

§2.2 Monte Carlo 模拟与重要抽样

本节开始我们讨论经典多粒子体系（特别是正则系综的）的重要抽样法，它是 Monte Carlo 方法在统计力学中应用的基础。在对 NVT 系综计算平均值时，构型空间应以 Boltzmann 因子 $\exp(-\beta H)$ 作为权重，由于积分空间是高维的，因此应该用重要抽样法的 Monte Carlo 积分（参看 1.3.2 节）。用 Boltzmann 权重因子在相空间中产生随机点的想法最初来源于 von Neumann 舍选抽样法。首先可以考虑在相空间中产生一些完全随机的序列，它们以几率 $\exp(-\beta H)$ 接受，式中的能量是构型的函数并且是正的。然而，对于随机构造的状态，能量较高的构型占绝大多数，但其接受几率很小，因此，我们可能花了很多计算时间在产生构型上，但多数都被舍去了，显然，这样的抽样效率很低。另外一种构造具有统计独立构型的方法可以是根据 Boltzmann 分布产生偏向于低能的随机构型，但这种方法很难实现。因此，1953 年 Metropolis 发展了一种相关的但却摒弃了构型的统计独立性的方法，随机构型是通过一种称为 Markov 链的方式构造出来的，其中新的构型依赖于之前的构型。本节阐述 Metropolis 方法以及必要的概率论基础。

2.2.1 随机过程

在 §1.1 中我们讨论了随机数的产生，这些随机数之间的自关联性应是很弱的，即任意两个随机数之间是有统计独立性的。现在我们讨论相反的问题，即对给定的自关联系数，如何产生相应的随机数序列。

2.2.1.1 条件概率

首先我们复习概率论中一些相关的基本概念。设 $\{x(t)\}$ 是时间变量 t 的函数的系综（例如一天中的温度曲线值，分子在热运动中的坐标值），对函数取值有几率 $P(X < x, t)$ 及几率密度 $p(x, t) = dP(x, t)/dx$ ，这样的—个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。当时间变量取离散化的时间序列 (t_1, t_2, \dots) 值时，则随机过程称为随机序列。

对应于离散化的时间序列值，上面的定义可以推广到高阶，几率和几率密度为 $P_n(x_1, \dots, x_n; t_1, \dots, t_n)$ ， $p_n(x_1, \dots, x_n; t_1, \dots, t_n) = d^n P_n(x_1, \dots, x_n; t_1, \dots, t_n)/dx_1 \cdots dx_n$ 。定义条件几率为

$$p(x_n | x_{n-1}, \dots, x_1; t_n, \dots, t_1) = p_n(x_1, \dots, x_n; t_1, \dots, t_n) / p_{n-1}(x_1, \dots, x_{n-1}; t_1, \dots, t_{n-1}), \quad (2.2.1.1-1)$$

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。

当满足条件 $p_n(x_1, \dots, x_n; t_1, \dots, t_n) = p_n(x_1, \dots, x_n; t_1 + t, \dots, t_n + t)$ 时，称随机过程是平稳的，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。

对于真正的随机过程（如前面所述的随机数产生序列），随机序列是无关联的链，即 n 个量 x_1, \dots, x_n 取特定值的几率是统计无关的，

$$p(x_1, \dots, x_n) = p_1(x_1) p_1(x_2) \cdots p_1(x_n), \quad (2.2.1.1-2)$$

即条件几率仅决定于该步独立取值的几率，与以前时刻无关：

$$p(x_n | x_{n-1}, \dots, x_1) = p_1(x_n). \quad (2.2.1.1-3)$$

2.2.1.2 Markov 链

我们称一个平稳的随机序列是 Markov 的，如果某一时刻 x 取值的条件几率是独立于上一时刻之前的所有 x 值的话，

$$p(x_n | x_{n-1}, \dots, x_1) = p(x_n | x_{n-1}). \quad (2.2.1.2-1)$$

该式表示，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。例如，SAW 路径就是非 Markov 序列，因为必需要记住历史上访问过的格点，因此，当前占位的几率与整个历史上占据点有关。定义 $W_{i \rightarrow j} = p_1(x_j | x_i)$ ，即将条件几率解释成从状态 x_i （变量 x_{n-1} 的一个特定取值）转移到状态 x_j （变量 x_n 的一个特定取值）的跃迁几率，它需满足跃迁几率的条件： $W_{i \rightarrow j} \geq 0$ ， $\sum_j W_{i \rightarrow j} = 1$ 。则由 (2.2.1.1-1) 式

$$\begin{aligned} p_n(x_1, \dots, x_n) &= p(x_n | x_{n-1}, \dots, x_1) p_{n-1}(x_1, \dots, x_{n-1}) = p(x_n | x_{n-1}) p_{n-1}(x_1, \dots, x_{n-1}), \\ &= p(x_n | x_{n-1}) \cdots p(x_2 | x_1) p_1(x_1) = p_1(x_1) W_{1 \rightarrow 2} \cdots W_{n-1 \rightarrow n} \end{aligned} \quad (2.2.1.2-2)$$

