

0.0.1 系综与公设

我们知道,一个由大量原子组成的系统,其宏观的平衡态只由少数几个状态变量所确定。只要系统处于某平衡态,则它的各状态函数性质的值就是确定的。例如,一个单组份体系的熵 $S = S(U, V, n)$, 由体系的内能 U 、体积 V 和摩尔数 n 所确定。然而,组成这个体系的 $N = nN_A$ 个原子,每一时刻 t 的位置 $\mathbf{r}_i(t)$ 和动量 $\mathbf{p}_i(t)$ 都不同。体系的瞬时状态由一组 $(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$ 的取值所确定。在体积为 V 、内能为 U 的限定下,这个系统仍然可以取很多组不同值的原子位置和动量。宏观上,我们认为体系已经处于 (U, V, n) 取某定值的平衡态,各宏观热力学性质已经恒定,不再随时间变化;但微观上,体系的原子位置和动量,一直在运动变化,只是它们的取值总满足 (U, V, n) 为一定值这一约束。

我们称体系由一组宏观状态变量 (X, Y, \dots) 所确定的热力学平衡态称作体系的一个宏观状态,这组状态变量称为宏观约束。体系由一组 $(\mathbf{r}^N, \mathbf{p}^N)$ 的取值所确定的状态称为体系的一个微观状态。其中记号 \mathbf{r}^N 是指 $\mathbf{r}_1, \dots, \mathbf{r}_N$, \mathbf{p}^N 也类似。在每一微观态下,也能计算体系的各个性质。例如,某孤立体系在平衡态下总能量不变。而对每一微观状态 $(\mathbf{r}^N, \mathbf{p}^N)$, 也能计算体系的总能量,它是这 N 个粒子的总动能和总势能之和,利用各粒子的位置和动量就能算出来。我们把体系宏观平衡态性质记为 $M = M(X, Y, \dots)$, 而由微观状态计算的相同性质计为 $\tilde{M} = \tilde{M}(\mathbf{r}^N, \mathbf{p}^N)$ 。性质 \tilde{M} 顶上的“ \sim ”表示这是微观状态的函数,以区别于同一物理量的宏观状态性质 M 。

体系可以取的所有微观状态组成一个概率论意义上的样本空间。例如非相对论经典力学中 N 个粒子的位置 $\mathbf{r}_i \in \mathbb{R}^3$ 和动量 $\mathbf{p}_i \in \mathbb{R}^3$ 可在整个 \mathbb{R}^{6N} 空间取值。但是在一定的宏观约束下,体系取不同微观状态的概率是不同的。例如,在恒定粒子数 N 、体积 V 和总能量 U 的宏观约束下,任一粒子位置处在这个体积区域之外的概率就无限小。如果我们给一个在某宏观约束 (X, Y, \dots) 下处于平衡态的体系不断拍摄快照,快照中能分辨该体系的瞬时微观状态。则这些微观状态均满足 $\tilde{X} = X, \tilde{Y} = Y, \dots$ 。随着拍照次数 \mathcal{N} 的增加,体系在这个宏观约束下能取的微观状态都出现过了,而且出现次数的直方图将趋于一个稳定的分布。这是概率论当中“随机试验”的思想。现在我们进一步假想,每次拍照不仅把这个系统当时处于的微观状态记录了下来,还对处于这个宏观约束的系统作了一分“活的复制”。我们讨论的快照次数 \mathcal{N} 很大,于是对系统的复制数也就很大。所谓“活的复制”,就是复制下来的每个假想系统都能够在相同的宏观约束下自行运动发展而不必保持其在被拍摄快照时的那个微观状态。如果我们要求,这一套复制的数量 \mathcal{N} 非常大,以保证平衡态下任一时刻只要我们把它们拿出来统计一番,其微观状态分布直方图都已经达到了那一个稳定的曲线(在概率论意义上我们其实要求 $\mathcal{N} \rightarrow \infty$);而且这些复制的系统完全相同且处于相同的宏观约束之外,但它们各自的运动变化是相互独立的,那么就把这样一个对实际系统的大量复制的集合称为系综(ensemble)。系综当然是一个假想的概念。它的出发点是无穷次随机试样反映统计规律的概率论思想,但又有超出了这个想法的内容(这些系统复制是“活的”)。

