

量子统计的基本原理

陈童

August 5, 2025

Contents

1 密度算符	2
1.1 期望值以及算符迹	2
1.2 混态	4
2 再谈密度算符与系综	7
3 量子熵	10
3.1 冯诺伊曼熵	10
3.2 相对熵	13
4 海森堡绘景与薛定谔绘景	14
5 粗粒化演化与热力学第二定律	17
6 三大热平衡系综	18

本章我们讲述量子统计力学的基本原理，也就是把前面第一章和第二章中为经典统计力学建立的理论框架推广到量子情形。为了完成这一推广，需要引入密度算符的概念，它是经典统计中相空间概念密度的量子对应物。此外，还需要把经典吉布斯熵和相对熵的概念推广到量子情形。最后，还需要把第二章中讨论的粗粒化演化和熵增加原理推广到量子系统。有了这些准备以后，建立量子系统的三大热平衡系综(微正则系综、正则系综以及巨正则系综)就是一件顺理成章的事情。

1 密度算符

1.1 期望值以及算符迹

测量与期望值

量子力学告诉我们：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

$$\bar{A} = \langle \psi | A | \psi \rangle, \quad (1)$$

$|\psi\rangle$ 为系统所处的量子态。量子力学原理也告诉我们，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

$$p_i = \langle \psi | i \rangle \langle i | \psi \rangle = \langle \psi | P_i | \psi \rangle. \quad (2)$$

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 $\langle \mathcal{O} \rangle$ ，

$$\langle \mathcal{O} \rangle = \langle \psi | \mathcal{O} | \psi \rangle. \quad (3)$$

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，在幺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

$$\text{Tr}(\mathcal{O}) = \sum_i \langle i | \mathcal{O} | i \rangle, \quad (4)$$

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幺正变换，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

$$\text{Tr}(AB) = \text{Tr}(BA). \quad (5)$$

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(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，(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

$$\langle \psi | \mathcal{O} | \psi \rangle = \text{Tr}(|\psi\rangle\langle\psi| \mathcal{O}) = \text{Tr}(\rho_\psi \mathcal{O}). \quad (6)$$

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。下面我們就是要将密度算符这一概念推广到更为一般的情形。

1.2 混态

密度算符与系综

上面的密度算符 ρ_ψ 只涉及一个量子态 $|\psi\rangle$ 。然而，有时候可能有多个量子态 $|a\rangle$, $a = 1, 2, \dots$, 而我们不能确定系统到底处于哪一个态，只知道系统处于 $|a\rangle$ 态的概率是 p_a ，这时候就需要将密度算符推广成如下更一般的形式

$$\rho = \sum_a p_a |a\rangle\langle a|. \quad (7)$$

显然， ρ 满足 $\text{Tr}(\rho) = \sum_a p_a = 1$ 。如果这样的概率不只分布在单个纯态上，那么相应的 ρ 就称作混态密度算符，我们称它描写的系统状态为混态。

总之，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：

- 它是一个厄密算符。
- 它的本征值都大于等于0，为正定算符。
- $\text{Tr}(\rho) = 1$ 。

有了密度算符以后，对任何物理可观测量 \mathcal{O} 期望值的计算就可以表达为

$$\langle \mathcal{O} \rangle = \text{Tr}(\rho \mathcal{O}). \quad (8)$$

由此可见，密度算符就是第二章中相空间概率密度的量子力学推广，而Tr就是相空间积分的量子力学对应物。

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

$$\rho^2 = \rho. \quad (9)$$

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

$$\text{Tr}(\rho^2) = \text{Tr}(\rho) = 1. \quad (10)$$

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

$$\rho = \sum_a p_a |a\rangle\langle a|. \quad (11)$$

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

$$\text{Tr}(\rho^2) < 1. \quad (12)$$

密度算符的(11)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(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)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

$$\langle \mathcal{O} \rangle = \sum_a p_a \langle a | \mathcal{O} | a \rangle = \text{Tr}(\rho \mathcal{O}). \quad (13)$$

