问题与离散态衰变

我们假设系统的末态是连续谱但是初态是一个离散态，记为 $|i\rangle$。这时候根据公式 (10.28)，单位时间之内，初态 i 跃迁到末态连续谱的总概率是 (注意这里末态不是一个态，而是一个连续谱, 我们要把所有的可能末态都积分起来)，

$$\Gamma_i = \frac{2\pi}{\hbar} \int_f |\langle \alpha_f | V | i \rangle|^2 \rho(E_f, \chi_f)|_{E_f = E_i} d\chi_f. \tag{10.114}$$

[1] 比方说，黑洞就可以看成是量子引力的高激发态，因此根据我们的这个观点，黑洞将是不稳定的，将会衰变。的确，霍金从数学上证明了这种衰变的存在，那就是著名的黑洞的霍金辐射。

表达式中的积分 $\int_f$ 表示对 $d\chi_f$ 积分，也就是对所有可能末态积分。

我们假设初态 $|i\rangle$ 和末态 $|\alpha\rangle$ 都是未扰动的哈密顿量 H_0 的本征态，即满足

$$H_0|i\rangle = E_i|i\rangle,\ H_0|\alpha\rangle = E_\alpha|\alpha\rangle. \tag{10.115}$$

我们假定 $|i\rangle$ 是一个离散态，为了简单起见我们不妨假定它不简并。假定 α 是连续指标，因此 E_α 是一个连续能谱，它的可能取值构成了实轴上的一个区间，并且这个区间足够长，也即是说，我们假定末态连续谱足够宽，为了简单起见不妨假定 E_α 连续取值在 $[0,+\infty)$ 区间上。当然，要实现末态连续谱与初态 $|i\rangle$ 的共振 (这里共振的含义后文会说清楚，注意，它不是初末态之间的振荡)，初态能量 E_i 也得落在这个区间之内。$|\alpha\rangle$ 归一化到 δ 函数上，满足 $\langle\alpha'|\alpha\rangle = \delta(\alpha'-\alpha)$。当然我们也常常将末态标记从归一化到 δ 函数的 α 指标变换到 (E,χ) 指标，E 其实就是 E_α，只不过刚才我们是以 α 为指标，现在是以 E 本身为指标，因此和前文中的 (10.21) 式一样，我们有

$$d\alpha = \rho(E,\chi)dEd\chi. \tag{10.116}$$

当然，离散谱 $|i\rangle$ 和连续谱 $|\alpha\rangle$ 是正交的，满足 $\langle i|\alpha\rangle = 0$. 并且，为了简单起见，我们将忽略除 $|i\rangle$ 之外的其它离散谱，因此 $|i\rangle$ 和 $|\alpha\rangle$ 构成封闭性关系

$$|i\rangle\langle i| + \int d\alpha|\alpha\rangle\langle\alpha| = 1. \tag{10.117}$$

现在，假设在系统中引入一个恒定微扰 V，这个微扰的作用是将离散谱 $|i\rangle$ 和连续谱 $|\alpha\rangle$ 耦合起来，因此一般来说，

$$\langle\alpha|V|i\rangle \neq 0. \tag{10.118}$$

另外，为了简单起见我们还进一步假设 V 满足

$$\langle i|V|i\rangle = \langle\alpha|V|\alpha\rangle = 0,\ \langle\alpha'|V|\alpha\rangle = 0, \tag{10.119}$$

这里第一个式子的假设是平凡的，它说的无非是 V 没有对角元，这一点总是可以通过将 V 的对角元吸收到 H_0 的定义中来满足。第二个式子 $\langle\alpha'|V|\alpha\rangle = 0$ 的假设会极大地简化我们的分析，它说的是连续谱中的态相互之间没有耦合。

以上看起来我们作了很多的假设，但这些假设大多都是为了简化数学分析，在物理上真正本质的假设其实主要是两条：第一，末态连续谱得足够宽。第二，$|i\rangle$ 态与这个连续谱在一个足够宽的范围内有耦合。

我们的出发点是前文在量子共振的一般性研究中所引入的方程 (10.66)。只不过现在需要把一些离散的态指标改成连续的态指标，相应的对离散态的求和也要改成连续的积分。同样我们假设初始时，即 $t_0 = 0$ 时，系统处在 $|i\rangle$ 态，并记 t 时刻系统跃迁到 $|\alpha\rangle$ 态的概率幅为 $b_\alpha(t)$，依然跃迁到 $|i\rangle$ 态本身的概率幅为 $b_i(t)$，即

$$b_\alpha(t) = \langle\alpha|U_I(t)|i\rangle,\ b_i(t) = \langle i|U_I(t)|i\rangle. \tag{10.120}$$

由于假设 $t=0$ 时系统处在 $|i\rangle$ 态，所以有初始条件 $b_i(0)=1, b_\alpha(0)=0$。因此根据方程 (10.66) 以及上面的那些假设，我们有

$$\begin{aligned}\frac{db_i}{dt} &= (-\frac{i}{\hbar})\int d\alpha e^{i(E_i-E_\alpha)t/\hbar}\langle i|V|\alpha\rangle b_\alpha(t)\\ \frac{db_\alpha}{dt} &= (-\frac{i}{\hbar})e^{i(E_\alpha-E_i)t/\hbar}\langle\alpha|V|i\rangle b_i(t).\end{aligned} \tag{10.121}$$

下面我们要做的就是在合理的近似下求解这个方程组。

首先，我们对方程组 (10.121) 中的第二个方程进行积分，注意到初始条件就可以得到 $b_\alpha(t)=(-\frac{i}{\hbar})\langle\alpha|V|i\rangle\int_0^t dt' e^{i(E_\alpha-E_i)t'/\hbar}b_i(t')$, 将之代入方程组 (10.121) 中的第一个方程，从而有

$$\begin{aligned}\frac{db_i}{dt} &= (-\frac{1}{\hbar^2})\int d\alpha\int_0^t dt' e^{i(E_i-E_\alpha)(t-t')/\hbar}|\langle\alpha|V|i\rangle|^2 b_i(t')\\ &= (-\frac{1}{\hbar^2})\int_0^t dt'\int dE d\chi e^{i(E_i-E)(t-t')/\hbar}|\langle E,\chi|V|i\rangle|^2\rho(E,\chi)b_i(t').\end{aligned} \tag{10.122}$$

式中第二个等号我们是将末态指标从 α 变换到了 (E,χ)。现在引入记号 $K(E)$, 其定义是

$$K(E)=\int d\chi|\langle E,\chi|V|i\rangle|^2\rho(E,\chi). \tag{10.123}$$

那么方程 (10.122) 就可以重写为

$$\frac{db_i}{dt}=(-\frac{1}{\hbar^2})\int_0^t dt'\int_0^{+\infty} dE e^{i(E_i-E)(t-t')/\hbar}K(E)b_i(t'). \tag{10.124}$$

现在我们假设微扰 V 的性质足够良好。以致于通过 (10.123) 式引入的函数 $K(E)$ 是一个足够平滑的函数，即它关于自变量 E 的变化足够缓慢。另外，根据我们的关于 $|i\rangle$ 态与连续谱在一个足够宽的范围内有耦合的假定，$K(E)$ 将在一个足够宽的能量范围内非 0。

根据上述 $K(E)$ 的性质可以知道，如果时间差 $t-t'$ 不足够小的话，那么方程 (10.124) 中的积分 $\int_0^{+\infty}dE e^{i(E_i-E)(t-t')/\hbar}K(E)$ 都会因为指数因子 $e^{i(E_i-E)(t-t')/\hbar}$ 而多次振荡，我们知道，这种指数振荡的求和 (积分) 是相互抵消的。因此这也就是说，除非 t' 足够接近 t，否则 $\int_0^{+\infty}dE e^{i(E_i-E)(t-t')/\hbar}K(E)$ 都将为 0。既然仅当 t' 离 t 足够近时方程 (10.124) 右边才不等于 0，那么我们就可以合理地将方程 (10.124) 右边的 $b_i(t')$ 近似为 $b_i(t)$，从而得到

$$\begin{aligned}\frac{db_i}{dt} &= (-\frac{1}{\hbar^2})\int_0^{+\infty}dE\int_0^t dt' e^{i(E_i-E)(t-t')/\hbar}K(E)b_i(t)\\ &= (-\frac{1}{\hbar^2})b_i(t)\int_0^{+\infty}dE K(E)\int_0^t d\tau e^{i(E_i-E)\tau/\hbar}.\end{aligned} \tag{10.125}$$

人们当然可以就任意的时间 t 分析求解方程 (10.125)。但是由于我们关心的是系统的长时间行为，所以我们还可以近似认为方程 (10.125) 中关于 τ 的积分上限是 $+\infty$，因此方程 (10.125) 中关于 τ 的积分变成 $\int_0^{+\infty}d\tau e^{i(E_i-E)\tau/\hbar}$，这个积分在上限处是无穷振荡的，为了让积分收敛，我们可以给 E_i-E 加上一个无穷小的正虚部 $i\varepsilon$，并在计算完成之后令 $\varepsilon\to 0$。因此我们要算的积分就变成

$$\int_0^{+\infty}d\tau e^{i(E_i-E+i\varepsilon)\tau/\hbar}=\frac{i\hbar}{E_i-E+i\varepsilon}. \tag{10.126}$$

利用数学公式

$$\lim_{\varepsilon\to 0}\frac{1}{x+i\varepsilon}=\mathscr{P}\frac{1}{x}-i\pi\delta(x), \tag{10.127}$$

其中 $\mathscr{P}$ 表示取主值部分。从而我们就有

$$\int_0^{+\infty}d\tau e^{i(E_i-E+i\varepsilon)\tau/\hbar}=i\hbar\mathscr{P}\frac{1}{E_i-E}+\hbar\pi\delta(E_i-E). \tag{10.128}$$

将这些近似和结果代入方程 (10.125) 就可以得到

$$\frac{db_i}{dt}=(-\frac{1}{\hbar})b_i(t)\left(\pi K(E_i)+i\mathscr{P}\int_0^{+\infty}dE\frac{1}{E_i-E}K(E)\right). \tag{10.129}$$

我们引入记号 ΔE_i，它的定义是

$$\begin{aligned}\Delta E_i&=\mathscr{P}\int_0^{+\infty}dE\frac{1}{E_i-E}K(E)\\&=\mathscr{P}\int dEd\chi\frac{|\langle E,\chi|V|i\rangle|^2}{E_i-E}\rho(E,\chi)\\&=\mathscr{P}\int d\alpha\frac{|\langle\alpha|V|i\rangle|^2}{E_i-E_\alpha}.\end{aligned} \tag{10.130}$$

式中第二个等号我们代入了 $K(E)$ 的定义 (10.123)。很显然，这个 ΔE_i 正是根据定态微扰论 $|i\rangle$ 能级能量的二级微扰修正。另一方面，根据函数 $K(E)$ 的定义 (10.123)，我们可以注意到 $\frac{2\pi}{\hbar}K(E_i)$ 正是我们前面在 (10.114) 式中定义的 Γ_i(只不过现在我们省略了 (10.114) 式中所有在末态上的记号 f)。如此一来我们就可以将方程 (10.129) 重写为

$$\frac{db_i}{dt}=\left(-\frac{\Gamma_i}{2}-i\frac{1}{\hbar}\Delta E_i\right)b_i(t). \tag{10.131}$$

方程 (10.131) 的解是

$$b_i(t)=e^{-\frac{\Gamma_i}{2}t}e^{-i\frac{\Delta E_i}{\hbar}t}. \tag{10.132}$$

将这个解代回最初的方程组 (10.121) 的第二个方程，又可以进一步得到

$$b_\alpha(t)=\langle\alpha|V|i\rangle\frac{1-e^{-\frac{\Gamma_i}{2}t}e^{i(E_\alpha-E_i-\Delta E_i)t/\hbar}}{(E_\alpha-E_i-\Delta E_i)+i\frac{\Gamma_i}{2}\hbar}. \tag{10.133}$$

这两个解就描写了系统是怎么渐渐从 $|i\rangle$ 态跃迁到末态连续谱的。尤其是这个解中没有出现振荡！这可以通过 t 时刻系统在 $|i\rangle$ 态上的概率 $p_i(t)$ 看得很清楚，

$$p_i(t)=|b_i(t)|^2=e^{-\Gamma_i t}. \tag{10.134}$$

这个结果与我们在第五章中用有效哈密顿量方法得到的结果完全一致，它说明与连续谱的耦合使得 $|i\rangle$ 态呈现出一个有限的寿命 τ_i, 其定义是

$$\tau_i=\frac{1}{\Gamma_i}. \tag{10.135}$$

我们也可以计算 t 时刻，系统跃迁到 $|\alpha\rangle$ 态上的概率密度 $p(\alpha,t)=|b_\alpha(t)|^2$。根据 (10.133) 式，当 $t\gg\tau_i$ 时，指数因子近似为 0，从而

$$p(\alpha,t)=|\langle\alpha|V|i\rangle|^2\frac{1}{(E_\alpha-E_i-\Delta E_i)^2+(\frac{\Gamma_i}{2}\hbar)^2}. \tag{10.136}$$

显然这个概率密度在 $E_\alpha=E_i+\Delta E_i$ 时达到最大，反映的是末态连续谱与初态的共振。

10.5.2 自电离以及散射共振态

下面举一个具体的例子以说明上述理论。假设我们考虑一个氦原子，其哈密顿量 H 为

$$\begin{aligned} H &= \frac{\mathbf{P}_1^2}{2m_e} + \frac{\mathbf{P}_2^2}{2m_e} - \frac{2e^2}{4\pi\varepsilon_0|\mathbf{x}_1|} - \frac{2e^2}{4\pi\varepsilon_0|\mathbf{x}_2|} + \frac{e^2}{4\pi\varepsilon_0|\mathbf{x}_1-\mathbf{x}_2|} \\ &= \frac{\mathbf{P}_1^2}{2m_e} + \frac{\mathbf{P}_2^2}{2m_e} - \frac{2\hbar\alpha c}{|\mathbf{x}_1|} - \frac{2\hbar\alpha c}{|\mathbf{x}_2|} + \frac{\hbar\alpha c}{|\mathbf{x}_1-\mathbf{x}_2|}. \end{aligned} \tag{10.137}$$

式中 $\mathbf{P}_1$, $\mathbf{P}_2$ 分别是氦原子两个电子的动量算符，$\mathbf{x}_1$, $\mathbf{x}_2$ 是这两个电子的位置矢量。式中 $\alpha = e^2/(4\pi\varepsilon_0\hbar c)$ 是精细结构常数，而且由于这里涉及的是非相对论物理，所以 αc 总是作为一个整体出现 (因为非相对论物理不涉及光速，而只有 αc 这样的整体才和光速 c 无关)，因此整个系统只涉及三个相互独立的物理量，m_e, $\hbar$, αc，为了数学分析的简洁我们很容易作无量纲化处理，即令

$$\hbar = 1,\ m_e = 1,\ \alpha c = 1. \tag{10.138}$$

如此一来所有的物理量都将无量纲了，要恢复量纲也很容易，人们只要做简单的量纲分析就可以了。因此无量纲化以后，氦原子的哈密顿量将是

$$H = \frac{1}{2}(\mathbf{P}_1^2 + \mathbf{P}_2^2) - \frac{2}{|\mathbf{x}_1|} - \frac{2}{|\mathbf{x}_2|} + \frac{1}{|\mathbf{x}_1-\mathbf{x}_2|}. \tag{10.139}$$

这个无量纲化的氦原子基态能量大约在 $E_g \approx -2.9$。

现在我们将氦原子的哈密顿量 H 分解成 $H_0 + V$，其中

$$H_0 = \frac{1}{2}(\mathbf{P}_1^2 + \mathbf{P}_2^2) - \frac{2}{|\mathbf{x}_1|} - \frac{2}{|\mathbf{x}_2|},\quad V = \frac{1}{|\mathbf{x}_1-\mathbf{x}_2|}. \tag{10.140}$$

V 代表的当然就是两个电子之间的库伦排斥相互作用，我们将它看成是对 H_0 的一个扰动。

H_0 可以精确求解，它就是两个相互独立的类氢原子哈密顿量之和，因此 H_0 的本征能量是两个没有相互作用的电子的能量之和。每个电子的能量谱又可以分成一个束缚态部分和一个散射谱部分，束缚态部分能级为

$$-\frac{2}{n^2},\ n \in \mathbb{Z}^+. \tag{10.141}$$

电子的散射谱部分可以表示为

$$\frac{1}{2}\mathbf{p}^2, \tag{10.142}$$

$\mathbf{p}$ 是这个电子在离核无穷远处的动量，这时候氦原子核的库伦场可以忽略，电子可以认为是自由的。

现在，让我们考察 H_0 的一个离散态 $|2,2\rangle$，它表示每个电子都处在 $n=2$ 的第二束缚态能级 (我们忽略电子的角动量量子数)，我们将它记作 $|R\rangle$ 态，即 $|R\rangle = |2,2\rangle$。$|R\rangle$ 态作为 H_0 的本征态，其能量本征值 E_R 为

$$E_R = -\frac{2}{2^2} - \frac{2}{2^2} = -1. \tag{10.143}$$

$|R\rangle$ 态当然只是 H_0 的本征态，考虑到两电子之间的相互作用 V, E_R 会发生能级移动。

让我们同时考察 H_0 的一组连续态 $|1,\mathbf{p}\rangle$, 它表示一个电子处于 $n=1$ 的第一束缚态能级，另一个电子电离出去，$\mathbf{p}$ 为这个电离出去的电子离氦原子核无穷远时的动量，因此 $|1,\mathbf{p}\rangle$ 描写的是一个 He^+ 离子加上一个电离出去的电子，我们记作 $He^+ + e$。$|1,\mathbf{p}\rangle$ 态的能量本征值我们记作 $E(\mathbf{p})$, 很显然它是

$$E(\mathbf{p}) = -\frac{2}{1^2} + \frac{1}{2}\mathbf{p}^2 = -2 + \frac{1}{2}\mathbf{p}^2. \tag{10.144}$$

显然 $E(\mathbf{p})$ 是一个连续谱，这个谱的取值范围是 $[-2,+\infty)$。注意，即使加上相互作用 V, $E(\mathbf{p})$ 也不会有能级移动，因为当一个电子被束缚在基态，另一个电子被电离到无穷远并在无穷远处有动量 $\mathbf{p}$ 时，两个电子之间的相互作用当然是可以忽略的。这实际上就是我们在散射的一般理论那一节中所描述的 H_0 的散射谱和 H 的散射谱之间的对应关系，而且从散射的一般理论中我们还知道 $|1,\mathbf{p}\rangle$ 态可以对应到 H 的两个散射定态 $|\psi^+_{1,\mathbf{p}}\rangle$ 和 $|\psi^-_{1,\mathbf{p}}\rangle$, $E(\mathbf{p})$ 同时也是这两个 H 的散射定态的本征值。值得一提的是，H 的基态 (即氦原子基态) 能量 ($E_g \approx -2.9$) 的确落在这个散射谱的范围之外。

从上面的分析我们清楚地看到，H_0 的这个离散能级 $|R\rangle$，其能量 E_R 落在连续谱 $[-2,+\infty)$ 的范围之内。如果没有电子间的库伦排斥相互作用 V, 那这也没有什么。但是库伦排斥相互作用 V 将 $|R\rangle$ 态和连续态 $|1,\mathbf{p}\rangle$ 耦合了起来，即

$$\langle 1,\mathbf{p}|V|R\rangle \neq 0. \tag{10.145}$$

假设系统最初处在离散能级 $|R\rangle = |2,2\rangle$ 态，那么根据我们在上一小节中的讨论，这一离散能级和连续谱的耦合将会导致束缚态 $|R\rangle = |2,2\rangle$ 变得不稳定，进而衰变到连续态 $|1,\mathbf{p}\rangle$ 之中。这一过程描述的就是，一个最初处在 $|2,2\rangle$ 束缚态的氦原子由于电子之间的库伦排斥相互作用自发地电离成 $He^+ + e$。这样的过程就叫做激发态原子的自电离过程。在物理上，这其实是因为 $|2,2\rangle$ 态的氦原子的某个电子自发地从 $n=2$ 能级跃迁到了 $n=1$ 能级，同时这一跃迁所释放的能量将另一个电子从 $n=2$ 能级电离出去的过程。也即是说，$|R\rangle = |2,2\rangle$ 态有一个有限的寿命 τ_R, 如果我们记单位时间 $|R\rangle$ 态自电离的概率为 Γ_R, 那么根据我们在上一小节中的结果，$\tau_R = 1/\Gamma_R$。

现在假设我们把这个过程反过来，我们用一个动量为 $\mathbf{p}$ 的自由电子作为入射粒子去和一个处在基态 (即 $n=1$ 能级) 的 He^+ 离子碰撞 (因此现在 $|1,\mathbf{p}\rangle$ 态是初态)。那么将会发生的是，由于静电吸引力，入射电子会短暂地被 He^+ 离子束缚，使得系统短暂地从 $|1,\mathbf{p}\rangle$ 态跃迁到 $|R\rangle = |2,2\rangle$ 态，但是 $|2,2\rangle$ 态是不稳定的，它很快又会衰变到另一个 $|1,\mathbf{p}'\rangle$ 态。因此 $|1,\mathbf{p}'\rangle$ 态才是散射的末态，它描述的是一个处在基态的 He^+ 离子，加上一个动量为 $\mathbf{p}'$ 的出射电子。我们当然可以测量整个从初态 $|1,\mathbf{p}\rangle$ 散射到末态 $|1,\mathbf{p}'\rangle$ 的散射截面 $d\sigma$。由散射的一般理论那一节中给出的公式 (10.92), 我们可以知道，

$$d\sigma \propto |\langle 1,\mathbf{p}'|T|1,\mathbf{p}\rangle|^2. \tag{10.146}$$

当然，决定散射截面的是系统从初态跃迁到末态的概率。而现在的末态 $|1,\mathbf{p}'\rangle$ 可以看作是由不稳定束缚态 $|2,2\rangle$ 衰变而来的, 因此根据我们在上一小节中的结果 (10.136), 这一

概率应该正比于 $\frac{1}{(E-E_R)^2+(\frac{\Gamma_R}{2})^2}$，这里 $E=E(\mathbf{p})=E(\mathbf{p}')$(因为散射过程的能量守恒)，$E_R$ 是 $|R\rangle=|2,2\rangle$ 态的能量，这里我们假设 E_R 中已经包含由于相互作用而产生的能级移动了 (因此这里的 E_R 不并等于没有相互作用时的 -1，而是修正后的结果)。这也就是说，

$$d\sigma \propto \frac{1}{(E-E_R)^2+(\frac{\Gamma_R}{2})^2}. \tag{10.147}$$

比较 (10.146) 式和 (10.147) 式这两个结果，我们会发现 T 矩阵在能量复平面上一定有如下形式的极点

$$\langle 1,\mathbf{p}'|T|1,\mathbf{p}\rangle \sim \frac{1}{E-E_R+i\frac{\Gamma_R}{2}}. \tag{10.148}$$

根据我们在散射的一般理论那一节中关于 T 矩阵解析性的讨论，这样的极点对应散射共振态。而在我们现在的例子中，这一极点来自于不稳定束缚态 $|R\rangle=|2,2\rangle$。这也就是说，不稳定束缚态在散射过程中就会成为散射共振态。虽然我们只是用一个具体例子来说明这一结论，但是回顾我们的分析过程，人们会发现完全类似的分析可以推广到一般情形。因此不稳定束缚态和散射共振态的这一联系完全是一般性的。

10.5.3 不稳定束缚态作为散射共振态

前面我们研究的是，由于和连续谱的耦合导致 H_0 的某个离散束缚态 $|i\rangle$ 变得不稳定，并最终衰变到连续谱 $|\alpha\rangle$ 的过程。但是在上一小节的最后，我们谈到如何将这个过程反过来，将连续的散射态 $|\alpha\rangle$ 作为初态，在 $|\alpha\rangle$ 态被相互作用 V 散射的过程中，系统可能短暂地跃迁到不稳定的 $|i\rangle$ 态，然后最终又衰变到连续的散射态 $|\beta\rangle$ 的过程，这时候这个不稳定束缚态 $|i\rangle$ 就成为散射的共振态，并对应于 T 矩阵在能量复平面下半平面的某个极点。这一小节我们就是要通过量子力学的推导来建立这个一般性的结论。

这一小节我们将沿用前面第 (10.5.1) 小节中的所有假设，唯一的不同在于，现在我们是要考察一个散射过程，初态是作为连续谱的散射态 $|\alpha\rangle$ 而不是不稳定束缚态 $|i\rangle$，散射的末态也是另一个散射态 $|\beta\rangle$，$|i\rangle$ 只是作为中间短暂的共振态而出现。我们首先要计算的就是单位时间的跃迁概率 $w_{\alpha\to\beta}$，通过将表达式 (10.91) 给出的概率密度 $w(\alpha\to\beta)$ 乘上末态数目 $d\beta=\rho(E,\chi)dEd\chi$ (式中 E 其实就是 E_β，只不过我们换成了以能量本身为末态指标)，并对末态能量积分，我们可以得到

$$w_{\alpha\to\beta}=\frac{2\pi}{\hbar}|\langle\beta|T|\alpha\rangle|^2\rho(E,\chi)|_{E=E_\alpha}d\chi. \tag{10.149}$$

显然，为了计算跃迁概率，我们首先需要计算 T 矩阵元 $\langle\beta|T|\alpha\rangle$。我们已经知道，$T$ 矩阵元可以通过求解散射定态 $|\psi_\alpha^+\rangle$ 来得到，当然一般来说，人们并不能精确求解出散射定态 $|\psi_\alpha^+\rangle$，但是在我们这里，由于沿用了第 (10.5.1) 小节中的许多简化性假设，散射定态 $|\psi_\alpha^+\rangle$ 实际上可以通过求解 Lippmann–Schwinger 方程 (10.110) 来得到。

具体来说，根据等式 (10.112)，T 矩阵元与散射定态的关系是，

$$\langle\beta|T|\alpha\rangle=\langle\beta|V|\psi_\alpha^+\rangle. \tag{10.150}$$

我们可以在 $\langle\beta|V|\psi_\alpha^+\rangle$ 中插入封闭性关系 (10.117)，并利用第 (10.5.1) 小节中关于连续态之间无耦合的假设，就可以得到

$$\langle\beta|T|\alpha\rangle=\langle\beta|V|i\rangle\langle i|\psi_\alpha^+\rangle. \tag{10.151}$$

由此可见，为了计算出 T 矩阵元以及相应的跃迁概率，我们只需求出散射定态 $|\psi_\alpha^+\rangle$ 在 $|i\rangle$ 态上的分量，即只需要求出 $\langle i|\psi_\alpha^+\rangle=\langle i|U_I(0,-\infty)|\alpha\rangle$，它实际上也就是入射的 $|\alpha\rangle$ 态在中间 $t=0$ 时刻暂时性地跃迁到不稳定束缚态 $|i\rangle$ 的概率幅。

正如前面所说的，我们可以通过求解 Lippmann–Schwinger 方程 (10.110) 来得到 $\langle i|\psi_\alpha^+\rangle$，我们将这个方程重写在下面

$$|\psi_\alpha^+\rangle=|\alpha\rangle+\frac{1}{E_\alpha-H_0+i\varepsilon}V|\psi_\alpha^+\rangle. \tag{10.152}$$

分别用 $|i\rangle$ 态以及连续态 $|\gamma\rangle$ 从左边和这个方程作内积，并注意到 $|i\rangle$ 和 $|\gamma\rangle$ 都是 H_0 的本征态，而且 $|i\rangle$ 和 $|\gamma\rangle$ 相互正交，从而就可以得到

$$\begin{aligned}\langle i|\psi_\alpha^+\rangle&=\frac{1}{E_\alpha-E_i+i\varepsilon}\langle i|V|\psi_\alpha^+\rangle\\&=\frac{1}{E_\alpha-E_i+i\varepsilon}\int d\gamma\langle i|V|\gamma\rangle\langle\gamma|\psi_\alpha^+\rangle\\\langle\gamma|\psi_\alpha^+\rangle&=\delta(\gamma-\alpha)+\frac{1}{E_\alpha-E_\gamma+i\varepsilon}\langle\gamma|V|\psi_\alpha^+\rangle\\&=\delta(\gamma-\alpha)+\frac{1}{E_\alpha-E_\gamma+i\varepsilon}\langle\gamma|V|i\rangle\langle i|\psi_\alpha^+\rangle.\end{aligned} \tag{10.153}$$

上式的第二行与第四行我们再一次在算符 V 和量子态 $|\psi_\alpha^+\rangle$ 的中间插入了封闭性关系 (10.117), 并利用了第 (10.5.1) 小节中的那些具体假设。

将方程组 (10.153) 的第二个方程代入第一个方程以消去 $\langle\gamma|\psi_\alpha^+\rangle$, 并稍作整理，我们就可以得到

$$\left[E_\alpha-E_i+i\varepsilon-\int d\gamma\frac{|\langle\gamma|V|i\rangle|^2}{E_\alpha-E_\gamma+i\varepsilon}\right]\langle i|\psi_\alpha^+\rangle=\langle i|V|\alpha\rangle. \tag{10.154}$$

式中我们利用了 V 的厄米性，即 $\langle i|V|\gamma\rangle=\langle\gamma|V|i\rangle^*$。

定义 $\Delta_i(E_\alpha)$ 和 $\Gamma_i(E_\alpha)$ 如下，

$$\Delta_i(E_\alpha)=\mathscr{P}\int d\gamma\frac{|\langle\gamma|V|i\rangle|^2}{E_\alpha-E_\gamma}, \tag{10.155}$$

$$\begin{aligned}\Gamma_i(E_\alpha)&=\frac{2\pi}{\hbar}\int d\gamma|\langle\gamma|V|i\rangle|^2\delta(E_\gamma-E_\alpha)\\&=\frac{2\pi}{\hbar}\int d\chi|\langle E,\chi|V|i\rangle|^2\rho(E,\chi)|_{E=E_\alpha}.\end{aligned} \tag{10.156}$$

则利用 $\frac{1}{E_\alpha-E_\gamma+i\varepsilon}=\mathscr{P}\frac{1}{E_\alpha-E_\gamma}-i\pi\delta(E_\alpha-E_\gamma)$, 人们可以验证方程 (10.154) 可以重写成

$$\left[E_\alpha-(E_i+\Delta_i(E_\alpha))+i\hbar\frac{\Gamma_i(E_\alpha)}{2}\right]\langle i|\psi_\alpha^+\rangle=\langle i|V|\alpha\rangle. \tag{10.157}$$

从 $\Delta_i(E_\alpha)$ 和 $\Gamma_i(E_\alpha)$ 的定义可以看出，当 $E_\alpha = E_i$ 时，$\Delta_i(E_\alpha)$ 就是能级 E_i 的二阶微扰修正，也就是第 (10.5.1) 小节中定义的 ΔE_i, 而这时候的 $\Gamma_i(E_\alpha)$ 就是 $|i\rangle$ 态单位时间的总衰变概率，也就是第 (10.5.1) 小节中定义的 Γ_i，即

$$\Delta_i(E_i) = \Delta E_i,\ \Gamma_i(E_i) = \Gamma_i. \tag{10.158}$$

到此为止，我们就已经可以解出 $\langle i|\psi_\alpha^+\rangle$ 了，将从 (10.157) 式解出的 $\langle i|\psi_\alpha^+\rangle$ 代入 T 矩阵元的表达式 (10.151)，并且记 $E_i + \Delta_i(E_\alpha) = E_{R,i}(E_\alpha)$(下标 R 代表共振态的意思)，我们就有

$$\langle\beta|T|\alpha\rangle = \frac{\langle\beta|V|i\rangle\langle i|V|\alpha\rangle}{E_\alpha - E_{R,i}(E_\alpha) + i\hbar\frac{\Gamma_i(E_\alpha)}{2}}. \tag{10.159}$$

显然，假设将 T 矩阵元看成是入射粒子能量 E_α 的函数，那么不稳定束缚态 $|i\rangle$ 的存在就对应于这个函数在能量复平面下半平面的一个极点，也就是对应于一个散射共振态。这就是不稳定束缚态与散射共振态之间的一般性联系。

人们很容易将上面的结果推广到 H_0 的多个离散能级与连续谱通过 V 相耦合的情形。这时候，$\Delta_i(E_\alpha)$ 需要推广成 $\Delta_{ij}(E_\alpha)$, $\Gamma_i(E_\alpha)$ 需要推广成 $\Gamma_{ij}(E_\alpha)$, 它们的定义如下，

$$\Delta_{ij}(E_\alpha) = \mathscr{P}\int d\gamma \frac{\langle i|V|\gamma\rangle\langle\gamma|V|j\rangle}{E_\alpha - E_\gamma}, \tag{10.160}$$

$$\begin{aligned}\Gamma_{ij}(E_\alpha) &= \frac{2\pi}{\hbar}\int d\gamma \langle i|V|\gamma\rangle\langle\gamma|V|j\rangle\delta(E_\gamma - E_\alpha)\\ &= \frac{2\pi}{\hbar}\int d\chi \langle i|V|E,\chi\rangle\langle E,\chi|V|j\rangle\rho(E,\chi)|_{E=E_\alpha}.\end{aligned} \tag{10.161}$$

很显然，它们都成了在离散态子空间里的矩阵，不妨引入离散态子空间里的算符 $\Delta(E_\alpha)$ 和 $\Gamma(E_\alpha)$，它们的定义是

$$\Delta(E_\alpha) = \sum_{i,j}\Delta_{ij}(E_\alpha)|i\rangle\langle j|, \quad \Gamma(E_\alpha) = \sum_{i,j}\Gamma_{ij}(E_\alpha)|i\rangle\langle j|. \tag{10.162}$$

假如我们同时将 $\sum_j |j\rangle\langle j|V|\alpha\rangle$ 看成是离散态子空间里的某个态 A_α, 同时将 $\sum_i\langle\beta|V|i\rangle\langle i|$ 看作 $\mathrm{A}_\beta^\dagger$，那么算符 $\Delta(E_\alpha)$ 和 $\Gamma(E_\alpha)$ 就可以写成

$$\Delta(E_\alpha) = \mathscr{P}\int d\gamma \frac{\mathrm{A}_\gamma\mathrm{A}_\gamma^\dagger}{E_\alpha - E_\gamma}, \tag{10.163}$$

$$\Gamma(E_\alpha) = \frac{2\pi}{\hbar}\int d\chi \mathrm{A}_{E,\chi}\mathrm{A}_{E,\chi}^\dagger\rho(E,\chi)|_{E=E_\alpha}. \tag{10.164}$$

而散射 T 矩阵的表达式现在就是，

$$\langle\beta|T|\alpha\rangle = \mathrm{A}_\beta^\dagger\frac{1}{E_\alpha - H_R(E_\alpha) + i\hbar\frac{\Gamma(E_\alpha)}{2}}\mathrm{A}_\alpha. \tag{10.165}$$

式中 $H_R(E_\alpha)$ 为离散态空间上的算符，其定义为 $H_R(E_\alpha) = \sum_i E_i|i\rangle\langle i| + \Delta(E_\alpha)$。

很显然，上面的所有结果都与第五章中用有效哈密顿量的方法得到的结果一致。而得到散射 T 矩阵 (10.165) 的最快方法其实也是应用第五章的有效哈密顿量。(这一段请读者

先去回顾一下第五章的相关结论) 首先, 注意到 $\langle\beta|T|\alpha\rangle = \langle\beta|P^\perp TP^\perp|\alpha\rangle$。其次, 根据 T 算符的一般公式 (10.98), $P^\perp T(z)P^\perp = P^\perp VP^\perp + P^\perp VG(z)VP^\perp$, 而按照我们的假设 $P^\perp VP^\perp = 0$, 所以有 $P^\perp T(z)P^\perp = P^\perp VPG(z)PVP^\perp$(这里在中间插入了 $P+P^\perp = 1$)。而 $PG(z)P$ 就是将 $G(z) = \frac{1}{z-H}$ 化约到离散态子空间 P 上, 利用第五章中的算符恒等式 $P\frac{1}{z-H}P = \frac{1}{z-H_{\text{eff}}(z)}$, 结果当然就是将哈密顿量 H 替换成 P 上的有效哈密顿量 $H_{\text{eff}}(z) = H_{PP} + \Delta(z) - \hbar\frac{i}{2}\Gamma(z)$, 从而最终我们就会得到

$$P^\perp T(z)P^\perp = V_{P^\perp P}\frac{1}{z - H_{PP} - \Delta(z) + \hbar\frac{i}{2}\Gamma(z)}V_{PP^\perp}. \tag{10.166}$$

将这个 T 算符写成 T 矩阵的形式，马上就能得到我们之前的结果 (10.165)。因此这一小节我们本质上是用求解 Lippmann–Schwinger 方程的方法印证了第五章有效哈密顿量的方法。

举例

下面我们考察上述理论的一个例子。假设我们考察单个入射粒子和一个固定靶散射，最后再作为末态出射的情形。不过，这个固定靶本身可以是一个有内部结构的系统，比方说，我们上一小节中的 He^+ 离子就是一个这样的固定靶。假设入射粒子在无穷远处是一个自由粒子，它在无穷远处的动量为 $\mathbf{p}$，出射粒子在无穷远处也是一个自由粒子，动量为 $\mathbf{p}'$。因此我们可以用 $|\mathbf{p}\rangle$ 来标记入射态 $|\alpha\rangle$, 用 $|\mathbf{p}'\rangle$ 来标记末态 $|\beta\rangle$。这样一个系统相应的连续能谱将可以写成

$$E_{\mathbf{p}} = E_0 + \frac{\mathbf{p}^2}{2m}. \tag{10.167}$$

式中 E_0 表示当散射的粒子与固定靶距离为无穷远时，固定靶的内能。散射过程的能量守恒意味着，入射粒子在无穷远处的动能必然与出射粒子在无穷远处的动能相等，因此 $|\mathbf{p}| = |\mathbf{p}'|$。正如我们早已知道的，这种连续谱的态密度 $\rho(E) = mp$, 式中 $p = |\mathbf{p}|$。对于这一类两体散射问题，我们可以计算微分散射截面，根据 (10.92) 式，现在它应该是，

$$d\sigma = (2\pi\hbar)^6\left(\frac{m}{2\pi\hbar^2}\right)^2|\langle\mathbf{p}'|T|\mathbf{p}\rangle|^2 d\Omega. \tag{10.168}$$

另外，假设系统有一个不稳定束缚态 $|i\rangle$, 下面我们将仅仅考虑这个不稳定束缚态与连续谱的耦合，因此就可以直接应用上面所发展的理论了。将态密度表达式 $\rho(E) = mp$ 代入 (10.156), 我们可以得到

$$\Gamma_i(E_{\mathbf{p}}) = \frac{2\pi}{\hbar}\int mpd\Omega|\langle\mathbf{p}|V|i\rangle|^2. \tag{10.169}$$

现在, 进一步假设相互作用 V 具有旋转不变性, 而且 $|i\rangle$ 态的角动量量子数 $l = 0$(因此没有简并)。这时候由于 $|i\rangle$ 和 V 均旋转不变, 所以矩阵元 $\langle\mathbf{p}|V|i\rangle$ 必定不依赖于动量 $\mathbf{p}$ 的方向, 即必有 $\langle\mathbf{p}|V|i\rangle = \langle p|V|i\rangle$。同样的道理, 也有 $\langle i|V|\mathbf{p}\rangle = \langle i|V|p\rangle$, 以及 $\langle\mathbf{p}'|V|i\rangle = \langle p|V|i\rangle$(这里我们用到了 $|\mathbf{p}'| = |\mathbf{p}| = p$)。将这些结果代入 (10.169) 就可以得到

$$\Gamma_i(E_{\mathbf{p}}) = 2\frac{(2\pi)^2}{\hbar}mp|\langle p|V|i\rangle|^2. \tag{10.170}$$

对于 T 矩阵元 $\langle \mathbf{p}'|T|\mathbf{p}\rangle$，根据 (10.159) 式，它应该是

$$\begin{aligned}\langle \mathbf{p}'|T|\mathbf{p}\rangle &= \frac{\langle \mathbf{p}'|V|i\rangle\langle i|V|\mathbf{p}\rangle}{E_{\mathbf{p}} - E_{R,i}(E_{\mathbf{p}}) + i\hbar\frac{\Gamma_i(E_{\mathbf{p}})}{2}} \\ &= \frac{|\langle p|V|i\rangle|^2}{E_{\mathbf{p}} - E_{R,i}(E_{\mathbf{p}}) + i\hbar\frac{\Gamma_i(E_{\mathbf{p}})}{2}} \\ &= \frac{1}{2(2\pi)^2 mp}\frac{\hbar\Gamma_i(E_{\mathbf{p}})}{E_{\mathbf{p}} - E_{R,i}(E_{\mathbf{p}}) + i\hbar\frac{\Gamma_i(E_{\mathbf{p}})}{2}}.\end{aligned} \tag{10.171}$$

