

目录

第一章 从实验引入量子力学	3
1.1 斯特恩-盖拉赫实验	3
1.2 级联斯特恩-盖拉赫实验	8
1.3 量子力学基本假设：量子态及其叠加原理	10
1.4 量子力学基本假设：投影测量	13
第二章 量子力学的理论结构	17
2.1 右矢	17
2.2 左矢与内积	19
2.3 基矢量	21
2.4 算符	25
2.5 算符的矩阵表示	27
2.6 量子力学基本假设：薛定谔方程	29
2.6.1 时间演化算符	29
2.6.2 薛定谔方程	32
2.7 量子力学基本假设：力学量算符	34
2.7.1 算符的本征值与本征态	34
2.7.2 力学量算符	35
2.7.3 算符的函数	38
2.7.4 相容力学量	39
2.7.5 简并与去简并	40
2.7.6 力学量完全集	42
2.8 不确定性关系	43
2.8.1 不确定性与不确定性关系	43
2.8.2 不确定性的物理意义	46

第一章 从实验引入量子力学

张越之

提起量子力学，你会想到什么？波粒二象性？不确定性？能量离散？观测效应？以波粒二象性为代表的一系列概念提出时量子力学还没有建立，那时可以称之为旧量子论的时代。待到量子力学在 1925 到 26 年间建立后，旧量子论的概念就必须重新审视。

什么是量子力学？让我们通过一个实验来开始讨论。

1.1 斯特恩-盖拉赫实验

1921 到 22 年间由奥托·斯特恩（Otto Stern）和瓦尔特·盖拉赫（Walther Gerlach）完成的斯特恩-盖拉赫实验（Stern-Gerlach experiment）是近代物理学史上最重要的实验之一，以下简称 SG 实验。SG 实验揭示了电子的固有自旋的存在，而级联 SG 实验又能体现量子力学的态叠加原理，SG 实验及其变种是讨论量子力学的最佳试验场之一，我们将由此迈进量子力学的大门。

实验者准备一个留有小孔的炉子，其中放有银材料，银被加热至沸点以上，剧烈的热运动使银原子逸出材料表面向各个方向飞散，一部分经孔洞离开炉子。在孔洞前放置数个带有小孔的挡板作为准直器，这样就得到了一束几乎沿同一个方向运动、准直性很好的银原子束。为排除各种干扰，整个实验装置要放置在高真空中。

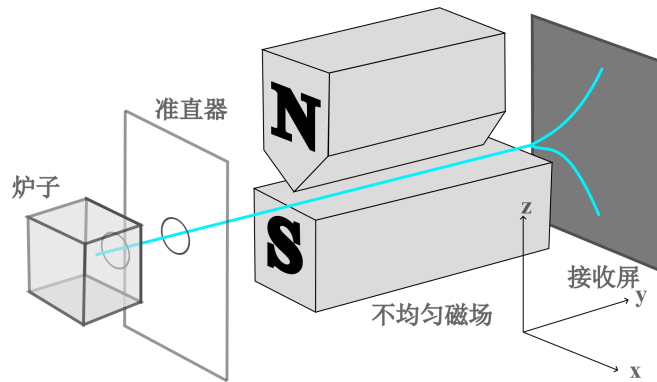


图 1.1. SG 实验装置示意图。实验装置的尺寸并不大，受益于 20 世纪初灯泡工业技术的发展，斯特恩和盖拉赫得以将实验装置整个置于真空中。你也许听说过，黑体辐射的研究最初就是由灯泡工业的需求推动的，谁能想到小小的灯泡竟映照出了一场物理学革命呢？

当时人们已经认识到粒子可以带有磁矩，且早在一个世纪前，安培已经由分子电流假说设想了类似的情况。根据经典电磁学，带有磁矩 $\vec{\mu}$ 、尺度足够小的物体¹，其在磁场 \vec{B} 中带有额外的能量

$$H = -\vec{\mu} \cdot \vec{B}, \quad (1.1)$$

这表明，为了尽可能降低能量以趋于稳定，磁矩会倾向于与磁场的方向相同。但这并不是像一块小磁铁被一块大磁铁吸引，把方向“扳过去”那样。实际上，带有磁矩的粒子²在磁场中受一个力矩

$$\vec{\mu} \times \vec{B}, \quad (1.2)$$

其使得磁矩方向绕着磁感线旋转，这称为拉莫尔进动（Larmor precession）。

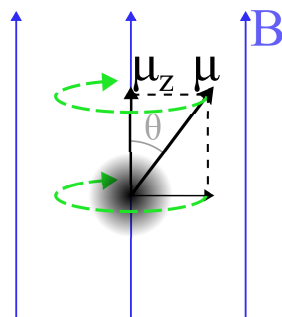


图 1.2. 拉莫尔进动示意图。带有磁矩 $\vec{\mu}$ 的粒子被置于 z 方向的匀强磁场 \vec{B} 中，“从上往下”看， $\vec{\mu}$ 将顺时针旋转。

¹足够小是指它在它占据的空间范围上，磁场的差异相对于磁场本身很小。这允许我们谈论物体所在位置的磁场是多少，对于尺度很大的物理问题会麻烦许多。

²正如“带有电荷的粒子”常常被简称为“电荷”一样，“带有磁矩的粒子”也可以简称为“磁矩”。当然，这种简称在我们的讨论超出电磁现象的范围时就不再合适了。

如图所示，当进动很快时（旋转周期远小于实验者的时间分辨能力），垂直于磁场的磁矩分量的效果平均而言为零，对外只显示平行于磁场的磁矩分量，即

$$\mu_z = |\vec{\mu}| \cos \theta. \quad (1.3)$$

这样看起来，磁矩的方向就和磁场相同了（但也变小了）。但是，携带着磁矩的物体本身的运动并不会受到影响，除非磁场是不均匀的。

当物体在空间不同位置可取得的能量不同时，就会感受到一个力，这个力会让物体倾向于往能让它能量更低的位置运动：

$$\vec{F} = -\nabla H = -\frac{\partial H}{\partial x} \vec{e}_x - \frac{\partial H}{\partial y} \vec{e}_y - \frac{\partial H}{\partial z} \vec{e}_z, \quad (1.4)$$

其中 $\vec{e}_x, \vec{e}_y, \vec{e}_z$ 为单位方向矢量。

例如，在地表附近的物体，其重力势能³在不同的高度不同：

$$H = mgz, \quad (1.5)$$

于是我们从能量的视角得到了重力：

$$\vec{F} = -\nabla H = -\frac{\partial H}{\partial z} \vec{e}_z = -mg\vec{e}_z. \quad (1.6)$$

你可能觉得这是徒增麻烦，“不是先有重力，后有重力势能吗？”，但在现代物理学中，能量要比力更加基本，而上面这样简单的换算也不再成立了。

说回正题。设银原子束的前进方向为 y 轴，并建立与其垂直的 z 轴和 x 轴。在前方设置一个方向与其垂直的非均匀磁场，如沿 z 轴正方向逐渐增强⁴：

$$\vec{B} = B_z \vec{e}_z, \quad \frac{\partial B_z}{\partial z} > 0, \quad (1.7)$$

对我们的实验设置而言，银原子束会受到一个平行于 z 轴的力：

$$F_z = \mu_z \frac{\partial B_z}{\partial z} = |\vec{\mu}| \cos \theta \frac{\partial B_z}{\partial z}, \quad (1.8)$$

这个力正比于磁场不均匀的程度 $\partial B_z / \partial z$ ，以及 μ_z ，或者说银原子磁矩 $\vec{\mu}$ 与 z 轴的夹角 θ ，使得银原子产生了平行于 z 轴的分速度，在 $y-z$ 平面上朝着 z 轴正方向或者负方向偏转。

³正如重力势能是物体与地球组成的系统共有的一样，磁矩在磁场中的能量也是其与磁场共有的。

⁴实际上，根据麦克斯韦方程组， $\nabla \cdot \vec{B} = \partial B_x / \partial x + \partial B_y / \partial y + \partial B_z / \partial z = 0$ ，一个物理上合理的磁场不可能仅在一个方向上非均匀。此处可以取 $\partial B_y / \partial y = 0$ ， $\partial B_x / \partial x = -\partial B_z / \partial z$ ，以及 $B_y = 0$ ，实验者需要用一些技巧来排除 x 方向不均匀性的干扰，以使我们能主要关心 z 分量的效果。

由于热运动， $\vec{\mu}$ 的方向是完全随机的， μ_z 应该在最大值（ $\theta = 0$ ，磁矩与磁场同向）和最小值（ $\theta = \pi$ ，磁矩与磁场反向）之间连续分布。每个银原子受到的力各不相同，因而偏转也各不相同。

在磁场出口放置一块冷却的玻璃板作为接收屏，让银原子冷凝在上面。持续“曝光”数小时后，用硫与其反应，产生黑色的硫化银痕迹，由此就可以读出银原子的偏转量。

按照上面的分析，我们预期会看到沿着 z 轴连续分布的痕迹，然而，实验中只看到了两道明确分开的痕迹。这只能解释为银原子磁矩分量 μ_z ，或者说 θ 仅有两种可能的取值，而非经典力学允许的无穷多种取值，这一现象称为角动量空间取向的量子化⁵。

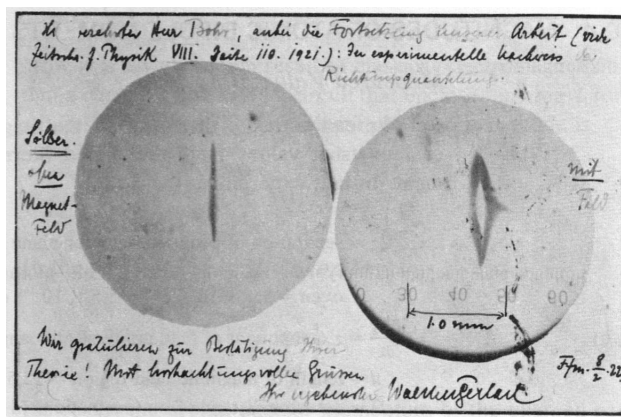


图 1.3. SG 实验的结果。实际的实验中使用了狭缝而非小孔来做准直器，左图为无磁场时银原子留下的痕迹，右图为加上不均匀磁场后银原子留下的痕迹。磁场方向在图中为左右方向，右图右边痕迹有一突出部分，这是磁铁最尖锐（磁场最不均匀）处造成的。图为 1922 年 2 月 8 日，盖拉赫寄给尼尔斯·玻尔（Niels Bohr）的明信片。

角动量，分子电流与磁矩

就像动量 \vec{p} 可以描述物体“包含多少平移运动”，角动量可以描述物体“包含多少旋转运动”。经典力学中，角动量定义为

$$\vec{L} := \vec{r} \times \vec{p}, \quad (1.9)$$

其中冒号和等号的组合 $:=$ 表示定义。连接物体与坐标原点的位置矢量 \vec{r} 的出现表明角动量是依赖于你选择的坐标原点的，即它不属于物体本身，而是物体和坐标系之间共有的。每当动量 \vec{p} 与 \vec{r} 不平行，物体就有绕原点旋转的趋势，不难想象吧？

基于安培的设想，我们可以将带磁矩的粒子想象为一个点电荷 e 绕着固定轴

⁵量子化（quantization）一词最初就是指物理量分立/离散的现象，这也是人们“量子 = 离散”这一刻板印象的来源。

做匀速圆周运动^a，角速度为 $\vec{\omega}$ ，位置矢量为 \vec{r} 。对轨道上的一个小截面，每隔周期 T 就通过电荷 e ，则平均电流为

$$I = \frac{e}{T} = \frac{\omega e}{2\pi}, \quad (1.10)$$

这样一个电流为 I ，半径为 r 的电流圈，其在磁场中受的合安培力为零，但是力矩非零：

$$\vec{M} = I\pi r^2 \vec{e}_n \times \vec{B}, \quad (1.11)$$

其中前半部分就称为磁矩， \vec{e}_n 为对电流用右手螺旋得到的单位矢量。即

$$\begin{aligned} \vec{\mu} &= I\pi r^2 \vec{e}_n \\ &= \frac{1}{2}e\omega r^2 \vec{e}_n, \end{aligned} \quad (1.12)$$

由

$$\begin{aligned} \vec{L} &= m\vec{r} \times (\vec{\omega} \times \vec{r}) \\ &= m\omega r^2 \vec{e}_n, \end{aligned} \quad (1.13)$$

得到

$$\vec{\mu} = \frac{e}{2m} \vec{L}. \quad (1.14)$$

即磁矩正比于角动量，比例系数称为**旋磁比**（gyromagnetic ratio）。磁矩和角动量的正比关系对服从量子力学的粒子也是成立的，但量子力学中角动量的含义已经和经典力学中有所不同。

上述的来自空间运动的角动量称为**轨道角动量**（orbital angular momentum），在建立量子力学的过程中，人们发现还存在一种与空间运动无关的角动量，称为**自旋角动量**（spin angular momentum）。

^a你也可以设想一个带电小球自转，但毕竟我们只是在做一个经典的玩具模型，在不引入量子力学的前提下增加细节并不会使我们更接近粒子的真实结构。

如果我们使用其它种类的原子进行实验，则痕迹条数会有不同（其原因在学习了角动量理论后会明白），但都是**分立的**。改变磁场的方向将改变痕迹分开的方向，但**痕迹的数目不变**。

今天的我们已经知道，银原子在实验中展现的磁矩完全由最外层电子的自旋产生⁶， μ_z 的两种可能取值对应电子自旋的两种状态，记为

$$|S_z+\rangle \text{ 和 } |S_z-\rangle. \quad (1.15)$$

⁶你也许了解一些化学知识，银的壳层结构是 $2|8|18|18|1$ ，由于球对称性，内层电子的磁矩都互相抵消了。银原子对外显示的净磁矩完全由最外层的 $5s$ 电子产生（原子核的磁矩远小于电子磁矩），而 s 轨道是没有轨道角动量的，因而只有自旋角动量有贡献。

实验测得电子自旋的两种取值分别为 $+\frac{\hbar}{2}$ 和 $-\frac{\hbar}{2}$ ，其中 \hbar 为约化普朗克常量，读作 h-bar，它和最初提出的普朗克常量的关系是⁷

$$\hbar = \frac{h}{2\pi} = 1.0546 \times 10^{-34} \text{J} \cdot \text{s}, \quad (1.16)$$

今天人们提到普朗克常量通常都指的是 \hbar 。

拿起一支笔，你可以将它指向随便什么方向，它和竖直方向的夹角在 0° 到 180° 之间任意取值，这支笔在竖直方向上的投影也就可以在整支笔的长度到零之间任意取值。然而 SG 实验告诉我们，一支表现得像电子自旋一样的笔将只能和竖直方向有两种夹角！“那好吧”，你可能以为这是因为竖直方向有某种特殊之处，“我看看这支笔现在指向什么方向，然后只要换一个参考的轴，就可以让笔和轴成任意角度了！”然而这是行不通的，下一节我们将看到，电子自旋不存在一个确定的方向——换言之，三个相互垂直方向上的分量是无法同时确定的！

