

# 当概率成为复数

## ——量子概率简介

张江(jakezj@163.com)

写于 2010-8-6

### 一、介绍

与其说量子力学是一门物理，还不如说它是一种数学。数学关心的是符号之间抽象的关系与结构，而不关心符号对应的实物。这门诞生于量子物理的抽象数学被人们称为量子概率。量子概率不一定非得描述微观粒子，而可以描述更一般的宏观系统，例如人类的认知现象。注意，这种应用并不是把人类认知还原为大量的微观粒子，而是将认知主体作为一个完整的系统，而运用量子概率去描述。正如随机过程作为一个独立的数学分支可以脱离它的原始物理背景而应用到包括金融、生物等更广阔的系统一样，量子概率也必将可以作为一套与经典概率不同的数学分支应用到各种领域之中。

本文绕开了那些繁杂的量子物理试验以及算符、波函数等量子力学的标准讲法，而是从概率幅可以取复数这个特点切入到量子概率的理论体系中。从这个角度看，量子概率可以说是将经典概率扩展到复数域之后的数学产物，因此整个量子概率体系的各种结论都可以自然得到。然而一个疑问是：究竟这样一种纯数学游戏般的概率扩充究竟有什么实际意义呢？本文试图指出：量子概率——经典概率在复数域的扩充其实刻画的是观察者对于外界信息的不确定性情况下的描述，这种不确定性要比经典概率不确定更不确定：一个是量子概率考虑到了观察对观测系统的影响，另一个是考虑到观察者处理信息能力的极限。

本文首先沿着将概率复数化的思路引入复数概率（即量子物理中的概率幅）的概念，之后为了看到将经典概率与复数概率的不同，我们从几何表示到不兼容属性对，再从不兼容属性对到复合系统到纠缠态再到贝尔不等式一路走来，其主要的目的就是展示复数概率与经典概率的区别。然而，这样做的代价是大家对复数概率的整个框架不甚了解。于是，我们在第四章，又重新用相对正规的数学定义出发，重新介绍了量子概率的体系框架。最后，我们将把量子概率体系如何应用到人类认知现象作了简单介绍。并进一步针对不兼容属性对以及纠缠和贝尔不等式的拓广进行了讨论。

### 二、如果概率可以取复数

#### 1、当 $i$ 进入物理学

所谓的复数就是形如下式所表示的数：

$$a + bi$$

其中  $a$  和  $b$  都是普通的实数，而  $i$  则表示  $\sqrt{-1}$ ，又称为虚数单位。 $i$  最早进入数学完全是偶然。在费罗的三次方程的求根公式中， $i$  会以计算的中间步骤出现，但是又消失在最终的计算中（参见《[虚数的故事](#)》）。因为人们无法想象一个平方等于-1的数是什么样子的，因此人们给它起了一个非常有诗意的名字：虚数（Imaginary number）。

虚数，顾名思义，它只存在于人们想象之中，而无法找到任何现实对应。所以，长期以来，虚数从未真正进入过物理学（尤其是 20 世纪以前的经典物理学）领域，因为虚数不可能表达**实实在在的**物理量。即使在交流电路分析中，工程师、物理学家们用复数来表示电流和电压，但那也仅仅是一种“数学技巧”，也就是说，如果我们不用复数，也能完成所有的运算，而只不过会让计算稍稍复杂一些罢了。

20 世纪以后的物理学开始变得让人匪夷所思了。先是爱因斯坦的相对论打破了人们旧有的时空观念，接下来量子理论又开始挑战人们头脑中的实在观念。有趣的是，这两个革命性理论的建立都与虚数  $i$  在物理学中的应用分不开。

首先，在相对论中，虚数  $i$  连带着时间  $t$  和光速  $c$  构成了一个可以与空间相提并论的四维闵可夫斯基空间，在这个奇妙的四维空间里，物体的匀速运动被转化成了绕原点的旋转。更有意思的是，与空间等价的坐标系不是时间  $t$ ，而是一个乘上了虚数  $i$  的虚时间。

如果说将  $i$  引入相对论中也不过是一种数学技巧（爱因斯坦建立的狭义相对论并不依赖于四维时空，因此也没有虚数  $i$ ）。那么，在量子力学这个更让人费解的物理学理论中引入虚数  $i$  则是一种不可避免的事情。虚数  $i$  对应的是量子状态的相位，虽然相位仍然不可测量，但是它却是使得量子系统表现为区别于经典系统的主要原因。进一步，量子力学中虚数  $i$  的引入，还使得人们对概率的认识过渡到了对**概率幅**这一可以取复数的类概率量（之所以称之为类概率量，是因为概率幅的运算法则与经典概率的运算法则完全一致）的阶段。这篇文章则是从复数概率幅这个角度来介绍量子概率理论的。

## 2、将概率复数化

当我们抛一枚硬币出去，那么该硬币就会以概率  $p(1)$  出现正面，以概率  $p(0)$  出现反面，并且  $p(1)+p(0)=1$ 。这句话的意思是，如果我抛掷硬币  $N$  次，出现正面的次数是  $n$  次，那么  $p(1)$  就解释为  $n/N$ ，而  $p(0)$  就是  $(N-n)/N$ 。显然，如果我直接将概率  $p(1)$  取为复数值，则没有任何意义，因为我们不再可能将  $p(1)$  解释为出现正面的次数占总的抛掷硬币次数的比例。

然而，经典概率论存在着一个致命的漏洞。在通常情况下，事件的概率并不是一个可观测的量。在每次抛硬币的过程中，我们实际上只能观测到正面或者反面，而无法观测概率  $p(1)$ 。只有当我们进行了大量的实验，并且能够保证这些实验都是对同一个随机事件（例如在同一个条件下抛掷同一枚硬币），我们才能渐进地观测到某一个事件出现的概率。然而，在更多的无法进行重复试验的场合下（例如明天是否会下雨），我们仍然可以定义概率，但是这里的概率就是一个彻头彻尾不可观测的量。

正是概率的这种不可直接观测性，使我们有机可乘，从而可以将概率变成复数。我们并不是直接把硬币出现正面的概率  $p(1)$  定义为复数  $a+bi$ ，而是定义一种叫做**概率幅**（在本文中又称为**复数概率**）的新量，即：

$$\psi = a + bi$$

并且，我们规定这个复数概率  $\psi$  可以按照下述规则转变为经典概率：

$$p = |\psi|^2 = \psi * \psi = (a - bi)(a + bi) = a^2 + b^2$$

$|\psi|$ 表示求复数 $\psi$ 的模， $\psi^*$ 表示 $\psi$ 的共轭复数，也就是说 $\psi$ 乘以它自己的共轭就得到了 $\psi$ 的模。当然，如果要求 $p$ 表示经典概率，那么必须要求 $a^2 + b^2 \in [0,1]$ 。正是因为在一次抛硬币随机试验中，我们原则上只能得到正面（1）或者反面（0）的观测，而不能得到关于概率 $p$ 的任何信息。所以，我们不妨在这个**观测过程**上面做文章，我们假设每一次观测会发生3件事：

首先，在观测前，假设硬币正面朝上可以用一个复数概率来表示，例如：

$$\psi(1) = \frac{1}{2} + \frac{1}{2}i$$

然后，在测量的瞬间，我们定义复数概率会自动转变成概率，即按照如下原则得到正面朝上的概率：

$$p(1) = |\psi(1)|^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}$$

最后，硬币会按照这个概率 $p(1)$ 随机地出现正面或者反面。我们看到，如果进行大量的实验，我们仍然可以得到出现正面的频率是 $1/2$ 。

总之，我们假设硬币出现正面这个随机事件不仅可以用概率 $p(1)$ 来描述，而且还可以用复数概率 $\psi(1)$ 进行描述。并且对一次随机试验的观测实际上是先从复数概率转变为经典概率，再由经典概率支配出现某一个观测值的比例。

因此，我们可以假设在概率背后还有一个更基本的复数概率制约着概率本身。正是因为概率本身并不是一个可以被直接观测到的量，所以我们再做一层复数概率的假设就不会引起实质的困难。尽管我们增加了理论的复杂性，但是我们终于可以成功地将概率复数化了。其实量子力学就是这么干的！

我们知道，某一个事件 $x$ 的经典概率与该事件的复数概率存在着对应关系：

$$P(X) = |\psi(X)|^2$$

然而，这个关系却不是对称的。从复数概率到经典概率的映射是一个多对一的映射。也就是说针对一个具体的概率值 $P(X)$ ，存在着无穷多个复数概率与其对应，可以证明这群复数概率满足：

$$\psi(X) = \sqrt{P(X)}(\cos\theta + i\sin\theta) = \sqrt{P(X)}\exp(i\theta)$$

其中 $\theta$ 为任意实数。这群复数落到了以原点为圆心，以 $\sqrt{P(X)}$ 为半径的圆上。这种多对一的关系使得复数概率完全可以涵盖经典概率的内容，甚至可以包含更丰富的信息。另一方面，复数概率具有更深的不可观测性，因为即使你对事件 $x$ 进行了大量的测量，也仅仅能够得到复数概率的模的信息，而不可能确定它的相角 $\theta$ ，这就为我们估算复数概率带来了挑战。正是因为这个相角的存在，复数概率和经典概率表现出了本质的区别。

### 3、分布与向量表示

在经典概率中，一个随机变量可以取多个不同的值，这些取值构成了两两互斥的随机事件。我们假设随机变量 $x$ 的可能取值为： $\{x_1, x_2, \dots, x_n\}$ ，那么我们可以定义 $n$ 个概率：

$$p(x_i) = P\{X = x_i\}, \text{ for } i = 1, 2, \dots, n$$

并且这些概率满足归一化条件：

$$\sum_{i=1}^n p(x_i) = 1$$

那么这一组概率  $p(x_i)$  就构成了  $X$  的一个概率分布。为了方便讨论，我们不妨将这一组概率写成一个向量的形式：

$$F(X) = p_1|x_1\rangle + p_2|x_2\rangle + \cdots + p_n|x_n\rangle$$

其中  $p_i = p(x_i)$ ， $p_i|x_i\rangle$  表示对应的  $X$  在取值为  $x_i$  的时候概率为  $p(x_i)$ 。我们知道， $F(X)$  这个向量相当于  $n$  维空间中的向量  $(p_1, p_2, \dots, p_n)$ ，其中  $p_i$  就是该向量在第  $i$  个维度上的坐标。这样， $|x_i\rangle$  就相当于是在这  $n$  维空间中的第  $i$  个基向量。所以  $F(X)$  又可以写成坐标乘以相应的基向量再求和的形式。

按照前面的将概率复数化的思想，我们也可以为每个随机事件  $X=x_i$  定义复数概率，这样  $n$  个复数概率也可以写成向量的形式：

$$\Psi(X) = \psi_1|x_1\rangle + \psi_2|x_2\rangle + \cdots + \psi_n|x_n\rangle$$

并且变量  $X$  取值  $x_i$  的经典概率就是：

$$P\{X = x_i\} = |\psi_i|^2$$

并且要求：

$$\sum_{i=1}^n |\psi_i|^2 = 1$$

在经典概率中，如果我们给定了随机变量  $X$  在各种可能值上的概率，也就是给定了向量  $F(X)$ ，那么系统的状态就给定了。因为，系统的任何性质都可以从概率分布中得到。同样的道理，在复数概率中，给定了向量  $\Psi(X)$ ，也就给定了系统的状态，因为系统的一切性质都可以从这个向量中导出。

在量子力学中向量  $\Psi(X)$  称之为一个量子系统的**叠加态**。这个状态与经典概率的分布  $F(X)$  还是很不一样的。在经典概率中，如果一个系统的概率分布是  $F(X)$ ，那么我们可以说变量  $X$  以  $p_1$  的概率取值  $x_1$ ，以  $p_2$  的概率出现  $x_2$ ，……。但是在状态为  $\Psi(X)$  的量子系统中，我们不能说系统以  $|\psi_1|^2$  的概率处于  $|x_1\rangle$  状态，以  $|\psi_2|^2$  的概率处于  $|x_2\rangle$  状态，……。这是因为这些状态之间会发生相互干涉影响，我们将在下一节中继续讨论经典概率与复数概率的区别。

