

によって拡散が進むという仮定に基づいている。つまり、拡散のある特定の方向に偏らせる外部あるいは内部の場は存在しないと考えている。たとえば、空孔の相関性のない拡散は、ランダムウォークによって記述できる。しかし、先端の空間モンテカルロシミュレーションでは、連鎖する事象が無相関でなかったり、または、方向の変化がエネルギー密度や化学的な濃度における、ある障壁とリンクした拡散プロセスやパーコレーション問題をも記述することもできる。

6.5 当たりはずれランダムサンプリング積分法

当たりはずれ手法は、定積分の近似計算を行うための最も古くで最も単純なモンテカルロ法である。ある制限範囲内ではデジタルコンピュータを使用しないで実行することもできる。当たりはずれ法では、ランダムに選択された関数の引数に対して多くの被積分関数値を直接計算して平均化する(6.6節参照)のではなく、任意次元の座標値をランダムに発生させ、その対応する点が与えられた被積分関数より上にあるか、下にあるかを調べることを行う。多数の試行に対して、試行回数の総和で当たり数を除した値が積分の相対的な近似値となる。

モンテカルロ当たりはずれ法を用いた単純なサンプリング手法の例は、円の積分の相対値によって示される(図6.5)。第一に、半径1の円を拡張した $-1 < x < 1, -1 < y < 1$ の領域の正方形に、両方の領域が完全に重なるように、つまり中心座標(0,0)が一致するようににはめ込む。第二に、上述の領域における座標値と解釈する一組の乱数を生じさせ、多くの試行を行う。試行を正方形で規定された領域内に制限するために、 x, y 座標ともに $(-1, +1)$ 内の一様分布より選ばれる。第三に、ランダムに選択された点が円の内部にあるのか外側にあるかを調べる。その積分の相対値は、当たり数(n_{hit})を試行回数(n_{all})で除することによって計算される。

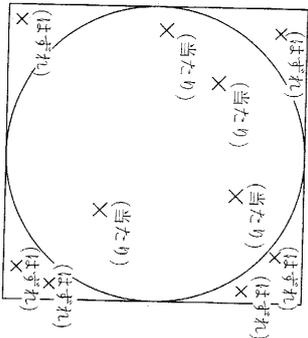


図 6.5 当たりはずれ法を用いた円の積分

$$\frac{\text{円の面積}}{\text{正方形の面積}} = \frac{n_{hit}}{n_{all}} \quad (6.2)$$

数値実験の対称性を利用して積分時間を減らすことができる。たとえば、上述の例では、

(0,1)の一様分布を用いて、全領域の4分の1に対してランダムな x, y 座標値を選ぶだけで十分である。その結果、その相対値は、

$$\frac{\text{円の面積}}{\text{正方形の面積}} = \frac{n'_{hit}}{n'_{all}} \quad (6.3)$$

となる。ここで、 n'_{hit} と n'_{all} はそれぞれ、縮小した領域における当たり数と試行回数である。この数値近似法の精度は試行回数に依存することは明らかである。わずか2回の試行では、その相対値は0%, 50%, 100%のいずれかになり、 $\pi/4 \approx 78.5\%$ からは程遠い。より複雑な関数を積分するときには、単純な当たりはずれ法は明らかに実用的ではない(図6.6)。たとえば、デルタ型関数(a)またはボルツマン関数(b)を数値積分する場合は効率的でない。なぜなら、ランダムな座標値を選ぶ際に、関心のある比較的狭い領域(積分の内側)からではなく、全領域(積分の外側)から選ぶからである。そのような場合には、関数が本質的に積分値に寄与する領域内には、試行のほんのわずかな部分だけが存在することになってしまう。

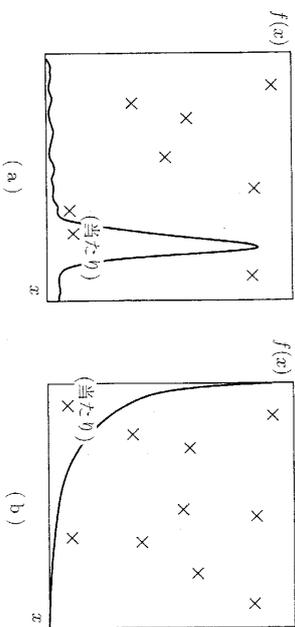


図 6.6 当たりはずれ法を用いた(a)デルタ型関数、および(b)ボルツマン型関数の積分。両方の場合に対して、単純な当たりはずれ法は明らかに非効率である。

したがって、確率的サンプリングの効率は、積分される関数と類似しているが必ずしも同じではない形の領域から乱数を選ぶことによって改善される。これは重点サンプリング法の基礎的な考え方である。これに関しては6.7節で扱う。

