

# The Fireside Chat on Elements of Statistical Physics

## Contents

<b>1</b>	<b>导语</b>	<b>2</b>
<b>2</b>	<b>统计物理学的基本原理</b>	<b>3</b>
2.1	相空间 . . . . .	3
2.2	量子统计基本原理 . . . . .	8
2.3	熵 . . . . .	13
<b>3</b>	<b>热力学</b>	<b>16</b>
3.1	温度，热力学第零、第二定律 . . . . .	16
3.2	绝热过程和压强 . . . . .	17
3.3	热力学第一定律，麦克斯韦关系的其他元素 . . . . .	20

## 1 导语

## 2 统计物理学的基本原理

### 2.1 相空间

对于拥有 $s$ 个自由度的系统，其任意时刻的状态，以及今后的状态，唯一地由如下向量表征：

$$(q_1, q_2, \dots, q_s, p_1, p_2, \dots, p_s) \quad (1)$$

或者等价地 $(q_1, q_2, \dots, q_s, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_s)$ ，区别在于前者是贴近哈密顿原理导出的，后者从拉格朗日量导出，通过勒让德变换可以联系二者。一般采取坐标加动量的描述，因为更容易进行量子物理的推广，以及分立性。这个向量所处的空间就是该系统的**相空间**。

整个系统随着时间的变化在它的相空间中画出一条曲线，称为它的相轨道。给定相轨道的出发点，它是唯一确定的。并且两条相轨道不会相交，否则和反演性矛盾。

对于整个系统的一部分，即一个**子系统**，以 $a$ 表征，它在长程中会经历各种可能的状态，考虑子系统的相空间，并将其切分为由诸多 $\Delta q \Delta p$ 的小块组成的整体空间。从长程角度而言，子系统对应的状态/相点在子系统相空间中历经各个子块，从平稳的角度而言，子系统相点出于某个特定子块的时间占比是确定的：

$$\omega(q, p) = \lim_{t \rightarrow \infty} \frac{\Delta t}{t} \quad (2)$$

其中 $\Delta t$ 是子系统相点在 $t$ 时间内处在 $(q \sim q + \Delta q, p \sim p + \Delta p)$ 相区域的时间，它也等同于在任意一个时间观测该子系统时，它正好处于该相区域的概率，即微分到理想情况时：

$$d\omega = \rho(p_1, p_2, \dots, p_s, q_1, q_2, \dots, q_s) d^s q d^s p \quad (3)$$

这里密度函数 $\rho$ 应该理解为联系两个微分 $d\omega$ 和 $d^s q d^s p$ 的函数，即 $\frac{d\omega}{d^s p d^s q}$ ，这里的符号取广

义微分的涵义，且密度函数应该满足（从广义微分的意义上而言）：

$$\int \rho d^s q d^s p = \int d\omega = 1 \quad (4)$$

对于一个物理量 $f$ 而言，它在长程意义上的平均值应该是：

$$\langle f \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t) dt \quad (5)$$

由于我们已经知道 $\omega$ 的物理意义，所以在时间上的平均可以改写为对于状态的平均：

$$\langle f \rangle = \int f(q, p) d\omega = \int f(q, p) \rho(q, p) d^s q d^s p \quad (6)$$

对于复数个子系统而言，不是一般性我们考虑两个子系统，并假设它们互相独立，这里互相独立的意思就是，一个系统的状态不会作用于另一个系统。换言之，如果同时考察两个系统所处的状态，跟分别地、独立地考察两个系统的状态，会得到一样的结果，即：

$$d\omega_{1,2} = d\omega_1 d\omega_2 \quad (7)$$

这可以从独立事件的概率满足的贝叶斯公式得出，类似地进行微分，可以继而得到：

$$\rho_{1,2} dq^{1,2} dp^{1,2} = \rho_1 dq^1 dp^1 \rho_2 dq^2 dp^2 \quad (8)$$

这里我们省略了 $\rho$ ...对应的参数，上式的概率意义依旧是明显的。当两个子系统独立时，

考虑一个联合系统的物理量，它是两个子系统对应量的乘积，即  $f_{1,2} = f_1 \cdot f_2$ ，则：

$$\begin{aligned}
 \langle f_{1,2} \rangle &= \int f_{1,2}(q_1, p_1, q_2, p_2) \rho_{1,2} dq^{1,2} dp^{1,2} \\
 &= \int f_1 f_2 \rho_1 \rho_2 dq^1 dp^1 dq^2 dp^2 \\
 &= \left( \int f_1 \rho_1 dq^1 dp^1 \right) \left( \int f_2 \rho_2 dq^2 dp^2 \right) \\
 &= \langle f_1 \rangle \langle f_2 \rangle
 \end{aligned} \tag{9}$$