由于概率是不可直接测量的量，因此，我们完全可以用复数概率来描述系统从而达到用经典概率描述等同的效果。例如，我们可以用经典概率分布：

$$F(X) = \frac{1}{2}|0\rangle + \frac{1}{2}|1\rangle$$

来表示一枚硬币处于正面（状态  $|1\rangle$ ）和反面（状态  $|0\rangle$ ）的概率各是 0.5。同样，我们也可以假设这枚硬币处于一个可以用复数概率描述的量子叠加态：

$$\Psi(X) = \left(\frac{1}{2} + \frac{1}{2}i\right)|0\rangle + \left(\frac{1}{2} + \frac{1}{2}i\right)|1\rangle$$

或者：

$$\Psi(X) = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

那么，按照从复数概率到经典概率的转换，这两种描述不会产生任何可观测到的区别。进一步，我们可以在复数概率体系下定义随机变量的均值，这里不详述了，参加附录 1。

### 三、多则不同

到此为止，量子概率和经典概率没有任何的不同之处。然而当我们考察两个以上事件或随机变量的时候，量子概率和经典概率才能表现出非常不同的特性。

#### 1、事件与希尔伯特空间

我们知道，事件这个概念是概率论的基础。在量子概率体系中，事件不再对应经典的集合，而是对应于线性空间（线性代数没学好的读者，建议先补休一下线性代数），这个空间又称为希尔伯特空间。

仍然以抛硬币为例，我们知道一个硬币在具体的测量之前可以处于一种用复数概率描写的量子力学叠加态：

$$\Psi(X) = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

这个状态可以恰好用一个二维的线性空间表达：

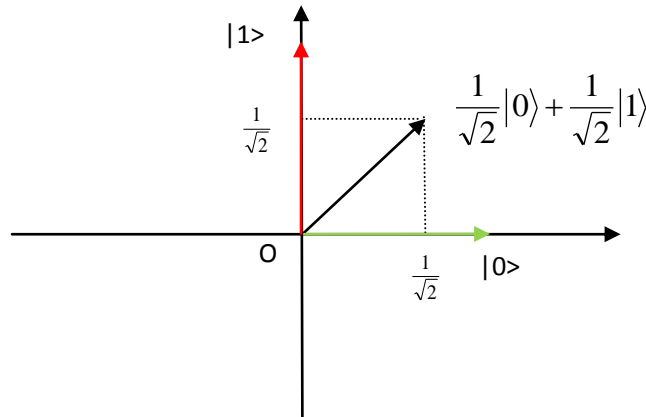


图 1 复数概率的向量表示

在图中，绿色的向量和红色的向量分别表示单位向量 $|0\rangle$ 和 $|1\rangle$

我们知道，在实验中， $0$  表示的是抛掷硬币得到反面， $1$  表示得到正面。出现正面或者反面实际上都是基本的事件。而在复数概率的向量表示中，这两个基本事件则变成了两个相互垂直，交于一点  $0$  的单位向量 $|0\rangle$ 和 $|1\rangle$ 。

这是一个非常关键的区别，在经典概率中，任何原子事件都可以用一个单元素的集合来表示，例如硬币正面朝上的事件可以用 $\{1\}$ 这个集合来表示，集合中的元素只有一个  $1$ 。但是，在复数概率中，这个原子事件则变成了向量 $|1\rangle$ 。读者一定会质疑，我们用向量表示原子事件又有什么好处呢？这与概率取复数值又有什么关系呢？答案就在于，只有这样扩展量子概率，我们才能表达不兼容属性。

## 2、不兼容属性

### (1)、不兼容性的复数概率表示

不兼容属性 (incompatible property) 是量子概率——将概率论扩充到复数域的最特别的概念，也是区别复数概率和经典概率的本质所在。

所谓的不兼容属性就是指一个客观物体所具备的两种属性，这两种属性不能同时得到确切的测量值。当你知道了属性 A 的确切值之后，你就不能获知 B 的确切值了。我们知道，量子力学中有很多测不准对，这些测不准对，例如同一个粒子的位置和动量，或者是能量和时间就是不兼容属性对。

我们先从抽象的数学形式上考察复数概率是如何可以很好地表征不兼容属性以及不确定性原理的 (Uncertainty Principle)。假设某一个系统具有属性 A，属性 A 的取值可以有 {U,D} (即上、下) 两种可能。另外，该系统还具有属性 B，B 的取值可以由 {L,R} (即左、右) 两种可能。这样，我们就可以构成 4 组不同的事件 {A 属性为 U, A 属性为 D, B 属性为 L 和 B 属性为 R}。我们知道，每个事件都可以用线性空间中的向量来表示。由于 A 属性要么取 U 要么取 D，所以前两个事件就可以用两个相互垂直的向量来表示。同样的道理，后两个事件 (B 为 L 或者 B 为 R) 也可以用两个相互垂直的向量表示。显然，前两个事件张成了一个平面，后两个事件又张成了另一张平面，问题的关键是，当这两张平面重合在了一起会如何呢？如下图所示：

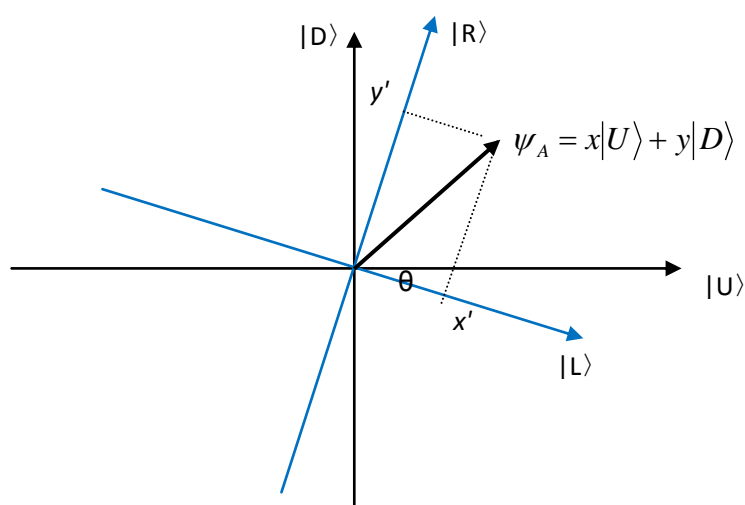


图 2 同一个状态向量在两组不同基下面的表示

如图所示前两个事件 (属性 A 取 U 或者 D 两个属性) 的两个基向量是黑色的坐标系 ( $|U\rangle$  和  $|D\rangle$ )。而后两个事件 (属性 B 取 L 或者 R 两个属性) 对应的基向量则用蓝色的坐标系 ( $|L\rangle$  和  $|R\rangle$ ) 表示。这两个坐标系重合在了同一张平面上，只不过它们之间存在了一个夹角  $\theta$  (在更一般的情况下该夹角可以取复数)。对于同一个向量，例如对图中的粗箭头来说，它在第一个坐标系下可以表示为：

$$\psi_A = x|U\rangle + y|D\rangle$$

那么，在第二个坐标系下 ( $|L\rangle$  和  $|R\rangle$  构成的坐标系) 的坐标就是：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{pmatrix} \quad (*)$$

因此，同样的向量可以表示为：

$$\psi_B = x'|L\rangle + y'|R\rangle$$

因此， $\psi_A$  和  $\psi_B$  是同一个向量在不同坐标系下的表示。回想该向量的物理意义表达的是发生 A 取值 U 事件的概率是  $|x|^2$ ，取值 D 事件的概率是  $|y|^2$ 。而  $\psi_B$  则表达的是 B 属性取值 L 事件的概率是  $|x'|^2$ ，取值 R 的概率是  $|y'|^2$ 。然而，由于  $x', y'$  与  $x, y$  存在着联系 (\*) 式，所以 A 属性上面的概率分布就会对 B 属性造成影响，反之亦然。因此，A 和 B 两个属性之间存在着一种强烈的联系。它们是不兼容性，即不能同时测准。这是因为，如果你确定地知道了 A 属性的取值就不可能完全知道 B 属性的取值了。

比如，假设我们确定地知道 A 属性的取值为 U，即发生 U 事件的概率为 1，则：

$$\psi_A = 1|U\rangle + 0|D\rangle$$

这样根据坐标变换 (\*) 式，同样的状态反映在 B 属性上就成了：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix}$$

这样，只要  $\theta$  不是 0 或者 90 度的整数倍，则我们必然得到 B 属性值是不确定的，它会以  $\cos^2\theta$  的概率取值 L，而以  $\sin^2\theta$  的概率取值 R。因此，我们说 A 和 B 是一对不兼容的属性，因为它们不能同时被测准。其实，在数学上存在着对 L 或者 R 发生的概率更简洁的运算方法，我们不在正文详述，具体请参见附录 2。

总体来说，所谓的不兼容属性对实际上就是两组处于同一个平面的坐标系，它表达的是两个随机变量分布之间的一种深刻的联系。也就是说只要确定了两个不兼容属性，那么它们之间的夹角  $\theta$  就确定下来，因此给定属性 A 上的复数概率分布，就必然会确定一组 B 上的复数概率分布。

## (2)、为什么一定要复数概率？

也许读者会争辩说，在经典概率中，我们也可以引入上述这些坐标系转换的概念。例如我们可以画出类似的图 2 的那种图。我们也可以把经典概率表达成向量：

$$F(X) = \frac{1}{2}|0\rangle + \frac{1}{2}|1\rangle$$

然后，同样把这个向量投影到蓝色的坐标系下，得到新的向量表示：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \cos\theta + \sin\theta \\ -\sin\theta + \cos\theta \end{pmatrix}$$

然而，不幸的是，这种坐标转换后得到的新“概率分布”却不总能满足概率分布的要求。

比如我们设  $\theta = \pi/4$ ，那么  $|L\rangle$  的概率就是  $\sqrt{2}/2$ ， $|R\rangle$  的概率则是 0。所以这对数不能构成一个合格的概率分布。

只有当考虑复数概率的时候，才能引入坐标转换的概念，这是因为，坐标转换本质上来

讲就是一种旋转操作，而旋转操作会保持向量的模（即向量长度）不改变。所以对于任何一个复数概率分布所对应的状态向量  $\psi_A = x|U\rangle + y|D\rangle$  来说，由于它的模必须是 1，所以无论你转换到哪一个坐标系，它都保持长度不变，也就能保证新的向量所对应的经典概率分布是一个良定义的概率分布，即各个分量上面的概率求和为 1。正是因为概率幅而非概率具有这种旋转模不变的性质，所以我们只能对复数概率进行坐标变换的定义。

## (2)、盒子中的萤火虫

下面我们用一个宏观系统的例子来说明不兼容属性对。这个例子最早是由 Cohen（参见：Cohen,D.W., An introduction to Hilbert space and quantum logic, Springer, New York,1989）引入来说明量子概率与经典概率区别的。我们引入这个例子的目的仅仅是说明复数概率的运算法则，有趣的是，如果你真的做出这样一个萤火虫系统，你是不会观察到复数概率预测的那些古怪性质的。因此，不要误把萤火虫系统看作是符合复数概率的量子系统看待。