6.6 単純サンプリング二重みづけのないモンテカルロ積分法

十分に多くの擬似乱数があれば、モンテカルロ法は多次元の定積分の近似値を求めるために効果的に用いることができる(Hammersley と Handscomb 1964, Allen と Tildesley 1989)。単純モンテカルロ法は、一様分布から乱数を選んで行われるあらゆる積分法を網羅している(Gardiner 1985, Binder と Stauffer 1987)。

単純な例として、次に示す一次元の定積分問題を考える。

$$I = \int_0^b h(x) dx \quad (6.4)$$

この積分の値は、区間 $[a, b]$ を n 個の短冊、すなわち $x_0 = a$ と $x_n = b$ となるような部分区間に等分し、対応する離散的な関数値を足し合わせるることによって求められる。真の積分値 J の近似値 \bar{J} は、

$$J \approx \bar{J} = (b-a) \frac{1}{n} \sum_{i=1}^n h(x_i) \quad (6.5)$$

となる。後に示すが(6.11節)、この積分の数値的近似における統計的誤差は、 $1/\sqrt{n}$ に比例して減少する。つまり、近似値と真値との偏差は、 n が無限大に近づくにつれて0に近づく。このことは、 n が非常に大きい数である場合に、積分 $\int_a^b h(x) dx$ の値は n 個の関数値 $h(x_1), h(x_2), \dots, h(x_n)$ の平均値で近似することができることを意味する。この手法は、区間 $[a, b]$ 内で等間隔で分布し、それぞれが等しい重みをもつ座標 x_1, x_2, \dots, x_n から被積分関数値をサンプリングすることによって積分を近似することに等しい。この積分法は決定論的サンプリング法とよばれる。

その一方で、同様に妥当な積分の近似値を得るために、区間 $[a, b]$ において規則正しく部分区間を選ぶのではなく、区間 $[a, b]$ における一様乱数からランダムに選ばれた m 個の座標 x_1, x_2, \dots, x_m における関数値をサンプリングし、平均化することも可能である(図6.7)。この手法では、調査する区間において m 個の無相関乱数による大きな列(ペルコフ連鎖)を必要とする。これらの点に対する被積分関数値は一組の統計的サンプルとみなすことができる。

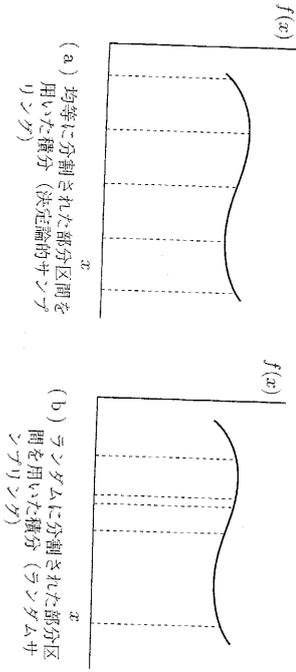


図 6.7 重みつけない数値積分法

$$J \approx \bar{J}_{MC} = (b-a) \frac{1}{m} \sum_{j=1}^m h(x_j) \quad (6.6)$$

ここで、 J は真の積分値、 \bar{J}_{MC} は統計的サンプリングによる近似値である。この手法で高効率を実現するためには、 $m \ll n$ であることが望ましい。ランダムサンプリングとよばれるこの手法の妥当性は、確率論の中心極限定理により確かめられる。ランダムサンプリングによる積分の数値的近似における統計的誤差は $1/\sqrt{m}$ に比例して減少する。このように、対象とする区間における一様分布から独立変数を選び、被積分関数値を計算するモンテカルロ法は、単純サンプリング法、または重みつけないランダムサンプリング法とよばれる。このランダムな数値サンプリング法は、低次元の積分の近似値を求める場合には、

