

第三章 双值量子系统

主要内容

1. \mathbb{C}^2 空间中量子态的表示, Bloch 向量.
2. 双值量子系统的观测量在本质上相当于 Pauli 矩阵在某个方向上的分量.
3. \mathbb{C}^2 空间上的酉变换等价于旋转变换.
4. 将 Born 规则用于双值量子系统.

凡是可以用两维复空间 \mathbb{C}^2 描述的量子系统都称为双值量子系统. 包括自旋 1/2 的粒子, 两能级系统等等. 双值指的是, 对这一类量子系统进行测量, 得到的结果有且仅有两个.

双值量子系统是最简单的量子系统.

1

纯态

纯态 $|\psi\rangle \in \mathbb{C}^2$, 且 $\langle\psi|\psi\rangle = 1$. 可以在自然基向量上表示为

$$|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle, \quad c_0, c_1 \in \mathbb{C}, \quad |c_0|^2 + |c_1|^2 = 1$$

有三个独立的实参数. 于是将 $|\psi\rangle$ 写为

$$|\psi\rangle = e^{i\gamma} \begin{pmatrix} \cos \frac{\theta}{2} e^{-i\frac{\phi}{2}} \\ \sin \frac{\theta}{2} e^{i\frac{\phi}{2}} \end{pmatrix} \quad (1)$$

其中 $e^{i\gamma}$ 是整体相因子. 把 $|\psi\rangle$ 表示为密度矩阵的形式,

$$\psi = |\psi\rangle\langle\psi| = \begin{pmatrix} \cos^2 \frac{\theta}{2} & \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\phi} \\ \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{i\phi} & \cos^2 \frac{\theta}{2} \end{pmatrix}$$

可以看到, 整体相因子不出现在密度矩阵中.

进而用 $\sigma_0 = \mathbb{1}, \sigma_1 = \sigma_x, \sigma_2 = \sigma_y, \sigma_3 = \sigma_z$ 展开, 有

$$\begin{aligned} \psi &= \frac{1}{2}(\mathbb{1} + \sigma_x \sin \theta \cos \phi + \sigma_y \sin \theta \sin \phi + \sigma_z \cos \theta) \\ &= \frac{\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{n}}{2} = \frac{\mathbb{1} + \sigma_n}{2} \end{aligned} \quad (2)$$

其中

$$\mathbf{n} = \begin{pmatrix} \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \end{pmatrix}$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \sigma_n = \sigma_x \sin \theta \cos \phi + \sigma_y \sin \theta \sin \phi + \sigma_z \cos \theta$$

通过 (2) 式, 有

$$\sigma_n |\psi\rangle = (2\psi - \mathbb{1}) |\psi\rangle = |\psi\rangle$$

表明 $|\psi\rangle$ 是 σ_n 的本征态, 相应的本征值为 $+1$. 也可以直接计算 σ_n 的本征向量, 其结果是

$$\sigma_n |n\pm\rangle = \pm |n\pm\rangle$$

其中

$$|n+\rangle = \begin{pmatrix} \cos \frac{\theta}{2} e^{-i\frac{\phi}{2}} \\ \sin \frac{\theta}{2} e^{i\frac{\phi}{2}} \end{pmatrix}, \quad |n-\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i\frac{\phi}{2}} \\ \cos \frac{\theta}{2} e^{i\frac{\phi}{2}} \end{pmatrix} \quad (3)$$

显然 $\langle n+|n-\rangle = 0$. 将 $|n+\rangle$ 与 (1) 式的 $|\psi\rangle$ 比较, 二者相差一个相因子 $e^{i\gamma}$.

可以将 σ_n 可以表示为

$$\sigma_n = |n+\rangle\langle n+| - |n-\rangle\langle n-|$$

结合 $\mathbb{1} = |n+\rangle\langle n+| + |n-\rangle\langle n-|$, 可以写出 $|n\pm\rangle$ 的密度矩阵形式,

$$|n\pm\rangle\langle n\pm| = \frac{\mathbb{1} \pm \sigma_n}{2}$$

形象地说, $|n+\rangle\langle n+|$ 对应于 \mathbb{R}^3 中的单位向量 \mathbf{n} , $|n-\rangle\langle n-|$ 对应于 \mathbb{R}^3 中的单位向量 $-\mathbf{n}$. 在 \mathbb{C}^2 空间中两个正交的向量对应于 \mathbb{R}^3 中两个方向相反的单位向量.

2

密度矩阵和 Bloch 向量

考虑两个纯态 $|\psi_1\rangle$ 和 $|\psi_2\rangle$ 分别以几率 p_1 和 p_2 混合, $p_1 + p_2 = 1$.

$$\psi_1 = |\psi_1\rangle\langle\psi_1| = \frac{1}{2}(\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{n}_1), \quad \psi_2 = |\psi_2\rangle\langle\psi_2| = \frac{1}{2}(\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{n}_2)$$

$$\rho = p_1\psi_1 + p_2\psi_2 = \frac{1}{2}[\mathbb{1} + \boldsymbol{\sigma} \cdot (p_1\mathbf{n}_1 + p_2\mathbf{n}_2)] = \frac{1}{2}(\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{r}), \quad |\mathbf{r}| < 1$$

其中 $\mathbf{r} = p_1\mathbf{n}_1 + p_2\mathbf{n}_2$, 是两个单位向量 \mathbf{n}_1 和 \mathbf{n}_2 的凸和 (convex sum).

进而可以知道, \mathbb{C}^2 中的任意量子态都可以用密度矩阵表示为

$$\rho = \frac{1}{2}(\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{r}), \quad \mathbf{r} \in \mathbb{R}^3, \quad |\mathbf{r}| \leq 1 \quad (4)$$

(4) 式中的 \mathbf{r} 称为 Bloch 向量. Bloch 向量分布在 \mathbb{R}^3 中的单位球的表面及内部. 这个单位球称为 Bloch 球. 球面上的点对应于 \mathbb{C}^2 中的纯态, 球内部的点对应于混合态. 球心对应于最大混合态 $\frac{1}{2}$. 球面上相对的两个点对应于两个正交的纯态, 例如前面 (3) 式中的 $|n+\rangle$ 和 $|n-\rangle$.

