

第二章 量子物理的基本概念与框架

第一节 波函数与薛定谔方程

一、波函数

1. 波函数与量子状态

通过上一章的学习，我们知道量子力学的核心思想是“波粒二象性”，即物质的运动应该用波来描述。考虑最简单的量子体系，即只有一个粒子的体系。一般情况下，我们用一个复函数来描述这个粒子的物质波的复振幅，记作 $\psi(x)$ ，称为该粒子的波函数。

波函数在空间某一点的取值，正比于“粒子在该点出现”这个事件的**几率幅**（回顾量子物理基本假设 1，所有事件发生的不确定性都由几率幅描述）。波函数在空间中某一点的模平方 $|\psi(x)|^2$ 正比于在这一点找到粒子的几率（回顾量子物理基本假设 1，事件发生的几率为几率幅的模平方）。所以物质波是一种几率波。后面我们会详细介绍，知道了体系的波函数，就能得到体系的各种性质，所以波函数描述了体系的量子状态。

由于粒子总会存在于空间中的某处，所以如果我们对它在空间各点出现的几率求和，结果一定是 1。这告诉我们，粒子在不同位置处出现的几率只取决于 $|\psi(x)|^2$ 的相对大小。换句话说，如果我们把波函数 $\psi(x)$ 乘以一个复常数 c ，虽然 $|\psi(x)|^2$ 变成了 $|c|^2|\psi(x)|^2$ ，但是波函数模平方在不同位置处的相对大小比例并没有发生变化，所以粒子在不同位置处出现的几率也没有发生变化。这告诉我们， $\psi(x)$ 和 $c\psi(x)$ 其实对应的是同一个量子状态。

总结一下：我们用波函数 $\psi(x)$ 描述粒子的量子状态，并且波函数 $\psi(x)$ 与 $c\psi(x)$ 描述的是同一个量子状态，其中 c 可以是任意的复常数。粒子在空间一点出现的几率正比于 $|\psi(x)|^2$ ，或者说，粒子在空间一点 x 附近出现的**绝对**几率密度为

$$\rho(x) = |\psi(x)|^2 / \int |\psi(x)|^2 dx \quad (1)$$

根据(1)式，我们有 $\int \rho(x) dx = 1$ ，这正是绝对几率密度所需要满足的要求。

以上我们只是考虑了一个粒子的体系，并且粒子运动的维度是一维。对于三维情况，类似的，波函数记为 $\psi(\mathbf{r})$ ，其中 \mathbf{r} 是粒子的位置矢量；对于多粒子体系，波函数写成 $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ 的形式，其中 \mathbf{r}_n 是第 n 个粒子的位置矢量， N 是体系的粒子个数。

2. 波函数的归一化

利用波函数 $\psi(x)$ 与 $c\psi(x)$ 描述的是同一个量子状态这一规则，我们可以对波函数进行归一化操作。具体来说，对于任意给定的波函数 $\psi(x)$ ，我们可以定义一个新的波函数

$$\phi(x) = \psi(x) / \sqrt{\int |\psi(x)|^2 dx}, \quad (2)$$

由于 $1/\sqrt{\int |\psi(x)|^2 dx}$ 是一个常数，这个新的波函数 $\phi(x)$ 与原来的波函数 $\psi(x)$ 对应于同一个量子状态。但是这个新的波函数有一个很好的性质，即它满足

$$\int |\phi(x)|^2 dx = 1, \quad (3)$$

这告诉我们， $|\phi(x)|^2$ 就是粒子出现在 x 位置的绝对几率密度。满足这个性质的波函数称为归一化波函数，而利用(2)式将波函数 $\psi(x)$ 变换为归一化波函数 $\phi(x)$ 的操作称为波函数的归

一化，而(3)式称为归一化条件。

事实上，(2)式并不是唯一的归一化方法，例如我们可以定义

$$\phi(x) = e^{i\theta} \psi(x) / \sqrt{\int |\psi(x)|^2 dx}, \quad (4)$$

其中 θ 为任意的实数，这样得到的 $\phi(x)$ 依然满足归一化条件。

3. 特殊的波函数

一般来说，一个波函数要能描述实际粒子的量子状态，它需要满足连续、单值、有限这三个条件。这样的波函数一般都是可以利用(2)式的方法进行归一化的。但是，存在一些特殊的波函数，它们不能进行这样的操作。一个典型的例子是平面波 $\psi(x) = e^{ikx}$ （注意这里我们没有考虑和时间相关的因子 $e^{-i\omega t}$ ，因为目前为止我们只考虑某一个确定的瞬间，粒子的状态）。如果我们把它代入(2)式，我们发现等号右边的分母为无穷大，从而 $\phi(x)$ 恒等于零。这个异常现象的出现是因为严格来说，平面波这类形式的波函数不能代表真实粒子的量子状态：平面波要求波函数扩展到无穷远处并且不发生衰减，这意味着在 $x=0$ 点和 $x=\infty$ 找到粒子的几率式相同的，这与实际情况不符，因为实际的粒子总是限定在一定的空间范围内运动的。虽然这类波函数不能代表实际粒子的量子状态，但是在量子力学中，我们依然大量使用它们，这是因为：第一，这类波函数可以看作实际波函数的一种极限或者说理想情况，例如我们可以限定粒子的运动范围为有限的区间 $-w < x < w$ ，并且得到粒子的波函数，这个波函数是可以代表粒子的真实状态并且可以通过(2)式归一化的，然后我们让 $w \rightarrow \infty$ ，就可以得到平面波；第二，这类波函数在我们处理实际问题时有很大的用处，我们将在后面逐渐体会到这一点。这与信号与系统中，我们采用傅里叶变换去分析信号一样，虽然一个形如 $e^{-i\omega t}$ 的函数不可能代表一个真实信号，但是我们利用这样的函数去对实际的信号展开，可以得到很多有意义的结论。对于这类波函数，我们可以采用另外一种方式归一化，我们后面再详细展开。

4. 态叠加原理

由于波函数表示了体系的量子状态，我们后面常常将波函数与量子状态（即量子态、状态、态）混用。但是同学们需要记住的是，波函数与量子态之间不是一对一的关系，而是多对一的关系（ $\psi(x)$ 和 $c\psi(x)$ 对应同一个量子态）。

量子力学要求物质具有波粒二象性的原因是如果物质可以看作一种波，那么就可以类似惠更斯原理，通过波的干涉（叠加）去给出最小作用原理一个逻辑自洽的解释。波的可叠加性要求波函数满足如下的叠加原理：如果波函数 $\psi_1(x)$ 和 $\psi_2(x)$ 对应的量子态是粒子可能的状态，那么对于任意复数 c_1 和 c_2 ， $\psi(x) = c_1\psi_1(x) + c_2\psi_2(x)$ 也同样对应粒子的可能状态，叫做叠加态。

与经典波不同的是，处于叠加态 $\psi(x)$ 的粒子，我们可以认为它既处于状态 $\psi_1(x)$ 中，又处于状态 $\psi_2(x)$ 中，并且几率分别是 $|c_1|^2$ 和 $|c_2|^2$ （几率幅分别是 c_1 和 c_2 ）（注意这种既存在于一种状态，又存在于另外一种独立的状态中的存在形式是经典物理中所不可能发生的事情）。

我们可以用第一章中引入的量子物理的基本假设去理解这一性质。我们认为粒子以某种方式存在是一个事件，如果这个事件可以由两种方式去完成，一种方式是粒子处于状态 $\psi_1(x)$ ，对应的几率幅是 c_1 ，另外一种方式是粒子处于状态 $\psi_2(x)$ ，对应的几率幅是 c_2 ；那么粒子出

现在空间中某一点 x 这件事，也将以两种方式发生：第一种方式是粒子在处于状态 $\psi_1(x)$ 的前提下，出现在 x 点，对应的几率幅是 $c_1\psi_1(x)$ （对应的几率是 $|c_1|^2|\psi_1(x)|^2$ ），另一种方式是粒子在处于状态 $\psi_2(x)$ 的前提下，出现在 x 点，对应的几率幅是 $c_2\psi_2(x)$ （对应的几率是 $|c_2|^2|\psi_2(x)|^2$ ）。从而我们知道，粒子出现在空间中某一点 x 这件事的几率幅 $\psi(x)$ 是这两种完成方式的几率幅的直接叠加，即 $\psi(x) = c_1\psi_1(x) + c_2\psi_2(x)$ 。

态叠加原理可以推广到多个态的情况：如果波函数 $\psi_1(x)$ 、 $\psi_2(x)$ 、... 对应的量子态是粒子可能的状态，那么对于任意复数 c_1 、 c_2 、...， $\psi(x) = c_1\psi_1(x) + c_2\psi_2(x) + \dots$ 也同样对应粒子的可能状态，叫做叠加态。反之，如果粒子的一个态 $\psi(x)$ 可以写成 $\psi(x) = c_1\psi_1(x) + c_2\psi_2(x) + \dots$ 的形式，我们可以说它处于 $\psi_1(x)$ 、 $\psi_2(x)$ 、... 的叠加态中，即既处于状态 $\psi_1(x)$ 中，又处于状态 $\psi_2(x)$ 中，... 并且几率分别是 $|c_1|^2$ 、 $|c_2|^2$ 、...（几率幅分别是 c_1 、 c_2 、...），这种把波函数 $\psi(x)$ 写成 $\psi_1(x)$ 、 $\psi_2(x)$ 、... 的线性叠加的形式的过程叫做波函数的分解（展开）。

5. 平面波展开与傅里叶分析

一类重要的展开是平面波展开。具体来说，我们知道平面波具有 e^{ikx} 的波函数形式，同时由德布罗意关系，我们知道波函数为平面波的粒子具有确定的动量 $p = \hbar k$ ，所以我们可以把平面波写作 $e^{\frac{i}{\hbar}px}$ 。不同的平面波具有不同的动量 p ，对应的波函数记作 $\psi_p(x)$ 。为了方便计算展开系数，我们要求不同的平面波之间满足下述归一化关系：

$$\int_{-\infty}^{+\infty} \psi_{p'}^*(x) \psi_p(x) dx = \delta(p - p') \quad (5)$$

这要求

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}px} \quad (6)$$

根据傅里叶分析的知识，我们知道 $\{\psi_p(x)\}$ 构成了一组正交归一完备的函数基矢，从而所有的波函数 $\psi(x)$ 都可以用 $\{\psi_p(x)\}$ 去展开，即

$$\psi(x) = \int_{-\infty}^{+\infty} c(p) \psi_p(x) dp \quad (7)$$

为了计算展开系数 $c(p)$ ，我们将上式的左右两边同时乘以 $\psi_{p'}^*(x)$ 并且对 x 积分，利用(5)式的归一化关系得到

$$\int \psi_{p'}^*(x) \psi(x) dx = \iint c(p) \psi_{p'}^*(x) \psi_p(x) dx dp = \int c(p) \delta(p - p') dp = c(p')$$

即

$$c(p) = \int_{-\infty}^{+\infty} \psi_p^*(x) \psi(x) dx \quad (8)$$

(8)式本质上就是傅里叶变换，而(7)式本质上就是相应的傅里叶逆变换。

对于三维情况，平面波定义为

$$\psi_p(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} \quad (9)$$

其中 $\frac{1}{(2\pi\hbar)^{3/2}}$ 是归一化因子，从而 $\{\psi_p(\mathbf{r})\}$ 满足归一化条件

$$\int \psi_{p'}^*(\mathbf{r}) \psi_p(\mathbf{r}) d^3r = \delta(\mathbf{p} - \mathbf{p}') \quad (10)$$

类似于一维情况，三维情况下波函数的平面波展开为

$$\psi(\mathbf{r}) = \int c(\mathbf{p}) \psi_{\mathbf{p}}(\mathbf{r}) d^3p \quad (11)$$

并且展开系数为

$$c(\mathbf{p}) = \int \psi_{\mathbf{p}}^*(\mathbf{r}) \psi(\mathbf{r}) d^3r \quad (12)$$

二、薛定谔方程

1. 薛定谔方程的形式

目前为止，我们对粒子状态的描述都是只考虑了某一个特定的时刻。然而我们知道实际上粒子是在不断运动的，或者说它的波函数是在不断随时间演化的，所以一个粒子的完整波函数应该既是位置的函数，也是时间的函数，即 $\psi(x, t)$ 。理解一个量子体系的运动方式的关键是理解波函数的演化方式，或者说如果已知了某个时刻的波函数 $\psi(x, t_0)$ ，如何知道下一时刻的波函数 $\psi(x, t_0 + \delta t)$ 。描述这一演化的方程叫做薛定谔方程：

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) + V(x) \psi(x, t) \quad (13)$$

其中 $V(x)$ 是粒子感受到的势能，我们这里暂时只考虑 V 与时间无关的情况。不同体系的不同之处体现在 $V(x)$ 的形式不同。

我们注意到薛定谔方程是关于时间的一阶微分方程，这告诉我们只要知道了一个初始条件，即 $\psi(x, t_0)$ ，我们就可以推导出后续任意时刻的波函数（当然也可以推导出之前任意时刻的波函数）。所谓的解薛定谔方程，就是当我们在已知 $\psi(x, t_0)$ 的前提下，利用(13)式得出任意时刻的 $\psi(x, t)$ 。

对于三维的情况，薛定谔方程的形式为

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}, t) + V(\mathbf{r}) \psi(\mathbf{r}, t) \quad (14)$$

对于多例子体系，假如粒子两两之间具有势能 $V_{mn}(\mathbf{r}_m, \mathbf{r}_n)$ ，那么该体系的薛定谔方程的形式为

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) = -\sum_n \frac{\hbar^2}{2m_n} \nabla_n^2 \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) + \frac{1}{2} \sum_{m,n} V_{mn}(\mathbf{r}_m, \mathbf{r}_n) \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) \quad (15)$$

其中 m_n 为第 n 个粒子的质量。

2. 平面波

我们在第一章介绍德布罗意关系的时候，引入了平面波 $e^{i(kx - \omega t)}$ 的概念，这里我们来证明，对于自由粒子来说，平面波确实是薛定谔方程的解。所谓自由粒子，就是粒子没有感受到任何外场的作用，所以该体系的薛定谔方程为

