

---

# Statistical Mechanics

---

赡科材

中国科学技术大学 化学物理系

2023 年 7 月 1 日

---



# 目录



# 前言

统计力学将热力学问题归结于 **概率与统计推断**，以等概率假设与遍历性假设为基础，建立起一整套关于微观-宏观相互联系的桥梁。通过一定的统计推断，我们可以从众多自由度中提取出一些重要信息，而忽视一些细节，或者无关紧要的部分 [?]

虽然我们打算忽略一些细节，但是还是得来看看这些细节都是怎么一回事。经典力学以相空间和哈密顿函数为基础，以 **正则变量** 的时间演化给出力学系统的时间演化。这些正则变量服从哈密顿方程：

$$\begin{aligned}\dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}\end{aligned}$$

不过，要想求解如此庞大的耦合微分方程系统是很困难的一件事——例如采样，而且是对  $10^{23}$  量级的粒子数！既然庞加莱早已证明，连最简单的三体问题都没有解析解，且对初值敏感，**我们又有信心认为对于  $10^{23}$  量级的结果是可靠的呢？**

好消息是，统计力学的遍历性假设将会让这些烦人的细节从我们的面前消失。考虑一个宏观小而微观大的子系统，它与周围的环境进行着复杂紊乱的相互作用，所以我们有理由认为：在足够长的时间间隔内子系统在自身所有可能的状态中经历足够多的次数 [?]。同时，这一假设也让我们相信：某个子系统的统计分布与同一系统的其他任意小部分的初始状态无关，因为从任何初始状态出发，这个系统总是将所有可能的状态都走了足够多遍，以至于这种初始状态的影响在足够长的时间内，被系统其余的、更为广大的部分的影响所完全消除。换句话说，**每一个状态都可以取为初始状态**。

所以我们看到，统计力学对一个物理系统的预言具有概率的特征。但必须指出的一点是，这样的概率绝非系统所固有，而是因为得到这些结论所依据的条件远比完整的力学描述所需要的少得多——例如我们可以仅用三个宏观可测的状态参量就能完全确定一个宏观系统，还能确定系统其他的状态参量的值，而这三个自由度实际上将天文数字的微观自由度都统摄其下。

另一个问题是 **微观可逆性与宏观不可逆性的矛盾**——如果用  $-t$  替代求解结果中的  $t$ （做一个时间反演，或者说把录像带倒着放）仍然是演化方程的一个可能的解，但我们从来没有观察到滴进池塘的墨水又重新聚集成刚滴进去时的形状。这样一来，**时间的流逝似乎就有了确定的方向**。这样

的“热力学时间”和我们其他地方作为正则变换参量（Hamilton-Jacobi 理论）或者酉变换参量（量子力学传播子）所引入的时间有什么差别？

总之，统计力学是作者本科以来最喜欢的一门理论课，它是一个关于“大”的理论。固然对“小”的研究也同样迷人（量子力学、第一性原理计算、单分子反应机理的探究等等），对这些细节的研究加深了我们对自然界的微观认识，使得我们对自然界究竟是怎样运转的有了更多的了解，甚至能在原子、电子、夸克尺度上让基本力服从我们的意志，但它能指引我们走多远呢？世界的本源有小有大，而我选择相信 **More Is Different**。具有巨大自由度数目的系统虽然系统与单一粒子遵循同样的力学规律，但是自由度的巨大却导致性质上的全新规律性。而这样的规律，是在微观层面意识不到的。

不识庐山真面目，只缘身在此山中。进入大数据时代，甚至现在的 new bing 和 ChatGPT 所引领的大语言模型时代，科学界的主要矛盾也在变化。这是统计模型大放光彩的时代，也是科技革命的浪潮之巅。

统计力学将会爆发出的能量，我想是不言而喻的。

# 第一章 系综理论的基本原理

经典力学通过相空间中的点  $(q, p)$  来描述一个系统的某个状态。它们的演化满足正则方程

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad (1.1)$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (1.2)$$

现在，我们要考虑的不是一个状态，而是一大群状态——在单一瞬间同时考虑大量系统，他们全部是给定系统的某种“思维复本”——其特性由与原系统一样的宏观态来表征，但极其自然地处在所有各种可能的微观态中 [?]

所以，所谓系综是这样一个集合，它是相空间中满足所有约束条件（粒子位置有限、总能量有限）的一个区域——这里默认是连续的，因为一个宏观态统摄下的微观态是如此之多，以至于两个相邻微观态之间的间隔足够小。

现在，这一大群具有与宏观态相同宏观性质的微观态随着时间的变化都会在相空间内画出一条条轨迹，同时这个集合的位置和形状也会随着时间不断变化。和在力学中所遵循的思路相似，我们希望找出这一演化的运动积分，而 刘维尔定理 则能完成这个任务。

## 1.1 刘维尔定理

虽然每一个宏观态统摄下的微观态都具有相同的宏观性质，但它们在系综所包括的区域内的分布可能不是均匀的，有疏密之分。这驱使我们定义 相点密度  $\rho(q, p, t)$  如下：

$$\rho = \lim_{\Delta V \rightarrow 0} \frac{\Delta N}{\Delta V} \quad (1.3)$$

这里的定义是离散的，但基于许多考虑，我们更喜欢一个连续、归一化的概率密度  $\rho(q, p, t)$ ，满足

$$\int_{\Omega} \rho(q, p, t) d\omega = 1 \quad (1.4)$$

使得物理量的系综平均可以表示为

$$\langle f \rangle = \int_{\Omega} \rho(q, p, t) f(q, p, t) d\omega \quad (1.5)$$

我们也可以这样理解相点密度  $\rho$ ：在系综内任意划定一个区域  $d\omega$ ，在某一相点在系综内进行遍历的过程中，在这个区域内找到这个相点的概率为  $\rho d\omega$ ，或者说在遍历过程中，系统的代表点在此处停留的时间与总时间的比  $\Delta t/T$  收敛于  $\rho d\omega$ 。

代表点的演化在相空间中对应于一个正则变换在时间上的连续伸展 [?], 而正则变换的 **Jacobi** 行列式的绝对值恒等于 1 (见附录)，所以我们期望：系综的时间演化服从流体的连续性方程

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (1.6)$$

事实也确实如此。代入哈密顿正则方程，即可得到著名的 **刘维尔定理**：

$$\frac{d\rho}{dt} \equiv \frac{\partial \rho}{\partial t} + [\rho, H] = 0 \quad (1.7)$$

根据这个定理，正如随同一个相点一道运动的观察者所看到的那样，相点的局部密度随时间保持恒定，是运动积分。因此，这群相点在相空间中的运动，从本质上说，跟不可压缩流体在物理空间中的运动是一样的！

### 反思与质疑

当看到式子  $[\rho, H]$ ，一些读者可能会有这样的疑问： $\rho$  是系综的局部性质， $\rho$  在某一点的值是与该点周围的相点分布情况有关，而  $H$  则是这个相点的哈密顿量。那么二者的对象不同，怎么能做 Poisson 括号呢？

对于这个问题，我觉得可以这样回答： $H$  并不只是某个相点的哈密顿量，因为根据正则方程，我们实际上在整个相空间内都定义了一个相场 [?]

$$\dot{\xi}_i = \Omega_{ij} \frac{\partial H}{\partial \xi_j} \equiv \Delta H, \quad \Omega = \begin{bmatrix} \mathbf{0} & I \\ -I & \mathbf{0} \end{bmatrix}$$

这个场穿过了系综内所有的相点，所有的相点都沿着这个场进行运动，那么这个场的对象当然就是整个系综了。

## 1.2 分布函数

从刘维尔定理可以知道，既然分布函数  $\rho$  是一个物理量，它必然是变量  $q, p, t$  的函数，并且以某一点所代表的子系统在运动时保持不变。所以，它应当可以被写作系统所有独立运动积分的一些组合。

实际上，分布函数并不依赖于所有的独立运动积分，我们只需要考虑到这样一件事：根据统计独立性，两个子系统的组合的分布函数  $\rho_{12}$  应当等于  $\rho_1$  与  $\rho_2$  的乘积，所以

$$\ln \rho_{12} = \ln \rho_1 + \ln \rho_2 \quad (1.8)$$

即分布函数的对数是可加性的量。所以我们可以得出结论：**分布函数不仅是运动积分，而且还是可加的运动积分**！而力学系统可加的运动积分只有七个：能量、动量的三个分量以及角动量的三个分



量，所以分布函数形式上应该能写为它们的线性组合

$$\ln \rho_a = \alpha_a + \beta E_a(q, p) + \gamma \cdot P_a(q, p) + \delta \cdot M_a(q, p) \quad (1.9)$$

总之，可加性的运动积分的值完全确定了系统的统计性质，也就是说完全确定了它的任何子系统的统计分布，因而同时也确定了子系统的任意物理量的平均值。**正是这七个独立的可加运动积分代替了在用力学方法处理问题所需要的多得不可想象的初始条件。**

上面的讨论使得我们可以直接构造出一个适用于描述系统的统计性质的分布函数——既然不可相加的运动积分的值不会对系统的统计性质造成影响，那么任意函数  $\rho$  只要它仅依赖于可加运动积分，并且满足刘维尔定理，就可以用来干这件事。

例如，我们总是可以通过设定合适的边界条件，或者选取合适的坐标系，使得系统的动量与角动量是一个固定的常数，从而可以合并为常数项  $\alpha_a$  中。事实上总是可以证明 [?]: 在热力学平衡状态下，闭合系统只可能整体做匀速的平动和转动。这样一来，密度分布函数就只依赖于系统的能量了。

现在，如果我们关注的是系统的某个稳定状态，那么自然而然地就要求  $\rho$  不明显地依赖于时间——这就导致  $[\rho, H] = 0$ 。不妨假设函数  $\rho$  对坐标  $(q, p)$  的依存关系只是通过对哈密顿函数  $H(q, p)$  来实现，即

$$\rho(q, p) = \rho[H(q, p)] \quad (1.10)$$

在随后对正则系综的讨论中我们将看到，在这类系综中最自然的选择，是以下的密度函数

$$\rho(q, p) \propto \exp \left[ -\frac{H(q, p)}{kT} \right] \quad (1.11)$$

其所定义的系统即为 **正则系综**。

当然，最简单的分布函数莫过于一个不依赖于任何坐标动量的常值函数了，它对应的就是接下来的主题：**微正则系综**。

## 1.3 微正则系综

### 1.3.1 热力学何以成为可能

任何知识的结构都由基本假设和逻辑推断组成。热力学作为自然科学的一部分也是如此。统计力学既要为热力学立法，要扫清以往物理化学 I 中的蒙昧，那就得提出使得这一切成为可能的 **先天综合判断** [?]

**一切的一切都迫使我们相信：从微观状态数  $\Omega$  的量值和它依赖于参数  $N, V$  和  $E$  的性质，我们可以推导出给定系统的全部热力学特性。**

也就是说：配容数是第一性的，是自然界的最优化目标函数。

我们后面的讨论都从这里出发。既然我们认为自然界为配容数选择的决策参量有三个，分别是  $N, V, E$ ：  $\ln \Omega(N, V, E)$ ，那么它的三个偏导数在优化过程中有非常重要的作用，它们代表了配容数上升的倾向有多大：

$$\beta = \left( \frac{\partial \ln \Omega}{\partial E} \right)_{N, V}, \quad \eta = \left( \frac{\partial \ln \Omega}{\partial V} \right)_{N, E}, \quad \zeta = \left( \frac{\partial \ln \Omega}{\partial N} \right)_{V, E} \quad (1.12)$$

这个最优化问题显然还受到一定的条件约束：总粒子个数、总体积、总能量不能变。这样一来，系综在相空间中所占据的区域就有了一定的边界。所以，现在我们写出自然界每时每刻都在“求解”的约束最优化问题：

$$\begin{aligned} & \max \Omega(N, V, E) \\ & \text{s.t. } \sum n_i = \mathcal{N}, \quad \sum n_i E_i = \mathcal{E} \end{aligned} \quad (1.13)$$

再回到热力学。热力学起初是人类根据完全的宏观经验总结出的一些规律（而没有涉及任何微观的分子原子或者能级之类的概念）。经过经年的实践，人类通过感官的认识，将配容数的三个重要偏导数所代表的内涵分别都取了名字：**温度、压强以及化学势**——回忆热力学第一定律与第二定律的联合方程  $dE = TdS - PdV + \mu dN$  以及各种热力学平衡条件。这些是人类最初认识“热及其动力学”(Thermodynamics) 的出发点。所以虽然它采用了很多数学记号，看起来具有坚实的数学基础，但它终究是一个完全唯像的理论，这些偏导数的正确性只能通过实验检验，而不是逻辑（全是综合判断，而少见分析判断）。

但情况有了一些变化。我们认识到熵这个量似乎和其他状态参量不同，因为自然界总是要熵增的方向发展，它的地位应当和其他状态参量有所差别。将熵还原为配容数之后，我们相信，或者说信仰：自然界总是想运动至配容数更大的地方。从这里出发，我们希望将统计学引入我们的理论体系中，并为热力学提供一个合理的解释。

