1 习题解答

1. 实对称 Wigner 矩阵 $A_n = (a_{ij})_{i,j=1}^n$ 满足条件:
(i)$\{a_{ij} : 1 \le i \le j \le n\}$ 相互独立，$\{a_{ij} : 1 \le i < j \le n\}$ 同分布，$\{a_{ii} : 1 \le i \le n\}$ 同分布; (ii)$\mathbb{E}(a_{11}) = 0, \mathbb{E}(a_{12}) = 0,\ \mathrm{Var}(a_{11}) < \infty, \mathrm{Var}(a_{12}) = 1$; (iii) a_{11} 和 a_{12} 的各阶矩均存在. 令

$$\| A_n \|_2 = \sup_{\|\vec{v}\|=1, \vec{v}\in\mathbb{R}^n} \| A_n \vec{v} \|,$$

试证明 $\forall \delta > 0$ 有

$$\lim_{n\to\infty} \mathbb{P}\left(\| A_n \|_2 \ge n^{\frac{1}{2}+\delta}\right) = 0 \ .$$

Proof. 考虑到

$$\max_i \lambda_i^{2k} \le \sum_i \lambda_i^{2k} = trA_n^{2k}, \tag{1}$$

我们有:

$$\mathbb{P}\left(\| A_n \|_2 \ge n^{\frac{1}{2}+\delta}\right) = \mathbb{P}(\max_i |\lambda_i|^{2k} \ge n^{k(\frac{1}{2}+\delta)}) \le \mathbb{P}(trA_n^{2k} \ge n^{2k(\frac{1}{2}+\delta)}) \le \frac{\mathbb{E}(trA_n^{2k})}{n^{k+2k\delta}}. \tag{2}$$

而我们知道 $trA_n^{2k} \sim C_k n^{k+1}$，故取 k 使得 $2k\delta > 1$ 即可. □

2. GUE 矩阵 (Gaussian Unitary Ensemble) 指复矩阵 $X = (x_{ij})_{i,j=1}^n$, 这里 $\{\mathrm{Re}x_{ij}, \mathrm{Im}x_{ij} : i,j = 1,\dots,n\}$ 为独立同标准正态分布. 令

$$H = \frac{1}{2}(X + X^*),$$

试证

(i) H 矩阵元联合密度为

$$f(H) = 2^{-\frac{n}{2}} \pi^{-\frac{n^2}{2}} e^{-\frac{1}{2}trH^2} \ .$$

(ii) (酉群不变性) 任给酉矩阵 $\mathrm{Q} \in \mathrm{U}(n)$ 有 QHQ^{-1} 与 H 同分布.

Proof. 作业中很多是通过证明 QX 和 X 同分布，进而得到 QHQ^{-1} 是同分布的。另一方面有一部分同学用 $trH^2 = tr(QHQ^{-1})^2$ 来说明，这是有问题的，换元时同时需要考虑 Jacobi 矩阵。

另外，由于我们考虑的是多重正态分布，他的分布完全由协方差矩阵所确定了，所以实际上我们只需要计算换元之后的协方差。设 $QHQ^{-1} = (a_{ij})$，由于 $Q^{-1} = \bar{Q}^t$ 则

$$a_{ij} = \sum_{k_1,k_2} q_{ik_1} a_{k_1k_2} \bar{q}_{jk_2}. \tag{3}$$

考虑协方差

$$\begin{aligned}Cov(a_{ij}, a_{st}) &= \mathbb{E}(\sum_{k_1,k_2} q_{ik_1} h_{k_1k_2} \bar{q}_{jk_2})(\sum_{k_3,k_4} q_{sk_3} h_{k_3k_4} \bar{q}_{tk_4}) \\ &= \sum_{k_1,k_2,k_3,k_4} \mathbb{E}(q_{ik_1} \bar{q}_{jk_2} q_{sk_3} \bar{q}_{tk_4} h_{k_1k_2} h_{k_3k_4})\end{aligned} \tag{4}$$