表达式 (4) 中的 Bloch 向量 \mathbf{r} 的三个分量是

$$r_x = \text{Tr}(\rho\sigma_x), \quad r_y = \text{Tr}(\rho\sigma_y), \quad r_z = \text{Tr}(\rho\sigma_z)$$

计算一下 r_x . 在 (4) 式的两端都乘以 σ_x , 有

$$\rho\sigma_x = \frac{1}{2}(\sigma_x + r_x\sigma_x^2 + r_y\sigma_y\sigma_x + r_z\sigma_z\sigma_x)$$

注意到 $\sigma_x^2 = \mathbb{1}$, $\sigma_y\sigma_x = -i\sigma_z$, $\sigma_z\sigma_x = i\sigma_y$. 再对上式两端求迹, 注意到三个 Pauli 矩阵的迹都等于零, 于是得到 $r_x = \text{Tr}(\rho\sigma_x)$. Bloch 向量的三个分量分别是三个 Pauli 矩阵在给定量子态 ρ 中的期望值, 它们都是实验可测的.

对于 \mathbb{C}^2 上的混合态 ρ , 它的 Bloch 向量位于单位球内, 有无穷多种分解形式. 图 1 画出了 Bloch 球在 xz 平面上的截面. 混合态对应于单位圆中的 P 点, Bloch 向量为 \mathbf{r} . 通过 Bloch 向量的直径是 P_+P_- , 这条直径对应于 ρ 的本征分解, 点 P_+ 和 P_- 分别表示 ρ 的两个本征值为 $+1$ 和 -1 的本征向量. 另外, 可以通过点 P 任意画一条弦 CD . 点 C 和点 D 虽然表示两个纯态, 但是它们不是 ρ 的本征向量. 又可以画一个凸多边形 (比如图中的 $\triangle EFG$), P 点在其内部, 这个凸多边形的各个顶点以一定的权重进行平均之后, 得到点 P .

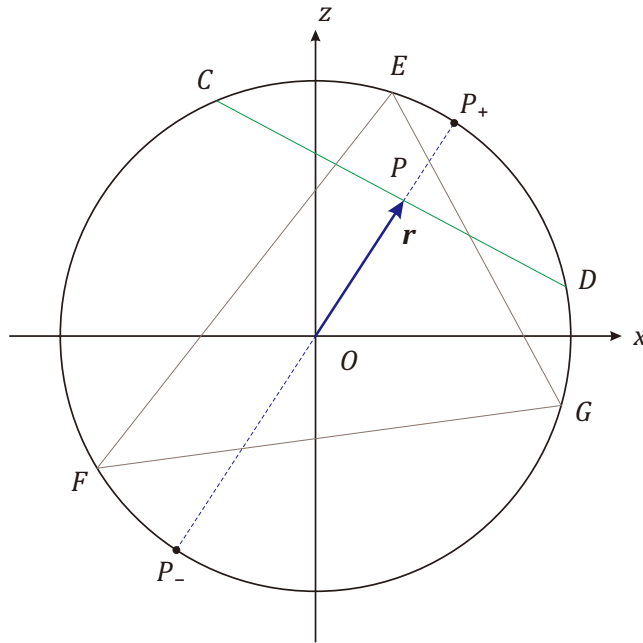


图 1

由 ρ 的表示 $\rho = \frac{1}{2}(\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{r})$, 可以算出它的本征值和本征向量,

$$\rho = \frac{1}{2}\mathbb{1} + \frac{1}{2}r\sigma_n, \quad \mathbf{r} = r\mathbf{n}, \quad \sigma_n = \boldsymbol{\sigma} \cdot \mathbf{n}$$

σ_n 的本征分解形式 $\sigma_n = |n+\rangle\langle n+| - |n-\rangle\langle n-|$

$$\therefore \rho \text{ 的本征值是 } \frac{1+r}{2}, \quad \frac{1-r}{2},$$

相应的本征向量是 $|n+\rangle, |n-\rangle$

$$\rho = \frac{1+r}{2} |n+\rangle\langle n+| + \frac{1-r}{2} |n-\rangle\langle n-|$$

还可以看到, 当 $r = 1$ 时, 本征值为 1 和 0, 对应于纯态, 当 $r = 0$ 时, 本征值是两个 $1/2$, 是最大混合态. Bloch 向量的长度可以表征量子态的混合程度, 或者说混乱度, 或者说纯度 (purity).

对于 \mathbb{C}^2 上的两个纯态 ψ_1 和 ψ_2 , 二者的内积的模方是

$$\begin{aligned} |\langle \psi_1 | \psi_2 \rangle|^2 &= \text{Tr}(\psi_1 \psi_2) \\ &= \text{Tr} \left[\frac{1}{2}(\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{n}_1) \frac{1}{2}(\mathbb{1} + \boldsymbol{\sigma} \cdot \mathbf{n}_2) \right] \\ &= \frac{1}{2} + \frac{1}{4} \text{Tr}[(\boldsymbol{\sigma} \cdot \mathbf{n}_1)(\boldsymbol{\sigma} \cdot \mathbf{n}_2)] \\ &= \frac{1}{2} + \frac{1}{2} \mathbf{n}_1 \cdot \mathbf{n}_2 \end{aligned}$$

其中最后一步用到

$$(\boldsymbol{\sigma} \cdot \mathbf{n}_1)(\boldsymbol{\sigma} \cdot \mathbf{n}_2) = \mathbb{1} \mathbf{n}_1 \cdot \mathbf{n}_2 + i \boldsymbol{\sigma} \cdot (\mathbf{n}_1 \times \mathbf{n}_2)$$

对于两个混合态 ρ_1 和 ρ_2 , 有

$$\text{Tr}(\rho_1 \rho_2) = \frac{1}{2} + \frac{1}{2} \mathbf{r}_1 \cdot \mathbf{r}_2 \quad (5)$$

这对应于两个矩阵的 Hilbert-Schmidt (HS) 范数 (亦称为 Frobenius 范数) 以及相应的内积.