设可能的状态总数目为 M ，则转移概率构成 $M \times M$ 矩阵 $\mathbf{W} = \{W_{ij}\}$ 。我们用矢量 $\mathbf{p}(1)$ 表示系统初始态构型的几率，那么下一个状态的几率分布为 $\mathbf{p}(2) = \mathbf{p}(1) \mathbf{W}$ ，则在 N 步之后，系统平衡态的几率分布为 $\mathbf{p}(N) = \lim_{N \rightarrow \infty} \mathbf{p}(1) \mathbf{W}^{N-1}$ 。

Markov 链的极限分布是与初始分布的选择无关的，仅决定于转移概率，这就保证了模拟的原子系统的平衡态是按照 Boltzmann 因子进行几率分布的。为了说明该点，我们考虑一个两能级系统，其第一和第二能级分布的 Boltzmann 因子是 2:1，因此，平衡态下的构型分布应该为矢量 $(2/3, 1/3)$ 。用下面的跃迁矩阵可以达到该极限分布，

$$\mathbf{W} = \begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}, \quad (2.2.1.2-3)$$

即 $W_{1 \rightarrow 1} = 0.5$ ， $W_{1 \rightarrow 2} = 0.5$ ， $W_{2 \rightarrow 1} = 1$ ， $W_{2 \rightarrow 2} = 0$ 。显然，按此转移概率，第一能级的最终占据几率要大。设初始分布为 $\mathbf{p}(1) = (1, 0)$ ，即第二能级完全没有被占据，乘以跃迁矩阵后可得 $\mathbf{p}(2) = (0.5, 0.5)$ ， $\mathbf{p}(3) = (0.75, 0.25)$ ，连续乘下去得到平衡态分布 $(2/3, 1/3)$ 。

达到平衡后，应有

$$\mathbf{p} = \mathbf{p} \mathbf{W}, \quad p_i = \sum_j p_j W_{ji}. \quad (2.2.1.2-4)$$

如两能级系统例子中,

$$\begin{pmatrix} 2/3 & 1/3 \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}. \quad (2.2.1.2-5)$$

2.2.1.3 主方程

因此我们的任务就是寻找到合适的转移概率,使之能给出平稳的分布 \mathbf{p} 。对于有 $M \times M$ 个元素的矩阵 \mathbf{W} 及有 M 个元素的 \mathbf{p} (M 可以不限于有限大小,但这里认为它是有限值), (2.2.1.2-4) 式的解有很多,而我们只要找到一个这样的解就行了,下面所述的 Metropolis 方法就是其一解。

现在我们进一步阐述 Markov 链的性质。用连续时间变量 t 推导主方程,它描述了 Markov 过程中几率密度 $p(x, t)$ 的变化。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: (1) x 的构型在 Δt 后变到 x' 构型; (2) x' 构型进入到 x 构型,因此有

$$p(x, t + \Delta t) - p(x, t) = - \int dx' W(x \rightarrow x') p(x, t) \Delta t + \int dx' W(x' \rightarrow x) p(x', t) \Delta t, \quad (2.2.1.3-1)$$

在无限小 Δt 时, 即为:

$$\frac{\partial p(x, t)}{\partial t} = \int dx' [-W(x \rightarrow x') p(x, t) + W(x' \rightarrow x) p(x', t)]. \quad (2.2.1.3-2)$$

该式称为主方程, 相当于几率守恒方程, 即对所有时间均有 $\int p(x, t) dx \equiv 1$ 。该方程还揭示了 Markov 过程的基本性质, 即 t 时刻的状态完全决定了系统的发展。也就是说, 系统在此之前的历史发展轨迹的记忆没有保存下来。

真正能用主方程描述的物理系统极少。但统计力学中在处理多粒子多自由度系统时, 如果只抽取某些重要的自由度而忽略其它的自由度, 即引入一类随机更新的动力学变量以补偿略去的微观细节, 则系统的行为可以用这种简单的过程描述。主方程的重要性在于, 在 Monte Carlo 模拟中起着巨大作用的重要抽样法 (Metropolis 方法) 可以解释成 Markov 链 - 随机行走。

由于我们要求的是平稳分布 (即平衡态分布), (2.2.1.3-2) 式左边取零后得,

$$\int dx' W(x \rightarrow x') p(x, t) = \int dx' W(x' \rightarrow x) p(x', t). \quad (2.2.1.3-3)$$

达到平衡后, 与时间无关, 故有

$$\frac{p(x)}{p(x')} = \frac{W(x' \rightarrow x)}{W(x \rightarrow x')}, \quad (2.2.1.3-4)$$

该式即为统计力学中的细致平衡解。对于它的物理意义, 我们可以用随机游动的醉鬼为例形象地加以说明: 设 x 附近醉鬼的密度为 $N(x)$, 对应于 (2.2.1.3-1) 式, 一些随机游动的醉鬼在不断地离开或进入 x 位置, 从而造成数目的变化,

$$\begin{aligned}\Delta N(x) &= -\int dx' \{W(x \rightarrow x')N(x) - W(x' \rightarrow x)N(x')\} \\ &= -\int dx' W(x \rightarrow x')N(x') \left\{ \frac{N(x)}{N(x')} - \frac{W(x' \rightarrow x)}{W(x \rightarrow x')} \right\},\end{aligned}\quad (2.2.1.3-5)$$

