

Mathematical and numerical methods in Physics (II)

计算物理

蔡子

上海交通大学

2025年春季



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



计算物理与数值方法: 蒙特卡洛方法

蒙特卡罗方法于20世纪40年代美国在第二次世界大战中研制原子弹的“曼哈顿计划”计划的成员S.M.乌拉姆和J.冯·诺伊曼首先提出。数学家冯·诺伊曼用驰名世界的赌城—摩纳哥的Monte Carlo—来命名这种方法，为它蒙上了一层神秘色彩。

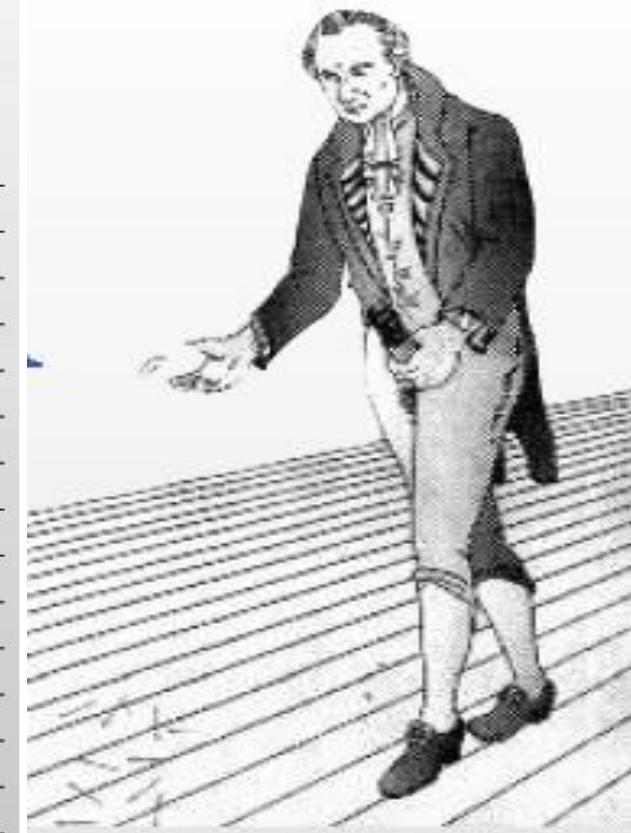
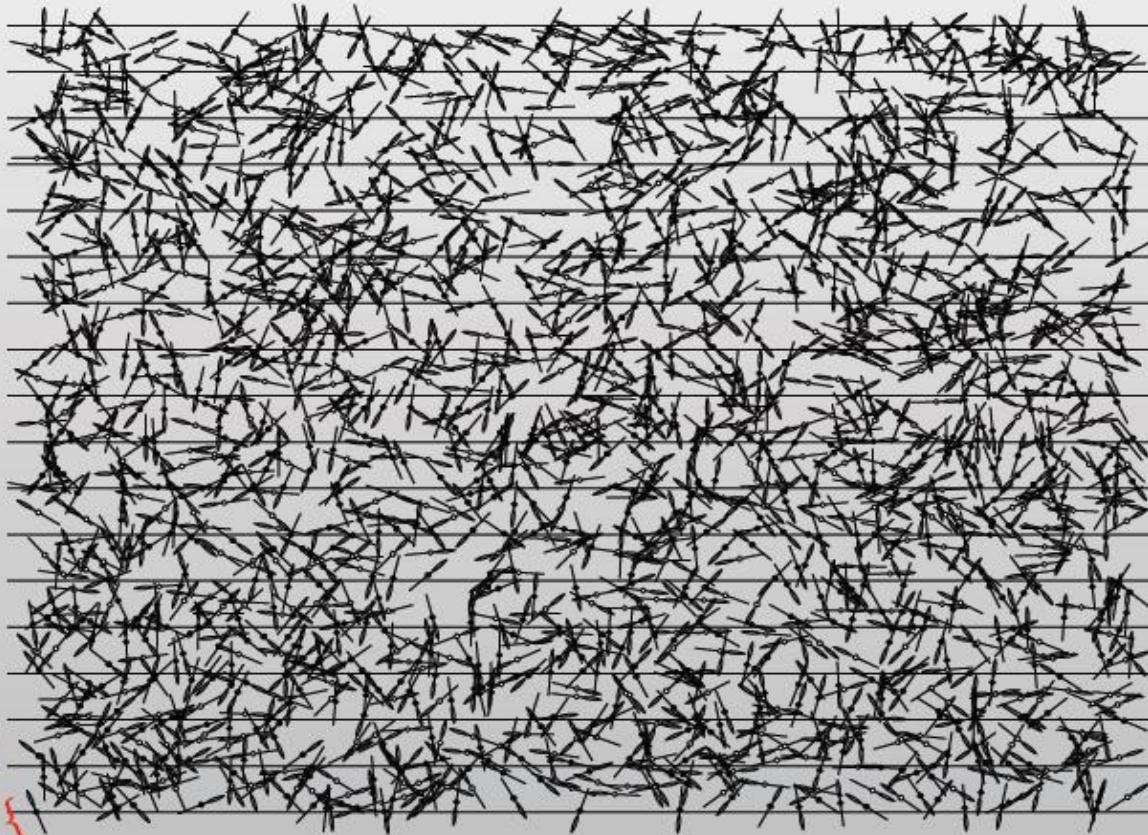




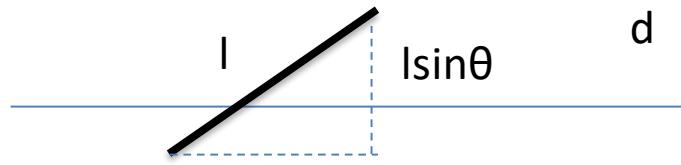
计算物理与数值方法: 蒙特卡洛方法

通常蒙特·卡罗方法通过构造符合一定规则的随机数来解决数学上的各种问题。对于那些由于计算过于复杂而难以得到解析解或者根本没有解析解的问题，蒙特·卡罗方法是一种有效的求出数值解的方法。

$$\langle N_{\text{hits}} \rangle = \frac{2}{\pi} \frac{\ell}{d}$$



Buffon 1777



$$\frac{1}{N} \sum_i \frac{l \sin \theta_i}{d} = P$$

$$P = \frac{1}{N} \sum_{i=1}^N \frac{l \sin \theta_i}{d} = \frac{l}{d\pi} \int_0^\pi d\theta \sin \theta = \frac{2l}{\pi d}$$

一般来说，蒙特卡洛模拟分为两种：

1. **直接蒙特卡洛模拟：**直接通过随机数来模拟自然界中的复杂过程。例如，在模拟一个城市的交通网络系统的过程中，单个车辆的行为可以通过随机过程来模拟。这类问题的核心在于如何用随机数来代表特定现象或物理过程。前面讲的投针法、布朗运动均属于这一类。
2. **Metropolis 蒙特卡洛模拟：**通过 Metropolis 算法，产生由一系列满足特定分布的随机数或者随机机构型组成的马尔可夫链。这类方法适用与求解经典或量子多体系统的平衡态性质。



赝随机数

赝随机数的产生

用数学的方法在计算机上产生真实的随机是不可能的，因为一个程序在运行中算得的结果完全是确定的，不具有任何随机性。但另一方面确实可以用确定的程序产生出可以重复的，满足随机数的各个检验要求的数列。为了与真实的随机数区别，我们通常把数学方法产生的随机数称为赝随机数。

LCG(linear congruential generator)代表了最好的伪随机数产生器算法。主要原因是容易理解，容易实现，而且速度快

$$X(n+1) = (a * X(n) + c) \% m$$

Source	m	a	c	种子位
Numerical Recipes	2^{32}	1664525	1013904223	

其中，a为"乘数"，b为"增量"，m为"模数",x0为"种子数"。

如果产生的是区间实在(0,1)之间的，则只需要每个数都除以m即可，即取

$$\zeta_n = x_n / m$$



中心极限定理

变量 x_i 的出现的概率为 $f(x_i)$, 则 n 个独立的随机变量 $\{x_i\}$ 的数学期望 E_x 、方差 V_x 和标准差 σ_x 分别为:

$$E_x = \sum_{i=1}^n f(x_i)x_i, \quad V_x = \sum_{i=1}^n (x_i - E_x)^2 f(x_i), \quad \sigma_x = \sqrt{V_x}, \quad (2.4)$$

中心极限定理: 大量独立的随机变量之和, 其分布会趋近于正态分布。具体来说, 设随机变量 X_1, X_2, \dots, X_n 是 n 个独立的随机变量, 它们均值为 μ , 方差为 σ^2 , 当 n 趋于无穷大时, 随机变量 $Y = (X_1 + X_2 + \dots + X_n)/n$ 会趋近于正态分布:

$$P(Y) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} e^{-\frac{(Y-\mu)^2}{2\tilde{\sigma}^2}}, \quad \tilde{\sigma} = \frac{\sigma}{\sqrt{n}}, \quad (2.5)$$



蒙特卡洛方法应用：数值积分

数值积分： $S = \int_0^1 f(x) dx,$

we can simply divide the region $[0, 1]$ evenly into M slices with $x_0 = 0$ and $x_M = 1$,

$$S = \frac{1}{M} \sum_{n=1}^M f(x_n) + O(h^2),$$

We can, on the other hand, select x_n with $n = 1, 2, \dots, M$ from a uniform random number generator in the region $[0, 1]$ to accomplish the same goal. If M is very large, we would expect x_n to be a set of numbers uniformly distributed in the region $[0, 1]$ with fluctuations proportional to $1/\sqrt{M}$. Then the integral can be approximated by the average:

$$S \simeq \frac{1}{M} \sum_{n=1}^M f(x_n),$$

Note that the possible error in the evaluation of the integral is now given by the fluctuation of the distribution x_n . If we use the standard deviation of statistics to estimate the possible error of the random sampling, we have:

$$(\Delta S)^2 = \frac{1}{M} (\langle f_n^2 \rangle - \langle f_n \rangle^2).$$



蒙特卡洛方法应用：数值积分

The simple Monte Carlo quadrature used in the above program does not show any advantage. For example, the trapezoid rule yields much higher accuracy with the same number of points. The reason is that the error from the Monte Carlo quadrature is

$$\Delta S \propto \frac{1}{M^{1/2}},$$

whereas the trapezoid rule yields an estimated error of

$$\Delta S \propto \frac{1}{M^2},$$

The true advantage of the Monte Carlo method comes in the evaluation of multidimensional integrals. For a d -dimensional space, the Monte Carlo quadrature will still yield the same error behavior, that is, it is proportional to $1/\sqrt{M}$, where M is the number of points sampled. However, the trapezoid rule has an error of

$$\Delta S \propto \frac{1}{M^{2/d}},$$

There are some specially designed numerical quadratures that would work still better than the Monte Carlo quadrature when d is slightly larger than 4. However, for a real many-body system, the dimensionality in an integral is $3N$, where N is the number of particles in the system. When N is on the order of 10 or larger, any other workable quadrature would perform worse than the Monte Carlo quadrature.