$$\|A\|_2 = \|A\|_{\text{HS}} = \sqrt{\text{Tr}(A^\dagger A)}$$

$$(A, B) = \text{Tr}(A^\dagger B)$$

注 1 $\|A\|_2$ 中下标 2 的意思是这样的. 对于 \mathbb{C}^n 上的某个矩阵 X (未必是厄密的), 矩阵 $X^\dagger X$ 一定是厄密的, 而且 $X^\dagger X \geq 0$. 将 $X^\dagger X$ 的本征值记作 λ_i , 它们都是非负的. 定义矩阵 X 的 singular values s_i

$$s_i = \sqrt{\lambda_i}$$

以此定义矩阵 X 的不同类型的范数

$$\|X\|_p = \left(\sum_j s_j^p \right)^{1/p}$$

例如 $\|X\|_1 = \sum_j s_j$, 又被称为 trace norm. 而 $\|X\|_2 = \left(\sum_j s_j^2 \right)^{1/2}$ 就是上面说的 HS norm. ■

对于两个纯态 $|\psi_1\rangle$ 和 $|\psi_2\rangle$, 它们的内积的模方 $|\langle \psi_1 | \psi_2 \rangle|^2$ 表示这两个态之间的重叠程度: 如果二者相同 (可以有整体相因子的差别), 那么 $|\langle \psi_1 | \psi_2 \rangle|^2 = 1$, 它们的重叠程度最大; 如果二者正交, 那么 $|\langle \psi_1 | \psi_2 \rangle|^2 = 0$, 二者没有重叠. 这种看法亦可用于混合态 ρ_1 和 ρ_2 , 不过需要稍作修正: 当它们对应的 Bloch 向量的长度给定的时候, 如果两个 Bloch 向量平行, 那么重叠最大, 如果反平行, 则重叠最小. 对于后者, 不能说 ρ_1 和 ρ_2 正交, 因为两个正交的混合态应该满足 $\text{Tr}(\rho_1 \rho_2) = 0$. 从 (5) 式可以看出, 双值系统的混合态不可能正交 —— 当 r_1 和 r_2 都小于 1 时, 不可能有 $\mathbf{r}_1 \cdot \mathbf{r}_2 = -1$.

双值系统量子态的演化可以用 Bloch 向量形象地表示, $\mathbf{r}(t)$. 需要注意的是, Bloch 向量只是量子态的一种表示形式, 虽然 $\mathbf{r} \in \mathbb{R}^3$, 但是这个 \mathbb{R}^3 与通常的三维欧氏空间 \mathbb{E}^3 有区别. 考虑正交向量的不同描述就可以感受到二者的不同 —— 对于 Bloch 向量而言, 方向彼此相反的两个单位向量对应于两个正交的量子纯态, 而在 \mathbb{E}^3 中, 正交的两个向量就是通常说的垂直, 夹角为 90° . 因此, Bloch 向量只是 \mathbb{C}^2 中的量子态的密度矩阵形式的形象描述. 而且, 还要指出的是, 在采用 Bloch 向量描述 $|\psi\rangle \in \mathbb{C}^2$ 的时候, 我们是不能知晓这个量子态的整体相位的. 在谈论测量结果的几率分布或者 Born 规则的时候, 量子态的整体相位不会对结果产生影响, 但是, 以后我们将会看到, 在量子态的演化过程中, 会出现一类不能用动力学过程解释的相位, 是为几何相 (或者说拓扑相), 在讨论这一类问题的时候, 整体相位不但是有意义的, 而且是很重要的.

厄密矩阵

设 A 是 \mathbb{C}^2 的厄密矩阵. 可以在 $\mathbb{1}, \sigma_i (i = 1, 2, 3)$ 上展开, 且展开系数均为实数.

$$A = a_0 \mathbb{1} + a_1 \sigma_1 + a_2 \sigma_2 + a_3 \sigma_3,$$

$$a_0 = \frac{1}{2} \text{Tr}(A), \quad a_i = \frac{1}{2} \text{Tr}(A \sigma_i)$$

其中的单位矩阵意义不大, 可以略去, 从而

$$A = a_1 \sigma_1 + a_2 \sigma_2 + a_3 \sigma_3 = a \sigma_n \quad (6)$$

这里 $a = \sqrt{a_1^2 + a_2^2 + a_3^2}$, σ_n 是 \mathbf{n} 方向上的泡利矩阵, 即 $\sigma_n = \mathbf{n} \cdot \boldsymbol{\sigma}$, 而方向 \mathbf{n} 是

$$\mathbf{n} = \left(\frac{a_1}{a}, \frac{a_2}{a}, \frac{a_3}{a} \right)$$

或者用球坐标中的角度 θ 和 ϕ 表示,

$$\mathbf{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$$

于是由 (6) 式表示的厄密矩阵 A 的两个本征值是 $+a$ 和 $-a$, 相应的本征向量是 $|n+\rangle$ 和 $|n-\rangle$,

$$|n+\rangle = \begin{pmatrix} \cos \frac{\theta}{2} e^{-i \frac{\phi}{2}} \\ \sin \frac{\theta}{2} e^{i \frac{\phi}{2}} \end{pmatrix}, \quad |n-\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i \frac{\phi}{2}} \\ \cos \frac{\theta}{2} e^{i \frac{\phi}{2}} \end{pmatrix}$$

厄密矩阵 A 又可以写为

$$A = a |n+\rangle \langle n+| - a |n-\rangle \langle n-|$$

这里看到这样一个事实: \mathbb{C}^2 上的厄密矩阵在本质上相对于 Pauli 矩阵在某个方向 \mathbf{n} 上的分量, 即 σ_n , 换句话说, σ_n 是 \mathbb{C}^2 上的厄密矩阵 (或者说观测量) 的典型代表, 而且, \mathbb{C}^2 中的任意一个向量都是某个 σ_n 的本征向量.

酉矩阵

前一章讨论过 \mathbb{C}^N 上的酉矩阵, 它们构成酉群 $U(N)$. 酉群 $U(N)$ 不是单群. 行列式等于 1 的酉矩阵构成特殊酉群 $SU(N)$, 它是单群. 酉群 $U(N)$ 可以分解为两个单群, $U(1)$ 和 $SU(N)$.

