

第五章 两体量子系统 III

优化 (majorization) 和偏序 (partial order)

有两个几率分布, 表示为几率向量的形式,

$$\mathbf{p} = (p_1, p_2, \dots, p_n), \quad p_i \geq 0, \quad \sum_{i=1}^n p_i = 1$$

$$\mathbf{q} = (q_1, q_2, \dots, q_n), \quad q_i \geq 0, \quad \sum_{i=1}^n q_i = 1$$

哪一个更为混乱, 哪一个更为有序?

用数学中优化的语言来说, 如果 \mathbf{q} 优于 \mathbf{p} (\mathbf{p} is majorized by \mathbf{q}), 记作 $\mathbf{p} \prec \mathbf{q}$, 那么几率分布 \mathbf{q} 更为有序. 下面介绍“优化”的概念.

考虑两个 n 维的非负分量的实向量 \mathbf{x}, \mathbf{y} , 它们的分量 $x_i \geq 0, y_i \geq 0$. 将它们分量以降序排列. 如果它们满足下列两个条件,

$$(i) \quad \sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i, \quad k = 1, 2, \dots, n$$

$$(ii) \quad \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

那么我们就说, 向量 \mathbf{y} 优于 \mathbf{x} (\mathbf{x} is majorized by \mathbf{y}), 记作 $\mathbf{x} \prec \mathbf{y}$.

如果我们所考虑的向量是几率向量, 所有分量的和为 1, 那么上述第 (ii) 个条件自然满足.

降序排列后的向量用 x^\downarrow 表示.

$$\left. \begin{aligned} x^\downarrow &\prec x^\downarrow \\ x^\downarrow &\prec y^\downarrow, y^\downarrow &\prec z^\downarrow &\implies x^\downarrow &\prec z^\downarrow \\ x^\downarrow &\prec y^\downarrow, y^\downarrow &\prec x^\downarrow &\implies x^\downarrow = y^\downarrow \end{aligned} \right\} \implies \text{partial order}$$

这个偏序结构中的最小和最大:

$$x_{(n)}^\downarrow = \left(\frac{1}{n} \quad \frac{1}{n} \quad \dots \quad \frac{1}{n}\right) \prec x^\downarrow \prec \left(1 \quad 0 \quad \dots \quad 0\right) = x_{(1)}^\downarrow$$

有一个容易证明的简单关系: 对于 $p \in [0, 1]$,

$$x_1^\downarrow \prec y^\downarrow \text{ and } x_2^\downarrow \prec y^\downarrow \implies (px_1^\downarrow + (1-p)x_2^\downarrow) \prec y^\downarrow$$

几率向量之间的优越关系描述了随机分布的有序性或者混乱程度. $x_{(1)}^\downarrow$ 是最为有序的, $x_{(n)}^\downarrow$ 是最为混乱的.

下面的定理给出了 $x < y$ 的充要条件.

Hardy, Littlewood and Pólya

$x < y$ 的充要条件是, 存在一个双随机矩阵 (bistochastic matrix) B , 使得 $x = By$.

随机矩阵 (stochastic matrix) 是这样的, 如果 n 行矩阵 $B = (b_{ij})$ 的矩阵元非负, 即 $b_{ij} \geq 0$, 并且每一列的矩阵元的和为 1, 即 $\sum_{i=1}^n b_{ij} = 1$, 那么 B 是随机矩阵. 进一步地, 如果一个 $n \times n$ 的随机矩阵 B 的每一行的矩阵元的和也为 1, 即 $\sum_{j=1}^n b_{ij} = 1$, 那么该矩阵是双随机矩阵.

想象用一个 n 行随机矩阵 B 作用于某个 n -分量的几率向量 p , 将其变换为 $p' = Bp$, 变换后的几率向量的某一个分量是

$$p'_i = \sum_{j=1}^n b_{ij} p_j$$

可以看到, 变换前后几率守恒, 即变换前 $\sum_{i=1}^n p_i = 1$, 变换后

$$\sum_{i=1}^n p'_i = \sum_{i,j=1}^n b_{ij} p_j = \sum_{j=1}^n \left(\sum_{i=1}^n b_{ij} \right) p_j = \sum_{j=1}^n p_j = 1$$

对于双随机矩阵, 不但保证几率守恒, 而且还可以保持 “最小的” $x_{(n)}^\downarrow$ 不变.

如果矩阵 B 的矩阵元 $b_{ij} = |u_{ij}|^2$, 其中 u_{ij} 是某个酉矩阵的矩阵元, 那么 $\sum_{i=1}^n b_{ij} = \sum_{j=1}^n b_{ij} = 1$, 表明矩阵 B 是双随机矩阵, 称为酉随机矩阵 (unistochastic matrix). 酉随机矩阵一定是双随机矩阵, 但是双随机矩阵不一定是酉随机矩阵.

密度矩阵的本征值

密度矩阵的本征值可以视作观测结果的几率. \mathbb{C}^n 中的量子态的密度矩阵 $\rho = \sum_{i=1}^n \lambda_i |\xi_i\rangle\langle\xi_i|$, 即本征值为 λ_i , 相应的本征向量为 $|\xi_i\rangle$.

如果观测量 X 与密度矩阵 ρ 对易, $[X, \rho] = 0$, 那么在基 $\{|\xi_i\rangle\}$ 上它们都是对角的, 即 $X = \sum_i x_i |\xi_i\rangle\langle\xi_i|$, 其中 x_i 是 X 的本征值. 测量 X , 得到结果 x_i 的几率是

$$p_i^X = \langle \xi_i | \rho | \xi_i \rangle = \lambda_i$$

于是有几率向量 p^X ,

$$p^X = \lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$$

现在, 我们想说明, 几率向量 p^X 是最优的.

设另外某个观测量 A 的本征值和本征向量分别是 a_j 和 $|\alpha_j\rangle$. 得到某个结果 a_j 的几率为 $p_j^A = \langle \alpha_j | \rho | \alpha_j \rangle$. 下面的过程给出了 p_j^A 和 p_i^X 之间的联系.

$$\begin{aligned} p_j^A &= \langle \alpha_j | \rho | \alpha_j \rangle \\ &= \sum_i |\langle \xi_i | \alpha_j \rangle|^2 \lambda_i \end{aligned}$$

$$= \sum_i |\langle \xi_i | \alpha_j \rangle|^2 p_i^X \quad (1)$$

用 \mathbf{p}^A 表示观测量 A 的测量结果的几率向量, 并定义矩阵 $D = (d_{ij})$, 矩阵元 $d_{ij} = |\langle \xi_i | \alpha_j \rangle|^2$, 因此有

$$\mathbf{p}^A = D\mathbf{p}^X = D\boldsymbol{\lambda}$$

显然矩阵 D 是酉随机矩阵, 也是双随机矩阵, 所以有

$$\mathbf{p}^A \prec \boldsymbol{\lambda} \quad (2)$$

这说明几率向量 $\boldsymbol{\lambda}$ 优于 \mathbf{p}^A .

Schur 凸函数

Schur 凸函数 (Schur-convex function)

对于 $\mathbf{x} \prec \mathbf{y}$, 如果实函数 f 满足 $f(\mathbf{x}) \leq f(\mathbf{y})$, 那么称之为 Schur-convex 函数.

一个可微函数 $f(x_1, x_2, \dots, x_n)$ 是 Schur 凸函数, 当且仅当 f 满足下面两个条件,

1. 函数 f 关于 $\{x_i\}$ 是交换不变的.
2. 对于所有的 $\mathbf{x} = (x_1, x_2, \dots, x_n)$, 有

$$(x_i - x_j) \left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_j} \right) \geq 0$$

详细的讨论和证明参见 T. Ando, Majorization, doubly stochastic matrices, and comparison of eigenvalues. Linear Algebra and its Applications **118**, 163-248 (1989).