蒙特卡洛方法应用：数值积分

如果我们在一个多维体积中均匀随机地选N个点： $x_1; x_2; x_3; \dots; x_N$ ，则

$$\sigma^2 = \left\langle \left(\frac{b-a}{N} \sum_{i=1}^N f_i \right)^2 \right\rangle - \left\langle \left(\frac{b-a}{N} \sum_{i=1}^N f_i \right) \right\rangle^2 = \frac{(b-a)^2}{N} (\bar{f}^2 - \bar{f}^2)$$

为了提高的精度，一种方法就是增大N，为了增加一位有效数字就要增加100倍的计算量

另一种提高精度的方法是减小方差，为此考虑不在积分域内均匀取点，而是按照某一分布取点，用 $g(x)$ 记这一分布：

$$I = \int_a^b dx f(x) = \int_a^b [g(x)dx] \left[\frac{f(x)}{g(x)} \right] \simeq \frac{1}{N} \sum_i \left[\frac{f(x'_i)}{g(x'_i)} \right]$$

定义新函数： $\tilde{f}(x) = f(x)/g(x)$

$$\sigma^2 = \frac{1}{N} \left[\sum_{i=1}^N \tilde{f}(x_i)^2 - \left(\sum_{i=1}^N \tilde{f}(x_i) \right)^2 \right]$$

可以构造一些简单而又满足上述要求的分布进行取样 $g(x) \sim f(x)$

满足特定概率的随机抽样（“重要”的地方抽样密度提高，“不重要”的地方抽样密度减少）得到的结果，其方差相比均匀随机采样得到的方差大大减小，这正是重要性抽样的基本思想



蒙特卡洛方法应用：数值积分

例子: $I = \int_0^{10} dx x^{68} \exp(-25x)$ $I^{\text{exact}} = 0.86416626117$

均匀随机抽样:

N	I	$\frac{ I - I^{\text{exact}} }{I^{\text{exact}}}$	$\sqrt{N} \frac{ I - I^{\text{exact}} }{I^{\text{exact}}}$	$\frac{\Delta I}{I}$	$\frac{\sqrt{N} \Delta I}{I}$
10^2	1.0757	0.245	2.44	0.235	2.35
10^3	0.83834	0.299×10^{-1}	0.945	0.874×10^{-1}	2.76
10^4	0.84711	0.197×10^{-1}	1.97	0.277×10^{-1}	2.77
10^5	0.86775	0.414×10^{-2}	1.31	0.867×10^{-2}	2.74
10^6	0.86343	0.850×10^{-3}	0.850	0.274×10^{-2}	2.75
10^7	0.86423	0.686×10^{-4}	0.217	0.868×10^{-3}	2.75
10^8	0.86380	0.420×10^{-3}	4.20	0.275×10^{-3}	2.75
10^9	0.86426	0.111×10^{-3}	3.50	0.868×10^{-4}	2.75
10^{10}	0.86420	0.383×10^{-4}	3.83	0.275×10^{-4}	2.75



蒙特卡洛方法应用：数值积分

为了减少方差，我们尝试用一个与被积函数比较接近的分布来取样，取此分布为

$$g(x) = 0.8920620583 \exp(-2.5(x - 2.8)^2)$$

然后，利用这个分布取样，并计算 $\frac{x^{68} \exp(-25x)}{0.8920620583 \exp(-2.5(x - 2.8)^2)}$

的平均值和误差。

为了实现对于 $g(x)$ 的抽样，我们利用如下Metropolis算法，关于这一算法的解释将在下一章给出。算法如下：

1. 任取 $x_0 \in [0, 10]$ 。
2. 设已经得到 x_i , 选 $y = x_i + \delta(ran() - 0.5)$. 其中 $ran()$ 给出 $(0, 1)$ 之间均匀分布的随机数， δ 为一可调参数。
3. 如果 $y < 0$ 或 $y > 10$, 则令 $x_{i+1} = x_i$. 否则，计算 $g(y)$, 产生随机数 $\xi = ran()$, 如果 $g(y)/g(x_i) \geq \xi$, 则 $x_{i+1} = y$, 否则 $x_{i+1} = x_i$.



蒙特卡洛方法应用：数值积分

$$I^{\text{exact}} = 0.86416626117$$

I	N	I	$\frac{ I - I^{\text{exact}} }{I^{\text{exact}}}$	$\sqrt{N} \frac{ I - I^{\text{exact}} }{I^{\text{exact}}}$	$\frac{\Delta I}{I}$	$\frac{\sqrt{N} \Delta I}{I}$
1.0757	10^2	0.86741	0.376×10^{-2}	0.038	0.346×10^{-1}	0.300
0.83834	10^3	0.86192	0.260×10^{-2}	0.082	0.117×10^{-1}	0.320
0.84711	10^4	0.86409	0.907×10^{-4}	0.009	0.365×10^{-2}	0.315
0.86775	10^5	0.86395	0.247×10^{-3}	0.078	0.116×10^{-2}	0.318
0.86343	10^6	0.86460	0.505×10^{-3}	0.505	0.367×10^{-3}	0.317
0.86423	10^7	0.86422	0.617×10^{-4}	0.195	0.116×10^{-3}	0.317
0.86380	10^8	0.86417	0.154×10^{-5}	0.015	0.367×10^{-4}	0.317
0.86426	10^9	0.86418	0.170×10^{-4}	0.538	0.116×10^{-4}	0.317
0.86420	10^{10}	0.86416	0.315×10^{-5}	0.315	0.367×10^{-5}	0.317



蒙特卡洛方法应用：逾渗

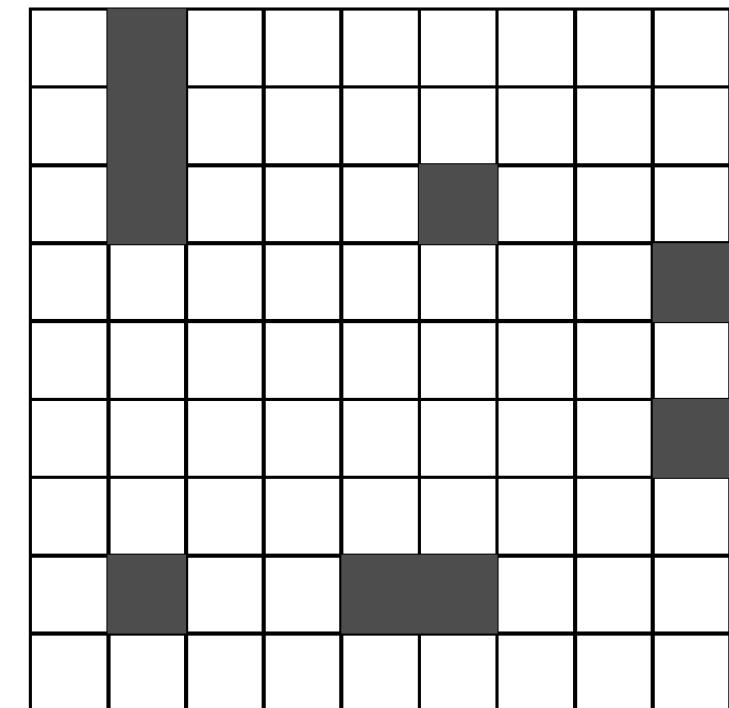
当系统的成分或某种意义上的密度变化达到一定值（称为逾渗阈值）时，在逾渗阈值处系统的一些物理性质会发生尖锐的变化，即在逾渗阈值处，系统的一些物理现象的连续性会消失（而从另一方面看，则是突然出现）。

现象或体系

- 多孔介质中流体的流动
- 群体中疾病的传播
- 通讯或电阻网络
- 导体和绝缘体的复合材料
- 超导体和金属复合材料
- 不连续的金属膜
- 螺旋状星系中恒星的随机形成
- 核物质中的夸克
- 表面上的液He薄膜
- 弥散在绝缘体中的金属原子
- 稀磁体
- 聚合物凝胶化，流化
- 玻璃化转变
- 非晶态半导体的迁移率
- 非晶态半导体中的变程跳跃

转变

- 堵塞/流通
- 抑制/流行
- 断开/联结
- 绝缘体/金属导体
- 正常导电/超导
- 绝缘体/金属导体
- 非传播/传播
- 禁闭/非禁闭
- 正常的/超流的
- 绝缘体/金属导体
- 顺磁性的/铁磁体的
- 液体/凝胶
- 液体/玻璃
- 局域态/扩展态
- 类似于电阻网络





蒙特卡洛方法应用：逾渗

逾渗模型：每一个格点可以处“占据”和“空”两种状态，设每一个格点被占据的概率为 p ，这一概率与格点的周围环境无关，这样一种模型称为点逾渗模型。

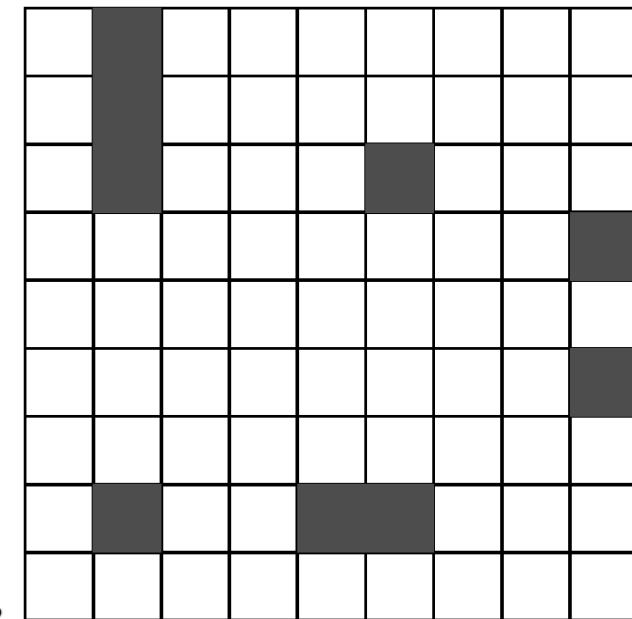
连接集团：对于给定的占据概率 p ，每一占据格点或者是孤立的，或者与最近邻占据格点形成集团。我们定义若二格点为最近邻且均被占据，则此二格点直接相连。若二格点之间可以找到一个通路，沿通路最近邻格点相连，则此二格点间接相连。集团定义为这样一组占据格点，其中的任何两个都直接或间接相连。集团的大小是逾渗问题中的一个关键量，如果存在一个大集团从格子的一边延伸到另一边，则逾渗样品将导通，否则为绝缘体。

几何相变：对于一个无穷大的格点，存在一个确定的临界概率 p_c ，当 $p > p_c$ 时，存在一个无穷大的跨越集团，当 $p < p_c$ 时，所有集团都是有限的，不存在跨越集团。逾渗的本征量是其连通性随着格点占据概率的连续变化，连通性在一个特定值 p_c 发生突变，这种由无跨越集团到有跨越集团的变化是一种相变，一般称为几何相变。