当进出达到平衡时,

$$\frac{N_e(x)}{N_e(x')} = \frac{W(x' \rightarrow x)}{W(x \rightarrow x')}. \quad (2.2.1.3-6)$$

对于偏离平衡的情况, 如过多的醉鬼进入 x , 则 $N(x)/N(x') > N_e(x)/N_e(x')$, 因此必有 $\Delta N(x) < 0$, 即过多的醉鬼会离开, 反之亦然。在醉鬼数目 N_0 足够多时, 其分布 $N(x)$ 最终会达到平衡态分布 $N_e(x) = N_0 p(x)$, 这就是统计力学的细致平衡原理。

(2.2.1.3-4) 式是可逆 Markov 链, 即由某步到达 x 并在下一步进入 x' 的几率和反方向进行的几率相同。我们可以验证, (2.2.1.3-4) 式就是 (2.2.1.2-4) 式。这说明, 几率密度守恒条件或主方程与 Markov 链存在对应关系。将变量 x 和 x' 分别用下标 i 和 j 替换, 则方程 (2.2.1.3-4) 为

$$p_i W_{ij} = p_j W_{ji}, \quad (2.2.1.3-7)$$

对下标 j 求和, 由转移概率的归一化式, $\sum_j W_{ij} = 1$, 故得 (2.2.1.2-4) 式。如前所述, 该方程的数目有 $2M$ 个, 而 \mathbf{W} 矩阵元有 M^2 个, 因此细致平衡下的主方程式不能唯一地确定跃迁矩阵。

2.2.1.4 Markov 过程的物理应用

除了 Metropolis 抽样是 Markov 链之外, 许多真实的物理系统的 Monte Carlo 方法模拟多是基于 Markov 过程, 因此可以说, Markov 链的概念是 Monte Carlo 模拟的关键。在多粒子体系的经典统计物理中, 动力学变量服从 Newton 方程, 相空间中系综代表点的几率密度按 Liouville 方程发展, 分子动力学方法就是按力学方程追踪相空间中单个代表点随时间变动的决定性轨迹。而在 Monte Carlo 方法中, 我们只是对相空间点进行随机抽样, 与物理时间无关, 对大量随机抽样点的平均与对大量轨迹的平均值相同, 这是 Monte Carlo 方法与分子动力学方法的主要区别之处, 除此之外, Markov 链是相空间的重要抽样, 只有对配分函数有重要贡献的区域才被更多地抽出。

因此, Markov 过程是不能应用于完全决定论的体系的。但更多情况下我们关心的不是体系中每个粒子的坐标和动量的时间演变, 而只是要求一个粗粒平均的图像。在粗粒平均中, 只有少数自由度保留下来, 而对那些刻画极为细微的空间和极快的时间上的细节均予抹去, 由此产生相关较慢的自由度之间的随机跃迁。这样, 在不同的时间尺度上有可能将 Liouville 方程改写成主方程。

例如, 我们现在考察 AB 二元合金在低温下原子间的扩散。固体金属中形成晶格, 每个格点或被 A 原子占据, 或被 B 原子占据, 或为空位。A 原子或 B 原子可以以几率 Γ_A 或 Γ_B 跳迁至最邻近空位, 许许多多这种随机的原子间的扩散使得浓度梯度最后降低。原子在可以占据的位子上的分布完全是随机的, 转移概率可以和近邻有关也可以是常数。考虑到固体中的势能, 跳迁事件是个热激发过程, $\Gamma \propto \exp(-\Delta E/k_B T)$, 势垒高度 ($\Delta E \approx 1\text{eV}$) 要远大于热运动能。因此, 空位从一个格点移到另外一个格点的时间要远大于晶格振动的时间(振动频率的倒数): 声子振动的时间在 10^{-13} 秒, 空位移动所需的时间要慢 10 倍。这样, 时间尺度上已经很好的分离了。在低温下原子振动的幅度仅有原子间距的 1% 量级。因此, 如果用数值求解 Newton 方程来计算体系中所有原子运动的分子动力学方法, 在有限的计算资源下采到的跳迁事件的样本点太少。而主方程允许我们用 Monte Carlo 方法来直接模拟重要的跳迁过程, 并完全忽略声子, 当然, 跳迁几率需作为计算的输入参数。许多物理过程可以用 Markov 过程描述, 如自旋玻璃中磁化的驰豫、溶液中巨分子的链团结构、粒子的随机 Brown 运动和扩散现象、物质中粒子的输运过程、表面有序单畴的结晶生长、多层膜的外延生长、逾渗和临界现象等等。