为了凸显配容数以及熵的第一性，我们将  $dS$  写到等式的左边：

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN \quad (1.14)$$

带入熵的定义式  $S \equiv k \ln \Omega$ ，对应可知

$$\beta = \frac{1}{kT}, \quad \eta = \frac{P}{kT}, \quad \zeta = -\frac{\mu}{kT} \quad (1.15)$$

请仔细品味与物理化学 I 或者其他哪本热力学教材的差别：这里是以微观状态数  $\Omega$  及其最优化为基础的，而能量、体积和粒子数则沦为了一个决策变量，温度、压强与化学势则沦为了对这三个决策变量的偏导数——不再像从前那样具有平等的地位（这里指的是任何单组分平衡态可以用任意三个状态参量完全确定，其中必须要有一个广度性质）。

由此可见，[统计力学之路，莫非陟降二途](#)。陟，意指向峰巅攀爬，即导出基本定律，同时厘清各种热力学概念；降，下山，即将基本原理应用于诸多情形。下面我们先来爬山。非常明显，基于以上的反思，这第一个台阶应是计算  $\Omega(N, V, E)$  的表达式。

### 1.3.2 微正则系综的研究方法

从前文可知，选择决策参量  $(N, V, E)$  来描述一个正在执行优化过程的系统是一件非常自然的事，因为我们的目标是用它们表示出配容数  $\Omega$ 。对于粒子数  $N$  和体积  $V$  的约束不太能够导出新的关系，但是对能量的限制却可以：对于一个有确定能量的系统，它有许多的简并态。在这些简并态中，我们没有理由认为其中哪一个是比另外几个更优越——因为极小的扰动都会使得系统从一个态变为能量差不多相等的另一个态。

所以，我们为以  $(N, V, E)$  标记的系综中的每一个态都赋予相等的概率，即系综的任何代表点处于相空间中所允许的区域内的概率都是完全相等的，这就是 [等概率假设](#)，这样的系综称为 [微正则系综](#)。

虽然在微正则系综中，系统的宏观态由分子数  $N$ 、体积  $V$  和能量  $E$  来确定，但现在我们要求能量  $E$  可以在一个范围内波动——从  $E - \Delta/2$  到  $E + \Delta/2$ 。这样一来，系综区域被限制在一个“超壳层”内。微正则系综的密度函数为

$$\rho(q, p) = \begin{cases} \text{const}, & \text{for } E - \frac{1}{2}\Delta \leq H(q, p) \leq E + \frac{1}{2}\Delta \\ 0, & \text{else} \end{cases} \quad (1.16)$$

不难看出，对于微正则系综而言，某个物理量的时间平均与其系综平均可以交换积分次序，也就是说这两种求平均值的过程可以相互颠倒。而对于长时间的平均值，根据遍历性假设，系综内每一个成员都做了几乎相同的事，那么取系综平均值又变得无关紧要 [?]

由此可见，[一个物理量在微正则系综上的时间平均等于系综平均](#)。

对于系统的其他状态函数，我们写出热力学第一定律与第二定律联合公式的变形

$$dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN$$

带入  $S = k \ln \Omega$

$$d \ln \Omega = \frac{1}{kT}dE + \frac{p}{kT}dV - \frac{\mu}{kT}dN \quad (1.17)$$

同时有

$$d \ln \Omega = \left( \frac{\partial \ln \Omega}{\partial E} \right)_{V, N} dE + \left( \frac{\partial \ln \Omega}{\partial V} \right)_{N, E} dV + \left( \frac{\partial \ln \Omega}{\partial N} \right)_{E, V} dN \quad (1.18)$$

所以对应即可，比如

$$\beta = \frac{1}{kT} = \left( \frac{\partial \ln \Omega}{\partial E} \right)_{V, N}, \quad \eta = \frac{p}{kT} = \left( \frac{\partial \ln \Omega}{\partial V} \right)_{N, E}, \quad \zeta = -\frac{\mu}{kT} = \left( \frac{\partial \ln \Omega}{\partial N} \right)_{E, V} \quad (1.19)$$

由于  $\Omega$  是  $(N, V, E)$  的函数，那么这些偏导数也是  $(N, V, E)$  的函数。这样一来，根据这些等式 (??) 我们就可以得到各种宏观状态参量之间的关系，也就爬到半山腰了。

## 1.4 经典理想气体

作为一个爬山的例子，我们现在来推导经典理想气体所满足的 [能量均分定理](#) 和 [理想气体状态方程](#)：

$$\begin{aligned} PV &= Nk_B T \\ E &= \frac{3}{2} Nk_B T \end{aligned}$$

### 1.4.1 一些直觉

经典理想气体是一个 [由无相互作用粒子组成、且不考虑粒子内部结构的经典系统](#)。之所以考虑这样一个系统，是因为它可以对  $\Omega(N, V, E)$  提供一个明确的计算方法（虽然这种方法是渐进性质的）。

不仅如此，在进行具体计算之前，就已经可以通过一些直觉性的洞察，来获得系统的状态函数之间所满足的一些关系。比如倘若粒子中任何一个处于资用空间特定区域内的概率与其他粒子的位置完全无关，则该系统  $N$  个粒子空间分布方式的总数，将简单地等于各个粒子能够独立容纳于资用空间内的分布方式数的乘积，也即概率论中的 [独立性](#)。

所以，系统分布方式的总数与  $V$  的  $N$  次幂成正比：

$$\Omega(N, V, E) \propto V^N \quad (1.20)$$

与前文所述进行对比，易得

$$\frac{P}{T} \equiv k \left( \frac{\partial \ln \Omega(N, V, E)}{\partial V} \right)_{N, E} = k \frac{N}{V} + C \quad (1.21)$$

其中  $C$  为常数，若  $C = 0$ ，则可得到理想气体状态方程

$$PV = nRT \quad (1.22)$$

### 1.4.2 $\Omega$ 的计算

考虑以  $a$  为边长的三维势箱，若只有一个粒子： $N = 1, V = a^3, E = \epsilon$ 。以下使用国际单位制考虑量级，是为了能够方便地看出某些项是否可以忽略。根据量子力学，系统的能级为

$$\epsilon_{n_x n_y n_z} = \frac{h^2}{8ma^2} (n_x^2 + n_y^2 + n_z^2), \quad n_x, n_y, n_z \in \mathbb{N} \quad (1.23)$$

很自然地定义三维态空间，系统的状态离散而均匀地分布于态空间中。注意，所有的态只分布在一个卦限中，所以这里只有八分之一球体，而这个八分之一球体的体积则对应于能量小于一定值的态的个数。

如果系统中有  $N$  个粒子，不考虑它们之间的相互作用，系统总能量应对  $3N$  个自由度求和：

$$E = \sum_{j=1}^{3N} \epsilon_j = \frac{h^2}{8ma^2} \sum_{j=1}^{3N} n_j^2 \implies \sum_{j=1}^{3N} n_j^2 = \frac{8ma^2}{h^2} E \equiv E^* \quad (1.24)$$

我们要得到的微观配容数  $\Omega$  就等于方程 (??) 落在  $(E - \frac{1}{2}\Delta E, E + \frac{1}{2}\Delta E)$  区间中的所有整数解。从理论上可以证明，这个数渐进地等于

$$\lim_{N \rightarrow \infty} \Omega = \frac{\partial \Sigma}{\partial E} \Delta E \quad (1.25)$$

其中  $\Sigma$  是能量小于  $E$  所有能态的数量，而这个数量同时也渐进等于  $3N$  维球体的体积！顺带一提，这件事和概率论中的累积分布函数和概率密度函数的关系很相似：

$$\Sigma(E) = \int_0^E \Omega(E) dE, \quad \Omega(E) = \Sigma'(E)$$

$3N$  维球体的体积由

$$V_{3N} = \frac{\pi^{3N/2}}{\Gamma(3N/2 + 1)} R^{3N} \quad (1.26)$$

确定，但是计算累计能态数时，注意需要除以卦限因子  $2^{3N}$  和不可分辨因子  $N!$ ，带入  $E^*$  则有

$$\Sigma(E) = \left( \frac{V}{h^3} \right)^N \frac{(2\pi m E)^{3N/2}}{N!(3N/2)!} \quad (1.27)$$

取对数，引入斯特林公式近似  $\ln N! \simeq n \ln N - N$  得到

$$\ln N = N \ln \left[ \frac{V}{Nh^3} \left( \frac{4\pi m E}{3N} \right)^{3/2} \right] + \frac{5}{2}N \quad (1.28)$$

故可以计算  $\Omega$ ：

$$\begin{aligned} \therefore \Omega &= \frac{\partial \Sigma}{\partial E} \Delta E = e^{\ln \Sigma} \frac{\partial \ln \Sigma}{\partial E} \Delta E \\ \therefore \ln \Omega &= \ln \Sigma + \ln \left( \frac{\partial \ln \Sigma}{\partial E} \Delta E \right) \\ \ln \Omega &= N \ln \left[ \frac{V}{Nh^3} \left( \frac{4\pi m E}{3N} \right)^{3/2} \right] + \frac{5}{2}N + \ln \frac{3}{2}N + \ln \frac{\Delta E}{E} \end{aligned}$$

由于我们总是希望  $N \rightarrow \infty$ ，所以后面两项和第一项相比总是可以忽略不计。能量分布在  $(E - \frac{1}{2}\Delta E, E + \frac{1}{2}\Delta E)$  之中的微观配容数就可以写为

$$\ln \Omega = N \ln \left[ \frac{V}{Nh^3} \left( \frac{4\pi m E}{3N} \right)^{3/2} \right] + \frac{5}{2}N \quad (1.29)$$

乍一看我们得到了一个看起来很奇怪的结果，即系统能量分布在  $E$  附近的配容数竟然等于从 0 开始一直数到  $E$  的总和，但这的确是合理的，因为系统分布函数的增加速度是如此迅速，连前面所

有的累积量都赶不上这一个小区间所造成的增长。同时我们也看到，如此庞大的数字，基本上可以将常用的任何量囫圇吞枣地吃掉。

还差一步。现在联用热力学基本公式，我们写出熵的表达式

$$S = k_B \ln \Omega = Nk_B \ln \left[ \frac{V}{Nh^3} \left( \frac{4\pi m E}{3N} \right)^{3/2} \right] + \frac{5}{2} Nk_B$$

与先前的结果进行对比，将  $S$  分别对  $E, V$  求导，再将导数和热力学公式对应，很容易就得到  $PV = Nk_B T$  和  $E = \frac{3}{2} Nk_B T$ ，我们也就爬到了半山腰。

### 反思与质疑

累计能态密度中需要除以的两个因子  $2^{3N}$  以及  $N!$  的原因分别是每个自由度只能取正整数，以及每个粒子不可分辨。并且由于理想气体空间上的稀薄导致的态占据、态分配上的稀薄，相同分子处于同一个态的几率很小，以至于可以忽略不计，所以不需要再添加一个组合因子。

那么为什么考虑  $3N$  个自由度时应当除以  $N!$  而不是  $(3N)!$  呢？那是因为这是这  $N$  个粒子不可区分，而不是这  $3N$  个自由度不可区分，对于每一个粒子来说， $x, y, z$  都是可以区分的。

不过，对于不可分辨粒子设定的合理性，则需要追溯至吉布斯佯谬中来。若认为组成理想气体的各个粒子是可以标号的，首先便会导致推出的熵公式不满足加和性。其次，在考虑同种粒子在同一压强下的混合时，也会推出此时的混合熵正定的矛盾[?]。基于逻辑上的考虑，认为这些粒子具有全同性会使得我们的理论自洽。

## 1.5 总结

微正则系综以其简单的最优化世界观为基础，帮助我们爬上了统计力学的半山腰。它以配容数  $\Omega$  及其对数  $S = k_B \ln \Omega$  为第一性的量，给了热力学一个完整的逻辑解释。对于经典理想气体，为了计算以  $(N, V, E)$  为标记的宏观态所能处于的可能的微观态数目，根据粒子能量所可能取值的特点，我们选择将微观态数目与  $3N$  维状态空间中的球体进行对应。而热力学极限让我们可以合法地使用微分代替寻找方程的整数解，从而得到  $\ln \Omega$  的渐进表达式。对于量子谐振子系统，则可以使用排列组合直接得到  $\Omega$  的显式表达式。

但如果继续向峰巅攀爬将会变得困难，因为微正则系综是一个非常形而上的理论，它对优化参量  $(N, V, E)$  准确度的要求让我们这些有限的生命望而却步——对于大多数物理系统来说，测量  $(N, V, E)$  将是一个非常艰巨的任务，更不要说计算出  $\Omega(N, V, E)$ 。

但微正则系综作为我们的理论起点，是值得肯定的，它是我们登山过程中的一条在山脊上延伸的小径。这里先对微正则系综研究方法进行一些总结，然后在下一节中，我们将换个方向——换到沿山谷的石板路上来。

