

初识量子力学

陈童

2024 年 8 月 28 日

目录

第一章 矩阵革命	5
1.1 玻尔的新观念	5
1.1.1 线索	6
1.1.2 玻尔的新观念	7
1.2 矩阵革命	13
1.2.1 跃迁元与跃迁元乘法	14
1.2.2 一些必要的数学	22
1.2.3 海森堡运动方程	27
1.2.4 基本对易关系	29
1.2.5 求解谐振子	33
1.2.6 符号的新发展	36
第二章 量子力学基本原理	39
2.1 可确定区分可能性完备集	40
2.2 玻恩定则, 不确定性与不可区分性	46
2.3 跃迁幅乘法	51
2.4 跃迁元乘法以及物理量的测量	55
2.5 表象无关的量子力学方程	61
第三章 态和算符	66
3.1 态叠加原理	67
3.2 算符	76
3.3 海森堡绘景、薛定谔绘景以及量子力学幺正性	82
3.3.1 海森堡绘景与薛定谔绘景	82

目录	2
3.3.2 量子力学幺正性	85
3.4 算符与物理测量	89
3.4.1 本征可能性完备集	89
3.4.2 不确定原理与算符	93
3.4.3 再谈物理量的测量	96
第四章 电子自旋以及相关模型	100
4.1 电子自旋	100
4.1.1 电子自旋的发现	100
4.1.2 泡利算符与自旋算符	102
4.1.3 泡利矩阵	105
4.2 两自旋耦合系统	108
4.2.1 海森堡模型	108
4.2.2 求解与讨论	109
第五章 量子纠缠	115
5.1 两体纠缠态	115
5.2 纠缠态与混态	120
5.3 用例子说明量子纠缠态的神奇	125
5.3.1 量子纠缠能实现超光速信息传递吗?	125
5.3.2 量子密集编码	131
5.3.3 量子隐形传态	132
5.3.4 GHZ态以及为什么爱因斯坦错了	134
第六章 波动方程与路径积分	137
6.1 薛定谔波动方程	137
6.1.1 坐标表象	137
6.1.2 德布罗意波和薛定谔波动方程	142
6.1.3 概率流密度	144
6.2 路径积分	146

前言

流行的构建现代量子力学理论体系的方式是从态叠加原理和算符开始，狄拉克的经典名著《量子力学原理》开始了这一处理方式，发展到今天，结合冯诺依曼的贡献，就成了从量子力学五大公设开始建立量子力学的处理方式，虽然这些公设未免有些是从天上掉下来的。

这种处理方式的问题是，它对态的演化过程和测量过程使用了二分法，态的演化是幺正演化，而测量则会导致非幺正的量子态坍缩。我想，任何一个有足够思维能力的人都会觉得这种二分法是疯狂的！

所以，为了解决这个问题，人们想了种种办法。其中最著名的莫过于多世界的想法。它坚持唯一只有态的幺正演化，而没有态坍缩。它用巧妙的量子纠缠观念来看待测量。但是，多世界的想法要能自洽地贯彻下去的话，就得假设世界随着观测一直在不断分裂，虽然这些分裂的世界依然是相干叠加的，但是对于处于某个世界分支的副本本人来说，他可无从感知这一相干叠加的量子现实。因此，这里有无穷无尽并依然在不断分裂的副本世界。

当然，多世界理论本身也有其它的一些难题，诸如推导出玻恩定则，为了解决这些问题，后人又想了种种办法。然而我觉得，多世界的原始想法对于任何有足够现实感的人来说，已经是疯狂且流于幻想的了。

有没有一种更物理更实际的想法可以更好地理顺量子力学的理论体系并重建整个现代量子力学体系呢？在长期的学习和思索之后，我认为回答是肯定的。出路就在于回到海森堡开创量子力学的那篇一个人的文章。海森堡是基础于量子跃迁的基本物理观念而引入他对物理量的新表示以及新乘法的。这种新表示就是矩阵，这种新乘法就是矩阵乘法。但是我更愿意将海森堡的新表示叫做跃迁元，把他引入得的新乘法叫做跃迁元乘法。我

发现，从跃迁元和跃迁元乘法开始完全有可能理顺量子力学的理论体系，只不过需要把这些概念在海森堡的基础上再往前推广。

实际上，费曼在其物理学讲义第三卷中已经朝着这种推广走了一大步。费曼遗留的问题是，他没有真正表明这是一种可以重建现代量子力学理论体系的新方式。费曼没有彻底摆脱狄拉克方式的影响，比如他没有从跃迁元和跃迁元乘法中构建量子态的相关概念，而是一开始就夹杂着这些概念。当然，费曼这本书中也没有使用跃迁元的说法，他也没有把他的想法和海森堡的原始想法联系起来。

本书就是结合海森堡的原始想法和费曼的发展的一个尝试。我们将表明，从这种结合中，可以发展出一种理顺量子力学理论并重建现代量子力学体系的系统化方式。在我们的这种方式中，没有与幺正性矛盾的量子态坍缩，也没有多世界。有的是物理根基非常深厚的量子跃迁，而且，只要你正确理解幺正性，那么它就和这种量子跃迁完全不矛盾。

本书的最终结论是，量子力学测量问题中，只有退相干是真正物理的，测量问题不多于退相干问题，至于量子态坍缩以及多世界之类的想法则可能是不必要的，是对一些人为的抽象构造做了过度的延伸。

另外，由于完全是从零开始处理这一切，尤其是不会有公设从天上掉下来，所以我们这本书也同样适合作为一本量子力学的入门书，所以才叫《初识量子力学》。作为入门书，在处理量子力学理论体系时，我们这本书当然和狄拉克味道的书很不一样，这一点前面已经大致阐述过了。但是，它也和通常侧重波动力学的量子力学书很不一样。实际上，波动力学只在本书的最后一章才出现，而且不是作为重点。作为一本新式的量子力学书，和通常量子力学书的另一个区别是，我们用了单独一整章来讨论量子纠缠，因为它很基本也越来越重要。另外就是，从我们的方式引入量子力学路径积分是相当自然的，这将构成本书最后面的那点内容。

由于本人学识之所限，本书的错漏在所难免，欢迎您的交流、探讨以及批评、指正。

第一章 矩阵革命

物理学是大自然最深层次的秘密，大自然从不会直接把这些秘密暴露给人类，但它偶尔会透露给我们一些蛛丝般细微的线索。而人类，往往能够凭借其天才，从这些细微的线索中，半猜测半推断地揭开大自然部分最迷人的秘密。量子力学的建立就是一个最为成功的例子。

本章当然会涉及到量子力学在历史上的一些发展，但是，值得说明的是，我们并不是要还原历史，我们借助历史的一些印迹，目的却是为了更加自然地建立起整个量子力学的理论体系，因此本章中的历史是一种经过重构以后的“历史”，它只是一个故事，或者是历史的一种可能性，而不能完全等同于真实的历史，想了解真实历史的读者应该去阅读物理学史方面的著作。

1.1 玻尔的新观念

前玻尔时代

我们的故事实际上要从玻尔才正式开始，但是在历史上，量子观念的引进却并不是从玻尔开始的，因此为了讲清楚玻尔的故事，我们需要一些前玻尔时代的背景，尤其是需要介绍一下普朗克和爱因斯坦的开创性工作。将这两个伟大物理学家的开创工作归入前玻尔时代并不意味着它们不如玻尔重要，只是从本书的理论逻辑来说，玻尔所引入的观念与我们的关系更加密切一些。不过，为了突出本书的所需要的核心观念，我们决定把对普朗克和爱因斯坦相关工作的简单介绍穿插在对玻尔工作的具体讨论中，这里只是提请读者注意，它们实际是发生在玻尔的工作之前。

1.1.1 线索

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

而那时候人们还发现了原子光谱的另一个令人吃惊的规律，这个规律我们今天已经很少提到，但它对于物理学的发展也许是极其重要的，这规律就是里兹组合规则，这规则是说，人们总是可以合适地将原子光谱的每一条分立谱线都对应到一个正整数对 (m, n) ，其中一个正整数用来标记谱线属于哪一个谱线系，另外一个整数用于指明它属于此系中的第几条谱线，人们发现，相应的谱线角频率 ω_{mn} 满足如下这条规则，

$$\omega_{mn} = \omega_{ml} + \omega_{ln}. \quad (1.1)$$

也就是说，假设有一条谱线角频率是 ω_{ln} ，另外还有一条谱线角频率是 ω_{ml} ，那么这两条谱线的组合 $\omega_{ml} + \omega_{ln}$ 也必定是一条谱线，对应于整数对 (m, n) 。这一规律就是所谓的里兹组合规则。并且人们发现可以进一步定义

$$\omega_{nm} = -\omega_{mn}, \quad (1.2)$$

从而引入谱线间的减法，以使得里兹组合规则对于所有的正整数对指标都成立。

历史的发展表明，里兹组合规则的这两个公式(1.1)和(1.2)是最为强有力的线索，它在指引人们建立正确的量子理论上有重要作用。当然，经典物理根本无法解释这两条规则。

经典物理的问题还不仅仅是解释不了原子的分立光谱和其里兹组合规则，它在逻辑上更大的问题是，按照经典物理，绕核运动的电子会不断地发出电磁辐射，电磁辐射本身是有能量的，因此核外电子就会不断地损失能量，这样它就会一边绕核运动一边逐步往原子核上掉，很快所有的核外电子都会完全掉落到原子核上。原子的大小是 $10^{-10}m$ 的量级，但是原子

核的大小只有 $10^{-15}m$ 左右，因此很快所有原子的尺寸都会坍缩5个数量级。我们的世界是由原子构成的，因此这也就意味着整个世界都会坍缩。这当然是不对的，可以说当时的物理学家面临的更大难题是，如何拯救世界。

这个拯救世界的人就是玻尔。但在讨论玻尔如何拯救了世界之前，我们还是先回到当时关于原子光谱的观测。当时关于各种原子的光谱有大量的观测数据，这些数据杂乱无章，除了里兹组合规则以外，没有人能理出更多头绪。但是不久之后，通过巴尔末和里德堡等人对观测数据的归纳整理，就最简单的原子，氢原子的光谱，人们逐渐理出了头绪。人们发现，所有氢原子的光谱线都可以写成下面这个漂亮的数学公式

$$\omega_{mn} = 2\pi R c \left(\frac{1}{m^2} - \frac{1}{n^2} \right), \quad (1.3)$$

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

首先，不难发现，巴尔末和里德堡的这个公式自动满足里兹组合规则的两个公式(1.1)和(1.2)，因为它把谱线表达成了两项之差。实际上，一般性的，对于任何原子的任何谱线，为了满足里兹组合规则的两个公式(1.1)和(1.2)，其角频率都应该是两项之差，不妨写作

$$\omega_{mn} = T_n - T_m. \quad (1.4)$$

只不过，除了氢原子光谱有漂亮的巴尔末-里德堡公式以外，人们并没有找到其它原子 T_n 的公式。不过，光谱总能够写成两项之差，这对于玻尔是一个巨大的启发。

正因为氢原子是最简单的原子，而且它的光谱满足上面那个奇迹般的公式。所以当时玻尔等物理学家想到，要从理论上解释原子光谱的数据，首先就应该研究氢原子，并推导出公式(1.3)。这就是玻尔的氢原子模型解决的问题。

1.1.2 玻尔的新观念

为了解释氢原子光谱，玻尔提出了三条假设，这三条假设就构成了玻尔的氢原子模型。首先，针对核外电子由于辐射损失能量而不能稳定运动，

要坍缩到原子核上的问题。玻尔提出了定态假设，玻尔说，考虑到当时刚刚兴起的量子物理以后，核外电子绕核运动的轨道不能是任意的，而是只能取某些特定的轨道，当电子在一个这样的特定轨道，比方说第 n 个特定轨道上运动时，它是不辐射电磁波的，因此它的能量就不会损失，而是会保持在一个确定的值 E_n ，核外电子在特定轨道上运动的这种状态就称之为定态。

定态是玻尔提出的第一个全新观念！定态之间没有连续过渡的轨道就意味着，定态的能量不是一个连续值，而是分立的 E_n ，或者说，原子能量是量子化的，原子能量按照从小到大形成一阶一阶的能量“阶梯”，也称为能级。当然，定态的概念，以及定态能量量子化的推论不仅适用于氢原子，而且普遍适用于所有原子！只不过，对于氢原子，核外电子只有一个，而其它原子则有多个核外电子，因此这些电子的定态轨道就更加复杂。

其实，早在玻尔之前，普朗克就提出过能量分立的假设，这就是著名的普朗克能量量子化假设，它在历史上第一次引入了量子的概念。普朗克的假设是为了解释当时的黑体辐射难题，具体内容和本书关系不大，所以我们略去，只说结论。总之，通过对黑体辐射的研究人们发现，原子辐射和吸收角频率为 ω 的电磁波时，其行为就好像一个角频率为 ω 的线性谐振子，而普朗克假设，一个线性谐振子的能量是量子化的，其必须是某个基本能量的整数倍，这个基本能量是 $\hbar\omega$ ，其中 \hbar 是一个常数，称作普朗克常数，其量纲是能量量纲乘以时间量纲，或者长度量纲乘以动量量纲(和角动量的量纲一样)，在国际单位制中，其值为 $1.05457266(63) \times 10^{-34} \text{ J} \cdot \text{s}$ ，非常小。换言之，普朗克假设，线性谐振子的能级 E_n 由下式给出

$$E_n = n\hbar\omega, \quad (1.5)$$

n 为整数。注意，普朗克并非任意作出这个假设的，而是，他发现非如此假设不足以解释实验上测出来的黑体辐射谱。

普朗克常数是自然界一个最基本的常数，其出现是量子效应的典型特征，而这个常数在国际单位制中是如此之小也正解释了为什么通常的日常生活中人们感受不到量子效应的原因，毕竟，国际单位制就是适用于日常生活的单位制。不过，单位制的选取毕竟是人为的，因此如果人们要考察的是一个纯粹量子的世界，那选取合适的单位以使得 $\hbar = 1$ 也许就更方便。

这就好像如果人们要考察的是一个高速的相对论性世界，选取光速 $c = 1$ 的单位制更方便一样。

总之，玻尔的定态假设并不是完全凭空产生的，而是在普朗克能量量子化假设的基础上又往前走了一步，一般性地将能量的量子化归结于定态的存在。

与普朗克不同的是，玻尔并不是直接假设氢原子能级 E_n 的具体表达式，毕竟氢原子能级比谐振子复杂好多，并不存在一个自然而又简单的假设。相反，玻尔是通过结合其它简单假设，最终推导出了 E_n 的表达式。

既然核外电子在定态上并不辐射，那么，为什么我们能观测到原子的光谱呢？为此，玻尔引入了第二条假设，跃迁假设。玻尔说，核外电子处在定态上的时候虽然不辐射，但是，它可能从一个定态跃迁到另一个定态，比方说从第 n 个定态跃迁到第 m 个定态，这时候电子就会辐射出特定频率 ω_{mn} 的电磁波，按照爱因斯坦此前提出的光量子假设，这个辐射出来的光子的能量就是 $\hbar\omega_{mn}$ ，因此由能量守恒，玻尔写出了如下方程，

$$\hbar\omega_{mn} = E_n - E_m. \quad (1.6)$$

同样，玻尔的这第二条假设也是普遍适用于所有的原子，而且，它自然地解释了为什么原子光谱可以写成两项之差！

这里简单解释一下爱因斯坦的光量子假设，它是爱因斯坦对普朗克量子论的发展，它提出，电磁波本身是量子化的，每一个电磁波的量子（电磁波的基本组成部分）就是一个光量子，也就是今天人们熟悉的光子。一个角频率为 ω 的光子，其能量为 $E = \hbar\omega$ 。爱因斯坦利用光量子假设成功地解释了光电效应，而且康普顿散射实验也验证了光量子的正确性。

玻尔的这两条假设一起，就自然地解释了原子光谱的分立性，原因就在于定态能级是分立的，而且公式(1.6)进一步告诉我们，每一条光谱线的确都对应两个正整数 m, n 。

至于原子为什么稳定，玻尔给出的最终解答是这样的，根据定态假设，原子的能量会量子化为能级，假设把这些能级按照能量从低到高排列，那么低能级的核外电子轨道必定更靠近原子核，因为从经典物理就可以知道，越靠近原子核的轨道能量就越低，主要原因是原子核与核外电子之间的库仑吸引能会负得越厉害，而如果核外电子掉到核上，那这个库仑能就是负无穷。但是，玻尔最终可以推导出来，总是有一个最低能级，其能量是有

限的，也就是说，这个最低能级对应的定态轨道并不是核外电子落到原子核上。因此，当核外电子发出辐射从高能级往低能级跃迁的时候，它最多跃迁到最低能级就不可能接着往下跃迁了，因此，核外电子永远也不会掉到原子核上去，原子是稳定的！

但是，如果想进一步推导出氢原子光谱的巴尔末-里德堡公式(1.3)，玻尔就必须进一步找到定态能量 E_n 的表达式。类比于行星绕太阳运动，玻尔知道决定核外电子轨道的是能量和角动量这两个守恒量，量纲分析又告诉他，考虑到量子效应，必然有 $J \sim \hbar$ 。因此玻尔又引入了第三条假设，称之为轨道角动量量子化假设，玻尔说，处在定态 n 上的电子的轨道角动量 J 是量子化的，它满足

$$J = rp = n\hbar. \quad (1.7)$$

式中 r 是核外电子离核的距离， p 是核外电子的动量。注意，这第三条假设是特殊的，它只适用于氢原子，因为其它原子的核外电子轨道涉及到多个电子。

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

$$m_e \frac{v^2}{r} = \frac{e^2}{4\pi\epsilon_0 r^2}, \quad (1.8)$$

式中 m_e 为电子质量， e 为电子电荷， ϵ_0 为真空的介电常数。不妨引入如下无量纲常数 α

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c}, \quad (1.9)$$

称作精细结构常数，它的数值大约是 $1/137$ 。从而，公式(1.8)也可以重写成，

$$m_e v^2 = \frac{\alpha c \hbar}{r}. \quad (1.10)$$

另外，由角动量量子化假设我们有

$$rm_e v = n\hbar. \quad (1.11)$$

由这两个结果我们容易推导出，定态的轨道半径 r 必然满足

$$r = n^2 \frac{\hbar}{m_e \alpha c}. \quad (1.12)$$

特别的，离原子核最近的第1个定态的轨道半径为 $\frac{\hbar}{m_e \alpha c} = a_0$, a_0 称之为玻尔半径，可以算得它的数值是 $0.52917721067(12) \times 10^{-10} m$ 。

另一方面，氢原子的能量是动能和势能的和，即 $E = \frac{1}{2}m_e v^2 - \frac{e^2}{4\pi\epsilon_0 r} = \frac{1}{2}m_e v^2 - \frac{\alpha ch}{r}$ ，由式(1.10)就可以得到， $E = -\frac{1}{2} \frac{\alpha ch}{r}$ (当然它也等于 $-\frac{1}{2}m_e v^2$)。将定态轨道半径的结果(1.12)代入这个能量的式子，我们就能得到第 n 个定态能量 E_n 的表达式

$$E_n = -\frac{1}{2}m_e \alpha^2 c^2 \frac{1}{n^2} = \frac{E_1}{n^2}. \quad (1.13)$$

式中 $E_1 = -\frac{1}{2}m_e \alpha^2 c^2$ 就是氢原子第1个定态的能量，它也就是氢原子最低能态(也称之为基态)的能量，代入相应物理量的值就可以算得 $E_1 = -13.6 eV$ 。 $E_1 = -\frac{1}{2}m_e \alpha^2 c^2$ 的这个结果很好记，由于氢原子的总能量等于 $-\frac{1}{2}m_e v^2$ ，而在基态的时候电子的速度 $v = \alpha c$ ，因此就有 $E_1 = -\frac{1}{2}m_e \alpha^2 c^2$ 。

更令人高兴的是，将公式(1.13)代入跃迁假设的公式(1.6)，玻尔就能导出氢原子的光谱，

$$\omega_{mn} = \frac{m_e \alpha^2 c^2}{2\hbar} \left(\frac{1}{m^2} - \frac{1}{n^2} \right), \quad (1.14)$$

这和神奇的巴尔末里德堡公式(1.3)形式完全一样，将它和巴尔末里德堡公式比较我们就可以得到，如果里德堡常数可以由下式给出

$$R = \frac{m_e \alpha^2 c}{4\pi\hbar}, \quad (1.15)$$

那么玻尔就正确地推导出了氢原子的光谱。代入物理常数的值人们容易算出公式(1.15)给出的里德堡常数和由光谱数据定出来的里德堡常数吻合得非常好，这充分说明了玻尔的模型的确解释了氢原子的光谱。

正如我们已经看到的，为了解决氢原子的难题，玻尔引入了一些革命性的新概念，但也可以看出玻尔理论的缺陷，在玻尔推导定态能量 E_n 表达式的过程中，他还需要求助于经典力学的牛顿方程，而他也没有用一个新的基本原理将牛顿方程和他新引入的三条假设统一起来。换言之，玻尔并没有证明牛顿方程和他的三条假设是相容的，相反，正如前面所说的，牛

顿方程和经典的电磁学理论结合起来会导致完全错误的结论，而玻尔的三条假设正是为了挽救这样的错误而引入的，因此这些假设实际上天然地与牛顿方程不相容。因此，玻尔的理论虽然取得了巨大的成功，但在理论结构上依然是不令人满意的，它遗留的问题几乎和它解决的问题一样大。

不过，玻尔关于定态和定态跃迁的全新观念将成为量子力学最为基本性的观念，也是本书的核心概念，值得我们作一些进一步的阐述。定态是核外电子的特定轨道，海森堡后来指出，电子轨道不可以观察，可观察的是光谱。所以，更恰当的看法是，**定态就是原子的一种确定可能性**，在这种可能性中，原子有一个分立的确定能量值。

要描述一个原子，只需要知道它的所有定态，因此，原子所有定态的集合就构成了原子的一种**可能性完备集**。由于原子存在一个最低能级，所以我们总是可以把不同定态按照能量从低到高排列，使 $E_1 \leq E_2 \leq E_3 \leq \dots$ ，相应的不同定态可能性就可以由正整数 n 来标记，它相应于第 n 能级。不妨将原子的第 n 个定态可能性记作 (n) ，从而原子的定态可能性完备集就是集合 $\{(n) | n = 1, 2, \dots\}$ 。

不过，对于氢原子，后来完整的量子力学理论表明，标记不同定态可能性的更合适方法是同时用四个数，除了决定定态能量的 n 以外，还需要 l, m, m_s 这三个数，其中 l 和 m 都是整数，而 $m_s = \pm 1/2$ ，所以，更合适的标记氢原子某个定态可能性的方法是将之记作 $(nlmm_s)$ ，当然，这个细节性的问题其实和建立量子力学的基本原理无关，只和它在原子物理中的应用有关，所以我们就不进一步阐述了。

下面讨论定态跃迁。为了更直观地阐述跃迁过程，我们以氢原子为例，画出示意图(1.1)，图中水平方向表示态可能性的空间，竖直方向代表时间，两条竖直直线分别代表定态可能性 (n) 和定态可能性 (m) 的电子，波浪线代表辐射出来的光子，这个图示意的就是电子从高能级 (n) 跃迁到低能级 (m) ，并辐射出一个光子的过程。

后来人们通过对原子发光过程的进一步研究指出，原子什么时候跃迁，跃迁到哪个定态上去，都完全是随机的。因此，**定态跃迁原则上就是随机的**，只能讨论跃迁的概率！这一点其实非常重要，虽然前文我们没怎么强调，但它是量子物理不同于经典物理的一个核心特征，即在量子物理里面，**我们只能讨论概率**，后来这成为玻恩的一个洞察，也是他得诺贝尔奖

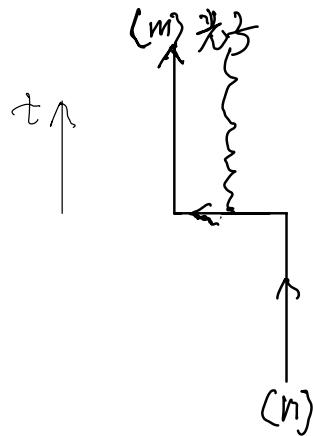


图 1.1: 水平方向表示态可能性的空间, 坚直方向代表时间, 两条直线分别代表定态可能性(n)和定态可能性(m)的电子, 波浪线代表辐射出来的光子。

的原因。

1.2 矩阵革命

差不多是夜里三点钟, 计算结果最终出来了。我深深地被震惊了。我很兴奋, 一点也不想睡。于是, 我离开房间, 坐在一块岩石上等日出。

——————海森堡

线索

在玻尔提出它的原子模型之后, 人们就已经完全理解里兹组合规则了,

$$\omega_{mn} = \omega_{ml} + \omega_{ln}, \quad \omega_{nm} = -\omega_{mn}. \quad (1.16)$$

之所以有这样的规则成立, 原因很简单, 因为根据玻尔的定态跃迁假设, 原子从第 n 个定态跃迁到第 m 个定态, 辐射(或者吸收)的光谱频率为 $\omega_{mn} = (E_n - E_m)/\hbar$, 其中 E_n 就是第 n 个定态的能量, 如此一来当然就有里兹组合规则了。对里兹组合规则(1.16)的解释很简单, 那就是核外电子

从 n 态跃迁到 m 态所放出的能量，等于它从 n 态跃迁到中间 l 态所放出的能量加上它从中间 l 态跃迁到 m 态放出的能量。问题是，给定初始的 n 态和末尾的 m 态，里兹组合规则中的中间 l 态可以是任意的，里兹组合规则对 l 没有任何限制！

1.2.1 跃迁元与跃迁元乘法

故事的下一个主角是海森堡，对于当时的海森堡来说，玻尔氢原子模型中所涉及到的核外电子轨道概念是没有道理的，因为当时的海森堡在物理学哲学上深受玻尔的影响，他坚持物理学应该用可观测量建立起来，或者换句话说，他认为物理学理论中所用到的量应该都是可观测的。我们这里不讨论这一哲学有没有道理，但是当时年轻的海森堡深信这一哲学，因此他觉得核外电子轨道的概念没有道理，因为不可观测。对于原子物理来说，可观测的是什么呢，是光谱，是谱线频率和谱线的强度，不是电子轨道，电子轨道不过是对太阳系模型的一种类比，而不是原子物理的观测事实。因此海森堡觉得关键是要研究谱线频率和谱线强度。但是海森堡也知道，原子光谱的分立性和里兹组合规则意味着，虽然核外电子轨道的概念没有道理，但是玻尔的定态和定态跃迁假设必定是对的。

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

前面说过，海森堡坚持原子物理应该回到可观测的量，也就是谱线频率和谱线强度。玻尔的原子模型当然对谱线频率的解释很不错，尤其是完全推导出了氢原子的所有光谱线。但是玻尔氢原子模型并没有涉及谱线的强度，因此海森堡想推进当时的物理学，那就要研究谱线的强度。注意到定态跃迁是随机的，只能谈论跃迁概率，而核外电子从 n 态跃迁到 m 态时所发出的这条谱线的强度当然取决于单位时间之内它从 n 态自发跃迁到 m 态的

概率 A_{mn} , 而这是一个当时的物理学家(爱因斯坦)已经提出来了物理量, 只是当时还没有人会计算这个量, 海森堡要做的, 就是研究如何计算出这个量。但是这个量涉及到原子和电磁场的相互作用, 而当时对电磁场的量子理论还只是知道一个光量子假设再加上爱因斯坦的受激辐射理论, 对于原子和电磁场如何相互作用的量子理论还一无所知, 毕竟, 量子力学的正确理论还正有赖于海森堡目前正在工作的这一工作呢。唯一知道的是, 辐射过程涉及的原子和电磁波的相互作用对于原本的原子而言是一个微扰。因此海森堡只能转而去看经典理论是如何计算电磁波辐射的, 通过经典电动力学我们知道, 一个加速运动的电子会辐射电磁波, 电动力学也推导出了其辐射功率为

$$P = \frac{e^2}{3\pi\epsilon_0 c^3} |\ddot{\mathbf{x}}|^2, \quad (1.17)$$

式中 $\ddot{\mathbf{x}}$ 就是电子的加速度。还没有学过电动力学从而不知道这个公式也没有关系, 把加速运动的电子会辐射出电磁波当成经典物理的一个基本事实来接受就可以了。

但是, 海森堡知道这个经典公式(1.17)肯定是对的, 不仅因为它的表达式中涉及到了核外电子轨道 $\mathbf{x}(t)$ 这样一个在海森堡看来无法观测的量, 更重要的是, 它和定态以及定态跃迁的概念是相矛盾的。当然, 玻尔引入定态和定态跃迁的原因之一本就是为了避免由于(1.17)这一经典公式而导致的原子稳定性问题。因此海森堡首先要做的就是改造这个经典公式(1.17), 使得它和定态以及定态跃迁的概念相容。

根据定态跃迁假设, 核外电子只在从一个定态 n 跃迁到另一个定态 m 时才辐射, 注意, 跃迁总是涉及到两个定态, 初态 n 和末态 m 。因此海森堡想到, 经典意义上的 $\mathbf{x}(t)$ 是不可观测的, 是没有意义的, 取而代之的应该是一个类似于 $[\mathbf{x}]_{mn} \equiv [\mathbf{x}_0]_{mn} e^{-i\omega_{mn} t}$ 这样的涉及两个定态 m, n 的量。其中 $[\mathbf{x}_0]_{mn}$ 与时间无关, 而海森堡将这个量对时间的依赖因子写成 $e^{-i\omega_{mn} t}$ 的原因在于, 通过普朗克对黑体辐射公式的推导, 海森堡知道, 原子定态跃迁发出电磁辐射的过程可以看成是一种角频率为 ω_{mn} 的简谐振动, 而一个作简谐振动的物理量, 其时间依赖就是这样的。至于写成指数形式而不是经典物理中更常见的只保留实部的 $\cos(\omega_{mn} t)$ 形式的原因是, 在经典物理中, 根据欧拉公式, 实的 \cos 形式总可以写成指数形式与其复共轭的和, 而海森堡认为在量子对应中, 这两部分(指数及其复共轭)是相互独立的, 海

森堡将定态 n 到定态 m 的跃迁对应到了 $[\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$, 其复共轭部分则对应到了反向的定态 m 到定态 n 的跃迁(即后面的厄米条件), 当然, 代价就是引入了复数, 而物理量的实数性就是通过额外规定厄米条件补回来的。更何况, 正如后文会看到的, 当将两个物理量相乘时, 如果这两个量对时间的依赖因子是cos形式, 那是不可能满足里兹组合规则的, 而写成指数形式则很轻易就可以满足这一点。由于相关的量变成了复数, 所以这时候(1.17)式中的矢量模长||当然就必须同时是求复数模长, 因为最终物理上可观测的辐射功率当然是实数。总之, 海森堡将 $[\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$ 代入公式(1.17)就得到

$$P(n \rightarrow m) = \frac{e^2 \omega_{mn}^4}{3\pi\epsilon_0 c^3} |[\mathbf{x}]_{mn}|^2. \quad (1.18)$$

注意, 现在的这个新公式(1.18)不仅用到了定态跃迁的概念, 它和定态假设也是相容的, 由于 $\omega_{nn} = 0$, 因此就有 $P(n \rightarrow n) = 0$, 这当然与定态假设一致(指定态上不辐射)。经典的轨道 $\mathbf{x}(t)$ 不可观测, 但是现在的 $[\mathbf{x}]_{mn}$ 是可以观测的, 这是因为, 从 n 态跃迁到 m 态的辐射功率等于单位时间辐射概率 A_{mn} 乘以辐射出来的光子能量 $\hbar\omega_{mn}$, 因此由(1.18)我们就可以得到 A_{mn} 的表达式, $A_{mn} \propto |[\mathbf{x}]_{mn}|^2$ 。 A_{mn} 当然可以观测, 因此 $[\mathbf{x}]_{mn}$ (严格来说是其模 $|[\mathbf{x}]_{mn}|$)就可以观测。另外, 前面说过, 辐射本身对原子是一个微扰, 很显然这个微扰与 $[\mathbf{x}]_{mn}$ 密切相关。

当然, 由于原子自发跃迁是从高能级跃迁到低能级, 所以 $[\mathbf{x}]_{mn} = [\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$ 这个量仅当 $E_n \geq E_m$ 时(即 $n \geq m$ 时)才有定义。不过, 海森堡注意到 $e^{-i\omega_{nm}t} = e^{i\omega_{mn}t} = [e^{-i\omega_{mn}t}]^*$, 所以他进一步规定

$$[\mathbf{x}_0]_{nm} = [\mathbf{x}_0]_{mn}^*. \quad (1.19)$$

这样一来, $[\mathbf{x}]_{mn}$ (也就是 $[\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t}$)就对任意的正整数指标 n, m 均有定义了。而海森堡规定的条件(称作厄米条件)就成为

$$[\mathbf{x}_0]_{nm}e^{-i\omega_{nm}t} = ([\mathbf{x}_0]_{mn}e^{-i\omega_{mn}t})^*, \quad (1.20)$$

也就是

$$[\mathbf{x}]_{nm} = [\mathbf{x}]_{mn}^*. \quad (1.21)$$

也就是说, 反向的 m 到 n 的跃迁等于正向 n 到 m 跃迁的复共轭。正如前面提过的, 这条件对应于经典物理量的实数性。

根据前文的讨论，在一个正比于物理量 \mathbf{x} 的微扰的影响下(这个微扰就是核外电子与电磁场之间的相互作用)，原子从定态 n 跃迁到定态 m 的概率(即前面的 A_{mn})将与 $|[\mathbf{x}]_{mn}|^2$ 成比例，这就是 $[\mathbf{x}]_{mn}$ 这个量的物理诠释。

这样一来，对应经典物理里面的位置坐标 $\mathbf{x}(t)$ ，海森堡就引入了一个新的物理量 $[\mathbf{x}]_{mn}$ 。在这个物理量中，如果给定 m 和 n ，那 $[\mathbf{x}]_{mn}$ 就是一个复数，但是，指标 m, n 也可以看作变量，跑遍所有正整数，所以，这时候 $[\mathbf{x}]_{mn}$ 就是一张有无穷行无穷列的数据表，排列出来就像下面这样

$$[\mathbf{x}]_{mn} = \begin{bmatrix} [\mathbf{x}]_{11} & [\mathbf{x}]_{12} & [\mathbf{x}]_{13} & \dots \\ [\mathbf{x}]_{21} & [\mathbf{x}]_{22} & [\mathbf{x}]_{23} & \dots \\ [\mathbf{x}]_{31} & [\mathbf{x}]_{32} & [\mathbf{x}]_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (1.22)$$

这个数据表可以向右方和下方无限延伸。 $[\mathbf{x}]_{mn}$ 的两个指标 mn 中，左边的指标 m 用来指示表中的元素处于第几行，因此称作行指标，而右边的指标 n 用来指示元素位于第几列，因此称作列指标。

读者不应该混淆符号 $[\mathbf{x}]_{mn}$ 的两种含义，当指标 m, n 看作可变时，它代表一整张数据表，而当指标 m, n 看作固定时，它又只代表这表中第 m 行第 n 列的元素，从而是一个复数。具体应该怎么看这两个指标则应该根据上下文。

请想象一张无穷无尽的数据表，任何纸张都是有限的，都写不下它，所以只能用省略号代之！而将位置坐标这样的概念变成一张这样的数据表，这本身就蕴含了最高的革命性。数学上，这样的数据表有一个专门的名称，叫做矩阵。不过，物理上，由于海森堡是在定态跃迁思想的指引下引入 $[\mathbf{x}]_{mn}$ 的，所以我们不妨称之为跃迁元。

有时候，我们也示意性地将 $[\mathbf{x}]_{mn}$ 记成如下样子，

$$[\mathbf{x}]_{mn} = \left\{ (m) \xleftarrow{\mathbf{x}} (n) \right\}. \quad (1.23)$$

其中 $(m) \xleftarrow{\mathbf{x}} (n)$ 表示从定态 n 到定态 m 的跃迁，箭头上面的 \mathbf{x} 表示跃迁是在物理量 \mathbf{x} 的作用下产生的，而外面的花括号则表示取这个跃迁的跃迁元。

而海森堡规定的条件(1.21)用矩阵的形式来写就成了,

$$\begin{bmatrix} [\mathbf{x}]_{11} & [\mathbf{x}]_{12} & [\mathbf{x}]_{13} & \dots \\ [\mathbf{x}]_{21} & [\mathbf{x}]_{22} & [\mathbf{x}]_{23} & \dots \\ [\mathbf{x}]_{31} & [\mathbf{x}]_{32} & [\mathbf{x}]_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} [\mathbf{x}]_{11}^* & [\mathbf{x}]_{21}^* & [\mathbf{x}]_{31}^* & \dots \\ [\mathbf{x}]_{12}^* & [\mathbf{x}]_{22}^* & [\mathbf{x}]_{32}^* & \dots \\ [\mathbf{x}]_{13}^* & [\mathbf{x}]_{23}^* & [\mathbf{x}]_{33}^* & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

也就是说, 这样的矩阵满足一个特殊的性质, 即将其行与列的相应元素对调(第1行与第1列对调, 第2行与第2列对调, 依次类推), 同时把每一个对调之后的元素换成其复数共轭, 结果得到的矩阵等于原来的矩阵。数学上称这样的矩阵为厄米矩阵!

习惯上也将整张跃迁元数据表 $[\mathbf{x}(t)]_{nm}$ 简记作 $\mathbf{X}(t)$, 也即是说

$$\mathbf{X}(t) \equiv [\mathbf{x}(t)]_{nm} \equiv [\mathbf{x}_0]_{nm} e^{-i\omega_{nm} t}. \quad (1.24)$$

在这个记号中 $\mathbf{X}(t)$ 就不再表示经典的位置矢量了, 而是表示与之相应的整张跃迁元数据表。

当我们有了位置物理量 $[\mathbf{x}(t)]_{mn} = [\mathbf{x}_0]_{mn} e^{-i\omega_{mn} t}$ 以后, 将它对时间求导当然就得到速度 $[\mathbf{v}]_{mn} = -i\omega_{mn} [\mathbf{x}_0]_{mn} e^{-i\omega_{mn} t}$, 进而就有动量 $[\mathbf{p}]_{mn} = m_e [\mathbf{v}]_{mn} = -im_e \omega_{mn} [\mathbf{x}_0]_{mn} e^{-i\omega_{mn} t}$, 这里 m_e 表示电子质量, 而且不难验证

$$[\mathbf{p}]_{nm} = [\mathbf{p}]_{mn}^*. \quad (1.25)$$

所以, 动量也变成了一张跃迁元数据表, 作为矩阵也是一个厄米矩阵。通常也将动量的跃迁元数据表简记如下

$$\mathbf{P}(t) \equiv [\mathbf{p}]_{nm} \equiv [\mathbf{p}_0]_{nm} e^{-i\omega_{nm} t}. \quad (1.26)$$

同样, 在这种记号中 $\mathbf{P}(t)$ 不再代表经典的动量变量, 而是代表与之相应的跃迁元数据表。

在海森堡想来, 一切物理量都应该变成跃迁元数据表, 只有跃迁元才是可观测的。由于任何物理量均是位置和动量的函数, 所以, 有了位置的跃迁元数据表和动量的跃迁元数据表以后, 原则上应该就可以构造出其它一切物理量的跃迁元数据表。由于定态跃迁可以看成是一种谐振动, 海森堡知道所有这些跃迁元应该都具有 $[A]_{mn} \equiv [A_0]_{mn} e^{-i\omega_{mn} t}$ 这样的形式, 式

中 $[A_0]_{mn}$ 与时间无关。当然，作为对经典物理量实数性的量子对应，这些跃迁元数据表也要满足类似的海森堡厄米条件

$$[A]_{mn} = [A]_{nm}^*. \quad (1.27)$$

从而作为矩阵也是厄米矩阵。和前面类似，厄米条件的含义是，反向的从 n 到 m 的跃迁元，是正向从 m 到 n 跃迁元的复数共轭，也可以示意性地记为

$$\left\{ (m) \xleftarrow{A} (n) \right\} = \left\{ (m) \xrightarrow{A} (n) \right\}^*. \quad (1.28)$$

而且，完全类似于前面对 $[x]_{mn}$ 的物理诠释， $[A]_{mn}$ 的物理诠释即是：如果设法引进一个正比于物理量 $A(t)$ 的微扰相互作用，使之引起原子系统发生定态跃迁，则原子从定态 n 跃迁到定态 m 的概率将与 $|[A]_{mn}|^2$ 成比例。

值得说明的是，习惯上常常将整个跃迁元数据表 $[A]_{mn}$ 简记作 A ，即

$$A \equiv [A]_{mn} \equiv [A_0]_{mn} e^{-i\omega_{mn}t}. \quad (1.29)$$

在这种记号规则中，同一个记号 A 既代表物理量，又代表与这个物理量相应的跃迁元数据表(或者说矩阵)。

对于一个矩阵 M ，当将其行与列的相应元素对调(第1行与第1列对调，第2行与第2列对调，依次类推)，同时把每一个对调之后的元素换成其复数共轭，我们就得到一个新的矩阵，称作 M 的厄米共轭矩阵，记作 M^\dagger ，所以

$$[M^\dagger]_{mn} = [M]_{nm}^*. \quad (1.30)$$

不难看出，厄米共轭再厄米共轭，一定等于本身，即

$$(M^\dagger)^\dagger = M. \quad (1.31)$$

有了厄米共轭的概念以后，物理量 A 的跃迁元数据表需要满足的海森堡条件(1.27)，就可以重记为

$$A = A^\dagger, \quad (1.32)$$

也即是说，厄米矩阵就是厄米共轭以后等于其自身的矩阵。

但是，这里有一个重要的问题，比方说动能 T ，它在经典物理里面是这样的 $T = \frac{\mathbf{p}^2}{2m_e}$ ，这里涉及到动量的平方。更一般地，假设物理量 C 等于两个物理量 A 和 B 的乘积，即 $C = AB$ 。那么根据海森堡的跃迁元思想，就需要将 C 的跃迁元 $[C]_{mn} \equiv [C_0]_{mn}e^{-i\omega_{mn}t}$ ，用 A 的跃迁元 $[A]_{mn} \equiv [A_0]_{mn}e^{-i\omega_{mnt}}$ 与 B 的跃迁元 $[B]_{mn} \equiv [B_0]_{mn}e^{-i\omega_{mnt}}$ 的乘积的形式表达出来。但现在 A 和 B 的所有跃迁元都分别构成了一张无穷行无穷列的数据表，两张这样的数据表该怎么乘，就是海森堡需要解决的大难题。

当然，学过线性代数的读者都知道(我们并不假定读者学过线性代数)，跃迁元数据表作为矩阵，可以按照标准的矩阵乘法来乘。但是，矩阵乘法只是数学家规定的一种运算，而现在我们的跃迁元数据表却是物理，因此物理上，你怎么知道跃迁元数据表应该按照矩阵乘法来乘呢？毕竟，两个数据表之间数据相乘的可能性可太多了，为什么一定要按照矩阵乘法来乘呢？更何况，在海森堡的时代，和今天不同，那时候物理学家并不学线性代数，所以海森堡根本就不知道存在矩阵这样的数学对象，更不知道矩阵乘法。所以，对于海森堡而言，他要处理的问题就是，给两个跃迁元数据表定义一种合适的乘法，使得它能够正确地把两个物理量乘在一块。

正是在这里，里兹组合规则(1.16)，或者等价地说玻尔的公式 $\omega_{mn} = (E_n - E_m)/\hbar$ 给了海森堡必要的指引。由定义， $[C_0]_{mn}e^{-i\omega_{mnt}} = [(AB)_0]_{mn}e^{-i\omega_{mnt}}$ ，而按照里兹组合规则， ω_{mn} 总是可以分裂成两部分的组合， $\omega_{mn} = \omega_{ml} + \omega_{ln}$ ，所以海森堡想到可以令 $[(AB)_0]_{mn}e^{-i\omega_{mnt}} = [A_0]_{ml}e^{-i\omega_{mlt}}[B_0]_{ln}e^{-i\omega_{lnt}}$ (从这里可以清楚地看到，如果不是将跃迁元数据表对时间的依赖写成 $e^{-i\omega_{mnt}}$ ，而是写成仅仅只有实部的cos形式，那就无法满足里兹组合规则了)，也就是说，应该有 $[(AB)_0]_{mn} = [A_0]_{ml}[B_0]_{ln}$ ，或者相应的 $[AB]_{mn} = [A]_{ml}[B]_{ln}$ ，看来问题就解决了。但是海森堡最天才的地方就在这里，他注意到里兹组合规则对中间的 l 态没有任何限制，也就是说 l 有无穷多种可能性，怎么对付这无穷多种可能性呢？海森堡的处理办法很简单，他说，既然我们不能确定中间的可能性 l 具体是哪个，那就应该平等地对待每一种可能性，因此最自然的处理办法就是把所有的可能性都加起来！因此海森堡给出

$$[C]_{mn} = [AB]_{mn} = \sum_l [A]_{ml} [B]_{ln}, \quad (1.33)$$

式中的求和是对所有的正整数 l 求和。我们可以示意性地把这种乘法记成如

下样子

$$\{(m) \xleftarrow{AB} (n)\} = \sum_l \{(m) \xleftarrow{A} (l) \xleftarrow{B} (n)\}. \quad (1.34)$$