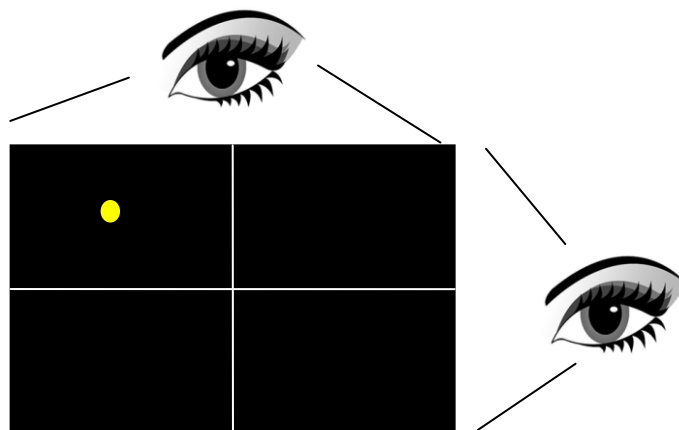


图 3 盒子中的萤火虫

假设在一个黑盒子中放了只萤火虫，萤火虫可以随机地分别穿梭于左上、右上、左下、右下四个方格中的某一个。然而，盒子上仅开了两个洞，洞 A 在盒子的右侧，另一个洞 B 在盒子的上侧。这样，观察者**仅仅能够通过这两个洞的某一个**来观测盒子中的萤火虫。当观察者通过右侧的 A 洞来观察的时候，他仅仅能得出萤火虫是在上 U 或者下 D 的结论，并且概率分别是 1/2 和 1/2。当他通过上方的 B 洞来观察萤火虫的时候，他仅仅能够得到左 L 或者右 R 的结论，并且概率分别是 3/8 和 5/8。但是，上、下的观测以及左、右的观测不能同时精确地做出。这样通过 A,B 两个洞来观测萤火虫就是一对不兼容的属性。

我们可以很好地用刚才讲述的复数概率运算法则来描述这个例子。假设观察者先从 A 洞观察萤火虫，这个时候，我们可以用一个复数概率来表示这种状态：

$$\psi_A = \frac{1}{\sqrt{2}}|U\rangle + \frac{1}{\sqrt{2}}|D\rangle$$

这可以表示为二维平面上的一个向量，如图 3 所示，只不过这个时候的 x,y 都取值  $1/\sqrt{2}$ ，它们表示萤火虫处于上、下两个状态的概率分别是 1/2。



下面我们再考虑从 B 洞观测萤火虫的观察者。我们知道尽管观察者变换了观察系统的视角，但是萤火虫本身并没有改变状态，因此萤火虫的状态表达不能改变。那么如何来表达观察者从 B 洞观测萤火虫的情况呢？我们知道，每一个属性值其实都是线性空间中的一个向量，这样具有两个不同值的属性就构成了一对相互垂直的单位向量。所以从 B 洞观测这个属性就可以表示为两个相互垂直的的向量  $|L\rangle$  和  $|R\rangle$ ，如图 2 表示。

因此，根据 (\*) 式，同样的向量在新基向量  $|L\rangle$  和  $|R\rangle$  下的坐标就是：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \cos\theta + \sin\theta \\ -\sin\theta + \cos\theta \end{pmatrix} \frac{1}{\sqrt{2}}$$

这样，只要我们取

$$\theta = \arccos\left(\frac{\sqrt{3} - \sqrt{5}}{4}\right)$$

就能使得新坐标是  $(\sqrt{\frac{3}{8}}, \sqrt{\frac{5}{8}})$ ，所以同样的状态向量也可以在新基 B 下表示为：

$$\psi_A = \sqrt{\frac{3}{8}}|L\rangle + \sqrt{\frac{5}{8}}|R\rangle$$

按照对复数概率的解释，我们就能以  $3/8$  的概率从 B 孔观测到萤火虫出现在左侧盒子，以  $5/8$  的概率出现在右侧的盒子。所以，我们就用复数概率很好地表达了盒子中的萤火虫这个例子。

仅仅用复数概率来表示萤火虫的例子还不能说明什么，更重要的是，给定了观察的属性对 A,B 两个孔后，那么  $\theta$  值就是固定的数了。假如我们让萤火虫在 U, D 上面的复数概率分布发生改变，例如新的状态是：

$$\psi_A = \frac{i}{2}|U\rangle + \frac{\sqrt{3}}{2}|D\rangle$$

那么，我们必然会得到 L 和 R 上的复数概率分布会发生改变为：

$$\frac{1}{8}((\sqrt{3} - \sqrt{5})i + \sqrt{6(4 + \sqrt{15})})|L\rangle + \frac{1}{8}(3 - \sqrt{15} - i\sqrt{2(4 + \sqrt{15})})|R\rangle$$

对应的概率就是 0.742 在 L 和 0.258 在 R。所以两个概率分布必然存在着一种制约关系。

这才是不兼容属性对的本质。

假设在盒子中有一个小妖精，它可以精确地知道萤火虫到底处于 LU,RU,LD,RD 之中的某一个的概率。比如说，萤火虫会以  $(1/8, 3/8, 1/4, 1/4)$  的概率分别处于 LU,RU,LD,RD 之中。这样，根据边际概率公式，我们就可以求出萤火虫处于 U 的概率是  $1/8 + 3/8 = 1/2$ ，处于 D 的概率是  $1/4 + 1/4 = 1/2$ ，处于 L 的概率是  $1/8 + 1/4 = 3/8$ ，处于 R 的概率是  $3/8 + 1/4 = 5/8$ ，刚好与我们上例中从 A 孔观察和 B 孔观察得到的概率相等。这样，我们用不着什么复数概率，只要通过联合概率  $(1/8, 3/8, 1/4, 1/4)$  就可以对这个问题进行完整的描述了。实际上，我们给出了属性 A

的概率分布和 B 的概率分布，我们总能给出一组联合概率分布数值满足要求。因此，在这一点上看，经典概率也可以模拟不兼容属性对的复数概率行为。

但不同的地方在于，我们的游戏规则规定了观察者只能从 A 或者 B 这两个孔的任何一个观察系统，而不能两个同时来做。

可是，凭什么不能同时测量 A,B 两个属性呢？我们完全可以这样做：让张三盯住 A 孔看，李四盯住 B 孔看。这样每次测量张三能够知道萤火虫在上还是下，李四则知道萤火虫在左还是由，只要把它们的测量结果综合到一起，不就可以获得该系统任意两个属性组合的概率了吗？

但是，仔细想想就会发现，其实能够让两个观察者对不兼容属性对观测的时候有一个前提，这就是要保证张三看到的萤火虫系统和李四看到的萤火虫系统是同一个系统。但是我们也可以这样说，当两个观察者同时观测一个系统的时候，系统就呈现出了两份拷贝，一份呈现在张三眼里，一份呈现在李四眼里，并且这两份拷贝还呈现高度相关，当第一只萤火虫在上的时候第二只也必然在上。那么实际上这种情况与两个观察者看同一个系统是无法分辨的。因此，我们便可以通过研究两个相互纠缠的系统来研究同时测量两个不相容的属性的时候会发生什么。

实际上，这种将不相容测量对归结为一种更底层的联合概率分布的想法就是量子力学中著名的隐变量理论。而让两个人同时测量一个系统的不兼容属性对，并且可以用两个相互纠缠的系统来隐含实现的想法正是著名的 EPR 试验的本质。但是，历史的发展告诉我们，不管是隐变量理论还是 EPR 试验都不能将复数概率行为归结为经典概率行为。

下面我们先引入如何将两个不同的系统复合成一个更大的系统，接下来，我们将介绍纠缠态、贝尔不等式等最能区别复数概率与经典概率不同之处的概念。

### 3、复合系统

所谓的复合系统就是指由两个系统组成的新系统。例如，在刚才的萤火虫的例子中，如果我们考察两个盒子中的两只萤火虫，那么这两只萤火虫就构成了一个复合的系统。假如每只萤火虫在各自的盒子中都有  $|U\rangle, |D\rangle$  两种状态，那么两只萤火虫的组合就有  $|UU\rangle, |UD\rangle, |DU\rangle$  和  $|DD\rangle$  这四种状态分别表示两只萤火虫都在上，一只上一只下，一只下一只上和两只都在下这四中不同的事件。所以，我们要指定这两个萤火虫系统的状态，就必须给定一个四维的空间的向量，分别表示两只萤火虫都在上、一只上一只下，一只下一只上和两只都在下的复数概率。

假如两只萤火虫的状态都是：

$$\psi = \alpha|U\rangle + \beta|D\rangle$$

表示萤火虫以  $(|\alpha|^2)$  的概率处于上，而以  $(|\beta|^2)$  的概率处于下方。那么两只处于该状态的萤火虫所构成的复合系统就对应状态：

$$\begin{aligned} \psi \otimes \psi &= (\alpha|U\rangle + \beta|D\rangle) \otimes (\alpha|U\rangle + \beta|D\rangle) = \alpha|U\rangle \otimes (\alpha|U\rangle + \beta|D\rangle) + \beta|D\rangle \otimes (\alpha|U\rangle + \beta|D\rangle) \\ &= \alpha^2|U\rangle \otimes |U\rangle + \alpha\beta|U\rangle \otimes |D\rangle + \beta\alpha|D\rangle \otimes |U\rangle + \beta^2|D\rangle \otimes |D\rangle \\ &= \alpha^2|UU\rangle + \alpha\beta|UD\rangle + \beta\alpha|DU\rangle + \beta^2|DD\rangle \end{aligned}$$

注意，我们这里定义了一个直积运算  $\otimes$ ，它的作用就是将两个系统粘合成一个新的复合系统，也就是把两个二维向量粘合成一个新的四维向量。它的运算规则符合分配率，并且该运算作用到复数上就是计算这两个复数的乘积。如果该运算作用到向量上，则表示将这两

个向量组合在一起形成新的向量。因此 $|U\rangle \otimes |D\rangle = |UD\rangle$ 。

新向量是一个四维空间中的向量，这个四维空间的基向量分别是 $|UU\rangle, |UD\rangle, |DU\rangle, |DD\rangle$ 。我们看到，按照直积运算得到的新向量 $\psi \otimes \psi$ 就是两个 $\psi$ 向量复合出来的系统状态。它表示以 $(|\alpha|^2)^4$ 的概率两只萤火虫都在上面，以 $(|\alpha|^2)(|\beta|^2)$ 的概率一只在上一只在下……。我们看到这个计算是符合经典概率运算准则的。在经典概率中，如果一个萤火虫以 $|\alpha|^2$ 的概率在上， $|\beta|^2$ 的概率在下，则两只都在上的概率就是 $(|\alpha|^2)^4$ ，一只上一只下的概率就是 $(|\alpha|^2)^2(|\beta|^2)^2$ 。

同样的，对于萤火虫在左侧 $|L\rangle$ 这个事件来说，我们也可以计算两只萤火虫都在左侧的概率。首先，我们需要计算两个 $|L\rangle$ 向量的直积：

$$\begin{aligned} |L\rangle \otimes |L\rangle &= (\cos\theta|U\rangle + \sin\theta|D\rangle) \otimes (\cos\theta|U\rangle + \sin\theta|D\rangle) \\ &= \cos^2\theta|UU\rangle + \cos\theta\sin\theta|UD\rangle + \cos\theta\sin\theta|DU\rangle + \sin^2\theta|DD\rangle \end{aligned}$$

它是四维空间中的一个向量。那么对于任何一个系统状态 $\psi \otimes \psi$ 来说，它在这个四维向量上的投影向量是：

$$\begin{aligned} (\psi \otimes \psi) \cdot (|L\rangle \otimes |L\rangle) |L\rangle \otimes |L\rangle &= (\alpha^*, \alpha\beta, \beta\alpha, \beta^2)^* (\cos^2\theta, \cos\theta\sin\theta, \cos\theta\sin\theta, \sin^2\theta) |L\rangle \otimes |L\rangle \\ &= (\alpha^* \cos^2\theta + 2\alpha^* \beta^* \sin\theta \cos\theta + \beta^* \sin^2\theta) |L\rangle \otimes |L\rangle = (\alpha^* \cos\theta + \beta^* \sin\theta)^2 |L\rangle \otimes |L\rangle \\ &= (\alpha^* \cos\theta + \beta^* \sin\theta)^2 (\cos^2\theta|UU\rangle + \cos\theta\sin\theta|UD\rangle + \cos\theta\sin\theta|DU\rangle + \sin^2\theta|DD\rangle) \end{aligned} \quad (**)$$