这也符合一般统计理论的结论，即独立变量函数的积的平均等于平均值的积。这一结论可以推导出中心极限的类似结论。考虑整体系统的一个物理量  $f$ ，它是各个子系统对应物理量的和  $f = \sum_{i=1}^N f_i$ ，其中把整个系统分为  $N$  个互相独立的子系统。这种可加性是很普遍的，譬如运动积分和许多热力学量，所以这种讨论是不失一般性的。首先有：

$$\begin{aligned}
 \langle f \rangle &= \int f(q^{1 \sim N}, p^{1 \sim N}) d\rho_{1 \sim N} q^{1 \sim N} dp^{1 \sim N} \\
 &= \int \sum_{i=1}^N f_i(q^i, p^i) \prod_{i=1}^N \rho_i dq^i dp^i \\
 &= \sum_{i=1}^N \int f_i \prod_{i=1}^N \rho_i dq^i dp^i \\
 &= \sum_{i=1}^N \int \left( \int f_i \rho_i dq^i dp^i \right) \prod_{j=1, j \neq i}^N \rho_j dq^j dp^j \\
 &= \sum_{i=1}^N \left( \int f_i \rho_i dq^i dp^i \right) \\
 &= \sum_{i=1}^N \langle f_i \rangle
 \end{aligned} \tag{10}$$

我们还可以继而考虑  $f$  的方均涨落，根据定义：

$$\langle (\Delta f)^2 \rangle = \langle \left( \sum_{i=1}^N \Delta f_i \right)^2 \rangle \tag{11}$$

但是对于  $i \neq j$ :

$$\langle \Delta f_i \Delta f_j \rangle = \langle \Delta f_i \rangle \langle \Delta f_j \rangle = 0 \quad (12)$$

根据独立性乘积变量的讨论，所以：

$$\langle (\Delta f)^2 \rangle = \sum_{i=1}^N \langle (\Delta f_i)^2 \rangle \quad (13)$$

导致：

$$\frac{\sqrt{\langle (\Delta f)^2 \rangle}}{\langle f \rangle} \propto \frac{1}{\sqrt{N}} \quad (14)$$

这是中心极限定理，或者弱大数定律在统计物理中的体现，也是统计物理正确性的支柱。

回到对于  $\rho$  的讨论，既然  $\rho$  是子系统相空间中的密度，就可以把它理解为一种相空间中的流体，并且满足流量的连续性方程：

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (15)$$

其中的  $\mathbf{v}$  应该理解为概率流体在相空间中的流速，这种流速的成因是子系统状态在时间中的不断变迁，展开散度项：

$$\nabla \cdot (\rho \mathbf{v}) = \sum_{i=1}^{2s} \frac{\partial}{\partial x_i} (\rho v_i) \quad (16)$$

其中  $s$  为子系统的自由度，每个  $x_i$  对应于某个广义坐标或者广义动量，代入它们的形式并

且利用哈密顿正则关系：

$$\begin{aligned}
 \nabla \cdot (\rho \mathbf{v}) &= \sum_{i=1}^{2s} \frac{\partial}{\partial x_i} (\rho v_i) \\
 &= \sum_{i=1}^s \left[ \frac{\partial}{\partial q_i} \rho \dot{q}_i + \frac{\partial}{\partial p_i} \rho \dot{p}_i \right] \\
 &= \sum_{i=1}^s \left[ \frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right] + \rho \sum_{i=1}^s \left[ \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right] \\
 &= \sum_{i=1}^s \left[ \frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right] \\
 &= \frac{d\rho}{dt} - \frac{\partial \rho}{\partial t}
 \end{aligned} \tag{17}$$

所以连续性条件化归为  $\frac{d\rho}{dt} = 0$ ，所以  $\rho$  只能是运动积分的函数，而独立子系统联合密度可改写为连乘的形式蕴含着，分布函数的对数应该是可加的运动积分，总共有7个备选项，但是动量，角动量可以通过整体移动来消除，所以不是一般性的：

$$\ln \rho_a = \alpha_a + \beta E_a(q, p) \tag{18}$$

其中  $a$  为子系统标记。换言之，子系统被观测处于某个状态  $(q, p)$  的概率，仅仅和  $(q, p)$  对应的能量有关。这继而引出一种理想情况，即子系统  $a$  的能量完全处于值  $E_0$  附近，即  $(E_0 - \Delta E, E_0 + \Delta E)$ ，则此时可以认为，在  $E_a \neq E_0$  对应的  $(q, p)$  处，子系统被观测到的概率均为0，此时子系统的概率密度平均地分布在  $E_0$  对应的所有可能的  $(q, p)$  值的集合上，而这个值的集合是该相空间中的低维流形（ $\lim \Delta E \rightarrow 0$  时），导致归一化条件失效。此时要么对  $\Delta E$  进行非零化并且尝试在计算末尾消去，要么使用狄拉克函数来确保积分的有效性，具体地：

$$\rho = C \cdot \delta(E_a - E_0) \tag{19}$$

这个分布，称为微正则（micro-canonical）分布显然满足平稳分布的条件，因为系统的  $E_0$  守恒不变，所以  $\rho$  在相空间中任一点对时间的微分都是零。