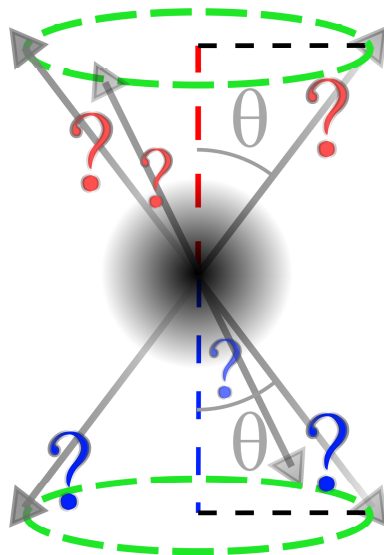


图 1.4. 电子磁矩与 z 轴只有两种可能的夹角，夹角约为 57.74° 。我们只能确定其在 z 轴（或其它轴）上的投影，无法确定其具体方向。

1.2 级联斯特恩-盖拉赫实验

接下来让我们通过级联（sequential）SG 实验进一步揭示电子自旋的性质。在前述的 SG 实验中，银原子束通过磁场后分裂为两束，若我们将向下偏转即 $|S_z - \rangle$ 的一束用接收屏吸收掉，让向上偏转即 $|S_z + \rangle$ 的另一束再次通过另一个同样的磁场，由于两个磁场之间并没有什么能改变电子磁矩方向的因素，可以预见，这束银原子仍会向上偏转。

⁷焦耳和秒乘起来就得到了角动量的单位，换言之能量乘时间和角动量的单位相同。熟悉物理量的单位将有助于记忆公式、检查公式是否有错误。角动量和作用量也具有相同的单位。

实际情况也确实如此，不同于初始从炉中射出的取向杂乱无章的银原子，我们用不均匀磁场和接收屏筛选得到了一束自旋 z 方向分量确定的银原子，可以用 $|S_z+\rangle$ 表示。

现在我们让 $|S_z+\rangle$ 通过一个 x 方向的不均匀磁场，会发生什么？

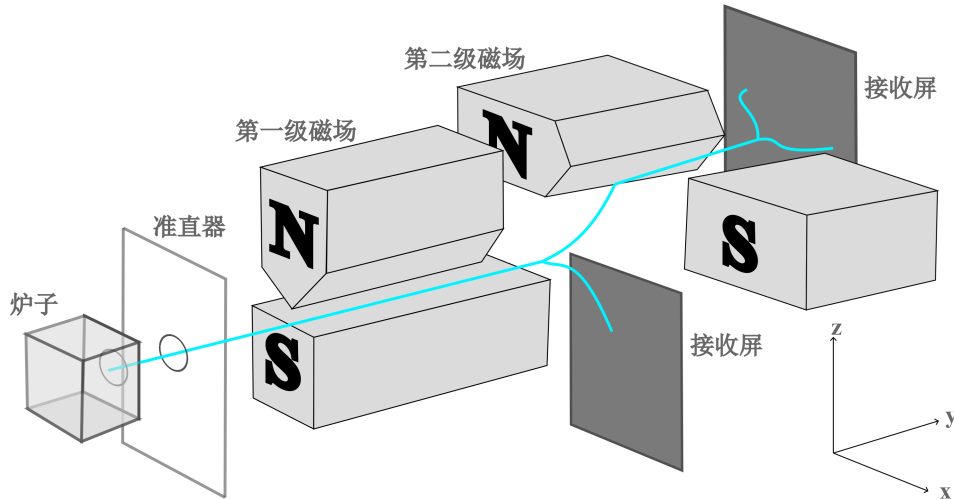


图 1.5. 级联 SG 实验示意图。 $|S_z+\rangle$ 束通过 x 方向不均匀磁场分裂为两束。

实验装置和结果如图1.5所示。 $|S_z+\rangle$ 银原子束在通过 x 方向不均匀磁场后，又分裂为两束，这说明它们表现出了不同的 μ_x 。两束的状态可分别记为 $|S_{x+}\rangle$ 和 $|S_{x-}\rangle$ 。你也许想将新分裂出的两束的状态分别记为

$$|S_z+, S_{x+}\rangle \text{ 和 } |S_z+, S_{x-}\rangle, \quad (1.17)$$

“毕竟， z 方向已经筛选过了啊！”然后你可能想，只需要再进行一次 y 方向的实验（当然，因为我们的原子束本就沿着 y 方向运动，做 y 方向实验需要一些技巧。），三个方向的自旋分量就都能确定，最终得到 $2 \times 2 \times 2 = 8$ 束各自取向都确定的银原子束。但这诱惑是无法兑现的。

让 $|S_{x+}\rangle$ 再次通过 z 方向不均匀磁场，会发生什么？倘若其状态果真是 $|S_z+, S_{x+}\rangle$ ，那就应该还向上偏转，但实际情况是银原子束再次分裂为上下两束，“本该被筛选消除掉的” $|S_z-\rangle$ 又出现了！

原则上，可以将级联 SG 实验做得无限复杂。如让经历了“ z 磁场- x 磁场- z 磁场”共三级得到的 $|S_z-\rangle$ 再通过 x 磁场，束仍会分裂。对 y 方向做实验，结果也是一样的。

一般地，对任意方向的不均匀磁场（不限于 x, y, z ），银原子束通过后总是分裂为两束，除非与上次经过的磁场方向相同。分裂的两束的强度（银原子数量）取决于这次磁场和上次磁场的夹角，当两次经过的磁场相互垂直时，两束强度相等，一比一分裂。

1.3 量子力学基本假设：量子态及其叠加原理

现在我们引入量子力学的一个基本假设来描述⁸ SG 实验：

量子态及其叠加原理

量子系统的状态（量子态）可以线性叠加（superposition），叠加的结果（叠加态）仍然是一个物理上合理的量子态。

总结级联 SG 实验的现象如下：

- $|S_z+\rangle \xrightarrow{z\text{磁场}} \text{仍为 } |S_z+\rangle$;
- $|S_z-\rangle \xrightarrow{z\text{磁场}} \text{仍为 } |S_z-\rangle$;
- $|S_z+\rangle \xrightarrow{x\text{磁场}} \text{等量分裂为 } |S_x+\rangle \text{ 和 } |S_x-\rangle$;
- $|S_z-\rangle \xrightarrow{x\text{磁场}} \text{也等量分裂为 } |S_x+\rangle \text{ 和 } |S_x-\rangle$;
- $|S_z+\rangle \xrightarrow{y\text{磁场}} \text{等量分裂为 } |S_y+\rangle \text{ 和 } |S_y-\rangle$;
- $|S_z-\rangle \xrightarrow{y\text{磁场}} \text{也等量分裂为 } |S_y+\rangle \text{ 和 } |S_y-\rangle$;
- $|S_x+\rangle \xrightarrow{z\text{磁场}} \text{等量分裂为 } |S_z+\rangle \text{ 和 } |S_z-\rangle$;
- $|S_x-\rangle \xrightarrow{z\text{磁场}} \text{也等量分裂为 } |S_z+\rangle \text{ 和 } |S_z-\rangle$;
- 其它组合同理;
- 当前后两级磁场方向不相互垂直，分裂不再等量。

这告诉我们，不可能通过级联 SG 实验不断筛选来确定电子自旋的三个分量，每当一个方向上的分量确定下来，其它方向上的分量就不再确定。 $|S_z+\rangle$ 银原子束分裂不是因为其中有些银原子处在 $|S_z+, S_x+\rangle$ 、有些银原子处在 $|S_z+, S_x-\rangle$ ， $|S_x+\rangle$ 和 $|S_x-\rangle$ 是以一种“潜在”的方式蕴含在 $|S_z+\rangle$ 中的，并非已经确定。

这就是量子力学中的叠加态，可以写为：

$$|S_z+\rangle = |S_x+\rangle + |S_x-\rangle. \quad (1.18)$$

即，电子自旋在 z 方向确定为正的态，同时也是电子自旋在 x 方向完全不确定（表现为会等量分裂）的态。另一方面也有

$$|S_x+\rangle = |S_z+\rangle + |S_z-\rangle. \quad (1.19)$$

⁸所谓“解释”大致有两种，一种是用人们更熟悉的另一种情况来类比，另一种是用“更基本”的语言来重新描述，物理学选择了后者。“更基本”的描述往往能囊括多种不同的情况，使得人们只需要很少的概念就能讨论很多现象。

你可能听说过，像式(1.18)等号右边这样的态，意味着有 50% 的概率得到 $|S_{x+}\rangle$ ，50% 的概率得到 $|S_{x-}\rangle$ ，“这只不过就像硬币有两面，两面的各有概率”，你可能会想。然而式(1.18)两边是等价的，这很关键！我们必须承认，**叠加态是确定的状态**，而不是“不知道什么状态”！等号左边有多确定，等号右边就同等地确定。

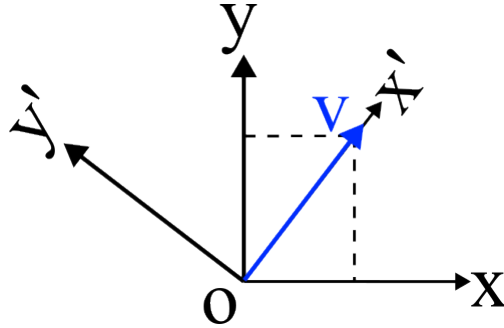


图 1.6. 矢量 \vec{v} 既不是 \vec{e}_x 方向，也不是 \vec{e}_y 方向，它有自己的方向——而不是“不知道什么方向”！在新坐标系 $\{x', y'\}$ 中， \vec{v} 平行于 x' 轴，相应地，平行于旧轴的矢量不平行于新轴。

这还告诉我们，“这个态是叠加态吗？”是一个缺乏意义的问题，当你关心的是 z 方向分量时，式(1.18)就不是叠加态，当你关心的是 x 方向分量时，它就是叠加态。**叠加态是相对而言的**⁹。这就像问一个矢量是否平行于坐标轴，是取决于你怎样建立什么样的坐标系一样。我不厌其烦地多次使用加粗字体，希望这能给你留下深刻的印象。

另一方面，对 $|S_z-\rangle$ 有

$$|S_z-\rangle = |S_{x+}\rangle - |S_{x-}\rangle. \quad (1.20)$$

比起式(1.18)两个叠加系数都是 +1，这里一个是 +1，另一个是 -1，为什么是这样？详细的原因要在正式学到自旋时才知道，在这里我们给出一个启发性的说明。将式(1.20)的尖头翻转，写成

$$\langle S_z-| = \langle S_{x+}| - \langle S_{x-}|. \quad (1.21)$$

然后将它和式(1.18)“相乘”，即

$$\begin{aligned} \langle S_z-|S_z+\rangle &= (\langle S_{x+}| - \langle S_{x-}|)(|S_{x+}\rangle + |S_{x-}\rangle) \\ &= \langle S_{x+}|S_{x+}\rangle + \langle S_{x+}|S_{x-}\rangle - \langle S_{x-}|S_{x+}\rangle - \langle S_{x-}|S_{x-}\rangle, \end{aligned} \quad (1.22)$$

我们约定，形如 $\langle B|A\rangle$ 的符号，表达的是“A 中有多少 B”。我们已经在式(1.18)中见过了“你中有我，我中有你”，那什么情况下才“你中无我”、互不相关呢？我们

⁹值得一提的是，对于最初从炉子中射出的银原子束，无论你探测哪个方向的分量，都会等量分裂。这些银原子不能由一个叠加态来描述，它们处于一种真正杂乱无章的状态。我们后面将会引入描述它们的工具：混合态。

约定，物理上可以完全区分的两个态是互不相关的，满足

$$\langle B|A\rangle = 0, \quad (1.23)$$

读作 B 态与 A 态**正交** (orthogonal)。比如 $|S_z+\rangle$ 和 $|S_z-\rangle$ 代表的两束银原子，最终在接收屏上的痕迹完全分开¹⁰，于是我们说

$$\langle S_z-|S_z+\rangle = 0. \quad (1.24)$$

同理有 $\langle S_x-|S_x+\rangle = 0$, $\langle S_y-|S_y+\rangle = 0$ 。

这样式(1.22)变为

$$0 = \langle S_x+|S_x+\rangle + 0 - 0 - \langle S_x-|S_x-\rangle, \quad (1.25)$$

A 中有多少 A 本身呢？当然是 100%，与 $\langle B|A\rangle = 0$ 相对应，我们最好约定 $\langle A|A\rangle = 1$ 。为此我们需要调整一下之前的叠加系数，即

$$\begin{aligned} |S_z+\rangle &= \frac{1}{\sqrt{2}} |S_x+\rangle + \frac{1}{\sqrt{2}} |S_x-\rangle, \\ |S_z-\rangle &= \frac{1}{\sqrt{2}} |S_x+\rangle - \frac{1}{\sqrt{2}} |S_x-\rangle. \end{aligned} \quad (1.26)$$

于是(1.22)变为

$$0 = 1 - 1 = 0, \quad (1.27)$$

说明我们的约定是**自洽的**。也就是说， $|S_z+\rangle$ 和 $|S_z-\rangle$ 都能等量分裂为 $|S_x+\rangle$ 和 $|S_x-\rangle$ 意味着两个叠加系数的**绝对值相等**，但符号不一样，这一符号差别保证了 $\langle S_z-|S_z+\rangle = 0$ 。而将系数绝对值从 1 改为 $\frac{1}{\sqrt{2}}$ 保证了 $\langle S_z+|S_z+\rangle = 1$ ，这种在保持相对大小不变的情况下，对系数的修改称为**归一化** (normalization)。

那么， $|S_y\pm\rangle$ 呢？正负号的“伎俩”已经用完了，我们该怎么表达 $|S_y\pm\rangle$ 与 $|S_z\pm\rangle$ 和 $|S_x\pm\rangle$ 之间的关系呢？答案是引入复数。

如果我们允许叠加系数是复数，那么可以写出（正式学到自旋后，我们会亲手求出来）

$$\begin{aligned} |S_y+\rangle &= \frac{1}{\sqrt{2}} |S_z+\rangle + \frac{i}{\sqrt{2}} |S_z-\rangle, \\ |S_y-\rangle &= \frac{1}{\sqrt{2}} |S_z+\rangle - \frac{i}{\sqrt{2}} |S_z-\rangle. \end{aligned} \quad (1.28)$$

此时，两个叠加系数就不再是简单的正负之差，而是在复平面上有角度之差，我们称之为**相位差** (phase difference)。

如对

$$|S_x-\rangle = \frac{1}{\sqrt{2}} |S_z+\rangle - \frac{1}{\sqrt{2}} |S_z-\rangle, \quad (1.29)$$

¹⁰严格来说，接收屏上的痕迹能否完全分开，直接体现的是接收屏上用于探测银原子的成分的状态是否正交，对自旋态的正交性的体现是间接的。

由欧拉公式 $e^{i\theta} = \cos \theta + i \sin \theta$ ，可以写为

$$|S_x-\rangle = \frac{e^{i0}}{\sqrt{2}} |S_z+\rangle + \frac{e^{i\pi}}{\sqrt{2}} |S_z-\rangle, \quad (1.30)$$

即右边两个系数相位差为 π ，而

$$|S_y+\rangle = \frac{e^{i0}}{\sqrt{2}} |S_z+\rangle + \frac{e^{i\frac{\pi}{2}}}{\sqrt{2}} |S_z-\rangle \quad (1.31)$$

