

Ch2 重要抽样 Monte Carlo 方法

Rainzor

- 1 系综理论（系综的概率密度分布、涨落、响应）
- 2 重要抽样（Markov链、主方程、细致平衡）
- 3 Ising模型（平均场理论、二级相变、临界涨落）
模拟退火法（自旋玻璃）

2.1 统计力学基础

系综是一个抽象概念，代表了大量性质相同的力学体系的集合，每个体系处于独立的运动状态(初始条件各不相同)。研究大量体系在相空间的分布，求其统计平均，是统计力学的基本任务

2.1.1 相空间理论

1. 相空间

经典统计力学考虑的是一个多自由度（原则上是无限多）的力学体系，这些自由度一般是粒子的坐标和动量，或者是磁矩即自旋。经典体系意指这些自由度是可对易的，**以这些自由度为坐标展开的空间即为相空间**。

如 N 个粒子的 $3N$ 个位置坐标 (q_1, \dots, q_{3N}) 和 $3N$ 个动量 (p_1, \dots, p_{3N}) 构成 $6N$ 维相空间，相空间中的一个点代表力学体系的微观状态，相应的 $6N$ 个坐标组成体系的一个构型。

每个坐标和动量的演变由经典力学的正则方程决定：

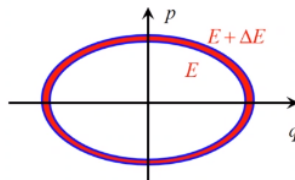
$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (1)$$

其轨迹的运动方向完全由速度矢量 $v = (\dot{q}, \dot{p})$ 给出，因而通过相空间中任一点的轨迹只能有一条。当力学体系从不同的初态出发时，在相空间中就沿着不同的轨迹而运动，这些轨迹是不相交的（否则自相交点出发有两条轨迹

对于能量守恒的保守系统，轨迹限于在相空间中由 $E = H(p, q)$ 确定的曲面上运动。如果总能处于 $(E, E + \Delta E)$ 的一个区域范围内，则轨迹限制于一个厚为 ΔE 的曲面壳层里：

例：一维谐振子

$$H(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 = E$$



2 统计系综 (Ensemble)

我们并不想知道所有粒子的坐标、动量、角速度等微观力学量，首先没有必要了解那么详细，另外这些量是不可测量的，只有平均的物理量如压力等才是可测的。测量结果是多个粒子在一段长时间内作用的平均效应，即使可以作瞬时测量，其瞬时值与平均值差别也是很小的（对于 $N \sim 10^{23}$ 的体系，涨落误差是 $1/\sqrt{N}$ ，但是**对于相变过程，涨落是不可以忽略的**）。

因此统计力学的中心思想是用几个宏观物理量（如粒子数 N 、体积 V 、温度 T 、压强 P 、能量 E 、化学势 μ 、比热 C 等）代替 $6N$ 个描述微观状态的自由度。

时间的平均：物理量的测量值

对物理量的测量本质上是对体系（相空间的代表点）随时间的演化的一条轨迹进行长时间平均：

$$\bar{A} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(q(t), p(t)) dt \quad (2)$$

系综概念

由大量性质完全相同的力学体系而构成的集合，每个体系各处在一运动状态而且是独立的。

密度分布

因为相空间中系综代表点有一定的密度分布，设其为 $\rho(p, q, t)$ ，求系综平均时须将它作为权重因子，因此宏观量的所有可能微观状态的系综平均值为

$$\langle A \rangle = \frac{\int A \rho dq dp}{\int \rho dq dp} \quad (3)$$

3 Liouville定理

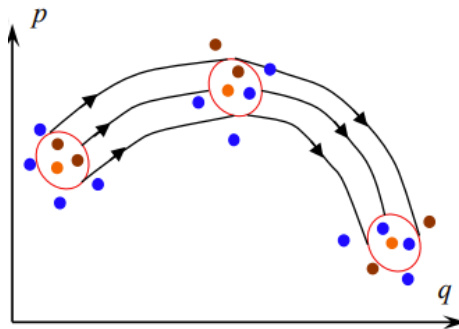
系综的几率密度在运动中不变

$$\frac{d\rho}{dt} = 0 \quad (4)$$

因为相空间中没有代表点的源或黑洞，代表点的总数应该是守恒的，按照正则方程证明，则是：

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + [\rho, H] \quad (5)$$

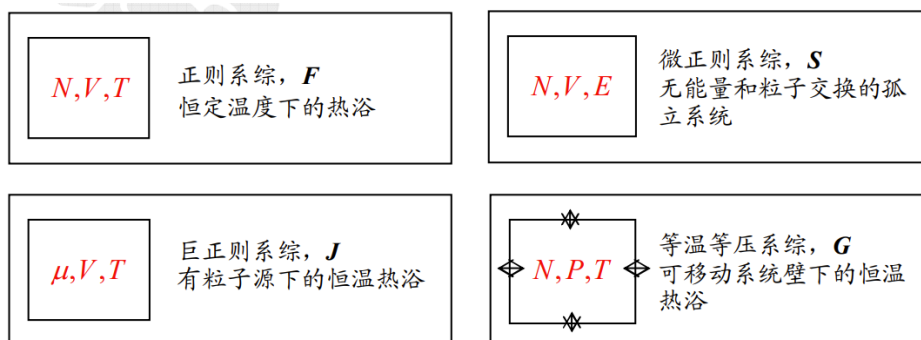
它的含义是，有一群代表点在一定的时间内由相空间中一个区域移到另一个区域，则移动前后各区域内的代表点密度保持不变，即随着代表点而运动的观察者来看，代表点的局域密度是不随时间变化的