在上式的最后一行中我们代入了 (10.170) 式。

将 (10.171) 的最终结果代入微分截面公式 (10.168), 并对所有的立体角积分，我们就能够得到总截面 $\sigma_T = \int d\sigma$，结果是

$$\sigma_T = \frac{4\pi}{k^2}\frac{[\hbar\Gamma_i(E_{\mathbf{p}})]^2}{4[E_{\mathbf{p}} - E_{R,i}(E_{\mathbf{p}})]^2 + [\hbar\Gamma_i(E_{\mathbf{p}})]^2}. \tag{10.172}$$

式中 $k = p/\hbar$。如果相互作用 V 的数学性质足够良好，以致于函数 $E_{R,i}(E_{\mathbf{p}})$ 和函数 $\Gamma_i(E_{\mathbf{p}})$ 随初态能量的变化都非常缓慢，那我们也可以将它们近似看成是常数，分别记为 $E_{R,i}$ 和 Γ_i，那么总截面的公式就可以简化为

$$\sigma_T = \frac{4\pi}{k^2}\frac{(\hbar\Gamma_i)^2}{4(E - E_{R,i})^2 + (\hbar\Gamma_i)^2}, \tag{10.173}$$

式中的 E 表示初态的能量，也就是之前的 $E_{\mathbf{p}}$。

将这个总截面 σ_T 看成是初态能量 E 的函数，函数图像大体如图 (10.7) 所示 (请注意，公式 (10.173) 中的 k 也是能量 E 的函数)。图中的共振峰位于 $E = E_{R,i}$ 的位置，对

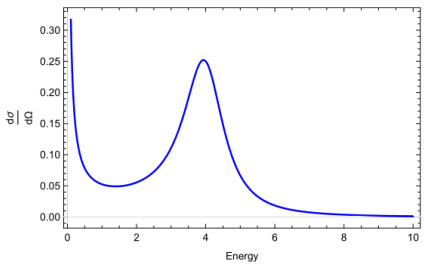

Figure 10.7: σ_T 作为初态能量 E 的函数

应于散射共振态的能量，峰的宽度大致由 $\hbar\Gamma_i/2$ 来衡量，Γ_i 越大峰越宽。图中我们画的是 $E_{R,i} = 4$, $(\hbar\Gamma_i)^2 = 2$ 的情形。也即是说，如果我们从实验上测量散射截面，然后发现有这么一个共振峰存在，那就意味着系统里必然有一个散射共振态 (或者说不稳定束缚态) 存在，而且由于 $\tau_i = 1/\Gamma_i$ 是这个不稳定态的寿命，因此如果共振峰越宽那就意味着相应的散射共振态的寿命越短，通过测量共振峰的宽度我们就能测出这个不稳定态的寿命。在粒子物理中，大量的粒子都是这么发现的，因为只有很少粒子是完全稳定的，大部分粒子实际上都不稳定，会衰变，因此会以散射共振态的形式在散射实验中显现。

10.6 习题

1. 假设在 $t_0=0$ 的初始时刻，某量子系统由哈密顿量 H_0 描述，假设我们已经完全了解了 H_0 的本征态，记为 $|n\rangle, n=1,2,3...$。设想我们缓慢地给系统加上一个微扰 W(W 是一个厄米算符)，使得 $t=+\infty$ 时系统的哈密顿量变成 H_0+W。假设整个过程可以由下面的哈密顿量 H 描述，

$$H=H_0+(1-e^{-t/\tau})W,\ 0\le t\le+\infty, \tag{10.174}$$

式中的时间尺度 τ 衡量了微扰加入的快慢。请用一阶含时微扰论计算 $t_0=0$ 时刻的 $|n\rangle$ 态在 $t=+\infty$ 时跃迁到 $|m\rangle$ 态的概率 $p_{n\to m}$。并请将 $\tau\to+\infty$ 时的答案与定态微扰论的计算进行比较。

2. 给一个线性谐振子 (哈密顿量为 $H_0=\hbar\omega(a^\dagger a+\frac{1}{2})$) 加上一个微扰 $V(t)=\hbar\omega\lambda(a^\dagger+a)e^{-t^2/\tau^2}$，请用一阶含时微扰计算出 $-\infty$ 时刻的 $|n\rangle$ 能级在 $+\infty$ 时刻跃迁到 $|m\rangle$ 能级的概率 $p_{n\to m}$。

3. 我们已经知道了 Jaynes-Cummings 哈密顿量 (JC 哈密顿量) 可以在 $|n,g\rangle$ 态和 $|n-1,e\rangle$ 态所张成的子空间里进行求解。

(1) 请在这个两态子空间里写出 JC 哈密顿量的矩阵形式，并求出它的两个本征态。

(2) 假设系统初始处在 $|n,g\rangle$ 态，请利用上一问求出来的两个本征态直接导出跃迁概率 $p_{g\to e}(t)$ 的公式。

4. 在磁共振中，我们常常让一个自旋 1/2 的粒子同时处在两个磁场中，一个是沿着 z 轴的磁场 B_0，这是一个不随时间变化的固定磁场，另外一个是在 $x-y$ 平面内旋转的磁场 $\vec{B}_1=B_1(\cos(\omega t),\sin(\omega t),0)$。系统的哈密顿量可以写成 $H=-\vec{\mu}\cdot(\vec{B}_0+\vec{B}_1)$, $\vec{\mu}$ 是例子的磁矩，它和粒子的自旋算符 $\vec{S}$ 之间的关系是，$\vec{\mu}=\gamma\vec{S}$, γ 就是所谓的旋磁比。因此这个哈密顿量也可以写成

$$H=-\gamma(B_0S_z+B_1S_x\cos(\omega t)+B_1S_y\sin(\omega t)). \tag{10.175}$$

请研究在磁场的作用下，这个粒子在自旋向上态与自旋向下态之间的拉比振荡行为，并推导出相关的公式。

5. 由于和环境或者与其它量子态的耦合，一个量子系统的定态其实总有一些不稳定性，这使得系统可以稳定地处在定态上的时间常常不是无穷大，而是有一个有限的典型时间，称之为相应定态的寿命。刻画定态寿命的一个最简单方法是给相应的定态能量加上一个负虚部，具体来说即是，考虑到有限的寿命 τ, 人们可以将一个能量为 E 的定态的能量

修正为 $E - i\frac{\hbar}{\tau}$。假设我们考察一个两能级系统，其基态 $|0\rangle$ 的能量为 E_0，寿命为无穷大，而激发态 $|1\rangle$ 的能量为 E_1，寿命为 τ。因此，以这两个定态为基，这个两能级系统的哈密顿量可以写成

$$\begin{pmatrix} E_1 - i\frac{\hbar}{\tau} & 0 \\ 0 & E_0 \end{pmatrix}. \tag{10.176}$$

现在，假设我们通过某种方式进一步将 $|0\rangle, |1\rangle$ 这两个态耦合起来，使得系统的哈密顿量变成

$$\begin{pmatrix} E_1 - i\frac{\hbar}{\tau} & W \\ W & E_0 \end{pmatrix}, \tag{10.177}$$

式中 W 是一个实数。

(1) 耦合之后开始计时，假设 $t = 0$ 时系统处在 $|0\rangle$ 态，请计算跃迁概率 $p_{0\to 0}(t)$ 以及 $p_{0\to 1}(t)$。

(2) 请就 W 很大和 W 很小两种情况画出跃迁概率随时间的变化曲线，并进行物理讨论。然后请想一想，这个系统的总概率守恒吗？为什么？



11. * 再论势散射

本章将详细地讨论一维和三维定态散射理论，并用来处理入射粒子在一个固定势场中的散射问题，称为势散射问题。我们讨论了如何从势散射的散射振幅中提取束缚态和共振态的信息。对于中心势场，我们也详细讨论了分波法。

在含时问题和散射问题这一章所发展的关于散射的一般理论中，我们说过，非微扰地分析散射幅和散射概率的关键在于求解哈密顿量 H 的散射定态。这里 H 是包含相互作用 V 的完整哈密顿量，散射定态是它的一组特殊本征态，这些散射定态的渐近形式中包含有我们需要的散射 T 矩阵信息。我们是在最一般的意义上达致这一结论的，因此这一理论框架的适用范围非常广，比方说，它适用于多粒子的散射，甚至适用于量子场论的分析，但这种一般性也使得它有些抽象。但是，如果限制于单个入射粒子在一个固定势场中的散射问题，那情况就会简化很多，这时候我们可以证明，前面关于散射定态的抽象理论其实是和我们的物理直观完全一致的。这种一致性就是我们这一章想要论述的第一个主题，我们将分别对于一维势散射问题和三维势散射问题，先从直观上给出散射定态的概念，并讨论如何从这些直观的散射定态中计算出散射概率，然后我们再分别证明这些直观的散射定态概念和前面含时问题与散射问题一章中关于散射定态的抽象定义完全一致。如此一来，一维势散射和三维势散射就成了前面关于散射的一般理论的一种具体应用。

当然，关于三维势散射的微扰理论和玻恩近似我们在前面的章节中已经系统地处理过了。但是，作为散射一般理论的应用，一维势散射和三维势散射问题能够很好地说明如何对散射问题进行非微扰分析。当然，这首先涉及到求解散射定态，其次还涉及到如何利用对称性，以及如何利用 S 矩阵的幺正性和解析性。我们将分别通过对宇称守恒的一维势散射，以及中心力场中的三维势散射的分析来说明这些非微扰分析方法的应用。在三维中心力场的情况下，这一方法通常被称为分波法。

值得说明的是，我们这里对一维势散射和三维势散射 (包括分波法) 的处理并不是最快捷的，这是因为我们想用这些特殊情形来说明更为一般性的散射问题中非微扰分析大致可以如何进行。正因为如此，对于很多结论，只要有可能我们就总是会采用一般性的有可能被推广的推导方法，而不仅仅只是为了处理一维势散射和三维势散射这样的特殊情形，如果只是为了后一目标，那这一课题当然可以被处理得更快捷一些。

11.1 一维势散射问题

所谓的一维势散射问题，我们指的是一个沿 x 轴运动的入射粒子受一个固定势场 $V(x)$ 散射的问题，在实际应用中产生这种散射势场的原因很多，比方说它可能反映的是一维材料在某个区域上的一些杂质，这时候势散射问题描写的其实就是入射粒子被这些杂质所散射的问题。我们对 $V(x)$ 的具体形式没有太多的限制，只是假定它只集中在一个有限的区域，即 $V(x)$ 在 $|x| \to \infty$ 时会以足够快的速度衰减为 0，如图 (11.1) 所示。因此在 $|x| \to \infty$

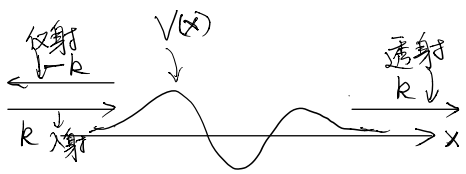

Figure 11.1: 一维势散射问题。入射粒子被势 $V(x)$ 散射，产生反射波和透射波。

的空间无穷远处，散射势可以忽略，这时候粒子可以看成是自由粒子，由自由粒子哈密顿

量 H_0 描述，$H_0=\frac{p^2}{2m}$，这里 p 为粒子的动量算符。当然，包含散射势在内的完整哈密顿量 $H=H_0+V(x)$。

对于自由粒子哈密顿量 $H_0=\frac{p^2}{2m}$。对于任何给定的大于 0 的能量 E，显然 H_0 有两个本征态，我们称之为右行波和左行波，分别记为 $\langle x|E,R\rangle$ 和 $\langle x|E,L\rangle$，它们由下式给出

$$\langle x|E,R\rangle=e^{ik_Ex},\quad \langle x|E,L\rangle=e^{-ik_Ex}, \tag{11.1}$$

式中，$k_E=\sqrt{\frac{2mE}{\hbar^2}}$(从而 $E=\frac{\hbar^2k_E^2}{2m}$)。如图 (11.2) 所示。很显然，右行波和左行波是正交的，

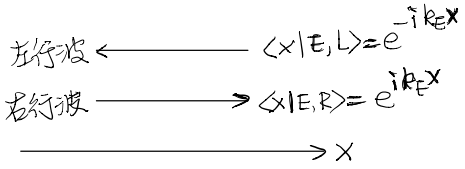

Figure 11.2: 自由粒子的右行波和左行波。

即 $\langle E,L|E,R\rangle=0$。另外，波函数当然可以相差一个整体归一化常数，而在公式 (11.1) 这种归一化选取下两个能量不同的右行波的正交归一性则是，$\langle E',R|E,R\rangle=2\pi\delta(k_{E'}-k_E)$，类似的也有左行波的正交归一性，$\langle E',L|E,L\rangle=2\pi\delta(k_{E'}-k_E)$。

11.1.1 直观的理论

如图 (11.1) 所示，假设我们考虑一个自由粒子从势 $V(x)$ 的左边入射，然后被势 $V(x)$ 所散射，一部分波反射回来，另一部分波从势中透射出去，我们所关心的就是粒子被反射的概率和透射的概率。下面我们会看到，这些概率可以通过求解定态薛定谔方程来得到，实际上，这就是求解一维势散射问题的散射定态。

我们要求解的方程是

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right)\psi(x)=E\psi(x), \tag{11.2}$$

由于我们这里关心的是散射定态，所以 $E>0$。当 $x\to-\infty$ 时，散射势 $V(x)$ 已经衰减为 0，所以粒子完全被自由粒子哈密顿量 H_0 所描述，这时候对于给定能量 $E>0$，定态解必定是右行波和左行波的某种线性叠加。从直观上，在 $x\to-\infty$ 的渐近区域，既有入射波 (这是一个右行波) 又有反射波 (这是一个左行波)，所以定态波函数必定有 $e^{ik_Ex}+re^{-ik_Ex}=\langle x|E,R\rangle+r\langle x|E,L\rangle$ 的渐近形式，这里 r 是一个复数，而且我们已经通过调整波函数 $\psi(x)$ 的整体常数系数将入射右行波前面的叠加系数设置为 1 了。类似的，在 $x\to+\infty$ 的渐近区域，粒子也成为一个自由粒子，但是，从图 (11.1) 的直观上我们可以知道，在这个区域只有右行的透射波没有左行波。因此，当 $x\to+\infty$ 时，定态波函数 $\psi(x)$ 必定渐近为 $te^{ik_Ex}=t\langle x|E,R\rangle$ 的形式。我们把在 $x\to\pm\infty$ 时具有刚才所说渐近形式的定态波函数记为 $\psi_{E,R}(x)=\langle x|\psi_{E,R}\rangle$，下标 R 表示它描述的是入射粒子从左边入射 (因此入射波是右行波) 的情形。

因此，我们现在要找的就是定态薛定谔方程 (11.2) 的某种特定类型的解 $\psi_{E,R}(x) = \langle x|\psi_{E,R}\rangle$，它具有如下渐近形式

$$\begin{aligned} x \to -\infty &: \ \langle x|\psi_{E,R}\rangle \sim \langle x|E,R\rangle + r\langle x|E,L\rangle \\ x \to +\infty &: \ \langle x|\psi_{E,R}\rangle \sim t\langle x|E,R\rangle. \end{aligned} \tag{11.3}$$

复系数 r 和 t 分别称作反射系数和透射系数，它们要通过求解满足上述渐近条件 (11.3) 的定态薛定谔方程 (11.2)，才能具体确定。

下面我们来看一下，复系数 r 和 t 是如何决定反射概率和透射概率的。我们定义反射概率 p_r 等于反射波的概率流密度的大小比上入射波概率流密度的大小 (在实验上，我们总是用一束入射粒子来做实验，因此这个定义其实说的就是，反射概率等于反射粒子流强度比上入射粒子流的强度)，反射波是 $r\langle x|E,L\rangle = re^{-ik_E x}$，很容易算得它的概率流密度大小为 $|r|^2\frac{\hbar k_E}{2m}$，类似的，入射波 $\langle x|E,R\rangle = e^{ik_E x}$ 的概率流密度为 $\frac{\hbar k_E}{2m}$，因此我们有

$$p_r = |r|^2. \tag{11.4}$$

类似的，我们定义透射概率 p_t 等于透射波概率流密度比上入射波概率流密度 (也就是透射流强比上入射流强)，最后我们也容易得出

$$p_t = |t|^2. \tag{11.5}$$

当然，由于入射粒子要么被反射，要么透射，总概率是守恒的，即入射的概率流大小等于出射 (包括反射和透射) 的概率流大小，因此必有

$$|r|^2 + |t|^2 = 1. \tag{11.6}$$

现在，假设我们让自由粒子从势 $V(x)$ 的右边入射，如图 (11.3) 所示。那这个时候入

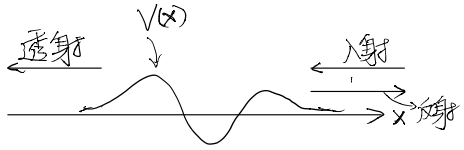

Figure 11.3: 自由粒子从势 $V(x)$ 的右边入射。

射波就是一个左行波，透射波也是一个左行波，而反射波成了一个右行波。相应的满足方程 (11.2) 的散射定态我们记作 $\psi_{E,L}(x) = \langle x|\psi_{E,L}\rangle$，它满足的渐近条件是

$$\begin{aligned} x \to +\infty &: \ \langle x|\psi_{E,L}\rangle \sim \langle x|E,L\rangle + r'\langle x|E,R\rangle \\ x \to -\infty &: \ \langle x|\psi_{E,L}\rangle \sim t'\langle x|E,L\rangle. \end{aligned} \tag{11.7}$$

同样，这时候的反射系数 r' 和透射系数 t' 也得满足概率守恒，

$$|r'|^2 + |t'|^2 = 1. \tag{11.8}$$

因此，一维势散射问题可以用四个复系数 r,t,r',t' 来刻画，这些复系数都是入射粒子能量 E 的函数。当然，独立的系数没有这么多，因为正如刚才我们所说的，它们要满足两个概率守恒关系。而且后文我们会进一步证明，实际上 r' 和 t' 并不是独立的，它们完全可以用 r,t 表达出来。

11.1.2 与散射的一般理论的联系

上一小节是我们从物理直观上得到的理论，尤其是，我们是根据直观物理图像写出散射定态所满足的渐近条件 (11.3) 和 (11.7) 的。而我们在前面的章节含时问题和散射问题中给出过散射定态的一般定义，我们知道散射定态 $|\psi_\alpha^+\rangle$ 就是哈密顿算符 H 的能写成如下形式的本征态

$$|\psi_\alpha^+\rangle = |\alpha\rangle + \frac{1}{E_\alpha - H_0 + i\varepsilon} T|\alpha\rangle. \tag{11.9}$$

那么上一小节的直观理论和这个一般性的形式化理论是不是一致呢？尤其是两者如何联系起来呢？这就是我们这一小节要探讨的内容。

首先，一般理论中的 $|\alpha\rangle$ 态是哈密顿量 H_0 的本征态，在我们这里当然就是 $|E,R\rangle$ 和 $|E,L\rangle$。不过，为了后面推导的方便，从现在开始我们改变一下 $|E,R\rangle$ 和 $|E,L\rangle$ 的归一化。具体来说就是，我们要求 $|E,R\rangle$ 归一化到能量 δ 函数上，即要求 $\langle E',R|E,R\rangle = \delta(E'-E)$ (在上一小节中我们是归一化到 $2\pi\delta(k_{E'}-k_E)$)，同样，现在我们也要求 $\langle E',L|E,L\rangle = \delta(E'-E)$。可以验证，我们只要将右行波 $\langle x|E,R\rangle$ 和左行波 $\langle x|E,L\rangle$ 作如下重新定义，就能满足这一新的归一化要求，

$$\langle x|E,R\rangle = \sqrt{\frac{m}{2\pi\hbar^2 k_E}} e^{ik_E x}, \quad \langle x|E,L\rangle = \sqrt{\frac{m}{2\pi\hbar^2 k_E}} e^{-ik_E x}. \tag{11.10}$$

现在，假设我们以指标 1 标记右行波，以指标 2 标记左行波，即令 $1=R$, $2=L$。那么自由粒子本征态的正交归一关系就可以写成

$$\langle E',i|E,j\rangle = \delta(E'-E)\delta_{ij}, \ i,j=1,2. \tag{11.11}$$

因此，这里的 (E,i) 这样的一对标记就相当于一般理论里的指标 α，相应的，一般理论中涉及的对 α 的积分在这里当然就要理解成既对连续指标 E 积分，同时也对离散指标 i 求和。

在散射的一般理论中，散射概率是由 S 矩阵和 T 矩阵决定的，它们满足

$$\langle\beta|S|\alpha\rangle = \delta(\beta-\alpha) - 2\pi i\delta(E_\beta - E_\alpha)\langle\beta|T|\alpha\rangle. \tag{11.12}$$

现在，在我们的一维势散射问题中，这一关系就是

$$\begin{aligned}\langle E',i|S|E,j\rangle &= \delta(E'-E)\delta_{ij} - 2\pi i\delta(E'-E)\langle E',i|T|E,j\rangle \\ &= \delta(E'-E)(\delta_{ij} - 2\pi i\langle E,i|T|E,j\rangle).\end{aligned} \tag{11.13}$$

$\delta(E'-E)$ 当然意味着散射初末态能量得守恒。公式 (11.13) 中的 $\delta(E'-E)$ 因子也意味着 S 矩阵关于 H_0 表象是“对角的”(关于 i,j 指标当然不对角),“对角元”是 $(\delta_{ij}-2\pi i\langle E,i|T|E,j\rangle)$, 我们常常将它简记作 $S_{ij}(E)$，即

$$\langle E',i|S|E,j\rangle=\delta(E'-E)S_{ij}(E),\ \ S_{ij}(E)=\delta_{ij}-2\pi i\langle E,i|T|E,j\rangle. \tag{11.14}$$

$S_{ij}(E)$ 构成了一个 2×2 的矩阵 $\widehat{S}(E)$，即

$$\widehat{S}(E)=\begin{pmatrix} S_{RR}(E) & S_{RL}(E)\\ S_{LR}(E) & S_{LL}(E)\end{pmatrix}. \tag{11.15}$$

不难验证，S 矩阵的幺正性现在就意味着这个 2×2 的矩阵是一个幺正矩阵，即满足

$$\widehat{S}^\dagger\widehat{S}=\widehat{S}\widehat{S}^\dagger=1. \tag{11.16}$$

由于 $\widehat{S}(E)$ 是 S 矩阵关于 H_0 表象的 “对角元”，所以在一维势散射问题中我们也常常直接把它叫做 S 矩阵。从 (11.14) 式我们还可以看到，这样定义的 S 矩阵实际上和标准的 T 矩阵 $\langle E,i|T|E,j\rangle$ 只差一个常数系数 $-2\pi i$ 以及一个相加常数 δ_{ij}。

由于我们只需要处理初末态能量相等 (均为 E) 的情形，所以后文我们也常常省略能量指标 E, 比如，后文我们常常将 $|E,i\rangle$ 态简写成 $|i\rangle$ 态，即将 $|E,R\rangle$ 简写成 $|R\rangle$, 将 $|E,L\rangle$ 简写成 $|L\rangle$。当这样简写时，我们默认粒子的能量为 E。

下面我们就是要阐明 S 矩阵 $\widehat{S}$ 和我们前面直观理论中的反射系数和透射系数之间的关系。直观上，由于这些系数代表的是出射波 (反射波和透射波都是出射波) 相对于入射波的幅度，而 S 矩阵按照定义就是入射的初态到出射的末态的跃迁幅，因此人们可以想见这些反射系数和透射系数大致应该就是 $\widehat{S}$ 的矩阵元。下面我们会利用散射定态的一般性形式理论证明这一点。由此我们也将上一小节关于散射定态的直观处理与一般性的形式定义 (11.9) 联系起来。

我们的出发点就是方程 (11.9)。在现在的一维势散射问题中，这一方程可以更具体地写成

$$\begin{aligned}|\psi_{E,R}\rangle&=|R\rangle+\frac{1}{E-H_0+i\varepsilon}T|R\rangle\\ |\psi_{E,L}\rangle&=|L\rangle+\frac{1}{E-H_0+i\varepsilon}T|L\rangle\end{aligned} \tag{11.17}$$

其中第一个式子描写的是粒子从 $V(x)$ 左边 $-\infty$ 处往右边入射的情形，第二个式子描写的粒子从 $+\infty$ 处往左边入射的情形。

根据上面的方程 (11.17), 我们有

$$\begin{aligned}\langle x|\psi_{E,R}\rangle&=\langle x|R\rangle+\langle x|\frac{1}{E-H_0+i\varepsilon}T|R\rangle\\ &=\langle x|R\rangle+\int_{-\infty}^{+\infty}dx'\langle x|\frac{1}{E-H_0+i\varepsilon}|x'\rangle\langle x'|T|R\rangle.\end{aligned} \tag{11.18}$$

接下来，我们来计算一下 $\langle x|\frac{1}{E-H_0+i\varepsilon}|x'\rangle$，利用动量本征态的封闭性关系 $\int dp|p\rangle\langle p|=1$，并注意到 $H_0|p\rangle=\frac{p^2}{2m}|p\rangle$，我们有

$$\begin{aligned}
&\langle x|\frac{1}{E-H_0+i\varepsilon}|x'\rangle=\int dp\langle x|\frac{1}{E-H_0+i\varepsilon}|p\rangle\langle p|x'\rangle\\
=&\int dp\frac{1}{E-\frac{p^2}{2m}+i\varepsilon}\langle x|p\rangle\langle p|x'\rangle=\int\frac{dp}{2\pi\hbar}\frac{1}{E-\frac{p^2}{2m}+i\varepsilon}e^{ip(x-x')/\hbar}\\
=&\int\frac{dk}{2\pi}\frac{e^{ik(x-x')}}{\frac{\hbar^2k_E^2}{2m}-\frac{\hbar^2k^2}{2m}+i\varepsilon}=-\frac{2m}{\hbar^2}\int_{-\infty}^{+\infty}\frac{dk}{2\pi}\frac{e^{ik(x-x')}}{k^2-k_E^2-i\varepsilon}.
\end{aligned}\tag{11.19}$$

为了算出上式最后那个对 k 的积分，我们首先注意到，$\frac{1}{k^2-k_E^2-i\varepsilon}$ 在 k 的复平面上有两个极点 $k_E+i\varepsilon$ 和 $-k_E-i\varepsilon$，如图 (11.4) 所示。另外，(11.19) 式最后对 k 的积分是沿着实轴积

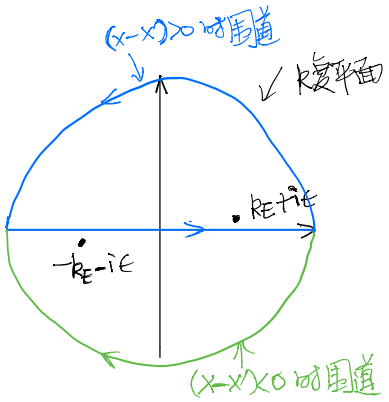

Figure 11.4: k 的极点与积分围道。

分，但是，当 $x-x'>0$ 时，我们可以把这个积分加上 k 复平面在上半平面无穷远处的一个半圆，形成一个闭合的围道积分，如图 (11.4) 所示。这是因为，当 $x-x'>0$ 时，(11.19) 式中的指数因子 $e^{ik(x-x')}$ 在复上半平面是指数衰减的 (因为这时候 k 有一个正虚部)，在无穷远处的半圆上衰减为 0，因此实际上这时候无穷大半圆上的积分贡献是 0，也即是说围道积分其实就等于沿着实轴的积分。类似的，当 $x-x'<0$ 时，我们可以把沿实轴的积分加上 k 复平面在下半平面无穷远处的一个半圆，形成在下半平面的一个闭合围道积分，如图 (11.4) 所示。利用围道积分的留数法，我们就可以将这两种情形时 (11.19) 式最后的积分结果都算出来，最后的结果可以概括如下

$$\begin{aligned}
&\langle x|\frac{1}{E-H_0+i\varepsilon}|x'\rangle\\
=&-2\pi i\left(\sqrt{\frac{m}{2\pi\hbar^2k_E}}\right)^2\left[\theta(x-x')e^{ik_E(x-x')}+\theta(x'-x)e^{-ik_E(x-x')}\right]\\
=&-2\pi i\left[\theta(x-x')\langle x|R\rangle\langle R|x'\rangle+\theta(x'-x)\langle x|L\rangle\langle L|x'\rangle\right].
\end{aligned}\tag{11.20}$$

式中 $\theta(x)$ 是阶跃函数，它的定义是，当 $x>0$ 时 $\theta(x)=1$，当 $x<0$ 时 $\theta(x)=0$。(11.20) 式的最后一个等号是利用了右行波 $\langle x|R\rangle$ 和左行波 $\langle x|L\rangle$ 的定义式 (11.10)。

将 (11.20) 的最终结果代入散射定态波函数的方程 (11.18), 我们就可以得到

$$\begin{aligned}&\langle x|\psi_{E,R}\rangle = \langle x|R\rangle + \\ &-2\pi i\left[\langle x|R\rangle\int_{-\infty}^{x}dx'\langle R|x'\rangle\langle x'|T|R\rangle + \langle x|L\rangle\int_{x}^{+\infty}dx'\langle L|x'\rangle\langle x'|T|R\rangle\right].\end{aligned} \tag{11.21}$$

因此，当 $x\to-\infty$ 时，我们有

$$\begin{aligned}\langle x|\psi_{E,R}\rangle &\sim \langle x|R\rangle - 2\pi i\left[\langle x|L\rangle\int_{-\infty}^{+\infty}dx'\langle L|x'\rangle\langle x'|T|R\rangle\right] \\ &= \langle x|R\rangle - 2\pi i\langle x|L\rangle\langle L|T|R\rangle.\end{aligned} \tag{11.22}$$

另一方面，根据 (11.14) 式，我们有 $S_{LR}=-2\pi i\langle L|T|R\rangle$, 因此我们有，$x\to-\infty$: $\langle x|\psi_{E,R}\rangle\sim\langle x|R\rangle+\langle x|L\rangle S_{LR}$。

类似的，根据 (11.21) 式可知，当 $x\to+\infty$ 时，我们有

$$\begin{aligned}\langle x|\psi_{E,R}\rangle &\sim \langle x|R\rangle - 2\pi i\left[\langle x|R\rangle\int_{-\infty}^{+\infty}dx'\langle R|x'\rangle\langle x'|T|R\rangle\right] \\ &= \langle x|R\rangle - 2\pi i\langle x|R\rangle\langle R|T|R\rangle.\end{aligned} \tag{11.23}$$

另一方面，根据 (11.14) 式，我们有 $S_{RR}=1-2\pi i\langle R|T|R\rangle$, 因此我们有，$x\to+\infty$: $\langle x|\psi_{E,R}\rangle\sim\langle x|R\rangle S_{RR}$。

简单归纳一下就是，通过上面两段的推导，我们得到了下面的结果

$$\begin{aligned}&x\to-\infty:\ \langle x|\psi_{E,R}\rangle\sim\langle x|R\rangle+\langle x|L\rangle S_{LR} \\ &x\to+\infty:\ \langle x|\psi_{E,R}\rangle\sim\langle x|R\rangle S_{RR}.\end{aligned} \tag{11.24}$$

将这个结果 (11.24) 与上一小节的 (11.3) 式进行比较，我们就有

$$S_{LR}=r,\ S_{RR}=t. \tag{11.25}$$

当然，读者可能注意到，上一小节的 (11.3) 式与这里的 (11.24) 式对右行波 $|R\rangle$ 以及对左行波 $|L\rangle$ 所采用的归一化是不同的。但这种不同完全不影响这两者的比较，因为相差一个整体的归一化常数对波函数 $\langle x|\psi_{E,R}\rangle$ 并不产生什么实质的影响。(11.25) 式这个结果非常好理解，S_{RR} 表示入射右行波到出射右行波的散射，这当然对应的是透射，因此等于 t，而 S_{LR} 表示入射右行波到出射左行波的散射，那当然对应的是反射，因此等于 r。

我们可以对入射粒子从 $V(x)$ 右边 $+\infty$ 处往左边入射的情形作完全类似的推导，最后会得到

$$\begin{aligned}&x\to-\infty:\ \langle x|\psi_{E,L}\rangle\sim\langle x|L\rangle S_{LL} \\ &x\to+\infty:\ \langle x|\psi_{E,L}\rangle\sim\langle x|L\rangle+\langle x|R\rangle S_{RL}.\end{aligned} \tag{11.26}$$

将这个式子与上一小节的 (11.7) 式比较，我们就有

$$S_{RL}=r',\ S_{LL}=t'. \tag{11.27}$$

综合 (11.25) 和 (11.27)，我们就有

$$\widehat{S}(E)=\begin{pmatrix} S_{RR} & S_{RL} \\ S_{LR} & S_{LL} \end{pmatrix}=\begin{pmatrix} t & r' \\ r & t' \end{pmatrix}. \tag{11.28}$$

可见，从散射的形式化理论开始推导出来的结果和上一小节的直观理论是完全一致的。形式化的推导虽然没有那么直接，但它可以看作是对上一小节的直观理论的一个数学证明。并且，通过这个推导，我们将反射系数和透射系数与 S 矩阵联系了起来，这就使得我们可以利用 S 矩阵的幺正性和解析性来进一步得到几个反射系数和透射系数之间的关系。而且，上面的推导过程具体地演示了如何从散射定态解中得到 T 矩阵和 S 矩阵的信息，而这样的处理不仅只适用于一维势散射问题，而是具有一般性的。形式化处理的最大好处也许是，它是一个统一性的理论，我们可以统一性地用它来处理所有的散射问题，不管是这里处理的一维势散射也好，还是后面将要讨论的三维势散射也好，还是其它更复杂的散射问题也好，我们都可以用同一套散射的一般理论来处理。

当然，我们可以考察一个更为一般的散射定态 $|\psi_E\rangle$，它是 $|\psi_{E,L}\rangle$ 和 $|\psi_{E,R}\rangle$ 的线性叠加，

$$|\psi_E\rangle=c_L|\psi_{E,L}\rangle+c_R|\psi_{E,R}\rangle. \tag{11.29}$$

那么根据 (11.24) 式和 (11.26) 式，我们可以看到，$|\psi_E\rangle$ 将满足如下渐近条件

$$\begin{aligned} &x\to-\infty:\ |\psi_E\rangle\sim c_R|R\rangle+(S_{LR}c_R+S_{LL}c_L)|L\rangle \\ &x\to+\infty:\ |\psi_E\rangle\sim c_L|L\rangle+(S_{RR}c_R+S_{RL}c_L)|R\rangle. \end{aligned} \tag{11.30}$$

很显然，这两个渐近表达式的第一项都是代表的入射波，而第二项都是代表的出射波 (比方说在 $x\to-\infty$ 的区域，也就是 $V(x)$ 的左边，入射波当然就是右行波，而左行波当然就是出射波)。(11.30) 式告诉我们，在 $\{|R\rangle,|L\rangle\}$ 表象下，入射波的列矢量 ψ_i 以及出射波的列矢量 ψ_o 可以分别表示为

$$\psi_i=\begin{pmatrix} c_R \\ c_L \end{pmatrix},\ \psi_o=\begin{pmatrix} S_{RR} & S_{RL} \\ S_{LR} & S_{LL} \end{pmatrix}\begin{pmatrix} c_R \\ c_L \end{pmatrix}=\widehat{S}\psi_i. \tag{11.31}$$

可见，S 矩阵正是把出射波的幅度与入射波的幅度联系起来的矩阵。而 S 矩阵的幺正性 (11.16) 所说的无非是，出射波的总概率等于入射波的总概率，也就是概率守恒。

(11.28) 告诉我们，反射系数和透射系数其实是 S 矩阵元，因此我们可以应用 S 矩阵的幺正性 (11.16)。由 $\widehat{S}^\dagger\widehat{S}=1$ 可以知道，

$$\begin{pmatrix} t^* & r^* \\ r'^* & t'^* \end{pmatrix}\begin{pmatrix} t & r' \\ r & t' \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{11.32}$$

从而必有

$$|t|^2+|r|^2=|t'|^2+|r'|^2=1,\ \frac{t'}{t^*}=-\frac{r'}{r}. \tag{11.33}$$

这个结果中的第一个式子是我们在上一小节中已经根据概率守恒得到过的，现在我们从 S 矩阵的幺正性再一次得到了它们。这是不令人意外的，因为正如我们在上一段所说的，幺正性的物理本质就是概率守恒。但是由于我们现在考虑的是更一般情形下的概率守恒，所以我们还额外得到一个关系式。不仅如此，假设我们进一步利用 $\widehat{SS}^{\dagger}=1$，那将还可以得到

$$|t|^2+|r'|^2=|t'|^2+|r|^2=1,\ \frac{t'}{t^*}=-\frac{r^*}{r'^*}. \tag{11.34}$$

比较 (11.33) 式的第一个式子和 (11.34) 式的第一个式子，我们有

$$|r'|^2=|r|^2,\ |t'|^2=|t|^2. \tag{11.35}$$

这告诉我们粒子从 $V(x)$ 右边入射的反射概率以及透射概率分别与粒子从左边入射的反射概率以及透射概率相等。当然，有了这个结果以后，那么 (11.33) 式的第二个式子和 (11.34) 式的第二个式子实际上就是一回事了。

不仅如此，注意到我们这里并没有涉及到粒子的自旋，尤其是没有涉及到自旋与外磁场耦合等破坏时间反演对称性的相互作用，因此我们的系统是时间反演不变的。在时间反演之下，粒子的动量 p 将变成 $-p$，即 $p\to -p$，这就告诉我们，在时间反演之下，左行波会变成右行波，而右行波会变成左行波，即 $|E,R\rangle\to|E,L\rangle$，$|E,L\rangle\to|E,R\rangle$。另外很显然的是，在时间反演之下，初末态也将调换，初态将变成末态，末态将变成初态。因此，在时间反演之下，$\langle E',R|S|E,R\rangle$(这里的 S 是散射一般理论中的 S 算符) 将变成 $\langle E,L|S|E',L\rangle$，$\langle E',L|S|E,R\rangle$ 将变成 $\langle E,L|S|E',R\rangle$，即 $\langle E',R|S|E,R\rangle\to\langle E,L|S|E',L\rangle$，$\langle E',L|S|E,R\rangle\to\langle E,L|S|E',R\rangle$。系统有时间反演对称性，那就意味着时间反演之后的结果和反演之前相等，即有

$$\langle E',R|S|E,R\rangle=\langle E,L|S|E',L\rangle,\ \langle E',L|S|E,R\rangle=\langle E,L|S|E',R\rangle. \tag{11.36}$$