$U(N)$ 的群元素有 N^2 个独立参数, $SU(N)$ 的群元素有 $N^2 - 1$ 个独立参数.

在最简单的 \mathbb{C}^2 空间中, 可以立即写出属于 $SU(2)$ 群的酉矩阵的一般形式,

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \quad a, b \in \mathbb{C}, \quad |a|^2 + |b|^2 = 1 \quad (7)$$

我们还说过, 考虑无穷小变换, 可以得到生成元. 生成元之间的对易关系构成了 Lie 代数.

4.1 单参数酉群

先看一下最简单的情形: 单参数酉群. 设酉算子 U 的参数是 s , 即 $U(s)$, 构成单参数 Lie 群. 有单位元, 满足乘法规则, 有逆元, 即 $U(0) = \mathbb{1}$, $U(s_1)U(s_2) = U(s_1 + s_2)$, $U^{-1}(s) = U(-s) = U^\dagger(s)$.

考虑无穷小变换, 令 $s \rightarrow 0$,

$$\begin{aligned}
 U(s) &= \mathbb{1} + \left. \frac{dU}{ds} \right|_{s=0} s + \dots \\
 UU^\dagger &= \mathbb{1} + \left[\left. \frac{dU}{ds} + \frac{dU^\dagger}{ds} \right] \Big|_{s=0} s + \dots \\
 \therefore \left. \frac{dU}{ds} \right|_{s=0} &\text{ 是反厄密的.} \\
 \left. \frac{dU}{ds} \right|_{s=0} &= iK, \quad K = K^\dagger.
 \end{aligned} \tag{8}$$

K 是 $U(s)$ 的生成元. 有限大小的酉变换 $U(s)$ 可以表示为

$$U(s) = e^{isK}$$

在以后的讨论中, K 可以是 Hamiltonian, 动量, 角动量等等.

4.2 $SU(2)$ 的生成元

酉变换的参数可以不只一个. (7) 式的酉矩阵有三个独立的实参数. 设想将它重新表示为

$$U = \begin{pmatrix} \cos \alpha e^{i\xi} & \sin \alpha e^{i\eta} \\ -\sin \alpha e^{-i\eta} & \cos \alpha e^{-i\xi} \end{pmatrix}, \quad \alpha \in \left[0, \frac{\pi}{2}\right] \tag{9}$$

接着, 如果将 (8) 式所示过程用于上述形式的 U 推导出生成元, 那么只能得到两个生成元 (自己算一下). 原因在于, 在参数空间的某些区域, (9) 式的参数化形式与群操作不是一一对应的. 单位变换对应于 $\alpha = 0, \xi = 0, 0 \leq \eta < 2\pi$. 因此, 不能从这一形式中得到变换的生成元.

为了得到 $SU(2)$ 变换的生成元, 采用如下形式的参数表示. 在 (7) 式中, $a = a_1 + ia_2, b = b_1 + ib_2$, 其中 a_1, a_2, b_1, b_2 都是实数, 并且 $a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1$, 令

$$\begin{aligned}
 a_2 &= \frac{1}{2}x_3, \quad b_1 = \frac{1}{2}x_2, \quad b_2 = \frac{1}{2}x_1 \\
 a_1 &= \sqrt{1 - \frac{1}{4}(x_1^2 + x_2^2 + x_3^2)}
 \end{aligned}$$

当然, 这些 x_i 需满足 $x_1^2 + x_2^2 + x_3^2 \leq 4$. 单位变换唯一地对应于 $x_1 = x_2 = x_3 = 0$.

接着可以得到如下三个生成元,

$$G_1 = \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad G_2 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad G_3 = \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

这几个生成元是反厄密的, 乘以 $-i$, 变为厄密的, 对应于 Pauli 矩阵, 对易关系是

$$\left[\frac{\sigma_x}{2}, \frac{\sigma_y}{2} \right] = i \frac{\sigma_z}{2}, \quad \dots\dots$$

\mathbb{C}^2 上的酉变换都可以视为旋转变换, 原因是:

- \mathbb{C}^2 上的任意一个厄密矩阵本质上是 $\mathbf{n} \cdot \boldsymbol{\sigma} = \sigma_n$.
- 酉变换 (9) 的三个生成元的对易关系满足角动量的三个分量之间的对易关系.

当空间的维数 > 2 , 就没有这样的结论了.

\mathbb{C}^2 上的酉变换都可以视为旋转变换, 变换的生成元是 $\frac{1}{2}\sigma_n$.

首先考虑一个相对简单的变换, 绕 z 轴转动角度 ϕ , 这句话的数学表示是 $U(z, \phi) = e^{-i\phi\frac{\sigma_z}{2}}$, 我们希望将它写为容易计算的矩阵形式

$$\begin{aligned} U(z, \phi) &= e^{-i\frac{\phi}{2}\sigma_z} \\ &= \mathbb{1} + (-i)\frac{\phi}{2}\sigma_z + \frac{(-i)^2}{2!}\left(\frac{\phi}{2}\right)^2\sigma_z^2 + \frac{(-i)^3}{3!}\left(\frac{\phi}{2}\right)^3\sigma_z^3 + \dots \\ &= \mathbb{1} \cos \frac{\phi}{2} - i\sigma_z \sin \frac{\phi}{2} \end{aligned}$$

然后, 一般地, 绕 \mathbf{n} 方向旋转角度 β ,

$$U(\mathbf{n}, \beta) = e^{-i\frac{\beta}{2}\sigma_n} = \mathbb{1} \cos \frac{\beta}{2} - i\sigma_n \sin \frac{\beta}{2}$$

这里, 注意到这样一个事实, 旋转角度 4π 才是单位变换,

$$U(\mathbf{n}, 4\pi) = \mathbb{1} \cos \frac{4\pi}{2} - i\sigma_n \sin \frac{4\pi}{2} = \mathbb{1}$$

而旋转 2π 的结果是 -1 , 会对量子态带来 π 的相位差. 这不同于刚体在 \mathbb{R}^3 空间中的旋转 —— 2π 的旋转使其位置形态回到原初.

