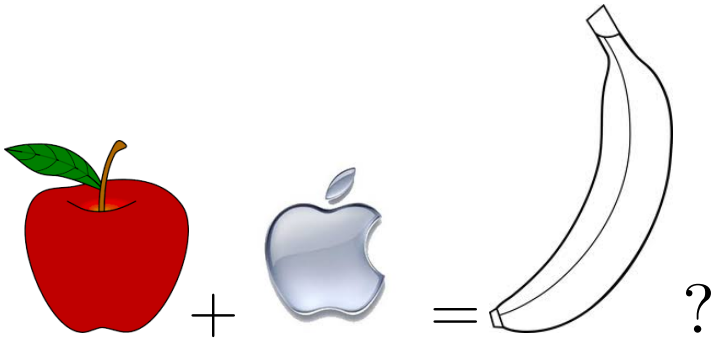
二态系统的量子力学

吴金闪



一个左边缺一口的苹果加上一个右边缺一口的苹果等于香蕉。（都换成自己的图片，采用 icon 的设计风格）

$$|\leftarrow_x\rangle + |\rightarrow_x\rangle = \sqrt{2}\,|\uparrow_z\rangle$$

物理学家给这个世界寻找合适的数学结构

打开一扇通往量子的世界的门

不学太多的具体知识，能做深入思考吗?

Teach Less, Learn More

2017 年 5 月 31 日

目录

第一部分　引论：为什么需要量子力学 17

第一章　量子系统的实验 21
1.1 经典粒子和经典波的实验 21
1.2 单个电子或者单个光子的实验 31
1.3 作业 44
1.4 本章小结 45

第二章　物理世界的数学模型 49
2.1 经典世界的确定性的数学模型 50
2.2 经典世界的随机性的数学模型 53
2.3 理论描述的目标：前后两次测量的关联 58
2.4 量子系统的确定性经典理论的尝试 61
2.5 量子系统的随机性经典理论的尝试 64
2.5.1 互斥构造 66
2.5.2 独立构造 68
2.6 作业 71
2.7 本章小结 71

第二部分　数学物理学上的准备 75

第三章　二维空间的线性代数：Dirac 符号、抽象算符与表象 79
3.1 数学是现实世界的结构和思维的语言 79
3.2 从分量形式的矢量和矩阵到抽象矢量和算符，Dirac 符号 . . . 82
3.2.1 实数域上的分量形式的矩阵和矢量，复习 82

3.2.2 Dirac 符号 83
3.2.3 矩阵的谱展开，Dirac 符号与抽象矢量 85
3.3 复数域上的抽象矢量与抽象矩阵算符 89
3.3.1 复数域上的矩阵和矢量运算 90
3.3.2 复数域上的左矢量的分量形式的定义 90
3.3.3 复数域上抽象算符的谱展开 92
3.3.4 抽象算符的函数与算符的代数 94
3.4 共同本征向量, 表象理论与线性变换 95
3.4.1 表象之间的变换 96
3.4.2 举例 98
3.4.3 相似变换作为算符 * 99
3.5 作业 101
3.6 本章小结 102

第四章 离散状态的概率论 105
4.1 古典概型概率论的 Dirac 符号形式 106
4.2 现代概率三元体 112
4.3 Dirac 符号作为概率论的语言 114
4.4 概率转移算符 * 117
4.5 作业 119
4.6 本章小结 120

第五章 经典力学精要 121
5.1 力学思想 121
5.2 Newton 力学的基本概念 123
5.3 势函数、保守力与能量守恒 128
5.4 从 Newtonian 力学到 Hamiltonian 力学 130
5.4.1 单个质点的分析力学形式 130
5.4.2 从单体到多体的分析力学形式 133
5.5 作业 135
5.6 本章小结 136

第三部分　单粒子系统的量子力学 137

第六章　量子系统状态的数学模型 141
6.1 有关量子状态和测量的公理 141
6.2 经典随机客体理论的公理形式 144
6.3 量子静态理论用于描述测量实验 146
6.3.1 自旋的测量和再测量 146
6.3.2 为什么自旋算符要用 $\hat{r}\cdot\vec{\hat{\sigma}}$? 152
6.3.3 光子的 which-way 实验 154
6.3.4 为什么偏振算符取公式 (6.46) 的形式? 156
6.3.5 自旋和光子 which-way 实验解释的补充：纠缠 158
6.4 粒子走哪一条路径和概率叠加原理 159
6.5 试试把测量分成三个步骤 164
6.6 完整地测量自旋的状态 166
6.7 量子与经典密度矩阵的区别 167
6.8 作业 170
6.9 本章小结 173

第七章　量子系统的演化 175
7.1 Schrödinger 方程 176
7.2 Schrödinger 绘景与 Heisenberg 绘景 179
7.3 作业 182
7.4 本章小结 184

第八章　单个谐振子的量子力学 185
8.1 位置表象 185
8.2 代数解法与能量表象 190
8.3 代数解法的更多细节：一切都是对易关系 * 194
8.4 作业 195
8.5 本章小结 195

第四部分　耦合量子系统：相互作用与纠缠 197

第九章　耦合量子系统的状态与测量 199

9.1 直积空间 199
9.2 经典关联态的测量 201
9.3 量子纠缠态的测量 203
9.3.1 计算视角一：一组对易算符的测量 203
9.3.2 计算视角二："先后"测量 205
9.3.3 计算视角三：可分辨性 208
9.4 作业 210
9.5 本章小结 210

第十章 耦合量子系统的演化与纠缠 211
10.1 演化以及演化导致的纠缠 211
10.2 测量导致的纠缠 213
10.3 作业 216
10.4 本章小结 217

第十一章 量子力学的确定性经典理论的再一次讨论 219
11.1 纠缠态上的量子测量的结果 219
11.2 经典关联态上的测量的 Bell 不等式 220
11.3 Bell 定理的实验检验以及 Die Hard 教授的吹毛求疵 223
11.4 前后两次测量的关联的量子力学计算 227
11.5 互斥构造和独立构造的经典理论的非定域性 228
11.6 作业 231
11.7 本章小结 231

第十二章 测量的量子力学？ 233
12.1 经典硬币的测量：克隆 233
12.2 波函数塌缩：非对角项的消失 235
12.3 经典和量子测量的统一的框架? 238
12.4 作业 240
12.5 本章小结 240

第五部分 量子信息简介 241

第十三章 远程传态和 Deutsch 算法 243

13.1 量子远程传态 243
13.2 Deutsch 算法 246
13.3 作业 250
13.4 本章小结 250

第十四章 量子博弈 253
14.1 经典客体上的操作和博弈 254
14.2 量子客体上的操作和博弈 260
14.3 经典和量子博弈的异同 265
14.4 作业 266
14.5 本章小结 266

参考文献 276

名词索引 277

人名与常用翻译 281

插图目录 284

举例目录 288

献给

我的母亲黄云弟

致谢

本书在内容和思想上受到Ballentine的《Quantum Mechanics – a modern development》、喀兴林的《高等量子力学》、裴寿镛的量子力学课程、Affleck的高等量子力学课程非常大的影响。在这里一并对这些深刻地影响了我对量子力学的理解的老师致谢。除了真实课堂，本书的逻辑结构很大程度上还受Feynman的《Feynman 物理学讲义》第三卷的影响，可惜没有能坐在Feynman的课堂里。

感谢阅读了本书的原稿并提出很多有意义的建议的师长尤其是裴寿镛、杨展如，感谢我在编写和使用本书的过程中学习我的《量子力学》课程的北京师范大学的本科生研究生，以及迫使我必须用更加简单的实验和语言来展示量子力学魅力的我的公众报告的公众和中小学生听众。

感谢裴寿镛对本书中的几乎每一句话的推敲，内容选择和结构安排上的批判性和建设性建议，提供的部分习题，以及参与听课、答疑，并对具体教学环节的指导。我记得很多次的讨论，需要从手机换成座机，需要从晚间延续到半夜。我记得很多次的争论，我声音提高面红耳赤，裴老师——尽管保持冷静——却也寸步不让。经过这些讨论和争论，再加上教学上的实验，有的时候我们能够达成一致的意见，有的时候就只能尊重保持各自的意见。我想裴老师是一个得道的“师者”——这个学生都毕业这么多年了，仍然需要负责给这个学生传道授业解惑。

感谢我的孩子们，吴逸兮和吴立心，你们给我很多的学习和教学上的启发。感谢我的夫人冯倩对于我做各种探索的支持。感谢我的岳母姚书君对孩子们的悉心照顾，使得我有更多的时间来做这些探索并完成本书。

本书的电子版可以从网页“吴金闪的书们”找到。如果你是实体书的读者，需要输入网址的话，它是：http://www.systemsci.org/jinshanw/books。

前言

通常的量子力学书籍从波函数开始讲起，需要比较多的背景知识和比较高的数学技能。然而，其在提高读者对量子力学的理解上，往往比较有限。很多时候，需要等到高等量子力学的课程中，用算符和向量的语言重新认识量子力学之后，读者才能形成对量子力学的一定的理解。最后，通过阅读研究性论文或文集，例如Bell的《Speakable and Unspeakable in Quantum Mechanics》[1]，以及von Nuemann的《Mathematical Foundations of Quantum Mechanics》[2]、Dirac的《The Principles of Quantum Mechanics》[3]、Shankar的《Principles of Quantum Mechanics》[4] 等强调量子力学基础的书籍，加上偏重量子力学基础的量子信息书籍，例如Nielsen和Chuang的《Quantum Computation and Quantum Information》[5]，以及Preskill的《Lecture Notes for Physics 229: Quantum Information and Computation》[6][1]之后，才能对量子力学形成比较完整的认识，了解其主要特点和可能的问题（或者说和经典世界的一般认识有严重冲突的地方，不一定是问题）。

同时，也有很少的量子力学书籍从最简单的二态系统开始介绍量子力学的最核心和最基本的理论，例如Feynman的《Feynman 物理学讲义第三卷》[7]、Ballentine的《Quantum Mechanics: a morden development》[8][2]、Sakurai的《Modern Quantum Mechanics》[9]。不过，《Feynman 物理学讲义第三卷》主要关注核心，讨论的内容比较有限，尤其没有深入讨论量子力学的基础的问题、测量的问题等等量子态的叠加性所带来的量子力学理解上的问题。后两者包含的内容又实在非常多。

因此，写一本用最简单的方法，来交代清楚量子力学的核心概念，能够帮助学生形成对量子力学的理解的书，就成了作者很长时间的梦想。一方面是上面提到的责任感的驱动，另一方面是作者曾经在三个学校跟物理学家、

[1] Preskill 的书在量子力学基础部分对量子力学理解上的可能的问题做了深入浅出的讨论，强烈推荐。

[2] Ballentine 的书非常深刻细致并具有启发性。在对量子力学的理解上 Ballentine 是有独到之处的。

数学家，其中包含裴寿镛、Ballentine等对量子力学有深刻的理解的老师，学习过 6 门量子力学和 3 门量子场论课程，自认为对量子力学的认识和学习有独到之处。再加上跟裴寿镛老师一起做过一点点关于量子力学基础——隐变量理论的可能性——的工作，认为对量子力学和经典力学的区别有相当深刻的认识，也做过一点量子博弈——把在经典硬币上的翻转游戏推广到自旋的转动的游戏——方面基础的工作，就自认为有一定的基础来跟大家分享对量子力学的认识和理解。还有一个原因是我认为不管是不是物理学专业的学生都需要学习量子力学，但是没有必要学太多除了对量子力学的理解之外的具体知识。正好，北京师范大学系统科学学院的本科生们给我提供了用我的新讲授方式来实践的机会。于是，就有了本书的初稿。选好例子又避开太多的具体知识，还要能够促进对量子力学的思考和理解，是一件困难的事情。不过，我很开心，我自认为，沿着Feynman在《Feynman 物理学讲义第三卷》[7] 的思路，我找到了一个好的解决方式。

着眼于增加对量子力学和量子世界的理解，本书最主要的目的就是说清楚经典系统的行为和量子系统的行为的差别和联系，然后比较经典力学和量子力学的异同。在尽量说清楚这个主题的同时，通过从量子系统的行为来构建量子理论的过程来体现数学和物理的关系，从而促进读者对什么是科学的理解：批判性思维在科学中的地位，观察和实验在科学中的地位，科学和数学的关系，可证伪性和可验证性的关系等。在学习方法上，本书非常强调对概念的深刻理解——所谓理解就是把握这个概念与其他概念之间的关系并且联系到实际应用或者实验。因此，也希望本书的读者在阅读和学习本书之后，在学习方法上也有所得，学会做理解型学习 [10]，而不是记忆型学习——做到“知其然，知其所以然，知其所以所以然”。也就是，明白本书主要说了什么（What），怎么说的（How），为什么说这些、为什么这样说（Why），以及还要问理解这些对我来说意味（Meaningful）着什么、读者“我”是否喜欢这样的选择和阐述方式。我称之为WHWM问题。本书对数学物理的基础和计算细节要求都不太高，但是对于“问到底”的精神对于不断地挑战自己思考的深度有非常高的要求。本书原则上对读者没有专业的限制，只要是甚至曾经是理科的学生老师而且善于思考，愿意了解量子系统的奇妙之处还愿意挑战自己的理解和思考，都可以来尝试阅读本书，而且有所收获。但是，警告：本书不是一本容易阅读的书，尤其在思考的深入程度上。

从结构上，见图 1，本书分成几个部分：第一部分，什么是科学，量子

系统的奇怪行为（实验），和这些行为的可能数学模型的尝试；第二部分，线性空间的矢量和算符、Dirac 符号，概率论和以 Hamiltonian 描述的经典力学；第三部分，量子系统的状态、测量和状态的演化；第四、五部分，量子纠缠和纠缠在量子信息上的应用。从整体思路上，第一部分是发展量子理论的动机，而第三部分是量子理论的核心。其它的，第二部分是数学物理基础，有基础的读者可以快速略读；第四、五部分主要用于开阔一下眼界，了解一下量子信息和量子博弈。同时，最后这一部分也强调一下纠缠在应用上的特殊意义，尽管其本质，将来读了这些章节之后可以看见，还是量子态的叠加原理。整本书，有一个主题贯穿始终，能不能用没有状态叠加原理的数学形式，来描述量子系统的行为，或者说，为什么量子系统的行为需要满足叠加原理的数学形式来描述。

希望读者在学习完了本书之后，第一，了解量子系统的行为以及行为和满足状态叠加原理的数学形式之间的关系；第二，思考和建立初步的对物理学（或者说科学）和数学的关系的认识；第三，学会通过阅读其它书籍和文献在核心概念的基础上自学的方法，以及提升自学的意愿。如果还能够对基于批判性思维和系联性思考的理解型学习有比较好的体会，就是额外有所得了。

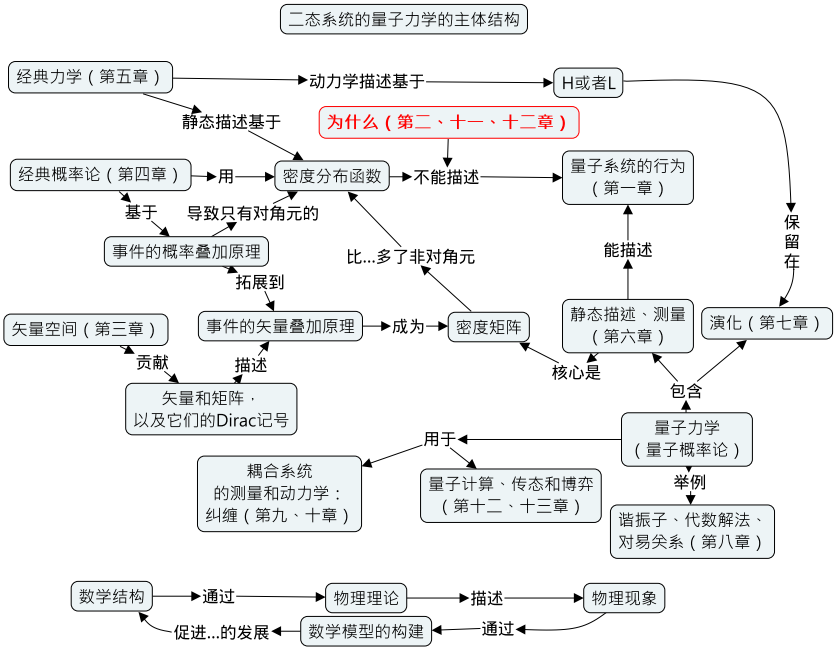


图 1: 本书的体系结构和大多数书不太一样，对于为什么量子系统的理论形式必须是基于矢量叠加原理的量子力学做了很多的讨论。因此，量子系统的实验行为部分在本书里面也占了很大的篇幅，而理论部分从“密度分布函数”到“密度矩阵”的转变以及转变的原因占了核心的地位。通过这样的结构，我们希望读者除了了解量子力学是什么之外，还能够思考为什么量子力学会这样并对此形成一定的理解。图中的下半部分是一个总结，请读者把它和上半部分联系起来，比如通过思考“例如”的关系：“数学结构”——“例如”——“矢量空间”。

第一部分

引论：为什么需要量子力学

科学家，尤其是物理学家，其主要目的就是给实际的世界提供一个可计算的可证伪的[3]理想模型，通过对这个理想模型做计算，人们能够得到实际世界的行为。为什么我们要学习量子力学呢？除了量子力学很有用——它是所有的半导体材料背后的理论（因此你的手机、电脑等依赖于它），它也是所有的化学反应背后的理论，现在它已经被广泛地应用于药物设计等等等等之外，对一个追求智力上深刻的愉悦或者是真正企图理解自然界的行为的人来说，根本的是因为原则上量子力学是所有（微观）客体运动变化的理论基础，并且量子客体的行为超过了我们的大脑所熟悉的直观的经典力学的世界能够解释的范畴。相对论对日常经验的突破是同时的相对性，量子力学对于日常经验的突破更大。这门课程的任务就是呈现给你这个突破是什么。这个突破我称之为一扇门：量子力学是一扇打开你用不同的眼光看世界的门。不打开这样一扇门，你的人生就是不完整的。至于打开之后是否就完整了，另说。

打开这扇门之后，看到了不同的现象，给物理对象找到了不同的数学结构，我们会产生如何来理解这样的数学结构的问题，以及这样的数学结构能够告诉我们什么新的没有想到过没有看见过没有计算出来过的现象。因此，除了量子力学本身，和打开一扇门的作用，学习量子力学还能够深刻地体会到什么是科学，科学与现实，科学与数学的关系。

下面的所有的具体内容，都是围绕着以上这个学习目的——我称之为“量子力学的大图景”，包括这是一扇什么样的门、打开这一扇门的意义、对什么是科学的认识（现实和数学结构的关系、数学结构的理解、理论的应用预测和检验）——来选择、设计和展开的。

在这个第一部分里，我们主要通过介绍几个量子系统的实验及其可能的经典力学的解释来展示：量子系统的实验结果经典理论不能解释。然后，我们还会稍微探索一下一个能够解释量子系统的实验现象的理论的数学结构有什么要求，这个数学结构的特殊之处。

本书第一部分推荐参考书：Feynman的《物理学讲义第三卷》[7]，Ballentine的《Quantum Mechanics – a morden development》[8]。Feynman的小册子《QED: The Strange Theory of Light and Matter》[12] 也非常值得一看。有兴趣得读者还可以看一下Bell的《Speakable and Unspeakable in Quantum Mechanics》[1]。不过更合适的阅读Bell这本书的时间点是在你差

[3]一个理论其所推导出来的结果原则上可以被实验和实践证明是错的，但是迄今为止，还没有被证明是错的就叫做可证伪的但是尚未被证伪的理论。更加深入的讨论见 [11]

不多理解了本书的第十一章之后。各个章节有它们自己的推荐阅读材料。

第一章　量子系统的实验

在这一章里面，我们将介绍一些典型的量子系统的行为，然后告诉大家用经典的粒子的理论（确定性或者随机性的经典力学）或者波的理论（经典波动力学）——两者的背后都是Newton 力学，不能解释这些实验结果。

介绍量子系统的实验之前，我们来讨论三个经典对象——水槽中的水波的双缝实验、进入相机镜头的光波和一把不太可靠的枪射出来的子弹的双缝实验——后者是一个理想实验。在讨论量子系统的行为的时候，这里我们用的是双缝实验以及关于偏振（自旋）的实验。实际上，Feynman 说过如下意思的话，量子力学的一切奇妙之处就在双缝实验。因此，所有的这些实验实际上展示的都是一样的：量子系统的行为到底在什么方面和经典系统的行为不一样，而后者可以通过经典粒子或者波的力学来解释。不过，我们将展示的各个实验确实还是会有每个实验容易理解和不容易理解的地方，相互可以补充，更好地启发你的思考，或者说更好地把你逼入思想的墙角。必须经历过这个被逼入墙角的过程，将来才能够对量子力学的理论体系为什么可以成为量子系统的数学模型有更好的体会，也了解“问题”所在。

所有的实验，我们都会企图用经典波或者经典粒子的理论来解释观测到的现象。在这里要注意，实际上经典介质波本质上还是经典粒子的理论：经典波实际上是一大群联系在一起的粒子的整体行为，两者背后都是 Newton 力学。我们还会讨论一些经典粒子和经典波都不能解释的现象。因此，在下面的实验中我们主要关心这个实验能否由经典波来解释，还是经典粒子来解释，还是两者都不能解释。

1.1　经典粒子和经典波的实验

实验 1.1 (**水波的双缝实验**)：一个较大的水槽，表面上有一个振动的小球。小球的振动形成的水波被一个固定的木板挡住，只能通过木板上的两个小

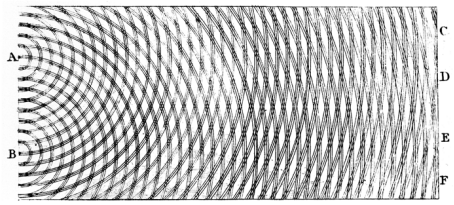


图 1.1: 来自于不同驱动源（A 点，B 点）的波在水槽里面的任意一点的效果都需要把来自于 AB 两点的矢量叠加。这个叠加产生了我们见到的相对稳定的明暗交替的波纹。图片转引自 wikipedia 的“Young氏干涉”条目[1]。

孔传到木板的另一侧。观察另一侧的水波的行为，我们可以看到有的地方振动的幅度很小，有的地方比较大。这个与单独打开一个小孔的时候不一样。单独开一个小孔的时候，如果不是看的范围很远以至于遇到容器壁或者衰减的太厉害的话，各个距离相同的点的振动幅度差不多，也就是成圆形散发的波动。打开两个小孔以后看到的有的地方振动幅度加强（与单个小孔的情形相比，图中明亮的条纹）有的地方减弱（图中比较暗的条纹）的现象就叫做干涉。整体还是呈圆形往外扩散，但是，条纹出现了强弱交替。如图 1.1 所示。

经典波动力学是这样解释的。木板另一侧的任何一个点受两个振动源——也就是那两个小孔——驱动。由于波的传播不是粒子真的沿着波的传播方向上在传波，而是每个粒子在做来回往复的振动，因此，**当同时受多个驱动源驱动的时候，自然其运动就是把多个源传过来的振动方式相叠加**（如图 1.2 所示）：在 t 时刻，第一个小孔传过来的振动希望 x 这个点的位置在 $z_1(x,t)$，第二个小孔传过来的振动希望 x 这个点的位置在 $z_2(x,t)$，那么合起来自然就是

$$z(x,t) = z_1(x,t) + z_2(x,t). \tag{1.1}$$

注意到 $z_i(x,t)$ 是类似于 $A\cos 2\pi\left(\frac{x}{\lambda} + \omega t + \phi_0\right)$ 余弦函数，所以上式经过三角函数和差化积（$\cos(\alpha) + \cos(\beta) = 2\cos\left(\frac{\alpha-\beta}{2}\right)\cos\left(\frac{\alpha+\beta}{2}\right)$）实际上就会成为一个不依赖于时间只依赖于位置的慢变部分和一个依赖于时间和位置的快变部分。其中这个慢变部分可以看成是快变部分的振幅，而这个振幅只依赖于位置。观察到的振动幅度的大小就是这个慢变部分的体现。在这个解释中，黑体字的部分是非常重要的，也是自然的——其背后是力的矢量叠加和

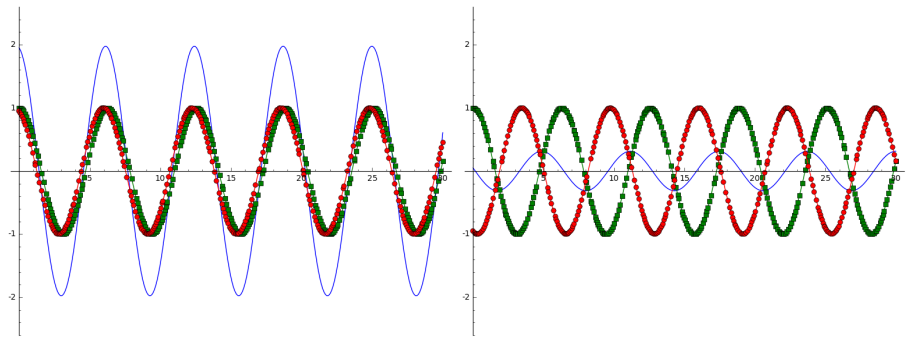


图 1.2: 横轴是时间 t，纵轴是振动高度 z。在同一个地点（x）振子受到来自于两个波源（记单独一个波源的时候振子的高度分别为 $z_1(t)$ 和 $z_2(t)$，下标表示波源，t 是时间）驱动的效果是两个波的效果的矢量叠加。通过这个叠加我们看到了振幅的相长（$\cos(t)+\cos(t+0.1\pi)$）相消（$\cos(t)+\cos(t+0.9\pi)$）现象。我们注意到，振幅相长还是相消取决于两个用来相加的 $\cos\theta_1+\cos\theta_2$ 函数之间的角度的差别 $\theta_1-\theta_2$。这个角度差，例如上式中的 0.1π，是由两个波源到 x 点的距离差造成的（为什么距离差造成这个角度差别可以参考光学或者波动力学的教材。在这里不讨论）。

Newton 定律。但是，如果我们的运动实际上就是粒子本身在沿着波传播的方向传播，例如光波，情况就不一样了，黑体字部分就不对了。以后我们还会仔细来考虑光波的干涉是如何解释的。暂时我们把光波也可以看作是在空间每一个点上有一个矢量这样的经典波的数学模型，因此上面的解释在水波和光波上都可以用。将来我们再讨论光波这样的**可以看做一个个粒子在沿着光波传播的方向传播**的波的干涉问题。

经典的波是介质波：一个点的介质的运动影响邻近点的介质的运动。一般来说介质点本身并不随着振动传播而传播开去。于是，如果有一个点的介质收到多个邻近点（不管在相同还是不同方向上）的振动的影响，那么其效果自然就是这些影响的叠加。这个叠加是 Newton 定律和力的叠加的结果。下面要讨论的光波和电子波等不再是介质波而是物质波：一个个的小个体在振动传播的方向上传播开去，它们不需要额外的介质来承担这样的运动。于是，一个粒子带动旁边的粒子运动这样的图景就不再适用了，一个波分成两束或者多束这样的图景也就不再适用了。至于我们如何肯定这样的小个体在传播过程中是否还是把自己分裂成很多份到了被观测的时候再一次合

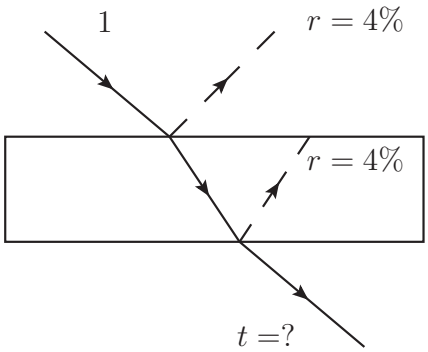


图 1.3: 光子经过一块玻璃，可以在玻璃的两个界面上发生反射和透射。那么从第一个界面出来的光就可能是第一次反射形成的，也可能是第一个界面和第二个界面反射并且在第一个界面透射以后形成的。这样互斥事件形成的光能够相互抵消吗？如果不能，我们给相机镜头镀膜——这里把膜看作新的一层玻璃——干什么？

起来的问题，我们将来再讨论。在这里，我们假设总是可以用一个在当前能量尺度上已经不能再裂开的粒子来做我们的实验，而且实际上，我们可以随时在实验的任何一个中间步骤通过探测器来检查这些粒子在运动过程中是否还是一个整体或者说分成好几个部分[2]。

实验 1.2 (**相机镜头上的贴膜形成的光的干涉现象**)：镜头上的贴膜就是让光尽可能多地通过镜头之前的膜而不是被这个膜反射太多 [13]，然后，也降低在膜和镜头——可以是玻璃或者塑胶——之间的反射。现在，我们来关心前半部分——如图 1.3 所示的光过膜的两个界面——的过程。为了简单起见，我们用 Feynman 在《QED: The Strange Theory of Light and Matter》[12] 中的光过玻璃的例子——把光过贴膜改成光过玻璃。概念上两者没有区别。

为什么通过这样一个两个界面的反射就会增加光的透射呢？我们知道在每一个界面上都有一部分光被反射出去，那么也就是说，不增加这样一层额外的镀膜应该有更多的光进入到镜头。是这样吗？如果不是这样的话，从第一个界面反射的光肯定是以某种方式与从第二个界面反射的光相互抵消从而降低了反射出去的部分。那么这个相消的过程是如何发生的呢？经典波

[2]其实，这个观测会引入其他的问题，我们留到后面再讨论。

动力学给出了如下的解释。

考虑两列从太阳或者某个远处的单一光源（于是近似看成平行入射光）过来的光波，一列经过一次反射，另一列经过两次反射一次透射以后遇到一起。既然遇到一起，那么就相当于两列光波在同时驱动相遇的这个点的振动，于是也需要把两者叠加起来。于是，我们就得到了与水波的干涉数学上完全相同的表达式，于是结果也相同。在这个解释中，**把光波看成是两个部分分开传播然后相遇——也就是同时到达某点——是非常重要的，矢量相加也是非常重要的**。

小结一下，经典波动力学对于干涉的解释要么依赖于一个点的粒子的振动状态同时受多个驱动力驱动，要么依赖于一个波可以分成多个波然后让这个多个波相遇，从而相当于一个点的振动同时受多个驱动力驱动。同时受多个驱动导致了状态的叠加关系，而状态的叠加关系形成了干涉。可以想见，如果我们需要考虑一个整个空间中的任何时刻，都仅存在一个粒子的情况，那么上面的解释就有可能会出问题。这个时候，在一个点上的粒子受其它粒子驱动的情形不能用了，一个波分成多个再合起来的情形也不能用了。以后我们将讨论这样的整个空间只有一个粒子的令人困惑的情形。

下面我们再运用光的偏振来做另一个同样——将来考虑单粒子以后——让人困惑的实验。光是横波，振动方程和传播方向垂直，因此在三维空间中可以分解成两个正交的振动方向。因此，光实际上有传播和偏振两个自由度。在这里，我们主要关注偏振这个内部自由度。偏振片是一个非常有意思的仪器：其内部有一个由制作偏振片的材料和方式决定的特定方向，通过偏振片的光其振动方向必然和这个内部方向一致。那部分振动方向不一致的光，就被完全挡住了。至于如何实现以及实现这样的选择的机理是什么，我们不关心。为了以后语言上的方便，我们称透过去的光为“透射”光，被挡住的为“反射”光。在物理上，是否这个被挡住确实是由于反射还是什么其他物理过程，在此我们不讨论。

实验 1.3 (Dirac **的光过三块偏振片**)：找三块完全相同的偏振片来做如图 1.4 所示的三个实验。先拿其中的两块做实验 A：一手拿一块偏振片，把两块偏振片完全平行地放在面前，透过镜片看物体。观察是否能够看到东西，还是基本不能透光？再拿这两块做实验 B：一手拿一块偏振片，把两块偏振片垂直地放在面前，透过镜片看物体。观察是否能够看到东西，还是基本不能透光？接着取出第三块镜片，做实验 C：把第三块镜片以某个角度——与之前的镜片都不平行——插入到实验 B 的两块之间（这时候你可能需要别

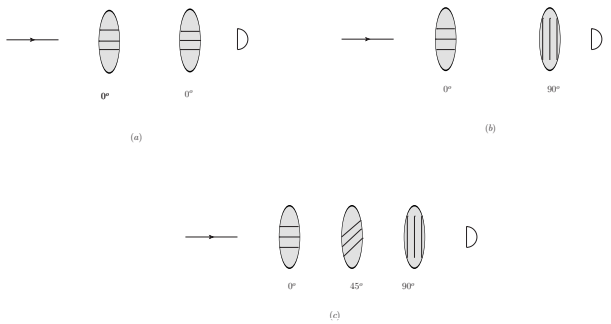


图 1.4: （a）. 光过两块平行偏振片（透过）。(b）. 光两块垂直偏振片（挡住）。(c）. 光过三块偏振片——最前最后两块相互垂直，中间的偏振片处于前后两者之间的某一角度（透过）。

人帮忙)。观察是否能够看到东西，还是基本不能透光？

实验结果是：A 通过两个平行镜片看物体，基本上没变（光的强度有变化，但是能够清楚地看到镜片之后的物体)；B 通过两个垂直镜片看物体，基本上看不到镜片之后的物体；C 增加一个镜片以后，之前不能看到的镜片之后的物体又能够看到了。

这个（将来我们会看到的）量子效应如此显著的实验，竟然是可以在家里做的，只要购买三个偏振片，例如偏振墨镜或者如图 1.5 的教具。

经典粒子理论的解释：第一块镜片选择让某一属性——其实是光子的特定偏振方向（例如硬币的向上态)——的粒子通过。如果第二块允许的方向完全与第一块平行，则光子不会再次被挡住——其性质已经是能够通过的状态（例如向上的硬币)，直接通过。如果第二块允许的方向完全与第一块垂直，则光子被挡住——其性质正好就是能够通过的状态的互补状态（例如向下的硬币)，完全不能通过。于是，我们了解了两个垂直的方向的光子正好就象向上和向下的硬币的关系一样。这个时候，我们再来增加第三块镜片。由于第一块和第三块——之前的第二块，已经把所有的可能都挡住了，因此，无论在中间如何改变光子（或者说硬币）的状态，都不可能有通过的光子。因此，经典粒子的模型不能解释这个实验。

为了强调这个实验多么的神奇不可理解，我经常用下面的类比：有一个屋子有两个门。前门只允许男人通过（进入这个屋子)，后门只允许女人通过（出这个屋子)。于是，在假设全世界只有男人和女人的前提下，没有人

图 1.5: 这个偏振片专门是开发了做教具的，非常便宜。

能够通过这个屋子——我们在后面观察不到人。这个时候，屋子里面增加了一道门，然后，我们就能够看到有人通过这个屋子了。这个是不可能的事情。屋子里面的那道门难道可以把男人变成女人，而且是本质上就变了而不仅仅是打扮成女人？我们称这个实验为“三道门实验”。

经典波动力学的解释：光波经过偏振片的过程可以用一个矢量投影来描述 $A_r = \vec{A} \cdot \hat{r}$，其中 $\vec{A}$ 是入射光子的状态，$\hat{r}$ 是镜片的方向，而且经过镜片之后的偏振状态是 $\hat{r}$，强度是 $|A_r|^2$。我们先不问这个经典波过偏振片的理论是从哪里来的，也不问为什么透过的光的强度和这个投影的计算有关系。于是，经过第一块镜片——记为 x 方向——之后，光子的状态就是 $\hat{x}$。如果第二块允许的方向完全与第一块平行，则光子不会再次被挡住——$1 = \hat{x} \cdot \hat{x}$，100% 完全直接通过。如果第二块允许的方向完全与第一块垂直——记为 y 方向，则光子不会再次被挡住——$0 = \hat{x} \cdot \hat{y}$，光子完全被挡住。如果有第三块镜片——记为 xy 平面上的某个角度 θ（$\theta \neq 0, \frac{\pi}{2}$）的方向，则通过前两块镜片的几率不为零，

$$\hat{x} \cdot \hat{r}_\theta = \cos(\theta) \neq 0. \tag{1.2}$$

过了中间那个镜片之后，其状态为 $\hat{r}_\theta$，于是过最后那个镜片的几率也不为零，

$$\hat{r}_\theta \cdot \hat{y} = \sin(\theta) \neq 0. \tag{1.3}$$

我们看到经典粒子不能解释实验现象，但是经典波似乎通过矢量投影——

通过 $\hat{r}$ 方向的出射光是 $\hat{r}$ 方向的偏振，其强度是 $|A_r|^2$，其中 $A_r = \vec{A}\cdot\hat{r}$——能够解释这个实验的现象。

实际上，确实可以运用经典的绳子上的波来实现这个实验。

实验 1.4 (**绳子上的波过三个缝**)：一根伸展开的绳子，记绳子伸展方向为 z 方向，水平和竖直方向分别为 x 和 y 方向。我们抖动绳子的一端，观察绳子另外一端的振动。当绳子上没有任何其他东西的时候，基本上主动的一端如何抖动则被动的一端就如何振动。现在我们来加上几个狭缝，如图 1.6(c)：这些狭缝在一个方向上限制振动传播到狭缝的另一面（绳子在那个方向上遇到边界），而允许另外一个方向上的振动传播到狭缝的另一面。我们作如图 1.6(a) 和 (b) 的实验，A：第一个狭缝沿着 x 方向，第二个沿着 y 方向；B：在实验 A 的基础上，再中间插入一个沿着 $x-y$ 平面 45^0 方向的斜着的狭缝。

在这里狭缝实际上就是“绳子波的偏振器”。在我们的实验中实际上我们用了如图 1.6(c) 的一个烧火的钳子来当做这个偏振器。我们看到了相同的实验现象。其实，和偏振片相对应的实验 A 现象也是一样的。那么，是否能够用相同的理论模型来解释呢?

经典波的解释：上面关于 Dirac 光过三个偏振片的经典波的解释完全就解释了这个绳子上的波过三个狭缝的实验。在那里最关键的就是经典波的情形下矢量投影

$$A_r = \vec{A}\cdot\hat{r} \tag{1.4}$$

当一个原来在某一个方向上的振动企图通过这个绳子偏振器的时候，可以想见大约矢量投影公式 (1.4) 还是正确的：振动方向和强度合起来是一个矢量，当其中某一个方向的振动受到限制从而只能够在另一个方向振动的时候，其在能够振动方向上的强度大约应该是通过在那个振动方向上做一个投影来得到。于是，A, B, C 的实验现象就得到了解释。

这个时候，我们可能觉得，光的行为，跟绳子上的波差不多。这个时候，我们利用绳子的模型再来思考，矢量投影公式 (1.4) 到底表示了什么含义。绳子上紧密与小缝相连的那个点的振动，受到小缝所在的点的驱动，本来应该是随着驱动点大概在同一个方向上运动。但是，当受狭缝限制以后，只能那个狭缝允许的方向上的振动被激发起来。这个时候，我们必须依赖于绳子上有很多个点，每个点都是一个可以做振动的物体这个事实，加上 Newton 定律，来理解矢量投影公式 (1.4) 的含义。

(a)

(b)

(c)

图 1.6: (a). 绳子上的波过两个垂直狭缝（挡住）。(b). 绳子上的波过三个狭缝——最前最后两个相互垂直，中间的狭缝处于前后两者之间的某一角度（透过）。(c). 实验用的火钳，非常便宜。

于是，在这个实验和解释中，我们看到：第一，矢量投影确实解释了实验现象；第二，矢量投影的含义依赖于多个可被激发出振动的小单元来理解，而这个理解的基础是 Newton 定律。然而，当将来在下一节中，我们考虑单个光子的传播的时候，它不再是介质上的波了，不能用 Newton 定律通过介质上的点的相互作用力导致传播导致矢量性来解释了。于是，第三，经典波动力学不能解释单光子的 Dirac 光过三个偏振片的实验。更多的讨论见下一节[3]。由于经典波理论能够解释以上实验行为的原因在于经典波满足叠加关系或者矢量投影，可分解，可叠加，我们可以自然地猜测，没准，这个叠加关系和矢量投影的数学结构，也就是状态的数学描述是矢量这一点，将来在量子系统的理论里面，也会保持下来。

现在，我们再来看另一个经典力学的例子。

实验 1.5 (**子弹的双缝实验**) ：如图 1.7 所示，一把准确度有问题的固定好的机枪在点射同一个地方，但是由于准确度有问题，射出的子弹以一定的概率 p ($(1-p)$) 到达目标上（下）方 $\frac{d}{2}$ 的地方。然后，在 $\pm\frac{d}{2}$ 的地方各有一个奇怪的小缝，这个小缝会把这个子弹在比较小的仰角 θ_0 的范围内以**均匀**（为简单计，我门假设打到屏上以后，子弹在一个小范围内均匀分布）的概率弹射到距离 L 之外的屏幕上。每一个打在屏幕上的子弹会在屏幕上留下一个荧光的斑点。实验中，我们观察屏幕上子弹留下的痕迹。

我们通过分析单个小缝的结果然后组合两个小缝的行为来分析这个实验。对于过第一个（第二个）缝然后出来的子弹，其在屏幕上形成的痕迹以 $\frac{d}{2}$ ($-\frac{d}{2}$) 处为中心宽度为 $2L\tan(\theta_0)$。当 $2L\tan(\theta_0) < d$ 的时候，两个斑点没有重叠的部分。就是分散的两个独立的斑点。可以想见，当 $2L\tan(\theta_0) > d$ 的时候，**在中间重叠的部分，由于能够获得来自于任意一条缝的子弹，其概率是两者**

[3]因此，如果有人给你展示这个 Dirac 的光过三偏振片实验，并且宣称他能够很好地理解和解释这个实验，而且这个实验展现了量子系统的独特行为，那么，他基本上是错的。第一、在不考虑单光子实验的情况下，这个展示和上面那个绳子上的波的实验是完全一样的，因此，展示的是波的特性——也就是矢量投影的解释能力，而不是量子系统的特性。第二、矢量投影在经典波的情形下为什么能够成立，才是需要思考和理解的重点，而不是先接受它然后用来解释现象。换句话说，**波的特性不是量子系统的特征，无介质的波的单元——单个粒子——也具有波的特性才是量子系统的特征**。如果有人给你演示这个实验，并且提醒你注意单光子的时候如何解释，而且告诉你解释起来有问题，才表明这个演示的人真的注意到了这个实验揭示的量子力学特性。经典波遵循的跟叠加关系类似的矢量投影解释了完全被两个具有互补关系的镜片相互挡住的光又一次重新出现的问题，但是将来我们考虑的对象是单个光子，没有介质，不能用基于 Newton 定律的经典波的解释。于是，看起来，一个单个的粒子要具有波的特性。那只能是几率或者几率幅的波了。沿着这个思路，我们还需要说明，几率的波不能解释量子系统的行为，只有几率幅的波才行。于是，我们才考虑下面的单电子过双缝、which-way 实验的问题。

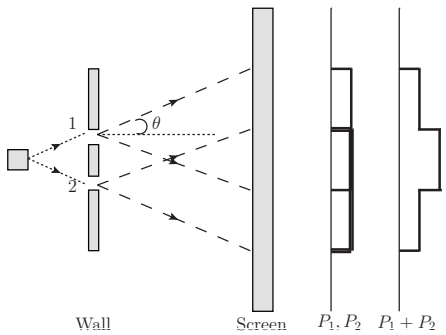


图 1.7: 子弹经过两个小孔。孔间距为 d，档板和屏幕之间的间距是 L。机枪覆盖仰角是 θ。后面的两个图是结果示意图。分别是两个小孔独自的结果和合起来的结果。在这里实际上我们忽略了一个“重新归一化”的计算细节，例如实际上应该是整体上除以 2 也就是 $\frac{1}{2}(P_1 + P_2)$。

之和，

$$P(z) = P\left(z\middle|\frac{d}{2}\right)P(\frac{d}{2}) + P\left(z\middle|-\frac{d}{2}\right)P(-\frac{d}{2}). \tag{1.5}$$

我们不知道这个实验是不是真的做过，如果做过小缝的行为是不是接近实验中的实际情况。但是，通过经典概率论的概率叠加原理（有的地方两个缝过来的子弹都能到，有的地方只能接收到一个缝过来的子弹）我们可以得到如下的大致结果，而且这个结果有一个重要的特征：在任意一个小缝的子弹辐射范围内，在两个小缝的情况下仍然有子弹的痕迹。如果小缝的行为比我们假设的简单情况复杂，例如均匀性的假设会不太对，实际的轨迹就会不完全一样，但是这个特征是不变的。一个随机变量的平均值只能够取值于其所有的可能取值之间，不能比其中最小的还要小。

1.2 单个电子或者单个光子的实验

上一节，我们提到，对于介质上的波，由于 Newton 定律的成立，并且可以考虑多个点的运动对某个点的驱动，**矢量叠加关系**和矢量投影能够解释上一节中提到的干涉现象，但是，在整个空间只有一个粒子并且没有介质的情况下，它们的成立的基础——介质上的粒子通过满足 Newton 定律的力来影响相互运动——似乎就没有了。另一方面，**概率叠加原理**——一个事件有多种互斥情况发生的可能，则其整体效果是这些不同可能的叠加——不

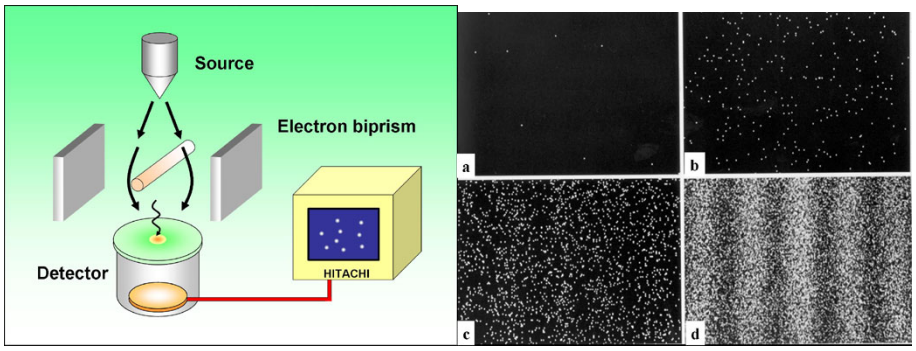


图 1.8: 来自于一个电子枪——绝大部分时间里面，这个电子枪每次发射一个电子——的电子经过一个障碍物隔开的两个小缝，到达屏幕上。电子在屏幕上的落点被记录下来，形成右侧的图样。注意和子弹的理想实验的区别，并思考其原因。装置示意图和实验结果图来自于 hitachi 公司的研究亮点网站上记得去获取授权。在这个网站http://www.hitachi.com/rd/portal/highlight/quantum上你还可以看到视频。这个工作由 Akira Tonomura （外村彰）及其合作者完成 [14]。

需要多粒子的图像，粒子一个一个过去的事件，仍然可以用概率叠加原理来描述。其所需要的仅仅是粒子的行为的多种可能，而不是多个粒子——多个粒子之间相互作用之类的。但是，这个概率叠加原理又不能展现出来矢量叠加关系和相干性：单缝情况下就有粒子到达的地方在打开双缝的情况下竟然会没有粒子到达。现在，我们来看一下量子系统的真实实验的结果。在这里，我们确实考虑在整个空间只有一个粒子的情况。我们来看看概率叠加原理能不能解释实验结果。

实验 1.6 (**电子的双缝干涉实验**) ：如图 1.8 所示，一个电子射线枪，射出与上面的情况完全相同的子弹——电子（其中，小缝的行为可能不是子弹的理想实验中的均匀分布的情况），我们观察电子在屏幕上留下的痕迹。我们发现，见图 1.8 中的实验结果和图 1.9 中的示意图，在某些地方——在一个小缝就能辐射到的范围内的某些地方，当两个小缝都打开的时候——电子几乎没有留下痕迹。

我们发现了一个新的特征：两个小缝都打开以后，电子打到某些地方的几率比任何一个小缝独自打开的时候都小。回到豌豆射手的类比，这个现象就好

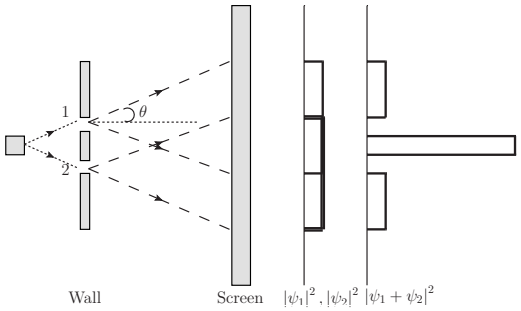


图 1.9: 基于 θ 角范围内均匀出射的干涉现象示意图。(a) 来自于一个单电子电子枪的电子经过一个障碍物隔开的两个小缝，到达屏幕上。考虑通过其中一个小缝的结果和同时打开两个缝的结果的差别：有的地方变得几乎没有电子到达，有的地方电子到达的几率变得非常大。而不仅仅是概率叠加。联系到图 1.7，在那里是遵循概率叠加的子弹。在这里实际上我们同样忽略了一个“重新归一化”的计算细节。

像是说，有的地方一个豌豆射手就能够做到射击覆盖，但是当增加一个豌豆射手的时候，那个地方反而不在射击范围之内了。这是不可思议的事情。这样的话，你玩植物大战僵尸游戏的时候就要更加小心了。

看起来，好像是，过两个小缝的电子之间存在着某种相互作用。但是，我们可以让电子枪点射，每次只能够射出一个电子。这样的话，相互作用的解释都非常不可能了。在整个我们关心的空间，每时每刻只有一个电子存在，它跟谁相互作用？还有一种可能：一个电子也能够象经典波（例如前面的光波或者水波）一样以某种方式劈开两半，然后同时过双缝，从而实现自己跟自己相互作用。我们不能否认这个可能性。一个简单的检验方法就是在两个小缝的地方各自放置一个探测器，然后看一下是否在某一个时刻可以在两个探测器上同时都探测到半个电子。实际上，这个实验对于光子的情形是做过的 [15]，我们从来没有探测到过半个光子或者两个光子。将来我们会知道，这个额外的探测器有可能带来一个额外的问题：先经过小缝然后再探测器以后的电子的状态有可能就不再是原来仅仅经过小缝而不经过探测器的状态了。因此真的靠这个额外的探测器来回答电子如何过两个小缝是有问题的。不过，在这里，我们关心的事情不是电子如何过双缝，而是是否电子在有的时候会分开两半。或者说，我们把整个空间看作是很多很多个小缝合起来的，难道说电子都需要分成这么多个部分？电子又如何知道需要分成

几个部分呢？所以分成几个部分一起过去的假说是不太合理的。当然，我们面对的对象是有独特行为的量子客体，我们不清楚哪些不合理是可以接受的。如果我们没有任何其它方式可以建立电子行为的数学模型和物理图景，那么不合理之处也是必须接受的。

幸好，我们还有别的选择。我们沿着这个想法继续走下去：电子在这个过程中没有与任何其他的东西发生未知的相互作用，除了电子从电子枪里面发射出来，然后在小缝处发生有可能是复杂的——但是行为是完全已知的——相互作用之外。例如之前我们所假设的在某个角度之内均匀出射，这个均匀出射（或者其他的概率分布的出射）是如何产生的我们暂时不管。也就是说，在仅有上（下）方小缝打开的时候，电子的痕迹如图 1.9$|\psi_1|^2$（$|\psi_2|^2$）所示——实际的图样可以和这个“在各自小孔对应的范围内均匀分布”的图样不一样，跟小缝的细节有关——但是，在两个小缝都打开的时候，成了图 1.9 中的 $|\psi_1+\psi_2|^2$ 的样子。关键就是：这个图中的 $|\psi_1+\psi_2|^2$ 绝对不是和 $p\,|\psi_1|^2+(1-p)\,|\psi_2|^2$ 这样的概率叠加能够得到的，包含任意的 p——概率性叠加永远也不会出叠加以后的值比叠加之前的值都小的情况。经典粒子的力学，甚至允许概率的经典力学，不能解释电子的双缝干涉现象。经典的波的力学，当考虑电子每次都是完整的一个而且是仅有一个的时候，也不能解释电子的干涉行为。如何理解这个现象？

其实，我们也可以控制前面那个光过玻璃的实验，让光的强度足够小[4]，以至于任何时刻，如果测量的话，整个光路中不会有一个以上的光的能量单位（将来我们会知道这个能量单位就叫做光子，现在我们就先在最小能量单位的意义上用这个名词）。这个时候，我们同样发现，基于把光波分成几份再合起来的经典波动力学的解释也不能用了。如何在单光子的情形下解释光子过玻璃然后反射部分相互抵消的现象？任何时刻，唯一的一个光子过第一个界面的时候可能反射也可能折射，第一种情况发生了反射，反射光就开始跑路了，注意光子跑得很快的，故事暂时结束。第二种情况，发生了折射，然后遇到第二个界面，假设这次发生了反射，光子又回到第一个界面，接着发生了折射。这些过程是需要时间的。这个时候，第一种情况下的光子，可能都已经跑到月球了。经典波动力学的数学表达式——注意——是能够解释干涉现象的，就是理解上需要把光波分成多个部分然后同时到达，而这一点，考虑单个光子的时候我们做不到。除非单个光子也是分成好多个部分来

[4]其实我们需要专门的技术来保证每次空间中只有一个光子，见例如 [15] 中所用的取一对下转换光子中的一个。

过玻璃的。我们说了，这个可能性以及加探测器来检验这一点的问题，以后再讨论。

现在，我们再按照经典粒子的图景来描述一下这个光子过玻璃的实验。我们总共只有一个光子，按照子弹的力学过程，第一个界面发生反射的可能性如果实现了，这个光子就跑掉了。然后在第二个界面可能再发生反射，于是透射的光波应该是两次反射之后剩下的那部分。也就是说光子发生透射的概率是在两个界面上都发生了折射，于是透射率是 $0.96 \times 0.96 \approx 0.92$。然而，我们的透射率大于 0.92 可以接近 0.99[12, 13] 的实验结果告诉我们，这两种可能一定要通过某种方式相互影响。可是，怎么相互影响呢？整个空间只有一个光子，如果第二种情形发生，那么第一种情形就不会发生：光子如果在第一个面就被反射了，怎么会在第二个面再一次被反射呢？因此，就算我们“理论上”想把这两种情形叠加起来，两种根本不会同时发生的事情要相互影响，也不行啊。我们仅仅能够按照经典概率论把两种可能的结果按照概率加起来。于是，不会出现被反射掉的光比其中任何一个都要小的情况。

经典粒子的力学和经典波的力学都不能解释这个光子过玻璃的实验。如何理解这个现象？

同样地，当我们考虑单光子过三个偏振片实验的时候，我们的矢量投影的基础——一个单元的受限振动由前一个单元驱动——就没有了。这个时候逼迫我们考虑完整的单个的光子的通过或者被挡住的行为。让我们跟着这个光子开始旅程：首先经过第一个偏振片，假设没有被挡住，那么它的偏振状态肯定是 x 方向。接着，如果遇到 y 方向的偏振片自然就完全被挡住了，如果遇到 x 方向的偏振片自然就完全过去了。关键是遇到 $\theta \neq 0, \frac{\pi}{2}$ 的时候怎么办？我们知道为了能够解释实验结果，我们需要在这个时候用矢量投影。可是这里的矢量投影好象是在说，就算一个已经被投影成为 x 方向分量的状态，里面竟然有 θ 方向的分量，而那个被投影成为 θ 方向的分量的状态，则竟然也还有 y 方向的分量。将来我们会看见，这个——“**一个看起来单纯一个方向的矢量，竟然，可以看作其它方向的矢量的叠加（也就是有那些方向的分量）**”——是量子系统的理论的特征，就算对于单个粒子。这个正是状态的叠加原理。

下面，我们再用两个类似的实验，来展示量子客体的行为的独特性。

光子分束器可以把光子以一定的概率反射或者透射。这里为了简单计，我们把反射和透射的几率设定为 0.5。在我们后面的习题中，我们会用到这样的仪器。光子偏振分束器（也称作极化分束器、偏振分束镜等）是一个能

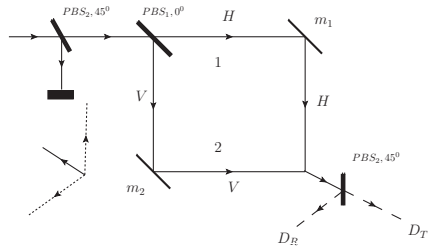


图 1.10: 光子 which-way 实验装置示意图。其中最后一步把两个不同路径来的光又合起来然后进入偏振分束器在实验中需要一个额外的仪器来完成。在这个示意图里面我们只需了解能够做到这样的合起来就可以了。注意偏振分束器的内部方向和图中的仪器的摆放方向不是一个东西。

够根据光的偏振状态来决定让光透射方向的仪器。这个仪器有一个内部的方向，这个方向是由制作该仪器的材料决定的。我们知道光波是横波，有两个可能的与传播方向垂直的振动方向。一般地，一个光波的电场分量可以写做，指定传播方向为 $\hat{z}$ 方向，

$$\vec{A} = A_0 \cos(kz - \omega t + \phi_0)\hat{x} + A_1 \cos(kz - \omega t + \phi_1)\hat{y}. \tag{1.6}$$

其中 $\hat{x}, \hat{y}$ 就是与 $\hat{z}$ 方向垂直的平面内的两个正交的方向矢量。我们这里不妨就把 $\hat{x}$ 指定成偏振分束器的内部方向。这个时候我们让上面的一般的沿着 $\hat{z}$ 方向传播的光经过这个分束镜，我们就会发发现 $\hat{x}$ 方向的光从一个角度透射过了这个分束镜，而 $\hat{y}$ 方向的光从另外一个角度从这个分束镜透射了（或者当是偏振片的时候是反射、完全挡住。这个如何实现的细节在这里不重要）：偏振分束器透射与其内部方向相同的光，反射与其内部方向垂直的光[5]。当然，如果让经过 $\hat{x}$ 方向分束镜透射的光再经过一个 $\hat{x}$ 方向分束镜，那么这个光会完全透射过去。注意，如果让一个透射过 $\hat{x}$ 分束镜的光当遇到另一面 $\hat{y}$ 向分束镜的时候，它会被完全反射，因为这个时候振动方向为 $\hat{x}$ 的光在后者看来正好是它的内部方向的垂直方向。下面，我们要用这样的偏振分束器来做一个有趣的实验。

[5]在这里其实垂直方向的光是从另外一角度透射出去。但是，按照我们之前的“透射”和“反射”术语的约定，由于牵涉到偏振片和一般分束器的问题，在这里，我们把偏振分束器中垂直方向的光的透射、偏振片的反射或者挡住、一般光子分束器的反射一概称为“反射”。再强调一下，根据仪器的不同，这个“反射”有可能是真的反射或者从另外一个角度透射

实验 1.7 (**光子** which-way **实验**)：如图 1.10 所示，一个光子经过第一面内部方向为 45^0（见图中所示的坐标系）偏振分束器之后，只允许透射光过去，反射光被完全挡住。然后这个光子继续经过一个 0^0 的偏振分束器。经过这个分束器的光子可能走两条路径。不管走哪一条，它都会被反射会到同一个位置（右下角），然后被引入到第三个内部方向为 45^0 偏振分束器。问：分束器之后的探测器 D_T 和 D_R 上都会有接收到光子的可能吗？

这个实验的结果[6]是只有 D_T 上有光子到达。

每次只打出一个光子，因此把光子分成很多份的经典波动力学的解释是不能用的。我们来看看经典粒子的力学加上概率论的解释。一方面，经过 45^0 方向的分束镜以后的光子经过 0^0 方向的分束镜分开可能的两条路径。然后每一条路径又被反射合起来，所以看起来这个 0^0 方向的分束镜的效果完全被两个反射镜消掉了。于是，在到达最后一个 45^0 方向分束镜的时候，光子的状态与刚刚出第一个 45^0 方向分束镜的时候一样。于是，在最后的 45^0 的分束镜上完全通过没有反射，只有 D_T 可以接收到光子。另一方面，我们可以问经过 0^0 的偏振分束器之后，光子是从哪一条路径过去的。这个时候，如果我们假设光子是从上面的路径过去的，代表 0^0 方向的偏振态。于是，这个偏振态遇到最后一个 45^0 方向的偏振分束器会在两个可能的方向上 D_T 和 D_R 有输出（尽管可能每次还是仅仅输出在一个方向上，多次合起来就会有两个可能）。同样的，如果我们假设光子是从下面的路经过去的，也就是 90^0 的状态，也是两个可能的输出。于是，按照概率叠加原理，这个事情的发生有两种可能——从上面或者是从下面的路径过去的，于是整个事件的结果就是这两个可能的概率平均。既然两者单独来说，D_T 和 D_R 上都可能有输出，那么，必然，其概率平均的结果是 D_T 和 D_R 上都有输出。可是，我们的实验结果是只有 D_T 能够接收到光子，而且前面第一个逻辑也给出这样的结果。这个现象怎么解释？我们的后一种解释哪里错了？我们区分光子可能走的两条路径，这件事情错了？

尽管我们已经强调经典波动力学的图景——由于需要把光波分成几个部分然后让这几个部分同时到达——在这里是不能用的，我们仍然把经典波动力学的解释写在下面，提供一个更清楚的说明。这个说明是针对学习过光学的读者来写的。我们直接利用数学表达式来更好地说明白。经过一个偏振

[6]实际的实验在最后一步探测的不是“哪一个方向上有粒子”，而是探测是否两个路径上过来的粒子“是否会出现干涉条纹”[16]。如果有干涉条纹则表示路径信息消失了，就不能问“粒子到底从哪一条路上过来”了。

分束器以后光子的偏振状态是 $\vec{A}_1 = A_0\cos(45^0)\hat{V} + A_0\sin(45^0)\hat{H}$，然后在经过第二个偏振分束器之后，路径 1 上的光子的状态是 $\vec{A}_2^{(1)} = A_0\cos(45^0)\hat{V}$，同时路径 2 上的光子的状态是 $\vec{A}_2^{(2)} = A_0\sin(45^0)\hat{H}$。经过两面反射镜到达第三面分束镜之前的状态是 $\vec{A}_3 = A_0\cos(45^0)\hat{V} + A_0\sin(45^0)\hat{H} = \vec{A}_1$，也就是 45^0 的偏振态。于是，最后完全通过第三个分束器。注意，这里在做把分开以后的光又合其来的运算的时候，我们默认的思考是：每一条路径上都走了一部分的光，于是把两个部分的光合起来自然就是两个代表其状态的矢量加起来；或者是每一条路径代表了对同时到达的那个点的振动的一种驱动方式，于是两种驱动方式的效果就是把两者的代表矢量叠加起来。这里，我们非常清楚地看见了分开两部分又同时到达的一个默认的物理图景。换句话说，对于矢量 $\vec{A}_1$ 的解释，我们默认是多个光子的整体矢量形式，其中的一部分遇到分束器以后可以走路径 1 于是成了状态为 $\hat{V}$ 的光子，另外的走路径 2 于是成了状态为 $\hat{H}$ 的光子。也就是说我们假设这些个光状态的矢量描述是多光子的状态，或者说一团光的状态，而这个一大团的光可以看作是一部分由路径 1 驱动，一部分由路径 2 驱动，于是整体的状态是两个驱动的效果的叠加。这个基于经典多源驱动的波的多光子图景是很容易接受的，可以不加思考地接受的。

在我们前面所展示的实验中，利用单光子来做实验就是挑战的这个多光子图景。也就是说，上面的基于经典波动力学的解释结果与实验符合，但是不能用于描述我们的单光子实验。经典概率论倒是可以用在单光子上，可又不能给出和实验结果一致的理论结果。两个理论都不对，怎么办？新的理论应该长什么样？在我们构造能够解释这些现象的理论之前，我们再来看一个电子的类似的实验。

有一种如图 1.11 所示的叫做Stern-Gerlach 装置的仪器（其内部主要是某个方向的一个磁场，不过现在我们不关心这个仪器的细节）有一个内部方向（其实由其内部的磁场决定，这个方向就好像是偏振片的内部方向一样），会按照电子的状态来改变经过这个仪器的电子的运动方向。这一点就好像是光子偏振分束器一样。例如，通常一束电子过这个仪器之后会在屏幕上产生两个斑点。这就是著名的Stern-Gerlach 实验。我们把上（下）方的斑点对应的路径称为向上（下）态的路径（以后我们会知道，实际上这个“上下”的命名不太准确，是倒过来的，应该是“下上”。不过，在这里无所谓）。如果我们把其中的一个方向的路径盖住，自然我们就得到一个斑点——对应着没有盖住的那条路径。现在，我们把两个内部方向相同的这样的仪器连起

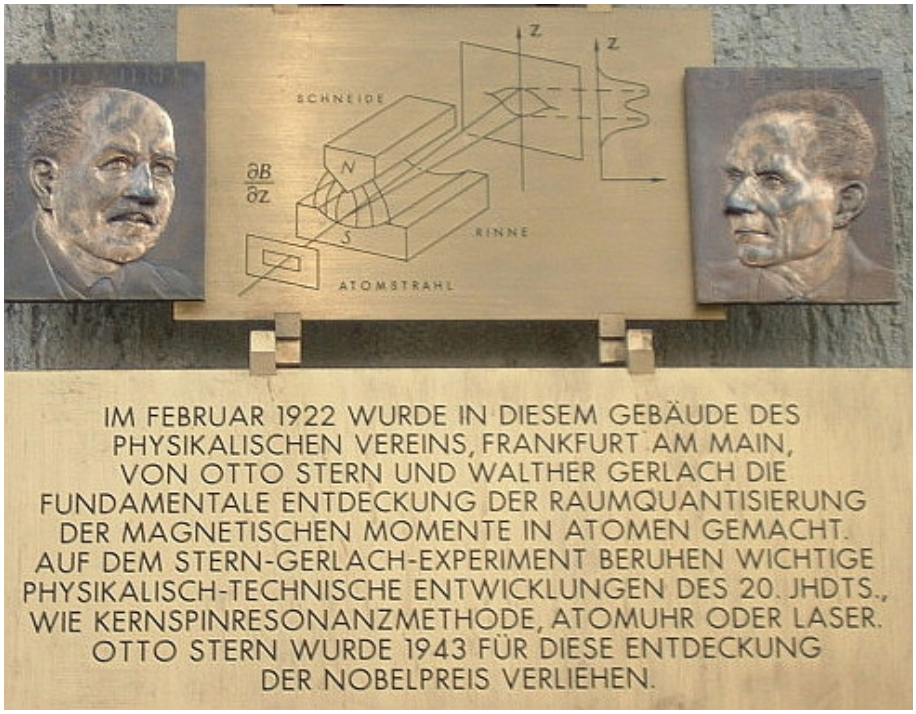


图 1.11: 来自于Stern和Gerlach的自旋Stern-Gerlach 装置示意图。图片来自于 Wikipedia 页面https://en.wikipedia.org/wiki/Stern-Gerlach_experiment。

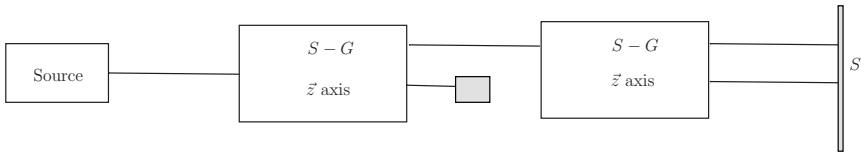


图 1.12: 自旋经过一个 Stern-Gerlach 装置——其内部就是一个磁场——之后挡住向下的输出，这样从装置出来的状态就是第一个装置的向上方向。接着让这个输出的自旋再一次经过同样方向的装置——得到仅有一个向上的输出结果。

来，如图 1.12，我们发现，向上态的路径经过第二个仪器之后还是向上态的。也就是说，这个向上态，看起来好像是经过这个仪器之后不变的状态——我们称其为这个仪器的本征态。向下态也是本征态。这个仪器以及这个电子在这个实验中的现象基本上可以通过一个随机硬币来解释。一个一般的硬币可能向上也可能向下，所以把这样的实验重复很多次以后，我们可以得到两个斑点。如果我们通过把向下的路径挡住来把向上的硬币挑出来，然后再让它进入仪器，自然只能得到向上的路径的斑点。这个不奇怪。

我们把电子的向上和向下态成为电子的自旋状态。我们已经看到电子具有空间和自旋两个自由度。在这里，我们主要关注电子的自旋自由度。

通过前面自旋的两个实验——原始的自旋通过单个磁场和两个同方向磁场的实验，我们已经看到，只需要把自旋看做一个具有两个离散状态的硬币，就可以理解这个实验。有的人喜欢这样说，原始的 Stern-Gerlach 实验证明了我们需要量子力学来描述这个世界。还有人的认为Einstein的光电效益实验、黑体辐射现象和Planck的黑体辐射公式也证明了我们需要量子力学来描述这个世界。当然，从量子力学发展的历史来说，这三个实验确实大大推动了量子力学的发展。但是，从是否真的体现了什么是量子性——完全不同于经典粒子和经典波的量子性，也就是状态的矢量叠加原理——对于描述这个世界的必要性来说，三个实验都是不够的。可以说，这三个实验揭示了：描述世界需要离散变量，而没有到达一定要矢量叠加原理的层次。Stern-Gerlach 实验表明同一束粒子出来可以分裂成两束，而且是分开的两束，因此后面肯定对应着一个两状态的变量。光电效益表明光子的能量是一份一份的，在保持光子的频率相同的条件下增加光子的数量也就是光强度，不能激发光电效益。原则上，可以做出来这样的实验：每次出射一个某种频率的光子（不再可分的光的最小单位），只要频率合适，仍然可以激发光电效益。于是，光子的能量，从这个实验我们推断，应该看做是一份一份的。黑体辐射对实验现象的解释依赖于把光子的能量（E）看成一份一份并且和频率（ν）的函数关系是 $E = h\nu$，加上Boltzmann热平衡分布（$p(E) = \frac{1}{Z}e^{-\beta E}$）[7]，就有了Planck的黑体辐射公式。其中的 h 称为Planck 常数。从连续的能量、连续的状态到离散的能量、离散的状态这一步的跳跃是非常具有创造性的，因此，这三个实验在量子力学的发展过程中的意义是非常大的。但是，它们离真正说明“状态的矢量叠加原理”的必要性还有很远。[8]

[7]实际上，还要把这个分布函数用到离散能量的粒子上去得到 Einstein-Bose 分布。这里就不介绍了。有兴趣的读者可以选择任何一本《统计物理学》的书来看。

[8]这一段来自于某一天和狄增如的讨论。本书里面还有很多地方来自于和其他老师，尤其是裴寿镛，或

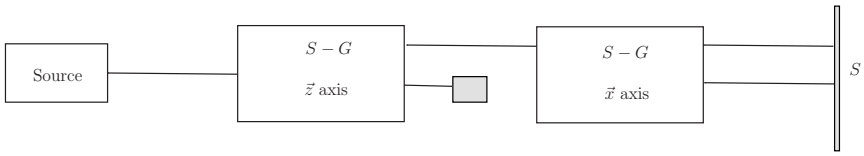


图 1.13: 电子过两个不同方向磁场实验装置示意图，先过 z 方向，接着 x 方向。

现在，我们来做一些稍微更有趣的实验，用来显示引入“状态的矢量叠加原理”的必要性。

实验 1.8 (**电子自旋过两个** Stern-Gerlach **仪器**) ：*我们让 $\hat{z}$ 方向的仪器出来的向上态的电子经过一个内部方向为 $\hat{x}$ 方向的仪器，如图 1.13，打到屏幕上。问：屏上有一个还是两个斑点？*

我们让从内部方向为 $\hat{z}$ 方向的仪器出来的向上态的电子经过一个内部方向为 $\hat{x}$ 方向的仪器，我们发现我们还是会得到两个斑点，而且单次实验只能得到一个斑点。甚至把 x 方向换成是一个不是 $\hat{z}$ 的一般的方向 $\hat{r}(\theta,\phi)$ 也行，结果也一样：一般来说多次实验的结果是两个斑点，但是每次单次实验只能得到两个斑点中的一个。注意，是两个分开的斑点，不是整个区域弥散开来的。这个特征很重要。一会儿，我们要用这个 $\hat{z}$ 方向向上的电子过 $\hat{x}$ 方向的仪器还会得到两个斑点的事实来做下面的实验。

注意，根据前一个实验的结果，在任何方向 $\hat{r}$（只要 $\hat{r}\neq\pm\hat{z}$）上做实验，我们都得到两个结果，记为 $|\uparrow_{\hat{r}})$ 和 $|\downarrow_{\hat{r}})$ 状态，分别对应落点在屏幕上部和下部的两个实验结果[9]。

这个的实验结果还是有两个输出。这件事情非常非常地不平庸：一个处于 $|\uparrow_z)$ 的电子，其状态如果从 $\hat{x}$ 方向来看，竟然，仍然包含两种可能 $|\uparrow_x)$ 和 $|\downarrow_x)$。这一点，我们粗略的记作，

$$|\uparrow_z) = |\uparrow_x) \oplus |\downarrow_x) . \tag{1.7}$$

者学生甚至其他朋友的讨论，但是当时没有记下来，也就搞不清楚了。身边有一群对基本概念的理解感兴趣的老师是一件幸福的事情。

[9]实际上这个记号有一点点问题，例如应该是出射方向和内部方向反过来的对应关系、纯态和混合态记号需要区分。这里的记号不是以后要学的 Dirac 记号。不过，这无关紧要，等到以后学到量子力学的 Dirac 记号再说。

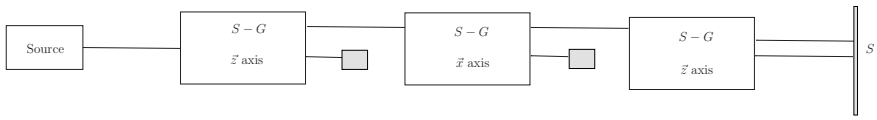


图 1.14: 电子过三个方向磁场实验装置示意图，先过 z 方向，接着 x 方向，然后 z 方向。

关于这个 $\oplus$ 的运算，这里，仅仅是“某种方式合起来”[10]的非常粗糙的意思。将来我们会看见，这个运算有两种可能的形式：概率叠加和矢量叠加。类似地，我们可以把实验装置的顺序倒过来，先 x 然后 z 方向，或者我们把参考系的方向重新定义一下（x 叫做新的 z，z 叫做新的 x），我们得到

$$|\uparrow_x\rangle = |\uparrow_z\rangle \oplus |\downarrow_z\rangle\,. \tag{1.8}$$

现在建立了这个状态之间的关系之后，我们来看下一个实验。

实验 1.9 (**电子自旋过三个** Stern-Gerlach **仪器**)：如图 1.14 所示，我们让 $\hat{z}$ 方向的仪器出来的向上态的电子经过一个内部方向为 $\hat{x}$ 方向的仪器。出来的向上态的电子再次经过 $\hat{z}$ 方向的仪器，打到屏幕上。问：屏上有一个还是两个斑点？

这个的实验结果还是有两个输出。这件事情非常非常非常地不平庸：一个 $\hat{z}$ 方向向下态先被挡住的电子，经过 $\hat{x}$ 的某种操作——其实就是挡住 x 方向的向下态（向上态结果也一样），竟然，再次出现了 $\hat{z}$ 方向向下态的可能。关于这个实验到底有多么不可思议，我们在 Dirac 的光过三个偏振片的实验中已经讨论。在此不再重复，仅再一次提一下结论：试验结果强烈暗示有可能 $|\uparrow_x\rangle$ 态里面包含 $|\uparrow_z\rangle$ 和 $|\uparrow_z\rangle$ 态，以及反过来 $|\downarrow_z\rangle$ 态里面包含 $|\uparrow_x\rangle$ 和 $|\downarrow_x\rangle$ 态，而且这个包含很可能是突破经典概率论的。我们再来看另一个实验。

实验 1.10 (**电子自旋** which-way **实验**)：这个理想实验来自于Ballentine的《Quantum Mechanics – a modern development》[8]。如图 1.15 所示，我们

[10]如果我们把 6 面的色子看作类似的 x, y, z 方向的向上和向下态，我们是不是会有类似的“某种方式合起来”的关系呢？例如 z 方向向上态可以看作“包含”x 方向向上和向下态？我们从来没有觉察到过这样的事情，因为在测量到 z 方向的向上态之后，如果我们马上再做一次观测，我们绝对不会见到 x 方向的向上或向下态。因此，我们也可以看到，上面的现象——测量完了一个方向得到结果接着测量另一个方向并且也能得到结果——是非常独特的，经典中不存在的。

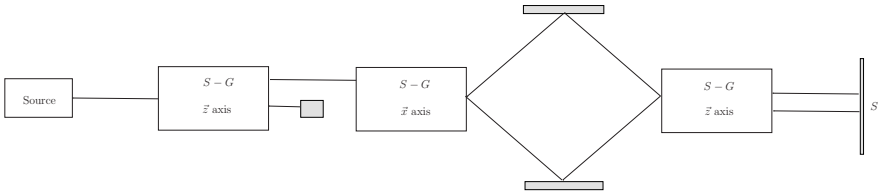


图 1.15: 电子 which-way 实验装置示意图

让 $\hat{z}$ 方向的仪器出来的向上态的电子经过一个内部方向为 $\hat{x}$ 方向的仪器，然后不打到屏幕上，而是通过另外一个神奇的仪器（就像光的反射镜）把可能已经分开成 $\hat{x}$ 两个方向的电子重新又合在一起，然后再经过一个 $\hat{z}$ 方向的仪器打在屏上。问：屏上有一个还是两个斑点？

每次只打出一个电子，因此把电子分成很多份的经典波动力学的解释是不能用的。我们来看看经典粒子的力学加上概率论的解释。一方面，经过 $\hat{z}$ 方向仪器以后的电子经过 $\hat{x}$ 方向的仪器分开可能的两条路径然后每一条路径又被反射合起来，所以看起来这个 $\hat{x}$ 方向的仪器的效果完全被两个反射镜消掉了，于是在到达最后一个 $\hat{z}$ 方向仪器的时候，电子的状态与刚刚出第一个 $\hat{z}$ 方向仪器的时候一样，也就是向上态。于是，只有一个斑点。

另一方面，如果要问经过 $\hat{x}$ 方向的仪器之后，电子是从哪一条路径过去的。这个时候，如果我们假设电子是从上面的路径过去的，代表 $\hat{x}$ 方向的向上状态，于是，这个状态遇到最后一个 $\hat{z}$ 方向的仪器会产生两个斑点。同样的，如果我们假设电子是从下面的路经过去的，代表 $\hat{x}$ 方向的向下状态，将来也是两个斑点。于是，按照概率叠加原理，这个事情的发生有两种可能——从上面或者是从下面的路径过去的，于是整个事件的结果就是这两个可能的概率平均。既然两者都给出两个斑点，那么，必然，其概率平均也给出两个斑点。

可是，我们的实验结果是只有一个斑点，而且前面第一个逻辑也给出一个斑点的解释。这个现象怎么解释？我们的后一种解释哪里错了？

单光子过玻璃的实验，单光子过三块偏振片，单电子的双缝干涉实验，单光子 which-way 实验，单电子自旋 which-way 实验，都告诉我们：第一、经典波动力学把波看成是分成多份经过不同路径又同时到达某处的解释有问题；第二、经典粒子理论——其中我们假定粒子选择了哪一条路径来走，然后加上概率论的互斥事件概率可加性——的预期与实验不符合。第一个

可能的出路，我们可以修改概率论的互斥事件概率可加性。以后我们会看到，这个互斥事件可加性是概率论的核心。我们需要尽可能的保留它：它在经典的时候确实是正确的，而且很容易理解——一件事情有不同的发生方式，那么其结果可以看成是所有方式的结果的某种叠加。第二个可能是我们需要放弃粒子选择了某一条路径这个概念。那么，是否可以替换成光子同时走了所有的路径呢？我们做的是单个粒子的实验，粒子也总是一个一个地被测到，没有测到过半个光子或者半个电子。粒子需要分开好多份一起走这个图景和测量的结果不符合。另外，如果我们假设粒子在测量之前分成许多份，那么测量的时候，在探测到的位置就要求发生某种奇怪的事情，导致其他部分的粒子瞬间到被测量的位置集合。这一点也是很难理解的：被测量到这件事情瞬间传给粒子的已经分开的所有的部分，并且一定意义上这些已经分开的部分在瞬时合回来。