## 2.2 量子统计基本原理

在量子物理的讨论范围内，我们不再能够直接地以（子）系统的广义坐标或者动量来唯一地表征系统，而需要预设（子）系统所处状态并继而推断它的坐标和动量表象。对于子系统，令它的哈密顿算符为 $\hat{H}$ ，那么系统的状态可以在 $\hat{H}$ 的特征状态张成的空间中进行唯一地表述：

$$|\psi\rangle = \sum_n c_n |\psi_n\rangle \quad (20)$$

其中 $|\psi_n\rangle$ 是 $\hat{H}$ 的对应特征值，即能量为 $E_n$ 的特征状态。此时子系统某个表象物理量 $f$ 的观测平均值为：

$$\begin{aligned} \langle f \rangle &= \langle \psi | \hat{f} | \psi \rangle \\ &= \sum_{n,m} c_n^* c_m \langle \psi_n | \hat{f} | \psi_m \rangle \\ &= \sum_{n,m} w_{m,n} f_{n,m} \end{aligned} \quad (21)$$

其中 $w_{m,n} = c_n^* c_m$ ，而 $f_{n,m}$ 是算符 $\hat{f}$ 在 $\hat{H}$ 的特征状态空间的基中的矩阵表达。 $w_{m,n}$ 也可以看做算符 $\hat{w} = |\psi\rangle\langle\psi|$ 在同一组基下的矩阵元素：

$$\begin{aligned} \langle \psi_m | \psi \rangle \langle \psi | \psi_n \rangle &= c_m c_n^* \\ &= w_{n,m}. \end{aligned} \quad (22)$$

现在可以进一步写出：

$$\begin{aligned} \langle f \rangle &= \sum_{n,m} w_{m,n} f_{n,m} \\ &= \sum_n \left[ \sum_m w_{m,n} f_{n,m} \right] \\ &= \sum_n \left[ \hat{f} \hat{w} \right]_n \\ &= \text{tr} \left[ \hat{f} \hat{w} \right] = \text{tr} \left[ \hat{w} \hat{f} \right] \end{aligned} \quad (23)$$

可以联想到这里的 $\hat{w}$ 和量子统计理论中的密度算符 $\hat{\rho}$ 的相似性，但是此处我们暂且局限于



纯态系综 (pure ensemble) 的讨论中。 $\hat{w}$ 的矩阵元素有一些其他限制, 考虑:

$$\begin{aligned}
 p(|\psi\rangle \text{ is observed to be in } |\psi_n\rangle) &= |\langle\psi_n|\psi\rangle|^2 \\
 &= \langle\psi_n|\psi\rangle\langle\psi|\psi_n\rangle \\
 &= w_{n,n}
 \end{aligned} \tag{24}$$

所以应有  $w_{n,n} \geq 0$ , 且  $\sum_n w_{n,n} = \text{tr}(\hat{w}) = 1$ 。

为了联系到相空间上, 需要进一步考虑  $|\psi\rangle$  如何体现在广义坐标或者动量上, 此时由于坐标和动量的不可交换性, 无法同时考察, 我们首先考虑坐标  $q$ , 这一关系由波函数引导:

$$\langle q|\psi\rangle = \sum_n c_n \langle q|\psi_n\rangle \tag{25}$$

所以以  $|\psi\rangle$  表征的子系统被发现处于  $q$  附近的概率为 (此处的  $q$  是一个广义坐标值, 也是  $|q\rangle$  在  $\hat{q}$  下对应的特征值):

$$\begin{aligned}
 |\langle q|\psi\rangle|^2 &= \left( \sum_n c_n^* \langle\psi_n|q\rangle \right) \left( \sum_m c_m \langle q|\psi_m\rangle \right) \\
 &= \sum_{n,m} w_{m,n} \langle\psi_n|q\rangle \langle q|\psi_m\rangle
 \end{aligned} \tag{26}$$

也可以记:

$$|\langle q|\psi\rangle|^2 = \langle q|\hat{w}|q\rangle \tag{27}$$

这两种表达显然是等价的：

$$\begin{aligned}
 \sum_{n,m} w_{m,n} \langle \psi_n | q \rangle \langle q | \psi_m \rangle &= \sum_{n,m} \langle \psi_m | \psi \rangle \langle \psi | \psi_n \rangle \langle \psi_n | q \rangle \langle q | \psi_m \rangle \\
 &= \sum_{n,m} \langle \psi | \psi_n \rangle \langle \psi_n | q \rangle \cdot \langle \psi | \psi_m \rangle \langle \psi_m | q \rangle \\
 &= \left( \sum_n \langle \psi | \psi_n \rangle \langle \psi_n | q \rangle \right) \left( \sum_m \langle \psi | \psi_m \rangle \langle \psi_m | q \rangle \right) \\
 &= \langle \psi | \left( \sum_n |\psi_n\rangle \langle \psi_n| \right) | q \rangle \cdot \langle \psi | \left( \sum_m |\psi_m\rangle \langle \psi_m| \right) | q \rangle \\
 &= \langle q | \psi \rangle \langle \psi | q \rangle
 \end{aligned} \tag{28}$$