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) \quad (16)$$

我们把 $\psi(x, t) = e^{i(kx - \omega t)}$ 代入上式，得到如下关系

$$\hbar\omega = \frac{\hbar^2 k^2}{2m} \quad (17)$$

这说明并不是所有的平面波都是自由粒子薛定谔方程的解，而是只有波矢 k 和圆频率 ω 之间满足关系(17)的平面波才是(16)式的解。(17)式称为平面波的色散关系，在第三章中我们还会再次遇到并且利用这一关系。根据德布罗意关系，我们知道 $\hbar\omega = E$ ， $\hbar k = p$ ，所以(17)式可以写为

$$E = \frac{p^2}{2m} \quad (18)$$

而这正是经典力学所要求的自由粒子的能量与动量之间的依赖关系。

3. 分离变量法与薛定谔方程的求解

对于非自由粒子，薛定谔方程的解不会像平面波这么简单。为了解薛定谔方程(13)，我们首先假设波函数可以写作如下形式

$$\psi(x, t) = X(x)T(t) \quad (19)$$

其中 $X(x)$ 和 $T(t)$ 分别是关于 x 和 t 的一元函数。代入(13)，我们有

$$i\hbar X(x)T'(t) = \left\{ \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] X(x) \right\} T(t)$$

在等式的左右两边同时除以 $\psi(x, t)$ ，我们有

$$i\hbar \frac{T'(t)}{T(t)} = \frac{\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] X(x)}{X(x)}$$

由于这个等式的左边是 t 的一元函数，右边是 x 的一元函数，左右两边相等的唯一可能是它们同时等于一个既不依赖于 t 也不依赖于 x 的待定常数，我们记这个常数为 E ，那么

$$i\hbar \frac{T'(t)}{T(t)} = \frac{\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] X(x)}{X(x)} = E \quad (20)$$

所以我们将薛定谔方程转化为了两个独立的常微分方程

$$i\hbar T'(t) = ET(t) \quad (21)$$

和

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] X(x) = EX(x) \quad (22)$$

(21)可以直接积分求解，得到

$$T(t) = T(0)e^{-\frac{i}{\hbar}Et} \quad (23)$$

而(22)的求解依赖于 $V(x)$ 的具体形式。(22)式称为定态薛定谔方程，我们后面的课程会用三章的篇幅对不同的 $V(x)$ 形式，分别求解定态薛定谔方程。这里我们先不具体求解该方程，而是直接引用数理方程中的“斯图姆-刘维尔定理”，这个定理告诉我们，对于我们关心的几乎任意情况来说，(22)式的全部解 $\{X_E(x)\}$ 都构成一组正交完备的函数基矢（注意这里 $X_E(x)$ 加入了下标 E ，是因为方程(22)的解依赖于 E 的取值，而 $\{X_E(x)\}$ 的含义是对于所有可能的 E 取值，我们都去求解方程(22)，如果解是存在的，我们就把它放到集合 $\{X_E(x)\}$ 中去）。这一结论使得我们可以构建一个解薛定谔方程(13)的一般步骤：

步骤一：我们首先解与薛定谔方程(13)对应的定态薛定谔方程(22)，得到正交完备函数基矢 $\{X_E(x)\}$ 。

步骤二：我们把已知的初始波函数对 $\{X_E(x)\}$ 展开，即

$$\psi(x, t_0) = \sum_E c_E^{(0)} X_E(x) \quad (24)$$

其中 $c_E^{(0)}$ 是展开系数。由于 $\{X_E(x)\}$ 的完备性，这一展开总是成立的。

步骤三：我们把任意时刻的波函数也对 $\{X_E(x)\}$ 展开，即

$$\psi(x, t) = \sum_E c_E(t) X_E(x) \quad (25)$$

由于 $\psi(x, t)$ 是一个二元函数，展开系数 $c_E(t)$ 是 t 的一元函数。我们把(25)代入(13)，得到

$$\sum_E [i\hbar \frac{d}{dt} c_E(t)] X_E(x) = \sum_E c_E(t) \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] X_E(x) = \sum_E E c_E(t) X_E(x) \quad (26)$$

其中第二个等号用到了(22)式。由于 $\{X_E(x)\}$ 的正交性，我们有

$$i\hbar \frac{d}{dt} c_E(t) = E c_E(t) \quad (27)$$

这就是方程(21)，我们利用初始条件 $c_E(t_0) = c_E^{(0)}$ ，解出

$$c_E(t) = c_E^{(0)} e^{-\frac{i}{\hbar} E(t-t_0)} \quad (28)$$

带回到(25)式中，我们有

$$\psi(x, t) = \sum_E c_E^{(0)} e^{-\frac{i}{\hbar} E(t-t_0)} X_E(x) \quad (29)$$

这就是薛定谔方程(13)的解。

步骤二和步骤三对于所有的薛定谔方程都是相同的，只是对初始波函数展开，然后在每一个展开项的系数上乘以一个 $e^{-\frac{i}{\hbar} E(t-t_0)}$ 因子。然而步骤一，定态薛定谔方程的解确是依赖于 $V(x)$ 的具体形式的，所以在后面的课程中（最后一章除外），我们不再去解完整的薛定谔方程(13)（也称为含时薛定谔方程），而是只去求解定态薛定谔方程(22)，因为只要解出了定态薛定谔方程，就可以按照上述标准步骤，得到含时薛定谔方程的解。

4. 几率守恒

由于薛定谔方程是告诉我们波函数怎么从一个时刻演化到下一时刻，一个直接的问题就是，如果一个波函数在起始时刻是满足归一化条件(3)的，按照薛定谔方程计算出来的后续时刻的波函数，还能不能满足归一化条件？答案是肯定的。

我们对薛定谔方程(13)取复共轭，有

$$-i\hbar \frac{\partial}{\partial t} \psi^*(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi^*(x, t) + V(x) \psi^*(x, t) \quad (30)$$

结合(13)和(30)式，我们有

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} w &= i\hbar \frac{\partial}{\partial t} |\psi|^2 = \psi^* \left(i\hbar \frac{\partial}{\partial t} \psi \right) - \psi \left(-i\hbar \frac{\partial}{\partial t} \psi^* \right) \\ &= -\frac{\hbar^2}{2m} (\psi^* \frac{\partial^2}{\partial x^2} \psi - \psi \frac{\partial^2}{\partial x^2} \psi^*) = \frac{\hbar^2}{2m} \frac{\partial}{\partial x} (\psi \frac{\partial}{\partial x} \psi^* - \psi^* \frac{\partial}{\partial x} \psi) \end{aligned} \quad (31)$$

其中 $w = |\psi|^2$ ，为相对几率密度，我们定义

$$j = \frac{i\hbar}{2m} (\psi \frac{\partial}{\partial x} \psi^* - \psi^* \frac{\partial}{\partial x} \psi)$$

为相对几率流密度，那么(31)可以写成

$$\frac{\partial}{\partial t} w + \frac{\partial}{\partial x} j = 0 \quad (32)$$

这代表了几率密度的守恒：空间某一处的几率密度变化，一定是要由它周边位置几率密度的

相反变化所补偿，而这种此消彼长的变换对应于几率密度的流动。

对于三维情况，我们有

$$\mathbf{j} = \frac{i\hbar}{2m} (\psi \nabla \psi^* - \psi^* \nabla \psi) \quad (33)$$

和

$$\frac{\partial}{\partial t} w + \nabla \cdot \mathbf{j} = 0 \quad (34)$$

将 w 对全空间积分，并且利用(34)式，我们有

$$\frac{\partial}{\partial t} \int_V w dv = \int_V \frac{\partial}{\partial t} w dv = - \int_V \nabla \cdot \mathbf{j} dv = - \oint_{\partial V} \mathbf{j} \cdot d\mathbf{s} \quad (35)$$

其中 V 代表了全空间，而 ∂V 代表了 V 的边界，可以想象为一个存在于无穷远处的闭合曲面，包围了整个空间。由于我们一般要求波函数在无穷远处衰减为零（即实际粒子不可能运动到无穷远处），根据(33)式， \mathbf{j} 在无穷远处也为 0，所以(35)式的积分为 0，即相对几率密度在全空间的积分 $\int_V w dv$ 不随时间变化。这证明了薛定谔方程保持归一化的特性：如果初始时刻

的波函数满足归一化条件 $\int_V w dv = 1$ ，那么通过薛定谔方程计算出的其余时刻的波函数也自动满足。

5. 薛定谔方程与最小作用原理

回顾第一章的内容，我们知道在经典力学中，牛顿第二定律等价于最小作用原理，而最小作用原理存在一个概念上的问题——质点如何提前知道哪条路径具有最小的作用量。为了解决这个问题，我们认为任何物质都具有波粒二象性，从而通过波的叠加来解释最小作用原理：对于宏观质点来说，给定运动曲线的起点 A 和终点 B ，它实际上是通过波传播的方式走过了所有可能的路径，只不过由于波的叠加特性（费曼路径积分），只有满足最小作用原理的路径才会与其周围的路径形成相长干涉（图 2-1）。

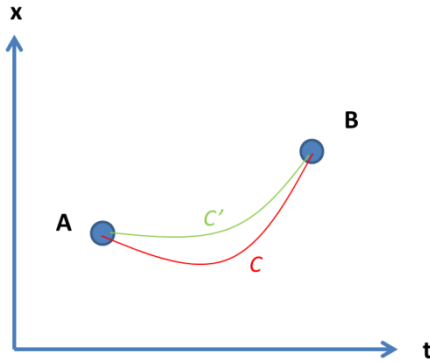


图 2-1 最小作用原理示意图

这一原理由第一章(14)式描述，即

$$u_B \sim \sum_C e^{\frac{i}{\hbar} S[C]}, \quad (36)$$

这里我们说明从这一原理出发，可以导出薛定谔方程。

我们从图 2-1 出发，已知粒子运动的起点 A 和终点 B，考虑其运动其实是物质波的传播，也即波函数随时间的演化。如图 2-2 所示，我们把 AB 之间的时间轴平均分割成一个个离散的时刻 t_1 、 t_2 、...、 t_N ，并且假设相邻的时刻相差一个趋于零的 Δt 。从一个时刻 t_n 到下一时刻 t_{n+1} ，波函数传播了 Δt 的时刻，从 $\psi(x, t_n)$ 演化为 $\psi(x, t_{n+1})$ 。根据惠更斯原理， t_n 时刻空间中任意一点的波函数都应该作为一个次级波源，将“振动”传播到下一时刻 t_{n+1} 空间中的所有点；而反之， t_{n+1} 时刻空间中任意一点的波函数，都是由 t_n 时刻空间中所有点的“振动”传播过来并且线性叠加的结果。用数学语言表达这一过程，即

$$\psi(x_{n+1}, t_{n+1}) = \int K(x_{n+1}, x_n, \Delta t) \psi(x_n, t_n) dx_n, \quad (37)$$

其中 $K(x_{n+1}, x_n, \Delta t)$ 代表了 t_n 时刻波函数在 x_n 处的复振幅对 t_{n+1} 时刻波函数在 x_{n+1} 处复振幅的贡献。直观上看，从每一个时刻到下一个时刻，相当于经历了一次多缝干涉，而波函数的整个演化过程，相当于无数个这样的多缝干涉级联的效果（图 2-2）。只不过这个干涉是发生在 t - x 平面上，而不是像之前我们介绍的电子双缝干涉，发生在 x - y 平面内。

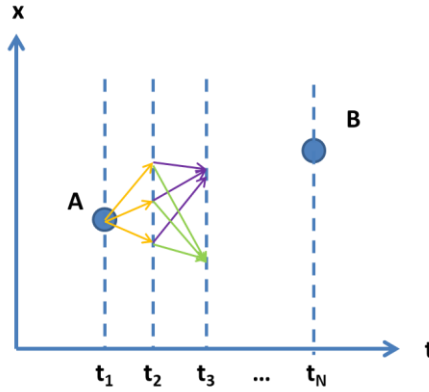


图 2-2 t - x 平面内的惠更斯原理

由于我们知道了运动曲线的起点为 A，即已知了 t_1 时刻粒子位于 x_1 点，或者说

$$\psi(x, t_1) = \delta(x - x_1), \quad (38)$$

我们可以把这一初值条件代入(37)式中，逐步推演出 t_2 、 t_3 、...、 t_N 时刻的波函数：

$$\begin{aligned} \psi(x_2, t_2) &= \int K(x_2, x, \Delta t) \psi(x, t_1) dx = K(x_2, x_1, \Delta t) \\ \psi(x_3, t_3) &= \int K(x_3, x_2, \Delta t) \psi(x_2, t_2) dx_2 = \int K(x_3, x_2, \Delta t) K(x_2, x_1, \Delta t) dx_2 \\ &\vdots \\ \psi(x_N, t_N) &= \int K(x_N, x_{N-1}, \Delta t) \dots K(x_4, x_3, \Delta t) K(x_3, x_2, \Delta t) K(x_2, x_1, \Delta t) dx_{N-1} \dots dx_3 dx_2, \end{aligned} \quad (39)$$

注意到(39)式对所有的中间时刻的位置 x_2 、 x_3 、...、 x_{N-1} 积分，相当于在固定初始时刻位置为 x_1 、终止时刻位置为 x_N 的前提下，遍历了 t - x 平面内所有可能的路径，这正是(36)式的求和所要求的。所以对比(39)式和(36)式，我们有

$$e^{\frac{i}{\hbar} S[C]} \sim K(x_N, x_{N-1}, \Delta t) \dots K(x_4, x_3, \Delta t) K(x_3, x_2, \Delta t) K(x_2, x_1, \Delta t), \quad (40)$$

其中路径 C 由点 (t_1, x_1) 、 (t_2, x_2) 、...、 (t_N, x_N) 连接而成。而我们根据 $S[C]$ 的定义，有

$$e^{\frac{i}{\hbar} S[C]} = e^{\frac{i}{\hbar} \int_{t_1}^{t_N} L dt} = e^{\frac{i}{\hbar} \sum_{n=1}^{N-1} L(x_n, \dot{x}_n) \Delta t} = \prod_{n=1}^{N-1} e^{\frac{i}{\hbar} L(x_n, \dot{x}_n) \Delta t}, \quad (41)$$

从而

$$K(x_{n+1}, x_n, \Delta t) \sim e^{\frac{i}{\hbar} L(x_n, \dot{x}_n) \Delta t} = e^{-\frac{i}{\hbar} V(x_n) \Delta t} e^{\frac{i}{\hbar} (\frac{1}{2} m \frac{\Delta x^2}{\Delta t})}, \quad (42)$$