也就是说，第一、粒子确实选择了某一条路径而不是同时走所有路径，尽管我们在观测之前不知道走了那一条路径，或者换一个表达方式：在任何时刻任何地方测量粒子，得到的结果都是这个地方测到了或者没有测到一个完整的粒子；第二、粒子的状态是粒子在所有可能路径——而不是某一个路径——上的某种平均或者某种取和，而且这种取和不能是概率性叠加。这个时时刻刻走一条路径和所有路径的非概率性叠加的矛盾并不在“一条”和“所有”上面，而是在“一条”和“所有的非概率性叠加”上面。我们已经看到，对于经典的随机子弹，注意是随机的子弹，其在任何时刻走的是“一条”路径，并且观测结果与“所有”的路径的结果的概率性叠加不矛盾。现在，我们需要提供一种数学模型，能够解决上面的这个“一条”和“所有的非概率性叠加”之间的矛盾，从而解决建立一个对量子粒子被测量的时候的行为的可计算的模型的问题。这个答案就是下面，在本书的第三部分里面。在那里，我们要学习和讨论的目前主流量子力学的理论框架，然后用来解释这一章里面所有没有得到解释的单粒子的“干涉”现象。

1.3 作业

习题 1.1 阅读Whiltehead的《教育的目的》[17]、吴金闪的《教的更少，学得更多——概念地图学习与教学方法》[10] 和Adler和van Doren的《如何读一本书》[18]，写下阅读体会。本书仅提供量子力学核心概念里面的核心概念，主要是启发读者的思考，以及给读者准备进一步学习的基础的。按照这

个设计，本书把大量的内容都放在了读者的自学上面，提供了相应的参考书和论文。因此，学会看书、学会思考，在这里非常重要。阅读报告要包含：总结（说了什么、怎么说的、为什么说这个这么说）、评价（我觉得对吗）、结合自身经验的体会（我喜欢吗，对我有什么意义），也就是问好WHWM这四个问题，目的是显示：我看了，我想了，还有点收获。形式上，读书报告要包含文字和概念地图。有关概念的内容请从《教的更少，学得更多——概念地图学习与教学方法》中学习。

习题 1.2 阅读 Feynman 的《Feynman 物理学讲义》[7] 第三卷，前三章，写一个报告。这个和本书的讲法类似。实际上本书的整体逻辑结构很大程度上受《Feynman 物理学讲义》第三卷的启发，尽管本书对于密度矩阵、矢量空间、概率论的处理还是比较独特和有新意的。这些异同以及这样做的原因通过阅读和比较这两本书能够有更深刻的体会。希望学习本书的读者能够明白本书内容、内容之间的逻辑关系，同时还明白为什么选择这些内容和这样来呈现它们之间的逻辑关系。我称之为“知其然，知其所以然，知其所以所以然”。这也是我在《教的更少，学得更多——概念地图学习与教学方法》一书里面提出来的学习方法。另外，本书的习题不多，大部分习题都是为了进一步学习设计的，有比较深远的意义的。请做题的时候多思考。

习题 1.3 搜索“Sidney Coleman 量子力学讲座”，听完，写一个报告。听不懂没关系，以后我们还要听很多遍，本课程的任务就是让你听懂这个讲座。

习题 1.4 搜索“Stanford 大学Susskind量子力学公开课”，听完第一次，写一个报告。听不懂没关系，以后我们还要听很多遍，本课程的任务就是让你听懂这个课程。

1.4 本章小结

经典的介质上的波由于有介质粒子之间的相互作用和 Newton 定律的存在，可以看作一束分成多束，并且波的叠加原理或者波矢量的分解原理成立。经典的粒子的状态满足概率叠加原理。然而，波的矢量叠加或者分解不能用于解释没有介质的单个粒子的行为，粒子的概率叠加原理也不能给出与上面所讨论的一系列实验结果一致的结果：单光子过玻璃的实验，单光子过三块偏振片，单电子的双缝干涉实验，单光子 which-way 实验，单电子自旋 which-way 实验。于是，我们要寻找一个能够解释这些实验的理论，同时

如果能够建立起来这样的理论，我们要问如何来理解这个理论，这个理论是否能够给出来其他的可实验检验的结果，并且进一步取做实验检验这些结果。

矢量叠加和分解原理能够形式上解释量子系统的行为提示我们我们需要构造的理论模型应该包含这样的数学，只不过，其基础不再是 Newton 定律。那我们的理论应该什么样呢？这样的理论看起来不能去问粒子到底走那一条路径然后做概率叠加。也就是数学上，象波——可以做矢量跌价：这样的理论包含类似于公式 (1.7) 这样的关系——某一个量的给定状态可以看作另外一个量的几个给定状态的和；物理上又能够用于单个粒子。我们将沿着这个方向做进一步的探索。

做为一个用概念地图来总结和整理知识的例子，这里，我们在图 1.16 中提供一张关于这一章主要内容的概念地图，以及在图 1.17 中的一张双缝实验的概念地图。注意，对知识的思考、理解和组织是概念地图反映的主要内容，而不是知识本身。也正是因为这个原因，以后的章节中，我们将不再提供概念地图，而是希望读者自己来思考和整理。更多的关于概念地图和概念地图思维方式、概念地图学习方法、概念地图方法和教学体系，请参考吴金闪的《教的更少，学得更多——概念地图学习与教学方法》[10]。

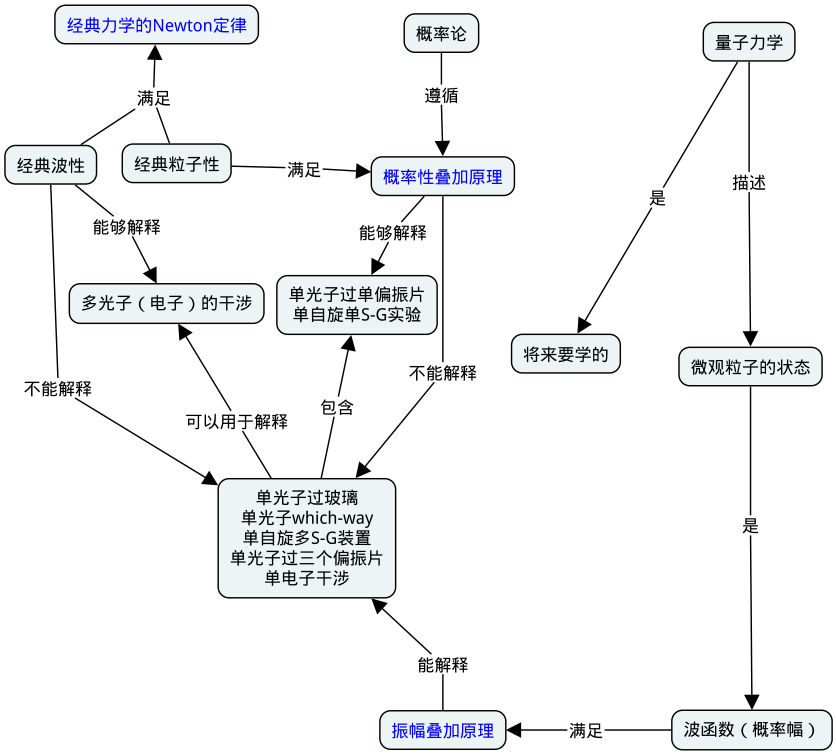


图 1.16: 为了让读者体会如何用概念地图来辅助做知识的深入思考、理解和组织，整理了这一章的主要内容。

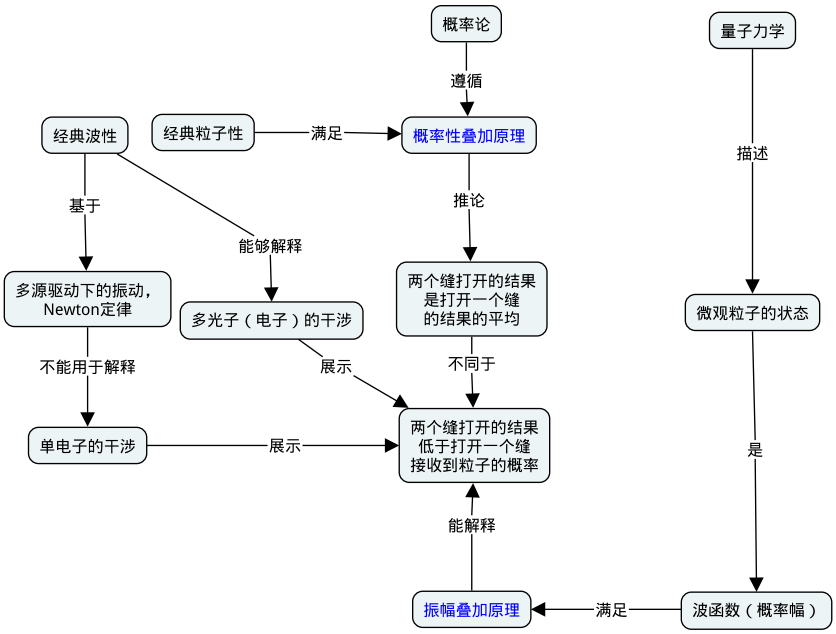


图 1.17: 关于双缝干涉这个实验的拥有更多细节的概念地图。

第二章　物理世界的数学模型

物理学，甚至整个科学的目标就是让我们的理念中的世界，心智中的世界符合物理的真实的世界。所谓符合物理的真实的世界就是构造一套简单的规则通过这些规则的操作给出与现实世界相符的结果。这一套规则本身是否能够得到验证，也就是说这套规则本身是不是就是世界运行的规则，不是科学关心的问题。这一套规则是否能够被我们的心智认可、理解、欣赏，也不是科学关心的问题。相符是不是就是可验证性呢？一般来说理论模型是不可以直接验证的，而且就算前面的 10000 个实验都相符，逻辑上也不能推断出来第 10001 个实验肯定也相符，于是就算通过计算得到的结果和实验相符，也很难说一个理论得到了验证。为了更好地来表述这个概念，Popper提出了可证伪性的说法 [11]。可证伪性就是指，我们的规则原则上允许不成立，而且一旦这个规则不成立，则相应的操作可以导出与这套规则成立的情况下不同的结果。如果这些规则不成立的情况下导出的结果被发现了，就称一个规则被证伪了。称为“科学”的规则就是原则上可以被证伪但是目前为止还没有被证伪的一套能够给出与现实世界相符的结果的规则。在实际科学研究中，用“和实验相符”粗糙地来代替“可证伪性”往往就够了。以后我们就简单地用“相符”。如果一定要在相符的前提下再提出一个更高的标准，可以认为是简单性。简单性就是如果有两个理论，一个简单（可以指的是规则本身更少，操作更简单）一个复杂，我们优先选择简单的理论；简单性也可以指如果有两套规则，在一套规则下，很多事情都能够通过一套理论来解释，而在另一套规则下，需要更多的理论来解释更多的事情，那么我们选择更加普适的理论。简单性还可以指系统性——用尽可能少的假设（公理）做为理论体系的基础来建立整个理论体系。

这样的一套规则往往是用数学的语言和结构描述的一套规则，称为数学模型。有的时候可能一套规则看起来用的不是数学的语言，但是实际上仍然是数学的结构，例如Faraday的电磁场后来就被Maxwell做了等价的数学化。

我们已经看到，量子世界的现象，尽管有的表现很像经典硬币（过 $\hat{z}$ 方向的 Stern-Gerlach 装置的电子），有的表现很象经典的波（过玻璃的光的经典波动力学解释），但是却不能够通过这些经典的对象来理解。实际上，我们在这里表达的意思是：这些量子的现象不能通过经典粒子背后的数学模型与经典波背后的数学模型来解释。那么，首先，我们来看一下什么是经典粒子和经典波后面的数学模型。

由于在下面的几节中我们会使用第二部分的数学符号和物理概念，请在继续阅读之前学习或者浏览本书的第二部分“数学物理学上的准备”。对于已经熟悉概率论、经典力学和 Dirac 符号的读者，可以直接阅读本章剩下的部分。或者，请读者先相信我们的这一章结论——量子系统的行为非常难以用经典理论来描述，然后跳过本章剩下的部分，将来等学习完了第二部分甚至第三部分以后，再回来挑战这一部分。

警告：初次思考本章剩余部分里面的问题的读者一般情况下会觉得这些问题比较难比较抽象。

2.1 经典世界的确定性的数学模型

经典粒子的状态以及状态变化的数学模型非常简单，就是一个平直的连续的欧几里德空间中的一个点 (x_1, x_2, x_3)，以及这个点随着时间——时间 t 是独立于一切之外的均匀流逝的——的变化。为了描述这个空间中的点，我们可以利用坐标系。原则上坐标系和距离有可能是不统一的，但是欧氏空间的直角坐标系就是利用距离来定义的，或者说坐标值之间的差的绝对值正好就是距离，

$$d_{PQ} = \sqrt{\sum_i \left(x_i^P - x_i^Q\right)^2}. \tag{2.1}$$

时间也是，

$$t_{PQ} = \sqrt{\left(t^P - t^Q\right)^2}. \tag{2.2}$$

于是，一个经典粒子的状态的完整的描述就成了一段时间内的 $\vec{r}(t) = \sum_i x_i(t)\,\hat{e}^i$ ——这段时间内的一条轨迹。

实际上，还存在着另外一种完整描述经典粒子状态的方式，就是同时确定某一个时间点 t_0 的位置和速度 $\left(\vec{r}(t_0), \dot{\vec{r}}(t_0)\right)$。有的时候这个组合也被替换成为确定一个时间点的位置和动量 $(\vec{r}(t_0), \vec{p}(t_0))$，其中 $\vec{p} = m\vec{v} = m\dot{\vec{r}}$。这种描述相对于前者——$\vec{r}(t)$——有很大的好处：前者需要确定整个时间段的整

个轨道，后者只需要确定一个时间点 t_0 的状态。之所以我们能够做样做是因为经典粒子的运动方程——Newton 方程是一个二阶微分方程，于是可以转变成为两个一阶方程，并且在由位置和动量构成的相空间内轨道不相交。

一般而言，状态的含义是当前时刻的而不是一段时间的。因此，我们更加喜欢前面两种状态表述中的第二种，采用一个时刻的位置和动量来代表系统的状态，而且这个状态还决定了这个系统的过去和将来。

如果我们对这样的一个经典粒子做测量的话，测量的图景非常简单。首先，我们需要确定一个测量的量 O，例如位置坐标、速度、能量等等。这些可以测量的量都是前面的位置和动量这两个基本量的函数。因此，这里我们就以位置坐标的某一个分量的测量为例 $O = x$。接着，我们需要有一个测量仪器 A，这个仪器能够根据位置的不同而得到不同的读数。也就是说，这个仪器与所测量的状态之间可以建立某种关联 $A_x = A(x)$。这个关联的建立本质上也应该是一个物理过程，但是我们在这里不再关心这个具体的物理过程，仅仅当作抽象的一个关联。最后，我们通过观察仪器上的读数 A_{x^*} 来确定所测量的量的值 $O = x^*$。例如，如果映射 $A(x^*)$ 可逆，则通过 $A^{-1}(A_x^*)$ 我们得到 x^*。

于是，我们发现，**在测量之前，系统的被观测量的值是 x^*，在测量之后是 x^*，测量得到的值也是 x^***。我们对这个“测量”过程的“理解”感到满意，因为看起来这真的就是一个测量：不改变系统的被观测量的状态，仅仅是得到这个被观测量的值，测量的之前之后都一样。在实际实验中，这样的完全不改变被侧量值的测量也是做得到的，尽管有的时候也会引起被测量值的一个小小的改变。后者不是原理上要求这样，仅仅是受到技术上的可能性的制约。通常通过把被测量量耦合到另外一个可以改变和破坏的自由度上去，通过观察代表那个自由度的变量的取值，我们是可以做到完全不影响被测量量的。你可能觉得所有的测量都应该满足上面的观测之前的量、观测所得记录和观测之后的量之间的关系，因此觉得上面的讨论没有必要。把这个疑虑先放在脑子里，我们很快就会回到这个问题。

现在我们已经讨论了经典粒子的状态和状态的测量的数学模型。为了完整地建立经典粒子的数学模型，其实我们还需要讨论经典粒子是如何演化的。不过，如何理解演化相对来说是一个平庸的问题。这里就不详细展开了。简单来说，就是粒子受到外界对它的影响或者作用，在这个影响下，粒子的状态 $(\vec{r}, \vec{p})$ 会发生变化。在影响因素和如何变化之间应该存在着一个关

系，这个关系通常就叫做运动方程，例如 Newton 方程，

$$(\vec{r}(t_0), \vec{p}(t_0)) \xrightarrow{\text{Newton Equation}} (\vec{r}(t), \vec{p}(t)). \tag{2.3}$$

Newton 方程的具体形式暂时我们不关心。

波动力学的图景是建立在上面这个经典粒子的状态和测量的图景的基础上的。经典波实际上是一系列粒子的运动的整体表现，例如 $\vec{r}_1(t), \vec{r}_2(t), \cdots, \vec{r}_N(t)$，以及从有限 N 个离散脚标变成连续的脚标 $\vec{r}(x,t)$。每一个在 x 点的粒子仅仅在它自己的平衡位置附近运动，不会沿着波的方向传播。波传播的不是粒子而是振动的状态，称为相位——把一个周期的振动看作一个 2π 的弧度，把其中任何一个点的振动状态对应于一个 $\theta \in [0, 2\pi]$。然后就是这个 θ 角的状态在沿着波传播——一般情况下写成 $\vec{r}(x,t) = \vec{A}\cos(\theta(x,t))$。有的时候 $\theta(x,t) = kx + \omega t$。例如，波谷的时候就相当于 $\theta = \pi$，波峰的时候就相当于 $\theta = 0$，然后这些波峰和波谷就可以沿着波传播。对于波上的任何一个单个的粒子来说，状态以及状态测量的问题自然就完全回到了上面的已经建立的单个粒子的状态和状态测量的物理图景。演化问题稍微复杂一点，因为这个时候，周围的粒子对这个“单个”粒子是有相互作用的。于是，其演化方程——称为波动方程，尽管实际上还是 Newton 方程，会出现看起来与 Newton 方程不一样的自己的独特的形式。

$$\vec{r}(x, t_0) \xrightarrow{\text{Wave Equation (Newton Equation)}} \vec{r}(x,t). \tag{2.4}$$

波动方程的具体形式暂时我们不关心。

对于这样的一个波来说，如果存在两个波源一起来驱动某个点——例如一条绳子的两端同时被两个人拿着然后上下振动——那么自然这个点的运动模式是两个源独立传播过来那个振动状态的叠加。具体来说，就是在特定的时刻 t^* 在这个特定的点 x^* 上的粒子的状态满足

$$\vec{r}(x^*, t^*) = \vec{A}_1 \cos(\theta_1(x^*, t^*)) + \vec{A}_2 \cos(\theta_2(x^*, t^*)). \tag{2.5}$$

这个称为波的叠加原理——实际上是不同的源驱动的波引起的同一个粒子的振动状态的位置矢量的叠加。这个叠加的物理基础是力的叠加和 Newton 定律：从

$$\vec{F} = \vec{F}_1 + \vec{F}_2, \tag{2.6}$$

得到

$$\vec{a} = \vec{a}_1 + \vec{a}_2. \tag{2.7}$$

于是，对于同一个点同一个初始条件，我们得到公式 (2.5)。

注意，这个叠加就好像是你的妈妈给你 A_1 的零花钱，你的爸爸再给你 A_2 的零花钱，合起来你得到的的零花钱是 $A_1 + A_2$。这个加法不代表这两件事情——妈妈给你钱和爸爸给你钱——可以加起来，也不代表这两份钱——例如两张十元的纸币（两张十元的纸币加起来肯定不是一张二十元纸币）——可以加起来，仅仅是钱的数值可以加起来。如果有一天，有人告诉你，这两张十块钱真的可以加起来，加起来以后得到一张钱，而且这张钱的面值是三块或者一百块，那么，你会觉得这个不是世界疯了就是那个这样告诉你的人疯了。而不久以后我们就要遇到这样的加法了：一个左边缺了一点的苹果加上一个右边缺了一点的苹果，竟然成了一个香蕉，竟然可以加起来。这就是本书封面图的来源。

有了这个基于位置矢量的数值的叠加原理的波的叠加原理，这个叠加之后的状态的点的测量就回到了单个粒子的测量的问题。所以，确定性的经典粒子和经典波的状态、状态的测量、状态的演化的数学模型就都归结到这一节一开始的时候我们介绍的图景：状态是欧氏空间的点，测量是一个状态、测量量、仪器和仪器读数这几个东西相互关联的过程——这个过程中，理论上，**测量前后不改变被测量量的状态仅仅读出被测量量的值**，演化就是状态的变化与系统内部各个部分之间相互影响以及外界对系统的影响的联系。

2.2 经典世界的随机性的数学模型

上一节，我们讨论了确定性经典粒子和经典波的状态、测量和状态演化的数学模型。这一节，我们来讨论随机性的经典粒子的数学模型。

让我们从一个简单的事件的数学模型开始：什么是一个色子的数学模型？对于一个色子，我们通常只关心其向上的面的状态，也就是那个向上的数字。我们知道这个数字必然是 $\{1\ 2\ 3\ 4\ 5\ 6\} \triangleq \Omega$ 之一。于是，我们知道了这个状态构成一个集合。然后，对于这个色子，在现实中，我们可以问各种各样的问题，例如状态是 1 的几率是多少。这个时候几率的含义往往是指，统计上，如果扔很多很多次，合起来看出现的频率。我们还可以问状态是奇数的几率是多少。其实，我们可以问任何一个 Ω 的子集所代表的事件出现的几率是多少这样的问题。于是，我们发现，我们感兴趣的有关色子的“事件”实际上就是 Ω 的所有子集的集合 $\{A \mid A \subseteq \Omega\} \triangleq \mathcal{T}$。进一步，我们关心的问题则是这样的事件 A 的几率是多少，也就是一个从 $\mathcal{T}$ 到 $[0,1]$ 的映射，

其把任何一个 A 对应到一个 $0,1$ 之间的数字 $P(A)$。

在这里，我们首先把有关色子的状态所有的问题数学化了。接着，我们要寻找这个数学问题的数学答案。然后，我们还可以接着问这个色子的进一步的问题，例如，如果这个色子的状态发生变化，如何用数学模型描述这个变化，以及如何找到这个变化发生的物理原因。后面的这个问题，我们暂时不讨论。在此，我们仅回答如何表述状态和如何形式上描述测量，也就是说，我们要给定一套规则，这套规则能够计算任意事件 A 的几率 $P(A)$。我们说这个答案非常简单，就是对于 Ω 中的任意元素 ω，先给定所有的 $P(\{\omega\})$，然后对于任意的集合 A，我们按照如下规则计算：$P(A)=\sum_{\omega\in A}P(\{\omega\})$。对于我们的理想的色子，我们取 $P(\{\omega\})=\frac{1}{6},\forall\omega\in\Omega$。如果你对以上关于色子的概率模型的集合论语言非常不熟悉，请参考本书第四章，或者Kolmogorov的《概率论导引》[19] 或者Ito的《伊藤清概率论》[20]。

总结一下，对于理想的色子——这个色子可能与真实世界的色子略有区别，不过这个区别可以通过修正下面的具体的数字来反映——我们的数学模型非常简单，

$$P(A)=\sum_{\omega\in A}P(\{\omega\}),\forall A\subseteq\Omega;P(\{\omega\})=\frac{1}{6},\forall\omega\in\Omega. \tag{2.8}$$

或者运用我们将要在第四章中学习到的密度矩阵的语言，

$$P(A)=tr\left(\hat{A}\rho^{c}\right),\hat{A}=\sum_{\omega\in A}|\omega\rangle\langle\omega|,\rho^{c}=\frac{1}{6}\sum_{\omega=1}^{6}|\omega\rangle\langle\omega|. \tag{2.9}$$

例如，当 $A=|1\rangle\langle 1|$ 的时候，我们问的是事件“色子状态是面 1 向上”的几率是多少。于是，按照一般概率论的语言，

$$P(1)=\sum_{\omega\in\{1\}}P(\{\omega\})=\frac{1}{6}, \tag{2.10}$$

或者按找密度矩阵的语言

$$P(|1\rangle\langle 1|)=tr(|1\rangle\langle 1|\rho^{c})=\langle 1|\rho^{c}|1\rangle=\frac{1}{6}. \tag{2.11}$$

更一般地，如果你需要测量一个可观测量 O，它是核心是一个 ω 的函数 $O(\omega)$（例如根据色子的每一面的值你可以赢得不同数量的钱），那么，按照通常概率论的语言

$$\langle O\rangle=\sum_{\omega}O(\omega)P(\{\omega\}), \tag{2.12}$$

或者相应地按照密度矩阵的语言，

$$\langle O\rangle = tr\left(\hat{O}\rho^{c}\right), \hat{O} = \sum_{\omega} O\left(\omega\right)|\omega\rangle\langle\omega|. \tag{2.13}$$

一句话，色子的状态就是

$$\rho^{c} = \frac{1}{6}\sum_{\omega=1}^{6}|\omega\rangle\langle\omega|, \tag{2.14}$$

可观测量是算符（例如前面的 $\hat{A}$ 和 $\hat{O}$），然后，观测以后得到的结果就按照

$$\langle O\rangle = tr\left(\hat{O}\rho^{c}\right) \tag{2.15}$$

来计算。

如果你问我“你怎么知道公式 (2.8) 或者等价的公式 (2.14) 这个数学描述就是色子的本质”，我拒绝回答。我仅仅能够告诉你，利用这个数学结构计算出来的任何结果，都可以通过实验的检验。如果你问我“你是否喜欢这个色子的数学模型”，我也拒绝回答。我还是仅仅能够告诉你，利用这个数学结构计算出来的任何结果，都可以通过实验的检验。

当然，我们的数学模型在几个层次都允许这个描述不成立的规则，例如，我们可以破坏公式 (2.8) 的后半部分，允许 $P\left(\{\omega\}\right) \neq \frac{1}{6}$。这个时候，我们需要实验证据来证明这个色子的行为其实更符合在这些新的不等于 $\frac{1}{6}$ 的数字下面得到的结果。如果成立，那么，我们放弃 $P\left(\{\omega\}\right) = \frac{1}{6}$。这样的话，我们的模型的细节被证伪了，我们修改，但是整个把概率结构作为色子的行为的描述的数学框架还是对的。破坏 $P\left(\{\omega\}\right) = \frac{1}{6}$ 还有别的方法，就是 $P\left(\{\omega\}\right)$ 不再是一个简单的数了，整个就不能用实数来表达了。实际上，一种这样的修改就是考虑 Ω 是一个连续集合而不是离散元素的集合。这个修改，如果我们色子的行为确实需要也是可以的。注意，到此为止，我门一直保留了公式 (2.8) 的前半部分，有关取和的部分，当然必要的时候可以成为积分。

我们也可以破坏公式 (2.8) 的前半部分，允许 $P\left(A\right) \neq \sum_{\omega\in A} P\left(\{\omega\}\right)$。这个修改的含义是，在我们保持 $P\left(\{\omega\}\right)$ 是一个简单的实数的前提下，我们的事件集合的概率不再是集合中每一种可能的事件的概率的和了。这样的修改在色子这个现实世界的物体上，是不可理解的。这个修改实际上就破换了整个概率论的核心——互斥事件的加法是概率论这个数学结构的基本公理之一。当然，如果我们的“色子”的行为确实不能够由前面的公式 (2.8) 数学模型描述，也不能在更一般的允许 $P\left(\{\omega\}\right)$ 取值的数学模型描述，也不能

由仅保留公式 (2.8) 的前半部分，把后半部分整个抛弃然后用连续变量来代替的更更一般的理论来描述，那么，我们也可以突破整个概率论的描述框架。不过，迄今为止，对于色子，我们总是可以保留整体框架，最多修改一下 $P(\{\omega\})$ 的取值，就可以描述其行为了。

讨论了状态和状态的测量之后，我们再来讨论一下测量后状态和状态的演化。首先，测量后状态，当测量值为 O^* 的时候，根据 $\omega^* = O^{-1}(O^*)$，为

$$\rho_{af}^c = |\omega^*\rangle\langle\omega^*| = |O^*\rangle\langle O^*|. \tag{2.16}$$

其次演化的过程肯定是从一个密度矩阵到另一个密度矩阵的映射，

$$\rho^c(t) = L(t,t_0)\rho^c(t_0). \tag{2.17}$$

其中 $L(t,t_0)$ 由外界对这个系统的相互作用以及这个系统的内部结构来决定。例如，在概率论和随机过程的课程里面，我们可能学到过概率转移矩阵 [21]，

$$P(t) = M(t,t_0)P(t_0). \tag{2.18}$$

这里，运用密度矩阵的语言（见本书第四章），有可能我们可以把演化过程表达成为

$$\rho^c(t) = U(t,t_0)\rho^c(t_0)U^\dagger(t,t_0). \tag{2.19}$$

下面举一个硬币翻转的例子来展示状态、状态的演化，测量和测量后状态。

例 2.1 (**硬币翻转的动力学过程的密度矩阵描述**)：有一个硬币每一个单位时间以后都会被翻转。写下这个动力学过程的普通概率论形式和密度矩阵形式的表达式。假设我们测量硬币的正面还是反面，然后当是硬币正面的时候赋值为 1（得到一个单位的钱），反之为 −1。写下这个测量的算符，测量值和测量后状态。

硬币的状态可以采用两种表示方法，概率矢量

$$P = \begin{bmatrix} p \\ 1-p \end{bmatrix}, \tag{2.20}$$

或者概率矩阵，

$$\rho^c = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}. \tag{2.21}$$

翻转操作是算符

$$X=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{2.22}$$

翻转过程可以表示为

$$P(t+1)=XP(t), \tag{2.23}$$

或者

$$\rho^c(t+1)=X\rho^c(t)X^\dagger. \tag{2.24}$$

测量的可观测量是算符

$$Z=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.25}$$

相当于

$$\hat{Z}=1\cdot|\uparrow\rangle\langle\uparrow|+(-1)\cdot|\downarrow\rangle\langle\downarrow|. \tag{2.26}$$

$|\uparrow\rangle\langle\uparrow|$ 和 $|\downarrow\rangle\langle\downarrow|$ 分别代表正面和反面。对于这个可观测量，有两个可能的观测值 1 和 -1，其概率分别为

$$P_\uparrow=tr\left(\rho^c|\uparrow\rangle\langle\uparrow|\right)=p, P_\downarrow=tr\left(\rho^c|\downarrow\rangle\langle\downarrow|\right)=1-p. \tag{2.27}$$

所以测量的平均值是，

$$z=1p+(-1)(1-p)=2p-1=tr\left(\hat{Z}\rho^c\right). \tag{2.28}$$

当我们观测到测量值的状态——例如 -1——的时候，测量后状态是

$$\rho^c_{af}=|\downarrow\rangle\langle\downarrow|. \tag{2.29}$$

这个就是作为经典随机对象的硬币的数学模型的所有的内涵。讨论完了色子和硬币这两个具体的模型，我们顺便指出，确定性的经典系统也可以采用概率模型来描述——只不过，这个时候，任何时候的分布函数就是一个 δ 分布：只在一个状态上取值，其它地方的概率取值都是 0。这个就好像是Dirac δ 函数或者Kronecker δ 记号一样。因此，通常认为，经典世界的数学模型的底线就是概率模型：有一个基本事件集合，这个集合的部分子集构成一个事件集合，然后从事件集合到 $[0,1]$ 有一个映射，这个映射遵循互斥事件的加法规则。其实，还有几个其它的要求，请参阅概率论的公理体系[20]，也可以参考本书第四章的概率论部分。

或者用密度矩阵的语言，经典世界的数学模型就是：状态是一个对角的密度矩阵 ρ^c，可观测量是对角的算符 $\hat{O}$，测量得到的平均值就是按照公式 (2.15) 计算的到的 $tr\left(\rho^c\hat{O}\right)$，观测到可观测量某一个值 o^* 的几率是 $tr\left(\rho^c|o^*\rangle\langle o^*|\right)$，当观测到这个具体的值之后，系统的观测后状态是 $|o^*\rangle\langle o^*|$。一句话来概括：**测量之前是 ρ^c（如果测量 O，可能得到多个结果），当测量得到的值是 o^* 的时候，测量之后的状态是** $|o^*\rangle\langle o^*|$；如果需要考虑演化，则 $\rho^c(t) = L(t,t_0)\rho^c(t_0)$。对于随机客体，**测量前后状态**至少数学形式上看起来**可以不同**，例如公式 (2.21) 和公式 (2.29)。这一点和经典确定性客体的数学模型不一样。不过，上面的数学模型可以统一描述确定性和随机的客体的状态、测量和演化。

有了这个确定性的和随机性的经典世界的数学模型，我们来看一看能不能构建一个量子世界的数学模型，让我们的数学结构能够描述量子系统的行为，包含典型的上一章介绍过的所有的量子实验的结果。

2.3 理论描述的目标：前后两次测量的关联

在第一章，我们已经展示了通常的经典粒子和经典波的理论都不能解释量子系统的行为。现在我们来尝试构造各种更加一般的经典理论，只要满足经典理论的底线——还是一个概率论理论——的经典理论，看看能不能有这样的理论可以解释量子行为。

我们希望找到一个数学模型来解释的现象是这样的。第一步，制备一个特定的状态：让自旋经过一个 z 方向 Stern-Gerlach 装置然后挡住向下的方向，仅仅让向上的方向出来的自旋进入下一步实验。第二步，做这个状态的 $\hat{r}_1$ 方向的测量：让自旋接着经过一个 $\hat{r}$ 方向 Stern-Gerlach 装置然后看看向上还是向下能够接收到自旋，记录单次的结果和多次以后的统计平均。第三步，再做这个从 $\hat{r}_1$ 出来的自旋的 $\hat{r}_2$ 方向的测量：让自旋接着经过一个 $\hat{r}_2$ 方向 Stern-Gerlach 装置然后看看向上还是向下能够接收到自旋，记录单次的结果和多次以后的统计平均。

这个里面其实有一些技术问题，例如经过 $\hat{r}_1$ 被测量的自旋如何再次被在 $\hat{r}_2$ 方向上测量，多少次才是足够多次以至于可以来计算统计平均。但是，在这里，我们假设这些技术问题总是可以解决的，而不关心它们。这个多次制备同样的系统来做同样的测量并计算均值的过程被称为系综测量。系综就是多个相同的系统的集合的意思。我们用系综测量来区分再次测量——

对于已经完成第一次测量的系统再一次在前一次测量结束之后的系统状态上马上做再一次的测量。用系综测量和再次测量这两个术语比用含义稍微模糊一点的重复测量要好。

将来构造出来的理论必须能够描述的实验结果是这样的。经过 $\hat{r}_1$ 方向的测量给出来的结果是，**单次测量的结果要么向上（记为 $s_{\hat{r}_1}=1$）要么向下（记为 $s_{\hat{r}_1}=-1$），多次统计平均得到向上和向下的几率分别是** $p_{\uparrow_{\hat{r}_1}}=\frac{1+\hat{z}\cdot\hat{r}_1}{2}, p_{\downarrow_{\hat{r}_1}}=\frac{1-\hat{z}\cdot\hat{r}_1}{2}$，也就是

$$\langle s_{\hat{r}_1}\rangle=\hat{z}\cdot\hat{r}_1,\ s_{\hat{r}_1}=\pm 1. \tag{2.30}$$

再做 $\hat{r}_2$ 方向的测量，如果之前测量得到的是 $s_{\hat{r}_1}=1$，那么给出来的结果是，单次测量的结果要么向上（记为 $s_{\hat{r}_2}=1$）要么向下（记为 $s_{\hat{r}_2}=-1$），多次统计平均得到向上和向下的几率分别是 $p_{\uparrow_{\hat{r}_2}}=\frac{1+\hat{r}_1\cdot\hat{r}_2}{2}, p_{\downarrow_{\hat{r}_1}}=\frac{1-\hat{r}_1\cdot\hat{r}_2}{2}$，也就是

$$\langle s_{\hat{r}_2}\rangle=\hat{r}_1\cdot\hat{r}_2,\ s_{\hat{r}_2}=\pm 1. \tag{2.31}$$

如果之前测量得到的是 $s_{\hat{r}_1}=-1$，那么给出来的结果是，单次测量的结果要么向上（记为 $s_{\hat{r}_2}=1$）要么向下（记为 $s_{\hat{r}_2}=-1$），多次统计平均得到向上和向下的几率分别是 $p_{\uparrow_{\hat{r}_2}}=\frac{1-\hat{r}_1\cdot\hat{r}_2}{2}, p_{\downarrow_{\hat{r}_2}}=\frac{1+\hat{r}_1\cdot\hat{r}_2}{2}$，也就是

$$\langle s_{\hat{r}_2}\rangle=-\hat{r}_1\cdot\hat{r}_2,\ s_{\hat{r}_2}=\pm 1. \tag{2.32}$$

把上面的两种情况合起来，计算一下关联函数（$\langle s_{\hat{r}_1}s_{\hat{r}_2}\rangle=p_{\uparrow_{\hat{r}_1}}1\times\hat{r}_1\cdot\hat{r}_2+p_{\downarrow_{\hat{r}_1}}(-1)\times(-\hat{r}_1\cdot\hat{r}_2)$），我们得到，

$$\langle s_{\hat{r}_1}s_{\hat{r}_2}\rangle=\hat{r}_1\cdot\hat{r}_2,\ s_{\hat{r}_1}s_{\hat{r}_2}=\pm 1. \tag{2.33}$$

注意到公式 (2.30)、公式 (2.31)、公式 (2.31) 和公式 (2.33) 之间具有内部相似性（例如，公式 (2.30) 左边默认了 $s_z=1$，公式 (2.32) 左边默认了 $s_{\hat{r}_2}=-1$），它们可以统一表达成

$$\langle s_{\hat{r}_1}s_{\hat{r}_2}\rangle=\hat{r}_1\cdot\hat{r}_2,\ s_{\hat{r}_i}=\pm 1. \tag{2.34}$$

这个结果涵盖了上面的初次测量的一次和系综测量的结果，也就是公式 (2.30)，以及再次测量的一次和系综测量的结果，也就是公式 (2.31)、公式 (2.31)。因此，我们所谓的构造一个理论模型希望能够解释量子系统的行为，就是指公式 (2.34)。注意，在这个统一公式的过程中，我们还把制备的

过程——先让系统通过 z 方向，选择向上的进入实验——和后面的测量 $\hat{r}_1$ 逻辑上等同了起来。

注意，对于已经学习过算符形式的量子力学的读者，这里 $s_{\hat{r}_1}, s_{\hat{r}_2}$ 仅仅表示做 $\hat{r}_1, \hat{r}_2$ 方向得到的测量结果的值 (± 1)，而不是这个方向的自旋算符。因此，$s_{\hat{r}_1} s_{\hat{r}_2}$ 也仅仅表示两个测量得到的值的乘积而不是两个算符的乘积。将来我们可以验算（见习题 3.8），如果是算符，公式 (2.34) 是不成立的。

我们所谓的构造一个理论模型就是给出来制备或者测量结束以后的状态的数学表达以及应该对这些状态做什么样的数学计算才能够得到测量的结果也就是公式 (2.34)，例如这样的形式理论，

$$\langle s_{\hat{r}_1} s_{\hat{r}_2} \rangle = \sum_{s_{\hat{r}_1}=\pm 1, s_{\hat{r}_2}=\pm 1} s_{\hat{r}_1} s_{\hat{r}_2} \rho\left(s_{\hat{r}_1}|0\right) \rho\left(s_{\hat{r}_2}|s_{\hat{r}_1}\right). \tag{2.35}$$

其中 $\rho\left(s_{\hat{r}_1}|0\right)$ 是给定某个初始状态 0 以后的 $s_{\hat{r}_1}$ 的分布函数，$\rho\left(s_{\hat{r}_2}|s_{\hat{r}_1}\right)$ 是给定第一次测量结果是 $s_{\hat{r}_2}$ 以后的 $s_{\hat{r}_2}$ 的分布函数。用密度矩阵的语言，上式可以写成，

$$\langle s_{\hat{r}_1} s_{\hat{r}_2} \rangle = \sum_{s_{\hat{r}_1}=\pm 1, s_{\hat{r}_2}=\pm 1} s_{\hat{r}_1} s_{\hat{r}_2} \left\langle s_{\hat{r}_1} \middle| \hat{\rho}(0) \middle| s_{\hat{r}_1} \right\rangle \left\langle s_{\hat{r}_2} \middle| \hat{\rho}\left(s_{\hat{r}_1}\right) \middle| s_{\hat{r}_2} \right\rangle. \tag{2.36}$$

现在我们的目标就是写下来经典密度矩阵 $\hat{\rho}(0)$ 和 $\hat{\rho}\left(s_{\hat{r}_1}\right)$ 的显式表达式来或者证明不可能写下来这样的经典密度矩阵。

很容易可以证明 $\hat{\rho}\left(s_{\hat{r}_1}\right)$ 必然依赖于 $s_{\hat{r}_1}$，否则，我们就会得到

$$\langle s_{\hat{r}_1} s_{\hat{r}_2} \rangle = \langle s_{\hat{r}_1} \rangle \langle s_{\hat{r}_2} \rangle, \tag{2.37}$$

而这个和实验不相符。这个说明前一次测量的结果改变了状态，而后一次测量的结果依赖于这个改变。这样才造成前后两次的关联。

将来在 11.4 节我们会看到，量子力学是能够给出来这些情况下的状态的描述以及应该做什么样的计算来得到公式 (2.34) 的，见公式 (11.12)。在那里（11.4 节）我们还将会对这个形式上的理论做进一步的讨论。

这个前后两次测量的关联性非常重要。这一章的主题相当于是讨论满足这个关联性的经典理论的可能性。将来在第十一章中，我们将对这个可能性做另外一个角度的讨论。在那里前后两次测量的关联被替换成两个“纠缠在一起”的自旋的关联。

除了直接把 $s_{\hat{r}_1}, s_{\hat{r}_2}$ 当做随机变量的经典理论，Einstein提出来还有可能可以考虑一个更复杂的“隐变量理论”——$s_{\hat{r}_1}, s_{\hat{r}_2}$ 不直接是随机变量，它们

是更基本的随机变量（记为 λ_1, λ_2）的函数。沿着这个思路，公式 (2.36) 可以成为，

$$\langle s_{\hat{r}_1} s_{\hat{r}_2} \rangle = \int\int d\lambda_1 d\lambda_2 s(\hat{r}_1, \lambda_1) s(\hat{r}_2, \lambda_2) \langle \lambda_1 | \hat{\rho}(0) | \lambda_1 \rangle \langle \lambda_2 | \hat{\rho}(\lambda_1) | \lambda_2 \rangle . \quad (2.38)$$

原则上，我们也需要讨论是否可以构建出来公式 (2.38) 中的 $\hat{\rho}(0), \hat{\rho}(\lambda_1)$ 来得到公式 (2.34)。其中 $s(\hat{r}_1, \lambda_1)$ 表示对这个随机变量做了 $\hat{r}_1$ 方向的测量，得到结果是 $s(\hat{r}_1, \lambda_1) \in \pm 1$，以及相应的随机变量在这个实现的样本中的取值是 λ_1。

讨论这个问题有两种方式：第一、构造出来这样的符合公式 (2.34) 的理论，然后讨论这个理论是否可以接受；第二、证明这样的理论在某些合理的要求下不可能符合公式 (2.34)。这一章我们先尝试前者，在第十一章我们来讨论后者。

2.4 量子系统的确定性经典理论的尝试

首先，我们来看一下概率理论的特殊情况——确定性的经典理论——是否能够描述量子系统的行为。所谓确定性在这里就是每时每刻系统的密度矩阵，或者说密度分布函数，都是一个 δ 函数：仅在某一个状态上取值，几率为 1，其它可能的状态的取值几率都为 0。我们以 Stern-Gerlach 装置的一系列实验为例。我们来看看一个未知状态的自旋通过 z 方向的磁场之后向上的那个状态的数学模型是什么？也就是取公式 (2.36) 中的 0 状态为 z 方向向上，然后想办法写下来 $\hat{\rho}(s_z = 1)$。或者说，我们取 $s_{\hat{r}_1} = 1, \hat{r}_1 = \hat{z}$，然后我们想写下来 $\hat{\rho}(s_z = 1)$。首先，我们认为这个时候自旋的状态是明确的（不表示是确定性的，仅仅表示不管什么形式，这个形式是已经给定的），客观的，记为状态 $|\uparrow_z)$。数学形式到底怎样，我们暂时不知道。这个状态就在那里，不管是否被测量，不管被测量的可观测量是什么。这个客观性要求是可以被挑战的。但是，我们暂时不打算放弃这个客观性要求。

为什么 $|\uparrow_z)$ 是明确的呢？因为如果我们对这个状态再做一个 z 方向自旋的测量（让这个自旋再一次进入一个 z 方向的 Stern-Gerlach 装置），我们得到的结果是确定的，还是得到而且仅得到一个向上方向上的斑点。这个结果是可重复的，而且你可以尝试改变很多很多的其他条件来测试是否这个状态还依赖于其它未知变量。只要你保证自旋先从第一个 z 方向的 Stern-Gerlach 装置的向上方向出来（挡住向下的），那么当这个自旋再一次进入

第二个 z 方向的 Stern-Gerlach 装置的时候，它还是从向上方向出来。按照我们经典世界的物理模型，这个自旋的 $|\uparrow_z)$，就应该是

$$\rho \triangleq |\uparrow_z) = |\uparrow_z\rangle\langle\uparrow_z| . \tag{2.39}$$

如果量子系统的行为就是这样的话，那么我们用确定性的经典力学就可以描述量子系统了，也就没有量子力学了。

幸好，或者不幸的是，量子系统的行为比这个远远复杂。我们说的量子系统的行为在这里指的是行为上符合公式 (2.34) 中 $\hat{r}_1 = \hat{z}$。我们希望写下来从 $\hat{z}$ 向上方向出来以后进入 $\hat{r}_2$ 之前的状态的数学描述。其实连具体的那些概率的数值我们都不太关心，主要关心定性上有几个输出。

如果我们让这个 $|\uparrow_z)$ 的自旋进入一个 x 方向的 Stern-Gerlach 装置，我们发现这个时候会出现两个可能性，而且每一次单次的实验只能得到一个斑点，多次合起来以后得到两个斑点。得到这两个斑点中的任意一个的频率是一样的。这个就好像是一个完全随机的硬币一样，而硬币的状态的数学模型是

$$\rho^c = p\,|\uparrow\rangle\langle\uparrow| + (1-p)\,|\downarrow\rangle\langle\downarrow| . \tag{2.40}$$

于是，从测量的结果来看，就算是对于一个明确的状态，$|\uparrow_z)$，我们需要放弃确定性的理论，而转投随机性的理论。例如，

$$\rho = p\,|\uparrow_x\rangle\langle\uparrow_x| + (1-p)\,|\downarrow_x\rangle\langle\downarrow_x| . \tag{2.41}$$

注意到之前我们用 $|\uparrow_z\rangle\langle\uparrow_z|$ 来表示这个状态，于是看起来

$$|\uparrow_z\rangle\langle\uparrow_z| = p\,|\uparrow_x\rangle\langle\uparrow_x| + (1-p)\,|\downarrow_x\rangle\langle\downarrow_x| . \tag{2.42}$$

实际上这个 x 方向可以任意取，实验都会得到两个可能的结果，只不过概率 p 的值依赖于所选择的方向。因此，这个关系需要对任何两个方向都对。这个可能吗?

$$|\uparrow_z\rangle\langle\uparrow_z| = p_{\hat{r}}\,|\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}| + (1-p_{\hat{r}})\,|\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}| . \tag{2.43}$$

尤其是如果我们取 z 方向向下当作进入 x 方向的光子的状态，也会得到 x 方向两个方向都有斑点，并且几率一样，也就是

$$|\uparrow_z\rangle\langle\uparrow_z| = p\,|\uparrow_x\rangle\langle\uparrow_x| + (1-p)\,|\downarrow_x\rangle\langle\downarrow_x| , \tag{2.44}$$

$$|\downarrow_z\rangle\langle\downarrow_z| = p\,|\uparrow_x\rangle\langle\uparrow_x| + (1-p)\,|\downarrow_x\rangle\langle\downarrow_x| . \tag{2.45}$$

这个可能吗？向上态就是向下态？将来在量子系统的量子力学模型里面我们会看到以上两行公式确实很像但是又不一样，同时和实验现象相符。目前的结果会导致看起来

$$|\uparrow_z\rangle\langle\uparrow_z| = |\downarrow_z\rangle\langle\downarrow_z| \tag{2.46}$$

这个可能吗？这里的 z 方向完全可以是任意方向。基于我们观察到有可能向上也有可能向下的两个方向的输出，我们想构建一个理论的目标也是将来给出一个某实验得到向上还是向下的结果。也就是说向上或者向下这两个方向的状态在我们的理论中是一定要区分的。如果我们沿着这个思路继续走下去，例如把 x 方向向上或者向下的状态相同代入到 $|\uparrow_z\rangle\langle\uparrow_z| = p|\uparrow_x\rangle\langle\uparrow_x| + (1-p)|\downarrow_x\rangle\langle\downarrow_x|$ 里面，我们还会得到 $|\uparrow_z\rangle\langle\uparrow_z| = |\uparrow_x\rangle\langle\uparrow_x|$。也就是任意方向的向上状态都相同。加上前面已经得到的某方向向上和向下状态相同，于是，任意方向任意状态都相同。这个可能吗？简单地说，给定 $|\uparrow_z\rangle\langle\uparrow_z|$ 和 $|\downarrow_z\rangle\langle\downarrow_z|$，对两者做测量得到的结果不一样；给定 $|\uparrow_z\rangle\langle\uparrow_z|$ 和 $|\uparrow_x\rangle\langle\uparrow_x|$，其结果也不一样。因此，这是不可能的。

上面的讨论是否说明了量子系统的确定性经典理论的不可能，同时，随机性的经典理论也不可能呢？

在回答这个问题之前，我们先停下来思考一下对经典随机性的认识角度：纯随机还是伪随机。实际上抛在空中以后落下来的硬币的状态，如果我们知道所有的信息——初始状态、空气阻力等等等等，其状态是可以通过经典力学计算出来的。因此，硬币的随机性实际上是一个本质上确定性的系统在信息不完全的条件下的有效形式，而不是最终形式。那么，是否上面这个形式上的随机性理论的要求，实际上也是由于信息不完全导致的呢？Einstein相信是这样的——“上帝不掷色子”。我们实验者看到了“色子”那是因为我们掌握的信息还不够充分。

也就是说，一个形式上的经典随机性理论，完全可能是一个经典确定性理论的信息不完全的时候的表现。实际上我们确实可以接受（或者不接受）这个“信念”。如果我们接受这个信念，那么关于量子系统的确定性的经典理论的寻求直接就成了“寻找一个形式上随机性的经典理论”。这个量子系统的随机性的经典理论的可能性就是我们下一节的主题，实际上这个可能性也是量子力学的隐变量理论研究的问题 [1]。如果你不接受这个信念，那么就把量子系统的确定性理论和量子系统的纯随机理论当作两个完全不同的可能的理论形式好了。在这个意义上，我们前面的讨论——给定确定的状态 $\rho = |\uparrow_z\rangle\langle\uparrow_z|$ 的情况下，测量得到的值和测量之后的状态可能是两个随机

出现的结果（$|\uparrow_x\rangle\langle\uparrow_x|$ 和 $|\downarrow_x\rangle\langle\downarrow_x|$）的一个——表明，量子系统的确定性理论是不可能的。同时，从实验结果推测出来 $|\uparrow_z\rangle\langle\uparrow_z| = |\downarrow_z\rangle\langle\downarrow_z|$ 也表明，如果我们的事件集合仅仅有 $\{|\uparrow_z\rangle\langle\uparrow_z|, |\downarrow_z\rangle\langle\downarrow_z|\}$，那么，随机的经典理论也是不可能描述量子系统的行为的。于是，至少，我们需要考虑更复杂的时间集合，例如 $\{|\uparrow_x\rangle\langle\uparrow_x|, |\downarrow_x\rangle\langle\downarrow_x| \ \cdots, |\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}|, |\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|, \cdots\}$ 以及 $\{|s_x \cdots s_{\hat{r}} \cdots\rangle\langle s_x \cdots s_{\hat{r}} \cdots|\}$。其中 $s_{\hat{r}} = \pm 1$。

关于量子力学是否能够由确定性理论描述的问题，除了企图构造出这样的理论之外，我们还可以从另一个角度来讨论：确定性的理论必须满足什么特征，这些特征是否被量子系统的行为所遵循。这个等到在第十一章讨论Bell 不等式的时候再来讨论。Mermin在一篇 Physics Today 的科普性文章 “Is the Moon There When Nobody Looks? Reality and the Quantum Theory” [22] 里面提供了另外一个经典确定性理论不可能描述量子系统的行为的证明：他设计了一个理想实验，构造了一个确定性经典模型，然后证明了无论这个经典模型的规则细节怎样，其理论结果永远不可能与其构造的理想实验相符。这是一篇非常值得一看的文章。然而，以后我们会看到，量子理论却能够得到与理想实验一致的结果。

2.5 量子系统的随机性经典理论的尝试

这一节我们来尝试构造量子系统的形式上的随机性的经典理论。我们已经看到：形式上，量子系统的行为，如果要用经典理论来描述，是要求随机性的经典理论的，确定性理论是不够的。当然，另外一方面，形式上的随机性的经典理论可以是本质上确定性的经典理论在信息不完全的条件下的表现。所以，在这个“信念”下，这一节的主题就是上一节的主题。

我们还是以 Stern-Gerlach 装置的一系列实验为例。首先，我们来看看一个未知状态的自旋通过 z 方向的磁场之后向上的那个状态的数学模型是什么。对于这个状态，我们知道如果让它通过另一个 z 方向的磁场，则它还是向上。于是，按照我们对于硬币的经验，这个状态应该是，

$$\rho_0^{c,z} = |\uparrow_z\rangle\langle\uparrow_z|. \tag{2.47}$$

这里，我们把通过这个仪器以后向上的就称为向上态 $|\uparrow_z\rangle\langle\uparrow_z|$。上一节我们已经提到，对于这个再一次通过 z 方向磁场的实验，这个数学模型就够了。测量的结果、测量后状态都能够很好的描述。

现在，我们让这个自旋通过 x 方向的磁场，而不是 z 方向的磁场。结果是我们已经了解的：有且仅有向上和向下两个可能。于是，我们自然想到实际上 ρ_0^c 应该是

$$\rho_0^{c,x} = \frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right). \tag{2.48}$$

如何让一个密度矩阵既满足公式 (2.47) 又满足公式 (2.48) 呢？一种可能就是，

$$|\uparrow_z\rangle\langle\uparrow_z| = \frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right). \tag{2.49}$$

关于这个关系如何导致——“任意方向的任意状态都相同”——的结果，我们在上一节已经讨论过。而这样的一个数学结果，肯定和实验不相符，和构建理论模型的目的不相符。我们构建量子系统的数学模型的目的就是为了对于给定的实验制备流程得到的状态能够写下来一个状态的数学表示，并且从这个表示能够做计算得到这个状态经过某个给定测量仪器得到的结果。现在，连某方向向上态和向下态数学上都不能区分了，怎么可能将来预测出来经过仪器以后是向上还是向下输出呢？因此，在这里，我们不再讨论这样的数学形式。但是，要注意，将来的量子系统的量子力学数学形式，却和这个数学形式非常像，

$$|\uparrow_z\rangle\langle\uparrow_z| = \frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x| + \Delta\right), \tag{2.50}$$

$$|\downarrow_z\rangle\langle\downarrow_z| = \frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x| - \Delta\right). \tag{2.51}$$

就是后面这个额外的“Δ”解决了所有的问题，也是我们所要学习的量子理论的核心。在学习这个额外的“Δ”之前，我们继续量子系统的经典理论的可能性的探索。

另一种可能是，要么

$$\rho_0^c = [|\uparrow_z\rangle\langle\uparrow_z|] + \left[\frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right)\right]. \tag{2.52}$$

要么

$$\rho_0^c = [|\uparrow_z\rangle\langle\uparrow_z|] \cdot \left[\frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right)\right]. \tag{2.53}$$

前者假设两个不同的方向的状态是互斥事件，于是整体的概率分布函数等于各自的概率分布函数之和；后者假设两个不同的方向的状态是两个相互独立的变量，于是联合分布函数等于独立分布函数的乘积。第一种情况相当于假设我们的基本事件集合是 $\{|\uparrow_x\rangle\langle\uparrow_x|, |\downarrow_x\rangle\langle\downarrow_x| \ \cdots, |\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}|, |\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|, \cdots\}$。

第二种情况相当于假设我们的基本事件集合是 $\{|s_x \cdots s_{\hat{r}} \cdots\rangle\langle s_x \cdots s_{\hat{r}} \cdots|\}$。对于第一种，实际上我们需要重新归一化。不过这个技术问题，这里，我们就假设总是可以简单处理的。

2.5.1 互斥构造

现在，我们来看看基于互斥事件的公式 (2.52) 是否能够解释实验结果。按照经典概率论，对于形如公式 (2.52) 的概率分布函数的测量可以得到所有的可能结果：这里包含 $|\uparrow_z\rangle\langle\uparrow_z|$、$|\uparrow_x\rangle\langle\uparrow_x|$，以及 $|\downarrow_x\rangle\langle\downarrow_x|$。也就是说，我们不能够再确定测量自旋的某一个方向——“我想测量自旋的 x 方向”，而是只能够做这样的测量：“我想测量一个自旋”。这一点与我们的实验是不相符的。在实验中，我们可以选择哪一个方向来测量。

做一个类比。这个问题就好像是一个硬币本来有正反面、颜色、形状等等等等的各种属于不同的空间的属性。现在，有一个硬币它的正反面、颜色、形状都是同等地位的在同一个空间中的属性，那么你自然就不能够在提出这样的问题：“我想测量这个硬币的颜色到底是绿的还是红的”，而仅仅能够题这样的问题：“我想测量这个硬币”。可是这个神奇的硬币竟然能够知道你的真实的企图，在你想测量颜色的时候给你颜色，形状的时候形状。看起来好像这个硬币和观测者的意识之间建立了某种联系[1]。

当然，你可以说，正是由于你选择了哪一个方向来测量，你选择了某种类型的仪器，而这个自旋通过跟这个被选择了的仪器相互作用，了解了“你的真实的想测量的方向”。于是，通过某种方式，“你的意识”跟实验结果建立了关联。尽管我们非常反对把实验者的意识和物理对象关联起来（因为我们希望物理对象的状态还是一个客观量），但是我们不能直接否定这个基本上不可证伪的理论。对于这样的辩护，一个实用主义的物理学家只能继续问你，所选择的仪器如何与自旋的状态建立起来的关联，从而实现了这个试验结果呢？一个完整的理论必须能够回答这个问题。当然，你也可以耍赖，说，这个问题现在还回答不了，但是每次这个未知的机制总是能够发挥作用。由于这个机制的问题没有给出答案，也暂时不可证伪。让我们暂时允许这样的理论，继续思考下去。

这样的一个仪器在经典随机理论的框架之内是可以实现的，而且非常

[1]实际上，有人就是这样来理解和看待量子力学的。究其原因，就是这样的人还在企图用经典力学的数学和来自于其自身对经典世界的经验来“理解”量子力学，就像咱们在这里做的一样。以后我们会看到，量子系统的行为从量子力学的数学框架来看，根本不要求这些神奇的理解或者说解释方式。

简单：把真正希望测量的那部分信息识别出来，把识别不出来的其它的信息扔掉。在硬币的例子上，你可以设计一个仪器，这个仪器只能够识别颜色，如果其它的可能状态过来了，这个仪器就显示一个探测不到的状态。也就是说，这个仪器只对颜色——这个反映真实测量意图的变量——敏感。这个仪器的缺陷是有的时候我们会得到得不到信号的结果。而且这个得不到信号的结果出现的概率由原始的公式 (2.52) 的概率分布函数决定。

现在，我们来看一看这个测不到信号的结果出现的几率。到现在为止，我们仅仅讨论了 z 方向和 x 方向的测量，现在我们来讨论 y 方向和任意 $\hat{r}(\theta,\phi)$ 方向的测量。为了满足 y 方向测量的结果（可能是两个结果，等几率，每一次单次的测量仅得到一个结果），我们需要修改公式 (2.52) 为，

$$\rho_0^c = [|\uparrow_z\rangle\langle\uparrow_z|] + \left[\frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right)\right] + \left[\frac{1}{2}\left(\left|\uparrow_y\right\rangle\left\langle\uparrow_y\right| + \left|\downarrow_y\right\rangle\left\langle\downarrow_y\right|\right)\right]. \quad (2.54)$$

现在，我们来归一化这个分布函数，

$$\rho_0^c = \frac{1}{3}[|\uparrow_z\rangle\langle\uparrow_z|] + \left[\frac{1}{6}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right)\right] + \left[\frac{1}{6}\left(\left|\uparrow_y\right\rangle\left\langle\uparrow_y\right| + \left|\downarrow_y\right\rangle\left\langle\downarrow_y\right|\right)\right]. \quad (2.55)$$

于是当我们用测量 z 方向的仪器做侧量的时候，我们有 $\frac{2}{3}$ 的几率得不到任何信号。可是，在我们的实验当中，这样的事情从来没有出现过。$\frac{2}{3}$ 是一个很大的几率，不可能和实验误差混在一起区分不开。

这个问题还会成为一个更大的问题。为了满足 $\hat{r}(\theta,\phi) = \sin\theta\cos\phi\hat{i} + \sin\theta\sin\phi\hat{j} + \cos\theta\hat{k}$ 方向的测量（可能是两个结果，向上的几率是 $\frac{1+\cos\theta}{2}$ 向下的几率是 $\frac{1-\cos\theta}{2}$，每一次单次的测量仅得到一个结果），我们是否需要再把这个 $\hat{r}$ 方向的分布函数添加到公式 (2.52) 里面呢？

是否可以把 $\hat{r}(\theta,\phi)$ 方向的测量看作是三个独立方向 x、y、z 的测量的结合呢？从平均值上来说，可以，

$$\langle\sigma_{\hat{r}}\rangle = \sin\theta\cos\phi\langle\sigma_x\rangle + \sin\theta\sin\phi\langle\sigma_x\rangle + \cos\theta\langle\sigma_z\rangle = \cos\theta. \quad (2.56)$$

但是，这样的话，单次测量的结果可能是

$$\pm\sin\theta\cos\phi \pm \sin\theta\sin\phi + \cos\theta \neq \pm 1. \quad (2.57)$$

所以，必须把所有的 $\hat{r}$ 方向的分布函数添加到公式 (2.52) 里面，也就是说，

$$\rho_0^c = \sum_{\hat{r}(\theta,\phi)} \rho^c(\theta,\phi), \quad (2.58)$$

其中

$$\rho^c(\theta,\phi)=\frac{1+\cos\theta}{2}|\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}|+\frac{1-\cos\theta}{2}|\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|. \tag{2.59}$$

如果你这个时候再来做一个归一化的话，任何一个单独的 $\rho(\theta,\phi)$ 都是无穷小，也就是说，如果你用一个测量 z 方向的仪器做侧量，绝大多数的情况下得不到任何信号。这一点我们从来都没有在实验上观察到过。

因此，除非有经典概率论之外的数学结构或者是让自旋了解实验者的意识的未知的机制——除非你写下来这个机制的数学模型，要不然相当于把量子系统的数学模型的问题转化成为这个机制的数学模型的问题。看起来，基于互斥事件构造起来的经典理论公式 (2.58) 不能描述量子系统的行为。还有没有其它的可能的经典理论呢?

2.5.2 独立构造

现在，我们来看看基于独立事件的公式 (2.53) 是否能够解释实验结果。按照经典概率论，对于两个独立变量 x_1,x_2 的联合分布函数 $\rho(x_1,x_2)=\rho^1(x_1)\rho^2(x_2)$，我们有

$$\begin{aligned}\left\langle A^1(x_1)\right\rangle&=\iint dx_1dx_2A^1(x_1)\rho(x_1,x_2)\\&=\int dx_1A^1(x_1)\rho^1(x_1)\int dx_2\rho^2(x_2)\\&=\int dx_1A^1(x_1)\rho^1(x_1).\end{aligned} \tag{2.60}$$

用密度矩阵的语言，这个就是

$$\left\langle A^1\right\rangle=tr^{1,2}\left(A^1\otimes I^2\rho^1\otimes\rho^2\right)=tr^1\left(A^1\rho^1\right). \tag{2.61}$$

这里 I^2 是第二个变量所在的空间的单位矩阵，$\otimes$ 表示矩阵或者是算符的直积。独立变量的联合分布保留了每一个变量自己的分布函数的所有的性质。于是，我们发现，公式 (2.53) 能够同时解释我们的两个不同方向上的实验结果。这个结果也是符合经典随机粒子的经验的。这里的 z 方向和 x 方向就好像是硬币的正反面和硬币的颜色一样是两个独立的变量。只要各自的分布函数满足各自的测量的结果的要求，那么整体来看必然是与所有的联合起来的实验结果一致的。

实际上，除了 x 方向和 z 方向，我们还需要考虑至少 y 方向，于是公式 (2.53) 应该成为如下的形式，

$$\rho_0^c = [|\uparrow_z\rangle\langle\uparrow_z|]\cdot\left[\frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right)\right]\cdot\left[\frac{1}{2}\left(\left|\uparrow_y\right\rangle\left\langle\uparrow_y\right| + \left|\downarrow_y\right\rangle\left\langle\downarrow_y\right|\right)\right]. \tag{2.62}$$

是不是还需要包含其它方向，例如任意的 $\hat{r}$ 方向呢？答案是需要。上面的密度矩阵的隐含假设是物理量 σ_x、σ_y、σ_z 是独立的，于是测量例如 σ_x 的时候，

$$\langle\sigma_x\rangle = tr\left(\rho_0^c\sigma_x\otimes I^y\otimes I^z\right) = tr\left(\rho_0^{c,x}\sigma_x\right). \tag{2.63}$$

类似的，如果测量 σ_y 的话，

$$\left\langle\sigma_y\right\rangle = tr\left(\rho_0^c I^x\otimes\sigma_y\otimes I^z\right) = tr\left(\rho_0^{c,y}\sigma_y\right). \tag{2.64}$$

对于任意一个方向的 Stern-Gerlach 装置的测量，我们需要做到单次测量必然是向上或者相下态的一个，多次测量得到两个斑点。于是，我们希望得到类似的，

$$\rho_0^{c,\hat{r}} = \left(\frac{1+\cos\theta}{2}|\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}| + \frac{1-\cos\theta}{2}|\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|\right). \tag{2.65}$$

其中，概率 $\frac{1\pm\cos\theta}{2}$ 的选取是为了保证（这一点量子系统的实验是符合的）

$$\langle\sigma_{\hat{r}}\rangle = \sin\theta\cos\phi\langle\sigma_x\rangle + \sin\theta\sin\phi\left\langle\sigma_y\right\rangle + \cos\theta\langle\sigma_z\rangle = \cos\theta. \tag{2.66}$$

为了实现这个公式 (2.65)，我们不得不让

$$\rho_0^c = \prod_{\hat{r}}\rho_0^c\left(\theta,\phi\right), \tag{2.67}$$

其中，

$$\rho^c\left(\theta,\phi\right)_0 = \frac{1+\cos\theta}{2}|\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}| + \frac{1-\cos\theta}{2}|\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|. \tag{2.68}$$

也就是说，我们把各个方向上的测量都看成是这个自旋的独立的属性。这样至少看起来对于所有的可能的测量就没有问题了。

现在我们来考虑让这个系统再经过一次测量，例如先过 z 方向挡住向下的部分，然后再过 x 方向挡住向下的部分，然后再让这部分的自旋经过一个 $\hat{r}$ 方向的测量。按照经典概率论，**对于一个变量的测量不改变其它独立变量的分布**。做一个类比，考虑两个独立的随机硬币的状态，测量其中一个，

另一个的状态的分布函数不会改变。于是我们猜测，经过第二次测量（先 z 后 x）以后的状态应该是，

$$\rho_1^c = \rho_1^c(z) \cdots \rho_2^c(x) \cdots \rho_0^c(\theta, \phi), \tag{2.69}$$

也就是说，其他的方向上都不变，但是 z 和 x 方向上成了新的形式，

$$\rho_1^c(z) = |\uparrow_z\rangle\langle\uparrow_z|, \ \rho_2^c(x) = |\uparrow_x\rangle\langle\uparrow_x|. \tag{2.70}$$

这个时候，如果我们再来测量这个状态的 $\hat{r}$ 方向。第一，$\hat{r}$ 就是 x 方向，结论是“只能得到向上”，符合实验结果。如果 $\hat{r}$ 就是 z 方向，因为 $\rho_1^c(z) = |\uparrow_z\rangle\langle\uparrow_z|$，结论是“只能得到向上”，与实验结果不相符。于是，我们发现，这个时候，其实 $\rho_1^c(z) = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\downarrow_z\rangle\langle\downarrow_z|$ 才能得到与实验相符的结果。也就是说，当测量 x 方向的时候，本来应该独立的 z 方向的随机变量的分布函数变了。同理，同时其他的所有的 $\hat{r}$ 方向的随机变量都要作相应的变化。

这个是什么样的独立变量的独立分布函数呀？**测量“符合独立分布的独立变量”中的一个，其它的都会发生改变！**测量后状态成了一个“神奇的世界”的代名词。测量后状态不再是经典概率论里面的图景：一个随机变量的测量后状态就是观测到的测量结果对应着的状态。我们被迫构造出一个这样的理论：一个系统的测量后状态不再是得到的观测值对应的状态，仅仅是一个这样的状态——它能够保证如果你再做一次任何方向上的测量，这个状态给出来的理论结果与实验结果相符。也就是说，在这个理论框架里面，测量以后系统是什么状态这个问题不是一个好问题，我们仅仅能够问：测量以后，如果我们再一次测量（同样的方向或者不同的方向），我们得到的结果是什么？当然，一个这样的理论也可以是一个可证伪的科学的理论。可是，这样一个把测量后状态的图景都破坏的理论，把分布函数中的所假设的独立变量的独立性都在测量中波坏的理论，是什么样的一个理论呀？

我们之所以尝试构造一个能够解释量子系统的行为的经典理论，是因为我们觉得经典理论，包含确定性的和随机性的经典理论，是一个我们能够理解的，觉得舒服的理论。可是，如果这个理论需要我们放弃经典理论中的那些能够理解的部分，例如状态是客观的，与观测者的意识没有关系，测量后状态就是简单地测量结果对应的状态，我们为什么还一定要追寻一个量子系统的经典理论呢？一个测量以后系统的状态不再是所观测到的状态，不在满足独立变量的测量互不干扰的理论，还是我们希望得到的理论吗？

为什么我们在这里仅仅考虑互斥构造和独立构造——在这里不管哪一种构造各个方向的相对地位一样，是否还应该考虑各个方向上原则上就相对地位不一样的构造？这是物理学的另外一个不愿意轻易突破的理念：各个方向是平权的。不过，退一万步，如果突破它就能够得到一个量子系统的经典理论，我想还是有物理学家愿意这样做的。

2.6 作业

习题 2.1 阅读 W.I.B. Beveridge 的《科学研究的艺术》(《The Art of Scientific Investigation》) [23]，写下阅读体会。

习题 2.2 阅读 Gowers T. 的《数学-牛津通识读本》(《Mathematics: A Very Short Introduction》) [24]，写下阅读体会。

习题 2.3 阅读 R. Feynman 的《物理定律的特性》(《The Character of Physical Law》) [25]，写下阅读体会。

习题 2.4 阅读 A. Einstein 和 L. Infeld 的《物理学的进化》(《The Evolution of Physics: The Growth of Ideas From Early Concepts to Relativity and Quanta》) [26]，写下阅读体会。

习题 2.5 通过 Google 和 Wikipedia 来了解和学习以下的一个主题，写一份学习报告：随机数生成器、博弈论、量子远程传态、量子多世界解释、量子力学的其它解释。学会自己寻找材料来学习，甚至自己寻找主题来学习，是研究性学习非常重要的一步。本书在很多地方匆匆而过，当然也在很多地方非常深入，需要你自己学习的地方非常多。Google 和 Wikipedia 是好朋友。

习题 2.6 从公式 (2.34) 出发，反推出来以下状态的几率：$p(s_2|s_1=1)$，它表示已知 $s_1=1$ 的时候，s_2 的可能取值和相应的几率。这个问题从另外一个角度证明了公式公式 (2.30)、公式 (2.31)、公式 (2.32) 实际上都包含在公式 (2.34) 中。

2.7 本章小结

由于本章的内容代表了非常深入的思考，这个思考的细节也有很多转折，我们这里提供一个比较详细的小结，而且把几个重要的公式和结论也都

重新写在这里。在本章里面，我们企图给自旋的状态构建一个经典的数学模型。所谓经典的就是遵循经典概率论的随机的或者确定性的。这个数学模型需要对各个方向的单次和再次——这里再次指的是对已经测量了一次的自旋做再一次的测量而不是重新制备和测量另一个自旋，后者我们称为系综测量——给出和实验相符的结果，也就是公式 (2.34)。

为了找到这样的一个理论，最关键的事情是写下来密度矩阵 $\hat{\rho}\left(s_{\hat{r}_1}\right)$ 或者说密度分布函数 $\left\langle s_{\hat{r}_2}\middle|\hat{\rho}\left(s_{\hat{r}_1}\right)\middle|s_{\hat{r}_2}\right\rangle$ 使得按照经典概率分布的平均值公式 (2.36) 的计算得到公式 (2.34) 的结果。注意，这里这些经典密度矩阵或者叫做密度分布函数都是对角的。其中 $s_{\hat{r}_1}$ 的含义是，对 $\hat{r}_1$ 方向的自旋状态做测量并且得到测量值为 $s_{\hat{r}_1} \in \pm 1$。因此，$\hat{\rho}\left(s_{\hat{r}_1}\right)$ 是这样的测量和测量结果以后的状态。

对于这个可能性，首先我们发现确定性的经典理论的问题归结为随机性的经典理论的问题：第一、量子系统的形式上的确定性的理论是不可能的；第二、本质上确定性的经典理论可以是由于信息不完全导致的经典随机理论。然后，我们讨论了量子系统的经典随机理论的可能性。我们发现，如果考虑某一个方向的测量，那么硬币的数学模型，就能够很好地描述量子系统。但是，当量子系统面对再次测量（需要同时符合单次测量和多次平均的得到的结果）的时候，尤其是再一次测量的方向和本次测量的方向不一样的时候，硬币的数学模型就不能用了。

因此，我们拓展了的硬币的经典理论，把各个方向的状态看作互斥事件或者是独立事件。对于测量了 $\hat{r}_1$ 并且得到 $s_{\hat{r}_1} \in \pm 1$ 其中之一的情况，互斥和独立两种方式构造了这个分布函数分别是

$$\rho^c = \frac{1}{\mathcal{N}} \sum_{\hat{r}_2} \rho^c \left(s_{\hat{r}_2} | s_{\hat{r}_1}\right), \tag{2.71}$$

和

$$\rho^c = \prod_{\hat{r}_2} \rho^c \left(s_{\hat{r}_2} | s_{\hat{r}_1}\right), \tag{2.72}$$

其中

$$\rho^c \left(s_{\hat{r}_2} | s_{\hat{r}_1}\right) = \frac{1 + s_{\hat{r}_1} \hat{r}_2 \cdot \hat{r}_1}{2} \left|\uparrow_{\hat{r}_2}\right\rangle \left\langle\uparrow_{\hat{r}_2}\right| + \frac{1 - s_{\hat{r}_1} \hat{r}_2 \cdot \hat{r}_1}{2} \left|\downarrow_{\hat{r}_2}\right\rangle \left\langle\downarrow_{\hat{r}_2}\right|. \tag{2.73}$$

自然对于更一般的情况，如果测量了 $\hat{r}_1$ 并且得到 $s_{\hat{r}_1} \in \pm 1$ 的两个结果都进入下一步实验，则数学描述需要做一个概率性叠加，也就是，

$$\rho^c \left(s_{\hat{r}_2} | s_{\hat{r}_1}\right) = \rho^c \left(s_{\hat{r}_2} | s_{\hat{r}_1} = 1\right) p_{s_{\hat{r}_1} = 1} + \rho^c \left(s_{\hat{r}_2} | s_{\hat{r}_1} = -1\right) p_{s_{\hat{r}_1} = -1}. \tag{2.74}$$

但是，这样的密度分布函数我们发现：第一、对于互斥构造，实验者将不能主动选择特定的观测方向，除非自旋能够有某种方法知道实验者的意图，否则只能够得到所有的可能的 $\hat{r}_2$ 中的一个。于是如果从满足实验者的意愿观测方向来说，会出现大多数时候观测不到想要的观测方向的事情。这一点在实验上没有观察到过。实际实验中，实验者总是可以选择特定的方向来观测的，而且观测到某个值的成功几率不是非常小，尽管也有观测不到结果的时候[2]。此外，在互斥构造中归一化常数 N 的取值也是一个问题。第二、对于独立构造，竟然会出现测量这些独立变量中的一个并且得到测量值会改变其他“独立”的变量的分布函数的事情。这样的约定好的独立变量，数学形式上也表现为独立变量的变量在测量的过程中相互影响，是基本不可以接受的事情。第三、自旋算符之间的代数关系

$$\sigma_{\hat{r}} = \sin\theta\cos\phi\sigma_x + \sin\theta\sin\phi\sigma_y + \cos\theta\sigma_z \tag{2.75}$$

也被这两个理论破坏了，各个方向上的自旋必须当做经典的互斥或者独立随机变量。因此，这两种密度矩阵的形式都不是好的经典理论的形式。于是，我们初步得到结论：经典理论看起来基本上不能解释量子系统的行为。

那么，还有没有其它的除了互斥构造和独立构造之外的可能的经典理论呢？当然可以有。但是，这样的理论就相当于假设这个物理世界的各个方向其实是有差别的。这一点，基本上是不可以接受的。

量子力学是量子系统的行为的数学模型。我们将来也需要来看一看是否这个理论保留了状态的客观性，保留了测量后状态与测量前状态的简单的联系，来看一看是否量子系统需要了解实验者的意识。我们的目标是建立一个尽量保持经典随机理论——甚至经典确定性理论——的各种数学结构的并且尽可能简单的理论。具体来说，这些数学结构和简单图景包含：状态的数学描述是什么，可观测量的数学描述是什么，测量的结果（单次测量和多次平均）是什么，测量以后的状态是什么，状态如何演化。量子系统的行为指的是本书第一章中的典型实验的结果——本质上就是**单个粒子**通过一个仪器（偏振片、Stern-Gerlach 装置等）或者多个仪器**的**各种**干涉现象**，或者以单个自旋的测量为例，就是公式 (2.34)。满足这个要求的量子理论的细节见第六章和 11.4 节。

另外，除了理论上构造出来这个经典密度矩阵，然后讨论这个理论的可

[2]量子力学认为，这是由于例如某一个光子被环境干扰了所以没有被探测到。当然，这样的几率随着仪器质量的提高，是很小的。也就是说，实验中观测到想要观测的方向的可能性远远大于 $\frac{1}{N}$，这个要多小就有多小的数。

接受性，还有一种方式来讨论量子系统的经典理论的可能性：证明一个所有的一定条件下的经典理论都必须满足的结论，然后检验量子系统的行为是否满足这个条件。在第十一章要介绍的 Bell 定理或者说Bell 不等式就沿着这样的方式展开讨论。我们认为，这两种方式是互补的。