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。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，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。

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

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (14)$$

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 $\rho = \frac{1}{2}(w_0 \cdot 1 + \mathbf{w} \cdot \vec{\sigma})$ ，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

$$\rho(\mathbf{w}) = \frac{1}{2}(1 + \mathbf{w} \cdot \vec{\sigma}). \quad (15)$$

\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

$$\mathbf{w}^2 \leq 1, \quad (16)$$

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

$$\rho(0) = \frac{1}{2} \cdot 1 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|). \quad (17)$$

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 $(\mathbf{n} \cdot \vec{\sigma})$ 的期望值，易得

$$\langle (\mathbf{n} \cdot \vec{\sigma}) \rangle_{\mathbf{w}} = \text{Tr}((\mathbf{n} \cdot \vec{\sigma})\rho(\mathbf{w})) = \mathbf{n} \cdot \mathbf{w}. \quad (18)$$

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 $\rho(\mathbf{n}) = \frac{1}{2}(1 + \mathbf{n} \cdot \vec{\sigma})$ 。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 $(\mathbf{n} \cdot \vec{\sigma})\rho(\mathbf{n}) = \rho(\mathbf{n})(\mathbf{n} \cdot \vec{\sigma}) = \rho(\mathbf{n})$, 容易看出, $|\psi(\mathbf{n})\rangle$ 必为 $(\mathbf{n} \cdot \vec{\sigma})$ 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

$$\rho(\mathbf{n}) = \begin{pmatrix} \cos^2(\theta/2) & \cos(\theta/2)\sin(\theta/2)e^{-i\phi} \\ \cos(\theta/2)\sin(\theta/2)e^{i\phi} & \sin^2(\theta/2) \end{pmatrix}. \quad (19)$$

由此可以得到

$$|\psi(\mathbf{n})\rangle = \begin{pmatrix} \cos(\theta/2)e^{-i\phi/2} \\ \sin(\theta/2)e^{i\phi/2} \end{pmatrix}. \quad (20)$$

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

$$t\rho(\mathbf{w}_1) + (1 - t)\rho(\mathbf{w}_2) = \rho(t\mathbf{w}_1 + (1 - t)\mathbf{w}_2), \quad (21)$$

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。

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(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中内连线段内，而只能处于连线段的端点，那这个点就称为极端点。比如图(1)中的 A, B, C, D, E 点就是极端点。

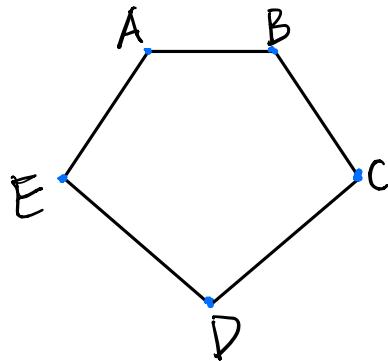


Figure 1: 两维实向量空间的一个凸子集， A, B, C, D, E 为极端点。

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

$$\rho = p_1\rho_1 + p_2\rho_2 + \dots + p_N\rho_N, \quad (22)$$

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，(22)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端点。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端点。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端点。这就完成了密度算符凸子集的极端点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端点，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，它的边界点和极端点是一回事，但是，图(1)告诉我们，对于一般的凸子集，它的极端点必定是边界点，但是反过来，它的边界点却不一定都是极端点。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端点。

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

$$\langle \mathcal{O} \rangle = \text{Tr}(\rho \mathcal{O}) = \sum_i p_i \text{Tr}(\rho_i \mathcal{O}) = \sum_i p_i \langle \mathcal{O} \rangle_i. \quad (23)$$

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(23)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

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|. \quad (24)$$

式中

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (25)$$

3 量子熵

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

$$\rho_{S_1}\rho_{S_2} = \rho_{S_1} \otimes \rho_{S_2}, \quad (26)$$

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

$$-\log(\rho). \quad (27)$$

额外加上的负号是为了使得这个算符为一个正定算符¹。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

$$S(\rho) = k_B \langle -\log \rho \rangle = -k_B \text{Tr}(\rho \log \rho). \quad (28)$$

