

目录

第十二章 *密度算符与熵	2
12.1 密度算符与量子力学原理	3
12.1.1 期望值以及算符迹	3
12.1.2 从量子纠缠态到混态	5
12.1.3 Bloch 球	10
12.1.4 哥本哈根诠释和多世界诠释	12
12.2 再谈密度算符与系综	14
12.3 量子熵	16
12.3.1 冯诺伊曼熵	16
12.3.2 相对熵	19
12.4 开放系统以及热力学第二定律	22
12.4.1 开放系统的时间演化	22
12.4.2 热力学第二定律	25

第十二章 *密度算符与熵

陈童

本章讲述密度算符的概念和其含义。本章也将从一个更高的角度重新审视量子力学的基本原理，并介绍量子力学的多世界诠释。

本章随之系统讲述了冯诺伊曼熵和量子相对熵，以及它们的各种性质。本章还会推导开放系统随着时间的演化方程，并证明热力学第二定律。

12.1 密度算符与量子力学原理

前面章节研究的量子力学系统都是所谓的封闭完整系统，也就是说，我们忽略了系统与外界的相互作用，并且我们研究的是整个系统而不是其中的一部分。但是，现实中的系统总是从属于一个更大的系统，是更大系统的一部分，而不是全部。现实中的系统也总是会与外界环境甚至包括我们的仪器相互作用。如何描述这样的现实系统就是本节要解决的核心问题。通过对这一问题的讨论，我们也得以从一个更高的视角来重新审视量子力学的基本原理。

12.1.1 期望值以及算符迹

下面我们先作一些准备。

测量与期望值

在第二章关于量子力学基本原理的讨论中，我们得出过一个重要公式：对某个物理量 A 多次重复实验测得的平均值 \bar{A} 等于算符 A 在态上的期望值，即

$$\bar{A} = \langle \psi | A | \psi \rangle, \quad (12.1)$$

$|\psi\rangle$ 为系统所处的量子态。我们也说过，物理量 A 的测量值是相应算符 A 的本征值，在单次实验中，我们测得哪个值是随机的，测得本征值 λ_i 的概率为 $p_i = |\langle i | \psi \rangle|^2$ ，式中 $|i\rangle$ 为与本征值 λ_i 相应的本征态。这样算出来的概率 p_i 当然也可以通过重复多次实验来检验，因此其计算公式应该也能写成期望值的形式，的确，假如定义投影算符 $P_i = |i\rangle\langle i|$ ，则人们很容易验证

$$p_i = \langle \psi | i \rangle \langle i | \psi \rangle = \langle \psi | P_i | \psi \rangle. \quad (12.2)$$

因此，不仅仅物理量的平均值，而且物理量值的概率分布，都可以表达成厄密算符期望值的形式。

以上讨论告诉我们，在量子力学中一切可以通过重复实验来进行检验的量都可以由一个适当厄密算符的期望值来计算，假设这个厄密算符是 \mathcal{O} ，习惯上，人们常将这个期望值简记为 $\langle \mathcal{O} \rangle$ ，

$$\langle \mathcal{O} \rangle = \langle \psi | \mathcal{O} | \psi \rangle. \quad (12.3)$$

各种厄密算符的期望值就是我们能够从一个量子态中提取的所有信息。

算符求迹

一个矩阵所有对角元的和称为矩阵的迹，在么正变换下矩阵的迹保持不变，这使得我们可以将迹的概念推广到线性算符。对于一个线性算符 \mathcal{O} ，我们记其迹为 $\text{Tr}(\mathcal{O})$ ，定义为

$$\text{Tr}(\mathcal{O}) = \sum_i \langle i | \mathcal{O} | i \rangle, \quad (12.4)$$

式中 $\{|i\rangle\}$ 为希尔伯特空间的一组正交归一矢量基。虽然为了计算算符的迹我们需要选取一个特定的表象，但由于不同的表象之间只相差一个么正变换，所以算符的迹实际上并不依赖于表象。由此可以知道，一个算符的迹就是它在任何一个表象中表示矩阵的迹，特别的，对于厄密算符，我们可以将这个表象选为它的本征表象，这时候算符的迹其实就是所有本征值的和。

假设有两个算符 A 和 B ，读者容易证明算符迹满足如下等式

$$\text{Tr}(AB) = \text{Tr}(BA). \quad (12.5)$$

不过，值得说明的是，对于无穷维希尔伯特空间，并非所有算符的迹都是定义良好的，因此在无穷维希尔伯特空间上应用(12.5)式时得特别小心，有时候它并不成立。比方说，对于坐标算符和动量算符的对易子来说，如果简单地应用这个公式将会得出 $\text{Tr}([X, P]) = 0$ ，但实际上 $\text{Tr}([X, P]) = i\hbar \text{Tr}(1) \neq 0$ 。在这个例子中，(12.5)式之所以不成立，正是因为式中各算符的迹都不是良好定义的。

利用算符的迹，我们可以把算符的期望值公式重写成

$$\langle \psi | \mathcal{O} | \psi \rangle = \text{Tr}(|\psi\rangle\langle\psi| \mathcal{O}) = \text{Tr}(\rho_\psi \mathcal{O}). \quad (12.6)$$

式中投影算符 $\rho_\psi = |\psi\rangle\langle\psi|$ 。要证明这个式子，我们只需注意到

$$\text{Tr}(|\psi\rangle\langle\psi| \mathcal{O}) = \sum_i \langle i | \psi \rangle \langle \psi | \mathcal{O} | i \rangle = \sum_i \langle \psi | \mathcal{O} | i \rangle \langle i | \psi \rangle = \langle \psi | \mathcal{O} | \psi \rangle.$$

可见，为了计算算符的期望值，我们并不需要知道量子态 $|\psi\rangle$ ，而是只需要知道厄密算符 ρ_ψ 。 $|\psi\rangle$ 和 ρ_ψ 的一个重要区别是， $|\psi\rangle$ 可以相差一个非物理的整体相位因子，变成 $e^{i\theta}|\psi\rangle$ ，而 ρ_ψ 在这种相位变换下不变，也即是说，它自动剔除了这一非物理的整体相位信息。

以上讨论告诉我们，从一个量子系统中能够提取出来的所有物理信息都包含在 $\rho_\psi = |\psi\rangle\langle\psi|$ 这一特殊厄密算符中。注意到 $\langle x|\rho_\psi|x\rangle = |\langle x|\psi\rangle|^2$ 正好是概率密度。即是说， ρ_ψ 在坐标表象中的对角元对应概率密度，所以人们常常称 ρ_ψ 为纯态密度算符，它和狄拉克符号 $|\psi\rangle$ 表示的量子态是对应的，这样的态也称为纯态。下面我们就是要将密度算符这一概念推广到更为一般的情形。

12.1.2 从量子纠缠态到混态

密度算符的一般概念和这样的问题密切相关，即对于一个由两部分组成的系统，假设我们仅仅关心其中的一部分，该如何描述我们所关心的对象？这一小节我们就是要解决这个问题。

一个例子

不妨先看一个例子。假设我们有两个量子比特，分别称作量子比特 A 和量子比特 B ，它们的希尔伯特分别为 \mathcal{H}_A 和 \mathcal{H}_B 。假设 A, B 两者共同处在 $|\psi\rangle$ 态，

$$|\psi\rangle = c_0|00\rangle + c_1|11\rangle. \quad (12.7)$$

则我们很容易写出与这个态相应的纯态密度算符

$$\rho_\psi = |c_0|^2|00\rangle\langle 00| + |c_1|^2|11\rangle\langle 11| + c_0c_1^*|00\rangle\langle 11| + c_1c_0^*|11\rangle\langle 00|. \quad (12.8)$$

我们称这个密度算符前面的两项为对角项，后面的两项为非对角项。非对角项相应于通常所说的干涉项，它描写了相互叠加的两个态之间的干涉。

假设我们忽略 B 而仅仅只关心 A ，对 A 进行各种测量。比方说我们测量 A 的某个物理量 \mathcal{O}_A ，因此我们就要计算如下期望值，

$$\langle\psi|\mathcal{O}_A|\psi\rangle = \text{Tr}_{AB}(\rho_\psi\mathcal{O}_A). \quad (12.9)$$

式中 Tr_{AB} 表示在整个系统的张量积希尔伯特空间 $\mathcal{H}_A \otimes \mathcal{H}_B$ 上求迹。注意到 $\mathcal{H}_A \otimes \mathcal{H}_B$ 的基矢量 $|ij\rangle_{AB} = |i\rangle_A|j\rangle_B$ ， $i, j = 0, 1$ ，所以 $\text{Tr}_{AB} = \text{Tr}_A\text{Tr}_B$ ，式中 Tr_A 和 Tr_B 分别表示在 \mathcal{H}_A 和 \mathcal{H}_B 上求迹。由于 \mathcal{O}_A 只作用在 \mathcal{H}_A 上，与 B 量子比特无关，所以 Tr_B 就直接作用在 ρ_ψ 上，得到

$$\rho_A = \text{Tr}_B(\rho_\psi) = |c_0|^2|0\rangle\langle 0| + |c_1|^2|1\rangle\langle 1|. \quad (12.10)$$