对于平衡态来说 $\partial \rho / \partial t = 0$ ，那么有 $[\rho, H] = 0$ ，系综处于定态

2.1.2 系综理论

相空间代表点的集合和几率密度分布一起规定了一个系综，它描述了在某种宏观约束条件下所有允许微观状态的概率，其约束条件可以由一组外加宏观参量来表示。



1 微正则系综

定义

微正则系综的定义：把 N 个粒子放入体积为 V 的盒子中，并固定总能量 E ，这样的一个独立体系即为微正则系综，其特征函数是熵 $S(N, V, E)$ 。

当 ρ 即不显含时间，也不依赖于坐标

$$\rho(q, p) = C \quad (6)$$

物理上表示，所选的系综在任意时间所有可能的微观状态都是均匀分布，则系综的每个成员是等概率地分布于所有微观态中，这就是微正则系综

$$\langle A \rangle = \frac{1}{\Omega} \int_{\Omega} A(q, p) d\Omega \quad (7)$$

等几率原理

平衡态统计物理的基本假设：对孤立体系的平衡态求统计平均时，认为相空间中能量曲面 $(E, E + \Delta E)$ 之间相等体积的几率相等。

等价于量子力学中不可区分理论

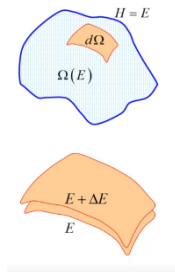
相空间体积

其实就是总代表点点数

$$\Omega(E) = \int_{H \leq E} d\Omega = \int \frac{d\mathbf{q}d\mathbf{p}}{h^{3N}N!} \quad (N \text{ 无量纲}) \quad (8)$$

处于厚度为 ΔE 的曲面壳层内

$$\int_{\Delta E} d\Omega = \Omega'(E) \Delta E \quad (9)$$



概率密度

由等概率原理得到，**概率密度分布**为

$$\rho(q, p) = \begin{cases} \frac{1}{\Omega'(E)\Delta E} & \text{if } E \leq H(q, p) \leq E + \Delta E \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

当 $\Delta E \rightarrow 0$ 密度分布成 $\rho = \frac{\delta(H(q, p) - E)}{Z_{NVE}}$

其中用于归一化的常数 Z 称为配分函数

那么对于系综的平均可以写成

$$\langle A \rangle = \int_{\Delta E} A(q, p) \rho(q, p) dq dp = \frac{1}{\Omega'(E)} \lim_{\Delta E \rightarrow 0} \frac{1}{\Delta E} \int_{\Delta E} A(q, p) d\Omega = \frac{1}{Z_{NVE}} \int \delta(H(q, p) - E) A(q, p) d\Omega \quad (11)$$

配分函数

系综里所有可能微观态的加权和，每个微观态的权重是它在系综中出现的概率，即

$$Z_{NVE} = \int \delta(H(q, p) - E) * 1 * d\Omega = g(E) \quad (12)$$

$$d\Omega = g(E) dE$$

可以把微正则系统的**分配函数** $g(E)$ 理解为：无限薄壳层的相空间面积， E 是球的半径；或理解为总能量恰为 E 的微观状态数。

$$g(E) = \Omega'(E) \quad (13)$$

特征函数

微正则系统的特征函数是熵： $S(N, V, E) = k \ln Z_{NVE}$

证明：

对于两个分布封闭独立的子系统；

$$N = N_1 + N_2, V = V_1 + V_2, E = E_1 + E_2 \quad (14)$$

体系的总微观状态数为(几率相乘)

$$g(N, V, E) = g_1(N_1, V_1, E_1)g_2(N_2, V_2, E_2) \quad (15)$$

平衡时是几率最大的状态，微观状态数也最大

$$\begin{aligned} dg &= g_1 dg_2 + g_2 dg_1 = 0 \\ d \ln g &= d \ln g_1 + d \ln g_2 = 0 \end{aligned} \quad (16)$$

封闭系统的熵： $S(N, V, E) = S_1(N_1, V_1, E_1) + S_2(N_2, V_2, E_2)$

	$dS = dS_1 + dS_2$	\rightarrow	$d \ln g = d \ln g_1 + d \ln g_2$
平衡态时熵极大:	$S = S_{\max}, \quad dS = 0$		$\ln g = \ln g_{\max}, \quad d \ln g = 0$

熵与 $\ln g$ 一致对应:

$$S(N, V, E) = k \ln g = k \ln Z_{NVE} \quad \text{不确定至一常数}$$

根据热力学第一定律 $dE = TdS - PdV + \mu dN$

$$T = \left(\frac{\partial S}{\partial E} \right)_{N, V}^{-1}, \quad P = T \left(\frac{\partial S}{\partial V} \right)_{E, N}, \quad \mu = -T \left(\frac{\partial S}{\partial N} \right)_{E, V}$$

所以当确定了特征函数，那么可以确定出体系中其他的物理量。

例 一维谐振子

例：一维谐振子 $H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^N \left(\frac{p_i^2}{2m} + \frac{1}{2} m \omega^2 q_i^2 \right) = E \quad x = m \omega q$

相空间体积:

$$\Omega = \frac{1}{(m \omega h)^N} \int_{\sum_{i=1}^N (x_i^2 + p_i^2) \leq 2mE} d^N x d^N p = \frac{1}{(m \omega h)^N} \frac{\pi^N}{N! \Gamma(N)} (2mE)^N = \frac{1}{N! \Gamma(N)} \left(\frac{E}{\hbar \omega} \right)^N$$

微观状态数、配分函数： $g = \Omega' = \frac{E^{N-1}}{\Gamma(N)} \left(\frac{1}{\hbar \omega} \right)^N$

特征函数： $S(N, V, E) = k \ln g = k \left[(N-1) \ln E - N \ln \hbar \omega - \ln \Gamma \right]$

$$\approx k \left[(N-1) \ln E - N \ln \hbar \omega - N \ln N + N \right] \approx Nk \left[1 + \ln \left(\frac{E}{N \hbar \omega} \right) \right]$$

$$T = \left(\frac{\partial S}{\partial E} \right)_{N, V}^{-1} = \left(\frac{Nk}{E} \right)^{-1}, \quad E = NkT, \quad \text{与能量均分定理结果一致}$$

$$P = T \left(\frac{\partial S}{\partial V} \right)_{E, N} = 0, \quad \mu = -T \left(\frac{\partial S}{\partial N} \right)_{E, V} = -kT \ln \frac{E}{N \hbar \omega} = -kT \ln \frac{kT}{\hbar \omega}$$

2 正则系综

定义

把 N 个粒子放入体积为 V 的盒子中，并将其置于温度恒为 T 的热浴中，这个体系即为正则系综，它是 Monte Carlo 方法模拟研究的典型体系，这时系综的总能量和压强不定，可能在一个平均值附近涨落。

当 ρ 是 Hamilton 量的显函数时:

$$\begin{aligned} [\rho, H] &= 0 \\ \rho(q, p) &= p[H(q, p)] \end{aligned} \quad (17)$$