我们如果能够知道体系在一定的宏观约束 (X, Y, \dots) 下达到平衡态时,满足该宏观约束的个微观状态的概率,那么对系统的某一性质 $M(X, Y, \dots)$ (若它有对应的微观状态计算方法所得到的 $\tilde{M}(\mathbf{r}^N, \mathbf{p}^N)$), 我们都可以按这个密度去计算 \tilde{M} 的期望。具体地,设满足宏观约束的第 j 个微观状态的概率质量函数是 P_j , 按这个微观状态计算的性质取值是 \tilde{M}_j , 则记

$$\langle \tilde{M} \rangle \equiv \sum_{j \in \{\text{满足宏观约束 } (X, Y, \dots) \text{ 的所有微观状态}\}} P_j \tilde{M}_j$$

称为性质 M 的系综平均(ensemble average)。其中,除了求和序号 j 遍历的是满足宏观约束 (X, Y, \dots) 的所有微观状态个数外,各微观状态的概率质量 $\{P_j\}$ 的取值是依赖宏观约束 (X, Y, \dots) 的,因此,尽管求完系综平均之后的 $\langle \tilde{M} \rangle$ 已经不依赖具体的微观状态了,但还依赖宏观体系的状态变量 (X, Y, \dots) 。

平衡态统计力学的第一个公设是系综平均公设:在平衡状态下,体系的宏观性质 M 等于相应宏观约束下体系的微观状态的系综平均,即 $M = \langle \tilde{M} \rangle$ 。

我们实际上不可能真给一个系统拍无穷多个快照然后统计出这个系统在某宏观约束下的微观状态概率分布函数。统计力学理论在给出系统平均公设之后, 必须还要先验地告诉我们, 这个概率分布是怎样的。

现在我们考虑宏观约束为 (N, V, U) 的情况。热力学中, 一个体系的平衡态由 (N, V, U) 确定的, 应属于一个体积为 V 的单组份孤立系统(且没有外场的作用)。我们把 (N, V, U) 约束下的系综称作微正则系综(microcanonical ensemble)。

平衡态统计力学的第二个公设是孤立系统的等先验概率原则(principle of equal apriori probabilities)。量子力学告诉我们, 一个体系在宏观约束 (N, V, U) 下, 其能量简并度 $\Omega(U)$ 是确定的。此时等先验概率原则要求, 实际体系处于这 $\Omega(U)$ 个微观状态中的任一个的概率相等。

0.0.2 正则系综

正则系综概率分布的形式推导

封闭系统允许了系统与环境的能量交换。达到平衡后, 系统与环境同处一个温度 T 。因此确定一个单组份简单封闭系统的状态变量是 (N, V, T) 。在热力学基础的章节我们已经解释过, 我们常常让系统跟一个温度就是 T 的“大热源”环境接触来实现控制系统下于我们想要的温度, 只要这个环境规模足够大, 它自己的温度变化就可以忽略, 近似于一个总能恒定在温度 T 的体系。

我们把宏观约束为 (N, V, T) 的体系的系综称为正则系综(canonical ensemble)。正则系综就是用来讨论上述这种封闭系统的情况的。从统计力学的角度, 我们需要知道, 在宏观约束 (N, V, T) 下, 体系每一微观状态概率, 以便我们能够对微观状态计算出来的性质进行系综平均, 来预测宏观性质。但是我们唯一拥有的, 能够赋予微观状态以概率的公设——等先验概率原则——是应用于一个孤立系统(微正则系综)的。而在 (N, V, T) 约束下, 系统的各微状态可能取不同的能量, 而显示出一个能量的分布谱。我们注意到, 系统与大热源环境一齐可视为一个达到了平衡态的孤立系统。我们将通过恰当地构造一个很大的孤立系统, 然后视其中一个子系统为与这个大系统的剩余部分热交换达到平衡态, 来解决这个子系统处于它的某微状态的概率问题。此时, 由于大系统满足微正则系综, 因而这个大系统的各个微观状态概率相等, 但在大系统的这么多微观状态当中, 只有部分是我们所关心的小系统恰好处于我们所关心的微观状态的。通过概率论可以从前者的概率去计算后者的概率, 从而得出这个小系统的微状态概率, 而这正是一个(小)封闭系统在热平衡时(即正则系综)的微观状态概率。