4.3 对量子态的变换

先是一个简单的例子. 让 $U(z, \phi)$ 作用于 $|0\rangle$,

$$U(z, \phi)|0\rangle = \begin{pmatrix} e^{-i\frac{\phi}{2}} & 0 \\ 0 & e^{i\frac{\phi}{2}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{-i\frac{\phi}{2}} \\ 0 \end{pmatrix} = e^{-i\frac{\phi}{2}}|0\rangle$$

变换前后, 有整体的相因子的差别. 如果写出密度矩阵的形式, 则看不出任何变化. 在密度矩阵的表示中, $|0\rangle$ 的 Bloch 向量是 z 方向上的单位向量, 绕 z 轴旋转, 不会有任何变化. 这也说明, Bloch 向量不能反映一个纯态的全貌.

接着让 $U(z, \phi)$ 作用于 $|x+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, 结果是

$$U(z, \phi)|x+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\frac{\phi}{2}} \\ e^{i\frac{\phi}{2}} \end{pmatrix}$$

写出结果的密度矩阵的形式, 容易看到, 它的 Bloch 向量位于 xy 平面内, 且与 x 轴的夹角是 ϕ . 所以, \mathbb{C}^2 上的酉变换可以形象地用 Bloch 向量在 \mathbb{R}^3 上的 $SO(3)$ 变换来描述. 不过提请注意的是, 后者不能体现整体相位的变化.

如果要考虑对混合态进行酉变换, 那么就涉及对矩阵的酉变换.

4.4 对矩阵的变换

在我们的叙述中, 矩阵的身份是算子, 对矩阵的变换应该就是超算子. 设 X 是 \mathbb{C}^N 上的一个矩阵, 那么对它的酉变换是

$$X \longrightarrow UXU^\dagger$$

看一个简单的例子, 用 $U(z, \phi)$ 作用于 σ_x ,

$$\begin{aligned} & U(z, \phi)\sigma_x U^\dagger(z, \phi) \\ &= \begin{pmatrix} e^{-i\frac{\phi}{2}} & 0 \\ 0 & e^{i\frac{\phi}{2}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{i\frac{\phi}{2}} & 0 \\ 0 & e^{-i\frac{\phi}{2}} \end{pmatrix} \\ &= \begin{pmatrix} 0 & e^{-i\phi} \\ e^{i\phi} & 0 \end{pmatrix} = \sigma_x \cos \phi + \sigma_y \sin \phi \end{aligned} \quad (10)$$

就像是将 x 方向上的向量绕 z 轴转动了角度 ϕ .

设 \mathbf{m} 是 \mathbb{R}^3 中的单位向量, 位于 xy 平面内,

$$\mathbf{m} = (\cos \phi \quad \sin \phi \quad 0)$$

在 \mathbf{m} 方向上的 Pauli 矩阵是

$$\sigma_m = \sigma_x \cos \phi + \sigma_y \sin \phi$$

又可以写为

$$\sigma_m = \sigma_x (\mathbb{1} \cos \phi + i \sigma_z \sin \phi) = (\mathbb{1} \cos \phi - i \sigma_z \sin \phi) \sigma_x$$

注意到

$$\mathbb{1} \cos \phi \pm i \sigma_z \sin \phi = e^{\pm i \phi \sigma_z}$$

于是有

$$\sigma_m = \sigma_x e^{i \phi \sigma_z} = e^{-i \phi \sigma_z} \sigma_x$$

再利用 σ_x 和 σ_z 之间的反对易关系, 有

$$\sigma_x e^{i \phi \sigma_z} = e^{-i \phi \sigma_z} \sigma_x = e^{-i \frac{\phi}{2} \sigma_z} \sigma_x e^{i \frac{\phi}{2} \sigma_z}$$

σ_m 可以表示为

$$\sigma_m = e^{-i \frac{\phi}{2} \sigma_z} \sigma_x e^{i \frac{\phi}{2} \sigma_z}$$

其中酉变换 $e^{-i \frac{\phi}{2} \sigma_z}$ 的意思很明确: 绕 z 轴转动角度 ϕ . 从 σ_x 到 σ_m 的变换就像是把 x 方向上的单位向量转到 \mathbf{m} .

更一般地, 考虑 $\sigma_n = \boldsymbol{\sigma} \cdot \mathbf{n}$, 其中方向矢量 \mathbf{n} 的球坐标是 θ 和 ϕ .

$$\sigma_n = e^{-\frac{i}{2} \phi \sigma_z} e^{-\frac{i}{2} \theta \sigma_y} \sigma_z e^{\frac{i}{2} \theta \sigma_y} e^{\frac{i}{2} \phi \sigma_z} \quad (11)$$

酉变换 $e^{-\frac{i}{2} \phi \sigma_z} e^{-\frac{i}{2} \theta \sigma_y}$ 的操作过程也是很形象的: 先绕 y 轴转动 θ , 然后绕 z 轴转动 ϕ . 如果在 \mathbb{R}^3 空间中, 这个变换将 \mathbf{e}_z 旋转到 \mathbf{n} .

(11) 式当然可以通过矩阵的相乘直接验证. 也可以利用 Baker-Hausdorff 公式:

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!}[A, [A, B]] + \frac{1}{3!}[A, [A, [A, B]]] + \dots$$

另外, 还要注意 e^{A+B} 一般不等于 $e^A e^B$. 但是, 如果 $[A, [A, B]] = [B, [A, B]] = 0$, 那么有

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A, B]}.$$

超算子的一种表示形式

酉变换 U 以超算子的身份作用于矩阵 X , 形式是

$$X \longrightarrow UXU^\dagger$$

现在, 把矩阵表示为一个厂向量, 记作 $|X\rangle$, 其构成是这样的: 把矩阵 X 的每一行按次序排成一个长的列向量. 例如, 对于 2×2 矩阵 X ,

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

相应的长向量 $|X\rangle$ 的形式是

$$|X\rangle = \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{pmatrix}$$

可以验证, 与 $X \longrightarrow UXU^\dagger$ 对应的向量形式是

$$|X\rangle \longrightarrow (U \otimes U^*) |X\rangle$$

考虑 σ_x 经历 $U(z, \phi)$ 变换, 变换后的向量形式是,

$$\begin{aligned} & (U(z, \phi) \otimes U^*(z, \phi)) |\sigma_x\rangle \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-i\phi} & 0 & 0 \\ 0 & 0 & e^{i\phi} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

将最后结果还原为矩阵形式, 就是 (10) 式给出的结果.