在坐标空间中某一点的观测概率就是前一章里的 $\rho$ ，类似地可以将该概率测度转换为长程时间测度：

$$dw_q = \langle q | \hat{w} | q \rangle dq \tag{29}$$

对于动量的测量完全同理：

$$dw_p = \langle p | \hat{w} | p \rangle dp \tag{30}$$

类似经典情况下的刘维尔定理，我们尝试导出量子情况下平稳状态的条件，此时应该有 $\hat{w}$ 的所有矩阵元素不变（前提是假定 $\hat{H}$ 不变，否则需要进入扰动理论进一步探讨），换言之： $\frac{\partial}{\partial t} w_{m,n}$ 应为零。回忆一下 $|\psi\rangle$ 在以 $\hat{H}$ 为哈密顿算符的情形下随时间的演变，众所周知地：

$$\begin{aligned}
 |\psi\rangle(t) &= \mathcal{U}(t) |\psi\rangle \\
 &= \exp \left\{ -\frac{it\hat{H}}{\hbar} \right\} |\psi\rangle \\
 &= \exp \left\{ -\frac{it\hat{H}}{\hbar} \right\} \sum_n c_n |\psi_n\rangle \\
 &= \sum_n c_n \exp \left\{ -\frac{itE_n}{\hbar} \right\} |\psi_n\rangle \\
 &= \sum_n c_n(t) |\psi_n\rangle
 \end{aligned} \tag{31}$$

现在我们可以显式地写出：

$$\begin{aligned}
 \frac{\partial}{\partial t} w_{m,n} &= \frac{\partial}{\partial t} [c_n^*(t) c_m(t)] \\
 &= \frac{\partial}{\partial t} \left[ c_n^* c_m \exp \left\{ \frac{it}{\hbar} (E_n - E_m) \right\} \right] \\
 &= c_n^* c_m \frac{i}{\hbar} (E_n - E_m) \exp \left\{ \frac{it}{\hbar} (E_n - E_m) \right\} \\
 &= \frac{i}{\hbar} (E_n - E_m) w_{m,n}
 \end{aligned} \tag{32}$$

可以将这个关系泛化到算符的层面：

$$\dot{\hat{w}} = \frac{i}{\hbar} (\hat{w} \hat{H} - \hat{H} \hat{w}) \tag{33}$$

证明方法是将上式左乘 $\langle \psi_m |$ ，右乘 $|\psi_n\rangle$ ，利用它们是 $\hat{H}$ 的特征状态的性质，以及哈密顿算符的特征状态基是完备的性质即可。如果想要 $\hat{w}$ 平稳，则等价地应该有：

$$[\hat{w}, \hat{H}] = 0 \tag{34}$$

即二者可交换，这继而蕴含二者可同时对角化，所以 $\hat{w}$ 的矩阵形式仅有对角元素非零，这就导致子系统物理量的平均值退化为：

$$\langle f \rangle = \sum_n w_{n,n} \langle \psi_n | \hat{f} | \psi_n \rangle \tag{35}$$

类比上一节中关于独立子系统的推导，并联系 $w$ 应和运动积分相关的事实，可以类似地得出，对于一个子系统而言：

$$\ln w_{n,n} = \alpha + \beta E_n \tag{36}$$

沿用微正则的假设，我们现在需要处理 $\hat{H}$ 的特征状态退化的问题，退化的存在使得 $w_{n,n}$ 上的分布需要进行重新累积，才能得到能量谱上的正确分布。这要求我们对于相空间以 $E$ 为参

照进行划分，譬如在以动量空间为相空间的情形下，这种划分就是同心球壳，但在有不均匀外场的情况下，情况会变得复杂起来。在微正则的假设下，所有满足 $E_{q,p} = E_0$ 的量子态都有相同的概率测度，而 $E_{q,p} \neq E_0$ 的量子态概率测度为零。以 $\Gamma(E)$ 记相空间中，能量在 $E$ 以下的量子态数量，则它在能量上引导了一个分布，在 $E$ 处微分，可得到在 $(E, E + \Delta E)$ 中的量子态数量为：

$$\Delta\Gamma = \frac{d\Gamma(E)}{dE} \Delta E \quad (37)$$

从已经可计算的 $w_{n,n}$ 出发，我们想要变换得到在能量上的观测密度 $W(E)$ ，直观上，这需要取出在 $(E, E + \Delta E)$ 中所有的量子态，并将这些量子态所对应的 $w$ 全部相加即可。考虑到 $w_{n,n}$ 又仅仅和 $E_n$ 相关，那么 $(E, E + \Delta E)$ 中所有的量子态，因为具有相同的能量，所以具有相同的 $w_{n,n}$ ，则求和转化为乘积，而这球壳中量子态的数量由之前一式给出：

$$W(E)\Delta E = w(E) \cdot \frac{d\Gamma(E)}{dE} \Delta E \quad (38)$$

在微正则假设下，全部的概率密度都必须集中在 $(E, E + \Delta E)$ 的壳内，所以：

$$W(E)\Delta E = w(E)\Delta\Gamma = 1 \quad (39)$$

到目前为止，关于量子态的讨论是脱离经典图景中的相空间的。二者之间的联系是：在以广义坐标和动量表征的相空间中，每个体积为 $(2\pi\hbar)^s$ 的微元空间对应一个量子态，所以可以在经典图景的相空间中对量子态进行计数，只需要在其中构造出等能量流形再划分即可。可以粗略地认为：

$$\Delta\Gamma = \frac{\Delta q \Delta p}{(2\pi\hbar)^s} \quad (40)$$

### 2.3 熵

沿用上一节末尾的讨论，一个子系统的熵定义为：

$$S = \ln \Delta\Gamma = \ln \frac{\Delta q \Delta p}{(2\pi\hbar)^s} \quad (41)$$