从根本上说式中玻尔兹曼常数 k_B 的引入完全是单位选取上的约定俗成，当然既然在经典吉布斯熵中引入了 k_B ，那这里为了和经典熵能对应上， k_B 就是必须的。显然， $S(\rho)$ 也是所有 $-\rho \log \rho$ 本征值之和，从而也有

$$S(\rho) = -k_B \sum_a p_a \log(p_a). \quad (29)$$

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

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

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，

$$S(\rho) \geq 0, \quad (30)$$

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(29)式中的指标 $a = 1, 2, 3, \dots, D$ 。则由(29)式可以证明，

$$S(\rho) \leq k_B \log D. \quad (31)$$

等于号当且仅当 $p_a = 1/D$ 时成立，这时候的密度算符为

$$\rho = \frac{1}{D} \sum_a |a\rangle\langle a| = \frac{1}{D} \cdot 1. \quad (32)$$

这当然是一个混态密度算符，相应的混态称之为最大混态。

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

$$\frac{d}{dt}S(\rho(t)) = -k_B \text{Tr}(\dot{\rho} \log \rho). \quad (33)$$

又注意到

$$\log \rho = \int_0^\infty ds \left(\frac{1}{s+1} - \frac{1}{s+\rho(t)} \right). \quad (34)$$

再注意到 $\ddot{\rho}(t) = 0$ ，则有

$$\frac{d^2}{dt^2}S(\rho(t)) = -k_B \int_0^\infty ds \text{Tr} \left(\dot{\rho} \frac{1}{s+\rho(t)} \dot{\rho} \frac{1}{s+\rho(t)} \right). \quad (35)$$

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

$$\sum_i p_i S(\rho_i) \leq S\left(\sum_i p_i \rho_i\right). \quad (36)$$

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

$$\chi = S\left(\sum_i p_i \rho_i\right) - \sum_i p_i S(\rho_i). \quad (37)$$

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

$$S(\rho \| \sigma) = k_B \langle -\log \sigma \rangle - k_B \langle -\log \rho \rangle = k_B \text{Tr}(\rho \log \rho - \rho \log \sigma). \quad (38)$$

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， $S(\rho \| \sigma)$ 也称作 σ 和 ρ 的量子相对熵。可以想见，量子相对熵 $S(\rho \| \sigma)$ 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

$$S(\rho \| \sigma) \geq 0, \quad (39)$$

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²。

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(式中 Tr_B 表示在子系统 B 的希尔伯特空间上求迹)，它只作用在子系统 A 的希尔伯特空间上， $\rho_B = \text{Tr}_A(\rho_{AB})$ 为子系统 B 的约化密度算符(Tr_A 表示在子系统 A 的希尔伯特空间上求迹)，它只作用在子系统 B 的希尔伯特

²A Mini-Introduction To Information Theory, Edward Witten, arXiv:1805.11965[hep-th]

空间上。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}

$$\sigma_{AB} = \rho_A \otimes \rho_B. \quad (40)$$

则我们这一猜测的额外无知为(注意 $\text{Tr}_{AB} = \text{Tr}_A \text{Tr}_B = \text{Tr}_B \text{Tr}_A$)

$$\begin{aligned} S(\rho_{AB} \parallel \sigma_{AB}) &= k_B \text{Tr}_{AB} (\rho_{AB} \log \rho_{AB} - \rho_{AB} \log \sigma_{AB}) \\ &= k_B \text{Tr}_{AB} (\rho_{AB} \log \rho_{AB} - \rho_{AB} \log \rho_A - \rho_{AB} \log \rho_B) \\ &= S_A + S_B - S_{AB}. \end{aligned} \quad (41)$$

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

$$I(A, B) = S_A + S_B - S_{AB}. \quad (42)$$

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

$$I(A, B) = S_A + S_B - S_{AB} \geq 0. \quad (43)$$

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。另外，不等式(43)有时候也称之为熵的次可加性(subadditivity)。