因此两个萤火虫都在左侧这个事件的发生概率就是上述向量的模平方。利用投影矩阵的方法，我们可以同样计算出这个向量，这里不再赘述，请参见附录3。总之，有了直积运算，我们就可以把多个系统粘合成一个大系统了。

## 4、纠缠态

下面，我们研究一类特殊的状态，它就是大名鼎鼎的纠缠态(entangled state)。让我们还是以那只可怜的萤火虫为例来说明。假设我们制备了两个萤火虫系统，只不过这两个系统之间存在着微妙的联系。如下图：

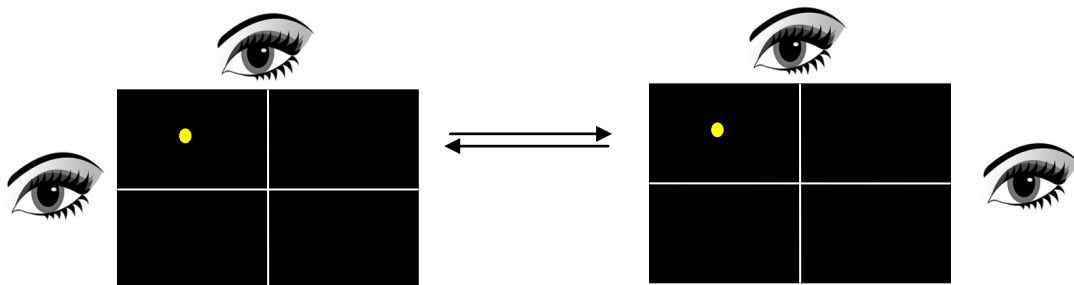


图4 一个处于纠缠态的两个萤火虫系统

显然这是两个萤火虫系统，然而，这两个系统却存在着微妙的联系，也就是左右两只萤火虫会存在着联动。当我们从左面的A孔观察萤火虫的时候，如果左侧那个盒子中的萤火虫在上侧，那么右盒的萤火虫也在上面；反之毅然。那么处于这种纠缠态的萤火虫系统如何

用复数概率表示呢？它是一种特殊的状态，可以表示为：

$$\xi = \alpha|UU\rangle + \beta|DD\rangle$$

首先，由于纠缠态是两个系统所表现出来的性质，因此我们必须在复合系统中来表示这种状态。其次，我们考察 $\xi$  状态的含义，按照复数概率到经典概率的映射，我们知道，系统将会以 $|\alpha|^2$  的概率处于状态 $|UU\rangle$ 之中，即两只萤火虫都处于上侧的概率是 $|\alpha|^2$ 。另外，我们知道在复合系统的中，任何一个向量都有 4 个基，然而上式仅仅出现了两个，这说明 $|UD\rangle$  和 $|DU\rangle$ 两个状态对应的复数概率都为 0。也就是说系统不可能出现一只在上一只在下的状态。

我们知道，对于任何两个萤火虫系统组成的复合系统的状态都可以写成：

$$(\alpha_1|U\rangle + \beta_1|D\rangle) \otimes (\alpha_2|U\rangle + \beta_2|D\rangle) = \alpha_1\alpha_2|UU\rangle + \alpha_1\beta_2|UD\rangle + \alpha_2\beta_1|DU\rangle + \beta_1\beta_2|DD\rangle$$

那么，假如纠缠态 $\xi$  就是两个特殊状态的复合的话，那么就要有：

$$\begin{cases} \alpha_1\beta_2 = 0 \\ \alpha_2\beta_1 = 0 \\ \alpha_1\alpha_2 = \alpha \\ \beta_1\beta_2 = \beta \end{cases}$$

同时成立。但我们知道这是不可能的，除非所有的系数都等于 0。所以，我们知道纠缠态是不能通过将两个单独的状态通过符合得到的。事实上，在量子力学中，人们就将**纠缠态定义为不能表示成两个状态的复合的状态**。

纠缠态被认为是量子系统中的一种与能量、信息同等重要的资源，这正是因为纠缠态的出现，才使得经典概率与复数概率发生了质的不同。下面我们举一个例子来说明。假设两萤火虫系统处于纠缠态：

$$\xi = \frac{1}{\sqrt{2}}|UU\rangle + \frac{1}{\sqrt{2}}|DD\rangle$$

即两只萤火虫会以 1/2 概率同时出现在上侧，并以 1/2 的概率同时出现在下侧。下面再定义从 B 孔观察萤火虫得出 L,R 性质的测量与水平轴成角是 60 度。这样，根据直积运算，我们可以计算出来左盒的萤火虫处于上（下）侧、右盒的萤火虫处于上（下）侧，左盒的萤火虫处于左（右）侧，右盒的萤火虫处于左（右）侧的概率都是 1/2。同时，还可以计算出来左盒的萤火虫处于上侧，而右盒的萤火虫处于左侧的概率是 1/8（参见附录 4）。

同样的道理，我们按照相似的方法计算出来从 A 孔观察左侧萤火虫，从 B 孔观察右侧萤火虫的各种组合的概率如下表：

表 1. 复数概率计算得到的伪联合概率分布

	U(左盒)	D(左盒)
L(右盒)	1/8	3/8
R(右盒)	1/8	3/8

大家注意，这个表得出来的正是对上一节讨论的对同一个系统的两个不兼容属性由两个观察者同时测量的对应物。只不过，在这里我们把同一个系统换成了两个相互纠缠的萤火虫系统了。这样左盒对应的就是左侧观察者看到的系统，右盒就是右侧观察者看到的系统。我们认为，实际上两个观察者无法分辨究竟他们同时测量的萤火虫系统是同一个系统，还是两个不同的系统，但发生着相互纠缠。因此我们便可把对不兼容属性对同时测量的问题归结为对两个相互纠缠系统的测量问题。

表 1 相当于一个联合概率分布。然而这个联合分布是通过测量左侧萤火虫的 A 属性，

同时测量右侧萤火虫的 B 属性得到的。根据这个联合分布，我们就能计算出单一属性的边际概率，我们知道萤火虫在左侧的概率就是  $1/8+3/8=1/2$ ，在右侧的概率也是  $1/2$ 。而萤火虫在上侧的概率是  $1/4$ ，在下侧的概率是  $3/4$ 。然而，之前我们已经计算出了萤火虫单独处于上侧的概率是  $1/2$ ，而非  $1/4$ 。出现了矛盾！

这个矛盾说明，我们不能将表 1 当作一个联合概率分布，也不能从它反过来构造某一个属性的边际概率分布，我们已经看出了复数概率系统与经典概率系统计算的差异。换句话说，一个复数概率系统中的不兼容属性对本质上是不能归结为背后的一种经典的联合概率分布模型的。

## 5、贝尔不等式——真正的破坏

然而，最本质地表征量子概率与经典概率的不同之处的还是贝尔不等式。贝尔不等式是科学家约翰·贝尔首先发现的经典概率系统中的一组不等式。然而，当采用复数概率的方法进行同样的计算的时候，这一不等式却被打破了，这说明复数概率法则存在着经典概率无法模仿的一面，从而证明了不存在经典的隐变量理论。进一步，人们实施了各种试验观察到电子的行为的确会破坏贝尔不等式，也就是说电子的行为不能够用隐含的经典概率来解释，而必须采用复数概率。

虽然原始的贝尔不等式非常复杂，但是我们可以简化它。一组最简单的贝尔不等式<sup>1</sup>（参见：<http://www.mtnmath.com/whath/node60.html>，原始的贝尔不等式讨论的是两个变量之间的相关性质，简化的贝尔不等式则集中到概率的性质，参见：[Quantum Probability-Quantum Logic](#)）就是下式：

$$1 \geq P(A) + P(B) - P(A \cap B) \geq 0 \quad (***)$$

有 A, B 两个事件，那么 A 事件的概率加上 B 事件的概率减去 A, B 两事件同时发生的概率必须在 [0,1] 之间，我们用集合文式图可以很清楚地表达出来这个不等式的含义：

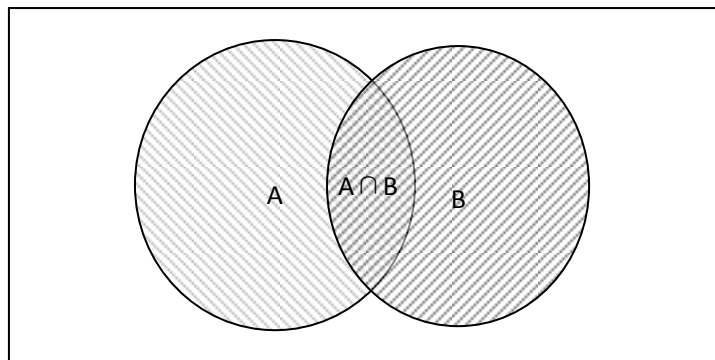


图 5 贝尔不等式的文氏图

如图所示，假设集合的面积对应的就是该事件的概率。全集的面积为 1。这样  $P(A \cap B)$  表示的就是图中方格阴影部分的面积，而  $P(A)+P(B)-P(A \cap B)$  就是斜影部分的面积。很显然当 A 和 B 完全重合的时候这部分面积最小为 0。当 A 和 B 完全不重合并且  $A \cup B$  为全集的时候这部分面积为最大为 1。所以对于任意给定的两个事件 A 和 B，这两个不等式必然成立。

<sup>1</sup> 严格来说，该不等式不能叫贝尔不等式，它是贝尔不等式的极其简化的版本。但为了叙述方便，我们仍称它叫作贝尔不等式。

然而，如果采用从复数概率出发得到的经典概率却有可能打破这两个不等式。让我们还是用萤火虫系统来说明。假设有两个黑箱子，它们分别关了两只萤火虫。并且这两只萤火虫是反关联的。也就是当左侧的萤火虫在上面，右侧的萤火虫就在下面，反之亦然。如下图式：

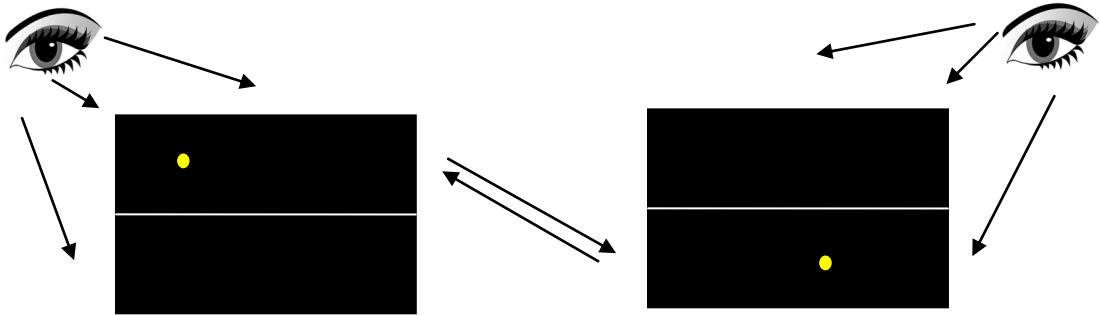


图 6 破坏贝尔不等式的量子纠缠态

这样，我们可以用一个纠缠态来表述这两个萤火虫构成的系统：

$$\xi = \frac{1}{\sqrt{2}}|UD\rangle - \frac{1}{\sqrt{2}}|DU\rangle$$

进一步，我们假设观察者在箱子上凿了三个小孔，透过这些孔可以观察到萤火虫是处于孔的左侧还是右侧。假设从其中任意一个孔观察萤火虫得到左或者右分别可以用下列的投影矩阵表示<sup>2</sup>：

$$M_{\theta}(L) = \frac{1}{2} \begin{pmatrix} 1 & e^{-i\theta} \\ e^{i\theta} & 1 \end{pmatrix}, \quad M_{\theta}(R) = \frac{1}{2} \begin{pmatrix} 1 & -e^{-i\theta} \\ -e^{i\theta} & 1 \end{pmatrix}$$