因为  $\Delta\Gamma \geq 1$ ，所以子系统的熵不会是负的。在微正则假设下：

$$S = \ln \Delta\Gamma = -\ln w(E) \quad (42)$$

在更一般的情形中，应该有：

$$\begin{aligned} S &= \langle -\ln w(E) \rangle \\ &= -\sum_n w_n \ln w_n \\ &= -\text{tr}(\hat{w} \ln \hat{w}) \end{aligned} \quad (43)$$

在这更一般地情形中，子系统的能量可能有所浮动，所以需要包括更多的一系列的特征状态。经典图景中类似地：

$$S = -\langle \ln [(2\pi\hbar)^s \rho] \rangle \quad (44)$$

现在考虑将一个整体的系统划分为一系列独立的子系统，分别对应统计权重  $\Delta\Gamma_1, \Delta\Gamma_2, \dots$ ，则母系统对应的量子态数量应为：  $\prod_a \Delta\Gamma_a$ ，而熵则相应是可加的。

现在将微正则假设应用于这个母系统，当各个子系统的能量在  $E_a^*$  附近进行微小的涨落  $\Delta E_a^*$ ，

并且不太影响母系统总能量时，它们总共对应的量子态数量为：

$$\begin{aligned}
 \prod_a \frac{d\Gamma_a}{dE_a}(E_a^*) \Delta E_a &= \prod_a \frac{\Delta\Gamma_a(E_a^*)}{\Delta E_a^*} \Delta E_a \\
 &= \exp\{S(E_1^*, E_2^*, \dots)\} \prod_a \frac{\Delta E_a}{\Delta E_a^*} \\
 &\propto \exp\{S(E_1^*, E_2^*, \dots)\} \prod_a \Delta E_a
 \end{aligned} \tag{45}$$

而我们又知道，每个子系统的能量都在 $E_a$ 为均值附近涨落，所以 $S(*)$ 应该在 $E_a^* = \bar{E}_a$ 附近取到最大值，因为这附近的区域，近似为微正则分布时，由 $w$ 给出最多的量子态。总结起来，母系统最可能处于的状态，对应该系统的熵 $S$ 最大的状态。这就是熵增长定律的一种表述。

另一个推论是，给定一个能量区间 $(E, E + \Delta E)$ 时，对应其中量子态数目 $\Delta\Gamma = \exp(S)$ ，则其中各个能量态之间的平均距离为：

$$D(E) = \Delta E \cdot \exp(-S(E)) \tag{46}$$

当系统很大时，系统的熵也相对更大，所以能量态之间的平均距离更小，能量谱更稠密。

统计物理里的熵和信息论中的信源熵可以进行很密切的关联，首先考虑微正则情况，此时熵可以通过相空间体积定义，也可以通过 $w$ 定义，假设此时子系统的能量为 $E_0$ ，且拥有总共 $\Delta\Gamma_0$ 个量子态：

$$S = \ln \Delta\Gamma_0 = -\ln w_0 \tag{47}$$

这个熵和一个拥有 $\Delta\Gamma_0$ 个符号，且各个符号输出概率相同的信源的信源熵相同：

$$\begin{aligned}
 S_I &= \sum_{i=1}^{\Delta\Gamma_0} -p_i \ln p_i \\
 &= \Delta\Gamma_0 \frac{1}{\Delta\Gamma_0} \ln \Delta\Gamma_0 \\
 &= S
 \end{aligned} \tag{48}$$

当子系统拥有不同的能级时，微正则情况中利用相空间体积的定义不在适用，考虑拥有两个能级 $E_1, E_2$ 的情况，且分别对应量子态数量 $\Delta\Gamma_1, \Delta\Gamma_2$ ，被观测概率分别为 $p_1, p_2$ 。且需知，任何对应能量为 $E_i$ 的特征状态被观测到的概率都必为 $p_i$ ，不会出现能量符合但是未被包含进入系综的情况，因为 $\ln w = \alpha + \beta E$ 。我们有归一化条件：

$$\Delta\Gamma_1 p_1 + \Delta\Gamma_2 p_2 = 1 \tag{49}$$

系统被发现在 $E_i$ 的概率为 $\Delta\Gamma_i p_i$ ，此时系统的熵，根据定义，应该为：

$$\begin{aligned}
 S &= - \sum_{i=1}^{\Delta\Gamma_1} p_1 \ln p_1 - \sum_{i=1}^{\Delta\Gamma_2} p_2 \ln p_2 \\
 &= -\Delta\Gamma_1 p_1 \ln p_1 - \Delta\Gamma_2 p_2 \ln p_2 \\
 &= p(E_1)(-\ln p_1) + p(E_2)(-\ln p_2) \\
 &= \mathbb{E}_{p(E_i)} [S_i] - [p(E_1) \ln p(E_1) + p(E_2) \ln p(E_2)]
 \end{aligned} \tag{50}$$

其中 $S_i$ 为对应能量为 $E_i$ 的微正则系综的熵。换言之，具有混合能量态的系统（非微正则系统）的热力学熵是一系列微正则系综的熵的期望，再加上选择能量的信息熵，引导该期望的分布是观测该系统坍塌到各个能量态的概率。从信息论的角度，这等同于一个混合信源，混合信源发信的过程由：选择子信源，子信源发信组成，则整体的熵，即不确定性由两个部分构成：选择信源的不确定性、各个信源的不确定性。