中两个系数相位差为 $\frac{\pi}{2}$ 。可见，当你把两个态叠加起来，你会得到什么态不仅取决于系数的绝对值（对复数来说，模长），还有系数之间的相位差。片面地用系数的绝对值（和概率有关）来描述一个态就会丢失关键的信息，毕竟， $|S_x-\rangle$ 和 $|S_y+\rangle$ 可是完全不同的！

1.4 量子力学基本假设：投影测量

我们已经通过 SG 实验认识了量子态及其叠加原理——虽然碍于知识准备不足，一些细节没有被论述。但还有一点对接下来的讨论至关重要，让我们再度考虑级联 SG 实验。

如对已经筛选出的 $|S_z+\rangle$ ，我们让它通过 x 方向不均匀磁场，如果加入接收屏筛掉 $|S_x\pm\rangle$ 其中之一，那么留下的银原子再次进入 z 磁场时就会分裂。如果我们在 x 磁场的出口紧接着再设置一个 x 方向反向的不均匀磁场，随后让银原子进入下一级 z 磁场，会发生什么？

按照经典力学的直觉，上述设置会将两束银原子合并为一束——虽然空间上沿着同一个方向运动，但是自旋取向却不同，一半是 $|S_x+\rangle$ ，另一半是 $|S_x-\rangle$ 。再进入 z 磁场，这两半都会等量分裂，就结果而言将在最终的接收屏上留下两道痕迹，就像最原始的 SG 实验一样。

然而出乎意料的事发生了：在最后的接收屏上，只有一道向上偏转的痕迹，银原子束没有分裂，就好像输入的 $|S_z+\rangle$ 跳过了中间的阶段一样。

这是怎么回事？用上一节引入的叠加态可以清楚地描述这个过程。入态为

$$|S_z+\rangle = \frac{1}{\sqrt{2}}(|S_x+\rangle + |S_x-\rangle), \quad (1.32)$$

进入 x 磁场后的态是什么呢？对于电子的自旋态而言，无论哪个方向，只有两种相互正交的可能，这称为完备性（completeness）。即，无论何时，电子的自旋态都能用 $|S_x+\rangle$ 和 $|S_x-\rangle$ 叠加出来，只不过两个系数不同。也就是说进入 x 磁场后，态变为

$$c_1(t) |S_x+\rangle + c_2(t) |S_x-\rangle, \quad (1.33)$$

其中 $c_1(t)$ 和 $c_2(t)$ 是两个复数，满足归一化 $|c_1(t)|^2 + |c_2(t)|^2 = 1$ 。以进入 x 磁场为 $t = 0$ 时刻，则

$$c_1(0) = c_2(0) = \frac{1}{\sqrt{2}}. \quad (1.34)$$

原则上，银原子离开第一级 x 磁场，进入第二级 x 反向磁场时的系数可以用量子力学计算出来。但这里我们只需要借助**电磁相互作用下的运动是可逆的**¹¹这一点，可知在银原子离开第二级磁场时，两个叠加系数又都变回 $\frac{1}{\sqrt{2}}$ ，也就是说态又变回了 $|S_z+\rangle$ ！这样，再进入 z 磁场时当然就只会有一束。

关键在于， $|S_x+\rangle$ 和 $|S_x-\rangle$ 必须始终保持叠加，或者说保持**相干性**（coherence，这里量子力学借用了光学的术语），即银原子束在 x 磁场中处于叠加态，是“一束”（尽管反直觉）而不是两束，这样最后才可能恢复为 $|S_z+\rangle$ 。

而如果在 x 磁场中，态“中间的加号被切断了”，银原子确实实地变成了两束，那就无法恢复为 $|S_z+\rangle$ ！我们称叠加被“切断”的现象为**退相干**（decoherence），如果能把“切断”后的态记录下来，我们就说发生了**测量**。

在 SG 实验中，重点是接收屏的存在，而不是“你”是否“看”实验装置。当然，为了见证实验结果，我们最终必须去看实验装置，但如果没有合适的接收屏，看也没有任何影响。怎样算合适呢？关键在于接收屏上的化学成分可以与银原子**局域地相互作用**以在屏上某处生成痕迹，因此接收屏具有**测量银原子位置的能力**。而通过不均匀磁场，自旋与银原子束的偏转、进而与接收屏上的位置**形成了关联**，这关联将电子自旋的信息**泄露了出去**，构成了一个合格的测量过程。我们今后会看到，这关联不是别的，正是**量子纠缠**（quantum entanglement）。由此，测量过程可以概括为“相互作用——建立纠缠——力学量的信息经关联泄露”。

测量是相互作用吗？只能说，相互作用是测量的必要条件。测量中涉及的必须是与待测的力学量有关的相互作用，而不是随便什么相互作用。如 SG 实验中，自旋与银原子束的偏转关联起来，所以我们要通过测量位置来间接测量自旋，这就要求使用接收屏——一个能和银原子局域地相互作用的装置。这里局域性是至关重要的，你必须预估银原子束的前进方向来放置接收屏，如果你将接收屏放置在八竿子打不着的地方，那么就无法测量，银原子束将保持相干性¹²。

还有一个问题是，根据量子力学计算，接收屏与银原子的相互作用将把接收屏这个仪器也“带入叠加态”，可以写成

$$|S_x+\rangle |\text{接收屏左侧发生化学反应}\rangle + |S_x-\rangle |\text{接收屏右侧发生化学反应}\rangle, \quad (1.35)$$

¹¹这一点对于经典力学和量子力学都适用。

¹²当然，你可能听说过，波动性让银原子有微小的可能出现在远离银原子束的地方，进而你可以期待摆错地方的接收屏仍能进行测量——但这可能性比中头奖还要远远低得多。只要不把银原子束的集中程度限制得太厉害，对于宏观的实验者来说就可以认为银原子束是在像经典力学的质点一样运动，从而可以预估。

即，相干性可以通过相互作用扩散到更大的系统中。但在实验室中，我们可以看到明确的结果态，这一经验事实与量子力学的结果发生了不协调。对这个问题，就像经典力学中的各种理想模型一样，量子力学的先驱们为测量也提出了一个理想模型：

测量投影假设

量子系统处于态 $|\psi\rangle$ ，则测量力学量 A 时，量子态会**随机地投影**到该力学量具有确定取值的某一个态上。设 A 有 N 种可能的取值 $\{a_n|n = 1, \dots, N\}$ ，则 $|\psi\rangle$ 测量投影到 $|a_n\rangle$ 的概率为 $|\langle a_n|\psi\rangle|^2$ 。

这一假设称为**投影假设**，其中概率的计算规则由马克斯·玻恩（Max Born）提出，称为**玻恩概率诠释**或**玻恩定则**¹³。

在许多实验，尤其是量子力学历史早期的实验中，人们不太关心（也缺乏技术手段去关心）测量的最终阶段发生了什么。接受这一假设意味着将测量当成一个黑箱——但是是一个行为可预期的黑箱，人们可以通过玻恩定则计算力学量在量子态上的**期望值**，即用概率加权的平均值：

$$\langle A \rangle = \sum_{n=1}^N |\langle a_n|\psi\rangle|^2 a_n. \quad (1.36)$$

有了期望值，人们就可以比较理论和实验了。包括投影假设在内的“标准量子力学”取得了辉煌的成就，成为了人类科学史上最精确的理论之一。

当然，有许多人不满足于此，对测量过程的研究称为**测量问题**。自 20 世纪 50 年代起，得益于量子光学与量子信息的理论和技术发展，许多过去只能模糊思辨的问题都能得到清楚的描述和确切的检验了。尽管测量问题的最终解决还未到来，但**连续量子测量**（continuous quantum measurement）、**量子轨线理论**（quantum trajectory theory）等新观念的诞生，大大深化了人们对测量问题的理解。弱测量，乃至“零作用测量”（interaction free measurement）等的实验实现，向人们揭示了一个远比先驱者们预想的更加广阔而充满潜力的量子自然。

¹³历史上，玻恩最初给出的是对位置测量的诠释，这里的推广到所有力学量的诠释也称为**广义统计诠释**（generalized statistical interpretation）。

练习

练习 1

判断下列说法的正误：

1. 接收屏上有两条痕迹只是因为磁性有 N, S 两级。
2. $|S_z+\rangle$ 表示自旋像箭头一样指向 z 轴正方向。
3. 可以通过级联 SG 实验确定自旋的状态。
4. 可以通过级联 SG 实验确定自旋的取向。

练习 2

回答下列问题：

1. 两个态的叠加是唯一的吗？与什么有关？
2. 利用正文中给出的 $|S_z\pm\rangle$ 与 $|S_x\pm\rangle, |S_y\pm\rangle$ 的关系，写出 $|S_x\pm\rangle$ 与 $|S_y\pm\rangle$ 的关系。
3. 自己设计一个级联 SG 实验，分析其过程和结果。

第二章 量子力学的理论结构

张越之

在进一步讨论其它量子现象前，让我们用量子力学的语言将自己武装起来，以使我们的讨论扎实可靠、不致把自己搞糊涂。相信我，掌握这一章的内容将使你受益无穷。

2.1 右矢

保罗·狄拉克 (Paul Dirac) 洞见到，无论是沃纳·海森堡 (Werner Heisenberg) 的矩阵力学，还是埃尔文·薛定谔 (Erwin Schrödinger) 的波动力学，态叠加原理允许我们统一用**矢量描述量子态**，称作**态矢量** (state vector)。

不同于我们熟悉的实数域上的矢量 (欧氏矢量)，态矢量是复数域上的矢量。而且，我们在经典力学中熟悉的都是有限维的矢量，而态矢量常常是**无穷维**的。上一章讨论的自旋态倒是只有二维，但因为系数可以是复数，所以需要不止两个实数来确定一个自旋态矢量¹。

狄拉克为此建立了一套优雅的符号，称为**狄拉克符号**或 bra-ket 符号 (取自英文单词 bracket, 即括号)。在这套符号下，一个量子态可由 $|\text{态的名字}\rangle$ 来表达，即用一条竖线和一个拉伸了的大于号将你给这个量子态取的名字包起来，如

$$|\psi\rangle, \quad |\text{第四激发态}\rangle, \quad |\text{真空态}\rangle, \quad |\uparrow\rangle, \quad |\text{随便一个名字}\rangle, \quad (2.1)$$

等等，这被称为**右矢 (ket)**，你可以将其看成像 \vec{v} 一样的东西。它允许我们较为“轻松”地表述物理，如著名的薛定谔的猫态：

¹两个复数系数意味着四个实数，加上归一化条件变成三个实数。实际上，由于接下来要讲的理由，只需要两个实数。

$$c_1 \left| \text{猫} \right\rangle + c_2 \left| \text{死猫} \right\rangle$$

图 2.1. 薛定谔的猫态示意。正如上一章提到的，叠加系数可以为任意满足归一化的复数，为了简单常常都设为 $\frac{1}{\sqrt{2}}$ 。实际上，它应该像(1.35)一样，此处为简化。

在上一章中我们已经“物理地”引入了这种符号，现在中我们将考察它作为矢量的数学性质。

在熟悉了狄拉克符号后，你会逐渐感到量子力学在数学上并不可畏，最终可以“像呼吸一样自然”地用狄拉克符号讨论量子力学问题，甚至带着它回头学习线性代数。

矢量性意味着两个右矢可以做加法，结果还是一个右矢：

$$|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle = |\gamma\rangle. \quad (2.2)$$

还可以做数乘，也得到一个右矢：

$$c|\alpha\rangle = |\delta\rangle, \quad (2.3)$$

其中 c 是任意一个复数，它写在 $|\alpha\rangle$ 的左边或右边没有区别。如果 $c = 0$ ，那结果就将是一个**零矢量**，可以记为 0 。零矢量满足

$$|\alpha\rangle + 0 = |\alpha\rangle, \quad \forall |\alpha\rangle, \quad (2.4)$$

它没有“方向”。在不致引起混淆的情况下，也可与标量零不加区分地都记为 0 。

和我们熟悉的欧氏矢量一样，右矢还满足加法结合律、数乘的结合律和分配律等：

$$\begin{aligned} |\alpha\rangle + |\beta\rangle + |\gamma\rangle &= |\alpha\rangle + (|\beta\rangle + |\gamma\rangle), \\ (c_1 + c_2)|\psi\rangle &= c_1|\psi\rangle + c_2|\psi\rangle, \\ c(|\alpha\rangle + |\beta\rangle) &= c|\alpha\rangle + c|\beta\rangle. \end{aligned} \quad (2.5)$$

任意一个集合，只要满足围绕加法、数乘的上述性质就称为**线性空间**，其中的元素就称为**矢量**。右矢所在的空间称为**希尔伯特（Hilbert）空间**，符号为 H ，为了不后面将引入的哈密顿量 H 混淆，也常改变字体写为 \mathcal{H} 等。

2.2 左矢与内积

除了加法和数乘两种基本运算，我们还可以进一步定义更高级的运算。在对 SG 实验的讨论中，我们引入了一种表示“你中有多少我”的符号：

$$\langle B|A \rangle. \quad (2.6)$$

当时我们说“尖头翻转，再‘相乘’”，并说明了作为一种符号游戏，它是自洽的，拥有正交性和可归一化两种性质——常合并称为**正交归一性**（orthonormality）。现在我们知道了右矢是矢量，那这种运算在数学上是什么呢？

对欧氏空间矢量，我们有点乘运算，如对两个三维矢量：

$$\begin{aligned} \vec{v} &= v_x \vec{e}_x + v_y \vec{e}_y + v_z \vec{e}_z, \\ \vec{u} &= u_x \vec{e}_x + u_y \vec{e}_y + u_z \vec{e}_z, \end{aligned} \quad (2.7)$$

它们的点乘为

$$\vec{v} \cdot \vec{u} = v_x u_x + v_y u_y + v_z u_z. \quad (2.8)$$

设 \vec{v} 和 \vec{u} 的夹角为 θ ，则有

$$|\vec{v} \cdot \vec{u}| = |\vec{v}| |\vec{u}| \cos \theta, \quad (2.9)$$

可见，点乘就是一种“你中有多少我”的运算，一个矢量除非和另一个矢量垂直，否则就能分出和另一个矢量平行的分量。在推广到更一般的对象时，我们将“点乘”改为“内积”，将“垂直”改为“正交”。 $\langle B|A \rangle$ 就称为**内积运算**。

为了给态矢量定义内积，我们需要引入**左矢 (bra)**。左矢是一种**对偶矢量 (dual vector)**，对偶矢量是输入矢量、输出标量的映射。在量子力学中，对每一个右矢 $|\alpha\rangle$ ，我们都引入一个与之**对偶**的左矢 $\langle\alpha|$ ²：

$$|\alpha\rangle \leftrightarrow \langle\alpha|. \quad (2.10)$$

全体左矢同样组成一个复线性空间，也有加法和数乘。

于是就可以定义内积了，要求 $|B\rangle$ 和 $|A\rangle$ 的内积，只需要把 $|B\rangle$ 对应的左矢 $\langle B|$ 和 $|A\rangle$ 放在一起，运算结果为一个标量（复数）：

$$\langle B|A \rangle = c \in \mathbb{C}. \quad (2.11)$$

²实际上，我们需要给态空间引入度规 g ，它是输入两个矢量，输出一个标量的映射： $g(|\psi\rangle, |\phi\rangle) = c$ 。如果它是正定的： $g(|\psi\rangle, |\psi\rangle) \geq 0$ ，那就可以称为内积。如果我们只输入一个矢量，将另一个输入留空，那么 $g(|\psi\rangle, \cdot)$ 就是一个只需要再输入一个矢量，就能输出标量的映射，这正是对偶矢量。一般而言，我们只能说有一个矢量空间，还有一个对偶矢量空间，而不能说一个矢量的对偶是什么。但有了度规，我们就可以让矢量和对偶矢量有一个自然的一一对应——两个空间有了**自然同构**，即 $\langle\psi| = g(|\psi\rangle, \cdot)$ 。