## 对微正则系综的总结

使用微正则系综处理问题可以遵循这样的顺序：

1. 微正则系综的核心量是配容数  $\Omega$
2. 根据等概率假设，通过对微观态进行简单计数得到  $\Omega(N, V, E)$
3.  $S = k_B \ln \Omega$
4. 联用热力学公式  $dS = \beta dE + \eta dV + \zeta dN$ ，偏导数对应相等





## 第二章 正则系综

前面我们从对热力学的反思中建立了系综理论的基本原理：以  $\Omega$  为最优化目标函数，广度性质  $(N, V, E)$  为决策变量，通过它们之间的依赖关系我们可以得到系统的所有热力学性质。现在，我们要对微正则系综理论进行反思，并找出一个可以弥补它缺点的解决方案。

使用  $(N, V, E)$  作为描述系统的最优化过程是一件非常自然的事，因为它们都是广度性质，而使用一些和系统大小成正比的量作为系统的状态参量无疑是非常直观的。然而，这样的直观反而使得我们的理论变得形而上，以至于不切实际：例如，我们很少测量一个物理系统的总能量，并且要想固定它也是很不容易的。

**问题的根源在于：对微正则系综的三个广度决策变量的测量，不具有实际的可行性。**

所以，我们希望找到一个容易测量的，并且和能量具有相同内涵的量来等价地替代  $E$  作为决策变量的地位——此时我们回想起了理论力学中的勒让德变换。所以我们的目标是：找到一个作用于决策变量  $E$  的勒让德变换，同时目标函数  $\Omega$  变为一个新的目标函数  $A$ ：

$$A = E \frac{\partial \Omega}{\partial E} - \Omega$$

此时我们再次回想起热力学公式

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN$$

显然，**温度  $T$  就是我们要找的变量**，并且固定一个系统的温度也是可以做到的，因为只需要把系统与一个足够大的大热库耦合即可。随后，再对我们想法的一些调整让它能够和热力学建立联系，我们将会看到：变换完全等价地将最优化问题 (??) 转换成另一个最优化问题：

$$\begin{aligned} \min A(N, V, T) \\ \text{s.t. } \sum n_i = \mathcal{N}, \quad T = T_0 \end{aligned} \tag{2.1}$$

其中  $A$  就是亥姆霍兹自由能，满足  $A = E - TS$ ，以之为核心的即是 **正则系综**。

### 2.1 最概然与正则分布

下面我们来将前面的反思落到实处，此时有必要将系综理论的基本原理再复述一遍：

在单一瞬间同时考虑大量系统，他们全部是给定系统的某种“思维复本”——其特性由与原系统一样的宏观态来表征，但极其自然地处在所有各种可能的微观态中 [?]

也就是说，我们考虑的不是“这个系统”，而是无数个“这个系统”所组成的更宏大的集合，但集合的元素具有一些关系，并且可能还共同分配了某个想象中的量，并且这个量的分配方式还决定了每个种类子集所占的比例。

我们考虑的就是这样一件事：把系统看成正则系综的一个成员，然后，由组成系综的  $\mathcal{N}$  个相同系统去分配该系综总能量  $\mathcal{E}$ ，进而研究这种分配过程的统计学——也就是计算在任意时刻，发现系统处于由能量  $E_r$  所表征的状态之一的概率  $P_r$  究竟有多大。

考虑由分享总能量  $\mathcal{E}$  的  $\mathcal{N}$  个全同系统组成的一个系综，令  $E_r$  表示这系统的能量本征值。倘若  $n_r$  表示在任意时刻  $t$  具有能量  $E_r$  的系统的数目，则这些数必须满足下面这些明显的条件：

$$\sum_r n_r = \mathcal{N} \quad (2.2)$$

$$\sum_r n_r E_r = \mathcal{E} = \mathcal{N}U, \quad \text{where } U \equiv \frac{\mathcal{E}}{\mathcal{N}} \quad (2.3)$$

显然，将  $\mathcal{N}$  个系统分成  $\{n_0, n_1, \dots\}$  份，分配方式服从多项分布：

$$W\{n_r\} = \frac{\mathcal{N}!}{n_0!n_1!n_2!\dots} \quad (2.4)$$

取对数，并引入斯特林近似则有

$$\ln W = \mathcal{N} \ln \mathcal{N} - \sum_r n_r \ln n_r \quad (2.5)$$

依照微正则系综的观点，我们要求解的是以下的约束最优化问题：

$$\begin{aligned} & \max \ln W, \\ & \text{s.t. } \sum_i n_i = \mathcal{N}, \quad \sum_r n_r E_r = \mathcal{E} \end{aligned} \quad (2.6)$$

我们将它转化为约束变分问题。既然要求  $\ln W$  在  $\{n_r\}$  这一组条件下取得最大值，那么其变分应当取为 0：

$$\delta(\ln W) = - \sum_r (\ln n_r + 1) \delta n_r = 0$$

并且，这些变分满足条件

$$\left. \begin{aligned} \sum_r \delta n_r &= 0 \\ \sum_r E_r \delta n_r &= 0 \end{aligned} \right\}$$

拉格朗日乘子法使得我们可以将约束变分问题转化为无约束变分问题：

$$\sum_r \{- (\ln n_r + 1) - \alpha - \beta E_r\} \delta n_r = 0 \quad (2.7)$$

所以，当且仅当对每一个  $\delta n_r$  来说系数都取 0，才有整体求和为 0。所以我们得到

$$\begin{aligned} \ln n_r &= -(\alpha + 1) - \beta E_r, \\ n_r &= C \exp(-\beta E_r), \end{aligned} \quad (2.8)$$

归一化就有

$$P_r = \frac{n_r}{\mathcal{N}} = \frac{\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)} \quad (2.9)$$

### 反思与质疑

微正则系综专注于具有同一本征能量的简并微观态，并认为在这些微观态中，没有哪一个是主导，所以引入等概率假设；正则系综则关注具有不同本征能量的各简并集合之间的联系——不同简并集合的大小不同，这些集合之间的比例也有所差异。

所以等概率假设与  $e^{\beta E_r}$  并不矛盾，因为一个描述的是简并集合内部的等概率性，一个描述的是这些简并集合之间的比例差异。并且，当我们说出“比例差异”的时候，也就自然地将问题理解为了一个几何概型：具有不同能量的各微观态仍然是等概率的，只是个数不一样。

## 2.2 正则系综的研究方法

### 2.2.1 统计量的物理含义

现在我们来赋予各个统计量以一定的物理含义，我们从能量开始。根据正则分布，定义系统的能量  $U$  是随机变量  $E$  的系综平均：

$$U = \frac{\sum_r E_r \exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)} = -\frac{\partial}{\partial \beta} \ln \left\{ \sum_r \exp(-\beta E_r) \right\} \quad (2.10)$$

并回忆那个以最小化亥姆霍兹自由能的热力学公式

$$dA = dU - TdS - SdT = -SdT - PdV + \mu dN$$

$$S = -\left(\frac{\partial A}{\partial T}\right)_{N,V}, \quad P = -\left(\frac{\partial A}{\partial V}\right)_{N,T}, \quad \mu = \left(\frac{\partial A}{\partial N}\right)_{V,T},$$

以及得到它的那个勒让德变换

$$U = A + TS = A - T \left(\frac{\partial A}{\partial T}\right)_{N,V} = -T^2 \left[ \frac{\partial}{\partial T} \left( \frac{A}{T} \right) \right]_{N,V} = \left[ \frac{\partial(A/T)}{\partial(1/T)} \right]_{N,V}$$

经过对比，我们可以将统计量与热力学量以这样的方式对应起来：

$$\beta = \frac{1}{kT}, \quad \ln \left\{ \sum_r \exp(-\beta E_r) \right\} = -\frac{A}{kT} \quad (2.11)$$

所以我们的最优化目标函数  $A$  就可以表示为  $(N, V, T)$  的函数了：

$$A(N, V, T) = -kT \ln Q_N(V, T) \quad (2.12)$$

其中  $Q_N(V, T)$  则称为 **正则配分函数**:

$$Q_N(V, T) = \sum_r \exp(-E_r/kT) \quad (2.13)$$

求和遍及所有可能的能量状态。 $Q$  对  $V, T$  的依赖关系是明显的, 而  $Q$  对粒子数  $N$  的依赖关系就体现在求和的项数中。然后联用新的热力学关系  $dA = -SdT - PdV + \mu dN$  就能得到各热力学量的统计诠释, 这里不再赘述。

现在, 有必要对我们所得的结果做一些评述。在推导的过程中, 我们实际上以微正则系综的配容数最大化为出发点, 并且微正则系综中最复杂的部分: 通过一定的数学方法计算配容数  $\Omega$  的过程则被更普适的变分方法 (??) 所替代, 所以我们可以放心大胆地宣称: **正则系综与微正则系综等价**。

我们给出著名的吉布斯熵公式

$$S = -k \langle \ln P_r \rangle = -k \sum_r P_r \ln P_r \quad (2.14)$$

由此我们可以看到, 一个系统的熵竟然与系统的具体性质无关, 只与系统在所能处于状态上的概率分布  $P_r$  有关, 换句话说: **熵是概率密度的函数**。而关于它的正确性, 信息论的证明更加简洁, 读者可自行参考。显然, 它与微正则系综的熵公式也是自洽的。

$$S = -k \sum_{r=1}^{\Omega} \left\{ \frac{1}{\Omega} \ln \left( \frac{1}{\Omega} \right) \right\} = k \ln \Omega$$

### 2.2.2 能量的涨落

正则系综将能量  $E$  视为一个取值为  $(0, \infty)$  的随机变量, 看来有必要研究一下能量的方差, 也就是能量涨落, 这可以通过对  $U$  的表达式求导来实现:

$$\begin{aligned} \frac{\partial U}{\partial \beta} &= - \frac{\sum_r E_r^2 \exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)} + \frac{\left[ \sum_r E_r \exp(-\beta E_r) \right]^2}{\left[ \sum_r \exp(-\beta E_r) \right]^2} \\ &= - \langle E^2 \rangle + \langle E \rangle^2, \end{aligned}$$

故能量的方差  $\text{Var}(E)$  就是

$$\text{Var}(E) \equiv \langle E^2 \rangle - \langle E \rangle^2 = - \left( \frac{\partial U}{\partial \beta} \right) = kT^2 \left( \frac{\partial U}{\partial T} \right) = kT^2 C_V \quad (2.15)$$

概率论能够确定能量相对均值  $U$  涨落的方差:

$$\text{Var} \left( \frac{E}{U} \right) = \frac{1}{U^2} \text{Var}(E) = \frac{kT^2 C_V}{U^2}$$

可见, 随者粒子数  $N$  的增长, 它的量级是  $O(N^{-1})$ , 所以我们期望: 在  $N \rightarrow \infty$  时, 概率密度的展宽将逐渐减小。为了确认这一点, 我们下面来直接研究  $E$  的概率密度函数  $P(E) = \Omega(E)e^{-\beta E}/Q$ , 其中  $\Omega(E)$  的含义是对于确定的能量, 也就是每一个微正则系综的状态数。

对概率取对数有

$$\frac{\partial \ln P(E)}{\partial E} = \frac{\partial \ln \Omega(E)}{\partial E} - \beta$$

在最大值处，有偏导数为 0，所以式中第一项在极值点处取值即为

$$\left( \frac{\partial \ln \Omega(E)}{\partial E} \right)_{E=U} = \beta$$

所以

$$\frac{\partial^2 \ln P(E)}{\partial E^2} = \frac{\partial \beta}{\partial E} = 1 \left/ \frac{\partial E}{\partial \beta} \right. = -\frac{1}{\text{Var} E} = -\frac{1}{\sigma_E^2}$$

那么现在对概率密度的对数在顶点  $E = U$  处做泰勒展开：

$$\ln P(E) = \ln P(U) - \frac{(E - U)^2}{2\sigma_E^2}$$

那么概率密度就可以写为 Gauss 分布的形式

$$P(E) = P(U) \exp \left[ -\frac{(E - U)^2}{2\sigma_E^2} \right] \quad (2.16)$$

显然，这个密度函数在热力学极限下收敛于  $\delta$  函数。这也表明：[正则系综 \(NVT\)](#) 在热力学极限下依概率收敛于微正则系综 (NVE)。

#### 反思与质疑

从前面的讨论中我们发现，正如我们在做勒让德变换时新旧变量具有一定联系那样，我们发现旧变量（广延量  $X$ ）的平均值是配分函数的自然对数对新变量（强度量  $S$ ）的偏导数：

$$\langle X \rangle = -\frac{\partial \ln Q}{\partial S}$$

而广延量  $X$  的方差（涨落）则可以写为

$$\text{Var}(X) = \frac{\partial^2 \ln Q}{\partial S^2} = -\frac{\partial \langle X \rangle}{\partial S}$$

