量子力学中的测量

（武家熙，中国科学技术大学，PB22000092）

【摘要】本文对量子力学的理论框架进行了梳理，并重新审视了量子力学中的核心概念——测量，特别对其动机进行了分析。本文首先回顾了经典物理中的测量概念，然后回顾了量子理论发展中产生的各个阶段性解释，最后讨论了重新审视测量对理论发展的意义。

【关键词】量子力学；近代物理；测量

1 经典量子力学的问题

量子力学是近代物理学最重要的成就之一，其在解释微观物理世界方面的成就奠定了近代物理几乎所有方面的基石。但是，正如量子力学的奠基人之一 Bohr 所言，“如果你第一次学量子力学时认为自己懂了，那么你事实上完全没有理解。”这不是物理专业的学生水平不足所致的，而是量子力学的固有属性：事实上，如果一个学生在初次学习量子力学（或者原子物理、量子物理等课程）时号称自己完全理解了，这一般不仅代表其没有理解量子力学，通常也意味着其对经典力学的理解具有一定的问题。

近代经典量子力学主要面临两方面的指责：其一，量子力学中的数学结构对于数学家来说并不清晰且严谨，且相较于其他物理理论而言，量子力学完全放弃了微分几何的语言，这令数学家和熟悉几何物理理论（如引力理论、经典力学和经典场论等）的物理学家感到不安；其二，量子力学中的测量的意义完全不同于经典物理，测量影响系统演化，使得系统遵循两套独立的演化规则。这逼迫人们重新审视与思考那些曾经当作习惯的物理直觉，特别是有关概率方面的，并思考其出现的原因。

回答第一个问题是数学家的任务，而需要指出的是，尽管物理学家往往并不习惯在量子力学中使用严谨的数学，但事实上在使用泛函分析的语言并将态空间扩大之后，量子力学中的绝大多数数学问题都能得到很好的解决。另一方面，物理学家从不拒绝逼迫数学家创造新的数学，如经典场论使得傅里叶分析和特殊函数论有了重要的发展，这些工具时至今日已经被当作严谨的数学看待。

而对于第二个问题，考虑到量子力学在实验上的巨大成就，没有人可以宣称这个理论是没有意义的，但是其相当令人不安。有的物理学家选择“Shut up and calculate.”，完全接受 Copenhagen 学派的解释，我们称之为“经典量子力学”。但是这种态度是令人遗憾的，因为其遗忘了物理的初衷是更好地描述世界并尝试理解，而当一个理论声称自己拥有从完备的第一性原理到实验结果的框架，但是需要庞大的补丁来修正，这只能说明这个理论在很大程度上是有待重新诠释的。另一些物理学家尝试给出自己的解释，例如 Bohr 的测量解释，Schrodinger 的猫态，von Neumann 的密度矩阵，直到现代的物理学家 Laloe 提醒人们重新审视对量子力学的解释，而这些对量子力学的解释尝试使得人们可能更加接近于真正的理解。

事实上，我们将量子力学与现代物理的另一个重要成就广义相对论（或者更严谨地说，Einstein 引力理论，因为数量可观的物理学家不承认其满足相对性原理）相比，量子力学的发展是令人惊奇的。广义相对论自建立起，Einstein 的基本原理就受到了各种意义上的攻击。事实上，Einstein 本人也屡次更改其基本原理，我们今天对其理解主要来自后世的解读。更重要的是，不乏与广义相对论同样有说服力的引力理论，如 Newton-Carton 引力、Lovelock 引力等理论，其甚至也使用了相似的数学语言，即微分几何；而这在很大程度上来自于人们对 Einstein 基本原理的理解进行了长久的讨论，并逐渐明确其地位。时至今日，修改引力仍然是一个很热门的物理领域。而对于量子力学，几乎没有人能提出替代的理论；量子力学确实发展出了其上层的理论，比如量子场论、量子色动力学、甚至于弦论，其中的一些理论，如格点量子色动力学，也在尝试解释一些量子理论建立过程中遗留的问题，但是援引杨振宁先生的观点，或许这些上层理论可以解释更多具体的问题，但是对于那些根本的问题，“它们不能比量子力学告诉我们更多。”这在很大程度上是源于至今仍然没有人可以号称自己理解了量子力学的基本原理，但是我们可以尝试更加理解，这对我们理解其上层理论也是有好处的。

我们将回顾近代的经典量子力学，重新审视

经典理论关于测量等问题的启发，进而引入量子力学的核心问题：量子意义下的测量与经典有何异同；然后我们将介绍近代和现代对量子理论做出重新解释的那些“非经典量子理论”。我们相信，这对理解量子理论是有益的。

2 数学基础

为了后续概念的引入，我们先来简要地介绍一些代数概念，而省略一些存在性的证明。这些内容不是必要的。对于数学系的学生而言，这些内容过于简略而缺乏严谨性；但是对于物理系的学生而言，这些内容或许有助于其理解后文中讨论物理问题时使用的数学工具。

2.1 线性空间

我们采用近世代数中标准的群（group）、环（ring）、域（field）、模（module）、映射的核（kernel）像（image）、同态（homomorphism）和同构（isomorphism）等定义与记法，特别地，记实数域为$\mathcal{R}$，复数域为$\mathcal{C}$。

域$\mathcal{F}$上的线性空间（linear space）V是一个集合上的模，即一个集合配上两个操作：加法与数乘：

$$+: V\times V\to V$$
$$\cdot: \mathcal{F}\times V\to V$$

其在加法下构成一个 Abel 群，对数乘满足结合律，对加法与数乘满足分配律（详细的定义见[1]）。线性空间中的元素被称为矢量（注意此处的矢量并非指其分量数组）。由于线性性，对与有限维的线性空间，存在一组完备的基矢量，其个数是唯一的，称为线性空间的维数，使得任何线性空间中的矢量都可以唯一展开为基的线性组合（采用 Einstein 求和约定，对同一式中出现的一组以相同字母表示的上下标对线性空间维数求和）：

$$x = x^i e_i, \forall x\in V$$

可以证明，同一个数域上的所有维数相同的线性空间彼此同构。我们一般不区分“同构”与“相等”。

下面简要地介绍代数系统（algebra system）的概念[1]，目的是在后面引入 Hilbert 空间系统。借用范畴论的定义，特别地，在集合范畴中，集合的直积（direct product）即卡氏积（Cartesian product），集合的直和（direct sum）即无交并（disjoint union）。

2.2 对偶空间

线性映射（linear mapping）f是同一个域上线性空间间的映射，满足线性性：

$$f: X\to Y$$
$$f(\alpha x+\beta y) = \alpha f(x) + \beta f(y)$$
$$\forall x,y\in X, \forall \alpha,\beta\in\mathcal{F}$$

两个线性空间间的所有线性映射构成一个线性空间，记为$\mathcal{L}^{(1)}(X;Y)$。

注意到域在其本身下形成一个线性空间，则称线性空间到其域的线性映射为线性函数，其集合$\mathcal{L}^{(1)}(X;\mathcal{F})$为$X$的对偶空间（dual space），记为$X^*$，其中的元素记为$x^*$，矢量的展开为：

$$x^* = x_i e^{*i}, \forall x^*\in X^*$$

其中对偶基是一组函数，其对原空间的基作用结果为 Kronecker 函数：

$$e^{*i}(e_j) = \delta^i_j$$

对于任意的线性空间，将作用视为一个二元运算，有：

$$X^{**}\cong X$$

由于我们不区分彼此同构的对象，记为：

$$X^{**} = X$$

在有限维的情况下，存在一个从线性空间到其对偶空间的同构（考虑基的作用，这是显然的）。

2.3 张量积空间。

张量积对于建立量子力学中的 Hilbert 空间是重要的，特别是多自由度的情形。一般的量子力学著作或课程中不会涉及到引入张量积的动机，甚至有些物理学家会认为张量积就是线性代数中的直和，我们在此给出一个动机明确的定义。