其中 $\theta$ 表示从不同的孔观察的参数。我们设这三个不同的孔对应的参数分别为 $0, \pi/4, \pi/2$ 。下面，我们来考察三个事件：

- 事件 1：从第 1 个孔看到左盒的萤火虫在左侧，并且从第 2 个孔看它在右侧；
- 事件 2：从第 2 个孔看到左盒的萤火虫在左侧，并且从第 3 个孔看它在右侧；
- 事件 3：从第 1 个孔看到左盒的萤火虫在左侧，并且从第 3 个孔看它在右侧。

我们记事件  $A_i$  为从第  $i$  个孔看萤火虫在左侧，且用  $\overline{A_i}$  上面加一横杠表示从第  $i$  个孔看萤火虫在右侧。那么，前面的三个事件就可以写作公式：

$$\text{事件 1: } A_1 \cap \overline{A_2}, \quad \text{事件 2: } A_2 \cap \overline{A_3}, \quad \text{事件 3: } A_1 \cap \overline{A_3}$$

这样事件 1 和 2 的交就是事件 3，所以应该符合下列概率不等式：

$$0 \leq P(1) + P(2) - P(3) \leq 1$$

另一方面，由于我们的设定，我们不能同时从两个孔看同一只萤火虫，但是我们可以利用两只萤火虫互相纠缠的性质，第一个观察者从第一个孔看左侧的萤火虫，第二个观察者在第 2 个孔看右侧的萤火虫。由于萤火虫反关联的性质，上面的三个事件又等价于：

- 事件 1'：从第 1 个孔看到左盒的萤火虫在左侧，并且从第 2 个孔看到右盒的萤火虫在左侧；
- 事件 2'：从第 2 个孔看到左盒的萤火虫在左侧，并且从第 3 个孔看右盒的萤火虫在左侧；

<sup>2</sup> 在原例中，是从 0,45 度，90 度这三个不同的角度测量电子的自旋，为了隐去问题的物理含义，我们采用盒子上面开了三个不同的小孔来比喻

事件 3'：从第 1 个孔看到左盒的萤火虫在左侧，并且从第 3 个孔看右盒的萤火虫在左侧。

因而我们可以定义这三个事件的投影矩阵，从而计算它们的概率，例如事件 1' 对应的投影矩阵是：

$$M_0(L) \otimes M_{\pi/4}(L) = \frac{1}{2} \begin{pmatrix} 1 & e^{-i0} \\ e^{i0} & 1 \end{pmatrix} \otimes \frac{1}{2} \begin{pmatrix} 1 & e^{-i\pi/4} \\ e^{i\pi/4} & 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & e^{-i\pi/4} & 1 & e^{-i\pi/4} \\ e^{i\pi/4} & 1 & e^{i\pi/4} & 1 \\ 1 & e^{-i\pi/4} & 1 & e^{-i\pi/4} \\ e^{i\pi/4} & 1 & e^{i\pi/4} & 1 \end{pmatrix}$$

将该矩阵乘以纠缠态的状态向量  $(\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}})$ ，再求所得向量的模平方就得到事件 1'

的概率是  $\sin^2(\pi/8)/2 \approx 0.073$ 。采用同样的方法可以计算出事件 2' 的概率是  $\sin^2(\pi/8)/2 \approx 0.0732$ ，事件 3' 的概率是  $\sin^2(\pi/4)/2 \approx 0.25$ 。这样这三件事的概率就不满足不等式：

$$P(1') + P(2') - P(3') \geq 0$$

从而破坏了这个最简单版本的贝尔不等式。这个例子说明，采用复数概率运算的纠缠系统是不符合经典概率法则的。

到此为止，我们才算真正领教了复数概率的威力，它在一些实例中能够破坏经典概率必须满足的贝尔不等式，这才是复数概率与经典概率最大的差别所在。我们看到，要想真正地破坏经典概率法则，纠缠态和不兼容的测量对都是必不可少的。在本例中，两个相互纠缠的萤火虫系统构成了测量对，而从三个不同的孔洞对萤火虫进行观察则构成了三个两两不相兼容的测量对，正是在这两个条件都具备的前提下才能真正破坏贝尔不等式。

假如你真的找来一只萤火虫，把它关进黑盒子里，再钻三个孔对它进行左、右的划分。你是否能看到按照复数概率法则运算得到的破坏贝尔不等式的情况呢？答案是否定的，我敢肯定，如果你做这样的试验，你肯定能够得到经典概率的结果。原因在于对于一切实存的系统来说，贝尔不等式(\*\*\*)都必须成立。而量子则是一类波、粒二象性共存的非实存系统。换句话说，如果你能找到一个破坏该不等式的例子，那么你就找到了一类非实存系统。你可以尝试构造这样的例子，你会发现，要真的打破该不等式是非常困难的。

## 四、量子概率体系

前面我们按照一种纯粹启发式的思路，也就是如何将经典的实数概率扩展到复数概率的角度论述了整个一套新的概率体系，在文献中这套心得概率体系就被称为量子概率。这种思路的好处在于能够让读者较快地接触到复数概率这套理论最核心的概念，以及它与经典概率的不同之处。然而，这种方法的弱点在于叙述过于零乱，而且不能保证数学的清晰性。为了弥补这个不足，我们特意从较粗糙的公理化角度浅显地论述整个量子概率公理体系，这套公理化体系主要是由 von Neumann 在他的名著《[Mathematical Foundations of Quantum Mechanics](#)》建立起来的。读者可以参考[这篇文章](#)获得更清晰的描述。

# 1、基本概念

## (1) 事件

在经典概率论中，我们用集合表示事件，用集合的交、并、补等运算来组合出更多的事件。同时，我们定义了两个特殊的事件，一个是不可能事件对应集合空集，另外一种事件为必然发生事件，它对应集合全集。

在量子概率中，我们用复线性空间（希尔伯特空间） $H$  作为基本事件的集合。任意一个事件都是该线性空间中的一个子空间（直线、平面或者超平面）。在量子概率中，不可能事件对应着  $0$  向量，即  $0$  维的线性空间。必然事件则对应着整个线性空间  $H$ 。

正如经典逻辑可以使用否、与、或等运算组合出各种新的事件一样。在量子概率中，我们也可以定义这三种运算。

**定义 1 (否运算):** 设事件  $A$  对应的子空间为  $L_A$ ，那么  $\sim A$  事件则对应着垂直于  $L_A$  的子空间记做  $L_A^\perp$ 。

例如，事件  $A$  对应的子空间  $L_A$  为三维空间中的一条直线  $l$ ，那么非  $A$  这个事件对应的子空间就是垂直于  $l$  的整个平面。

**定义 2 (与运算):** 设事件  $A$  对应的子空间  $L_A$ ，事件  $B$  对应的子空间  $L_B$ ，那么事件  $A$  且  $B$  对应的子空间就是这两个线性子空间的交集，即

$$L_{A \cap B} = L_A \wedge L_B$$

例如  $L_A$  和  $L_B$  都是三维空间中的两张平面，那么  $L_{A \cap B}$  就是这两张平面的交线。如果  $L_A$  和  $L_B$  都是三维空间中的直线，则它们的交  $L_{A \cap B}$  必然是一个点，这个点就是  $0$  维线性空间，因此是不可能事件。

**定义 3 (或运算):** 设事件  $A$  对应的子空间  $L_A$ ，事件  $B$  对应的子空间  $L_B$ ，那么事件  $A$  或  $B$  对应的子空间就是由这两个空间所张成的更大空间。即：

$$L_{A \cup B} = \text{Span}(L_A, L_B)$$

注意，量子概率中的或运算与经典概率中事件的或运算是本质不同的，它不是两个子空间的并集，而是线性扩展  $\text{Span}$ 。例如假如  $L_A$  和  $L_B$  分别是两条相交于一点直线，那么  $L_{A \cup B}$  就是这两条直线所张成的平面。

量子概率的运算性质大多与经典事件的运算性质相同，但有一个本质的区别，这就在于量子概率中的事件运算不一定满足分配率。即：

$$L_A \wedge (L_B \vee L_C) \neq (L_A \wedge L_B) \vee (L_A \wedge L_C)$$

但我们知道集合的运算是满足这个性质的。那么量子事件的运算是如何打破分配率的呢？让我们来看一个例子，如图：



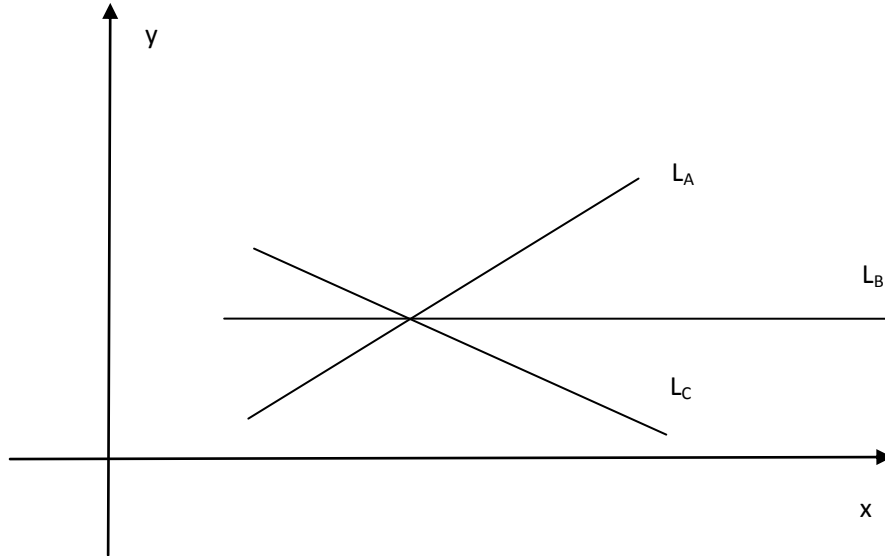


图 7 破坏分配律法则的量子事件运算

如图中所示的空间中有三条同在一个平面内的直线  $L_A, L_B, L_C$  分别对应事件  $A, B, C$ ，则  $L_B \vee L_C$  就是整个平面，而  $L_A \wedge (L_B \vee L_C)$  就是  $L_A$  这根直线。但反过来，由于  $L_A \wedge L_B$  和  $L_A \wedge L_C$  都是点，所以  $(L_A \wedge L_B) \vee (L_A \wedge L_C)$  还是点，即不可能事件。所以量子事件不一定满足分配律。

## (2) 互斥事件

一枚硬币要么朝上，要么朝下，不可能同时发生，每抛掷一次必然有一个发生，那么这两个事件构成了互斥事件。从数学上我们可以表达成：

**定义 4 (互斥事件)** 给定一组事件  $A_1, A_2, \dots, A_n$ ，它们分别对应子空间  $L_1, L_2, \dots, L_n$ ，则这些子空间满足：

$$L_i \wedge L_j = O, \text{ and } L_i \perp L_j, \text{ for } \forall i, j,$$

并且：

$$\text{Span}(L_1, L_2, \dots, L_n) = H$$

我们称这组事件互斥。

一组互斥事件往往对应同一个属性取不同的属性值。例如硬币只有正反两个可能取值。有趣的是，在经典概率中，我们通常可以用一条线上的不同点对应不同取值，例如一个人的身高有多种可能取值，这样所有的身高可以表示为一条直线。但是，在量子概率中，每一个属性值都对应着一条直线，这样  $n$  组属性值就对应着  $n$  条直线，并且这些属性值都是两两互斥的，也就是说这些直线两两垂直，而且所有这些可能的直线张成了一个  $n$  维希尔伯特空间。如下图，经典事件和量子事件表示的区别：