例如, 下列函数是 Schur 凸函数,

$$f(\mathbf{x}) = \sum_i x_i \log x_i, \quad f(\mathbf{x}) = \sum_i x_i^k \text{ (for } k \geq 1), \quad f(\mathbf{x}) = -\prod_i x_i, \quad \text{etc.}$$

如果 $f(\mathbf{x})$ 是 Schur 凸函数, 那么 $-f(\mathbf{x})$ 是 Schur 凹函数. 比如 $f(\mathbf{x}) = -\sum_i x_i \log x_i$ 是 Schur 凹函数.

Shannon 熵和 von Neumann 熵

“熵”是 Clausius 为了讨论物理系统的热力学行为而引入的概念, 在 Boltzmann, Gibbs 等人的工作之后, von Neumann 和 Shannon 又对熵理论做出了进一步贡献. 他们二人的着眼点是不同的, von Neumann 的出发点在于量子力学, Shannon 则是奠定了经典通讯理论.

熵是系统状态的不确定性和随机性的度量. Shannon 认为, 物理系统承载着的 uncertainty 可以被视作系统携带的信息. 通过适当的操作, 我们可以从物理系统中获得一定量的信息. 如果物理系统有着较大的 uncertainty, 那么获得的信息也较多.

设经典随机变量 X 有 n 个可能的取值 x_i , 相应的概率是 p_i , 且 $\sum_i p_i = 1$. 经典情形下的 Shannon 熵是

$$H(X) = - \sum_i p_i \log p_i$$

其中, 对数的底可以是 2, 得到 bit, 也可以是 e , 得到 nat.

在统计力学中, 熵是微观状态数的度量, 即 $S = k_B \ln \Omega$, 其中, k_B 是 Boltzmann 常数, Ω 是微观状态数. 考虑 Boltzmann 分布, N 个可区分的粒子放入 k 个盒子中, 并且在第 i 个盒子中放入 N_i 个粒子, $\sum_{i=1}^k N_i = N$, 那么状态数是

$$W(N_1, N_2, \dots, N_k) = \frac{N!}{N_1! N_2! \dots N_k!}$$

用概率的语言描述, 粒子放入第 i 个盒子的概率是 p_i , 那么对于 $N \gg 1$, 第 i 个盒子中的粒子数是 $N_i = p_i N$, 状态数改写为

$$W(p_1, p_2, \dots, p_k) = \frac{N!}{(p_1 N)! (p_2 N)! \dots (p_k N)!}$$

利用 Stirling 公式 $\ln N! = N \ln N - N + \mathcal{O}(\ln N)$, 可以看到

$$\left| \frac{1}{N} \ln W(p_1, p_2, \dots, p_k) - H(p_1, p_2, \dots, p_k) \right| = \mathcal{O}\left(\frac{\ln N}{N}\right)$$

于是 Shannon 熵与统计力学中的熵有着紧密联系, 所以有时候也把 Shannon 熵称为 Boltzmann-Gibbs-Shannon 熵.

量子情形下, 有量子熵, 又叫 von Neumann 熵.

von Neumann 熵

量子态 ρ 的 von Neumann 熵是

$$S(\rho) = - \text{Tr}(\rho \log \rho) = - \sum_i (\lambda_i \log \lambda_i)$$

其中 λ_i 是密度算子 ρ 的本征值.

纯态的量子熵等于 0. \mathbb{C}^n 上的最大混合态的量子熵等于 $\log n$.

von Neumann 熵和统计物理中的熵在一定程度上有所对应. 设想两种不同的理想气体 A 和 B, 分别有 N_1 和 N_2 个气体分子, 热力学状态分别是 (P, V_1, T) 和 (P, V_2, T) , 也就是说, 它们有相同的压强和温度. 在相同的温度 T 下将它们混合, 混合后的体积是 $V = V_1 + V_2$. 混合前后熵的改变是

$$\Delta S = k_B N_1 \log \frac{V}{V_1} + k_B N_2 \log \frac{V}{V_2}$$

用几率 p_1 和 p_2 表示,

$$N_1 = p_1 N, \quad N_2 = p_2 N, \quad V_1 = p_1 V, \quad V_2 = p_2 V$$

而且 $p_1 + p_2 = 1$. 熵的增量改写为

$$\Delta S = -k_B N p_1 \log p_1 - k_B N p_2 \log p_2$$

再引入一些量子力学的描述. 假设混合前气体 A 的分子处于 ψ_1 , 气体 B 的分子处于状态 ψ_2 , 这两个状态是正交的. 混合前, 气体 A 的熵记作 $S(\psi_1, p_1 N)$, 气体 B 的熵记作 $S(\psi_2, p_2 N)$. 整体的熵是

$$S(\psi_1, p_1 N) + S(\psi_2, p_2 N)$$

混合后系统的状态是 $\rho = p_1 \psi_1 + p_2 \psi_2$, 混合后气体的熵记作 $S(\rho, N)$,

$$\begin{aligned} S(\rho, N) &= S(\psi_1, p_1 N) + S(\psi_2, p_2 N) + \Delta S \\ &= S(\psi_1, p_1 N) + S(\psi_2, p_2 N) - k_B N p_1 \log p_1 - k_B N p_2 \log p_2 \end{aligned} \quad (3)$$

假定熵 S 正比于粒子数,

$$S(\psi_1, N_1) = N_1 S(\psi_1), \quad S(\psi_2, N_2) = N_2 S(\psi_2), \quad S(\rho, N) = N S(\rho)$$

在 (3) 式的两端同除以 N ,

$$S(\rho) = p_1 S(\psi_1) + p_2 S(\psi_2) - k_B p_1 \log p_1 - k_B p_2 \log p_2$$

而 ψ_1 和 ψ_2 都是纯态, 所以它们的熵应该等于零, $S(\psi_1) = S(\psi_2) = 0$, 进而有

$$S(\rho) = -k_B p_1 \log p_1 - k_B p_2 \log p_2$$

令 Boltzmann 常数 $k_B = 1$, 得到 von Neumann 熵的表达式¹.

von Neumann 熵是“最优的”

密度矩阵 ρ 的本征值构成几率向量

$$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$$

某个观测量 A 的测量结果的几率构成另一个几率向量

$$\boldsymbol{p}^A = (p_1^A, \dots, p_n^A)$$

(2) 式表明, $\boldsymbol{\lambda}$ 优于 \boldsymbol{p}^A . 注意到 Shannon 熵函数是 Schur 凹函数, 所以有

$$S(\rho) = H(\boldsymbol{\lambda}) \leq H(\boldsymbol{p}^A)$$

von Neumann 熵是所有观测量测量结果的 Shannon 熵中的最小值.

混合态的纯化

如果两体 (以至于多体) 量子系统处于某个纯态 $\Psi = |\Psi\rangle\langle\Psi|$, 那么某个子系统的局部量子态是用约化密度矩阵描述的, 子系统的量子态是混合态.

¹ 气体混合前后熵的改变, 这个话题涉及 Gibbs 佯谬. 考虑到量子情形下全同粒子不可区分, Gibbs 佯谬这个古老的问题仍然有很多不清楚的地方, 参看 S. Saunders, The Gibbs Paradox. Entropy **20**, 552 (2018).

现在考虑相反的过程, 设某个量子系统 Q 处于混合态 ρ^Q , 我们可以认为 ρ 是处于纯态的两体量子系统的关于某个子系统的约化密度矩阵.