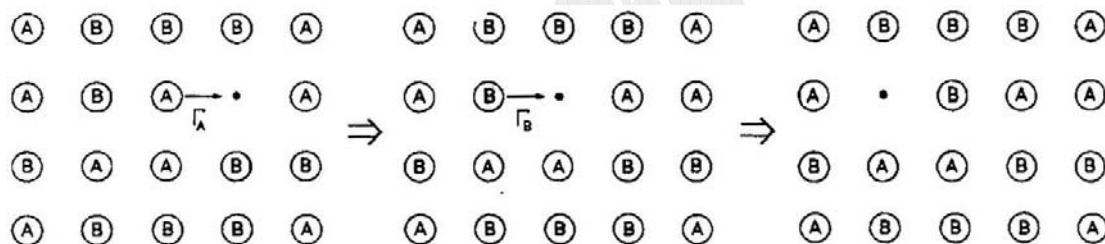


图 1.11 二元合金中空位的内扩散过程。

2.2.2 Metropolis 方法

第一章中所述的按给定分布产生随机抽样的方法很难推广到高维下复杂分布函数的情况, 因此还需要其它办法来对这样的分布进行抽样。Metropolis 等人正在进行中子穿透材料的模拟时发现了一个重要抽样方法可以改进 Monte Carlo 计算的平均值, 由此法产生的随机数满足细致平衡条件, 也就是说, 随机数的分布服从给定的平衡态分布, 或者等效的是按平衡态分布产生的。Metropolis 方法现已成为计算物理的一个重要基础, 这不仅是因为, 对积分变量在一给定值下的任意分布函数, 用此方法即可产生该分布的随机数或重要抽样, 尤其是, 它对于统计物理问题中经常遇到的分布最为有用, 因为正则系综的配分函数可以是系综中所有坐标的复杂函数分布, 即使是非统计物理问题也可以通过 Boltzmann 型因子转化成统计物理问题进行求解, 如量子 Monte Carlo 的路径积分方法那样。

2.2.2.1 抽样规则

Metropolis 提出的 (2.2.1.3-7) 式的解是一个非对称的。设由 $i \rightarrow j$ 的跃迁几

率为 $W_{ij} = T_{ij}A_{ij}$ ，其中 T_{ij} 是由 x_i 选择步进至 x_j 的几率， A_{ij} 是接受该步的几率。可以假定 T 矩阵是对称的， $T_{ij} = T_{ji}$ ，即选择步进时 x_i 和 x_j 是等价的。按 Metropolis 提出的规则，接受几率是非对称的，根据待满足的几率分布 p 而定，转移概率表示式为：

$$W_{ij} = \begin{cases} T_{ij}, & \text{if } p_j > p_i; \\ T_{ij}(p_j/p_i), & \text{if } p_j < p_i, \end{cases} \quad (2.2.2.1-1)$$

式中给出的态下标 i 和 j 是不同的。当它们相同时，应该根据转移概率的归一化式得出：

$$W_{ii} = 1 - \sum_{j \neq i} W_{ij} \quad (2.2.2.1-2)$$

可以证明，由 (2.2.2.1-1) 式给出的抽样规则满足细致平衡条件 (2.2.1.3-7) 式。另外一种对称的 Barker 抽样规则同样满足细致平衡条件：

$$W_{ij} = T_{ij} \frac{p_j}{p_j + p_i}, \quad i \neq j \quad (2.2.2.1-3)$$

(2.2.2.1-1) 式仅是解法的数学抽象表示，下面我们具体阐述抽样的方法。仍用变量 x 表示物理系统的所有动力学变量，对 N 个粒子系统，它代表了 $3N$ 个坐标分量和 $3N$ 个动量分量。Metropolis 方法的实际做法有点像舍选法，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 $r = p(x_t)/p(x_n)$ ：

- 1、如果 $r > 1$ 则选取，即 $x_{n+1} = x_t$ ($r > 1$)。
- 2、否则，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 $x_{n+1} = x_t$ ($r < 1$) 的选取几率为 r 。否则，放弃；即 $x_{n+1} = x_n$ 。

图 2.2.2.1-1 为其抽样过程的示意图。该方法是一种重要抽样法，即抽样得到

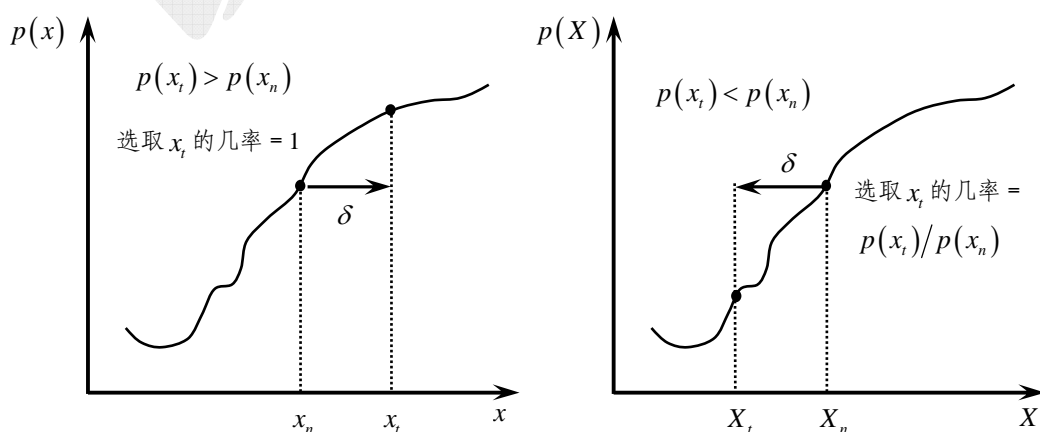


图 2.2.2.1-1 Metropolis 抽样方法的过程。

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, 可以验证, 相关系数 $C(l) \neq 0$ 。Markov 链现在可以看成是有偏压的随机行走, 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。如对积分所做的那样, 只要抽样出来的点分布与被积函数曲线形态一致, 我们并不关心点之间是如何形成分布的。这正是 Metropolis 方法用于统计力学分布抽样的关键, 而在理想随机数的应用时我们却是要避免随机数之间的关联性。