具体地, 我们假想有 \mathcal{N} 个数量为 N 、体积为 V 的小系统, 拼接成一个大的系统。这些小系统之间的分隔不允许物质交换, 体积也恒定, 但允许热交换。我们先把如此拼接而成的大系统放到一个更加大的、温度为 T 的大热源环境中达到热平衡, 此时这个大系统中的各个小系统都处于 (N, V, T) 状态了。然后我们把这个大系统从大热源环境中取出并隔绝, 变成一个孤立系统。

由于这个大系统是孤立系统, 它能取的微观状态均等概率。但是我们要进一步讨论大系统的微观状态与小系统的关系, 以便得出某一个小系统处于其某个微观状态的概率。

首先, 这个大系统的总粒子数应该是 $\mathcal{N}N$ 、总体积应是 $\mathcal{N}V$ 。由于每个子系统的粒子数 (N, V) 是确定的, 因此每个子系统可取的能级(量子本征值)以及每个能级的简并度也是确定的(见量子力学基础)。我们把每个子系统可取的能级值列为 $E_1, E_2, \dots, E_j, \dots$ 。这里的各个能级值不是一个能级列表; 如果某能级有简并度, 我们就把这一能级按其简并数列为 j 不同的多个取值相同的 E_j 。也就是说 j 遍历的是, 每个小系统可取的所有微观状态; 当我们说一个小系统取能量 E_j 时, 不仅说出了它所取能量的大小是 E_j 的值, 还说出它取了哪个能级下的哪个简并态。

然后我们再考虑, 这个大系统在当初与大热源环境处于热平衡时, 体系取了某总能量 \mathcal{E} 。现在它被孤立出来了, 总能量 \mathcal{E} 保持不变。也就是说它是一个宏观约束为 $(\mathcal{N}N, \mathcal{N}V, \mathcal{U})$ 的系统。但是它的各个小系统之间可以发生热交换, 所以这些小系统, 谁取什么能级, 可以变来变去, 只要满足它们加起来的总能量恒定为 \mathcal{E} 。

我们可以先作一个类别区分。记有 n_j 个小系统取能量值 E_i 。那么不管 n_j 怎么取值, 以下总成立

$$\sum_j n_j = \mathcal{N}, \quad \sum_j n_j E_j = \mathcal{E} \quad (1)$$

满足这个规定的 $\{n_1, n_2, \dots\}$ 的取值可以有很多套。我们把某一套取值简记为向量 $\mathbf{n} = (n_1, n_2, \dots)$ 。

第三步, 我们要注意到, 在我们所考虑的这个问题中, 各个小系统之间是可区分的。因为我们最后要关心某一个选定的小系统的微状态, 这种可选定性说明, 我们能够分辨小系统。这样的话, 某一满足约束(1)的 \mathbf{n} , 就仅规定了有 n_1 个小系统能量为 E_1 、 n_2 个小系统能量为 E_2 、……而没有具体区分, 到底哪 n_1 个小系统能量为 E_1 、哪 n_2 个小系统能量为 E_2 、……。如果我们想要讨论, 各小系统能量值, 在满足约束(1)下的取法一共有多少种, 那么就相当于要从 \mathcal{N} 个可分辨的小球中, 选 n_1 个放到 E_1 盒子里、选 n_2 个放到 E_2 盒子里、……这样的放法还只是给定某 \mathbf{n} 时的情况。记它的方法数是 $W(\mathbf{n})$, 则

$$W(\mathbf{n}) = \frac{\mathcal{N}!}{\prod_j (n_j!)}$$

注意到, 满足约束(1)的 \mathbf{n} 还有好多个。所以总共有 $\sum_{\{\mathbf{n}\}} W(\mathbf{n})$ 种满足约束(1)情况。这里面的每种情况, 既是大系统的一个微观状态, 其中各小系统又都处在了其某一微观状态(取了某 E_j)。