引入一个辅助量子系统 (通常称作 ancilla), 记作 A . 系统 Q 和辅助系统 A 构成两体量子系统. 描述 Q 和 A 的 Hilbert 空间分别是 \mathcal{H}^Q 和 \mathcal{H}^A . 描述两体系统的 Hilbert 空间是 $\mathcal{H} = \mathcal{H}^Q \otimes \mathcal{H}^A$.

构造一个纯态 $|\Psi\rangle \in \mathcal{H}$, 使得

$$\rho^Q = \text{Tr}_A \Psi, \quad \Psi = |\Psi\rangle \langle \Psi|$$

从 \mathcal{H}^Q 上的 ρ^Q 到 \mathcal{H} 中的 $|\Psi\rangle$ 就是混合量子态的纯化 (purification).

标准纯化形式

首先写出 ρ^Q 的本征分解形式

$$\rho^Q = \sum_{i=1}^n \lambda_i |\xi_i\rangle \langle \xi_i|$$

这里假设 \mathcal{H}^Q 的维数是 n , 且 ρ^Q 是满秩的. λ_i 是 ρ^Q 的本征值, $\lambda_i > 0$, $\sum_{i=1}^n \lambda_i = 1$.

将 \mathcal{H}^A 的维数选定为 n , 即 $\dim \mathcal{H}^Q = \dim \mathcal{H}^A = n$. 将 \mathcal{H}^A 的基向量选择为 $\{|\varphi_i\rangle\}$. 构造 $|\Psi\rangle$ 如下,

$$|\Psi\rangle = \sum_{i=1}^n \sqrt{\lambda_i} |\xi_i\rangle \otimes |\varphi_i\rangle \quad (4)$$

容易验证

$$\text{Tr}_A \Psi = \rho^Q$$

考虑一下非标准的纯化. 设想 ρ^Q 是系综 $\{p_i, \psi_i\}_{i=1, \dots, m}$ 的平均量子态, 这里 ψ_i 是纯态 $|\psi_i\rangle$ 的密度矩阵, 不同的 $|\psi_i\rangle$ 不一定正交, 且 m 也未必等于 Hilbert 空间 \mathcal{H}^Q 的维数. 系统 Q 的密度矩阵是

$$\rho^Q = \sum_{i=1}^m p_i \psi_i, \quad \psi_i = |\psi_i\rangle \langle \psi_i|$$

引入辅助系统 A , $\dim(\mathcal{H}^A) = m$, $|\varphi_i\rangle \in \mathcal{H}^A$, 且 $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$, 我们可以把 ρ^Q 的纯化形式表示为

$$|\Psi\rangle = \sum_{i=1}^m \sqrt{p_i} |\psi_i\rangle \otimes |\varphi_i\rangle$$

分析一下 \mathcal{H}^Q 和 \mathcal{H}^A 维数相同情况下的纯化, $\dim(\mathcal{H}^Q) = \dim(\mathcal{H}^A) = n$. 标准的纯化形式由 (4) 给出. 现在考虑在 \mathcal{H}^A 的另一组基向量上表示 $|\Psi\rangle$,

$$\{|\varphi_i\rangle\} \xrightarrow{V} \{|\eta_i\rangle\}$$

其中 V 是酉矩阵, 它的矩阵元是 $v_{ij} = \langle \varphi_i | \eta_j \rangle$.

在新的表象中, $|\Psi\rangle$ 写为,

$$\begin{aligned} |\Psi\rangle &= \sum_i \sqrt{\lambda_i} |\xi_i\rangle \otimes \sum_j |\eta_j\rangle \langle \eta_j | \varphi_i \rangle \\ &= \sum_j \left(\sum_i \sqrt{\lambda_i} v_{ij}^* |\xi_i\rangle \right) \otimes |\eta_j\rangle \end{aligned} \quad (5)$$

如果令 $|\tilde{\psi}_j\rangle = \sum_i \sqrt{\lambda_i} v_{ij}^* |\xi_i\rangle$, 那么有

$$|\Psi\rangle = \sum_{j=1}^n |\tilde{\psi}_j\rangle \otimes |\eta_j\rangle$$

其中 $|\tilde{\psi}_j\rangle$ 既不归一, 也不彼此正交, (5) 式是非标准的纯化形式. 这表明, 当 $\dim(\mathcal{H}^Q) = \dim(\mathcal{H}^A)$ 时, 标准纯化形式和非标准纯化形式的差别仅仅是对辅助系统的局部酉变换. 实际上, 这个说法含有模糊不清的地方, 详细的讨论可参看 L. P. Hughston, R. Jozsa, W. K. Wootters, A complete classification of quantum ensembles having a given density matrix. Physics Letters A **183**, 14-18 (1993).

Mach-Zehnder 干涉仪

Which way

图 1 是 Mach-Zehnder 干涉仪的示意图. 我们将分束器 BS (Beam Splitter) 的作用视作一个 Hadamard 变换,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \text{Hadamard 变换}$$

$$|0\rangle \longrightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |1\rangle \longrightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

反射镜的变换是

$$U_{\text{mirror}} = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

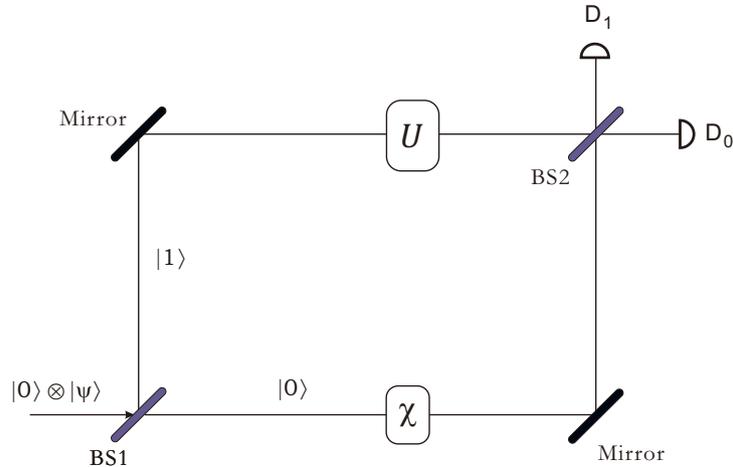


图 1

已经知道, 探测器 D_0 和 D_1 记录到粒子的几率分别是

$$p_0 = \frac{1}{2} [1 + \text{Re}(e^{-i\chi} \langle \psi | U | \psi \rangle)],$$

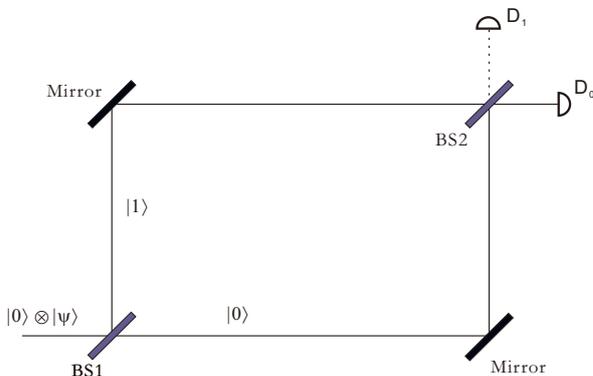
$$p_1 = \frac{1}{2} [1 - \text{Re}(e^{-i\chi} \langle \psi | U | \psi \rangle)]$$

对于单粒子源, 经常询问这样的问题: 粒子通过了哪一条路径?