古典的な決定論的手法と比較してはるかに効率的でない。たとえば、 m 個のサンプルをとった場合、台形則では $1/m^2$ の誤差を生じるが、これに對してモンテカルロ近似法では $1/\sqrt{m}$ の誤差を生じる。しかし、ランダムサンプリング法の実際の利点は、多体の動力学や統計力学の分野で現れるような、多次元の積分を近似する際に生まれる(第5章参照)。多次元系において、モンテカルロ法による積分誤差は $1/\sqrt{m}$ に比例することは変わらないが、台形則による積分誤差は $1/m^2/D$ となる (D は次元数)。

ランダムサンプリング法による積分の概念は、 k 次元の関数まで容易に拡張できる。

$$\int_{a_1=0}^{b_1=1} \int_{a_2=0}^{b_2=1} \dots \int_{a_k=0}^{b_k=1} h(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k \approx \frac{1}{s} \sum_{i=1}^s h(x_{i1}, x_{i2}, \dots, x_{ik}) \quad (6.7)$$

ここで、 $(x_{i1}, x_{i2}, \dots, x_{ik}) = x_i$ は s 個の無相関なサンプリングベクトルによる集合の i 番目のものである。各ベクトル x_i は、ランダムに選ばれ一様に重みづけされた k 個の成分からなる。このベクトルは、積分によって記述される多次元体におけるサンプリング点を表している。

各次元において積分範囲の長さが1ではなく、たとえば $[a_1, b_1], [a_2, b_2], \dots, [a_k, b_k]$ (a_i, b_i は任意の値)の範囲となっている場合、区間 $[a_1, b_1], [a_2, b_2], \dots, [a_k, b_k]$ に成分をもつ s 個のランダムに分布するベクトルに関する $h(x_{i1}, x_{i2}, \dots, x_{ik})$ の平均は、積分の単純な近似値ではなく、積分値をそれが占める多次元領域で割った値に対する近似値となる。つまり、

$$\frac{1}{s} \sum_{i=1}^s h(x_{i1}, x_{i2}, \dots, x_{ik}) \approx \frac{1}{(b_1 - a_1)(b_2 - a_2) \dots (b_k - a_k)} \int_{a_1}^{b_1} \int_{a_2}^{b_2} \dots \int_{a_k}^{b_k} h(x_1, x_2, \dots, x_k) dx_1 dx_2 \dots dx_k \quad (6.8)$$

となる。多次元のモンテカルロ積分法の典型的な応用例は、原子スケールの多体シミュレーションの分野にみられる。この分野では、関連する巨視的なデータは、統計力学を用いて多くの微視的なデータから導かれなければならない。たとえば、離散的なエネルギーの場合には、物性値 q に対する正準集合平均 $\langle q \rangle_{NVT}$ は、

$$\langle q \rangle_{NVT} = \frac{\int \rho_{NVT}(\mathbf{r}) q(\mathbf{r}) d\mathbf{r}}{\int \rho_{NVT}(\mathbf{r}) d\mathbf{r}} \quad (6.9)$$

のように表される。ここで、 $\mathbf{r} = r_1, \dots, r_N, p_1, \dots, p_N$ は位相空間配置あるいは位相空間での点、演算子 $\rho_{NVT}(\mathbf{r})$ は正準確率あるいは位相空間の分布関数である(第5章参照)。この方程式は正準確率密度 $\omega_{NVT}(\mathbf{r})$ を用いて書き換えられる。

$$\langle q \rangle_{NVT} = \frac{\int \mathbf{r} \omega(\mathbf{r}) q(\mathbf{r}) d\mathbf{r}}{\int \mathbf{r} \omega(\mathbf{r}) d\mathbf{r}} \quad (6.10)$$

s 個の離散的なサンプリングベクトル \mathbf{r}_i を用いて、この方程式をモンテカルロ積分のために定式化しなおすと、

$$\langle q \rangle_{NVT} \approx \langle q \rangle_{MC} = \frac{\sum_{i=1}^s \omega(\mathbf{T}_i) q(\mathbf{T}_i)}{\sum_{i=1}^s \omega(\mathbf{T}_i)} = \frac{\sum_{i=1}^s \exp(-\beta H(\mathbf{T}_i)) q(\mathbf{T}_i)}{\sum_{i=1}^s \exp(-\beta H(\mathbf{T}_i))} \quad (6.11)$$