前面我们引入了线性映射的概念，当涉及到多个线性空间$X_1, X_2, ..., X_n$上的映射时，可以类似地定义多重线性映射（multilinear mapping）：

$$f: X_1\times X_2\times ......\times X_n\to Z$$

$$\begin{aligned}&f(x_1,x_2,...,ax_i+by_i,...,x_n)\\&= af(x_1,...,x_i,...,x_n)\\&+bf(x_1,...,y_i,...,x_n), \forall i\end{aligned}$$

即其对每一个位置上的线性空间都是线性映射。但是，考虑上式左边的 Cartesian 空间作为一个线性空间，其线性性应表现为：

$$\lambda(x_1,x_2,...,x_n) = (\lambda x_1, \lambda x_2, ..., \lambda x_n)$$

但是一个多重线性映射：

$$f(\lambda x_1,\lambda x_2,...,\lambda x_n) = \lambda^n f(x_1,x_2,...,x_n)$$

在$\lambda\neq 1$时不表现线性性，即多重线性映射不是线性映射，而线性映射具有我们希望的性质，

为此引入张量积空间：

$$X_1 \otimes X_2 \otimes ... \otimes X_n$$

其使得 Cartesian 空间到其的映射为多重线性映射，而其到目标空间为线性映射。可以证明，线性空间的张量积空间是存在的，且在同构意义下是唯一的[2]：

$$X_1 \otimes ... \otimes X_n \cong \mathcal{L}^{(n)}(X_1^*, ..., X_n^*; \mathcal{K})$$

$$X_1^* \otimes ... \otimes X_n^* \cong \mathcal{L}^{(n)}(X_1, ..., X_n; \mathcal{K})$$

记其中元素为：

$$x_1 \otimes ... \otimes x_n \in X_1 \otimes ... \otimes X_n, x_i \in X_i$$
$$x_1^* \otimes ... \otimes x_n^* \in X_1^* \otimes ... \otimes X_n^*, x_i^* \in X_i^*$$

张量积满足结合律。

张量积可以交换，但是需要同时交换作用顺序使得每一个函数的作用对象不变。

2.4 张量代数

考虑同一个线性空间及其对偶空间的张量积，记

$$X_s^r = \bigotimes_{k=1}^{r} X \otimes \bigotimes_{q=1}^{s} X^*$$

在线性空间的直和下，称

$$T(X) = \bigoplus_{r=0}^{\infty} X_0^r$$

为协变张量代数，因为其在张量积下形成一个结合代数；称

$$T^*(X) = \bigoplus_{s=0}^{\infty} X_s^0$$

为协变张量代数，其也是一个结合代数。更一般的张量代数是：

$$\mathcal{T}(X) = \bigoplus_{r=0,s=0}^{\infty} X_s^r$$

这是张量代数的标准形式。

2.5 泛函分析基础

一个任意的集合A上的度量是一个满足一定条件的映射：

$$\mathrm{d} : A \times A \to \mathcal{R}^+$$

$$1.\mathrm{d}(u,v) \leq \mathrm{d}(u,w) + \mathrm{d}(v,w)$$
$$2.\mathrm{d}(u,v) = \mathrm{d}(v,u)$$
$$3.\mathrm{d}(u,v) = 0 \ iff \ u = v$$

线性空间X的内积是一个满足一定条件的映射：

$$g : X \times X \to \mathcal{K}$$

$$1.g(u,u) \in \mathcal{R}^+,\ g(u,u) = 0 \ iff \ u = 0$$
$$2.g(u,v) = \overline{g(v,u)}$$
$$3.g(u, \lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 g(u, v_1) + \lambda_2 g(u, v_2)$$
$$\forall \lambda_1, \lambda_2 \in \mathcal{K}$$

线性空间X的范数是一个满足一定条件的映射：

$$||\cdot|| : X \to \mathcal{R}^+$$

$$1.||u+v|| \leq ||u|| + ||v||$$
$$2.||\lambda u|| = |\lambda|\,||u||, \forall \lambda \in \mathcal{R}$$
$$3.||u|| = 0 \ iff \ u = 0$$

由此可见，一个线性空间的内积定义了一个范数，一个范数定义了一个度量。若一个空间的 Cauchy 列均收敛于其自身内，则称其为完备的（complete）。一个完备的赋范空间被称为 Banach 空间，而一个完备的内积空间被称为 Hilbert 空间。

3 经典物理的启示

对量子力学的认知往往重点在于对态矢量和算符的代数运算和对波函数的分析。事实上，如同经典力学的重点是建立描述物理系统的数学空间，我们认为对量子体系的认识是一个量子理论的核心，而非进行计算。对经典物理的回顾有益于给我们关于建立量子理论的启示。

3.1 经典力学的启示

经典力学描述宏观低速物体的运动。

经典力学最基本的原理，往往也是最容易被忽视的，被称为 Newton-Laplace 决定性原理（determinacy principle），表述为一个系统在给定时刻的状态完全决定了它在所有时刻 t（过去和将来）的运动。这在 Newton 力学和经典的分析力学中都是适用的，且对我们研究的系统做出了一定的约束（例如，其阻止了我们使用分析力学描述 Norton 穹顶之类的问题）。

Lagrange 和 Hamilton 的经典力学是一切现代理论物理学的基础，其相较于 Newton 力学的进步远不止是数学形式上的简洁，而且具有更经得起考究的数学结构，也更富有物理启发性。众所周知，Heisenberg 和 Schrodinger 等人建立的量子理论受启发于 Hamilton 力学，而 Feynman 的路径积分理论受启发于 Lagrange 力学。

一个问题在于，理论上更简洁、与量子力学

对应更明显的 Hamilton 力学对应于更复杂的几何结构，即辛几何（sympletic）；而几何结构更简洁的 Lagrange 力学缺乏简洁的几何与物理学间的对应。

由于我们关心的是经典力学中对建立量子理论有启发的部分，我们不对经典力学的具体细节做出详细叙述，被省略的细节部分见参考文献[3][4]。这种省略是非常令人不安心的，但是其有助于凸显我们关注的重点。

Lagrange 力学中，系统状态的描述是由位形空间（configuration space）承担的。考虑由N个三维质点组成的经典力学体系，其间存在$3N-n$个完整约束，则系统位置是由一个$\mathcal{R}^{3N}$空间的n维子流形描述的，这个子流形被称为位形空间M。任何系统可能的状态均唯一地由其中的点来描述。由于我们考虑的位形空间是光滑流形，我们可以考虑其切丛TM作为一个$2n$维的流形，其上可以被赋以光滑结构，进而有向量场。我们熟知的 Euler-Lagrange 方程：

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}-\frac{\partial L}{\partial q}=0$$

$$L:TM\to\mathcal{R}, L=L(t;\vec{q},\dot{\vec{q}})$$

实际上描述了位形空间切丛上的一个向量场。

我们记 Lagrange 力学为(M,L)，以提醒我们一套经典力学理论应包含两部分，其一是如何描述状态，其二是如何给出运动学方程。

Hamilton 力学和 Lagrange 力学是等价的，其可以被视作 Legendre 变换的结果。略去数学上的细节，一个流形M的 Legendre 变换是从其切丛到余切丛的一个映射：

$$\tau_L:TM\to T^*M$$

其满足一些条件[5]。如果这是一个微分同胚，则其给出 Hamiltonian 及其满足的运动方程：

$$\dot{\vec{p}}=-\frac{\partial H}{\partial q},\dot{\vec{q}}=\frac{\partial H}{\partial p}$$

$$H:T^*M\to\mathcal{R}, H=H(t;q,p)$$

被称为 Hamilton 正则方程，描述了一个流形余切丛上的向量场。我们记作(M,H)。

考虑余切丛作为一个偶数维的流形，Hamilton 正则方程（准确地说是相流）限制了标准辛形式（canonical sympletic form），进而使其成为一个辛流形。记其上的函数场为$\mathcal{F}(T^*M)$

辛形式定义了标准 Poisson 括号，这是一个二元的内部操作：

$$[-,-]_C:\mathcal{F}(T^*M)\times\mathcal{F}(T^*M)\to\mathcal{F}(T^*M)$$

其满足斜对称、Leibniz 律和 Jacobi 恒等式。由于经典的函数场是交换的，其在 Poisson 括号下构成一个 Poisson 代数。这是启发正则量子化的经典内涵。