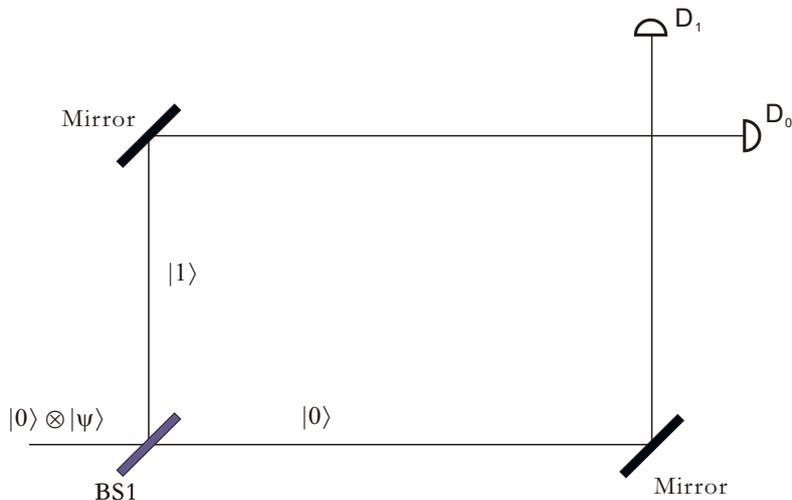
如果上下两条路径上没有任何变换, 即 $\chi = 0, U = 1$, 那么有

$$p_0 = 1, \quad p_1 = 0.$$

只有探测器 D_0 有响应, 而 D_1 不会记录到任何粒子. 对此, 我们可能会说: “粒子通过了两条路径.”



另一方面, 如果没有分束器 BS_2 , 那么从单粒子源出发的一个粒子经过了 MZ 干涉仪之后, 探测器 D_0 或 D_1 有响应的几率均为 $1/2$. 如同讨论多次的 SG 实验结果.



此时, 我们可能会说: “粒子通过了上面或者下面的路径.”

John Archibald Wheeler

Quantum Theory and Measurement,

edited by John A. Wheeler and Wojciech H. Zurek

- With the final half-silvered mirror in place the photodetector D_0 goes click-click as the successive photons arrive but the adjacent counter (*i.e.*, D_1) registers nothing. This is **evidence of interference** between the upper and the lower beams; or, in photon language, evidence that each arriving light quantum has

arrived by **both routes**. In such experiments, Einstein originally argued, it is unreasonable for a single photon to travel simultaneously two routes.

- Remove the half-silvered mirror, and one will find that the one counter goes off, or the other. Thus the photon has traveled **only one route**.
- It travels only one route, but it travels both routes; it travels both routes, but it travels only one route. What nonsense! How obvious it is that quantum theory is inconsistent!
- In our own day we have learned to state the point even more sharply by way of a so-called **delayed-choice experiment**. There we make the decision whether to put the final half-silvered mirror in place or to take it out at the very last picosecond, after the photon has already accomplished its travel.
- The dependence of what is observed upon the choice of experimental arrangement made Einstein **unhappy**. It conflicts with the view that the universe exists “out there” independent of all acts of observation.
- **In contrast** Bohr stressed that we confront here an inescapable new feature of nature, to be welcomed because of the understanding it gives us. In struggling to make clear to Einstein the central point as he saw it, Bohr found himself forced to introduce the word “**phenomenon**.”
- In today’s words Bohr’s point — and the central point of quantum theory — can be put into a single, simple sentence. “**No elementary phenomenon is a phenomenon until it is a registered (observed) phenomenon.**”
- It is wrong to speak of the “route” of the photon in the experiment of the beam splitter. It is wrong to attribute a tangibility to the photon in all its travel from the point of entry to its last instant of flight.
- A phenomenon is not yet a phenomenon until it has been brought to a close by an irreversible act of amplification such as the blackening of a grain of silver bromide emulsion or the triggering of a photodetector.
- How can one contemplate indeterminism, complementarity and “phenomenon” without being reminded of the words of Gertrude Stern about modern art?
 “It looks strange and it looks strange and it looks very strange; and then **suddenly it doesn’t look strange at all** and you can’t understand what made it look strange in the first place.”

Zeilinger, A., Weihs, G., Jennewein, T. and Aspelmeyer, M.

Happy centenary, photon,
 Nature **433**, 230-238 (2005)

- When analysing quantum interference we can fall into all kinds of traps. The general conceptual problem is that we tend to reify — to take too realistically — concepts like wave and particle. Indeed if we consider the quantum state representing the wave simply as a calculational tool, problems do not arise.
- In this case, we should not talk about a wave propagating through the double-slit setup or through a Mach-Zehnder interferometer; the quantum state is simply a tool to calculate probabilities.

- Probabilities of the photon being somewhere? No, we should be even more cautious and only talk about probabilities of a photon detector firing if it is placed somewhere.
- One might be tempted, as was Einstein, to consider the photon as being localized at some place with us just not knowing that place. But, whenever we talk about a particle, or more specifically a photon, we should only mean that which a “click in the detector” refers to.

波粒二相性

在上一小节的 MZ 干涉仪中, 令 $U = \mathbb{1}$, 但保留关于相位 χ 的变换.

粒子的空间自由度, 即路径, 记作系统 S . 粒子的内禀自由度, 即自旋, 记作系统 Q .

设 S 的初态是

$$\varphi = \frac{1}{2}(\mathbb{1} + \mathbf{s} \cdot \boldsymbol{\sigma}).$$

$\mathbf{s} = (s_1, s_2, s_3)$ 是 φ 的 Bloch 向量. Q 的初态为 ψ .

两体系统 SQ 经历的酉变换是

$$(H \otimes \mathbb{1})(\sigma_x \otimes \mathbb{1})U(\chi)(H \otimes \mathbb{1}). \quad (6)$$

其中

$$U(\chi) = \begin{pmatrix} e^{i\chi} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}.$$

变换 (6) 使得

$$\varphi \otimes \psi \longrightarrow \Psi.$$

在探测器 D_0 观测到粒子的几率是

$$p_0 = \text{Tr}[\Psi (|0\rangle \langle 0| \otimes \mathbb{1})] = \frac{1}{2}(1 - s_2 \sin \chi + s_3 \cos \chi). \quad (7)$$

几率 p_0 的最大值和最小值是

$$p_0^{\max/\min} = \frac{1}{2}(1 \pm \sqrt{s_2^2 + s_3^2}).$$

于是可见度 V 等于

$$V = \sqrt{s_2^2 + s_3^2}.$$

当 S 的初态为纯态, 并且 Bloch 向量 \mathbf{s} 位于 yz 平面内的时候, 得到最大可见度 1. 如果 Bloch 向量 \mathbf{s} 位于 x 轴, 则可见度为零.

注意到以上过程中没有对 Q 的操作, 因此在形式上可以简化为以下过程.

$$\varphi \longrightarrow \varphi' = H \sigma_x \begin{pmatrix} e^{i\chi} & 0 \\ 0 & 1 \end{pmatrix} H \varphi \text{ H.C.}$$

其中 H.C. 表示 φ 左侧的酉变换的厄米共轭.

对 φ' 进行 σ_z 测量, 得到结果 +1 的几率是

$$p_0 = \langle 0 | \varphi' | 0 \rangle.$$

结果与 (7) 式相同.

在下面的讨论中, 我们将忽略系统 Q , 仅仅考虑粒子的空间自由度, 即系统 S .

现在分析路径的可区分性.

如果没有分束器 BS_2 , 那么探测器 D_0 有响应表明粒子来自路径 $|0\rangle$; 探测器 D_1 有响应表明粒子来自路径 $|1\rangle$. 这就是对路径的探测.

我们把粒子进入 BS_2 之前时刻 S 的状态记作 φ'' , 即

$$\varphi' = H\varphi''H^\dagger = H\varphi''H.$$

对 φ'' 进行 σ_z 测量就是对路径的探测.