和式里面的项只有在 $(k_1, k_2) = (k_4, k_3)$ 的时候才不为 0。所以上式等于

$$\begin{aligned}Cov(a_{ij}, a_{st}) &= \sum_{k_1,k_2,k_3,k_4} \mathbb{E}(q_{ik_1} \bar{q}_{jk_2} q_{sk_3} \bar{q}_{tk_4} h_{k_1k_2} h_{k_3k_4}) \\ &= \sum_{k_1 \neq k_2} q_{ik_1} \bar{q}_{tk_1} \bar{q}_{jk_2} q_{sk_2} + \sum_{k_1} q_{ik_1} \bar{q}_{tk_1} \bar{q}_{jk_1} q_{sk_1} \\ &= (\sum_{k_1} q_{ik_1} \bar{q}_{tk_1})(\sum_{k_2} \bar{q}_{jk_1} q_{sk_1}) = \delta_{it} \delta_{js}.\end{aligned} \tag{5}$$

同样再考虑

$$Cov(a_{ij}, \bar{a}_{st}) = \delta_{is} \delta_{jt}. \tag{6}$$

即可证明不变。 □

3. Wishart 模型 $X = (x_{ij})_{1 \le i \le p, 1 \le j \le n}$，矩阵元 $\{x_{ij}\}$ 为独立同 $N(0,1)$，对固定的 $\alpha = n - p \ge 0$，证明当 $p \to \infty$ 时

$$\frac{1}{p} \mathbb{E}\left[tr \left(\frac{XX^t}{p} \right)^m \right] \to C_m := \frac{1}{m+1} \binom{2m}{m} .$$

如果困难的话，只需验证 $m = 2\ 3$ 特殊情形即可.

Proof. 见上次习题课的 Wishart 矩阵一节。因为这时候 $\frac{n}{p} \to 1$, 所以是半圆率的情况。 □

4. 厄米 Wigner 矩阵 $A_n = (a_{ij})_{i,j=1}^n$ ，$A_n = A_n^*$.

   - ▷ 实 $\{a_{ii}\}$ 独立同，与 Y 同分布；
   - ▷ $\{\mathrm{Re} a_{ij}\ , \mathrm{Im} a_{ij}\}_{i<j}$ 独立同，与 Z 同分布；
   - ▷ $\{a_{ii} : 1 \le i \le n\} \cup \{\mathrm{Re} a_{ij}, \mathrm{Im} a_{ij} :\ 1 \le i < j \le n\}$ 独立；
   - ▷ $\mathbb{E}[Y] = 0$, $\mathbb{E}[Z] = 0$, $var(Y) < \infty$, $var(Z) = \frac{1}{2}$;
   - ▷ $\forall k \ge 3$, $\mathbb{E}[|Y|^k]$,$\mathbb{E}[|Z|^k] < \infty$.

   试证明 Wigner 半圆律同样成立，即有

$$\frac{1}{n} \mathbb{E}\left[tr \left(\frac{A_n}{\sqrt{n}} \right)^k \right] \to \gamma_k = \int x^k \frac{1}{2\pi} \sqrt{1 - x^2} dx .$$

Proof. 注意到在实 Wigner 矩阵的证明过程中，主要路径的边都是一来一回的，这样就可以把 a_{ij}^2 变成 $|a_{ij}|^2$. □

2 熵与物理

2.1 热力学中的熵

熵是一个神秘而有趣的概念，关于它能讲的故事数不胜数，但最为容易理解的解释还是从物理出发。为此，我们要回顾热学中熵的定义，而在之后的统计力学之中，我们便能给予热力学这一现象科学（包括熵）以一定的概率/数学解释。