第四步, 我们现在关心, 某个小系统取 E_j 的概率。由于大系统遵循微正则系综, 因而其 $\sum_{\{\mathbf{n}\}} W(\mathbf{n})$ 个微状态概率均等。但在些大系统微观状态当中, 每一个都可能不同数量的小系统的状态是 E_j 。或者说, 每个大系统的微观状态中, 取状态 E_j 的小系统个数占比 n_j/\mathcal{N} 是不同的。如果我们要问, 某一选定的小系统取 E_j 的概率是多少, 那就要问, 大系统的这么多微观状态中, 平均有几分之几的小系统取 E_j , 这也就是要问 n_j/\mathcal{N} 在大系统的这么多个微状态之中的平均值是多少。现在大系统的微状态概率均等, 这个平均值是能算的。关键是要表出, 大系统取每一微观状态时, 到底有多少个小系统取 E_j 。大系统取每一微观状态时, 有多少个小系统取 E_j 呢? 只要这个微观状态是属于某给定 \mathbf{n} 的 $W(\mathbf{n})$ 个之一, 那看看这个 \mathbf{n} 中的 n_j 是几, 就有几个小系统取 E_j 。我们记 $\pi_j(\mathbf{n})$ 为取向量 \mathbf{n} 的第 j 个分量值的函数。则某给定 \mathbf{n} 的 $W(\mathbf{n})$ 个大系统微状态中的每一个, 都有 $\pi_j(\mathbf{n})$ 个小系统取 E_j 。这些大系统微观状态中的第一个中, 取 E_j 的小系统占比就都是 $\pi_j(\mathbf{n})/\mathcal{N}$ 。其他 \mathbf{n} 的情况也依此类推。因此, 想求取 E_j 的小系统占比的平均值, 就要把每一个大系统微观状态下取 E_j 的小系统占比加起来, 除以大系统微观状态数总数, 它就是:

$$\frac{\overline{n_j}}{\mathcal{N}} = \frac{\sum_{\{\mathbf{n}\}} W(\mathbf{n}) (\pi_j(\mathbf{n})/\mathcal{N})}{\sum_{\{\mathbf{n}\}} W(\mathbf{n})}$$

上式就是我们最终想求的, 一个与温度为 T 的大热源保持热平衡的封闭系统, 取某一微观状态的概率。

经过这四步考虑, 我们完成了正则系综概率分布的形式推导。我们需要具体地把这个表达式计算为更有用的形式。这是需要取 $\mathcal{N} \rightarrow \infty$ 极限才能做到的。附录 xxx(待补充)将会证明, 当 $\mathcal{N} \rightarrow \infty$ 时, 一个满足多项分布 $W(\mathbf{n})$ 的随机变量 \mathbf{n} 的均值 $\bar{\mathbf{n}}$ 将趋于那个使分布概率密度最大的那个值 \mathbf{n}^* , 且分布的宽度收敛。所以 \mathcal{N} 越大, 拿 \mathbf{n}^* 代替 $\bar{\mathbf{n}}$ 的误差就越小。虽然在 $\mathcal{N} \rightarrow \infty$ 的同时 $\mathbf{n}^* \rightarrow \infty$ (意思是均值向量 \mathbf{n}^* 的各分量 n_j^* 都发散), 但我们关心的比例 n_j^*/\mathcal{N} 保持有限的定值。因此, 在 $\mathcal{N} \rightarrow \infty$, 正则系综概率质量函数