还有一条重要的要求，叠加态右矢的对偶**不是**相应左矢的直接叠加，而是系数取复共轭后的叠加：

$$c_1 |\alpha\rangle + c_2 |\beta\rangle \leftrightarrow c_1^* \langle\alpha| + c_2^* \langle\beta|, \quad (2.12)$$

这称为**共轭线性**。

让我们用一个特例看看共轭线性的后果，考虑

$$\begin{aligned} |\psi\rangle &= c_1 |\alpha\rangle + c_2 |\beta\rangle, \\ |\phi\rangle &= c_3 |\alpha\rangle + c_4 |\beta\rangle, \end{aligned} \quad (2.13)$$

其中 $\langle\alpha|\beta\rangle = 0$ ，且都已归一化。这就像是两个平面矢量，但是系数是复的。则有

$$\begin{aligned} \langle\phi|\psi\rangle &= c_3^* c_1 + c_4^* c_2, \\ \langle\psi|\phi\rangle &= c_1^* c_3 + c_2^* c_4, \end{aligned} \quad (2.14)$$

可见，为了自洽应有

$$\langle B|A\rangle = \langle A|B\rangle^*, \quad (2.15)$$

即，态矢量的内积交换顺序，结果差一个复共轭，这与我们熟悉的实数域上的欧氏矢量不同。

为什么？我们希望，一个矢量和自身的内积能表示其大小，如

$$|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}, \quad (2.16)$$

由于只有实数可以比大小，自身内积的结果应该是实数。考虑：

$$|\psi\rangle = c_1 |A\rangle + c_2 |B\rangle, \quad (2.17)$$

其中， $|A\rangle, |B\rangle$ 归一但不一定正交。由共轭线性性，相应的左矢应为

$$\langle\psi| = c_1^* \langle A| + c_2^* \langle B|, \quad (2.18)$$

做内积有：

$$\begin{aligned} & (c_1^* \langle A| + c_2^* \langle B|)(c_1 |A\rangle + c_2 |B\rangle) \\ &= |c_1|^2 \langle A|A\rangle + c_1^* c_2 \langle A|B\rangle + c_2^* c_1 \langle B|A\rangle + |c_2|^2 \langle B|B\rangle \\ &= |c_1|^2 + |c_2|^2 + c_1^* c_2 \langle A|B\rangle + (c_1^* c_2 \langle A|B\rangle)^* \\ &= |c_1|^2 + |c_2|^2 + 2 \operatorname{Re}(c_1^* c_2 \langle A|B\rangle), \end{aligned} \quad (2.19)$$

其中 Re 表示取实部。可见，上述要求保证了 $\langle\psi|\psi\rangle$ 是一个实数。

实际上，自身内积不仅要是实数，还应当是一个非负实数，当且仅当 $|\psi\rangle$ 为零矢量时为零。只有这样，我们的符号才能容纳玻恩的概率诠释——自身内积即总概率，应当非负才能放缩到 $[0, 1]$ 区间中。

这一节总结起来就是：

- $c|\psi\rangle \leftrightarrow \langle\psi|c^*$;
- $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$;
- $\langle\psi|\psi\rangle \geq 0$ 。

你大可以先将这三条作为游戏规则接受下来，上手应用，日后再仔细体会。

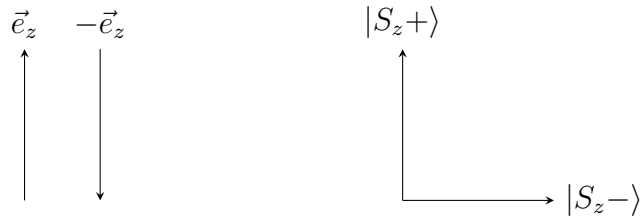


图 2.2. $z+$ 和 $z-$ 两个方向在“实际空间”中互相平行，而在态空间中，对应的态相互正交。

回看 SG 实验，当时我们说

$$\langle S_z- | S_z+ \rangle = 0, \quad (2.20)$$

可见，虽然对我们行动于其中的欧氏空间， $z+$ 和 $z-$ 是两个相反的方向，指向这两个方向的矢量平行/线性相关，但相应的态空间中的态 $|S_z+\rangle$ 和 $|S_z-\rangle$ 却是正交（垂直）的。

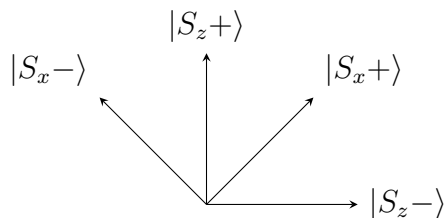


图 2.3. 由式(1.20)画出的 $|S_z\pm\rangle$ 与 $|S_x\pm\rangle$ 关系示意图，由于复数的存在，我们无法这样画上 $|S_y\pm\rangle$ 。

至于常见的另一种矢量运算，叉乘，其只在三维空间中有良好定义，而态矢量的空间往往并非三维（对自旋态是二维，而我们接下来会经常见到无穷维！），因此对右矢/左矢并没有叉乘运算。今后为了区分起见，我们将对态矢量的运算专门称为内积，而对欧氏矢量的运算专门称为点乘。

2.3 基矢量

对欧氏平面矢量，可以选择两个正交归一的基矢 \vec{e}_x, \vec{e}_y ，从而任意矢量都可以由它们展开。我们现在可以给态空间也选择一组正交归一的基矢，而彼此独立的基矢的数目就是态空间的维度。

一个态空间有多少维？这取决于所考虑系统的物理特性，至少有某个物理量可能的不同取值那么多。如在 SG 实验中，电子自旋 z 轴投影 S_z 有两种可能的取值，相应的自旋态空间是二维的。而之后会学到的描述位置和动量所要用的态空间则是无穷维的——有无穷多不同的取值。

但很多时候，某个物理量可以一个取值对应多个不同的态，我们称这个取值存在**简并**（degenerate）。存在简并时态空间的维数就比物理量可能的取值更多。在经典力学中就存在这样的情况，如对做一维运动的质点，动能的一个取值可以对应向两个相反方向运动的状态，这叫**二重简并**或者说**简并度**为二。

态空间的**维数还可能变化**，如当粒子产生、湮灭时。

总之，任意一个态总可以用一组正交归一基矢展开：

$$|\psi\rangle = \sum_{n=1}^N \psi_n |e_n\rangle, \quad (2.21)$$

其中 N 可以为 ∞ 。基矢的正交归一性可以表达为

$$\langle e_m | e_n \rangle = \delta_{mn}, \quad m, n = 1, \dots, N, \quad (2.22)$$

其中 δ_{mn} 称为**克罗内克符号**或 delta 符号，当 $m = n$ 时为 1， $m \neq n$ 时为 0，即基矢自身内积为 1，不同基矢内积为 0。当 δ_{mn} 的角标被求和时，可以将式子中的一个指标变成另一个指标。举个例子：

$$\sum_{i=1}^3 \delta_{i1} = \delta_{11} + \delta_{21} + \delta_{31} = \delta_{11} = 1. \quad (2.23)$$

我们将第 n 个基矢前的展开系数记作 ψ_n 以提醒我们它们是从 $|\psi\rangle$ 而来的，它们可以用内积求得：

$$\psi_n = \langle e_n | \psi \rangle. \quad (2.24)$$

$\{|e_n\rangle | n = 1, \dots, N\}$ 并不比 $\{\vec{e}_n | n = x, y, z\}$ 更抽象，尝试接受它！但也许你想要让事情更具体些，如

$$\vec{e}_x \stackrel{?}{=} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad (2.25)$$

这里能画等号么？实际上，右边表达的是

$$\vec{e}_x = 1 \cdot \vec{e}_x + 0 \cdot \vec{e}_y + 0 \cdot \vec{e}_z, \quad (2.26)$$

即 \vec{e}_x 这个矢量在 $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ 这组基矢上的展开系数分别是 1, 0, 0。也就是说，对于一个抽象的矢量，我们可以通过它在一组基矢上的展开系数来让它看起来更具

体。由于其具体形式是取决于你选择的基矢的，所以两边不能直接画等号，让我们换一个符号：

$$\vec{e}_x \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad (2.27)$$

在等号上面加一点，读作 \vec{e}_x 在 $\{\vec{e}_n | n = x, y, z\}$ 下表示为 $(1, 0, 0)^T$ （上标 T 表示转置）。于是，对式(2.21)，我们可以写：

$$|\psi\rangle \doteq \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \quad (2.28)$$

由于基矢是正交归一的，为了 $|\psi\rangle$ 归一，系数的模方之和为 1。

粗看起来， N 个基矢意味着要 N 个复数系数，即 $2N$ 个实数来表达一个态。将复数用模长-相位的形式表达为 $\psi_n = |\psi_n|e^{i\theta_n}$ ，即需要 N 个模长， N 个相位。而只有 $N - 1$ 个模长是独立的，剩下一个可以用归一化条件

$$\sum_{n=1}^N |\psi_n|^2 \quad (2.29)$$

求出。于是我们只需要 $2N - 1$ 个实数： $N - 1$ 个模长， N 个相位。

实际上，我们还可以进一步限制。回想平面矢量的情况，决定一组平面矢量的几何关系的是它们的模长和**夹角**——在坐标系中的绝对角度是无关紧要的！将一组平面矢量整体旋转，绝对角度改变了，但内禀的几何性质不会改变。

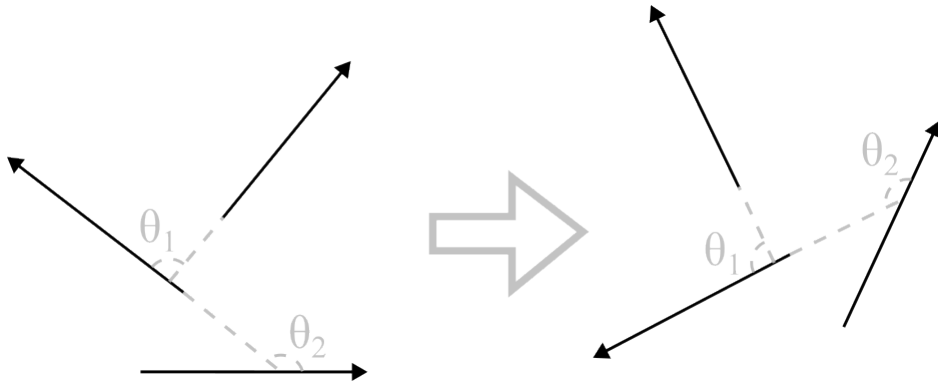


图 2.4. 对一组矢量的几何性质，重要的是相对角度。

对态矢量也是一样，重要的是**相对相位**或者说相位差，我们可以将所有的态都改变一个**整体相位**，即乘一个单位复数 $e^{i\delta}$ ，而不会改变物理现象。也就是说，我

们可以对 ψ_n 提出一个公共相位因子：

$$\begin{aligned}
 (\psi_1, \psi_2, \dots, \psi_N)^T &= (|\psi_1|e^{i\theta_1}, |\psi_2|e^{i\theta_2}, \dots, |\psi_N|e^{i\theta_N})^T \\
 &= e^{i\theta_1} (|\psi_1|, |\psi_2|e^{i(\theta_2-\theta_1)}, \dots, |\psi_N|e^{i(\theta_N-\theta_1)})^T \\
 &\xrightarrow{\text{消去整体相位}} (|\psi_1|, |\psi_2|e^{i(\theta_2-\theta_1)}, \dots, |\psi_N|e^{i(\theta_N-\theta_1)})^T,
 \end{aligned} \tag{2.30}$$

这样，我们总可以让一个系数直接是实数（这里选了 $\psi_1 \rightarrow |\psi_1|$ ），其它系数是复数，即只需要总共 $2N - 2$ 个实数就能描述一个 N 维态矢量。为了清楚起见，对整体相位只在初始时修改一次，如果态在后续的变化中又出现了整体相位，则我们加以保留，以对比其和初始状态之间的相位差。

当然，右矢有了基矢，左矢也有相应的基矢。既然右矢表示为列向量（列矩阵），那我们就应该把左矢表达为行向量（行矩阵），于是抽象的内积运算变成了行向量和列向量间的矩阵乘法。比如

$$\langle \psi | \doteq (\psi_1^*, \psi_2^*, \dots, \psi_N^*), \tag{2.31}$$

别忘了共轭线性！于是有

$$\begin{aligned}
 \langle \psi | \psi \rangle &= \left(\sum_{m=1}^N \psi_m^* \langle e_m | \right) \left(\sum_{n=1}^N \psi_n | e_n \rangle \right) \\
 &= \sum_{m=1, n=1}^N \psi_m^* \psi_n \langle e_m | e_n \rangle \\
 &= \sum_{m=1, n=1}^N \psi_m^* \psi_n \delta_{mn} \\
 &= \sum_{n=1}^N |\psi_n|^2 \\
 &= \begin{pmatrix} \psi_1^* & \psi_2^* & \dots & \psi_N^* \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} \\
 &= 1.
 \end{aligned} \tag{2.32}$$

注意，第一行中两个展开式的求和指标必须不同，一个是 m ，一个是 n 。想一想，如果指标相同意味着什么？如果你想不清楚，就写一个简单的情况（如二维）看一看！

练习

练习 3

考虑欧氏二维矢量空间，设一组基矢为 $\{\vec{e}_x, \vec{e}_y\}$ ，其正交归一，另一组基矢为 $\{\vec{e}_{x'} = \vec{e}_x - \vec{e}_y, \vec{e}_{y'} = \vec{e}_x + \vec{e}_y\}$ 。

1. 带撇基矢是否正交归一？
2. $\vec{A} = 2\vec{e}_x + 3\vec{e}_y$ ，求 \vec{A} 的在两组基矢下展开系数。
3. $\vec{B} = -7\vec{e}_x + 5\vec{e}_y$ ， $\vec{C} = \vec{A} - \vec{B}$ ，求 \vec{C} 的在两组基矢下展开系数。

练习 4

考虑二维态空间，正交归一基矢为 $\{|e_1\rangle, |e_2\rangle\}$ 。

1. 用 $\{|e_1\rangle, |e_2\rangle\}$ 表达出另一组正交归一基矢。
2. $|\psi\rangle = |e_1\rangle + 2i|e_2\rangle$ ，要将其归一化应乘以什么数？
3. $|\psi\rangle$ 对应的左矢是 $\langle e_1| + 2i\langle e_2|$ 么？
4. 提出一个整体相位，使得 $|\phi\rangle = (a + bi)|e_1\rangle + ce^{id}|e_2\rangle$ 的第一个系数为实数。此时刚才的 $|\psi\rangle$ 会变成什么？
5. 计算 $\langle\phi|\psi\rangle$ 。

2.4 算符

对着一个固定不变的态矢量能有多少意思呢？为了有趣的事发生，我们需要做变换，从一个态矢量得到另一个态矢量。如已知初始时刻的态矢量，想要求后续某时刻的态矢量。或者改变视角，态也会改变——如旋转坐标轴来改变自旋态。