该系统是正则系综，其概率密度分布函数是Boltzmann分布

$$\rho(q, p) \propto \exp[-H(q, p)/kT] \quad (18)$$

注：对于速度 $p(v) \propto v^2 \exp[-mv^2/2kT]$

概率密度

正则系综中，原则上体系的总能在零至无穷之间变化，我们的问题是要找到系统在任意时间处于能量为 E 的几率。将体系（总能量 E ）与其浸入的热浴（总能量 E_h ）合起来看成是一个大的力学体系并对这个大体系应用微正则系综，大体系的总能量为 $E_0 = E_h + E$ 并保持恒定

大体系的相体积元为 $d\Omega d\Omega_h$

$$\begin{aligned} \langle A \rangle &= \frac{1}{g(E_0)} \lim_{\Delta E \rightarrow 0} \frac{1}{\Delta E} \int A(\mathbf{q}, \mathbf{p}) d\Omega d\Omega_h \\ &= \frac{1}{g(E_0)} \int A(\mathbf{q}, \mathbf{p}) d\Omega \lim_{\Delta E \rightarrow 0} \frac{1}{\Delta E} \int d\Omega_h \end{aligned}$$

$d\Omega_h$ 的积分限为： $E_0 \leq E_h + E \leq E_0 + \Delta E$

$$\int_{\Delta E} d\Omega_h = \Omega_h(E_h + \Delta E) - \Omega_h(E_h) = \frac{d\Omega_h(E_h)}{dE_h} \Delta E$$

$$\langle A \rangle = \frac{1}{g(E_0)} \int g_h(E_h) A(\mathbf{q}, \mathbf{p}) d\Omega \rightarrow \rho(\mathbf{q}, \mathbf{p}) = \frac{g_h(E_h)}{g(E_0)}$$

假定 $E_h \gg E$ ，那么 $E/E_0 = 1 - E_h/E \approx 0$ ，对 $g_h(E_h)$ 做对数形式的泰勒展开

$$\ln g_h(E_h) = \ln g_h(E_0) + \frac{\partial \ln g_h}{\partial E_h} (E_h - E_0) \quad (19)$$

由于相互接触的两个热力学系统热平衡时，体系温度相等，故有

$$\frac{1}{kT} = \frac{1}{k} \frac{\partial S}{\partial E} = \frac{\partial \ln g}{\partial E} = \frac{\partial g_h}{\partial E_h} = \beta = \text{const} \quad (20)$$

于是可以得到 $g_h(E_h)$ 的表达式

$$\begin{aligned} \ln g_h(E_h) &= \ln g_h(E_0) - \beta E \\ g_h(E_h) &= g_h(E_0) \exp(-\beta E) \end{aligned} \quad (21)$$

所以概率密度 ρ 为：

$$\rho(q, p) = \frac{g_h(E_h)}{g(E_0)} = \frac{g_h(E_0)}{g(E_0)} \exp(-\beta E) = \frac{g_h(E_0)}{g(E_0)} \exp(-\beta H) \quad (22)$$

当对 ρ 归一化后为Boltzmann分布，且对于系综的平均可以写成

$$\begin{aligned} \rho_{NVT} &= \frac{\exp(-\beta H)}{Z_{NVT}} \\ \langle A \rangle &= \frac{\int A(q, p) e^{-\beta H(q, p)} d\Omega}{Z_{NVT}} \end{aligned} \quad (23)$$

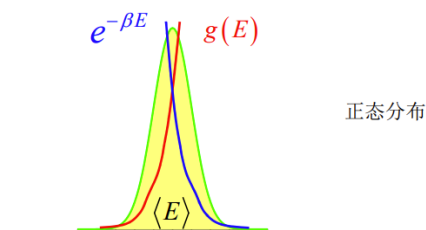
分配函数

正则分配函数为

$$Z_{NVT} = \int \exp[-\beta H(q, p)] d\Omega = \int \exp(-\beta E) g(E) dE = \int \exp(-\beta E) Z_{NVE} dE \quad (24)$$

正则配分函数的积分可表为对所有给定能量的微观状态数的积分

$$Z_{NVT} = \int \exp(-\beta E) d\Omega = \int \exp(-\beta E) g(E) dE = \int \exp(-\beta E) Z_{NVE} dE$$



微正则配分函数是随能量而迅速增加的函数，但Boltzmann分布则是快速递减的函数，故两个函数的乘积是在某一个值 $\langle E \rangle$ 附近有一尖锐的分布。体系在大多数时间内的能量状态都是逗留此值的附近。因此，正则系综与能量严格为 $\langle E \rangle$ 的微正则系综几乎等价，正则系综中的能量涨落不会很大。

热力学极限定理：当粒子数和体积均为无穷大时，两系综的平均相等。

特征函数

正则系综的特征函数是Helmholtz自由能： $F(N, V, T) = -kT \ln Z_{NVT}$

证明： $\rho_\alpha = \frac{1}{Z} \exp(-\beta E_\alpha), \quad Z = \sum_\alpha \exp(-\beta E_\alpha)$

任意物理量 A 的期待值： $\langle A \rangle = \sum_\alpha A_\alpha \rho_\alpha = \frac{1}{Z} \sum_\alpha A_\alpha e^{-\beta E_\alpha}$

内能： $U = \langle E \rangle = \frac{1}{Z} \sum_\alpha E_\alpha e^{-\beta E_\alpha} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}$

比热（单位温度变化时内能的变化量）： $C_v = \frac{\partial U}{\partial T} = -k\beta^2 \frac{\partial U}{\partial \beta} = k\beta^2 \frac{\partial^2 \ln Z}{\partial \beta^2}$

根据热力学： $C_v = T \frac{\partial S}{\partial T} = -\beta \frac{\partial S}{\partial \beta}$

$\frac{\partial S}{\partial \beta} = -k\beta \frac{\partial^2 \ln Z}{\partial \beta^2}$ 两边积分 \rightarrow $S = -k\beta \frac{\partial \ln Z}{\partial \beta} + k \ln Z$

自由能： $F = U - TS = -\frac{\partial \ln Z}{\partial \beta} - T \left(-k\beta \frac{\partial \ln Z}{\partial \beta} + k \ln Z \right)$
 $= -kT \ln Z = -\frac{1}{\beta} \ln Z$

特征函数是Helmholtz自由能： $F(N, V, T) = -kT \ln Z_{NVT}$

特征函数的意义是其它热力学变量均可通过该函数而得到，例如：

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

热力学中： $dU = TdS - PdV + \mu dN$ $dF = d(U - TS) = dU - TdS - SdT$
 $= -SdT - PdV + \mu dN + \dots$

