第二章　量子态　力学量　Born 规则 I

主要内容

1. 量子态, pre-probability, 被表示为 Hilbert 空间中归一化的向量.
2. 观测量, 或者力学量, 是在特定的观测过程中我们所关心的系统的某些性质. 观测量的数学形式是 Hilbert 空间上的厄密算子.
3. 在有限维空间 $\mathbb{C}^n$ 中量子态的数学表示, 不同表象之间的关系.
4. 在 $\mathbb{C}^n$ 上的矩阵, 尤其是厄密矩阵和酉矩阵.
5. 量子力学的测量假说, 即 Born-Lüders 规则.

1 量子态

量子态属于量子理论的形式系统中的概念. 不过, 我们要从实际的观测结果的角度说明量子态, 于是, 对量子态的定义也将涉及对应规则, 涉及量子理论的解释系统.

- 量子态本身不是几率, 将要在具体的观测中给出观测结果的几率分布.
- 对所有可能的测量有所响应.
  "所有可能的测量", 这使得量子态的定义显得非常繁复. 当然, 更实际的说法是, 就某种类型而言的所有可能的测量. 例如, 在 SG 实验中, 我们关心的是像是经典的"磁矩"一类的性质, 所用的实验装置也是为此服务的, 所以, 在这种场合中, "所有可能的测量"指的是所有可能的磁场指向, 而不需要考虑关于能级跃迁之类的测量.

这样的叙述赋予了量子态的操作意义. 它不仅仅是一个数学符号, 而且将在具体的实验中体现出真实的现象.

1.1 初态的制备

为了使量子态的操作意义更为严格, 需要有量子态的制备过程 (preparation).

最简单的制备过程就是对测量结果作选择, 或者说, 选择性的测量.

例如, 让一束经过准直的银原子通过 SG(z) —— 磁场沿 z 方向的 SG 装置, 然后在出射粒子中选择向 $+z$ 方向偏转的那一束, 我们就说制备了量子态 $|\uparrow\rangle$, 有时我们也把这个量子态标记为 $|0\rangle$, $|z+\rangle$.

制备过程是用经典的语言描述的. 制备过程的描述实际上就是实验操作流程. 在这个描述中, 你需要告诉别人用到的实验参数是什么, 对结果进行了怎样的选择. 比如在 SG 实验中, 需要指明磁场的指向, 磁场的非均匀度 (梯度), 磁场区域

的大小等等.

所谓的 "制备出了大量的处于某个特定状态的粒子", 指的是, 这些粒子来自相同的制备过程 —— 相同的用经典语言描述的实验过程.

因为这个实验过程是可控的, 在理想情形下, 可以认为制备出来的粒子处于相同的状态 —— 关于该制备过程而言的相同的状态. 例如, 在 SG 实验中, 我们可以控制粒子的偏转角度, 又选择了向上偏转的粒子束, 就可以认为, 我们得到了关于自旋自由度的 $|0\rangle$ 态, 甚至还可以说, 得到了粒子的空间位置波函数 $\psi(\boldsymbol{r})$, 但是, 我们不知道银原子的能级, 不知道电子处于怎样的原子结构中的量子态.

制备过程把一些经典信息植入了量子态, 或者说植入了量子系统.

经典情形下的正交性可以被植入到非正交的量子态中.

量子态的制备过程以及对制备过程的描述为我们提供了量子态的一些经典 "性质". 例如, 如果我们让制备好的处于 $|\uparrow\rangle$ 状态的粒子通过 SG(z), 那么所有的出射粒子都将偏向 $+z$ 方向, 确定地而不是随机地.

制备初态就是在特定的表象中制备确定的结果. 这些制备好的量子系统将面临以后的测量和操作.

1.2 量子态的数学表示

复 Hilbert 空间中的归一化的向量.

用到的类比是强度和振幅之间的关系. 如今是几率和几率幅. 有一些素材可以帮助我们建立这样的类比.

波粒二象性

在某些情况下, 微观世界中发生的事情在经典世界中表现出粒子性; 而在另一些情况下, 又可以表现出波动性.

对于波动性, 类比于光的电磁理论, 强度 I 与电场 $\mathbf{E}$ 之间的关系是 $I \sim |\mathbf{E}|^2$. 于是设想几率也有类似的对应, 即几率幅.

SG 序列

波动性还体现在包含三个 SG 装置的观测结果中. 考虑一个自旋为 1/2 的粒子, 通过接连放置的 SG 装置, 如图 1 所示.

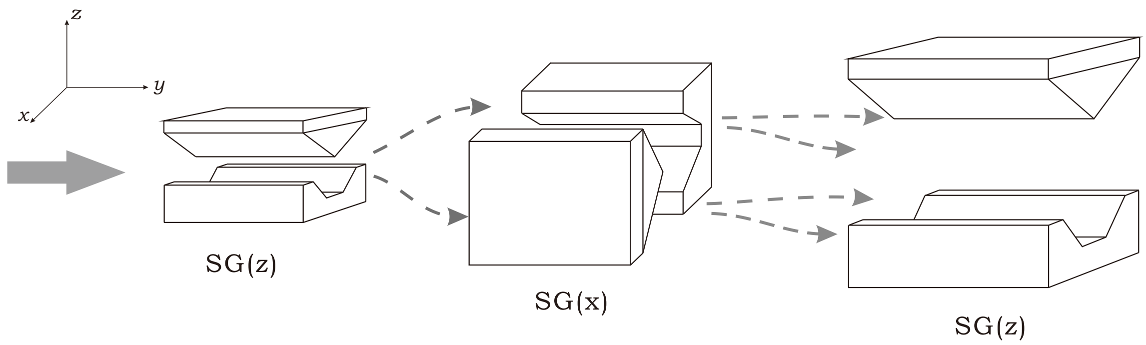

图 1

图中的从左到右的第一个和第三个 SG 装置的梯度磁场的方向为 z 方向, 记作 SG(z). 第二个 SG 装置的梯度磁场方向为 x 方向, 记作 SG(x). 图中的虚线表示假想的反事实的 (counterfactual) 的 "路径". 唯一的接收屏放在第三个 SG 装置的右侧. 观测到的实验结果如图 2 所示.

虽然我们并不能确定粒子在各个 SG 装置之间的路径, 但是图 2 的结果符合通常的直觉.