一般说来，我们可以在态空间上指定态与态之间的映射规则，这就是**算符** (operator)。用狄拉克符号写为

$$X|\alpha\rangle = |\beta\rangle. \quad (2.33)$$

算符像一台机器，投入一个右矢作为原料，就产出一个右矢作为产品。映射规则可以是非常任意的，而我们尤为关注其中那些能描述物理的规则。

有两个简单的算符例子。如果一个算符无论输入什么，都原样输出，则称为**恒等算符**或**单位算符**，记作 I （在不致引起混淆的情况下也可记为 1）：

$$I|\alpha\rangle = |\alpha\rangle, \quad \forall |\alpha\rangle. \quad (2.34)$$

而如果无论输入什么，都输出零矢量，则称为**零算符**：

$$X|\alpha\rangle = 0, \quad \forall |\alpha\rangle, \quad (2.35)$$

零算符本身不常见，没有赋予专门的符号。

如果两个算符相减为零算符，那就说这两个算符**相等**——是的，算符之间可以运算。两个算符相加还是一个算符，定义 $Z = X + Y$ 对右矢的作用为

$$Z|\alpha\rangle := X|\alpha\rangle + Y|\alpha\rangle. \quad (2.36)$$

等号右边无非是右矢的加法，满足交换律和结合律，可知算符相加也满足交换律和结合律。

算符的作用有线性性，对叠加态 $c_1|\alpha\rangle + c_2|\beta\rangle$ 的作用等于对其成分先作用再叠加：

$$X(c_1|\alpha\rangle + c_2|\beta\rangle) = c_1(X|\alpha\rangle) + c_2(X|\beta\rangle), \quad (2.37)$$

物理上看，这是态叠加原理的要求。

就像函数可以复合一样，算符也可以复合，毕竟，算符作用右矢后还是一个右矢，当然可以再次被算符作用。由此我们可以定义算符的乘法，定义 $Z = XY$ 对态的作用为

$$Z|\alpha\rangle := X(Y|\alpha\rangle), \quad (2.38)$$

算符的乘法满足结合律

$$XYZ = (XY)Z, \quad (2.39)$$

但不满足交换律

$$XY \neq YX, \quad (2.40)$$

假如 $XY = YX$ ，我们就说 X 和 Y **对易** (commute)。这和复合函数一样，如 $f(x) = x^2$ 和 $g(x) = \sin x$ 之间复合

$$f(g(x)) = (\sin x)^2, \quad g(f(x)) = \sin(x^2), \quad f(g(x)) \neq g(f(x)), \quad (2.41)$$

而对 $f(x) = x^2$ 和 $g(x) = x^3$ ， $f(g(x)) = g(f(x)) = x^6$ 。

如果两个算符的复合为单位算符，则称二者互为**逆算符**，符号为算符的“ -1 次方”：

$$AB = I \Rightarrow B = A^{-1}. \quad (2.42)$$

当然，算符也可以输入左矢、输出左矢³，算符作用左矢时将算符写在左矢的右边：

$$\langle\delta|X = \langle\gamma|. \quad (2.43)$$

³你还可以设想具有更复杂的输入-输出关系的映射，数学上即**张量** (tensor)。我们讨论的算符是 $(1,1)$ 型张量。

$X|\alpha\rangle$ 整体作为一个右矢，它的名字可以自然地叫做“ $X\alpha$ ”。与 $|X\alpha\rangle$ 对偶的左矢是 $\langle X\alpha|$ ，它与 $\langle\alpha|$ 有什么关系？我们将能把 $\langle\alpha|$ 变成 $\langle X\alpha|$ 的算符记作 X^\dagger ：

$$\langle\alpha|X^\dagger = \langle X\alpha| \leftrightarrow |X\alpha\rangle, \quad \forall |\alpha\rangle \leftrightarrow \langle\alpha|, \quad (2.44)$$

或者说

$$X|\alpha\rangle \leftrightarrow \langle\alpha|X^\dagger. \quad (2.45)$$

符号 \dagger 读作 dagger。数学上称 X^\dagger 为 X 的厄米共轭，但你不必太在乎这个名字！一般而言 $\langle\alpha|X^\dagger \neq \langle\alpha|X$ ，即 $X \neq X^\dagger$ ，也就是说算符为右矢和左矢指定的映射关系通常不“对称”。数乘作为特殊的算符给出了一个具体的例子：

$$c|\alpha\rangle \leftrightarrow \langle\alpha|c^*. \quad (2.46)$$

如果 $X = X^\dagger$ ，则称 X 为**厄米算符** (Hermitian operator)，后面我们会看到，厄米算符对于量子力学至关重要。

2.5 算符的矩阵表示

在给定了一组基矢 $\{|n\rangle | n = 1, \dots, N\}$ 之后，右矢可以表示为列向量，左矢可以表示为行向量，这让事情看起来更加具体。那么算符能不能表达得具体些呢？让我们先考虑看似无关的另一个问题。

考虑将式(2.24)代入式(2.21)，有

$$|\psi\rangle = \sum_{n=1}^N \langle e_n|\psi\rangle |e_n\rangle, \quad (2.47)$$

既然 $\langle e_n|\psi\rangle$ 是个数，我们可以把它移动一下，将上式写成更具启发性的形式：

$$|\psi\rangle = \sum_{n=1}^N |e_n\rangle \langle e_n|\psi\rangle, \quad (2.48)$$

我们可以将内积分开，像普通的乘法一样，将与求和指标 n 无关的 $|\psi\rangle$ 作为公因式提到求和号外面：

$$|\psi\rangle = \left(\sum_{n=1}^N |e_n\rangle \langle e_n| \right) |\psi\rangle, \quad (2.49)$$

我们可以先计算括号里的求和，再和括号外面的 $|\psi\rangle$ 。但是等号左边不就是 $|\psi\rangle$ 吗？我们并没有说 $|\psi\rangle$ 是某个特殊的态，可见，括号内的求和可以将任意的态 $|\psi\rangle$ 变换为它自己，这正是单位算符：

$$\sum_{n=1}^N |e_n\rangle \langle e_n| = I. \quad (2.50)$$

这就是说，右矢和左矢尖头对尖头放在一起，就可以形成算符——我们将看到，所有的算符都可以表达成这种形式，与内积相对，这种运算可以称为**外积**。像式(2.50)这

样，正交归一基矢外积再求和得到单位算符，称为**完备性**。完备性关系是任意态都可以用基矢展开这一点的另一种表达方式，这正是我们在 SG 实验中提到的所有自旋态都能由 $|S_z \pm\rangle$ 叠加出来的那个完备性。

正如我们刚才操作中隐含的，内积和外积之间有结合律。同样地，内外积和算符的作用之间也有结合律。如

$$|\psi\rangle\langle\phi| \Rightarrow \begin{cases} \text{左乘 } \langle\chi| \rightarrow \langle\chi|\psi\rangle\langle\phi| \text{ 得到左矢,} \\ \text{右乘 } |\xi\rangle \rightarrow \langle\phi|\xi\rangle|\psi\rangle \text{ 得到右矢.} \end{cases} \quad (2.51)$$

狄拉克符号让这些运算在符号上非常直观，体现了其的强大表达能力。

有了完备性关系，我们就可以“变魔术”了。显然，任意算符 X 和单位算符复合后仍为 X ，我们左复合一个，右复合一个：

$$X = IXI, \quad (2.52)$$

代入完备性关系，有（注意求和指标要不同）：

$$\begin{aligned} X &= \left(\sum_{m=1}^N |e_m\rangle\langle e_m| \right) X \left(\sum_{n=1}^N |e_n\rangle\langle e_n| \right) \\ &= \sum_{m=1, n=1}^N |e_m\rangle\langle e_m| X |e_n\rangle\langle e_n|, \end{aligned} \quad (2.53)$$

运用结合律，我们发现 $\langle e_m|X|e_n\rangle$ 是一个数，它称为算符 X 在 $\{|e_n\rangle\}$ 这组基下的**矩阵元**，记为 X_{mn} ，将它移动一下有

$$X = \sum_{m=1, n=1}^N X_{mn} |e_m\rangle\langle e_n|. \quad (2.54)$$

这样，我们发现算符也可以用基矢展开，将展开系数按两个指标依次排列，就得到了**算符的矩阵表示**：

$$X \doteq \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1N} \\ X_{21} & X_{22} & \cdots & X_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ X_{N1} & X_{N2} & \cdots & X_{NN} \end{pmatrix}. \quad (2.55)$$

具体写一下就明白了，考虑任意算符 X 与任意态 $|\psi\rangle$ ，将 $X|\psi\rangle$ 展开：

$$\begin{aligned} X|\psi\rangle &= \sum_{m=1, n=1}^N X_{mn} |e_m\rangle\langle e_n| \sum_{l=1}^N \psi_l |e_l\rangle \\ &= \sum_{m=1, n=1, l=1}^N X_{mn} \psi_l |e_m\rangle \langle e_n|e_l\rangle \\ &= \sum_{m=1, n=1, l=1}^N X_{mn} \psi_l \delta_{nl} |e_m\rangle \\ &= \sum_{m=1}^N \left(\sum_{n=1}^N X_{mn} \psi_n \right) |e_m\rangle, \end{aligned} \quad (2.56)$$

即， $X|\psi\rangle$ 是一个以 $\sum_{n=1}^N X_{mn}\psi_n$ 为第 m 展开系数的态，它表示为列向量正是

$$\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1N} \\ X_{21} & X_{22} & \cdots & X_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ X_{N1} & X_{N2} & \cdots & X_{NN} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}! \quad (2.57)$$

这样，我们就将算符之间、算符和左右矢之间的抽象的作用表示为了看起来更具体的矩阵乘法。

另一方面，对 X^\dagger 有

$$X^\dagger = \sum_{m=1, n=1}^N X_{nm}^* |e_m\rangle\langle e_n|, \quad (2.58)$$

注意是 X_{nm}^* ，下标交换了。即， X^\dagger 的矩阵表示是 X 的矩阵表示的**共轭转置**。显然，厄米算符的矩阵表示，对角元必然是实数，而关于对角线对称的非对角元互为共轭。参考上面对 X 的处理，运用 X^\dagger 的定义(2.45)，动手验证一下！

练习

练习 5

考虑二维态空间，若对正交归一基矢 $\{|e_1\rangle, |e_2\rangle\}$ ，若算符 X 的作用为

$$X|e_1\rangle = |e_1\rangle - |e_2\rangle, \quad X|e_2\rangle = |e_1\rangle + |e_2\rangle,$$

用 $\{|e_1\rangle, |e_2\rangle\}$ 表示出 X 。若算符 Y 的作用为

$$Y|e_1\rangle = |e_1\rangle, \quad Y|e_2\rangle = |e_1\rangle,$$

用 $\{|e_1\rangle, |e_2\rangle\}$ 表示出 Y ，以及 XY 。

2.6 量子力学基本假设：薛定谔方程

2.6.1 时间演化算符

我担心你会对过多的代数感到疲惫，在进一步讨论有关算符的一般性质之前，让我们讨论一点物理——引入算符的一个重要的实例，**时间演化算符**（time-evolution operator）。

设系统在 t_0 时刻的态为 $|\psi(t_0)\rangle$ ，另一时刻 t 的态为 $|\psi(t)\rangle$ ，态是怎样从一个变成另一个的？我们想要的无非是“输入一个态，输出一个态”，可以用一个算符

来作为桥梁：

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle, \quad (2.59)$$

$U(t, t_0)$ 称为时间演化算符，简称演化算符，其有两个输入参数，表示演化从 t_0 开始，到 t 结束。对右矢而言算符的作用是从右到左的，故两个时间参数也从右到左写，注意不要搞错。

这样抽象的关系式能告诉我们什么呢？让我们通过一系列物理上的要求将其具体化。首先，如果我们待在同一个时刻，即 $t = t_0$ ，那么显然

$$U(t_0, t_0) = I; \quad (2.60)$$

其次，一段时间演化可以分成多步，如：

$$\begin{aligned} & |\psi(t_0)\rangle \\ & \Downarrow \\ & |\psi(t_1)\rangle = U(t_1, t_0) |\psi(t_0)\rangle \\ & \Downarrow \\ & |\psi(t)\rangle = U(t, t_1) |\psi(t_1)\rangle, \end{aligned} \quad (2.61)$$

即，演化算符的复合仍然是演化算符，参数“首尾相接”：

$$U(t, t_0) = U(t, t_1)U(t_1, t_0), \quad (2.62)$$

其中我们并不需要限制这几个参数的大小关系，这可以容纳反向演化⁴——已知后来的态，要求之前的态，如：

$$U(t_0, t_0) = I = U(t_0, t_1)U(t_1, t_0) \Rightarrow U(t_0, t_1) = U^{-1}(t_1, t_0), \quad (2.63)$$

即，交换演化算符的两个参数将得到逆算符。相应地，由式(2.45)让左矢从 t_0 演化到 t 的算符应为

$$U^\dagger(t, t_0). \quad (2.64)$$

人们往往将演化开始的参考时刻 t_0 设为 0，此时演化算符中可以只写出终止时刻 t ： $U(t)$ ，问题中不考虑多个终止时刻时甚至直接写为 U 。

考虑态的内积如何随时间演化还将得到另一个重要的要求，对任意初始态 $|\psi(t_0)\rangle$ ，归一化则初始内积为 1，由于自身内积有总概率这一物理意义，则任意时刻态的自身内积都应为 1：

$$\begin{aligned} \langle\psi(t)|\psi(t)\rangle &= 1 = \langle\psi(t_0)|U^\dagger(t, t_0)U(t, t_0)|\psi(t_0)\rangle \\ &= \langle\psi(t_0)|\psi(t_0)\rangle, \end{aligned} \quad (2.65)$$

⁴注意此处不是时间反演，时间反演对应的是另一个算符。

即，一次归一化，始终归一化。这告诉我们

$$U^\dagger(t, t_0)U(t, t_0) = I \Rightarrow U^\dagger(t, t_0) = U^{-1}(t, t_0), \quad (2.66)$$

演化算符取厄米共轭就得到逆算符，取 \dagger 等于取逆称为算符的**么正性**（unitary）。因此，人们常将量子力学中的时间演化称为**么正演化**。我们立刻看到么正演化的后果是，不仅态的自身内积，任意内积都不随时间变化⁵：

$$\langle \psi | \phi \rangle = \langle \psi | U^\dagger U | \phi \rangle, \quad (2.67)$$

这一性质时常称为**信息守恒**。

与么正演化针锋相对的则是投影测量，投影测量从现象上看起来会丢失信息，许多人将这视为我们宇宙中**不可逆性**的起源（之一）。当然，投影测量只是一个理想化的黑箱，测量过程的物理实质仍有待商榷。

通过考虑所谓的**无穷小演化**，我们可以得到演化算符的最具体形式。设从 t_0 时刻演化到与其邻近的 $t_0 + \delta t$ 时刻，演化算符为 $U(t_0 + \delta t, t_0)$ ，可对其形式上泰勒展开：

$$U(t_0 + \delta t, t_0) = I - i\delta t \Omega, \quad (2.68)$$

其中算符 Ω 刻画了这个无穷小演化算符相对于单位算符的偏离程度。是否考虑更高阶的项不影响最终结论，考虑一阶就已足够。么正性要求

$$\begin{aligned} I &= U^\dagger(t_0 + \delta t, t_0)U(t_0 + \delta t, t_0) \\ &\cong I - i\delta t \Omega + i\delta t \Omega^\dagger, \end{aligned} \quad (2.69)$$