当然学过线性代数的读者容易知道，这的确就是标准的矩阵乘法，但是海森堡并没有学过矩阵乘法，对他来说，之所以要这么乘，是因为有无穷多种中间可能性都满足要求，我们需要把所有的中间可能性都加起来！今天看来，这是一个极其原创性的想法，因为它把握住了量子力学相干叠加的本质！

一般来说，两个物理量相乘，结果当然是一个新的物理量。但是，有一种特别重要的物理量例外，那就是单位1，1和任何物理量A相乘，结果当然还是A。因此，相对应的，1的跃迁元数据表和任何跃迁元数据表A相乘，结果必须还是数据表A。按照我们的记号规则，1的跃迁元数据表应该记作 $[1]_{mn}$ ，其物理含义就是在不受任何微扰作用(1的意思就是不受任何微扰)的情况下，系统自发从n态跃迁到m态的跃迁元。因此，对于任何A，均有

$$[A]_{mn} = \sum_l [1]_{ml} [A]_{ln}. \quad (1.35)$$

不难验证满足这个条件的 $[1]_{mn}$ 必定为

$$[1]_{mn} = \delta_{mn}, \quad (1.36)$$

式中 δ_{mn} 称作克罗内克符号，它的定义就是当 $m = n$ 时取1，否则都取0，从而(1.35)式对l的求和中，只有一项有非零贡献，就是 $l = m$ 的那项，这时候 $[1]_{ml}$ 当然取1，而 $[A]_{ln}$ 就成了 $[A]_{mn}$ ，从而(1.35)式右边等于左边。 δ_{mn} 排成矩阵的话就是

$$\delta_{mn} = \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.37)$$

这种矩阵称作单位矩阵，它只在对角线上的元素为1，其它元素都是0。单位矩阵乘以任何矩阵A都还等于A。人们也常常简单地把单位矩阵就记作1，

根据上下文不难把它和数字1相区分开来。不难验证，对于任何矩阵 A ，不仅 $1 \cdot A = A$ ，而且也有 $A \cdot 1 = A$ 。

根据跃迁元的概率诠释， $|[1]_{mn}|^2$ 应该正比于系统不受任何微扰作用时自发从 n 态跃迁到 m 态的概率，但是， $[1]_{mn}$ 不是等于0就是等于1，无论是0还是1，其模方都是本身。因此这就说明 $[1]_{mn} = \delta_{mn}$ 本身就正比于系统不受任何微扰作用时自发从 n 跃迁到 m 的概率。又由于给定初态 n ，这种跃迁只有一种可能性，即仅当末态 $m = n$ 时，才有 $[1]_{mn}$ 等于1，其它 $m \neq n$ 的情况都为零。这说明，当系统不受任何微扰作用时，给定它的一个初态 n ，你去测量它末态能跃迁到所有定态可能性中的哪个，则会发现它只能百分之百从 n 跃迁到 n 自身，而跃迁到一个与初态不同的 $m \neq n$ 的末态的概率为零。当然， n 态跃迁到 n 态自身这种说法，实际上已经稍微拓宽玻尔的定态跃迁概念了。之所以作这种拓展，是因为它会带来语言表达上的适当简化。

以上讨论了海森堡引入的新思想，海森堡的想法是如此自然，却又如此新鲜大胆，问题是，这想法起作用吗？这个问题我们稍后再来解决，稍后我们将通过用海森堡的想法推导出线性谐振子的能量量子化假设的那个），进而验证这想法的确是起作用的。现在，让我们先发展一点必要的数学，已经熟悉相关数学的读者可以略过，或者只需简单浏览一下这个数学部分。

1.2.2 一些必要的数学

可能有些读者并没有学过线性代数，因此并不熟悉海森堡所制定的这种奇怪乘法规则

$$[AB]_{mn} = \sum_l [A]_{ml} [B]_{ln}, \quad (1.38)$$

所以我们对其作进一步的阐述。首先，看这个式子各项涉及的指标的时候，都是从右往左读， $[AB]_{mn}$ 描述的是从初始 n 态到末尾 m 态的跃迁元，它可以拆分成先从 n 出发在物理量 B 的作用下跃迁到中间可能性 l ，再在物理量 A 的作用下从中间可能性 l 跃迁到 m ，中间可能性 l 要求和。其次，注意等式(1.38)右边，数据表 A 和 B 相乘， A 的列指标和 B 的行指标是一样的，都是中间可能性 l ，而且它是要被求和的。这样被求和的指标我们称之为哑标，

所谓哑标，就是无需与等式(1.38)左边的指标对应存在的指标(等式左边的指标里面没有 l)，而剩下的两个指标 m 和 n 在等式(1.38)的两边是要对应存在的，称作自由指标。

不妨进一步举例说明这种矩阵乘法是如何进行的。为了简单起见，假设我们考虑的不是无穷行无穷列的数据表，而是两行两列的数据表，比方说

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}. \quad (1.39)$$

那么按照乘法规则(1.38)， AB 就是

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}. \quad (1.40)$$

有时候人们也说， AB 第1行第1列的元素是 A 第1行乘以 B 第1列的结果，同样， AB 第1行第2列的元素是 A 第1行乘以 B 第2列的结果， AB 第 m 行第 n 列的元素是 A 第 m 行乘以 B 第 n 列的结果。其中，当将 A 的某一行抽出来乘以 B 的某一列时，我们是像下面这样乘的，

$$[a \quad b] \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd, \quad (1.41)$$

也就是要把行和列中对应的各元素相乘起来，并求和。

注意，当我们考察的不是无穷行无穷列的数据表时，通常我们只会考察行数等于列数的数据表，比方 N 行 N 列的数据表，通常称之为 $N \times N$ 矩阵。而且，必须要确保两个数据表(矩阵)的行列数相同(比方两者都是 $N \times N$)，我们才能够将它们进行运算，比如乘起来。我们不能够将一个 $N \times N$ 的数据表与一个 $M \times M$ (设 $N \neq M$)的数据表乘起来，因为在这种情况下乘法规则(1.38)无法定义(乘法规则要求 A 的列数与 B 的行数相同)。

不难验证跃迁元数据表的乘法满足结合律，具体来说，假设有三个跃迁元数据表 $[A]_{mn}$ 、 $[B]_{mn}$ 以及 $[C]_{mn}$ ，则

$$[(AB)C]_{mn} = \sum_{l,k} [A]_{ml} [B]_{lk} [C]_{kn} = [A(BC)]_{mn}. \quad (1.42)$$

但是，跃迁元数据表的乘法，或者说矩阵乘法，通常不满足交换律。为了看清楚这一点，不妨以两个两行两列的数据表为例。假设

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \quad (1.43)$$

则不难按照上面的乘法规则算出

$$AB = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad BA = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}. \quad (1.44)$$

很显然， $AB \neq BA$ 。所以，一般来说，跃迁元数据表的乘法并不满足乘法交换律！而如果两个跃迁元数据表 X 和 Y 乘法可交换，即满足 $XY = YX$ ，则我们就称它们对易，反之就称为不对易。

两个物理量除了可以相乘当然还可以相加，因此我们还要定义相应跃迁元数据表的加法，相应的定义是显然的，如下

$$[A + B]_{mn} = [A]_{mn} + [B]_{mn}. \quad (1.45)$$

不妨以 2×2 (两行两列)的数据表为例，说明这个加法是如何进行的，

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}. \quad (1.46)$$

有时候还需要将一个常数乘以一个物理量，因此还需要定义常数与跃迁元数据表的乘法，叫做数乘，相应的定义也是显然的，如下

$$[\lambda A]_{mn} = \lambda[A]_{mn}, \quad (1.47)$$

式中 λ 为某个常数。不妨以 2×2 的数据表为例进行具体说明，

$$\lambda \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{bmatrix}. \quad (1.48)$$

定义了加法、乘法，以及数乘，这样一个原子系统所有可能跃迁元数据表的集合就构成了一个数学上所谓的代数。问题是，与物理量对应的跃迁元数据表还需要额外满足海森堡的条件，即它们必须是厄米矩阵，但是正如稍后将会说明的，所有厄米矩阵的集合并不是乘法封闭的，也就是两

个厄米矩阵的乘积可能不是厄米矩阵。所以，为了数学上的方便性，我们常常扩大考察范围，即不仅仅只考察厄米的跃迁元数据表，而是考察所有数学上可能的跃迁元数据表，即使它不厄米，因为这样就可以形成一个乘法封闭代数。当然，真正物理的(厄米的)跃迁元数据表只是这个乘法封闭代数中的一个子集。

为了说明为什么厄米矩阵的集合并不乘法封闭，我们来考察两个矩阵 A 和 B ，它们不一定厄米，我们考察 AB 的厄米共轭矩阵 $(AB)^\dagger$ ，按照厄米共轭矩阵的定义，它应该满足

$$\begin{aligned} [(AB)^\dagger]_{mn} &= [AB]_{nm}^* = \left(\sum_l [A]_{nl} [B]_{lm} \right)^* = \sum_l [A]_{nl}^* [B]_{lm}^* \\ &= \sum_l [A^\dagger]_{ln} [B^\dagger]_{ml} = \sum_l [B^\dagger]_{ml} [A^\dagger]_{ln} = [B^\dagger A^\dagger]_{mn}, \end{aligned} \quad (1.49)$$

推导过程中我们多次使用了乘法规则。最终一头一尾的相等告诉我们

$$(AB)^\dagger = B^\dagger A^\dagger. \quad (1.50)$$

这是一个很有用的公式。下面我们假设 A, B 均为厄米矩阵，即满足 $A^\dagger = A$, $B^\dagger = B$ ，则利用上面的公式我们有 $(AB)^\dagger = B^\dagger A^\dagger = BA$ ，但是，由于矩阵乘法一般来说没有交换律，所以，一般来说， $BA \neq AB$ ，也即是 $(AB)^\dagger \neq AB$ ，从而一般来说 AB 不是厄米矩阵。

前面说过，矩阵(数据表)的乘法通常不满足交换律，因此两个矩阵 A, B 相乘，通常有 $AB \neq BA$ ，除非这两个矩阵对易。为了衡量两个矩阵的不对易程度，人们常常引入一个新的概念，叫做两个矩阵的对易子，记作 $[A, B]$ ，它也是一个矩阵，定义是

$$[A, B] \equiv AB - BA. \quad (1.51)$$

很显然，当且仅当 A, B 对易的时候，其对易子才等于零。根据定义显然有

$$[A, B] = -[B, A]. \quad (1.52)$$

很显然，任何矩阵都必定和自己对易，即 $[A, A] = 0$ 。另外，根据单位矩阵的性质，任何矩阵都必定和单位矩阵对易，即 $[A, 1] = 0$ 。

现在，假设有三个矩阵 A, B, C ，则有

$$[A, BC] = [A, B]C + B[A, C]. \quad (1.53)$$

为了证明这个结果，我们把等式左边按照定义展开，得

$$ABC - BCA, \quad (1.54)$$

另一方面，我们将等式右边按照定义展开，得

$$(AB)C - (BA)C + B(AC) - B(CA) = ABC - BCA, \quad (1.55)$$

推导的过程中我们用到了结合律(但是，请特别注意，你不能用交换律)，很显然等式右边这个最终结果与等式左边一样，所以原等式得证。类似的，还可以证明如下雅可比恒等式

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \quad (1.56)$$

证明过程大体如下

$$[A, [B, C]] = [A, BC - CB] = [A, BC] - [A, CB], \quad (1.57)$$

然后进一步利用上面的恒等式(1.53)，得到(详细过程请读者自己写出来)

$$[A, [B, C]] = [[A, B], C] + [B, [A, C]], \quad (1.58)$$

不难发现，这就等价于要证明的(1.56)式(请读者自己思考一下)。

对易子运算的一个好处是，若 A, B 均为厄米矩阵，则 $i[A, B]$ 也必定为厄米矩阵(请读者自己证明，用到 $(C + D)^\dagger = C^\dagger + D^\dagger$ ，以及 $(\lambda D)^\dagger = \lambda^* D^\dagger$)，所以，所有厄米矩阵的集合虽然在矩阵乘法运算下不封闭，但在求对易子再乘以虚数单位*i*的运算下是封闭的，进一步考虑加法和数乘(限制于乘以实数，厄米矩阵乘以实数依然得到厄米矩阵)，当然也还是运算封闭的(即对一些厄米矩阵进行这几种运算的结果依然是厄米矩阵)，因此厄米矩阵的集合在这几种运算下构成一个代数，而且由于求对易子的运算满足雅可比恒等式，求对易子再乘以*i*的运算当然也满足雅可比恒等式，所以数学家也常常称它为李代数(满足雅可比恒等式的代数就叫李代数)。由于厄米矩阵对应着物理可观测量的跃迁元数据表，所以也称这个李代数为物理可观测量的李代数，实际上，这就是为什么对易子运算在量子力学中这么重要的原因。

1.2.3 海森堡运动方程

经典物理里有一个最重要的量，那就是系统的哈密顿量 H ，也就是经典系统的能量(对哈密顿量相关内容的初步介绍，请读者参阅我的另一个简短讲义，《初识经典力学》)，等于动能加势能。但是现在，一方面系统的能量值量子化了，另一方面，作为一个物理量，哈密顿量 H 当然也要变成跃迁元数据表。与哈密顿量相应的跃迁元数据表应该是什么呢？根据海森堡的一般性想法，它当然应该是 $[H]_{mn} \equiv [H_0]_{mn} e^{-i\omega_{mn}t}$ ，其中 $[H_0]_{mn}$ 与时间无关。注意到 $\omega_{nn} = 0$ ，相应的 $e^{-i\omega_{nn}t} = 1$ ，所以这个跃迁元数据表应该是这样

$$H = \begin{bmatrix} [H_0]_{11} & [H_0]_{12}e^{-i\omega_{12}t} & [H_0]_{13}e^{-i\omega_{13}t} & \dots \\ [H_0]_{21}e^{-i\omega_{21}t} & [H_0]_{22} & [H_0]_{23}e^{-i\omega_{23}t} & \dots \\ [H_0]_{31}e^{-i\omega_{31}t} & [H_0]_{32}e^{-i\omega_{32}t} & [H_0]_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.59)$$

当然，它还要满足厄米条件，这里不写了。很显然，除了对角线上的元素(称作对角元)不依赖于时间之外，其它非对角元都依赖于时间。

现在一个关键的假设来了，我们知道，一个封闭的经典力学系统总能量是守恒的，其哈密顿量不会随时间演化。我们假设：量子化以后，能量守恒定律依然成立，因此封闭系统的跃迁元数据表不能依赖于时间。从而，根据这条假设，我们知道， H 非对角元的那些 $[H_0]_{mn}$, $m \neq n$ 必定都为零，即 $[H_0]_{mn} = 0$, $m \neq n$ ，因为否则， H 就会依赖于时间。所以哈密顿量的跃迁元数据表必定为

$$H = \begin{bmatrix} [H_0]_{11} & 0 & 0 & \dots \\ 0 & [H_0]_{22} & 0 & \dots \\ 0 & 0 & [H_0]_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.60)$$

这样的，只有对角元非零，其它元素都为零的矩阵称作对角矩阵。厄米条件进一步告诉我们 $[H_0]_{nn} = [H_0]_{nn}^*$ ，所以这些对角元必定都是实数！然而，在另一方面，我们知道系统的能量值量子化为了 E_n ，所以，最自然的可能性就是，这些对角元刚好就是这些量子化能量，即 $[H_0]_{nn} = E_n$ 。因此我们

知道，哈密顿量的跃迁元数据表最自然的可能性就是

$$H = \begin{bmatrix} E_1 & 0 & 0 & \dots \\ 0 & E_2 & 0 & \dots \\ 0 & 0 & E_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (1.61)$$

不妨称之为哈密顿矩阵，在不引起混淆的时候，也常常简称为哈密顿量。

我们知道，任意一个物理量 A 所对应的跃迁元数据表可以写成 $A = [A]_{mn} = [A_0]_{mn}e^{-i\omega_{mn}t}$ ，式中 $[A_0]_{mn}$ 与时间无关，而 $\omega_{mn} = (E_n - E_m)/\hbar$ 。很显然，数据表 A 是随时间演化的，现在我们想知道的是它满足什么演化微分方程。为此，我们把它对时间求导，得

$$\frac{d}{dt}[A]_{mn} = -i\omega_{mn}[A]_{mn} = \frac{-i}{\hbar}([A]_{mn}E_n - E_m[A]_{mn}). \quad (1.62)$$

排列成数据表的形式，即有

$$\begin{aligned} i\hbar \frac{dA}{dt} &= \begin{bmatrix} [A]_{11}E_1 & [A]_{12}E_2 & \dots \\ [A]_{21}E_1 & [A]_{22}E_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} - \begin{bmatrix} E_1[A]_{11} & E_1[A]_{12} & \dots \\ E_2[A]_{21} & E_2[A]_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \\ &= \begin{bmatrix} [A]_{11} & [A]_{12} & \dots \\ [A]_{21} & [A]_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} E_1 & 0 & \dots \\ 0 & E_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} - \begin{bmatrix} E_1 & 0 & \dots \\ 0 & E_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} [A]_{11} & [A]_{12} & \dots \\ [A]_{21} & [A]_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \end{aligned}$$

很明显，利用哈密顿矩阵的定义(1.61)，这个最终结果可以很整洁地写成

$$i\hbar \frac{dA}{dt} = [A, H]. \quad (1.63)$$

这就是著名的海森堡运动方程。

不难验证，海森堡运动方程与跃迁元数据表的乘法是相容的。即假设有两个物理量，其跃迁元数据表分别为 A 和 B ，它们均满足海森堡运动方程，则可以证明，两者的乘积 AB 也必定满足海森堡运动方程。证明如下(注意，推导的过程中不能使用乘法交换律)

$$i\hbar \frac{d}{dt}(AB) = i\hbar \left(\frac{dA}{dt}B + A \frac{dB}{dt} \right) = [A, H]B + A[B, H] = [AB, H], \quad (1.64)$$

证明完毕！其中第2个等于号用到了 A, B 均满足海森堡运动方程，最后一个等于号是用到了矩阵对易子的恒等式(1.53)。另外， A, B 的和 $A + B$ 满足海森堡运动方程更是容易证明的。进而不难得到，所有满足海森堡运动方程的跃迁元数据表的集合在加法、乘法，以及数乘运算下是封闭的，从而构成一个代数。

以上证明的一个推论是，若 A, B 均满足海森堡运动方程，则它们的对易子 $[A, B]$ 也必定满足海森堡运动方程。进而，所有满足海森堡运动方程的跃迁元数据表在求对易子的运算(再加上加法运算和数乘运算)下也构成一个封闭代数，由于求对易子运算满足雅可比恒等式，所以这也是一个李代数。显然，如果将数乘限制于乘以实数，将求对易子运算再乘以 i ，那所有物理可观测量的跃迁元矩阵(必然厄米且满足海森堡运动方程)的集合也是一个满足海森堡运动方程的李代数。

将海森堡运动方程中的 A 取成哈密顿量 H 本身，即有 $i\hbar \frac{dH}{dt} = [H, H] = 0$ ，很显然，不过是又一次验证了最终的方程会自动给出能量守恒。更一般的，如果一个物理量 A (所以满足厄米条件)与哈密顿量对易， $[A, H] = 0$ ，从而 $\frac{dA}{dt} = 0$ ，我们就称 A 为一个守恒量。显然，如果物理量 A, B 均为守恒量，则 AB 也必定为守恒量(虽然可能不满足厄米条件，从而不是真正的物理量)，同样， $i[A, B]$ 也必定为守恒量，且满足厄米条件。所以，一个量子系统所有物理守恒量的集合在求对易子再乘以虚数单位*i*等几种运算下也构成一个封闭的李代数。

作为物理量，核外电子位置矢量的跃迁元数据表**X**以及动量的跃迁元数据表**P**当然也满足海森堡运动方程，从而即有

$$i\hbar \frac{d\mathbf{X}}{dt} = [\mathbf{X}, H], \quad i\hbar \frac{d\mathbf{P}}{dt} = [\mathbf{P}, H]. \quad (1.65)$$

1.2.4 基本对易关系

下面要推导一个重要的结论，这个结论在一维系统中看得比较清楚。另外，下面我们要同时考虑经典的变量和相应的量子跃迁元数据表，所以，我们要再解释一下符号约定：我们以小写的 x 和 p 分别表示一维经典系统的位移变量和动量，因此， x 和 p 都是通常的数，不是矩阵，而相应的量子跃迁元数据表，我们分别记作大写的 X 和 P 。

对于这样的一维系统，方程(1.65)就变成了

$$i\hbar \frac{dX}{dt} = [X, H], \quad i\hbar \frac{dP}{dt} = [P, H]. \quad (1.66)$$

为了得到哈密顿矩阵的表达式，我们先写出经典系统的哈密顿量 $\frac{p^2}{2m_e} + V(x)$ ，式中 $V(x)$ 为势能函数， m_e 为粒子的质量。然后，我们将表达式中的 x, p 分别替换成相应的跃迁元矩阵，即 $x \rightarrow X, p \rightarrow P$ ，从而就得到

$$H = \frac{P^2}{2m_e} + V(X), \quad (1.67)$$

一个细节的问题是，由于现在 X 是跃迁元数据表，而不是普通的数，所以其函数 $V(X)$ 需要一个合适的定义，一种最简单的定义是，先把普通的函数 $V(x)$ 泰勒展开成如下幂级数

$$V(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots \quad (1.68)$$

再把 $x \rightarrow X$ ，就可以了，也即是说

$$V(X) = a_0 + a_1X + a_2X^2 + a_3X^3 + \dots, \quad (1.69)$$

当然常数项 a_0 应该理解成 a_0 乘以单位矩阵 1。很显然这个幂级数的每一项都只涉及到矩阵乘法，因此是有定义的。

现在，我们立即可以从方程(1.66)得出一个推论：即 X 和 P 不可能对易！具体的推理是这样的：首先 $[P, P] = 0$ ，从而 $[P, P^2] = 0$ ，这是因为 $[P, P^2] = [P, PP]$ ，利用恒等式(1.53)，即有 $[P, PP] = P[P, P] + [P, P]P = 0$ 。同样， $[X, X] = 0$ ，从而必有 $[X, V(X)] = 0$ （想想为什么）。类似的，如果 X 和 P 对易，则 X 与 P^2 也必定对易，因为根据恒等式(1.53)， $[X, P^2] = P[X, P] + [X, P]P$ 。另外，如果 P 和 X 对易，则 P 与 $V(X)$ 也必定对易（想想为什么）。但，如此一来（即假如 $[X, P] = 0$ ），那 X 就和 H 中的每一项都对易，从而 $[X, H] = 0$ ，同样 $[P, H] = 0$ ，那方程(1.66)就会给出 $\frac{dX}{dt} = 0, \frac{dP}{dt} = 0$ ，也即是说，任何一维系统的运动方程都是这个，而与系统的势能无关！这当然是不可能的！所以，必定有 $[X, P] \neq 0$ 。

关键是 $[X, P]$ 等于什么？不妨暂记

$$[X, P] = C. \quad (1.70)$$

所以，关键是要导出矩阵 C 具体是什么？

为了找到这个答案，我们先回到熟悉的经典系统，这时候相应的方程当然不是海森堡运动方程，而是哈密顿正则方程(请参阅我的《初识经典力学》)，

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} = \frac{p}{m_e}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x} = -V'(x), \quad (1.71)$$

式中我们代入了经典的 $H = \frac{p^2}{2m_e} + V(x)$, $V'(x)$ 表示 $V(x)$ 对自变量 x 的导数，用上面的幂级数也就是

$$V'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots \quad (1.72)$$

下面我们作一个重要的假设：即量子系统要和相应的经典系统能够对应得上！换言之，方程(1.66)最终给出的结果，应该等价于将上面经典系统运动方程做 $x \rightarrow X, p \rightarrow P$ 替换之后的结果，也就是应该等价于

$$\frac{dX}{dt} = \frac{P}{m_e}, \quad \frac{dP}{dt} = -V'(X). \quad (1.73)$$

下面我们来看(1.66)式的第一步，由于 $[X, V(X)] = 0$, 所以这个式子告诉我们

$$\begin{aligned} \frac{dX}{dt} &= \frac{1}{i\hbar}[X, H] = \frac{1}{i\hbar}\left[X, \frac{P^2}{2m_e}\right] = \frac{1}{i\hbar}\frac{1}{2m_e}([X, P]P + P[X, P]) \\ &= \frac{1}{m_e}\frac{1}{i\hbar}\frac{1}{2}(CP + PC). \end{aligned} \quad (1.74)$$

而我们想要的结果是 $\frac{dX}{dt} = \frac{1}{m_e}P$, 不难看出，为了达到这个想要的结果，最简单的就是取 $C = i\hbar \cdot 1$, 式中1当然表示单位矩阵，常常省略。所以我们试探性地取

$$[X, P] = i\hbar. \quad (1.75)$$

下面来看这种取法能否使得(1.66)式的第二个式子也同时达到我们想要的结果，也就是等价于 $\frac{dP}{dt} = -V'(X)$ 。为此，我们注意到 $[P, P^2] = 0$, 所以，(1.66)式的第二个式子告诉我们

$$\frac{dP}{dt} = \frac{1}{i\hbar}[P, H] = \frac{1}{i\hbar}[P, V(X)], \quad (1.76)$$

为了计算 $[P, V(X)]$, 我们代入 $V(X)$ 的级数展开, 从而

$$[P, V(X)] = a_1[P, X] + a_2[P, X^2] + a_3[P, X^3] + \dots \quad (1.77)$$

不断利用恒等式(1.53), 以及我们的假设 $[P, X] = -[X, P] = -i\hbar$, 不难计算出结果为(具体计算留给读者作练习)

$$[P, V(X)] = -i\hbar(a_1 + 2a_2X + 3a_3X^2 + \dots) = -i\hbar V'(X). \quad (1.78)$$

代入(1.76)式, 即有

$$\frac{dP}{dt} = -V'(X). \quad (1.79)$$

正是我们想要的结果。

所以, (1.75)式的确是正确的! 不过, 我们不妨进行最后一个检验, 即注意到(1.75)式右边是一个与时间无关的常数, 而左边的 X, P 都是随时间演化的, 所以我们还要检验左边求对易子的结果不随时间演化。这的确是成立的, 验证如下

$$\begin{aligned} \frac{d}{dt}([X, P]) &= \dot{X}P + X\dot{P} - \dot{P}X - P\dot{X} \\ &= (PP/m_e - XV'(X) + V'(X)X - PP/m_e) = 0. \end{aligned} \quad (1.80)$$

式中第二行的等于号是代入了方程(1.73), 而最后的等号是利用了 $[X, V'(X)] = 0$ 。

以上推导过程也说明了一件事情, 即海森堡运动方程(1.66)就是经典哈密顿力学哈密顿正则方程的量子对应物, 如果读者学习过经典哈密顿力学的泊松括号运算的话, 将会进一步看清两者的对应, 大体来说就是经典的泊松括号运算, 在量子里面就对应求对易子运算。当然, 这不是我们想强调的内容, 因此就不再展开了。

以上得到的结果(1.75)在量子力学中有基本的重要性, 所以也叫做量子力学基本对易关系, 在历史上, 它最早是玻恩等人得到的。实际上在我看来, 海森堡、玻恩和若尔当三个人对这个结果都有重要的贡献, 海森堡的原始论文实际上暗含了推导的方向, 当然, 最终确立这个结果的是玻恩和若尔当。不过, 玻恩等人用的不是我们这里给出的推导方法。

以上考察的是一个自由度的一维问题，下面把重要结果(1.75)推广到三个自由度的三维情形。这时候按照之前的符号约定，大写的 \mathbf{X} 和 \mathbf{P} 分别代表粒子的位置和动量的跃迁元矩阵。当然，它们不仅是矩阵，同时还是三维空间的向量，有三个分量，因此实际上可以记作

$$\mathbf{X} = (X_1, X_2, X_3), \quad \mathbf{P} = (P_1, P_2, P_3). \quad (1.81)$$

当然，每一个分量的位置 X_a 和动量 P_a ($a = 1, 2, 3$)本身都是跃迁元矩阵。

注意到运动的不同自由度是相互独立的，也就是说，不同自由度的这些基本变量之间应该和普通数的乘法一样，是相互对易的，因此有，当 $a \neq b$ 时， $[X_a, X_b] = 0$, $[P_a, P_b] = 0$, 以及 $[X_a, P_b] = 0$ 。另外，上面对单自由度的讨论又告诉我们，对于同一个自由度， $[X_a, P_a] = i\hbar$, $[X_a, X_a] = 0$, $[P_a, P_a] = 0$ 。综上，我们可以把结果概括如下

$$[X_a, P_b] = i\hbar\delta_{ab}, \quad [X_a, X_b] = 0, \quad [P_a, P_b] = 0, \quad a, b = 1, 2, 3 \quad (1.82)$$

式中 δ_{ab} 就是克龙内克符号，当 $a = b$ 时取1，当 $a \neq b$ 时取0。由于任何物理可观测量都可以用一些 X_a 和一些 P_a 通过数乘、加法以及矩阵乘法生成(当然，同时要满足厄米条件)，因此任何两个物理可观测量进行求对易子再乘以*i*运算的结果都可以通过上面的基本结果(1.82)计算出来，我们称(1.82)式生成了物理可观测量的李代数。通常称这个李代数为海森堡代数。这个结果也可以推广到任意的*n*自由度量子系统，形式上这个推广很简单，只需要让上面的指标*a, b*从1取到*n*(而不是只取1, 2, 3)就可以了。

1.2.5 求解谐振子

然而，这一切对吗？海森堡的想法虽然并不牵强，但是，他真是在玩一个疯狂的游戏，经典物理里那些习以为常的物理量怎么就变成无穷大的矩阵了呢？怎么它们就要按照那种奇怪的乘法乘起来呢？是的，对于没有学过相关数学知识的物理学家而言，海森堡发明的乘法虽然并不牵强，但它可太奇怪了！这就是具有最高创造性的工作的一个特征，很自然，同时也很疯狂。

然而，这一切正确吗？

最直接的检验当然就是把它应用到氢原子上，然后解出氢原子能级，也就是哈密顿矩阵(正如我们讲过的，在到此为止的处理中，它都必然是一个对角矩阵)的对角元 E_n ，甚至解出每一个 $[x]_{mn}$ ，进而算出谱线的强度。然而，这在数学上太难了一点，海森堡解不出来，这个问题是稍后泡利解决的。

因此海森堡想到将他的思想用到一个稍微简单一点的问题上，也就是线性谐振子上。也即是说，假设发出电磁辐射的是一个带电线性谐振子(而不是更复杂的氢原子)，一个线性谐振子从一个定态能级跃迁到另外一个定态能级从而辐射出电磁波。根据普朗克对黑体辐射公式的推导，线性谐振子的能量应该是量子化的，而且其量子化能级具有 $E_n = n\hbar\omega$ 这样的形式。但是，普朗克是直接假设了这个结果(非如此假设不足以解释黑体辐射的实验曲线)，他可没有具体推导出这个结果，假设海森堡能用它的矩阵想法推导出这个，那不就是他理论正确的强烈证据吗？

当时的海森堡的确做了这件事情，并且他发现，结果真是正确的！

下面我们来重复海森堡的工作。我们要考察的是单自由度的所谓一维谐振子，其哈密顿矩阵为

$$H = \frac{P^2}{2m_e} + \frac{1}{2}m_e\omega^2 X^2. \quad (1.83)$$

其中 $\frac{1}{2}m_e\omega^2 X^2$ 为谐振子的势能(弹簧振子的势能)。由运动方程(1.73)，可以得到

$$\dot{X} = P/m_e, \quad \dot{P} = -m_e\omega^2 X, \quad (1.84)$$

两式联立，即有 $\ddot{X} + \omega^2 X = 0$ ，代入 $X = [X_0]_{mn}e^{-i\omega_{mn}t}$ ，即有

$$(-\omega_{mn}^2 + \omega^2)[X_0]_{mn} = 0. \quad (1.85)$$

以上结果表明，仅当 $\omega_{mn} = \pm\omega$ 时， $[X_0]_{mn}$ 才不为零，即除了两个矩阵元，其它矩阵元均为零，对定态进行适当编号，从而可设 $\omega_{m,m\pm 1} = \pm\omega$ ，进而 $[X_0]_{m,m\pm 1}$ 非零，其余矩阵元都为零。

由量子力学基本对易关系 $XP - PX = i\hbar$ ，知 $m_e(X\dot{X} - \dot{X}X) = i\hbar$ ，即

$$m_e([X\dot{X}]_{mn} - [\dot{X}X]_{mn}) = i\hbar\delta_{mn}. \quad (1.86)$$

当 $n = m$ 时，利用矩阵乘法规则，即有

$$\begin{aligned} m_e \sum_l (-[X_0]_{ml}\omega_{lm}[X_0]_{lm} + \omega_{ml}[X_0]_{ml}[X_0]_{lm}) &= \hbar \\ \Rightarrow 2m_e \sum_l (\omega_{ml}[X_0]_{ml}[X_0]_{lm}) &= \hbar. \end{aligned} \quad (1.87)$$

利用厄米条件 $[X_0]_{lm} = [X_0]_{ml}^*$ ，即有

$$2m_e \sum_l (\omega_{ml}|[X_0]_{ml}|^2) = \hbar. \quad (1.88)$$

但是我们知道，仅仅当 $l = m \pm 1$ 时， $[X_0]_{ml}$ 才非零，所以上式的求和中，仅仅只有这两项有贡献，从而

$$2m_e\omega(|[X_0]_{m,m+1}|^2 - |[X_0]_{m,m-1}|^2) = \hbar. \quad (1.89)$$

注意到 $[X_0]_{m,m+1} = [X_0]_{m+1,m}^*$ ，也即是说，我们有递推公式

$$|[X_0]_{m+1,m}|^2 = |[X_0]_{m,m-1}|^2 + \hbar/(2m_e\omega). \quad (1.90)$$

假设把最低能级编号为第0能级(虽然前文的习惯是把最低能级编码为第1能级)，由于不存在-1能级，因此当然有 $[X_0]_{0,-1} = 0$ 。进而上面的递推公式就给出

$$|[X_0]_{m+1,m}|^2 = (m+1)\hbar/(2m_e\omega). \quad (1.91)$$

通常可以进一步取 $[X_0]_{m+1,m}$ 为实数，进而即有 $[X_0]_{m+1,m} = [X_0]_{m,m+1} = \sqrt{m+1}\sqrt{\frac{\hbar}{2m_e\omega}}$ 。不妨将 $[X_0]_{mn}$ 的矩阵具体排出来(注意， m, n 从零开始数的)，即是

$$[X_0]_{mn} = \sqrt{\frac{\hbar}{2m_e\omega}} \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (1.92)$$

谐振子的能级就是哈密顿矩阵的对角元，从而

$$\begin{aligned} E_n &= [H]_{nn} = \frac{1}{2}m_e([\dot{X}^2]_{nn} + \omega^2[X^2]_{nn}) \\ &= \frac{1}{2}m_e\left(\sum_l[\dot{X}]_{nl}[\dot{X}]_{ln} + \omega^2\sum_l[X]_{nl}[X]_{ln}\right). \end{aligned} \quad (1.93)$$

注意到 $[\dot{X}]_{nl} = -i\omega_{nl}[X]_{nl}$ 以及厄米条件，并注意到 $|[X]_{mn}|^2 = |[X_0]_{mn}|^2$ ，即有

$$\begin{aligned} E_n &= \frac{1}{2}m_e\sum_l(\omega_{nl}^2 + \omega^2)|[X]_{nl}|^2 \\ &= m_e\omega^2(|[X_0]_{n,n+1}|^2 + |[X_0]_{n,n-1}|^2) \\ &= (n + \frac{1}{2})\hbar\omega. \end{aligned} \quad (1.94)$$

上式最后一行我们代入了(1.91)式。

所以，根据海森堡的推导，线性谐振子的能级 $E_n = (n + \frac{1}{2})\hbar\omega = n\hbar\omega + \frac{1}{2}\hbar\omega, n = 0, 1, 2, \dots$ 与普朗克的量子化假设是完全吻合的。实际上，相比于普朗克的假设，结果只多了一个整体的常数 $\frac{1}{2}\hbar\omega$ ，而定态跃迁辐射光的过程只涉及能级差，因此这个整体常数当然是普朗克无从预料的。所以，海森堡知道，他对了！

1.2.6 符号的新发展

跃迁元的概念，以及跃迁元的乘法规则(1.38)当然是海森堡引领的这场矩阵革命中最具创造性，也是最核心的成果。这一小节，我们再次回到它们，我们将用一种新的符号来书写它们，这种新的符号是狄拉克所引入的。你可能觉得，只不过是一种不同的记号而已，不是什么本质的东西，然而情况并非如此，不同的符号系统有不同的表达能力，事实证明，狄拉克所引入的这种新记号在量子力学后来的发展中有更加强大的表达能力。

为了引入这一新的符号系统，让我们考察跃迁元矩阵 $[A]_{mn}$ ，它表示在物理量 A 的微扰之下，系统从定态可能性(n)(还记得吧，前面说过，定态就是系统的一种特定可能性)跃迁到定态可能性(m)的跃迁元，根据前文说过的，这个跃迁元可以示意性地记作

$$[A]_{mn} = \left\{ (m) \xleftarrow{A} (n) \right\}. \quad (1.95)$$

现在，我们把这个示意性的符号再简化一下，使之成为一种更好用的正式记号，我们将跃迁元 $[A]_{mn}$ 重记为

$$[A]_{mn} = \left\{ (m) \xleftarrow{A} (n) \right\} = \langle m | \hat{A} | n \rangle, \quad (1.96)$$

$\langle m | \hat{A} | n \rangle$ 就是这个新的正式记号，它的物理意义还是和 $[A]_{mn}$ 一样，只不过记号变了。现在，我们用 $|n\rangle$ 来表示初始的定态可能性，用 $\langle m|$ 来表示末尾的定态可能性，用加了一个“帽子”的 \hat{A} 来表示物理量 A 对系统的作用(通过微扰引入)，注意，新记号区分了初态可能性和末态可能性，而且我们将初态可能性写在了 \hat{A} 的右边，将末态可能性写在了它的左边。因此，跃迁元 $[A]_{mn}$ 就重记成了 $\langle m | \hat{A} | n \rangle$ ，读作系统从初态 $|n\rangle$ 经过物理量 A 的作用跃迁到末态 $|m\rangle$ ，注意，从右往左读的。

用这种新的记号，厄米共轭矩阵的定义 $[M^\dagger]_{mn} = [M]_{nm}^*$ 就应该重写为

$$\langle m | \hat{M}^\dagger | n \rangle = \langle n | \hat{M} | m \rangle^*. \quad (1.97)$$

而物理量的厄米条件 $[A]_{mn} = [A]_{nm}^*$ 就应该重写为

$$\langle m | \hat{A} | n \rangle = \langle n | \hat{A} | m \rangle^*. \quad (1.98)$$

它的含义当然还是，交换初末态之后反向的从 (n) 到 (m) 的跃迁元，等于正向的从 (m) 到 (n) 的跃迁元的复数共轭，新记号由于明显地区分了初末态，所以这个含义就更明显了。

按照这种新的记号，跃迁元的乘法规则(1.38)就应该重写为

$$\langle m | \hat{A} \hat{B} | n \rangle = \sum_l \langle m | \hat{A} | l \rangle \langle l | \hat{B} | n \rangle. \quad (1.99)$$

它表示，系统在物理量 AB 的影响下从 (n) 跃迁到 (m) 的跃迁元，等于系统先在 B 的影响下从 (n) 跃迁到中间可能性 (l) 的跃迁元，再乘以，在 A 的影响下从 (l) 跃迁到 (m) 的跃迁元，注意，对于 B 来说， (l) 是一个末态可能性，但是，对于 A 来说， (l) 是一个初态可能性。当然，由于中间可能性不能确定，所以在乘的时候我们要把所有的中间可能性都平等地加起来。值得说明的是，从上面的解释可以清楚看到，这个式子的正确读法也是倒着的，即从最右边往左边读。

特别的，按照新的记号，物理量1的跃迁元 $[1]_{mn}$ 就应该记作 $\langle m|1|n\rangle$ ，不过，由于1的意思就是不受任何微扰，所以在书写中常常省略，就把这个跃迁元记作 $\langle m|n\rangle$ ，它表示系统在不受微扰时从可能性 (n) 跃迁到 (m) 的跃迁元。前面我们说过， $[1]_{mn} = \delta_{mn}$ ，所以，用新的记号即是

$$\langle m|n\rangle = \delta_{mn}, \quad (1.100)$$

它表示系统在不受微扰时从 (n) 跃迁到 (n) 自身的概率是1，而跃迁到其它可能性的概率都是零。我们称这种性质为：**系统的任何两个不同定态可能性都可以确定地相区分！**因此，系统定态可能性的完备集(即所有定态可能性的集合)是一个可确定区分可能性完备集。

本章所讨论的跃迁元都是定态可能性之间的跃迁元，非常依赖于定态这一概念，这实际上是一个很大的局限，下一章我们将从这种局限中获得解放。

第二章 量子力学基本原理

上一章按照历史发展，从定态和定态的量子跃迁出发，通过引入跃迁元和跃迁元乘法规则，最终建立起了量子力学的初步理论。但是，这里有一个巨大的局限，就是我们的一切都基础于定态可能性，而定态虽然重要，但却是比较特殊的，它限制了量子力学理论的灵活应用。这一章我们将从这种局限中解放出来，获得更大的自由。

熟悉态矢量、希尔伯特空间、量子态内积以及算符等概念的读者，请先假装忘记这些东西，它们对于本章的展开是不需要的，反而会干扰我们理解本章的概念体系。作为本章的自然推论，这些东西会是下一章的主题。

不过，在具体开始本章之前，还是让我们重写一下上一章得到的跃迁元乘法规则，任给物理量 A, B ，按照上一章最后介绍的新记号，我们有

$$\langle E_m | \hat{A} \hat{B} | E_n \rangle = \sum_l \langle E_m | \hat{A} | E_l \rangle \langle E_l | \hat{B} | E_n \rangle. \quad (2.1)$$

式中物理量符号加一个帽子 $\hat{\cdot}$ 表示此物理量对系统的作用，它由一个正比于这个物理量的微扰产生。相对于上一章，我们在记号上作了一点小改动，我们用 (E_n) 而不是 (n) 来表示定态可能性，它表示的就是能量取确定的量子化值 E_n 的这么一个定态。

从上一章我们知道，一个量子系统所有定态可能性的集合构成一个可确定区分可能性完备集，不妨记作 \mathcal{E} ，其中的可能性满足如下可确定区分性

$$\langle E_m | E_n \rangle = \delta_{mn}. \quad (2.2)$$

式中 $\langle E_m | E_n \rangle$ 表示物理量 1 的跃迁元，即 $\langle E_m | E_n \rangle = \langle E_m | 1 | E_n \rangle$ 。

当然，作为物理可观测量， A, B 的跃迁元均满足厄米条件，以物理量 A 为例，

$$\langle E_m | \hat{A} | E_n \rangle = \langle E_n | \hat{A} | E_m \rangle^*. \quad (2.3)$$

或者示意性地记作

$$\left\{ (E_m) \xleftarrow{A} (E_n) \right\} = \left\{ (E_m) \xrightarrow{A} (E_n) \right\}^*. \quad (2.4)$$

表示从 (E_n) 到 (E_m) 的跃迁元等于从 (E_m) 到 (E_n) 跃迁元的复数共轭。不过，正如上一章说过的，虽然 A, B 均是物理量，均满足厄米条件，但它们的乘积 AB 可不一定满足厄米条件。

本章我们将一般性地定义可确定区分可能性完备集的概念，并进一步推广跃迁元，以及推广跃迁元的乘法规则。本章开始于对一类物理学实验的考察，即所谓的施特恩-格拉赫实验。原始版本的施特恩-格拉赫实验的确是物理学家真正做过而且重要的物理学实验，不过，本章考察的是它的一种改版，这是费曼在其物理学讲义第三卷中引入的，可能并没有物理学家真正做过这个费曼版的施特恩-格拉赫实验，因此，本章的考察实际上是对一种假想实验的考察。之所以我们相信这种对假想实验的考察，是因为我们相信量子力学的基本原理，比如本章将会着重讨论的不确定性和不可区分性原理，这些原理在其它真实实验中是经受过检验的。实际上，本章就是想通过假想实验来说清楚量子力学的基本原理。

2.1 可确定区分可能性完备集

在量子力学中，除了能量的量子化以外，另一种同样重要的量子化是所谓的角动量量子化，它最早由玻尔在其氢原子模型中作为假设提出。不过，玻尔的这条假设是半量子半经典的，真正完全符合量子力学原理的角动量量子化则要复杂一些，本节将以一个例子对此展开讨论。

我们要讨论的例子是核外电子总角动量为1的原子，角动量为1只是通常的一种简化说法，其具体意思我们马上就会讨论。由于核外电子带电，其角动量就会产生一个磁矩，可以和磁场耦合，因此将这种原子放到一个外磁场中，它就会受到磁场的作用，如果这外磁场是一个非均匀磁场，那