我们知道，热学中广为流传的一条定律为：

定律 2.1. *在绝热体系（孤立体系）中，系统总是朝着熵增加的状态演化。*

我们要将这一定理形式化进行描述，需要回顾热学里一些基础的概念与定义。这一部分的表述读者可以参考统计物理的教材 [1]。

定律 2.2. *(热力学第一定律) 设系统的内能为 U，我们考虑系统经历的一个无穷小过程，则我们有如下等式：*

$$dU = đQ + đW \tag{7}$$

其中 $đQ$ 为该过程中系统从外界吸收的热量，$đW$ 为该过程外界对系统所做的功。

上面这个定律实际上是说明热量是能量的一种以及其满足能量守恒定律。

而在对热机的研究中人们发现在两个恒温热源工作的热机的热效率是是有上限的，设热机在做功过程中，从温度 T_1 的恒温热源吸收 Q_1 热量，再在 $T_2(< T_1)$ 热源中放出 Q_2 热量，经过对卡诺热机的分析，我们可以得到热效率 η 会满足：

$$\eta = 1 - \frac{Q_2}{Q_1} \leq 1 - \frac{T_2}{T_1} \tag{8}$$

等号在理想可逆热机中才能取到，再给 Q 根据吸收放出加上符号，我们得到克劳修斯不等式

定理 2.1 (克劳修斯不等式)**.**

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} \leq 0. \tag{9}$$

考虑多个热源和连续的情况就有：

$$\sum_{i=1}^{n} \frac{Q_i}{T_i} \leq 0, \tag{10}$$

$$\oint \frac{dQ}{T} \leq 0 \tag{11}$$

从而我们得到一个只减不增的过程量。而对于两个状态 A, B 间的可逆过程，这个积分与路径无关，从而可以得到一个只与状态有关的态函数–熵。

定义 2.1. *我们如下定义熵 S_A，其中 A 为系统状态。*

$$S_B - S_A = \int_A^B \frac{dQ}{T} \tag{12}$$

其中积分对任一从 A 到 B 的可逆过程。可逆过程中微分形式为:

$$dS = \frac{dQ}{T}. \tag{13}$$

对于任意过程有:

$$S_B - S_A \geq \int_A^B \frac{dQ}{T} \tag{14}$$

$$\text{đ}S \geq \frac{dQ}{T}. \tag{15}$$

这样我们知道在系统不与外界发生热量交换时，也就是系统热量 $dQ = 0$，我们有 $dS \geq 0$，也就是熵不会减少。说明使熵值最大的状态是该系统的典型状态（平衡/稳定状态）。(通过熵可以唯象解释气体的三相相变)

定理 2.2. 孤立绝热系统处在稳定平衡状态的充分必要条件为 $\Delta S < 0$。即熵值是局部最大点。

同时为了方便寻找其他条件下系统的典型状态，我们通过熵的不减得到其他单调量。以气体为例，我们假设

$$dW = -pdV \tag{16}$$

从而 $dU = dQ - pdV$，我们定义其他一系列重要物理量

定义 2.2. *(i)* 定义焓 $H = U + pV$，从而

$$dH = dU + pdV + Vdp = dQ + Vdp \leq TdS + Vdp, \tag{17}$$

意味着在常压常熵条件下系统向焓值变小演化。

(ii) 定义自由能 $F = U - TS$，则

$$dF \leq -SdT - pdV, \tag{18}$$

表示在恒温恒容条件下系统向自由能变小演化。

(iii) 定义吉布斯自由能 $G = U - TS + pV$，则

$$dG \leq -SdT + Vdp \tag{19}$$

恒温恒压下系统朝吉布斯自由能变小演化。

2.2 统计力学中的熵

熵在热学之中有着非常重要的地位，但是同样也具有神秘色彩。统计物理给了我们一种全新视角去从微观审视宏观物理现象，解释在大量的微观粒子聚集的时候，系统之外的宏观属性和现象如何涌现而出。

统计物理有一个基本假设，等概率原理

定律 2.3 (等概率原理)**.** 对于处于平衡态的孤立系，系统各个可能的微观状态出现的概率相等。

而每个状态（比如说某个能量）会有一定的可能状态数目，可能的微观状态数目最多的那个就是最概然分布/热学里的平衡态。我们以独立经典粒子的 Boltzmann 统计为例解释这一原理。