其中我们只保留了一阶项，但若从一开始就保留高阶项则仍能得到同样结论。要让等式成立，就要求

$$\Omega = \Omega^\dagger, \quad (2.70)$$

即 Ω 是厄米算符。

这意味着什么呢？实际上，我们先前对演化算符的要求恰好是数学上对**群**（group）的要求，态空间上全体演化算符构成了一个群，**时间平移群**（time translation group），上面引入的算符 Ω 就是这个群的**生成元**（generator）。我们并不需要你熟知群论，重要的是，艾米·诺特（Emmy Noether）提出的定理⁶告诉我们，时间平移的生成元就代表能量！

准确地说， Ω 和**哈密顿量算符**（Hamiltonian）成正比，而**能量是哈密顿算符的本征值**，我们将在2.7.1中具体讨论。由于 U, I 都没有单位——无量纲，则与 δt

⁵注意这里内积的两个态都是要演化的，或者说**含时的**，如果有一个不含时——“按住不动”则不成立。

⁶诺特第一定理：物理系统的每一种**连续对称性**都对应一种守恒量。

相乘的 Ω 应该有时间倒数（频率）的量纲。为了量纲平衡， Ω 等于哈密顿算符 H 除以一个有角动量量纲的常量，提到量子力学中有角动量量纲的常量，我们自然想到普朗克常量：

$$\Omega = \frac{H}{\hbar}, \quad (2.71)$$

无数量子力学实验验证了这确实和我们在 SG 实验中遇到的是同一个常量。

2.6.2 薛定谔方程

一般地，考虑 $U(t + \delta t, t_0)$ 和 $U(t, t_0)$ 之差：

$$\begin{aligned} U(t + \delta t, t_0) - U(t, t_0) &= U(t + \delta t, t)U(t, t_0) - U(t, t_0) \\ &= (I - i\delta t \frac{H}{\hbar})U(t, t_0) - U(t, t_0) \\ &= -i\delta t \frac{H}{\hbar}U(t, t_0), \end{aligned} \quad (2.72)$$

我们得到

$$i\hbar \frac{\partial U(t, t_0)}{\partial t} = HU(t, t_0), \quad (2.73)$$

这就是**演化算符的薛定谔方程**（Schrödinger equation for the time-evolution operator）。

薛定谔方程

量子系统的时间演化由一个么正的时间演化算符给出，演化算符通过薛定谔方程由系统的哈密顿量决定。

演化算符的薛定谔方程，或者简单说，么正演化是量子力学的基本假设，无论是否考虑相对论、考虑的对象是粒子或是场，这个最广义的薛定谔方程都成立。薛定谔方程的线性性再次强调了量子力学是一个线性理论。在方程(2.73)两边同时右乘初始态 $|\psi(t_0)\rangle$ ，注意到这是个常矢量，可以放进时间求导内，有

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H |\psi(t)\rangle, \quad (2.74)$$

这就是（稍微狭义的）**态矢量的薛定谔方程**。它告诉我们态的时间演化**不能突变**，于是可以回答一个常见的问题：“测量投影后会怎么样？”答案是以投影得到的末态为新的初态代入上述方程，结果是态会“一点一点地”（关于时间必须可导！）偏离这个新的初态，如果短时间（相对于态演化的快慢）内再次测量，则大概率会投影到上一次的结果——态还没来得及偏离太多。

而薛定谔最初给出的狭义的薛定谔方程：

$$i\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t} = [-\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r})] \psi(\vec{r}, t), \quad (2.75)$$

则是非相对论单粒子的位置波函数的方程，其具体意义后面将会介绍。

为了求解微分方程(2.73)得到 U ，我们需要知道 H 。我们可以先考虑和我们要处理的量子系统类似的经典系统，其哈密顿量常常可以诠释为动能加势能 $H = T + V$ ，并且其中只包含位置和动量两个基本变量。如对一个自由运动的电子，考虑一个自由运动的经典质点

$$H = \frac{p^2}{2m}, \quad (2.76)$$

按这个表达式将动量替换为**动量算符**（我们将在后面引入）就得到了哈密顿算符⁷。当然，也有一些没有经典类似的量子系统，那时我们就要另想办法。

注意到

$$\begin{aligned} H &= i\hbar \frac{U(t_0 + \delta t, t_0) - U(t_0, t_0)}{\delta t} \\ &= i\hbar \left. \frac{dU(t, t_0)}{dt} \right|_{t=t_0}, \end{aligned} \quad (2.77)$$

作为演化算符本身的导数， H 本身也可以随时间变化，甚至不同时刻的 H 可以不对易。我们暂且考虑简单的情况：

$$\frac{\partial H}{\partial t} = 0, \quad (2.78)$$

此时我们只需要不断地进行无穷小演化，就可以得到完整的有限大演化，具体来说，将演化时长 $t - t_0$ 分成 N 份，则

$$U(t, t_0) = U(t, t_0 + \frac{N-1}{N}(t-t_0)) \cdots U(t_0 + \frac{1}{N}(t-t_0), t_0), \quad (2.79)$$

取 $N \rightarrow \infty$ 的极限，则有

$$U(t, t_0) = \lim_{N \rightarrow \infty} (1 - \frac{i}{\hbar} \frac{t-t_0}{N} H)^N, \quad (2.80)$$

仿照实数集上的极限 $\lim_{x \rightarrow \infty} (1 + \frac{1}{x})^x = e$ ，我们将上式右边的这样的极限定义为**算符的指数**，即

$$U(t, t_0) = \exp \left[-i \frac{(t-t_0)H}{\hbar} \right]. \quad (2.81)$$

求出上述极限就求出了演化算符，由于 H 不变，只需要调整其中的初末时刻，作用到态上，就可以得到我们想要的时间演化。

⁷这里出现的另一个量，质量 m 为什么不换成算符呢？一种回答是，经典哈密顿力学中只考虑正则坐标和正则动量，既然量子力学继承其概念，那我们可以合理认为只有位置和动量需要算符化。但实际上有**质量算符**！比如对中微子我们就引入了质量算符，为什么？因为我们观测到了**中微子振荡**——中微子处于不同质量的叠加态——的现象。换言之，这很大程度上取决于**经验事实**的驱动。

2.7 量子力学基本假设：力学量算符

2.7.1 算符的本征值与本征态

算符对态矢的改变往往是很“剧烈”的——新的态矢的大小和方向都与旧的态矢不同，但有时比较“温和”，新的态矢与旧的态矢平行，只差一个复常数：

$$X|\psi\rangle = c|\psi\rangle, \quad (2.82)$$

其中态 $|\psi\rangle$ 为算符 X 的**本征态** (eigenstate)，相应的常数 c 为**本征值** (eigenvalue)，这一关系就称为本征方程。

正如我们提到过的，对同一个常数 c ，还可能存在态 $|\phi\rangle, |\chi\rangle, \dots$ ，都满足

$$X|\sim\rangle = c|\sim\rangle, \quad \sim = \psi, \phi, \chi, \dots, \quad (2.83)$$

即 X 的本征值 c 出现了**简并**，这种情况我们留到2.7.5中再讨论。现在我们只考虑非简并的情况，于是我们可以用本征值去标记本征态。可以用大写字母表示算符，用相应的小写字母表示本征值⁸：

$$X|x_n\rangle = x_n|x_n\rangle, \quad n = 1, \dots, N, \quad (2.84)$$

这就给出了一系列的本征方程，我们可以将其解读为：**算符可以从其本征态中提取信息**，进而我们可以希望用算符来表达物理可观测量。

另一方面，利用共轭线性有

$$\langle x_n|X^\dagger = \langle x_n|x_n^*. \quad (2.85)$$

即， X 的本征态的对偶是 X^\dagger 的本征态，本征值取共轭。

既然在2.5中我们已经将算符表示成了矩阵，那就可以利用线性代数中的结论来求解本征值和本征态（的展开系数）。即，设本征值为 x ，让算符矩阵减去 x 倍的单位矩阵，再求行列式为零的方程：

$$|(X) - xI| = 0, \quad (2.86)$$

其中 (X) 代表 X 的矩阵表示，以和 X 本身相区分。求出本征值后再代入矩阵形式的本征方程，就能求出本征态对应的列向量。你可能并不熟悉线性代数，但没关系，让我们看一个具体例子。

⁸对于如何表达算符及其矩阵表示、与普通的数量相区分，有许多约定，如在算符上面加 hat（帽子）： \hat{X} 。许多文献干脆不在符号上做区分，而是根据上下文来确定哪些是算符，实践表明，这样并不会带来多少混淆。

考虑一个二维态空间上的算符 Y ，设它对正交归一基矢 $\{|e_1\rangle, |e_2\rangle\}$ 的作用规则为

$$Y|e_1\rangle = i|e_2\rangle, \quad Y|e_2\rangle = -i|e_1\rangle, \quad (2.87)$$

则有

$$Y = -i|e_1\rangle\langle e_2| + i|e_2\rangle\langle e_1|, \quad (2.88)$$

(动手验证一下!) 于是矩阵表示为

$$(Y) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (2.89)$$

求解本征值:

$$\begin{aligned} \begin{vmatrix} -y & -i \\ i & -y \end{vmatrix} &= 0, \\ \Rightarrow y^2 - 1 &= 0, \\ \Rightarrow y_1 = 1, y_2 &= -1, \end{aligned} \quad (2.90)$$

设 (Y) 的本征列向量为 $(\alpha, \beta)^T$ ，则有

$$\begin{aligned} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \\ \Rightarrow \begin{pmatrix} -i\beta \\ i\alpha \end{pmatrix} &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \\ \Rightarrow \alpha &= -i\beta, \\ \xrightarrow{|\alpha|^2 + |\beta|^2 = 1} \alpha &= \frac{1}{\sqrt{2}}, \beta = \frac{i}{\sqrt{2}}, \end{aligned} \quad (2.91)$$

则 Y 的与本征值 $y_1 = 1$ 相对应的本征态为

$$\frac{1}{\sqrt{2}}|e_1\rangle + \frac{i}{\sqrt{2}}|e_2\rangle, \quad (2.92)$$

同理，与本征值 $y_2 = -1$ 相对应的本征态为

$$\frac{1}{\sqrt{2}}|e_1\rangle - \frac{i}{\sqrt{2}}|e_2\rangle. \quad (2.93)$$

2.7.2 力学量算符

在2.7.1中我们已经看到，有一种从抽象的态中获得具体的值——提取信息的方式，那就是算符的本征方程。我们可以希望用算符来代表可观测量——但并不是随便什么算符，而是厄米算符。

2.7.1最后举例的 Y 满足 $Y = Y^\dagger$ ，就是一个厄米算符，你注意到了吗？虽然它的展开式/矩阵表示中含有虚数，但它的本征值是实数。后面我们会看到，它其实就是 y 方向自旋算符 S_y （差个常量），而 $|e_1\rangle, |e_2\rangle$ 则是 $|S_z+\rangle, |S_z-\rangle$ 。

可以证明，厄米算符的本征值必然是实数，并且不同本征值对应的本征态必然正交。由我们讲过的内积的性质(2.15)，有

$$\begin{aligned}\langle\beta|X|\alpha\rangle &= \langle\beta|(X|\alpha\rangle) \\ &= \langle\alpha|X^\dagger|\beta\rangle^*,\end{aligned}\tag{2.94}$$

那么对厄米算符 $A = A^\dagger$ 有

$$\langle\beta|A|\alpha\rangle = \langle\alpha|A|\beta\rangle^*,\tag{2.95}$$

设 A 有 N 个非简并本征值 a_n ，相对应的本征态 $|a_n\rangle$ ，则有

$$\langle a_m|A|a_n\rangle = \langle a_m|A^\dagger|a_n\rangle,\tag{2.96}$$

由(2.85)有

$$(a_m^* - a_n)\langle a_m|a_n\rangle = 0.\tag{2.97}$$

分两种情况，第一，如果 $m = n$ ，那就推出

$$a_n^* = a_n,\tag{2.98}$$

证明了厄米算符的本征值为实数。第二，如果 $m \neq n$ ，利用刚刚得到的结论， $a_m^* = a_m \neq a_n$ ，那就推出

$$\langle a_m|a_n\rangle = 0,\tag{2.99}$$

证明了厄米算符的不同本征值对应的本征态相互正交。

这两点对描述物理意义重大！本征值为实数意味着可以比较大小，而对物理量的测量实际上就是和物理量的单位比大小——如把某个国王的臂展作为长度的单位，则测量一根杆的长度就意味着取这根杆能容下多少个这位国王的臂展；而不同本征值对应的本征态相互正交，意味着不同的本征值物理上可以完全区分，这正是我们需要的！

力学量算符假设

力学量可以用厄米算符来表示。不同于经典力学量，仅当系统处于本征态上时力学量才有确定的取值，即相应的本征值。

我们将那些要表达成算符的物理量专门称为**力学量**，那些不表达成算符的物理量可以称为**参数**。比如，在“量子力学”一词狭义所指的范围内，粒子的质量和电荷往往是参数而不是算符。

在今天看来，可观测量不应仅仅限制为厄米算符，有许多可观测量效应并不能简单地由厄米算符描述，比如干涉实验中实际上测量的是相位差，现代物理中经常视为可观测量的关联函数也不一定是实的。

物理上，我们抱有一个信念：厄米算符的本征态系是完备的⁹，因而可以作为一组基矢：

$$\sum_{n=1}^N |a_n\rangle\langle a_n| = I. \quad (2.100)$$

这称为选择了一个表象 (representation)：A 表象 (自旋表象、能量表象……)。这样，之前介绍的略显抽象的基矢就有了物理对应，力学量的不同取值对应的本征态——“所有可能性”——可以自然地用来展开任意的态。

我们在引入投影测量假设时已经隐含了这一点，如果完备性不成立，那系统的态在测量时就可能投影到一个没有取值——无论是本征值还是期望值——的成分上，而这在物理上看是不可思议的。之所以说是信念是因为这一点虽然对物理很重要，但只在一些简单的情况下能得到数学证明。自量子力学诞生以来，物理学家总是默认它成立：拿到一个态，用力学量的本征态展开，至今还没有遇到不可克服的困难。

我们说过，算符可以用任意的正交归一基矢展开，而对力学量有一套特殊的基矢，那就是它的本征态组成的基矢。对力学量 A 有

$$\begin{aligned} A &= A \sum_{n=1}^N |a_n\rangle\langle a_n| \\ &= \sum_{n=1}^N a_n |a_n\rangle\langle a_n|, \end{aligned} \quad (2.101)$$

也可以从式(2.54)出发得到这一结果。这就是说，矩阵的对角元 ($m = n$) 就是 A 的各个本征值，非对角元为零，A 在自身表象下为对角矩阵：

$$A \doteq \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_N \end{pmatrix}. \quad (2.102)$$

当力学量算符作用于任意态时，有

$$\begin{aligned} A|\psi\rangle &= A \sum_{n=1}^N \langle a_n|\psi\rangle |a_n\rangle \\ &= \sum_{n=1}^N a_n \langle a_n|\psi\rangle |a_n\rangle \\ &\neq c|\psi\rangle, \end{aligned} \quad (2.103)$$