另一方面，根据矩阵 $\widehat{S}(E)$ 的定义 (11.14)，$\langle E',i|S|E,j\rangle=\delta(E'-E)S_{ij}(E)$，因此上面的式子就告诉我们

$$S_{RR}(E)=S_{LL}(E). \tag{11.37}$$

用透射系数来写，这个结果即是 $t'=t$。结合 (11.33) 式的第二个式子，我们有

$$t'=t,\ r'=-r\frac{t}{t^*}. \tag{11.38}$$

正因为这个结果，我们前面才说，系数 r',t' 可以用 r,t 表达出来。

11.1.3 空间反演对称性

上一小节最后通过对时间反演对称性的分析，我们已经看到对称性如何可以限制 S 矩阵。这一小节我们进一步考虑空间反演对称性可以如何简化一维势散射问题。

我们假定势函数具有 $V(-x)=V(x)$ 的形式，如图 (11.5) 所示。习惯上，人们常常定义空间反演算符 P，

$$P|x\rangle=|-x\rangle,\ \ P^2=1. \tag{11.39}$$

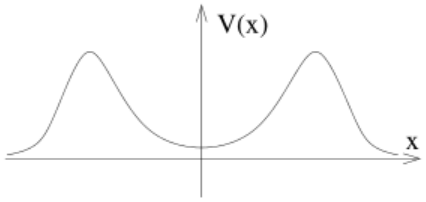

Figure 11.5: 具有空间反演对称性的势函数。

假设我们将系统的哈密顿算符记作 $H(x)=-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)$, 则由于 $V(-x)=V(x)$, 因此 $PH(x)=H(-x)P=H(x)P$, 另外由于 $H_0=-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$ 也是 x 的偶函数, 所以很显然也有 $PH_0=H_0P$, 即

$$[P,H_0]=[P,H]=0. \tag{11.40}$$

现在让我们来考察空间反演算符在左行波和右行波上的作用。由于 $P|x\rangle=|-x\rangle$, 则 $\langle x|P=\langle -x|P$。现在让我们来考察 $P|R\rangle$, 由于 $\langle x|P|R\rangle=\langle -x|R\rangle=\sqrt{\frac{m}{2\pi\hbar^2k_E}}e^{-ik_Ex}=\langle x|L\rangle$, 因此即有 $P|R\rangle=|L\rangle$, 类似的 $P|L\rangle=|R\rangle$, 即

$$P|R\rangle=|L\rangle,\quad P|L\rangle=|R\rangle. \tag{11.41}$$

下面我们进一步来考察 P 在散射定态 $|\psi_{E,R}\rangle$ 和 $|\psi_{E,L}\rangle$ 上的作用。根据我们在散射的一般理论中的讨论, 散射定态满足一个形式化的定义 $|\psi_\alpha^+\rangle=U_I(0,-\infty)|\alpha\rangle$, 在我们这里即有 $|\psi_{E,R}\rangle=U_I(0,-\infty)|R\rangle$, $|\psi_{E,L}\rangle=U_I(0,-\infty)|L\rangle$。另一方面, 根据定义, 时间演化算符 $U_I(t_2,t_1)=e^{i\frac{H_0}{\hbar}t_2}U(t_2,t_1)e^{-i\frac{H_0}{\hbar}t_1}=e^{i\frac{H_0}{\hbar}t_2}e^{-i\frac{H}{\hbar}(t_2-t_1)}e^{-i\frac{H_0}{\hbar}t_1}$。由于空间反演算符 P 与 H_0 以及 H 均对易 (11.40), 因此很显然 $PU_I(t_2,t_1)=U_I(t_2,t_1)P$。由此即有 $P|\psi_{E,R}\rangle=PU_I(0,-\infty)|R\rangle=U_I(0,-\infty)P|R\rangle=U_I(0,-\infty)|L\rangle=|\psi_{E,L}\rangle$, 同理有 $P|\psi_{E,L}\rangle=|\psi_{E,R}\rangle$, 即

$$P|\psi_{E,R}\rangle=|\psi_{E,L}\rangle,\quad P|\psi_{E,L}\rangle=|\psi_{E,R}\rangle. \tag{11.42}$$

由于 P 与时间演化算符 U_I 对易, 而 S 算符是一个特殊的时间演化算符 $U_I(+\infty,-\infty)$, 所以 $PS=SP$。由此即有 $\langle E',R|S|E,R\rangle=\langle E',L|PSP|E,L\rangle=\langle E',L|SP^2|E,L\rangle=\langle E',L|S|E,L\rangle$, $\langle E',R|S|E,L\rangle=\langle E',R|SP|E,R\rangle=\langle E',R|PS|E,R\rangle=\langle E',L|S|E,R\rangle$, 也即是 $\langle E',R|S|E,R\rangle=\langle E',L|S|E,L\rangle$, $\langle E',R|S|E,L\rangle=\langle E',L|S|E,R\rangle$。类似的也有 $\langle E',L|S|E,R\rangle=\langle E',R|S|E,L\rangle$。由于 $\langle E',i|S|E,j\rangle=\delta(E'-E)S_{ij}(E)$, 因此即有

$$S_{RR}=S_{LL},\quad S_{RL}=S_{LR}. \tag{11.43}$$

或者用反射系数和透射系数来写即是

$$t'=t,\quad r'=r. \tag{11.44}$$

更方便的处理方式是重新定义基矢 $|+\rangle$ 和 $|-\rangle$，它们的定义是

$$|+\rangle = \frac{1}{\sqrt{2}}(|R\rangle + |L\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|R\rangle - |L\rangle). \tag{11.45}$$

根据 (11.41)，很显然 $|+\rangle$、$|-\rangle$ 均为空间反演算符 P 的本征态，

$$P|+\rangle = |+\rangle, \quad P|-\rangle = -|-\rangle. \tag{11.46}$$

即 $|+\rangle$ 为正宇称态，$|-\rangle$ 为负宇称态。

在新基矢下，有 $\langle E',-|S|E,+\rangle = \langle E',-|SP|E,+\rangle = \langle E',-|PS|E,+\rangle = -\langle E',-|S|E,+\rangle$，因此必有 $\langle E',-|S|E,+\rangle = 0$，同理 $\langle E',+|S|E,-\rangle = 0$。这就意味着在这个新的基矢下 $S_{+-}(E) = S_{-+}(E) = 0$，$S$ 矩阵的非 0 矩阵元为 $S_{++}(E)$ 和 $S_{--}(E)$。新基矢下 S 矩阵的幺正性就变成

$$|S_{++}(E)|^2 = |S_{--}(E)|^2 = 1. \tag{11.47}$$

这就意味着 S_{++} 和 S_{--} 都是一个相因子，有时侯也记作 $S_{++}(E) = e^{2i\delta_+(E)}$，$S_{--}(E) = e^{2i\delta_-(E)}$。当然，根据基矢之间的变换关系，我们也有

$$S_{++} = S_{RR} + S_{RL} = t + r, \quad S_{--} = S_{RR} - S_{RL} = t - r. \tag{11.48}$$

同样，我们可以定义新的散射定态 $|\psi_{E,+}\rangle$ 和 $|\psi_{E,-}\rangle$，它们的定义是

$$|\psi_{E,+}\rangle = (|\psi_{E,R}\rangle + |\psi_{E,L}\rangle), \quad |\psi_{E,-}\rangle = (|\psi_{E,R}\rangle - |\psi_{E,L}\rangle). \tag{11.49}$$

根据 (11.42) 式可知，它们也是 P 的本征态，满足

$$P|\psi_{E,+}\rangle = |\psi_{E,+}\rangle, \quad P|\psi_{E,-}\rangle = -|\psi_{E,-}\rangle. \tag{11.50}$$

则根据散射定态解的渐近形式 (11.30) 式，我们有

$$\begin{aligned} x \to -\infty &: \ |\psi_{E,+}\rangle \sim |R\rangle + (S_{LR} + S_{LL})|L\rangle = |R\rangle + S_{++}|L\rangle \\ x \to +\infty &: \ |\psi_{E,+}\rangle \sim |L\rangle + (S_{RR} + S_{RL})|R\rangle = |L\rangle + S_{++}|R\rangle. \end{aligned} \tag{11.51}$$

同理也有

$$\begin{aligned} x \to -\infty &: \ |\psi_{E,-}\rangle \sim |R\rangle + (S_{LR} - S_{LL})|L\rangle = |R\rangle - S_{--}|L\rangle \\ x \to +\infty &: \ |\psi_{E,-}\rangle \sim -|L\rangle + (S_{RR} - S_{RL})|R\rangle = -|L\rangle + S_{--}|R\rangle. \end{aligned} \tag{11.52}$$

也即是说，通过研究有确定宇称的散射定态的渐近形式我们就可以得到 S 矩阵的信息。

11.1.4 一维势散射举例

让我们考虑一个如图 (11.6) 所示的对称方形势阱的例子。让我们来考察它的偶宇称散射定态 $|\psi_{E,+}\rangle$，相应的波函数记为 $\psi_{E,+}(x)$。则根据 (11.51) 式可知

$$\begin{aligned} x \le -\frac{a}{2} &: \ \psi_{E,+}(x) = e^{ik_E x} + S_{++}(E)e^{-ik_E x} \\ x \ge +\frac{a}{2} &: \ \psi_{E,+}(x) = e^{-ik_E x} + S_{++}(E)e^{ik_E x}. \end{aligned} \tag{11.53}$$

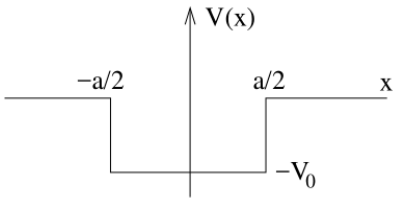

Figure 11.6: 对称的方形势阱。

另外，很显然，$[-a/2,+a/2]$ 区间上的偶函数解必定有如下形式

$$x\in[-a/2,+a/2]:\ \psi_{E,+}(x)=A(e^{iq_Ex}+e^{-iq_Ex}),\tag{11.54}$$

式中 $q_E>0$ 满足 $\frac{\hbar^2q_E^2}{2m}=E+V_0$。利用波函数和波函数的一阶导数在 $x=-a/2$(或者 $x=a/2$)处的连续性条件，我们可以得到

$$S_{++}(E)=e^{-ik_Ea}\frac{1+i(q_E/k_E)\tan(q_Ea/2)}{1-i(q_E/k_E)\tan(q_Ea/2)}.\tag{11.55}$$

类似的，人们也可以考察奇宇称的散射定态 $|\psi_{E,-}\rangle$, 并进而得到

$$S_{--}(E)=e^{-ik_Ea}\frac{1+i(k_E/q_E)\tan(q_Ea/2)}{1-i(k_E/q_E)\tan(q_Ea/2)}.\tag{11.56}$$

求出 S_{++} 和 S_{--} 以后，人们当然也可以利用 (11.48) 式求出反射系数和透射系数。

束缚态

上面求出来的 S 矩阵 (11.55) 和 (11.56) 适用于 $E>0$ 情形。但是假如我们将它们解析延拓到 $E<0$ 的情形，或者等价的，将 k_E 解析延拓到正虚轴上，之所以是正虚轴而不能是负虚轴，是因为 $k_E=\sqrt{\frac{2mE}{\hbar^2}}$，而且我们的解析延拓是从能量复平面正实轴割线的上沿开始的 (即是从 $E^+=E+i\varepsilon$ 开始的)，因此当 E 延拓到负实轴上的时候，相应的 k_E 应该延拓到正虚轴。这样解析延拓以后，我们就可能发现 $S_{++}(E)$ 和 $S_{--}(E)$ 的极点。根据我们在散射的一般理论中的论述，T 矩阵在能量复平面实轴上的极点对应于哈密顿量 H 的束缚态。由于我们这里的 S 矩阵和 T 矩阵其实是密切相关的。因此这也就是说，$S_{++}(E)$ 和 $S_{--}(E)$ 在负实轴上的极点对应的是系统的束缚态。

为了加深对这种 S 矩阵在能量负实轴上的极点与系统束缚态之间的对应关系的理解，我们现在通过简单考察散射定态在解析延拓下的行为来作进一步说明。假设我们将能量 E 延拓到负实轴，相应的 k_E 就解析延拓为 $k_E=i\lambda_E(\lambda_E>0)$。根据 (11.53) 式，这时候散射定态波函数在 $x\geq+\frac{a}{2}$ 区间上将延拓为

$$e^{\lambda_Ex}+S_{++}(E)e^{-\lambda_Ex},\tag{11.57}$$

很显然，一般来说这样的波函数将是无法归一化的，从而解析延拓后的散射定态一般来说不对应任何物理态。但是，有一种情形例外，那就是如果延拓后的能量刚好是 $S_{++}(E)$ 的

极点的话，那么 (11.57) 式的第二项将有一个无穷大的幅度，因此不可归一化的第一项相比就可以忽略 (为了看得更清楚，人们可以将表达式 (11.57) 整体除以一个常数 $S_{++}(E)$，从而看到在极点处 (11.57) 的第一项相对来说为 0)，从而使得解析延拓后的散射定态在 $x \geq +\frac{a}{2}$ 区间上实际上正比于 $e^{-\lambda_E x}$，这样一个在无穷远处衰减为 0 的波函数所描述的当然就是一个束缚态。

假设我们将 k_E 解析延拓为 $k_E = i\lambda_E$，因此 $E = -\frac{\hbar^2\lambda_E^2}{2m} < 0$，从而 $\frac{\hbar^2 q_E^2}{2m} = E + V_0 = -\frac{\hbar^2\lambda_E^2}{2m} + V_0$。则从表达式 (11.55) 很容易看出，$S_{++}(E)$ 的极点由下面的方程决定

$$\lambda_E = q_E \tan(q_E a/2), \quad \lambda_E^2 + q_E^2 = \frac{2mV_0}{\hbar^2}. \tag{11.58}$$

这组方程的解对应于偶宇称的束缚态。同样的道理，根据 $S_{--}(E)$ 的极点，我们也可以得到奇宇称束缚态所满足的方程

$$q_E = -\lambda_E \tan(q_E a/2), \quad \lambda_E^2 + q_E^2 = \frac{2mV_0}{\hbar^2}. \tag{11.59}$$

不难验证，通过直接求解束缚态情形下的定态薛定谔方程，我们也正好可以得到 (11.58) 和 (11.59)，这就说明了，H 的束缚态的确对应于 S 矩阵解析延拓以后在能量复平面实轴上的极点。

我们可以用画图的方法来求解方程组 (11.58) 和方程组 (11.59)，如图 (11.7) 所示，方程组 (11.58) 对应左图，方程组 (11.59) 对应右图。从图中很容易看出，对于偶宇称情形，

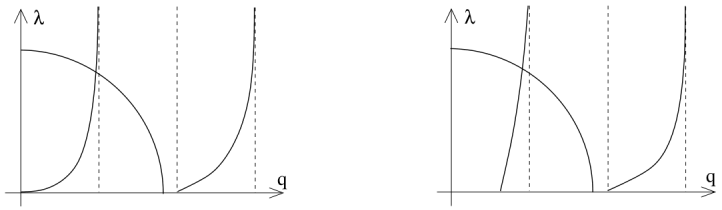

Figure 11.7: 束缚态解。左图对应偶宇称束缚态，右图对应奇宇称束缚态。

无论 $V_0 > 0$ 具体取多少，至少存在一个束缚态。但是，对于奇宇称情形，如果 $V_0 > 0$ 的值过小的话，奇宇称束缚态有可能不存在。

11.1.5 共振态

上一小节我们讨论了如何通过将 S 矩阵解析延拓进而找到系统的束缚态。然而，如果势函数 $V(x)$ 具有的是如图 (11.5) 所示的形状，那么严格的束缚态是不存在的。因为这时候粒子虽然可以暂时性地被束缚在势阱之中，但是，它也有一定的概率通过量子隧穿效应隧穿到势阱外面去。因此，这时候没有稳定的束缚态存在，只要等待的时间足够长，粒子必定会出现在势阱之外，成为一个自由粒子。但是，如果粒子最初处在势阱之内，并且我们等待的时间也并不太长，那么直观上很显然的是，粒子将有更大的概率依然被束缚在势阱之内。也即是说，在一定的意义上，束缚态的概念在这里依然适用，不过，这时候的束缚态不是一个稳定束缚态，而是一个会随时间衰变的束缚态。

上一章在散射的一般理论中我们说过，这种不稳定的束缚态对应的是 T 矩阵在能量复平面上实轴之外的极点，更严格地说是在能量复平面下半平面上的极点。在我们这里也就是，不稳定束缚态对应 S 矩阵在能量复平面下半平面的极点。以偶宇称的不稳定束缚态为例，这时候，它们对应于 $S_{++}(E)$ 在下半平面的极点。比方说，假设其中一个这样的极点为 $E_R - i\frac{\Gamma}{2}\hbar$, 则由于 $S_{++}(E)$ 本身是一个相因子，因此在此极点附近必有

$$S_{++}(E) \sim \frac{E - E_R - i\frac{\Gamma}{2}\hbar}{E - E_R + i\frac{\Gamma}{2}\hbar}. \tag{11.60}$$

根据我们在上一章中的讨论，$\tau = 1/\Gamma$ 就是这个不稳定束缚态的典型寿命。直观上，这是因为，这个不稳定束缚态的能量是一个复数 $E = E_R - i\frac{\Gamma}{2}\hbar$, 它随时间的演化将满足

$$e^{-i\frac{E}{\hbar}t} = e^{-i\frac{E_R}{\hbar}t} e^{-\frac{\Gamma}{2}t}. \tag{11.61}$$

因此，$t = 0$ 时刻处于这一态上的粒子在 t 时刻依然处于这个态上的概率将为 $e^{-\Gamma t} = e^{-\frac{t}{\tau}}$。

这种不稳定束缚态也称作散射共振态。因为，在散射实验中，假设入射粒子能量接近 E_R, 那么它将和这个不稳定束缚态共振，从而使得入射粒子短暂地处于这个不稳定束缚态上，然后再衰变为出射自由粒子态。

举例

下面我们举一个会出现这种散射共振态的具体例子。假设我们考虑双 δ 函数势

$$V(x) = V_0 \frac{\hbar^2}{2m}\big[\delta(x-1) + \delta(x+1)\big], \tag{11.62}$$

式中 $V_0 > 0$ 为这个 δ 函数势的强度。

假设我们考虑偶宇称散射定态 $|\psi_{E,+}\rangle$。则在 $|x| \geq 1$ 的地方，散射定态波函数将具有 (11.53) 的形式。而在 $x \in [-1,1]$ 区间上，散射定态波函数将具有 $\psi_{E,+}(x) = A(e^{ik_E x} + e^{-ik_E x})$ 的形式。在 $x = \pm 1$ 处，波函数依然是连续的，但是波函数的一阶导数并不连续，比方说，将相应的定态薛定谔方程在 $x = 1$ 附近积分我们可以得到

$$\psi'(1^+) - \psi'(1^-) = V_0 \psi(1), \tag{11.63}$$

式中 $\psi'(x)$ 表示 $\psi(x)$ 对 x 的一阶导数，$1^+ = 1 + \varepsilon$, $1^- = 1 - \varepsilon$，$\varepsilon > 0$ 为一个无穷小量。

由波函数的连续性，以及波函数的一阶导数所满足的方程 (11.63), 我们就可以得到

$$S_{++}(E) = e^{-2ik_E} \frac{1 + i(\tan k_E - \frac{V_0}{k_E})}{1 - i(\tan k_E - \frac{V_0}{k_E})}. \tag{11.64}$$

现在，将能量延拓到复平面上，或者等价地将 k_E 延拓到复平面上，则我们可以找到 S_{++} 的极点，它们满足 $1 - i\tan k_E + i\frac{V_0}{k_E} = 0$，整理一下即是

$$e^{2ik_E} = -\left(1 - i\frac{2k_E}{V_0}\right). \tag{11.65}$$

很容易证明，对于 V_0 取一个足够大的有限值情形，在 k_E 复平面上沿着实轴或者沿着虚轴，都找不到这个方程的解。因此，一般来说，(11.65) 的解是一个不为纯虚数的复数，对应的能量 $E=\frac{\hbar^2k_E^2}{2m}$ 当然也必定是复数。也即是说，$S_{++}(E)$ 的极点不在能量复平面的实轴上，因此相应于不稳定束缚态。

假设 V_0 足够大，那我们可以将 (11.65) 的解按照 $1/V_0$ 进行级数展开，人们容易验证，下面给出的 k_E 直到 $(1/V_0)^2$ 阶为止都满足方程 (11.65)

$$k_E=\frac{\pi}{2}(1-\frac{1}{V_0}+\frac{1}{V_0^2})-i(\frac{\pi}{2V_0})^2. \tag{11.66}$$

相应的能量极点 $E=\frac{\hbar^2k_E^2}{2m}$ 为

$$E=E_R-i\frac{\hbar}{2}\left(\frac{\hbar\pi^3}{4mV_0^2}\right), \tag{11.67}$$

式中 E_R 的近似表达式我们留给读者自己去推导。这个式子告诉我们，直到 $(1/V_0)^2$ 阶，这个不稳定束缚态的 $\Gamma=\frac{\hbar\pi^3}{4mV_0^2}$。

11.2 三维势散射问题

所谓的三维势散射问题，指的是一束入射粒子被一个固定势场 $V(\mathbf{x})$ 散射的问题，如图 (11.8) 所示。我们要求势场 $V(\mathbf{x})$ 在无穷远处衰减得足够快，以使得势场的有效作用范围有限。整个系统的哈密顿量 H 为，$H=H_0+V(\mathbf{x})$，其中 $H_0=-\frac{\hbar^2}{2m}\nabla^2$ 为自由粒子哈密

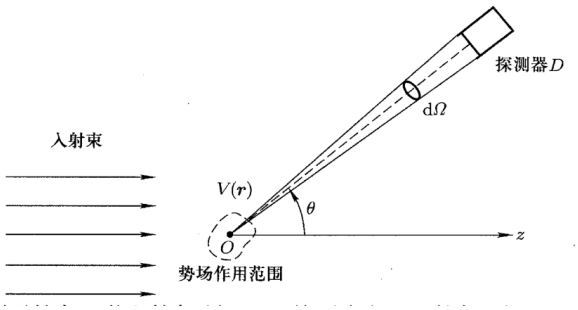

Figure 11.8: 入射粒子流受固定势场散射。

顿量。

首先让我们来回顾一下散射截面的概念。假设如图 (11.8) 所示，入射粒子束沿着水平的 z 轴入射，每单位时间单位横截面积上的入射粒子通量为 F_i，显然，F_i 等于入射粒子总数与入射粒子概率流密度 J_i 的乘积，$F_i=J_iN$。假设势场 $V(\mathbf{x})$ 定域在坐标原点 O 附近。在 θ,φ 方向上，在远离散射势场的地方，我们放置有一个探测器，它对 O 点的张角为立体角 $d\Omega$。经过势场的散射以后，入射粒子有一定的概率被散射到这个立体角 $d\Omega$ 里面去，从而被探测到。假设单位时间探测器探测到的粒子数为 dN，很显然 dN 等于入射粒子总数 N 与跃迁概率 $w_{i\to f}$ 的乘积，$dN=w_{i\to f}N$，式中 $w_{i\to f}$ 表示单个粒子单位时间之内从入射

初态跃迁 (散射) 到 $d\Omega$ 所表示的连续末态的概率。根据我们以前的定义，微分散射截面 $d\sigma = w_{i\to f}/J_i = w_{i\to f}N/(J_iN) = dN/F_i$, 即等于探测器单位时间之内探测到的粒子数与入射粒子通量的比值。很显然，探测器探测到的粒子数将正比于 $d\Omega$, 因此 $d\sigma$ 也正比于 $d\Omega$, 常常也写作

$$d\sigma = \sigma(\theta,\varphi)d\Omega. \tag{11.68}$$

人们有时候也把 $\sigma(\theta,\varphi)$ 称为微分散射截面。

为了后文的推导，让我们先来研究一下如果散射势不存在，粒子完全自由时的行为。很显然，这时候给定能量 $E>0$, 设 $E=\frac{\hbar^2k_E^2}{2m}(k_E>0)$, 则自由粒子所满足的定态薛定谔方程为

$$(\nabla^2+k_E^2)\phi_E(\mathbf{x})=0, \tag{11.69}$$

式中 $\phi_E(\mathbf{x})$ 为能量为 E 的自由粒子定态波函数。方程 (11.69) 有一类非常平凡的平面波解

$$\phi_E(\mathbf{x}) = e^{i\mathbf{k}_E\cdot\mathbf{x}}, \tag{11.70}$$

式中 $\mathbf{k}_E$ 为任意一个满足 $|\mathbf{k}_E|=k_E$ 的波矢量。很显然，这个解描述的是一个动量为 $\hbar\mathbf{k}_E$ 的自由粒子。

为了寻找方程 (11.69) 其它形式的解，我们利用球坐标中的关系式

$$\nabla^2 = \frac{1}{r}\partial_r^2 r - \frac{\mathbf{L}^2}{\hbar^2r^2}, \tag{11.71}$$

式中，$r=|\mathbf{x}|$ 为离散射中心的距离，$\mathbf{L}$ 为角动量算符。下面我们仅仅关心 $r\to+\infty$ 时方程 (11.69) 在球坐标中的解。由 (11.71) 式可知，当 $r\to+\infty$ 时，$\frac{\mathbf{L}^2}{\hbar^2r^2}$ 可以忽略，因此方程 (11.69) 渐近为

$$r\to+\infty:\ (\frac{1}{r}\partial_r^2 r+k_E^2)\phi_E(\mathbf{x})=0. \tag{11.72}$$

这个方程的解为

$$r\to+\infty:\ \phi_E(\mathbf{x})\sim g(\theta,\varphi)\frac{e^{\pm ik_Er}}{r}. \tag{11.73}$$

很明显，这两个解分别描写的是外行和内行的球面波，式中 $g(\theta,\varphi)$ 是一个任意的角度部分波函数。

11.2.1 直观理论

考虑到固定势场 $V(\mathbf{x})$, 则完整的定态薛定谔方程就应该是

$$\left(-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})\right)\psi_E(\mathbf{x})=E\psi_E(\mathbf{x}), \tag{11.74}$$

式中 $\psi_E(\mathbf{x})$ 为定态波函数。不过，由于我们是在研究散射问题，所以我们关心的其实是这个定态薛定谔方程的散射定态解。这些散射定态解除了要满足这个定态薛定谔方程之外，还需要满足一些额外的限制。

为了找到散射定态所需满足的额外限制，我们来考察方程 (11.74) 在 $r=|\mathbf{x}|\to+\infty$ 时的渐近行为。根据我们的假设，这时候势场 $V(\mathbf{x})$ 完全可以忽略，被散射的粒子将成为一个自由粒子，从而由自由粒子定态薛定谔方程 (11.69) 所描述。这告诉我们两点：第一，对于散射定态，能量 E 可以定义为粒子在无穷远处的动能，因此必定可以取成，$E=\frac{\hbar^2k_E^2}{2m}(k_E>0)$。第二，在 $r\to+\infty$ 的无穷远处，定态波函数 $\psi_E(\mathbf{x})$ 必定趋近于某个自由粒子定态波函数 $\phi_E(\mathbf{x})$。

直观上，当 $r\to+\infty$ 时，定态波函数 $\psi_E(\mathbf{x})$ 应该包含两部分，其中一部分描写入射自由粒子态 (注意入射粒子是从 $z\to-\infty$ 的无穷远处入射的)，另外一部分描写散射以后的出射粒子态。由于粒子是沿着 z 轴水平入射，所以入射自由粒子态显然可以取成平面波 e^{ik_Ez}。但是，经过势场散射以后，粒子将沿着四面八方出射，所以在 $r\to+\infty$ 时，出射粒子态应该由一个外行的球面波来描写。这也就是说，在 $r\to+\infty$ 时，散射定态 $\psi_E(\mathbf{x})$ 应该具有如下渐近行为

$$r\to+\infty:\ \psi_E(\mathbf{x})\sim e^{ik_Ez}+f(\theta,\varphi)\frac{e^{ik_Er}}{r}. \tag{11.75}$$

其中函数 $f(\theta,\varphi)$(有时候也称为散射振幅) 的具体形式要通过求解定态薛定谔方程 (11.74) 才能得到。

为了求出散射定态解 $\psi_E(\mathbf{x})$ 如何决定微分散射截面 $d\sigma$，我们可以计算入射波 e^{ik_Ez} 的概率流密度，结果为 $J_i=\frac{\hbar k_E}{m}$，因此单位时间单位横截面积上的入射粒子通量 $F_i=\frac{\hbar k_E}{m}N$。然后我们再计算单位时间之内探测器在无穷远处探测到的粒子数 dN，很显然，它等于总粒子数 N 与出射波概率流密度 J_r 在面积元 $r^2d\Omega$ 上的通量的乘积，即 $dN=NJ_rr^2d\Omega$。对于出射波 $f(\theta,\varphi)\frac{e^{ik_Er}}{r}$，我们很容易计算出它沿着径向的概率流密度 J_r，结果为 $J_r=\frac{\hbar k_E}{m}\frac{|f(\theta,\varphi)|^2}{r^2}$。由此我们可以得到

$$d\sigma=dN/F_i=|f(\theta,\varphi)|^2d\Omega. \tag{11.76}$$

可见，微分散射截面完全由散射定态解 $\psi_E(\mathbf{x})$ 在 $r\to+\infty$ 时的渐近形式决定。

11.2.2 与散射的一般理论的联系

上一小节我们给出的是在三维势散射问题中散射定态的直观理论。但是我们还有一个关于散射定态的一般性理论。那么上面的直观理论和一般性形式化理论给出来的结果是否一致呢？这就是这一小节将要研究的内容，通过这个研究，我们将会看到，三维势散射问题同样是一般性散射理论的一个特例。

为了看清楚理论的一般结构，现在我们将三维势散射的散射定态 $\psi_E(\mathbf{x})$ 重记为 $|\psi_{\mathbf{k}_i}\rangle$，式中 $\mathbf{k}_i$ 为入射波的波矢量 (它沿着水平的 z 轴)，因此 $|\mathbf{k}_i|=k_E$。根据散射定态的一般理论，$|\psi_{\mathbf{k}_i}\rangle$ 将满足方程

$$|\psi_{\mathbf{k}_i}\rangle=|\mathbf{k}_i\rangle+\frac{1}{E-H_0+i\varepsilon}T|\mathbf{k}_i\rangle. \tag{11.77}$$

式中 T 为散射理论的 T 算符，$|\mathbf{k}\rangle$ 为自由粒子平面波态，我们取

$$\langle\mathbf{x}|\mathbf{k}\rangle=e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{11.78}$$

因此它们满足归一化条件

$$\langle \mathbf{k}|\mathbf{k}'\rangle = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}'). \tag{11.79}$$

与这个归一化相配套的，对 $\mathbf{k}$ 态的求和现在也应该是积分 $\int \frac{d^3\mathbf{k}}{(2\pi)^3}\cdots$。

根据 (11.77) 式，我们有 $\langle \mathbf{x}|\psi_{\mathbf{k}_i}\rangle = \langle \mathbf{x}|\mathbf{k}_i\rangle + \int d^3\mathbf{x}'\langle \mathbf{x}|\frac{1}{E-H_0+i\varepsilon}|\mathbf{x}'\rangle\langle \mathbf{x}'|T|\mathbf{k}_i\rangle$。为了进一步计算这个散射定态波函数，我们首先需要计算出 $\langle \mathbf{x}|\frac{1}{E-H_0+i\varepsilon}|\mathbf{x}'\rangle$，具体的推导过程和一维情形完全类似

$$\begin{aligned}
&\langle \mathbf{x}|\frac{1}{E-H_0+i\varepsilon}|\mathbf{x}'\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\langle \mathbf{x}|\frac{1}{E-H_0+i\varepsilon}|\mathbf{k}\rangle\langle \mathbf{k}|\mathbf{x}'\rangle \\
=&\int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{E-\frac{\hbar^2\mathbf{k}^2}{2m}+i\varepsilon}\langle \mathbf{x}|\mathbf{k}\rangle\langle \mathbf{k}|\mathbf{x}'\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}}{\frac{\hbar^2k_E^2}{2m}-\frac{\hbar^2\mathbf{k}^2}{2m}+i\varepsilon} \\
=&\int_0^{+\infty}\frac{k^2dk}{(2\pi)^3}\int_0^{\pi}\sin\theta d\theta\int_0^{2\pi}d\varphi\frac{e^{ik|\mathbf{x}-\mathbf{x}'|\cos\theta}}{\frac{\hbar^2k_E^2}{2m}-\frac{\hbar^2k^2}{2m}+i\varepsilon} \\
=&\frac{m}{2\pi\hbar^2}\frac{1}{|\mathbf{x}-\mathbf{x}'|}\frac{i}{\pi}\left(\int_0^{+\infty}kdk\frac{e^{ik|\mathbf{x}-\mathbf{x}'|}-e^{-ik|\mathbf{x}-\mathbf{x}'|}}{k^2-k_E^2-i\varepsilon}\right) \\
=&\frac{m}{2\pi\hbar^2}\frac{1}{|\mathbf{x}-\mathbf{x}'|}\frac{i}{\pi}\left(\int_{-\infty}^{+\infty}kdk\frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{k^2-k_E^2-i\varepsilon}\right).
\end{aligned} \tag{11.80}$$

最后这个对 k 的积分和一维情形中表达式 (11.19) 最后的积分类似，在 k 复平面上同样有两个极点，$k_E+i\varepsilon$ 以及 $-k_E-i\varepsilon$。但是，现在由于指数因子 $e^{ik|\mathbf{x}-\mathbf{x}'|}$ 只在 k 的上半复平面收敛，并在上半复平面无穷远处的半圆上衰减为 0，所以现在的积分围道只有一种选择，就是选择由实轴和上半复平面无穷远处的半圆所构成的围道，这个围道只包含极点 $k_E+i\varepsilon$，由留数定理就可以得到

$$\langle \mathbf{x}|\frac{1}{E-H_0+i\varepsilon}|\mathbf{x}'\rangle = -\frac{m}{2\pi\hbar^2}\frac{e^{ik_E|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}. \tag{11.81}$$

所以散射定态 $\langle \mathbf{x}|\psi_{\mathbf{k}_i}\rangle$ 可以由下面的方程给出

$$\langle \mathbf{x}|\psi_{\mathbf{k}_i}\rangle = \langle \mathbf{x}|\mathbf{k}_i\rangle - \frac{m}{2\pi\hbar^2}\int d^3\mathbf{x}'\frac{e^{ik_E|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}\langle \mathbf{x}'|T|\mathbf{k}_i\rangle. \tag{11.82}$$

下面我们来考察散射定态 $\langle \mathbf{x}|\psi_{\mathbf{k}_i}\rangle$ 在 $r=|\mathbf{x}|\to+\infty$ 时的渐近行为。如图 (11.9) 所示，这时候

$$|\mathbf{x}-\mathbf{x}'| \approx r - \mathbf{e}_r\cdot\mathbf{x}', \tag{11.83}$$

式中 $\mathbf{e}_r = \mathbf{x}/|\mathbf{x}|$ 为径向单位矢量。因此我们有

$$\frac{e^{ik_E|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|} \approx \frac{1}{r}e^{ik_Er}e^{-i\mathbf{k}_f\cdot\mathbf{x}'} = \frac{e^{ik_Er}}{r}\langle \mathbf{k}_f|\mathbf{x}'\rangle, \tag{11.84}$$

式中

$$\mathbf{k}_f = k_E\mathbf{e}_r, \tag{11.85}$$

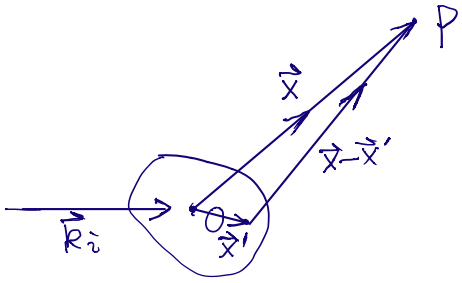

Figure 11.9: $r = |\mathbf{x}| \to +\infty$ 时的渐近行为。

它可以理解为出射粒子的波矢量。因此我们可以得到 (11.82) 式在 $r \to +\infty$ 时的渐近形式

$$\begin{aligned}\langle \mathbf{x}|\psi_{\mathbf{k}_i}\rangle &\sim \langle \mathbf{x}|\mathbf{k}_i\rangle - \frac{m}{2\pi\hbar^2}\frac{e^{ik_E r}}{r}\int d^3\mathbf{x}'\langle \mathbf{k}_f|\mathbf{x}'\rangle\langle \mathbf{x}'|T|\mathbf{k}_i\rangle \\ &= e^{ik_E z} - \frac{e^{ik_E r}}{r}\frac{m}{2\pi\hbar^2}\langle \mathbf{k}_f|T|\mathbf{k}_i\rangle. \end{aligned} \tag{11.86}$$

将这个结果与上一小节直观理论中的结果 (11.75) 进行比较，我们就可以得到，只要令

$$f(\theta,\varphi) = -\frac{m}{2\pi\hbar^2}\langle \mathbf{k}_f|T|\mathbf{k}_i\rangle, \tag{11.87}$$

那么散射定态的直观理论给出来的结果与散射一般性理论推导出来的结果就是一致的。正因为散射振幅 $f(\theta,\varphi)$ 具有 (11.87) 的形式，我们也常常将之记为 $f(\mathbf{k}_f,\mathbf{k}_i)$。

到此为止我们就建立起了三维势散射的直观理论与散射一般性理论之间的联系。不仅如此，将上面的 (11.87) 式代入微分截面公式 (11.76)，我们还可以得到 $d\sigma = \left(\frac{m}{2\pi\hbar^2}\right)^2 |\langle \mathbf{k}_f|T|\mathbf{k}_i\rangle|^2 d\Omega$。这个结果与我们在上一章散射的一般理论中给出的截面公式完全吻合。这已经足以说明，三维势散射问题完全可以看成是散射的一般理论的一个特例。

11.2.3 时间反演对称性和空间反演对称性

不妨让我们讨论一下时间反演对称性和空间反演对称性将会如何限制散射振幅。首先讨论时间反演对称性，如果系统具有这种对称性，则时间反演算符 Θ 将与 H_0 以及相互作用 V 均对易。由于 Θ 是一个反线性算符，所以必有 $\Theta T(E^+)\Theta^{-1} = T(E^-) = [T(E^+)]^\dagger$，式中 $E^- = E - i\varepsilon$，而 $T(E^+)$ 就是标准的 T 算符，我们一般简记为 T。也即是说，对于具有时间反演对称性的系统，我们必有

$$\Theta T \Theta^{-1} = T^\dagger. \tag{11.88}$$

记 $|\phi\rangle = |\mathbf{k}'\rangle$, $|\psi\rangle = T|\mathbf{k}\rangle$, 则利用时间反演算符的性质我们有

$$\langle \Theta\psi|\Theta\phi\rangle = \langle \phi|\psi\rangle = \langle \mathbf{k}'|T|\mathbf{k}\rangle. \tag{11.89}$$

另一方面，由于在时间反演之下动量会反向，即 $\Theta|\mathbf{k}\rangle = |-\mathbf{k}\rangle$，所以我们也有

$$|\Theta\psi\rangle = \Theta T\Theta^{-1}\Theta|\mathbf{k}\rangle = T^\dagger|-\mathbf{k}\rangle. \tag{11.90}$$

当然也有 $|\Theta\phi\rangle = \Theta|\mathbf{k}'\rangle = |-\mathbf{k}'\rangle$。将这些代入 (11.89) 式，即有

$$\langle -\mathbf{k}|T|-\mathbf{k}'\rangle = \langle \mathbf{k}'|T|\mathbf{k}\rangle. \tag{11.91}$$

也即是说，对于具有时间反演对称性的系统，必有 $f(\mathbf{k}',\mathbf{k}) = f(-\mathbf{k},-\mathbf{k}')$。在前面的一维散射问题中，我们也得到过类似的结果。

假如系统除了具有时间反演对称性之外，同时还具有空间反演对称性，则由于动量会在空间反演之下反向，我们将有

$$\langle -\mathbf{k}|T|-\mathbf{k}'\rangle = \langle \mathbf{k}|T|\mathbf{k}'\rangle. \tag{11.92}$$

结合前面的 (11.91) 式，即有

$$f(\mathbf{k}',\mathbf{k}) = f(\mathbf{k},\mathbf{k}'). \tag{11.93}$$

由于散射振幅的模方是微分散射截面，即 $\frac{d\sigma}{d\Omega}(\mathbf{k}\to\mathbf{k}') = |f(\mathbf{k}',\mathbf{k})|^2$，所以这就告诉我们，对于同时具有时间反演对称性和空间反演对称性的系统，我们有

$$\frac{d\sigma}{d\Omega}(\mathbf{k}\to\mathbf{k}') = \frac{d\sigma}{d\Omega}(\mathbf{k}'\to\mathbf{k}). \tag{11.94}$$

这有时候也称为细致平衡方程。

11.2.4 全同粒子散射

前面我们处理的是单个入射粒子在一个固定势场中的散射问题，但是我们也说过，同样的处理也适用于两个粒子“相互碰撞”的问题。这时候，我们需要选取两粒子的质心参考系，然后研究一个粒子相对于另一个粒子的运动，如果我们设粒子 1 的坐标为 $\mathbf{x}_1$，粒子 2 的坐标为 $\mathbf{x}_2$，那这种相对运动就由相对坐标 $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$ 来刻画，这时候散射势就是两粒子之间的相互作用势，相对运动的波函数将依然具有如下渐近形式

$$e^{i\mathbf{k}\cdot\mathbf{x}} + f(\theta,\varphi)\frac{e^{ikr}}{r}. \tag{11.95}$$

但是，如果这两个粒子是两个无自旋的全同玻色子，那这时候如图 (11.10) 所示，波函数在 D 处的渐近形式将由图中两种情形共同贡献。注意到在球坐标中 $\mathbf{x}\to-\mathbf{x}$ 相应于