集团尺寸 s ：定义为集团内的格点数。

考虑一个大小为 s 的集团的回转半径

$$R_s^2 = \frac{1}{s} \sum_{i=1}^s (\mathbf{r}_i - \bar{\mathbf{r}})^2$$



平均集团尺寸分布 $n_s(p)$: 尺寸为 s 的集团的数目, sn_s 为尺寸为 s 的集团所占据的格点的总数

随机抽取一个占据格点时, 该格点处于某一尺寸为 s 的集团的概率: $w_s = \frac{sn_s}{\sum_s sn_s}$

平均集团尺寸: $S = \sum_s sw_s = \frac{\sum_s s^2 n_s}{\sum_s sn_s}.$ $S(p) \sim |p - p_c|^{-\gamma}$

平均连接长度 $\xi(p)$: 可以把非跨越集团回转半径的平均值定义为 ξ , 也定义为最大的非跨越集团的回转半径。当 $p < p_c$ 时 ξ 随 p 的增加而增加, 当 $p > p_c$ 时 ξ 随 p 的减小而增加。 ξ 在 $p = p_c$ 处发散, 当 $|p - p_c| \ll 1$ 时我们称系统处于临界区域, ξ 的发散行为是非解析的, 满足如下的指数关系

$$\xi(p) \sim |p - p_c|^{-\nu}$$

跨越集团概率 P_∞ : 跨越集团内包含的格点总数与占据格点总数之比. 对于一个无穷大的格点, 当 $p < p_c$ 时 $P_\infty = 0$, 而当 $p = 1$ 时, $P_\infty = 1$ 。在 $p = p_c$, P_∞ 由零变为非零且随 p 增加而增加。在临界区域, 我们期望当 $p > p_c$ 时: $P_\infty \sim (p - p_c)^\beta$

这样我们定义了三个临界指数, 对于二维问题, 这些临界指数可以解析求出, 其值为 $\beta = 5/36$, $\gamma = 43/18$, $\nu = 4/3$. 对于三维问题, 只能用各种近似方法来得到, 其结果为 $\beta \approx 0.4$, $\gamma \approx 1.8$, $\nu \approx 0.9$. 这些指数具有很大的普适性, 与晶格结构以及与关联长度的不同定义均无关. 除此之外, 它们之间还满足一定的关系, 如

$$2\beta + \gamma = \nu d$$



蒙特卡洛方法应用：逾渗的计算机模拟

Hoshen-Kopelman算法：

- 从左下角开始，并从左到右从下到上进行编号，因格点(1,1)被占据，我们给它赋予集团标号1，下一个没有占据，无需赋值。下一个占据格点是(3,1)由于其左边的格点为空，我们给它赋予下一个集团标号2。格点(4,1)左边占据，与其左边的格点同属一个集团，也赋予2。
- 然后我们到第二行，从左到右检查这一行，如果一个格点为空，则跳过该格点，如果格点被占据，则考察其左边和下边的最近邻，若此2邻均空，则给此格点赋予下一个集团标号。如果只有一个近邻被占据，则给此格点赋予其近邻的集团标号。如果格点的左和下两个近邻均被占据，则意味着原标定为不同标号的两个集团实际上是一个集团，我们必须进行合并并重新编号。但是在后面可能还会有进一步的重复和合并，我们把这一更新过程推迟到后面再进行

7	7					
7				8	3	3
	6	3			3	3
6	6	5	5	3	3	3
		5		4	3	3
1	1			4	3	3
1		2	2		3	

7	7					
7				3	3	3
	3	3			3	3
3	3	3	3	3	3	3
		3		3	3	3
1	1			3	3	3
1		2	2		3	



蒙特卡洛方法应用：逾渗的计算机模拟

3.为了记录这种合并信息，我们引入一个正规标定数组 n ，数组的每一个元素与一个集团标号相联系，对独立的集团标号，数组的对应元素赋值为该数组的下标，否则赋值对应将要合并的集团的标号。对于本例，在到达格点(6,2)前，分别有

$$np(1) = 1, \quad np(2) = 2, \quad np(3) = 3, \quad np(4) = 4.$$

而到达格点(6,2)，标号为4的集团实际上与标号为3的集团是同一集团，为了记住这一信息，我们有 $np(4)=3$.这一赋值告诉我们标号4是非正规的，而且与正规标号3相连。

4.这样一个过程仍然有不确定性，当我们到达格点(5,4)，此格点的左边标号为5下边标号为4，我们已经知道标号4为非正规标号（因为 $np(4)=3$ ），此时我们并不是按照上面的规则为(5,4)标号4，而是标号为较小标号所指的正规标号，这里为3（因为 $np(4)=3$ ）同时我们知道标号5也是非正规的，其所对应的正规编号为3，故应令 $np(5)=3$

5.当检查完所有格点后，我们可以根据数组 np 的值对非正规标号进行更新。

7	7					
7			8	3	3	
	6	3		3	3	
6	6	5	5	3	3	3
		5		4	3	3
1	1			4	3	3
1		2	2		3	

7	7					
7				3	3	3
	3	3			3	3
3	3	3	3	3	3	3
		3		3	3	3
1	1			3	3	3
1		2	2		3	



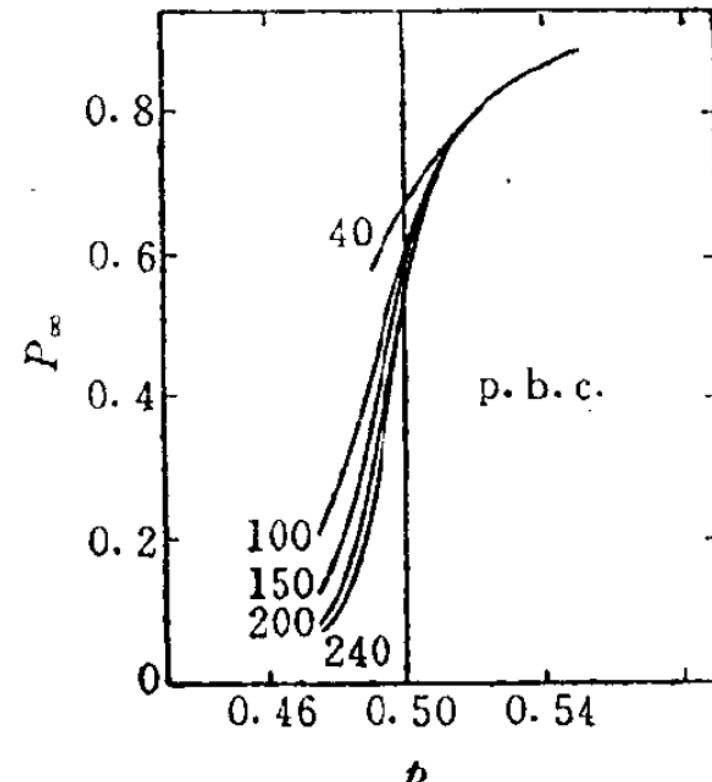
蒙特卡洛方法应用：逾渗的计算机模拟

系统平均：实际计算时，我们总是分析有限的格点，对每一个给定大小的格点和在给定 p 后生成的逾渗样本，计算得到的 P_∞ 一般并不相同，因此必需对多个样本进行平均，这样计算得到的结果与所取样本大小有关。

临界指数：为了计算临界指数，我们可在 p_c 附近对不同的 p 计算 P_∞

$$\ln P_\infty = \beta \ln(p - p_c) + C$$

有限尺寸标度：为了得到无穷大格点的结果，应对不同大小的格点进行计算，然后外推到无穷大样本。



统计物理

信息熵

假设在一个事件集合（样本集） $\{X\}$ 中，事件 x 发生的概率为 $P(x)$ 。考虑到一个事件的所携带的信息量是用来衡量该事件带来的“惊奇”程度这一事实，信息量应该满足：

1. 一个事件发生的概率越大，则该事件的信息量越小。考虑一个极端情况，如果一个事件确定要发生（例如太阳从东方升起），则该事件不会带来任何惊奇，其信息量为零。
2. 两次独立事件的信息量应该是这两次事件各自信息量之和。。

为满足上面两个性质，我们可以定义事件 x 所包含的信息量为：

$$I(x) = -\ln P(x) \quad (12.23)$$

这一信息量只针对单个事件。为了定义满足概率分布 $P(x)$ 的整个事件集合 $\{X\}$ 所包含的信息量（不确定度），我们采用 Shannon 信息熵：

$$S(X) = \mathbb{E}_{x \sim \{X\}_P} I(x) = -\mathbb{E}_{x \sim \{X\}_P} \ln P(x) = -\sum_x P(x) \ln P(x) \quad (12.24)$$

这里 $\mathbb{E}_{x \sim \{X\}_P}$ 代表对满足概率分布 $P(x)$ 的事件集合 $\{X\}$ 中所有事件求和。 \sum_x 表示对于满足均匀概率分布的 x 求和。

最大熵原理

当我们面对不确定时，应该尽可能地保留所有的可能性。例如，对一个骰子，在没有任何额外信息时，最合理的猜测是每种点数的概率均为 $1/6$ 。这一猜测的依据就是最大熵原理，即对于随机事件的概率进行预测时，我们的预测应该满足全部已知条件，同时对未知情况不做任何主观假设，这样得到的概率分布对应的信息熵最大。

考虑一个具体的例子：假设一个超市内只有 A、B、C 三种商品，其价格分别为 10、20、30 元。已知某位顾客在该超市内的平均消费为 16 元，求该顾客购买这三种商品的概率。在这里“平均消费为 16 元”为一个已知的确定性条件，如果没有这一条件，我们对该顾客的信息完全无知，可以应该合理假设该顾客购买三种商品的概率均为 $1/3$ ，但是这时平均消费为 20 元（而不是 16 元）。因此，下面我们要要求在满足约束条件“平均消费为 16 元”下，对应的可能（信息熵）最大的概率分布。

$$P_A + P_B + P_C = 1 \quad 10P_A + 20P_B + 30P_C = 16$$

$$S(P_A, P_B, P_C) = -(P_A \ln P_A + P_B \ln P_B + P_C \ln P_C)$$

利用拉格朗日乘子法，引入拉格朗日乘子 λ_0 、 λ_1 ，然后对

$$\tilde{S}(P_A, P_B, P_C) = S(P_A, P_B, P_C) + \lambda_0(P_A + P_B + P_C - 1) + \lambda_1(10P_A + 20P_B + 30P_C - 16)$$