这是一个只作用在 \mathcal{H}_A 上的厄密算符。而期望值公式(12.9)就变成

$$\langle \psi | \mathcal{O}_A | \psi \rangle = \text{Tr}_A(\rho_A \mathcal{O}_A). \quad (12.11)$$

以上讨论告诉我们，如果仅仅只关心 A 量子比特，那就无需知道 $|\psi\rangle$ ，只需知道 ρ_A ，忽略 B 以后所有能从 A 量子比特提取出来的信息都包含在厄密算符 ρ_A 中。我们称 ρ_A 为 A 量子比特的密度算符，当然它不是前面引入的纯态密度算符，因为现在这个密度算符不能写成 $|\phi_A\rangle\langle\phi_A|$ 这样的形式，我们称这种密度算符为混态密度算符，简称混态。很明显， ρ_A 满足

$$\text{Tr}_A(\rho_A) = |c_0|^2 + |c_1|^2 = 1. \quad (12.12)$$

在这个例子中，系统所处的 $|\psi\rangle = c_0|00\rangle + c_1|11\rangle$ 态显然是一个纠缠态，上面的推导过程告诉我们，如果忽略 B 量子比特(数学上就是将 ρ_ψ 对 B 求迹，称作trace掉 B)，我们将得到一个 A 的混态密度算符 ρ_A ，称作 ρ_ψ 在 A 上的约化密度算符。从(12.10)式可以看得很清楚，在密度算符 ρ_A 中，量子比特0态与1态间的干涉项消失了，这就叫做退相干。在这个例子中， A 的0态与1态退相干的原因有两个：首先， A 和 B 发生了量子纠缠，其次，我们忽略了 B 。

从(12.10)式还可以看到，退相干的 A 量子比特对于忽略了 B 的我们来说，完全可以看作是以 $|c_0|^2$ 的概率处于 $|0\rangle$ 态，以 $|c_1|^2$ 的概率处于 $|1\rangle$ 态，这分别对应(12.10)式中的 $|c_0|^2|0\rangle\langle 0|$ 和 $|c_1|^2|1\rangle\langle 1|$ 。

量子擦除

我们知道，一个量子比特如果处在 $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ 态，那它的 $|0\rangle$ 态和 $|1\rangle$ 态之间将会有干涉，两者的相对位相可以观测，通过对相对位相的观测我们可以把 $|+\rangle$ 态和 $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ 态100%地区分开来。但是，从前面的讨论中我们也知道，如果将这个量子比特作为 A 与另一个量子比特 B 纠缠起来，使得 AB 整体处于 $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2} = (|++\rangle + |--\rangle)/\sqrt{2}$ 态，那么当我们忽略 B 的时候， A 的 $|0\rangle$ 态和 $|1\rangle$ 态之间将会退相干，这时候 A 将由混态密度算符 $\frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -|)$ 描述， $|0\rangle, |1\rangle$ 之间的相对位相将会变得不可观测。这种不可观测是因为这时候 A 量子比特取值是0还是1的信息通过与 B 的纠缠泄露出去了。

但是，假如我们不是完全忽略 B 的信息，那就有一种方法可以恢复 A 量子比特 $|0\rangle$ 态和 $|1\rangle$ 态之间的相干叠加性。这方法就是，假设我们不

是在 $\{|0\rangle, |1\rangle\}$ 基, 而是在 $\{|+\rangle, |-\rangle\}$ 基中观测 B 量子比特, 并根据观测结果将所有的 A 量子比特分成 $+$, $-$ 两类。 $+$ 类 A 量子比特相应于对 B 的观测结果是 $|+\rangle$, $-$ 类 A 量子比特则相应于对 B 的观测结果是 $|-\rangle$ 。那么这时候 $+$ 类 A 量子比特的 $|0\rangle, |1\rangle$ 态之间将是相干叠加的, $-$ 类也是如此。这是因为, AB 整体的量子态是 $|\phi^+\rangle = (|++\rangle + |--\rangle)/\sqrt{2}$, 因此当 B 的观测结果是 $|+\rangle$ 时, 对相应 A 量子比特的描述将不是混态密度算符 $\frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$, 而是纯态 $|+\rangle$, 同样, 当 B 的观测结果是 $|-\rangle$ 时, 对相应 A 量子比特的描述也是纯态 $|-\rangle$ 。当我们根据对 B 的观测结果将这两类 A 量子比特分开时, 它们各自当然都是相干的。

上面描述的这种恢复 A 量子比特相干性的方法就叫做量子擦除, 因为无论对 B 的测量结果是 $|+\rangle$ 还是 $|-\rangle$, 我们都得不到任何取值是0还是1的信息。因此这种测量就从 B 量子比特中擦除了通过纠缠泄露出去的 A 取值是0还是1的信息。

两体纠缠的刻画-施密特分解

前面两量子比特的例子所揭示的量子纠缠态与子系统混态之间的联系是一般性的, 为了看清楚这一点, 让我们更仔细地考察一下两体量子纠缠。

对于一个由 A 、 B 两部分组成的系统, 假设这两部分的希尔伯特空间分别为 \mathcal{H}_A 和 \mathcal{H}_B , 则任意量子态 $|\psi\rangle$ 必定可以写成

$$|\psi\rangle = \sum_{i_A, i_B} \psi_{i_A i_B} |i_A\rangle |i_B\rangle, \quad (12.13)$$

式中 $\{|i_A\rangle\}$ 和 $\{|i_B\rangle\}$ 分别为 \mathcal{H}_A 和 \mathcal{H}_B 的正交归一矢量基。式中的叠加系数 $\psi_{i_A i_B}$ 有两个指标, 因此可以看成是一个矩阵, 根据矩阵的奇异值分解定理¹, $\psi_{i_A i_B}$ 必定可以分解成如下形式

$$\psi_{i_A i_B} = \sum_a \lambda_a u_{i_A}^{(a)} v_{i_B}^{(a)}, \quad (12.14)$$

式中 $u_{i_A}^{(a)}$ 和 $v_{i_B}^{(a)}$ 分别满足正交归一条件 $\sum_{i_A} (u_{i_A}^{(a)})^* u_{i_A}^{(b)} = \delta_{ab}$ 和 $\sum_{i_B} (v_{i_B}^{(a)})^* v_{i_B}^{(b)} = \delta_{ab}$, 并且我们总可以通过调节 $u_{i_A}^{(a)}$ 和 $v_{i_B}^{(a)}$ 之间的相对相位使得 $\lambda_a \geq 0$, 因此我们可以令 $\lambda_a = \sqrt{p_a}$, $p_a \geq 0$ 。现在记 $|a\rangle_A = \sum_{i_A} u_{i_A}^{(a)} |i_A\rangle$, $|a\rangle_B =$

¹如果将这个矩阵看成是某个算符的表示矩阵, 那矩阵的奇异值分解定理其实就来源于算符的奇异值分解定理。

$\sum_{i_B} v_{i_B}^{(a)} |i_B\rangle$, 则由 $u_{i_A}^{(a)}$ 和 $v_{i_B}^{(a)}$ 的正交归一性可以证明, $|a\rangle_A$ 和 $|a\rangle_B$ 分别正交归一, 即 ${}_A\langle a|b\rangle_A = {}_B\langle a|b\rangle_B = \delta_{ab}$. 将(12.14)式代入(12.13)就可以得到

$$|\psi\rangle = \sum_a \sqrt{p_a} |a\rangle_A |a\rangle_B. \quad (12.15)$$

进一步利用 $|\psi\rangle$ 态的归一化条件, 容易得到

$$\sum_a p_a = 1. \quad (12.16)$$

(12.15)式就叫做两体系统量子态的施密特分解, 其中非零的 p_a 数目就称作量子态的施密特数。显然, 施密特数最小等于1, 最大等于 \mathcal{H}_A 和 \mathcal{H}_B 维数中较小的那个。如果施密特数等于1, 则由(12.15)式很容易看出, 这时候量子态 $|\psi\rangle$ 必定分解成 A, B 量子态乘积的形式, 反过来, 如果施密特数大于1, 那就必定没有这样简单的分解, 从而 $|\psi\rangle$ 态就必定为一个纠缠态。即是说, $|\psi\rangle$ 不是纠缠态的充要条件是施密特数等于1, 因此施密特数多少反映了两体纠缠的状况, 施密特数越大说明 A, B 纠缠得越厉害。

密度算符与系综

一个量子系统通常总是处于某个环境之中, 假设我们把系统看作 A , 把环境看作 B , 那么整个世界就可以看成是一个由 A, B 两部分组成的封闭完整系统, 其 A, B 两部分通常会纠缠起来。下面我们对整个世界的量子态应用施密特分解公式(12.15), 我们记整个世界的密度算符为 $\rho_\psi = |\psi\rangle\langle\psi|$, 假设我们忽略 B 部分, 那就会得到 A 的约化密度算符 $\rho_A = \text{Tr}_B(\rho_\psi)$, 我们从 A 中能够提取出来的所有信息都包含在 ρ_A 中, 并可以通过计算算符期望值 $\text{Tr}_A(\rho_A \mathcal{O}_A)$ 来得到, 因此我们也称 ρ_A 描写了子系统 A 的一个量子态, 只不过这个量子态不一定是纯态, 而可以是混态。