么原子就会受到一个作用力。这个作用力和原子角动量在磁场取向方向上的分量，以及磁场梯度的方向都有关系。施特恩和格拉赫的著名实验发现，如果把一束这样的非极化原子(所谓的非极化原子，就是每个原子的总角动量取向完全随机的一束原子)沿着水平方向通过如图(2.1)所示的一个装置，它就会分裂成分立的三束(而不是连续的一片)。装置沿着水平方向(称作 y 方向)放置，装置内有一个指向 z 方向的非均匀磁场，并且磁场的梯度也是沿着 z 方向。

对这种分裂成量子化的分立三束的解释是：原子角动量的 z 分量 J_z (也就是沿着磁场方向的分量)有三种量子化的可能性，不妨分别记作 $+, 0, -$ 吧(更通常的说法是， J_z 量子化为 $+\hbar, 0, -\hbar$)，由于磁场存在梯度，原子会受到一个 z 方向的作用力，但是不同的量子化可能性受到的力大小不同，所以通过装置以后，原子就分裂成了三束。

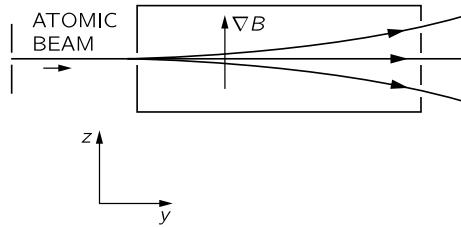


图 2.1: 在施特恩-格拉赫实验中，角动量为1的原子会分裂成三束。

当然，原子除了有角动量以外，还有动量等等物理量，但是，从现在开始，假设我们只关心原子的角动量，而忽略其动量等其它一切信息。这时候，我们就可以说， z 方向出现的 $+, 0, -$ 这三种可能性，构成了原子的一个可确定区分可能性完备集，所谓完备，意思就是，原子束分裂的时候只有这三种可能性，只会分裂成三束，再没有其它可能性了，而所谓可确定地区分，当然是因为这三种可能性的角动量 J_z 值不同，因此在分裂成理想的三束的情况下可以百分之百地区分它们。这里我们看到，除了上一章着重研究的定态可能性完备集以外，原子还可以有其他可确定区分可能性完备集，比如这里沿着 z 方向的 $+, 0, -$ (假设只关心角动量)。

实验上怎么检验这三种可能性是否真的可确定地区分呢？为了更方便研究这个问题，以及类似这样的其它一些问题，下面我们修改一下实验装置。图(2.2)(a)就是改版之后的实验装置简图，它包括一连串三个非均匀磁

场，左边的第一个就是通常的施特恩-格拉赫装置的磁铁，它把一束入射的原子分裂成三束，第二个磁铁和第一个磁铁类似，但是磁场方向和第一个相反，因此它施加在原子上的力也是反向的，会倾向于使得分裂的原子束反向运动(即向着轴线方向运动)，并且我们设想这第二个磁铁比较长，比方说是第一个磁铁的两倍那么长，第三个磁铁和第一个完全一样，它的作用是使得分裂的原子束在从装置中出来的时候恰好重新汇聚在一起。最终，原子束的运动路径如图(2.2)(b)所示。

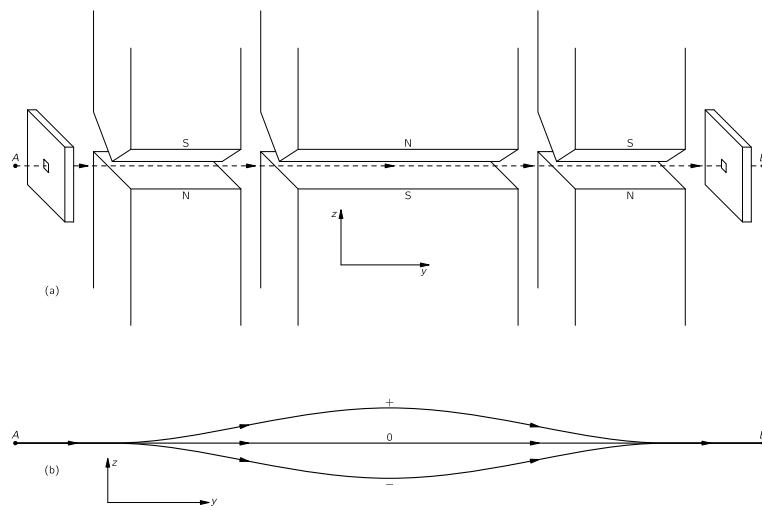


图 2.2: (a)改版之后的施特恩-格拉赫实验装置。(b)角动量为1的原子的路径。

我们设想原子束是从图(2.2)(a)左端的A点由静止开始运动的，经过左端的准直孔入射到装置中，最终从右端的准直孔出射，然后在某种作用下减速，并停止在右端的B点。由于最开始和最终了原子都是静止的，所以这就使得我们可以忽略对原子运动的关注，而把注意力集中在原子的角动量上。

我们把一台如图(2.2)(a)这样的，磁场在 z 方向上的装置称作装置 S ，相应的可能性完备集记作 $\{(iS), i = +, 0, -\}$ 。后文我们还会考察旋转一个适当角度(比如绕着水平 y 轴旋转一个适当角度)之后的同样装置，我们将称旋转以后的装置为装置 T ，原子通过它的时候当然也会在磁场方向(旋转后的)上分裂成三束，相应的可能性完备集记作 $\{(iT), i = +, 0, -\}$ 。

为了检验 $+, 0, -$ 三种可能性是否真的可确定地区分，我们需要做一个如图(2.3)所示的实验。我们要把两台 S 装置串联起来，使通过第一台装置的原子束接着通过第二台装置。我们在第一台 S 装置里面放入挡板，挡去三条分裂原子束中的两条，以确定从第一台装置中出来的原子束是三种可能性中的哪一种，比如挡掉了 $0, -$ 可能性，那我们就能确定从第一台装置出来的原子束处于 $+$ 可能性。我们再让从第一台装置出来的原子束接着通过第二台一样的 S 装置，然后看原子束可以从原来的可能性跃迁到第二台装置的哪种可能性，由此检验三种可能性的可确定区分性。比如像

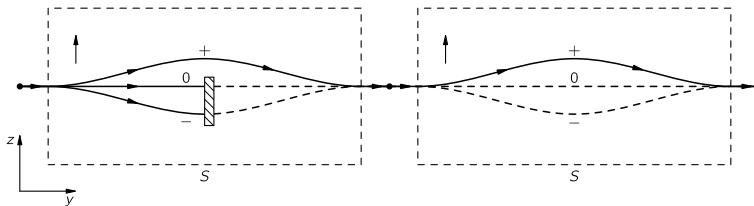


图 2.3: 串联两台 S 装置。

图(2.3)中所示的情形，在第一台装置中，通过利用挡板，过滤出了 $+$ 可能性的原子，因此再通过第二台装置的时候，这些原子就只会走最上面那条路线。而如果第一台装置的挡板过滤出来的是 0 可能性的原子，那再通过第二台装置时，它们就只会走中间的线路。同样，如果第一台装置过滤出来的是 $-$ 可能性原子，那它们再通过第二台装置时，就只会走最下面那条线路。而像图(2.3)中第一台装置那样过滤出一种确定可能性的装置，我们就叫它过滤器。

为了说得更具体一点，我们仿照费曼将改版以后的施特恩-格拉赫装置示意性地记作

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S , \quad (2.5)$$

下标 S 表示这是一台 S 装置。进一步，我们在这个记号中引入 \parallel ，表示挡板。比如我们挡住 S 装置的 $0, -$ 两种可能性(从而构成一台 $+S$ 过滤器)，那就记

作

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S . \quad (2.6)$$

图(2.4)总结了这个示意性记号的各种可能性。如果我们将两台装置串联，

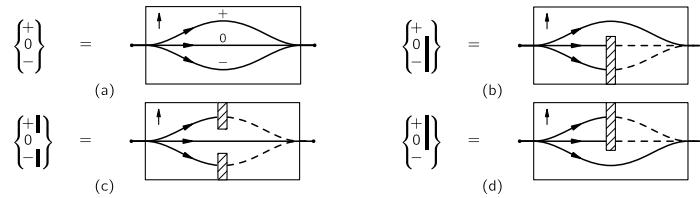


图 2.4: 施特恩-格拉赫装置的示意性记号.

那我们就把两个这样的示意性记号紧接在一起，像这样

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S . \quad (2.7)$$

注意，在这种写法中，原子束依然从左边往右边通行的，也就是先通过左边的装置，接着再通过右边的装置。

下面我们来讨论 $+, 0, -$ 三种可能性可确定区分性的检验。首先，对于下面这种情况，

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S , \quad (2.8)$$

从上面的讨论可知，这时必定有，从第一台装置通出的原子，最终也从第二台装置通出的概率为百分之百。这时候我们说，原子从第一台装置的 $(+S)$ 可能性跃迁到第二台装置 $(+S)$ 可能性的概率为百分之百，由于这种跃迁不是在任何微扰作用下发生的，所以对应物理量1的跃迁元(当然，这里跃迁和跃迁元的概念都被拓展了)，因此我们就可以把这个结果记成

$$\langle +S | +S \rangle = \langle +S | 1 | +S \rangle = 1. \quad (2.9)$$

和第一章中一样，物理量 1 常常从记号中省略。类似的，当然也有

$$\langle 0S|0S\rangle = \langle -S| -S\rangle = 1. \quad (2.10)$$

其次，对于下面的情况，以及其它类似的情况，

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S, \quad \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S, \quad (2.11)$$

很显然，最终将没有任何原子通出来，也即是说，这时候原子从第一台装置确定下来的可能性跃迁到第二台装置确立的可能性的概率为零。

综上，即有

$$\langle iS|jS\rangle = \delta_{ij}, \quad i, j = +, 0, -. \quad (2.12)$$

式中，右边的 $|jS\rangle$ 叫做初态可能性，左边的 $\langle iS|$ 叫做末态可能性。类似的分析不仅适用于 S 装置，也适用于任何 T 装置，即对任何 T 装置，同样有

$$\langle iT|jT\rangle = \delta_{ij}, \quad i, j = +, 0, -. \quad (2.13)$$

很显然，这两个式子是第一章中描述定态可能性可确定区分性的(2.2)式的推广。

可确定区分可能性完备集

现在，我们可以把可确定区分可能性完备集的概念推广到任意的量子系统。

首先，在量子力学的世界中，一个系统有多种可能性，我们将系统的一组可以相互确定地区分的可能性完备集记作 $\mathcal{C} = \{(i), i = 1, 2, 3, \dots\}$ ¹。所谓可能性的完备集，我们是指，在确保能收集到这些可能性所列举的信息的同时，你不可能再额外收集到系统更多的信息，因此，只要你采用适当

¹这里我们假设这些可能性是分立的，量子化的，当然，可确定区分可能性完备集的概念可以推广到连续的非量子化的可能性，但这在数学表述上会带来一点复杂性，因此为了理论表述的简洁性，我们限制于考虑分立的可能性。

的方式观测这个系统，你总会发现它一定处在这些可能性中的某一个，从某种意义上来说，这一完备集穷尽了你的观测所能收集到的所有可能性。

而所谓可以确定地区分，是指从 \mathcal{C} 的任意一种可能性出发，在不受任何微扰作用的情况下，系统只能百分之百地跃迁到这种可能性本身，跃迁到 \mathcal{C} 中任何别的不同可能性的概率都为零，也即满足

$$\langle j|i \rangle = \delta_{ij}, \quad i, j \in \mathcal{C}. \quad (2.14)$$

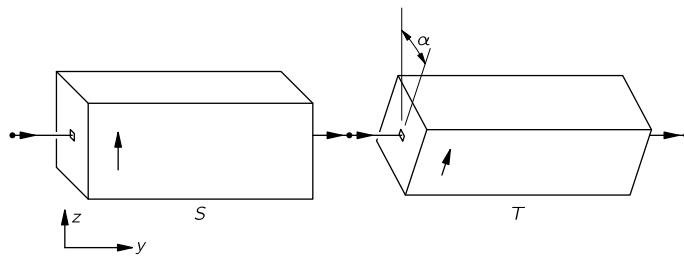
换言之，假设一个系统处于 \mathcal{C} 中的某些可能性，让你检测，而你检测到系统处于可能性 (j) ，那你就能百分之百地断言，检测之前系统就处于可能性 (j) ，因为它不可能从一个不同的可能性跃迁到 (j) 从而让你检测到。因此，通过检测，你总是可以100%地将 \mathcal{C} 的这些可能性中的任何两个区分开来。

比如上面串联 S 装置中的第一台，就是通过引入适当的挡板过滤，从而制备出处于可能性集 $\{(iS), i = +, 0, -\}$ 中的原子，而第二台 S 装置就是对这些可能性进行检测。通过检测，我们能百分之百地知道第一台装置制备出来的是哪种可能性。

当然，第一章讨论的原子的定态可能性完备集就是这样一个可确定区分可能性完备集。再比方说，在仅关心原子角动量时，一个角动量为1的原子相对于 z 方向的 $+, 0, -$ 三种可能性也构成了一个可确定区分可能性完备集。因此可确定区分可能性完备集可能是一个无限集，比如上一章的定态可能性完备集，但也可能是有限集，比如本节的 $\{+, 0, -\}$ 。而且，一个系统的可确定区分可能性完备集不是唯一的，而是有无穷多组，比方说本节的例子，对应装置 S 的 $\{(iS), i = +, 0, -\}$ 是一个可确定区分可能性完备集，但是对于任何一个装置 T 的 $\{(iT), i = +, 0, -\}$ 也都是可确定区分可能性完备集。

2.2 玻恩定则，不确定性与不可区分性

回到改版的施特恩-格拉赫实验，现在，假设我们把一台 S 装置和一台 T 装置串联起来，那会发生什么？为了明确起见，我们假设所考虑的 T 装置相对于原来的 S 装置是绕着水平 y 轴旋转了一个适当角度 α ，如图(2.5)所示。

图 2.5: 串联一台 S 装置和一台 T 装置。

假设我们做下面三组实验

$$\begin{Bmatrix} + \\ 0\parallel \\ -\parallel \end{Bmatrix}_S \begin{Bmatrix} + \\ 0\parallel \\ -\parallel \end{Bmatrix}_T, \quad \begin{Bmatrix} + \\ 0\parallel \\ -\parallel \end{Bmatrix}_S \begin{Bmatrix} +\parallel \\ 0 \\ -\parallel \end{Bmatrix}_T, \quad \begin{Bmatrix} + \\ 0\parallel \\ -\parallel \end{Bmatrix}_S \begin{Bmatrix} +\parallel \\ 0\parallel \\ - \end{Bmatrix}_T. \quad (2.15)$$

我们会发现，每一组实验都有原子从整个装置的右端出来，即使最初每次从左端通进来的只有一个原子，在每一组实验中，它从右端出来的概率也都不是零。这就说明，给定一个 $+S$ 过滤器过滤出来的确定处于 $+S$ 可能性的原子，它跃迁到 $+T, 0T, -T$ 这三种可能性的概率都不是零！

问题是相应的跃迁概率该怎么计算呢？为了回答这个问题，我们需要推广一下物理量 1 的跃迁元概念，我们给这三组实验分别写下一个跃迁元，为 $\langle +T | +S \rangle, \langle 0T | +S \rangle, \langle -T | +S \rangle$ ，进一步，受上一章讲过的跃迁元的模方决定跃迁概率的启发，我们宣称，这三种情况的跃迁概率分别由这三个跃迁元的模方给出。从而，由于这三种情况的跃迁概率均非零，所以我们知道 $\langle +T | +S \rangle \neq 0, \langle 0T | +S \rangle \neq 0, \langle -T | +S \rangle \neq 0$ 。推而广之，即是，原子从 (iS) 可能性跃迁到 (jT) 可能性的概率 $P(iS \rightarrow jT)$ ，由下式给出

$$P(iS \rightarrow jT) = |\langle jT | iS \rangle|^2, \quad i, j = +, 0, -. \quad (2.16)$$

注意，这个结论是从大量实验中归纳出来的，叫做玻恩定则，它不能从别的地方推导出来。

完全类似于上面三组实验， $0S$ 可能性以及 $-S$ 可能性到 $+T, 0T, -T$ 三种可能性的跃迁概率也都不是零，从而，任给 $i, j = +, 0, -,$ 均有 $P(iS \rightarrow jT) = |\langle jT | iS \rangle|^2 \neq 0$ ，从而 $\langle jT | iS \rangle \neq 0$ 。给定初态可能性 iS ，末态必定跃

迁到 $+T, 0T, -T$ 三种可能性之一，也就是说， iS 分别跃迁到这三种可能性的概率总和必定等于1，从而有

$$\sum_j |\langle jT|iS \rangle|^2 = 1. \quad (2.17)$$

又由于任给 $j = +, 0, -$ 均有 $\langle jT|iS \rangle \neq 0$ ，因此还能推论出 $\langle jT|iS \rangle \neq 1$ (三个正数加起来等于1，必定每个分别都不等于1)。

$\langle jT|iS \rangle$ 当然是物理量1的跃迁元，但是请注意，和之前讨论的跃迁元都不一样，现在这个跃迁元的初态可能性和末态可能性属于不同的可确定区分可能性完备集，而不是属于同一个完备集。而且，也正由于这里涉及到两套不同的可确定区分可能性完备集，所以1的跃迁元可以不是 δ_{ij} 。1当然是一种很特殊的物理量，所以也常常称相应的 $\langle jT|iS \rangle$ 这样的跃迁元为跃迁幅。所以，玻恩定则告诉我们的就是，在不受任何微扰的情况下，跃迁概率由相应跃迁幅的模方给出来。

确定原子的一种 S 可能性，比如像上面三组实验那样，通过 $+S$ 过滤器过滤出确定的 $+S$ 可能性，那么由于 $\langle jT|iS \rangle \neq 0, 1$ ，那原子的 T 可能性就是不确定的！因为如果 T 可能性确定，那原子当然就只会百分之百地跃迁到这种确定的 T 可能性，但现在原子由 $+S$ 跃迁到 $+T, 0T, -T$ 三种 T 可能性的概率均非零，所以 T 可能性是不确定的。换言之，你把一个处于 $+S$ 可能性的原子输入一台 T 过滤器，你会发现， $+T, 0T, -T$ 三种情况原子通出的概率都不是零，就像上面描述的那三组实验告诉我们的那样。所以，确定原子的一种 S 可能性，那原子的 T 可能性就将不能确定。

一个问题是，假设我们先让原子通过一台 $+S$ 过滤器，接着再通过一台 $0T$ 过滤器，那么我们是不是就能同时确定原子的 S 可能性和 T 可能性了呢？当然，我们知道，这时候过滤出来的原子其 T 可能性一定是 $0T$ (因为后通过的是 $0T$ 过滤器)，问题是，它的 S 可能性还是不是原来确定的 $+S$ 呢？为此，我们接着把从 $0T$ 过滤器出来的原子分别通入 $+S, 0S, -S$ 过滤器，即

做如下三组实验

$$\begin{aligned} & \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S , \\ & \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S , \\ & \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S . \end{aligned}$$

如果原子的 S 可能性同时保持为第一台过滤器过滤出来的 $+S$, 那这三组实验将只有第一组原子最终从整个装置右边出来的概率非零。但是, 事实是, 三组实验, 原子从右边出来的概率都非零。也即是说, 原子的 S 可能性和 T 可能性无法同时确定!

用量子跃迁很容易解释这个实验结果, 首先, 原子以非零的概率 $|\langle 0T | +S \rangle|^2$ 从第一台装置的 $+S$ 可能性跃迁到第二台装置的 $0T$ 可能性。但是, 一旦跃迁到中间装置的 $0T$ 可能性, 原来 $+S$ 可能性的信息就失去了! 这时候, 再让原子通过第三台装置, 它就会接着从 $0T$ 可能性跃迁到第三台装置的各可能性, 而原子分别跃迁到第三台装置的 $+S, 0S, -S$ 的概率都不是零, 按照玻恩定则, 这些概率分别由 $|\langle +S | 0T \rangle|^2, |\langle 0S | 0T \rangle|^2, |\langle -S | 0T \rangle|^2$ 给出。当然, 从这里我们也能推断出 $\langle +S | 0T \rangle \neq 0, \langle 0S | 0T \rangle \neq 0, \langle -S | 0T \rangle \neq 0$ 。假如将中间的 $0T$ 过滤器换成 $+T$ 过滤器或者 $-T$ 过滤器, 那么类似的实验还将告诉我们 $\langle jS | iT \rangle \neq 0$ (任给 $i, j = +, 0, -$), 进一步由 $\sum_j |\langle jS | iT \rangle|^2 = 1$ 还能推论出 $\langle jS | iT \rangle \neq 1$ 。

由于我们一直是用量子跃迁的思维来思考这个实验, 所以你可能没反应过来以上实验有多么神奇, 它是经典物理完全无法解释的! 在经典物理上, 如果你先过滤出一种可能性 A , 那 A 就确定了, 在这个基础上, 你再过滤出可能性 B , 那 A, B 就同时确定了, 这时候你再用第一种装置再去测试 A 可能性, 你将会发现它一定还是原来确定的 A 。经典的可能性是可以同时确定的。

但是，在我们现在描述的量子实验中， S 可能性和 T 可能性无法同时确定！

不仅如此， S 可能性和 T 可能性还无法确定地区分。为了说清楚这是什么意思，让我们设想有一些原子，只知道每一个均可能处于三个 S 可能性之一，或者处于三个 T 可能性之一，我们想通过检测确定它们到底处于哪种可能性。为此，我们分别将这些原子通入 $+S$, $0S$, $-S$ 过滤器(通入 T 过滤器的分析是一样的)，然后比方说吧，我们检测到通入 iS 过滤器的原子从装置右边出来了，由此我们能不能百分之百地推断出检测之前这个原子处于哪一个可能性呢？回答是，推断不了，因为它既可能是由原来的 iS 可能性跃迁来的(iS 跃迁到 iS 的概率是1)，也可能是由原来的任何 jT 可能性跃迁来的，因为前面说过 $\langle iS|jT \rangle \neq 0$ ，从而 jT 可能性跃迁到 iS 可能性的概率也不是零。换言之，对于我们的检测而言， iS 可能性与三个 jT 可能性无法百分之百地区分，或者说无法确定地区分！从这个分析也可以看出，无法确定区分的原因，就在于 $\langle iS|jT \rangle \neq 0, 1$ (或者 $\langle iT|jS \rangle \neq 0, 1$)。

玻恩定则，不确定性与不可区分性

下面把本节的讨论推广到任意量子系统。假设有一个任意的量子系统，我们考察它的两个不同的可确定区分可能性完备集，分别记作 $\mathcal{C} = \{(i), i = 1, 2, 3, \dots\}$ 以及 $\mathcal{D} = \{(i'), i' = 1', 2', 3', \dots\}$ 。所谓 \mathcal{C} 和 \mathcal{D} 不同，其意思就是，存在某些可能性 $(m) \in \mathcal{C}$ ，以及某些可能性 $(n') \in \mathcal{D}$ ，使得

$$\langle n'|m \rangle \neq 0, 1. \quad (2.18)$$

这么定义 \mathcal{C} 和 \mathcal{D} 不同的原因在于，如果它们是同一个可确定区分可能性完备集，则由可确定区分性的定义，就必然有，对于任何 $(i) \in \mathcal{C}$ 和 $(j') \in \mathcal{D}$ ，均有 $\langle j'|i \rangle$ 要么等于0要么等于1，这时候，集合 \mathcal{D} 实际上就只是集合 \mathcal{C} 的一个重排而已。

首先，玻恩定则告诉我们，系统在不受任何微扰时由 $(i) \in \mathcal{C}$ 可能性跃迁到 $(j') \in \mathcal{D}$ 可能性的概率由下式给出

$$P(i \rightarrow j') = |\langle j'|i \rangle|^2, \quad (2.19)$$

式中 $\langle j'|i \rangle = \langle j'|1|i \rangle$ 称作由 (i) 到 (j') 的跃迁幅。而由总概率等于1，我们还有

$$\sum_{j' \in \mathcal{D}} |\langle j'|i \rangle|^2 = 1. \quad (2.20)$$

其次，完全类似于上面对改版施特恩-格拉赫实验的讨论，也必然有，如果 $\langle n'|m \rangle \neq 0, 1$ ，则可能性 $(m) \in \mathcal{C}$ 与可能性 $(n') \in \mathcal{D}$ 不可以同时确定，并且，不存在检测手段可以将它们确定地区分开来。这就是所谓的不确定性与不可区分性原理。它也可以表述成：如果 \mathcal{C}, \mathcal{D} 是系统两个不同的可确定区分可能性完备集，则总体上，分属于这两个完备集的可能性不可以同时确定，并且没有检测手段可以将它们百分之百地区分。所谓总体上，就是不排除某些可能性可以同时确定以及百分之百地区分，但是，不可能所有可能性都可以这样。

比如，一个粒子所有具有确定位置坐标的可能性构成粒子的一个可确定区分可能性完备集，同样，一个粒子所有具有确定动量的可能性也构成粒子的一个可确定区分可能性完备集。但是，这两个完备集是两个不同的完备集，因此根据上面的不确定性和不可区分性原理，粒子的位置和动量不可以同时确定，并且不存在检测手段能百分之百地检测出粒子是具有确定的位置还是具有确定的动量。

2.3 跃迁幅乘法

继续我们的改版施特恩-格拉赫实验。假设我们在 S 装置和 T 装置的基础上，再引入一个与这两者都不同(转过了一个不同的角度)的 R 装置。假设我们来分析下面的实验

$$\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_S \begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_T \begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_R . \quad (2.21)$$

注意原子是从左边往右边通行的。实验的结果是发现，中间的 T 装置并没有起到影响实验的作用，换言之，

$$\begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_S \begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_R = \begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_S \begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_T \begin{Bmatrix} + \\ 0 \\ - \end{Bmatrix}_R . \quad (2.22)$$

我们发现以上实验结果是符合跃迁元乘法的，虽然要将这乘法稍作推广。下面我们来分析原子从左边 $+S$ 过滤器的 $+S$ 可能性跃迁到右边 $+R$ 可能

性的跃迁幅(跃迁概率是它的模方), 即分析 $\langle +R| +S \rangle$ 。仿照第一章海森堡引进跃迁元乘法的思想, 我们发现这可以分解成 $+S$ 先跃迁到中间*iT*可能性的跃迁幅, 再乘以从中间*iT*可能性跃迁到末尾 $+R$ 可能性的跃迁幅, 并且, 由于整个装置完全不能具体确定中间可能性是什么, 所以, 我们要对所有中间可能性求和, 从而我们写出

$$\langle +R| +S \rangle = \sum_i \langle +R|iT \rangle \langle iT| +S \rangle. \quad (2.23)$$

这就是(2.22)告诉我们的一切。当然, 这个式子涉及到的三个跃迁幅都是物理量1的跃迁元, 物理量1乘以物理量1, 结果还是1, 所以, 跃迁幅乘以跃迁幅, 结果依然是跃迁幅, 只是等式的右边要插入中间可能性并对之求和。因此, 这个式子反映的就是一种跃迁元乘法, 只不过, 和第一章不同, 现在的初态可能性、中间可能性以及末态可能性, 分别属于三个不同的可确定区分可能性完备集, 而不是属于同一个完备集。

以上结论当然可以推广, 比如我们可以写出下式

$$\langle jR|iS \rangle = \sum_k \langle jR|kT \rangle \langle kT|iS \rangle. \quad (2.24)$$

它描述的依然是跃迁幅乘法, 并且可以用完全类似于上面的实验来检验。

记跃迁幅 $\langle jR|iS \rangle = \psi$, 我们发现它是三项的和, 分别记这三项为 $\phi_+ = \langle jR| +T \rangle \langle +T|iS \rangle$, $\phi_0 = \langle jR|0T \rangle \langle 0T|iS \rangle$, $\phi_- = \langle jR| -T \rangle \langle -T|iS \rangle$, 从而由(2.24)式,

$$\psi = \phi_+ + \phi_0 + \phi_-. \quad (2.25)$$

式中 ϕ_+ 描写的是原子跃迁的时候从中间 $+T$ 可能性经过的跃迁幅, ϕ_0, ϕ_- 的含义依此类推。

下面, 我们来计算*iS*跃迁到*jR*的概率 P , 也即 $P = |\langle jR|iS \rangle|^2 = |\psi|^2$, 根据(2.25), 有

$$\begin{aligned} P = & |\phi_+ + \phi_0 + \phi_-|^2 = |\phi_+|^2 + |\phi_0|^2 + |\phi_-|^2 \\ & + (\phi_+^* \phi_0 + \phi_0^* \phi_+) + (\phi_0^* \phi_- + \phi_-^* \phi_0) + (\phi_-^* \phi_+ + \phi_+^* \phi_-). \end{aligned} \quad (2.26)$$

我们称 $|\phi_+|^2$ 为原子经过中间可能性 $+T$ 跃迁到末态的概率, 类似的, $|\phi_0|^2$ 为原子经过中间 $0T$ 跃迁到末态的概率, $|\phi_-|^2$ 为原子经过中间 $-T$ 跃迁到末

态的概率，但是，(2.26)式清楚地告诉我们，原子经过所有中间可能性跃迁的概率 P ，并不等于这三个概率的和！而是多出来6个两不同 ϕ 的交叉项。通常称这些交叉项为干涉项，它描述了原子跃迁的时候，三个分跃迁幅 ϕ_+, ϕ_0, ϕ_- 之间的交叉干涉。

干涉项的存在，是量子力学不同于经典物理的本质特征之一，它来自于哪里呢？我们将整个推导过程回溯一下就会发现，它正来自于跃迁幅乘法的对所有中间可能性求和！注意，这个求和不是把 $|\phi_+|^2, |\phi_0|^2, |\phi_-|^2$ 这三个概率求和，而是在模方算概率之前，对跃迁幅的几种可能性求和，正是因此，才导致最终出现量子干涉项。

诡异的地方在于，你可能觉得，原子从初态可能性跃迁到末态可能性的时候，总是要么经过中间可能性 $+T$ ，要么经过 $0T$ ，要么经过 $-T$ ，此外并没有其它可能了。但，这样你就是在用经典物理进行思维，按照这种思维，最终的概率 P 就应该等于这三种可能性的概率之和（要么要么要么，三个要么的概率不就是相加么），从而即有 $P = |\phi_+|^2 + |\phi_0|^2 + |\phi_-|^2$ ，那就没有量子干涉项了，而实验的结果是验证了干涉之存在的。

问题出在哪里呢？回答是，出在当你像(2.21)的装置那样安排实验的时候，你对中间可能性是完全不能确定的，换言之，你根本无从知道原子是如何经过中间可能性的。你说，难到不是要么经过 $+T$ ，要么经过 $0T$ ，要么经过 $-T$ 吗？但你也看到了，这样想是错的，这样想就无法解释量子干涉的存在了。这样想错就错在，你没有对中间可能性进行观测，因此，你就甚至连“要么经过 $+T$ ，要么经过 $0T$ ，要么经过 $-T$ ”都不能说！

你可能说，不就是没有观测吗？那观测一下好了，因此你把(2.21)的实验装置改装成

$$\left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} +\odot \\ 0\odot \\ -\odot \end{array} \right\}_T \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_R . \quad (2.27)$$

其中符号 \odot 代表一个探测器，比方说一盏灯，它的工作机制是，如果原子从 T 装置中最上面的路径经过（即 $+T$ 可能性），那最上面的那盏灯就闪一下，原子从 T 装置中间的路径经过（ $0T$ 可能性），那中间的灯就闪一下，同样，原子从最下面的路径经过，最下面的灯就闪一下。所以，你只要观测到底是哪盏灯闪了，就能知道原子在 T 装置中到底是经过了哪个中间可能性。

实验的结果是，的确，每次有一个原子成功从第一台装置出来的时候，接着我们总是观察到中间的T装置要么最上面的灯闪了，要么中间的灯闪了，要么最下面的灯闪了，从来不会有两盏灯同时闪。这当然因为每次通过的只有一个原子，它并没有分身术。所以，现在，每一个原子从中间通过的时候，通过观察三盏灯中哪一盏闪了，你的确就可以知道它通过的是哪一种中间可能性了，因此你的确可以说原子“要么经过 $+T$ ，要么经过 $0T$ ，要么经过 $-T$ ”了，那最终还存在量子干涉吗？谢天谢地，就和之前我们推理的那样，量子干涉不存在了(否则就真无法解释实验了)！这时候总概率的确就是 $P = |\phi_+|^2 + |\phi_0|^2 + |\phi_-|^2$ 。

总结一下就是：如果我们原则上就不能确定原子是如何经过中间可能性的，那我们就是按照跃迁幅乘法(2.24)来计算最终的跃迁幅，需要把跃迁幅的中间可能性都加起来。但是，如果我们能够确定原子是如何经过中间可能性的，那就不能按照跃迁幅乘法(2.24)来计算，而是要分别把每一种可能性的概率(而不是跃迁幅)加起来，就像下面这样

$$P = \sum_k |\langle jR|kT\rangle|^2 |\langle kT|iS\rangle|^2. \quad (2.28)$$

所以，这一节我们得到了两条规则，跃迁幅乘法规则(2.24)，以及概率求和规则(2.28)。上面的讨论也说清楚了什么情况下该使用哪一条规则。当然，我们也能把这两条规则推广到任意的量子系统，这个推广是容易的，这里就不再赘述了。

跃迁幅的一条性质

对于任意一个量子系统，考察它任意三个可确定区分可能性完备集，分别记作 $\mathcal{I} = \{(i), i = 1, 2, 3, \dots\}$ ， $\mathcal{J} = \{(j'), j' = 1', 2', 3', \dots\}$ ，以及 $\mathcal{K} = \{(k''), k'' = 1'', 2'', 3'', \dots\}$ 。则根据跃迁幅的乘法，有

$$\langle k''|i\rangle = \sum_{j' \in \mathcal{J}} \langle k''|j'\rangle \langle j'|i\rangle. \quad (2.29)$$

特别的，取 $\mathcal{K} = \mathcal{I}$ ，注意到 $\langle i|i\rangle = 1$ (由集合 \mathcal{I} 的可确定区分性)，从而也有

$$1 = \langle i|i\rangle = \sum_{j' \in \mathcal{J}} \langle i|j'\rangle \langle j'|i\rangle. \quad (2.30)$$

另一方面，由总概率守恒方程(2.20)，又有

$$1 = \sum_{j' \in \mathcal{J}} |\langle j'|i\rangle|^2 = \sum_{j' \in \mathcal{J}} \langle j'|i\rangle^* \langle j'|i\rangle. \quad (2.31)$$

比较(2.30)式和(2.31)式，不难发现，能够使得这两个式子总是同时成立的最简单条件是

$$\langle i|j'\rangle = \langle j'|i\rangle^*. \quad (2.32)$$

即，从(j')到(i)的跃迁幅等于逆向(i)到(j')的跃迁幅的复数共轭，或者示意性地记作

$$\left\{ (i) \xleftarrow{1} (j') \right\} = \left\{ (i) \xrightarrow{1} (j') \right\}^*. \quad (2.33)$$

(2.32)式是跃迁幅的一条最重要的性质。

2.4 跃迁元乘法以及物理量的测量

下面让我们一般性地推广跃迁元的概念，以及跃迁元乘法规则。还是从改版的施特恩-格拉赫实验开始，

$$\begin{Bmatrix} + \\ 0\parallel \\ -\parallel \end{Bmatrix}_S \begin{Bmatrix} A \end{Bmatrix} \begin{Bmatrix} +\parallel \\ 0 \\ -\parallel \end{Bmatrix}_R. \quad (2.34)$$

中间的 A 是指任何一个施特恩-格拉赫装置的复杂组合，其中磁场可以转向任何方向，甚至可以有额外的电场，也可以有各种挡板等等....总之，你可以想象放进去任何东西。当然，原子是从整个装置的左边往右边通行的。问题是，如何用一个量合适地描写这个实验从事的事情呢，即如何描写一个由左边 S 过滤器过滤出来的 $+S$ 原子，经过中间的复杂装置 A 的作用，最终跃迁到右边 R 过滤器的 $0R$ 可能性呢？受第一章跃迁元概念的启发，我们可以写下如下量

$$\langle 0R|\hat{A}|+S\rangle, \quad (2.35)$$

注意，这个式子是从右往左读的。式中 \hat{A} 表示装置 A 对原子的作用，在这个符号上加上 \wedge ，是为了把 A 在原子上的作用与装置 A 本身区分开来。

很明显，以上量可以推广到任意初始的 S 可能性，和任意末尾的 R 可能性，进而即有

$$\langle jR|\hat{A}|iS\rangle, \quad i, j = +, 0, - \quad (2.36)$$

我们称之为物理量 A （或者装置 A ）的跃迁元！描写的是原子从初始(iS)可能性出发，经过物理量 A 的作用，最终跃迁到(jR)可能性的过程（读式子的时候，从右边往左边读）。跃迁的概率就取决于这个量的模方。特别的， A 可以是物理量 1，这时候它就表示中间的装置对原子不起任何作用。

为了考察跃迁元的乘法，让我们假设装置 C 是由装置 A 和 B 串联而成，原子先通过 B ，接着再通过 A ，即

$$\left\{ C \right\} = \left\{ B \right\} \left\{ A \right\}, \quad (2.37)$$

示意性地可以写作

$$\hat{C} = \hat{A}\hat{B}, \quad (2.38)$$

即 C 的作用等于 B 的作用再乘以 A 的作用（从右往左读）。

关键问题是 $\hat{A}\hat{B}$ 这个乘法该怎么定义？为此，我们回到实验，也就是要回到对量子跃迁过程的分析，结果表明

$$\begin{aligned} & \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} B \\ \vdots \end{array} \right\} \left\{ \begin{array}{c} A \\ \vdots \end{array} \right\} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_R \\ &= \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_S \left\{ \begin{array}{c} B \\ \vdots \end{array} \right\} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_T \left\{ \begin{array}{c} A \\ \vdots \end{array} \right\} \left\{ \begin{array}{c} + \\ 0 \\ - \end{array} \right\}_R. \end{aligned} \quad (2.39)$$

根据这个结果，我们可以仿照海森堡的跃迁元乘法，写下

$$\langle 0R|\hat{A}\hat{B}|+S\rangle = \sum_k \langle 0R|\hat{A}|kT\rangle \langle kT|\hat{B}|+S\rangle. \quad (2.40)$$

注意，原子是从左到右通行，但是式子的读法却是从右往左。

将上面的式子稍作推广，就得到

$$\langle jR|\hat{A}\hat{B}|iS\rangle = \sum_k \langle jR|\hat{A}|kT\rangle \langle kT|\hat{B}|iS\rangle. \quad (2.41)$$

这就是跃迁元乘法规则，它的关键就是在等式右边插入中间可能性并求和。式中涉及的 S 装置、 R 装置、以及 T 装置，都是任意的，特别的，三者当然也可以选成同一种装置。相比于海森堡原始版本的跃迁元乘法规则，现在的规则中，初态可能性、中间可能性，以及末态可能性，三者可以属于三个不同的可确定区分可能性完备集。特别的，取 $\hat{A} = \hat{B} = 1$ ，那么这条规则给出来的就是前面的跃迁幅乘法规则(2.24)。

实际应用中，常常将上式涉及的 S, R, T 三种装置取成同一种，比方都取成 T 装置(这时候符号 T 就可以省略了)，那就有

$$\langle j|\hat{A}\hat{B}|i\rangle = \sum_k \langle j|\hat{A}|k\rangle \langle k|\hat{B}|i\rangle. \quad (2.42)$$

这时候，人们习惯把 $\langle j|\hat{A}|i\rangle$ 这样的量排成矩阵，如下

$$\begin{bmatrix} \langle +|\hat{A}|+\rangle & \langle +|\hat{A}|0\rangle & \langle +|\hat{A}|-\rangle \\ \langle 0|\hat{A}|+\rangle & \langle 0|\hat{A}|0\rangle & \langle 0|\hat{A}|-\rangle \\ \langle -|\hat{A}|+\rangle & \langle -|\hat{A}|0\rangle & \langle -|\hat{A}|-\rangle \end{bmatrix}. \quad (2.43)$$

这时候相应的乘法(2.42)就变成了标准的矩阵乘法。

推广到任意量子系统

对于任意一个量子系统，考察它任意两个可确定区分可能性完备集，分别记作 $\mathcal{I} = \{(i), i = 1, 2, 3, \dots\}$, $\mathcal{J} = \{(j'), j' = 1', 2', 3', \dots\}$ 。则系统的任何一个物理量 A (或者装置 A)都对应一组跃迁元，记作 $\langle j'|\hat{A}|i\rangle$ 。

在集合 \mathcal{I}, \mathcal{J} 的基础上，再添上一个新的可确定区分可能性完备集 $\mathcal{K} = \{(k''), k'' = 1'', 2'', 3'', \dots\}$ ，并考虑两个物理量 A, B ，则它们乘积的跃迁元由下式给出

$$\langle k''|\hat{A}\hat{B}|i\rangle = \sum_{j' \in \mathcal{J}} \langle k''|\hat{A}|j'\rangle \langle j'|\hat{B}|i\rangle. \quad (2.44)$$

这就是跃迁元乘法规则，正如前面说过的，其关键就是在等式右边的合适地方插入中间可能性，并对之求和。特别的，取 $A = 1, B = 1$ ，利用 $1 \cdot 1 = 1$ ，即得一般性的跃迁幅乘法规则

$$\langle k''|i\rangle = \sum_{j' \in \mathcal{J}} \langle k''|j'\rangle \langle j'|i\rangle. \quad (2.45)$$

而只取 $B = 1$ ，则根据 $\hat{A} \cdot 1 = \hat{A}$ ，即有

$$\langle k''|\hat{A}|i\rangle = \sum_{j' \in \mathcal{J}} \langle k''|\hat{A}|j'\rangle \langle j'|i\rangle. \quad (2.46)$$

注意，所有这些结果中的初态可能性、中间可能性以及末态可能性，三者可以分属三个不同的可确定区分可能性完备集。不过，实际应用中往往不需要这么大的任意性，常常人们会将这三个可能性完备集取成同一个，比方说都取成集合 \mathcal{I} ，那这时候上面的乘法规则就变成

$$\langle k|\hat{A}\hat{B}|i\rangle = \sum_{j \in \mathcal{I}} \langle k|\hat{A}|j\rangle \langle j|\hat{B}|i\rangle. \quad (2.47)$$

而物理量的跃迁元也就变成了 $\langle j|\hat{A}|i\rangle$ 这种形式，这时候人们常常把它排成下面的矩阵形式

$$\begin{bmatrix} \langle 1|\hat{A}|1\rangle & \langle 1|\hat{A}|2\rangle & \dots \\ \langle 2|\hat{A}|1\rangle & \langle 2|\hat{A}|2\rangle & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (2.48)$$

这时候常常称可能性完备集 \mathcal{I} 为 \mathcal{I} 表象，上面的矩阵就称为物理量 A 在 \mathcal{I} 表象中的矩阵表示。而跃迁元乘法(2.47)当然也就成了矩阵乘法。

比如我们上一章所考察的定态可能性完备集 \mathcal{E} ，就构成了所谓的能量表象，上一章的所有讨论都是在能量表象中进行的。上一章的所有跃迁元数据表都是相应物理量在能量表象下的表示矩阵。在历史上，海森堡最先引入的，就是能量表象。但是，我们也可以用任何表象来处理量子力学问题，因为正如马上我们就会看到的，量子力学的基本方程是表象无关的。虽然上一章我们是在能量表象中推导出这些方程的，但结果却是，只要稍作推广，就可以使得它们不依赖于任何具体的表象。

物理量的值

那么能量表象特殊在哪里呢？回答是，特殊在哈密顿量在这个表象中的表示矩阵为对角矩阵，用现在的符号来写即是(哈密顿量 H 对物理系统的作用当然就记作 \hat{H})

$$\begin{bmatrix} \langle E_1 | \hat{H} | E_1 \rangle & \langle E_1 | \hat{H} | E_2 \rangle & \dots \\ \langle E_2 | \hat{H} | E_1 \rangle & \langle E_2 | \hat{H} | E_2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} E_1 & 0 & \dots \\ 0 & E_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (2.49)$$

或者简单地写作

$$\langle E_m | \hat{H} | E_n \rangle = E_n \delta_{nm}. \quad (2.50)$$

也即是说，定态可能性(E_n)通过这个 H 的跃迁元只能跃迁到(E_n)自身，跃迁到其它不同定态可能性的概率都是零。从而，当系统处于定态可能性(E_n)时，其能量有确定值 E_n 。这就是任意定态可能性($E_n \in \mathcal{E}$)在物理上的特殊性。

现在如果我们让系统处在一个($i \in \mathcal{I}$)可能性，我们问，系统的能量值为多少呢？回答是，能量值是不确定的，当你去测量能量值的时候，系统会随机地从(i)可能性跃迁到某个定态可能性(E_n)，一旦它跃迁到(E_n)，那你测得的能量值就是 E_n ，而根据玻恩定则，(i)跃迁到(E_n)的概率是

$$P(i \rightarrow E_n) = |\langle E_n | i \rangle|^2. \quad (2.51)$$

所以，当系统处于(i)可能性时，如果你把系统复制多份(即制备很多份同样处于(i)的系统)，然后分别去测量它们的能量值，则，每一次你测到的值都可能不一样，你会随机地(以概率 $|\langle E_n | i \rangle|^2$)得到不同能量值 E_n 。

由此也可以知道，在任意的 \mathcal{I} 表象下，哈密顿量的表示矩阵可不是对角矩阵！只有在能量表象中，哈密顿量的表示矩阵才是对角矩阵。

完全类似的，任何物理量 A 也都有一个这样特殊的表象，其中 A 的表示矩阵是对角化的。不过，和此前讨论的表象不一样的是，由于物理量 $A(t)$ 在随着时间演化(这与哈密顿量不同，由于能量守恒，哈密顿量并不随时间演化)，所以其对角化的表象也必定是相对于参考时刻 t 而定义的。

不妨把这个相应的可确定区分可能性完备集记作 $\mathcal{A}_t = \{(a, t), a = 1, 2, 3, \dots\}$ ((a, t) 表示相对于参考时刻 t 定义的可能性 (a))，则根据定义，它满足

$$\langle b, t | \hat{A} | a, t \rangle = \lambda_a \delta_{ab}, \quad (2.52)$$

从而，当系统处于可能性 (a, t) 时，其物理量 $A(t)$ 有确定值 λ_a 。但是，如果系统处于任意的 $(i) \in \mathcal{I}$ 可能性，你去测量物理量 $A(t)$ 的值，你会发现，这个值是不确定的，你会以概率

$$P(i \rightarrow a, t) = |\langle a, t | i \rangle|^2 \quad (2.53)$$

随机地得到值 λ_a ，而一旦你测得 λ_a ，系统也就从可能性 (i) 跃迁到了可能性 (a, t) 。

由于物理量 $A(t)$ 在随时间演化，所以一个自然的问题是，假设系统在 t_1 时刻处于可能性 (a, t_1) ，从而其物理量 $A(t_1)$ 的值为 λ_a ，那么它在 t_2 时刻跃迁到可能性 (b, t_2) (从而其物理量 $A(t_2)$ 的值为 λ_b) 的概率是多少？很显然，这个概率将由下式计算

$$P(a, t_1 \rightarrow b, t_2) = |\langle b, t_2 | a, t_1 \rangle|^2. \quad (2.54)$$

特别的，物理上很重要的一种情形是，让 $t_1 \rightarrow -\infty$ ，让 $t_2 \rightarrow +\infty$ ，即考察系统从无穷远过去的 (a) 可能性跃迁到无穷远将来的 (b) 可能性的概率，

$$P(a, -\infty \rightarrow b, +\infty) = |\langle b, +\infty | a, -\infty \rangle|^2. \quad (2.55)$$

这时候 $\langle b, +\infty | a, -\infty \rangle$ 这样的跃迁元就被称为 S 矩阵 (和本章讨论的 S 装置没有关系)，它在讨论粒子从无穷远过去入射，经过相互作用的散射，在无穷远将来再出射出去，这样的粒子散射问题时很常用。

假设系统处于 $(i) \in \mathcal{I}$ 可能性上，我们把它复制很多份，然后多次重复测量物理量 $A(t)$ 的值，我们想知道， $A(t)$ 的平均值是多少？

记 $A(t)$ 的平均值为 $\langle A \rangle$ ，由于测量时我们是以概率 $|\langle a, t | i \rangle|^2$ 随机得到 λ_a ，所以很显然，

$$\langle A \rangle = \sum_a \lambda_a |\langle a, t | i \rangle|^2 = \sum_a \lambda_a \langle a, t | i \rangle^* \langle a, t | i \rangle, \quad (2.56)$$

利用跃迁幅的基本性质 $\langle a, t | i \rangle^* = \langle i | a, t \rangle$ 即有

$$\langle A \rangle = \sum_a \lambda_a \langle i | a, t \rangle \langle a, t | i \rangle = \sum_{a,b} \langle i | b, t \rangle \lambda_a \delta_{ab} \langle a, t | i \rangle. \quad (2.57)$$

代入(2.52)式的 $\langle b, t | \hat{A} | a, t \rangle = \lambda_a \delta_{ab}$, 即有

$$\langle A \rangle = \sum_{a,b} \langle i | b, t \rangle \langle b, t | \hat{A} | a, t \rangle \langle a, t | i \rangle, \quad (2.58)$$

下面, 对中间可能性 (a, t) 和 (b, t) 分别应用跃迁元乘法规则, 具体来说, 就是应用与 $1 \cdot \hat{A} \cdot 1 = \hat{A}$ 对应的跃迁元乘法规则, 即有

$$\langle A \rangle = \langle i | \hat{A} | i \rangle. \quad (2.59)$$

这就是我们最终要得到的公式, 它告诉我们, 对一个处于 $(i) \in \mathcal{I}$ 可能性的系统进行重复测量, 得到的物理量 A 的平均值等于跃迁元 $\langle i | \hat{A} | i \rangle$ 。

2.5 表象无关的量子力学方程

一个量子系统的可能表象有无穷多个, 如果每个表象中的量子力学方程都不同, 那量子力学的复杂性就会让我们绝望。好在, 利用跃迁元乘法可以证明, 量子力学方程具有一种表象无关性, 并且, 表象无关的方程可以从我们上一章在能量表象下得到的量子力学方程推广而来。量子力学中的表象就好比广义相对论中的局部坐标系, 换表象就好比换局部坐标系, 而表象无关就好比广义相对论中的广义协变性。

厄米条件

让我们先来把下面物理量 A 的厄米条件推广成表象无关的形式,

$$\langle E_m | \hat{A} | E_n \rangle = \langle E_n | \hat{A} | E_m \rangle^*. \quad (2.60)$$

为此, 我们任取系统的两个可确定区分可能性完备集, $\mathcal{I} = \{(i), i = 1, 2, 3, \dots\}$ 和 $\mathcal{J} = \{(j'), j' = 1', 2', 3', \dots\}$, 并考察物理量 A 的跃迁元 $\langle j' | \hat{A} | i \rangle$.