例 1. 设 $\{\epsilon_l\}$ 为粒子具有的一系列能级，每个能级 ϵ_l 有 ω_l 个简并态（和这个能级非常接近的状态数目），我们设该系统在各个能级有 $\{a_l\}$ 个粒子。则对于 $\{a_l\}$ 的情况，系统总共有 $\omega_l^{a_l}$ 中可能的微观状态。我们把这 N 个粒子分成 $\{a_l\}$ 有 $\frac{N!}{\Pi_l a_l!}$，故总状态数为

$$\Omega = \frac{N!}{\Pi_l a_l!}\Pi_l \omega_l^{a_l}.$$

我们现在要考虑在一定条件约束下微观状态数目最多的 $\{a_l\}$，限制条件为：

$$\sum_l a_l = N,\ \ \sum_l a_l\epsilon_l = E. \tag{20}$$

取对数做近似

$$\begin{aligned}\log\Omega &= \log N! - \sum_l \log a_l! + \sum_l a_l\log\omega_l \\ &\approx N(\log N - 1) - \sum_l a_l(\log a_l - 1 + \log\omega_l).\end{aligned} \tag{21}$$

用拉格朗日乘子法求最大值，得到：

$$\begin{aligned}\delta\log\Omega &= -\sum_l \delta a_l\log a_l + \delta a_l + \delta a_l(\log\omega_l - 1) - \alpha\sum_l\delta a_l - \beta\sum_l\epsilon_l\delta a_l \\ &= \sum_l \delta a_l(-\beta\epsilon_l - \alpha + \log\omega_l - \log a_l) = 0\end{aligned} \tag{22}$$

从而

$$a_l = \omega_l e^{-\beta\epsilon_l - \alpha}. \tag{23}$$

可以知道 β 和温度有关，α 和温度与化学势（粒子数）有关。

由于 a_l 这种精细的东西都能够在统计力学的框架下清楚，我们想要由此重构出热学中的物理量同时依此确定 α,β 物理意义，我们定义配分函数：

定义 2.3 (配分函数)**.** 我们定义配分函数如下：

$$Z = \sum_l \omega_l e^{-\beta\epsilon_l}. \tag{24}$$

我们有

$$N = \sum_l a_l = \sum_l \omega_l e^{-\alpha-\beta\epsilon_l} = e^{-\alpha}Z \tag{25}$$

可以将内能重写如下：

$$\begin{aligned}U &= \sum_l \epsilon_l a_l \\ &= e^{-\alpha}\sum_l \epsilon_l \omega_l e^{-\beta\epsilon_l} \\ &= \frac{N}{Z}(-\frac{\partial}{\partial\beta}\omega_l e^{-\beta\epsilon_l}) \\ &= \frac{N}{Z}(-\frac{\partial Z}{\partial\beta}) = -N\frac{\partial}{\partial\beta}\log Z.\end{aligned} \tag{26}$$

而同样可以写出压强，体积微小改变会导致 ϵ_l 的微小改变：

$$\begin{aligned}p &= -\sum_l \frac{\partial\epsilon_l}{\partial V}a_l \\ &= -\sum_l \frac{\partial\epsilon_l}{\partial V}\omega_l e^{-\alpha-\beta\epsilon_l} \\ &= e^{-\alpha}\frac{1}{\beta}\frac{\partial\epsilon_l}{\partial V}\sum_l \omega_l e^{-\beta\epsilon_l} \\ &= \frac{N}{Z}(\frac{1}{\beta}\frac{\partial}{\partial V})Z = \frac{N}{\beta}\frac{\partial}{\partial V}\log Z.\end{aligned} \tag{27}$$

同时热量可以如下定义：

$$dQ = dU - Ydy = -Nd(\frac{\partial Z}{\partial\beta}) - \frac{N}{\beta}\frac{\partial\log Z}{\partial V}dV \tag{28}$$

用 β 乘上式得到

$$\beta(dU - Ydy) = -N\beta d(\frac{\partial Z}{\partial\beta}) - N\frac{\partial\log Z}{\partial V}dV \tag{29}$$

而对 Z 我们有：

$$d\log Z = \frac{\partial\log Z}{\partial\beta}d\beta + \frac{\partial\log Z}{\partial V}dV \tag{30}$$

比较可知

$$\beta(dU + pdV) = Nd(\log Z - \beta\frac{\partial}{\partial\beta}\log Z) \tag{31}$$