其中 $\Delta x = x_{n+1} - x_n$ ，并且由于 Δt 很小，我们假设了从 (t_n, x_n) 到 (t_{n+1}, x_{n+1}) 的这一小段运动曲线是直线段。把上式代入(37)式，有

$$\psi(x_{n+1}, t_{n+1}) \sim \int e^{-\frac{i}{\hbar} V(x_n) \Delta t} e^{\frac{i}{\hbar} (\frac{1}{2} m \frac{\Delta x^2}{\Delta t})} \psi(x_n, t_n) dx_n,$$

由于当 Δt 很小时， $e^{\frac{i}{\hbar} (\frac{1}{2} m \frac{\Delta x^2}{\Delta t})}$ 只有当 Δx 才能避免相位因子的快速震荡从而抵消，所以上式的积分实际上是在 x_{n+1} 附近的一个小区间对 x_n 积分，所以我们可以认为 $V(x_n)$ 是一个常数从而提到积分号外面，并且将积分变量等价地写成 Δx ，即

$$\psi(x_{n+1}, t_{n+1}) \sim e^{-\frac{i}{\hbar} V(x_{n+1}) \Delta t} \int e^{\frac{i}{\hbar} (\frac{1}{2} m \frac{\Delta x^2}{\Delta t})} \psi(x_{n+1} - \Delta x, t_n) d\Delta x, \quad (43)$$

由于 Δx 很小，我们可以对 $\psi(x_{n+1} - \Delta x, t_n)$ 泰勒展开（保留到二阶），即

$$\psi(x_{n+1} - \Delta x, t_n) = \psi(x_{n+1}, t_n) - \psi'(x_{n+1}, t_n) \Delta x + \frac{1}{2} \psi''(x_{n+1}, t_n) (\Delta x)^2,$$

并代入到(43)的积分中，从而得到

$$\begin{aligned} \psi(x_{n+1}, t_{n+1}) &\sim \left[e^{-\frac{i}{\hbar} V(x_{n+1}) \Delta t} \int e^{\frac{i}{\hbar} (\frac{1}{2} m \frac{\Delta x^2}{\Delta t})} d\Delta x \right] \psi(x_{n+1}, t_n) \\ &\quad + \frac{1}{2} \left[e^{-\frac{i}{\hbar} V(x_{n+1}) \Delta t} \int (\Delta x)^2 e^{\frac{i}{\hbar} (\frac{1}{2} m \frac{\Delta x^2}{\Delta t})} d\Delta x \right] \psi''(x_{n+1}, t_n) \\ &\sim \sqrt{\frac{2i\pi\hbar\Delta t}{m}} e^{-\frac{i}{\hbar} V(x_{n+1}) \Delta t} \left[1 + \frac{i\Delta t}{\hbar} \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right] \psi(x_{n+1}, t_n), \\ &\sim \sqrt{\frac{2i\pi\hbar\Delta t}{m}} \left[1 - \frac{i\Delta t}{\hbar} V(x_{n+1}) + \frac{i\Delta t}{\hbar} \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right] \psi(x_{n+1}, t_n), \end{aligned}$$

由于当 $\Delta t = 0$ 时， $\psi(x_{n+1}, t_{n+1})$ 应等于 $\psi(x_{n+1}, t_n)$ ，上式右边应该除以因子 $\sqrt{\frac{2i\pi\hbar\Delta t}{m}}$ 从而确保这一关系，即

$$\psi(x_{n+1}, t_{n+1}) = \left[1 - \frac{i\Delta t}{\hbar} V(x_{n+1}) + \frac{i\Delta t}{\hbar} \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right] \psi(x_{n+1}, t_n), \quad (44)$$

当 $\Delta t \rightarrow 0$ 时，这就是薛定谔方程。

所以从这个角度来说，薛定谔方程是我们引入物质波去解释最小作用原理的必然要求。回顾第一章的内容，我们现在可以得出这样一个表格，它反映了我们之前学过的经典力学、经典光学和我们这里学习的量子力学之间的内在逻辑联系。

	光学		力学	
局域理论	折射反射定律	几何光学	牛顿第二定律	经典力学
整体理论	费马原理		最小作用原理	
整体理论	惠更斯原理	波动光学	费曼路径积分	量子力学 (波动力学)
局域理论	麦克斯韦方程组		薛定谔方程	

第二节 力学量与算符

一、力学量

物理是一门实验科学，理论要能够对实验进行描述和预测，所以量子物理需要对于实验中观测的力学量进行描述。我们通过第一节，告诉大家波函数描述了系统完整的量子状态，那么这意味着，如果我们知道了粒子的波函数，应该也可以从波函数得到其各种力学量的取值。我们下面对于常见的力学量分别举例说明。

1. 位置

粒子的位置是最常见的力学量之一，对于一维情况，我们用 x 表示它。

位置的取值蕴含在波函数的定义中：假设一个粒子的量子状态由波函数 $\psi(x)$ 表示，那么粒子出现在 x 处的相对几率密度为 $|\psi(x)|^2$ ；如果波函数 $\psi(x)$ 可以归一化，那么粒子出现在 x 处的绝对几率密度为 $\rho(x) = |\psi(x)|^2 / \int |\psi(x)|^2 dx$ （即式(1)）；并且位置的期望值为

$$\langle x \rangle = \frac{\int \psi^*(x) x \psi(x) dx}{\int \psi^*(x) \psi(x) dx} \quad (45)$$

【补充说明】

但是需要注意的是，正如我们在上一节中提到的，存在一些特殊的波函数，它们是实际波函数的极限情况，对于这些波函数，并不能用通常的方式归一化。这里我们以 δ 函数为例进行说明。如果一个粒子的波函数是 δ 函数，即

$$\psi_{x_0}(x) = \delta(x - x_0), \quad (46)$$

那么根据 δ 函数的定义，这个 $\psi_{x_0}(x)$ 只在 x_0 取值为无穷大，在其余所有位置取值为 0。这意味着在此量子状态下，粒子具有确定的位置，即 x_0 。与平面波的归一化(5)式类似， δ 函数满足如下的归一化关系

$$\int_{-\infty}^{+\infty} \psi_{x_1}^*(x) \psi_{x_2}(x) dx = \delta(x_1 - x_2) \quad (47)$$

这使得我们可以方便地把任意波函数用 δ 函数展开：

$$\psi(x) = \int_{-\infty}^{+\infty} \psi(x_0) \psi_{x_0}(x) dx_0 \quad (48)$$

如果粒子的波函数为(46)式，那么根据波函数的定义，这个粒子位于 x_0 的几率为 1，并且其位置期望值一定也是 x_0 。但是如果我们利用(45)式，有

$$\langle x \rangle = \frac{\int \delta(x - x_0) x \delta(x - x_0) dx}{\int \delta(x - x_0) \delta(x - x_0) dx} = \frac{x_0 \int \delta(x - x_0) \delta(x - x_0) dx}{\int \delta(x - x_0) \delta(x - x_0) dx} \quad (49)$$

从而分子分母都是无穷大。但是我们知道 δ 函数是实际波函数的理想极限，所以在取极限的意义下，我们可以用罗必塔法则，得到(49)取值为 x_0 。

对于位置的任意一元函数 $f(x)$ （例如势能函数 $V(x)$ ），我们可以类似的得到其期望值为

$$\langle f(x) \rangle = \frac{\int \psi^*(x) f(x) \psi(x) dx}{\int \psi^*(x) \psi(x) dx} \quad (50)$$

2. 动量

如果已知一个粒子的波函数为 $\psi(x)$ ，我们该如何得知它的动量取值？这需要借用上一节介绍的平面波展开，由于任意波函数一定可以写成形式(7)的叠加形式，同时根据态叠加原理，(7)式意味着粒子有 $|c(p)|^2$ 的相对几率密度处于状态 $\psi_p(x)$ ，又由于处于状态 $\psi_p(x)$ 的粒子具有确定的动量 p ，我们最终得知，波函数为 $\psi(x)$ 的粒子，具有动量 p 的相对几率密度为 $|c(p)|^2$ ，绝对几率密度为

$$\rho(p) = |c(p)|^2 / \int |c(p)|^2 dp, \quad (51)$$

并且动量的期望值为

$$\langle p \rangle = \frac{\int c^*(p) p c(p) dp}{\int c^*(p) c(p) dp} \quad (52)$$

对于位置的任意一元函数 $f(p)$ （例如动能 $E = \frac{p^2}{2m}$ ），我们可以类似的得到其期望值为

$$\langle f(p) \rangle = \frac{\int c^*(p) f(p) c(p) dp}{\int c^*(p) c(p) dp} \quad (53)$$

【补充说明】

与位置类似，对于一些特殊的波函数，(53)式也要在极限的意义下进行计算。这里我们以平面波为例进行说明。如果一个粒子的波函数是平面波

$$\psi_{p_0}(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p_0 x}, \quad (54)$$

那么这个粒子具有动量 p_0 的几率为1，并且其动量期望值一定也是 p_0 。但是如果我们利用(8)式，有

$$c(p) = \delta(p - p_0), \quad (55)$$

带入(52)式，有

$$\langle p \rangle = \frac{\int \delta(p - p_0) p \delta(p - p_0) dp}{\int \delta(p - p_0) \delta(p - p_0) dp} = \frac{p_0 \int \delta(p - p_0) \delta(p - p_0) dp}{\int \delta(p - p_0) \delta(p - p_0) dp} \quad (56)$$

从而分子分母都是无穷大。但是我们知道平面波是实际波函数的理想极限，所以在取极限的意义下，我们可以用罗必塔法则，得到(56)取值为 p_0 。

3. 位置与动量的二元函数

但是对于位置与动量的二元函数，例如角动量

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (57)$$

以及能量

$$E = \frac{p^2}{2m} + V(\mathbf{r}) \quad (58)$$

我们该如何从波函数得到这些力学量的取值呢？波函数在某一点 x 的取值，只告诉我们粒子位于这一点 x 的几率，并不会告诉我们粒子在位于这一点的前提下，具有动量 p 的几率，因为 $c(p)$ 的确定需要波函数在全空间的取值（式(8)）；同理，波函数的平面波展开系数 $c(p)$ ，只告诉我们粒子具有动量 p 的几率，却并没有告诉我们粒子位于位置 x 的几率，因为 $c(p)$ 的获得是对波函数在全空间的取值进行了积分得到的。为了处理这些力学量，我们需要引入算符的概念。

二、算符

1. 算符的定义与性质

算符(operator)可以理解为作用在函数上面的一种操作, 经过这个操作, 一个函数被变换为另外一个波函数。例如“把任意函数与实数 5 相乘”这个操作可以看作一个算符, “对任意函数求导”也可以看作一个算符, “把任意函数取平方”也可以看作一个算符。

我们用记号 \hat{F} 表示算符, 用 u 表示函数, 用 $\hat{F}u$ 表示 u 经过 \hat{F} 作用后, 得到的新的函数。我们可以把算符作为研究对象, 取定义它们之间的各种运算。

(1) 算符的相等: 如果两个算符对于任意函数作用的结果都是相同的, 那么我们称两个算符相等, 用数学符号来表示, 就是: 如果 $\forall u$, 都有 $\hat{F}_1 u = \hat{F}_2 u$, 那么我们称 $\hat{F}_1 = \hat{F}_2$ 。

(2) 两个算符的加法: 对于任意的两个算符 \hat{F}_1 和 \hat{F}_2 , 我们定义一个新的算符 \hat{F} 作为 \hat{F}_1 和 \hat{F}_2 的和, 这个算符对任意函数 u 的作用效果为 $\hat{F}u = \hat{F}_1 u + \hat{F}_2 u$, 或者我们把 \hat{F} 记为 $\hat{F}_1 + \hat{F}_2$, 并且定义 $(\hat{F}_1 + \hat{F}_2)u = \hat{F}_1 u + \hat{F}_2 u$ 。

可以很容易地验证, 算符的加法满足交换律和结合律。

(3) 单位算符我们把不做任何操作的操作称为单位算符 \hat{I} , 即 $\hat{I}u = u$ 。

(4) 两个算符的乘积: 对于任意的两个算符 \hat{F}_1 和 \hat{F}_2 , 我们定义一个新的算符 \hat{F} 作为 \hat{F}_1 和 \hat{F}_2 的积, 这个算符对任意函数 u 的作用效果为 $\hat{F}u = \hat{F}_1(\hat{F}_2 u)$, 即先对 u 进行 \hat{F}_2 操作, 再对得到的结果进行 \hat{F}_1 操作。或者我们把 \hat{F} 记为 $\hat{F}_1 \hat{F}_2$, 并且定义 $\hat{F}_1 \hat{F}_2 u = \hat{F}_1(\hat{F}_2 u)$ 。

任意算符 \hat{F} 与单位算符的乘法满足关系: $\hat{F}\hat{I} = \hat{I}\hat{F} = \hat{F}$

但是需要注意的是, 算符的乘法一般并不满足交换律, 即对于任意两个算符 \hat{F}_1 和 \hat{F}_2 , 一般来说 $\hat{F}_1 \hat{F}_2 \neq \hat{F}_2 \hat{F}_1$ 。

例如我们取算符 \hat{F}_1 为“把任意函数与实数 5 相乘”, \hat{F}_2 为“把任意函数取平方”, 那么 $\hat{F}_1 \hat{F}_2 u = 5u^2$, 而 $\hat{F}_2 \hat{F}_1 u = 25u^2$, 两者显然不相等。

由于这一现象, 我们可以进一步定义两个算符的对易算符和反对易算符:

对易算符: $[\hat{F}_1, \hat{F}_2] \equiv \hat{F}_1 \hat{F}_2 - \hat{F}_2 \hat{F}_1$

反对易算符: $\{\hat{F}_1, \hat{F}_2\} \equiv \hat{F}_1 \hat{F}_2 + \hat{F}_2 \hat{F}_1$

如果两个算符的对易算符为零, 我们称这两个算符对易, 例如任意算符与单位算符总是对易的; 如果两个算符的反对易算符为零, 我们称这两个算符反对易。

(5) 逆算符: 如果 $\hat{F}_1 \hat{F}_2 = \hat{I}$, 称 \hat{F}_1 和 \hat{F}_2 互为逆算符, 记 $\hat{F}_1^{-1} = \hat{F}_2$, $\hat{F}_2^{-1} = \hat{F}_1$, 且 $\hat{F}_2 \hat{F}_1 = \hat{I}$ 自动满足。