我们将来在第十一章 11.5 节还会看到我们构造的这个互斥随机变量和独立随机变量的理论还有另外一个的问题——非定域性。也就是说，即使如果你愿意接受一个具有上面提到的这么多问题的数学模型作为量子系统的理论，那么，它还有非定域性的问题：大概来说就是两个遥远的不同系统上的测量之间单次的结果各自应该是独立的——没有信息能够相互交流来建立起来这个关联。我们将来还会看到这个定域性的更准确的数学形式的表达。顺便，既然我们提到定域性的问题，我们将仅仅讨论和展示前面所构造的经典理论的“非定域性”，而不是将来要构建的量子系统的量子力学的非定域性。

一句话总结一下：尽管不能完全否定，但是看起来用经典理论描述量子系统非常地不可能。因此本章的真正目的，除了讨论一个量子系统的经典描述的可能性这个问题本身之外，其实，是为了回答为什么我们需要新的数学理论来描述量子系统的行为，从而为了学习量子系统的量子力学做准备。

第二部分

数学物理学上的准备

为了构造能够描述上一部分中的量子系统的行为的数学模型——将来我们称之为量子力学或者量子理论，我们需要做一些数学和物理上的准备。这个准备包含矢量空间、概率论和经典力学。矢量空间是将来量子理论的核心。概率论是将来的量子理论的经典对应，或者说，量子理论就是经典概率论的一种拓展。通过经典力学的学习，我们可以了解用 Hamiltonian 来描述物理系统的方式。另外，抽象矢量记号——Dirac 符号将会是我们以后的模型使用的主要的数学语言，也要在这里先学习一下。此外，对经典概率和经典随机客体的测量的理解也是一个问题：单个硬币是随机的吗？在学习的时候要注意这些内容和将来的内容之间的联系，也要思考有助于深刻理解这些内容的问题。

这一部分的主要参考书，Landau的《力学》[27]，喀兴林的《高等量子力学》[28]，吴兆颜的《高等量子力学》[29]。Kolmogorov的《概率论导引》[19] 也非常值的一看。

第三章　二维空间的线性代数：Dirac 符号、抽象算符与表象

为了学习量子系统的数学模型，我们需要一点点数学上的准备。这里，数学上的准备分成两个层次，计算上和理解上的。在本章中，我们会把计算上的数学内容控制在非常简单的二维矩阵和二维矢量的运算上。但是，对于这些矩阵和矢量的理解要比通常的线性代数高很多，需要接触到一点点近世代数的内容。当然，我们会企图把整个概念上比较复杂的部分都尽量地通过二维空间的形式来讨论。

另一方面，在本章中，我们会讨论数学和物理学，数学和科学的关系，以及在理解这个关系的基础上如何学习数学。

这一部分的推荐阅读材料是Einstein和Infeld的《物理学的进化（The Evolution of Physics）》[26]、Feynman的《物理定律的本性（The Character of Physical Law）》[25]、Beveridge的《科学研究的艺术（The Art of Scientific Investigation）》[23]、Gowers的《数学是什么（Mathematics: A Very Short Introduction）》[24]、Axler的《线性代数应该这样学（Linear Algebra Done Right）》[30]。前四本是关于什么是物理学和什么是科学，以及科学和数学之间的关系的。后一本是线性代数方面的具体知识的。其主要关注点在于线性空间的矢量和算符，和本书的思路类似，跟大多数线性代数的书不一样，值得推荐。学有余力的还可以参考喀兴林的《高等量子力学》[28] 的第一章。

3.1　数学是现实世界的结构和思维的语言

数学本质上是对事物之间的联系的描述。事物通常用集合以及集合里面的元素来描述。而元素之间的关系，有的时候也称作结构，则通过映射来描述。把数学理解成结构的语言，而不仅仅是一堆运算规则，对于创造和使

用数学是非常有意义的。例如，我们从小就会计算“$1+1=2$”，“一个苹果加上一个苹果等于两个苹果”。看起来，我们是在做“苹果的加法”。我们来仔细看一看，我们在计算什么东西的加法。定义一个苹果的集合 P，为简单计，我们假设总数是有限个，但是是一个很大的数。所谓映射，就是把一个集合中的元素与另外一个集合中的元素联系起来。现在，我们猜测，“一个苹果加上一个苹果等于两个苹果”，是集合 $P\otimes P$ 到 P 的映射（需要两个来自于 P 的元素，得到一个 P 里面的元素）。真的是这样吗？我们得到的结果，也就是“两个苹果”，是集合 P 中的元素吗？注意，集合 P 的元素是一个又一个的苹果，尽管颜色大小酸甜等都可以不一样，甚至有一个特大的苹果可以等于两个其中的小的苹果那么大，但是，两个苹果绝对不是集合 P 中的元素。也就是说，我们实际上，不是在计算苹果的加法。

实际上，我们计算的是苹果的数量的加法。从集合 P 开始，构造一个 P 的所有子集的集合——幂集 $\mathcal{T}=\{Q|Q\subseteq P\}$。这个时候我们发现，取 $\mathcal{T}$ 中的元素 t_1 和 t_2，如果 $t_1\cap t_2=\phi$，也就是不相交没有共同的元素，则 $t=t_1\cup t_2$ 必然是 $\mathcal{T}$ 中的元素，也就是 $t\in\mathcal{T}$，并且，如果定义集合的大小，也就是集合中元素的个数，$|t|$，我们发现

$$|t|=|t_1|+|t_2|. \tag{3.1}$$

这个才是我们计算的加法。也就是说，加法是定义在苹果集合 P 的幂集上的计算，并且只考虑集合大小这个特性。或者说，苹果的加法实际上在计算苹果集合 P 的子集的并的运算，并且要求用来计算的子集没有交集。简答地说，“苹果的加法”是没有共同元素的“集合的并运算”，而不是直接在苹果集合上的运算——我们**不能把两个苹果相加得到另一个苹果**。

通过上面这个例子，我们看到，第一，集合和映射的语言可以把问题说清楚；第二，苹果之间不存在直接的加法，通常的加法实际上是集合的并运算；第三，数学是描述事物之间关系的语言。关于数学的这个理解——数学是对事物的关系的描述（事物就是集合，关系就是映射）[24]，对于我们以后的学习是非常重要的。

在这里，为了强调数学是对事物之间的关系和结构的描述，我们再来讨论另外一个例子：矢量加法。中学物理我们就学习过矢量的叠加和分解。我们知道位移、速度、加速度和力是矢量。现在，我们来思考一下，为什么我们需要矢量这个数学对象来描述位移、速度、加速度和力这些物理概念，同时我们也思考一下这个问题：位置是不是矢量。首先，在运动问题中，我们注意到运动物体不仅有速度的大小，还有速度的方向。在某些力的作用下，

速度的方向可以发生改变，例如拉着一根绳子转一个小球。而且，更进一步，速度的大小和方向的改变，和力的大小和方向有联系。于是，我们就发现，需要强度和方向两个变量来描述位移、速度、加速度和力这些物理量。我们暂时称这样的一个东西为“矢量”，将来我们看一看它是不是真的是满足矢量运算——例如矢量加法——的矢量。这个“矢量”暂时记作 (r,θ)。其中前者表示强度，后者表示方向——实际上可以是多个变量，例如用多个角度来表示 3 维空间的矢量。那下一步，就是看这样的数学对象的集合上面可以定义什么操作。以位移“矢量”为例，我们可以把前后两个的位移合起来讨论合起来的位移。以力“矢量”为例，我们可以讨论同时作用在一个物体上的两个力的效果，看其是否等于这两个力各自的效果的叠加。经过物理学的观察和实验，我们发现，前后两个时间段的位移合起来确实还是位移，同时作用在物体上的两个力的效果可以看成一个合力的效果，而这个合成遵循平行四边形法则——也就是矢量的各个方向上的分量可以相互加起来，但是，不是上面的两个坐标 r 和 θ 的值直接加起来。这个“矢量”的分量形式（注意不一定是坐标形式）加起来的操作，以后我们会看到，正好就是矢量空间的矢量的加法。因此，矢量就成了位移、速度、加速度和力这些物理量的数学模型。

现在，我们来讨论第二个问题：位置是不是矢量。给定一个位置，给定一个新的位置，我们能够通过一个自然——也就是和某种我们所感兴趣的物理过程直接相关——的运算，得到一个新的位置吗？例如一个处于平面上 (x_p,y_p) 位置的 P 点，和一个处于平面上 (x_q,y_q) 位置的 Q 点，用某种方式加起来，会给我们一个处于例如平面上某个 (x,y) 的 R 点吗？你可能会觉得 R 在 (x_p+x_q,y_p+y_q) 点很自然。其实，如果是这样我们计算的是从原点开始的两个“位移矢量”的加法而不是两个“位置矢量”的加法。实际上，对于位置来说，它只有坐标，连矢量都不是，更何况加法。至少，在考虑物体运动的范围内，我们完全没有对应着这种运算的物理过程。回到前面的苹果集合上加法的例子——加法实际上不是定义在苹果上，而是苹果集合的幂集，而且仅考虑幂集元素（也就是原始苹果集合的子集）的大小，现在我们发现位置这个集合上的加法运算也不是定义在位置集合上的，而是定义在位移集合上的。

物理学家，或者更一般地来说，科学家（甚至数学家），学习数学，要注意思考数学概念提出的动机、数学概念的实例、数学定理在物理或者现实中的含义。从创造性地使用数学结构和创造数学结构目的来说，了解数学结

构提出的动机和一些实例是很有好处的。

3.2 从分量形式的矢量和矩阵到抽象矢量和算符，Dirac 符号

我们先来复习一下实数作为元素的线性代数，然后学习 Dirac 符号，接着用 Dirac 符号来学习复数作为元素的线性代数。真正从集合映射开始来学习矢量和对偶矢量已经超出本书的范围。有兴趣的学习者可以参考喀兴林的《高等量子力学》[28]，以及吴金闪的《系统科学导引》[31]。

3.2.1 实数域上的分量形式的矩阵和矢量，复习

通常的线性代数课程定义矩阵就是一个由数字——通常是实数——排成 M 行 N 列的方块。对于这个方块，我们可以计算行列式，我们可以定义这个方块之间的乘法、加法和数乘。方块和列向量之间的乘法是方块和方块之间乘法的一个特例。这里我们重复一下这种方块形式的运算的定义。现在这些定义看起来好像是需要死记硬背的。先不要着急，将来我们会更加清楚其中的道理。对于数字方块 $A = (A_{ij})_{M\times N}$ 和 $B = (B_{ij})_{N\times L}$，我们定义其乘法得到的方块为：

$$(AB)_{ij} = \sum_{k=1}^{N} A_{ik}B_{kj}. \tag{3.2}$$

如果要计算乘法必须要求左边矩阵的列数等于右边矩阵的行数。而且，通过这个定义我们就看到，一对能够做 AB 乘积的矩阵不一定能够做 BA 乘积。所以，一般地来说，矩阵乘法是不可交换的。甚至当两个矩阵都是同样维数的方阵的时候也是这样。为了描述这种不可交换的关系，我们有关于对易子的定义，

$$[A, B] = AB - BA. \tag{3.3}$$

$[A, B]$ 称为矩阵 A 和 B 的对易子。如果要计算矩阵加法，那么这两个矩阵必须有同样的行数和列数，

$$(A + B)_{ij} = A_{ij} + B_{ij}. \tag{3.4}$$

数乘的定义非常简单，一个矩阵 A 和一个实数 a 的乘积定义如下，

$$(aA)_{ij} = aA_{ij}. \tag{3.5}$$

基本的定义仅包含乘法、加法和数乘。如果这部分不是复习，那么，请回去先学习线性代数再来学习本书，或者继续——如果你认为你已经准备好了接受挑战，本书的内容也是自足的。

由于在以后我们的讨论中我们基本上只关心方阵，所以，在不加说明的时候，以下的部分说到矩阵的时候就是指 $N\times N$ 的方阵。相应的，在提到矢量的时候，一般指的是 N 个实数排成一列的列向量。有的时候，我们也用左矢量，或者行向量的语言。它们就是 N 个数字排成一行的那个东西。对于一个行向量来说，其维数是 $1\times N$，所以只能够乘在一个 $N\times N$ 矩阵的左边。这就是为什么它还被叫做左向量。同样的原因，列向量有的时候也被称作右矢量。

两个右矢量可以讨论它们的点积或者叫做内积，定义如下，

$$(u,v)=\sum_j u_j v_j \triangleq u^T v. \tag{3.6}$$

这个矢量的转置的定义就是把一个列矢量的元素按顺序排成一个行矢量，或者反过来。这个转置也可以定义在矩阵上，

$$(A^T)_{ij}=A_{ji}. \tag{3.7}$$

可以证明 $(AB)^T=B^TA^T$。一个数，也就是 1×1 的矩阵的转置就是它自己。如果两个矢量的内积为零，则我们称这两个矢量相互正交。一组两两相互正交的矢量可以用来作为这个空间的基矢。

在我们后面的计算中，我们偶尔会用到矩阵的行列式。这里我们仅仅给出 2×2 矩阵的行列式的定义，

$$\det\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}=A_{11}A_{22}-A_{12}A_{21}. \tag{3.8}$$

这一节，我们复习了矩阵（包含矩阵的特例——行矢量和列矢量）的乘法、加法、数乘、转置的定义，以及右矢量的内积，还有 2×2 方阵的行列式。

3.2.2 Dirac 符号

在线性代数的课程里面，我们通常把一个形如 $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ 的列矢量称为矢量。在多元微积分或者矢量微积分里面，我们也常常把一个位置矢量写作

$\vec{r}=\begin{bmatrix} x \\ y \\ z \end{bmatrix}$，然后我们说 x 是这个矢量 $\vec{r}$ 的 x 方向上的分量。有的时候，我们也把这个关系写作，

$$\vec{r}=x\hat{i}+y\hat{j}+z\hat{k}. \tag{3.9}$$

在后面的表达式里面，$\hat{i}$、$\hat{j}$、$\hat{k}$ 是矢量，x，y，z 是数，$x\hat{i}$ 表示一个数乘。这个时候，理解上是有所不一样的：$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ 不再是矢量，而是 $\vec{r}$ 这个矢量在 $\left\{\hat{i},\hat{j},\hat{k}\right\}$ 这三个矢量构成的基矢下的分量。在这样一个理解下，很自然地我们就可以讨论，如果存在另外一套不一样的基矢 $\{\hat{e}_1,\hat{e}_2,\hat{e}_3\}$，那么我们自然可以通过代入 $\hat{i}$（同时也需要 $\hat{j},\hat{k}$）在新的基矢 $\{\hat{e}_1,\hat{e}_2,\hat{e}_3\}$ 下的表达式——形如 $\hat{i}=s_1^1\hat{e}_1+s_2^1\hat{e}_2+s_3^1\hat{e}_3$，从而得到矢量 $\vec{r}$ 在新的基矢下的分量形式。这样，我们自然就完成了基矢转换的过程。

我们说这样的一个对矢量的理解是非常重要的：$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ 不再是矢量，$x\hat{i}+y\hat{j}+z\hat{k}$ 才是矢量。为了节省空间，以后我们把列矢量写作行矢量的转置，例如 $[x,y,z]^T$。$[x,y,z]^T$ 仅仅是抽象矢量 $\vec{r}$ 在某一套基矢——这里是 $\left\{\hat{i},\hat{j},\hat{k}\right\}$——下的分量形式。在这里，“抽象”的含义是，我们不能写出它的一般表达式，我们也想象不出来，除非给我们一组基矢。也就是说，更一般地来说，$\hat{i}$ 以及 $\hat{j}$，$\hat{k}$ 的含义可以是位置矢量、速度矢量、力矢量，也可以是更一般的矢量。例如，函数可以看作矢量，在那里，一定条件下三角函数可以看作基矢。这就是 Fourier 级数和 Fourier 变换背后的数学结构。

为了在一般的矢量空间，而不仅仅是三维实空间矢量，做这个分量形式和抽象矢量的区分，我们引入Dirac 符号。我们把上面的表达式写成，

$$|r\rangle=x|i\rangle+y|j\rangle+z|k\rangle, \tag{3.10}$$

也就是把所有的带着矢量符号的东西都用半个括号（叫做 ket，右括号）来代替。类似的，行向量 $[x,y,z]^T$ 就可以记为

$$\langle r|=x\langle i|+y\langle j|+z\langle k|. \tag{3.11}$$

其中 $\langle r|$ 被称为左矢量（bra，左括号）。在这里，转置是一个从右矢量到左矢量的线性映射。具体 $\langle i|$ 等的含义是什么，我们一会儿会回来讨论。在这里，

我们仅仅就是换了一套记号。将来，我们会发现这套记号有额外的好处，但是目前阶段，就是使得我们的表达式写起来更复杂一点。为了让大家早点熟悉这个记号，下面所有的讨论，我们都将采用 Dirac 符号。

利用 Dirac 符号的左矢和右矢，我们可以把公式 (3.6) 中的点积改写成，

$$(u, v) = u^T v = \langle u \,|\, v \rangle . \tag{3.12}$$

同样的，当这个点积为零的时候，我们称矢量 $|u\rangle$ 和 $|v\rangle$ 相互正交。

3.2.3 矩阵的谱展开，Dirac 符号与抽象矢量

我们假设我们的读者从线性代数已经学到了本征值和本征向量的计算。一个矩阵 A 的右本征向量 v 的定义是满足下面这个表达式的非零向量，

$$Av = \lambda v \Longleftrightarrow A\,|v\rangle = \lambda\,|v\rangle . \tag{3.13}$$

这个方程也可以被写作

$$(A - \lambda)\, v = 0 \Longleftrightarrow (A - \lambda)\,|v\rangle = 0. \tag{3.14}$$

于是，这个方程有非零解的条件是

$$\det (A - \lambda I) = 0. \tag{3.15}$$

这里 I 表示对角的而且对角元素都是 1 的单位矩阵。我们假设我们的读者能够计算任意一个对称的元素为实数的 2×2 的矩阵的本征值和本征向量。

这一节我们从本征矢量开始证明一个关于矩阵本征矢量的最重要的的定理。下一节，我们会讨论矩阵和左矢量的抽象含义。再一次强调，我们的计算都很简单，但是我们寻求的对于概念的理解是不简单的。

假设我们已经得到了一个实对称 2×2 矩阵 A 的本征值 α 和本征向量 $|\alpha\rangle$，我们知道，这样的本征值 α 最多就两个，因为相应的行列式的方程是二阶的，最多有两个根。我们要证明的是当两个根不一样的时候，相应的两个本征矢量相互正交。很容易就能证明，当两个根一样的时候，这个矩阵 A 就是单位矩阵的常数倍，$A = \alpha I$[1]。对于这个特殊情况，我们只能够得到一

[1] 一种证明的方法是把二维实对称矩阵的本征值方程 $\left\| \begin{array}{cc} A_{11} - \lambda & A_{12} \\ A_{21} = A_{12} & A_{22} - \lambda \end{array} \right\| = 0$ 写下来，得到一元二次方程，然后让这个一元二次方程的判别式 $\Delta = 0$。我们就会发现，这个时候必须满足 $A_{11} = A_{22}$ 而且 $A_{12} = A_{21} = 0$。

个重根本征值，本征向量的计算也会有点小问题（任意二维矢量都是这个矩阵的本征向量）。我们把这种特殊情况叫做简并本征值。为了以后使用方便，这个时候本征向量的通常的选择方法是选择任意两个相互正交的矢量。现在，我们来证明这个重要的定理。

定理 3.1 实对称矩阵 $A^T = A$ 的不同本征值对应的本征向量相互正交。

假设 N 维实对称矩阵 $A, A^T = A$，的本征值集合为 $\left\{\alpha_j\right\}$，相应的本征向量集合为 $\left\{\left|\alpha_j\right\rangle\right\}$。对于这个集合中的任何一个元素 $\left|\alpha_j\right\rangle$，我们有 $A\left|\alpha_j\right\rangle = \alpha_j\left|\alpha_j\right\rangle$。于是，

$$\begin{aligned}\alpha_j\left\langle\alpha_i \mid \alpha_j\right\rangle &= \left\langle\alpha_i|A|\alpha_j\right\rangle = \left(\left\langle\alpha_j\left|A^T\right|\alpha_i\right\rangle\right)^T \\ = \left(\left\langle\alpha_j|A|\alpha_i\right\rangle\right)^T &= \left(\alpha_i\left\langle\alpha_j \mid \alpha_i\right\rangle\right)^T = \alpha_i\left\langle\alpha_i \mid \alpha_j\right\rangle \\ &\Longrightarrow (\alpha_i - \alpha_j)\left\langle\alpha_i \mid \alpha_j\right\rangle = 0. \end{aligned} \tag{3.16}$$

因此，要么 $\alpha_i = \alpha_j$，要么 $\left\langle\alpha_i \mid \alpha_j\right\rangle = 0$ 也就是 $|\alpha_i\rangle$ 和 $\left|\alpha_j\right\rangle$ 两个矢量正交。

其实，我们还可以证明实对称矩阵的本征值必然是实数，不会是复数。不过，这个留到我们介绍复数域上的矩阵的时候再来证明。有了本征矢量的正交性，对于我们的下一步讨论就够用了。

在以上的证明过程中，我们用到了转置的性质，$\left(\left\langle\alpha_i|A|\alpha_j\right\rangle\right)^T = \left\langle\alpha_j\left|A^T\right|\alpha_i\right\rangle$。严格说来，这一条是需要我们先证明的。我们能够这么做是因为在通常的线性代数中 $(u^T A v)^T = v^T A^T u$。实际上，这个逻辑是不完全正确的。既然我们希望建立抽象矢量而不是分量形式的矢量的概念，那么，分量形式的矢量和矩阵的公式就不应该在这里使用。然而，如果我们真的要从头开始建立抽象矢量的概念，我们需要花很大的精力，从矢量和矢量空间的抽象定义开始。这当然也是可行的，但是一般来说会等到更高级的量子力学和代数的课程里面。在这里，我们的目的和思路比较简单，我们希望读者学习完了这个部分之后从概念上懂得抽象矢量和分量形式的区别，但是也要看见以前所学的分量形式的计算都是对的，而且明白为什么是对的。所以，有的时候我们通过分量形式的公式来猜测抽象形式的公式的形式是没有问题的。如果真的希望采用严格的证明，我们需要先定义矢量（右矢量）空间和对偶矢量（左矢量）空间，以及把矢量空间上的算符对应到对偶矢量空间上的算符 [28, 31]。

对于任意一个非零矢量 $|u\rangle$，我们总是可以做一个归一化，$|\tilde{u}\rangle = \frac{|u\rangle}{\sqrt{\langle u|u\rangle}}$，于是 $\langle\tilde{u}\,|\,\tilde{u}\rangle = 1$。从现在开始，我们把本征矢量就看做既相互正交又自己归

一的一个矢量集合。有了这个矢量集合，我们来探讨把这个集合作为基矢，我们如何来表达其他的矢量。对于本征值出现多重根，也就是简并的情况，我们假设已经通过某种方式选择了两个相互正交的矢量来作为基矢。所以，以后我们不再区分是否出现多重根。也就是说，通过求解某个矩阵本征值和本征向量的方法，我们已经得到了一组 N 个正交归一的矢量，N 就是矩阵所在的矢量空间的维数。这样的一组矢量的集合也被称为完备的正交归一基矢，其中的矢量的两两之间的内积满足

$$\langle\beta\,|\,\alpha\rangle = \delta_{\alpha,\beta}. \tag{3.17}$$

其中的 $\delta_{\alpha,\beta}$ 是Kronecker δ 记号，

$$\delta_{\alpha,\beta} = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{otherwise} \end{cases}. \tag{3.18}$$

我们从这里开始。

定理 3.2 *给定一组完备的正交归一基矢的 $\{|\alpha\rangle\}$，矢量空间中的任意矢量 $|u\rangle$ 可以按照如下方式展开，*

$$|u\rangle = \sum_\alpha u_\alpha\,|\alpha\rangle, \tag{3.19}$$

而且，

$$u_\alpha = \langle\alpha|\,u\rangle. \tag{3.20}$$

证明非常简单，第一部分，$|u\rangle$ 是这个 N 维空间中的一个矢量，$\{|\alpha\rangle\}$ 是相互正交的，也就是线性无关的 N 个矢量，于是必然这个 $N+1$ 个矢量不可能线性无关。所以有了第一个表达式。然后，我们把这个表达式的两边与矢量 $|\beta\rangle$ 做内积，于是

$$\langle\beta\,|\,u\rangle = \langle\beta|\sum_\alpha u_\alpha\,|\alpha\rangle = \sum_\alpha u_\alpha\,\langle\beta\,|\,\alpha\rangle = \sum_\alpha u_\alpha\delta_{\alpha,\beta} = u_\beta. \tag{3.21}$$

有了这个定理，我们来看一下，左矢量的作用。左矢量 $\langle\alpha|$ 的第一个作用是把一个矢量 $|u\rangle$ 对应的在 $|\alpha\rangle$ 上的分量 u_α 取出来。它把一个矢量映射成为一个数。特别地，当我们的 $|u\rangle = |\alpha\rangle$ 的时候，得到的数是 1。从这个意义上说，左矢量 $\langle\alpha|$ 就好像是右矢量 $|\alpha\rangle$ 的一个镜像。左矢量的这两条性质：从矢量空间到数的映射，右矢量的镜像，是非常重要的性质。以后，我们经常会利用这两条性质来理解其它的数学对象。

现在，我们就用这个理解来看看矩阵到底是什么。考虑一个有两个正交归一基矢这样构成的对象 $M_{\alpha\beta} = |\alpha\rangle\langle\beta|$。我们来看看 $M_{\alpha\beta}$ 作用到任意一个矢量上得到什么。

$$M_{\alpha\beta}|u\rangle = |\alpha\rangle\langle\beta|\sum_j u_j|j\rangle = u_\beta|\alpha\rangle. \tag{3.22}$$

这个结果的含义是 $M_{\alpha\beta}$ 把一个矢量 $|u\rangle$ 的 β 分量 u_β 单独拿出来，然后把这个分量转动了一个方向作为 α 方向上的分量。也就是说，$M_{\alpha\beta}$ 是一个把右矢量变成另一个右矢量的映射，而且这个映射看起来像是转动，或者更严格一点说，转动的一部分。这样的一个把右矢量变成另一个右矢量的映射，叫做算符。

接着，我们来看一下这样一个算符是什么：$\hat{A} = \sum_{\alpha,\beta} A_{\alpha\beta}|\alpha\rangle\langle\beta|$。其中 $A_{\alpha\beta}$ 是实数。这个对象就是相当于一堆 $|\alpha\rangle\langle\beta|$ 的线性组合，也就是把好多个 $|\alpha\rangle\langle\beta|$ 加起来。自然，它还是一个把右矢量变成另一个右矢量的映射，而且是“转动”。

$$\hat{A}|u\rangle = \sum_{\alpha,\beta} A_{\alpha\beta}|\alpha\rangle\langle\beta|\sum_j u_j|j\rangle = \sum_{\alpha,\beta} A_{\alpha\beta}u_\beta|\alpha\rangle, \tag{3.23}$$

也就是说，

$$\left(\hat{A}u\right)_\alpha = \sum_\beta A_{\alpha\beta}u_\beta. \tag{3.24}$$

这个正好就是矩阵与矢量的乘法的定义，如果记 $v = Au$，则 $v_\alpha = \sum_\beta A_{\alpha\beta}u_\beta$。从这里，我们看到，所谓矩阵就是这个算符 $\hat{A}$ 的在完备正交归一基矢集合 $\{|\alpha\rangle\}$ 下展开的系数。于是，我们把 $|\alpha\rangle\langle\beta|$ 看作算符，然后把 $A_{\alpha\beta}$ 看作这些算符相互叠加起来的系数，记作

$$\hat{A} = \sum_{\alpha,\beta} A_{\alpha\beta}|\alpha\rangle\langle\beta|. \tag{3.25}$$

就好像我们不再把 $[x, y, z]^T$ 看作矢量，而是看做矢量 $|r\rangle$ 在某套基矢下的分量一样，我们不再把矩阵 $(A_{ij})_{N\times N}$ 看作一个矩阵算符，而是看作算符 $\hat{A}$ 在某套基矢下的分量形式。特别地，我们来看一看单位矩阵，或者说不变算符 $\hat{I}$ 的展开形式，

$$\left(\hat{I}u\right)_\alpha = u_\alpha \Longrightarrow I_{\alpha\beta} = \delta_{\alpha\beta}, \tag{3.26}$$

于是

$$\hat{I} = \sum_{\alpha,\beta} \delta_{\alpha\beta} |\alpha\rangle \langle\beta| = \sum_{\alpha} |\alpha\rangle \langle\alpha| . \tag{3.27}$$

最后这个表达式非常重要，有一个专门的名字，叫做完全性关系。

把抽象的矢量和矩阵算符与排成一列的数字和排成一个方块的数字之间的区别开来，是我们这一节最重要的目标。我们是通过运用 Dirac 符号和左矢量来达到这个目的的。没有 Dirac 符号和左矢量的帮助，我们很难说清楚两者之间的区别和联系。同时，我们学习了谱展开的一些关键概念：对称矩阵的不同本征值对应的本征向量相互正交，通过对于简并本征值的特殊处理之后可以构成完备的正交归一基矢。利用这组基矢以及这组基矢的完全性关系，我们可以得到各个抽象矢量和抽象算符在这组基矢下的分量形式。在得到分量形式之后，以前在线性代数课程中学到过的所有的运算都是正确的有效的。也就是说，在这个意义——线性代数的矢量和矩阵是抽象矢量和算符在某一套给定的基矢下的分量形式——上，认为列矢量是矢量，一个方块的数字是矩阵，没有问题。

没有抽象矢量记号之前，矢量的定义是从坐标变换下的分量的变换形式定义的，说坐标变换下分量不变叫做标量，分量如何变化的叫作矢量，如何如何变化的叫做张量。这当然，也是正确的定义。但是，就好像你识别出一个朋友的脸，你肯定不是通过他/她的脸如何变化来识别的，而是通过某些内在特征来识别的。有了这个记号，我们终于可以说：矢量在坐标变换下是不变的，标量和张量在坐标变换下也是不变的，改变的仅仅是分量形式。这个抽象矢量记号对于我们后面的内容非常重要，所有的矢量将来都会写成这个抽象形式。只有这样，我们认为，才能够形成足够深刻的认识：**矢量不是数**，数仅仅是矢量在某坐标系下的分量形式。

因此，这个部分的学习目的不是学到新的计算，而是学到新的理解。这个不是很容易。请试着不依赖于任何一套基矢给我一个矢量。如果你觉得左矢量不太好直观想象举例子，那么，右矢量其实同样不太好直观想象举例子，矩阵也是一样的。

3.3 复数域上的抽象矢量与抽象矩阵算符

前面，我们讨论了实数域上的矢量和矩阵的加法、乘法、数乘、点积、转置、本征值和本征向量、抽象形式与分量形式的关系，现在我们来平行地讨论复数域上的矢量和矩阵的加法、乘法、数乘、点积、转置、本征值和本

征向量、抽象形式与分量形式的关系。同样地，我们先学会计算，这个简单的任务，然后来讨论对于这些计算的理解。

3.3.1 复数域上的矩阵和矢量运算

加法、乘法、数乘的定义完全一样。我们仅给出不一样的点积的定义。这个时候通常称为内积。两个复数域上的列矢量，这个时候是分量形式的，u 和 v 的内积定义为

$$(v,\mu) = v^{\dagger}\mu = \sum_j v_j^* \mu_j, \tag{3.28}$$

其中 $v^{\dagger}$ 称为 v 的 Hermitian 共轭。其一般的定义是，

$$\left(A^{\dagger}\right)_{ij} = A_{ji}^*. \tag{3.29}$$

这个操作与实数域上的转置操作略有区别，可以写做，

$$A^{\dagger} = \left(A^*\right)^T, \tag{3.30}$$

相当于先做元素的复共轭再做矩阵转置。这些都是分量形式的定义。我们只能够在给定基矢集合以后做这样的运算。下面，我们讨论抽象的定义。当然，我们的做法还是建立分量形式和抽象形式之间的联系，而不是真正的直接讲授抽象矢量。后者，对于学习更复杂的对象是有必要的，但是我们仅仅关系二能级体系上的量子力学，懂得抽象矢量的存在及其与分量形式的联系就够了。

3.3.2 复数域上的左矢量的分量形式的定义

在实数域上，左矢量的分量形式很好定义，就是把右矢量的分量形式——那个列向量，横过来变成行向量就行了。现在，复数域上的矢量内积的定义比实数域上的定义多了一个共轭操作，于是右矢量和左矢量的关系稍微的复杂了一点点，但是左矢量是从右矢量到数的映射，左矢量是右矢量的一个镜像，这个是不变的。我们希望保持下面这个表达式，

$$(v,\mu) \triangleq \langle v|\,\mu\rangle \tag{3.31}$$

满足内积的性质，例如可以用内积来定义距离，于是[2]

$$\langle v|\,v\rangle \geq 0. \tag{3.32}$$

[2]内积其实还有其他要求，见例如 [28]。

我们还希望保持

$$\hat{I} = \sum_{\alpha} |\alpha\rangle \langle\alpha| . \tag{3.33}$$

其中 $\{|\alpha\rangle\}$ 是某一套完备正交归一的基矢。如果这样的一套基矢还能够与某个算符的本征矢量相互联系起来，就更好了。这些就是我们的目标，然后我们来看看，我们的左矢量应该如何定义。

我们定义其分量形式为，把同样基矢下的右矢量所对应的分量形式的列向量，横过来，然后给每一个元素做一下复共轭，也就是如果

$$|\nu\rangle \Rightarrow [\nu_1, \nu_2, \cdots]^T , \tag{3.34}$$

则

$$\langle\nu| \Rightarrow [\nu_1^*, \nu_2^*, \cdots] . \tag{3.35}$$

这个定义有一个后果，

$$|a\nu\rangle = a\,|\nu\rangle , \tag{3.36}$$

则

$$\langle a\nu| = a^* \langle\nu| . \tag{3.37}$$

这里，a 是一个复数。这一点和实数域上的情况略有不同，那里 $\langle a\nu| = a\langle\nu|$，因为 $a^* = a$。也就是说，在实数域上左矢量和右矢量形式上完全相同就是一个是行矢量一个是列矢量，但是在复数域上，其数字的值也不一样。为什么会这样定义呢?

我们试着用这个行向量(左矢量)$[\nu_1^*, \nu_2^*, \cdots]$ 跟列矢量(右矢量)$[\nu_1, \nu_2, \cdots]^T$ 做一个相乘，我们就发现，刚好我们得到了公式 (3.32) 中定义的结果 ($\nu_1^*\nu_1 \geq 0$，但是 $\nu_1\nu_1$ 没有这个性质)，也就是说，公式 (3.31) 得到了满足。所以，复数域上的内积的定义，以及保持分量形式的行列矢量相乘的运算与内积一致的要求，使得我们选择了这个左矢量的定义。

这个行矢量和列矢量之间的关系，就好像实数域上的转置，在复数域上被称为复共轭，或者 Hermitian 共轭，记为

$$(|\nu\rangle)^\dagger = \langle\nu| . \tag{3.38}$$

正如转置可以定义在矩阵上，Hermitian 共轭也可以定义到矩阵上，分量形式的定义就是公式 (3.29)。

3.3.3 复数域上抽象算符的谱展开

现在，我们回过头来讨论复数域上的抽象左右矢量和算符。让我们假设抽象矢量的定义和它的 Hermitian 共轭的定义是明确的。实际上，我们是通过分量形式的定义公式 (3.34) 和公式 (3.35) 来理解抽象左右矢量的。我们说过，这个是不完整的，有必要的话，可以通过直接定义抽象的矢量来弥补。例如，矢量之所以成为矢量是可以对矢量做加法、数乘、旋转、拉伸。不过，我们的处理方式是让这些抽象定义在更高层次的量子力学和代数中学习。所以，在这里我们接受通过分量形式辗转定义的左矢量，而且我们不回答右矢量是什么的问题。反正给定一个右矢量，我们在某一套完备正交归一的基矢下面写下分量形式，然后写下这个分量形式的 Hermitian 共轭，这个共轭的结果就是这一套基矢下面这个右矢量对应的左矢量。换句话说，给定一套基矢 $\left\{\left|\alpha_j\right\rangle\right\}$，先别管这套基矢哪里来的，如果我们的右矢量是

$$|\mu\rangle = \sum_j \mu_j \left|\alpha_j\right\rangle, \tag{3.39}$$

那么我们的相应的左矢量就是

$$\langle\mu| = \sum_j \mu_j^* \left\langle\alpha_j\right|. \tag{3.40}$$

然后，我们来讨论形如 $|\alpha\rangle\langle\beta|$ 的对象是什么。很容易验证，就像在实数域上的情况一样，$|\alpha\rangle\langle\beta|$ 把一个右矢量变成另一个右矢量，所以是一个算符。而且，是一个线性算符。一个一般的线性算符 L 的含义是，当作用在对象 $a\mu + b\nu$ 上，我们有

$$L(a\mu + b\nu) = aL(\mu) + bL(\nu). \tag{3.41}$$

其中 a, b 是数，μ, ν 是对象。可以验证 $|\alpha\rangle\langle\beta|$ 满足这个线性性的要求。于是，一个一般的作用在所有右矢上的线性算符必然是

$$\hat{A} = \sum_{\alpha\beta} A_{\alpha\beta} |\alpha\rangle\langle\beta|. \tag{3.42}$$

这个结论与实数域上的矢量的相应结论一模一样。现在，我们来看这个算符的 Hermitian 共轭的问题。我们把算符 $\hat{A}$ 的每一个部分都做一个 Hermitian 共轭，

$$\hat{A}^\dagger = \sum_{\alpha\beta} A_{\alpha\beta}^* |\beta\rangle\langle\alpha| = \sum_{\alpha\beta} A_{\beta\alpha}^* |\alpha\rangle\langle\beta|. \tag{3.43}$$

其中，我们用到了左（右）矢量的 Hermtian 共轭是右（左）矢量。于是，

$$\hat{A}^{\dagger} = \hat{A} \Longleftrightarrow A^{*}_{\alpha\beta} = A_{\beta\alpha}. \tag{3.44}$$

我们把满足这个性质的算符称为 Hermitian 算符。

有了 Hermitian 算符，我们来考察这样的算符的本征值和本征向量。这个定理是本小节最重要的结果。

定理 3.3 Hermitian 算符本征向量正交性：Hermitian 算符 $\hat{A}$ 的所有的本征值（满足 $A\left|\alpha_j\right\rangle = \alpha_j\left|\alpha_j\right\rangle$）为实数 $\alpha_j \in \mathcal{R}$，且不同本征值对应的本征向量相互正交：$\left\langle \alpha_i \mid \alpha_j \right\rangle = 0$，其中 $\alpha_i \neq \alpha_j$。

大致证明如下：先证明本征值为实数，

$$\begin{aligned}
\alpha_j \left\langle \alpha_j \middle| \alpha_j \right\rangle = \left\langle \alpha_j \middle| A \middle| \alpha_j \right\rangle &= \left(\left\langle \alpha_j \middle| A \middle| \alpha_j \right\rangle^{*}\right)^{*} = \left\langle \alpha_j \middle| A^{\dagger} \middle| \alpha_j \right\rangle^{*} \\
&= \left\langle \alpha_j \middle| A \middle| \alpha_j \right\rangle^{*} = \alpha_j^{*} \left\langle \alpha_j \middle| \alpha_j \right\rangle \\
&\Longrightarrow \left(\alpha_j - \alpha_j^{*}\right)\left\langle \alpha_j \mid \alpha_j \right\rangle = 0 \\
&\Longrightarrow \alpha_j = \alpha_j^{*}.
\end{aligned} \tag{3.45}$$

假设 $\alpha_i \neq \alpha_j$，

$$\begin{aligned}
\alpha_j \left\langle \alpha_i \mid \alpha_j \right\rangle = \left\langle \alpha_i \middle| A \middle| \alpha_j \right\rangle &= \left(\left\langle \alpha_i \middle| A \middle| \alpha_j \right\rangle^{*}\right)^{*} = \left\langle \alpha_j \middle| A^{\dagger} \middle| \alpha_i \right\rangle^{*} \\
&= \left\langle \alpha_j \middle| A \middle| \alpha_i \right\rangle^{*} = \alpha_i^{*} \left\langle \alpha_j \mid \alpha_i \right\rangle^{*} = \alpha_i \left\langle \alpha_i \mid \alpha_j \right\rangle \\
&\Longrightarrow (\alpha_i - \alpha_j)\left\langle \alpha_i \mid \alpha_j \right\rangle = 0 \\
&\Longrightarrow \left\langle \alpha_i \mid \alpha_j \right\rangle = 0. \blacksquare
\end{aligned} \tag{3.46}$$

与实数域上的情形类似，如果算符 $\hat{A}$ 的本征值都不相同，那么只要我们把每一个本征矢量各自归一化，我们就从 Hermitian 算符 $\hat{A}$ 的本征矢量得到了空间 V 的正交归一基矢，于是我们有完全性展开关系，

$$I = \sum_j \left|\alpha_j\right\rangle\left\langle\alpha_j\right|. \tag{3.47}$$

存在相同本征值对应不同本征向量，这个称为简并本征值，的问题技术上更复杂（需要在相同本征值对应的本征向量的子空间内正交化矢量），结论是一样的，在此我们不讨论。

在以上的证明过程中，我们也用了分量形式的 Hermitian 共轭的关系 $\left\langle \alpha_i \middle| A \middle| \alpha_j \right\rangle^{*} = \left\langle \alpha_j \middle| A^{\dagger} \middle| \alpha_i \right\rangle$。这个表达式的用法也和实数域上的情形一致：尽管

原则上我们可以通过抽象矩阵和抽象矢量来代替这个基于分量形式的表达式，但是，在这里我们就直接用了分量形式的结论。

总结，复数域上的抽象矢量和抽象矩阵，与实数域上的基本一致，除了在左矢量的定义上，我们需要把右矢量做一个转置以后在给每一个元素做一个复共轭，而不仅仅是转置。其他方面，完全是对应的。

另外，完全性关系的价值可能我们还没有看到。这一小节的最后，我们展示一下。可以想见 $A_{\alpha\beta}$ 的意义就是抽象算符 $\hat{A}$ 在 $\{|\alpha\rangle\}$ 这一套基矢下的分量形式，但是，由于我们的算符的定义直接就是由 $A_{\alpha\beta}$ 通过公式 (3.42) 给出的，这个倒过来如何从 $\hat{A}$ 得到 $A_{\alpha\beta}$ 的问题不容易看见，现在，我们说，实际上，

$$A_{\alpha\beta} = \langle\alpha|A|\beta\rangle, \tag{3.48}$$

怎么看出来？让我们做以下计算，

$$\begin{aligned} A = IAI &= \sum_{\alpha} |\alpha\rangle\langle\alpha| A \sum_{\beta} |\beta\rangle\langle\beta| \\ &= \sum_{\alpha\beta} (\langle\alpha| A |\beta\rangle) |\alpha\rangle\langle\beta|, \end{aligned} \tag{3.49}$$

于是 $A_{\alpha\beta}$ 自然就是 $\langle\alpha|A|\beta\rangle$。

3.3.4 抽象算符的函数与算符的代数

有了抽象算符的谱展开，我们就可以通过数的函数来定义抽象算符的函数。例如我们知道算符 A，就可以知道算符 A^2，

$$A^2 = AA \sum_{\alpha} |\alpha\rangle\langle\alpha| = A \sum_{\alpha} \alpha |\alpha\rangle\langle\alpha| = \sum_{\alpha} \alpha^2 |\alpha\rangle\langle\alpha|. \tag{3.50}$$

于是，$A^2 = \sum_{\alpha} \alpha^2 |\alpha\rangle\langle\alpha|$。更一般地，$A^n = \sum_{\alpha} \alpha^n |\alpha\rangle\langle\alpha|$。那么一般的算符 A 的多项式 $f(A) = \sum_{\alpha} f(\alpha) |\alpha\rangle\langle\alpha|$，例如

$$e^{\lambda A} = \sum_{\alpha} e^{\lambda\alpha} |\alpha\rangle\langle\alpha|. \tag{3.51}$$

另外，算符的加减乘除等运算跟通常的数的运算有所不同。例如对于数 $e^{a+b} = e^a e^b$ 但是，对于算符通常 $e^{A+B} \neq e^A e^B$。其根本原因是一般来说 $AB \neq BA$。同样的，一般来说 $e^A B e^{-A} \neq B$。对这些有兴趣的同学可以进一步通过学习高等量子力学的书籍 [28] 来进一步了解。这里我们给出一个很有意思的关于 $e^A B e^{-A}$ 的讨论。

例 3.1 (Baker–Campbell–Hausdorff **公式的一个变体**) ： 讨论算符 $e^A B e^{-A} \neq B$ 的显式形式，也就是去掉指数算符以后的形式。

定义算符

$$f(A,B,\lambda) = e^{\lambda A} B e^{-\lambda A}. \tag{3.52}$$

已知 $f(A,B,\lambda=0)=B$。现在，我们来计算算符 $f(A,B,\lambda)$ 的几个关于 λ 的导数。思路是这样的：反正 $f(A,B,\lambda)=B$ 是 λ 的函数，于是无论如何，

$$f(A,B,\lambda) = \sum_{n=0}^{\infty} C_n \lambda^n. \tag{3.53}$$

当然，我们还必须证明为什么不考虑 λ^{-n} 之类的项。这里暂时不管这个问题。如果上式成立，那么，我们就可以想办法通过计算 $\left.\frac{\partial^n f(A,B,\lambda)}{\partial \lambda^n}\right|_{\lambda=0}$ 来得到所有的 C_n。现在，我们来试试前面几阶。

$$\frac{\partial}{\partial \lambda} f(A,B,\lambda) = A e^{\lambda A} B e^{-\lambda A} - e^{\lambda A} B e^{-\lambda A} A = e^{\lambda A}[A,B] e^{-\lambda A} = f(A,[A,B],\lambda). \tag{3.54}$$

于是，

$$C_0 = f(A,B,\lambda=0) = B. \tag{3.55}$$

接着，

$$C_1 = \left.\frac{\partial f(A,B,\lambda)}{\partial \lambda}\right|_{\lambda=0} = f(A,[A,B],\lambda=0) = [A,B]. \tag{3.56}$$

类似地

$$C_2 = \frac{1}{2!}\left.\frac{\partial^2 f(A,B,\lambda)}{\partial^2 \lambda}\right|_{\lambda=0} = \frac{1}{2!} f(A,[A,[A,B]],\lambda=0) = \frac{1}{2!}[A,[A,B]]. \tag{3.57}$$

同理我们可以得到各阶的系数。于是计算 $f(A,B,\lambda=1)$ 我们得到

$$e^A B e^{-A} = B + \frac{1}{1!}[A,B] + \frac{1}{2!}[A,[A,B]] + \frac{1}{3!}[A,[A,[A,B]]] + \cdots. \tag{3.58}$$

通过引入 λ 再让 $\lambda=0$ 以及 $\lambda=1$ 来求得算符是一个很巧妙的技巧。通过这个例子，我们也看见了算符和数之间的区别造成的后果。

3.4 共同本征向量, 表象理论与线性变换

如果我们有两个不同的 Hermitian 算符 $\hat{A},\hat{B}$，我们问这个时候它们对应的正交归一基矢是否相同?

定理 3.4 Hermitian 算符 $\hat{A},\hat{B}$ 存在共同本征向量如果 $\hat{A},\hat{B}$ 对易：$\left[\hat{A},\hat{B}\right]=\hat{A}\hat{B}-\hat{B}\hat{A}=0$。

同样我们假设本征值没有简并的条件下来讨论。取 $\hat{A}$ 的本征向量 $|\alpha\rangle$，

$$\hat{A}\left(\hat{B}|\alpha\rangle\right)=\hat{A}\hat{B}|\alpha\rangle=\hat{B}\hat{A}|\alpha\rangle=\alpha\left(\hat{B}|\alpha\rangle\right)\Longrightarrow\hat{B}|\alpha\rangle=b_\alpha|\alpha\rangle. \tag{3.59}$$

于是 $|\alpha\rangle$ 也是算符 $\hat{B}$ 的本征向量。证毕 ■

反过来，很简单就是可说明，如果两个算符 $\hat{A},\hat{B}$ 的所有的本征向量都完全一样，则这两个算符必然对易。直观地说，在这个共同本征态所形成的基矢下，这两个算符都只有对角矩阵元，例如

$$\hat{A}=\sum_\alpha\hat{A}|\alpha\rangle\langle\alpha|=\sum_\alpha\alpha|\alpha\rangle\langle\alpha|, \tag{3.60}$$

完全对角，而对角的矩阵永远是相互对易的。

通过这个定理，我们看见，如果我们需要讨论的算符不是所有的都相互对易，那么我们就可以选择不同的 Hermitian 算符所代表的本征矢量来作为基矢。这样，我们就有很多种基矢的选择。于是，一个算符或者一个向量，它们在不同的基矢下就有不同的表现形式。这些不同的表现形式之间有什么样的关系呢?

我们称某一个相容算符的集合 X——这个集合可以包含多个算符，这些算符两两对易——所对应的共同本征向量为一个 X 表象。这些共同本征向量可以作为表达其他所有矢量和矩阵的基矢。一个矢量和矩阵在这一套基矢下的分量形式就称为这个矢量或者矩阵的 X 表象下的形式。现在，一个自然的问题就是，如果我们有两个相容算符的集合 X 和 Y，但是 X 和 Y 之间不相容，那么这个矢量或者矩阵在两个表象下的形式有什么区别和联系。这个问题就是矢量和矩阵的线性变换的问题。这就好像是一个人不管哪一天出现，总是会有一个外在的形象，尽管这个外在的形象可能每次都不一样，那么我们自然会关心这些外在的形象之间有什么关系。表象就是把一个抽象的矢量或者矩阵呈现在某一套基矢下面，就是把抽象的人的具体的外在形象呈现出来。

3.4.1 表象之间的变换

为了简单性，我们这里假设 X（X'）表象只有一个算符 $\hat{X}$（$\hat{X}'$），记本征矢量为 $|x\rangle$（$|x'\rangle$）。现在我们来考察这两个表象之间的联系。一个矢量 $|\mu\rangle$

可以在表象 X 中也可以在表象 X' 中表达，分别是

$$|\mu\rangle = \sum_x \mu_x |x\rangle \tag{3.61}$$

和

$$|\mu\rangle = \sum_{x'} \mu'_{x'} |x'\rangle . \tag{3.62}$$

于是两者之间可以通过以下方式相联系，

$$\mu'_{x'} = \langle x' | \mu\rangle = \sum_x \mu_x \langle x' | x\rangle . \tag{3.63}$$

记 $S_{x'x} = \langle x' | x\rangle$，于是

$$\mu'_{x'} = \sum_x S_{x'x}\mu_x. \tag{3.64}$$

在这里，$S_{xx'} = \langle x | x'\rangle$ 的集合其实仅仅是一个集合，S 不是一个把一个矢量映射成为另一个矢量的矩阵或者算符。强行按照矩阵运算的规则，上式可以记作，

$$\mu' = S\mu. \tag{3.65}$$

这样，看起来好像两个表象之间的分量存在一个相似变换。当然，我们还需要证明这个形式上的矩阵 S 满足 $S^\dagger S = I$。这个证明非常简单，留给大家做作业。

我们再来讨论一下矩阵的不同表象的联系。对于一个算符 $\hat{A}$，我们有

$$\hat{A} = \sum_{xy} |x\rangle \langle x| \hat{A} |y\rangle \langle y| = \sum_{x'y'} |x'\rangle \langle x'| \hat{A} |y'\rangle \langle y'| , \tag{3.66}$$

于是

$$A_{x'y'} = \sum_{xy} \langle x' | x\rangle \langle x| \hat{A} |y\rangle \langle y| y'\rangle \Longrightarrow A' = SAS^\dagger. \tag{3.67}$$

我们发现，联系着两个表象之间的是一个相性变换矩阵

$$S = (\langle x' | x\rangle)_{N\times N} . \tag{3.68}$$

到这里，S 还是一个矩阵，不是一个算符，也就是排成一个方块的遵循矩阵运算规则的数字，而不是一个把矢量映射成为矢量的算符。

3.4.2 举例

下面我们举例来熟悉表象和线性变换，而且要把表象变换和你所熟悉的线性代数中的线性变换联系起来。

例 3.2 (**二维 $\sigma_x, \sigma_y, \sigma_z$ 在不同表象下的形式和抽象形式**)：通常这三个算符分别写作 $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$。请问这个是在什么表象下写出来的。计算三个算符在 σ_x 表象下的形式，讨论这个表象下的形式与通常表象下的形式之间的变换。

我们发现 σ_z 矩阵在这个表象下的本征矢量就是 $|\uparrow_z\rangle = [1,0]^T$，$|\downarrow_z\rangle = [0,1]^T$。正好就是这个空间的基矢。所以，这套算符是在 σ_z 表象下写出来的。现在，我们来转到 σ_x 表象。我们需要先求出来 σ_x 的本征向量，有 $|\uparrow_x\rangle = \frac{\sqrt{2}}{2}[1,1]^T$，$|\downarrow_x\rangle = \frac{\sqrt{2}}{2}[1,-1]^T$。下一步我们需要把 σ_z 表象的基矢转化成 σ_x 表象的基矢，

$$|\uparrow_z\rangle = \frac{\sqrt{2}}{2}\left(|\uparrow_x\rangle + |\downarrow_x\rangle\right),$$
$$|\downarrow_z\rangle = \frac{\sqrt{2}}{2}\left(|\uparrow_x\rangle - |\downarrow_x\rangle\right).$$

于是，

$$\sigma_y = -i|\uparrow_z\rangle\langle\downarrow_z| + i|\downarrow_z\rangle\langle\uparrow_z| = i|\uparrow_x\rangle\langle\downarrow_x| - i|\downarrow_x\rangle\langle\uparrow_x|.$$

如果写成矩阵就是，

$$\sigma_y = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}.$$

我们还可以写下来表象变换的矩阵形式。

通过这个例子，我们看到同样的算符在不同的表象下可以写成不同的矩阵的形式。

例 3.3 (**一元偏微分方程的二维子空间**)：求解微分方程 $\frac{\partial^2}{\partial t^2}\psi(x,t) - \nabla^2\psi(x,t) = 0$，边界条件 $\psi(0,t) = 0 = \psi(1,t)$，初始条件 $\psi(x,0) = 0.3\sin(\pi x) + 0.5\sin(2\pi x)$。

首先我们注意到这是一个线性方程，于是任意两个解的线性叠加，仍然是这个方程的解。因此，这个方程的解构成一个线性空间。至于这个空间的维数和基矢我们需要通过下面的计算来搞清楚。对方程做分离变量，假设

$\psi(x,t) = X(x)T(t)$（也可以做 Fourier 变换来求解），于是

$$\frac{X(x)}{X''(x)} = \frac{T''(t)}{T(t)}.$$

左边是 x 的函数，右边是 t 的函数，两边相等的唯一可能就是不是依赖于 (x,t) 的常数。于是方程变成如下两个，

$$X''(x) = \lambda X(x),$$
$$T''(t) = \lambda T(t).$$

这两个方程的解都是 e 指数的形式(e^{kx} 或者 e^{ikx},或者等价的 $\sin(kx), \cos(kx)$ 或者 $\sinh(kx), \cosh(kx)$)。考虑到边界条件，我们采用 $\sin(kx)$ 的形式，并且 $k = n\pi$。在时间侧的方程上，习惯上用 $e^{-i\omega t}$。于是

$$\psi(x,t) = \sum_n a_n \sin(n\pi x) e^{-in\pi t}.$$

根据初始条件，我们得到

$$\psi(x,t) = 0.3\sin(\pi x)e^{-i\pi t} + 0.5\sin(2\pi x)e^{-i2\pi t}.$$

例 3.4 (Fourier **级数的矢量空间解释**)：结合上面的例子，思考一个函数的 Fourier 级数的形式到底做了什么？进一步思考，通常的函数形式是在写在什么表象下的？那 Fourier 变换呢？

在上面的计算中，我们先找到满足边界条件的一组解，$\sin(n\pi x)e^{-in\pi t}$，然后把一般解写成这组解的组合的形式，最后通过和初始条件对比来确定组合的系数。因此，实际上，这里先做了一个基矢分解，然后留下系数待确定。更一般地来说，Fourier 级数就是把正弦和余弦函数当作函数空间的基矢量，Fourier 变换就是把 $e^{\pm ikx}$ 当作函数空间的基矢量，来求得函数在这个基矢空间的表象形式。那么，当我们写下来一个一般的 $f(x)$ 的时候，把什么当作基矢呢？提示一下，试试这样来看一个函数，

$$\hat{f} = \int_{-\infty}^{\infty} dx f(x) |x\rangle \langle x|.$$

3.4.3 相似变换作为算符 *

本节为选读部分。

现在，我们以 S 矩阵为基础来定义一个把矢量映射成为矢量的算符。为了和上一节区别开来，我们用新的符号来表示基矢：X 表象的记作 $\left|e_j\right\rangle$，X'

表象的记作 $\left|e'_j\right\rangle$。现在，我们来定义一个新的矢量，

$$|\mu'\rangle = \sum_j \left|e_j\right\rangle \left\langle e'_j\right| \cdot |\mu\rangle . \tag{3.69}$$

注意，这个我们已经默认建立了一个在 $\left|e_j\right\rangle$ 和 $\left|e'_j\right\rangle$ 之间的顺序对应关系：把 X 表象的第 j 个基矢与 X' 表象的第 j 个基矢相互联系起来。这个时候 $\sum_j \left|e_j\right\rangle \left\langle e'_j\right|$ 就真的是一个算符了。现在，我们来看一下这个算符与前面的 S 矩阵联系。

我们来看一下，这个新的矢量 $|\mu'\rangle$ 在 X 表象下的分量形式，

$$\langle e_k| \cdot |\mu'\rangle = \sum_j \delta_{jk} \left\langle e'_j\right| \cdot |\mu\rangle = \left\langle e'_k\right| \cdot |\mu\rangle . \tag{3.70}$$

新矢量在旧表象下的分量形式等于旧矢量在新表象下的分量形式。进一步，右侧等于

$$\langle e_k \, |\mu'\rangle = \langle e'_k \, |\mu\rangle = \sum_l \langle e'_k \, |e_l\rangle \langle e_l| \, \mu\rangle = \sum_l S_{kl} \langle e_l| \, \mu\rangle . \tag{3.71}$$

于是

$$|\mu'\rangle = \sum_{kl} S_{kl} \, |e_k\rangle \langle e_l| \, \mu\rangle . \tag{3.72}$$

定义

$$\hat{S} = \sum_{kl} S_{kl} \, |e_k\rangle \langle e_l| , \tag{3.73}$$

我们得到

$$|\mu'\rangle = \hat{S} \, |\mu\rangle . \tag{3.74}$$

这个时候，$\hat{S}$ 真正的成了一个算符。进一步，我们可以把 $\hat{S}$ 写作，

$$\hat{S} = \sum_{kl} \left\langle e'_k\right| e_l\rangle \, |e_k\rangle \langle e_l| = \sum_k |e_k\rangle \left\langle e'_k\right| \tag{3.75}$$

正好就是我们前面定义的把旧矢量 $|\mu\rangle$ 变成新矢量 $|\mu'\rangle$ 的映射。于是，我们发现，相似变换也可以看作是一个算符。只不过，再看做这个算符的过程中，我们需要把旧矢量在新坐标系下的分量形式，看成是一个新矢量的变换之前的坐标系下的分量形式。这个视角的转变，说起来有点拗口，实际上，是我们以前做坐标变换的时候经常使用的一个角度：把变换以后的矢量看作

是一个新的矢量，而不是原来的矢量在新的坐标下的形式。这两种视角其实是相通的。当然，从抽象矢量的角度来说，我们更加喜欢把矢量看成不变的东西，把不同的数字看作是同一个矢量在不同坐标系下的分量。

需要指出的是，没有 Dirac 抽象矢量记号，我们几乎就不可能说明白，认识清楚这些问题。把抽象的矢量和具体的分量形式分开来，这个是说清楚这些问题的最重要的一步。

3.5 作业

习题 3.1 本课程到现在为止主要讲了那些内容，它们之间有什么关系，用概念地图来表达，看不见几个模块之间的联系的话，可以分成好几片。

习题 3.2 从 σ_z 表象下的 Pauli 矩阵开始，求 3 个二维 Pauli 矩阵的本征值本征向量，并在 σ_x 表象——也就是用 σ_x 的本征矢量当作基矢——中表达三个 Pauli 矩阵。

习题 3.3 在 σ_y 表象中表达三个 Pauli 矩阵。

习题 3.4 记 $\sigma_x, \sigma_y, \sigma_z$ 分别为 $\sigma_1, \sigma_2, \sigma_3$。定义 ϵ_{ijk} 为一个顺序和乱序记号（它有自己的名字叫做 Levi-Civita 记号）：当 $ijk = 123, 231, 312$ 的时候它等于 1，当 $ijk = 132, 213, 321$ 的时候它等于 -1。在 σ_z 和 σ_y 两个表象中，证明 $\sigma_i\sigma_j = \delta_{ij}I + i\epsilon_{ijk}\sigma_k$。其中 δ_{ij} 就是Kronecker δ 记号，i, j 相等的时候等于 1，否则等于 0。无论在哪个表象中，这三个算符之间的这个关系都成立是一个很重要的事实。

习题 3.5 从某个表象下这个关系成立开始，运用表象变换证明 $\sigma_i\sigma_j = \delta_{ij}I + i\epsilon_{ijk}\sigma_k$ 在无论在哪个表象中都成立。（选做）

习题 3.6 定义 $\sigma_r = \vec{\sigma} \cdot \hat{r}$，其中 $\hat{r}$ 是三维空间方向矢量，$\vec{\sigma} = [\sigma_x, \sigma_y, \sigma_z]^T$ 就是把 Pauli 矩阵看作一个三维空间矢量。在 σ_r 表象中表达三个 Pauli 矩阵。

习题 3.7 计算 3 个二维 Pauli 矩阵的对易关系，计算 $\sigma_z, \sigma_x + i\sigma_y, \sigma_x - i\sigma_y$ 之间的对易关系。可以在某个表象下计算，也可以用前面证明的任意表象下都成立的一般关系 $\sigma_i\sigma_j = \delta_{ij}I + i\epsilon_{ijk}\sigma_k$。我鼓励你两种方式都做一下。

习题 3.8 定义 $\sigma_r = \vec{\sigma} \cdot \hat{r}$。$\sigma_{r_1}$ 和 σ_{r_2} 是对应着两个方向 $\hat{r}_1$ 和 $\hat{r}_2$ 的算符。计算 $\sigma_{r_1}\sigma_{r_2}$。可以在某个表象下计算也可以运用前面证明的任意表象下都

成立的一般关系 $\sigma_i\sigma_j = \delta_{ij}I + i\epsilon_{ijk}\sigma_k$。为了计算简单，两个方向矢量都可以取 $\phi_1 = 0 = \phi_2$。看一看计算得到的算符是否还是某一个自旋算符 σ_{r_0}。接着计算这个算符在 σ_z 本的两个征态上的均值，也就是 $\langle\uparrow_z|\sigma_{r_1}\sigma_{r_2}|\uparrow_z\rangle$ 和 $\langle\downarrow_z|\sigma_{r_1}\sigma_{r_2}|\downarrow_z\rangle$。

习题 3.9 用矩阵运算的形式，求 $\sigma_z\otimes\sigma_x$、$\sigma_z\otimes I$ 和 $I\otimes\sigma_y$ 的本征值与本征向量，其中 $\otimes$ 为矩阵直积（不明白的话 google 之，wikipedia 之）。明确写出你写下来的矩阵的基矢是什么。注意任意一个矩阵都是抽象算符在某一套基矢下的分量形式。注意不同的基矢下计算的复杂程度不一样。

习题 3.10 用抽象矢量记号，也就是 Dirac 符号的形式，求 $\sigma_z\otimes\sigma_x$、$\sigma_z\otimes I$ 和 $I\otimes\sigma_y$ 的本征值与本征向量。

习题 3.11 尝试用矩阵运算和抽象矢量记号，求 $\sigma_x\otimes\sigma_x+\sigma_y\otimes\sigma_y+\sigma_z\otimes\sigma_z$ 的本征值与本征向量。如果用矩阵运算，注意你所用的基矢。

习题 3.12 计算 $e^{is_x\theta}$，$e^{is_y\theta}$ 和 $e^{is_z\theta}$，其中 $s_j = \frac{1}{2}\sigma_j$。提示：考虑 Euler 公式，考虑正弦余弦函数的多项式展开。

习题 3.13 计算 $e^{is_r\theta}$，其中 $s_r = \frac{1}{2}\sigma_r$。

习题 3.14 独立地（不看书，或者看完书以后）用 Dirac 符号证明复对称矩阵（Hermitian Matrices）不同本征值对应着的本征向量相互正交。

习题 3.15 用 Dirac 符号、本征向量、本征值表示矩阵的迹。

习题 3.16 有 Hermitian 算符 A 和 B，请证明 $tr(AB) = tr(BA) = \sum_\alpha \alpha B_\alpha$，其中 α 是 A 的本征值，$B_\alpha = \langle\alpha|B|\alpha\rangle$。

3.6 本章小结

在这一章里面，我们通过区分左右矢量定义了抽象的矢量和矩阵，然后利用表象——某个相容算符集合的所有相互正交的本征矢量——把抽象矢量与分量形式联系了起来。两个表象之间的联系本身不是一个自然定义的算符，但是，通过左右矢量，我们仍然可以把这个表象之间的变换看作是算符。从计算上来说，这一章唯一的新的东西是复数域上的列矢量转置成行矢量的时候，需要加一个复共轭。只有这个复共轭和转置的联合起来的

Hermitian 共轭操作能够保持复数域上矢量的内积。从概念上来说，矢量和算符是不依赖于分量形式的存在，它们可以在不同的表象下表现为不同的分量形式。

本章一开始所讨论的数学与科学的关系，数学与物理学的关系，以及如何学习数学，需要你好好体会。这个体会有助于创造性地运用数学和创造数学。数学是对结构的描述，是我们进行思考的语言。所谓科学，就是一个“管用”的现实世界的心智模型，而心智模型中最主要的就是数学模型。在这里，管用，就是指，能够用来理解和解释现实世界。

我们已经提到本书的习题少而精，一定要做。这里再强调一遍，本章的习题不仅要尝试着做，一定要都会做（除了标明选做的题目）。这一点非常非常重要。Dirac 符号、谱展开这些在将来的学习中非常非常的重要。

第四章　离散状态的概率论

我们不讨论这个世界到底是决定性的还是随机的，我们把确定性的事件也看成随机事件的一个特例，服从 δ 分布的一个特例。所谓 δ 分布，以后会介绍，就是在给定时刻这个“随机”变量仅有一个可能取值。在这个视角下，所有的经典世界都由概率论描述。

在这一章里面，我们要解决的问题是如何描述随机事件。从古典概型，我们已经知道随机事件可以用随机变量来描述，离散随机变量的每一个可能取值对应着一个概率，连续随机变量可以用概率密度分布函数描述，同时离散和连续随机变量都可以用累积概率分布函数来描述。这些还不够，我们要讨论一下随机事件的定义，经典概率论最核心的性质是什么（将来我们会看到这个核心是概率叠加原理：互斥事件的加法），然后利用上一章学到的矩阵和 Dirac 符号来介绍一种新的描述概率的方法——密度矩阵。接着，我们讨论一下离散时间的随机过程——这部分推荐给学有余力的读者。

最后，我们会简短地讨论一下随机变量的测量。这是一个基本上所有的教材都不讨论的问题。我们认为其实关于这个问题的认识非常重要。例如，我们考虑这样一个事件：扔出色子并且色子的某一面——例如 6 这一面——向上之后，观测这个正面的值，然后我们说这个时候这个随机事件的概率分布就是 6 这一面 100%，其他取值的概率为零。那么是否意味着在观测之前的概率分布——每一面都是 $\frac{1}{6}$——是错误的呢？我们认为，如果认为是错误的，则是对概率论和随机变量的测量的非常错误的理解。如果认为是正确的，那问题来了，测量之后之后的状态不一样，中间只发生了测量：难道测量改变了状态？

在原来的这个硬币的状态到后来的这个硬币的状态之间，存在着一个操作：测量。测量的前后，硬币的状态确实原则上可以发生改变的。在测量之后，确实硬币的状态是 100% 取值为 6，但是并不意味着测量之前的状态的描述也发生了改变。大多数经典客体的状态在测量前后不发生改变。其

实，这一点没有必要对于随机客体还保持。原则上，这个世界上可以存在着真正随机的客体。然而，如果真的存在，那么我非常怀疑，一个没有深刻理解概率论的学者，可以理解这个世界的任何部分。

这个部分的推荐阅读材料是：Kolmogorov的《概率论导引》[19]，王梓坤的《概率论基础及其应用》[32]，何书元的《概率引论》[33]，Chung的《A Course in Probability Theory》[34]。

4.1 古典概型概率论的 Dirac 符号形式

古典概型对于随机事件有一个基本的假设：存在着某一个基本的简单事件的集合，其中每一个基本简单事件的概率是明确的甚至是相等的[1]，然后复杂事件的概率总是可以通过分解成基本简单事件的方式来计算出来。在这里，我们先通过几个古典概型的例子，来了解一下概率论，介绍一下古典概型的 Dirac 符号形式。然后，我们会展示这种基于简单基本事件的概率论存在一些问题，于是我们需要构造更加一般的概率论。

例 4.1 (**对称色子的几率问题**)：计算两个六面的对称的色子出现的正面向上的数字之和的几率。

扔一个色子的事件有 6 个可能结果 $x \in \{1,2,3,4,5,6\}$，每一面出现的可能性都是 $P(x) = \frac{1}{6}$。于是，计算扔两个色子（$y = x_1 + x_2$）的所有可能性，我们的到 $y \in \{2, \cdots, 12\}$，注意各个数值不再是等几率的了。我们需要数一数每种可能出现的次数。例如 $2 = 1 + 1$ 只有一种可能，于是其几率等于

$$P(y=2) = P(x_1 = 1, x_2 = 1) = P(x_1 = 1)\, P(x_2 = 1) = \frac{1}{36}.$$

再如 $10 = 4 + 6 = 6 + 4 = 5 + 5$，于是

$$P(y=10) = P(x_1 = 4, x_2 = 6) + P(x_1 = 6, x_2 = 4) + P(x_1 = 5, x_2 = 5) = \frac{3}{36}.$$

其他的都可以类似的计算出来。

这里我们把最基本的事件找出来，确定它们的几率，然后用他们来计算复杂的事件的几率。其中我们用到了互斥事件的概率是各个事件的概率的叠加，也用到了独立事件的概率是各个事件概率的乘积。存在基本事件，可以找到

[1]尽管有一般的情况，但是很多时候古典概型从某些等概率的基本事件出发来构造更复杂的事件的概率。

它们的几率，然后运用概率叠加原理和独立事件原理来计算复杂事件的概率是古典概型的特征。

例 4.2 (**不太守时的约会**) ：Alice 和 Bob 约会，约定三点钟见面，可是这两个人都不太守时，耐心也不太好，会在三点钟的前后 5 分钟出现，而且出现以后只等对方 5 分钟。问他们能够实现这个约会的几率是多少？

首先，我们确定事件的描述，Alice 和 Bob 到达的时间，记为 t_A 和 t_B。它们的取值范围是 $t \in [2.55, 3.05]$。然后我们再来确定这样一个事件 (t_A, t_B) 的几率 $P(t_A, t_B)$。我们发现这个几率对于任何的在取值范围内的取值都是一样的 (或者事先给定的形式，例如集中在刚好 3 点的几率大一些)，可是我们写不出来显式表达式。一对 (t_A, t_B) 是在所有取值空间 $[2.55, 3.05]\times[2:55, 3:05]$ 中的一个点，如图 4.1 所示。把一个点和一个正方形区间相比较，我们不会做这样的计算。不过，换一个方式，我们问，在 (t_A, t_B) 附近一个小小的正方形区域 Δ 之内的几率是多少，我们就能够写出来了：$\rho(t_A, t_B)\cdot\Delta$。于是，我们发现，实际上所有的信息都可以放到这个 $\rho(t_A, t_B)$ 中去。这个东西被称为概率密度函数。如果已知概率密度函数，那么任意一个点附近的任意一个区域的几率都可以计算出来。下面，我们以 3:00 为时间原点，以分钟为单位写下 t_A, t_B 的值。这里，我们考虑各个点等几率，因此 $\rho(t_A, t_B) = \frac{1}{100}$。这里，我们用到了归一化条件，

$$1 = \int_{t_A, t_B \in [-5,5]} dt_A dt_B \rho(t_A, t_B).$$

接着，我们来看，Alice 和 Bob 约会成功的条件是什么。由于等待时间是 5 分钟，所以相遇的条件是 $|t_A - t_B| \le 5$。

于是，他们约会成功的可能性是，

$$P = \int_{|t_A - t_B| \le 5} dt_A dt_B \rho(t_A, t_B) = \frac{1}{100}\int_{|t_A - t_B| \le 5} dt_A dt_B = \frac{3}{4}.$$

从概率分布函数到概率密度函数是从离散随机变量到连续随机变量的概率的描述方式的自然的转变。同时，归一化因子的计算也是概率论问题中重要的一步。现在，我们利用 Dirac 符号，把以上的概率分布函数形式上写成矩阵的形式。

我们把一个对称的色子的分布函数记作，

$$\hat{\rho} = \frac{1}{6}|1\rangle\langle 1| + \frac{1}{6}|2\rangle\langle 2| + \frac{1}{6}|3\rangle\langle 3| + \frac{1}{6}|4\rangle\langle 4| + \frac{1}{6}|5\rangle\langle 5| + \frac{1}{6}|6\rangle\langle 6|. \tag{4.1}$$

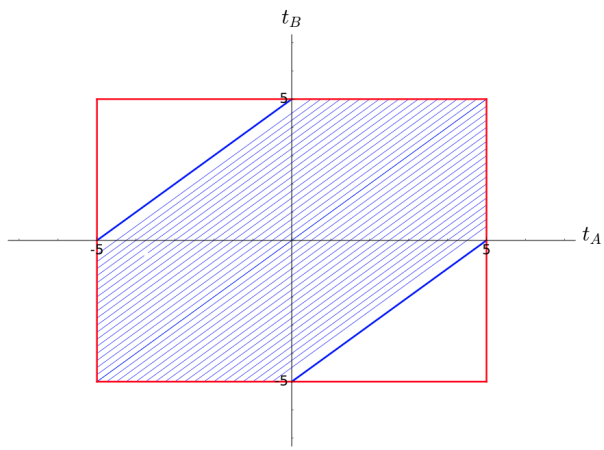


图 4.1: Alice 和 Bob 的约会可能时间和约会成功的区域（阴影部分）。图是用免费数学软件 SageMath 做的。SageMath 的详细情况可以访问sagemath.org的主页。

其中，出现在 Dirac 符号中的就是状态，前面的系数就是这个状态出现的几率。更一般地，可以写做

$$\hat{\rho} = \sum_j P(j) |j\rangle\langle j| . \tag{4.2}$$

在这里，暂时我们整体认读 $|j\rangle\langle j|$，表示事件 j。在其它的概率论的教材里面，这样的一个色子的状态可以写成一个形式上的“矢量”——注意由于归一化因子的存在这个矢量的加法和数乘不遵循一般的矢量的加法和数乘——形如，

$$p = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \end{bmatrix}, \tag{4.3}$$

或者一个列表

$$p=\begin{cases}\frac{1}{6} & \text{if } x=1\\ \frac{1}{6} & \text{if } x=2\\ \frac{1}{6} & \text{if } x=3\\ \frac{1}{6} & \text{if } x=4\\ \frac{1}{6} & \text{if } x=5\\ \frac{1}{6} & \text{if } x=6\end{cases}. \tag{4.4}$$

这三个表达方式的含义完全是一样的。我们更喜欢公式 (4.2) 中的矩阵形式的表达式。由于这个矩阵表示的含义是一个概率分布函数，或者概率密度分布函数，我们称这个矩阵为密度矩阵。我们采用密度矩阵的语言来描述概率论，因为：第一，这个形式在可以原封不动地用在量子力学里面；第二，这个形式的所有的计算的规则和矩阵是一样的。一般地来说一个密度矩阵要满足以下条件，

$$\langle x|\rho|x\rangle \geq 0, tr(\rho)=1. \tag{4.5}$$

前者表示概率或者概率密度应该大于零，后者就是归一化条件。

实际上，我们也可以写下来 Alice 的状态的密度矩阵，例如

$$\rho^A=\int dt_A \frac{1}{10}|t_A\rangle\langle t_A|. \tag{4.6}$$

但是，我们所有的量子系统的讨论都面对离散变量。因此，连续变量的密度矩阵形式就不在此讨论了。

在概率论中，我们关注的往往是某个可观测量的某种形式的平均值。因此，除了概率分布函数，可观测量也是概率论的重要组成部分。例如，色子的面值就是一个可观测量。除了色子的面值，我们也可以制定另一个可观测量，例如色子的面值是偶数就是 1，是奇数就是 −1。我们记色子的面值的可观测量为 X，记后者为 O。如果我们需要计算 O 的平均值，我们可以这样做，

$$\langle O\rangle=\sum_j O(j)P(j) \tag{4.7}$$

其中，对于我们的例子

$$O(j)=(-1)^j. \tag{4.8}$$

当然

$$X(j) = j. \tag{4.9}$$

现在，有了 Dirac 符号形式的分布函数，我们还需要 Dirac 符号形式的可观测量，才能完全替代原来的概率论的符号体系。我们定义，

$$\hat{O} = \sum_j O(j) |j\rangle \langle j| \tag{4.10}$$

而且平均值的计算按照以下公式，

$$\langle O \rangle = tr\left(\hat{O}\hat{\rho}\right). \tag{4.11}$$

现在，我们来验证一下是否公式 (4.11) 等价于公式 (4.7)。

$$\begin{aligned}
\langle O \rangle = tr\left(\hat{O}\hat{\rho}\right) &= tr\left(\sum_j O(j) |j\rangle \langle j| \sum_n P(n) |n\rangle \langle n|\right) \\
&= tr\left(\sum_{j,n} O(j) |j\rangle \delta_{jn} P(n) \langle n|\right) \\
&= tr\left(\sum_j O(j) P(j) |j\rangle \langle j|\right) \\
&= \sum_j O(j) P(j).
\end{aligned}$$

完全等价，如果我们假设

$$\langle j | n \rangle = \delta_{jn}. \tag{4.12}$$

这个基本事件之间的正交关系是我们这套 Dirac 符号体系额外需要的假设。其含义是这两个事件不可能同时发生：一个色子不可能同时编号为 j 的面向上又有编号为 n 的面向上。

经典离散随机变量的概率的密度矩阵形式的定义式公式 (4.2)，经典离散随机变量的可观测量的密度矩阵形式的定义式公式 (4.10)，观测量平均值的计算公式公式 (4.11)，以及古典概型的基本事件之间的正交关系公式 (4.12) 是我们的符号体系中的核心定义，用来替代原来的概率论的列表形式的分布函数以及求平均的公式。这个替代除了能够显式地写下来分布函数，而不用写一个长长的列表之外，目前看起来，没有什么好处。以后我们会看到这样做的价值。