となる。ここで、 $\beta = 1/k_B T$ 、 H は系のハミルトニアン、また、 $\langle q \rangle_{NVT}$ は $\langle q \rangle_{MC}$ の真の値に対するモンテカルロ法による近似値である。

図6.8は、被積分関数が位相空間において滑らかな関数でない場合には、ランダムサンプリング法が効果的でないことを表している。この例では、離散確率密度関数 $\omega_{NVT}(\mathbf{T}_i)$ である被積分関数にはいくつかのピークがあるが、ランダムサンプリング法では検出されない部分がある(斜線部)。この例を一般化すると次のことが示される。座標 \mathbf{T}_j における統計集合重み係数 $\omega_{ens}(\mathbf{T}_j) / \sum_i \omega_{ens}(\mathbf{T}_i)$ のある局所的な値が平均値から十分離れている場合には、単純サンプリング法は効果的ではない、とくに統計力学において、正準ボルツマン密度関数 $\omega_{ens}(\mathbf{T}) = \exp(-\beta H(\mathbf{T}))$ が配置空間の大部分の領域において非常に小さな値をとる場合が典型的な例である。重点ランダムサンプリング法の導入によってこの欠点を解消できることが次節の内容である。

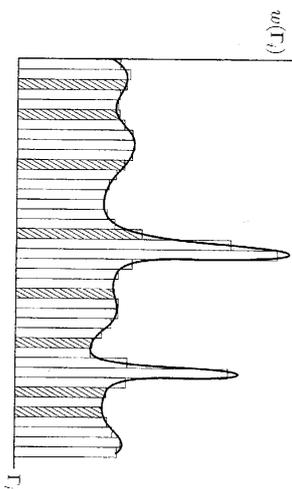


図 6.8 重みづけのないサンプリング法(単純サンプリング法)による離散統計集合確率密度関数 $\omega_{ens}(\mathbf{T}_i)$ の積分(斜線で示された被積分関数値から計算された平均値は真の積分値の低い近似値となる)

6.7 重点サンプリング-重点モンテカルロ積分

多体系の状態を表現する多くの平均量は積分形で表される(第5章参照)。適切な積分法を選択するためには、そのような平均量の特性を考慮すべきである。第一に、平均量は N 個の粒子の独立した座標値と運動量ベクトル(原子ガスでは $6N$ 個の変数)に依存するのうに、被積分関数は数桁の範囲で変化することがある。このことは、積分値に大きな影響を与える配置がある一方で、無視できるような重要度しかない配置もあることを意味する。第三に、選択された方法は一般的に、デルタ型関数のような急勾配の関数でも積分可能であることが求められる。そのような関数は、たとえば小正準集合の位相空間密度である。

従来の数値積分法は、そのような条件下では多くの場合、十分に高速でも正確でもない。上述した単純サンプリングモンテカルロ積分法では、限られた数の一様に分布した被積分関数値を用いるが、この方法でさえ、そのような場合には十分ではない。なぜなら、この方法では、極端に変化する位相密度関数を正確に積分することが困難だからである。その一方で、サンプリングによる積分は、なめらかな関数を積分する場合にはとても効果的であることを考慮するべきである。たとえば、被積分関数が定数である場合、単純サンプリング積分から正確な積分値を求めるには、ただ一点のみが必要とされる。同様の考え方は、被積分関数になめらかな一定値に近い場合にも適用される。このような場合には、少数のサンプリングされた被積分関数値からよい近似値が求められる。

さまざまな勾配をもつ被積分関数を非常になめらかな関数に変換し、積分範囲を調整するという考え方が重点サンプリング法の主要な背景である。いいかえれば、単純サンプリングから重点サンプリング法への拡張においては、重みづけされた被積分関数を用いるという大きな改善が導入されている。これは、選択された擬似乱数は一様ではない分布から選択されることを意味する。この手順により、積分値に重要な寄与をする空間領域に関数評価を集中させることができる。

積分 J を概算するための重み関数 $g(x)$ の導入により、以下のように書ける。

$$J = \int_a^b h(x) dx = \int_a^b \left[\frac{h(x)}{g(x)} \right] g(x) dx \quad (6.12)$$