4.5 对密度矩阵的变换

前面已经看到, 对于 \mathbb{C}^2 的厄密算子的典型代表 σ_n , 在酉矩阵的作用下, 变换过程有形象的几何描述: 就像是与 σ_n 相关的单位向量 \mathbf{n} 在 \mathbb{R}^3 空间上的旋转变换. 由此可以给出 \mathbb{C}^2 中的量子态的密度矩阵形式在酉变换下的几何描述, 即 Bloch 向量的旋转变换.

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma})$$

$$\begin{aligned} &\rightarrow U(\mathbf{n}, \beta)\rho U^\dagger(\mathbf{n}, \beta) \\ &\sim \text{SO}(3) \text{ rotation, } R(\mathbf{n}, \beta)\mathbf{r} \end{aligned}$$

最后强调指出, 虽然在 \mathbb{C}^2 上的酉变换可以当作旋转变换来考虑, 但是, 在更高维的空间中, 这种看法是不成立的. 以 \mathbb{C}^3 空间为例. 特殊酉变换构成 $SU(3)$ 群, 它有 9 个生成元, 可以表示为 3×3 的厄密矩阵. 其中一种表示形式是 Gell-Mann 矩阵, 形成 $\mathfrak{su}(3)$ 代数. 在这九个生成元中, 有三个构成了其中的子代数, 它们之间的对易关系就是我们说的旋转变换的生成元应该满足的代数关系. 显然, 在这种情形下, 酉变换是更为一般的变换, 旋转变换虽然是酉变换, 但只是其中的一个子群.

在以后讨论自旋为 1 的粒子的时候, 会说到它的自旋角动量的表示. 三个分量的矩阵形式是

$$S_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad S_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (12)$$

如果采用 Gell-Mann 矩阵, 则是

$$S_x = \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad S_y = \hbar \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad S_z = \hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (13)$$

它们的形式不同, 原因是采用了不同的表象, 而且, 不论是 (12) 式还是 (13) 式的表示, 都满足

$$[S_x, S_y] = i\hbar S_z, \quad \dots\dots$$

5

\mathbb{C}^2 空间中的量子测量

这一节的内容是 Born 规则在对双值量子系统测量中的具体应用. 这里仍然不讨论量子测量的物理过程, 只是提供数学形式上的计算.

双值量子系统的观测量可以表示为 \mathbb{C}^2 空间上的厄密矩阵, 这些厄密矩阵与 σ_n 并无本质上的差别, 所以我们只需要讨论对 σ_n 的测量.

设双值量子系统的状态为

$$|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle$$

这里的 $|0\rangle$ 和 $|1\rangle$ 是自然基向量, 也是 σ_z 的本征向量.

如果被测力学量是 σ_z , 那么得到结果 ± 1 (记作 $\sigma_z = \pm 1$) 的几率是

$$p(\sigma_z = +1) = |\langle 0|\psi\rangle|^2 = |c_0|^2, \quad p(\sigma_z = -1) = |\langle 1|\psi\rangle|^2 = |c_1|^2$$

非选择测量后, 系统的状态是混合态

$$|c_0|^2 |0\rangle\langle 0| + |c_1|^2 |1\rangle\langle 1|$$

用密度矩阵的形式表示测量前后系统状态的变化,

$$\begin{pmatrix} |c_0|^2 & c_0 c_1^* \\ c_0^* c_1 & |c_1|^2 \end{pmatrix} \rightarrow \begin{pmatrix} |c_0|^2 & 0 \\ 0 & |c_1|^2 \end{pmatrix}$$

测量后系统量子态的密度矩阵的非对角项变为零, 即所谓的退相干过程.

$$\psi \rightarrow \sum_k \Pi_k^{\sigma_z} \psi \Pi_k^{\sigma_z}$$

如果测量 σ_x , 那么有

$$p(\sigma_x = +1) = |\langle x+ | \psi \rangle|^2 = \frac{1}{2} |c_0 + c_1|^2, \quad p(\sigma_x = -1) = |\langle x- | \psi \rangle|^2 = \frac{1}{2} |c_0 - c_1|^2$$

也可以在 σ_x 表象中表示 $|\psi\rangle$,

$$\begin{aligned} |\psi\rangle &= (|x+\rangle\langle x+| + |x-\rangle\langle x-|)(c_0 |0\rangle + c_1 |1\rangle) \\ &= \frac{c_0}{\sqrt{2}} |x+\rangle + \frac{c_1}{\sqrt{2}} |x+\rangle + \frac{c_0}{\sqrt{2}} |x-\rangle - \frac{c_1}{\sqrt{2}} |x-\rangle \\ &= \frac{c_0 + c_1}{\sqrt{2}} |x+\rangle + \frac{c_0 - c_1}{\sqrt{2}} |x-\rangle \end{aligned}$$

展开系数的模方给出几率 $p(\sigma_x = \pm 1)$.

测量后的状态是

$$p(\sigma_x = +1) |x+\rangle\langle x+| + p(\sigma_x = -1) |x-\rangle\langle x-|$$

在 σ_x 表象中是对角的, 但是在 σ_z 表象中则不是对角的.

再让我们用密度矩阵和投影算子的形式重新表示 $p(\sigma_x = \pm 1)$.

$$\begin{aligned} p(\sigma_x = +1) &= |\langle x+ | \psi \rangle|^2 \\ &= \langle x+ | \psi \rangle \langle \psi | x+ \rangle \\ &= \text{Tr}(|\psi\rangle\langle\psi| |x+\rangle\langle x+|) \\ &= \text{Tr}(\psi \Pi_+^{\sigma_x}) \end{aligned}$$

其中 $\psi = |\psi\rangle\langle\psi|$, $\Pi_+^{\sigma_x} = |x+\rangle\langle x+|$.