我们来运用 Dirac 符号形式的分布函数做一个计算，熟悉一下这套形式。

例 4.3 (**盒子里放小球的几率**)：一个可以容纳无穷多个小球的盒子，每放入一个小球，需要外界输入小球一份能量 ϵ，反之也成立——每提供一份 ϵ 的能量盒子内的小球就增加一个。现在假设这个能量份数 n 是一个随机变量，符合如下分布，

$$P(n) = \frac{1}{Z} e^{-\beta n \epsilon}, \tag{4.13}$$

求得其中的归一化常数 Z，以及这个盒子里面的平均粒子数。

状态空间就是盒子里面小球的数量的所有可能取值，利用 Dirac 符号就是以下的事件集合 $\{|n\rangle\langle n|\}$。相应的密度矩阵就是，

$$\rho = \sum_{n=0}^{\infty} \frac{1}{Z} e^{-\beta n \epsilon} |n\rangle \langle n|, \tag{4.14}$$

正定性非常明显，我们来看看归一化。

$$1 = tr(\rho) = \sum_m \langle m| \sum_{n=0}^{\infty} \frac{1}{Z} e^{-\beta n \epsilon} |n\rangle \langle n| m\rangle = \frac{1}{Z} \sum_{mn} e^{-\beta n \epsilon} \delta_{mn} = \frac{1}{Z} \sum_n e^{-\beta n \epsilon}. \tag{4.15}$$

于是，

$$Z = \sum_n e^{-\beta n \epsilon} = \frac{1}{1 - e^{-\beta \epsilon}}. \tag{4.16}$$

这样，我们就得到了盒子里的小球数量的状态的密度矩阵的显式表达式，

$$\rho = \left(1 - e^{-\beta \epsilon}\right) \sum_{n=0}^{\infty} e^{-\beta n \epsilon} |n\rangle \langle n|. \tag{4.17}$$

现在我们来解决第二个任务，盒子里面的平均粒子数。我们先写下粒子数算符，

$$\hat{n} = \sum_n n |n\rangle \langle n|. \tag{4.18}$$

于是，从公式 (4.11) 我们得到可观测量平均值，

$$\langle n\rangle = tr(\hat{n}\rho) = \left(1 - e^{-\beta \epsilon}\right) \sum_n n e^{-\beta n \epsilon} = -\frac{1}{\epsilon} \frac{\partial}{\partial \beta} \ln Z = \frac{1}{e^{\beta \epsilon} - 1}. \tag{4.19}$$

这一章我们的主要任务就是得到古典概型的 Dirac 符号形式。这个主要任务我们已经完成。对于我们后面的学习来说，稍微了解一下概率空间的一般的定义也是有好处的。下面，我们来介绍一般的概率空间的定义。

4.2 现代概率三元体

我们先来看一看基于基本事件以及基本事件的几率来计算复杂事件的（等）几率的古典概型有什么问题。

例 4.4 (Bertrand **圆内弦长的问题**)：单位元内任意取一条弦，其长度小于 $\sqrt{3}$ 的几率是多少？

考虑弦是这样来画的：选择圆上任意一点作为弦的一个端点，然后利用所对应的圆周角 θ 的弧度来决定另外一个端点。这个圆周角的取值范围为 $[0,2\pi]$。然后，当 $\frac{2\pi}{3}<\theta<\frac{4\pi}{3}$ 的时候，弦长大于 $\sqrt{3}$。于是，概率为 $\frac{2\pi}{3}\frac{1}{2\pi}=\frac{1}{3}$。

另外，我们再考虑这个弦的中点，一条弦完全由这个中点确定（给定这个中点有且只有一条弦，除了这个中点正好就是圆心）。我们发现，当且仅当这个中点在半径为 $\frac{1}{2}$ 的圆以内的时候，弦长大于 $\sqrt{3}$。满足这个条件的中点有 $\frac{\pi\left(\frac{1}{2}\right)^2}{\pi}$ 这么多，于是概率为 $\frac{1}{4}$。

其实还有另外一种确定弦的方式，可以得到别的答案。这里就不介绍了。这样一个问题就有了多个答案，每一个都挺有道理。那么，哪一个是对的呢？这就有问题了。这个问题的更多讨论见 Wikipedia 的“概率论中的 Betrand 谬论”页面，https://en.wikipedia.org/wiki/Bertrand_paradox_(probability)。

在这个例子中，我们发现存在着多种可以假设为等概率的基本简单事件的选择，不同的选择导致同一个问题的不同的答案。于是，基于基本简单事件的概率或者概率密度的古典概型，就有问题了。现代概率论通过下面的概率公理化定义来解决这个问题。我们先看一下这个公理化定义是什么，然后在做进一步的关于这个样的定义的含义和动机的讨论。

定义 4.1 概率空间：概率三元体 $(\Omega,\mathcal{F},P)$，其中集合 Ω，集合元素对应简单事件记号 $\omega\in\Omega$，Ω 的子集 A 构成集合 $\mathcal{F}$ 是 Ω 上的 σ 代数，即 $\mathcal{F}$ 满足

1. F 至少包含 Ω：$\Omega\in\mathcal{F}$；
2. 对集合的补集操作封闭：$\bar{A}\in\mathcal{F}$，$\forall A\in\mathcal{F}$；
3. 对可数个集合的并集封闭：$\cup_{j=1}^{\infty}A_j\in\mathcal{F}$，$\forall A_j\in\mathcal{F}$。

从 $\mathcal{F}$ 到 $[0,1]$ 的映射 P，满足下列条件

1. 完全性：

$$P(\Omega)=1; \tag{4.20}$$

2. 可列可加性：对于可数个不相交的集合（互斥事件，$A_i \cap A_j = \phi, \forall A_i, A_j$）

$$P\left(\cup_j A_j\right) = \sum_j P\left(A_j\right). \tag{4.21}$$

$\mathcal{F}$ 是 σ 代数的要求是为了保证对于任意两个概率有定义的集合，它们的集合操作——交并补——都有定义。也就是说，在概率的意义上，我们从两个集合的概率开始，我们可以讨论这两个集合对应着的事件合起来（“或”关系，集合并运算）的事件的几率，或者是都发生（“与”关系，集合交运算）的几率，或者是一个发生另一个不发生（“非”，集合补运算）的几率，或者两者都不发生的几率。映射 P 保证在 $[0,1]$ 之间也是很好理解的，概率最好是大于零的，而且最大最大就是 1。完全性也很好理解，所有的事件的整体的概率显然应该是 1。定义中的所有这些都容易理解，而且相对平凡。于是，所有的概率论的定义的核心的一条，就是为了保证“可列可加性”：对于互斥的事件，其整体的概率等于各个互斥部分的概率之和。我们称这一条为“概率性叠加原理”。也就是说，就是这个概率性叠加原理体现了概率论的所有的意义。以后，我们会有更深刻的体会。对于同一个 Ω，不同的 F，不同的 P 都会导致不同的概率的定义。实际上 Betrand 园内弦长的问题就是定义了不同的 P——从复合事件到简单事件的拆分有多种。

古典概型的定义是基于基本简单事件以及基本简单事件的（等）几率来定义复合事件的几率的。其中概率叠加原理和独立事件原理非常重要。相当于利用基本简单事件的几率加上概率叠加原理就定义了所有的其它的事件的几率——就看看这个复合事件最后分解称为什么样的基本简单事件就行了。但是，我们通过园内弦长的问题已经看到基于基本简单事件的几率的定义会出问题。于是，概率空间的公理化定义直接就把概率叠加原理做为最基本的定义。这样，假设我们希望找到某个复合事件的几率，实际上，我们需要保证的，就是，不管这个复合事件可以如何分解成相互不相交的子集合（分解的方式可能不唯一，$\mathcal{F}$ 保证分解以后的各个子集合还是可以讨论其概率的），只要我们能够保证分解以后按照概率叠加原理合起来之后得到的值是一致的，那么这个映射 P 描述的，就是一种概率现象。所以，我们说，概率叠加原理就是概率论的核心。

在这里，我们先来展示一下，通过这一条概率性叠加原理来定义一下任意两个集合 A，B 的并的几率，如果这两个集合 A 和 B 已经确定是 $\mathcal{F}$ 的元素的话。我们先来看 $A \cap B$，如果它等于空集，则 A 和 B 是互斥事件，于是 $P(A \cup B) = P(A) + P(B)$，解决。如果 $A \cap B \neq \phi$，我们把 $A \cup B$ 分解

成三个互斥的集合 $A\cap\bar{B}$，$\bar{A}\cap B$ 和 $A\cap B$。于是，$P(A\cup B)=P(A\cap\bar{B})+P(\bar{A}\cap B)+P(A\cap B)$。当然，我们可以进一步讨论 $A\cap\bar{B}$ 的概率，如果它不是已知的话。对于任意的事件 A，如果我们找到一组 A 的互斥分解并且分解以后的时间的概率已知，则我们就可以通过概率叠加性原理来求得 A 的概率。并且，如果存在多种这样的分解，所得到的结果一致。

事实上，古典概型的等概率假设就指定了一种事件分解的方式。回到我们的圆内弦长的问题，可以证明，从我们前面的任何一种等概率假设都可以构造一个满足上述要求的概率空间。

实际上，再进入下一部分内容之前，我们还应该完成这样一件事情：从概率论的公理化定义出发，找出适当的条件，在这个条件下随机变量的分布函数存在。但是，这个任务超过本书的水平，在我们发现一个好的能够在本书的概念和方法的范畴内讨论这个问题之前，我们直接就用分布函数来描述概率论。

4.3 Dirac 符号作为概率论的语言

现在，让我们用概率论的 Dirac 符号语言来看一下概率空间定义的最重要的性质：概率性叠加原理，在我们的语言下如何表达。概率叠加原理说，如果事件 A,B 互斥，则

$$P(A\cup B)=P(A)+P(B). \tag{4.22}$$

我们希望第一，能够把 $A\cup B$ 写成，

$$A\cup B\Leftrightarrow\hat{A}+\hat{B}, \tag{4.23}$$

也就是事件的加法有意义，而且正好就是集合取并操作的含义；第二，能够有

$$P\left(\hat{A}+\hat{B}\right)=P\left(\hat{A}\right)+P\left(\hat{B}\right), \tag{4.24}$$

也就是事件之间的加法与概率之间的加法保持一致。现在，我们来做一个检验。

在这里我们还是面对古典概型。给定一个密度矩阵，

$$\hat{\rho}=\sum_j p_j\left|j\right\rangle\left\langle j\right|, \tag{4.25}$$

我们来考虑一个事件 A 的概率。记事件 A 的每一个基本简单事件的元素为 α_j，则

$$\hat{A} = \sum_{\alpha_j \in A} \left|\alpha_j\right\rangle \left\langle\alpha_j\right|, \tag{4.26}$$

从而通过公式 (4.11) 我们得到事件 A 发生的概率

$$P\left(\hat{A}\right) = tr\left(\hat{A}\hat{\rho}\right) = \sum_{\alpha_j} p_{\alpha_j}. \tag{4.27}$$

这里，我们把事件集合 A 作为观测量 $\hat{A}$，需要用到示性函数 $I_A(\omega)$，

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}. \tag{4.28}$$

也就是判断元素 ω 是否属于集合 A。于是，

$$\hat{A} = \sum_j I_A(j) \left|j\right\rangle \left\langle j\right| = \sum_{\alpha_j \in A} \left|\alpha_j\right\rangle \left\langle\alpha_j\right|. \tag{4.29}$$

如果有一个与 A 互斥的事件 $\hat{B} = \sum_{\beta_j \in B} \left|\beta_j\right\rangle \left\langle\beta_j\right|$，也就是任意 α_j, β_k 满足

$$\left\langle\alpha_j \,|\, \beta_k\right\rangle = 0. \tag{4.30}$$

也就是集合 A, B 没有相同的元素。事件 B 发生的概率为，

$$P\left(\hat{B}\right) = \sum_{\beta_j} p_{\beta_j}. \tag{4.31}$$

现在，我们知道了

$$P\left(\hat{A}\right) + P\left(\hat{B}\right) = \sum_{\alpha_j \in A} p_{\alpha_j} + \sum_{\beta_k \in B} p_{\beta_k}. \tag{4.32}$$

我们也知道了

$$\hat{A} + \hat{B} = \sum_{\alpha_j \in A} \left|\alpha_j\right\rangle \left\langle\alpha_j\right| + \sum_{\beta_k \in B} \left|\beta_k\right\rangle \left\langle\beta_k\right| \tag{4.33}$$

还有事件 $A \cup B$。我们想看看事件 $A \cup B$ 的 Dirac 符号形式是不是正好就是 $\hat{A} + \hat{B}$，而且 $P\left(\hat{A} + \hat{B}\right) = P\left(\hat{A}\right) + P\left(\hat{B}\right)$ 是否成立。第一部分非常简单，记 $C = A \cup B$，于是

$$\hat{C} = \sum_j I_C(j) \left|j\right\rangle \left\langle j\right| = \sum_j I_A(j) \left|j\right\rangle \left\langle j\right| + \sum_j I_B(j) \left|j\right\rangle \left\langle j\right| = \hat{A} + \hat{B}. \tag{4.34}$$

注意，如果 $A \cap B \neq \phi$ 上面的等式的中间一步是不对的。我们再来看第二部分。由于公式 (4.11)（$\langle O \rangle = tr\left(\hat{O}\hat{\rho}\right)$）是线性的，于是只要第一步正确，第二布自然就是正确的。

于是，通过 Dirac 符号，我们把事件 A, B 写成了矩阵 $\hat{A}, \hat{B}$ 的形式，同时，互斥事件的加法真的可以表达成矩阵加法运算的形式，而且对于这个加法运算，求观测量平均值的公式公式 (4.11) 仍然适用。对于不是互斥的事件，我们总是可以把它分解成互斥的部分。因此，所有的的事件之间的加法运算，以及这个加法运算的概率含义，尤其是概率叠加原理公式 (4.22)，都被保留了下来。这个就是我们引入概率论的 Dirac 符号的目的。在 Dirac 符号的形式下，互斥集合取并的操作直接就是算符的加法运算，求可观测量平均值的定义是线性的，这两点直接就保证了概率叠加原理的成立。因此，Dirac 语言是概率论的一种自然的语言。

最后，我们注意到，这一章定义的所有的可观测量矩阵都是对角的，不存在这样的可观测量，$O_{ij} \neq 0$ 当 $i \neq j$，也就是一般的

$$\hat{O} = \sum_{i,j} O_{ij} \left|i\right\rangle \left\langle j\right| . \tag{4.35}$$

由于有观测量的这个性质，我们发现所有的观测量矩阵都是对易的，也不用做表象变换，因为在这个自然的表象下，矩阵本来就是对角的。这一点，实际上就表明，我们的矩阵符号有点多余，实际上写下对角元就够了。逻辑上这个符号系统是自洽的。在经典概率论的层次，这套新的符号系统实际上不带来新的结果，仅仅是理解上的不同。

以 N 个离散状态的系统为例，总结一下 Dirac 符号的概率论和通常的概率论的不同：

- 状态：$N \times N$ 矩阵 $(\rho)_{N\times N}$，N 维矢量 $(P)_{1\times N}$；
- 可观测量：$N \times N$ 矩阵 $(O)_{N\times N}$，N 维矢量 $(O)_{1\times N}$；
- 测量结果的计算：求迹 $tr\,(A\rho)$，内积 $O^T P$；
- 测量过程的理解（测量前状态，测量所得值，测量后状态）：Dirac 符号形式和通常的概率分布矢量形式一样，都表现出来测量前后分布函数有变化。

4.4 概率转移算符 *

本节为选读部分。

有了表示状态和可观测量的 Dirac 符号形式，我们就可以用这样的算符记号来表示概率转移算符 [21] 了。对随机状态的合理的一个操作要满足一些要求，例如操作之后得到的状态还必须是一个概率分布（满足归一化、非负、互斥事件可加性等）。如果我们同时还要求被操作的客体的状态空间（状态的集合）不变，仅仅改变其状态，也就相当于给状态的概率做了一个重新分配，那么，这样的操作必须是状态矩阵 ρ 上的一个线性算符（记为 M），也就是 $M(a_1\rho_1 + a_2\rho_2) = a_1 M(\rho_1) + a_2 M(\rho_2)$。按照线性性以及仍然非负归一的要求，我们可以写下来这个转移矩阵的一般形式。

概率转移算符：在传统矢量记号下

$$P(t+1) = MP(t), \tag{4.36}$$

其中矩阵 $\left(M^i_j\right)_{N\times N}$ 满足 $\sum_i M^i_j = 1$；在 Dirac 记号下，

$$\rho(t+1) = U\rho(t)U^\dagger, \tag{4.37}$$

其中矩阵 $\left(U^i_j\right)_{N\times N}$ 满足 $UU^\dagger = 1 = U^\dagger U$。在这里我们按照物理学家的习惯把转移矩阵作用在分布函数上表示成矩阵在左矢量在右，而且概率分布是一个列矢量。需要注意这个习惯和数学家的不一样。在这里状态变化的表示是一个算符作用 U 以及一个相应的算符 $U^\dagger$ 左右乘在一个密度矩阵上。这大概可以这样来理解：ρ 总是可以看作 $\rho = \rho_{ij}|i\rangle\langle j|$，这样的右矢量和左矢量的乘积项的叠加，如果一个算符作用在右矢量上的表达式大约是 $U|i\rangle$ 的话，那么，相应的左矢量的形式就是 $\langle i|U^\dagger$，于是 $|i\rangle\langle j|$ 在这个操作下的变化就是 $U|i\rangle\langle j|U^\dagger$。

用这个记号我们描述以下两个过程。其中第二个过程稍微复杂一些。

例 4.5 (**硬币确定翻转的算符**)：有一个硬币每一个单位时间以后都会被翻转。写下这个动力学过程的普通概率论形式和密度矩阵形式的表达式。

硬币的状态可以采用两种表示方法，概率矢量

$$P = \begin{bmatrix} p \\ 1-p \end{bmatrix}, \tag{4.38}$$

或者概率矩阵，

$$\rho^c = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}. \tag{4.39}$$

翻转操作是算符

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{4.40}$$

可以验证翻转过程可以表示为

$$P(t+1) = XP(t), \tag{4.41}$$

或者

$$\rho^c(t+1) = X\rho^c(t)X^\dagger. \tag{4.42}$$

例 4.6 (**硬币随机翻转的算符**)：一个硬币每一个单位时间以后都会被决定是否翻转。决定的过程如下：抛另外一个完全随机的硬币，如果得到硬币向上则翻转，否则不翻。写下这个动力学过程的普通概率论形式和密度矩阵形式的表达式。

硬币的状态可以采用两种表示方法，概率矢量

$$P = \begin{bmatrix} p \\ 1-p \end{bmatrix}, \tag{4.43}$$

或者概率矩阵，

$$\rho^c = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}. \tag{4.44}$$

翻转操作和不翻转操作分别是是算符

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.45}$$

可以验证翻转过程可以表示为

$$P(t+1) = 0.5XP(t) + 0.5IP(t) = \frac{I+X}{2}P(t), \tag{4.46}$$

或者

$$\rho^c(t+1) = 0.5X\rho^c(t)X^\dagger + 0.5I\rho^c(t)I^\dagger. \tag{4.47}$$

注意，最后的表达式公式 (4.47) 不能下成算符取和以后再作用的形式，也就是，

$$\rho^c(t+1) \neq \frac{I+X}{2}\rho^c(t)\frac{I+X}{2}. \tag{4.48}$$

甚至这个加起来的算符都不满足约束 $UU^\dagger = 1$。因此，两个随机翻转“合”起来的运算在传统矢量语言下面表现为“取和”——$M = \frac{I+X}{2}$，在 Dirac 符号下面不表现为“取和”——$U \neq \frac{I+X}{2}$。在本书第十四章关于量子博弈的讨论中，我们会看到，这个差别是有深远的意义的。

4.5 作业

习题 4.1 写下一个向上的硬币的状态的完整描述。一个完整的描述包含：状态的数学描述、所有的可观测量的数学描述（还要思考这样的可观测量的含义），以及这些可观测量在这个状态下的可能取值，还有取值的概率，以及均值。

习题 4.2 写下一个随机的硬币的状态的完整描述。

习题 4.3 Alice 和 Bob 正在打赌：硬币的状态向上则 Alice 从 Bob 那里获得 10 元，否则 Alice 给 Bob10 元。现在对于一个向上的硬币，分别用 Dirac 符号和矩阵的形式写下来，Alice 和 Bob 的输赢状态。如果是一个随机的硬币呢？

习题 4.4 Alice 和 Bob 正在打赌：硬币的状态向上则 Alice 从 Bob 那里获得 10 元，否则 Alice 给 Bob10 元。这个硬币初始处于向上的状态。然后，Alice 和 Bob 可以分别去操作这个硬币。他们可以选择的操作包含翻转和不翻转这个硬币这两个操作。

1. 分别用 Dirac 符号和矩阵的形式写下来，Alice 和 Bob 的可能的操作。提示：考虑 Alice 和 Bob 可能选择一个翻转和不翻转这两个操作的某种组合，例如一个这两个操作的概率分布，当然他们之中的任何一人都不能选择同时做这两个操作——同时翻转和不翻转硬币的操作不知道是什么意思。

2. 计算给定这样的概率分布的情况下，Alice 和 Bob 可能获得的钱的数量及其分布，还有平均值。

3. 尝试构造一个映射 H^a（H^b），满足 $H^a(S^a, S^b)$（$H^b(S^a, S^b)$）的值就是 Alice（Bob）的收益。其中，S^a 是 Alice 的操作，S^b 是 Bob 的操作。

4. 把这个得到的映射按照 Dirac 符号的语言写成一个矩阵。注意思考这个时候这个矩阵的基矢是什么。然后验证，这样的矩阵满足 $E^a = tr(\rho^a \otimes \rho^b H^a)$（$E^b = tr(\rho^a \otimes \rho^b H^b)$）。

4.6 本章小结

本章主要介绍了 Dirac 符号形式的概率论。我们发现，概率分布函数可以用一个密度矩阵来表达，见公式 (4.2) 。其中各个基本简单事件所对应的基矢相互正交，见公式 (4.12)。同时，我们引入了把随机变量的函数做为可观测量的概念，见公式 (4.10)。于是，概率论用来求可观测量平均值的基本公式就成了公式 (4.11)。

这套符号本身在概率论的层次不带来新的东西，除了把事件之间的集合运算变成了事件所对应的算符的加法运算而且自然地表示了事件的概率叠加原理。而概率叠加原理，通过一般的概率空间的讨论，我们注意到就是整个概率论定义的核心。

另外，通过整理和对比矢量语言和密度矩阵语言下的状态、可观测量和算符，我们发现，在两套语言下，状态和可观测量都是对角的，一般的算符在密度矩阵的语言下可以是非对角的。也就是说，密度矩阵所在的空间可能比概率矢量更大。这个不一致性实际上就启发我们：是否我们需要讨论一个允许前两者——状态和可观测量——也存在非对角元的理论。以后我们会看到，这个理论就是量子系统的数学模型——量子力学。

第五章　经典力学精要

在这里，我们学习力学的目的是给学习量子力学做准备。为了这个目的，我们需要学习一些概念，例如位形空间、坐标系、力、能量、保守系统等。最关键的是学习通过 Hamiltonian 来描述一个力学系统，得到这个系统的所有的行为。然而，实际上我们应该从力学学习到更加重要的东西。从思想上，从经典力学我们了解什么是物理学——物理学就是描述物体状态（包含运动状态和结构状态等）的变化，以及寻找导致物体状态变化的原因的科学，以及了解把事物抽象成理想模型的观念。这个研究思路以及这个观念不仅仅是物理学的思想，也是整个科学的思想。当然，为了达成这个目的，我们就需要真正地学习一门力学的课程。这个不是我们这一章的任务。所以，在这一章里，尽管我们会尽可能地让你体会这个研究思路和这个观念，我们的主要精力还是会放在如何利用 Hamiltonian 来描述一个力学系统这一点上。已经熟悉 Hamltonian 力学的读者可以跳过这一章。

本章推荐参考书：Landau的《Mechanics》[27]，Feynman的《Feynman 物理学讲义第一卷》[35]，方励之的《力学概论》[36]，漆安慎的《力学》[37]。

5.1　力学思想

常见的力学的基本研究对象是日常生活所见到的桌子以及桌子在一个推动的力作用下运动起来这个现象，或者一个乒乓球以及这个乒乓球在拍子挥动的条件下发生运动这个现象，或者一个汽车以及这个汽车在发动机驱动下和方向盘的控制下发生运动这个现象，或者天上运动的星星以及它们的运动。对于这个研究对象，宏观的物体以及它们的运动，第一个要解决的问题是如何描述他们的运动状态。在这里，物理学引入了一个理想模型——质点，一个有质量没有体积没有形状没有其他任何属性的点。为什么除了质量以外其他的属性都不重要，这是一个很重要的问题。没有体积和形状就要

求我们只关心这个物体作为一个整体的运动，我们暂时不关心这个物体的自转等运动形式。这个运动的形式是所有的运动形式里面最简单的一种。我们称之为物体的平动。其他运动形式在我们这个精要课程里面不再涉及。解决了运动的描述的问题，下面的问题是就是运动状态是否会发生变化，变化的原因是什么。状态描述、状态变化和变化原因这三个问题被称作力学的世界观，也就是力学看问题的时候总是要关心的三个问题。

基于生活经验，我们知道，力的作用是使运动的状态发生改变。我们知道当我们推一张桌子时候，如果我们希望我们停下来之后桌子动得快一点，我们需要推的时间长一点或者推的力气大一点。甚至在我们推动桌子的过程中，我们也可以通过改变这个推力的大小来达到控制其速度的目的。也就是说，当我们把力作用在一个质点上的时候，我们会改变质点的运动速度。当然，我们还不知道力和速度的改变这两者之间具体的关系是什么，和质点的什么属性相关，例如大小、形状、质地、生产厂家等等。不过，有了这个思想认识上的突破之后，剩下的突破就容易多了。其实通过现实生活的经验，除了猜测“力是改变物体运动状态的原因”之外，还有可能得到“力是维持运动的原因”的直观结论。例如，我们松开推着的桌子桌子就不动了。

解决这个到底力是维持还是改变运动的原因，以及力和运动状态的改变具体什么关系的问题依赖于物理学的另一个，也是最重要的，思想：做实验，做测量。这个来自于Galileo的思想使得物理学真正从思辩的层次独立出来，成为科学。Galileo的实验很简单，把一个物块放在一个斜面上，然后让这个物块滑下来到达一个平面上。接着，改变这个平面的光滑程度。物块从斜面下来以后，我们看看这个物块走得多远。这个远近是可以测量的。我们发现，随着光滑程度的不同物块运动的距离是不一样的，而且越光滑的表面物块运动的越远。在物块接触到平面之前，所有的过程是一样的，因此（其实你也可以怀疑这个因此），物块刚刚接触到平面上的状态，不管哪一种光滑程度的平面，都是一样的。于是远近不同只能是平面的光滑程度导致的。Galileo根据他的实验猜测，如果有一个无限光滑的平面，则这个物块会永远运动下去。也就是说，光滑程度的不同可以看作是某种力的不同，而这种力使得物块从运动变成静止。运动状态本身不需要力来维持。对于无限光滑的表面，这种力为零，于是物体的运动状态不发生改变，一直运动下去。现在，我们当然知道这个力就是摩擦力。但是，当时把这个运动状态发生变化的原因归结为摩擦力是革命性的。通过实验，我们还可以研究运动状态与哪些因素有关，例如，如果我们能够给力的大小定下一个度量的标尺，我们就可以

研究给定大小的力对不同的物体的运动状态的改变的异同，例如这些不同的物体可以有不同的形状、颜色、质地等等。当然，最后我们发现这个因素仅仅是物体所包含的物质的量的多少的过程，是一个非常非平庸的过程。

小结一下，理想化模型化，实验与测量，是物理学几个核心的重要的思想；问状态表述、是否变化、变化的原因是典型的力学思考。质点是重要的理想模型之一。物理学很多时候都通过理想化模型化来抓住主要因素暂时忽略次要因素。

5.2 Newton 力学的基本概念

通过质点这个理想模型，我们已经有了用空间位置来描述运动的简单想法。下面，我们把这个想法实现。首先，为了准确地描述这个空间位置，我们需要有一个坐标系：一个原点，三个独立的方向，一个测量单位。通常这个坐标系的原点需要建立在一个参考系上：一个假设不动的东西。确定这个所谓的不动的东西，不是简单的事情。地球在自转和公转，所以在地球上，原则上所有的点都不能作为这个参考系。同样的，太阳系、甚至整个宇宙都在运动，没有一个点是真正意义上不动的。那我们如何选取我们的参考系，然后建立坐标系呢？如果我们的对象是一个在地球表面运动的汽车，那么在汽车的时间尺度内自转和公转造成的影响与汽车的车程相比比较小，于是，地球就是一个好的参考系。以这个汽车出发的时候的地球上的一个点作为原点，是可以接受的。如果我们考虑的是导弹的运动，那么，自转的影响可能就需要考虑。这个时候，参考系可能就要选取地球在某个时刻的点，然后考虑经过这段时间以后地球的运动，或者考虑非惯性参考系的问题。所谓惯性参考系，就是一个“不动”的参考系，或者说一个做匀速运动的参考系。也就是说，这个参考系在相对于另一个参考系运动速度不变或者运动速度为零。所谓非惯性参考系，就是以速度变化的物体为中心建立起来的坐标系。

在下面的讨论中，我们假设通过对所关心的过程的考察，讨论这个问题所需要的惯性参考系以及相应的坐标系已经建立起来。那么这个时候，我们的运动的轨迹的描述就是一个 $\vec{x}(t)$ 的时间过程。知道在这个过程中任意一个 t 时刻质点的位置 $\vec{x}(t)$ 由哪些因素决定，对于给定的问题计算出 $\vec{x}(t)$ 来，就是我们的理论的目标。通过微积分的学习，我们知道这个 $\vec{x}(t)$ 实际上是空间的一条带参数的曲线，如果我们知道这个曲线在任何时候的切线，我们

就可以把这些切线连（积分）起来得到这条曲线。也就是说 $\vec{x}(t)$ 和 $\dot{\vec{x}}(t)$ 加上 t_0 时刻的原点 $\vec{x}_0 = \vec{x}(t_0)$ 是完全等价的。这个 $\dot{\vec{x}}(t)$ 被称为这个质点的速度，记为 $\vec{v}(t)$。类似地，我们可以定义加速度 $\dot{\vec{v}}(t) = \vec{a}(t)$，以及更高阶的时间导数。给定了任何一阶导数的时间过程加上所有的初始条件完整地描述了这个运动过程。

Newton 力学告诉我们，给定一个力的位置和速度的依赖函数 $\vec{F}(\vec{x},\vec{v})$（由于某些我们这门课程还不能讨论的原因，这个力可以依赖于位置、速度和时间，但是不能依赖于更高阶的时间导数。为了讨论和记号简单，我们进一步假设力不依赖于时间），运动过程由以下方程决定，

$$\ddot{\vec{x}} = \vec{a} = \frac{\vec{F}(\vec{x},\vec{v})}{m}. \tag{5.1}$$

这个方程被称为Newton 第二定律。我们还可以把这个方程改写成一阶封闭的微分方程的形式，

$$\begin{cases} \dot{\vec{x}} = \vec{v}, & (5.2a) \\ \dot{\vec{v}} = \dfrac{\vec{F}(\vec{x},\vec{v})}{m}. & (5.2b) \end{cases}$$

在进一步讨论这个方程之前，我们先来通过几个例子熟悉一下这个方程。

例 5.1 (**自由落体运动**)：质量为 m 的小球，在地球表面的高度为 h 的地方落下来，初始速度为 0，问经过多长时间落地。

先做受力分析：小球受到竖直方向向下的地球的吸引力[1]。为了将来计算方便，我们要规定好坐标系，为了建立坐标系，我们需要选择参考系——这个参考系最好还是一个做匀速直线运动的物体。在这里，相对于我们所关心的问题来说，地球是一个很好的不动的东西——尽管它实际上在自转和公转，因此选为参考系。在这个参考系下，我们规定竖直方向向下为我们的 y 轴正方向，地面为零点。于是，有了这个参考系和坐标系以后，我们就可以把思考转化成数学运算了。于是，

$$\vec{F} = mg\hat{j}.$$

[1]为什么这个用“竖直”而不用“垂直”呢？主要是考虑到将来还有斜面的问题，这个时候，垂直就容易造成垂直于斜面的误解。因此，指向地球中心的方向（由于地球在所讨论的问题的尺度下看来是平的，非常大）就习惯上称作竖直方向。

接着，运用Newton 运动定律，$\vec{F} = m\vec{a}$，得到

$$\vec{a} = \frac{\vec{F}}{m} = g\hat{j}.$$

最后，运用加速度和速度以及位置之间的积分关系以及初始条件，我们得到，

$$y = y_0 + \frac{1}{2}gt^2.$$

其中 $y_0 = -10\,(m)$，于是通过

$$-y_0 = \frac{1}{2}g\tau^2,$$

可以求得落地得时间。

例 5.2 (**经典一维谐振子（连着弹簧的一个小球）**)：质量为 m 的小球，连在一根弹性系数为 k 原长为 l 的水平方向的弹簧上。初始时刻，拉开小球距离为 A_0，求松开以后小球运动的轨迹。

同样，先建立参考系坐标系，选择弹簧得固定的一端做为参考系和坐标零点，从零点指向弹簧的方向 x 轴为正方向。于是，我们得到初始条件，$x_0 = l + A_0$，$v_0 = 0$。弹簧所受到的力在水平方向上是

$$\vec{F} = -k\,(x - l)\,\hat{i}$$

竖直方向的力不管，因为这个小球仅仅在水平方向运动。

接着，运用Newton 运动定律，得到

$$\ddot{x}\hat{i} = \vec{a} = \frac{\vec{F}}{m} = -\frac{k}{m}\,(x - l)\,\hat{i}.$$

变量替换，$\tilde{x} = x - l$，我们得到，

$$\ddot{\tilde{x}} = -\frac{k}{m}\tilde{x}.$$

这个微分方程存在通解，

$$\tilde{x} = A\cos\,(\omega t + \phi_0),$$

其中 $\omega = \sqrt{\frac{k}{m}}$。代入初始条件，我们得到

$$x = A_0\cos\,(\omega t) + l.$$

例 5.3 (**经典一维谐振子（连着弹簧的两个小球）**)：质量为 m 的两个小球，连在一根弹性系数为 k 原长为 l 的水平方向的弹簧上。初始时刻，拉开两个小球距离为 $l+A$，求松开以后两个小球运动的轨迹。

先建立参考系和坐标系并做受力分析。以左边的小球的初始位置为原点，从左到右为 x 轴正方向。运动开始以后，假设两个小球的位置分别是 x_1 和 x_2，则左边的小球受到的力是 $F_{2\to1}=k\left(x_2-x_1-l\right)$，右边的小球受到的力是 $F_{1\to2}=k\left(x_1+l-x_2\right)$。

接着用Newton 运动定律，得到

$$\ddot{x}_1=-\frac{k}{m}\left(x_2-x_1-l\right),$$
$$\ddot{x}_2=-\frac{k}{m}\left(x_1+l-x_2\right).$$

把这两个方程相加和相减，我们得到，

$$\ddot{x}_1+\ddot{x}_2=0,$$
$$\ddot{x}_1-\ddot{x}_2=-2\frac{k}{m}\left(x_1-x_2-l\right). \tag{5.3}$$

这两个方程的解都能够直接写出来。我们就不再继续算下去了。

我们发现，这个弹簧连着两个小球的运动，就好像是整体 x_1+x_2 在匀速运动，然后，相对距离 x_1-x_2，在做简谐振动。

实际上，在第一个例子中的地球也是这个问题的参与者，地球与小球之间的引力相互作用才是导致这个运动的原因。但是，由于小球的运动对地球的状态产生的影响基本可以忽略不计，所以我们把地球作为参考系，用与地球状态无关的方向和大小都不变的重力来代替实际上随着位置变化的引力。第二个例子中，弹簧的另一端，也就是墙，实际上也有运动，但是幅度小很多。于是，我们把墙看作参考系，然后把墙与小球之间的相互作用看成与墙的状态无关的弹簧的弹力。在第三个问题中，我们把墙替换成了另一个小球，这个时候，其运动也成了我们需要考察的对象，两个小球之间的力与两个小球的状态都有关系。从这个角度来看，第一个例子考察的是地球与小球之间的相互作用，第二个例子是小球与墙壁之间的相互作用，第三个是两个小球之间的相互作用。实际上，我们相信，各种各样的力，都是物体与物体之间的相互作用造成的。有的时候能够看见相互作用的各个主体有的时候只能看见一个主体或者一部分主体，是我们在考虑问题的时候根据需要划分的系统的不同造成的。当一个系统之外的外界对系统的影响不是很大的

时候，或者说倒过来，当这个系统对系统的外界影响不大的时候，我们可以把外界的主体忽略，当成一个不变的影响或者直接忽略。

这个由一个质点或者多个质点的位置的集合所构成的空间称为一个系统的位形空间。除了位形空间，我们常常还在相空间里面讨论问题。相空间就是由一个质点或者多个质点的位置加上速度的集合构成的空间。我们可以把公式 (5.1) 看作是位形空间的方程，把公式 (5.2) 看成是相空间的方程。我们知道两个方程是等价的，那为什么要引入相空间呢？因为位形空间的轨道可以相交，而相空间的轨道不能相交。很容易理解为什么位形空间的轨道可以相交，撞车就是这样发生的。那为什么相空间的轨道就不会相交呢？一个相空间的点对应了确定的位置与速度。而速度决定了下一个时刻的位置，因此，这个时刻的相空间中的一个点完全决定了下一个时刻这个点在哪里。于是，如果两个轨道在某一个点相交，则将来这两个轨道完全一样。同样的道理，过去也完全一样。于是他们就是同一条轨道。也可以用下面的例子来说明。考察一个一维系统的两条相空间轨道。一维系统的相空间是二维的，我们来证明有两条不同的相空间轨道不可能在某点相交。

假设有 1, 2 两条轨道在 t_0 时刻相交于相空间的一个点 Q。公式 (5.1) 和公式 (5.2) 都是一个确定性微分方程，在给定初始条件的情况下解是唯一的。它们都需要两个初始条件来确定一个唯一解。在这个交点 Q，我们有两个初始条件，$x_1(t_0) = x_2(t_0) = x_Q, p_1(t_0) = p_2(t_0) = p_Q$。于是，根据唯一性，通过这个交点的轨道只能有一条。因此，不同轨道不能在一点相交，如果两个轨道在相空间相交于一点，实际上，它们必定是相同的轨道。另一方面，给定位形空间的一个点只能给定一个初始条件，$x_1(t_0) = x_2(t_0) = x_Q$，不能唯一确定一个解。因此，存在多条轨道通过这个给定的点。由于相空间的这个好性质，在以后的讨论中，我们都会用相空间。很多时候，速度会替换成动量 $\vec{p} = m\vec{v}$，两者相差一个常数（对于给定的系统，质量是一个常数）。所以相空间也经常指由位置和动量坐标联合确定的空间。

作为力学本身来说，主要概念和主要结构就只有这些，空间、时间、参考系、坐标、位置、速度（动量）、加速度、力、位形空间、相空间，还有Newton运动定律。这些概念描述了运动，把运动与运动变化的原因，也就是力联系起来。但是，什么是力的问题，力学是回答不了的。力学回答的问题是，给定了力，也就是物体之间的相互作用之后，运动将会如何发生。给定的力到底怎样的问题，由研究具体的力的学科来回答，例如电磁学研究电磁力，引力科学研究引力，核物理学研究核力等等。

有一种类型的力称为保守力或者有势力在物理学里面特别重要。我们甚至相信所有的力的背后，如果我们把我们的系统做的足够大，把足够多的外界包含在我们的系统之中，所有的力都应该是保守力。

5.3 势函数、保守力与能量守恒

保守力是指满足以下关系的力 $\vec{F}(\vec{x},\vec{v})$，存在一个函数 $V(\vec{x})$ 满足，

$$\vec{F}(\vec{x},\vec{v}) = -\vec{\nabla}V(\vec{x}). \tag{5.4}$$

其中 $\vec{\nabla}$ 是位置坐标的导数。所以，这样的力，不能是速度 $\vec{v}$ 的函数，而仅仅是位置 $\vec{x}$ 的函数。

前两个例子中的力，就是保守力，也被称为有势力，也就是可以写成一个势函数然后把力看成是这个势函数的梯度的力；后面的例子中的力，一般看成是来自于相互作用，而这个相互作用的也可以写成势函数的形式。

例 5.4 (**势函数举例**) ：前面两个例子中的势函数分别是 $V_1 = -mgx$，$V_2 = \frac{1}{2}k(x-l)^2$。可以非常简单地验证通过这两个势函数能够得到与前面的例子中的力。第三个的相互作用是函数实际上是 $V_3 = \frac{1}{2}k(x_1 - x_2 - l)^2$。这个可以通过计算 $\vec{F}_1 = -\vec{\nabla}_1 V(\vec{x}_1,\vec{x}_2) = -\frac{\partial}{\partial x_1}V_3\hat{i} = -k(x_1 + l - x_2)\hat{i}$。这个和受力分析的结果完全一致。

当方程中所有的力都是保守力的时候，我们来看一下方程公式 (5.1) 的一个变形，

$$\begin{aligned}
m\frac{d}{dt}\vec{v} = -\vec{\nabla}V(\vec{x}) &\Longrightarrow \vec{v}\cdot\frac{d}{dt}\vec{v} = -\vec{v}\cdot\vec{\nabla}V(\vec{x}) \\
\Longrightarrow mv_x\frac{d}{dt}v_x + mv_y\frac{d}{dt}v_y + mv_z\frac{d}{dt}v_z &= -\left(v_x\frac{\partial}{\partial x} + v_y\frac{\partial}{\partial y} + v_z\frac{\partial}{\partial z}\right)V(\vec{x}) \\
\Longrightarrow \frac{1}{2}m\frac{d}{dt}\left(v_x^2 + v_y^2 + v_z^2\right) &= -\left(\frac{dx}{dt}\frac{\partial}{\partial x} + \frac{dx}{dt}\frac{\partial}{\partial y} + \frac{dx}{dt}\frac{\partial}{\partial z}\right)V(\vec{x}) \\
\Longrightarrow \frac{d}{dt}\left(\frac{1}{2}mv^2 + V(\vec{x})\right) &= 0
\end{aligned}$$

也就是

$$\frac{1}{2}mv^2 + V(\vec{x}) = E. \tag{5.5}$$

这里 E 是一个不依赖于时间的常数。这个等式被称为能量守恒，其中第一项被称为动能 T，第二项被称为势能 V。

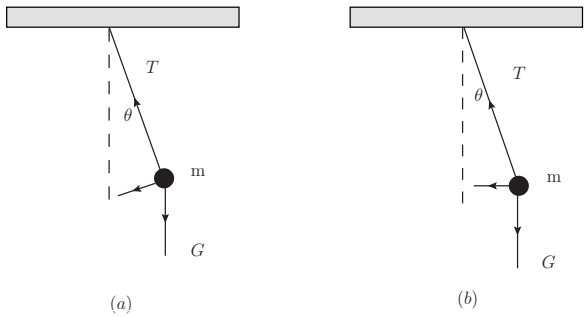


图 5.1: （a）单摆的运动的受力分解；（b）同样的装置放在一个加速运动的汽车里面。

例 5.5 (**能量守恒的应用**) ：利用能量守恒求解第一个例子中小球落地以后的速度，以及第二、第三个例子中弹簧被压缩的最大长度。

这里，仅仅以第二个问题为例。首先这里能量包含动能和弹簧上积累的势能，于是任何一个时刻的总能量是 $E_T = \frac{1}{2}mv^2 + \frac{1}{2}k(x-l)^2$。根据能量守恒，

$$\frac{1}{2}mv^2 + \frac{1}{2}k(x-l)^2 = \frac{1}{2}kA_0^2.$$

其中，我们代入了初始状态的能量。于是，$(x-l)$ 的极值必然发生在 $v=0$ 的时候，

$$\left|(x-l)\right|_{max} = A_0.$$

于是，最大压缩长度是 A_0。

现在，我们已经完成了Newton 力学的主要内容的学习，我们来试着求解一下下面这个问题。

例 5.6 (**单摆的周期**) ：固定在天花板上的轻质绳子连着一个质量为 m 小球，拉开绳子一小段距离后松开，求小球再次回到松开位置的时间。

在用 Newton 力学来求解问题的过程中，最关键的一步是受力分析，也就是合力的计算。考虑一个跟单摆非常类似的问题：把同样的装置放在一个在加速的汽车里面。这两个问题的情境如图 5.1 所示。我们知道按照图 5.1(a) 的分解和合成方式，合力 $F_T = mg\sin(\theta)$，按照图 5.1(b) 的分解和合成方式，合力 $F_T = mg\tan(\theta)$。为什么同样的装置，不同的情形，受力分析不一样呢？实际上，是因为，我们知道（整体或者某个运动的方向上的）合

力的方向肯定是加速度的方向，而加速度的方向，在速度为零的时候，肯定就是下一个时刻运动的方向。于是，由于我们对于这两种情况下的运动有一个预期，(a) 下一个时刻不会在径向运动只能在切线方向运动，于是按照切线和径向方向来分解；(b) 会沿着水平方向运动，于是沿着水平方向和竖直方向分解。一旦做出来这个分解，那么，按照 Newton 第二定律，我们得到

$$mg\sin(\theta) = -ml\ddot{\theta} \Rightarrow \ddot{\theta} = -\frac{g}{l}\sin(\theta) \propto -\frac{g}{l}\theta. \tag{5.6}$$

最后一步是一个 θ 很小的时候的近似。

通过这个例子，我们发现，做受力分析需要对运动形式有一定的前瞻性，洞察力：最终合力的方向与可能产生的运动方向的相互关联。如果我们再考虑更复杂一点的平面摆，多级摆，这个受力分析就会更困难。一个复杂的机器可能由很多部分构成。这个时候，受力分析的方法基本上就不能用了。Lagrange 和 Hamilton 发明了一个不用做受力分析的求解力学问题的方法——分析力学。这里我们仅要介绍 Hamiltonian 力学。

5.4 从 Newtonian 力学到 Hamiltonian 力学

Hamiltonian 力学，也称为分析力学的 Hamiltonian 形式，解决的问题是，只要我们能够写下动能 T 和势能 V 作为自变量 $\vec{x}$ 以及 $\vec{p}$ 的函数，那么，所有的力学问题都可以直接变成求解微分方程的技术问题。我们先来看看逻辑上为什么可行，然后举几个例子学会如何运用，最后再讨论这个方法的意义。

5.4.1 单个质点的分析力学形式

对单个质点的守保守力作用的系统定义 Hamiltonian，

$$H = T(\vec{x},\vec{p}) + V(\vec{x},\vec{p}) = \frac{1}{2m}(\vec{p})^2 + V(\vec{q}), \tag{5.7}$$

同时记位置坐标 $\vec{x}$ 为 $\vec{q}$，把自变量 $\dot{\vec{x}}$ 替换成 $\frac{\vec{p}}{m}$. Newton 方程与下列方程等价，

$$\begin{cases} \dfrac{d}{dt}q_j = \dfrac{\partial H}{\partial p_j}, & (5.8a) \\ \dfrac{d}{dt}p_j = -\dfrac{\partial H}{\partial q_j}. & (5.8b) \end{cases}$$

其中 x_j 是 $\vec{x}$ 的 j 分量的坐标。这个方程被称为 Hamilton 方程。

对于单个质点的系统，我们可以简单验证这个方程确实给出 Newton 方程。

$$\begin{cases} \dfrac{d}{dt}q_j = \dfrac{\partial H}{\partial p_j} \Longrightarrow \dot{q}_j = \dfrac{p_j}{m}, & (5.9a) \\ \dfrac{d}{dt}p_j = -\dfrac{\partial H}{\partial q_j} \Longrightarrow \dot{p}_j = -\dfrac{\partial}{\partial x_j}V = F_j. & (5.9b) \end{cases}$$

于是 $ma = m\ddot{x}_j = m\ddot{q}_j = F_j$。

例 5.7 (**谐振子的** Hamiltonian **力学**) ： $V = \frac{1}{2}k(q-l)^2$，从 Hamiltonian 推导出其运动方程。

知道势能，我们可以写下来 $H = \frac{p^2}{2m} + \frac{1}{2}k(q-l)^2$，于是方程就成了，

$$\frac{d}{dt}q = \frac{\partial H}{\partial p} = \frac{p}{m} \tag{5.10}$$

$$\frac{d}{dt}p = -\frac{\partial H}{\partial q} = k(q-l) \tag{5.11}$$

把第二式代入第一式，我们得到

$$\frac{d^2}{dt^2}q = \frac{\dot{p}}{m} = \frac{k}{m}(q-l). \tag{5.12}$$

这个方程与通过 Newton 力学得到的完全一致。

例 5.8 (**单摆的** Hamiltonian **力学**) ： 固定在天花板上的轻质绳子连着一个质量为 m 小球，写出这个小球的运动方程。

对于单摆，我们可以写下来 $H = \frac{p^2}{2m} - mgl\cos(\theta)$，但是这个有一个小小的问题，我们用了角度当作自变量，于是将来计算偏导数的时候需要协调一下。第一个处理的方式是这样，利用 $q = l\theta$，

$$l\frac{d}{dt}\theta = \frac{d}{dt}q = \frac{\partial H}{\partial p} = \frac{p}{m} \tag{5.13}$$

$$\frac{d}{dt}p = -\frac{\partial H}{\partial q} = -\frac{1}{l}\frac{\partial H}{\partial \theta} = -mg\sin(\theta) \tag{5.14}$$

于是把第二式代入第一式，我们得到

$$\frac{d^2}{dt^2}\theta = \frac{\dot{p}}{lm} = -\frac{g}{l}\sin(\theta). \tag{5.15}$$

这个方程与通过 Newton 力学得到的完全一致。

另外一个方式是理解角动量的定义以后，直接用角度变量 θ，角速度 $\dot{\theta}$ 和角动量 $p_\theta \triangleq \vec{r} \otimes m\vec{v} = ml^2\dot{\theta}$，或者定义式 $p_\theta \triangleq I\dot{\theta} = ml^2\dot{\theta}$。其中 $I = ml^2$ 叫做转动惯量。于是通过 $H = \frac{1}{2}ml^2\dot{\theta}^2 - mgl\cos(\theta)$，得到

$$H = \frac{1}{2}ml^2\left(\frac{p_\theta}{ml^2}\right)^2 - mgl\cos(\theta) = \frac{p_\theta^2}{2ml^2} - mgl\cos(\theta) \tag{5.16}$$

于是通过 Hamilton 方程得到

$$\frac{d}{dt}\theta = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{ml^2} \tag{5.17}$$

$$\frac{d}{dt}p_\theta = -\frac{\partial H}{\partial \theta} = -mgl\sin(\theta) \tag{5.18}$$

于是把第二式代入第一式，我们得到

$$\frac{d^2}{dt^2}\theta = \frac{\dot{p}_\theta}{ml^2} = -\frac{g}{l}\sin(\theta). \tag{5.19}$$

这个方程与通过 Newton 力学得到的完全一致。

在这个例子中，无论是角动量的处理方式，还是把角度通过半径变成长度以后再利用一般的动量（称为线动量）来计算，得到结果是一样的。

更一般的，如果了解一点点 Lagrangian 力学[2]，任意一个坐标对应着的角动量，可以通过 $p = \frac{\partial L}{\partial \dot{q}}$ 来定义。这里 q 是广义坐标，可以是通常的直角坐标的分量值，也可以是角度，等等。而一般情况下，通过 q 和 $\dot{q}$ 直接写出 Lagrangian 比较简单。例如，对于我们的单摆的例子，

$$L = T - V = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos(\theta), \tag{5.20}$$

于是

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = ml^2\dot{\theta}. \tag{5.21}$$

与我们通过角动量的定义得到的关系一致。然后，通过 Hamiltonian 和 Lagrangian 之间的如下关系——这个关系的证明需要用到Legendre 变换[3]，请参考 Landau 的《力学》[27]——来得到 Hamiltonian，

$$H = \dot{q}(p,q)\,p - L(q,\dot{q}(p,q)), \tag{5.22}$$

[2]这一段的内容需要关于 Langangian 力学的知识。详情请参阅 Landau 的《力学》[27]

[3]一般的数学变化在改变自变量的时候，要么维持函数形式不变，要么维持函数值不变。但是，Legendre的这个Legendre 变换是一个神奇的变换，自变量函数值函数形式都变化，但是维持推导出来的运动方程不变。

其中 $\dot{q}(p,q)$ 需要从 $p=\frac{\partial L}{\partial \dot{q}}$ 中反解出来。例如，对于上面的谐振子的例子，我们有

$$\dot{\theta}=\frac{p_\theta}{ml^2}. \tag{5.23}$$

于是，

$$H=\frac{p_\theta}{ml^2}p_\theta-\frac{1}{2}ml^2\left(\frac{p_\theta}{ml^2}\right)^2-mgl\cos(\theta)=\frac{p_\theta^2}{2ml^2}-mgl\cos(\theta). \tag{5.24}$$

在这里一定要注意 L 的自变量是 $(q,\dot{q})$，而 H 的自变量是 (q,p)。

通常从 Lagrangian 开始，而不是直接从 Hamiltonian 开始的原因是，一般情况下以坐标和坐标的导数，也就是速度，为自变量写下 T 和 V，要比以坐标和动量为自变量写下 T 和 V 更容易。更多的细节请参看 Landau 的《力学》[27]。

5.4.2 从单体到多体的分析力学形式

我们花了额外的本来用于学习量子力学的时间学习了 Hamilton 方程，不仅仅是因为它们可以从受力分析中解脱出来，还因为，它们的形式非常容易推广到多体相互作用的系统。一个质点的动力学完全取决于 $V(\vec{x})$，而这个 $V(\vec{x})$，我们说过，也是由于和隐藏起来的不在系统之内的**外界**物体相互作用产生的。那么，当有多个质点，而且这个多个质点之间存在相互作用的时候，我们的理论的这个形式还能不能用呢？这里，所谓相互作用就是从受力分析的角度，一个质点受到的力不仅仅和这个质点的位置相关，也跟其他的质点的位置相关。从势能的角度来说，就是系统的势能同时依赖于多个质点的位置，而不仅仅是一个依赖于单个质点的位置的势能函数的和。用数学表达式来说明就是，

$$V(\vec{x}_1,\vec{x}_2)\neq V_1(\vec{x}_1)+V_2(\vec{x}_2) \tag{5.25}$$

从 Newton 力学来说，单个质点的系统和多个质点的系统没有本质区别，但是受力分析会复杂很多很多，任意两个质点之间的受力都需要考虑，然后最后叠加起来。幸好，分析力学的 Hamilton 形式，可以直接从单个质点推广到多个质点，仅仅需要包含一个相互作用项。以两个质点的系统为例

$$H(\vec{q}_1,\vec{q}_2;\vec{p}_1,\vec{p}_2)=\frac{\vec{p}_1^2}{2m_1}+\frac{\vec{p}_2^2}{2m_2}+V_1(\vec{q}_1)+V_2(\vec{q}_2)+V_{12}(\vec{q}_1,\vec{q}_2). \tag{5.26}$$

这个方便的形式在处理多体相互作用的系统的时候非常有用。其中 $V_{12}(\vec{q}_1,\vec{q}_2)$ 就是相互作用项。这一项的取值同时依赖于这个系统内两个质点的位置，而不是一个。

例 5.9 (**一根弹簧连着的两个小球的** Hamiltonian **力学**) ： 一根水平方向的弹簧连着的两个小球，从 Hamiltonian 开始，求出运动方程，并做比较。

首先，我们写下来这个系统的 Hamiltonian，

$$H=\frac{p_1^2}{2m}+\frac{p_2^2}{2m}+\frac{1}{2}k\left(q_1-q2-l\right)^2$$

然后，带入 Hamilton 方程，得到

$$\begin{aligned}\dot{x}_1&=\frac{p_1}{m},\\ \dot{x}_2&=\frac{p_2}{m},\\ \dot{p}_1&=-k\left(x_1+l-x_2\right),\\ \dot{p}_2&=-k\left(x_2-x_1-l\right).\end{aligned}\tag{5.27}$$

这个方程和之前得到的方程完全一致。

例 5.10 (**二级摆的** Hamiltonian **力学**) ： 这个问题会用到 Lagrangian 力学和 Hamiltonian 力学的关系。固定在天花板上的轻质硬杆连着另一个轻质硬杆，第二个轻质硬杆连着一个质量为 m 小球。假设这个系统就在一个平面内运动（拉开小球到一定位置以后释放，这样避免由于初始速度的问题，我们需要考虑在三维空间中的运动），写出这个小球的运动方程。

轻质表示质量可以忽略。硬杆表示长度不可伸长。以天花板上的轻杆固定点做为零点，建立参考系坐标系，向下为 y 轴正方向，向右为 x 轴正方向。选取自变量为第一个杆和竖直方向的逆时针方向夹角 α 和第二个杆和竖直方向的逆时针方向夹角 β。我们先想办法写下来动能和势能。

$$\begin{aligned}x&=l\sin(\alpha)+l\sin(\beta),\\ y&=l\cos(\alpha)+l\cos(\beta).\end{aligned}$$

于是，

$$\begin{aligned}V_x&=l\cos(\alpha)\dot{\alpha}+l\cos(\beta)\dot{\beta},\\ V_y&=-l\sin(\alpha)\dot{\alpha}-l\sin(\beta)\dot{\beta}.\end{aligned}$$

于是，

$$E_T=\frac{1}{2}m\left(V_x^2+V_y^2\right)=\frac{1}{2}ml^2\left(\dot{\alpha}^2+\dot{\beta}^2+2\cos(\alpha-\beta)\dot{\alpha}\dot{\beta}\right),$$

$$E_P=-mgl\cos(\alpha)-mgl\cos(\beta).$$

于是，

$$L=\frac{1}{2}ml^2\left(\dot{\alpha}^2+\dot{\beta}^2+2\cos(\alpha-\beta)\dot{\alpha}\dot{\beta}\right)+mgl\cos(\alpha)+mgl\cos(\beta).$$

借助于 Lagrange 方程，这个 L 可以直接用来推导出来运动方程。这里，我们希望回到 Hamilton 方程，于是我们想求出来以动量和位置为自变量的 H。按照定义，

$$\begin{aligned}p_\alpha&=\frac{\partial L}{\partial\dot{\alpha}}=\frac{1}{2}ml^2\left(2\dot{\alpha}+2\cos(\alpha-\beta)\dot{\beta}\right),\\p_\beta&=\frac{\partial L}{\partial\dot{\beta}}=\frac{1}{2}ml^2\left(2\dot{\beta}+2\cos(\alpha-\beta)\dot{\alpha}\right).\end{aligned}$$

然后，从这里，我们反解出来 $\dot{\alpha},\dot{\beta}$ 做为 p_α,p_β 的函数，然后代入到

$$H=\dot{\alpha}p_\alpha+\dot{\beta}p_\beta-L$$

就可以得到 H。最后通过 Hamilton 方程得到运动方程。具体的计算这里就省略了。

这个问题，如果你打算用受力分析来做，是非常非常麻烦的。如果需要包含杆的质量，以及杆的长度的变化，我们仍然可以通过直接写下 L 来做，而那个时候，通过受力分析来做，就几乎没有希望了。

分析力学的形式还能够帮助我们简化约束的处理，例如，考虑一个在曲面或者曲线上运动的物体。这里我们就不再讨论，有兴趣的读者可以去看一下 Dirac 的约束运动 [38]。

当然，除了分析更加简单了，分析力学形式的好处主要在于，现在，只要给定位置（x）和动量（p）做为自变量的 $H=H(x,p)$，或者位置（x）和速度（$\dot{x}$）做为自变量的 $L=L(x,\dot{x})$，那么，系统的一切行为都确定了。在这个意义上，我们说，这样的 $H(x,p)$ 或者这样的 $L(x,\dot{x})$ 是一个系统的完整的描述。注意，一般来说，$p=\frac{\partial}{\partial\dot{x}}L(x,\dot{x})\neq m\dot{x}$。

5.5 作业

习题 5.1 N 个耦合谐振子系统，$L=\sum_{j=0}^{N}\left[\frac{m}{2}\dot{q}_j^2-\frac{k}{2}(q_j-q_{j+1})^2\right]$，其中 $q_0=q_{N+1}=0$。写出 H，还有运动方程。对于 $N=2$，找出独立变量使得所有的方程只依赖于自身的变量。求出来的这些“独立变量”被称为这个系统的正则动量和正则坐标。

习题 5.2 一跟硬杆连着的两个小球：两个小球通过一根硬杆（不能伸长和收缩）相连，可以在整个三维空间中运动，写下这个系统的运动方程。

习题 5.3 三跟硬杆连着的三个小球：三个小球通过三根硬杆（不能伸长和收缩）首尾相连，放在一个光滑平面上，写下这个系统的运动方程。

习题 5.4 三根弹簧连着的三个小球：三个小球通过三根弹簧首尾相连，放在一个光滑平面上，写下这个系统的运动方程。

5.6 本章小结

通过力学的学习，我们已经了解到物理学的基本思考方式是描述事物的状态，以及探索事物状态发生变化的原因。Newton 力学关心的是物体的时空状态——在任何时间点上物体在什么位置、速度是多少，以及状态发生变化的原因。利用力的概念，Newton 力学把原因与状态联系起来，也就是 Newton 第二定律。我们还了解到这个问题完全由物体的 Hamiltonian 描述，从其中我们可以得到运动方程，然后在一定的初始条件下求解运动方程我们可以得到物体任意时刻的状态，不管过去还是未来。最后，我们还了解到运用受力分析和 Newton 第二定律来处理相互作用的多体系统以及受约束的系统是很困难的，但是 Hamilton 力学使得这个问题更容易处理。

写下了一个经典系统的 Hamiltonian 就是知道了这个系统的所有信息，配上初始条件就可以得到这个系统在相空间的轨道。

实际上，除了我们已经了解的经典力学的 Newton 力学形式、Hamilton 力学形式，我们还有 Lagrange 力学形式。一方面，它们都是等价的。另一方面，我们还可以找出一个更加基本的原理，从这个原理来推出这些力学形式，而且都可以认为是从这个更加基本的原理得到的结果。这样的一个更加基本的原理就叫做最小作用量原理。关于这个部分更详细的推导请参考 Landau 的《力学》[27]。追求更加统一的方式来描述自然是物理学的终极目标之一。

最后提醒一个技术上要注意的细节：确定一个系统需要多少个变量来描述，是一个非常重要的步骤。不同的选择有可能会非常大地影响分析计算的复杂度。实际上，建立任何一个数学模型的第一步都是描述变量的选择。

第三部分

单粒子系统的量子力学

我们把量子力学——也就是量子系统的数学模型——的主题内容分成单体的量子力学和耦合的量子力学两个部分。其实，其理论完全就是一致的——完全就是单体系统的量子力学。或者说，任意维数的量子系统，我们都可以当作一个单体系统来研究。我们这里所做的区分仅仅是因为所计算分析的量上的差别，而不是计算分析所用的理论或者工具的不同。在耦合系统里面，我们关心纠缠态、纠缠态的形成、纠缠态的测量等问题。

这个部分的推荐阅读材料是：Feynman的《Feynman 物理学讲义》第三卷 [7]、Ballentine的《Quantum Mechanics – a modern development》[8]、喀兴林的《高等量子力学》[28]。

第六章　量子系统状态的数学模型

这一章我们先构建描述量子系统的给定状态的数学模型。这个数学模型必须能够对于任意的物理量的测量，给出与实验结果相符的理论推断。下一章，我们再来讨论状态的演化的问题。

6.1　有关量子状态和测量的公理

公理 1 (**量子态公理**) 量子系统的状态由 Hilbert 空间上的密度矩阵 ρ 表示，或者在特殊情况下，由 Hilbert 空间 V 的归一化的矢量 $|\mu\rangle$ 表示。

密度矩阵是 Hilbert 空间上的正定归一的 Hermitian 算符，满足

$$\rho^{\dagger} = \rho, tr(\rho) = 1, \tag{6.1}$$

并且其本征值 p_j 满足，

$$0 \leq p_j \leq 1. \tag{6.2}$$

当且仅当其中一个 $p_\mu = 1$（此时其他的本征值都是 0），

$$\rho = |\mu\rangle\langle\mu|. \tag{6.3}$$

这个时候也可以认为量子态的表示就是 $|\mu\rangle$。密度矩阵的形式更一般，当有多个非零本征值的时候表示混合态。

区分纯态和混合态在以后的讨论中会有用。这里介绍一下纯态的概念。测量得到一个系统的密度矩阵之后，我们把比较特殊的，满足

$$\rho = |\psi\rangle\langle\psi|, \tag{6.4}$$

的密度矩阵称为纯态，不满足公式 (6.4) 的状态称为混合态。也就是说纯态就是上面提到的密度矩阵有且仅有一个非零本征值的特殊情况。可以证明，纯态满足

$$tr\left(\rho^2\right) = tr\left(\rho\right) = 1, \tag{6.5}$$

而混合态必然有

$$tr\left(\rho^2\right) < 1. \tag{6.6}$$

Hilbert 空间矢量存在加法、数乘、内积的操作。按照这个公理这些操作也成了量子系统的状态上的相应操作。至于各个操作对于量子系统的意义，例如什么情况下我们用两个矢量的加法 $a\left|\mu\right\rangle + b\left|\nu\right\rangle$，于是相当于 $a^*a\left|\mu\right\rangle\left\langle\mu\right| + a^*b\left|\nu\right\rangle\left\langle\mu\right| + ab^*\left|\mu\right\rangle\left\langle\nu\right| + b^*b\left|\nu\right\rangle\left\langle\nu\right|$，什么情况下我们用两个密度矩阵的加法 $p_1\rho_1 + p_2\rho_2$，于是大概相当于 $a^*a\left|\mu\right\rangle\left\langle\mu\right| + b^*b\left|\nu\right\rangle\left\langle\nu\right|$，我们还需要以后进一步来了解。在这里先给出一个形式上的说明，具体含义到将来的具体实验中会更清楚。

当一个状态可能由两个（也可以是多个，可以简单推广）方式发生的时候，如果这两个方式不可区分——也就是我们根本不可能了解到底哪一种情况发生了的时候，在量子状态上，我们采用矢量叠加，

$$\rho = a_1^*a_1\left|\psi_1\right\rangle\left\langle\psi_1\right| + a_1^*a^2\left|\psi_2\right\rangle\left\langle\psi_1\right| + a_1a_2^*\left|\psi_1\right\rangle\left\langle\psi_2\right| + a_2^*a_2\left|\psi_2\right\rangle\left\langle\psi_2\right|, \tag{6.7}$$

相当于

$$\left|\psi\right\rangle = a_1\left|\psi_1\right\rangle + a_2\left|\psi_2\right\rangle. \tag{6.8}$$

其中，我们假设方式一对应着状态 $\left|\psi_1\right\rangle$（$\rho_1 = \left|\psi_1\right\rangle\left\langle\psi_1\right|$），方式二对应着状态 $\left|\psi_2\right\rangle$（$\rho_2 = \left|\psi_2\right\rangle\left\langle\psi_2\right|$）。当这两种发生的方式可以区分的时候——后面在具体实验中我们会更清楚地解释什么是可区分什么是不可区分[1]，我们用概率性叠加，

$$\rho = p_1\left|\psi_1\right\rangle\left\langle\psi_1\right| + p_2\left|\psi_2\right\rangle\left\langle\psi_2\right|. \tag{6.9}$$

将来我们会看到，在量子力学里面，矢量叠加和概率性叠加的区别和适用条件是非常重要的。

公理 2 (量子观测量公理) 量子系统的可观测量由 Hilbert 空间的 Hermitian 算符矢量 O 表示。

[1]或者说，任何时候都当作不可区分，于是都在用矢量叠加，直到遇到测量于是需要做部分迹的时候：这个时候，从矢量叠加数学形式的密度矩阵开始，经过部分迹运算，合适的条件下自然就会得到概率性叠加的密度矩阵。

是不是所有的 Hermitian 算符矢量都可以成为量子系统的可观测量是另一个问题。理论上应该可以，但是实际测量过程中，满足要求的仪器不一定能够设计的出来。Hilbert 空间的线性 Hermitian 算符存在算符的加法、数乘，线性变换等操作。按照这个公理，这些操作也成了对量子系统的可观测量的相应的操作。这些操作 $\hat{O}_1 + \hat{O}_2$ 对于量子系统的含义我们也需要进一步来了解。

公理 3 (**量子测量公理**) 对于给定状态 ρ，测量可观测量 O 的所有可能结果是 O 的本征值 o 的集合，每一个状态 o 出现的几率满足

$$P_o = \langle o|\rho|o\rangle . \tag{6.10}$$

因此，测量的平均值满足

$$\langle O\rangle = \sum_o oP_o = tr\,(O\rho) \tag{6.11}$$

其中我们用到了 $\sum_o oP_o = \sum_o o\,\langle o|\rho|o\rangle = \sum_o \langle o|O\rho|o\rangle$，$\langle o|$ 以及 $|o\rangle$ 是 Hermitian 算符 O 的本征向量。对于 $\rho = |\mu\rangle\langle\mu|$ 的特殊情况，$\langle O\rangle = \langle\mu|O|\mu\rangle$。

公理 4 (**量子测量后状态公理**) 对于给定状态 ρ，测量可观测量 O 并观测到状态 o^*，则测量以后，系统的状态成为特定的观测到的状态 $\rho = |o^*\rangle\langle o^*|$。

测量以后的状态就是所测量到的状态。这一点，在经典力学的世界里头不用特别指出：这本来就是测量的含义——测量就是测得系统当时的状态，因此系统在测量之后的当时，必然就是测的的结果反映的状态。在量子的世界里面，这一点不是那么平庸。我们把这个公理单独拿出来。通常的教科书把它与上一个公理合在一起。

关于这一条公理，还有值得注意的地方。在实际测量过程中，对一个系统的可观测量 O 的测量总是通过引入另外一个叫做测量仪器的系统 A，有的时候这个仪器可以是同一个量子系统的不同的自由度——实际上相当于不同的系统，来实现的。这个时候，记录下来的往往是测量仪器 A 的状态 α^*，当然这个 α^* 与被测量的物理量的取值 o^* 的对应关系是测量之前就已经知道的知识。所以，实际的测量过程还牵涉到如何在 A 和 O 之间建立这样的关联。这个问题，在这里我们不讨论。仅仅当作实现公理3和公理4的背后的暂时不需要关心的一种机制。

下面，我们用这四条来解释上面提到的量子力学的实验。所谓解释就是我们给出一张配方（菜谱），在实验的每一个阶段，你会知道对状态进行物

理操作应该对应着什么样的数学操作，当时的状态是什么，测量得到的结果是什么，而且这样的配方对所有的量子系统都适用。至于理解这个配方，是以后的事情。为了和经典概率论相区别，这里我们也会尝试用经典概率论来给出一个量子行为的配方。我们可以把经典概率论写成以下这个和上面的量子理论非常相像的形式。这并不意味着我们认为物理世界存在两种完全不同的数学描述——经典和量子描述，那样的话问题就成了哪些系统应该用经典描述哪些用量子描述，两者的边界在哪里，系统的什么特征使得其更加适合其中的哪一种来描述。同样的，前面这句附加的说明也不意味着我们认为系统本质上就是量子的——那样的话，我们还需要回答其在什么样的特殊情况下表现为经典的。关于这个世界到底是经典还是量子的问题我们在 6.4 节和 6.5 节还会再一次讨论。在这里，我们只是形式上把经典随机客体的理论独立提出来，并整理成和量子公理形式上一致的形式，然后尝试用这个理论来描述量子系统的行为。

6.2 经典随机客体理论的公理形式

在这里，我们的经典概率论局限在有限的离散随机变量的范畴之内。实际上，下面这个体系也可以拓展到可数无限的甚至连续取值的随机变量上面去。

公理 5 (**经典状态公理**) 经典系统的状态由事件集合上的密度矩阵 $\rho = \sum_\mu p_\mu |\mu\rangle\langle\mu|$ 表示。

经典事件以及经典密度矩阵可以通过 $a\rho_1 + b\rho_2$ 的形式相加，但是不能通过 $a|\mu\rangle + b|\nu\rangle$ 形式相加。如果允许，就会出现 $|\mu\rangle\langle\nu|$ 这样的项，而经典密度矩阵不能出现这样的项。

公理 6 (**经典可观测量公理**) 经典系统的可观测量由事件集合上的对称算符矢量 $O = \sum_\mu O_\mu |\mu\rangle\langle\mu|$ 表示。

算符的加法和数乘是指这样的操作 $\hat{O}_1 + \hat{O}_2 = \sum_\mu (O_{1,\mu} + O_{2,\mu}) |\mu\rangle\langle\mu|$。其中这里的基矢和密度矩阵的基矢都对应着基本简单事件 μ。也就是说，基矢变换在这里没有意义。算符和密度矩阵完全对角，而且存在自然的基矢集合。

公理 7 (经典测量公理) 对于给定状态 ρ，测量可观测量 O 的所有可能结果是 O 的本征值 o 的集合，每一个状态 o 出现的几率满足

$$P_o = \langle o|\rho|o\rangle . \tag{6.12}$$

因此，测量的平均值满足

$$\langle O\rangle = \sum_o oP_o = tr\left(O\rho\right). \tag{6.13}$$

公理 8 (经典测量后状态公理) 对于给定状态 ρ，测量可观测量 O 并观测到状态 o^*，则测量以后，系统的状态成为特定的观测到的状态记为 $\rho = |o^*\rangle\langle o^*|$。

我们看到经典概率论与量子理论的形式完全一致，唯一的区别就是量子理论允许矢量相加（$|\mu\rangle = a|\alpha\rangle + b|\beta\rangle$），于是写成密度矩阵的形式时候（$\rho = |\mu\rangle\langle\mu| = a^*a|\alpha\rangle\langle\beta| + a^*b|\beta\rangle\langle\alpha| + ab^*|\alpha\rangle\langle\beta| + b^*b|\beta\rangle\langle\beta|$）存在非对角元（$a^*b|\beta\rangle\langle\alpha| + ab^*|\alpha\rangle\langle\beta|$），而经典密度矩阵只存在对角元素（$\rho = p_\alpha|\alpha\rangle\langle\beta| + p_\beta|\beta\rangle\langle\beta|$）。

为了简化记忆，所有的这些公理，唯一需要记住的公式就是

$$\langle O\rangle = tr\left(O\rho\right), \tag{6.14}$$

不管经典还是量子。从这个公式（公式 (6.14)）可以得到公式 (6.10) 和相应的公式 (6.12)。取可观测量 $\hat{P}_o = |o\rangle\langle o|$，其中 $|o\rangle\langle o|$ 是原来的可观测量的所有的可能取值的一个，于是，

$$P_o = \left\langle \hat{P}_o \right\rangle = tr\left(|o\rangle\langle o|\rho\right) = \langle o|\rho|o\rangle . \tag{6.15}$$

也就是说，无论经典概率论还是量子力学，整个静态描述的问题，完全就是一个公式——公式 (6.14)。当然，两者还是有区别的，再强调一次：量子力学的理论建立在矢量空间的基础上，因此允许矢量加法；经典概率论的理论建立在事件集合的基础上，因此只允许概率叠加，没有矢量加法。前者导致一般来说密度矩阵存在非对角元，后者导致密度矩阵完全对角，而且对角的表象就是把简单事件作为基矢。因此，在经典系统中，我们必然候有

$$\langle\mu|\,\nu\rangle = \delta_{\mu\nu}, \tag{6.16}$$

而在量子系统中，我们可能有

$$|\langle\mu|\,\nu\rangle| \in [0,1] . \tag{6.17}$$

有了这些数学结构和以上的数学和物理之间关系公理，我们就可以来讨论如何用量子理论来描述量子系统的行为了。

6.3 量子静态理论用于描述测量实验

我们首先来看一下自旋的测量的实验。看看经典概率论和量子理论能否描述这个实验。然后我们会讨论光子 which-way 实验。光子过玻璃、双缝干涉、三个偏振片的魔术，这几个实验就留给读者来解释了。

6.3.1 自旋的测量和再测量

实验装置如图 6.1(a) 所示，让一个自旋先通过一个 z 方向的磁场，然后挡住 z 方向向下状态的出口，仅仅让 z 方向向上状态出射。然后让射出的自旋接着再通过 z 方向的磁场。测量这个时候有几个输出。实验结果我们已经知道，只有一个输出，就是向上。我们来构造这个实验的理论，也就是回答每一个时刻的状态是什么，测量的物理量是什么，测量结果是什么的问题。

我们先来给这个系统选择一个 Hilbert 空间。将来我们会回答为什么这样选择。基于量子力学公理，针对自旋的情况，整个状态和测量描述的配方是这样的：对于某个方向的磁场 $\hat{r}$ 出来的自旋翻转的仪器，我们写下来一个算符

$$\sigma_r = \hat{r} \cdot \vec{\sigma}, \tag{6.18}$$

其中

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_x = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{6.19}$$

从这个仪器的向上（向下）方向出来的自旋处于这个算符的向上（向下）的本征态 $|\uparrow_r\rangle$（$|\downarrow_r\rangle$），如果从两个方向出来的自旋用某种方式不可区分并且合起来，那么，状态将是

$$|\psi\rangle = \alpha|\uparrow_r\rangle + \beta|\downarrow_r\rangle \tag{6.20}$$

其中 α,β 如何确定留待以后再说。这就确定了这个系统的 Hilbert 空间：所有的这样的 $|\psi\rangle$ 的集合。

有了状态 $|\psi\rangle$ 以后，我们这样来构建测量：考虑一个测量的仪器其内部磁场方向为 $\hat{r}_0$，先得到算符 σ_{r_0}，并且得到相应的本征向量 $\left|\uparrow_{r_0}\right\rangle$ 和 $\left|\downarrow_{r_0}\right\rangle$，然后测量结果和测量后状态按照公理3和公理4确定。

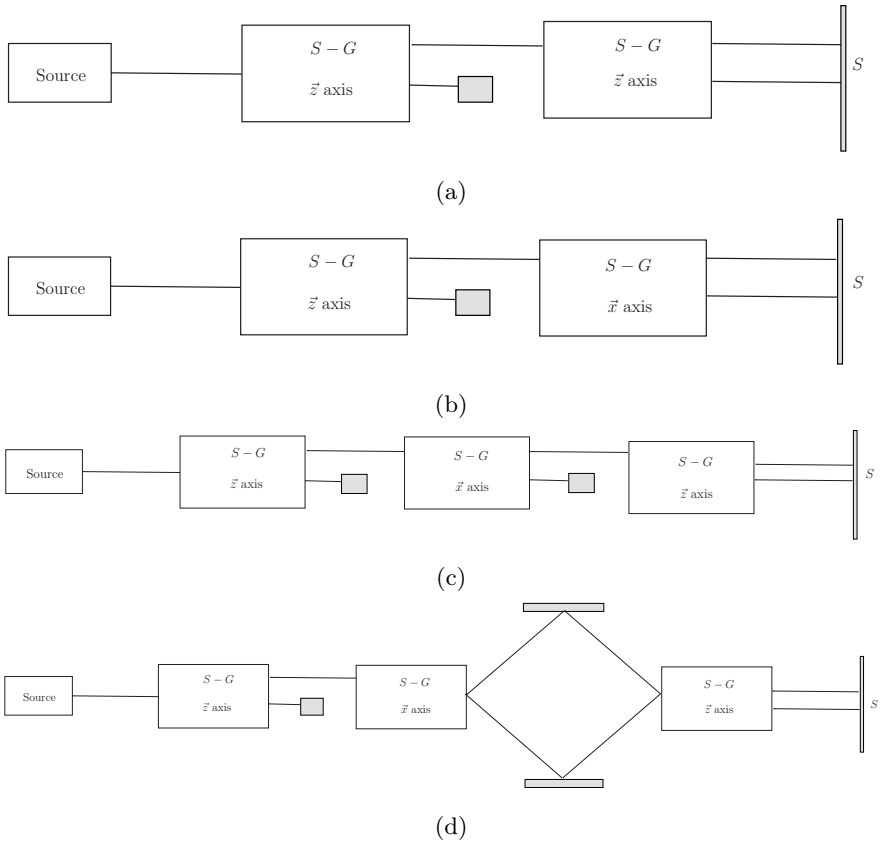


图 6.1: (a) 自旋经过一个 z 方向 Stern-Gerlach 装置之后挡住向下的输出，让向上的输出再一次经过 z 方向的装置。(b) 自旋经过一个 z 方向 Stern-Gerlach 装置之后挡住向下的输出，让向上的输出经过 x 方向的装置。(c) 自旋经过一个 z 方向 Stern-Gerlach 装置之后挡住向下的输出，让向上的输出经过 x 方向的装置，挡住向下的输出，让向上的输出进入 z 方向的装置。(d) 自旋经过一个 z 方向 Stern-Gerlach 装置之后挡住向下的输出，让向上的输出经过 x 方向的装置，“反射”之后合起来，再次进入 z 方向的装置。