⁹本征态系即本征态的集合的简称。

可见，一般而言力学量没有确定的取值，但我们先前引入的玻恩定则可以定义其期望值：

$$\langle A \rangle = \sum_{n=1}^N |\langle a_n | \psi \rangle|^2 a_n, \quad (2.104)$$

即，力学量 A 的取值 a_n ，在 $|\psi\rangle$ 态中测得的概率为 $|\langle a_n | \psi \rangle|^2$ ，于是 A 在 $|\psi\rangle$ 上的期望值就是用这概率对本征值做加权平均。改写一下上式，我们有

$$\begin{aligned} \langle A \rangle &= \sum_{n=1}^N \langle a_n | \psi \rangle \langle \psi | a_n \rangle a_n \\ &= \sum_{n=1}^N \langle \psi | a_n | a_n \rangle \langle a_n | \psi \rangle \\ &= \langle \psi | \left(\sum_{n=1}^N a_n | a_n \rangle \langle a_n | \right) | \psi \rangle \\ &= \langle \psi | A | \psi \rangle, \end{aligned} \quad (2.105)$$

这是力学量期望值的另一种表达，它的优越性在于只涉及力学量 A 和给定的态 $|\psi\rangle$ ，而没有涉及选取什么样的基矢——理应如此！选取什么基矢纯粹取决于理论计算的方便与否，大自然并不在乎。**期望值独立于基矢的选择**，是基矢变换下的**不变量**（invariant），不变量在现代物理学中有举足轻重的地位。

2.7.3 算符的函数

你可能仍对2.6.2最后的指数心存疑虑，“算符的指数？看上去是很像，但具体是什么？”首先，我们早已引入的算符的复合以及逆算符允许我们轻易定义算符的整数次幂：

$$\begin{aligned} X^n &:= \underbrace{XX \cdots X}_{n \uparrow X}, \quad n > 0, \\ X^{-n} &:= \underbrace{X^{-1}X^{-1} \cdots X^{-1}}_{n \uparrow X^{-1}}, \quad n > 0. \end{aligned} \quad (2.106)$$

再约定 $X^0 = I$ ，我们就可以写出算符的多项式函数，如

$$f(X) = X^2 + X. \quad (2.107)$$

当有多个不同算符时，就需要注意算符之间一般不对易这一点，如

$$(X + Y)(X - Y) = X^2 + YX - XY - Y^2 \neq X^2 - Y^2. \quad (2.108)$$

有了多项式函数，我们就可以把实数函数的展开式（原本是推论）拿来作为算符函数的定义。如算符的 e 指数即为：

$$e^X := I + X + \frac{X^2}{2} + \frac{X^3}{3!} + \cdots, \quad (2.109)$$

当然，数学上还要提出展开式的收敛性等问题。物理上最常见的算符的函数就是多项式和指数函数，它们往往都有良好的收敛性——指在理论预言与实验对比时给出了合理的结果。

特别地，对于力学量算符的函数，我们有一种更自然的定义方式。我们定义力学量 A 的函数 $f(A)$ 满足

$$f(A) |a_n\rangle := f(a_n) |a_n\rangle, \quad (2.110)$$

于是，利用力学量本征态系的完备性(2.100)， $f(A)$ 对任意态 $|\psi\rangle$ 的作用就明确了：

$$\begin{aligned} f(A) |\psi\rangle &= f(A) \left(\sum_{n=1}^N |a_n\rangle \langle a_n| \right) |\psi\rangle \\ &= \left(\sum_{n=1}^N f(a_n) |a_n\rangle \langle a_n| \right) |\psi\rangle, \end{aligned} \quad (2.111)$$

即

$$f(A) = \sum_{n=1}^N f(a_n) |a_n\rangle \langle a_n|. \quad (2.112)$$

可以证明，这与上述的展开式定义是相容的。

现在，我们就可以具体给出2.6.2最后的指数了，首先求解其本征方程

$$H |E_n\rangle = E_n |E_n\rangle, n = 1, 2, \dots, N, \quad (2.113)$$

这称为定态薛定谔方程。由上述算符函数定义，演化算符为

$$U(t, t_0) = \sum_{n=1}^N \exp \left[-\frac{i(t-t_0)E_n}{\hbar} \right] |E_n\rangle \langle E_n|, \quad (2.114)$$

对任意初始态 $|\psi(t_0)\rangle$ 的作用为

$$U(t, t_0) |\psi(t_0)\rangle = \sum_{n=1}^N \exp \left[-\frac{i(t-t_0)E_n}{\hbar} \right] \langle E_n | \psi(t_0) \rangle |E_n\rangle, \quad (2.115)$$

即，要求 t 时刻的态，只需要把 $|\psi(t_0)\rangle$ 用能量本征态去展开，再对每个展开系数都乘一个 e 指数 $e^{-i(t-t_0)E_n/\hbar}$ 即可。或者说，在能量表象下（以 $\{|E_n\rangle\}$ 为基），(2.114)的矩阵表示为对角矩阵，乘到同样在能量表象下¹⁰ 写出的 $|\psi(t_0)\rangle$ 对应的列向量上，就得到了 $|\psi(t)\rangle$ 对应的列向量。

我们得到了态的薛定谔方程(2.74)的以 $|\psi(t_0)\rangle$ 为初始条件的解。

2.7.4 相容力学量

我们说过，算符之间一般不对易，我们定义一种运算来衡量两个算符之间不对易的程度：

$$[A, B] := AB - BA, \quad (2.116)$$

¹⁰在表象中做计算时，务必要确保算符和态矢是在同一个表象下。

这称为 A 和 B 的**对易子**，在量子力学中有着重要的作用。

如果两个力学量对易，则称它们**相容**，两个相容力学量有**共同本征态系**，不相容则没有。如不同方向的自旋算符就不相容，所以没有共同本征态系，而是有像式(1.20)这样的关系。而某个方向的自旋算符，和我们还没介绍的**总自旋算符**则是相容的。

证明如下，设 $[A, B] = 0$ （零算符），其矩阵元显然都为零，如以 A 的本征态系为基：

$$\langle a_m | [A, B] | a_n \rangle = 0, \quad (2.117)$$

那么就有

$$\begin{aligned} 0 &= \langle a_m | AB | a_n \rangle - \langle a_m | BA | a_n \rangle \\ &= (a_m - a_n) \langle a_m | B | a_n \rangle \\ \Rightarrow &\begin{cases} m \neq n, & \langle a_m | B | a_n \rangle = 0, \\ m = n, & \langle a_m | B | a_n \rangle \text{ 不一定为 } 0, \end{cases} \end{aligned} \quad (2.118)$$

换句话说

$$B_{mn} = \delta_{mn} \langle a_n | B | a_n \rangle. \quad (2.119)$$

由式(2.54)展开 B ，利用上式有

$$\begin{aligned} B &= \sum_{m=1, n=1}^N B_{mn} |a_m\rangle \langle a_n| \\ &= \sum_{m=1, n=1}^N \delta_{mn} \langle a_n | B | a_n \rangle |a_m\rangle \langle a_n| \\ &= \sum_{n=1}^N \langle a_n | B | a_n \rangle |a_n\rangle \langle a_n|, \end{aligned} \quad (2.120)$$

即， B 在 $\{|a_n\rangle\}$ 下也是对角的， $|a_i\rangle$ 是 A 和 B 的**共同本征态**，可以记作 $|a_i, b_i\rangle$ ：

$$\begin{aligned} B |a_i, b_i\rangle &= \sum_{n=1}^N B_{nn} |a_n\rangle \langle a_n | a_i, b_i \rangle \\ &= B_{ii} |a_i, b_i\rangle = b_i |a_i, b_i\rangle, \\ b_i &\equiv \langle a_i | B | a_i \rangle. \end{aligned} \quad (2.121)$$

从 B 出发同理，试一试！

2.7.5 简并与去简并

了解了相容力学量的上述性质后，我们就可以讨论简并这一有趣现象了。我们曾由经典一维运动质点为例引入了简并度的概念，现在让我们明确其意义。设 A

的某个本征值 a_i 存在简并，你也许会想把 a_i 对应的多个本征态一个个罗列出来，再加上一个指标来标记，如：

$$|a_i^{(1)}\rangle, |a_i^{(2)}\rangle, \dots \quad (2.122)$$

但是态叠加原理，或者说线性性告诉我们，任意两个这样态的线性组合仍然是 a_i 的本征态，所以直接这样标记是不可行的。所幸，从它们中可以找出一组相互正交的态，这样就可以像上面这样一个个标记了。

假设我们已经找到这样一组相互正交的简并态，既然 A 作用于任意一个由它们叠加出的态都仍然得到 a_i ，而和其它的 $a_{n \neq i}$ 无关，它们就形成了一个独立的小天地，我们称为 a_i 对应的**简并子空间**。回看之前的证明(2.97)，在考虑简并时变为

$$(a_m^* - a_n) \langle a_m^{(k)} | a_n^{(l)} \rangle = 0, \quad (2.123)$$

即 a_m^* 和 a_n 都可以对应多个不同的本征态，取 $m = n$ 一般无法推出 $\langle a_n^{(k)} | a_n^{(l)} \rangle = 1$ ，也就无法证明本征值为实数。但若找到相互正交的简并态那就得证了，进而 $m \neq n$ 时

$$(a_m - a_n) \langle a_m^{(k)} | a_n^{(l)} \rangle = 0, \quad (2.124)$$

可以推出 $\langle a_m^{(k)} | a_n^{(l)} \rangle = 0$ ，不同本征值对应的简并子空间正交，可见子空间这个名字是恰如其分的。

此时完备性关系应该写为

$$\sum_{n=1, n \neq i, j, \dots}^N |a_n\rangle\langle a_n| + \sum_{m_i=1}^{M_i} |a_i^{(m_i)}\rangle\langle a_i^{(m_i)}| + \sum_{m_j=1}^{M_j} |a_j^{(m_j)}\rangle\langle a_j^{(m_j)}| + \dots = I. \quad (2.125)$$

称彼此正交的简并态的数目为简并度。即在 A 的 N 个不同本征值中，有 ω 个简并本征值 a_i, a_j, \dots ，简并度分别为 M_i, M_j, \dots ，可以知道态空间应该有 $N - \omega + M_i + M_j + \dots$ 维。

再回看在2.7.4中我们对共同本征态系的证明，现在若对 $[A, B] = 0$ ， A 有简并而 B 非简并，则有

$$\begin{aligned} 0 &= \langle a_m^{(k)} | [A, B] | a_n^{(l)} \rangle \\ &= (a_m - a_n) \langle a_m^{(k)} | B | a_n^{(l)} \rangle \\ &\Rightarrow \begin{cases} m \neq n, & \langle a_m^{(k)} | B | a_n^{(l)} \rangle = 0, \\ m = n, & \langle a_m^{(k)} | B | a_n^{(l)} \rangle \text{ 不一定为0,} \end{cases} \end{aligned} \quad (2.126)$$

即

$$\begin{aligned} B_{mn}^{(kl)} &= \delta_{mn} \langle a_n^{(k)} | B | a_n^{(l)} \rangle, \\ B &= \sum_{n=1}^N \sum_{k,l=1}^{M_n} B_{nn}^{(kl)} |a_n^{(k)}\rangle\langle a_n^{(l)}|, \end{aligned} \quad (2.127)$$

可见, 此时 $|a_i^{(s)}\rangle$ 一般不是 B 的本征态

$$\begin{aligned} B|a_i^{(s)}\rangle &= \sum_{k,l=1}^{M_i} B_{ii}^{(kl)} |a_i^{(k)}\rangle \langle a_i^{(l)}|a_i^{(s)}\rangle \\ &= \sum_k B_{ii}^{(ks)} |a_i^{(k)}\rangle \\ &\neq c |a_i^{(s)}\rangle. \end{aligned} \quad (2.128)$$

实际上, B 此时表示为分块对角矩阵, 如对仅有 a_i 一个二重简并本征值的情况:

$$B \doteq \begin{pmatrix} B_{11} & & & & \\ & \ddots & & & \\ & & B_{ii}^{(11)} & B_{ii}^{(12)} & \\ & & B_{ii}^{(21)} & B_{ii}^{(22)} & \\ & & & \ddots & \\ & & & & B_{N+1,N+1} \end{pmatrix}. \quad (2.129)$$

但是反过来, 我们发现非简并的 B 的本征态一定是 A 的本征态! 为此只需将2.7.4最后的证明中的 Aa 和 Bb 交换即可。于是, 我们只需要用 $\{b_n\}$ 作为基, 就可以同时对角化 A 和 B ! $\{b_n\}$ 对 B 本身的对角化自不必说, 对 A 有

$$A|a_i, b_j\rangle = a_i |a_i, b_j\rangle, A|a_i, b_{j+1}\rangle = a_i |a_i, b_{j+1}\rangle, \dots, \quad (2.130)$$

在所有可能的相互正交的 $\{|a_i^{(1)}\rangle, |a_i^{(2)}\rangle, \dots\}$ 中, $\{|a_i, b_j\rangle, |a_i, b_{j+1}\rangle, \dots\}$ 是最自然的, 使得

$$\begin{aligned} A &= \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \langle a_i, b_j | A | a_i, b_j \rangle | a_i, b_j \rangle \langle a_i, b_j |, \\ B &= \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \langle a_i, b_j | B | a_i, b_j \rangle | a_i, b_j \rangle \langle a_i, b_j |. \end{aligned} \quad (2.131)$$

其中 N_A, N_B 为 A, B 不同本征值数, N_B 即实际的态空间维度。称 B 的本征态对角化了 A 或者对 A 去简并。举一个经典的例子, 做一维运动的质点, 仅用动能描述则有二重简并, 但如果考虑了动量, 那就能去除简并。

2.7.6 力学量完全集

如果 B 也有简并呢? 当然, B 的简并不一定和 A 的简并“重合”, 运气好的话两者可以互相对角化。但一般而言, 我们总可以引入更多与它们相容的力学量 C, D, \dots , 满足

$$[A, B] = [B, C] = [A, C] = \dots = 0, \quad (2.132)$$

直到再也不能引入更多为止, 用这样一组力学量, 我们就必然可以用它们的共同本征态来去除其中任一个力学量的简并:

$$|a_n, b_n, c_n, \dots\rangle. \quad (2.133)$$

物理上看, 当我们测量 A , 得到简并本征值 a_i 时, 我们并不确定系统到底处于什么态, 只知道处于某种简并态的线性组合:

$$\sum_{m=1}^M |a_{im}\rangle, \quad (2.134)$$

如果我们在此基础上再测量与 A 相容, 且没有简并的 B , 那我们就明确地知道系统会投影到

$$|a_i, b_i\rangle, |a_i, b_{i+1}\rangle, \dots, |a_i, b_{i+M-1}\rangle, \quad (2.135)$$

其中之一, 对这些态再测量 A , 当然还会得到 a_i , 于是系统的量子态就确定了。如果 B 也有简并, 那我们就再引入 C , 以此类推。最终我们就能完全确定系统的量子态了, 故将彼此相容的、不能再扩充的力学量集合 $\{A, B, C, \dots\}$ 称为**力学量完全集** (complete set of commuting observable, CSCO)。