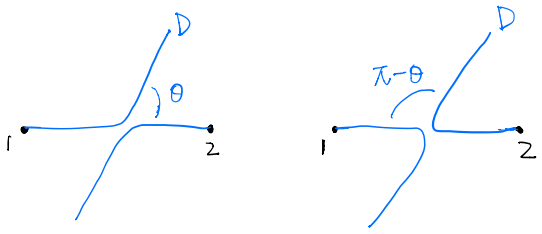

Figure 11.10: 全同粒子散射

$(\theta,\varphi)\to(\pi-\theta,\pi+\varphi)$，因此这时候波函数的渐近形式将为

$$e^{i\mathbf{k}\cdot\mathbf{x}} + e^{-i\mathbf{k}\cdot\mathbf{x}} + \big[f(\theta,\varphi) + f(\pi-\theta,\pi+\varphi)\big]\frac{e^{ikr}}{r}. \tag{11.96}$$

因为只有这样的波函数才是一个关于 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 的对称波函数。因此最终的微分散射截面就应该是

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= |f(\theta,\varphi)+f(\pi-\theta,\pi+\varphi)|^2 \\ &= |f(\theta,\varphi)|^2+|f(\pi-\theta,\pi+\varphi)|^2+2\mathrm{Re}\big[f(\theta,\varphi)f^*(\pi-\theta,\pi+\varphi)\big].\end{aligned} \tag{11.97}$$

很显然，如果 $f(\theta,\varphi)=f(\theta)$ 而与 φ 无关 (比如相互作用势为中心势)，则 $\theta,=\pi/2$ 处是干涉加强的。

但是，如果这两个粒子是两个自旋 $1/2$ 的全同费米子，并且假设它们之间的相互作用与自旋无关。那这时候总的波函数就应该是一个反对称波函数，而这又有两种可能性：第一，这两个粒子处于自旋单态，即自旋波函数反对称，那这时候空间波函数就应该对称，这和上面的情形类似。第二，这两个粒子处于自旋三重态，即自旋波函数对称，那这时候空间波函数就应该反对称了，这和上面的情形不同。为了具体起见，不妨进一步假设入射的这两个粒子是非极化的，它们以相等的概率处于各种可能的自旋本征态，即有 $1/4$ 的概率处于自旋单态，$3/4$ 的概率处于自旋三重态，那这时候的微分散射截面就应该是

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \frac{1}{4}|f(\theta,\varphi)+f(\pi-\theta,\pi+\varphi)|^2+\frac{3}{4}|f(\theta,\varphi)-f(\pi-\theta,\pi+\varphi)|^2 \\ &= |f(\theta,\varphi)|^2+|f(\pi-\theta,\pi+\varphi)|^2-\mathrm{Re}\big[f(\theta,\varphi)f^*(\pi-\theta,\pi+\varphi)\big].\end{aligned} \tag{11.98}$$

很显然，如果 $f(\theta,\varphi)=f(\theta)$ 与 φ 无关，那这时候 $\theta=\pi/2$ 处是干涉减弱的。

11.2.5 卢瑟福公式的非微扰精确求解

回到固定势场散射情形。通过前面的讨论我们看到，如果能求出 H 的散射定态解，那么从散射定态的渐近形式中我们就能读出散射振幅的信息，进而就可以算出散射截面。当然一般来说散射定态的精确解往往都得不到。但是，有一些特殊的散射势是存在精确解的，比方说，库伦势的散射定态就可以精确求解。这一小节我们主要就是要讨论如何从库伦势的精确散射定态中得到卢瑟福截面公式。

在上一章中我们已经用玻恩近似得到了卢瑟福截面公式。但是玻恩近似是一个微扰论公式，而且只计算到了微扰级数的第一阶，因此如果考虑到高阶微扰修正，那之前我们得到的卢瑟福公式就有可能要作修正。但是，通过这一小节的精确求解我们将看到，由于库伦势的特殊性 (它除了有中心力场的旋转对称性以外，还有更多的隐藏对称性)，精确的卢瑟福截面公式实际上和我们前面用玻恩近似得出来的结果完全一样，高阶微扰修正只是改变了散射振幅的相因子，而这个相因子在散射截面中并不会反映出来 (因为散射截面正比于散射幅的模方)。当然，正如我们将要看到的，散射振幅的相因子也包含了重要的信息，比方说，它里面隐藏了库伦吸引势的束缚态信息，而这些信息在微扰论的计算下是无法得到的。

库伦势本身也有一点额外的复杂性，那就是它在无穷远处衰减得不够快，因此实际上，它的散射定态解并没有 (11.75) 式那样的渐近形式。这一点我们会在库伦势散射定态精确解的渐近形式中清楚地看到。也就是说，本节前几个小节所发展的理论对于库伦势情形严

格来说是不成立的。好在，前几个小节的大致框架依然是成立的，只需要在细节上作一些修正。最终我们依然可以从散射定态的渐近形式中得到散射振幅，进而可以推导出卢瑟福截面公式。

我们可以将库伦势写成

$$V(r)=\frac{V_0}{r}, \tag{11.99}$$

式中 V_0 表示库伦势的强度，$V_0>0$ 为排斥势，$V_0<0$ 为吸引势，对于两个原子核的散射，$V_0=\frac{Z_1Z_2e^2}{4\pi\varepsilon_0}=Z_1Z_2e_s^2$，对于氢原子和类氢离子 $V_0=-\frac{Ze^2}{4\pi\varepsilon_0}=-Z\hbar\alpha c$($\alpha$ 为精细结构常数)。则我们需要求解的定态薛定谔方程即是

$$\left(-\frac{\hbar^2}{2m}\nabla^2+\frac{V_0}{r}\right)\psi_E(\mathbf{x})=E\psi_E(\mathbf{x}). \tag{11.100}$$

当然，我们关心的是散射定态，因此入射粒子能量 $E=\frac{\hbar^2k_E^2}{2m}$，从而我们可以将上面的定态薛定谔方程重写为

$$\left(\nabla^2+k_E^2-\frac{2\gamma_Ek_E}{r}\right)\psi_E(\mathbf{x})=0. \tag{11.101}$$

式中 $\gamma_E=mV_0/(\hbar^2k_E)$ 是一个无量纲参数，它依赖于入射能量 E。

假设我们记相应的入射粒子波矢量为 $\mathbf{k}$，和前文一样，$\mathbf{k}$ 沿着水平的 z 轴，大小等于 k_E，因此 $\mathbf{k}=k_E\mathbf{e}_z$ (式中 $\mathbf{e}_z$ 为 z 轴方向的单位矢量)，则相应的散射定态解是

$$\psi_E(\mathbf{x})=e^{i\mathbf{k}\cdot\mathbf{x}}e^{-\pi\gamma_E/2}\Gamma(1+i\gamma_E)\,{}_1F_1\big(-i\gamma_E;1;i(k_Er-\mathbf{k}\cdot\mathbf{x})\big). \tag{11.102}$$

其中 ${}_1F_1(a;b;w)$ 是合流超几何函数，它可以定义为下面的无穷级数

$${}_1F_1(a;b;w)=1+\sum_{n=1}^{+\infty}\frac{(a)_n}{(b)_n}\frac{w^n}{n!}, \tag{11.103}$$

式中 $(a)_n=a(a+1)...(a+n-1)$，叫做 Pochhammer 符号。当然，由于 $\mathbf{k}$ 沿着 z 轴，所以实际上式中的 $\mathbf{k}\cdot\mathbf{x}=k_Ez$，$k_Er-\mathbf{k}\cdot\mathbf{x}=k_E(r-z)$。我们只是直接给出这个精确解，至于具体是如何找到它的，这不是我们这里要关注的重点，读者可以参阅朗道《量子力学》第十七章第 135 节的相关推导。

下面我们来考察这个解 (11.102) 在 $r\to+\infty$ 时的渐近形式。对于散射问题来说，重要的是考察 $\mathbf{x}$ 不在水平 z 轴上的情形，这时候 $r\to+\infty$ 就相当于 $r-z\to+\infty$。因此我们需要用到合流超几何函数 ${}_1F_1(a;b;w)$ 在 $|w|\to+\infty$ 时的渐近展开 (参见朗道《量子力学》的数学附录)，它是

$$\begin{aligned}{}_1F_1(a;b;w)=&\,e^{i\pi a}\frac{\Gamma(b)}{\Gamma(b-a)}w^{-a}\left(1+\sum_{n=1}^{+\infty}(-1)^n\frac{(a)_n(a-b+1)_n}{n!}w^{-n}\right)\\&+e^ww^{a-b}\frac{\Gamma(b)}{\Gamma(a)}\left(1+\sum_{n=1}^{+\infty}\frac{(b-a)_n(1-a)_n}{n!}w^{-n}\right).\end{aligned} \tag{11.104}$$

由此我们可以得到散射定态解 (11.102) 在 $r-z\to+\infty$ 时的渐近形式

$$\begin{aligned}\psi_E(\mathbf{x}) &\sim e^{ik_Ez}[k_E(r-z)]^{i\gamma_E}-\frac{\gamma_E}{k_E(r-z)}\frac{\Gamma(1+i\gamma_E)}{\Gamma(1-i\gamma_E)}e^{ik_Er}[k_E(r-z)]^{-i\gamma_E}+....\\ &=e^{ik_Ez+i\gamma_E\log(k_E(r-z))}-\frac{\gamma_E}{k_E(r-z)}\frac{\Gamma(1+i\gamma_E)}{\Gamma(1-i\gamma_E)}e^{ik_Er-i\gamma_E\log(k_E(r-z))}+....\\ &=e^{ik_Ez+i\gamma_E\log(k_E(r-z))}+f(\theta)\frac{e^{ik_Er-i\gamma_E\log(k_E(r-z))}}{r}+...\end{aligned}\tag{11.105}$$

式中

$$f(\theta)=-\frac{\gamma_E}{k_E}\frac{\Gamma(1+i\gamma_E)}{\Gamma(1-i\gamma_E)}\frac{r}{r-z}=-\frac{\gamma_E}{2k_E}\frac{\Gamma(1+i\gamma_E)}{\Gamma(1-i\gamma_E)}\frac{1}{\sin^2(\theta/2)},\tag{11.106}$$

式中 θ 是 $\mathbf{x}$ 与水平 z 轴的夹角，也就是散射角。

从 (11.105) 的最终结果来看，库伦势情形的散射定态解并没有 (11.75) 式那样标准的渐近形式,但是很像,只需要我们修正一下入射波和出射球面波的相位,将 $e^{ik_Ez+i\gamma_E\log(k_E(r-z))}$ 当作入射波，将 $e^{ik_Er-i\gamma_E\log(k_E(r-z))}/r$ 当作出射球面波。因此 (11.106) 式给出的 $f(\theta)$ 就是精确的散射振幅表达式，它的模方将给出精确的微分散射截面，即

$$\frac{d\sigma}{d\Omega}=|f(\theta)|^2=\frac{V_0^2}{16E^2\sin^4(\theta/2)},\tag{11.107}$$

式中我们代入了 $\gamma_E=mV_0/(\hbar^2k_E)$，以及 $E=\hbar^2k_E^2/(2m)$，最终的这个微分散射截面就是著名的卢瑟福截面，不过，我们现在给出来的是精确结果，但是很显然它和玻恩近似给出来的结果是一样的。实际上，散射振幅 $f(\theta)$(根据 (11.87) 式，它在相差一个常数系数的意义上也就是 T 矩阵元) 的精确解 (11.106) 只是比玻恩近似的结果多出了一个相位 $e^{i\delta(E)}$

$$e^{i\delta(E)}=\frac{\Gamma(1+i\gamma_E)}{\Gamma(1-i\gamma_E)},\tag{11.108}$$

这个相位对散射截面没有贡献，但是，它其实包含了很重要的信息。

为了看清楚相位 $e^{i\delta(E)}$ 中额外包含的信息。我们将入射粒子能量 E 解析延拓到 $E<0$ 的情形，相应的 $k_E=\sqrt{2mE/\hbar^2}$ 就解析延拓到正虚轴，进一步，如果我们的库伦势是一个吸引势，即 $V_0<0$, 则这时候 $\gamma_E\propto V_0/k_E$ 也会解析延拓到正虚轴，不妨记 $\gamma_E=i\nu_E(\nu_E>0)$。当然，这时候由 (11.108) 式给出来的 $e^{i\delta(E)}$ 就不再是一个相位了，而会变成

$$\frac{\Gamma(1-\nu_E)}{\Gamma(1+\nu_E)}.\tag{11.109}$$

伽马函数 $\Gamma(w)$ 在复平面上并没有 0 点，但是有极点，$w=0,-1,-2,-3,...$ 就是伽马函数的极点。这意味着当 ν_E 满足下式时，会对应于解析延拓以后的 T 矩阵元在能量复平面上的极点，

$$1-\nu_E=-(n-1),\quad n=1,2,3,....\tag{11.110}$$

显然，这时候 $\nu_E=n$, 即 $\gamma_E=in$, 即 $k_E=-imV_0/(\hbar^2n)$(注意，$V_0<0$)，即

$$E=\frac{\hbar^2k_E^2}{2m}=-\frac{mV_0^2}{2\hbar^2}\frac{1}{n^2}.\tag{11.111}$$

如果代入 $V_0=-Z\hbar\alpha c$, 那这个结果给出来的就是著名的氢原子能级公式。从这个例子我们再一次清楚地看到，T 矩阵元在能量负实轴上的极点对应于系统的束缚态能级。

但是，如果是库伦排斥势情形，也即是 $V_0>0$ 的情形，那当我们将能量解析延拓到 $E<0$ 时，k_E 还是会解析延拓到正虚轴，但这时候 $\gamma_E\propto V_0/k_E$ 将会解析延拓到负虚轴，不妨记 $\gamma_E=-i\nu_E(\nu_E>0)$。那这时候由 (11.108) 式给出来的 $e^{i\delta(E)}$ 就会解析延拓成

$$\frac{\Gamma(1+\nu_E)}{\Gamma(1-\nu_E)}. \tag{11.112}$$

由于 $\nu_E>0$，因此这个表达式并没有极点。也就是说，这时候 T 矩阵元在能量负实轴上并没有极点，因此系统当然也就没有束缚态。这非常符合我们的直观，因为库伦排斥势的势能大于 0，当然是不可能有束缚态存在的。

11.3 中心势场的分波法

对于三维势散射问题，通过前面的论述我们已经知道，散射的 T 矩阵元可以从散射定态解的渐近形式中读出来。当然，通常我们并不能得到散射定态的精确解，然而如果系统有一些对称性的话，那么我们将可以进行进一步的非微扰分析。下面我们将集中于系统有三维空间旋转不变性的情形，具体来说就是，我们将考察球对称的中心势 $V(r)$ 的情形。

11.3.1 对称性分析

对于球对称的中心力场，旋转不变性告诉我们 $[\mathbf{L},V(r)]=0$, 式中 $\mathbf{L}$ 为粒子的角动量算符。另外，自由粒子哈密顿量 H_0 显然是旋转不变的，从而我们有

$$[\mathbf{L},H_0]=[\mathbf{L},H]=0. \tag{11.113}$$

又由于时间演化算符 $U_I(t_2,t_1)=e^{i\frac{H_0}{\hbar}t_2}e^{-i\frac{H}{\hbar}(t_2-t_1)}e^{-i\frac{H_0}{\hbar}t_1}$，因此 $[\mathbf{L},U_I(t_2,t_1)]=0$，特别的，$\mathbf{L}$ 与 $S=U_I(+\infty,-\infty)$ 对易，即

$$[\mathbf{L},S]=0. \tag{11.114}$$

这告诉我们，S 算符是一个标量算符。

前面在第 (11.2.2) 小节中，我们的分析都是在平面波矢量基 $|\mathbf{k}\rangle$ 下进行的。但是对于旋转不变的中心力场, 引入球对称的矢量基 $|E,l,m\rangle$ 将是非常有帮助的。按照定义，$|E,l,m\rangle$ 是 H_0、$\mathbf{L}^2$、以及 L_z 共同的本征态，本征值分别为 E、$l(l+1)\hbar^2$、以及 $m\hbar$。这里的 m 表示磁量子数，而不是散射粒子的质量，这两者我们用了相同的符号，希望读者可以根据上下文作出区分。因此，这里的 $|E,l,m\rangle$ 就相当于散射的一般理论中用抽象的指标表示的 $|\alpha\rangle$，它们的归一化如下，

$$\langle E',l',m'|E,l,m\rangle=\delta(E-E')\delta_{l'l}\delta_{m'm}, \tag{11.115}$$

(在散射的一般理论中是抽象地写作 $\langle\beta|\alpha\rangle=\delta(\beta-\alpha)$)。

我们来考察在这个球对称表象 $\{|E,l,m\rangle\}$ 下的 S 矩阵元 $\langle E',l',m'|S|E,l,m\rangle$。由于散射过程的能量守恒，可以知道 $\langle E',l',m'|S|E,l,m\rangle \propto \delta(E'-E)$。进一步我们可以利用旋转不变性，由 $[\mathbf{L}^2,S]=0$，有

$$\begin{aligned} 0 &= \langle E',l',m'|[\mathbf{L}^2,S]|E,l,m\rangle \\ &= \langle E',l',m'|\mathbf{L}^2 S|E,l,m\rangle - \langle E',l',m'|S\mathbf{L}^2|E,l,m\rangle \\ &= [l'(l'+1)-l(l+1)]\hbar^2\langle E',l',m'|S|E,l,m\rangle, \end{aligned} \tag{11.116}$$

即 $\langle E',l',m'|S|E,l,m\rangle \propto \delta_{l'l}$。而由 $[L_z,S]=0$，我们有

$$\begin{aligned} 0 &= \langle E',l,m'|[L_z,S]|E,l,m\rangle \\ &= \langle E',l,m'|L_z S|E,l,m\rangle - \langle E',l,m'|SL_z|E,l,m\rangle \\ &= (m'-m)\hbar\langle E',l,m'|S|E,l,m\rangle, \end{aligned} \tag{11.117}$$

即 $\langle E',l,m'|S|E,l,m\rangle \propto \delta_{m'm}$。又由 $[L_+,S]=0$，我们有

$$\begin{aligned} 0 &= \langle E',l,m|[L_+,S]|E,l,m-1\rangle \\ &\propto \big(\langle E',l,m-1|S|E,l,m-1\rangle - \langle E',l,m|S|E,l,m\rangle\big), \end{aligned} \tag{11.118}$$

(式中第二个等号我们利用了 $L_-|l,m\rangle = \hbar\sqrt{l(l+1)-(m-1)m}|l,m-1\rangle$ 的共轭转置 $\langle l,m|L_+ = \hbar\sqrt{l(l+1)-(m-1)m}\langle l,m-1|$), 即

$$\langle E',l,m-1|S|E,l,m-1\rangle = \langle E',l,m|S|E,l,m\rangle. \tag{11.119}$$

将上一段的所有对称性分析结果概括起来，即有

$$\langle E',l',m'|S|E,l,m\rangle = \delta(E'-E)S_l(E)\delta_{l'l}\delta_{m'm}. \tag{11.120}$$

实际上，如果引用第七章中讨论过的 Wigner-Eckart 定理的话，那上一段的推导完全可以省略，我们可以直接写出这个结果。也就是说, 在 $\{|E,l,m\rangle\}$ 表象中，S 矩阵变成了一个对角矩阵，$S_l(E)$ 是它的对角元，因此 S 矩阵的幺正性就变成

$$|S_l(E)|^2 = 1. \tag{11.121}$$

由此可以将 $S_l(E)$ 写成

$$S_l(E) = e^{2i\delta_l(E)}, \tag{11.122}$$

式中的 $\delta_l(E)$ 称作相移。

类似的可以对 T 矩阵进行分析，有 $\langle E,l',m'|T|E,l,m\rangle = T_l(E)\delta_{l'l}\delta_{m'm}$。散射的一般理论告诉我们，$S$ 矩阵和 T 矩阵之间有如下关系

$$\langle E',l',m'|S|E,l,m\rangle = \delta(E'-E)\delta_{l'l}\delta_{m'm} - 2\pi i\delta(E'-E)T_l(E)\delta_{l'l}\delta_{m'm}. \tag{11.123}$$

因此结合上一段的结果，我们有

$$S_l(E) = 1 - 2\pi i T_l(E), \tag{11.124}$$

这个表达式就决定了 T 矩阵与相移的关系。

11.3.2 分波法与散射振幅

从上一节的前两个小节我们知道，对于三维势散射问题，关键是求出散射振幅 $f(\theta,\varphi)$，或者等价的求出动量表象下的 T 矩阵 $\langle \mathbf{k}_f|T|\mathbf{k}_i\rangle$。但是，在本节上一小节中我们又看到，对于球对称的中心势场情形，分波法的表象 $\{|E,l,m\rangle\}$ 对于对称性分析来说更方便。这一节我们将要把这两个表象联系起来，即利用分波法表象下的 T 矩阵对角元 $T_l(E)$ 来表示出动量表象下的 T 矩阵元 $\langle \mathbf{k}_f|T|\mathbf{k}_i\rangle$。

首先我们要求出分波法表象和动量表象之间的变换矩阵 $\langle \mathbf{k}|E,l,m\rangle$，式中 $E=\hbar^2k_E^2/(2m_\mu)$（这里的 m_μ 表示散射粒子质量），$\mathbf{k}$ 为一个任意的波矢量。我们注意到，给定 $\mathbf{k}$，它的矢量方向 $\widehat{\mathbf{k}}$（$\widehat{\mathbf{k}}$ 为单位矢量）就在空间指定了一个确定方向，假设我们在波矢量空间取球坐标，那这个方向也即相应于单位球面上的一个点，不妨记作 $\widehat{\mathbf{k}}=(\theta,\varphi)$。而概率幅 $\langle \mathbf{k}|E,l,m\rangle$ 表示的就是，对于一个能量为 E，角动量为 $|l,m\rangle$ 的粒子，我们发现它的波数为 $k=|\mathbf{k}|$，同时发现它处在单位球面的 $\widehat{\mathbf{k}}$ 点的概率幅。我们知道，对于一个 $|l,m\rangle$ 态的粒子，我们在单位球面上某点 (θ,φ) 发现它的概率幅就是球谐函数 $Y_{lm}(\theta,\varphi)$（因为球谐函数就是具有确定角动量 $|l,m\rangle$ 的粒子在单位球面上的波函数）。由此我们知道 $\langle \mathbf{k}|E,l,m\rangle$ 必然正比于 $Y_{lm}(\widehat{\mathbf{k}})$。另外，由于 $|E,l,m\rangle$ 态具有确定的能量 E，因此可知除非 $E_{\mathbf{k}}=\hbar^2\mathbf{k}^2/(2m_\mu)=\hbar^2k^2/(2m_\mu)=E$，否则概率幅 $\langle \mathbf{k}|E,l,m\rangle$ 必然等于 0。概括这两点，我们就必然有

$$\langle \mathbf{k}|E,l,m\rangle = g_l(E)\delta(E_{\mathbf{k}}-E)Y_{lm}(\widehat{\mathbf{k}}). \tag{11.125}$$

为了进一步确定这个表达式中未定的 $g_l(E)$，我们可以利用正交归一关系 $\langle E',l',m'|E,l,m\rangle=\delta(E'-E)\delta_{l'l}\delta_{m'm}$。具体来说，$\langle E',l',m'|E,l,m\rangle=\int\frac{d^3\mathbf{k}}{(2\pi)^3}\langle E',l',m'|\mathbf{k}\rangle\langle \mathbf{k}|E,l,m\rangle=g_l^2(E)\int\frac{d^3\mathbf{k}}{(2\pi)^3}\delta(E_{\mathbf{k}}-E')\delta(E_{\mathbf{k}}-E)Y_{lm}(\widehat{\mathbf{k}})Y^*_{l'm'}(\widehat{\mathbf{k}})=g_l^2(E)\delta(E'-E)\frac{m_\mu}{\hbar^2k_E}\frac{k_E^2}{(2\pi)^3}\int d\Omega_{\widehat{\mathbf{k}}}Y_{lm}(\widehat{\mathbf{k}})Y^*_{l'm'}(\widehat{\mathbf{k}})=g_l^2(E)\frac{m_\mu k_E}{\hbar^2(2\pi)^3}\delta(E'-E)\delta_{l'l}\delta_{m'm}$，（式中第三个等号我们利用了 $\delta(E_{\mathbf{k}}-E)=\frac{m_\mu}{\hbar^2k_E}\delta(k-k_E)$，而最后一个等号则利用了球谐函数的正交归一性 $\int d\Omega_{\widehat{\mathbf{k}}}Y_{lm}(\widehat{\mathbf{k}})Y^*_{l'm'}(\widehat{\mathbf{k}})=\delta_{l'l}\delta_{m'm}$）。将这个结果与正交归一关系 $\langle E',l',m'|E,l,m\rangle=\delta(E'-E)\delta_{l'l}\delta_{m'm}$ 比较，我们有 $g_l(E)=(2\pi)^{3/2}\hbar/\sqrt{m_\mu k_E}$，即

$$\langle \mathbf{k}|E,l,m\rangle=(2\pi)^{3/2}\frac{\hbar}{\sqrt{m_\mu k_E}}\delta(E_{\mathbf{k}}-E)Y_{lm}(\widehat{\mathbf{k}}). \tag{11.126}$$

现在我们可以用 T 矩阵分波 $T_l(E)$ 来表示 $\langle \mathbf{k}_f|T|\mathbf{k}_i\rangle$ 了（注意，假设我们以 E 来表示入射粒子能量，则由于散射过程的能量守恒，所以 $E_{\mathbf{k}_f}=E_{\mathbf{k}_i}=E$，即 $|\mathbf{k}_f|=|\mathbf{k}_i|=k_E$）。

$$\begin{aligned}
\langle \mathbf{k}_f|T|\mathbf{k}_i\rangle &= \sum_{l',m',l,m}\int dE'dE''\langle \mathbf{k}_f|E',l',m'\rangle\langle E',l',m'|T|E'',l,m\rangle\langle E'',l,m|\mathbf{k}_i\rangle \\
&= (2\pi)^3\frac{\hbar^2}{m_\mu k_E}\sum_{l',m',l,m}Y_{l'm'}(\widehat{\mathbf{k}}_f)Y^*_{lm}(\widehat{\mathbf{k}}_i)\langle E,l',m'|T|E,l,m\rangle \\
&= (2\pi)^3\frac{\hbar^2}{m_\mu k_E}\sum_{l',m',l,m}Y_{l'm'}(\widehat{\mathbf{k}}_f)Y^*_{lm}(\widehat{\mathbf{k}}_i)T_l(E)\delta_{l'l}\delta_{m'm} \\
&= (2\pi)^3\frac{\hbar^2}{m_\mu k_E}\sum_{l,m}Y_{lm}(\widehat{\mathbf{k}}_f)Y^*_{lm}(\widehat{\mathbf{k}}_i)T_l(E) \\
&= (2\pi)^3\frac{\hbar^2}{m_\mu k_E}\sum_l T_l(E)\frac{2l+1}{4\pi}P_l(\cos\theta).
\end{aligned} \tag{11.127}$$

式中最后一行我们利用了球谐函数的加法定理

$$\sum_m Y_{lm}(\widehat{\mathbf{k}}_f)Y_{lm}^*(\widehat{\mathbf{k}}_i)=\frac{2l+1}{4\pi}P_l(\cos\theta). \tag{11.128}$$

式中 θ 为 $\widehat{\mathbf{k}}_f$ 与 $\widehat{\mathbf{k}}_i$ 之间的夹角，也就是散射角。

将上面的 T 矩阵表达式 (11.127) 代入公式 (11.87)(注意这两个式子的散射粒子质量用了不同的符号，但其实是同一个质量)，我们就有

$$\begin{aligned} f(\theta)&=-\frac{\pi}{k_E}\sum_l(2l+1)T_l(E)P_l(\cos\theta)\\ &=\frac{1}{2ik_E}\sum_l(2l+1)\big(S_l(E)-1\big)P_l(\cos\theta)\\ &=\frac{1}{k_E}\sum_l(2l+1)e^{i\delta_l(E)}\sin(\delta_l)P_l(\cos\theta). \end{aligned} \tag{11.129}$$

式中我们代入了方程 (11.124) 以及方程 (11.122)。

利用我们在散射的一般理论中推导出来的光学定理 $\frac{k_\alpha}{4\pi}\sigma_T=-\frac{m}{2\pi\hbar^2}\mathrm{Im}T_{\alpha\alpha}=\mathrm{Im}f(\theta)|_{\theta=0}$(这里的 $\theta=0$ 是因为对于朝前散射幅来说，$\mathbf{k}_f=\mathbf{k}_i$，即散射角为 0)，以及 $P_l(\cos 0)=P_l(1)=1$，我们可以得到总截面 $\sigma_T=\int d\sigma$ 的公式，

$$\sigma_T=\frac{4\pi}{k_E^2}\sum_l(2l+1)\sin^2(\delta_l). \tag{11.130}$$

11.3.3 如何确定相移

通过上一小节我们已经看到，对于中心势场，散射振幅完全由相移 $\delta_l(E)$ 决定。那么如何求解这些相移呢？回答是，通过求解分波形式的散射定态。这一小节我们将详细地阐明这一点。

首先，根据 (11.75) 我们已经知道，当 $r\to+\infty$ 时，散射定态波函数 $\psi_E(\mathbf{x})$ 有如下渐近形式

$$r\to+\infty:\ \psi_E(\mathbf{x})\sim e^{ik_Ez}+f(\theta)\frac{e^{ik_Er}}{r}. \tag{11.131}$$

下面我们要做的就是将这个渐近形式用分波表示出来。为此，我们首先注意到下面的恒等式 (证明参见朗道《量子力学》第五章第 34 节)

$$e^{ik_Ez}=e^{ik_Er\cos\theta}=\sum_l(2l+1)i^l j_l(k_Er)P_l(\cos\theta). \tag{11.132}$$

式中 $j_l(k_Er)$ 为球贝塞尔函数。其次，当 $r\to+\infty$ 时，球贝塞尔函数有如下渐近形式

$$j_l(k_Er)\sim\frac{e^{i(k_Er-\frac{l\pi}{2})}-e^{-i(k_Er-\frac{l\pi}{2})}}{2ik_Er}. \tag{11.133}$$

将这些结果代入 (11.131) 式，并代入散射振幅 $f(\theta)$ 的分波表达式 (11.129)，我们就有

$$r\to+\infty:\ \psi_E(\mathbf{x})\sim\sum_l(2l+1)P_l(\cos\theta)\frac{1}{2ik_Er}\left[S_l(E)e^{ik_Er}-e^{-i(k_Er-l\pi)}\right]. \tag{11.134}$$

由这个表达式可以看出，散射势的存在完全体现在将出射波前面的系数从 1 替换成了相因子 $S_l(E)$。这个表达式也告诉我们，我们只需要求解出分波形式下的散射定态解，然后从其渐近形式中就能读出相因子 $S_l(E)$ 的具体取值，从而就可以得到相移 $\delta_l(E)$。

公式 (11.134) 还告诉我们一个证明 $|S_l(E)|^2=1$ 的不同方法。首先从公式 (11.134) 我们可以知道，对于给定的分波 l，在半径为 r 的球面上，外行出射波的概率流密度正比于 $|S_l(E)|^2$，而内行入射波的概率流密度正比于 $|e^{il\pi}|^2=1$。概率守恒和角动量守恒告诉我们，对于每一个分波 l，球面上内行的概率通量必定等于外行的概率通量，从而必有 $|S_l(E)|^2=1$。前面我们是利用 S 矩阵的幺正性得到这个结果的，但在物理本质上，S 矩阵的幺正性本身反映的就是概率守恒。所以这两种不同的推导方式在物理本质上其实是一样的，只不过从 S 矩阵的幺正性得到 $|S_l(E)|^2=1$ 这一推导更加具有一般性。

我们还有另外一个更好用的方法可以表示结果 (11.134)。为此我们注意到，由于角动量守恒，所以算符 $H,\mathbf{L}^2,L_z$ 两两可对易，因此我们可以求解它们共同的本征态 $\psi_{E,l,m}(\mathbf{x})$，由于中心对称性，很显然 $\psi_{E,l,m}(\mathbf{x})$ 必定可以写成

$$\psi_{E,l,m}(\mathbf{x})=R_{E,l}(r)Y_{lm}(\theta,\varphi)=\frac{1}{r}u_{E,l}(r)Y_{lm}(\theta,\varphi). \tag{11.135}$$

其中 $u_{E,l}(r)$ 满足径向方程

$$\left[-\frac{\hbar^2}{2m_\mu}\frac{d^2}{dr^2}+\frac{l(l+1)\hbar^2}{2m_\mu r^2}+V(r)\right]u_{E,l}(r)=Eu_{E,l}(r). \tag{11.136}$$

当然我们关心的不是束缚态，而是 $E=\frac{\hbar^2k_E^2}{2m_\mu}>0$ 时的解。

当 $r\to+\infty$ 时，势能 $V(r)$ 已经衰减为 0，同样离心势 $\frac{l(l+1)\hbar^2}{2m_\mu r^2}$ 也已经衰减为 0，因此方程 (11.136) 渐近为

$$r\to+\infty:\ \left[\frac{d^2}{dr^2}+k_E^2\right]u_{E,l}(r)=0. \tag{11.137}$$

解这个渐近方程，从而就有

$$r\to+\infty:\ u_{E,l}(r)\sim A_l(E)e^{ik_Er}+B_l(E)e^{-ik_Er}. \tag{11.138}$$

同样，根据概率守恒，在半径为 r 的球面上内行的概率通量应该等于外行的概率通量，从而必有

$$|A_l(E)|^2=|B_l(E)|^2. \tag{11.139}$$

当然 $A_l(E)$ 和 $B_l(E)$ 的具体值要通过求解完整的径向方程 (11.136) 才能得出。

另一方面，散射定态 $\psi_E(\mathbf{x})$ 当然一定能写成 $\psi_{E,l,m}(\mathbf{x})$ 的线性组合形式。不仅如此，由于散射问题绕着水平的 z 轴旋转不变，所以散射定态必定与角度 φ 无关，因此这也就是说，散射定态必定能写成 $\psi_{E,l,0}(\mathbf{x})=Y_{l0}(\theta,\varphi)\propto\frac{1}{r}u_{E,l}(r)P_l(\cos\theta)$ 的线性组合。很显然，我们前面得到的散射定态的渐近形式 (11.134) 的确满足这一点。不仅如此，根据渐近形式 (11.134)，我们还可以知道，在相差一个整体归一化常数的意义上，必有

$$r\to+\infty:\ u_{E,l}(r)\sim S_l(E)e^{ik_Er}-e^{-i(k_Er-l\pi)}. \tag{11.140}$$

比较 $u_{E,l}(r)$ 的这两个渐近形式 (11.138) 和 (11.140)(两者可以相差一个整体常数) 中出射波与入射波的幅度之比, 我们可以得到

$$S_l(E) = (-)^{l+1}\frac{A_l(E)}{B_l(E)}. \tag{11.141}$$

当然，根据 (11.139) 式，结果必定是一个相因子。

以上就是我们最终求解 $S_l(E)$ 的方法。即我们首先求解径向方程 (11.136)，并得出解 $u_{E,l}(r)$ 在 $r\to+\infty$ 时的渐近形式 (11.138)，根据 (11.141)，从这个渐近形式中我们就能读出 $S_l(E)$。

11.3.4 束缚态和共振态

假设我们将能量 E 解析延拓到 $E<0$, 从而 k_E 就解析延拓到正虚轴，不妨记 $k_E = i\lambda_E$。从而渐近形式 (11.138) 就解析延拓为

$$r\to+\infty:\ u_{E,l}(r)\sim A_l(E)e^{-\lambda_E r}+B_l(E)e^{\lambda_E r}. \tag{11.142}$$

由于 $\lambda_E>0$，所以一般来说，这样的解当然是不可归一化的，从而也就不是物理的。除非相应的能量 E 刚好是函数 $B_l(E)$ 的零点，这时候 (11.142) 中不可归一化的指数增长项实际上是 0，相应的波函数在无穷远处实际上是指数衰减的，这当然就是一个束缚态。所以，系统的束缚态相应于函数 $B_l(E)$ 在能量复平面负实轴上的零点。根据 (11.141) 式，这也就是说，系统的束缚态相应于 $S_l(E)$ 在能量复平面负实轴上的极点。这和我们在散射的一般理论中通过考察 S 矩阵和 T 矩阵的解析性得到的结论相一致。

类似的，$S_l(E)$ 在能量复平面下半平面的极点对应于散射共振态 (或者说不稳定束缚态)。由于 $S_l(E)$ 是一个相因子，因此当能量 E 靠近某个共振态极点 $E_R - i\Gamma\hbar/2$ 时，必有

$$S_l(E) = e^{2i\delta_l^{(0)}(E)}\frac{E-E_R-i\Gamma\hbar/2}{E-E_R+i\Gamma\hbar/2}. \tag{11.143}$$

式中 $\delta_l^{(0)}(E)$ 表示远离共振态的通常散射所贡献的相移。为了将注意力集中在共振态上，我们不妨忽略所有的通常散射贡献，从而令 $\delta_l^{(0)}(E)=0$。这时候取 $S_l(E)=e^{2i\delta_l(E)}$ 的实部，可以得到 $\cos(2\delta_l)=\frac{(E-E_R)^2-(\Gamma\hbar)^2/4}{(E-E_R)^2+(\Gamma\hbar)^2/4}$, 进而可以得到 $\sin^2\delta_l=\frac{(\Gamma\hbar)^2}{4(E-E_R)^2+(\Gamma\hbar)^2}$，将这个结果代入总截面 σ_T 的公式 (11.130), 可以得到

$$\sigma_T(E)\sim\frac{4\pi}{k_E^2}(2l+1)\frac{(\Gamma\hbar)^2}{4(E-E_R)^2+(\Gamma\hbar)^2}. \tag{11.144}$$

这个公式叫做 Breit-Wigner 公式，它告诉我们散射共振态的存在会在散射截面 (作为入射粒子能量 E 的函数) 中引入一个共振峰，这个峰的位置在 E_R 处，宽度由 $\Gamma\hbar$ 来衡量。由于散射共振态的寿命 $\tau=1/\Gamma$，因此共振峰越宽的共振态寿命就越短。

在粒子物理中，大多数微观粒子都是不稳定的，都会衰变，从而可以作为散射截面的共振峰而被发现。这时候由于爱因斯坦质能关系 $E_R=mc^2$，共振峰的位置 E_R 就对应这些不稳定粒子的质量。图 (11.11) 和图 (11.12) 是两个散射截面的例子。图 (11.11) 是正反电子对撞的散射截面，从图中可以看出，这个截面在 $E_R\approx 91GeV$ 处有一个共振峰，其宽度

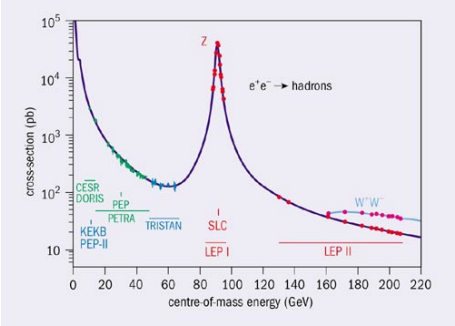

Figure 11.11: Z 玻色子的散射截面。

为 $\Gamma\hbar \approx 2.5GeV$。这个共振峰对应 Z 玻色子，根据其宽度可以算得其寿命约为 $3\times10^{-25}s$。图 (11.12) 所示的是 2012 年发现上帝粒子 (Higgs 玻色子) 的截面数据。从图中共振峰的位置可以读出 Higgs 玻色子的质量约为 $125GeV$。

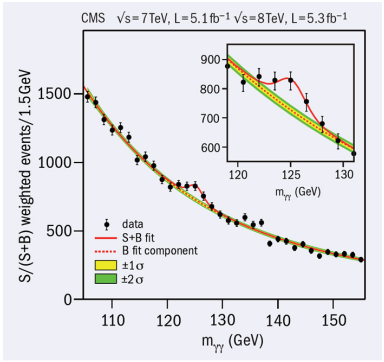

Figure 11.12: 上帝粒子的散射截面。

11.3.5 关于相移的一般性讨论

相移 δ_l 的正负号

在前文中，我们已经看到，为了确定相移，关键是要求解径向方程 (11.136)。利用 $E=\hbar^2k_E^2/(2m)$(这里 m 表示散射粒子的质量)，我们可以将这个方程重写为

$$\begin{aligned} &\frac{d^2}{dr^2}u_{E,l}(r)+\left[k_E^2-2mV(r)/\hbar^2-\frac{l(l+1)}{r^2}\right]u_{E,l}(r)=0 \\ \Leftrightarrow\quad &\frac{d^2}{dr^2}u_{E,l}(r)+k^2(r)u_{E,l}(r)=0. \end{aligned} \tag{11.145}$$