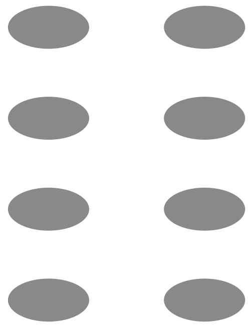

图 2

现在, 逐步减弱 SG(x) 的磁场强度, 直至为零.

如果把微观粒子当做经典粒子看待, 如果把 counterfactual 的推断当作事实本身, 那么将预言图 3 所示的结果.

实际结果如图 4 所示. 当 SG(x) 中的磁场强度减为零时, 结果是图 5. 图 4 和图 5 体现了干涉现象, 这类现象是不能用 "粒子通过了某条路径" 这类描述来解释的.

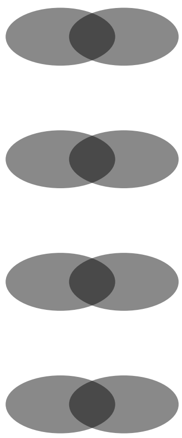

图 3

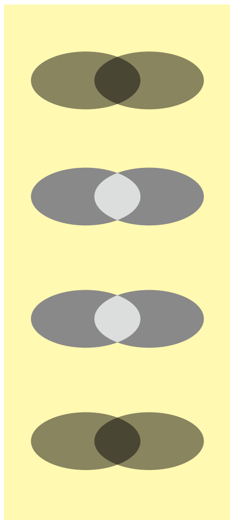

图 4

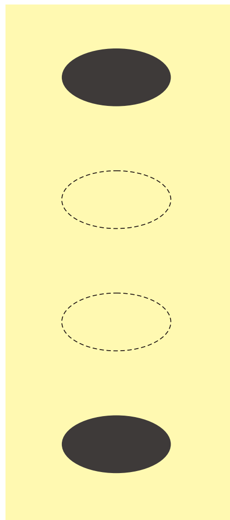

图 5

类比于以上素材, 将量子态表示为向量形式, 更具体地说, 是 Hilbert 空间中的向量. 空间的维数等于可以严格区分的现象的最大个数. 在 SG 实验中, 可以严格区分的现象只有两个, 所以需要最简单的两维复空间 $\mathbb{C}^2$ 描述量子态. 两维实空间不能胜任, 参看 Sakurai 书.

2 观测量

观测量被表示为 Hilbert 空间中的厄密算子, 在有限维情形下, 表示为厄密矩阵. 这是量子力学的一条基本假设. 考虑一下有哪些素材能够帮助我们理解这一假设.

2.1 基本物理量是时空变换的生成元

非相对论情形下, Galilei 时空变换有 10 个生成元, 分别对应于位置, 动量, 角动量和哈密顿量. 这些生成元支配了不同的变换, 变换将作用于量子态. 例如, 哈密顿量支配了时间演化, 以后将要讨论的时间演化算子的形式是 $e^{-iHt/\hbar}$.

量子态被表示为 Hilbert 空间中的向量, 而变换作用于向量, 则变换和生成元具有算子的形式.

2.2 厄密算子的数学性质

厄密算子的本征值是实数, 本征向量彼此正交. 厄密算子的这些基本性质可以用来描述测量过程.

测量结果应该用实数表示, 这可以对应于厄密算子的本征值. 不同的测量结果可以严格区分, 这对应于厄密算子的彼此正交的本征向量.

例如, 在 SG 实验中, 根据看到的现象, 为微观粒子构造出了一个叫做*磁矩*或*自旋角动量*的观测量. 在理想情形下, 屏幕上亮点的位置对应于观测量的本征值. 落在不同位置上的亮点可以严格区分, 在量子力学的数学形式中, 它们对应于彼此正交的量子态.

至此, 我们谈论了量子力学的两条基本假设:

1. 量子态被表示为 Hilbert 空间中的向量.
2. 观测量被表示为 Hilbert 空间中的厄密算子.

进一步的讨论需要线性空间的基本知识, 下面作简单介绍, 并借此引入 Dirac 符号.

3

有限维复空间

这里及以后, 我们将有限维和无限维两种情形都统称为 Hilbert 空间. 在以后的位置表象及空间波函数的讨论中叙述无限维 Hilbert 空间, 目前, 仅关注有限维 Hilbert 空间.

3.1 线性空间的基本概念

数域 $\mathbb{F}$, 空间 V. 这里及以后, $\mathbb{F} = \mathbb{C}$, 即, 讨论复数域上的线性空间.

如果数 $a, b \in \mathbb{C}$, 向量 $\psi, \varphi, \chi \in V$, 那么有

- $a\psi + b\varphi \in V$,
- $\psi + \varphi = \varphi + \psi$, $\psi + (\varphi + \chi) = (\psi + \varphi) + \chi$,
- $a(b\psi) = (ab)\psi$,
- $a(\psi + \varphi) = a\psi + b\varphi$, $(a + b)\psi = a\psi + b\varphi$,
- 存在零向量 $\mathbf{0}$, 对于任意的 $\psi \in V$, 有 $\mathbf{0} + \psi = \psi$,
- 数域中存在 0 和 1, 对于任意的 $\psi \in V$, 有 $0 \cdot \psi = \mathbf{0}$, 以及 $1 \cdot \psi = \psi$.

范数 (norm)

对于任意的 $\psi \in V$, 定义相应的一个实数 $\|\psi\|$, 称之为 ψ 的范数, 它满足

- 对于所有的 $\psi \neq \mathbf{0}$, $\|\psi\| > 0$. 当且仅当 $\psi = \mathbf{0}$, $\|\psi\| = 0$.
- 用 a 数乘 ψ, 得到 $a\psi$, 它的范数是
$$\|a\psi\| = |a|\|\psi\|$$

内积 (inner product)

对于任意两个向量 $\psi, \varphi \in V$, 定义一个数 $(\psi, \varphi) \in \mathbb{C}$, 称为 ψ 和 φ 之间的内积, 须满足如下要求:

- $(\psi, \psi) \geqslant 0$, 其中当且仅当 $\psi = \mathbf{0}$, 等号成立.
- $(\psi, \varphi + \chi) = (\psi, \varphi) + (\psi, \chi)$.
- $(\psi, a\varphi) = a(\psi, \varphi)$.
- $(\psi, \varphi) = (\varphi, \psi)^*$

由最后两个条件可以推出 $(a\psi, \varphi) = a^*(\psi, \varphi)$.