ここで、積分変数を x から、たとえば y にかえるならば、

$$g(x) = \int_0^x g(x') dx' \quad (6.13)$$

となり、積分は以下のようになる。

$$J = \int_{g(x=a)}^{g(x=b)} \frac{h(g(x))}{g'(g(x))} dg \quad (6.14)$$

この積分のモンテカルロ法による近似は、上述のように進む。すなわち、 $h(g(x))/g'(g(x))$ の離散値を、区間 $[a, b]$ に対応する y の範囲で一様に分布する n 点のランダムサンプリングで平均化する。

$$\tilde{J} \approx (y(x=b) - y(x=a)) \frac{1}{n} \sum_{i=1}^n \frac{h(g(x)_i)}{g'(g(x)_i)} \quad (6.15)$$

重み関数のモンテカルロ法への導入は、簡単な解析例により明らかにになる。関数 $h(x)$ の a から b までの積分 J は以下のように書ける。

$$J = \int_a^b h(x) dx \quad (6.16)$$

$h(x) = \exp(-x/2)$ とすると、式(6.16)は以下のようになる。

$$J = \int_a^b \exp\left(-\frac{x}{2}\right) dx \quad (6.17)$$

はじめに、対象領域において、もとの関数 $h(x)$ と同様の分布をもつ2番目の関数 $g(x)$ を見つける。つまり、 $h(x) \approx g(x)$ であるような関数 $g(x)$ である。この例では、指数関数 $\exp(-x/2)$ は $1-x/2+x^2/2-x^3/6 \dots \approx 1-x/2$ の形の級数展開により近似される。区間 $[a=0, b=1]$ において、代入関数 $g(x) = 1-x/2$ を用いると、もとの積分は以下のように書ける。

$$\int_a^b \frac{h(x)}{g(x)} g(x) dx = \int_{g(a)}^{g(b)} \frac{h(y)}{g(y)} dy = \int_{y(a=a)}^{y(b=b)} \frac{\exp(-\frac{x}{2})}{1-\frac{x}{2}} dy \quad (6.18)$$

ここで、以下の2式を用いる。

$$y(x) = \int_0^x g(x') dx' = \int_0^x \left(1 - \frac{x'}{2}\right) dx' = x - \frac{x^2}{4} \quad (6.19)$$

$$x = 2(1 + \sqrt{1-y}) \quad (6.20)$$

もとの積分は結局以下のように書ける。

$$\int_a^b \frac{h(x)}{g(x)} g(x) dx = \int_{y=0}^{y=3/4} \frac{\exp(-\frac{x}{2})}{1-\frac{x}{2}} dy \quad (6.21)$$

この例から、変数の置き換えにより積分範囲と積分形が修正されることがわかる。それゆえ、 a から b までの範囲に一樣に分布する x に対する乱数を選択し、もとの関数 $h(x)$ の単純サンプリングモンテカルロ積分法にそれらを用いるかわりに、 y に対する一樣乱数を選択し、重み関数 $h(x)/g(x)$ を積分する。 y に対する乱数値が十分に分散している場合は、1に近い関数 $h(x)/g(x)$ を積分するという比較的単純な作業となる。

次節では、状態関数に対して重み関数を用いることに基づいたメトロポリスモンテカルロ法について、一定の化学組成をもつ正準集合の場合、およびさまざまな化学組成をもつ大正準集合の場合に対して説明する。これらの古典的な場合には、重み関数は通常、ボルツマン型の形式を仮定する。このことは、非量子粒子の確率分布はそのエネルギーに指数的に依存するという事実を反映している。さらに、メトロポリスアルゴリズムのさまざまな応用について議論する。近年、さまざまな発展形のメトロポリス法が提案されている。これらのなかでとくに原子ボテンシャルに関連したモンテカルロアルゴリズム (Bacher ら 1992)、およびいわゆる逆モンテカルロ法 (Livet 1987, Desplais ら 1996) がますます盛んになってきている。

6.8 メトロポリスモンテカルロ法

6.8.1 アルゴリズムの基礎

メトロポリスモンテカルロ法のアルゴリズムは、ある特定の統計的な集合における系の状態を生成するための、重みづけされたランダムサンプリング法、すなわち重点ランダムサンプリング法である。この系は熱平衡状態 $\rho_{\text{ens}}(\mathbf{I})$ において制約を受けた確率分布によって特徴づけられる (第5章参照)。一定の化学組成をもつ系 (正準集合, 小正準集合,