$$P_j = \frac{\overline{n_j}}{\mathcal{N}} = \frac{n_j^*}{\mathcal{N}}$$

我们再复述一次 n_j^* 是什么。 n_j^* 是使大系统满足约束(1)的某 \mathbf{n} 中, 使 $W(\mathbf{n})$ 最大的一个(记为 \mathbf{n}^*)的第 j 个分量。我们可以用拉格朗日乘子法去解, 在满足约束(1)下使函数 $W(\mathbf{n})$ 最大的这一最值问题(见附录 xxx 待补充)。结果得到, \mathbf{n}^* 的每个分量都取:

$$n_j^* = \mathcal{N} e^{-\alpha} e^{-\beta E_j}, \quad j = 1, 2, \dots$$

时满足这一要求。其中 α, β 是拉格朗日乘子。利用 $\sum_j n_j^* = \mathcal{N}$ 可以把 α 消掉, 从而, 正则系综的概率质量函数的形式变为

$$P_j = \frac{e^{-\beta E_j(N,V)}}{\sum_i e^{-\beta E_i(N,V)}}, \quad j = 1, 2, \dots$$

式中我们明显地提醒, 系统可取的能量状态数是依赖 (N, V) 约束而确定的。

从热力学关系确定 β

照理, α 和 β 的值需要代回到原函数中, 在令原函数取最值时解出来。但是在我们现在的问题中, $W(\mathbf{n})$ 的值是不可知的。要继续讨论, 就需要用到统计力学的第一条公设了。注意到, 在没有各种外场作用时, 一个系统的总能量就是它的热力学能(即内能)。这就能够联系函数 $W(\mathbf{n})$ 与体系的热力学状态函数。 α 和 β 就能表示成与热力学函数有关的形式。

具体地, 我们现在直接就能得到, 无外场作用下一个分子数为 N 、体积为 V 、温度为 T 的单组分简单封闭系统的内能 $U = E = \langle \tilde{E} \rangle$, 且

$$\langle \tilde{E} \rangle = \frac{\sum_j P_j E_j}{\sum_j P_j} = \frac{\sum_j E_j e^{-\beta E_j}}{\sum_j e^{-\beta E_j}}$$

我们看到 $\langle \tilde{E} \rangle$ 是 $(N, V, \{P_j\})$ 的函数, 它的全微分可以写成

$$\begin{aligned} d\langle \tilde{E} \rangle &= \sum_j E_j dP_j + \sum_j P_j dE_j \\ &= -\beta^{-1} \sum_j (\ln P_j + \ln Q) dP_j + \sum_j P_j \left. \frac{\partial E_j}{\partial V} \right|_N dV \end{aligned}$$

其中我们记 $Q \equiv \sum_j e^{-\beta E_j}$, Q 不依赖 P_j 。利用 P_j 的归一化性质,

$$\sum_j P_j = 1, \quad \sum_j dP_j = 0$$

含 Q 的项就没了。而我们看到

$$-p_j = \left. \frac{\partial E_j}{\partial V} \right|_N$$

是一个压强(参照(?)), $-p_j dV$ 是当系统的能量为 E_j 时, 使系统的体积变化 dV 所做的功。它是系统微状态的性质, 故有

$$\sum_j P_j \left. \frac{\partial E_j}{\partial V} \right|_N dV = -\langle \tilde{p} \rangle dV$$

因此全微分式给出

$$-\beta^{-1} d \left(\sum_j P_j \ln P_j \right) = d\langle \tilde{E} \rangle + \langle \tilde{p} \rangle dV$$

由第一公设, $\langle \tilde{E} \rangle$ 就是体系的内能 U , $\langle \tilde{p} \rangle$ 就是体系的压强 p 。而由热力学第一定律(??)(单组份封闭体系)我们得到

$$-\beta^{-1} d \left(\sum_j P_j \ln P_j \right)$$

就是 TdS 。