式中，$k(r)=\sqrt{k_E^2-2mV(r)/\hbar^2-l(l+1)/r^2}$ 是 $u_{E,l}(r)$ 随着径向坐标的振荡频率。很显然，当散射势不存在时，$k(r)=k_0(r)=\sqrt{k_E^2-l(l+1)/r^2}$，当散射势是一个吸引势时，即 $V(r)<0$ 时，$k(r)>k_0(r)$，即在吸引势场中，$u_{E,l}(r)$ 将比没有势场时振荡得稍微快一些，反过来，当

散射势是一个排斥势时，即 $V(r)>0$ 时，$u_{E,l}(r)$ 将比没有势场时振荡得稍微慢一些，即 $k(r)<k_0(r)$。

另一方面，根据 (11.140) 式我们也知道，在远离散射势的区域，$u_{E,l}(r)$ 渐近为 $u_{E,l}(r)\sim S_l(E)e^{ik_Er}-e^{-i(k_Er-l\pi)}=e^{2i\delta_l(E)}e^{ik_Er}-e^{-i(k_Er-l\pi)}=e^{i(l\pi/2+\delta_l(E))}\cdot\left(e^{i(k_Er-l\pi/2+\delta_l(E))}-e^{-i(k_Er-l\pi/2+\delta_l(E))}\right)\propto \sin(k_Er-l\pi/2+\delta_l(E))$，即在相差一个整体常数因子的意义上，$u_{E,l}(r)$ 在远离散射区的渐近形式也可以重写为

$$u_{E,l}(r)\sim\sin\big(k_Er-l\pi/2+\delta_l(E)\big). \tag{11.146}$$

当然，如果散射势不存在，则相移 $\delta_l(E)=0$，所以从渐近式 (11.146) 可以知道，相移 $\delta_l(E)$ 其实就是由于散射势 $V(r)$ 的存在而使得振荡函数 $u_{E,l}(r)$ 在散射区额外累积出来的相位。由于当 $V(r)<0$ 时，$u_{E,l}(r)$ 在散射区将振荡得比没有散射势时快，因此它额外累积出来的相位应该为正，即 $\delta_l(E)>0$。反过来，当 $V(r)>0$ 时，$u_{E,l}(r)$ 在散射区将振荡得比没有散射势时慢，因此它额外累积出来的相位应该为负，即 $\delta_l(E)<0$。可见，**一般来说，相移 $\delta_l(E)$ 的正负号总是和散射势的正负号相反的**。

当然，上两段的分析结论仅在 $|V(r)|$ 比较小时才成立。很显然，$|V(r)|$ 越大，则 $k(r)$ 和 $k_0(r)$ 的差别就越大，从而 $|\delta_l(E)|$ 也就会越大。特别的，当 $|V(r)|$ 大到一个适当值时，有可能使得某个分波 l 的 $|\delta_l(E)|=\pi$。这时候根据散射振幅 $f(\theta)$ 的公式 (11.129)，这个 l 分波实际上反而是没有散射的！

低能散射

下面我们简单地讨论一下 δ_l 对角动量量子数 l 的依赖关系。首先，我们假定势场 $V(r)$ 的有效作用范围为 $r<a$，当 $r>a$ 时势场的作用近似可以忽略。其次，一般来说，入射粒子不可能完全对准靶粒子 (或者说散射中心)，散射中心 O 与入射粒子运动直线之间的垂直距离就是所谓的瞄准距离，通常记为 b。因此，如果 $b>a$，则势场的存在就可以忽略。根据经典力学，一个动量为 $\hbar k_E$，瞄准距离为 b 的入射粒子，其角动量为 $L=b\hbar k_E$，而根据量子力学 $L\sim l\hbar$，从而 $l\sim bk_E$，或者说 $b\sim l/k_E$，也即是说，给定入射粒子能量，l 越大的散射分波，其瞄准距离就越大。很显然，当 $l>ak_E$ 时，$b>a$，从而散射势的存在可以忽略。这也即是说，对于一个力程为 a 的散射势，当 $l>ak_E$ 时，相应的 $\delta_l(E)\sim0$。即我们仅仅只需要考察 $l\le ak_E$ 的分波的相移。特别的，如果入射粒子能量非常低，以致于 $k_E\to0$，那么这时候我们将只需要考察 $l=0$ 的 s 分波的散射。

这种低能散射的一个重要例子是所谓的 Ramsauer–Townsend 效应。我们知道，对于惰性气体氩、氪、氙而言，其核外电子的电荷分布通常相当对称。因此这些惰性气体原子在比最外层电子轨道半径 a 更大的距离上所产生的电场可以忽略 (原子核的电场和核外电子的电场几乎完全抵消)。因此当用低能电子和这些惰性气体原子相互碰撞的时候，如果瞄准距离 $b>a$，那散射效应就可以忽略。这也就是说，如果入射电子能量足够低，使得 $ak_E<1$，那我们将仅仅只需考察 $l=0$ 的 s 分波的散射。但是，前面我们也说了，对于一个合适大小的散射势 $|V(r)|$，如果入射粒子能量 E 取一个合适的值 (依然足够低)，那将有可能使得相移 $|\delta_0(E)|=\pi$，那么这时候 s 分波实际上也没有散射。因此，对于氩、氪、氙

这样的惰性气体，如果入射低能电子的能量取一个合适的值，它将可能几乎没有散射，使得散射截面出现一个极小值！实验发现，对于氩原子这个极小值大约出现在 $0.8eV$ 的能量上。这就是 Ramsauer–Townsend 效应，它是一个在量子力学出现之前人们就已经发现的效应，但值到量子力学理论建立起来以后人们才正确地解释了它。

我们已经知道，当入射粒子能量足够低时，我们将只需要考虑最低 l 的那些分波的散射，现在我们将更细致地考察这个结论。首先，由 (11.87) 式我们知道，散射振幅 $f(\theta,\varphi)$ 可以看成是入射粒子波矢量 $\mathbf{k}_i$ 以及出射粒子波矢量 $\mathbf{k}_f$ 的函数，因此我们可以将之重记为 $f(\mathbf{k}_f,\mathbf{k}_i)$, 即 $f(\mathbf{k}_f,\mathbf{k}_i)=f(\theta,\varphi)$。其次，对于中心力场，从 $f(\mathbf{k}_f,\mathbf{k}_i)$ 的分波表达式 (11.129) 可以看到，$f(\mathbf{k}_f,\mathbf{k}_i)$ 实际上是入射粒子能量 E 以及 $\mathbf{k}_i$ 与 $\mathbf{k}_f$ 的夹角 θ 的函数，而由于散射过程的能量守恒 $|\mathbf{k}_i|=|\mathbf{k}_f|=k_E$，由此我们可以知道，$f(\mathbf{k}_f,\mathbf{k}_i)$ 关于其两个自变量 $\mathbf{k}_i$ 和 $\mathbf{k}_f$ 完全是对称的。

假设我们取定 $\mathbf{k}_i$ 沿着水平的 z 轴，则 θ 角就是 $\mathbf{k}_f$ 与 z 轴的夹角，从而完全由 $\mathbf{k}_f$ 的方向决定。另外，我们知道球谐函数 $|\mathbf{k}_f|^l Y_{lm}(\theta,\varphi)$ 当从球坐标变换到直角坐标后，它其实是关于 $\mathbf{k}_f$ 的直角坐标分量的 l 阶齐次函数。而 $Y_{l0}=\sqrt{(2l+1)/(4\pi)}P_l(\cos\theta)$, 因此当然 $|\mathbf{k}_f|^l P_l(\cos\theta)$ 也是关于 $\mathbf{k}_f$ 的直角坐标分量的 l 阶齐次函数。由于自变量 $\mathbf{k}_i$ 和 $\mathbf{k}_f$ 完全对称，因此 $|\mathbf{k}_i|^l|\mathbf{k}_f|^l P_l(\cos\theta)=k_E^{2l}P_l(\cos\theta)$ 将同时是 $\mathbf{k}_i$ 和 $\mathbf{k}_f$ 的直角坐标分量的 l 阶齐次函数。

当入射粒子能量足够低时，$|\mathbf{k}_i|=|\mathbf{k}_f|=k_E$ 将足够小，也即是说，入射波和出射波的波长都将足够长，因此这些波探测不到散射势 $V(r)$ 在小尺度上的任何剧烈变化，对于入射波和散射波而言，散射势看起来是足够平滑的。因此这也就是说，函数 $f(\mathbf{k}_f,\mathbf{k}_i)$ 在 $\mathbf{k}_i\sim 0$, $\mathbf{k}_f\sim 0$ 时是一个足够光滑的函数，因此我们可以在直角坐标中将函数 $f(\mathbf{k}_f,\mathbf{k}_i)$ 在 $(\mathbf{k}_i,\mathbf{k}_f)=(0,0)$ 处进行泰勒展开。我们知道，对于多元函数，泰勒级数的每一阶都将是一个齐次函数，同时又注意到 $f(\mathbf{k}_f,\mathbf{k}_i)$ 关于 $\mathbf{k}_i$ 和 $\mathbf{k}_f$ 完全对称，从而我们可以知道，$f(\mathbf{k}_f,\mathbf{k}_i)$ 的展开式必定可以写成，$f(\mathbf{k}_f,\mathbf{k}_i)=\sum_l H_l(\mathbf{k}_f,\mathbf{k}_i)$, 式中 $H_l(\mathbf{k}_f,\mathbf{k}_i)$ 表示某个关于 $\mathbf{k}_i$ 的直角坐标分量的 l 阶齐次函数，当然同时它也是关于 $\mathbf{k}_f$ 的直角坐标分量的 l 阶齐次函数。将这个泰勒展开的结果代入 $f(\mathbf{k}_f,\mathbf{k}_i)$ 的分波表达式 (11.129) 的等号左边，并在等号右边利用上一段的分析结果，注意到等号左右两边的 l 阶齐次函数必定是一样的，从而即有 $H_l(\mathbf{k}_f,\mathbf{k}_i)\propto k_E^{2l}P_l(\cos\theta)$。进而再一次根据分波表达式 (11.129) 我们就可以知道，当能量足够低时必有

$$e^{i\delta_l(E)}\sin(\delta_l(E))\sim k_E^{2l+1}. \tag{11.147}$$

由于能量足够低时 k_E 足够小，这时候相移 $\delta_l(E)$ 实际上是很小的，因此上式又可以重写为

$$k_E\to 0:\quad \delta_l(E)\sim k_E^{2l+1}. \tag{11.148}$$

这就是低能散射相移的普遍规律，从这个规律也可以清楚地看到，当 k_E 足够小时，我们将仅仅只需考虑 $l=0$ 分波的相移，其它分波的相移都是高阶小量。

对于足够低能的散射，我们将仅仅只需考察 s 分波，根据结果 (11.148), 我们知道 $\delta_0\sim k_E$。当然，由于 δ_0 是无量纲的，而 k_E 是长度的负一次方量纲，因此在足够低能时实

际上我们必定可以将 δ_0 写成如下形式

$$\delta_0(E) \sim -k_E a_s. \tag{11.149}$$

式中 $|k_E a_s| \ll 1$，a_s 可正可负，是某个长度量纲的量，通常称作散射长度。进一步根据总截面的公式 (11.130) 可以知道，在足够低能时，

$$\sigma_T \sim 4\pi a_s^2. \tag{11.150}$$

散射长度的概念对于刻画散射势的低能散射行为非常有用。

高能散射

下面我们简单地讨论一下高能散射情形。具体来说，对于一个力程为 a 的散射势，我们来考察一下入射粒子能量足够高，以致于 $k_E a$ 足够大时的散射。根据我们前面的分析，所有 $l \leq k_E a$ 的分波这时候都需要考虑，因此这时候将有大量的分波需要考虑，根据总截面的公式 (11.130)，这时候 $\sigma_T \to \frac{4\pi}{k_E^2}\sum_{l=0}^{k_E a}(2l+1)\sin^2\delta_l$。由于有大量的 l 需要考虑，因此 δ_l 将随着 l 的变化而在 0 到 2π 的范围内近似于连续地变化，呈现出一种周期性变化规律。因此，我们可以将总截面公式中的 $\sin^2\delta_l$ 近似地用其周期平均值 $\frac{1}{2}$ 来替换，从而有 $\sigma_T \to \frac{2\pi}{k_E^2}\sum_{l=0}^{k_E a}(2l+1)$，当 $k_E a$ 足够大时，求和的最终结果将趋近于 $2\pi a^2$。因此，最终人们发现，高能散射的总截面将趋于

$$\sigma_T \to 2\pi a^2. \tag{11.151}$$

这和低能散射时的 $4\pi a_s^2$ 有根本性的不同。

对于一个半径为 a 的硬球。经典的散射截面当然就是这个硬球的横截面积 πa^2，也即是说，所有碰到这个横截面积上的入射粒子都会被散射。但是在量子的层次上，对于低能散射，我们将得到总截面 $4\pi a^2$，是经典散射截面的 4 倍。当然，低能意味着长波，这时候散射粒子的波动性非常重要，所以我们得到的结果与经典截面有偏离并不出乎意料。但是，即使对于高能散射 (粒子性比较强)，根据我们上一段的分析，量子散射的总截面也将是 $2\pi a^2$，也不同于经典的 πa^2，而是它的两倍。这是因为，在量子的层次上，散射波其实可以分成两部分，其一是反射部分，这一部分产生的截面就是 πa^2。但是，在量子层次上，散射波还有一个朝前散射波部分，这个朝前散射波在散射角 $\theta = 0$ 时最强 (所以叫朝前波)。朝前散射波会和入射波干涉相消，从而在硬球的背后留下一个阴影 (其实并不是完全干涉相消，因为在阴影正中心还会有一个泊松亮斑)。朝前波是必须存在的，否则硬球这一障碍物的背后将不会有任何阴影，这当然不符合我们的物理直觉。实际上，这个朝前散射波对截面的贡献也是 πa^2，因此最终的总截面就是 $2\pi a^2$。

11.3.6 零能束缚态与散射长度：一个例子

这一节我们将考察一个简单的例子。假设我们考察一个入射粒子在下面这样一个球对称势阱 $V(r)$ 上的散射，

$$V(r) = \begin{cases} -V_0, & r < a \\ 0, & r > a \end{cases} \tag{11.152}$$

为了简单起见，我们将仅限于考察 $l=0$ 的 s 波。相应的波函数 $u_E(r)=r\psi_E(r)$ 将满足下面的径向方程

$$\begin{aligned}\frac{d^2}{dr^2}u_E(r)+(k_E^2+k_0^2)u_E(r)=0, \quad r<a\\ \frac{d^2}{dr^2}u_E(r)+k_E^2u_E(r)=0, \quad r>a.\end{aligned} \tag{11.153}$$

式中 $k_E=\sqrt{2mE/\hbar^2}$, $k_0=\sqrt{2mV_0/\hbar^2}$。由于波函数 $\psi_E(r)$ 在 $r=0$ 处有限，所以 $u_E(r)$ 得满足 $u_E(0)=0$ 的边界条件。由此我们可以得到方程 (11.153) 的解，为

$$\begin{aligned}u_E(r)=A\sin(\sqrt{k_E^2+k_0^2}r), \quad r<a\\ u_E(r)=B\sin(k_Er+\delta_0), \quad r>a\end{aligned} \tag{11.154}$$

式中 δ_0 就是 s 波相移。

根据 $u_E(r)$ 以及它的一阶导数 $u_E'(r)$ 在 $r=a$ 处的连续性, 即 $u_E'(a)/u_E(a)$ 的连续性。我们可以得到

$$\tan(k_Ea+\delta_0)=\frac{k_E}{\sqrt{k_E^2+k_0^2}}\tan(a\sqrt{k_E^2+k_0^2})=f(k_E). \tag{11.155}$$

由这个式子我们又可以得到 s 分波的 S 矩阵 $S_0(E)=e^{2i\delta_0}$ 为，

$$S_0(E)=e^{-2ik_Ea}\frac{1+if(k_E)}{1-if(k_E)}. \tag{11.156}$$

假设我们将注意力集中在低能散射上，即考察 $k_E^2\ll k_0^2$, 同时 $k_Ea\ll 1$ 的情形。这时候方程 (11.155) 可以近似为

$$\frac{\tan(k_Ea)+\tan(\delta_0)}{1-\tan(k_Ea)\tan(\delta_0)}\approx\frac{k_E}{k_0}\tan(k_0a). \tag{11.157}$$

利用 $\tan(k_Ea)\approx k_Ea\ll 1$ 我们可以进一步得到

$$\delta_0\approx\tan(\delta_0)\approx k_Ea\left(\frac{\tan(k_0a)}{k_0a}-1\right). \tag{11.158}$$

根据散射长度 a_s 的定义 (即 $\delta_0\approx -k_Ea_s$) 我们可以知道

$$a_s=a-\frac{\tan(k_0a)}{k_0}. \tag{11.159}$$

很显然，当 V_0 很小，也就是 k_0 很小时，这个散射长度为负 (从而 δ_0 为正)，注意到我们这里考察的是一个吸引势，因此这与我们前面得到的 δ_l 的正负号与 $V(r)$ 相反是一致的。但是，随着 k_0 的增加，显然 a_s 将会从 0 开始沿着负方向逐渐趋于 $-\infty$(当 $k_0a=\pi/2$ 时)，之后如果 k_0 进一步增加，使得 $k_0a=\pi/2+\varepsilon$(ε 为一个正的无穷小量)，则 a_s 会从 $-\infty$ 跳到 $+\infty$。之后，k_0 再增加，a_s 就会逐渐减小，直至再次趋于 $-\infty$，然后再次跳到 $+\infty$，如此不断循环。总之，随着 V_0 的逐渐增加，散射长度 a_s 将会不断地在 $-\infty$ 到 $+\infty$ 之间振荡。

我们想问的是，散射长度的这一奇异振荡行为起因于哪里？特别的，类似于上面的推导人们不难发现，如果散射势是一个排斥势，即 $V_0 < 0$，那相应的散射长度 a_s 将始终大于 0，并且有限，因此不会有 $-\infty$ 到 $+\infty$ 之间的振荡。

理解吸引势时 a_s 的这种发散和振荡行为的关键在于考察系统的束缚态。为此我们将入射粒子能量 E 解析延拓到 $E < 0$ 情形，等价的，这也就是将 k_E 解析延拓到正虚轴，即取 $k_E = i\lambda_E (\lambda_E > 0)$。束缚态相应于 S 矩阵的极点，因此根据 (11.156) 式，束缚态能级 E_n 将满足 $f(k_{E_n}) = -i$。进一步根据 $f(k_E)$ 的定义式 (11.155), 以及解析延拓关系 $k_{E_n} = i\lambda_{E_n}$，这也就是要满足

$$\tan(a\sqrt{k_0^2 - \lambda_{E_n}^2}) = -\frac{\sqrt{k_0^2 - \lambda_{E_n}^2}}{\lambda_{E_n}}. \tag{11.160}$$

我们可以引入 $q_E = \sqrt{k_0^2 - \lambda_E^2}$, 从而将这个方程写为

$$q_{E_n} = -\lambda_{E_n} \tan(a q_{E_n}), \qquad q_{E_n}^2 + \lambda_{E_n}^2 = k_0^2. \tag{11.161}$$

这个方程组是我们前面在一维势散射求束缚态的例子中碰到过的，我们可以用画图的方法求解这组方程，我们知道，当势阱足够深，从而 V_0 和 k_0 足够大时，这组方程有解，从而束缚态存在。

散射长度 a_s 的发散行为与一种特殊的束缚态密切相关。一般来说，束缚态的能量 E_n 当然小于 0，但是，如果 V_0 取一个合适值，将有可能使得最高束缚态能级的能量趋于 0，从而出现一种特殊的束缚态，我们称之为零能束缚态。为了求出什么时候会出现零能束缚态，我们直接在方程 (11.160) 中令 $\lambda_{E_n} = 0$，从而得到零能束缚态的条件 $\tan(k_0 a) = \infty$(或者等价的 $k_0 a = (n + \frac{1}{2})\pi$)。很显然，正是在零能束缚态出现的时候，散射长度 a_s 变得发散。随着 k_0 的增加，在零能束缚态出现之前，$\tan(k_0 a)$ 逐渐趋于 $+\infty$，这时散射长度 a_s 逐渐趋于 $-\infty$，但是，从方程 (11.160) 很容易看到，一旦出现零能束缚态，$\tan(k_0 a)$ 将立即从 $+\infty$ 跳到 $-\infty$，相应的散射长度 a_s 就将从 $-\infty$ 跳到 $+\infty$。可见，散射长度的发散以及从 $-\infty$ 到 $+\infty$ 的跳变行为正是由于零能束缚态引起的。随着 k_0 的增加，不断会有新的零能束缚态出现，因此散射长度 a_s 就周期性地从 $-\infty$ 跳变到 $+\infty$。

为了进一步理解散射长度 a_s 在零能束缚态附近的发散行为，我们不妨记最高束缚态能级为 $E_0 < 0$, 最后我们将考察 $E_0 \to 0$ 的零能束缚态极限 (这时候 $\lambda_{E_0} \to 0$)。根据定义，$i\lambda_{E_0}$ 是 S 矩阵 $S_0(E)$ 的一个极点，在这个极点附近，$S_0(E)$ 可以近似成

$$S_0(E) = e^{2i\delta_0} \sim \frac{i\lambda_{E_0} + k_E}{i\lambda_{E_0} - k_E}. \tag{11.162}$$

由于我们考察的是低能散射，所以 $k_E \ll \lambda_{E_0}$，由此我们近似地可以得到 $\delta_0 \sim -k_E/\lambda_{E_0}$, 从而即有

$$a_s \sim \frac{1}{\lambda_{E_0}}. \tag{11.163}$$

很显然，当 E_0 趋于一个零能束缚态时，即 $\lambda_{E_0} \to 0$ 时，a_s 将发散。

11.4 附注

关于一维势散射的两个具体求解例子，关于卢瑟福公式的精确推导，以及关于零能束缚态和散射长度的那个例子，均参考了 D. Tong 的讲义《Applications of Quantum Mechanics》。



12. * 密度算符与熵

本章讲述密度算符的概念和其含义。本章也将从一个更高的角度重新审视量子力学的基本原理，并介绍量子力学的多世界诠释。

本章随之系统讲述了冯诺伊曼熵和量子相对熵，以及它们的各种性质。本章还会推导开放系统随着时间的演化方程，并证明热力学第二定律。

12.1 密度算符与量子力学原理

前面章节研究的量子力学系统都是所谓的封闭完整系统，也就是说，我们忽略了系统与外界的相互作用，并且我们研究的是整个系统而不是其中的一部分。但是，现实中的系统总是从属于一个更大的系统，是更大系统的一部分，而不是全部。现实中的系统也总是会与外界环境甚至包括我们的仪器相互作用。如何描述这样的现实系统就是本节要解决的核心问题。通过对这一问题的讨论，我们也得以从一个更高的视角来重新审视量子力学的基本原理。

12.1.1 期望值以及算符迹

下面我们先作一些准备。

测量与期望值

在第二章关于量子力学基本原理的讨论中，我们得出过一个重要公式：对某个物理量 A 多次重复实验测得的平均值 $\overline{A}$ 等于算符 A 在态上的期望值，即

$$\overline{A} = \langle\psi|A|\psi\rangle, \tag{12.1}$$

$|\psi\rangle$ 为系统所处的量子态。我们也说过，物理量 A 的测量值是相应算符 A 的本征值，在单次实验中，我们测得哪个值是随机的，测得本征值 λ_i 的概率为 $p_i = |\langle i|\psi\rangle|^2$，式中 $|i\rangle$ 为与本征值 λ_i 相应的本征态。这样算出来的概率 p_i 当然也可以通过重复多次实验来检验，因此其计算公式应该也能写成期望值的形式，的确，假如定义投影算符 $P_i = |i\rangle\langle i|$，则人们很容易验证

$$p_i = \langle\psi|i\rangle\langle i|\psi\rangle = \langle\psi|P_i|\psi\rangle. \tag{12.2}$$

因此，不仅仅物理量的平均值，而且物理量值的概率分布，都可以表达成厄密算符期望值的形式。

以上讨论告诉我们，在量子力学中一切可以通过重复实验来进行检验的量都可以由一个适当厄密算符的期望值来计算，假设这个厄密算符是 $\mathscr{O}$，习惯上，人们常将这个期望值简记为 $\langle\mathscr{O}\rangle$，

$$\langle\mathscr{O}\rangle = \langle\psi|\mathscr{O}|\psi\rangle. \tag{12.3}$$

各种厄密算符的期望值就是我们能够从一个量子态中提取的所有信息。

算符求迹

一个矩阵所有对角元的和称为矩阵的迹，在幺正变换下矩阵的迹保持不变，这使得我们可以将迹的概念推广到线性算符。对于一个线性算符 $\mathscr{O}$，我们记其迹为 $\mathrm{Tr}(\mathscr{O})$，定义为

$$\mathrm{Tr}(\mathscr{O}) = \sum_i \langle i|\mathscr{O}|i\rangle, \tag{12.4}$$

式中 $\{|i\rangle\}$ 为希尔伯特空间的一组正交归一矢量基。虽然为了计算算符的迹我们需要选取一个特定的表象，但由于不同的表象之间只相差一个幺正变换，所以算符的迹实际上并不

依赖于表象。由此可以知道，一个算符的迹就是它在任何一个表象中表示矩阵的迹，特别的，对于厄密算符，我们可以将这个表象选为它的本征表象，这时候算符的迹其实就是所有本征值的和。

假设有两个算符 A 和 B，读者容易证明算符迹满足如下等式

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA). \tag{12.5}$$

不过，值得说明的是，对于无穷维希尔伯特空间，并非所有算符的迹都是定义良好的，因此在无穷维希尔伯特空间上应用 (12.5) 式时得特别小心，有时候它并不成立。比方说，对于坐标算符和动量算符的对易子来说，如果简单地应用这个公式将会得出 $\mathrm{Tr}([X,P]) = 0$，但实际上 $\mathrm{Tr}([X,P]) = i\hbar\mathrm{Tr}(1) \neq 0$。在这个例子中，(12.5) 式之所以不成立，正是因为式中各算符的迹都不是良好定义的。

利用算符的迹，我们可以把算符的期望值公式重写成

$$\langle\psi|\mathscr{O}|\psi\rangle = \mathrm{Tr}\big(|\psi\rangle\langle\psi|\mathscr{O}\big) = \mathrm{Tr}\big(\rho_\psi\mathscr{O}\big). \tag{12.6}$$

式中投影算符 $\rho_\psi = |\psi\rangle\langle\psi|$。要证明这个式子，我们只需注意到

$$\mathrm{Tr}\big(|\psi\rangle\langle\psi|\mathscr{O}\big) = \sum_i \langle i|\psi\rangle\langle\psi|\mathscr{O}|i\rangle = \sum_i \langle\psi|\mathscr{O}|i\rangle\langle i|\psi\rangle = \langle\psi|\mathscr{O}|\psi\rangle.$$

可见，为了计算算符的期望值，我们并不需要知道量子态 $|\psi\rangle$，而是只需要知道厄密算符 ρ_ψ。$|\psi\rangle$ 和 ρ_ψ 的一个重要区别是，$|\psi\rangle$ 可以相差一个非物理的整体相位因子，变成 $e^{i\theta}|\psi\rangle$，而 ρ_ψ 在这种相位变换下不变，也即是说，它自动剔除了这一非物理的整体相位信息。

以上讨论告诉我们，从一个量子系统中能够提取出来的所有物理信息都包含在 $\rho_\psi = |\psi\rangle\langle\psi|$ 这一特殊厄密算符中。注意到 $\langle x|\rho_\psi|x\rangle = |\langle x|\psi\rangle|^2$ 正好是概率密度。即是说，ρ_ψ 在坐标表象中的对角元对应概率密度，所以人们常常称 ρ_ψ 为纯态密度算符，它和狄拉克符号 $|\psi\rangle$ 表示的量子态是对应的，这样的态也称为纯态。下面我们就是要将密度算符这一概念推广到更为一般的情形。

12.1.2 从量子纠缠态到混态

密度算符的一般概念和这样的问题密切相关，即对于一个由两部分组成的系统，假设我们仅仅关心其中的一部分，该如何描述我们所关心的对象？这一小节我们就是要解决这个问题。

一个例子

不妨先看一个例子。假设我们有两个量子比特，分别称作量子比特 A 和量子比特 B，它们的希尔伯特分别为 $\mathscr{H}_A$ 和 $\mathscr{H}_B$。假设 A,B 两者共同处在 $|\psi\rangle$ 态，

$$|\psi\rangle = c_0|00\rangle + c_1|11\rangle. \tag{12.7}$$

则我们很容易写出与这个态相应的纯态密度算符

$$\rho_\psi = |c_0|^2|00\rangle\langle 00| + |c_1|^2|11\rangle\langle 11| + c_0c_1^*|00\rangle\langle 11| + c_1c_0^*|11\rangle\langle 00|. \tag{12.8}$$

我们称这个密度算符前面的两项为对角项，后面的两项为非对角项。非对角项相应于通常所说的干涉项，它描写了相互叠加的两个态之间的干涉。

假设我们忽略 B 而仅仅只关心 A, 对 A 进行各种测量。比方说我们测量 A 的某个物理量 $\mathscr{O}_A$, 因此我们就要计算如下期望值，

$$\langle\psi|\mathscr{O}_A|\psi\rangle = \mathrm{Tr}_{AB}\big(\rho_\psi \mathscr{O}_A\big). \tag{12.9}$$

式中 Tr_{AB} 表示在整个系统的张量积希尔伯特空间 $\mathscr{H}_A\otimes\mathscr{H}_B$ 上求迹。注意到 $\mathscr{H}_A\otimes\mathscr{H}_B$ 的基矢量 $|ij\rangle_{AB} = |i\rangle_A|j\rangle_B$, $i,j=0,1$, 所以 $\mathrm{Tr}_{AB} = \mathrm{Tr}_A\mathrm{Tr}_B$, 式中 Tr_A 和 Tr_B 分别表示在 $\mathscr{H}_A$ 和 $\mathscr{H}_B$ 上求迹。由于 $\mathscr{O}_A$ 只作用在 $\mathscr{H}_A$ 上，与 B 量子比特无关，所以 Tr_B 就直接作用在 ρ_ψ 上，得到

$$\rho_A = \mathrm{Tr}_B(\rho_\psi) = |c_0|^2|0\rangle\langle 0| + |c_1|^2|1\rangle\langle 1|. \tag{12.10}$$

这是一个只作用在 $\mathscr{H}_A$ 上的厄密算符。而期望值公式 (12.9) 就变成

$$\langle\psi|\mathscr{O}_A|\psi\rangle = \mathrm{Tr}_A\big(\rho_A \mathscr{O}_A\big). \tag{12.11}$$

以上讨论告诉我们，如果仅仅只关心 A 量子比特，那就无需知道 $|\psi\rangle$，只需知道 ρ_A，忽略 B 以后所有能从 A 量子比特提取出来的信息都包含在厄密算符 ρ_A 中。我们称 ρ_A 为 A 量子比特的密度算符，当然它不是前面引入的纯态密度算符，因为现在这个密度算符不能写成 $|\phi_A\rangle\langle\phi_A|$ 这样的形式，我们称这种密度算符为混态密度算符，简称混态。很明显，ρ_A 满足

$$\mathrm{Tr}_A\big(\rho_A\big) = |c_0|^2 + |c_1|^2 = 1. \tag{12.12}$$

在这个例子中，系统所处的 $|\psi\rangle = c_0|00\rangle + c_1|11\rangle$ 态显然是一个纠缠态，上面的推导过程告诉我们，如果忽略 B 量子比特 (数学上就是将 ρ_ψ 对 B 求迹，称作 trace 掉 B)，我们将得到一个 A 的混态密度算符 ρ_A，称作 ρ_ψ 在 A 上的约化密度算符。从 (12.10) 式可以看得很清楚，在密度算符 ρ_A 中，量子比特 0 态与 1 态间的干涉项消失了，这就叫做退相干。在这个例子中，A 的 0 态与 1 态退相干的原因有两个: 首先，A 和 B 发生了量子纠缠，其次，我们忽略了 B。

从 (12.10) 式还可以看到，退相干的 A 量子比特对于忽略了 B 的我们来说，完全可以看作是以 $|c_0|^2$ 的概率处于 $|0\rangle$ 态，以 $|c_1|^2$ 的概率处于 $|1\rangle$ 态，这分别对应 (12.10) 式中的 $|c_0|^2|0\rangle\langle 0|$ 和 $|c_1|^2|1\rangle\langle 1|$。

量子擦除

我们知道，一个量子比特如果处在 $|+\rangle = \big(|0\rangle + |1\rangle\big)/\sqrt{2}$ 态，那它的 $|0\rangle$ 态和 $|1\rangle$ 态之间将会有干涉，两者的相对位相可以观测，通过对相对位相的观测我们可以把 $|+\rangle$ 态和 $|-\rangle = \big(|0\rangle - |1\rangle\big)/\sqrt{2}$ 态 100% 地区分开来。但是，从前面的讨论中我们也知道，如果将这个量子比特作为 A 与另一个量子比特 B 纠缠起来，使得 AB 整体处于 $|\phi^+\rangle = \big(|00\rangle + |11\rangle\big)/\sqrt{2} = \big(|++\rangle + |--\rangle\big)/\sqrt{2}$ 态，那么当我们忽略 B 的时候，A 的 $|0\rangle$ 态和 $|1\rangle$

态之间将会退相干，这时候 A 将由混态密度算符 $\frac{1}{2}\big(|0\rangle\langle 0|+|1\rangle\langle 1|\big)=\frac{1}{2}\big(|+\rangle\langle +|+|-\rangle\langle -|\big)$ 描述，$|0\rangle,|1\rangle$ 之间的相对位相将会变得不可观测。这种不可观测是因为这时候 A 量子比特取值是 0 还是 1 的信息通过与 B 的纠缠泄露出去了。

但是，假如我们不是完全忽略 B 的信息，那就有一种方法可以恢复 A 量子比特 $|0\rangle$ 态和 $|1\rangle$ 态之间的相干叠加性。这方法就是，假设我们不是在 $\{|0\rangle,|1\rangle\}$ 基，而是在 $\{|+\rangle,|-\rangle\}$ 基中观测 B 量子比特，并根据观测结果将所有的 A 量子比特分成 $+,-$ 两类。$+$ 类 A 量子比特相应于对 B 的观测结果是 $|+\rangle$，$-$ 类 A 量子比特则相应于对 B 的观测结果是 $|-\rangle$。那么这时候 $+$ 类 A 量子比特的 $|0\rangle,|1\rangle$ 态之间将是相干叠加的，$-$ 类也是如此。这是因为，AB 整体的量子态是 $|\phi^+\rangle=\big(|++\rangle+|--\rangle\big)/\sqrt{2}$, 因此当 B 的观测结果是 $|+\rangle$ 时，对相应 A 量子比特的描述将不是混态密度算符 $\frac{1}{2}\big(|0\rangle\langle 0|+|1\rangle\langle 1|\big)$，而是纯态 $|+\rangle$，同样，当 B 的观测结果是 $|-\rangle$ 时，对相应 A 量子比特的描述也是纯态 $|-\rangle$。当我们根据对 B 的观测结果将这两类 A 量子比特分开时，它们各自当然都是相干的。

上面描述的这种恢复 A 量子比特相干性的方法就叫做量子擦除，因为无论对 B 的测量结果是 $|+\rangle$ 还是 $|-\rangle$, 我们都得不到任何取值是 0 还是 1 的信息。因此这种测量就从 B 量子比特中擦除了通过纠缠泄露出去的 A 取值是 0 还是 1 的信息。

两体纠缠的刻画-施密特分解

前面两量子比特的例子所揭示的量子纠缠态与子系统混态之间的联系是一般性的，为了看清楚这一点，让我们更仔细地考察一下两体量子纠缠。

对于一个由 A、B 两部分组成的系统，假设这两部分的希尔伯特空间分别为 $\mathscr{H}_A$ 和 $\mathscr{H}_B$，则任意量子态 $|\psi\rangle$ 必定可以写成

$$|\psi\rangle=\sum_{i_A,i_B}\psi_{i_Ai_B}|i_A\rangle|i_B\rangle, \tag{12.13}$$

式中 $\{|i_A\rangle\}$ 和 $\{|i_B\rangle\}$ 分别为 $\mathscr{H}_A$ 和 $\mathscr{H}_B$ 的正交归一矢量基。式中的叠加系数 $\psi_{i_Ai_B}$ 有两个指标，因此可以看成是一个矩阵，根据矩阵的奇异值分解定理 [1]，$\psi_{i_Ai_B}$ 必定可以分解成如下形式

$$\psi_{i_Ai_B}=\sum_a\lambda_a u_{i_A}^{(a)}v_{i_B}^{(a)}, \tag{12.14}$$

式中 $u_{i_A}^{(a)}$ 和 $v_{i_B}^{(a)}$ 分别满足正交归一条件 $\sum_{i_A}(u_{i_A}^{(a)})^*u_{i_A}^{(b)}=\delta_{ab}$ 和 $\sum_{i_B}(v_{i_B}^{(a)})^*v_{i_B}^{(b)}=\delta_{ab}$, 并且我们总可以通过调节 $u_{i_A}^{(a)}$ 和 $v_{i_B}^{(a)}$ 之间的相对相位使得 $\lambda_a\geq 0$, 因此我们可以令 $\lambda_a=\sqrt{p_a}$, $p_a\geq 0$。现在记 $|a\rangle_A=\sum_{i_A}u_{i_A}^{(a)}|i_A\rangle$, $|a\rangle_B=\sum_{i_B}v_{i_B}^{(a)}|i_B\rangle$，则由 $u_{i_A}^{(a)}$ 和 $v_{i_B}^{(a)}$ 的正交归一性可以证明，$|a\rangle_A$ 和 $|a\rangle_B$ 分别正交归一，即 ${}_A\langle a|b\rangle_A={}_B\langle a|b\rangle_B=\delta_{ab}$。将 (12.14) 式代入 (12.13) 就可以得到

$$|\psi\rangle=\sum_a\sqrt{p_a}|a\rangle_A|a\rangle_B. \tag{12.15}$$

[1] 如果将这个矩阵看成是某个算符的表示矩阵，那矩阵的奇异值分解定理其实就来源于算符的奇异值分解定理。

进一步利用 $|\psi\rangle$ 态的归一化条件，容易得到

$$\sum_a p_a = 1. \tag{12.16}$$

(12.15) 式就叫做两体系统量子态的施密特分解，其中非零的 p_a 数目就称作量子态的施密特数。显然，斯密特数最小等于 1，最大等于 $\mathscr{H}_A$ 和 $\mathscr{H}_B$ 维数中较小的那个。如果施密特数等于 1，则由 (12.15) 式很容易看出，这时候量子态 $|\psi\rangle$ 必定分解成 A,B 量子态乘积的形式，反过来，如果施密特数大于 1，那就必定没有这样简单的分解，从而 $|\psi\rangle$ 态就必定为一个纠缠态。即是说，$|\psi\rangle$ 不是纠缠态的充要条件是施密特数等于 1，因此施密特数多少反映了两体纠缠的状况，施密特数越大说明 A,B 纠缠得越厉害。

密度算符与系综

一个量子系统通常总是处于某个环境之中，假设我们把系统看作 A，把环境看作 B，那么整个世界就可以看成是一个由 A, B 两部分组成的封闭完整系统，其 A, B 两部分通常会纠缠起来。下面我们对整个世界的量子态应用施密特分解公式 (12.15)，我们记整个世界的密度算符为 $\rho_\psi = |\psi\rangle\langle\psi|$，假设我们忽略 B 部分，那就会得到 A 的约化密度算符 $\rho_A = \mathrm{Tr}_B(\rho_\psi)$, 我们从 A 中能够提取出来的所有信息都包含在 ρ_A 中，并可以通过计算算符期望值 $\mathrm{Tr}_A\big(\rho_A\mathscr{O}_A\big)$ 来得到，因此我们也称 ρ_A 描写了子系统 A 的一个量子态，只不过这个量子态不一定是纯态，而可以是混态。

由 (12.15) 式容易有

$$\rho_A = \sum_a p_a |a\rangle_A\langle a|. \tag{12.17}$$

显然，ρ_A 满足 $\mathrm{Tr}_A\big(\rho_A\big) = \sum_a p_a = 1$。由于我们没有对 A 和 B 进行任何限制，所以可以想见，(12.17) 应该能代表 A 的一个任意量子态。从 (12.17) 可以看出，施密特分解的 p_a 刚好是密度算符 ρ_A 的本征值，施密特分解的 $|a\rangle_A$ 则是相应的本征态。同样的，我们也可以忽略 A 部分，进而得到 B 部分的约化密度算符 $\rho_B = \mathrm{Tr}_A\left(\rho_\psi\right)$，为

$$\rho_B = \sum_a p_a |a\rangle_B\langle a|. \tag{12.18}$$

如果原来 AB 整体的量子态 $|\psi\rangle$ 不是一个纠缠态，也就是说施密特数等于 1，假设非零的 p_a 为 $a=1$ 的 p_1, 则由于 $\sum_a p_a = 1$，必有 $p_1 = 1$，由 (12.17) 和 (12.18) 的表达式可见，这时候必有

$$\rho_A = |1\rangle_A\langle 1|, \quad \rho_B = |1\rangle_B\langle 1|. \tag{12.19}$$