由(12.15)式容易有

$$\rho_A = \sum_a p_a |a\rangle_A \langle a|. \quad (12.17)$$

显然, ρ_A 满足 $\text{Tr}_A(\rho_A) = \sum_a p_a = 1$. 由于我们没有对 A 和 B 进行任何限制, 所以可以想见, (12.17)应该能代表 A 的一个任意量子态。从(12.17)可以看出, 施密特分解的 p_a 刚好是密度算符 ρ_A 的本征值, 施密特分解的 $|a\rangle_A$ 则是

相应的本征态。同样的，我们也可以忽略A部分，进而得到B部分的约化密度算符 $\rho_B = \text{Tr}_A(\rho_\psi)$ ，为

$$\rho_B = \sum_a p_a |a\rangle_B \langle a|. \quad (12.18)$$

如果原来AB整体的量子态 $|\psi\rangle$ 不是一个纠缠态，也就是说施密特数等于1，假设非零的 p_a 为 $a = 1$ 的 p_1 ，则由于 $\sum_a p_a = 1$ ，必有 $p_1 = 1$ ，由(12.17)和(12.18)的表达式可见，这时候必有

$$\rho_A = |1\rangle_A \langle 1|, \quad \rho_B = |1\rangle_B \langle 1|. \quad (12.19)$$

显然这是两个纯态密度算符。反过来，如果原来的 $|\psi\rangle$ 是一个纠缠态，施密特数大于1，那 ρ_A, ρ_B 就是混态密度算符。所以整个完整系统AB的纠缠态和子系统A或者B的混态是相对应的。

当然，我们可以完全不提B部分，直接定义一个系统的密度算符，而不管它是否源于系统与环境的纠缠。一个系统的密度算符 ρ 包含了所有我们能从对这个系统的观测中提取出来的信息，根据上面的讨论可以知道，**密度算符 ρ 应该满足如下性质**：1. 它是一个厄密算符。2. 它的本征值都大于等于0，为正定算符。3. $\text{Tr}(\rho) = 1$ 。有了密度算符以后，对任何物理可观测测量 \mathcal{O} 期望值的计算就可以表达为

$$\langle \mathcal{O} \rangle = \text{Tr}(\rho \mathcal{O}). \quad (12.20)$$

如果一个密度算符能够写成 $\rho = |\psi\rangle \langle \psi|$ 的形式，就称为纯态密度算符，它描述的系统状态就是纯态，否则就是混态密度算符，描述的状态就是混态。很显然，纯态密度算符额外满足

$$\rho^2 = \rho. \quad (12.21)$$

反过来，如果一个密度算符额外满足上式，则它的本征值必为0或1，由于密度算符所有本征值的和要等于1，因此它就必定能写成 $\rho = |\psi\rangle \langle \psi|$ 的形式($|\psi\rangle$ 为 ρ 的本征态)，从而必为纯态密度算符。对于纯态密度算符，我们必定有

$$\text{Tr}(\rho^2) = \text{Tr}(\rho) = 1. \quad (12.22)$$

对于一个任意的密度算符 ρ ，我们必可以通过求解它的本征方程将它对角化成如下形式

$$\rho = \sum_a p_a |a\rangle \langle a|. \quad (12.23)$$

密度算符的正定性告诉我们 $p_a \geq 0$, $\text{Tr}(\rho) = 1$ 的条件则告诉我们 $\sum_a p_a = 1$, 所以 $p_a \leq 1$, 等号仅当 ρ 为纯态密度算符时才成立。因此, 对于混态密度算符, 我们必有 $\text{Tr}(\rho^2) = \sum_a p_a^2 < \sum_a p_a = 1$, 即对于混态密度算符, 必有

$$\text{Tr}(\rho^2) < 1. \quad (12.24)$$

所有系统都可以看作某个更大的 AB 系统的 A 部分, 并且(12.23)的形式以及前面关于施密特分解的知识进一步告诉我们, 所有的密度算符都可以看成由这个更大系统的某个具有施密特分解(12.15)形式的纯态 $|\psi\rangle$ 通过忽略 B 部分来得到, (12.15)就称作密度算符(12.23)的纯化。

然而, 密度算符的概念其实并不需要通过它的纯化来理解。密度算符的(12.23)形式启发我们可以从系综的角度解释密度算符。也即是说, 我们可以认为密度算符描述的是大量系统的集合, 其中每一个系统以概率 p_a 处于 $|a\rangle$ 态。这是因为, 这一系综中任何物理量 \mathcal{O} 的期望值同样由下式给出

$$\langle \mathcal{O} \rangle = \sum_a p_a \langle a | \mathcal{O} | a \rangle = \text{Tr}(\rho \mathcal{O}). \quad (12.25)$$

密度算符的系综解释和纯化解释是有所不同的, 纯化解释可以用于单个系统, 而系综解释只适用于大量系统的系综。但是, 由于对任何物理量期望值的测量都需要多次重复实验, 这相当于引入了大量系统, 因此在实验上我们实际上无法区分密度算符的纯化解释和系综解释。

不过, 根据系综解释, 我们可以将密度算符中出现的概率 p_a 看成是源于我们对系统信息的缺失, 它造成了我们对系统的某种无知。在纯化解释中, 系统作为 A 部分与辅助的 B 纠缠成了一个整体, 在这个整体中 A 和 B 的“个性”都已经失去, 而 p_a 所反映的信息缺失就源于我们对 B 部分的忽略。总之, 密度算符描述中包含了对系统的某种无知, 这种无知来源于对系统的信息缺失, 而系综诠释的好处是, 它允许这种信息缺失和纠缠无关, 任何信息缺失都可以对应于一个系综, 从而给出一个密度算符。然而, 如何量化密度算符中包含的无知呢? 这个问题我们后面再来讨论。

12.1.3 Bloch 球

下面我们详细地讨论一下单个量子比特的密度算符 ρ 。这时候希尔伯特空间是2维的, 为了具体起见我们可以将基矢量 $\{|0\rangle, |1\rangle\}$ 表示成

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (12.26)$$

如此一来，密度算符 ρ 就应该表示成一个 2×2 的厄密矩阵，称为密度矩阵。但是，任何 2×2 的厄密矩阵都可以用 $\{1, \sigma_x, \sigma_y, \sigma_z\}$ 这四个厄密矩阵展开为 $\rho = \frac{1}{2}(w_0 \cdot 1 + \mathbf{w} \cdot \sigma)$ ，式中 $\{w_0, \mathbf{w}\}$ 均为实数。注意到泡利矩阵的迹为零，则由 $\text{Tr}(\rho) = 1$ 可知， $w_0 = 1$ ，从而单个量子比特的密度矩阵 ρ 必定可以表示成如下形式

$$\rho(\mathbf{w}) = \frac{1}{2}(1 + \mathbf{w} \cdot \sigma). \quad (12.27)$$

\mathbf{w} 称为量子比特的极化矢量，所以单量子比特的密度矩阵完全由其极化矢量刻画。另外，容易算出 $\text{Tr}(\rho^2) = \frac{1}{2}(1 + \mathbf{w}^2)$ ，由 $\text{Tr}(\rho^2) \leq 1$ 可知

$$\mathbf{w}^2 \leq 1, \quad (12.28)$$

等号当且仅当 $\rho(\mathbf{w})$ 为一个纯态密度矩阵时才成立。

由上面的讨论可知，单个量子比特的所有可能密度矩阵一一对应于3维极化矢量空间中的一个单位球体 $\mathbf{w}^2 \leq 1$ ，称为Bloch球，这个球体的表面对应于单量子比特的可能纯态，球内部分描写的当然就是混态。特别的，在Bloch球体中心， $\mathbf{w} = 0$ 处的那个混态，就是通常所谓的最大混态，很显然，它是

$$\rho(0) = \frac{1}{2} \cdot 1 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|). \quad (12.29)$$

之所以说 $\rho(\mathbf{w})$ 表达式中的 \mathbf{w} 是极化矢量，是因为，假设将单量子比特看成电子自旋， $|0\rangle$ 态看成自旋向上态， $|1\rangle$ 态看成自旋向下态，则我们可以计算 \mathbf{n} 方向泡利算符 $(\mathbf{n} \cdot \sigma)$ 的期望值，易得

$$\langle (\mathbf{n} \cdot \sigma) \rangle_{\mathbf{w}} = \text{Tr}((\mathbf{n} \cdot \sigma)\rho(\mathbf{w})) = \mathbf{n} \cdot \mathbf{w}. \quad (12.30)$$

这告诉我们， \mathbf{n} 方向自旋期望值是矢量 \mathbf{w} 在 \mathbf{n} 方向上的投影，可见 \mathbf{w} 正好代表自旋极化。