(6) 算符的本征值与本征函数：给定算符 \hat{F} ，如果存在一个数 f 和函数 $\psi_f(x)$ ，满足

$$\hat{F}\psi_f(x) = f\psi_f(x), \quad (59)$$

那么 f 称为算符 \hat{F} 的本征值， $\psi_f(x)$ 称为算符 \hat{F} 的本征函数，与本征值 f 相对应的本征函数。

2. 线性算符

如果一个算符 \hat{F} 满足条件：对任意的函数 u_1, u_2 和常数 c_1, c_2 ，都有

$$\hat{F}(c_1u_1 + c_2u_2) = c_1\hat{F}u_1 + c_2\hat{F}u_2 \quad (60)$$

那么我们称这个算符为线性算符。换句话说，线性算符能够保持矢量之间的线性叠加关系。

例如算符“把任意函数与实数 5 相乘”是线性算符，“对任意函数求导”也是线性算符，“把任意函数取平方”则不是线性算符。

3. 算符与力学量

量子力学中，对于任意的力学量，我们都可以找到某个线性算符与之对应。我们依然从位置和动量这两个基本力学量出发。

(1) 位置：与位置 x 对应的算符是“把任意函数与函数 $f(x) = x$ 相乘”，或者记作

$$\hat{x} = x, \quad (61)$$

称为位置算符。位置算符 \hat{x} 对任意波函数 $\psi(x)$ 的作用表达为

$$\hat{x}\psi(x) = x\psi(x), \quad (62)$$

另外我们注意到 δ 函数是位置算符的本征函数，因为

$$\hat{x}\delta(x - x_0) = x\delta(x - x_0) = x_0\delta(x - x_0), \quad (63)$$

同时 $\delta(x - x_0)$ 也是力学量 x 具有确定取值的状态，取值为相应的本征值 x_0 。

而力学量 x 的期望值(45)式显然也可以写作

$$\langle x \rangle = \frac{\int \psi^*(x)\hat{x}\psi(x)dx}{\int \psi^*(x)\psi(x)dx}. \quad (64)$$

对于三维情况，位置 \mathbf{r} 是一个矢量，我们可以分别考虑它的三个分量：

$$\hat{x} = x, \hat{y} = y, \hat{z} = z$$

或者直接写成

$$\hat{\mathbf{r}} = \mathbf{r} \quad (65)$$

它表示位置算符 $\hat{\mathbf{r}}$ 作用到任意波函数 $\psi(\mathbf{r})$ 上，将得到一个矢量函数 $\mathbf{r}\psi(\mathbf{r})$ 。

(2) 动量：与动量 p 对应的算符是“把任意函数对 x 求导，再乘以 $-i\hbar$ ”，或者记作

$$\hat{p} = -i\hbar \frac{\partial}{\partial x} \quad (66)$$

称为动量算符。动量算符 \hat{p} 对任意波函数 $\psi(x)$ 的作用表达为

$$\hat{p}\psi(x) = -i\hbar \frac{\partial}{\partial x}\psi(x) \quad (67)$$

另外我们注意到平面波 $\psi_{p_0}(x)$ 是动量算符的本征函数，因为

$$\hat{p}\psi_{p_0}(x) = -i\hbar \frac{\partial}{\partial x} \left[\frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}p_0x} \right] = p_0\psi_{p_0}(x) \quad (68)$$

同时平面波 $\psi_{p_0}(x)$ 也是力学量 p 具有确定取值的状态，取值为相应的本征值 p_0 。

利用傅里叶变换关系(7)和(8)，很容易证明(53)式也可以写为

$$\langle p \rangle = \frac{\int \psi^*(x) \hat{p} \psi(x) dx}{\int \psi^*(x) \psi(x) dx} = \frac{\int \psi^*(x) (-i\hbar \frac{\partial}{\partial x}) \psi(x) dx}{\int \psi^*(x) \psi(x) dx}. \quad (69)$$

对于三维情况，动量 \mathbf{p} 是一个矢量，我们可以分别考虑它的三个分量：

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}, \hat{p}_y = -i\hbar \frac{\partial}{\partial y}, \hat{p}_z = -i\hbar \frac{\partial}{\partial z}$$

或者直接写成

$$\hat{\mathbf{p}} = -i\hbar \nabla \quad (70)$$

它表示位置算符 $\hat{\mathbf{p}}$ 作用到任意波函数 $\psi(\mathbf{r})$ 上，将得到一个矢量函数 $-i\hbar \nabla \psi(\mathbf{r})$ 。

(3) 一般的力学量

由于一般的力学量 F 总可以写成 x 和 p 的函数 $f(x, p)$ ，我们可以把通过 x 和 p 换成算符，得到相应的该力学量的算符形式 $F(\hat{x}, \hat{p})$ （注意由于 \hat{x} 和 \hat{p} 不对易，这个形式不一定是直接代换），然后找到这个算符 \hat{F} 的所有本征值 $\{f\}$ 和相应的本征函数 $\{\psi_f(x)\}$ （(59)式），这些本征函数 $\psi_f(x)$ 代表了粒子的力学量 F 取确定值 f 的状态。

我们有如下结论：对于任意的波函数 $\psi(x)$ ，我们把它用这些本征函数 $\{\psi_f(x)\}$ 展开，

$$\psi(x) = \int_{-\infty}^{+\infty} c(f) \psi_f(x) df \quad (71)$$

展开系数的模平方 $|c(f)|^2$ 就对应于力学量 F 取 f 的相对几率密度，对所有的取值进行加权平均，我们就得到了力学量 F 的期望值

$$\langle f \rangle = \frac{\int c^*(f) f c(f) df}{\int c^*(f) c(f) df} \quad (72)$$

在后面两节中，我们会详细论述这一点。

4. 算符的对易关系

我们来计算一下位置 and 动量的对易算符： $[\hat{x}, \hat{p}]$ ，考虑它对任意波函数 $\psi(x)$ 的作用：

$$\begin{aligned} [\hat{x}, \hat{p}] \psi(x) &= \hat{x} \hat{p} \psi(x) - \hat{p} \hat{x} \psi(x) = x \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(x) - \left(-i\hbar \frac{\partial}{\partial x} \right) [x \psi(x)] \\ &= i\hbar \psi(x) \end{aligned}$$

从而我们有

$$[\hat{x}, \hat{p}] = i\hbar \quad (73)$$

这是量子力学的基本对易关系。

对于三维的情况，通过类似的推导，我们有

$$\begin{aligned} [\hat{x}, \hat{p}_x] &= [\hat{y}, \hat{p}_y] = [\hat{z}, \hat{p}_z] = i\hbar \\ [\hat{x}, \hat{p}_y] &= [\hat{x}, \hat{p}_z] = [\hat{y}, \hat{p}_x] = [\hat{y}, \hat{p}_z] = [\hat{z}, \hat{p}_x] = [\hat{z}, \hat{p}_y] = 0 \\ [\hat{x}, \hat{y}] &= [\hat{y}, \hat{z}] = [\hat{z}, \hat{x}] = [\hat{p}_x, \hat{p}_y] = [\hat{p}_y, \hat{p}_z] = [\hat{p}_z, \hat{p}_x] = 0 \end{aligned} \quad (74)$$

对于角动量算符，我们根据定义(57)，有

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \quad \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z, \quad \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \quad (75)$$

并且可以证明

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y \quad (76)$$

我们可以进一步定义

$$\hat{L}^2 = \hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z \quad (77)$$

它具有如下对易性质

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}^2, \hat{L}_y] = [\hat{L}^2, \hat{L}_z] = 0 \quad (78)$$

5. 哈密顿算符与薛定谔方程

我们定义哈密顿算符为

$$\hat{H} = \hat{T} + \hat{V} = \frac{\hat{p}^2}{2m} + V(\hat{x}) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \quad (79)$$

那么薛定谔方程(13)可以写为

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = \hat{H} \psi(x, t) \quad (80)$$

或者我们可以写作

$$\hat{H} = i\hbar \frac{\partial}{\partial t} \quad (81)$$

所以薛定谔方程相当于告诉我们，哈密顿算符既是能量算符(79)，又是与时间演化相关的算符(81)。而求定态薛定谔方程的解，就是求解哈密顿算符(79)的本征方程。

第三节 基矢与表象

前面两节介绍了量子力学中波函数的概念，并且说明了它和量子状态及力学量之间的关系。由于态叠加原理，量子状态是可以做线性叠加的，这告诉我们可以用线性代数的语言对量子力学进行描述。本节中，我们将会发现，在线性代数的语言下，量子力学的规律将会变得非常简洁。

一、复线性空间（酉空间）

1. 酉空间的定义：

如果对一个集合 $\{\mathbf{u}\}$ 中的元素定义了加法与数乘两种运算，并且这些运算之间满足下述关系：

- 加法和数乘运算封闭： $\mathbf{u} + \mathbf{v}$, $c\mathbf{u}$
- 元素加法满足交换律和结合律
- 对于加法运算，有零元素（记为 $|0\rangle$ ）和逆元素（记为 $-\mathbf{u}$ ）
- 数乘满足结合律： $(c_1 c_2)\mathbf{u} = c_1(c_2\mathbf{u})$
- 数乘和加法满足分配律： $(c_1 + c_2)\mathbf{u} = c_1\mathbf{u} + c_2\mathbf{u}$, $c(\mathbf{u}_1 + \mathbf{u}_2) = c\mathbf{u}_1 + c\mathbf{u}_2$
- 矢量数乘1得到自己： $1\mathbf{u} = \mathbf{u}$

那么这个集合称为线性空间，或者矢量空间、向量空间，记作 L ；而这个集合中任意的元素 u 称为矢量或者向量。

这里我们允许数乘所采用的数为复数（我们跳过欧式空间，因为对欧式空间，大家都非常熟悉了），所以这个又可以称为复线性空间。

例如所有的函数 $f(x)$ 组成的空间，就是一个线性空间。

线性空间的定义使得我们可以把矢量看作数一样进行加法和减法（ $u - v$ 定义为 $u + (-v)$ ）运算，并且可以像数一样线性叠加。但是数与数之间还可以有乘法运算，特别的，对于复数 c_1 和 c_2 ，我们可以定义一种特殊的乘法： $c_1 \odot c_2 = \text{Re}(c_1^* c_2)$ （其中 Re 表示取实部），这样任意复数的模长就可以写作 $|c| = \sqrt{c \odot c}$ ，有了模长，我们就可以给复数赋予几何意义：每个复数对应于一个矢量箭头；如果 $c_1 \odot c_2 = 0$ ，我们称它们之间垂直（或者正交）；如果 $c_1 \odot c_2 \neq 0$ ，那么我们称它们之间有个夹角 θ ，并且满足 $\cos \theta = \frac{(c_1 \odot c_2)}{\sqrt{c_1 \odot c_1} \sqrt{c_2 \odot c_2}}$ 。

我们希望矢量之间也可以定义类似的运算，这样我们至少可以给每个矢量赋予一个模长的几何意义，也可以建立矢量之间正交的概念。于是对于任意两个矢量 u 和 v 的组合，我们让其对应于一个复数 $\langle u, v \rangle$ ，或者记作 $\langle u | v \rangle$ ，这个复数称为 u 和 v 的内积，前提是这一对应关系需要满足下述条件：

- a. $\langle u, v \rangle = \langle v, u \rangle^*$ ，或者记作 $\langle u | v \rangle = \langle v | u \rangle^*$ ；即如果颠倒组合中两个矢量的次序，对应的复数（内积）要取复共轭
- b. $\langle u, u \rangle > 0 \quad \forall u \neq 0$ ，或者记作 $\langle u | u \rangle > 0 \quad \forall u \neq 0$ ；即非零矢量和自己的内积是实数且必须大于零
- c. $\langle u, c_1 v_1 + c_2 v_2 \rangle = c_1 \langle u, v_1 \rangle + c_2 \langle u, v_2 \rangle$ ，或者记作 $\langle c_1 u_1 + c_2 u_2, v \rangle = c_1^* \langle u_1, v \rangle + c_2^* \langle u_2, v \rangle$ ，即内积能够保持线性叠加关系

我们稍微解释一下这些要求：我们定义矢量内积的驱动力是为了让矢量有模长的概念，所以这要求 $\langle u, u \rangle$ 一定得是实数。这样的话，条件 a 就成为了一个自然的要求，有了它，对于任意的 u ，都有 $\langle u, u \rangle = \langle u, u \rangle^*$ ，即 $\langle u | u \rangle$ 是实数。这样的话我们就可以把任意矢量 $|u\rangle$ 的模长定义为 $|u| = \sqrt{\langle u, u \rangle}$ 。条件 b 实际上说的是，只有零矢量的模长才为 0，所有非零矢量的模长都大于 0。而 c 是为了使得内积这个运算保持线性叠加关系，即我们把 $\langle u, v \rangle$ 看作是对 v 的一个操作，如果 v 原本是两个矢量 v_1 和 v_2 的线性叠加，那么对 v 操作的结果也得是对 v_1 和 v_2 分别操作的结果的线性叠加，且叠加系数保持不变。而 c 中的第二行是第一行与条件 a 结合的必然要求。

定义了这样一种内积运算的线性空间就称为酉空间。

如果两个矢量的内积为零，我们称这两个矢量正交。如果一个矢量的模长为 1，我们称这个矢量为单位矢量。

2. 酉空间中矢量的基矢展开

由于线性空间中的矢量总是满足叠加封闭性，我们可以从某个线性空间 L 的全体矢量集合 $\{u\}$ 当中随机去依次挑选一些元素，放到一个新的集合 U 中：

首先任意挑选一个非零矢量 u_1 ，放到 U 中，即让 $U = \{u_1\}$ ；

然后再去随机挑选第二个元素 u_2 ，并且判断，如果 u_2 能够写成 u_1 的线性组合的形式，即 $u_2 = c_1 u_1 \exists c_1$ ，那么把它扔掉重新随机挑选，直到选出来的 u_2 不能写成 u_1 的线性组合的形式为止，即 $u_2 \neq c_1 u_1 \forall c_1$ ，我们把这样的 u_2 放入 U 中，即让 $U = \{u_1, u_2\}$ ；

然后我们再去挑选第三个元素 u_3 ，并且判断，如果 u_3 能够写成 u_1 和 u_2 的线性组合的形式，即 $u_3 = c_1 u_1 + c_2 u_2 \exists c_1, c_2$ ，那么把它扔掉重新随机挑选，直到选出来的 u_3 不能写成 u_1 和 u_2 的线性组合的形式为止，即 $u_3 \neq c_1 u_1 + c_2 u_2 \forall c_1, c_2$ ，我们把这样的 u_3 放入 U 中，即让 $U = \{u_1, u_2, u_3\}$ ；

.....