最后一个等式具有的物理含义是非常重要的，因为强度量常常是环境施加给系统的一种外部作用。如果系统的涨落恰好等于系统的广延量对外界的响应，这就说明：[系统对外界响应的强度取决于自身涨落的大小](#)。

人也是如此。老师给你讲了半天，响应强的那就表明你学会了，说明你的涨落很强。但如果是对牛弹琴，牛是很稳定的，那就没有多少响应，它还是自顾自地吃草。

### 2.3 配分函数的性质

自从正则系综被引入后，事情似乎变得简单了起来。现在，我们再次来到了起初使用微正则系综时所到达过的统计力学高峰的半山腰，可以停下休息，并回顾之前的上山路径了。在爬山的过程中，我们证明了微正则系综与正则系综实际上是对同一物理系统的两种等价表述。

然而到目前为止，我们并没有提到系综理论与经典力学对一个系统状态的描述方式：相空间有什么紧密的联系——无论是对理想气体的推导还是正则系综的配分函数的计算过程，我们都只是默认了“求和遍及所有的微观态”这件事，而这些微观态是离散的。那么这样的一个离散集合，与相空间中的点有什么联系呢？

### 2.3.1 微观态与相空间

很自然的，我们期望这样一件事情的发生：相空间中一定区域的体积对应于一个微观态，我们要找到“等效于一个微观态”的基本体积  $\omega_0$ 。而事实也确实如此，只需要考虑这样一件事：系统代表点所占据的相空间体积为

$$\omega = \int' \dots \int' (d^{3N}q d^{3N}p) \quad (2.17)$$

而积分区域是在能量  $E$  的附近，所以

$$\iiint_{(E-\Delta) \leq \sum_{i=1}^{3N} (p_i^2/2m) \leq E} d^{3N}p = \iiint_{(E-\Delta) \leq \sum_{i=1}^{3N} y_i^2 \leq 2mE} d^{3N}y$$

换到球坐标系，容易看出这是  $3N$  维球壳的体积，其结果为

$$\Delta \left( \frac{m}{2E} \right)^{1/2} \left\{ \frac{2\pi^{3N/2}}{[(3N/2) - 1]!} (2mE)^{(3N-1)/2} \right\}$$

由此得出

$$\omega \simeq \frac{\Delta}{E} V^N \frac{(2\pi mE)^{3N/2}}{[(3N/2) - 1]!}$$

所以渐进的有

$$(\omega/\Gamma)_{\text{asyp}} \equiv \omega_0 = h^{3N} \quad (2.18)$$

更一般地说，倘若系统有  $N$  个自由度，则所期望的转换系数为  $\omega_0 = h^N$ 。

对于量子系统，由于不确定性原理，我们再也无法使用经典的相空间来描述系统的状态了——这看起来与我们“使用相空间中的一定区域来代表一个量子态”的想法不谋而合：在相空间任一点  $(q, p)$  的周围，存在着一定的体积，在这个体积内不可能准确地确定代表点的位置。换句话说，这些相体积可以和系统的量子态一一对应。而 Bohr 从这里出发，建立了旧量子论的一个重要假设：角动量子化。

通过简单的计算，对于一维量子系统我们可以同样说明，与一个量子态对应的“相空间体积”是  $\omega_0 = 2\pi\hbar = h$ 。

现在，我们发现自己处于这样一种境地：既然相空间中确定的体积对应于一个微观态，那么对微观态的求和则可以转化为对正则变量  $(q, p)$  的积分，而且这个转换系数是常数！这无疑会为计算正则配分函数提供一种新的视角和随之而来的新方法。作为一个有意义的暗示，如果我们能对系统哈密顿量进行变量分离，那么对相空间的积分也会是变量分离的。

哈密顿函数的变量分离，说明被分离的系统互不干扰，也即是说：**大系统可以分解为若干无关联系统的加和。**

### 反思与质疑

事实上，Bohr 氢原子模型中的角动量量子化实际上是作用变量量子化，它发轫自经典力学的作用变量-角变量理论——周期运动的系统总是可以通过一定的正则变换，从而使得新哈密顿函数依赖于作用变量  $J$  与角变量  $\Theta$ ： $H = H(J, \Theta)$ ，其中  $J$  是守恒量，它的定义为

$$J = \frac{1}{2\pi} \oint p dq$$

环路积分的含义是针对一个周期。此时可以证明： $\Theta$  则总是随着时间线性变化，并且在一个周期内的变化量为  $2\pi$ ：

$$\Theta = \Theta_0 + 2\pi \times \frac{t}{\tau}$$

而 Bohr 提出，周期运动的电子只能取一些分立的值作为作用变量。所以，这里“角动量的量子化”实际上是作用变量的量子化，只是为了让学界容易接受而使用角动量做了等价表述。

### 2.3.2 连续形式的配分函数

**为了让这种求和-积分的转变更加丝滑，先来考虑配分函数的连续形式。**

回想在研究能量涨落时，我们不假思索地将概率密度写为了  $P(E) = \Omega(E)e^{-\beta E}/Q$ ， $\Omega(E)$  像是能量的简并因子。我们就来完整地对它进行讨论。

多数情况下，一个物理系统的能级是简并的，记能级  $E_i$  的简并度为  $g_i$ ，则配分函数可写为

$$Q_N(V, T) = \sum_i g_i \exp(-\beta E_i) \quad (2.19)$$

由于每个简并能级中的微观态都是等概率的，所以这只是将数钱的方式从“按顺序数”换到了“按面值数”而已。系统处于具有能量  $E_i$  的任何状态的概率  $P_i$  的相应表达式就是

$$P_i = \frac{g_i \exp(-\beta E_i)}{\sum_i g_i \exp(-\beta E_i)} \quad (2.20)$$

系统的能量在热力学极限下趋于连续，所以将概率质量函数换为概率密度函数，再做一番归一化，我们就得到

$$P(E)dE = \frac{\exp(-\beta E)g(E)dE}{\int_0^\infty \exp(-\beta E)g(E)dE} \quad (2.21)$$

在这种情形下，配分函数也从求和转变为了积分

$$Q_N(V, T) = \int_0^\infty e^{-\beta E} g(E) dE \quad (2.22)$$



由于  $\beta > 0$ ，所以我们总是可以认为配分函数  $Q$  是能态密度  $g$  的 **拉普拉斯变换**，其逆变换公式为

$$g(E) = \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} e^{\beta E} Q(\beta) d\beta \quad (\beta' > 0) \quad (2.23)$$

积分路径沿着平行于虚轴的右边进行，即沿着直线  $\text{Re}\beta = \beta' > 0$ 。当然，只要积分收敛，路径可以连续变形。

也许现在看来，这样做的价值仅是通过拉普拉斯变换的唯一性说明了态密度和配分函数的等价性。但的确有些场合，对于给定系统配分函数的计算，以及随后进行其态密度的计算，确实变得十分简单，而直接从第一性原理出发来进行态密度的计算却是相当复杂的，例如相对论气体。

为了这一方法的使用更加顺畅（不至于积不出来），这里有一个相关的积分公式：

$$\frac{1}{2\pi i} \int_{s' - i\infty}^{s' + i\infty} \frac{e^{sx}}{s^{n+1}} ds = \begin{cases} \frac{x^n}{n!} & \text{for } x \geq 0 \\ 0 & \text{for } x \leq 0 \end{cases} \quad (2.24)$$

### 2.3.3 无关联系统

**现在，我们将从配分函数的性质出发，将统计力学推上巅峰。**

前面讲到，我们希望将相空间的表述形式也囊括进系综理论的范围之内——因为系综是相空间中的一定区域。而物理量  $f(q, p, t)$  在这一区域上的平均值则是

$$\langle f \rangle = \frac{\int f(q, p) \rho(q, p) d^{3N}q d^{3N}p}{\int \rho(q, p) d^{3N}q d^{3N}p} \quad (2.25)$$

其中函数  $\rho(q, p)$  则是相空间一点附近的代表点密度，是在相点  $(q, p)$  附近找到一个代表点的概率的一种量度。而我们早已说明：在正则系综里，

$$\rho(q, p) \propto \exp\{-\beta H(q, p)\} \quad (2.26)$$

所以  $\langle f \rangle$  可以取为

$$\langle f \rangle = \frac{\int f(q, p) \exp(-\beta H) d\omega}{\int \exp(-\beta H) d\omega} \quad (2.27)$$

该表达式的分母与配分函数直接相关。根据我们已经得到的微观态与相体积的转换关系：在相空间中的体积元  $d\omega$  对应于系统的

$$\frac{d\omega}{h^{3N}}$$

这么多的量子态，因而配分函数的精确表达式就是

$$Q_N(V, T) = \frac{1}{h^{3N}} \int e^{-\beta H(q, p)} d\omega \quad \text{or} \quad Q_N(V, T) = \frac{1}{N! h^{3N}} \int e^{-\beta H(q, p)} d\omega \quad (2.28)$$

有时视情况需要再除以一个不可分辨因子  $N!$ ，但这并不会对我们的思路造成什么影响。



这是变量代换的一小步，但却是统计力学的一大步。它的一个显著的优点前面已经提到过了，即如果系统可以分解为若干无关联子系统，则哈密顿函数变量分离，配分函数的积分计算也变量分离。而积分计算的分离导致系统总配分函数是各个子配分函数的乘积：

$$Q_N(V, T) = [Q_1(V, T)]^N \quad \text{or} \quad Q_N(V, T) = \frac{1}{N!} [Q_1(V, T)]^N \quad (2.29)$$

这也叫做 **配分函数的因子化**。显然，即使分子具有内部自由度，这个情况也不会发生任何改变。

#### 反思与质疑

首先，这和概率论是一致的。因为配分函数的因子化其实相当于随机变量的独立性： $p(X = x, Y = y) = p(X = x)p(Y = y)$ ，所以条件期望满足  $E(X) = E[YE(X|Y = y)] = E(X)E(Y)$ 。

其次，这和热力学是一致的。由于热力学量均是对配分函数的对数求偏导，而导数算符的线性性保证了无关联系统热力学量的加和性。

这时有必要对本节的历程做一个简单的评述。我们从一个简单的想法（系综求和转换为相空间的积分）出发，推导出了宏观配分函数的等价形式(??)式，它还可以被无关联性分解为(??)的形式，即单个粒子的微观性质。所以，我们可以自豪的宣布：**配分函数，连同以之为巅峰的统计力学，是宏观与微观相互联系的桥梁。**

## 2.4 实例

前面的讨论已经使我们站上了统计力学的巅峰。现在可以下山了。在下山的过程中，有必要给物理学一点小小的统计震撼。

### 2.4.1 再论理想气体状态方程

作为一个简单例子，我们先来讨论理想气体状态方程： **$PV = NkT$  是经典不可区分子系统无关联性的直接结果。**

首先，这样一个系统的配分函数可以写为

$$Q = \frac{1}{N!} Q_1^N$$

引入斯特林近似，易得亥姆霍兹自由能可以写作如下形式

$$A = -kT \ln Q = -kTN \ln Q_1 + kT(N \ln N - N)$$

联用热力学公式， $P$  的表达式为

$$P = -\frac{\partial A}{\partial V} = NkT \frac{\partial \ln Q_1}{\partial V}$$

然而, 回顾 (??) 一节, 在那里我们通过独立性给出了无相互作用粒子的配容数对体积的依赖关系。其实配分函数对体积的依赖关系也是一样的, 只要单粒子哈密顿量只依赖于正则动量  $\mathbf{p}$  即可:

$$Q_1 = \frac{1}{1! \times h^3} \int dq_1 dq_2 dq_3 \int e^{H(p_1, p_2, p_3)} dp_1 dp_2 dp_3 = \frac{V}{h^3} \int e^{H(p_1, p_2, p_3)} dp_1 dp_2 dp_3$$

所以  $Q_1 \propto V$ , 则

$$P = NkT \frac{\partial \ln Q_1}{\partial V} = NkT \times \frac{1}{V}$$

#### 反思与质疑

我们反复地以理想气体为例, 通过不同的方法得到了这一状态方程。为了彻底地理解这一等式, 不妨放弃压强、体积、温度等量的具体或是力学的、或是感性的实际意义, 而上升为最优化问题的决策变量和偏导数。这时, 无关联性要求

$$-\left(\frac{\partial A}{\partial V}\right)_{N,T} V = Nk \left/ \left(\frac{\partial S}{\partial E}\right)_{N,V} \right.$$

### 2.4.2 有限能级系统与负温度

不失一般性, 考虑  $N$  个粒子组成无关联的系统, 每个粒子有两个能级, 能量分别为 0 (基态) 和  $\epsilon$  (激发态)。这样的系统可以用占据数  $\{n_i\}$  进行描述。 $n_i$  描述第  $i$  个粒子占据能级的情况, 若粒子占据  $\epsilon = 0$  的能级, 则  $n_i = 0$ , 否则  $n_i = 1$ 。容易得到系统的配分函数与能量如下:

$$Q = (1 + e^{-\epsilon/kT})^N, \quad U = \frac{N\epsilon}{e^{\epsilon/kT} + 1} \quad (2.30)$$