设 ψ 的 Bloch 向量为 \mathbf{n} , 而 $\Pi_+^{\sigma_x}$ 的 Bloch 向量是简单的 $(1, 0, 0)$, 将几率 $p(\sigma_x = +1)$ 用 Bloch 向量表示, 结果是

$$p(\sigma_x = +1) = \text{Tr} \left[\frac{1}{2} (\mathbb{1} + \mathbf{n} \cdot \boldsymbol{\sigma}) \cdot \frac{1}{2} (\mathbb{1} + \sigma_x) \right] = \frac{1}{2} (1 + n_x)$$

这里 n_x 是 \mathbf{n} 的 x 分量.

一般地, 设量子态 ρ 的 Bloch 向量是 \mathbf{r} , 系统的被测力学量是 $\sigma_m = \boldsymbol{\sigma} \cdot \mathbf{m}$, 其中 \mathbf{m} 是 \mathbb{R}^3 中的单位向量, 测量结果为 ± 1 的几率是

$$p(\sigma_m = \pm 1) = \frac{1}{2} (1 \pm \mathbf{r} \cdot \mathbf{m})$$

再来看看 Bloch 向量的意义. 三个分量可以表示为

$$r_x = \text{Tr}(\rho\sigma_x) = p(\sigma_x = +1) - p(\sigma_x = -1)$$

$$r_y = \text{Tr}(\rho\sigma_y) = p(\sigma_y = +1) - p(\sigma_y = -1)$$

$$r_z = \text{Tr}(\rho\sigma_z) = p(\sigma_z = +1) - p(\sigma_z = -1)$$

这表明 Bloch 向量的本质是测量结果的几率分布.

在实际的 SG 实验中, 不论梯度磁场指向什么方向, 观测到的实验现象都是沿磁场梯度的正反方向分为两束, 据此可以认为, 银原子的内禀状态可以用最大混合态 $\frac{1}{2}\mathbb{1}$ 描述.

5.1 分析一个实验序列

实验过程如图 2 所示. 自旋为 1/2 的粒子的初态是 $\psi = |\psi\rangle\langle\psi|$, 先后经过 SG(z), SG(n) 和 SG(x) 测量仪器, 每一次测量都给出了明确的实验现象.



图 2

将纯态形式的初态表示为

$$|\psi\rangle = \alpha |z+\rangle + \beta |z-\rangle = \alpha |0\rangle + \beta |1\rangle$$

将三次测量的观测量分别记作 $A = \sigma_z$, $B = \sigma_n$, $C = \sigma_x$. 将它们的测量结果分别记作 a, b, c , 它们的可能取值均为 ± 1 .

粒子经过了 SG(z) 装置后, 得到结果 a 的几率是

$$p(a) = \langle\psi|\Pi_a^A|\psi\rangle = \text{Tr}(\psi\Pi_a^A)$$

这里 $\Pi_a^A = |a\rangle\langle a|$, 而 $|a\rangle$ 是 A 的对应于本征值 a 的本征向量.

第一次测量得到结果 a , 第二次测量得到结果 b , 这两件事情的联合几率是

$$p(a, b) = \|\Pi_b^B \Pi_a^A |\psi\rangle\|^2$$

这里 $\Pi_b^B = |b\rangle\langle b|$ 是对应于测量结果 b 的投影算子. 上式又可以写为

$$p(a, b) = \langle\psi|\Pi_a^A \Pi_b^B \Pi_a^A |\psi\rangle = \text{Tr}(\Pi_a^A \Pi_b^B \Pi_a^A \psi) \quad (14)$$

其中 $\psi = |\psi\rangle\langle\psi|$ 是纯态 $|\psi\rangle$ 的密度算子. 下面我们来解释并推导这个结果.

联合几率 $p(a, b)$ 可以表示

$$p(a, b) = p(b|a)p(a)$$

其中 $p(b|a)$ 是条件几率: 在第一次测量得到结果 a 的条件下, 第二次测量得到结果 b 的几率. 这相对于将对应于第一次测量结果 a 的量子态作为初态输入到测量仪器 SG(n) 中, 并计算测量结果 b 出现的几率.

对应于结果 a 的量子态是 $|a\rangle$, 又可以写作 $|z_{\pm}\rangle$, 或者写作 $|0\rangle, |1\rangle$. 但是, 我们不想采用这些形式, 而是用初态和投影算子表示, 其目的是为了推导过程具有一般性.

未归一的态矢量是 $\Pi_a^A |\psi\rangle$, 归一化之后的态矢量是

$$\frac{1}{\sqrt{p(a)}} \Pi_a^A |\psi\rangle$$

然后, 对如此形式的量子态测量力学量 B , 得到结果 b 的几率就是条件几率

$$\begin{aligned} p(b|a) &= \left| \langle b | \frac{1}{\sqrt{p(a)}} \Pi_a^A |\psi\rangle \right|^2 \\ &= \frac{1}{p(a)} \left| \langle b | \Pi_a^A |\psi\rangle \right|^2 \\ &= \frac{1}{p(a)} \langle \psi | \Pi_a^A |b\rangle \langle b | \Pi_a^A | \psi \rangle \\ &= \frac{1}{p(a)} \langle \psi | \Pi_a^A \Pi_b^B \Pi_a^A | \psi \rangle \end{aligned}$$

于是联合事件 (a, b) 发生的几率是

$$p(a, b) = p(b|a)p(a) = \langle \psi | \Pi_a^A \Pi_b^B \Pi_a^A | \psi \rangle$$

从操作的角度说, 系统的初态 $\psi = |\psi\rangle\langle\psi|$ 经历了如下演化过程,

$$\psi \longrightarrow \Pi_a^A \psi \Pi_a^A \longrightarrow \Pi_b^B \Pi_a^A \psi \Pi_a^A \Pi_b^B$$

对最后结果求迹得到 (14) 式.