进一步取定态可能性完备集 \mathcal{E} 为中间可能性集，并应用与 $\widehat{A} = 1 \cdot \widehat{A} \cdot 1$ 相应的跃迁元乘法，则有

$$\langle j'|\widehat{A}|i\rangle = \sum_{m,n} \langle j'|E_m\rangle \langle E_m|\widehat{A}|E_n\rangle \langle E_n|i\rangle, \quad (2.61)$$

利用跃迁幅的基本性质(2.32式) $\langle j'|E_m\rangle = \langle E_m|j'\rangle^*$, $\langle E_n|i\rangle = \langle i|E_n\rangle^*$, 以及能量表象中的厄米条件(2.60), 即可以得到

$$\begin{aligned} \langle j'|\widehat{A}|i\rangle &= \left(\sum_{m,n} \langle E_m|j'\rangle \langle E_n|\widehat{A}|E_m\rangle \langle i|E_n\rangle \right)^* \\ &= \left(\sum_{m,n} \langle i|E_n\rangle \langle E_n|\widehat{A}|E_m\rangle \langle E_m|j'\rangle \right)^* \\ &= \langle i|\widehat{A}|j'\rangle^*. \end{aligned} \quad (2.62)$$

最后一个等于号我们再次利用了跃迁元乘法规则。注意，现在 \widehat{A} 左右两边的可确定区分可能性完备集是任意的两个，甚至，这两个集都可以不同。正是在这个意义上，我们说，最终的厄米条件(2.62)是表象无关的！最终的这个结果也可以示意性地记作

$$\left\{ (j') \xleftarrow{A} (i) \right\} = \left\{ (j') \xrightarrow{A} (i) \right\}^*. \quad (2.63)$$

如果取某一个特定的表象 \mathcal{I} , 那相应的厄米条件就成为

$$\langle j|\widehat{A}|i\rangle = \langle i|\widehat{A}|j\rangle^*. \quad (2.64)$$

排列成矩阵的话，即是

$$\begin{bmatrix} \langle 1|\widehat{A}|1\rangle & \langle 1|\widehat{A}|2\rangle & \dots \\ \langle 2|\widehat{A}|1\rangle & \langle 2|\widehat{A}|2\rangle & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} \langle 1|\widehat{A}|1\rangle^* & \langle 2|\widehat{A}|1\rangle^* & \dots \\ \langle 1|\widehat{A}|2\rangle^* & \langle 2|\widehat{A}|2\rangle^* & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (2.65)$$

根据第一章中关于厄米矩阵的定义可知，这样的矩阵就是厄米矩阵。所以，在一个确定的表象中，满足厄米条件的物理量的表示矩阵必定是厄米矩阵。

海森堡运动方程

下面我们来证明海森堡运动方程以及量子力学基本对易关系均可以推广成表象无关的形式。为此，我们还需要先定义一下表象无关的跃迁元加法(虽然它是比较显然的)，具体来说就是，任给物理量 A, B ，它们对系统的作用分别为 \hat{A} 和 \hat{B} ，我们要定义这两个作用的和 $\hat{A} + \hat{B}$ 。定义如下：任给可能性 $(i) \in \mathcal{I}$ ，以及可能性 $(j') \in \mathcal{J}$ ，定义

$$\langle j' | (\hat{A} + \hat{B}) | i \rangle \equiv \langle j' | \hat{A} | i \rangle + \langle j' | \hat{B} | i \rangle. \quad (2.66)$$

进而也可以定义这两个作用的对易子为

$$\langle j' | [\hat{A}, \hat{B}] | i \rangle \equiv \langle j' | \hat{A} \hat{B} | i \rangle - \langle j' | \hat{B} \hat{A} | i \rangle. \quad (2.67)$$

更为显然的是 \hat{A} 的数乘，它可以定义为

$$\langle j' | \lambda \hat{A} | i \rangle \equiv \lambda \langle j' | \hat{A} | i \rangle, \quad (2.68)$$

式中 λ 为任意常数。

此外还有一些准备工作要做。首先，任给物理量 A ，按照上一章的老记号，它在能量表象中的表示矩阵依然是记为 A (矩阵 A)，从而，用上一章的老记号，即有 $\langle E_m | \hat{A} | E_n \rangle = [A]_{mn}$ 。另一方面，利用跃迁元乘法，我们有

$$\begin{aligned} \langle j' | \hat{A} | i \rangle &= \sum_{m,n} \langle j' | E_m \rangle \langle E_m | \hat{A} | E_n \rangle \langle E_n | i \rangle \\ &= \sum_{m,n} \langle j' | E_m \rangle [A]_{mn} \langle E_n | i \rangle. \end{aligned} \quad (2.69)$$

类似的，根据老记号有 $\sum_l \langle E_m | \hat{A} | E_l \rangle \langle E_l | \hat{B} | E_n \rangle = \sum_l [A]_{ml} [B]_{ln} = [AB]_{mn}$ ，另一方面，再利用跃迁元乘法，即有

$$\begin{aligned} \langle j' | \hat{A} \hat{B} | i \rangle &= \sum_{m,l,n} \langle j' | E_m \rangle \langle E_m | \hat{A} | E_l \rangle \langle E_l | \hat{B} | E_n \rangle \langle E_n | i \rangle \\ &= \sum_{m,n} \langle j' | E_m \rangle [AB]_{mn} \langle E_n | i \rangle. \end{aligned} \quad (2.70)$$

进而当然也有对易子之间的关系

$$\langle j' | [\hat{A}, \hat{B}] | i \rangle = \sum_{m,n} \langle j' | E_m \rangle [A, B]_{mn} \langle E_n | i \rangle, \quad (2.71)$$

假设记 $C = [A, B]$ 为矩阵 A, B 的对易子，那么等式右边的 $[A, B]_{mn}$ 实际上就是 $[C]_{mn}$.

下面来推广海森堡运动方程，对于任意随时间演化的物理量 A ，利用方程(2.69)，即有

$$i\hbar \frac{d}{dt} \langle j' | \hat{A} | i \rangle = \sum_{m,n} \langle j' | E_m \rangle \left(i\hbar \frac{d}{dt} [A]_{mn} \right) \langle E_n | i \rangle, \quad (2.72)$$

进而根据上一章推导出来的能量表象中的海森堡运动方程 $i\hbar \frac{d}{dt} [A]_{mn} = [A, H]_{mn}$ ，即有

$$\begin{aligned} i\hbar \frac{d}{dt} \langle j' | \hat{A} | i \rangle &= \sum_{m,n} \langle j' | E_m \rangle [A, H]_{mn} \langle E_n | i \rangle \\ &= \langle j' | [\hat{A}, \hat{H}] | i \rangle. \end{aligned} \quad (2.73)$$

式中第二行的等于号是倒过来应用了方程(2.71)。如果在形式上把上面结果的左边写成 $i\hbar \frac{d}{dt} \langle j' | \hat{A} | i \rangle = \langle j' | (i\hbar \frac{d}{dt} \hat{A}) | i \rangle$ (毕竟可能性 (i) 和 (j') 又不依赖于时间)，那最终的方程就可以写成

$$\langle j' | i\hbar \frac{d}{dt} \hat{A} | i \rangle = \langle j' | [\hat{A}, \hat{H}] | i \rangle. \quad (2.74)$$

这就是表象无关的海森堡运动方程！所谓表象无关，是指它对任意的初态可能性和任意的末态可能性均成立。

为了省事，也可以把初末态的这两个可能性省略不写(因为它们是任意的)，进而把上面方程写作

$$i\hbar \frac{d}{dt} \hat{A} = [\hat{A}, \hat{H}]. \quad (2.75)$$

当然，就本章而言，这样写只有形式上的意义，但是下一章我们会重新解释物理装置对系统的作用，从而使得这种写法在数学上获得真正的意义。

假设在海森堡运动方程(2.74)中取末态可能性与初态可能性一样，即取 $(j') = (i)$ ，则根据物理量平均值的计算公式(2.59)，不难看出

$$i\hbar \frac{d}{dt} \langle \hat{A} \rangle = \langle [\hat{A}, \hat{H}] \rangle. \quad (2.76)$$

量子力学基本对易关系

为了把量子力学基本对易关系推广成表象无关的形式，和上一章一样，我们依然先考察一维系统，记粒子坐标为 X ，动量为 P ，它们在量子系统上的作用则分别记为 \hat{X}, \hat{P} 。则，根据方程(2.71)，立即有

$$\langle j' | [\hat{X}, \hat{P}] | i \rangle = \sum_{m,n} \langle j' | E_m \rangle [X, P]_{mn} \langle E_n | i \rangle. \quad (2.77)$$

而又由上一章得到的 $[X, P] = i\hbar \cdot 1$ ，即有 $[X, P]_{mn} = i\hbar \delta_{mn}$ ，代入上面这个式子，即有

$$\begin{aligned} \langle j' | [\hat{X}, \hat{P}] | i \rangle &= i\hbar \sum_{m,n} \langle j' | E_m \rangle \delta_{mn} \langle E_n | i \rangle \\ &= i\hbar \sum_n \langle j' | E_n \rangle \langle E_n | i \rangle \\ &= i\hbar \langle j' | i \rangle = i\hbar \langle j' | 1 | i \rangle. \end{aligned} \quad (2.78)$$

第三行的等于号是再一次利用了跃迁幅乘法。同样，可以省略初末态可能性，写作

$$[\hat{X}, \hat{P}] = i\hbar \cdot 1. \quad (2.79)$$

这就是表象无关的量子力学基本对易关系。

推广到任意有限自由度情形都是容易的，比方说推广到 n 自由度系统，这时候就有如下代数关系

$$[\hat{X}_a, \hat{P}_b] = i\hbar \delta_{ab}, \quad [\hat{X}_a, \hat{X}_b] = [\hat{P}_a, \hat{P}_b] = 0, \quad a, b = 1, 2, \dots, n. \quad (2.80)$$

第三章 态和算符

从根本上来说，量子态是一个虚构的概念，态塌缩则更是。另一方面，量子跃迁是量子力学的基本特征，反而幺正演化是构造出来的。——作者

本章讨论如何从跃迁元以及跃迁元乘法中构造出态矢量、希尔伯特空间以及算符等概念。从本章开始，我们省略 \hat{A} 这类符号上面的 $\hat{\cdot}$ ，如果偶尔有和矩阵符号混淆的情况，则会另作说明。

我们已经看到，量子力学最核心的概念就是跃迁元，物理量 A (或者说物理装置 A)的跃迁元可以记作

$$\langle j' | A | i \rangle, \quad (3.1)$$

其中右边的 $|i\rangle$ 叫作系统的初态可能性，它来自于一个可确定区分可能性完备集 \mathcal{I} ，左边的 $\langle j'|$ 叫做系统的末态可能性，它来自于另一个可确定区分可能性完备集 \mathcal{J} ，中间的符号 A 代表装置 A 对系统的作用。所有的物理量都要用跃迁元来表示，只有跃迁元是物理的。

为了测量一个跃迁元，通常我们需要三台装置，第一台装置用来把系统制备到可能性 (i) 上，第二台装置就是我们希望分析的装置 A ，第三台装置是用来检测系统是否处于可能性 (j') ，由这三台装置我们可以测出系统从初态 (i) 跃迁到末态 (j') 的概率，这个概率正好由跃迁元 $\langle j' | A | i \rangle$ 的模方 $|\langle j' | A | i \rangle|^2$ 决定。

两个不同物理量 A, B 相乘在物理上就是把它们的跃迁元相乘，跃迁元的乘法规则是

$$\langle j' | AB | i \rangle = \sum_{k'' \in \mathcal{K}} \langle j' | A | k'' \rangle \langle k'' | B | i \rangle, \quad (3.2)$$

其中 k'' 称作中间可能性，它来自第三个可确定区分可能性完备集 \mathcal{K} ，跃迁元相乘的时候要对 \mathcal{K} 中所有的中间可能性求和。

最为特殊的跃迁元就是物理量 1 的跃迁元，也就是中间装置对系统不起任何作用时的跃迁元，这时候我们常把相应的跃迁元称作跃迁幅，并省略中间的 1，记为 $\langle j' | i \rangle$ 。跃迁幅的模方 $|\langle j' | i \rangle|^2$ 就等于在中间装置不起任何作用时(或者根本没有任何中间装置时) 系统从 (i) 跃迁到 (j') 的概率(注意，是等于而不仅仅是决定)，换言之，也就是你把系统制备到可能性 (i) 上，然后却在可能性 (j') 上检测到它的概率。也正因为这个原因，其它量子力学书常常称跃迁幅为概率幅(虽然也有叫跃迁幅的)。它当然也满足相应的跃迁元乘法，不仅如此，而且由于 $1 \cdot 1 = 1$ ，所以跃迁幅乘以跃迁幅，结果依然是跃迁幅。

通常的量子力学几大公设里的做法是把态矢量、希尔伯特空间以及算符作为第一公设。而本章的主要目的，就是要表明态矢量、希尔伯特空间以及算符，这三者都可以自然地从跃迁元以及跃迁元乘法规则中构造出来。采用量子跃迁以及跃迁元作为第一概念的好处是，它们和真实的物理关系更密切，尤其是，我们会发现用这种概念体系来思考量子力学，则量子态塌缩的概念自动就是不必要的，毕竟，态矢量本身已经只是一种有用的构造而不是第一概念了！

值得强调的是，从历史的角度来看，本书的这种做法并不是异端，正如在本书第一章中读者已经看到的，我们这种做法反而是更符合海森堡的原创精神的，不仅如此，费曼在其物理学讲义第三卷中也曾经沿着这条路子走了一大步，虽然费曼并没有彻底认识到这是一种构建现代量子力学的不同方式，而我们只是把它进一步落实了下来，从中构建整个现代量子力学的理论体系。

3.1 态叠加原理

以电子自旋为例

还是让我们从上一章介绍的改版施特恩-格拉赫实验开始讲起。不过，这一次我们考察一种甚至更简单的情况，我们假设通过施特恩-格拉赫装置

的原子只分裂成两束(也就是所谓的角动量等于 $1/2$ 的情形, 这种角动量由核外电子的自旋贡献)! 因此相应的可确定区分可能性完备集中只有两种可能性, 同样, 我们假设装置沿着水平的 y 轴放置, 磁场的指向我们主要考虑两种不同情况: 如果磁场指向 z 轴, 相应的两种可能性我们记作 $S = \{\uparrow, \downarrow\}$, 分别称作电子自旋向上和自旋向下, 也就是电子自旋方向平行于 z 轴和逆着 z 轴; 而如果将装置绕着水平 y 轴转90度, 使得磁场指向 x 轴, 那相应的两种可能性我们就记作 $T = \{\rightarrow, \leftarrow\}$, 分别称作电子自旋向右和自旋向左。

实验测得, 如下两种情况, 原子的跃迁概率均为 $1/2$ 。

$$\begin{Bmatrix} \uparrow \\ \downarrow \parallel \end{Bmatrix} \begin{Bmatrix} \rightarrow \\ \leftarrow \parallel \end{Bmatrix}, \quad \begin{Bmatrix} \uparrow \\ \downarrow \parallel \end{Bmatrix} \begin{Bmatrix} \rightarrow \parallel \\ \leftarrow \end{Bmatrix} \quad (3.3)$$

用跃迁元的记号来写即是

$$|\langle \rightarrow | \uparrow \rangle|^2 = \frac{1}{2}, \quad |\langle \leftarrow | \uparrow \rangle|^2 = \frac{1}{2}. \quad (3.4)$$

完全类似的实验也告诉我们

$$|\langle \rightarrow | \downarrow \rangle|^2 = \frac{1}{2}, \quad |\langle \leftarrow | \downarrow \rangle|^2 = \frac{1}{2}. \quad (3.5)$$

当然, 由于跃迁幅性质 $\langle j' | i \rangle = \langle i | j' \rangle^*$, 进而 $|\langle j' | i \rangle|^2 = |\langle i | j' \rangle|^2$, 所以我们也有

$$|\langle \uparrow | \rightarrow \rangle|^2 = |\langle \uparrow | \leftarrow \rangle|^2 = |\langle \downarrow | \rightarrow \rangle|^2 = |\langle \downarrow | \leftarrow \rangle|^2 = \frac{1}{2}. \quad (3.6)$$

为了满足(3.4)式, 通常可取

$$\langle \rightarrow | \uparrow \rangle = \langle \uparrow | \rightarrow \rangle = \sqrt{\frac{1}{2}}, \quad \langle \leftarrow | \uparrow \rangle = \langle \uparrow | \leftarrow \rangle = \sqrt{\frac{1}{2}}. \quad (3.7)$$

另一方面由于集合 S 的可确定区分性, 以及跃迁幅乘法(等式右边插入中间可能性完备集 T), 我们有

$$0 = \langle \downarrow | \uparrow \rangle = \langle \downarrow | \rightarrow \rangle \langle \rightarrow | \uparrow \rangle + \langle \downarrow | \leftarrow \rangle \langle \leftarrow | \uparrow \rangle. \quad (3.8)$$

所以, 作了(3.7)式的选取以后, 为了同时满足(3.6)式和(3.8)式, $\langle \downarrow | \rightarrow \rangle$ 和 $\langle \downarrow | \leftarrow \rangle$ 的选取就有所受限, 不过显然可以取为

$$\langle \downarrow | \rightarrow \rangle = \langle \rightarrow | \downarrow \rangle = \sqrt{\frac{1}{2}}, \quad \langle \downarrow | \leftarrow \rangle = \langle \leftarrow | \downarrow \rangle = -\sqrt{\frac{1}{2}}. \quad (3.9)$$

为了更容易看清楚，不妨将上面得到的结果归纳如下

$$\begin{bmatrix} \langle \uparrow | \rightarrow \rangle & \langle \uparrow | \leftarrow \rangle \\ \langle \downarrow | \rightarrow \rangle & \langle \downarrow | \leftarrow \rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (3.10)$$

一个完整的跃迁元总是同时包含初态可能性与末态可能性，然而，有时候我们并不关心系统最终跃迁到了哪种可能性，我们可以把这种情形用如下装置来表示，

$$\left\{ \begin{array}{c} \uparrow \\ \downarrow \parallel \end{array} \right\} \left\{ \dots \right\}, \quad (3.11)$$

其中省略号表示我们并不关心的末态可能性。与之相应的跃迁元就应该写成

$$\langle \dots | \uparrow \rangle. \quad (3.12)$$

利用跃迁幅乘法(即插入中间可能性)，即有

$$\langle \dots | \uparrow \rangle = \langle \dots | \rightarrow \rangle \langle \rightarrow | \uparrow \rangle + \langle \dots | \leftarrow \rangle \langle \leftarrow | \uparrow \rangle. \quad (3.13)$$

当然，这个式子的含义就是，从初态可能性(\uparrow)跃迁到某个末态可能性的跃迁幅，等于通过两个不同中间可能性的跃迁幅的相干叠加。但是，如果我们代入 $\langle \rightarrow | \uparrow \rangle = \langle \leftarrow | \uparrow \rangle = 1/\sqrt{2}$ ，则有

$$\langle \dots | \uparrow \rangle = \langle \dots | \rightarrow \rangle \frac{1}{\sqrt{2}} + \langle \dots | \leftarrow \rangle \frac{1}{\sqrt{2}}. \quad (3.14)$$

由于我们并不关心末态可能性，所以上面的(3.13)式对任意的末态可能性均成立。一个自然的想法是，将(3.13)式各项中的末态可能性符号 $(\dots |$ 干脆省略不写，从而得到一个跃迁元不完整的式子

$$| \uparrow \rangle = | \rightarrow \rangle \frac{1}{\sqrt{2}} + | \leftarrow \rangle \frac{1}{\sqrt{2}}. \quad (3.15)$$

注意，其中 $| \uparrow \rangle$ 这样的表达式应该理解成不完整的跃迁元，或者向末态开放的跃迁元，也就是说，要在心中想象给它左边配上某个末态可能性才完整，只是由于整个等式每一项所配的末态可能性都一样，而且任意，所以我们干脆将它统统省略了。

但在另一方面，如果我们将注意力集中在这个省略了末态可能性的式子(3.15)，我们发现，它也可以解读成是在初态可能性之间引入了某种叠加原理，即，将可能性 $|\rightarrow\rangle$ 和可能性 $|\leftarrow\rangle$ 以 $1/\sqrt{2}$ 为权重进行叠加，即得到可能性 $|\uparrow\rangle$ 。当然，也需要说明，这只是某种有用的解读方式，因为(3.15)式真正完整的写法应该是(3.13)式，而它的含义是，不同中间可能性的跃迁幅的叠加，而不是初态可能性本身的叠加。

然而，可以在可能性之间引入某种叠加原理，这就是量子可能性与经典可能性完全不同的地方！为了让读者看清这种不同，我们来考虑著名的薛定谔的猫：假设有一只猫，它有两种可能性，(死猫)和(活猫)，如果这是一只人们熟悉的经典猫，那这两种可能性显然是不能进行叠加的，(死猫)+(活猫)没有任何意义。但是，如果这是一只量子的猫。那类比于上面自旋的例子，我们就可以在相应的两种量子可能性之间引入叠加，也即是说，我们可以考虑如下可能性

$$|\Delta\text{猫}\rangle = |\text{死猫}\rangle \frac{1}{\sqrt{2}} + |\text{活猫}\rangle \frac{1}{\sqrt{2}}. \quad (3.16)$$

处在这种可能性的猫是一只什么猫呢？它当然不是死猫，但也不是活猫，也不是半死不活的猫，甚至也不是既死又活的猫，所有这些说法都是误导人的。实际上，它是日常概念无法描述的一只猫，因为我们的日常概念是描述经典世界的，而这是猫的一种完全量子的可能性。当然，我们可以说猫的这种可能性是死猫和活猫的叠加，但是，请注意，可能性的叠加只是我们对跃迁元和跃迁元乘法的一种解读，它的真实含义是，处在这种可能性的猫跃迁到(死猫)的概率为 $1/2$ ，跃迁到(活猫)的概率也为 $1/2$ 。

回到自旋的例子。不难明白，和前面类似的考虑，也有

$$|\downarrow\rangle = |\rightarrow\rangle\langle\rightarrow|\downarrow\rangle + |\leftarrow\rangle\langle\leftarrow|\downarrow\rangle = \frac{1}{\sqrt{2}}(|\rightarrow\rangle - |\leftarrow\rangle). \quad (3.17)$$

而如果我们反过来，取 $|\rightarrow\rangle, |\leftarrow\rangle$ 为初态可能性，那就会有

$$\begin{aligned} |\rightarrow\rangle &= |\uparrow\rangle\langle\uparrow|\rightarrow\rangle + |\downarrow\rangle\langle\downarrow|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \\ |\leftarrow\rangle &= |\uparrow\rangle\langle\uparrow|\leftarrow\rangle + |\downarrow\rangle\langle\downarrow|\leftarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle). \end{aligned} \quad (3.18)$$

其中我们代入了(3.10)式的结果。这些结果告诉我们，可能性 $|\uparrow\rangle, |\downarrow\rangle$ (S可能性)是可能性 $|\rightarrow\rangle, |\leftarrow\rangle$ (T可能性)叠加的结果，反过来也一样， $|\rightarrow\rangle, |\leftarrow\rangle$

$|S\rangle$ 也是 $|\uparrow\rangle, |\downarrow\rangle$ 叠加的结果。这正好从另一方面反映了 S 可能性与 T 可能性的不可以同时确定也不可以百分之百地区分的性质。

上面的(3.15)式、(3.17)式以及(3.18)式有一个自然的几何解释。即将 $|\uparrow\rangle, |\downarrow\rangle$ 以及 $|\rightarrow\rangle, |\leftarrow\rangle$ 看成是两维向量空间中的向量，而这几个叠加关系就自然地成为这些向量之间的合成关系，如图(3.1)所示。不仅如此，可确

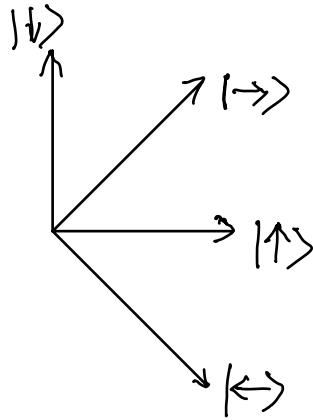


图 3.1: S 可能性与 T 可能性相互作为对方的向量合成

定区分性给出的 $\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1$ 可以解释成向量 $|\uparrow\rangle$ 和向量 $|\downarrow\rangle$ 的模长均为1，而 $\langle\downarrow|\uparrow\rangle = \langle\uparrow|\downarrow\rangle = 0$ 则可以解释成向量 $|\uparrow\rangle$ 和向量 $|\downarrow\rangle$ 正交。类似的解释同样可以用于向量 $|\rightarrow\rangle$ 和向量 $|\leftarrow\rangle$ 。关于跃迁幅的这些解释后文会进一步展开讨论。

类似的讨论也可以用于末态可能性，比方说我们有

$$\begin{aligned}\langle\rightarrow| &= \langle\rightarrow|\uparrow\rangle\langle\uparrow| + \langle\rightarrow|\downarrow\rangle\langle\downarrow| = \frac{1}{\sqrt{2}}(\langle\uparrow| + \langle\downarrow|) \\ \langle\leftarrow| &= \langle\leftarrow|\uparrow\rangle\langle\uparrow| + \langle\leftarrow|\downarrow\rangle\langle\downarrow| = \frac{1}{\sqrt{2}}(\langle\uparrow| - \langle\downarrow|).\end{aligned}$$

当然，这里我们是省略了每个式子右边的初态可能性。所以，这些末态可能性也能进行叠加，因此也有类似的向量解释。

态叠加原理

现在把以上例子阐述的事情推广到一般情况。假设我们考察的是一个任意的量子系统，记 $\mathcal{I} = \{(i), i = 1, 2, 3, \dots\}$ 为它一个任意的可确定区分可

能性完备集。另设 ψ 为属于系统某个可确定区分可能性完备集(一般来说不是 \mathcal{I})的可能性。假设我们把系统制备到 ψ 上，并且不关心它最终会跃迁到哪个末态可能性，因此相应的跃迁幅可以写成 $\langle \dots | \psi \rangle$ ，利用跃迁幅乘法插入中间可能性，即有

$$\langle \dots | \psi \rangle = \sum_{i \in \mathcal{I}} \langle \dots | i \rangle \langle i | \psi \rangle. \quad (3.19)$$

进一步，和上面例子中的相关处理一样，既然这个式子对于任意末态可能性均成立，而我们又不关心末态可能性，那就可以从表达式中将它省略，于是即有

$$|\psi\rangle = \sum_{i \in \mathcal{I}} |i\rangle \langle i | \psi \rangle. \quad (3.20)$$

$\langle i | \psi \rangle$ 当然是复数，不妨记为 a_i ，即 $a_i = \langle i | \psi \rangle$ (给定 ψ , a_i 就给定了)，因此即有

$$|\psi\rangle = \sum_{i \in \mathcal{I}} |i\rangle a_i = \sum_{i \in \mathcal{I}} a_i |i\rangle. \quad (3.21)$$

我们可以把这个式子解读成， \mathcal{I} 中的各可能性以复数 a_i 为权重进行叠加，即得到可能性 $|\psi\rangle$ 。注意，之前都是先给定 ψ , a_i 是由 ψ 决定的，但是，我们也可以反过来，通过任意调节 a_i ，进而由(3.21)式得到相应的 $|\psi\rangle$ 。我们称这样的定义了叠加运算的量子可能性为**量子态**。

既然可以以 a_i 为权重进行叠加，当然也可以以 λa_i 为权重进行叠加，结果是，因此，

$$\sum_{i \in \mathcal{I}} (\lambda a_i) |i\rangle = \lambda \sum_{i \in \mathcal{I}} a_i |i\rangle = \lambda |\psi\rangle. \quad (3.22)$$

这种将 $|\psi\rangle$ 乘以某个复数 λ 的运算称作量子态的数乘运算。任给 $\lambda \neq 0$ ，在物理上， $|\psi\rangle$ 和 $\lambda|\psi\rangle$ 描述的都是系统的同一个可能性(同一个量子态)，因为这两个东西描述的到任意(i)态的跃迁概率都是 $|\langle i | \psi \rangle|^2$ (这是因为总概率是要归一的，因此一个整体的常数 λ 并不影响具体的概率分布情况)。总之， $|\psi\rangle$ 和 $\lambda|\psi\rangle$ 在物理上等价。

假设 ψ_1 为系统的一个量子态， ψ_2 也为系统的一个量子态，且它们分别由下式给出

$$|\psi_1\rangle = \sum_i a_i |i\rangle, \quad |\psi_2\rangle = \sum_i b_i |i\rangle. \quad (3.23)$$

则我们当然也可以把 a_i 和 b_i 相加，假设将集合 \mathcal{I} 中的各量子态以 $a_i + b_i$ 为权重进行叠加，结果即是

$$\sum_i (a_i + b_i)|i\rangle = \sum_i a_i|i\rangle + \sum_i b_i|i\rangle = |\psi_1\rangle + |\psi_2\rangle. \quad (3.24)$$

也即是说，任意两个量子态之间可以进行加法运算，结果是一个新的量子态。既有加法又有数乘，因此 $|\psi_1\rangle$ 和 $|\psi_2\rangle$ 的任意线性叠加 $c_1|\psi_1\rangle + c_2|\psi_2\rangle$ (c_1, c_2 为两个复数)也就有了定义。所以，所有可能的 $|\psi\rangle$ 之间可以进行任意的线性叠加运算！数学上，这样的可以进行任意线性叠加的东西，就叫做矢量，所以 $|\psi\rangle$ 是矢量，称作态矢量！所有态矢量所构成的矢量空间，就叫做希尔伯特空间。

比如说前面电子自旋例子中， $|\uparrow\rangle, |\downarrow\rangle, |\rightarrow\rangle, |\leftarrow\rangle$ 都是态矢量，而且正如前面已经看到的，它们都是二维矢量空间里的矢量。但是，对于更一般的量子系统，其态矢量可不是二维空间的矢量，当然更不是通常的三维欧氏空间的矢量，态矢量是一种抽象的矢量，一般来说，它的矢量空间可以是任意维的，甚至无穷维的。稍后我们会进一步讨论如何刻画这个希尔伯特空间的维数。

注意，一个量子态 ψ 可以对应到任意的 $\lambda|\psi\rangle$ ，只要 $\lambda \neq 0$ ，这些态矢量构成态矢量空间中一条过原点的“直线”(原点要扣掉)¹！所以，一个量子态对应希尔伯特空间中一条过原点的“直线”(原点要扣掉)，因此，量子态的空间就是所谓的射影希尔伯特空间。不过，在通常的表达中，我们可能不会很注意区分量子态和态矢量这两个概念，虽然它们之间严格来说是一对多的关系。

还是对于可能性 ψ ，如果我们让它作为系统的初态可能性，则按照上面讲的，有

$$|\psi\rangle = \sum_i a_i|i\rangle, \quad a_i = \langle i|\psi\rangle. \quad (3.25)$$

但是，假设我们不是把系统制备到 ψ 上，而是检测到系统处于 ψ ，也就是说，它是作为一个末态可能性. 则就有

$$\langle\psi| = \sum_i \langle\psi|i\rangle\langle i|. \quad (3.26)$$

¹之所以打引号，是因为同一方向的不同倍数是一个复数倍数，不是实数，所以这是一条复直线。

代入 $\langle\psi|i\rangle = \langle i|\psi\rangle^* = a_i^*$ (跃迁幅的性质)即有

$$\langle\psi| = \sum_i a_i^* \langle i|. \quad (3.27)$$

这就是与初态可能性 $|\psi\rangle$ 对应的末态可能性，其中叠加权重要取复数共轭。

同样的，也可以在末态可能性之间定义叠加运算，使得它们成为态矢量，不过为了和初态可能性对应的态矢量进行区分，通常我们称 $|\psi\rangle$ 这样的态矢量为右矢，而称与之相应的 $\langle\psi|$ 为它的共轭左矢。注意，左矢和右矢是以一一对应的，只不过，右矢用来描述初态可能性，而左矢则用来描述末态可能性。而且，右矢与右矢可以叠加，左矢与左矢也可以叠加，但是，右矢和左矢是不能叠加的！

任给一个右矢 $|\psi\rangle$ ，一个左矢 $\langle\phi|$ ，我们称从 ψ 到 ϕ 的跃迁幅为这两个矢量之间的一种乘法，称为内积，依然记作 $\langle\phi|\psi\rangle$ 。因此，根据跃迁幅乘法，我们有

$$\langle\phi|\psi\rangle = \sum_i \langle\phi|i\rangle \langle i|\psi\rangle. \quad (3.28)$$

另外，根据跃迁幅的性质，当然也有

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*. \quad (3.29)$$

不过，这里有一个细节值得一提，如果真是上一章处理的跃迁幅，那 ψ 和 ϕ 都必定是系统的某个可能性，必定分别属于某两个可确定区分可能性完备集，从而作为可确定区分性的要求，必定有

$$\langle\psi|\psi\rangle = \langle\phi|\phi\rangle = 1. \quad (3.30)$$

但是现在，作为态矢量 $|\psi\rangle$ 可以乘以任意的倍数 λ ，从而成为 $\lambda|\psi\rangle$ ，相应的共轭左矢则是 $\lambda^*\langle\psi|$ ，左矢和右矢的内积就成了 $|\lambda|^2\langle\psi|\psi\rangle$ ，由于 λ 任意，它当然不一定等于1！换言之，态矢量内积的概念其实相比原来跃迁幅的概念略有拓展，因为它可以允许一个态矢量与其共轭左矢的内积不等于1，这时候我们就称 $\langle\psi|\psi\rangle$ 为矢量 $|\psi\rangle$ 的模长平方。不过，物理上，由于 $\lambda|\psi\rangle$ 和 $|\psi\rangle$ 等价，所以我们总是可以通过适当调节这个 λ ，使得 $\langle\psi|\psi\rangle = 1$ ，这个过程就叫将态矢量进行归一化。相应的，(3.30)式就称为归一化条件。归一化以后的态矢量就是模长为1的态矢量。

归一化条件并不能唯一确定态矢量，因为如果 $|\psi\rangle$ 满足归一化条件，则 $e^{i\theta}|\psi\rangle$ 显然也满足归一化条件，而 $|\psi\rangle$ 与 $e^{i\theta}|\psi\rangle$ 在物理上则是等价的。通常称这个定不下来的 $e^{i\theta}$ 为相因子，因此，归一化条件只能把态矢量确定到相差一个任意的相因子。

内积运算还有一条基本的性质，为了说清楚它，依然让我们考虑量子态 $|\psi\rangle$ 和 $|\phi\rangle$ ，利用跃迁幅乘法规则并将中间可能性取在完备集 \mathcal{I} 中，有

$$\langle\phi|\psi\rangle = \sum_{i \in \mathcal{I}} \langle\phi|i\rangle \langle i|\psi\rangle. \quad (3.31)$$

记 $\langle i|\psi\rangle = a_i$ ，从而 $|\psi\rangle = \sum_{i \in \mathcal{I}} a_i|i\rangle$ ，代入上式左边，即有

$$\langle\phi| \sum_{i \in \mathcal{I}} a_i|i\rangle = \sum_{i \in \mathcal{I}} \langle\phi|i\rangle a_i = \sum_{i \in \mathcal{I}} a_i \langle\phi|i\rangle. \quad (3.32)$$

左边第一个式子是先进行线性叠加然后再进行内积运算，而右边最后一个式子则是先进行内积运算再进行线性叠加，因此这个结果清楚地告诉我们，对于右矢而言，内积运算和线性叠加运算可以交换顺序。有时候也称这个性质为，内积运算关于右矢是线性的。

严格来说，希尔伯特空间不仅是指态矢量所构成的矢量空间，还要求态矢量之间能够进行内积！只有定义了内积的态矢量空间才是希尔伯特空间。通常记一个量子系统的希尔伯特空间为 \mathcal{H} 。

正如前面自旋的例子中提到过的，任给两个态矢量 $|\psi\rangle, |\phi\rangle$ ，如果 $\langle\phi|\psi\rangle = 0$ ，我们就称这两个态矢量正交！按照上一章讲过的不可区分原理，对于任意两个不同量子态 $|\psi\rangle, |\phi\rangle$ ，当且仅当它们正交时，两者才可以百分之百地区分，或者说才可以确定地区分。因此用现在的术语来说，可确定区分可能性完备集 \mathcal{I} 的可确定区分性 $\langle i|j\rangle = \delta_{ij}$ 就叫做正交归一性，也就是集合中任意两个不同的态矢量正交，同一个态矢量则归一。用现在的术语，常常也将与 \mathcal{I} 中那些可能性(i)对应的态矢量 $|i\rangle$ 称作满足正交归一关系的基矢量。根据(3.21)式，任意态矢量 $|\psi\rangle$ 都可以由这些基矢量叠加而成。

当集合 \mathcal{I} 为一个有限集时，可以证明其元素的个数(也就是正交归一基矢量的个数)是一个独立于 \mathcal{I} 的不同选取的不变量²称之为希尔伯特空间的维数，记作 $\dim(\mathcal{H})$ 。而当集合 \mathcal{I} 为无限集时，则称希尔伯特空间为无穷维的，即 $\dim(\mathcal{H}) = +\infty$ 。

²为了证明这个结论，不妨任选两个可确定区分可能性完备集 \mathcal{I}, \mathcal{J} ，设 \mathcal{I} 的元素个数为 I ，

所以，通过对跃迁幅以及跃迁幅乘法进行适当的解读，我们构造出了量子态的概念，也构造出了态矢量和希尔伯特空间的概念。但是，我们也别忘了，真正有物理意义的，依然是跃迁幅和跃迁幅乘法。从这个意义上说，态矢量是虚构的东西，如此看来，通常人们所谓的量子态塌缩当然也是虚构的！

3.2 算符

以上我们考察了如何从跃迁幅以及跃迁幅乘法中构造出态矢量以及希尔伯特空间。下面来进一步考察由更具一般性的跃迁元和跃迁元乘法中能构造出什么。

为此，考察给定初态可能性 ψ 经物理量 A 作用的结果，同样，假设我们不关心最终跃迁到的末态可能性，从而可以把这个过程的跃迁元写成

$$\langle \dots | A | \psi \rangle, \quad (3.35)$$

作为一个复数，假设这个跃迁元可以等效地由跃迁幅 $\langle \dots | \phi \rangle$ 给出，从而即有

$$\langle \dots | A | \psi \rangle = \langle \dots | \phi \rangle. \quad (3.36)$$

由于我们不关心末态可能性，所以干脆，我们将它省略不写，从而即有如下不完整的开放的方程

$$A|\psi\rangle = |\phi\rangle. \quad (3.37)$$

这个方程可以自然地解读成，任给一个量子态 $|\psi\rangle$ ，则在物理量 A 的作用之下，就会得到一个新的量子态 $|\phi\rangle$ ，一般来说 $|\psi\rangle$ 不同，则最终得到的 $|\phi\rangle$ 也不

$\sum_{i \in \mathcal{I}, j' \in \mathcal{J}} \langle i | j' \rangle \langle j' | i \rangle$

的元素个数为 J ，下面考虑 $\sum_{i \in \mathcal{I}, j' \in \mathcal{J}} \langle i | j' \rangle \langle j' | i \rangle$ ，由跃迁幅乘法，有

$$\sum_{i \in \mathcal{I}, j' \in \mathcal{J}} \langle i | j' \rangle \langle j' | i \rangle = \sum_{i \in \mathcal{I}} \langle i | i \rangle = \sum_{i \in \mathcal{I}} 1 = I, \quad (3.33)$$

式中我们用到 $\langle i | i \rangle = 1$ (可确定区分性)。另一方面，同一个量调整一下相乘两因子的顺序，则有

$$\sum_{i \in \mathcal{I}, j' \in \mathcal{J}} \langle j' | i \rangle \langle i | j' \rangle = \sum_{j' \in \mathcal{J}} \langle j' | j' \rangle = \sum_{j' \in \mathcal{J}} 1 = J. \quad (3.34)$$

既然这是同一个量的两种不同计算方式，所以 $I = J$. 证明完毕.