如果两个向量的内积为零, 那么它们彼此正交.

可以通过内积定义范数:
$$\|\psi\| = (\psi, \psi)^{1/2}.$$

不过, 需要知道的是, 范数的定义可以不依赖于内积.

两个向量之间的距离

利用范数定义两个向量 ψ 和 φ 之间的距离,
$$d(\psi, \varphi) = \sqrt{(\psi - \varphi, \psi - \varphi)} = \|\psi - \varphi\|.$$

两个不等式

线性空间中的向量满足如下两个不等式,
$$|(\psi, \varphi)| \leqslant \|\psi\|\,\|\varphi\|, \qquad \text{Cauchy-Schwartz 不等式},$$
$$\|\psi + \varphi\| \leqslant \|\psi\| + \|\varphi\|, \quad \text{三角形不等式}.$$

完备性

内积空间 V 中存在一组正交归一的向量, 记作 $\{e_i\}$, $i = 1, 2, \cdots, N$, N 是空间 V 的维数, 即 V 中线性无关的向量的最大个数.

$\{e_i\}$ 构成了空间 V 的一组基, 基向量的正交归一性表示为
$$(e_i, e_j) = \delta_{ij}$$

所谓的完备性, 可以这么说: V 中的任意一个向量都可以在给定的基上展开:

$$\psi = \sum_{i=1}^{N} c_i e_i$$

复数 c_i 是向量 ψ 在基向量 e_i 上的分量,

$$c_i = (e_i, \psi)$$

有限维复空间一定是完备的内积空间, 因此, 在有限维情形下, 可以不提完备性. 但是, 在无限维情形下, 只有完备的内积空间才是 Hilbert 空间.

3.2 直和与直积

两个 Hilbert 空间, $\mathcal{H}_1$ 和 $\mathcal{H}_2$, 它们的维数分别为 M 和 N. 它们的直和 $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ 是一个 $M+N$ 的空间.

设 $\psi = \sum_i \psi_i e_i \in \mathcal{H}_1$, $\varphi = \sum_j \varphi_j f_j \in \mathcal{H}_2$, $\{e_i\}$ 和 $\{f_j\}$ 分别是 $\mathcal{H}_1$ 和 $\mathcal{H}_2$ 的基.

直和 $\Psi = \psi \oplus \varphi \in \mathcal{H}$,

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_M \end{pmatrix} \oplus \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_N \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_M \\ \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_N \end{pmatrix}.$$

$\mathcal{H}_1$ 和 $\mathcal{H}_2$ 的直积空间 $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, 其中的向量是 (以 $\mathbb{C}^3 \otimes \mathbb{C}^2$ 为例):

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} \otimes \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \begin{pmatrix} \psi_1\varphi_1 \\ \psi_1\varphi_2 \\ \psi_2\varphi_1 \\ \psi_2\varphi_2 \\ \psi_3\varphi_1 \\ \psi_3\varphi_2 \end{pmatrix}.$$

直积空间将用来表示两体 (以至多体) 量子系统的量子态.

3.3 子空间, 正交子空间

设 $\mathcal{M}$ 是 $\mathcal{H}$ 的一个子空间, 也就是说, $\mathcal{M}$ 中的向量的任意形式的线性组合都属于 $\mathcal{M}$.

整个空间 $\mathcal{H}$ 可以表示为

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}.$$

这里, $\mathcal{M}^{\perp}$ 是 $\mathcal{H}$ 的另一个子空间, $\mathcal{M}^{\perp}$ 中的任意一个向量与 $\mathcal{M}$ 中的任意向量是正交的, 即

$$(\psi, \varphi) = 0, \quad \forall\ \psi \in \mathcal{M}, \varphi \in \mathcal{M}^{\perp}.$$

$\mathcal{M}^{\perp}$ 是 $\mathcal{M}$ 的正交子空间.

$\mathcal{H}$ 的中的任意一个向量 Ψ 可以表示为

$$\Psi = \psi \oplus \varphi, \quad \text{其中}\ \psi \in \mathcal{M}, \varphi \in \mathcal{M}^{\perp}.$$

4

线性算子

4.1 线性算子构成线性空间

线性算子 T 将 Hilbert 空间 $\mathcal{H}$ 中的向量变换为另一个 Hilbert 空间 $\mathcal{H}'$ 中的向量. 这些线性算子也构成了一个线性空间, 记作 $\mathcal{L}(\mathcal{H}, \mathcal{H}')$, 其线性性表现为

$$T(a\psi + b\varphi) = aT(\psi) + bT(\varphi).$$

算子 T 的范数有多种定义, 例如,

$$\|T\| = \sup_{\|\psi\|=1} \|T(\psi)\|_{\mathcal{H}'}.$$

上式右端的意思是, 原先属于 $\mathcal{H}$ 的向量 ψ 被变换到 $\mathcal{H}'$ 中, 其范数为 $\|T(\psi)\|_{\mathcal{H}'}$. 考虑所有属于 $\mathcal{H}$ 的单位向量, 最大的 $\|T(\psi)\|_{\mathcal{H}'}$ 即是 T 的范数.

如果 $\mathcal{H} = \mathcal{H}'$, 那么 $\mathcal{L}(\mathcal{H}, \mathcal{H}')$ 可以记作 $\mathcal{L}(\mathcal{H})$, 而 T 就被称为 $\mathcal{H}$ 上的线性算子.

有限维复空间 $\mathbb{C}^N$ 上的线性算子具有矩阵形式, 是 $N \times N$ 的矩阵.

4.2 对偶空间

一个特殊的情形: $\mathcal{L}(\mathcal{H}, \mathbb{C})$, 即, $\mathcal{H}$ 中的向量被变成了复数, 这是 $\mathcal{H}$ 上的线性泛函. $\mathcal{L}(\mathcal{H}, \mathbb{C})$ 构成的空间被称为 $\mathcal{H}$ 对偶空间, 记作 $\mathcal{H}^{\dagger}$.