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。注意在计算系综统计平均时, 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

$$\langle A \rangle = \frac{1}{n-m} \sum_{i=m+1}^n A_i. \quad (2.2.2.1-4)$$

2.2.2.2 证明

下面接着要证明的是由 Metropolis 法抽样产生的随机点的平衡态分布 $N_e(x)$ 确实正是 $p(x)$ 。设 $W(x \rightarrow y) = T(x \rightarrow y)A(x \rightarrow y)$, T 是由 x 作试探步进至 y 的几率, 它对于 x 和 y 是对称的: $T(x \rightarrow y) = T(y \rightarrow x)$, A 是接受该步的几率, 即由 Metropolis 抽样法则中规定的。因此, 细致平衡条件 (2.2.1.3-6) 或 (2.2.1.3-7) 式可写成,

$$N_e(x)/N_e(y) = A(y \rightarrow x)/A(x \rightarrow y), \quad (2.2.2.2-1)$$

如果 $p(y) > p(x)$, 由法则(1), 该步被接受, $A(x \rightarrow y) = 1$ 。由法则(2), $A(y \rightarrow x) = p(x)/p(y)$ 。如果 $p(y) < p(x)$, $A(x \rightarrow y) = p(y)/p(x)$, 且有 $A(y \rightarrow x) = 1$ 。故 (2.2.2.2-1) 式在任何情况下都有 (图 2.2.2.2-1)

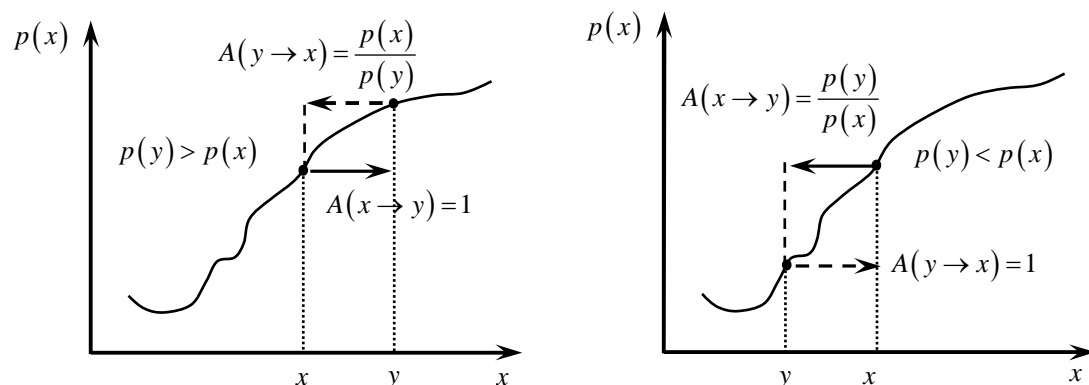


图 2.2.2.2-1 Metropolis 方法抽样中的接受概率。

$$N_e(x)/N_e(y) = p(x)/p(y). \quad (2.2.2.2-2)$$

[作业]: 设体系的能量为 $H = x^2/2\sigma_x^2 + y^2/2\sigma_y^2$ (以 $k_B T$ 为单位), 采用 Metropolis 抽样法计算 $\langle x^2 \rangle, \langle y^2 \rangle, \langle x^2 + y^2 \rangle$, 并与解析结果进行比较。抽样时在 2 维平面上依次标出 Markov 链点分布, 从而形象地理解 Markov 链。

2.2.3 不同系综的 Monte Carlo 方法

在分子动力学模拟中, 总能量 E 和总动量 \mathbf{P} 是守恒量, 因此系综平均和微正则系综 (即恒 $NVE - \mathbf{P}$) 中的相似。一般来说, Monte Carlo 模拟研究的是正则系综。系综的不同导致分子动力学和 Monte Carlo 计算统计平均的不同, 但差别在数百个粒子的体系中已经变得很小, 而且存在许多与不同系综涨落相关的方法。例如, 现在可以在恒温、恒压等条件下进行分子动力学模拟。对 Monte Carlo 来说, 系综的选择更广泛: 等压、等温、等张力等温、微正则、巨正则, 除此之外, 还增加了一种研究多相共存的 Gibbs 系综技术。尽管各系综方法有所差异, 但基本思想仍是采用重要抽样法, 用 Monte Carlo 技术对系综平均有重要贡献的相空间区域产生随机行走, 并选择接受规则, 以使按相应的概率分布产生构型。