求极值，并将结果代入上述约束条件，得到 $P_A = 0.554$ 、 $P_B = 0.292$ 、 $P_C = 0.154$ 。

考虑一般情况，假设随机变量 X 取 $X = x$ 的概率为 $P(x)$ 除此以外，一共有 l 个约束：

$$\sum_x P(x)f_k(x) = \tau_k, \quad k = 0 \dots l$$

例如归一化条件对应 $f_0(x) = 1$ ，固定平均值的约束条件对应 $f_1(x) = x$

在这些约束条件下，求信息熵的极大值，即对：

$$\sum_x P(x) \ln P(x) + \sum_{k=0}^l \lambda_k (\sum_x P(x)f_k(x) - \tau_k)$$

求极值（其中 λ_k 为第 k 个约束条件对应的拉格朗日乘子）。我们可以得到：

$$P(x) = \frac{1}{Z} \exp \left(- \sum_{k=1}^l \lambda_k f_k(x) \right)$$

其中归一化因子：

$$Z = \sum_x \exp \left(- \sum_{k=1}^l \lambda_k f_k(x) \right)$$

上面的结果可以直接推广到 x 连续取值的情况。例如，只有一个约束: $f_1(x) = x$, 即知道变量的平均值为 τ_1 , 可以求得:

$$P(x) = \frac{1}{\tau_1} \exp\left(-\frac{x}{\tau_1}\right) \quad (12.39)$$

如果存在两个约束: $f_1(x) = x$, $f_2(x) = x^2$, 即我们不仅知道随机变量的平均值 τ_1 , 还知道 x^2 的平均值 τ_2 , 我们可以求出在这两个约束条件下, 满足最大熵的概率分布为:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \tau_1)^2}{2\sigma^2}\right) \quad (12.40)$$

该分布正好是一个均值为 τ_1 , 方差为 $\sigma^2 = \tau_2 - \tau_1^2$ 的高斯分布。这也为自然界中高斯分布的广泛存在提供了一个解释: 即该分布为满足均值和方差一定的情况下最大概然分布。

近独立粒子系统： 系统中粒子相互作用可以忽略， 系统能量可以表达为单个粒子能量之和

$$E = \sum_{i=1}^N \varepsilon_i$$

微观状态 $(q_{ix}, q_{iy}, q_{iz}, p_{ix}, p_{iy}, p_{iz})$

宏观状态：对于一个孤立系统，可以用粒子数N，体积V，和能量E等宏观量来标准系统的平衡态，统计物理认为，宏观物质的特性是大量微观粒子运动的集体表现，宏观物理量是相应微观量的统计平均值，为了研究系统的宏观特性，没有必要，也没有可能追随微观状态的复杂变化，只要知道各个微观状态出现的概率，就可以用统计方法求围观两的统计平均值。

确定各个微观状态出现的概率是统计物理的根本问题！

对于一个孤立系统，在确定的宏观条件下，系统可能的微观状态是大量的（且微观状态不断发生着极其复杂的变化）

等概率原理：对于处于平衡态的孤立系统，系统各个可能的微观状态出现的概率是相等的！（对于有相同的N, E, V的各种微观状态，没有理由认为那一个微观状态出现的概率应该更大一些）

分布和微观状态：

能级 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l, \dots$	$\sum_l a_l = N$
粒子数 $a_1, a_2, \dots, a_l, \dots$	$\sum_l a_l \varepsilon_l = E$

(为简便起见，省略简并度)

分布和微观状态是两个不同的概念，给定了一个分布 $\{a_l\}$ ，只是确定了每一个能级 ε_l 上，
在给定分布后，为了确定微观状态，还需要确定在各个能级 ε_l 上是哪 a_l 个粒子

对于玻尔兹曼系统（经典粒子），粒子可分辨。与分布 $\{a_l\}$ 相对应的系统的微观状态数是：

$$\Omega = \frac{N!}{\prod_l a_l!}$$

在所有满足宏观条件 $\sum_l a_l = N, \sum_l a_l \varepsilon_l = E$

的分布中，微观状态最多的分布，出现的概率最大，称为最概然分布（玻尔兹曼分布）

斯特林公式： $\ln m! \approx m(\ln m - 1), \text{ for } m \gg 1$

$$\ln \Omega \approx N \ln N - \sum_l a_l \ln a_l \quad \text{取极大值:} \quad \delta[\ln \Omega] = -\sum_l \delta a_l \ln a_l = 0$$

但是这些 δa_l 不是彼此独立的，它们必须满足条件： $\sum_l \delta a_l = 0, \sum_l \delta a_l \varepsilon_l = 0$

在约束条件下求极值： $-\sum_l \delta a_l \ln a_l - \alpha \sum_l \delta a_l - \beta \sum_l \varepsilon_l \delta a_l = 0$

$$\ln a_l + \alpha + \beta \varepsilon_l = 0, \text{ for } l = 1, 2, 3 \dots$$

$$a_l = e^{-\alpha - \beta \varepsilon_l} \quad (\text{玻尔兹曼分布})$$

拉格朗日乘子 α, β 由约束条件: $N = \sum_l e^{-\alpha - \beta \varepsilon_l}$ $E = \sum_l \varepsilon_l e^{-\alpha - \beta \varepsilon_l}$

在所有满足宏观条件 $\sum_l a_l = N, \sum_l a_l \varepsilon_l = E$ 的分布中, 最概然分布相应的微观状态数的极大值非常陡, 其他分布的微观状态数相比于最概然分布几乎为0!

热力学量的统计表达

内能:
$$U = \sum_l a_l \varepsilon_l = \sum_l \varepsilon_l e^{-\alpha - \beta \varepsilon_l}$$

配分函数:
$$Z = \sum_l e^{-\beta \varepsilon_l} \quad N = \sum_l e^{-\alpha - \beta \varepsilon_l} \quad Z = Ne^\alpha$$

$$U = \sum_l a_l \varepsilon_l = \sum_l \varepsilon_l e^{-\alpha - \beta \varepsilon_l} = \frac{N}{Z} \sum_l \varepsilon_l e^{-\beta \varepsilon_l}$$

例子: 磁场中的单自旋 $H=-h$ S ($s=1$, or -1)

配分函数:
$$Z = \sum_s e^{-\beta E_s} = e^{-\beta h} + e^{\beta h}$$

磁化:
$$\langle S \rangle = \frac{1}{Z} \sum_s s e^{-\beta E_s} = \frac{e^{\beta h} - e^{-\beta h}}{e^{\beta h} + e^{-\beta h}} = \tanh \beta h$$



伊辛模型

Ising 模型：

$$\mathcal{H} = -J \sum_{\langle ij \rangle}^N s_i s_j - B \sum_{i=1}^N s_i,$$

$$m = \frac{1}{N} \sum_{i=1}^N \langle s_i \rangle, \quad m = \frac{1}{Z} \sum_{\sigma} s_{\sigma} e^{-\mathcal{H}_{\sigma}/k_B T}, \quad \text{where } s_{\sigma} = S_{\sigma}/N, \text{ with } S_{\sigma} = \sum s_i$$

The summation s over all the possible configurations. Here Z is the partition function given by

$$Z = \sum_{\sigma} e^{-\mathcal{H}_{\sigma}/k_B T}.$$

The average of a physical quantity, can be obtained from

$$m \simeq \frac{1}{M} \sum_{\sigma=1}^M s_{\sigma},$$

with $\sigma = 1, 2, \dots, M$ indicating the configurations sampled according to the distribution function

$$\mathcal{W}(S_{\sigma}) = \frac{\exp[-\mathcal{H}_{\sigma}/k_B T]}{Z}.$$

Ising 模型的平均场方法:

Mean field:

i	A	1	3	2	6	4	8
B	4	7	4	5	0	1	

$$\langle AB \rangle = \frac{1}{6} \sum_i A_i B_i = 11.8$$

$$\langle AB \rangle \approx \langle A \rangle \langle B \rangle = 14$$

A	4	5	3	3	4	5
B	4	7	4	5	0	1

$$\langle AB \rangle = \frac{1}{6} \sum_i A_i B_i = 13.9.$$

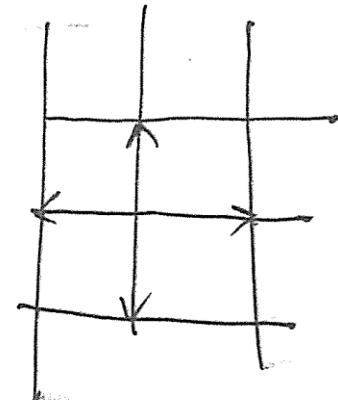
$$\langle AB \rangle \approx \langle A \rangle \langle B \rangle = 14.$$

3.1. Mean field theory of Ising model:

$$H = -\frac{1}{2} \sum_{i=1}^N \sum_{\delta} J S_i S_{i+\delta} \quad S_i = \pm 1$$

Weiss: Spin S_i see an average of neighbor S_j :

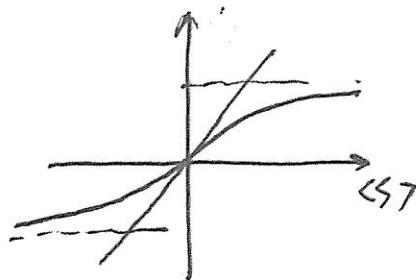
$$H_W = \sum_{i=1}^N S_i \left[\sum_{\delta} -\frac{J}{2} S_{i+\delta} \right] \approx \sum_{i=1}^N S_i \left[-\frac{JC}{2} \langle S \rangle \right]$$



$$Z_w \approx \sum_{\{S_i\}} e^{-\beta H_w}$$

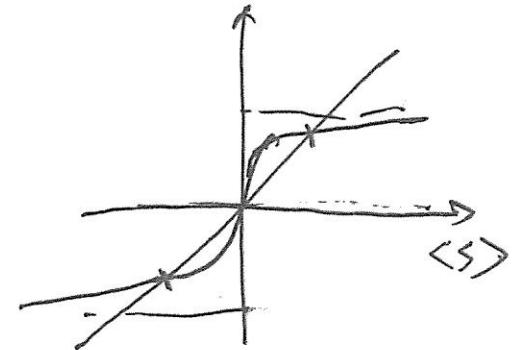
$$\langle S \rangle^w = \frac{\sum_{S_i=\pm 1} e^{-\beta H_w(S_i)}}{\sum_{S_i=\pm 1} e^{-\beta H_w(S_i)}} = \frac{e^{\frac{C_B J}{2} \langle S \rangle} - e^{-\frac{C_B J}{2} \langle S \rangle}}{e^{\frac{C_B J}{2} \langle S \rangle} + e^{-\frac{C_B J}{2} \langle S \rangle}} = \tanh\left(\frac{C_B J}{2} \langle S \rangle\right).$$

$$\frac{C_B J}{2} < 1$$



$$\boxed{T_c = \frac{C_J}{2}}$$

$$\frac{C_B J}{2} > 1$$

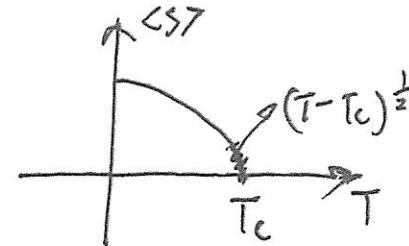


$$T = T_c(1+t)$$

$$t \langle S \rangle + \frac{\langle S^2 \rangle}{3} = 0 \quad \langle S \rangle = \pm \sqrt{3}t.$$

$$F = -k_B T \ln Z_w = -k_B T \ln 2. \quad \langle S \rangle = 0$$

$$< -k_B T \ln 2 \quad \text{if } \langle S \rangle \neq 0$$





伊辛模型

现在我们构造一个过程, 从系统的某一微观状态出发, 并在过程的每一步转移到一个新的状态. 为了确定起见, 下面用 \mathbf{x}_i 代表系统的微观状态, 如果从 \mathbf{x}_0 出发, 则这一过程产生一系列状态 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots$, 这一系列状态构成一个链. 所谓Markov 过程, 是指这样一种过程, 在过程的每一步所达到的状态只与前一状态有关, 从一状态 r 到另一状态 s 的转移通过一转移概率 $w(\mathbf{x}_r \rightarrow \mathbf{x}_s)$ 来实现. 由Markov 过程产生的一系列状态所构成的链称为Markov 链. 只要取 $w(\mathbf{x}_r \rightarrow \mathbf{x}_s)$ 满足如下条件, 就可达到我们的要求.