如果放上 BS_2 , 分束器的变换 H 是针对系统 S 的, 使得

$$|0\rangle \longrightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |x+\rangle, \quad |1\rangle \longrightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |x-\rangle.$$

由于 H 既是厄密的, 又是酉的, 所以

$$|x+\rangle \xrightarrow{H} |0\rangle, \quad |x-\rangle \xrightarrow{H} |1\rangle,$$

$$\varphi'' = H\varphi'H.$$

于是, 对 φ'' 测量 σ_z 就相当于对 φ' 测量 σ_x . 例如, 对 φ' 测量 σ_x , 得到结果 $+1$ 的几率可以写为

$$\langle x_+ | \varphi' | x_+ \rangle = \langle 0 | H\varphi'H | 0 \rangle = \langle 0 | \varphi'' | 0 \rangle$$

类似地, 对 φ' 测量 σ_z (这是为了探测干涉现象), 等价于对 φ'' 测量 σ_x .

所以, 从 φ'' 的角度来说,

测量 φ'' 的 $\sigma_z \sim$ 路径探测.

测量 φ'' 的 $\sigma_x \sim$ 干涉现象探测.

或者, 从 φ' 的角度说,

测量 φ' 的 $\sigma_z \sim$ 干涉现象探测.

测量 φ' 的 $\sigma_x \sim$ 路径探测.

因此, 在 BS_2 存在的情况下, 我们可以采用两种不同的方式来测量 φ' . 根据 σ_z 测量的几率分布来定义干涉现象的可见度,

$$\mathcal{V} = \text{Tr} \left[\varphi' \frac{\mathbb{1} + \sigma_z}{2} \right] = (s_2^2 + s_3^2)^{1/2}. \quad (8)$$

用 σ_x 测量来定义路径的可区分性, 记作 \mathcal{P} ,

$$\mathcal{P} = |p(\sigma_x = +1|\varphi') - p(\sigma_x = -1|\varphi')| = \left| \frac{1}{2}(1 - s_1) - \frac{1}{2}(1 + s_1) \right| = s_1.$$

显然有

$$\mathcal{P}^2 + \mathcal{V}^2 \leq 1. \quad (9)$$

- 为了在实验中得到 \mathcal{P} 和 \mathcal{V} , 需要两种不同类型的测量. 一个是测量 σ_x , 另一个是测量 σ_z .
- 虽然 \mathcal{P} 可以用来描述路径的可区分性, 但是我们并没有在干涉仪的任何一条路径上放置任何用以探测路径的仪器, 因此, \mathcal{P} 所描述的路径的可区分性是基于 σ_x 的测量结果的一种预言, 在这个意义上, \mathcal{P} 的意思是关于路径的 predictability.
- (9) 式是波粒二相性的数学形式, 是互补性的体现. 但是在本质上, 类似于 Maassen-Uffink 给出的熵形式的不等式,

$$H(A|\rho) + H(B|\rho) \geq -2 \log c, \quad c = \max_{j,k} |\langle \alpha_j | \beta_k \rangle|,$$

其中 $|\alpha_j\rangle$ 和 $|\beta_k\rangle$ 分别是观测量 A 和 B 的本征向量.

以下内容可以不看

现在, 引入路径探测器, 记作 D . 考虑两体量子系统 SD . 初态是

$$\rho_{in} = \varphi^S \otimes \omega^D.$$

其中 φ^S 是系统 S 的初态, $\varphi^S = \frac{1}{2}(1 + \mathbf{s} \cdot \boldsymbol{\sigma})$, 暂不指定探测器 D 的初态 ω^D 具体的形式. 系统 S 和路径探测器 D 之间的相互作用发生在第二个分束器 BS_2 之前.

在下面的计算值, 省略上标 S 和 D .

经过 BS_1 之后, 有

$$\begin{aligned} \rho_{in} = \varphi \otimes \omega &\longrightarrow (H \otimes 1)(\varphi \otimes \omega)(H \otimes 1) \\ &= \frac{1}{2}(\mathbb{1} + s_3\sigma_1 - s_2\sigma_2 + s_1\sigma_3) \otimes \omega \\ &= \frac{1}{2} \begin{pmatrix} \omega & 0 \\ 0 & \omega \end{pmatrix} + \frac{s_3}{2} \begin{pmatrix} 0 & \omega \\ \omega & 0 \end{pmatrix} - \frac{s_2}{2} \begin{pmatrix} 0 & -i\omega \\ i\omega & 0 \end{pmatrix} + \frac{s_1}{2} \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix}. \end{aligned}$$

接着考虑相位变换,

$$\begin{pmatrix} e^{i\chi} & 0 \\ 0 & 1 \end{pmatrix} \otimes \mathbb{1} = \begin{pmatrix} e^{i\chi} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}.$$

以及平面镜的反射变换,

$$\sigma_x \otimes \mathbb{1} = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

经过了相位变换和平面镜反射之后, SD 的状态是

$$\begin{aligned} &\frac{1}{2} \begin{pmatrix} \omega & 0 \\ 0 & \omega \end{pmatrix} - \frac{s_1}{2} \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix} \\ &- \frac{s_2}{2} \begin{pmatrix} 0 & i e^{-i\chi} \omega \\ -i e^{i\chi} \omega & 0 \end{pmatrix} + \frac{s_3}{2} \begin{pmatrix} 0 & e^{-i\chi} \omega \\ e^{i\chi} \omega & 0 \end{pmatrix}. \end{aligned} \tag{10}$$

然后, S 和 D 之间产生了相互作用, 二者经历的酉变换是控制酉变换,

$$U = |0\rangle\langle 0| \otimes U_0 + |1\rangle\langle 1| \otimes U_1 = \begin{pmatrix} U_0 & 0 \\ 0 & U_1 \end{pmatrix}.$$

将酉变换 U 作用于 (10) 式, 得到

$$\begin{aligned} & \frac{1}{2} \begin{pmatrix} U_0 \omega U_0^\dagger & 0 \\ 0 & U_1 \omega U_1^\dagger \end{pmatrix} - \frac{s_1}{2} \begin{pmatrix} U_0 \omega U_0^\dagger & 0 \\ 0 & -U_1 \omega U_1^\dagger \end{pmatrix} \\ & + \frac{s_2}{2} \begin{pmatrix} 0 & -ie^{-i\chi} U_0 \omega U_1^\dagger \\ ie^{i\chi} U_1 \omega U_0^\dagger & 0 \end{pmatrix} + \frac{s_3}{2} \begin{pmatrix} 0 & e^{-i\chi} U_0 \omega U_1^\dagger \\ e^{i\chi} U_1 \omega U_0^\dagger & 0 \end{pmatrix} \end{aligned}$$

然后进入第二个分束器 BS_2 , 从 BS_2 出来之后的末态是

$$\begin{aligned} \rho_f &= \frac{1-s_1}{4} (\mathbb{1} + \sigma_x) \otimes U_0 \omega U_0^\dagger \\ &+ \frac{1+s_1}{4} (\mathbb{1} - \sigma_x) \otimes U_1 \omega U_1^\dagger \\ &+ \frac{s_3 - is_2}{4} e^{-i\chi} (\sigma_z - i\sigma_y) \otimes U_0 \omega U_1^\dagger \\ &+ \frac{s_3 + is_2}{4} e^{i\chi} (\sigma_z + i\sigma_y) \otimes U_1 \omega U_0^\dagger \end{aligned} \quad (12)$$

在末态 ρ_f 中, 路径探测器 D 的量子态是

$$\rho_f^D = \text{Tr}_S \rho_f = \frac{1-s_1}{2} U_0 \omega U_0^\dagger + \frac{1+s_1}{2} U_1 \omega U_1^\dagger. \quad (13)$$