假设在单位球面上取一个单位矢量 $\mathbf{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ ， $\mathbf{n}^2 = 1$ ，则纯态密度矩阵就是 $\rho(\mathbf{n}) = \frac{1}{2}(1 + \mathbf{n} \cdot \sigma)$ 。假设记相应的纯态为 $|\psi(\mathbf{n})\rangle$ ，即 $\rho(\mathbf{n}) = |\psi(\mathbf{n})\rangle\langle \psi(\mathbf{n})|$ ，则由于 $(\mathbf{n} \cdot \sigma)\rho(\mathbf{n}) = \rho(\mathbf{n})(\mathbf{n} \cdot \sigma) = \rho(\mathbf{n})$ ，容易看出， $|\psi(\mathbf{n})\rangle$ 必为 $(\mathbf{n} \cdot \sigma)$ 的本征值为1的本征态。

利用3个泡利矩阵的表达式，并将单位矢量 \mathbf{n} 的分量形式代入 $\rho(\mathbf{n})$ ，容易得到

$$\rho(\mathbf{n}) = \begin{pmatrix} \cos^2(\theta/2) & \cos(\theta/2) \sin(\theta/2) e^{-i\phi} \\ \cos(\theta/2) \sin(\theta/2) e^{i\phi} & \sin^2(\theta/2) \end{pmatrix}. \quad (12.31)$$

由此可以得到

$$|\psi(\mathbf{n})\rangle = \begin{pmatrix} \cos(\theta/2)e^{-i\phi/2} \\ \sin(\theta/2)e^{i\phi/2} \end{pmatrix}. \quad (12.32)$$

假设有两个密度矩阵 $\rho(\mathbf{w}_1)$ 和 $\rho(\mathbf{w}_2)$, 则对于任意 $0 \leq t \leq 1$ 均有

$$t\rho(\mathbf{w}_1) + (1-t)\rho(\mathbf{w}_2) = \rho(t\mathbf{w}_1 + (1-t)\mathbf{w}_2), \quad (12.33)$$

也即是说, 两个密度矩阵按照 $t\rho(\mathbf{w}_1) + (1-t)\rho(\mathbf{w}_2)$ 的形式组合以后结果依然是一个密度矩阵, 而且这个密度矩阵的极化矢量刚好是 $t\mathbf{w}_1 + (1-t)\mathbf{w}_2$ 。实际上, 假设将 $0 \leq t \leq 1$ 看成可变参数, 那方程 $t\mathbf{w}_1 + (1-t)\mathbf{w}_2$ 描写的就是 \mathbf{w}_1 与 \mathbf{w}_2 的连线线段, 这条线段当然依然在3维空间的单位球体之内, 因此连线上的每一点当然都对应一个密度矩阵。

12.1.4 哥本哈根诠释和多世界诠释

在第二章的“量子比特与薛定谔猫”一节中, 我们谈到测量会导致态的坍缩, 这种说法其实只是对量子力学的一种诠释, 而对于量子测量过程发生了什么, 其实有多种诠释。现在, 让我们更细致地考察一下量子测量过程, 并介绍两种不同的主要诠释, 哥本哈根诠释和多世界诠释。

大致来说, 目前对量子测量过程的认识是下面这样的: 简言之, 测量过程就是一个被测对象和观察者(包括仪器)以及环境发生纠缠的过程。以测量一个电子的自旋 z 分量为例, 这时候我们可以把一个任意的电子自旋态 $|\psi\rangle$ 分解成自旋向上态 $|\uparrow\rangle$ 和自旋向下态 $|\downarrow\rangle$ 的叠加, 即 $|\psi\rangle = c_\uparrow|\uparrow\rangle + c_\downarrow|\downarrow\rangle$ 。假设测量之前电子处在 $|\psi\rangle$, 观察者(包括仪器)处在 $|m_0\rangle$ 态, 环境处在 $|e_0\rangle$ 态。那么测量之前整个世界的量子态 $|W_0\rangle$ 就是 $|W_0\rangle = |\psi\rangle|m_0\rangle|e_0\rangle = c_\uparrow|\uparrow\rangle|m_0\rangle|e_0\rangle + c_\downarrow|\downarrow\rangle|m_0\rangle|e_0\rangle$ 。

现在对电子的自旋 z 分量进行测量, 当测到电子自旋向上时, 我们记观察者处于 $|m_\uparrow\rangle$ 态, 记这时候环境的状态为 $|e_\uparrow\rangle$, 而当测到电子自旋向下时, 记观察者处于 $|m_\downarrow\rangle$ 态, 记这时候环境的状态为 $|e_\downarrow\rangle$ 。

而电子、观察者、环境这三者的整体也是一个量子系统, 因此在测量过程中这个整体的量子态必然是保持线性叠加并么正演化的, 因此测量以后, 世界的量子态将由测量前的 $|W_0\rangle$ 么正演化到 $|W\rangle$, 由于要保持量子态的线性叠加, $|W\rangle$ 必定可以写成 $|W\rangle = c_\uparrow|\uparrow\rangle|m_\uparrow\rangle|e_\uparrow\rangle + c_\downarrow|\downarrow\rangle|m_\downarrow\rangle|e_\downarrow\rangle$ 。很显然, 与测量前的 $|W_0\rangle$ 态不同, 在 $|W\rangle$ 态中, 电子、观察者、以及环境这三者纠缠成了一个整体。

请注意, $|W\rangle$ 态依然是一个确定的量子态, 因此这时候依然是不谈概率的。

但是, 简化一点来说, 在测量时我们关心的当然只是作为观测对象的电子, 以及仪器上的读数和观察者看到某个读数时的心情, 我们会完全忽视环境, 实际上我们也根本无法收集到环境的信息。而由于电子、观察者、以及环境纠缠成了一个整体, 因此当我们忽视环境的信息时, 电子和观察者的量子力学描述就不再是一个纯态了, 而是混态 $|c_\uparrow|^2 |\uparrow, m_\uparrow\rangle\langle\uparrow, m_\uparrow| + |c_\downarrow|^2 |\downarrow, m_\downarrow\rangle\langle\downarrow, m_\downarrow|$ 。也就是说, 原来的相干叠加消失了, 概率出现了。观察者有 $|c_\uparrow|^2$ 的概率处在 $|m_\uparrow\rangle$ 态并发现电子自旋向上, 有 $|c_\downarrow|^2$ 的概率处在 $|m_\downarrow\rangle$ 态并发现电子自旋向下。这个过程就是我们前面说过的退相干。很显然, 退相干是因为量子信息泄露到了环境中去而产生的。

到此为止都还没有涉及到量子态的坍缩, 量子力学的不同诠释对上面的内容基本上还是有共识的。麻烦就出在, 一旦我们在某次测量中观察到电子自旋向上, 在我们看来整个世界就应该处在 $|W_\uparrow\rangle = |\uparrow\rangle|m_\uparrow\rangle|e_\uparrow\rangle$ 态, 相反, 如果在某次测量中我们观察到电子自旋向下, 那在我们看来, 这时候整个世界就应该处在 $|W_\downarrow\rangle = |\downarrow\rangle|m_\downarrow\rangle|e_\downarrow\rangle$ 态。

这时候就有不同的观点了, 哥本哈根诠释的观点是, 一次测量只发生了一个可能的结果, 因此量子态坍缩了, 当我们观察到电子自旋向上时, 世界的量子态就从 $|W\rangle$ 坍缩到了 $|W_\uparrow\rangle$, 而当我们观察到电子自旋向下时, 世界的量子态就从 $|W\rangle$ 坍缩到了 $|W_\downarrow\rangle$ 。至于坍缩过程是如何进行的, 目前物理学家们依然在研究之中。

这种诠释的问题在于, 我们可以设想有一个外在于整个世界的“朋友”, 他根本不关心我们的电子自旋是向上还是向下, 他根本就没有进行观测, 因此对于他来说, 世界的量子态应该依然是 $|W\rangle$, 根本就没有什么量子态的坍缩。

另一种不同的诠释通常人们叫做多世界诠释。这种诠释的观点是, 一次测量, 所有可能的结果都发生了, 但是, 世界本身分裂了, 不同的可能结果发生在分裂以后的不同平行世界里。也就是说, 在一个世界里你测到电子自旋向上, 而在另一个世界里你测到电子自旋向下。但是, 不同的平行世界只是同一个量子现实的不同分支, 而整个量子现实依然是由相干叠加的 $|W\rangle = c_\uparrow|W_\uparrow\rangle + c_\downarrow|W_\downarrow\rangle$ 态来描写的。根据这种多世界诠释, 从来就没有什么量子态坍缩, 有的不过是世界分裂。

这种诠释的问题是, 分裂以后的两个平行世界怎么能够保持相干叠加

从而共同构成同一个量子现实呢?