$$P(\mathbf{x}_r)w(\mathbf{x}_r \rightarrow \mathbf{x}_s) = P(\mathbf{x}_s)w(\mathbf{x}_s \rightarrow \mathbf{x}_r)$$

式中 $P(\mathbf{x})$ 为正则分布(6.4.10). 这一式子又称为细致平衡条件.

考虑从 r 态到 s 态的转移, 若二状态的能量差为 $\delta H \equiv H(\mathbf{x}_s) - H(\mathbf{x}_r)$

Metropolis算法:

$$w(\mathbf{x}_r \rightarrow \mathbf{x}_s) = \begin{cases} \exp[-\beta \delta H] & \text{如果 } \delta H > 0 \\ 1 & \text{如果 } \delta H \leq 0 \end{cases}$$

w 的选择并不唯一, 只要满足细致平衡的要求即可, 但不同的 w 收敛速度往往差别很大, 如何选择合适的 w 以达到尽可能快的收敛速度和尽可能高的计算精度仍然是当前Monte Carlo算法研究的前沿课题之一.

Ising 模型的蒙特卡洛模拟:

一般来说Markov链中的每一个状态与其前一个状态相差应较小,因为如果两个状态相差过大,其能量差亦较大,从而转移概率太小。计算很容易陷入相空间中初态附近一个很小的子空间内,通常有两种作法:一种是一次翻转一个自旋,这是一种不保持总自旋守恒的计算。另一种是每次交换一对相邻自旋,这种计算将保持总自旋守恒。

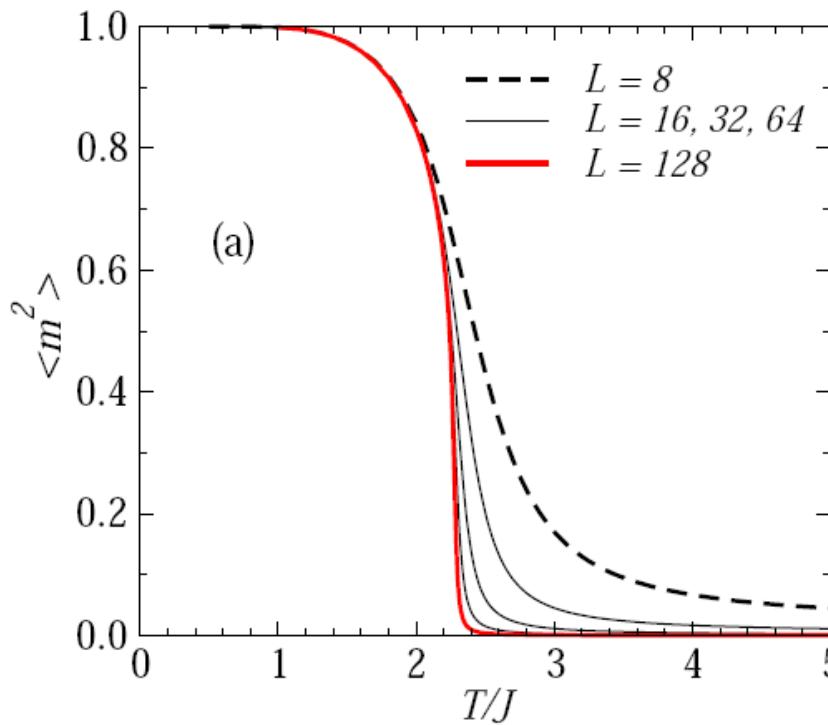
重复如下步骤多次:

- 1.选择一个格点 i , 其自旋将考虑作翻转 $s_i \rightarrow -s_i$;
- 2.计算与此翻转相联系的能量变化 H ,并计算这一翻转的转移概率 w ;
- 3.产生一在[0,1]之间均匀分布的随机数 z ;
- 4.如果 $z < w$, 则翻转该自旋,否则保持不变,**不论何种情况,其结果都作为一新的状态**;
- 5.分析该状态, 并计算平均值收集数据。

上述过程的第5步不必对每次自旋都进行,而是每间隔一个或数个蒙卡步数进行一次
另外,如在前面已经指出过的,前面若干个蒙卡步数应该舍弃。

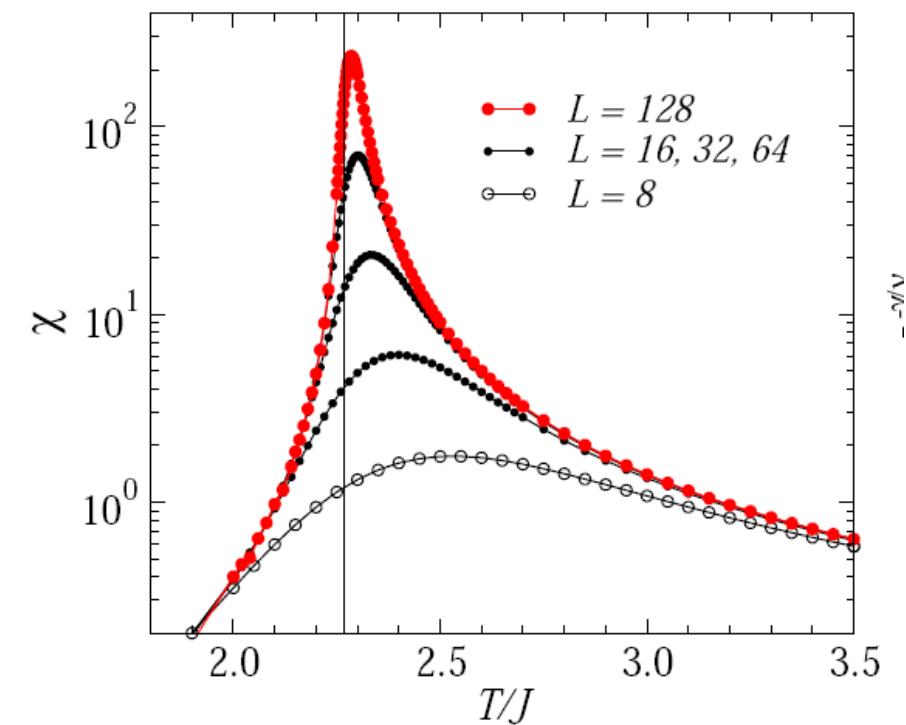
伊辛模型

Ising 模型的蒙特卡洛模拟:



$$m = \frac{1}{N} \sum_{i=1}^N \sigma_i.$$

$$m \sim |t|^\beta, \quad \beta = 1/8$$



$$\chi = \frac{d\langle m \rangle}{dh} \Big|_{h=0} = \frac{N}{T} (\langle m^2 \rangle - \langle |m| \rangle^2),$$

$$\chi \sim |T - T_c|^{-\gamma} \quad \gamma = 7/4$$



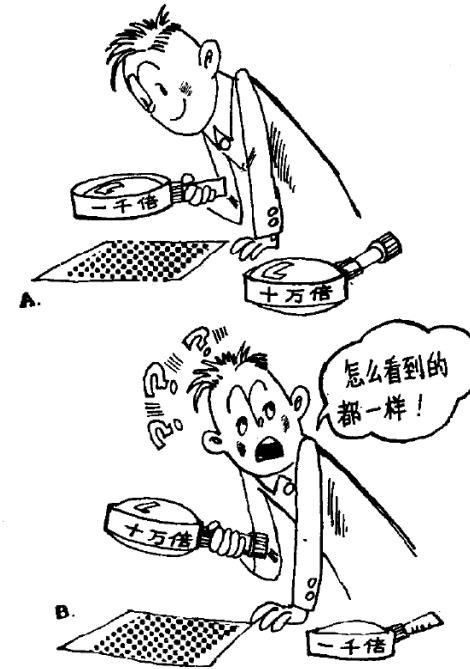
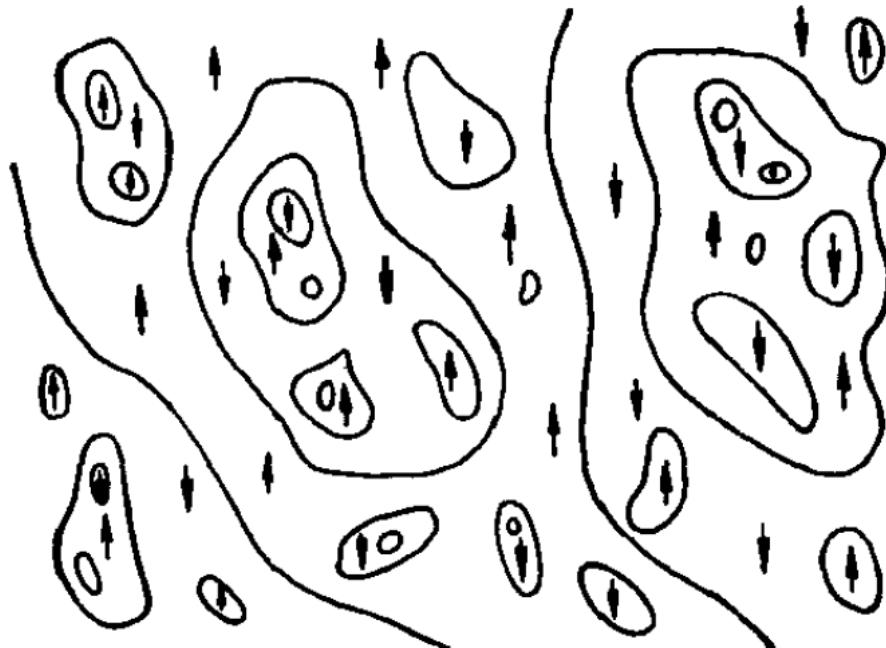
长程关联和长程序