在 ρ_f 中, 路径 S 的量子态是 $\rho_f^S = \text{Tr}_D \rho_f$. 令

$$s_3 + is_2 = s_{23} e^{-i\theta}, \quad C = \text{Tr}(U_0 \omega U_1^\dagger),$$

其中 $s_{23} = \sqrt{s_2^2 + s_3^2}$. 可以将 ρ_f^S 表示为

$$\begin{aligned} \rho_f^S &= \frac{1}{2} \left[\mathbb{1} - s_1 \sigma_x + \frac{s_{23}}{2} (-iC e^{i(\theta-\chi)} + iC^* e^{-i(\theta-\chi)}) \sigma_y \right. \\ &\quad \left. + \frac{s_{23}}{2} (C e^{i(\theta-\chi)} + C^* e^{-i(\theta-\chi)}) \sigma_z \right]. \end{aligned}$$

为了看到看到干涉现象, 需要对 ρ_f^S 测量 σ_z , 得到结果 +1 的几率是

$$\begin{aligned} p_0 &= \langle 0 | \rho_f^S | 0 \rangle \\ &= \frac{1}{2} + \frac{s_{23}}{4} [C e^{i(\theta-\chi)} + C^* e^{-i(\theta-\chi)}] \\ &= \frac{1}{2} + \frac{1}{2} s_{23} \text{Re}[C e^{i(\theta-\chi)}] \end{aligned}$$

p_0 的最大和最小值分别是

$$p_0^{\max} = \frac{1}{2}(1 + s_{23}|C|), \quad p_0^{\min} = \frac{1}{2}(1 - s_{23}|C|).$$

于是得到干涉现象的可见度

$$\mathcal{V} = s_{23}|C| \leq s_{23} = \sqrt{s_2^2 + s_3^2}. \quad (14)$$

这表明, 由于路径探测器的存在和影响, 干涉现象的可见度有所降低, 不再是先前的由 (8) 式给出的结果.

再来看路径的可区分性. 在 SD 的末态 (12) 式和 D 的末态 (13) 中, 令

$$\rho_0^D = U_0 \omega U_0^\dagger, \quad \rho_1^D = U_1 \omega U_1^\dagger. \quad (15)$$

从 (12) 式可以看到, 路径探测器 D 的状态 ρ_0^D 与 S 的状态 $|x+\rangle$ 联系在一起, 而 D 的状态 ρ_1^D 与 S 的状态 $|x-\rangle$ 联系在一起. 前面的讨论告诉我们, S 的量子态 $|x\pm\rangle$ 分别对应于路径 0 和 1. 因此, D 的状态 ρ_0^D 和 ρ_1^D 之间的可区分性就是路径的可区分性.

令 $q_0 = \frac{1-s_1}{2}$, $q_1 = \frac{1+s_1}{2}$, 重写 (13) 式,

$$\rho_f^D = q_0 \rho_0^D + q_1 \rho_1^D.$$

我们可以这么说: ρ_f^D 是系综 $\{q_0, \rho_0^D; q_1, \rho_1^D\}$ 的平均量子态.

于是问题转化为: 两个量子态, ρ_0^D 和 ρ_1^D , 分别以几率 q_0 和 q_1 出现 ($q_0 + q_1 = 1$), 如何有效地区分它们? 更具体地, 可以设想, 有某个人分别以几率 q_0 和 q_1 制备了处于量子态 ρ_0^D 和 ρ_1^D 的量子系统, 然后让观测者回答这样的问题: 现在发送给你的量子态是 ρ_0^D 还是 ρ_1^D ?

下面的叙述中我们省略量子态的上标 D .

这里我们假设, 允许观测者知道 ρ_0 和 ρ_1 的具体形式 (或者说, 可以告知观测者量子态的制备过程), 也允许观测者知道几率 q_0 和 q_1 的值, 并且, 观测者可以对接收到的量子系统进行任意形式的操作或测量.

现在, 观测者开始对他接收到的量子系统进行测量, 被测力学量是 A , 它的本征值和相应的本征向量分别是 a_j 和 $|\alpha_j\rangle$. 当他得到某个结果 a_j 的时候, 该如何判断测量前的量子态是 ρ_0 还是 ρ_1 呢? 因为观测者知道 ρ_0 和 ρ_1 的形式, 他可以事先计算出一个几率表,

$$\mathbf{p}^{(0)} = \{p_1^{(0)}, p_2^{(0)}, \dots, p_n^{(0)}\},$$

$$\mathbf{p}^{(1)} = \{p_1^{(1)}, p_2^{(1)}, \dots, p_n^{(1)}\}.$$

这里, $\mathbf{p}^{(0)}$ 表示对 ρ_0 进行观测量 A 的测量, 得到不同的结果 a_j 的几率分布, 即 $p_j^{(0)} = \langle \alpha_j | \rho_0 | \alpha_j \rangle$; $\mathbf{p}^{(1)}$ 是对 ρ_1 的测量结果的几率分布, $p_j^{(1)} = \langle \alpha_j | \rho_1 | \alpha_j \rangle$. 注意到 ρ_0 和 ρ_1 出现的几率分别是 q_0 和 q_1 , 所以, 观测到某个结果 a_j 的几率是

$$p(a_j) = q_0 p_j^{(0)} + q_1 p_j^{(1)}.$$

对于特定的测量结果 a_j , 观测者采用这样的策略进行判断:

- 如果 $q_0 p_j^{(0)} > q_1 p_j^{(1)}$, 那么就回答: 测量前的量子态是 ρ_0 .
- 如果 $q_0 p_j^{(0)} < q_1 p_j^{(1)}$, 那么就回答: 测量前的量子态是 ρ_1 .
- 如果 $q_0 p_j^{(0)} = q_1 p_j^{(1)}$, 那么就随便挑选 ρ_0 或者 ρ_1 作为答案.

简言之, 挑选几率较大的作为答案.

于是, 如果观测者针对某个观测量 A 进行测量, 并采用以上策略判断测量前的量子态, 那么可以用下面的量描述判断正确的可能性,

$$\mathcal{L}_A = \sum_j \max \{q_0 p_j^{(0)}, q_1 p_j^{(1)}\} = \frac{1}{2} + \frac{1}{2} \sum_j |q_0 p_j^{(0)} - q_1 p_j^{(1)}|. \quad (16)$$

为了得到上式的第二步, 需注意到 $\sum_j q_0 p_j^{(0)} + q_1 p_j^{(1)} = 1$.

令 $\Pi_j^A = |\alpha_j\rangle\langle\alpha_j|$, 把 \mathcal{L}_A 改写为

$$\mathcal{L}_A = \frac{1}{2} + \frac{1}{2} \sum_j |\text{Tr}[(q_0 \rho_0 - q_1 \rho_1) \Pi_j^A]|. \quad (17)$$

这只是对特定的观测量 A 进行测量得到的结果, 还需要考虑所有可能的观测量. 例如, 测量另一个力学量 B , 相应的投影算子是 Π_k^B , 这时, 得到 \mathcal{L}_B , 它的形式与 \mathcal{L}_A 类似, 只是把 (17) 式中的 Π_j^A 换成 Π_k^B . 在所有可能的测量方式中进行优化, 即, 求 \mathcal{L}_A 的最大值, 有

$$\mathcal{L}_{\text{opt}} = \max_{\text{All } A} \mathcal{L}_A.$$

下标 opt 表示 optimal. \mathcal{L}_{opt} 可以表示为

$$\mathcal{L}_{\text{opt}} = \frac{1}{2}(1 + \mathcal{D}), \quad (18)$$

其中

$$\mathcal{D} = \text{Tr} \{ |q_0 \rho_0 - q_1 \rho_1| \}. \quad (19)$$

这里, 在算子 $q_0 \rho_0 - q_1 \rho_1$ 的两端加上了绝对值符号, 它的意思是, $|X| = \sqrt{X X^\dagger}$. 对 $|X|$ 求迹, 实际上就是 trace norm.