12.2 再谈密度算符与系综

这一节我们主要是推广系综的概念。

密度算符集合的凸性

首先我们定义一下什么是向量空间的凸子集。对于向量空间的某个子集，如果它满足集合中任意两个点的连线段完全处于集合内部，那么这个集合就是向量空间的一个凸子集。比如图(12.1)所示的多边形就是两维平面的凸子集。如果凸子集中的某点永远不能处在子集内连线段内，而只能处于连线段的端点，那这个点就称为极 endpoint。比如图(12.1)中的 A, B, C, D, E 点就是极 endpoint。

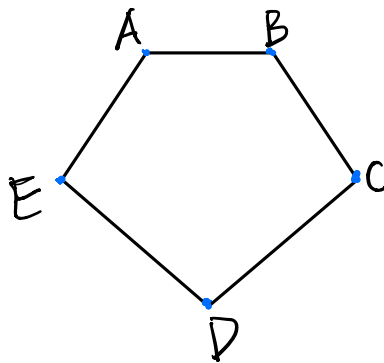


图 12.1: 两维实向量空间的一个凸子集， A, B, C, D, E 为极 endpoint。

另外，由于任何两个厄密算符的实系数线性组合依然是厄密算符。所以一个量子系统所有厄密算符的集合构成了一个实向量空间，系统每一个可能的厄密算符都对应这个实向量空间里的一个向量。

之所以讨论以上两个概念，是因为，一个量子系统的所有可能密度算符的集合构成了厄密算符向量空间的一个凸子集。具体来说，假设 $\rho_1, \rho_2, \dots, \rho_N$ 是系统 N 个可能的密度算符，则下式给出的 ρ 也必定是系统可能的密度算符

$$\rho = p_1\rho_1 + p_2\rho_2 + \dots + p_N\rho_N, \quad (12.34)$$

式中, $0 \leq p_i \leq 1$, 且 $\sum_{i=1}^N p_i = 1$ 。要证明 ρ 是密度算符, 只需验证它满足密度算符必须满足的3条性质。根据所给条件, ρ 的厄密性以及 $\text{Tr}(\rho) = 1$ 这两条都很明显, 唯一需要说明的是 ρ 的正定性。对于一个厄密算符的正定性, 我们有两种等价的判定方法, 1. 所有本征值大于等于0; 2. 这个算符在任何量子态 $|\psi\rangle$ 上的期望值都大于等于0。根据第2种判定正定性的方法, (12.34)式给出来的厄密算符 ρ 的正定性也是显然的(注意我们已知每一个 ρ_i 都正定)。

上一段我们刻画了密度算符集合的凸性, 实际上更符合凸子集定义的刻画方法是: 假设 ρ_1, ρ_2 为系统两个可能的密度算符, 则这两者的连线段 $\rho(t) = t\rho_1 + (1-t)\rho_2$ 上每一个密度算符 $\rho(t)$ 也必为系统可能的密度算符, 这里 $0 \leq t \leq 1$ 。这种刻画凸性的方式不仅是上一段给出来的刻画方式的特例, 而且从它出发通过使用数学归纳法我们也能反过来导出上一段的刻画方式。所以对凸性的这两种刻画方式实际上是等价的。

不仅如此, 我们还可以证明, 纯态密度矩阵相应于这个密度算符凸子集的极 endpoint。为此, 我们假设 $\rho = |\psi\rangle\langle\psi|$, 我们用反证法, 设存在密度算符 ρ_1 和 ρ_2 , 使得 $\rho = t\rho_1 + (1-t)\rho_2$, $0 < t < 1$ 。任意取一个与 $|\psi\rangle$ 正交的 $|\psi_\perp\rangle$ 态, 它满足 $\langle\psi|\psi_\perp\rangle = 0$, 则有 $0 = \langle\psi_\perp|\rho|\psi_\perp\rangle = t\langle\psi_\perp|\rho_1|\psi_\perp\rangle + (1-t)\langle\psi_\perp|\rho_2|\psi_\perp\rangle$ 。进一步根据密度算符的正定性, 必有 $\langle\psi_\perp|\rho_1|\psi_\perp\rangle = \langle\psi_\perp|\rho_2|\psi_\perp\rangle = 0$ 。但是 $|\psi_\perp\rangle$ 是任意与 $|\psi\rangle$ 正交的态, 因此刚才的结果就意味着 $\rho_1 = \rho_2 = |\psi\rangle\langle\psi| = \rho$ 。这就证明了纯态必为极 endpoint。反过来, 对于混态密度算符 ρ , 由于总有 $\rho = \sum_a p_a |a\rangle\langle a|$, 式中 $|a\rangle\langle a| = \rho_a$ 均为纯态密度算符, 这就说明混态密度算符必定不是极 endpoint。这就完成了密度算符凸子集的极 endpoint 与纯态一一对应的证明。

比方说对前面讨论过的单量子比特密度算符 Bloch 球体的例子, 整个 Bloch 球体就是一个凸子集, 球体表面的每一个点都是这个凸子集的极 endpoint, 因此必定相应于纯态。这正好是我们前面已经知道的结论。不过, Bloch 球体是一个特例, 它的边界点和极 endpoint 是一回事, 但是, 图(12.1)告诉我们, 对于一般的凸子集, 它的极 endpoint 必定是边界点, 但是反过来, 它的边界点却不一定是极 endpoint。具有3维以上希尔伯特空间的量子系统的密度算符凸子集也是这样, 它的边界点并不都是极 endpoint。

系综概念的推广

根据凸性, 假设某个密度算符 ρ 可以写成 $\rho = \sum_i p_i \rho_i$, $\sum_i p_i = 1$ 。则很

显然

$$\langle \mathcal{O} \rangle = \text{Tr}(\rho \mathcal{O}) = \sum_i p_i \text{Tr}(\rho_i \mathcal{O}) = \sum_i p_i \langle \mathcal{O} \rangle_i. \quad (12.35)$$

这就意味着，对于一个量子系统，假设我们能够制备出所有的 ρ_i 态，则我们就可以制备一个系综，让这个系综中的每个系统以 p_i 的概率处在 ρ_i 态，而这个系综就可以由密度算符 $\rho = \sum_i p_i \rho_i$ 描述。密度算符 ρ 既可以看成是描述一个系统的量子态，也可以看成是描述刚才所说的系综，等式(12.35)告诉我们，在物理上我们并不能区分这两者。

很明显，除了纯态以外，任何一个密度算符 ρ 都有多种办法分解成 $\rho = \sum_i p_i \rho_i$, $\sum_i p_i = 1$ 。因此与密度算符 ρ 所对应的系综远不是唯一的，而是有无穷多的可能性。比方说，对于单量子比特Bloch球体内部的任何一点(对应密度算符 ρ)，我们都可以作无数条过这一点的线段与球面上的两点(对应于两个纯态密度算符)相交，每一根这样的线段都意味着我们可以将密度算符 ρ 表示成球面上的这两个纯态密度算符的线性组合，而 ρ 就可以看成是描述了相应的纯态系综，由于这样的线段有无数条，所以 ρ 可以描述无数多个不同的纯态系综，在物理上这些纯态系综不可区分。比方说，对于Bloch球心的那个最大混态密度算符 $\rho = \frac{1}{2} \cdot 1$ ，我们就有无数种表达它的方式，比如下面的两种

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|. \quad (12.36)$$

12.3 量子熵

12.3.1 冯诺伊曼熵

前面我们说过，混态密度算符可以写成 $\rho = \sum_a p_a |a\rangle\langle a|$ 的形式， p_a 可以看成概率，是一种不确定性，它反映了我们对系统的某种无知。然而如何衡量这种无知呢？

一个自然的想法是将这种无知定义成某个算符的期望值，不妨称这个算符为“无知算符”，下面我们就是要找到它的表达式。由于我们的无知反映在系统的密度算符 ρ 中，所以这个“无知算符”应该依赖于 ρ (而通常的物理量算符是与态 ρ 无关的)。而且，如果 ρ 是一个纯态密度算符，那它就没有任何不确定性，从而相应的“无知算符”应该为零。

另外，从直观上我们知道，如果两个系统 S_1 和 S_2 相互独立，那我们对它们的无知应该是相加的，或者说相应的“无知算符”应该是相加的。另一方面，假设我们记 S_1 的密度算符为 ρ_{S_1} ，记 S_2 的密度算符为 ρ_{S_2} ，则由于这两个系统相互独立，所以两者整体的密度算符应该是 ρ_{S_1} 和 ρ_{S_2} 的乘积，不过， ρ_{S_1} 只作用在 S_1 的希尔伯特空间 \mathcal{H}_{S_1} 上， ρ_{S_2} 只作用在 S_2 的希尔伯特空间 \mathcal{H}_{S_2} 上，为了强调这一点我们常将这个乘积写成如下形式

$$\rho_{S_1}\rho_{S_2} = \rho_{S_1} \otimes \rho_{S_2}, \quad (12.37)$$

它作用在两者整体的希尔伯特空间 $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2}$ 上。为了将密度算符的这种相乘关系转化为相加关系，我们定义与密度算符 ρ 相应的“无知算符”如下

$$-\log(\rho). \quad (12.38)$$

额外加上的负号是为了使得这个算符为一个正定算符²。显然，这样定义的“无知算符”的确是相加的，满足 $-\log(\rho_{S_1}\rho_{S_2}) = -\log(\rho_{S_1}) - \log(\rho_{S_2})$ 。