4 海森堡绘景与薛定谔绘景

在量子力学中，刻画一个量子系综需要一组物理量算符以及描述系综量子态的一个密度算符，不妨记其中一个任意的物理量算符为 \mathcal{O} ，记系综的密度算符为 ρ ，则实验可观测的量总是由某物理量算符 \mathcal{O} 在 ρ 上的期望值 $\langle \mathcal{O} \rangle_\rho$ 给出，其中

$$\langle \mathcal{O} \rangle_\rho \equiv \text{Tr}(\rho \mathcal{O}). \quad (44)$$

当然，量子系统会随着时间演化，这种演化由系统的哈密顿量 H 决定，它反映为期望值 $\langle \mathcal{O} \rangle_\rho$ 会随着时间演化。但是关于 $\langle \mathcal{O} \rangle_\rho$ 随时间的演化是由物理量算符 \mathcal{O} 的演化引起还是由密度算符 ρ 的演化引起，我们有两种常用的不同观点，分别称之为，海森堡绘景和薛定谔绘景。

海森堡绘景

让我们先来看海森堡绘景，它认为是物理量算符在随时间演化，而密度算符则是不变的。为了讲得更清楚具体一点，我们定义时间演化算符 $U(t)$

$$U(t) = \exp(-iHt/\hbar). \quad (45)$$

显然，它满足

$$U(0) = 1, \quad U(t)U(s) = U(t+s). \quad (46)$$

换言之， $U(t)$ 构成一个阿贝尔群。将 $U(t)$ 对时间 t 求导，可以得到其满足的微分方程，为

$$i\hbar \frac{d}{dt}U(t) = HU(t). \quad (47)$$

海森堡绘景认为：描述量子态的密度算符 ρ 从不随时间演化，随时间演化的是物理量算符，它其实应该为如下 \mathcal{O}_t ，

$$\mathcal{O}_t \equiv U^\dagger(t)\mathcal{O}U(t), \quad (48)$$

注意，式中 \mathcal{O} 是不随时间演化的，但它只是 $t = 0$ 时刻的物理量算符，即 $\mathcal{O}_0 \equiv \mathcal{O}_{t=0} = \mathcal{O}$ 。根据 \mathcal{O}_t 的定义和方程(47)，不难得到如下海森堡运动方程

$$i\hbar \frac{d}{dt}\mathcal{O}_t = [\mathcal{O}_t, H]. \quad (49)$$

这个微分方程的初始条件为

$$\mathcal{O}_0 = \mathcal{O}. \quad (50)$$

进而物理量的期望值应该表达为

$$\langle \mathcal{O}_t \rangle_\rho \equiv \text{Tr}(\rho\mathcal{O}_t). \quad (51)$$

薛定谔绘景

如果我们将 \mathcal{O}_t 的定义(48)代入期望值公式(51), 即有

$$\mathrm{Tr}(\rho\mathcal{O}_t) = \mathrm{Tr}(\rho U^\dagger(t)\mathcal{O}U(t)) = \mathrm{Tr}(U(t)\rho U^\dagger(t)\mathcal{O}). \quad (52)$$

由此, 我们也完全可以把时间演化算符与态的密度算符结合起来, 而不是和物理量算符结合起来, 即可以定义随时间演化的态 ρ_t

$$\rho_t \equiv U(t)\rho U^\dagger(t). \quad (53)$$

从而即可以将上面的期望值公式重写成

$$\langle \mathcal{O}_t \rangle_\rho \equiv \mathrm{Tr}(\rho\mathcal{O}_t) = \mathrm{Tr}(\rho_t\mathcal{O}) \equiv \langle \mathcal{O} \rangle_{\rho_t}. \quad (54)$$