Riesz 引理

对于每一个 $T \in \mathcal{H}^{\dagger}$, 有唯一的 $\varphi_T \in \mathcal{H}$, 使得对于所有的 $\psi \in \mathcal{H}$, $T(\psi) = (\varphi_T, \psi)$.

一般来说, 将向量变为数的变换有多种形式, 但是 Riesz 引理告诉我们, 不论什么形式的选择, 到后来总能落实为内积的形式.

换一个说法: 对于 $\mathcal{H}$ 中的每一个向量 φ, 总可以定义一个连续的线性泛函 $T_{\varphi} \in \mathcal{H}^{\dagger}$,

$$T_{\varphi}(\psi) = (\varphi, \psi).$$

对偶空间将用来定义左矢空间.

4.3 Self-adjoint, 或者简单地说, 厄密共轭

设 ψ 和 φ 是 $\mathscr{H}$ 中任意两个向量, X 是 $\mathscr{H}$ 上的线性算子, 即 $X \in \mathcal{L}(\mathscr{H})$. 用 X 作用于 ψ, 将 $X(\psi)$ 简记为 $X\psi$, 显然 $X\psi \in \mathscr{H}$. 考虑内积 $(\varphi, X\psi)$, 线性算子 X 的厄密共轭 $X^\dagger$ 定义为

$$(X^\dagger\varphi, \psi) = (\varphi, X\psi) \tag{1}$$

如果 $X = X^\dagger$, 那么 X 是厄密算子.

在有限维空间中, 容易验证

$$X^\dagger = (X^*)^T$$

厄密共轭的存在性

上面给出了厄密共轭的定义, 还需要证明确实存在满足该定义的算子. Riesz 定理指出, 内积运算可以视作泛函,

$$T_\varphi(\cdot) = (\varphi, \cdot)$$

其中 φ 是 Hilbert 空间 $\mathscr{H}$ 中的向量, $\varphi \in \mathscr{H}$. 与 φ 一一对应的 T_φ 是对偶空间 $\mathscr{H}^\dagger$ 中的泛函, $T_\varphi \in \mathscr{H}^\dagger$. 空间 $\mathscr{H}^\dagger$ 也可以称为泛函空间.

现在, 线性算子 X 作用于 $\mathscr{H}$ 中的向量 ψ, 得到 $X\psi \in \mathscr{H}$, 内积 $(\varphi, X\psi)$ 可以表示为 $T_\varphi(X\psi)$. 换一种方式说, $T_\varphi(X\psi)$ 可以看作对 ψ 进行两个线性变换, 先 X, 然后 T_φ, 即

$$T_\varphi(X\psi) = (T_\varphi \circ X)\psi$$

因此, $T_\varphi(X\psi)$ 给出了一个新的泛函, 记作 F_φ,

$$F_\varphi(\psi) = (\varphi, X\psi)$$

注意, 泛函 T_φ 的定义是 $T_\varphi(\psi) = (\varphi, \psi)$.

还是根据 Riesz 定理, 对于泛函 F_φ, 一定存在一个向量 $\chi \in \mathscr{H}$, 使得

$$F_\varphi(\psi) = (\chi, \psi) = T_\chi(\psi) \tag{2}$$

在算子 X 给定的前提下, 向量 χ 由 φ 唯一确定, 因此, 一定存在一个线性算子, 记作 $X^\dagger$, 使得 $\chi = X^\dagger\varphi$. 于是 (2) 式中的两个泛函有如下关系,

$$F_\varphi = T_{X^\dagger\varphi}$$

这也就是说

$$(\varphi, X\psi) = F_\varphi(\psi) = T_{X^\dagger\varphi}(\psi) = (X^\dagger\varphi, \psi)$$

即 (1) 式.

上述过程中, 画波浪线的地方表明, X 的厄密共轭是存在的.

5 Dirac 符号

在这一节中, 引入 Dirac 符号, 将上述线性空间的基本内容重述一遍.

5.1 右矢和左矢

ψ 是 $\mathscr{H}$ 的中向量, 现在用 Dirac 符号记作 $|\psi\rangle$, 这便是右矢 (ket). 右矢 $|\psi\rangle$ 与 ψ 没有没有本质上的区别, 仅仅是换了一种写法.

为了说明左矢 (bra) 和左矢空间, 我们回到内积的定义. 对于 $\psi, \varphi \in \mathscr{H}$, 二者的内积是

$$(\varphi, \psi) \in \mathbb{C}$$

根据 Riesz 定理, 内积运算可以视作一个操作, 它是一个把向量 ψ 变成复数的线性泛函.

$$(\varphi, \psi) \iff T_\varphi(\psi)$$

T_φ 是一个关于 ψ 的线性泛函, 即 $T_\varphi \in \mathcal{L}(\mathscr{H}, \mathbb{C})$. 或者说, T_φ 是对偶空间 $\mathscr{H}^\dagger$ 中的一个 "向量", $T_\varphi \in \mathscr{H}^\dagger$. Riesz 定理还告诉我们, $\mathscr{H}$ 中的 φ 与 $\mathscr{H}^\dagger$ 中的 T_φ 是一一对应的, $\mathscr{H}$ 和 $\mathscr{H}^\dagger$ 是同构的 (isomorphic).

用左矢形式 $\langle\varphi|$ 表示 T_φ, 这实际上是泛函 T_φ 的另一种写法, 于是有

$$(\varphi, \psi) \iff T_\varphi(\psi) \iff \langle\varphi|\psi\rangle \in \mathbb{C}$$

在有限维空间中, 可以把右矢 $|\psi\rangle$ 看成是列向量, 把左矢 $\langle\psi|$ 看成行向量, 并且行向量 $\langle\psi|$ 中的每一个分量都是列向量 $|\psi\rangle$ 中相应的分量的复共轭.

在有限维空间 $\mathbb{C}^N$ 中, 存在一组正交归一的基向量, $|e_i\rangle$, $i = 1, 2, \cdots, N$, 正交归一性体现为

$$\langle e_i|e_j\rangle = \delta_{ij}$$

其中 $\langle e_i|$ 是 $|e_i\rangle$ 的左矢形式.