很容易验证，对于纯态密度算符的确有 $-\log(\rho) = 0$ 。这是因为，纯态密度算符满足 $\rho^2 = \rho$ ，从而 ρ 的任何幂次都依然等于 ρ 。从而对于任何一个幂级数函数 $f(x)$ ，必有 $f(\rho) = f(1)\rho$ 。现在我们按照 $-\log(x) = -\log(1+x-1)$ 关于 $x-1$ 的泰勒展开级数来定义算符函数 $-\log(\rho)$ ，从而当然就有 $-\log(\rho) = -\log(1)\rho = 0$ 。

下面我们可以将密度算符 ρ 中包含的无知定义为“无知算符”的期望值，称作冯诺伊曼熵，简称熵，记为 $S(\rho)$ ，则

$$S(\rho) = \langle -\log \rho \rangle = -\text{Tr}(\rho \log \rho). \quad (12.39)$$

显然， $S(\rho)$ 也是所有 $-\rho \log \rho$ 本征值之和，从而也有

$$S(\rho) = -\sum_a p_a \log(p_a). \quad (12.40)$$

这正好是香农关于经典概率分布 $\{p_a\}$ 信息熵的表达式，经典信息熵衡量的是获取信息以后能够消除的不确定性，换言之，信息熵衡量的正是获取具体信息之前我们对它的无知，所以冯诺伊曼熵 $S(\rho)$ 的确衡量的是无知。很明显，

$$S(\rho) \geq 0, \quad (12.41)$$

²因为 ρ 的本征值 $0 \leq p_a \leq 1$ ，所以恒有 $-\log(\rho)$ 的本征值 $-\log(p_a) \geq 0$ 。

等于号仅当 ρ 为纯态密度算符时才能取到。

另外, 假设系统的希尔伯特空间为 D 维, 即(12.40)式中的指标 $a = 1, 2, 3, \dots, D$ 。则由(12.40)式可以证明,

$$S(\rho) \leq \log D. \quad (12.42)$$

等于号当且仅当 $p_a = 1/D$ 时成立, 这时候的密度算符为

$$\rho = \frac{1}{D} \sum_a |a\rangle\langle a| = \frac{1}{D} \cdot 1. \quad (12.43)$$

这当然是一个混态密度算符, 相应的混态称之为最大混态。

对于一个由 A 和 B 两部分组成的封闭系统, 它的整体(我们记作系统 AB)必定处于一个纯态, 我们可以对它进行施密特分解, 进而能够导出前面的(12.17)式和(12.18)式, 即 A 和 B 两个子部分的密度算符 ρ_A, ρ_B 分别为

$$\rho_A = \sum_a p_a |a\rangle_A \langle a|, \quad \rho_B = \sum_a p_a |a\rangle_B \langle a|. \quad (12.44)$$

从而必有

$$S(\rho_A) = - \sum_a p_a \log(p_a) = S(\rho_B). \quad (12.45)$$

即两个通过纠缠形成一个整体纯态的 A, B 系统, 其两个子部分的冯诺伊曼熵是相等的。正因为这个美妙的性质, 人们常常将这种情况下子部分的冯诺伊曼熵看成是对两体纠缠度的一种衡量。

假设 A, B 两部分的希尔伯特空间维数相等, 均为 D 。那前面关于最大混态的讨论告诉我们, 两体纠缠度最大为 $\log D$, 这时候 $p_a = \frac{1}{D}$, 从而根据施密特分解, AB 整体的量子态 $|\psi\rangle$ 为

$$|\psi\rangle = \frac{1}{\sqrt{D}} \sum_a |a\rangle_A |a\rangle_B. \quad (12.46)$$

由于纠缠度最大, 这被称为最大纠缠态。

凹性

前面我们说过, 一个系统所有可能密度算符的集合是一个凸集, 因此冯诺伊曼熵 $S(\rho)$ 就是这个凸集上的函数, 实际上, 它是一个凹函数。为了

证明这一点, 任取两个不同的密度算符 ρ_1, ρ_2 , 然后构造 $\rho(t) = t\rho_1 + (1-t)\rho_2, 0 \leq t \leq 1$, 我们要证明的是 $\frac{d^2}{dt^2}S(\rho(t)) \leq 0$ 。为此, 我们首先注意到

$$\frac{d}{dt}S(\rho(t)) = -\text{Tr}(\dot{\rho} \log \rho). \quad (12.47)$$

又注意到

$$\log \rho = \int_0^\infty ds \left(\frac{1}{s+1} - \frac{1}{s+\rho(t)} \right). \quad (12.48)$$

再注意到 $\dot{\rho}(t) = 0$, 则有

$$\frac{d^2}{dt^2}S(\rho(t)) = - \int_0^\infty ds \text{Tr} \left(\dot{\rho} \frac{1}{s+\rho(t)} \dot{\rho} \frac{1}{s+\rho(t)} \right). \quad (12.49)$$

注意这个积分式的被积函数具有 $\text{Tr}(A^2)$ 的形式, 式中 $A = (s+\rho(t))^{-1/2} \dot{\rho} (s+\rho(t))^{-1/2}$ 为厄密算符, 所以 A^2 必定是正定算符, 从而 $\text{Tr}(A^2) \geq 0$, 从而即有 $\frac{d^2}{dt^2}S(\rho(t)) \leq 0$ 。

由凹函数的熟知性质, 我们有

$$\sum_i p_i S(\rho_i) \leq S\left(\sum_i p_i \rho_i\right). \quad (12.50)$$

式中 $0 \leq p_i \leq 1$, 且 $\sum_i p_i = 1$ 。即是说, 将多个密度算符“混”起来会增加冯诺伊曼熵。这个熵的增量通常称作Holevo 信息, 记作 χ

$$\chi = S\left(\sum_i p_i \rho_i\right) - \sum_i p_i S(\rho_i). \quad (12.51)$$

12.3.2 相对熵

相对熵和互信息

假设一个量子系统实际由密度算符 ρ 描写, 但是我们不知道, 我们猜测了一个描写它的密度算符 σ , 那我们这个猜测中额外包含的无知就应该是

$$S(\rho \parallel \sigma) = \langle -\log \sigma \rangle - \langle -\log \rho \rangle = \text{Tr}(\rho \log \rho - \rho \log \sigma). \quad (12.52)$$

式中 $\langle -\log \sigma \rangle = -\text{Tr} \rho \log \sigma$ 为我们的猜测中总共包含的无知, $\langle -\log \rho \rangle$ 为系统密度算符中原本包含的信息缺失, $S(\rho \parallel \sigma)$ 也称作 σ 和 ρ 的量子相对熵。

可以想见, 量子相对熵 $S(\rho \parallel \sigma)$ 这种额外的无知一定大于等于零(等号当且仅当 $\sigma = \rho$ 时才成立), 即

$$S(\rho \parallel \sigma) \geq 0, \quad (12.53)$$

事实也的确如此, 证明并不复杂, 不过证明过程和我们将要进行的讨论关系不大, 因此我们推荐感兴趣的读者去阅读文献³。

特别的, 假设我们考虑一个由 A 和 B 两个子系统组成的复合系统 AB , 记 AB 的密度算符为 ρ_{AB} , 记 $\rho_A = \text{Tr}_B(\rho_{AB})$ 为子系统 A 的约化密度算符, $\rho_B = \text{Tr}_A(\rho_{AB})$ 为子系统 B 的约化密度算符。当然, 一般来说, 由于相互作用, A 和 B 之间会存在关联, 因此 $\rho_{AB} \neq \rho_A \otimes \rho_B$ 。但是假设我们误以为 A 和 B 没有关联, 并将 $\rho_A \otimes \rho_B$ 猜测为 AB 的密度算符 σ_{AB}

$$\sigma_{AB} = \rho_A \otimes \rho_B. \quad (12.54)$$

则我们这一猜测的额外无知为

$$\begin{aligned} S(\rho_{AB} \parallel \sigma_{AB}) &= \text{Tr}_{AB}(\rho_{AB} \log \rho_{AB} - \rho_{AB} \log \sigma_{AB}) \\ &= \text{Tr}_{AB}(\rho_{AB} \log \rho_{AB} - \rho_{AB} \log \rho_A - \rho_{AB} \log \rho_B) \\ &= S_A + S_B - S_{AB}. \end{aligned} \quad (12.55)$$