および等温・等圧集合) では、確率分布は系のハミルトニアン関数である。変化する化学組成をもつ系 (大正準集合) では、確率分布は化学ポテンシャルの関数である。

N 個の古典的な粒子からなる系の熱力学的状態は、位置ベクトル $\mathbf{r}_1, \dots, \mathbf{r}_N$ の3成分と、粒子の運動量ベクトル $\mathbf{p}_1, \dots, \mathbf{p}_N$ の3成分により特徴づけられる。これらの $6N$ 個の成分は $6N$ 次元位相空間におけるベクトル \mathbf{I} の成分とみなせる。

二つの集合状態 \mathbf{I}_i と \mathbf{I}_j は、系が状態 i から状態 j に移行する確率の大きさを示す遷移確率 π_{ij} により関係づけられる。系の i 番目の状態に対する確率密度関数の値は $\rho_{\text{ens}}(\mathbf{I}_i)$ で与えられ、 j 番目の状態に対する値は、 $\rho_{\text{ens}}(\mathbf{I}_j)$ で与えられる。この遷移行列 π の成分に対するメトロポリスモンテカルロ解法は以下のようなになる。

$$\pi_{ij} = \begin{cases} \alpha_{ij}, & (\rho_{\text{ens}}(\mathbf{I}_j) \geq \rho_{\text{ens}}(\mathbf{I}_i)), \quad i \neq j \\ \alpha_{ij} (\rho_{\text{ens}}(\mathbf{I}_j) / \rho_{\text{ens}}(\mathbf{I}_i)), & (\rho_{\text{ens}}(\mathbf{I}_j) < \rho_{\text{ens}}(\mathbf{I}_i)), \quad i \neq j \\ 1 - \sum_{i \neq j} \pi_{ij}, & (\rho_{\text{ens}}(\mathbf{I}_j) = \rho_{\text{ens}}(\mathbf{I}_i)), \quad i = j \end{cases} \quad (6.22)$$

ここで、 α は対称で確率論的なマルコフ連鎖の行列である。 $\rho_{\text{ens}}(\mathbf{I}_j) < \rho_{\text{ens}}(\mathbf{I}_i)$ である場合に対して π の成分を計算するための規則は、メトロポリスモンテカルロ法のアルゴリズムでは、比 $\rho_{\text{ens}}(\mathbf{I}_j) / \rho_{\text{ens}}(\mathbf{I}_i)$ がわかっているならばよく、統計集合の分配関数 $Z_{NV\tau}$ に関しては未知でかまわないことを示している (第5章参照)。行列 α の対称性を利用すると、以下のことが示される。

$$\sum_j \pi_{ij} = 1 \quad (6.23)$$

$$\rho_{\text{ens}}(\mathbf{I}_i) \pi_{ij} = \rho_{\text{ens}}(\mathbf{I}_j) \pi_{ji} \quad (6.24)$$

式 (6.23) は、各状態 \mathbf{I}_i が位相空間のどこかに存在していなければならないことを本質的には反映している。式 (6.24) は微視的可逆性の原理を示している。これはマルコフ連鎖に従った分布が確率密度分布関数 $\rho_{\text{ens}}(\mathbf{I})$ であることを示すために使うことができる。

メトロポリスアルゴリズムはランダムに、または体系的に新しい試状態 \mathbf{I}_j を選び、

$$\alpha = \frac{1}{N} \quad (6.25)$$

に応じて確率論的行列をサンプリングし (N は可能な粒子位置の数)、そして確率密度 $\rho_{\text{ens}}(\mathbf{I}_j)$ と $\rho_{\text{ens}}(\mathbf{I}_i)$ の比 $\rho_{\text{ens}}(\mathbf{I}_j) / \rho_{\text{ens}}(\mathbf{I}_i)$ に応じて配置エネルギー変化を評価することによって進行する。次の6.8.2項では、遷移行列に対するメトロポリス解が通常、どのようにして一定の化学組成または変化する化学組成をもつ統計集合に対して同定されるのかを示す。

6.8.2 正準集合, 小正準集合に対するメトロポリス法

メトロポリスモンテカルロ法では、 N 個の統計集合原子それぞれに初期位置が割り当てられる。続いて、この配置に対するハミルトニアンが計算される。開始時の系の再配置は、規定された巨視的な値によって定まる。正準系および小正準系では基本的な構成要素の数は規定されている。