一直不停地这样做下去，我们会发现最终当穷尽了 L 中所有的元素之后，我们得到了一个完整的集合 $U = \{u_1, u_2, \dots, u_N\}$ 。集合 U 的元素个数 N （可以是有限，也可以是无限，我们这里假设 N 有限）就叫做这个线性空间 L 的维度，集合 U 自身叫做 L 的一组基矢。根据基矢的选择方式，我们知道 L 中的任意一个元素 $|u\rangle$ ，都可以写成这组基矢的线性叠加：

$$u = \sum_{n=1}^N c_n u_n$$

由于我们从 L 得到 U 的过程是经过了大量的任意选取，对于给定的线性空间 L 来说，它的基矢选取不是唯一的。只要得到的 U 中，所有的元素都是线性无关的，并且 L 中所有的元素都可以表示成 U 中元素的线性组合，我们就认为 U 是 L 的一组基矢。但是两组不同的基矢，它们的元素个数一定是相同的，都是 N （线性空间的维度定理）。

对于任意一组基矢 $U = \{u_1, u_2, \dots, u_N\}$ ，我们一定可以对它进行处理，得到一组正交归一的基矢，这个过程叫做施密特正交化：

$$\beta_1 = u_1, e_1 = \beta_1 / |\beta_1|$$

$$\beta_2 = u_2 - (e_1, u_2) e_1, e_2 = \beta_2 / |\beta_2|$$

$$\beta_3 = u_3 - (e_1, u_3) e_1 - (e_2, u_3) e_2, e_3 = \beta_3 / |\beta_3|$$

.....

依次类推，我们从 $U = \{u_1, u_2, \dots, u_N\}$ 得到了 $E = \{e_1, e_2, \dots, e_N\}$ ，可以直接代入证明得

$$(e_m, e_n) = \delta_{mn} \quad (82)$$

并且 E 也是 L 的一组基矢，我们称这组基矢为正交归一基矢。

我们把任意矢量写成这组正交归一基矢的线性叠加形式

$$u = \sum_{n=1}^N u_n e_n \quad (83)$$

其中 u_n 是展开系数。用 e_m 与上式的左右两边同时做内积，我们有

$$u_m = (e_m, u) \quad (84)$$

这体现了正交归一基矢的方便之处：任意矢量的展开系数，通过简单的内积计算就可以求出。

结合(83)和(84)式，我们有

$$u = \sum_{n=1}^N (e_n, u) e_n = \sum_{n=1}^N e_n (e_n, u)$$

注意上式中最右边的表达式 $e_n (e_n, u)$ 表示矢量 e_n 与复数 (e_n, u) 做数乘，即我们定义 $cu = uc$ ，它们都表示矢量 u 和复数 c 的数乘运算。由于上式对所有的矢量 u 都成立，我们可以写作

$$\sum_{n=1}^N \mathbf{e}_n (\mathbf{e}_n, \mathbf{e}_n) = 1 \quad (85)$$

这个等式与“ $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N\}$ 是正交归一完备的基矢组”这件事是等价的。

3. 矢量的线性代数形式

由于任意矢量都可以表达成基矢展开的形式(83)，所以如果给定了基矢 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N\}$ ，矢量和它的展开系数是等价的，我们干脆把矢量记作它的展开系数罗列的形式，即

$$|u\rangle = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} \quad (86)$$

并且称 $|u\rangle$ 为 \mathbf{u} 的右矢量（列矢量）形式。

我们看一下矢量的内积在基矢展开下应该怎么表达。我们考虑任意两个矢量 \mathbf{u} 和 \mathbf{v} 的内积，并且把它们都写成(83)式的形式：

$$(\mathbf{u}, \mathbf{v}) = (\sum_{n=1}^N u_n \mathbf{e}_n, \sum_{m=1}^N v_m \mathbf{e}_m)$$

利用内积定义中的条件 c，有

$$(\mathbf{u}, \mathbf{v}) = \sum_{m,n=1}^N u_n^* v_m (\mathbf{e}_n, \mathbf{e}_m) = \sum_{m,n=1}^N u_n^* v_m \delta_{nm} = \sum_{n=1}^N u_n^* v_n$$

这告诉我们如果对任意矢量 \mathbf{u} ，定义与之对应的左矢量（行矢量）

$$\langle u| = [u_1^* \quad u_2^* \quad \dots \quad u_N^*] \quad (87)$$

那么矢量之间的内积就是行矢量和列矢量在线性代数规则下的乘法运算：

$$(\mathbf{u}, \mathbf{v}) = \langle u|\mathbf{v}\rangle = [u_1^* \quad u_2^* \quad \dots \quad u_N^*] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \quad (88)$$

这也是为什么我们在定义内积的时候，就说“ (\mathbf{u}, \mathbf{v}) 也记作 $\langle u|\mathbf{v}\rangle$ ”。

从线性代数运算的角度，左矢量可以看作右矢量的共轭转置。如果一个矢量 \mathbf{u} 是另外两个矢量 \mathbf{u}_1 和 \mathbf{u}_2 的线性叠加，即 $\mathbf{u} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2$ ，那么 \mathbf{u} 对应的右矢量 $|u\rangle$ 也是 $|u_1\rangle$ 和 $|u_2\rangle$ 的同样方式的线性叠加，即 $|u\rangle = c_1 |u_1\rangle + c_2 |u_2\rangle$ ；而 \mathbf{u} 对应的左矢量 $\langle u|$ 也是 $\langle u_1|$ 和 $\langle u_2|$ 的线性叠加，但是叠加的系数应该取复共轭，即 $\langle u| = c_1^* \langle u_1| + c_2^* \langle u_2|$ ，因为只有这样能与内积的条件 c 相吻合。

基矢本身也可以表达成左矢量和右矢量的形式，因为它们的展开系数中只有一项为 1，其余项都是 0：

$$|e_1\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, |e_2\rangle = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, |e_N\rangle = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (89)$$

$$\langle e_1| = [1 \ 0 \ \dots \ 0], \langle e_2| = [0 \ 1 \ \dots \ 0], \dots, \langle e_N| = [0 \ 0 \ \dots \ 1]$$

而矢量的基矢展开表达式(83)也可以写作左矢量和右矢量的形式（请自行验证其正确性）：

$$\begin{aligned} |u\rangle &= \sum_{n=1}^N u_n |e_n\rangle = \sum_{n=1}^N |e_n\rangle u_n \\ \langle u| &= \sum_{n=1}^N u_n^* \langle e_n| = \sum_{n=1}^N \langle e_n| u_n^* \end{aligned} \quad (90)$$

有了左矢量和右矢量的形式，我们可以把(85)式重新写作

$$\sum_{n=1}^N |e_n\rangle \langle e_n| = 1 \quad (91)$$

它作用在任何右矢量 $|u\rangle$ 上，得到右矢量 $|u\rangle$ 自己：

$$\sum_{n=1}^N |e_n\rangle \langle e_n| u \rangle = \sum_{n=1}^N |e_n\rangle u_n = \sum_{n=1}^N u_n |e_n\rangle = |u\rangle$$

它作用在任何左矢量 $\langle u|$ 上，也得到左矢量 $\langle u|$ 自己：

$$\langle u| \sum_{n=1}^N |e_n\rangle \langle e_n| = \sum_{n=1}^N \langle u| e_n \rangle \langle e_n| = \sum_{n=1}^N \langle e_n| u \rangle^* \langle e_n| = \sum_{n=1}^N u_n^* \langle e_n| = \langle u|$$

所以(91)式可以插入到任何一个表达式的中间，我们在后面将大量利用这一点。

4. 线性映射与矩阵

第二节中，我们曾经引入了算符的概念，并且定义它是作用在函数上面的一种操作。更一般地，对于线性空间中的矢量，我们也可以引入类似的“操作”的概念，或者称之为映射 \hat{f} ：把一个矢量 $|u\rangle$ 映射到另外一个矢量 $|v\rangle = \hat{f}|u\rangle$ （从这里开始，我们用右矢或者左矢的记号来表示矢量，不再使用 u 的记号，因为它们本质上是等价的）。我们同样可以定义映射之间的加法和乘法运算，并且显然可以得出映射的加法满足交换律，而乘法不满足。

线性代数中，我们主要感兴趣的也是其中一类特殊的映射，即线性映射，这类映射保持矢量之间的线性叠加关系，即如果 $|u\rangle = c_1|u_1\rangle + c_2|u_2\rangle$ ，那么 $\hat{f}|u\rangle = c_1\hat{f}|u_1\rangle + c_2\hat{f}|u_2\rangle$ 。由于线性映射保持线性叠加关系，我们只需要把线性映射对基矢的作用搞清楚，就能把它对所有矢量的作用搞清楚。考虑 \hat{f} 对任意基矢 $|e_n\rangle$ 的作用， $\hat{f}|e_n\rangle = |g_n\rangle$ ，即它把 $|e_n\rangle$ 映射成了某个矢量 $|g_n\rangle$ 。由于任意矢量都可以用基矢展开，我们把 $|g_n\rangle$ 写成(90)的形式：

$$|g_n\rangle = \hat{f}|e_n\rangle = \sum_{m=1}^N |e_m\rangle \langle e_m| \hat{f}|e_n\rangle$$

这相当于在 \hat{f} 和 $|e_n\rangle$ 之前插入了(91)式。注意到如果我们把所有的 $N \times N$ 个系数 $\langle e_m| \hat{f}|e_n\rangle$ 都找到了，那么基矢组 $\{|e_1\rangle, |e_2\rangle, \dots, |e_N\rangle\}$ 的映射结果就都清楚了，这样的话任意矢量的映射也清楚了。我们把这 $N \times N$ 个系数写成矩阵的形式，并且记为

$$[f] = \begin{bmatrix} f_{11} & \cdots & f_{1N} \\ \vdots & \ddots & \vdots \\ f_{N1} & \cdots & f_{NN} \end{bmatrix} \quad (92)$$

其中矩阵元 $f_{mn} = \langle e_m| \hat{f}|e_n\rangle$ 。

这样的话，线性映射对任意矢量 $|u\rangle$ 的作用，都可以写成矩阵 $[f]$ 和列矢量 (u) 的乘积：

$$\hat{f}|u\rangle = \begin{bmatrix} f_{11} & \cdots & f_{1N} \\ \vdots & \ddots & \vdots \\ f_{N1} & \cdots & f_{NN} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} = [f](u) \quad (93)$$

我们来验证一下这是正确的：我们把 $\hat{f}|u\rangle$ 这个矢量做基矢展开，即在 $\hat{f}|u\rangle$ 之前插入(91)式，有

$$\hat{f}|u\rangle = \sum_{m=1}^N |e_m\rangle \langle e_m| \hat{f}|u\rangle$$

我们进一步再对上式中的 $|u\rangle$ 做基矢展开，即在 \hat{f} 和 $|u\rangle$ 之间插入(91)式，有

$$\hat{f}|u\rangle = \sum_{m=1}^N |e_m\rangle \langle e_m| \hat{f}|u\rangle = \sum_{m,n=1}^N |e_m\rangle \langle e_m| \hat{f}|e_n\rangle \langle e_n| u \rangle = \sum_{m=1}^N [\sum_{n=1}^N f_{mn} u_n] |e_m\rangle$$

即 $\hat{f}|u\rangle$ 的第 m 个展开系数等于 $\sum_{n=1}^N f_{mn} u_n$ ，这正是(93)所表达的含义。

而两个线性映射的加法，就对应于它们所对应的矩阵的加法；两个线性映射的乘法，就对应于它们所对应的矩阵的乘法；这一点可以直接验证，我们不再推导。

上面我们只考虑了 \hat{f} 对右矢量的作用，对于左矢量，因为它是右矢量的共轭转置，我们可以很容易得到，如果 $\hat{f}|u\rangle = |v\rangle$ ，那么

$$\langle v| = [v_1^* \quad \dots \quad v_N^*] = [u_1^* \quad \dots \quad u_N^*] \begin{bmatrix} f_{11}^* & \dots & f_{N1}^* \\ \vdots & \ddots & \vdots \\ f_{1N}^* & \dots & f_{NN}^* \end{bmatrix} = (u)^+ [f]^+ \quad (94)$$

其中我们用+号表示共轭转置。由于线性映射与矩阵之间是一一对应的关系，我们把矩阵 $[f]^+$ 所代表的线性映射称为 \hat{f} 的厄米共轭映射，记作 \hat{f}^+ 。

所以行矢量和列矢量中的元素是矢量在给定基矢下的展开系数，而矩阵中的元素是线性映射在给定基矢下的展开系数，而线性代数里面的矩阵运算，实际上本质上是表示了矢量和线性映射之间的各种关系。

5. 基矢变换（幺正变换）

通过前面的讨论，我们知道线性映射把正交归一基矢组 $\{|e_1\rangle, |e_2\rangle, \dots, |e_N\rangle\}$ 映射成了一组新的矢量 $\{|g_1\rangle, |g_2\rangle, \dots, |g_N\rangle\}$ 。如果这一组新的矢量依然是一组正交归一基矢，那么我们称这个映射 \hat{U} 为幺正变换。我们看一下幺正变换的矩阵元需要满足什么条件。

首先，根据(92)式， \hat{U} 的矩阵元为

$$U_{mn} = \langle e_m | \hat{U} | e_n \rangle = \langle e_m | g_n \rangle \quad (95)$$

由于 $\{|g_1\rangle, |g_2\rangle, \dots, |g_N\rangle\}$ 依然是一组正交归一基矢，我们有

$$\langle g_m | g_n \rangle = \delta_{mn}$$

在上式中间插入(91)，有

$$\sum_{l=1}^N \langle g_m | e_l \rangle \langle e_l | g_n \rangle = \delta_{mn}$$