由此可见, 要想将所有的粒子尽可能的激发到更高的能级上去, 也就是说让  $U$  尽可能靠近  $N\epsilon$ , 必须要求

$$\frac{1}{T} \rightarrow -\infty, \quad T \rightarrow -0$$

对于一个正常系统, 这相当于温度从正方向通过绝对零度。事实上, 实际系统的能级不可能只有两个——最简单的谐振子也具有无限的能级可供填充。这里出现负温度, 完全是能级受限, 以至于态占据的受限所致, 而施加这样的限制, 无疑会消耗环境的功。

此外, 我们注意到系统的热容  $C_V$  为

$$C_V = N\epsilon^2 \frac{e^{\epsilon/kT}}{kT^2(1 + e^{\epsilon/kT})^2} \quad (2.31)$$

可见, 热容在温度趋于零和趋于无穷时均为零。温度趋于零时, 热容以指数衰减趋于零, 基态和最低激发态之间具有能量差的系统都有这个特征; 高温时热容趋于零则是饱和效应的体现, 能量随微观状态数变化而出现极值的系统常常有这样的特征。

此外, 使用微正则系综分析不难发现, 只要系统的能级是有限的, 那么配容数  $\Omega$  服从多项分布, 其必有极大值。也就是说负温度对于有限能级系统来说是不可避免的。这迫使我们相信, 实际系统都至少存在可数个能级。

## 反思与质疑

至于现实中为什么观察不到负温度，我们做如下解释：设将系统分割为大量很小的部分，用  $M_a, E_a, P_a$  标记这些部分的质量、能量与体积。由于熵是内能的单调函数，而内能等于总能量  $E_a$  减去整体运动的能量

$$S = \sum_a S_a \left( E_a - \frac{P_a^2}{2M_a} \right) \quad (2.32)$$

考虑一个整体静止的闭合系统，现在假设温度可以是负的，那么熵就会随着括号内量的减小而增大。但由于熵是自发增大的，所以为了保持总动量不变，物体将会自发瓦解，相互飞散，以使得每一个括号内的数值都更小。所以，在  $T < 0$  的情形下，根本不可能有平衡物体存在。

## 2.4.3 谐振子系统

先来看经典谐振子。哈密顿量如下：

$$H(q_i, p_i) = \frac{1}{2}m\omega^2 q_i^2 + \frac{1}{2m}p_i^2 \quad (i = 1, \dots, N) \quad (2.33)$$

由于粒子是无关联的，所以单粒子配分函数

$$\begin{aligned} Q_1(\beta) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left\{ -\beta \left( \frac{1}{2}m\omega^2 q^2 + \frac{1}{2m}p^2 \right) \right\} \frac{dqdp}{h} \\ &= \frac{1}{h} \left( \frac{2\pi}{\beta m \omega^2} \right)^{1/2} \left( \frac{2\pi m}{\beta} \right)^{1/2} = \frac{1}{\beta \hbar \omega} = \frac{kT}{\hbar \omega} \end{aligned} \quad (2.34)$$

不妨设每个粒子可区分，所以系统总配分函数以及亥姆霍兹自由能为

$$Q_N(\beta) = [Q_1(\beta)]^N = (\beta \hbar \omega)^{-N} = \left( \frac{kT}{\hbar \omega} \right)^N, \quad A \equiv -kT \ln Q_N = NkT \ln \left( \frac{\hbar \omega}{kT} \right) \quad (2.35)$$

易得熵的表达式

$$S = Nk \left[ \ln \left( \frac{kT}{\hbar \omega} \right) + 1 \right] \quad (2.36)$$

能态密度  $g(E)$  是配分函数的逆拉普拉斯变换

$$g(E) = \frac{1}{(\hbar \omega)^N} \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} \frac{e^{\beta E}}{\beta^N} d\beta \quad (\beta' > 0) \quad (2.37)$$

根据公式 (??)，积分结果为

$$g(E) = \begin{cases} \frac{1}{(\hbar \omega)^N} \frac{E^{N-1}}{(N-1)!} & \text{for } E \geq 0 \\ 0 & \text{for } E \leq 0 \end{cases} \quad (2.38)$$

通过  $g(E)$  我们同样可以计算系统的熵

$$S(N, E) = k \ln g(E) \approx Nk \left[ \ln \left( \frac{E}{N \hbar \omega} \right) + 1 \right] \quad (2.39)$$

不难发现和 (??) 式是等价的, 只要将平衡能量代换为平衡温度。这同时也表明, 微正则系综与正则系综在热力学极限下是等价的。

再来看量子谐振子。所有可能能级的能量为

$$\varepsilon_n = \left(n + \frac{1}{2}\right) \hbar\omega; \quad n = 0, 1, 2, \dots \quad (2.40)$$

配分函数

$$Q_1(\beta) = \sum_{n=0}^{\infty} e^{-\beta(n+1/2)\hbar\omega} = \frac{\exp(-\frac{1}{2}\beta\hbar\omega)}{1 - \exp(-\beta\hbar\omega)}, \quad Q_N(\beta) = [Q_1(\beta)]^N = e^{-(N/2)\beta\hbar\omega} \{1 - e^{-\beta\hbar\omega}\}^{-N} \quad (2.41)$$

所以亥姆霍兹自由能为

$$A = NkT \ln \left[ 2 \sinh \left( \frac{1}{2} \beta \hbar \omega \right) \right] = N \left[ \frac{1}{2} \hbar \omega + kT \ln \{1 - e^{-\beta \hbar \omega}\} \right] \quad (2.42)$$

则系统能量

$$U = \frac{1}{2} N \hbar \omega \coth \left( \frac{1}{2} \beta \hbar \omega \right) = N \left[ \frac{1}{2} \hbar \omega + \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \right] \quad (2.43)$$

这表明量子谐振子不服从能量均分定理, 每个自由度的对能量的贡献总是大于  $\frac{1}{2}kT$  的, 只有在高温极限下这个现象才会逐渐消失。

#### 2.4.4 极端相对论气体

对于极端相对论气体, 动量与能量的关系为  $\epsilon = |\mathbf{p}|c$ , 其中  $c$  是光速。利用配分函数的连续形式

$$Q_1(V, T) = \int g(\epsilon) e^{-\beta \epsilon} d\epsilon$$

同时  $g(\epsilon)d\epsilon$  又可写为

$$\begin{aligned} g(p)dp &= \frac{V}{h^3} 4\pi p^2 dp = \frac{4\pi V}{h^3} \frac{\epsilon^2}{c^2} \frac{d\epsilon}{c} = g(\epsilon)d\epsilon \\ \therefore g(\epsilon) &= \frac{4\pi V}{(hc)^3} \epsilon^2 \end{aligned} \quad (2.44)$$

所以配分函数

$$\therefore Q_1(V, T) = \int_0^{\infty} g(\epsilon) d\epsilon = \frac{4\pi V}{(hc)^3} \int_0^{\infty} \epsilon^2 e^{-\beta \epsilon} d\epsilon = 8\pi V \left( \frac{kT}{hc} \right)^3 \quad (2.45)$$

总配分函数

$$Q_N(V, T) = \frac{1}{N!} \left\{ 8\pi V \left( \frac{kT}{hc} \right)^3 \right\}^N \quad (2.46)$$

能态密度

$$\begin{aligned} g(E) &= \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} e^{\beta E} Q(\beta) d\beta \\ &= \frac{(8\pi V)^N}{N! (hc)^{3N}} \text{Res} \left[ \frac{e^{\beta E}}{\beta^{3N}} \right]_{\beta=0} \\ &= \frac{(8\pi V)^N E^{3N-1}}{N! (3N-1)! (hc)^{3N}} \end{aligned} \quad (2.47)$$

微正则系综也可以解决这个问题。对于极端相对论气体，它所有可能的能量状态为

$$\varepsilon(n_x, n_y, n_z) = \frac{hc}{2L} (n_x^2 + n_y^2 + n_z^2)^{1/2} \quad (2.48)$$

不过，计算  $\Omega$  就有些麻烦了。

## 2.5 总结

正则系综作为微正则系综的等价方法，极大简化了对物理系统的研究与计算。它以简并度  $g(E)$  整合了具有相同能量的各个简并微观态，并通过拉普拉斯变换后的配分函数  $Q_N(V, T)$  为核心，与亥姆霍兹自由能建立联系。不仅如此，我们看到这样的配分函数可以因子化，只要系统的哈密顿量变量分离。

总之，正则系综开辟了通向统计力学顶峰的另一条道路。

### 对正则系综的总结

使用正则系综处理问题可以遵循这样的顺序：

1. 正则系综的核心量是配分函数  $Q$
2. 直接求和，或者转换为相空间的积分，得到  $Q_N(V, T)$
3.  $S = -k_B \sum P_i \ln P_i$
4. 联用热力学公式  $dA = -SdT - PdV + \mu dN$ ，偏导数对应相等



## 第三章 巨正则系综

在上一章中，我们基于微正则系综  $(N, V, E)$  的三个广度决策变量的测量没有实际可行性的考虑，首先对能量  $E$  通过勒让德变换将其等价转换为了熵对能量的导数：也就是强度量  $T$ ，根据我们的实际经验，将它命名为“温度”。基于此，我们将使用  $(N, V, T)$  作为决策变量的这一方法称为正则系综。这种理论处理方案的有效性，从已经讨论过的许多例子就可以看得很清楚。然而，对于一些物理和化学问题，正则系综表述形式的适用性毕竟还是有所限制的——不仅系统的能量难以直接测量得到，粒子数目  $N$  亦是如此。

现在，我们在对能量  $E$  勒让德变换后，继续对另一个广度性质  $N$  做勒让德变换。再次回想起以熵  $S$  为核心的热力学公式

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN$$

所以，和在正则系综中的思路一样，与  $N$  等价的强度量为  $\mu$ ，这也就是本章的主题：[巨正则系综](#)。

由于确定热力学系统的一个平衡态至少需要一个广度性质，所以最多只能使用两次勒让德变换。如果不选择  $N$  而是对  $V$  做变换，就会得到 [等温等压系综](#)。

### 3.1 巨正则分布与配分函数

根据前面的想法，巨正则分布的推导过程应当与正则系综类似。这里我们不加证明地写出巨正则分布与相应的配分函数。

不过，“想象由  $\mathcal{N}$  个全同系统组成的系综，令  $n_{r,s}$  表示在任意时刻  $t$  具有粒子数  $N_r$  和能量值  $E_s$  的系统的数目”这样的表述和前面没什么区别，不妨换一个角度：从巨正则系综中取出一个成员，则这个成员具有粒子数  $N_r$  和能量值  $E_s$  的概率为

$$P(N = N_r, E = E_s) = \frac{\exp(-\alpha N_r - \beta E_s)}{\sum_r \sum_s \exp(-\alpha N_r - \beta E_s)} \quad (3.1)$$

我们注意到巨正则分布 (??) 是二维离散分布, 那么各基本物理量的系综平均值如下:

$$\bar{N} = \frac{\sum_{r,s} N_r \exp(-\alpha N_r - \beta E_s)}{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)} \equiv -\frac{\partial}{\partial \alpha} \left\{ \ln \sum_{r,s} \exp(-\alpha N_r - \beta E_s) \right\} \quad (3.2)$$

$$\bar{E} = \frac{\sum_{r,s} E_s \exp(-\alpha N_r - \beta E_s)}{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)} \equiv -\frac{\partial}{\partial \beta} \left\{ \ln \sum_{r,s} \exp(-\alpha N_r - \beta E_s) \right\} \quad (3.3)$$

既然是二维分布, 那么就可以引入条件分布和边缘分布了。容易看出  $N$  与  $E$  不是独立的:

$$P(N = N_i | E = E_j) = \frac{\exp(-\alpha N_i - \beta E_j)}{\sum_r \exp(-\alpha N_r - \beta E_j)}$$

$$P(N = N_i) = \frac{\sum_s \exp(-\alpha N_i - \beta E_s)}{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)}$$

通过巨正则系综, 一些更加精细的量也被允许计算平均值, 例如某一能级  $\epsilon$  的平均粒子数  $\langle n_\epsilon \rangle$ , 这也是下一章要讨论的内容。例如某一个能级  $E_i$  上的平均占据数  $\langle n_i \rangle$  应当可以写为

$$\langle n_i \rangle = \frac{\sum_r N_r \exp(-\alpha N_r - \beta E_i)}{\sum_{r,s} \exp(-\alpha N_r - \beta E_s)} \equiv -\frac{1}{\beta} \frac{\partial}{\partial E_i} \left\{ \ln \sum_{r,s} \exp(-\alpha N_r - \beta E_s) \right\} \quad (3.4)$$