显然这是两个纯态密度算符。反过来，如果原来的 $|\psi\rangle$ 是一个纠缠态，施密特数大于 1，那 ρ_A, ρ_B 就是混态密度算符。所以整个完整系统 AB 的纠缠态和子系统 A 或者 B 的混态是相对应的。

当然，我们可以完全不提 B 部分，直接定义一个系统的密度算符，而不管它是否源于系统与环境的纠缠。一个系统的密度算符 ρ 包含了所有我们能从对这个系统的观测中提取出来的信息，根据上面的讨论可以知道，**密度算符 ρ 应该满足如下性质**：1. 它是一个

厄密算符。2. 它的本征值都大于等于 0，为正定算符。3. $\mathrm{Tr}(\rho)=1$。有了密度算符以后，对任何物理可观测量 $\mathscr{O}$ 期望值的计算就可以表达为

$$\langle\mathscr{O}\rangle=\mathrm{Tr}(\rho\mathscr{O}). \tag{12.20}$$

如果一个密度算符能够写成 $\rho=|\psi\rangle\langle\psi|$ 的形式，就称为纯态密度算符，它描述的系统状态就是纯态，否则就是混态密度算符，描述的状态就是混态。很显然，纯态密度算符额外满足

$$\rho^2=\rho. \tag{12.21}$$

反过来，如果一个密度算符额外满足上式，则它的本征值必为 0 或 1，由于密度算符所有本征值的和要等于 1，因此它就必定能写成 $\rho=|\psi\rangle\langle\psi|$ 的形式 ($|\psi\rangle$ 为 ρ 的本征态)，从而必为纯态密度算符。对于纯态密度算符，我们必定有

$$\mathrm{Tr}(\rho^2)=\mathrm{Tr}(\rho)=1. \tag{12.22}$$

对于一个任意的密度算符 ρ，我们必可以通过求解它的本征方程将它对角化成如下形式

$$\rho=\sum_a p_a|a\rangle\langle a|. \tag{12.23}$$

密度算符的正定性告诉我们 $p_a\geq 0$, $\mathrm{Tr}(\rho)=1$ 的条件则告诉我们 $\sum_a p_a=1$，所以 $p_a\leq 1$，等号仅当 ρ 为纯态密度算符时才成立。因此，对于混态密度算符，我们必有 $\mathrm{Tr}(\rho^2)=\sum_a p_a^2<\sum_a p_a=1$，即对于混态密度算符，必有

$$\mathrm{Tr}(\rho^2)<1. \tag{12.24}$$

所有系统都可以看作某个更大的 AB 系统的 A 部分，并且 (12.23) 的形式以及前面关于施密特分解的知识进一步告诉我们，所有的密度算符都可以看成由这个更大系统的某个具有施密特分解 (12.15) 形式的纯态 $|\psi\rangle$ 通过忽略 B 部分来得到，(12.15) 就称作密度算符 (12.23) 的纯化。

然而，密度算符的概念其实并不需要通过它的纯化来理解。密度算符的 (12.23) 形式启发我们可以从系综的角度解释密度算符。也即是说，我们可以认为密度算符描述的是大量系统的集合，其中每一个系统以概率 p_a 处于 $|a\rangle$ 态。这是因为，这一系综中任何物理量 $\mathscr{O}$ 的期望值同样由下式给出

$$\langle\mathscr{O}\rangle=\sum_a p_a\langle a|\mathscr{O}|a\rangle=\mathrm{Tr}(\rho\mathscr{O}). \tag{12.25}$$

密度算符的系综解释和纯化解释是有所不同的，纯化解释可以用于单个系统，而系综解释只适用于大量系统的系综。但是，由于对任何物理量期望值的测量都需要多次重复实验，这相当于引入了大量系统，因此在实验上我们实际上无法区分密度算符的纯化解释和系综解释。

不过，根据系综解释，我们可以将密度算符中出现的概率 p_a 看成是源于我们对系统信息的缺失，它造成了我们对系统的某种无知。在纯化解释中，系统作为 A 部分与辅助的 B 纠缠成了一个整体，在这个整体中 A 和 B 的“个性”都已经失去，而 p_a 所反映的信息缺失就源于我们对 B 部分的忽略。总之，密度算符描述中包含了对系统的某种无知，这种无知来源于对系统的信息缺失，而系综诠释的好处是，它允许这种信息缺失和纠缠无关，任何信息缺失都可以对应于一个系综，从而给出一个密度算符。然而，如何量化密度算符中包含的无知呢？这个问题我们后面再来讨论。

12.1.3 Bloch 球

下面我们详细地讨论一下单个量子比特的密度算符 ρ。这时候希尔伯特空间是 2 维的，为了具体起见我们可以将基矢量 $\{|0\rangle,|1\rangle\}$ 表示成

$$|0\rangle=\begin{pmatrix}1\\0\end{pmatrix}\quad |1\rangle=\begin{pmatrix}0\\1\end{pmatrix}. \tag{12.26}$$

如此一来，密度算符 ρ 就应该表示成一个 2×2 的厄密矩阵，称为密度矩阵。但是，任何 2×2 的厄密矩阵都可以用 $\{1,\sigma_x,\sigma_y,\sigma_z\}$ 这四个厄密矩阵展开为 $\rho=\frac{1}{2}(w_0\cdot 1+\mathbf{w}\cdot\sigma)$, 式中 $\{w_0,\mathbf{w}\}$ 均为实数。注意到泡利矩阵的迹为零，则由 $\mathrm{Tr}(\rho)=1$ 可知，$w_0=1$, 从而单个量子比特的密度矩阵 ρ 必定可以表示成如下形式

$$\rho(\mathbf{w})=\frac{1}{2}\big(1+\mathbf{w}\cdot\sigma\big). \tag{12.27}$$

$\mathbf{w}$ 称为量子比特的极化矢量，所以单量子比特的密度矩阵完全由其极化矢量刻画。另外，容易算出 $\mathrm{Tr}(\rho^2)=\frac{1}{2}(1+\mathbf{w}^2)$, 由 $\mathrm{Tr}(\rho^2)\le 1$ 可知

$$\mathbf{w}^2\le 1, \tag{12.28}$$

等号当且仅当 $\rho(\mathbf{w})$ 为一个纯态密度矩阵时才成立。

由上面的讨论可知，单个量子比特的所有可能密度矩阵一一对应于 3 维极化矢量空间中的一个单位球体 $\mathbf{w}^2\le 1$，称为 Bloch 球，这个球体的表面对应于单量子比特的可能纯态，球内部分描写的当然就是混态。特别的，在 Bloch 球体中心，$\mathbf{w}=0$ 处的那个混态，就是通常所谓的最大混态，很显然，它是

$$\rho(0)=\frac{1}{2}\cdot 1=\frac{1}{2}\big(|0\rangle\langle 0|+|1\rangle\langle 1|\big). \tag{12.29}$$

之所以说 $\rho(\mathbf{w})$ 表达式中的 $\mathbf{w}$ 是极化矢量，是因为，假设将单量子比特看成电子自旋，$|0\rangle$ 态看成自旋向上态，$|1\rangle$ 态看成自旋向下态，则我们可以计算 $\mathbf{n}$ 方向泡利算符 $(\mathbf{n}\cdot\sigma)$ 的期望值，易得

$$\langle(\mathbf{n}\cdot\sigma)\rangle_{\mathbf{w}}=\mathrm{Tr}((\mathbf{n}\cdot\sigma)\rho(\mathbf{w}))=\mathbf{n}\cdot\mathbf{w}. \tag{12.30}$$

这告诉我们，$\mathbf{n}$ 方向自旋期望值是矢量 $\mathbf{w}$ 在 $\mathbf{n}$ 方向上的投影，可见 $\mathbf{w}$ 正好代表自旋极化。

假设在单位球面上取一个单位矢量 $\mathbf{n}=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$, $\mathbf{n}^2=1$，则纯态密度矩阵就是 $\rho(\mathbf{n})=\frac{1}{2}(1+\mathbf{n}\cdot\sigma)$。假设记相应的纯态为 $|\psi(\mathbf{n})\rangle$，即 $\rho(\mathbf{n})=|\psi(\mathbf{n})\rangle\langle\psi(\mathbf{n})|$，则由于 $(\mathbf{n}\cdot\sigma)\rho(\mathbf{n})=\rho(\mathbf{n})(\mathbf{n}\cdot\sigma)=\rho(\mathbf{n})$，容易看出，$|\psi(\mathbf{n})\rangle$ 必为 $(\mathbf{n}\cdot\sigma)$ 的本征值为 1 的本征态。

利用 3 个泡利矩阵的表达式，并将单位矢量 $\mathbf{n}$ 的分量形式代入 $\rho(\mathbf{n})$，容易得到

$$\rho(\mathbf{n})=\begin{pmatrix}\cos^2(\theta/2) & \cos(\theta/2)\sin(\theta/2)e^{-i\phi}\\ \cos(\theta/2)\sin(\theta/2)e^{i\phi} & \sin^2(\theta/2)\end{pmatrix}. \tag{12.31}$$

由此可以得到

$$|\psi(\mathbf{n})\rangle=\begin{pmatrix}\cos(\theta/2)e^{-i\phi/2}\\ \sin(\theta/2)e^{i\phi/2}\end{pmatrix}. \tag{12.32}$$

假设有两个密度矩阵 $\rho(\mathbf{w}_1)$ 和 $\rho(\mathbf{w}_2)$, 则对于任意 $0\le t\le 1$ 均有

$$t\rho(\mathbf{w}_1)+(1-t)\rho(\mathbf{w}_2)=\rho(t\mathbf{w}_1+(1-t)\mathbf{w}_2), \tag{12.33}$$

也即是说，两个密度矩阵按照 $t\rho(\mathbf{w}_1)+(1-t)\rho(\mathbf{w}_2)$ 的形式组合以后结果依然是一个密度矩阵，而且这个密度矩阵的极化矢量刚好是 $t\mathbf{w}_1+(1-t)\mathbf{w}_2$。实际上，假设将 $0\le t\le 1$ 看成可变参数，那方程 $t\mathbf{w}_1+(1-t)\mathbf{w}_2$ 描写的就是 $\mathbf{w}_1$ 与 $\mathbf{w}_2$ 的连线线段，这条线段当然依然在 3 维空间的单位球体之内，因此连线上的每一点当然都对应一个密度矩阵。

12.1.4 哥本哈根诠释和多世界诠释

在第二章的”量子比特与薛定谔猫”一节中，我们谈到测量会导致态的坍缩，这种说法其实只是对量子力学的一种诠释，而对于量子测量过程发生了什么，其实有多种诠释。现在，让我们更细致地考察一下量子测量过程，并介绍两种不同的主要诠释，哥本哈根诠释和多世界诠释。

大致来说，目前对量子测量过程的认识是下面这样的：简言之，测量过程就是一个被测对象和观察者（包括仪器）以及环境发生纠缠的过程。以测量一个电子的自旋 z 分量为例，这时候我们可以把一个任意的电子自旋态 $|\psi\rangle$ 分解成自旋向上态 $|\uparrow\rangle$ 和自旋向下态 $|\downarrow\rangle$ 的叠加，即 $|\psi\rangle=c_\uparrow|\uparrow\rangle+c_\downarrow|\downarrow\rangle$。假设测量之前电子处在 $|\psi\rangle$，观察者（包括仪器）处在 $|m_0\rangle$ 态，环境处在 $|e_0\rangle$ 态。那么测量之前整个世界的量子态 $|W_0\rangle$ 就是 $|W_0\rangle=|\psi\rangle|m_0\rangle|e_0\rangle=c_\uparrow|\uparrow\rangle|m_0\rangle|e_0\rangle+c_\downarrow|\downarrow\rangle|m_0\rangle|e_0\rangle$。

现在对电子的自旋 z 分量进行测量，当测到电子自旋向上时，我们记观察者处于 $|m_\uparrow\rangle$ 态，记这时候环境的状态为 $|e_\uparrow\rangle$，而当测到电子自旋向下时，记观察者处于 $|m_\downarrow\rangle$ 态，记这时候环境的状态为 $|e_\downarrow\rangle$。

而电子、观察者、环境这三者的整体也是一个量子系统，因此在测量过程中这个整体的量子态必然是保持线性叠加并幺正演化的，因此测量以后，世界的量子态将由测量前的 $|W_0\rangle$ 幺正演化到 $|W\rangle$，由于要保持量子态的线性叠加，$|W\rangle$ 必定可以写成 $|W\rangle=c_\uparrow|\uparrow\rangle|m_\uparrow\rangle|e_\uparrow\rangle+c_\downarrow|\downarrow\rangle|m_\downarrow\rangle|e_\downarrow\rangle$。很显然，与测量前的 $|W_0\rangle$ 态不同，在 $|W\rangle$ 态中，电子、观察者、以及环境这三者纠缠成了一个整体。

请注意，$|W\rangle$ 态依然是一个确定的量子态，因此这时候依然是不谈概率的。

但是，简化一点来说，在测量时我们关心的当然只是作为观测对象的电子，以及仪器上的读数和观察者看到某个读数时的心情，我们会完全忽视环境，实际上我们也根本无法收集到环境的信息。而由于电子、观察者、以及环境纠缠成了一个整体，因此当我们忽视环境的信息时，电子和观察者的量子力学描述就不再是一个纯态了，而是混态 $|c_\uparrow|^2|\uparrow,m_\uparrow\rangle\langle\uparrow,m_\uparrow|+|c_\downarrow|^2|\downarrow,m_\downarrow\rangle\langle\downarrow,m_\downarrow|$。也就是说，原来的相干叠加消失了，概率出现了。观察者有 $|c_\uparrow|^2$ 的概率处在 $|m_\uparrow\rangle$ 态并发现电子自旋向上，有 $|c_\downarrow|^2$ 的概率处在 $|m_\downarrow\rangle$ 态并发现电子自旋向下。这个过程就是我们前面说过的退相干。很显然，退相干是因为量子信息泄露到了环境中去而产生的。

到此为止都还没有涉及到量子态的坍缩，量子力学的不同诠释对上面的内容基本上还是有共识的。麻烦就出在，一旦我们在某次测量中观察到电子自旋向上，在我们看来整个世界就应该处在 $|W_\uparrow\rangle=|\uparrow\rangle|m_\uparrow\rangle|e_\uparrow\rangle$ 态，相反，如果在某次测量中我们观察到电子自旋向下，那在我们看来，这时候整个世界就应该处在 $|W_\downarrow\rangle=|\downarrow\rangle|m_\downarrow\rangle|e_\downarrow\rangle$ 态。

这时候就有不同的观点了，哥本哈根诠释的观点是，一次测量只发生了一个可能的结果，因此量子态坍缩了，当我们观察到电子自旋向上时，世界的量子态就从 $|W\rangle$ 坍缩到了 $|W_\uparrow\rangle$，而当我们观察到电子自旋向下时，世界的量子态就从 $|W\rangle$ 坍缩到了 $|W_\downarrow\rangle$。至于坍缩过程是如何进行的，目前物理学家们依然在研究之中。

这种诠释的问题在于，我们可以设想有一个外在于整个世界的“朋友”，他根本不关心我们的电子自旋是向上还是向下，他根本就没有进行观测，因此对于他来说，世界的量子态应该依然是 $|W\rangle$，根本就没有什么量子态的坍缩。

另一种不同的诠释通常人们叫做多世界诠释。这种诠释的观点是，一次测量，所有可能的结果都发生了，但是，世界本身分裂了，不同的可能结果发生在分裂以后的不同平行世界里。也就是说，在一个世界里你测到电子自旋向上，而在另一个世界里你测到电子自旋向下。但是，不同的平行世界只是同一个量子现实的不同分支，而整个量子现实依然是由相干叠加的 $|W\rangle=c_\uparrow|W_\uparrow\rangle+c_\downarrow|W_\downarrow\rangle$ 态来描写的。根据这种多世界诠释，从来就没有什么量子态坍缩，有的不过是世界分裂。

这种诠释的问题是，分裂以后的两个平行世界怎么能够保持相干叠加从而共同构成同一个量子现实呢?

12.2 再谈密度算符与系综

这一节我们主要是推广系综的概念。

密度算符集合的凸性

首先我们定义一下什么是向量空间的凸子集。对于向量空间的某个子集，如果它满足集合中任意两个点的连线段完全处于集合内部，那么这个集合就是向量空间的一个凸子集。比如图 (12.1) 所示的多边形就是两维平面的凸子集。如果凸子集中的某点永远不能处在子集内连线段内，而只能处于连线段的端点，那这个点就称为极端点。比如图 (12.1) 中的 A,B,C,D,E 点就是极端点。

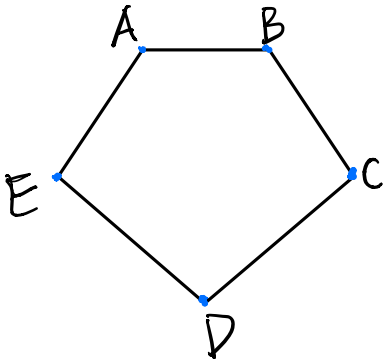

Figure 12.1: 两维实向量空间的一个凸子集，A,B,C,D,E 为极端点。

另外，由于任何两个厄密算符的实系数线性组合依然是厄密算符。所以一个量子系统所有厄密算符的集合构成了一个实向量空间，系统每一个可能的厄密算符都对应这个实向量空间里的一个向量。

之所以讨论以上两个概念，是因为，一个量子系统的所有可能密度算符的集合构成了厄密算符向量空间的一个凸子集。具体来说，假设 $\rho_1,\rho_2,...,\rho_N$ 是系统 N 个可能的密度算符，则下式给出的 ρ 也必定是系统可能的密度算符

$$\rho = p_1\rho_1 + p_2\rho_2 + ... + p_N\rho_N, \tag{12.34}$$

式中，$0 \le p_i \le 1$, 且 $\sum_{i=1}^{N} p_i = 1$。要证明 ρ 是密度算符，只需验证它满足密度算符必须满足的 3 条性质。根据所给条件，ρ 的厄密性以及 $\mathrm{Tr}(\rho) = 1$ 这两条都很明显，唯一需要说明的是 ρ 的正定性。对于一个厄密算符的正定性，我们有两种等价的判定方法，1. 所有本征值大于等于 0; 2. 这个算符在任何量子态 $|\psi\rangle$ 上的期望值都大于等于 0。根据第 2 种判定正定性的方法，(12.34) 式给出来的厄密算符 ρ 的正定性也是显然的 (注意我们已知每一个 ρ_i 都正定)。

上一段我们刻画了密度算符集合的凸性，实际上更符合凸子集定义的刻画方法是: 假设 ρ_1,ρ_2 为系统两个可能的密度算符，则这两者的连线段 $\rho(t) = t\rho_1 + (1-t)\rho_2$ 上每一个密度算符 $\rho(t)$ 也必为系统可能的密度算符，这里 $0 \le t \le 1$。这种刻画凸性的方式不仅是上一段给出来的刻画方式的特例，而且从它出发通过使用数学归纳法我们也能反过来导出上一段的刻画方式。所以对凸性的这两种刻画方式实际上是等价的。

不仅如此，我们还可以证明，纯态密度矩阵相应于这个密度算符凸子集的极端点。为此，我们假设 $\rho = |\psi\rangle\langle\psi|$, 我们用反证法，设存在密度算符 ρ_1 和 ρ_2，使得 $\rho = t\rho_1 + (1-t)\rho_2$, $0 < t < 1$。任意取一个与 $|\psi\rangle$ 正交的 $|\psi_\perp\rangle$ 态，它满足 $\langle\psi|\psi_\perp\rangle = 0$, 则有 $0 = \langle\psi_\perp|\rho|\psi_\perp\rangle = t\langle\psi_\perp|\rho_1|\psi_\perp\rangle + (1-t)\langle\psi_\perp|\rho_2|\psi_\perp\rangle$。进一步根据密度算符的正定性，必有 $\langle\psi_\perp|\rho_1|\psi_\perp\rangle = \langle\psi_\perp|\rho_2|\psi_\perp\rangle = 0$。但是 $|\psi_\perp\rangle$ 是任意与 $|\psi\rangle$ 正交的态，因此刚才的结果就意味着 $\rho_1 = \rho_2 = |\psi\rangle\langle\psi| = \rho$。这就证明了纯态必为极端点。反过来，对于混态密度算符 ρ，由于总有 $\rho = \sum_a p_a|a\rangle\langle a|$，式中 $|a\rangle\langle a| = \rho_a$ 均为纯态密度算符，这就说明混态密度算符必定不是极端点。这就完成了密度算符凸子集的极端点与纯态一一对应的证明。

比方说对前面讨论过的单量子比特密度算符 Bloch 球体的例子，整个 Bloch 球体就是一个凸子集，球体表面的每一个点都是这个凸子集的极端点，因此必定相应于纯态。这正好是我们前面已经知道的结论。不过，Bloch 球体是一个特例，它的边界点和极端点是一回事，但是，图 (12.1) 告诉我们，对于一般的凸子集，它的极端点必定是边界点，但是反过来，它的边界点却不一定是极端点。具有 3 维以上希尔伯特空间的量子系统的密度算符凸子集也是这样，它的边界点并不都是极端点。

系综概念的推广

根据凸性，假设某个密度算符 ρ 可以写成 $\rho=\sum_i p_i\rho_i$, $\sum_i p_i=1$。则很显然

$$\langle\mathscr{O}\rangle=\mathrm{Tr}(\rho\mathscr{O})=\sum_i p_i\mathrm{Tr}(\rho_i\mathscr{O})=\sum_i p_i\langle\mathscr{O}\rangle_i. \tag{12.35}$$

这就意味着，对于一个量子系统，假设我们能够制备出所有的 ρ_i 态，则我们就可以制备一个系综，让这个系综中的每个系统以 p_i 的概率处在 ρ_i 态，而这个系综就可以由密度算符 $\rho=\sum_i p_i\rho_i$ 描述。密度算符 ρ 既可以看成是描述一个系统的量子态，也可以看成是描述刚才所说的系综，等式 (12.35) 告诉我们，在物理上我们并不能区分这两者。

很明显，除了纯态以外，任何一个密度算符 ρ 都有多种办法分解成 $\rho=\sum_i p_i\rho_i$, $\sum_i p_i=1$。因此与密度算符 ρ 所对应的系综远不是唯一的，而是有无穷多的可能性。比方说，对于单量子比特 Bloch 球体内部的任何一点 (对应密度算符 ρ)，我们都可以作无数条过这一点的线段与球面上的两点 (对应于两个纯态密度算符) 相交，每一根这样的线段都意味着我们可以将密度算符 ρ 表示成球面上的这两个纯态密度算符的线性组合，而 ρ 就可以看成是描述了相应的纯态系综，由于这样的线段有无数条，所以 ρ 可以描述无数多个不同的纯态系综，在物理上这些纯态系综不可区分。比方说，对于 Bloch 球心的那个最大混态密度算符 $\rho=\frac{1}{2}\cdot 1$，我们就有无数种表达它的方式，比如下面的两种

$$\rho=\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|1\rangle\langle 1|=\frac{1}{2}|+\rangle\langle +|+\frac{1}{2}|-\rangle\langle -|. \tag{12.36}$$

12.3 量子熵

12.3.1 冯诺伊曼熵

前面我们说过，混态密度算符可以写成 $\rho=\sum_a p_a|a\rangle\langle a|$ 的形式，p_a 可以看成概率，是一种不确定性，它反映了我们对系统的某种无知。然而如何衡量这种无知呢?

一个自然的想法是将这种无知定义成某个算符的期望值，不妨称这个算符为“无知算符”，下面我们就是要找到它的表达式。由于我们的无知反映在系统的密度算符 ρ 中，所以这个“无知算符”应该依赖于 ρ(而通常的物理量算符是与态 ρ 无关的)。而且，如果 ρ 是一个纯态密度算符，那它就没有任何不确定性，从而相应的“无知算符”应该为零。

另外，从直观上我们知道，如果两个系统 S_1 和 S_2 相互独立，那我们对它们的无知应该是相加的，或者说相应的“无知算符”应该是相加的。另一方面，假设我们记 S_1 的密度算符为 ρ_{S_1}，记 S_2 的密度算符为 ρ_{S_2}，则由于这两个系统相互独立，所以两者整体的密度算符应该是 ρ_{S_1} 和 ρ_{S_2} 的乘积，不过，ρ_{S_1} 只作用在 S_1 的希尔伯特空间 $\mathscr{H}_{S_1}$ 上，ρ_{S_2} 只作

用在 S_2 的希尔伯特空间 $\mathscr{H}_{S_2}$ 上，为了强调这一点我们常将这个乘积写成如下形式

$$\rho_{S_1}\rho_{S_2}=\rho_{S_1}\otimes\rho_{S_2}, \tag{12.37}$$

它作用在两者整体的希尔伯特空间 $\mathscr{H}_{S_1}\otimes\mathscr{H}_{S_2}$ 上。为了将密度算符的这种相乘关系转化为相加关系，我们定义与密度算符 ρ 相应的“无知算符”如下

$$-\log(\rho). \tag{12.38}$$

额外加上的负号是为了使得这个算符为一个正定算符[2]。显然，这样定义的“无知算符”的确是相加的，满足 $-\log(\rho_{S_1}\rho_{S_2})=-\log(\rho_{S_1})-\log(\rho_{S_2})$。

很容易验证，对于纯态密度算符的确有 $-\log(\rho)=0$。这是因为，纯态密度算符满足 $\rho^2=\rho$，从而 ρ 的任何幂次都依然等于 ρ。从而对于任何一个幂级数函数 $f(x)$，必有 $f(\rho)=f(1)\rho$。现在我们按照 $-\log(x)=-\log(1+x-1)$ 关于 $x-1$ 的泰勒展开级数来定义算符函数 $-\log(\rho)$, 从而当然就有 $-\log(\rho)=-\log(1)\rho=0$。

下面我们可以将密度算符 ρ 中包含的无知定义为“无知算符”的期望值，称作冯诺伊曼熵，简称熵，记为 $S(\rho)$，则

$$S(\rho)=\langle-\log\rho\rangle=-\mathrm{Tr}\big(\rho\log\rho\big). \tag{12.39}$$

显然，$S(\rho)$ 也是所有 $-\rho\log\rho$ 本征值之和，从而也有

$$S(\rho)=-\sum_a p_a\log(p_a). \tag{12.40}$$

这正好是香农关于经典概率分布 $\{p_a\}$ 信息熵的表达式，经典信息熵衡量的是获取信息以后能够消除的不确定性，换言之，信息熵衡量的正是获取具体信息之前我们对它的无知，所以冯诺伊曼熵 $S(\rho)$ 的确衡量的是无知。很明显，

$$S(\rho)\geq 0, \tag{12.41}$$

等于号仅当 ρ 为纯态密度算符时才能取到。

另外，假设系统的希尔伯特空间为 D 维，即 (12.40) 式中的指标 $a=1,2,3,...,D$。则由 (12.40) 式可以证明，

$$S(\rho)\leq\log D. \tag{12.42}$$

等于号当且仅当 $p_a=1/D$ 时成立，这时候的密度算符为

$$\rho=\frac{1}{D}\sum_a|a\rangle\langle a|=\frac{1}{D}\cdot 1. \tag{12.43}$$

这当然是一个混态密度算符，相应的混态称之为最大混态。

[2]因为 ρ 的本征值 $0\leq p_a\leq 1$, 所以恒有 $-\log(\rho)$ 的本征值 $-\log(p_a)\geq 0$。

对于一个由 A 和 B 两部分组成的封闭系统，它的整体 (我们记作系统 AB) 必定处于一个纯态，我们可以对它进行施密特分解，进而能够导出前面的 (12.17) 式和 (12.18) 式，即 A 和 B 两个子部分的密度算符 ρ_A,ρ_B 分别为

$$\rho_A = \sum_a p_a |a\rangle_A \langle a|, \quad \rho_B = \sum_a p_a |a\rangle_B \langle a|. \tag{12.44}$$

从而必有

$$S(\rho_A) = -\sum_a p_a \log(p_a) = S(\rho_B). \tag{12.45}$$

即两个通过纠缠形成一个整体纯态的 A,B 系统，其两个子部分的冯诺伊曼熵是相等的。正因为这个美妙的性质，人们常常将这种情况下子部分的冯诺伊曼熵看成是对两体纠缠度的一种衡量。

假设 A，B 两部分的希尔伯特空间维数相等，均为 D。那前面关于最大混态的讨论告诉我们，两体纠缠度最大为 $\log D$，这时候 $p_a = \frac{1}{D}$, 从而根据施密特分解，AB 整体的量子态 $|\psi\rangle$ 为

$$|\psi\rangle = \frac{1}{\sqrt{D}} \sum_a |a\rangle_A |a\rangle_B. \tag{12.46}$$

由于纠缠度最大，这被称为最大纠缠态。

凹性

前面我们说过，一个系统所有可能密度算符的集合是一个凸集，因此冯诺伊曼熵 $S(\rho)$ 就是这个凸集上的函数，实际上，它是一个凹函数。为了证明这一点，任取两个不同的密度算符 ρ_1, ρ_2，然后构造 $\rho(t) = t\rho_1 + (1-t)\rho_2$, $0 \le t \le 1$，我们要证明的是 $\frac{d^2}{dt^2} S(\rho(t)) \le 0$。为此，我们首先注意到

$$\frac{d}{dt} S(\rho(t)) = -\mathrm{Tr}\big(\dot{\rho} \log \rho\big). \tag{12.47}$$

又注意到

$$\log \rho = \int_0^\infty ds \big(\frac{1}{s+1} - \frac{1}{s+\rho(t)} \big). \tag{12.48}$$

再注意到 $\ddot{\rho}(t) = 0$，则有

$$\frac{d^2}{dt^2} S(\rho(t)) = -\int_0^\infty ds \mathrm{Tr}\Big(\dot{\rho} \frac{1}{s+\rho(t)} \dot{\rho} \frac{1}{s+\rho(t)} \Big). \tag{12.49}$$

注意这个积分式的被积函数具有 $\mathrm{Tr}(A^2)$ 的形式，式中 $A = (s+\rho(t))^{-1/2} \dot{\rho} (s+\rho(t))^{-1/2}$ 为厄密算符，所以 A^2 必定是正定算符，从而 $\mathrm{Tr}(A^2) \ge 0$，从而即有 $\frac{d^2}{dt^2} S(\rho(t)) \le 0$。

由凹函数的熟知性质，我们有

$$\sum_i p_i S(\rho_i) \le S(\sum_i p_i \rho_i). \tag{12.50}$$

式中 $0 \le p_i \le 1$, 且 $\sum_i p_i = 1$。即是说，将多个密度算符 “混” 起来会增加冯诺伊曼熵。这个熵的增量通常称作 Holevo 信息，记作 χ

$$\chi = S(\sum_i p_i \rho_i) - \sum_i p_i S(\rho_i). \tag{12.51}$$

12.3.2 相对熵

相对熵和互信息

假设一个量子系统实际由密度算符 ρ 描写，但是我们不知道，我们猜测了一个描写它的密度算符 σ, 那我们这个猜测中额外包含的无知就应该是

$$S(\rho \parallel \sigma) = \langle -\log\sigma\rangle - \langle -\log\rho\rangle = \mathrm{Tr}(\rho\log\rho - \rho\log\sigma). \tag{12.52}$$

式中 $\langle -\log\sigma\rangle = -\mathrm{Tr}\rho\log\sigma$ 为我们的猜测中总共包含的无知，$\langle -\log\rho\rangle$ 为系统密度算符中原本包含的信息缺失，$S(\rho \parallel \sigma)$ 也称作 σ 和 ρ 的量子相对熵。可以想见，量子相对熵 $S(\rho \parallel \sigma)$ 这种额外的无知一定大于等于零 (等号当且仅当 $\sigma = \rho$ 时才成立)，即

$$S(\rho \parallel \sigma) \geq 0, \tag{12.53}$$

事实也的确如此，证明并不复杂，不过证明过程和我们将要进行的讨论关系不大，因此我们推荐感兴趣的读者去阅读文献[3]。

特别的，假设我们考虑一个由 A 和 B 两个子系统组成的复合系统 AB，记 AB 的密度算符为 ρ_{AB}, 记 $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ 为子系统 A 的约化密度算符，$\rho_B = \mathrm{Tr}_A(\rho_{AB})$ 为子系统 B 的约化密度算符。当然，一般来说，由于相互作用，A 和 B 之间会存在关联，因此 $\rho_{AB} \neq \rho_A \otimes \rho_B$。但是假设我们误以为 A 和 B 没有关联，并将 $\rho_A \otimes \rho_B$ 猜测为 AB 的密度算符 σ_{AB}

$$\sigma_{AB} = \rho_A \otimes \rho_B. \tag{12.54}$$

则我们这一猜测的额外无知为

$$\begin{aligned} S(\rho_{AB} \parallel \sigma_{AB}) &= \mathrm{Tr}_{AB}\left(\rho_{AB}\log\rho_{AB} - \rho_{AB}\log\sigma_{AB}\right) \\ &= \mathrm{Tr}_{AB}\left(\rho_{AB}\log\rho_{AB} - \rho_{AB}\log\rho_A - \rho_{AB}\log\rho_B\right) \\ &= S_A + S_B - S_{AB}. \end{aligned} \tag{12.55}$$

式中 $S_A = S(\rho_A)$ 为 A 的熵, $S_B = S(\rho_B)$ 为 B 的熵, 而 $S_{AB} = S(\rho_{AB})$ 为 AB 的熵。类似这样的记号我们后面还会用，将不再进行说明。人们通常称 $S_A + S_B - S_{AB}$ 为 A 和 B 的量子互信息，记为

$$I(A,B) = S_A + S_B - S_{AB}. \tag{12.56}$$

则 $S(\rho_{AB} \parallel \sigma_{AB}) \geq 0$ 就意味着量子互信息总是大于等于零的，即

$$I(A,B) = S_A + S_B - S_{AB} \geq 0. \tag{12.57}$$

不过，值得说明的是，量子互信息这个概念完全是类比于经典信息论中的互信息概念而引入的，但它其实并没有经典互信息的内涵。正因为如此我们这里才没有介绍经典互信息的具体含义。另外，不等式 (12.57) 有时候也称之为熵的次可加性 (subadditivity)。

[3] A Mini-Introduction To Information Theory, Edward Witten，arXiv:1805.11965[hep-th]

相对熵的单调性

令人吃惊的是，Lieb 和 Ruskai 还证明量子相对熵具有一种强大的单调性，不过这个证明很难，因此我们这里不会讨论，感兴趣的读者可以阅读原始论文[4], 或者查阅量子信息方面的专业书籍。

为了说清楚什么是量子相对熵的单调性，我们还是考察一个由 A,B 两部分组成的系统，记为系统 AB, 假设它有两个密度算符 ρ_{AB} 和 σ_{AB}, 记 $\rho_A = \mathrm{Tr}_B\rho_{AB}$, $\sigma_A = \mathrm{Tr}_B\sigma_{AB}$ 为相应的两个在 A 上的约化密度算符。量子相对熵的单调性说的是

$$S(\rho_{AB} \parallel \sigma_{AB}) \geq S(\rho_A \parallel \sigma_A). \tag{12.58}$$

即对一个系统的一部分求迹 (称作部分迹) 会降低量子相对熵。

不妨让我们给出对 (12.58) 式的一个直观理解 (不是证明！)，其实这个式子说的无非是，当将我们对一个大系统 (这里即 AB 系统) 密度算符的猜测用于它的一部分 (这里即是 A 部分) 时，我们的额外无知将单调减少。直观上这当然很好理解，难的是证明。

下面我们将量子相对熵的这种单调性用于一个三部分的系统 ABC, 设其密度算符为 ρ_{ABC}，当然我们有约化密度算符 $\rho_A = \mathrm{Tr}_{BC}\rho_{ABC}$, $\rho_{BC} = \mathrm{Tr}_A\rho_{ABC}$。我们再引入第二个密度算符

$$\sigma_{ABC} = \rho_A \otimes \rho_{BC}. \tag{12.59}$$

这样我们就可以计算 $S(\rho_{ABC} \parallel \sigma_{ABC})$ 了，

$$S(\rho_{ABC} \parallel \sigma_{ABC}) = S(\rho_{ABC} \parallel \rho_A \otimes \rho_{BC}) = I(A,BC) = S_A + S_{BC} - S_{ABC}. \tag{12.60}$$

为了应用单调性，现在我们对 C 求迹，从而有

$$\rho_{AB} = \mathrm{Tr}_C\rho_{ABC}, \quad \sigma_{AB} = \mathrm{Tr}_C\sigma_{ABC} = \rho_A \otimes \rho_B. \tag{12.61}$$

因此我们又可以计算求完迹以后的 $S(\rho_{AB} \parallel \sigma_{AB})$，

$$S(\rho_{AB} \parallel \sigma_{AB}) = S(\rho_{AB} \parallel \rho_A \otimes \rho_B) = I(A,B) = S_A + S_B - S_{AB}. \tag{12.62}$$

量子相对熵的单调性告诉我们

$$S(\rho_{ABC} \parallel \sigma_{ABC}) \geq S(\rho_{AB} \parallel \sigma_{AB}). \tag{12.63}$$

从而我们可以得到

$$S_{AB} + S_{BC} \geq S_B + S_{ABC}. \tag{12.64}$$

这个不等式称作熵的强次可加性 (strong subadditivity)。

[4]E. H. Lieb and M. B. Ruskai, "Proof Of The Strong Subadditivity Of Quantum Mechanical Entropy," J. Math. Phys. 14 (1973) 1938.

强次可加性有一些有趣的等价形式，比方说，假设我们对 ABC 系统引入一个辅助的 D 部分，使得密度算符 ρ_{ABC} 在这个 $ABCD$ 系统中被纯化，则利用我们前面证明过的美妙性质 (12.45)，我们将有 $S_{AB}=S_{CD}$(将 $ABCD$ 看成是由 AB 和 CD 两部分组成)，以及 $S_{ABC}=S_D$ (将 $ABCD$ 看成是由 ABC 和 D 两部分组成)，则不等式 (12.64) 可以重写成

$$S_{CD}+S_{BC}\geq S_B+S_D. \tag{12.65}$$

习惯上人们会引入 $S(C|D)=S_{CD}-S_D$，以及 $S(C|B)=S_{BC}-S_B$ 这两个量，称作条件熵，则上面这个不等式又可以重写成

$$S(C|D)+S(C|B)\geq 0. \tag{12.66}$$

这个式子告诉我们 $S(C|D)$ 和 $S(C|B)$ 这两个条件熵不能同时为负。

但是量子条件熵的确是可以取负值的，比方说，如果 AB 处于最大纠缠态 (当然是纯态)，则 $S_{AB}=0$, 而 S_B 达到最大，那这时候条件熵 $S(A|B)=S_{AB}-S_B$ 当然就为负！因此 (12.66) 就告诉我们，C 不能与 D 和 B 两者同时处于最大纠缠态，这称为量子纠缠的单配性 (monogamy of entanglement)。

12.4 开放系统以及热力学第二定律

12.4.1 开放系统的时间演化

我们知道，一个封闭量子系统随时间是幺正演化的，纯态 $|\psi\rangle$ 将演化为 $|\psi\rangle\to U|\psi\rangle$，式中 U 为幺正的时间演化算符。从而密度算符 ρ 将演化为

$$\rho\to U\rho U^\dagger. \tag{12.67}$$

幺正演化意味着概率和信息的守恒。但是，对于一个与环境相互作用的开放系统而言，其信息可以不断泄露到环境中去，从而其演化就不可能是幺正演化。那么，该怎么讨论开放系统的时间演化呢?