现在, 我们来解释一下 (18) 式, 也就是说, 为什么有

$$\max_{\text{All } A} |\text{Tr}[(q_0 \rho_0 - q_1 \rho_1) \Pi_j^A]| = \text{Tr} \{ |q_0 \rho_0 - q_1 \rho_1| \}. \quad (20)$$

首先, 让我们看看根据 trace norm 定义的距离, 称为“迹距离” (trace distance). 设 X 和 Y 是两个厄米算子, 定义二者之间的 trace distance 是 $D(X, Y) = \frac{1}{2} \text{Tr} |X - Y|$. 我们举一个例子看看这个定义的具体结果. 设 X 和 Y 是一个两能级系统的两个不同的量子态的密度矩阵, 即 ρ_1 和 ρ_2 . 我们把它们表示为

$$\rho_1 = \frac{1}{2}(\mathbb{1} + \mathbf{r}_1 \cdot \boldsymbol{\sigma}), \quad \rho_2 = \frac{1}{2}(\mathbb{1} + \mathbf{r}_2 \cdot \boldsymbol{\sigma}).$$

\mathbf{r}_1 和 \mathbf{r}_2 分别是 ρ_1 和 ρ_2 的 Bloch 向量.

$$D(\rho_1, \rho_2) = \frac{1}{2} \text{Tr} |\rho_1 - \rho_2| = \frac{1}{4} \text{Tr} |(\mathbf{r}_1 - \mathbf{r}_2) \cdot \boldsymbol{\sigma}|.$$

注意到 $(\mathbf{r}_1 - \mathbf{r}_2) \cdot \boldsymbol{\sigma}$ 的本征值是 $\pm |\mathbf{r}_1 - \mathbf{r}_2|$. 所以 $|(\mathbf{r}_1 - \mathbf{r}_2) \cdot \boldsymbol{\sigma}|$ 的迹就是 $2|\mathbf{r}_1 - \mathbf{r}_2|$. 于是

$$D(\rho_1, \rho_2) = \frac{1}{2} |\mathbf{r}_1 - \mathbf{r}_2|.$$

因此, 两个量子态 ρ_1 和 ρ_2 之间的距离就是相应的两个 Bloch 向量之间的距离的一半.

为了说明 (20), 我们来看 (16) 式的含义. 在这个表达式中, 有一个比较大小的过程, 即比较 $q_0 p_j^{(0)}$ 和 $q_1 p_j^{(1)}$ 哪个大哪个小, 然后选择较大的那一个. 从更为一般的角度说, 观测者做的事情无非是, 根据观测结果做出一个决定: 选择 ρ_0 还是选择 ρ_1 . 因此, 我们可以用以下描述简化 (16) 式.

观测者采用一般意义上的广义的量子测量, 效果算子是 E_0 和 E_1 , 对应的观测结果记作 0 和 1. 这里, $E_0 \geq 0$, $E_1 \geq 0$, 并且 $E_0 + E_1 = \mathbb{1}$. 如果观测者得到结果 0, 那么他就选择 ρ_0 作为答案; 如果观测者得到结果 1, 那么就选择 ρ_1 作为答案. 所以, 对于特定的一组效果算子 E_0 和 E_1 , 我们用如下形式的 $\mathcal{L}(E_0, E_1)$ 衡量猜测成功的可能性,

$$\mathcal{L}(E_0, E_1) = q_0 \text{Tr}(\rho_0 E_0) + q_1 \text{Tr}(\rho_1 E_1). \quad (21)$$

$\mathcal{L}(E_0, E_1)$ 是 \mathcal{L}_A 的等价形式. 最佳的测量方式将给出 $\mathcal{L}(E_0, E_1)$ 的最大值,

$$\mathcal{L}_{\text{opt}} = \max_{\{E_0, E_1\}} \mathcal{L}(E_0, E_1).$$

下面我们来求 \mathcal{L}_{opt} , 并且给出最佳的测量方式, 对应的效果算子记作 E_0^o 和 E_1^o , 上标 o 表示 optimal.

将 $\mathcal{L}(E_0, E_1)$ 改写为

$$\begin{aligned} \mathcal{L}(E_0, E_1) &= q_0 \text{Tr}(\rho_0 E_0) + q_1 \text{Tr}[\rho_1(1 - E_0)] \\ &= q_1 + \text{Tr}[(q_0 \rho_0 - q_1 \rho_1) E_0]. \end{aligned}$$

q_1 是不改变的. 因此, 求 $\mathcal{L}(E_0, E_1)$ 的最大值就相当于在所有可能的 E_0 中求 $\text{Tr}[(q_0 \rho_0 - q_1 \rho_1) E_0]$ 的最大值. 令

$$\Gamma = q_0 \rho_0 - q_1 \rho_1.$$

Γ 是厄密矩阵, 但是它的本征值可正可负. 将它的本征分解形式表示为

$$\Gamma = \sum_j \gamma_j |\gamma_j\rangle \langle \gamma_j|,$$

即, Γ 的本征值是 γ_j , 对应的本征向量是 $|\gamma_j\rangle$. 考虑 $\mathcal{L}(E_0, E_1)$ 中的 $\text{Tr}(\Gamma E_0)$,

$$\text{Tr}(\Gamma E_0) = \sum_j \gamma_j \langle \gamma_j | E_0 | \gamma_j \rangle.$$

其中 $0 \leq \langle \gamma_j | E_0 | \gamma_j \rangle \leq 1$, 所以有

$$\text{Tr}(\Gamma E_0) \leq \sum_j' \gamma_j. \quad (22)$$

其中, 求和符号上的撇号表示仅仅针对非负的 γ_j 求和. (22) 式给出了 $\text{Tr}(\Gamma E_0)$ 的上限, 我们还需要证明这个上限是可以达到的. 也就是说, 存在最优的 E_0^o , 使得 $\text{Tr}(\Gamma E_0^o) = \sum_j' \gamma_j$.

我们可以在 $\{|\gamma_j\rangle\}$ 表象中构造这个最优的 E_0^o :

$$\left. \begin{aligned} \langle \gamma_j | E_0^o | \gamma_j \rangle &= 0, & \text{当 } \gamma_j < 0, \\ \langle \gamma_j | E_0^o | \gamma_j \rangle &= 1, & \text{当 } \gamma_j > 0, \\ \langle \gamma_j | E_0^o | \gamma_j \rangle &= 0, & \text{当 } \gamma_j = 0. \end{aligned} \right\} \quad (23)$$

实际上, 当 $\gamma_j = 0$ 时, 我们可以将 $\langle \gamma_j | E_0^o | \gamma_j \rangle$ 设定为 $[0, 1]$ 中的任意实数, 为了使 E_0^o 的形式更简洁, 我们选择了零.

不过, 在 (23) 中, 我们只是指明了在 $\{|\gamma_j\rangle\}$ 表象中 E_0^o 的对角项, 尚没有明确它的非对角项. 下面我们要说明的是, 非对角项一律为零.