2.8 不确定性关系

2.8.1 不确定性与不确定性关系

在引入具体的物理模型之前, 让我们再最后讨论一个一般的性质: 不相容力学量之间的**不确定性关系** (uncertainty relation)。

我们已经给出了力学量的期望值(2.105), 为了构成一个较为完整的统计描述, 我们还需要引入不确定性。考虑力学量 A 和其期望值的差

$$A - \langle A \rangle, \quad (2.136)$$

准确地说, 是 $A - \langle A \rangle I$, 我们希望这个差算符能刻画 A 偏离期望值的程度。直接求这个差的期望是平庸的:

$$\langle A - \langle A \rangle \rangle = \langle A \rangle - \langle \langle A \rangle \rangle = 0. \quad (2.137)$$

我们转而求它的平方的期望:

$$\begin{aligned} \langle (A - \langle A \rangle)^2 \rangle &= \langle A^2 - 2\langle A \rangle A + \langle A \rangle^2 \rangle \\ &= \langle A^2 \rangle - \langle A \rangle^2, \end{aligned} \quad (2.138)$$

即, A 平方的期望减去期望的平方, 在统计学中这称为**方差**。我们再将其开方以得到量纲和 A 相同的量:

$$\sigma_A := \sqrt{\langle A^2 \rangle - \langle A \rangle^2}, \quad (2.139)$$

这就是 A （相对于态 $|\psi\rangle$ ）的**不确定性（标准差）**。许多文献中将其写为 ΔA ，而我们将 Δ 专门用来表达两个量直接做差， σ 专门表达不确定性。

当我们考虑的态是 A 的本征态时， $\sigma_A = 0$ ：

$$\begin{aligned}
 \langle A \rangle &= \sum_{i=1}^M c^{(i)*} \langle a_n^{(i)} | A \sum_{j=1}^M c^{(j)} | a_n^{(j)} \rangle \\
 &= \sum_{i=1}^M a_n |c^{(i)}|^2 \\
 &= a_n, \\
 \langle A^2 \rangle &= \sum_{i=1}^M c^{(i)*} \langle a_n^{(i)} | A^2 \sum_{j=1}^M c^{(j)} | a_n^{(j)} \rangle \\
 &= \sum_{i=1}^M a_n^2 |c^{(i)}|^2 \\
 &= a_n^2, \\
 \sigma_A &= \sqrt{a_n^2 - a_n^2} = 0,
 \end{aligned} \tag{2.140}$$

其中我们考虑了可能本征态是简并态的叠加， $\{a_n^{(i)}\}$ 为相互正交的简并态——非简并情况你容易自己写出。这是显然的，毕竟，本征态就是力学量能提取出确定的值的态（确定值态），简并也并不意味着这一点。今后在计算期望值时只需要考虑简并子空间整体的权重，而不需要考虑子空间内的态的相对关系。

对一般的叠加态 $|\psi\rangle = \sum_{n=1}^N \psi_n |a_n\rangle$ 则有

$$\sigma_A = \sqrt{\sum_{n=1}^N |\psi_n|^2 a_n^2 - \left(\sum_{m=1}^N |\psi_m|^2 a_m \right)^2}, \tag{2.141}$$

其中关系不容易看出，让我们换一种写法：

$$\begin{aligned}
 \sigma_A &= \sqrt{\langle \psi | A A | \psi \rangle - \langle \psi | A | \psi \rangle^2} \\
 &= \sqrt{\langle \psi | A A | \psi \rangle \langle \psi | \psi \rangle - |\langle \psi | A | \psi \rangle|^2},
 \end{aligned} \tag{2.142}$$

其中用到 $\langle \psi | \psi \rangle = 1$ 以及 $\langle \psi | A | \psi \rangle$ 为实数。我们引入一个不等式，施瓦茨不等式：

$$\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \geq |\langle \alpha | \beta \rangle|^2, \tag{2.143}$$

对此我们不作证明，你可以类比欧氏矢量的 $|\vec{a}|^2 |\vec{b}|^2 \geq |\vec{a} \cdot \vec{b}|^2$ 来理解。将 $A|\psi\rangle$ 当作 $|\alpha\rangle$ ， $|\psi\rangle$ 当作 $|\beta\rangle$ ，于是证得

$$\sigma_A \geq 0, \tag{2.144}$$

取等条件即是 $|\psi\rangle$ 为 A 的本征态。

现在我们给出不确定性关系：

不确定性关系

两个力学量 A 和 B 的不确定性的乘积有一个下限，这个下限和它们的对易子的期望值有关：

$$\sigma_A \sigma_B \geq \frac{1}{2} |\langle [A, B] \rangle|. \quad (2.145)$$

这一关系最早由海森堡给出，它告诉我们不相容力学量一般无法同时取确定值，即不相容力学量不存在共同本征态系的另一种说法。值得注意的是，不确定性乘积的下限是**对易子的期望值**决定的，对不相容力学量，对易子非零，但可能存在使对易子期望为零的态，于是在这样的态上不相容力学量也可以同时确定。

不确定性关系(2.145)可以证明如下，令

$$|\alpha\rangle = (A - \langle A \rangle) |\psi\rangle, |\beta\rangle = (B - \langle B \rangle) |\psi\rangle, \quad (2.146)$$

由施瓦茨不等式(2.143)，

$$\begin{aligned} \sigma_A \sigma_B &= \sqrt{\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle} \\ &\geq \sqrt{|\langle \alpha | \beta \rangle|^2} = |\langle \alpha | \beta \rangle| \\ &= \sqrt{(\operatorname{Re} \langle \alpha | \beta \rangle)^2 + (\operatorname{Im} \langle \alpha | \beta \rangle)^2} \\ &\geq |\operatorname{Im} \langle \alpha | \beta \rangle| \\ &= \frac{1}{2} |\langle \alpha | \beta \rangle - \langle \beta | \alpha \rangle|, \end{aligned} \quad (2.147)$$

其中 Im 为取虚部。计算 $\langle \alpha | \beta \rangle$ ：

$$\begin{aligned} \langle \alpha | \beta \rangle &= \langle \psi | AB - \langle A \rangle B - \langle B \rangle A + \langle A \rangle \langle B \rangle | \psi \rangle \\ &= \langle \psi | AB | \psi \rangle - \langle \psi | A | \psi \rangle \langle \psi | B | \psi \rangle, \\ \langle \beta | \alpha \rangle &= \langle \alpha | \beta \rangle^* \\ &= \langle \psi | BA | \psi \rangle - \langle \psi | A | \psi \rangle \langle \psi | B | \psi \rangle. \end{aligned} \quad (2.148)$$

于是证得不确定性关系：

$$\sigma_A \sigma_B \geq \frac{1}{2} |\langle [A, B] \rangle|. \quad (2.149)$$

显然，随着时间演化，不等式右边的下限可以随时间变化。一个下限固定的例子是最为人熟知的，**位置与动量的不确定性关系**¹¹：

$$\sigma_q \sigma_p \geq \frac{\hbar}{2}, \quad (2.150)$$

其对任意态成立，反推得知

$$|[q, p]| = \hbar I. \quad (2.151)$$

¹¹这里使用 q 代表位置。

实际上, 这一点告诉我们描述位置和动量的算符(态空间)必然是无穷维的。以两个二维力学量为例, 在其中一个量的表象下矩阵表示为

$$\begin{aligned} & \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} c & d \\ d^* & e \end{pmatrix} - \begin{pmatrix} c & d \\ d^* & e \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \\ &= \begin{pmatrix} 0 & ad - bd \\ -ad^* + bd^* & 0 \end{pmatrix}, \end{aligned} \quad (2.152)$$

我们发现, 对易子矩阵的迹——对角元之和——为零, 我们将算符本身的迹也定义为对角元之和, 迹也是基矢变换下的不变量:

$$\begin{aligned} \text{tr } A &= \sum_{a_n} \langle a_n | A | a_n \rangle \\ &= \sum_{a_n} \langle a_n | \sum_{b_i} | b_i \rangle \langle b_i | A \sum_{b_j} | b_j \rangle \langle b_j | a_n \rangle \\ &= \sum_{a_n} \sum_{b_i} \sum_{b_j} \langle b_j | a_n \rangle \langle a_n | b_i \rangle \langle b_i | A | b_j \rangle \\ &= \sum_{b_i} \sum_{b_j} \langle b_j | b_i \rangle \langle b_i | A | b_j \rangle \\ &= \sum_{b_i} \langle b_i | A | b_i \rangle. \end{aligned} \quad (2.153)$$

任意有限维力学量的对易子, 迹都为零, 因而必然不能正比于单位算符, 不确定性乘积的下限也就不可能是固定的常数。

2.8.2 不确定性的物理意义

上文定义的不确定性的物理意义是很清楚的。不确定性, 进而不确定性关系, 是发端于态叠加原理的、基于态矢量和算符的量子力学理论结构的 **必然结果**。

为了实际观测不确定性, 以及更早引入的期望值, 我们需要引入系综 (ensemble)。理论上, 只要得知了系统的态 $|\psi\rangle$, 那任意力学量 A 的 $\langle A \rangle, \sigma_A$ 都容易由定义算出, 但实验中我们往往不知道 $|\psi\rangle$ 。如果测量系统, 我们能确定的只是测量后的态是什么, 而无法得知测量前的态——除非我们准备大量都由 $|\psi\rangle$ 描述的系统, 称为系综¹²。系综中每个系统都随机地投影, 统计其中各个测量结果的频率, 我们就能逐渐逼近理论概率 $\langle a_n | \psi \rangle$, 进而得到 $\langle A \rangle, \sigma_A$ 。量子力学中所说的重复测量往往都是指对系综的测量。

¹²这种系综称为纯系综。还有一种系综称为混合系综, 其中系统不是由同一个态矢量 $|\psi\rangle$ 描述, 而是由同一组可能的态矢量 $\{|\psi_1\rangle, \dots, |\psi_N\rangle\}$ 和一组概率 $\{p_1, \dots, p_N\}$ 描述, 每个系统都有概率由其中一个态矢量描述。这里说的概率是经典概率, 没有相干性, 故也称为非相干叠加。

于是，不确定性在理论上是量子态的内禀属性，在实验上则是系综的属性。由于力学量之间的不相容性，一个态不可能是所有力学量的共同本征态，它必然对某些力学量是叠加态，也就伴随着相应的内禀不确定性。

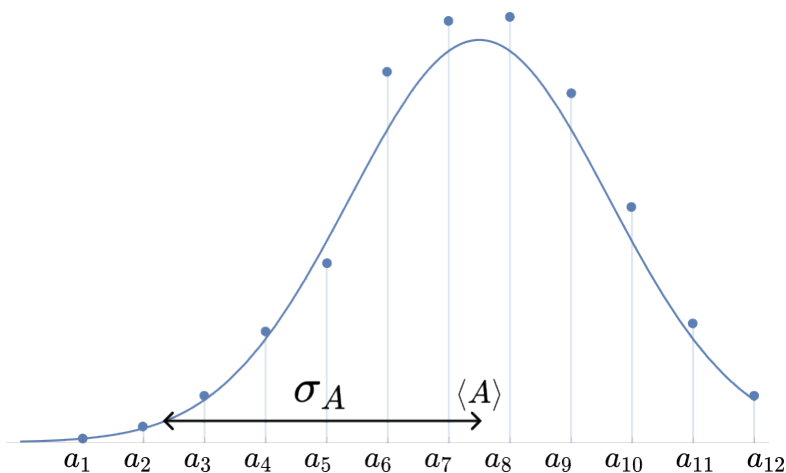


图 2.5. 本征值、期望值、不确定性示意图。图中曲线由理论概率连成，散点为实验测得频率。

我们应当将这种内禀不确定性与实验误差区分开来，图2.5中可以看到，不确定体现为理论概率曲线的“宽度”。而实验误差在这里体现为什么呢？设系综中有些系统投影到 $|a_n\rangle$ ，理论上我们的测量结果就是 a_n ，但误差使得这些投影到 $|a_n\rangle$ 的系统，有些输出值比 a_n 大一点，有些则小一点——实验测量结果落在以 a_n 为中心的一个区间。最终， $\langle A \rangle_{\text{实验}}$ 和理论值也就略有不同。

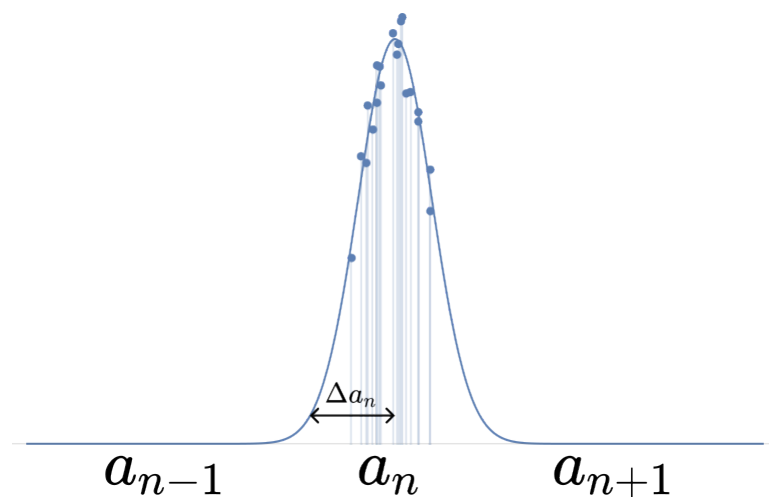


图 2.6. 测量本征值 a_n 的误差示意图，图2.5中一个点的放大。若误差 Δa_n 太大，则 a_n 与 a_{n-1}, a_{n+1} 将无法区分开，测量也不再能看成理想的投影测量。

对不相容力学量 A, B ，图2.7直观地体现了不确定性关系。当不确定性乘积（曲面下面积）相对于所关心的尺度¹³很小时，就可以近似认为 A, B 同时确定，取

¹³准确地说，相对于作用量—— \hbar 既是角动量的自然单位，也是作用量的自然单位。

值即为期望值 $\langle A \rangle, \langle B \rangle$ 。如对加速器中的粒子，我们之所以可以像经典质点一样研究其运动轨迹就是因为不确定性小于分辨能力。

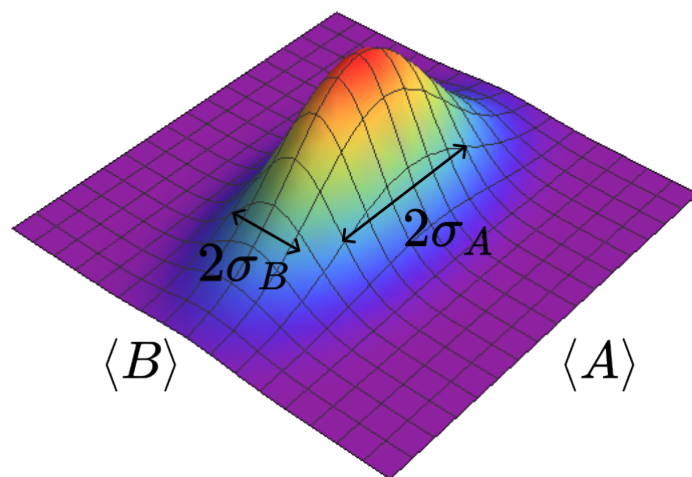


图 2.7. 不确定性关系示意图。不确定性关系禁止我们将图中曲面缩小为一个只占据一点的尖峰。