也就是说 $\beta(dU - Ydy)$ 是一个全微分，而我们知道 $dS = \frac{dQ}{T}$ 是一个全微分（熵是只与状态有关函数），从而积分因子 β 与 $1/T$ 成正比。同时得到

$$\begin{aligned}S &= kN\log Z - \beta\frac{\partial}{\partial\beta}\log Z \\ &= kN(\log N + \alpha) + k\beta U \\ &= k[N\ln N + \sum_l(\alpha + \beta\epsilon_l)a_l] \\ &= k[N\ln N + \sum_l a_l\log\omega_l - a_l\log a_l] \quad 利用 a_l = \omega_l e^{-\alpha-\beta\epsilon_l}\end{aligned} \tag{32}$$

与推导 Boltzmann 分布时式子比较：

$$\log \Omega = N \log N + \sum_l a_l \log \omega_l - \sum_l a_l \log a_l. \tag{33}$$

可以得到

定理 2.3 (Boltzmann 关系)**.**

$$S = k \log \Omega. \tag{34}$$

从而可以知道熵的统计意义：熵是系统混乱程度的量度。
同时读者可以自行验证如下命题

定理 2.4. *对于经典粒子的 Boltzmann 分布，熵函数可以表示为：*

$$S = -Nk \sum_s P_s \log P_s, \tag{35}$$

其中 P_s 是粒子处在状态 s 的概率，$P_s = \frac{e^{-\alpha-\beta\epsilon_s}}{N} = \frac{e^{-\beta\epsilon_s}}{Z_1}, \sum_s$ 为对所有状态求和。

这个命题实际上也告诉了我们物理熵和信息熵的联系。

2.3 信息中的熵

我们知道熵系统混乱程度的度量，这种解释也给我们在信息论角度一种模糊的观点。1948 年 Shannon 在 A Mathematical Theory of Communication 一文中引入了信息熵并证明了三个信息领域的重要定理 Shannon's source coding theorem, Noisy-channel coding theorem, Shannon–Hartley theorem. 分别描述了无失真条件下信息压缩的最大程度；传输过程中信道有噪音时信道信息传输率的上界；有失真情况下失真率较低编码的存在性。

我们可以介绍第一个定理无失真信源编码定理的证明。这部分内容可以参考 [2] 的第五章。

定义 2.4. *我们定义离散随机变量的信息熵如下：*

$$H(X) = -\sum_i p_i \log_2 p_i. \tag{36}$$

Shannon's source coding theorem 说的是这样一件事：

定理 2.5. *N 个 i.i.d. 随机变量 $\{X_n\}$ 在 $N \to \infty$ 时至多可以被压缩到比 $NH(X)$ 更长的长度而以极低的概率出现信息损失；如果他们被压缩到少于 $NH(X)$ 那么信息几乎一定会被丢失。*

证明这个定理我们需要一些关于编码的基本公式。

定理 2.6. 对于即时码，我们有 *Kraft* 不等式：

$$\sum D^{-l_i} \leq 1 \tag{37}$$

对于唯一可译码，我们有 *McMillan* 不等式

$$\sum D^{-l_i} \leq 1 \tag{38}$$

其中 D 是用来编码的字母表的字母数量，比如 *2* 截止就是 *2*，l_i 是第 i 个信息对应编码的码长。其中即时码指的是能够边读入边翻译的编码方式，即要求对任意的 $i \neq j$，i 的编码不能是 j 的前缀。唯一可译码指的是原来的数据串到字母表的字母序列是单射，也就是逆总是存在，总是可解。

Proof. 只要对唯一可译码证明即可，令 $l(x)$ 为 x 对应的编码长度，那么字符串 $(x_1, x_2, ..., x_k)$ 的长度为

$$l(x_1, x_2, ..., x_k) = \sum_{i=1}^{k} l(x_i), \tag{39}$$

我们要证明的不等式为：

$$\sum_x D^{-l(x)} \leq 1. \tag{40}$$

我们考虑左边的 n 次幂，有：

$$\begin{aligned}
(\sum_x D^{-l(x)})^n &= \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} D^{-l(x_1)} D^{-l(x_2)} \ldots D^{-l(x_n)} \\
&= \sum_{x_1, x_2 \ldots, x_n} D^{-l(x_1, x_2, \ldots, x_n)} \\
&= \sum_{m=1} a(m) D^{-m}
\end{aligned} \tag{41}$$