$E \propto -XY, \quad X = -\frac{\partial F}{\partial Y} = -\frac{\partial E}{\partial Y}$

$E \propto -MH, \quad M = -\frac{\partial F}{\partial H}$ Y 是外加“场”，强度量
 X 是对应的共轭力学量，广延量

从上式可以看出，对于能量，总可以表现出 强度量 \times 广延量的形式， Y 是广义的场， X 是广义的磁矩。

根据特征函数的关系，可以得到内能（广延量）的涨落与比热（强度量）成正相关

$$\langle E^2 \rangle = \frac{1}{Z} \sum_{\alpha} E_{\alpha}^2 e^{-\beta E_{\alpha}} = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}$$

$$\langle E^2 \rangle - \langle E \rangle^2 = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} - \left(\frac{\partial \ln Z}{\partial \beta} \right)^2 = \frac{\partial^2 \ln Z}{\partial \beta^2} = \frac{C_v}{k\beta^2} = kT^2 C_v$$

$$\text{方差 (自相关函数): } C(0) = \langle (\delta A(0))^2 \rangle = \langle (\delta A)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 = \text{var}\{A\}$$

$$\text{比热正比于能量涨落: } C_v = k\beta^2 (\langle E^2 \rangle - \langle E \rangle^2) = k\beta^2 \langle (\delta E)^2 \rangle = k\beta^2 \text{var}(E)$$

这是微观统计力学推得的结果，不可能由宏观热力学定律得到。

例：1升气体中约有 3×10^{22} 个气体分子，比热 $\approx 1 \text{ J/K}$ （常温常压）

$$\text{var}(E) \approx 10^{-18} \text{ J}, \quad U \approx 10^2 \text{ J} \quad \frac{\sqrt{\text{var}(E)}}{\langle E \rangle} \approx 10^{-20}$$

涨落、自相关函数、响应函数

$$\text{比热和内能是广延量: } C_v \propto V \propto N, \quad U \propto V \propto N$$

$$\text{能量的涨落: } \frac{\sqrt{\text{var}(E)}}{\langle E \rangle} \propto \frac{\sqrt{V}}{V} \propto \frac{1}{\sqrt{V}} \propto \frac{1}{\sqrt{N}} \rightarrow 0$$

热力学极限：非常大的体系 ($V \rightarrow \infty, N \rightarrow \infty$)

$$\text{如一维谐振子: } E = NkT \quad C_v = Nk \quad \frac{\sqrt{\text{var}(E)}}{\langle E \rangle} = \frac{1}{\sqrt{N}}$$

实际的凝聚态物质：无穷大（热力学极限成立），涨落很小；

MC 模拟中的体系：有限大，结果有涨落，需要模拟尽可能大的体系

$$\text{热力学关系: } E \propto -XY, \quad X = -\frac{\partial F}{\partial Y}$$

$$\begin{aligned} \text{统计力学中: } \langle X \rangle &= -\frac{\partial F}{\partial Y} = \frac{1}{\beta} \frac{\partial \ln Z}{\partial Y} = \frac{1}{\beta Z} \frac{\partial Z}{\partial Y} = \frac{1}{\beta Z} \frac{\partial}{\partial Y} \sum_{\alpha} e^{-\beta E_{\alpha}} \quad (E_{\alpha} \propto -X_{\alpha} Y) \\ &= \frac{1}{Z} \sum_{\alpha} X_{\alpha} e^{-\beta E_{\alpha}} \end{aligned}$$

$$\langle X^2 \rangle = \frac{1}{Z} \sum_{\alpha} X_{\alpha}^2 e^{-\beta E_{\alpha}}$$

$$\text{var}(X) = \langle X^2 \rangle - \langle X \rangle^2 = \frac{1}{Z} \sum_{\alpha} X_{\alpha}^2 e^{-\beta E_{\alpha}} - \frac{1}{\beta^2 Z^2} \left(\frac{\partial Z}{\partial Y} \right)^2 = -\frac{1}{\beta} \frac{\partial^2 F}{\partial Y^2} = \frac{1}{\beta} \frac{\partial}{\partial Y} \langle X \rangle$$

上述方程说明了：

物理量 X 的平均值 $\langle X \rangle$ 可由自由能关于外场 Y 的一阶导数得到

物理量 X 的涨落 $\langle (\delta X)^2 \rangle$ 可由自由能关于外场 Y 的二阶导数得到

定义广义的“磁化率 χ ”：体系物理量 X 对外场的 Y 变化的响应，称作线性响应定理

$$\chi \equiv \frac{\partial \langle X \rangle}{\partial Y} = \beta \text{var}(X) \quad \text{线性响应定理}$$

实验测量 MC计算

外加磁场时: $E \propto -MH$

$$\chi = \frac{\partial M}{\partial H} = \beta \text{var}(M) = \beta (\langle M^2 \rangle - \langle M \rangle^2) = \frac{\langle (\delta M)^2 \rangle}{kT}$$

$$F = U - TS$$

$$C_v = \frac{\partial U}{\partial T} = T \frac{\partial S}{\partial T} = k\beta^2 \text{var}(E) = k\beta^2 (\langle E^2 \rangle - \langle E \rangle^2) = \frac{\langle (\delta E)^2 \rangle}{kT^2}$$

要点: 比热是内量的涨落, 磁化率是磁化强度的涨落

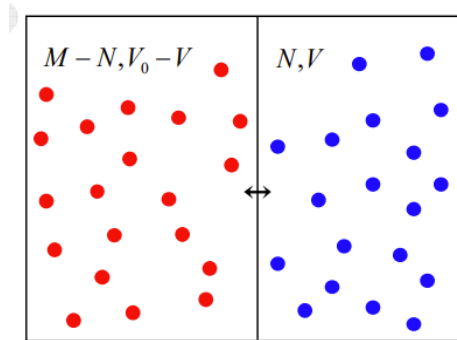
3 等温等压系综

定义

把 N 个粒子放入可移动边壁的盒子中使其压强 P 为固定值, 并将其置于温度恒为 T 的热浴中, 这个体系即为等温等压系综, 一般是在 Monte Carlo 方法模拟中加以实现, 这时系综的总能量和体积不定, 可以在一个平均值附近涨落。对于磁性系统, P 和 V 分别用磁矩和外磁场替代。