也即是说, 我们可以完全等价地认为是量子态在按照 ρ_t 随时间演化, 而物理量算符是不随时间演化的, 永远为 \mathcal{O} 。

以上这种等价观点就是薛定谔绘景。根据 ρ_t 的定义不难看出, 它满足如下薛定谔方程

$$i\hbar \frac{d}{dt}\rho_t = [H, \rho_t]. \quad (55)$$

这个微分方程的初始条件为

$$\rho_0 = \rho. \quad (56)$$

很显然, 海森堡绘景是经典哈密顿绘景的量子力学对应物, 薛定谔绘景则是经典刘维尔绘景的量子对应物。而 ρ_t 所满足的薛定谔方程其实就是经典刘维尔方程的量子版本。

根据(53)式, 以及 $U^\dagger U = 1$ 从而 $f(U\rho U^\dagger) = Uf(\rho)U^\dagger$, 不难有

$$\begin{aligned} S(\rho_t) &= -k_B \mathrm{Tr}\rho_t \log \rho_t = -k_B \mathrm{Tr}[U(\rho \log \rho)U^\dagger] \\ &= -k_B \mathrm{Tr}[U^\dagger U(\rho \log \rho)] = -k_B \mathrm{Tr}[\rho \log \rho] = S(\rho). \end{aligned} \quad (57)$$

式中第二行利用了 $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. 以上结果告诉我们, 一个量子系统的冯诺依曼熵并不随着时间演化。当然这里考虑的冯诺依曼熵是一种精细熵, 所以这个结论就是量子系统的精细熵不随时间演化。

5 粗粒化演化与热力学第二定律

量子系统的精细化熵不随时间演化，但是粗粒化熵却不是这样。不过为了讨论量子系统的粗粒化熵，我们需要把第二章《统计物理的基本原理》中的粗粒化演化的概念推广到量子情形。

同样，我们把时间离散化成步长为 τ 的离散时间步，因此 $t_{i+1} = t_i + \tau$ ，这里 t_i 为第*i*步的时间。我们记 $\rho_{t_i}^c$ 为第*i*步的粗粒化密度算符，则可以定义第*i+1*步在粗粒化之前的密度算符为

$$\rho_{t_{i+1}} \equiv U(\tau) \rho_{t_i}^c U^\dagger(\tau). \quad (58)$$

根据关于精细化熵不变的讨论，很显然，这一步是保持冯诺依曼熵不变的，即有

$$S(\rho_{t_{i+1}}) = S(\rho_{t_i}^c). \quad (59)$$

为了定义粗粒化之后的 $\rho_{t_{i+1}}^c$ ，我们把系统的希尔伯特空间 \mathcal{H} 分解成若干小的希尔伯特子空间 \mathcal{H}_α 的直和，即 $\mathcal{H} \equiv \bigoplus_\alpha \mathcal{H}_\alpha$ ，并假设每个子空间 \mathcal{H}_α 的维数均为 d ，即

$$d = \dim(\mathcal{H}_\alpha) = \text{Tr}_{\mathcal{H}_\alpha} 1. \quad (60)$$

进而可以定义

$$\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha} \equiv \frac{1_{\mathcal{H}_\alpha}}{d} \cdot \text{Tr}_{\mathcal{H}_\alpha}(\rho_{t_{i+1}}|_{\mathcal{H}_\alpha}), \quad (61)$$

式中 $\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha}$ 表示将算符 $\rho_{t_{i+1}}^c$ 限制在子空间 \mathcal{H}_α 上，等式右边的 $1_{\mathcal{H}_\alpha}$ 表示子空间 \mathcal{H}_α 上的单位算符。迭代上面的离散时间步，就完成了对粗粒化演化的定义。

根据上面的定义，不难有 $\text{Tr}_{\mathcal{H}_\alpha} \rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha} = \text{Tr}_{\mathcal{H}_\alpha}(\rho_{t_{i+1}}|_{\mathcal{H}_\alpha})$ ，进而有

$$\begin{aligned} \text{Tr}_{\mathcal{H}_\alpha} [\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha} \log (\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha})] &= \text{Tr}_{\mathcal{H}_\alpha} [\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha} \log (\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha})] \\ &= \text{Tr}_{\mathcal{H}_\alpha} [\rho_{t_{i+1}}|_{\mathcal{H}_\alpha} \log (\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha})], \end{aligned} \quad (62)$$