在基向量 $|e_i\rangle$ 上, 或者说在表象 $\{|e_i\rangle\}$ 中, 任意向量 $|\psi\rangle$ 可以展开为

$$|\psi\rangle = \sum_{i=1}^{N} \psi_i |e_i\rangle, \quad \psi_i = \langle e_i|\psi\rangle \tag{3}$$

于是将向量 $|\psi\rangle$ 表示为

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} \tag{4}$$

与之对偶的左矢的形式是

$$\langle\psi| = \begin{pmatrix} \psi_1^* & \psi_2^* & \cdots & \psi_N^* \end{pmatrix} \tag{5}$$

设另有一个向量 $|\varphi\rangle$ 可以表示为

$$|\varphi\rangle = \begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_N \end{pmatrix}^T$$

其中上标 T 表示转置. 两个向量 $|\psi\rangle$ 和 $|\varphi\rangle$ 之间的内积是

$$\langle\psi|\varphi\rangle = \sum_{i=1}^{N} \psi_i^* \varphi_i = \langle\varphi|\psi\rangle^*$$

关于向量 $|\psi\rangle$ 的形式 (3) 和 (4)

在基向量 $|e_i\rangle$ 上, 将向量 $|\psi\rangle$ 表示为 (3), 这是完整而准确的形式, 其中有分量 ψ_i, 也明确指明了基向量 $|e_i\rangle$. 虽然形式 (4) 更为形象直观, 但没有体现基向量, 故而有时会有引起误解.

在具体计算中, 尽可以使用 (4) 式, 但须牢记该形式是在什么基向量上展开的.

对于左矢及其行向量的形式, 也是如此.

5.2 算子的表示

设 $\{|e_i\rangle\}$ 是 Hilbert 空间 $\mathscr{H}$ 的一组基, X 是 $\mathscr{H}$ 上的线性算子, $X \in \mathcal{L}(\mathscr{H})$. X 可以表示为

$$X = \sum_{i,j} x_{ij} \, |e_i\rangle\langle e_j| \tag{6}$$

对于有限维空间, 上式对 i, j 的求和是从 1 到 $N = \dim \mathscr{H}$; 对于无限维空间, 求和的上限是 ∞.

(6) 表明, 算子 X 在第 i 行第 j 列的的矩阵元是 x_{ij}, 该矩阵元又可以表示为

$$x_{ij} = \langle e_i|X|e_j\rangle$$

为了说明这一点, 可以考虑有限维空间 $\mathbb{C}^N$, 引入如下形式的自然基向量,

$$|e_1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |e_2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad |e_N\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

一般地, 我们并不建议把基向量展开成某个具体的形式, 这是因为, 必须先有了一组基, 然后才能说展开形式. 基向量本来就是为了展开其它向量而设的, 原则上不存在用来展开基向量的更为基本的向量了. 但是, 在很多具体场合中, 我们还是写出了基向量的具体形式, 使得计算过程更清楚, 但这只是一种实用手段.

现在考虑两个自然基向量头对头的形式, $|e_i\rangle\langle e_j|$. 该形式是一个矩阵, 第 i 行第 j 列的矩阵元为 1, 其余皆为零. 所以, $|e_i\rangle\langle e_j|$ 构成了 $\mathbb{C}^N$ 上的矩阵的基.

以 $\mathbb{C}^2$ 为例

$\mathbb{C}^2$ 的两个自然基向量是

$$|e_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |e_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

它们的头对头的形式是,

$$|e_1\rangle\langle e_1| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad |e_1\rangle\langle e_2| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad |e_2\rangle\langle e_1| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad |e_2\rangle\langle e_2| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

给定某个矩阵 $X = (x_{ij})$, 第 i 行第 j 列的矩阵元是 x_{ij}, 可以表示为

$$X = \sum_{i,j=1}^{N} x_{ij} \, |e_i\rangle\langle e_j|$$

单位阵可以表示为

$$\mathbb{1}_N = \sum_{i=1}^{N} |e_i\rangle\langle e_i|$$

我们将经常利用这个形式, 例如

$$|\psi\rangle = \mathbb{1}_N\,|\psi\rangle = \sum_{i}^{N} |e_i\rangle\langle e_i|\ |\psi\rangle = \sum_{i}^{N} \psi_i\,|e_i\rangle$$

$$\langle\varphi| = \langle\varphi|\,\mathbb{1}_N = \sum_{i}^{N} \langle\varphi|e_i|\varphi|e_i\rangle\,\langle e_i| = \sum_{i}^{N} \varphi_i^*\,\langle e_i|$$

$$X = \mathbb{1}\,X\,\mathbb{1} = \cdots\ \cdots$$

求迹运算 (trace)

对于 $N \times N$ 的矩阵 X, 迹 $\mathrm{Tr}(x)$ 是对角元的和,

$$\mathrm{Tr}(X) = \sum_{i=1}^{N} x_{ii}$$

用 Dirac 符合表示,

$$\begin{aligned}\mathrm{Tr}(X) &= \sum_{i=1}^{N} \langle e_i|X|e_i\rangle \\ &= \sum_{i=1}^{N} \langle e_i|\left(\sum_{m,n=1}^{N} x_{mn}\,|e_m\rangle\langle e_n|\right)|e_i\rangle\end{aligned}$$

考虑了基向量的正交归一性之后, 得到 $\sum_{i=1}^{N} x_{ii}$.

转置

矩阵 X 的转置记作 X^T.

$$\begin{aligned}X^T &= \left(\sum_{ij} x_{ij}\,|e_i\rangle\langle e_j|\right)^T \\ &= \sum_{ij} x_{ij}\,|e_j\rangle\langle e_i| = \sum_{ij} x_{ji}\,|e_i\rangle\langle e_j|\end{aligned}$$

厄密共轭

(1) 式给出了厄密共轭的定义, 即 $(\varphi, X\psi) = (X^\dagger\varphi, \psi)$, 现在要用 Dirac 符号重新描述.

首先有

$$X\psi \Longleftrightarrow X\,|\psi\rangle$$

当然, $X\,|\psi\rangle \in \mathcal{H}$. 接着有

$$(\varphi, X\psi) \Longleftrightarrow \langle\varphi|\big(X\,|\psi\rangle\big) = \langle\varphi|X|\psi\rangle \tag{7}$$

这个形式的意思是, 算子 X 向右作用于右矢 $|\psi\rangle$, 得到另外某个向量, 然后再与 $|\varphi\rangle$ 作内积.