概率密度

设想一个理想气体体系有 $M - N$ 个粒子, 体积为 $V_0 - V$, 我们的力学体系通过一活动边壁与理想气体保持接触, 总体系的粒子数 M 和体积 V_0 恒定。



边壁自由活动时, 力学体系的体积 V 会有涨落。总体系的配分函数为两个体系的配分函数之积,

$$Z_{MV_0T} = \int \exp(-\beta H) d\Omega \int \exp(-\beta H_i) d\Omega_i \quad (25)$$

$$\int \exp(-\beta H_i) d\Omega_i = \left(\int dV \int \exp(-\beta p^2/2m) d\mathbf{p} \right)^{M-N} = \left[(V_0 - V)(2\pi mkT)^{3/2} \right]^{M-N} \quad (2.1.2.3-2)$$

由于体积 V 可变, 则求系综平均值时应对体积积分, 力学体系的几率密度为

$$\rho = \frac{(V_0 - V)^{M-N} \exp(-\beta H)}{\int (V_0 - V)^{M-N} dV \int \exp(-\beta H) d\Omega} \quad (2.1.2.3-3)$$

现在取极限 $V_0 \rightarrow \infty, M \rightarrow \infty$, 使得理想气体的密度为 $\rho_0 \rightarrow (M - N)/V_0$ 。故在极限 $V/V_0 \rightarrow 0$ 下, 有

$$\begin{aligned} (V_0 - V)^{M-N} &\rightarrow V_0^{M-N} \left(1 - \frac{V}{V_0} \right)^{M-N} \rightarrow V_0^{M-N} \exp \left[-(M - N) \left(\frac{V}{V_0} \right) \right], \\ &\rightarrow V_0^{M-N} \exp(-\beta PV) \end{aligned} \quad (2.1.2.3-4)$$

式中已代入理想气体状态方程 $\rho_0 = \beta P$, 这里 P 是常数。故 (2.1.2.3-3) 式为

$$\rho = \frac{\exp[-\beta(H + PV)]}{\int \exp(-\beta PV) dV \int \exp(-\beta H) d\Omega} \quad (2.1.2.3-5)$$

配分函数

对上式归一化

$$Z_{NPT} = \iint \exp\{-\beta[H(\mathbf{q}, \mathbf{p}) + PV]\} d\Omega dV = \int dV \exp(-\beta PV) Z_{NVT}, \quad (2.1.2.3-6)$$

系综平均为

$$\langle A \rangle = Z_{NPT}^{-1} \iint A(\mathbf{q}, \mathbf{p}) e^{-\beta[H(\mathbf{q}, \mathbf{p}) + PV]} d\Omega dV. \quad (2.1.2.3-7)$$

特征函数

体系的特征函数是总体系的自由能与理想气体的自由能之差, 即 Gibbs 自由能 $G(N, P, T)$,

$$G(N, P, T) = -kT \ln Z_{NPT}, \quad (2.1.2.3-8)$$

它与其它热力学量的关系为

$$G = F + PV = E - TS + PV, \quad (2.1.2.3-9)$$

由此可得,

$$S = -\left(\frac{\partial G}{\partial T} \right)_{N,P}, \quad V = \left(\frac{\partial G}{\partial P} \right)_{N,T}, \quad \mu = \left(\frac{\partial G}{\partial N} \right)_{P,T}. \quad (2.1.2.3-10)$$

4 巨正则系综

定义

巨正则系综中具有确定温度 T 、给定化学势 μ 及体积 V , 但是粒子数可以是变化的, 如在化学反应中那样。巨正则系综一般是通过 Monte Carlo 方法模拟加以实现, 这时系综的总能量、压强和粒子数存在涨落。

概率密度

和推导正则系综一样, 考虑浸入在一个热浴中的力学体系。力学体系和热浴之间不仅有能量还有粒子交换, 由力学体系和热浴组成的大体系的总能量 $E_0 = E_h + E$ 以及总粒子数 $N_0 = N_h + N$ 保持恒定, 并且力学体系的能量和粒子数远远小于热浴的。微观状态数 Ω 不仅是依赖于能量, 也是粒子数的函数。

$$\text{正则} \quad \ln g_h(E_h) = \ln g_h(E_0) + \left(\frac{\partial \ln g_h}{\partial E_h} \right)_{E_h=E_0} (E_h - E_0) + \dots$$



$$\text{巨正则} \quad \ln g_h(E_h, N_h) = \ln g_h(E_0, N_0) + \left(\frac{\partial \ln g_h}{\partial E_h} \right)_{E_h=E_0} (-E) + \left(\frac{\partial \ln g_h}{\partial N_h} \right)_{N_h=N_0} (-N) + \dots$$

两个具有粒子交换的体系达到相平衡时化学势相等：

分配函数

$$\frac{\partial \ln g_h}{\partial N_h} = \frac{\partial \ln g}{\partial N} = \beta \mu$$

$$g_h(E_h, N_h) \approx g_h(E_0, N_0) \exp[-\beta(\mu N + E)]$$

$$\rho(\mathbf{q}, \mathbf{p}) = \frac{g_h(E_h, N_h)}{g(E_0, N_0)} = \frac{g_h(E_0, N_0)}{g(E_0, N_0)} \exp[-\beta(\mu N + E)] \propto \exp[-\beta(\mu N + E)]$$

巨正则配分函数是，

$$Z_{\mu VT} = \sum_N \frac{1}{N!} \int \exp\{-\beta[\mu N + H(\mathbf{q}, \mathbf{p})]\} d\Omega, \quad (2.1.2.4-6)$$

式中已经将全同粒子的归一化因子 $N!$ 放入。如果有多种粒子的话，则上两式应改为求和式

$$\rho_{\mu VT} = Z_{\mu VT}^{-1} \left(\prod_i N_i! \right)^{-1} \exp\left[-\beta\left(\sum_i \mu_i N_i + H\right)\right], \quad (2.1.2.4-7)$$

$$Z_{\mu VT} = \sum_{N_i} \left(\prod_i \frac{e^{-\beta \mu_i N_i}}{N_i!} \right) \int \exp[-\beta H(\mathbf{q}, \mathbf{p})] d\Omega, \quad (2.1.2.4-8)$$

系综平均为，

$$\langle A \rangle = Z_{\mu VT}^{-1} \sum_{N_i} \left(\prod_i \frac{e^{-\beta \mu_i N_i}}{N_i!} \right) \int A(\mathbf{q}, \mathbf{p}) e^{-\beta H(\mathbf{q}, \mathbf{p})} d\Omega. \quad (2.1.2.4-9)$$