事实上，我们可以用前文提到的张量代数形式写出该理论的更严谨的版本，比如用阶化算子、内乘算子等工具构建描述运动的向量场。

在这个过程中，有三件事是富有启发性的。

其一，一套力学理论应当包括两个部分：第一，如何描述系统状态；第二，如何描述力学的结果，即如何给出运动。我们可观测的结果总是以数值的方向呈现，而我们的理论中必须包含其背后的描述；与数学理论不同的是，我们需要足够有说服力的结论来支持我们的观点，即其应给出可测的运动规律。举例来说，对于一个引力理论，第一部分被称为时空观，而第二部分被称为场方程；特别地对于广义相对论，第一部分为一个伪 Riemann 流形，而第二部分为 Einstein 场方程。我们会在后面反复提到有关理论的两个部分，且我们认为一个量子理论也应当给出类似的东西。

其二，系统状态由一个微分流形上的点描述，而广义的可观测量（包括 Lagrangian 等在经典意义下没有物理意义的量）被视为一个由位形空间的切丛到实数的映射，即不同的状态对应着物理量的不同值。将物理量视为一种映射而非一个数学意义下的数，这是我们在经典力学中就在做但从未被注意到的，而这对我们转向量子理论是极其重要的。

其三，关于我们如何看待时间与演化：首先可以明确的是，在非相对论情形下，时间不作为与其他广义坐标平权的一维，我们考虑的流形总是偶数维的。其次，由于切丛投影算符的存在，我们总是可以考虑位形空间的切矢量场；而完备的光滑切矢量场总是唯一对应于一簇积分曲线[6]，进而诱导一个微分同胚变换（略去关于完备性的条件）。时间作为积分曲线的参数参与变换的过程。注意到此处存在两种视角：主动视角和被动视角。主动视角中，时间演化相当于一种单参数微分同胚变换，描述状态的空间本身发生变化。被动视角中，微分同胚变换不是作用于状态空间本身，而是诱导了一种坐标变换，其使得其上可观察量的表示（representation）发生了变化。这两种视角的物理结果是相同的。对这两种视角的理解启发了我们对量子理论中不同图像的理

解：Schrodinger 图像对应于主动视角，而 Heisenberg 图像对应于被动视角。图像的选择不应影响量子理论中的任何物理，不存在任何可被观察的物理对象可以使得我们推断哪一种图像是物理的而哪一种不是，其在物理上是等价的。需要注意的是，这也能向我们说明为什么时间在经典量子力学中不是一个可观测量，比如不确定性关系对于可观测量由可观测量在某态下的不确定度定义，而在时间和能量情形下需要借助 Ehrenfest 定理给出定义。

3.2 经典热力学的启示

经典平衡态热力学描述微观系统的宏观行为。其统计思想对量子理论的建立极具启发性，也将是我们建立量子理论的基础。

简单地说，热力学研究的是包含大量微观粒子的系统在宏观上可以被观测到的行为。

我们沿用上面的思路来分析热力学系统[7]：类似地，一个热力学系统被相空间中的一个点描述。与力学系统不同的是，对于一个包含n个粒子的热学系统，我们不会使用一个$6n$维的流形描述其状态，因为热力学关注的主体，即作为理论的第二部分，热力学的观测结果是那些宏观可测的东西，而热力学中单个粒子的运动是宏观不可测的。所以，我们选择使用系统的广延量（extensive）作为流形的标准坐标。

一般的热学书会声明广延量指正比于系统摩尔数一次方的量，而相应地，强度量（intensive）指正比于系统摩尔数零次方的量，我们在此不选择这种声明作为理论的基础，而是将其作为一种推论。

我们定义系统的广延量指那些在满足相同平衡条件下多个系统的合并中数值为各个系统之和的量，其包括内能U和各组分的摩尔数N等，在流体系统（如气体）中还包括体积V、在磁介质系统中还包括磁化矢量的大小m等。这构成了热力学系统的状态表示，即我们在经典力学中定义的“物理理论的第一部分”。

定义中的“满足相同条件”的描述是一件困难的事，严格地说，在数学上是一种循环定义，但是在物理上的逻辑并无问题。我们称那些所有强度量都相同的系统构成一个“同状态系统集合”，这是一个想象的集合，这个集合中不同的点对应了广延量拥有不同值的系统。回顾 2.1 中的定义，广延量的可加性对应着这个集合的一个线性结构，即固定强度量的系统集合是一个$\mathcal{R}$上的一维线性空间。考虑到严格的平衡态热力学下的系统合并应当在两系统达到平衡下进行，这是不难理解的。另一方面，这意味着对于一个均匀的平衡态系统，我们可以自由地将其划分成任意的几个子系统的并，并分别讨论他们的平衡；这将成为我们讨论平衡判据和稳定判据的基础。

通过以上论述可知，考量系统性质时，广延量不起作用；即，当所有强度量确定时，所有广延量均呈正比关系。注意此时我们考虑的是一种假象的视角，其并不是关于某一个给定系统的状态变化，而是关于所有可能的平衡系统的集合；上述讨论同样告诉我们，一个平衡态系统是不可能在不改变任何强度量的情况改变广延量的。

我们讨论以上讨论中平衡概念合法性，即强度量的存在性。这由推广的热力学第零定律保证：当第一个体系与第二个体系的平衡、第二个体系与第三个体系的平衡分别达到时，第一个体系和第三个体系也达到平衡。根据我们前面的讨论，各种平衡对应于各种强度量的相等，这给了我们强度量的物理对应：平衡对应一种物理上的等价关系，所有物理系统按这种等价关系分为不同的等价类，标记各种等价类的参量即为相应的强度量。例如热学平衡对应着一种名为温度的等价关系，力学平衡对应着一种名为压强的等价关系，化学平衡对应着一种名为化学势的等价关系。这给了我们给予强度量物理意义的依据，其标记了上述等价关系对应的等价类。

另一方面，我们指出各强度量的平衡是相互独立的。以两个参量的系统集合为例，我们考虑这样的一个系统集合和A、B两个平衡标准。首先我们将集合中所有的系统用A标准判断平衡（平衡中假设只有A类接触而隔绝B类，这是假象的平衡判断而不改变系统状态），这给出了一种等价关系，我们划分出其等价类，我们记每个等价类为X_A；其次，我们在每个等价类中（注意每个等价类对于B标准仍然构成不同状态系统的集合），我们对其进行B类判断，将在A的等价类元中给出B等价关系，进而给出等价类(X_A, X_B)，每个(X_A, X_B)对应一组此二标准都平衡的系统集合。只考虑这两种平衡时，(X_A, X_B)即我们前面讨论的“满足相同条件的系统集合”，即广延量的可加性意味着(X_A, X_B)上的线性结构。我们改变上述判断平衡的顺序，即先判断B类平衡，再判断A类平衡，给出等价类(Y_B, Y_A)。我们说平衡是独立的，就是说：

$$(X_A, X_B) = (Y_B, Y_A)$$

即对各平衡标准的判别是独立的，具体地说，

热平衡、力学平衡、化学平衡等都是独立的。这就是说，态流形不是一个一维的流形。

以上讨论给出了热力学系统平衡态的定义，但是我们尚未知晓如何描述一个给定的热力学系统状态的变化，我们接下来关注如何描述一个给定的热力学系统。根据以上讨论，对于给定所有强度量的系统集合，其中的一个系统不可能通过仍在这个系统集合中的连续变化变成另一个。这启发了我们：反过来考虑，是否任何系统都唯一对应着一个给定强度量的系统集合中的一点呢？不考虑相变的情况下，均匀系统的答案是肯定的。也就是说，我们可以使用广延量来标记一个系统，而强度量则标记了其上的场。这回到了我们最初的定义，也说明了这是恰当的。