这将会是玻色-爱因斯坦系统以及费米-狄拉克系统中的重要参量。

## 3.2 巨正则系综的研究方法

### 3.2.1 统计量的物理含义

### 3.2.2 粒子数与能量的涨落

## 3.3 实例

## 3.4 总结

### 反思与质疑

似乎我们的理论在巨正则系综这里走到了尽头, 但是不要忘记: 最开始的三个广度性质并不一定要选择  $(N, V, E)$ , 因为广度性质的背后, 是系统的独立可加运动积分, 而这样的可加运动积分一共有七个: 能量、动量的三个分量以及角动量的三个分量 (参见 (??) 节)。所以在原则上, 我们完全可以选择可加运动积分的线性组合作为描述系统的状态参



量，也就是最优化问题的决策变量。

力学系统运动积分是由哈密顿函数确定的，但哈密顿力学本身并不局限于物理。它对称共轭的形式使得人们可以将一些本来和力学相距甚远的问题（比如捕食者-猎物系统中种群数量随时间的变化）通过合适的变换统一在哈密顿体系下。

而系综理论，作为哈密顿力学的统计延伸，同样不受物理的束缚。也就是说，如果今天有了一个现实系统，而经验告诉我们：为了确定系统的状态需要四个状态参量，其中要有一个广度性质。那么这时，巨正则系综是不是还可以更进一步，再做一组勒让德变换呢？



## 第四章 简单气体理论

前面三章我们分别对三种系综分别进行了一定的讨论。它们在“包含系统全部的运动信息”的意义上都是等价的，区别仅在将不易测量的广度决策变量通过勒让德变换转化为了容易测量的强度决策变量，而这一变换过程对于热力学系统而言，停留在了巨正则系综上。这标志着我们研究的基本原理与基本框架已经成型，现在我们来将它运用于一个最简单的系统，正如它的名称所标明的那样：简单气体。

首先，研究这样的理想模型对于理解自然界中的实际气体是有帮助的：对简单气体模型进行一些必要的修正（例如范德瓦尔斯经验公式），我们可以让它更接近实际气体的行为。其次，我们将对“气体”概念做必要的推广，它的含义并不局限于“无形状有体积的可压缩和膨胀的流体”，用来描述可以用感官感受到的分子气体（箱子里的氦气、水蒸气以及相平衡）。这样的“气体”可以看不见摸不着（光子气体），甚至能够不是实体，仅存在于想象中（声子气体）。

显然，被称为“气体”的系统应当满足“气体”概念的核心性质——**稀薄、无相互作用**。

不过，在正式讨论气体理论之前，我们首先需要解决一个遗留问题：**微观粒子的全同性与相容性**。这将会引出三种统计方法：麦克斯韦-玻尔兹曼统计（称为经典统计）、玻色-爱因斯坦统计以及费米-狄拉克统计（称为量子统计）。

### 4.1 量子统计

在(??)一节我们特别说明了全同性如何影响配分函数的形式（究竟需不需要除以不可分辨因子 $N!$ ），现在我们继续这个话题，即研究全同性对微观态计数的影响。

### 4.2 量子力学系综中的理想气体

#### 4.2.1 配分函数

由于两种量子统计的引入，其对能级占有数的限制将会为正则配分函数 $Q_N$ 的计算带来麻烦，**所以除了经典系统适合使用正则系综描述以外，玻色系统与费米系统更推荐使用巨正则系综来研究。**

下面的推导将会明确地说明正则配分函数遇到的困难以及怎样转向巨正则配分函数  $\Xi$ 。

设系统包含  $N$  个全同粒子（也就是说每个粒子所能够触及的能级都相同，为  $\{\varepsilon\} = \{\varepsilon_0, \varepsilon_1 \dots\}$ ）设系统总能量值  $E$  可以由单粒子的能量  $\varepsilon$  来表示，例如

$$E = \sum_{\varepsilon} n_{\varepsilon} \varepsilon$$

这里对不同能级求和不会导致歧义，所以省去自然数下标。这些  $n_{\varepsilon}$  满足

$$\sum_{\varepsilon} n_{\varepsilon} = N$$

所以正则配分函数可以写为

$$Q_N(V, T) = \sum_{\{n_{\varepsilon}\}} g\{n_{\varepsilon}\} e^{-\beta \sum_{\varepsilon} n_{\varepsilon} \varepsilon} \quad (4.1)$$

其中  $g\{n_{\varepsilon}\}$  是系统一个具有  $E$  能量的微观态的个数。

#### 反思与质疑

由于全同性，前面两个约束条件只能将  $\{n_{\varepsilon}\}$  确定到每个分量的大小  $\{n_1, n_2 \dots\}$ ，但系统仍然可以让不同的粒子去占据，也就是说这里的“组态”一词  $\{n_{\varepsilon}\}$  实际上给一些微观态打了包而已。在 (??) 节中我们使用自然数为这些集合标号，而这里需要使用一组自然数为它们标号。

不过，这里的  $g\{n_{\varepsilon}\}$  有点复杂：

$$g_{\text{B.E.}}\{n_{\varepsilon}\} = 1, \quad (4.2)$$

$$g_{\text{F.D.}}\{n_{\varepsilon}\} = \begin{cases} 1 & \text{if all } n_{\varepsilon} = 0 \text{ or } 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

将它们带入正则配分函数 (??) 式，发现对求和号的限制条件太多了，它的显式计算相当麻烦；而另一方面，巨正则配分函数却变得容易处理，我们有

$$\begin{aligned} \Xi(z, V, T) &= \sum_{N=0}^{\infty} \left[ z^N \sum_{\{n_{\varepsilon}\}} e^{-\beta \sum_{\varepsilon} n_{\varepsilon} \varepsilon} \right] \\ &= \sum_{N=0}^{\infty} \left[ \prod_{\varepsilon} z^{n_{\varepsilon}} \sum_{\{n_{\varepsilon}\}} \prod_{\varepsilon} e^{-\beta \varepsilon n_{\varepsilon}} \right] \\ &= \sum_{N=0}^{\infty} \sum_{\{n_{\varepsilon}\}} \prod_{\varepsilon} (z e^{-\beta \varepsilon})^{n_{\varepsilon}}. \end{aligned} \quad (4.4)$$

这样一来，双重求和  $\sum_{N=0}^{\infty} \sum_{\{n_{\varepsilon}\}}$  实际上等效于彼此无关地按照  $n_{\varepsilon}$  的所有可能值求和。这也可以理解

为将求和约束完全释放：

$$\begin{aligned}\Xi(z, V, T) &= \sum_{n_0, n_1, \dots} \left[ (ze^{-\beta\epsilon_0})^{n_0} (ze^{-\beta\epsilon_1})^{n_1} \dots \right] \\ &= \left[ \sum_{n_0} (ze^{-\beta\epsilon_0})^{n_0} \right] \left[ \sum_{n_1} (ze^{-\beta\epsilon_1})^{n_1} \right] \dots\end{aligned}\quad (4.5)$$

#### 反思与质疑

这样的处理实际上也是概率的分离，它将总体事件的概率还原为了每个基本事件的概率。那么全同性所导致的繁琐计数也就被轻而易举地消除了——这与在第??章中的因子化方法在概率论的意义上是一致的。

现在，非常显然：

$$\Xi(z, V, T) = \begin{cases} \prod_{\epsilon} \frac{1}{(1 - ze^{-\beta\epsilon})} & \text{in the Bose-Einstein case, with } ze^{-\beta\epsilon} < 1, \\ \prod_{\epsilon} (1 + ze^{-\beta\epsilon}) & \text{in the Fermi-Dirac case.} \end{cases}\quad (4.6)$$

如果将经典情况也包含进来，这三种统计可以用一个式子表示：

$$\ln \Xi \equiv \frac{PV}{kT} = \frac{1}{a} \sum_{\epsilon} \ln (1 + aze^{-\beta\epsilon})\quad (4.7)$$

其中  $a = -1, 1, 0$  分别对应玻色系统、费米系统以及经典系统。对于经典系统，在  $a \rightarrow 0$  情形下很容易得到

$$\ln \Xi = z \sum_{\epsilon} e^{-\beta\epsilon} = zQ_1$$

这的确和经典情形是一致的。而系统的粒子数均值和能量均值则为

$$\begin{aligned}\bar{N} &\equiv z \left( \frac{\partial \ln \Xi}{\partial z} \right)_{V, T} = \sum_{\epsilon} \frac{1}{z^{-1}e^{\beta\epsilon} + a} \\ \bar{E} &\equiv - \left( \frac{\partial \ln \Xi}{\partial \beta} \right)_{z, V} = \sum_{\epsilon} \frac{\epsilon}{z^{-1}e^{\beta\epsilon} + a}\end{aligned}\quad (4.8)$$

根据期望的线性性，很容易从中提取出某一能级的平均占有数  $\langle n_{\epsilon} \rangle$  应当采取下面的形式：

$$\langle n_{\epsilon} \rangle = \frac{1}{z^{-1}e^{\beta\epsilon} + a}\quad (4.9)$$

而基于对巨正则配分函数的观察，也会导致相同的结果。

#### 反思与质疑

将式(??)中的双重求和合并一下，并从求积中提出  $z$ ：

$$\ln \Xi = \sum z^N \prod_{\epsilon} e^{-\beta\epsilon n_{\epsilon}}$$

可以看到它求和的每一项实际上仍然是  $e^{(\dots)}$ ，和最原始的巨正则系综配分函数是一致的。

那么要想求某个量的系综平均，只需要对括号内和这个量相乘的其他量求导即可，即

$$\langle n_\varepsilon \rangle = \frac{\partial \ln \Xi}{\partial (-\beta \varepsilon)} = \frac{1}{z^{-1} e^{\beta \varepsilon} + a}$$

#### 4.2.2 经典极限

#### 4.2.3 占有数的统计

### 4.3 总结

到这里为止，我们可以将所有的知识总结为统计力学的一个主题：**涨落**。而在下一章中，我们将转向统计力学的另一个主题：**相变**。

#### 平衡系统三部曲

能谱：

$$E_\nu = E_\nu \{n_1, n_2, \dots, n_N\}$$

配分函数：

$$Q = \sum_\nu \exp(-\beta E_\nu)$$

$$\Xi = \sum_N \sum_\nu \exp(-\alpha N - \beta E_\nu)$$

微观到宏观：

$$\begin{aligned} S &= k \ln \Omega, & \frac{PV}{kT} &= \ln \Xi \\ A &= -kT \ln Q, & S &= -k \sum p_\nu \ln p_\nu \end{aligned}$$

## 第五章 相变的统计力学

在前面的章节中，我们主要处理的是无相互作用粒子组成的系统。对于这类情况，只需知道系统各个组分的能级，就可以直接获悉整个系统的热力学函数。比热、平衡常数、能谱分布等都是属于这一类型的主要现象。我们发现这一类这些系统的热力学函数都有一个显著的特征——都是光滑连续的（除了玻色-爱因斯坦凝聚）。

但世界上还有另一类系统等待着统计力学，那就是 **相互作用系统**。对于这类情况，当我们研究系统的热力学函数时，会遇到解析上的不连续性和奇异性，并且粒子之间的相互作用再也无法通过坐标变换移除，以至于整个系统的能级与各单一组分的能级之间也就不再有简单的加和关系。在适当条件下，系统大量微观组分之间的相互作用还会变得非常强，从而各组分间呈现出 **合作** 行为。当系统的温度达到一个临界  $T_c$  时，这样的合作行为还会具有宏观效应。

### 反思与质疑

光子气体的玻色-爱因斯坦凝聚的原因并不是相互作用，而是微观粒子的关联性。玻色子的交换对称性使得它们倾向于占据较为靠近的态，而费米子的交换反对称性使得它们倾向于占据较为远离的态（这里比较的对象是既不交换对称也不交换反对称的一般波函数）。这种关联性的结果可以等效为一个吸引势或者排斥势，但并不属于这里讨论的相互作用范畴。

另一个问题是：热力学函数是配分函数的偏导数，而配分函数求和的各项很明显是无穷阶可微的，怎么就奇异了呢？如果系统可及的微观态数量有限，那么导数算符的线性和可微函数的加和性确实可以保证热力学函数处处连续可微，但问题在于有的时候求和的项是可数无限的，此时就会出现函数项级数的求导-求和什么时候可交换的问题，也就是配分函数的一致收敛问题。

此外，这样的合作现象还具有一些本质的、不依赖于微观细节的特性。例如，系统的响应函数  $\chi$  在临界点的奇异性，以及系统的各种热力学函数在临界点附近的非解析性。这些特性是相互作用系统的普遍特征，而不是某一特定系统的特征。这类系统的研究是统计力学的一个重要分支，称为 **临界现象** 的统计力学。

**相变** 集中体现着这些特点。为了简化因相互作用而导致的复杂问题，我们需要引入一些模型，

这样的模型需要对相互作用做一定的简化，但仍然保留关于组分间合作行为的基本特性。至于模型能够将相互作用简化至何种程度，就取决于自然界所能呈现的最简单、最容易理解的相互作用是什么。