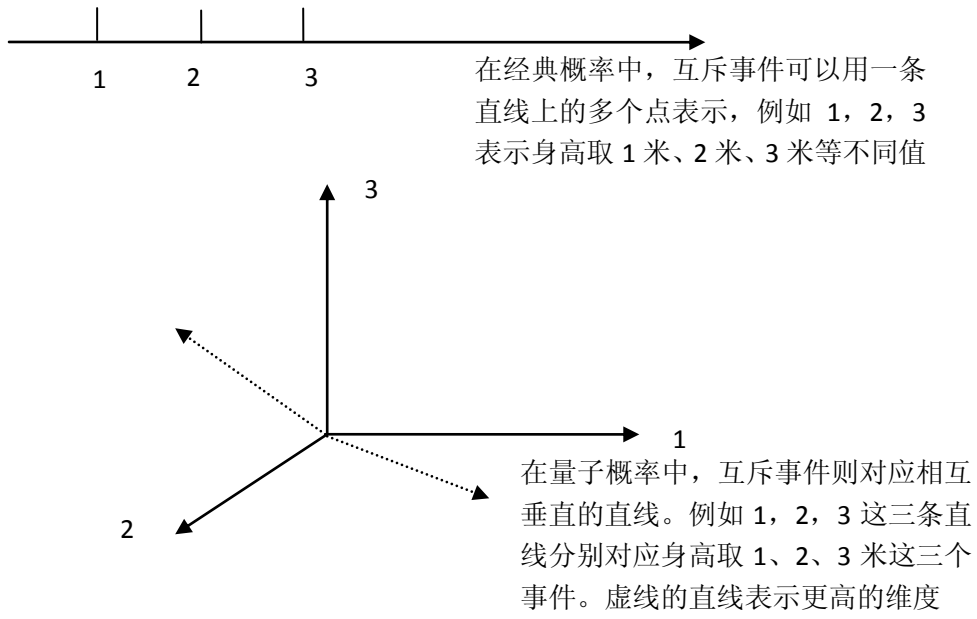


图 8 经典概率与量子概率在互斥事件上表示的不同

## (2) 概率与测量

在量子概率理论中，我们可以采用三个步骤来计算任意一个事件的概率。首先，我们要确定一个状态来表示系统所处的环境和条件。我们用  $|z\rangle$  来表示该状态，它是希尔伯特空间  $H$  中的一个向量，并且这个向量的长度必须是 1。即：

$$\langle z | z \rangle = 1$$

这里， $\langle z |$  表示向量  $|z\rangle$  的共轭，而  $\langle z | z \rangle$  则表示向量  $\langle z |$  和  $|z\rangle$  的内积。例如如果  $|z\rangle = (1+i, 2-i)$ ，则  $\langle z | = (1-i, 2+i)$ ，那么  $\langle z | z \rangle = (1+i)(1-i) + (2-i)(2+i) = 2+5=7$ 。这是量子力学中的狄拉克符号记法。

其次，我们需要定义投影算子的概念。

**定义 5: (投影算子)** 每个事件  $A$  都对应了一个投影算子  $P_A$ ，我们可以把状态向量  $|z\rangle$  通过  $P_A$  的作用投射到事件  $A$  对应的子空间  $L_A$  上，因此，利用数学语言，投影算子可以表达为：

$$P_A : H \rightarrow L_A$$

在几何上，投影算子就是求向量  $|z\rangle$  到子空间  $L_A$  上的投影向量，因此，这个投影就构成了一个向量：

$$P_A |z\rangle$$

这个概念可以用下列图形清晰地表达：

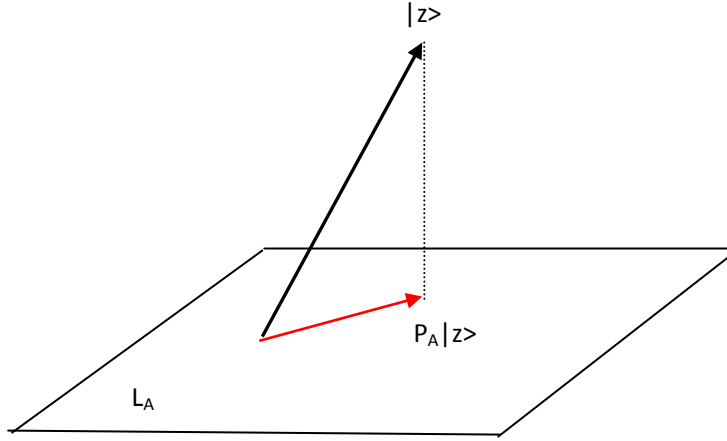


图 9. 向量的投影

在图中，原向量为  $|z\rangle$ ，投影空间  $L_A$  为一个平面，则  $|z\rangle$  在其上的投影为红色向量。

如果我们给  $H$  选定了一组基向量，并且  $|z\rangle$  也用基向量表示，那么投影算子  $P_A$  就对应了一个矩阵称之为投影矩阵（参见附录 2），这个矩阵满足性质：

$$P_A P_A = P_A$$

从几何意义上很容易理解这个性质：任何一个向量往事件  $A$  对应的线性子空间上投影之后得到的向量再往  $A$  上投影则必然保持不变。

最后，我们计算该投影向量（红色向量）的模平方即为事件  $A$  的发生概率：

$$\Pr(A) = |P_A|z\rangle|^2 = \langle z|P_A|z\rangle$$

这里介绍的计算事件概率的算法与前面讲到的复数概率到经典概率的映射相同。在量子概率中，测量是一个十分重要的过程，它不仅仅决定了一个事件发生的概率，而且也改变了系统的状态。也就是说，如果我们测量过状态是  $z$  的量子系统，并且得到事件  $A$  发生的结果，那么该量子系统的状态将变成：

$$z' = \frac{P_A z}{|P_A z|}$$

也就是说，我们将用  $A$  事件的概率  $|P_A z|$  去重新归一化向量  $P_A z$ ，这样测量之后的状态  $z'$  又是一个长度为 1 的单位向量了。假如我们对状态  $z'$  再进行一次完全相同的测量，就会得到：

$$\Pr(A) = |P_A z'|^2 = |P_A \frac{P_A z}{|P_A z|}|^2 = \frac{|P_A P_A z|^2}{|P_A z|^2} = \frac{|P_A z|^2}{|P_A z|^2} = 1$$

所以，一旦测量  $P_A$  得到  $A$  发生的概率之后，系统将会一直呈现出  $A$  事件保持不变。有趣的是，在量子概率中，我们将测量与投影算子密切地联系起来。一次测量，就是将某一个投影算子作用到一个状态向量上，测量的结果是让观察者得到一个确定的事件，同时也会让被测向量完成一次投影，形成一个新的向量。

我们可以将量子概率的运算法则应用到一组互斥事件上，这就相当于将一个高维空间中的向量  $|z\rangle$  分别向一组相互垂直的直线投影，这样，可以验证量子概率满足测度的性质，即如果  $C=A \vee B$ ，并且  $A \wedge B=O$ ，那么  $\Pr(C)=\Pr(A)+\Pr(B)$ 。

### (3) 不相容属性对

我们说过，量子概率与经典概率最大的不同就是存在着相互不相容的属性对，也就是不能同时进行测量的两个属性。下面，我们给出不相容属性对的定义：

**定义 6: (不相容属性对)** 假设属性  $M$  有  $m$  个不同的属性值  $\{M_1, M_2, \dots, M_m\}$ ，这些属性值张成了  $m$  维线性空间  $H$ ，并且每两个属性对所对应的子空间彼此垂直。另有一个属性  $N$ ，它也有  $m$  个属性值  $\{N_1, N_2, \dots, N_m\}$ ，这些属性值对应的子空间也张成一个  $m$  维的线性空间  $H'$ ，如果  $H=H'$ ，则称  $M$  和  $N$  这两个属性不相容。

我们仍然从图形的角度来理解这个概念，假设  $m=2$ ，这样  $M$  这个属性的两个属性值就张成了一个平面  $H$ 。那么  $N$  这个属性的两个属性值张成的二维空间  $H'$  也是  $H$ ，这就意味着， $N$  对应的是  $H$  中的另外一个坐标系。如图 10 所示，黑色的坐标系就表示  $M$  属性，蓝色的坐标系表示  $N$  属性。

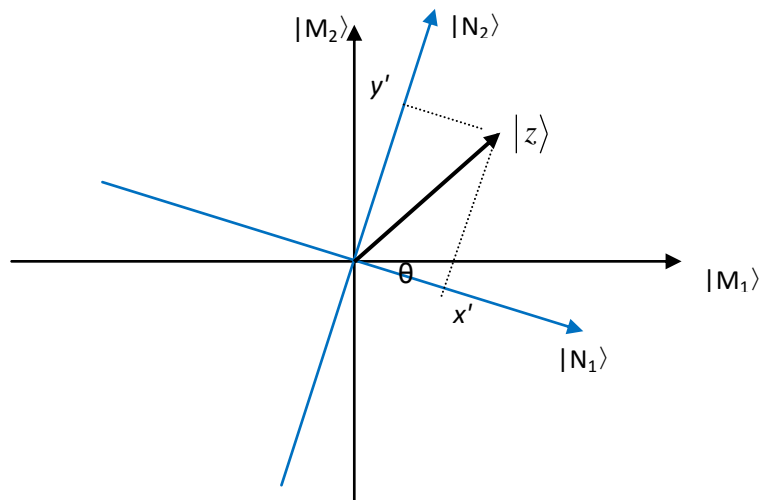


图 10 不兼容属性对的测量

下面我们考虑两组不同的测量，第一组是先测量事件  $M_1$  是否发生，再测量事件  $N_1$  是否发生，第二组测量是先测量  $N_1$ ，再测量  $M_1$ 。按照前面定义的测量规则，第一组测量相当于把向量  $|z\rangle$  先投影到  $M_1$  直线上，然后将  $M_1$  上的单位向量投影到  $x'$  上。这两种结果最终得到的系统状态是完全不同的，第一组测量最终得到的是  $|N_1\rangle$  上的单位向量，而第二组测量将得到  $|M_1\rangle$  上的单位向量。这不同的测量顺序导致不同的结果也会体现在投影算符上，即存在不等式：

$$P_{Ni}P_{Mj} \neq P_{Mj}P_{Ni}$$

我们看到，不同的测量顺序会导致不同的测量结果，这正是量子概率不兼容属性的一种特别的性质之一。

### (4) 相容属性对

如果两个属性  $M$  和  $N$  可以被同时测量，则它们就构成了相容的属性对。在数学上，相容的属性对可以用复合系统来表示。

**定义 7 (相容属性对):** 假设属性  $M$  有  $m$  个不同的属性值  $\{M_1, M_2, \dots, M_m\}$ ，另有一个属性  $N$ ，

它有  $n$  个属性值  $\{N_1, N_2, \dots, N_n\}$ ，那么如果这两组属性值可以张成  $m \times n$  维空间  $H$ ，并且每两个属性对的复合，即：

$$M_i \otimes N_j, \text{ for any } i, j$$

都构成了  $m \times n$  维空间  $H$  中的一组基，那么这两个属性值就是相互兼容的。

所以，这  $m \times n$  维的希尔伯特空间就可以写成：

$$H = \text{Span}(\dots, M_i \otimes N_j, \dots), \forall i < m, j < n$$

下面，我们来研究一下相容属性对的测量。假设我有两个相容属性对，每个属性都具有 2 个可能的属性值（我们都用 01 来表示）。这样，全空间  $H$  就可以写成这两个属性对每两个属性值的组合张成的空间。

$$H = \text{Span}(|00\rangle, |01\rangle, |10\rangle, |11\rangle)$$

在这些基中，写在左侧的表示第一个属性值。下面我们对第一个属性值=0 进行测量，而当第一个属性=0，第二个属性还有 0 和 1 两种可能，因此剩下的属性就张成了一个二维的线性空间，并且  $|00\rangle$  和  $|01\rangle$  构成了这个二维子空间的基向量。所以，对第一个属性值测量就相当于把一个四维空间中的向量投影到一个二维的平面上。