如果让 X 向左看, 那么它遇到的是左矢 $\langle\varphi|$, 而左矢对应于泛函. 在这种情况下, X 应该被理解为泛函空间 $\mathscr{H}^\dagger$ 上的线性算子, 即 $X \in \mathcal{L}(\mathscr{H}^\dagger)$. X 向左作用于左矢 $\langle\varphi|$, 得到另一个左矢 (或者说泛函), 记作 $\langle\chi|$, 即

$$\langle\chi| = \langle\varphi| X$$

Riesz 定理说, 在右矢空间 $\mathscr{H}$ 中, 一定有一个右矢 $|\varphi\rangle$ 与左矢 $\langle\varphi|$ 一一对应, 一定有一个右矢 $|\chi\rangle$ 与左矢 $\langle\chi|$ 一一对应. 既然 $\langle\varphi|$ 在线性线性算子 X 的作用下变为 $\langle\chi|$, 那么在 $\mathscr{H}$ 上一定有一个线性算子将它们各自对应的 $|\varphi\rangle$ 和 $|\chi\rangle$ 联系起来, 将这个 $\mathscr{H}$ 上线性算子记作 $X^\dagger$, 有

$$|\chi\rangle = X^\dagger |\varphi\rangle$$

上述关系体现在图 6 中.

$$\begin{tikzcd}
\langle\varphi| \arrow[rrr, leftrightarrow, "\text{Riesz 定理}"] \arrow[d, "X"'] & & & |\varphi\rangle \arrow[d, "X^\dagger"] \\
\langle\chi| \arrow[rrr, leftrightarrow, "\text{Riesz 定理}"] & & & |\chi\rangle
\end{tikzcd}$$

图 6

至此, 从 "X 向左看" 的观点出发, 得到了左矢空间中的 $\langle\chi|$ 以及它所对应的右矢空间中的 $|\chi\rangle$. 右矢 $|\chi\rangle$ 和 $|\psi\rangle$ 的内积是

$$\langle\chi|\psi\rangle = \langle\psi|\chi\rangle^* = \langle\psi|X^\dagger|\varphi\rangle^*$$

这个内积应该等于 "X 向右看" 给出的内积, 即 (7) 式, 从而有

$$\langle\psi|X^\dagger|\varphi\rangle^* = \langle\varphi|X|\psi\rangle$$

这就是用 Dirac 符号表示的算子的厄密共轭.

算子的厄密共轭有如下性质: 对于 $X, Y \in \mathcal{L}(\mathscr{H})$, 有

$$(cX)^\dagger = c^* X^\dagger, \quad c \in \mathbb{C}$$

$$(X+Y)^\dagger = X^\dagger + Y^\dagger$$

$$(XY)^\dagger = Y^\dagger X^\dagger$$

有限维空间上线性算子的厄密共轭很简单 —— 转置再取复共轭,

$$X^\dagger = \left(X^T\right)^*, \quad \text{for all } X \in \mathcal{L}(\mathbb{C}^N)$$

右矢和左矢头对头的形式, $|\psi\rangle\langle\varphi|$, 是一个算子, 它的厄密共轭是

$$\left(|\psi\rangle\langle\varphi|\right)^\dagger = |\varphi\rangle\langle\psi|$$

如果矩阵 H 满足 $H = H^\dagger$, 那么它是厄密矩阵, 并且

$$\langle\varphi|H|\psi\rangle = \langle\psi|H|\varphi\rangle^*$$

5.3 厄密矩阵的本征值和本征向量

根据线性代数的知识, 我们知道, 厄密矩阵的本征值是实数, 本征向量彼此正交, 进一步可以将本征向量归一化, 因此可以说, 厄密矩阵的本征向量是正交归一的.

以下讨论非简并情形, 就是说, 本征值和本征向量一一对应 (本征向量可以有单位复数的差别, 或者说相因子的差别).

设 A 是厄密矩阵, $A = A^\dagger$. 本征值和相应的本征向量分别是 a_i 和 $|\alpha_i\rangle$, 即

$$A\,|\alpha_i\rangle = a_i\,|\alpha_i\rangle\,,\quad a_i \in \mathbb{R},\quad i = 1,\cdots,N$$

本征向量正交归一, 即

$$\langle\alpha_i|\alpha_j\rangle = \delta_{ij}$$

$|\alpha_i\rangle$ 可以作为 $\mathbb{C}^N$ 的基向量 [1]. 选择 $\{|\alpha_i\rangle\}$ 作为基, 就是选择了 A 表象.

可以将 A 表示为

$$A = \sum_i a_i\,|\alpha_i\rangle\langle\alpha_i| \tag{8}$$

厄密矩阵 A 在其自身的表象中具有对角矩阵的形式. (8) 式又称为 A 的本征分解形式, 或谱分解形式.

A 可以有自然基向量上的表示, 即

$$A = \sum_{ij} a_{ij}\,|e_i\rangle\langle e_j|$$

其中 a_{ij} 是矩阵元, 区别于本征值 a_i. 将 A 写成 (8) 式, 就是对角化的过程.

在数学上, 表象是基向量的选择方式. 在物理上, 表象对应于为了观测特定的物理量而设定的测量方式. 如果量子系统的某个物理量可以通过厄密矩阵 A 描述, 那么为了分析 A 性质, 就要置身于 A 表象.

在 A 表象中, 任何一个向量 $|\psi\rangle$ 可以展开为

$$|\psi\rangle = \sum_i c_i\,|\alpha_i\rangle\,,\quad c_i = \langle\alpha_i|\psi\rangle$$

如果希望讨论另一个性质 —— 用另一个厄密矩阵 B 表示的性质, 那么就要选择 B 表象.

$$B = \sum_j b_j\,|\beta_j\rangle\langle\beta_j|$$

$$|\psi\rangle = \sum_j d_j\,|\beta_j\rangle\,,\quad d_j = \langle\beta_j|\psi\rangle$$

同一个向量 $|\psi\rangle$ 在不同的表象中的表示是不同的. 涉及到表象的变换, 酉变换.