我们设系统可以达到n种平衡，对应n个强度量，那么我们说这个热力学系统被一个n维流形M描述。考虑其上的平凡坐标域的映射φ，记$\varphi(M)\subset\mathcal{R}^n$的各坐标分量为各广延量，自然地，其余切丛为$T^*M$。则热力学平衡条件为：存在一个态空间上的函数熵，平衡时其取极值：

$$\exists S: M\to\mathcal{R}, \Delta S\equiv 0$$

回顾经典力学理论的第二部分运动学，Lagrange 方程描述了切空间的向量场，与 Hamilton 方程描述了余切空间的向量场。熵取极值事实上是热力学态空间余切丛空间的一个微分方程：在局域上，这要求：

$$dS=0$$

其描述了一个微分 1-形式场：以双组分流体体系为例：

$$S=S(U,V,N_1,N_2)$$

$$dS=(\frac{\partial S}{\partial U})_{V,N_1,N_2}dU+(\frac{\partial S}{\partial V})_{U,N_1,N_2}dV$$
$$+(\frac{\partial S}{\partial N_1})_{U,V,N_2}dN_1+(\frac{\partial S}{\partial N_2})_{U,V,N_2}dN_2$$

注意到偏导数的下标不是必须的，其在热力学教材中的普遍存在是因为人们在表象变换中总是不记得将熵看作函数；如总是记得指明每一处的熵是哪一个微分流形上的函数，则偏导操作的意义由外微分算子严格定义，也就无需指明下标了。

上式是一个微分 1-形式的方程，而每个广延量外微分系数是微分 0-形式即流形上的标量场，按照这种方式获得的标量场被称为强度量：

$$\frac{1}{T}=\frac{\partial S}{\partial U}$$

$$\frac{p}{T}=\frac{\partial S}{\partial V}$$

$$-\frac{\mu_i}{T}=\frac{\partial S}{\partial N_i}$$

注意此时左侧整体被称为强度量，其形式的诡异是由于我们希望那些人们习惯于测量的量拥有简洁的形式。特别注意此时强度量不具有任何物理意义，其只是具有我们给定的形式。

则平衡条件改写为：

$$dS=\frac{1}{T}dU+\frac{P}{T}dV-\frac{\mu_1}{T}dN_1-\frac{\mu_2}{T}dN_2$$

即著名的热力学第一定律。事实上，我们的体系下热力学第一定律的准确表述是：热力学平衡条件确定的强度量场与判断平衡时的强度量相等。因此，我们不区分这两者。

注意到熵是广延量的函数，进一步地，其是一个多重线性函数，则可以对上式积分：

$$S=\frac{U}{T}+\frac{PV}{T}-\frac{\mu_1 N}{T}-\frac{\mu_2 N}{T}$$

这被称为 Euler 关系。对上式取外微分再与热力学第一定律相减，即 Gibbs-Duhem 关系：

$$Ud(\frac{1}{T})+Vd(\frac{P}{T})-N_1d(\frac{\mu_1}{T})-N_2d(\frac{\mu_2}{T})=0$$

其中的每一个强度量外微分都是其作为广延量函数的微分而获得的 1-形式场。

事实上，广延量的选取具有一个任意性：以上以熵作为平衡判据被称为熵表象，而例如可以将熵作为最初的广延量，而内能作为判断平衡的“运动方程”，我们也能发展出一套完全等价的热力学理论，且强度量将具有更为人们所熟知的形式如温度、压强等，这被称为内能表象。但是需要注意的是，熵在非统计的热力学部分令人难以理解的物理学意义，正是这个理论存在的核心：它如同 Lagrangian 一样，表征了广义的“体系的运动”。

热力学的理论框架如上基本建立，其也如同经典力学一样可以进行 Legendre 变换，其中熵表象中变换后的函数（类似 Hamiltonian）被称为热 Massieu 函数，而内能表象在变换的函数后被称为热力学势。

作为一个系统参数是广延量的应用，我们考虑均匀系统的稳定性：对于一个均匀系统，其平衡时意味着熵取极大值。由于系统参数为广延量，可以将系统划分为任意多个子系统之并，使得子系统的所有广延量之和均为原系统的值，且强度量均相等。稳定性条件给出：若系统是稳定的，

改变各子系统的广延量值，但保持和不变，则任意的改变一定造成其总熵下降。这说明系统参量是广延量意味着稳定系统倾向于均匀。

3.3 经典理论的启示

综上所述，通过对经典理论的重写，我们发现事实上已经获得了很多熟知的量子力学假设，有些甚至形式一致。当我们建立量子理论时，我们需要两个部分：第一，我们如何界定并描述一个量子系统。第二，我们如何描述量子系统应当遵守的广义运动方程，其应当是一个微分方程。事实上还有一个关于测量的问题：理论的第二部分应当是可以被观测的，我们应当给出其被测量的方式。

4 经典测量

在正式开始考察量子理论之前，我们首先考虑有关经典意义下的测量的一些概念。其动机来自于两个部分：其一，概率是量子理论的核心内容，而对于概率的理解极度依赖于测量；其二，Copenhagen 学派认为测量使得态发生跃变，这使得测量与非测量情形下的系统演化规律不同，同时带来了奇异的时间反演，这是其与经典理论最大的区别。

4.1 经典力学中的测量

注意到我们目前为止从未讨论过经典力学意义下的测量。事实上，这远比其直观上困难。

我们关注的测量是实际上可以发生的测量，而非抽象意义上的测量，即物理可测而非数学意义上的测量（当然物理可测的必要条件是数学可测）。在数学的意义下，当我们确定了一个系统的状态流形，并指定了其中的一个点，声称其代表了一个粒子的位置，则此时测量已经完成，因为没有任何不确定的东西。但是这在物理上没有任何意义，也无法被用来描述任何物理行为。

通过我们对热学的讨论，我们可以定义：测量行为是状态集合等价类集到实数的映射。由于状态到其等价类的映射是显然的，这在直观上也符合测量的定义。

我们讨论经典力学的测量。对于经典力学，我们援引一段 Einstein 对广义协变性的点一致性（point-coincidence）论述来说明这一点[8]：如果两个点在某一个图的坐标系中坐标分量相等，则其在同一图册中另一图的坐标系中坐标分量也相等。需要注意的是，虽然这段论述是用来说明其广义协变性的合法性，但是其事实上被认为并不能有力地证明 Einstein 对于广义相对论的广义协变性的观点。尽管如此，这对于经典力学是正确的。这意味着两件事：其一，我们使用实数、或与实数轴拓扑同胚的流形（注意不是 Lie 群，因为我们无法声明坐标分量的加法）来标记状态是合法的，而不使用复数或是别的数学结构，也不是一个群。其二，经典力学的坐标具有规范自由度，即对于一个质点，其坐标首先被一个图的坐标标记，同时被另一个与之相容的图标记，若质点始终处于这两个图的交中，则这两个图的坐标可以等价地互相转换。我们在此再次强调：经典物理的状态流形是没有群结构的，更没有线性结构。

因此，当我们在谈论经典力学中的测量时，首先需要指定一个图并确定一个坐标架：在理论框架中，经典力学理论具有规范对称性，因此不必在一个满足相容性条件的图册中指定图；但是谈论测量时，我们需要操作局部坐标，所以需要指定图。其次我们指出最重要的一点：经典力学中的测量是依赖直觉的测量：首先我们假设时空各向同性且均匀，以使得我们可以认为任何时间任何地点都可以进行相同的测量；其次测量本质是一种依赖人直觉的判断：判断现实中的对象对应于哪一个理论物理模型中的状态。判断木棍与刻度是否对其、小球落下时是否刚好掐表、粒子通过加速器是否发生。最后一步是读取数据，即人们通常意识到的测量。但是请注意，可能的随机性并不来自于读取数据，即物理状态到实数的映射；而是来自于最关键的第二步：判断一个物理状态是人们想象中的一个状态，而这一步是无论如何也要依赖人主观的判断的，即依赖于直觉。即使是粒子物理中最先进的实验装置，其测量手段也是由人决定的。