事实上，这时候我们可以把环境包括进来，将我们关心的系统看成 A 系统，而将环境看成 B 系统，两者共同构成一个封闭的 AB 系统。下面当我们单独说到系统时，指的都是我们关心的 A 系统，而对于 AB 整体，我们将称之为复合系统。

假设环境最初处于 $|e_0\rangle$ 态，而我们关心的系统最初处于 $|\psi\rangle$ 态，两者没有纠缠。由于整个复合系统是封闭的，所以复合系统将满足幺正演化，

$$|e_0\rangle|\psi\rangle\to U\big(|e_0\rangle|\psi\rangle\big). \tag{12.68}$$

下面我们取环境希尔伯特空间的一组正交归一矢量基 $\{|e_s\rangle\}$，然后将复合系统的量子态 $U\big(|e_0\rangle|\psi\rangle\big)$ 按这组基展开成

$$U\big(|e_0\rangle|\psi\rangle\big)=\sum_s|e_s\rangle|\psi_s\rangle. \tag{12.69}$$

注意，我们并没有说 $|\psi_s\rangle$ 正交归一。下面我们引入一组作用在 A 系统上的算符 E_s，定义为 $|\psi_s\rangle = E_s|\psi\rangle$，则

$$|e_0\rangle|\psi\rangle \to U\big(|e_0\rangle|\psi\rangle\big) = \sum_s |e_s\rangle E_s|\psi\rangle. \tag{12.70}$$

利用 U 的幺正性，我们可以得到

$$\sum_s E_s^\dagger E_s = 1. \tag{12.71}$$

所以开放系统从初态到 t 时刻态的映射完全由这组算符 E_s 决定。

将上面的推导照搬到密度算符上。假设系统初始的密度算符为 ρ，环境初始的密度算符为 $|e_0\rangle\langle e_0|$, 则整个复合系统初始的密度算符为这两者的乘积，不过，为了强调 ρ 只作用在系统的希尔伯特空间，而 $|e_0\rangle\langle e_0|$ 只作用在环境的希尔伯特空间，我们常常将这个乘积写成 $|e_0\rangle\langle e_0|\otimes\rho$，并简记为 $\widehat{\rho}$。则类似于上面的推导我们将有

$$\widehat{\rho} \to \widehat{\rho}' = U\widehat{\rho}U^\dagger = \sum_{s,s'} |e_s\rangle\langle e_{s'}| \otimes E_s\rho E_{s'}^\dagger. \tag{12.72}$$

由于我们并不关心环境，所以我们忽略它，即对环境的希尔伯特空间求迹，从而得到

$$\rho \to \mathrm{Tr}_B\big(\widehat{\rho}'\big) = \sum_s E_s\rho E_s^\dagger. \tag{12.73}$$

我们常常把这个方程写作

$$\rho \to \mathscr{E}(\rho) = \sum_s E_s\rho E_s^\dagger, \quad \sum_s E_s^\dagger E_s = 1. \tag{12.74}$$

式中定义的映射 $\mathscr{E}$ 称作量子映射 (quantum map)，可以验证它将把密度算符映射到密度算符，比方说很容易看出 $\mathrm{Tr}\big(\mathscr{E}(\rho)\big) = \mathrm{Tr}(\rho) = 1$。反过来可以证明，任何将密度算符映射到密度算符的线性映射，只要满足一些自然的条件，就必定可以表示成 (12.74) 的形式，这就是所谓的 Kraus 表示定理，(12.74) 式中的 E_s 就称作 Kraus 算符。(12.74) 式这样的量子映射有时候又称作保迹全正映射 (trace preserving, completely positive (TPCP) map)。很显然，封闭系统的时间演化只是 (12.74) 式的一个特例，这时候只有一个非零 Kraus 算符，即幺正演化算符 U。

但是，一般来说，(12.74) 式的量子映射只是把初始时的密度算符 ρ 映射到 t 时刻的密度算符 $\mathscr{E}(\rho)$，它却不能用于考察任意 t 时刻的密度算符如何继续随着时间演化，这是因为，在推导 (12.74) 式时我们有一个重要的假定，即最初的系统和环境没有关联，两者的密度算符是直接乘积的关系。但这个条件却不适用于任意 t 时刻，因为随着系统与环境的相互作用，一般来说 t 时刻系统和环境一定是纠缠和关联的。正因为如此，(12.74) 式其实不能真正看作开放系统的时间演化方程。

刘维尔方程和 Lindblad 方程

为了真正得到一个开放系统的时间演化方程，我们现在假定环境是巨大的，就好像统计物理中的恒温热源那样，以至于系统和环境的任何关联都会在某个非常短的微观时间尺

度 τ 之内丧失。假定我们关心的是系统在远大于微观时间 τ 的宏观时间尺度上的演化。这时候我们就可以推导出一个真正的时间演化方程。

根据上一段的假设，我们来考察密度算符在一个宏观无穷小时间间隔 δt 上的演化，当然，作为一个宏观时间，我们依然有 $\delta t \gg \tau$。按照假设，我们可以对任意初始时刻 t 利用 (12.74) 式，从而有相应的演化方程

$$\rho(t+\delta t) = \sum_s E_s(t+\delta t,t)\rho(t)E_s^\dagger(t+\delta t,t). \tag{12.75}$$

对于幺正演化的封闭系统 (它当然与环境没有关联)，我们只有一个非零 Kraus 算符，即为时间演化算符，记作

$$E_0(t+\delta t,t) = U(t+\delta t,t) = 1 - iH\delta t/\hbar + ... \tag{12.76}$$

代入 (12.75) 式，即得到标准的刘维尔方程

$$\hbar\frac{d\rho}{dt} = -i[H,\rho]. \tag{12.77}$$

对于与环境相互作用的开放系统，这时候就可能有许多非零 Kraus 算符。但是可以设想，对于一个无穷小的 δt，$E_0(t+\delta t,t)$ 将依然是主要的，按照 δt 进行泰勒展开依然有 $E_0(t+\delta t,t) = 1 + O(\delta t)$，不过当然，$E_0(t+\delta t,t)$ 不再幺正，从而我们应该假设

$$E_0(t+\delta t,t) = 1 - i\big(H - i\frac{\Gamma}{2}\big)\delta t/\hbar + ... \tag{12.78}$$

式中 Γ 也是一个厄密算符。E_0 的这一假设形式显然是受前面章节对不稳定态研究的启发。

其余 Kraus 算符都是次要的，我们假设它们 $E_m(t+\delta t,t) = O(\sqrt{\delta t})$，$m = 1,2,3...$。具体来说，我们假设

$$E_m(t+\delta t,t) = L_m\sqrt{\delta t}/\sqrt{\hbar} + ..., m = 1,2,3... \tag{12.79}$$

由于 Kraus 算符要满足 $\sum_s E_s^\dagger E_s = 1$，从而我们有

$$-\Gamma + \sum_{m=1} L_m^\dagger L_m = 0 \Rightarrow \Gamma = \sum_{m=1} L_m^\dagger L_m. \tag{12.80}$$

将 Kraus 算符的这些展开形式代入时间演化方程 (12.75) 就有

$$\hbar\frac{d\rho}{dt} = -i[H,\rho] - \frac{1}{2}\big(\Gamma\rho(t) + \rho(t)\Gamma\big) + \sum_{m=1} L_m\rho(t)L_m^\dagger. \tag{12.81}$$

这就是开放系统密度算符的时间演化方程，称作 Lindblad 方程，它可以看作是量子版本的 Fokker-Planck 方程，其中的算符 L_m 称作 Lindblad 算符，它描写系统与环境的相互作用。

12.4.2 热力学第二定律

前面冯诺伊曼熵和量子相对熵的公式最后都是一个求迹，因此它们在幺正变换下都将保持不变，以量子相对熵为例，在幺正变换 $\rho \to \rho' = U\rho U^\dagger$, $\sigma \to \sigma' = U\sigma U^\dagger$ 的作用下，将有

$$S(\rho' \parallel \sigma') = S(\rho \parallel \sigma). \tag{12.82}$$

特别的，一个封闭量子系统的时间演化就是一个幺正量子变换，因此，一个封闭量子系统的冯诺伊曼熵和量子相对熵在时间演化下都将保持不变。

上面结论好像和我们熟悉的热力学第二定律相矛盾，其实不然，因为冯诺伊曼熵是所谓的精细熵，而通常的热力学熵是一种粗粒化熵，孤立系统的精细熵保持不变，但是其粗粒化熵是增长的，这就是热力学第二定律。这一节我们就是要对与一个大的恒温热源密切接触的量子系统证明热力学第二定律。由于与恒温热源之间的能量和信息交换，这时候的量子系统当然不是孤立系统，甚至也不是封闭系统，在统计物理中我们知道，这时候的热力学第二定律其实表现为系统自由能随着时间的不断降低，这一节就是要证明这个结论。

由于我们要处理的不是封闭系统，所以需要用到量子映射。我们将要用到的关键结果就是，量子相对熵在量子映射下单调降低。具体来说，假设 ρ 和 σ 为系统的两个密度算符，在量子映射 $\mathscr{E}$ 之下，它们变换为

$$\rho \to \rho' = \sum_s E_s \rho E_s^\dagger, \quad \sigma \to \sigma' = \sum_s E_s \sigma E_s^\dagger. \tag{12.83}$$

则我们有

$$S(\rho \parallel \sigma) \geq S(\rho' \parallel \sigma'). \tag{12.84}$$

为了证明这一点，我们回顾一下前面对量子映射的推导过程，首先我们需要将密度算符与环境的 $|e_0\rangle\langle e_0|$ 乘起来，由于 $|e_0\rangle\langle e_0|$ 是一个纯态，熵是零，所以这一步对相对熵 $S(\rho \parallel \sigma)$ 没有影响，第二步是对整个复合系统用幺正演化，根据前面所说，这一步也不会改变相对熵，最后一步是对环境求部分迹，根据量子相对熵的单调性，这一步将使得量子相对熵单调减少，因此综合起来我们就有 (12.84) 式。

下面我们取 σ 为与恒温热源达到热平衡时的密度算符，根据统计物理知识，

$$\sigma = \frac{1}{Z} e^{-\beta H}, \tag{12.85}$$

式中 $\beta = 1/T$, T 为恒温热源的温度，式中配分函数 Z 也可以写

$$Z = e^{-\beta F(\sigma)}, \tag{12.86}$$

$F(\sigma)$ 为这个热平衡态的自由能。

假设 ρ 为这个与恒温热源密切接触的系统任意时刻的密度算符。我们可以计算量子

相对熵 $S(\rho \parallel \sigma)$，有

$$\begin{aligned} S(\rho \parallel \sigma) &= \langle -\log \sigma \rangle - S(\rho) \\ &= \langle \beta H \rangle - \beta F(\sigma) - S(\rho) \\ &= \beta E(\rho) - S(\rho) - \beta F(\sigma) \\ &= \beta \big[F(\rho) - F(\sigma) \big]. \end{aligned} \tag{12.87}$$

式中 $E(\rho) = \langle H \rangle = \mathrm{Tr}(\rho H)$，而 $F(\rho) = E(\rho) - TS(\rho)$ 为 ρ 态的自由能。

下面我们将与系统接触的恒温热源看成系统所处的环境，它的存在定义了一个量子映射 $\mathscr{E}$，由于热平衡态 σ 在这样的恒温环境下将保持不变，因此这个量子映射 $\mathscr{E}$ 将把 σ 映射到 σ 本身。因此由 (12.84) 式和 (12.87) 式，我们即有

$$\begin{aligned} & S(\rho \parallel \sigma) \geq S(\rho' \parallel \sigma) \\ \Rightarrow \quad & \beta \big[F(\rho) - F(\sigma) \big] \geq \beta \big[F(\rho') - F(\sigma) \big] \\ \Rightarrow \quad & F(\rho) \geq F(\rho'). \end{aligned} \tag{12.88}$$

也即是说，这个系统的自由能是随着时间递减的。这就完成了我们的证明。

13. * 狄拉克方程介绍

本章讲述狄拉克方程的最基本知识，包括狄拉克方程是如何引入的，以及狄拉克是如何提出正电子的。

之后我们着重讲述了如何通过取狄拉克方程的非相对论近似得出正确的电子自旋磁矩，以及导出正确的自旋轨道耦合项。这体现了复杂的物理效应如何可以从简洁的方程中推导出来，充分体现了狄拉克方程的威力。

我们知道电子有自旋，也许你常将自旋想象成一个小球绕着自身的转轴转动。这是自旋的经典图像，然而电子的自旋是相对论量子力学的，经典的自旋图像并不能真正解释电子的自旋，比方说，电子的自旋磁矩比按照经典图像算出来的磁矩大 2 倍。既然经典图像不对，那么，电子的自旋是从哪儿来的呢?

电子自旋的起源，以及自旋磁矩的那个神秘的 2 倍都曾经让物理学家伤透了脑筋。令人意想不到的是，纯粹从方程的数学之美出发，狄拉克得到了一个很简单的方程，从中自动就能得出自旋，尤其是能够得出那个神秘的 2。这个方程就是著名的狄拉克方程。本章是关于狄拉克方程的一个介绍，在本章中我们取 $\hbar = 1, c = 1$ 的自然单位制。本章我们使用求和约定，即对方程中重复出现的指标默认求和。

13.1 狄拉克方程

13.1.1 狄拉克方程的引入

让我们从单粒子的薛定谔方程开始，也就是

$$i\frac{\partial}{\partial t}\psi(\mathbf{x},t) = H\psi(\mathbf{x},t). \tag{13.1}$$

这个方程有两个最基本的特征，第一，这是一个线性偏微分方程，因此满足量子力学最基本的原理，态的线性叠加原理。第二，这个方程里只出现时间的一阶导数，这使得我们可以定义一个正定的守恒概率 $\int d^3\mathbf{x}|\psi(\mathbf{x},t)|^2$.

薛定谔在发表这个方程时取的是非相对论形式，也就是将哈密顿算符取成 $H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x})$, 其中 $\mathbf{p} = -i\nabla$ 是粒子的动量算符。哈密顿算符对应的当然是粒子的能量，然而，考虑到相对论以后，自由粒子的能量与动量之间的关系应该是

$$E^2 = \mathbf{p}^2 + m^2. \tag{13.2}$$

因此，考虑到相对论，我们就应该将自由粒子哈密顿算符取成 $H = \sqrt{\mathbf{p}^2 + m^2}$。这里涉及到算符 $\mathbf{p}^2 + m^2$ 的开方，一个可行的定义是利用泰勒展开，但是这一方面会使得方程变得丑陋，另一方面，在一般性的相对论情形下，算符 $\mathbf{p}^2$ 与 m^2 的大小可能相当，这样泰勒展开就没有良好的数学定义。

看起来绕过这个困难的唯一办法是，放弃开方，直接使用相对论关系 (13.2)，也就是说，将自由粒子薛定谔方程推广成 $-(\partial^2/\partial t^2)\psi(\mathbf{x},t) = (\mathbf{p}^2 + m^2)\psi(\mathbf{x},t)$。这样一来我们得到的就是一个关于时间的二阶微分方程。这就引起了新的问题，因为这样就破坏了原来薛定谔方程所满足的第二条特征，这使得我们再也不能定义一个守恒并且正定的概率。而一个正定的守恒概率，是整个量子力学理论的基石之一。

看起来往前往后都走不通，我们已经走投无路了!

然而，狄拉克意识到还有一条新的道路，一方面相对论当然要满足，但在另一方面方程应该仅仅只含时间的一阶导数这个特征也应该保持。由于在相对论中时间和空间可以相互转换，因此这两点结合起来的逻辑推论就是，我们应该让方程对于动量算符 (即对空间的导数) 也是一阶的。而要保持一阶就必须将下面的开方开出来

$$H = \sqrt{\mathbf{p}^2 + m^2}. \tag{13.3}$$

这怎么可能呢？然而狄拉克说这是可能的，只要你放弃对普通的数的坚持，转入到奇妙的反对易数世界里去，也就是像我们在讨论泡利算符时所做的那样。

现在，假定你已经将开方 (13.3) 开出来了，写成

$$H = \vec{\alpha}\cdot\mathbf{p} + \beta m, \tag{13.4}$$

这里的 $\vec{\alpha}$ 和 β 就像四元数里的 i,j,k 一样，是抽象的代数。为了满足相对论的关系式 (13.2), 我们再将 (13.4) 两边平方，就得到

$$\begin{aligned} E^2 &= \alpha_i\alpha_j p_i p_j + m(\alpha_i\beta + \beta\alpha_i)p_i + m^2\beta^2, \\ &= \frac{1}{2}(\alpha_i\alpha_j + \alpha_j\alpha_i)p_i p_j + m(\alpha_i\beta + \beta\alpha_i)p_i + m^2\beta^2. \end{aligned} \tag{13.5}$$

式中指标 $i,j = 1,2,3$。与 (13.2) 比较，我们发现 $\vec{\alpha}$，β 必须满足下面的抽象代数，

$$\alpha_i\alpha_j + \alpha_j\alpha_i = 2\delta_{ij} \tag{13.6}$$

$$\alpha_i\beta + \beta\alpha_i = 0,$$

$$\beta^2 = 1. \tag{13.7}$$

β 和 $\vec{\alpha}$ 的这种反对易代数很像四元数代数，两者同样奇妙，然而狄拉克得到的这个代数并不是四元数代数，而是数学家研究过的另一种称为 Clifford 代数的东西。现在，利用 $\vec{\alpha}$ 和 β 的 Clifford 代数，并注意到 (13.4), 我们就可以将自由电子的相对论波动方程写成，

$$i\partial_t\Psi = (-i\vec{\alpha}\cdot\vec{\partial} + \beta m)\Psi, \tag{13.8}$$

这就是著名的狄拉克方程，它所描述的粒子我们称为狄拉克粒子，电子就是一个狄拉克粒子，式中 $\vec{\partial} = \nabla$。由于相互不对易，式中的 $\vec{\alpha}, \beta$ 应该理解为与坐标和动量无关的算符，而波函数 $\Psi(\mathbf{x},t)$ 关于坐标变量是一个通常的波函数，但对于算符 $\vec{\alpha}, \beta$ 来说却应该理解成一个狄拉克右矢。

如果我们定义概率密度 $\rho(\mathbf{x},t) = \Psi^\dagger\Psi$, 定义概率流密度

$$\mathbf{j} = \Psi^\dagger\vec{\alpha}\Psi. \tag{13.9}$$

则利用 (13.8) 式容易验证下面的概率守恒方程

$$\frac{\partial\rho}{\partial t} + \vec{\partial}\cdot\mathbf{j} = 0. \tag{13.10}$$

13.1.2 矩阵表示

从 (13.4) 式容易看出，为了保证哈密顿算符的厄密性，$\vec{\alpha}$ 和 β 都应该是厄密算符，不妨令 $\alpha_4 = \beta$, 则对于任何 $a = 1,2,3,4$，应该有

$$\alpha_a^\dagger = \alpha_a. \tag{13.11}$$

另外，$\vec{\alpha}$ 和 β 的代数关系可以概括成

$$\alpha_a\alpha_b + \alpha_b\alpha_a = 2\delta_{ab}, \quad a,b = 1,2,3,4. \tag{13.12}$$

即这四个厄密算符两两反对易，且每个厄密算符的平方都等于 1。我们可以将这些厄密算符表示成矩阵，这一小节就是要具体解决什么样的矩阵能够表示它们?

首先，根据 (13.12) 的代数关系，这些厄密算符的平方都等于 1，因此它们只有 ± 1 的本征值，一般来说，每个本征值都可能有简并，我们首先要解决 $+1$ 本征值和 -1 本征值分别有几重简并。为此我们定义算符 $\alpha_{ij} = \alpha_i \alpha_j$, 根据 (13.12) 式，我们有

$$[\alpha_1, \alpha_{12}] = 2\alpha_2, \quad [\alpha_2, \alpha_{12}] = -2\alpha_1. \tag{13.13}$$

由此我们容易有 $\mathrm{Tr}(\alpha_1) = -\frac{1}{2}\mathrm{Tr}([\alpha_2, \alpha_{12}]) = 0$, $\mathrm{Tr}(\alpha_2) = \frac{1}{2}\mathrm{Tr}([\alpha_1, \alpha_{12}]) = 0$。类似的，利用 α_{34}，我们也可以证明 $\mathrm{Tr}(\alpha_3) = \mathrm{Tr}(\alpha_4) = 0$。概括起来即有

$$\mathrm{Tr}(\alpha_a) = 0, \quad a = 1,2,3,4. \tag{13.14}$$

因此每个 α_a 的 $+1$ 本征值的重数必定与 -1 本征值的重数相等，不妨设这个重数为 n，则 α_a 的矩阵表示必定为 $2n \times 2n$ 的厄密矩阵。

最简单的情形是 $n = 1$，即表示为 2×2 的厄密矩阵。我们知道，独立的 2×2 厄密矩阵刚好有四个，3 个泡利矩阵，再加上单位矩阵，3 个泡利矩阵当然两两反对易，但是它们都和第四个矩阵，即单位矩阵，对易。因此这就无法满足 4 个厄密矩阵两两反对易的要求。所以 $n = 1$ 不行。

实际上，能满足要求的最小 n 是 $n = 2$，相应的即是把 α_a 表示成 4×4 厄密矩阵。不妨让我们选取 $\beta = \alpha_4$ 对角的表象，取前两个分量对应 $+1$ 本征空间，后两个分量对应 -1 本征空间，即把 β 表示成

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{13.15}$$

式中 1 表示 2×2 的单位矩阵。由于 α_i, $i = 1,2,3$ 均与 β 反对易，所以它们的表示矩阵必定为如下形式

$$\alpha_i = \begin{pmatrix} 0 & v_i^\dagger \\ v_i & 0 \end{pmatrix}, \tag{13.16}$$

式中 v_i 均为 2×2 矩阵。由于 α_i 各自的平方等于 1，且两两反对易，我们容易有

$$v_i^\dagger v_j + v_j^\dagger v_i = v_i v_j^\dagger + v_j v_i^\dagger = 2\delta_{ij}. \tag{13.17}$$

很显然，只要取 $v_i = v_i^\dagger = \sigma_i$ 就能满足这些要求，这里 σ_i 表示 3 个泡利矩阵。所以我们有

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}. \tag{13.18}$$

式 (13.15) 和式 (13.18) 合起来就构成厄密算符 α_a 最简单的矩阵表示。在这个表示下自由粒子哈密顿算符 (13.4) 就是 4×4 的算符矩阵，

$$H = \begin{pmatrix} m & \vec{\sigma} \cdot \mathbf{p} \\ \vec{\sigma} \cdot \mathbf{p} & -m \end{pmatrix}. \tag{13.19}$$

而相应的波函数 $\Psi(\mathbf{x})$ 就应该表示成 4 分量的所谓旋量波函数。实际上，α_a 的更高 n 值的表示总可以约化成多个上述 4×4 矩阵的直和。所以我们只需考虑这种四分量表示就足够了。

13.1.3 反粒子

假定我们考虑一个动量确定的自由狄拉克粒子，即考察动量本征态，这时候动量算符 $\mathbf{p}$ 其实就取其本征值，我们依然记作 $\mathbf{p}$，所以现在 $\mathbf{p}$ 不再是算符了，而是普通的数。但是，$\vec{\alpha}, \beta$ 依然是算符，所以这时候狄拉克粒子的哈密顿量 (13.4) 依然是算符。而由于 $H^2=\mathbf{p}^2+m^2$ 为普通的数，所以，这样的动量本征态同时也是 H^2 的本征态。很显然，在这样的态空间上，$H=\vec{\alpha}\cdot\mathbf{p}+\beta m$ 可以取 $\pm\sqrt{\mathbf{p}^2+m^2}$ 这两个不同的本征值。注意，这也就是说，狄拉克粒子可以处于本征值为 $-\sqrt{\mathbf{p}^2+m^2}$ 这样的负能态。

所以狄拉克方程存在负能量态，它的能量本征值没有下界。通常来说，这是一个很大的问题，因为能量最低原理告诉我们，粒子总是更喜欢待在能量更低的态上。如果一个系统的能量本征值没有下界的话，那粒子就会持续不断地往更低的能态跃迁，这样的系统就不可能稳定存在。

为了解决这个问题，狄拉克注意到电子是一个费米子，它的每一个能态上只能占据一个电子，狄拉克想象所有的负能态都已经被电子占据满了，即是说真空其实是一片填满负能电子的海洋，称为狄拉克海。由于负能电子都被占满了，所以真空中的电子就不再能够向负能态跃迁了，它只能待在正能态上，这样就解决了稳定性问题。

奇妙的是，狄拉克进一步想象负能海中的某个电子被激发到了正能态。这时候真空中就产生了两种结果，一是多出了一个正能电子，二是负能海空出了一个未被占据的态，称之为空穴。负能海填满了电子当然带负电，但作为一种时空背景，我们无法感受到这个负能海的电荷。但是，如果负能海空出了一个态，那在我们的感受中，这个空穴就应该表现为正电荷！而且由于空穴是空出来的一个电子态，所以它在负能海中移动时其惯性应该和电子一样。即是说空穴也会表现出一个质量，并且和电子的质量一样。这就是说，负能海中的空穴在所有方面的表现都和电子一样，除了电荷相反，是正电荷。因此我们完全可以把负能海中的空穴当成是一种基本粒子，由于电荷相反，狄拉克称之为正电子。所以激发负能海的结果就是在真空中产生了一对粒子，电子和正电子。反过来，假设我们把正电子和电子放在一起，则由于正电子是空穴，所以这时候正能的电子就会向下跃迁到这个空穴上，留下一个填满的负能海真空，同时辐射出光子。我们称这个过程为电子和正电子的湮灭。

所以，电子和正电子可以成对产生，也可以成对湮灭！正电子除了电荷相反之外，其它表现都和电子相同，通常我们称之为电子的反粒子。正电子这种反粒子的存在是狄拉克的一个伟大预言，而它很快就在实验中观测到了！不仅电子有反粒子，实际上一切微观粒子都有反粒子，不过有些粒子，比如光子，由于完全中性，不带任何荷，所以它的反粒子其实就是它本身。不过，系统地处理反粒子的问题需要用到量子场论。

13.1.4 协变形式

狄拉克方程更常见的写法是这样的，我们定义 $\gamma^0=\beta$, $\gamma^i=\beta\alpha_i$, 它们组成一个四维时空的矢量 γ^μ，$\mu=0,1,2,3$，这样 (13.8) 就可以重写成

$$(i\gamma^\mu\partial_\mu-m)\Psi=0. \tag{13.20}$$

式中 $\partial_\mu=\frac{\partial}{\partial x^\mu}=(\partial_t,\vec{\partial})$, $x^\mu=(t,\mathbf{x})$ 为坐标四矢量。现在 γ^μ 满足如下 Clifford 代数，

$$\gamma^\mu\gamma^\nu+\gamma^\nu\gamma^\mu=2\eta^{\mu\nu}, \tag{13.21}$$

式中 $\eta^{\mu\nu}=\mathrm{diag}\{1,-1,-1,-1\}$ 为四维闵可夫斯基时空的度规张量，而任何形如 $a^\mu b_\mu$ 的式子其实都可以理解成 $\eta_{\mu\nu}a^\mu b^\nu$。很容易看出来，如果在方程 (13.20) 的左边乘上 $(i\gamma^\mu\partial_\mu+m)$，那么得到的就是克莱因-戈登方程 $(\partial^\mu\partial_\mu+m^2)\Psi=0$, 因此，狄拉克方程可以看成是克莱因-戈登方程的开方。

习惯上常常定义 $\overline{\Psi}=\Psi^\dagger\gamma^0$, 则很显然概率密度 $j^0=\rho$ 可以重写成

$$j^0=\Psi^\dagger\Psi=\overline{\Psi}\gamma^0\Psi. \tag{13.22}$$

而概率流密度则可以重写成

$$\mathbf{j}=\Psi^\dagger\vec{\alpha}\Psi=\overline{\Psi}\vec{\gamma}\Psi. \tag{13.23}$$

概率密度和概率流密度合起来构成一个洛伦兹四矢量 j^μ，

$$j^\mu=\overline{\Psi}\gamma^\mu\Psi. \tag{13.24}$$

而概率守恒方程则可以写成如下协变形式

$$\partial_\mu j^\mu=0. \tag{13.25}$$

注意，γ^0 是厄密的，而 γ^i 则反厄密，即 $\vec{\gamma}^\dagger=(\beta\vec{\alpha})^\dagger=\vec{\alpha}^\dagger\beta^\dagger=\vec{\alpha}\beta=-\beta\vec{\alpha}=-\vec{\gamma}$。关于这些 gamma 矩阵，除了上一段所用的表象之外，还有一个常用的表象是取

$$\gamma^0=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \vec{\gamma}=\begin{pmatrix}0&\vec{\sigma}\\-\vec{\sigma}&0\end{pmatrix}. \tag{13.26}$$

读者很容易验证这样的选取方式的确满足代数关系 (13.21)。为了看清楚这是一个什么表象，我们不妨计算 $\gamma^5=\gamma^0\gamma^1\gamma^2\gamma^3$，很容易得到

$$\gamma^5=\begin{pmatrix}i&0\\0&-i\end{pmatrix}, \tag{13.27}$$

式中 i 表示 2×2 的常数矩阵。很显然，这个表象是 γ^5 的对角表象。γ^5 是一个反厄密算符，其平方为 -1，$\pm i$ 是它的本征值。我们可以将四分量的旋量波函数按照 γ^5 的本征值进行分解，对应本征值 i 的旋量叫左手旋量，对应本征值 $-i$ 的旋量叫右手旋量。

13.2 从狄拉克方程到自旋磁矩以及 LS 耦合

这一节我们将通过取狄拉克方程的非相对论近似导出正确的电子自旋磁矩，以及正确的自旋轨道耦合。为此，请读者回想一下我们在第三章第 3.4 节中的有关讨论。在那里我们看到，从自旋 1/2 带电粒子的泡利哈密顿量出发我们能得到正确的电子自旋磁矩，但问题是，为什么有泡利哈密顿量呢？本节我们将看到，泡利哈密顿量其实是狄拉克方程的非相对论近似结果。不仅如此，我们还将看到，进一步研究这个非相对论近似还能导出正确的自旋轨道耦合。

13.2.1 有效哈密顿量方法回顾

为了讲清楚狄拉克方程的非相对论近似，我们先来回顾一下第五章第 5.2 节中讨论过的有效哈密顿量方法。根据 5.2 节的讨论，如果我们能将矩阵按照 1 和 2 两个子空间进行正交分解，则对于一个如下算符矩阵形式的哈密顿量 H，

$$H=\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}, \tag{13.28}$$

(式中子矩阵 $H_{11},H_{12},H_{21},H_{22}$ 同时也是算符)，我们可以将其本征矢量 Ψ 按照两个子空间分解成

$$\Psi=\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \tag{13.29}$$

通过消去子空间 2 里的分量 ψ_2，我们可以得到子空间 1 里的有效哈密顿量 $H_{\text{eff}}(E)$

$$H_{\text{eff}}(E)=H_{11}+H_{12}\frac{1}{E-H_{22}}H_{21}, \tag{13.30}$$

并将原来的本征方程 $H\Psi=E\Psi$ 转换成子空间 1 里的有效本征方程

$$H_{\text{eff}}(E)\psi_1=E\psi_1. \tag{13.31}$$

由于狄拉克方程中 β 乘在静止能量 m 上，它的本征值反映的是粒子的静止能量本征值，而非相对论极限下静止能量远大于动能。因此下面我们的基本思路就是取 β 的本征值为 1 的本征空间为子空间 1，本征值为 -1 的本征空间为子空间 2，然后将狄拉克方程的哈密顿算符分解成 (13.19) 这种形式，将四分量狄拉克旋量相应分解成如下形式

$$\Psi=\begin{pmatrix} u \\ v \end{pmatrix}, \tag{13.32}$$

u,v 均为两分量旋量波函数。然后再通过消去静止能量本征值为负的 $v(\mathbf{x})$ 分量推导出通常的正静止能量分量 $u(\mathbf{x})$ 所满足的有效本征方程。

13.2.2 泡利哈密顿量

为了得到电子磁矩，我们需要加上磁场，然后考察电子自旋与磁场的耦合。电子与磁场的耦合可以通过将动量算符 $\mathbf{p}$ 替换成力学动量算符 $\vec{\pi}=\mathbf{p}-q\mathbf{A}$ 来得到，这里 $\mathbf{A}$ 为电磁场矢量势，$q=-e$ 为电子电量。这样替换以后，相对论电子的哈密顿量 (13.19) 就变成了，

$$H=\begin{pmatrix} m & \vec{\sigma}\cdot\vec{\pi} \\ \vec{\sigma}\cdot\vec{\pi} & -m \end{pmatrix}. \tag{13.33}$$

按照 (13.30) 式，$\beta=1$ 的本征子空间里的有效哈密顿量 $H_{\text{eff}}(E)$ 为

$$H_{\text{eff}}(E)=m+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{E+m}. \tag{13.34}$$

相应的有效本征方程为

$$\left[m+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{E+m}\right]u(\mathbf{x})=Eu(\mathbf{x}). \tag{13.35}$$

在非相对论近似下，能量本征值的零级近似为静止能量 m, 记相应的非相对论修正为 ε, 将能量的零级近似 m 代入方程 (13.35) 左边, 则右边就可以给出非相对论修正 $E=m+\varepsilon$, 很显然

$$\left[\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}\right]u(\mathbf{x})=\varepsilon u(\mathbf{x}). \tag{13.36}$$

因此

$$\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}=\frac{[\vec{\sigma}\cdot(\mathbf{p}-q\mathbf{A})]^2}{2m} \tag{13.37}$$

就是电子在磁场中的非相对论哈密顿量。很显然，这正是泡利给出来的泡利哈密顿量，前面第三章 3.4 节中我们已经看到，这个哈密顿量能够给出正确的电子自旋磁矩。所以，对电子自旋磁矩那个 2 倍因子的正确解释最终需要用到狄拉克方程，它是狄拉克方程自动的一个结果。

13.2.3 LS 耦合

下面我们进一步假设有一个电势 $\phi(\mathbf{x})$ 作用在电子上，则相对论电子的哈密顿量就应该写成

$$H=\begin{pmatrix} m & \vec{\sigma}\cdot\vec{\pi} \\ \vec{\sigma}\cdot\vec{\pi} & -m \end{pmatrix}+q\phi(\mathbf{x})\cdot 1=\begin{pmatrix} m+q\phi & \vec{\sigma}\cdot\vec{\pi} \\ \vec{\sigma}\cdot\vec{\pi} & -m+q\phi \end{pmatrix}. \tag{13.38}$$

式中 $q\phi$ 为电势能。则 $\beta=1$ 的本征子空间里的有效哈密顿量 $H_{\text{eff}}(E)$ 现在就应该为

$$H_{\text{eff}}(E)=m+(\vec{\sigma}\cdot\vec{\pi})\frac{1}{E+m-q\phi}(\vec{\sigma}\cdot\vec{\pi}). \tag{13.39}$$

相应的有效本征方程就是

$$\left[m+(\vec{\sigma}\cdot\vec{\pi})\frac{1}{E+m-q\phi}(\vec{\sigma}\cdot\vec{\pi})\right]u(\mathbf{x})=Eu(\mathbf{x}). \tag{13.40}$$

为了取非相对论近似，令 $E=m+\varepsilon$, 对静止能量的修正 ε 来自于动量 $\vec{\pi}$ 和电势能 $q\phi$。注意到

$$\frac{1}{E+m-q\phi}=\frac{1}{2m(1+\frac{\varepsilon-q\phi}{2m})}\simeq\frac{1}{2m}(1-\frac{\varepsilon-q\phi}{2m}), \tag{13.41}$$

代入上面的 (13.40) 式，并将所有的 ε 项移到方程右边，则有

$$\Big[q\phi+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{q}{4m^2}(\vec{\sigma}\cdot\vec{\pi})\phi(\vec{\sigma}\cdot\vec{\pi})\Big]u(\mathbf{x})=\varepsilon\big[1+\frac{1}{4m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big]u(\mathbf{x}). \tag{13.42}$$

我们想把修正能量 ε 看成是某个厄密哈密顿量 H' 的本征值，但是上面这个方程显然不具有这种定态薛定谔方程的形式，为了将之改写成这种形式，我们重新定义一个两分量旋量波函数 $\psi(\mathbf{x})$,

$$\psi(\mathbf{x})=\big[1+\frac{1}{8m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big]u(\mathbf{x}). \tag{13.43}$$

并利用 (忽略 $\vec{\pi}/m$ 的四次方阶)

$$\begin{aligned}&\big[1+\frac{1}{4m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big]\simeq\big[1+\frac{1}{8m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big]^2\\&\big[1-\frac{1}{8m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big]\big[1+\frac{1}{8m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big]\simeq 1.\end{aligned} \tag{13.44}$$

则我们就能把方程 (13.42) 改写成标准的

$$H'\psi(\mathbf{x})=\varepsilon\psi(\mathbf{x}), \tag{13.45}$$

式中厄密哈密顿量 H' 为

$$H'\simeq\big[1-\frac{1}{8m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big]\Big[q\phi+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{q}{4m^2}(\vec{\sigma}\cdot\vec{\pi})\phi(\vec{\sigma}\cdot\vec{\pi})\Big]\big[1-\frac{1}{8m^2}(\vec{\sigma}\cdot\vec{\pi})^2\big].$$

展开上面 H' 的表达式就能得到

$$H'=q\phi+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{q}{8m^2}(\vec{\sigma}\cdot\vec{\pi})\big[\phi,\vec{\sigma}\cdot\vec{\pi}\big]+\frac{q}{8m^2}\big[\vec{\sigma}\cdot\vec{\pi},\phi\big](\vec{\sigma}\cdot\vec{\pi})+.... \tag{13.46}$$

省略号表示 $\vec{\pi}/m$ 的四次方项和更高次项。利用

$$[\vec{\pi},\phi]=[\mathbf{p},\phi]=-i\nabla\phi, \tag{13.47}$$

可以进一步得到

$$H'=q\phi+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{iq}{8m^2}\Big[(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\nabla\phi)-(\vec{\sigma}\cdot\nabla\phi)(\vec{\sigma}\cdot\vec{\pi})\Big]+... \tag{13.48}$$

利用恒等式

$$(\vec{\sigma}\cdot\vec{F})(\vec{\sigma}\cdot\vec{G})=(\vec{F}\cdot\vec{G})+i\vec{\sigma}\cdot(\vec{F}\times\vec{G}), \tag{13.49}$$

又可以得到 (注意 $(\nabla\phi)\times\vec{\pi}=-\vec{\pi}\times(\nabla\phi)$)

$$\begin{aligned}H' &= q\phi+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{iq}{8m^2}\Big[\vec{\pi}\cdot(\nabla\phi)-(\nabla\phi)\cdot\vec{\pi}\Big]-\frac{q}{4m^2}\big[\vec{\pi}\times(\nabla\phi)\big]\cdot\vec{\sigma}+...\\ &= q\phi+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{q}{8m^2}\nabla^2\phi-\frac{q}{4m^2}\big[\vec{\pi}\times(\nabla\phi)\big]\cdot\vec{\sigma}+...\end{aligned}\tag{13.50}$$

或者也可以记 $q\phi(\mathbf{x})=V(\mathbf{x})$ 为电势能，记 $\vec{\sigma}/2=\mathbf{S}$ 为自旋算符，进而将 H' 写成

$$H'=V(\mathbf{x})+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{1}{2m^2}\big[(\nabla V)\times\vec{\pi}\big]\cdot\mathbf{S}++\frac{1}{8m^2}\nabla^2V+...\tag{13.51}$$

通过量纲分析恢复常数 $\hbar$ 和 c，则 (13.51) 就应该是

$$H'=V(\mathbf{x})+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}+\frac{1}{2m^2c^2}\big[(\nabla V)\times\vec{\pi}\big]\cdot\mathbf{S}++\frac{\hbar^2}{8m^2c^2}\nabla^2V+...\tag{13.52}$$

相比于泡利哈密顿量，(13.52) 右边多出两项，其中 $\frac{\hbar^2}{8m^2c^2}\nabla^2V$ 是所谓的达尔文振颤项，它来自于电子不是经典粒子，它的位置会在康普顿波长 $\hbar/(mc)$ 的范围内随机振颤，达尔文项就是这种振颤导致的能量修正。而 (13.52) 右边的

$$\frac{1}{2m^2c^2}\big[(\nabla V)\times\vec{\pi}\big]\cdot\mathbf{S}\tag{13.53}$$

描写的就是自旋轨道耦合。特别的，假设忽略磁场，从而 $\vec{\pi}=\mathbf{p}$，并且考察中心力场, 从而 V 仅仅依赖于径向半径 r, 即有 $\nabla V=\mathbf{e}_r\frac{\partial V}{\partial r}=\frac{\mathbf{x}}{r}\frac{\partial V}{\partial r}$, 则这一自旋轨道耦合项就可以重写成

$$\frac{1}{2m^2c^2}\big[\frac{\mathbf{x}}{r}\frac{\partial V}{\partial r}\times\mathbf{p}\big]\cdot\mathbf{S}=\frac{1}{2m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}\mathbf{L}\cdot\mathbf{S},\tag{13.54}$$

式中 $\mathbf{L}=\mathbf{x}\times\mathbf{p}$ 是电子的轨道角动量算符。对于氢原子 $V(r)=-\frac{e^2}{4\pi\varepsilon_0 r}$, 从而容易得到相应的自旋轨道耦合项为

$$\frac{e^2}{8\pi\varepsilon_0}\frac{1}{m^2c^2r^3}\mathbf{L}\cdot\mathbf{S}.\tag{13.55}$$

之所以有这样的自旋轨道耦合项，是因为从电子的角度来看，是氢原子核在围绕着它转动，这种转动会在电子位置产生一个磁场，这个磁场正比于轨道角动量 $\mathbf{L}$, 电子的自旋磁矩会和这个磁场耦合，这就产生了自旋轨道耦合，简言之，这一项来源于电子自旋磁矩与它感受到的磁场的耦合。但是，想从这一物理图像得到系数完全正确的自旋轨道耦合项，人们必须考虑到一个相对论效应，即所谓的 Thomas 进动，否则结果会大了两倍。狄拉克方程由于本身是相对论的，将它取非相对论近似时，Thomas 进动的相对论效应自动就包含进来了，所以我们上面才能导出正确的自旋轨道耦合。



14. 附录: 线性代数概要

这个数学附录的内容是为了帮助初学者理解线性代数的基本概念和思维方式，并不是要给出完备的严格数学证明。因此下面讲述的很多内容在数学严格性上肯定是有待补充的。

14.1 向量空间

通常的三维空间 $\mathbb{R}^3$ 由如下有序三元实数组构成

$$\mathbb{R}^3 = \{[x,y,z] : x,y,z \in \mathbb{R}\}.$$

我们可以记 $\mathbf{x} = [x,y,z]$，称之为 $\mathbb{R}^3$ 空间的向量，任取 $\mathbf{x}_1 = [x_1,y_1,z_1]$, $\mathbf{x}_2 = [x_2,y_2,z_2]$，我们可以定义向量的加法，$\mathbf{x}_1 + \mathbf{x}_2 = [x_1+x_2, y_1+y_2, z_1+z_2]$。显然，这样定义的向量加法满足加法交换律和结合律。同时我们还可以定义标量乘法 $\lambda\mathbf{x} = [\lambda x, \lambda y, \lambda z], \lambda \in \mathbb{R}$。显然，它满足分配律，即 $\lambda(\mathbf{x}_1 + \mathbf{x}_2) = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2$。

类似的代数结构我们也可以在 n 维复空间 $\mathbb{C}^n$ 中发现，$\mathbb{C}^n$ 由如下有序 n 元复数组构成

$$\mathbb{C}^n = \{[x_1,x_2,...,x_n] : x_1,x_2,...,x_n \in \mathbb{C}\}.$$

记 $\mathbf{x} = [x_1,x_2,...,x_n]$, $\mathbf{y} = [y_1,y_2,...,y_n]$, 称为 $\mathbb{C}^n$ 中的两个向量。类似地可以定义向量加法，$\mathbf{x}+\mathbf{y} = [x_1+y_1, x_2+y_2, ..., x_n+y_n]$, 很显然，它同样满足交换律和结合律。同样可以定义标

量乘法，$\lambda\mathbf{x}=[\lambda x_1,\lambda x_2,...,\lambda x_n]$，不过现在标量 $\lambda\in\mathbb{C}$。很明显，这样的标量乘法也满足分配律。

利用向量的加法和标量乘法，我们很容易将 k 个向量线性组合起来，得到一个新的向量，

$$\mathbf{x}=c_1\mathbf{x}_1+c_2\mathbf{x}_2+...+c_k\mathbf{x}_k.$$

式中向量 $\mathbf{x}$ 就是 $\mathbf{x}_1$,$\mathbf{x}_2$,...,$\mathbf{x}_k$ 的线性组合。可以进行任意的线性组合就是向量的核心特征。反过来，具有这个特征的东西都叫向量。

比方说我们考察 $[0,1]$ 区间上的连续复数值函数，我们把所有这些函数的集合记为 $C_{[0,1]}$。任给 $f(x),g(x)\in C_{[0,1]}$，很显然我们可以定义一个新的 $C_{[0,1]}$ 集合内的函数 $(f+g)(x)$，$(f+g)(x)=f(x)+g(x)$。这就给 $C_{[0,1]}$ 集合内的函数定义了加法，它显然满足交换律和结合律。我们也可以定义标量乘法，$(\lambda f)(x)=\lambda f(x)$，它显然满足分配律。有了这两种运算，我们就可以把 $C_{[0,1]}$ 内的函数任意线性组合，比方说 $c_1f_1(x)+c_2f_2(x)+...+c_kf_k(x)$ 显然还是 $C_{[0,1]}$ 内的函数，$c_1,c_2,...,c_k$ 就是组合系数，按照标量乘法的定义，它们当然都是普通的复数。可见，我们完全可以把 $C_{[0,1]}$ 内的函数称作向量!