5.4 酉变换

$\mathbb{C}^n$ 上的酉变换的定义是, 对于 $\psi, \varphi \in \mathbb{C}^n$, 以及 $U \in \mathcal{L}(\mathbb{C}^n)$, 如果

$$\big(U(\varphi), \psi\big) = \big(\varphi, U^{-1}(\psi)\big) \tag{9}$$

[1] 如果存在简并, 则考虑简并子空间中的正交归一化.

那么, 算子 U 描述了一个酉变换.

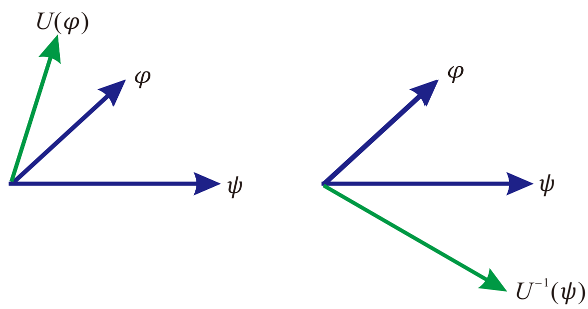

图 7: (9) 式的形象描述.

或者简单地说, 满足 $UU^\dagger = U^\dagger U = \mathbb{1}$ 的矩阵 U 是酉矩阵.

如果 $\mathcal{U}$ 是有限维空间 $\mathbb{C}^n$ 上的酉矩阵的集合, 那么

- 对于 $c \in \mathbb{C}$, $|c| = 1$ 和 $U \in \mathcal{U}$, 有 $cU \in \mathcal{U}$.
- 如果 $U, V \in \mathcal{U}$, 那么 $UV \in \mathcal{U}$.
- 如果 $U \in \mathcal{U}$, 那么 $U^{-1} \in \mathcal{U}$.
- U 是酉矩阵, 当且仅当它将一组正交归一基向量变为另一组正交归一的基向量.
- 如果 U 是酉矩阵, 那么它的本征值的模为 1.
- 如果 U 是酉矩阵, 那么它的行向量或者列向量是正交归一的.
- 酉矩阵 U 的行列式满足 $|\det(U)| = 1$.

集合 $\mathcal{U}$ 具有群结构, 是 U(n) 群.

$\mathbb{C}^n$ 上的酉矩阵有 n^2 个独立的实参数.

$$n \times n \text{ 矩阵}, n^2 \text{ 个矩阵元}, 2n^2 \text{ 个实数}$$

$$\text{每一行或者每一列归一} \implies n \text{ 个限制条件}$$

$$\text{不同行或不同列正交} \implies \frac{n(n-1)}{2} \times 2 \text{ 个限制条件}$$

$$\therefore\ 2n^2 - n - \frac{n(n-1)}{2} \times 2 = n^2$$

$\mathbb{C}^2$ 上的酉矩阵可以表示为

$$U = e^{i\gamma} \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}$$

其中 $\gamma \in \mathbb{R}$, $a, b \in \mathbb{C}$, 且 $|a|^2 + |b|^2 = 1$. 单位复数 $e^{i\gamma}$ 也具有群结构, 是 U(1) 群.

$$\mathrm{U}(2) \simeq \mathrm{U}(1) \times \mathrm{SU}(2)$$

SU(2) 是 special unitary group. 对于 $U \in \mathrm{SU}(2)$, 有 $\det(U) = 1$.

Wigner 定理

对于变换 $|\psi'\rangle = U|\psi\rangle$, $|\varphi'\rangle = U|\varphi\rangle$, 如果 $|\langle\psi|\varphi\rangle| = |\langle\psi'|\varphi'\rangle|$, 那么 U 是酉的 (线性的) 或者反酉的 (反线性的).

6

不同表象之间的变换

在空间的基向量 (或者说表象) 改变的时候量子态和力学量的表示如何变化, 这是被动观点. 主动观点说的是量子态的演化, 或者说, 用某个变换直接作用于量子态. 量子态的演化分为两种类型: 酉演化和非酉演化. 前者是由 Schrödinger 方程决定的, 后者涉及开放量子系统的演化 (典型的例子是量子测量过程).

目前讨论的是被动观点.

先复习线性代数. 向量 v 可以在两组不同的基向量上展开.

$$v = \sum_i v_i e_i = \sum_j v'_j f_j.$$

其中 $\{e_i\}$ 是一组基向量, $\{f_j\}$ 是另一组基向量. 现在需要考虑的是, 在两组不同的基上, 分量 v_i 和 v'_j 之间有怎样的联系?

设两组基向量之间的关系是

$$(f_1, f_2, \cdots, f_N) = (e_1, e_2, \cdots, e_N)\ T$$

$$f_j = \sum_i T_{ij}\, e_i \tag{10}$$

$$\vec{f} = \vec{e}\,T, \quad \vec{e} = \vec{f}\,T^{-1}. \tag{11}$$

这里假设线性变换 T 是满秩的, 存在逆变换 T^{-1}.

将上述形式代入 v 的表达式

$$\begin{aligned} v &= \sum_{i,j} v_i\ f_j \left(T^{-1}\right)_{ji} \\ &= \sum_j \left[\sum_i \left(T^{-1}\right)_{ji} v_i\right] f_j \\ &= \sum_j v'_j f_j. \end{aligned}$$

$$v'_j = \sum_i \left(T^{-1}\right)_{ji} v_i.$$

这就是在基向量 $\{f_j\}$ 上 v 的表示.

简单地说,

$$v = \begin{pmatrix} e_1 & e_2 & \cdots & e_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & \cdots & e_n \end{pmatrix} T\, T^{-1} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix}$$

$$= \sum_j v'_j f_j$$

例如, 在 $\mathbb{R}^2$ 中, 坐标架 xy 和 $x'y'$ 之间的关系是

$$e'_x = e_x \cos\varphi + e_y \sin\varphi$$

$$e'_y = -e_x \sin\varphi + e_y \cos\varphi$$

也就是

$$\begin{pmatrix} e'_x & e'_y \end{pmatrix} = \begin{pmatrix} e_x & e_y \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

在 $x'y'$ 坐标架中, 向量的分量的变换是由矩阵 $\left(\begin{smallmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{smallmatrix}\right)$ 决定的.