特征函数

$$J(\mu, V, T) = -kT \ln Z_{\mu VT} \quad (2.1.2.4-10)$$

它与其它热力学量的关系为，

$$J = F - \mu N = E - TS - \mu N, \quad (2.1.2.4-11)$$

代入式

$$E = TS - PV + \mu N, \quad (2.1.2.4-12)$$

故有，

$$J = -PV, \quad (2.1.2.4-13)$$

$$S = -\left(\frac{\partial J}{\partial T} \right)_{V, \mu}, \quad P = -\left(\frac{\partial J}{\partial V} \right)_{\mu, T}, \quad N = -\left(\frac{\partial J}{\partial \mu} \right)_{V, T}. \quad (2.1.2.4-14)$$

2.2 Monte Carlo 模拟与重要抽样

物理量的平均

$$\langle A \rangle = \frac{\int A(q, p) \rho(q, p, t) dq dp}{\int \rho(q, p, t)} \quad (26)$$

在NVT系统计算平均值时，以Boltzmann因子 $\exp(-\beta H)$ 作为权重。

归一化因子 Z

$$Z_{NVT} = \int \exp[-\beta H(q, p)] dq dp \quad (27)$$

上式归一化因子难以解析计算。

简单抽样

考虑在相空间中一些完全随机的序列，以Boltzmann因子作为接受几率。对于随机构造的状态，能量高的构型占绝大多数，其接受几率小，因此计算效率低

重要抽样

根据Boltzmann分布产生不具有统计独立性的构型，其偏向于能量低的最可几随机构型，此称作Metropolis重要抽样方法，随机构型则是通过一种称为Markov链的方式构造出来的，其中新的构型仅取决于之前的构型。抽样不依赖于归一化因子。

2.2.1 随机过程

1 随机序列

$\{x(t)\}$ 是时间变量 t 的函数集合，对 x 的取值有几率分布，成为一个随机过程。

时间变量离散化 (t_1, t_2, \dots, t_n) 时，随机过程成为随机序列。

可以理解为有大量的的粒子 N ，在 t_i 时刻某个粒子出现在 x_j 有概率密度 $p(x = x_j; t = t_i)$

这样的时间函数的系综称为一个随机过程，系综中的一个某一具体的函数值即是随机过程的实现

当时间演化时，对应于离散化的时间序列值，拓展到高阶

$$p_n(x_1, \dots, x_n; t_1, \dots, t_n) \quad (28)$$

2 条件概率

$$p(x_n | x_{n-1}, \dots, x_1; t_1, \dots, t_n) = \frac{p_n(x_1, \dots, x_n; t_1, \dots, t_n)}{p_{n-1}(x_1, \dots, x_{n-1}; t_1, \dots, t_{n-1})} \quad (29)$$

它表示，在 $t = t_n$ 时刻 x 取值 x_n ，但在之前的时间内 x 取值分别为 x_{n-1}, \dots, x_1 的几率

3 Markov链

我们称一个平稳的随机序列满足： $p_n(x_1, \dots, x_n; t_1, \dots, t_n) = p_n(x_1, \dots, x_n; t_1 + t, \dots, t_n + t)$ ，这时时间的起点是无关紧要的。时间序列仅仅表示 x 取值的先后。下面我们考虑平稳过程，并略去时间序列值。

若条件几率密度独立于上一步之前的所有 x 值：

$$p_n(x_n | x_{n-1}, \dots, x_1) = p(x_n | x_{n-1}) \quad (30)$$

某一步的结果仅仅依赖于上一步，与更前面的历史无关。对应的态序列 (x_1, \dots, x_n) 称为 Markov 链

转移几率

x 取离散值时 $x_n = x_j$, $x_{n-1} = x_i$ ，其中 x_n 是变量， x_j 是取值

转移几率定义为，将条件几率解释成从状态 x_i 转移到状态 x_j 的跃迁几率，比如 p_1 为走的第一步概率

$$W_{i \rightarrow j} = p_1(x_j | x_i) \quad (31)$$

满足条件

$$\begin{aligned} W_{i,j} &\geq 0 \\ \sum_j W_{i,j} &= 1 (\text{行之和为1}) \end{aligned} \quad (32)$$

对于大量的步骤

$$p(N+1) = Wp(N) = \dots = W^N p(1) \quad (33)$$

最终时，平衡态的分布与初态 $p(1)$ 无关，仅仅取决于转移几率，系统对历史发展轨迹没有保留记忆。这也保证了最终的抽样结果仅仅取决于最终的分布，而与初始状态无关

达到平衡时,符合“本征方程”

$$\mathbf{p} = \mathbf{p}\mathbf{W}, \quad p_i = \sum_j p_j W_{ji} \quad (p \text{是行向量}) \quad (34)$$

现实中，往往在平衡时， p 是给定已知的，而 W 是待求解的。

我们的任务即是寻找到合适的转移概率 W ，使之能给出平稳的分布 p

当给定平衡态分布后，矩阵 W 待确定的参数有 M^2 个，本征方程数有 M 个，归一化方程有 M 个。往往 $M^2 > 2M$ ，所以并不能唯一确定转移矩阵， M 有多种选择。并且实际上本征方程与归一化方程也并不独立

4 主方程 Master equation

对于各态历经的链来说，当 t 很大时 $p(x,t)$ 与时间无关。在 Δt 时间内， $p(x,t)$ 的变化是由于：

- x 的构型在 Δt 后变到 x' 构型，转出
- x' 的构型在 Δt 后进入到 x 构型，转入

因此有

$$p(x, t + \Delta t) - p(x, t) = - \int dx' W(x \rightarrow x') p(x, t) \Delta t + \int dx' W(x' \rightarrow x) p(x', t) \Delta t \quad (35)$$

取无限小 Δt , 得到主方程表达式

$$\frac{\partial p(x, t)}{\partial t} = - \int dx' W(x \rightarrow x') p(x, t) + \int dx' W(x' \rightarrow x) p(x', t) \quad (36)$$

离散化取值后为：

$$\frac{\partial p_i}{\partial t} = \sum_j (p_i W_{ij} - p_j W_{ji}) \quad (37)$$