事实上，Einstein 对这点认识得很清楚：人们学习广义相对论时，往往会将注意力集中于矢量的类时类空分解，并惊叹于时空间区分的非良定的神奇；但是人是三维的观测者，人拥有在局域坐标架中观测的能力这一点看上去平平无奇的结论也是值得思考的。举例来说，我们知道有质量粒子的世界线是类时曲线的像。在双生子实验中，我们判断的时间差是各自固有时的差。这意味着一个观测者可以观测到自己的固有时。但是实际实验中我们使用的观测工具一般是原子钟，并认为原子钟的读数是固有时。这就很有趣了：众所周知，简单来说原子钟测量的是可以原子吸收的光的频率，其可以测量时间的本质是原子能级的量子化，是一种典型的量子效应；而广义相

对论的局域极限是狭义相对论，低速极限是Newton力学,而这二者和量子力学都是不兼容的。所以,我们事实上是做了这样的假设：Plank常数趋于0的量子力学经典极限，和光速趋于无穷的狭义相对论经典极限的结果是兼容的，这就是直觉上的物理世界。

更进一步，我们指出对于测量需要依赖直觉并不是一个新颖的概念。在很多理论中，都有着时间定向的概念；在广义相对论中，时间定向是一个流形上的非零类时矢量场；在宇宙学意义上，时间定向意味着宇宙演化的方向；在统计力学中，时间定向意味着孤立系统熵增加的方向。但是，最特别的是在心理学或哲学意义上，人可以感受到时间的方向性：人可以知晓过去，但是只能推测未来。确定性的一侧就是时间定向的反向。事实上，现代统计物理、或是量子统计的基本假设就包括统计力学意义上的时间定向等价于心理学或哲学意义上的时间定向[9]。

以上讨论不是为了展现物理学的测量有多么不可靠，而是为了强调：在物理学实验中，直觉永远不可能被排除在框架之外，人的存在是观测存在结果的必然要求。在经典意义上，我们往往过于依赖直觉而忽视了对直觉的依赖，因此在面对量子假设时会感到不安。事实上，我们从未在经典物理中排除自己的存在。因此，我们也不应对将自己置于量子理论中的特殊位置而感到不适。

4.2 经典物理中的概率

有了上述关于测量中直觉作用的讨论，我们应当更加警惕如何看待观测结果：借用Pauli的话，我们应当警惕区分什么是我们观测到的结果，而什么是我们做的假设。那么经典物理中什么是可观测的结果呢？时间均匀和空间均匀的假设使得我们可以做同条件重复实验，并假设每一次的测量结果对于同一个物理量来说都是平权的。因此，测量结果应当是一系列实数。

我们首先考虑一个最常见的理解概率的方式：系综理论。回忆上面对于热学系统的讨论，系统的强度量标记了等价类，而一个广延量唯一确定了一个系统。系综的概念是说：对于由上确定的一个系统，即所有广延量与强度量均相等，考虑无穷多个宏观上完全相同的系统构成的序列，其拥有独立的微观状态。其宏观量将不再是一个实数，而是构成了一个序列；我们假定每个系统平权。在此意义上，我们假定的宏观量应当指的是最可几值或均值（顺便指出，宏观热力学的适用范围即此分布最可几值等于均值，Landau相变的失效的统计解释就是在临界点附近最可几值不再等于均值）。我们观测的结果是针对系综中的某一个系统的，因此观测结果也将是一个分布。

由于系综中系统的独立性与观测的独立性，我们可以使用Liapunov中心极限定理：

设$\{x_n, n\in\mathcal{N}^*\}$为相互独立的随机变量序列，$E(X_n)=0$，若

$$\exists\,\delta>0, \frac{1}{s_n^{2+\delta}}\sum_{k=1}^{n}E(|X_k|^{2+\delta})\to 0\ (n\to\infty)$$

则：

$$P\circ(\frac{S_n}{s_n})^{-1}\Rightarrow N(0,1)$$

其中

$$S_n=\sum_{k=1}^{n}X_k,\ s_n=(\mathrm{var}(S_n))^{\frac{1}{2}}$$

即对系综的观测结果应当是一个正态分布。这也是数理统计中处理实验结果的核心思想。

但是自然而然地，我们会发出疑问：以热力学为例，理论框架中的强度量标记了等价类，那么对于两次测量得到了两个不同的结果，我们理应将其放入不同的等价类，而不是将等价类的假设改为均值：毕竟按照热力学的观点，这就是两个不同的系统，或是根本就不满足平衡态的定义。举例来说，考虑连续测量一盒气体的温度，得到了一组数据，那我们怎么假定这是来自一个系综的测量结果，而不是实验条件发生了变化？事实上我们的做法往往是反过来的：当我们得到的数据是正态分布时，我们说它们来自一个系综；否则是实验本身出了问题（例如在上面的实验中，如果测得的温度越来越高，则没有人会说这种偏离是来自于系综的分布）。

问题来自于我们的假设：宏观状态相同而微观状态独立。宏观结果是我们测得的，而微观状态是我们的假设（在热学中是如此）。因此，如果我们不强调热学的定义而考虑所有的物理，这种观测分布的出现本质是我们舍弃了区分一些信息。经典的部分仍然存在：如果我们将所有测量结果视为分布，则当我们了解一个物理系统的所有信息后，测得的分布是一个Dirac delta函数，测量结果回归到了我们熟悉的情形，同时也是前面提到的数学意义上的测量。而当我们放弃区分一些信息而将系统混合，我们就得到了状态系综；

由于中心极限定理，我们测量的结果分布会是正态的。

因此，我们可以说：经典物理中的概率来源于信息的缺失。更进一步地，我们有统计物理基本假设：等几率假设。狭义地说，其说明对于热学系统，微观态概率相同。广义地说，其说明我们通过舍弃信息构造出的物理模型概率是一种 Bayes 概率。

另外，对于数学系的学生来说，更熟悉的说法或许是：物理状态定义了一种测度，系综中的物理状态构成了一个σ-代数。这不是我们讨论的重点，故在此略去；但值得说明的是，我们同样可以用测度的观点看待量子测量；事实上，测度（measurement）和测量（measurement）本来就是同一个概念，只是翻译将其区分开了。

最后我们指出：既然我们前面已经讨论了测量过程中人存在的必要性，那么 Dirac 函数的概率分布就不可能被观测到；人作为无法被移除的信息缺失始终存在。

至此，我们已经重新审视了经典意义下的测量。最后，我们讨论非理想测量的问题。测量概念强依赖于探测子的概念，即其与被测物之前存在一种特殊的相互作用，其可以被被测物影响，而被测物无法受到其影响。热学实验中热容为 0 的温度计、电磁学实验中的试探电荷与相对论实验中的理想观测者都是这个概念。比如类时曲线是有质量粒子的世界线，作为一个观测者，其质量分布必然造成引力场改变（甚至可能造成非稳态的时空解），但是我们假定的理想观测者在类时世界上运动而不弯曲引力场。无论是否理想，观测意味着一种相互作用；而在很多情况下，量子力学中是没有探测子的，观测意味着一种更特殊的相互作用。

5 量子力学中的测量

我们开始讨论量子力学中的测量概念。首先，我们讨论其中最古老的看法：Copenhagen 学派。

5.1 经典量子力学的公理化体系

我们首先介绍近代经典量子力学的基本理论框架。我们略去那些量子力学建立过程中的问题，将重点放在一些较为公认的量子力学理论上，并给出其基本原理或假设。

历史上，由 Heisenberg 建立的矩阵力学与由 Schrodinger 建立的波动力学形成了从 Bohr 的旧量子论到近代公理化量子力学的桥梁。而随着 Dirac 建立了 Dirac 符号[10]，这两套理论在物理内涵和数学形式上被统一，沿用至今的量子力学记号和表述被建立。经典量子力学认为，波动力学与矩阵力学是完全等价的。

量子力学的基本假设一般被认为包含五条，其中三条描述了一个量子系统，被称为态假设（State Postulate）、算符假设（Operator Postulate）和测量假设（Measurement Postulate）；一条描述了其随时间连续演化，被称为动力学假设（Dynamics Postulate）；一条描述了其对称性，被称为全同粒子假设（Identical Particle Postulate）。