现在，我们先按照这个配方写下状态的数学表达式，然后验证其是否正确。

通过第一个磁场之前的状态未知，我们不写了。通过第一个装置之后向下方向被挡住，仅留下向上方向的自旋作进一步实验。这个时候自旋的状态是

$$\rho_1 = |\uparrow_z\rangle\langle\uparrow_z|, \tag{6.21}$$

第二个装置代表所测量的物理量是

$$\sigma_z = |\uparrow_z\rangle\langle\uparrow_z| - |\downarrow_z\rangle\langle\downarrow_z|. \tag{6.22}$$

其本征向量是 $|\uparrow_z\rangle$ 和 $|\downarrow_z\rangle$。于是按照我们的理论，测量的结果是，

$$\begin{cases} p_{\uparrow_z} = \langle\uparrow_z|\rho_1|\uparrow_z\rangle = 1 \\ p_{\downarrow_z} = \langle\downarrow_z|\rho_1|\downarrow_z\rangle = 0 \end{cases}. \tag{6.23}$$

这个结果与实验结果一致——只有一个输出，向上。

这表明，在这个实验上，我们的理论能够解释实验。但是，这里的密度矩阵和可观测量都是对角的，因此，这个理论既可能是量子的，也可能是经典的。也就是说，这个实验可以用经典的，也可以用量子的理论来解释。下面，我们再来看一个稍微有一点差别的实验。

让一个自旋先通过一个 z 方向的磁场，然后挡住 z 方向向下状态的出口，仅仅让 z 方向向上状态出射。然后让射出的自旋接着再通过一个 x 方向的磁场。测量这个时候有几个输出。实验结果我们已经知道，有两个输出。我们来构造这个实验的理论，也就是回答每一个时刻的状态是什么，测量的物理量是什么，测量结果是什么的问题。

注意，经过第一个装置以后的状态，不应该随着第二个装置的改变而改变。这个是一个很重要的逻辑假定。因为第二个装置可以是在很远很远的将来很远很远的地方做的测量，逻辑上很难接受，自旋经过第一个装置的状态取决于很远的将来在很远的地方发生的事情。这个叫做因果律。一般情况下，物理学需要尊重因果律。

于是，

$$\rho_1 = |\uparrow_z\rangle\langle\uparrow_z|, \tag{6.24}$$

第二个装置代表所测量的物理量是

$$\sigma_x = |\uparrow_x\rangle\langle\uparrow_x| - |\downarrow_x\rangle\langle\downarrow_x|. \tag{6.25}$$

为了得到与实验一致的理论结果，我们正好有

$$\begin{cases} p_{\uparrow_x} = \langle\uparrow_x|\rho_1|\uparrow_x\rangle = \langle\uparrow_x|\uparrow_z\rangle\langle\uparrow_z|\uparrow_x\rangle = \frac{1}{2} \\ p_{\downarrow_x} = \langle\downarrow_x|\rho_1|\downarrow_x\rangle = \langle\downarrow_x|\uparrow_z\rangle\langle\uparrow_z|\downarrow_x\rangle = \frac{1}{2} \end{cases} . \tag{6.26}$$

这个结果用到了 $\langle\uparrow_x|\uparrow_z\rangle = \frac{\sqrt{2}}{2}$ 这个两个矢量之间的内积关系[2]。

$\langle\uparrow_x|$ 和 $|\uparrow_z\rangle$ 分别是来自于 σ_x 算符和 σ_z 算符的本征向量，我们还需要这两个向量之间的内积的模方处于 $[0,1]$ 之间。这一点仅仅在相互不对易的算符之间能够实现。于是，我们发现，简单的各个物理量都是数（于是对易）的经典理论不能描述量子系统的行为，必须是物理量是算符（于是不对易）的量子理论才可以。当然，更加复杂的经典理论的可能性我们暂时不讨论，可见第二章和第十一章。

现在，我们来进一步讨论下一个更加复杂的自旋实验。我们挡住上一个实验中经过第二个装置以后的 x 方向向下的自旋，仅仅让 x 方向向上的自旋出去。这个时候，我们再让这个出射的自旋经过一个 z 方向的磁场，问有几个输出。实验结果是有两个输出。我们来看，我们的理论应该如何构造。

首先，按照我们的配方，这个先经过 z 方向向上再从 x 方向向上出射的自旋与直接从从 x 方向向上出射的自旋的状态是一样的。在这里，所谓一样的就是，无论做什么样的测量，结果都一样。也就是说，这个状态都是

$$\rho_2 = |\uparrow_x\rangle\langle\uparrow_x| . \tag{6.27}$$

接着，这个状态需要做第三个装置也就是 σ_z 的测量，于是正好

$$\begin{cases} p_{\uparrow_z} = \langle\uparrow_z|\rho_2|\uparrow_z\rangle = \langle\uparrow_z|\uparrow_x\rangle\langle\uparrow_x|\uparrow_z\rangle = \frac{1}{2} \\ p_{\downarrow_z} = \langle\downarrow_z|\rho_2|\downarrow_z\rangle = \langle\downarrow_z|\uparrow_x\rangle\langle\uparrow_x|\downarrow_z\rangle = \frac{1}{2} \end{cases} , \tag{6.28}$$

和实验相符。到这里为止，这个实验的量子力学解释看起来好像和前面一个没有什么区别，没有额外费劲。

[2]我们已经说过，经典概率论的任意简单事件需要满足 $\langle\mu|\nu\rangle = \delta_{\mu\nu}$，因此这个结果就表明经典概率论是不能描述量子系统的行为的。当然，实际上更一般的经典概率论还是有可能的，例如尽管 $|\uparrow_z\rangle$ 代表简单事件但是 $|\uparrow_x\rangle$ 代表了复合事件。简单事件和复合事件之间的内积就可以允许各种可能了。对于这样的更一般的经典理论用来描述量子系统的行为的可能性的讨论可见第二章和第十一章。在这里我们就不展开讨论了。简单地说，一个把 $|\uparrow_z\rangle$ 看作简单事件把 $|\uparrow_x\rangle$ 看作复合事件的理论是有缺陷的。这两个状态之间应该是完全对称的，其差别仅仅是把仪器转动了一个方向而已。例如，我们可以让第一个装置取 x 方向，第二个装置取 z 方向。那样的话，我们的经典理论就必须把 $|\uparrow_x\rangle$ 看成简单事件，把 $|\uparrow_z\rangle$ 看作复合事件。很难想象，这样的理论是可行的。不过，实在有必要这样的理论也是允许的。

但是，注意，我们来看一下这个最终结果：一开始 z 方向向下的成分被挡住了，去掉了，然后经过某个操作（实际上是过 x 方向以后挡住向下的成分），z 方向向下的成分又回来了。这个结果，在经典世界里面，就好像是说一个盒子的红球和黑球，首先去掉所有红球，接着去掉这些球里面玻璃做的留下塑料的，然后，红球又出现了。这是一件多么不可思议的事情啊。但是，你看，我们的量子理论可以描述这件事情，只要我们允许 $\langle\uparrow_z|\uparrow_x\rangle = \frac{\sqrt{2}}{2} \neq 0 \neq 1$。

现在，我们来进一步讨论下一个更加更加复杂的自旋实验。我们挡住把上一个实验中经过第二个装置以后的 x 方向向上和向下的自旋用磁场转弯以后合起来[3]，然后让合起来的自旋经过一个 z 方向的磁场，问有几个输出。实验结果是有一个输出。我们来看，我们的理论应该如何构造。

首先，按照我们的配方，这个先经过 z 方向向上的自旋状态是，

$$\rho_1 = |\uparrow_z\rangle\langle\uparrow_z|. \tag{6.29}$$

这个自旋再分裂成两个可能的 x 方向的自旋本征态，

$$\rho_1 = |\uparrow_z\rangle\langle\uparrow_z| = \frac{|\uparrow_x\rangle + |\downarrow_x\rangle}{\sqrt{2}}\frac{\langle\uparrow_x| + \langle\downarrow_x|}{\sqrt{2}}. \tag{6.30}$$

也就是说，相当于这样两个矢量的叠加，

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}|\uparrow_x\rangle + \frac{1}{\sqrt{2}}|\downarrow_x\rangle. \tag{6.31}$$

在这个实验的下一步，一组仪器使得这样分开的两个 x 方向的自旋状态又重新合起来。对于最后一步进入 z 方向磁场并且从里面出来的自旋来说，我们根本不可能区分它是走上面的路，也就是 $|\uparrow_x\rangle$ 来的，还是走下面的路，也就是 $|\downarrow_x\rangle$ 来的。

这个时候就有了一个如何把两条路径的状态合起来的问题。在这里，我们先暂时通过补充一个“公理”的形式来解决怎么合起来的问题，但是注意，将来我们会发现这个“公理”并不独立于之前的量子力学公理之外。因此，也就不是真的需要这一条公理。只不过，从前面的公理“推导”出来这一条公理需要用到后面有关纠缠态和测量的知识，见 6.3.5 节和 6.4 节。于是，我们先拿来当公理用。

[3]实际如何实现，这个实验比之前的几个实验要复杂，但是也是能够实现的。磁场可以使带电粒子转弯。只要这个自旋在一个带电粒子上就可以设计出来这样的仪器。当然，实际真的设计及和做出来这样的仪器是另外的问题。真实的实验实现还是见实验 (1.7) 光子版本的这个 which-way 实验。

关于不可区分状态的临时公理：当一个系统存在由于多种不可区分的方式导致的不可区分的状态的时候，这个系统处于这些不可区分状态的矢量叠加态；如果两个状态可以区分，则处于这两个状态的概率叠加态。

以两个不可区分状态为例，就是，系统可能处于以下两个不可区分状态，$\rho_1 = |\psi_1\rangle\langle\psi_1|$ 和 $\rho_2 = |\psi_2\rangle\langle\psi_2|$，则系统处于

$$\rho = \frac{1}{N}|\alpha\psi_1 + \beta\psi_2\rangle\langle\alpha\psi_1 + \beta\psi_2|, \tag{6.32}$$

其中 α,β 是代表两个状态所占振幅比例的系数，N 是归一化常数。当两者所占你振幅比例完全一样的时候，我们取 $\alpha = \frac{\sqrt{2}}{2} = \beta$，于是 $N = 1$。如果两个状态可区分，则系统处于

$$\rho = p_1\rho_1 + p_2\rho_1. \tag{6.33}$$

其中 p_1, p_2 代表两个状态出现的可能性。注意，在这里，我们没有明确解释什么是“可区分”什么是“不可区分”。这个等到在 6.4 节以后从前面的公理推导出来这一条临时公理的时候再来讨论。

按照我们加了临时公理的配方，对于不可区分的状态，我们要做矢量叠加，于是我们重新得到，

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}|\uparrow_x\rangle + \frac{1}{\sqrt{2}}|\downarrow_x\rangle. \tag{6.34}$$

也就是说，

$$\rho_2 = \frac{|\uparrow_x\rangle + |\downarrow_x\rangle}{\sqrt{2}}\frac{\langle\uparrow_x| + \langle\downarrow_x|}{\sqrt{2}} = |\uparrow_z\rangle\langle\uparrow_z|. \tag{6.35}$$

接着，这个状态需要做 σ_z（它是下一个装置）的测量，于是

$$\begin{cases} p_{\uparrow_z} = \langle\uparrow_z|\rho_2|\uparrow_z\rangle = 1 \\ p_{\downarrow_z} = \langle\downarrow_z|\rho_2|\downarrow_z\rangle = 0 \end{cases}. \tag{6.36}$$

因此，只有一个输出。现在我们已经用基于量子力学公理的配方解释了第一章中的所有的关于量子自旋的实验。我们发现这个配方能够解释所有的实验结果。同时，我们也再一次提醒，经典力学不能解释这一系列实验。

注意，从这几个实验的解释我们看到量子力学区别与经典力学的基本特征是，在量子情形取任意两个状态 μ,ν，$\langle\mu|\nu\rangle = \delta_{\mu\nu}$ 不一定满足，而经典满足。

6.3.2 为什么自旋算符要用 $\hat{r}\cdot\vec{\sigma}$?

在这一节的最后，我们来回答为什么自旋算符要用 $\hat{r}\cdot\vec{\sigma}$，为什么我们要把 ρ_1、ρ_2 写成这个样子。在这里，我们假设通过用经典概率论来描述量子系统的行为的尝试，我们已经知道了，量子系统的行为必须用算符的形式来描述。我们仅仅来解释为什么算符和密度矩阵要这样写。

首先，我们说一个任意的 $\hat{r}$ 方向的自旋算符在自己的本征矢量当做基矢的坐标系下面，我们肯定有

$$\sigma_{\hat{r}} = |\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}| - |\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|\,. \tag{6.37}$$

这个表示对于任意方向的测量，我们得到两个结果——$\hat{r}$ 方向向上或者 $\hat{r}$ 方向向下的输出。为了能够区分这样两个方向的输出，我们记其中一个方向的结果为 $+1$（本征值），另一个为 -1（本征值）。于是，按照这个记号，一个 $\hat{r}$ 方向向上的输出（也就是观察到可观测量的值为 $+1$ 的时候的测量后状态）的状态应该记作，

$$\rho = |\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}|\,. \tag{6.38}$$

于是剩下的问题就成了按照实验结果来定义合适的 $\langle\uparrow_z|\uparrow_{\hat{r}}\rangle$,$\langle\downarrow_z|\uparrow_{\hat{r}}\rangle$ 了。假设我们知道了这两个数，那么按照完全性关系，我们就可以得到，

$$|\uparrow_{\hat{r}}\rangle = |\uparrow_z\rangle\langle\uparrow_z|\uparrow_{\hat{r}}\rangle + |\downarrow_z\rangle\langle\downarrow_z|\uparrow_{\hat{r}}\rangle\,. \tag{6.39}$$

我们可以先尝试一下 $\langle\uparrow_z|\uparrow_{\hat{r}}\rangle = 0,1$ 这样的行不行。我们发现这个和实验不相符，取 $\hat{r}=\hat{x}$ 为例，明明一个 x 方向的本征态和可以从一个 z 方向的本征态里面出现，以及反过来。进一步实验我们甚至可以发现这些内积的数值，例如 $|\langle\uparrow_z|\uparrow_x\rangle|^2 = \frac{1}{2} = |\langle\downarrow_z|\uparrow_x\rangle|^2$。于是，取这些概率幅的模（复数的相位角部分该取多少以后再说）我们得到，

$$|\uparrow_x\rangle = \frac{\sqrt{2}}{2}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right). \tag{6.40}$$

我们直接在二维空间里面根据这个矢量来求得一个跟它完全正交的矢量

$$|\downarrow_x\rangle = \frac{\sqrt{2}}{2}\left(|\uparrow_z\rangle - |\downarrow_z\rangle\right). \tag{6.41}$$

于是

$$\sigma_x = |\uparrow_x\rangle\langle\uparrow_x| - |\downarrow_x\rangle\langle\downarrow_x| = |\uparrow_z\rangle\langle\downarrow_z| + |\downarrow_z\rangle\langle\uparrow_z|\,. \tag{6.42}$$

同样的道理，通过实验，我们可以把任意的 $\sigma_{\hat{r}}$ 表达成为 $|\uparrow_z\rangle, |\downarrow_z\rangle$ 的形式。于是，我们就得到了 σ_z 表象下的任意 $\sigma_{\hat{r}}$ 自旋算符的形式。

关于从概率得到复数相位角的问题，实际上可以做一个循环来尝试，例如可以假设 z 方向和 x 方向本征矢量内积的复数相位角，然后再假设 z 方向和 y 方向本征矢量内积的复数相位角，于是原则上，就有了 x 方向和 y 方向本征矢量内积的复数相位角。而这个推导出来的 x 方向和 y 方向本征矢量内积的复数相位角是否正确是可以实验检验的：让通过 x 方向磁场的自旋接着通过 y 方向就可以。更具体的情况就不展开计算了。另外，这里的“为什么”是在采用量子力学的数学模型的前提下的为什么，也就是说，假设物理量是算符之后，问自旋算符为什么这样，而不是更基本的量子系统的数学模型为什么必须是算符和矢量的那个为什么。

在这里，为了更加熟悉这些算符，我们把自旋算符的本征态的问题当做例题做一下。

例 6.1 (**自旋算符的表象**)：已知 $\sigma_{\hat{r}} = \vec{\sigma}\cdot\hat{r} = \sigma_x \sin(\theta)\cos(\phi) + \sigma_y \sin(\theta)\sin(\phi) + \sigma_z \cos(\theta)$，其中 $\sigma_x, \sigma_y, \sigma_z$ 为 Pauli 矩阵。求解矢量 $|\uparrow_{\hat{r}}\rangle$ 和 $|\downarrow_{\hat{r}}\rangle$ 在 σ_z 表象下的形式。$\hat{r}$ 方向由球坐标的两个角度 θ 和 ϕ 描述。

按照上面 $\sigma_{\hat{r}}$ 的定义，以及 Pauli 矩阵的 σ_z 表象下的表达式，我们有，

$$\sigma_{\hat{r}} = \begin{bmatrix} \cos(\theta) & \sin(\theta)e^{-i\phi} \\ \sin(\theta)e^{i\phi} & -\cos(\theta) \end{bmatrix}. \tag{6.43}$$

于是，得到本征向量（本征值肯定是 ±1），

$$|\uparrow_{\hat{r}}\rangle = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right)e^{i\phi} \end{bmatrix}, \tag{6.44a}$$

$$|\downarrow_{\hat{r}}\rangle = \begin{bmatrix} \sin\left(\frac{\theta}{2}\right) \\ -\cos\left(\frac{\theta}{2}\right)e^{i\phi} \end{bmatrix}. \tag{6.44b}$$

运用抽象矢量记号，我们可以把上面的矢量形式写作，

$$|\uparrow_{\hat{r}}\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow_z\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|\downarrow_z\rangle, \tag{6.45a}$$

$$|\downarrow_{\hat{r}}\rangle = \sin\left(\frac{\theta}{2}\right)|\uparrow_z\rangle - \cos\left(\frac{\theta}{2}\right)e^{i\phi}|\downarrow_z\rangle. \tag{6.45b}$$

可以验证 $\sigma_r = |\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}| - |\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|$。

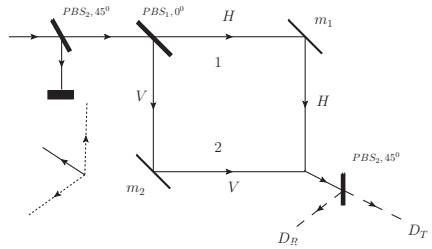


图 6.2: 光子的 which-way 实验装置示意图

顺便，在一般的物理书里面，在自旋物理量算符和这个 Pauli 矩阵之间存在着一个转换常数，$S_{\hat{r}} = \frac{\hbar}{2}\sigma_{\hat{r}}$。这个跟Planck 常数联系在一起的常数 $\hbar = \frac{h}{2\pi}$ 不会造成本征向量的改变，因此所计算出来的所有的测量结果的概率都是一样的，仅仅对本征值做了重新标度。实际上，我们应该称 $\hat{S}_{\hat{r}}$ 为自旋算符而不是 $\sigma_{\hat{r}}$。不过，在这里，我们大多数时候不再区分它们。

6.3.3 光子的 which-way 实验

我们再来看一下光子 which-way 实验，如图 6.2。同样地，描述一个实验我们需要描述任意时刻的状态，所测量的可观测量，测量的结果的所有可能性和相应的几率。

关于光子的偏振，我们的配方是这样的。我们假定光子传播方向是 z 方向，则定义算符[4]，

$$\begin{aligned}\hat{P}^{\theta} = &\left|H\cos(\theta) + V\sin(\theta)\right\rangle\left\langle H\cos(\theta) + V\sin(\theta)\right| \\ &-\left|H\sin(\theta) - V\cos(\theta)\right\rangle\left\langle H\sin(\theta) - V\cos(\theta)\right|. \end{aligned} \tag{6.46}$$

其中 θ 是在 $x-y$ 平面内和 x 方向的一个夹角，H，V 分别是水平和竖直方向的偏振。这个夹角的含义稍后实验中会看到。然后，经过一个内部方向为 θ 方向的偏振片以后，透射出去的光的状态是这个算符的一个本征态 $|H_\theta\rangle = \left|H\cos(\theta) + V\sin(\theta)\right\rangle$，反射出去的光的状态是另一个本征态 $|V_\theta\rangle = \left|H\sin(\theta) - V\cos(\theta)\right\rangle$。形式上，这两个本征态就相当于自旋算符里面的 $|\uparrow_r\rangle$ 和 $|\downarrow_r\rangle$。这一步确定了通过仪器以后光子的状态。下一步我们给出描

[4]实际上偏振片实现了光路方向和偏振的耦合。这个物理过程的真正的算符形式比下面的算符复杂一些，参考第十章。

述测量结果的配方。对于给定的状态 ρ，经过一个 $\hat{P}^{\phi}$ 的仪器以后，观测得到 $\left|H_{\phi}\right\rangle$ 态的几率是，

$$P^{\phi}=\left\langle H_{\phi}\right|\rho\left|H_{\phi}\right\rangle, \tag{6.47}$$

并且测量后状态如果观测得到结果就是 $\left|H_{\phi}\right\rangle$ 的话，就是 $\left|H_{\phi}\right\rangle$ 状态。如果光子不知道从两个分别对应着本征态 $\left|H_{\phi}\right\rangle$ 和 $\left|V_{\phi}\right\rangle$ 的路径的哪一个路径过来，则需要做两个状态的矢量叠加。注意，这一部分的配方，实际上是完全类比之前的自旋部分的配方写下来的。

我们来看以上的配方如何用于解释光子的 which-way 实验。经过第一个透过 45^0 偏振光的偏振分束器以后（之前的我们不关心）的光子的状态为 45^0 偏振，于是相应的算符

$$\hat{P}^{45^0}=\frac{1}{2}\left(|H+V\rangle\langle H+V|-|H-V\rangle\langle H-V|\right). \tag{6.48}$$

这个算符（仪器）的透射光状态是

$$\rho_1=\left|45^0\right\rangle\left\langle 45^0\right|=\frac{1}{2}\left(|H+V\rangle\langle H+V|\right). \tag{6.49}$$

如果我们在 0^0 的偏振分束镜后面，也就是 m_1 和 m_2 的地方分别放两个探测器，测量水平 H 和竖直 V 方向的偏振分量。我们来看一下，按照我们的理论，我们会得到什么结果。我们的可观测量这时候是

$$O_1=\hat{P}^{0^0}=|H\rangle\langle H|-|V\rangle\langle V|, \tag{6.50}$$

这里本征值 ±1 代表屏幕的荧光点：水平方向（这个时候是光路方向，不是偏振方向，记为 h，偏振方向为 H）过来的光达到屏幕上记为 $+1$，竖直方向（记为 v，偏振方向为 V）过来的光达到屏幕上记为 -1。按照公理3，$P_h=\langle H|\rho_1|H\rangle, P_v=\langle V|\rho_1|V\rangle$，我们有测量结果的概率分布，

$$\rho_1^c=\frac{1}{2}|h\rangle\langle h|+\frac{1}{2}|v\rangle\langle v|=\frac{1}{2}|+\rangle\langle+|+\frac{1}{2}|-\rangle\langle-|, \tag{6.51}$$

也就是说水平方向、竖直方向收到光子的概率都是 $\frac{1}{2}$，或者说屏幕上记录一个 $+1$ 或者 -1 的可能性都是 $\frac{1}{2}$。这个预测与真实的实验结果完全吻合。实际上，这里量子态的测量的结果最后是通过粒子到底走哪一条路径于是到达哪一个探测器来确定的。在我们现在的数学形式中，我们特意避免了讨论光子的偏振是如何和光子的路径纠缠起来的。稍后我们就会回到这个问题。实际上，以后，我们还会多次回到这个问题。

现在，我们用同样的配方，来看一下在第二个偏振分束镜做的测量。由于经过第一个分束镜以后，两个光路被反射镜又完全合到了一起不再能够区分，我们知道这个时候要重新做矢量叠加，于是

$$\psi_2 = \frac{\sqrt{2}}{2}\left(|H\rangle + |V\rangle\right), \tag{6.52}$$

也就是，

$$\rho_2 = \frac{1}{2}\left(|H+V\rangle\langle H+V|\right). \tag{6.53}$$

光子的偏振状态完全与入射第一个分束镜的时候相同。第二个分束镜是按照 45^0 来分开光束，也就是

$$O_2 = \hat{P}^{45^0} = \frac{1}{2}|H+V\rangle\langle H+V| - \frac{1}{2}|H-V\rangle\langle H-V|. \tag{6.54}$$

于是，我们得到可能观测结果的几率分布，

$$\rho_2^c = 1\cdot|h\rangle\langle h| + 0\cdot|v\rangle\langle v| = 1\cdot|+\rangle\langle +| + 0\cdot|-\rangle\langle -|, \tag{6.55}$$

也就是说探测器 D_T 所在的透射方向、探测器 D_R 所在的反射方向收到光子的概率分别是 1、0。这个结果也与实验结果完全吻合。

6.3.4 为什么偏振算符取公式 (6.46) 的形式？

下面，我们来解释为什么偏振算符取公式 (6.46) 的形式（也就解释了为什么 $|H+V\rangle$ 代表 45^0 的方向，$|H-V\rangle$ 代表 135^0 的方向），为什么重新合起来的光子状态是 ρ_2。同样的，这个解释不是为了理解，而是配方层次的解释。

我们先规定 $|H\rangle$ 代表 0^0 的方向，$|V\rangle$ 代表 90^0 的方向的偏振光。那么 45^0 的偏振光如何表示呢？按照我们的公理1，所有的各个方向的偏振状态的集合是一个 Hilbert 空间，而且这个空间还是两维的：原因是无论哪里一个方向的偏振光过来，经过偏振分束镜以后只有两种可能——要么走水平方向要么走竖直方向。因此，$\left|45^0\right\rangle$ 必然能够表达成 $|H\rangle, |V\rangle$ 的线性叠加，也就是说，

$$\left|45^0\right\rangle = \alpha|H\rangle + \beta|V\rangle = \cos\theta|H\rangle + \sin\theta e^{i\phi}|V\rangle, \tag{6.56}$$

其中我们用归一化条件找出了独立变量 θ, ϕ。对于高维的系统，这样的独立变量会更难找。对于这个状态，我们重复前面 0^0 和 90^0 偏振的测量。如果

我们只关心这两个状态出现的几率，则，理论上，在仅仅考虑这个 0^0 和 90^0 偏振两个事件的几率的前提下（其他方向的状态的测量暂时不考虑），观测得到的所有可能值的分布函数是，去掉 $\left|45^0\right\rangle\left\langle 45^0\right|$ 中的非对角元，仅仅保留对角元，

$$\rho_{45^0}^c = \cos^2\theta\,|h\rangle\langle h| + \sin^2\theta\,|v\rangle\langle v|\,. \tag{6.57}$$

进一步通过实验我们发现这里的两个方向偏振光束的出射光是等几率的，于是

$$\cos^2\theta = \sin^2\theta = \frac{1}{2}. \tag{6.58}$$

于是，

$$\left|45^0\right\rangle = \frac{\sqrt{2}}{2}\left(|H\rangle + |V\rangle\right). \tag{6.59}$$

在这里其实还有到底是 $\theta = 45^0$ 还是 $\theta = 135^0$ 的问题，还有 ϕ 的取值的问题。前者人为规定一个就可以，不如直接就取成和已知的 45^0 相同。否则，也不过就是语言换一下凡是写在前面的角度的 θ，实际上代入到公式 (6.46) 的时候，都要用 $\theta - 90^0$。ϕ 的确定更加复杂一些，需要考虑多组偏振片的组合，就像确定自旋算符的形式的过程中的那个复数相位值一样。我们暂时采取最简单的 $\phi = 0$。

有了 45^0 偏振态的矢量表达式，再加上 135^0 偏振态的表达式，一个 45^0 方向偏振片的算符形式不过就是这两个本征态的叠加，于是，

$$\hat{P}^{45^0} = \frac{1}{2}\left(|H\rangle + |V\rangle\right)\left(\langle H| + \langle V|\right) - \frac{1}{2}\left(|H\rangle - |V\rangle\right)\left(\langle H| - \langle V|\right) \tag{6.60}$$

类似地，我们可以通过实验加上以上的推导得到所有的 $\hat{P}_\theta$ 的算符形式。

用以上配方，我们可以类似地解释双缝干涉实验，以及光过玻璃的实验。再一次强调，这里所谓的解释，其实是一个预测量子系统状态的机器，一个数学模型：从一个最开始的状态出发（这个状态本身在给定实验条件的情况下有固定的规则写下来），测量的结果是什么（可观测量原则上由给定的实验仪器与步骤确定），测量以后的状态是什么；然后测量结果与测量以后的状态能够通过进一步的实验检验。

物理学的理论，可以认为是对世界如何运行的描述，甚至有的时候可以回答为什么的问题。但是，其基本任务是回答是什么，怎么样的问题。也就是说，如果有一个理论，对于给定的实验条件、实验仪器、实验步骤的系统的实验结果能够给出正确的计算结果，那么这个理论就是这个现象的好的

理论。当然，一个更好的理论还要求能够解释更多的同类的现象。但是，只要能够达到以上的目标，一个理论就能够建立起来。

在量子力学的理论中，路径积分量子化的物理图像就与我们这里的正则量子化的图像有所不同。一个好的物理学家，对于核心的现象，头脑中常常是多于一个理论的，而且还要能够在这些不同形式的理论之间自由的转换。关于路径积分量子力学的图景可以参阅 [39]。

6.3.5 自旋和光子 which-way 实验解释的补充：纠缠

进入下一部分量子系统的演化之前，最后的问题：在自旋 which-way 实验中，为什么 $\rho_1 = (|\uparrow_z\rangle\langle\uparrow_z|)$ 经过分束镜，然后经过反射镜，合回来的光子状态还是 $\rho_2 = (|\uparrow_z\rangle\langle\uparrow_z|)$；在光子 which-way 实验中，为什么 $\rho_1 = \frac{1}{2}(|H+V\rangle\langle H+V|)$ 经过分束镜，然后经过反射镜，合回来的光子状态还是 $\rho_2 = \frac{1}{2}(|H+V\rangle\langle H+V|)$。这两个问题一样，我们用光子的实验为例。这个问题也就是关于矢量叠加态的临时公理的内容。因此，这一小节的讨论相当于企图给这个临时公理找一个逻辑上的基础。这一小节和下一节6.4合起来是对这个临时公理比较完整的说明或者说逻辑推导。

考虑光子的偏振和光子走的光路构成的整个系统，我们称先走水平方向然后反射回到竖直方向的光路为光路 1，先走竖直方向然后反射回到水平方向的光路为光路 2。在光子进入偏振分束镜之前，整个系统的状态为

$$\rho_0^{P,T} = \frac{1}{2}(|H+V\rangle\langle H+V|)\otimes|0\rangle\langle 0|, \tag{6.61}$$

其中 P,T 分别表示偏振（Polarization）和路径（Trojectory），0 表示在光子经过分束镜之前的光路。此时它不区分光路 1, 2。经过分束镜进入反射镜之前，

$$\rho_1^{P,T} = \frac{1}{2}(|H\otimes 1+V\otimes 2\rangle\langle H\otimes 1+V\otimes 2|). \tag{6.62}$$

这表示偏振为 H（V）的光子走的是路径 1（2）。经过反射镜之后到达最后的分束镜之前，

$$\begin{aligned}\rho_2^{P,T} &= \frac{1}{2}(|H\otimes 0+V\otimes 0\rangle\langle H\otimes 0+V\otimes 0|)\\ &= \frac{1}{2}(|H+V\rangle\langle H+V|)\otimes|0\rangle\langle 0|.\end{aligned} \tag{6.63}$$

这表示偏振为 H 和 V 的光子走的路径重新不可区分（关于这个不可区分性下一节6.4有更详细的解释。这里表示路径重新合起来之后，到底从哪一条

路径过来的已经不可能知道了）。对于这个状态，如果我们只关心偏振状态而不是光子的路径的话，那么这个系统的状态就跟如下状态等价

$$tr^T\left(\rho_2^{P,T}\right)=\frac{1}{2}\left(|H+V\rangle\langle H+V|\right). \tag{6.64}$$

我们得到的状态正好是前面的 ρ_2。这里，我们用了部分迹的运算 $tr^T(\cdot)$，它表示对变量 T 的自由度求和。当我们需要计算的可观测量只与 P 的自由度有关的时候，按照公理3，

$$\left\langle O^P\right\rangle=tr\left(O^P\rho^{P,T}\right)=tr^P\left(O^P tr^T\left(\rho^{P,T}\right)\right). \tag{6.65}$$

因此，$tr^T(\rho^{P,T})$ 就好像是这种情况下系统在 P 自由度下的状态，称为约化密度矩阵。在经典概率论中，对应的数学概念就是约化密度分布函数。

注意，如果我们直接对公式 (6.62) 做部分迹的计算，则得到

$$tr^T\left(\rho_1^{P,T}\right)=\frac{1}{2}\left(|H\rangle\langle H|+|V\rangle\langle V|\right). \tag{6.66}$$

这是一个混合态，概率性混合。这个和对公式 (6.63) 做部分迹得到的结果也就是公式 (6.64)——这里还是纯态还是矢量叠加——不一样。其根本原因是公式 (6.63) 里面包含了偏振自由度和路径自由度的纠缠在一起的信息，而公式 (6.62) 里面两者已经独立。这里独立的含义是给定一个两个随机变量（可以更一般更多个随机变量）的密度分布函数，满足如下的独立事件概率相乘的性质，

$$\rho^{12}=tr^2\left(\rho^{12}\right)tr^1\left(\rho^{12}\right). \tag{6.67}$$

光子经过 0^0 偏振分束器以后偏振自由度和路径自由度纠缠在了一起，但是经过反射，这两个自由度又独立出来了，于是，经过部分迹之后，$\rho_2=(|\uparrow_z\rangle\langle\uparrow_z|)=\rho_1$。自旋的情况与此相同。

6.4 粒子走哪一条路径和概率叠加原理

现在，我们来回答第一章当中提出的一个问题：为了描述量子系统的行为，我们不能同时保留“粒子走哪一条路径”和“概率叠加原理”，那么我们保留什么，去掉什么。在我们这一章中已经建立的量子理论告诉我们，首先，我们需要用“量子事件叠加原理”（更多的时候被称作“量子矢量叠加原理”或者“态叠加原理”）来代替“概率叠加原理”。现在，我们再来看一

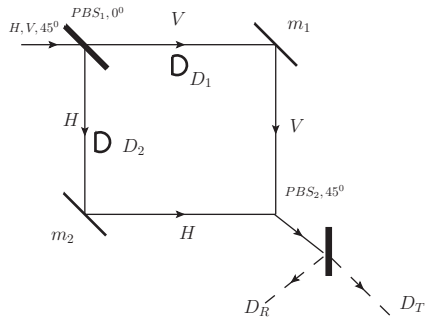


图 6.3: 加了探测器的光子的 which-way 实验装置示意图

下，是不是我们就可以留着“粒子走哪一条路径”了呢，是不是“概率叠加原理”在量子力学就不能用了呢？讨论清楚了这个问题，我们也就明白了如何从前面的公理来推导出来关于不可区分状态的叠加这一条临时公理。

关于不可区分状态的叠加的临时公理已经说了对于实验中不能区分的两条路径，我们用“量子事件叠加原理”；对于实验中能够区分的两条路径，我们用“概率叠加原理”。也就是说，粒子走一条路径或者走另一条路径的说法仅仅在“概率叠加原理”成立的情形下有意义，对于必须要用到“量子事件叠加原理”的情形，因为我们不再能够区分粒子到底走哪一条路，我们不在采用“粒子走一条路径或者走另一条路径”的说法，也不再问是否“粒子同时走两条路”，而是简单粗暴地把两条路径所代表的状态直接按照“量子事件叠加原理”加起来。也就是说，当我们能够问和回答“粒子走哪一条路径”的时候，我们需要用“概率叠加原理”，否则，当我们从根本上就不能回答“粒子走哪一条路径”的时候，我们需要用“量子事件叠加原理”。那么，为什么这样？能够回答“粒子走哪一条路径”是什么意思。这个和纠缠有关。

下面我们采用如图 6.3 加了探测器的光子的 which-way[5]来给这两个叠加原理做一个比较，给“能否回答粒子走哪一条路径”做一个说明。

首先，我们考虑入射光的偏振状态

$$\rho_1^P = \left|45^0\right\rangle\left\langle 45^0\right| = \frac{1}{2}\left(|H+V\rangle\langle H+V|\right). \tag{6.68}$$

[5]实际上这个探测器可以通过在光路的后面增加一个叫做量子消除器（Quantum Erasor）的偏振片来实现 [40]。

加上光路状态以后，我们有

$$\rho_1^{P,T} = \frac{1}{2}\left(|H+V\rangle\langle H+V|\right)\otimes|0\rangle\langle 0|. \tag{6.69}$$

其中 $|0\rangle\langle 0|$ 表示进入第一个偏振分束器之前的光路。接着，我们考虑经过第一个偏振分束器之后的状态，

$$\rho_2^{P,T} = \frac{1}{2}\left(|H\otimes 1+V\otimes 2\rangle\langle H\otimes 1+V\otimes 2|\right). \tag{6.70}$$

下一步，我们考虑增加探测器的效果，

$$\begin{aligned}\rho_{2D}^{P,T} = \frac{1}{2}&\left(|H\otimes 1\otimes D_1+V\otimes 2\otimes D_2\rangle\langle H\otimes 1\otimes D_1+V\otimes 2\otimes D_2|\right)\\ &= \frac{1}{2}\left(|H\otimes 1\otimes D_1\rangle\langle H\otimes 1\otimes D_1|\right.\\ &\quad +|H\otimes 1\otimes D_1\rangle\langle V\otimes 2\otimes D_2|\\ &\quad +|V\otimes 2\otimes D_2\rangle\langle H\otimes 1\otimes D_1|\\ &\quad \left.+|V\otimes 2\otimes D_2\rangle\langle V\otimes 2\otimes D_2|\right).\end{aligned} \tag{6.71}$$

我们称这个状态为纠缠态，光子偏振自由度的状态和哪一个探测器被激活这两个变量之间存在着相互联系。关于纠缠态的更进一步的讨论可见第九章。

这个时候，我们来做一件在经典概率论中经常做的事情，约化分布函数：对于两个随机变量的联合分布函数，如果我们只关心其中一个变量的分布函数，我们取

$$\rho^1 = tr^2\left(\rho^{12}\right) = tr^{-1}\left(\rho^{12}\right). \tag{6.72}$$

有的时候我们也用后面的记号，表示把系统 1 的变量留下来，其它都取和或者积分掉。现在，我们利用这个公式对探测器的状态取和，也就是说，把探测器的状态取和以后的状态是

$$\begin{aligned}\rho_{2\not{D}}^{P,T} &= \langle D_1|\rho_{2D}^{P,T}|D_1\rangle + \langle D_2|\rho_{2D}^{P,T}|D_2\rangle\\ &= \frac{1}{2}\left(|H\otimes 1\rangle\langle H\otimes 1|+|V\otimes 2\rangle\langle V\otimes 2|\right).\end{aligned} \tag{6.73}$$

这个时候我们发现，非对角元全部消失了！这个状态下，如果我们计算输出的结果的几率，我们就会发现，有两个输出，而且几率相等。其中 $\not{D}$ 表示把探测器变量求和掉。

于是，我们发现，如果加入探测器，那么，实验结果是有两个输出，和经典概率论的解释一致，而且也与实验结果一致 [41]。也就是说，加入了探

测器之后，粒子到底走哪一条路径，成了可以分辨的了，于是，我们应该采用两条路径的状态的“概率叠加”。其结果正好就是公式 (6.73)。那么什么时候得到的状态会回到形如公式 (6.64) 的“量子态叠加”呢？在一种情况下可能：探测器 D_1 和 D_2 完全不准确，不管粒子过哪一条路径，探测器都会发光，或者都不发光。例如原来准确的探测器如果发现路径 1（2）有光子就显示红（绿）光，但是现在的探测器不管那个光路上有光子都显示白光。于是实际上

$$D_1 = D_2 = D \tag{6.74}$$

那么，这个时候再对公式 (6.71) 经过部分迹求和之后，我们得到

$$\rho_{2\not{D}}^{P,T} = \frac{1}{2}\left(|H \otimes 1 + V \otimes 2\rangle\langle H \otimes 1 + V \otimes 2|\right), \tag{6.75}$$

也就是公式 (6.64)。

在上面的计算过程中，我们没有用到那条临时公理，我们依靠的是前面的公理，以及部分迹这个数学操作对于纠缠态的作用结果。我们发现探测器完全准确则得到概率叠加态，完全不准确则得到矢量叠加态。

所以，实际上我们的“路径能否区分”的概念在这里就相当于探测器是否准确。对于完全准确的探测器，我们采用“概率叠加”，对于完全不准确的探测器，我们采用“量子态叠加”。对于粒子走哪一条路径的问题，在完全准确的探测器的情况下，我们可以问。在完全不准确的探测器的情况下，我们不能问这个问题。我们也不能问是否粒子同时走两条路径的问题。因为只要你问这样的问题，为了回答它们，我们就需要准确的探测器。然后，只要有准确的探测器，实际上我们就不得不放弃“量子态叠加”而采用“概率叠加”，整个得到的状态就不再是加入探测器之前的状态了。从前面的准确与否的探测器的问题，我们可以看到，如果探测器准确，则光子的偏振状态和路径之间建立起来了一一对应关系（将来我们会专门有一个词来描述这样的对应——纠缠），则做完部分迹以后更像经典态；如果探测器完全不准确，则做完部分迹之后的状态更像量子态。因此，有可能由前面的经典随机客体的公理所描述的经典世界，是由量子公理所描述的世界的特殊情形。当然，也有可能，我们本来就有经典和量子两个不同的世界。

更一般地来说，状态是否可区分总是对应着一个系统本身状态的自由度和某个外界用来标志这个状态的另一个自由度的“纠缠”，然后，我们需要对这个这个纠缠之后的状态做部分迹。部分迹之后可能得到矢量叠加态也可能得到概率叠加态，取决于纠缠的程度：完全没有纠缠的状态在部分迹

之后得到满足矢量叠加的纯态，完全纠缠则得到完全概率叠加的混合态，一般的纠缠态得到两者之间的混合态。因此，临时公理也就不需要了，只需要在考虑一个系统的时候，把可能的和这个系统纠缠的其他自由度也考虑进来，并且在仅仅关心系统的时候，做部分迹，就自然会得到正确的矢量叠加态或者概率叠加态。

当然，在你明白量子系统的量子性被探测器破坏之前需要采用“量子态叠加”之后，那么你就算采用“粒子同时走两条路径”或者“粒子走的路径有两个可能”的说法的时候——只要你确实采用“量子态叠加”的方式——就没有问题。路径积分量子力学就建立在这样一个图景之上。现在我们强调在量子力学里面“粒子到底走哪一条路径”的问题不能问（问了就需要测量，测量了就改变了状态），但是路径积分形式的量子力学就采用这样一种视角 [39]：每一条路径，记为 $\mathcal{L}$，的贡献可以写成一个复数形式的振幅 $\mathcal{A}(\mathcal{L})\,e^{iS(\mathcal{L})}$，然后我们需要把所有的路径的贡献采用“量子态叠加”的方式加起来，得到的结果就是整体的量子态。

在本节的最后，我们交代一个我们故意隐藏起来的问题：为什么从两个随机变量的联合密度矩阵通过部分迹运算我们得到的是其中一个变量的密度矩阵，或者说更好的问题是——如果我们承认这个部分迹操作是从联合密度矩阵到部分变量的密度矩阵的定义的话，那么我们需要回答什么时候我们要做这样的操作，什么时候不做。例如，就算我们已经得到了公式 (6.75)，如果我们对它做一个光路的部分迹，我们就会得到

$$\rho^{P}_{2,false} = \frac{1}{2}\left(|H\rangle\langle H| + |V\rangle\langle V|\right). \tag{6.76}$$

而这个密度矩阵给出的结果是将来有两个输出，是不符合实验结果的。于是，我们的问题是，为什么现在我们不做这个光路部分自由度的部分迹操作，而对于探测器的问题，我们需要做部分迹操作呢？什么东西使得探测器成了特殊的一个东西，凡是看见它们以后都需要做一个部分迹呢？

你也许猜测是因为路径的信息可以通过反射镜再一次抹掉，也就是路径 1,2 重新变成路径 0，但是探测器记录到了光子经过了那就不能再抹去了。那么，接着问，能和不能抹去的差别是什么物理特质造成的？如果我们想办法把探测器得到的记录毁掉不去看，算不算抹去？

这个“什么时候用部分迹”的问题我们现在还回答不了。但是，这是一个将来需要回答的问题。如果你想了解更多，你可以在 Google[6]上检索“退

[6]google 网页搜索服务，http://www.google.com，2017 年 5 月 31 日访问。

相干”或者“decoherence”，以及等你来做更深入的研究。

现在，我们坚持对纠缠在一起的两个或者多个自由度，或者两个或者多个系统，当我们仅仅关注其中的一个自由度或者一个系统的状态的时候，我们需要对其做部分迹，把不关心的自由度消掉。在逻辑上，一旦我们这样来看，那么，关于态叠加的临时公理就不再是独立的了，它可以从量子力学的其他公理中推导出来。

习题 6.1 加了探测器的双缝干涉：阅读《Feynman 物理学讲义》第三卷关于双缝干涉以及加了探测器以后的双缝干涉的部分，用在里学到的 Dirac 符号密度矩阵的语言，以及这里的部分迹，来重新推导和表述书里面的主要论证和结论。

6.5 试试把测量分成三个步骤

有了上面的实验解释和部分迹的铺垫，在这一节，我们来尝试给测量建立一个更加清楚的图景，也交代对测量的理解上的问题所在。

对于测量，我们首先要有一个目标状态和测量物理量，前者表现为一个密度矩阵 ρ_0^q，后者表现为一个算符 O。其实我们还需要一个测量仪器 m。我们知道测量的形式理论，也就是测量公理是这样表达的：测量得到的结果是算符 O 的某一个本征值 o，其几率（可以理解为如果同样的 ρ_0^q、O 和同样的 m 重复多次做系综测量得到的统计规律）是 $p_0 = tr\left(|o\rangle\langle o|\rho_0^q\right)$；测量后状态，在观测到 o^* 本征值的条件下是 $|o^*\rangle\langle o^*|$。当我们问测量的物理过程到底是怎样，而不仅仅是形式理论的时候，我们关心的是 q 系统和 m 仪器是如何相互作用的，以至于导致了上面的形式理论结果。关于这一点，我们还需要了解“退相干”，还需要进一步学习量子态的演化尤其是第十章。甚至，还有需要进一步研究的地方。在这一节，我们仅仅给这方面的思考和探索开一个头。

我们把测量分成如下三个步骤，看看是否能够比形式理论走得更深一步，

$$\rho_0^q \otimes \rho_0^m \xrightarrow{①} \rho^{qm} \xrightarrow{②} \rho^m \xrightarrow{③} |m^*\rangle\langle m^*| \xrightarrow{④} |o^*\rangle\langle o^*|. \tag{6.77}$$

其中第 4 步完全就是逻辑上的推导：测量仪器上显示是状态 m^* 则必然被测量系统的状态就是相应的 o^*。这样一个状态相互对应的体系是设计测量仪器的时候就要明确的，例如指针向上（向下）就是自旋 z 方向向上（向下）

态之类的。因此，这里我们关心前面三步。我们已经看到对于自旋状态和自旋的路径，光子偏振和光子的光路，第一步往往是这样的，

$$
\begin{aligned}
&(\alpha|H\rangle+\beta|V\rangle)(\alpha^*\langle H|+\beta^*\langle V|)\otimes|0\rangle\langle 0| \\
\xrightarrow{①}\ &(\alpha|H\otimes 1\rangle+\beta|V\otimes 2\rangle)(\alpha^*\langle H\otimes 1|+\beta^*\langle V\otimes 2|)\,.
\end{aligned}
\tag{6.78}
$$

这一步的主要任务就是在系统的状态和测量仪器的状态之间建立起来一一对应的关联。第二步往往是通过部分迹来完成的。在概率论上，部分迹的含义是当有多个变量的时候，如果我们忽略某些变量，仅仅关注一部分变量，则需要运用部分迹。

$$
\begin{aligned}
&(\alpha|H\otimes 1\rangle+\beta|V\otimes 2\rangle)(\alpha\langle H\otimes 1|+\beta\langle V\otimes 2\rangle|) \\
&\xrightarrow{②}(\alpha^*\alpha|1\rangle\langle 1|+\beta^*\beta|2\rangle\langle 2|)
\end{aligned}
\tag{6.79}
$$

第三步实际上就是从这个概率分布中抽取一个样本的过程，也就是，

$$
(\alpha^*\alpha|1\rangle\langle 1|+\beta^*\beta|2\rangle\langle 2|)\xrightarrow{③}|1\rangle\langle 1| \tag{6.80}
$$

或者

$$
(\alpha^*\alpha|1\rangle\langle 1|+\beta^*\beta|2\rangle\langle 2|)\xrightarrow{③}|2\rangle\langle 2| \tag{6.81}
$$

分别以概率 $\alpha^*\alpha$ 和概率 $\beta^*\beta$。

这里无论哪个步骤都是挺神奇的：第一步，被测量系统和仪器构成的整体系统从没有关联的状态演化成为一个关联的状态；第二步，整体系统中被测量系统的变量被求和掉；第三步，从一个概率分布函数中抽样得到具体的其中的某一个状态。其中第一步我们将来会在第七章量子系统的演化中讨论。另外的两步，逻辑上好像也很自然。但是，我们已经提到，还存在问题：第一，什么情况下我们要做这个部分迹求和？第二，从概率分布中抽样得到一个具体的值真的是一件容易被理解的事情吗？假设经典纯随机客体存在，例如纯随机的硬币，我们如何理解看之前这个硬币可能有两个状态，但是看了之后必然处于其中的某一个这样一件事情呢？这个硬币的状态改变了吗？更重要的，这个硬币的状态客观吗，还是受观察者的“意识”影响从而改变了状态？在这个问题上，强烈推荐大家再一次去看一下 Coleman 的讲座《Quantum Mechanics in Your Face》[7]。关于这个问题的思考，我们以后

[7]见http://media.physics.harvard.edu/video/?id=SidneyColeman_QMIYF

在还会继续，暂时先停在这里。请大家思考，经典随机性客体的测量也可以表达成上面的三步的过程吗?

6.6 完整地测量自旋的状态

由于量子测量的几率特征，见公理3和公理4，单次测量给我们的信息有限。例如如果一次测量给我们的结果是某个本征值出现了，我们仅仅能够说：这个被测量的状态和出现这个本征值代表的状态不正交，包含了这一部分分量。那么，这个时候，我么自然就要问这样一个问题：我们如何通过测量明确地知道一个量子态呢？例如，对于一个从 z 方向向上射出的自旋，是不是我们只有采用了 z 方向的测量，而且一直看不见向下的地方有输出，我们才知道这个自旋的状态确实是 $|\uparrow_z\rangle$ 呢？如果是这样的话，我们如何从所有的方向中，刚好选到了这个 z 方向呢？我们不仅仅对给定的方向要做很多次测量，还需要选择所有的方向都尝试那么多次的测量。如果是这样的话，通过测量来确定量子态就是一个非常不可能完成的任务了。

实际上，我们有更简单的办法。任意的自旋的状态，可以表达成，

$$\rho = \begin{bmatrix} p & q \\ q^* & 1-p \end{bmatrix}. \tag{6.82}$$

总共就是三个变量（复数 q 算两个），也就是说，如果我们把实验设计好了，我们只需要三个试验结果就可以得到这个状态的一般表达式。我们来试试测量 σ_x、σ_y、σ_z 的平均值。例如，

$$\langle\sigma_z\rangle = tr\left(\rho\sigma_z\right) = 2p - 1. \tag{6.83}$$

也就是说，如果我们测量得到 σ_z 的平均值，那么我们也就知道了参数 p。接着，我们有

$$\langle\sigma_x\rangle = q + q^*, \left\langle\sigma_y\right\rangle = i\left(q - q^*\right). \tag{6.84}$$

于是，我们只需要把这三个自旋的平均值都测量一下，当然，这三个中的每一个量的测量还是需要做系综测量才能得到平均，而不是对所有的方向都需要做系综测量，就可以得到自旋的状态了。

6.7 量子与经典密度矩阵的区别

这一部分，我们来尝试用经典密度矩阵来解释双缝干涉现象、光子 which-way 实验和电子 which-way 实验。在那之前，我们尝试来解释一下更简单的一个或者两个 Stern-Gerlach 装置的 Stern-Gerlach 实验。

先考虑一个只有一个 Stern-Gerlach 装置的实验，由于初始自旋状态没有完全确定，这个实验通常会出现两个斑点，就好像在原始的Stern和Gerlach的实验的结果一样。这样的一个结果，其实是可以简单地用一个两状态的经典随机客体——就好像一个完全随机的硬币——来描述的，其状态大概可以这样，

$$\rho^c = p_\uparrow |\uparrow\rangle\langle\uparrow| + p_\downarrow |\downarrow\rangle\langle\downarrow| . \tag{6.85}$$

当考虑两个 Stern-Gerlach 装置的实验的时候，我们还是可以写下来经典的密度矩阵来解释实验。例如，先通过 z 方向，堵住向下的自旋仅仅让向上的自旋进入 z 方向的下一个 Stern-Gerlach 装置，看看有几个输出。我们已经知道实验结果是一个输出。能够解释实验结果的纯随机客体的经典密度矩阵可以写成这样，

$$\rho^c = |\uparrow\rangle\langle\uparrow| . \tag{6.86}$$

于是，可以计算得到这个测量的结果只有一种可能（重复多次的话，屏幕上会有一个斑点）。

再例如，先通过 z 方向，堵住向下的自旋仅仅让向上的自旋进入 x 方向的下一个 Stern-Gerlach 装置，看看有几个输出。我们已经知道实验结果是两个输出。能够解释实验结果的纯随机客体的经典密度矩阵可以写成这样，

$$\rho^c = \frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right) . \tag{6.87}$$

可以计算得到这个测量的结果有两种可能（重复多次的话，屏幕上会有两个斑点），这两种可能的几率都是 $\frac{1}{2}$。

那么这样的经典密度矩阵是不是也能够解释三个 Stern-Gerlach 装置的实验例如先过 z 挡住向下的自选再过 x 挡住向下的自选然后再过 z 的那个 $S_z S_x S_z$、双缝干涉现象、光子 which-way 实验和电子 which-way 实验呢？我们试试。我们仅仅用电子自旋 which-way 实验（图 6.1d）为例，解释其他的几个实验所需要的实际上背后的数学是一样的。

还是按照前面的配方，第一步，到达 σ_x 磁场之前的状态是

$$\rho_0^{S,T;c} = \frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right)\otimes|0\rangle\langle 0|. \tag{6.88}$$

经过 σ_x 磁场之后，到达最后的 σ_z 磁场之前，

$$\rho_1^{S,T;c} = \frac{1}{2}\left(|\uparrow_x\otimes 1\rangle\langle\uparrow_x\otimes 1| + |\downarrow_x\otimes 2\rangle\langle\downarrow_x\otimes 2|\right). \tag{6.89}$$

经过最后的 σ_z 磁场之前，

$$\rho_2^{S,T;c} = \frac{1}{2}\left(|\uparrow_x\rangle\langle\uparrow_x| + |\downarrow_x\rangle\langle\downarrow_x|\right)\otimes|0\rangle\langle 0|, \tag{6.90}$$

回到入射自旋的状态。然后我们需要计算测量得到 $|\uparrow_z\rangle$ 和 $|\downarrow_z\rangle$ 方向自旋的几率。原则上，为了计算这个，我们需要代表这两个方向自旋的算符 σ_x^c 和 σ_z^c，然后分别求出来它们的本征向量，再利用 $P_{s_z^c} = \left\langle s_z^c\middle|\rho\middle|s_z^c\right\rangle$ 算出这样的自旋方向的几率。实际上，经典力学框架之内，我们找不到这样的算符的数学形式。我们只能够通过实验来给一个说法。

首先，我们注意到，对于我们的具体情况，

$$\begin{aligned} P_{s_z^c} = \left\langle s_z^c\middle|\rho_2^{P,T;c}\middle|s_z^c\right\rangle &= \frac{1}{2}\left\langle s_z^c\middle|\rho_{2,\uparrow_x}^{S,T;c}\middle|s_z^c\right\rangle + \frac{1}{2}\left\langle s_z^c\middle|\rho_{2,\downarrow_x}^{S,T;c}\middle|s_z^c\right\rangle \\ &= \frac{1}{2}P_{s_z^c}(\uparrow_x) + \frac{1}{2}P_{s_z^c}(\downarrow_x). \end{aligned} \tag{6.91}$$

其中，$\rho_{2,\uparrow_x}^{S,T;c}$ 代表 $\rho_2^{P,T;c}$ 的第一部分（关于 $|\uparrow_x\rangle\langle\uparrow_x|$ 的部分），$\rho_{2,\downarrow_x}^{S,T;c}$ 代表 $\rho_2^{P,T;c}$ 的第二部分（关于 $|\downarrow_x\rangle\langle\downarrow_x|$ 的部分）。这个公式表明，不管 $\sigma_{\hat{r}}^c$ 是什么，概率性叠加公式是成立的：某事件有两种发生的可能，则其概率是这两种事件发生概率按照全概率公式取和。现在，我们需要知道 $P_{s_z^c=1}(\uparrow_x), P_{s_z^c=-1}(\uparrow_x)$ 和 $P_{s_z^c=1}(\downarrow_x), P_{s_z^c=-1}(\downarrow_x)$。因为没有算符的具体形式，我们通过如下的实验来推算这几个数值。制备自旋状态 $|\uparrow\rangle\langle\uparrow|$，然后做过 z 方向磁场的测量。也就是让上面的实验中的路径 1 的自旋直接通过 z 方向磁场，同时把路径 2 的自旋挡住。实验告诉我们，过了 z 方向磁场之后有两个可能的输出光斑，多次实验告诉我们这两个光斑的强度在误差范围内完全相同。于是我们知道，

$$P_{s_z^c=1}(\uparrow_x) = \frac{1}{2} = P_{s_z^c=-1}(\uparrow_x) \tag{6.92}$$

类似的，我们可以通过让路径 2 的自旋单独通过 z 方向磁场，得到

$$P_{s_z^c=1}(\downarrow_x) = \frac{1}{2} = P_{s_z^c=-1}(\downarrow_x) \tag{6.93}$$

于是，代入公式 (6.91) 得到

$$P_{s_z^c=1} = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2} = P_{s_z^c=-1}. \tag{6.94}$$

最终我们得到：按照经典力学的几率表达式公式 (6.88)，我们会得到两个可能的输出光斑，而且其多次实验得到的强度相同。这个计算结果与实验完全不相符。

对比公式 (6.35) 和公式 (6.87)，我们看到其区别就是

$$\Delta\rho = \frac{1}{2}\left(|\uparrow_x\rangle\langle\downarrow_x| + |\downarrow_x\rangle\langle\uparrow_x|\right). \tag{6.95}$$

这个部分只有在量子的密度矩阵中存在，经典的没有。然而这部分对于解释量子系统的实验行为是非常重要的。由于这些项的存在，我们能够用量子力学的公理来“计算出来”量子系统的测量结果，而不能用没有这些项的经典概率论的公理来“计算出来”量子系统的测量结果。

由于量子系统的状态由 Hilbert 空间的矢量描述，而矢量可以用基矢的叠加来表示（例如 $|\mu\rangle = \mu_1|v_1\rangle + \mu_2|v_2\rangle$），当我们把这样的矢量转变成为密度矩阵的时候（$\rho = |\mu\rangle\langle\mu|$），必然会带来以上形式的交叉项。我们粗略地称之为密度矩阵的非对角元。之所以“粗略”是因为这些非对角元在其它的表象下，可以成为对角元。而且，因为密度矩阵是 Hermitian 的，这样的表象总是存在的，尽管这个表象和具体实验中的可观测量的表象经常不一样。如果两者一样，那么在这个一样的表象下，密度矩阵和可观测量都是对角的。这个时候，整个系统的数学形式就完全回到了经典概率论的形式。至少在这个可观测量的测量的角度来说，量子性（相干相消或者相干相长）消失了。当然，同样的系统，如果我们考虑另外一个可观测量，只要这个可观测量对应的算符与前一个可观测量不能对易（可对易算符有共同本征矢量集合），那么这个时候量子性又会体现出来。因此，也可以说量子力学的量子性体现在算符的非对易性上：物理量不再是数，而是矩阵或者说算符。

尽管我们一再指出，通过量子实验表现出来的量子系统的行为使得我们有必要引入事件之间的加法操作，才是量子力学的数学采用矩阵的形式的根本原因，其实，形式上，我们完全可以从对易关系开始构建整个量子力学。这个就是正则量子化。正则量子化的内容我们会在第八章中介绍。

在此之前，我们需要学习一点点量子系统的演化的描述。在上面的例子中，我们有一个小小的未完成的步骤，为什么状态是 $\rho_0^{P,T}$ 光子经过偏振分束镜就能够变成状态 $\rho_1^{P,T}$ 了呢？这个问题与量子系统的演化有关。在解释它之前，让我们来学习一下量子系统的演化。

6.8 作业

习题 6.2 一个经典的完全随机的无偏硬币 c 当做被测量系统——你的大脑里面的一个硬币的映像模型 m（例如不太准确地说，当硬币向上的时候某些神经元点亮，向下的时候其他神经元点亮）——所测量。请你参考本章量子测量的三个步骤（其实是四个）写下来这个经典硬币的测量的可能的可以分解来看的步骤。

习题 6.3 构建一个自旋经过 z 方向的 Stern-Gerlach 装置之后，挡住向下的输出把向上的输出送到一个 $\hat{r}$ 方向的 Stern-Gerlach 装置之中得到的实验结果的数学描述。一个这样的数学描述包含：任意时刻的状态，测量的可观测量，测量各种结果及其几率等的数学模型。

习题 6.4 计算一个自旋经过 z 方向的 Stern-Gerlach 装置之后，挡住向下的输出把向上的输出送到一个 $\hat{r}$ 方向的 Stern-Gerlach 装置之中，挡住向下的输出，再送到一个 z 方向的 Stern-Gerlach 装置之后的实验结果。试验结果的描述需要包含所有的可能结果以及各个结果的几率。

习题 6.5 坐标变换对自旋状态的作用：考虑一个三维位置空间的矢量，一个绕 z 抽转动 α 角度的操作可以写成一个矩阵形式（先试着写下来试试，找找感觉）。现在，我们来考虑同样的转动对于自旋状态的效果。取任意一个自旋，看一看经过这个转动之后自旋的状态的数学表达式成了什么样，然后尝试用矩阵运算的方式把前后两个自旋状态联系起来。问这个时候那个联系着之前的和之后的自旋状态的矩阵是什么？对于任意的转动，我们可以写下来这个相应的矩阵吗？对三维空间的矢量是可以的。对于任意的转动，你不需要明确写下这个表达式。但是如果通过查阅资料能够找到，也可以写下来，不计分。提示：对于自旋，尝试考虑幺正变换，也就是 $|\psi'\rangle = U|\psi\rangle$ 的形式；可以考虑一个原来是 $\hat{r}(\theta,\phi)$ 方向向上状态的自旋矢量，经过坐标轴旋转之后成了哪一个方向的矢量，然后通过这个写下新的矢量的表达式；不一定要按照提示的思路。

习题 6.6 计算如图 6.4 一个自旋经过 z 方向的 Stern-Gerlach 装置之后，挡住向下的输出把向上的输出送到一个 x 方向的 Stern-Gerlach 装置之中，接着把两个可能的输出经过反射再次合起来，送到一个 z 方向的 Stern-Gerlach 装置之后的实验结果。实验结果的描述需要包含所有的可能结果以及各个结果的几率。

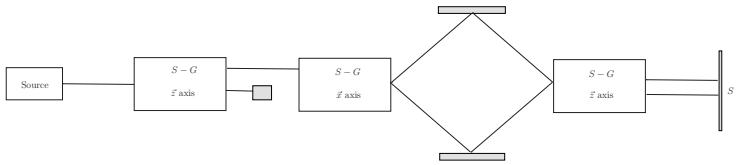


图 6.4: 一个自旋经过 z 方向的 Stern-Gerlach 装置之后，挡住向下的输出把向上的输出送到一个 x 方向的 Stern-Gerlach 装置之中，接着把两个可能的输出经过反射再次合起来，送到一个 z 方向的 Stern-Gerlach 装置

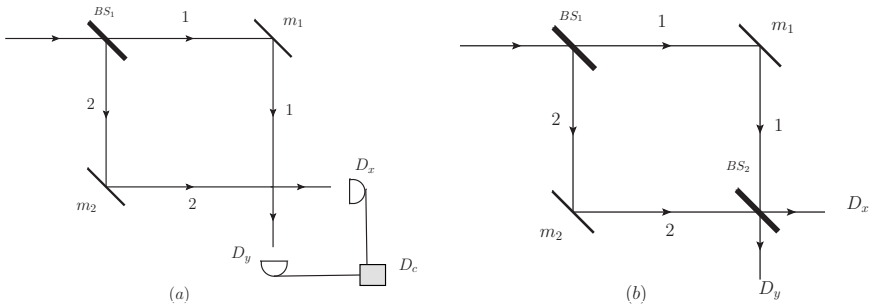


图 6.5: (a) 单光子进入 50 : 50 的分束器 BS_1，探测器 D_x 和 D_y 分别探测被 BS_1 反射的光子和透射的光子。符合计数器 D_c 用来检查 D_x 和 D_y 是否同时接收到光子。结果发现两个探测器从未同时接收到光子（Aspect Anti-coincidences 实验）。(b) 单光子进入 Mach-Zehnder 干涉仪，光子要经过两个分束器，然后被 D_x 和 D_y 探测。

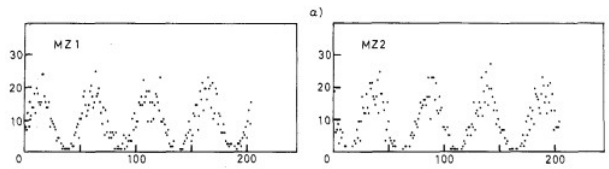


图 6.6: 图 6.5 中的实验 (b)，可以通过调整右上方那个可移动反射镜的位置，来改变两条光路的光程差。本图展示的是按照光程差的不同得到实验结果：D_x（MZ_1）和 D_y（MZ_2）上接收到的光子的数量。实验结果图来自于 Aspect 等人的原始文章 [15]（记得去获取授权）。

习题 6.7 图 6.5 是 Aspect 等人 1986 年完成的实验 [15]，目的是检验光子的波粒二象性。请回答以下问题：

1. 请计算第一种情况下探测器 D_x, D_y 接收到光子的概率。
2. 请计算第二种情况下探测器 D_x, D_y 接收到光子的概率，并和实验结果图 6.6 相比较。

习题 6.8 仍然考虑上一题的实验，如果在光子经过第一个分束器 BS_1 之后，到达计数器之前，再决定是否放置第二个分束器 BS_2，（图 6.5(a) 相当于不放置 BS_2, 图 6.5(b) 相当于放置 BS_2）请你预测结果会如何？这个想法是 1978 年 Wheeler 提出的，称作推迟选择实验。实际的推迟选择实验可以参看 Hellmuth 等人的 “Delayed-choice experiments in quantum interference” [42] 或者 Jacques 等人的 “Experimental Realization of Wheeler’s Delayed-Choice Gedanken Experiment” [43]。如果你愿意就这个问题做深入的思考，你可以收集和阅读这个方面的文献，做一个综述和评论。

习题 6.9 单个粒子自旋的计算：考虑一个在 $(z-x)$ 平面内的单位向量 $\hat{r}_\alpha = \sin(\alpha)\,\hat{i} + \cos(\alpha)\,\hat{k}$。这个 $\hat{r}$ 方向的电子自旋算符 S_α 是什么，计算它的本征值和本征向量。对于处于这个方向向上状态的电子，如果测量在 $\hat{r}_\beta$ 方向的自旋 S_β，得到哪些可能的结果，这些结果的概率是多少？进一步求出 S_β 在 S_α 方向向上状态下的期望值。

习题 6.10 求自旋的 σ_z 的本征态在 σ_x 表象的密度矩阵。在 σ_x 表象求 σ_z 本征态的密度矩阵。

习题 6.11 在 σ_z 表象，系统的的密度矩阵为 $\rho = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}$，求测量 σ_x 得到的所有可能取值的概率，并求均值。能否利用题6.10的结果直接计算 σ_z 的均值。提示：这个 ρ 可以看作什么，是纯态还是混合态？

习题 6.12 给定一个二维系统的密度矩阵 ρ，做以下三个物理量的测量，

$$A = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, C = \begin{bmatrix} 0 & -2i \\ 2i & 0 \end{bmatrix},$$

分别得到期望值

$$\langle A \rangle = 2, \langle B \rangle = \frac{1}{2}, \langle C \rangle = 0.$$

求 ρ。

6.9 本章小结

在这一章里面，我们学习了量子系统的数学模型的主要结构：状态是密度矩阵（或者特殊情况下，矢量）——还要注意什么时候用矢量叠加什么时候用概率性叠加或者说永远用矢量叠加直到遇到测量导致的部分迹于是自然退化成为概率性叠加、可观测量是算符——可能非对易、测量相当于得到一个概率分布、测量后状态就是被测量量算符的本征态——对应着所观测到的被测量量的值的那个本征态。然后，我们用这些基本数学结构描述了之前的量子系统的实验。

第七章　量子系统的演化

上一章，我们讨论了对于一个量子系统的给定的状态，我们如果测量某一个物理量，我们得到的结果是什么，描述这个“过程”的数学模型是什么。实际上在这里，对这个“过程”的数学描述并不是一个真实的发生在某一段时间内的过程。我们把所有的时间因素都去掉了。至于这个过程是否一定需要发生在一段时间之内，我们暂时没有讨论。在这一章里面，我们来讨论一个系统的一个时间段的前后两个状态之间的联系。我们学习过经典力学，我们知道力学的基本问题就是状态如何描述（上一章已经解决），状态如何发生变化以及发生变化的原因是什么。量子力学也是力学，也需要问后一个问题。经典力学中所有的这部分的信息放在了 Hamiltonian 里面，对应的方程是 Hamilton 方程，现在，我们来看看量子力学的情况。

如果我们需要考察前后两个状态之间是否有联系，那么我们可以给定实验装置之后，制备一个初始状态（例如让自旋通过 z 方向的 Stern-Gerlach 装置以后挡住向下的部分仅仅让向上的通过），然后让这个初始状态在同样的实验装置（例如 x 方向的磁场，为什么这样选一会儿就会知道）里面呆不同的时间，最后来测量（见 6.6 节完整地测量自旋的状态）从这个装置里面出来的状态，然后考察这个状态和时间的关系。我们还可以制备不同的初始状态的自旋，通过同样的仪器同样的时间，从而考察初始状态与末状态之间的关系。当然，为了验证我们的仪器是一个完整的确定了的仪器，我们需要通过让同样的初始状态的自旋通过同样的装置同样的时间来确定，我们得到的状态是同一个状态。

有了以上的实验结果，我们就可以构造演化过程的可能的数学理论了。实际上，上面的这些实验都是做过的。但是，并不完全是在量子理论提出来之前，有些实验是在量子理论提出之后做的。所以，尽管我们今天对于下面构造出来的量子力学的演化过程的数学理论可以按照这个逻辑方式来理解，但是实际上量子力学的演化过程的数学理论是 Schrödinger 猜出来的。等价

的形式还有 Heisenberg 的矩阵力学，Feynman 的路径积分。这里我们先介绍 Schrödinger 的形式。

本章在任何一本量子力学教材上都可以找到。如果一定要推荐的话，可以参考喀兴林的《高等量子力学》[28] 和Ballentine的《Quantum Mechanics – a modern development》[8]。

7.1 Schrödinger 方程

公理 9 (**量子态的演化公理**) 对于给定初始状态 $\rho(0)$ 的系统，系统自身动力学完全由系统的 Hamiltonian（H）通过如下方程决定，

$$i\frac{\partial}{\partial t}\rho(t) = [H, \rho(t)]. \tag{7.1}$$

或者在特定情况下 $\rho(0) = |\psi(0)\rangle\langle\psi(0)|$ 的时候，

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{7.2}$$

这个方程被称为 Schrödinger 方程。有时候，前者也被称为 Liouville–von Neumann 方程。可以验证前者在所指出的特殊情况下确实就是后者。如果 H 不显含时间，形式上以上的两个方程可以解开如下，

$$\rho(t) = e^{-iHt}\rho(0)e^{iHt}. \tag{7.3}$$

或者特殊情况下

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle \tag{7.4}$$

利用 H 的本征值和本征向量（记为 $E_n, |n\rangle$），我们还可以得到

$$U(t) = e^{-iHt} = e^{-iHt}\sum_n |n\rangle\langle n| = \sum_n e^{-iE_n t}|n\rangle\langle n|. \tag{7.5}$$

也就是说，如果我们能够得到 H 的本征值和本征向量，那么，从给定的初始状态出发，任何一个时刻的状态我们都可以得到。

得到新的时间点的状态之后，如果我们需要在这个状态上对某一个可观测量做测量，我们在利用前面的公理3来解决测量结果的问题。现在，我们用一个例子来展示一下这个公理的应用。

例 7.1 (**自旋演化**)：经过 z 方向磁场以后，在其正方向出射（反方向完全被挡主）的 $\frac{1}{2}$-自旋系统，经过 x 方向磁场 B 中 $t=\frac{\pi}{2B\mu}$ 时间以后出射。求正方向和反方向接收到自旋的几率。

在这里，对于这个具体问题，我们要做一个如何构造 Hamiltonian 的说明。对于所有的自旋系统

$$H=-\mu_0\vec{S}\cdot\vec{B}=-\frac{\hbar}{2}\mu_0\vec{\sigma}\cdot\vec{B}\triangleq-\mu\vec{\sigma}\cdot\vec{B}. \tag{7.6}$$

在最后的表达式中，我们把合起来的常数当作 μ。以后，我们就一直用这个符号体系：对于 $\frac{1}{2}$ 的自旋，我们取

$$H=-\mu\vec{\sigma}\cdot\vec{B}. \tag{7.7}$$

在跟实际实验对比的时候，我们需要把这些常数还原。但是在这里我们不关心这个细节。

有了这个 Hamiltonian 我们就可以利用公理9求出演化算符，

$$U(t)=\sum_n e^{-iE_nt}|n\rangle\langle n|=e^{-i\mu Bt}|\uparrow_{\hat{r}}\rangle\langle\uparrow_{\hat{r}}|+e^{i\mu Bt}|\downarrow_{\hat{r}}\rangle\langle\downarrow_{\hat{r}}|. \tag{7.8}$$

其中 $\hat{\sigma}_{\hat{r}}$ 的本征矢量 $|\uparrow_{\hat{r}}\rangle$ 和 $|\downarrow_{\hat{r}}\rangle$ 我们以前都计算过。这里 $\hat{r}=\hat{x}$，因此，

$$\begin{aligned}\left|\psi(t)\right\rangle=U(t)\left|\psi(0)\right\rangle&=e^{-i\mu Bt}|\uparrow_x\rangle\langle\uparrow_x|\uparrow_z\rangle+e^{i\mu Bt}|\downarrow_x\rangle\langle\downarrow_x|\uparrow_z\rangle\\&=\frac{e^{-i\frac{\pi}{2}}|\uparrow_x\rangle+e^{i\frac{\pi}{2}}|\downarrow_x\rangle}{\sqrt{2}}.\end{aligned} \tag{7.9}$$

在这个例子中，我们从一个 $|\uparrow_z\rangle=\frac{|\uparrow_x\rangle+|\downarrow_x\rangle}{\sqrt{2}}$ 出发，由于 x 方向磁场的作用我们得到了末状态 $\frac{e^{i\frac{\pi}{2}}|\uparrow_x\rangle+e^{-i\frac{\pi}{2}}|\downarrow_x\rangle}{\sqrt{2}}=i\frac{|\uparrow_x\rangle-|\downarrow_x\rangle}{\sqrt{2}}=i|\downarrow_z\rangle$。然后，我们就可以通过测量这个末状态（见 6.6 节）来验证它确实是这样的。

这个例题中的演化算符和相应的 Hamiltonian 具有特殊的地位：这个演化算符被称为 NOT 算符，其作用是把 $|\uparrow_z\rangle$ 和 $|\downarrow_z\rangle$ 相互翻转，就好像是一个计算机逻辑运算中的非运算一样，把 0 和 1 互换。

例 7.2 (**神奇的偏振分束镜**)：初始状态是 $\rho_0^{P,T}$ 光子经过偏振分束镜就能够变成状态 $\rho_1^{P,T}$：一个 3 维 Hilbert 空间和 2 维 Hilbert 空间直积构成的空间，定义 Hamiltonian

$$H=\Sigma\otimes\sigma_z, \tag{7.10}$$

其中 $\Sigma = i\frac{\sqrt{3}}{3}\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$ 是测量仪器部分的 Hamiltonian，σ_z 是光子部分的 Hamiltonian，满足 $\sigma_z|H\rangle = |H\rangle$，$\sigma_z|V\rangle = -|V\rangle$。定义 $U = e^{-iH\tau}$，其中 $\tau = \frac{4\pi}{3}$，证明

$$\rho_1^{P,T} = U\rho_0^{P,T}U^\dagger. \tag{7.11}$$

实际上，我们需要证明

$$U|0\rangle \otimes |H\rangle = |1\rangle \otimes |H\rangle, U|0\rangle \otimes |V\rangle = |2\rangle \otimes |V\rangle. \tag{7.12}$$

有了这个很容易就能够证明公式 (7.11)。我们以证明第一部分为例。首先，我们做一个关于算符 Σ 的说明。第一，为什么它是 3 维的。我们需要把一个一开始的状态 $|0\rangle$ 按照不同的情况分别映射到 $|1\rangle$ 和 $|2\rangle$ 态，于是这个空间至少是 3 维的，而且 3 维也就够了，因为我们考虑的对象总共也就这三种可能。定义 $|0\rangle$ $|1\rangle$ $|2\rangle$ 分别作为这个空间的基矢 $[1,0,0]^T, [0,1,0]^T, [0,0,1]^T$。第二，我们再来看一下这个 Σ 算符的在这个表象下的三个本征态，本征值分别为 $\sigma = 0$，$\sigma = -1$，$\sigma = 1$，

$$|\sigma = 0\rangle = \frac{\sqrt{3}}{3}[1,1,1]^T \tag{7.13}$$

$$|\sigma = -1\rangle = \frac{\sqrt{3}}{3}\left[-1, e^{-i\frac{\pi}{3}}, e^{i\frac{\pi}{3}}\right]^T \tag{7.14}$$

$$|\sigma = 1\rangle = \frac{\sqrt{3}}{3}\left[-1, e^{i\frac{\pi}{3}}, e^{-i\frac{\pi}{3}}\right]^T \tag{7.15}$$