这里我们注意到 $\rho_{t_{i+1}}^c|_{\mathcal{H}_\alpha}$ 是一个常数算符。通过利用上式两边对 α 求和的结果，进而即有

$$\begin{aligned}\delta S &\equiv S(\rho_{t_{i+1}}^c) - S(\rho_{t_i}^c) = S(\rho_{t_{i+1}}^c) - S(\rho_{t_{i+1}}) \\&= k_B \left[-\text{Tr} \rho_{t_{i+1}}^c \log \rho_{t_{i+1}}^c + \text{Tr} \rho_{t_{i+1}} \log \rho_{t_{i+1}} \right] \\&= k_B \left[-\text{Tr} \rho_{t_{i+1}} \log \rho_{t_{i+1}}^c + \text{Tr} \rho_{t_{i+1}} \log \rho_{t_{i+1}} \right] = S(\rho_{t_{i+1}} \parallel \rho_{t_{i+1}}^c) \geq 0.\end{aligned}\quad (63)$$

由此可见，粗粒化演化过程是熵增加的！所有的处理和推导都完全平行于第二章中关于粗粒化演化的处理。因此，类似于第二章，取热力学极限之后就有量子系统的热力学第二定律，或者说熵增加原理。

有了孤立系统的熵增加原理之后，接下来的讨论就完全和第二章中的相关讨论一样了。比方说，同样可以把孤立系统的热平衡态定义为给定宏观条件下的极大熵状态。

6 三大热平衡系综

微正则系综

不妨先考察孤立系统的热平衡态，也就是所谓的微正则系综。记能量本征值为 E 的本征子空间为 \mathcal{H}_E ，则由于孤立系统的能量守恒，我们所考察的密度算符应该限制在这个本征子空间之内。考虑到在热力学极限下，一个典型量子系统的能谱通常来说为连续谱，所以，在粗粒化之后，我们就需要考虑本征值在 $[E, E + \delta E]$ 区间之内的本征子空间，其中 δE 为小量。由于能量守恒，所考察的密度算符只在这样的本征子空间之内才非零。因此，我们可以设系统的密度算符 ρ 为

$$\rho = \begin{cases} \sigma \frac{1}{\delta E}, & \text{能量本征值落在 } [E, E + \delta E] \text{ 区间之内} \\ 0, & \text{否则} \end{cases} \quad (64)$$

式中 σ 为某个厄密算符，它只作用在本征子空间 \mathcal{H}_E 上。

根据密度算符的归一化条件，有

$$\text{Tr} \rho = \text{Tr}_{\mathcal{H}_E}(\sigma) \cdot \delta N(E) / \delta E = 1. \quad (65)$$

式中 $\delta N(E)$ 表示本征值在 $[E, E + \delta E]$ 区间之内的能量本征态的数目。不妨记 $\delta N(E)/\delta E \equiv \Omega(E)$ 为态密度，则以上结果可以重写成

$$\mathrm{Tr}_{\mathcal{H}_E}(\sigma) = \frac{1}{\Omega(E)}. \quad (66)$$

相应的冯诺依曼熵为

$$\begin{aligned} S(\rho) &= -k_B \frac{\delta N(E)}{\delta E} \mathrm{Tr}_{\mathcal{H}_E}[\sigma \log(\sigma/\delta E)] \\ &= -k_B \Omega(E) \mathrm{Tr}_{\mathcal{H}_E}(\sigma \log \sigma) + k_B \log \delta E. \end{aligned} \quad (67)$$

其中 $k_B \log \delta E$ 项的贡献是非广延的，因此在热力学极限下相比于前面的广延项它可以忽略。如此一来，极大化 $S(\rho)$ 就等价于极大化如下泛函

$$S_G(\sigma) = -k_B \Omega(E) \mathrm{Tr}_{\mathcal{H}_E}(\sigma \log \sigma), \quad (68)$$