我们在此采用一种较为令人容易接受的表述[11][12]，至于一些教材坚持以波函数作为论述的起点，我深感遗憾。这如同在学习群论之前就学习了群表示论，分别地学习了矩阵表示和线性变换表示，然后声称其二者的等价性并开始记号滥用。这对理解一个理论的思想内涵显然是毫无好处的。

我们首先关注前三条：态假设的表述是：量子系统的态（state）被 Hilbert 空间中的矢量表示，其配有常义下 Hilbert 空间的内积。算符假设的表述是：力学量（或译可观测量，observable）被 Hilbert 空间上的厄米算符（hermitian operator）表示。测量假设是 Copenhagen 学派的核心内容，其表述为：一、对力学量的观测得到的结果一定是其本征值之一；二、由于力学量算符的本征值构成 Hilbert 空间的一组正交基，在某一个态下测量到某一个本征值的概率是（归一的）该态与本征值对应本征矢的内积的模平方；三、测量会导致系统坍缩（collapse）到测量结果对应的本征态。

以上内容实际上已经足够，但是其仅支持我们进行代数的运算，为了进行发展分析的理论，人们发展了表象理论（representation，注意其和代数表示论的表示是同一个词，事实上其也可以被视作一种代数表示论；因此，或许翻译为表示更合理）。简单地说，表象理论就是使用某一个算符的本征值作为 Hilbert 空间的基，并在该基下对态矢量和算符进行展开，其展开系数被称为波函数（wavefunction）；由于算符是厄米的，这是始终可以做到的。然后，我们可以处理作为函数的展开系数而不是态矢量本身。需要注意的是，这个过程中没有引入任何新的假设。量子力学的课本往往声称算符的本征矢（或在一些高等量子力学课程中被称为谱（spectrum））分为离散的（discrete）和连续的（continuous）两种，其对应了不同的正交归一化（orthonormalization）方法与本征值，更进一步地定义了不同的波函数。事实上，这意味

着我们要时刻注意谈论的算子作用于哪一个线性空间。当谈论算符作用于非自己定义域的线性空间时，我们事实上是在谈论一个作用于张量积/空间的张量积算符，其分量算符是原算符。

注意到测量假设的一个隐藏内涵：对一个可观测量的观测一定有结果，这是我们使用“概率”这种语言的要求。这意味着，波函数（无论是离散的还是连续的）都必须满足归一化条件；作为波函数，其应当是平方可积的。我们有时会谈论那些平方可积函数空间外的波函数，但是其不单独具有物理意义；或者说，其对应态矢量属于更大的空间，即装备 Hilbert 空间（rigged Hilbert space）。在讨论这些波函数的完备性时需要小心。

动力学假设的基本内容是给出了随时间连续演化的量子系统的幺正演化方程，Heisenberg 图像中的演化方程对应于经典力学中的 Hamiltonian 正则方程。图像（或译绘景，picture）描述了我们看待演化的方式，其类似微分流形中同胚变换的主动观点与被动观点。注意此处的动力学演化都是连续的，即对于时间反演变换是不会出现奇异性的。

全同粒子假设一般被表述为当我们交换体系中的一些粒子的态时，体系的变化是不变或是取反。简单地说，其类似 Noethor 定理的量子力学版本，描述了一种量子力学所特有的对称性。这表明我们无法区分两个粒子，进而在量子统计中，计算可能状态时只能谈论粒子数目而无法谈论某一个粒子与其它相同粒子的排列。这会影响 Bayes 概率的基本事件数目。

5.2 Copenhagen 学派的测量

我们详细讨论 Copenhagen 学派的测量概念。我们介绍 von Neumann 的密度算符理论，用一种更确的方式理解。

令$\mathcal{H}$是一个复数域上的有限维 Hilbert 空间。注意其是一个确定的空间，而非一般的量子力学课本中所指的 Hilbert 空间，那一般意味着一个张量代数：

$$\mathrm{H}=\bigoplus_{r=0}^{\infty}\mathcal{H}^r,\dim(\mathcal{H}^r)=r$$

我们此处谈论的 Hilbert 空间有确定的维数且是有限维的，且其配有的内积是 hermite 线性的。我们说一个给定的量子系统是一个 Hilbert 空间。

由于其是有限维的，我们可以合法地谈论其矩阵表示。同时，其同构于其对偶空间也是显然的：

$$\mathcal{H}^*\cong\mathcal{H}$$

记其全体自同态为$\mathrm{End}(\mathcal{H})$。由于对偶的同构，我们有

$$A\in\mathrm{End}(\mathcal{H})\ iff\ A^*\in\mathrm{End}(\mathcal{H})$$

我们考虑其子空间。使用群的记法（考虑到线性空间首先是一个 Abel 群，这个记法是合理的），考虑其子空间$V\leq\mathcal{H}$，其补空间为：

$$V^{\perp}=\{v\in\mathcal{H}\ iff\ (v,u)=0\ \forall u\in\mathcal{H}\}$$

显然

$$\mathcal{H}=V\oplus V^{\perp}$$

则我们定义正交投影算符

$$\mathrm{P}_V\in\mathrm{End}(\mathcal{H}):\mathrm{P}_V(v)=v_{\parallel}$$

$$\forall v=v_{\perp}+v_{\parallel}\ where\ v_{\perp}\in V^{\perp},v_{\perp}\in V^{\perp}$$

我们说态（state）是一个正定的自同态：

$$\mathcal{S}=\{M\in\mathrm{End}(\mathcal{H}),M^*=M,\mathrm{tr}(M)=1,M\geq 0\}$$

其中迹为 1 的条件保证了其归一化，正定的定义为：

$$(u,M(u))\geq 0,\forall u\in\mathcal{H}$$

可以证明，由于其 hermitian，这等价于其可以分解为投影算符的线性组合（即泛函分析中的谱定理）[5]：

$$M=\sum_{i=1}^{k}m_iP_i,m_i\geq 0$$

其中迹为 1 的条件保证了：

$$\sum_{i=1}^{k}m_i\dim(V_i)=1$$

其中每个投影子空间的维数$\dim(V_i)$被称为简并度（degeneracy）。

我们指出：此时的态矢量包括纯态（pure state）和混合态（mixed state），前面的经典量子力学公理化中谈论的态均为纯态。为了区分量子态和 Hilbert 空间中的矢量，常见将此处的态矢量称为密度算符。纯态的条件是上述求和只有一项：

$$M=m\mathrm{P}_V$$

请注意虽然此处讨论的态矢量空间更大（包括混态），但是仍是正统的 Copenhagen 学派，其重点在于下面的测量假设。有的物理学家通过引入混合态的概念来改变量子力学基本原理，这是徒劳的，其只是将问题重新说了一遍。

力学量是 hermite 算符：

$$\mathcal{A}=\{A\in\mathrm{End}(\mathcal{H}),A^*=A\}$$

类似的，其有谱分解：

$$A=\sum_{i=1}^{k}\lambda_i \mathrm{P_i}$$

然后是测量，即著名的 Born 规则：在某一个态对某一个力学量的测量定义了一种实数上的概率测度。通过我们前面对经典测量的讨论，我们知道经典概率来源于信息缺失（或使用 Schrodinger 的语言，无知（ignorance）），但是量子力学中的概率不来源于此，而是量子理论的基本属性。但是，既然我们要谈论概率，又不能借用习以为常的信息理论，我们需要给概率以新的定义。在数学上，概率被概率测度定义，而一个测度来源于一个σ-代数。我们需要在量子理论的基本假设中给出有关概率的完备定义，这是通过测量假设给出的。另外在此指出，之所有上面使用如此晦涩而数学的语言改写量子力学基本假设，就是为了给概率一个良好的定义。有趣的是，有关谱定理等很多数学定理是由 von Neumann 证明的，而上述密度算符的理论也是由 von Neumann 提出的，这二者之间存在显然的关系。

量子测量是一个如下的映射：

$$\mathcal{A}\times\mathcal{S}\to\mu_{-}^{-}:E\to[0,1],\forall E\in\mathcal{B}(\mathcal{R})$$