关联函数: $C(r) = \frac{1}{N} \sum_i \langle \sigma_i \sigma_{i+r} \rangle$

在低温有序相, 系统具有长程关联, 意味着其关联函数 $C(r \rightarrow \infty) \rightarrow m^2$ 。

在高温相, 系统关联函数随距离增加指数衰减。 $\tilde{C}(r) \sim e^{-r/\xi}$ $\xi \sim |T - T_c|^{-\nu}$

在相变点上, 关联函数在长距离上呈现幂指数衰减: $C(r) \sim r^{-(d-2+\eta)}$



有限尺寸效应

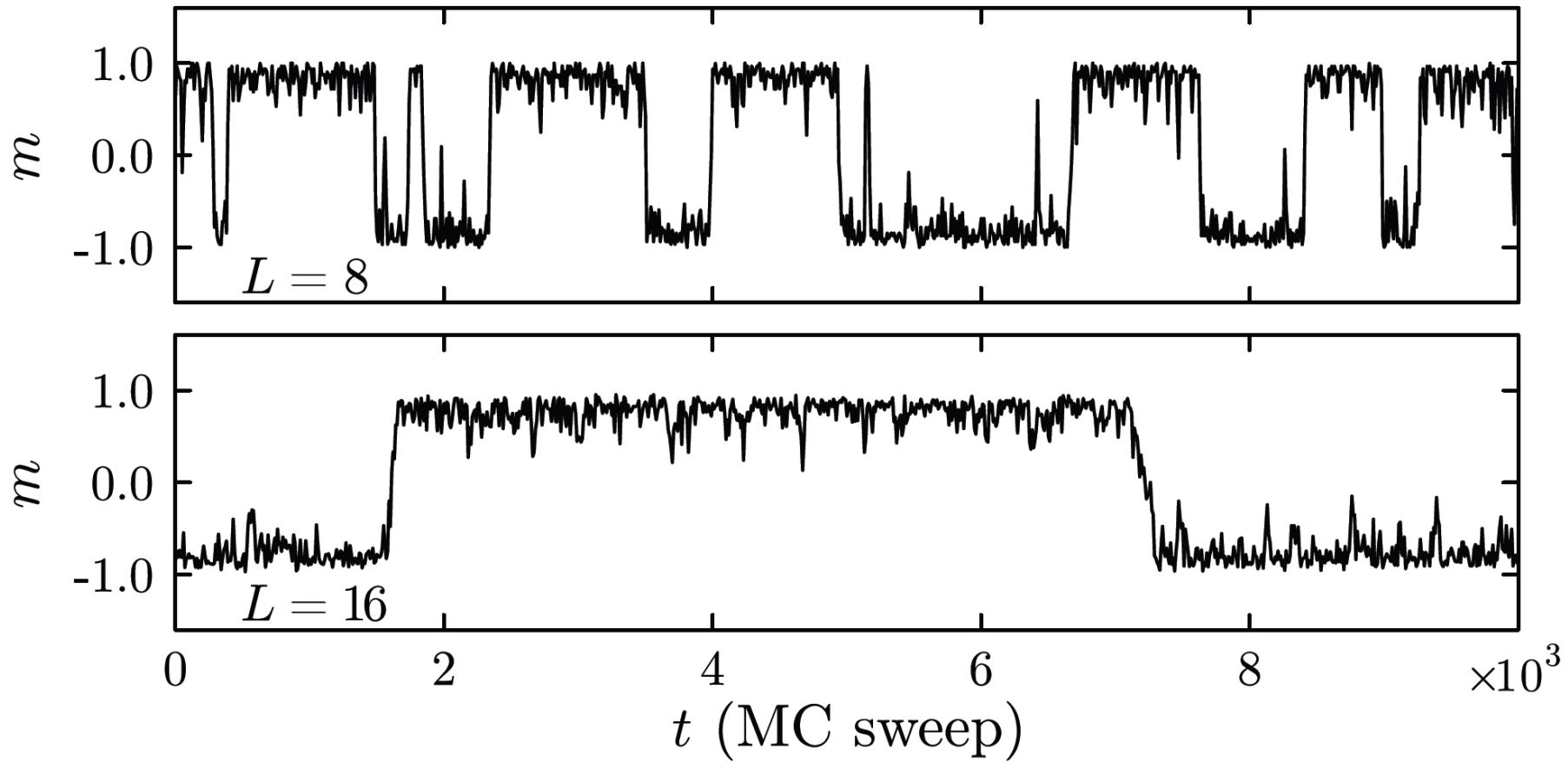
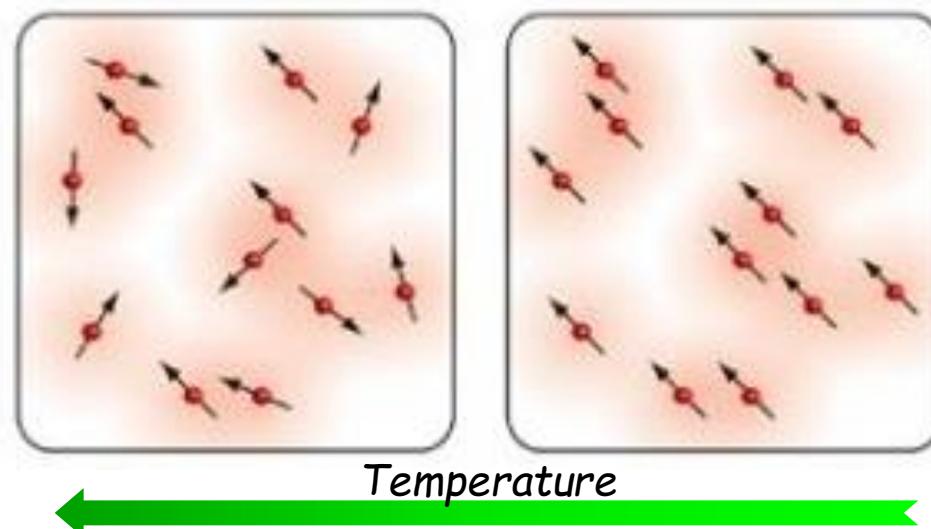


图 9.6: 在相变温度以下 ($T=2.2J$), 不同尺寸下 ($L = 8$ 和 $L = 16$) 的二维伊辛模型中, 平均磁化 m 随蒙特卡洛演化时间的变化。

Spontaneous symmetry breakings in equilibrium physics



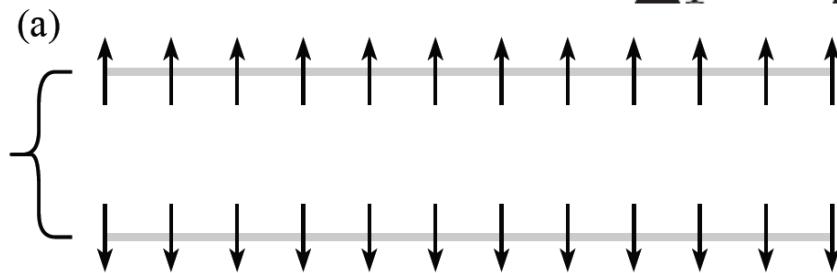
$$\rho_s = e^{-\beta H_s} \quad F = E - TS$$



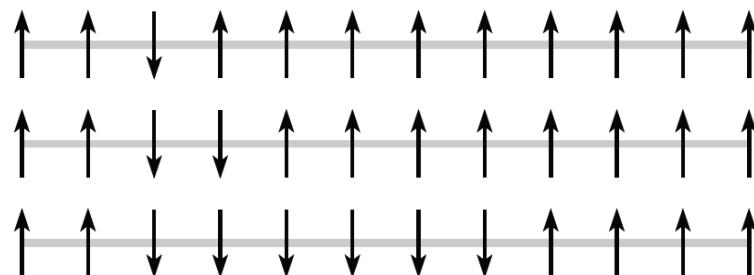
SSB in equilibrium system

- ★ Minimize (free) energy
- ★ Determined by interaction and temperature
- ★ Competition between energy and entropy

$$\Delta F = \Delta U - T\Delta S.$$



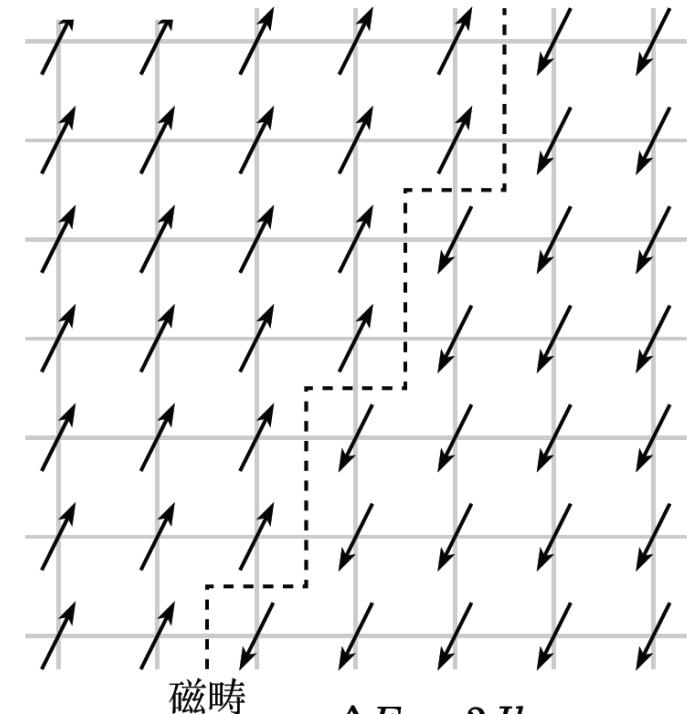
$$\text{基态 } E_g = -JL$$



$$\text{第一激发态 } E_1 = E_g + \Delta E, \Delta E = 4J$$

$$\Delta U = 4J,$$

$$\Delta S \sim \ln L,$$



$\Delta E_l = 2Jl$ 为长度为 l 的畴壁导致的能量增加。

我们可以估计其长度 l 的畴壁对应的微观状态 Ω : $2^l < \Omega < 3^l$,

因此其带来的熵的增加范围为 $\Delta S = k_B \ln \Omega \in [k_B l \ln 2, k_B l \ln 3]$.