于是，代入演化算符的定义以及 $\tau = \frac{4\pi}{3}$，我们有

$$e^{-i\Sigma\tau} = \sum_\sigma e^{-i\sigma\tau}|\sigma\rangle\langle\sigma| = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{7.16}$$

于是

$$e^{-i\Sigma\tau}|0\rangle = |1\rangle, e^{i\Sigma\tau}|0\rangle = |2\rangle. \tag{7.17}$$

现在，我们可以来讨论 $U|0\rangle \otimes |H\rangle$ 了，

$$U|0\rangle \otimes |H\rangle = e^{-i\Sigma\otimes\sigma_z\tau}|0\rangle \otimes |H\rangle = e^{-i\Sigma\tau}|0\rangle \otimes |H\rangle = |1\rangle \otimes |H\rangle. \tag{7.18}$$

同理，

$$U|0\rangle\otimes|V\rangle = e^{-i\Sigma\otimes\sigma_z\tau}|0\rangle\otimes|V\rangle = e^{i\Sigma\tau}|0\rangle\otimes|V\rangle = |2\rangle\otimes|V\rangle. \tag{7.19}$$

这个例子说明，测量仪器的状态和被测量物体的状态之间的关联的建立，实际上是通过两个系统之间的相互作用来完成的，而这个相互作用的过程可以用量子系统的演化方程——Schrödinger 方程来描述。

到此为止，量子力学的基本框架的内容就有了。我们的基本任务是帮助读者理清思路，建立最基本的框架，掌握最核心的概念，学会运用最基本的计算分析技术。到这里，我们的基本任务就完成了。可以看到，非常神奇地，通常的量子力学书，例如曾谨言的《量子力学》[44] 的用了大量篇幅的核心内容，在我们这里仅仅就是一小节，共 4 页。对于理解量子力学来说，静态描述是什么以及为什么需要这样的静态描述是更加重要的事情。下一节，我们介绍与 Schrödinger 方程等价的 Heisenberg 方程。

7.2 Schrödinger 绘景与 Heisenberg 绘景

坐标变换实际上相当于对量子状态也做了一个相应的变换。例如围绕 z 轴的 $\frac{\pi}{2}$ 旋转，记为 $R_z\left(\frac{\pi}{2}\right)$，把原来的 y 轴变成了新的 x' 轴，把新的 y' 建立在了原来的 x 轴的负方向。那么原来的例如 x 轴的向上方向的矢量就成了在新的坐标下的 y' 轴的正方向的矢量，也就是 $|\downarrow_x\rangle \to \left|\uparrow_{y'}\right\rangle$，这相当于，在 σ_z 表象下，

$$\frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \to \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ i \end{bmatrix} = S\left(R_z\left(\frac{\pi}{2}\right)\right)\frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{7.20}$$

我们把这个由于坐标变换 $R_z\left(\frac{\pi}{2}\right)$ 导致的自旋的状态的变换记为 $S\left(R_z\left(\frac{\pi}{2}\right)\right)$。

我们暂时不去关心 $R_z\left(\frac{\pi}{2}\right)$ 和 $S\left(R_z\left(\frac{\pi}{2}\right)\right)$ 的具体形式。对于有兴趣的读者这些具体形式的问题称为群的表示理论：类似于 $R_z\left(\frac{\pi}{2}\right)$ 这样的所有的坐标变换构成一个群。了解坐标变换会导致一个自旋状态的矢量的分量形式的变换这件事情在这个阶段比我们找出具体的这些变换的算符的数学形式重要。

在做坐标变换的时候，我们希望有一种东西是不变的，例如，对于这里的旋转，我们希望我们的这个自旋状态，初始是 $|\downarrow_x\rangle$，后来为 $\left|\uparrow_{y'}\right\rangle$ 仅仅是记号上的改变，实际上，这个矢量的意义没有变换，就是那个如果在原来的坐标下面看来，是 $|\downarrow_x\rangle$，在后来的坐标下看起来是 $\left|\uparrow_{y'}\right\rangle$ 的那个自旋矢量。那么

这个不变的东西用什么样的数学形式来表达呢？答案是，这个矢量在任何一个可观测量下的观测值（包含本征值和本征值出现的几率）是不变的，也就是说，

$$tr\left(A\left|\downarrow_x\right\rangle\left\langle\downarrow_x\right|\right)=tr\left(A'\left|\uparrow_{y'}\right\rangle\left\langle\uparrow_{y'}\right|\right). \tag{7.21}$$

实际上，不仅仅是这个矢量，我们要求对于任意的自旋矢量，上式都成立，

$$tr\left(A\left|\mu\right\rangle\left\langle\mu\right|\right)=tr\left(A'\left|\mu'\right\rangle\left\langle\mu'\right|\right), \tag{7.22}$$

其中，$|\mu\rangle$ 是老坐标下的矢量，$|\mu'\rangle$ 是新坐标下的矢量。用我们的记号，

$$\left|\mu'\right\rangle=S\left|\mu\right\rangle, \tag{7.23}$$

那么只要

$$A'=SAS^{\dagger}, \tag{7.24}$$

并且

$$S^{\dagger}S=I=SS^{\dagger}, \tag{7.25}$$

上式就成立。满足这个性质的变换 S，分别按照相应的形式作用在矢量上和算符上，称为幺正变换。幺正变换在量子力学里面具有特殊的地位：由于这样的变换不改变矢量和算符的含义，仅仅改变两者的形式，它们被称为不改变物理的变换，仅仅是换一个观测的角度。物理学的规律对于仅仅改变观测的角度的变换是不变的——包含所有的公式的形式不变，所有的观测值不变——被认为是物理学的基本假设。

现在，我们用这个幺正变换以及其不变性的角度来提出一种与 Schrödinger 方程完全等价的量子力学体系——Heisenberg 力学。之前我们已经学习的描述量子系统演化的方式——密度矩阵发生演化可观测量算符不变的描述，称为 Schrödinger 绘景。现在，我们介绍另一种让可观测量算符的发生演化，分布函数保持不变的描述方法，称为 Heisenberg 绘景。这样的一个视角的转换——从关心态的演化到关心算符的演化——有的时候可以方便求解某些问题。以后我们会了解到，谐振子就是这样的一个例子。

我们已经知道，对于 H 不显含时间 t 的量子系统，其的演化可以表达为

$$\rho\left(t\right)=e^{-iHt}\rho\left(0\right)e^{iHt}, \tag{7.26}$$

物理量 A^S（为了表示 Schrödinger 绘景的算符与将要定义的 Heisenberg 绘景的算法的区别，我们把之前的算符的记号稍作修改，加上一个上标 ^{S}）的观测值 α 和相应的几率 p_α，由如下方程决定

$$p_\alpha = tr\left(\left|\alpha^S\right\rangle\left\langle\alpha^S\right|\rho\left(t\right)\right) = \left\langle\alpha^S\left|\rho\left(t\right)\right|\alpha^S\right\rangle = \left\langle\alpha^S\left|e^{-iHt}\rho\left(0\right)e^{iHt}\right|\alpha^S\right\rangle. \tag{7.27}$$

现在，我们定义

$$A^H = e^{iHt}A^S e^{-iHt}, \tag{7.28}$$

$$\rho^H = e^{iHt}\rho^S\left(t\right)e^{-iHt} = \rho^S\left(0\right). \tag{7.29}$$

然后记相应的 A^H 的本征向量为 $\left|\alpha^H\left(t\right)\right\rangle$，则

$$\left|\alpha^H\left(t\right)\right\rangle = e^{iHt}\left|\alpha^S\left(t\right)\right\rangle \tag{7.30}$$

因此，

$$tr\left(A^H\rho^H\right) = tr\left(\left|\alpha^H\left(t\right)\right\rangle\left\langle\alpha^H\left(t\right)\right|\rho\left(0\right)\right) = \left\langle\alpha^S\left|e^{-iHt}\rho\left(0\right)e^{iHt}\right|\alpha^S\right\rangle = p_\alpha. \tag{7.31}$$

也就是说，计算各种观测量的方式和得到的结果都不变。这种保持变换前后基本公式的形式和计算结果都不变的变换，在物理学里，有非常重要的地位。于是，我们得到量子力学的另一种形式，

$$A^H = e^{iHt}A^S e^{-iHt}, \tag{7.32}$$

$$\rho^H = \rho^S\left(0\right), \tag{7.33}$$

$$\langle A\rangle = tr\left(A^H\rho^H\right). \tag{7.34}$$

这里，最后一个公式中的算符 A 可以是任何一个可观测量的算符。因此，如果我们用这个绘景来讨论物理问题，则我们需要求解所有的物理量对应的算符的演化方程。这个，有可能比求解密度矩阵的演化（只有一个方程）来的复杂，除非我们能够找到一个最小的固定的算符的集合，所有的其它算符都可以用这个集合里面的算符来表达。对于任意算符，我们还可以得到一般的算符的演化方程。对 A^H 求时间 t 的导数，我们有

$$i\frac{\partial}{\partial t}A^H = \left[A^H, H\right]. \tag{7.35}$$

这个方程被称为 Heisenberg 方程。实际上，在量子力学的进一步发展中，我们会看到，很多时候 Heisenberg 绘景反而使问题变得更容易处理，因为这

个最小而固定的算符集合经常能够被找到。它们有的时候被称为产生-湮灭算符。这个，在谐振子问题中以及将来有机会学习量子场论的过程中，我们还会看见。

例 7.3 (**产生湮灭算符的演化**)：假设我们有一个系统，它的 Hamiltonian 是 $H = \omega a^\dagger a$，其中 $\left[a, a^\dagger\right] = 1$，求解 $a^H(t)$。

首先，$a^{\dagger,H} = e^{iHt} a^\dagger e^{-iHt}$，$a^H = e^{iHt} a e^{-iHt}$，于是，

$$\left[a^H, a^{\dagger,H}\right] = e^{iHt}\left[a, a^\dagger\right] e^{-iHt} = 1. \tag{7.36}$$

其次，$H = \omega e^{iHt} a^\dagger a e^{-iHt} = \omega e^{iHt} a^\dagger e^{-iHt} e^{iHt} a e^{-iHt} = \omega a^{\dagger,H} a^H$。接着，把算符 $a^{\dagger,H} a^H$ 代入 Heisenberg 方程，我们得到

$$i\frac{\partial}{\partial t} a^H = \left[a^H, \omega a^{\dagger,H} a^H\right] = \omega a^H \tag{7.37}$$

求解这个方程，我们得到

$$a^H(t) = e^{-i\omega t} a^S. \tag{7.38}$$

类似的我们可以得到，

$$a^{\dagger,H}(t) = e^{i\omega t} a^{\dagger,S}. \tag{7.39}$$

于是，如果我们已知初始状态，任意一个能够由 a^S 和 $a^{\dagger,S}$ 表达的算符 $\left\langle A\left(a^{\dagger,S}, a^S\right)\right\rangle$ 的平均值 $\left\langle A^H\left(a^{\dagger,H}, a^H\right)\right\rangle$ 自然也就知道了。把上面这句话里面的平均值换成所有的可能结果以及这些结果的相应几率也是正确的。

这个问题的物理背景和以上计算的物理意义在下一章中会解释。目前阶段，请把这个举例仅仅当作 Heisenberg 方程的一个练习。至少，通过这个练习，我们学会了计算 Heisenberg 方程，而且我们了解到能够在 Schrödinger 绘景下回答的问题也能够在 Heisenberg 绘景下来回答。

7.3 作业

习题 7.1 看Coleman的讲座《Quantum Mechanics in Your Face》，用概念地图的方式做一个包含总结体会和评价的报告。

习题 7.2 看Susskind的网上课程《Quantum Mechanics》的前两课，用概念地图的方式做一个包含总结体会和评价的报告。

习题 7.3 看 glsname:Feynman 的《Feynman 物理学讲义》第三卷的前三章，用概念地图的方式做一个包含总结体会和评价的报告。

习题 7.4 一个自旋系统，给定 Hamiltonian 为

$$H = \frac{1}{4+2\sqrt{2}}\begin{pmatrix} 1 & -1-\sqrt{2} \\ -1-\sqrt{2} & 3+2\sqrt{2} \end{pmatrix}. \tag{7.40}$$

在这里，我们先不管这样的 Hamiltonian 是如何得来的，对应着什么样的外界对自旋的相互作用。验算经过时间 π 之后，系统的状态由演化如下算符决定，

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{7.41}$$

这个算符有一个专门的名字——Hadamard算符，或者叫做Hadamard 变换。看看这个算符把 $|\uparrow_z\rangle$ 和 $|\downarrow_z\rangle$ 分别映射成什么？

习题 7.5 自旋的演化与测量，$H = -\mu\vec{B}\cdot\vec{\sigma}$，$\rho_0 = |\uparrow_z\rangle\langle\uparrow_z|$，求 t 时刻做 σ_y 测量得到的结果。其中 $\vec{B}$ 是一个一般的三维矢量，可以考虑用 (θ,ϕ) 来表达。

习题 7.6 两个自旋构成的一个系统，给定 Hamiltonian 为

$$H = -J\sigma_z^1\sigma_z^2. \tag{7.42}$$

系统的初始状态是一个纯态 $|\uparrow_x\rangle_1|\downarrow_x\rangle_2$，求 $t = \frac{\pi}{2J}$ 时刻系统的状态。

习题 7.7 Unitary（幺正）演化不会使得状态在纯态和混合态之间转化：对于给定的状态 ρ，请证明 $tr(\rho)$ 和 $tr(\rho^2)$ 是不变量。由于纯态的 $tr(\rho^1) = 1$，混合态的 $tr(\rho^1) < 1$。因此，Unitary（幺正）演化不会使得状态在纯态和混合态之间转化。

习题 7.8 Unitary（幺正）演化不改变熵：对于给定的状态 ρ，定义熵

$$S = -tr(\ln(\rho)\rho) \tag{7.43}$$

请证明：第一、量子力学的 Unitary（幺正）演化不改变熵；第二、纯态的熵等于零，混合态的熵大于零。因此，Unitary（幺正）演化不会使得状态在纯态和混合态之间转化。

7.4 本章小结

在这一章里面，我们学习了如何描述量子系统的状态的演化，并且用 Schrödinger 绘景——状态变化算符不变——和 Heisenberg 绘景——状态不变算符变化——两种方式来描述。两种方式都离不开演化算符 $U(t)$。

第八章　单个谐振子的量子力学

整个这一本量子力学教材中，仅仅这一章，和通常的量子力学教材是类似的。在这一章中，我们从二维 Hilbert 空间的量子系统的讨论，跳到位置表象下的量子力学系统的讨论。然后，我们又会回到抽象的 Hilbert 空间的形式。其实，位置表象下的波函数也是抽象的 Hilbert 矢量空间的矢量。为了和经典力学相联系，我们需要介绍一下正则量子化。正则量子化是把一个经典系统变成一个相应的量子系统的经常采用的办法。

本章推荐参考书喀兴林《高等量子力学》[28] 和 Ballentine 的《Quantum Mechanics – a modern development》[8]。

8.1　位置表象

正则量子化的基本思想是，找到一个经典系统的正则坐标 q 和正则动量 p，然后把 q 和 p 都升级成算符 $\hat{q}$ 和 $\hat{p}$，并假设其对易关系为

$$[\hat{q}, \hat{p}] = i. \tag{8.1}$$

这两个物理量之间的对易关系有时候被称为正则对易关系。然后其它物理量，它们是 q, p 的函数，就成了 $\hat{q}, \hat{p}$ 的算符。其中，对应关系的细节（例如 $q^2 p$ 应该成为 $\hat{q}^2\hat{p}$ 还是 $\hat{q}\hat{p}\hat{q}$）还需要进一步讨论，但是有了正则对易关系之后，其他的物理量都成了一般情况下不相互对易的算符。于是，共同本征矢量不存在了，密度矩阵的表述需要非对角元了，量子干涉现象就出现了。我们的下一个例子就是一维谐振子的量子化。在符号上，这一节我们把位置算符 $\hat{q}$ 重新还原成 $\hat{x}$，特指位置表象，而不是更一般的广义坐标表象。

在讨论这个具体例子之前，我们来看一下，位置表象是什么，位置表象下的算符，包含动量、能量等，又如何在位置表象下通过一定的基矢表达出

来。首先，我们来看位置算符 $\hat{x}$ 的本征值 x 和本征向量 $|x\rangle$，

$$\hat{x}|x\rangle = x|x\rangle. \tag{8.2}$$

这个时候我们需要知道 x 的取值范围。

从物理问题上，我们就可以知道，这个问题的答案是所有的实数，因为在实验上我们可以观察粒子到达任何一个点的这样一个事件。假设将来有一个状态 $|\psi\rangle = \sum_x \psi(x)|x\rangle$，或者说 $|\psi\rangle = \int dx\psi(x)|x\rangle$（为了记号简单，连续变量的积分符号和离散变量的求和符号不再区分），其中 $|x\rangle$ 表示基矢，$\psi(x)$ 这个基矢下的分量，然后我们做一个位置 $\hat{x}$ 的测量，按照量子力学的一般框架就应该有

$$P_x = \langle x|\left(|\psi\rangle\langle\psi|\right)|x\rangle = \left|\psi(x)\right|^2. \tag{8.3}$$

于是，既然各个点原则上都有测得粒子的概率，那么，至少 $\psi(x)$ 原则上应该有可能不为零的值。我们再从数学上来看 x 的取值范围。

先假设我们有某一个本征值 x_0 和本征向量 $|x_0\rangle$，然后我们来构造一个本征向量 $|x+x_0\rangle$：定义一个矢量 $e^{-ix\hat{p}}|x_0\rangle$。我们考虑 $\hat{x}e^{-ix\hat{p}}|x_0\rangle$。运用公式 (8.1)，我们可以写出如下表达式，

$$\begin{aligned}
\hat{x}e^{-ix\hat{p}}|x_0\rangle &= \hat{x}\sum_n \frac{(-ix)^n}{n!}\hat{p}^n|x_0\rangle \\
&= \sum_n \frac{(-ix)^n}{n!}\hat{x}\hat{p}^n|x_0\rangle \\
&= \sum_n \frac{(-ix)^n}{n!}\hat{x}\hat{p}\hat{p}^{n-1}|x_0\rangle \\
&= \sum_n \frac{(-ix)^n}{n!}(i+\hat{p}\hat{x})\hat{p}^{n-1}|x_0\rangle \\
&= \sum_n \frac{(-ix)^n}{n!}\left(i\hat{p}^{n-1}+\hat{p}\hat{x}\hat{p}^{n-1}\right)|x_0\rangle \\
&= \sum_n \frac{(-ix)^n}{n!}\left(in\hat{p}^{n-1}+\hat{p}^n\hat{x}\right)|x_0\rangle \\
&= (x+x_0)e^{-ix\hat{p}}|x_0\rangle
\end{aligned} \tag{8.4}$$

于是 $e^{-ix\hat{p}}|x_0\rangle$ 也是算符 $\hat{x}$ 的本征向量，而且其本征值为 $(x+x_0)$。从这里，我们看到，由于任意 x 的值，$e^{-ix\hat{p}}|x_0\rangle$ 都是 $\hat{x}$ 的本征向量，于是算符 $\hat{x}$ 的取

值空间是所有的实数（联系到测量的实验，位置空间的点是实数）。换过来说，把位置和动量算符的对易关系取为公式 (8.1)，与我们可以测量位置空间中的任意一个点粒子到达的几率这件事情没有矛盾。这个计算的倒数第三式到倒数第二式的推导用到了对易关系公式 (8.1)，并且把这个对易关系重复了足够多次来把算符 $\hat{x}$ 移到表达式的最右边。也就是

$$\hat{x}\hat{p}^n = in\hat{p}^{n-1} + \hat{p}^n\hat{x}. \tag{8.5}$$

有了位置空间的所有的本征矢量 $\{|x\rangle\}$ 构成的矢量空间之后，我们来研究算符 $\hat{x}, \hat{p}$ 在这个空间中的表达式。对于算符 $\hat{x}$，我们有

$$\hat{x}\,|x\rangle = x\,|x\rangle\,. \tag{8.6}$$

正交归一关系表达为，

$$\langle x|\,x'\rangle = \delta\left(x - x'\right). \tag{8.7}$$

完全性关系表达为

$$\sum_x |x\rangle\langle x| = I. \tag{8.8}$$

当算符的本征值为连续变量的时候，实际上，归一化和完全性关系都会有问题[1]。在此，我们按照离散取值的本征值来看待连续取值情形。更深入的讨论见 Ballentine 的《Quantum Mechanics – a modern development》[8]。

于是

$$\hat{x} = \sum_x \hat{x}\,|x\rangle\langle x| = \sum_x x\,|x\rangle\langle x|\,. \tag{8.9}$$

我们来看看满足以上对易关系的 $\hat{p}$ 会是什么？我们先给出答案，再给出一个直观的说明。答案是在 $|x\rangle$ 的表象下，任意一个函数 $\psi(x)$ 成了一个如下的矢量，

$$|\psi\rangle = \sum_x \psi(x)\,|x\rangle\,, \tag{8.10}$$

算符 $\hat{x}$ 作用在 $|\psi\rangle$ 可以表达成其如何做用在分量 $\psi(x)$ 上：

$$\hat{x}\,|\psi\rangle = \sum_x \hat{x}\psi(x)\,|x\rangle = \sum_x x\psi(x)\,|x\rangle\,, \tag{8.11}$$

于是，从分量形式来看，

$$\hat{x}\psi(x) = x\psi(x)\,. \tag{8.12}$$

[1]例如，模方代表概率，概率应该归一，但是，$\int dx\,|\langle x|\,x'\rangle|^2 = \int dx\left|\delta\left(x - x'\right)\right|^2 = \int dx\delta^2\left(x - x'\right) = \infty$

那么，在这个意义上，从分量形式来看，

$$\hat{p}\psi(x) = -i\frac{\partial}{\partial x}\psi(x). \tag{8.13}$$

可以验证如下：

$$(\hat{x}\hat{p} - \hat{p}\hat{x})\psi(x) = -ix\frac{\partial}{\partial x}\psi(x) + i\frac{\partial}{\partial x}x\psi(x) = i\psi(x). \tag{8.14}$$

于是，

$$\left[x, -i\frac{\partial}{\partial x}\right] = i. \tag{8.15}$$

现在我们来给一个更数学的说明（更严格的证明见喀兴林的《高等量子力学》[28]）。从正则对易关系出发，

$$\langle x'|i|x\rangle = \left\langle x'\left|(\hat{x}\hat{p} - \hat{p}\hat{x})\right|x\right\rangle = \left\langle x'\left|\hat{p}(x'-x)\right|x\right\rangle = (x'-x)\langle x'|\hat{p}|x\rangle \tag{8.16}$$

因此

$$\langle x'|\hat{p}|x\rangle = i\frac{\langle x'\mid x\rangle}{(x'-x)} = i\frac{\delta(x'-x)}{(x'-x)} = -i\frac{\partial}{\partial x'}\delta(x'-x). \tag{8.17}$$

最后一步，我们利用了Dirac δ 函数的性质，$x\delta'(x) = -\delta(x)$ 或者说 $-\delta'(x) = \frac{\delta(x)}{x}$（这两个表达式可以用分步积分证明。更严格的数学证明可以避开这两个等式 [28]。）。这个就是算符 $\hat{p}$ 在 $\hat{x}$ 的本征矢量的表象下的分量形式。这个表象称为位置表象。如果我们愿意，我们可以验证

$$\langle x|\hat{p}|\psi\rangle = \sum_{x'}\langle x|\hat{p}(|x'\rangle\langle x'|)|\psi\rangle = \sum_{x'} -i\frac{\partial}{\partial x}\delta(x-x')\psi(x') = -i\frac{\partial}{\partial x}\psi(x). \tag{8.18}$$

以后，在位置表象下，当我们把一个状态矢量 $|\psi\rangle$ 的投影函数 $\psi(x)$ 看作我们所关注的状态的时候（默认基矢是 $|x\rangle$），我们就简单把动量算符 $\hat{p}$ 记作

$$\hat{p} = -i\frac{\partial}{\partial x}. \tag{8.19}$$

现在，我们准备好了来写下和求解位置空间的 Schrödinger 方程了：用 $H(\hat{x},\hat{p})$ 来代替 $H(x,p)$，其中，在位置表象下 $\hat{x} = x, \hat{p} = -i\frac{\partial}{\partial x}$。通常的量子力学书从这个方程开始学习量子力学。

例 8.1 (**一维谐振子的量子化与求解**)：经典一维谐振子的 Langrangian 是 $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega_0^2x^2$，求解对应的量子系统的 Hamiltonian，以及最小能量值的本征态对应的本征函数，以及其时间演化。

首先，我们需要求出正则坐标和动量，其次要把它们看成算符，在位置表象下把这两个算符的具体分量形式写下来，然后我们把这些分量形式代入 Schrödinger 方程，求解本征函数以及时间演化。

取 $q=x$，则 $p=\frac{\partial}{\partial\dot{q}}=m\dot{x}\Rightarrow\dot{q}=\frac{p}{m}$，于是 $H=p\dot{q}-L=\frac{p^2}{2m}+\frac{1}{2}m\omega_0^2x^2$ 成了 q,p 的函数。然后，我们需要把 q,p 看作算符 $\hat{q},\hat{p}$，并且让其对易关系 $[\hat{q},\hat{p}]=i$。按照这个对易关系，我们已经知道在位置表象下，位置算符和动量算符的形式。于是，对于一维谐振子系统，其在位置表象下的形式是，

$$i\frac{\partial}{\partial t}\psi(x,t)=\left(-\frac{1}{2m}\frac{\partial^2}{\partial x^2}+\frac{1}{2}m\omega_0^2x^2\right)\psi(x,t). \tag{8.20}$$

在位置表象下，我们求解的程序还是和一般的抽象的算符，或者任何其他表象，一样的。首先，求得 H 的本征值 E_n 和本征向量 $\phi_n(x)$，这里称为本征函数；然后，利用这些个本征函数加上时间项组合起来（$\left|\psi(t)\right\rangle=\sum_n c_n\left|\phi_n\right\rangle e^{-iE_nt}$，其中 $c_n=\left\langle\phi_n\right|\psi(0)\rangle$，或者合起来 $\left|\psi(t)\right\rangle=\sum_n\left|\phi_n\right\rangle e^{-iE_nt}\left\langle\phi_n\right|\psi(0)\rangle$），得到，

$$\psi(x,t)=\sum_n c_n\phi_n(x)e^{-iE_nt}, \tag{8.21}$$

其中

$$c_n=\int_{-\infty}^{\infty}dx\phi_n^*(x)\psi(x,0), \tag{8.22}$$

$\psi(x,0)$ 为初始状态对应着的波函数。

现在，我们来看看谐振子的本征波函数，其满足

$$\left(-\frac{1}{2m}\frac{d^2}{dx^2}+\frac{1}{2}m\omega_0^2x^2\right)\phi_n(x)=E_n\phi_n(x). \tag{8.23}$$

这个方程当然是可以想办法直接求解的。我们这里，作为一个例子，我们仅仅想考虑这个方程最简单的解（除了 $\phi_n(x)=0$ 以外）。我们看到，如果 $\phi_n(x)$ 的二阶导数刚好给出 x^2 项，那么方程左边的系数就会不依赖于 x，这样整个方程就简单了。也就是说大概 $\phi(x)=e^{\alpha x^2}$ 就行：$\frac{d^2}{dx}\phi(x)=2\alpha e^{\alpha x^2}+4\alpha^2x^2e^{\alpha x^2}$。于是 $\alpha^2=\frac{1}{4}m^2\omega_0^2\Longrightarrow\alpha=-\frac{1}{2}m\omega_0$（实际上正的 α 是没有物理意义的，在 x 很大的时候发散。物理系统一般要求 $\phi(x)$ 有限），而这个时候的本征能量 $E_0=\frac{1}{2}\omega_0$（能量下标 0 是有含义的。这里，我们不讨论这个问题）。因此，

$$\phi_0(x)=e^{-\frac{1}{2}m\omega_0x^2},E_0=\frac{1}{2}\omega_0. \tag{8.24}$$

系统如果处于这个状态，那么其时间演化为

$$\psi(x,t)=e^{-\frac{1}{2}m\omega_0x^2}e^{-i\frac{1}{2}\omega_0t}. \tag{8.25}$$

在这个例子中，我们讨论了位置表象下的 $\hat{x}$ 和 $\hat{p}$，然后利用他们的形式，求解了一维谐振子的基态波函数（我们没有证明这个是基态，在下一个例子中，我们会顺便补充这个证明）。我们还以谐振子为例讨论了正则量子化的一般手续。

下面，我们再举一个例子，还是求解一维谐振子的 Hamiltonian 的本征值和本征态，但是不在任何表象中求解，而是直接利用对易关系。这个解法也因此被称为代数解法。这个解法是科学发展历史上的一件艺术品，是很多量子力学的进一步理论的发展的基础。科学中有大美——用简单和统一的理论来描述这个世界，也有精巧的小美——妙手偶得的有时候还意义深远的优雅小品。

8.2 代数解法与能量表象

例 8.2 (**一维谐振子的代数解法**)：经典一维谐振子的 Hamiltonian 是 $H = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega_0^2\hat{x}^2$，求解其本征能量。

上一节我们已经利用 Shrödinger 绘景的运动方程——Shrödinger 把谐振子的本征值的问题转化为一个微分方程的本征函数的问题，见公式 (8.23)。这一节，我们用 Heisenberg 绘景的运动方程——Heisenberg 方程来求解同一个问题。在上一节中，我们已经能够注意到所有的计算都从最基本的对易关系，公式 (8.1) 开始。现在，我们还是从这里开始。

首先，让我们来写下算符 $\hat{x}^H$ 和 $\hat{p}^H$ 的演化方程。为了记号简单，对于算符 $\hat{x}^H, \hat{p}^H$，我们就直接用 x, p 了。注意，在这里它们是时间的函数。

$$\dot{x} = -i\left[x, H\right] = -i\frac{1}{2m}\left[x, p^2\right] = -i\frac{1}{2m}\left(p\left[x, p\right] + \left[x, p\right]p\right) = \frac{1}{m}p, \tag{8.26}$$

$$\dot{p} = -i\left[p, H\right] = -i\frac{1}{2}m\omega_0^2\left[p, x^2\right] = -m\omega_0^2 x. \tag{8.27}$$

联合这两个方程，我们得到

$$\ddot{x} = -\omega_0^2 x. \tag{8.28}$$

考虑到初始条件

$$x\left(0\right) = x^S, \tag{8.29}$$

$$p\left(0\right) = p^S, \tag{8.30}$$

我们有

$$x(t) = x^S \cos(\omega_0 t) + \frac{p^S}{m\omega_0} \sin(\omega_0 t), \tag{8.31}$$

$$p(t) = p^S \cos(\omega_0 t) - x^S m\omega_0 \sin(\omega_0 t). \tag{8.32}$$

于是，任何其他物理量都可以通过这两个算符得到。例如，

$$\begin{aligned} x^2(t) = {} & x^S x^S \cos^2(\omega_0 t) + \frac{p^S p^S}{m^2\omega_0^2} \sin^2(\omega_0 t) \\ & + \cos(\omega_0 t)\sin(\omega_0 t)\frac{p^S}{m\omega_0}\left(x^S p^S + p^S x^S\right). \end{aligned} \tag{8.33}$$

如果我们需要把理论的结果于实验的对比，那么，我们就可以从 $\langle x^s \rangle = \langle \psi(0) | x^s | \psi(0) \rangle$、$\langle p^s \rangle$、$\langle x^s x^s \rangle$、$\langle p^s p^s \rangle$ 等等得到所有的可观测量的理论值了。

到这里，量子力学的问题似乎也完全解决了。但是，我们却没有解开这个例题所问的问题：本征能量是多少？解决这个问题需要一些额外的技巧。

定义湮灭算符（这个名字之后会解释）

$$a = \frac{1}{\sqrt{2m\omega_0}}(m\omega_0 x + ip), \tag{8.34}$$

则产生算符（这个名字之后会解释）

$$a^\dagger = \frac{1}{\sqrt{2m\omega_0}}(m\omega_0 x - ip). \tag{8.35}$$

可以验证，

$$H = \omega_0\left(a^\dagger a + \frac{1}{2}\right), \tag{8.36}$$

而且

$$\left[a, a^\dagger\right] = 1. \tag{8.37}$$

按照我们上一节的经验，公式 (8.36) 和公式 (8.37) 就完整地定义了一个量子力学系统——**对于量子系统，我们只需要一个 Hamiltonian 和一个对易关系**。现在，我们开始讨论公式 (8.36)——这个新的形式下的谐振子。

第一步，我们来看一看产生湮灭算符的方程，

$$\dot{a} = -i\omega_0 a, \tag{8.38}$$

$$\dot{a}^\dagger = i\omega_0 a^\dagger. \tag{8.39}$$

于是，其解就是

$$a = e^{-i\omega_0 t} a^S, \tag{8.40}$$

$$a^\dagger = e^{i\omega_0 t} a^{\dagger,S}. \tag{8.41}$$

这一步，如果你仔细看的话，实际上就是把 x, p 满足的方程做了对角化。结果就是，得到的算符的方程非常简单，方程中仅仅包含自己这个算符，不包含其它算符。这个特征——不相互耦合的方程——对于求解方程非常重要。到这里，我们也就是再一次求解了一次 Heisenberg 方程。关键在于下面的步骤。

第二步，我们来证明：如果 $a^\dagger a$（与 H 仅仅相差一个常数）的本征值是 n 的话（相应的本征向量为 $|n\rangle$），$a|n\rangle$ 以及 $a^\dagger|n\rangle$ 也是 $a^\dagger a$ 的本征向量，仅仅本征值不一样。

$$a^\dagger a \cdot a|n\rangle = \left(a^\dagger a - aa^\dagger + aa^\dagger\right) \cdot a|n\rangle = (-1+n)\, a|n\rangle, \tag{8.42}$$

$$a^\dagger a \cdot a^\dagger|n\rangle = a^\dagger \left(aa^\dagger - a^\dagger a + a^\dagger a\right) \cdot a|n\rangle = (1+n)\, a^\dagger|n\rangle. \tag{8.43}$$

于是，看起来

$$a|n\rangle \propto |n-1\rangle, \tag{8.44}$$

$$a^\dagger|n\rangle \propto |n+1\rangle. \tag{8.45}$$

我们可以求得系数，

$$\left(\langle n| a^\dagger\right)(a|n\rangle) = n \Rightarrow a|n\rangle = \sqrt{n}\,|n-1\rangle. \tag{8.46}$$

以及类似的

$$a^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle. \tag{8.47}$$

现在，我们知道，H 的本征值就是这些 n，考虑到常数和系数，我们有

$$E_n = \omega_0\left(n + \frac{1}{2}\right). \tag{8.48}$$

第三步，我们来证明这些 n 只能够是非负整数。从某一个 n 开始，我们用算符 a，可以得到所有的本征值小于 n 的所有的本征向量，这个本征向量可以是无穷多个，

$$|n-1\rangle = \frac{1}{\sqrt{n}} a|n\rangle. \tag{8.49}$$

总有一天，这个本征向量的本征值会小于零，例如 $\left|n-[n+1]\right\rangle \propto a^{[n+1]}\left|n\right\rangle$。只要这个本征值小于零，那么

$$\left(\left\langle n-[n+1]\right| a^{\dagger}\right)\left(a\left|n-[n+1]\right\rangle\right)=n-[n+1]<0. \tag{8.50}$$

但是，$\left(\left\langle n-[n+1]\right| a^{\dagger}\right)\left(a\left|n-[n+1]\right\rangle\right)\geq 0$ 是一个向量自身的内积。于是，这个把本征值不停地减少的过程就不能一直做下去，必须在大于等于零的地方自然截断。唯一满足着个要求的可能就是 n 是非负整数，于是，$\sqrt{n-[n]}=0$。于是，如果继续减少本征值，就会出现

$$\left|-1\right\rangle \propto a\left|0\right\rangle = \sqrt{0}\left|-1\right\rangle = 0. \tag{8.51}$$

结论，H 的本征值是 $\omega_0\left(n+\frac{1}{2}\right)$，其中 n 为非负整数。

这个问题回答完了，但是这个求解方法具有一般的意义。算符 a 的含义是把本征值不停地减少，算符 $a^{\dagger}$ 的含义是把本征值不停地增加。如果我们把本征向量 $\left|n\right\rangle$ 看做是一个有 n 个能量粒子的状态，那么算符 a 的含义就是消灭粒子，算符 $a^{\dagger}$ 的含义是产生粒子。所以，它们被叫做湮灭和产生算符。更进一步，我们可以通过真空状态 $\left|0\right\rangle$ 和产生算符来定义所有的本征态，

$$\left|n\right\rangle = \frac{\left(a^{\dagger}\right)^{n}}{\sqrt{n!}}\left|0\right\rangle. \tag{8.52}$$

在能量本征态 $\{\left|n\right\rangle\}$ 构成的表象下，我们来计算几个算符。例如

$$H=\sum_{n} E_n\left|n\right\rangle\left\langle n\right| = \sum_{n}\left(n+\frac{1}{2}\right)\omega_0\left|n\right\rangle\left\langle n\right|, \tag{8.53}$$

$$a=\sum_{n} a\left|n\right\rangle\left\langle n\right| = \sum_{n}\sqrt{n}\left|n-1\right\rangle\left\langle n\right|, \tag{8.54}$$

而

$$a^{\dagger}=\sum_{n} a^{\dagger}\left|n\right\rangle\left\langle n\right| = \sum_{n}\sqrt{n+1}\left|n+1\right\rangle\left\langle n\right|. \tag{8.55}$$

注意，现在这些算符都成了矩阵。利用 $\hat{x},\hat{p}$ 和 $a,a^{\dagger}$ 算符的关系，我们就可以得到能量表象下算符 $\hat{x},\hat{p}$ 的形式。于是，$\hat{x},\hat{p}$ 也成了矩阵。

8.3 代数解法的更多细节：一切都是对易关系 *

这一节为选读内容。

回到我们一开始的 $\hat{x},\hat{p}$ 算符的形式，我们发现在谐振子的问题中的任意一个算符都能够用 $\hat{x},\hat{p}$ 来表达，而 $\hat{x},\hat{p}$ 可以用 $a,a^\dagger$ 来表达，例如

$$\hat{A}=A\left(a,a^\dagger\right)=\sum_{m=0,n=0}^{\infty}A_{m,n}\left(a^\dagger\right)^m a^n. \tag{8.56}$$

后者的时间演化又是完全已知的。于是，

$$\hat{A}^H(t)=\sum_{m=0,n=0}^{\infty}e^{-i(n-m)\omega_0 t}A_{m,n}\left(a^\dagger\right)^m a^n \tag{8.57}$$

给定任何一个初始状态——波函数或者密度矩阵，这里以波函数 $\left|\psi(0)\right\rangle$ 为例，肯定可以写成本征向量的叠加，也就是

$$\left|\psi(0)\right\rangle=\sum_n C_n\left|n\right\rangle=\sum_n C_n\frac{\left(a^\dagger\right)^n}{\sqrt{n!}}\left|0\right\rangle. \tag{8.58}$$

于是，算符 $\hat{A}$ 在任意时刻的平均值可以写成，

$$\begin{aligned}&\left\langle\psi(0)\right|\hat{A}^H(t)\left|\psi(0)\right\rangle\\=&\sum_i C_i^*\frac{a^i}{\sqrt{i!}}\left\langle 0\right|\sum_{m=0,n=0}^{\infty}e^{-i(n-m)\omega_0 t}A_{m,n}\left(a^\dagger\right)^m a^n\sum_j C_j\frac{\left(a^\dagger\right)^j}{\sqrt{j!}}\left|0\right\rangle\\=&\sum_{ijmn}e^{-i(n-m)\omega_0 t}C_i^*C_jA_{m,n}\frac{1}{\sqrt{i!j!}}\left\langle 0\right|a^i\left(a^\dagger\right)^m a^n\left(a^\dagger\right)^j\left|0\right\rangle.\end{aligned} \tag{8.59}$$

这样所有的物理量的计算的问题就成了计算

$$\left\langle 0\right|a^i\left(a^\dagger\right)^m a^n\left(a^\dagger\right)^j\left|0\right\rangle \tag{8.60}$$

的问题，而这个问题可以完全通过对易来求解——交换 a^n 和 $\left(a^\dagger\right)^j$ 的顺序，我们得到形如 $a\left|0\right\rangle$ 的表达式，它就是零；交换 a^i 和 $\left(a^\dagger\right)^m$ 的顺序，我们得到形如 $\left\langle 0\right|a^\dagger$ 的表达式，它就是零。于是，$\left\langle 0\right|a^i\left(a^\dagger\right)^m a^n\left(a^\dagger\right)^j\left|0\right\rangle$ 的非零部分就是通过一系列的对易关系计算以后得到的剩余项。

通过这个讨论，我们得到一个认识：关于一维谐振子的任意物理量在任何时刻的平均值都可以通过一系列算符 $a,a^\dagger$ 的对易关系得到。这个结论是一个非常强的结论，以后的进一步的量子理论例如量子场论的学习都从这里开始。这一节的内容有一定的难度，计算也跳过了很多步骤，是选学内容。

8.4 作业

习题 8.1 对于 $L=\frac{1}{2}m\dot{x}^2-V(x)$ 的系统，推导出来位置表象下的 Schrödinger 方程。

习题 8.2 对于 $L=\frac{1}{2}m\dot{x}^2-V(x)$ 的系统，推导出来动量表象下的 Schrödinger 方程。

习题 8.3 对于满足如下形式的 $V(x)$ 的量子系统，求解位置表象下的 Schrödinger 方程（求出本征向量，本征值，给定一般的初始条件的本征态组合系统的确定方式），

$$V(x)=\begin{cases}0 & |x|\leq 1\\ \infty & \text{otherwise}\end{cases} \tag{8.61}$$

这样的 V 被称作一维无线深势阱。

习题 8.4 独立（不看书，或者看完书以后）完成一维谐振子的代数解法。产生湮灭算符对于深入理解和进一步学习量子力学非常的重要。

习题 8.5 一维谐振子的算符 $\hat{x}$ 和 $\hat{p}$，还有 H 对应的 $A_{m,n}$ 的形式。

8.5 本章小结

这一章，我们学习了位置表象中的量子力学算符、态矢量和演化方程的形式，以及它们与一般的抽象的不在某个表象下的量子力学的形式的联系。

在这里，我们还发现，一个量子力学系统的最核心的内容是基本算符的对易关系——位置算符和动量算符的对易子不为零。基于这个对易关系的从经典力学到量子力学的过度叫做正则量子化。通过量子实验与经典概率论的对比，我们已经知道密度矩阵的非对角元对于解释量子现象的重要性。如果所有的算符都对易，那么它们存在共同本征向量，在这个本征向量构成的基矢下面，所有的算符只有对角元。因此，非对角元就没有存在的必要了。于是，量子力学的实验就不可能得到解释。在这个意义上，有的人说，量子力学的核心就是算符的非对易关系。当然，我们更加强调，量子力学的核心是，在一般的表象下，密度矩阵的非对角元的存在。两者是一致的。

除了正则量子化，一个经典系统的对应的量子理论还可以通过路径积分量子化等其它方法来构建。这里就不再讨论关于路径积分量子化的方法，

尽管正像 Feynman 说的，任何一个严肃的物理学家都应该在脑子里对一种理论存在多种等价的表述方式。我们这里提到的表象理论，是指一个抽象定义的算符，通常仅仅通过对易关系定义的算符，可以在某一个算符的本征矢量做基矢的情况下表示出来（把各个分量求出来）。这个时候，有可能一个算符在一套基矢下的形式和另外一套基矢下的形式完全不一样，但是它们表示同一个算符。更一般地来说，一个理论的不同形式，也可以看成是这个理论的不同表象。

再次强调，具体采用什么算符的本征态当作表象的基矢不重要，抽象算符之间的对易关系最重要。

第四部分

耦合量子系统：相互作用与纠缠

第九章　耦合量子系统的状态与测量

我们已经学习了量子力学的基本内容：状态是 Hilbert 空间内的矢量或者跟一般地密度矩阵，物理量是这个空间上的算符，测量就是建立一个从这个密度算符到一个以被测量算符的本征值和本征态构成的事件空间的概率分布的映射，演化就是给定一个外界对系统的作用形式 H 以后得到状态的一个以时间为参数的幺正变换。一直到现在，我们考虑的系统都是一个二维的可以看作是一个叫自旋的东西。现在，在这个部分，我们来考虑两个自旋构成的耦合系统。

按照经典概率论，从一个硬币的状态到两个硬币状态，概率论还是原来的概率论，唯一需要额外考虑的东西就是两个自旋的关联，也就是 $P(S_1, S_2)$ 是否等于 $P(S_1)P(S_2)$。对于量子系统，我们来看一下，这个时候需要额外考虑的是什么。

9.1　直积空间

正如两个硬币的状态是 $2^2 = 4$ 个（正正、正反、反正、反反），两个自旋的状态的空间维数也是两个各自的自旋的状态的维数的乘积 $4 = 2^2$，基本的状态——也就是空间的一组自然的基矢也可以写成如下四个，

$$|\uparrow_z\uparrow_z\rangle, |\uparrow_z\downarrow_z\rangle, |\downarrow_z\uparrow_z\rangle, |\downarrow_z\downarrow_z\rangle, \tag{9.1}$$

其中写在第一和第二位置的自旋状态符号分别表示第一个和第二个自旋的状态。这 4 个基矢中的每一个都是原来的二维空间的基矢的直积形式，因此有的时候被称作直积基矢。由于量子系统状态矢量的可叠加性（硬币没有），

这个 4 维空间的基矢还可以换成另外的一组，例如，

$$|\uparrow_z\uparrow_z\rangle, \frac{|\uparrow_z\downarrow_z\rangle + |\downarrow_z\uparrow_z\rangle}{\sqrt{2}}, |\downarrow_z\downarrow_z\rangle \ \frac{|\uparrow_z\downarrow_z\rangle - |\downarrow_z\uparrow_z\rangle}{\sqrt{2}}. \tag{9.2}$$

这一组基矢的特点是，当我们交换两个自旋的编号的时候，前面的 3 个不变，最后的那个改变一个符号。这样的特殊的对称性在很多问题中有特别的意义。前面三个被称为对称基矢，后面的一个被称为反称基矢。有的时候，我们也会采用一组叫做 Bell 态的基矢，

$$\frac{|\uparrow_z\uparrow_z\rangle + |\downarrow_z\downarrow_z\rangle}{\sqrt{2}}, \frac{|\uparrow_z\uparrow_z\rangle - |\downarrow_z\downarrow_z\rangle}{\sqrt{2}}, \frac{|\uparrow_z\downarrow_z\rangle + |\downarrow_z\uparrow_z\rangle}{\sqrt{2}}, \frac{|\uparrow_z\downarrow_z\rangle - |\downarrow_z\uparrow_z\rangle}{\sqrt{2}}. \tag{9.3}$$

这组基矢的对称性与上一组一样。在下面的讨论中，我们会用到这组基矢。有的时候，我们用一组特定的符号来表示这四个 Bell 态，分别记为 $\left|\Phi^+\right\rangle$、$|\Phi^-\rangle$、$\left|\Psi^+\right\rangle$、$|\Psi^-\rangle$。

这四个基矢构成的 4 维的 Hilbert 空间实际上是两个 2 维的 Hilbert 空间的直积，$H = H^1 \otimes H^2$。一个一般的状态矢量可以是，

$$|\psi\rangle = \alpha\left|\Phi^+\right\rangle + \beta\left|\Phi^-\right\rangle + \gamma\left|\Psi^+\right\rangle + \eta\left|\Psi^-\right\rangle, \tag{9.4}$$

其中 $\left|\alpha^2\right| + \left|\beta^2\right| + \left|\gamma^2\right| + \left|\eta^2\right| = 1$。这个空间的状态也可以用第一组直积形式的基矢来表达，例如

$$|\psi\rangle = \alpha'\,|\uparrow_z\uparrow_z\rangle + \beta'\,|\uparrow_z\downarrow_z\rangle + \gamma'\,|\downarrow_z\uparrow_z\rangle + \eta'\,|\downarrow_z\downarrow_z\rangle\,. \tag{9.5}$$

这样的一个状态一般不能写成两个二维空间矢量的直积，也就是

$$|\psi\rangle \neq \left|\psi^1\right\rangle\left|\psi^2\right\rangle. \tag{9.6}$$

当 $\alpha' = 1$，其它分量等于零的时候，

$$|\psi\rangle = |\uparrow_z\uparrow_z\rangle, \tag{9.7}$$

是一个直积状态。当 α' 和 η' 不等于零，其它等于零的时候，就不是直积状态。在 Bell 态作为基矢的展开形式下，就算其中的一个分量唯一，其它都为零——也就是状态就是 Bell 态本身——也不是直积状态。这个直积态和非直积态的区别与经典概率论的独立分布函数和关联分布函数非常像，但

是我们很快会看见它们的区别。更一般地来说，用密度矩阵的语言，满足下面等式的状态就叫做直积态，否则称为非直积态，

$$\rho = \rho^1 \otimes \rho^2, \tag{9.8}$$

其中 $\rho^i = tr^{(-i)}(\rho)$，而 $tr^{(-i)}$ 表示对除了 i 系统的自由度之外的其它自由度都求迹。

在经典概率论中，非直积态被称为关联态。在量子力学中，非直积态有的时候也被称为关联态，或者纠缠态。在有的物理学家的语言中，纠缠态是一个更严格的概念，代表了更小的状态集合。这个更狭义的纠缠态的定义我们会在本节的稍后一点介绍。

有了耦合系统的状态的一般概念之后，我们按照我们之前讨论量子系统的惯例，来讨论耦合量子系统的测量和状态演化。在此之前，我们先来回顾一下经典耦合系统的测量。

9.2 经典关联态的测量

我们从最简单的开始，测量两个独立的完全对称的硬币，其状态为

$$\rho^{12,C} = \frac{1}{4}|\uparrow\uparrow\rangle\langle\uparrow\uparrow| + \frac{1}{4}|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + \frac{1}{4}|\downarrow\uparrow\rangle\langle\downarrow\uparrow| + \frac{1}{4}|\downarrow\downarrow\rangle\langle\downarrow\downarrow|. \tag{9.9}$$

测量得到的可能的状态是

$$|\uparrow\uparrow\rangle\langle\uparrow\uparrow|, |\uparrow\downarrow\rangle\langle\uparrow\downarrow|, |\downarrow\uparrow\rangle\langle\downarrow\uparrow|, |\downarrow\downarrow\rangle\langle\downarrow\downarrow| \tag{9.10}$$

这四个，得到这些状态的几率分别是

$$P_{\uparrow\uparrow}, P_{\uparrow\downarrow}, P_{\downarrow\uparrow}, P_{\downarrow\downarrow}. \tag{9.11}$$

它们的取值都是 $\frac{1}{4}$。

于是，当我们发现第一个硬币是正面的时候，我们得到第二个硬币的状态为 S^2 概率为，

$$P\left(S^2 \mid S^1 = \uparrow\right) = \frac{P(S^2, S^1 = \uparrow)}{P(S^1 = \uparrow)} = \frac{1}{2}. \tag{9.12}$$

也就是说，知道了关于第一个硬币的知识不增加我们任何关于第二个硬币的知识。这也就是独立硬币的含义。

现在，我们来考虑两个用一根神奇的“红线”连接在一起的硬币，它们俩总是显示相同的一面（我们暂时忽略在经典硬币上是如何实现这个“总是显示相同”这个问题），也就是

$$\rho^{12,C} = \frac{1}{2}|\uparrow\uparrow\rangle\langle\uparrow\uparrow| + \frac{1}{2}|\downarrow\downarrow\rangle\langle\downarrow\downarrow|. \tag{9.13}$$

测量得到的可能的状态是

$$|\uparrow\uparrow\rangle\langle\uparrow\uparrow|, |\downarrow\downarrow\rangle\langle\downarrow\downarrow| \tag{9.14}$$

这两个，得到这些状态的几率分别是

$$P_{\uparrow\uparrow}, P_{\downarrow\downarrow}. \tag{9.15}$$

它们的取值都是 $\frac{1}{2}$。

于是，当我们发现第一个硬币是正面的时候，我们得到第二个硬币的状态为 S^2 概率为，

$$P\left(S^2 = \uparrow \mid S^1 = \uparrow\right) = \frac{P\left(S^2 = \uparrow, S^1 = \uparrow\right)}{P\left(S^1 = \uparrow\right)} = \frac{P\left(S^2 = \uparrow, S^1 = \uparrow\right)}{P\left(S^2 = \uparrow, S^1 = \uparrow\right)} = 1. \tag{9.16}$$

在这里我们发现，当知道第一个硬币的状态是向上的时候，第二个硬币的状态必然向上，不可能向下。这个也就是完全相同（两个状态百分之一百关联在一起）的硬币的含义。考虑两个完全相反的硬币，我们可以得到类似的结果。

这样的完全关联在一起的随机变量是一个可以想象的东西，而且，想象这样一个东西的难度不会比想象一个随机的硬币的难度高，尽管想象一个随机的硬币已经不是一件简单的事情。如果有一个随机的硬币，我们找出来其中的随机的核心——例如一个随机数生成器，然后利用这个生成器产生的数据同时来控制这两个硬币的状态就可以了。因此，如果我们假设一份经典随机数可以用来控制两个系统的状态，那么，想象两个关联在一起的硬币的难度，完全就是想象一个随机硬币的难度。对于存在一个真随机的硬币，我们没有觉得多么不可思议。于是，对于存在两个完全关联的硬币，我们也没有觉得多么不可思议。顺便说一下，这个“一份随机数可以用来控制两个系统”的假设，实际上依赖于经典信号的可复制性——给定一个经典信号 0 或者 1，我们可以在不用观察得到这个信号的真值的条件下，复制一个相同的信号。以后，我们还会回到这个问题。

下面，我们来看耦合起来的两个量子自旋的测量。

9.3 量子纠缠态的测量

我们来考虑两个处于下面这个状态的自旋的测量，

$$\rho^{12,Q} = \frac{1}{2}\left(|\uparrow_z\uparrow_z\rangle + |\downarrow_z\downarrow_z\rangle\right)\left(\langle\uparrow_z\uparrow_z| + \langle\downarrow_z\downarrow_z|\right). \tag{9.17}$$

例如，我们考虑测量第一个自旋的 z 方向，第二个自旋的 z 方向。对于这个测量，我们的相当于测量一组相互对易的两个算符的本征值和相应的几率，

$$O_1 = \sigma_z^1 \otimes I^2, O_2 = I^1 \otimes \sigma_z^2. \tag{9.18}$$

其本征向量分别为如下四个（这里我们取了相互对易的两个算符的共同本征态，证明这些状态就是这样的共同本征态的任务留作练习），

$$|\uparrow_z\uparrow_z\rangle, |\uparrow_z\downarrow_z\rangle, |\downarrow_z\uparrow_z\rangle, |\downarrow_z\downarrow_z\rangle. \tag{9.19}$$

于是这四个状态出现的相应的几率分别是，

$$P_{\uparrow_z\uparrow_z} = \langle\uparrow_z\uparrow_z|\rho^{12,Q}|\uparrow_z\uparrow_z\rangle = \frac{1}{2}, \tag{9.20a}$$

$$P_{\uparrow_z\downarrow_z} = \langle\uparrow_z\downarrow_z|\rho^{12,Q}|\uparrow_z\downarrow_z\rangle = 0, \tag{9.20b}$$

$$P_{\downarrow_z\uparrow_z} = \langle\downarrow_z\uparrow_z|\rho^{12,Q}|\downarrow_z\uparrow_z\rangle = 0, \tag{9.20c}$$

$$P_{\downarrow_z\downarrow_z} = \langle\downarrow_z\downarrow_z|\rho^{12,Q}|\downarrow_z\downarrow_z\rangle = \frac{1}{2}. \tag{9.20d}$$

也就是得到的结果是两个自旋都向上的可能性是 $\frac{1}{2}$，都向下的可能性也是 $\frac{1}{2}$，但是不会有不相同的时候。这个结论和上面那个测量两个完全关联的硬币是一样的。我们还可以得到更有意思的结果。

9.3.1 计算视角一：一组对易算符的测量

假设我们测量一个自旋的 x 方向，第二个自旋的 x 方向。对于这个测量，我们的相当于测量一组相互对易的两个算符的本征值和相应的几率，

$$O_1 = \sigma_x^1 \otimes I^2, O_2 = I^1 \otimes \sigma_x^2. \tag{9.21}$$

其本征向量分别为如下四个，

$$|\uparrow_x\uparrow_x\rangle, |\uparrow_x\downarrow_x\rangle, |\downarrow_x\uparrow_x\rangle, |\downarrow_x\downarrow_x\rangle. \tag{9.22}$$

于是这四个状态出现的相应的几率分别是，

$$P_{\uparrow_x\uparrow_x} = \langle\uparrow_x\uparrow_x|\rho^{12,Q}|\uparrow_x\uparrow_x\rangle = \frac{1}{2}, \tag{9.23a}$$

$$P_{\uparrow_x\downarrow_x} = \langle\uparrow_x\downarrow_x|\rho^{12,Q}|\uparrow_x\downarrow_x\rangle = 0, \tag{9.23b}$$

$$P_{\downarrow_x\uparrow_x} = \langle\downarrow_x\uparrow_x|\rho^{12,Q}|\downarrow_x\uparrow_x\rangle = 0, \tag{9.23c}$$

$$P_{\downarrow_x\downarrow_x} = \langle\downarrow_x\downarrow_x|\rho^{12,Q}|\downarrow_x\downarrow_x\rangle = \frac{1}{2}. \tag{9.23d}$$

也就是得到的结果是两个自旋都向上的可能性是 $\frac{1}{2}$，都向下的可能性也是 $\frac{1}{2}$，但是不会有不相同的时候。这个结果和同时测量两个自旋的 z 方向完全一致。

我们举例计算其中的第二个等式，公式 (9.23b)，

$$P_{\uparrow_x\downarrow_x} = \langle\uparrow_x\downarrow_x|\rho^{12,Q}|\uparrow_x\downarrow_x\rangle = \frac{1}{2}\langle\uparrow_x\downarrow_x|\left(|\uparrow_z\uparrow_z\rangle\langle\uparrow_z\uparrow_z| + |\downarrow_z\downarrow_z\rangle\langle\downarrow_z\downarrow_z|\right)|\uparrow_x\downarrow_x\rangle \tag{9.24a}$$

$$+\frac{1}{2}\langle\uparrow_x\downarrow_x|\left(|\uparrow_z\uparrow_z\rangle\langle\downarrow_z\downarrow_z| + |\downarrow_z\downarrow_z\rangle\langle\uparrow_z\uparrow_z|\right)|\uparrow_x\downarrow_x\rangle \tag{9.24b}$$

$$= \frac{1}{4} - \frac{1}{4} = 0 \tag{9.24c}$$

其中第一部分和第二部分相互抵消。

假设，我们只有第一部分那个对角项公式 (9.24a)，也就是

$$\rho^{12} = \frac{1}{2}\left(|\uparrow_z\uparrow_z\rangle\langle\uparrow_z\uparrow_z| + |\downarrow_z\downarrow_z\rangle\langle\downarrow_z\downarrow_z|\right). \tag{9.25}$$

那么，我们就会得到

$$P_{\uparrow_x\downarrow_x} = \langle\uparrow_x\downarrow_x|\rho^{12}|\uparrow_x\downarrow_x\rangle = \frac{1}{4}. \tag{9.26}$$

也就是说，如果让我们的两个自旋的量子态长的跟两个经典硬币的关联状态一样，那么，我们的测量结果就会完全不一样。在量子态 $\rho^{12,Q}$ 的情形，对于任何一个方向的对两个自旋的同时同方向测量（这里我们只计算了同时测量 z 和同时测量 x）我们都得到两个自旋的方向完全相同的结果。在量子态 ρ^{12}（看起来很像很像 $\rho^{12,C}$）的情形，对于两个自旋的同时同方向测量我们有的时候得到两个自旋的方向完全相同的结果（例如测量 z 方向），有的时候得到两个自旋的方向不同的结果（例如测量 x 方向）。

在得到这个结果的计算过程中，我们注意到最关键的地方在于矢量之间的内积 $\langle\uparrow_x\downarrow_x|\downarrow_z\downarrow_z\rangle$ 不等于非 0 即 1，而是由某个线性叠加关系给出，例

如 $|\uparrow_x\rangle = \frac{1}{\sqrt{2}}|\uparrow_z\rangle + \frac{1}{\sqrt{2}}|\downarrow_z\rangle$。这个叠加关系正是我门一直以来强调的量子力学和经典力学的不同：量子系统的状态之间允许做线性叠加。

也就是说，**状态 $\rho^{12,Q}$ 具有对两个自旋做各方向的配对测量都得到相同状态的性质——两个自旋的状态在各方向上的同时测量的结果都完全相同**，但是ρ^{12} **没有这个性质——两个自旋可以有相反的状态**，而两者之间的差别仅仅是非对角元 $\frac{1}{2}(|\uparrow_z\uparrow_z\rangle\langle\downarrow_z\downarrow_z| + |\downarrow_z\downarrow_z\rangle\langle\uparrow_z\uparrow_z|)$ 存在与否。$\rho^{12,Q}$ 这种在各个方向上都完全相关的量子态被称为纠缠态。ρ^{12} 这种在某个方向上完全相关的量子态被称为关联态。能够表达成两个系统的状态的密度矩阵的直积的状态称为直积态，或者独立状态。实际上，纠缠态可以比这个各个方向完全相关广泛。甚至可以定义纠缠度——给定一个状态，就算出来纠缠的程度是多少。更一般地来说，非直积态都可以被认为是相关态，所有的相关态中除去经典关联态（可以表达成为 ρ^{12} 这样的在各个自旋的某种基矢的乘积作为整个系统的基矢的表象下完全对角的状态），都是纠缠态。本书中，我们只关心 $\rho^{12,Q}$ 这样的完全纠缠态。

历史上，当大家发现完全纠缠态的这个性质——在任意同一个方向上做量子自旋的测量得到状态相同的结果——的时候，很多科学家都认为这是一件非常神奇的事情。现在，我们看见，这件事情的本质，仍然是非对角元，仍然是量子态的叠加原理。如果仅仅从关联的建立的角度，两个粘在一起的硬币也能够得到观测值完全相同（或者相反，取决于如何粘）的结果。最关键的地方不是存在着关联的测量结果，而是“在任何一个方向上”都可以得到关联。因此，量子纠缠的特性，还是来源于量子相干性，或者说态叠加原理，或者说非对角元的存在，或者说算符的非对易性，或者说 Hilbert 空间允许的矢量加法。这些表述都是等价的。纠缠仅仅是这些个等价的原理的结果和表现。

9.3.2 计算视角二：“先后”测量

在上面测量结果的计算中，对于观测者一——称为 Alice——测量第一个自旋的 σ_x 方向，观测者二——称为 Bob——测量第二个自旋的 σ_x 方向这件事情，我们采用了一组对易的关于两个自旋系统的算符 O_1, O_2 来整体描述。现在，我们用另一个视角——先让 Alice 测量第一个自旋，然后再让 Bob 测量第二个自旋——来看待这件事情，看看算出来这个测量会得到什么结果。也就是说，我们先看 $O_1 = \sigma_x^1 \otimes I^2$ 的测量结果，然后再来看在这个测量结果的基础上，做 σ_x^2 的测量的结果。对于这个计算，显然我们的依据

是量子力学的公理3 和公理4，而按照这两个公理，只要我们把密度矩阵 ρ 写成所测量的算符——这里主要是 σ_x^1——的本征向量的展开，我们就可以直接把测量结果和测量后状态读出来。现在，我们来做一下这个。为了简化记号，由于我们的密度矩阵 ρ^{12} 实际上是一个纯态，也就是说存在 $|\Psi\rangle$ 使得 $\rho^{12} = |\Psi\rangle\langle\Psi|$，我们直接在 $|\Psi\rangle$ 上做计算。另外，为了简化符号，x（z）方向的本征态，我们用 $|\pm\rangle$（$\uparrow\downarrow$）来表示。

$$
\begin{aligned}
\left|\Psi^{12}\right\rangle = & \frac{\sqrt{2}}{2}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right) \\
= & \frac{\sqrt{2}}{2}\left(\frac{\sqrt{2}}{2}\left(|+\rangle + |-\rangle\right)|\uparrow\rangle + \frac{\sqrt{2}}{2}\left(|+\rangle - |-\rangle\right)|\uparrow\rangle\right) \\
= & \frac{\sqrt{2}}{2}\left(|+\rangle\,\frac{\sqrt{2}}{2}\left(|\uparrow\rangle + |\downarrow\rangle\right) + |-\rangle\,\frac{\sqrt{2}}{2}\left(|\uparrow\rangle - |\downarrow\rangle\right)\right) \\
= & \frac{\sqrt{2}}{2}\left(|+\rangle|+\rangle + |-\rangle|-\rangle\right). \qquad (9.27)
\end{aligned}
$$

于是，由于测量的算符是 σ_x^1，我们发现得到的结果有两个：$|+\rangle$ 或者 $|-\rangle$，各自的几率是 $\frac{1}{2}$ 和 $\frac{1}{2}$。如果观测得到 $|+\rangle$（$|-\rangle$），则系统在测量之后的状态是

$$
|+\rangle|+\rangle\left(|-\rangle|-\rangle\right). \qquad (9.28)
$$

接着在这个状态上测量第二个自旋的 x 方向，我们得到测量结果 $100\%|+\rangle$（$100\%|-\rangle$）。

在这个计算里面，我们不去寻找 O_1, O_2 算符的共同本征态，而是先看看 Alice 测量的结果，再来看 Bob 在 Alice 测量之后的状态上测量的结果。我们可以反过来，先看 Bob，再看 Alice。结果是一样的。这个给我们一个启示，在上面的计算中所谓的先后，仅仅是逻辑上的先后，不是实际物理过程的先后。实际上，Bob 测量的方向不一定和 Alice 的一致，这个时候上面的计算（或者看作两个对易算符 O_1, O_2 的计算），也能够给出与实验相符的结果。这个留做作业。如果看作是物理过程的先后，我们上面的计算“可以这样来理解”：第二个自旋“了解到”Alice 测量的方向，然后按照 Alice 的结果向上还是向下来“决定”呈现给 Bob 状态 $|+\rangle$ 还是状态 $|-\rangle$。于是，看起来好像第二个自旋跟 Alice 之间可能存在某种信息的交流。或者反过来，第一个自旋和 Bob 之间存在着某种信息交流。在真实的实验中 [45]，两个自旋可以相距非常远，两个观测者做出观测的时间间隔可以非常小，以至于，如果存在这样的交流，这个交流必须是以超光速的方式来传递的。当然，我

们看到，这个看起来像个问题的问题，并不是量子力学的理论结构导致的，如果把 Alice 和 Bob 的测量看作是整体系统的测量（一组对易算符 O_1, O_2），量子理论完全解释了实验观察。这个问题是我们企图用先分析一个观测者然后再分析另外一个观测者，而且企图通过这个先后测量来“理解”这个计算结果导致的。实际上，量子力学给出来的对测量结果的预测是一样的，跟谁先谁后无关。

也就是说，不管从整体测量的角度来运用公理3和公理4，还是从先考虑某一个观测者然后另外一个的角度，给出来的结果都符合实验观察。存在问题的地方在于我们去“理解”这个量子现象的企图。顺便，我们指出来，在经典关联态的测量中，如果我们从真随机的经典客体的角度来“理解”，那么，看起来第二个硬币也需要先了解第一个观测者的结果，因此，也存在这个信息交流的问题。当然，在经典的情形下，第二个硬币只需要知道测量的结果——向上还是向下，而在量子的情形下，第二个自旋需要知道 Alice 测量的方向——所测量的自旋的空间方向 $\hat{r}$，以及测量的结果——向上还是向下。

这个多出来的需要知道 Alice 所测量的自旋的空间方向 $\hat{r}$ 的根本原因，还是量子态可以做表象变换，也就是线性叠加运算，于是对于一个给定方向的本征态的测量可以是在任意方向上的。于是，纠缠态区别与经典关联态的地方还是态叠加原理。

我们已经看到，除了上经典关联态已经具有的神奇之处——长程关联，纠缠态的独特特性——对两个自旋做任意一对方向上的测量结果之间都存在关联——的根本原因还是量子态的叠加原理，或者说表象变换，或者说不是所有算符都对易。当然，以后我们会看到，这样的纠缠，在具体应用上确实有独特的地方，可以成为很多量子信息问题和量子计算问题的基础。这一章的主要任务是说明在理论上纠缠是什么。纠缠态的上面这个属性和测量的时候两个自旋的距离没有关系，可以相距很远很远。这个时候，有的人就会问那么如何来理解一个自旋知道另外一个自旋已经被测量而且得到某个状态的结果呢，“因为这个自旋必须知道这个情况才能决定在自己被测量的时候显示什么状态呀”？例如，“第二个自旋怎么知道第一个的自旋被测量的是 z 方向并且得到结果向上，于是把自己的状态调整为向上态”。这其实是一个不好的问题。没有什么理由这个自旋必须知道另一个自旋被测量的情况才能决定自身的被侧量结果。在经典关联硬币的例子中，两个硬币，只要关联没有被破坏，也可以相距很远很远。如果这个时候，你不觉得理解上有

问题，那么在量子自旋的情况，你也不应该觉得有问题。真正量子和经典的不同不在于出现与否长程关联，而在于量子的纠缠态能够保证任何一组相同的方向的测量都有关联的结果，而经典的情形仅仅能够保证在特定的方向上有关联的结果。而给定一个状态，对任何一个方向上的测量都能够给出正确的结果，这件事情，量子力学是通过引入态叠加（等价于表象变换、算符非对易、用有非对角元的密度矩阵代替密度分布函数）来实现的。

9.3.3 计算视角三：可分辨性

在本小结的最后，我们用另一个方式来运用公理3和公理4对纠缠态的测量给出一个计算结果。对于同一个问题，不同的呈现方式有的时候可以给我们不同的启发。

对于 $O_1 = \sigma_x^1 \otimes I^2$ 的测量，我们来计算这个算符的本征向量。由于简并的本征值（一个本征值对应着多个本征向量）我们需要认为引入额外的记号来区分这些简并态：$|+\rangle|r_+\rangle, |+\rangle|r_-\rangle, |-\rangle|r_+\rangle, |-\rangle|r_-\rangle$。其中 $|r_\pm\rangle$ 是第二个自旋的状态空间中的任意一组相互正交的态。可以验证，这四个矢量确实是算符 O_1 的本征矢量。按照公理3，我们先来计算概率，

$$P_{+r_+} = \langle r_+|\langle +|\rho^{12}|+\rangle|r_+\rangle = \frac{1}{2}\langle r_+|\left(|+\rangle\langle +|\right)|r_+\rangle \tag{9.29}$$

$$P_{+r_-} = \langle r_-|\langle +|\rho^{12}|+\rangle|r_-\rangle = \frac{1}{2}\langle r_-|\left(|+\rangle\langle +|\right)|r_-\rangle \tag{9.30}$$

$$P_{-r_+} = \langle r_+|\langle -|\rho^{12}|-\rangle|r_+\rangle = \frac{1}{2}\langle r_+|\left(|-\rangle\langle -|\right)|r_+\rangle \tag{9.31}$$

$$P_{-r_-} = \langle r_-|\langle -|\rho^{12}|-\rangle|r_-\rangle = \frac{1}{2}\langle r_-|\left(|-\rangle\langle -|\right)|r_-\rangle. \tag{9.32}$$

毫无疑问，这四个数都是非负实数，但是，具体的值，依赖于我们选择的 $|r_\pm\rangle$。看起来，我们的测量结果依赖于一个任意的选择！

实际上，当我们关注 Alice 的测量的时候，我们关心的是

$$P_+ = P_{+r_+} + P_{+r_-}. \tag{9.33}$$

于是，

$$P_+ = \frac{1}{2}\langle r_+|\left(|+\rangle\langle +|\right)|r_+\rangle + \frac{1}{2}\langle r_-|\left(|+\rangle\langle +|\right)|r_-\rangle \tag{9.34}$$

$$= \frac{1}{2}tr\left(|+\rangle\langle +|\right)$$

$$= \frac{1}{2}. \tag{9.35}$$

这个是与实验相符的结果。也就是说，对于 Alice 测量的结果来说，我们不关心第二个自旋处于什么状态。

下面来计算 Alice 观测结果得到向上的时候的测量后状态：当我们观测结果是 $+r_+$ 的时候，我们得到相应状态，$|+\rangle\left|r_+\right\rangle$；当我们观测结果是 $+r_-$ 的时候，我们得到相应状态，$|+\rangle|r_-\rangle$。现在的问题是如何把这两个结果合起来，下面的两个密度矩阵的哪一个？是

$$\begin{aligned}\rho^f &= P_{+r_+}\left|+r_+\right\rangle\left\langle+r_+\right| + P_{+r_-}|+r_-\rangle\langle+r_-| \\ &= (|+\rangle\langle+|)\left(P_{+r_+}\left|r_+\right\rangle\left\langle r_+\right| + P_{+r_-}|r_-\rangle\langle r_-|\right),\end{aligned} \tag{9.36}$$

还是，

$$\begin{aligned}\left|\Psi^f\right\rangle &= \left\langle r_+\right| +\rangle\left|+r_+\right\rangle + \langle r_-| +\rangle|+r_-\rangle \\ &= (|+\rangle)\left(\left|r_+\right\rangle\left\langle r_+\right| +\rangle + |r_-\rangle\langle r_-| +\rangle\right) \\ &= (|+\rangle)(|+\rangle).\end{aligned} \tag{9.37}$$

按照实验的结果，我们知道当然公式 (9.37) 与实验相符——当 Alice 测量得到向上的时候，整体系统的状态是 $(|+\rangle)(|+\rangle)$。公式 (9.36) 和公式 (9.37) 理论上的区别在于：前者是概率叠加——把两个密度矩阵相加，不存在交叉项，后者是相干性叠加——把两个态矢量相加，于是转化成密度矩阵的语言的时候存在交叉项。那为什么这里相干性叠加给出正确的结果，概率性叠加不能呢？因为在讨论 Alice 的测量结果的时候，第二个自旋的状态是不可区分的。这个就好像是在讨论 which-way 实验的时候一样，对于不能区分的状态，我们用相干性叠加。

我们看到这三个计算的视角不一样，但是结果是完全一样的。给我们的启发也不一样：对于完全确定的没有简并的本征态的耦合系统上的测量算符集合（就算看起来是单个系统的测量也需要写成耦合系统上的测量算符），直接运用公理3和公理4给出结果就行；如果我们把耦合系统的测量看作是逻辑上先后的测量，那么这个先后不可能是物理上的时间因果过程，仅仅是逻辑上的，因为无论哪个为先哪个当后，给出的测量结果预测是一样的；对于存在简并本征态的耦合系统上的测量算符，在运用公理4写下测量后状态的时候，需要考虑测量结果的是否可分辨。不可分辨的时候测量后状态需要用矢量叠加，可分辨的时候用概率叠加。

总而言之，第一，对于耦合系统，公理3和公理4仍然给出与实验相符的结果；第二，量子纠缠态的各个方向的测量都有关联的特性是态叠加原理的

结果；第三，有的时候“理解”量子力学是个问题，尤其是用熟悉了经典世界的人的理解能力来理解的时候。

9.4 作业

习题 9.1 验证 $|\psi\rangle = \alpha' |\uparrow_z\uparrow_z\rangle + \eta' |\downarrow_z\downarrow_z\rangle$ 不可能是直积状态。

习题 9.2 对易算符的本征向量：计算 $O_1 = \sigma_z^1 \otimes I^2, O_2 = I^1 \otimes \sigma_z^2$ 的共同本征向量。给定状态 $|\Phi_-\rangle$，按照这些共同本征态来计算 Alice 测量 σ_z，Bob 测量 σ_z 的结果。

习题 9.3 对易算符的本征向量：计算 $O_1 = \sigma_z^1 \otimes I^2, O_2 = I^1 \otimes \sigma_x^2$ 的共同本征向量。给定状态 $|\Phi_-\rangle$，按照这些共同本征态来计算 Alice 测量 σ_z，Bob 测量 σ_x 的结果。

习题 9.4 先后测量：给定状态 $|\Phi_-\rangle$，计算 Alice 先测量 σ_z，Bob 再测量 σ_x 的结果，以及反过来计算 Bob 先测量 σ_x，Alice 再测量 σ_z 的结果。

习题 9.5 纠缠态的非对角元：计算 $\langle\uparrow_x\downarrow_x|\rho^{12,Q}|\uparrow_x\downarrow_x\rangle$ 和 $\langle\uparrow_x\downarrow_x|\rho^{12}|\uparrow_x\downarrow_x\rangle$。其中 $\rho^{12,Q} = \frac{1}{2}\left(|\uparrow_z\uparrow_z\rangle + |\downarrow_z\downarrow_z\rangle\right)\left(\langle\uparrow_z\uparrow_z| + \langle\downarrow_z\downarrow_z|\right)$，$\rho^{12} = \frac{1}{2}\left(|\uparrow_z\uparrow_z\rangle\langle\uparrow_z\uparrow_z| + |\downarrow_z\downarrow_z\rangle\langle\downarrow_z\downarrow_z|\right)$。

9.5 本章小结

本章中我们主要介绍了两个自旋的状态空间是什么——4 维 Hilbert 空间，这个空间和两个自旋的本身的空间的关系是什么——直积空间。然后我们介绍了经典关联态的测量和量子纠缠态的测量。其中，Bell 态的定义需要了解，在以后的章节中我们会用到。对于纠缠态的测量，我们用不同的计算方式，根据公理3和公理4，得到了与实验相符的结果。同时，通过这些计算，我们发现，纠缠态的除了与经典关联一样的具有长程关联之外的特殊性质——对两个自旋做**任意一对方向**上的测量，其结果都有关联，就是态叠加原理的结果。

第十章　耦合量子系统的演化与纠缠

这一章，我们来回答耦合系统的状态如何在这个 4 维 Hilbert 空间里面演化，这样的演化如何计算，一个系统能够从非纠缠的状态到达纠缠的状态吗？量子力学的基本理论不会随着 Hilbert 空间维数的变化而变化，所有之前学习过的量子力学的一般理论都成立。

10.1　演化以及演化导致的纠缠

量子力学的一般理论告诉我们从系统的 Halmitonian 开始，演化算符是 $U = e^{-iHt}$，而系统的 Hamiltonian 对应着外界对系统或者系统的各个部分之间的相互作用。我们还是利用具体的例子来展示一个独立状态如何可以演化成为一个纠缠态。

例 10.1 (**演化导致的纠缠**)：在基矢集合 $|\uparrow_z\uparrow_z\rangle$，$|\downarrow_z\downarrow_z\rangle$，$|\uparrow_z\downarrow_z\rangle$，$|\downarrow_z\uparrow_z\rangle$ 的表象（按照顺序）下，给定如下 Hamiltonian，

$$H = \begin{pmatrix} H_0 & 0 \\ 0 & H_0 \end{pmatrix}, \tag{10.1}$$

其中

$$H_0 = \frac{1}{4+2\sqrt{2}} \begin{pmatrix} 1 & -1-\sqrt{2} \\ -1-\sqrt{2} & 3+2\sqrt{2} \end{pmatrix}. \tag{10.2}$$

计算 $t = \pi$ 时刻的演化算符，并讨论这个演化算符把独立状态 $|\uparrow_z\uparrow_z\rangle$，$|\downarrow_z\downarrow_z\rangle$，$|\uparrow_z\downarrow_z\rangle$，$|\downarrow_z\uparrow_z\rangle$ 分别映射成为什么状态。