同。所以，我们可以把物理量 A 看作一台“机器”，每给这台机器输入一个 $|\psi\rangle$ ，则它就会输出一个 $|\phi\rangle$ ，这样的“机器”就叫做算符。所以，物理量对系统的作用可以解释成算符对量子态的作用。

回到上面的(3.36)式，假设我们在等式左边应用 $A = A \cdot 1$ 的跃迁元乘法，并且在可确定区分可能性完备集 \mathcal{I} 中取中间可能性，则有

$$\langle \dots | A | \psi \rangle = \sum_{i \in \mathcal{I}} \langle \dots | A | i \rangle \langle i | \psi \rangle. \quad (3.38)$$

记 $\langle i | \psi \rangle = a_i$ ，从而 $|\psi\rangle = \sum_{i \in \mathcal{I}} a_i |i\rangle$ ，代入上式左边，即有

$$\langle \dots | A \sum_{i \in \mathcal{I}} a_i | i \rangle = \sum_{i \in \mathcal{I}} \langle \dots | A | i \rangle a_i = \sum_{i \in \mathcal{I}} a_i \langle \dots | A | i \rangle. \quad (3.39)$$

从这个式子中省略末态可能性，即有

$$A \sum_{i \in \mathcal{I}} a_i | i \rangle = \sum_{i \in \mathcal{I}} a_i (A | i \rangle). \quad (3.40)$$

等式左边是先进行量子态的线性叠加运算，再用算符 A 进行作用，而等式右边则是先用算符 A 进行作用，再做线性叠加运算。所以，这个结果告诉我们，算符 A 的作用和线性叠加运算可以交换顺序！

这样的算符就称作线性算符。由于上面考察的 A 是任意的，所以这就说明，在量子力学中，每一个物理量都会对应一个线性算符。这个线性算符是作用在态矢量上的，或者说是定义在希尔伯特空间 \mathcal{H} 上的。

比如在上一节的自旋例子中，下面的方程就定义了一个线性算符 X (请和粒子的坐标 X 相区分，这根据上下文不难做到)

$$X | \uparrow \rangle \equiv | \downarrow \rangle, \quad X | \downarrow \rangle \equiv | \uparrow \rangle. \quad (3.41)$$

根据这个定义，可以计算出 X 在任意线性叠加态 $c_1 | \uparrow \rangle + c_2 | \downarrow \rangle$ 上的作用，如下

$$X(c_1 | \uparrow \rangle + c_2 | \downarrow \rangle) = c_1 X | \uparrow \rangle + c_2 X | \downarrow \rangle = c_1 | \downarrow \rangle + c_2 | \uparrow \rangle. \quad (3.42)$$

其中第一个等于号是线性算符的性质。

为了具体计算算符 A 在量子态 $|\psi\rangle$ 上的作用，即具体应用方程 $A|\psi\rangle = |\phi\rangle$ ，我们常常需要选取一个表象 \mathcal{I} ，并考虑 $\langle i|A|\psi\rangle = \langle i|\phi\rangle$ ，在等式左边应用跃迁元乘法，即有

$$\sum_{j \in \mathcal{I}} \langle i|A|j\rangle \langle j|\psi\rangle = \langle i|\phi\rangle. \quad (3.43)$$

人们常常把这个方程排成如下矩阵形式

$$\begin{bmatrix} \langle 1|A|1\rangle & \langle 1|A|2\rangle & \dots \\ \langle 2|A|1\rangle & \langle 2|A|2\rangle & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \vdots \end{bmatrix} = \begin{bmatrix} \langle 1|\phi\rangle \\ \langle 2|\phi\rangle \\ \vdots \end{bmatrix}. \quad (3.44)$$

为了计算等式左边的矩阵乘法，我们需要把第一个矩阵的第*i*行和第二个矩阵这单独的一列相乘，将结果作为第三个矩阵第*i*行的那个元素。单独只有一列的矩阵也叫做列矢量。以上式子告诉我们，为了计算算符在量子态上的作用，我们只需要先计算出算符在某个表象 \mathcal{I} 中的表示矩阵，再把相应的量子态表示成列矢量，然后计算算符的这个矩阵与这个列矢量相乘的结果。这种方法有一个巨大的好处，就是你只需要计算一次算符的表示矩阵，然后当考虑它在不同量子态上的作用时，你就把这同一个矩阵与对应量子态的不同列矢量相乘就可以了。

比如上面对于自旋例子定义的 X 算符，其在 $\{\uparrow, \downarrow\}$ 表象中的表示矩阵即是

$$\begin{bmatrix} \langle \uparrow |X|\uparrow\rangle & \langle \uparrow |X|\downarrow\rangle \\ \langle \downarrow |X|\uparrow\rangle & \langle \downarrow |X|\downarrow\rangle \end{bmatrix} = \begin{bmatrix} \langle \uparrow|\downarrow\rangle & \langle \uparrow|\uparrow\rangle \\ \langle \downarrow|\downarrow\rangle & \langle \downarrow|\uparrow\rangle \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad (3.45)$$

式中第一个等于号利用了 X 的定义，第二个等于号利用了基矢量的正交归一性。

特别的，物理量1对应的算符就称作单位算符，依然记作1，不难证明，在任何表象中，单位算符的表示矩阵都是单位矩阵(请读者想想为什么)。

前面定义算符的时候，我们都是让算符作用在右矢上，因为我们是通过忽略末态可能性来定义算符的。但是完全一样的，也可以忽略跃迁元的初态可能性，进而定义算符在左矢上的作用，也即是说， $\langle \phi|A$ 这样的表达式也是有意义的，它依然是一个左矢，其含义就在于

$$(\langle \phi|A)|\psi\rangle \equiv \langle \phi|(A|\psi\rangle) = \langle \phi|A|\psi\rangle. \quad (3.46)$$

也即是说，跃迁元中的算符，无论理解成是作用在右矢上，还是理解成是作用在左矢上，结果是等价的。

至于算符 A 与算符 B 的乘积 AB ，其定义就是先用算符 B 对任意量子态 $|\psi\rangle$ 进行作用，然后接着再用算符 A 进行作用。具体计算的时候当然还是根据跃迁元乘法来乘，或者在选定了某个表象以后，算符的相乘就转换成了相应表示矩阵的矩阵相乘。这些我们在上一章实际都已经交待过了，只是那时候没提算符的概念。上面提到单位算符，当然，单位算符1乘以任何算符，结果仍然是这个算符。

另外，上一章的最后实际上也已经定义了算符的加法，以及算符的数乘，即对两个算符 A, B 定义了 $A + B$ ，以及 λA ，虽然那时候还没引入算符的概念，但是我们已经通过跃迁元定义这两种运算了，由于算符的作用不过就是跃迁元省略掉末态可能性，所以那里的定义当然可以照搬到这里。

有了加法，数乘以及算符乘法，那希尔伯特空间上所有线性算符在这几种运算之下就构成一个封闭的算符代数。

任给一个线性算符(不一定是物理量算符) \mathcal{O} ，都有一个相应的厄米共轭算符，记作 \mathcal{O}^\dagger ，其定义如下

$$\langle \phi | \mathcal{O}^\dagger | \psi \rangle = \langle \psi | \mathcal{O} | \phi \rangle^*. \quad (3.47)$$

式中 ψ, ϕ 是系统的任意两个可能性(任意两个态)。或者示意性地记作

$$\left\{ \phi \xleftarrow{\mathcal{O}^\dagger} \psi \right\} = \left\{ \phi \xrightarrow{\mathcal{O}} \psi \right\}^*. \quad (3.48)$$

因此不难看出，

$$\left\{ \phi \xleftarrow{(\mathcal{O}^\dagger)^\dagger} \psi \right\} = \left\{ \phi \xrightarrow{\mathcal{O}^\dagger} \psi \right\}^* = \left\{ \phi \xleftarrow{\mathcal{O}} \psi \right\}. \quad (3.49)$$

式中最后一个等于号用到复数共轭的复数共轭等于本身。这个结果清楚地告诉我们

$$(\mathcal{O}^\dagger)^\dagger = \mathcal{O}. \quad (3.50)$$

即厄米共轭的厄米共轭等于原算符本身。

根据厄米共轭算符的这个定义，可以得到，如果 $\langle \phi |$ 是与右矢 $|\phi\rangle$ 对应的共轭左矢，则 $\langle \phi | \mathcal{O}^\dagger$ 就是 $\mathcal{O} | \phi \rangle$ 的共轭左矢。这是因为，根据厄米共轭算

符的定义式(3.47), 其等式左边可以理解成 $\langle\phi|\mathcal{O}^\dagger$ 与 $|\psi\rangle$ 的内积, 而等式右边则可以理解成 $\langle\psi|\mathcal{O}|\phi\rangle$ 的内积的复数共轭, 根据内积符号的性质, 它也就是 $\mathcal{O}|\phi\rangle$ 的共轭左矢与 $|\psi\rangle$ 的内积, 等式左边等于等式右边, 因此 $\langle\phi|\mathcal{O}^\dagger$ 就是 $\mathcal{O}|\phi\rangle$ 的共轭左矢。

任给两个算符 $\mathcal{O}_1, \mathcal{O}_2$, 不难证明

$$(\mathcal{O}_1 \mathcal{O}_2)^\dagger = \mathcal{O}_2^\dagger \mathcal{O}_1^\dagger. \quad (3.51)$$

证明过程其实和第一章就厄米共轭矩阵相应结果的证明很类似。具体证明如下

$$\begin{aligned} \langle\phi|(\mathcal{O}_1 \mathcal{O}_2)^\dagger|\psi\rangle &= \langle\psi|\mathcal{O}_1 \mathcal{O}_2|\phi\rangle^* = \sum_{i \in \mathcal{I}} \langle\psi|\mathcal{O}_1|i\rangle^* \langle i|\mathcal{O}_2|\phi\rangle^* \\ &= \sum_{i \in \mathcal{I}} \langle i|\mathcal{O}_1^\dagger|\psi\rangle \langle\phi|\mathcal{O}_2^\dagger|i\rangle = \sum_{i \in \mathcal{I}} \langle\phi|\mathcal{O}_2^\dagger|i\rangle \langle i|\mathcal{O}_1^\dagger|\psi\rangle = \langle\phi|\mathcal{O}_2^\dagger \mathcal{O}_1^\dagger|\psi\rangle. \end{aligned} \quad (3.52)$$

式中第一行的最后一个等于号是利用跃迁元乘法, 第二行的第一个等于号是利用厄米共轭算符的定义, 最后一个等于号是再一次利用跃迁元乘法。

另一方面, 根据物理量 A 的厄米条件 $\langle\phi|A|\psi\rangle = \langle\psi|A|\phi\rangle^*$ (式中 ψ, ϕ 为分别取在某两个可确定区分可能性完备集中的可能性), 与刚才关于厄米共轭算符的定义比较可知, 这个厄米条件可以表达成

$$A^\dagger = A. \quad (3.53)$$

这样的算符就叫做厄米算符, 即其厄米共轭算符是自身的那种算符。所以, 物理量的厄米条件告诉我们, 与物理量相对应的算符必定为厄米算符。

另一方面, 由于定义了乘法、加法以及数乘, 所以两个算符的对易子也就有定义了, 即是 $[A, B] \equiv AB - BA$ 。不难证明, 算符对易子满足如下恒等式

$$\begin{aligned} [A, BC] &= [A, B]C + B[A, C] \\ [A, [B, C]] + [B, [C, A]] + [C, [A, B]] &= 0. \end{aligned} \quad (3.54)$$

这两个恒等式形式上和第一章给出的矩阵的相应恒等式一模一样, 证明也完全一样。其中第二个恒等式就叫做雅可比恒等式。不难验证, 如果 A 为厄米算符, B 也为厄米算符, 则 $i[A, B]$ 也必定为厄米算符! 所以, 所有厄米

算符在求对易子并乘以虚数单位*i*的运算，以及加法和数乘运算之下，构成一个封闭的李代数。总之，第一章关于矩阵代数(跃迁元数据表的代数)的很多相应叙述都可以照搬到算符上来，这里不再赘述。

前面我们说过，左矢乘以右矢就是希尔伯特空间内积。那么，右矢乘以左矢是什么呢？也就是说，下面这样的式子含义是什么呢

$$|u\rangle\langle v|, \quad (3.55)$$

式中 $|u\rangle$ 为某个右矢， $\langle v|$ 为某个左矢，注意在这个式子中，右矢在左边，左矢在右边，这个顺序不能颠倒。我们发现，对这种式子的自然解释是，把它理解成一个算符，为此我们定义

$$(|u\rangle\langle v|)|\psi\rangle = |u\rangle(\langle v| \cdot |\psi\rangle) = |u\rangle(\langle v|\psi\rangle), \quad (3.56)$$

我们把上面这样的推导过程看作是利用了左矢、右矢混合乘积的结合律！由于最后式中的 $\langle v|\psi\rangle$ 为一个复数，所以最后的式子当然是一个正比于 $|u\rangle$ 的量子态。所以这个式子的含义就是， $|u\rangle\langle v|$ 作用在量子态 $|\psi\rangle$ 上，最终得到了一个正比于 $|u\rangle$ 的量子态，既然是把量子态映射到量子态，那 $|u\rangle\langle v|$ 就是一个算符，而且，由于内积运算关于右矢的线性性质，不难验证，这个算符还是一个线性算符。

特别的，前面根据跃迁幅乘法写过如下式子

$$|\psi\rangle = \sum_{i \in \mathcal{I}} |i\rangle\langle i|\psi\rangle, \quad (3.57)$$

那现在，将结合律用上，这个式子就可以重新解释为

$$|\psi\rangle = \left(\sum_{i \in \mathcal{I}} |i\rangle\langle i| \right) |\psi\rangle. \quad (3.58)$$

也即是说，算符($\sum_{i \in \mathcal{I}} |i\rangle\langle i|$)作用在任何态 $|\psi\rangle$ 上，结果都还是 $|\psi\rangle$ 本身，那这个算符就是一个单位算符。所以，我们有

$$\sum_{i \in \mathcal{I}} |i\rangle\langle i| = 1. \quad (3.59)$$

上面这个式子可以让我们重新理解跃迁元(以及跃迁幅)乘法，比方说根据跃迁元乘法，我们有

$$\langle \phi | AB | \psi \rangle = \sum_{i \in \mathcal{I}} \langle \phi | A | i \rangle \langle i | B | \psi \rangle, \quad (3.60)$$

现在，我们可以把它重新理解成

$$\langle \phi | AB |\psi \rangle = \langle \phi | A \left(\sum_{i \in \mathcal{I}} |i\rangle \langle i| \right) B |\psi \rangle, \quad (3.61)$$

由于中间插入的这个算符是单位算符，所以这个式子当然是成立的。换言之，跃迁元乘法可以理解成是在算符 A, B 中间插入了单位算符1，然后再把单位算符1用(3.59)式来表达！值得强调的是，虽然从数学形式上这是一种完全等价的看待问题的方式，但是，本书的逻辑是完全反过来的，我们是先从海森堡的原创想法中引入了跃迁元乘法，然后在第二章中从物理上把这种乘法推广成普遍形式，然后本章我们再表明，通常的那些数学形式，包括量子态概念、量子态叠加、希尔伯特空间、以及算符的概念、尤其是以及(3.59)式，这一切都可以从物理上的跃迁元概念以及跃迁元乘法中自然地引导出来。

构造抽象态矢量和算符的最大好处是，它们深刻地把握住了量子力学内在数学结构的表象无关性，代价就是如果过度延伸这些抽象构造的话，就必然会引入一些非物理的东西。

3.3 海森堡绘景、薛定谔绘景以及量子力学幺正性

3.3.1 海森堡绘景与薛定谔绘景

利用算符的概念，就可以把上一章最后得到的，与表象无关的海森堡运动方程以及量子力学基本对易关系重新用算符的形式写出来，其实上一章已经写了，不过由于那时候我们还没有定义算符的概念，所以只是在形式上写的，现在有了算符的概念，这些写法才获得了真正数学上的意义，这里不妨重写一下。海森堡运动方程是

$$i\hbar \frac{dA}{dt} = [A, H]. \quad (3.62)$$

式中 H 就是哈密顿算符。量子力学基本对易关系是

$$[X, P] = i\hbar. \quad (3.63)$$

海森堡运动方程(3.62)有一个形式上的通解，我们先写出这个通解，之后再进行解释和验证

$$A(t) = e^{iHt/\hbar} A_0 e^{-iHt/\hbar}, \quad (3.64)$$

式中 A_0 是一个不随时间演化的算符。特别的，单位算符 1 自动满足海森堡运动方程(3.62)，它当然不随时间演化，所以，在上式中代入 $A_0 = A(t) = 1$ ，即有

$$1 = e^{iHt/\hbar} e^{-iHt/\hbar}. \quad (3.65)$$

首先，我们解释一下把哈密顿算符放到指数上是什么意思，也就是 e^B (式中 B 为某个算符)这样的表达式是什么含义，其含义就是

$$e^B \equiv 1 + \frac{1}{1!}B + \frac{1}{2!}B^2 + \frac{1}{3!}B^3 + \dots \quad (3.66)$$

也即是说，这样的式子是按照指数函数的级数展开来理解的。根据这个定义，不难验证(请读者自己动手一下)

$$e^B e^{-B} = 1. \quad (3.67)$$

从而就具体验证了(3.65)式。另外，也不难验证 $B e^B = e^B B$ ，从而

$$e^B B e^{-B} = B. \quad (3.68)$$

并且，根据上面那个定义，也不难得到

$$\frac{d}{dt} e^{iHt/\hbar} = \frac{i}{\hbar} H e^{iHt/\hbar} = \frac{i}{\hbar} e^{iHt/\hbar} H. \quad (3.69)$$

从而将表达式(3.64)对 t 求导，根据求导的莱布尼兹法则，即可以得到

$$\begin{aligned} i\hbar \frac{dA}{dt} &= -e^{iHt/\hbar} H A_0 e^{-iHt/\hbar} + e^{iHt/\hbar} A_0 H e^{-iHt/\hbar} \\ &= -e^{iHt/\hbar} H e^{-iHt/\hbar} e^{iHt/\hbar} A_0 e^{-iHt/\hbar} + e^{iHt/\hbar} A_0 e^{-iHt/\hbar} e^{iHt/\hbar} H e^{-iHt/\hbar} \\ &= -H A(t) + A(t) H = [A, H]. \end{aligned} \quad (3.70)$$

式中第二行的等于号是利用了(3.67)式(从而 $e^{-iHt/\hbar} e^{iHt/\hbar} = 1 = e^{iHt/\hbar} e^{-iHt/\hbar}$)，第三行则是利用了(3.68)式(从而 $e^{iHt/\hbar} H e^{-iHt/\hbar} = H$)。所以，我们就验证了(3.64)式的确是(3.62)式的通解。

另外，根据指数算符的级数展开定义，不难得到

$$(e^B)^\dagger = e^{B^\dagger}. \quad (3.71)$$

由于哈密顿算符是厄米算符 $H^\dagger = H$ ，而且 $i^\dagger = -i$ ，所以

$$(e^{-iHt/\hbar})^\dagger = e^{iHt/\hbar}. \quad (3.72)$$

到此为止，在我们的观点中，都是认为量子态并不随时间演化（因为最开始引入量子态的概念，即量子系统的可能性，本来就不依赖于时间），随时间演化的是代表物理量的算符。在量子力学中，这种观点叫做海森堡绘景。实际上，还有一种与之等价的观点，即，认为物理量算符不随时间演化，而是量子态在随时间演化，这种观点就叫做薛定谔绘景。

为了从海森堡绘景过渡到薛定谔绘景，我们考察如下跃迁元（注意，只有跃迁元才是物理的，量子态和算符都是人为构造出来的），

$$\langle \phi | A(t) | \psi \rangle. \quad (3.73)$$

代入(3.64)式即有

$$\langle \phi | e^{iHt/\hbar} A_0 e^{-iHt/\hbar} | \psi \rangle. \quad (3.74)$$

很显然，如果我们定义如下随时间演化的量子态

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi\rangle, \quad |\phi(t)\rangle = e^{-iHt/\hbar} |\phi\rangle. \quad (3.75)$$

则(3.74)式就可以写成 $\langle \phi(t) | A_0 | \psi(t) \rangle$ （注意 $|\phi(t)\rangle = e^{-iHt/\hbar} |\phi\rangle$ 的共轭左矢 $\langle \phi(t) |$ 是 $\langle \phi | (e^{-iHt/\hbar})^\dagger = \langle \phi | e^{iHt/\hbar}$ ），换言之，

$$\langle \phi | A(t) | \psi \rangle = \langle \phi(t) | A_0 | \psi(t) \rangle. \quad (3.76)$$

也即是说，同一个物理的跃迁元，我们既可以认为，量子态不随时间演化，是物理量算符 $A(t)$ 在随时间演化，也可以等价地认为，物理量算符始终是不随时间演化的 A_0 ，随时间演化的是量子态，它们在按照(3.75)式随时间演化，后一种观点就叫做薛定谔绘景。

总之，在薛定谔绘景中，量子态按照 $|\psi(t)\rangle = e^{-iHt/\hbar}|\psi\rangle$ 的规律随时间演化，而算符则是不演化的 A_0 ，它和海森堡绘景算符的关系由(3.64)式给出。不难将薛定谔绘景的量子态对时间求导，并得到如下方程

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle. \quad (3.77)$$

这就是著名的薛定谔方程。

具体应用时，人们常常会选取一个表象 \mathcal{I} ，并把薛定谔方程在这个表象中表示出来，即

$$i\hbar \frac{d}{dt}\langle i|\psi(t)\rangle = \sum_{j \in \mathcal{I}} \langle i|H|j\rangle \langle j|\psi(t)\rangle. \quad (3.78)$$

假设记 $C_i(t) = \langle i|\psi(t)\rangle$ ，并把上述方程排列成矩阵的形式即

$$i\hbar \frac{d}{dt} \begin{bmatrix} C_1(t) \\ C_2(t) \\ \vdots \end{bmatrix} = \begin{bmatrix} \langle 1|H|1\rangle & \langle 1|H|2\rangle & \dots \\ \langle 2|H|1\rangle & \langle 2|H|2\rangle & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} C_1(t) \\ C_2(t) \\ \vdots \end{bmatrix}. \quad (3.79)$$

根据海森堡绘景算符与薛定谔绘景算符之间的关系，不难得出，在薛定谔绘景中，量子力学基本对易关系为

$$[X_0, P_0] = i\hbar. \quad (3.80)$$

式中 X_0, P_0 分别为薛定谔绘景中的坐标算符和动量算符。

3.3.2 量子力学幺正性

人们常常记

$$U(t) = e^{-iHt/\hbar}, \quad (3.81)$$

并称之为时间演化算符，因为在薛定谔绘景中，它作用在0时刻的量子态 $|\psi\rangle$ 上，就会得到演化到t时刻的量子态 $|\psi(t)\rangle$ ，即 $|\psi(t)\rangle = U(t)|\psi\rangle$ 。很显然， $U(t)$ 并不是一个厄米算符，其厄米共轭 $U^\dagger(t) = e^{iHt/\hbar}$ 。根据(3.67)式，不难得得到

$$U^\dagger U = UU^\dagger = 1. \quad (3.82)$$

这样的算符就称作幺正算符。因此人们也常常说，量子系统是随时间幺正演化的。量子系统的这一性质就称之为幺正性。

那么量子系统的幺正性意味着什么呢？回答是意味着量子态的希尔伯特空间内积不随时间变化，也就是任给 $|\psi(t)\rangle = U(t)|\psi\rangle$, $|\phi(t)\rangle = U(t)|\phi\rangle$, 则

$$\langle\phi(t)|\psi(t)\rangle = \langle\phi|U^\dagger U|\psi\rangle = \langle\phi|\psi\rangle. \quad (3.83)$$

特别的，在(3.83)式中取 $|\phi\rangle = |\psi\rangle$, 并假设 $|\psi\rangle$ 已经归一化，则(3.83)式告诉我们

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi|\psi\rangle = 1, \quad (3.84)$$

即时间演化会保持量子态的归一化，初始归一的量子态将永远自动归一。而如果在(3.83)式中取 $\langle\phi|\psi\rangle = 0$, 即两态正交，则(3.83)式告诉我们这两个态将永远保持正交。由于正交性在物理上意味着可确定区分，所以这个结果意味着，两个可确定区分的量子态任何时刻都将是可确定区分的！这两点可以概括成，幺正性意味着时间演化保持量子系统任何可确定区分可能性完备集的可确定区分性。

反过来，如果时间演化保持任何可确定区分可能性完备集的可确定区分性，那就必然意味着相应的时间演化算符是幺正算符。为了说明这一点，假设任取一个可确定区分可能性完备集 \mathcal{I} , 可确定区分性意味着

$$\langle i|j\rangle = \delta_{ij}, \quad i, j \in \mathcal{I}. \quad (3.85)$$

另一方面，在薛定谔绘景中， $|i\rangle$ 将演化成 $U(t)|i\rangle$, 因此如果时间演化保持可确定区分性，那就意味着

$$\langle i|U^\dagger(t)U(t)|j\rangle = \delta_{ij}, \quad (3.86)$$

这也就是说，算符 $U^\dagger(t)U(t)$ 在表象 \mathcal{I} 中的表示矩阵是一个单位矩阵，但是，只有单位算符的表示矩阵才是单位矩阵。所以，这反过来说明 $U^\dagger(t)U(t)$ 是一个单位算符，即

$$U^\dagger(t)U(t) = 1. \quad (3.87)$$

这就说明，时间演化算符 $U(t)$ 必定是一个幺正算符³。

其实，如果是在海森堡绘景中，那么(3.83)式就特别简单，因为它描述的不过就是跃迁幅不随时间变化，而跃迁幅不过就是物理量1的跃迁元，而在海森堡绘景中，物理量1本来就不随时间演化，它的跃迁元当然也不随时间变化。

实际上，在海森堡绘景中，幺正性常常还指另一件事情。为了说清楚这件事情，我们需要取两个不同的可确定区分可能性完备集 $\mathcal{I} = \{(i), i = 1, 2, 3, \dots\}$, $\mathcal{I}' = \{(i'), i' = 1', 2', 3', \dots\}$, 考虑跃迁幅 $\langle j'|i\rangle$, 则根据跃迁幅乘法，我们有

$$\begin{aligned}\sum_{j' \in \mathcal{I}'} \langle k|j'\rangle \langle j'|i\rangle &= \langle k|i\rangle = \delta_{ki}, \\ \sum_{j \in \mathcal{I}} \langle k'|j\rangle \langle j|i'\rangle &= \langle k'|i'\rangle = \delta_{k'i'}.\end{aligned}\quad (3.88)$$

记 $S_{j'i} = \langle j'|i\rangle$, 并注意到 $\langle i|j'\rangle = \langle j'|i\rangle^* = S_{j'i}^*$, 就能把上面的结果写成

$$\sum_{j' \in \mathcal{I}'} S_{j'k}^* S_{j'i} = \delta_{ki}, \quad \sum_{j \in \mathcal{I}} S_{k'j} S_{i'j}^* = \delta_{k'i'}.\quad (3.89)$$

$S_{j'i}$ 当然描述的是 \mathcal{I} 到 \mathcal{I}' 的变换关系，在海森堡绘景中，上面这个关于这一变换关系的方程也叫做幺正性。这是因为，如果把 $S_{j'i}$ 看成一个矩阵的话，那么上面这个方程说的就是，这个变换矩阵是一个幺正矩阵！

量子不可克隆定理

下面我们利用薛定谔绘景讲述一个重要而有趣的结果，叫做量子不可克隆定理。它大体说的是，不可能克隆(复制)一个任意的未知量子态，比如说，如果不能事先确定你的量子态，那就不可能克隆出一个一模一样的你，不是技术无法达到，而是在物理原理上就不可能。

为了讲清楚这个定理，我们需要做一些准备工作。第一个准备就是如下引理：如果 $|\psi\rangle, |\phi\rangle$ 是系统任意两个已经归一化的量子态，则

$$|\langle\phi|\psi\rangle| \leq 1,\quad (3.90)$$

³严格来说，为了进一步导出幺正算符需要满足的另一个条件 $UU^\dagger = 1$, 我们还需要假设时间演化可逆。

等于号仅当 $|\phi\rangle$ 和 $|\psi\rangle$ 是同一个量子态时才成立。

下面给出一个物理上的证明，或者不叫证明，叫物理上的论证吧。由于 $|\psi\rangle, |\phi\rangle$ 都已经归一化，所以我们可以将它们所描写的量子可能性分别归入两个可确定区分可能性完备集，比方说 $\psi \in \mathcal{I} = \{(i), i = 1, 2, 3, \dots\}$, $\phi \in \mathcal{J} = \{(j'), j' = 1', 2', 3', \dots\}$, 当然 \mathcal{I}, \mathcal{J} 可能相同。则根据上一章讲述的总概率等于1可知

$$\sum_{j' \in \mathcal{J}} |\langle j'|i\rangle|^2 = 1, \quad (3.91)$$

根据 \mathcal{I}, \mathcal{J} 的取法， $|\langle \phi | \psi \rangle|^2$ 必然是上面这个求和中对应某个*i*和某个*j'*的其中一项，既然求和中的每一项都大于等于零，而对所有*j'*的求和才等于1，那这其中的一项当然必定小于等于1，从而

$$|\langle \phi | \psi \rangle|^2 \leq 1 \Leftrightarrow |\langle \phi | \psi \rangle| \leq 1. \quad (3.92)$$

当等于号成立时，那就说明(3.91)式的求和中只有一项有贡献，相应的 $|i\rangle$ (即是 $|\psi\rangle$)和这一项之外的其余 $|j'\rangle$ (即 $|\phi\rangle$ 之外的 $|j'\rangle$)均正交，根据集合 \mathcal{J} 的可确定区分性，这就说明 $|\psi\rangle$ 也属于集合 \mathcal{J} ，而且实际上和 $|\phi\rangle$ 描述的是同一个量子可能性，所以，这时候 $|\phi\rangle$ 和 $|\psi\rangle$ 是同一个量子态。

我们要做的第二个准备工作是，要说清楚什么叫克隆(复制)。所谓的克隆，就是有两个量子系统，我们把其中一个系统的量子态复制到另一个系统上。比方说，让这两个系统中的第一个处于量子态 $|\psi\rangle$ ，让第二个最初处于某个参考量子态，不妨记作 $|0\rangle$ 吧。因此最初整个复合系统的量子态可以记作

$$|\psi\rangle|0\rangle. \quad (3.93)$$

所谓的复制，就是通过对整个复合系统进行某个操作，让它演化到如下量子态

$$|\psi\rangle|\psi\rangle. \quad (3.94)$$

既然复制从本质上是整个复合系统的一种量子演化，所以复制操作当然对应某个公正的时间演化算符 U_C ，虽然与这个时间演化算符相应的哈密顿量

可能很复杂，但我们并不需要知道具体的哈密顿量，只需要知道 U_C 是一个幺正算符。

现在，假设有系统1的两个未知量子态 $|\psi\rangle, |\phi\rangle$ (设已经归一化)，如果能够同时复制它们，也就是能够在同一个 U_C 的作用下(同一个复制操作下)，使得

$$U_C(|\psi\rangle|0\rangle) = |\psi\rangle|\psi\rangle, \quad U_C(|\phi\rangle|0\rangle) = |\phi\rangle|\phi\rangle. \quad (3.95)$$

由于幺正演化保持量子态的内积不变，从而必定有

$$(\langle\phi|\langle\phi|)(|\psi\rangle|\psi\rangle) = (\langle\phi|\langle 0|)(|\psi\rangle|0\rangle). \quad (3.96)$$

注意到系统1的态只能和系统1的态进行内积，系统2的态也是和系统2的态进行内积，并利用 $\langle 0|0\rangle = 1$ ，即可以得到

$$(\langle\phi|\psi\rangle)^2 = \langle\phi|\psi\rangle. \quad (3.97)$$

但这就意味着 $\langle\phi|\psi\rangle$ 等于0或者等于1，如果 $\langle\phi|\psi\rangle = 0$ ，那就说明 $|\psi\rangle$ 和 $|\phi\rangle$ 正交，而如果 $\langle\phi|\psi\rangle = 1$ ，则根据前面的引理，就说明 $|\psi\rangle$ 和 $|\phi\rangle$ 是同一个量子态。

因此，如果 $|\phi\rangle$ 和 $|\psi\rangle$ 既不是同一个量子态，也不正交，那这样的复制算符 U_C 就不可能存在！所以，不可能同时复制两个不同且不正交的未知量子态！这就是量子不可克隆定理，值得说明的是，如果这两个量子态正交，也就是可以确定地区分，那它们是可以同时被复制的！

3.4 算符与物理测量

3.4.1 本征可能性完备集

如果一个态矢量 $|u_n\rangle$ 满足如下方程

$$A|u_n\rangle = \lambda_n|u_n\rangle, \quad (3.98)$$

则 $|u_n\rangle$ 就称之为算符 A 的一个本征态，相应的 λ_n 就称之为算符 A 的本征值，而这个方程就称作算符 A 的本征方程。算符 A 很可能有多个本征态，因此我

们以指标 n 来区分, $|u_n\rangle$ 就叫做 A 的第 n 个本征态, λ_n 就叫做本征态 $|u_n\rangle$ 对应的本征值。这里的本征态和本征值当然是一个数学概念, 稍后我们会把这些概念和物理联系起来。

如果选取一个表象 \mathcal{I} , 那么以上方程就可以表示成(通过配上左矢 $\langle i|$ 并利用跃迁元乘法)

$$\sum_{j \in \mathcal{I}} \langle i | A | j \rangle \langle j | u_n \rangle = \lambda_n \langle i | u_n \rangle. \quad (3.99)$$

如果排列成矩阵, 那就是

$$\begin{bmatrix} \langle 1 | A | 1 \rangle & \langle 1 | A | 2 \rangle & \dots \\ \langle 2 | A | 1 \rangle & \langle 2 | A | 2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \langle 1 | u_n \rangle \\ \langle 2 | u_n \rangle \\ \vdots \end{bmatrix} = \lambda_n \begin{bmatrix} \langle 1 | u_n \rangle \\ \langle 2 | u_n \rangle \\ \vdots \end{bmatrix}. \quad (3.100)$$

这就是矩阵的本征方程。

特别的, 哈密顿算符的本征方程, 就叫做定态薛定谔方程, 因为正如稍后将看到的, 它的本征态就是能量有确定取值的定态, 因此可以将定态薛定谔方程写成

$$H |E_n\rangle = E_n |E_n\rangle. \quad (3.101)$$

物理上最重要的算符就是厄米算符, 也就是满足 $A^\dagger = A$ 的算符。这是因为厄米算符可以表示物理量。前面说过, 厄米算符满足如下厄米条件

$$\langle \phi | A | \psi \rangle^* = \langle \psi | A | \phi \rangle. \quad (3.102)$$

厄米算符有一些很重要的特性, 首先, 厄米算符的本征值一定是实数, 其次, 厄米算符不同本征值的本征态一定正交。

为了证明这两个性质, 我们记厄米算符 A 的第 i 个本征态为 $|i\rangle$, 相应的本征值记为 λ_i 。因此,

$$\lambda_j \langle i | j \rangle = \langle i | A | j \rangle = \langle j | A | i \rangle^* = \lambda_i^* \langle j | i \rangle^* = \lambda_i^* \langle i | j \rangle, \quad (3.103)$$

这里第一个等号和第三个等号是利用本征态的定义, 即 $A|i\rangle = \lambda_i |i\rangle$, 第二个等号是利用方程(3.102)。推导过程(3.103)告诉我们 $\lambda_j \langle i | j \rangle = \lambda_i^* \langle i | j \rangle$,

取 $i = j$, 就可以知道 $\lambda_i^* = \lambda_i$, 因此厄米算符的本征值必为实数. 由此我们进一步得到, $(\lambda_j - \lambda_i)\langle i|j\rangle = 0$, 因此当 $\lambda_i \neq \lambda_j$ 时, 必有 $\langle i|j\rangle = 0$, 即相应的两个本征态必定正交。这其实就已经完成了我们需要的证明。

但有一个问题是, 如果某两个本征态 $|i\rangle, |j\rangle$ 有相同的本征值, 即如果 $\lambda_i = \lambda_j$, 这时候怎么办, 这时候我们就不能证明这两个态正交了。这种情况, 我们就叫做这两个本征态简并, 如果只有 $|i\rangle, |j\rangle$ 两个本征态相简并, 我们就叫做本征值 λ_i 有二重简并(当然, 如果有 N 个本征态都简并, 我们就叫做 N 重简并)。碰到这种二重简并的情况, 我们注意到

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

最后一个等号我们用到了简并条件 $\lambda_i = \lambda_j$ 。这个推导过程告诉我们, 如果 $|i\rangle, |j\rangle$ 简并, 则它们的任意线性叠加将依然是算符 A 的本征值为 λ_i 的本征态, 也就是说, $\{|i\rangle, |j\rangle\}$ 实际上张成了一个两维的线性子空间, 这个子空间里的任何态矢量都是 A 的本征值为 λ_i 的本征态, 这样一个线性子空间就称之为简并子空间。而在一个两维的线性子空间(可以将这样的两维矢量空间想象成一个两维平面) 上我们总是可以重新选择两个相互正交的基矢量, 我们可以将这两个正交的基矢量重新定义为新的 $\{|i\rangle, |j\rangle\}$ 态。类似的操作当然也可以推广到多重简并的情况。

因此, 这也就是说, 碰到本征态简并的情况, 我们总是可以对这些简并的本征态进行重新选择, 使得重新选择以后的本征态相互正交。因此, 对于厄米算符, 我们总是可以选取一组两两正交的本征态, 加上归一化条件, 就有

$$\langle i|j\rangle = \delta_{ij}. \quad (3.104)$$

一般来说, 对于在量子力学中出现的某个厄米算符 A , 由于它总是某个物理量的表示, 因此这组正交归一的本征态的线性叠加一定能构造出希尔伯特空间里的任何态矢量, 也就是说, 它们是完备的。同时满足正交归一性和完备性, 这就意味着厄米算符 A 的这一组本征态 $\{|i\rangle, i = 1, 2, 3, \dots\}$ 可以构成整个希尔伯特空间的一组正交矢量基(常常也称本征态的线性叠加结果为叠加态)。或者用物理的话说, 相应的可能性完备集 $\{(i), i = 1, 2, 3, \dots\}$ 构成了一组可确定区分可能性完备集, 也称之为物理量 A 的**本征可能性完备集**, 相应的可能性 (i) 就称之为物理量 A 的**本征可能性**。

很明显，如果我们把物理量 A 以它的本征可能性完备集为表象表示出来，则由于 $A|i\rangle = \lambda_i|i\rangle$ 以及正交归一关系 $\langle i|j\rangle = \delta_{ij}$ ，所以一定有

$$\begin{bmatrix} \langle 1|A|1\rangle & \langle 1|A|2\rangle & \dots \\ \langle 2|A|1\rangle & \langle 2|A|2\rangle & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \dots \\ 0 & \lambda_2 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (3.105)$$

也就是说，这个表示矩阵一定为对角矩阵！

进而由上一章的相关讨论可以知道。如果量子系统处在物理量 A 的本征可能性(i)上，那 A 就有确定的取值 λ_i ，这时候去测量物理量 A ，就会得到值 λ_i (特别的，哈密顿量的本征可能性就是能量有确定取值的可能性，因此也就是定态可能性)。而如果系统处在一个任意的初态可能性 ψ 上，那 A 就没有确定取值，这时候去测量 A 的值(也就是对系统的量子可能性进行适当的检测)，那系统就会以概率 $|\langle i|\psi\rangle|^2$ 随机地从 ψ 跃迁到某个本征可能性(i)，一旦它跃迁到(i)，当然就会测到值 λ_i 。

这也反过来可以说明为什么物理量要用厄米算符来表示。原因在于，某个物理量有不同确定取值的量子态(本征可能性)之间，由于可以根据这个物理量取值的不同来确定地区分，因此这些态必定是两两正交的。而厄米算符刚好有类似的性质，因为根据我们在前面的数学证明，厄米算符不同本征值的本征态必定是相互正交的。而如果系统不是处于物理量的本征可能性，那物理量就没有确定取值。另一方面，厄米算符的本征值当然只在本征态上有定义，如果系统处在某个厄米算符本征态的叠加态上，那当然就谈不上本征值，而且厄米算符本身是一个抽象算符而不是一个实数，因此也就是说，这时候厄米算符的值是没有定义。这些都刚好吻合物理量所必须遵循的基本原理。因此，可以把物理量的本征可能性当作某个厄米算符的本征态，把物理量的值当成某个厄米算符的本征值，也就是说，把物理量用厄米算符来表示！而且，我们前面也证明过，厄米算符的本征值刚好是实数，因此完全可以当成物理量的值，如果厄米算符的本征值可以是复数的话，那将物理量表示成厄米算符就不合适了。

归纳一下就是，对于任何一个量子系统，如果对它某个物理量的值进行检测，则测量值一定是相应算符的本征值。但除非系统原来就处在这个物理量的本征可能性上，否则测得的值是不确定的，而是会随着量子跃迁的结果，以概率 $|\langle i|\psi\rangle|^2$ 随机地得到本征值 λ_i 。

注意，没有什么测量引起的态塌缩，只有系统在初末两种量子可能性之间的量子跃迁，而且量子跃迁与量子力学幺正性并不矛盾，相反，正如前文已经表明的，包括态矢量的概念以及算符与幺正性的概念，都是从跃迁元以及跃迁元乘法中构造出来的！从这个角度上说，量子跃迁才是量子力学无可避免的根本特征，它和量子力学幺正性并没有什么冲突，之所以通常人们认为有冲突，是因为过度延伸了人为构造出来的概念体系，包括态矢量、态矢量的幺正演化，以及尤其是量子态的塌缩！

本章最后一小节会进一步讨论量子力学中的测量问题，尤其是讨论其与幺正性之间的关系问题。

3.4.2 不确定原理与算符

还是让我们从电子自旋的例子开始，前面定义了算符 X (请与粒子坐标相区分)，即

$$X|\uparrow\rangle = |\downarrow\rangle, \quad X|\downarrow\rangle = |\uparrow\rangle. \quad (3.106)$$

另一方面，根据前面的(3.18)式，我们有

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle), \quad |\leftarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle). \quad (3.107)$$

从而不难验证， $|\rightarrow\rangle, |\leftarrow\rangle$ 均为算符 X 的本征态，本征值分别为 $+1, -1$ ，即有

$$X|\rightarrow\rangle = |\rightarrow\rangle, \quad X|\leftarrow\rangle = -|\leftarrow\rangle. \quad (3.108)$$

类似的，我们还可以定义一个算符 Z ，它以 $|\uparrow\rangle, |\downarrow\rangle$ 为本征态，本征值也是 $+1, -1$ ，即

$$Z|\uparrow\rangle = |\uparrow\rangle, \quad Z|\downarrow\rangle = -|\downarrow\rangle. \quad (3.109)$$

所以，算符 X 和算符 Z 的本征态是完全不同的， X 有确定取值时， Z 的取值就完全不确定，反之亦然。另一方面，不难验证 X 和 Z 不对易，比如 $XZ|\uparrow\rangle = X|\uparrow\rangle = |\downarrow\rangle$ ，而 $ZX|\uparrow\rangle = Z|\downarrow\rangle = -|\downarrow\rangle$ ，即 $XZ|\uparrow\rangle = -ZX|\uparrow\rangle$ ，类似的也有 $XZ|\downarrow\rangle = -ZX|\downarrow\rangle$ ，所以实际上

$$XZ = -ZX. \quad (3.110)$$

这个例子让我们想到，算符的不对易，和相应两物理量的值不能同时确定，这两者之间有没有什么必然的联系呢？

另一方面，如果两个物理量的值总是可以同时确定（请注意总是这个词），那就意味着，不管系统处在任何初态可能性，每次检测这两个物理量时，系统所跃迁到的末态可能性必然同时是这两个物理量的本征可能性。并且，由于我们假设对任何初态可能性进行检测总能同时得到这两个物理量的某个值，因此这也意味着这些共同的本征可能性必然可以张成希尔伯特空间的一组矢量基（即任何量子态都可以写成这组共同本征态的线性叠加），因此，如果两个物理量的值总是可以同时确定，那就意味着它们必然有一组共同的本征可能性完备集。

但是，我们已经看到，物理量是用厄米算符来表示的，那么从厄米算符的角度来看，两个物理量的值总是可以同时确定的充要条件是什么呢？答案很简单，是这两个算符的乘积要可以交换顺序，也就是这两个算符可对易。

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

$$\begin{aligned} (AB - BA)|\psi\rangle &= \sum_i c_i (AB - BA)|i\rangle \\ &= \sum_i c_i (\beta_i A|i\rangle - \alpha_i B|i\rangle) = \sum_i c_i (\beta_i \alpha_i - \alpha_i \beta_i)|i\rangle = 0. \end{aligned} \quad (3.111)$$

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

下面证明充分性（充分性的证明略微有一点复杂，并且纯粹是一个数学定理，读者第一遍读的时候可以先略过）。即要证明如果两个物理量的厄米算符满足 $AB = BA$ ，则它们必有一组共同的本征态矢量基。首先，前文已经说过，对于量子力学中考察的厄米算符而言，其本征态必定可以取作希尔伯特空间的正交矢量基。所以真正需要证明的是，我们总是可以把这些

本征态取成这两个算符共同的本征态。为此我们假设 $\{|i\rangle, i = 1, 2, 3\dots\}$ 为算符A的一组本征态，本征值分别为 α_i (注意， α_i 可以有简并)，则由于

$$A(B|i\rangle) = AB|i\rangle = BA|i\rangle = \alpha_i B|i\rangle, \quad (3.112)$$

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

但是，如果算符A的本征值为 α_i 的本征态有简并，那这时候由(3.112)式我们就不能得出 $B|i\rangle$ 正比于 $|i\rangle$ 的结论了。比方说假设有二重简并(多重简并的证明是类似的)，也即是说有某个 $\alpha_j = \alpha_i$ ，从而相应的 $|j\rangle$ 态和 $|i\rangle$ 态有相同的本征值 α_i 。这就和我们前面证明总是可以将厄米算符的本征态取成正交的时候碰到的简并情况是一样的。 $|i\rangle$ 和 $|j\rangle$ 简并的一个显然结果是，它们的任意线性叠加 $c_1|i\rangle + c_2|j\rangle$ 都将是A的本征态，本征值都是 α_i ，因此这就张成了一个二维的简并子空间，我们可以把这个简并子空间记为 \mathcal{H}_{α_i} 。那这时候我们由(3.112)式就只能得出 $B|i\rangle$ 一定是这个二维的简并子空间里的某个态矢量。不仅如此，由于

$$\begin{aligned} A[B(c_1|i\rangle + c_2|j\rangle)] &= B(c_1A|i\rangle + c_2A|j\rangle) \\ &= B(c_1\alpha_i|i\rangle + c_2\alpha_j|j\rangle) = \alpha_i[B(c_1|i\rangle + c_2|j\rangle)], \end{aligned} \quad (3.113)$$

这里最后一个等号我们利用了简并条件 $\alpha_j = \alpha_i$ 。这个推导过程(3.113)就告诉我们， $[B(c_1|i\rangle + c_2|j\rangle)]$ 依然是A的某个本征值为 α_i 的本征态，这也就是说，B作用在简并子空间 \mathcal{H}_{α_i} 里的任意态上，结果都依然还是简并子空间里的态。因此我们就总可以在这个二维的简并子空间 \mathcal{H}_{α_i} 里求解算符B的本征方程，从而得到B的两个本征态 $|i'\rangle$ 和 $|j'\rangle$ 。前文引入厄米算符本征态时的相关证明告诉我们，无论这两个态对应的B本征值相同还是不同，我们总是可以让它们正交，从而成为二维简并子空间 \mathcal{H}_{α_i} 新的正交基(因此原来的 $|i\rangle$ 和 $|j\rangle$ 也都可以反过来写成这两个新的基矢量的线性叠加)。由于 $|i'\rangle$ 和 $|j'\rangle$ 都属于A的简并子空间 \mathcal{H}_{α_i} ，所以它们当然同时也是A的本征态，这样它们就是A和B共同的本征态。也就是说，碰到这种二重简并的情况，

我们只需要把原来的 $|i\rangle, |j\rangle$ 态替换成新的 $|i'\rangle$ 和 $|j'\rangle$ 态就找到了这时候两个算符的共同本征态。多重简并情形的推理和结论都是类似的。因此这就证明了, A 和 B 的共同本征态总是可以找到的。这样就完成了我们对充分性的证明。这个充分性的证明告诉我们, 如果两个算符可对易, 那么它们相应的物理量必然有一组共同的本征可能性完备集, 从而这两个物理量的值总是可以同时确定。

这个充分性命题的逆否命题就是, 不可同时确定的两个物理量, 它们相应的算符必定不对易。

另外, 必要性命题的逆否命题是, 如果两个算符不对易, 那么它们相应的物理量就不可能有一组共同的本征可能性完备集。一个典型的例子就是前面的自旋算符 X (注意和粒子坐标相区分)和 Z , 由于 $XZ = -ZX$, 因此两者不对易, 从而正如我们已经看到的, 它们没有共同的本征态。另一个典型的例子就是位置算符和动量算符, 即 X, P , 这两个算符的对易子就是, $[X, P] = i\hbar$, 即对易子是个常数, 因此 X 和 P 必定没有共同的本征态。

3.4.3 再谈物理量的测量

前面讨论物理量的值, 也就是相应算符本征值的时候, 我们忽略了一个细节。那就是, 在海森堡绘景中, 物理量算符是要随着时间演化的, 这个时候定义本征态的时候就需要考虑到时间因素。

具体来说即是, 给定一个海森堡绘景中的物理量算符 $A(t)$, 我们需要在某个任意给定的参考时刻 t_0 来定义它的本征态, 因此相应的本征方程是

$$A(t_0)|a, t_0\rangle = \lambda_a|a, t_0\rangle, \quad a = 1, 2, \dots \quad (3.114)$$

式中 λ_a 表示 $A(t_0)$ 的本征值, $|a, t_0\rangle$ 就是相应的本征态, 它依赖于参考时刻 t_0 .