2.2.3.1 正则系综

正则系综的若干具体应用将在 §2.3 中给出, 由于该系综的几率分布函数就是 Boltzmann 分布, $p(x) \propto \exp(-\beta H)$, 因此, 直接代入 Metropolis 抽样方法即可。以一个相互作用经典粒子体系为例, 粒子间的相互作用势能为 $U(x)$, 正则配分函数中可以对粒子的动能项积分, 因此我们仅需要考虑粒子体系中各粒子的空间坐标, 各坐标

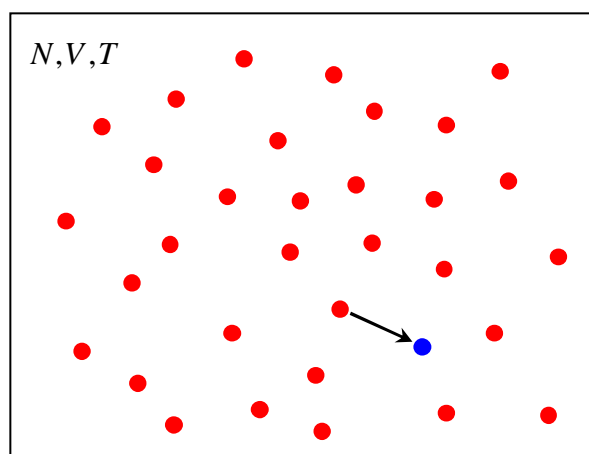


图 2.2.3.1-1 正则系综中粒子数、体积和温度是固定的, 一个粒子空间位置的移动即为 Monte Carlo 模拟的一步。

取值的一个集合构成一个构型： $x \rightarrow (\mathbf{r}_1, \dots, \mathbf{r}_N)$ ，改变任一个粒子坐标值将产生一个新构型。因此模拟方法为：

在固定的空间体积范围内，对给定的粒子数任意确定它们的坐标值（不考虑每个粒子的动能），形成一个初始构型 x_0 ，设置固定位置改变步长 Δ 值（可根据模拟结果调整）。计算该构型下的相互作用能量 $U(x_0)$ ，它是粒子对之间距离的函数，根据相互作用势能形状选择一截断距离，对每个粒子计数截断距离内的粒子对，计算每对之间的相互作用势并求和，注意求和中一对粒子只能计数一次。随机选择一个粒子 i ，给此粒子一随机位移， $\mathbf{r}_i \rightarrow \mathbf{r}_i + \delta = \mathbf{r}_i + (\xi - 0.5)\Delta$ ，其中 ξ 代表关于直角坐标的三个随机数，这样生成了一个新构型 x_1 。计算该构型下的相互作用能量 $U(x_1)$ 。接受此粒子移动的几率为

$$p_a = \min \{1, \exp[-\beta(U_1 - U_0)]\}, \quad (2.2.3.1-1)$$

用另一新的随机数 ξ 和该值比较进行选择。如果 $\xi > p$ ，则拒受该构型，仍保持原构型不变，但在进行平均值计算的 (2.2.2.1-4) 式中要计数一次。将热化阶段的初始 m 个构型扔去，每产生一个新构型，计算待求物理量在其中的取值。

2.2.3.2 微正则系综

大多数实验是在 NPT 下进行的，有时也是在 μVT 或 NVT 下。而在固定能量的 NVE 下的实验是很少的。对稠密液体或固体，几乎不可能进行 NVE 系综的 Monte Carlo 模拟，设想采用正则系综中的模拟方式，我们将使 N 和 V 保持恒定，随机地改变粒子体系的动力学变量并尽可能多地计数微观状态的抽样，但是只能选择那些能保持总能固定的微观状态进行系综平均。显然，在大量的微观状态中，只有极少数的态才能有固定的能量，而绝大多数状态的能量在平均值 $\langle E \rangle$ 的上下进行涨落，涨落幅度可以很大，因此几乎所有的状态都被舍去了，效率极低。

Creutz 等开发了一个效率较好的微正则系综 Monte Carlo 模拟方法，它不使用随机数来确定某步 Monte Carlo 移动是否接受，而是用以下方式来实现。将所考虑的体系附加一个外部自由度，由于历史的原因该外部自由度被称为 demon（小恶魔、守护神），小鬼在力学体系中来回游荡，和体系交换能量从而改变体系的力学变量。如果体系的能量减小的话，小鬼拿走这部分能量；体系的能量要增加的话，如果小鬼手里握有足够的能量，它把这部分能量又交还给体系。这样保持了体系总能的固定，唯一需要满足的条件是小鬼的能量不能为负。

对于经典粒子体系，其具体算法是：设体系的初始构型是 x_0 ，势能是 $U(x_0)$ ，把此体系的总能量固定在某个值 E 上，附加自由度是体系能量的剩余部分（小鬼的能量）， $E_d = E - U$ ，它应总是非负值，这时开始 Monte Carlo 步骤：

- 1、随机选择一个粒子进行尝试移动，形成一个新构型 x_1 ，计算体系的能量变化， $\Delta U = U(x_1) - U(x_0)$ 。
- 2、如果 $\Delta U < 0$ ，即体系的能量降低的话，接受该步移动，将 $|\Delta U|$ 交给 demon，即 $E_d \rightarrow E_d + |\Delta U| = E_d - \Delta U$ 。