类似地, 考虑最后一次测量得到结果 c , 联合事件 (a, b, c) 对应的的演化过程是

$$\psi \longrightarrow \Pi_c^C \Pi_b^B \Pi_a^A \psi \Pi_a^A \Pi_b^B \Pi_c^C$$

相应的几率是

$$p(a, b, c) = \langle \psi | \Pi_a^A \Pi_b^B \Pi_c^C \Pi_b^B \Pi_a^A | \psi \rangle$$

从联合几率可以到边缘几率

$$\begin{aligned} \sum_c p(a, b, c) &= \text{Tr} \left(\psi \Pi_a^A \Pi_b^B \sum_c \Pi_c^C \Pi_b^B \Pi_a^A \right) \\ &= \text{Tr} (\psi \Pi_a^A \Pi_b^B \Pi_b^B \Pi_a^A) \\ &= p(a, b) \end{aligned}$$

类似地, 有

$$\sum_{b,c} p(a, b, c) = p(a)$$

以上过程给出的计算结果不仅仅局限于两维复空间, 可以用于一般情况下的有限维复空间.

现在回到 \mathbb{C}^2 情形. 第一次测量给出 ± 1 的几率是

$$p(a) = \langle \psi | \Pi_a^A | \psi \rangle = \begin{cases} |\alpha|^2 & a = +1 \\ |\beta|^2 & a = -1 \end{cases}$$

接着考虑第二次测量. 当 $a = +1$ 时, 系统的测量后的状态是 $|z+\rangle$, 或者说 $|0\rangle$. 测量 $B = \sigma_n$, 这里假设 \mathbf{n} 的方向是 $\mathbf{n} = (\sin \theta, 0, \cos \theta)$. 得到结果 $b = +1$ 的几率是

$$p(b = +1 | a = +1) = |\langle n+ | z+\rangle|^2 = \cos^2 \frac{\theta}{2}$$

联合事件 $(a = +1, b = +1)$ 的几率是

$$p(a = +1, b = +1) = |\alpha|^2 \cos^2 \frac{\theta}{2}$$

这个结果来自于 $p(a, b) = p(b|a)p(a)$. 如果直接利用 (14) 式, 有

$$\begin{aligned} p(a = +1, b = +1) &= \langle \psi | \Pi_+^A \Pi_+^B \Pi_+^A | \psi \rangle \\ &= \langle \psi | z+\rangle \langle z+ | n+\rangle \langle n+ | z+\rangle \langle z+ | \psi \rangle \\ &= |\alpha|^2 \cos^2 \frac{\theta}{2} \end{aligned}$$

结论是一样的.

考虑最后测量 $C = \sigma_x$, 得到结果 $c = +1$ 的几率. 需要注意的是, 这个结果有四个来源, 如图 3 所示.

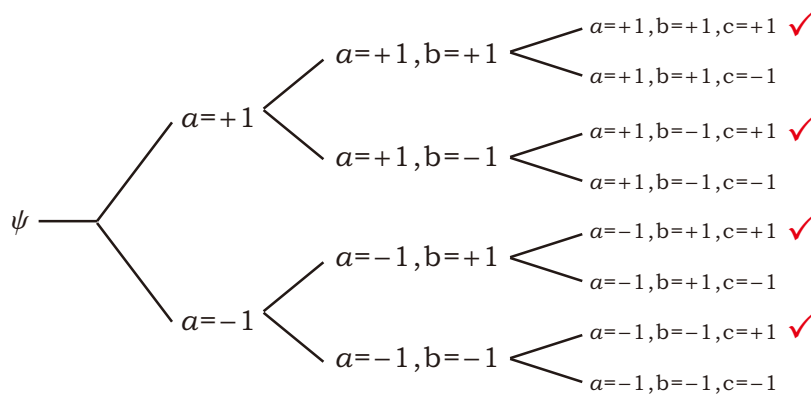


图 3

分别计算上图中的四个几率.

$$p(+1, +1, +1) = \frac{1}{2} |\alpha|^2 (1 + \sin \theta) \cos^2 \frac{\theta}{2}$$

$$p(+1, -1, +1) = \frac{1}{2} |\alpha|^2 (1 - \sin \theta) \sin^2 \frac{\theta}{2}$$

$$p(-1, +1, +1) = \frac{1}{2} |\beta|^2 (1 + \sin \theta) \sin^2 \frac{\theta}{2}$$

$$p(-1, -1, +1) = \frac{1}{2} |\beta|^2 (1 - \sin \theta) \cos^2 \frac{\theta}{2}$$

将上面四个几率相加, 得到

$$p(c = +1) = \frac{1}{2}[1 + (|\alpha|^2 - |\beta|^2) \sin \theta \cos \theta] \quad (15)$$

如果没有前两次测量, 而是直接对 $|\psi\rangle$ 测量 C , 那么得到 $c = +1$ 的几率是

$$|\langle x+|\psi\rangle|^2 = \frac{1}{2}|\alpha + \beta|^2$$

这与 (15) 有大不同.

5.2 前选择和后选择

$p(a, c)$ 是前选择 a , 后选择 c 的几率,

$$p(a, c) = \sum_b p(a, b, c)$$

在条件 (a, c) 下, 中间过程测量 B 得到结果 b 的几率是

$$p(b|a, c) = \frac{p(a, b, c)}{p(a, c)}$$

一个具体的结果:

$$p(\sigma_n = +1 | \sigma_z = +1, \sigma_x = +1) = \cos^2 \frac{\theta}{2} \frac{1 + \sin \theta}{1 + \sin \theta \cos \theta}$$

没有这样的量子态会给出这样的结果.

测量过程的对称性. 初态设为 $|a\rangle$, 或者前选择 a , 接着测量 B , 最后测量 C , 后选择 c ,

$$p(a, b, c) = \text{Tr}(\Pi_a^A \Pi_b^B \Pi_c^C \Pi_b^B)$$

又可以写为

$$p(a, b, c) = \text{Tr}(\Pi_c^C \Pi_b^B \Pi_a^A \Pi_b^B) = p(c, b, a)$$

上式右端相对于前选择 c , 接着测量 B , 最后测量 A , 并后选择 a .