**磁学系统**是一个比理想气体添加中心势还要简单的模型——自旋只有一上一下两个状态，相邻自旋方向相同和相反将给出两个不同的相互作用能。如果将这些自旋磁子在一维、二维、三维空间中摆放整齐，并将各自旋之间除了最近邻相互作用能之外的项统统忽略，我们就将得到那个著名、普适、以及优雅的 **Ising 模型**。

### 5.1 磁学系统的热力学

对于以往的理想气体、声子、光子系统，它们的热力学我们是见怪不怪的，也容易理解它们确实具有一定的“热”力学（和环境可以进行热交换，或者我们可以直观地看到、感受到它们的热量）。但对于磁学系统，我们从前从未听说过有什么热力学性质（压强？熵？焓？），只知道有一个 **居里温度** 是顺磁性与铁磁性转变的临界温度。

基于我们对磁学系统了解不深的现状，在此有必要建立 **磁学系统的热力学**，也就是要为 (??) 一节中提出的各种偏导数赋予独属于磁学系统的物理意义。

为了从宏观的角度描述一个磁学系统，我们发现需要三个参量：温度  $T$ 、磁场  $H$ 、磁矩  $M$ 。可以看到温度 and 外加磁场均为强度性质，而磁矩则是广度性质。这样一来，对于磁学系统而言，巨正则系综应当作为微观与宏观的桥梁。

首先我们应当从能量守恒定律推导出磁学系统的热力学第一定律，这里不加证明地给出结论：

$$dU = TdS + HdM \quad (5.1)$$

可以看到磁学系统对外做功  $dW = -HdM$ ，代替了理想气体的  $dW = PdV$ 。从这里开始，对第一定律做各种勒让德变换，就可以分别得到磁学系统的各种热力学函数，比如焓、亥姆霍兹自由能、吉布斯自由能等等。这里只给出磁学系统的吉布斯自由能，因为对于我们即将发展的磁学系统的统计力学而言，它是最为重要的热力学函数：

$$dG = -SdT - MdH \quad (5.2)$$

以及从吉布斯自由能出发得到的系统磁矩表达式：

$$M = - \left( \frac{\partial G}{\partial H} \right)_T \quad (5.3)$$

磁学系统的基本热力学函数大概就是以上这些，但一个物理系统我们非常关心这些热力学函数对外界变化的响应，而磁学系统对变化外场  $H$  的响应就是磁感应系数  $\chi$ ：

$$\chi = \frac{kT}{N} \left( \frac{\partial M}{\partial H} \right)_T = \frac{1}{N} \langle \delta M^2 \rangle \quad (5.4)$$



如 (??) 节中所讲到的那样，响应函数正比于系统涨落。

## 5.2 Ising 模型

### 5.2.1 基本模型

Ising 模型是一个非常简单的模型，它的基本假设是：自旋只有两个状态，即  $S_i = \pm 1$ ，并且自旋之间的相互作用只有最近邻的相互作用能  $J$ ，其余的相互作用能统统忽略。这样一来，Ising 模型一个微观态  $\nu$  的能量分为两部分：外场作用和相互作用：

$$E_\nu = - \sum_i \mu H S_i - J \sum_{\langle i,j \rangle} S_i S_j \quad (5.5)$$

其中  $\langle i,j \rangle$  表示求和只对最近邻自旋进行。这里我们假设外场  $H$  是均匀的，即  $\mu H$  是一个常数，并且  $J > 0$ ，表明自旋倾向于平行排列。

那么很容易写出 Ising 模型的配分函数：

$$Q = \sum_{\nu} e^{-\beta E_\nu} = \sum_{\{S_i\}} e^{\mu\beta H \sum_i S_i + \beta J \sum_{\langle i,j \rangle} S_i S_j} \quad (5.6)$$

在后面的讨论中，我们设  $\mu\beta H = K$ ， $\beta J = K'$ 。从配分函数的表达式中很容易看出系统的磁矩  $M$  可以写为：

$$\langle M \rangle = \left( \frac{\partial \ln Q}{\partial K} \right)_{K'} \quad (5.7)$$

同时回顾 (??) 式，我们就能建立起热力学函数与配分函数的关系：

$$G = -kT \ln Q(P, H) \quad (5.8)$$

### 5.2.2 熵焓竞争与对称性破缺

有了以上的讨论，我们在外场  $H = 0$  的条件下先来说明 **熵焓竞争** 怎样影响磁学系统的相变。总体而言，一维 Ising 模型不能相变的原因是熵占优，而二维 Ising 模型能相变的原因是焓占优，并且这样的竞争关系的根源在于边界的维数。我们的思路非常简单：假设系统处于有序状态，然后考虑打乱一部分后体系的自由能变化，如果自由能下降，那么系统就会倾向于无序状态，熵占优，无相变；反之则倾向于有序状态，焓占优，有相变。

**对于周期边界的一维 Ising 链**，一个区域无论大小，永远只有两个端点，所以如果我们将一致向上的  $N$  个自旋磁子取出一部分磁畴翻转下去，那么相对最初的状态，能量变化为  $4J$ 。由于这样做有  $N$  个可能的构型，那么熵的变化为  $k \ln N$ 。而

$$\delta G = \delta E - T\delta S = -4J + kT \ln N$$

显然，我们可以定义一个临界温度  $T_c$ ，在  $2J = kT_c \ln N$  就是系统相变发生的地方：

$$T_c = \frac{4J}{k \ln N} \quad (5.9)$$

可以看到，在热力学极限下， $T_c = 0$ ，也就是说，一维 Ising 模型在任何温度下都不会发生相变， $T = 0$  是一个平凡解。

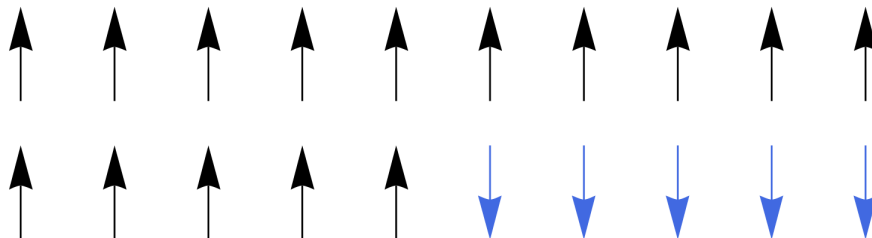


图 5.1: 一维 Ising 模型的熵焓竞争

对于周期边界的二维 Ising 链，一个区域的边界大小是随着区域大小的增加而增加的，所以如果我们将一致向上的  $N$  个自旋磁子取出一部分磁畴翻转下去，那么相对最初的状态，能量变化  $\delta E \propto N^{1/2}$ 。但由于最多只能有  $N^\alpha$  种可能的方式达成这一点，那么熵的变化为  $\alpha k \ln N$ 。这个时候，再写出自由能的变化：

$$\delta G = \delta E - T\delta S = -N^{1/2} + \alpha k T \ln N$$

在热力学极限下，第一项就会超过第二项，也就是能量占优，磁畴的翻转变得困难，所以更倾向于保持同向，也就是相变。

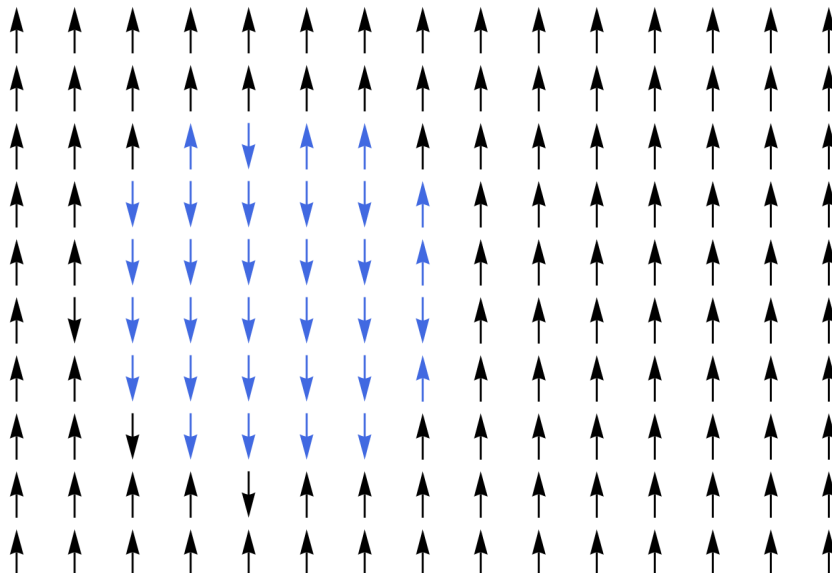


图 5.2: 二维 Ising 模型的熵焓竞争

这也就解释了为什么我们实际观察到体系的对称性要小于系统哈密顿量的对称性，也就是 **对称破缺** 的问题。磁学系统哈密顿量虽然具有自旋翻转的对称性（系统所有自旋全部翻转，总能量不变），

但要想在实际中体现这一点，[必须要求这个态要能够自由地变换到另一个对称态](#)。但如前所见，这样的翻转动作将会面临巨大的能垒，且在热力学极限下能垒高度为正无穷，所以动力学上不能发生。用一个公式来表示就是：

$$\langle M \rangle = \lim_{H \rightarrow 0} \lim_{N \rightarrow \infty} \left( -\frac{\partial G}{\partial H} \right)_T \neq 0 \quad (5.10)$$

在  $N$  有限的情况下，系统可以随意翻越能垒，所以磁矩可以为 0；现在为系统增加一个外场，那么系统将会聚集在某一边。现在让  $N \rightarrow \infty$ ，把门关上，系统就被困在了一个状态，纵使外场消失，系统即使想回去，也回不去了。所以宏观磁矩不为 0，这就是对称性破缺的本质。

### 5.2.3 一维 Ising 模型的严格解

下面我们正式进入一维 Ising 模型的严格解。我们先来考虑无外场周期性边界一维 Ising 模型的配分函数：

$$Q_N = \sum_{S_1, S_2, S_3, \dots, S_N} \exp [K'(S_1 S_2 + S_2 S_3 + \dots + S_N S_1)] \quad (5.11)$$

我们将它改写为如下形式：

$$Q_N = \sum_{S_1, S_2, S_3, \dots, S_N} \exp(K' S_1 S_2) \exp(K' S_2 S_3) \dots \exp(K' S_N S_1) \quad (5.12)$$

这样一来，我们就可以将配分函数视为 [张量缩并](#)。为了看出这一点，将  $\exp(K' S_1 S_2)$  写成矩阵的形式：

$$T(S_1, S_2) = \exp(K' S_1 S_2) = \begin{pmatrix} e^{K'} & e^{-K'} \\ e^{-K'} & e^{K'} \end{pmatrix} \quad (5.13)$$

这叫做 [转移矩阵](#)，再把被求和的  $\{S_1, S_2, \dots, S_N\}$  视为哑标，那么配分函数就可以写成：

$$Q_N = \sum_{S_1, S_2, S_3, \dots, S_N} T(S_1, S_2) T(S_2, S_3) \dots T(S_N, S_1) = \text{Tr}(T^N) = \lambda_1^N + \lambda_2^N \quad (5.14)$$

其中  $\lambda_1, \lambda_2$  是转移矩阵的两个本征值，易得

$$\lambda_{1,2} = e^{K'} \pm e^{-K'} \quad (5.15)$$

所以配分函数可以写成

$$Q_N = (2 \cosh K')^N + (2 \sinh K')^N \quad (5.16)$$

如果有外场，那么我们需要对外场所导致的额外能量做一点变换，因为要想构成二维张量，需要两个自旋：

$$E = - \sum_{i=1}^N J S_i S_{i+1} - \mu H \sum_{i=1}^N S_i = - \sum_{i=1}^N J S_i S_{i+1} - \frac{\mu H}{2} \sum_{i=1}^N (S_i + S_{i+1})$$

其中  $S_{N+1} = S_1$ 。转移矩阵就变成了

$$T(S_1, S_2) = \exp\left(\frac{1}{2}K(S_1 + S_2) + K' S_1 S_2\right) = \begin{pmatrix} e^{K+K'} & e^{-K'} \\ e^{-K'} & e^{-K+K'} \end{pmatrix} \quad (5.17)$$

本征值

$$\lambda_{\pm} = e^{K'} \left( \cosh K \pm \sqrt{\sinh^2 K + e^{-4K'}} \right) \quad (5.18)$$

在热力学极限下，较小的特征值  $\lambda_2$  将会被忽略，所以配分函数可以写成

$$Q = e^{NK'} \left( \cosh K + \sqrt{\sinh^2 K + e^{-4K'}} \right)^N \quad (5.19)$$