其中 $a(m)$ 是长为 n 字符串的编码长度为 m 的字符串个数，由于是唯一可译码，每个字符串至多对应一个编码，所以 $a(m) \leq D^m$. 从而

$$(\sum_x D^{-l(x)})^n \leq \sum_{m=1}^{n \max l(x)} D^m D^{-m} = n \max_i l_i. \tag{42}$$

令 $n \to \infty$ 即可证明。 □

我们计算在这个限制条件下，期望码长最小长度是多少，也就是要求 $\sum_i p_i l_i$ 在条件 $\sum_i D^{-l_i} \leq 1$ 下的最小值，其中 p_i 为第 i 个字符出现的概率。

简单的拉格朗日乘子法就能知道，$D = 2$ 时这个最小值就是信息熵 $H(X)$. 也就是说期望最小值就是 $H(X)$，定理的另外两半就是通过大偏差 (Large Deviation Principle) 和构造编码来证明，这里就不再赘述。

这些理解我们对熵可以有更多直观解释。比如说 $H(X, Y) \leq H(X) + H(Y)$ 等。

2.4 例题

依照如下的步骤证明 Boltzmann 分布：

（1）将粒子的动量空间和能量空间分为 $\hbar^3$ 的小的小块，近似看成每一个小块中包含一个量子态（由不确定性原理一个方向上的空间不确定度和动量不确定度乘积约为 $\hbar$）。

（2）根据单个粒子能量为 $\epsilon = \frac{m}{2}(v_x^2 + v_y^2 + v_z^2) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)$ 得到在 $[x+\Delta x] \times [y+\Delta y] \times [z+\Delta z] \times [p_x+\Delta p_x] \times [p_y+\Delta p_y] \times [p_z+\Delta p_z]$ 这一区间内粒子数正比于 $e^{-\frac{1}{kT}\Delta x\Delta y\Delta z\Delta p_x\Delta p_y\Delta p_z}$

（3）对空间坐标积分得到气体体积，对动量坐标做换元写成能量形式，得到 Boltzmann 分布：

$$f(v) \propto v^2 e^{-\frac{mv^2}{2kT}}, \tag{43}$$

其中 $f(v)$ 为粒子速度在 v 附近的概率。

3 Ising model

本节主要介绍 Ising Model 的定义和 $<\sigma_0\sigma_n>$ 的意义，以及一维 Ising model 的转移矩阵解法。

2 维 Ising model 是物理中出现相变最简单的 toy model，具体来说是在热力学极限下（粒子数趋向于无穷）时，一些物理量如熵、自由能等相对于温度变化出现一个突变（例如气体的三相，在相变点附近状态的熵值突变，所以需要额外吸热），而这种突变也能体现在配分函数 Z 之上。而要精细刻画这种相变，我们需要考虑不同温度下收敛到的随机场（random field）是如何不同，我们可以考虑 k 点关联函数 $<\sigma_{x_1}\sigma_{x_2}...\sigma_{x_k}>$(k-point correlation function) 来刻画这个随机场。在维数 $d \geq 5$ 时，我们知道 Ising model 在 critical 温度下对应的随机场收敛到 Gaussian free field，而 $d = 4$ 时我们知道他的边分布也是高斯的（Michael Aizenman, Hugo Duminil-Copin 2021 年发表在 Annals of mathematics 上的结果）。

这一节中我们用转移矩阵方法考虑 1 维的 Ising model，并且计算他的配分函数和点关联函数 $<\sigma_0\sigma_n>$. 具体可以参看 stanford 的lecture note.