上式作为非微正则系统的熵，写出了以熵的相空间定义导出的一般性情况。

### 3 热力学

#### 3.1 温度，热力学第零、第二定律

我们从系统整体的熵增加原理开始，考虑将整个系统划分为两个子部分，分别具有 $E_1, E_2, S_1, S_2$ ，而 $S = S_1 + S_2$ ， $E = E_1 + E_2$ ，给定 $E$ ，并且最大化熵，这要求：

$$\begin{aligned}\frac{dS}{dE_1} &= \frac{dS_1}{dE_1} + \frac{dS_2}{dE_1} \\ &= \frac{dS_1}{dE_1} + \frac{dS_1}{dE_2} \frac{dE_2}{dE_1} \\ &= \frac{dS_1}{dE_1} - \frac{dS_2}{dE_2} = 0\end{aligned}\tag{51}$$

定义系统的（绝对）温度为：

$$\frac{1}{T} = \frac{dS_1}{dE_1} = \frac{dS_2}{dE_2}\tag{52}$$

是属于这个系统平衡时，本征的状态。如果一个系统的两个子系统未达到平衡，则根据熵的增加性，应该有系统总熵对时间的微商大于零：

$$\begin{aligned}\frac{dS}{dt} &= \frac{dS_1}{dE_1} \frac{dE_1}{dt} + \frac{dS_2}{dE_2} \frac{dE_2}{dt} \\ &= \left( \frac{1}{T_1} - \frac{1}{T_2} \right) \frac{dE_1}{dt} > 0\end{aligned}\tag{53}$$

所以如果 $T_1 > T_2$ ，则应有 $\frac{dE_1}{dt} < 0$ ，即能量从子系统1流向子系统2，直到二者的温度相等。以上两个结论分别对应热力学第零、第一定律。

值得回顾一下熵如何作为内能的函数 $S(E)$ ，在量子的场景中，如果子系统的内能已知，则其可以近似为以 $E$ 为能量的微正则系统，此时通过在相空间中对 $\Gamma$ 对能量在 $E$ 处微分，就可以得到 $\Delta\Gamma$ ，继而得到熵。在经典场景中，一样可以以 $E$ 引导分布 $\rho$ 并代入熵在经典场景中的定义式即可。



### 3.2 绝热过程和压强

绝热过程是这样一种过程，外界对于系统产生缓慢的影响，其影响造成改变的速度小于系统达到即时平衡的速度，此时系统的哈密顿可以写成：

$$E(q, p, \lambda(t)) \quad (54)$$

绝热过程的熵是不变的，考虑熵的变化率应该取决于外场的变化率：

$$\frac{dS}{dt} = f\left(\frac{d\lambda}{dt}\right) \quad (55)$$

将未知的取决关系按照级数展开：

$$\frac{dS}{dt} = C_0 + C_1 \frac{d\lambda}{dt} + C_2 \left(\frac{d\lambda}{dt}\right)^2 + \dots \quad (56)$$

而其中  $C_0 = C_1 = 0$ ，因为如果  $\dot{\lambda}$  为零，则平衡时应有  $\dot{S} = 0$ ，同时  $\dot{\lambda}$  可能小于零，但是  $\dot{S} \geq 0$ ，所以当外界变化缓慢时，保留二次项的近似：

$$\frac{dS}{d\lambda} = C_2 \frac{d\lambda}{dt} \quad (57)$$

故  $\dot{\lambda} \rightarrow 0$  时， $\frac{dS}{d\lambda} \rightarrow 0$ ，即系统的熵在绝热过程中不变。利用绝热过程的哈密顿量形式，考察  $\dot{E}$ ，其中  $E$  是平均的哈密顿量，利用正则关系：

$$\frac{dE(q, p, \lambda)}{dt} = \frac{\partial E(q, p, \lambda)}{\partial t} = \frac{\partial E(q, p, \lambda)}{\partial \lambda} \frac{\partial \lambda}{\partial t} \quad (58)$$

所以：

$$\begin{aligned}
 \frac{dE}{dt} &= \frac{\int E(q, p, \lambda(t)) \rho(q, p, \lambda(t)) dq dp}{dt} \\
 &= \int \frac{\partial E(q, p, \lambda)}{\partial \lambda} \frac{\partial \lambda}{\partial t} \rho(q, p, \lambda(t)) dq dp \\
 &= \dot{\lambda} \frac{\int E(q, p, \lambda(t)) \rho(q, p, \lambda(t)) dq dp}{\partial \lambda} \\
 &= \dot{\lambda} \frac{\partial \bar{E}(\lambda)}{\partial \lambda}
 \end{aligned} \tag{59}$$

其中 $E(\lambda)$ 为外场参数为 $\lambda$ 时能量的平均值。另一方面，能量 $E$ 是熵 $S$ 和参量 $\lambda$ 的函数，所以：

$$\frac{dE}{dt} = \frac{\partial E}{\partial S} \frac{\partial S}{\partial t} + \frac{\partial E}{\partial \lambda} \frac{\partial \lambda}{\partial t} \tag{60}$$