相容属性对的测量是不区分顺序的。比如我们先对第一个属性值=0 进行测量，这样就把 4 维投影到  $|00\rangle$  和  $|01\rangle$  构成的二维平面上，然后再对第二个属性值是否为 0 进行测量，这样就把向量投影到  $|00\rangle$  这个向量上面。反过来，如果先对第二个属性值是否为 0 进行测量，就把 4 维向量投影到  $|00\rangle, |10\rangle$  这个平面上，然后再对第一个属性值是否为 0 进行测量，就把向量又投影到  $|00\rangle$  直线上，所以，这两种不同顺序的测量是完全相同的。也就是说兼容的属性对是不区分顺序的。

实际上，假如我们用  $n$  个不兼容属性对构成了一个庞大的系统，那么对这些属性的测量就是一个降维的过程，我们会把一个  $2^n$  维中的向量逐渐压缩到一条直线上。然而，对不兼容属性的测量则不会产生降维的现象。如图 10 所示，我们将一条二维空间中的直线投影到不兼容属性值的某一条直线上，则这个向量仍然在二维空间中，并且这个投影还可以再次投影到这个空间中的另外一个不兼容属性上。

## 2、与经典概率的区别

我们看到，量子概率无非提供了另外一套不同于经典概率的概率运算法则。那么究竟量子概率与经典概率有什么区别和联系呢？

首先，针对一个属性的测量来说，量子概率和经典概率会得到完全相同的结果。进一步，对于兼容的属性对来说，量子概率与经典概率仍然没有区别。换句话说，量子概率中的兼容性运算已经能够涵盖所有经典概率论的内容。

但是，针对不兼容属性对来说，很多经典概率的运算法则都不再成立。同时量子概率还预言了一些新的性质是经典概率中没有的。我们列出下表来表示这些区别和联系。

性质	经典概率运算	不兼容属性的事件 A 和 B 的量子概率运算
联合概率： $\Pr(A \wedge B)$	$\Pr(A \wedge B) = \Pr(B \wedge A)$	无定义

条件概率: $\Pr(A B)$	$\Pr(A B)=\Pr(A \wedge B)/P(B)$	$\frac{ P_A P_B z\rangle ^2}{ P_B z\rangle ^2}$
条件概率: $\Pr(B A)$	$\Pr(B A)=\Pr(A \wedge B)/P(A)$	$\frac{ P_B P_A z\rangle ^2}{ P_A z\rangle ^2}$
全概率公式	$\Pr(A) = \sum_i \Pr(A   B_i) \Pr(B_i)$	不成立
条件概率的互易性: $\Pr(A B)=\Pr(B A)$	不满足	满足
双向随机性 $\sum_i \Pr(A_i   B_j) = \sum_j \Pr(A_i   B_j) = 1$	不满足	满足

既然经典概率系统的性质都可以在量子概率的兼容属性测量中找到对应，但是反过来，量子概率中的不兼容属性对具备的性质则具有特殊性，所以可以说量子概率是比经典概率蕴含了更多东西的概率运算系统。

## 五、应用与展望

### 1、量子认知

我们知道，量子概率中一个重要的性质就是测量在整个运算过程中的应用。测量是将量子的虚幻世界转变为现实世界的一条重要途径。深刻理解这个测量过程会发现，它与我们人类的自由意志存在着某种深刻的联系(参见: Stapp H.P. *Mindful Universe: Quantum Mechanics and the Participating Observer*. Springer, 2007)。因为，人类的意识本身就是在将各种不同的可能性转变成现实。这就使得很多大科学家和哲学家认为量子概率不仅能够描述微观粒子世界，也能够描述人类的意识、认知行为(例如彭罗斯所著的《皇帝新脑》)。虽然早期的应用充满了各种哲学思辨，但是最近，人们已经将量子概率应用到认知行为的解释方面做出了一些工作。

比如，J. R. Busemeyer 和 Z. Wang 就发现人类的决策就是一种量子行为(请参看[这个PPT](#))，即可以用量子概率描述的宏观系统。他们设计了一组人类行为学试验，让人类被试作出决策。

首先，他们让每个人类被试观察下列一组头像图片：



图 11 Busemeyer 和 Wang 的试验

然后让被试在两组不同的实验条件下完成判断：是否会把这个人当做自己的朋友。在第一组实验中，试验者要求被试直接作出判断，会不会把这张脸当作自己的朋友。第二组试验则要求被试先把所有的脸归为两类：好人或者坏人，然后再根据这个归类的结果判断是否把这人当作自己的朋友。

下面我们考虑用量子概率的方法对人类的行为进行建模。首先，我们利用是否是朋友这个属性作为基向量建立一个二维的线性空间，如下图：

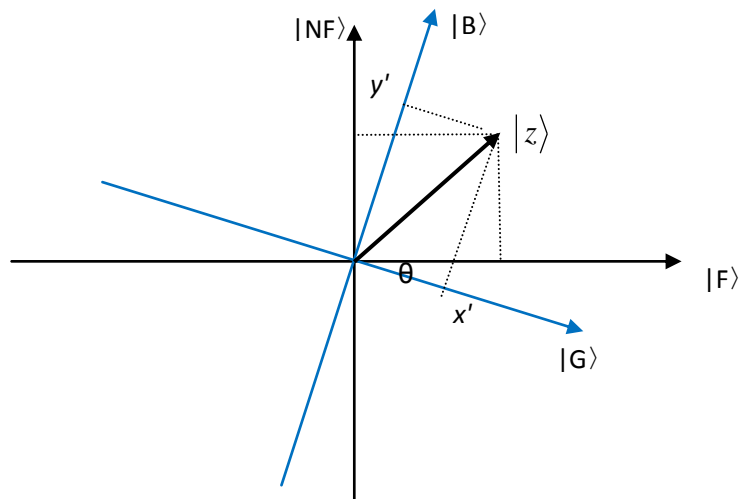


图 12 人类行为试验的量子概率表示

其次，我们假设被试初始的状态可以用向量 $|z\rangle$ 来表示。那么判断这张脸是属于好人还是坏人则是一个与是否为朋友这个属性的不兼容的属性，用蓝色的坐标系表示。

第一组试验要求被试直接做出朋友或者不是朋友的选择，而不管被试是把这张脸分类成好人还是坏人。这就相当于直接把向量 $|z\rangle$ 往 $|F\rangle$ 上进行投影，假设这个投影的长度也就是判断这张脸是朋友的概率为  $P(F)$ 。

在第二组情况下，试验员相当于进行了两次连续的测量，首先让被试做出分类为好人还是坏人的测量，之后再作出是否把此人当作朋友的测量。这就相当于将向量 $|z\rangle$ 先投影到 $|G\rangle$ 或者 $|B\rangle$ 上，然后再将它投影到 $|F\rangle$ 上面，我们设在这种情况下测得的朋友概率是  $P'(F)$ 。根据投影测量的性质，我们知道， $P(F)$ 是不等于  $P'(F)$ 的，并且究竟哪一个概率大，需要由 $|z\rangle$ 以及两个坐标轴夹角决定。

假如人类的决策是按照经典概率运算的法则，那么  $P(F)$ 就应该等于  $P'(F)$ 。但是 J. R. Busemeyer 和 Z. Wang 却发现试验结果是  $P(F) > P'(F)$ 。因此，人类决策行为的这种对测量的敏感性可以很好的用量子概率来解释。(参考[此文](#))。

我们知道，行为学家们告诉我们，人类的决策行为有很多反常规的特性，这些特性被称为人类决策的悖论（参见著名经济学家 Kanneman 的相关研究工作）。近年来，有些人将量子概率用于解释这些决策悖论取得了较好的效果。

然而，笔者认为，要验证人类认知、决策行为是否符合量子概率法则的一个重要因素就是考察贝尔不等式 (\*\*\*) 是否能够被打破。如果我们在人类决策的例子中的确找到了打破贝尔不等式的例子，那么就有十分充足的理由说人类的决策行为不可以用经典概率描述。这里的困难之处在于，我们如何去构造类似纠缠态这样的人类行为状态呢？这里还有很长的路要走，我们将另写文章评述。

## 2、展望

当我们领教了复数概率以及量子概率体系的威力之后，冷静地思考，其实最核心的内容就在于不兼容属性对。我们能否将量子概率应用宏观体系中的关键就在于是否能够在宏观世界中找到不兼容的属性对。

然而，一旦我们试图去寻找，就会发现这实际上非常困难。有什么系统的性质会展现出类似电子那样速度和位置的不能同时被测准的性质呢？

也许在寻找这种不兼容属性对之前，应该深入理解不确定原理的含义。我们说 A 和 B 这两个属性是不兼容的，实际上是说 A 这个属性上各种值的概率分布与 B 属性的概率分布存在着紧密的联系。注意这种联系并不是说 A 和 B 两个属性的直接联系，而是说它们概率分布的间接联系。也就是说，如果我们说 A 表示人的身高，并且 B 表示此人的腿长，那么 A 和 B 通常会成正比关系，这两个属性就不是不兼容属性。用数学语言来说，不兼容属性对描述的是  $p(B)=f(p(A))$  这种关系，而非  $B=f(A)$  的关系，也就是一种对两个属性概率分布关系的间接描述。

更进一步，这种间接联系描述的是一种不确定性的制衡。按照不确定原理（Uncertainty Principle），我们对 A 属性知道的越多，对 B 属性知道得越少。这种知道的多和少既可以用误差  $\Delta A$  来表示，也可以用熵  $S(p(A))$  来表示。所以，也有人证明了不兼容属性对之间也存在用熵来表达的不确定原理，形如： $S(A)+S(B)>const$ （请参见：<http://arxiv.org/abs/1001.4668>）。

对概率的理解可以用客观频率法和主观概率法两种手段。如果从客观频率法的角度来理解不兼容属性对的确非常困难，那么我们不妨从主观概率的角度来理解。按照主观学派，我们对谋事件的发生概率的解释不是该事件的频率，而是主观认为的对该事件发生的信念程度。按照这种理解，不确定原理在说什么呢？它其实在说，我们对两个属性 A 和 B 的认识存在着局限性，如果我们对 A 取某个特定值的信念程度越高，我们对 B 取值的看法就会变得非常模糊不清。

从这个角度看，不确定原理就会变得非常容易理解。由于我们人类的观察、信息处理能力非常有限，所以，在某一个特定时刻，我们掌握了 A 属性的性质，也就必然会伴随着对 B 属性的不掌握。例如，我们看到了月球的正面，就必然看不到月球的反面。因此，正反面就构成了一对不兼容的属性，它们会呈现出不确定的关系。

也许你会争论说，不对，虽然我不能同时看到正反两面，但是当我看正面的时候，反面也必然同时在那里啊，因为它是一个客观存在的事物啊！问题就出在这里，的确，如果坚持事物的客观实在性，我们便发现不兼容属性对似乎只能在量子那里找到答案。

但如果我们反过来看，认为世界的本原就不是唯物的，当我们看不到月亮背面的时候它就什么都不是，那么我们会发现，几乎任何两个我们不能同时观察的属性或者事物都会呈现出兼容的性质。

正如文章《[系统中的观察者](#)》提到的，笔者正在试图将观察者作为第一性的事物出发，重新构筑公理体系。在物质世界中，量子系统中的不确定原理和不兼容属性对是隶属于物质世界的一种蹩脚的性质，但也许它恰恰构成了精神世界中，万事万物的起点：观察者的一种第一原理（First principle）。换句话说，本来所有的属性都是不兼容的，都会呈现出测不准性质，而所谓的兼容恰恰是一种特例。其原因是，人类观察的性质就是在任意一个时间点，只能处理有限的信息。所以，在任意一个时刻，只能有一个我存在，在每个时刻，我只能打下一个汉字，你也只能阅读一个单词，在任何一场球赛中只能有一个冠军。因此不确定关系： $S(A)+S(B)>const$  就是一个必然的要求。