假设存在临界温度 T_c $\Delta F = \Delta U - T_c \Delta S = 0$,

估计临界温度 T_c 的范围为 $1.821J < k_B T_c < 2.885J$,

严格解的结果告诉我们 $k_B T_c = 2.27J$,



临界慢化与动力学普适类

蒙特卡洛模拟时间 v.s. 真实演化时间

在统计物理中，有一类重要的非平衡现象，称为弛豫现象：将一个系统和一个处于热力学平衡态的大热库耦合，经过一段时间后，热库会使得系统弛豫到热力学平衡态，其温度和热库的温度相同。这里“时间”这一概念指的是自然界中真实的物理时间。原则上，要模拟这一动力学过程，需要求解这一多体系统的动力学方程，同时合理的考虑热库对动力学方程的影响。一个重要的例子是朗之万方程，通过在牛顿运动方程中加入耗散和噪声唯象地考虑热库对布朗粒子的影响。

在利用Metropolis算法计算二维伊辛模型的过程中，我们从一个初态出发，经过足够的蒙特卡洛演化步数之后（“热化”时间），系统达到热力学平衡态，之后我们才开始统计物理量的平均值。

对某些更新算法（例如Metropolis算法），蒙特卡洛模拟的热化过程可以反映一个真实物理系统耦合上热库后的弛豫过程的动力学普适类，但是对另一些更新算法（例如集团更新算法），可能在自然界的真实系统中找不到其对应的动力学普适类}。



临界慢化与动力学普适类

弛豫动力学和临界慢化

当温度接近相变温度时，系统的弛豫时间呈现幂指数发散，这一现象称之为临界慢化。

当温度正好处于相变温度时，弛豫时间随着系统尺寸 L 呈现幂指数发散：

$$\tau \sim L^z, \quad (9.46)$$

其中 z 为动力学临界指数，是不同于上一节所有临界指数的一个新的临界指数。它不仅和系统哈密顿量有关，还和系统的更新方式有关（对于上述单自旋翻转 Metropolis 的更新方法， $z = 2$ ，见图 9.10）。

在相变点附近，不仅系统的关联长度会发散，弛豫时间也会发散。弛豫时间随温度靠近相变温度时也会呈现处幂指数发散：

$$\tau_0 \sim |T - T_c|^{-\nu z} \quad (9.47)$$

利用式(9.34)我们可以得到关联长度和弛豫时间之间的关系：

$$\tau_0 \sim \xi^z \quad (9.48)$$

Glauber 动力学 v.s. Kawasaki 动力学

1. Glauber 算法与前面讨论的单自旋翻转 Metropolis 算法完全相同：在每个蒙特卡洛演化步中，随机选一个格点，翻转其自旋，计算翻转前后能量改变量 ΔE ，根据概率 $P = \min[1, e^{-\beta\Delta E}]$ 保留或拒绝该次更新[39]。
2. 在 Kawasaki 算法中，每次在系统中随机选择一对近邻自旋，如果这对自旋相同，则拒绝这次更新；如果相反，则交换他们的自旋，计算交换前后能量改变量 ΔE ，更具概率 $P = \min[1, e^{-\beta\Delta E}]$ 保留或拒绝该次更新[40]。

比较上面两种更新方法，我们发现一个根本的区别是Glauber动力学中，系统总自旋不守恒，而在Kawasaki动力学中，由于只涉及自旋交换，系统的总自旋在蒙特卡洛模拟过程中守恒。

在 Kawasaki 动力学中，弛豫时间和系统尺寸满足： $\tau \sim L^3$

因此，当总自旋守恒时，动力学临界指数变成 $z=3$ 。这也意味着，由于守恒条件的限制，Kawasaki动力学相比于Glauber动力学弛豫更慢。这是一个动力学过程中的守恒量定性改变系统动力学普适类的例子。事实上，不同的守恒量可能导致不同的动力学普适类，依据不同守恒量对动力学普适类的分类被称为“字母分类”。



蒙特卡洛方法应用：集团算法

细致平衡: $\sum_{\nu} p_{\mu} P(\mu \rightarrow \nu) = \sum_{\nu} p_{\nu} P(\nu \rightarrow \mu).$

Given that we wish the equilibrium distribution to be the Boltzmann distribution, clearly we want to choose the values of p_{μ} to be the Boltzmann probabilities:

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{p_{\nu}}{p_{\mu}} = e^{-\beta(E_{\nu} - E_{\mu})}.$$

We break the transition probability down into two parts: $P(\mu \rightarrow \nu) = g(\mu \rightarrow \nu) A(\mu \rightarrow \nu)$.

The quantity $g(\mu \rightarrow \nu)$ is the selection probability, which is the probability, given an initial state μ that our algorithm will generate a new target state ν ,

$A(\mu \rightarrow \nu)$ is the acceptance ratio (sometimes also called the "acceptance probability"). The acceptance ratio says that if we start off in a state μ and our algorithm generates a new state ν from it, we should accept that state and change our system to the new state ν a fraction of the time $A(\mu \rightarrow \nu)$. The rest of the time we should just stay in the state .



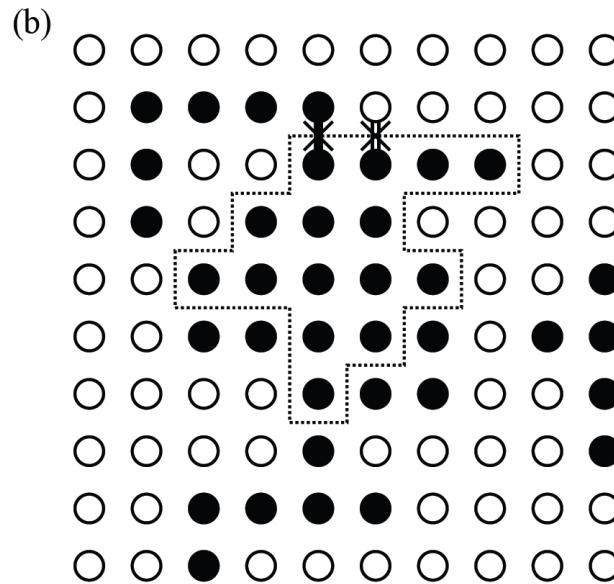
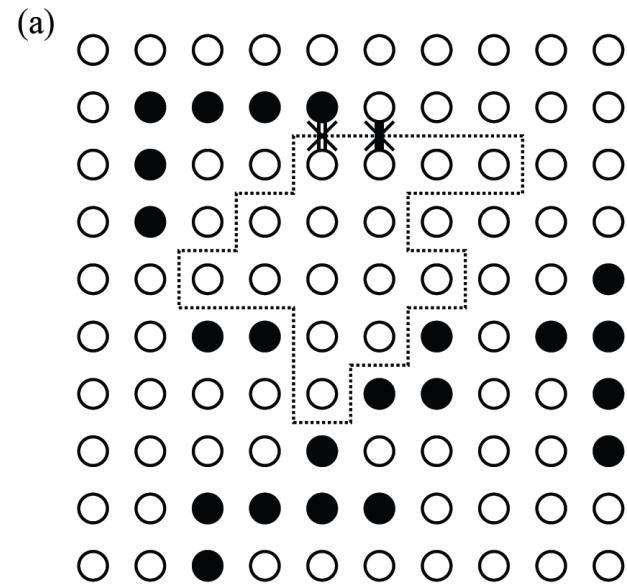
蒙特卡洛方法应用：wolff算法

自关联时间（metropolis 算法）：

$$\tau_{\text{CPU}} \sim L^{d+z},$$

Wolff算法： $\tau \sim L^{0.25}$

1. Choose a seed spin at random from the lattice.
2. Look in turn at each of the neighbours of that spin. If they are pointing in the same direction as the seed spin, add them to the cluster with probability $P_{\text{add}} = 1 - e^{-2\beta J}.$
3. For each spin that was added in the last step, examine each of its neighbours to find the ones which are pointing in the same direction and add each of them to the cluster with the same probability P_{add} (Notice that as the cluster becomes larger we may find that some of the neighbours are already members, in which case obviously we don't have to consider adding them again. Also, some of the spins may have been considered for addition before, as neighbours of other spins in the cluster, but rejected. In this case, they get another chance to be added to the cluster on this step.) This step is repeated as many times as necessary until there are no spins left in the cluster whose neighbours have not been considered for inclusion in the cluster.
4. Flip the cluster.



Notice that in each of the two states, some of the spins just outside the cluster are pointing the same way as the spins in the cluster. The bonds between these spins and the ones in the cluster have to be "broken" when the cluster is flipped.

Suppose that, for the forward move, there are m bonds which have to be broken in order to flip the cluster. These broken bonds represent pairs of similarly oriented spins which were not added to the cluster by the algorithm. The probability of not adding such a spin is $1 - P_{\text{add}}$. Thus the probability of not adding all of them, which is proportional to the selection probability $g(\mu \rightarrow \nu)$ for the forward move, is $(1 - P_{\text{add}})^m$. If there are n bonds which need to be broken in the reverse move then the probability of doing it will be $(1 - P_{\text{add}})^n$. condition of detailed balance,

$$\frac{g(\mu \rightarrow \nu) A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu) A(\nu \rightarrow \mu)} = (1 - P_{\text{add}})^{m-n} \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-\beta(E_\nu - E_\mu)},$$

$$E_\nu - E_\mu = 2J(m - n).$$

$$\frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = [e^{2\beta J}(1 - P_{\text{add}})]^{n-m}.$$

$$P_{\text{add}} = 1 - e^{-2\beta J}$$

With this choice, we can make the acceptance ratios for both forward and backward moves unity



蒙特卡洛方法应用：Swendsen-Wang算法

The Swendsen-Wang algorithm divides the entire lattice into clusters in exactly the same way, with this same probability P_{add} of making a link. But then, instead of flipping just one cluster, all the clusters are flipped with probability 0.5. That is, for each cluster in turn, we decide independently with probability 0.5 whether to flip it or not.

dimension d	Metropolis	Wolff	Swendsen Wang
2	2.167 ± 0.001	0.25 ± 0.01	0.25 ± 0.01
3	2.02 ± 0.02	0.33 ± 0.01	0.54 ± 0.02
4	-	0.25 ± 0.01	0.86 ± 0.02

TABLE 4.1 Comparison of the values of the dynamic exponent z for the Metropolis, Wolff and Swendsen-Wang algorithms in various dimensions. The values are taken from Coddington and Baillie (1992), Matz *et al.* (1994), and Nightingale and Blöte (1996). To our knowledge, the dynamic exponent of the Metropolis algorithm has not been measured in four dimensions.