而绝热过程中熵不变，联立上两式：

$$\frac{\partial \bar{E}(\lambda)}{\partial \lambda} = \left( \frac{\partial E}{\partial \lambda} \right)_S \tag{61}$$

一个应用该式的例子就是求系统中的压强，考虑系统中微元面积 $d\mathbf{s}$ 上的作用力，根据力的定义，它应该等于：

$$\mathbf{F} = - \frac{\partial E(q, p, \mathbf{r})}{\partial \mathbf{r}} \tag{62}$$

所以平均的压力为：

$$\begin{aligned}
 \bar{\mathbf{F}} &= - \frac{\partial \bar{E}(q, p, \mathbf{r})}{\partial \mathbf{r}} \\
 &= - \left( \frac{\partial E}{\partial \mathbf{r}} \right)_S \\
 &= - \left( \frac{\partial E}{\partial V} \frac{\partial V}{\partial \mathbf{r}} \right)_S \\
 &= - \left( \frac{\partial E}{\partial V} \right)_S \frac{\partial V}{\partial \mathbf{r}} \\
 &= - \left( \frac{\partial E}{\partial V} \right)_S d\mathbf{s}
 \end{aligned} \tag{63}$$

定义压强为：

$$P = - \left( \frac{\partial E}{\partial V} \right)_S \tag{64}$$

则帕斯卡定律，即面元上的压力指向法向，并且和面元大小成正比成立。

根据熵和能量的可加性，应该知道，给定能量（熵）时，熵（能量）仅和体积相关。

一个利于理解这一点的方法是联想微正则状态中的量子气体，对于量子气体而言，如果给定能量，则其熵只和该能量对应的量子态的数量相关，而这一数量，在自由气体的情况下，在动量上有所限制（在球壳上），但是在位置上的限制完全由气体的体积给定。相反地，如果给定熵，则系统的量子态数量被限制了，而给定不同的体积，就可以类似地影响动量，继而影响能量。

【可以这样联系一种理想图景，三维空间的三个独立坐标对应动量的三个维度，量子态数量就是球壳的体积，而宏观气体的体积，以理想势垒的形式，在这个球壳中规定一个天顶角。】

在这一关系可以被理解以后，自然地有：

$$\begin{aligned} dE &= \frac{\partial E}{\partial S} dS + \frac{\partial E}{\partial V} dV \\ &= T dS - P dV \end{aligned} \quad (65)$$

对于相接触的两个系统，可以得出它们在平衡时，平衡处压强相等的结论，分别记二者的体积为 $V_1, V_2$ 且 $V_1 + V_2 = V$ ，则 $V_1$ 应变化直到总体的熵取极大值：

$$\begin{aligned} \frac{\partial S}{\partial V_1} &= \frac{\partial S_1}{\partial V_1} + \frac{\partial S_2}{\partial V_1} \\ &= \frac{\partial S_1}{\partial V_1} - \frac{\partial S_2}{\partial V_2} = 0 \end{aligned} \quad (66)$$

但是对于子系统而言有：

$$\frac{\partial S_i}{\partial V_i} = \frac{P_i}{T_i} \quad (67)$$

因为：

$$dS = \frac{1}{T} dE + \frac{P}{T} dV \quad (68)$$

所以热力学第零定律给出压强的均衡性。

### 3.3 热力学第一定律，麦克斯韦关系的其他元素

微元功的定义从直观出发：

$$dR = -PdV \quad (69)$$

其中约定外界对系统做功时 $R$ 是正的，此时 $dV$ 是一个负值。系统内能的改变一方面来自于功，剩余的部分被定义为热量：

$$dE = dR + dQ \quad (70)$$

这就是热力学第一定律，它直接给出：

$$dQ = dE + PdV \quad (71)$$

再利用 $dE = TdS - PdV$ ，就有：

$$dQ = TdS \quad (72)$$

比热的定义是升高温度时吸收的热量，也即热量对温度的微商，通常分别考虑系统体积一定和压强一定的情况：

$$C_V = T \left( \frac{\partial S}{\partial T} \right)_V, C_P = T \left( \frac{\partial S}{\partial T} \right)_P, \quad (73)$$

在实际的非理想热过程中，除了物理上能量的交换以外，可能还有其他形式（譬如化学）能量的变化，导致实际的熵增速率比吸热带来的更多，这对应于不可逆过程：

$$\delta Q < T\delta S \quad (74)$$

这一点也可以从卡诺热机的理论出发来证得。【\*\*\*\*\*】

对于定容过程，即 $dV = 0$ ，有 $dQ = dE$ ，而对于恒压过程，即 $dP = 0$ ，可把热量增量写

成某个宏观量的微分：

$$\begin{aligned}
 dQ &= dE + PdV + VdP \\
 &= d(E + PV) \\
 &= dW
 \end{aligned} \tag{75}$$

其中 $W = E + PV$ 是系统的焓，在恒压变化中，焓的变化就是系统吸收的热量。依旧可以利用 $dE = TdS - PdV$ 化约：

$$dW = TdS + VdP \tag{76}$$

即 $W$ 可以写成熵和压强的函数，两个偏导数分别给出系统的温度和体积。如果恒压变换同时还是热绝缘的（这是很容易实现的，只需要将绝热材料以恒定的压强封闭气体系统），则系统的焓是一个常量。