根据前面讲述的海森堡运动方程的通解, 我们有

$$A(t_0) = e^{iHt_0/\hbar} A_0 e^{-iHt_0/\hbar}, \quad (3.115)$$

式中 A_0 就是参考时刻为0时刻所对应的算符, 当然它也是相应的薛定谔绘景中的算符。根据这个关系式, 我们可以构造出前面的本征态 $|a, t_0\rangle$ 。为此, 假设 $|a\rangle$ 为 A_0 的本征态, 满足如下本征方程

$$A_0|a\rangle = \lambda_a|a\rangle, \quad (3.116)$$

则根据(3.115)式，不难验证 $e^{iHt_0/\hbar}|a\rangle$ 为 $A(t_0)$ 的本征值为 λ_a 的本征态，换言之，

$$|a, t_0\rangle = e^{iHt_0/\hbar}|a\rangle. \quad (3.117)$$

而且这个构造也说明了 $A(t_0)$ 的本征值 λ_a 本身并不依赖于 t_0 。

注意， $|a, t_0\rangle$ 并不是一个随时间演化的量子态，它只是一个在参考时刻 t_0 定义的量子态，本身依然是不随时间演化的，否则就和薛定谔绘景中的态演化概念相冲突了。

现在，假设任给系统的一个初态可能性 $|\psi\rangle$ ，我们在参考时刻 t_0 去检测物理量 $A(t_0)$ 的值，当然，系统就会以跃迁幅 $\langle a, t_0 | \psi \rangle$ 从初态可能性跃迁到 $A(t_0)$ 的某个本征可能性 $\langle a, t_0 |$ （作为检测的末态可能性），一旦我们确定系统跃迁到了 $\langle a, t_0 |$ ，我们就能推测出 $A(t_0)$ 的值为 λ_a 。也就是说，我们会以概率 $|\langle a, t_0 | \psi \rangle|^2$ 随机地测得值 λ_a 。

这里简单地提一下薛定谔绘景中是如何看待 $|\langle a, t_0 | \psi \rangle|^2$ 这样的概率的。很显然，由于 $|a, t_0\rangle = e^{iHt_0/\hbar}|a\rangle$ ，所以 $\langle a, t_0 | = \langle a | e^{-iHt_0/\hbar}$ ，所以

$$|\langle a, t_0 | \psi \rangle|^2 = |\langle a | e^{-iHt_0/\hbar} | \psi \rangle|^2 = |\langle a | \psi(t_0) \rangle|^2, \quad (3.118)$$

式中最后一个等于号用到了薛定谔绘景量子态随时间的演化关系。

回到海森堡绘景。当然，参考时刻 t_0 并不是时间的终点，所以 $|a, t_0\rangle$ 会成为新的初态可能性继续往下进行量子跃迁，直到某个最终的末态可能性 $|\phi\rangle$ 。所以，纵观从初态可能性 $|\psi\rangle$ 到最终的末态可能性 $|\phi\rangle$ 的全过程， (a, t_0) 就只是一个中间可能性。然而，由于我们能够确定系统在 t_0 时刻跃迁到了这个中间可能性，所以根据第二章讲述的量子力学基本原理，整个过程的概率将由下式给出

$$|\langle \phi | a, t_0 \rangle|^2 |\langle a, t_0 | \psi \rangle|^2, \quad (3.119)$$

注意，这时候并不能使用跃迁幅乘法规则，因为我们能够确定具体的中间可能性是什么。相反，如果有多种能够确定的中间可能性都能导致相同的最终末态，那我们就应该把上面的概率(3.119)(而不是跃迁幅)对这些中间可能性求和。这当然就是第二章中讲述过的基本原理之一。

可见，整个分析并没有量子态么正演化以及量子态塌缩什么事，当然也就没有这两者之间的矛盾冲突！有的只是计算概率的两种不同方法：如

果原则上就不能确定中间可能性，当然也就无从推测物理量 A 的值，相应的就应该对中间可能性使用跃迁幅乘法规则来计算完整的跃迁幅 $\langle\phi|\psi\rangle$ ，并将这个跃迁幅模方得到概率。正如第二章中讲述过的，这时候各种中间可能性之间就存在量子干涉。而如果原则上能够确定具体中间可能性是什么，当然就能推测出这个中间时刻物理量的值是多少，但这时候就不能对这些中间可能性使用跃迁幅乘法了，而应该是把各种不同中间可能性的概率(而不是跃迁幅)相加，当然，这时候也就不存在量子干涉了。

如果想进一步统一这两种计算概率的方法，或者问，在对中间可能性进行了测量的情况下，量子相干性是如何消失的？那就是另一个问题了，这时候可能就需要把测量仪器以及观察者都包括进量子系统之中了。实际上，这就是著名的退相干问题，对这个问题的研究构成了一个专门的研究领域。退相干当然是一个物理问题，虽然不在本书的讨论范围之内。本书只是想说，通常人们所认为的量子力学基本问题或者测量问题，其实不等于退相干，退相干是物理的，但是量子态塌缩、多世界之类的想法则可能是不必要的，是对一些人为的抽象构造进行了过度的延伸。

那么幺正性是怎么回事呢？按照本书观点的回答如下：如果原则上不能确定中间可能性(比如不对中间可能性进行检测)，那就有完整的跃迁幅 $\langle\phi|\psi\rangle$ ，幺正性说的无非就是这个跃迁幅不随考察时刻 t 而变化，这在海森堡绘景中是平庸的。而如果原则上能确定中间可能性，那 $\langle\phi|\psi\rangle$ 就没有定义了！这时候只能分别考察跃迁幅 $\langle a, t_0 | \psi \rangle$ 和跃迁幅 $\langle \phi | a, t_0 \rangle$ ，它们当然也都不随考察时刻 t 而变化。这时候你可以定义比方说 $\langle\phi|\psi\rangle_{a,t_0} \equiv \langle\phi|a, t_0\rangle\langle a, t_0|\psi\rangle$ ，但是原来的能够对中间可能性使用跃迁幅乘法的 $\langle\phi|\psi\rangle$ 是没有定义的。在能确定具体中间可能性的情况下，如果硬要定义完整的跃迁幅 $\langle\phi|\psi\rangle$ ，那就只能把测量仪器、观察者、甚至整个世界，都包括进量子系统之中了，但是这样一来，初态可能性就不再是 $|\psi\rangle$ 了，而是整个世界的初态可能性，末态可能性也一样，不再是 $\langle\phi|$ 了，而是整个世界的末态可能性。

但是，由于一定意义上的测量无时无刻不在进行，无时无刻不在引起被测系统的退相干，所以，为了定义完整的跃迁幅，唯一的办法也许真是考察整个世界的初态可能性和整个世界的末态可能性。

什么是整个世界的初态可能性呢？当然就是在整个世界的无穷远过去定义的量子可能性，也就是在 $t_1 = -\infty$ 的参考时刻定义的可能性，

人们可以把它取在某个可确定区分可能性完备集 $\mathcal{I}_{-\infty} = \{(i, -\infty), i = 1, 2, 3, \dots\}$ 中。同样，所谓整个世界的末态可能性，当然就是在整个世界的无穷远将来定义的量子可能性，也就是在 $t_2 = +\infty$ 的参考时刻定义的可能性，类似的，也可以取在某个可确定区分可能性完备集 $\mathcal{I}_{+\infty} = \{(f, +\infty), f = 1', 2', 3', \dots\}$ 中。

进而，整个世界完整的跃迁幅当然就是 $\langle f, +\infty | i, -\infty \rangle$ ，沿用粒子物理里面的说法，可以称之为整个世界的 S 矩阵，记为 $S_{fi} = \langle f, +\infty | i, -\infty \rangle$ 。那么，所谓整个世界的幺正性，指的就是这个 S 矩阵(也就是 $\mathcal{I}_{-\infty}$ 到 $\mathcal{I}_{+\infty}$ 之间的变换矩阵)必须是一个幺正矩阵，即满足

$$\sum_{f \in \mathcal{I}_{+\infty}} S_{fk}^* S_{fi} = \delta_{ki}, \quad \sum_{i \in \mathcal{I}_{-\infty}} S_{gi} S_{fi}^* = \delta_{gf}. \quad (3.120)$$

如果跃迁幅 $\langle f, +\infty | i, -\infty \rangle$ 的定义和计算完全没有问题，那这个幺正性必定是自动满足的，因为它是跃迁幅乘法的结果。

第四章 电子自旋以及相关模型

前面几章着重在建立现代量子力学的理论体系，本章将研究两个具体例子。首先就是研究电子自旋，电子自旋可以说是最简单，同时又最具有量子力学特征的例子了，而且上一章中我们已经涉及到电子自旋了，但是上一章并没有系统研究电子自旋的描述，本章将从最基本的实验事实出发，利用量子力学的基本理论，建立起对电子自旋的完整描述。其次，本章还要研究一个和电子自旋相关的物理模型，具体来说就是两自旋耦合的模型。这依然是用来作为上一章理论体系的一个具体例子。

4.1 电子自旋

4.1.1 电子自旋的发现

这一小节让我们更仔细地研究一下施特恩和格拉赫(Stern and Gerlach)的原版实验。

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

$$\vec{\mu}_L = \frac{q}{2m} \mathbf{L}. \quad (4.1)$$

当然，我们也知道，磁矩会和磁场耦合，如果存在一个磁场 \mathbf{B} ，那磁矩 $\vec{\mu}$ 和

磁场相互作用的势能 V 将是

$$V = -\vec{\mu} \cdot \mathbf{B}. \quad (4.2)$$

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

施特恩和格拉赫的实验装置如图(4.1)a所示。在这个实验中，施特

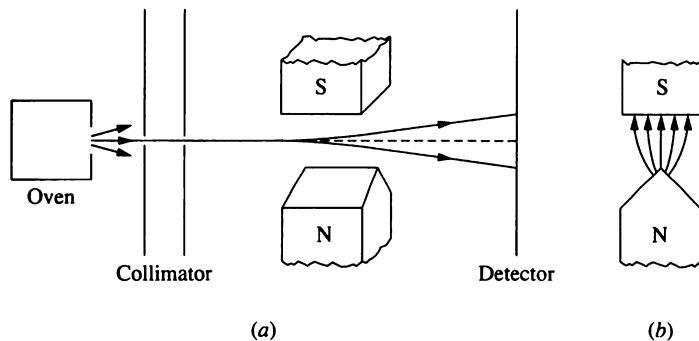


图 4.1: 施特恩-格拉赫实验。

恩和格拉赫让一束经过准直的银原子束通过一个非均匀磁场，然后再打在一个屏幕上。这个磁场如图(4.1)b所示，它在竖直的 z 方向有一个梯度 $\frac{\partial B_z}{\partial z}$ 。因此如果银原子有一个磁矩 μ 的话，它在 z 方向将受到一个力 F_z ， $F_z = -\frac{\partial V}{\partial z} = \vec{\mu} \cdot \frac{\partial \mathbf{B}}{\partial z} \simeq \mu_z \frac{\partial B_z}{\partial z}$ 。从图(4.1)b可以看到，在这里磁场的梯度 $\frac{\partial B_z}{\partial z}$ 其实是负的(下方的磁场更强，上方的磁场更弱)，因此如果 μ_z 为负的话，那银原子的受力将沿着 z 轴的正方向。

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

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

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

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

4.1.2 泡利算符与自旋算符

根据施特恩-格拉赫实验，通过施特恩-格拉赫装置(假设磁场沿着 z 方向)的电子会分裂成两束，因此电子磁矩的 z 分量 μ_z 必然取量子化的两个分立值，不妨记为 $\pm \mu_B$ 。当然， z 方向并没有任何特殊，我们同样可以将施特恩-格拉赫装置转向 x 方向和 y 方向(即转动转置使得磁场分别沿着 x 方向和 y 方向)，进而可知，电子磁矩的 x 分量 μ_x 以及 y 分量 μ_y 也都只能取 $\pm \mu_B$ 两个量子化的分立值。同样，可以把施特恩-格拉赫装置转向任意**n**方向(**n**为单位矢量)，进而即知 $\mu_n = \vec{\mu} \cdot \mathbf{n}$ 也只能取 $\pm \mu_B$ 两个分立值。

另一方面，在量子力学中，电子的磁矩当然也要用厄米算符来描述，

仍然记电子的磁矩算符为 $\vec{\mu}$, 则根据上一段的论述可知, 对于任意单位矢量 \mathbf{n} , 算符 $\mu_{\mathbf{n}} = \vec{\mu} \cdot \mathbf{n}$ 都只有 $\pm \mu_B$ 两个本征值, 换言之

$$(\vec{\mu} \cdot \mathbf{n})^2 = \mu_B^2. \quad (4.3)$$

不妨引入一个不带量纲的算符 $\vec{\sigma}$ (即所谓的泡利算符), 它和电子磁矩算符的联系是

$$\vec{\mu} = \mu_B \vec{\sigma}. \quad (4.4)$$

将这个式子代入上面的(4.3), 即有

$$(\vec{\sigma} \cdot \mathbf{n})^2 = 1. \quad (4.5)$$

或者说, 对于任意 \mathbf{n} , $\vec{\sigma} \cdot \mathbf{n}$ 的本征值总是 ± 1 。

下面记 $\vec{\sigma}$ 的三个分量为 $\sigma_a, a = x, y, z$ (它们的本征值当然也都是 ± 1), 记单位矢量 \mathbf{n} 的三个分量为 $n_a, a = x, y, z$, 并记

$$\vec{\sigma} \cdot \mathbf{n} = \sigma_a n_a, \quad (4.6)$$

式中我们使用了所谓的求和约定, 即对表达式中重复出现两次的指标默认对它的所有可能性求和, 式中也就是默认对 a 求和, 后文也采用同样约定。根据这些约定, 我们就能把(4.5)式重写为

$$(\sigma_a n_a)(\sigma_b n_b) = 1. \quad (4.7)$$

式中 $a, b = x, y, z$ 。由这个式子即有

$$\sigma_a \sigma_b n_a n_b = 1 \Rightarrow \frac{1}{2} (\sigma_a \sigma_b + \sigma_b \sigma_a + [\sigma_a, \sigma_b]) n_a n_b = 1. \quad (4.8)$$

对于式中的 $[\sigma_a, \sigma_b] n_a n_b$, 通过交换指标 a, b (反正它们都是要求和的指标), 即有 $[\sigma_a, \sigma_b] n_a n_b = [\sigma_b, \sigma_a] n_b n_a = -[\sigma_a, \sigma_b] n_a n_b$, 一个式子等于负的它本身, 那这个式子必定等于零, 从而 $[\sigma_a, \sigma_b] n_a n_b = 0$ 。从而(4.8)式可以简化成

$$\frac{1}{2} (\sigma_a \sigma_b + \sigma_b \sigma_a) n_a n_b = 1. \quad (4.9)$$

与单位矢量的归一条件 $\mathbf{n}^2 = n_a n_a = \delta_{ab} n_a n_b = 1$ 比较, 并由于单位矢量 n_a 的任意性, 即知

$$\sigma_a \sigma_b + \sigma_b \sigma_a = 2\delta_{ab}. \quad (4.10)$$

满足类似这组式子的抽象代数就叫Clifford代数，数学家已经系统地研究过这类代数的矩阵表示了，后文我们将演示给大家这样的表示可以如何进行。

(4.10)式也可以更清楚地写成

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1. \quad (4.11)$$

以及

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0, \quad \sigma_y\sigma_z + \sigma_z\sigma_y = 0, \quad \sigma_z\sigma_x + \sigma_x\sigma_z = 0. \quad (4.12)$$

也就是说，这三个泡利算符是两两反交换的，或者说反对易的，所谓两个算符反对易即是指 $AB = -BA$ 。

根据上面的反对易性，不难验证算符 $\sigma_x\sigma_y\sigma_z$ 和三个泡利算符 $\sigma_x, \sigma_y, \sigma_z$ 均对易！但是只有常数才可能和所有的 σ_a 均对易，所以必定有 $\sigma_x\sigma_y\sigma_z$ 为常数。另一方面，根据上面反对易关系(4.12)式，通过不断将两个相同的 σ_a 分量对易到一起，进而利用(4.11)式，从而不难得到

$$(\sigma_x\sigma_y\sigma_z)^2 = -1. \quad (4.13)$$

从而 $\sigma_x\sigma_y\sigma_z = \pm i$ ，由于只要将 $-\sigma_a$ 重新定义成 σ_a 就能交换式中的正负号，所以这两种可能性本质上是等价的，通常取正号，所以有

$$\sigma_x\sigma_y\sigma_z = i. \quad (4.14)$$

将这个式子两边分别乘以 $\sigma_x, \sigma_y, \sigma_z$ ，并利用(4.11)式和(4.12)式，即能得到

$$\sigma_x\sigma_y = i\sigma_z, \quad \sigma_y\sigma_z = i\sigma_x, \quad \sigma_z\sigma_x = i\sigma_y, \quad (4.15)$$

这三个式子涉及的三个指标刚好构成轮换关系。后文我们将尤其要用到

$$\sigma_x\sigma_y = i\sigma_z. \quad (4.16)$$

电子磁矩正比于电子自旋，同时也正比于泡利算符，所以我们可以利用泡利算符来定义电子的自旋算符 \mathbf{S} ，它的三个分量分别定义为

$$S_x = (\hbar/2)\sigma_x, \quad S_y = (\hbar/2)\sigma_y, \quad S_z = (\hbar/2)\sigma_z. \quad (4.17)$$

从这个定义可以知道, S_a 的本征值必定为 $\pm\hbar/2$ 。

利用泡利算符的反对易性以及(4.15)式, 不难验证泡利算符满足如下代数关系

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \quad [\sigma_y, \sigma_z] = 2i\sigma_x, \quad [\sigma_z, \sigma_x] = 2i\sigma_y. \quad (4.18)$$

进而根据上面的定义, 可得自旋算符的代数关系

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y. \quad (4.19)$$

式中涉及的三个指标刚好构成轮换关系。实际上, 这组代数关系不仅适用于电子自旋算符, 也可以推广到任何角动量算符。

由于自旋算符任何两个分量均不对易, 所以根据上一章讲述的定理, 电子自旋的任何两个分量都不可能同时确定, 比方说, 电子自旋的 S_x 和 S_y 不可能同时有确定值。这和经典物体的自转是完全不同的, 经典的角动量三个分量都可以同时有确定值。

4.1.3 泡利矩阵

因此, 要将电子自旋算符表示成矩阵, 我们只需要等价地将泡利算符表示成矩阵就可以了。而这可以按照Clifford代数的标准表示理论进行。

下面的推导在数学上是完全严格的。它完全是从泡利算符的抽象代数出发, 也即是从方程(4.11), (4.12)以及方程(4.16)出发, 然后构造出一切。尤其是, 我们没有隐含地假设任何算符的本征态的存在性, 而是直接构造出了这些本征态。

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

$$\sigma = (\sigma_x - i\sigma_y)/2, \quad \sigma^\dagger = (\sigma_x + i\sigma_y)/2. \quad (4.20)$$

则利用方程(4.11)和(4.12), 很容易得到

$$\sigma^2 = \sigma^{\dagger 2} = 0. \quad (4.21)$$

而且, 由于 $\sigma^\dagger\sigma = (\sigma_x + i\sigma_y)(\sigma_x - i\sigma_y)/4 = (\sigma_x^2 + \sigma_y^2 + i\sigma_y\sigma_x - i\sigma_x\sigma_y)/4 = (1 + \sigma_z)/2$, 这里最后一个等号我们利用了 σ_x 和 σ_y 的反对易以及 $\sigma_x\sigma_y = i\sigma_z$,

类似的，也可以算得 $\sigma\sigma^\dagger = (1 - \sigma_z)/2$ ，归纳一下即有

$$\sigma^\dagger\sigma = (1 + \sigma_z)/2, \quad \sigma\sigma^\dagger = (1 - \sigma_z)/2. \quad (4.22)$$

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

$$\sigma|\downarrow\rangle = 0. \quad (4.23)$$

这是因为，任取一个自旋量子态 $|\psi\rangle$ ，如果它满足 $\sigma|\psi\rangle = 0$ ，那这个 $|\psi\rangle$ 就是我们要找的态，如果 $\sigma|\psi\rangle \neq 0$ ，那我们就可以令 $\sigma|\psi\rangle = |\phi\rangle$ ，这时候将有 $\sigma|\phi\rangle = \sigma^2|\psi\rangle = 0$ (后一个等号是由于 $\sigma^2 = 0$)，因此 $|\phi\rangle$ 就将是我们要找的态。总之，满足方程(4.23)的量子态 $|\downarrow\rangle$ 总能找到。对于这个态我们必有 $0 = \langle\downarrow|\sigma^\dagger\sigma|\downarrow\rangle = \langle\downarrow|(1 + \sigma_z)|\downarrow\rangle/2$ ，即有 $\langle\downarrow|(1 + \sigma_z)|\downarrow\rangle = 0$ ，由于 σ_z 的本征值为 ± 1 ，因此这个结果就告诉我们 $|\downarrow\rangle$ 必定是 σ_z 的本征值为 -1 的本征态，即

$$\sigma_z|\downarrow\rangle = -|\downarrow\rangle. \quad (4.24)$$

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

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

$$|\uparrow\rangle = \sigma^\dagger|\downarrow\rangle. \quad (4.25)$$

很显然 $\langle\uparrow|\sigma\sigma^\dagger|\uparrow\rangle = \langle\uparrow|\sigma\sigma^{\dagger 2}|\downarrow\rangle = 0$ ，利用 $\langle\uparrow|\sigma\sigma^\dagger|\uparrow\rangle = \langle\uparrow|(1 - \sigma_z)/2|\uparrow\rangle$ ，即有 $\langle\uparrow|(1 - \sigma_z)|\uparrow\rangle = 0$ ，从而 $|\uparrow\rangle$ 必为 σ_z 的本征值为 $+1$ 的本征态，即

$$\sigma_z|\uparrow\rangle = |\uparrow\rangle. \quad (4.26)$$

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

将方程(4.25)共轭转置，就有 $\langle\uparrow| = \langle\downarrow|\sigma$ ，因此， $\langle\uparrow|\downarrow\rangle = \langle\downarrow|\sigma|\downarrow\rangle = 0$ (后一个等号是由于方程(4.23))，也即

$$\langle\uparrow|\downarrow\rangle = 0. \quad (4.27)$$

因此自旋向上态和自旋向下态必定是两个正交的量子态，这正符合这两个态可以确定地区分的物理要求。而且也可以算得 $\langle \uparrow | \uparrow \rangle = \langle \downarrow | \sigma\sigma^\dagger | \downarrow \rangle = \langle \downarrow | (1 - \sigma_z)/2 | \downarrow \rangle = \langle \downarrow | \downarrow \rangle$ ，因此如果我们将最开始时选的 $|\downarrow\rangle$ 态归一化，那由方程(4.25)定义的 $|\uparrow\rangle$ 也将自动是归一化的，即

$$\langle \downarrow | \downarrow \rangle = \langle \uparrow | \uparrow \rangle = 1. \quad (4.28)$$

由于同时满足正交归一性，因此我们构造出来的 $\{|\downarrow\rangle, |\uparrow\rangle\}$ 可以作为电子自旋量子态的两个基矢量。

从物理上来说，这两个基矢量当然应该是完备的。但是，我们能够从泡利算符的代数出发从数学上证明这一点吗？回答很简单，可以。比方说我们观察刚才的两个基矢量 $|\downarrow\rangle, |\uparrow\rangle = \sigma^\dagger |\downarrow\rangle$ ，我们发现 $|\uparrow\rangle$ 是用 σ^\dagger 作用在 $|\downarrow\rangle$ 上得到的，那我们能用 σ^\dagger 再重复作用一次从而得到第三个态吗？答案是不能，因为 $\sigma^{\dagger 2} = 0$ ，因此 σ^\dagger 的重复作用只能得到0，即

$$\sigma^\dagger |\uparrow\rangle = 0. \quad (4.29)$$

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

$$\sigma |\uparrow\rangle = |\downarrow\rangle. \quad (4.30)$$

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

以 $\{|\uparrow\rangle, |\downarrow\rangle\}$ 为矢量基（现在要注意这两个基矢量的排列顺序），就能够分别计算出泡利算符 $\sigma_z, \sigma, \sigma^\dagger$ 在这个基中的表示矩阵仍然记作 $\sigma_z, \sigma, \sigma^\dagger$ （也就是说，我们使用了相同的符号来表示算符以及它们的表示矩阵，读者不难根据上下文进行区分）。

$$\sigma_z = \begin{bmatrix} \langle \uparrow | \sigma_z | \uparrow \rangle & \langle \uparrow | \sigma_z | \downarrow \rangle \\ \langle \downarrow | \sigma_z | \uparrow \rangle & \langle \downarrow | \sigma_z | \downarrow \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (4.31)$$

$$\sigma = \begin{bmatrix} \langle \uparrow | \sigma | \uparrow \rangle & \langle \uparrow | \sigma | \downarrow \rangle \\ \langle \downarrow | \sigma | \uparrow \rangle & \langle \downarrow | \sigma | \downarrow \rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad (4.32)$$

$$\sigma^\dagger = \begin{bmatrix} \langle \uparrow | \sigma^\dagger | \uparrow \rangle & \langle \uparrow | \sigma^\dagger | \downarrow \rangle \\ \langle \downarrow | \sigma^\dagger | \uparrow \rangle & \langle \downarrow | \sigma^\dagger | \downarrow \rangle \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (4.33)$$

在计算中，需要使用方程(4.23)，(4.24)，(4.25)，(4.26)，(4.27)，(4.28)，(4.29)，(4.30)，当然，这些方程中的每一个其实都很简单。

而由 σ, σ^\dagger 的定义式(4.20)，我们又可以进一步得到 σ_x 的表示矩阵 $\sigma_x = \sigma + \sigma^\dagger$ 以及 σ_y 的表示矩阵 $\sigma_y = i(\sigma - \sigma^\dagger)$ 。这三个矩阵 $\sigma_x, \sigma_y, \sigma_z$ 就是著名的泡利矩阵，我们将它们归纳在下面，

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (4.34)$$

实际上，泡利算符 σ_x 就是上一章定义过的算符 X ，泡利算符 σ_z 就是上一章定义的算符 Z 。 X, Z 这样的记号在量子信息和量子计算中用得多一些，泡利算符的记号则在其它物理领域用得多一些。

4.2 两自旋耦合系统

4.2.1 海森堡模型

铁，钴，镍等物质在很小的外磁场影响下，就能产生远大于其他物质的磁化效应，这就是铁磁性。这些物质之所以有铁磁性，是因为在它们的每一个小区域之内都会自发磁化形成磁畴。而自发磁化的原因，是因为铁磁物质不同原子的磁矩产生了相互平行的指向，就像这样(↑↑↑↑↑↑↑↑↑↑↑↑↑↑↑↑)。这些原子的磁矩从哪儿来呢？来自于铁磁原子未配对的核外电子的自旋，已经配对的那些电子由于泡利不相容原理，它们的自旋一定是相反的，因而磁矩总是抵消为0。那么不同原子的这些未配对电子的自旋为什么会平行地指向呢？为什么只有铁磁性物质会这样平行指向呢？为了对这一问题进行理论研究，海森堡提出，这是因为铁磁性物质邻近原子的未配对电子存在自旋与自旋间的相互作用，为了反映这种自旋相互作用是如何导致自发磁化的，海森堡提出了一个简单的模型，这就是今天依然有大量研究者在研究的著名的海森堡模型，这个模型对于铁磁性的研究，甚至对于整个凝聚态物理来说都非常重要，因此我们想介绍一下。海森堡模型的哈密顿算符是下面这样的

$$H = J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j, \quad (4.35)$$

式中的 $\langle i, j \rangle$ 表示两个邻近格点位置*i*和*j*, \vec{S}_i, \vec{S}_j 分别表示格点位置*i, j*上的原子的自旋, 式中的求和表示将每一对相互邻近的原子的贡献都加起来, 另外, J 的大小表示两自旋之间的耦合强度, 更常见的是将之写成 $-J$, 不过后面我们将对 J 取正和取负两种情况都会进行讨论。

对于一个任意的二维或三维格点, 求解海森堡模型不是一件容易的事情, 因此我们当然不可能在这里讨论这样的课题。以上所说主要是给我们提供一个物理背景, 我们真正将要研究的, 是只考虑两个原子, 每个原子只有一个未配对电子, 这两个电子的自旋之间按照海森堡模型的形式进行耦合。这时候, 由于电子的自旋算符可以用泡利算符来进行表达, 所以我们也可以将相应的模型写成下面的形式,

$$H = J(\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y + \sigma_1^z \sigma_2^z). \quad (4.36)$$

(4.36)才是我们这一节将要进行求解的问题。我们将会看到, 即使这样一个简单的一个模型也能告诉我们, 什么时候系统会出现铁磁性, 也就是两电子的自旋相平行, 什么时候系统又会出现反铁磁性, 也就是两电子的自旋反平行。不过, 首先我们想说一下, 什么是求解出这个系统? 我们这里所谓的求解出哈密顿算符(4.36), 实际上指的是求解出它的本征值, 也就是系统的本征能量 $\{E_n\}$, 以及相应的本征态 $\{|E_n\rangle\}$

$$H|E_n\rangle = E_n|E_n\rangle. \quad (4.37)$$

正如我们已经知道的, 由于在哈密顿算符的本征态 $|E_n\rangle$ 上, 系统有确定的能量取值 E_n , 所以这样的本征态也叫做定态, 相应的本征能量的集合也称作系统的能谱或者能级。为了求解这个定态薛定谔方程, 我们都需要取定一个表象(也就是选取一组正交归一的矢量基), 取定表象以后, 这个方程当然就变成了相应的哈密顿矩阵的本征方程。

4.2.2 求解与讨论

先总结一下上一节得到的关于泡利算符的一些有用结果。首先, 通常称上一节定义的 σ 为自旋降算符, 称 σ^\dagger 为自旋升算符, 这是因为 σ 的作用会把自旋向上态降为自旋向下, 而 σ^\dagger 的作用则会把自旋向下升为自旋向上,

如下

$$\sigma^\dagger |\downarrow\rangle = |\uparrow\rangle, \quad \sigma |\uparrow\rangle = |\downarrow\rangle. \quad (4.38)$$

其次, $|\downarrow\rangle$ 无法进一步再降了, 所以在 σ 的作用下必定为零, 同样 $|\uparrow\rangle$ 无法再升了, 所以在 σ^\dagger 的作用下也必定为零, 即

$$\sigma |\downarrow\rangle = 0. \quad (4.39)$$

$$\sigma^\dagger |\uparrow\rangle = 0. \quad (4.40)$$

公式(4.38)以及公式(4.39)、(4.40)这几个结果再加上下面的两个本征方程, 这些都是上一节得到的基本结论, 现在归纳在这里

$$\sigma^z |\uparrow\rangle = |\uparrow\rangle, \quad \sigma^z |\downarrow\rangle = -|\downarrow\rangle. \quad (4.41)$$

下面我们就可以开始求解(4.36)给出来的两自旋耦合系统了。首先, 两个不同电子的自旋当然是相互独立的, 因此任何两个属于不同电子的泡利算符都必定相互对易。其次, 类似于上一节的做法, 我们分别引入两个不同电子的自旋降算符和自旋升算符, $\sigma_1 = \frac{1}{2}(\sigma_1^x - i\sigma_1^y), \sigma_2 = \frac{1}{2}(\sigma_2^x - i\sigma_2^y), \sigma_1^\dagger = \frac{1}{2}(\sigma_1^x + i\sigma_1^y), \sigma_2^\dagger = \frac{1}{2}(\sigma_2^x + i\sigma_2^y)$ 。则很容易验证, 原来的哈密顿量(4.36)可以重写成

$$H = 2J(\sigma_1\sigma_2^\dagger + \sigma_1^\dagger\sigma_2) + J(\sigma_1^z\sigma_2^z). \quad (4.42)$$

现在, 我们给系统选定四个正交归一基矢量 $\{|\uparrow\uparrow\rangle, |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\downarrow\rangle\}$ (请注意这四个基矢量的排列顺序, 后面的处理要与这个顺序一致), 以基矢量 $|\downarrow\uparrow\rangle$ 为例, 它两个标记符号的前一个 \downarrow 表示第1个电子的自旋状态, 后一个符号 \uparrow 表示第2个电子的自旋状态, 其余三个基矢量也是用类似办法进行标记的。

为了求解我们的系统, 我们需要将哈密顿算符在这四个基矢量构成的矢量基中表示出来。为了做到这一点, 让我们首先注意一个简单的数学事实, 即, 如果有归一化的量子态 $|u\rangle, |v_1\rangle, |v_2\rangle$, 并且 $|v_1\rangle$ 和 $|v_2\rangle$ 正交, 若某算符 A 将 $|u\rangle$ 映射到 $|v_1\rangle$ 和 $|v_2\rangle$ 的某个叠加态, 比如 $A|u\rangle = \alpha_1|v_1\rangle + \alpha_2|v_2\rangle$ (α_1, α_2 为叠加系数), 那么利用 $|v_1\rangle$ 和 $|v_2\rangle$ 的正交关系可知必有 $\langle v_1|A|u\rangle = \alpha_1, \langle v_2|A|u\rangle = \alpha_2$ 。

因此为了求出哈密顿算符在我们的矢量基中的表示矩阵，我们只需要求出它对四个基矢量的作用。为了求出这样的作用，我们只需要反复利用上面归纳的简单公式(4.38)、(4.39)、(4.40)、以及(4.41)。比如， $H|\uparrow\uparrow\rangle = 2J(\sigma_1\sigma_2^\dagger + \sigma_1^\dagger\sigma_2)|\uparrow\uparrow\rangle + J(\sigma_1^z\sigma_2^z)|\uparrow\uparrow\rangle = J|\uparrow\uparrow\rangle$ (注意， $\sigma_1^\dagger|\uparrow\rangle = 0$)，即 $H|\uparrow\uparrow\rangle = J|\uparrow\uparrow\rangle$ 。类似的可以求出， $H|\downarrow\uparrow\rangle = 2J|\downarrow\uparrow\rangle - J|\downarrow\uparrow\rangle$ ， $H|\uparrow\downarrow\rangle = 2J|\downarrow\uparrow\rangle - J|\uparrow\downarrow\rangle$ ，以及 $H|\downarrow\downarrow\rangle = J|\downarrow\downarrow\rangle$ 。因此就容易写出 H 在矢量基 $\{|\uparrow\uparrow\rangle, |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\downarrow\rangle\}$ 中的表示矩阵 H (我们依然用 H 表示哈密顿算符的表示矩阵)，

$$H = J \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4.43)$$

很容易求出这个矩阵的四个归一化本征矢量

$$\psi_{1,+1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \psi_{1,0} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \psi_{1,-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \psi_{0,0} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}. \quad (4.44)$$

对应的四个本征值分别为 $J, J, J, -3J$ 。没有学过线性代数的读者，你只需要验证这四个本征矢量是上面给出的哈密顿矩阵 H 的本征矢量就可以了。

注意到，在任意表象中态与表示它的列矢量的关系是 $|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle$ ，具体到我们这里即是

$$|\psi\rangle = \begin{bmatrix} |\uparrow\uparrow\rangle & |\downarrow\uparrow\rangle & |\uparrow\downarrow\rangle & |\downarrow\downarrow\rangle \end{bmatrix} \begin{bmatrix} \langle\uparrow\uparrow|\psi\rangle \\ \langle\downarrow\uparrow|\psi\rangle \\ \langle\uparrow\downarrow|\psi\rangle \\ \langle\downarrow\downarrow|\psi\rangle \end{bmatrix}. \quad (4.45)$$

因此，分别以上面求出来的这四个本征矢量作为叠加系数，相应的就有原来的哈密顿算符 H 的四个本征态

$$|1, +1\rangle = |\uparrow\uparrow\rangle, \quad |1, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \quad |1, -1\rangle = |\downarrow\downarrow\rangle, \quad (4.46)$$

以及

$$|0,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad (4.47)$$

其中前面三个本征态是简并的，对应的本征值都为 J ，这三个简并态称作两 $1/2$ 自旋耦合系统的自旋三重态，最后一个本征态 $|0,0\rangle$ 对应的本征能量是 $-3J$ ，它称作两自旋耦合系统的自旋单态。

人们很容易看到，自旋三重态的三个态有一个共同点，那就是它们对于两个电子来说是对称的，即如果把电子1和电子2进行交换，那么这三个态都将保持不变，而且后面我们将看到它们的总自旋大小是两电子贡献相加的，是1，因此我们常常称三重态为自旋平行态。相反，自旋单态关于两个电子是反对称的，如果将电子1和电子2进行交换，那相应的自旋单态就会出一个负号，而且它的总自旋两电子抵消为0了，因此我们常常称单态为自旋反平行态。

在量子力学中，一个系统的能量最低的本征态又称作系统的基态，基态的能量就是量子系统可能具有的最低能量，比基态能量更高的定态就称作激发态。基态(有时候也包括低激发态)对于研究一个系统的性质而言尤其重要，原因在于，对于一个宏观系统而言，只要环境的温度足够低，那么它就会处在基态附近，因此基态和低激发态往往决定了一个系统的宏观行为。

那么我们的两电子自旋耦合系统的基态是什么呢？很显然，答案依赖于 J 是正还是负。如果 J 大于0，那么由于自旋单态的能量为 $-3J$ ，为最低能量，因此系统的基态就将是这样一个自旋反平行态，这时我们称系统处在反铁磁相。想反，如果 $J < 0$ ，那么能量为 J 的自旋三重态就将是系统的基态，注意，这时候系统的基态是简并的， $|1,+1\rangle, |1,0\rangle, |1,-1\rangle$ 张成了一个简并子空间，任何一个给定的系统当然不可能同时处在这三个态，所以它就必须从这三个简并基态所张成的简并子空间中随机选取一个态。另一方面，从我们原来的哈密顿算符(4.36)来看，三个空间方向地位完全平等，因此我们的系统当然是空间旋转不变的。但是，这个简并子空间中的任意一个态都有一个特定的总自旋方向，比方说， $|1,+1\rangle$ 态总自旋沿着 z 轴向上， $|1,-1\rangle$ 态总自旋反着 z 轴向下，因此不管系统从简并子空间中选择了哪个态，都意味着它选定了一个特定的空间方向，因此就破坏了空间旋转不

变性，由于这种对空间旋转不变性的破坏完全是系统自发的，而不是因为我们在系统的哈密顿算符中引入了某个特定指向的磁场之类的东西，所以人们常常称这种旋转对称性的自发破坏为自发破缺(注意我们不用破坏这个词，而是用自发破缺)！但是，不管系统破缺到简并子空间中的哪一个态，它都是一个自旋平行态(比如， $|1, 0\rangle$ 态的两电子也是自旋平行的，只不过平行的方向不是在 z 轴上而已)，这时候我们就称系统处在铁磁相。

从这个例子我们可以想见，更具一般性的海森堡模型(4.35)到底是铁磁的还是反铁磁的，这取决于 J 的符号，要从理论上解释某个材料的铁磁性，那我们就必须从理论上计算出一个小于0的 J 。但这是一件相当困难的事情，很少有人能从理论上计算出某个材料的 J 。因此人们通常采用的理论研究方案是，从某一个对系统的更基本描述出发，然后看看有没有一个机制能够让这个系统的有效描述可以是一个 J 为负的海森堡铁磁模型。如果理论上找到了一个这样的机制，那么人们就会设计实验来实现这样的机制，并看它是否能和理论推导一样使得系统处于铁磁相。总之，即使从海森堡提出他的模型到现在已经有很长时间了，关于铁磁性的研究依然一直是理论凝聚态的前沿领域之一，就是因为这种研究并不容易。

两自旋耦合系统的总自旋

让我们再次回到两自旋耦合的哈密顿算符(4.36)，很显然，我们可以用泡利算符的矢量形式将它重写成 $H = J\vec{\sigma}_1 \cdot \vec{\sigma}_2$ 。而且，我们可以引入一个总泡利算符 $\vec{\sigma} = \vec{\sigma}_1 + \vec{\sigma}_2$ ，它和系统总自旋 \vec{S} 之间的关系是 $\vec{S} = \hbar\vec{\sigma}/2$ ，所以研究总泡利算符等价于研究系统总自旋。让我们先来考察一下 $(\vec{\sigma})^2 = (\vec{\sigma}_1)^2 + (\vec{\sigma}_2)^2 + 2\vec{\sigma}_1 \cdot \vec{\sigma}_2$ ，而由于 $(\vec{\sigma}_1)^2 = (\sigma_1^x)^2 + (\sigma_1^y)^2 + (\sigma_1^z)^2 = 3$ ，同样 $(\vec{\sigma}_2)^2 = 3$ ，所以我们可以知道 $\vec{\sigma}_1 \cdot \vec{\sigma}_2 = (\vec{\sigma})^2/2 - 3$ ，因此原来的哈密顿算符就可以重写成 $H = J[(\vec{\sigma})^2/2 - 3]$ 。