阻挫磁性系统：自旋玻璃



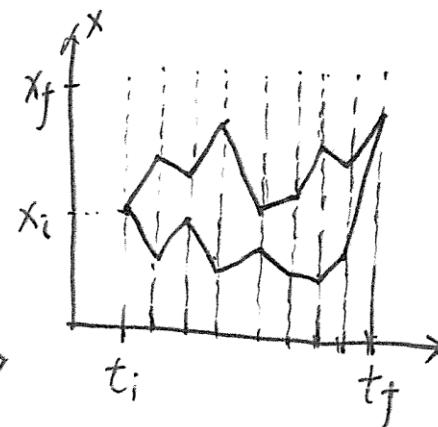
蒙特卡洛方法应用：量子系统

5.3 经典-量子 对应

* 路径积分：

$$G(x_i, t_i; x_f, t_f) = \langle x_i | e^{-i\hat{H}(t_f - t_i)} | x_f \rangle$$

$$= \sum_{\{ |x_n \rangle \}} \langle x_i | e^{-iHst} | x_1 \rangle \langle x_1 | e^{-iHdt} | x_2 \rangle \langle x_2 | \cdots \langle x_n | e^{-iHdt} | x_f \rangle$$



* 统计物理：

经典： $Z = \sum_{\{g\}} e^{-\beta H[g]}$

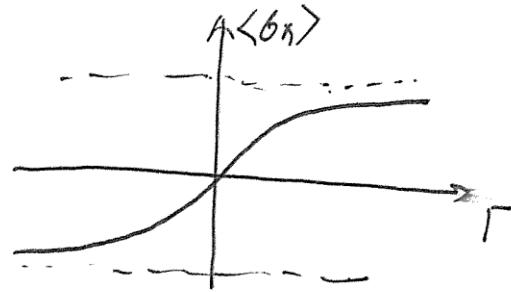
$$\boxed{\beta = M \cdot \Delta \tau}$$

量子： $Z = \text{Tr } e^{-\beta \hat{H}} = \sum_{|g_0\rangle} \langle g_0 | e^{-\beta \hat{H}} | g_0 \rangle = \sum_{|g_0\rangle} \sum_{\{|g_n\rangle\}} \langle g_0 | e^{-\beta \hat{H}} | g_1 \rangle \langle g_1 | \cdots \langle g_n | e^{-\beta \hat{H}} | g_0 \rangle$

single spin system: $H = -\frac{\hbar \hat{G}_z}{H_0} - \frac{\Gamma \hat{G}_x}{H_1}$ matrix $\begin{pmatrix} -h & -\Gamma \\ -\Gamma & +h \end{pmatrix}$

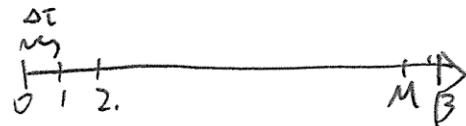
$$Z = \text{Tr} e^{-\beta \hat{H}} = 2 \cosh(\beta \sqrt{\Gamma^2 + h^2})$$

$$\langle G_x \rangle = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \Gamma}$$



$$e^{-\Delta\tau \hat{H}} \approx e^{-\Delta\tau \hat{H}_0} e^{-\Delta\tau \hat{H}_1} + O(\Delta\tau^2)$$

$$e^{-\Delta\tau \hat{H}_0} |G_i\rangle = e^{-\Delta\tau G_i} |G_i\rangle$$



$$\langle G_i | \hat{H}_1 | G_{i+1} \rangle = \langle G_i | e^{\Gamma \hat{G}_x \Delta\tau} | G_{i+1} \rangle = A e^{B G_i G_{i+1}} \quad [A = \frac{1}{2} \cosh(2\Delta\tau \Gamma) \quad e^{-2B} = \tanh(\Delta\tau \Gamma)]$$

$$Z = \sum_{\{G_i\}} e^{\beta h G_i^z + \beta B G_i^z G_{i+1}^z}$$



量子蒙特卡洛方法简介

$$\text{Tr } e^{-\beta H} = \text{Tr } e^{-\beta H_0} \underbrace{e^{\beta H_0} e^{-\beta H}}_{\hat{A}}, \quad \hat{A} = \hat{H}_0 + \hat{H}_1$$

$$A(\tau) = e^{\tau H_0} e^{-\tau H}$$

$$\begin{aligned} \frac{dA(\tau)}{dt} &= e^{\tau H_0} (-H_1) e^{-\tau H} = e^{\tau H_0} (-H_1) e^{-\tau H_0} e^{\tau H_0} e^{-\tau H} \\ &\quad \xrightarrow{\text{相互作用框架}} \\ &= - \underbrace{H_1(\tau)}_{\text{相互作用框架}} A(\tau) \end{aligned}$$

$$\begin{aligned} A(\tau) &= e^{H_1(\tau)\Delta\tau} e^{H_1(\tau-\Delta\tau)\Delta\tau} \dots e^{H_1(0)} A(0) \\ &= \underbrace{\text{Tr}}_{\text{偏时乘积}} e^{-\int_0^\tau d\tau' H_1(\tau')} A(0) \end{aligned}$$



量子蒙特卡洛方法简介

$$Z = \text{Tr} e^{-\beta \hat{H}} = \text{Tr} [e^{-\beta \hat{H}_0} T_C e^{-\int_0^\beta d\tau H_1(\tau)}]$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \dots \int_0^\beta d\tau_k T_C \text{Tr} [e^{-\beta \hat{H}_0} H_1(\tau_1) \dots H_1(\tau_k)]$$

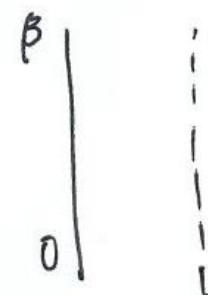
$$= \sum_{k=0}^{\infty} (-1)^k \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 \dots \int_{\tau_{k-1}}^\beta d\tau_k T_C \text{Tr} [e^{-(\beta - \tau_k) \hat{H}_0} \hat{H}_1 e^{-\beta (\tau_k - \tau_{k-1}) \hat{H}_0} \dots \hat{H}_1 e^{-\tau_1 \hat{H}_0}]$$

$$= \sum_{k=0}^{\infty} (-1)^k \int_0^\beta d\tau_1 \dots \int_{\tau_{k-1}}^\beta d\tau_k \sum_{b_0} \langle b_0 | e^{-(\beta - \tau_k) \hat{H}_0} \hat{H}_1 | b_1 \rangle \langle b_1 | e^{-(\tau_k - \tau_{k-1}) \hat{H}_0} \dots \langle b_N | \hat{H}_1 | e^{\tau_1 \hat{H}_0} | b_0 \rangle$$



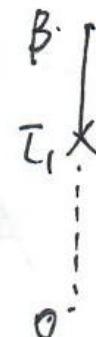
量子蒙特卡洛方法简介

$K=0.$

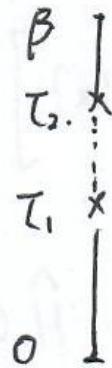


$$Z = e^{-\beta h} + e^{\beta h}$$

$$K=1 \quad (-\Gamma) \int_0^\beta d\tau_1 \sum_{b_0} \langle b_0 | e^{-(\beta - \tau_1) h b_0} \hat{H}_1 e^{\tau_1 b_0} | b_0 \rangle$$

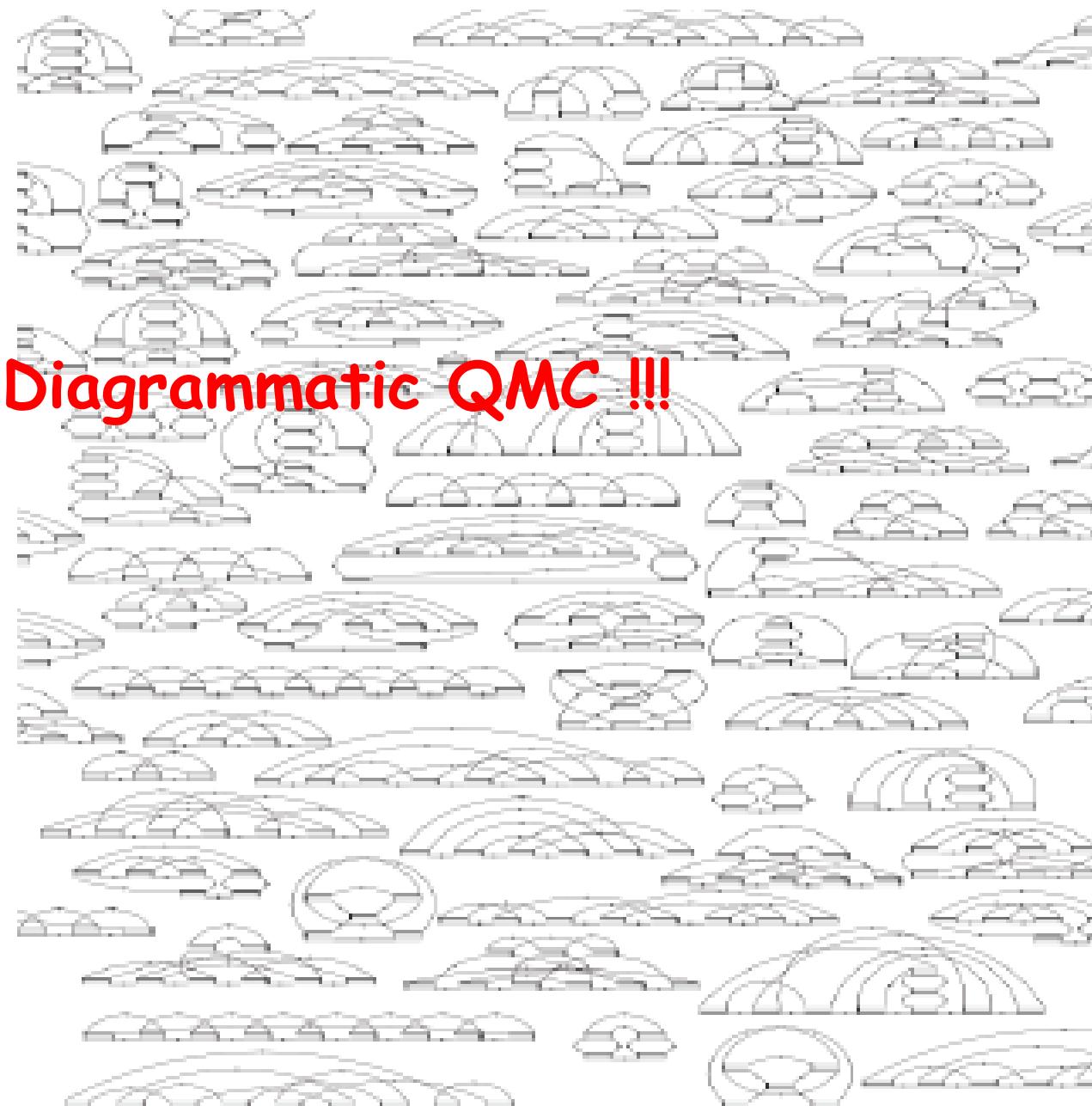


$$K=2 \quad (-\Gamma)^2 \int_0^\beta d\tau_1 \int_0^\beta d\tau_2 \sum_{b_0} \langle b_0 | e^{-(\beta - \tau_1) h b_0} \hat{H}_1 | b_1 \rangle \langle b_1 | e^{-(\tau_2 - \tau_1) h_1} \hat{H}_1 | b_0 \rangle$$



$$Z = (K=0) + \Gamma^2 (K=2) + \cdots \Gamma^K (K=t) + \dots$$

Yes, they were real thing, not a cartoon!



Diagrammatic QMC !!!