对主方程 dx 积分

$$\begin{aligned} \frac{\partial}{\partial t} \int p(x, t) dx &= - \int dx \int dx' W(x \rightarrow x') p(x, t) + \int dx \int dx' W(x' \rightarrow x) p(x', t) = 0 \\ &= \frac{\partial(1)}{\partial t} = 0 \end{aligned} \quad (38)$$

说明该式与几率守恒方程 $\int p(x, t) dx = 1$ 等价。

细致平衡条件

我们要求的是平稳分布（即平衡态分布），在足够长的时间下， p 趋于稳定达到平衡，与 t 无关，那么则有

$$\begin{aligned} \int dx' W(x \rightarrow x') p(x, t) &= \int dx' W(x' \rightarrow x) p(x', t) \\ \Rightarrow \frac{p(x)}{p(x')} &= \frac{W(x' \rightarrow x)}{W(x \rightarrow x')} \end{aligned} \quad (39)$$

该式即为统计力学中的细致平衡解。

离散化表达后为

$$p_i W_{ij} = p_j W_{ji} \quad (40)$$

当对下标 j 求和，则得到 $p_i = \sum_j p_j W_{ji}$ ，即本征方程，这说明了细致平衡条件下主方程与本征方程是等价的，没有增加更多的约束条件。

同时为了得到上式，利用了转移概率归一化的条件 $\sum_j W_{ij} = 1$

在某种程度上说明了 本征方程与概率归一化条件并不独立。

但只要 W 满足了细致平衡条件，则可以说明，最终稳定是达到 p 的分布

2.2.2 Metropolis 重要抽样方法

由于转移概率不能唯一确定，但可以设计各种满足细致平衡条件的方法，来增加独立方程的个数

设从 i 到 j 的几率分解： $M_{ij} = T_{ij}A_{ij}$

建议分布： T_{ij} 是由 x_i 选择步进到 x_j 的几率。一般选择 $T_{ij} = T_{ji}$ 对称矩阵

接受分布： A_{ij} 是接受该步的几率

1 Barker 抽样规则

对称T，非对称A，根据待满足的几率分布 p 的形状而定

$$W_{ij} = T_{ij} \frac{p_i}{p_j + p_i}, A_{ij} = \frac{1}{1 + p_i/p_j}, i \neq j$$
$$W_{ii} = 1 - \sum_{j \neq i} W_{ij} \quad (41)$$

考虑到归一化等条件，最终的结果为

$$p = (p_1 \quad p_2 \quad \cdots \quad p_M), \quad W = \begin{pmatrix} p_1 & p_2 & \cdots & \cdots & \cdots & p_M \\ p_1 & p_2 & \cdots & \cdots & \cdots & p_M \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ p_1 & p_2 & \cdots & \cdots & \cdots & p_M \end{pmatrix} \quad \text{Barker抽样规则}$$

2 Metropolis (Rosenbluth) 抽样规则

对称T，非对称A，根据待满足的几率分布 p 的形状而定

$$W_{i,j} = T_{ij} \times \begin{cases} 1, & \text{if } p_j > p_i \\ p_j/p_i, & \text{otherwise} \end{cases}$$
$$A_{ij} = \min\{1, p_j/p_i\}$$
$$W_{ii} = 1 - \sum_{j \neq i} W_{ij} \quad (42)$$

而建议分布 T 一般是任意对称的条件概率，比如正态分布 $T(x \rightarrow x') = T_{ij} = N(x_i - x_j | \mu = 0, \sigma)$ ，代表着建议往附近的点转移的概率大，往远处的概率小。

解释上式为：设 $p(x)$ 为所考虑的几率密度分布，并且已经产生了 x_1, x_2, \dots, x_n 个抽样点，现在的问题是如 何产生下一个抽样点 x_{n+1} 。可以在上一个点附近构造一个试探解， $x_t = x_n + \delta$ ， δ 是试探步长（可正可负，例如可取 $\delta = (\xi - 0.5)\Delta x$ ， Δx 是固定步长， $\xi \in (0,1)$ 是 均匀分布的随机数），该点是否被选取决定于比值 $r = p(x_t)/p(x_n)$ ：

- 如果 $r > 1$ 则接受，即 $x_{n+1} = x_t$
- $r < 1$ ，产生 $[0, 1]$ 区间内均匀分布的随机数 ξ ，如果 $\xi < r$ 则选取；也即使得接受选取几率为 r 。否则，舍去 $x_{n+1} = x_n$ 保持在原点不动。

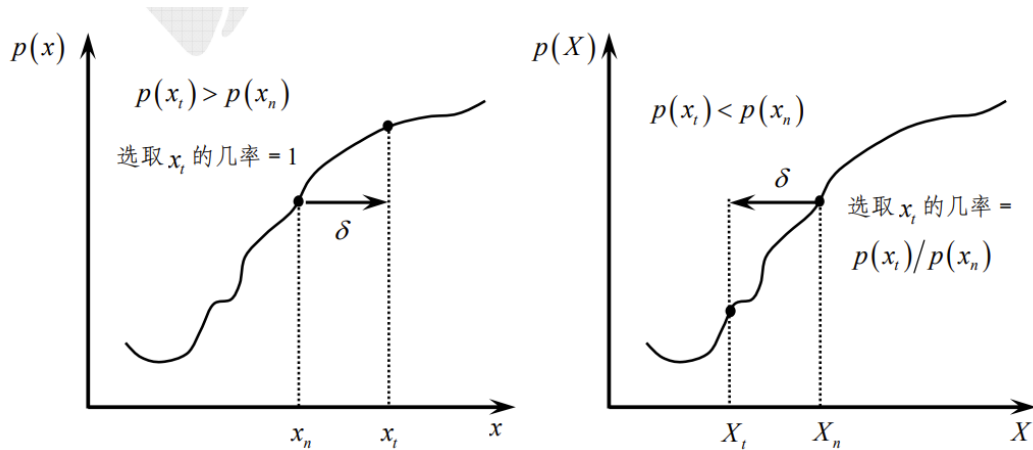


图 2.2.2.1-1 Metropolis 抽样方法的过程。

该方法是一种**重要抽样法**，即抽样得到的 x 的值倾向于落在分布 $p(x)$ 取较大值的区域。由它产生的抽样点序列 x_{n+1}, x_{n+2}, \dots 是强关联的。

特别在序列起始点 x_1 附近，因为很难选择一个合适的起始点。因此，如果不去特别选择的话，通常是任意给定一个起始点产生，然后舍去序列的初始段。同样可以将序列中任一部分的抽样点舍去以减小抽样的强关联性，这个步骤称为分布的热化处理。