结合(95)式，上式相当于

$$[U]^+ [U] = [\delta_{mn}] = [I] \quad (96)$$

满足这个关系的矩阵叫做幺正矩阵。这告诉我们，把一组正交归一基矢映射成另一组正交归一基矢的线性映射，它的矩阵是幺正矩阵。这也是为什么这个映射叫做幺正变换。

我们看一下在幺正变换下，任意矢量 $|u\rangle$ 的展开系数怎么变化。由(84)式，我们知道在老基矢 $\{|e_1\rangle, |e_2\rangle, \dots, |e_N\rangle\}$ 下，矢量 $|u\rangle$ 的第 m 个展开系数为

$$u_m = \langle e_m | u \rangle$$

我们在 $\langle e_m |$ 和 $|u\rangle$ 之间插入 $\sum_{n=1}^N |g_n\rangle \langle g_n| = 1$ （因为 $\{|g_1\rangle, |g_2\rangle, \dots, |g_N\rangle\}$ 也是正交归一基矢），有

$$u_m = \sum_{n=1}^N \langle e_m | g_n \rangle \langle g_n | u \rangle = \sum_{n=1}^N \langle e_m | g_n \rangle u'_n = \sum_{n=1}^N U_{mn} u'_n$$

其中第二个等号利用了 $\langle g_n | u \rangle$ 就是 $|u\rangle$ 在新基矢 $\{|g_1\rangle, |g_2\rangle, \dots, |g_N\rangle\}$ 下的第 n 个展开系数 u'_n ，所以我们可以把上式写成矩阵形式

$$(u) = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} = [U_{mn}] \begin{bmatrix} u'_1 \\ \vdots \\ u'_N \end{bmatrix} = [U](u)' \quad (97)$$

或者在上式左右两边同时乘以 $[U]^+$ ，得到

$$(u)' = \begin{bmatrix} u'_1 \\ \vdots \\ u'_N \end{bmatrix} = [U_{mn}]^+ \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} = [U]^+ (u) \quad (98)$$

这就是新老基矢下，同一个矢量的展开系数之间的变换关系。

我们再看一下同一个线性映射，在新老基矢下的展开系数（即矩阵元）之间的变换关系。我

们考虑映射 \hat{f} ，它作用在任意矢量 $|u\rangle$ 上得到 $|v\rangle$ ，即 $\hat{f}|u\rangle = |v\rangle$ 。由于在么正变换（即正交归一基矢变换）下， $|u\rangle$ 和 $|v\rangle$ 的展开系数发生了变化，为了依然使得 $\hat{f}|u\rangle = |v\rangle$ ， \hat{f} 的矩阵元要发生相应的变化。我们把这个等式写成矩阵的形式，在老基矢下，它写作

$$(v) = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix} = \begin{bmatrix} f_{11} & \cdots & f_{1N} \\ \vdots & \ddots & \vdots \\ f_{N1} & \cdots & f_{NN} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} = [f](u) \quad (99)$$

但是在新基矢下，它应该写作

$$(v)' = \begin{bmatrix} v'_1 \\ \vdots \\ v'_N \end{bmatrix} = \begin{bmatrix} f'_{11} & \cdots & f'_{1N} \\ \vdots & \ddots & \vdots \\ f'_{N1} & \cdots & f'_{NN} \end{bmatrix} \begin{bmatrix} u'_1 \\ \vdots \\ u'_N \end{bmatrix} = [f]'(u)' \quad (100)$$

我们把(97)代入到(99)式中，得到

$$(v) = [U](v)' = [f](u) = [f][U](u)'$$

在上式两边同时乘以 $[U]^+$ ，我们有

$$[U]^+(v) = [U]^+[U](v)' = (v)' = [U]^+[f][U](u)'$$

将上式最后一个等号与(100)式做比较，我们有

$$[f]' = [U]^+[f][U] \quad (101)$$

这就是同一个线性映射 \hat{f} ，在新老基矢下的展开系数（即矩阵元）之间的变换关系。

在中学阶段的学习中，我们知道了可以通过建立坐标系，将欧式空间中的矢量用坐标进行表示；而如果我们改变了坐标系，每个矢量的具体坐标也会发生变化，但是矢量与矢量之间的关系保持不变：例如“矢量 \mathbf{u} 等于矢量 \mathbf{v} ”这件事与坐标系的选取（即 \mathbf{u} 和 \mathbf{v} 的具体坐标值）无关；再例如“把矢量 \mathbf{u} 顺时针旋转 45 度得到矢量 \mathbf{v} ”这件事与坐标系的选取（即 \mathbf{u} 和 \mathbf{v} 的具体坐标值）也无关。

同样的，在酉空间中，矢量是个绝对的量，而它的分量取决于基矢的选择；矢量之间的关系是绝对的关系，从而只要某些矢量的关系在一个基矢展开下成立，即使我们又选择了不同的基矢去展开这些矢量，在新的基矢下这些矢量的关系依然成立。例如在某个正交归一基矢下，我们有矩阵表达式（表达成矩阵形式）

$$F((u), (v), \dots; [f], [g], \dots) = 0 \quad (102)$$

其中 $(u), (v), \dots$ 是矢量， $[f], [g], \dots$ 是线性算符（矩阵），且 F 是算符和矢量之间的任意加、乘等操作，那么当我们改变了基矢的选择，变换到另外一组正交归一基矢下时，同样的表达式一定依然成立，只不过矢量和线性算符的展开系数要做相应的么正变换，即

$$F((u)', (v)', \dots; [f]', [g]', \dots) = 0 \quad (103)$$

依然成立。

6. 厄米矩阵

如果一个矩阵的共轭转置等于它自己，那么这个矩阵叫做厄米矩阵，即厄米矩阵 $[f]$ 满足

$$[f]^+ = [f] \quad (104)$$

它所代表的线性映射 \hat{f} ，也相应地称为厄米映射。

厄米矩阵的重要之处在于它的本征值和本征矢量。如果一个线性映射 \hat{f} 对某个非零矢量 $|u\rangle$ 操作的结果是使得矢量 $|u\rangle$ 与某个复数 λ 进行数乘，即

$$\hat{f}|u\rangle = \lambda|u\rangle \quad (105)$$

或者写成矩阵形式

$$[f](u) = \lambda(u) \quad (106)$$

那么这个复数 λ 叫做线性映射 \hat{f} 的本征值；矢量 $|u\rangle$ 叫做与本征值 λ 对应的， \hat{f} 的本征矢量；式(105)或者(106)叫做 \hat{f} 的本征方程。

如果 \hat{f} 是厄米映射，那么我们在(106)的左右两边同时乘以 $(u)^+$ ，有

$$(u)^+[f](u) = (u)^+\lambda(u) = \lambda(u)^+(u) \quad (107)$$

而根据(94)式， $(u)^+[f]$ 其实是 $\hat{f}^+|u\rangle$ 的行矢量形式，而由于(104)式， $(u)^+[f]$ 也是 $\hat{f}|u\rangle$ 的行矢量形式，又由于(105)式，所以 $(u)^+[f]$ 也是 $\lambda|u\rangle$ 的行矢量形式，即 $(u)^+[f] = \lambda^*(u)^+$ 。把此式带入(107)的左边，有

$$(u)^+[f](u) = \lambda^*(u)^+(u) = \lambda(u)^+(u)$$

由于 $|u\rangle$ 是非零矢量， $(u)^+(u) = |u|^2 > 0$ ，所以上式要求 $\lambda = \lambda^*$ ，即厄米矩阵（厄米算符）的本征值一定是实数。

如果 \hat{f} 有两个不同的本征值 λ_1 和 λ_2 ，并且分别有本征矢量 $|u_1\rangle$ 和 $|u_2\rangle$ ，即

$$\begin{aligned} [f](u_1) &= \lambda_1(u_1) \\ [f](u_2) &= \lambda_2(u_2) \end{aligned} \quad (108)$$

将第一个等式左右两边同时左乘 $(u_2)^+$ ，有

$$(u_2)^+[f](u_1) = (u_2)^+\lambda_1(u_1) = \lambda_1(u_2)^+(u_1) \quad (109)$$

将第二个等式做共轭转置，有

$$(u_2)^+[f]^+ = \lambda_2^*(u_2)^+$$

利用 $[f]^+ = [f]$ 和 λ_2 为实数，有

$$(u_2)^+[f] = \lambda_2(u_2)^+$$

将此时左右两边同时右乘 (u_1) ，有

$$(u_2)^+[f](u_1) = \lambda_2(u_2)^+(u_1)$$

将此式与(109)对比，由于我们假定了 $\lambda_1 \neq \lambda_2$ ，所以有

$$\langle u_2|u_1\rangle = (u_2)^+(u_1) = 0 \quad (110)$$

即厄米矩阵的属于不同本征值的本征矢量一定互相正交。

根据这个性质，对于厄米映射 \hat{f} ，我们总可以用它的本征矢量组成一组正交归一完备基矢。这是因为矩阵 $[f]$ 的本征值 λ 一定要满足方程

$$\det([f] - \lambda[I]) = 0 \quad (111)$$

而这个方程是关于 λ 的 N 阶代数方程，根据代数基本定理，它一定有 N 个复数解，即存在 N 个本征值 λ （我们把它记作 $\{f_1, f_2, \dots, f_N\}$ ）和相应的本征矢量（我们把它记作 $\{|f_1\rangle, |f_2\rangle, \dots, |f_N\rangle\}$ ），满足

$$\begin{aligned} \hat{f}|f_1\rangle &= f_1|f_1\rangle \\ \hat{f}|f_2\rangle &= f_2|f_2\rangle \\ &\vdots \\ \hat{f}|f_N\rangle &= f_N|f_N\rangle \end{aligned} \quad (112)$$

如果 $f_i \neq f_j$ ，那么根据(110)， $|f_i\rangle$ 和 $|f_j\rangle$ 必定正交；如果进一步， \hat{f} 所有的本征值都互不相同（我们称这些本征值为非简并的），那么(112)式中的每个方程正好只有一个线性无关的解，并且 $\{|f_i\rangle\}$ 自动形成了一组正交完备基矢，我们只要再把每个 $|f_i\rangle$ 归一化，它们就形成了正交归一完备的基矢组。

与之相反，如果某几个本征值相同，例如 $f_{i1} = f_{i2} = \dots = f_{iM} = \lambda^0$ ，我们称 λ^0 这个本征值是 M 重简并的。这时候，可以证明本征方程 $\hat{f}|f\rangle = \lambda^0|f\rangle$ 将有 M 个线性无关的解 $\{|f_{i1}\rangle, |f_{i2}\rangle, \dots, |f_{iM}\rangle\}$ ，并且这 M 个解的任意线性组合也一定是本征方程 $\hat{f}|f\rangle = \lambda^0|f\rangle$ 的解，换句话说，本征值 λ^0 对应的本征矢量也构成了一个小的线性空间（这个空间称为原来的 N 维线性空间的子空间）。于是我们可以对 $\{|f_{i1}\rangle, |f_{i2}\rangle, \dots, |f_{iM}\rangle\}$ 做施密特正交化，使它们成为这个子空间的正交归一基矢。由于大家都是与 λ^0 对应的本征矢量，我们既然可以记作

$$\begin{aligned}\hat{f}|f_{i1}\rangle &= f_{i1}|f_{i1}\rangle \\ \hat{f}|f_{i2}\rangle &= f_{i2}|f_{i2}\rangle \\ &\vdots \\ \hat{f}|f_{iM}\rangle &= f_{iM}|f_{iM}\rangle\end{aligned}$$

的形式。

对所有存在简并的本征值进行上述处理，我们最终可以在保持(112)式的前提下，使得 $\langle f_m|f_n\rangle = \delta_{mn}$ ，即用 \hat{f} 的本征矢量集合构成一组正交归一基矢。

如果我们用 $\{|f_1\rangle, |f_2\rangle, \dots, |f_N\rangle\}$ 做基矢，那么根据 5 中的内容，所有矢量的分量形式都要发生变化，所有线性映射的矩阵元也要发生变化，具体变换的方式由么正变换给出。

但是有一组矢量的分量形式在基矢 $\{|f_1\rangle, |f_2\rangle, \dots, |f_N\rangle\}$ 下变得特别简单，这就是 $\{|f_1\rangle, |f_2\rangle, \dots, |f_N\rangle\}$ 自己：在基矢 $\{|f_1\rangle, |f_2\rangle, \dots, |f_N\rangle\}$ 下，

$$(f_1) = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, (f_2) = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, (f_N) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (113)$$

同时有一个线性映射的矩阵元也变得非常简单，这就是 \hat{f} 自己：它在基矢 $\{|f_1\rangle, |f_2\rangle, \dots, |f_N\rangle\}$ 下成为：

$$[f] = \begin{bmatrix} f_1 & 0 & 0 & 0 \\ 0 & f_2 & \dots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_N \end{bmatrix} = \text{diag}(f_1, f_2, \dots, f_N) \quad (114)$$

同学们可以自行验证这两个等式。

需要注意的是，这一部分对酉空间的论述都是假设了空间的维度 N 是有限值。但是显然，这些基本结论在 N 趋于无限的时候依然是成立的。

二、量子力学的线性代数表述

有了酉空间的概念作为基础，我们重新表示量子力学的基本概念（下面的红字部分是量子力学理论所做的假设，可以认为是**物理**的部分；而黑字的部分是酉空间的数学结果，可以认为是数学的必然结论）。

【1】对于给定的粒子体系，它的所有可能量子态构成了一个酉空间。我们用右矢量的记号 $|\psi\rangle$ 表示这个空间中的任意一个元素，并且称之为态矢。

由于酉空间对复系数线性叠加是封闭的，所以任意两个量子状态的复系数线性叠加一定也是一个可能的量子状态，这就是态叠加原理。正如我们在第一节介绍的，量子力学给态叠加原理赋予的物理含义是：**【2】如果一个态矢 $|\psi\rangle = \sum_n c_n |\phi_n\rangle$ ，那么当粒子处于态矢 $|\psi\rangle$ 时，等价于它同时处于态矢 $\{|\phi_n\rangle\}$ 中，并且处于态矢 $|\phi_n\rangle$ 的几率幅（相对）为 c_n 。**

而线性算符 \hat{f} 是这个酉空间中的线性映射，它们把一个态矢映射成另一个态矢，并且保持它们之间的线性叠加关系。

【3】可观测的力学量 \hat{A} 对应于厄米映射；它的本征方程写作

$$\hat{A}|a_n\rangle = a_n|a_n\rangle \quad (115)$$

这个厄米映射的全部本征值 $\{a_n\}$ 就是这个力学量所有可能的观测值；（由于厄米映射的本征值一定是实数，这说明可观测力学量的测量值一定是实数）

当体系处于本征矢量（也称为本征态） $|a_n\rangle$ 时，体系的力学量 \hat{A} 具有确定的取值 a_n 。

另外，由于厄米映射（可观测力学量）的本征矢量总可以组成一组正交归一基矢，我们可以首先指定一个可观测力学量 \hat{A} ，然后找到它的全部本征矢量，并且用这些本征矢量组成一组正交归一基矢，从而把所有的量子状态 $|\psi\rangle$ 和所有的线性算符 \hat{f} 展开成列矢量（行矢量）和矩阵的形式。这种与一开始 \hat{A} 的选择有关的，具体的展开形式称为 \hat{A} 表象。