3、如果 $\Delta U > 0$ ，即体系的能量增加的话，检查 demon 是否能提供这部分能量（即 $E_d > \Delta U$ ？）：是，则接受该步移动，且 $E_d \rightarrow E_d - \Delta U$ ；否，则拒受此尝试，系统保持原构型不变， E_d 值也不变。

可见，除了试探移动是随机的以外，上述接受步骤没有使用随机数。在足够多的步数之后，力学体系和 demon 各自达到平衡， E_d 的分布也是服从 Boltzmann 几率分布的， $p(E_d) \propto \exp(-\beta E_d)$ ，而大体系的总能保持为 E 不变。由于力学体系是个多自由度的系统，相比而言 demon 只有一个自由度，因此力学体系能量的涨落非常小，为 $1/\sqrt{N}$ 量级，可认为是很好地代表了一个微正则系综。

[作业]: 考虑一维经典粒子组成的理想气体，由于无相互作用，各粒子的能量不依赖于其位置，只需考虑它的动能，因此体系的构型即是各粒子速度坐标值的集合。给定粒子的质量、初始速度、总粒子数、总能、demon 能，模拟足够多步后达到平衡时的粒子速度分布。微正则系综中没有定义温度，其数值由 $\frac{1}{2}kT = \frac{1}{2}m\langle v^2 \rangle$ 给出，求平衡时的温度值。

2.2.3.3 等温等压系综

等温等压系综在 Monte Carlo 模拟中有广泛的应用，这是因为大多数实验是在给定压力和温度的条件下进行的，而且恒 NPT 的模拟可以用来计算体系的状态方程，另外还可以模拟临近一阶相变的体系。恒压下的 Monte Carlo 模拟首先是针对二维硬盘势的模型提出的，后经改进可用于有任意连续势能表达式（如 Lennard-Jones 势）的体系。

考虑相互作用的经典粒子体系，配分函数（2.1.2.3-6）式中对粒子的动量坐标的积分可以求出并归为一相乘因子，只需要各粒子的空间坐标积分和体积积分。为此，设体系是处在边长为 $L = V^{1/3}$ 的立方盒子中，并将空间坐标 \mathbf{r} 用 L 进行约化得到标度坐标 \mathbf{s} ，即

$$\mathbf{r}_i = L\mathbf{s}_i \quad (i=1, \dots, N), \quad (2.2.3.3-1)$$

则配分函数为

$$Z_{NPT} = \int V^N \exp(-\beta PV) dV \int \exp[-\beta U(\mathbf{s}^N; L)] d\mathbf{s}^N. \quad (2.2.3.3-2)$$

在给定体积 V 内包含 N 个粒子的一个特定构型中，体系的几率密度分布为

$$\begin{aligned} \rho_{NPT}(\mathbf{s}^N; V) &\propto V^N \exp(-\beta PV) \exp[-\beta U(\mathbf{s}^N; L)] \\ &= \exp\left\{-\beta\left[U(\mathbf{s}^N; V) + PV - N\beta^{-1} \ln V\right]\right\}. \end{aligned} \quad (2.2.3.3-3)$$

在恒 NPT 系综的 Monte Carlo 模拟中， V 被处理为一个附加坐标。于是，体积改变的尝试移动和粒子位置 \mathbf{s} 改变中的尝试移动必须满足相同的准则，即 Markov 链的要求。设体积的尝试移动是从 $V_0 \rightarrow V_1 = V_0 + \delta V$ ，其中 δV 可正可负，其值用随机抽样 $\delta V = (\xi - 0.5)\Delta V$ ， ΔV 是给定最大移动步长（可根据模拟结果的接受率进行调整），可对系统的边长 L 改变而得到。按 Metropolis 方法，随机

体积变化的接受几率为

$$p_a = \min \left[1, \exp \left\{ -\beta \left[U(\mathbf{s}^N; V_1) - U(\mathbf{s}^N; V_0) + P(V_1 - V_0) - N\beta^{-1} \ln(V_1/V_0) \right] \right\} \right], \quad (2.2.3.3-4)$$

一般来说, 一次体积尝试移动需要重新计算新的约化坐标下的所有粒子间的相互作用, 因此体积移动相当于进行 N 次粒子位置的移动。通常是对 N 个粒子位置全部尝试移动一次后, 在下一个循环之间插入一次体积尝试移动。另外还可以以 $1/N$ 概率的体积移动插入到粒子位置移动中。

如果粒子体系的相互作用势能具有下述距离的函数形式的话,

$$U = \varepsilon \sum_{i < j} (\sigma/r_{ij})^n = \varepsilon \sum_{i < j} [\sigma/(Ls_{ij})]^n, \quad (2.2.3.3-5)$$

则当体系的线性维数从 L_0 变化至 L_1 时, 势能的变化很易计算, 因为

$$U(L_1) = (L_0/L_1)^n U(L_0). \quad (2.2.3.3-6)$$

这时没有必要再重新计算新的约化坐标下的所有粒子间的相互作用。当势能是 (2.1.3.2-5) 型的线性组合时 (如 Lennard-Jones 势), 也是如此, 即对每一项有 (2.1.3.2-6) 式。故这样的体积尝试移动可以用和粒子位置尝试移动一样的频率。