除此之外，试探步长 δ 或 Δx 的选取对抽样效率和结果分布也很重要，它不能取得过大或过小。设想 x_n 最可能取的值是在 $p(x)$ 最大处， δ 过大时许多试探被舍去，过小时序列集中在此点的附近，不能覆盖 x 取值的区间，其结果不能很好的代表分布 $p(x)$ 。通常 δ 的选取是使得接受的效率为一半左右。

该抽样过程产生的离散 x 值可以作为满足特定分布 $p(x)$ 的随机数，当它与前面所述的随机数不同之处在于，现在的随机数序列是强烈相关的，因为某一个点总是在上个点的附近产生的，两个顺序点之间相隔很近。尽管有着强的相关性，但是细致平衡原理保证了，只要抽样点数足够多，就能得到平衡分布。

例如

统计力学中 Boltzmann 因子为正则系综的几率分布 $p(x) = \exp[-\beta H(x)]$ ，其中能量是构型 x 的函数，构型变量 x 视具体问题而定。Metropolis 抽样规则中只要求知道 p 的比值，因此分母中的配分函数可以消去（如果我们已知了配分函数，则根本就无需作抽样计算了），则我们可以根据该抽样规则产生大量的离散 x 值，构成 Markov 链 $x_k, k = 1, \dots, m, \dots, n$ ，其中序列中热化过程的前 m 个构型被舍去。而剩下的点列在某种程度上可以看成是“电子云”

注意在计算系统统计平均时，需要将所有有效步数统计在内（不包括热化阶段）而不能只保留选取成功的步数和构型，如第 i 个构型计算出物理量的值为 A_i ，则系综平均为

$$\langle A \rangle = \frac{1}{n - m} \sum_{i=m+1}^n A_i \quad (43)$$

3 Metropolis-Hasting 抽样规则

更一般地，建议分布和接受几率都是非对称的，接受几率根据待满足的几率分布 p 形状而定

取 $A_{ij} = \min\{1, \frac{p_j T_{ji}}{p_i T_{ij}}\}$ ，根据细致平衡条件得

$$W_{ij} = \begin{cases} T_{ij}, & \text{if } p_i T_{ji} > p_j T_{ij} \\ \frac{p_j T_{ji}}{p_i T_{ij}}, & \text{otherwise} \end{cases} \quad (44)$$

$$W_{ii} = 1 - \sum_{j \neq i} W_{ij}$$

根据建议分布 T 进行初始抽样， $T(x \rightarrow x')$ 一般是任意的条件概率分布，最好与 p 有接近的形状。比如 $H = H_0 + H(t), T = \exp(-\beta H_0)$

4 Metropolis思想

Markov链可看作相空间中有偏压的随机行走所形成的一串随机序列，尽管链条之间两部不独立，有很强的相关性，但细致平衡条件原理保证了足够多的抽样后，总能够达到平衡分布。

我们不关心点之间是否独立，以及分布是如何形成的，只要抽样出来的点与几率分布曲线形态一致即可。

2.2.3 各种系综的 Monte Carlo 的抽样方法

2.3 正则系综的统计力学模型

在真实的物理系统中，涉及的物理问题往往是多粒子之间的相互作用的，而且体系的宏观物理性质不仅与微观的相互作用有关，也依赖于环境如温度等变量。特别是，许多物理系统中有一种共性，就是由短程相互作用引起系统的长程有序。

例如，分子间的作用力是短程的，但无数个分子聚集成物质材料却可以具有一种集体的效应，如铁磁性。磁性的来源本质上是由于电子的自旋和轨道运动，因此一些原子和分子甚至有机分子都具有弱的磁性，但是原子间相互作用的区域仅为一个纳米左右，因此在铁磁材料中，这些原子必须以集体协调的方式配合行动才能形成宏观的磁性。

这种相互作用的多粒子体系可以产生一种重要的物理现象，即相变，相变问题在物理学的众多领域的研究中一直扮演着重要角色

2.3.1 Ising 模型

1 自旋与磁性

物质在外磁场H中的磁化强度M（单位体积中的总磁矩）为

$$\mathbf{M} = \chi \mathbf{H} \quad (45)$$

Ising 模型中，每一近邻自旋对之间有相互作用，系统的能量(Hamilton 量)E为

$$E = - \sum_{\langle i,j \rangle=1}^N J_{ij} \sigma_i \sigma_j - \mu_B H \sum_{i=1}^N \sigma_i \quad (46)$$

其中下标 $\langle i, j \rangle$ 表示近邻自旋对， J 是交换积分常数，度量了自旋—自旋相互作用的强弱， μ_B 是 Bohr 磁矩， H 是磁场强度

如果一对自旋方向相同，则能量为 $-J$ ，相反为 J 。因此定性来看：

当 $J > 0$ ，体系更趋向于把所有自旋取向排成一致方向以使得能量最低，**在没有磁场情况下产生了自发磁化，这是铁磁性。**

当 $J < 0$ ，自旋对的取向相反才可能使能量最低，**宏观不表现磁性，但当加上磁场后逼迫自旋取向相同，产生磁化，这是反铁磁性。**

当温度升高时，热激发效应使得某些自旋取向随机反转，逐步使系统无序化，在足够高的温度下自发磁化消失成为顺磁性。

2 统计力学分布

假设 $H=0$ ，那么哈密顿量E为

$$E = -J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1} \quad (47)$$

令 $K = J/k_B T$ ，则系统的配分函数为

$$Z = \sum_a \exp(-E_a/k_B T) = 2(2 \cosh K)^{N-1} \quad (48)$$

3 一维模拟：热平衡

4 二维 Onsager 解

纯数学解析解，缺乏物理图像理解

5 Weiss平均场理论

物理图像更加直观，类比 van der Waals 真实气体状态方程

$$(P + a \frac{N^2}{V^2})(V - Nb) = NkT \quad (49)$$

假设自旋受到周围相邻磁矩产生的内磁场 H_{in} 与外磁场 H 影响，定义参量 H_{in}

$$H_{in} = a \langle \sigma \rangle \quad (50)$$

则系统的哈密顿量重写作

$$E = -\frac{1}{2} \sum_i zJ \langle \sigma \rangle \sigma_i - \quad (51)$$

6 二维 Monte Carlo 模拟

2.3.2 自旋自相关函数