式中 $S_A = S(\rho_A)$ 为 A 的熵, $S_B = S(\rho_B)$ 为 B 的熵, 而 $S_{AB} = S(\rho_{AB})$ 为 AB 的熵。类似这样的记号我们后面还会用, 将不再进行说明。人们通常称 $S_A + S_B - S_{AB}$ 为 A 和 B 的量子互信息, 记为

$$I(A, B) = S_A + S_B - S_{AB}. \quad (12.56)$$

则 $S(\rho_{AB} \parallel \sigma_{AB}) \geq 0$ 就意味着量子互信息总是大于等于零的, 即

$$I(A, B) = S_A + S_B - S_{AB} \geq 0. \quad (12.57)$$

不过, 值得说明的是, 量子互信息这个概念完全是类比于经典信息论中的互信息概念而引入的, 但它其实并没有经典互信息的内涵。正因为如此我们这里才没有介绍经典互信息的具体含义。另外, 不等式(12.57)有时候也称之为熵的次可加性(subadditivity)。

³A Mini-Introduction To Information Theory, Edward Witten, arXiv:1805.11965[hep-th]

相对熵的单调性

令人吃惊的是, Lieb 和Ruskai还证明量子相对熵具有一种强大的单调性, 不过这个证明很难, 因此我们这里不会讨论, 感兴趣的读者可以阅读原始论文⁴, 或者查阅量子信息方面的专业书籍。

为了说清楚什么是量子相对熵的单调性, 我们还是考察一个由A,B两部分组成的系统, 记为系统AB, 假设它有两个密度算符 ρ_{AB} 和 σ_{AB} , 记 $\rho_A = \text{Tr}_B \rho_{AB}$, $\sigma_A = \text{Tr}_B \sigma_{AB}$ 为相应的两个在A上的约化密度算符。量子相对熵的单调性说的是

$$S(\rho_{AB} \parallel \sigma_{AB}) \geq S(\rho_A \parallel \sigma_A). \quad (12.58)$$

即对一个系统的一部分求迹(称作部分迹)会降低量子相对熵。

不妨让我们给出对(12.58)式的一个直观理解(不是证明!), 其实这个式子说的无非是, 当我们将一个大系统(这里即AB系统)密度算符的猜测用于它的一部分(这里即是A部分)时, 我们的额外无知将单调减少。直观上这当然很好理解, 难的是证明。

下面我们将量子相对熵的这种单调性用于一个三部分的系统ABC, 设其密度算符为 ρ_{ABC} , 当然我们有约化密度算符 $\rho_A = \text{Tr}_{BC} \rho_{ABC}$, $\rho_{BC} = \text{Tr}_A \rho_{ABC}$ 。我们再引入第二个密度算符

$$\sigma_{ABC} = \rho_A \otimes \rho_{BC}. \quad (12.59)$$

这样我们就可以计算 $S(\rho_{ABC} \parallel \sigma_{ABC})$ 了,

$$S(\rho_{ABC} \parallel \sigma_{ABC}) = S(\rho_{ABC} \parallel \rho_A \otimes \rho_{BC}) = I(A, BC) = S_A + S_{BC} - S_{AB} \quad (12.60)$$

为了应用单调性, 现在我们对C求迹, 从而有

$$\rho_{AB} = \text{Tr}_C \rho_{ABC}, \quad \sigma_{AB} = \text{Tr}_C \sigma_{ABC} = \rho_A \otimes \rho_B. \quad (12.61)$$

因此我们又可以计算求完迹以后的 $S(\rho_{AB} \parallel \sigma_{AB})$,

$$S(\rho_{AB} \parallel \sigma_{AB}) = S(\rho_{AB} \parallel \rho_A \otimes \rho_B) = I(A, B) = S_A + S_B - S_{AB}. \quad (12.62)$$

量子相对熵的单调性告诉我们

$$S(\rho_{ABC} \parallel \sigma_{ABC}) \geq S(\rho_{AB} \parallel \sigma_{AB}). \quad (12.63)$$

⁴E. H. Lieb and M. B. Ruskai, "Proof Of The Strong Subadditivity Of Quantum Mechanical Entropy," J. Math. Phys. 14 (1973) 1938.

从而我们可以得到

$$S_{AB} + S_{BC} \geq S_B + S_{ABC}. \quad (12.64)$$

这个不等式称作熵的强次可加性(strong subadditivity)。

强次可加性有一些有趣的等价形式, 比方说, 假设我们对 ABC 系统引入一个辅助的 D 部分, 使得密度算符 ρ_{ABC} 在这个 $ABCD$ 系统中被纯化, 则利用我们前面证明过的美妙性质(12.45), 我们将有 $S_{AB} = S_{CD}$ (将 $ABCD$ 看成是由 AB 和 CD 两部分组成), 以及 $S_{ABC} = S_D$ (将 $ABCD$ 看成是由 ABC 和 D 两部分组成), 则不等式(12.64)可以重写为

$$S_{CD} + S_{BC} \geq S_B + S_D. \quad (12.65)$$

习惯上人们会引入 $S(C|D) = S_{CD} - S_D$, 以及 $S(C|B) = S_{BC} - S_B$ 这两个量, 称作条件熵, 则上面这个不等式又可以重写为

$$S(C|D) + S(C|B) \geq 0. \quad (12.66)$$

这个式子告诉我们 $S(C|D)$ 和 $S(C|B)$ 这两个条件熵不能同时为负。

但是量子条件熵的确是可以取负值的, 比方说, 如果 AB 处于最大纠缠态(当然是纯态), 则 $S_{AB} = 0$, 而 S_B 达到最大, 那这时候条件熵 $S(A|B) = S_{AB} - S_B$ 当然就为负! 因此(12.66)就告诉我们, C 不能与 D 和 B 两者同时处于最大纠缠态, 这称为量子纠缠的单配性(monogamy of entanglement)。

12.4 开放系统以及热力学第二定律

12.4.1 开放系统的时间演化

我们知道, 一个封闭量子系统随时间是么正演化的, 纯态 $|\psi\rangle$ 将演化为 $|\psi\rangle \rightarrow U|\psi\rangle$, 式中 U 为么正的时间演化算符。从而密度算符 ρ 将演化为

$$\rho \rightarrow U\rho U^\dagger. \quad (12.67)$$

么正演化意味着概率和信息的守恒。但是, 对于一个与环境相互作用的开放系统而言, 其信息可以不断泄露到环境中去, 从而其演化就不可能是么正演化。那么, 该怎么讨论开放系统的时间演化呢?

事实上, 这时候我们可以把环境包括进来, 将我们关心的系统看成 A 系统, 而将环境看成 B 系统, 两者共同构成一个封闭的 AB 系统。下面

当我们单独说到系统时，指的都是我们关心的A系统，而对于AB整体，我们将称之为复合系统。

假设环境最初处于 $|e_0\rangle$ 态，而我们关心的系统最初处于 $|\psi\rangle$ 态，两者没有纠缠。由于整个复合系统是封闭的，所以复合系统将满足么正演化，

$$|e_0\rangle|\psi\rangle \rightarrow U(|e_0\rangle|\psi\rangle). \quad (12.68)$$

下面我们取环境希尔伯特空间的一组正交归一矢量基 $\{|e_s\rangle\}$ ，然后将复合系统的量子态 $U(|e_0\rangle|\psi\rangle)$ 按这组基展开成

$$U(|e_0\rangle|\psi\rangle) = \sum_s |e_s\rangle|\psi_s\rangle. \quad (12.69)$$

注意，我们并没有说 $|\psi_s\rangle$ 正交归一。下面我们引入一组作用在A系统上的算符 E_s ，定义为 $|\psi_s\rangle = E_s|\psi\rangle$ ，则

$$|e_0\rangle|\psi\rangle \rightarrow U(|e_0\rangle|\psi\rangle) = \sum_s |e_s\rangle E_s|\psi\rangle. \quad (12.70)$$

利用 U 的么正性，我们可以得到

$$\sum_s E_s^\dagger E_s = 1. \quad (12.71)$$

所以开放系统从初态到 t 时刻态的映射完全由这组算符 E_s 决定。

将上面的推导照搬到密度算符上。假设系统初始的密度算符为 ρ ，环境初始的密度算符为 $|e_0\rangle\langle e_0|$ ，则整个复合系统初始的密度算符为这两者的乘积，不过，为了强调 ρ 只作用在系统的希尔伯特空间，而 $|e_0\rangle\langle e_0|$ 只作用在环境的希尔伯特空间，我们常常将这个乘积写成 $|e_0\rangle\langle e_0| \otimes \rho$ ，并简记为 $\hat{\rho}$ 。则类似于上面的推导我们将有

$$\hat{\rho} \rightarrow \hat{\rho}' = U\hat{\rho}U^\dagger = \sum_{s,s'} |e_s\rangle\langle e_{s'}| \otimes E_s\rho E_{s'}^\dagger. \quad (12.72)$$

由于我们并不关心环境，所以我们忽略它，即对环境的希尔伯特空间求迹，从而得到

$$\rho \rightarrow \text{Tr}_B(\hat{\rho}') = \sum_s E_s\rho E_s^\dagger. \quad (12.73)$$

我们常常把这个方程写作

$$\rho \rightarrow \mathcal{E}(\rho) = \sum_s E_s\rho E_s^\dagger, \quad \sum_s E_s^\dagger E_s = 1. \quad (12.74)$$