另外，在观察者体系中一个十分重要的问题就是如何从多个不同的观察者眼中看到一个



完全相同的客观世界。在[《系统中的观察者》](#)一文中，我们把这个问题归结为一个通讯问题，也就是说只要多个观察者可以找到一种公共媒介，那么从他们的沟通之中就能构造出来一个共同的“客观世界”出来。

也许你已经看到了，其实这个问题和我们前面讲到的纠缠态存在着深刻的联系。纠缠态解决的恰恰就是不同的观察者对两个处于相互纠缠的系统进行观察并保持通信的能力。换句话说，如果我们能够真正构造出两个相互纠缠的虚拟世界呈现给不同的观察者，那么他们将会“误认为”他们看到的是一个客观世界！如果相互纠缠系统牵扯的属性全部是兼容的属性对，那么这种“客观世界”的伪装会不成问题。但是，一旦把不兼容属性对融入其中，就会发现致命的漏洞：对贝尔不等式的破坏！所以纠缠和不兼容属性对放在一起才是量子概率体系中最宝贵的资源。

冯诺依曼曾说，概率论中存在着一个漏洞（loophole），生命正是利用了这个漏洞而完成了自我复制。现在，我非常怀疑这个漏洞跟贝尔不等式有着深刻的联系。

## 附录 1、用复数概率计算数学期望

在经典概率中，一种非常重要的运算操作就是求某一个函数的数学期望，假设  $x$  是一个随机变量，它的分布由  $F(x)$  刻画， $f(x)$  是任意一个函数，那么随机变量  $f(x)$  的数学期望是：

$$E(f(X)) = \sum_{i=1}^{i=n} f(x_i) p(x_i) = U(X) \cdot F(X)$$

其中后一个表达式采用向量内积的方式书写， $U(x) = (f(x_1), f(x_2), \dots, f(x_n))$ 。同样的道理，我们也可以为复数概率定义数学期望。假设一个随机变量  $x$ ，它可以用复数概率向量  $\Psi(x)$  来描述， $f(x)$  是任意一个函数，那么  $f(x)$  的数学期望就是：

$$E(f(X)) = \sum_{i=1}^{i=n} f(x_i) |\psi(x_i)|^2$$

当然，我们也可以把它写成两个向量内积的形式：

$$E(f(X)) = U(X) \cdot (|\psi_1|^2, |\psi_2|^2, \dots, |\psi_n|^2)$$

然而，我们知道

$$|\psi_i|^2 = \psi_i \psi_i^*$$

所以数学期望的后半部分又可写作一对共轭复数向量内积的形式。

$$E(f(X)) = U(X) \cdot (\Psi(X) \cdot \Psi^*(X))$$

这让我们联想到了线性代数中的二次型，也就是说，我们完全可以把上式改写成：

$$E(f(X)) = \Psi \cdot A \cdot \Psi^{*T}$$

其中

$$\Psi^{*T} = (\psi_1^*, \psi_2^*, \dots, \psi_n^*)^T = \begin{pmatrix} \psi_1^* \\ \psi_2^* \\ \vdots \\ \psi_n^* \end{pmatrix}$$

即原始向量 $\Psi$ 的专置共轭矩阵。

$$A = \begin{pmatrix} f(x_1) & 0 & \dots & 0 \\ 0 & f(x_2) & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & f(x_n) \end{pmatrix}$$

即  $A$  是一个对角阵，对角线元素分别为  $f(x_1), f(x_2), \dots, f(x_n)$ ，可以验证采用矩阵和向量相乘的方法同样可以计算出随机变量  $f(x)$  的平均值。

## 附录 2 投影矩阵

仍然以图 2 为例，我们看到，在几何上，发生  $|L\rangle$  事件的概率就相当于是原始向量  $\psi_A$  在  $|L\rangle$  这个向量上的投影值  $x'$  的模的平方  $|x'|^2$ 。对于任何一个基向量  $|V\rangle$ ：

$$|V\rangle = a|U\rangle + b|D\rangle$$

我们都可以构造出一个称之为**投影矩阵**的矩阵：

$$|V\rangle\langle V| = \begin{pmatrix} a^*a & ab^* \\ a^*b & b^*b \end{pmatrix}$$

其中  $|V\rangle\langle V|$  是量子力学中的记法，表示从向量  $|V\rangle$  构造的投影矩阵。这样该矩阵与原始向量的乘积就是原向量在  $|V\rangle$  上面的投影向量：

$$M = |V\rangle\langle V| \cdot \psi_A = \begin{pmatrix} a^*a & ab^* \\ a^*b & b^*b \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a^*ax + ab^*y \\ a^*bx + b^*by \end{pmatrix}$$

该向量的模平方就是发生  $|V\rangle$  事件的概率。大家可以将  $a = \cos\theta$  ,  $b = \sin\theta$  代入进行验证。

因此，我们看到，只要我们知道了投影矩阵，我们便能计算投影矩阵对应的事件的概率。因此，在复数概率中，我们可以用投影矩阵来表示某一个确定的事件。

## 附录 3 投影矩阵的复合

我们知道，每一个事件都对应了一个投影矩阵，那么  $|L\rangle$  这个事件的投影矩阵（在  $|U\rangle, |D\rangle$  坐标下）是：

$$|L\rangle\langle L| = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix}$$

那么两只萤火虫都在左侧这个事件的投影矩阵（在四维的基向量组  $|UU\rangle, |UD\rangle, |DU\rangle, |DD\rangle$  下表示）就是：

$$\begin{aligned}
|L\rangle\langle L| \otimes |L\rangle\langle L| &= \begin{pmatrix} \cos^2 \theta |L\rangle\langle L| & \sin \theta \cos \theta |L\rangle\langle L| \\ \sin \theta \cos \theta |L\rangle\langle L| & \sin^2 \theta |L\rangle\langle L| \end{pmatrix} \\
&= \begin{pmatrix} \cos^2 \theta \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} & \sin \theta \cos \theta \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} \\ \sin \theta \cos \theta \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} & \sin^2 \theta \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} \end{pmatrix} \\
&= \begin{pmatrix} \cos^4 \theta & \sin \theta \cos^3 \theta & \sin \theta \cos^3 \theta & \sin^2 \theta \cos^2 \theta \\ \sin \theta \cos^3 \theta & \sin^2 \theta \cos^2 \theta & \sin^2 \theta \cos^2 \theta & \sin^3 \theta \cos \theta \\ \sin \theta \cos^3 \theta & \sin^2 \theta \cos^2 \theta & \sin^2 \theta \cos^2 \theta & \sin^3 \theta \cos \theta \\ \sin^2 \theta \cos^2 \theta & \sin^3 \theta \cos \theta & \sin^3 \theta \cos \theta & \sin^4 \theta \end{pmatrix}
\end{aligned}$$

运算法则是先将右侧的矩阵分块放到左侧矩阵中的每一个元素上，然后再用左侧矩阵的每一个元素相乘，最后合并成一个 4\*4 的矩阵。那么状态  $\psi \otimes \psi$  在这个新矩阵上的投影就是：

$$\begin{aligned}
|L\rangle\langle L| \otimes |L\rangle\langle L| \cdot (\psi \otimes \psi) &= \begin{pmatrix} \cos^4 \theta & \sin \theta \cos^3 \theta & \sin \theta \cos^3 \theta & \sin^2 \theta \cos^2 \theta \\ \sin \theta \cos^3 \theta & \sin^2 \theta \cos^2 \theta & \sin^2 \theta \cos^2 \theta & \sin^3 \theta \cos \theta \\ \sin \theta \cos^3 \theta & \sin^2 \theta \cos^2 \theta & \sin^2 \theta \cos^2 \theta & \sin^3 \theta \cos \theta \\ \sin^2 \theta \cos^2 \theta & \sin^3 \theta \cos \theta & \sin^3 \theta \cos \theta & \sin^4 \theta \end{pmatrix} \begin{pmatrix} \alpha^2 \\ \alpha\beta \\ \alpha\beta \\ \beta^2 \end{pmatrix} \\
&= \begin{pmatrix} \alpha^2 \cos^4 \theta + 2\alpha\beta \sin \theta \cos^3 \theta + \beta^2 \sin^2 \theta \cos^2 \theta \\ \alpha^2 \sin \theta \cos^3 \theta + 2\alpha\beta \sin^2 \theta \cos^2 \theta + \beta^2 \sin^3 \theta \cos \theta \\ \alpha^2 \sin \theta \cos^3 \theta + 2\alpha\beta \sin^2 \theta \cos^2 \theta + \beta^2 \sin^3 \theta \cos \theta \\ \alpha^2 \sin^2 \theta \cos^2 \theta + 2\alpha\beta \sin^3 \theta \cos \theta + \beta^2 \sin^4 \theta \end{pmatrix} = \begin{pmatrix} \cos^2 \theta (\alpha \cos \theta + \beta \sin \theta)^2 \\ \sin \theta \cos \theta (\alpha \cos \theta + \beta \sin \theta)^2 \\ \sin \theta \cos \theta (\alpha \cos \theta + \beta \sin \theta)^2 \\ \sin^2 \theta (\alpha \cos \theta + \beta \sin \theta)^2 \end{pmatrix}
\end{aligned}$$

因此，这个向量的模平方就是发生 LL 事件的概率。我们看到这个计算结果与\*\*的计算结果是一致的。所以采用矩阵的方法完全可以完成同样的运算。当我们将两个系统做成复合系统的时候，所有的状态向量、矩阵也都可以进行同样的扩充。

## 附录 4 利用投影矩阵计算两盒子中萤火虫处于各种状态的概率

萤火虫处于左侧这个事件对应的投影矩阵（参见附录 2）就是：

$$|L\rangle\langle L| = \begin{pmatrix} \cos^2 \pi/3 & \sin \pi/3 \cos \pi/3 \\ \sin \pi/3 \cos \pi/3 & \sin^2 \pi/3 \end{pmatrix} = \begin{pmatrix} 1/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 3/4 \end{pmatrix}$$

当两个系统复合成一个更大的系统之后，我们可以在复合系统中执行下面的测量，仅观察左侧系统中的萤火虫处于左还是右，而无论右侧萤火虫的状态如何，这相当于执行如下投影矩阵所表示的测量：

$$|L\rangle\langle L| \otimes I = \begin{pmatrix} 1/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 3/4 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/4 & 0 & \sqrt{3}/4 & 0 \\ 0 & 1/4 & 0 & \sqrt{3}/4 \\ \sqrt{3}/4 & 0 & 3/4 & 0 \\ 0 & \sqrt{3}/4 & 0 & 3/4 \end{pmatrix}$$

通过把这个矩阵和纠缠态向量  $\xi = \frac{1}{\sqrt{2}}|UU\rangle + \frac{1}{\sqrt{2}}|DD\rangle$  相乘，再计算模的平方能计算出

测量得到左侧萤火虫处于左室的概率是  $1/2$ 。

类似地，我们还可以计算出左盒的萤火虫处于上（下）侧、右盒的萤火虫处于上（下）侧，左盒的萤火虫处于左（右）侧，右盒的萤火虫处于左（右）侧的概率都是  $1/2$ 。

下面，我们计算观察到左盒的萤火虫在上侧，同时右盒的萤火虫在左侧舱室中的概率。这相当于求状态向量在下列投影矩阵上的投影长度：

$$(|U\rangle\langle U|) \otimes (|L\rangle\langle L|) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 3/4 \end{pmatrix} = \begin{pmatrix} 1/4 & \sqrt{3}/4 & 0 & 0 \\ \sqrt{3}/4 & 3/4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

从而得到投影的模平方即经典概率是  $1/8$ 。也就是说左盒的萤火虫在上，同时右盒的萤火虫在左的概率就是  $1/8$ 。