对于连续位置空间中的粒子体系，我们一般选择“位置”这个力学量，它对应的线性映射记为 \hat{x} ，并且它的本征方程写作

$$\hat{x}|x\rangle = x|x\rangle \quad (116)$$

由于 \hat{x} 是可观测力学量，（适当归一化的） $\{|x\rangle\}$ 可以构成正交归一基矢，即

$$\langle x_1|x_2\rangle = \delta(x_1 - x_2) \quad (117)$$

注意这个关系是(82)式的连续版本，当 m, n 下标成为连续量 x_1, x_2 时， δ_{mn} 成为 $\delta(x_1 - x_2)$ 。

我们用基矢 $\{|x\rangle\}$ 去展开所有的量子态 $|\psi\rangle$ （即用 \hat{x} 表象去表述量子状态），有

$$|\psi\rangle = \int dx |x\rangle \langle x|\psi\rangle \quad (118)$$

这是(90)式即 $|u\rangle = \sum_{n=1}^N |e_n\rangle \langle e_n|u\rangle$ 得连续版本，当下表 n 成为连续量 x 时，对 n 的求和变为对 x 的积分。**【4】其中 $\langle x|\psi\rangle$ 是 $|\psi\rangle$ 的展开系数，它也可以看成是 x 的函数，我们把它记作 $\psi(x)$ ，称为波函数。**这就是2.1节引入的波函数。

由于(118)式是 $|\psi\rangle$ 的展开形式，根据**【2】**，有当粒子的波函数为 $\psi(x)$ 时，它处于状态 $|x\rangle$ 的几率幅（相对）为 $\langle x|\psi\rangle = \psi(x)$ ，或者说粒子位于 x 处的几率幅（相对）为 $\psi(x)$ ，这就是第一节中我们引入的波函数的物理意义。

在 \hat{x} 表象下，两个态矢 $|\phi\rangle$ 和 $|\psi\rangle$ 的内积应该按照(88)式的方式展开，即

$$\langle \phi|\psi\rangle = \sum_n \phi_n^* \psi_n = \sum_n \langle \phi|e_n\rangle \langle e_n|\psi\rangle = \int \phi^*(x) \psi(x) dx \quad (119)$$

知道了任意量子状态 $|\psi\rangle$ 的展开系数，我们再去看看力学量 \hat{A} 的展开系数。

我们首先依然考察最简单的力学量 \hat{x} 。根据(114)，如果 x 取离散值，力学量 \hat{x} 在它自己的表象下成为对角矩阵 $\text{diag}(x_1, x_2, \dots)$ ，或者写成矩阵元的形式为 $x_{mn} = x_m \delta_{mn}$ 。但是 x 实际上是连续取值的，所以我们要把 δ_{mn} 换成 δ 函数，即

$$x_{x_1, x_2} = x_1 \delta(x_1 - x_2) \quad (120)$$

而映射 \hat{x} 对矢量 $|\psi\rangle$ 的作用，根据(93)式，或者等价的

$$\langle e_m | \hat{f} | u \rangle = \sum_{n=1}^N f_{mn} \langle e_n | u \rangle \quad (121)$$

我们有（用 \hat{x} 替换 \hat{f} ，用 $|\psi\rangle$ 替换 $|u\rangle$ ，用 $|x_1\rangle$ 替换 $|e_m\rangle$ ，用 $|x_2\rangle$ 替换 $|e_n\rangle$ ，用 x_1 替换 m ，用 x_2 替换 n ，用 $\int dx_2$ 替换 \sum_n ）

$$\langle x_1 | \hat{x} | \psi \rangle = \int dx_2 x_{x_1, x_2} \langle x_2 | \psi \rangle = \int dx_2 x_1 \delta(x_1 - x_2) \psi(x_2) = x_1 \psi(x_1) \quad (122)$$

这就是(62)式，即如果把 \hat{x} 看成是对函数的微分算符，那么 $\hat{x} = x$ 。

我们再去考察“动量”这个力学量。**【5】量子力学的一个基本假设是 $[\hat{x}, \hat{p}] = i\hbar$** ，即(73)式。

我们把这一要求写成矩阵元的形式，即

$$[\hat{x}, \hat{p}]_{x_1, x_3} = i\hbar \delta(x_1 - x_3) \quad (123)$$

利用矩阵乘法，有

$$\begin{aligned} [\hat{x}, \hat{p}]_{x_1, x_3} &= [\hat{x}\hat{p}]_{x_1, x_3} - [\hat{p}\hat{x}]_{x_1, x_3} = \int dx_2 x_{x_1, x_2} p_{x_2, x_3} - \int dx_2 p_{x_1, x_2} x_{x_2, x_3} \\ &= \int dx_2 x_1 \delta(x_1 - x_2) p_{x_2, x_3} - \int dx_2 p_{x_1, x_2} x_2 \delta(x_2 - x_3) \\ &= x_1 p_{x_1, x_3} - p_{x_1, x_3} x_3 = (x_1 - x_3) p_{x_1, x_3} \end{aligned}$$

代入(123)，有

$$p_{x_1, x_3} = i\hbar \frac{\delta(x_1 - x_3)}{x_1 - x_3} = -i\hbar \delta'(x_1 - x_3) = -i\hbar \frac{\partial}{\partial x_1} \delta(x_1 - x_3) \quad (124)$$

我们进一步把它写成(121)式的矩阵乘法形式，有（用 \hat{p} 替换 \hat{f} ，用 $|\psi\rangle$ 替换 $|u\rangle$ ，用 $|x_1\rangle$ 替换 $|e_m\rangle$ ，用 $|x_2\rangle$ 替换 $|e_n\rangle$ ，用 x_1 替换 m ，用 x_2 替换 n ，用 $\int dx_2$ 替换 \sum_n ）

$$\langle x_1 | \hat{p} | \psi \rangle = \int dx_2 p_{x_1, x_2} \langle x_2 | \psi \rangle = -i\hbar \int dx_2 \frac{\partial}{\partial x_1} \delta(x_1 - x_2) \psi(x_2) = -i\hbar \frac{\partial}{\partial x_1} \psi(x_1) \quad (125)$$

这就是(67)式，即如果把 \hat{p} 看成是对函数的微分算符，那么 $\hat{p} = -i\hbar \frac{\partial}{\partial x}$ 。

最后，**【6】如果一个量子系统与宏观系统没有相互作用，那么其态矢随时间的演化应满足薛定谔方程**

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \quad (126)$$

其中 $|\psi(t)\rangle$ 表示态矢随时间变化，所以是时间的“函数”。而(80)式只是(126)式在坐标表象下的矩阵乘法形式。

这里，我们相当于把第一节和第二节中的内容用态矢的语言重新定义了一遍。这样做的好处是，这里所有的**红色表达式**都不再依赖于具体基矢的选择，所以是对于任何基矢都成立的。而第一节和第二节的内容只是这些红色表达式在位置表象下的具体展开形式。

三、表象变换

根据上一部分的内容，我们可以任意选择基矢，把红字部分的内容进行基矢展开，这里我们举一个例子，就是用动量表象去展开这些结论。

我们首先要把动量算符 \hat{p} 的本征态矢作为基矢，根据**【3】**，这些本征态矢 $|p\rangle$ 满足本征方程

$$\hat{p} |p\rangle = p |p\rangle \quad (127)$$

并且可构成正交归一基矢，即

$$\langle p_1 | p_2 \rangle = \delta(p_1 - p_2) \quad (128)$$

动量表象下，任意态矢 $|\psi\rangle$ 应该用基矢 $\{|p\rangle\}$ 去展开，即

$$|\psi\rangle = \int dp |p\rangle \langle p | \psi \rangle = \int dp |p\rangle c(p) \quad (129)$$

其中展开系数 $c(p) = \langle p | \psi \rangle$. 我们在 $\langle p | \psi \rangle$ 的中间插入关系（即(91)的连续形式）

$$1 = \int dx |x\rangle \langle x| \quad (130)$$

有

$$c(p) = \langle p | \psi \rangle = \int dx \langle p | x \rangle \langle x | \psi \rangle = \int \psi_p^*(x) \psi(x) dx \quad (131)$$

这就是(8)式。其中 $\psi_p^*(x) = \langle p | x \rangle$ ，或者说 $\psi_p(x) = \langle x | p \rangle$ ，是态矢 $|p\rangle$ 在位置表象下的波函数，由(6)式给出。所以我们看到，波函数 $\psi(x)$ 的平面波展开系数，就是 $\psi(x)$ 所代表的态矢 $|\psi\rangle$ 在动量表象下的基矢展开系数。而平面波展开，或者说傅里叶变换，就相当于从坐标表象到动量表象的么正变换。

由于(129)式是态矢对基矢 $\{|p\rangle\}$ 的展开，根据【2】，有当粒子的波函数为 $\psi(x)$ 时，或者说它的态矢为 $|\psi\rangle$ 时，它具有动量 p 的几率幅（相对）为 $c(p)$ ，这与我们前面的内容是一致的。

与上一部分完全类似，两个态矢的内积为

$$\langle \phi | \psi \rangle = \int \langle \phi | p \rangle \langle p | \psi \rangle dp = \int c_\phi^*(p) c_\psi(p) dp \quad (132)$$

动量算符 \hat{p} 在自己的表象下成为对角矩阵

$$p_{p_1, p_2} = p_1 \delta(p_1 - p_2) \quad (133)$$

并且根据【5】，可以推导出位置算符 \hat{x} 在 p 表象下成为如下矩阵

$$x_{p_1, p_3} = -i\hbar \frac{\delta(p_1 - p_3)}{p_1 - p_3} = i\hbar \delta'(p_1 - p_3) = i\hbar \frac{\partial}{\partial p_1} \delta(p_1 - p_3) \quad (134)$$

我们再举一个比较简单的例子，就是能量表象。我们首先把哈密顿算符 \hat{H} 的本征态矢作为基矢，根据【3】，这些本征态矢满足本征方程（我们假设本征值是离散的，连续的情况可以类似得出）

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle \quad (135)$$

并且可构成正交归一基矢，即

$$\langle \psi_m | \psi_n \rangle = \delta_{mn} \quad (136)$$

能量表象下，任意态矢 $|\psi\rangle$ 应该用基矢 $\{|\psi_n\rangle\}$ 去展开，即

$$|\psi\rangle = \sum_n c_n |\psi_n\rangle = \sum_n \langle \psi_n | \psi \rangle |\psi_n\rangle \quad (137)$$

或者写成列矢量的形式

$$|\psi\rangle = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix} \quad (138)$$

其中 $c_n = \langle \psi_n | \psi \rangle$.

而哈密顿量在自己的表象下成为对角矩阵

$$\hat{H} = \begin{bmatrix} E_1 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & \ddots \end{bmatrix} \quad (139)$$

其中对角元是各本征能量。

在能量表象下，含时薛定谔方程也变得特别简洁，即

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} c_1(t) \\ c_2(t) \\ \vdots \end{bmatrix} = \hat{H} \begin{bmatrix} c_1(t) \\ c_2(t) \\ \vdots \end{bmatrix} = \begin{bmatrix} E_1 c_1(t) \\ E_2 c_2(t) \\ \vdots \end{bmatrix} \quad (140)$$

直接解得：

$$|\psi(t)\rangle = \begin{bmatrix} c_1(t=0)e^{-\frac{i}{\hbar}E_1 t} \\ c_2(t=0)e^{-\frac{i}{\hbar}E_2 t} \\ \vdots \end{bmatrix} \quad (141)$$

第四节 测量与不确定原理

一、测量结果的预测

我们首先看一下，如果已知粒子的量子状态，怎么预测测量任意力学量的结果。

如果我们希望观测力学量 A ，那么根据【3】，与力学量 A 所对应的厄米映射（厄米算符） \hat{A} 的全部本征值 a_n 是测量的可能取值，并且相应的本征矢量 $|a_n\rangle$ 可以组成正交归一完备的基矢。当体系处于本征矢量 $|a_n\rangle$ 时，该力学量具有确定的取值 a_n ，或者说观测的结果一定 100% 是 a_n 。

如果体系并不是处于 \hat{A} 本征状态，而是处于一个任意的状态 $|\psi\rangle$ ，我们怎么去预测测量结果呢？由于我们总可以把 $|\psi\rangle$ 用正归一完备的 $\{|a_n\rangle\}$ 展开，即

$$|\psi\rangle = \sum_n c_n |a_n\rangle = \sum_n \langle a_n | \psi \rangle |a_n\rangle \quad (142)$$

这样的话，利用态叠加原理，我们知道粒子相当于同时处于不同的状态 $\{|a_n\rangle\}$ 中，并且处于状态 $|a_n\rangle$ 的几率幅（相对）为 $c_n = \langle a_n | \psi \rangle$ 。由于 $|a_n\rangle$ 代表了测量 A 的结果确定为 a_n 的状态，所以我们可以说，当粒子处于状态 $|\psi\rangle$ 时，如果 A 不存在简并，那么测量 A 的结果（确定）为 a_n 的几率幅（相对）为 $c_n = \langle a_n | \psi \rangle$ ，或者说几率（相对）为 $|c_n|^2 = |\langle a_n | \psi \rangle|^2$ ，即我们预测了测量结果的几率分布。

如果 A 存在简并，那么(142)的展开中，某些项，例如 $|a_{i1}\rangle$ 、 $|a_{i2}\rangle$ 、...、 $|a_{iM}\rangle$ 对应的是同一个本征值 $a_{i1} = a_{i2} = \dots = a_{iM} = \lambda^0$ 。这时根据态叠加原理，(142)式依然表示粒子同时处于不同的状态 $\{|a_n\rangle\}$ 中，并且处于状态 $|a_n\rangle$ 的几率幅（相对）为 $c_n = \langle a_n | \psi \rangle$ 。但是由于 $|a_{i1}\rangle$ 、 $|a_{i2}\rangle$ 、...、 $|a_{iM}\rangle$ 对应的是同一个本征值 $a_{i1} = a_{i2} = \dots = a_{iM} = \lambda^0$ ，所以如果我们取预测测量 A 的结果为 λ^0 的几率（相对），那么这个几率需要把不同的方式进行几率叠加（量子物理基本假设第三条），即

$$P(\lambda^0) \sim \sum_{m=1}^M |c_{im}|^2 \quad (143)$$

通过上述几率分布，我们可以得到测量结果的期望值：

$$\langle A \rangle = \frac{\sum_n |c_n|^2 a_n}{\sum_n |c_n|^2}$$

我们也可以把这个表达式写成矩阵的形式：

$$\langle A \rangle = \frac{[c_1^*, c_2^*, \dots] \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & \dots \\ 0 & \vdots & \ddots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix}}{[c_1^*, c_2^*, \dots] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix}}$$