由此可得磁学系统的各热力学函数：

$$\begin{aligned} G &= -kT \ln Q = -NkT \ln(e^{\beta J} + e^{-\beta J}), \\ S &= - \left( \frac{\partial G}{\partial T} \right)_H = Nk [\ln(e^{\beta J} + e^{-\beta J}) - \beta J \tanh \beta J], \\ U &= G + TS = NJ \tanh \beta J, \\ C_H &= T \left( \frac{\partial S}{\partial T} \right)_H = Nk\beta^2 J^2 \operatorname{sech}^2 \beta J, \\ \langle M \rangle &= \frac{\sinh \mu\beta H}{\sqrt{\sinh^2 \mu\beta H + e^{-4\beta J}}}, \\ \chi &= \left( \frac{\partial \langle M \rangle}{\partial H} \right)_T = \frac{\beta e^{-2\beta J}}{\sqrt{\sinh^2 \mu\beta H + e^{-4\beta J}}} \end{aligned}$$

在  $\beta \rightarrow 0$  时， $S \rightarrow Nk \ln 2$ ，表明系统有  $2^N$  的配容数；在  $\beta \rightarrow \infty$  时， $S \rightarrow 0$ ，表明系统处于绝对零度。

可以看到，在  $H \neq 0$  时，响应函数在温度趋于 0 时收敛为 0，而在  $H = 0$  时发散。

#### 反思与质疑

容易看出，转移矩阵方法其实也可以用在非周期条件上，只不过不再是张量的求和缩并为迹，而是张量的求和缩并为张量。这样一来，我们就可以将一维 Ising 模型的配分函数写成

$$Q_N = \sum_{S_1, S_2, S_3, \dots, S_N} T(S_1, S_2) T(S_2, S_3) \cdots T(S_{N-1}, S_N) = \sum_{S_1, S_N} T^{N-1} \quad (5.20)$$

也就是矩阵  $T^{N-1}$  的所有项之和。

当然，转移矩阵并不一定是 2 维的，所以如果系统有三个可能状态， $S_i$  的可能取值为 1, 0, -1，那么无外场转移矩阵为

$$T = \begin{pmatrix} e^{K'} & 1 & e^{-K'} \\ 1 & 1 & 1 \\ e^{-K'} & 1 & e^{K'} \end{pmatrix} \quad (5.21)$$

而实对称阵必能正交相似对角化。

### 5.2.4 二维 Ising 模型概要

二维 Ising 模型的解析解由杨振宁完成，这里只给出一些基本结果，不做推导。

在  $H = 0$  的条件下，二维 Ising 模型的配分函数为

$$Q(N, \beta, 0) = [2 \cosh(2\beta J e^I)]^N \quad (5.22)$$

其中

$$I = \frac{1}{2\pi} \int_0^\pi \ln \left[ \frac{1}{2} \left( 1 + \sqrt{1 - K(J)^2 \sin^2 \theta} \right) \right] d\theta,$$

$$K(J) = \frac{2 \sinh 2\beta J}{\cosh^2 2\beta J}$$

当  $K(J) = 1$  时，系统可能出现奇异性，也即相变点满足

$$\sinh 2\beta_c J = 1 \quad (5.23)$$

数值结果约为

$$kT_c \approx 2.269J \quad (5.24)$$

而每个自旋的平均值为

$$m = \frac{\langle M \rangle}{N} = \begin{cases} 0 & T > T_c \\ [1 - \sinh^{-4} 2\beta J]^{1/8} & T < T_c \end{cases} \quad (5.25)$$

## 5.3 平均场理论

Ising 模型的解析解的求解难度随着维数的增加而增加，三维 Ising 模型没有解析解，只能通过数值模拟的方法完成。这表明 Ising 模型本身仍然需要做一定的近似，使得高维 Ising 模型也有近似的解析解。

**平均场理论** 就是这样一种近似方法。回想我们在无相互作用系统中的高光时刻，如果我们能将每个 Ising 磁子受到的相互作用以一个等价的平均外场来替代，那么我们就可以将 Ising 模型转化为一个无相互作用的模型，配分函数瞬间就变为了单粒子配分函数的  $N$  次方。

### 5.3.1 基本结果

现在我们来将单粒子近似用到统计力学里来（Hartree-Fork 理论或者金属能带论都有涉及）。既然是平均场，那么自旋势必需要做一点平均：

$$\langle S_i \rangle = \frac{\langle M \rangle}{N} \equiv m \quad (5.26)$$

这里的  $m$  有一个特殊的名称：[序参量](#)。将字面意义与物理意义结合起来看，这个量刻画了系统的有序程度。设每个磁子周围都有  $z$  个其他磁子，那么磁子  $i$  的能量为

$$E_i = -\mu H S_i - J \sum_j S_i S_j = -\mu H S_i - z J m S_i \quad (5.27)$$

单粒子配分函数与  $S_i$  的平均值为

$$Q_1 = \sum_{S_i=\pm 1} e^{\beta \mu H S_i + \beta z J m S_i} = 2 \cosh(\beta \mu H + \beta z J m), \quad (5.28)$$

$$\langle S_i \rangle = \frac{\partial \ln Q_1}{\partial (\beta \mu H)} = \tanh(\beta \mu H + \beta z J m) \quad (5.29)$$

然而，系统的每个磁子都是等价的，所以  $S_i$  本身的平均值也应当是  $m$ 。换言之，为了使得我们的理论自洽，必须有如下的[自洽场方程](#)：

$$\begin{aligned} m &= \tanh(\beta \mu H + \beta z J m) \\ \beta &= \frac{1}{2zJm} \ln \frac{1+m}{1-m} \end{aligned} \quad (5.30)$$

在外场为 0 的条件下，显然  $m = 0$  为平凡解。要想得到非平凡解，根据函数的性质，必须要求

$$\left. \frac{\partial \tanh(\beta z J m)}{\partial m} \right|_{m=0} > 1 \Rightarrow \beta z J > 1 \Rightarrow T_c = \frac{zJ}{k} = \begin{cases} \frac{2J}{k} & \text{one dimation} \\ \frac{4J}{k} & \text{two dimation} \\ \frac{6J}{k} & \text{three dimation} \end{cases} \quad (5.31)$$

### 反思与质疑

三维 Ising 模型的临界温度的模拟结果为  $4J/k$ ，可见平均场理论总是高估相变温度。这是因为它仅仅考虑一个磁子在平均场下的涨落，而忽略了其他磁子的涨落——将取向一致的一碗米粒稍微摇晃一下，米在碗中的高度将会继续下降。

### 5.3.2 临界行为

### 5.3.3 平均场的变分方法

## 5.4 重正化方法

[响应函数  \$\chi\$](#)  的发散究竟反映了什么物理？

“系统本身的涨落越大，那么对外界的响应就越灵敏”这句话，实际上蕴含了这样一则惊天秘闻：系统对外界的响应是一种宏观性质，当整个系统都“合作起来”对外界变化产生响应时，意味着能参与合作的范围，也就是[关联尺度](#)，开始发散。一个例子便是液体的临界乳光现象，透明的水在临界点附近变为乳白色，这表明本应当从水分子的间隙中漏过大部分的自然光，却大部分被散射了。这

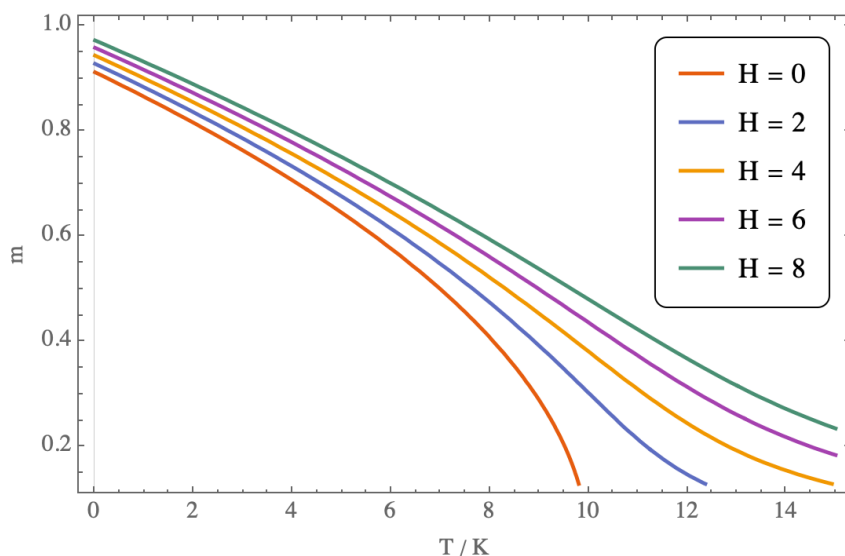


图 5.3: 平均场下一维 Ising 模型序参量随温度的变化情况

样的现象只有一种解释：水分子不再是孤立的，而是形成了一个的团簇，这些团簇的尺度越来越大，最后甚至大到与光的波长相当，从而产生了散射——把渔网加密，就能捞到更小的鱼。

那么问题来了：如果水分子的关联尺度发散，那么多少水分子能算一个团簇？这相当于询问临界水的 **特征尺度** 是多少。仔细想想，我们发现这个问题无法回答，因为关联尺度发散意味着一个水分子和无穷远的另一个水分子都有相互作用，这个特征尺度是不存在的。

特征尺度一旦不存在，我们发现系统将会呈现出 **自相似** 的特征（例如分形），这也是相变非常典型的特征：用不同尺度的放大镜来观察相变，得到的结果都是相似的。

这样一来，在相变点附近，多自由度系统能够合理合法地约化到低自由度系统，并且相变点系统的形貌在 **尺度变换** 下不变。

也就是说：**相变点是尺度变换的不动点。**

## 5.5 响应函数与关联函数

自旋作为两个随机变量，它们之间的关联性，可以用 **协方差** 来描述：

$$\text{Cov}(S_1, S_2) = \langle S_1 S_2 \rangle - \langle S_1 \rangle \langle S_2 \rangle \quad (5.32)$$

平均场近似认为各自旋变量都是独立的，所以自然有  $\text{Cov}(S_1, S_2) = 0$ 。下面我们来讨论响应函数和关联函数之间的联系，只需指出：

$$\begin{aligned}
 \chi &= \frac{1}{N} \langle (M - \langle M \rangle)^2 \rangle \\
 &= \frac{\mu^2}{N} \left\langle \left[ \sum_i^N S_i - \langle S_i \rangle \sum_j^N S_j - \langle S_j \rangle \right] \right\rangle \\
 &= \frac{\mu^2}{N} \sum_{i,j} \text{Cov}(S_i, S_j) = \mu^2 \sum_j^N \text{Cov}(S_1, S_j)
 \end{aligned} \tag{5.33}$$

那么响应函数是关联函数的总和，响应函数要想发散，各处的关联函数必须全都大于 0，这也表明关联尺度发散。

### 5.5.1 重正化群 RG



## 附录 A 附录

### A.1 无穷小正则变换行列式的绝对值等于 1

将系统的状态参量  $(q_i, p_i), i = 1, 2, \dots, n$  统一表达为  $\xi_\alpha, \alpha = 1, 2, \dots, 2n$ ，其中前半是坐标  $q$ ，后半是动量  $p$ 。引入求和约定，则经典 Poisson 括号可以写为

$$[f, g] = \frac{\partial f}{\partial \xi_\alpha} \Omega_{\alpha\beta} \frac{\partial g}{\partial \xi_\beta}$$

其中

$$\Omega = \begin{bmatrix} & -I \\ I & \end{bmatrix}_{2n \times 2n}$$

设正则变换将  $\{\xi\}$  变换为  $\{\xi'\}$ ，雅可比矩阵  $M$  应为

$$M = \begin{bmatrix} \frac{\partial \xi'_1}{\partial \xi_1} & \frac{\partial \xi'_1}{\partial \xi_2} & \dots & \frac{\partial \xi'_1}{\partial \xi_{2n}} \\ \frac{\partial \xi'_2}{\partial \xi_1} & \frac{\partial \xi'_2}{\partial \xi_2} & \dots & \frac{\partial \xi'_2}{\partial \xi_{2n}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial \xi'_{2n}}{\partial \xi_1} & \frac{\partial \xi'_{2n}}{\partial \xi_2} & \dots & \frac{\partial \xi'_{2n}}{\partial \xi_{2n}} \end{bmatrix}_{2n \times 2n}$$

由线性代数知识可知， $\det(M\Omega M^T) = \det(M)^2 \det(\Omega) = -\det(M)^2$ ，同时它又可以写为

$$\begin{aligned} (M\Omega M^T)_{ij} &= M_{i\alpha} \Omega_{\alpha\beta} M_{\beta j}^T = M_{i\alpha} \Omega_{\alpha\beta} M_{j\beta} \\ &= \frac{\partial \xi'_i}{\partial \xi_\alpha} \Omega_{\alpha\beta} \frac{\partial \xi'_j}{\partial \xi_\beta} \\ &= [\xi'_i, \xi'_j]_\xi \\ &= \Omega_{ij} \end{aligned}$$

最后一个等式用到了正则变换的基本 Poisson 括号不变性。故  $\det(M)^2 = 1$ ，证明完毕。

根据系统演化的连续性，我们还可以说明：对于系统经历的每一个无穷小变换的 Jacobi 矩阵，其行列式保持同号。