我们首先定义哈密顿量 (Hamiltonian，这里可以看成是能量)：

$$H = -J\sum_j \sigma_j\sigma_{j+1}. \tag{44}$$

则每一个态出现的概率正比于 $e^{-\beta H}$，注意到 $H = \sum_j E(\sigma_j, \sigma_{j+1})$. 我们有

$$e^{-\beta H} = \Pi_j e^{-\beta E(\sigma_j,\sigma_{j+1})} =<\sigma_1|t|\sigma_2><\sigma_2|t|\sigma_3> ... <\sigma_n|t|\sigma_{n+1}> \tag{45}$$

其中 $\sigma_i = 1$ 和 $\sigma_i = -1$ 可以分别由矩阵

$$|\sigma = 1> \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{46}$$

以及

$$|\sigma = -1> \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{47}$$

来表示。而 t 可以看成是这样一个矩阵：

$$\begin{pmatrix} e^{\beta} & e^{-\beta} \\ e^{-\beta} & e^{\beta} \end{pmatrix} \tag{48}$$

同时我们记 $<\sigma| = |\sigma>^t$. 则有 $<\sigma_i|t|\sigma_{i+1}>= e^{-E(\sigma_i,\sigma_j)}$. 同时可以看到

$$\sum_{\sigma} |\sigma><\sigma| = I_2. \tag{49}$$

我们考虑边界条件 σ_1, σ_{n+1} 确定时的配分函数：

$$\begin{aligned} Z_{\sigma_1,\sigma_{n+1}}(\beta) &= \sum_{\sigma_i}^{n} e^{-\beta H} \\ &= \sum_{\sigma_i} \Pi_{j=1}^{n} <\sigma_j|t|\sigma_{j+1}> \\ &=<\sigma_1|t\ \Pi_{j=1}^{n-1}(\sum_{\sigma_j=\pm 1} |\sigma_{j+1}><\sigma_{j+1}|t)\ |\sigma_{n+1}> \\ &=<\sigma_1|t\ \Pi(tI_2)^{n-1}|\sigma_{n+1}>=<\sigma_1|t^n|\sigma_{n+1}>. \end{aligned} \tag{50}$$

这时候我们只要将矩阵 t 对角化，就能得到边界条件 σ_1,σ_{n+1} 确定时的配分函数了。而边界条件不确定时对边界条件求和就能确定整体的配分函数 $Z_n(\beta) \sum_{\sigma_1,\sigma_{n+1}} Z_{\sigma_1,\sigma_{n+1}}$.

而

$$<\sigma_1\sigma_{n+1}>= \frac{1}{Z_n(\beta)} \sum_{\sigma_1,\sigma_{n+1}} Z_{\sigma_1,\sigma_{n+1}}\sigma_1\sigma_{n+1}. \tag{51}$$

4 Laplace method/Saddle point method

我们再介绍一下 Laplace method 或者说 saddle point method。其主要关心

$$I(x) = \int_0^b e^{-xt} q(t) dt \tag{52}$$

在 $x \to \infty$ 的渐进行为，其中 $q(t)$ 是连续函数。可以想象他的渐进行为在 0 以外的区域都是指数级衰减。所以他的渐近行为只与 $q(t)$ 在 0 附近的渐近展开相关。我们有如下定理 (参考这篇文章)：

定理 4.1 (Watson's Lemma (Real version))**.** *Consider integrals of the form*

$$I(x) = \int_0^b f(t) e^{-xt} dt, \quad b > 0, b \neq \infty.$$

Assume that $f(t) \in C[0,b]$ *and* $f(t)$ *has the asymptotic expansion*

$$f(t) \sim t^{\alpha} \sum_{n=0}^{\infty} a_n t^{\beta n} \quad as\ t \longrightarrow 0^+$$

with $\alpha > -1$ and $\beta > 0$ so that the integral converges at $t = 0$. Then

$$I(x) \sim \sum_{n=0}^{\infty} a_n \frac{\Gamma(\alpha + \beta n + 1)}{x^{\alpha+\beta n+1}} \qquad \text{as } x \longrightarrow \infty.$$