式中定义的映射 \mathcal{E} 称作量子映射(quantum map), 可以验证它将把密度算符映射到密度算符, 比方说很容易看出 $\text{Tr}(\mathcal{E}(\rho)) = \text{Tr}(\rho) = 1$ 。反过来可以证明, 任何将密度算符映射到密度算符的线性映射, 只要满足一些自然的条件, 就必定可以表示成(12.74)的形式, 这就是所谓的Kraus表示定理, (12.74)式中的 E_s 就称作Kraus算符。(12.74)式这样的量子映射有时候又称作保迹全正映射(trace preserving, completely positive (TPCP) map)。很显然, 封闭系统的时间演化只是(12.74)式的一个特例, 这时候只有一个非零Kraus算符, 即么正演化算符 U 。

但是, 一般来说, (12.74)式的量子映射只是把初始时的密度算符 ρ 映射到 t 时刻的密度算符 $\mathcal{E}(\rho)$, 它却不能用于考察任意 t 时刻的密度算符如何继续随着时间演化, 这是因为, 在推导(12.74)式时我们有一个重要的假定, 即最初的系统和环境没有关联, 两者的密度算符是直接乘积的关系。但这个条件却不适用于任意 t 时刻, 因为随着系统与环境的相互作用, 一般来说 t 时刻系统和环境一定是纠缠和关联的。正因为如此, (12.74)式其实不能真正看作开放系统的时间演化方程。

刘维尔方程和Lindblad 方程

为了真正得到一个开放系统的时间演化方程, 我们现在假定环境是巨大的, 就好像统计物理中的恒温热源那样, 以至于系统和环境的任何关联都会在某个非常短的微观时间尺度 τ 之内丧失。假定我们关心的是系统在远大于微观时间 τ 的宏观时间尺度上的演化。这时候我们就可以推导出一个真正的时间演化方程。

根据上一段的假设, 我们来考察密度算符在一个宏观无穷小时间间隔 δt 上的演化, 当然, 作为一个宏观时间, 我们依然有 $\delta t \gg \tau$ 。按照假设, 我们可以对任意初始时刻 t 利用(12.74)式, 从而有相应的演化方程

$$\rho(t + \delta t) = \sum_s E_s(t + \delta t, t) \rho(t) E_s^\dagger(t + \delta t, t). \quad (12.75)$$

对于么正演化的封闭系统(它当然与环境没有关联), 我们只有一个非零Kraus算符, 即为时间演化算符, 记作

$$E_0(t + \delta t, t) = U(t + \delta t, t) = 1 - iH\delta t/\hbar + \dots \quad (12.76)$$

代入(12.75)式, 即得到标准的刘维尔方程

$$\hbar \frac{d\rho}{dt} = -i[H, \rho]. \quad (12.77)$$

对于与环境相互作用的开放系统，这时候就可能有许多非零Kraus算符。但是可以设想，对于一个无穷小的 δt ， $E_0(t + \delta t, t)$ 将依然是主要的，按照 δt 进行泰勒展开依然有 $E_0(t + \delta t, t) = 1 + O(\delta t)$ ，不过当然， $E_0(t + \delta t, t)$ 不再么正，从而我们应该假设

$$E_0(t + \delta t, t) = 1 - i\left(H - i\frac{\Gamma}{2}\right)\delta t/\hbar + \dots \quad (12.78)$$

式中 Γ 也是一个厄密算符。 E_0 的这一假设形式显然是受前面章节对不稳定态研究的启发。

其余Kraus算符都是次要的，我们假设它们 $E_m(t + \delta t, t) = O(\sqrt{\delta t})$ ， $m = 1, 2, 3, \dots$ 。具体来说，我们假设

$$E_m(t + \delta t, t) = L_m\sqrt{\delta t}/\sqrt{\hbar} + \dots, m = 1, 2, 3, \dots \quad (12.79)$$

由于Kraus算符要满足 $\sum_s E_s^\dagger E_s = 1$ ，从而我们有

$$-\Gamma + \sum_{m=1} L_m^\dagger L_m = 0 \Rightarrow \Gamma = \sum_{m=1} L_m^\dagger L_m. \quad (12.80)$$

将Kraus算符的这些展开形式代入时间演化方程(12.75)就有

$$\hbar \frac{d\rho}{dt} = -i[H, \rho] - \frac{1}{2}(\Gamma\rho(t) + \rho(t)\Gamma) + \sum_{m=1} L_m\rho(t)L_m^\dagger. \quad (12.81)$$

这就是开放系统密度算符的时间演化方程，称作Lindblad方程，它可以看作是量子版本的Fokker-Planck方程，其中的算符 L_m 称作Lindblad算符，它描写系统与环境的相互作用。

12.4.2 热力学第二定律

前面冯诺伊曼熵和量子相对熵的公式最后都是一个求迹，因此它们在么正变换下都将保持不变，以量子相对熵为例，在么正变换 $\rho \rightarrow \rho' = U\rho U^\dagger$ ， $\sigma \rightarrow \sigma' = U\sigma U^\dagger$ 的作用下，将有

$$S(\rho' \parallel \sigma') = S(\rho \parallel \sigma). \quad (12.82)$$

特别的，一个封闭量子系统的时间演化就是一个么正量子变换，因此，一个封闭量子系统的冯诺伊曼熵和量子相对熵在时间演化下都将保持不变。

上面结论好像和我们熟悉的热力学第二定律相矛盾，其实不然，因为冯诺伊曼熵是所谓的精细熵，而通常的热力学熵是一种粗粒化熵，孤立系统的精细熵保持不变，但是其粗粒化熵是增长的，这就是热力学第二定律。这一节我们就是要对与一个大的恒温热源密切接触的量子系统证明热力学第二定律。由于与恒温热源之间的能量和信息交换，这时候的量子系统当然不是孤立系统，甚至也不是封闭系统，在统计物理中我们知道，这时候的热力学第二定律其实表现为系统自由能随着时间的不断降低，这一节就是要证明这个结论。

由于我们要处理的不是封闭系统，所以需要用到量子映射。我们将要用到的关键结果就是，量子相对熵在量子映射下单调降低。具体来说，假设 ρ 和 σ 为系统的两个密度算符，在量子映射 \mathcal{E} 之下，它们变换为

$$\rho \rightarrow \rho' = \sum_s E_s \rho E_s^\dagger, \quad \sigma \rightarrow \sigma' = \sum_s E_s \sigma E_s^\dagger. \quad (12.83)$$

则我们有

$$S(\rho \parallel \sigma) \geq S(\rho' \parallel \sigma'). \quad (12.84)$$

为了证明这一点，我们回顾一下前面对量子映射的推导过程，首先我们需要将密度算符与环境的 $|e_0\rangle\langle e_0|$ 乘起来，由于 $|e_0\rangle\langle e_0|$ 是一个纯态，熵是零，所以这一步对相对熵 $S(\rho \parallel \sigma)$ 没有影响，第二步是对整个复合系统用么正演化，根据前面所说，这一步也不会改变相对熵，最后一步是对环境求部分迹，根据量子相对熵的单调性，这一步将使得量子相对熵单调减少，因此综合起来我们就有(12.84)式。

下面我们取 σ 为与恒温热源达到热平衡时的密度算符，根据统计物理知识，

$$\sigma = \frac{1}{Z} e^{-\beta H}, \quad (12.85)$$

式中 $\beta = 1/T$ ， T 为恒温热源的温度，式中配分函数 Z 也可以写

$$Z = e^{-\beta F(\sigma)}, \quad (12.86)$$

$F(\sigma)$ 为这个热平衡态的自由能。

假设 ρ 为这个与恒温热源密切接触的系统任意时刻的密度算符。我们可

以计算量子相对熵 $S(\rho \parallel \sigma)$, 有

$$\begin{aligned}
 S(\rho \parallel \sigma) &= \langle -\log \sigma \rangle - S(\rho) \\
 &= \langle \beta H \rangle - \beta F(\sigma) - S(\rho) \\
 &= \beta E(\rho) - S(\rho) - \beta F(\sigma) \\
 &= \beta [F(\rho) - F(\sigma)].
 \end{aligned} \tag{12.87}$$

式中 $E(\rho) = \langle H \rangle = \text{Tr}(\rho H)$, 而 $F(\rho) = E(\rho) - TS(\rho)$ 为 ρ 态的自由能。

下面我们将与系统接触的恒温热源看成系统所处的环境, 它的存在定义了一个量子映射 \mathcal{E} , 由于热平衡态 σ 在这样的恒温环境下将保持不变, 因此这个量子映射 \mathcal{E} 将把 σ 映射到 σ 本身。因此由(12.84)式和(12.87)式, 我们即有

$$\begin{aligned}
 S(\rho \parallel \sigma) &\geq S(\rho' \parallel \sigma) \\
 \Rightarrow \beta [F(\rho) - F(\sigma)] &\geq \beta [F(\rho') - F(\sigma)] \\
 \Rightarrow F(\rho) &\geq F(\rho').
 \end{aligned} \tag{12.88}$$

也即是说, 这个系统的自由能是随着时间递减的。这就完成了我们的证明。