显然，算符 $(\vec{\sigma})^2$ 和哈密顿算符 H 的本征态是一样的，由前面求出来的 H 的本征值和本征态我们容易知道，自旋三重态 $|1, +1\rangle, |1, 0\rangle, |1, -1\rangle$ 和自旋单态 $|0, 0\rangle$ 分别是 $(\vec{\sigma})^2$ 的本征值为8和本征值为0的本征态。如果用总自旋 $(\vec{S})^2 = (\vec{\sigma})^2\hbar^2/4$ 来考虑的话，那么自旋三重态就是 $(\vec{S})^2$ 的本征值为 $2\hbar^2$ 的本征态，通常我们说这种态的总自旋为1，这就是三重态的记号中前一个1所代表的含义，不过，我们不用管为什么把这样的量子态称之为

自旋为1，我们只需要将它理解成一个约定，其含义就是，这种态对应于 $(\vec{S})^2 = 2\hbar^2$ ，也就是说，这三个态的总自旋大小是一样的，都是 $\sqrt{2}\hbar$ 。当然，用 $(\vec{S})^2$ 来考虑，自旋单态就对应于 $(\vec{S})^2 = 0$ ，也即是说，自旋单态是总自旋为0的态，这就是 $|0, 0\rangle$ 这个记号中前一个0的含义。

很明显， $(\vec{S})^2$ 的本征值不足以将两自旋耦合系统的四个本征态都区分出来，因为自旋三重态的那三个态对于 $(\vec{S})^2$ 来说是简并的。因此，人们通常进一步考察总自旋的z分量 $S_z = \sigma^z \hbar/2$ 。由于 $(\vec{\sigma})^2 = 6 + 2\vec{\sigma}_1 \cdot \vec{\sigma}_2 = 6 + 2(\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y + \sigma_1^z \sigma_2^z)$ ，我们很容易利用泡利算符的代数关系(4.18)验证 $[\sigma^z, (\vec{\sigma})^2] = 0$ ，这也即是说， S_z 和 $(\vec{S})^2$ 对易，

$$[(\vec{S})^2, S_z] = 0. \quad (4.48)$$

因此，根据上一章讲述的定理， S_z 和 $(\vec{S})^2$ 可以有共同的本征态，实际上，由于 $S_z = \frac{\hbar}{2}\sigma^z = \frac{\hbar}{2}(\sigma_1^z + \sigma_2^z)$ ，人们很容易验证： $S_z|1, +1\rangle = \hbar|1, +1\rangle$ （即 $|1, +1\rangle$ 是 S_z 的本征值为1倍 \hbar 的本征态，这就是 $|1, +1\rangle$ 记号中后一个+1的含义）， $S_z|1, 0\rangle = 0|1, 0\rangle$ （这就是本征态记号中的那个0的含义）， $S_z|1, -1\rangle = -\hbar|1, -1\rangle$ （即本征值为-1倍的 \hbar ，这就是记号 $|1, -1\rangle$ 中那个-1的含义），类似的， $S_z|0, 0\rangle = 0|0, 0\rangle$ （这就是本征态记号中第二个0的含义）。

简单归纳一下，即，自旋单态总自旋大小为0，当然 S_z 也为0，自旋三重态总自旋为1，其 S_z 的值分别为 $+1, 0, -1$ （以 \hbar 为单位）。

第五章 量子纠缠

前面的章节都是成熟的理论体系，但是本章不同，量子纠缠依然是前沿研究的热门内容，因此本章只能算是一个初步的介绍。值得强调的是，对量子纠缠的深入研究无疑将处于整个物理学研究的核心。

5.1 两体纠缠态

有一条量子力学的基本原理，我们在前面的章节中已经凭借着直觉使用了，但却未加深思和检验，本章我们将看到这条原理会带来多少不可思议的事情。

这条原理涉及到两个子系统的复合系统，我们称这两个子系统为 A 和 B 。假设任意为这两个系统分别取一组可确定区分可能性完备集 $\mathcal{I}_A = \{(i_A), i_A = 1, 2, 3, \dots\}$ 以及 $\mathcal{I}_B = \{(i_B), i_B = 1, 2, 3, \dots\}$ ，注意， i_A 和 i_B 是两个相互独立的指标。则可以构造整个复合系统(记作 AB 系统)的一组可确定区分可能性完备集 $\mathcal{I} = \{(i_A, i_B), i_A \in \mathcal{I}_A, i_B \in \mathcal{I}_B\}$ ，由可确定区分性，我们有如下关系

$$\langle j_A, j_B | i_A, i_B \rangle = \delta_{i_A j_A} \delta_{i_B j_B}, \quad (5.1)$$

其中 $i_A, j_A \in \mathcal{I}_A$, $i_B, j_B \in \mathcal{I}_B$. 当然，这个关系可以解释成复合系统希尔伯特空间矢量基 $|i_A, i_B\rangle$ 的正交归一性。

假设再取 A 和 B 的另一组可确定区分可能性完备集 \mathcal{I}'_A 和 \mathcal{I}'_B ，则又可以构造复合系统 AB 的另一组可确定区分可能性完备集 $\mathcal{I}' = \{(i'_A, i'_B), i'_A \in \mathcal{I}'_A, i'_B \in \mathcal{I}'_B\}$. 现在我们可以考虑复合系统从可能性 (i_A, i_B) 到可能性 (j'_A, j'_B) 的跃迁过程，很显然，它实际上就是系统 A 从 (i_A) 跃迁到了 (j'_A) ，同时与之

相独立的，系统B从(i_B)跃迁到了(j'_B)，根据独立事件概率相乘的性质，我们有

$$P(i_A, i_B \rightarrow j'_A, j'_B) = P(i_A \rightarrow j'_A)P(i_B \rightarrow j'_B). \quad (5.2)$$

而跃迁概率是由跃迁幅的模方给出的，所以满足上面这一跃迁概率性质的最自然的跃迁幅关系是

$$\langle j'_A, j'_B | i_A, i_B \rangle = \langle j'_A | i_A \rangle \langle j'_B | i_B \rangle. \quad (5.3)$$

实际上，(5.1)式是(5.3)式的特殊情况，为了看出这一点，我们取 $\mathcal{I}'_A = \mathcal{I}_A$, $\mathcal{I}'_B = \mathcal{I}_B$, 从而(5.3)式告诉我们

$$\langle j_A, j_B | i_A, i_B \rangle = \langle j_A | i_A \rangle \langle j_B | i_B \rangle = \delta_{i_A j_A} \delta_{i_B j_B}, \quad (5.4)$$

正是(5.1)式。

对(5.3)式的一种自然解读是

$$|i_A, i_B\rangle = |i_A\rangle |i_B\rangle, \quad (5.5)$$

当然，同样也有 $|j'_A, j'_B\rangle = |j'_A\rangle |j'_B\rangle$ ，进而在算希尔伯特空间内积 $\langle j'_A, j'_B | i_A, i_B \rangle$ 时，应该分别将A的态与A的态进行内积，把B的态与B的态进行内积，然后再把结果乘起来，也就有了(5.3)式。这实际上也是对 $|i_A\rangle |i_B\rangle$ 这种乘法的一个定义，这种乘法我们在第三章讲量子不可克隆定理时就见过了，当时我们只是凭直觉来理解它，现在才开始追溯这种乘法到底是怎么来的。文献中通常称这种乘法为张量积，并记作 $|i_A\rangle \otimes |i_B\rangle$ ，本书采用更自然(但是数学上稍微含糊一点)的记号 $|i_A\rangle |i_B\rangle$ 。(5.5)式告诉我们，复合系统的矢量基可以由两个子系统矢量基的张量积给出，其物理含义就是，复合系统的可能性(i_A, i_B)由两个子系统相互独立的可能性(i_A), (i_B)并置而成。

\mathcal{I} 作为复合系统的一组可确定区分可能性完备集，则根据跃迁幅乘法，复合系统的任何量子态 $|\psi\rangle$ 当然都能用 \mathcal{I} 来展开

$$|\psi\rangle = \sum_{(i_A, i_B) \in \mathcal{I}} |i_A, i_B\rangle \langle i_A, i_B | \psi \rangle. \quad (5.6)$$

记跃迁幅 $\langle i_A, i_B | \psi \rangle = \psi_{i_A i_B}$ (为一个复数)，则根据(5.5)式可知

$$|\psi\rangle = \sum_{(i_A, i_B) \in \mathcal{I}} \psi_{i_A i_B} |i_A\rangle |i_B\rangle. \quad (5.7)$$

不难看出，当且仅当 $\psi_{i_A i_B}$ 具有 $\psi_{i_A i_B} = \phi_{i_A} \varphi_{i_B}$ 的形式时，复合系统AB的量子可能性 $|\psi\rangle$ 才能写成A和B两个子系统可能性的乘积(张量积)，此时即有

$$|\psi\rangle = \sum_{(i_A, i_B) \in \mathcal{I}} \phi_{i_A} \varphi_{i_B} |i_A\rangle |i_B\rangle = \left(\sum_{i_A \in \mathcal{I}_A} \phi_{i_A} |i_A\rangle \right) \left(\sum_{i_B \in \mathcal{I}_B} \varphi_{i_B} |i_B\rangle \right) = |\phi\rangle |\varphi\rangle,$$

式中 $|\phi\rangle = \sum_{i_A \in \mathcal{I}_A} \phi_{i_A} |i_A\rangle$, $|\varphi\rangle = \sum_{i_B \in \mathcal{I}_B} \varphi_{i_B} |i_B\rangle$. 这时候就称可能性 $|\psi\rangle$ 为可分解可能性，因为它可以按照子系统进行因式分解，或者称 $|\psi\rangle$ 为可分量子态。换言之，复合系统的可分量子可能性是由两个子系统相互独立的可能性 $|\phi\rangle$ 和 $|\varphi\rangle$ 并置而成。

如果 $\psi_{i_A i_B}$ 不具有 $\phi_{i_A} \varphi_{i_B}$ 的形式，那可能性 $|\psi\rangle$ 就不能按照子系统进行因式分解，可以称之为不可分解的整体可能性(整体就是指AB整体，子系统A和B都是这个整体的部分)，不过，由于历史原因，文献中通常称这时候的 $|\psi\rangle$ 为量子纠缠态。不可分解的整体可能性不能理解成A,B两个部分各自局部可能性的简单并置，所以它是整体性的。

假设记系统A的希尔伯特空间为 \mathcal{H}_A ，设其维数为 $\dim(\mathcal{H}_A) = N_A$ ，记系统B的希尔伯特空间为 \mathcal{H}_B ，其维数为 $\dim(\mathcal{H}_B) = N_B$. 再记整个复合系统的希尔伯特空间为 \mathcal{H} ，则由于 \mathcal{H} 的基矢量为所有可能的 $|i_A, i_B\rangle$ ，很显然一共有 $N_A N_B$ 个不同的基矢量，所以

$$\dim(\mathcal{H}) = N_A N_B = \dim(\mathcal{H}_A) \dim(\mathcal{H}_B). \quad (5.8)$$

通常称复合系统的希尔伯特空间为两个子系统希尔伯特空间的张量积，记作 $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. 它的意思就是， \mathcal{H} 中的所有态矢量，都具有(5.7)的形式，这样的态矢量由复数 $\psi_{i_A i_B}$ 参数化，这样的复数有 $N_A N_B$ 个。考虑到态矢量相差一个整体复数因子所描述的是同一个量子态，也就是 $\lambda|\psi\rangle$ 与 $|\psi\rangle$ 在物理上等价，即 $\lambda\psi_{i_A i_B} \sim \psi_{i_A i_B}$ ，所以， \mathcal{H} 的所有量子态由 $N_A N_B - 1$ 个复数参数化。另一方面，根据上面的讲述可知，所有可分量子态由复数 ϕ_{i_A} 和复数 φ_{i_B} 共同参数化，很显然这样的复数只有 $N_A + N_B$ 个，但是，考虑到 $\lambda\phi_{i_A} \sim \phi_{i_A}$ 、 $\eta\varphi_{i_B} \sim \varphi_{i_B}$ 所以实际上只有 $N_A + N_B - 2$ 个复数参数化了不同的可分量子态。

很显然，只要 $N_A > 1, N_B > 1$, 那么就有 $N_A N_B - 1 > N_A + N_B - 2$, 所以，复合系统的所有量子可能性中有大量的不可分解整体可能性，或者换句话说，希尔伯特空间 \mathcal{H} 中有大量的量子纠缠态。尤其是当 N_A, N_B 都很大时，量子纠缠态的数目将远远比可分量子态的数目多！

可分量子态和纠缠态有一个本质的区别，那就是 A, B 两个部分局部的任何操作都无法把一个可分量子态转化为一个纠缠态。为了看清楚这一点，我们假设有一个可分量子态 $|\phi\rangle|\varphi\rangle$ ，我们在 A, B 两个部分上分别进行一个任意的操作，使得它们分别经历一个幺正演化(也叫幺正变换)，演化算符分别为 U_A 和 U_B ， U_A 只作用在希尔伯特空间 \mathcal{H}_A 上， U_B 则只作用在 \mathcal{H}_B 上，则很显然，经过这个演化以后原来的可分量子态变为 $U_A U_B(|\phi\rangle|\varphi\rangle) = U_A|\phi\rangle U_B|\varphi\rangle = |\phi'\rangle|\varphi'\rangle$ ，其中 $|\phi'\rangle = U_A|\phi\rangle$ 依然是一个 A 的量子态，同样 $|\varphi'\rangle = U_B|\varphi\rangle$ 也依然是一个 B 的量子态。操作的最终结果 $|\phi'\rangle|\varphi'\rangle$ 显然依然是可分量子态。所以，局部的操作无法创造纠缠态，为了创造出整体的纠缠态，你需要让 A, B 两部分进行相互作用，进而产生一个整体的演化。

对纠缠态的刻画可以进一步简化，为此我们将注意力集中在(5.7)式上，应用线性代数的奇异值分解定理(这里略过，感兴趣的读者请自行参考线性代数的书籍)可以证明，对于每一个给定的 $|\psi\rangle$ 态，总可以通过合适地选取 \mathcal{I}_A 以及 \mathcal{I}_B ，使得(5.7)式简化成如下形式

$$|\psi\rangle = \sum_i \lambda_i |i\rangle|i\rangle. \quad (5.9)$$

注意到 $|i\rangle_B \sim e^{i\theta_i}|i\rangle_B$ ，所以我们只需要合适地调节系统 B 的这些相位 $e^{i\theta_i}$ ，就可以让上式中的 $\lambda_i \geq 0$ ，进而根据归一化条件 $\langle\psi|\psi\rangle = 1$ ，即可以得到 $\sum_i \lambda_i^2 = 1$ 。人们通常令 $\lambda_i^2 = p_i \geq 0$ ，所以

$$\sum_i p_i = 1, \quad (5.10)$$

并进而将(5.9)式写成

$$|\psi\rangle = \sum_i \sqrt{p_i} |i\rangle|i\rangle. \quad (5.11)$$

(5.11)式就叫做 A, B 两体复合系统量子态的**施密特分解**，其中非零的 p_i 数目就称作量子态 $|\psi\rangle$ 的**施密特数**。很显然，施密特数最小为1，这时候只有一个非零的 p_i ，不妨设为 p_1 ，根据(5.10)式，必有 $p_1 = 1$ ，从而这时候(5.11)式就相当于

$$|\psi\rangle = |1\rangle|1\rangle, \quad (5.12)$$

从而这时候 $|\psi\rangle$ 必为一个可分量子态(注意A的 $|1\rangle$ 和B的 $|1\rangle$ 不是一回事)。反过来也一样，可分量子态必定施密特数为1。所以，量子态 $|\psi\rangle$ 为可分量子态的充要条件是，其施密特数为1。从而，施密特数大于1的量子态都是纠缠态！从(5.11)式可以看出，施密特数最大为 $\dim(\mathcal{H}_A)$ 和 $\dim(\mathcal{H}_B)$ 中较小的那个。

因此，施密特数多少反映了两体纠缠的状况，施密特数越大就说明纠缠得越厉害。特别的，如果 $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$, 而且所有的 p_i 均相等，从而 $p_i = 1/d$, 这时候相应的施密特分解即为

$$|\Phi_d^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle|i\rangle, \quad (5.13)$$

这样的纠缠态就称作**最大纠缠态**。

多体量子纠缠

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

当然也可以仿照着前面关于两体纠缠定义多体量子纠缠态，不过，组成整体的部分一多，就会带来更多的复杂性，因此多体量子纠缠的刻画和

研究远没有两体量子纠缠那么成熟，因此，我们将不进行一般性的介绍，仅在后文以一个三体量子纠缠态的例子来说明多体纠缠的神奇之处。

5.2 纠缠态与混态

本节我们讲述纠缠态最神奇的一点特性，即如果AB整体处于纠缠态，那么你可以完全了解这个整体，但对A和B这两个部分却必然都带有无知。即，纠缠态是一种可以完全了解整体却无法同时完全了解部分的东西。这种特性当然是经典物理里完全没有的，对于经典的事物，你完全了解了整体必然意味着你同时了解了它的所有部分。但是，量子纠缠态不是这样的。

下面我们先作一些准备。

测量与期望值

在第二章关于量子力学基本原理的讨论中，我们得出过一个重要公式：对某个物理量A多次重复实验测得的平均值 $\langle A \rangle$ 等于算符A在态上的期望值，即

$$\langle A \rangle = \langle \psi | A | \psi \rangle, \quad (5.14)$$

$|\psi\rangle$ 为系统所处的量子态。我们也说过，物理量A的测量值是相应算符A的本征值，在单次实验中，我们测得哪个值是随机的，测得本征值 λ_i 的概率为 $p_i = |\langle i | \psi \rangle|^2$ ，式中 $|i\rangle$ 为与本征值 λ_i (请和上文施密特分解中的 λ_i 相区分)相应的本征态。这样算出来的概率 p_i (请和上文施密特分解中的 p_i 区分)当然也可以通过重复多次实验来检验，因此其计算公式应该也能写成期望值的形式，的确，假如定义投影算符 $P_i = |i\rangle\langle i|$ ，则人们很容易验证

$$p_i = \langle \psi | i \rangle \langle i | \psi \rangle = \langle \psi | P_i | \psi \rangle. \quad (5.15)$$

因此，不仅仅物理量的平均值，而且物理量值的概率分布，都可以表达成厄密算符期望值的形式。

以上讨论告诉我们，在量子力学中一切可以通过重复实验来进行检验的量都可以由一个适当厄密算符的期望值来计算，各种厄密算符的期望值就是我们能够从一个量子态中提取的所有信息。

算符求迹

一个矩阵所有对角元的和称为矩阵的迹，我们可以将迹的概念推广到线性算符。对于一个线性算符 \mathcal{O} ，我们记其迹为 $\text{Tr}(\mathcal{O})$ ，定义为

$$\text{Tr}(\mathcal{O}) = \sum_i \langle i | \mathcal{O} | i \rangle, \quad (5.16)$$

式中 $\{|i\rangle\}$ 为希尔伯特空间的一组正交归一矢量基。虽然为了求出算符的迹我们需要选取一个特定的表象，但是利用跃迁元乘法，不难验证算符的迹与具体表象的选取无关，具体证明我们留给读者自行思考。由此可以知道，一个算符的迹就是它在任何一个表象中表示矩阵的迹，特别的，对于厄密算符，我们可以将这个表象选为它的本征表象，这时候算符的迹其实就是所有本征值的和。

假设有两个算符 A 和 B ，利用跃迁元乘法，容易证明算符迹满足如下等式

$$\text{Tr}(AB) = \text{Tr}(BA). \quad (5.17)$$

证明如下， $\text{Tr}(AB) = \sum_i \langle i | AB | i \rangle = \sum_{i,j} \langle i | A | j \rangle \langle j | B | i \rangle = \sum_{i,j} \langle j | B | i \rangle \langle i | A | j \rangle = \sum_j \langle j | BA | j \rangle = \text{Tr}(BA)$ 。不过，值得说明的是，对于无穷维希尔伯特空间，并非所有算符的迹都是定义良好的，因此在无穷维希尔伯特空间上应用(5.17)式时得特别小心，有时候它并不成立。比方说，对于坐标算符和动量算符的对易子来说，如果简单地应用这个公式将会得出 $\text{Tr}([X, P]) = 0$ ，但实际上 $\text{Tr}([X, P]) = i\hbar\text{Tr}(1) \neq 0$ 。在这个例子中，(5.17)式之所以不成立，正是因为式中各算符的迹都不是良好定义的。

利用算符的迹，我们可以把算符的期望值公式重写成

$$\langle \psi | \mathcal{O} | \psi \rangle = \text{Tr}(|\psi\rangle\langle\psi| \mathcal{O}) = \text{Tr}(\rho_\psi \mathcal{O}). \quad (5.18)$$

式中 $\rho_\psi = |\psi\rangle\langle\psi|$ 。要证明这个式子，我们只需注意到

$$\text{Tr}(|\psi\rangle\langle\psi| \mathcal{O}) = \sum_i \langle i | \psi \rangle \langle \psi | \mathcal{O} | i \rangle = \sum_i \langle \psi | \mathcal{O} | i \rangle \langle i | \psi \rangle = \langle \psi | \mathcal{O} | \psi \rangle.$$

可见，为了计算算符的期望值，我们并不需要知道量子态 $|\psi\rangle$ ，而是只需要知道厄密算符 ρ_ψ 。 $|\psi\rangle$ 和 ρ_ψ 的一个重要区别是， $|\psi\rangle$ 可以相差一个非物理的

整体相位因子，变成 $e^{i\theta}|\psi\rangle$ ，而 ρ_ψ 在这种相位变换下不变，也即是说，它自动剔除了这一非物理的整体相位信息。

以上讨论告诉我们，从一个量子系统中能够提取出来的所有物理信息都包含在 $\rho_\psi = |\psi\rangle\langle\psi|$ 这一特殊厄密算符中。通常称这样的 ρ_ψ 为一个纯态密度算符。假设记 U 为时间演化算符，则由于 $\rho_\psi = |\psi\rangle\langle\psi|$ ，并且 $|\psi\rangle$ 会演化为 $|\psi\rangle \rightarrow U|\psi\rangle$ (这里选取的是薛定谔绘景)，从而不难看出 ρ_ψ 的时间演化为

$$\rho_\psi \rightarrow U\rho_\psi U^\dagger. \quad (5.19)$$

纠缠态与混态

下面进入本节的正题。假设我们有一个 AB 复合系统，它由 A, B 两个子系统复合而成，两个子系统的希尔伯特分别为 \mathcal{H}_A 和 \mathcal{H}_B 。假设整个复合系统处在 $|\psi\rangle$ 态，它当然也会对应一个纯态密度算符 ρ_ψ 。假设我们忽略 B 而仅仅只关心 A ，对 A 进行各种测量。比方说我们测量 A 的某个物理量 \mathcal{O}_A ，因此我们就要计算如下期望值，

$$\langle\psi|\mathcal{O}_A|\psi\rangle = \text{Tr}_{AB}(\rho_\psi \mathcal{O}_A). \quad (5.20)$$

式中 Tr_{AB} 表示在整个系统的张量积希尔伯特空间 $\mathcal{H}_A \otimes \mathcal{H}_B$ 上求迹。注意到 $\mathcal{H}_A \otimes \mathcal{H}_B$ 的基矢量 $|ij\rangle_{AB} = |i\rangle_A |j\rangle_B$ ，所以 $\text{Tr}_{AB} = \text{Tr}_A \text{Tr}_B$ ，式中 Tr_A 和 Tr_B 分别表示在 \mathcal{H}_A 和 \mathcal{H}_B 上求迹。由于 \mathcal{O}_A 只作用在 \mathcal{H}_A 上，与子系统 B 无关，所以 Tr_B 就直接作用在 ρ_ψ 上，根据施密特分解(5.11)式，即得到

$$\rho_A = \text{Tr}_B(\rho_\psi) = \sum_i p_i |i\rangle_A \langle i|. \quad (5.21)$$

这是一个只作用在 \mathcal{H}_A 上的厄密算符。而期望值公式(5.20)就变成

$$\langle\psi|\mathcal{O}_A|\psi\rangle = \text{Tr}_A(\rho_A \mathcal{O}_A). \quad (5.22)$$

从(5.21)式可以看出，施密特分解的 p_i 刚好是算符 ρ_A 的本征值，施密特分解的 $|i\rangle_A$ 则是相应的本征态，因为很容易验证相应的本征方程成立。

以上讨论告诉我们，如果仅仅只关心子系统 A ，那就无需知道整体的 $|\psi\rangle$ ，只需知道 ρ_A ，忽略 B 以后所有能从子系统 A 提取出来的信息都包含在厄密算符 ρ_A 中。我们称 ρ_A 为系统 A 的密度算符，也称之为 ρ_ψ 在 A 上的约化

密度算符，当然它不是纯态密度算符，因为一般来说现在这个密度算符不能写成 $|\phi\rangle_A\langle\phi|$ 这样的形式。很明显， ρ_A 满足

$$\mathrm{Tr}_A(\rho_A) = \sum_i p_i = 1. \quad (5.23)$$

完全类似的，如果我们是忽略 A 而仅关心 B ，那就会有

$$\rho_B = \mathrm{Tr}_A(\rho_\psi) = \sum_i p_i |i\rangle_B\langle i|. \quad (5.24)$$

由于 $\rho_A = \mathrm{Tr}_B(\rho_\psi)$ ，所以不难看出， B 部分局部的任何幺正操作 U_B 都无法影响子系统 A ，因为操作完成之后依然有 $\mathrm{Tr}_B(U_B\rho_\psi U_B^\dagger) = \mathrm{Tr}_B(U_B^\dagger U_B \rho_\psi) = \mathrm{Tr}_B(\rho_\psi) = \rho_A$.

如果原来 AB 整体的量子态 $|\psi\rangle$ 不是一个纠缠态，也就是说施密特数等于1，则仅有 $p_1 = 1$ 非零，则，由(5.21)可见，这时候必有

$$\rho_A = |1\rangle_A\langle 1|. \quad (5.25)$$

显然这是一个纯态密度算符， ρ_B 也一样。反过来，如果原来的 $|\psi\rangle$ 是一个纠缠态，施密特数大于1，则从(5.21)式可以看到，对于忽略了 B 的我们来说，子系统 A 完全可以看作是以 p_i 的概率随机地处于 $|i\rangle_A$ 态，由于施密特数大于1，那就有多个这样的 $|i\rangle_A$ ，所以我们说这时候系统 A 处于一个由 ρ_A 描述的混态，这时候 ρ_A 就是混态密度算符。同样，这时候 ρ_B 也为混态密度算符。所以，整个完整系统 AB 的纠缠态和子系统 A 以及 B 的混态是相对应的。特别的，不难看到与最大纠缠态对应的子系统混态密度算符为 $\frac{1}{\dim(\mathcal{H}_A)} \sum_i |i\rangle\langle i|$ ，这时候相应的混态就叫做最大混态。

混态是什么意思呢？由于它意味着以某个概率分布随机地处于多个态的“混杂”，所以，混态就意味着我们对系统有某种无知！这里多个量子可能性以一定的概率混杂就反映了我们的无知。另一方面，正如刚刚看到的，整个复合系统处于纠缠态就意味着作为部分的子系统必定为混态。这就说明，当整体处于纠缠态时，即使我们完全了解了整体，我们对每一个部分也都必然含有无知。特别的，如果整体处于最大纠缠态，则我们对部分就最无知。可以完全了解整体却无法同时完全了解部分，甚至可以对部分完全无知(最大混态)，这就是量子纠缠最奇怪的地方。

密度算符

当然，我们可以完全不提 B 部分，直接定义一个系统的密度算符，而不管它是否源于与另一个系统 B 的纠缠。一个系统的密度算符 ρ 包含了所有我们能从对这个系统的观测中提取出来的信息，根据上面的讨论可以知道，密度算符 ρ 应该满足如下性质：1. 它是一个厄密算符。2. 它的本征值都大于等于0，为正定算符。3. $\text{Tr}(\rho) = 1$ 。有了密度算符以后，对任何物理可观测量 \mathcal{O} 期望值的计算就可以表达为

$$\langle \mathcal{O} \rangle = \text{Tr}(\rho \mathcal{O}). \quad (5.26)$$

如果一个密度算符能够写成 $\rho = |\psi\rangle\langle\psi|$ 的形式，就称为纯态密度算符，它描述的系统状态就是纯态，否则就是混态密度算符，描述的状态就是混态。很显然，纯态密度算符额外满足

$$\rho^2 = \rho. \quad (5.27)$$

反过来，如果一个密度算符额外满足上式，则它的本征值必为0或1，由于密度算符所有本征值的和要等于1，因此它就必定能写成 $\rho = |\psi\rangle\langle\psi|$ 的形式($|\psi\rangle$ 为 ρ 的本征态)，从而必为纯态密度算符。对于纯态密度算符，我们必定有

$$\text{Tr}(\rho^2) = \text{Tr}(\rho) = 1. \quad (5.28)$$

对于一个任意的密度算符 ρ ，我们必可以通过求解它的本征方程将它对角化成如下形式

$$\rho = \sum_i p_i |i\rangle\langle i|. \quad (5.29)$$

密度算符的正定性告诉我们 $p_i \geq 0$ ， $\text{Tr}(\rho) = 1$ 的条件则告诉我们 $\sum_i p_i = 1$ ，所以 $p_i \leq 1$ ，等号仅当 ρ 为纯态密度算符时才成立。因此，对于混态密度算符，我们必有 $\text{Tr}(\rho^2) = \sum_i p_i^2 < \sum_i p_i = 1$ ，即对于混态密度算符，必有

$$\text{Tr}(\rho^2) < 1. \quad (5.30)$$

5.3 用例子说明量子纠缠态的神奇

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

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle). \quad (5.31)$$

这个量子态当然是一个量子纠缠态，而且是最大纠缠态。这就意味着处在 $|\phi^+\rangle$ 态上的两个电子共同形成了一个不可分解的整体，而这个不可分割的整体与两个电子之间的距离没有关系，也即是说，即使你将其中一个电子放在南昌，另一个放到天边，这两者遥遥相隔，但是它们却依然处于同一个不可分解的整体之中。

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

5.3.1 量子纠缠能实现超光速信息传递吗？

为了分析量子纠缠态能否实现超光速传递信息的问题，让我们假设某个实验室中制备了 N 对电子，每一对都处在纠缠态 $|\phi^+\rangle$ ，你和你的她分别持有每一对电子中的一个，你待在南昌，而她去了天边，因此你们分享着 N 个纠缠对，但是你们之间不能通信。在这种情况下，你能用你们之间分享的纠缠对来给她瞬时传递信息吗？比方说，你和她约定，如果你们在南昌养的那只猫死了，你就会对你的电子进行测量，而你的测量立即就会

影响她持有的另一个电子，因此看起来只要她接着对自己持有的那个电子进行测量就能获知猫死的信息。情况真是这样的吗？

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

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

为了下一步的分析，我们首先来看一下 x 方向的测量和 z 方向的测量有什么不同。量子力学的基本原理告诉我们，如果我们沿着 x 轴测量电子的自旋态，那么被测的电子就会跃迁到 x 方向上的两个自旋本征态 $|\uparrow_x\rangle \equiv |\rightarrow\rangle$ 和 $|\downarrow_x\rangle \equiv |\leftarrow\rangle$ 中的某一个，根据第三章中的相关讲述，这两

个本征态和 z 方向本征态的关系是

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle + |\downarrow_x\rangle), \quad |\downarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle - |\downarrow_x\rangle). \quad (5.32)$$

当然，反过来也有

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle), \quad |\downarrow_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle). \quad (5.33)$$

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

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\uparrow_x\rangle + |\downarrow_x\downarrow_x\rangle). \quad (5.34)$$

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

现在假设你沿着 x 轴测量，之后她再沿着 z 轴进行测量，我们来看她是否能得到不同的结果。由于纠缠态 $|\phi^+\rangle$ 也可以写成(5.34)的形式，所以你沿着 x 轴的测量结果将是，大约有一半你所持有的电子会跃迁到 $|\uparrow_x\rangle$ 态，当然由于纠缠态的性质(5.34)，这时候她相应的另一个电子也会立即跃迁到 $|\uparrow_x\rangle$ 态，同样，你另半数电子会跃迁到 $|\downarrow_x\rangle$ 态，这时候她的相应电子也会立即跃迁到 $|\downarrow_x\rangle$ 态，当然这种跃迁是完全随机的。因此在你的测量完

成以后，她所持有的每一个电子有 $1/2$ 的概率处在 $|\uparrow_x\rangle$ 态， $1/2$ 的概率处在 $|\downarrow_x\rangle$ 态，当然由于她还不知道你的测量方式，所以对于这个结果她是并不知情的。她只是选择沿着 z 轴进行她的测量，由(5.33)式可以知道，这时候如果她测的电子处在 $|\uparrow_x\rangle$ 态，那她的测量将会使得这个电子以 $1/2$ 的可能性跃迁到 $|\uparrow\rangle$ ，还有 $1/2$ 的可能性跃迁到 $|\downarrow\rangle$ ，如果她测的电子处在 $|\downarrow_x\rangle$ 态，结论也是一样的，因此总的来说，对于她所持有的每一个电子，她都有 $1/2$ 的概率测到它自旋向上， $1/2$ 的概率测到它自旋向下。因此，当她完成所有的测量以后，她同样发现，在她所持有的电子中，大约有一半自旋向上，有一半自旋向下，而且这个结果是完全随机的。你已经看到了，她所得到的这个结果和你沿着 z 轴进行测量时她所得到的结果完全一样。也即是说，远在天边的她根本无从推断出你的测量方式，因此当然也就无法得知猫的死活。因此，用这种方式同样无法实现信息的超光速传递。

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

假设你们之间能够正常通信，那么即使你不直接告诉她你的测量方式，你也还可以告诉她你的测量结果，也就是说，你也还可以告诉她你所测的每一个电子自旋是沿着你测量轴的正方向，还是沿着你测量轴的反方向，而她得知你的这些信息就能进一步推断出你是沿着 z 轴进行测量，还是沿着 x 轴进行测量。不妨假设你沿着 x 轴进行测量，因此在你的测量完成之后所有电子都跃迁到了某个 x 方向的本征态，这时候如果她也沿着 x 轴进行测量，那么由于纠缠的性质，她会发现她所测的每一个电子结果都和你的测量结果一样(指电子自旋是沿着测量轴还是反着测量轴的结果。注意，你已经告诉她这个测量结果了)。但是，如果她是沿着 z 轴进行测量，这时候由于她的电子已经处在 x 方向的某个本征态，而 x 方向的本征态跃迁到 z 方向的本征态是完全随机的，所以她会发现她所测的电子到底是沿着测量轴的正方向还是反方向，与你告诉她的测量结果之间完全没有关联。所以，

通过这种将她的测量结果和你的测量结果相比较的方式，她就能推断出你的测量方式与她的是否一样，从而也就知道了你是沿着 z 轴测量，还是沿着 x 轴进行测量。这样，关于你的测量方式的信息就成功传递给她了。

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

如何提取纠缠态中的信息

在量子计算和量子信息科学中，人们将电子自旋这样的例子抽象成了量子比特模型，所谓的量子比特，就是一个只有0和1这两种可确定区分可能性的量子系统，因此，一个量子比特的希尔伯特空间就是两维的，其正交归一基矢量可以记作 $\{|0\rangle, |1\rangle\}$ 。当然，我们可以用电子自旋来实现量子比特，比方说取自旋向上态 $|\uparrow\rangle$ 代表 $|0\rangle$ 态，自旋向下态 $|\downarrow\rangle$ 代表 $|1\rangle$ 态。这样定义的话，那之前第三章中对于电子自旋定义的算符 X 和 Z (其实也就是泡利算符 σ_x 和 σ_z)现在就成了，

$$\begin{aligned} X|0\rangle &= |1\rangle, & X|1\rangle &= |0\rangle \\ Z|0\rangle &= |0\rangle, & Z|1\rangle &= -|1\rangle. \end{aligned} \quad (5.35)$$

X 和 Z 当然都是厄米算符，不仅如此，还容易验证 $X^2 = Z^2 = 1$ (泡利算符的性质)，所以 X 和 Z 还是么正算符。同样也可以验证 ZX 为么正算符。

因此，之前我们所讨论的两电子自旋的纠缠对 $|\phi^+\rangle$ 就可以重写成两个量子比特的纠缠对 $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ 。但是两个量子比特的希尔伯特空间是 $2 \times 2 = 4$ 维的，它应该有4个正交归一的基矢量，假设我们将 $|\phi^+\rangle$ 选作其中一个基矢量，那么其余三个基矢量可以怎么选择呢？事实上我们可以将其余三个基矢量也选作纠缠态(并且是最大纠缠态)，

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \quad (5.36)$$

$$|\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \quad (5.37)$$

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

现在，假定你和你远在天边的她是同一个导师的学生，你们的导师将两比特的重要信息(这就有四种可能性)用这四个贝尔态(它们构造四个可确定区分的不可分解整体可能性)来编码，每种可能性对应一个贝尔态，并且导师将这个贝尔纠缠态中的第一个量子比特发给了南昌的你，而将与之纠缠的另一个量子比特发给了远在天边的她。即是说，如果你们想获知导师给出的具体是什么信息，你们就得确定共享的是四个贝尔态中的哪一个。

你们如何能确定这一点呢？由于相隔遥远，所以你们首先想到的可能就是各自独立地对手中的那个量子比特进行测量。但是，显然的是，只要你们不相互通信，那你们就不可能从这种测量中获取任何信息，因为从这四个贝尔态的表达式(5.37)可以知道，不管你们共享的是哪一个贝尔态，你的测量将总是以 $1/2$ 的可能性得到0，以 $1/2$ 的可能性得到1，结果完全是随机的，你完全无法从中得知贝尔态的任何具体信息，当然她也一样。这是贝尔态和通常的单个量子比特量子态的根本性不同，对于通常的量子比特，如果你存进一比特的信息，比如将它存为 $|0\rangle$ 态，那你只要测量这个量子比特的值，就一定能把这个信息读取出来。但是，现在你导师明明存进了两个比特的信息，但你和她只要相互保持独立不进行沟通，就什么信息也读取不出来。也即是说，信息是隐藏的，是存储在你们共享的纠缠对的整体之中。

你可能会问，那这四个贝尔态到底如何区分呢？答案很简单，只要同时测量算符 Z_1Z_2 和算符 X_1X_2 的值，这里写在算符上的下标表示这个算符仅对相应的量子比特进行作用，下标1就表示这个算符仅作用在纠缠对的第1个量子比特上，下标2就表示这个算符仅作用在纠缠对的第2个量子比特上，当然，作用在不同量子比特上的算符是相互对易的。利用 $ZX = -XZ$ ，人们很容易验证 Z_1Z_2 和 X_1X_2 是对易的，因此可以有共同的本征态，实际上，贝尔基的四个量子态就是它们共同的本征可能性完备集。从贝尔态的表达式(5.37)可以很容易验证， Z_1Z_2 的两个不同本征值可以用来区分四个贝尔态是属于 $|\phi\rangle$ 类型还是属于 $|\psi\rangle$ 类型，而 X_1X_2 的本征值可以用来进一步区分它们在这两个类型中的±号。但是，正如你已经看到

的，要完成这种区分，就需要对纠缠对中的两个量子比特进行联合测量，比如说，对算符 Z_1Z_2 的值进行测量。注意，这和你测 Z_1 ，她同时测 Z_2 有根本性的不同，因为在联合测量中可以仅仅测得 Z_1Z_2 的值而不必同时知道 Z_1 、 Z_2 分别是多少。实际上，要同时对 X_1X_2 和 Z_1Z_2 进行这种联合测量，就必须首先将你和她分别持有的量子比特放到一起来。也即是说，只要你们依然分别持有纠缠对中的一个量子比特，那就无法进行这样的联合测量，从而也就无法完整地读出导师存储的两比特信息。

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

5.3.2 量子密集编码

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

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

你如何做到这一点呢？首先，你们之间得事先约定好一种如何用四个贝尔态对应两比特经典信息的编码方式。其次，你注意到可以对自己所持有的量子比特进行四种不同的幺正变换，分别为恒等变换 I ，变换 Z_1 ，变换 X_1 ，以及变换 Z_1X_1 （利用泡利算符的性质，你很容易验证这四个算符的

么正性), 变换的结果你也很容易算出来,

$$1|\phi^+\rangle = |\phi^+\rangle, Z_1|\phi^+\rangle = |\phi^-\rangle, X_1|\phi^+\rangle = |\psi^+\rangle, Z_1X_1|\phi^+\rangle = |\psi^-\rangle. \quad (5.38)$$

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

5.3.3 量子隐形传态

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

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

那么你们有没有什么绝对安全的量子信息发送方法呢？有。答案就在于充分利用你们之间共享的纠缠对 $|\phi^+\rangle_{12}$ 。为了说清楚这一点，首先由贝尔态的定义式(5.37)，我们很容易得到

$$|00\rangle = \frac{1}{\sqrt{2}}(|\phi^+\rangle + |\phi^-\rangle), \quad |11\rangle = \frac{1}{\sqrt{2}}(|\phi^+\rangle - |\phi^-\rangle) \quad (5.39)$$

$$|01\rangle = \frac{1}{\sqrt{2}}(|\psi^+\rangle + |\psi^-\rangle), \quad |10\rangle = \frac{1}{\sqrt{2}}(|\psi^+\rangle - |\psi^-\rangle). \quad (5.40)$$

其次，我们注意到整个系统的量子态可以重写成

$$\begin{aligned} |\Psi\rangle_3|\phi^+\rangle_{12} &= (\alpha|0\rangle_3 + \beta|1\rangle_3)\frac{1}{\sqrt{2}}(|00\rangle_{12} + |11\rangle_{12}) \\ &= \frac{1}{\sqrt{2}}[\alpha|00\rangle_{31}|0\rangle_2 + \alpha|01\rangle_{31}|1\rangle_2 + \beta|10\rangle_{31}|0\rangle_2 + \beta|11\rangle_{31}|1\rangle_2] \\ &= \frac{1}{2}[(\alpha|0\rangle_2 + \beta|1\rangle_2)|\phi^+\rangle_{31} + (\alpha|0\rangle_2 - \beta|1\rangle_2)|\phi^-\rangle_{31}] \\ &\quad + \frac{1}{2}[(\alpha|1\rangle_2 + \beta|0\rangle_2)|\psi^+\rangle_{31} + (\alpha|1\rangle_2 - \beta|0\rangle_2)|\psi^-\rangle_{31}], \end{aligned} \quad (5.41)$$

式中第三个等号我们利用了公式(5.40)，并且由贝尔态的正交归一性可以知道，第三个等号右边的最终表达式中的4个态相互正交。因此，你只需要将你的第3个量子比特和你所持有的第1个量子比特放在一起，并同时对它们进行 Z_3Z_1 和 X_3X_1 的联合测量，测量的结果是，第3个量子比特和第1个量子比特所构成的系统将会跃迁到它们相应的四个贝尔态中的某一个。然后，你再把你的测量结果用经典的方式发送给天边的她，她如果得知你的结果是 $|\phi^+\rangle_{31}$ ，那就什么也不需要做，因为从式(5.41)可以知道，这时候她的量子比特2的量子态已经变成 $\alpha|0\rangle_2 + \beta|1\rangle_2 = |\Psi\rangle_2$ 态了，如果她得知你的结果是 $|\phi^-\rangle_{31}$ ，根据(5.41)式，那她就对她的量子比特进行 Z_2 的幺正变换，由于 $Z_2(\alpha|0\rangle_2 - \beta|1\rangle_2) = (\alpha|0\rangle_2 + \beta|1\rangle_2) = |\Psi\rangle_2$ ，因此变换以后她也将得到正确的量子态，同样的，如果你的结果是 $|\psi^+\rangle_{31}$ ，那她就对自己的量子比特进行 X_2 的幺正变换，根据(5.41)式，结果也将是正确的量子态，而如果你的结果是 $|\psi^-\rangle_{31}$ ，根据(5.41)式，那她就进行 Z_2X_2 的幺正变换，同样会得到正确的量子态。总之，在得知你的测量结果以后，她总可以让自己的量子比特变成 $|\Psi\rangle_2$ 态。这样一来，你的第三个量子比特的未知量子态就成功传送给她了。

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

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

5.3.4 GHZ态以及为什么爱因斯坦错了

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

在量子力学刚刚发展起来的时候，爱因斯坦认为量子力学理论是不完备的。因为他认为一个物理量的值总是存在的，或者说任何时候任何物理量总会有一个确定的值，虽然可能因为种种原因你不能测到这个值，但它总存在，爱因斯坦所谓的“即使你没有看月亮，月亮也存在”就是这个意思，由此他指出量子力学里面物理量的值之所以不确定，之所以我们只能测得物理量取值的一个概率分布，是因为我们还缺失了一些信息，爱因斯