把这些要素抽象出来，我们就可以给出向量空间的一般性定义：*向量空间 V 是一个集合，它里面的元素 (我们称作向量) 可以是任何东西，只要这些东西之间定义了加法和标量乘法运算，并且这些运算满足交换律、结合律以及分配律。简言之，只要这些东西可以进行任意的线性组合，组合系数是普通的数。当然，还有很重要的一条，无论如何线性组合，结果依然得是集合 V 中的一个向量，也即是说，集合 V 在加法和标量乘法运算下得具有封闭性。*

向量和向量空间是抽象的。然而读者通常都能按照 3 维空间里面的我们从高中就已经熟悉的普通向量来想象它们，因为不管向量空间多么抽象，它和 3 维向量空间都有同样的本质，那就是里面的向量可以任意线性组合。所以，通常人们也把**向量空间称作线性空间**。

既然向量是抽象的,我们最好给它们发明一个通用的抽象的记号。我们把向量记作 $|\psi\rangle$，读作向量 ψ。比方说在前面的 $\mathbb{R}^3$ 空间情形，$|\psi\rangle$ 就是三元组 $[x,y,z]$，所以 $|\psi\rangle=[x,y,z]$。在 $\mathbb{C}^n$ 情形，$|\psi\rangle=[x_1,x_2,...,x_n]$。而在 $C_{[0,1]}$ 情形，$|\psi\rangle=f(x)$。

在量子力学中我们关心的都是组合系数为复数的向量空间，称为复向量空间。根据向量空间的定义，对于任何复向量空间 V，假如向量 $|\psi_1\rangle,|\psi_2\rangle,...,|\psi_k\rangle\in V$，则它们的任意复系数线性组合也必属于 V，即

$$c_1|\psi_1\rangle+c_2|\psi_2\rangle+...+c_k|\psi_k\rangle\in V.$$

子空间与直和

如果向量空间 V 的一个子集 U 本身也是向量空间，即 U 本身就在加法和标量乘法下保持封闭，则 U 就称为 V 的一个**子空间**。比如很容易验证 $\{[x_1,x_2,0]:x_1,x_2\in\mathbb{C}\}$ 就是 $\mathbb{C}^3$ 的一个子空间。同样，通常 3 维空间 $\mathbb{R}^3$ 中任何一个过原点的平面也构成 $\mathbb{R}^3$ 的一个子空间。实际上，人们通常都可以把向量空间的子空间想象成 “一个过原点的超平面”。

设 U 和 W 均是向量空间 V 的子空间，并且如果对于 $|u\rangle \in U$, $|w\rangle \in W$，当且仅当 $|u\rangle = |w\rangle = 0$ 时才满足 $|u\rangle + |w\rangle = 0$。则我们就可以定义 U 和 W 的**直和**, 记为 $U \oplus W$，它也是 V 的子空间，$U \oplus W$ 的要点在于其里面的任何向量 $|\psi\rangle$ 均可以唯一分解成如下形式

$$U \oplus W = \{|\psi\rangle = |u\rangle + |w\rangle : |u\rangle \in U, |w\rangle \in W\}.$$

这种分解的唯一性很容易证明，设另有 $|u'\rangle \in U$, $|w'\rangle \in W$ 也满足 $|\psi\rangle = |u'\rangle + |w'\rangle$, 则将 $|\psi\rangle$ 的这两种分解方式相减就可以得到 $(|u'\rangle - |u\rangle) + (|w'\rangle - |w\rangle) = 0$，根据前面的假设必有 $|u'\rangle - |u\rangle = 0$, $|w'\rangle - |w\rangle = 0$，这就证明了分解的唯一性。

举个例子, 假设向量空间 $\mathbb{R}^3$ 的两个子空间 U 和 W 的定义如下, $U = \{[x,y,0] : x,y \in \mathbb{R}\}$, $W = \{[0,0,z] : z \in \mathbb{R}\}$, 则 $U \oplus W = \mathbb{R}^3$。

直和的概念当然可以推广。设 $U_1, U_2, ..., U_m$ 均是向量空间 V 的子空间，并且对于 $|u_1\rangle \in U_1, ..., |u_m\rangle \in U_m$, 当且仅当 $|u_1\rangle = |u_2\rangle = ... = |u_m\rangle = 0$ 时才有 $|u_1\rangle + |u_2\rangle + ... + |u_m\rangle = 0$。那么我们就可以定义这 m 个子空间的直和, 记为 $U_1 \oplus U_2 \oplus ... \oplus U_m$。具体来说, $U_1 \oplus ... \oplus U_m$ 也是 V 的一个子空间，它里面的向量可以唯一分解成如下形式

$$U_1 \oplus ... \oplus U_m = \{|u_1\rangle + |u_2\rangle + ... + |u_m\rangle : |u_1\rangle \in U_1, ..., |u_m\rangle \in U_m\}.$$

举个例子。设 U_j 是 $\mathbb{C}^n$ 中除第 j 个坐标以外其余坐标全是 0 的那些向量所组成的子空间，则 $U_1 \oplus U_2 \oplus ... \oplus U_n = \mathbb{C}^n$。

14.2 有限维向量空间

张成空间与线性无关

复向量空间 V 中的一组向量 $|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$ 的所有可能线性组合所构成的集合称为 $|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$ 的张成空间，记为 $\text{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle)$, 即

$$\text{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle) = \{c_1|v_1\rangle + c_2|v_2\rangle + ... + c_m|v_m\rangle : c_1, c_2, ..., c_m \in \mathbb{C}\}.$$

可以证明 $\text{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle)$ 是包含 $|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$ 的最小子空间。

有限维向量空间就是可以由向量空间中一组有限数目的向量张成的向量空间。除了有限维向量空间之外的向量空间都叫无限维的。虽然量子力学中的向量空间常常为无限维，但其基本数学结构都可以从有限维的情况得到理解。因此，我们这个附录将主要限于讨论有限维向量空间。而且，我们将主要讨论有限维复向量空间。

举个例子。很明显, 由 n 个向量组成的向量组 $|e_1\rangle = [1,0,...,0], |e_2\rangle = [0,1,0,...,0], ..., |e_n\rangle = [0,...,0,1]$ 张成了向量空间 $\mathbb{C}^n$。

再举个例子, 任何两个不超过 n 次的复多项式的和依然是一个不超过 n 次的复多项式，所以所有不超过 n 次的复多项式 $p(z)$ 的集合构成了一个复向量空间, 定义这个向量空间的向量 $|p\rangle = p(z)$。很显然, 这个向量空间可以由如下向量组张成, $|0\rangle = 1, |1\rangle = z^1, ..., |n\rangle = z^n$。这是因为，任何不超过 n 次的复多项式都可以写成 $p(z) = c_0 + c_1 z + c_2 z^2 + ... + c_n z^n$，用向量的形式来写即是 $|p\rangle = c_0|0\rangle + c_1|1\rangle + c_2|2\rangle + ... + c_n|n\rangle$。

对于一组向量 $|v_1\rangle,|v_2\rangle,...,|v_m\rangle$，如果 $c_1|v_1\rangle+c_2|v_2\rangle+...+c_m|v_m\rangle=0$, 当且仅当所有的系数都为 0, 即 $c_1=c_2=...=c_m=0$, 则我们就称 $|v_1\rangle,|v_2\rangle,...,|v_m\rangle$ **线性无关**。否则就称它们线性相关。容易证明，如果 $|v_1\rangle,|v_2\rangle,...,|v_m\rangle$ 线性无关，则任何 $|u\rangle\in\text{Span}(|v_1\rangle,|v_2\rangle,...,|v_m\rangle)$ 的线性组合形式 $|u\rangle=a_1|v_1\rangle+a_2|v_2\rangle+...+a_m|v_m\rangle$ 都必定是**唯**一的。因为否则，假设 $|u\rangle$ 另有一种不同的线性组合 $|u\rangle=b_1|v_1\rangle+b_2|v_2\rangle+...+b_m|v_m\rangle$, 则将 $|u\rangle$ 的两种线性组合相减就会得到 $(a_1-b_1)|v_1\rangle+(a_2-b_2)|v_2\rangle+...+(a_m-b_m)|v_m\rangle=0$, 其中组合系数不全为 0，但这就和 $|v_1\rangle,|v_2\rangle,...,|v_m\rangle$ 线性无关的假设矛盾了。

如果 $|v_1\rangle,|v_2\rangle,...,|v_m\rangle$ 线性相关,那这就说明对于张成子空间 $W=\text{Span}(|v_1\rangle,|v_2\rangle,...,|v_m\rangle)$ 来说，向量组 $|v_1\rangle,|v_2\rangle,...,|v_m\rangle$ 有冗余, 我们完全可以从中删去某些向量，使得剩余的向量依然能够张成 W。比方说，向量组 $|e_1\rangle=[1,0],|e_2\rangle=[0,1],|e_3\rangle=[1,1]$ 张成了向量空间 $\mathbb{C}^2$, 但是由于 $|e_1\rangle+|e_2\rangle-|e_3\rangle=0$，因此这个向量组是线性相关的，我们完全可以从中删除向量 $|e_3\rangle$，剩下的 $|e_1\rangle,|e_2\rangle$ 依然张成了向量空间 $\mathbb{C}^2$。

所以对于张成一个固定的子空间来说，线性无关向量组是最小的向量组。由此我们可以得到下面的概念。

基和维数

向量空间 V 的基是任何一组既线性无关又能张成 V 的向量, 基中的每一个向量叫作基向量。比方说, 对于所有不超过 n 次的复多项式所成的向量空间, $|0\rangle=1,|1\rangle=z^1,...,|n\rangle=z^n$ 就是它的一组基。再比方说, 对于向量空间 $\mathbb{C}^n$, $|e_1\rangle=[1,0,...,0],|e_2\rangle=[0,1,0,...,0],...,|e_n\rangle=[0,...,0,1]$ 就是它的一组基。

根据基的定义我们知道，给定向量空间的一组基，则空间中任何向量都可以写成这组基的线性组合形式，并且组合系数是唯一的。在有限维向量空间 V 上选定一组基 $|e_1\rangle,|e_2\rangle,...,|e_n\rangle$ 以后，V 中的任意向量 $|v\rangle$ 就可以唯一性地用它的组合系数来表示。具体来说，我们可以唯一性地将 $|v\rangle$ 写成，

$$|v\rangle=v_1|e_1\rangle+v_2|e_2\rangle...+v_n|e_n\rangle, \tag{14.1}$$

式中 $v_1,v_2,...,v_n\in\mathbb{C}$ 为复组合系数。进一步将 (14.1) 式重写成

$$|v\rangle=(|e_1\rangle,|e_2\rangle,...,|e_n\rangle)\begin{pmatrix}v_1\\v_2\\...\\v_n\end{pmatrix}, \tag{14.2}$$

等式右边的表达式表示将行与列的各个分量对应地乘起来再求和，结果刚好由 (14.1) 式给出。从 (14.2) 式我们很容易看出，向量 $|v\rangle$ 可以唯一性地表示成列矢量 $v=\begin{pmatrix}v_1\\v_2\\...\\v_n\end{pmatrix}$。人们也很容易验证，向量的加法就刚好对应相应列矢量的加法，向量的标量乘法就对应列矢量的标量乘法。所以任何向量空间在取定基以后都可以表示成列矢量空间。正是在这个意义上，我们完全可以将向量 $|v\rangle$ 想象成一个列矢量。

但是，一个向量空间的基本身不是唯一的。比方说，对于向量空间 $\mathbb{C}^n$。$|e_1\rangle=[1,0,...,0],|e_2\rangle=[0,1,0,...,0],...,|e_n\rangle=[0,...,0,1]$ 是它的一组基，但是将这组基中的第 n 个向量 $|e_n\rangle$ 替换成 $|e_n'\rangle=[0,...,0,1,1]$ 结果依然是一组基。

然而由于基是最小的张成向量组，所以对于有限维向量空间来说，基矢量的数目是唯一的。有限维向量空间 V 的基向量的数目就称为 V 的维数，记作 $\dim(V)$。比方说，$\dim(\mathbb{C}^n)=n$。

设 V 为有限维向量空间，再设它能分解成子空间 $U_1,U_2,...,U_m$ 的直和，即 $V=U_1\oplus...\oplus U_m$，则可以证明

$$\dim(V)=\dim(U_1)+\dim(U_2)+...+\dim(U_m).$$

14.3 线性算符

向量空间 V 上的线性算符 T 就是 V 到自身的一个映射，$T:V\to V$。对于 V 的任意向量 $|u\rangle$，T 将会把它映射到 V 的一个新向量 $|v\rangle$，记作 $|v\rangle=T|u\rangle$，并将 $T|u\rangle$ 读作算符 T 在向量 $|u\rangle$ 上的作用。并且，这种作用得保持向量空间 V 的线性结构，即保持向量空间的加法和标量乘法。具体来说就是，对于任意 $|u\rangle,|w\rangle\in V$，必有

$$T(|u\rangle+|w\rangle)=T|u\rangle+T|w\rangle.$$

另外，对于任意 $\lambda\in\mathbb{C}$，$|u\rangle\in V$，有

$$T(\lambda|u\rangle)=\lambda(T|u\rangle).$$

比方说，对于所有复多项式构成的向量空间，如下定义的运算 T 是一个线性算符，$Tp(z)=zp(z)$，式中 $p(z)$ 为任意一个多项式，T 的作用是将这个多项式乘以 z，结果显然还是一个多项式。同样，下面的求导运算 D 也是复多项式向量空间上的线性算符，$Dp(z)=\frac{dp(z)}{dz}$。

再比方说，容易验证下式定义的 T 是向量空间 $\mathbb{C}^2$ 上的线性算符，$T([x,y])=[2x-y,7x+5y];x,y\in\mathbb{C}$。

有一个特殊的线性算符值得单独解释一下，那就是恒等算符，也叫做单位算符，我们常常将它简记为 1，它在任何向量上的作用结果依然是这个向量本身，即 $1|u\rangle=|u\rangle$。

我们可以按照下面的式子分别定义两个线性算符 A，B 的和 $A+B$，以及它们的乘积 AB，

$$(A+B)|u\rangle=A|u\rangle+B|u\rangle,$$
$$(AB)|u\rangle=A(B|u\rangle).$$

读者容易验证，$A+B$ 和 AB 都依然是线性算符。

给定线性算符 T，如果向量 $|u\rangle$ 满足 $T|u\rangle=0$，则称之为算符 T 的零向量。很明显，如果 $|u\rangle,|v\rangle$ 均为 T 的零向量，则 $|u\rangle+|v\rangle$ 也为 T 的零向量，而 $\lambda|u\rangle$ 这样的 $|u\rangle$ 乘上标量常数 λ 以后的向量也是 T 的零向量。也即是说，T 的所有零向量所构成的空间 $\mathrm{Null}(T)$ 在

加法和标量乘法下封闭。因此 Null(T) 构成整个向量空间的一个子空间，称为算符 T 的零空间。

比方说，由于常数的导数为 0，前面我们定义的作用在所有复多项式的向量空间上的线性算符 D 的零空间就是 0 次多项式空间，即 $\text{Null}(D)=\{a:a\in\mathbb{C}\}$。

对于线性算符 A，若存在某个算符 B，使得 $AB=BA=1$，则我们说，A,B 互为逆算符，记作 $A=B^{-1}$, $B=A^{-1}$。

如果算符 A 有逆算符，我们就说它可逆。很容易证明，A 可逆的必要条件是 $\text{Null}(A)=\{0\}$。这是因为，如果 A 可逆，则我们就可以用 A^{-1} 去乘零向量的方程 $A|u\rangle=0$，从而得到 $|u\rangle=0$。实际上可以进一步证明，如果 A 是有限维向量空间上的线性算符，则 $\text{Null}(A)=\{0\}$ 其实是 A 可逆的充要条件。

线性算符与矩阵

在 n 维向量空间 V 上取定一组基 $|e_1\rangle,|e_2\rangle,...,|e_n\rangle$，注意到 V 中的任何向量都可以用基向量来展开，进而设线性算符 A 在基向量上的作用结果可以用基展开为

$$A|e_j\rangle=\sum_{i=1}^{n}|e_i\rangle a_{ij},\quad j=1,2,..,n \tag{14.3}$$

式中 a_{ij} 为复组合系数，请注意式中 a_{ij} 下标的顺序。人们常常将所有 a_{ij} 排成一个矩阵，记为 $\widehat{A}$,

$$\widehat{A}=\begin{pmatrix}a_{11} & a_{12} & ... & a_{1n}\\ a_{21} & a_{22} & ... & ...\\ ... & ... & ... & ...\\ a_{n1} & ... & ... & a_{nn}\end{pmatrix}.$$

由此我们很容易注意到 (14.3) 式 $A|e_j\rangle$ 的结果为，行 $(|e_1\rangle,|e_2\rangle,...,|e_n\rangle)$ 与矩阵 $\widehat{A}$ 的第 j 列各分量对应相乘并求和。进而我们就可以将 (14.3) 式重新写成

$$\begin{aligned}A(|e_1\rangle,|e_2\rangle,...,|e_n\rangle)&=(A|e_1\rangle,A|e_2\rangle,...,A|e_n\rangle)\\ &=(|e_1\rangle,|e_2\rangle,...,|e_n\rangle)\begin{pmatrix}a_{11} & a_{12} & ... & a_{1n}\\ a_{21} & a_{22} & ... & ...\\ ... & ... & ... & ...\\ a_{n1} & ... & ... & a_{nn}\end{pmatrix}.\end{aligned}$$

也即是说，取定了基以后，线性算符 A 可以唯一性地表示成矩阵 $\widehat{A}$。

下面我们再来考察 A 在向量 $|v\rangle$ 上的作用 $A|v\rangle=|u\rangle$ 可以如何表示。为此我们将 $|v\rangle$ 在基中展开为 $|v\rangle=\sum_{j=1}^{n}v_j|e_j\rangle$，注意到 $A|v\rangle=A(\sum_{j=1}^{n}v_j|e_j\rangle)=\sum_{j=1}^{n}v_jA|e_j\rangle=\sum_{j,i=1}^{n}|e_i\rangle a_{ij}v_j=\sum_{i=1}^{n}(\sum_{j=1}^{n}a_{ij}v_j)|e_i\rangle$。显然最后结果中的 $(\sum_{j=1}^{n}a_{ij}v_j)$ 为矩阵 $\widehat{A}$ 的第 i 行各分量与列矢量

$v=\begin{pmatrix}v_1\\v_2\\...\\v_n\end{pmatrix}$ 各分量对应相乘并求和。记 $u_i=\sum_{j=1}^n a_{ij}v_j$，则有 $A|v\rangle=\sum_{i=1}^n u_i|e_i\rangle=|u\rangle$，其中

$$\begin{pmatrix}u_1\\u_2\\...\\u_n\end{pmatrix}=\begin{pmatrix}a_{11}&a_{12}&...&a_{1n}\\a_{21}&a_{22}&...&...\\...&...&...&...\\a_{n1}&...&...&a_{nn}\end{pmatrix}\begin{pmatrix}v_1\\v_2\\...\\v_n\end{pmatrix}.$$

很显然，最终的这个结果说明，取定了基以后，我们可以将向量表示成列矢量，将线性算符表示成矩阵，而算符在向量上的作用就正好对应于矩阵与列矢量相乘！也正是在这个意义上，我们可以粗略地将抽象的向量想象成列矢量，将抽象的线性算符想象成矩阵。

不仅如此，我们还可以证明，取定了基以后，线性算符的和可以表示成相应矩阵的和，而线性算符的乘积可以表示成相应矩阵的矩阵乘积。

以算符乘积为例。假设有线性算符 A 和 B，其中 A 在基向量上的作用如前所述，而 B 在基向量上的作用为 $B|e_j\rangle=\sum_{i=1}^n|e_i\rangle b_{ij}$，其中 b_{ij} 为复矩阵 $\widehat{B}$ 的第 i 行第 j 列。则 $AB|e_k\rangle=A(B|e_k\rangle)=A(\sum_{j=1}^n|e_j\rangle b_{jk})=\sum_{j=1}^n A|e_j\rangle b_{jk}=\sum_{i,j=1}^n|e_i\rangle a_{ij}b_{jk}=\sum_i^n|e_i\rangle(\sum_{j=1}^n a_{ij}b_{jk})$。进而根据算符与矩阵之间的对应关系我们可以知道，$AB$ 乘积算符的矩阵 $\widehat{AB}$ 的第 i 行第 k 列必为

$$(\widehat{AB})_{ik}=\sum_{j=1}^n a_{ij}b_{jk}. \tag{14.4}$$

这个方程告诉我们，矩阵 $\widehat{AB}$ 的第 i 行第 k 列由矩阵 $\widehat{A}$ 的第 i 行各分量与矩阵 $\widehat{B}$ 的第 k 列各分量对应相乘并求和。通常人们把这个矩阵乘法规则归纳为如下方程

$$\widehat{AB}=\begin{pmatrix}a_{11}&a_{12}&...&a_{1n}\\a_{21}&a_{22}&...&...\\...&...&...&...\\a_{n1}&...&...&a_{nn}\end{pmatrix}\begin{pmatrix}b_{11}&b_{12}&...&b_{1n}\\b_{21}&b_{22}&...&...\\...&...&...&...\\b_{n1}&...&...&b_{nn}\end{pmatrix}=\widehat{A}\widehat{B}.$$

由此可见，算符的乘积正好表示成了相应矩阵的矩阵乘积。正是在这个意义上，我们有时候会说，线性算符和矩阵本质上是一回事。

14.4 不变子空间、本征值与本征向量

为了更好地理解向量空间 V 上的线性算符 A，我们常常将 V 分解为若干个真子空间的直和，

$$V=U_1\oplus U_2\oplus...\oplus U_m. \tag{14.5}$$

然后我们可以将算符 A 限制在某个真子空间 U_j 上来考虑，记为 $A|_{U_j}$，它表示仅仅将 A 作用在子空间 U_j 内的向量上。这样做是因为，限制以后的 $A|_{U_j}$ 往往会比原来的整个算符 A 更简单，更好理解。

但是，$A|_{U_j}$ 不一定是 U_j 上的线性算符，也就是说，虽然 $A|_{U_j}$ 作用在 U_j 上，但它不一定将 U_j 映射到 U_j 本身，有可能有某些 U_j 里的向量会被映射到 U_j 外面去！如果是这种情形，那单独研究 $A|_{U_j}$ 其实就没有任何简化。所以为了能真正简化问题，我们通常要求上面分解出来的每一个 U_j 在 $A|_{U_j}$ 的作用下都封闭，即任何一个 U_j 中的向量在 A 的作用下都依然映射到 U_j 中的某个向量。或者说，$A|_{U_j}$ 将把 U_j 映射到 U_j。这时候，我们就称这些子空间为算符 A 的**不变子空间**。可见，不变子空间的本质就是要在算符 A 的作用之下保持封闭。

例如 $\mathrm{Null}(A)$ 就是 A 的一个不变子空间。这是因为，首先零向量 0 显然属于 $\mathrm{Null}(A)$。其次，根据 $\mathrm{Null}(A)$ 的定义，对于任意 $|u\rangle \in \mathrm{Null}(A)$，必有 $A|u\rangle = 0 \in \mathrm{Null}(A)$，从而 $\mathrm{Null}(A)$ 在 A 的作用下的确封闭。

对于某个非零向量 $|v\rangle$，如果一维子空间 $U_v = \{c|v\rangle : c \in \mathbb{C}\}$ 是算符 A 的不变子空间，那么就称向量 $|v\rangle$ 为算符 A 的**本征向量**。根据不变子空间的定义，这时候必有 $A|v\rangle \in U_v$，从而也即是说，$A|v\rangle$ 必定具有 $\lambda|v\rangle$ 的形式，即

$$A|v\rangle = \lambda|v\rangle.$$

这个方程就称为算符 A 的本征方程，其中的 λ 就称为 A 的**本征值**。

当然我们也可以在向量空间上选定一组基，进而把一切都翻译成矩阵。比方说，取基以后，我们可以将向量 $|v\rangle$ 表示成列矢量 v，把算符 A 表示成矩阵 $\widehat{A}$，那上面的本征方程就可以表示成一个矩阵方程，$\widehat{A}v = \lambda v$，也即是

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix}.$$

人们也称这个方程为矩阵 $\widehat{A}$ 的本征方程。

我们可以把算符 A 的本征方程改写成 $(A - \lambda \cdot 1)|v\rangle = 0$, 其中 $|v\rangle$ 为非零向量。可见，λ 为算符 A 本征值的充要条件是，$\mathrm{Null}(A - \lambda \cdot 1)$ 中包含有非零向量。对于有限维向量空间，根据算符可逆的充要条件，这等价于 $A - \lambda \cdot 1$ 不可逆。所以，对于有限维向量空间，λ **为算符 A 的本征值等价于 $A - \lambda \cdot 1$ 不可逆**。

但是，给定一个本征值 λ，可能有多个线性无关的向量满足同样的本征方程。也即是说，方程 $(A - \lambda \cdot 1)|v\rangle = 0$ 可能有多个线性无关的本征向量解。不过很显然，所有这些本征向量解的全体就构成了 $\mathrm{Null}(A - \lambda \cdot 1)$，人们也常常称之为 A 的本征值为 λ 的本征空间，记为 $E(\lambda, A)$。因此，

$$E(\lambda, A) = \mathrm{Null}(A - \lambda \cdot 1).$$

14.5 内积空间

内积

对于高中学过的三维向量空间 $\mathbb{R}^3$ 中的向量 $\mathbf{x}=[x_1,x_2,x_3]$，我们可以定义它的模长 $|\mathbf{x}|$ 为 $|\mathbf{x}|^2=x_1^2+x_2^2+x_3^2$。如果另有一个向量 $\mathbf{y}=[y_1,y_2,y_3]$，那我们还可以定义两个向量的内积 $\mathbf{x}\cdot\mathbf{y}$ 为 $\mathbf{x}\cdot\mathbf{y}=x_1y_1+x_2y_2+x_3y_3$。

推广到 n 维实向量空间 $\mathbb{R}^n$ 中的两个向量 $\mathbf{x}=[x_1,x_2,...,x_n]$ 和 $\mathbf{y}=[y_1,y_2,...,y_n]$，我们可以将它们的内积定义为 $\mathbf{x}\cdot\mathbf{y}=x_1y_1+x_2y_2+...x_ny_n$。特别的，$\mathbf{x}$ 的模长 $|\mathbf{x}|$ 定义为 $|\mathbf{x}|^2=\mathbf{x}\cdot\mathbf{x}$。

但是，在量子力学中，我们主要关心的是复向量空间。对于最简单的复向量空间 $\mathbb{C}^n$ 中的向量 $\mathbf{z}=[z_1,z_2,...,z_n]$ 和 $\mathbf{w}=[w_1,w_2,...,w_n]$ 我们需要把内积的概念推广成，$\overline{w}_1z_1+\overline{w}_2z_2+...+\overline{w}_nz_n$，式中 $\overline{w}$ 表示复数 w 的复共轭，有时候也写成 w^*。为什么要在内积的定义里面多加一个复数共轭我们很快就会清楚。但很显然，这样定义以后，作内积的两个向量就不完全对称了。我们可以利用前面发明的抽象向量记号 $|z\rangle=\mathbf{z},|w\rangle=\mathbf{w}$ 来反映这种不对称性，具体来说，我们将 $|z\rangle$ 和 $|w\rangle$ 的内积记为 $\langle w|z\rangle$，它由下式计算

$$\langle w|z\rangle=\overline{w}_1z_1+\overline{w}_2z_2+...+\overline{w}_nz_n.$$

注意这个公式，我们需要对内积符号左边的那个向量进行复数共轭。

之所以将左边的向量加上一个复共轭，原因在于向量的模长由向量与其自身的内积确定。而只有这样定义以后，向量 $|z\rangle$ 的模长公式 $\big||z\rangle\big|^2=\langle z|z\rangle$ 才是一个大于零的正实数 (因为复数 z 只有和它的复共轭 $\bar{z}$ 相乘结果才是正实数)。

将以上的例子作一个一般性的抽象，我们就可以在任何复向量空间 V 上定义内积运算。所谓的内积，就是一个函数，它把 V 中的任何一对向量 $|u\rangle,|v\rangle$ 映射到一个复数 $\langle u|v\rangle$。并且这个映射满足：(1). **正定性**，即对于任何 $|v\rangle\in V$，$\langle v|v\rangle\geq 0$, 等于号仅当 $|v\rangle=0$ 时才成立。(2). **线性性**，即 $\langle u|v\rangle$ 对于符号右边的向量来说是线性的。具体来说即是满足 $\langle u|\big(c_1|v_1\rangle+c_2|v_2\rangle\big)=c_1\langle u|v_1\rangle+c_2\langle u|v_2\rangle$。(3). **共轭对称性**，即 $\langle u|v\rangle^*=\langle v|u\rangle$。定义了内积的复向量空间就称为内积空间。

对于高中学过的三维向量，我们知道，两个向量的内积为零又称作这两个向量正交。这个概念同样可以推广到任何内积空间。如果 $\langle u|v\rangle=0$, 我们就称 $|u\rangle$ 和 $|v\rangle$ **正交**。类似的，类比于单位向量的概念，如果一个向量 $|u\rangle$ 模长为 1，即满足 $\langle u|u\rangle=1$，我们就说它是归一的。

正交归一基与向量表示

对于一个有限维内积空间 V, 如果它的一组基 $|e_1\rangle,|e_2\rangle,...,|e_n\rangle$ 满足

$$\langle e_i|e_j\rangle=\delta_{ij},i,j=1,2,...,n.$$

我们就称这一组基为正交归一基。这里克龙内克符号 δ_{ij} 当两个指标相同时取 1，不同时取 0。

取定了正交归一基以后，我们就可以把向量 $|v\rangle$ 展开成，$|v\rangle=\sum_{i=1}^n v_i|e_i\rangle$。通过利用正交归一关系计算内积 $\langle e_i|v\rangle$ 我们容易发现，

$$v_i=\langle e_i|v\rangle.$$

由此我们也可以得到 $v_i^* = \langle e_i|v\rangle^* = \langle v|e_i\rangle$(利用了内积的共轭对称性)。

现在我们来考察 $\langle u|v\rangle$, 我们有 $\langle u|v\rangle = \langle u|\sum_{i=1}^n v_i|e_i\rangle = \sum_{i=1}^n v_i\langle u|e_i\rangle$, 假设把 $|u\rangle$ 展开成 $|u\rangle = \sum_{i=1}^n u_i|e_i\rangle$，则利用 $u_i^* = \langle u|e_i\rangle$，我们可以得到

$$\langle u|v\rangle = \sum_{i=1}^n u_i^* v_i. \tag{14.6}$$

前面我们已经知道，向量 $|u\rangle, |v\rangle$ 在基中可以分别表示为列矢量 u, v, 它们由下式给出

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \cdots \\ u_n \end{pmatrix} = \begin{pmatrix} \langle e_1|u\rangle \\ \langle e_2|u\rangle \\ \cdots \\ \langle e_n|u\rangle \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{pmatrix} = \begin{pmatrix} \langle e_1|v\rangle \\ \langle e_2|v\rangle \\ \cdots \\ \langle e_n|v\rangle \end{pmatrix}.$$

则很显然，上面的 (14.6) 式其实就是

$$\langle u|v\rangle = \begin{pmatrix} u_1^* & u_2^* & \cdots & u_n^* \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{pmatrix} = u^\dagger v. \tag{14.7}$$

式中 $u^\dagger$ 表示列矢量 u 的复共轭再转置 (称之为厄密共轭)，它显然是行矢量。(14.7) 式告诉我们，可以将内积符号左边的向量想象成一个行矢量，它是相应列矢量的厄密共轭。

基于以上观察，我们干脆可以定义一种左矢 $\langle u|$，其定义就是

$$\langle u| = |u\rangle^\dagger.$$

左矢可以想象成行矢量，是列矢量的厄密共轭。因此，内积运算最终的公式形式 (14.7) 就是行矢量乘以列矢量。注意，行矢量在左边，列矢量在右边。

正交补空间

最后我们再给出一个简单的概念。设 V 为内积空间，U 为 V 的一个子空间。则 U 的正交补空间也是一个子空间，记为 $U^\perp$, 它的定义是

$$U^\perp = \{|v\rangle \in V : \text{对每个 } |u\rangle \in U \text{ 均有 } \langle u|v\rangle = 0\}.$$

也即是说，$U^\perp$ 由 V 中所有与 U 正交的向量构成。从而，整个向量空间 V 可以正交分解成 U 和它的正交补 $U^\perp$ 的直和，即

$$V = U \oplus U^\perp.$$

关于内积空间以及内积空间中的线性算符的更多讨论，由于和量子力学有直接的关系，所以我们放在本书的正文中进行。

参考文献

我将首先列举一些我参考过的，或者阅读过的量子力学教材和讲义，然后再针对本书讨论过的一些话题推荐进一步的阅读材料。

量子力学书

三本量子力学入门书：1. Quantum Mechanics，作者：Leonard Susskind and Art Friedman. 2. 《费曼物理学讲义》第三卷，作者：[美] 理查德 · 费曼，上海科学技术出版社。3.《量子力学概论》，格里菲斯 (D J. Griffiths) 和施勒特 (D F. Schroeter)，机械工业出版社。

三本很好的提高书：1.《量子力学》两卷，作者：C. Cohen-Tannoudji /B. Diu /F. Laloe，高等教育出版社。2.《现代量子力学》，作者：J. J. Sakurai，世界图书出版公司。3.《量子力学原理》，作者：山卡（R. Shankar），世界图书出版公司。

其它值得参考的书：1. Lectures on Quantum Mechanics，作者：Steven Weinberg, 出版社：Cambridge University Press。2. James Binney and David Skinner, The Physics of Quantum Mechanics, 3. Quantum Mechanics - A Modern Development, 作者：Leslie E. Ballentine, World Scientific Publishing Company. 4. David Tong 的在线讲义，Lectures on Topics in Quantum Mechanics, http://www.damtp.cam.ac.uk/user/tong/topicsinqm.html 5. 最后推荐一下我本科量子力学老师张永德老师的《量子力学 (第 2 版)》，作者：张永德，科学出版社。

专题文献

关于量子信息与量子计算，推荐 Preskill 超级好的在线讲义 Preskill, J., Lecture Notes for Physics 229: Quantum Information and Computation. California Institution of Technology, Pasadena. http://theory.caltech.edu/ preskill/ph229/

关于超导量子比特，推荐综述，P. Krantz, M. Kjaergaard, F. Yan, T.P. Orlando, S. Gustavsson, and W. D. Oliver, A Quantum Engineer' s Guide to Superconducting Qubits, arXiv:1904.0656

关于量子霍尔效应的更多讨论参见 David Tong 的在线讲义"Lectures on the Quantum Hall Effect"，http://www.damtp.cam.ac.uk/user/tong/qhe.html

关于任意子和相关理论的更多讨论，推荐 Steven H. Simon 的在线讲义 "Topological Quantum"，http://oxfordtopquantum.tiddlyspot.com/

关于拓扑量子计算，下面的综述是极好的 Nayak, C., Simon, S. H., Stern, A., Freedman, M., Sarma, S. D.. (2007). Non-abelian anyons and topological quantum computation. Review of Modern Physics, 80(3), 1083-1159.

关于二次量子化方法在玻色-爱因斯坦凝聚 (BEC) 中的应用，我们推荐下面这篇综述。Dalfovo, F. , Giorgini, S. , Pitaevskii, L. P. , Stringari, S. . (1998). Theory of bose-einstein condensation in trapped gases. Review of Modern Physics, 71(3), 463-512. 更多二次量子化和场论方法对 BEC 和超流的讨论可以参见 A. Altland, B. Simons，Condensed Matter Field Theory，chapter 6, Cambridge University Press 2010。关于量子场论方法在凝聚态物理中的应用也都推荐这本书。