2.2.3.4 巨正则系综

用巨正则系综研究的体系是粒子数可变的, 例如在气体和吸附剂组成的体系中那样, 我们需要得到与外部环境有关的体系内粒子平均数目的信息, 达到平衡时, 吸附剂内气体和外部气体的温度和化学势相同。

将 (2.2.3.3-1) 式中的标度坐标应用到 (2.1.2.4-6) 式的巨正则配分函数中 (只考虑一种粒子的体系), 与上同样将粒子动量积分求出, 定义 $\Lambda = \sqrt{2\pi mkT}$,

$$\begin{aligned} Z_{\mu VT} &= \sum_N \frac{1}{N!} \int \exp \left\{ -\beta [\mu N + H(\mathbf{q}, \mathbf{p})] \right\} d\Omega \\ &= \sum_N \frac{\exp(-\beta\mu N) V^N}{\Lambda^{3N} N!} \int \exp[-\beta U(\mathbf{s}^N)] d\mathbf{s}^N, \end{aligned} \quad (2.2.3.4-1)$$

相应的体系的几率密度分布为

$$\rho_{\mu VT}(\mathbf{s}^N; N) \propto \frac{V^N \exp(-\beta\mu N)}{\Lambda^{3N} N!} \exp[-\beta U(\mathbf{s}^N)]. \quad (2.2.3.4-2)$$

在 Monte Carlo 模拟中, 我们必须根据上式的几率分布进行随机抽样。按照 Metropolis 方法, 确定可接受的尝试移动为:

- 1、粒子空间位置的移动: 随机选择一粒子, 将构型改变, $(\mathbf{s}_0^N) \rightarrow (\mathbf{s}_1^N)$, 该步移动的接受几率为

$$p_a(\mathbf{s}_0^N \rightarrow \mathbf{s}_1^N) = \min \left[1, \exp \left\{ -\beta [U(\mathbf{s}_1^N) - U(\mathbf{s}_0^N)] \right\} \right]. \quad (2.2.3.4-3)$$

- 2、粒子的增加和削除: 在任意随机位置增加或削除一个粒子, 相应的接受

几率分别为:

$$p_a(N \rightarrow N+1) = \min \left[1, \frac{V}{\Lambda^3(N+1)} \exp \left\{ \beta \left[\mu - U(\mathbf{s}^{N+1}) + U(\mathbf{s}^N) \right] \right\} \right]; \quad (2.2.3.4-4)$$

$$p_a(N \rightarrow N-1) = \min \left[1, \frac{\Lambda^3 N}{V} \exp \left\{ -\beta \left[\mu + U(\mathbf{s}^{N-1}) - U(\mathbf{s}^N) \right] \right\} \right]. \quad (2.2.3.4-5)$$

如同证明 (2.2.2.2-1) 式那样, 我们可证明上述的接受几率确实满足细致平衡条件 (2.2.1.3-4) 式:

$$\frac{p(N \rightarrow N+1)}{p(N+1 \rightarrow N)} = \frac{W(N+1 \rightarrow N)}{W(N \rightarrow N+1)} = \frac{p_t(N+1 \rightarrow N)}{p_t(N \rightarrow N+1)} \frac{p_a(N+1 \rightarrow N)}{p_a(N \rightarrow N+1)}. \quad (2.2.3.4-6)$$

由于增加和削除一个粒子的转移概率是相等的, $p_t(N+1 \rightarrow N) = p_t(N \rightarrow N+1)$, 故代入 (2.2.3.2-2) 后得,

$$\begin{aligned} \frac{p(N \rightarrow N+1)}{p(N+1 \rightarrow N)} &= \frac{\exp[\beta \mu(N+1)] V^{N+1}}{\Lambda^{3(N+1)} (N+1)!} \exp[-\beta U(\mathbf{s}^{N+1})] \\ &\quad \times \frac{\Lambda^{3N} N!}{V^N} \exp(-\beta \mu N) \exp[\beta U(\mathbf{s}^N)] \\ &= \frac{V}{\Lambda^3(N+1)} \exp \left\{ \beta \left[\mu - U(\mathbf{s}^{N+1}) + U(\mathbf{s}^N) \right] \right\} \\ &= \frac{p_a(N+1 \rightarrow N)}{p_a(N \rightarrow N+1)}. \end{aligned} \quad (2.2.3.4-7)$$

注意在最后一步中, (2.2.3.4-4) 或 (2.2.3.4-5) 式中必取一个因子 1。

参考文献

- [1] D.P. Landau and K. Binder, A Guide to Monte Carlo Simulations in Statistical Physics (Cambridge University Press, Cambridge, 2000) (该书有世界图书出版公司中国重影版)。
- [2] D. Frenkel and B. Smit, Understanding Molecular Simulation: From Algorithms to Applications (Elsevier, 1996) (中文译本: Frenkel & Smit (著), 汪文川 等 (译), 《分子模拟 - 从算法到应用》, 化学工业出版社, 2002年)。