而我们注意到在 A 表象下, $|\psi\rangle = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \end{bmatrix}$, $\langle\psi| = [c_1^*, c_2^*, \dots]$, $\hat{A} = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & \dots \\ 0 & \vdots & \ddots \end{bmatrix}$, 所以上式在 A

表象下可以写作 $\langle A \rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}$, 又由于不同的表象只是态矢和算符的具体展开系数不同, 而态矢、算符之间的关系式应该与表象无关, 所以我们有

$$\langle A \rangle = \frac{\sum_n |c_n|^2 a_n}{\sum_n |c_n|^2} = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} \quad (144)$$

这一等式对于所有表象都成立。我们可以验证(64)和(69)都是这一表达式的特例。

二、测量对量子状态的影响——塌缩

所谓测量, 就是我们通过一个宏观的仪器与被观测的对象(量子体系)形成相互作用, 从而利用仪器的读数(例如指针的指向)去判断量子体系某力学量的取值。所以测量一旦发生, 就意味着量子体系与经典体系发生了相互作用。

上一节的内容告诉我们, 在测量发生(量子体系与宏观仪器产生相互作用)之前, 我们可以通过体系量子状态 $|\psi\rangle$ 的展开系数(以待测力学量 A 的本征函数为基矢)预测测量结果的几率分布。而一旦测量发生并且完成, 如果 A 不存在简并, 这些测量结果中的某一个, 例如 a_n 就被选择了, 宏观仪器的指针也相应地指向与 a_n 对应的位置。与此相对应地, **【7】体系的量子状态在测量完成时, 也被选择为(142)展开式中相应的那一项,**

$$|\psi\rangle = \langle a_n | \psi \rangle |a_n\rangle \quad (145)$$

即量子状态在测量完成时发生了塌缩。量子状态从(143)塌缩到(144)的过程类似于一个矢量在某个坐标轴上投影的过程。由于 $\langle a_n | \psi \rangle |a_n\rangle$ 其实就是 $|a_n\rangle$ 数乘一个系数, 两者本质上是同一个量子状态, 我们也可以直接说体系的量子状态在测量完成时塌缩为 $|a_n\rangle$ 。

我们可以通过一个假想的连续测量过程去理解塌缩: 如果我们在对力学量 A 的测量完成之后, 立即再去测量 A 力学量, 由于宏观的仪表不应该在间隔为零的时间内发生转动, 这要求第二次测量的结果必须和第一次相同。如果第一次测量 A 得到了结果 a_n , 这相当于要求第二次测量得到结果依然为 a_n 的几率是 100%, 而这个要求只有当体系的量子状态为 $|a_n\rangle$ 时才能符合, 即第一次测量之后, 系统的状态要塌缩为 $|a_n\rangle$ 。

如果 A 存在简并, 例如 $|a_{i1}\rangle, |a_{i2}\rangle, \dots, |a_{iM}\rangle$ 对应的是同一个本征值 $a_{i1} = a_{i2} = \dots = a_{iM} = \lambda^0$, 并且测量 A 的结果恰好得到 λ^0 , 那么这一测量结果同时对应于(142)中的 $|a_{i1}\rangle, |a_{i2}\rangle, \dots, |a_{iM}\rangle$ 那些项, 体系的量子状态会按照下述方式塌缩

$$|\psi\rangle = \sum_{m=1}^M \langle a_{im} | \psi \rangle |a_{im}\rangle \quad (146)$$

前面我们讨论量子状态随时间的演化时, 假设了体系是一个纯的量子体系, 即其与经典宏观体系没有相互作用, 这时它的演化由薛定谔方程给出。与此相反, 测量导致的塌缩反映了当体系与宏观体系发生相互作用时, 宏观体系的影响对量子体系大到如此的地步, 以至于量子状态会发生塌缩。

三、不确定原理

根据我们前面的论述。如果对一个体系，在几乎同时的时间范围内反复去测量同一个力学量 A ，除了第一次测量得到不同的结果是几率性的以外，以后的测量都会确定地得到与第一次测量结果相同的结果，这是因为体系的状态在第一次测量完成时发生了塌缩。

现在我们考虑另外一个问题，我们先进行力学量 A 的测量，然后再进行另一个不同的力学量 B 的测量，并且假设这两次测量之间时间间隔非常短，量子状态还来不及发生除塌缩以外的演化。我们知道，如果测量之前体系处于一个任意的量子状态，那么第一次测量（即对 A 的测量）的结果一般是几率性的；但是第二次测量（即对 B 的测量），是几率性的还是确定性的呢？

事实上，这取决于算符（矩阵） \hat{A} 与 \hat{B} 是否对易。我们假设 \hat{A} 的本征值不存在简并，即

$$\begin{aligned}\hat{A}|a_1\rangle &= a_1|a_1\rangle \\ \hat{A}|a_2\rangle &= a_2|a_2\rangle \\ &\vdots \\ \hat{A}|a_N\rangle &= a_N|a_N\rangle\end{aligned}$$

且 a_1 、 a_2 、...、 a_N 互不相同。如果 \hat{A} 与 \hat{B} 对易，那么 $\hat{A}\hat{B} = \hat{B}\hat{A}$ ，于是

$$\hat{A}(\hat{B}|a_n\rangle) = \hat{B}\hat{A}|a_n\rangle = \hat{B}a_n|a_n\rangle = a_n(\hat{B}|a_n\rangle)$$

这告诉我们 $\hat{B}|a_n\rangle$ 也是 \hat{A} 的与本征值 a_n 对应的本征矢量。由于我们假设了 \hat{A} 不简并，所以方程 $\hat{A}|\psi\rangle = a_n|\psi\rangle$ 只有一个线性无关的解，即 $|a_n\rangle$ 与 $\hat{B}|a_n\rangle$ 只能线性相关，或者说

$$\hat{B}|a_n\rangle = c|a_n\rangle$$

其中 c 是某个常数。而这意味着 $|a_n\rangle$ 也同时是 \hat{B} 的本征矢量，或者说 $|a_n\rangle$ 是 \hat{A} 与 \hat{B} 的共同本征矢量。由于上式中，一开始取不同的 n ，我们一般会得到不同的 c ，所以我们可以不失一般性地把 c 记作 b_n ，即

$$\begin{aligned}\hat{A}|a_1\rangle &= a_1|a_1\rangle & \hat{B}|a_1\rangle &= b_1|a_1\rangle \\ \hat{A}|a_2\rangle &= a_2|a_2\rangle & \hat{B}|a_2\rangle &= b_2|a_2\rangle \\ &\vdots & &\vdots \\ \hat{A}|a_N\rangle &= a_N|a_N\rangle & \hat{B}|a_N\rangle &= b_N|a_N\rangle\end{aligned}\tag{145}$$

这时，我们再回到刚才的问题，当我们对力学量 A 的测量完成时，根据第二部分的内容，如果测量结果为 $A = a_n$ ，那么体系的量子状态塌缩为 $|a_n\rangle$ ；此时我们立即测量力学量 B ，根据第一部分的结果，由于体系的状态 $|a_n\rangle$ 已经是 \hat{B} 的本征态了，所以测量的结果将会 100% 的是 $B = b_n$ 。

总结一下，如果算符 \hat{A} 与 \hat{B} 对易，那么它们将存在一组共同的（正交归一完备）本征态集合（可以证明，这个结论对于存在简并的情况也是成立的），对这个集中的任何一个量子态进行 A 或者 B 的测量，都会得到确定的结果；如果体系初始的状态是一个任意的量子态，那么经过上述 A 测量之后（如果 A 是简并的，那么经过依次测量 A 和 B 之后），体系总会塌缩到它们的某个共同本征态上，此后再立即进行 A 或 B 的任意测量，都会得到确定的结果。我们称 A 和 B 是可以共同测量的。

反之，如果 \hat{A} 与 \hat{B} 不对易，那么它们就不存在共同的（正交归一完备）本征态集合，从而一般来说， A 测量之后再去测量 B ，结果总是几率的， B 测量之后再去测量 A ，结果也是几率的。我们称 A 和 B 是不可以共同测量的。

对易算符（力学量）的例子（参考(74)式）包括三维空间中的三个位置分量， \hat{x} , \hat{y} 和 \hat{z} ，它们的共同本征函数为 $\delta(x-x_0)\delta(y-y_0)\delta(z-z_0)$ ；三个动量分量也是互相对易的， \hat{p}_x , \hat{p}_y 和 \hat{p}_z ，它们的共同本征函数为 $\frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}p_x x}\frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}p_y y}\frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}p_z z}$ ； \hat{x} , \hat{y} 和 \hat{p}_z 也是相互对易的，它们的共同本征函数为 $\delta(x-x_0)\delta(y-y_0)\frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}p_z z}$ ；等等。

不对易的算符，最典型的就是同一个方向上的位置和动量， $[\hat{x}, \hat{p}_x] = i\hbar \neq 0$ 。同一个量子状态，如果对力学量 x 的取值越确定，那么这个状态对 p_x 的取值就越不确定：例如波函数 $\delta(x-x_0)$ ，其力学量 x 的取值是 100%确定为 x_0 的，但是它的平面波展开系数模平方为 $|c(p_x)|^2 = \frac{1}{2\pi\hbar}$ 与 p_x 无关，这说明所有的动量 p_x 都具有相同的取值几率，即 p_x 的取值具有最大的不确定度；再例如波函数 $\frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}p_0 x}$ ，它的动量具有确定的取值 p_0 ，但是它的波函数模平方恒为 $|\psi(x)|^2 = \frac{1}{2\pi\hbar}$ ，这说明所有的位置 x 都具有相同的取值几率，即位置的取值具有最大的不确定度。更一般地，我们可以验证，如果一个粒子的波函数为高斯波包 $\psi(x) \sim e^{-ax^2}$ ，那么它的平面波展开系数也是一个高斯波包 $c(p) \sim e^{-bp^2}$ ，并且这两个高斯波包的宽度乘积为一个常数。换句话说，如果波函数的波包宽度越宽，对应于粒子位置的不确定性越大，那么它的平面波展开系数的宽度越窄，对应于粒子动量的不确定性越小；反之，如果波函数的波包宽度越窄，对应于粒子位置的不确定性越小，那么它的平面波展开系数的宽度越宽，对应于粒子动量的不确定性越大。

更一般地，可以证明，如果 \hat{A} 与 \hat{B} 不对易，并且我们把它们的对易算符写作

$$[\hat{A}, \hat{B}] = i\hat{C}$$

（注意由于 $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$ ，所以 \hat{C} 一定是一个厄米算符），那么有

$$\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq \frac{1}{4} \langle C \rangle^2 \quad (146)$$

其中 $\langle \Delta A^2 \rangle$ 和 $\langle \Delta B^2 \rangle$ 分别表示对于某个确定的量子状态，力学量 A 和 B 测量值的均方偏差。由于测量值的均方偏差可以作为不确定性的衡量，这个式子告诉我们，只要两个力学量不对易，那么它们的不确定性乘积一般要大于等于某个正数，而不会同时为 0。这就是不确定原理。

对于位置和动量，直接带入(146)式，我们有

$$\langle \Delta x^2 \rangle \langle \Delta p^2 \rangle \geq \frac{1}{4} \hbar^2 \quad (147)$$

第五节 本章总结

我们总结一下量子力学的基本概念和理论：

1. 量子系统的状态由量子态 $|\psi\rangle$ 表示，它是 Hilbert 空间（一种可以对矢量做微分的复线性空间）中的矢量（称为态矢量，或者态矢），或者说一个体系所有可能的量子态构成一

个复线性空间（也叫做这个体系的态矢空间）。由于所有量子态构成一个线性空间，那么如果 $|\psi\rangle$ 是体系可能的量子态，则任意 $c|\psi\rangle$ 一定也是体系可能的量子态； $|\psi\rangle$ 与任意 $c|\psi\rangle$ 对应于同一个物理状态，或者说 Hilbert 空间用于表示体系的物理状态是存在冗余的。

2. 由于所有量子态构成一个线性空间，那么如果 $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_N\rangle$ 都是体系可能的量子态，则任意 $c_1|\psi_1\rangle + c_2|\psi_2\rangle + \dots + c_N|\psi_N\rangle$ 一定也是体系可能的量子态；量子态叠加原理给这种叠加赋予了物理含义：如果一个态矢 $|\psi\rangle = \sum_n c_n |\psi_n\rangle$ ，那么当粒子处于态矢 $|\psi\rangle$ 时，等价于它同时处于各态矢 $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_N\rangle$ 中，并且处于态矢 $|\psi_n\rangle$ 的几率幅（相对）为 c_n 。
3. 力学量对应于线性映射（算符），更具体地说任意可以测量的力学量 A 对应于某个相应的厄米映射 \hat{A} ，它的本征方程写作

$$\hat{A}|a_n\rangle = a_n|a_n\rangle$$

\hat{A} 的全部本征值 $\{a_n\}$ 就是这个力学量 A 所有可能的观测值。当体系处于本征态 $|a_n\rangle$ 时，力学量 A 具有确定的取值 a_n 。

4. 当体系处于任意的状态 $|\psi\rangle$ 时，对力学量 A 的测量一般具有几率特性：无简并情况下， A 测量值为 a_n 的相对几率为 $|\langle a_n|\psi\rangle|^2$ ，并且测量完成时，体系的状态塌缩为 $\langle a_n|\psi\rangle|a_n\rangle$ ；有简并情况下，测量值的相对几率要对简并的各项进行求和，塌缩后的状态也要对简并的各项进行求和。
5. 位置和动量之间满足对易关系 $[\hat{x}, \hat{p}] = i\hbar$ 。
6. 如果一个量子系统与宏观系统没有相互作用，那么其态矢随时间的演化应满足薛定谔方程

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle.$$