假设 E_0^o 的本征分解形式是

$$E_0^o = \sum_k e_k |e_k\rangle \langle e_k|.$$

根据 (23) 式, 当 $\gamma_j \leq 0$ 时, 有

$$0 = \langle \gamma_j | E_0^o | \gamma_j \rangle = \sum_k e_k |\langle e_k | \gamma_j \rangle|^2.$$

于是, 对于 $e_k \neq 0$, 一定有 $\langle e_k | \gamma_j \rangle = 0$. 由此给出, 对于 Γ 的具有非正的本征值的两个本征向量, $|\gamma_\ell\rangle$ 和 $|\gamma_j\rangle$, 有

$$\langle \gamma_\ell | E_0^\circ | \gamma_j \rangle = \sum_k e_k \langle \gamma_\ell | e_k \rangle \langle e_k | \gamma_j \rangle = 0,$$

其中, $\ell \neq j$ 并且 $\gamma_\ell \leq 0$, $\gamma_j \leq 0$. 上式表明, 在对应于 Γ 的非正本征值的本征向量张开的子空间中, E_0° 是对角的.

还需要证明在 Γ 的正本征值 ($\gamma_j > 0$) 的本征向量张开的子空间中, E_0° 也是对角的. 为了达到这个目的, 回到 (21) 式, 把它改写为

$$\mathcal{L}(E_0, E_1) = q_0 - \text{Tr}[(q_0 \rho_0 - q_1 \rho_1) E_1] = q_0 - \text{Tr}(\Gamma E_1).$$

于是, 求 $\mathcal{L}(E_0, E_1)$ 的最大值就相当于在所有可能的 E_1 中求 $\text{Tr}(\Gamma E_1)$ 的最小值. 沿用对 E_0° 的分析过程, 可以得到结论: 在 Γ 的正本征值 ($\gamma_j > 0$) 的本征向量张开的子空间中, E_1° 也是对角的. 并且 $\text{Tr}(\Gamma E_1)$ 的最小值可以写为

$$\text{Tr}(\Gamma E_1) \geq \sum_j'' \gamma_j. \quad (24)$$

这里, 求和符号上的两个撇号表示对所有的负的 γ_j 求和.

最后, 根据 $E_0^\circ + E_1^\circ = 1$, 可以看出 E_0° 和 E_1° 都是对角的, 且非零的对角项均为 1. 实际上, E_0° 和 E_1° 都是投影算子, 当 Hilbert 空间的维数大于 2 时, 它们是子空间上的投影算子.

至此, 我们可以写出 $\mathcal{L}(E_0, E_1)$ 的最大值,

$$\begin{aligned} \mathcal{L}_{\text{opt}} &= q_1 + \sum_j' \gamma_j = q_0 - \sum_k'' \gamma_k \\ &= \frac{1}{2} \left(q_0 + q_1 + \sum_j' \gamma_j - \sum_k'' \gamma_k \right) \\ &= \frac{1}{2} \left(1 + \sum_j |\gamma_j| \right) \\ &= \frac{1}{2} (1 + \text{Tr} |\Gamma|). \end{aligned} \quad (25)$$

这就是 (18) 式和 (19) 式. 现在我们回到这两个式子.

我们把 \mathcal{D} 称为路径的可区分性 (度量). 最后的结论是

$$\mathcal{D}^2 + \mathcal{V}^2 \leq 1. \quad (26)$$

我们要证明这个不等式.

首先, 设路径探测器 D 的初态是纯态, 即 $\omega = |\omega\rangle \langle \omega|$. 在这种情况下, 由 (15) 式定义的 ρ_0^D 和 ρ_1^D 也都是纯态. 于是需要计算 trace distance

$$\mathcal{D} = \text{Tr} \{ |q_0 \rho_0^D - q_1 \rho_1^D| \}.$$

$q_0 \rho_0^D$ 对应于 Hilbert 空间中的一个未归一的向量 $\sqrt{q_0} |\xi_0\rangle$, 其中 $|\xi_0\rangle = U_0 |\omega\rangle$, 且 $q_0 \rho_0^D = \sqrt{q_0} |\xi_0\rangle \langle \xi_0| \sqrt{q_0}$; 而 $q_1 \rho_1^D$ 对应于另一个未归一的向量 $\sqrt{q_1} |\xi_1\rangle$, 其中 $|\xi_1\rangle = U_1 |\omega\rangle$, 且 $q_1 \rho_1^D = \sqrt{q_1} |\xi_1\rangle \langle \xi_1| \sqrt{q_1}$. 我们可以把 $|\xi_1\rangle$ 表示为

$$|\xi_1\rangle = |\xi_0\rangle \cos \frac{\theta}{2} + |\xi_0^\perp\rangle \sin \frac{\theta}{2}.$$

其中 $|\xi_0^\perp\rangle$ 表示某个与 $|\xi_0\rangle$ 正交的并且已经归一化的向量. 这里可以有整体相因子的差别, 但不影响最后的结果. 于是, 所有的计算就变成了在一个由 $\{|\xi_0\rangle, |\xi_0^\perp\rangle\}$ 张开的两维子空间中的计算. 不失一般性地将 $|\xi_0\rangle$ 记作 $|0\rangle$, 将 $|\xi_0^\perp\rangle$ 记作 $|1\rangle$,

$$q_0\rho_0^D - q_1\rho_1^D = \begin{pmatrix} q_0 - q_1 \cos^2 \frac{\theta}{2} & -q_1 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ -q_1 \sin \frac{\theta}{2} \cos \frac{\theta}{2} & -q_1 \sin^2 \frac{\theta}{2} \end{pmatrix}.$$

它的本征值是

$$\frac{1}{2} \left(q_0 - q_1 \pm \sqrt{1 - 4q_0q_1 \cos^2 \frac{\theta}{2}} \right).$$

将本征值的绝对值相加, 得到 $\text{Tr} \{|q_0\rho_0^D - q_1\rho_1^D|\}$, 即

$$\mathcal{D} = \sqrt{1 - 4q_0q_1 \cos^2 \frac{\theta}{2}} = \sqrt{1 - 4q_0q_1 |\langle \xi_0 | \xi_1 \rangle|^2}.$$

再来看干涉现象的可见度 \mathcal{V} , 由 (14) 式给出. 其中的 C 是

$$C = \text{Tr}(U_0 \omega U_1^\dagger) = \text{Tr}(|\xi_0\rangle \langle \xi_1|) = \langle \xi_1 | \xi_0 \rangle.$$

注意到 $q_0 = \frac{1-s_1}{2}$, $q_1 = \frac{1+s_1}{2}$, 最终, 我们有

$$\begin{aligned} \mathcal{D}^2 + \mathcal{V}^2 &= 1 - 4q_0q_1 |\langle \xi_0 | \xi_1 \rangle|^2 + (s_{23}|C|)^2 \\ &= 1 - (1 - s_1^2) |\langle \xi_0 | \xi_1 \rangle|^2 + (s_2^2 + s_3^2) |\langle \xi_0 | \xi_1 \rangle|^2 \\ &= 1. \end{aligned}$$

这表明, 只要选择路径探测器 D 的初态为纯态, 那么 $\mathcal{D}^2 + \mathcal{V}^2$ 的值一定是 1.

如果 D 的初态是混合态, 那么 $\mathcal{D}^2 + \mathcal{V}^2$ 的值小于 1. 关于这个结论的证明就略去不再叙述了. 一般地, 我们有

$$\mathcal{D}^2 + \mathcal{V}^2 \leq 1. \quad (27)$$

这个不等式描述的是, 在使用路径探测器的情况下, 路径的可区分性与干涉现象的可见度之间存在一种此消彼长的关系. 这也是波粒二相性的定量描述, 而且, 在关于粒子性 (与路径探测相关) 的描述中, 不仅仅根据反事实 (counterfactual) 的推断, 而且确实用到了具体的探测手段.

另外, 关于 interaction-free measurements, 参看以下文献.

- Elitzur, A. and Vaidman, L.
Quantum mechanical interaction-free measurements,
Foundations of Physics **23**, 987-997 (1993)
- Vaidman, L.
The Meaning of the Interaction-Free Measurements,
Foundations of Physics **33**, 491-510 (2003)