Proof. We first split $I(x)$ as follows:

$$I(x) = \underbrace{\int_0^{\varepsilon} f(t)e^{-xt}dt}_{I(x;\varepsilon)} + \int_{\varepsilon}^{b} f(t)e^{-xt}dt$$

for some $\varepsilon > 0$. The second integral introduces exponentially small errors for any $\varepsilon > 0$:

$$\left| \int_{\varepsilon}^{b} f(t)e^{-xt}dt \right| \leqslant \|f\|_{\infty} \int_{\varepsilon}^{b} e^{-xt}dt = \|f\|_{\infty} \left(\frac{e^{-\varepsilon x} - e^{-bx}}{x} \right)$$

which converges to 0 exponentially for $x \longrightarrow \infty$. Next, we substitute the asymptotic expansion (3.4) into $I(x;\varepsilon)$ to obtain:

$$I(x;\varepsilon) = \sum_{n=0}^{N} \int_0^{\varepsilon} a_n t^{\alpha+\beta n} e^{-xt}dt + \left(I(x;\varepsilon) - \sum_{n=0}^{N} \int_0^{\varepsilon} a_n t^{\alpha+\beta n} e^{-xt}dt \right)$$

In particular, we can choose ε sufficiently small such that

$$\left| f(t) - t^{\alpha} \sum_{n=0}^{N} a_n t^{\beta n} \right| \leqslant K t^{\alpha} t^{\beta(N+1)} \qquad \text{for every } t \in [0, \varepsilon],$$

for some constant $K > 0$. Thus,

$$\begin{aligned}
\left| I(x;\varepsilon) - \sum_{n=0}^{N} \int_0^{\varepsilon} a_n t^{\alpha+\beta n} e^{-xt}dt \right| &\leqslant \int_0^{\varepsilon} \left| f(t) - t^{\alpha} \sum_{n=0}^{N} a_n t^{\beta n} \right| e^{-xt}dt \\
&\leqslant K \int_0^{\varepsilon} t^{\alpha+\beta(N+1)} e^{-xt}dt \\
&\leqslant K \int_0^{\infty} t^{\alpha+\beta(N+1)} e^{-xt}dt \\
&= \frac{K}{x^{\alpha+\beta(N+1)+1} \int_0^{\infty} s^{\alpha+\beta(N+1)} e^{-s}ds} \\
&= K \left(\frac{\Gamma(\alpha + \beta(N+1) + 1)}{x^{\alpha+\beta(N+1)+1}} \right) \\
&= O\left(\frac{1}{x^{\alpha+\beta(N+1)+1}} \right) \text{ as } x \longrightarrow \infty
\end{aligned}$$

Finally,

$$
\begin{aligned}
\int_0^{\varepsilon} a_n t^{\alpha+\beta n} e^{-xt} dt &= \int_0^{\infty} a_n t^{\alpha+\beta n} e^{-xt} dt - \int_{\varepsilon}^{\infty} a_n t^{\alpha+\beta n} e^{-xt} dt \\
&= \frac{a_n}{x^{\alpha+\beta n+1}} \int_0^{\infty} s^{\alpha+\beta n} e^{-s} ds + \underbrace{O\left(e^{-\varepsilon x}\right)}_{\text{as } x \to \infty} \qquad \text{Let } s = xt.] \\
&= \frac{a_n \Gamma(\alpha+\beta n+1)}{x^{\alpha+\beta n+1}} + \underbrace{O\left(e^{-\varepsilon x}\right)}_{\text{as } x \to \infty}
\end{aligned}
$$

Combining all the estimates leads to

$$
I(x) - \sum_{n=0}^{N} a_n \frac{\Gamma(\alpha+\beta n+1)}{x^{\alpha+\beta n+1}} \sim O\left(\frac{1}{x^{\alpha+\beta(N+1)+1}}\right) = o\left(\frac{1}{x^{\alpha+\beta(N+1)}}\right) \qquad \text{as } x \longrightarrow \infty.
$$

The desired result follows since N was arbitrary in the asymptotic representation.

□

References

[1] 汪志诚，热力学 · 统计物理（第五版），高等教育出版社，(2013).

[2] Thomas M. Cove 著，阮吉寿，张华译，信息论基础，机械工业出版社，(2008).