关系 $dE = TdS - PdV$ 还允许我们写出：

$$\begin{aligned}
 C_V &= T \left( \frac{\partial S}{\partial T} \right)_V \\
 &= T \left( \frac{1}{T} \frac{\partial E}{\partial T} \right)_V \\
 &= \left( \frac{\partial E}{\partial T} \right)_V
 \end{aligned} \tag{77}$$

类似地 $C_P = \left( \frac{\partial W}{\partial T} \right)_P$ ，因为：

$$\begin{aligned}
 C_P &= T \left( \frac{\partial S}{\partial T} \right)_P \\
 &= T \left( \frac{1}{T} \frac{\partial W}{\partial T} \right)_P \\
 &= \left( \frac{\partial W}{\partial T} \right)_P
 \end{aligned} \tag{78}$$

从 $dE = TdS - PdV$ 和 $dW = TdS + VdP$ 中可以认为焓之于恒压过程，类比于能量之于

定容过程。从能量到焓的过程可以认为是一种勒让德变换：

$$\begin{aligned} dE &= TdS - PdV \\ dE &= TdS - [d(PV) - VdP] \\ d(E + PV) &= TdS + VdP \end{aligned} \tag{79}$$

注意这一变换对于热过程的宏观条件（绝热、等压、定容等等）没有任何预设。除了对 $(P, V)$ 应用勒让德变换，也可以对于 $(T, S)$ 应用变换：

$$\begin{aligned} dE &= TdS - PdV \\ dE &= [d(TS) - SdT] - PdV \\ d(E - TS) &= SdT - PdV \\ d(E - TS + PV) &= SdT + VdP \end{aligned} \tag{80}$$

其中 $F = E - TS$ 和 $\Phi = E - TS + PV$ 分别是系统的自由能（又称Helmholtz自由能）和热力学势（又称Gibbs势能）

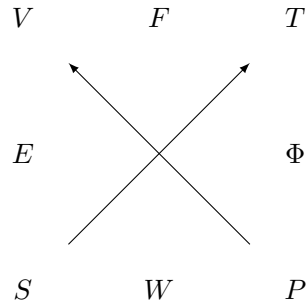


Figure 1: 四个宏观热力学量和四个全微分

利用简单的微分就可以得出它们之间进一步的联系，譬如：

$$\begin{aligned}
 E &= F + TS \\
 &= F - T \left( \frac{\partial F}{\partial T} \right)_V \\
 &= -T^2 \left( \frac{\partial}{\partial T} \frac{F}{T} \right)_V
 \end{aligned} \tag{81}$$

完全类似地：

$$W = -T^2 \left( \frac{\partial}{\partial T} \frac{\Phi}{T} \right)_P \tag{82}$$

当系统所处在外场有其他的一系列参量 $\{\lambda_i\}_i$ 时，其哈密顿量的微分有其他的余项：

$$dE = TdS - PdV + \sum_i \Lambda_i d\lambda_i \tag{83}$$

而对于 $(P, V)$ 或者 $(S, T)$ 的勒让德变换不影响这些额外参数，所以有：

$$\frac{\partial \bar{E}(q, p, \lambda)}{\partial \lambda} = \left( \frac{\partial E}{\partial \lambda} \right)_{S, V} = \left( \frac{\partial F}{\partial \lambda} \right)_{T, V} \tag{84}$$

可以改变 $E$ 或者 $F$ 为其他的全微分，并更改对应的不变量。

举例而言，若已知体系自由能的表达式，要求粒子的平均动能，我们首先有：

$$\frac{\partial E(q, p, m)}{\partial m} = -\frac{1}{m} K(p) \tag{85}$$

因为能量中势能的部分和质量无关，而动能部分 $K = \frac{p^2}{2m}$ ，所以对于平均动能：

$$\bar{K}(p) = -m \cdot \frac{\partial \bar{E}(q, p, m)}{\partial m} = -m \cdot \left( \frac{\partial F}{\partial m} \right)_{T, V} \tag{86}$$

当 $\lambda$ 的变化不大时，四个全微分的增量近似都相等：

$$\delta E_{S,V} = \delta F_{T,V} = \dots \quad (87)$$

这一性质叫做**小增量定理**。

几个全微分形式允许我们在不同的热过程中挑选合适的量来简化表达，联立 $dQ \leq TdS$ ， $dQ = dE + PdV$ ，有：

$$dE \leq TdS - PdV \quad (88)$$

这一不可逆过程带来的不等关系可以带来：

$$\frac{dE}{dt} \leq T \frac{dS}{dt} - P \frac{dV}{dt} \quad (89)$$

所以对于 $(T, V)$ 恒定的热过程，有：

$$\frac{dF}{dt} = \frac{d(E - TS)}{dt} = \dot{E} - T\dot{S} - S\dot{T} \leq 0 \quad (90)$$

即对于 $(T, V)$ 恒定的热过程，系统的自由能会自发地下降到极小值。

类似地，对于 $(T, P)$ 恒定的热过程，系统的热力学势会自发地下降到极小值。

但是对于能量和焓而言，想要确保其自发地下降需要确保系统的熵恒定，这从实验的角度较难保证。