坦称之为有一些隐变量，如果我们能进一步掌握这些缺失的信息，那量子力学将和经典力学一样，没有任何不确定性。总之，爱因斯坦认为物理量的值总是存在的，量子力学中的不确定性和概率的起源与我们日常概念中的概率起源一样，都是因为我们缺失了一些信息，因此他强烈反对玻尔和海森堡等人的不确定性和概率是世界的内在属性的观念，认为不确定性在量子力学理论中的存在只不过反映了量子力学理论的不完备性，也即是说量子力学理论没有把所有隐变量都包括进来。

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

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

$$|GHZ\rangle_{123} = \frac{1}{\sqrt{2}}(|000\rangle_{123} + |111\rangle_{123}). \quad (5.42)$$

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

$$Z_1Z_2, \quad Z_2Z_3, \quad X_1X_2X_3, \quad (5.43)$$

利用泡利算符的反对易性，人们很容易验证这三个算符两两对易。现在我们进一步引入泡利算符 σ_y ，在这里写作Y，根据泡利算符的乘法规则我们有 $ZX = iY$ 。由于 $YY_1Y_2X_3 = -Z_1X_1Z_2X_2X_3 = -(Z_1Z_2)(X_1X_2X_3)$ (式中我们已经利用了不同量子比特的算符相互对易的性质)，因此我们可以知道， $|GHZ\rangle_{123}$ 也是 $Y_1Y_2X_3$ 的本征值为-1的本征态。类似的，我们可以得到，当

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

$$Y_1Y_2X_3 = -1, \quad X_1Y_2Y_3 = -1, \quad Y_1X_2Y_3 = -1, \quad X_1X_2X_3 = 1. \quad (5.44)$$

下面，让我们来假设爱因斯坦的观点也成立，即任何情况下，每一个物理量的值总是存在，我们将三个泡利算符的值分别记作 $v(X), v(Y), v(Z)$ ，当然由于泡利算符的平方等于1，所以它们的值只能取±1。请注意，作为物理量的值， $v(X), v(Y), v(Z)$ 都是普通的数，因此它们相互之间当然都相互对易。如果 $v(X), v(Y), v(Z)$ 总是存在，虽然可能由于我们不了解隐变量，缺失了信息，导致我们不能完全确定这些取值到底是多少，但按照爱因斯坦的观点，它们总是存在的，总是有定义的。如此一来，根据(5.44)，对于 $|GHZ\rangle_{123}$ 态，我们就必定有

$$v(Y_1)v(Y_2)v(X_3) = -1, \quad v(X_1)v(Y_2)v(Y_3) = -1, \quad (5.45)$$

$$v(Y_1)v(X_2)v(Y_3) = -1, \quad v(X_1)v(X_2)v(X_3) = 1. \quad (5.46)$$

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

因此，这也就是说，要么量子力学是错的，从而(5.44)式不对，要么爱因斯坦就是错的(也就是说我们不能假定 $v(X)$ 这样的量总是存在)，两者必居其一。到底谁错了呢？实验发现，爱因斯坦错了，(5.44)式是成立的。因此，这就从实验上否定了物理量的取值必定存在的观点。也从侧面证明了为什么在量子力学中物理量只能表示成算符，因为算符的值当然只在本征态上有定义，对于叠加态，谈算符的取值是没有意义的，因为它根本就不存在。

第六章 波动方程与路径积分

本书是从海森堡引入的矩阵力学开始，然后通过跃迁元概念以及跃迁元乘法，过渡到量子力学的基本原理，再过渡到量子力学的现代理论体系。然而，在历史上，矩阵力学出现不久薛定谔就提出了一种看起来完全不同，但实际却是等价的波动力学。在波动力学里，波函数成了基本概念，跃迁元通常是作为导出概念而引入的。但是，既然我们已经有了完整的量子力学理论体系，那逻辑当然也可以反过来，我们当然可以从量子力学的一般原理出发，进而导出薛定谔的波动力学。这也就是本章要做的事情之一。

本章的另一个任务是介绍量子力学的路径积分表述，这种表述是费曼在量子力学理论已经成熟之后才找到的，实际上，它可以作为跃迁幅乘法的一个漂亮应用。

本章主要以单自由度的一维运动为例，推广到三维，甚至推广到任意多自由度情形都是直接了当的。

6.1 薛定谔波动方程

本节都是在薛定谔绘景中考虑问题，因此所有的物理量算符均不随时间演化，我们将这些算符记作 A (而不是 A_0)这样的形式，注意，和之前的记号相比，现在省略了下标0。相应的，这些算符在坐标表象中的表示则记作 \hat{A} (请注意和第一第二章中的 \hat{A} 相区别)。

6.1.1 坐标表象

对于一个非相对论粒子来说，它的基本物理量就是坐标和动量，相应的算符分别为 X 和 P 。但是 X 和 P 并不对易，因此它们没有共同的本征表

象。所以，要么考虑算符 X 的本征表象，要么考虑算符 P 的本征表象，前者叫做坐标表象，后者称为动量表象。相比来说，坐标表象的用处更大一些，物理上也更直观一些，所以我们主要考虑这种表象。

首先，我们要取 X 的所有本征态 $|x\rangle$ 来构成一个可确定区分可能性完备集， $|x\rangle$ 满足如下本征方程

$$X|x\rangle = x|x\rangle, \quad (6.1)$$

它描述粒子处于位置 x 的这种可能性。这些量子可能性当然也要满足可确定区分性。

另一方面，我们已经熟知，可确定区分可能性完备集要满足 $\langle i|j\rangle = \delta_{ij}$ ，但是，这里的可能性 i, j 都是离散的可能性。而现在 x 则是连续的可能性，因此我们需要将克龙内克符号 δ_{ij} 推广到指标连续的情况，这就是所谓的狄拉克德尔塔函数，记作 $\delta(x - x')$ 。与 δ_{ij} 相比， i, j 不等时 δ_{ij} 取零，所以现在 $x \neq x'$ 时， $\delta(x - x')$ 等于零。但是现在， $x = x'$ 时， $\delta(x - x')$ 可不是等于1，而是等于 $+\infty$ ，稍后我们将看到，这是连续和离散的本质不同所带来的必然结果。

δ_{ij} 的一个基本性质是有如下方程成立

$$\sum_j \delta_{ij} f_j = f_i, \quad (6.2)$$

即所有对 j 的求和中，只有 $j = i$ 这一项有非零贡献。推广到连续指标就需要把求和换成积分，因此相应的推广就是 $\int dx' \delta(x - x') f_{x'} = f_x$ ，式中的积分区间是从负无穷到正无穷，习惯上常把 f_x 记作 $f(x)$ (因为它们均表示这个量对 x 有依赖)，因此就有

$$\int dx' \delta(x - x') f(x') = f(x). \quad (6.3)$$

特别的，取 $f(x) = 1$ ，即有

$$\int dx' \delta(x - x') = 1. \quad (6.4)$$

但是，前面说了 $x \neq x'$ 时， $\delta(x - x')$ 均为零，现在它的积分得等于1，因此就说明 $x = x'$ 时， $\delta(x - x')$ 得为 $+\infty$ 。所以，克龙内克符号推广到连续指标

的这个 $\delta(x - x')$ 是一个很奇怪的函数，假设取定 x' ，让 x 变动，那这个函数描述的是 $x = x'$ 位置的一个无限尖锐的脉冲，并且这个脉冲的积分等于1。实际上，严格来说， $\delta(x - x')$ 并不是一个函数，而是一种被称为广义函数的东西，它的最核心特征，就是满足方程(6.3)。

有了 δ_{ij} 的推广以后，我们就可以把可确定区分性的 $\langle i|j \rangle = \delta_{ij}$ 推广到连续情形，因此就有

$$\langle x|x' \rangle = \delta(x - x'). \quad (6.5)$$

这就是位置本征可能性的可确定区分性，或者也叫做位置本征态的正交归一关系。

对于任意量子态 $|\psi\rangle$ ，根据跃迁幅乘法我们有 $|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle$ ，现在把中间可能性换成位置本征可能性，相应的当然就要把离散的求和换成积分，因此即有 $|\psi\rangle = \int dx|x\rangle\langle x|\psi\rangle$ ，其中展开系数 $\langle x|\psi\rangle$ 就叫做微观粒子的波函数 $\psi(x)$ ，即

$$\psi(x) = \langle x|\psi\rangle. \quad (6.6)$$

所以波函数就是一种特殊的跃迁幅，它表示粒子从可能性 ψ 跃迁到位置可能性(x)的跃迁幅。

利用波函数的概念，我们就可以计算两个量子态之间的内积，根据跃迁幅乘法

$$\langle \phi|\psi \rangle = \int dx\langle \phi|x \rangle\langle x|\psi \rangle = \int dx\phi^*(x)\psi(x) = (\phi, \psi), \quad (6.7)$$

其中第二个等号是利用了 $\langle \phi|x \rangle = \langle x|\phi \rangle^* = \phi^*(x)$ 。另外， (ϕ, ψ) 是坐标表象下的一个常用记号，它就定义成 $(\phi, \psi) = \int dx\phi^*(x)\psi(x)$ ，通常称作两个波函数的内积，很显然，它也满足

$$(\phi, \psi)^* = (\psi, \phi). \quad (6.8)$$

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

说，都具有 $\mathcal{O}(X, P)$ 这样的形式。这里 $\mathcal{O}(X, P)$ 表示某个 X 和 P 的函数，当然我们还要求它有厄米性。

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

$$\hat{A}\psi(x) = \langle x|A|\psi\rangle, \quad (6.9)$$

式中 $|\psi\rangle$ 是任意一个量子态， $\psi(x)$ 是它的波函数。很显然，如果抽象算符 A 作用在态 $|\psi\rangle$ 上得到 $|\phi\rangle$ ，即 $A|\psi\rangle = |\phi\rangle$ 。那么根据我们的定义(6.9)，就有 $\hat{A}\psi(x) = \langle x|A|\psi\rangle = \langle x|\phi\rangle = \phi(x)$ ，即有 $\hat{A}\psi(x) = \phi(x)$ 。也就是说， \hat{A} 在波函数上的作用与抽象算符 A 在相应狄拉克符号上的作用完全对应，因此它的确够成了抽象算符 A 的表示，也因此，只要 A 是一个线性算符，那么相应的 \hat{A} 也必然是一个线性算符。假设 A 的本征方程为 $A|u_n\rangle = \lambda_n|u_n\rangle$ ，那么相应就有 \hat{A} 的本征方程， $\hat{A}u_n(x) = \lambda_n u_n(x)$ ，两者的本征值和本征矢量都完全对应。

不仅如此，我们还可以计算任意算符 A 在任意两个态 $|\psi\rangle$ 和 $|\phi\rangle$ 之间的跃迁元 $\langle\phi|A|\psi\rangle$ ， $\langle\phi|A|\psi\rangle = \int dx \langle\phi|x\rangle \langle x|A|\psi\rangle = \int dx \phi^*(x) \hat{A}\psi(x) = (\phi, \hat{A}\psi)$ ，即有

$$\langle\phi|A|\psi\rangle = (\phi, \hat{A}\psi). \quad (6.10)$$

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

$$\hat{A}^\dagger\psi(x) = \langle x|A^\dagger|\psi\rangle. \quad (6.11)$$

因此我们同样有 $\langle\psi|A^\dagger|\phi\rangle = (\psi, \hat{A}^\dagger\phi)$ ，而由厄米共轭算符的定义 $\langle\psi|A^\dagger|\phi\rangle^* = \langle\phi|A|\psi\rangle$ ，因此 $(\psi, \hat{A}^\dagger\phi)^* = (\phi, \hat{A}\psi)$ 。又由于根据波函数内积所满足的(6.8)， $(\psi, \hat{A}^\dagger\phi)^* = (\hat{A}^\dagger\phi, \psi)$ ，因此我们就有

$$(\hat{A}^\dagger\phi, \psi) = (\phi, \hat{A}\psi). \quad (6.12)$$

这就是我们定义的 \hat{A}^\dagger 所满足的基本关系式，它和人们在波函数空间中对一个算符的厄米共轭算符的通常定义完全一样。根据这个结果我们可以知

道，任何一个抽象的厄米算符，其在坐标表象中的表示也必然是一个厄米算符！特别的，位置算符 X 的表示 \hat{X} 必然是一个厄米算符，动量算符 P 的表示 \hat{P} 也必然是一个厄米算符。我们下面要做的，就是根据我们对算符表示的定义(6.9)，求出 \hat{X} 和 \hat{P} 的具体形式。

但是首先我们需要知道，根据上面的论述，坐标表象中的算符所构成的算符代数应该和原来抽象算符的算符代数完全对应¹。特别的，由 $[X, P] = i\hbar$ ，必然有

$$[\hat{X}, \hat{P}] = i\hbar. \quad (6.13)$$

我们先来看位置算符。由于 $|x\rangle$ 是位置本征态，相应的坐标值是 x ，即 $X|x\rangle = x|x\rangle$ ，也即 $\langle x|X = x\langle x|$ 。因此根据我们的定义(6.9)， $\hat{X}\psi(x) = \langle x|X|\psi\rangle = x\langle x|\psi\rangle = x\psi(x)$ ，这也就是说，在坐标表象中，位置算符 \hat{X} 其实就是给波函数乘上 x ，你也可以简记为 $\hat{X} = x$ 。

为了求出坐标表象中的动量算符 \hat{P} ，我们注意到 $\hat{P}(\hat{X}\psi(x)) = \hat{P}(x\psi(x))$ ，另一方面，根据(6.13)式， $\hat{P}\hat{X}\psi(x) = \hat{X}\hat{P}\psi(x) - i\hbar\psi(x)$ ，也即 $\hat{P}(\hat{X}\psi(x)) = x(\hat{P}\psi(x)) - i\hbar\psi(x)$ ，与前面的结果比较，即有

$$\hat{P}(x\psi(x)) = -i\hbar\psi(x) + x(\hat{P}\psi(x)). \quad (6.14)$$

这让我们很自然地猜测

$$\hat{P} = -i\hbar\partial_x, \quad (6.15)$$

式中 ∂_x 表示对 x 求偏导。

由上面的猜测，我们容易得到 $[\hat{X}, \hat{P}]\psi(x) = (-i\hbar)(x\partial_x\psi(x) - \partial_x(x\psi(x))) = (-i\hbar)(x\partial_x\psi(x) - \psi(x) - x\partial_x\psi(x)) = i\hbar\psi(x)$ ，即 $[\hat{X}, \hat{P}]\psi(x) = i\hbar\psi(x)$ 。由于 $\psi(x)$ 为任意波函数，因此这就说明，上面的猜测的确能使得 $[\hat{X}, \hat{P}] = i\hbar$ 成立²，所以它是一个正确的猜测。因此，在坐标表象中，动量算符正比于一个对坐标 x 的偏导运算。

¹数学上的专业术语叫做两个算符代数同构

²值得说明的是，使得这个式子成立的 \hat{P} 并不唯一，以上猜测只是最简单的一种可能。不过， \hat{P} 的不同选取并不会导致不同的量子力学，因为这些不同选取之间只差一个非物理的规范变换。

6.1.2 德布罗意波和薛定谔波动方程

当然，我们也可以考察动量本征态(也就是有确定的动量取值的量子可能性)，即满足 $P|p\rangle = p|p\rangle$ 的本征态 $|p\rangle$ ，它们当然也满足正交归一性(注意到 p 同样是一个连续变量)

$$\langle p|p'\rangle = \delta(p - p'). \quad (6.16)$$

另一方面， $|p\rangle$ 在坐标表象中的波函数当然是 $\langle x|p\rangle$ ，根据定义，它应该满足 $\hat{P}\langle x|p\rangle = \langle x|P|p\rangle = p\langle x|p\rangle$ ，代入 $\hat{P} = -i\hbar\partial_x$ ，这也就是微分方程

$$-i\hbar\partial_x(\langle x|p\rangle) = p\langle x|p\rangle. \quad (6.17)$$

很容易积分求出这个微分方程的解，即是

$$\langle x|p\rangle = c \exp(ipx/\hbar), \quad (6.18)$$

式中 c 为积分常数。

另外，由正交归一关系 $\langle p'|p\rangle = \delta(p - p')$ ，利用跃迁幅乘法插入中间可能性，即有 $\delta(p - p') = \int dx \langle p'|x\rangle \langle x|p\rangle = \int dx \langle x|p'\rangle^* \langle x|p\rangle = \int dx |c|^2 e^{i(p-p')x/\hbar} = \hbar |c|^2 \int dy e^{i(p-p')y}$ ，式中最后一个等于号是利用了变量代换 $x = y\hbar$ 。代入数学公式(请自行查阅这个公式) $\int dx e^{ikx} = (2\pi)\delta(k)$ ，即有 $\delta(p - p') = |c|^2 (2\pi\hbar)\delta(p - p')$ ，两边比较，显然可以取 $c = 1/\sqrt{2\pi\hbar}$ ，从而即可以把(6.18)式写成

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp(ipx/\hbar). \quad (6.19)$$

上面这个式子告诉我们，一个动量为 p 的粒子，其到位置本征可能性的跃迁幅可以看作是一个平面波，这个波的波长是

$$\lambda = (2\pi\hbar)/p. \quad (6.20)$$

这一点不难验证， $\exp(ip(x + 2\pi\hbar/p)/\hbar) = \exp(ipx/\hbar) \exp(2\pi i) = \exp(ipx/\hbar)$ ，即这个波在空间上的确以上面给出的 λ 为周期。

粒子同时可以用一个波来描写！这是德布罗意的伟大想法，它拓展了爱因斯坦关于光的波粒二象性，进而认为所有的微观粒子都具有波粒二象

性。一个动量为 p 的粒子，其德布罗意波的波长正由上面的(6.20)式给出，这个波长就称之为德波罗意波长！因此，所有的微观粒子应该都可以进行波的干涉和衍射实验，实际上，德波罗意提出他的波粒二象性以后，这一点很快就被实验证实了。

(6.20)式告诉我们，德布罗意波长反比于粒子动量，动量越大则波长越短。而我们知道，当用一个波(比方说光波)去观察一个对象时，我们的观察有一个分辨率，这个分辨率取决于波长，比波长更小的空间尺度是无法分辨的。因此，如果我们想看清更小的更微观的结构，那就得用波长更短的波去观测。通常的可见光波长在几百个纳米，是原子尺寸的上千倍，因此如果用可见光去观测，那我们连原子都不可能看清。基于这个原因，整个20世纪物理学的很大一部分都在追求驾驭波长越来越短的光，或者其它微观粒子的德布罗意波，比方说电子的波布罗意波长比可见光就要短得多，因此可以用来观察晶体微观的晶格结构等等。当然，波长越短动量就越大，而动量越大就意味着能量越大，因此为了用微观粒子来看清世界更小更微观的结构，我们就得把这些粒子加速到很高的能量，这就是为什么要建造强大的粒子加速器的原因。因为，我们想看清在最基本的最渺不可见的尺度上，我们这个世界是怎么构成的。

当然，也可以定义动量空间波函数，为 $\tilde{\psi}(p) \equiv \langle p|\psi \rangle$ ，根据这个定义，它和坐标空间波函数 $\psi(x)$ 的关系为

$$\tilde{\psi}(p) \equiv \langle p|\psi \rangle = \int dx \langle p|x \rangle \langle x|\psi \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dx e^{-ipx/\hbar} \psi(x), \quad (6.21)$$

式中我们利用跃迁幅乘法插入了中间可能性(x)。完全类似的，通过插入中间可能性(p)，我们也有

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int dp e^{ipx/\hbar} \tilde{\psi}(p). \quad (6.22)$$

由此可见，动量空间波函数与坐标空间波函数的关系是互为傅里叶变换与逆变换的关系。

一维非相对论粒子的总能量等于动能与势能之和，相应的哈密顿算符应该是 $H = \frac{P^2}{2m} + V(X)$ ，式中 $V(X)$ 为势能算符。在薛定谔绘景中，量子态按照如下薛定谔方程随时间演化

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle. \quad (6.23)$$

假设记哈密顿算符 H 在坐标表象下的表示为 \hat{H} , 那么也就有 $\hat{H} = \frac{\hat{P}^2}{2m} + V(\hat{X})$, 代入上面找到的 \hat{X} 和 \hat{P} 的表达式, 即有 $\hat{H} = -\frac{\hbar^2}{2m}\partial_x^2 + V(x)$ 。又由 $i\hbar\partial_t\psi(x, t) = i\hbar\partial_t\langle x|\psi(t)\rangle = \langle x|i\hbar\partial_t|\psi(t)\rangle = \langle x|H|\psi(t)\rangle = \hat{H}\psi(x, t)$ (式中第三个等号我们使用了一般形式的薛定谔方程(6.23)), 我们就有 $i\hbar\partial_t\psi(x, t) = \hat{H}\psi(x, t)$ 。由此就可以知道一维非相对论粒子在坐标表象下的薛定谔方程为

$$i\hbar\frac{\partial}{\partial t}\psi(x, t) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi(x, t). \quad (6.24)$$

这就是最广为人知的薛定谔波动方程。这样, 我们就从最一般形式的薛定谔方程(6.23)推导出了薛定谔波动方程。

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

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x}, t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right]\psi(\mathbf{x}, t). \quad (6.25)$$

它描写微观粒子的波函数 $\psi(\mathbf{x}, t)$ 如何随着时间演化。

6.1.3 概率流密度

在第三章中我们已经看到, 量子态随时间的幺正演化一定会保持希尔伯特空间的内积不变, 特别的, 量子态的归一关系在时间演化下会保持不变, 即对任意 t 恒有 $\langle\psi(t)|\psi(t)\rangle = 1$ 。在中间插入位置可能性, 即有 $1 = \int dx\langle\psi(t)|x\rangle\langle x|\psi(t)\rangle = \int dx|\langle x|\psi(t)\rangle|^2$ 。注意到 $|\langle x|\psi(t)\rangle|^2$ 为 ψ 可能性的粒子在参考时刻 t 跃迁到位置可能性 x 的概率密度, 简称 t 时刻在 x 位置找到粒子的概率密度, 也就是单位空间体积内的概率(之所以是概率密度而不是概率, 是因为现在 x 是连续变量, 我们最终是要对 x 积分的), 通常记作 $\rho(x, t)$, 即 $\rho(x, t) = |\langle x|\psi(t)\rangle|^2$, 当然由于 $\psi(x, t) \equiv \langle x|\psi(t)\rangle$ 为粒子的波函数, 所以也可以把局域概率密度写成 $\rho(x, t) = |\psi(x, t)|^2$ 。所以, 归一化关系永远为1就意味着总概率 $\int dx\rho(x, t)$ 永远为1, 也就是总概率守恒。

一个自然的问题是, 除了总概率守恒之外, 在单粒子波动力学中有没有局域的概率守恒? 换言之, 我们想知道存不存在一个概率流密度 $J(x, t)$,

使得下面的连续性方程成立

$$\partial_t \rho + \partial_x J = 0. \quad (6.26)$$

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

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

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \psi(x, t) &= \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x, t). \\ -i\hbar \frac{\partial}{\partial t} \psi^*(x, t) &= \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi^*(x, t). \end{aligned} \quad (6.27)$$

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

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} (\psi^* \psi) &= \left(-\frac{\hbar^2}{2m} \right) \left[\psi^* \frac{\partial^2}{\partial x^2} \psi - \psi \frac{\partial^2}{\partial x^2} \psi^* \right] \\ &= \frac{\partial}{\partial x} \left[-\frac{\hbar^2}{2m} (\psi^* \partial_x \psi - \psi \partial_x \psi^*) \right]. \end{aligned} \quad (6.28)$$

很显然，最终的这个方程的确能够写成连续性方程(6.26)的形式，只需要我们取

$$\begin{aligned} J(x, t) &= \frac{-i\hbar}{2m} (\psi^* \partial_x \psi - \psi \partial_x \psi^*) \\ &= \frac{1}{2m} (\psi^* \hat{P} \psi - \psi \hat{P} \psi^*). \end{aligned} \quad (6.29)$$

这就是一维非相对论粒子概率流密度的表达式。

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

$$\partial_t \rho + \nabla \cdot \mathbf{J} = 0. \quad (6.30)$$

为了得到 $\mathbf{J}(\mathbf{x}, t)$ 的表达式，人们只需将(6.29)式中的一维动量算符替换成三维动量算符 $\hat{\mathbf{P}} = -i\hbar \nabla$ ，从而即有

$$\mathbf{J}(\mathbf{x}, t) = \frac{-i\hbar}{2m} (\psi^* \nabla \psi - \psi \nabla \psi^*). \quad (6.31)$$

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

为了看清楚局域概率守恒方程(也就是所谓的连续性方程)(6.30)的含义，我们将它在一个任意的三维区域 V 上积分，进而即有

$$-\partial_t \int_V d^3\mathbf{x} \rho(\mathbf{x}, t) = \int_V d^3\mathbf{x} \nabla \cdot \mathbf{J} = \oint_{\partial V} d\mathbf{S} \cdot \mathbf{J}, \quad (6.32)$$

式中 ∂V 表示三维区域 V 的边界， $\oint_{\partial V}$ 表示在这个边界上积分，这个式子的最后一个等号是用了多变量微积分中的高斯定理，其中 $d\mathbf{S}$ 表示边界上的面积元。上面这个最终结果的含义很清楚，那就是，区域 V 内总概率随着时间的减少率，等于单位时间之内通过区域边界流出去的概率。这就是局域概率守恒的真正含义。

6.2 路径积分

跃迁幅乘法最漂亮的应用之一，是可以用来导出量子力学的第三种表述形式，也就是费曼所建立起来的路径积分表述，本节就是要给出这一推导。这时候在海森堡绘景中进行讨论更加方便一点，所以与上一节不同，本节默认使用海森堡绘景。

路径积分在现代物理中的重要性是怎么强调都不为过的，比方说它可以很漂亮地建立起量子场论，尤其是量子的规范场论，而如果用传统的正则量子化办法就会困难得多。路径积分的一大好处是，它直接和作用量联系起来，而作用量是明显保持对称性的，所以路径积分量子化容易明显地保持对称性，尤其是在考虑了狭义相对论以后，它可以明显地保持狭义相对论的庞加莱对称性，这也是它广泛应用于相对论量子场论的原因之一。

下面开始推导。在前面的章节中我们引入过海森堡绘景，对于薛定谔绘景中的算符 \mathcal{O} ，其相应海森堡绘景中的算符 $\mathcal{O}(t)$ 为

$$\mathcal{O}(t) = \exp(iHt/\hbar)\mathcal{O}\exp(-iHt/\hbar). \quad (6.33)$$

另外，引入海森堡绘景的位置本征态 $|x, t\rangle$ 对于我们来说是方便的，其定义是

$$|x, t\rangle = \exp(iHt/\hbar)|x\rangle. \quad (6.34)$$

很容易验证， $|x, t\rangle$ 满足 $X(t)|x, t\rangle = x|x, t\rangle$ ，式中 $X(t)$ 为海森堡绘景中的位置算符。

记 x_i 为粒子的初始位置， x_f 为最终位置。下面我们将注意力集中在粒子从0时刻的初态可能性 $(x_i, 0)$ 跃迁到 T 时刻的末态可能性 (x_f, T) 的跃迁幅

$$\langle x_f, T | x_i, 0 \rangle. \quad (6.35)$$

通过在任意中间 t 时刻插入位置本征态的完备集，然后利用跃迁幅乘法，我们可以把上面的跃迁幅改写成

$$\langle x_f, T | x_i, 0 \rangle = \int dx \langle x_f, T | x, t \rangle \langle x, t | x_i, 0 \rangle. \quad (6.36)$$

这是跃迁幅满足的一个基本方程。现在，我们如图(6.1)中所示的那样，将

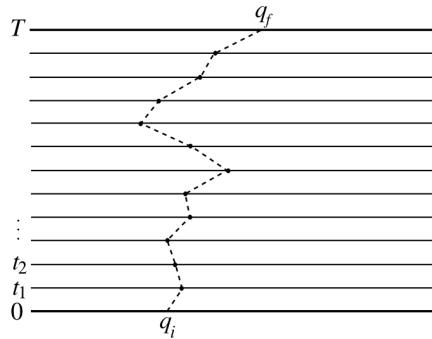


图 6.1：图中的 q_i, q_f 实际上分别是 x_i, x_f ，水平方向表示坐标 x ，竖直方向为时间。

整个时间区间 T 等分成 N 个间距为 $\epsilon = T/N$ 的小区间，其中各分割时刻分别为

$$t_m = m\epsilon. \quad (6.37)$$

则通过在每一个中间分割时刻都插入相应位置本征态的完备集并利用跃迁幅乘法，我们有

$$\langle x_f, T | x_i, 0 \rangle = \int dx_{N-1} \dots dx_1 \prod_{m=0}^{N-1} \langle x_{m+1}, t_{m+1} | x_m, t_m \rangle. \quad (6.38)$$

式中 $x_0 = x_i$, $x_N = x_f$ 。其含义就是 $(x_i, 0) \rightarrow (x_f, T)$ 的跃迁幅等于所有形如 $(x_i, 0) \rightarrow (x_1, t_1) \rightarrow (x_2, t_2) \rightarrow \dots \rightarrow (x_{N-1}, t_{N-1}) \rightarrow (x_f, T)$ 的逐时跃迁(如图(6.1))的跃迁幅之和(积分的本质就是求和), 这里的求和是指对所有 t_1, t_2, \dots, t_{N-1} 时刻的中间可能性求和(求积分)。

根据定义, 我们有 $|x_m, t_m\rangle = \exp(iHt_m/\hbar)|x_m\rangle$, 当然也有 $\langle x_{m+1}, t_{m+1}| = \langle x_{m+1}|\exp(-iHt_{m+1}/\hbar) = \langle x_{m+1}|\exp(-iH\epsilon/\hbar)\exp(-iHt_m/\hbar)$, 式中利用了 $t_{m+1} = t_m + \epsilon$, 从而即有

$$\begin{aligned}\langle x_{m+1}, t_{m+1}|x_m, t_m\rangle &= \langle x_{m+1}|\exp(-iH\epsilon/\hbar)|x_m\rangle \\ &= \int dp_m \langle x_{m+1}|p_m\rangle \langle p_m| \exp(-iH\epsilon/\hbar)|x_m\rangle.\end{aligned}\quad (6.39)$$

式中哈密顿算符 H 是一个关于坐标算符和动量算符的函数, $H(P, X)$, 而式中的 $|p_m\rangle$ 为动量本征态。通过算符对易关系我们总是可以将所有的 P 都放到 H 表达式的左边, 而将所有的 X 都对易到右边, 从而有(注意, 由 $P|p_m\rangle = p_m|p_m\rangle$, 有 $\langle p_m|P = p_m\langle p_m|$)

$$\langle p_m|H(P, X)|x_m\rangle = H(p_m, x_m)\langle p_m|x_m\rangle. \quad (6.40)$$

从而根据(6.39)式我们有(精确到 ϵ 的一次方阶)

$$\begin{aligned}\langle x_{m+1}, t_{m+1}|x_m, t_m\rangle &= \int dp_m \exp[-iH(p_m, x_m)\epsilon/\hbar] \langle x_{m+1}|p_m\rangle \langle p_m|x_m\rangle \\ &= \int \frac{dp_m}{2\pi\hbar} \exp\left\{-i[H(p_m, x_m)\epsilon - p_m(x_{m+1} - x_m)]/\hbar + O(\epsilon^2)\right\} \\ &= \int \frac{dp_m}{2\pi\hbar} \exp\left\{\frac{i}{\hbar}[p_m\dot{x}_m - H(p_m, x_m)]\epsilon\right\}.\end{aligned}\quad (6.41)$$

式中我们代入了动量本征波函数的表达式(6.19), 并且最后一行利用了 $x_{m+1} - x_m = \dot{x}(t_m)\epsilon$, 并忽略了 ϵ 的高阶项。

将(6.41)式代入(6.38)式, 就可以得到

$$\begin{aligned}\langle x_f, T|x_i, 0\rangle &= \int \frac{dp_{N-1}dx_{N-1}}{2\pi\hbar} \dots \frac{dp_1dx_1}{2\pi\hbar} \frac{dp_0}{2\pi\hbar} \\ &\times \exp\left\{\frac{i}{\hbar} \sum_{m=0}^{N-1} [p_m\dot{x}_m - H(p_m, x_m)]\epsilon\right\}.\end{aligned}\quad (6.42)$$

注意，随着离散时间的进行 $t_0 \rightarrow t_1 \rightarrow \dots \rightarrow t_{N-1}$ ，各离散时刻的坐标变量和动量变量也随之而演化为 $(x_i, p_0) \rightarrow (x_1, p_1) \rightarrow \dots \rightarrow (x_{N-1}, p_{N-1})$ ，也就是在 (x, p) 的空间中演化，这个空间当然也就是经典力学里熟知的相空间。不妨用相空间的直线把 (x_m, p_m) 与 (x_{m+1}, p_{m+1}) 连起来，如此一来，随着离散时间的演化，我们就得到一条在相空间中的折线路径， t_m 时刻路径经过相空间坐标 (x_m, p_m) 。当然，注意到(6.42)是要对所有 (x_m, p_m) ($N-1 \geq m \geq 1$) 积分的，这就相当于对所有可能的折线路径进行积分！

下面，我们取 $N \rightarrow +\infty$, $\epsilon \rightarrow 0$ 的极限，同时保持 $N\epsilon = T$ 固定。则在这个极限下，上面的相空间折线路径当然就变成了一条任意的路径，这时候(6.42)式中的复杂多重积分(实际上已经变成了无穷重积分) 就相当于对所有以 $x(0) = x_i, x(T) = x_f$ 为端点(端点处的动量并不固定)的相空间路径 $x(t), p(t)$ 进行积分。在这个极限之下，(6.42)式就趋近于

$$\rightarrow \langle x_f, T | x_i, 0 \rangle = \int [\mathcal{D}p(t) \mathcal{D}x(t)] \exp \left\{ \frac{i}{\hbar} \int_0^T dt [p\dot{x} - H(p, x)] \right\} \quad (6.43)$$

最后这个表达式中的 $\int [\mathcal{D}p(t) \mathcal{D}x(t)]$ 是对(6.42)式中的无穷重积分测度的简记符号。其中 $S[p(t), x(t)] = \int_0^T dt [p\dot{x} - H(p, x)]$ 就是所谓的相空间作用量，它是相空间路径的泛函。也即是说，跃迁幅 $\langle x_f, T | x_i, 0 \rangle$ 可以由所有以 $x(0) = x_i, x(T) = x_f$ 为端点(端点处的动量并不固定)的相空间路径 $x(t), p(t)$ 给出，每一条相空间路径的贡献是 $\exp(iS[p(t), x(t)]/\hbar)$ ，将所有这些相空间路径的贡献都加起来(积分无非就是加起来)，即得到跃迁幅 $\langle x_f, T | x_i, 0 \rangle$ 。这就是所谓的相空间路径积分。

如果我们是求算符 $\mathcal{O}(X, P)$ 的跃迁元，那相应于(6.43)式，就有

$$\begin{aligned} & \langle x_f, T | \mathcal{O}(X(t), P(t)) | x_i, 0 \rangle \\ &= \int [\mathcal{D}p(t) \mathcal{D}x(t)] \mathcal{O}(x(t), p(t)) \exp \left\{ \frac{i}{\hbar} \int_0^T dt [p\dot{x} - H(p, x)] \right\}, \end{aligned} \quad (6.44)$$

注意，等式右边路径积分中插入的是算符所对应的经典物理量 $\mathcal{O}(x(t), p(t))$ ，其中 x, p 都不是算符。

对于非相对论粒子，哈密顿量 $H(p, x)$ 通常可以写成如下形式

$$H(p, x) = \frac{p^2}{2m_e} + V(x), \quad (6.45)$$

式中 $V(x)$ 为势能, m_e 是粒子质量。因此(6.41)式的积分是

$$\langle x_{m+1}, t_{m+1} | x_m, t_m \rangle = \int \frac{dp_m}{2\pi\hbar} \exp \left\{ -\frac{i}{\hbar} \left[\frac{p_m^2}{2m_e} - p_m \dot{x}_m + V(x_m) \right] \epsilon \right\}. \quad (6.46)$$

这是一个所谓的高斯积分, 利用后面附录给出的高斯积分公式, 可以算出结果为

$$\langle x_{m+1}, t_{m+1} | x_m, t_m \rangle = \sqrt{\frac{m_e}{2\pi\hbar i\epsilon}} \exp \left(\frac{i}{\hbar} \left[\frac{1}{2} m_e \dot{x}_m^2 - V(x_m) \right] \epsilon \right). \quad (6.47)$$

将(6.47)式代入(6.38)式就可以得到

$$\langle x_f, T | x_i, 0 \rangle = \int dx_{N-1} \dots dx_1 \left(\sqrt{\frac{m_e}{2\pi\hbar i\epsilon}} \right)^N \exp \left(\frac{i}{\hbar} \sum_{m=0}^{N-1} \left[\frac{1}{2} m_e \dot{x}_m^2 - V(x_m) \right] \epsilon \right).$$

与前面相空间路径积分的处理类似, 保持 $N\epsilon = T$ 固定, 取 $N \rightarrow +\infty$, $\epsilon \rightarrow 0$ 的极限则这个结果就变成

$$\langle x_f, T | x_i, 0 \rangle = \int [\mathcal{D}x(t)] \exp \left(\frac{i}{\hbar} \int_0^T dt L(x, \dot{x}) \right), \quad (6.48)$$

式中 $L(x, \dot{x}) = \frac{1}{2} m_e \dot{x}^2 - V(x)$ 为坐标空间路径 $x(t)$ 的拉格朗日量, 式中积分测度 $\int [\mathcal{D}x(t)]$ 的定义为

$$\int [\mathcal{D}x(t)] \dots = \lim_{\epsilon \rightarrow 0, N \rightarrow +\infty} \int dx_{N-1} \dots dx_1 \left(\sqrt{\frac{m_e}{2\pi\hbar i\epsilon}} \right)^N \dots \quad (6.49)$$

可以理解为是对坐标空间中连接 $(x_i, 0)$ 与 (x_f, T) 两个端点的所有路径进行积分. 所以上面最终的结论(6.48)又叫做坐标空间(或者位形空间)的路径积分。

(6.48)式的含义可以大致理解为: 跃迁幅 $\langle x_f, T | x_i, 0 \rangle$ 由所有以 $x(0) = x_i$, $x(T) = x_f$ 为端点的坐标空间路径 $x(t)$ 给出, 每条路径的贡献是 $\exp \left(iS[x(t)]/\hbar \right)$ (其中 $S[x(t)] = \int_0^T dt L(x, \dot{x})$ 是坐标空间的作用量泛函), 将所有这些路径的贡献都加起来(积分无非就是加起来), 即得到跃迁幅 $\langle x_f, T | x_i, 0 \rangle$ 。可以示意性地写作

$$\langle x_f, T | x_i, 0 \rangle = \sum_{x(t)} e^{\frac{i}{\hbar} S[x(t)]}. \quad (6.50)$$

我们可以利用(6.50)式这个路径积分公式讨论量子力学与经典力学间的对应关系。由于两条邻近路径的作用量之差近似为一阶变分 δS , 而从(6.50)式的最后结果可以看出, 当 $\hbar \rightarrow 0$ 时, 一般来说相邻路径的相位差 $\frac{\delta S}{\hbar}$ 是随着路径的微小变动快速振荡的, 因此在 $\hbar \rightarrow 0$ 时, 相邻路径通常总是干涉相消的, 除非我们考虑的是 $\delta S = 0$ 这条路径的邻近路径。对于这条 $\delta S = 0$ 的路径, 它和邻近路径的相位差近似为0, 从而是干涉加强的。很显然, 这就是最小作用量原理, 这条干涉加强的路径就是所谓的经典路径。也即是说, 当 $\hbar \rightarrow 0$ 时, 坐标空间的路径积分即给出经典物理里的最小作用量原理。不难明白, 如果我们是对相空间路径积分进行这样的讨论, 那结果就是, 当 $\hbar \rightarrow 0$ 时, 相空间路径积分即给出经典物理里的相空间最小作用量原理。从这些讨论我们看到, 经典物理是量子物理在 $\hbar \rightarrow 0$ 时的极限。

即使从数学上看, 路径积分也是非常新颖的, 因为它涉及到无穷重的积分, 或者说在一个无穷维的空间上进行多重积分。这是此前(费曼之前)数学家所没有想象过的。正如读者从上面的推导中看到的, 这种积分的定义不是一件容易的事情, 尤其是到了量子场论, 相应路径积分的严格定义是一个重大的数学难题。虽然缺乏严格定义, 然而物理学家已经应用路径积分得到了大量的重要结果, 甚至包括一些困难的纯数学结果。所以, 这里的问题就有点像牛顿-莱布尼兹时代的微积分, 物理学家能用它得到大量结果, 但是微积分的严格理论基础在牛顿时代是缺乏的。路径积分, 尤其是场的路径积分, 其严格的数学基础今天也是缺乏的。某种意义上来说, 这也就是如何从数学上严格定义量子场论的问题, 当前只有很少一部分相对简单些的量子场论模型已经有了严格的数学基础, 大多数量子场论在数学上都是不严格的。但是, 如果能严格定义路径积分, 那大量量子场论模型的严格定义问题就解决了。

附录：高斯积分

所谓高斯积分, 是指形如下式的积分

$$\int dx \exp\left(-\frac{1}{2}ax^2 + bx\right), \quad (6.51)$$

式中的积分范围是从 $-\infty$ 到 $+\infty$. 换言之, 高斯积分的指数上是一个二次项

系数小于零的二次函数。

为了算出这个积分，我们首先把指数进行配方，即有

$$\begin{aligned} \int dx \exp\left(-\frac{1}{2}ax^2 + bx\right) &= \int dx \exp\left(-\frac{1}{2}a(x - b/a)^2 + \frac{1}{2}\frac{b^2}{a}\right) \\ &= e^{\frac{1}{2}\frac{b^2}{a}} \int dx \exp\left(-\frac{1}{2}ax^2\right) = e^{\frac{1}{2}\frac{b^2}{a}} \frac{1}{\sqrt{a}} \int dx \exp\left(-\frac{1}{2}x^2\right), \end{aligned} \quad (6.52)$$

式中第二行的第一个等于号是进行了 $x - b/a \rightarrow x$ 的变量代换，第二等于号是进行了 $\sqrt{ax} \rightarrow x$ 的变量代换。所以，最终我们真正要算的只有

$$I = \int dx \exp\left(-\frac{1}{2}x^2\right). \quad (6.53)$$

为了算出这个结果，我们考虑 I^2 ，很显然它等于

$$\begin{aligned} I^2 &= \int dx \exp\left(-\frac{1}{2}x^2\right) \int dy \exp\left(-\frac{1}{2}y^2\right) \\ &= \int \int dxdy \exp\left(-\frac{1}{2}(x^2 + y^2)\right). \end{aligned} \quad (6.54)$$

我们可以将这看成是二维 (x, y) 平面上的两重积分。为了算出它，我们可以在 (x, y) 平面上取极坐标 (r, θ) ，从而 $dxdy = r dr d\theta$, $x^2 + y^2 = r^2$, 从而上面的积分变成

$$\begin{aligned} I^2 &= \int \int r dr d\theta e^{-\frac{1}{2}r^2} \\ &= 2\pi \int_0^{+\infty} r dr e^{-\frac{1}{2}r^2} = 2\pi \int_0^{+\infty} dt e^{-t} = 2\pi. \end{aligned} \quad (6.55)$$

其中第二行利用了 $\frac{1}{2}r^2 = t$ 的变量代换。所以最终得到 $I = \sqrt{2\pi}$. 代入上面的(6.52)式，即得

$$\int dx \exp\left(-\frac{1}{2}ax^2 + bx\right) = \sqrt{\frac{2\pi}{a}} \exp\left(\frac{1}{2}\frac{b^2}{a}\right). \quad (6.56)$$

如果常数 a, b 中出现虚数，那你还是照常代上面的公式(6.56)。至于为什么这个公式可以推广到虚数的情况，则需要复变函数的知识才能证明，我们这里略去。

参考文献

给本书最大启发的，是海森堡开创量子力学的第一篇论文，《运动学与力学关系的量子理论重新诠释》，也就是那篇著名的“一个人的文章”。此外，玻恩和若尔当紧接着的“两个人的文章”，也就是《关于量子力学》，我也曾阅读过，虽然对本书的直接影响并不大。

影响本书的另一份文献是，《费曼物理学讲义第三卷》，前言中已经说过了，这对本书影响蛮大的。另外，跃迁元的概念借用自费曼的《量子力学与路径积分》，虽然其他人可能也用过这个概念。

当然狄拉克的《量子力学原理》也影响了本书，虽然本书对狄拉克符号的处理和狄拉克原始的处理并不同。

可确定区分性以及不可区分性等概念当然来自量子信息的文献，比如Charles H. Bennett 和Peter W. Shor 1998年那份简短的综述Quantum Information Theory. 不过，可确定区分可能性完备集这个概念的确是本书提出来的。

另外，本书的内容选材略微受到L. Susskind 《理论最小值：量子力学》的影响。当然本书不只是入门讲义，如果单纯当作入门讲义来看的话，那本书可能比Susskind 这本书要难一些。

本书的原始想法是我在写作《量子力学新讲》的时候产生的，当时只概括写成了一节，这里系统性地发展成了本书的前三章。另外，本书后三章约半数内容是在《量子力学新讲》中出现过的。