演化算符 $U = e^{-iHt}$ 可以通过求解 H 的本征值和本征向量得到。我们先来计算 H_0 的本征向量，得到，

$$|E_1 = 0\rangle = \frac{1}{\sqrt{4+2\sqrt{2}}}\begin{bmatrix} 1+\sqrt{2} \\ 1 \end{bmatrix}, \tag{10.3a}$$

$$|E_2 = 1\rangle = \frac{1}{\sqrt{4+2\sqrt{2}}}\begin{bmatrix} 1 \\ -1-\sqrt{2} \end{bmatrix}. \tag{10.3b}$$

于是，H_0 所对应的演化算符 U_0 可以表达成为，

$$\begin{aligned} U_0 &= e^{-iE_1 t}|E_1\rangle\langle E_1| + e^{-iE_2 t}|E_2\rangle\langle E_2| \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \end{aligned} \tag{10.4}$$

于是，整体的演化算符是

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \tag{10.5}$$

这个算符作用在 $[1\ 0\ 0\ 0]^T$ 得到 $\frac{1}{\sqrt{2}}[1\ 1\ 0\ 0]^T$，也就是。

$$U|\uparrow_z\uparrow_z\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\uparrow_z\rangle + |\downarrow_z\downarrow_z\rangle\right) \tag{10.6}$$

我们看到这个演化算符把独立状态映射成为纠缠态。如果我们知道如何实现 H_0 那么，我们就能够实现 H 也就能够实现 U 了。注意到这里的 U_0 其实是Hadamard 变换。

在这个计算中，我们特意采用了交换了顺序的一组基矢。更加常用的基矢其实是 $|\uparrow_z\uparrow_z\rangle$，$|\uparrow_z\downarrow_z\rangle$，$|\downarrow_z\uparrow_z\rangle$，$|\downarrow_z\downarrow_z\rangle$。如果采用这一组基矢，我们的 Hamiltonian 需要定义为，

$$H = \frac{1}{4+2\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & -1-\sqrt{2} \\ 0 & 1 & -1-\sqrt{2} & 0 \\ 0 & -1-\sqrt{2} & 3+2\sqrt{2} & 0 \\ -1-\sqrt{2} & 0 & 0 & 3+2\sqrt{2} \end{pmatrix}. \tag{10.7}$$

读者不妨验证一下，是否如公式 (10.6) 一样，这个 Hamiltonian 对应着的演化算符确实把独立状态映射到相应的纠缠态。

这个 H 实际上反映了两个自旋的相互作用。

10.2 测量导致的纠缠

上一节我们已经看到通过演化，某些形式的 Hamiltonian 可以把独立状态转化成为纠缠态。现在，我们试试用另一个角度来构造纠缠态。这个角度就是测量，因为我们已经知道测量和演化都可以使得量子系统的状态发生变化。我们也通过一个例子来展示通过测量建立纠缠。假设我们已经得到了两个光子的纠缠态，现在我们想得到四个光子的纠缠态。当然，我们可以构造四个光子的相互作用来通过演化得到四个光子的纠缠态。这里，我们用另一种方式，来自于文献 [46]。

考虑两对已经分别实现纠缠的光子，其中 12 和 34 都处于如下纠缠态（在实际实验中，经常用另一个 Bell 态 $\left|\Psi_-^{12}\right\rangle = \frac{1}{\sqrt{2}}\left(|HV\rangle - |VH\rangle\right)$，为了描述简单，这里我们采用下面这个 Bell 态），

$$\left|\Psi^{12}\right\rangle = \frac{1}{\sqrt{2}}\left(|HH\rangle + |VV\rangle\right) \triangleq \left|\Phi^+\right\rangle = \left|\Psi^{34}\right\rangle, \tag{10.8}$$

那么整个系统的状态就是

$$\begin{aligned}\left|\Psi^{1234}\right\rangle &= \left|\Phi^+\right\rangle\left|\Phi^+\right\rangle \\ &= \frac{1}{2}\left(|HHHH\rangle + |HHVV\rangle + |VVHH\rangle + |VVVV\rangle\right).\end{aligned} \tag{10.9}$$

这里我们用了 Bell 态的记号。测量导致纠缠的基本思想就是从上面的四个状态里面，我们想办法把不需要的状态去掉，通过测量只留下需要的状态。例如，在这里，我们希望去掉这个部分，

$$|HHVV\rangle + |VVHH\rangle. \tag{10.10}$$

我们如何去掉这个部分呢，选择留下测量结果的一部分来做下一步的实验。

实验 10.1 4 光子纠缠实验：从两对已经实现纠缠的光子中选择其中一个经过反射之后到达内部方向是 0^0 的偏振分束器——它透射 0^0（标记为 H）的光，反射 90^0（标记为 V）的光。按照偏振分束器的出射光子的情况来决定是否让光子进入下一步的实验：如果有两个方向的输出的情况，继续后续的

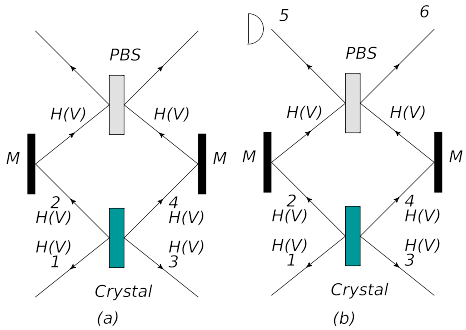


图 10.1: (a) 4 光子纠缠实验装置示意图：仅保留从偏振分束器两侧都出现光子的情况进入后续的实验，如果实验中出现分束器一侧观察到光子的情形，实验终止，重新开始。(b) 3 光子纠缠实验装置示意图：仅保留在 5 出口的探测器上接收到光子，并且 6 出口有光子输出的情况进入后续的实验。记得去获取授权

实验，否则终止。问：对于能够进入后续实验的例子来说，它们处于什么状态？在实验中，实际上，我们还需要，制备纠缠态的技术，以及探测光子从偏振分束器的哪一边出来又不破坏光子的偏振态的技术。在此，不讨论这些技术如何实现。

两对光子的状态组合有四种 (HH, HH)、(HH, VV)、(VV, HH)、(VV, VV)。下面我们会看到通过选择出射光子的情况（丢弃两个光子都从偏振分束器左侧或者右侧出射的情况），实际上进入后续实验的光子的状态肯定是 $\frac{|HHHH\rangle+|VVVV\rangle}{\sqrt{2}}$。

这里，我们让这样的两对光子的各自的其中一个光子，我们称为光子 2、4，经过一个偏振分束器（PBS），如图 10.1。我们发现，这个时候有三个可能的结果，两个光子都从左侧出来，两个光子都从右侧出来，两个光子分别从左侧和右侧出来。前两个结果对应着的光子的状态分别是 $|HV\rangle$，$|VH\rangle$。例如 $|HV\rangle$ 会让光子 2 透射，光子 4 反射，于是它们一起出现在右边。同理 $|VH\rangle$ 会让光子同时出现在左边。

对于两边都有光子的情形，$|HH\rangle$ 和 $|VV\rangle$ 都有可能，并且不可区分。于是，我们将来需要采用矢量叠加。当光子 2、4 处于 $|HH\rangle$ 态的时候，由于 1、3 分别和 2、4 的纠缠，整体必然这个时候是 $|HHHH\rangle$ 态。同理，当光子 2、4 处于 $|VV\rangle$ 态的时候，整体必然这个时候是 $|VVVV\rangle$ 态。两者不可区分

而且概率幅相同。于是，

$$\left|\Psi^{1234}\right\rangle = \frac{1}{\sqrt{2}}\left(|HHHH\rangle + |VVVV\rangle\right). \tag{10.11}$$

实际上我们为了确定偏振分束器的两侧都接收到光子，我们需要做一个不改变光子自旋状态却能够确定光子从偏振分束器的哪一侧出射的测量，然后才能把都从右侧出来或者都从左侧出来的结果排除在后续的实验中。具体技术细节就不在这里讨论了。一个简单的后续实验可以这样：例如，在四个方向上分别安装探测器的话，同时四个方向都有探测到光子的结果就可以验证我们得到了目标纠缠态。另外，实际上，实验中用的状态是另一种纠缠，$\left|\Psi_{-}^{12}\right\rangle = \frac{1}{\sqrt{2}}\left(|HV\rangle - |VH\rangle\right)$，但是原理是一样的。

实际上，这个方案是从另外一个更早一点的工作 [47] 发展过来的。在那个工作中，两对纠缠的光子被用来制备三个纠缠的光子，其中的一个光子在测量的时候用掉了。其方案非常简单，这里再作进一步的简化。我们用同样的仪器，但是在偏振分束器（PBS，内部方向 0^0，透射水平偏振光反射竖直偏振光）后面的左侧增加一个探测器。然后，我们按照这个探测器上的结果来作进一步的实验。这个探测器有三种可能的结果，得到两个光子，得到一个光子，没有得到任何一个光子。对于得到两个光子和没有得到光子的情形，我们忽略所有可能的下一步的实验得到的结果。于是，我们的下一步实验仅仅记录这个探测器得到一个光子的情形。现在，我们来思考得到一个光子的情形对应的光子的状态。相比于上面的 4 光子纠缠实验，这个 3 光子的实验更加容易实现：只要探测器能够区分收到 $0, 1, 2$ 三个光子的情形就可以，而不用一定需要探测光子经过但是不破坏其偏振状态的技术。

这个光子可能是反射过来的光子 2，于是其状态为 V，于是光子 1 的状态也是 V。同时，由于光子 4 没有投射过来，也就是被反射了，于是 4 的状态就是 V，因此 3 的状态也是 V。也就是说，134 的光子会进入下一步的实验，其状态是 VVV。这个光子可能是透射过来的光子 4，于是其状态为 H，于是光子 3 的状态也是 H。同时，由于光子 2 没有反射过来，也就是被透射了，于是 2 的状态就是 H，因此 1 的状态也是 H。也就是说，123 的光子会进入下一步的实验，其状态是 HHH。如果这两种可能的状态哪一种发生了完全不能从实验过程上区分（注意不管哪一种状态其出口都是 136），那么系统的状态就是，

$$\left|\Psi^{136}\right\rangle = \frac{1}{\sqrt{2}}\left(|HHH\rangle + |VVV\rangle\right). \tag{10.12}$$

于是，通过测量和选择，我们把两对双光子纠缠态转化成了三光子纠缠态。

在实际实验中，双光子的纠缠比较容易实现，多光子的纠缠如果通过相互作用来实现有一定的难度。因此，这种通过测量来实现多光子纠缠的方法经常被使用。

10.3 作业

习题 10.1 给定如下 Hamiltonian，

$$H_0 = \frac{1}{4+2\sqrt{2}}\begin{pmatrix} 1 & -1-\sqrt{2} \\ -1-\sqrt{2} & 3+2\sqrt{2} \end{pmatrix}. \tag{10.13}$$

请用自旋算符，例如 $I, \sigma_x, \sigma_y, \sigma_z$ 之类的形式写出这个 H_0。有了上面这个题的经验之后，来思考下面的问题。给定如下 Hamiltonian，

$$H = \frac{1}{4+2\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & -1-\sqrt{2} \\ 0 & 1 & -1-\sqrt{2} & 0 \\ 0 & -1-\sqrt{2} & 3+2\sqrt{2} & 0 \\ -1-\sqrt{2} & 0 & 0 & 3+2\sqrt{2} \end{pmatrix}. \tag{10.14}$$

请用自旋算符，例如 $\sigma_x^1\sigma_x^2, \sigma_x^1\sigma_y^2$ 之类的形式写出这个 H。

习题 10.2 CNOT 量子逻辑门的 Hamiltonian：给定如下 Hamiltonian，

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5 \\ 0 & 0 & -0.5 & 0.5 \end{pmatrix}, \tag{10.15}$$

计算 $t=\pi$ 时刻的演化算符，并讨论这个演化算符把独立状态 $|\uparrow_z\uparrow_z\rangle$，$|\uparrow_z\downarrow_z\rangle$，$|\downarrow_z\uparrow_z\rangle$，$|\downarrow_z\downarrow_z\rangle$ 分别映射成为什么状态。这个演化算符有一个自己的名字——CNOT 逻辑门，实现控制非（Controlled-NOT）计算。

习题 10.3 两个自旋的关联 1：假设氢原子电离之后，电子-质子处于如下的自旋态（仅仅考虑自旋的自由度）$|\uparrow_\alpha\downarrow_\theta\rangle$ 状态。其中 $|\uparrow_\theta\rangle$（$|\downarrow_\theta\rangle$）表示 $\hat{r}_\theta = \sin(\theta)\hat{i} + \cos(\theta)\hat{k}$ 方向的向上（向下）本征态，也就是自旋算符 $S_\theta = \vec{\hat{S}}\cdot\hat{r}_\theta$ 的向上本征态。

1. 求在该自旋态测量电子自旋 S_α^e 得到的各个可能取值的概率。测量之后，假设得到这个方向的向上状态，该系统的自旋状态是什么？对电子自旋的测量是否改变了质子的自旋状态？

2. 计算在该自旋态下，测量质子自旋 S_β^p 得到的各个可能取值的概率。

3. 计算在该自旋态下，“先”测量电子自旋 S_α^e 得到某个可能的取值，“后”在所得到的状态下测量质子自旋 S_β^p 得到的各个可能取值的概率。

4. 计算上面的测量得到的结果的关联系数 $E(\alpha,\beta)=\frac{\langle S_\alpha^e S_\beta^p\rangle-\langle S_\alpha^e\rangle\langle S_\beta^p\rangle}{\Delta S_\alpha^e \Delta S_\beta^p}$。其中 ΔS_α^e 表示测量电子的 $\hat{r}_\alpha$ 方向的自旋得到的结果的标准差。

习题 10.4 两个自旋的关联 2：假设氢原子电离之后，电子-质子处于如下的自旋态（仅仅考虑自旋的自由度）$\frac{1}{\sqrt{2}}(|\uparrow_\alpha\downarrow_\theta\rangle-|\downarrow_\alpha\uparrow_\theta\rangle)$ 状态。重复以上计算。

10.4 本章小结

本章主要展示了在耦合系统中，利用相互作用和测量——这两种改变系统状态的方法——可以得到纠缠态。纠缠态的独特的性质在上一章已经讨论过。在下一章，以及量子信息部分，我们还会再一次见到具有这样的性质的纠缠态能够用来做什么。

在耦合系统的最后我们提一下从单个粒子到耦合粒子系统理论上需要考虑的真正多出来的一个因素：全同粒子的状态的问题。我们一直没有讨论这个问题。在我们之前的讨论中，我们总是能够区分粒子 1 和粒子 2 的。如果在某些问题中，粒子是不可以区分的，量子系统的实验告诉我们，这样的粒子的状态只能够在直积空间的某些子空间内取值。例如费米子（Fermion）只能够在交换粒子编号反对称的子空间取值，而波色子（Boson）只能够在交换粒子编号对称的子空间取值。这个问题相当于对量子系统提出了一个额外的约束，或者说在粒子之间建立了额外的关联。这个问题具有非常深刻的根源和非常深刻的影响 [48]。在本书中，我们不再讨论。

下一章，以及后续的章节是纠缠态——我们也会看到其本质还是量子态叠加原理——的应用。

第十一章　量子力学的确定性经典理论的再一次讨论

在引论部分，我们已经讨论了量子系统的测量行为，以及用建立在确定性经典理论和随机性经典理论——这个随机性可以使内秉真随机性也可以是由于信息不完备导致的伪随机性，在后者的情况下随机性经典理论等价于确定性经典理论——来描述量子系统的测量行为的可能性。我们已经基本否定了这样的可能性，所以在后面的章节中转而讨论允许使用非对角元，突破经典概率论框架的数学模型来描述量子系统的测量行为。在这一章，我们再次回到这个量子系统的经典理论的可能性的问题，只是稍微改变一下我们的测量对象，使得这个测量得到的结果看起来更加具有特殊性，并且来思考什么样的经典理论可能可以描述这样的特殊行为。这个测量就是一个纠缠态的测量。

本章推荐阅读材料：J. Bell 《Speakable and Unspeakable of Quantum Mechanics》[1]。

11.1　纠缠态上的量子测量的结果

考虑一个纠缠态，

$$\left|\Psi^-\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\downarrow_z\rangle - |\downarrow_z\uparrow_z\rangle\right), \tag{11.1}$$

计算测量 $\sigma_z^1\sigma_r^2$ 的平均值。

$$\begin{aligned}\left\langle\sigma_z^1\sigma_r^2\right\rangle &= \frac{1}{2}\left(\langle\uparrow_z\downarrow_z| - \langle\downarrow_z\uparrow_z|\right)\sigma_z^1\sigma_r^2\left(|\uparrow_z\downarrow_z\rangle - |\downarrow_z\uparrow_z\rangle\right)\\ &= \frac{1}{2}\left(\langle\downarrow_z|\sigma_r^2|\downarrow_z\rangle - \langle\uparrow_z|\sigma_r^2|\uparrow_z\rangle\right)\\ &= -\cos\theta,\end{aligned} \tag{11.2}$$

其中 θ 就是 $\hat{r}$ 和 z 方向的夹角。

现在量子系统的经典理论的可能性的问题就成了是否存在一种经典概率分布，和一组对应的测量，其结果与上面的两个不同方向的自旋测量的结果一致。这里我们只关心测量的平均值。在引论部分第二章的讨论中，我们关心单次测量的结果及其平均值，还有测量以后马上再测量的单次的结果及其平均值。这一章里面，我们就不再关心单次的结果而仅仅关心平均值，也不再关心测量以后再次测量的单次结果和平均值的问题了。因此，本章在理论要符合哪些实验结果要求比引论部分的讨论要低。需要注意的是，如果我们只关心测量的平均值，对于单个的自旋，满足测量平均值与量子理论和量子实验得到的结果一样的经典概率论是能够构造出来的。在仅仅考虑平均值的情况下，Bell 就提供了一个与二维量子系统一致的隐变量理论[1]。在引言部分，我们的结论是在非常宽泛的条件下，考虑了再测量之后，与量子实验的结果相一致的经典理论基本上是不存在的。所以，我们在引言部分的讨论和这里的讨论不是重复的，尽管企图解决的问题是同一个：我们希望说明满足公式 (11.2) 测量结果关联性的经典对象是不存在的。当然，在这个时候，我们就需要对什么系统是经典对象有一个界定：凡是由经典概率论——也就是概率三元体和经典概率论公理5-8——所描述的系统，我们称为经典对象。

11.2 经典关联态上的测量的 Bell 不等式

现在，我们来展示符合经典概率论的状态是不可能遵循上面的这个测量结果之间的关联关系的。考虑这样一个系统，这个系统的状态由某一个隐藏起来的不能直接测量和确定的随机变量 λ 决定。在这个系统上，我们可以测量两个量 $A(\vec{a},\lambda)$ 和 $B\left(\vec{b},\lambda\right)$，例如 $\vec{a}=\hat{z}$，$\hat{b}=\hat{r}$，A、B 就是着两个方向上的自旋。我们想知道，如果上面的两个自旋的量子系统用经典概率分布来描述的话，这样的所测量到的关联值满足什么约束。然后，我们考察量子系统的上述结果公式 (11.2) 是否满足这个约束。

下面，我们来计算按照经典概率分布的理论这个每一次测量之后的得到的 AB 乘积的平均值。注意，在这里，我们假设 λ 取什么值和选择的测量的方向 $\vec{a}$、$\vec{b}$ 是没有关系的。这个假设实际上是要求我们的测量和被测量物理对象的状态之间存在着某种客观性：被测量物理对象的状态是独立于测量者希望来测量的物理量之外的。违反这一条的理论相当于允许被测量物理对象了解测量者的企图。这样的理论，尽管没有什么原理说不能成立，不是物理学理论追求的目标——试想一下让一个自旋了解实验者到底要测量“它自己”的什么方向然后这个自旋甚至会做相应的调整这件事情有多么的不可接受。

在这个假设下，当然这个 λ 就更加不能是同时依赖于 $\vec{a},\vec{b}$ 的函数了。不管实验者选择的测量方向是什么，被测量的物理对象的状态是由一个预先已经给定概率分布给出的某个样本决定的，而不是由某一个实验者，或者是所有的实验者所选择的方向决定的。

在这个条件下，我们来求 AB 乘积的平均值，

$$\left\langle A\left(\vec{a},\lambda\right)B\left(\vec{b},\lambda\right)\right\rangle=\int d\lambda\rho\left(\lambda\right)A\left(\vec{a},\lambda\right)B\left(\vec{b},\lambda\right). \tag{11.3}$$

这个表达式“表示”：系统的状态是由参数 λ 决定的，λ 自己符合某个概率分布，然后，给定某一个 λ 的值以后，任何测量——例如 A 测量 $\vec{a}$ 方向的自旋以及 B 测量 $\vec{b}$ 方向的自旋——结果是确定的；在这个随机客体的测量的解释下，我们需要保证测量结果之间的关联与实验观察相符。在这里，我们隐含着一个要求，**λ 和 $\vec{a},\vec{b}$ 没有关系**。λ 代表了客观系统的状态，尽管未知，尽管 λ 是隐藏变量。$\vec{a}$（$\vec{b}$）代表第一（二）个观测者 Alice（Bob）的意图。一个好的经典理论，我们希望客观系统的状态和观测者的意图是分开的，也就是说，一定程度上客观实在性是存在的。同时这个表达式还有另一重假设**$\vec{a}$ 和 $\vec{b}$ 应该是独立的，不能通过 λ 建立起来某种联系**。我们也可以形式上写下来一个 Alice 和 Bob 的整体测量结果表达式，例如

$$\left\langle AB\left(\vec{a},\vec{b};\lambda\right)\right\rangle=\int d\lambda\rho\left(\lambda\right)AB\left(\vec{a},\vec{b};\lambda\right). \tag{11.4}$$

公式 (11.3) 相当于是假设了

$$AB\left(\vec{a},\vec{b};\lambda\right)=A\left(\vec{a},\lambda\right)B\left(\vec{b},\lambda\right). \tag{11.5}$$

这个假设被称为Einstein 定域性[8]。

当然，在实验中，我们可以不仅仅要求测量结果的关联性与实验相符，还可以要求 $A(\vec{a},\lambda)$ 必须是 ± 1 等其它条件得到满足。在下面的讨论中，我们仅关心这个测量结果之间的统计关联性，因为这个是我们将要讨论的Bell 不等式所关注的量。因此，这一章的主要任务就是检验，包含隐变量的经典理论计算出来的关联公式 (11.3) 是否和量子系统测量得到的关联公式 (11.2) 存在矛盾。

为了这个目的，我们来考查这样的关联所满足的约束，

$$C\left(\vec{a},\vec{b}\right) \triangleq \left\langle A(\vec{a},\lambda)\, B\left(\vec{b},\lambda\right)\right\rangle. \tag{11.6}$$

以下计算的过程如果太抽象，想着这里我们关心的是两个不同方向的自旋的测量值的关联，也就是 $\left\langle s(\vec{a},\lambda)\, s\left(\vec{b},\lambda\right)\right\rangle$。

以下的讨论来自于 Ballentine 的《Quantum Mechanics – a modern development》[8]。首先，对于任意一个方向的单个自旋的测量，我们有

$$\left|A(\vec{a},\lambda)\right| \le 1, \left|B\left(\vec{b},\lambda\right)\right| \le 1. \tag{11.7}$$

然后，我们来证明下面的结果——称为Bell 不等式

$$\left|C\left(\vec{a},\vec{b}\right) - C\left(\vec{a},\vec{b}'\right)\right| + \left|C\left(\vec{a}',\vec{b}'\right) + C\left(\vec{a}',\vec{b}\right)\right| \le 2. \tag{11.8}$$

证明如下，

$$\begin{aligned} C\left(\vec{a},\vec{b}\right) - C\left(\vec{a},\vec{b}'\right) = & \int d\lambda \rho(\lambda)\left[A(\vec{a},\lambda)\, B\left(\vec{b},\lambda\right) - A(\vec{a},\lambda)\, B\left(\vec{b}',\lambda\right)\right] \\ = & \int d\lambda \rho(\lambda)\, A(\vec{a},\lambda)\, B\left(\vec{b},\lambda\right)\left[1 \pm A(\vec{a}',\lambda)\, B\left(\vec{b}',\lambda\right)\right] \\ - & \int d\lambda \rho(\lambda)\, A(\vec{a},\lambda)\, B\left(\vec{b}',\lambda\right)\left[1 \pm A(\vec{a}',\lambda)\, B\left(\vec{b},\lambda\right)\right], \end{aligned}$$

得到，

$$\begin{aligned} \left|C\left(\vec{a},\vec{b}\right) - C\left(\vec{a},\vec{b}'\right)\right| \le & \int d\lambda \rho(\lambda)\left[1 \pm A(\vec{a}',\lambda)\, B\left(\vec{b}',\lambda\right)\right] \\ + & \int d\lambda \rho(\lambda)\left[1 \pm A(\vec{a}',\lambda)\, B\left(\vec{b},\lambda\right)\right], \\ = & 2 \pm \left[C\left(\vec{a}',\vec{b}'\right) + C\left(\vec{a}',\vec{b}\right)\right]. \end{aligned} \tag{11.9}$$

现在，我们取 $\vec{a}\,\vec{b},\vec{a}',\vec{b}'$ 都在一个平面内，并且，$\vec{a}$ 在零度，$\vec{b} = \vec{a}'$ 在 θ 的角度，$\vec{b}'$ 在 2θ 的角度，我们得到，

$$\left|C(\theta) - C(2\theta)\right| + \left|C(\theta) + C(0)\right| \le 2. \tag{11.10}$$

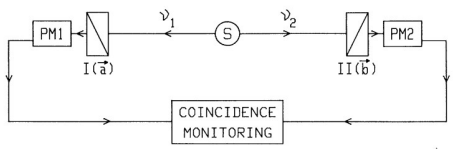


图 11.1: 一对纠缠光子经过偏振片 I 和 II 以后，再做测量。图取自 [51]。记得去获取授权

对于量子系统，我们有 $C(\theta) = -\cos(\theta)$，于是，我们考察以下不等式是否永远满足，

$$2\cos(\theta) - \cos(2\theta) \leq 1. \tag{11.11}$$

我们发现，当例如 $\theta = \frac{\pi}{3}$ 的时候，左边等于 $\frac{3}{2} > 1$。于是，量子系统不满足Bell 不等式，而Bell 不等式则是经典概率论的直接推论。

于是，如果实验证明，量子系统的行为符合公式 (11.2)，但是不符合Bell 不等式，则，证明符合经典概率论的量子系统的理论是不可能的。下一节，我们就来看量子系统的实验结果。

11.3 Bell 定理的实验检验以及 Die Hard 教授的吹毛求疵

始于 1935 年的爱因斯坦和玻尔关于量子力学对物理实在的描述是否完备，是否存在定域隐变量等问题的论争，长期停留在思辨性的层面 [1]。1964 年，Bell为这场论证导出一个简洁的、实验可行的不等式作为谁是谁非的判据，论证于是转化为由实验结果裁定的时代，物理学家开始努力完成验证Bell 不等式的各种类型的实验。验证Bell 不等式的实验可以大致分为 4 代。

第一代是在 1970 年代完成的几个实验 [49–51]，典型的实验装置如下图 11.1。这基本上就是的纠缠光子对实现 Einstein-Podolsky-Rosen-Bohm 思想实验的原理图。光源 S 利用原子级联跃迁产生的纠缠光子对，发出一对处于纠缠态的光子 ν_1 和 ν_2，分别通过两个线偏振片 I 和 II，光子经过 I 之后的极化方向为 $\vec{a}$，光子经过 II 之后的极化方向为 $\vec{b}$ 。它们分别通过两个光电倍增管 PM1 和 PM2，把微弱的光子信号放大，最后到达符合计数器 CM。符合计数法的作用是对接收到的一对光子的极化关联进行测量。

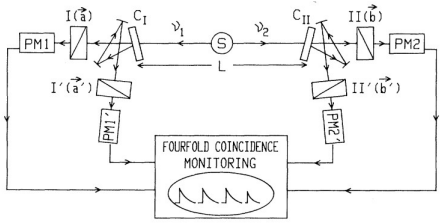


图 11.2: 一对纠缠光子经过偏振片 I（或者 I′）和 II（或者 II′）以后，再做测量。其中到底经过 I，II 还是 I′，II′ 由开关在光子飞行途中决定。图取自 [51]。记得去获取授权

这几个实验的结果都给出来量子系统的行为违反Bell 不等式或者CHSH 不等式（后者是Bell 不等式的另一种形式，更便于实验的验证，见 [52]），支持量子力学的预言。但是这类实验有一些漏洞，主要的漏洞是：实验采用两个固定于实验室内的单路偏振片，它们预先设定了偏振方向，这样尚不足以完全否定量子力学是定域实在的理论。粗略地来说，由于偏振片 I 和 II 的方向事先就给定，存在者一种可能的经典理论，能够得到满足所测量到的关联性的结果。关于更多的这个漏洞的讨论可见“Wikipedia 上Bell 不等式的检验”的条目[1]以及“Bell 不等式的检验的漏洞”的条目[2]。

第二代实验开始于 1982 年 Aspect 等人的实验。最重要的改进是把第一代的单路偏振片改为双路偏振片（见下图 11.2），在 ν_1 光子这一方，两个偏振片分别为 I 和 I′，偏振方向分别为 $\vec{a}$ 和 $\vec{a}'$；在 ν_2 光子这一方，两个偏振片分别为 II 和 II′，偏振方向分别为 $\vec{b}$ 和 $\vec{b}'$。开关 C_1 和 C_2 分别连接在每一方的两个偏振片组之间。在光子由光源飞向偏振片的途中，开关随机地选择偏振的取向（$\vec{a}$ 或者 $\vec{a}'$，$\vec{b}$ 或者 $\vec{b}'$），这样做是为了避免在光源和偏振片之间，或测量光子偏振的两个事件之间，可以用任何等于或小于光信号的速度建立联系，这样做更为接近爱因斯坦等人的理想的实验方案。这个实验的结果明显违反Bell 不等式，支持量子力学。

第三代实验始于 1980 年代后期，仍然是用纠缠光子完成的，让紫外光子通过非线性晶体下转换为一对纠缠光子，这样光源具有更高的效率。两路偏振探测器的距离拉大。1998 年，Gisin等人的实验也是采用下转换光子做到了 10 公里的距离 [53]。1998 年，Zeilinger等人的实验两个偏振片组之间

[1]https://en.wikipedia.org/wiki/Bell_test_experiments

[2]https://en.wikipedia.org/wiki/Loopholes_in_Bell_test_experiments

的距离为 400 米（1982 年，Aspect等人的实验为 12 米），检验的对象从Bell 不等式换成了CHSH 不等式[54]。这些实验尽量避免了所有可能的漏洞，向爱因斯坦等人的理想中的实验更前进一步。

以上三代实验均是用纠缠光子对作为对象，验证量子力学幽灵般的超距作用是否存在。这些实验虽然不断改进，但是存在两个共同的漏洞—探测漏洞。探测漏洞指实验不能探测到所有用于实验的光子，甚至漏掉 80%。这样的取样是不“公正的”，实验者只能假定所探测到的光子的性质代表了全部光子群体。

第四代实验，是荷兰Delft 科学技术大学的 Hanson等人利用纠缠电子完成的。它避免了使用光子的探测漏洞。为了避免另一种漏洞—通信漏洞（即电子相距很远时，纠缠态会遭到破坏），Hanson采用了纠缠交换技术（entanglement swapping ）将光子和电子的优点结合在一起。这被认为是没有漏洞的Bell 不等式的验证实验，完全实现了爱因斯坦等人的假想实验。但是，实验结果不支持爱因斯坦的定域实在论中的观点。他们利用Delft 科学技术大学的两个实验室，把两个电子分别存放在这两个实验室内的金刚石 A 和 B 中，A 和 B 相距 1280 米。这两个电子处于非纠缠态，但是它们在分别与两个光子纠缠，而这两个光子都被发送到第三个实验室 C，在第 C 让这两个光子纠缠，这就导致了与光子纠缠的两个电子也处于纠缠态。实验装置如图 11.3 所示。

在 9 天内，该实验组共产生了 245 对纠缠电子，最终测量结果表明，两个电子之间的相干性超过了Bell 不等式的上限支持量子力学。由于电子易于检测，所以避免了探测漏洞，又由于两个电子的距离足够远，所以避免了通讯漏洞。到这个实验为止，应该说，就算是 DieHard 教授，也应该可以接受了：量子系统的行为不满足Bell 不等式。然而，所有的定域实在经典理论都满足Bell 不等式。因此，量子系统的理论不可能是定域实在的经典理论。也就是说，量子系统的理论有必要采用目前的基于包含了非对角元的密度矩阵的量子力学的理论，或者说基于可线性叠加的矢量和非对易算符的量子力学的理论。

实际上，作者对于量子力学的基本问题的兴趣，也是从这个利用最一般的经典理论来描述量子系统的行为的可能性开始的。可以看到，本书在结构上和内容上，一再回到这个主题：是否能够用经典概率论来描述量子系统的行为。实际上，作者在开始讨论这个问题的时候，已经有大量的Bell 不等式检验的实验，当然更加已经有了定域实在经典系统的 Bell 定理。但是，那

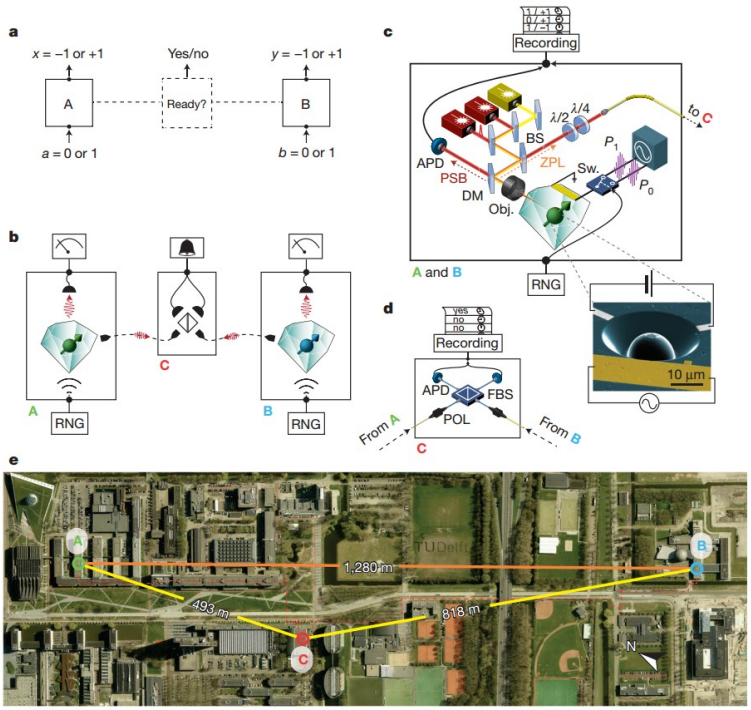


图 11.3: 利用电子代替了光子避免了其他Bell 不等式检验的漏洞。图取自[55]。记得去获取授权

个时候，作者提出来的思路是，如果我们不再进行进一步的实验，是否现在已经知道的量子系统的行为，就能够告诉我们：不可能构造一个能够解释一个自旋的单次测量的结果及其平均值，以及测量以后马上再次测量得到的单次结果及其平均值的实验结果的经典理论。我们的研究发现，这样的经典理论会非常的难以理解，需要破坏大量的一个理论应该有的自然的合理的性质。甚至，我们必须放弃以下的测量的“直观理解”：测量得到一个状态表示这个系统当前处于这个状态，而代之以：测量的到一个状态表示如果接着马上再一次重复同样的测量我们将获得一样的结果。这是弱很多的一个条件。在这个条件下，我们将不再能够问这个时候的状态是什么的问题，仅仅能够回答这个时候的这个状态下做那样的测量得到的结果是什么。也就是说，量子状态和量子测量不再是能够分开的问题。因此，实际上，在这个工作中，我们得到：不满足矢量叠加原理的以经典概率论为最一般的形式的经典理论不能描述量子系统的行为 [56]。

11.4 前后两次测量的关联的量子力学计算

在 2.3 节中，我们企图构建一个能够复现量子系统前后两次测量的关联的经典理论，现在，我们来写下来相应的量子理论，然后再一次来讨论这个经典理论的可能性。

首先，我们写下来量子力学对公式 (2.34) 的计算。

$$
\begin{aligned}
\langle s_{\hat{r}_1} s_{\hat{r}_2} \rangle = & \sum_{s_{\hat{r}_1}=\pm 1, s_{\hat{r}_2}=\pm 1} s_{\hat{r}_1} s_{\hat{r}_2} \langle s_{\hat{r}_1} | \rho_0 | s_{\hat{r}_1} \rangle \langle s_{\hat{r}_2} | \left(| s_{\hat{r}_1} \rangle \langle s_{\hat{r}_1} | \right) | s_{\hat{r}_2} \rangle \\
= & \sum_{s_{\hat{r}_1}} \langle s_{\hat{r}_1} | \rho_0 | s_{\hat{r}_1} \rangle \left[\sum_{s_{\hat{r}_2}} s_{\hat{r}_1} s_{\hat{r}_2} \langle s_{\hat{r}_2} | \, s_{\hat{r}_1} \rangle \langle s_{\hat{r}_1} \, | s_{\hat{r}_2} \rangle \right] \\
= & \sum_{s_{\hat{r}_1}} \langle s_{\hat{r}_1} | \rho_0 | s_{\hat{r}_1} \rangle \hat{r}_1 \cdot \hat{r}_2 \\
= & \hat{r}_1 \cdot \hat{r}_2. \qquad (11.12)
\end{aligned}
$$

其中 ρ_0 是任意一个初始状态，$s_{\hat{r}_1} \langle s_{\hat{r}_1} | \rho_0 | s_{\hat{r}_1} \rangle$ 表示测量到第一个自旋的取值为 $s_{\hat{r}_1} \in \pm 1$ 并且其的几率为 $\langle s_{\hat{r}_1} | \rho_0 | s_{\hat{r}_1} \rangle$，$s_{\hat{r}_2} \langle s_{\hat{r}_2} | \left(| s_{\hat{r}_1} \rangle \langle s_{\hat{r}_1} | \right) | s_{\hat{r}_2} \rangle$ 表示对于以上测量后的状态 $\left(| s_{\hat{r}_1} \rangle \langle s_{\hat{r}_1} | \right)$ 来说，测量到 $s_{\hat{r}_2} \in \pm 1$ 及其几率为 $\langle s_{\hat{r}_2} | \left(| s_{\hat{r}_1} \rangle \langle s_{\hat{r}_1} | \right) | s_{\hat{r}_2} \rangle$。因此，我们看到很难构造出来经典概率分布函数来符合的实验结果公式 (2.34)，在量子力学的数学框架里面，用了量子力学的公

理3和公理4，是一件非常简单的事情：测量 $\hat{r}_1$ 方向的自旋得到取值为 $s_{\hat{r}_1}$ 以后，系统的密度矩阵是，

$$\hat{\rho}\left(s_{\hat{r}_1}\right)=\left|s_{\hat{r}_1}\right\rangle\left\langle s_{\hat{r}_1}\right|, \tag{11.13}$$

这个时候如果测量 $\hat{r}_2$ 方向的自旋，得到取值为 $s_{\hat{r}_2}$ 的几率是

$$p\left(s_{\hat{r}_2}|s_{\hat{r}_1}\right)=\left\langle s_{\hat{r}_2}\right|\left(\left\langle s_{\hat{r}_1}\right|\left|s_{\hat{r}_1}\right\rangle\right)\left|s_{\hat{r}_2}\right\rangle. \tag{11.14}$$

11.5 互斥构造和独立构造的经典理论的非定域性

现在让我们用第二章所构建的基于互斥随机变量或者独立随机变量的量子系统的经典理论来描述处于纠缠态的两个自旋的测量。我们先给出量子力学的描述。状态为

$$\left|\Psi^{-}\right\rangle=\frac{1}{\sqrt{2}}\left(|\uparrow_z\downarrow_z\rangle-|\downarrow_z\uparrow_z\rangle\right). \tag{11.15}$$

测量结果的计算可以直接计算两个所测量的算符的乘积的平均值，也就是

$$\begin{aligned}
\left\langle\sigma_z^1\sigma_r^2\right\rangle &= \frac{1}{2}\left(\langle\uparrow_z\downarrow_z|-\langle\downarrow_z\uparrow_z|\right)\sigma_z^1\sigma_r^2\left(|\uparrow_z\downarrow_z\rangle-|\downarrow_z\uparrow_z\rangle\right) && \text{(11.16a)}\\
&= \frac{1}{2}\left(\langle\uparrow_z\downarrow_z|\sigma_z^1\sigma_r^2|\uparrow_z\downarrow_z\rangle+\langle\downarrow_z\uparrow_z|\sigma_z^1\sigma_r^2|\downarrow_z\uparrow_z\rangle\right) && \text{(11.16b)}\\
&\quad -\frac{1}{2}\left(\langle\uparrow_z\downarrow_z|\sigma_z^1\sigma_r^2|\downarrow_z\uparrow_z\rangle+\langle\downarrow_z\uparrow_z|\sigma_z^1\sigma_r^2|\uparrow_z\downarrow_z\rangle\right) && \text{(11.16c)}\\
&= \frac{1}{2}\left(\langle\downarrow_z|\sigma_r^2|\downarrow_z\rangle-\langle\uparrow_z|\sigma_r^2|\uparrow_z\rangle\right) && \text{(11.16d)}\\
&= -\hat{r}\cdot\hat{z}. && \text{(11.16e)}
\end{aligned}$$

注意公式 (11.16b) 和公式 (11.16c) 中的每一项，我们都可以展开成乘积项，也就是形如，

$$\langle\uparrow_z\downarrow_z|\sigma_z^1\sigma_r^2|\uparrow_z\downarrow_z\rangle=\left(\langle\uparrow_z|\sigma_z^1|\uparrow_z\rangle\right)\left(\langle\downarrow_z|\sigma_r^2|\downarrow_z\rangle\right). \tag{11.17}$$

这个就是量子版本的 $AB\left(\vec{a},\vec{b};\lambda\right)=A\left(\vec{a},\lambda\right)B\left(\vec{b};\lambda\right)$，也就是公式 (11.5) 的Einstein 定域性。因此，量子力学的状态和测量的公理3和公理4，不仅能够描述量子纠缠系统的测量，还能够在数学形式上满足Einstein 定域性。用量子力学来解释纠缠系统测量的结果还可以不采用算符乘积形式而是看做逻辑上“先后”测量。我们已经在第九章的 9.3 节做过这一个计算。

现在，我们来看一看第二章所构建的基于互斥随机变量或者独立随机变量的量子系统的经典理论能否给出同样和实验相符的结果并且数学形式上满足公式 (11.5) 的Einstein 定域性。我们先尝试着从公式 (2.71)、公式 (2.72) 和公式 (2.73) 写下来两个自旋的经典理论，

$$\rho^c = \frac{1}{\mathcal{N}} \sum_{\hat{r}_1, \hat{r}_2} \rho^c \left(s_{\hat{r}_1}, s_{\hat{r}_2} \right), \tag{11.18}$$

和

$$\rho^c = \prod_{\hat{r}_1, \hat{r}_2} \rho^c \left(s_{\hat{r}_1}, s_{\hat{r}_2} \right), \tag{11.19}$$

其中

$$\begin{aligned} \rho^c \left(s_{\hat{r}_1}, s_{\hat{r}_2} \right) = & \, p_{\uparrow\uparrow} \left(\hat{r}_1, \hat{r}_2 \right) \left| \uparrow_{\hat{r}_1} \uparrow_{\hat{r}_2} \right\rangle \left\langle \uparrow_{\hat{r}_1} \uparrow_{\hat{r}_2} \right| \\ & + p_{\uparrow\downarrow} \left(\hat{r}_1, \hat{r}_2 \right) \left| \uparrow_{\hat{r}_1} \downarrow_{\hat{r}_2} \right\rangle \left\langle \uparrow_{\hat{r}_1} \downarrow_{\hat{r}_2} \right| \\ & + p_{\downarrow\uparrow} \left(\hat{r}_1, \hat{r}_2 \right) \left| \downarrow_{\hat{r}_1} \uparrow_{\hat{r}_2} \right\rangle \left\langle \downarrow_{\hat{r}_1} \uparrow_{\hat{r}_2} \right| \\ & + p_{\downarrow\downarrow} \left(\hat{r}_1, \hat{r}_2 \right) \left| \downarrow_{\hat{r}_1} \downarrow_{\hat{r}_2} \right\rangle \left\langle \downarrow_{\hat{r}_1} \downarrow_{\hat{r}_2} \right|. \end{aligned} \tag{11.20}$$

这里

$$\begin{aligned} p_{\uparrow\uparrow} \left(\hat{r}_1, \hat{r}_2 \right) &= \frac{1 - \hat{r}_1 \cdot \hat{r}_2}{4}, \\ p_{\uparrow\downarrow} \left(\hat{r}_1, \hat{r}_2 \right) &= \frac{1 + \hat{r}_1 \cdot \hat{r}_2}{4}, \\ p_{\downarrow\uparrow} \left(\hat{r}_1, \hat{r}_2 \right) &= \frac{1 + \hat{r}_1 \cdot \hat{r}_2}{4}, \\ p_{\downarrow\downarrow} \left(\hat{r}_1, \hat{r}_2 \right) &= \frac{1 - \hat{r}_1 \cdot \hat{r}_2}{4}. \end{aligned} \tag{11.21}$$

可以验证如果实际上 Alice 和 Bob 选择了观测 $\hat{r}_1, \hat{r}_2$，则得到 $s_{\hat{r}_1} = \pm 1, s_{\hat{r}_2} = \pm 1$ 的四个结果的概率和量子力学预测的相符，和量子系统的实验相符。这四个观测导致的合起来的关联函数也正好是 $-\hat{r}_1 \cdot \hat{r}_2$。看起来我们的经典理论和实验结果相符啊！

当然，你可以说我作弊，实际上我在构造经典理论的时候的所有的这些概率都是先通过量子力学计算出来然后写成经典理论的形式的。但是，至少我有了一个能够给出和实验结果相符的经典理论啊。当然，它有一切单个自旋的经典理论的毛病：破坏自旋算符之间的内部关系（各个方向的自旋成为独立或互斥的随机变量）、实验者不再能够决定希望观测的自旋方向大多数时候观测不到想要的方向、测量独立随机变量当中的一个会改变其他“独立”随机变量的分布函数。但是，万一，Die Hard 宁愿接受这些也不愿意接

受量子力学呢？量子力学的密度矩阵的非对角元确实超越经典概率论，而且也正是这些非对角元才导致了对量子系统的测量看起来会改变其状态，不再是一个克隆，或者说一个传播。我们来揭示这样的理论的另外一个有可能 Die Hard 也不想接受的性质：非定域性。

如果Einstein 定域性成立，我们应该有公式 (11.5)，在这里测量的是 $s_{\hat{r}_1}, s_{\hat{r}_2}$，于是

$$s_1 s_2 \left(\hat{r}_1, \hat{r}_2; \lambda\right) = s_1 \left(\hat{r}_1, \lambda\right) s_2 \left(\hat{r}_2, \lambda\right). \tag{11.22}$$

而且，λ 和 $\hat{r}_1, \hat{r}_2$ 没有关系，是独立于被测量方向的描述系统状态的客观实在的隐变量。从公式 (11.17) 我们也已经看到，量子力学的数学形式是满足这一条的。

然后，在我们现在构造的描述量子纠缠态的经典理论中，我们注意到，实际上测量单个自旋对应着的随机变量根本没有意义，我们的基本事件是有一**对**随机变量 $\hat{r}_1, \hat{r}_2$ 所标记的事件 $\left|\uparrow_{\hat{r}_1}\uparrow_{\hat{r}_2}\right\rangle\left\langle\uparrow_{\hat{r}_1}\uparrow_{\hat{r}_2}\right|$ 等。因此，在求平均的过程中真正发挥作用的是，

$$s_1 s_2 \left(\hat{r}_1, \hat{r}_2; \lambda_{\hat{r}_1\hat{r}_2}\right) = s_1 \left(\hat{r}_1, \lambda_{\hat{r}_1\hat{r}_2}\right) s_2 \left(\hat{r}_2, \lambda_{\hat{r}_1\hat{r}_2}\right). \tag{11.23}$$

也就是不仅被测量系统和观测者之间的独立的客观实在不存在，两个自旋的独立的客观实在也不存在。现在，我们终于清楚了，为什么我们能够写下来量子系统的经典理论了，除了前面已在提到的这个理论的问题，还有一个很大的问题：不满足Einstein 定域性。当然，你可以继续追问，为什么我们希望我们的理论满足Einstein 定域性呢？这个和物理学关于客观实在的梦想有关：被观测系统总在那里（有某个状态）不管观测者是否观测这个系统。破坏这个梦想，就相当于让被测量系统具有某种渠道可以了解观测者的意图：看了它才现实相应的状态，才在那里。这总是一件很奇怪的事情。能够不牵扯这一步还是不要牵扯了。而且，量子力学在数学形式满足这个要求，不需要破坏非定域性。

在本章和第二章中，我们对于量子系统的经典理论的可能性做了非常深入和有一定难度的讨论。我们希望读者能够跟着这个思路想一想，甚至超越这个思路再想一想这个问题。其实，这是量子力学的魅力的一部分，非常重要的一部分。

11.6 作业

习题 11.1 补充公式 (11.12) 的推导。具体计算的时候为了方便（但是要交代为什么可以）可以取 $\hat{r}_1 = \hat{z}$。你可能会用到三角函数的倍角公式或者积化和差、和差化积公式。

习题 11.2 从量子力学推导公式 (11.20) 中的形如 $p_{\uparrow\uparrow}(\theta_1, \phi_1; \theta_2, \phi_2)$ 的四个概率的表达式。具体计算的时候为了方便（但是要交代为什么可以）可以取 $\hat{r}_1 = \hat{z}$。你可能会用到三角函数的倍角公式或者积化和差、和差化积公式。

习题 11.3 检索和阅读 “GHZ 状态和 GHZ 实验”，做一个综述，包含整理、总结、评价。

习题 11.4 阅读 David Mermin 的《Is the moon there when nobody looks? Reality and the quantum theory》，做一个阅读报告，包含整理、总结、评价。

11.7 本章小结

这一章，我们用纠缠自旋态的两个不同方向上的自旋的测量关联值来讨论，量子系统的行为——它满足量子的理论预测 $-\cos\theta$ 的关联——是否能够用满足经典概率分布的理论模型来描述。我们发现，至少数学结果上，这个经典关联必须满足一个不等式的约束：Bell 不等式，而量子关联数学上不满足这个约束。实验上，系统是否遵循量子关联，还是经典关联就是另外一个问题了，尽管看起来量子系统的行为不满足Bell 不等式，除了一些吹毛求疵的漏洞问题。尤其是 2015 年的Hanson等人的实验 [55] 被认为是没有漏洞的实验。

不过从 Die Hard 教授的角度来说，漏洞就是漏洞。是否真的就不能找出来Hanson等人的实验的漏洞也是一个问题。因此，仅仅从Bell 不等式的实验检验的角度来考察量子系统是否就真的不能用满足经典概率的理论（而且这个经典概率可以看作是信息不完全的时候的确定性理论的表现）来描述，还是一个还可以继续研究的问题。

然而，在引论部分，我们考虑了满足量子系统单次测量的结果及其系综测量的平均值以及测量后再次测量的单次结果及其系综测量的平均值的经典理论的可能性。从这个角度来说，第二章构建的能够给出和量子系统的实

验相符的结果的经典理论和本章最后两节对这个经典理论的讨论，一定程度上回答了这个问题：如果非要这样的理论不可，代价相当大——非定域、破坏自旋算符之间的内部关系（各个方向的自旋成为独立或互斥的随机变量）、实验者不再能够决定希望观测的自旋方向大多数时候观测不到想要的方向、测量独立随机变量当中的一个会改变其他“独立”随机变量的分布函数。相比于这个经典理论。量子系统的量子理论具有以下特征：定域性，自旋算符之间的内部关系得到了保持，实验者可以选择观察方向，独立变量的测量真的独立。

在思路上，Bell 不等式的思路是证明符合经典概率论的理论的一般性质，然后通过实验来检验这样的约束是否得到满足；引论中的讨论则是，企图显式地构造出来满足量子单次和再次测量实验的单次结果和平均值的实验结果的经典概率论形式的理论，然后来考察这样的理论是否存在是否比现有的量子理论看起来更加合理。从证明不可能的严谨性来说，Bell 不等式的思路更加合理，因为显式构造的思路上可能没有办法穷尽所有可能的理论的形式。不过，我认为，后者这样的构造性思路更加能够促进对一个理论的理解。

第十二章　测量的量子力学？

在我们已经学习过的目前的量子理论中，测量是通过形式化的公理——关于测量得到的结果的公理3、关于测量后状态的公理4——来表达的，而不涉及实际的测量如何实现。实际上，在量子力学的实验中，测量是通过具体的装置来实现的。也就是说，原则上，测量的过程，是通过我们所关心的量子系统和装置的相互作用来实现的。于是，自然地，我们可以问这样的问题：是否，测量过程，也应该是量子力学相互作用导致的演化过程的一部分？到底是否存在一个关于测量的动力学的描述？在这一章里面，我们来讨论一下这个问题。这个问题，目前为止，还没有很好的答案。

12.1　经典硬币的测量：克隆

我们先来看经典状态的测量。我们还是关心最简单的经典系统：二态系统的测量。一个二态经典系统 c，例如硬币，的一般状态由如下密度矩阵描述，

$$\rho^c = p\,|\uparrow\rangle\langle\uparrow| + (1-p)\,|\downarrow\rangle\langle\downarrow|, \tag{12.1}$$

其中 $p \in [0,1]$ 是唯一的变量。我们如果测得 p 的值，也就知道了这个系统的状态。

为了完成这件事情，我们可以考虑引入一个三个状态——状态 0，1，−1——的系统（例如人的眼睛）m。其初始处于某个状态 0。然后我们设计一种机制，让这个 m 的末状态和 ρ^c 的状态之间建立起来关联：例如当 c 处于 $\uparrow$ ($\downarrow$) 态的时候，m 会相应地处于 1 (−1) 态。这样的一种机制，其实本质上必须是某个动力学过程的结果，例如，人的眼睛看到硬币的状态这样的一个过程。实际写出来这个动力学过程通常是比较困难的[1]。如果我们假设这样

[1]例如，我们需要关心如下过程牵涉到的所有的基础的物理的、化学的，甚至神经动力学层次的动力学

的过程存在，那么我们的讨论会边的简单很多。也就是，某个过程 U 满足，

$$\rho_0^{cm} = \rho^c \otimes |0\rangle\langle 0| \longrightarrow p\,|\uparrow,1\rangle\langle\uparrow,1| + (1-p)\,|\downarrow,-1\rangle\langle\downarrow,-1|. \tag{12.2}$$

我们统一了经典和量子演化的形式 $\rho(t) = U(t)\rho(0)U^\dagger(t)$。于是，我们相当于寻找这样的 U，使得

$$U(t)\rho^c \otimes |0\rangle\langle 0| U^\dagger(t) = p\,|\uparrow,1\rangle\langle\uparrow,1| + (1-p)\,|\downarrow,-1\rangle\langle\downarrow,-1|. \tag{12.3}$$

可以验证，形式上，如下算符就可以满足要求，

$$\tilde{U}(t) = \frac{1+\sigma_z^c}{2}|1^m\rangle\langle 0^m| + \frac{1-\sigma_z^c}{2}|-1^m\rangle\langle 0^m|. \tag{12.4}$$

可惜这个算符不满足幺正性，于是，我们构造，

$$\begin{aligned} U(t) = -i\big(&|\uparrow,1\rangle\langle\uparrow,0| + |\uparrow,0\rangle\langle\uparrow,1| + |\downarrow,-1\rangle\langle\downarrow,0| + |\downarrow,0\rangle\langle\downarrow,-1| \\ &+ |\uparrow,-1\rangle\langle\uparrow,-1| + |\downarrow,1\rangle\langle\downarrow,1|\big). \end{aligned} \tag{12.5}$$

可以验证 $U^\dagger U = I = UU^\dagger$。那下一步的问题是，什么样的相互作用 H 会导致这样的 $U(t)$。具体计算过程不难但是比较复杂：求出来 U 的本征向量，然后凑出合适的 H。这里我们直接给出答案，

$$H = |\uparrow,1\rangle\langle\uparrow,0| + |\uparrow,0\rangle\langle\uparrow,1| + |\downarrow,-1\rangle\langle\downarrow,0| + |\downarrow,0\rangle\langle\downarrow,-1|, \tag{12.6}$$

并且

$$t = \frac{\pi}{2}. \tag{12.7}$$

反过来，很容易验证 $U(t) = e^{-iHt}$ 得到上面的演化算符。至于我们的眼睛如何实现这样的 H，甚至是否我们的眼睛真的是这样实现的，就是另外一个问题了。这里仅仅通过这个例子说明，经典客体的状态的测量可以看作这样一个过程，

$$\rho_0^{cm} \longrightarrow \rho_f^{cm}, \tag{12.8}$$

其中，末状态包含经典关联。在我们的例子中，建立起来的经典关联态是，

$$\rho_f^{cm} = p\,|\uparrow,1\rangle\langle\uparrow,1| + (1-p)\,|\downarrow,-1\rangle\langle\downarrow,-1|. \tag{12.9}$$

方程：先把 c 系统通过眼珠这个透镜在视网膜成像，再把像转化成电信号由神经元传给大脑来处理成对大脑有意义的信号。当然，原则上，写下这些方程来也不是不可能的。

这个状态有一个很好的性质，

$$tr^m\left(\rho_f^{cm}\right) = p\left|\uparrow\right\rangle\left\langle\uparrow\right| + (1-p)\left|\downarrow\right\rangle\left\langle\downarrow\right| = \rho_0^c, \tag{12.10a}$$

$$tr^c\left(\rho_f^{cm}\right) = p\left|1\right\rangle\left\langle 1\right| + (1-p)\left|-1\right\rangle\left\langle -1\right| \sim \rho_0^c. \tag{12.10b}$$

其中，我们运用了部分求迹的操作。我们知道这个部分求迹的操作的含义，在概率论中，是忽略某一个变量的取值，而只看留下来那个变量的情况。于是，从上面的部分求迹的结果我们发现，ρ_f^{cm} 这个状态有如下的性质：当我们只关心 c 系统而不关心 m 系统的时候，看起来就是 c 系统的初始状态；当我们只关心 m 系统而不关心 c 系统的时候，看起来还是像是 c 系统的初始状态，只不过被“复制”到了 m 系统上。后面的状态，只要我们把 c 系统和 m 系统认同，那么，就直接就是 ρ_0^c。这也就是，

$$tr^m\left(\rho_f^{cm}\right) = \rho_0^c, \tag{12.11a}$$

$$tr^c\left(\rho_f^{cm}\right) = \rho_0^c. \tag{12.11b}$$

我们发现经典测量的结果满足这样的性质。从这个意义上说，经典的测量总是制备了一个所测量的系统的克隆。

12.2 波函数塌缩：非对角项的消失

现在，我们再来看量子系统的测量。我们也考虑最简单的，经典 $\frac{1}{2}$ 自旋的测量。

由于“克隆”这个词在量子信息的语言中已经被运用，我们先来说一下，量子信息语言中的“克隆”的含义。对于一个 q 系统的纯态 $|\psi^q\rangle$，我们希望能够得到的克隆定义为这样的一个状态，

$$\left|\psi_0^q\right\rangle \otimes \left|\phi_0^m\right\rangle \longrightarrow \left|\psi_0^q\right\rangle \otimes \left|\phi_f^m\right\rangle : \left|\phi_f^m\right\rangle \sim \left|\psi_0^q\right\rangle. \tag{12.12}$$

其中，$\left|\phi_f^m\right\rangle \sim \left|\psi_0^q\right\rangle$ 的含义是，如果我们把 m 和 q 系统的相应本征态重新编号以后一一对应起来，则 $\left|\phi_f^m\right\rangle = \left|\psi_0^q\right\rangle$，两个状态同构。我们来看看通过量子系统的演化，是否能够实现这样的“克隆”。

假设我们可以克隆任意的 $\left|\psi_0^q\right\rangle$。我们选择两个状态 $\left|\psi_1^q\right\rangle$ 和 $\left|\psi_2^q\right\rangle$，分别得到满足要求的 $\left|\psi_1^q\right\rangle \otimes \left|\phi_1^m\right\rangle$ 和 $\left|\psi_2^q\right\rangle \otimes \left|\phi_2^m\right\rangle$。让我们来计算这两个状态的内积。

由于量子演化过程内积不变，我们有

$$\begin{aligned}\left(\left\langle\psi_1^q\right|\otimes\left\langle\phi_0^m\right|\right)\left(\left|\psi_2^q\right\rangle\otimes\left|\phi_0^m\right\rangle\right)&=\left(\left\langle\psi_1^q\right|\otimes\left\langle\phi_1^m\right|\right)\left(\left|\psi_2^q\right\rangle\otimes\left|\phi_2^m\right\rangle\right)\\ \Rightarrow\left\langle\psi_1^q\middle|\psi_2^q\right\rangle&=\left\langle\psi_1^q\middle|\psi_2^q\right\rangle\left\langle\phi_1^m\middle|\phi_2^m\right\rangle\\ \Rightarrow\left\langle\psi_1^q\middle|\psi_2^q\right\rangle\left(1-\left\langle\phi_1^m\middle|\phi_2^m\right\rangle\right)&=0\end{aligned}\tag{12.13}$$

于是，我们得到要么 $\left\langle\psi_1^q\middle|\psi_2^q\right\rangle=0$ 或者 $\left\langle\phi_1^m\middle|\phi_2^m\right\rangle=1$。也就是说，对于任意状态我们做不到克隆，除非被克隆的状态要么是一样的，要么是正交的。而对于经典客体的状态，我们知道这一点是满足的：任意两个状态要么相同要么正交，也就是

$$\left\langle\psi_i^c\middle|\psi_j^c\right\rangle=\delta_{ij}.\tag{12.14}$$

但是量子态不满足这个要求。因此，这个量子不可克隆，经典可以克隆的定理，很好地反应了经典和量子的区别。正是因为这个原因，量子不可克隆定理在量子信息的研究中有一定的地位。未知的经典信号可以通过电路来复制，未知的量子状态是不能复制的。

但是，到此为止，不是克隆的全部。量子不可克隆定理仅仅告诉我们未知的纯态，通过幺正演化的方式，是不能克隆的。那么，未知的混合态呢？如果我们允许一般的过程而不仅仅是幺正演化呢？这个时候，我们必须搞清楚两件事情：经典的混合态是否可以克隆[2]？所谓一般的过程指的是什么？

通过上一节经典测量的讨论，我们发现，经典的混合态可以克隆，其含义是，对于任意 ρ_0^c，存在经典幺正演化过程，满足

$$\rho_0^c\otimes\rho_0^m\longrightarrow\rho_f^{cm}:tr^m\left(\rho_f^{cm}\right)=\rho_0^c,tr^c\left(\rho_f^{cm}\right)=\rho_0^c.\tag{12.15}$$

可以验证，纯态的克隆，仅仅是上面这个一般的克隆的一个特例。那么，在量子的情形，我们相当于问，是否存在量子的幺正演化过程，满足

$$\rho_0^q\otimes\rho_0^m\longrightarrow\rho_f^{qm}:tr^m\left(\rho_f^{qm}\right)=\rho_0^q,tr^q\left(\rho_f^{qm}\right)=\rho_0^q.\tag{12.16}$$

实际上，幺正演化过程可以放宽为演化加上部分求迹的过程，并且大多数一般的量子过程都可以看作先和一个外界系统合起来演化，然后部分求迹的过程 [5]。当然，也存在一些量子过程不能看作联合演化再部分求迹的过程，见 [5] 第 8.5 节“量子算符表示的局限”。如果我们限定在“先联合演化再

[2]在这个混合态的复制的意义下，量子信息领域把这样的“克隆”称为“传播”

部分求迹”的一般过程的范围内，就是相当于问是否存在这样的一般过程满足上面的条件。关于这个问题的答案，请参考“量子不可传播定理”的证明[57]：这个形如公式 (12.16) 的混合态（也包含纯态）的“克隆”在量子信息的语言里面叫做“转播”。在这里，我们只关心，经典的测量过程正好制备了这样的 ρ_f^{cm}，那么是否量子的测量过程，能够制备类似的满足要求的 ρ_f^{qm}。如果不能，为什么？

既然是理想中的测量，那么，对于那些所要测量的物理量的本征态，记为 $|l\rangle_q$，肯定是要建立起来一一对应的关系的，也就是 m 的状态必须是 $\left|j(l)\right\rangle_m$，$j(l)$ 是 l 的一对一（one-one）函数（不一定到上，onto）。在这个一一对应条件下，我们先来看一个非常一般的状态 ρ^{qm} 的部分求迹 $tr^m(\rho^{qm})$。

$$\rho^{qm} = \sum_{ln} \rho_{ln}^q |l\rangle_q \left|j(l)\right\rangle_m \langle n|_q \left\langle k(n)\right|_m . \tag{12.17}$$

通过部分迹，我们得到，

$$tr^m(\rho^{qm}) = \sum_l \rho_{ll}^q |l\rangle_q \langle l|_q . \tag{12.18}$$

我们发现，那些 ρ^q 中的非对角元素 $\sum_{l,n\neq l} \rho_{ln}^q |l\rangle_q \langle n|_q$ 都消失了[3]。对于一个任意的状态，非对角元 ρ_{ln}^q 一般不等于零。于是，对于这样的状态，我们不可能做到

$$tr^m\left(\rho_f^{qm}\right) = \rho_0^q. \tag{12.19}$$

到这里，我们得到结论：如果我们希望对于所要测量算符的本征态满足“克隆”的要求——由于它们完全正交这样的克隆是做得到的，那么，对于不是本征态的那些一般的状态，通过测量过程得到的末状态不能看作是这样的状态的“克隆”，除非，这些状态所对应的密度矩阵在这一套所测量量的本征态下的没有非对角元。也就是说，没有非对角元的密度矩阵是可以“克隆”的，有非对角元的密度矩阵不可以。

回到经典和量子的对比，两个密度矩阵的区别正好就是，在一般的基矢下，量子密度矩阵存在非对角元，而经典密度矩阵不存在非对角元。这个差别的产生的根本原因是量子状态满足矢量叠加原理，因此写成密度矩阵以后存在交叉项。或者说，其原因是量子算符不可对易，因此一般不可能用共

[3]同时，我们可以证明，这个结果不依赖于用来求迹的基矢正好就是测量指针态 $\left|j(l)\right\rangle_m$ 的事实。假设我们在另一组基矢下面求部分迹，$tr^m(\rho^{qm}) = \sum_d \sum_{ln} \rho_{ln}^q |l\rangle_q \langle d|_m \; j(l)\rangle_m \langle n|_q \langle k(n)_m |d\rangle_m = \sum_d \sum_{ln} \rho_{ln}^q |l\rangle_q \langle n|_q \langle k(n)_m |d\rangle_m \langle d|_m \; j(l)\rangle_m = \sum_{ln} \rho_{ln}^q |l\rangle_q \langle n|_q \left\langle k(n)\right|_m \; j(l)\rangle_m = \sum_l \rho_{ll}^q |l\rangle_q \langle l|_q$。

同本征向量构成完整的基矢。于是，一个表象下对角的状态，必然在另一个表象下非对角。因此，量子测量不能得到“克隆”态[4]的根本原因，还是态叠加原理。

这个经典测量相当于制备了“克隆”态，量子测量不能看做制备“克隆”态，非对角元被消灭，的事实是一个非常重要的有深远含义的结果。对于经典测量，我们一般认为比较平庸，没有对系统做什么，实际上就是因为，这个经典测量过程就可以看作给所测量的系统做了一个备份，当然，什么改变都没有做。同时，对于量子系统，由于本质上就不能不改变所测量的系统的状态——除非正好就是所测量物理量的本征态或者本征态的概率性叠加，我们就会觉得量子的测量更加神秘——因为它对所测量的系统做了点什么，消灭了所有的非对角元。

12.3 经典和量子测量的统一的框架？

可是，换一个角度来看，我们前面讨论的经典和量子测量，其实都遵循了统一的框架。首先，被测量系统有一个初始状态 ρ_0^o（不管是 ρ_0^c 还是 ρ_0^q），测量仪器也有一个初始状态 ρ_0^m。接着，通过某个相互作用 H^{om}（导致的演化算符 U^{om}），在两个系统之间建立起来关联，得到状态 ρ_f^{om}。这一个步骤完全符合量子或者经典的演化过程。我们对于这个末状态密度矩阵有一个要求：如果 ρ_0^o 对应着纯态基矢 $|l\rangle_o$，则相应地 ρ_f^{om} 对应着纯态基矢 $|l\rangle_o \left|j(l)\right\rangle_m$——也就是在给定的一组纯态基矢上的测量必须一一对应。得到状体 ρ_f^{qm} 以后，我们来计算部分迹，得到测量仪器的一个密度矩阵 ρ_f^m。然后，这个密度矩阵做概率解释：任何一次实验的结果相当于从一个概率分布中随机取一个样本 j_*，接着反过来运用一一函数 $j(l)$ 得到所测量的系统的状态是 l_*。

$$\rho_0^o \otimes \rho_0^m \xrightarrow{U} \rho_f^{qm} \xrightarrow{tr^o} \rho_f^m \xrightarrow{sampling} j_* \xrightarrow{j^{-1}(l)} l_*. \tag{12.20}$$

在测量的形式语言——也就是不关心测量的可能的动力学——描述里面，人们把公式 (12.20) 的整个过程——忽略测量仪器系统的话看起来好像我们从一个一般的可以是叠加态或者混合态的 ρ_0^o 开始得到了一个测量量的本征态 $|l_*\rangle$——称为波函数的塌缩：整个过程好像是一个叠加态波函数“塌缩”到了其中的一个基矢状态上去一样。

实际上，在以上所有步骤中，$\xrightarrow{U}$ 属于经典力学或者量子力学演化过程，

[4]或者按照量子信息的习惯称为“传播态”

$\xrightarrow{sampling}$ 在经典随机客体的状态中就是这样，$\xrightarrow{j^{-1}(l)}$ 就是一个逆函数的计算。因此，唯一神奇的地方就在于计算部分迹的过程 $\xrightarrow{tr^{o}}$。可是，在经典随机客体的测量中，我们也要运用这样的部分迹啊：它表示忽略所测量系统的状态，我们仅关心测量仪器的状态。可是这个看起来平庸的部分迹引入了一个对这个系统的非常大的改变——消灭了观测量本征态表象下的所有的非对角元。如果说，被观测系统已开始是一个纯态的话，那么，经过这个部分迹，其状态就是一个混合态。这个过程不是满足幺正演化条件的量子力学和经典力学的演化能够描述的[5]。

那么，是否存在着一个不需要求部分迹这一步骤的对测量过程的描述？这个是一个没有得到回答的问题，尽管答案看起来是“否”，因为在我们的描述里面，就算经典客体状态的测量也需要这个求部分迹的步骤。如果这样一个超越幺正演化的步骤确实需要的话，在真实的实验过程中，什么时候需要我们来做这样的一个部分迹——从而消灭非对角元也就是量子相干性，什么时候留着非对角元——从而保持着量子相干性？例如，在 Stern-Gerlach 实验中，这个部分迹到底是在电子过了磁场以后来做呢，还是电子到达屏幕的时候来做呢？为什么？把通过磁场分开的电子重新通过合起来的实验（见本书实验 (1.10) 以及 [8] 的第八、九章）我们看到在分开路径之后没有到达屏幕之前，实际上相干性都被保持着，重新合起来之后仍然有相干性。于是，一个自然的问题，就是，屏幕的什么特性使得这样的相干性消失了呢？如果把屏幕换成是 Schrödinger 的猫呢？换成是细菌呢？换成了 C_{60} 球呢？这个什么时候需要做部分迹计算从而消灭相干性的问题仍然没有答案，尽管我自己的态度是这个问题不需要回答。我认为（不是科学实验或者科学理论计算的结果，仅仅是认为）只要牵涉到多个子系统构成的整体系统的问题，当我们关心子系统的状态的时候，就一定要做部分迹，不管是量子还是经典系统。也就是说，这个部分迹的操作不是一个物理过程，不发生在实际的时间和空间里面，而是在逻辑上。结果到底得到一个平庸的复制还是得到不平庸的混合态，完全看当时整体系统的状态：纠缠比较弱（整体系统接近一个直积态）则得到的态更像复制态，比较强（整体系统偏离直接态非常大）则得到可能的混合态。

本章讨论的测量都是投影测量，关于更一般的测量 [5] 的动力学的讨论可以建立在这个投影测量的基础上。

[5]很容易证明幺正演化保证纯态到纯态，混合态到混合态，见习题 7.7，习题 7.8。

12.4 作业

习题 12.1 给定一个 GHZ 状态，$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\uparrow\rangle - |\downarrow\downarrow\downarrow\rangle)$，计算”先后”测量 $\sigma_x^1, \sigma_x^2, \sigma_x^3$ 的结果，”先后”测量 $\sigma_x^1, \sigma_y^2, \sigma_y^3$ 的结果。按照这个结果，你可以做一些关于 GHZ 实验的拓展阅读。

习题 12.2 再一次看 Coleman 的讲座《Quantum Mechanics in Your Face》，更新你以前的报告。

习题 12.3 再一次看 Susskind 的网上课程《Quantum Mechanics》的前两课，更新你以前的报告。

习题 12.4 再一次看 Feynman 的《Feynman 物理学讲义》第三卷的前三章，更新你以前的报告。

12.5 本章小结

在本章中，我们企图去构建一个量子测量的描述过程，而不再仅仅通过运用形式化的关于测量的公理3和4来讨论问题。我们发现，测量的过程的第一步其实就是建立起来从被测量系统到测量仪器的状体之间的关联。而这个建立关联的步骤完全可以放到量子力学的演化过程这个量子力学公理9来描述。测量过程的其他步骤却不能从量子力学的其他公理推导出来。如果所有的步骤都可以用其他公理来代替，那么，我们也就不需要独立的关于测量的公理3和4了。

建立起来关联之后，我们发现，不管是经典还是量子的测量，都需要做一个“求部分迹”的运算。其中，对于经典状态，这个运算不引起任何的改变；对于量子状态，这个运算消灭了被测量量本征态表象下的非对角元。完成这一步计算之后的状态的解释完全可以通过经典随机变量的测量来理解。因此，量子测量的问题就成了：为什么和什么时候需要用这个“求部分迹”运算？关于这个问题，量子力学还没有好的答案。

第五部分

量子信息简介

第十三章　远程传态和 Deutsch 算法

优秀的量子计算的算法通常是巧妙地利用了纠缠态：考虑给对象系统配上一个辅助的系统，建立起来关联，然后通过部分测量的方式把状态在对象系统和辅助系统之间转换。所谓的“量子并行性”——量子算符能够一次性地作用在叠加态上从而相当于作用在大量的基矢上，不能直接用来加快计算速度。

实际上，我们还会看见量子远程传态，和量子计算类似，也是对纠缠态的一种巧妙运用。我们先来看看在远程传态中量子纠缠的作用，以及状态在对象系统和辅助系统之间的转换。然后再来看一个量子计算的例子。

本章的例子都选了最简单的，最不需要额外学习一些基础的问题。这样的选择不能非常好地反映这一章的主题——纠缠的作用和状态在不同系统之间转换的非凡价值。因此，本章真的仅仅是一个入门，有兴趣的读者需要做一些自学。量子信息和量子计算的书籍已经不少，但是，经典的 Nelson 和 Chuang 的《Quantum Computation and Quantum Information》[5] 和 Preskill 的《Lecture Notes for Physics 229: Quantum Information and Computation》[6] 还是最推荐的两本。

13.1　量子远程传态

我们已经知道，对于一对经典关联起来的处于 $\rho = \frac{1}{2}\left(|11\rangle\langle 11| + |00\rangle\langle 00|\right)$ 的真随机硬币，任何时候的观察都得到相同的结果：例如都是正面——记为 $|11\rangle\langle 11|$ 或者反面——$|00\rangle\langle 00|$ 中的一个。对于完全纠缠起来的处于 $\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right)$ 一对自旋，如果我们测量的是同样的方向——可以是任意的 $\hat{r}$ 方向但是两个自旋的被测量方向要一致——则也得到都是正面

或者都是反面的结果。甚至，当两个方向不相同的时候，在这个状态上做的量子测量得到的结果也具有特定的关联，见第二章和第十一章。这样的量子的完全纠缠态的性质比经典的更好，对于无论哪一个方向的测量这个结果都成立。于是，我们来思考，是否能够使用这样的状态来实现状态的远程传输：能不能把一个纯态从 Alice 传给距离非常远的 Bob?

最愚蠢的办法是 Alice 先对这个状态做一个测量，得到完整的描述。例如自旋的纯态需要两个参数（$\cos(\theta)|\uparrow\rangle + \sin(\theta)e^{i\phi}|\downarrow\rangle$），每一个参数需要通过大量的测量来确定。然后，把这个测量出来的参数通过经典渠道，例如电话，传给 Bob。接着 Bob 设计一个实验仪器，把这样的状态制备出来。如果是自旋的纯态，那么选择磁场方向和作用时间，能够实现这样的参数的任意的状态。那远程传态的问题是：如果 Alice 不去做大量的测量，能否把这个状态正确地传给 Bob 呢?