用 Dirac 符号叙述. $\{|\alpha_i\rangle\}$ 是一组基向量, 称为 A 表象. $\{|\beta_j\rangle\}$ 是另一组基向量, 称为 B 表象. 从 A 表象变换到 B 表象. 建立两组基向量之间的联系.

$$|\beta_j\rangle = \sum_i u_{ij} |\alpha_i\rangle, \quad U = (u_{ij}) \tag{12}$$

$$u_{ij} = \langle\alpha_i|\beta_j\rangle, \quad U = \sum_i |\beta_i\rangle\langle\alpha_i| \tag{13}$$

基向量的变换矩阵 U 是在 A 表象中书写的. 应该如此, 我们应该知道将要变换到的基向量的具体形式, 而且是在 A 表象中的形式.

基向量变换的另一种等价形式

从 $\{|\alpha_i\rangle\}$ 到 $\{|\beta_j\rangle\}$ 的变换也可以写为

$$U|\alpha_j\rangle = |\beta_j\rangle$$

U 的矩阵元可以表示为

$$u_{ij} = \langle\alpha_i|U|\alpha_j\rangle = \langle\alpha_i|\beta_j\rangle$$

与 (13) 一致. 这里, 可以采取实用的观点, 将 $|\alpha_i\rangle$ 和 $|\beta_i\rangle$ 看作列向量. 例如, 假设 $|\alpha_j\rangle$ 是自然基向量 $|e_j\rangle$. 将 $|e_j\rangle$ 视作列向量, 其中第 j 行为 1, 其余都是 0.

$$U|e_j\rangle = \begin{pmatrix} u_{1j} \\ u_{2j} \\ \vdots \\ u_{nj} \end{pmatrix} = \sum_i u_{ij} |e_i\rangle = |\beta_j\rangle$$

与 (10) 一致.

容易证明 $UU^\dagger = U^\dagger U = \mathbb{1}$, 变换 U 是酉变换.

形式 $U = \sum_i |\beta_i\rangle\langle\alpha_i|$ 的意思很明确: 将 $|\alpha_i\rangle$ 映射为 $|\beta_i\rangle$.

重申: 这里的 U 是对基向量的变换, 在新的表象中, 向量的分量的变换是由 $U^\dagger$ 决定的.

$$
\begin{aligned}
|\psi\rangle &= \sum_i c_i |\alpha_i\rangle \\
&= \sum_{i,j} c_i \; |\beta_j\rangle\langle\beta_j| \; |\alpha_i\rangle \\
&= \sum_j \Big(\sum_i u_{ij}^* c_i \Big) \; |\beta_j\rangle \\
&= \sum_j \Big[\sum_i \left(U^\dagger\right)_{ji} c_i \Big] \; |\beta_j\rangle
\end{aligned}
$$

$\mathbb{C}^2$ 空间中的一个例子

在 $\mathbb{C}^2$ 空间中, 定义如下形式的两个观测量,

$$
\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
$$

我们之所以能写出观测量的矩阵形式, 是因为已经为 $\mathbb{C}^2$ 空间设定了基向量. 实际上, 它们正是 σ_z 的两个本征向量.

σ_z 的两个本征值是 $+1$ 和 -1, 对应的本征向量分别是

$$
|z+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |z-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
$$

从实用的观点来说, 它们可以看作自然基向量.

如果想变换到 σ_x 表象, 那么首先需要了解 σ_x 的本征向量. 容易看出, σ_x 的本征值也是 $+1$ 和 -1, 相应的本征向量分别是

$$
|x+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |x-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}
$$

从 σ_z 表象变换到 σ_x 表象, 变换矩阵是

$$
U = |x+\rangle\langle z+| + |x-\rangle\langle z-| = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \text{Hadamard matrix}
$$

在 σ_x 表象中向量的分量由 $U^\dagger$ 决定. 例如, 假设某个量子态在 σ_z 表象中的形式是

$$
|\psi\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}
$$

注意, 这里的 $|0\rangle$ 不是基向量. 该量子态在 σ_x 表象中的形式变为

$$U^\dagger \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

这里, 重申以前说过的话: 将量子态表示为列向量的时候, 一定要注意所处的表象. 上式右端的形式绝不是 $|x+\rangle$, 而是 $|x\rangle$ 和 $|x-\rangle$ 的叠加, 即 $\frac{1}{\sqrt{2}}(|x+\rangle + |x-\rangle)$.

仅仅从列向量的形式看, $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ 当然不能等于 $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, 但是, 它们是同一个向量在不同表象中的表示形式, 即

$$|\psi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\sigma_z} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\sigma_x}$$

其中的下标表示所处的表象.

再考虑表象变换下力学量的表述形式的变换. 在 A 表象中, 矩阵 X 写为

$$X = \sum_{ij} x_{ij} \, |\alpha_i\rangle\langle\alpha_j| = X^{(A)}$$

写出在 B 表象中的形式.

$$\begin{aligned} X &= \sum_{\substack{i,j \\ m,n}} x_{ij} \, |\beta_m\rangle\langle\beta_m| \, |\alpha_i\rangle\langle\alpha_j| \, |\beta_n\rangle\langle\beta_n| \\ &= \sum_{\substack{i,j \\ m,n}} u_{im}^* \, x_{ij} \, u_{jn} \; |\beta_m\rangle\langle\beta_n| \end{aligned}$$

在 B 表象中的矩阵形式记作 $X^{(B)}$.

$$X^{(B)} = U^\dagger X^{(A)} U.$$

例如, 将 σ_z 表象中的 σ_x 变换到 σ_x 表象中,

$$\sigma_x^{(X)} = U^\dagger \sigma_x^{(Z)} U = \mathrm{diag}(1, \ -1)$$

在自身的表象中是对角的, 这也就是在线性代数中说的厄密矩阵的对角化.

当然, 对于一般形式的矩阵, 酉变换未必能使其对角化. 介绍下面的两种分解形式.

第一种, polar decomposition. X 是任意的方阵.

$$X = PU \tag{14}$$

其中 P 是 (半) 正定矩阵, U 是酉矩阵. 这可以类比于将复数 z 表示为 $z = re^{i\phi}$. 既然 P 是半正定的厄密矩阵, 总可以通过酉变换将其对角化, 于是有下面的分解形式.

第二种, Singular value decomposition.

$$X = V \Lambda W^\dagger.$$

Λ 是对角的实数矩阵, 对角元非负. V 和 W 是酉的.