相应的约束条件就是上面的(66)式。对 σ 进行变分并利用拉格朗日乘子法，不难得到极大值对应如下方程

$$\Omega(E) \mathrm{Tr}_{\mathcal{H}_E} [(-k_B(\log \sigma + 1) + \lambda)\delta\sigma] = 0, \quad (69)$$

式中 λ 为拉格朗日乘子。进而不难得到相应的 σ 必定为常数算符，代入归一化条件(66)，立即有(不妨设 \mathcal{H}_E 本身的简并度为1)

$$\sigma \equiv \frac{1_{\mathcal{H}_E}}{\Omega(E)}. \quad (70)$$

这时候，相应的熵为

$$S = k_B \log \Omega(E). \quad (71)$$

这两个结果就确立了微正则系综的等概率原理和玻尔兹曼熵公式。很显然，这些结论与经典情形是完全对应的。

正则系综以及巨正则系综

与经典版本相比，正则系综的量子版本几乎没有什么新的东西。同样是先引入自由能泛函

$$F(\rho) \equiv E(\rho) - TS(\rho) = \langle H \rangle_\rho - TS(\rho), \quad (72)$$

这里与经典版本相比唯一的不同在于，这里的 ρ 是密度算符，这里的 $S(\rho)$ 是冯诺依曼熵。完全类似的，我们也有恒温系统的自发过程必定导致自由能减少(假设系统与外界之间没有力学功的交换)，进而恒温系统的热平衡态就对应自由能泛函 $F(\rho)$ 的极小。

为了求出自由能泛函 $F(\rho)$ 的极小，需要对密度算符 ρ 进行变分，并要通过拉格朗日乘子法考虑到 ρ 的归一化条件，最终可以得到正则系综的密度算符为

$$\rho = \frac{1}{Z} e^{-\beta H}, \quad (73)$$

很显然，这里的一切都和经典情形完全一样，区别只在于现在哈密顿量 H 是算符，从而 ρ 也是算符。归一化条件 $\text{Tr}\rho = 1$ 告诉我们

$$Z = \text{Tr}(e^{-\beta H}), \quad (74)$$

其它各热力学量的求法均与经典情形完全一致，这里不再赘述。

量子版本的巨正则系综甚至比经典版本更为简单，因为现在不必对每一不同粒子数的情形都分别引入一个概率密度，只需对系统(无论其粒子数为多少)笼统地引入一个粒子数算符 N (注意，现在是厄密算符了)，以及笼统地引入描写量子态的密度算符 ρ 。系统的巨热力学势泛函可以定义为

$$\Phi(\rho) \equiv \langle H \rangle_\rho - TS(\rho) - \mu \langle N \rangle_\rho, \quad (75)$$

式中 T, μ 分别为热库的温度以及化学势。完全类似的，热平衡态相应于 $\Phi(\rho)$ 的极小值。通过变分不难求出巨正则系综的密度算符为

$$\rho = \frac{1}{\Xi} e^{-\beta(H-\mu N)}. \quad (76)$$

利用归一化条件，有

$$\Xi = \text{Tr}(e^{-\beta(H-\mu N)}). \quad (77)$$

不过，值得注意的是，由于现在 H 和 N 都是厄密算符，所以这里涉及到这两个算符作用顺序的问题。通常人们要求这两个算符可交换顺序，即 $[H, N] = 0$ ，因为只有这样才能满足 $[\rho, H] = 0$ (这是热平衡的要求之一，相应于经典统计里的 $\{\rho, H\} = 0$)。换言之，要求粒子数算符是系统的守恒

量。直观上这是因为，在巨正则系综的考察中，粒子并没有凭空消失或产生，它们只是和粒子源之间产生了交换。

在量子场论的应用中，粒子数常常不是守恒量，成为守恒量的是诸如电荷量这样的量。那这时候就不能用粒子数算符来定义巨正则系综了，而应该相应地用电荷算符。具体的定义和粒子数算符情形完全类似，这里不再赘述。