量子不可克隆定理已经保证未知的纯态不能被复制，因此，如果 Alice 从拿着一个状态为 $|\phi^o\rangle$ 的量子客体 o 开始，如果最后 Alice 和 Bob 都有一份 $|\phi\rangle$ 的话，就违反了不可克隆定理。因此，传态的结果不可能是复制。由于相对论的限制，也不能是是传物——纠缠态的关联是瞬间的，因此不能指望瞬间的“过程”能够把一个物体送到远处。那到底传的什么呢？把一个量子客体的状态在未知的情况下传给远处的另外一个客体。其特点第一是远处，第二是未知。我们下面还会发现，对纠缠态的运用和把量子态矢量按照系统的不同部分来重新分解和合成，是另一个特点。而且，这个特点，在很多量子计算问题中都有体现。在这三个特点中，其实远程和未知传态在经典系统中能够，通过运用经典关联态，来实现。不能实现的是状态的分解和合成。

下面，我们仔细来考察一个未知状态

$$|\phi^o\rangle = \cos(\theta)|\uparrow\rangle + \sin(\theta)e^{i\phi}|\downarrow\rangle \tag{13.1}$$

是如何被 Alice 传送给 Bob 的。当 Alice 拿到这个状态 $|\phi^o\rangle$ 之前，Alice 和 Bob 已经分别拿到了纠缠起来的两个自旋中的一个，其纠缠态是，

$$\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right). \tag{13.2}$$

现在，三个自旋合起来的状态是

$$|\Phi\rangle = |\phi^o\rangle\left|\Phi^+\right\rangle = \left(\cos(\theta)|\uparrow\rangle + \sin(\theta)e^{i\phi}|\downarrow\rangle\right)\frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right). \tag{13.3}$$

下一步就是神奇的一步，Alice 对自旋 o 和她自己的那个自旋合起来，做 Bell 态的测量，也就是按照测量所得到的结果对 $|\Phi\rangle$ 做一个部分投影（并且重新归一化）

$$\langle\Phi_+|\,\Phi\rangle = \cos(\theta)\,|\uparrow\rangle + \sin(\theta)e^{i\phi}\,|\downarrow\rangle \tag{13.4a}$$

$$\langle\Phi_-|\,\Phi\rangle = \cos(\theta)\,|\uparrow\rangle - \sin(\theta)e^{i\phi}\,|\downarrow\rangle \tag{13.4b}$$

$$\langle\Psi_+|\,\Phi\rangle = \cos(\theta)\,|\downarrow\rangle + \sin(\theta)e^{i\phi}\,|\uparrow\rangle \tag{13.4c}$$

$$\langle\Psi_-|\,\Phi\rangle = \cos(\theta)\,|\downarrow\rangle - \sin(\theta)e^{i\phi}\,|\uparrow\rangle\,. \tag{13.4d}$$

上面这 4 个状态分别对应着测量可能得到的 4 个结果，编号为结果 $1,2,3,4$。这个编号是 Alice 和 Bob 的共同知识。于是，我们发现，如果我们测量得到的结果是 $|\Phi_+\rangle$，Bob 的自旋的状态直接就是我们所希望得到的 $|\phi^o\rangle$。如果我们得到的状态是 $|\Psi_-\rangle$，我们需要把 $\cos(\theta)\,|\downarrow\rangle - \sin(\theta)e^{i\phi}\,|\uparrow\rangle$ 变成我们需要的 $|\phi^o\rangle$。这个非常容易实现，例如

$$\sigma_y\Big(\cos(\theta)\,|\downarrow\rangle - \sin(\theta)e^{i\phi}\,|\uparrow\rangle\Big) = -i\Big(\cos(\theta)\,|\uparrow\rangle + \sin(\theta)e^{i\phi}\,|\downarrow\rangle\Big). \tag{13.5}$$

也就是说，Alice 做完测量结果之后，需要告诉 Bob 得到哪一个结果，然后 Bob 按照这个信息来决定对自己的自旋做一个怎样的操作。例如，得到结果编号为 1 的结果——对应着 $|\Phi_+\rangle$，那么 Bob 什么都不用做——对应于算符 I。这个经典信息传递包含 4 个值，过程需要两个位来编码。这个传递过程是经典的，不能超光速。

如果你看其他的量子信息的书籍，例如 [5]，由于其中用到的 Alice 的测量是在通常基矢 $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$ 上，因此过程会稍微复杂一点。相当于是需要建立起来 Bell 态和通常基矢之间的一个映射。从思想上说是一样的：在这里，理论上对整体状态我们做了一个部分系统状态的分解与合成；实验上，我们测量了部分系统 o 和 Alice 的自旋。经过这个部分系统的状态的分解与合成，我们就把 o 系统状态前面的系数转给了 Bob 的自旋。这一点，在量子信息和量子计算中是非常重要的。对于这一步的完成，很多时候，我们需要纠缠的纯态。

因此，我们发现，第一，部分系统的测量非常重要；第二，纠缠是重要的通过部分系统的测量来实现状态的传送、改变的重要工具。

下面，我们来演示一个量子计算的算法问题，从另一个角度来说明以上两点——部分系统的测量和纠缠的重要意义。

13.2 Deutsch 算法

为了和 [5] 的符号一致，这一节我们用 $|0,1\rangle$ 来代表 $|\uparrow\downarrow\rangle$。

在演示这个量子算法之前，我们来讨论一下所谓的“量子并行性”的问题。由于量子态的可叠加性，很多人认为量子计算区别于或者说优于经典计算的根本原因就是“并行性”——在一个量子态上的计算就相当于计算了某个基矢集合下的所有的矢量。例如，考虑计算一个函数 $f(x)$，而 $x=0,1,...,2^n-1$。为了简单性，我们假设 $f(x)$ 的取值也是 0 到 2^n-1 之间的整数。在经典计算上，我们需要做一个循环，一个接一个来计算，或者采用多个计算核心把任务分开来计算。在量子计算上，确实，我们有可能直接把所有的这样的数 $f(0),f(1),...,f(2^n-1)$ 一次性地计算出来。例如，我们可以考虑让 $\hat{f}$ 所对应的量子操作作用到下面这样一个矢量上去，

$$\hat{f}|\phi\rangle=\sum_{j=0}^{2^n-1}\phi_j\hat{f}|j\rangle=\sum_{j=0}^{2^n-1}\phi_j\left|f(j)\right\rangle. \tag{13.6}$$

为了更具体，我们甚至可以让 $\phi_j=\frac{1}{\sqrt{2^n}}$。于是，看起来，一次计算，我们就实际上做了所有的 $f(x)$ 的计算。在有的问题中，我们确实希望能够做到一次性把 $f(x)$ 算出来，并且做后续的操作，例如在量子搜索问题 [5] 中我们想看一看哪一个对应于零，也就是寻找 x_0 使得，

$$f(x_0)=0. \tag{13.7}$$

但是，如果我们需要把这样的计算结果取出来并检查，我们需要做测量。例如，我们可以在这样的初始状态上得到的末状态上做一个测量。但是，每次从这个末状态测量得到的唯一的结果（考虑投影测量，也就是仅得到所有 2^n 个状态中的一个，而不是全部）我们不知道实际上得到的状态会是 $\hat{f}$ 作用在哪一个态 $|j\rangle$ 上的结果。因此，都算一遍，很容易实现，但是一次性取出所有结果，不可能。

上面这个计算过程确实一次性计算了所有结果但是取不出来。就算取出一个来，也不知道是哪一个初始状态的计算结果：各个状态出现的几率是一样的。于是，我们自然思考是否可以把来自于哪一个初始状态记录下来，并且检验是否这样就可以一次性地把所有的计算结果都取出来。在这里，我们要的思路是，如果要否定一个东西，要尽量在这个东西成立的条件都满足的情况下，来否定。

现在，我们来构造一个记住初始状态的过程。如果对于 CNOT 门还有映像，它实现了如下的计算（符号采用 [5]），

$$U_{CNOT}\left|x,y\right\rangle = \left|x,y\oplus x\right\rangle, \tag{13.8}$$

其中运算 $\oplus$ 是加起来然后再除以 2 所得到的余数。这个运算当 $x=0$ 时，$y\oplus x=y$；这个运算当 $x=1$ 时，$y\oplus x=1-y$，翻转 y。我们也可以让 $y=0$，于是

$$U_{CNOT}\left|x,0\right\rangle = \left|x,x\right\rangle. \tag{13.9}$$

我们就有了 x 的复制。不过，这个时候，$x=0,1$ 两个正交状态而不是一般的未知态，不违反不可克隆定理。由于这个状态保留着一个自旋的状态 x，于是，我们希望能够这个运算来实现

$$U_f\left|x,y\right\rangle = \left|x,y\oplus f(x)\right\rangle. \tag{13.10}$$

如果成功，那么，

$$U_f\frac{\left|0,0\right\rangle+\left|1,0\right\rangle}{\sqrt{2}} = \frac{\left|0,f(0)\right\rangle+\left|1,f(1)\right\rangle}{\sqrt{2}}. \tag{13.11}$$

在这个状态下，每次测量的结果是某一个状态 $\left|x,f(x)\right\rangle$。于是，不仅函数值得到了，自变量也得到了。不过，即使如此，我们每一次测量的结果仍然仅仅得到公式 (13.11) 中的一项，而不是全部。因此，我们发现，即使引入映射 U_f，仍然不能一次性取出所有信息。

下面我们来补充说明，这样的 U_f 原则上是可以存在的。U_f 是一个幺正变换。例如可以验证 $U_fU_f\left|x,y\right\rangle=\left|x,y\right\rangle$[1]。量子操作最好保持可逆性，这样容易通过找到某个 H 来实现 $U=e^{-iHt}$。实际上之前在第十章的导致纠缠的演化算符就是这样构造出来，然后反过来做为本书的习题的。当然，必要的时候量子操作也可以通过演化加上部分系统的测量来实现，不过这个时候的量子计算是不可逆的。如果允许这样的不可逆操作，那么，我们实现 $\left|x,x\right\rangle\to\left|x,f(x)\right\rangle$ 的方式就更多了。例如，保留第一个 x 的自旋，让第二个 x 的自旋经过经典计算量子化以后的过程。关于经典计算的量子化对应，可以从 [5] 找到。

既然不能一次性取出来所有信息——尽管它们包含在 U_f 的作用结果里面，愚蠢的方法就是，按照经典的思路，按顺序制备一次一个不同的初态，

[1] 可以验证，$\left|x\right\rangle\to\left|f(x)\right\rangle$ 一般是不可逆的。于是，在这里，辅助系统 y 的引入是重要的。

然后放到这个代表 $\hat{f}$ 的相互作用 H 中去，得到唯一的末状态，然后做测量，得到确实处于这个末状态，识别出来。这样的话，我们需要重复 2^n 次制备、过 H、测量的过程。也可以总是从类似于 $\frac{|0,0\rangle+|1,0\rangle}{\sqrt{2}}$ 的状态出发、经过 H、做 2^n 次测量，得到完整的 $f(x)$ 信息。这样的过程的复杂度完全和经典计算的复杂度是一样的。而且，这样的做法，幸好是经典算符 $f(x)$，给定 x 其取值仅是 2^n 个整数中的一个。如果量子的情形，可以存在叠加态的话，那么，针对每一个初态，我们都需要大量的过 H 和测量。后面的这个量子的情形的问题被称作量子过程解析（Quantum Process Tomography）的问题 [5]。

因此，我们看到了，如何把确实已经包含在“量子并行性”里面的信息拿出来是关键问题，“量子并行性”本身不解决问题。而这个拿出来信息的过程，很多时候，就需要依靠前面说的两条——部分系统的测量和纠缠。

现在，我们来看一个通过一次计算把某个函数的某种函数 $F(f(x))$ 计算出来的问题：有一个确定的函数 $f(x)$，$x=0,1$（可以更一般，只要考虑多个 qubit 版本的算法），我们希望计算 $f(0)\oplus f(1)$。这个问题本身基本上没有实际意义，但是，我们会看到运用量子力学，Deutsch 确实找到了办法只需要调用一次 $f(x)$ 的计算。经典算法的话，需要调用两次函数 f 的计算过程，分别得到 $f(0)$ 和 $f(1)$，然后把两者用来计算 $f(0)\oplus f(1)$。其中所用到的用 CNOT 门来建立对象系统和辅助系统之间的关联的做法，把状态的系数从一个系统写到另一个系统上去的看问题的角度——也就是状态在不同系统之间的分解和合成，是具有一般意义的。同时，Deutsch 也启发了更加一般的应用非常广泛的量子 Fourier 变换 [5] 的提出。

这个算法依赖于一个外部函数 $f(x)$。假设这个函数已经有一个单独的量子操作模块提供。如何实现这个外部函数不是我们讨论的内容。我们要揭示的是一个算法通过调用这个专门的模块一次来计算 $f(0)\oplus f(1)$——记住，经典的情况要计算两次。更一般的 $F(f(0),f(1))$ 也不是我们讨论的内容。我们要利用上面构造的映射 U_f，

$$U_f|x,y\rangle=\big|x,y\oplus f(x)\big\rangle. \tag{13.12}$$

在构造这个映射的过程中，我们引入了辅助系统 y 而不仅仅是所计算的系统 x。有了这个映射，剩下的就简单了。

从初始状态

$$|\phi\rangle=\frac{|0\rangle+|1\rangle}{\sqrt{2}}\frac{|0\rangle-|1\rangle}{\sqrt{2}} \tag{13.13}$$

开始，我们把 U_f 作用上去，

$$\begin{aligned} U_f|\phi\rangle = & \frac{1}{2}\left(\left|0,f(0)\right\rangle - \left|0,1-f(0)\right\rangle + \left|1,f(1)\right\rangle - \left|1,1-f(1)\right\rangle\right) \\ = & \frac{1}{2}\left(|0\rangle\left(\left|f(0)\right\rangle - \left|1-f(0)\right\rangle\right) + |1\rangle\left(\left|f(1)\right\rangle - \left|1-f(1)\right\rangle\right)\right) \\ = & \frac{|0\rangle + (-1)^{f(0)\oplus f(1)}|1\rangle}{\sqrt{2}}\frac{\left|f(0)\right\rangle - \left|1-f(0)\right\rangle}{\sqrt{2}} \end{aligned} \tag{13.14}$$

我们对辅助系统的状态仔细观察，发现，其实，它没有变，肯定还是，

$$\frac{\left(\left|f(0)\right\rangle - \left|1-f(0)\right\rangle\right)}{\sqrt{2}} = \pm\frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{13.15}$$

整体乘上的常数可以忽略不计。我们发现原来 U_f 改变的是辅助系统的状态，但是经过我们的再次分解和组合，我们发现，实际上，我们改变了对象系统的状态，

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \to \frac{|0\rangle + (-1)^{f(0)\oplus f(1)}|1\rangle}{\sqrt{2}}. \tag{13.16}$$

于是，如果我们进行 x 方向本征态的测量，也就是 $\frac{|0\rangle\pm|1\rangle}{\sqrt{2}}$ 基矢下的测量，我们得到，

$$\begin{cases} |\uparrow_x\rangle & \text{if } f(0)\oplus f(1) = 0, \\ |\downarrow_x\rangle & \text{if } f(0)\oplus f(1) = 1. \end{cases} \tag{13.17}$$

于是，按照我们测量所得到的结果，我们就知道了 $f(0)\oplus f(1)$ 的值。

当然，这个计算的实际意义有限，而且实现一个一般的函数 $f(x)$ 的 U_f 是有一定的难度的，更何况一般的函数的函数 $F(f(x))$。从这个算法我们需要看见的事情是，对于某些函数，我们可以构造一个比经典算法效率更高的量子算法，并且，在这个构造的过程中，引入辅助系统并且把状态在对象系统和辅助系统之间分解和合成是非常重要的。另外，我们计算的初始状态（上一节的远程传态中是 Bell 态，这一节的 Deutsch 算法中是 x 方向的本征态），总是计算基矢（指的是 $|0,1\rangle$ 或者 $|00,01,10,11\rangle$ 这样的态）的叠加态。在一点，也是量子计算的独特之处。不管是状态在系统之间的分解和合成，还是计算基矢的叠加态，其能够成立的条件都是量子态的叠加原理。

因此，所谓量子计算，就是对某些问题创造性地运用好量子态的叠加原理。

13.3 作业

习题 13.1 检索和阅读“量子不可传播定理”，做一个综述，包含整理、总结、评价。

习题 13.2 检索和阅读“量子 Fourier 变换”，做一个综述，包含整理、总结、评价。

习题 13.3 验证书中从公式 (13.3) 到公式 (13.5) 的计算。考虑如果得到编号为 2 的结果，Bob 应该如何做。

13.4 本章小结

在这一章里面，我们讨论了量子远程传态和 Deutsch 算法。实际上在这一章里面的所有的讨论都很大程度上借鉴了 [5]。连本章强调的深入理解的地方都是和 [5] 一样的：量子并行性不能直接提高量子计算的效率、纠缠态以及更根本的状态的分解和叠加对于实现量子计算和远程传态等操作是非常重要的。那么，本章的意义在哪里呢？第一，我们还强调了把整个系统的状态从一个子系统的状态分解合成到另一个子系统的状态上面——相当于对整个系统的某一部分的操作可以看作对另一部分的操作。第二，我们选择了最简单的例子来体现以上的思考和理解。这个是本书基本选材原则的又一次体现：选得少，理解的透彻，还希望启发读者思考为什么这样选择。这一章，甚至包含下一章，我们其实都在体现这样一个主题：量子态的叠加原理使得我们能够做经典做不到的计算或通讯（以及下一章的博弈）。

任何一本书，任何一门课，其核心任务，绝对不是传授知识，而是传授对知识的组织和理解，中间当然经常需要通过传授知识来体现。因此，在内容选择上，一定要有背后的思考，一定要有统领全文的目标。书和课程有这些思考和目标，读者和学生才会有思考和目标。进一步，这些被选择的知识一定要尽量地少 [17]，在体现这些思考和目标的前提下。这个就是我们一直的呼吁和实践的“Teach Less, Learn More” [10] 学习和教学体系的精神，见图图 13.1。

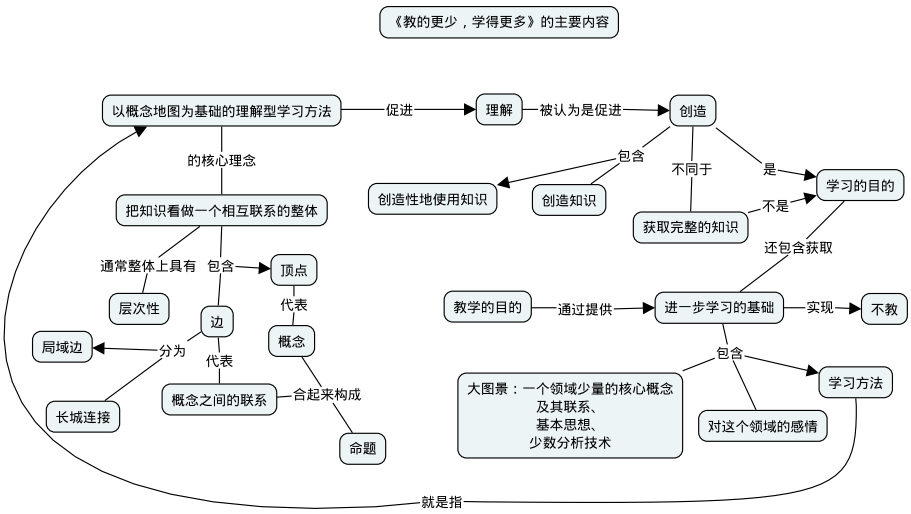


图 13.1: 学习的目的是为了创造知识和创造性地使用知识，为了这个目的，我们要强调帮助学生提升对知识的理解和组织，提升对这个学科的感情，掌握学习方法，而不是知识本身。

第十四章　量子博弈

在本书中，我们一再强调，量子系统和经典系统的区别是：经典系统的状态由一个概率分布函数描述，量子系统由一个密度矩阵描述，后者具有非对角元，也就是相干项。通过前面的学习，我们还知道了这个数学形式上的差别是由量子系统的行为和经典系统的行为的不同导致的。现在我们来看看这个数学结构上的区别导致的人类决策行为上的区别。

顺便，我们也提一下所谓的量子决策理论——在那里人们形式上用波函数，也就是密度矩阵，来描述人的决策。这个形式上的描述和我们下面要讨论的量子博弈是不一样的。在我们这里，用密度矩阵来描述决策状态具有坚实的理论基础。当然，这不是说，形式上采用波函数的量子决策理论就没有意义，仅仅是说其没有坚实的理论和实践基础。

这个问题，还牵涉到，到底什么时候，我们用经典的模型或者量子的模型来解释实际系统的行为的问题。一般来说，尽管我自己相信原则上一切都是量子的，这取决于理论和行为的比较：如果经典模型已经大部分能够解释观察到的行为，那么，对于这个现象，我们就不需要量子模型来解释了。那么，下一个问题就是，如果这样的话，量子的现象和模型从哪里来呢？一方面，当然是来自于已经出现的超过经典模型的现象，例如当年的光电效益、黑体辐射、电光子干涉、Stern–Gerlach 实验等等。另一方面，理论上，我们也可以通过量子化经典模型、经典现象的方式来启发从经典对象中构造类似的相对应的量子对象，然后通过实验来研究这样的量子系统——它们是通过量子化经典系统得来的。

当然，以上都是“一般来说”。既然如此，就还有“特殊情形”。我们说过，科学就是一个（成体系的）能够用于解释和理解现实的还没有被证伪的理论模型。因此，理论模型本身是否具有坚实的可以理解的基础，是无所谓的，尽管一般情况下，我们希望它有基础、可理解。于是，一个理论，只要它能够用来计算分析，又能够得到跟实际观察相符的结果，总是有价值的。

如果这个结果还没有其它的理论来解释，那就更有价值了。到底量子决策，还有量子博弈，属于那一类，一般来说的还是特殊情况的，还是说其实不是科学的，就要看后续的研究了。

本章用到了一点点博弈论，一点点李群。尽管真的是一点点，但是有对这两个内容的一点点比较深刻的理解——了解什么是博弈、博弈理论解决什么样的基本问题、博弈的数学描述怎样、什么是群、群和操作或者说算符的关系、李群是同时有群员之间的乘法和加法（甚至可以做微分，无穷小量的加法）的群，就够了——能够对这一章的内容有更好的理解。

14.1 经典客体上的操作和博弈

我们先来看经典博弈。翻硬币游戏是一个典型的经典博弈。我们就从这里开始。

翻硬币游戏是这样的。有一个硬币，初始正面向上。有两个参与者，我们称为 Alice 和 Bob。他们可以选择去“翻”（记为 X）或者“不翻”（记为 I）这个硬币。选择完成以后，先把 Alice 的操作作用到硬币上去，然后再把 Bob 的操作作用到硬币上，然后，得到硬币的末状态。最后，按照这个末状态，来给 Alice 和 Bob 收益。例如，如果硬币还是向上则 Alice 得到 1 个单位的收益，Bob 失去一个单位，否则倒过来。

当我们把这个博弈的结构搞得更加清楚的时候，我们就会发现，实际上，我们可以把硬币这个具体的东西忘了，仅仅考虑双方的策略。例如我们得到：当 Alice 选择 I，Bob 选择 I 的时候，Alice 得到收益 1，Bob 得到收益 -1，简记为 $E^1(I,I)=1=-E^2(I,I)$。用这个记号，我们可以得到，$E^1(I,X)=-1=-E^2(I,X)$，$E^1(X,I)=-1=-E^2(X,I)$，$E^1(X,X)=1=-E^2(X,X)$。写成一个“矩阵”的形式，有

$$G^{1,2}=\begin{bmatrix}1,-1 & -1,1\\ -1,1 & 1,-1\end{bmatrix}, \tag{14.1}$$

或者用更加复杂一点的记号来表示上面的翻硬币博弈，

2 \ 1	I	X
I	$1,-1$	$-1,1$
X	$-1,1$	$1,1$

,

其核心还是上面的矩阵 $G^{1,2}$。因此，我们以后就用这个 $G^{1,2}$ 这样的矩阵来表示一个经典博弈。在这个矩阵的层次，策略 I, X 实际上是代表了什么样的做用于经典客体的操作这件事情，我们已经完全看不见了，也不再关心了。这样的矩阵，称为收益矩阵，就是经典博弈论的基本数学形式，也称为经典博弈的抽象定义。

然后，用列矢量来代表博弈者 Alice 和 Bob 的策略，分别用 $\mathcal{P}^1 = [p^1, 1-p^1]^T$（表示以概率 p^1 取策略 I）和 $\mathcal{P}^2 = [p^2, 1-p^2]^T$（表示以概率 p^2 取策略 I）来表示，那么在给定的任何一组策略组合的情况下，Alice 和 Bob 的收益由以下公式决定，

$$E^{1,2} = \left(\mathcal{P}^1\right)^T G^{1,2}\mathcal{P}^2. \tag{14.2}$$

有了这个定义的一般形式以后，任何的经典博弈的定义，最关键的就是找到这个矩阵 $G^{1,2}$。而有了这个 $G^{1,2}$ 之后，博弈论理论的任务就是对于给定的任何一种 $G^{1,2}$ 找到任意博弈者 i 的预期行为 $\mathcal{P}^{i,*}$。

从理论上说，给定一个博弈，只要能够求出来这样的任意一个博弈者 i 确实按照 $\mathcal{P}^{i,*}$ 来采取行动的解，就解决了博弈的所有的问题，也就能够做理论和实际博弈者行为之间的比较了。也就能够对现实社会的情况来构造合适的博弈模型来解释现实和指导规则的制定了。Nash就提出了这样的一个博弈的解的概念，称为Nash 均衡$\mathcal{P}_{eq}^{1,2,\cdots,N}$。对于非合作博弈，当系统中的所有的博弈者 $i, i = 1, 2, \cdots, N$ 都处于 $\mathcal{P}_{eq}^{i}$ 以至于合起来系统处于 $\mathcal{P}_{eq}^{1,2,\cdots,N}$ 的时候，任何一个博弈者独自的偏离都不会给这个博弈者带来额外的收益，也就是，

$$E^i\left(\mathcal{P}^i, \mathcal{P}_{eq}^{-i}\right) \le E^i\left(\mathcal{P}_{eq}^{i}, \mathcal{P}_{eq}^{-i}\right), \forall \mathcal{P}^i. \tag{14.3}$$

并且Nash证明了Nash 定理[58]：对于任何有限策略的非合作博弈，这样的Nash 均衡总是存在的。于是，看起来博弈论的问题就成了如何把这样的Nash 均衡求解出来的问题了。

当然，实际上，这个解的计算比较困难，有的时候存在多个这样的解，还有的时候所有的这样的理论解和实际博弈者的行为存在比较大的差别。因此，实际上，经典博弈本身也是一个有很多研究工作可以做的学科。在这里，我们不关心这些经典博弈论的问题，而是来看一下，如果我们的操作不是针对硬币的而是针对量子自旋的，会怎么样？理论的基本形式和基本结构还会一样吗？

为了解决这个问题，我们先把上面讨论过的经典博弈的抽象形式和操作性定义，明确地写下来。

首先，我们有一个经典客体的初始状态 ρ_0^c。接着，对于这个初始状态，我们可以进行操作 $L\left(U^{(1)},U^{(2)},\cdots,,U^{(N)}\right)$——这个操作本身由所有的博弈者的选择确定，操作完了得到末状态 ρ_f^c。这个从所有博弈者的选择到整体操作算符的映射可以是简单地以某种顺序把各个博弈者选择的操作相乘，例如 $L\left(U^{(1)},U^{(2)},\cdots,,U^{(N)}\right)=U^{(1)}U^{(2)}\cdots$。最后，按照这个末状态，我们来决定每一个博弈者的收益，其中 $E^i=tr\left(P^i\rho_f^c\right)$。在这里，我们用记号 P^i 表示博弈者 i 的支付方式。下面给出一般的定义和举例。

定义 14.1 经典博弈的纯策略操作性定义 $\Gamma^{op}\left(\rho_0^c,O,\left\{P^i\right\}\right)$：经典客体 C 的初始状态为 ρ_0^c，其上相应的算符集合记为 O。博弈者 i 选择操作 $U^i\in O$，从而构成所有博弈者整体的操作 $L\left(U^{(1)},U^{(2)},\cdots,,U^{(N)}\right)=U^{(j_1)}U^{(j_2)}\cdots U^{(j_N)}$，其中 $(j_1,\cdots,j_N)$ 是一个 $(1,2,\cdots,N)$ 的全排列。利用这个整体操作，得到客体 C 的末状态，

$$\rho_f^c=L\rho_0^cL^\dagger. \tag{14.4}$$

于是，博弈者 i 的收益，由这个末状态还有收益支付规则 P^i 决定，

$$E^i=tr\left(P^i\rho_f^c\right). \tag{14.5}$$

当这个经典客体是一个由每一个博弈者控制一个自己的“硬币”构成的时候，如果我们把博弈者的策略集合看作是自己控制的个体的算符集合的时候，L 是一个直积运算 $L\left(U^{(1)},U^{(2)},\cdots,,U^{(N)}\right)=U^{(1)}\otimes U^{(2)}\otimes\cdots$。但是，如果这个时候，我们把博弈者的策略看作是所有的硬币构成的符合经典客体所在的操作集合里面的元素，那么，自然，$U^{(1)}\to U^{(1)}\otimes I^{(2)}\otimes\cdots$（注意到直积算符之间的乘积就是把相应空间的算符乘在一起，然后直积），于是，上面的直积运算，自然回到了定义里面的乘积运算。因此，从现在开始，我们不再区分这个经典客体是一个单纯的经典客体，还是一个由所有博弈者自己所控制的经典客体综合起来的复合经典客体。同时，我们把各个博弈者选择的操作看作是这个整体的经典客体——不管是单纯的还是复合的——上的操作集合里面的元素。

例 14.1 (**翻硬币游戏的操作性定义**)：硬币的初始状态 $\rho_0^c=\begin{bmatrix}1&0\\0&0\end{bmatrix}$，其上的操作集合 $\{I,X\}$。Alice 和 Bob 选择其中的一个操作，记为 $U^{(1)}$ 和 $U^{(2)}$。

作用到初始状态上得到末状态，$\rho_f^c = U^{(2)}U^{(1)}\rho_0^c\left(U^{(1)}\right)^\dagger\left(U^{(2)}\right)^\dagger$，相当于 $L\left(U^{(1)},U^{(2)}\right) = U^{(2)}U^{(1)}$。得到末状态之后，按照以下收益支付方式来决定收益，$P^1 = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = -P^2$。

下面，我们来看这个操作性定义是如何转化成为一般的经典博弈的抽象定义的。

定义 14.2 经典博弈的纯策略抽象定义：经典博弈者 i 从 M 个元素的可选操作集合 O 选择一个操作 $U^{(i)}$，从而构成所有博弈者的整体策略状态 $U^{(1)},U^{(2)},\cdots,U^{(N)} \in O^N$。每一个博弈者 i 的收益由映射 G^i 给出，

$$G^i\left(U^{(1)},U^{(2)},\cdots,U^{(N)}\right)\Big|_{O^N\to R}. \tag{14.6}$$

由于两个定义要给出相同的收益，我们有，

$$G^i\left(U^{(1)},\cdots,U^{(N)}\right) = tr\left(P^iL\left(U^{(1)},\cdots,,U^{(N)}\right)\rho_0^cL^\dagger\left(U^{(1)},\cdots,,U^{(N)}\right)\right). \tag{14.7}$$

例 14.2 (**翻硬币游戏的纯策略抽象定义**)：Alice 和 Bob 的操作集合 $\{I,X\}$。Alice 和 Bob 选择其中的一个操作，记为 $U^{(1)}$ 和 $U^{(2)}$ 得到的收益由公式 (14.1) 中的收益矩阵决定，

$$G^{1,2} = \begin{bmatrix}1,-1 & -1,1\\ -1,1 & 1,-1\end{bmatrix}. \tag{14.8}$$

例如，当两者都选择 I 的时候，Alice 的收益是 1，Bob 是 -1。

接着，我们给出从纯策略抽象定义到混合策略的抽象定义。所谓混合策略就是一个经典策略集合里面的策略按照某种概率叠加起来。

定义 14.3 经典博弈的混合策略抽象定义：经典博弈者 i 从 M 个元素的可选操作集合 O 中以一定的概率 $p_j^{(i)}$ 选择一个操作 U_j。每一个博弈者 i 的收益由映射 G^i 按照概率平均给出，

$$E^i = \sum_{j_1=1\cdots M;\cdots} p_{j_1}^{(1)}\cdot p_{j_2}^{(2)}\cdots p_{j_N}^{(N)}G^i\left(U_{j_1},\cdots,U_{j_N}\right). \tag{14.9}$$

其中取和操作要让所有的 j_k 都跑遍 $1,2,\cdots M$。

例 14.3 (**翻硬币游戏的混合策略抽象定义**)：Alice 和 Bob 的操作集合 $\{I, X\}$。假设 Alice（Bob）的混合策略是以 p^1（p^2）的几率选择 I，那么，其收益由按照概率叠加计算得到

$$\begin{aligned} E^1 &= p^1 p^2 G^1(I,I) + \left(1-p^1\right) p^2 G^1(X,I) \\ &+ p^1\left(1-p^2\right) G^1(I,X) + \left(1-p^1\right)\left(1-p^2\right) G^1(X,X) \\ &= p^1 p^2 - \left(1-p^1\right) p^2 - p^1\left(1-p^2\right) + \left(1-p^1\right)\left(1-p^2\right) = -E^2. \end{aligned} \tag{14.10}$$

现在我们运用概率论的 Dirac 抽象矢量记号语言来把混合策略的经典博弈改写为密度矩阵形式。原则上，这里，就像把概率论改成密度矩阵形式，没有任何新的东西。这样做，仅仅是为了将来介绍量子客体上的博弈的方便——就好像从经典概率论的密度矩阵过渡到量子态的密度矩阵一样。

定义 14.4 密度矩阵形式的经典博弈的抽象定义 $\Gamma^{ab}\left(\left\{\rho^i\right\}, \left\{G^i\right\}\right)$：经典博弈者 i 从 M 个元素的可选操作集合 O 中选择策略的密度矩阵是对角的矩阵 ρ^i，其中对角元素 $\rho^i_{j_i j_i} = p^i_{j_i}$ 代表博弈者 i 选择策略 $j_i \in O$ 的概率。所有博弈者的策略密度矩阵可以用 ρ 表示，

$$\rho = \Pi_{i=1}^N \otimes \rho^i. \tag{14.11}$$

博弈者 i 的收益由以下公式给出，

$$E^i = tr\left(\rho H^i\right). \tag{14.12}$$

其中对角矩阵 H^i 叫做支付矩阵，其非零元素

$$H^i\left(U^{(1)}_{j_1}, U^{(2)}_{j_2}, \cdots, U^{(N)}_{j_N}; U^{(1)}_{j_1}, U^{(2)}_{j_2}, \cdots, U^{(N)}_{j_N}\right) = G^i\left(U^{(1)}_{j_1}, U^{(2)}_{j_2}, \cdots, U^{(N)}_{j_N}\right). \tag{14.13}$$

其中 $j_k = 1, 2, \cdots, M$。

可以称明，公式 (14.12) 给出的结果和公式 (14.9) 的结果完全一样。例如，见下面的例子。

例 14.4 (**翻硬币游戏的密度矩阵形式的抽象定义**)：Alice 和 Bob 的操作集合 $\{I, X\}$。假设 Alice（Bob）的混合策略是以 p^1（p^2）的几率选择 I，那么，Alice 和 Bob 的策略状态的密度矩阵是，

$$\rho^{1,2} = \begin{bmatrix} p^1, p^2 & 0,0 \\ 0,0 & 1-p^1, 1-p^2 \end{bmatrix}. \tag{14.14}$$

于是整体系统的策略状态是

$$\rho=\rho^1\otimes\rho^2=\begin{bmatrix} p^1p^2 & 0 & 0 & 0 \\ 0 & p^1\left(1-p^2\right) & 0 & 0 \\ 0 & 0 & \left(1-p^1\right)p^2 & 0 \\ 0 & 0 & 0 & \left(1-p^1\right)\left(1-p^2\right) \end{bmatrix}. \quad (14.15)$$

整体系统的抽象支付矩阵可以写成，

$$H^{1,2}=\begin{bmatrix} 1,-1 & 0 & 0 & 0 \\ 0 & -1,1 & 0 & 0 \\ 0 & 0 & -1,1 & 0 \\ 0 & 0 & 0 & 1,-1 \end{bmatrix}. \quad (14.16)$$

其收益由按照 $E^i=tr\left(H^{1,2}\rho\right)$ 计算得到

$$E^1=p^1p^2-\left(1-p^1\right)p^2-p^1\left(1-p^2\right)+\left(1-p^1\right)\left(1-p^2\right)=-E^2. \quad (14.17)$$

运用密度矩阵形式的语言 Nash 均衡可以重新写成下面的形式: 给定收益矩阵 $\left\{H^{(1)},\cdots,H^{(N)}\right\}$，Nash 均衡就是这样的一组密度矩阵 $\rho^{i,*}$，满足

$$tr\left(\rho^{1,*}\otimes\cdots\rho^i\otimes\cdots\rho^{N,*}H^i\right)\leq tr\left(\rho^{1,*}\otimes\cdots\rho^{i,*}\otimes\cdots\rho^{N,*}H^i\right),\forall\rho^i. \quad (14.18)$$

其中左边表达式用一个一般的状态 ρ^i 代替了右边表达式中的特定状态 $\rho^{i,*}$。小于等于号（≤）表示这样的替换不会增加博弈者 i 的收益。于是，原则上，如果 i 是理性的，就不会主要去偏离这个特定的 $\rho^{i,*}$。这就是为什么这样的博弈的策略被称为均衡解，被称为 Nash 均衡。Nash 首先提出来的这样的均衡的概念并且证明了其存在性 [58, 59]。注意到更一般地关联密度矩阵，而不一定是直积密度矩阵——代表独立概率分布，我们还可以定义，这样的“均衡” 态 ρ^*，

$$tr\left(\rho^i\otimes tr^i\left(\rho^*\right)H^i\right)\leq tr\left(\rho^*H^i\right),\forall\rho^i. \quad (14.19)$$

ρ^* 包含独立均衡态也包含关联均衡态。在这里我们不展开讨论。这里的部分求迹运算 $tr^i\left(\rho^*\right)$ 是从整体分布函数中去掉关于博弈者 i 的策略的信息，然后在 $\rho^i\otimes tr^i\left(\rho^*\right)$ 中，我们利用直积算符又把这部分信息加回来，只不过这个时候加进来的分布是任意的，而且是独立于其他博弈者的。这个就是独自的偏离的含义。

小结一下，在经典博弈上，通过引入经典客体和博弈的操作性定义，我们把支付函数变成了对角的收益算符，把概率“矢量”变成了对角的密度矩阵，同时收益的计算方式也从公式 (14.2) 变成了公式 (14.12)。这些，运用我们之前介绍过的，概率论的密度矩阵的语言，都是非常自然的事情。

现在，我们有了经典博弈上，从操作性定义到密度矩阵形式的抽象定义的道路，我们现在来看量子博弈的操作性定义和密度矩阵形式的抽象定义。以下内容部分来自于文献 [60]。

14.2 量子客体上的操作和博弈

让我们从具体例子开始——一个自旋上的博弈问题。

例 14.5 (**自旋翻转游戏的操作性定义**)：有一个自旋，其初始状态是 ρ_0^q。Alice 和 Bob 都可以通过量子仪器来操作这个自旋。为了简单起见，我们规定他们所有的可能的操作的集合是所有的 2×2 的幺正演化算符，也就是所有的自旋算符，它们构成 $SU(2)$ 群。当 Alice 和 Bob 选择完了他们的操作 $U^{(1)}$ 和 $U^{(2)}$ 以后，我们先构成一个整体算符 $L = U^{(2)}U^{(1)}$，然后这个算符作用到自旋的初态上去得到末态，$\rho_f^q = L\rho_0^q L^\dagger$。最后，根据这个末态，我们计算 Alice 和 Bob 的收益，$E^{1,2} = tr\left(P^{1,2}\rho_f^q\right)$。

对照例14.1，我们发现，除了用 $SU(2)$ 代替了集合 $\{I, X\}$，也就是群 Z_2 之外，这个经典和量子的操作性定义，完全一样。于是，一个自然的猜测就是在将来的抽象定义中，两者也没有任何不同，仅仅是集合扩大了，而且是从离散群变成了连续群。如果确实这个策略集合扩大，从离散策略变成连续策略，那么，这个世界上就不存在理论上的“量子博弈”，因为，经典博弈也可以定义在连续策略上，例如考虑买卖产品的价格的博弈。我们以后还会回到这个问题。我们会发现，问题不是这么简单。我们很快会看见其原因：$SU(2)$ 是一个李群，其存在元素之间的加法，而不仅仅是群乘；在离散群 Z_2 中，我们仅仅有群乘操作，没有元素之间的加法。从量子力学的学习，我们以讲看见，有加法的集合（量子态的集合）和没有加法的集合（经典态的集合）有什么不一样。我们即将看见同样的这个不同在博弈问题上的表现：头加法运算的量子策略集和没有加法运算的经典策略集会导致整个理论框架的不同。

按照上面这个例子，我们写下来一般的量子客体上的博弈的操作性定义。

定义 14.5 量子博弈的操作性定义 $\Gamma^{op}\left(\rho_0^q, \mathcal{U}, \left\{P^i\right\}\right)$: 有一个量子客体，其初始状态是 ρ_0^q，其上的幺正算符空间记为 $\mathcal{U}$。博弈者 i 可以选择任意的 $U^{(i)} \in \mathcal{U}$，从而构成一个整体算符 $L = U^{(j_1)} \cdots U^{(j_N)}$。其中 $(j_1, \cdots, j_N)$ 是 $(1, \cdots, N)$ 的一个全排列。然后这个算符 L 作用到自旋的初态上去得到末态，$\rho_f^q = L\rho_0^q L^\dagger$。最后，根据这个末态，我们决定博弈者 i 的收益，$E^i = tr\left(P^i \rho_f^q\right)$。

现在，我们来考虑抽象定义。

定义 14.6 密度矩阵形式的量子博弈的抽象定义 $\Gamma^{ab}\left(\left\{\rho^i\right\}, \left\{H^i\right\}\right)$: 博弈者 i 可以选择 $\mathcal{U}$ 上的一个密度矩阵 ρ^i 来表示其策略，从而构成一个所有博弈者状态的密度矩阵 $\rho = \Pi_{i=1}^N \otimes \rho^i$。然后，根据这个密度矩阵，博弈者 i 的收益由下面的公式确定，

$$E^i = tr\left(\rho H^i\right). \tag{14.20}$$

为了保证这个抽象定义得到的收益和上面的操作性定义一致，我们要求 H^i 的元素满足，对于某一组给定的 $\mathcal{U}$ 的基矢 $\left\{\left|U_j\right\rangle\right\}$

$$\left\langle U_{j_1}^{(1)}, \cdots, U_{j_N}^{(N)} \middle| H^i \middle| V_{j_1}^{(1)}, \cdots, V_{j_N}^{(N)} \right\rangle = tr\left(P^i L_U \rho_0^q L_V^\dagger\right). \tag{14.21}$$

其中 L_U 是 $L\left(U_{j_1}^{(1)} \cdots U_{j_N}^{(N)}\right)$ 的简化记号。注意到公式 (14.13) 和公式 (14.21) 的不同就是前者只有对角项也就是左右都是 U 而后者具有左边是 U 右边是 V 的非对角项。

这里，算符空间 $\mathcal{U}$ 的基矢 $\left\{\left|U_j\right\rangle\right\}$ 是指把 $\mathcal{U}$ 看作一个矢量空间（加法数乘很容易验证满足矢量空间的定义），并定义算符矢量——有的时候也被称作超矢量——的内积为 $\langle A|\, B\rangle = tr\left(A^\dagger B\right)$。我们很快会看到，这个时候，一般情况下，无论是 ρ 还是 H^i 都会存在非零的非对角元，而且，当两者都在一组自然基矢下变成对角的时候，量子博弈自然过渡到经典博弈。

下面，我们来证明两者给出的收益，无论在纯策略还是混合策略的情形都是一致的。

首先，我们看到 H^i 在定义所用的基矢下面的对角元是对的——与操作性定义给出的收益一致，也就是对于给定的一组 $U_{j_1}^{(1)}, \cdots, U_{j_N}^{(N)}$，这个时候

$$\rho = \left|U_{j_1}^{(1)}, \cdots, U_{j_N}^{(N)}\right\rangle \left\langle U_{j_1}^{(1)}, \cdots, U_{j_N}^{(N)}\right|, \tag{14.22}$$

于是，

$$E^i = tr\left(\rho H^i\right) = \left\langle U_{j_1}^{(1)}, \cdots, U_{j_N}^{(N)} \middle| H \middle| U_{j_1}^{(1)}, \cdots, U_{j_N}^{(N)} \right\rangle = tr\left(P^i L_U \rho_0^q L_U^\dagger\right) \tag{14.23}$$

接着，我们来证明对于所有的纯策略 H^i 给出的收益与操作性定义一致。为简单起见，我们假设仅仅第一个博弈者的策略是给定纯策略基矢的叠加，其它博弈者还是采用给定的纯策略基矢，也就是，

$$|S\rangle = \left(\sum_{j_1} s_{j_1}^1 \left|U_{j_1}^{(1)}\right\rangle\right) \otimes \cdots \otimes \left|U_{j_N}^{(N)}\right\rangle. \tag{14.24}$$

更一般的纯策略情况仅仅是这个特殊情况的简单扩展。于是，相应的密度矩阵是，

$$\rho = |S\rangle\langle S| = \left(\sum_{j_1,k_1} s_{j_1}^1 s_{k_1}^{1,*} \left|U_{j_1}^{(1)}\right\rangle\left\langle U_{k_1}^{(1)}\right|\right) \otimes \cdots \otimes \left|U_{j_N}^{(N)}\right\rangle\left\langle U_{j_N}^{(N)}\right|. \tag{14.25}$$

代入到 $E^i = tr\,(\rho H^i)$ 我们有，

$$\begin{aligned} E^i = \langle S|\,H\,|S\rangle &= \sum_{j_1,k_1} s_{j_1}^1 s_{k_1}^{1,*} \left\langle U_{j_1}^{(1)} \cdots U_{j_N}^{(N)}\right| H^i \left|U_{k_1}^{(1)} \cdots U_{j_N}^{(N)}\right\rangle \\ &= \sum_{j_1,k_1} s_{j_1}^1 s_{k_1}^{1,*} tr\left(P^i L\left(U_{j_1}^{(1)} \cdots U_{j_N}^{(N)}\right) \rho_0^q L\left(U_{k_1}^{(1)} \cdots U_{j_N}^{(N)}\right)\right) \\ &= tr\left(P^i L\,(S)\,\rho_0^q L\,(S)\right). \end{aligned} \tag{14.26}$$

其中，最后一步，我们用到了 L 是一个线性映射这个性质，得到了操作性定义的收益计算形式。

最后一步，我们来证明，对于混合策略，$E^i = tr\,(\rho H^i)$ 给出的结果与操作性定义还是一致的。为简单起见，我们假设这个混合策略仅仅是第一个博弈者在采用，而且所用的混合策略就是已经选定的纯策略基矢下的混合策略，也就是

$$\rho = \sum_{j_1} p_{j_1}^1 \left|U_{j_1}^{(1)}\right\rangle \otimes \cdots \otimes \left|U_{j_N}^{(N)}\right\rangle\left\langle U_{j_1}^{(1)}\right| \otimes \cdots \otimes \left\langle U_{j_N}^{(N)}\right|. \tag{14.27}$$

按照操作性定义，对于混合策略，我们有

$$E^i = \sum_{j_1} p_{j_1}^1 \left(P^i L\left(U_{j_1}^{(1)} \cdots U_{j_N}^{(N)}\right) \rho_0^q L^\dagger \left(U_{j_1}^{(1)} \cdots U_{j_N}^{(N)}\right)\right). \tag{14.28}$$

这个正好是 $E^i = tr\,(\rho H^i)$ 给出的结果。

到此，我们证明了量子博弈的操作性定义和密度矩阵形式的抽象定义给出的收益是完全一致的。下一步的问题是讨论对于这样的博弈，均衡解是

否存在，以及如何计算，是否能够用于解释实验的问题。在此，我们就不再讨论了。为了更加直观地体现经典和量子博弈的异同，我们来看一个具体的例子。

例 14.6 (**两人翻自旋博弈的收益矩阵**)：Alice 和 Bob 同时控制一个自旋，其纯策略集合是 $SU(2)$。Alice 和 Bob 还可以采用这个集合上的混合策略。以 $I, \sigma_x, \sigma_y, \sigma_z$ 为基矢，也就是

$$u^{(i)} = a_I^{(i)} I + i a_x^{(i)} \sigma_x + i a_y^{(i)} \sigma_y + i a_z^{(i)} \sigma_z, \tag{14.29}$$

或者记为抽象纯策略矢量

$$\left|u^{(i)}\right) = a_I^{(i)} \left|I\right) + i a_x^{(i)} \left|\sigma_x\right) + i a_y^{(i)} \left|\sigma_y\right) + i a_z^{(i)} \left|\sigma_z\right), \tag{14.30}$$

或者记为

$$\left|u^{(i)}\right) = \left[a_I^{(i)}, i a_x^{(i)}, i a_y^{(i)}, i a_z^{(i)}\right]^T. \tag{14.31}$$

其混合策略如下：

$$\rho^{q,S} = \begin{bmatrix} \rho_{ii}^{q,1} & \rho_{ix}^{q,1} & \rho_{iy}^{q,1} & \rho_{iz}^{q,1} \\ \rho_{xi}^{q,1} & \rho_{xx}^{q,1} & \rho_{xy}^{q,1} & \rho_{xz}^{q,1} \\ \rho_{yi}^{q,1} & \rho_{yx}^{q,1} & \rho_{yy}^{q,1} & \rho_{yz}^{q,1} \\ \rho_{zi}^{q,1} & \rho_{zx}^{q,1} & \rho_{zy}^{q,1} & \rho_{zz}^{q,1} \end{bmatrix} \otimes \begin{bmatrix} \rho_{ii}^{q,2} & \rho_{ix}^{q,2} & \rho_{iy}^{q,2} & \rho_{iz}^{q,2} \\ \rho_{xi}^{q,2} & \rho_{xx}^{q,2} & \rho_{xy}^{q,2} & \rho_{xz}^{q,2} \\ \rho_{yi}^{q,2} & \rho_{yx}^{q,2} & \rho_{yy}^{q,2} & \rho_{yz}^{q,2} \\ \rho_{zi}^{q,2} & \rho_{zx}^{q,2} & \rho_{zy}^{q,2} & \rho_{zz}^{q,2} \end{bmatrix}. \tag{14.32}$$

纯策略写成混合策略的方法和量子力学完全一致，也就是，

$$\rho = \left|u\right) \left(u\right|. \tag{14.33}$$

其中，由于 u 是一个算符或者说超矢量而不是量子态矢量，我们用圆括号的 Dirac 符号代替了尖括号的 Dirac 符号。于是，当这个混合策略就是表示前面的纯策略的时候，

$$\rho^{q,i} = \begin{bmatrix} a_I^{(i)} \\ i a_x^{(i)} \\ i a_y^{(i)} \\ i a_z^{(i)} \end{bmatrix} \left[a_I^{(i)}, i a_x^{(i)}, i a_y^{(i)}, i a_z^{(i)}\right]. \tag{14.34}$$

其收益矩阵如下：

$$
H^{(1)} = \begin{bmatrix}
1 & & & 1 & & 1 & -i & & & i & 1 & & 1 & & & 1 \\
 & -1 & i & & -1 & & & -1 & i & & & i & & 1 & -i & \\
 & -i & -1 & & -i & & & -i & -1 & & & -1 & & i & 1 & \\
1 & & & 1 & & 1 & -i & & & i & 1 & & 1 & & & 1 \\
 & -1 & i & & -1 & & & -1 & i & & & i & & 1 & -i & \\
1 & & & 1 & & 1 & -i & & & i & 1 & & 1 & & & 1 \\
i & & & -i & & -i & 1 & & & -1 & -i & & -i & & & -i \\
 & -1 & i & & -1 & & & -1 & i & & & i & & 1 & -i & \\
 & -i & -1 & & -i & & & -i & -1 & & & -1 & & i & 1 & \\
-i & & & -i & & -i & -1 & & & 1 & -i & & -i & & & -i \\
1 & & & 1 & & 1 & -i & & & i & 1 & & 1 & & & 1 \\
 & -i & -1 & & -i & & & -i & -1 & & & -1 & & i & 1 & \\
1 & & & 1 & & 1 & -i & & & i & 1 & & 1 & & & 1 \\
 & 1 & -i & & 1 & & & 1 & -i & & & -i & & -1 & i & \\
 & i & 1 & & i & & & i & 1 & & & 1 & & -i & -1 & \\
1 & & & 1 & & 1 & -i & & & i & 1 & & 1 & & & 1
\end{bmatrix},
$$

并且 $H^{(2)} = -H^{(1)}$。如果要计算给定策略组合的收益，则

$$
E^{(i)} = tr\left(H^{(i)}\rho^{q,S}\right). \tag{14.35}
$$

我们看到了到处都有非对角元。

我们看到，从抽象定义的数学框架上，经典博弈是对角的密度矩阵和对角的收益矩阵，而量子博弈是有非对角元的密度矩阵和有非对角元的收益矩阵。我们不仅仅用一个可能非对角的矩阵来表示策略还用一个可能非对角的矩阵来表示收益。这个是经典博弈所没有的，而且，这个差别，正好就是经典概率论到量子力学的区别。因此，量子博弈和经典博弈的关系，正好就是，量子力学之于经典力学的关系。我们还应该看到，经典博弈里面所谓的操作，基本上可以用抽象的策略来代替，如何实现这些操作基本上不是问题。然而，在量子博弈中，这些操作是通过对量子客体的实验操作来实现的。例如，在自旋上，我们是通过控制自旋通过的磁场的方向和通过的时间来实现的。这样的操作，属于幺正演化。我们还可以考虑更一般的操作，例如对给定的自旋做一个测量，甚至广义测量，这个时候，我们的算符的集合

可以是非幺正的，因此超过了 $\mathcal{U}$。这个的更大的算符集合对理论上提出的挑战是什么是一个有意思的问题，例如是否这样的策略还是能够由策略的密度矩阵 ρ 所描述。在此，我们就不再讨论了。

14.3 经典和量子博弈的异同

在量子博弈刚提出来的时候，有一个有意思的例子展示了运用量子策略的人可以在博弈中获得优势。这个博弈的例子大概是这样的 [61]：有一个 z 方向向上的自旋，博弈者 Alice 可以先对这个自旋做一个操作，然后博弈者 Bob 再对这个自旋做一个操作，接着 Alice 还有一次操作这个自旋的机会。三次操作完成以后，按照这个自旋的状态来决定两个博弈者的收益。例如，如果最后还是 z 方向向上，则 Alice 赢。在这个具体的例子中，如果 Alice 和 Bob 都不懂量子力学，那么，他们就只能够把自旋看作硬币，在翻转 (X) 和不翻转 (I) 两个操作中选择。于是，我们会发现，Alice 和 Bob 没有什么好方法保证赢的次数更多。但是，如果 Alice 懂得量子力学，而 Bob 不懂，那么，Alice 可以通过运用可逆算符 $S = \frac{I+iX}{\sqrt{2}}$ 的方式来让 Bob 的选择不发挥作用——经过这个算符 S 作用之后，Bob 无论选择翻转还是不翻转都不改变自旋的状态——然后再一次使用 $S^\dagger$ 算符来保证获胜。

这个博弈的例子在推动量子博弈的发展上发挥了重大的作用：你看，懂得量子力学的博弈者确实可以赢得更多。但是，我们随后就发现，这个是因为一件非常自然的事情：量子博弈者的策略选择空间变大了，于是，自然有可能，可以赢得更多。也就是说，量子博弈的创新之处不在博弈上，而在量子上。整个博弈的数学描述没有任何变化，仅仅是策略空间变大了。

由于这个观察，博弈的研究者们开始大力批评量子博弈的研究——你们只不过使得策略空间变大了，变成连续的策略了。但是，在经典博弈中，我们也可以采用连续策略呀，例如价格的选择，于是，量子博弈仅仅是量子系统上的经典博弈，而不是真正的数学结构不同的量子博弈——就好像从经典力学的密度分布函数到达量子力学的密度矩阵这样的数学形式的不同。

后来，大量的量子博弈的研究，确实仅仅在更大的而且连续的策略空间这个方面取得了很多进展。甚至有人采用这个连续得策略空间上的概率分布函数来描述量子混合策略，并且开展实验 [62]。

一直到我们自己的研究工作指出来，其实，由于量子博弈的策略构成李群——李群是连续群并且除了乘法还有加法和微分运算，实际上，我们的

“量子混合策略”的描述需要**从更大的连续的策略空间上的对角的概率分布函数变成李群的基矢上的可以有非对角元的密度矩阵**。这样，量子博弈和经典博弈之间的关系就好像是量子力学和经典力学之间的关系。这个方向的工作，还有待于进一步展开。

14.4 作业

习题 14.1 考虑如下的博弈：如果 Alice 和 Bob 各自都拿着一个硬币，初始两个硬币都向上，然后两人选择操作——包含翻转和不翻转自己的硬币——而得到末状态。按照硬币的末状态来决定收益：如果两个硬币的末状态一样，则 Alice 获得 1 元否则输掉 1 元（反过来，如果硬币状态不一样则 Bob 赢得 1 元）。写出这个具体和博弈的抽象定义。

习题 14.2 考虑如下的博弈：如果 Alice 和 Bob 各自都拿着一个自旋，初始两个自旋都 z 方向向上，然后两人选择操作——包含转动自己的自旋——而得到末状态。按照自旋的末状态来决定收益：测量两个硬币的 z 方向，如果两个硬币的末状态一样，则 Alice 获得 1 元否则输掉 1 元（反过来，如果硬币状态不一样则 Bob 赢得 1 元）。写出这个具体和博弈的抽象定义。

习题 14.3 考虑如下的博弈：如果 Alice 和 Bob 各自都拿着一个自旋，初始两个自旋处于纠缠态 $|\Phi_-\rangle\langle\Phi_-|$，然后两人选择操作——包含转动自己的自旋——而得到末状态。按照自旋的末状态来决定收益：测量两个硬币的 z 方向，如果两个硬币的末状态一样，则 Alice 获得 1 元否则输掉 1 元（反过来，如果硬币状态不一样则 Bob 赢得 1 元）。写出这个具体和博弈的抽象定义。

14.5 本章小结

在这一章里面，我们介绍了经典博弈，然后把经典博弈看成是经典客体上的博弈。于是，自然，我们就可以问量子客体上的博弈是什么。接着，我们把这个量子客体上的博弈改造成由抽象策略和抽象策略空间的语言描述的博弈。我们发现，这个正好就是把经典博弈的概率分布函数的数学框架，变成了密度矩阵的数学框架。这个经典博弈和量子博弈之间的关系，正好就是经典力学和量子力学之间的关系。以此，我们也突破了博弈学者对量子博弈研究的批评——我们真的有数学结构上的不同，而不仅仅是策略空间变大了，连续了。

一般而言，数学结构上的突破总会带来新的结果上的突破。

结束语

整本书到此结束。本书在内容选择甚至具体内容的展开上都和大多数书不太一样。希望这个企图做到不一样的努力能够对于学习者的学习和理解有点效果（make a difference）。

内容上贯穿全书的是从各个方面阐述量子的行为和理论和经典的行为和理论的差异。思想上贯穿全书的是什么是科学（批判性思维、科学和实验的关系、科学和数学的关系、科学对简单性和统一性的追求）。学习方法上贯穿本书的是一定要做理解型学习（对所学的内容 make sense，系联性思考）并且主要依靠自学，老师的作用仅仅是提点一下整体思路整体方向感以及提点一下少数难点。在这几个方面，量子力学都是好素材：通过学习量子力学，我们有可能在内容上、思想上、学习方法上，甚至好奇心学习态度上，都有所得。希望本书多多少少完成了这个任务。

谢谢你作为学习者付出的时间和努力，希望你喜欢这个深入思考的过程。

参考文献

[1] Bell J S. Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy [M]. 2nd ed. Cambridge University Press, 1982.

[2] von Neumann J. Mathematical Foundations of Quantum Mechanics [M]. translation - from german ed. Princeton University Press, 1955.

[3] Dirac P A M. The Principles of Quantum Mechanics [M]. 4th ed. Clarendon Press, 1982.

[4] Shankar R. Principles of Quantum Mechanics [M]. Springer, 1994.

[5] Nielsen M A, Chuang I L. Quantum Computation and Quantum Information [M]. Cambridge University Press, 2004.

[6] Preskill J. Lecture notes for physics 229: Quantum Information and Computation [M/OL]. CreateSpace Independent Publishing Platform, 2015. http://www.theory.caltech.edu/people/preskill/ph229/index.html.

[7] Feynman R P, Leighton R B, Sands M. The Feynman lectures on physics. III [M]. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1965.

[8] Ballentine L E. Quantum mechanics: A modern development [M]. World scientific, 1998.

[9] Sakurai J J. Modern Quantum Mechanics [M]. Addison Wesley, 1993.

[10] 吴金闪. 教的更少, 学得更多——概念地图学习与教学方法 [M/OL]. Unpublished, retrieved March 2016, 2016. http://www.systemsci.org/jinshanw/books/.

[11] Popper K. The logic of scientific discovery [M]. London/New York: Routledge, 2005.

[12] Feynman R P. QED: The Strange Theory of Light and Matter [M]. Princeton University Press, 1988.

[13] Moghal J, Kobler J, Sauer J, et al. High-Performance, Single-Layer Antireflective Optical Coatings Comprising Mesoporous Silica Nanoparticles [J]. ACS Applied Materials & Interfaces, 2012, 4 (2): 854–859.

[14] Tonomura A, Endo J, Matsuda T, et al. Demonstration of single-electron buildup of an interference pattern [J]. American Journal of Physics, 1989, 57 (2): 117–120.

[15] Grangier P, Roger G, Aspect A. Experimental Evidence for a Photon Anticorrelation Effect on a Beam Splitter: A New Light on Single-Photon Interferences [J/OL]. EPL (Europhysics Letters), 1986, 1 (4): 173. http://stacks.iop.org/0295-5075/1/i=4/a=004.

[16] Lukishova S G, Carlos R Stroud J, Bissell L, et al. Teaching Experiments on Photon Quantum Mechanics [C/OL]. In Frontiers in Optics 2008/Laser Science XXIV/Plasmonics and Metamaterials/Optical Fabrication and Testing, 2008: SThD3. http://www.osapublishing.org/abstract.cfm?URI=FiO-2008-SThD3.

[17] Whitehead A N. The Aims of Education and Other Essays（中译本《教育的目的》） [M]. Free Press; Reissue edition, 1967.

[18] Adler M, Van Doren C. How to Read a Book（《如何阅读一本书》）[M]. New York: Touchstone, a Divisition of Simon and Schuster, 2011.

[19] Kolmogorov A. 概率论导引 [M]. 哈尔滨工业大学出版社, 2012.

[20] Ito K. 概率论 [M]. 人民邮电出版社, 2011.

[21] Ross S M. Introduction to Probability Models [M]. 9th ed. Academic Press, 2007.

[22] Mermin N D. Is the Moon There When Nobody Looks? Reality and the Quantum Theory [J/OL]. Physics Today, 1985, 38 (4): 38–47. http://scitation.aip.org/content/aip/magazine/physicstoday/article/38/4/10.1063/1.880968.

[23] Beveridge W I B. The art of scientific investigation [M]. Melbourne: Heinemann, 1953.

[24] Gowers T. Mathematics: A Very Short Introduction [M]. Oxford: Oxford University Press, 2002.

[25] Feynman R P. The character of physical law [M]. London: British Broadcasting Corporation, 1965.

[26] Einstein A, Infeld L. The evolution of physics: the growth of ideas from early concepts to relativity and quanta [M]. Simon and Schuster, 1938.

[27] Landau L D, Lifshitz E M. Mechanics, Third Edition: Volume 1 (Course of Theoretical Physics) [M]. Butterworth-Heinemann, 1976.

[28] 喀兴林. 高等量子力学 [M]. 北京: 高等教育出版社, 2001.

[29] 吴兆颜. 高等量子力学 [M]. 吉林: 吉林大学出版社, 2008.

[30] Axler S. Linear algebra done right [M]. Cham: Springer, 2015.

[31] 吴金闪. 系统科学导引 [M/OL]. Unpublished, retrieved Jan. 2017, 2017. http://www.systemsci.org/jinshanw/books/.

[32] 王梓坤. 概率论基础及其应用 [M]. 北京师范大学出版社, 2007.

[33] 何书元. 概率引论 [M]. 高等教育出版社, 2011.

[34] Chung K L. A course in probability theory [M]. Boston: Academic Press, 1974.

[35] Feynman R P, Leighton R B, Sands M. The Feynman lectures on physics. I [M]. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1965.

[36] 方励之, 李淑娴. 力学概论 [M]. 安徽科学技术出版社, 1987.

[37] 漆安慎, 杜婵英. 力学 [M]. 高等教育出版社, 2005.

[38] Dirac P. Generalized Hamiltonian Dynamics [J/OL]. Can. J. Math., 1950, 2: 129–148. http://dx.doi.org/10.4153/CJM-1950-012-1.

[39] Feynman R P, Hibbs A R. Quantum mechanics and path integrals [M]. New York: McGraw-Hill, 1965.

[40] Schneider M B, LaPuma I A. A simple experiment for discussion of quantum interference and which-way measurement [J/OL]. American Journal of Physics, 2002, 70 (3): 266–271. http://scitation.aip.org/content/aapt/journal/ajp/70/3/10.1119/1.1450558.

[41] Frabboni S, Gazzadi G C, Pozzi G. Ion and electron beam nanofabrication of the which-way double-slit experiment in a transmission electron microscope [J/OL]. Applied Physics Letters, 2010, 97 (26): 263101. http://dx.doi.org/10.1063/1.3529947.

[42] Hellmuth T, Walther H, Zajonc A, et al. Delayed-choice experiments in quantum interference [J/OL]. Phys. Rev. A, 1987, 35: 2532–2541. http://link.aps.org/doi/10.1103/PhysRevA.35.2532.

[43] Jacques V, Wu E, Grosshans F, et al. Experimental Realization of Wheeler's Delayed-Choice Gedanken Experiment [J/OL]. Science, 2007, 315 (5814): 966–968. http://science.sciencemag.org/content/315/5814/966.

[44] 曾谨言. 量子力学教程 [M]. 科学出版社, 2014.

[45] Peng C-Z, Yang T, Bao X-H, et al. Experimental Free-Space Distribution of Entangled Photon Pairs Over 13 km: Towards Satellite-Based Global Quantum Communication [J/OL]. Phys. Rev. Lett., 2005, 94: 150501. http://link.aps.org/doi/10.1103/PhysRevLett.94.150501.

[46] Zhao Z, Chen Y-A, Zhang A-N, et al. Experimental demonstration of five-photon entanglement and open-destination teleportation [J/OL]. Nature, 2004, 430: 54–58. http://dx.doi.org/10.1038/nature02643.

[47] Zeilinger A, Horne M A, Weinfurter H, et al. Three-Particle Entanglements from Two Entangled Pairs [J/OL]. Phys. Rev. Lett., 1997, 78: 3031–3034. http://link.aps.org/doi/10.1103/PhysRevLett.78.3031.

[48] Levin M, Wen X-G. Colloquium : Photons and electrons as emergent phenomena [J/OL]. Rev. Mod. Phys., 2005, 77: 871–879. http://link.aps.org/doi/10.1103/RevModPhys.77.871.

[49] Aspect A, Grangier P, Roger G. Experimental Tests of Realistic Local Theories via Bell's Theorem [J/OL]. Phys. Rev. Lett., 1981, 47: 460–463. http://link.aps.org/doi/10.1103/PhysRevLett.47.460.

[50] Aspect A, Grangier P, Roger G. Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment : A New Violation of Bell's Inequalities [J/OL]. Phys. Rev. Lett., 1982, 49: 91–94. http://link.aps.org/doi/10.1103/PhysRevLett.49.91.

[51] Aspect A, Dalibard J, Roger G. Experimental Test of Bell's Inequalities Using Time- Varying Analyzers [J/OL]. Phys. Rev. Lett., 1982, 49: 1804–1807. http://link.aps.org/doi/10.1103/PhysRevLett.49.1804.

[52] Cirel'son B S. Quantum generalizations of Bell's inequality [J/OL]. Letters in Mathematical Physics, 1980, 4 (2): 93–100. http://dx.doi.org/10.1007/BF00417500.

[53] Tittel W, Brendel J, Zbinden H, et al. Violation of Bell Inequalities by Photons More Than 10 km Apart [J/OL]. Phys. Rev. Lett., 1998, 81: 3563–3566. http://link.aps.org/doi/10.1103/PhysRevLett.81.3563.

[54] Weihs G, Jennewein T, Simon C, et al. Violation of Bell's Inequality under Strict Einstein Locality Conditions [J/OL]. Phys. Rev. Lett., 1998, 81: 5039–5043. http://link.aps.org/doi/10.1103/PhysRevLett.81.5039.

[55] Hensen B, Bernien H, Dreau A E, et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres [J/OL]. Nature, 2015, 526: 682–686. http://dx.doi.org/10.1038/nature15759.

[56] Wu J, Pei S. Could a Classical Probability Theory Describe Quantum Systems? [J/OL]. arXiv:quant-ph/0503093, 2005. https://arxiv.org/abs/quant-ph/0503093.

[57] Barnum H, Caves C M, Fuchs C A, et al. Noncommuting Mixed States Cannot Be Broadcast [J/OL]. Phys. Rev. Lett., 1996, 76: 2818–2821. http://link.aps.org/doi/10.1103/PhysRevLett.76.2818.

[58] Nash J F J. Equilibrium points in n-person games [J/OL]. Proceedings of the National Academy of Sciences of the United States of America, 1950, 36 (1): 48–49. http://www.pnas.org/content/36/1/48.full.

[59] Dixit A K, Skeath S, Reiley D H J. Games of Strategy [M]. W. W. Norton & Company, 2015.

[60] Wu J. Hamiltotian Formalism of Game Theory [J/OL]. arXiv:quant-ph/0501088, 2005. https://arxiv.org/abs/quant-ph/0501088.

[61] Meyer D A. Quantum Strategies [J/OL]. Phys. Rev. Lett., 1999, 82: 1052–1055. http://link.aps.org/doi/10.1103/PhysRevLett.82.1052.

[62] Du J, Li H, Xu X, et al. Experimental Realization of Quantum Games on a Quantum Computer [J/OL]. Phys. Rev. Lett., 2002, 88: 137902. http://link.aps.org/doi/10.1103/PhysRevLett.88.137902.

名词索引

B | D | E | F | G | H | J | K | L | N | P | S | T | W | Y | Z

B

Bell **不等式** Bell's Inequality. 64, 74, 222–226, 231, 232
CHSH **不等式** CHSH Inequality. 224, 225

D

DelftU Delft University of Technology（Delft 科学技术大学）. 225
Dirac δ **函数** Dirac δ function. 57, 188
Dirac **符号** Dirac 发明的用于表示矢量、对偶矢量和算符的符号. 84

E

Einstein **定域性** Einstein Locality. 221, 228–230

F

反射 光在两种介质的界面处发生方向的改变，回到与入射光相同的介质中的行为。但是，在本书中，我们不太严格地使用“透射”和“反射”这一组词：有的时候，我们把从另一个方向出来的光叫做“反射光”，同时从这个仪器一个方向出来的光叫做“透射光”，而不去管背后的物理过程到

底是否就是狭义上的反射和透射。用这样的不太严格的术语，我们统一来描述偏振片、半反半透镜、偏振分束器等光学仪器. 25, 36

G

概率叠加原理 互斥事件的和的概率是各自概率的和. 31, 114
概念地图 Concept Mapping/Concept Map. 45
干涉 合起来的波的强度不等于这个波的每一部分的强度之和的现象. 22, 24
光电效益 Photoelectric effect. 40
光子分束器 把光子分开成不同出射口出射的仪器，通常通过反射和透射来实现. 35

H

Hadamard **变换** 有的时候指 Hadamard 变换，有的时候指实现 Hadamard 变换的量子逻辑门. 183, 212
黑体辐射 Black-body radiation. 40

J

简单性 Simplicity. 49
介质波 介质点的振动形成的波，传播的是波形而不是介质点. 23
经典力学 Classical Mechanics. 21, 30
经典波动力学 以把 Newton 定律用于连续介质得到的波动方程为基础的力学. 21, 22, 27, 43
经典概率论 经典粒子的力学可以写成 Newton 方程的确定性形式，也可以运用经典概率论写成相空间中的随机分布函数的 Liouville 方程。因此，在这里我们把经典概率论描述的力学看作经典力学的子概念. 105

经典粒子理论 把 Newton 定律用于质点的力学，可以表现为确定的形式或者随机分布函数的形式. 26, 43
Newton **第二定律** $\vec{F} = m\vec{a}$. 124–127
Newton **力学** 经典力学的 Newton 方程形式. 21, 124, 129

K

可证伪性 指一个理论或者猜想，其本身能够通过实验来检验是否正确. 49
Kronecker δ **记号** Kronecker δ notation. 57, 87, 101

L

Legendre **变换** 自变量、函数值、函数形式都变化，维持运动方程不变. 132

N

Nash **定理** Nash Theorem. 255
Nash **均衡** Nash Equilibrium. 255

P

偏振 Polarization. 25
偏振分束器 按照光子偏振方向的不同来把光子按照不同的出射口分开的仪器. 35
Planck **常数** Planck Constant. 40, 154

S

三道门实验 Three-Gate experiment. 27
Dirac **的光过三块偏振片** pass through three polarized beam splitter. 25
绳子上的波过三个缝 ing wave pass through three gates. 28
双缝实验 Double-slit experiment. 21

水波的双缝实验 Water wave double-slit experiment. 21
子弹的双缝实验 Bullet double-slit experiment. 21, 30
Stern-Gerlach **实验** Stern-Gerlach experiment. 38
Stern-Gerlach **装置** Stern-Gerlach apparatuses. 38, 39

T

透射 通常指光从一种介质穿另外一种介质又回到第一种介质的行为——这种情况背后的物理过程可能是两次折射。但是，在本书中，我们不太严格地使用“透射”和“反射”这一组词：有的时候，我们把光穿过一个仪器的过程都叫做透射；有的时候，我们把从这个仪器一个方向出来的光叫做“透射光”，同时从另一个方向出来的光叫做“反射光”，而不去管背后的物理过程到底是否就是狭义上的反射和透射。用这样的不太严格的术语，我们统一来描述偏振片、半反半透镜、偏振分束器等光学仪器. 25, 36

W

WHWM 理解型学习、阅读和写作过程中要问的四个重要问题：表达什么（What），怎么表达（How），为什么表达、这个为什么这样表达（Why），对读者有什么意义（Meaningful）. 14, 45
物质波 不通过介质传播而是物质自身在传播的波. 23

Y

隐变量理论 Hidden Variable Theory. 60, 63, 220

Z

矢量叠加原理 量子状态的和满足矢量加法，因此违反概率叠加. 159

人名与常用翻译

A | B | C | D | E | F | G | H | I | K | L | M | N | P | Q | S | V | W | Y | Z

A

Adler	Mortimer J. Adler （艾德勒）. 44
Affleck	Ian Affleck（阿夫雷克）. 11
Aspect	Alain Aspect（阿斯佩）. 225
Axler	Sheldon Axler（阿克斯勒）. 79

B

Ballentine	Leslie Ballentine（巴楞廷（暂译））. 11, 13, 14, 19, 42, 139, 176
Bell	John Bell（贝尔）. 13, 19, 223
Beveridge	William Ian Beardmore Beveridge（贝弗里奇）. 79
Boltzmann	Ludwig Boltzmann（玻尔兹曼）. 40

C

Chung	Kai Lai Chung（钟开莱）. 106
Coleman	Sidney Coleman（科尔曼）. 45, 182

D

Dirac	Paul A.M. Dirac（狄拉克）. 13
狄增如	北京师范大学系统科学学院教师. 40

E

Einstein	Albert Einstein（爱因斯坦）. 40, 60, 63, 79

F

方励之	美国亚利桑那大学物理系教师. 121
Faraday	Michael Faraday（法拉第）. 49
Feynman	Richard Feynman（费曼）. 11, 13, 14, 19, 79, 121, 139

G

Galileo	Galileo Galilei（伽利略）. 122
Gerlach	Walter Gerlach（盖拉赫）. 39, 167
Gisin	Nicolas Gisin（基辛）. 224
Gowers	Timothy Gowers（高尔）. 79

H

Hadamard	Jacques Hadamard（阿达马）. 183
Hanson	Ronald Hanson（汉森）. 225, 231

I

Infeld	Leopold Infeld（英菲尔德）. 79
Ito	Kiyoshi Ito（伊藤清）. 54

K

喀兴林	北京师范大学物理系教师. 11, 77, 79, 139, 176
Kolmogorov	Andrey Nikolaevich Kolmogorov（柯尔莫戈洛夫）. 54, 77, 106

L

Landau	Lev Davidovich Landau（朗道）. 77, 121
Legendre	Adrien-Marie Legendre（勒让德）. 132

M

Maxwell	James Maxwell（麦克斯韦）. 49
Mermin	David Mermin（墨敏（暂译））. 64

N

Nash	John Nash（纳什）. 255
Nielsen	Michael Nielsen（尼尔森）. 13

P

裴寿镛	北京师范大学物理系教师. 11, 14, 40
Planck	Max Planck（普朗克）. 40
Popper	Carl Popper（波普尔）. 49
Preskill	John Preskill（普利斯吉尔（暂译））. 13

Q

漆安慎	北京师范大学物理系教师. 121

S

Sakurai	Jun Sakurai（樱井纯）. 13
Shankar	Ramamurti Shankar（尚卡尔）. 13
Stern	Otto Stern（斯特恩）. 39, 167
Susskind	Leonard Susskind（萨斯金（暂译））. 45, 182

V

van Doren	Charles van Doren（范多伦）. 44
von Nuemann	John von Nuemann（冯 · 诺伊曼）. 13

W

王梓坤	北京师范大学数学科学院教师. 106
Whitehead	Alfred Whitehead（怀特海）. 44
吴金闪	北京师范大学系统科学学院教师. 44
吴兆颜	吉林大学物理学院教师. 77

Y

杨展如	北京师范大学物理系教师. 11
Young	Thomas Young（杨）. 22

Z

Zeilinger	Anton Zeilinger（蔡林格）. 224
曾谨言	北京大学物理学院教师. 179

插图

1 本书结构图 16

1.1 Young 亲手绘制的水波干涉图 22
1.2 经典波的叠加 23
1.3 Feynman 的光过玻璃示意图 24
1.4 Dirac 的光三块偏振片 26
1.5 便宜的偏振片 27
1.6 绳子上的波过三个缝 29
1.7 子弹过小孔理想实验示意图 31
1.8 单电子过双缝实验装置示意图和实验结果 32
1.9 单电子过双缝实验结果示意图 33
1.10 光子 which-way 实验装置示意图 36
1.11 自旋的 Stern-Gerlach 装置 39
1.12 自旋过 Stern-Gerlach 装置示意图 39
1.13 电子过两个不同方向磁场 41
1.14 电子过三个方向磁场 42
1.15 电子 which-way 实验装置示意图 43
1.16 量子系统的行为一章的概念地图 47
1.17 双缝干涉的概念地图 48

4.1 不太守时的约会 108

5.1 单摆的受力分析 129

6.1 自旋过多个 Stern-Gerlach 装置示意图 147
6.2 光子的 which-way 实验装置示意图 154

6.3 加了探测器的光子的 which-way 实验装置示意图 160
6.4 先分开再合起来的自旋 SG 装置实验 171
6.5 Aspect 单光子分束器实验 171
6.6 Aspect 单光子分束器实验结果 172

10.1 4 光子和 3 光子纠缠实验装置示意图 214

11.1 Bell 不等式的直接检验装置示意图 223
11.2 Bell 不等式的动态检验装置示意图 224
11.3 没有漏洞的 Bell 不等式的检验 226

13.1 教得少学得多的基本思想 251

举例

2.1 例 (硬币翻转的动力学过程的密度矩阵描述) 56

3.1 例 (Baker–Campbell–Hausdorff 公式的一个变体) 95
3.2 例 (二维 $\sigma_x, \sigma_y, \sigma_z$ 在不同表象下的形式和抽象形式) 98
3.3 例 (一元偏微分方程的二维子空间) 98
3.4 例 (Fourier 级数的矢量空间解释) 99

4.1 例 (对称色子的几率问题) 106
4.2 例 (不太守时的约会) 107
4.3 例 (盒子里放小球的几率) 111
4.4 例 (Bertrand 圆内弦长的问题) 112
4.5 例 (硬币确定翻转的算符) 117
4.6 例 (硬币随机翻转的算符) 118

5.1 例 (自由落体运动) 124
5.2 例 (经典一维谐振子（连着弹簧的一个小球）) 125
5.3 例 (经典一维谐振子（连着弹簧的两个小球）) 126
5.4 例 (势函数举例) 128
5.5 例 (能量守恒的应用) 129
5.6 例 (单摆的周期) 129
5.7 例 (谐振子的 Hamiltonian 力学) 131
5.8 例 (单摆的 Hamiltonian 力学) 131
5.9 例 (一根弹簧连着的两个小球的 Hamiltonian 力学) 134
5.10 例 (二级摆的 Hamiltonian 力学) 134

6.1 例 (自旋算符的表象) 153

7.1 例 (自旋演化) 177
7.2 例 (神奇的偏振分束镜) 177
7.3 例 (产生湮灭算符的演化) 182

8.1 例 (一维谐振子的量子化与求解) 188
8.2 例 (一维谐振子的代数解法) 190

10.1 例 (演化导致的纠缠) 211

14.1 例 (翻硬币游戏的操作性定义) 256
14.2 例 (翻硬币游戏的纯策略抽象定义) 257
14.3 例 (翻硬币游戏的混合策略抽象定义) 258
14.4 例 (翻硬币游戏的密度矩阵形式的抽象定义) 258
14.5 例 (自旋翻转游戏的操作性定义) 260
14.6 例 (两人翻自旋博弈的收益矩阵) 263