其中μ_A^M是一个概率测度，$\mathcal{B}(\mathcal{R})$被称为 von Neumann-Bole 代数，可以理解为可测集的集合。

具体地说，对于任何$\mathcal{R}$的子集，给出的概率测度满足：

$$\mu_A^M(E)=\mathrm{tr}(\mathrm{P}_A(E)M),\ \forall E\in\mathcal{B}(\mathcal{R})$$

可以证明，这确实给出了一个概率测度[5]。

而测量后的态矢量变为：

$$M\Rightarrow \mathrm{P}_A(E)M$$

即坍缩向力学量本征态。

在非测量时，态矢量演化满足 Schrodinger 方程：

$$M\Rightarrow M(t)=U(t)MU(t)^{-1},\forall M\in\mathcal{S}$$

其中$U(t)$是一个由 Hamilton 量定义的幺正（unitary）算符：

$$U(t)=e^{-\frac{i}{\hbar}tH},\forall t\in\mathcal{R}$$

至此，Copenhagen 学派给出了不依赖于经典物理信息理论的概率定义。但是同时，其立即带来了一个最大的问题，即系统遵循两套独立的、不相容的演化规则：不进行测量时，其满足幺正演化；测量时，其遵循 Born 规则，即态矢量塌缩。事实上，这是量子力学与经典物理最大的差别。

另外，我们注意到幺正演化一定是可逆的，而投影演化一般是不可逆的。对于不可逆的演化，我们无法合法地谈论时间反演的对称性。

我们立即产生三个疑问：首先，我们如何区分一个系统是在被测量、或是不被测量？这是否依赖于人的意志？其次，既然非测量过程的演化与测量过程不同，这是否意味着我们只能对其进行推断而非观察？最后，按照经典物理的理论，测量是通过某种相互作用获得信息，而我们总能通过将测量仪器和被测量物放在一起看作一个更大的系统，使得系统整体遵循与被测量物单独存在时相同的测量规律。但是在量子物理中，这是否无论如何都是办不到的？或者说，这是否意味着量子物理中的测量一定是非物理的？

完全解释这些问题远远超过我们、以及任何物理学家的能力。但是，我们有能力总结前人对这些问题的看法。

我们想再次强调，这些问题不仅是不显然的，而且也是尚无定论的。从 Copenhagen 学派到 EPR 理论，没有哪一种理论如同 Newton 三定律或是 Maxwell 方程那般令人感到舒适，或许是因为不存在这样一个能级，使得在其之内量子力学是完备、定域且自洽的。正因如此，也没有哪个自洽的理论是一定错误的，无论其看上去多么复杂。以上繁复的讨论，从经典物理到量子力学，足以提醒我们谨慎地看待每一个观点，包括我们自己看待实验结果的方式。正如 Laloe 著名的量子力学著作的书名，应当时刻提醒自己，Do we really understand quantum mechanics?

5.3 Bohr 的理解

作为量子力学的早期学者，Bohr 对演化的理解与 von Neumann 不同。其理论主要可以被概括为：量子力学的测量分为两个阶段：准备阶段与被测量阶段。当考虑连续测量时，即使用的仪器不变，在第二次测量前发生的一切演化均属于第二次测量的准备阶段；所以，询问两次测量间的系统状态演化是没有意义的，包括第一次测量在内的所有行为都被打包进了第二次测量的准备中。按照这种观点，人们无需考虑两种演化规则的差异（因为根本就不存在两种演化）。准备、然后测量，仅此而已。因此，我们也无需要考虑态矢量的坍缩。

但是等价的，这带来了新的问题：是什么使得我们关注的那次测量变得如此不同？我们将如何区分我们关注的那次测量的测量仪器、与其他的任何准备过程？Bohr 没有解释这些问题。

回到 Copenhagen 学派的观点。Copenhagen 学派事实上使用了一种很简单的方法来区分测量仪器与被测量对象，其被称为 Heisenberg 截断（Heisenberg cut）。在截断之下的物理规律是量子的，而在截断之上的测量仪器。对此，有几个著名的思想实验说明这种截断的困难。

5.4 von Neumann 无限链

von Neumann 任何可以被观测的结果都是宏观的（macroscopic），而被预言可观测的量子效应都是微观的（microscopic）。因此，我们的量子观测仪器应该是一种从微观到宏观的映射。由于 Schrodinger 方程是线性的：

$$M \Rightarrow M(t) = U(t)MU(t)^{-1}, \forall M \in \mathcal{S}$$

假设我们的测量仪器也是线性的，则其相当于将一个微观态变成一个（较为）宏观的指针。而对于足够宏观的指针，我们总可以用上面在经典测量中讨论的等价关系等方法，使用经典测量将其映射为一个实数。另外，当我们这个仪器的跨越足够小时，我们总能保证其映射的线性。

对于那些处于力学量本征态的态矢量：

$$[M, A]_Q = 0$$

这种测量总能给出一个确定的指针位置，进而给出一个确定的宏观观测结果与实数。但是对于那些不在本征态上的态矢量：

$$[M, A]_Q \neq 0$$

由于我们假设的测量仪器是线性的，这意味着其给出的指针没有一个确定的位置，也是某种线性叠加。由于宏观上不存在叠加态，这意味着我们的指针仍是微观的，其可以被看作一个量子系统。我们再次使用一个更宏观的指针来标记我们的指针，并使得其也是线性的。这样的行为总是可以继续下去的，使得最终我们获得一个宏观的指针，这种观测方式被称为 von Neuman 链（chain）或 von Neumann 退后（regress）。

但是，由于我们始终无法观察到宏观意义上的叠加现象，这种链一定包含无限个过程。那么，Heisenberg 截断出现在什么地方？

一种观点是无限过程使得量子效应涌现为经典的，其衍生了退相干理论（decoherence）；另一种观点是宏观世界具有特殊性，而其衍生了可能是物理学中最著名的思想实验：Schrodinger 的猫态。

5.5 Schrodinger 的猫态

Schrodinger 的猫实验的主要内容是：考虑一个完全封闭的盒子（这杜绝了系统-环境耦合），其中有一个放射性的原子拥有固定的衰变概率密度，当原子衰变，则猫死亡。这个盒子中不存在任何观测者。则在打开盒子之前，猫可以被认为处于叠加态吗？

这个实验非常著名，且广受讨论，但是事实上我们指出，Schrodinger 的原意非常简单：与 von Neumann 的无限链中的指针类似，其也构建了一个可能存在的处于宏观的叠加态。那么我们可以认为这种宏观态是合法的、甚至认为其演化满足 Schrodinger 方程吗？问题在于，如果我们认可这一点，那么 Heisenberg 截断将变成一个完全宏观的截断。进而，宏观量子效应（不是量子效应的宏观体现）一定可观测。但是实验证明宏观叠加是不存在的。这进一步说明了 Heisenberg 截断的困难。

5.6 Wigner 的朋友

最后，我们介绍一个有关自由意志的 Schrodinger 猫态思想实验，其由 Wigner 提出。

同样假设一个隔绝了系统-环境耦合的封闭合理，不同的是这次物理学家 Wigner 将其朋友至于其中，并令其做一个量子力学实验，如 Stern-Gerlach 实验。这个朋友拥有和 Wigner 通讯的能力，但是只能告知实验是否有结果，而不能告知自旋是向上或是向下。对于 Wigner，这事实上无法判断实验结果，进而其没有获取任何信息。那么我们考虑时间反演过程：对于 Wigner，这没有任何问题；但是对于他那个在盒子里做实验的朋友，其遇到了奇异性，因为其做过观测并导致了坍缩。那么这种奇异-非奇异的现象是有什么导致的呢？换句话说，如果盒子里没有 Wigner 的朋友，而是只是有一堆原子，那么盒外和盒内的时间反演都能各自进行。是否是盒内存在 Wigner 的朋友这一自由意志导致了结果的不同？我们发现在这个实验中，Heisenberg 截断存在于我们的脑中，这更加不可思议。

以上我们讨论了 Copenhagen 学派内的测量概念，我们可以发现其事实上面临的主要问题是如何解释 Heisenberg 截断。另一方面，所有理论都基于一个假设：可以考虑一个封闭的系统。这意味着即使存在环境扰动，这种扰动也应该是可以被忽略的。对于经典理论，这种假设总是显见的；但是对于量子理论，一些人认为这未必成立。因此，退相干理论在 20 世纪 40 年代被提出，其基本思路是使用环境耦合的解释来使得不必假设一个 Heisenberg 截断的必要性，而是否认经典理论的主体地位，将其视作量子理论的一种涌现。

最后我们对 Heisenberg 截断对当今物理的意

义进行讨论。在 Copenhagen 学派提出 Heisenberg 截断的时代，包括在 Schrodinger、甚至更晚的一些时代，宏观与微观的区分是拥有不言自明的鸿沟的，猫与原子不可能被用同一套理论叙事讨论。然而如今越来越多的原子纠缠被实现，直接探测的技术也越来越趋向量子极限，这种界限有时容易被忽略。比如如今当某实验室制备出更多的原子纠缠态时，“一个更大的猫态被制备”之类的描述常常被使用。但是或许这种描述在其领域有意义，对于理论量子科学而言，我们应记住这是错误的。Schrodinger 的猫态指的是一种无法被制备的东西：宏观纠缠态。这与技术能力无关，而是一种对理论层面上完备性的描述， 是对 Heisenberg 截断存在的必要性的描述。而当我们还在使用 Copenhagen 学派、或是上面讨论的一些“后 Copenhagen 学派”的解释时，就不应该论宏观纠缠；通过上面的讨论我们也能发现，Heisenberg 截断不是通过制备更宏观的纠缠就能解决的（通过对 von Neuman 无限链的理解，这不过是将边界推前了几步）。另一方面，我们也可以重新审视能级概念对于理论的意义：理论总是有自己存在的能级，但是对于量子力学，其意义尤为重大。现在的量子信息技术是不足以改变其存在的能级的。

6 EPR 理论简述

我们已经回顾了经典量子力学的核心概念，并考察了历史上为解释其做出的努力。接下来我们考虑来自另一个方面的对其的攻击，即 EPR 理论。其来自于 Einstein、Podolsky 和 Rosen 于 1935 年发表于 Physics Review 上的一篇文章[15]，这也被认为是量子信息领域的奠基之作之一。这篇文章是 PR 上被引用最多的文章之一，同时也是被误解最多的（即使是 Born 这样的物理学家）。事实上，Einstein 等人的这篇文章的动机十分明确，其不是为了寻找一个悖论来攻击量子力学（其有时被不恰当地称为 EPR paradox），也不是为了嘲讽其同事，或是像上面的诠释一样探讨一些心理学或哲学层面的原理。相反，这篇文章的出发点是纯粹逻辑合理的，其目的是为了阐述量子理论的定域性和实在性不可同时存在；从今天的视角看，他们的理论相当成功。

简单地说，EPR 理论的内容是：如果量子力学的预测正确、特别是对远程相互纠缠的粒子正确，且我们相信物理理论可以完备且可分地描述局域的世界，那么量子力学一定不是一个完备的理论，其至少忽略了一些构成完备性的信息。

我们简单回顾广义相对论对因果的论述：一点的编时未来和编时过去由其未来定向和过去定向的类时曲线可以达到的点构成，在满足强因果条件时，其诱导了一个时空流形上的一个拓扑。对于一个不存在闭合类时曲线的时空流形，我们称其满足编时条件。可以证明，在此意义上，广义相对论是一个满足定域条件的完备物理理论。

将其运用于量子力学，我们可以发现：如果继续采用上面的假设，则量子力学中要么无法使用波函数完备的描述一个物理体系（那些不可对易的物理量），要么不具有局域性（违反因果关系）：他们用作例子的实验是：考虑两个波函数描述系统，其在某个时刻前相互作用，在其之后各自分离；其分离后不再相互作用。在其分离后，我们测量第一个系统的位置；则第二个系统的状态即为其位置本征矢，这意味着我们在未对其扰动的情况下获取了其信息。这是量子纠缠概念的提出。

事实上，量子纠缠已经被实验证实，但是这并不能被用来反驳 EPR 理论；相反，这说明他们的假设是对的：量子力学确实无法同时具有完备性与定域性。关于此的讨论较为复杂，在此无法详细展开。但是我们可以知道：这确实为我们寻找量子理论的解释提供了另外的思路。

7 重新审视量子测量的意义

正如我们在最开始说的那样，量子力学可能是世界上最成功的物理理论，但同时也是最难被理解的。有的理论难以被理解是因为描述它所需要的数学过于庞大，比如广义相对论一般需要用微分几何来描述。诚然，量子力学也需要测度论、概率论和泛函分析这样的数学工具，但是通过以上所有讨论我们可以发现：数学工具的艰深不是我们理解一个理论的困难，相反，使用更恰当而完备的数学工具往往可以帮助我们理解一个理论，正如我们使用量子测度取代 Bayes 理论来理解概率可以帮助我们理解量子意义下的概率。真正困难的是物理上的理解。

那么，为什么我们一定得理解量子力学不可呢？要知道，学习量子场论、量子统计、量子色动力学甚至是弦论等建立在量子力学上的学科都不需要像我们一样拷问量子力学的基本原理，只要接受就好了，毕竟物理学家最喜欢的事就是只要实验结果正确，那多么丑陋的理论都是可接受的。而且，量子力学在技术和工业上的成就可

能也是有史以来所有物理理论中最大的，那么为什么我们还要在此讨论其基本原理呢？

我认为，重要的是动机（motivation）。借用我最喜欢的量子力学有关著作 *Do we really understand quantum mechanics?* 中的话：One should nevertheless realize that the most important and urgent question at the time was not so much to explain the fine details of the properties of interactions between radiation and matter, or the peculiarities of the blackbody radiation. It was more general: to understand the origin of the stability of atoms, that is, of all matter which surrounds us and of which we are made! 那么量子力学成功地解释了这个问题吗？很遗憾，没有，因为其做出的假设不比其要解决的问题更容易被人接受。或者说，不比直接将其作为基本假设更显然。这是我们要考虑量子力学基本假设的原因：不管是改变基本假设或是改变我们的直觉，我们绝不接受一个不符合直觉的假设，我们不接受一个理论只能被用来计算而不能被用来理解。

现代科学的主流哲学其实是现象学分析。特别是凝聚态领域，基本所有人都秉持 *More is Different* 的思路，认为从基本假设或是公理出发的理论是特别局限的，理论越唯象就越实用。或许这是对的。又或许这是因为我们的基本假设经不起检验。无论如何，我们相信，物理学终究是为了更好地理解这个世界，为此，我们不应放弃解构与追问为什么，至少应当允许有人在这么做。毕竟，当人类第一次仰望星空的时候，物理学的大厦就已经建立百分之四十了。

8 参考文献

[1] 丘维声，高等代数学, 2nd. Ed. 清华大学出版社, 2019.
[2] 曹利明，引力理论, draft, 2021.
[3] 朱界杰，理论力学 A 讲义, draft, 2023.
[4] 埃格先生，微分几何与分析力学初步, draft, 2018.
[5] Leon A. Takhtajan. Quantum Mechanics for Mathematicians, 1st. Ed. American Mathematical Society.
[6] 陈省身，陈维桓，微分几何讲义, 2nd. Ed. 北京大学出版社, 2001.
[7] Herbert B. Callen, Thermodynamics and an Introduction to Thermostatistics, 2nd, Ed. University of Pennsylvania.
[8] Albert Einstein, Foundation of General Relativity, 1916.
[9] R. Landauer, Irreversibility and Heat Generation in the Computing Process, IBM J. Res. Dev. 5, 183 (1961).
[10] P.A.M. Dirac, The Principles of Quantum Mechanics. 4th. Ed. Vol. Oxford University Press, 1967.
[11] Si Li, Quantum Mechanics: Basic Theory. draft. 2024.
[12] J.J. Sakurai, Jim Napolitano, Modern Quantum Mechanics, 2nd. Ed. Addison-Wesley, 2012.
[13] Frank Laloe, Do We Really Understand Quantum Mechanics, 2nd. Ed. Cambridge University Press, 2019.
[14] A. Einstein, B. Podolsky, N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete, Physics Review, 1935. 5.